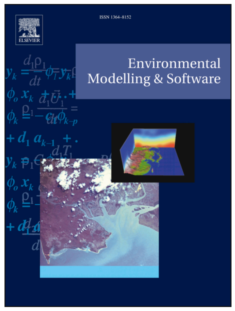
Iwanaga, T., Wang, H.-H., Hamilton, S.H., **Grimm, V.**, Koralewski, T.E., Salado, A., Elsawah, S., Razavi, S., Yang, J., Glynn, P., Badham, J., Voinov, A., Chen, M., Grant, W.E., Peterson, T.R., **Frank, K.**, Shenk, G., Barton, C.M., Jakeman, A.J., Little, J.C. (2021):
Socio-technical scales in socio-environmental modeling: managing a system-of-systems modeling approach
*Environ. Modell. Softw.* **135** , art. 104885

Journal Pre-proof




Socio-technical scales in socio-environmental modeling: managing a system-of-systems modeling approach

Takuya Iwanaga, Hsiao-Hsuan Wang, Serena H. Hamilton, Volker Grimm, Tomasz E. Koralewski, Alejandro Salado, Sondoss Elsawah, Saman Razavi, Jing Yang, Pierre Glynn, Jennifer Badham, Alexey Voinov, Min Chen, William E. Grant, Tarla Rai Peterson, Karin Frank, Gary Shenk, C. Michael Barton, Anthony J. Jakeman, John C. Little

PII: S1364-8152(20)30942-7

DOI: https://doi.org/10.1016/j.envsoft.2020.104885

Reference: ENSO 104885

To appear in: *Environmental Modelling and Software*

Accepted Date: 29 September 2020

Please cite this article as: 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., Socio-technical scales in socio-environmental modeling: managing a system-of-systems modeling approach, *Environmental Modelling and Software*, https://doi.org/10.1016/j.envsoft.2020.104885.

# Socio-technical scales in socio-environmental modeling: managing a system-of-systems modeling approach

Takuya Iwanaga[1], Hsiao-Hsuan Wang[2], Serena H. Hamilton[1,3], Volker Grimm[4], Tomasz E. Koralewski[2], Alejandro Salado[5], Sondoss Elsawah[1,6], Saman Razavi[7], Jing Yang[8], Pierre Glynn[9], Jennifer Badham[10], Alexey Voinov[11], Min Chen[12], William E. Grant[2], Tarla Rai Peterson[13], Karin Frank[4], Gary Shenk[14], C. Michael Barton[15], Anthony J. Jakeman[1], John C. Little[16]

[1]Institute for Water Futures and Fenner School of Environment and Society, The Australian National University, Canberra, Australia
[2]Ecological Systems Laboratory, Department of Ecology and Conservation Biology, Texas A&M University, College Station, TX 77843, USA
[3]CSIRO Land & Water, Canberra, Australia
[4]Helmholtz Centre for Environmental Research - UFZ, Department of Ecological Modelling, Leipzig, Germany and University of Potsdam, Plant Ecology and Nature Conservation, Potsdam, Germany
[5]Grado Department of Industrial and Systems Engineering, Virginia Tech, Blacksburg, VA, 24061, USA
[6]School of Electrical Engineering and Information Technology, University of New South Wales, Australian Defence Force Academy, Canberra, ACT, Australia
[7]Global Institute for Water Security, School of Environment and Sustainability, Department of Civil, Geological, and Environmental Engineering, University of Saskatchewan, Saskatoon, Saskatchewan, Canada
[8]National Institute of Water and Atmospheric Research, New Zealand
[9]U.S. Department of the Interior, U.S. Geological Survey, Reston, Virginia, USA
[10]Centre for Research in Social Simulation, University of Surrey, Guildford, GU2 7XH, United Kingdom
[11]Center on Persuasive Systems for Wise Adaptive Living (PERSWADE), Faculty of Engineering & IT, University of Technology Sydney, Australia
[12]Key Laboratory of Virtual Geographic Environment (Nanjing Normal University), Ministry of Education, Nanjing, 210023, China
[13]Environmental Science and Engineering Program, University of Texas at El Paso, El Paso, TX 79968, USA
[14]U.S Geological Survey, Chesapeake Bay Program, Annapolis, MD, 21403, USA
[15]Center for Social Dynamics & Complexity, School of Human Evolution & Social Change, Arizona State University, Tempe Arizona
[16]Department of Civil and Environmental Engineering, Virginia Tech, Blacksburg, VA, 24061, USA

