Recognizing political influences in participatory social-ecological systems modeling

Theodore C. Lim[1*], Pierre D. Glynn[2], Gary W. Shenk[3], Patrick Bitterman[4], Joseph H.A. Guillaume[5], John C. Little[6], and D.G. Webster[7]

[1] School of Public and International Affairs, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

[2] Arizona State University Consortium for Science, Policy and Outcomes, Tempe, AZ, USA and United States Geological Survey, Science and Decisions Center

[3] United States Geological Survey, Virginia/West Virginia Water Science Center, Richmond, VA, USA

[4] School of Global Integrative Studies, University of Nebraska, Lincoln, USA

[5] Institute for Water Futures and Fenner School of Environment & Society, Australian National University, Canberra, Australia

[6] Department of Civil and Environmental Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

[7] Department of Environmental Studies, Dartmouth College, Hanover, NH, USA

Abstract

Stakeholder participation in social-ecological systems (SES) modeling is increasingly considered a desirable way to elicit diverse sources of knowledge about SES behavior and to promote inclusive decision-making in SES. Understanding how participatory modeling processes function in the context of long-term adaptive management of SES may allow for better design of participatory processes to achieve the intended outcomes of inclusionary knowledge, representativeness, and social learning, while avoiding unintended outcomes. Long-term adaptive management contexts often include political influences -- attempts to shift or preserve power structures and authority, and efforts to represent the political and economic interests of stakeholders -- in the computer models that are used to shape policy making and implementation. In this research, we examine a period that included a major transition in the watershed model used for management of the Chesapeake Bay in the United States. The Chesapeake Bay watershed model has been in development since the 1980s, and is considered by many to be an exemplary case of participatory modeling. We use documentary analysis and interviews with participants involved in the model application and development transition to reveal a variety of ways in which participatory modeling may be subject to different kinds of political influences, some of which resulted in unintended outcomes, including: perceptions of difficulty updating the model in substantive ways, "gaming" of the model/participatory process by stakeholders, and increasing resistance against considering uncertainty in the system not captured by the model. This research suggests unintended or negative outcomes may be associated with both participatory decision-making and stakeholder learning even though they are so often touted as the benefits of participatory modeling. We end with a hypothesis that further development of a theory of computer model governance to bridge model impact and broader theories of environmental governance at the science-policy interface may result in improved SES modeling outcomes.

Keywords

participatory modeling; science policy; evidence-based policy; boundary objects; watershed management

Correspondence:
Contact T. Lim at tclim@vt.edu

Cite this article as:
Lim, T.C., Glynn, P.D., Shenk, G.W., Bitterman, P., Guillaume, J.H.A., Little, J.C., & Webster, D.G.
Recognizing political influences in participatory social-ecological systems modeling
Socio-Environmental Systems Modelling, vol. 5, 18509, 2023, doi:10.18174/sesmo.18509

Socio-Environmental Systems Modelling
An Open-Access Scholarly Journal
http://www.sesmo.org

Code and Data availability
All primary meeting documents (minutes, presentations, and reports) are available for download through the Chesapeake Bay Program website. Code to automatically download all documents and count word frequencies (as used in the preliminary stages of this research) are available on GitHub: https://github.com/theochli/Chesapeake-Bay-Program-minutes-scrape.

1. Introduction

Stakeholder participation in social-ecological systems (SES)[1] modeling is increasingly considered a desirable way to elicit diverse sources of knowledge about social-ecological system behavior and to promote inclusive decision-making. In the context of social-ecological systems management, it is not only important that models are realistic and scientifically sound; users of the model need to *perceive* models as salient, credible, and legitimate (Cash et al., 2003). A growing literature on "participatory modeling" reflects the normative idea that stakeholders should be involved in the development of the model itself, especially if models are to play critical roles in subsequent decision- and policy-making (Gray et al., 2018; Voinov and Bousquet, 2010). Motivations for doing so include: that participation in the model's development reflects more inclusionary, representative, and therefore democratic, decision-making (Falconi and Palmer, 2017); that with stakeholder participation, important local sources of knowledge will improve the model's representation of reality (Corburn, 2009); and that stakeholder involvement in the modeling process leads to greater buy-in to subsequent policy based on the model's outputs (Gray et al., 2018).

Participatory modeling, defined as a "purposeful learning process for action that engages the implicit and explicit knowledge of stakeholders to create formalized and shared representation(s) of reality", is an umbrella term[2] that spans methods ranging from qualitative approaches to problem definition -- such as cognitive mapping and causal loop diagrams-- to numerical quantitative modeling -- such as systems dynamics models and agent based modeling (Voinov et al., 2018). Participatory modeling methods have been applied to SES models in a variety of contexts, scales, and geographies (Hedelin et al., 2021). However, a review of research on participatory modeling has shown that many applications of participatory modeling methods tend to be in the problem formulation phase, academically-led, and relatively short-lived (Hedelin et al., 2021). Therefore, these applications of participatory modeling do not embody the same kinds of organizational and power-relations factors that affect models used for policy-making and long-term environmental management. In these contexts, participation in modeling may also have unintended negative consequences, including: the risk of delegitimization of models, risk of over-legitimization of models, and misrepresentation of consensus through models (Korfmacher, 2001). These undesired consequences are well-theorized and empirically demonstrated in the literature around the science-policy interface, but a gap persists in connecting the normative motivations of participatory modeling with its empirical evaluation in real-life policy and management contexts.

In this paper, we define "political influences" as the value-laden actions where multiple interests and objectives use power relationships to come to collective decisions. Politics are inextricably related to SES policy-making and have been extensively researched from many theoretical angles -- from discourses and values, to organizational structures and networks, to systems processes -- to name a few (Cash et al., 2006; Dryzek, 2013; Layzer, 2011; Meadowcroft, 2002; Wesselink et al., 2013; Wyborn et al., 2019). For this paper, the most important areas of this wide literature include: the incorporation of diverse sources of knowledge in collective decision-making (sometimes referred to as "co-production of knowledge") and the nature of the relationship between science and policy (often referred to as the "science-policy interface"). Our focus is specifically on collective decision making about computer models, which are increasingly used within policy contexts, and we therefore also draw on theories that explicate the roles and legitimacy that these objects enjoy.

The aim of this research is to present an empirical case in which participatory modeling, and its application to policy making, policy implementation, and environmental management have converged for the Chesapeake Bay Watershed in the eastern United States. In this case, more than three decades of continuous model development showed a trend of increasing physical process representation in the models (Linker et al., 2002;

[1] For an overview of social-ecological systems (SES) the reader is referred to Colding and Barthel (2019)
[2] Other related terms include: "companion modeling" and "modeling with stakeholders." Readers are referred to Voinov et al. (2016) for more detail.

Shenk and Linker, 2013). In late 2012 however, this trend reversed for the watershed model when the decision was made to change the model from a dynamic model to a static model. The events, meetings, and perspectives leading up to this seemingly abrupt change are the focus of this paper. We chose to focus on the period 2008 to 2012, hypothesizing that this change in approach would reveal important factors and themes based on the documents and reflections surrounding the science-policy interface from key informants. We seek to illustrate the kinds of political influence embedded in the modeling process, and therefore in a computational model itself, and hence raise awareness of potential unintended consequences of participation that may arise without more explicit governance of the computer model development process.

2. Literature Review

The models that are used in SES management are the products of scientific research, and often represent the application of scientific knowledge to specific contexts. They are therefore an important locus of the boundary constantly being negotiated between political interests and science. In this section, we review the literature related to the science-policy interface, and apply the literature to computer models of socio-ecological systems.

The academic discipline of science and technology studies has long framed the information and technology systems used in government and corporate organizations as a kind of "knowledge infrastructure" (Hewitt, 1988; Latour, 1987; Star, 1999), and computer models as "boundary objects" (Sundberg, 2007; White et al., 2010) — objects which "inhabit several intersecting social worlds… and satisfy the informational requirements of each of them." (Star and Griesemer, 1989, p. 393). To be an effective boundary object in environmental management contexts, model development is dependent on the contextual knowledge and habits of actors involved. The term "boundary object" is typically used in a positive sense: if computer models sufficiently capture the aspects of the social-ecological problem that are most salient to diverse stakeholders, then the boundary object can successfully facilitate cross-boundary communication and coordination toward decision-making. Participatory modeling engages the idea of a model as a boundary object that aids social learning across the diverse interests often involved in social-ecological problems (Voinov et al., 2018).

However, others have argued that just as computer models may act as boundary objects that facilitate communication and knowledge sharing, they may also become sites of contestation and objects that are used to leverage power. Stakeholders' environmental conflicts can become conflicts about the parameters and structures of computer models. For example, Sundberg showed how modelers and empiricists in climate science view the purpose of model parameterizations in fundamentally different ways, and that only when there is effective translation between their viewpoints is useful knowledge created (Sundberg, 2007). Another example showed how environmental modeling used in a policy context was an instance of science "going public" -- a description of when scientific knowledge leaves laboratories and becomes subject to political, social, and economic maneuvering (Latour, 1987). In the case of the Chesapeake Bay modeling systems, examined in this research, "going public" ignited challenges to the credibility and legitimacy of the computer models themselves, complete with a proposal for an alternative model whose outputs better represented certain stakeholders' interests as the basis for decision-making (Paolisso et al., 2015). However, more understanding about the processes of how interests and power are incorporated into computer models and what the effects of this socio-political process might be, could allow for more precise prescriptions for how computer models *should* go public, and what policy supports may be needed in order to achieve the intended outcomes of going public.

The mutual and bidirectional influences between science and policy are referred to as the "science-policy interface." And, while the relationship between science and policy is often idealized as the linear transmission of scientific knowledge into "objective" or "data-driven" policy, empirical evidence rarely supports this ideal (van den Hove, 2007; Wesselink et al., 2013). In fact, some criticize the term "evidence-based policy" as being a naive, techno-rational imaginary, itself unsubstantiated by empirical evidence (French, 2019; Funtowicz and Ravetz, 1993). The desire to have clear boundaries separating science from policy so as to not delegitimize scientific facts, evidence, and methods, is referred to as the "demarcation model." (Glynn et al., 2017; Liberatore and Funtowicz, 2003) However this ideal is often criticized as being based on the "twin myth of rationality": (1) that policy is based on an accumulation of scientific facts and uncertainty reduction, and (2) that science can provide objective facts to adjudicate controversies (Saltelli and Giampietro, 2017). Critics of the demarcation model contend that in reality, instead of evidence-based policy, the science-policy interface often more resembles "policy-based evidence," in which policy goals are substantiated and legitimized through methods or sources associated with science (Strassheim and Kettunen, 2014). According to this conceptualization, rather than science adjudicating policy, science may be leveraged, or even weaponized, by political factions seeking to bolster their positions, resulting in considerable erosion of public trust in science, but also in the models that are used to make very consequential public policy decisions (Sarewitz, 2004). In other cases, models may not be purposefully weaponized, yet still embody non-epistemic, subjective values, which may be either legitimate or illegitimate (Intemann, 2015; van der Sluijs et al., 2008). The idea of "post-normal science" reflects the impossibility of separating either "evidence-based policy" from "policy-based evidence", or "facts" from "values." (Funtowicz and Ravetz, 1993). Post-normal science frames science as just one component of a social process of collective problem-solving.

Computational models are used to numerically represent the causes and effects of SES challenges that science is called upon to provide solutions to, and are therefore ideal objects of study for understanding the science-policy interface. Past and current examples of computer models used to inform public policy have included: the impact of pesticides on bees, the impacts of carbon dioxide emissions on global climate change, the impact of endocrine disruptors on human health, and the benefits of shale gas fracking (Saltelli and Giampietro, 2017; Sarewitz, 2004). In the context of policy, models' ability to produce "crisp numbers" and the appearance of precision play useful social functions-- for example, projecting leadership and confidence in the short-term (Saltelli et al., 2015; van der Sluijs, 2005). However, these "crisp numbers" do not capture uncertainty of physical systems themselves, nor can they quantify the uncertainty that surrounds the policy implementation process itself (Saltelli and Funtowicz, 2014).

Protocols, such as the Structured Decision Making process, have been developed about how stakeholders should be involved in defining natural resource problems, management objectives, the kinds of information, data, and models that should be used, identification of tradeoffs, and how to make decisions and take action (Gregory et al., 2012; Martin et al., 2009). These protocols are based on explicitly and transparently acknowledging the role that subjective stakeholder values play in decision processes.

Biases, beliefs, heuristics, and values (BBHV), including social norms, condition all political influences, such as how prioritizations are made (both externally and internally) during participatory modeling processes and governance of SES (Glynn et al., 2017, 2018), and also in integrated modeling (Glynn, 2014; Glynn et al., 2017). They affect all: individuals (including scientists), groups, and institutions. Norms and values drive how political decisions are made: they tend to reflect the priorities, prevalent beliefs, and values of stakeholder communities and of their representatives (Akerlof, 1997; Elster, 1989). BBHV and human and social factors also affect how information is perceived, valued, communicated, and often manipulated (or gamed) during decision-making processes, and also in the construction and application of models (Glynn et al., 2022a, 2022b). Lastly, decision-making, participatory modeling processes, and model construction and applications rarely, if ever, fully acknowledge the sources of knowledge that they involve (often innately), or even more so the knowledge gaps, oversimplifications, and knowledge fragmentation that is more consciously being applied (Sterman, 2012).

