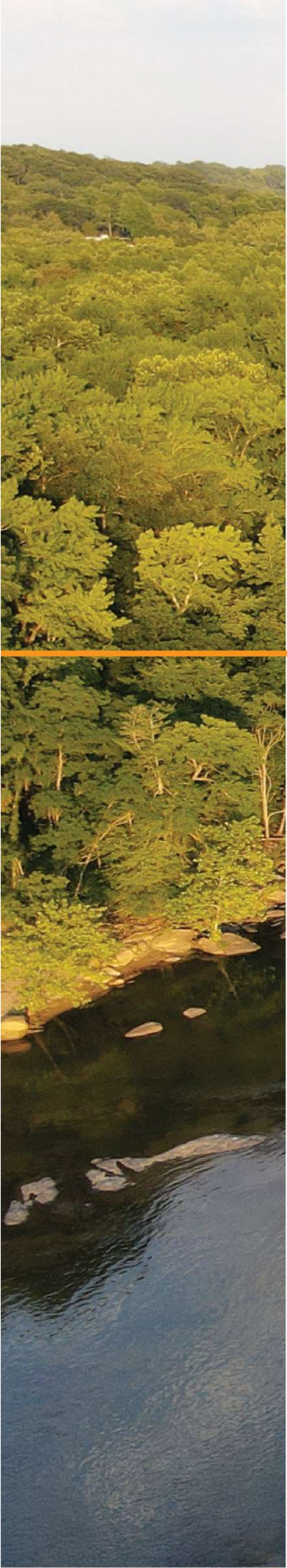


Achieving Water Quality Goals in the Chesapeake Bay: A Comprehensive Evaluation of System Response

An Independent Report from the Scientific and Technical Advisory Committee (STAC)
Chesapeake Bay Program
Annapolis, MD

May 2023

Scientific and Technical Advisory Committee (STAC)

About the Scientific and Technical Advisory Committee

The Scientific and Technical Advisory Committee (STAC) provides scientific and technical guidance to the Chesapeake Bay Program (CBP) on measures to restore and protect the Chesapeake Bay. Since its 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 workshops, and (5) interaction between STAC members and the CBP. Through professional and academic contacts and organizational networks of its members, STAC ensures close cooperation among and between the various research institutions and management agencies represented in the watershed. For additional information about STAC, please visit the STAC website at http://www.chesapeake.org/stac.

Publication Date: May 9, 2023

Publication Number: 23-006

Suggested Citation:
Scientific and Technical Advisory Committee (STAC). (2023). *Achieving water quality goals in the Chesapeake Bay: A comprehensive evaluation of system response* (K. Stephenson & D. Wardrop, Eds.). STAC Publication Number 23-006, Chesapeake Bay Program Scientific and Technical Advisory Committee (STAC), Edgewater, MD. 129 pp.

Report Editors
Kurt Stephenson, *Virginia Tech*
Denice Wardrop, *Chesapeake Research Consortium and Penn State University*

Writer Team
William Dennison, *University of Maryland Center for Environmental Science*
Zachary Easton, *Virginia Tech*
Mark Monaco, *National Oceanic and Atmospheric Administration, National Centers for Coastal Ocean Science*
Kenneth Rose, *University of Maryland Center for Environmental Science*
Leonard Shabman, *Resources for the Future*
Kurt Stephenson, *Virginia Tech*
Jeremy Testa, *University of Maryland Center for Environmental Science*
Denice Wardrop, *Chesapeake Research Consortium and Penn State University*

Steering Committee
Brian Benham, *Virginia Tech*
Anthony Buda, *U. S. Department of Agriculture, Agricultural Research Service*
William Dennison, *University of Maryland Center for Environmental Science*
Zachary Easton, *Virginia Tech*
Ellen Gilinsky, *Ellen Gilinsky, LLC*
Andy Miller, *University of Maryland Baltimore County*
Mark Monaco, *National Oceanic and Atmospheric Administration, National Centers for Coastal Ocean Science*
Kenneth Rose, *University of Maryland Center for Environmental Science*
Leonard Shabman, *Resources for the Future*
Kurt Stephenson, *Virginia Tech*
Jeremy Testa, *University of Maryland Center for Environmental Science*

Contributing STAC Members
William Ball, *Johns Hopkins University*
Lee Blaney, *University of Maryland Baltimore County*
Craig Beyrouty, *University of Maryland*
J. K. Böhlke, *U. S. Geological Survey*
Kathy Boomer, *Foundation for Food and Agriculture Research*
Charles Bott, *HRSD*
Chris Brosch, *Delaware Department of Agriculture*
Katherine Bunting-Howarth, *New York Sea Grant and Cornell Cooperative Extension*
Shirley Clark, *Penn State University*
Amy Collick, *Morehead State University*
Alix Dowling Fink, *Longwood University*
Celso Ferreira, *George Mason University*
Lara Fowler, *Penn State University*
Deidre Gibson, *Hampton University*

Kirk Havens, *Virginia Institute of Marine Science*
Ben Hayes, *Bucknell University*
Jason Hubbart, *West Virginia University*
Thomas Ihde, *Morgan State University, Patuxent Environmental & Aquatic Research Laboratory*
Thomas Johnson, *U. S. Environmental Protection Agency*
Brandon Jones, *National Science Foundation*
Hamid Karimi, *District of Columbia Department of Energy and Environment*
Jennifer Keisman, *U. S. Geological Survey*
Elliot Kellner, *U. S. Department of Agriculture, Natural Resources Conservation Service*
Catherine Kling, *Cornell University*
Scott Knoche, *Morgan State University, Patuxent Environmental & Aquatic Research Laboratory*
Ellen Kohl, *St. Mary's College of Maryland*
Erin Letavic, *Herbert, Rowland & Grubic, Inc.*
Martin Lowenfish, *U. S. Department of Agriculture, Natural Resources Conservation Service*
Chanceé Lundy, *Nspiregreen, LLC*
Dave Martin, *The Nature Conservancy*
Gregory Noe, *U. S. Geological Survey*
Efeturi Oghenekaro, *District of Columbia Department of Energy and Environment*
Leah Palm-Forster, *University of Delaware*
Michael Runge, *U. S. Geological Survey*
Larry Sanford, *University of Maryland Center for Environmental Science*
Adel Shirmohammadi, *University of Maryland*
Eric Smith, *Virginia Tech*
Jay Stauffer, *Penn State University*
Tess Thompson, *Virginia Tech*
Leon Tillman, *U. S. Department of Agriculture, Natural Resources Conservation Service*
Lisa Wainger, *University of Maryland Center for Environmental Science*
Gene Yagow, *Virginia Tech*
Weixing Zhu, *Binghamton University*

Acknowledgements
Carl Hershner, *Virginia Institute of Marine Science*
Lew Linker, *U. S. Environmental Protection Agency*
Gary Shenk, *U. S. Geological Survey*
Peter Tango, *U. S. Geological Survey*
Qian Zhang, *University of Maryland Center for Environmental Science*

STAC Staff
Meg Cole, *Chesapeake Research Consortium*
Tou Matthews, *Chesapeake Research Consortium*

Executive Summary

In 1983, the governors of Maryland, Pennsylvania, and Virginia, the mayor of District of Columbia, and the U.S. Environmental Protection Agency (EPA) administrator signed the first Chesapeake Bay Agreement. The one-page agreement acknowledged the "historical decline in the living resources of the Chesapeake Bay" and committed to addressing a major cause of the decline by pledging "to fully address the extent, complexity, and sources of pollutants entering the Bay." Subsequent Bay agreements have expanded the number of partners and the number of restoration goals, but reducing two key pollutants, nitrogen (N) and phosphorus (P), has remained a centerpiece of every subsequent Bay agreement.

Over four decades, water quality and pollutant reduction goals have been established and refined. Under authority provided by the Clean Water Act, the Bay jurisdictions and EPA adopted Bay water quality standards in 2003. The water quality standards identified living resources as the designated use of the Bay and defined numeric water quality criteria deemed necessary to support the designated use. Numeric water quality criteria were set for dissolved oxygen, water clarity, and chlorophyll *a* across five different Bay habitats: shallow water (submerged aquatic vegetation [SAV]), open water (fish and shellfish), deep water (seasonal fish and shellfish), deep channel (seasonal refuge), and migratory fish spawning and nursery.

Nutrient reduction goals were first written into the 1987 Bay agreement (and quantified in 1992 amendments). When nutrient reduction efforts failed to attain Bay water quality standards, EPA developed the country's most expansive total maximum daily load (TMDL) in 2010. The TMDL set nutrient and sediment load targets for the Bay that, if met, were predicted to achieve the water quality standards. The TMDL established that all management actions needed to achieve the target pollutant loads (214.9 million lb/yr of N, 13.3 million lb/yr of P, and 18,587 million lb/yr of sediment) should be in place by 2025. The Bay states and District of Columbia wrote watershed implementation plans (WIPs), which were approved by EPA, describing approaches to reduce nutrients and sediment to meet the load targets.

There has been progress in addressing nutrients since the first Bay agreement. The Chesapeake Bay Program (CBP) watershed model estimated that N loads to the Bay were reduced from 370 million lb/yr in 1985 to approximately 258 million lb/yr in 2021 and that P loads were reduced from 29 million lb/yr in 1985 to approximately 15 million lb/yr in 2021. Wastewater treatment plant upgrades provided the majority of these reductions. According to CBP estimates, the TMDL sediment limits have been met. Achieving these pollutant reductions in the face of significant population growth and development throughout the watershed is a noteworthy accomplishment.

However, modeling and monitoring evidence indicates that current efforts to reduce nutrient loads will not meet the TMDL targets. In addition, the CBP's ambient water quality monitoring program indicates that estuary water quality has been slow to respond to realized nutrient and sediment reductions in many regions of the Bay. The CBP has estimated that 27% of the Bay area met the water quality standards in 1985. By 2020, that figure had only risen to the mid-30% range. The consequences for living resources have not been fully evaluated.

This report summarizes the Scientific and Technical Advisory Committee (STAC) evaluation of why progress toward meeting the TMDL and water quality standards has been slower than expected and offers options for how progress can be accelerated. The report evaluates the effectiveness of current actions to reduce pollutants (N, P, and sediment) from wastewater treatment point sources and from farms and developed lands (nonpoint sources). Chapter 4 provides results from the evaluation of the water quality response in the estuary (dissolved oxygen, water clarity/SAV) to the realized nutrient and sediment reductions. Finally, chapter 5 summarizes what is known about the response of fish, shellfish, and other living organisms to changed water quality conditions. Decision-relevant uncertainties at each stage of program implementation and assessment are identified and their implications for progress considered.

Three overarching conclusions emerged from these evaluations. First, achieving pollutant reduction and water quality improvements is proving more challenging than expected. Second, the Bay system faces permanent and ongoing changes in land use, climate change, population growth, and economic development that will challenge notions of restoration based on recreating historical conditions. Third, opportunities to meet these challenges exist but efforts require changes and new approaches to implementation, planning, and decision-making. Specific findings of this evaluation supporting these conclusions and associated policy implications are summarized here.

Achieving the pollutant targets of the Bay TMDL

Finding: Existing implementation actions to reduce nonpoint sources of nutrients are insufficient to achieve the TMDL.

Meeting the TMDL depends to a significant degree on reducing nonpoint sources of pollutants. Agriculture is the largest remaining source of nutrient loads to the Bay, and urban nonpoint sources are the fastest growing. To date, CBP partner efforts to reduce nonpoint sources of nutrients have not produced sufficient levels of best management practice (BMP) implementation to meet the TMDL, and the implementation that has occurred may not be producing the pollutant reductions expected.

The CBP acknowledges the challenges of generating enough nonpoint source BMP adoption to meet nutrient reduction goals, particularly for N. Tens of millions of pounds of N reductions are needed to achieve the TMDL goal, but a decade of implementation since 2010 has produced only 3 million lb/yr of nonpoint source N reductions (as estimated by the CBP watershed model). The difference between water quality practices implemented and practices needed is termed an implementation gap and has multiple potential causes. Nonpoint source incentive programs are generally designed to encourage voluntary adoption of BMPs by covering a portion of the costs of installation. While successful at encouraging the adoption of practices that generate benefits to landowners (e.g., enhanced soil productivity), such programs do not provide sufficient incentives for adoption of practices with the largest pollutant reduction potential. Evidence also suggests that nutrient load reduction gains that have come from BMP implementation efforts are being partially offset by regional increases in imported nutrients. For example, increases in livestock numbers mean more N and P are imported into a region in the form of animal feed without corresponding increases in exports of animal products or by-products (i.e., manure). This nutrient mass imbalance leads to an accumulation of nutrients in the watershed that in turn may be transported to the Bay in runoff.

Evidence also suggests that the nonpoint source pollutant control efforts may not be as effective at producing nutrient reductions as expected by the CBP, resulting in a response gap. The existence of a response gap means that less progress is being made in meeting TMDL pollutant targets than represented by current accounting systems, and more nonpoint source controls will be needed to produce needed pollutant reductions. The response gap for phosphorus may be particularly large. While CBP modeling suggests that P reductions targeted by the TMDL are nearly achieved, analysis of water quality at riverine monitoring stations finds limited evidence of observable reductions in P concentrations. Nutrient response gaps have many potential causes, including long lag times for actions taken on the ground to produce reductions at water quality monitoring stations. However, response gaps could have a variety of other causes, including incomplete understanding of how people use nutrients on the landscape (particularly animal manures), overestimating nonpoint source practice effectiveness, incomplete or inaccurate information about nutrient inputs, landscape changes, and insufficient monitoring. Identifying and addressing response gaps is challenging, and this challenge is exacerbated by the TMDL accounting framework that tasks water quality managers with counting practices implemented and thereby diverts attention from the question of whether those practices generate the predicted pollutant reductions.

Together the implementation and response gaps represent significant challenges to the CBP's ability to achieve the nonpoint source pollutant load reductions as required by the TMDL. Uncertainty and complexity of nonpoint source-generating behaviors and processes confound assessment of these gaps. These challenges are not unique to the CBP, with many large-scale eutrophication management efforts (e.g., Great Lakes, Gulf of Mexico, Baltic Sea) also facing similar challenges for reducing watershed-scale nonpoint source pollutants.

Policy implication: There are opportunities to further reduce nutrients from nonpoint sources, but changes to programs and policies need to be considered.

Additional funding of existing implementation efforts is unlikely to produce the intended nutrient reduction outcomes. Achieving and sustaining substantial nonpoint pollutant reductions will likely require development and adoption of new implementation programs and tools.

Nonpoint source implementation efforts could be improved by shifting the focus from a census of implemented practices to an accounting of load reductions. Finer spatial scale modeling and monitoring could further identify high nutrient loss areas and operations and be used to consider more effective treatment options. Additionally, new financial incentive programs such as pay-for-performance or pay-for-success programs offer opportunities to reward treatment of high-loss areas or operations and to encourage adoption of highly effective practices that land managers may not consider under standard cost-share programs. These approaches would provide both the identification of high-value opportunities and the incentives for landowners to take advantage of them.

Achieving large-scale reductions in nonpoint sources of nutrients depends on adequately addressing regional nutrient mass imbalances. Many regions of the Bay watershed exhibit mass imbalances, where nutrient imports to a region (animal feed, fertilizer, atmospheric deposition) exceed nutrient exports from the region (agricultural products harvested, manure transport). The problem is particularly acute in areas of intensive livestock production. A variety of options is available to address these imbalances, including implementing technologies that reduce nutrient inputs, improving manure distribution (from surplus to deficit areas), and exporting nutrients from the watershed.

Most nonpoint source policies are based on allowing land managers to decide whether or how to reduce nonpoint source pollution. Such an approach is often reasonable given the number and diversity of people involved in producing nonpoint pollutants. However, the extensive history of nonpoint policy illustrates the limits of relying on voluntary actions. New and refined requirements in case-specific circumstances may be necessary to achieve substantial progress in reducing nonpoint source loads. Such requirements need not be overly costly to land managers if land managers are given flexibility in how to meet the pollutant control requirements and are provided financial assistance (similar to how some states are upgrading wastewater treatment plants).

Given uncertainties around the complexity and diversity of nonpoint source pollutant processes, not all alternatives will work as expected. Nevertheless, program change, innovation, and experimentation are needed. Institutional innovation could be facilitated by considering ideas such as sandboxing. Sandboxing is a formalized way to test and evaluate the efficacy of new rules and programmatic approaches to nonpoint source or water quality management without disrupting the operation of existing implementation efforts. Sandboxing also requires a commitment from management agencies to make larger programmatic changes if the sandboxed change demonstratively improves outcomes.

Achieving the water quality standards

Finding: Preliminary analyses suggest that nutrient load reductions have not produced the expected level of improvement in estuary water quality, and this response gap is particularly pronounced in the Bay's deep channel.

Evidence indicates that the nutrient and sediment load reductions realized to date have led to improved water quality conditions in some portions the Bay, but these nutrient load reductions have not produced the expected level of increased dissolved oxygen in most of the Bay's habitats. This shortfall, or water quality response gap, is particularly pronounced in the Bay's deeper waters and could have significant consequences because of the large nutrient reductions required to achieve the dissolved oxygen criteria in the deep water and deep channel habitats. Quantification of a response gap for water clarity is not possible because of the absence of a formal predictive model, but progress in improving water clarity and expanding SAV remains below the stated goal.

A variety of factors may explain the response gap in water quality conditions to pollutant load reduction. For example, recent studies suggest that higher water temperatures offset roughly 6–34% of the water quality improvement from N reductions. Furthermore, Bay water quality response will differ across habitats and may be nonlinear in some. That is, water quality response to pollutant loads may occur fairly slowly until conditions are sufficient to accelerate improvements. The thresholds where conditions more rapidly improve are often called tipping points.

Identifying response gaps and their potential causes is limited by the design of the CBP monitoring networks. The current estuary monitoring program is more attuned to assessing attainment of water quality criteria than understanding processes underlying water quality response. For example, nonlinear interactions (tipping points) have been identified at the scale of subsystems in the Bay, but monitoring to determine the thresholds associated with either degradation or restoration is not currently done. Monitoring designs may need to be modified, and coupled with research and modeling efforts, to better understand the range of conditions and relationships between stressors and water quality standards attainment.

Policy implication: Additional nutrient reductions will improve water quality, but water quality criteria may be unattainable in some regions of the Bay under existing technologies.
The CBP is trying to achieve water quality standards in a highly altered environment that will continue to change in ways with no historical precedent. Climate change is producing increases in water temperature and changing precipitation patterns that confound efforts to achieve water quality goals. The deep channel dissolved oxygen level has proven to be relatively intransigent to load reduction efforts, but this area often serves as the primary policy focus for CBP work. This reality may necessitate assessing the costs and tradeoffs of attaining numeric water quality criteria in specific situations and locations and adapting numeric goals if desired.

Managing water quality to enhance living resources

Finding: Significant enhancement of living resources can be achieved through additional management actions without complete achievement of water quality standards across all habitats.
The Bay water quality criteria were selected based on chemical and physical conditions (dissolved oxygen and water clarity) necessary, but not sufficient alone, to support fish and invertebrate species living in different habitats and at different life stages. For instance, the presence of adequate dissolved oxygen in a habitat does not guarantee that organisms will fully populate that habitat. Direct evidence of the impact of water quality changes on various classes of living resources is mixed, partly because of the confounding multiple changes occurring and complex ecological interactions, and partly because there have not been substantial system-wide changes in some criteria, like dissolved oxygen. As a result, quantifying living resource

responses to any specific management and restoration action is a significant analytical challenge. While the CBP employs a suite of models to predict the impacts of management actions on chemical conditions in the estuary (particularly nutrient levels and dissolved oxygen), the CBP does not use models to relate changes in dissolved oxygen and habitat to the composition or abundance of living resources.

Living resource benefits may occur without full attainment of water quality criteria across all habitats and in every region of the Bay. The location and timing of water quality improvements will influence the composition and abundance of living resources. The five habitats demonstrate different patterns and trajectories of attainment of water quality criteria, and attaining the criteria is expected to be most difficult in the deep channel habitat. The shallow water and open water habitats, however, more directly influence the life cycles of most fish species. Habitat types also differ in their sensitivity to local management actions that can enhance living resource response. For example, actions in shallow waters such as creating living shorelines and improving benthic habitat can greatly increase the living resource response to water quality conditions.

The living resource outcomes that can be expected from incremental attainment of water quality criteria depend greatly on a host of other factors. Structural aquatic habitat, nearshore habitat (wetlands, shoreline), commercial and recreational harvest, disease, and water conditions (temperature, salinity) are all significant drivers of the composition and abundance of living resources. Research points to the importance of specific habitats (particularly shallow water) and nearshore conditions for many important species. Improvements in dissolved oxygen may not increase the abundance of desirable fish species if these other factors are already limiting populations. Thus, focusing investments on these other factors could improve composition and abundance of living resources for any given level of water quality improvement.

Policy implication: The legal requirements of the Clean Water Act (the water quality goal) divert attention away from considering multiple means of improving living resources (support of aquatic life as the designated use) as articulated in the Chesapeake Bay Watershed Agreement.

The TMDL framework presents challenges to focusing management attention and resources on improving living resource outcomes. The Chesapeake Bay Watershed Agreement lists 10 goals and 31 desired outcomes. Water quality is only one of the goals, but it is the only legally enforceable goal (under the Clean Water Act). This means that the benefits of restoration actions tend to be expressed primarily in terms of nutrient reductions rather than benefits for living resources. For example, benefits of restoring wetlands or living shorelines are often framed in terms of the TMDL rather than improved habitat. Yet these investments can substantially improve Bay living resources. A broader policy challenge for the CBP is how to allocate restoration funds and efforts to generate the largest living resource impacts for the most stakeholders.

Policy implication: Opportunities exist to adjust approaches to prioritize management actions that improve living resource response.

Possible changes to TMDL implementation could help prioritize water quality investments that have greater and more immediate impacts on living resources. The TMDL as currently structured directs management attention toward meeting an aggregate nutrient load limit that is largely driven by dissolved oxygen conditions in the Bay. A tiered approach to TMDL implementation would identify the locations or habitats expected to achieve pollutant reduction limits first. Shallow water habitats in specific regions of the Bay may offer significant opportunities to produce living resource responses. These are also areas with significant stakeholder engagement because of their status as primary areas of recreational use, their cultural significance, and their visibility as iconic Chesapeake landscapes. Reevaluation of water quality criteria may also include consideration of new criteria (e.g., water temperature, toxic and emerging contaminants of concern) or new frameworks for devising criteria (e.g., indicators of resilience). Exploring such policy options would be enhanced with additional analytical capacities and analyses capable of more fully articulating potential living resource responses to water quality management.

Enhancing adaptive management

Finding: The Chesapeake Bay Program's current portfolio of adaptive management processes is inadequate to address the uncertainties and response gaps described in this report.

Moving forward, the CBP enters a new era of management. The Bay of the future will be different from the Bay of the past because of permanent and ongoing changes in land use, climate change, population growth, and economic development. Refining restoration goals over time should be considered as knowledge evolves about what future conditions are possible, what local communities and the partnership at-large see as priorities, and what is required to attain those possible futures. Uncertainty is inherent in each of these.

The CBP has built a sophisticated TMDL implementation and accounting process premised on the use of deterministic predictive and planning models to secure a desired pollutant reduction and water quality response. The CBP's decision framework and associated Strategy Review System (SRS) assesses and evaluates progress toward achievement of specific CBP goals and adjusts implementation based on these assessments. The water quality goal also adds an accountability framework. However, the deterministic models providing single estimates of pollutant loads for all inputs, land uses, and management actions are not well suited for evaluating and addressing uncertainty. Such modeling approaches make it difficult to assess the performance risk of different BMPs, inform decision makers of uncertainties, or assess the robustness of management actions to underlying assumptions or changing environmental conditions. The CBP has limited capacity to assess the potential of management alternatives for improving living resources. The critical question is not simply: Are planned actions being undertaken? Rather, are the actions producing load reductions and improved estuary conditions?

Policy implication: Expanding the scope of adaptive management could address critical uncertainties and response gaps.

A formalized adaptive management approach currently exists: the SRS, complemented by the TMDL accountability framework. It is used to refine the existing implementation programs and accounting structure. However, the system does not provide adequate insights into potentially necessary policy changes at multiple levels ranging from devising new programs, rules, and accountability systems to making budgetary and funding decisions and revising goals. Enhancing adaptive management for water quality improvement suggests the CBP consider ways to include more decision makers who have influence in broader scale potential changes to the programs and policies.

To respond effectively to the issues raised in this report, the current adaptive management process for water quality could be enhanced in several ways. Decision science offers processes to integrate complex technical analyses with the planning processes used by those with the authority to make choices about goals, programs, and budgets. A number of tools and processes are available to identify and reduce decision-relevant uncertainties. Such approaches aim to identify those uncertainties that pose the greatest risk to achieving management objectives, identify how much a given outcome could be improved if a given uncertainty was resolved, and identify the cost of error. These tools can be used for a variety of purposes including supporting program design, implementation, and prioritization of research needs.

Conclusion

Four decades of efforts to manage nutrient and sediment pollutants have improved water quality conditions in some portions of the Chesapeake Bay, but results are mixed. Additionally, changing conditions from population growth, land use, and climate will make future restoration more challenging. However, opportunities exist to improve the effectiveness of pollution reduction efforts and accelerate improvements in living resources by building on the data, knowledge, and experience gained over decades of effort. Capitalizing on these opportunities will require adoption of new policies, procedures, and programs and expanded capacities to address uncertainties around system response in decision-making. Finally, achieving reductions in pollutants and realizing improvements in water quality and living resources in a system as large, diverse, and complex as the Bay watershed and estuary calls for patience as changes are planned and implemented and the system responds.

Contents

Figures

Foreword

In January 2003, the Chesapeake Bay Program's Scientific and Technical Advisory Committee (STAC) produced an independent report entitled *Chesapeake Futures* that captured the state of knowledge of the restoration effort and presented a likely set of outcomes, or scenarios, based on that knowledge and projected trends. The report detailed outcomes for land use and development, forests, agriculture, and the Bay and its fisheries under each of three scenarios: under recent trends, if the objectives of agreements in place at that time were met, and if feasible alternatives were put in place. Many of the scenarios were prescient in their predicted outcomes under the current agreement objectives and proposed alternatives that were feasible but only partially implemented. *Chesapeake Futures* did not propose specific policy recommendations but instead offered the report as constructive advice from a body of experts in the hope of informing the decision-making of the partnership engaged in restoration of the Bay. It inspired reflection, thought, discussion, and debate, and outlined a series of feasible innovations, all in the hope of moving the work forward. This report aspires to stand on the shoulders of this previous work.

Chesapeake Futures represented the work of many scientists and their collective expertise, all under the auspices of STAC. This report does the same, but with a specific focus on synthesizing the robust amount of data and discovery amassed over the last 30 to 40 years of the Chesapeake Bay Program (CBP) related to Bay water quality management and assessing the implications for future policy. The effort began as a STAC independent initiative in March 2019, after Kurt Stephenson, Zach Easton, and Brian Benham proposed the idea of a report that would identify gaps and uncertainties in system response—physical, chemical, biological, and socioeconomic—that impact efforts designed to attain water quality standards in Chesapeake Bay. STAC agreed to the challenge, and as STAC Chair at the time, Benham facilitated the development of a collaborative process that would engage the entire committee.

As a first step in approaching the long causal chain that links management actions to their eventual impact on water quality and living resources, workgroups were formed around the subsystems of this chain: nutrient and sediment reductions (watershed), water quality response to nutrient and sediment reductions (estuary) and living resource response to water quality (living resources). Each of these workgroups generated an independent document with a self-determined scope (i.e., workgroups were afforded flexibility to address issues beyond the original objectives). Because the content of each document was both unique and substantial, STAC chose to publish them as stand-alone documents with authorship attribution (see references provided below). In the second step, a steering committee developed a series of framing questions to guide the preparation of this report that would meet the original objective of identifying gaps and uncertainties in achieving the Bay total maximum daily load (TMDL) and water quality standards. Coeditors Stephenson and Wardrop, supported by a subgroup of the steering committee (the "Writer Group" that included Leonard Shabman, Zach Easton, Jeremy Testa, William Dennison, Kenny Rose, and Mark Monaco), were tasked with assembling ideas and contributions to write a single draft text, drawing material from the aforementioned resource documents, STAC and CBP reports, the scientific literature, and a limited amount of additional analyses performed in collaboration with CBP scientists. Zach Easton made significant contributions in drafting the initial framing of the report, and Leonard Shabman was particularly helpful in assisting coeditors Stephenson and Wardrop in summarizing the main findings and conceptualizing implications within the context of water policy. The resulting report (this document) was then submitted for several reviews by steering committee members, the STAC membership at-large, and the U.S. Geological Survey to produce a consensus report.

This report represents the thought and analyses of many beyond those who formally put pen to paper, via participation in the critical thinking and discernment processes that are the foundation of a science-based approach. They are listed as contributors. The editors also wish to recognize various members of the Chesapeake Bay partnership who provided additional expertise. We cannot begin to add up and properly attribute the contributions and commitment that this effort represents. Thank you for allowing us to stand on the shoulders of the scientific expertise that you represent and take in the view; it's an extraordinary and hopeful landscape.

Resource documents:

Easton, Z., Stephenson, K., Benham, B., Böhlke, J. K., Buda, A., Collick, A., Fowler, L., Gilinsky, E., Hershner, C., Miller, A., Noe, G., Palm-Forster, L., & Thompson, T. (2023). *Evaluation of watershed system response to nutrient and sediment policy and management.* STAC Publication Number 23-003, Chesapeake Bay Program Scientific and Technical Advisory Committee (STAC), Edgewater, MD. 55 pp.

Testa, J. M., Dennison, W. C., Ball, W. P., Boomer, K., Gibson, D. M., Linker, L., Runge, M. C., & Sanford, L. (2023). *Knowledge gaps, uncertainties, and opportunities regarding the response of the Chesapeake Bay estuary to proposed TMDLs*. STAC Publication Number 23-004, Chesapeake Bay Program Scientific and Technical Advisory Committee (STAC), Edgewater, MD. 61 pp.

Rose, K., Monaco, M. E., Ihde, T., Hubbart, J., Smith, E., Stauffer, J., & Havens, K. J. (2023). *Proposed framework for analyzing water quality and habitat effects on the living resources of Chesapeake Bay*. STAC Publication Number 23-005, Chesapeake Bay Program Scientific and Technical Advisory Committee (STAC), Edgewater, MD. 52 pp.

1. Introduction

The 2014 Chesapeake Bay Watershed Agreement (CBWA) (CBP, 2014) contains 10 broad management goals, most of which are directed at supporting viable populations of living resources throughout the Bay watershed. Federal, state, and local programs to secure those living resource goals, as measured by policy attention and resource commitment, have been directed primarily toward achieving water quality standards (WQS) in the tidal waters of the Bay. The standards follow from the Clean Water Act (CWA) and begin with specification of designated uses (DUs) to support specific living resources, and that support is manifested in ambient numeric water quality criteria (WQC) for dissolved oxygen (DO), chlorophyll *a* (Chl *a*), water clarity, and submerged aquatic vegetation (SAV), that in turn are expected to be attained by limiting nitrogen (N), phosphorus (P), and sediment loads to the Bay. After 2010 this load limit, the total maximum daily load (TMDL), or "pollution diet", became a legal obligation for the states with areas that drain into the Bay and for the District of Columbia (Linker et al., 2013; USEPA, 2010).

The 2010 Chesapeake Bay TMDL-driven nutrient and sediment load targets were based on the predicted response of the numeric criteria in the tidal tributaries and in the Bay itself. The predictions were made using the suite of computer models developed and refined over the years since the Chesapeake Bay Program (CBP) was created (Hood et al., 2021). Bay jurisdictions develop watershed implementation plans (WIPs) that describe water quality improvement actions they will take to meet the TMDL. The CBP modeling suite translates actions into nutrient and sediment loads to the Bay and then calculates whether the numeric WQC will be met.

Over the past 40 years (including two decades of effort prior to the TMDL), the Chesapeake Bay jurisdictions have made notable progress in reducing nutrient and sediment loads to the Bay. Based on reports of point and nonpoint source load reduction practices put in place from 1985 through 2021, the CBP watershed model CAST (Chesapeake Assessment Scenario Tool, the primary tool for calculating load reductions from the practices installed to meet the TMDL) predicts that the current level of implementation is sufficient to achieve 73%, 91%, and 100% of the N, P, and sediment load reductions specified in the TMDL (CBP, n.d.-a). To date, the majority of nutrient reductions have been attributed to wastewater treatment plant (WWTP) upgrades (point sources) and to large and sustained reductions in atmospheric N deposition (CBP, n.d.-a). Nonpoint sources have contributed a smaller share of the needed reductions. The CAST model calculates that total urban and agricultural N and P nonpoint loads decreased 15% and 29%, respectively, between 1985 and 2021. Nonpoint sources, however, are the single largest remaining source of nutrients to the Bay. Of the controllable nutrient loads in 2021 (excluding loads from natural sources), approximately three-quarters originate from agricultural and urban nonpoint sources.

However, the CBP's extensive ambient water quality monitoring system in the estuary indicates that water quality has been slow to respond to improvement efforts. Water quality standards have been attained in some portions of the Bay. Some species of SAV, an important Bay living resource and water quality criteria, have expanded in several regions of the Bay in recent years (Lefcheck et al., 2018). Nonetheless, the reported 30 years of load reductions appear to have had limited effect on achievement of WQS. A composite index developed by the CBP summarizes WQS attainment across the entire Bay (Zhang, Murphy, et al., 2018). Using this index, the CBP reported in 2020 that 29.6% of the Bay attained WQS during 2018–20, only up from 26.5% in 1985–87 (CBP, n.d.-e). This is an average annual improvement rate in WQS attainment of approximately 0.25% over the 30 years since Bay restoration activities began, although there is year-to-year variability (figure 1.1). At this rate of progress, full attainment of the WQS is an uncertain and distant possibility.

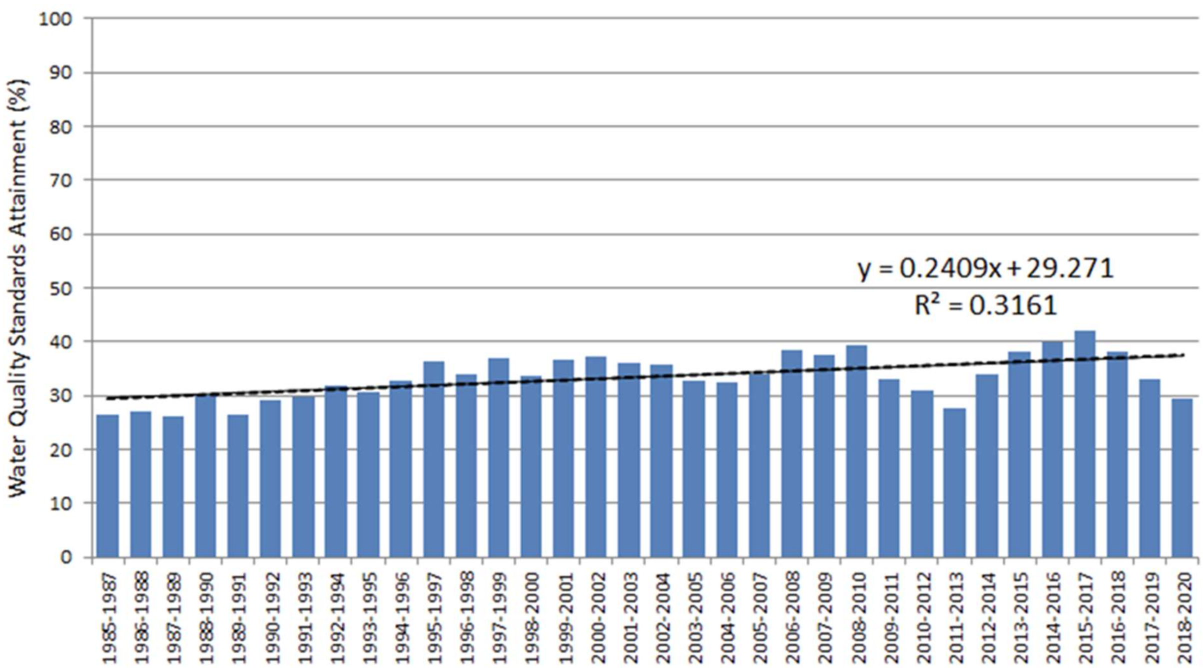


FIGURE 1.1.—Achievement of Chesapeake Bay WQS, 1985–2020 (Source: CBP, n.d.-e).

This report provides the results of the CBP Scientific and Technical Advisory Committee (STAC) effort to evaluate whether and why improvements in water quality have failed to meet expectations (called gaps here). One possible cause for a gap is that water quality improvements could just take time. Well-recognized lags occur in the system that lengthen the time between when actions are taken to reduce pollutants and when pollutant reductions and water quality outcomes are fully realized. However, other explanations also are considered. For example, the physical, chemical, biological, or socioeconomic systems may not be responding as scientists and water quality managers expect. The divergence between expected and realized outcomes can arise from incomplete or incorrect data or knowledge of system relationships, ranging from how policy translates into behavior and then pollutant control effectiveness to how estuary water quality responds to changing nutrient loads and how living resources respond to changes in water quality conditions. If water quality response to pollution control efforts is limited by more than just lag times, then policy changes may be needed.

With the 2025 TMDL WIP implementation deadline approaching, the CBP STAC undertook this effort to identify key areas of uncertainty and improve learning about efforts to attain the WQS underlying the TMDL process. The goal is to identify where CBP partnership water quality programs and policies may not be yielding anticipated system responses and to identify possible reasons for these gaps. Consequently, this report is purposefully unbalanced: we focus on response gaps and opportunities to improve our collective understanding of how to achieve water quality load targets and standards and improve living resource response, rather than celebrate what we already know or have already achieved. Recognizing and actively addressing response gaps and uncertainties can improve our understanding of system response and facilitate the development of new approaches and techniques to improve the effectiveness of CBP water quality improvement efforts.

The specific objectives of this report are to:

- identify gaps between the expected and realized physical, chemical, biological, and socioeconomic responses to management actions, and identify recent scientific developments that can advance efforts to attain WQS;
- characterize the critical uncertainties in system response to management actions, and identify strategies that improve understanding of system response relevant to the attainment of WQS; and
- identify strategies for better integrating scientific and technical analysis into management efforts in order to aid decision-making under uncertainty.

In pursuing these objectives, STAC strives to direct science in service of policy. The scientific community should not be expected to define what restoration means, decide what water quality goals are pursued, or decide which management actions and policies should be used to achieve those goals. However, the work of the scientific community must be relied upon to improve our collective understanding of the consequences of both the means and ends of policy, how the system responds to efforts to achieve a specific water quality goal, and what can be gained and at what cost from different WQS. Consequently, the report offers implications and options, rather than recommendations, for policies and programs regarding setting and achieving water quality goals in the face of uncertainty. For example, the report does not recommend a particular pollutant reduction policy be pursued but rather aims to frame the discussion as: if *X* pollutant reduction result is desired, then *Y* and *Z* options may improve the chances of achieving *X*. This report also focuses attention on those uncertainties (called decision-relevant uncertainties) and system responses that have direct and immediate implications for achieving the TMDL and specifying WQS.

The main body of the report is organized around system response. Chapter 2 summarizes Bay water quality policy and describes the terms and approach used in the report to evaluate system responses to policy goals and efforts. Chapter 3 describes Bay management efforts to achieve the TMDL and summarizes our understanding of how nutrient and sediment loads are responding to those management efforts. Chapter 4 reviews how the numeric measures of Bay water quality used to set the TMDL are being measured and how the measures are responding to reduced nutrient and sediment loads. The Bay water quality measures and the TMDL limit set to achieve these measures are expected to support populations of specific living resources. Chapter 5 reviews the available analyses of living resource response, noting that the link between meeting the current WQS and the abundance of specific living resources is not explicitly identified by the CBP and is highly uncertain. Chapter 6 describes the implications of the findings for future Bay water quality management, including identifying options for supporting Bay living resources and for research and monitoring to support those options.

2. Approach to Evaluating System Response to Water Quality Management Efforts

The organization of the report follows the conceptual logic underlying attainment of Chesapeake Bay WQS (see figure 2.1). The left side of figure 2.1 shows the Bay restoration policy goals and programs as stated in the 2014 CBWA (dark blue box). This report focuses on the specific water quality goal that became legally enforceable under the CWA in 2010 when the Bay was listed as an impaired waterbody (from sediment and nutrients P and N) and the Bay TMDL was developed. Reading from top to bottom in figure 2.1, the CWA planning process for the Bay began with specifying WQS, including DUs with the numeric WQC deemed necessary to support each DU. Because monitoring and assessment found that the WQS were not being met, a TMDL was prepared specifying the reductions in pollutant stressors N, P, and sediment needed to achieve the WQS. What followed was the design of federal, state, and local regulations and incentive programs to secure nutrient and sediment reductions from point and nonpoint sources. Bay states and the District of Columbia outlined general plans in WIPs. As actions called for in the plan were taken, progress toward implementing practices to meet load reductions was tracked and reported.

How social, physical, chemical, and biological systems respond to the water policy is shown on the right side of figure 2.1. STAC divides the causal chain that links pollutant reduction actions to achievement of WQS into three groups (reading from bottom to top in figure 2.1). First, CBP models (text box 2.1) predict how the management actions taken in response to regulations and incentives, such as wastewater treatment at point sources or best management practices (BMPs) on farm, forest, and urban lands, reduce Bay pollutants (total N, total P, and total suspended sediment, or TN, TP, and TSS) (orange box in figure 2.1). Second, nutrient and sediment reductions to the Bay are predicted to change water quality conditions (e.g., DO) at different locations throughout the estuary (light blue box in figure 2.1). Third, improvements in estuary water quality (as measured by the WQC) are expected to *support* fish, shellfish, and other living organisms in the Bay (green area in figure 2.1), although specifically how living resources will be supported by improved water quality is currently not predicted by the CBP.

Section 2.1 elaborates on the current structure of the Bay water quality program (left side of figure 2.1). Section 2.2 describes the general approach used to assess how the system is responding to CBP water quality policy across the causal chain (right side of figure 2.1). This background forms the basis of chapters 3, 4, and 5 (corresponding to each box on the right side of figure 2.1), which describe challenges to generating sufficient behavioral change and practice adoption (implementation gaps), identifying where the system may not be responding as predicted (response gaps), and addressing the uncertainties associated with the causes of those gaps.

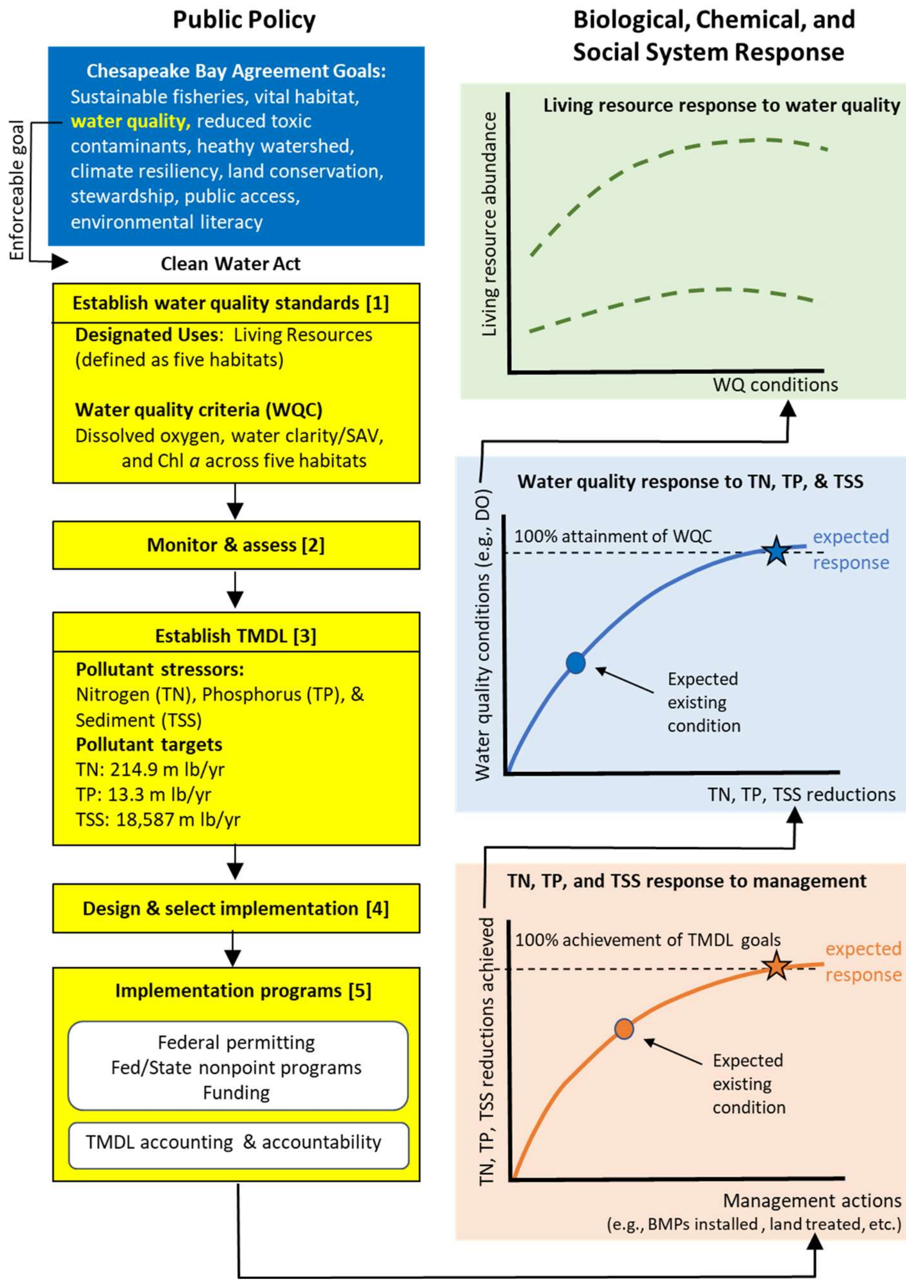


Figure 2.1.—Conceptual representation of system response to Chesapeake Bay water quality goals and management efforts

2.1. Water quality policy in the Chesapeake Bay: WQS, TMDL, and TMDL implementation

Chesapeake Bay states and responsible federal agencies have established numerous restoration goals for the Bay through a series of agreements since the inception of the CBP. For instance, the 1987 Chesapeake Bay Agreement established general restoration goals (non-numeric) including protection and enhancement of living resources (specifically identified as aquatic vegetation, habitat, shellfish, wetlands, and waterfowl/wildlife), water quality, public education, and public access (CBP, 1987). The most recent agreement, signed in 2014, includes 10 overall goals and 31 outcomes under those goals (see dark blue box in figure 2.1) (CBP, 2014). The focus of this report is the Chesapeake Bay water quality goal, but it is only one among a larger set of Bay restoration goals.