**Abstract:**
System-of-systems approaches for integrated assessments have become prevalent in recent years. Such approaches integrate a variety of models from different disciplines and modeling paradigms to represent a socio-environmental (or social-ecological) system aiming to holistically inform policy and decision-making processes. Central to the system-of-systems approaches is the representation of systems in a multi-tier framework with nested scales. Current modeling paradigms, however, have disciplinary-specific lineage, leading to inconsistencies in the conceptualization and integration of socio-environmental systems. In this paper, a multidisciplinary team of researchers, from engineering, natural and social sciences, have come together to detail socio-technical practices and challenges that arise in the consideration of scale throughout the socio-environmental modeling process. We identify key paths forward, focused on explicit consideration of scale and uncertainty, strengthening interdisciplinary communication, and improvement of the documentation process. We call for a grand vision (and commensurate funding) for holistic system-of-systems research that engages researchers, stakeholders, and policy makers in a multi-tiered process for co-creation of knowledge and solutions to major socio-environmental problems.

**Keywords:** social-ecological modeling, interdisciplinary modeling, integrated modeling, scale issues, system-of-systems approach

# 1. Introduction

Socio-environmental systems (SES) function across a range of inter-related scales that collectively represent a system of systems (SoS). The term SoS has been used since the 1950s and various definitions exist (Nielsen et al., 2015). In this paper, we distinguish between an SoS as a collection of human and natural systems, and SoS models which are engineered representations of an SoS. The former is defined as an interconnected collection of multiple, heterogeneous, distributed systems that collectively may give rise to emergent behavior, where each system represents a process or set of processes. In the modeling of SoS, we follow Little et al. (2019) who define SoS models as "a collection of independent constituent systems, in which each fulfills its own purpose while acting jointly towards a common goal." (p. 84). In environmental modeling, SoS models may take the form of Integrated Assessment Models (IAMs) or, more generally, Integrated Environmental Models (IEMs), which are commonly applied to inform environmental management processes (Ewert et al., 2011; Iwanaga et al., 2020; Letcher (Kelly) et al., 2013; Matott et al., 2009).

Central to SoS modeling is the view of system representations as a multi-tier structure with different levels of abstraction, where systems and indicators at lower levels can be scaled up to higher levels. These representations capture processes that operate at different scales (e.g. temporal, spatial, organizational) in contrast to 'single-system' approaches, which assume such drivers to be exogenous and, crucially, do not account for any feedback mechanisms between the represented systems. This view also sets the focus on how to integrate knowledge from the different disciplines involved and coordinate information exchange among these in a consistent and meaningful way. Knowledge integration is not limited to the technical coupling of models, but to integration among multi-scale stakeholder and expert processes. This combined socio-technical focus makes scale issues and their treatment a core consideration of SoS modeling.

## 1.1. The need for a holistic treatment of scale

A crucial ingredient in SoS modeling is attending to the socio-technical processes involved. Representation of scales is defined by modelers for a particular purpose (Meadows, 2008) and is ultimately subject to human processes. Accordingly, the representation of an SoS is the end-product of what the people involved implicitly or explicitly have chosen to represent, and how they implemented their choices. These then influence the model structure and uncertainties embedded, and the consideration of its different dimensions, analyses conducted, and data and methods used (Glynn et al., 2017; Gorddard et al., 2016; Voinov et al., 2018). Such choices are subject to the available knowledge, experiences, biases, beliefs, heuristics and social values, as well as the perceived purpose(s) of the modeling.

A key scale issue in SoS modeling is the development of a consistent and defensible characterization of scale (Elsawah et al., 2020). Existing systems analysis and modeling approaches tend to come from entrenched disciplinary paradigms and so with a specific focus on their scales and facets, and embedded language and terms. Inconsistencies then manifest in the conceptualization and treatment of scale in SoS, which prevent researchers from: (1) understanding the implications of scale choices; (2) formulating, implementing and validating models that are relevant to the questions of interest; (3) predicting future SoS responses in support of decision making (Elsawah et al., 2020; Little et al., 2019; Razavi et al., 2020); and (4) communicating modeling results in ways that help identify trade-offs and synergies within an SoS and among the systems under investigation (Fridman and Kissinger, 2019; Miyasaka et al., 2017). Addressing issues that arise from the conceptualization and representation of multiple scales are often omitted or left for future discussion (Ayllón et al., 2018).