Social psychologists have studied peoples' general desire to appear more knowledgeable than they are and a general unwillingness to admit incompetence. This condition has been well studied in social psychology (e.g. Kruger and Dunning, 1999; Wu and Dunning 2018). This is known as the "Dunning-Kruger" effect. This well-documented non-epistemic human and social behavioral bias may contribute to how computer models may be improperly used in political contexts, ultimately eroding public trust in science.

In the view of the authors of the present article, research to understand the nature and structure of participation in decision-making, including behavioral and social processes, and how computer models are used in relation to decision-making and societal engagement could improve environmental governance processes. Arnstein's ladder of participation (Arnstein, 1969) is a well-known conceptualization of participation, which positions "full citizen control" at the top of the ladder, followed by various degrees of stakeholder tokenism (e.g.: placation, consultation, informing), and non-participatory processes at the bottom of the ladder (e.g.: therapy, manipulation). The implication of the hierarchical rungs of the ladder is that full citizen control results in better outcomes. Other theories of participation focus more on the contextual and governance factors that may influence positive and negative outcomes of participatory processes (Fung, 2006; Newig and Fritsch, 2009). For example, Cooke et al. (2001) describe potential problems of participation as "tyrannies" that can result from (1) facilitators overriding legitimate concerns of the group ("tyranny of control"), (2) group dynamics that reinforce the positions of the already powerful ("tyranny of the group"), and (3) participatory methods that result in decision outcomes that may not be as good as if other methods were used ("tyranny of method") (Cooke et al., 2001). The role of group dynamics, group authority, and group inertia contribute to some of these "tyrannies" of participation (Everett et al., 2015; Lindner and Strulik, 2008). Groupthink (Janis, 1982) and social proofing (Cialdini and Goldstein, 2004) add more subtle reinforcements to the group dynamics in the establishment of what stands for accepted truths offered by the model.

As is shown from the above review of the literature, such political processes may be intentional attempts to embed subjective interests and objectives into the model (tyranny of control/ tyranny of the group); or they may be unintentional outcomes of the decision-making process (tyranny of the group/tyranny of method). We seek to understand how such political processes exhibit themselves in a participatory modeling process.

3. Methods

This research is an in-depth, qualitative, historical case study of the period between 2008 and 2012 in the Chesapeake Bay Program (CBP), the regional partnership charged with restoring water quality in the Chesapeake Bay (CBP, 2023d). This period was chosen because the research team expected to observe many decision-making processes and events leading up to the decisions to simplify the Chesapeake Bay watershed model into what became known as its Phase 6 form[3], also known as the Chesapeake Assessment Scenario Tool. Qualitative historical research relies on gathering a variety of primary and (if available) secondary sources that may provide information about the events during the period in question. These sources are then evaluated for authenticity and accuracy by the research team, and the researchers analyze the sequence of events to develop a narrative exposition of the findings (Berg and Lune, 2011, p. 311). The goal of qualitative historical research is to build an argument and tell a story. In our case, the contribution of the story of the 2008 - 2012 period in the Chesapeake Bay Program's watershed model's development was motivated by important present-day questions about the nature of participation in scientific and technical modeling, and more broadly, the nature of the interface between science and policy, which is often facilitated by computer models, as reviewed in the above sections.

Our qualitative historical research includes both primary and secondary sources. Our primary sources included: meeting minutes, reports, and documents publicly available on the Chesapeake Bay Program's website (CBP, 2023d), as well as interviews conducted with key informants who experienced or witnessed the events in this time period. Key informants included participants identified in meeting minutes, as well as others that were recommended by initial interviewees. Secondary sources include other histories that have been published summarizing the sequence of events during this period and interpretations of these events. Interviews with key informants were conducted under a protocol that received an exemption from Virginia Tech Institutional Review Board (Virginia Tech IRB #19-873). This protocol included the researchers' arrangement with the interviewees of confidentiality, and make it such that anonymized transcripts of the interviews cannot be made publicly available, as interviewees' identities would be easily discernible from the full transcripts. Because of this, interviews will not appear as bibliographic entries in this paper, but rather as footnotes.

We conducted content analysis of the above-described primary and secondary sources. The purpose of case study-based research is often different from the goals of generalizability associated with other research methods. First, we use in-depth case study to reveal evidence that builds "falsification," a form of critical reflexivity in social sciences that is critical to theory building (Ormerod, 2009). The value of "falsification" is illustrated by the notable example statement "all swans are white" to which it was proposed that an observation of just one black swan would fundamentally change the theory (Popper, 2002). By analogy, this research interrogates the maxim that "greater participation and social learning lead to better outcomes" with respect to computer model development in SES. Second, carefully chosen case studies are often used in the development of key scientific theories. For example, identification of "critical cases" -- extreme cases, or paradigmatic cases -- can provide information that is likely generalizable about intermediary cases (Flyvbjerg, 2006). As will be discussed in the following section, the CBP and its history and emphasis on voluntary, consensus-based, and participatory bottom-up decision-making qualified it as both an extreme and paradigmatic case, that is expected to benefit other cases that may not have such a history, but may intend to move toward incorporating more participation within the modeling process. Although a singular case study, this research aims to motivate and

[3] Major development phases of the watershed model were numbered sequentially starting in the early 1990s.

contribute to a theory of "computer model governance," which we address in the Discussion and Conclusion sections of the paper.

3.1 Case description

The Chesapeake Bay, located in the eastern United States, has a drainage area that includes six states and the District of Columbia (166,000 km^2, population living within watershed: 18M) (CBP, 2023c). The Chesapeake Bay estuary, which President Reagan called a "productive recreational area and a special national resource," (Reagan, 1984) has experienced harmful algal blooms and deterioration due to rapid increases in urbanization and intensification of agricultural activity in its watershed (Ernst, 2003; Layzer, 2011). The CBP was established in 1983 as a partnership between the six states and the District of Columbia, funded by federal dollars, to improve water quality. Governance in the Chesapeake Bay Program is based on consensus, with each of the seven jurisdictions (six states and DC) having political appointee representation as well as subject matter experts present on goal implementation teams and executive decision-making levels of the organization. The Chesapeake Bay Program's governance documents establish protocols for handling disagreements. Jurisdictions are directed to resolve conflicts through negotiation and collaboration to the best of their ability (CBP, 2021). Conflicts that cannot be resolved are escalated to higher levels of decision-making authority within the organization.

The collaborative and consensus-based nature of decision-making of the CBP, which also has a strong history of reliance on voluntary action, combined with a federal mandate enforced by the U.S. Environmental Protection Agency (henceforth "EPA") gives governance within the Chesapeake Bay Program a hybrid polycentric-hierarchical governance structure (Cumming, 2016). The hierarchical structure was strengthened in 2009 when President Obama issued an executive order (Executive Office of the President, 2009) for the federal government to renew the effort to protect and restore the watershed, and then again in 2010, with the establishment of a Total Maximum Daily Load (TMDL) by the EPA (U.S. EPA, 2010). The TMDL is a limit set by the federal government for how much pollutants can be discharged into the Chesapeake Bay.

Since its inception, the CBP has relied on a complex modeling system to guide its efforts. The CBP system includes four linked models--an estuary model, watershed model, airshed model, and a land use change model-- that are used to determine where pollutants are coming from and to estimate the impacts that each source has on pollutant loading to the Chesapeake Bay estuary (Hood et al., 2021; Shenk and Linker, 2013). In addition to assigning responsibility for the Chesapeake Bay's remediation, the CBP's models are used by the jurisdictions to plan scenarios to reduce pollutant loadings (CBP, 2014). For more detail on the technical aspects of the modeling system, the reader is referred to Hood et al. (2021) and Shenk and Linker (2013). In this article the modeling system is classified as an SES model. Examples of environmental processes represented include pollutant fate and transport and dispersion of nutrients in the Chesapeake Bay. Examples of social-ecological processes represented in the model include: the estimation of fertilizer application on land within the watershed, crop selection, and land use change-- for example, modeling of agricultural land converted to urban land.

The Chesapeake Bay watershed model has been through continuous rounds of development and application since the 1980s, and is considered by many to be an exemplary case of participatory modeling because of the input that the organization constantly elicits from jurisdictions on decisions regarding the development and improvement of the model (Hood et al., 2021). Which processes are represented in the Chesapeake Bay watershed and estuary models, at what spatial or temporal resolutions, and even the numerical values chosen for specific parameters potentially each have significant impacts on the costs of jurisdictions' mitigation responsibilities. Because statutory requirements, planning, and budgeting all hinge on the outputs of the model, it is critical both that the scientific basis of the model be sound, and that jurisdictions accept its legitimacy and its capability to represent their most significant concerns in the planning process.

There are several key groups within the organization that have had particular influence on model-related decisions. The Water Quality Goal Implementation Team (WQGIT) "works to evaluate, focus and accelerate the implementation of practices, policies and programs that will restore water quality in the Chesapeake Bay and its tributaries to conditions that support living resources and protect human health" (CBP, 2023b). WQGIT meetings often evaluate scientific information, data, and model-related decisions within the policy context, and therefore are often a key site exhibiting science-policy interface dynamics. The Modeling Workgroup (MODWG) "provide[s] state-of-the-art decision-support modeling tools that are built through community and participatory principles" (CBP, 2023a). Both the WQGIT and MODWG are composed of members from the seven jurisdictions, the federal government and other CBP partners. They are supported by staff from the CBP office that coordinate the group activities and work directly on the model. CBP office staff are employees of EPA, USGS, academic institutions working under EPA grants, and contractors, and work in collaboration with spatially distributed, federal and academic researchers and contractors. Lastly, the Scientific and Technical Advisory Committee (STAC) "enhances scientific communication and outreach throughout the Chesapeake Bay watershed and beyond" (Scientific and Technical Advisory Committee, 2023). STAC is composed of academic researchers and scientists and political appointees. STAC is responsible for conducting reviews of data, models, and scientific priorities of the CBP, and advises the governors/mayor of the jurisdictions.

3.2 Documentary analysis, timeline visualizations, and semi-structured interviews

Tracing processes supports understanding how factors such as internal and external political interests and power relations develop in relation to the science-policy interface. Process tracing uses multiple sources of information--for example documentary analysis and interviews-- to triangulate interpretation of events, as has been demonstrated effectively in related studies (Girod et al., 2009; Morrison, 2017).

We collected meeting minutes, presentations, and briefs associated with CBP meetings that occurred between January 2008 and December 2012. All documents were publicly accessible through the Chesapeake Bay Program's website (base URL: https://www.chesapeakebay.net). To get an initial overview of which meetings the model was being discussed and modeling decisions were being made, we wrote a Python (Python Software Foundation, 2023) script that scraped all meetings recorded on the CBP website's historical calendar between the study dates, opened all attachments on the webpage, and counted the incidence of the words "model" and "modeling" appearing in minutes, presentations, and briefs. We found that the majority of mentions of these words occurred during meetings of the WQGIT, STAC, and the MODWG. We then downloaded minutes documents associated with these groups' meetings during the study period. This resulted in a total of 106 meeting minutes that formed the primary basis of the documentary analysis. Later, this set of documents was supplemented with more publicly available documents hosted on the CBP website, or referenced by interviewees (see below). For example, if meeting minutes referenced a specific report, memo or letter, that document was added to the analysis. Python code illustrating how CBP website pages were scraped can be publicly accessed (see Code and Data Availability). In the Results section below, we will refer to analyses based on specific primary documents obtained in the above-described process.

Meeting minutes were analyzed using the qualitative data analysis software Dedoose (Dedoose, 2021). We tracked key shifts in topics discussed in meetings over time. We used codes to track themes, following a two-stage "grounded theory approach" that allows for free-coding of significant topics, and subsequent consolidation of preliminary codes into larger categories (Charmaz, 2014), rather than having *a priori* determined codes. The first author of this paper reviewed all documents, identifying chunks of text large enough to understand the context of the content, then assigning multiple codes to the chunk (Ryan and Bernard, 2003). After initial coding of all documents, the initial "open" codes were then consolidated into broader "axial" categories by reviewing initial code frequencies, and co-occurrences (Corbin and Strauss, 1990). The broader categories were interpreted as themes of discussions emerging from meeting minutes.

From the meeting minutes, we reconstructed timelines that visualized the timing of discussions around major model improvements (numbered phases in the Chesapeake Bay Program) (for more detail on the numbered phases, see: Linker et al., 2002), discrete events--such as the releases of major reports, passage of policies, and important meetings where significant decisions were made or important topics discussed (Kolar et al., 2017). The meeting minutes sometimes revealed conflicts and disagreements between attendees at the meetings. Consequently, a series of semi-structured interviews was conducted to provide or improve contextual understanding and reveal possible dynamics that may not have been captured in the official meeting minutes (Smith, 1995). Timelines were first produced that included only events, and events and themes coded from the meeting minutes. The timelines were then sequentially used as prompts following a semi-structured interview protocol. In historical and sociological research, the use of visual timelines has been shown to improve recall accuracy of interview subjects (van der Vaart, 2004). See the Supplementary Material A for the semi-structured interview protocol and timelines that were shown to key informants.