Since CBP inception, water quality improvement has been the primary focus of Bay restoration efforts. The 1987 Agreement stated the water quality goal as targets for nutrient load reduction, 40% reduction in nutrients N and P entering the Bay from 1985 baseline levels, which were to be achieved by the year 2000. The current Bay water quality goals that focus on the expected water quality conditions were first developed in the early 2000s (Tango & Batiuk, 2013; USEPA 2003a, 2003b). Prior to the required development of the Bay TMDL in 2010, the WQS were aspirational. In 2010 achievement of the WQS was formally incorporated into a legal and regulatory structure under Section 303d of the CWA. Once baywide water quality management came under the requirements of the CWA planning process, the Bay program had to adapt the CWA water quality planning process and then a TMDL process to the Bay as a whole.

The WQS consist of DUs for the estuary and numeric WQC to determine achievement of the DUs (yellow box 1 in figure 2.1). While DUs could include recreation, navigation, or water supply, the DU chosen for the Bay program was the protection of aquatic living resources. However, because a variety of habitats across the Bay support differing assemblages of aquatic organisms, the U.S. Environmental Protection Agency (EPA) convened a process through which state, federal, academic, and multistakeholder representatives agreed that the DUs would be made specific to major habitat types occurring across all of the Chesapeake Bay and its tidal tributaries (USEPA, 2003a). This meant that the DU of supporting aquatic living resources would be applied to five different habitats: shallow water (Bay grass use), open water (fish and shellfish use), deep water (seasonal fish and shellfish use), deep channel (seasonal refuge use), and migratory fish (spawning and nursery use) (figure 2.2).

Text box 2.1. Chesapeake Bay models referenced in this report

Three Chesapeake Bay models are referenced in this report as those used by the CBP to generate the expected estimates in pollutant loads and water quality outcome responses (Hood et al., 2021). The CBP uses two versions of the watershed model to generate estimates of nutrient loads flowing into the Bay from the watershed. CAST estimates average annual loads that would be generated under 10 years of typical weather conditions (typical defined as 1991–2000). The CBP uses CAST to set TMDL planning targets, design implementation plans, and track implementation progress. The dynamic version of the watershed model estimates daily nutrient loads and is used to provide loads estimates to the estuary model (Hood et al., 2021; CBP, n.d.-c). The CBP uses the estuarine model to estimate attainment of tidal Bay DO, Chl *a*, and water clarity criteria under different nutrient and sediment loads. The estuarine model is made up of a hydrodynamics component that measures transport in the estuary and biogeochemistry component that translates nutrients into water conditions.

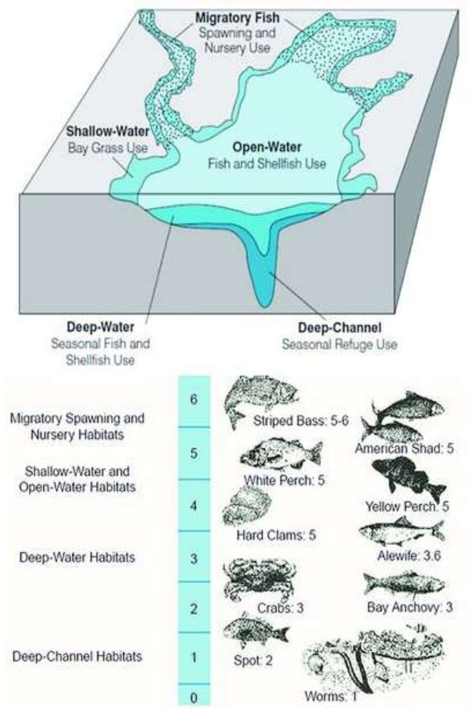


FIGURE 2.2.—Chesapeake Bay DUs (top) and DO WQC for five DU habitats (Source: USEPA, 2003a).

Numeric WQC are defined to secure attainment of the DUs in each habitat type. These WQC include numeric values for DO, water clarity, and Chl *a* for the designated habitats (Tango & Batiuk, 2013). Dissolved oxygen criteria have been developed for each of the five habitats throughout the Bay, while water clarity/SAV and Chl *a* apply to specific habitats (i.e., shallow water and select regions of the open water habitat, respectively). Open water habitat is expected to achieve a 30-day mean DO concentration of 5 mg/L, a 7-day mean of 4 mg/L, and an instantaneous minimum of 3.2 mg/L. These levels are necessary to support different life stages of fish and shellfish species. For deep channel habitat, an instantaneous minimum DO of 1 mg/L is designed to support benthic organisms (i.e., worms, clams) (figure 2.2). Water quality criteria for some habitats, such as deep water and deep channel, are defined based on seasonal values (DO levels in the summer months, June–Sept.).

Note that Bay WQC are largely based on chemical and physical conditions (e.g., DO and water clarity). With the exception of the SAV criteria for water clarity, the WQC are not based on achievement of specific living resource species or populations (such as a specific assemblage of fish, shellfish, or aquatic animals). Rather the numeric criteria were identified based on species tolerances for regions throughout the Bay (Monaco et al., 1998; Tango & Batiuk, 2013; USEPA 2003a, 2003b). For example, the 5 mg/L DO criteria in the open water habitat is designed to support finfish species such as striped bass. This approach ensures that adequate conditions are present to support organisms in different habitats, but it does not necessarily translate into actual population and food web responses because of the many other factors that also affect species abundances. Essentially, the WQC are necessary, but not sufficient, conditions to support aquatic organisms populating different habitats.

A series of specific water quality assessment monitoring protocols have been developed to determine the degree of attainment of the WQS corresponding to each WQC (Tango & Batiuk, 2013; USEPA 2003a, 2004, 2017) (yellow box 2 in figure 2.1). For assessment purposes, the Bay is divided into 92 distinct areas called segments (figure 2.3). Each segment typically contains multiple habitats and their associated WQC. For example, one segment may be assessed for open water (DO, and in some regions Chl *a*), deep water (DO), and shallow water habitats (DO, clarity/SAV). The CBP uses three years of data to determine whether the temporal and spatial distributions of water quality outcomes within a segment have achieved WQS attainment (compliance) (Batiuk et al., 2009).

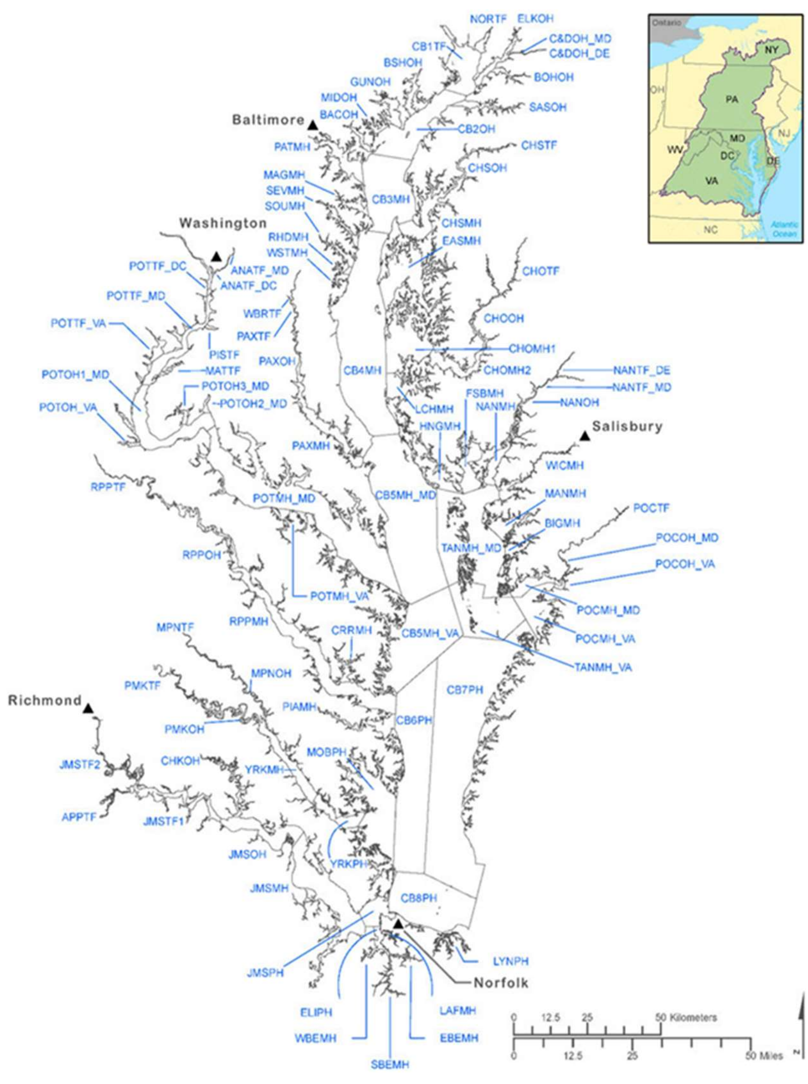


FIGURE 2.3.—The 92 water segments of the Chesapeake Bay TMDL (Source: CBP, 2008).

The CWA water quality management process expects that once WQC are established a process of continual monitoring and assessment follows to establish whether WQS are being met. Thus, full attainment of the WQS requires meeting multiple DO criteria in all 92 segments of the Chesapeake Bay with up to 5 different habitats in each segment. The Chl *a* and water clarity criteria must be met in a subset of those segments and habitats. The combinations of segments, habitats, and individual criteria represent a total of 1,052 unique conditions to be met before reaching full attainment (see figure 2.4). In addition, some of the WQS for the Bay cannot actually be assessed because of insufficient monitoring. For example, one-day and instantaneous minimums for DO cannot be assessed with existing monitoring frequencies in open water DUs.

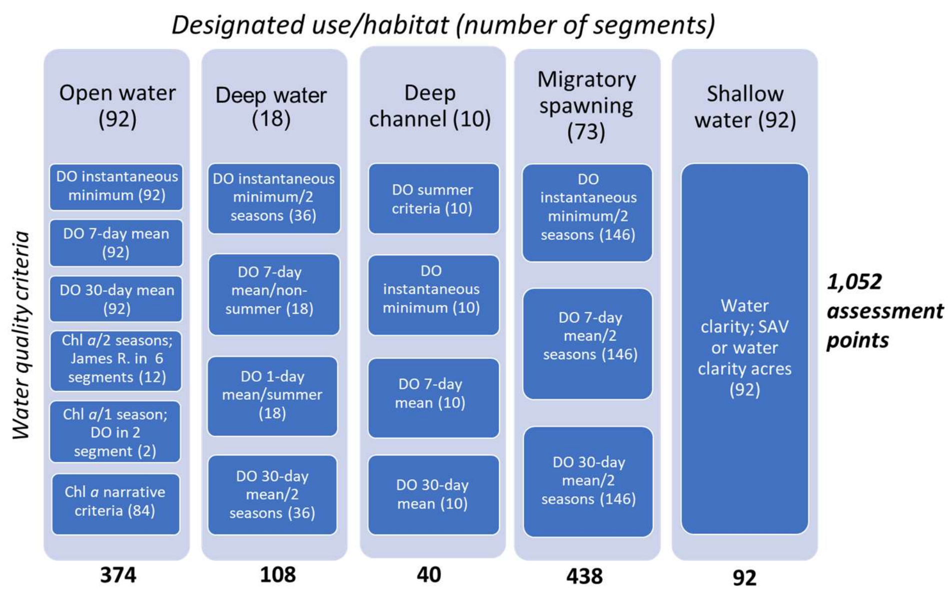


FIGURE 2.4.—Chesapeake Bay Program WQC and water quality assessment points across five DU habitats.

By 2010 the Bay was deemed impaired, and the CWA Section 303d process then required defining a limit on the allowable pollutants causing the impairment—the TMDL (yellow box 3 in figure 2.1). When the Chesapeake Bay TMDL was finalized, it was (and remains) the only TMDL for a water body as large and complex as the Bay. The Bay TMDL sets total annual limits on TN, TP, and TSS. If the TMDL nutrient and sediment load targets are reached, water quality is predicted to attain the WQS for all 1,052 assessment points in figure 2.4.

The TMDL was set using the CBP estuary water quality model that predicts how water quality conditions (e.g., DO) respond to changes in nutrient and sediment loads. As required by CWA planning regulations, when setting the TMDL modelers were expected to include "conservative" model assumptions to accommodate model prediction uncertainty. The TMDL TN, TP, and TSS load targets are 214.9, 13.3, and 18,587 million lb/yr, respectively (CBP, n.d.-e). The CBP estimates that achieving the nutrient and sediment reductions required to meet the

TMDL could not result in attainment of the criteria in every segment and habitat. Initial water quality modeling indicated that at least one segment (lower Chester River segment) could not meet the deep water DO criteria under the TMDL pollution diet given economically feasible pollution control options. Consequently, EPA approved a Maryland state “restoration variance” for that segment (USEPA, 2010). Since that time, EPA has granted restoration variances for five deep water or deep channel segments and nine open water segments.

Chesapeake Bay Program partners established a 2025 deadline for Bay jurisdictions to meet the TMDL pollution control obligations. The deadline does not require that all WQS be met or that pollutant reductions be realized by 2025, but rather the management practices that CBP models predict will be sufficient to meet the TMDL pollutant target loads for nutrients and sediment must be in place (USEPA 2010).

As part of the TMDL establishment, the jurisdictions designed plans for implementation of actions predicted to result in load reductions from CWA-regulated point sources and unregulated nonpoint sources (yellow box 4 in figure 2.1). The approach implemented is broadly similar to approaches taken in most TMDLs. State and federal authorities rely on a variety of regulatory and voluntary policies and programs (yellow box 5 in figure 2.1) for both point and nonpoint sources. Jurisdictions develop, and EPA reviews, WIPs that describe the specific types of practices that will be used to meet the TMDL.

For point sources, jurisdictions impose numeric nutrient effluent limitations on municipal and industrial WWTPs above a certain size under the CWA permitting program. Urban nonpoint source loads are the fastest growing category of nutrient and sediment loads. To reduce these loads, several state governments established numeric nutrient and sediment permits for municipal separate storm sewer systems (called MS4s). These permits are unique in that states place numeric limits on urban nonpoint source loads. Numeric MS4 permits represent a significant departure from traditional MS4 permits which were based on narrative rather than numeric requirements. Agricultural nonpoint source pollution is the single largest contributor of nutrient and sediment loads to the Bay, but the CWA explicitly prohibits federal permitting requirements for controlling most agricultural runoff. The jurisdictions rely on a mix of education, financial assistance, and technical assistance programs to induce agricultural producers to voluntarily adopt BMPs. States operate permitting programs for certain agricultural operations such as concentrated animal feeding operations (CAFOs) over a certain size.

The CBP uses the CAST watershed model to inform development of WIPs, credit progress toward meeting TMDL nutrient and sediment load reduction targets, and track implementation progress (Hood et al., 2021). State and local jurisdictions’ TMDL obligations for 2025 are considered met when CAST calculates that enough pollutant reduction practices have been installed to meet the TMDL, not by whether nutrient and sediment delivery to the Bay reaches a specific level. The presumption is that if CBP partners implement needed nutrient and sediment pollution control measures as estimated by CAST by 2025, then pollutant reductions called for under the TMDL will be realized at some unspecified future date. The difference between when pollutant control measures are installed and when pollutant reductions are achieved is attributed to the time lag associated with nutrient and sediment movements through the watershed and to the Bay.

For nonpoint sources, EPA expects the states to offer “reasonable assurance” that nonpoint load reductions will be achieved. A state offers reasonable assurance by including in its WIP only those land use practices (BMPs) that meet three criteria: the practices must (1) exist, (2) be technically feasible at a level required to meet allocations, and (3) have a high likelihood of implementation.

The CBP also implements a version of adaptive management for the CBWA outcomes through a decision framework (CBP, n.d.-b). This process is intended facilitate decision-making under uncertainty by continuously monitoring and evaluating progress toward achievement of specific CBP program outcomes and adjusting implementation based on these assessments. The CBP’s Goal Implementation Teams (GITs) are responsible for implementing the decision framework. More generally, the GITs are responsible for overseeing and promoting the implementation of plans to achieve the goals of the CBP. The implementation of the decision framework at the GIT level is articulated in the 2-year cycle Strategy Review System (SRS) that establishes the link between the factors that could impact the partnership’s ability to achieve an outcome and the actions it is taking to manage them. For water quality goals under the CBWA, the CBP partnership has an accountability process for the Bay TMDL that is described in the management strategy for the 2025 WIP and standards attainment outcome (CBP, 2020). While these outcomes come under the SRS adaptative management reviews, the TMDL accountability process has additional steps to address regulatory requirements.

2.2. Evaluating system response to nutrient and sediment control efforts

The overarching goals of this report are to evaluate how the Chesapeake Bay system and its physical, biological, and social subsystems are responding to water quality improvement efforts and to help identify ways that system response to management efforts could be improved, when appropriate. To conduct this evaluation, STAC relied on both the state of current knowledge as reflected in the scientific literature and review of portions of extensive CBP data. In addition, STAC has evaluated many of the issues in previous workshops, reports, and requested reviews, and this accumulated body of work is a critical part of this overall evaluation. Based on this body of evidence, this report discusses the gaps and uncertainties associated with watershed, estuary, and living resource system response.

Federal, state, and local implementation policies are intended to change people’s behavior in ways that produce nutrient and sediment reductions (orange graph in figure 2.5). The solid orange line, labeled expected response, represents the CBP’s predicted nutrient and sediment reductions in response to implementation of management actions and is largely derived from

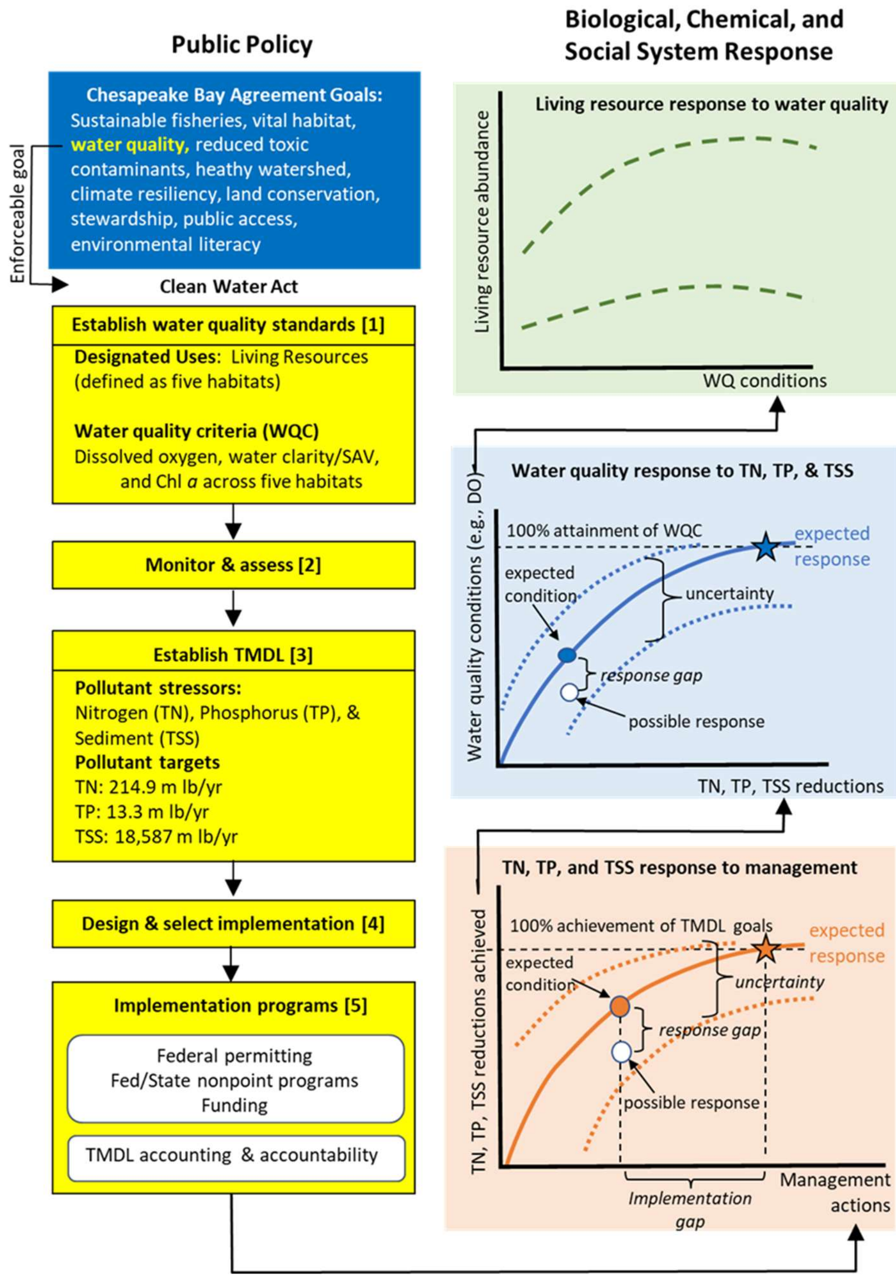


FIGURE 2.5.—Conceptual representation of uncertainty and gaps in system response to Chesapeake Bay water quality policy.

the CAST model. The orange star represents 100% achievement of the TMDL N, P, and sediment targets. Conceptually, the pollutant reduction response to management actions (i.e., practices and technologies installed) eventually begins to diminish (the orange line begins to flatten), indicating that more effort will be required to achieve each additional gain in pollutant reduction (i.e., the actions that produce large reductions in pollutant loads have already been undertaken). For example, an initial upgrade of a single WWTP might result in a one-million-pound reduction in N, but achieving subsequent reductions from nonpoint sources could require treating tens of thousands of acres of farmland or the installation of tens of thousands of urban stormwater BMPs.

As the 2025 TMDL deadline approaches, possible ***implementation gaps*** exist. For purposes of this report, an implementation gap is the difference between current implementation (expected current condition) and the implementation goal (desired condition). Conceptually, an implementation gap is how much additional implementation is needed to achieve the pollutant reduction goal (under the assumption that the effectiveness of the implementation is known and correct). In the orange graph in figure 2.5, the orange dot represents the expected N, P, and sediment reduction response given the level of implementation achieved to date. The horizontal distance between the orange dot and the star represents the implementation gap that impedes achievement of the TMDL targets.

Reductions in nutrient and sediment loads to the Chesapeake Bay are expected to produce improvements in Bay water quality (light blue graph in figure 2.5), specifically defined by the DO, water clarity, and Chl *a* criteria. Estuary science and modeling of the relationships between these criteria and water quality produce expected water quality response to reductions in nutrient and sediment loads (solid blue line in figure 2.5). The implementation gap, as defined above, affects the achievement of the WQS (distance between blue dot and star). The shape of the curve represents the expectation that the initial increments of nutrient and sediment reductions are expected to bring more areas (segments) of the Chesapeake Bay into compliance than later increments. In other words, larger pollutant reductions will be needed to bring the remaining few segments into attainment. For instance, CBP modeling indicates that millions of additional pounds of nutrient reductions are needed to achieve DO criteria in parts of four (out of 92) segments (deep waters and deep channel habitats in CB3MH, CB4MH, CB5MH, and POTMH segments [USEPA, 2010]). Complete (100%) achievement of the DO criteria is indicated by the blue star.

Finally, there is a general public expectation that as progress is made toward achieving the WQS, living resources, particularly specific species of fish and shellfish, will increase in abundance (increasing response on the green graph in figure 2.5). How living resources will respond to improved oxygen levels and improved water clarity is not specifically described by the WQS or by the CBP, however. As noted above, the WQC were designed to avoid specifying a living resource outcome because of the numerous other factors that affect living resources besides water quality. Thus, there is no formal expected response defined by the CBP between improving water quality conditions and living resources (no expected response curve in figure 2.5). The green graph shows possible living resource responses to improving water quality conditions (discussed in more detail in chapter 5). The possible living resource responses represented in figure 2.5 can be thought of as the key benefit responses to the costs incurred in reducing Bay pollutant stressors in the orange graph.

Achieving pollutant reduction and water quality goals will also be challenging if emerging scientific data and analyses find evidence that biological, physical, and behavioral outcomes do not match expected/predicted outcomes, called a ***response gap*** in this report. For example, experts may estimate the average nutrient removal effectiveness of specific BMPs to be 75%, but emerging scientific evidence may strongly suggest that overall BMP effectiveness is less than 75%. An example of such a response gap is illustrated in the orange graph of figure 2.5 as the difference between a realized outcome represented by the white dot and the expected outcome represented by the orange dot. While a response gap could be positive or negative, this report is most concerned with situations when actual response appears to be falling short of expected (predicted) response. Such a response gap means that any given level of management effort is producing less reduction than is being claimed under the TMDL and less improvement in Bay water quality and living resources.

Uncertainty in response

Uncertainty surrounds how management actions translate into pollutant reductions, water quality improvements, and changes in living resources and makes evaluating the causes of implementation and response gaps challenging. Translating policy into appropriate behavioral change, nutrient and sediment load reductions, and water quality response is a complex biological, physical, and social process. Implementation policy generates changes in the behaviors of large numbers of people that are often difficult to observe and sometimes generates unintended behavioral consequences. The pollution reduction effectiveness of people's actions is based on a wide variety of factors including the removal effectiveness of management actions, BMP maintenance, soil conditions, slope, vegetative cover, distance to water, and weather conditions that make assessing pollution control effectiveness challenging. Nutrients and sediments in the Bay interact with changing physical habitat, biological activity, and climate conditions to produce changes in DO and water clarity. Living resources respond to these conditions, but the size of that response is influenced by other changes that affect fish and shellfish abundance, such as water temperature, habitat, and harvest rates. In other words, the causal chain from TMDL policy to living resource response is long and contains factors that are occurring simultaneously, not always easily observable or measurable, and often outside of management control.

Thus, this report emphasizes that Bay water quality management decisions will always be made under uncertainty. The dotted lines above and below the expected response lines in figure 2.5 illustrate that the actual (realized) system response is not precisely known. Sources of uncertainty exist throughout the system, including, for example, how people respond to programs designed to improve water quality (e.g., incentive programs that promote BMP implementation), the effectiveness of load reduction control practices in specific locations, how water quality measures at different places in the system will respond to changing nutrient loads, and how different living resources will respond to changes in estuarine water quality.

Two general sources of uncertainty are relevant to making decisions about water quality policy goals. The first is natural variation or stochasticity (i.e., aleatory uncertainty). Weather and climate are examples where natural variation generates fluctuations in nutrient loads, estuarine water quality responses, and living resource population dynamics, and are difficult to account for. Characterizing natural variation and, to the best of our ability, incorporating it into management decisions are critically important in accounting for uncertainty.

Uncertainty associated with missing, incomplete, or imperfect knowledge (i.e., so-called epistemic uncertainty) is distinguished from uncertainty associated with natural variation. Epistemic uncertainty describes the limits to scientific or technical understanding of relationships or processes impacting water quality. For example, there may be uncertainty about the relative magnitude of N removal pathways (e.g., for surface water, groundwater, or atmospheric sources), the effectiveness of BMPs in different settings, or people's nutrient use behavior. In other cases, an underlying conceptual relationship may be well understood, but data availability or modeling capacity presents a barrier to applying the understanding more broadly.

Both types of uncertainty are contained in the representation of the dotted lines in figure 2.5. The dotted lines are illustrative because epistemic uncertainty presents challenges with even identifying the precise bounds around expected responses.

In this report, we also strive to delineate decision-relevant uncertainties rather than all uncertainties. A decision-relevant is defined as an uncertainty that, if resolved, may change a management or policy decision (e.g., what pollutant control technology to implement or how to express water quality goals). The restoration effort will always require decision makers to make decisions under uncertainty and assess the results in order to learn, meaning that we knowingly recognize that one kind of risk (i.e., that our selected management and policy actions may need to be improved or revised) is being accepted to avoid another (i.e., the outcomes of continuing to make choices with incorrect information). Assessing and managing this balance requires that we formally assess the efficacy of our actions and their unintended consequences.

Uncertainty is an inescapable reality of future Bay water quality management (Hershner, 2011). Recognizing the uncertainty associated with expected responses or any gaps is not a reason to delay efforts to improve water quality. Rather, recognizing response gaps and their potential causes, as well as the uncertainty that surrounds all expected responses, offers opportunities to improve our understanding of how the system works, improve policy effectiveness, and accelerate progress toward meeting goals. With this understanding, recognizing and reducing uncertainty and gaps and, in response, changing water quality management and policy decisions (yellow boxes in figure 2.5) are critical to future program success.

3. Nutrient and Sediment Response to Management Efforts

A key issue confronting the CBP as the 2025 TMDL implementation deadline approaches is whether the physical and social systems are responding to implementation in ways sufficient to meet nutrient and sediment reduction targets. This chapter addresses the following questions:

- Is the physical-social system responding to management efforts to meet TMDL N, P, and sediment targets in ways consistent with expectations?
- What are possible gaps and uncertainties confronting efforts to reduce N, P, and sediment delivered to the Chesapeake Bay?

To date, investments in wastewater treatment have produced tens of millions of pounds of point source nutrient reductions, and atmospheric deposition of N has decreased steadily for several decades. These are noteworthy achievements (Lyerly et al., 2014). Despite these successes, achieving the Bay TMDL depends primarily on additional agricultural and urban nonpoint source load reductions.

Achieving nonpoint source reductions has proven more challenging than anticipated when the first nutrient reductions targets were established in the early 1990s. The challenge is twofold. First, voluntary nonpoint source programs struggle to produce the scale of behavioral change and practice adoption necessary to achieve water quality goals, i.e., an implementation gap. Second, the nonpoint source programs and practices implemented may not be as effective as expected at reducing nonpoint source pollution, i.e., a response gap.

Two key objectives of this report are to (1) identify the gaps and uncertainties that inhibit TMDL attainment and to (2) identify management options that can improve pollutant control efforts in the face of uncertainty. Existing policies have limited capacity to induce the behavioral change required to meet TMDL nutrient targets. Meeting TMDL targets will require new approaches to nonpoint source management. How much additional nonpoint source reduction can be achieved and at what cost is uncertain. A major challenge confronting the CBP partnership is addressing a complex system in the face of great uncertainty (Freedman et al., 2008; Hershner, 2011; Shabman et al., 2007).

3.1. Nonpoint source implementation policies in the Chesapeake Bay watershed

Assessing the effectiveness of nonpoint source programs requires an understanding of the CBP nonpoint policies themselves. The CBP approach to nonpoint policies is built on two foundations. First, Chesapeake Bay jurisdictions implement programs designed to induce agricultural and urban land managers (including landowners and those who manage lands) to change behavior to reduce nutrient and sediment loads. Outside of MS4s and CAFOs, which are regulated to various extents, these programs primarily rely upon incentives to encourage land managers to voluntarily adopt BMPs. These BMPs are intended to reduce nutrient inputs, retain nutrients and sediment on the landscape, or transform nutrients into less damaging forms (e.g., biologically available forms of N converted to inert N). The second foundation is the CBP TMDL modeling and accounting system that calculates pollution reductions and assigns pollutant reduction credit to state and local jurisdictions for implementing BMPs. The two foundations of nonpoint source policy are not independent. The structure of the TMDL modeling and accounting system influences people's incentives and behavior.

Agricultural nonpoint source policy generally allows land managers to decide how to manage their operations and whether or how to reduce nonpoint source pollution (Pannell & Claassen, 2020; Shortle et al., 2021). Information and education programs inform agricultural managers about BMPs and the availability of government funding to encourage their adoption. Because most BMPs are costly, and in many cases reduce producer incomes (i.e., private BMP costs outweigh private BMP benefits), state and federal programs cover a portion of the costs for implementing BMPs (Shortle et al., 2021). Most state and federal financial assistance takes the form of cost sharing, which typically pays for a portion of the cost of BMP installation and, in limited circumstances, some annual operation and maintenance costs (Ribaudo, 2001; Ribaudo & Shortle, 2019).

For over three decades, federal and state governments have been committed to funding agricultural financial assistance programs to encourage BMP implementation in hopes of meeting nonpoint source reduction targets. Federal and state efforts have been successful at increasing funding to support these programs.

In most states, urban nonpoint source programs assign municipalities and larger industrial sites responsibility for reducing N, P, and sediment loads from urban lands under their jurisdiction through MS4 permits (details differ by states).[1] Municipal separate storm sewer systems can meet these requirements using a variety of BMPs. They can implement BMPs on public land, often upgrading existing stormwater infrastructure and stream restoration projects, or in private developments using common stormwater BMPs (Gonzalez et al., 2016). Given the competition for urban land, BMP implementation opportunities are often limited and exceedingly expensive.

The CBP CAST model credits nonpoint source load reductions by either BMP nutrient removal efficiencies, land use change, or nutrient source reductions. The CAST model calculates nutrient load reductions via BMP removal efficiencies by multiplying model-based estimates of average nutrient runoff (lb/ac) by an assigned BMP removal efficiency (e.g., practice reduces runoff by 60%) and the number of acres treated by the BMP (see figure 3.1). Expected runoff is based on average precipitation over a baseline 3-year period. The model calculates nutrient loads as an average over relatively large areas of approximately 20,000 acres, called land-river segments, for different land use types within each segment (e.g., row crop production, hay production, etc.). These area averages are based on estimates of commercial fertilizers (purchases), animal manure production (based on estimates of livestock numbers), atmospheric N deposition

[1] States also operate stormwater permit systems for industrial facilities. Some, but not all, urban stormwater loads originate from lands managed by permitted sources.

(provided by an airshed model and reanalysis of observations), and segment physiography. The CAST model estimates long-term loads delivered to the Bay.

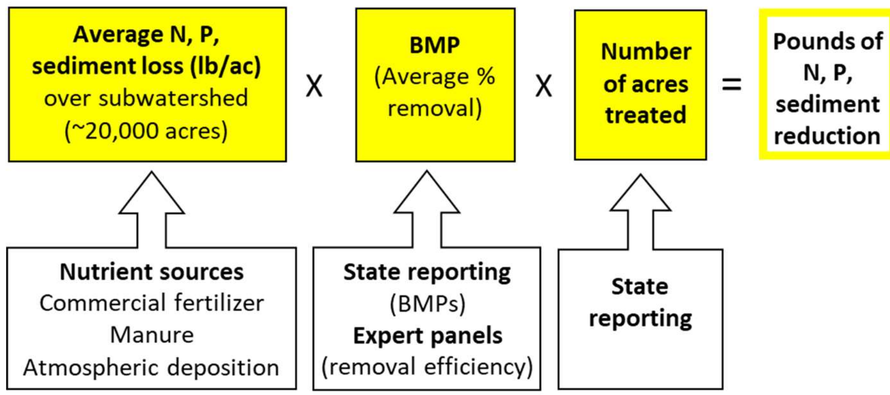


FIGURE 3.1—Chesapeake Bay Program modeling of a typical nonpoint source BMP load reduction.

The BMP pollutant removal efficiencies are generally a single number (e.g., 30% N removal for a vegetative buffer strip) averaged across the watershed or very large geographic regions. BMP removal efficiencies are generated based on input from a group of subject area experts (Stephenson et al., 2018). The CBP currently lists nearly 300 specific agricultural and urban BMPs, although several BMPs include a number of variations (e.g., CBP lists over 100 cover crop BMPs). BMPs must be vetted by the expert panels and approved by the CBP before any nutrient or sediment reductions can be assigned for implementation of these practices. To get credit toward TMDL compliance, jurisdictions report the number and types of BMPs installed and the number of acres treated by the BMPs.

3.2. Nutrient nonpoint source pollutant response to TMDL implementation policy

Implementation gap

The current TMDL planning targets for the Phase III WIPs for N, P, and sediment are 214.9, 13.3, and 18,587 million lb/yr, respectively.[2] The expected N and P response to implementation of practices for the TMDL is shown in figure 3.2. As of 2021, CAST predicts that sufficient practices have been installed to meet the sediment goal and 91% of the needed P reductions to achieve TMDL load targets (from a 1985 baseline). Figure 3.2 illustrates the N reduction goal is the most difficult pollutant target to achieve; CAST estimates that only 73% of the N reduction has been achieved. Note: figure 3.2 does not reflect any lag times between BMP implementation and observed changes in pollutant loads.

---

[2] The N load target assigned to the Chesapeake Bay jurisdictions is 199 million lb/yr, with the difference between 199 and 214.9 attributed to atmospheric N deposition.

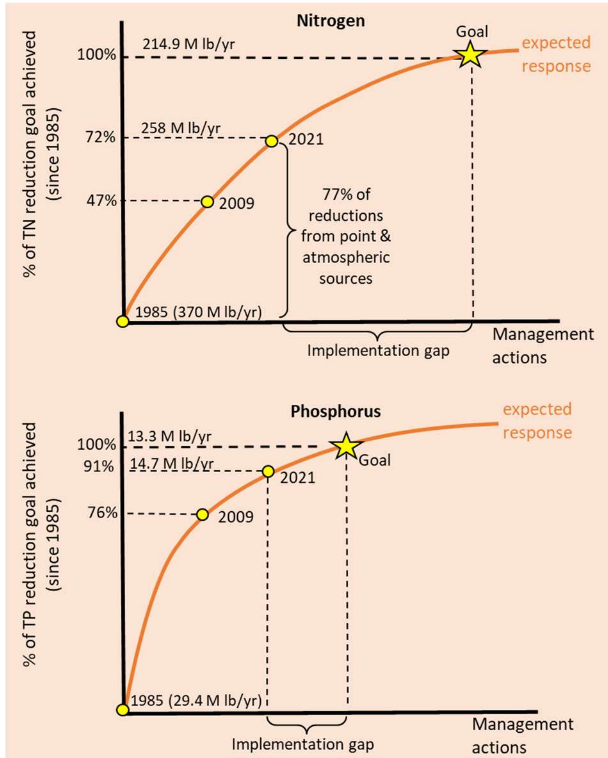


FIGURE 3.2.—Progress toward meeting TMDL nutrient reduction goals and expected response to watershed management efforts as calculated by the CBP CAST model (Source: CBP, n.d.-a).

A substantial implementation gap exists for N. In 2021, the difference between expected N reductions and the TMDL target was estimated by CAST to be about 43 million lb/yr (see figure 3.2). Closing this N implementation gap will be difficult. Since 1985, the CAST model estimates that annual N loads to the Bay have been reduced by 112 million pounds (370 to 258 million lb/yr, see figure 3.2). Point source WWTPs and reduced atmospheric deposition of N are responsible for three-quarters of those reductions. Since most large WWTPs are operating at or near the limits of technology, point sources can make only a modest contribution to closing the N implementation gap. As of 2021, point sources and atmospheric sources account for only 18% of the total N load to the Bay. The remaining N load is from either nonpoint sources (64%) or natural background sources (18%).

To what extent the N implementation gap can be closed depends largely on reducing nonpoint sources, and nonpoint source reductions have historically been difficult to achieve. The CAST model calculates that it took 36 years to reduce agricultural and urban nonpoint source N loads by 27 million lb/yr (1985 through 2021). Recently, nonpoint source reductions have been more difficult to achieve. Despite concerted efforts since the TMDL was adopted in 2010, the CAST model estimates that the annual amount of nonpoint source N loads reaching the Bay has been reduced by only 3.5 million pounds during the period 2009–2021. The 3.5-million-pound nonpoint source reduction represents just 9% of the total 40 million lb/yr of N reductions achieved over the same period from all sources.

Finally, the N implementation gap is an underestimate of the total reductions needed to meet the TMDL. The 43 million lb/yr gap does not include additional needed reductions associated with the infill of Conowingo reservoir (roughly 6 million lb/yr of N) or additional reductions needed to meet the DO criteria because of climate change (initial estimates of 9 million lb/yr) (Shenk et al., 2021). In addition, the CBP has also recently discovered unaccounted for sources of nutrients (i.e., undercounting millions of animals and missing fertilizer sales, discussed below) that add millions of pounds of additional N loads to CAST estimates (Blankenship, 2022).

Response gap

Estimates of achieving TMDL targets generally assume that expected pollutant responses to BMP implementation are accurate. Given the complexity and uncertainties of reducing pollutant loads across a large watershed, this will rarely be the case. A critical technical and policy question to address is the extent to which nonpoint source policy and management actions are as effective as expected. Emerging evidence suggests that nonpoint source management actions have had mixed success in reducing nutrient and sediment loads and that response gaps may exist between expected and observed pollutant reductions (Ator et al., 2019, 2020; Chanat & Yang, 2018; Fanelli et al., 2019; Fisher et al., 2021; Keisman, Blomquist, et al., 2018; Kleinman et al., 2019; Moyer & Blomquist, 2017; Noe et al., 2020; Roland et al., 2022). As illustrated in figure 3.3, a response gap would mean that the realized nonpoint source response (red curve) might be less than what is expected by the CBP CAST estimates (orange curve). Also, the magnitude of the possible response gap could differ by pollutant. Evidence suggests that the possible response gap for P is particularly large. A response gap would imply that additional management effort and implementation would be needed to achieve any given level of pollutant reduction.

Analysis of ambient water quality trends suggests the possibility of nonpoint source response gaps (Easton et al., 2023). The U.S. Geological Survey (USGS) monitors and analyzes water quality through a series of monitoring networks in the watershed. Table 3.1 and figure 3.4 illustrate the long-term (1985–2021) loads and trends in TN, TP, and TSS at nine river input monitoring (RIM) stations at or near the mouth of major Chesapeake Bay tributaries. The CAST model estimates that these nine tributaries contribute roughly two-thirds of the nutrients to the Bay, while point and nonpoint sources discharging into the tidal portions of the Bay comprise the other third. The numbers in table 3.1 are the average annual loads over the 1985–2021 time period; the colors indicate trends over those time periods. For example, the yellow shading for TP in the Susquehanna River at Conowingo indicate no detectable trend. The green shading for the Potomac illustrates the declining trends for TN, TP, and TSS. Long-term trends show N loads declining in the four largest tributaries, but P loads are decreasing in just three of the nine tributaries. Figure 3.4 shows flow-normalized TN, TP and TSS load estimates (gold dots) and trends (black lines) on a per acre basis. The tributaries that show the most consistent and sustained decreasing trends in nutrient loads are the Potomac and Patuxent rivers. These two tributaries also had the highest initial proportion of nutrient loads coming from point sources. Tributaries where point source nutrient loads comprise less than 10% of total nutrient loads

(Choptank, Appomattox, Mattaponi, Rappahannock, and Pamunkey) generally show increasing or no trend in nutrient loads over the period 1985–2020. Similar mixed results can be found throughout the watershed (Hyer et al., 2021).

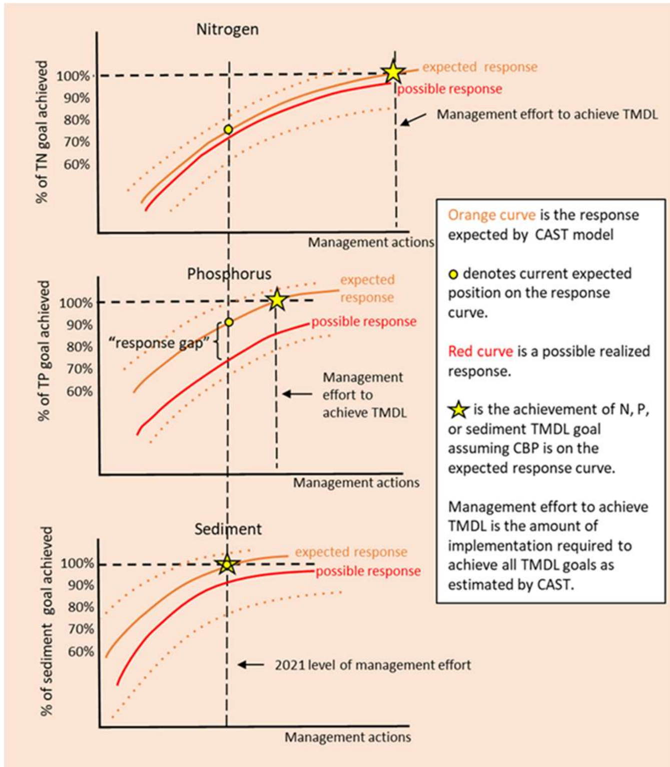


FIGURE 3.3.—Conceptual illustration of possible response gaps and response uncertainty in achieving TMDL nutrient and sediment targets.

TABLE 3.1—Average load (million lb/yr, 1985–2021) and long-term flow-normalized trends in load of TN, TP, and TSS at RIM stations (Source: Mason & Soroka, 2022).

	TN	TP	TSS
Susquehanna River at Conowingo MD	135	5.44	3,533
Potomac River, Chain Bridge at Washington, DC	48.6	3.27	2,452
James River at Cartersville, VA	11.0	2.21	1,543
Rappahannock River, near Fredericksburg, VA	4.32	0.64	478
Appomattox River at Matoaca, VA	1.46	0.14	39.4
Pamunkey River near Hanover, VA	1.41	0.16	85.0
Mattaponi River near Beulahville, VA	0.65	0.06	14.8
Patuxent River near Bowie, MD	1.51	0.11	49.0
Choptank River near Greensboro MD	0.55	0.04	5.12
Green shaded cells indicate long-term declining loads; red shaded cells indicate increasing loads; and yellow shaded cells indicate no statistical trend in loads.			