Discrepancies in the treatment of scale can be addressed firstly by developing a shared understanding of the system(s) being analyzed through a holistic interdisciplinary process (Thompson, 2009; White et al., 2019). There is increasing recognition that holistic approaches are necessary to enable an integrated assessment of scale issues in socio-environmental (social-ecological) systems (Schlüter et al., 2019a, 2019b; Hoekstra et al., 2014). The rise of inter/multidisciplinary fields, such as socio-hydrology (Elshafei et al., 2014; Sivapalan et al., 2012) and eco-hydrology (Hannah et al., 2004; Porporato and Rodriguez-Iturbe, 2002), gives further credence to this need. For SoSs in particular, it is necessary to additionally acknowledge the socio-technical influences on their modeling. Explicit inclusion of the socio-technical perspective pushes beyond traditional modeling approaches, as it advocates assimilation of not only the data and mechanistic processes across different systems, but also includes the knowledge and information held in the social institutions involved in the modeling.

## 1.2. Purpose

The purpose of this paper is to advance knowledge and implementation of interdisciplinary SoS modeling by identifying and articulating the practices, issues and challenges involved with respect to issues of scale. Central to this interdisciplinary lens is making concrete the multidimensional nature of scale issues and the interplay among these. Here, the term “interdisciplinary” is favored over trans- or multi-disciplinary as the focus is on the “blending” of disciplinary knowledge (White et al., 2019) and avoiding implications of the extent to which they are blended.

The primary audience of the paper is modelers, albeit in different domains and scientific disciplines with an interest in adopting an SoS approach as a methodological framework in SES modeling. In the following (Section 2), we first provide definitions for the key terminology used throughout this paper. These definitions are not intended to be universal but are provided to contextualize and aid in communication given the range of disciplines involved in SES modeling. In Section 3, we explore issues of scale which need to be considered throughout the modeling. We then describe in Section 4 the long-term challenges towards resolving such scale issues and suggest paths to be taken in the shorter-term.

# 2. Concepts and definitions of scale

## 2.1. The Process of Defining Scales

SoS models principally provide a representation of the interactions that occur between the systems involved. Holistic integration of knowledge from the various disciplines involved is necessary so that the implications of the different methodological choices on scale can be understood (Elsawah et al., 2020). To this end, a three-day workshop was held in October 2019 in which a culturally and disciplinary diverse group of 20 participants convened to share their knowledge. An additional 3 contributed in complementary ways to the drafting of this paper. Contributors originated from Europe, North America and the Asia-Pacific and included engineers, economists, social scientists, mathematicians, physicists, hydrologists, computer scientists and ecologists.

To prevent miscommunication, we developed a set of terms (outlined in the next section) to build a shared language (Rubin et al., 2010; Spitzberg and Cupach, 1989; Thompson, 2009). Although prior definitions of “scale” are available (see for example Cash et al., 2006; Gibson et al., 2000), it was considered useful to develop a shared, empathetic understanding of each other’s perspectives (Banerjee et al., 2019; Thomas and McDonagh, 2013). The process additionally served to break down cognitive constraints (MacLeod and Nagatsu, 2018), which may otherwise blind researchers to relevant notions of scale allowing disciplinary bias to creep in and knowledge gaps to form. The range of disciplines involved in SES modeling often makes addressing cognitive constraints difficult, as there are different notions of scale, and related terms are used in different ways depending on context. This variance has been observed in the use of common terms with conflicting definitions between (and sometimes within) disciplinary fields (Bridle et al., 2013).

## 2.2. Scale Terminology in SoS modeling

Defining the terminology associated with scales was an arduous process at first, owing to the diversity amongst workshop participants. A brief overview of the resulting primary terms used in this paper is provided in Table 1. For the discussion here, “scale” is taken to have an expansive definition, covering the scope of work to be conducted in the treatment and representation of system processes. Aspects of scale that had unanimous consensus included the *commensurability* of the choice of scale within the *purpose of the modeling,* and the consistency of *spatial* and *temporal* scales across models. It was also acknowledged that scale can mean many things beyond the spatial and temporal, for example the less tangible such as treatment of ethical considerations within the modeling process (e.g. Häyhä et al., 2016). Regardless of definitions, treatment of scales - and the choices made in this treatment - influences the model uncertainties and the outcomes of the modeling.

*Commensurability* refers to the appropriateness of the selected approaches and methods for the SoS modeling purpose. Broadly speaking, these approaches can be described as being subject to *socio-technical* considerations, which are the focus of the discussion in this paper. The social (human) aspect of modeling includes the circumstances of collaboration, project management and participatory processes, as well as those settings influencing the technical aspects, including modeling and computational considerations.