Key informants were identified through snowball sampling, with the seeds identified through meeting minutes (Parker et al., 2019). Once semi-structured interviews began to have significant overlap with previous interviews, we stopped the snowball sampling process. In total, we conducted interviews with 12 key informants. All interviews were conducted over Zoom (Zoom Video Communications, Inc.) and lasted between 30 and 60 minutes. The interviews were recorded and transcribed. After each interview, we used new content to revise interpretations of the timeline of events and themes, checking with subsequent interviewees that revisions were appropriate. Among the key informants were: CBP office staff, EPA administrators, at-large and appointed members of STAC, WQGIT, and MODWG, and one journalist who had attended and reported on many of the open meetings of the CBP.

The data from the interviews are presented anonymously without attribution or more details about the interviewee's position. Confidentiality was a mutually agreed upon condition between the researchers and the interviewees and as a condition of the ethics guidelines at the authors' institutions (VT-IRB #19-873). Anonymity helps prevent interviewees from potential retaliation for revealing controversial information. Lastly, full confidentiality allows interviewees to feel they can speak candidly (Punch, 1986). Information from the interviews was complemented with information from meeting minutes and other documents from the CBP. However, readers should understand the interviews provide data that otherwise would not be observable through official documentation and are reflective of the interviewees' perceptions and memories of what occurred and the importance of what occurred during the time period.

4. Results

We present our results in two sections. Section 4.1 outlines the reconstructed timeline of events and themes that emerged around the use of the watershed model. Section 4.2 specifically identifies how political influences shaped collective decisions about how a computer model should be adapted, using the "tyrannies" of participation by Cooke et al. (2001) as an organizing framework.

4.1 Timeline of events

Figure 1 shows the final timeline of events that were relevant to the shift in modeling approach for Phase 6 of the watershed model that was created from the documentary analysis and refined through interviews with key informants. We divide the events into two main eras: pre-TMDL (2010-12-29), and post-TMDL.

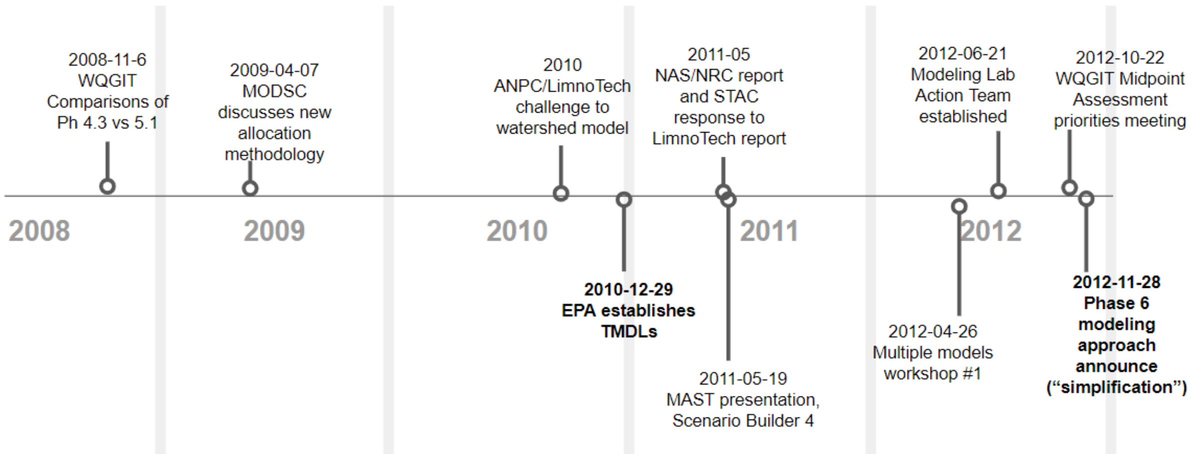

Figure 1: Timeline of key and illustrative events leading to the 2012-11-28 announcement to shift to a simplified approach for the watershed model, which reversed a decades-long trend toward increasing detail and resolution in the model. On 2010-12-29, the U.S. Environmental Protection Agency (EPA) established its Total Maximum Daily Loads (TMDLs), legislative limits to pollutant loading allocated to each jurisdiction in the watershed. The reconstruction of the timeline of events and thematic analysis revealed that this was a key turning point in the function of the model. Prior to the TMDL, the watershed model was used to determine the allocations; after the TMDL, the purpose of the model shifted to planning and tracking progress. After the shift in purpose, the authority of the model was challenged by some actors, but subsequently strengthened through top-down action. However, the shift toward a "simpler" model was deemed necessary for the new purpose of scenario-based planning and progress tracking.

Pre-TMDL (2008 – 2010)

During this period leading up to the release of the TMDL, the watershed model was primarily being used to determine nutrient allocations for each of the jurisdictions, answering questions about how much of the Chesapeake Bay's pollutant load (nitrogen, phosphorous, and sediment) was originating from what areas (which states would be responsible) and what sectors needed to be targeted (for example, agriculture vs urban wastewater and stormwater). A significant theme during this time was agreement of "fair and feasible" extreme endpoint scenarios used in the quantification of the allocation principles established in 2003, which included that (U.S. EPA, 2010):

- Basins that contribute the most to the problem must do the most to resolve the problem.
- States that benefit most from the Chesapeake Bay recovery must do more.
- All reductions in nutrient loads are credited toward achieving final assigned loads.

The watershed model needed to be able to operationalize these allocation principles and represent them quantitatively by establishing a level-of-effort metric of the distance between the "E3 Scenario", which was meant to represent a case in which Everyone, Everywhere, did Everything, and the "No Action" scenario, which was meant to represent a case in which no best management practices (BMPs) were implemented to improve water quality and previously existing ones were removed (U.S. EPA 2010; Section 7). The difference in nutrient and sediment loads between these two scenarios represents the range of possible levels of effort and is used in the calculation that assigns reduction responsibilities. An "allocation line" was developed that established a relationship between a jurisdiction's contribution to the Chesapeake Bay's pollution and the expected level of effort (U.S. EPA 2010; Section 6). Between 2008 and 2010, there was much discussion about the construction of these two scenarios. For example, on 2009-08-10, states were asked whether they "could live with" the way E3 and No Action were defined in the model. Below is an example of how state representatives responded (***emphasis added*** *to highlight language expressing of fairness and feasibility within the model*):

> ***[West Virginia representative]:*** *[we] can live with both [scenarios], [but the] devil is in details when it comes to the slope of the allocation line.*
>
> ***[Virginia representative]:*** *[We] can live with both point sources and air [definitions]. [We have] more concerns with E3. If E3 is going to [require an]* ***unreal level of effort****, [then it] needs to be [recognized as] equally unreal. [We should not have] a design flow concept with urban stormwater… [Non-point sources are] not equal[ly feasible]... Without knowing the lines, we are uncomfortable. Can live with no action, for E3 we cannot live with it, unless there is significantly more than 10% difference between [point source] and [nonpoint source]* ***lines to make-up for this lack of equity*** *it may be livable. Major problems include: achievability of point source… [it is] unreal to think that urban practices would get [the water quality benefits of] forest back, [or] pasture conversion to wetlands for agriculture.*
>
> ***[Pennsylvania representative]:*** *E3 is not acceptable because it would take* ***our point sources*** *down to 3 mg/l and there is* ***no possible way*** *that load could be addressed by [nonpoint source BMPs][4].*

This excerpt from the meeting minutes shows how different interpretations of fairness and feasibility are conflated in the discussions of these scenarios. For more detailed explanation on the Scenarios and their operationalization of fairness, see the Supplementary Material B.

Discussions such as the one above[5] illustrated how the model functioned as a boundary object through operationalization of key scenarios. The "boundaries" being crossed through definition of these scenarios

[4] CBP. 2009. "Minutes of the Water Quality Goal Implementation Team." https://www.chesapeakebay.net/channel_files/18331/wqgit_8-10_minutes_(2).pdf.

[5] The below meetings contained instances of "fairness" of model operationalization options being discussed.

CBP. 2008a. "Minutes of the Water Quality Steering Committee." http://archive.chesapeakebay.net/pubs/calendar/WQSC_01-22-08_Minutes_1_9043.pdf.

———. 2008. "Minutes of the Water Quality Steering Committee." http://archive.chesapeakebay.net/pubs/calendar/WQSC_04-22-08_Minutes_1_9376.pdf.

———. 2008. "Minutes of the Water Quality Steering Committee." http://archive.chesapeakebay.net/pubs/calendar/WQSC_06-16-08_Minutes_1_9533.pdf.

included disparate understandings of fairness and feasibility, what jurisdictions viewed as an undue burden, politically or economically feasible/infeasible, and how baselines and starting points (some states had made improvements prior to the upcoming TMDL faster than other states) should be taken into account (Linker et al., 2013). For example, some of the jurisdictions, such as Maryland and Virginia have coastlines directly adjacent to the Chesapeake Bay, and therefore might derive more economic benefit from its restoration -- from improved aquaculture, to more opportunities for recreation and tourism. States like West Virginia and New York would see few of these benefits. Another example was the balance between urban and rural land uses, and how to quantify the feasibility of reducing nutrients from point sources such as wastewater treatment plants, which were usually associated with urban and suburban areas, versus the feasibility of reducing nutrients from non-point sources, such as runoff from agricultural fields from thousands of individual land owners. The differences in socio-political impacts on the jurisdictions influenced their assertions of what was fair and feasible, which in turn, influenced model parameterization and testing.

As these concerns were discussed in WQGIT meetings, the model was updated between meetings to show the effects of changes to model parameters and data inputs on the model outputs. During the pre-TMDL period, the model had been under development for decades, and incorporated so much detail and so many processes that properly prepping the input data and running scenarios took the CBP office staff over two weeks (Devereux and Rigelman, 2014; Lim, 2021). Each time there was a change to input data, model parameter values, or even

[5] [*continued*]

CBP. 2008. “Minutes of the Water Quality Steering Committee.” http://archive.chesapeakebay.net/pubs/calendar/WQSC_07-21-08_Minutes_1_9534.pdf.

———. 2008. “Minutes of the Water Quality Steering Committee.” http://archive.chesapeakebay.net/pubs/calendar/WQSC_09-29-08_Minutes_1_9536.pdf.

———. 2008. “Minutes of the Water Quality Steering Committee.” http://archive.chesapeakebay.net/pubs/calendar/WQSC_10-20-08_Minutes_1_9773.pdf.

———. 2008. “Minutes of the Water Quality Steering Committee.” http://archive.chesapeakebay.net/pubs/calendar/WQSC_11-06-08_Minutes_1_9795.pdf.

———. 2009. “Minutes of the Water Quality Steering Committee.” http://archive.chesapeakebay.net/pubs/calendar/WQSC_01-12-09_Minutes_1_9901.pdf.

———. 2009. “Minutes of the Water Quality Steering Committee.” https://www.chesapeakebay.net/channel_files/13148/wqgit_minutes_080811.pdf.

———. 2009. “Minutes of the Water Quality Steering Committee.” http://archive.chesapeakebay.net/pubs/calendar/WQSC_02-09-09_Minutes_1_9904.pdf.

———. 2009. “Minutes of the Water Quality Steering Committee.” http://archive.chesapeakebay.net/pubs/calendar/WQSC_04-15-09_Minutes_1_10097.pdf.

———. 2009. “Minutes of the Water Quality Steering Committee.” http://archive.chesapeakebay.net/pubs/calendar/WQSC_05-18-09_Minutes_1_10241.pdf.

———. 2009. “Minutes of the Water Quality Steering Committee.” http://archive.chesapeakebay.net/pubs/calendar/WQSC_06-22-09_Minutes_1_10242.pdf.

———. 2000. “Minutes of the Water Quality Steering Committee.” http://archive.chesapeakebay.net/pubs/calendar/WQSC_07-06-09_Minutes_1_10345.pdf.

———. 2009. “Minutes of the Water Quality Goal Implementation Team.” https://www.chesapeakebay.net/channel_files/18331/wqgit_8-10_minutes_(2).pdf.

———. 2000. “Minutes of the Water Quality Goal Implementation Team.” https://www.chesapeakebay.net/channel_files/13572/wqgit_091409_minutes-f.pdf.

———. 2009. “Minutes of the Water Quality Goal Implementation Team.” https://www.chesapeakebay.net/channel_files/18326/wqgit_9-29.30_re-abridged_minutes.pdf.

———. 2009. “Minutes of the Water Quality Goal Implementation Team.” https://www.chesapeakebay.net/channel_files/18321/wqgit_121409_minutes-f.pdf.

———. 2010. “Minutes of the Water Quality Goal Implementation Team.” https://www.chesapeakebay.net/channel_files/13442/wqgit_041910_minutes-f.pdf.

———. 2010. “Minutes of the Water Quality Goal Implementation Team.” https://www.chesapeakebay.net/channel_files/13421/wqgit_051710_minutes-f.pdf.

———. 2011. “Minutes of the Water Quality Goal Implementation Team.” https://www.chesapeakebay.net/channel_files/13264/wqgit_2-14-11_minutes.pdf.

———. 2011. “Minutes of the Water Quality Goal Implementation Team.” https://www.chesapeakebay.net/channel_files/13238/wqgit_031411_minutes.pdf.

———. 2011. “Minutes of the Water Quality Goal Implementation Team.” https://www.chesapeakebay.net/channel_files/13148/wqgit_minutes_080811.pdf.

historical reference periods, the differences in model output could imply changes in allocations that would cost jurisdictions millions of dollars to mitigate. This in turn, would lead to debates and questioning by the jurisdictions that would last for weeks or months afterwards, asking for a ballooning number of model runs of the CBP office staff to pinpoint the exact cause of the changes[6]. This carried an extremely high transaction cost for the CBP office modeling team[7].