Analysis of ambient water quality trends generally shows mixed evidence of nonpoint source pollution reduction effectiveness. Using the empirical Spatially Referenced Regression on Watershed Attributes (SPARROW) model that relates ambient pollutant levels to landscape characteristics and nutrient sources, Ator et al. (2019) found little evidence that agricultural nonpoint source loads declined between 1992 and 2012. Another statistical analysis of monitoring data found that while P loads were declining in some regions of the Bay watershed, those improvements were offset by increases in agricultural P sources in other areas (Fanelli et al., 2019). A STAC workshop report summarized: "Current research suggests that the estimated effects of conservation practices [BMPs] have not been linked to water quality improvements in most streams" (Keisman, Blomquist, et al., 2018, p. 9).

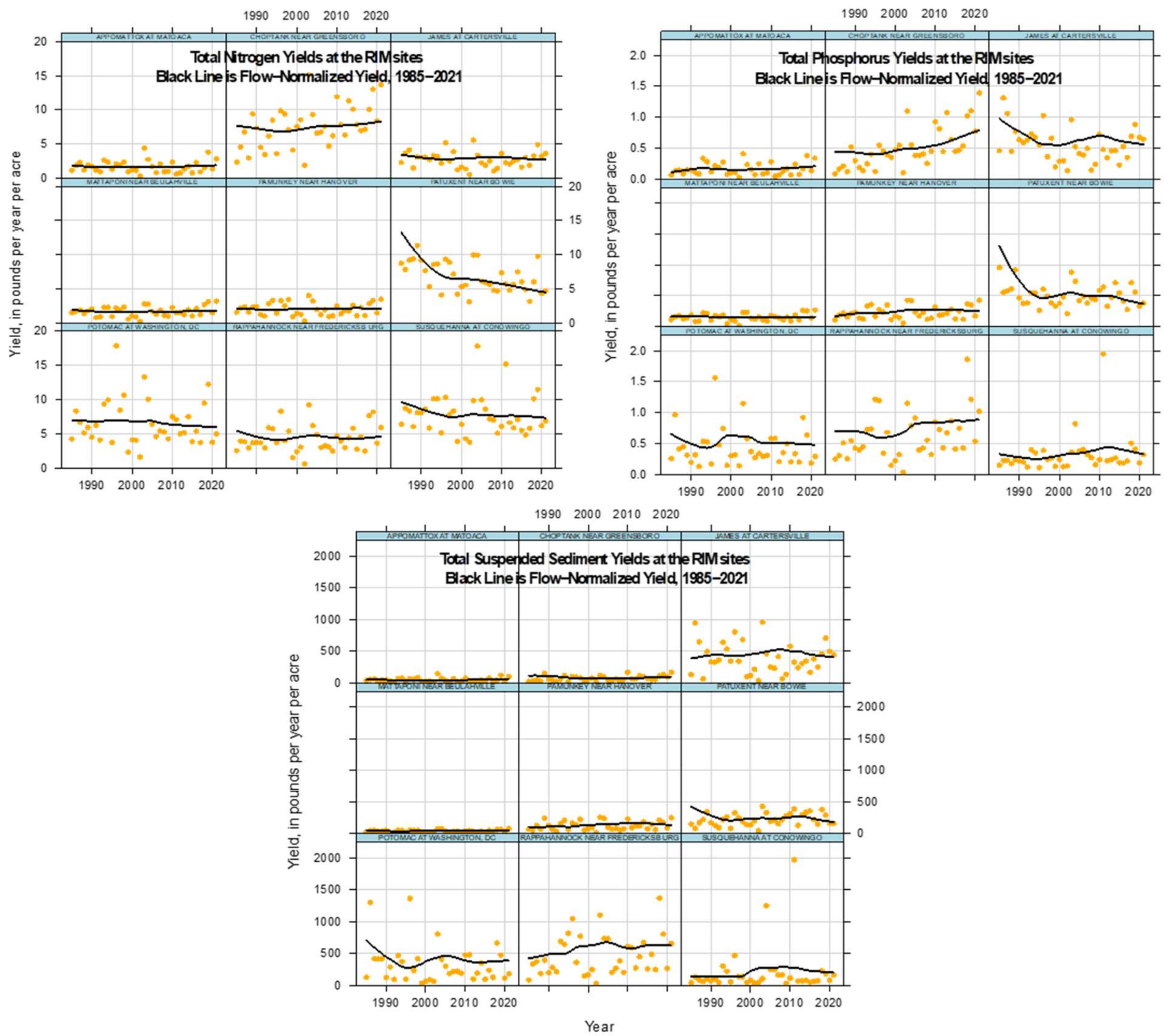


FIGURE 3.4.—Flow-normalized TN, TP, and TSS yields at RIM stations for the period 1985–2020. Figure created by Chris Mason, US Geological Survey Virginia and West Virginia Water Science Center, Richmond, VA (Data source: Mason & Soroka, 2022).

Figure 3.5, adapted from Ator et al. (2020), presents a comparison between CAST estimates of delivered N and P loads and an empirical analysis using SPARROW. CAST estimates and SPARROW's empirically-driven estimates generally agree that total N delivery to the Bay declined between 1992 and 2012, driven to a large degree by reductions in point source loads. Phosphorus loads estimated by CAST, however, differ markedly from empirical SPARROW estimates. CAST suggests consistent reductions in delivered P loads, with the reductions occurring across most source sectors. SPARROW suggests that delivered P loads were increasing and that those increases were largely attributable to both agricultural and urban nonpoint source pollution (Fanelli et al., 2019).

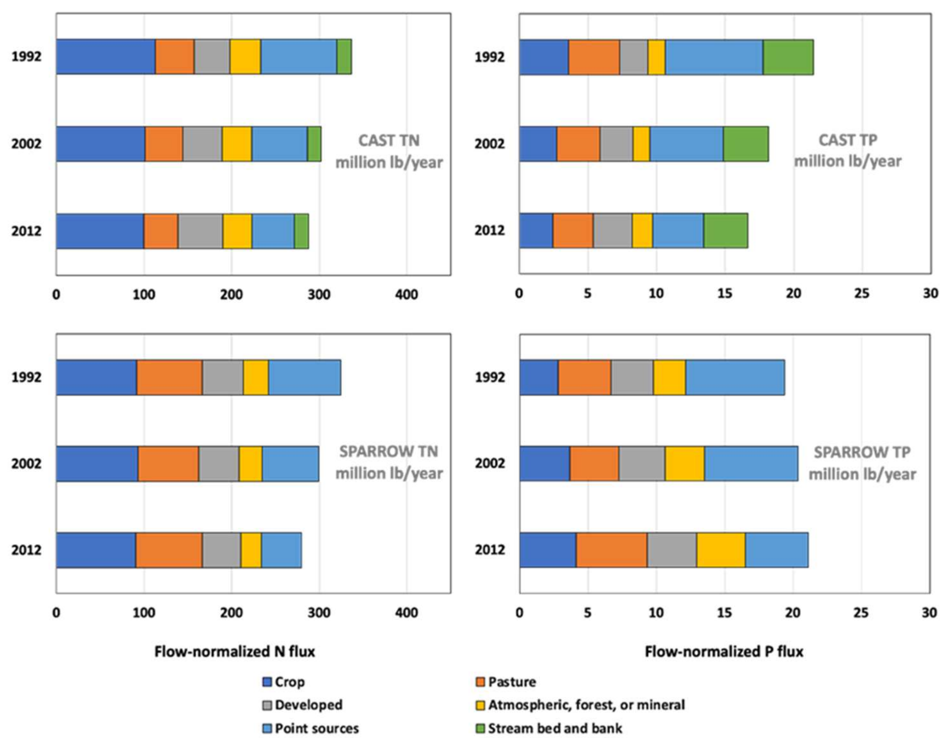


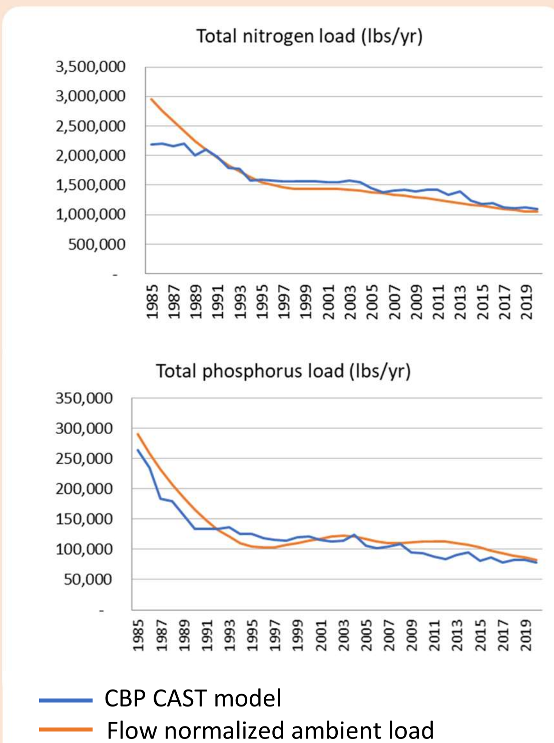
FIGURE 3.5.—Flow-normalized total and source sector TN and TP fluxes to the Chesapeake Bay for 1992, 2002, and 2012 estimated with the Chesapeake Bay CAST model and SPARROW model (Source: adapted from Ator et al. [2020] by Easton et al. [2023]).

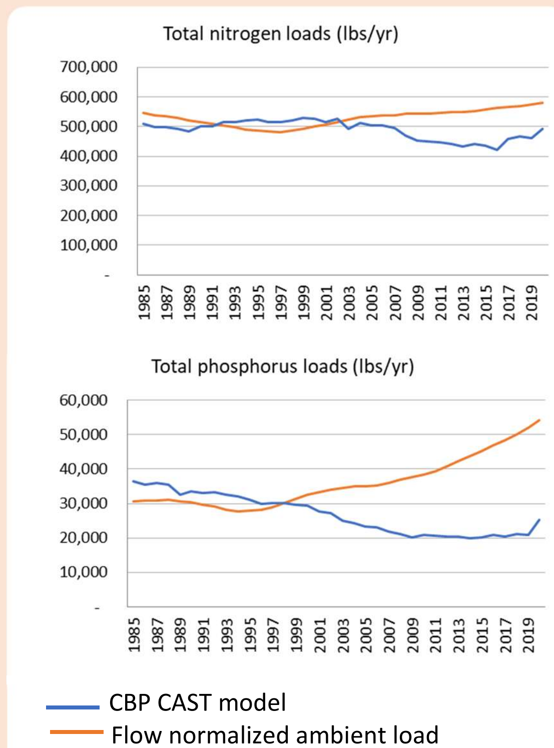
The P results have policy-relevant implications because state and local management decisions designed to meet the WQS are based on CAST results. Program managers are currently focused largely on the N implementation gap, but analysis of ambient monitoring data suggests significant response gaps exist, particularly with respect to P reduction goals.

Text box 3.1 illustrates nutrient trends and expected responses for two watersheds, the Patuxent and Choptank. In 1985, point sources were the dominant source of nutrient loads in the Patuxent while nonpoint sources contributed the vast majority of nutrient loads in the Choptank. Patuxent has seen large reductions (exceeding 80%) in point source nutrient loads since 1985, and these reductions are reflected in the declining nutrient loads measured in the river and a close correspondence between expected (modeled response, blue line) and observed (orange lines) outcomes. Nutrient loads have generally increased or remained relatively constant in the Choptank, despite considerable effort to reduce nonpoint source loads (Fox et al., 2021). The Choptank experience also illustrates the uncertainties and challenges for realizing pollutant reduction in many nonpoint source-dominated watersheds.

Text box 3.1. Watershed nutrient load trends (1985–2020) in a point and nonpoint source dominated watershed, CAST model and flow normalized ambient estimates

Patuxent	**Choptank**
Point source-dominated	**Nonpoint source-dominated**
Point source share of 1985 load: 45-55%	Point source share of 1985 load: <1%





Trends

Steadily declining N and P loads, driven by 80%+ reduction in point source loads since 1985	Observed nutrient loads (orange line) constant or increasing since 1985

Response gap

Close correspondence between CBP CAST model response (blue line) and observed outcomes (orange line)	Differences between CBP CAST model response (blue line) and observed outcome (orange line), particularly for P

(Source: Data provided by Chesapeake Bay Program)

The information in text box 3.1 also shows a large potential P response gap in the Choptank. The CBP CAST model estimates that P loads have declined since 1985, but ambient water quality monitoring and the associated trends indicate that P loads have increased steadily since the mid-1990s.

Uncertainty also surrounds attainment of sediment targets. While CAST estimates that sediment load reduction targets have been achieved, long-term trend data (based on monitoring) in most tributaries indicate increasing sediment loads since 1985 (table 3.1), although more recent trends (2011–2020) indicate some tributaries are improving with respect to sediment goal attainment. Previous reports have concluded that sediment itself is not the primary contributor to Bay water quality problems, and nutrients should be the focus of pollution reduction efforts (Miller et al., 2019).

A variety of explanations for response gaps exist. One possible explanation is lag times, which is the time between when nutrient and sediment reduction efforts (i.e., BMPs, reduced nutrient applications, etc.) are initiated and when pollutant reductions are detected through monitoring. Some BMPs may take years or even decades to fully deliver pollutant load reductions to the Bay given both the time required for some BMPs to become fully functional (e.g., riparian buffers) and the time required for nutrients and sediment to migrate from their point of origin to the Bay (Easton et al., 2023). The CBP readily acknowledges, and scientific studies support, that lag times influence water quality response to reduction of nonpoint sources (STAC, 2013). One analysis, using the CBP's dynamic watershed model, estimated that a considerable portion of the N response gap could be attributable to lag times. Lag times explain less of the P response gap, providing evidence that simply waiting will not eventually generate the reductions expected from management actions (Shenk et al., 2022).[3] Response gaps, however, could also be caused by incomplete or incorrect understanding and representations of behavioral and physical responses to management efforts or the limitations of the existing monitoring system to accurately reflect trends that may be occurring (Ator et al., 2020).

The challenge of realizing and sustaining large reductions in nonpoint source loads is not unique to the CBP. Studies of individual BMPs, or studies of BMP nonpoint source reduction efforts conducted at a fine scale (i.e., edge of field, headwater basin) with intensive monitoring, have shown BMP implementation can reduce nutrient and sediment loads to streams (Ator et al., 2020; Böhlke & Denver, 1995; Denver et al., 2010; Jefferson et al., 2017; Li et al., 2017; Staver & Brinsfield, 1998). Demonstrating the effectiveness of nonpoint source control efforts at larger watershed scales has proven more difficult (Lintern et al., 2020; Osmond et al., 2012; Sprague & Gronberg, 2012; Tomer & Locke, 2011). A recent synthesis review of agricultural conservation programs concluded that "there has been little evaluation of the incentives-adoption-outcome chain: that is, how well different incentives promote adoption, whether adoption leads to

[3] Note, the ability to accurately model lag times is uncertain. Hood et al. (2021) recommended improvements in the CBP watershed model to better account for lag times.

meaningful and measurable changes in outcomes, and what factors shape these links" (Pineiro et al., 2021, p. 1).

3.3. Assessment of gaps and uncertainties in efforts to reduce nonpoint source loads

Understanding the possible reasons why nonpoint source reductions have proven so challenging is of critical importance (Ator et al., 2020; Easton et al., 2023). Three possible causes that can delay or inhibit achievement of the CBP TMDL nutrient and sediment reduction targets are: (1) lag times, (2) response gaps, and (3) implementation gaps. If lag times are the sole reason for limited observed nonpoint source response, then jurisdictions can simply wait until an unspecified future date for the expected reductions to be observed in the watershed and the Bay. However, improved understanding of system responses to nonpoint source implementation programs could enhance effectiveness of pollutant control efforts. The presence of response gaps suggests that just focusing on installing BMPs to meet TMDL targets does not ensure actual reductions of nutrient and sediment loads. Implementation gaps also inhibit future reductions and achievement of TMDL targets. Even if response gaps are eliminated, nonpoint source programs have yet to induce the type and scale of practice implementation needed to achieve the TMDL.

The uncertainty that characterizes such a complex system complicates our ability to diagnose the multitude of possible explanations for the limited observed response in nonpoint source loads. Figure 3.6 illustrates the major elements in the Chesapeake Bay watershed system that

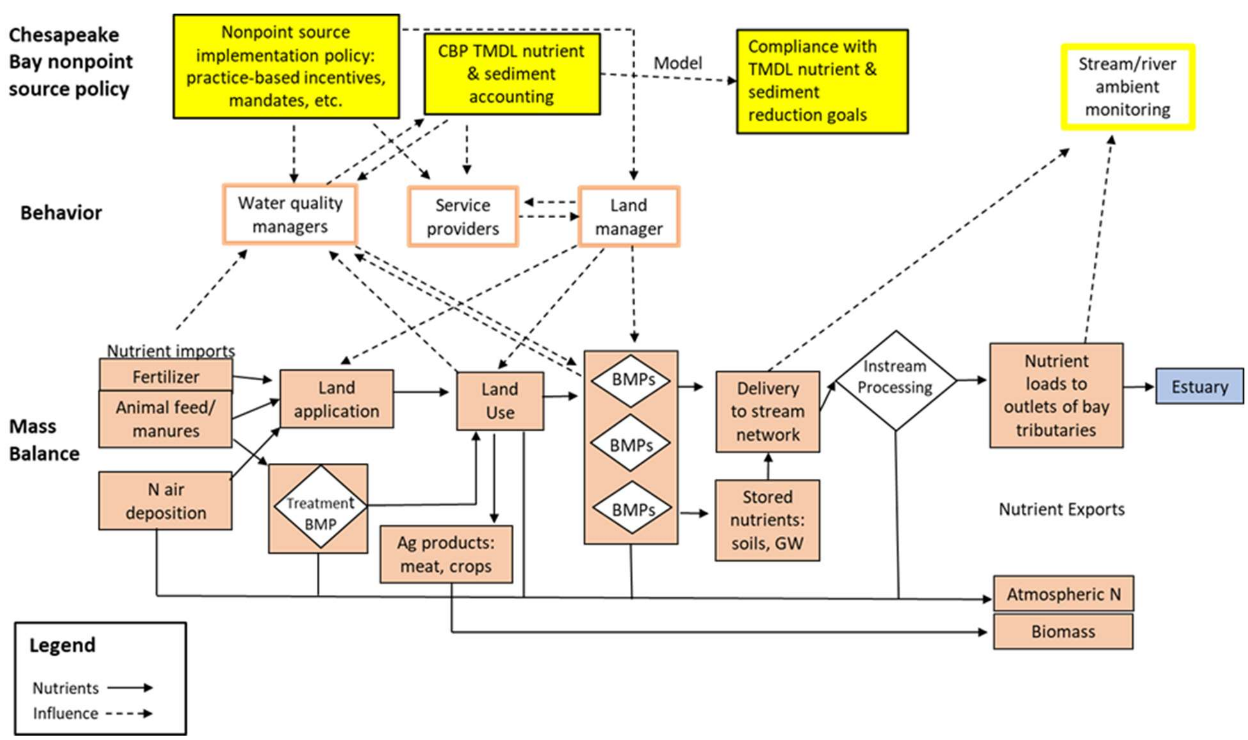


FIGURE 3.6.—Behavioral and physical system response to Chesapeake Bay nonpoint source control policy.

may be characterized by lag times, response gaps, and implementation gaps. In figure 3.6, CBP nonpoint source implementation policies (yellow boxes) are designed to induce behavioral changes (open orange boxes) that will ultimately reduce nutrient and sediment loads delivered to the physical system of the Bay (shaded orange boxes). Behavioral changes are expected among agricultural and urban land managers as well as federal, state, and local program staff (water quality managers and technical conservation service providers) who are tasked with TMDL implementation. The physical system is represented by a simplified watershed nutrient mass balance portion of the diagram.

Nutrient mass balances quantify the (1) inputs to a system (e.g., livestock feed, fertilizer, atmospheric deposition), (2) the outputs from the system (e.g., biomass exports such as agricultural products, losses to water or air), and the (3) changes in nutrient storage in the system (e.g., in soils, floodplains, or groundwater). The flow of nutrients is represented by solid arrows in figure 3.6. Movement of nutrients and sediment through the system is influenced by land application of fertilizers and manures, land use, weather patterns, BMPs, and processing and transport through soils, groundwater, and stream networks. Ultimately the TMDL objective is to reduce nutrient and sediment loads delivered to the estuary.

3.4. Lag times and legacy nutrients and sediment

A National Research Council report cautioned that achieving Bay water quality goals could be significantly delayed by lag times and legacy nutrients (NRC, 2011). The accumulated stores of legacy nutrients and sediment have been identified as important contributors to the lack of observable water quality improvement (Basu et al., 2022; Kleinman et al., 2019; Noe et al., 2020; Sharpley et al., 2013; Stackpoole et al., 2019; Van Meter et al., 2021). Legacy nutrients result from excess nutrient inputs that accumulate and are stored in soils and groundwater (Van Meter et al., 2016) and legacy sediment in floodplains (Noe et al., 2020, 2022). Reducing nutrient inputs and implementing certain types of BMPs designed to remove or transform nutrients may produce immediate reductions at the point of BMP implementation, but the collective benefits of these BMPs can take years or even decades to find their way through the coupled surface water-soil-groundwater system, resulting in significant lag times between BMP implementation and downstream water quality response (Böhlke, 2002; STAC, 2013). Thus, legacy nutrients continue to be a source of nutrients to surface water bodies even as contemporary nutrient loads are reduced (presumably due to BMP implementation).

Legacy P presents a substantial water quality management challenge (Kleinman et al., 2019; Staver et al., 2014). Legacy P is stored primarily in soils, but it can also be stored in groundwater (Holman et al., 2008; Meinikmann et al., 2015). In areas with intensive livestock production (e.g., poultry, dairy), animal manures are typically applied to land as fertilizer. Animal manures contain more P than N relative to plant needs; as a result of applying manures to meet crop needs for N, P has historically been applied at rates that exceed crop needs, creating a buildup of P in soils. Most P management has focused on so-called “soil conservation” strategies, as P is typically tightly bound to sediment and travels to surface water with eroded sediments carried in runoff. In soils with high P levels, however, P loads to surface water may also be in biologically-available dissolved forms (Kleinman et al., 2019). The increasing importance of dissolved P losses from legacy P in soils creates challenges when P management strategies focus solely on preventing sediment loss. The challenge of remediating legacy P is that significant P stores in soils can serve as a constant source contributing to dissolved P losses (Kleinman et al., 2019; Sharpley et al., 2013). Removal of P from soils by crops can take decades even with no additional applications of P (Fiorellino et al., 2017).

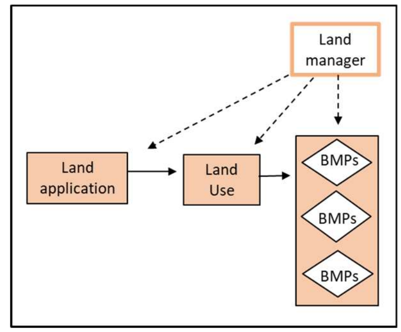
Legacy N exists in both groundwater and soils. Groundwater modeling of the Bay watershed shows elevated concentrations of N (in the form of nitrate) stored in groundwater in several regions (Greene et al., 2005). Legacy N is an important element of the contemporary N load. For instance, in the Susquehanna River Basin, legacy N in groundwater (greater than one year residence time) was found to contribute nearly 50% of the N load entering the Bay (Van Meter et al., 2017). The travel time for groundwater discharged into surface water ranges from less than a year to more than 50 years (Lindsey et al., 2003; Meals et al., 2010; Phillips & Lindsey, 2003; Sanford & Pope, 2013; STAC, 2013), indicating that stores of N will continue to contribute to surface water loads for decades. Fertilizer nitrates can persist in soils for decades (Sebilo et al., 2013).

Legacy sediment also introduces potential lag times in system response. Legacy sediment is defined as eroded sediments from land-disturbing activities stored in uplands, flood plains, stream channels, and impoundments (Miller et al., 2019). It is estimated that historical erosion rates through the 19th century were considerably higher than contemporary erosion rates due to improved soil conservation practices and lower rates of land conversion during the 20th century (Noe et al., 2020). A large portion of that legacy sediment is still stored on hillslopes, footslopes, and valley floors throughout the Chesapeake Bay watershed (Jacobson & Coleman, 1986; Smith & Wilcock, 2015; Walter & Merritts, 2008). The extent to which legacy sediment can be remobilized and transported to the Bay is highly variable over both time and space (Miller et al., 2019). More legacy sediments may be mobilized during heavier storms (Opalinski et al., 2022), and the region is already seeing heavier precipitation, a trend predicted to continue as a result of climate change (Mallakpour & Villarini, 2017). Sediment can be stored for long periods of time (years to millennia) and be subject to multiple mobilization and deposition events before entering the estuary (Pizzuto et al., 2014). In some cases this may mean that continued erosion of legacy sediment will generate elevated sediment loads well into the future, despite reduced sediment inputs from upland areas (Jackson et al., 2005).

3.5. Nonpoint source response gaps

The limited nonpoint source response observed to date may also be because nonpoint source programs are not as effective at producing pollutant reductions as estimated. Possible response gaps may result from inadequate understanding of behavioral responses to policy, overestimation of BMP effectiveness, and limited or incorrect model input data, among other causes. Sections of figure 3.5 are reproduced in the following discussion to highlight parts of the system implicated by each possible cause.

Nutrient use and conservation behavior



Relatively little is known about how conservation planning and behavior varies across agricultural land managers (Pannell & Claassen, 2020; Reimer & Prokopy, 2014). The BMP adoption literature consistently refers to the challenges in identifying factors that explain farmer adoption of conservation practices or lack thereof (Patterson et al., 2013; Prokopy et al., 2019; Ranjan et al., 2019; Reimer et al., 2014). Reimer et al. (2014) argued that understanding of landowner concerns and adoption behaviors lags understanding of physical processes associated with estimating nonpoint source loads.

There is limited information on how, when, and where land managers actually apply fertilizer and manure to farms and fields in the Bay watershed (Yagow et al., 2016). To calculate nutrient loads under the TMDL, the CBP makes assumptions about the methods, quantity, and timing of fertilizer and livestock manure applications across the watershed. The CBP collects information about total fertilizer sales and livestock numbers (from which manure production is estimated) at state and county levels. Because of the lack of information on farm level nutrient use, the CBP then assumes manure and fertilizer are applied within counties according to a complex formula that considers crop needs and nutrient management plans (CBP n.d.-c, section 3).

The CBP provides credits for land managers who develop different types of nutrient management plans that outline the rate and timing of nutrient applications. How nutrients are actually applied given farmers' risk preferences, changing on-farm constraints, and market opportunities is unknown, though evidence suggests nutrient management plans and actual use diverge (Osmond et al., 2015). The potential for some land managers to over-apply nutrients is particularly high for manure sources because manures are costly to transport, are concentrated in specific regions of the watershed, and have more uncertain nutrient content than commercial fertilizers. Agronomists have long recognized the risks of overapplying P given that the N:P ratios of manures (roughly 2:1) are incommensurate with plant needs (roughly 8:1) (Kleinman et al., 2017). If manures are applied to meet crops' N needs, then P will accumulate in soils.

Voluntary, incentive-based agricultural BMP implementation programs require that land managers self-select into the program. If nutrient use and nutrient or sediment losses differ among BMP adopters, then CBP accounting could overestimate the effectiveness of nonpoint source programs. For instance, if land managers interested in conservation practices have lower nutrient losses overall, then additional BMPs added by this group will produce lower than average nutrient reductions. This challenge is illustrated in figure 3.7. In this example, estimated N load reductions are calculated based on a subwatershed-wide average of 20 lb/ac N in runoff. The watershed (200 total acres) has a willing adopter who has already reduced N losses to 10 lb/ac, but the neighboring land manager has N losses of 30 lb/ac. If BMPs with a 50% removal efficiency are applied by land managers with high adoption rates (10 lb/ac N in runoff), only 5 lb/ac of N are reduced (10 lb/ac x 50%), but 10 lb/ac N reduction credit would be awarded (20 lb/ac x 50%).

There is limited research about how nutrient loads and BMP adoption vary across land managers. Some research has found that nutrient use can vary widely across land managers. Pearce and Maguire (2020) estimated P mass balance for 58 dairy farms (a nonrandom sample) in the Shenandoah Valley and found that surplus P (i.e., P imports minus P exported in products and manure) ranged from -27.6 to +87.1 lb/ac. High-loss operations with below average adoption rates could be partially responsible for the nonpoint source response gap.

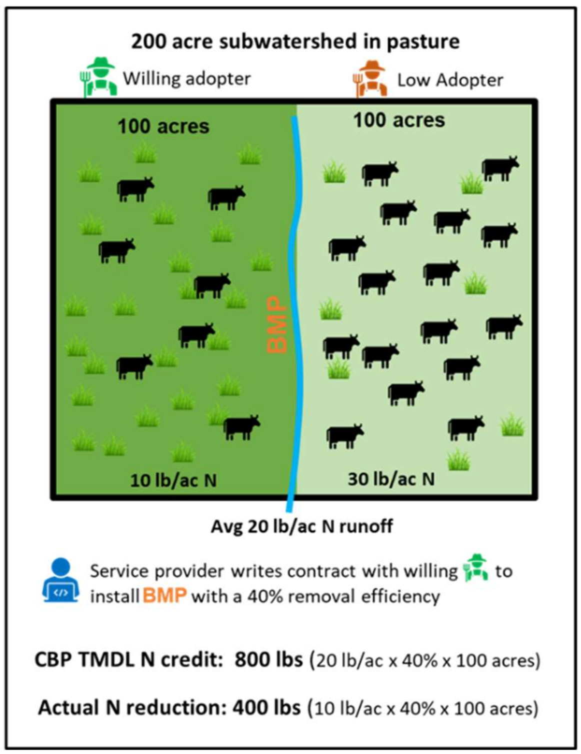
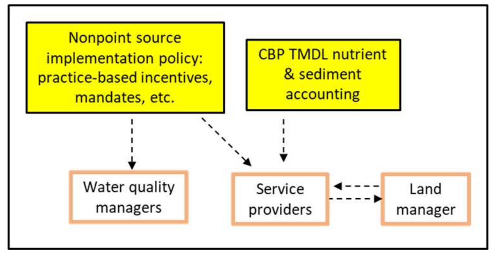


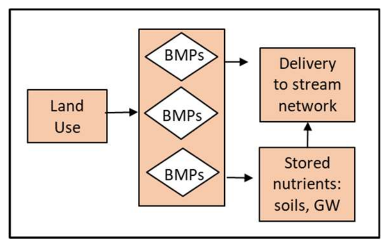
FIGURE 3.7.—Self-selection challenge under CBP TMDL accounting system resulting from spatial variability in baseline nutrient losses.

Program incentives of service providers and water quality managers



Technical service providers are a conduit between the entity funding BMP implementation and producers or land managers, providing engineering, installation, and maintenance assistance, and facilitating financial assistance. Agency and program incentives center around installing practices creditable within CAST (getting practices on the ground). Because all practices and operations are generally counted the same within a geographic area, there are no Chesapeake TMDL-related programmatic incentives for service providers and water quality program managers to locate areas and operators who generate disproportionately high loads. Within a land-river segment, a practice adopted by a low-loss or high-loss operator counts the same, while the amount of effort required to gain adoption may differ. This often means that, within agricultural areas, service providers such as Natural Resources Conservation Service or Soil and Water Conservation Districts may only work with cooperative land managers, who may, as a group, already have relatively low nutrient and sediment loss rates, reinforcing the self-selection bias described above (Stephenson et al., 2022). Inadequate staffing levels of technical service providers would compound the challenges of devoting time and staff resources to high-loss operators.

BMP effectiveness



The CBP may systematically overestimate BMP effectiveness. There is evidence that estimating BMP effectiveness with models tends to overestimate BMP performance (Liu et al., 2017, 2018). Osmond et al. (2012) found that models used in the U.S. Department of Agriculture (USDA) Conservation Effects Assessment Project consistently overestimated BMP effectiveness, and a recent review article (Lintern et al., 2020) noted that most studies showing water quality improvements used model-generated estimates, while field and monitoring studies showed mixed or little to no improvement from BMP implementation. Text box 3.2 illustrates these challenges for one of the most common agricultural BMPs, conservation tillage.

Easton et al. (2023) summarized a variety of possible reasons why BMP effectiveness might be overestimated.

> **Gaps in the literature.** The CBP uses expert panels to generate estimates of average BMP nutrient and sediment removal efficiencies (Stephenson et al., 2018). The panels rely on existing studies, and there are often gaps in the scientific literature. For example, some pollutants and loss pathways are better studied than others. Surface runoff is better characterized than groundwater leaching, and the long-term fate of stored

Text box 3.2. Tradeoffs and uncertainties in nutrient removal effectiveness of conservation tillage practices

Conservation tillage is often touted as a win-win that improves both water quality and agricultural profitability. Indeed, the CBP found that conservation tillage reduces surface runoff of N by 2 to 14% and surface runoff of P by 6 to 72% compared to conventional tillage, depending on type of reduced tillage method and geographic region (Thomason et al., 2016). Research consistently shows that conservation tillage practices dramatically reduce soil erosion.

Research also demonstrates the complex tradeoffs and loss pathways surrounding one of the most widely accepted BMPs (Kleinman et al., 2022). Overall, nutrient removal effectiveness of conservation tillage depends critically on both site-specific soil/topography/environmental conditions and nutrient management behavior, all of which can vary significantly across sites and land managers. Some forms of N, such as nitrate, are highly soluble and can be readily leached through soils to groundwater. Practices like no-till can enhance water-holding capacity and infiltration rates in soils, which has been shown to increase nutrient loss risk by facilitating the development of soil macropores that increase leaching (Thomason et al., 2016).

A number of recent watershed-level studies highlight some of the uncertainties surrounding the relationship between water quality and conservation tillage. Ator (2019) found that N loads in river networks in the northeastern United States are positively correlated with increased use of no-till practices. Using a statistical model (SPARROW), Roland et al. (2022) found that increasing conservation tillage and cover crop practices by 50% in the Chesapeake Bay watershed would reduce P loads approximately 6% but would increase N loads to the Bay by 0.9%. In a review of 43 studies, Daryanto et al. (2017) found mixed and inconclusive results for the overall effectiveness of conservation tillage to reduce nitrate leaching and runoff losses in agricultural systems.

Conservation tillage can also pose risks of increasing P runoff. Since P tends to bind to soil particles, reducing soil erosion will reduce the loss of P attached to soil particles. However, evidence shows how conservation tillage practices can increase subsurface loss of dissolved P (Duncan et al., 2019; Kleinman et al., 2011, 2015, 2020). By reducing or minimizing soil disturbance, conservation tillage systems can contribute to P stratification in soils. In stratified soils, P becomes concentrated in the top layer of soil, reducing potential of P to bind to soil particles and allowing dissolved P to be lost via surface runoff or groundwater leaching. Stratification can occur relatively quickly, especially when surface application of fertilizers or manures is combined with no-till. Dissolved P has greater potential to contribute to eutrophication problems than P attached to soil particles.

Substantial research documents the potential for conservation tillage practices to increase dissolved P losses (Smith et al., 2019). In a study of research findings on P loss from tillage practices, Daryanto et al. (2017) found that conservation tillage reduced sediment P losses by reducing erosion but generally increased dissolved P losses. In some cases, increases in dissolved P can offset reductions in losses of sediment P (Duncan et al., 2019). Recent increases in harmful algae blooms in Lake Erie have been associated with increases in the amount of dissolved P arising from increases in the use of no-till practices (Jarvie et al., 2017; Macrae et al., 2021; Michalak et al., 2013). Water quality analysts have also observed increases in dissolved P levels in multiple river systems in the Chesapeake Bay watershed that are linked to agriculture (Fanelli et al., 2019; Kleinman et al., 2019).

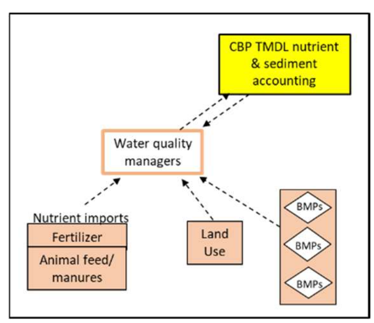
nutrients is difficult to characterize within the time frame of typical studies (Jefferson et al., 2017).

Operation and maintenance behavior. Assumptions must be made about how people will actually operate and maintain a BMP over time. BMPs must be operated and maintained according to design specifications, and most BMPs require some type of routine maintenance to sustain their performance. BMPs that are the subjects of scientific studies are likely to be operated and maintained differently than in real world settings. Furthermore, evidence suggests that lack of or inconsistent maintenance is a frequent problem for both urban and agricultural BMPs (Aguilar & Dymond, 2019; Jackson-Smith et al., 2010). Additional research is needed to better characterize BMP operation and maintenance under actual conditions (Liu et al., 2017).

Long-term performance. Evaluations of BMP performance typically assess performance a few years after installation. For many BMPs, performance varies over time. Many BMPs function by storing nutrients or sediment, either in the soil or in plant biomass. Yet the storage capacity of a BMP is limited, and the fate of those stored nutrients or sediments over time is not well characterized (LeFevre et al., 2015). If stored nutrients are more likely to be mobilized as the BMP ages, then BMP performance will be overestimated over time (Hopkins et al., 2020; Selbig & Bannerman, 2008).

Behavioral biases. Behavioral research consistently finds that people are prone to systematic errors in decision-making in certain situations (Ariely, 2008; Kahneman et al., 2021). Expert panels are commonly confronted with literature that contains wide ranges of findings (e.g., from BMPs with removal efficiencies of over 90% to BMPs that increase nutrient losses under some conditions). In situations with considerable uncertainty, people (experts included) tend to assign causal explanations to randomly produced outcomes (Stephenson et al., 2018). This possibility is increased in settings with limited data, situations most often faced by expert panels.

Watershed data and monitoring



The CBP collects and uses a tremendous amount of data about the watershed to generate estimates of nonpoint source loads, evaluate scenarios, and monitor TMDL compliance. Data are collected on land use, soils, weather, physiography, nutrient inputs, and BMPs installed. In many cases, data are of high quality, and numerous processes have been put in place to improve the quality of data. Land use data are an example. However, collecting data in a highly complex system is expensive, necessarily selective, and subject to error.

Uncertainty exists regarding the accuracy of input data for the CAST model. For instance, some jurisdictions express concerns that the CBP undercounts BMPs, therefore underestimating nonpoint source reductions (Royer et al., 2016). To estimate nutrient inputs, the CBP collects data on aggregate commercial fertilizer sales across the watershed and estimates the quantity of manure produced in the watershed based on estimates of animal numbers reported from National Agricultural Statistics Service (NASS) surveys. Recent evidence suggests NASS survey data can lead to undercounting livestock, and the CBP undercounted total fertilizer sales for recent years (Blankenship, 2022).

In addition, limited information is collected about key nutrient generating processes. As already discussed, little data exist on how, when, and where land managers actually apply fertilizer and manure in the Bay watershed at the farm- or field-scale (Yagow et al., 2016). Soil P levels are an important intermediate indicator of the effectiveness of management actions because changes in soil P are indicative of changes in water quality (Kleinman et al., 2019), but farm-level data on soil nutrient levels (particularly P) are limited because of privacy concerns. Artificial drainage (surface ditch and subsurface tile drainage) of agricultural fields can increase nutrient loss by providing direct pathways to water bodies (Bryant et al., 2019; Kleinman et al., 2019) but is not explicitly accounted for in the CAST modeling system (Bryant et al., 2019). Artificial drainage systems for agricultural fields are common in coastal regions of the watershed, and their use is increasing (Bryant et al., 2019). Bryant et al. (2019) estimated that hundreds of miles of agricultural drainage have been installed by farmers without cost-share assistance, so the extent of these additions is largely unknown.

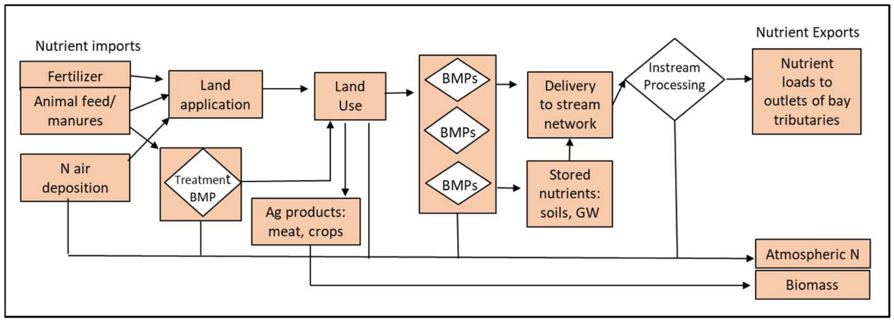
Finally, ambient water quality monitoring data are often not collected at the scale that can provide insights into the effectiveness of management actions (Easton et al., 2023; Ator et al. 2020). The CBP Nontidal Monitoring Network (NTN) was established to assess watershed response to nutrient and sediment reduction efforts (CBP, 2022). Selection of sites was driven largely by the location of historical monitoring stations. Furthermore, the NTN monitoring sites tend to be located on larger streams and rivers (with drainage areas ranging from 50 to over 1000 square miles), thus limiting the ability to distinguish between multiple factors affecting changes in monitoring data at more localized scales. Jurisdiction partners have limited incentives to invest in additional and finer-scale ambient instream monitoring.

3.6. Implementation gaps in achieving nonpoint source pollutant reductions

Several technical and management issues affect the ability to secure the type and level of implementation necessary to achieve nonpoint reduction goals and TMDL targets. Important factors contributing to implementation gaps include (1) nutrient mass imbalances, (2) limited participation in voluntary incentive programs, and (3) limited spatial targeting of efforts to encourage BMP adoption.

Nutrient mass imbalances

From a system perspective, nutrient mass balances are critical determinants of the nutrient status of a given region (Easton et al., 2023). If inputs (e.g., fertilizers and livestock feed brought in) exceed outputs (e.g., crops and livestock harvested and removed), then excess nutrients may find their way to surface water or groundwater.



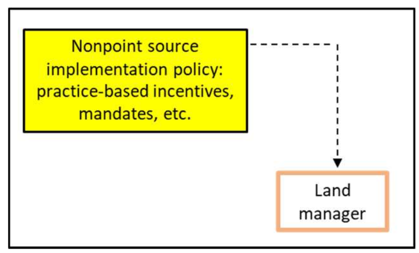
Available data indicate that overall nutrient inputs from fertilizer and livestock feed, two primary nutrient sources, have declined across the entire watershed since the 1980s (Clune et al., 2021; Keisman, Devereux, et al., 2018; Sabo et al., 2022). Declines have been driven largely by reductions in fertilizer use generated by improvements in nutrient use efficiency. Reductions in P fertilizer inputs were particularly pronounced between 1980 and 2012 (Keisman, Devereux, et al., 2018). Areas with substantial declines in nutrient inputs are closely associated with improving water quality trends in nonpoint source-dominated watersheds (Sabo et al., 2022).

Nutrient inputs are unevenly distributed across the watershed, however, and in many places, nutrients have become more concentrated over time (Keisman, Devereux et al., 2018). Regions with high imports of nutrients in the form of fertilizer and animal feed are associated with mass imbalances, and these are also areas with high nonpoint source loads. While mass imbalances can also be high in many urban areas, the most severe and extensive mass imbalances are in regions dominated by intensive agriculture (Keisman, Devereux, et al., 2018). Increases in these areas that already receive higher than average nutrient inputs are being driven largely by increases in livestock numbers (primarily poultry) and agricultural intensification (Ator & Denver, 2015; Kleinman et al., 2012, 2019). According to USDA Census of Agriculture data, poultry numbers in the Lower Susquehanna basin and on the Delmarva Peninsula increased from 2002 to 2017 by 20.5 million (from 37 million, a 64% increase) and 70.8 million (from 143 million, a 66% increase), respectively (Easton et al 2023). Sabo et al. (2022) found surplus amounts of nutrients in some agricultural areas of the Bay watershed have increased from 2009 through 2019, reversing longer-term declines since the 1980s. Across the Bay watershed, livestock produce about 10 times more excrement by mass volume than the human population (Kleinman et al., 2012), and yet law requires some form of wastewater treatment for all human waste to reduce nutrient content.

Mass nutrient imbalances limit the amount of long-term progress that can be made in reducing nonpoint source loads (Ator et al., 2020; Beegle, 2013). Addressing existing mass imbalances would require reducing the import of feed and fertilizer, increasing the uptake of nutrients and their export out of the region or watershed in food and fiber, hauling manure out of the watershed, adopting manure conversion technologies, or implementing appropriate wastewater treatment mechanisms (Flynn et al., 2023; Kleinman et al., 2012; Sharpley et al., 2013; Spiegal et al., 2020). Appreciable reductions in nutrient loads cannot be achieved unless regional mass imbalances are successfully addressed.

Many conventional agricultural BMPs that are integral elements of WIPs do not appreciably change the mass balance. Regional mass imbalances in areas of intensive livestock production are particularly problematic for P. Phosphorus removal efficiencies of cover crops, conservation tillage, and riparian buffers tend to result from altering transport pathways, increasing temporary nutrient storage, or altering the form of nutrients (Kleinman et al., 2022). Drawing down soil P levels within a reasonable time frame requires reducing P inputs and some sort of P mining (e.g., by P-scavenging plants).