The *spatial* and *temporal* features of a system are often the primary aspects around which scale is traditionally considered and framed. These define the time and space of interest (both their horizons and discretization) and the events and processes that are considered important to represent (Cash et al., 2006). The spatial scales selected may be influenced by the temporal scales of interest, and vice versa. Their dependence can be intensified by the fact that spatio-temporal scales are often influenced by factors outside their defined boundaries. Such influences may be important but may not be well understood or ignored (Zhang et al., 2014b, 2014a).

*Resolution* defines the *granularity* of system representation and refers to the unit of spatial/temporal scale represented in each system. Resolution may be spatial or temporal in nature but extends in other ways such as to social units (individuals to families to communities, etc.) and thus may be represented so as to conform to a semantic or conceptual *hierarchy* (Cash et al., 2006). Choice of resolution is highly dependent on the modeling context, generally informed by the availability of data, the needs of the model (including for numerical stability, sensitivity and model identifiability), and model purpose.

*Hierarchy* and their respective *levels of organization* relate to the representation of nested relationships among systems (Ostrom, 2007). For example, various governance systems may co-exist at a range of scales with separate administrative or institutional concerns (Daniell and Barreteau, 2014). Team-based organizations are one example where the hierarchical scales may not be constrained to specific locations, with members performing a variety of roles within an organization that may be geographically spread across different time zones.

*Actors* influence and define the aspects of scale that are considered and may be both human and non-human entities which affect or influence one another. The term has its roots in the social sciences (an example may be found in Wessells, 2007). Actors have roles and carry out one or more activities in the system and can be represented individually or collectively. Human actors have attributes such as values, goals and mental models, which influence their behavior (Pahl-Wostl, 2007). Non-human actors are defined here in its literal sense (i.e. not an individual biological person) such that organizations, flora and fauna are non-human actors but may still exhibit collective culture and personalities (Hobday et al., 2018; Schneider et al., 2013). A system can encapsulate many actors and may be an actor itself.

The different types of system modeling encompass many terms that are often used interchangeably across the sciences. As alluded to in the introduction we are guided by, but do not directly adopt, definitions as applied in system-of-systems engineering (cf. Dahmann and Baldwin, 2008). Here, a single-system model targets a specific system, for instance an agricultural system without explicit representation of the hydrological dynamics or climatic influences. Consequently, single-system models constrain themselves to the concerns and considerations of a single sector. Models concerned with a single system may, of course, use several models internally (e.g. crop growth, soil water properties, etc.) and these are referred to here as *component models*.

A direct approach to representing additional systems can be accomplished by applying, albeit separately, a selection of single-system models for a given problem domain. In such cases, knowledge gained in the application of a model may inform the use of another. Data from one model may be fed into another, and vice versa, typically via manual processes. For example, a weather forecast model may be used to provide inputs to an agricultural model to determine seasonal effects on crops, and the agricultural model may provide land surface boundaries to the weather forecast model.

Multi-system representations can be *integrated* by *coupling* models together such that data interoperation occurs in an automated fashion. Individual “system level” models are then referred to as *constituent models*. The advantage of multi-system models over their single system relatives is that the impacts and feedback mechanisms can be represented across/between their individual scales (Elag et al., 2011; Tscheikner-Gratl et al., 2019; Wang et al., 2019). Multi-system models, with their explicit representation of system interactions, are therefore capable of providing more holistic assessment compared to the use of individual models in isolation (Kelly (Letcher) et al., 2013). Component-based modeling stems from Component-Based Software Engineering (Vale et al., 2016; Hutton et al., 2020) and common usage in environmental modeling typically makes no distinction between constituent and component models (e.g. Malard et al., 2017). A conscious decision has been made here to adopt the term “constituent” from the systems engineering field (Nielsen et al., 2015) to convey this distinction.

It is important to note that “integrated” and “multi-system” models could then equally apply to both single-system models with several component or constituent models. The requirement for a model to be regarded as “integrated” is that its (component or constituent) models are coupled together through the use of a common automated infrastructure to facilitate data interoperation (see for example, Malard et al., 2017 and Whelan et al., 2014). By necessity, multi-system integrated models are more complex and may involve a variety of modeling paradigms (e.g. Bayesian networks, agent-based, system dynamics, etc.) and their combinations.