TMDL and beyond (2010 – 2013)

Beginning slightly before the official release of the Federal government's TMDL allocations on December 29, 2010, the importance of the outputs of the model became more evident and its implications more urgent. In addition to setting the allocations that each jurisdiction would be responsible for, the TMDL required each jurisdiction to prepare plans for how they would comply with the allocations, called Watershed Implementation Plans (WIPs) (U.S. EPA, 2010). As had been the case in recent years because the model had reached such high levels of complexity and detail, jurisdictions struggled to engage in the "what if" scenario testing necessary to evaluate alternatives that would be included in their plans, because running scenarios required so many computational resources and so much pre-processing and CBP office staff time. In fact, informal efforts to create statistical emulators and simplified versions of the watershed model had existed for about 10 years prior to the TMDL's passage, for use in education (Crouch et al., 2008) and decision support (Lim, 2021)[8]. However, as scenarios were required in planning documents, specific states contracted with local consultants to develop emulators such as the Virginia Approximation Scenario Tool (VAST) and the Maryland Approximation Scenario Tool (MAST) (described in Kaufman et al. 2021)[9]. These emulators generated wider attention among the

---

[6] In our analysis, "requests for more information" were mentioned in the following meetings:

CBP. 2008. "Minutes of the Water Quality Steering Committee." http://archive.chesapeakebay.net/pubs/calendar/WQSC_01-22-08_Minutes_1_9043.pdf.

———. 2008. "Minutes of the Water Quality Steering Committee." http://archive.chesapeakebay.net/pubs/calendar/WQSC_04-22-08_Minutes_1_9376.pdf.

———. 2008. "Minutes of the Water Quality Steering Committee." http://archive.chesapeakebay.net/pubs/calendar/WQSC_11-06-08_Minutes_1_9795.pdf.

———. 2009. "Minutes of the Water Quality Steering Committee." http://archive.chesapeakebay.net/pubs/calendar/WQSC_07-06-09_Minutes_1_10345.pdf.

———. 2009. "Minutes of the Water Quality Goal Implementation Team." https://www.chesapeakebay.net/channel_files/18331/wqgit_8-10_minutes_(2).pdf.

———. 2000. "Minutes of the Water Quality Goal Implementation Team." https://www.chesapeakebay.net/channel_files/18329/wqgit_090909_minutes_f.pdf.

———. 2009. "Minutes of the Water Quality Goal Implementation Team." https://www.chesapeakebay.net/channel_files/13572/wqgit_091409_minutes-f.pdf.

———. 2010. "Minutes of the Water Quality Goal Implementation Team." https://www.chesapeakebay.net/channel_files/13264/wqgit_2-14-11_minutes.pdf.

———. 2010. "Minutes of the Water Quality Goal Implementation Team." https://www.chesapeakebay.net/channel_files/13454/wqgit_040510,040610_minutes-f.pdf.

———. 2010. "Minutes of the Water Quality Goal Implementation Team." https://www.chesapeakebay.net/channel_files/13442/wqgit_041910_minutes-f.pdf.

———. 2010. "Minutes of the Water Quality Goal Implementation Team." https://www.chesapeakebay.net/channel_files/13421/wqgit_051710_minutes-f.pdf.

———. 2010. "Minutes of the Water Quality Goal Implementation Team." https://www.chesapeakebay.net/channel_files/13410/wqgit_060110_minutes-f.pdf.

———. 2011. "Minutes of the Water Quality Goal Implementation Team." https://www.chesapeakebay.net/channel_files/13238/wqgit_031411_minutes.pdf.

———. 2011. "Minutes of the Water Quality Goal Implementation Team." https://www.chesapeakebay.net/channel_files/13194/wqgit_5-23-11_minutes.pdf.

[7] Five anonymous informants in discussion with Theodore Lim, [4/4/2022, 4/5/2022, 4/8/2022, and 6/7/2022], transcript and recording.

[8] Three anonymous informants in discussion with Theodore Lim, [4/4/2022, 4/8/2022], transcript and recording.

[9] Additional evidence of these tools' development and use by jurisdictions are provided:

Devereux, Olivia. 2015. "Chesapeake Bay Program Decision Support Tools." Presented at the STAR Team Meeting: IAN Seminar, February 26. https://d18lev1ok5leia.cloudfront.net/chesapeakebay/documents/star_ian_seminar.pdf.

———. 2016. "Advanced Analysis of BayFAST, CAST, and VAST Scenario Results." February 17. http://www.chesapeake.org/stac/presentations/251_Devereux_BayFAST_20160217.pdf.

Loudon County Department of Building and Development. 2014. "Upper Broad Run Watershed Management Plan (Chapter 6: Modeling)." Leesburg, VA. https://www.loudoun.gov/DocumentCenter/View/108680/Chapter-06--Modeling?bidId=.

jurisdictions. For example, during a May 19, 2011 meeting of the Modeling Quarterly Review, MAST was presented to illustrate the convenience of jurisdictions being able to test and learn from running their own scenarios to support WIP development[10]. However, because they weren't complete representations of the CBP modeling system, the computational challenges still remained for evaluating the draft plans submitted to the CBP (Kaufman et al., 2021).

The capability to quickly run scenarios and immediately see results through a web-based interface increased in salience to users after the passage of the TMDL[11]. The early prototypes VAST and MAST were precursors to CAST, the Chesapeake Assessment and Scenario Tool, the simplified, web-based Phase 6 model that was officially released in 2017, which produces model output in seconds, compared to the weeks required by the Phase 5 model (Devereux and Rigelman, 2014). By the time of its release, CAST was no longer an emulator and therefore the "A" now stands for "Assessment" rather than "Approximation" (more detail on this development can be found in Lim, 2021). However, increases in computational time and the convenience of web-based scenario runs were not the only factors that led to the reversal of the decades-long trend of increasing detail to the more simplified model. As stated above, the WQGIT and the CBP office staff modelers had just emerged from a multiyear period of intense calibrations, debates about parameters, fairness and equity, and model intercomparisons, much of which led to confusion and frustration in determining what changes to the complex model were causing changes seen in model outputs[12].

In 2010, a challenge to the validity of the watershed model itself was led by the agricultural interest group, the Agricultural Nutrient Policy Council (ANPC), who hired the engineering firm Limnotech to analyze an alternative scientific model developed and applied by USDA to show that the USDA model (USDA, 2011) had different assumptions, data, and structure, and resulted in different agricultural acreage and loadings to the Chesapeake Bay compared to the CBP watershed model (LimnoTech, 2010). The implication of this challenge was that the basis for TMDL allocations was not scientifically credible, and that as a result, further action should be paused until credibility in the model could be restored. In response to the challenge, the scientists and technical experts of CBP's STAC deployed quickly to evaluate the claims made by ANPC/Limnotech and established that the reason for the discrepancy between the alternative model and the CBP's watershed model was actually a misuse in ANPC/Limnotech's application of the USDA model; once that mistake was corrected the outputs of both models were similar. STAC's response report was released in mid-2011 (CBP STAC, 2011). From the report:

> *LimnoTech based its recommendations on unrealistic criteria for watershed model performance, inappropriate expectations for agreement between watershed models developed for different objectives, selective interpretation of the findings of the CB-CEAP report, and errors in the interpretation of the models and their results. LimnoTech failed to acknowledge that fundamental differences in models (such as the input data, assumptions, and process representations) are unavoidable because of the different objectives of the models and differences in the data and resources available to support each effort. …When LimnoTech's errors in interpretation of model results are corrected, the results of the two models are more similar to each other than reported by LimnoTech. The corrected results indicate that the model predictions of loads are in approximate agreement despite the differences in model objectives, assumptions, input data, model frameworks, and spatial and temporal details. More importantly, the results of the two models are similar in their assessment of the need for implementing more management practices on cropland.*

Though the issue was resolved in less than a year, the challenge sent reverberations through the organization. The challenge demonstrated that stakeholders would not simply use the model for generating scenarios to

---

[9] [*continued*]

Maryland Department of the Environment. n.d. "Maryland Assessment and Scenario Tool (MAST)." Accessed February 12, 2023. https://mde.maryland.gov/programs/water/TMDL/TMDLImplementation/Pages/MASTInformation.aspx.

[10] Maryland Department of the Environment. 2011. "MAST: Maryland's Assessment and Scenario Tool." Presented at the Chesapeake Bay Modeling Quarterly Review, May 19. https://archive.chesapeakebay.net/pubs/calendar/11636_05-19-11_Presentation_6_11278.pdf.

[11] Four anonymous informants in discussion with Theodore Lim, [4/4/2022 4/8/2022] transcript and recording.

[12] Five anonymous informants in discussion with Theodore Lim, [4/4/2022, 4/5/2022, 4/8/2022, and 6/7/2022], transcript and recording.

inform mitigation actions. Rather, they would also question the legitimacy of its use[13]. STAC had long-called for uncertainty of the CBP's models to be evaluated due to the model's complexity at that point (Band et al., 2008). It was recognized that there were many parameters that were often referred to as "levers" that could be pulled to tweak model output, and this was sometimes referred to as evaluations of model sensitivity or robustness, and had often been in relation to the need to evaluate the effects of varied parameterizations on model output. This had led to calls for greater examination of the idea of "multiple models" to be used in the management of the Chesapeake Bay (e.g., Band et al. 2008). Now, stakeholders (in this case from the agricultural sector) were demanding explanations of why this model, why these parameters, as well as assessments of model skill (goodness-of-fit) compared to alternative scientific models. As a result, a workshop on "multiple models" was held in 2013, in which scientists, modelers, and stakeholders discussed the meanings of "multiple models" and how they could fit into the regulatory framework of the CBP (Weller et al., 2013).

Simultaneously, the ANPC/Limnotech challenge also heightened perceptions for the need for the science of modeling to be separated from the messiness of politics[14]. In 2012, the "Modeling Lab Action Team" was established in response to a National Academies report (National Research Council, 2011) to explore the proposition of having a standalone modeling lab that would have authority over model development and evaluation and increase transparency of the models used in the management of the Chesapeake Bay and that would be separate from political actors and interest groups.

In October 2012, a Midpoint Assessment priorities meeting was held, in which the CBP solicited feedback from all the workgroups on what their priorities were for the following phase of the program[15]. The first two priorities submitted by the Agriculture Workgroup both had to do with the model: (1) "Modeling Baseline/Input Data Assumptions/Needs"; and (2) "Scenario Builder and Chesapeake Bay Program Modeling Suite Transparency, Accuracy, and Confidence", reflecting the lessons learned from the ANPC/Limnotech challenge. Listed under the second priority was the recommendation to use "PQUAL" for the watershed model simulations. PQUAL is the simpler of two options to simulate nutrients within HSPF (Hydrological Simulation Program - FORTRAN) (Bicknell et al., 1997), the model code that the CBP had traditionally used. This recommendation for preferring less complex simulation was what ultimately led to the vast simplification of the Phase 6 CBP Watershed model as documented in the interviews conducted in this study[16].

On November 28, 2012, the CBP office modeling team presented its new approach to the watershed model: Phase 6, the Chesapeake Assessment and Scenario Tool (CAST). Shortly after that, the recommendations of the Modeling Lab Action Team to form a standalone entity to oversee model development were rejected by the Management Board of the CBP, and the action team was dissolved. These decisions were the culmination of the previously described events of 2008 − 2012 (and also shown in Figure 1). To summarize, those events included:

- Multiple, iterative discussions about how to fairly represent fairness and feasibility.
- High-stakes model changes that would result in multiplying questions about the causes of changed model outputs.
- High computational and staff time costs of completing model runs, and needs for jurisdictions to be able to iteratively design scenarios by which they could meet their TMDL allocations.
- Issues about the transparency of the model and difficulty explaining complex model output.
- Pressure to justify the legitimacy of the "crisp numbers" output from one model on which the TMDL was based in the face of intense debate about multiple, or alternative models, and questions about the legitimacy of the science behind the model.

---

[13] Five anonymous informants in discussion with Theodore Lim, [4/4/2022, 4/5/2022, 4/8/2022, and 4/11/2022], transcript and recording.

[14] Four anonymous informants in discussion with Theodore Lim, [4/5/2022, 4/8/2022, and 4/11/2022], transcript and recording.

[15] CBP. 2012. "Minutes of the Water Quality Goal Implementation Team." https://www.chesapeakebay.net/what/event/water_quality_goal_implementation_team_meeting; also see post-meeting summary presentation: https://d18lev1ok5leia.cloudfront.net/chesapeakebay/documents/atte3-summary_of_decisions_and_working_updates_to_schedule_11_6_12.pdf

[16] Seven anonymous informants in discussion with Theodore Lim, [4/4/2022, 4/6/2022, 4/8/2022, 4/11/2022], transcript and recording.

Supplementary Material C shows code frequencies by year for some of the emergent themes from the documents that helped to construct the timeline of events.