Behavioral response to nonpoint source reduction policy



While local success stories exist, voluntary, conventional (cost sharing) incentive-based programs targeting the agricultural sector have not consistently generated the type or scale of BMP implementation sufficient to produce the reductions needed to achieve WQS (Liu et al., 2020; Patterson et al., 2013; Prokopy et al., 2019; Reimer & Prokopy, 2014; Ribaudo & Shortle, 2019; Shortle et al., 2012, 2021; Stephenson et al., 2022). Multiple reasons exist that can limit adoption of effective pollutant control practices.

The structure of practice-based cost-share incentive programs can limit choices about the type and extent of adoption. Research generally finds that conventional financial assistance programs encourage the adoption of practices with significant private benefits and low upfront costs (Claassen et al., 2014; Pineiro et al., 2021), such as conservation tillage (which lowers operating costs) and cover crops (which increase long-term soil productivity). Perceived risk to farm profitability is a major factor limiting adoption (Duke et al., 2022; Wilson et al., 2014). Unsurprisingly, adoption rates tend to be much lower for practices with higher upfront costs and few private on-farm benefits, such as stream buffers, denitrifying bioreactors, stream fencing, and manure treatment. Many such BMPs, however, can generate substantial nutrient reductions at relatively low per unit ($/lb) costs to the public (Easton et al., 2019; Price et al., 2021). Yet, if financial returns are a primary motivator, a land manager will not install and operate a BMP with few agronomic benefits, even if a portion of the costs are shared (Prokopy et al., 2019; Ribaudo, 2015). There is limited evidence that temporary cost-share payments lead to sustained adoption (Pannell & Claassen, 2020). Similar logic applies to landowner adoption of structural urban stormwater controls (Gonzalez et al., 2016). Sharing the costs of a practice with high upfront costs is often insufficient to generate voluntary investments in BMPs with high public benefits. The problem is compounded when the purported public benefits accrue downstream.

The adoption of BMPs can also have unintended consequences for nonpoint source pollution. Taking land out of production to install riparian buffers may induce more intensive upland cropping practices, partially offsetting the pollutant reduction gains from the buffer (Bonham et al., 2006). Practices and cost-share payments that improve financial outcomes may expand and intensify agricultural production, potentially offsetting the pollutant reductions gained from BMP implementation (Fleming et al., 2018; Lichtenberg & Smith-Ramírez, 2011). Cover crops are an example. Research consistently finds that cover crops (winter ground cover) improve soil conditions, significantly reduce soil erosion, and potentially reduce nutrient runoff (Blanco-Canqui, 2018; Koudahe et al., 2021). However, the nutrient reduction potential of cover crops is contingent on the land manager's nutrient application behavior. Experience and success with cover crops may lead farmers to plant more cash cover crops (winter wheat, for example), intensifying production and encouraging additional nutrient applications, partially offsetting some of the water quality benefits associated with cover crops.

Research also suggests other factors that limit adoption (Easton et al., 2023). Cultural norms are one example. Members of plain sect communities in Pennsylvania and Virginia, a significant portion of the agricultural communities in these regions, do not participate in government assistance programs, including cost-share programs. These farmers, as well as small, land-constrained, or resource-limited agricultural operations often have systematically lower BMP adoption rates, at least under typical cost-share programs (Prokopy et al., 2019). Finally, many agricultural land managers lease or rent land, which also lowers BMP adoption (Collins et al., 2022; Ranjan et al., 2019).

Targeting nonpoint source investment

The CBP offers few tools or incentives to identify and treat site-specific high-loss areas (Easton et al., 2019; Stephenson et al., 2022). While CAST can identify high-loading areas at a relatively coarse spatial scale (averaged over thousands of acres, see figure 3.1), it does not reflect localized high-loss areas at field or farm scale (Easton et al., 2020; Lintern et al., 2020). Researchers have noted that areas of high nutrient and sediment loss are site-specific and highly localized (Easton et al., 2017; Easton et al., 2008), and many studies suggest that 5–20% of the land area generates 50–90% of runoff and nonpoint source loads (Heathwaite et al., 2000; Qui, 2009; Rao et al., 2009; Wagena & Easton, 2018; White et al., 2009; Xu et al., 2019). Improving the ability to spatially identify and target BMPs to sites and operations with higher pollution potential could improve water quality and reduce costs of pollution reduction efforts (Choi et al., 2020; Giri et al., 2012; Kast et al., 2021; Khanna et al., 2003; Lintern et al., 2020; Xu et al., 2019; Yang & Weersink, 2005).

Incentives and ability to reach nonadopters

Existing cost-share program rules also can limit the program effectiveness of service providers (Collins et al., 2022). Highly localized hotspots may mean that a large percentage of loads could be treated on a relatively small portion of the farm operation. Most cost-share program rules require whole farm plans and would prevent targeted treatment for land managers reluctant to treat the entire operation.

Similarly, strict technical standards for BMP installation have been cited as a barrier to BMP adoption (Collins et al., 2022). For instance, land managers willing to fence livestock away from streams or install riparian buffers are ineligible for program benefits if their efforts do not meet all technical criteria, even if pollutant control impacts are positive. Also, cost-share funding rules prevent service providers from offering additional financial assistance to reluctant or resource-limited land managers even if such assistance could produce large quantities of low-cost pollutant reductions. While such programmatic conditions serve several important functions, they also represent a barrier to increasing adoption among reluctant, skeptical, or resource-limited land managers.

Incentives for innovation in nonpoint source load reductions

Technological innovations to reduce nutrient inputs, increase efficiency of nutrient use, and treat existing sources can accelerate progress toward meeting nonpoint source reduction goals. However, barriers limit development and adoption of such technologies (Stephenson et al., 2022). In CBP nonpoint source programs, all BMPs must be approved by the CBP through the BMP protocol process to be credited toward meeting the TMDL. For new BMPs, the approval process can take years, creating a high cost of entry for new, and potentially innovative, BMPs.

In addition, the existing accounting framework and incentive systems provide limited incentives for people to invest in actions to improve the certainty of achieving nonpoint source pollutant outcomes. Typically, practices are assigned a fixed pollutant reduction credit, but actual pollutant removal performance is often both highly variable and uncertain. A number of technologies offer opportunities to directly or indirectly measure pollution outcomes (Rose et al., 2015; Stephenson et al., 2018; Stephenson & Shabman, 2017a), but their application is rare because TMDL accounting automatically credits reductions associated with installed practices. Activities that provide certainty (e.g., measurement of nutrient load from a practice or intermediate indicators of effectiveness such as changes in nutrient levels in soils) add costs to practices. The structure of incentives, however, provides people little reason to pay for greater assurances that pollutant reductions occur.

Climate change

Climate change poses an array of challenges to meeting the Bay TMDL. Some of these challenges, such as increased precipitation and streamflow, are widely recognized for their potential to increase nutrient and sediment delivery to the Bay (Hanson et al., 2022; Ryberg et al., 2018; Sinha et al., 2017). Indeed, the TMDL Phase III WIPs now require all Bay jurisdictions to account for the additional nutrient and sediment loading expected from climate change.

The primary climate-related drivers affecting the Bay watershed are air temperature, precipitation, and sea-level rise (Najjar et al., 2010). Changes in these drivers are expected to alter key processes within the Chesapeake Bay and its watershed, including the duration, frequency, and magnitude of precipitation; evapotranspiration; soil moisture; streamflow; and temperature. Climate change will also affect watershed water quality by indirect means, with changes in agricultural land use (e.g., different cropping mixes, increases in artificial drainage) and increased agricultural intensification (e.g., double cropping) in response to longer growing seasons and changes in rainfall patterns. This could also fundamentally alter nutrient mass balances, and consequently the cycling and export of nutrients, in ways that are not fully understood.

3.7. Conclusions and implications

Achieving nutrient and sediment reduction targets under the TMDL is going to be more difficult than is generally acknowledged. Because of the success of efforts to reduce point source N and P and atmospheric N deposition, agricultural and urban nonpoint sources are the largest remaining sources of controllable nutrient loads. Tens of millions of pounds of nutrient reductions are needed to meet the TMDL, but a decade of intensive management efforts have produced only modest (as predicted by CAST) reductions in nonpoint source loads. Statistical analyses of monitoring data provide some evidence that nonpoint source practices are not producing the reductions expected. Given the growth of animal production in the watershed, conservation practices and efficiency improvements in agricultural production have helped prevent *significant* increases in nutrient pollutants. Despite successes of point source and atmospheric reductions, the TMDL targets are not likely to be achieved without increasing efforts to address nonpoint source loads—in particular the gaps that limit progress toward achieving nonpoint source load reductions.

Nonpoint source policy relies heavily on voluntary, practice-based incentive and education programs to induce behavioral change. Empirical evidence suggests that nonpoint source management programs have not produced the type and scale of behavioral changes needed to close the implementation gap.

Data indicate that efforts to reduce nonpoint source loads are not as effective as expected. These response gaps make achieving the Bay TMDL more difficult. Phosphorus is of particular concern, and while CBP reports significant progress toward achieving TMDL P goals, statistical analyses of ambient monitoring data often show no change or increasing P loads. Dissolved P loss, a form of P that contributes the most to algae growth, is increasing in many regions of the watershed, particularly in regions associated with nutrient mass imbalances. The difference between realized and expected nonpoint source outcomes could be due to the large stores of legacy nutrients and sediment in the watershed that delay fully realizing nonpoint source reductions. Evidence also suggests a variety of reasons why our understanding and characterization of the effectiveness of nonpoint source control programs may be incomplete or inaccurate.[4]

STAC concludes that additional effort and funding through existing programs alone is unlikely to produce the type and magnitude of changes needed to meet TMDL targets. Given the scale of the challenge, additional amounts of technical assistance and effort will be necessary to help close the implementation and response gaps. However, new and innovative options need to be considered to improve the effectiveness of nonpoint source programs, especially toward programmatic and behavioral changes.

Ways to address these gaps are not always obvious and surely will not be easy. Indeed, the challenge is not unique to the CBP; nonpoint source management has long been a “wicked” problem (Patterson et al., 2013; Shortle and Horan, 2017; Wiering et al., 2020). The inability to produce nonpoint source reductions on the scale sought by water quality management programs is one of the most fundamental and common challenges confronting large-scale efforts to address eutrophication (Boesch, 2019; Wiering et al., 2020). Effective nonpoint source control at large watershed scales has proven difficult for water quality programs in the United States (Lintern et al., 2020; Osmond et al., 2012; Sprague & Gronberg, 2012; Tomer & Locke, 2011).

Recognizing these challenges, there are still opportunities to improve program effectiveness (Easton et al., 2023). Given uncertainties, not all alternatives are guaranteed to work as expected, but adaptive management, innovation, and experimentation are likely needed to address the nonpoint source challenge (NRC, 2001; NRC, 2011). Below are some opportunities to improve program outcomes.

Improved targeting of nonpoint source investments. Research consistently highlights the potential of targeting nonpoint source investments to reduce nonpoint source loads and costs (Choi et al., 2020; Easton et al., 2019; Giri et al., 2012; Kast et al., 2021; Khanna et al., 2003; Lintern et al., 2020; Xu et al., 2019; Yang & Weersink, 2005). Distribution of nonpoint source loads is highly uneven across watersheds, but existing tools, TMDL accounting rules, budget allocations, and incentives limit options for targeting nutrient hotspots. Targeting investments to areas of the landscape that produce disproportionately large loads offers opportunities for increasing program effectiveness. This could be accomplished in a number of ways. Developing and using finer-scale (field-level) tools to identify high-loss areas is one example. Additionally, developing new financial incentive programs, such as pay-for-performance or pay-for-success programs, could reward treatment of high-loss areas or operations and encourage adoption of potentially more effective practices (Talberth et al., 2015). Nonpoint source program outcomes

[4] Ranking or isolating the relative contributions of the various explanations for the response gap is beyond the scope of this report.

could be improved by designing programs and incentives to shift attention from installing practices to maximizing reductions.

Focus on nutrient mass imbalances. More effective and systematic approaches to addressing nutrient mass balance issues offer opportunities for substantial, sustained reductions in nonpoint source nutrient loads. A mass balance approach describes inputs (e.g., fertilizer and feed) and outputs (e.g., grain or meat export, loss to water bodies) to and from the system, reactions or transformations (e.g., denitrification), and storages (e.g., build-up of P in soil) in the system. Most BMPs do not substantially alter mass balances. Evidence suggests that policies designed to alter regional mass balances have proven particularly effective in improving water quality. The largest regional mass imbalances, particularly P, are associated with intensive livestock agriculture. Given the trends towards increasing animal numbers and more intensive livestock operations, traditional BMPs like cover crops or no-till that do not substantially alter inputs or transformations are unlikely to substantially alter nutrient losses. Reducing the nutrient content of livestock feed through precision feed management as a way to manage nutrient imports can address mass imbalances (Cerosaletti et al., 2004). Manure utilization, treatment, and conversion technologies (e.g., thermochemical or microbial transformation) offer the potential to significantly alter the mass balance by transforming manure-derived nutrients into less biologically available forms or converting manure into forms more easily transportable. Manure transport programs that export excess nutrients to nutrient deficit areas can be strengthened, for example by moving to more restrictive P-based nutrient management plans (Saha et al., 2022).

Requiring more. A key element of many nonpoint source policies is that land managers voluntarily decide whether or how to participate in nonpoint source programs. Such an approach is often reasonable given the diversity and number of people involved in producing nonpoint pollutants. However, the extensive history of nonpoint policy illustrates the limits to voluntary adoption, regardless of the type of incentive programs used. To make substantial progress in reducing nonpoint source loads, new and refined policies in strategic nonpoint source sectors may be required. Such requirements need not be overly costly to the land manager if well-designed and accompanied by financial assistance.

Encourage innovation. Finally, improving implementation effectiveness could be facilitated by encouraging and providing opportunities for innovation. Given that exemptions in the CWA limit federal authority over nonpoint sources, responsibility for addressing nonpoint sources rests to a large degree with state and local governments. The complexity and challenges of nonpoint source management also imply there is no one best way to manage nonpoint sources. The CBP could encourage innovations in nonpoint source management by supporting state and local governments to develop new approaches for managing nonpoint sources. Sandboxing is a formalized way to test and evaluate the efficacy of new rules and programmatic approaches to nonpoint source management (Higgins & Male, 2019; O'Sullivan, 2021). In a sandbox, a government agency allows people to develop and implement an innovative rule or approach to achieving a common objective without disrupting the operation of existing programs.

Government agencies must grant permissions for sandboxing and be willing to provide technical support to adequately test the efficacy of new approaches.

4. Estuary Water Quality Responses to Nutrient and Sediment Load Reductions

In this chapter, we examine the progress, ongoing challenges, and future opportunities in the evaluation of estuary water quality responses to reductions of nutrient and sediment loads from the watershed. We address whether nutrient and sediment load reductions from the watershed initiate and sustain the physical, chemical, and biological responses necessary to meet the stated WQS and, ultimately, protect the aquatic living resources for which the WQS were developed. This chapter addresses the following questions:

- Is estuary water quality responding in ways consistent with the expected response to nutrient and sediment load reductions achieved to date?
- What are the major uncertainties in assessing attainment of Bay WQS, including DO, water clarity/SAV, and algal biomass (Chl *a*)?
- What are the major uncertainties in efforts to attain Bay WQS (DO, water clarity/SAV, Chl *a*)?

The expected response curve in figure 4.1 anticipates the response of estuary water quality to nutrient and sediment load reductions. The expected response curve rises to 100% attainment of WQC (represented by the black dashed line at the top of the graph) when the expected load reductions set by the TMDL are achieved. Expected responses are determined by the CBP estuarine models (Hood et al., 2021), which predict variations in Bay circulation and water quality due to changes in input loads, atmospheric forcing, and climate change effects (temperature, precipitation, and sea-level rise). Given the complexity of estuary biological and chemical processes, changing environmental conditions, and diversity of habitats, significant scientific uncertainty exists around the ability to explain (from monitoring results) and predict (from the estuarine models) the level of WQC attainment from any given reduction of nutrient and sediment loads to the Bay (dotted curves in figure 4.1).

Chapter 3 established that we have not yet realized the nutrient and sediment load reductions from the watershed required to meet the TMDL targets due to implementation and response gaps. In this chapter a response gap (figure 4.1) is defined as the difference between the expected response of estuary water quality to a given reduction of nutrient and sediment loads (yellow dot in figure 4.1) and the realized water quality response to those reductions (red dot in figure 4.1). If actual water quality response is below the expected response (as shown in figure 4.1), attainment is more difficult; likewise, if realized response is above the expected response, attainment is easier than current science suggests. Accordingly, management concern is focused on response gaps that may make achieving the WQC more difficult.

4.1. Water quality criteria and conditions for attainment

In developing the WQS for Chesapeake Bay, eutrophication was identified as the primary challenge to achieving conditions suitable for aquatic life. Figure 4.2 shows primary degradation factors for water quality that result from eutrophic (nutrient-rich) conditions. (Note that the figure does not include climate change effects such as sea level rise and water temperature

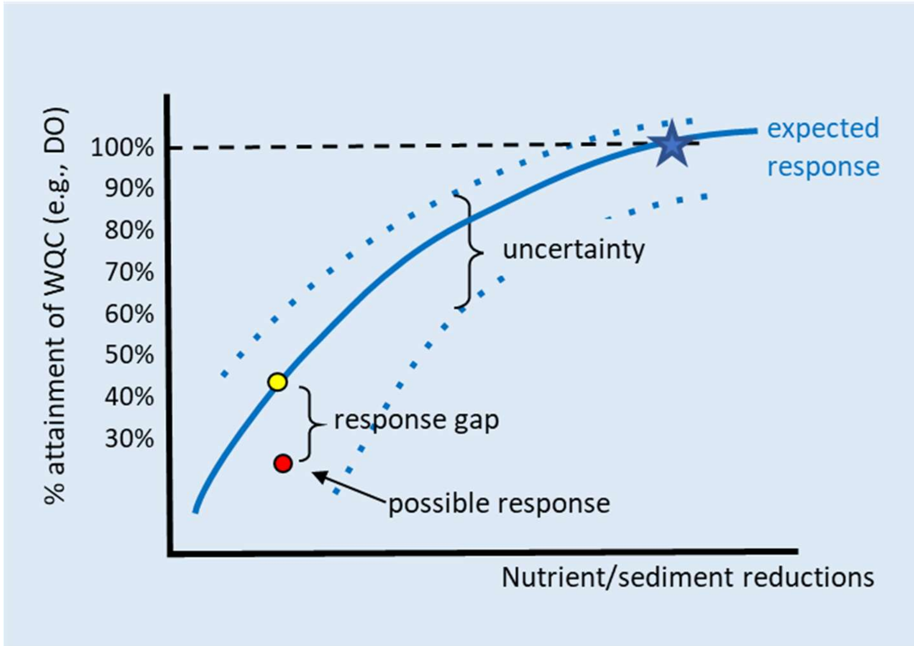


FIGURE 4.1.—Conceptual water quality response to nutrient and sediment reductions, where attainment of WQC is a function of nutrient or sediment load reductions.

increase). The Bay WQC reflect the need for more water clarity, more SAV, more DO, and less Chl *a* in the interest of achieving the designated use of the Bay, which is to support aquatic life. Reducing nutrient and sediment loads improves water clarity and, thus, improves light conditions for SAV. Chlorophyll *a* is an indicator of algae growth, which must be reduced to address both the DO and water clarity criteria. The WQC are predicted to be achieved if the TMDL is met (USEPA, 2010).

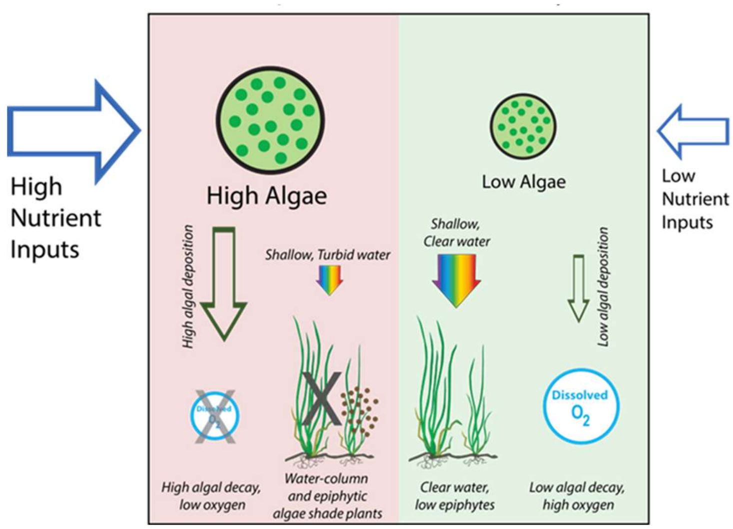


FIGURE 4.2.—Conceptual diagram of primary degradation factors for water quality in eutrophic (nutrient-rich) conditions as compared to low nutrient input conditions (Source: Testa et al., 2018).

4.2. Estuary response to realized nutrient and sediment loads

The factors illustrated in figure 4.2 underlie the responses of the Chesapeake Bay estuarine system to nutrient and sediment load reductions, and figure 4.3 illustrates specific processes associated with these factors. The shaded orange box in figure 4.3 represents the nutrient and sediments loads entering the Bay from the watershed, as generally described in Chapter 3. CBP implementation policies (working through the processes illustrated in figure 3.5) are responsible for generating load reductions, and the CBP partners are responsible for monitoring and assessing Bay water quality (yellow boxes in figure 4.3). The blue boxes in figure 4.3 represent the physical, biological, and chemical processes that affect or are affected by nutrient and sediment loads.

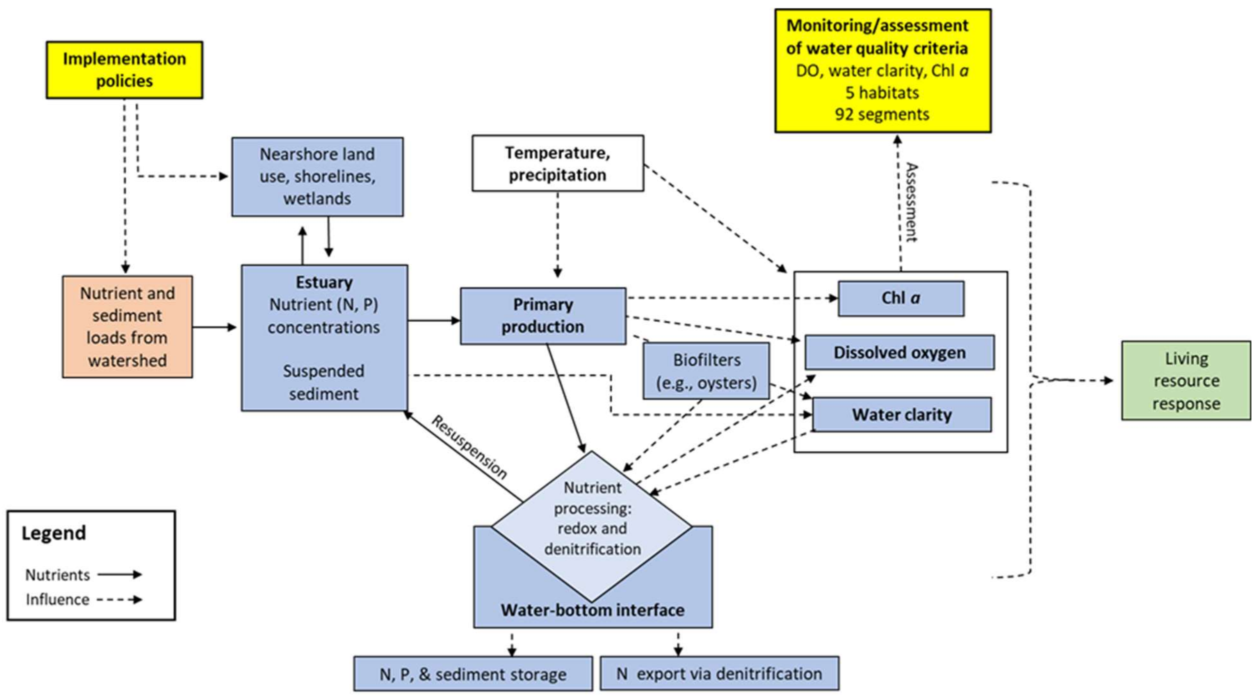


FIGURE 4.3.—Estuarine physical system response to CBP implementation policy.

Figure 4.4 shows N and P loads for the period 1985–2020. Data on nutrient and sediment loads to the Bay come from four main sources. Most of the nutrient and sediment load is monitored via the RIM stations located on the nine major tributaries in the watershed, and operated via a partnership between USGS, MD Department of Natural Resources, and VA Department of Environmental Quality. The flow-weighted annual loads are estimated by USGS using the monitoring results from these RIM stations. This load, termed “river input”, represents loads emanating from about 78% of the watershed. In order to estimate total annual loads of N and P delivered to the estuary from the remaining 22% of the watershed (located below the RIM stations), the CBP adds measured loads of N and P contributed by WWTPs and computer-simulated estimates of N and P loads from nonpoint sources located below these RIM stations. Remaining nutrient load data comes from estimates of the atmospheric deposition of N falling

FIGURE 4.4.—Measured and estimated N (panel a) and P (panel b) loads and annual flow by October 1–September 30 water year (Source: CBP, n.d.-e).

on tidal waters. Annual loads to the estuary vary considerably due to annual variations in precipitation. Within this variability, declines in N loads to the Bay over the past 40 years have been relatively modest, based on the source data for figure 4.4. For example, mean annual N loads averaged 368 million lb/yr over the period 1985–94 and fell to 320 million lb/yr for 2011–20, representing a 15% decline; this same data shows smaller declines for P and sediment.[5]

Realized reductions in loads entering the Bay are generally translating into lower nutrient concentrations throughout the estuary. The majority of estuary monitoring stations have seen reductions in surface water N and P concentrations when measured over time periods of 10, 20, or more than 30 years (especially since 1985) as shown by the downward pointing blue triangles in figure 4.5. Based on long term trends, 82% and 79% of stations saw significant decreases in N and P concentrations, respectively (Murphy et al., 2022).

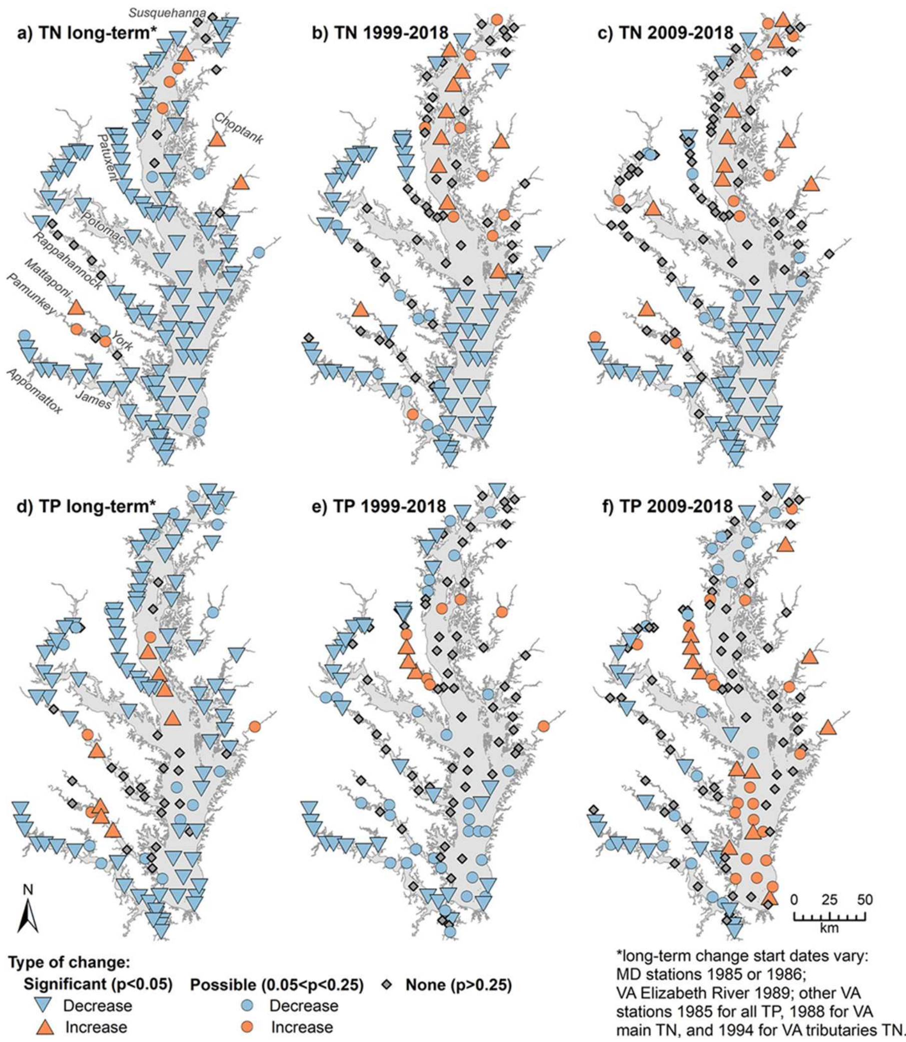


FIGURE 4.5.—Mean change in surface TN (a-c) and TP (d-f) over three time periods at tidal monitoring stations, calculated using Generalized Additive Modeling approach (Source: Murphy et al., 2022).

[5] The 15% decline in N loads between the time periods 1985–1994 and 2011–2020 calculated from this data source represents the best estimate of what the Bay "sees" in terms of nutrient loads. In contrast, simulated loads from CAST and jurisdiction-reported data on wastewater discharges suggest a 30% reduction in N loads (370 to 258.9 million lb/yr, 1985–2020) (see figure 3.2). The difference between estimated and realized loads to the Bay can be attributed to possible lag times and nonpoint source response gaps (as discussed in chapter 3).

Despite the complexities of developing accurate estimates of nutrient and sediment loads, spatial variation in TN and TP concentration trends in tidal waters generally reflect the assumption "where nutrient loads decline, estuarine concentrations will decline." To test this, an analysis by Testa et al. (2018) compared trends of both computed average annual loads (kg/yr) and estuarine concentrations (mg/L) between two periods (1989–91 and 2012–14) selected to represent historic and recent periods with similar average annual flows and also representative of long-term average conditions (this effectively removed the effects of variable river flow from the comparison of recent and historical loads). Concentrations of TN and TP in the estuary were derived from the monitoring results from estuarine water quality monitoring stations (n=140) and were aggregated into each of the 92 CB segments, and pollutant load estimates were compiled from the same sources as described for figure 4.4 for each segment. Points representing a segment's annual TN or TP concentration trend and annual load trend were placed in one of four quadrants (figure 4.6) where the bottom left quadrants show results for segments where both N or P loads and concentrations in the estuary have declined (green dots), and the upper right quadrant shows results for segments where both loads and concentrations increased (red dots). Segment loading changes correctly predicted the direction of change in estuary concentration in 88% (TN) and 66% (TP) of cases (upper right and lower left quadrants), but the remaining cases show unexpected patterns (gray dots in figure 4.6). These results illustrate the challenges in directly relating load reductions to decreasing concentrations.

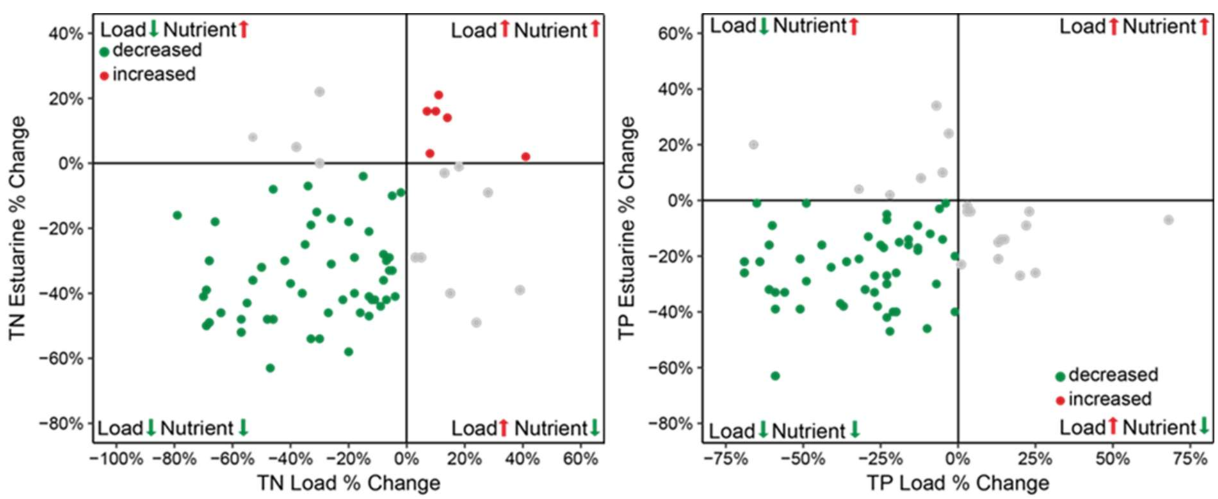


FIGURE 4.6.—Percent change in estuarine TN and TP loads and concentrations, late 1980s to mid-2010s, where each dot represents a Bay segment (Source: Testa et al., 2018).

While trends in N and P concentrations show correlations with trends in loads, these do not necessarily translate into improving trends in constituents represented by WQC. The CBP, MD Department of Natural Resources, and VA Department of Environmental Quality have sampled water quality on a bi-monthly or monthly basis at more than 130 stations located throughout the mainstem of the Chesapeake Bay and the tidal portions of numerous tributaries on the western and eastern shores since the mid-1980s. These data are used to produce estimates of conditions (improving/no change/degrading) for nutrients, DO, Secchi depth (a measure of water clarity), Chl *a*, and water temperature. Scientists evaluate short- and long-term trends using a Generalized Additive Modeling approach. The most recent report shows that, over the long term, most stations have seen no change or degrading conditions in DO in the bottom layer (82% of stations), annual Secchi depth (84% of stations), and spring season surface layer Chl *a* (76% of stations), indicating that load reductions are not translating into improved water quality conditions in the deep waters of the Bay (CBP, n.d.-d). However, there is substantial spatial variability in the results, as evidenced in the long-term mean DO concentrations depicted in figure 4.7.

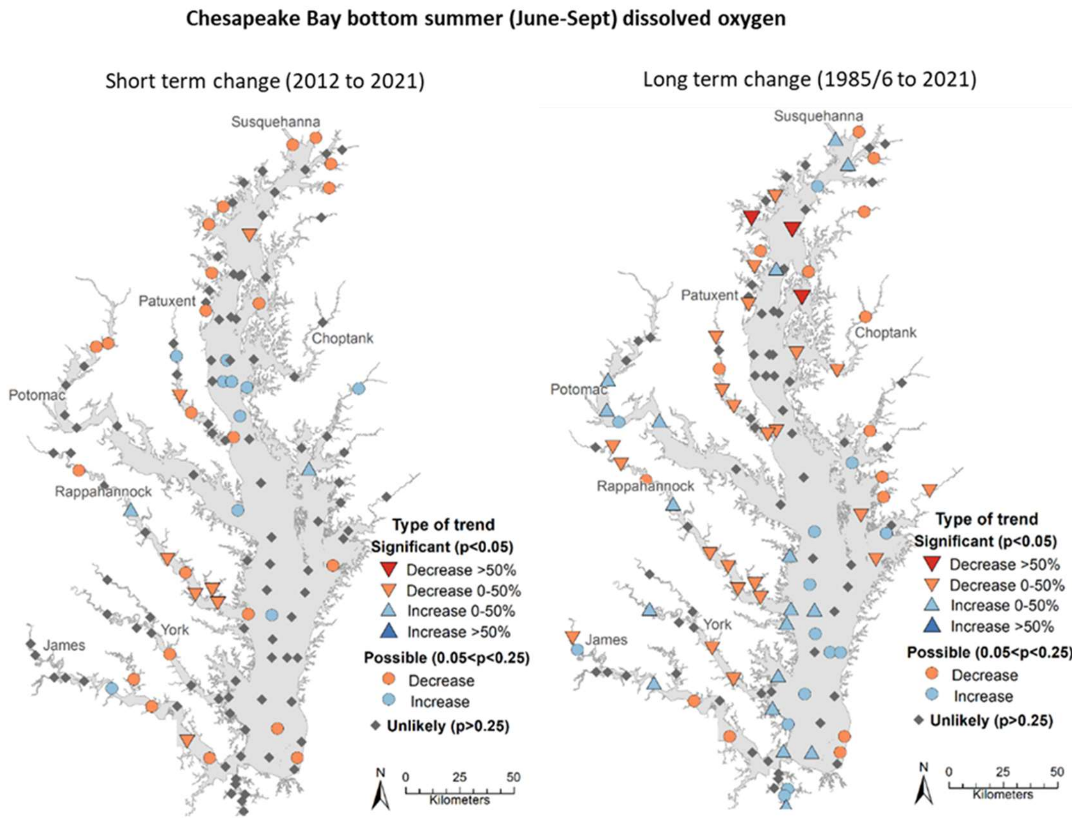


FIGURE 4.7.—Changes in DO in bottom water layer measured during June—September, short term (left pane) and long term (right panel); starting dates for long-term measurements vary (Source: CBP, n.d.-d).

The question of whether these reduced nutrient concentrations are being manifested in the attainment of WQS remains. In order to measure progress in meeting the WQS, a multimetric attainment indicator was developed by Zhang, Murphy, et al. (2018) to estimate combined standards attainment (DO, Chl *a*, and water clarity/SAV). Given the spatial variability in water quality trends, examination of this indicator at finer spatial scales yields additional insights into the connection between load reduction efforts and water quality response. Figure 4.8 presents the estimated percent attainment of indicators across different habitats of the Bay over the period 1985–2018 and shows differential patterns over time. Indicator attainment in the migratory and spawning habitat (blue line) was closer to 80% from mid-1980s through about

2000, but has declined afterwards towards attainment ranging from 60% to 80%. Indicator attainment in the adjacent open water DU (orange line) reveals a pattern of attainment below 50% before the early 1990s, but with values ranging from 50% to 70% since. Indicator attainment in the deep water (gray line) ranges from 20% to 40%, while attainment in the deep channel is lowest (ranging from zero to just over 10%). Increased frequency of non-zero values for the DO attainment indicator in the deep channel in figure 4.8 is notable, since it may suggest that nutrient load reduction efforts are beginning to initiate an ecosystem response. Figure 4.8 also shows that attainment of clarity/SAV in the shallows has slowly improved, with very few areas attaining before about 2001, but some increasing closer to 10% afterwards. Finally, Chl *a* in open water has no discernable pattern and ranges from zero (the majority of the time) with a few higher values over the period of assessment. It is especially instructive to compare attainment of DO criteria alone across the habitats; for the 30-year period of 1985–87 to 2014–16, annual changes in attainment of DO criteria are estimated to be -0.66% in migratory and spawning habitat, 0.00% in the deep channel, 0.10% in deep water, and 0.61% in open water (Zhang, Murphy, et al., 2018). These patterns suggest that specific habitats might differ in the timing of their attainment of WQS; for example, the progress of attainment depicted in figure 4.8 would suggest that the deep channel may be the last to reach attainment of DO standards.

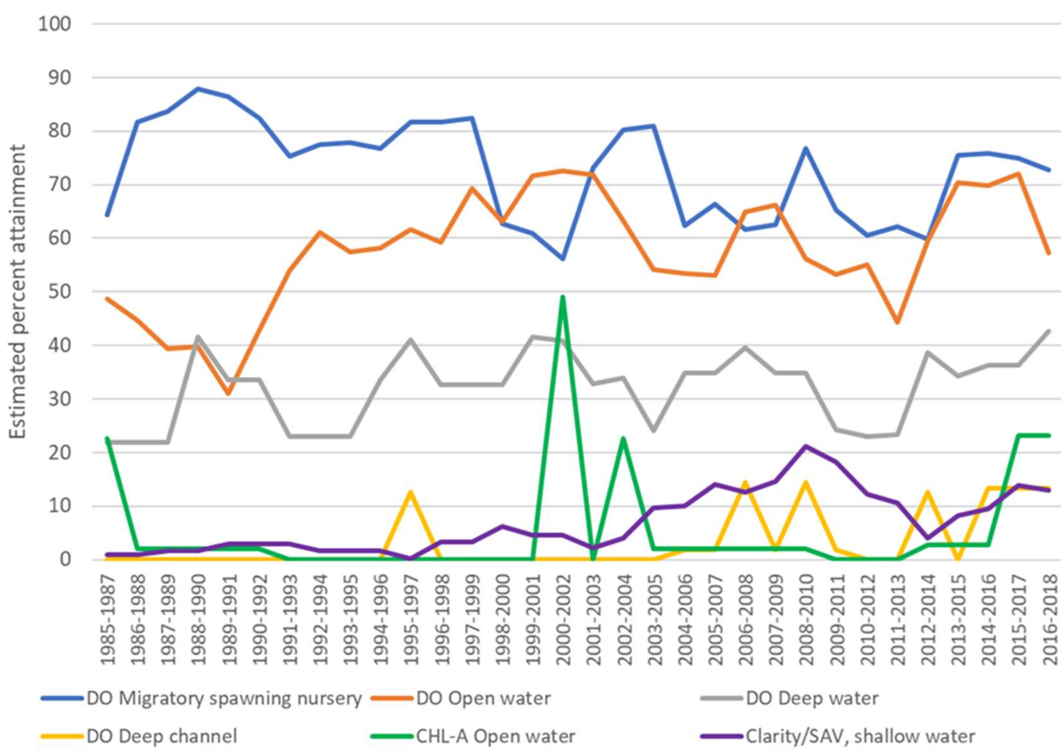


FIGURE 4.8.—Time series of area-weighted estimated WQC attainment for the five DUs, 1985–87 through 2016–18 (Source: Zhang, Murphy, et al. [2018] with updated data).

The lack of indicator attainment, especially that of DO WQS, across all habitats leads to the question of whether estuary water quality is responding in ways consistent with expected response to nutrient and sediment reductions achieved to date. Figure 4.9 illustrates one way to explore this question, albeit with a few notable caveats. In figure 4.9, the orange circles show estimates of expected DO standard attainment across a range of N loads for open water, deep water, and deep channel regions and represent expected responses of DO WQS attainment to load reductions provided by the CBP estuary model.[6] Load estimates were generated for both N and P, but figure 4.9 shows only the N values for simplicity. The observed response (attainment of DO WQS from monitoring results at a given estimated load) is represented by the blue diamonds and yellow squares in figure 4.9. The blue diamonds represent 3-year running mean of percent attainment of DO standards within each habitat during the period 1985–2020 (y value) and the corresponding 3-year running mean of annual watershed N load (x value) as obtained from the data sources referenced in figure 4.4 and described above. Yellow squares represent the same data in 10-year averages.

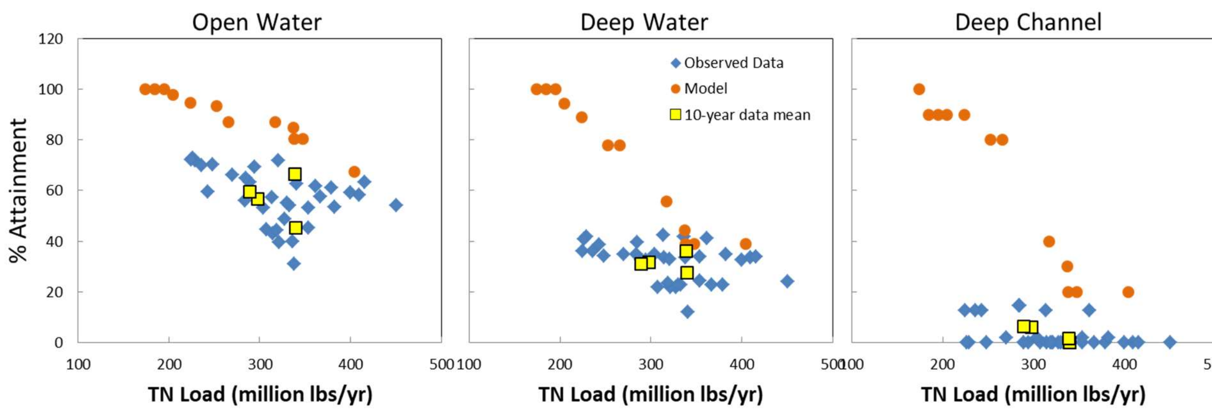


FIGURE 4.9.—Expected and realized relationships between TN loads and DO criteria attainment for open water, deep water, and deep channel habitat, calculated as 3-year running mean observed values (blue diamonds) and expected responses from estuary model (orange dots) for the same time periods. Yellow squares are 10-year means of the observed data.

The data in figure 4.9 suggest that the observed levels of DO standards attainment at any given N load are generally below the expected response (blue and yellow points are below the orange points), a potential estuary response gap. Three patterns are significant. First, a potential response gap exists in each of the three habitats (open water, deep water, and deep channel).

[6]These estimates were generated during preparation of the Phase III WIP planning targets (CBP Principals Staff Committee, December 19-20, 2017). They represent loads resulting from each of the 12 management planning scenarios and model year assumptions used in the TMDL development (USEPA, 2010, Section 6) but updated using the Phase 6 model. Points representing Model and Observed Data use somewhat different assumptions including: (1) Model points represent the estimated attainment (y-axis) for the hydrology period of 1993–1995, while the corresponding estimated load (x-axis) are for the hydrology period of 1991–2000; (2) Observed points represent the estimated attainment for the actual hydrology of each 3-year period while the Model is always the 1991–2000 hydrology (that is, the Observed is showing the response to management and weather, while the Model is just management response in wet years); (3) Temperature is rising over the Observed period, but not for the Model; and (4) Phosphorus load assumptions are not the same for Observed and Model.

Second, the response gap between expected (orange dots) and observed attainment for 3-year averages (blue diamonds) increases in size at lower N load levels. Third, the response gap at lower N load levels is significantly larger for the deep water and deep channel habitats as compared to the open water habitat, with the deep channel habitat exhibiting the largest response gap and percent attainment remaining below 20%. This suggests responses in DO attainment will vary by habitat as loads are reduced and need to be better understood. The lack of progress towards attainment in the deep channel suggests that management assumptions regarding the magnitude, speed, and timing of this response should be reconsidered.