An *SoS model* is then regarded here as an integrated model with constituent models. Each constituent model may be a single-system or another SoS model such that a tiered network of relationships between models is formed, with each representing a layer of abstraction. In SoS modeling, each constituent model may operate across different spatial/temporal scales, hierarchical levels, and resolutions to incorporate multiple aspects of distinctly separate (disciplinary or sectoral) domains and modeling paradigms. An SoS perspective allows, but does not prescribe, consideration of complex system properties including nonlinearities, interdependencies, feedback loops, thresholds and emergence.

*Table 1. Brief descriptions of the primary terms defined in this paper and relevant literature. Where no references are provided, the terms are assumed to be generic and widely known.*

| Term | Definition | Relevant Literature |
|---|---|---|
| Spatial/temporal | Spatial and temporal aspects define, respectively, the bounds or horizons over the space and time frame of the events and processes of interest as well as their discretization in a model. | N/A |

| | | |
|---|---|---|
| Multi-system model | A catch-all term referring to any model that represents multiple systems. | N/A |
| Emergence or emergent behavior/simplicity/complexity | Here, emergence relates to the behavior of the system and can span from simple to complex. Emergent complexity describes the complex, possibly chaotic, behavior that arises from the collective interactions of simple constituent systems, whereas emergent simplicity is the opposite. | (Bar-Yam, 1997) |
| System and System of systems | At its core a "system" refers to a collection of processes and mechanisms that may interact depending on context.<br><br>A system of systems is represented as a collection of autonomous constituent systems that give rise to collective behavior. A constituent model may, itself, be a system-of-systems model. A system-of-systems model then is an interconnected, tiered, network of models. | (Eusgeld et al., 2011; Little et al., 2019; Tranquillo, 2019) |
| Integrated model | A model itself consists of two or more separate and separable models, connected through a common computational framework to allow automated interactions between models to occur. | (van Ittersum et al., 2008; Voinov and Shugart, 2013; Whelan et al., 2014) |
| Resolution/Granularity | The represented unit of scale at which a system component is modeled (e.g. unit of distance, volume, time, social unit, etc.) | (Ewert et al., 2011; Groen et al., 2019; Neumann et al., 2019) |
| Actor | Actors are entities, both human and non-human (e.g. objects, biota, flora and fauna, institutions, and organizations), which influence the modeling, the pathways taken throughout the modeling process, and their representations within a model.<br><br>Actors may themselves be composed of actors, such that a system is an actor within a larger system (e.g. engine in a car, team within a company, etc.). Actors may influence one another through a network of relationships and be modeled as such. Actors may embody collective culture and personalities, as may be the case with teams and organizations. | (Cresswell et al., 2010; Macy and Willer, 2002; Tate, 2013; Hobday et al., 2018; Schneider et al., 2013) |
| Hierarchy/Level | The ordered linkage crossing scales, which may be spatial/temporal (neighborhood to city) or virtual/conceptual (employee and employer), and these may be nested within one another. | (Ostrom, 2007; Schweiger et al., 2020; Steinhardt and Volk, 2001) |

# 3. Scale issues to consider

Models are developed through a life cycle of various phases, each with specific considerations and steps (the "modeling cycle"; Grimm and Railsback, 2012; Hamilton et al., 2015; Jakeman et al., 2006). SoS modeling is more complex compared to 'single-system' models due to the number of people and disciplines involved as well as the dependencies between the constituent models. Similarly, management of the modeling process is made more complex, as there is not a single modeling cycle, but multiple cycles occurring asynchronously. Each actor and model may have separate objectives and purposes, priorities and differing levels of available resources not to mention the need to consider the availability of resources for the SoS modeling as a whole.

The sections below are adapted from the modeling phases identified in Badham et al. (2019) and Hamilton et al. (2015), wherein the actions undertaken in each modeling phase are described. In contrast, we identify the relevant phases within an SoS context and outline the considerations with respect to scale issues. Figure 1 depicts the high-level considerations/objectives within each phase. While the sections below are presented in a sequential manner, we stress that modeling is an iterative and concurrent process.

