

<https://helda.helsinki.fi>

Black Boxes and the Role of Modeling in Environmental Policy Making

Maeda, Eduardo

2021-03-23

Maeda , E , Haapasaari , P , Helle , I , Lehikoinen , A , Voinov , A & Kuikka , S 2021 , ' Black Boxes and the Role of Modeling in Environmental Policy Making ' , Frontiers in Environmental Science , vol. 9 , 629336 . <https://doi.org/10.3389/fenvs.2021.629336>

<http://hdl.handle.net/10138/328444>

<https://doi.org/10.3389/fenvs.2021.629336>

cc_by

publishedVersion

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.



Black Boxes and the Role of Modeling in Environmental Policy Making

Eduardo Eiji Maeda^{1*}, Päivi Haapasaaari^{2,3}, Inari Helle^{4,5,6}, Annukka Lehikoinen^{2,3}, Alexey Voinov^{7,8} and Sakari Kuikka^{2,3}

¹Department of Geosciences and Geography, University of Helsinki, Helsinki, Finland, ²Ecosystems and Environment Research Programme, Faculty of Biological and Environmental Sciences, University of Helsinki, Helsinki, Finland, ³Kotka Maritime Research Center, Kotka, Finland, ⁴Organismal and Evolutionary Biology Research Programme, Faculty of Biological and Environmental Sciences, University of Helsinki, Helsinki, Finland, ⁵Helsinki Institute of Sustainability Science, University of Helsinki, Helsinki, Finland, ⁶Natural Resources Institute Finland (Luke), Helsinki, Finland, ⁷Center on Persuasive Systems for Wise Adaptive Living, Faculty of Engineering and IT, University of Technology Sydney, Sydney, NSW, Australia, ⁸Faculty of Engineering Technology, University of Twente, Enschede, Netherlands

OPEN ACCESS

Edited by:

Tracylee Clarke,
California State University Channel
Islands, United States

Reviewed by:

Michael Soltys,
California State University Channel
Islands, United States

Scott Feister,
California State University Channel
Islands, United States

*Correspondence:

Eduardo Eiji Maeda
eduardo.maeda@helsinki.fi

Specialty section:

This article was submitted to
Science and Environmental
Communication,
a section of the journal
Frontiers in Environmental Science

Received: 14 November 2020

Accepted: 17 February 2021

Published: 23 March 2021

Citation:

Maeda EE, Haapasaaari P, Helle I, Lehikoinen A, Voinov A and Kuikka S (2021) Black Boxes and the Role of Modeling in Environmental Policy Making. *Front. Environ. Sci.* 9:629336. doi: 10.3389/fenvs.2021.629336

Modeling is essential for modern science, and science-based policies are directly affected by the reliability of model outputs. Artificial intelligence has improved the accuracy and capability of model simulations, but often at the expense of a rational understanding of the systems involved. The lack of transparency in black box models, artificial intelligence based ones among them, can potentially affect the trust in science driven policy making. Here, we suggest that a broader discussion is needed to address the implications of black box approaches on the reliability of scientific advice used for policy making. We argue that participatory methods can bridge the gap between increasingly complex scientific methods and the people affected by their interpretations

Keywords: artificial intelligence, modeling, environmental policy, decision making, stakeholder engagement

INTRODUCTION

Computer algorithms and models are critical in modern science. They have been embraced by academic circles for decades, permeating influence all the way from policy-making to people's everyday life. Today, any person can check the weather forecast in the local news or on a cell phone, without noticing the colossal cyber infrastructure behind the information provided. From classical physics to complex models and algorithms run on supercomputers, the issue of understanding and trusting them is irrelevant for someone who simply wants to choose which clothes to wear tomorrow. People are happy to believe that the forecast is based on the best available knowledge. In this case, science has become part of people's daily life and is well accepted, despite the end user's superficial understanding of the methods and tools involved.

In many other cases, the increasing complexity of scientific methods and models tends to disconnect the general public from science, as the underlying logic and theories become more difficult to understand even for other scientists. When it comes to artificial intelligence, for example, some researchers argue that the entire field is becoming too complex for anyone to understand (Voosen, 2017). In fields such as machine learning, even top scientists sometimes struggle to make sense of the mechanisms driving their own results.