The extent to which attainment of the DO WQC in the deep channel is lagging behind attainment in the open water and deep water can be seen in figure 4.10, which plots the attainment deficit (Zhang, Tango, et al., 2018) for these three habitats over three-year assessment periods from 1985 to 2021. While figures 4.7 and 4.8 provide a binary classification of attainment (full attainment or nonattaining), figure 4.10 provides additional understanding of actual conditions since it quantifies the nonattainment of DO standards over both space and time for all tidal segments. Among the three habitats, only open water showed a statistically significant long-term decrease in nonattainment (0.04% annually). In the deep water and deep channel habitats, neither the long-term nor short-term trends were statistically significant. In addition, the relationships between attainment of WQS in one habitat and its influence on the attainment of WQS in another are poorly understood. For example, water clarity in the shallows may allow for a benthic algae response that contributes to attainment of DO standards in the deep channel. Such a response could imply that attainment of WQS in the shallow habitats may accelerate response in other habitats including the deep channel.

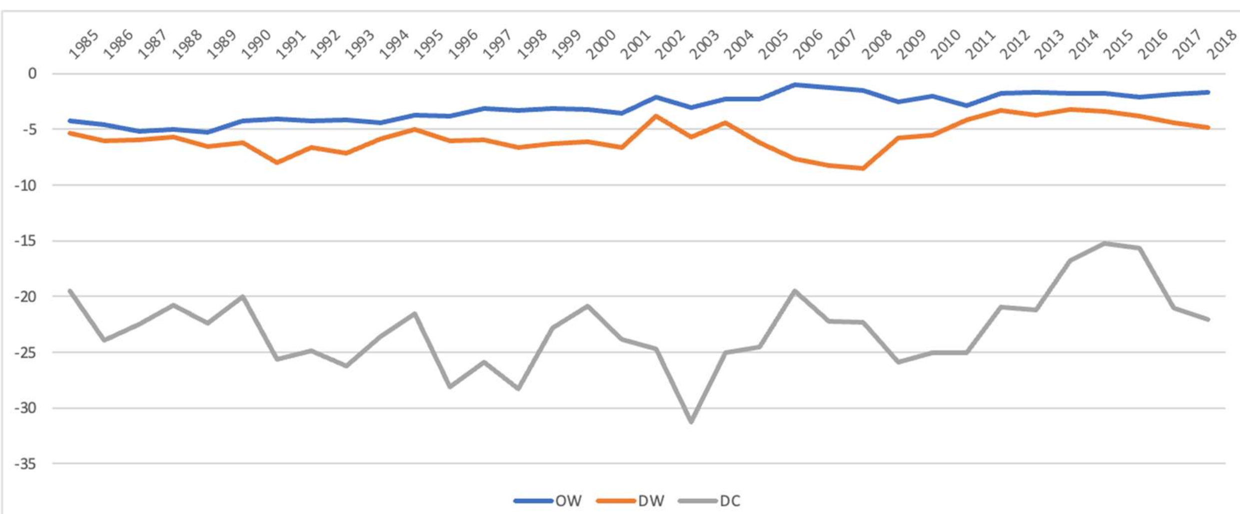


FIGURE 4.10.—Attainment deficit (0%=full attainment, -100%=full non-attainment) for DO WQC in three DU habitats, 1985–2020, calculated as 3-year running means. Year shown represents the first year of a 3-year assessment period (Source: Q. Zhang, personal communication).

Pronounced water quality improvements associated with nutrient reductions have occurred in localized regions of the Bay, however. For example, clear signs of substantial water quality improvement in some Chesapeake Bay regions have been associated with upgrades to WWTPs in specific portions of the Back River, Potomac River, James River, Patuxent River, and Patapsco River basins that have provided measurable and substantial reductions in nutrient loads, ambient nutrient concentrations, and algal biomass (Boynton et al., 2014; Fisher et al., 2021; Testa et al., 2022). The effects of these point source improvements were observed in waters local to the WWTP facilities, which are generally sited in brackish and tidal freshwater regions of tributaries. The identification of a response gap is much harder for water clarity and SAV because of the lack of a modeled expected response. The SAV goals were based on the historical presence of SAV for most areas. Baywide SAV trends for the tidal fresh, mesohaline, and polyhaline salinity zones are presented in figure 4.11; long-term decreases in nutrient loads have led to a significant increase in SAV, despite a high amount of interannual variability in loads and SAV coverage (Lefcheck et al., 2018; VIMS, n.d.-c). Most notably, SAV has shown clear recoveries in tidal fresh (salinity levels <0.5 parts per thousand [ppt]) and oligohaline (0.5 to 5.0 ppt) regions of the Bay (e.g., Susquehanna Flats) over the last several decades (Gurbisz & Kemp, 2014; Gurbisz et al., 2017) and more recently in mesohaline (5.0 to 18.0 ppt) regions of the Bay. The intense precipitation of 2018–2019, combined with declines in eelgrass, interrupted this positive trajectory in the polyhaline (18.0 to 30.0 ppt) Bay regions that are experiencing limited light availability and warming water temperatures (Lefcheck et al., 2017).

The variability in SAV cover in different salinity zones over the past several years (figure 4.11) illustrates the difficulty in trying to measure progress on shorter time scales. The SAV declines observed in 2019 were followed by gains in 2020–2021. The recent gains have not offset the 2019 loss, so there is not yet evidence that SAV has returned to a positive trajectory. Nevertheless, current SAV coverage remains well below the WQC goal of 185,000 acres. According to the 2021 SAV Report (Patrick et al., 2022), 27,528 hectares (68,025 acres) of underwater grasses were mapped in the Chesapeake Bay in 2021, representing 52% of the CBP 2025 restoration target of 130,000 acres and 37% of the 185,000-acre goal. While this does not represent a quantitative gap between an expected and realized response (i.e., the response gap as defined previously), it does represent a significant gap between current and desired SAV levels.

In summary, response gaps appear to be likely for DO in the open water, deep water, and deep channel habitats, with the largest response gap in the deep channel, and may be inhibiting progress toward achieving WQS. Quantification of a response gap for water clarity/SAV is not possible because of the absence of a formal predictive model, but progress remains well below the stated goal. Furthermore, while clear examples of TMDL-associated nutrient load reductions across the watershed have led to well-documented and substantial declines in estuarine nutrient availability and improvement in some measures of habitat (SAV) and water quality (DO), many tidal waters throughout the Bay show limited DO and water clarity response despite substantial reductions in regional and local stressors. The following section explores these response challenges.

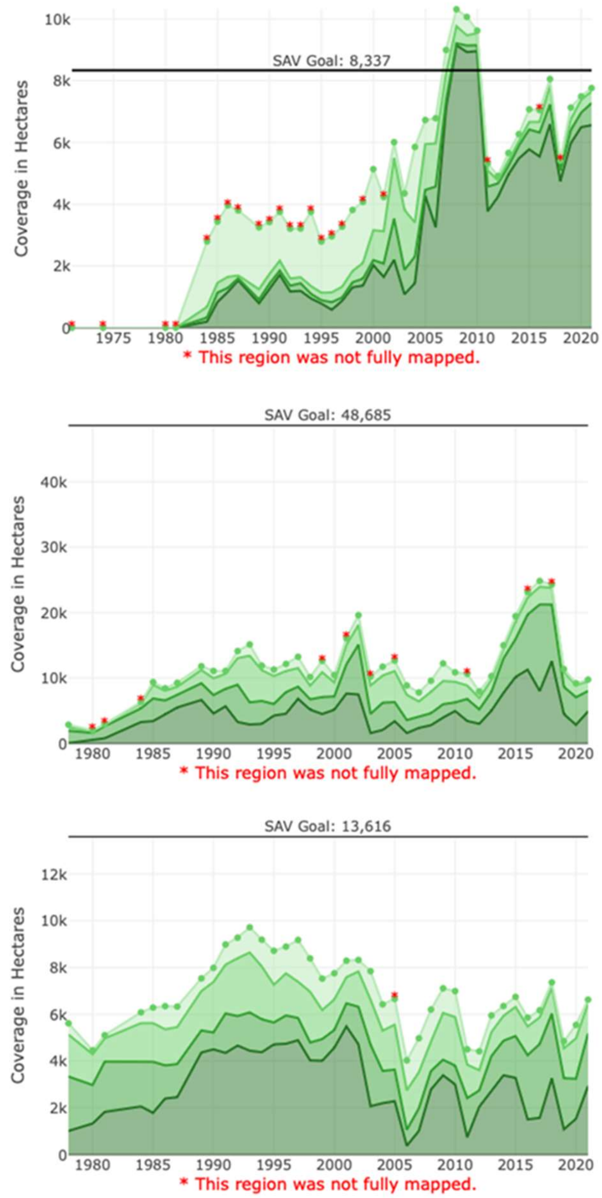


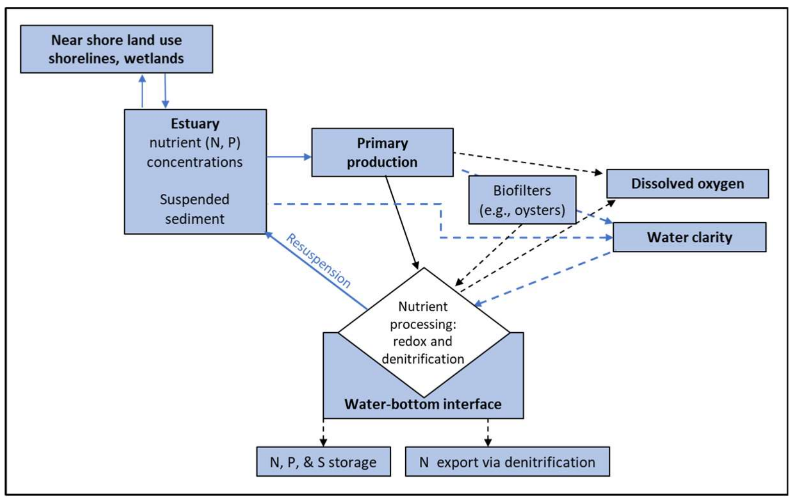
FIGURE 4.11.—Seagrass area (hectares) by salinity zone and in coastal bays from 1971 through 2021. Amount of seagrass abundance in tidal fresh (top panel), moderately salty (mesohaline—middle panel), and full salinity (polyhaline—bottom panel) areas. Density of cover is indicated by degree of shading: very dense (70–80% cover), dense (40–70% cover), sparse (10–40% cover), and very sparse (<10% cover) (Source VIMS, n.d.-b).

4.3. Response gaps and uncertainties in achieving WQS

The lack of sustained progress toward achieving WQS across all habitats strongly suggests that estuary water quality is not responding to load reductions achieved to date in ways consistent with current assumptions. This led STAC to examine why these gaps exist and what learning may help us close the gaps (i.e., improve the effectiveness of water quality actions). Response gaps, identified quantitatively for DO and qualitatively for SAV/water clarity, may result from limited ability to accurately assess the degree of attainment due to gaps in monitoring, problems estimating or measuring loads, or failure to account for other factors that may be confounding the desired response (including complex estuary processes or the influence of climate change). Each of these is addressed in the following sections, and parts of figure 4.3 are reproduced to highlight where in the estuarine system each challenge may be manifested.

Confounding factors and nonlinear interactions

Uncertainties around underlying estuary processes may also be partly responsible for water quality response gaps. Processes presented in figures 4.2 and 4.3 can operate differentially between habitats and within specific locations of each. The variability in attainment of WQC shown in figure 4.8 suggests that the relationship between nutrient and sediment load reductions and achievement of WQS is unique to each habitat.



Nonlinear interactions may explain response gaps, especially when internal processes in the Bay create patterns where water quality improvements are slowed. Interactions between nutrients and a host of other factors, such as light availability, algae, turbidity, DO concentrations, redox conditions, and ecosystem engineers like oysters and SAV habitats (as shown in figure 4.3) may slow or speed up responses to nutrient reductions (Cerco & Noel, 2007; Kemp et al., 2005; Newell, 1988; Testa & Kemp, 2012). For example, the depletion of DO leads to changes in chemistry that slow nutrient removal and speed up nutrient recycling, essentially increasing the number of times that any N or P molecule can be used to sustain eutrophication. In these cases, oxygen depletion supports a self-sustaining cycle; Testa and Kemp (2012) and Ni et al. (2020) associated this cycling with a potential doubling of the volume of hypoxic water generated per unit of TN load in the past half century. Feedback mechanisms such as this may help to explain particular spatial and temporal patterns, even during years of reduced N loading.

In addition, the abundance and condition of living resources (e.g., oysters) and habitats in the terrestrial-aquatic interface (e.g., tidal marshes) add additional controls and feedback mechanisms to both water clarity and nutrient cycling processes. The absorption or removal of nutrients from tidal waters by both oyster communities and tidal marshes has been documented (Cerco & Noel, 2007; Parker & Bricker, 2020). Losses of these elements from the Bay ecosystem over time would make the Bay more sensitive to a given nutrient load from the watershed. Furthermore, recovery of biological communities within the ecosystem that modulate the timing and amount of available nutrients and organic carbon will affect nutrient concentrations (Basu et al., 2022).

A significant consideration in conceptualizing response curves by habitat is that feedback mechanisms illustrated in figure 4.2 can result in ecological tipping points (or thresholds), generally defined as ecosystem states where small changes in environmental conditions result in large or rapid shifts in ecological status or function. When such tipping points exist, the relationship between nutrient and sediment reductions and attainment of WQS will be nonlinear and may create wide intervals of uncertainty around expected response. Tipping points are a reality at the scale of localized regions of the Bay and have been demonstrated in water clarity, DO, and SAV relationships (Boynton et al., 2009; Ganju et al., 2020; Gruber & Kemp, 2010; Kemp et al., 1990; Testa & Kemp, 2012). For a specific example, see the case study of Mattawoman Creek in text box 4.1, where a multiyear lag existed between nutrient load reduction and ecosystem recovery, but once the recovery began, Mattawoman Creek proceeded rapidly from a degraded algae-dominated system to a restored, SAV-dominated system.

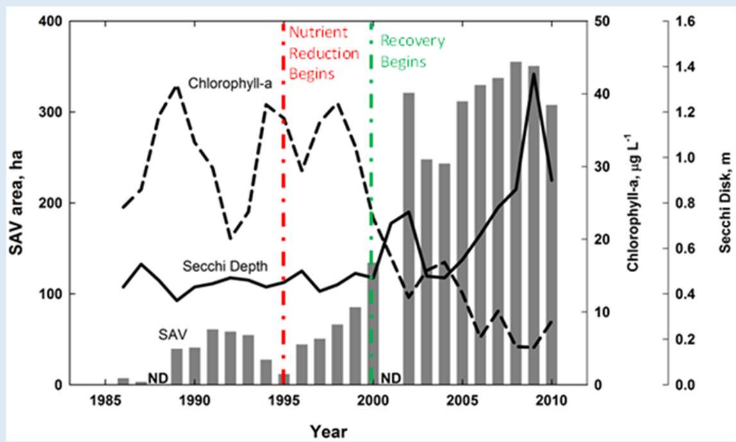
Tipping points can indicate progressive conditions on either a degradation or restoration trajectory (for an excellent summary, see Kemp et al. [2005]), and factors that can contribute to the type of nonlinear response seen in the recovery in Mattawoman Creek are shown in figure 4.12. The left side of figure 4.12 illustrates how positive feedback interactions tend to reinforce and accelerate the eutrophication process (i.e., explain response gaps), while the right side shows how they can reinforce the restoration process by enhancing water quality improvements once they are initiated. Exceeding critical thresholds in the direction of restoration is of particular interest now. A recent analysis (Lefcheck et al., 2018) linked nutrient reductions to the resurgence of SAV in the Chesapeake Bay, providing a lesson on staying the course, allowing time for benthic communities to recover and cross tipping points that enable self-sustaining processes to initiate. We note that figure 4.12 includes only one impact of climate change, namely sea level rise. It does not account for water temperature increases or hydrologic changes from either climate change or land cover shifts; it also does not include oyster biology.

An implication of all these uncertainties and challenges is that closing the response gaps identified in this report may not be feasible in all of the habitats and segments of the estuary. For example, the very low DO levels in the deep channel may be a historic feature, and DO expectations for this habitat may not align with the dynamics of the rest of the Bay ecosystem. Changing conditions for SAV, including temperature shifts discussed below, result in species shifts due to habitat suitability ranges and may make expectations of SAV cover that are based on historical conditions no longer attainable.

Text box 4.1. Shallow waters and tipping points at work: Mattawoman Creek

Mattawoman Creek, a small tributary branching off the upper Potomac River, has shown a dramatic restoration in recent decades. Nutrient levels were historically high in this shallow tributary due to high nutrient inputs from the adjacent Potomac River and small WWTPs within the Mattawoman Creek watershed. As a result, water clarity was poor (Secchi depth < 0.6 meters), SAV was absent, and the water was filled with an abundance of microscopic algae. In the mid- to late-1990s, nitrogen reductions began in earnest, and an extended drought period in 1999–2002 contributed to drops in N loads. This extended period of reduced nutrient loads produced a decline in algal biomass and a correlated increase in water clarity. The increase in water clarity supported the resurgence of SAV, assisted by the presence of an invasive exotic species (*Hydrilla*) which can take advantage of short-term periods of water clarity for establishment (Boynton et al., 2014).

The improvements in the ecosystem of this shallow creek following sustained nutrient reductions reveal how rapid recovery can occur in regions of Chesapeake Bay where modest improvements in water clarity support SAV expansion, while large-scale recovery may lag several years following nutrient load reductions. This case study is an example of a water clarity feedback as expressed in figure 4.12 (Kemp et al., 2005). Comparable recoveries occurred in other regions of Chesapeake Bay, including the enormous Susquehanna Flats SAV bed.



Long-term patterns of chlorophyll *a*, Secchi depth, and SAV coverage in Mattawoman Creek, 1985–2010. (Source: Boynton et al., 2014).

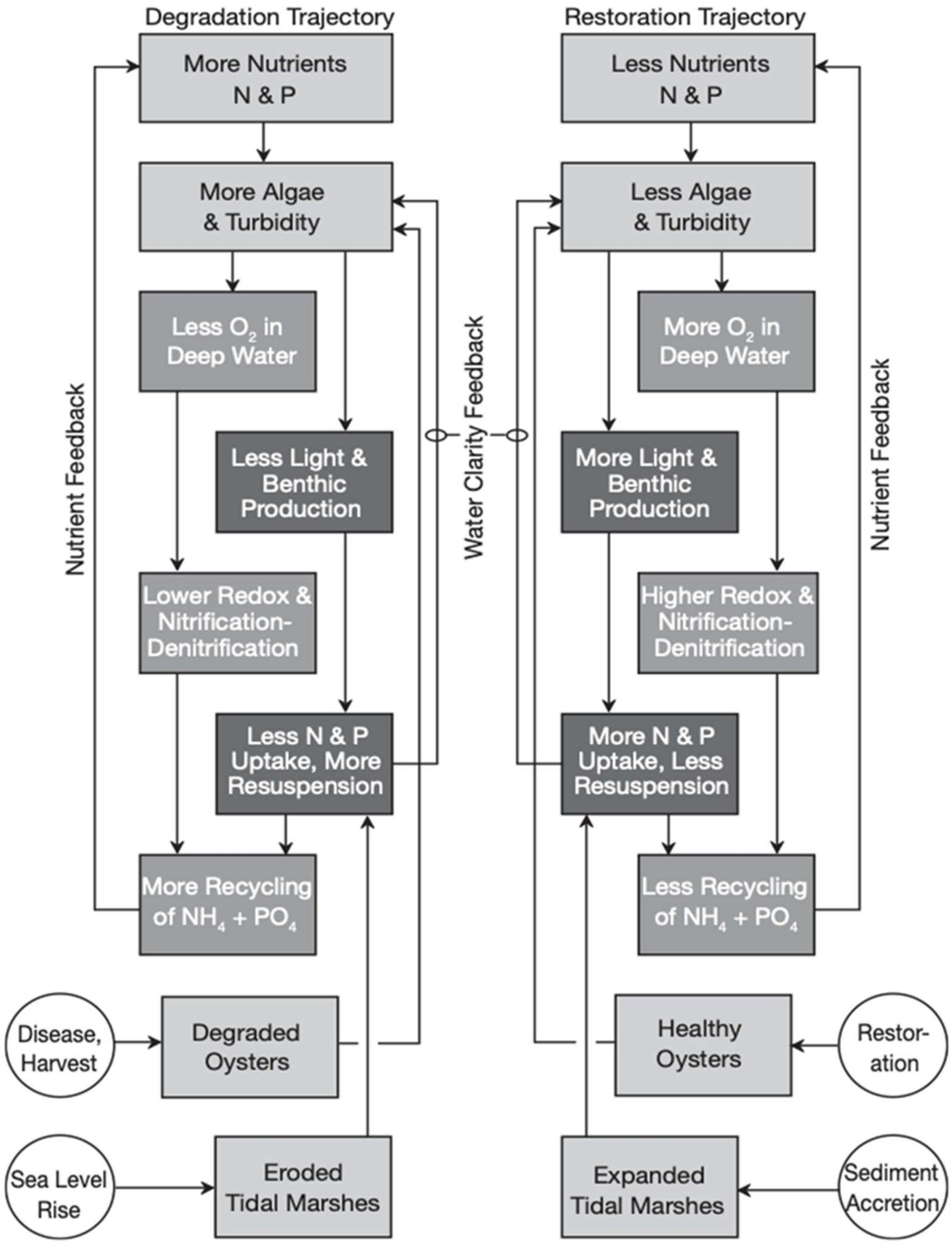
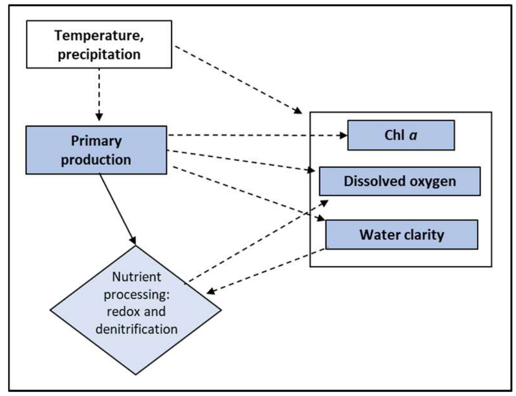


FIGURE 4.12.—Effects of N and P additions on physical, chemical, and biological elements of the estuarine system, including algal biomass, bottom water oxygen, and nutrient recycling. Effects of climate change (save sea level rise) and land cover change are not included (Source: Kemp et al., 2005).

Climate change

Climate change is a global challenge for ecosystem restoration with ever-present effects in the Chesapeake Bay. Climate change is no longer a future threat: the Bay has already warmed, has higher sea levels, and is receiving altered patterns and magnitudes of precipitation (Batiuk et al., 2023; Fleming et al., 2020; Hinson et al., 2022; Najjar et al, 2010). Climate change effects interact with nutrient management challenges to further test our understanding of system response.



Dissolved oxygen concentrations are highly sensitive to climatic processes both directly and indirectly, and climate change effects are likely to make it increasingly difficult to meet DO criteria. For example, changes in wind direction, precipitation, and water temperature may result in climatically forced variability in stratification and the associated supply of DO to bottom waters (Du et al., 2018; Hinson et al., 2022; Scully, 2010); none of these factors are explicitly recognized in figure 4.9. Additionally, it appears that warming over the past 35 years has limited the otherwise positive oxygen response to the TMDL (Frankel et al., 2022; Ni et al., 2020). A study by Frankel at al. (2022) quantified the impact of watershed N reductions on Bay hypoxia during a recent period including both average discharge and extremely wet years (2016–2019) (figure 4.13) and concluded that if the N reductions since 1985 had not occurred, annual hypoxic regions (i.e., areas where O_2 < 3 mg/L) would have been much greater. The discrepancy between our expected responses and current conditions may be largely accounted

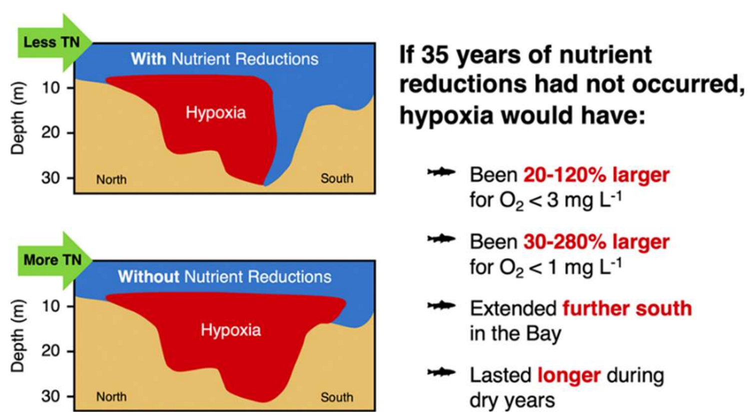


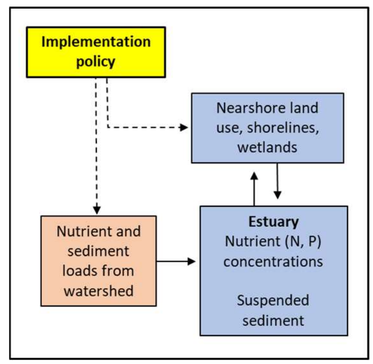
FIGURE 4.13.—Estimated extent of Chesapeake Bay hypoxia with and without 35 years of nutrient reductions (Source: Frankel et al., 2022).

for by warming in the Bay that has offset roughly 6–34% of the improvement from N reductions (Frankel et al., 2022). The effects of climate change on the Bay will be felt far beyond oxygen levels and will include shoreline erosion, alteration of productivity, disease impacts, species migrations, among many other effects (Batiuk et al., 2023).

Additional factors

Climate change, nonlinear actions, and other confounding factors will continue to challenge predictions of what water quality conditions will be in the future. Furthermore, some habitats are more sensitive and vulnerable to these changes than others. For example, shallow water habitats may be particularly sensitive to changes in warming. Shallow water habitats may also be more capable of achieving nonlinear responses (tipping points) from management actions. Given the need for enhanced understanding of the impacts of both nonlinear interactions and climate change and the importance of these habitats to living resources, nearshore shallow water habitats may be particularly effective test beds for addressing uncertainties and improving water quality response (text box 4.2).

Ability to estimate or measure loads delivered to Chesapeake Bay



The exercise of assessing response gaps, such as those suggested by figure 4.9, highlights the importance of accurately measuring nutrient and sediment loads. Without accurate measures of loads, prediction of responses is more difficult. The size and complexity of the Bay make accurate determination of loads a challenge: monitoring does not allow all loads to be measured directly. Several pieces of information are used to estimate total loads to the estuary, including the RIM results, discharges from WWTP, and estimates of non-point source loads. Estimates of the loads originating from the watershed above the RIM stations are judged to be fairly accurate. RIM loads are estimated to account for about 60% of the total load of nutrients and sediments to the estuary. Additionally, discharge data are available from WWTP point sources below the RIM stations. From these two information sources, it is estimated that 84% of the load is observed through measurement of flow and effluent concentration. The remaining nonpoint sources of nutrients and sediments (occurring below the RIM stations) are provided using the CBP watershed model. In addition, nutrients and sediment may be transported from coastal terrestrial environments into adjacent aquatic environments during tidal inundation; a recent modeling effort estimated that the amount of dissolved inorganic N contributed during a perigean spring tide (seasonally high tide or king tide) event in one Bay segment exceeded its annual N load allocation (as specified in the appropriate WIP) by 30% (Macías-Tapia et al., 2021).

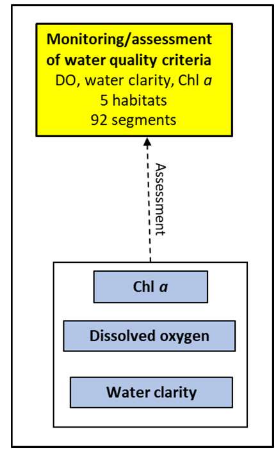
Text box 4.2. Nearshore shallow water habitats as test beds for restoration and addressing uncertainty

One habitat deserves particular attention because of its positive response to stressor reductions, its potentially outsized role in Bay restoration, and its relatively high attainment of applicable WQC (Figure 4.8). This habitat encompasses the land-sea interface, including the nearshore regions of the watershed, tidal marshes, and shallow, nearshore habitats in tidal waters, including much of the migratory and spawning habitat; for the purposes of this report we combine these regions under the umbrella term of nearshore shallow water. This habitat supports extensive expanses of SAV beds, oyster reefs, and wetlands, and is heavily influenced by an area termed the terrestrial-estuarine transition zone, or T-zone. The T-zone is defined as "the area of existing and predicted future interactions among tidal and terrestrial or fluvial processes that result in mosaics of habitat types, assemblages of plant and animal species, and sets of ecosystem services that are distinct from those of adjoining estuarine, riverine, or terrestrial ecosystems" (Goals Project, 2015, p. 59). Notably, the T-zone extends much farther inland than coastal wetland-upland boundaries and includes areas affected by tidal surges, well inland of tidal freshwater wetlands. On the western shore of the Chesapeake Bay, the T-zone extends to the fall line, whereas on the eastern shore, the flat topography imposes limited constraints to the T-zone's inland migration as sea level rise occurs. On the western side of the Bay, the T-zone currently extends more than 30 miles inland from the mouths of the major tributaries.

These transitional zones have recently been recognized as critical but overlooked landscape areas connecting watersheds and coastal waters because their waterways and adjacent river corridors represent critical hydrologic connections that concentrate watershed discharge, regulate exchange, and influence the distributions of living resources (Boomer et al., 2019). The dynamic hydrochemical gradients created by shifting circulation patterns and changes in flow velocities, along with a cascade of biogeochemical constraints along the T-zone waterways, create a diversity of habitats that influence nutrient and sediment delivery to the Bay's mainstem. Most of these processes primarily affect the dynamics of P, arguably the most critical element driving algae blooms and eutrophication in the low salinity regions of estuaries (Carpenter, 2008; Froelich, 1988), and there is clear evidence for benthic recycling hotspots in eutrophic triblets (Boynton et al., 2009; Testa et al., 2019). While we have a conceptual understanding of how these hydro-chemical gradients influence biogeochemical processes and nutrient dynamics, less is known about when and where they occur and the direction and magnitude of biogeochemical processes within a complex, nested estuarine and triblet system like the Chesapeake Bay (Statham, 2012; Wang et al., 2016; Ward et al., 2020).

These combined T-zone/shallow tidal habitats can serve as test beds for integrating the land-sea interface into investigations of tipping points and climate change using monitoring, modeling, and research approaches. The planned CBP advances in the Phase 7 modeling suite, which include increases in spatial resolution and process representation in these shallow systems, will facilitate this new research and management focus. Coupling this emphasis with increased engagement with stakeholders who spend appreciable amounts of time interacting with these habitats can help to identify a wider range of restoration outcomes to meet WQS and prioritize living resources, recreational uses, and habitat restoration.

Ability to accurately assess degree of attainment



The sheer amount of monitoring required to measure WQS at all 1,052 assessment points (figure 2.4) presents an enormous challenge. The traditional CBP partnership's long-term water quality monitoring and assessment program was deemed "marginal" for measuring and reporting on all published criteria to produce a complete accounting of Bay health (USEPA, 2003b), and efforts since then have included adjustments to the monitoring network and analyses to estimate attainment from available data. However, despite involvement of multiple partners, the monitoring and assessment program has never been focused at the temporal and spatial scales necessary to provide a full assessment for ANY of the 92 segments in the Bay (CBP, 2022). Thus, the program has essentially been unable to effectively monitor and assess all required WQS, and their responses to CBP implementation, for all habitats associated with the TMDL (the yellow boxes in figure 4.3).

A multimetric indicator was developed (Hernandez Cordero et al., 2020) to estimate attainment of WQC using available monitoring information until a more complete accounting could be supported by more comprehensive monitoring and assessment protocols. However, without additional monitoring information and associated research efforts, the CBP cannot fully assess progress towards (1) attaining WQS; (2) understanding the relationship between load reduction efforts and achieving WQS; and (3) achieving additional water quality conditions to support crabs, oysters, and other fisheries in the estuary. Additional monitoring efforts would be needed to also make connections to other habitat conditions across management-relevant temporal and spatial assessment scales needed by diverse CBP partner interests and GITs (i.e., Water Quality Goal Team; Sustainable Fisheries Goal Team; Scientific, Technical Assessment and Reporting Climate Resiliency Workgroup) (CBP, 2022).

In addition, water quality monitoring is not carried out evenly across habitat types, leading to differential understanding of the processes controlling responses to nutrient and sediment reductions. For example, 17 of the 31 outcomes expressed in the 2014 CBWA (CBP, 2014) are related to shallow water habitats, yet that habitat type is not adequately monitored, modeled, or studied. Construction and implementation of a new estuarine model with improved capability in shallow waters, generally defined as the upper 2 meters of open water habitat, is underway. However, monitoring of shallow waters is challenging. First, there is limited boat or road access to smaller waterways (termed triblets) draining catchments that typically range from 50 to 150 square kilometers in size and support extensive expanses of SAV, oyster reefs, and wetlands. Second, the sheer number of triblets, along with the variability of physiochemical conditions within each triblet system, presents a challenge for developing a compelling framework for collecting and interpolating observation data. Recent technological advances along with a model-based sampling framework could address the lack of data for these systems.

Monitoring protocols are based primarily on status and trends concepts, and thus are more attuned to accountability objectives surrounding achievement of WQS than to understanding ecosystem processes that link water quality to load reduction efforts. In other words, current monitoring may answer the question of whether we have met criteria for a given nutrient or DO concentration, but if we haven't met those criteria, the monitoring program does not provide measurements that would allow a greater understanding of why.

4.4. Conclusions and Implications

Annual loads of nutrients and sediment to the estuary vary considerably due to variations in precipitation, but declines in N, P and sediment loads over the past 40 years (though modest in some cases) have generally resulted in lower nutrient concentrations throughout the estuary. Even without sustained periods of low loads, the majority of estuary monitoring stations have seen reductions in surface water N and P concentrations over the long term, and spatial variation in TN and TP concentrations in tidal waters generally reflect "where nutrient loads decline, estuarine concentrations will decline." However, these trends in N and P concentrations do not necessarily translate into improving trends in WQC variables; most stations have seen no change or degrading conditions in DO in the bottom layer, annual Secchi depth, and spring-season surface layer Chl *a*. When attainment of WQS is examined at the scale of individual habitats, clear differences emerge, with attainment of WQS relatively intransigent for DO in the deep water and deep channel, suggesting that these habitats may be the last to reach attainment of their DO standards.

The lack of attainment of WQS across all habitats has implications about the assumptions of how estuary water quality will respond to the nutrient and sediment reduction targets. Based on the likely estuary response gaps for DO in the open water, deep water, and deep channel habitats, the current load reduction targets may not be adequate to attain WQS in all areas of the estuary. The largest potential response gap is in the deep channel, which implies additional nutrient reductions may be required to attain any given level of DO response. Further, the CBP could consider putting additional focus on nutrient and sediment reduction in shallow and open waters since the likely response gap is smaller in these habitats, and because attainment of WQS in the shallow habitats may accelerate response in other habitats including the deep channel. Quantification of a response gap for water clarity/SAV is not possible because of the absence of a formal predictive model, but progress remains well below the stated goal. Various assessments of progress toward WQS attainment indicate that complete attainment is not likely to happen soon.

A critical element in the resolution of these uncertainties is monitoring information. The monitoring program has never achieved assessment of WQS at the temporal and spatial scales necessary to provide an assessment of attainment for any of the 92 segments in the Bay. Ongoing additions and revisions to the monitoring program are aimed toward being able to assess attainment in selected segments in 2026 (CBP, 2022). Even with these changes, the monitoring program would not provide answers to questions about how the system is working. Current monitoring protocols are more focused on accountability objectives than on understanding processes. For example, the existence of tipping points in subsystems of the Bay is well-documented, but monitoring to determine the thresholds associated with either degradation or restoration is not currently done.

The CBP could consider enhancing the estuary monitoring network to better document estuary processes in addition to improved assessment of WQS attainment. The enhanced monitoring, coupled with associated research, could cover the range of local conditions where proximate and distant factors can be characterized and causal relationships between stressor reductions and WQS attainment can be determined. In addition, monitoring protocols would need to be aligned with modeling approaches. The most effective approach for identifying relationships between stressor reduction efforts and WQS attainment may be to structure monitoring around subsegment or smaller spatial scales. At these smaller scales, clear signs of successful water quality remediation in some Chesapeake regions, such as reductions in nutrient concentrations and algal biomass, have been associated with upgrades to WWTPs, but other areas with similar reductions do not demonstrate the same type of response. Similarly, recent patterns in TN load and concentration reductions have been linked to a resurgence of SAV in many small-scale regions of the Bay (Lefcheck et al., 2018), but the relationships between stressor reduction and living resource response are unique to salinity regimes, SAV species distributions, and water temperature.

The largest source of uncertainty about the relationship between nutrient reductions and WQS attainment in each habitat may be the existence of tipping points, which have been demonstrated in localized regions in the Bay. Tipping points exist in water clarity, DO, and SAV relationships, and can indicate progressive conditions on either a degradation or restoration trajectory; crossing the implied thresholds in the direction of restoration is of particular interest now.

The shallow waters present an opportunity for the CBP to reduce critical uncertainties about the effects of load reductions on living resources in an area of high engagement by stakeholders who live or recreate in these areas. In these habitats, DO and SAV respond to many variables beyond nutrient and sediment reductions, and understanding these dynamics is critical for both identifying effective management actions in the shallow waters themselves and understanding the relationships between shallow water habitats and the attainment of WQS in other habitats (e.g., deep water and deep channel DO). The CBP could put more emphasis on the shallow waters, both in terms of load reductions and scientific understanding, to accelerate and better understand attainment of WQS and benefits to living resources.

5. Living Resource Response to Water Quality Conditions

The support and enhancement of aquatic living resources (the designated use of the Chesapeake Bay, in the language of the CWA) is the ultimate goal of efforts to reduce nutrient and sediment loads and achieve WQS in the Bay. As described in the previous chapters, most regions of the Chesapeake Bay have seen long-term declines in nutrient levels but more mixed results in achieving long-term improvements in water quality. Chapter 4 described potential response gaps and deficits in attaining the WQS, particularly DO standards in the deep channel. An important question at the center of Bay water quality management efforts is: How will Bay living resources respond (e.g., increased abundance, diversity, and resilience of desired aquatic organisms) to efforts to improve water quality conditions? Answering this question would enable us to provide insight into the direct policy-relevant question: What water quality management actions could accelerate improvement in Bay living resources?

The DUs in the WQS are generally stated as supporting "recreationally, commercially and ecologically important species" in the various Bay habitats (USEPA, 2003a).[7] The response of living resources to water quality conditions is most easily expressed for those species with relatively simple life cycles and for which WQS can be easily derived based on their most basic requirements; an example is SAV and its companion WQS of water clarity. For the remaining species of living resources (e.g., blue crabs, oysters, finfish), the responses are more complicated.

The influence of multiple factors and stressors makes detecting and isolating living resource responses to changing water quality conditions challenging. In recognition of this challenge when establishing the Chesapeake Bay WQS, drafters organized the WQS around the water quality conditions (e.g., criteria for DO) necessary to support an individual of a species in specific habitats. It was understood that meeting the specific WQS would, in turn, provide water quality conditions to support specific living resources, but generally not to generate a specific living resource response.[8] The specific numeric WQC (measures) selected for the Bay were based on laboratory studies of species tolerances and habitat affinities for regions throughout the Bay (Monaco et al., 1998; Tango & Batiuk, 2013; USEPA 2003a, 2003b). Bay WQS can be viewed as necessary but not inclusive of all the conditions needed to support different groups of species in different habitats.

This chapter begins by briefly identifying some basic conceptual ideas that help structure our understanding of, and frame public expectations about, how much improvement in living resources can be expected from the water quality changes targeted by the WQS. Unlike for pollutant loads and estuary water quality conditions, the CBP does not use a comprehensive

[7] Specifically, the DUs aim to support particular finfish (including migratory and nonmigratory species, game fish, prey species across multiple habitats), crabs, oysters (open water, shallow water, deep water), and underwater bay grasses (shallow water) and to maintain populations of sediment-dwelling worms and small clams (deep channel).

[8] The exception is SAV criteria.

analytical framework to systematically evaluate the relationship between water quality management actions and living resource response. The next section provides a brief description of the conceptual relationship between living resource response and CBP water quality management efforts. This is followed by illustrative examples of statistical and ecological modeling analyses that have examined the Bay living resource response to water quality and other factors. The chapter concludes by noting analytical options for more fully examining responses of Chesapeake Bay living resources to water quality, habitat, climate, and other factors, and other additional management actions.

5.1. Conceptual model of living resource response to water quality conditions

Figure 5.1 shows a conceptual representation of the possible responses of living resource abundance to attainment of Bay WQS (DO, Chl *a*, and water clarity). The horizontal axis shows the percent attainment of multiple WQS across the 5 habitats and all 92 segments in the Bay. The overwhelming majority of those assessment points involve DO. Furthermore, existing empirical evidence and water quality models suggest that increases in attainment (moving left to right on the horizontal axis) will not occur evenly across the estuary and across habitats. As described in chapter 4, as attainment moves toward 100%, the last remaining nonattainment areas will likely be the deeper water habitats (deep water and deep channel). The vertical axis shows one conceptual metric of living resource response, overall abundance. Other living resource metrics, such as biological diversity, could also be used to illustrate similar conceptual relationships.

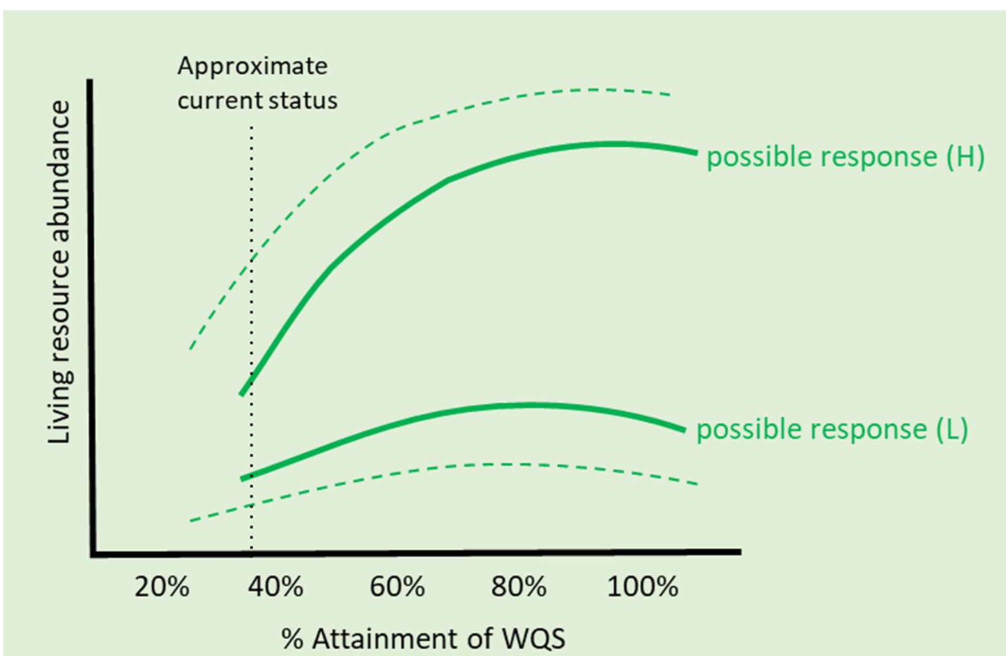


FIGURE 5.1.—Conceptual living resource response to attainment of WQC.

Figure 5.1 reflects the general expectation that achievement of WQS will increase living resource abundance over a certain range. While the CBP does not identify a specific living resource response from water quality changes (i.e., no *expected response* as the term is used in this report), reducing areas of low DO expands habitat, improves organism growth rates and reproductive success, and can decrease mortality (Breitburg et al., 2018). The pattern of attainment discussed in chapter 4, in which the percent attainment is currently highest for the open water habitat and lagging most for the deep channel, suggests that percent attainment of WQS may proceed across habitats somewhat sequentially; that is, a high level of attainment will occur sooner for the open water habitat, and attainment in the deep channel habitat will occur last. Since some regions and habitats of the Bay are used more extensively by aquatic organisms than others (shallow water habitats, for example, support a wide range of fish species across key life stages), the rate of improvement in living resources is likely to change as attainment progresses and nonattainment areas shrink toward the deep-water habitats (in figure 5.1, moving left to right on the horizontal axis). Of note, living resource abundance may eventually decrease as nutrient loads continue to decrease and WQS approach full attainment. For example, lower nutrient loads needed to achieve 100% attainment of DO standards may restrict primary production at the base of the food chain in the more productive shallow water areas and, at some point, act to limit or diminish the biomass of some species further up the food chain (Breitburg, Craig, et al., 2009; Breitburg, Hondorp, et al., 2009).

Considerable uncertainty will accompany any effort to predict how fish and shellfish populations respond to changes in water quality alone (represented by dashed green lines in figure 5.1). Estimating the abundance of various species throughout the Bay ecosystem is challenging and subject to measurement error. Complex biological systems respond to variation caused by stochastic events like weather as well as a host of other factors besides water quality, and accounting for all the major sources of variation in living resource response is challenging.

Figure 5.1 shows two possible responses of living resource abundance to achievement of Bay WQS. The response may resemble curve H, which shows relatively rapid improvements in living resources in response to water quality improvement, while response L shows a more limited living resource response to improving water quality conditions. Whether the living resource response resembles H or L depends on a number of other biological, physical, and management factors besides water quality.