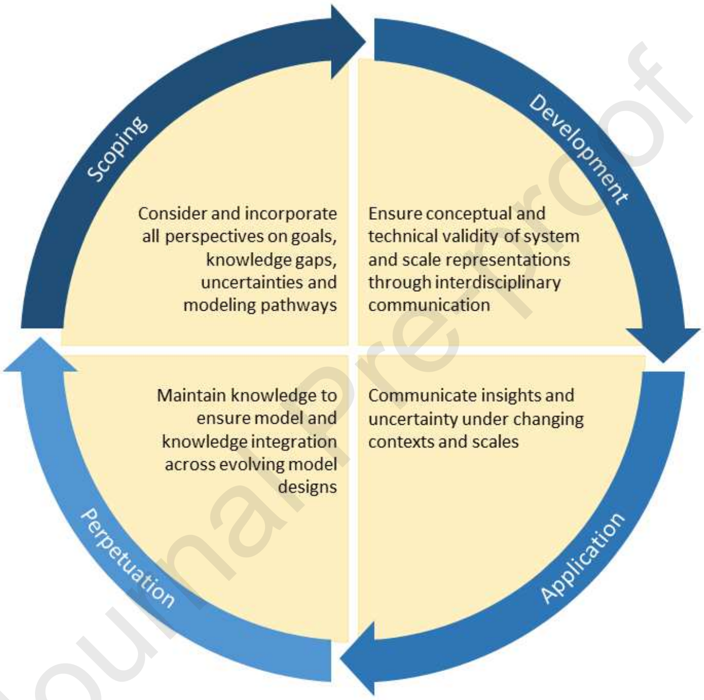



*Figure 1. The phases in the modeling cycle (adapted from Badham et al., 2019, and Hamilton et al., 2015) with key considerations within each phase.*

## 3.1. Scoping Phase

In this phase, the objectives of the modeling are clarified by defining the problem and how modeling is intended to address it. Examples of model (or modeling) purpose could be to fill gaps in knowledge, to support learning and communication processes, to validate current understandings and assumptions, to predict what might happen in the future, or to carry out scenario analysis (Badham et al., 2019; Kelly (Letcher) et al., 2013). Ideally, this scoping phase results in a clear understanding of the model types and components that need to be developed or, in later iterations, their limitations with respect to the model purpose and how to address these.

### 3.1.1. Problem definition and scoping

While the overarching purpose of the SoS model may be known, the specifics may be less clear at the outset. Development of a consistent and shared view of the scales to be considered involves communication of the scope and interactions across the constituent systems between all involved (see Figure 2). This process can aid in identifying and addressing areas that require reconciliation of different views that often exist across the stakeholders. Awareness of the scale issues will likely evolve as the modeling progresses through the iterations. The choice of modeling pathways and methodological framework employed is heavily informed by this awareness (MacLeod and Nagatsu, 2018).

Involvement of stakeholders, including domain experts, through participatory processes can inform the identification of relevant scales in the face of uncertainty and (poor) data availability (Hamilton et al., 2015; Kragt et al., 2013). Stakeholders can also play a role in selecting and combining data, furthering holistic consideration of system actors and aid in developing the model purpose. The relationship between actors and their roles in framing the scale, scope and purpose of the modeling has been previously recognized (Kragt et al., 2013; Refsgaard et al., 2007) and is further explored in the next subsection.

Insufficient consideration or agreement regarding the overarching purpose of the SoS model may ultimately affect model performance and outcomes (Connor et al., 2019). The higher number of actors in SoS modeling increases the difficulty in reconciling different or mismatched perspectives, requirements and purposes. This is a “problem of heterogeneity” (O’Connell and Todini, 1996) and is not restricted to any single discipline. Often, and by necessity, the scale of the modeling is to be commensurate with its purpose, including the level of certainty being sought, and the available resources.

Purpose and use of constituent models may be mismatched if conflicting perspectives over the scope of the modeling are not addressed. Modelers that have different goals in mind may only consider scales relevant to their immediate (and often discipline-specific) concerns, leading to an improper selection of constituent models. There is potential for a high degree of mismatch between constituent models even if modelers coordinate their efforts. Unexpected cascades of effects through scales is commonplace in complex systems (Tranquillo, 2019), and could arguably be taken as the rule rather than the exception.