When the internal logic of a model is hidden, it is often referred to as a black box model. The term "black box" traditionally comes from empirical models that simply relate inputs and outputs, without knowledge or assumptions about the internal functions or processes. Similar black box concepts can be used to assist modular approaches, in which the inner workings of smaller components are temporarily ignored, so interactions in the higher structure can be better understood. Nonetheless, today even

process based models become similar to black box ones because the complexity of the processes and interactions included makes them too difficult to understand, especially by policy makers and stakeholders with no special training. Indeed, as scientific understanding of the system deepens, as new methods emerge, and computational power increases, our access to the internal logic of models tends to become more obscure. This does not make such models less accurate or useful in comparison to more transparent causal or process-based models. However, the question is how can we maintain trust in and usability of such black box models in environmental policy-making?

MODEL COMPLEXITY AND POLICY-MAKING

Socio-ecological systems are inherently—and increasingly—complex, involving interactions of numerous factors, stakeholder groups and sectors at multiple temporal and spatial scales (Voinov et al., 2016). Given such complexity, observations alone are incapable of informing the management of environmental problems or the use of natural resources (Schmolke et al., 2010). In addition to observations, an understanding of the system's logic is needed. This leads us to an intrinsic paradox in the use of models to support environmental policymaking. By definition, a model is a simplified representation of the key processes and factors that explain the behavior of a system in a certain context within certain boundaries. Replacing a system with a model and then analyzing it, provides justified predictions for the future and allows the exploration of alternative policies and management scenarios (Schmolke et al., 2010). However, with the advent of modern artificial intelligence tools that enable, for example, increasingly efficient machine learning and integrated modeling studies, the scientific approaches are now heading to the opposite direction with increasing model complexity.

Already three decades ago, scientists warned that models providing results that are too abstract and distant from the realities of political challenges are almost certain to fail as tools for policy-making (Kraemer and King, 1986). We argue that this may be also the case when the processes driving modeling become inaccessible for non-scientists, e.g., due to high complexity. A well-known example can be seen in the field of climate change mitigation, where scientists have not yet succeeded in convincing all policy makers and the public in either the problem or the results produced by climate models. It is notorious that the lobbying by climate change deniers can influence funding for climate research, as well as affect the implementation of policies (Stern et al., 2016). Climate change is, of course, a complex example, where multiple conflicting interests are in place, and science itself is by far not the only player. However, when it is difficult for the general public to understand the theory behind climate models it certainly weakens the scientific arguments underlying the policy advice (Stern et al., 2016; Visschers, 2018).

Achieving public confidence and trust is thus a major challenge for models designed to support environmental policy. Problem framing defines the questions a model can answer, and leads to identifying relevant variables and their interactions. Hence, the conceptual phase of model design is

highly dependent on the participation of experts and stakeholders altogether. In such a context, the less transparent are the models, the bigger is the challenge for integrating stakeholders into the policy making process. Current protocols, as the ones set by the European Commission guidelines for better regulation, warns that successful modeling requires communicating to decision makers how a model works and the strengths and limitations of a chosen modeling approach. However, official guidelines do not yet foresee how fields such as artificial intelligence will affect public confidence in scientific advice.

The challenges imposed by the lack of transparency and understandability of emerging technologies should not, however, create a barrier to the advancement of science. Instead of closing themselves in their black boxes, researchers should invest in alternative strategies to keep policy makers and stakeholders involved, in order to maintain their trust independently of their access to the models, by explaining why the models behave as they do. Studies have shown that stakeholders involved in modeling do not necessarily need to understand the scientific method as such, if their role is, for example, to provide practical information for the modeler (Ritzema et al., 2010; Mäntyniemi et al., 2013; Voinov et al., 2016). To maintain stakeholders' trust in a model, the results of which will affect their lives, modelers need to clarify the assumptions and inference logic behind the models and justify their choices (Voinov et al., 2018).

STRATEGIC ROLE OF COMMUNICATION

Using scientific advice for policy making involves transmitting knowledge from researchers to non-scientists and/or non-specialists and back. If this link is poorly maintained, the level of complexity in the scientific methods becomes irrelevant, as even the simplest approaches can be criticized and discarded. Transmitting the right message is therefore important for extending the reach of scientific advice, given that numerous barriers are in place for effectively communicating scientific results to policy makers (Cartwright et al., 2016). For example, communicating the uncertainties intrinsic to a model is essential (Stern et al., 2016; Visschers, 2018). As demonstrated in fisheries science, misunderstanding and misinterpretation of model uncertainties can reduce confidence in the policy advice and lead to decision paralysis in situations when stakeholders cannot or do not want to take appropriate actions (Cartwright et al., 2016). Obstacles in the communication of model results are expected, given that scientific management advice often concerns controversial, complex and uncertain matters. Likewise, it is important for end-users to develop a critical view toward scientific methods, so that scientific advice is not blindly accepted as infallible but, instead, realistic expectations from scientific inputs are created. There are, of course, guidelines and best practices for mitigating or overcoming many of these communication bottlenecks (Gluckman, 2014). Nonetheless, communication barriers will likely be magnified when scientific methods become more complex and less digestible, making scientific advice ever more challenging.