Figure 5.2 shows a simplified system diagram of the major factors that influence the abundance and composition of aquatic living resources. Starting at the left, a number of estuarine water quality conditions impact the aquatic habitat of living resource communities. The factors in yellow are explicitly managed under Bay WQS (DO, water clarity, Chl *a*). Water temperature can have a large influence on the type and abundance of living resources (Ihde & Townsend, 2017). Water temperatures are increasing in the Bay, and increases are projected to continue with climate change (Batiuk et al., 2023) Water temperature increases are most pronounced in shallow water habitats, and these areas are important to forage fish and many fish life cycles.

Other major water quality conditions, such as salinity and pH, can also have a pronounced influence on the habitats used by living resources.

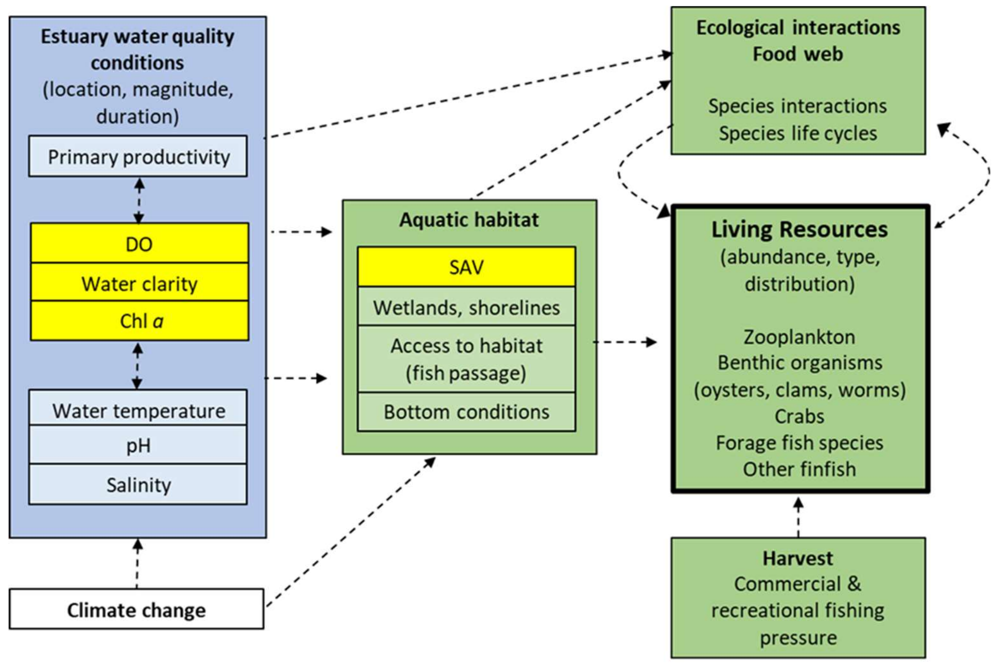


FIGURE 5.2.—Factors affecting the composition and abundance of living resources.

The water and nearshore structural habitat conditions also directly impact living resources (middle section of figure 5.2). For example, coastal wetlands benefit a wide range of aquatic animals.[9] Conversely, shoreline hardening is associated with diminished abundance of a number of important recreational and forage fish species (Kornis et al., 2017). Fabrizio et al. (2021) found that shallow water habitats are particularly important for a number of key forage fish species. In addition to habitat, additional factors affect living resources (right side of figure 5.2). One important factor impacting living resource abundance is fisheries harvest. For important commercial and recreational fish species, harvest is often a primary driver. Ecological and food-web interactions will also influence living resources.

Species respond differently to water quality and habitat conditions, and tradeoffs between different species occur as these conditions change. Species, as well as their specific life stages, exhibit varying tolerances across different water quality conditions (DO, water temperature) and habitat conditions. For example, habitat modeling studies suggest some species (striped bass, bay anchovy, oysters) may be much less sensitive to changes in the volume of hypoxic waters than bottom dwelling Atlantic sturgeon (Schlenger et al., 2022).

[9] The CBWA includes numeric wetland restoration goals but not specifically for tidal wetlands.

5.2. Evidence and analysis of water quality impact on living resources in Chesapeake Bay

Figure 5.3 shows historical indices of fish abundance for four well-known species in the Bay: blue crab, striped bass, summer flounder, and bay anchovy. The graphs show the average weight of fish caught per standardized unit of fishing effort (i.e., trawl) from 1988 through 2016. The data series shows significant annual variation in these species and different trends across each. For example, the population of striped bass varies annually, and there is also a long-term trend showing a decline in population abundance. However, there is significant analytical challenge associated with isolating the causal factors that could explain observed changes in living resources and to predict the impact of the various factors that influence the composition and abundance of living resources (Hood et al., 2021; Rose et al., 2023). Extensive data exist to investigate various causal factors and aquatic organism abundance, and analytical approaches generally fall into one of three categories: habitat models, statistical models, and process-based models.[10] Examples of each category applied specifically to the Chesapeake Bay follow.

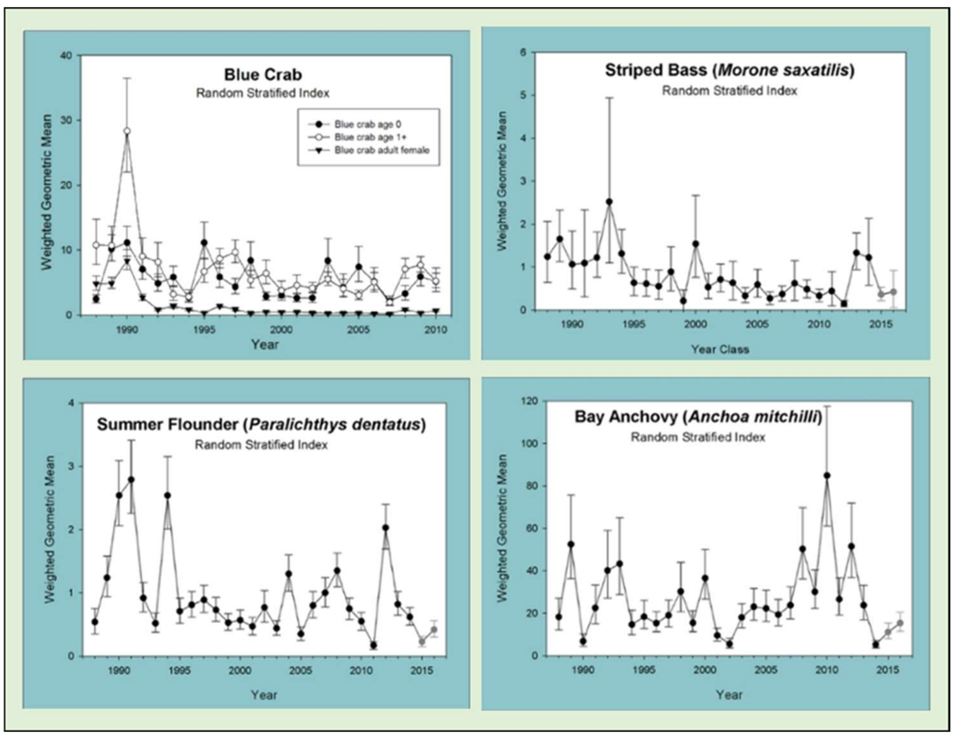


FIGURE 5.3.—Indices of Chesapeake Bay fish abundance, various species over time (Source: VIMS, n.d.-a).

[10] Habitat models may use statistical methods for fitting, but they predict habitat quality (typically in units of area or volume) as the response variable. We use the term "statistical models" to refer to models that predict response variables about biological quantities, such as organism abundances, growth rate, average weight per individual, or survival rate, from their correlations to explanatory variables. Process models are simulation (not statistical) approaches and predict biological quantities from assumed process rates.

Habitat-based assessments estimate the relative suitability of a habitat for a species based on the observed relationship between the species and environmental conditions. A variety of habitat-based models have been and can be used to relate living resource response to water quality conditions. For example, Schlenger et al. (2022) used a habitat volume model to estimate the sensitivity of 12 species in the Bay (at different life stages) to habitat changes resulting from changes in water temperature, salinity, and DO. The study differentiated between required habitat (the region outside of which mortality would occur due to salinity, temperature, and/or DO levels) and optimal habitat (the region outside of which salinity, temperature, and/or DO levels would cause physiological stress, leading to decreased growth, production, or impairments to other metabolic processes). The modeling effort provided quantitative understanding of the factors that limit habitat and their importance to individual species/life stages. Salinity and temperature were responsible for the largest changes in optimal habitat, while the results for DO were mixed across optimal vs. required habitat. Dissolved oxygen had a major impact on the required habitat for some species/life stages (e.g., Atlantic sturgeon) but had a relatively minor influence on optimal habitat across other species. The impact of DO improvement on habitat may vary with water temperature; a habitat model for Atlantic sturgeon found that achieving the Bay DO criteria would increase total habitat by 13% for an average weather year, but a 1 ^{o}C increase in water temperature would reduce habitat by 65% (Niklitschek & Secor, 2005).

Recent work illustrates how statistical analyses of monitoring data can be used to provide evidence linking living resources to water quality. Fabrizio et al. (2021) examined changes in habitat and the abundance of four common Bay forage species: bay anchovy, juvenile spot, juvenile weakfish, and juvenile spotted hake. Habitat conditions included distance to shore, sediment, water temperature, salinity, and DO. Habitat conditions such as water depth, water current, and temperature were identified as key habitat conditions for all four species, while DO delineated suitable habitats for juvenile spot in the winter and bay anchovy in the winter and spring. While the study highlighted the importance of shallow water, nearshore habitats for these forage fish, it found mixed and limited statistical evidence of a relationship between specific aspects of habitat conditions and observed seasonal abundance over 17 years (Fabrizio et al., 2021). Woodland et al. (2021) conducted statistical analyses of the relationships between abundance of classes of forage species (fish and benthic species) and water quality variables such as DO, water temperature, and salinity but did not include measures of nearshore habitat conditions. This study found that forage species abundance was consistently related to water temperature, particularly the rate of warming in the spring and summer (gradual warming is associated with increased abundance). More complex and varied responses were found to water quality conditions associated with WQS. Abundance of different classes of forage species were both positively and negatively correlated to Chl *a* levels. The abundance of most forage groups increased with DO levels in the mainstem section of the Bay, but forage group abundance showed a more varied response (both positive and negative relationships) to DO in the Bay tributaries (Woodland et al., 2021).

A variety of living resource process-based models have been developed to predict ecological response to different environmental conditions. Network analysis and Ecopath with Ecosim models have been developed or explored for the Chesapeake Bay (Baird & Ulanowicz, 1989; Christensen et al., 2009; Ma et al., 2010; Monaco & Ulanowicz, 1997; Townsend, 2014). Ecopath with Ecosim is the most-applied trophic ecosystem model in the world and is used in other locations to assess responses to restoration (Vasslides et al., 2017). These simulations allow quantitative connections to be made between water quality and commercially and recreationally important species, and they can be used to assess the benefits of and trade-offs between water quality restoration goals and fisheries management goals, with some restrictions (Townsend, 2014). Townsend (2014) describes the limitations and challenges of coupling Ecopath with Ecosim-based fishery models to existing Bay water quality models to assess the effect of water quality management on managed fish stocks.

Another process-based model that couples predator-prey relationships with biochemistry conditions is the Chesapeake Atlantis Model (Fulton et al., 2011). The Chesapeake Atlantis Model has been used to estimate the higher trophic level impacts of fully achieving the TMDL requirements for N and sediment under present day climate conditions, as well as under warmer water temperatures, and simulations were also combined with habitat loss and gain (restoration) scenarios (Ihde et al., 2016; Ihde & Townsend, 2017). Results suggest water temperature is a major factor in explaining abundance and that “achievement of the TMDL-prescribed reductions of N and sediment had only a moderate effect on modeled guilds” (Ihde & Townsend, 2017, p. 7). Fulford et al. (2010) used a network simulation model to examine the differential impacts of increased oyster production and decreased nutrient loads on phytoplankton, zooplankton, and species of forage fishes. The model suggested a 50% reduction in nutrient loads would increase benthic habitat (reduced hypoxia), decrease phytoplankton production, but have negligible predicted effects on bay anchovy and benthic fish biomass.

The degree to which specific aspects of water quality and habitat (e.g., DO, wetlands) are identified as the causes of detected changes in the living resources varies widely among these analyses. While many of these can be utilized in specific instances, collating results of independent analytical efforts is difficult because (1) the species, statistical and modeling methods, spatial coverage (e.g., regions of the Bay), and temporal coverage (e.g., which years) vary greatly across analyses; (2) existing analyses addressed study-specific questions and hypotheses but application of results to assess specific restoration actions have been mostly, at best, suggestive (i.e., speculative); and (3) in situ analyses specific to water quality and habitat rarely are specifically designed to advance the understanding of the role of the TMDL and other CBWA-related restoration actions in affecting the living resources response. To date, there has not been a comprehensive examination of living resources responses in situ that also attempts to relate the responses to CBP actions.

5.3. Conclusions and future directions

Research shows that the timing, location, and extent of changes in DO conditions influence both the species which are impacted and the magnitudes of the impacts. Direct evidence of the impact of water quality changes on various classes of living resources varies and is frequently unclear, partly because of the confounding multiple changes and their effects across complex ecological interactions, and partly because there have not been substantial system-wide changes in some criteria (e.g., DO).

Future conditions in the Bay will be different from historical conditions. The Bay is warming, and precipitation patterns appear to be changing (Hinson et al., 2022; Najjar et al, 2010; Fleming et al., 2020). Research consistently reports that water temperature (Batiuk et al., 2023) and salinity will have major influences on living resource communities, independent of water quality conditions such as DO. Thus, the living resource outcomes that can be expected from incremental achievement of WQS are conditional on these other factors. Research also highlights key roles that other factors, many which can be directly managed, have on explaining abundance, composition, and distribution of aquatic species. For example, research points to the importance of habitat (particularly shallow water) quality and nearshore conditions for living resources.

The CBP has devoted substantial resources to the development of analytical and statistical modeling of pollutant and water quality outcomes. By comparison, quantitative assessment of living resource responses to water quality and other factors has received less attention outside of reporting of annual indicators of population health. Ecosystem-level models have been developed for the Bay system but have not been used to their fullest potential to assess how different water quality management actions might impact living resources (Hood et al., 2021; Rose et al., 2023).

While numerous scientific studies have been conducted, particularly for individual species, they collectively offer limited help with analyzing how different classes of living resources and their interactions might respond to future and alternative water quality and nearshore conditions (Rose et al., 2023). Specifically, additional analytical and modeling capacity that focuses on living resource responses could shed light on questions such as:

- What is the expected (projected) response of living resources to the level of attainment of WQS and habitat conditions in the Bay: (a) without the TMDL and habitat and sustainable fisheries targets, (b) if present conditions of attainment and partial achievement of the habitat and fisheries goals continue, and (c) if water quality targets under the TMDL and/or the vital habitat and sustainable fisheries targets are fully achieved?
- What combination of conditions, in addition to and including achievement of existing WQS, would need to occur to achieve desired responses in living resources?
- What are the decision-relevant uncertainties of existing analyses, and how can they help guide future monitoring and modeling efforts? This information could be used to inform

natural resource managers about future research needs and their decisions about quantitatively defensible indicators and measures for tracking living resource responses.

Rose et al. (2023) described a modeling framework that could help address these questions for the Chesapeake Bay. This approach differs from the present approach by using predicted in situ responses of populations and food webs, rather than tolerances and preferences of individual organisms, to inform water quality and habitat targets. To date, restoration progress has been determined by achievement of the WQS (which were informed by living resource information) and status and trend indicators as part of the CBWA (CBP, n.d.-e). Without attribution to specific causative factors, documented changes in living resources (positive and negative trends) become difficult to interpret as reflective of restoration efforts or not. The approach described by Rose et al. (2023) could enable broader conclusions and more refined statements about the role of the TMDL and other restoration actions in the CBWA. The analyses would also provide information on realistic expectations to managers and the public about species responses to changes in water quality and habitat. Grounding expectations is especially important when the responding species (like many in the Chesapeake Bay) are long-lived, have complex life cycles, and are affected by multiple factors. Further analyses would also be useful as a guide for interpreting living resource responses to the TMDL and habitat actions, designing future indicators, informing adaptive management, and possibly refining restoration goals.

6. Findings and Implications

Progress toward achieving Bay pollutant reductions and WQS has been slower and more challenging than expected. The question is: Why? Policy and management actions are intended to reduce nutrient and sediment pollution to improve Bay water quality conditions and ultimately enhance Bay living resources. Key objectives of this report were (1) to identify gaps and uncertainties in the causal chain of responses that link policy and management actions to these goals, and (2) to offer approaches to address these gaps and uncertainties going forward.

This chapter briefly summarizes the findings from chapters 2–5 and offers options for how progress can be accelerated. Efforts have fallen short of achieving TMDL nutrient targets. In many cases, implementation and response gaps confound efforts to reduce nonpoint sources. Pollutant load reductions , if successful, still take time to work through soils, aquifers, and stream channels of a large watershed, leading to time lags between implementation and observed reductions in loads reaching tidewater. If and when loads reaching the Bay are reduced, the Bay water quality response indicated by standard metrics (DO, water clarity/SAV, and Chl *a*) is mixed. For example, nutrients have shown signs of improvement, but water quality responses have been less than expected. The fact that water quality conditions have not deteriorated given significant economic and population growth, land use change, and a changing climate in the past three decades is a notable accomplishment.

The remainder of this chapter offers suggestions on what can be done to improve system response. What options and opportunities exist for water quality management efforts to get more nutrient reductions from implementation actions (particularly nonpoint sources), and what actions might improve the response of living resources to water quality improvement efforts? STAC does not offer prescriptive solutions, but rather describes choices facing CBP partners and identifies promising options for improving system response to management efforts. STAC offers multiple options to improve the effectiveness of nonpoint source programs, prioritize water quality and TMDL management efforts that have the greatest potential to boost fish and shellfish habitat and populations, and enhance decision-making about these options when outcomes are uncertain.

The findings and implications are offered based on three themes:

First, achieving load reductions and water quality improvements is proving more challenging than expected. To date, efforts to reduce nonpoint sources have not produced sufficient levels of BMP implementation to meet the TMDL, and the implementation that has occurred may not be producing the pollutant reductions expected. Evidence indicates that the nutrient and sediment load reductions realized to date have led to improved water quality conditions in some areas of the Bay, especially in areas with substantial localized reductions, but other areas have mixed results.

Second, the Chesapeake Bay system observed in the past will not be the same system in the future. Future conditions will be influenced by permanent and ongoing changes in land use, climate change, population growth, and economic development. Recognizing these changing conditions will challenge notions of restoration that are based on recreating historical conditions.

Third, new approaches to water quality management will be required to adequately address uncertainty, changing future conditions, and response gaps. Opportunities exist to improve pollutant management, water quality, and living resource outcomes, but many require changes and new approaches to implementation, planning, and decision-making. Given uncertainty, change requires a willingness to accept the risk that some efforts will fail and a desire to enhance capacities to learn from successes and failures. Nothing in this report should be interpreted as suggesting "backsliding" or retreating from commitments to improve water quality in the estuary. However, water quality policy will need to evolve over time based on an understanding of what future conditions are possible, what local communities and the partnership at-large see as priorities, and what is required to attain those possible futures.

6.1. Summary of findings

Four major findings emerge from chapters 2–5 and are summarized as follows.

The slow rate of water quality change in the estuary suggests that achievement of WQS in the Bay is uncertain and remains in the distant future. To date, attainment of the Bay WQS stands at about 30%. Evidence indicates that load reductions achieved to date have led to improved conditions in some portions of the estuary, but overall results are mixed. Evidence also indicates that the nutrient reductions achieved to date have not produced the magnitude of DO response expected. For example, deep channel and deep water habitats of the Bay exhibit significant potential response gaps for DO, while SAV coverage in shallow and open water habitats falls short of stated goals. Multiple possible reasons exist for limited progress in achieving the WQS, including climate change and insufficient reductions to stimulate an accelerated response (tipping point).

Existing efforts to reduce nonpoint sources of nutrients are likely insufficient to achieve the TMDL. The TMDL and its nutrient and sediment reduction targets are the primary policy drivers to achieve Bay WQS. Progress has been made in reducing nutrient loads, particularly in relation to point source discharges and atmospheric N deposition. However, the CBP jurisdictions have identified in their WIPs that additional load reductions will be focused on nonpoint sources in agricultural and urban areas. These pollutant reduction efforts must induce sufficient and effective behavioral change from hundreds of thousands of people and land managers within the watershed whose behaviors contribute to nonpoint source loads from agriculture (the largest nonpoint source) and developed land (the fastest-growing nonpoint source). Making substantial progress toward meeting the TMDL will require an increased rate of agricultural and urban nonpoint source load reductions relative to what has been achieved in the past.

The pathway to achieving TMDL nonpoint reduction targets is unclear, but options are available. The CBP acknowledges the challenges of generating enough nonpoint source BMP adoption to meet nutrient reduction goals, particularly for N. This implementation gap, the difference between control practices implemented and practices needed, is widely acknowledged for N. In addition, statistical analyses of ambient monitoring data indicate that nonpoint source management actions (BMPs) are not generating the amount of reductions expected (response gap) for N and P. The P response gap is especially large and has many potential causes, most of which are associated with regions of the watershed with significant mass P imbalances (which, in turn, are associated with intensive livestock and human populations).

Additional funding alone is unlikely to achieve desired nutrient and sediment reductions. Achieving and sustaining substantial future pollutant reductions will require a willingness to develop and adopt new implementation approaches and technologies. The basic structure of the CBP's implementation programs has been in place for several decades: numeric effluent limitations for permitted sources, voluntary practice-based cost-share programs for agriculture and unregulated urban sources, and a TMDL accounting system that tracks success through a central modeling system. This report describes the accumulated evidence of implementation and response gaps that limit our ability to secure reductions in nonpoint source loads within the current structure. Making substantial progress in reducing nonpoint source discharges will require changes in program structures, incentives, and requirements.

Improving water quality to meet the Bay WQS may not be sufficient to generate desired changes in the composition and abundance of Bay living resources. Water quality criteria were chosen to support and enhance Bay aquatic living resources. The question of what type of living resource response can be expected from water quality improvements is complex and uncertain. Compared to water quality assessment, the CBP has devoted much less effort and resources to assessing the impacts of water quality on living resources in different habitats (open water, shallow water, deep water, etc.). However, a major advantage with the Chesapeake Bay is that much of the data and information (e.g., physical and biological information) needed to expand this capacity are available, well-established, and vetted. Expectations for how living resources will respond to water quality changes should be conditioned upon the range of additional factors that also impact their composition and abundance, including climate change, land use and economic change, harvest, human population growth, and natural variations outside the influence of management. For example, not only does Bay warming make attainment of DO goals more difficult, warming water temperatures will affect habitat and aquatic species mix and ecosystem interactions. Changes in nearshore habitat and land use will also affect fish composition and abundance in a myriad of ways beyond just altering nutrient and sediment loads.

Current CBP adaptive decision-making processes have limited capacity to effectively address the full range of questions required for effective water quality policy. Effective water quality policy answers the following questions: (1) what is; (2) what if; and (3) what should be. The CBP has built a sophisticated adaptive decision-making process focused on TMDL implementation, the SRS, and an accounting process premised on the use of predictive, deterministic planning models. The current SRS process, TMDL accountability framework, monitoring program, and modeling are aimed toward answering questions 1 and 2. The existing analytical models and planning processes have limited capacities to systematically address the critical uncertainties and causes of the response gaps for these questions. Current adaptive management processes have limited decision-making scope to address question 3.

6.2. Implications: Policy options for improving effectiveness of water quality management
The approaching 2025 federal TMDL deadline offers an opportunity to reassess and enhance the potential of policy to improve water quality and living resource response to pollution control efforts in the face of uncertainty. The remainder of this section provides three overarching areas that could be considered to improve water quality management effectiveness.

Refocusing water quality management efforts on improving living resource response. Current WQS focus on achieving water quality conditions (DO and water clarity) that broadly *support* living resources. This focus does not necessarily mean that progress toward achieving WQS will directly translate into enhancing living resource populations and responses. Options are discussed to address the question: Can Bay water quality goals and the way attainment is measured be revised to increase attention to, and potential for, other water quality investments to improve living resources?

Improving effectiveness of nonpoint source management. Options are discussed to address the question: What policy and implementation options offer potential to deliver substantial and sustained reductions in nonpoint source loads?

Enhancing adaptive management to improve the CBP's ability to learn and respond to uncertainties and response gaps. Given the knowledge, uncertainties, and complexities described throughout this report associated with assessing and responding to response gaps, options are discussed that address the question: What processes and analytical approaches are available to improve learning, especially as it pertains to the first two areas?

The practical necessity of TMDL implementation confines management efforts to making decisions and choices within established accounting rules, operational programs, and well-defined programmatic goals (in figure 6.1, see feedback arrow 5). The 2025 TMDL deadline, however, is an opportunity to consider a broader set of options to improve water quality policy (figure 6.1, feedback arrows 1 through 4) This chapter considers options for designing and implementing new pollutant control approaches and TMDL policies (figure 6.1, feedback arrow 4). Changes may also include policy options for how water quality goals could be amended (e.g., changing criteria, assessments, etc.) to increase the potential for enhancing abundance, diversity, and resilience of desired aquatic species (arrows 3, 2, and 1 in figure 6.1). Discussions of options and alternatives emphasize the need to consider options beyond existing programs and goals, but ones that are still possible within the legal and programmatic constraints of the CWA.

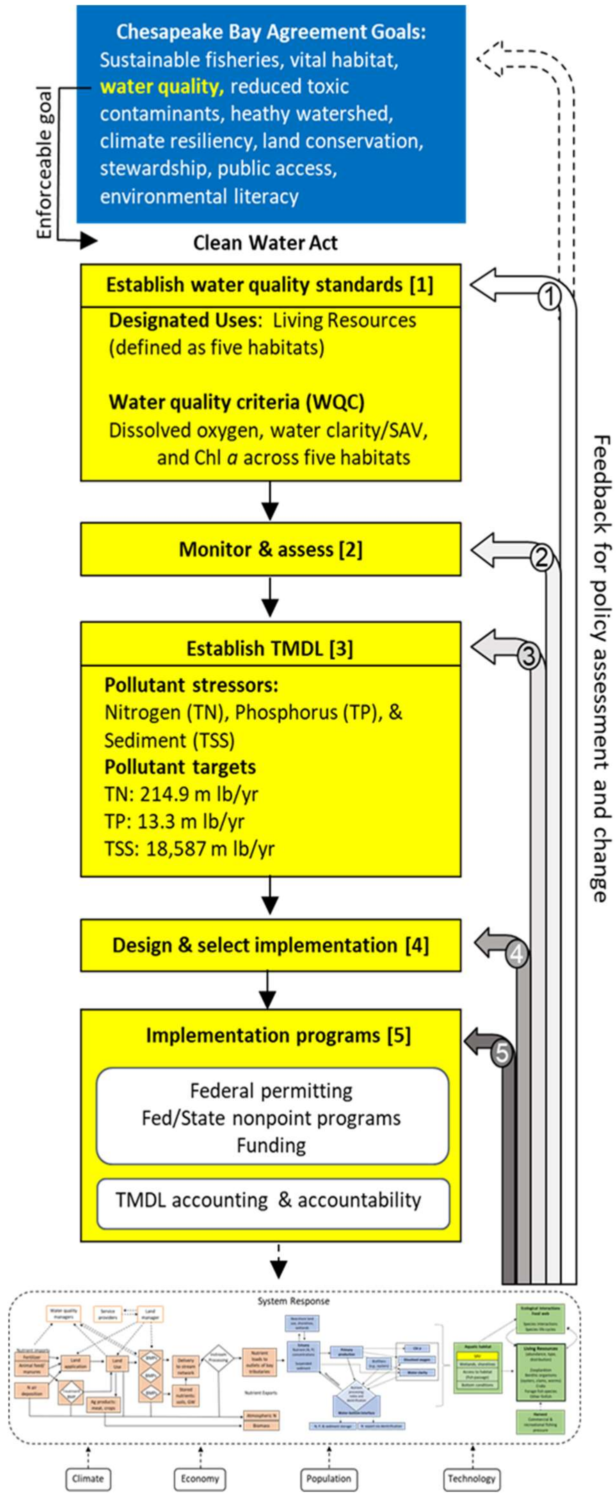


FIGURE 6.1.—Levels of policy feedback and learning in expansion of adaptive management.

6.3. Water quality goals and living resource response

The objective of the WQS and associated TMDL is to improve Bay living resources. Water quality standards are largely expressed in terms of achieving key water properties (DO and water clarity) to support living resources, but not directly in terms of enhancing species populations. The approach of the WQS means management attention is directed to reducing pollutants to achieve the measurable criteria (e.g., DO). The policy challenge that needs to be considered is this: What can be done to improve living resources from our water quality management efforts given limited public resources and budget constraints?

Figure 6.2 shows the conceptual cost and benefit tradeoffs associated with different levels of attainment of Bay WQS. Panel A of figure 6.2 shows a conceptual representation of costs to attain the 1,052 combinations of WQC (DO, water clarity, and Chl *a*), estuary segments (92 spatial segments), and habitats (up to 5 different habitats per segment). The cost to bring a greater portion of the Bay into attainment with the WQS is expected to increase at an increasing rate for a couple of reasons. First, more pollutant reductions will be necessary to bring additional segments and habitats of the Bay into attainment. For example, an additional 20 million lb/yr of N reduction may bring many open water and shallow water habitats into compliance, but another 20 million lb/yr reduction may bring only a few additional, generally deep water, habitats into attainment (attainment of DO criteria in deep water segments/habitats is expected to take the longest).

Second, the pollutant control cost to achieve each additional unit of pollutant reduction ($/lb) is increasing as low-cost control options are used first (Kaufman et al., 2021). Kaufman et al. (2014) estimated that the costs to reduce agricultural nonpoint source loads in Pennsylvania would increase sharply (exceeding $100/lb/yr for N) when approaching TMDL targets. Urban nonpoint source controls are often estimated to be an order of magnitude more expensive ($100 to over $1,000/lb/yr for N) than other nonpoint source controls and will still be needed to meet TMDL targets (Price et al., 2021). Maryland estimates a cost of $1.195 billion for municipal stormwater systems to achieve the required additional reduction of 85,000 lb/yr of N, or approximately $1,100/lb/yr (MDE, 2019).[11] As a relative comparison, past point source reductions at municipal WWTPs have been achieved for less than $40/lb/yr (Stephenson & Shabman, 2017b).

Figure 6.2 (panel A) also shows the implication of both uncertainties and gaps described in chapters 3 and 4. The costs to obtain any given level of WQS attainment are not precisely known given uncertainties in producing pollutant reductions and attaining WQS (dotted blue curves). The existence of estuary and nutrient management response gaps means that more reductions will be needed to achieve any given WQS. Given changes in the watershed, climate change, and the presence of implementation and response gaps identified in this report, full achievement of the existing WQS may be technically infeasible

[11] Annualizing $1.19 billion over 20 years at 5%.

(particularly for the deep water and deep channel habitats), regardless of the amount of management effort and resources spent (i.e., the blue cost curve may not cross 100% attainment). Further, pursuing an unattainable water quality goal can divert resources from other actions that could be more impactful in terms of the responses of living resources.

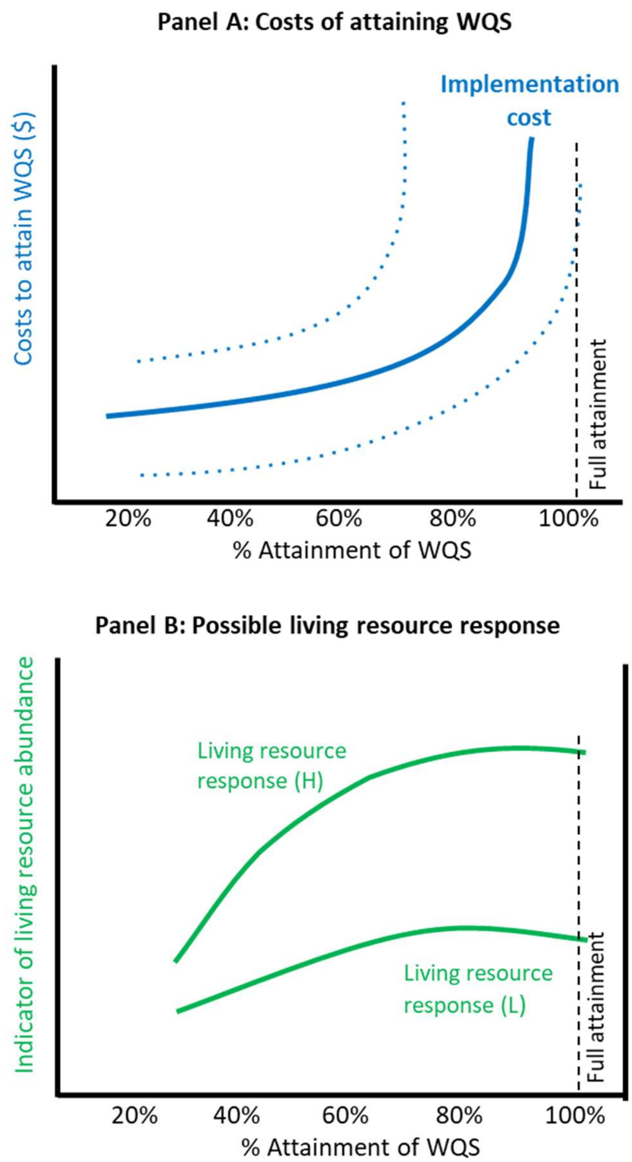


FIGURE 6.2.—Conceptual relationship between costs, attainment of WQS, and living resource response.

As implementation costs increase, so does public interest in what improvements will be achieved with this expenditure of resources. The public benefits associated with achieving WQS arise largely from the potential improvements in aquatic living resources (designated uses) (see panel B in figure 6.2).

As described in chapter 5, living resources are expected to respond favorably to improved and expanded habitat conditions (e.g., more usable areas with adequate DO), but the amount of improvement is uncertain. The location and timing of the water quality improvements will affect living resource response. For example, improvements in DO in shallow water habitats that support both nursery habitat and forage fish may generate larger living resource responses than similar levels of water quality improvement in deeper water habitats. Currently, the TMDL management focuses on achieving a single baywide reduction goal for N, P, and sediment. The final Bay TMDL load targets were established to bring all segments and habitats into compliance, and these targets were driven by the deep water DO criteria, the most difficult to attain. Thus, costs are expected to increase sharply and living resource improvements diminish as nonattainment area shrinks to the deeper water habitats.

The living resource response to improvements in DO and water clarity depends on many other factors that also impact the abundance and composition of living resources. Whether living resource response to water quality improvement will follow the H or L response curve in figure 6.2, panel B, depends on a variety of factors such as water temperature, salinity, structural habitat, disease, and harvest. The response of living resources to DO improvements may be modest (response L) if these other factors are already limiting carrying capacity for fish, shellfish, and crustacean populations. Achieving a greater living resource response to water quality improvements will require appropriate types and levels of investments in these additional factors, some of which are represented as goals and outcomes in the CBWA (e.g., sustaining fisheries and recovering habitats). Supporting and increasing progress towards these other CBWA goals could be constrained if management and legal attention becomes too rigidly and narrowly focused on the TMDL.

Finally, some of these factors that will significantly influence the composition and abundance of living resources are beyond CBP management approaches and policies. Climate change will increase Bay water temperatures and alter historical salinity patterns and habitat (e.g., tidal wetlands affected by sea level rise; replacement of historic SAV species with new ones). These factors should be considered and evaluated in any future discussions surrounding the TMDL and water quality standards.

6.4. Options for improving living resource response

This section examines options for improving potential living resource response to water quality management efforts. Several options are presented, beginning with options that would require the least amount of change in practice implementation and associated policies. First, changes to the way the TMDL is implemented may enhance living resource benefits without requiring any changes to the WQS (feedback arrow 4 in figure 6.1). Second, modification of and refinements to the existing WQS (feedback arrow 1 in figure 6.1) may also be pursued.

Consider a tiered approach to structuring the TMDL and achieving WQS. Currently, the overriding approach driving water quality management is planning and implementing practices to reach the pollution targets for N, P, and sediment, which would meet WQS in all habitats, in all 92 segments. The current focus is on achieving criteria in the deep water habitats, with the assumption that attainment there will be the most difficult to achieve. The CBP could consider where pollutant reductions should occur first to accelerate the potential for living resource improvements.

Tiered attainment of the TMDL and WQS based on critical habitats (such as shallow and open waters) or locations could provide early and more substantial living resource benefits and still contribute to baywide water quality improvements. Chapter 5 established living resource response likely varies by the location and timing of water quality improvement, so response could potentially be accelerated by prioritizing water quality improvement efforts in critical habitats. The implication is that with a focus on shallow and open water habitats full attainment of the WQS (including the deep water habitats) in the near future may not be necessary to achieve significant potential gains in living resource response.

There are several approaches to prioritizing TMDL implementation and water quality attainment. The following list is illustrative, but not comprehensive. Achievement of TMDL targets could be prioritized according to location (segments) or habitat type for most living resources. Establishing different deadlines and staggered TMDL pollutant reduction targets may focus implementation efforts in areas projected to have the largest potential for improving living resource conditions. Granting temporary or permanent variances for segments or habitats where achievement is technically infeasible or exceptionally costly could allow resources to be shifted to areas where attainment is possible and with a higher living resource impact.[12]

It should be stressed, however, that the CBP has devoted limited analytical effort to inform such policy discussions about ecosystem living resource responses to TMDL pollutant management (Hood et al., 2021). Additional investment in analytical capabilities and models could make important contributions to a tiered or prioritized approach to the TMDL (for more details, see the discussion of analytical tools in section 6.6 and Rose et al. [2023]).

Consider revisions to existing WQC. Potential revisions to the existing WQC (DO, Chl *a*, water clarity) could strengthen the link between the criteria and living resources. Reevaluation of WQC may also include consideration of new criteria or new frameworks for devising criteria. Criteria with potential impacts on living resources could include localized water temperature anomalies, fine sediment versus total suspended sediment loads, or impacts from chemicals of emerging concern for living resources. For example, increasing water temperatures will likely alter ecosystem structure and function, such as a shift in aquatic ecosystems from diatom-dominated to green algae- or cyanobacteria-dominated (Batiuk et al., 2023). Excessive fine-grained sediment is responsible for degrading habitat in both the watershed and estuary,

[12] EPA has granted a number of water quality variances for specific habitat segments. See chapter 2 for a brief discussion.

including underwater grasses (Kemp et al., 2004, Lefcheck et al., 2017). The potential effects of toxic contaminants and contaminants of emerging concern on the abundance and composition of living resources may be significant, and yet information on sources and occurrence has been limited.

New frameworks for devising criteria could emphasize system resilience, particularly for the biological criteria such as SAV. Much of the management of Chesapeake Bay, and estuarine systems around the world, has traditionally focused on ecosystem conditions. But in the face of changes due to human population growth and land use change, combined with the impacts of climate change, this focus will necessitate a shift to managing for resilience. Resilience is defined as the ability to recover from the disturbances that are increasing in frequency and magnitude, and the concept is applicable to both human and natural systems. Criteria that reflect resilience characteristics can be developed and would provide a more appropriate target than historical condition. The SAV criteria were established based on historical estimates of SAV acreage (which may no longer reflect SAV potential in a warmer Bay). Other SAV criteria, besides total area, could consider the recovery of SAV in the wake of episodic events or the speed of recovery. Ultimately, consideration of other designated uses such as “fishable and swimmable” or recreational use could imply the development of new WQC.

6.5. Improving nonpoint source management

Improving the approaches to reduce nonpoint sources of nutrients is the central TMDL challenge for achieving WQS under the current TMDL framework. Nonpoint sources are diverse, involving tens of thousands of people making land use and nutrient use decisions across agricultural and urban landscapes. Tracking nutrient changes across the landscape is often challenging, and isolating cause and effect is difficult. Inducing behavioral and technical changes of sufficient type and scale to achieve nonpoint pollutant reduction goals is a fundamental challenge faced by large-scale eutrophication reduction efforts worldwide (e.g., Great Lakes, Gulf of Mexico, Baltic Sea, etc.).

An overriding theme of possible policy reforms described below is to shift incentives and behavior away from counting practices toward crediting pollutant reduction and water quality outcomes. The current CBP partner approach for nonpoint source reduction is built on a practice-based, cost-share incentive structure and the CBP modeling and accounting system to credit nonpoint source load reductions. While providing accountability, current voluntary incentive systems focus attention on installing practices to get credit within the CBP TMDL modeling framework. This system enables EPA to count jurisdictions’ contributions towards TMDL compliance, but it separates reporting implementation of practices from assessment of pollutant reductions. Chapter 3 described multiple examples of the potential to improve nonpoint source reductions through a better understanding of the sources and locations of high nutrient loss areas, directing management attention and assistance to who does what and where to reduce those pollutant loads, and rewarding pollutant reduction outcomes. A focus on practice installation can limit the willingness and ability of water quality managers to capitalize on these opportunities.

Policy reform can take many different forms (Alberini & Segerson, 2002; Ribaudo & Shortle, 2019; Shortle et al., 2021). Given the diversity of people, production systems, and land use decisions involved in nonpoint source pollution generation, policy change and reform will not be a simple answer or single approach. Policies to improve nutrient reduction effectiveness for agricultural regions with large nutrient mass imbalances will look different from policies targeting fertilizer applications by urban homeowners. Opportunities for incremental adjustments to nonpoint management programs exist, including better model input data (e.g., animal numbers, fertilizer applications, etc.), increasing cost-share amounts, and increasing the number of service providers to encourage greater BMP implementation (Chesapeake Bay Commission, 2017; Collins et al., 2022). However, making substantial progress in nonpoint source management will require programmatic and policy change.

The following are illustrative, but not exhaustive, policy and programmatic options that offer promise for reducing nutrients from nonpoint sources.

Address mass imbalances. More effective and systematic approaches to addressing nutrient mass balance issues offer opportunities for substantial, sustained reductions in nonpoint source nutrient loads.[13] Only limited progress can be made if mass imbalances are not adequately addressed. Most BMPs do not substantially alter mass balances. Many traditional BMPs (e.g., cover crops, no-till, many stormwater practices) that do not substantially reduce nutrient inputs have limited long-term capacities to alter nutrient losses in areas with mass imbalances. Evidence suggests that policies designed to alter regional mass balances have proven particularly effective in improving water quality. The P detergent ban and wastewater treatment technology investments to increase denitrification are examples from point source management.

The most significant regional mass imbalances are associated with intensive animal agriculture and row crop production (Beegle, 2013; Kleinman et al., 2012; Spiegal et al., 2020). Concentrated livestock production and the subsequent land application of manure is a contributor to the pronounced response gap surrounding the effectiveness of P reduction efforts. Evidence suggests intensive livestock operations are closely linked to increasing P (particularly dissolved P) . Addressing regional mass imbalances related to livestock manure involves improving program implementation and technologies by reducing nutrient inputs, improving manure distribution (i.e., from surplus to deficit areas), and exporting manure from the watershed (Flynn et al., 2023; Saha et al., 2022). Specific efforts that could be strengthened include reducing the nutrient content of livestock feed, increasing feed efficiency, improving manure transport programs, siting of livestock facilities, using P-based nutrient management,

[13] A mass balance approach describes nutrient inputs (e.g., fertilizer and feed) to and outputs (e.g., grain or meat export, loss to water bodies) from the system, reactions or transformations (e.g., denitrification), and storages (e.g., buildup of P and N in soil and groundwater) in the system.

expanding manure treatment and conversion technologies, and enhancing advanced nutrient management incentives.

Improve incentives and capacity for identifying who does what when and where. Improving the incentives and capacities to better identify where nutrient-reducing actions are needed could improve program effectiveness (reducing both implementation and response gaps). Chapter 3 noted that the distribution of nutrient losses is uneven and can vary significantly across the landscape and across land managers. Existing nonpoint source programs and CBP accounting systems generally do not provide incentives or the capacity to adequately address and capitalize on this diversity. Research consistently shows that improving targeting of nonpoint source control investments can generate more, and lower cost, reductions (Choi et al., 2020; Giri et al., 2012; Easton et al., 2019; Kast et al., 2021; Khanna et al., 2003; Lintern et al., 2020; Xu et al., 2019; Yang & Weersink, 2005). Improved technical tools and new incentive programs could better direct efforts to identify low-cost, high-impact opportunities for nutrient and sediment reduction (Easton et al., 2019; Fleming et al., 2022).

New outcome-based incentive programs could be designed to place more emphasis on achievement of pollutant reduction outcomes (Collins et al., 2022; Easton et al., 2019; Fleming et al., 2022). Payment for environmental service programs (e.g., pay for performance, pay for success, rewards for success) compensate land managers for quantifiable pollutant reductions or for attainment of observable benchmarks directly or indirectly linked to pollutant reduction outcomes (e.g., soil nutrient levels, ambient water quality conditions, etc.). Such systems reward land managers for pollutant reduction achieved rather than paying a portion of costs to install a BMP. These alternative financial incentives could encourage agricultural producers and service providers to search out high-impact, low-cost options for nutrient and sediment reductions. For example, several nutrient removal technologies have high up-front costs but produce potentially large and low cost-per-unit nutrient reductions (e.g., manure treatment, bioreactors, riparian buffers, etc.). Pay-for-performance programs could fully compensate land managers for implementation.[14]

Pay-for-performance programs depend critically on how performance is defined (Fleming et al., 2022). A variety of options exist. Using CAST predictions is one way to define nutrient and sediment reduction performance, but chapter 3 described a number of limits to CAST's ability to identify fine-scale, high-loss areas and land managers and incentivize treatment. Other measures of performance should be explored and developed, including finer-scale models to identify critical nutrient and sediment sources. Payments or reward programs could be based on quantifiable changes in indicators or proxies of nutrient and sediment reduction performance (e.g., reductions in soil P levels). Indeed, some pilot programs have established reward payments to groups of landowners for achieving improvements in ambient water quality outcomes (Maille et al., 2009).