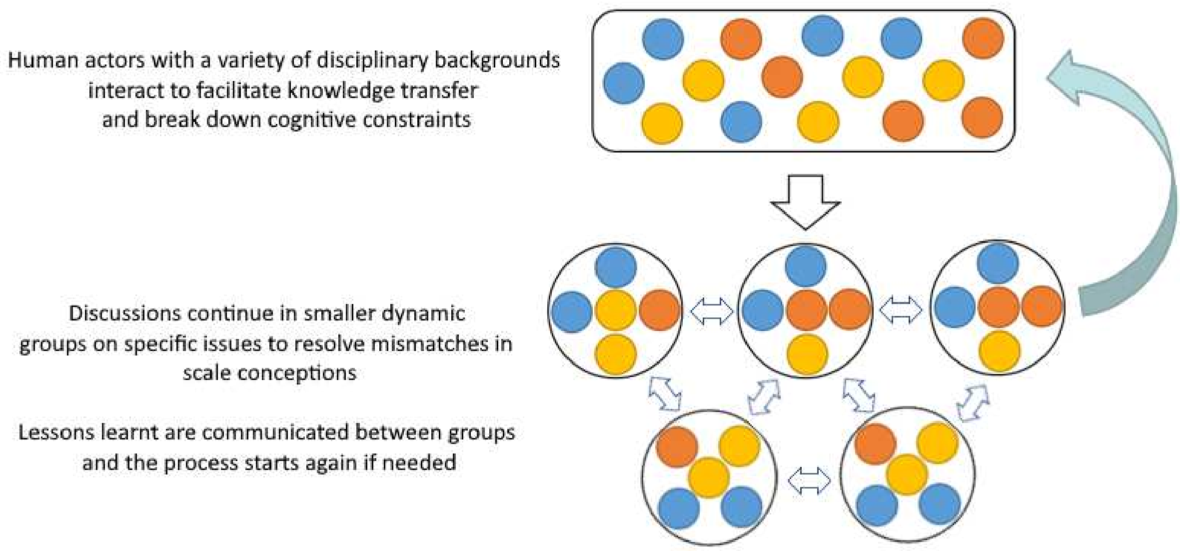


*Figure 2. Continuous and repeated interactions between human actors (domain experts, stakeholders, modelers, etc. represented by the different colored circles), and between their social groups, are necessary throughout the modeling process to ensure mismatches in system conceptualization and constituent model scales are avoided.*

Change in scale may also occur during the modeling process, due to new information that triggers a necessary change in model context. The scale of model interactions to be represented can also influence the number and type of constituent models included, and overall system complexity. The choices regarding scale then have implications for how well interactions among systems can be represented with respect to the model purpose. Scope creep, wherein the scale of the modeling is continually extended to cover contexts not originally envisioned (cf. Barton and Shan, 2017), may eventually compromise modeling efforts, as available resources get stretched too thinly to achieve sufficient progress (Sarosa and Tatnall, 2015).

Choice of scales is further compounded in cases where system bounds cannot be clearly and definitively defined. Coastal zones, atmospheric systems, and natural resource management systems are examples of systems with ambiguous system boundaries. Social systems and their dynamic structures are another example that do not have clear boundaries yet place important, even governing, conditions on system behavior. Such social systems, and their influences, are so far under-represented in current integrated assessment efforts (Zare et al., 2017). The lack of clear boundaries of such systems are often considered to be part of the problem (Voinov and Bousquet, 2010).

Reconciling conceptual differences and perspectives between human actors can be demanding but not insurmountable. There are various methods available for group decision making, such as the Delphi technique (Gokhale, 2001), which can be used to help the group reach agreement on the definition of the problem and/or the system boundaries. The subsequent modeling itself can be used to combine and reconcile different views among stakeholders, and may be useful in cross-cultural or particularly contentious settings (cf. Potter et al., 2016). The influence of modeler and stakeholder bias can also be constrained such as by using numerical optimization and/or exploratory modeling processes (Martin et al., 2017; Reichert, 2020)., The influence of personal preferences is restricted by using the exploratory approach as it focuses on identifying the relevant scales and conditions (or combinations of conditions) that normally lead to desirable outcomes.

### 3.1.2. Stakeholder Planning

Here, “stakeholder” refers to the individual or groups that may affect or be affected by the modeling or have an interest in its outcomes (Freeman, 2010). Thus, in this context, the modelers (and teams of modelers) are also stakeholders. There is a plethora of stakeholder-focused approaches (e.g. in integrated modeling, participatory modeling), but these methodologies are still limited in their capacity to deal with scale-specific questions and challenges brought by SoS modeling (Jordan et al., 2018). Generally, participatory approaches aim to bring together the multiple goals, issues, and concerns of interest from multiple scales and governance systems by developing a mutually beneficial relationship between stakeholders (Thompson, 2009). Thoughtful consideration of transparency, traceability and governance issues in engagement and participatory processes (Cockerill et al., 2019; Glynn et al., 2017) will be essential for optimizing saliency, legitimacy, ultimately, and credibility of the SoS modeling (Cash et al., 2003).