4.2 "Tyrannies" resulting from participation in the modeling process

In the literature review section above, we alluded to the "three tyrannies" that Cooke et al. use to explain how the process participation might result in unintended consequences (Cooke et al., 2001):

1. Tyranny of the group: group dynamics that reinforce the positions of the already powerful
2. Tyranny of control: facilitators overriding legitimate concerns of the group
3. Tyranny of method: participatory methods that result in decision outcomes that may not be as good as if other methods were used

Table 1 summarizes categories of unintended political influences observed in the study period -- categorized by Cooke and colleagues' three tyrannies.

Table 1: Political influences on decision-making around CBP watershed model (2008 ~ 2012), categorized by Cooke and colleagues' "three tyrannies" (Cooke et al., 2001) of participation.

Tyranny of the Group	Tyranny of Control	Tyranny of Method
Strategic Use. Jurisdictions/interest groups attempting to "game" the model	**Prioritization.** External political agendas at federal level and state level influence recommendations for model priorities. **Centralizing authority.** The model is used to strengthen hierarchical authority for the CBP	**Inertia.** Consensus-based governance makes status quo easier than change **Encouraging easy agreements.** Superficial participation through easy changes are preferred to more complex or substantial changes **Over-legitimization of the model.** Politics motivate a reliance on the model as the singular source of truth

4.2.1 Tyranny of the Group

Strategic Use

The challenge led by ANPC/Limnotech described above is an example of this type of political influence. As is shown in the following quotation, there was also evidence that jurisdictions and interest groups attempted to participate in model development simply in order to get the model to produce output that would be favorable to their interests (e.g., show greater pollutant reductions from past or planned actions). This quotation was made in reference to users of the model, usually watershed managers. However, interviewees[17] also referenced working groups within the CBP governance structure that they asserted were formed specifically to find and legitimize the parameter values that could show the largest reductions in the model.

> *But to me it's really odd … how much effort goes into gaming the model and us trying to do everything we possibly can to keep [it based on] sound science…The more [managers] learn, the more they know how to game the model, and the further away we get from science…. [managers] want to learn [how to use the model] because they know honestly, if they can get the model to show greater reductions than what may actually be occurring, they look good, and they meet their goals and things. Yeah, it can cost a couple hundred thousand dollars as salaries to game the model compared to literally hundreds of millions, if not billions of dollars, to actually do the work on the ground.*[18]

4.2.2 Tyranny of Control

Prioritization

Several interviewees commented that external political agendas influenced prioritizations for the CBP, which in turn, translated to the kinds of changes that could be incorporated into the models. The issue of climate change

[17] Three anonymous informants in discussion with Theodore Lim, [4/4/2022, 4/11/2022], transcript and recording.
[18] Anonymous informant in discussion with Theodore Lim, 4/4/2022, transcript and recording.

was given as an example multiple times. The watershed model's outputs for TMDL allocations were based on meteorological data (e.g., precipitation amounts) from a specific 10-year period that was deemed by the WQGIT to be representative of the historical record for the region. However, climate change is changing precipitation patterns and making extreme events more frequent. This affects how pollutants are transported to the Chesapeake Bay, and jurisdictions expected that incorporating the effects of climate change would increase their mitigation responsibilities. One interviewee said:

> *[The CBP leadership] were very sensitive to perceptions about the role of climate change at the White House level and at the state governor level. And so they did not want to make an already complicated issue harder to negotiate… When we [STAC] were providing input, we'd provide it to the CBP director or his immediate delegate, generally together, or to the governors… And so that is inherently connected to a state policy position*.[19]

Another confirmed:

> *I know the Chesapeake Bay Program was not really keen on trying to incorporate [climate change] at the time. It's in the middle of the political debate on climate change. Also, incorporating climate into [the model], [would result in] changes in allocations, which would cause some issues with the jurisdictions.*[20]

STAC's policies and procedures aim to position STAC to act as an objective, "honest broker" of scientific facts to the CBP. However, external political factors, such as receptibility and expected political pushback directly factored into their choice of recommendations among available priorities based on scientific facts. The collaborative, and good faith-based governance arrangement of the CBP resulted in efforts made to keep political leaders engaged in the process, not making radical changes or putting them in positions they might perceive as politically dangerous.

Centralizing authority

The problem of "multiple models" essentially was that multiple models were expected to lead to multiple model outputs, and authorities believed this could decrease the legitimacy of the policy goal: specific numbers specified in the TMDL to reduce nutrient loads in the Chesapeake Bay to 199 million lbs of nitrogen, 13 million lbs of phosphorous, and 19 billion lbs of sediment (U.S. EPA, 2010). This revealed fundamental disconnects between how scientists and technical staff viewed model outputs:

> *So from a modeler's or the engineer's perspective, you're like, Oh my gosh, they [the jurisdictions] are trading back and forth in what would be like the decimal dust of the model*.[21]

This quotation illustrates how much trust in the model was needed by jurisdictions in order to bargain with the allocations, and how this level of trust perhaps exceeded the precision of what the model could be expected to provide. In reference to having multiple models, another interviewee said:

> *It gets a little complicated in a regulatory environment, because in the end, you still need to end up with one number. The TMDL is still based on one number no matter how many models you have. So at some point you have to have a process that resolves all those different models and still gets to one number.*[22]

From most of the interviewees' perspectives, multiple models would have delegitimized that single number. However, the Multiple Models Workshops revealed the path forward chosen by the CBP. One interviewee described the impact of the workshops on their understanding of "multiple models" and how the concept could be used to strengthen the authority of the model, rather than weaken its authority:

> *[We realized] you can do multiple models, different ways. You can have multiple independent models. Just like a hurricane model. You can have one model that is the model. And then you have other models that are essentially used to validate that model. One way to do that is [to show that] academic models are generally finding the same thing [as our regulatory model],*

[19] Anonymous informant in discussion with Theodore Lim, 4/11/2022, transcript and recording.
[20] Anonymous informant in discussion with Theodore Lim, 4/5/2022, transcript and recording.
[21] Anonymous informant in discussion with Theodore Lim, 4/5/2022, transcript and recording.
[22] Anonymous informant in discussion with Theodore Lim, 4/4/2022, transcript and recording.

> *so we feel good about [it]. And then the third way was actually [to have] one model [that is made of] different components that anyone can mess with. So the regulated industry scientists, academics, anybody can mess with them. And they submit this revised module. And they have a bunch of different versions of the modules. And then they test those multiple versions of the module to see which one is best. And that becomes the regulatory model. [We realized that if] we could build a simple model then its parameters can be informed by multiple models....So we still have one model at the end of the day. A deterministic, one input, one output for CAST. But one of the big benefits of multiple models is that we will have the best parameter estimation using different multiple models to build them… [After Limnotech,] we wanted to be able to say that we had already incorporated [information that challengers might bring up] into the model*.[23]

This interpretation of "multiple models", while somewhat addressing scientific concerns about sensitivity, was chosen to provide more opportunities to strengthen the authority of the model itself, through subsuming the contribution of other models. While, as noted in Easton et al. (2017), the structure of CAST provides a better opportunity to perform uncertainty quantification compared to previous models, for which uncertainty quantification was minimal or often non-existent.

4.2.3 Tyranny of Method

Inertia

Paradoxically, the same consensus-based governance framework adopted by the CBP that is meant to prevent the kind of political gaming of the model mentioned above, also results in much more difficulty in pushing for changes to the model that many would consider positive[24]. On consensus, one interviewee said:

> *[Consensus] can really shut things down though because the way it works is, if there is a stop [a party in disagreement], it's kind of dead in the water. And it only takes one stop. So I think that approach really does hamper some of the changes in the model, sometimes that's a good thing because the proposal is just, and I'll be frank here, just kind of gaming the model. And not much science behind the change. But in some cases it's the change that would really enhance the model scientifically.*[25]

Encouraging easy agreements

Logistically, consensus-based decision-making caused model changes to be largely limited to input on parameter values. One interviewee described the process of making changes to the model:

> *There was no singular authority which you could call up and say, Hey, person X, please make a change. [Changes were] a product of engagement and persuasion. To get stuff done, you could have a workshop, requesting STAC resources. [STAC] would fund a workshop. We would [write a] workshop report. Sometimes that led to parametric changes [within the model]. There were opportunities for peer review. There's opportunities for dialogue. But, structurally my recollection is it's a professionalized model run by a cadre of professionals. At the end of the day, there were limited opportunities for non feds to make discrete and specific changes in those tools.*[26]

Another interviewee described their long-term observation of the difficulty of jurisdictions offering feedback on the model during meetings, because their level of knowledge and experience with the model was much more limited compared to the modeling staff:

> *CBP staff looks for data to use in the model. They present it at meetings to get signoff. But, at the meetings, it is hard to object to it. And they usually get signoff. Only after the jurisdictions start to apply the model over years do they find out that they sometimes get counterintuitive*

[23] Anonymous informant in discussion with Theodore Lim, 4/8/2022, transcript and recording.

[24] Consensus-based decision-making is usually thought of as a deliberative process to enhance voice and representation that is meant to prevent more powerful/majority actors from always getting their way. It is contrasted to aggregative processes, such as voting (Fung, 2006).

[25] Anonymous informant in discussion with Theodore Lim, 4/4/2022, transcript and recording.

[26] Anonymous informant in discussion with Theodore Lim, 4/11/2022, transcript and recording.

> *results. The process for making a change to the data doesn't actually work, because consensus is needed. The default is to keep using the problematic data. Consensus keeps people from gaming the system, but also from fixing the model.*[27]

Over-legitimization of the model

Over-legitimization of participatory models is one of the unintended consequences of participatory modeling mentioned by Korfmacher (2001). It refers to the tendency of decision-makers to place too much weight on technical data produced by scientists. In the case of the Chesapeake Bay however, over-legitimization also occurred because of the long history of the model's development and all that had been invested in it -- financially and emotionally because of the participatory process[28].

One interviewee referred to the model as the "lingua franca" of the CBP in the sense that it was a common basis of understanding that had been forged through years of effort, and that as a reason why changes were often resisted[29]. Over time, so much has been invested in the model and its outputs, and there is a preference built up around the model as the source of truth, to which everyone is bought in. Models were seen as attractive because they can offer a single, discrete number as an output, compared to the noisiness of empirically collected data[30]. One interviewee expressed frustration at the position that the model has gained as the single source of truth for the CBP over the decades, essentially describing how they perceived that the importance of the model's outputs has surpassed the importance of action in the Chesapeake Bay:

> *Instead of dedicating more resources to [modeling] maybe those resources could be better used for monitoring to support the model or for implementation… is the goal to build the best model? That's one of the objectives, but the goal is to clean up the Bay. The modeling tool gives us some good direction on who should do how much and to make policy decisions and helps us understand what our actions are likely going to accomplish. But it ultimately doesn't restore the Bay*.[31]

5. Discussion

The case of the Chesapeake Bay between 2008 and 2012 offers a nuanced example of how political factors influence participatory modeling. On the one hand, for a case that is frequently held up as an exemplar of participatory modeling in a policy context, the analysis revealed both relatively superficial means of participation as well as unintended consequences of participation. On the other hand, the model has gone through a major structural change in its Phase 6 form, has successfully withstood political challenges, still continues to support the Chesapeake Bay's efforts, and is still accepted as the basis of decision-making.

Interviews and documentary analysis supported the idea that the majority of "participation" in the model was not structural, but parameter-based. Parameter-based participation, in which the values of specific coefficients can be changed, is more superficial than discussions about what components are included in the model and how they are represented. However, rather than a purposeful withholding of power by those with privilege from those without (Arnstein, 1969), superficial participation was more the result of the difficulty of truly engaging participants in a model that by this time period already had so much legacy political and technical complexity built into it. In fact, the willingness of EPA officials to truly engage in consensus-based decision-making both around the model and the watershed itself, and the clear devolved organizational structure of the CBP, reflected elements of two of the three upper rungs of Arnstein's "ladder of citizen participation." Participation at the "partnership" rung (where power is redistributed amongst participants) and at the "delegated control" rung (exemplified by negotiations between citizens and public officials) are present in the approach to computer model development in the CBP. The uppermost rung in Arnstein's ladder, that of full citizen control, was not present in the CBP process of modeling and policy development. This is evidenced by the rejection of the standalone Modeling Lab proposal. Yet, despite what seem to be well-constructed governance arrangements

[27] Anonymous informant in discussion with Theodore Lim, 6/7/2022, transcript and recording.

[28] Four anonymous informants in discussion with Theodore Lim, [4/8/2022, 4/11/2022, 6/2/2022, 6/7/2022], transcript and recording.

[29] Anonymous informant in discussion with Theodore Lim, 6/2/2022, transcript and recording

[30] Two anonymous informants in discussion with Theodore Lim, [6/2/2022, 6/7/2022], transcript and recording.

[31] Anonymous informant in discussion with Theodore Lim, 4/8/2022, transcript and recording.

for decision-making about the model, interviewees still felt that governance arrangements -- such as needing to reach consensus -- prohibited some scientifically valuable changes to the model from being implemented.

All three the tyranny categories listed by Cooke et al. (2001) were observed in the time period analyzed in this research. The inertia caused by consensus-based decision-making was an example of tyranny of method. The use of participation as a means to strengthen authority exhibited the tyranny of control. Some jurisdictions' and stakeholders' willingness and ability to spend more on the salaries of skilled modelers to challenge the outputs and premise of the model illustrates how access to resources might result in tyranny of the group -- in which the positions of the already powerful are further reinforced -- although in this case, tyranny of the group was ultimately overridden by the power associated with the tyranny of control (top down authority of the EPA) when it rejected calls for a standalone modeling lab.