ACTIONABLE RECOMMENDATIONS

As the increasing complexity of models starts to influence policy making, it is important for scientists to create new approaches to communicate their underlying assumptions, reasoning, data and methods to stakeholders. Here, we define a stakeholder as a person, group, or organization involved in or affected by a course of action (in this case, actions guided by models). These may also include scientists from different disciplines, as complex socio-environmental problems are by nature transdisciplinary and require input from various disciplines to provide relevant scientific advice. When models serve decision making that directly affects people's lives, as in the case of many natural resource management actions, it is important to make the stakeholders part of the modeling process.

For instance, some of the tools also categorized under artificial intelligence, such as graphical Bayesian Networks (Korb and Nicholson, 2010), can help users to better perceive and understand complex and uncertain systems. Visualization of the framing and causal structures of inference combined with the possibility to interactively adjust scenarios, update the level of knowledge and observe the system's reactions in terms of dynamically updating visual probability distributions, improves the understanding about the conditional aspects and impact mechanisms of potential management interventions under consideration (Carriger et al., 2018; Kaikkonen et al., 2021). Bayesian methods can be designed to use the input from stakeholders as an intrinsic part of the model, in which case local knowledge can be used to provide priors for the model, leading to more realistic outputs (i.e., posteriors). In general, approaches that allow end-users to manipulate model parameters can be a powerful learning tool, as even if the internal logic of the model is hidden, users can still understand how changes in the system affect the final results.

Participatory modeling (PM) is an interactive learning process that employs the knowledge and/or values of stakeholders to create modeling and communication tools (Voinov et al., 2018). The level of engagement of stakeholders in PM can vary. The most intense participation occurs when the people affected by the results are engaged in all stages of a modeling process, from problem identification and framing, model designing, parameter selection, data collection, parameter estimation, model validation, and communication of the results, all the way to using the model for policy advice (Voinov et al., 2016). Most often, however, stakeholders only participate in some of the modeling stages. When it comes to models that from the stakeholder perspective are seen as "black boxes", a full participation of stakeholders, as described above, may not be feasible as the model neither includes clear model structures nor understandable parameters. Still, stakeholder involvement in defining appropriate questions, key indicators and/or management objectives and their prioritization is not only possible (McCabe and Halog, 2018; Laurila-Pant et al., 2019), but highly desirable for keeping the stakeholders informed and involved and for improving the quality of the policy making process. Stakeholder engagement is also essential to explain how uncertainties are treated in modeling.

Participatory approaches that bring together complex models and stakeholders have been implemented in a large variety of applications dealing with environmental or natural resource policy. For example, during the next years, European fisheries are transitioning to new rules aiming at reducing discards and mandatorily bringing all catches to land. These new rules will significantly affect fishery activities in Europe. As it is impossible to completely avoid unwanted catches, developing models to identify the most suitable fishing grounds to avoid them will be critical for maintaining the competitiveness as well as acceptability of the European fishing industry. To improve the performance of such models while at the same time allowing an easier assimilation and stronger legitimacy of the tools at local levels, researchers have created a framework for combining stakeholders' knowledge with a Bayesian model of fishery discards (Maeda et al., 2017). This framework allows fishermen to input their prior knowledge of the fishing grounds into the model (i.e., to indicate in which areas they think by-catches are high), to then combine it with observed data and provide posterior probabilities of fishery discards in the form of maps. In this case, although the core algorithms are mathematically demanding and remain a "black box" from the perspective of the fishermen, the credibility of the overall framework is strengthened as the end users are involved in the modeling process from the beginning and see that their knowledge is used and preferences are taken into account, before receiving the final product.

PM has also been applied for supporting policies related to oil spill accidents, water resources management, and agricultural practices, among many others (Vayssières et al., 2011; Lehtikoinen et al., 2013; Basco-Carrera et al., 2017; Hedelin et al., 2017). A common ground among these applications is the capability for integrating local knowledge and views, leading to a stronger transfer of values and perspectives among modelers and stakeholders. Such benefits have increased trust in scientific information during the process of environmental decision-making (Hedelin et al., 2017; Ulibarri, 2018). Given the requirement to gain stakeholders' and policy makers' trust in science for policy, the importance of PM may be growing in environmental policy making as models are even more often seen as black boxes.