---

[14] Performance incentive systems are based on the premise that land managers can profit from water quality actions (financial returns above costs). This represents a fundamental departure from traditional cost-share programs.

Differential crediting of nonpoint source investments would facilitate the identification and treatment of high-loss areas and agricultural operations (Easton et al., 2019). The CBP accounting framework generally counts and credits loads based on averages (e.g., average loading rates, average BMP effectiveness, average nutrient application rates). Differential crediting could be accomplished in several ways, including using a finer-scale targeting or modeling system or granting flexibility to state and local partners to develop, test, evaluate, and quantify the performance of BMP alternatives that target high-loss areas or operations. For example, new CBP stream restoration crediting protocols support greater TMDL credit for projects located where streambank erosion is greater. This approach will necessitate finer-scale modeling and measurement capabilities that can identify areas of the landscape contributing disproportionately high nonpoint source nutrient and sediment loads.

More efforts are needed to close the implementation gap. New incentive systems can help improve program effectiveness, but the extensive history of nonpoint source policy illustrates that limits to voluntary adoption exist (Liu et al., 2020; Ribaudo, 2015; Ribaudo & Shortle, 2019; Shortle et al., 2012). Achieving larger nonpoint source reductions may necessitate additional requirements on nonpoint source management in selective situations (Kling, 2013). Such requirements can be accompanied with financial assistance and do not necessarily have to be broad, inflexible, or exceedingly costly. Performance-based requirements, for example, establish clear outcomes to be achieved but grant flexibility in how to achieve those outcomes. In vertically integrated production systems, such as poultry and swine, responsibility for manure management that will address mass imbalances in a watershed could be shared between integrators and their contract producers rather than, as is now the case, borne just by the producers. This shared responsibility would incentivize integrators to work with producers in seeking and then implementing innovative manure treatment, transport, and use strategies for the relevant watersheds.

Encourage institutional innovation through sandboxing. Improving implementation effectiveness will require institutional innovation in addition to innovation in pollution control methods, modeling, and monitoring. Institutional innovations can range from incremental improvements within existing programs to consideration of new programs, rules, and accountability systems. This report's suggestions for making substantial progress in reducing nonpoint source loads will require exploring more than incremental adjustments, but major changes in implementation approaches cannot and should not be made based only on conceptual arguments. Each of the policy options described above, while offering promise, also faces numerous design questions and implementation challenges (Howard, 2020; Palm-Forster et al., 2016; Shortle et al., 2021).

Sandboxing is a formalized process that begins with conceptual development of new rules and programmatic approaches to nonpoint source load reduction or water quality improvement, followed by trial and evaluation. Sandboxing allows refinement of institutional reform details before any programmatic changes are made and can proceed without disrupting the operation of existing programs and operations (Higgins & Male, 2019; O'Sullivan, 2021).

The regulatory agency or programmatic organization would work with the sandbox proposer to operate in an isolated environment of sufficient size and autonomy to implement and test the efficacy of a potential change. A sandbox would require a commitment of time and resources by the CBP and its partners to effectively operate the trial (sandbox). The authorizing agency or organization would grant permission to form the sandbox under the condition that the proposer provides both a well-developed plan and a way to evaluate outcomes from the new rule or programmatic change being considered. To promote innovation, the authorizing agency or organization would also make commitments to pursue reforms if the sandbox trial produces demonstrative improvements in the agreed-upon outcomes. Two examples of possible sandboxing applications are described in text box 6.1.

Supporting innovation and change recognizes that some proposed innovations may sound attractive in concept but will prove to be ineffective or too difficult to implement in practice. The possibility of failure should not discourage pursuit of such opportunities, with the caveat that learning occurs as a result (i.e., the reason for failure is understood). The most effective learning often starts with failure. The potential reward, however, is the discovery of otherwise unrealized improvements in water quality outcomes.

6.6. Expanding adaptive decision-making and improving program learning

As described in chapter 2, the CBP has implemented a decision framework intended to continuously monitor and evaluate progress toward achievement of specific CBP program goals of the CBWA. The SRS process and the TMDL accountability framework adjusts implementation based on defining implementation goals, describing factors influencing goal attainment, and assessing management gaps. Adaptive management processes have occurred primarily within the implementation phase (feedback arrow 5 in figure 6.1) and have been justifiably focused on assessing whether planned and implemented management actions are being undertaken to achieve the TMDL, including implementing and tracking practices, refining modeling and accounting tools, and assessing staff and budgetary requirements to implement practices (CBP, 2020).

In order to respond effectively to the issues raised in this report, the partnership will need to expand the current sphere and capacity of adaptive management processes for the TMDL and attainment of WQS to include the scope and scale of program changes described in this report. This includes consideration of changes across the entire range of water quality policy management (figure 6.1, arrows 4, 3, 2, and 1), from understanding how living resources may respond to water quality management to addressing uncertainties in water quality management, particularly with respect to nonpoint sources. The question is not simply: Are the jurisdictions implementing the planned actions? Rather, the questions must be asked: Are the actions producing the anticipated pollution reductions and attainment of WQS? If so, why? If not, why not? This shift in thinking will require the CBP to reconsider how decision makers are to be involved and how to apply a more expansive adaptive management process. Some options to evolve the adaptive management processes are described in the rest of this section.

Text box 6.1. Illustrations of possible sandboxing applications for the Chesapeake Bay Program

Illustration 1. Nonpoint source accounting and compliance sandbox

Challenge: The Bay TMDL operates under a single accounting framework for counting and tracking BMPs and assigning nonpoint source reductions to those practices. As outlined in chapter 3, the modeling system averages nonpoint source nutrient loads over fairly large geographic areas and makes general assumptions about land managers' nutrient use behaviors. State and local water quality managers receive credit for installing BMPs that are assigned pollution reduction credits by the CAST model. Such a system may limit the ability and incentives for state and local water quality managers to search for and treat high-loss areas or operations which are not reflected in the model.

Sandbox idea: Design of an appropriately scaled watershed sandbox could allow Bay jurisdictions the flexibility to deviate from TMDL accounting rules to test alternative nonpoint source accounting and compliance systems. For example, a local jurisdiction could be exempt for a specified period of time from the TMDL modeling framework if the jurisdiction can produce improvements in the way to quantify nutrient reductions through finer-scaled models and more intensive monitoring of localized outcomes (Easton et al., 2019). Such an alternative compliance strategy could reward local actions demonstrating improvements in local water quality rather than simply installing practices.

Illustration 2. Innovative financial incentives for achieving reductions

Challenge: Existing nonpoint source cost-share programs do not award financial incentives based on the magnitude of the water quality benefit. Under existing programs, land managers have few financial incentives to install nutrient removal technologies that can generate large pollutant reductions but produce little benefit to the land owner. Cost-share program rules may also fail to provide sufficient incentives to land managers who have limited financial resources to share the costs of installing effective BMPs to address substantial pollution problems.

Sandbox idea: Grant a local conservation district flexibility and a dedicated budget to develop and administer alternative financial incentive payments to secure cost-effective nonpoint source reductions. Supplemental incentive payments could be awarded based on the size of the pollutant reductions that could be achieved on a particular operation. Reward payments could be offered for achieving observable farm-level milestones related to nutrient reduction success. Many different payment options exist, but questions exist about which incentives are most effective at stimulating behavioral change. A sandbox would create an opportunity to test the intended and unintended consequences of different alternatives without committing the entire program to a change.

Expand participation in adaptive management. An expanded adaptive management process will involve increasing the participation of decision makers across all levels of the program. The jurisdictions and EPA are responsible for implementing these activities for the TMDL. However, the policy and management opportunities for improving approaches to meet the TMDL that are outlined in this report involve greater resource and management commitments. The consideration and implementation of changes to WIPs, funding support, and water quality goals will require expanded involvement of those participants who have authority for authorizing and funding programs. Within the Chesapeake Bay organizational hierarchy, this would include members of the Chesapeake Executive Council and the Principals' Staff Committee, which include leaders of federal and state programs.

Use decision science for enhanced adaptive management. Enhancing adaptive management requires addressing the challenge of how to bring technical, programmatic, and scientific knowledge about the issues raised in this report to the people with the authority to make the choices regarding water quality goals, funding, and regulatory and administrative changes. Decision science can be used to effectively present and consider policy alternatives and options which produce complex and uncertain outcomes. Decision scientists have developed processes under a variety of labels (e.g., structured decision making, shared vision planning, collaborative modeling, etc.) that effectively combine planning processes and technical analyses (Bourger, 2011; Gregory et al., 2012; Runge et al., 2020). This includes the analytical tools and processes needed to identify decision-relevant uncertainties and the most-probably-effective actions given the uncertainties. Because the resources for pollution control implementation are limited, strategy development also requires prioritization of uncertain actions with the objective of maximizing potential outcomes at lowest possible cost.

Expand analytical tools to support decision-making under uncertainty. The CBP has limited capacity to systematically evaluate uncertainties and response gaps described in this report. Deterministic models providing single estimates of pollutant loads for all inputs, land uses, and management actions are not well suited for evaluating and addressing uncertainty. Such modeling approaches make it difficult to assess the performance risk of different BMPs, inform decision makers of uncertainties, or assess the robustness of management actions to underlying assumptions or changing environmental conditions.

An expanded adaptive management process will need to better incorporate tools and processes to identify and reduce decision-relevant uncertainties (Marchau et al., 2019). For example, the expected value of information (Runge et al., 2011) can be used to identify which uncertainties in nonpoint source management are most important to resolve in order to improve water quality. Such approaches aim to identify those uncertainties that pose the greatest risk of not achieving management objectives, identify how much a given outcome could be improved if a given uncertainty were resolved, and identify the cost of error. Robust decision-making and modeling tools seek to identify solutions and management strategies that perform well under many possible assumptions, rather than optimal solutions that minimize risk given a stringent set of assumptions. Modeling of watershed processes that involves more explicit characterization of uncertainties in system process and model parameters can be used for a variety of purposes, including supporting program design, implementation, and prioritization of research needs.

Improve capacities to assess living resource response. Opportunities may exist to accelerate and improve living resource response to water quality management actions. Development of additional analytical and modeling capabilities would enable broader conclusions and more informed policy discussions about the role of the TMDL in improving living resources (Rose et al., 2023). The process of constructing models of multiple species' (ecological) responses to policy-relevant water quality conditions would generate useful information about identifying and prioritizing among living resource goals and targets. Analyses to address these management questions can be used to identify the tradeoffs among species (or classes of species) for alternative combinations of restoration actions, provide an assessment of progress for species of interest and their food webs, and inform what magnitude of changes in water quality and habitat are needed for certain sets of responses. Such analyses would be useful to inform discussions about refining the WQC and help establish clear public expectations about what living resource responses might be achieved from water quality improvement efforts.

If it is determined that more comprehensive and robust answers to the management questions about living resource response are needed, obtaining those answers will involve changes for assessing living resource responses. Such changes would require a concerted and coordinated effort, and options are summarized in Rose et al. (2023). Multiple issues will need to be considered, such as which species to address, what data are available, what models to use in new analyses, and how to leverage the extensive analyses already done. Rose et al. (2023) offer a framework for guiding the decision about how to assess living resource responses that can also be used to formulate a strategic analysis plan and interpret results.

Target monitoring and research to support adaptive management. Monitoring networks have been used primarily to identify water quality trends, inform and calibrate CBP watershed models, and track attainment of WQS. Ambient monitoring networks have not been designed to address critical uncertainties regarding nonpoint source program effectiveness, to assess the efficacy of particular technology or policy approaches, or to address critical uncertainties in estuary response.

There is widespread acknowledgment that additional monitoring is necessary (CBP, 2022), but an effective adaptive management process needs to identify key questions and appropriate monitoring metrics to assess system response and inform future actions. This is difficult because most water quality improvement strategies incorporate multiple actions, and determining effects of specific actions is difficult. Designing monitoring networks and associated research efforts for better evaluating current program efficacy and the implications for future program goals will be a critical undertaking. While acknowledging these challenges, additional monitoring and research are critical to assess the following situations.

Finer-scale watershed monitoring could improve assessment of the efficacy and uncertainties related to nonpoint source management, for example, finer-scale water quality monitoring and monitoring of intermediate indicators of change (e.g., groundwater, soil conditions, etc.). Monitoring for BMP implementation effectiveness to improve water quality would also need to include systematic evaluation of people's nutrient use, BMP adoption, and land use behavior. For example, nutrient use behavior and BMP maintenance are either assumed or poorly understood but have important influence on water quality changes. Relatively little has been invested in understanding behavioral change under different implementation programs.

For better understanding estuary response, elucidating relationships between stressor reduction efforts and achievement of WQS on a subsegment or smaller spatial scale is particularly important for shallow water habitats. It is notable that 17 of the 31 outcomes in the CBWA refer to conditions in shallow waters, which is an area of high engagement by stakeholders. The shallow waters present an opportunity to resolve critical uncertainties in the relationship of load reductions to living resource impacts. Dissolved oxygen in these habitats responds to a number of variables beyond nutrient and sediment reductions, and understanding these dynamics is critical for both identifying effective management actions in the shallow waters themselves and understanding the relationships between shallow water habitats and WQS in other habitats (e.g., deep water DO).

Tipping points in water clarity, DO, and SAV relationships have been shown to exist at the scale of subsystems in the Bay. Monitoring and research to either discern tipping points or determine the thresholds associated with the various tipping points can indicate progressive conditions on either a degradation or restoration trajectory; crossing the thresholds in the direction of restoration is of particular interest now.

List of Abbreviations

BMP	best management practice
CBP	Chesapeake Bay Program
CAFO	concentrated animal feeding operation
CAST	Chesapeake Assessment Scenario Tool
CBWA	Chesapeake Bay Watershed Agreement
Chl *a*	Chlorophyll *a*
CWA	Clean Water Act
DO	dissolved oxygen
DU	designated use
EPA	U.S. Environmental Protection Agency
GIT	goal implementation team
MS4	municipal separate storm sewer system
N	nitrogen
NASS	National Agricultural Statistics Service
NTN	Nontidal Monitoring Network
P	phosphorus
RIM	river input monitoring
SAV	submerged aquatic vegetation
SPARROW	Spatially Referenced Regression on Watershed Attributes
SRS	strategy review system
STAC	Scientific and Technical Advisory Committee
TMDL	total maximum daily load
TN	total nitrogen
TP	total phosphorus
TSS	total suspended sediment
USDA	U.S. Department of Agriculture
USGS	U.S. Geological Survey
WIP	watershed implementation plan
WQC	water quality criteria
WQS	water quality standards
WWTP	wastewater treatment plant

References

Aguilar, M. F., & Dymond, R. L. (2019). Evaluation of stormwater control measure performance uncertainty. *Journal of Environmental Engineering*, 145(10), 04019060. https://doi.org/10.1061/(ASCE)EE.1943-7870.0001590

Alberini, A., & Segerson, K. (2002). Assessing voluntary programs to improve environmental quality. *Environmental and Resource Economics*, 22, 157–184. https://doi.org/10.1023/A:1015519116167

Ariely, D. (2008). *Predictably irrational: The hidden forces that shape our decisions*. Harper Collins.

Ator, S. W. (2019). *Spatially referenced models of streamflow and nitrogen, phosphorus, and suspended-sediment loads in streams of the northeastern United States*. U.S. Geological Survey Scientific Investigations Report 2019-5118. https://doi.org/10.3133/sir20195118

Ator, S. W., Blomquist, J. D., Webber, J. S., & Chanat, J. G. (2020). Factors driving nutrient trends in streams of the Chesapeake Bay watershed. *Journal of Environmental Quality*, 49(4), 812–834. http://doi.org/10.1002/jeq2.20101

Ator, S. W., & Denver, J. M. (2015). *Understanding nutrients in the Chesapeake Bay watershed and implications for management and restoration—The Eastern Shore*. U.S. Geological Survey Circular 1406. http://dx.doi.org/10.3133/cir1406

Ator, S. W., Garcia, A. M., Schwarz, G. E., Blomquist, J. D., & Sekellick, A. J. (2019). Toward explaining nitrogen and phosphorus trends in Chesapeake Bay tributaries, 1992–2012. *JAWRA Journal of the American Water Resources Association,* 55(5), 1149–1168. https://doi.org/10.1111/1752-1688.12756

Baird, D., & Ulanowicz, R. E. (1989). The seasonal dynamics of the Chesapeake Bay ecosystem. *Ecological Monographs*, 59(4), 329–364. https://doi.org/10.2307/1943071

Basu, N. B., Van Meter, K. J., Byrnes, D. K., Van Cappellen, P., Brouwer, R., Jacobsen, B. H., Jarjo, J., Cunha, M. C., Nelson, N., Bhattacharya, R., Destouni, G., & Olsen, S. B. (2022). Managing nitrogen legacies to accelerate water quality improvement. *Nature Geoscience*, 15(2), 97–105. https://doi.org/10.1038/s41561-021-00889-9

Batiuk, R. A., Breitburg, D. L., Diaz, R. J., Cronin, T. M., Secor, D. H., & Thursby, G. (2009). Derivation of habitat-specific dissolved oxygen criteria for Chesapeake Bay and its tidal tributaries. *Journal of Experimental Marine Biology and Ecology*, 381, S204–S215. https://doi.org/10.1016/j.jembe.2009.07.023

Batiuk, R., Brownson, K., Dennison, W., Ehrhart, M., Hanson, J., Hanmer, R., Landry, B., Reichert-Nguyen, J., Soueidan, J., Tassone, S., & Vogt, B. (2023). *Rising watershed and bay water temperatures: Ecological implications and management responses*. STAC Publication Number 23-001, Chesapeake Bay Program Scientific and Technical Advisory Committee (STAC), Edgewater, MD. 505 pp.

Beegle, D. (2013). Nutrient management and the Chesapeake Bay. *Journal of Contemporary Water Research & Education*, 151(1), 3–8. https://doi.org/10.1111/j.1936-704X.2013.03146.x

Blanco-Canqui, H. (2018). Cover crops and water quality. *Agronomy Journal*, 110(5), 1633–1647. https://doi.org/10.2134/agronj2018.02.0077

Blankenship, K. (2022). Chesapeake Bay & the case of the missing chickens. *Bay Journal*, July 13. https://www.bayjournal.com/news/pollution/chesapeake-bay-the-case-of-the-missingchickens/article_a514fd52-02aa-11ed-891f-dfd96bff1134.html

Boesch, D. F. (2019). Barriers and bridges in abating coastal eutrophication. *Frontiers in Marine Science*, 6, 123. https://doi.org/10.3389/fmars.2019.00123

Böhlke, J. K. (2002). Groundwater recharge and agricultural contamination. *Hydrogeology Journal*, 10(1), 153–179. https://doi.org/10.1007/s10040-001-0183-3

Böhlke, J. K., & Denver, J. M. (1995). Combined use of groundwater dating, chemical, and isotopic analyses to resolve the history and fate of nitrate contamination in two agricultural watersheds, Atlantic coastal plain, Maryland. *Water Resources Research*, 31(9), 2319–2339. https://doi.org/10.1029/95WR01584

Bonham, J. G., Bosch, D. J., & Pease, J. W. (2006). Cost-effectiveness of nutrient management and buffers: Comparisons of two spatial scenarios. *Journal of Agricultural and Applied Economics*, 38(1), 17–32. https://doi.org/10.1017/S1074070800022045

Boomer, K., Boynton, W., Muller, A., Muller, D., & Sellner, K. (2019). *Revisiting coastal land-water interactions: The triblet connection*. STAC Publication Number 19-005, Chesapeake Bay Program Scientific and Technical Advisory Committee (STAC), Edgewater, MD. 37 pp. https://www.chesapeake.org/stac/wp-content/uploads/2019/11/STAC-Report_Triblets-FINAL.pdf

Bourger, L. (Ed.). (2011). *Converging waters: Integrating collaborative modeling with participatory processes to make water resources decisions*. U.S. Army Corps of Engineers, Institute for Water Resources. https://hdl.handle.net/11681/44821

Boynton, W. R., Hodgkins, C. L. S., O'Leary, C. A., Bailey, E. M., Bayard, A. R., & Wainger, L. A. (2014). Multi-decade responses of a tidal creek system to nutrient load reductions: Mattawoman Creek, Maryland USA. *Estuaries and Coasts*, 37, 111–127. https://doi.org/10.1007/s12237-013-9690-4

Boynton, W. R., Testa, J. M., & Kemp, W. M. (2009). *An ecological assessment of the Corsica River estuary and watershed: Scientific advice for future water quality management*. Report to Maryland Department of Natural Resources. University of Maryland Center for Environmental Science Technical Report Series No. TS-587-09. https://gonzo.cbl.umces.edu/documents/Publications/FULL%20CORSICA%20TEXT%20FINAL%2023OCT09.pdf

Breitburg, D. L., Craig, J. K., Fulford, R. S., Rose, K. A., Boynton, W. R., Brady, D. C., Ciotti, B. J., Diaz, R. J., Friedland, K. D., Hagy, J. D., Hart, D. R., Hines, A. H., Houde, E. D., Kolesar, S. E., Nixon, S. W., Rice, J. A., Secor, D. H., & Targett, T. E. (2009). Nutrient enrichment and fisheries exploitation: Interactive effects on estuarine living resources and their management. *Hydrobiologia*, 629(1), 31–47. https://doi.org/10.1007/s10750-009-9762-4

Breitburg, D. L, Hondorp, D. W., Davias, L. A., & Diaz, R. J. (2009). Hypoxia, nitrogen, and fisheries: Integrating effects across local and global landscapes. *Annual Review of Marine Science*, 1(1), 329–349. https://doi.org/10.1146/annurev.marine.010908.163754

Breitburg, D., Levin, L. A., Oschlies, A., Grégoire, M., Chavez, F. P., Conley, D. J., Garçon, V., Gilbert, D., Gutiérrez, D., Isensee, K., Jacinto, G. S., Limburg, K. E., Montes, I., Naqvi, S.

W. A., Pitcher, G. C., Rabalais, N. N., Roman, M. R., Rose, K. A., Siebel, B. A., … Zhang, J. (2018). Declining oxygen in the global ocean and coastal waters. *Science*, 359(6371). https://doi.org/10.1126/science.aam7240

Bryant, R., Baldwin, A., Cahall, B., Christianson, L., Jaynes, D., Penn, C., & Schwartz, S. (2019). *Best management practices for agricultural ditch management in the phase 6 Chesapeake Bay Watershed Model*. L. Collins, J. Hanson & C. Gill (Eds.). CBP/TRS-326-19. https://d18lev1ok5leia.cloudfront.net/chesapeakebay/documents/ag_ditches_bmp_panel_report_draft_for_cbp_review_04sep2019.pdf

Carpenter, S. R. (2008). Phosphorus control is critical to mitigating eutrophication. *Proceedings of the National Academy of Sciences*, 105(32), 11039–11040. https://doi.org/10.1073/pnas.0806112105

Cerco, C.F., & Noel, M.R. (2007). Can oyster restoration reverse cultural eutrophication in Chesapeake Bay? *Estuaries and Coasts*, 30, 331–343. https://doi.org/10.1007/BF02700175

Cerosaletti, P. E., Fox, D. G., & Chase, L. E. (2004). Phosphorus reduction through precision feeding of dairy cattle. *Journal of Dairy Science*, 877, 2314–2323. https://doi.org/10.3168/jds.S0022-0302(04)70053-3

Chanat, J. G., & Yang, G. (2018). Exploring drivers of regional water-quality change using differential spatially referenced regression—A pilot study in the Chesapeake Bay watershed. *Water Resources Research*, 54(10), 8120–8145. https://doi.org/10.1029/2017WR022403

Chesapeake Bay Commission. (2017). *Boots on the ground: Improving technical assistance to farmers*. https://www.chesbay.us/library/public/documents/Policy-Reports/CBC-TA-Report-Boots-on-the-Ground.pdf

Chesapeake Bay Program (CBP). (n.d.-a). *2025 Watershed Implementation Plans*. Chesapeake Progress. https://www.chesapeakeprogress.com/clean-water/watershed-implementation-plans

Chesapeake Bay Program (CBP). (n.d.-b). *Chesapeake Decisions* https://www.chesapeakebay.net/what/what-guides-us/decisions

Chesapeake Bay Program (CBP). (n.d.-c). *Model documentation*. https://cast.chesapeakebay.net/Documentation/ModelDocumentation

Chesapeake Bay Program (CBP). (n.d.-d). *Tidal water quality change: 2021 results*. Integrated trends analysis team. https://d18lev1ok5leia.cloudfront.net/chesapeakebay/documents/2021-Tidal-Trends-Summary-Document.pdf

Chesapeake Bay Program (CBP). (n.d.-e). *Water quality standards attainment and monitoring*. Chesapeake Progress. https://www.chesapeakeprogress.com/clean-water/water-quality

Chesapeake Bay Program (CBP). (1987). *Chesapeake Bay Agreement*. https://www.chesapeakebay.net/what/publications/chesapeake-bay-agreement-1987

Chesapeake Bay Program (CBP). (2008). *Chesapeake Bay segmentation scheme (for 303d listing—92 segments)*. Maps. https://www.chesapeakebay.net/what/maps/tag/segments

Chesapeake Bay Program (CBP). (2014). *Chesapeake Bay Watershed Agreement*. https://www.chesapeakebay.net/what/what-guides-us/watershed-agreement

Chesapeake Bay Program (CBP). (2022). *Enhancing the Chesapeake Bay program monitoring networks—A report to the Principals' Staff Committee*. https://d18lev1ok5leia.cloudfront.net/chesapeakebay/documents/FINAL_Enhancing_the_Chesapeake_Bay_Program_Monitoring_Networks_A-Report_to_the_Principals_Staff_Committee_10.13.22-1_2022-10-18-211409_qtmr.pdf

Chesapeake Bay Program (CBP). (2020). *2025 WIP and Water Quality Standards Attainment and Monitoring Outcomes.* https://d18lev1ok5leia.cloudfront.net/chesapeakebay/documents/wq_revised_clean_ms_2021-22_v1.6.2021.pdf

Choi, D. S., Ready, R. C., & Shortle, J. S. (2020).Valuing water quality benefits from adopting best management practices: A spatial approach. *Journal of Environmental Quality*, 49(3), 582–592. https://doi.org/10.1002/jeq2.20005

Christensen, V., Beattie, A., Buchanan, C., Ma, H., Martell, S. J. D., Latour, R., Preikshot, D., Sigrist, M., Uphoff, J. H., Walters, C. J., Wood, R. J., & Townsend, H. (2009). *Fisheries ecosystem model of the Chesapeake Bay: Methodology, parameterization, and model exploration*. U.S. Department of Commerce NOAA Technical Memorandum NMFS-F/SPO-106. https://repository.library.noaa.gov/view/noaa/4180

Claassen, R., Horowitz, J., Duquette, E., & Ueda, K. (2014). *Additionality in US agricultural conservation and regulatory offset programs*. USDA-ERS Economic Research Report 170. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2502846

Clune, J. W., & Capel, P. D. (Eds.). (2021). *Nitrogen in the Chesapeake Bay watershed—A century of change, 1950–2050 (ver. 1.1, December 2021)*. U.S. Geological Survey Circular 1486. https://doi.org/10.3133/cir1486

Collins, L., Stephenson, K., Palm-Forster, L., Power, L., Gibson, A., Arbuckle, J., Handen, A., & Read, D. (2022). *Overcoming the hurdle: Addressing implementation of agricultural best management practices (BMPs) through a social science lens*. STAC Publication Number 22-002, Chesapeake Bay Program Scientific and Technical Advisory Committee (STAC), Edgewater, MD. 107 pp. https://www.chesapeake.org/stac/wp-content/uploads/2022/10/FINAL-Report_Overcoming-the-Hurdle_22-002.pdf

Daryanto, S., Wang, L., & Jacinthe, P. A. (2017). Meta-analysis of phosphorus loss from no-till soils. *Journal of Environmental Quality*, 46(5), 1028–1037. https://doi.org/10.2134/jeq2017.03.0121

Denver, J. M., Tesoriero, A. J., & Barbaro, J. R. (2010). Trends and transformation of nutrients and pesticides in a coastal plain aquifer system, United States. *Journal of Environmental Quality*, *39*(1), 154–167. https://doi.org/10.2134/jeq2009.0107

Du, J., Shen, J., Park, K., Wang, Y. P., & Yu, X. (2018). Worsened physical condition due to climate change contributes to the increasing hypoxia in Chesapeake Bay. *Science of The Total Environment*, 630, 707–717. https://doi.org/10.1016/j.scitotenv.2018.02.265

Duke, J. M., Johnston, R. J., Shober, A. L., & Liu, Z. (2022). Barriers to cover crop adoption: Evidence from parallel surveys in Maryland and Ohio. *Journal of Soil and Water Conservation*, 77(2), 198–211. https://doi.org/10.2489/jswc.2022.00062

Duncan, E. W., Osmond, D. L., Shober, A. L., Starr, L., Tomlinson, P., Kovar, J. L., Moorman, T. B., Peterson, H. M., Fiorellino, N. M., & Reid, K. (2019). Phosphorus and soil health management practices. *Agricultural and Environmental Letters*, 4(1), 190014. https://doi.org/10.2134/ael2019.04.0014

Easton, Z. M., Bock, E., & Stephenson, K. (2019). Feasibility of using woodchip bioreactors to treat legacy nitrogen to meet Chesapeake Bay water quality goals. *Environmental Science & Technology*, 53(21), 12291–12299. https://doi.org/10.1021/acs.est.9b04919

Easton, Z., Scavia, D., Alexander, R., Band, L., Boomer, K., Kleinman, P., Martin, J., Miller, A., Pizzuto, J., Smith, D., & Welty, C. (2017). *Scientific and Technical Advisory Committee review of the phase 6 Chesapeake Bay Watershed Model*. STAC Publication Number 17-007, Chesapeake Bay Program Scientific and Technical Advisory Committee (STAC), Edgewater, MD. 47 pp. https://www.chesapeake.org/pubs/379_Easton2017.pdf

Easton, Z., Stephenson, K., Benham, B., Böhlke, J. K., Buda, A., Collick, A., Fowler, L., Gilinsky, E., Hershner, C., Miller, A., Noe, G., Palm-Forster, L., & Thompson, T. (2023). *Evaluation of Management Efforts to Reduce Nutrient and Sediment Contributions to the Chesapeake Bay Estuary*. STAC Publication Number 23-003, Chesapeake Bay Program Scientific and Technical Advisory Committee (STAC), Edgewater, MD. 53 pp.

Easton, Z. M., Stephenson, K., Collick, A., Fleming, P. M., Kellner, E., Martin, J., Ribaudo, M., & Shenk, G. (2020). *Increasing effectiveness and reducing the cost of non-point source best management practice implementation: Is targeting the answer?* STAC Publication Number 20-002, Chesapeake Bay Program Scientific and Technical Advisory Committee (STAC), Edgewater, MD. 53 pp. https://www.chesapeake.org/stac/wp-content/uploads/2020/02/FINAL_STAC-Report_BMP-Targeting.pdf

Easton, Z. M., Walter, M. T., & Steenhuis, T. S. (2008). Combined monitoring and modeling indicate the most effective agricultural best management practices. *Journal of Environmental Quality*, 37(5), 1798–1809. https://doi.org/10.2134/jeq2007.0522

Fabrizio, M. C., Tuckey, T. D., Bever, A. J., & MacWilliams, M. L. (2021). The extent of seasonally suitable habitats may limit forage fish production in a temperate estuary. *Frontiers in Marine Science*, 8, 706666. https://doi.org/10.3389/fmars.2021.706666

Fanelli, R. M., Blomquist, J. D., & Hirsch, R. M. (2019). Point sources and agricultural practices control spatial-temporal patterns of orthophosphate in tributaries to Chesapeake Bay. *Science of The Total Environment*, 652, 422–433. https://doi.org/10.1016/j.scitotenv.2018.10.062

Fiorellino, N., Kratochvil, R., & Coale, F. (2017). Long-term agronomic drawdown of soil phosphorus in mid-Atlantic coastal plain soils. *Agronomy Journal*, 109(2), 455–461. https://doi.org/10.2134/agronj2016.07.0409

Fisher, T. R., Fox, R. J., Gustafson, A. B., Koontz, E., Lepori-Bui, M, & Lewis, J. (2021). Localized water quality improvement in the Choptank estuary, a tributary of Chesapeake Bay. *Estuaries and Coasts*, 44, 1274–1293. https://doi.org/10.1007/s12237-020-00872-4

Fleming, B. J., Archfield, S. A., Hirsch, R. M., Kiang, J. E., & Wolock, D. M. (2020) Spatial and temporal patterns of low streamflow and precipitation changes in the Chesapeake Bay watershed. *JAWRA Journal of the American Water Resources Association*, 57(1), 96–108. https://doi.org/10.1111/1752-1688.12892

Fleming, P. M., Lichtenberg, E., & Newburn, D. A. (2018). Evaluating impacts of agricultural cost sharing on water quality: Additionality, crowding in, and slippage. *Journal of Environmental Economics and Management*, 92, 1–19. https://doi.org/10.1016/j.jeem.2018.08.007

Fleming, P. M., Stephenson, K., Collick, A., & Easton, Z. (2022). Targeting for nonpoint source pollution reduction: A synthesis of lessons learned, remaining challenges, and emerging opportunities. *Journal of Environmental Management*, 308, 114649. https://doi.org/10.1016/j.jenvman.2022.114649

Flynn, K. C., Spiegal, S., Kleinman, P. J. A., Meinen, R. J., & Smith, D. R. (2023). Manureshed management to overcome longstanding nutrient imbalances in US agriculture. *Resources, Conservation and Recycling*, 188, 106632. https://doi.org/10.1016/j.resconrec.2022.106632

Fox, R. J., Fisher, T. R., Gustafson, A. B., Koontz, E. L., Lepori-Bui, M., Kvalnes, K. L., Bunnell-Young, D. E., Gardner, J. R., Lewis, J., Winsten, J. R., Fisher, K. A., & Silaphone, K. (2021). An evaluation of the Chesapeake Bay management strategy to improve water quality in small agricultural watersheds. *Journal of Environmental Management*, 299, 113478. https://doi.org/10.1016/j.jenvman.2021.113478

Frankel, L. T., Friedrichs, M. A. M., St-Laurent, P., Bever, A. J., Lipcius, R. N., Bhatt, G., & Shenk, G. W. (2022). Nitrogen reductions have decreased hypoxia in the Chesapeake Bay: Evidence from empirical and numerical modeling. *Science of The Total Environment*, 814, 152722. https://doi.org/10.1016/j.scitotenv.2021.152722

Freedman, P. L., Reckhow, K. H., Shabman, L., Benaman, J., Schwer, R., & Stiles, T. (2008). A new approach to adaptive implementation in the TMDL program. *Water Practice*, 2(1), 1–9.

Froelich, P. N. (1988). Kinetic control of dissolved phosphate in natural rivers and estuaries: A primer on the phosphate buffer mechanism. *Limnology and Oceanography*, 33(4, part 2), 649–668. https://doi.org/10.4319/lo.1988.33.4part2.0649

Fulford, R. S., Breitburg, D. L., Luckenbach, M., & Newell, R. I. E. (2010). Evaluating ecosystem response to oyster restoration and nutrient load reduction with a multispecies bioenergetics model. *Ecological Applications*, 20(4), 915–934. https://doi.org/10.1890/08-1796.1

Fulton, E. A., Link, J. S., Kaplan, I. C., Savina-Rolland, M., Johnson, P., Ainsworth, C., Horne, P., Gorton, R., Gamble, R. J., Smith, A. D. M., & Smith, D. C. (2011). Lessons in modelling and management of marine ecosystems: The Atlantis experience. *Fish and Fisheries*, 12(2), 171–188. https://doi.org/10.1111/j.1467-2979.2011.00412.x

Ganju, N. K., Testa, J. M., Suttles, S. E., & Aretxabaleta, A. L. (2020). Spatiotemporal variability of light attenuation and net ecosystem metabolism in a back-barrier estuary. *Ocean Science*, 16(3), 593–614. https://doi.org/10.5194/os-16-593-2020

Giri, S., Nejadhashemi, A. P., & Woznicki, S. A. (2012). Evaluation of targeting methods for implementation of best management practices in the Saginaw River watershed. *Journal of Environmental Management*, 103, 24–40. https://doi.org/10.1016/j.jenvman.2012.02.033

Goals Project. (2015). *The Baylands and climate change: What we can do. Baylands ecosystem habitat goals science update 2015*. San Francisco Bay Area Wetlands Ecosystem Goals

Project, California State Coastal Conservancy. https://www.sfei.org/sites/default/files/biblio_files/Baylands_Complete_Report.pdf

Gonzalez, G., Mosley, A., & Stephenson, K. (2016). An analysis of stormwater utility incentive programs in the Chesapeake Bay. *Watershed Science Bulletin*, December. http://owl.cwp.org/mdocs-posts/an-analysis-of-stormwater-utility-incentive-programs-in-the-chesapeake-bay/

Greene, E. A., LaMotte, A. E., & Cullinan, K.-A. (2005). *Ground-water vulnerability to nitrate contamination at multiple thresholds in the mid-Atlantic region using spatial probability models*. U.S. Geological Survey Scientific Investigations Report 2004-5118. https://doi.org/10.3133/sir20045118

Gregory, R., Failing, L., Harstone, M. Long, G., McDaniels, T., & Ohlson, D. (2012). *Structured decision making: A practical guide to environmental management choices*. Wiley-Blackwell.

Gruber, R. K., & Kemp, W. M. (2010). Feedback effects in a coastal canopy-forming submersed plant bed. *Limnology and Oceanography*, 55(6), 2285–2298. https://doi.org/10.4319/lo.2010.55.6.2285

Gurbisz, C., & Kemp, W. M. (2014). Unexpected resurgence of a large submersed plant bed in Chesapeake Bay: Analysis of time series data. *Limnology and Oceanography*, 59(2), 482–494. https://doi.org/10.4319/lo.2014.59.2.0482

Gurbisz, C., Kemp, W. M., Cornwell, J. C., Sanford, L. P., Owens, M. S., & Hinkle, D. C. (2017). Interactive effects of physical and biogeochemical feedback processes in a large submersed plant bed. *Estuaries and Coasts*, 40(6), 1626–1641. https://doi.org/10.1007/s12237-017-0249-7

Hanson, J., Bock, E., Asfaw, B., & Easton, Z. M. (2022). *A systematic review of Chesapeake Bay climate change impacts and uncertainty: Watershed processes, pollutant delivery and BMP performance*. CBP/TRS-330-22. https://bit.ly/BMP-CC-synth

Heathwaite, L., Sharpley, A., & Gburek, W. (2000). A conceptual approach for integrating phosphorus and nitrogen management at watershed scales. *Journal of Environmental Quality*, 29(1), 158–166. https://doi.org/10.2134/jeq2000.00472425002900010020x

Hernandez Cordero, A. L., Tango, P. M., & Batiuk, R. A. (2020). Development of a multimetric water quality indicator for tracking progress towards the achievement of Chesapeake Bay water quality standards. *Environmental Monitoring and Assessment*, 192, 94. https://doi.org/10.1007/s10661-019-7969-z

Hershner, C. (2011). The certainty of confronting uncertainty in the Chesapeake Bay restoration effort. *Choices*, 26(3).

Higgins, P., & Male, T. (2019). *Sandboxing nature: How regulatory sandboxes could help restore species, enhance water quality and build better habitats faster*. Environmental Policy Innovation Center. https://www.policyinnovation.org/blog/sandbox?rq=sandboxing%20nature

Hinson, K. E., Friedrichs, A. M., St-Laurent, P., Da, F., & Najjar, R. G. (2022). Extent and causes of Chesapeake Bay warming. *JAWRA Journal of the American Water Resources Association*, 58(6), 805–825. https://doi.org/10.1111/1752-1688.12916

Holman, I. P., Whelan, M. J., Howden, N. J. K., Bellamy, P. H., Willby, N. J., Rivas-Casado, M., & McConvey, P. (2008). Phosphorus in groundwater—An overlooked contributor to eutrophication? *Hydrological Processes*, 22(26), 5121–5127. https://doi.org/10.1002/hyp.7198

Hood, R. R., Shenk, G. W., Dixon, R. L., Smith, S. M., Ball, W. P., Bash, J. O., Batiuk, R, Boomer, K., Brady, D. C., Cerco, C., Claggett, P., de Mutsert, K., Easton, Z. M., Elmore, A. J., Friedrichs, M. A. M., Harris, L. A., Ihde, T. F., Lacher, L., Li, L., Linker, L. C., … Zhang, Y. J. (2021). The Chesapeake Bay Program modeling system: Overview and recommendations for future development. *Ecological Modeling*, 456, 109635. https://doi.org/10.1016/j.ecolmodel.2021.109635

Hopkins, K. G., Bhaskar, A. S., Woznicki, S. A., & Fanelli, R. M. (2020). Changes in event-based streamflow magnitude and timing after suburban development with infiltration-based stormwater management. *Hydrological Processes*, 34(2), 387–403. https://doi.org/10.1002/hyp.13593

Howard, G. (2020). Additionality violations in agricultural payment for service programs: Experimental evidence. *Land Economics*, 96(2), 244–264. https://doi.org/10.3368/le.96.2.244

Hyer, K. E., Phillips, S. W., Ator, S. W., Moyer, D. L., Webber, J. S., Felver, R., Keisman, J. L., McDonnell, L. A., Murphy, R., Trentacoste, E. M., Zhang, Q., Dennison, W. C., Swanson, S., Walsh, B., Hawkey, J., & Taillie, D. (2021). *Nutrient trends and drivers in the Chesapeake Bay watershed*. U.S. Geological Survey Fact Sheet 2020-3069. https://doi.org/10.3133/fs20203069

Ihde, T. F., Kaplan, I. C., Fulton, E. A., Gray, I. A., Hasan, M., Bruce, D., Slacum, W., & Townsend, H. M. (2016). *Design and parameterization of the Chesapeake Bay Atlantis Model: A spatially explicit end-to-end ecosystem model*. U.S. Department of Commerce NOAA Technical Memorandum NMFS-F/SPO-166. https://repository.library.noaa.gov/view/noaa/17357

Ihde, T. F., & Townsend, H. M. (2017). Accounting for multiple stressors influencing living marine resources in a complex estuarine ecosystem using an Atlantis model. *Ecological Modelling*, 365, 1–9. https://doi.org/10.1016/j.ecolmodel.2017.09.010

Jackson, C. R., Martin, J. K., Leigh, D. S., & West, L. T. (2005). A southeastern piedmont watershed sediment budget: Evidence for a multi-millennial agricultural legacy. *Journal of Soil and Water Conservation*, 60(6), 298–310.

Jackson-Smith, D. B., Halling, M., de la Hoz, E., McEvoy, J. P., & Horsburgh, J. S. (2010). Measuring conservation program best management practice implementation and maintenance at the watershed scale. *Journal of Soil and Water Conservation*, 65(6), 413–423. https://doi-org.proxy2.cl.msu.edu/10.2489/jswc.65.6.413

Jacobson, R. B., & Coleman, D. J. (1986). Stratigraphy and recent evolution of Maryland Piedmont floodplains. *American Journal of Science*, 286(8), 617–637. https://doi.org/10.2475/ajs.286.8.617

Jarvie, H. P., Johnson, L. T., Sharpley, A. N., Smith, D. R., Baker, D. B., Bruulsema, T. W., & Confesor, R. (2017). Increased soluble phosphorus loads to Lake Erie: Unintended consequences of conservation practices? *Journal of Environmental Quality*, 46(1), 123–132. https://doi.org/10.2134/jeq2016.07.0248

Jefferson, A. J., Bhaskar, A. S., Hopkins, K. G., Fanelli, R., Avellandeda, P. M., & McMillan, S. K. (2017). Stormwater management network effectiveness and implications for urban watershed function: A critical review. *Hydrological Processes*, 31(23), 4056–4080. https://doi.org/10.1002/hyp.11347

Kahneman, D., Sibony, O., & Sunstein, C. R. (2021). *Noise: A flaw in human judgement*. Little, Brown Spark.