Further, while learning is typically considered a positive outcome of participatory processes, including participatory modeling (Hedelin et al., 2021; Jordan et al., 2018; Sterling et al., 2019), in our case, we observed that learning about ways to influence the model resulted in perceptions that the model was being usurped, or "gamed" by some participants. One distinction that could be made is that participatory modeling includes methods that range from more qualitative methods (and thus more similar to the suggested increased attention to "framing" by critics of rationalist approaches to the use of science in policy contexts) to highly quantitative methods (Voinov et al., 2018). The types of learning that occur alongside different types of modeling are important to distinguish. Lim (2021), for example, demonstrated in the Chesapeake Bay that modeling facilitated learning about the watershed, learning about the management of the watershed, and learning about the model itself. In different stages of policy-making and implementation, different types of learning may have different levels of utility. Understanding these different types of learning will help structure social learning opportunities that result in collective problem-solving, rather than individual gaming. Social learning may have both positive and negative effects on collective governance that need to be considered.

Voinov et al. point out the factors that researchers and practitioners should consider when designing participatory modeling processes, including: the number, diversity, backgrounds and skills of participants; the level and intensity of stakeholder participation; the timing and stages of participation; and power asymmetries (Voinov et al., 2018). In the case of the 2008 − 2012 period, and especially post-TMDL (2010), many of these considerations were not in the control of the CBP. This might be considered a kind of "path dependence" of technology that limits the ways in which subsequent changes can take place, given both the technological configurations of the past and the social and political interests that they embody (Dolfsma and Leydesdorff, 2009; Haasnoot et al., 2014; Lahtinen et al., 2017). In particular, this study period revealed the power of the model to legitimize itself as shown in the Over-legitimization theme of this study, through its resistance to change and the primacy of its outputs in the policy process. This finding aligns with that of scholars in the field of science and technology studies who point to the conceptualization of "momentums" embodied by technology outside of their human actors (Hughes, 1987; Latour, 2005).

The analysis period in our study also illustrated the process of "displacement" described by Rayner, in which models are used by organizations to socially construct ignorance, especially when dealing with uncomfortable knowledge (Rayner, 2012). Rayner (2012) defines displacement as: "the process by which an object or activity, such as a computer model designed to inform management of a real-world phenomenon, actually becomes the object of management." We observe displacement occurring most directly in the case when an interviewee expressed dismay at their perception that watershed managers are expending more of their mental efforts in constructing model runs to demonstrate progress than they are actually implementing mitigation projects on the ground. However, the case also adds to Rayner's "displacement" by illustrating why. In Rayner's analysis, an organization must find a way to deal with the uncomfortable knowledge provided by a model. In the case of the Chesapeake Bay watershed monitoring and modeling, the uncomfortable knowledge may be that empirical data can show inconclusive evidence about water quality improvements. The model offers a singular number that is more convenient than dealing with uncertainty. And yet, the model has become a critical tool for supporting organizational authority derived through decades of negotiation and deliberation in a polycentric governance arrangement. This was the product of both the tyranny of control (centralizing authority) and the tyranny of method (over-legitimization of the model).

The political influences in participatory modeling in the CBP between 2008 and 2012 illustrate both "evidence-based policy" and "policy-based evidence." The computer model is a chimera (Edmonds et al., under review, this Special Issue) in that its existence has been motivated by an effort to have "objective" science be the basis for difficult policy negotiations, but, at the same time, the organizational need to have a useful boundary object that is meant to encourage continued collaborative engagement from political actors. The former purpose aspires to be the basis of evidence-based policy. The latter is pragmatic, settling to produce policy-based evidence. The development of the CBP watershed model supports Strassheim and Kettunen's conceptualization of "expertise and evidence" as "socially embedded in authority relations and cultural contexts" (Strassheim and Kettunen, 2014). In its pragmatism, participatory modeling in a policy implementation context relies on simplification to increase legibility, counterintuitively enabling state control over the process, complexity, and open-endedness that is typically associated with participatory processes (Scott, 1999). The case illustrates another dimension of the balance between simplicity and complexity in model selection (Klosterman, 2012; Lim, 2021; Zellner et al., 2022) -- one that introduces hierarchical power relations and the wider governance context.

In addition, in the process of operationalization, models necessarily constrain both problem and solution spaces, as models are simplified versions of reality (Merry, 2016; Oreskes, 2003; Porter, 1996). Critics of how computational scientific models are used in public policy suggest that premature use and adoption of computational models and the primacy they assume in guiding or leading, which are ultimately political processes -- value-laden decisions where multiple interests and objectives must be addressed -- can be counter-productive. This is because computational and quantitative models eliminate potentially useful problem framings and associated solution spaces. Saltelli and Giampietro (2017) have called the resulting focus and reliance on precise (though perhaps inaccurate), quantitative models as "hypocognition." Notably, the resulting state of "hypocognition" is the exact opposite of the learning and communicative knowledge sharing that is a central aspiration of participatory modeling and the idea of models as boundary objects that facilitate cooperation across disciplines and perspectives (Glynn et al., 2017, 2018; Wu and Dunning, 2018).

Lastly, the case of the Chesapeake Bay's TMDL has a particular policy and governance context, in which governance is prescribed to be consensus-based and collaborative amongst the stakeholders, yet the premise of the TMDL itself is competitive and top-down. Allocation of a "nutrient diet" pits the jurisdictions against each other in the process because attempting to reduce one's own pollution mitigation responsibility implies others' responsibilities would increase. The combination of a single, authoritative model that produced "crisp numbers" to support the "crisp numbers" of the allocation might be considered to be a particularly American form of bureaucratism, born of a culture that paradoxically, is often simultaneously suspicious of expertise, yet insistent that administrative decisions are not political (Porter, 1996, p 195). In his influential work Trust in Numbers: The Pursuit of Objectivity in Science and Public Life, Theodore Porter describes a lack of trust specific to the culture of American bureaucracy, which "has inspired Congress to impose rules on every agency, dictating how to award contracts or hire and fire employees, as well as how to carry out its central mission. It sometimes even imposes such standards on itself… As currently practiced, it is a distinctive achievement of American political culture" (Porter, 1996, p 195). The problems observed in the case therefore are both (1) specific to the governance and policy context of the TMDL and CBP, and (2) reflective of the broader socio-cultural context in which the governance arrangement and TMDL policy were formed.

5.1 Toward a Theory of Model Governance

As mentioned in the Literature Review section, participatory modeling is motivated by gaps between science and action. Others have identified a need for more explicit attention to the modeling process and maximizing a model's impact to inform social-ecological challenges. For example, efforts to outline best practices in modeling has included model reporting and review, proper identification of purpose, and revision (Jakeman et al., 2006). As awareness of the connectedness of biophysical and social systems has grown, the modeling process is also being called upon to manage interdisciplinary knowledge integration at multiple scales (Iwanaga et al., 2021), and to take into account modeling as a social process that should be evaluated within its social context to improve a community environmental asset (Hamilton et al., 2019). Arnold et al. (2020) make the case for the governance of numerical models, which they define as "the managerial procedures by which technical modeling projects are governed."

While the above literature focuses on improving model impact, our research emphasizes a need for additional attention to the long-term institutional arrangements around a model and modeling program. We suggest that the concept of *computer model governance* -- which we define as: the rules, norms, processes, and structures (formally or informally) adopted within an organization to shape individual or collective actions around the development or use of a computer model -- may help distinguish the circumstances in which socio-political influences on modeling, including participation, are appropriate. In this paper, our use of the term “tyranny” to describe unintended impacts of participatory processes suggests that questioning both participation and social learning, and explicitly recognizing foundational dynamics, such as exercise of power and control may support a process that yields more of the intended outcomes of collective problem solving. Some have begun to examine subjectivity, perspective, and interests embedded in modeling processes and how socio-behavioral dynamics could lead to path dependency in model development (Lahtinen et al., 2017). There are opportunities to pair such analyses with broader theories of environmental governance, including those that more explicitly recognize the roles of power, scales and hierarchies, networks, conflict, and adaptation (Armitage et al., 2008; Booher and Innes, 2010; Cumming et al., 2006; Lubell and Segee, 2013; Morrison et al., 2019; Ostrom, 2011).

We believe that a theory of computer model governance may contribute to understanding the factors that influence public acceptance of computer models and the policies that are based on them. Such an understanding would be generalizable to other kinds of SES modeling contexts, including, for example: the nitrogen discussions in European Union countries, models of fisheries and other ecological systems, the development of global climate models with national policies, and models of pandemics, such as those used during Covid-19, many of which have already exhibited problems at the science-policy interface (Saltelli et al., 2020). More in-depth study of the nature of participation (for example: consultative, deliberative, or aggregative), socio-cultural context, and policy and legal structures shaping each case and thus a more robust theory of computer model governance may support understanding of the factors influencing public acceptance of computer models and the policies that are based on them. In addition, such a theory might also include an alternative to the “evidence-based policy”/ “policy-based evidence” dichotomy. A robust theory of computer model governance might attempt to elucidate the limitations of models and be explicit about policy goals, and guide both towards an interactive middle ground.

6. Conclusion

In this article, we examined the process of model development during a key transitionary period of the Chesapeake Bay Program’s watershed model, 2008 − 2012. Our research contributes examples of how political factors can directly and indirectly influence model development processes in an SES regulatory context. While participatory modeling is generally considered a desirable approach that may increase representation, incorporation of diverse knowledge sources, and acceptance of model outputs for policy-making, we illustrate that participation can sometimes result in unintended consequences. Future research to advance understanding of specific contexts of participation and model governance arrangements may be used to promote positive outcomes of participation in model development -- such as broader and more inclusive engagement, perspective sharing and learning, and collective action -- while minimizing unintended negative outcomes.

A more generalizable “theory of computer model governance” may support both: (1) how to best structure participation at various points of the model-policy development process so that unintended consequences -- such as “gaming” of the model, displacement, and organizational hypocognition -- are minimized, while positive outcomes -- such as broader and more inclusive engagement, perspective sharing and learning, and collective action -- are maximized; and (2) how to understand how the social process of modeling fits within its broader social and political context with both formal and informal institutions. Better understanding of computer model governance may provide the vocabulary and theory to describe that context and different options for how the modeling process can fit within it to make robust recommendations at any point in the process.

Acknowledgments

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

We would like to thank the 12 interviewees whose recollections of events between 2008 and 2012 were crucial to this study. We also thank two anonymous reviewers for *Socio-Ecological Systems Modelling,* and Mike Runge and Jeni Keisman who provided internal reviews for the U.S. Geological Survey; their comments and suggestions greatly improved the paper.

This material is based upon work supported by the National Science Foundation under Grant No. 2238552, awarded to Theodore Lim. The work was also supported by Virginia Polytechnic Institute and State University's Institute of Society, Culture and Environment.

References

Akerlof, G.A., 1997. Social Distance and Social Decisions. Econometrica 65, 1005–1027. https://doi.org/10.2307/2171877

Armitage, D., Marschke, M., Plummer, R., 2008. Adaptive co-management and the paradox of learning. Glob. Environ. Change 18, 86–98. https://doi.org/10.1016/j.gloenvcha.2007.07.002

Arnold, T., Guillaume, J.H.A., Lahtinen, T.J., Vervoort, R.W., 2020. From ad-hoc modelling to strategic infrastructure: A manifesto for model management. Environ. Model. Softw. 123, 104563. https://doi.org/10.1016/j.envsoft.2019.104563

Arnstein, S.R., 1969. A Ladder Of Citizen Participation. J. Am. Inst. Plann. 35, 216–224. https://doi.org/10.1080/01944366908977225

Band, L., Dillaha, T., Duffy, C., Reckhow, K., Welty, C., 2008. Scientific and Technical Advisory Committee: Chesapeake Bay Watershed Model Phase V Review (No. 08–003).

Berg, B.L., Lune, H., 2011. Qualitative Research Methods for the Social Sciences, 8th edition. Pearson, Boston.

Bicknell, B.R., Imhoff, J.C., Kittle, J.L., Donigan, A.S.Jr., Johanson, R.C., 1997. Hydrological Simulation Program--Fortran: User's manual for version 11 (No. EPA/600/R-97/080). U.S. Environmental Protection Agency, Athens, GA.

Booher, D.E., Innes, J.E., 2010. Governance for Resilience: CALFED as a Complex Adaptive Network for Resource Management. Ecol. Soc. 15.

Cash, D., Adger, W.N., Berkes, F., Garden, P., Lebel, L., Olsson, P., Pritchard, L., Young, O., 2006. Scale and Cross-Scale Dynamics: Governance and Information in a Multilevel World. Ecol. Soc. 11. https://doi.org/10.5751/ES-01759-110208

Cash, D.W., Clark, W.C., Alcock, F., Dickson, N.M., Eckley, N., Guston, D.H., Jäger, J., Mitchell, R.B., 2003. Knowledge systems for sustainable development. Proc. Natl. Acad. Sci. 100, 8086–8091. https://doi.org/10.1073/pnas.1231332100

CBP, 2014. Chesapeake Watershed Agreement. As amended Jan 24, 2020. https://www.chesapeakebay.net/documents/FINAL_Ches_Bay_Watershed_Agreement.withsignatures-HIres.pdf

CBP, 2021. Governance and Management Framework for the Chesapeake Bay Program. October 7, 2021. https://d18lev1ok5leia.cloudfront.net/chesapeakebay/documents/CBP_GDOC_Version_4.0.pdf

CBP, 2023a. How We're Organized: Scientific, Technical Assessment and Reporting (STAR): Modeling Workgroup [WWW Document]. URL https://www.chesapeakebay.net/who/group/modeling-team (accessed 2.12.23).