Despite the benefits of PM, engaging stakeholders in scientific projects is not easy (Lavery, 2018). Among the challenges is the poor understanding of the outcomes that participatory approaches can produce, or the lack of a clear definition of who should be involved as a stakeholder (Lavery, 2018). This, as well as lack of evidence in the needs for stakeholder engagement, can significantly affect resource allocation in projects involving PM implying that budgets for stakeholder participation remain too small. We believe that addressing these challenges will become increasingly important, as the complexity of scientific methods together with the need for participatory approaches continues to grow. Thus, the next challenge for the scientific community is to find an appropriate balance between the prediction capacity and understandability of the models used.

The popularity of PM studies has grown remarkably in recent years (Voinov et al., 2016) and proper communication practices

become a key component in science (Gluckman, 2014). There are, therefore, good opportunities for reconciling advanced technologies and people benefiting from them. While scientific advice may be, in the future, increasingly based on complex models, the environmental policies benefiting from these tools should head into the direction of more transparent and collaborative processes. It may be difficult to communicate the computational details of a black box model, but it is still necessary to communicate the logic behind the results, as well as the policy process that uses the models (Röckmann et al., 2012; Glynn et al., 2018; Jordan et al., 2018). To keep the participatory processes up with the advancements of the fast-developing computation capacity and modeling technology, substantial effort should be spent on developing tools and skills to support stakeholder participation and engagement.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

REFERENCES

- Basco-Carrera, L., Warren, A., van Beek, E., Jonoski, A., and Giardino, A. (2017). Collaborative modelling or participatory modelling? A framework for water resources management. *Environ. Model. Softw.* 91, 95–110. doi:10.1016/j.envsoft.2017.01.014
- Carriger, J. F., Dyson, B. E., and Benson, W. H. (2018). Representing causal knowledge in environmental policy interventions: Advantages and opportunities for qualitative influence diagram applications. *Integr. Environ. Assess. Manag.* 14, 381–394. doi:10.1002/ieam.2027
- Cartwright, S. J., Bowgen, K. M., Collop, C., Hyder, K., Nabe-Nielsen, J., Stafford, R., et al. (2016). Communicating complex ecological models to non-scientist end users. *Ecol. Model.* 338, 51–59. doi:10.1016/j.ecolmodel.2016.07.012
- Gluckman, P. (2014). Policy: the art of science advice to government. *Nature* 507, 163–165. doi:10.1038/507163a
- Glynn, P. D., Voinov, A. A., Shapiro, C. D., and 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”. *Earth's Future* 6, 762–769. doi:10.1002/2018ef000819
- Hedelin, B., EversAlkan-Olsson, M., Alkan-Olsson, J., and Jonsson, A. (2017). Participatory modelling for sustainable development: key issues derived from five cases of natural resource and disaster risk management. *Environ. Sci. Policy* 76, 185–196. doi:10.1016/j.envsci.2017.07.001
- Jordan, R., Gray, S., Zellner, M., Glynn, P. D., Voinov, A., Hedelin, B., et al. (2018). Environmental synthesis center participatory modeling pursuit working GroupTwelve questions for the participatory modeling community. *Earth's Future* 6, 1046–1057. doi:10.1029/2018ef000841
- Kaikkonen, L., Parviainen, T., Rahikainen, M., Uusitalo, L., and Lehikoinen, A. (2021). Bayesian networks in environmental risk assessment: a review. *Integr. Environ. Assess. Manag.* 17, 62–78. doi:10.1002/ieam.4332
- Korb, K. B., and Nicholson, A. E. (2010). *Bayesian Artificial Intelligence*. Cleveland, OH: CRC Press
- Kraemer, K. L., and King, J. L. (1986). OR practice-computer-based models for policy making: uses and impacts in the United States Federal government. *Operations Res.* 34, 501–512. doi:10.1287/opre.34.4.501
- Laurila-Pant, M., Mäntyniemi, S., Venesjärvi, R., and Lehikoinen, A. (2019). Incorporating stakeholders' values into environmental decision support: a

AUTHOR CONTRIBUTIONS

All authors contributed equally with discussions, editings and writing, to produce the final version of the manuscript. EM wrote the initial version of the manuscript.