Kast, J. B., Kalcic, M., Wilson, R., Jackson-Smith, D., Breyfogle, N., & Martin, J. (2021). Evaluating the efficacy of targeting options for conservation practice adoption on watershed-scale phosphorus reductions. *Water Research*, 201, 117375. https://doi.org/10.1016/j.watres.2021.117375

Kaufman, Z., Abler, D., Shortle, J., Harper, J., Hamlett, J., & Feather, P. (2014). Agricultural costs of the Chesapeake Bay total maximum daily load. *Environmental Science & Technology*, 48(24), 14131–14138. https://doi.org/10.1021/es502696t

Kaufman, D. E., Shenk, G., Bhatt, G., Asplen, K., Devereux, O., Rigelman, J., Ellis, H., Hobbs, B. F., Bosch, D. J., VanHoutven, G. L., McGarity, A., Linker, L. C., & Ball, W. P. (2021). Supporting cost-effective watershed management strategies for Chesapeake Bay using a modeling and optimization framework. *Environmental Modeling & Software*, 144, 105141. https://doi.org/10.1016/j.envsoft.2021.105141

Keisman, J., Blomquist, J., Böhlke, J. K., Davis-Martin, J., Dennison, W., Friedrichs, C., Murphy, R., Phillips, S., Testa, J., Trentacoste, E., & Weller, D. (2018). *Integrating recent findings to explain water-quality change: Support for the mid-point assessment and beyond*. STAC Publication Number 18-005, Chesapeake Bay Program Scientific and Technical Advisory Committee (STAC), Edgewater, MD. 27 pp. http://www.chesapeake.org/pubs/394_Keisman2018.pdf

Keisman, J., Devereux, O. H., LaMotte, A. E., Sekellick, A., & Blomquist, J. (2018). *Manure and fertilizer inputs to land in the Chesapeake Bay watershed, 1950–2012*. U.S. Geological Survey Scientific Investigations Report 2018-5022. https://doi.org/10.3133/sir20185022

Kemp, W. M., Sampou, P., Caffrey, J., Mayer, M., Henriksen, K., & Boynton, W. R. (1990). Ammonium recycling versus denitrification in Chesapeake Bay sediments. *Limnology and Oceanography*, 35(7), 1545–1563. https://doi.org/10.4319/lo.1990.35.7.1545

Kemp, W. M., Batleson, R., Bergstrom, P., Carter, V., Gallegos, C. L., Hunley, W., Karrh, L., Koch, E. W., Landwehr, J. M., Moore, K. A., Murray, L., Naylor, M., Rybicki, N. B., Stevenson, J. C., & Wilson, D. J. (2004). Habitat requirements for submerged aquatic vegetation in Chesapeake Bay: Water quality, light regime, and physical-chemical factors. *Estuaries*, 27, 363–377. https://doi.org/10.1007/BF02803529

Kemp, W. M., Boynton, W. R., Adolf, J. E., Boesch, D. F., Boicourt, W. C., Brush, G. Cornwell, J. C., Fisher, T. R., Glibert, P. M., Hagy, J. D., Harding, L. W., Houde, E. D., Kimmel, D. G., Miller, W. D., Newell, R. I. E., Roman, M. R., Smith, E. M., & Stevenson, J. C. (2005). Eutrophication of Chesapeake Bay: Historical trends and ecological interactions. *Marine Ecology Progress Series*, 303, 1–29. https://doi.org/10.3354/meps303001

Khanna M., Yang, W., Farnsworth, R., & Önal, H. (2003). Cost-effective targeting of land retirement to improve water quality with endogenous sediment deposition coefficients. *American Journal of Agricultural Economics*, 85(3), 538–553. https://doi.org/10.1111/1467-8276.t01-1-00454

Kleinman, P., Blunk, K. S., Bryant, R., Saporito, L., Beegle, D., Czymmek, K., Ketterings, Q., Sims, T., Shortle, J., McGrath, J., Coale, F., Dubin, M., Dostie, D., Maguire, R., Meinen, R., Allen, A., O'Neill, K., Garber, L., Davis, M., Clark, B., … Smith, M. (2012). Managing manure for sustainable livestock production in the Chesapeake Bay watershed. *Journal of Soil and Water Conservation*, 67(2), 54A–61A. https://doi.org/10.2489/jswc.67.2.54A

Kleinman, P. J. A., Fanelli, R. M., Hirsch, R. M., Buda, A. R., Easton, Z. M., Wainger, L. A., Brosch, C., Lowenfish, M., Collick, A. S., Shirmohammadi, A., Boomer, K, Hubbart, J. A., Bryant, R. B., & Shenk, G. W. (2019). Phosphorus and the Chesapeake Bay: Lingering issues and emerging concerns for agriculture. *Journal of Environmental Quality*, 48(5), 1191–1203. https://doi.org/10.2134/jeq2019.03.0112

Kleinman, P. J., A., Osmond, D. L., Christianson, L. E., Flaten, D. N., Ippolito, J. A., Jarvie, H. P., Kaye, J. P., King, K. W., Leytem, A. B., McGrath, J. M., Nelson, N. O., Shober, A. L., Smith, D. R., Staver, K. W., & Sharpley, A. N. (2022). Addressing conservation practice limitations and trade-offs for reducing phosphorus loss from agricultural fields. *Agricultural & Environmental Letters*, 7(2), e20084. https://doi.org/10.1002/ael2.20084

Kleinman, P. J. A., Sharpley, A. N., Buda, A. R., Easton, Z. M., Lory, J. A., Osmond, D. L., Radcliffe, D. E., Nelson, N. O., Veith, T. L., & Doody, D. G. (2017). The promise, practice, and state of planning tools to assess site vulnerability to runoff phosphorus loss. *Journal of Environmental Quality*, 46(6), 1243–1249. https://doi.org/10.2134/jeq2017.10.0395

Kleinman, P. J., A., Sharpley, A. N., McDowell, R. W., Flaten, D. N., Buda, A. R., Tao, L., Bergstrom, L., & Zhu, Q. (2011). Managing agricultural phosphorus for water quality protection: Principles for progress. *Plant and Soil*, 349, 169–182. https://doi.org/10.1007/s11104-011-0832-9

Kleinman, P. J. A., Sharpley, A. N., Withers, P. J. A., Bergström, L., Johnston, L. T., & Doody, D. G. (2015). Implementing agricultural phosphorus science and management to combat eutrophication. *AMBIO*, 44(Suppl 2), 297–310. https://doi.org/10.1007/s13280-015-0631-2

Kleinman, P. J. A., Spiegal, S., Liu, J., Holly, M., Church, C., & Ramirez-Avila, J. (2020). Managing animal manure to minimize phosphorus losses from land to water. In H. M. Waldrip, P. H. Pagliari & Z. He (Eds.), *Animal manure: Production, characteristics, environmental concerns, and management*, (pp. 201–228). American Society of Agronomy. https://doi.org/10.2134/asaspecpub67.c12

Kling, C. L. (2013). State level efforts to regulate agricultural sources of water quality impairment. *Choices*, 28(3), 1–4.

Kornis, M. S., Breitburg, D., Balouskus, R., Bilkovic, D. M., Davias, L. A., Giordano, S., Heggie, K., Hines, A. H., Jacobs, J. M., Jordan, T. E., King, R. S., Patrick, C. J., Seitz, R. D., Soulen, H., Targett, T. E., Weller, D. E., Whigham, D. F., & Uphoff, J. (2017). Linking the abundance of estuarine fish and crustaceans in nearshore waters to shoreline hardening and land

cover. *Estuaries and Coasts*, 40, 1464–1486. https://doi.org/10.1007/s12237-017-0213-6
Koudahe, K., Allen, S. C., & Djaman, K. (2021). Critical review of the impact of cover crops on soil properties. *International Soil and Water Conservation Research*, 10(3), 343–354. https://doi.org/10.1016/j.iswcr.2022.03.003
Lefcheck, J. S., Orth, R. J., Dennison, W. C., Wilcox, D. J., Murphy, R. R., Keisman, J., Gurbisz, C., Hannam, M., Landry, J. B., Moore, K. A., Patrick, C. J., Testa, J., Weller, D. E., & Batiuk, R. A. (2018). Long-term nutrient reductions lead to the unprecedented recovery of a temperate coastal region. *Proceedings of the National Academy of Sciences*, 115(14), 3658–3662. https://doi.org/10.1073/pnas.1715798115
Lefcheck, J. S., Wilcox, J. S., Murphy, R. R., Marion, S. R., & Orth, R. J. (2017). Multiple stressors threaten the imperiled coastal foundation species eelgrass (Zostera marina) in Chesapeake Bay, USA. *Global Change Biology*, 23(9), 3474–3483. https://doi.org/10.1111/gcb.13623
LeFevre, G. H., Paus, K. H., Natarajan, P., Gulliver, J. S., Novak, P. J., & Hozalski, R. M. (2015). Review of dissolved pollutants in urban storm water and their removal and fate in bioretention cells. *Journal of Environmental Engineering*, 141(1), 04014050. https://doi.org/10.1061/(ASCE)EE.1943-7870.0000876
Li, C., Fletcher, T. D., Duncan, H. P., & Burns, M. J. (2017). Can stormwater control measures restore altered urban flow regimes at the catchment scale? *Journal of Hydrology*, 549, 631–653. https://doi.org/10.1016/j.jhydrol.2017.03.037
Lichtenberg, E., & Smith-Ramírez, R. (2011). Slippage in conservation cost sharing. *American Journal of Agricultural Economics*, 92(1), 113–129. https://doi.org/10.1093/ajae/aaq124
Lindsey, B. D., Phillips, S., Donnelly, C. A., Speiran, G. K., Plummer, N., Böhlke, J. K., Focazio, M. J., Burton, W. C., & Busenberg, E. (2003). *Residence times and nitrate transport in ground water discharging to streams in the Chesapeake Bay watershed*. U.S. Geological Survey Water Resources Investigations Report 2003-4035. https://doi.org/10.3133/wri034035
Linker, L. C., Batiuk, R. A., Shenk, G. W., & Cerco, C. F. (2013). Development of the Chesapeake Bay watershed total maximum daily load allocation. *JAWRA Journal of the American Water Resources Association*, 49(5), 986–1006. https://doi.org/10.1111/jawr.12105
Lintern, A., McPhillips, L., Winfrey, B., Duncan, J., & Grady, C. (2020). Best management practices for diffuse nutrient pollution: Wicked problems across urban and agricultural watersheds. *Environmental Science & Technology*, 54(15), 9159–9174. https://doi.org/10.1021/acs.est.9b07511
Liu, H., Zhang, W., Irwin, E., Kast, J., Aloysius, N., Martin, J., & Kalcic M. (2020). Best management practices and nutrient reduction: An integrated economics-hydrologic model of the western Lake Erie Basin. *Land Economics*, 96(4), 510–530. https://doi.org/10.3368/wple.96.4.510
Liu, Y., Engel, B. A., Flanagan, D. C., Gitau, M. W., McMillan, S. K., & Chaubey, I. (2017). A review on effectiveness of best management practices in improving hydrology and water quality: Needs and opportunities. *Science of The Total Environment*, 601–602, 580–593. https://doi.org/10.1016/j.scitotenv.2017.05.212

Liu, Y., Engel, B. A., Flanagan, D. C., Gitau, M. W., McMillan, S. K., Chaubey, I., & Singh, S. (2018). Modeling framework for representing long-term effectiveness of best management practices in addressing hydrology and water quality problems: Framework development and demonstration using a Bayesian method. *Journal of Hydrology*, 560, 530–545. https://doi.org/10.1016/j.jhydrol.2018.03.053

Lyerly, C. M., Hernández Cordero, A. L., Foreman, K. L., Phillips, S. W., & Dennison, W. C. (2014). *New insights: Science-based evidence of water quality improvements, challenges, and opportunities in the Chesapeake Bay*. https://ian.umces.edu/site/assets/files/11070/new-insights-science-based-evidence-of-water-quality-improvements-challenges-and-opportunities-in-the-chesapeake.pdf

Ma, H., Townsend, H., Zhang, X., Sigrist, M., & Christensen, V. (2010). Using a fisheries ecosystem model with a water quality model to explore trophic and habitat impacts on a fisheries stock: A case study of the blue crab population in the Chesapeake Bay. *Ecological Modelling*, 221(7), 997–1004. https://doi.org/10.1016/j.ecolmodel.2009.01.026

Macías-Tapia, A., Mulholland, M. R., Selden, C. R., Loftis, J. D., & Bernhardt, P. W. (2021). Effects of tidal flooding on estuarine biogeochemistry: Quantifying flood-driven nitrogen inputs in an urban, lower Chesapeake Bay sub-tributary. *Water Research*, 201, 117329. https://doi.org/10.1016/j.watres.2021.117329

Macrae, M., Jarvie, H., Brouwer, R., Gunn, G., Reid, K., Joosse, P., King, K. Kleinman, P., Smith, D., Williams, M., & Zwonitzer, M. (2021). One size does not fit all: Toward regional conservation practice guidance to reduce phosphorus loss risk in the Lake Erie watershed. *Journal of Environmental Quality*, 50(3), 529–546. https://doi.org/10.1002/jeq2.20218

Maille, P., Collins, A. R., & Gillies, N. (2009). Performance-based payments for water quality: Experiences from a field experiment. *Journal of Soil and Water Conservation*, 64(3), 85A–87A. https://doi.org/10.2489/jswc.64.3.85A

Mallakpour, I., & Villarini, G. (2017). Analysis of changes in the magnitude, frequency, and seasonality of heavy precipitation over the contiguous USA. *Theoretical and Applied Climatology*, 130(1), 345–363. https://doi.org/10.1007/s00704-016-1881-z

Marchau, V. A. W. J., Walker, W. E., Bloemen, P. J. T. M., & Popper, S. W. (Eds.). (2019). *Decision making under deep uncertainty*. Springer.

Maryland Department of the Environment (MDE). 2019*. Maryland's Phase III Watershed Implementation Plan to Restore Chesapeake Bay by 2025*. https://mde.maryland.gov/programs/Water/TMDL/TMDLImplementation/Pages/Phase3WIP.aspx

Mason, C. A., & Soroka, A. M. (2022). *Nitrogen, phosphorus, and suspended-sediment loads and trends measured at the Chesapeake Bay River Input Monitoring stations: Water years 1985–2021*. U.S. Geological Survey data release. https://doi.org/10.5066/P90CZJ1Y

Meals, D. W., Dressing, S. A., & Davenport, T. E. (2010). Lag time in water quality response to best management practices: A review. *Journal of Environmental Quality*, 39(1), 85–96. https://doi.org/10.2134/jeq2009.0108

Meinikmann, K., Hupfer, M, & Lewandowski, J. (2015). Phosphorus in groundwater discharge—A potential source for lake eutrophication. *Journal of Hydrology*, 524, 214–226. https://doi.org/10.1016/j.jhydrol.2015.02.031

Michalak, A. M., Anderson, E. J., Beletsky, D., Boland, S., Bosch, N. S., Bridgemen, T. B., Chaffin, J. D., Cho, K., Confesor, R. Daloğlu, I., DePinto, J. V., Evans, M. A., Fahnenstiel, G. L., He, L., Ho, J. C., Jenkins, L., Johengen, T. H., Kuo, K. C., LaPorte, E., … Zagorski, M. A. (2013). Record-setting algal bloom in Lake Erie caused by agricultural and meteorological trends consistent with expected future conditions. *Proceedings of the National Academy of Sciences*, 110(16), 6448-6452. https://doi.org/10.1073/pnas.1216006110

Miller, A., Baker, M., Boomer, K., Merritts, D., Prestegaard, K., & Smith, S. (2019). *Legacy sediment, riparian corridors, and total maximum daily loads*. STAC Publication Number 19-001, Chesapeake Bay Program Scientific and Technical Advisory Committee (STAC), Edgewater, MD. 64 pp. http://www.chesapeake.org/pubs/399_Miller2019.pdf

Monaco, M. E., & Ulanowicz, R. E. (1997). Comparative ecosystem trophic structure of three U.S. mid-Atlantic estuaries. *Marine Ecology Progress Series*, 161, 239–254. https://doi.org/10.3354/meps161239

Monaco, M. E., Weisberg, S. B., & Lowery, T. A. (1998). Summer habitat affinities of estuarine fish in U.S. mid-Atlantic coastal systems. *Fisheries Management and Ecology*, 5, 161–171.

Moyer, D. L., & Blomquist, J. D. (2017). *Summary of nitrogen, phosphorus, suspended-sediment loads and trends measured at the Chesapeake Bay Nontidal Network stations: Water year 2016 update*. https://www.usgs.gov/media/files/summary-nitrogen-phosphorus-and-suspended-sediment-loads

Murphy, R. R., Keisman, J., Harcum, J., Karrh, R. R., Lane, M., Perry, E. S., & and Zhang, Q. (2022). Nutrient improvements in Chesapeake Bay: Direct effect of load reductions and implications for coastal management. *Environmental Science & Technology*, 56(1), 260–270. https://doi.org/10.1021/acs.est.1c05388

Najjar, R. G., Pyke, C. R., Adams, M. B., Breitburg, D., Hershner, C., Kemp, M., Howarth, R., Mulholland, M. R., Paolisso, M., Secor, D., Sellner, K., Wardrop, D., & Wood, R. (2010). Potential climate-change impacts on the Chesapeake Bay. *Estuarine, Coastal and Shelf Science*, 86(1), 1–20. https://doi.org/10.1016/j.ecss.2009.09.026

National Research Council (NRC). (2001). *Assessing the TMDL approach to water quality management*. National Academies Press.

National Research Council (NRC). (2011). *Achieving nutrient and sediment reduction goals in the Chesapeake Bay: An evaluation of program strategies and implementation*. National Academies Press.

Newell, R. I. E. (1988). Ecological changes in Chesapeake Bay: Are they the result of overharvesting the American oyster, Crassostrea virginica." In M. P. Lynch & E. C. Krome (Eds.), *Understanding the estuary: Advances in Chesapeake Bay research* (pp. 536–546). Chesapeake Research Consortium Publication 129, CBP/TRS 24/88. http://www.oyster-restoration.org/wp-content/uploads/2012/06/Newell-1988-filtering.pdf

Ni, W., Li, M., & Testa, J. M. (2020). Discerning effects of warming, sea level rise and nutrient management on long-term hypoxia trends in Chesapeake Bay. *Science of The Total Environment*, 737, 139717. https://doi.org/10.1016/j.scitotenv.2020.139717

Niklitschek, E. J., & Secor, D. H. (2005). Modeling spatial and temporal variation of suitable nursery habitats for Atlantic sturgeon in the Chesapeake Bay. *Estuarine, Coastal and Shelf Science*, 64(1), 135–148. https://doi.org/10.1016/j.ecss.2005.02.012

Noe, G. B., Cashman, M., Skalak, K., Gellis, A., Hopkins, K., Moyer, D., Webber, J., Benthem, A., Maloney, K., Brakebill, J., Sekellick, A., Langland, M., Zhang, Q., Shenk, G., Keisman, J., & Hupp, C. (2020). Sediment dynamics and implications for management: State of the science from long-term research in the Chesapeake Bay watershed, USA. *Wiley Interdisciplinary Reviews: Water*, 7(4), e1454. https://doi.org/10.1002/wat2.1454

Noe, G. B., Hopkins, K. G., Claggett, P. R., Schenk, E. R., Metes, M. J., Ahmed, L., Doody, T. R., & Hupp, C. R. (2022). Streambank and floodplain geomorphic change and contribution to watershed material budgets. *Environmental Research Letters*, 17, 064015. https://doi.org/10.1088/1748-9326/ac6e47

Opalinski, N., Schultz, D., Veith, T. L., Royer, M., & Preisendanz, H. E. (2022). Meeting the moment: Leveraging temporal inequality for temporal targeting to achieve water-quality load-reduction goals. *Water*, 14(7), 1003. https://doi.org/10.3390/w14071003

Osmond, D. L., Hoag, D. L. K., Luloff, A. E., Meals, D. W., & Neas, K. (2015). Farmers' use of nutrient management: Lessons from watershed case studies. *Journal of Environmental Quality*, 44(2), 382–390. https://doi.org/10.2134/jeq2014.02.0091

Osmond, D., Meals, D., Hoag, D., Arabi, M., Luloff, A., Jennings, G., McFarland, M., Spooner, J., Sharpley, A., & Line, D. (2012). Improving conservation practices programming to protect water quality in agricultural watersheds: Lessons learned from the National Institute of Food and Agriculture—Conservation Effects Assessment Project. *Journal of Soil and Water Conservation*, 67(5), 122A–127A. https://doi.org/10.2489/jswc.67.5.122a

O'Sullivan, A. (2021). *Expanding regulatory sandboxing to fast-track innovation*. The James Madison Institute. https://jamesmadison.org/wp-content/uploads/2021/01/Regulatory_Sandbox_1.27.21.pdf

Palm-Forster, L. H., Swinton, S. M., Lupi, F., & Shupp, R. S. (2016). Too burdensome to bid: Transaction costs and pay-for-performance conservation. *American Journal of Agricultural Economics*, 98(5), 1314–1333. https://doi.org/10.1093/ajae/aaw071

Pannell, D. J., & Claassen, R. (2020). The roles of adoption and behavior change in agricultural policy. *Applied Economic Perspectives and Policy*, 42(1), 31–41. https://doi.org/10.1002/aepp.13009

Parker, M., & Bricker, S. (2020). Sustainable oyster aquaculture, water quality improvement, and ecosystem service value potential in Maryland Chesapeake Bay. *Journal of Shellfish Research*, 39(2), 269–281. https://doi.org/10.2983/035.039.0208

Patrick, C. J., Wilcox, D. J., Whiting, J. R., Kenne, A. K., & Smith, E. R. (2022). *2021 distribution of submerged aquatic vegetation in Chesapeake Bay and coastal bays*. Virginia Institute of Marine Science. https://www.vims.edu/research/units/programs/sav/reports/2021/index.php

Patterson, J. J., Smith, C., & Bellamy, J. (2013). Understanding enabling capacities for managing the ‘wicked problem’ of nonpoint source water pollution in catchments: A conceptual framework. *Journal of Environmental Management*, 128, 441–452. https://doi.org/10.1016/j.jenvman.2013.05.033

Pearce, A., & Mcguire, R. (2020). The state of phosphorus balance on 58 Virginia dairy farms. *Journal of Environmental Quality*, 49(2), 324–334. https://doi.org/10.1002/jeq2.20054

Phillips, S. W., & Lindsey, B. D. (2003). *The influence of ground water on nitrogen delivery to the Chesapeake Bay*. U.S. Geological Survey FS-091-03. https://pubs.usgs.gov/fs/2003/0091/report.pdf

Pineiro, V., Arias, J., Elverdin, P., Ibanez, A. M., Opazo, C. M., Prager, S., & Terero, M. (2021). *Achieving sustainable agricultural practices: From incentives to adoption and outcomes*. International Food Policy Research Institute Policy Brief. https://cdm15738.contentdm.oclc.org/utils/getfile/collection/p15738coll2/id/134262/filename/134472.pdf

Pizzuto, J., Schenk, E. R., Hupp, C. R., Gellis, A., Noe, G., Williamson, E., Karwan, D. L., O’Neal, M., Maquard, J., Aalto, R, & Newbold, D. (2014). Characteristic length scales and time-averaged transport velocities of suspended sediment in the mid-Atlantic region, USA. *Water Resources Research*, 50(2), 790–805. https://doi.org/10.1002/2013WR014485

Price, E., Flemming, T. H., & Wainger, L. (2021). *Cost analysis of stormwater and agricultural practices for reducing nitrogen and phosphorus runoff in Maryland*. University of Maryland Center for Environmental Science Technical Report #TS-772-21. https://doi.org/10.13140/RG.2.2.28896.74246/1

Prokopy, L. S., Floress, K., Arbuckle, J. G., Church, S. P., Eanes, F. R., Gao, Y., Gramig, B. M., Ranjan, P., & Singh, A. S. (2019). Adoption of agricultural conservation practices in the United States: Evidence from 35 years of quantitative literature. *Journal of Soil and Water Conservation*, 74(5), 520–534. https://doi.org/10.2489/jswc.74.5.520

Qui, Z. (2009). Assessing critical source areas in watersheds for conservation buffer planning and riparian restoration. *Environmental Management*, 44, 968–980. https://doi.org/10.1007/s00267-009-9380-y

Ranjan, P., Church, S. P., Floress, K., & Prokopy, L. S. (2019). Synthesizing conservation motivations and barriers: What have we learned from qualitative studies of farmers’ behaviors in the United States? *Society and Natural Resources*, 32(11), 1171–1199. https://doi.org/10.1080/08941920.2019.1648710

Rao, N. S., Easton, Z. M., Schneiderman, E. M., Zion, M. S., Lee, D. R., & Steenhuis, T.S. (2009). Modeling watershed-scale effectiveness of agricultural best management practices to reduce phosphorus loading. *Journal of Environmental Management*, 90(3), 1385–1395. https://doi.org/10.1016/j.jenvman.2008.08.011

Reimer, A. P., & Prokopy, L. S. (2014). Farmer participation in U.S farm bill conservation programs. *Environmental Management*, 53(2), 318–332. https://doi.org/10.1007/s00267-013-0184-8

Reimer, A., Thompson, A., Prokopy, L. S., Arbuckle, J. G., Genskow, K., Jackson-Smith, D., Lynne, G., McCann, L., Morton, L. W., & Nowak, P. (2014). People, place, behavior, and context: A research agenda for expanding our understanding of what motivates farmers'

conservation behaviors. *Journal of Soil and Water Conservation*, 69(2), 57A–61A. https://doi.org/10.2489/jswc.69.2.57A
Ribaudo, M. (2001). Non-point source pollution control policy in the USA. In J.S. Shortle & D. Abler (Eds.), *Environmental policies for agricultural pollution control* (pp. 123–150). CABI Publishing.
Ribaudo, M. (2015). The limits of voluntary conservation programs. *Choices*, 30(2), 1–5.
Ribaudo, M., & Shortle, J. (2019). Reflections on 40 years of applied economics research on agriculture and water quality. *Agricultural and Resource Economics Review*, 48(3), 519–530. https://doi.org/10.1017/age.2019.32
Roland, V. L., Garcia, A. M., Saad, D. A., Ator, S. W., Robertson, D., & Schwarz, G. (2022). Quantifying regional effects of best management practices on nutrient losses from agricultural lands. *Journal of Soil and Water Conservation*, 77(1), 15–29. https://doi.org/10.2489/jswc.2022.00162
Rose, K., Monaco, M. E., Ihde, T., Hubbart, J., Smith, E., Stauffer, J., & Havens, K. J. (2023). *Proposed framework for analyzing water quality and habitat effects on the living resources of Chesapeake Bay*. STAC Publication Number 23-005, Chesapeake Bay Program Scientific and Technical Advisory Committee (STAC), Edgewater, MD. XX pp.
Rose, K. A., Sable, S., DeAngelis, D. L., Yurek, S., Trexler, J. C., Graf, W., & Reed, D. J. (2015). Proposed best modeling practices for assessing the effects of ecosystem restoration on fish. *Ecological Modelling*, 300, 12–29. https://doi.org/10.1016/j.ecolmodel.2014.12.020
Royer, M., Shortle, J., & Cook, A. (2016). *An analysis of the Pennsylvania farm conservation practices inventory for purposes of reporting practices to the Chesapeake Bay Program*. Penn State University College of Agricultural Sciences. https://files.dep.state.pa.us/water/ChesapeakeBayOffice/Farm%20Survey%20Report%20Final%20121516.pdf
Runge, M. C., Converse, S. J., & Lyons, J. E. (2011). Which uncertainty? Using expert elicitation and expected value of information to design an adaptive program. *Biological Conservation*, 144(4), 1214–1223. https://doi.org/10.1016/j.biocon.2010.12.020
Runge, M. C., Converse, S. J., Lyons, J. E., & Smith, D. R. (Eds.). (2020). *Structured decision making: Case studies in natural resource management*. Johns Hopkins Press.
Ryberg, K. R., Blomquist, J. D., Sprague, L. A., Sekellick, A. J., & Keisman, J. (2018). Modeling drivers of phosphorus loads in Chesapeake Bay tributaries and inferences about long-term change. *Science of The Total Environment*, 616-617, 1423–1430. https://doi.org/10.1016/j.scitotenv.2017.10.173
Sabo, R. D., Sullivan, B., Wu, C., Trentacoste, E., Zhang, Q., Shenk, G. W., Bhatt, G., & Linker, L. C. (2022). Major point and nonpoint sources of nutrient pollution to surface water have declined throughout the Chesapeake Bay watershed. *Environmental Research Communications*, 4(4), 045012. https://doi.org/10.1088/2515-7620/ac5db6
Saha, A., Saha, G. K., Cibin, R., Spiegal, S., Kleinman, P. J. A., Veith, T. L., White, C. M., Drohan, P. J., & Tsegaye, T. (2022). Evaluating water quality benefits of manureshed management in the Susquehanna River Basin. *Journal of Environmental Quality*, forthcoming. https://doi.org/10.1002/jeq2.20429

Sanford, W. E., & Pope, J. P. (2013). Quantifying groundwater's role in delaying improvements to Chesapeake Bay water quality. *Environmental Science & Technology*, 47(23), 13330–13338. https://doi.org/10.1021/es401334k

Schlenger, A., North, E. W., Li, Y., Li, M., & Kemp, W. M. (2022). Applying a three-dimensional habitat volume model to estimate sensitivity of Chesapeake Bay living resources to environmental change: A proof-of-concept exercise. *Estuaries and Coasts*, 45, 393–412. https://doi.org/10.1007/s12237-021-00967-6

Scully, M. E. (2010). The importance of climate variability to wind-driven modulation of hypoxia in Chesapeake Bay. *Journal of Physical Oceanography*, 40(6), 1435–1440. https://doi.org/10.1175/2010JPO4321.1

Sebilo, M., Mayer, B., Nicolardot, B., Pinay, G., & Mariotti, A. (2013). Long-term fate of nitrate fertilizer in agricultural soils. *Proceedings of the National Academy of Sciences*, 110(45), 18185–18189. https://doi.org/10.1073/pnas.1305372110

Selbig, W. R., & Bannerman, R. T. (2008). *A comparison of runoff quantity and quality from two small basins undergoing implementation of conventional and low-impact-development (LID) strategies: Cross Plains, Wisconsin, water years 1999–2005*. U.S. Geological Survey Scientific Investigations Report 2008–5008. https://doi.org/10.3133/sir20085008

Shabman, L., Reckhow, K., Beck, M. B., Benaman, J., Shapra, S., Freedman, P., Nellor, M., Rudek, J., Schwer, D., Stiles, T., & Stow, C. (2007). *Adaptive implementation of water quality improvement plans: Opportunities and challenges*. Nicholas Institute for Environmental Policy Solutions Report #NI-R-07-03, Duke University.

Sharpley, A., Jarvie, H. P., Buda, A., May, L., Spears, B., & Kleinman, P. (2013). Phosphorus legacy: Overcoming the effects of past management practices to mitigate future water quality impairment. *Journal of Environmental Quality*, 42(5), 1308–1326. https://doi.org/10.2134/jeq2013.03.0098

Shenk, G., Zhang, Q., Bhatt, G., Bertani, I., Mason, C., Moyer, D., & Tango, P. (2022). *Proposed integrated watershed TMDL indicator and dashboard products*. Presentation to Chesapeake Bay Program Water Quality Goal Implementation Team, November 14. https://d18lev1ok5leia.cloudfront.net/chesapeakebay/documents/2022-11-14-Monitored-TMDL-Indicator-gshenk-WQGIT.pdf

Shenk, G., Blatt, G., Tian R., Cerco, C. F., Bertani, I., & Linker, L. C. (2021). *Modeling climate change effects on Chesapeake Bay water quality standards and development of 2025 planning targets to address climate change*. CBPO Publication Number 328-21. https://cast-content.chesapeakebay.net/documents/P6ModelDocumentation%2FClimateChangeDocumentation.pdf

Shortle, J. S., & Horan, R. D. (2017). Nutrient pollution: A wicked challenge for economic instruments. *Water Economics and Policy*, 3(2), 1650033. https://doi.org/10.1142/S2382624X16500338

Shortle, J., Ollikainen, M., & Iho, A. (2021). *Water quality and agriculture: Economics and policy for nonpoint source water pollution*. Palgrave MacMillan.

Shortle, J. S., Ribaudo, M., Horan, R. D., & Blandford, D. (2012). Reforming agricultural nonpoint pollution policy in an increasingly budget-constrained environment. *Environmental Science & Technology*, 46(3), 1316–1325. https://doi.org/10.1021/es2020499

Sinha, E., Michalak, A. M., & Balaji, V. (2017). Eutrophication will increase during the 21st century as a result of precipitation changes. *Science*, 357(6349), 405–408. https://doi.org/10.1126/science.aan2409

Smith, D. R., Macrae, M. L., Kleinman, P. J. A., Jarvie, H. P., King, K. W., & Bryant, R. B. (2019). The latitudes, attitudes, and platitudes of watershed phosphorus management in North America. *Journal of Environmental Quality*, 48(5), 1176–1190. https://doi.org/10.2134/jeq2019.03.0136

Smith, S. M. C., & Wilcock, P. R. (2015). Upland sediment supply and its relation to watershed sediment delivery in the contemporary mid-Atlantic Piedmont (U.S.A). *Geomorphology*, 232, 33–46. https://doi.org/10.1016/j.geomorph.2014.12.036

Spiegal, S., Kleinman, P. J. A., Endale, D. M., Bryant, R. B., Dell, C., Goslee, S., Meinen, R. J., Flynn, K. C., Baker, J. M., Browning, D. M., McCarty, G., Bittman, S., Carter, J., Cavigelli, M., Duncan, E., Gowda, P., Li, X., Ponce-Campos, G. E., Cibin, R., Silveira, M. L., … Yang Q. (2020). Manuresheds: Advancing nutrient recycling in US agriculture. *Agricultural Systems*, 182, 102813. https://doi.org/10.1016/j.agsy.2020.102813

Sprague, L. A., & Gronberg, J. A. M. (2012). Relating management practices and nutrient export in agricultural watersheds of the United States. *Journal of Environmental Quality*, 41(6), 1939–1950. https://doi.org/10.2134/jeq2012.0073

STAC (Chesapeake Bay Program Scientific and Technical Advisory Committee). (2013). *Incorporating lag-times into the Chesapeake Bay Program*. STAC Publication Number 13-004, Chesapeake Bay Program Scientific and Technical Advisory Committee (STAC), Edgewater, MD. 66 pp. https://www.chesapeake.org/pubs/305_Hirsch2013.pdf

Stackpoole, S. M., Stets, E. G., & Sprague, L. A. (2019). Variable impacts of contemporary versus legacy agricultural phosphorus on US river water quality. *Proceedings of the National Academy of Sciences*, 116(41), 20562–20567. https://doi.org/10.1073/pnas.1903226116

Statham, P. J. (2012). Nutrients in estuaries—An overview and the potential impacts of climate change. *Science of The total environment*, 434, 213–227. https://doi.org/10.1016/j.scitotenv.2011.09.088

Staver, K., Kleinman, P., Ator, S., Buda, A., Ketterings, Q., Sims, J. T., & Meisinger, J. (2014). *A review of agricultural P-dynamics in the Chesapeake Bay Watershed Model*. STAC Publication Number 14-005, Chesapeake Bay Program Scientific and Technical Advisory Committee (STAC), Edgewater, MD. 20 pp. https://www.chesapeake.org/stac/wp-content/uploads/2021/01/Final-STAC-Phosphorus-Review-Report-8.26.14.pdf

Staver, K. W. and Brinsfield, R. B. (1998). Using cereal grain winter cover crops to reduce groundwater nitrate contamination in the mid-Atlantic coastal plain. *Journal of Soil and Water Conservation*, *53*(3), 230-240.

Stephenson, K., Hershner, C., Benham, B., Easton, Z., Hanson, J., Julius, S., & Hinrichs, E. (2018). *Consideration of BMP performance uncertainty in Chesapeake Bay program implementation*. STAC Publication Number 18-003, Chesapeake Bay Program Scientific and Technical Advisory Committee (STAC), Edgewater, MD. 33 pp. http://www.chesapeake.org/pubs/390_Stephenson2018.pdf

Stephenson, K., & Shabman, L. (2017a). Nutrient assimilation services for water quality trading programs: A comparative analysis with nonpoint source credits. *Coastal Management*, 45(1), 24–43. https://doi.org/10.1080/08920753.2017.1237240

Stephenson, K., & Shabman, L. (2017b). Where did the agricultural nonpoint source trades go? Lessons from Virginia water quality trading programs. *JAWRA Journal of the American Water Resources Association*, 53(5), 1178–1194. https://doi-org.proxy2.cl.msu.edu/10.1111/1752-1688.12565

Stephenson, K., Shabman, L., Shortle, J., & Easton, Z. (2022). Confronting our agricultural nonpoint source control policy problem. *JAWRA Journal of the American Water Resources Association*, 58(4), 496–501. https://doi.org/10.1111/1752-1688.13010

Talberth, J., Selman, M., Walker, S., & Gray, E. (2015). Pay for performance: Optimizing public investments in agricultural best management practices in the Chesapeake Bay watershed. *Ecological Economics*, 118, 252–261. https://doi.org/10.1016/j.ecolecon.2015.07.033

Tango, P. J., & Batiuk, R. A. (2013). Deriving Chesapeake Bay water quality standards. *JAWRA Journal of the American Water Resources Association*, 49(5), 1007–1024. https://doi.org/10.1111/jawr.12108

Testa, J. M., Boynton, W. R., Hodgkins, C. L., Moore, A. L., Bailey, E. M., & Rambo, J. (2022). Biogeochemical states, rates, and exchanges exhibit linear responses to large nutrient load reductions in a shallow, eutrophic urban estuary. *Limnology and Oceanography*, *67*(4), 739–752. https://doi.org/10.1002/lno.12037

Testa, J. M., Dennison, W. C., Ball, W. P., Boomer, K., Gibson, D. M., Linker, L., Runge, M. C., & Sanford, L. (2023). *Knowledge gaps, uncertainties, and opportunities regarding the response of the Chesapeake Bay estuary to proposed TMDLs*. STAC Publication Number 23-004, Chesapeake Bay Program Scientific and Technical Advisory Committee (STAC), Edgewater, MD. XX pp.

Testa, J .M., Harris, L, Boynton, W. R, Bukaveckas, P., Eshleman, K., Murphy, R., Keisman, J., Blomquist, J., Lane, M., Karrh, R., & Lyubchich, V. (2018). *Progress toward the restoration of Chesapeake Bay in time and space*. Chesapeake Biological Laboratory Technical Series TS-716-18.

Testa, J. M., & Kemp, W. M. (2012). Hypoxia-induced shifts in nitrogen and phosphorus cycling in Chesapeake Bay. *Limnology and Oceanography*, 57(3), 835–850. https://doi.org/10.4319/lo.2012.57.3.0835

Testa, J. M., Lyubchich, V., & Zhang, Q. (2019). Patterns and trends in Secchi disk depth over three decades in the Chesapeake Bay estuarine complex. *Estuaries and Coasts*, 42, 927–943. https://doi.org/10.1007/s12237-019-00547-9

Thomason, W., Duiker, S., Ganoe, K., Gates, D., McCollum, B., & Reiter, M. (2016). *Conservation tillage practices for use in phase 6.0 of the Chesapeake Bay Program Watershed Model*. CBP/TRS-308-16. https://d18lev1ok5leia.cloudfront.net/chesapeakebay/documents/CT_6.0_Conservation_Tillage_EP_Revised_Full_Report_12-14-16.2_FINAL_NEW_TEMPLATE.pdf

Tomer, M. D., & Locke, M. A. (2011). The challenge of documenting water quality benefits of conservation practices: A review of USDA-ARS's conservation effects assessment project watershed studies. *Water Science & Technology*, 64(1), 300–310. https://doi.org/10.2166/wst.2011.555

Townsend, H. (2014). Comparing and coupling a water quality and a fisheries ecosystem model of the Chesapeake Bay for the exploratory assessment of resource management strategies. *ICES Journal of Marine Science*, 71(3), 703–712. https://doi.org/10.1093/icesjms/fst060

U.S. Environmental Protection Agency (USEPA). (2003a). *Ambient water quality criteria for dissolved oxygen, water clarity and chlorophyll-a for the Chesapeake Bay and its tidal tributaries*. EPA 903-R-03-002. https://nepis.epa.gov/Exe/ZyPDF.cgi/P100YKPQ.PDF?Dockey=P100YKPQ.PDF

U.S. Environmental Protection Agency (USEPA). (2003b). *Technical support document for identification of Chesapeake Bay designated uses and attainability*. EPA 903-R-03-004. https://nepis.epa.gov/Exe/ZyPURL.cgi?Dockey=P1001WE5.txt

U.S. Environmental Protection Agency (USEPA). (2004). *Ambient water quality criteria for dissolved oxygen, water clarity and chlorophyll-a for the Chesapeake Bay and its tidal tributaries: 2004 addendum*. EPA 903-R-04-005. https://nepis.epa.gov/Exe/ZyPURL.cgi?Dockey=P1001WB8.txt

U.S. Environmental Protection Agency (USEPA). (2010). *Chesapeake Bay total maximum daily load for nitrogen, phosphorus and sediment*. https://www.epa.gov/chesapeake-bay-tmdl/chesapeake-bay-tmdl-document

U.S. Environmental Protection Agency (USEPA). (2017). A*mbient water quality criteria for dissolved oxygen, water clarity and chlorophyll-a for the Chesapeake Bay and its tidal tributaries: 2017 technical addendum*. EPA 903-R-17-002. https://nepis.epa.gov/Exe/ZyPURL.cgi?Dockey=P100YL1C.txt

Van Meter, K. J., Basu, N. B., & Van Cappellen, P. (2017). Two centuries of nitrogen dynamics: Legacy sources and sinks in the Mississippi and Susquehanna River Basins. *Global Biogeochemical Cycles*, 31(1), 2–23. https://doi.org/10.1002/2016GB005498

Van Meter, K. J., Basu, N. B., Veenstra, J. J., & Burras, C. L. (2016). The nitrogen legacy: Emerging evidence of nitrogen accumulation in anthropogenic landscapes. *Environmental Research Letters*, 11(3), 035014. https://doi.org/10.1088/1748-9326/11/3/035014

Van Meter, K., McLeod, M. M., Liu, J., Tenkouano, G. T., Hall, R. I, Van Cappellen, P., & Basu, N. B. (2021). Beyond the mass balance: Watershed phosphorus legacies and the evolution of the current water quality policy challenge. *Water Resources Research*, 57(10), e2020WR029316. https://doi.org/10.1029/2020WR029316

Vasslides, J. M., de Mutsert, K., Christensen, V., & Townsend, H. (2017). Using the Ecopath with Ecosim modeling approach to understand the effects of watershed-based management actions in coastal ecosystems. *Coastal Management*, 45(1), 44–55. https://doi.org/10.1080/08920753.2017.1237241

Virginia Institute of Marine Science (VIMS). (n.d.-a). *Indices of abundance*. https://www.vims.edu/research/departments/fisheries/programs/juvenile_surveys/data_products/indices/index.php

Virginia Institute of Marine Science (VIMS). (n.d.-b). *SAV monitoring and restoration*. https://www.vims.edu/research/units/programs/sav/index.php

Wagena, M. B., & Easton, Z. M. (2018). Agricultural conservation practices can help mitigate the impact of climate change. *Science of The Total Environment*, 635, 132–143. https://doi.org/10.1016/j.scitotenv.2018.04.110.

Wang, P., Wang, H., Linker, L., & Tian, R. (2016). Effects of cross-channel bathymetry and wind direction on destratification and hypoxia reduction in the Chesapeake Bay. *Estuarine, Coastal and Shelf Science*, 178, 168–188. https://doi.org/10.1016/j.ecss.2016.05.009

Walter, R. C., & Merritts, D. J. (2008). Natural streams and the legacy of water-powered mills. *Science*, 319(5861), 299–304. https://doi.org/10.1126/science.1151716

Ward, N. D., Megonigal, J. P., Bond-Lamberty, B., Bailey, V. L., Butman, D., Canuel, E. A., Diefenderfer, H., Ganju, N. K., Goni, M. A., Graham, E. B., Hopkinson, C. S., Khangaonkar, T., Langley, J. A., McDowell, N. G., Myers-Pigg, A. N., Neumann, R. B., Osburn, C. L., Price, R. M., Rowland, J., … Windham-Myers, L. (2020). Representing the function and sensitivity of coastal interfaces in Earth system models. *Nature Communications*, 11, 2458. https://doi.org/10.1038/s41467-020-16236-2

White, M. J., Storm, D. E., Busteed, P. R., Stoodley, S. H., & Phillips, S. J. (2009). Evaluating nonpoint source critical source area contributions at the watershed scale. *Journal of Environmental Quality*, 38(4), 1654–1663. https://doi.org/10.2134/jeq2008.0375

Wiering, M., Liefferink, D., Boezeman, D., Kaufmann, M., Crabbé, A., & Kurstjens, N. (2020). The wicked problem the water framework directive cannot solve. The governance approach in dealing with pollution of nutrients in surface water in the Netherlands, Flanders, Lower Saxony, Denmark and Ireland. *Water*, 12, 1240. https://doi.org/10.3390/w12041240

Wilson, R. S., Howard, G., & Burnett, E. A. (2014). Improving nutrient management practices in agriculture: The role of risk-based beliefs in understanding farmers' attitudes toward taking additional action. *Water Resources Research*, 50(8), 6735–6746. https://doi.org/10.1002/2013WR015200

Woodland, R. J., Buchheister, A., Latour, R. J., Lozano, C., Houde, E., Sweetman, C. J., Fabrizio, M. C., & Tuckey, T. D. (2021). Environmental drivers of forage fishes and benthic invertebrates at multiple spatial scales in a large temperate estuary. *Estuaries and Coasts*, 44, 921–938. https://doi.org/10.1007/s12237-020-00835-9

Xu, Y., Bosch, D. J., Wagena, M. B., Collick, A. S., & Easton, Z. M. (2019). Meeting water quality goals by spatial targeting of best management practices under climate change. *Journal of Environmental Management*, 63, 173–184. https://doi.org/10.1007/s00267-018-01133-8

Yagow, G., Collick, A., Ribaudo, M., Thomason, W., & Veith, T. (2016). *Scientific and technical advisory committee review of nutrient input estimation for the Chesapeake Bay Watershed Model*. STAC Publication Number 16-005, Chesapeake Bay Program Scientific and Technical Advisory Committee (STAC), Edgewater, MD. 49 pp. http://www.chesapeake.org/pubs/358_Yagow2016.pdf

Yang W., & Weersink, A. (2005). Cost-effective targeting of riparian buffers. *Canadian Journal of Agricultural Economics*, 52(1), 17–34. https://doi.org/10.1111/j.1744-7976.2004.tb00092.x

Zhang, Q., Murphy, R. R., Tian, R., Forsyth, M. K., Trentacoste, E. M., Keisman, J, & Tango, P. J. (2018). Chesapeake Bay's water quality condition has been recovering: Insights from a multimetric indicator assessment of thirty years of tidal monitoring data. *Science of The Total Environment*, 637–638, 1617–1625. https://doi.org/10.1016/j.scitotenv.2018.05.025

Zhang, Q., Tango, P. J., Murphy, R. R., Forsyth, M. K., Tian, R., Keisman, J., & Trentacoste, E. M. (2018). Chesapeake Bay dissolved oxygen criterion attainment deficit: Three decades of temporal and spatial patterns. *Frontiers in Marine Science*, 5, 422. https://doi.org/10.3389/fmars.2018.00422