CBP, 2023b. How We're Organized: Water Quality Goal Implementation Team (GIT 3) [WWW Document]. URL https://www.chesapeakebay.net/who/group/water-quality-goal-implementation-team (accessed 2.12.23).

CBP, 2023c. Learn the Issues: Population Growth [WWW Document]. Chesapeake Bay Program. URL https://www.chesapeakebay.net/issues/threats-to-the-bay/population-growth (accessed 3.10.23).

CBP, 2023d. The Chesapeake Bay Program [WWW Document]. Chesapeake Bay Program. URL https://www.chesapeakebay.net/

CBP STAC, 2011. Review of the LimnoTech Report "Comparison of Load Estimates for Cultivated Cropland in the Chesapeake Bay Watershed" (No. 11– 02). Edgewater, MD.

Charmaz, K., 2014. Constructing Grounded Theory, Second edition. SAGE Publications Ltd, London; Thousand Oaks, Calif.

Cialdini, R.B., Goldstein, N.J., 2004. Social Influence: Compliance and Conformity. Annu. Rev. Psychol. 55, 591–621. https://doi.org/10.1146/annurev.psych.55.090902.142015

Colding, J., Barthel, S., 2019. Exploring the social-ecological systems discourse 20 years later. Ecol. Soc. 24. https://doi.org/10.5751/ES-10598-240102

Cooke, B., Cooke, P.B., Kothari, U., 2001. Participation: The New Tyranny? Zed Books.

Corbin, J., Strauss, A., 1990. Grounded theory research: Procedures, canons, and evaluative criteria 19.

Corburn, J., 2009. Cities, Climate Change and Urban Heat Island Mitigation: Localising Global Environmental Science. Urban Stud. 46, 413–427. https://doi.org/10.1177/0042098008099361

Crouch, J.R., Shen, Y., Austin, J.A., Dinniman, M.S., 2008. An educational interactive numerical model of the Chesapeake Bay. Comput. Geosci. 34, 247–258. https://doi.org/10.1016/j.cageo.2007.03.017

Cumming, G.S., 2016. Heterarchies: Reconciling Networks and Hierarchies. Trends Ecol. Evol. 31, 622–632. https://doi.org/10.1016/j.tree.2016.04.009

Cumming, G.S., Cumming, D.H.M., Redman, C.L., 2006. Scale mismatches in social-ecological systems: Causes, consequences, and solutions. Ecol. Soc. 11, 14.

Dedoose, 2021. Version 9.0.17, cloud application for managing, analyzing, and presenting qualitative and mixed method research data (2021). Los Angeles, CA: SocioCultural Research Consultants, LLC www.dedoose.com. Devereux, O., Rigelman, J.R., 2014. CAST: An Online Tool for Facilitating Local Involvement in Watershed Implementation Plans for the Chesapeake Bay Total Maximum Daily Load. J. Water Manag. Model. 1–8. https://doi.org/10.14796/JWMM.C364

Dolfsma, W., Leydesdorff, L., 2009. Lock-in and break-out from technological trajectories: Modeling and policy implications. Technol. Forecast. Soc. Change 76, 932–941. https://doi.org/10.1016/j.techfore.2009.02.004

Dryzek, J.S., 2013. The Politics of the Earth: Environmental Discourses, 3rd edition. Oxford University Press, Oxford.

Easton, Z., Scavia, D., Alexander, R., Band, L., Kleinman, P., Martin, J., Miller, A., Pizzuto, J., Smith, D., Welty, C., Tech, V., 2017. Scientific and Technical Advisory Committee Chesapeake Bay Watershed Model Phase 6 Review (No. 17– 007). CBP STAC.

Edmonds, B., Hofstede. G.J., Koch, J., le Page, C., Lim, T., Lippe, M., Nöldeke, B., van Delden, H., under review. Chimaera Modelling – when the modellers have to reconcile inconsistent elements. Socio-Environmental Systems Modelling.

Elster, J., 1989. Social Norms and Economic Theory. J. Econ. Perspect. 3, 99–117. https://doi.org/10.1257/jep.3.4.99

Ernst, H.R., 2003. Chesapeake Bay Blues: Science, Politics, and the Struggle to Save the Bay, 1st Edition, Rowman & Littlefield Publishers, Inc., Lanham.

Everett, J.A.C., Faber, N.S., Crockett, M., 2015. Preferences and beliefs in ingroup favoritism. Front. Behav. Neurosci. 9.p 15

Executive Office of the President, 2009. Chesapeake Bay Protection and Restoration.Federal Register, Executive Order 13508, May 12, 2009. https://www.federalregister.gov/documents/2009/05/15/E9-11547/chesapeake-bay-protection-and-restoration

Falconi, S.M., Palmer, R.N., 2017. An interdisciplinary framework for participatory modeling design and evaluation—What makes models effective participatory decision tools? Water Resour. Res. 53, 1625–1645. https://doi.org/10.1002/2016WR019373

Flyvbjerg, B., 2006. Five Misunderstandings About Case-Study Research. Qual. Inq. 12, 219–245. https://doi.org/10.1177/1077800405284363

French, R.D., 2019. Is it time to give up on evidence-based policy? Four answers. Policy Polit. 47, 151–168. https://doi.org/10.1332/030557318X15333033508220

Fung, A., 2006. Varieties of Participation in Complex Governance. Public Adm. Rev. 66, 66–75. https://doi.org/10.1111/j.1540-6210.2006.00667.x

Funtowicz, S.O., Ravetz, J.R., 1993. The Emergence of Post-Normal Science, in: Von Schomberg, R. (Ed.), Science, Politics and Morality. Springer Netherlands, Dordrecht. https://doi.org/10.1007/978-94-015-8143-1

Girod, B., Wiek, A., Mieg, H., Hulme, M., 2009. The evolution of the IPCC's emissions scenarios. Environ. Sci. Policy 12, 103–118. https://doi.org/10.1016/j.envsci.2008.12.006

Glynn, P.D., 2014. W(h)ither the Oracle? Cognitive Biases and Other Human Challenges of Integrated Environmental Modeling. International Environmental Modelling and Software Society (iEMSs) 7th Intl Congrees on Env. Modelling and Software, San Diego, CA, USA.https://www.iemss.org/society/index.php/iemss-2014-proceedings

Glynn, P.D., Chiavacci, S.J., Rhodes, C.R., Helgeson, J.F., Shapiro, C.D., Straub, C.L., 2022a. Value of Information: Exploring Behavioral and Social Factors. Front. Environ. Sci. 10.p 437

Glynn, P.D., Rhodes, C.R., Chiavacci, S.J., Helgeson, J.F., Shapiro, C.D., Straub, C.L., 2022b. Value of Information and Decision Pathways: Concepts and Case Studies. Front. Env. Sci 10, 805214.

Glynn, P.D., Voinov, A.A., Shapiro, C.D., White, P.A., 2018. Response to Comment by Walker et al. on "From Data to Decisions: Processing Information, Biases, and Beliefs for Improved Management of Natural Resources and Environments." Earths Future 6, 762–769. https://doi.org/10.1002/2018EF000819

Glynn, P.D., Voinov, A.A., Shapiro, C.D., White, P.A., 2017. From data to decisions: Processing information, biases, and beliefs for improved management of natural resources and environments. Earths Future 5, 356–378. https://doi.org/10.1002/2016EF000487

Gray, S., Voinov, A., Paolisso, M., Jordan, R., BenDor, T., Bommel, P., Glynn, P., Hedelin, B., Hubacek, K., Introne, J., Kolagani, N., Laursen, B., Prell, C., Olabisi, L.S., Singer, A., Sterling, E., Zellner, M., 2018. Purpose, processes, partnerships, and products: four Ps to advance participatory socio-environmental modeling. Ecol. Appl. 28, 46–61. https://doi.org/10.1002/eap.1627

Gregory, R., Failing, L., Harstone, M., Long, G., McDaniels, T., Ohlson, D., 2012. Structured Decision Making: A Practical Guide to Environmental Management Choices. John Wiley & Sons.

Haasnoot, M., van Deursen, W.P.A., Guillaume, J.H.A., Kwakkel, J.H., van Beek, E., Middelkoop, H., 2014. Fit for purpose? Building and evaluating a fast, integrated model for exploring water policy pathways. Environ. Model. Softw. 60, 99–120. https://doi.org/10.1016/j.envsoft.2014.05.020

Hamilton, S.H., Fu, B., Guillaume, J.H.A., Badham, J., Elsawah, S., Gober, P., Hunt, R.J., Iwanaga, T., Jakeman, A.J., Ames, D.P., Curtis, A., Hill, M.C., Pierce, S.A., Zare, F., 2019. A framework for characterising and evaluating the effectiveness of environmental modelling. Environ. Model. Softw. 118, 83–98. https://doi.org/10.1016/j.envsoft.2019.04.008

Hedelin, B., Gray, S., Woehlke, S., BenDor, T.K., Singer, A., Jordan, R., Zellner, M., Giabbanelli, P., Glynn, P., Jenni, K., Jetter, A., Kolagani, N., Laursen, B., Leong, K.M., Schmitt Olabisi, L., Sterling, E., 2021. What's left before participatory modeling can fully support real-world environmental planning processes: A case study review. Environ. Model. Softw. 143, 105073. https://doi.org/10.1016/j.envsoft.2021.105073

Hewitt, C., 1988. Offices Are Open Systems, in: Bond, A.H., Gasser, L. (Eds.), Readings in Distributed Artificial Intelligence. Morgan Kaufmann, pp. 321–329. https://doi.org/10.1016/B978-0-934613-63-7.50036-X

Hood, R.R., Shenk, G.W., Dixon, R.L., Smith, S.M.C., 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., Miller, A., Moriarty, J., Noe, G.B., Onyullo, G., Rose, K., Skalak, K., Tian, R., Veith, T.L., Wainger, L., Weller, D., Zhang, Y.J., 2021. The Chesapeake Bay program modeling system: Overview and recommendations for future development. Ecol. Model. 456, 109635. https://doi.org/10.1016/j.ecolmodel.2021.109635

Hughes, T., 1987. Evolution of Large Technical Systems, in: The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology. MIT Press, Cambridge, Mass.; London, pp. 51–82.

Intemann, K., 2015. Distinguishing between legitimate and illegitimate values in climate modeling. Eur. J. Philos. Sci. 5, 217–232. https://doi.org/10.1007/s13194-014-0105-6

Iwanaga, T., Wang, H.-H., Hamilton, S.H., Grimm, V., Koralewski, T.E., Salado, A., Elsawah, S., Razavi, S., Yang, J., Glynn, P., Badham, J., Voinov, A., Chen, M., Grant, W.E., Peterson, T.R., Frank, K., Shenk, G., Barton, C.M., Jakeman, A.J., Little, J.C., 2021. Socio-technical scales in socio-environmental modeling: Managing a system-of-systems modeling approach. Environ. Model. Softw. 135, 104885. https://doi.org/10.1016/j.envsoft.2020.104885

Jakeman, A.J., Letcher, R.A., Norton, J.P., 2006. Ten iterative steps in development and evaluation of environmental models. Environ. Model. Softw. 21, 602–614. https://doi.org/10.1016/j.envsoft.2006.01.004

Janis, I.L., 1982. Groupthink: Psychological Studies of Policy Decisions and Fiascoes, 2nd edition. Cengage Learning, Boston.

Jordan, R., Gray, S., Zellner, M., Glynn, P.D., Voinov, A., Hedelin, B., Sterling, E.J., Leong, K., Olabisi, L.S., Hubacek, K., Bommel, P., BenDor, T.K., Jetter, A.J., Laursen, B., Singer, A., Giabbanelli, P.J., Kolagani, N., Carrera, L.B., Jenni, K., Prell, C., National Socio-Environmental Synthesis Center Participatory Modeling Pursuit Working Group, 2018. Twelve Questions for the Participatory Modeling Community. Earths Future 6, 1046–1057. https://doi.org/10.1029/2018EF000841

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

Klosterman, R.E., 2012. Simple and Complex Models. Environ. Plan. B Plan. Des. 39, 1–6. https://doi.org/10.1068/b38155

Kolar, K., Ahmad, F., Chan, L., Erickson, P.G., 2017. Timeline Mapping in Qualitative Interviews: A Study of Resilience with Marginalized Groups. Int. J. Qual. Methods 14, 13–32. https://doi.org/10.1177/160940691501400302

Korfmacher, K.S., 2001. The Politics of Participation in Watershed Modeling. Environ. Manage. 27, 161–176. https://doi.org/10.1007/s002670010141

Kruger, J., Dunning, D., 1999. Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments. J. Pers. Soc. Psychol.77 (6) p 1121

Lahtinen, T.J., Guillaume, J.H.A., Hämäläinen, R.P., 2017. Why pay attention to paths in the practice of environmental modelling? Environ. Model. Softw. 92, 74–81. https://doi.org/10.1016/j.envsoft.2017.02.019

Latour, B., 2005. Reassembling the social: an introduction to actor-network-theory. Oxford University Press, Oxford; New York.