FUNDING

EM was funded by the Academy of Finland (decision numbers 318252 and 319905). IH was funded by the Helsinki Institute of Sustainability Science (HELSUS), University of Helsinki. SK and AL were supported by the Strategic Research Council (SRC) at the Academy of Finland, under the project “Creative adaptation to wicked socioenvironmental disruptions” (WISE) (decision number 312627), and SK also under the SRC project “SmartSea” (Grant No. 292 985) and under the Lloyd's Register Foundation (Agreement May 28, 2013). The Lloyd's Register Foundation supports the advancement of engineering-related education and funds research and development that enhances the safety of life at sea, on land, and in the air.

- Bayesian Belief Network approach. *Sci. Total Environ.* 697, 134026. doi:10.1016/j.scitotenv.2019.134026
- Lavery, J. V. (2018). Building an evidence base for stakeholder engagement. *Science* 361 (6402), 554–556. doi:10.1126/science.aat8429
- Lehikoinen, A., Luoma, E., Mäntyniemi, S., and Kuikka, S. (2013). Optimizing the recovery efficiency of Finnish oil combating vessels in the gulf of Finland using bayesian networks. *Environ. Sci. Technol.* 47, 1792–1799. doi:10.1021/es303634f
- Maeda, E. E., Mäntyniemi, S., Despoti, S., Musumeci, C., Vassilopoulou, V., Stergiou, K. I., et al. (2017). A Bayesian model of fisheries discards with flexible structure and priors defined by experts. *Ecol. Model.* 366, 1–14. doi:10.1016/j.ecolmodel.2017.10.007
- Mäntyniemi, S., Haapasaaari, P., Kuikka, S., Parmanne, R., Lehtiniemi, M., and Kaitaranta, J. (2013). Incorporating stakeholders' knowledge to stock assessment: central Baltic herring. *Can. J. Fish. Aquat. Sci.* 70, 591–599. doi:10.1139/cjfas-2012-0316
- McCabe, A., and Halog, A. (2018). Exploring the potential of participatory systems thinking techniques in progressing SLCA. *Int. J. Life Cycle Assess.* 23, 739–750. doi:10.1007/s11367-016-1143-4
- Ritzema, H., Froebrich, J., Raju, R., Sreenivas, C., and Kselik, R. (2010). Using participatory modelling to compensate for data scarcity in environmental planning: a case study from India. *Environ. Model. Softw.* 25, 1450–1458. doi:10.1016/j.envsoft.2010.03.010
- Röckmann, C., Ulrich, C., Dreyer, M., Bell, E., Borodzicz, E., Haapasaaari, P., et al. (2012). The added value of participatory modelling in fisheries management—what has been learnt?. *Mar. Policy* 36, 1072–1085. doi:10.1016/j.marpol.2012.02.027
- Schmolke, A., Thorbek, P., DeAngelis, D. L., and Grimm, V. (2010). Ecological models supporting environmental decision making: a strategy for the future. *Trends Ecol. Evol. (Amst)* 25 (8), 479–486. doi:10.1016/j.tree.2010.05.001
- Stern, P. C., Perkins, J. H., Sparks, R. E., and Knox, R. A. (2016). Global change research. The challenge of climate-change neoskepticism. *Science* 353, 653–654. doi:10.1126/science.aaf6675
- Ulibarri, N. (2018). Collaborative model development increases trust in and use of scientific information in environmental decision-making. *Environ. Sci. Policy* 82, 136–142. doi:10.1016/j.envsci.2018.01.022
- Vayssières, J., Vigne, M., Alary, V., and Lecomte, P. (2011). Integrated participatory modelling of actual farms to support policy making on sustainable intensification. *Agric. Syst.* 104, 146–161. doi:10.1016/j.agry.2010.05.008

- Vischers, V. H. M. (2018). Public perception of uncertainties within climate change science. *Risk Anal.* 38, 43–55. doi:10.1111/risa.12818
- Voinov, A., Jenni, K., Gray, S., Kolagani, N., Glynn, P. D., Bommel, P., et al. (2018). Tools and methods in participatory modeling: selecting the right tool for the job. *Environ. Model. Softw.* 109, 232–255. doi:10.1016/j.envsoft.2018.08.028
- Voinov, A., Kolagani, N., McCall, M. K., Glynn, P. D., Kragt, M. E., Ostermann, F. O., et al. (2016). Modelling with stakeholders—next generation. *Environ. Model. Softw.* 77, 196–220. doi:10.1016/j.envsoft.2015.11.016
- Voosen, P. (2017). The AI detectives. *Science* 357 (6346), 22–27. doi:10.1126/science.357.6346.22

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Maeda, Haapasaari, Helle, Lehtikainen, Voinov and Kuikka. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.