Latour, B., 1987. Science in Action: How to Follow Scientists and Engineers Through Society. Harvard University Press.

Layzer, J.A., 2011. Ecosystem-Based Management in the Chesapeake Bay, in: The Environmental Case: Translating Values Into Policy. CQ Press, Washington, D.C.

Liberatore, A., Funtowicz, S., 2003. 'Democratising' expertise, 'expertising' democracy: What does this mean, and why bother? Sci. Public Policy 30, 146–150. https://doi.org/10.3152/147154303781780551

Lim, T.C., 2021. Model emulators and complexity management at the environmental science-action interface. Environ. Model. Softw. 135, 104928. https://doi.org/10.1016/j.envsoft.2020.104928

LimnoTech, 2010. Comparison of Draft Load Estimates for Cultivated Cropland in the Chesapeake Bay Watershed, Prepared for the Agricultural Nutrient Policy Council. Ann Arbor, MI.

Lindner, I., Strulik, H., 2008. Social Fractionalization, Endogenous Appropriation Norms, and Economic Development. Economica 75, 244–258. https://doi.org/10.1111/j.1468-0335.2007.00614.x

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 J. Am. Water Resour. Assoc. 49, 986–1006. https://doi.org/10.1111/jawr.12105

Linker, L.C., Shenk, G., Wang, P., Hopkins, K.J., Pokharel, S., 2002. A Short History of Chesapeake Bay Modeling and the Next Generation of Water Shed and Estuarine Models, in: Proceedings of the Water Environment Federation. Presented at the Water Environment Federation, Alexandria, VA, pp. 569–582.

Lubell, M., Segee, B., 2013. Conflict and Cooperation in Natural Resource Management, in: Vig, N.J., Kraft, M.E. (Eds.), Environmental Policy: New Directions for the Twenty-First Century. CQ Press, Thousand Oaks,CA.

Martin, J., Runge, M.C., Nichols, J.D., Lubow, B.C., Kendall, W.L., 2009. Structured decision making as a conceptual framework to identify thresholds for conservation and management. Ecol. Appl. 19, 1079–1090. https://doi.org/10.1890/08-0255.1

Meadowcroft, J., 2002. Politics and scale: some implications for environmental governance. Landsc. Urban Plan., Scaling and Environmental Understanding 61, 169–179. https://doi.org/10.1016/S0169-2046(02)00111-1

Merry, S.E., 2016. The Seductions of Quantification: Measuring Human Rights, Gender Violence, and Sex Trafficking, Illustrated edition. The University of Chicago Press.

Morrison, T.H., 2017. Evolving polycentric governance of the Great Barrier Reef. Proc. Natl. Acad. Sci. 114, E3013–E3021. https://doi.org/10.1073/pnas.1620830114

Morrison, T.H., Adger, W.N., Brown, K., Lemos, M.C., Huitema, D., Phelps, J., Evans, L., Cohen, P., Song, A.M., Turner, R., Quinn, T., Hughes, T.P., 2019. The black box of power in polycentric environmental governance. Glob. Environ. Change 57, 101934. https://doi.org/10.1016/j.gloenvcha.2019.101934

National Research Council, 2011. Achieving Nutrient and Sediment Reduction Goals in the Chesapeake Bay: An Evaluation of Program Strategies and Implementation. National Academies Press, Washington, D.C.

Newig, J., Fritsch, O., 2009. Environmental governance: participatory, multi-level – and effective? Environ. Policy Gov. 19, 197–214. https://doi.org/10.1002/eet.509

Oreskes, N., 2003. The role of quantitative models in science, in: Canham, C.D., Cole, J.J., Laurenroth, W.K. (Eds.), Models in Ecosystem Science. Princeton University Press, NJ, USA, pp. 13–31.

Ormerod, R.J., 2009. The history and ideas of critical rationalism: the philosophy of Karl Popper and its implications for OR. J. Oper. Res. Soc. 60, 441–460. https://doi.org/10.1057/palgrave.jors.2602573

Ostrom, E., 2011. Background on the Institutional Analysis and Development Framework. Policy Stud. J. 39, 7–27. https://doi.org/10.1111/j.1541-0072.2010.00394.x

Paolisso, M., Trombley, J., Hood, R.R., Sellner, K.G., 2015. Environmental Models and Public Stakeholders in the Chesapeake Bay Watershed. Estuaries Coasts 38, 97–113. https://doi.org/10.1007/s12237-013-9650-z

Parker, C., Scott, S., Geddes, A., 2019. Snowball Sampling. SAGE Res. Methods Found.

Popper, K., 2002. The Logic of Scientific Discovery. Routledge. https://doi.org/10.4324/9780203994627

Porter, T.M., 1996. Trust in Numbers: The Pursuit of Objectivity in Science and Public Life, Trust in Numbers. Princeton University Press. https://doi.org/10.1515/9781400821617

Punch, M., 1986. The politics and ethics of fieldwork. Beverly Hills, CA: Sage, 1986. 93 p.

Python Software Foundation, 2023. Python Language Reference.

Rayner, S., 2012. Uncomfortable knowledge: the social construction of ignorance in science and environmental policy discourses. Econ. Soc. 41, 107–125. https://doi.org/10.1080/03085147.2011.637335

Reagan, R., 1984. Address Before a Joint Session of the Congress on the State of the Union.

Ryan, G.W., Bernard, H.R., 2003. Techniques to Identify Themes. Field Methods 15, 85–109. https://doi.org/10.1177/1525822X02239569

Saltelli, A., Bammer, G., Bruno, I., Charters, E., Di Fiore, M., Didier, E., Nelson Espeland, W., Kay, J., Lo Piano, S., Mayo, D., Pielke Jr, R., Portaluri, T., Porter, T.M., Puy, A., Rafols, I., Ravetz, J.R., Reinert, E., Sarewitz, D., Stark, P.B., Stirling, A., van der Sluijs, J., Vineis, P., 2020. Five ways to ensure that models serve society: a manifesto. Nature 582, 482–484. https://doi.org/10.1038/d41586-020-01812-9

Saltelli, A., Funtowicz, S.O., 2014. When All Models Are Wrong. Issues Sci. Technol. 30, 79–85.

Saltelli, A., Giampietro, M., 2017. What is wrong with evidence based policy, and how can it be improved? Futures, Post-Normal science in practice 91, 62–71. https://doi.org/10.1016/j.futures.2016.11.012

Saltelli, A., Stark, P.B., Becker, W., Stano, P., 2015. Climate Models as Economic Guides: Scientific Challenge or Quixotic Quest? Issues Sci. Technol. 79–84.

Sarewitz, D., 2004. How science makes environmental controversies worse. Environ. Sci. Policy, Science, Policy, and Politics: Learning from Controversy Over The Skeptical Environmentalist 7, 385–403. https://doi.org/10.1016/j.envsci.2004.06.001

Scientific and Technical Advisory Committee, 2023. Chesapeake Bay Scientific and Technical Advisory Committee [WWW Document]. URL https://www.chesapeake.org/stac/

Scott, J.C., 1999. Seeing like a State: How Certain Schemes to Improve the Human Condition Have Failed. Yale University Press, New Haven, Conn.

Shenk, G.W., Linker, L.C., 2013. Development and Application of the 2010 Chesapeake Bay Watershed Total Maximum Daily Load Model. JAWRA J. Am. Water Resour. Assoc. 49, 1042–1056. https://doi.org/10.1111/jawr.12109

Smith, J.A., 1995. Semi structured interviewing and qualitative analysis, in: Smith, J.A., Harre, R., Van Langenhove, L. (Eds.), . Sage Publications, pp. 9–26.

Star, S.L., 1999. The Ethnography of Infrastructure. Am. Behav. Sci. 43, 377–391. https://doi.org/10.1177/00027649921955326

Star, S.L., Griesemer, J.R., 1989. Institutional Ecology, `Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. Soc. Stud. Sci. 19, 387–420. https://doi.org/10.1177/030631289019003001

Sterling, E.J., Zellner, M., Jenni, K.E., Leong, K., Glynn, P.D., BenDor, T.K., Bommel, P., Hubacek, K., Jetter, A.J., Jordan, R., Olabisi, L.S., Paolisso, M., Gray, S., 2019. Try, try again: Lessons learned from success and failure in participatory modeling. Elem. Sci. Anthr. 7, 9. https://doi.org/10.1525/elementa.347

Sterman, J.D., 2012. Sustaining Sustainability: Creating a Systems Science in a Fragmented Academy and Polarized World, in: Weinstein, M.P., Turner, R.E. (Eds.), Sustainability Science. Springer New York, New York, NY, pp. 21–58. https://doi.org/10.1007/978-1-4614-3188-6_2

Strassheim, H., Kettunen, P., 2014. When does evidence-based policy turn into policy-based evidence? Configurations, contexts and mechanisms. Evid. Policy 10, 259–277. https://doi.org/10.1332/174426514X13990433991320

Sundberg, M., 2007. Parameterizations as Boundary Objects on the Climate Arena. Soc. Stud. Sci. 37, 473–488. https://doi.org/10.1177/0306312706075330

U.S. EPA, 2010. Chesapeake Bay Total Maximum Daily Load for Nitrogen, Phosphorous and Sediment. https://www.epa.gov/chesapeake-bay-tmdl/chesapeake-bay-tmdl-document

USDA, 2011. Assessment of the Effects of Conservation Practices on Cultivated Cropland in the Chesapeake Bay Region, Conservation Effects Assessment Project (CEAP).

van den Hove, S., 2007. A rationale for science–policy interfaces. Futures 39, 807–826. https://doi.org/10.1016/j.futures.2006.12.004

van der Sluijs, J., 2005. Uncertainty as a monster in the science–policy interface: four coping strategies. Water Sci. Technol. 52, 87–92. https://doi.org/10.2166/wst.2005.0155

van der Sluijs, J.P., Petersen, A.C., Janssen, P.H.M., Risbey, J.S., Ravetz, J.R., 2008. Exploring the quality of evidence for complex and contested policy decisions. Environ. Res. Lett. 3, 024008. https://doi.org/10.1088/1748-9326/3/2/024008

van der Vaart, W., 2004. The Time-line as a Device to Enhance Recall in Standardized Research Interviews: A Split Ballot Study. J. Official Stat. 20, 301.

Voinov, A., Bousquet, F., 2010. Modelling with stakeholders. Environ. Model. Softw., Thematic Issue - Modelling with Stakeholders 25, 1268–1281. https://doi.org/10.1016/j.envsoft.2010.03.007

Voinov, A., Jenni, K., Gray, S., Kolagani, N., Glynn, P.D., Bommel, P., Prell, C., Zellner, M., Paolisso, M., Jordan, R., Sterling, E., Schmitt Olabisi, L., Giabbanelli, P.J., Sun, Z., Le Page, C., Elsawah, S., BenDor, T.K., Hubacek, K., Laursen, B.K., Jetter, A., Basco-Carrera, L., Singer, A., Young, L., Brunacini, J., Smajgl, A., 2018. Tools and methods in participatory modeling: Selecting the right tool for the job. Environ. Model. Softw. 109, 232–255. https://doi.org/10.1016/j.envsoft.2018.08.028

Voinov, A., Kolagani, N., McCall, M.K., Glynn, P.D., Kragt, M.E., Ostermann, F.O., Pierce, S.A., Ramu, P., 2016. Modelling with stakeholders – Next generation. Environ. Model. Softw. 77, 196–220. https://doi.org/10.1016/j.envsoft.2015.11.016

Weller, D.E., Benham, B., Friedrichs, M., Naijar, R., Paolisso, M., Pascual, P., Shenk, G., Sellner, K., 2013. Multiple Models for Management in the Chesapeake Bay (No. 14– 004). Edgewater, MD.

Wesselink, A., Buchanan, K.S., Georgiadou, Y., Turnhout, E., 2013. Technical knowledge, discursive spaces and politics at the science–policy interface. Environ. Sci. Policy, SI: Environmental and Developmental Discourses: Technical knowledge, discursive spaces and politics 30, 1–9. https://doi.org/10.1016/j.envsci.2012.12.008

White, D.D., Wutich, A., Larson, K.L., Gober, P., Lant, T., Senneville, C., 2010. Credibility, salience, and legitimacy of boundary objects: water managers' assessment of a simulation model in an immersive decision theater. Sci. Public Policy 37, 219–232. https://doi.org/10.3152/030234210X497726

Wu, K., Dunning, D., 2018. Hypocognition: Making Sense of the Landscape beyond One's Conceptual Reach. Rev. Gen. Psychol. 22, 25–35. https://doi.org/10.1037/gpr0000126

Wyborn, C., Datta, A., Montana, J., Ryan, M., Leith, P., Chaffin, B., Miller, C., van Kerkhoff, L., 2019. Co-Producing Sustainability: Reordering the Governance of Science, Policy, and Practice. Annu. Rev. Environ. Resour. 44, 319–346. https://doi.org/10.1146/annurev-environ-101718-033103

Zellner, M.L., Milz, D., Lyons, L., Hoch, C.J., Radinsky, J., 2022. Finding the Balance Between Simplicity and Realism in Participatory Modeling for Environmental Planning. Environ. Model. Softw. 157, 105481. https://doi.org/10.1016/j.envsoft.2022.105481