The participation of a higher diversity of stakeholders in such processes allows for a more holistic representation to be developed, covering potential ‘blind-spots’ in the system conceptualization and avoiding the “siloing” of knowledge (Hoekstra et al., 2014). Including further perspectives may increase the complexity of the modeling, and so requires careful management of individual expectations and biases (Martin et al., 2017). As the social processes influence the socio-technical, effective management of an SoS may at times be predicated on effective management of stakeholders and their level (and capacity) of involvement (Ostrom, 2007; Boone and Fragaszy, 2018).

Increases in the variety of perspectives also increases potential for conflict - defined here as disagreements of any degree - between teams, team members and/or stakeholders. On the one hand, there is evidence that conflict plays a positive role in learning and effective teamwork (Tjosvold et al., 2003). Such positive benefits, however, may only occur in cases where there are high levels of pre-existing trust within the group, and when the conflict is task-related rather than interpersonal (De Dreu, 2008). Power dynamics within teams and stakeholders therefore need to be considered (National Research Council, 2013). Identification and focus on objectives that require participants to work together (known as goal interdependence) is an identified foundation towards project success and may additionally help in avoiding conflict (Knight et al., 2001; Lee et al., 2015; Tjosvold et al., 2003). Careful design and management of interactions between teams and stakeholders requires an explicit consideration of how the multiple, and at times contradictory, objectives might align or connect. Approaches to conflict resolution and prevention (e.g. boundary critiquing, Midgley and Pinzón, 2011) are promising, but still under-utilized techniques.

Effective stakeholder engagement will in practice be impacted by geographic spread (Allen and Henn, 2006), as the realities of scheduling rarely allow all stakeholders to be engaged at the same time and place. Additionally, a diversity of stakeholders (e.g. policy makers, scientists, and the public) mean material and modes of communication may need to be tailored for each. Use of online participation platforms and technologies extends the reach to participants and are appealing for their asynchronous and distributed modes of engagement (Yearworth and White, 2018). These relatively new technologies are simply tools, however, and a capacity to both use and leverage their advantages is also required (Cooke et al., 2015). Regardless of how interactions are to occur, without documenting a Record of Engagement and Decision-making (RoED, Cockerill et al., 2019), the original purpose, assumptions, and social and biophysical context of the engagement and resulting model choices might be lost, leading to mismatches in understanding, conceptualization, and implementation. The literature is still limited on the effectiveness of using different participatory methods for different purposes and audiences (Voinov et al., 2018). Nevertheless, plans for stakeholder engagement for SoS modeling should explicitly consider the scaling challenges, and devise strategies to deal with these.

### 3.1.3. Preliminary Conceptual Model

The preliminary conceptual model represents the current understanding of the system and the relationship between constituents, including identification of key drivers, interactions and outputs of interest (Badham et al., 2019). In describing and capturing the essence of the system, development of the conceptual model helps with the design of the subsequent (computational) model as well as making concrete the model purpose. Two scale-specific aspects are to be considered here: the approach used for conceptual model development (see Table 2 for a general overview) and the formal representation (e.g. equations, technical specifications, etc.). The processes that are included or excluded based on actors' perceptions, priorities, beliefs, and values under the SoS context will inevitably influence the data leveraged, the properties of the computational model, and therefore the paths taken.

Few mapping techniques exist that focus on illustrating multi-scale representations. Scale separation maps (Hoekstra et al., 2007) or Stommel diagrams (Scholes et al., 2013) represent the scales of the constituent systems on a two dimensional space-time map. System diagrams, such as the representations used in van Delden et al. (2011) and Oxley and ApSimon (2007), organize the system components according to their spatial and/or temporal scales, and show the interactions between these components. On the other hand, coupling diagrams (Falcone et al., 2010) show the flow of data between models.

A further approach is to use the ODD protocol, named after its three blocks: Overview, Design concepts, and Details (Grimm et al., 2006). The original purpose of the ODD protocol was to describe and enable transparent communication of agent-based models (ABMs) to ensure their replication and the reproducibility of results based solely on the model description (Grimm et al., 2020). The conceptualization involved in the Overview block mandates identifying the scales of the processes or system components to ensure a shared understanding of the system being modeled. This is further complemented with the identification of relevant resolutions and spatial/temporal bounds. At this stage, the bounds can be vaguely defined (e.g. local, regional, global). This initial assessment of the scales involved may be revised throughout the modeling process as understanding improves. The ODD protocol is under continual development, and planned additions extend its consideration and applicability of use to other areas not previously considered (as outlined in Grimm et al., 2020).

If differences in conceptual understanding of the scales and their interactions cannot be reconciled at this stage, it is possible to create multiple alternative models representing the different hypotheses which can be tested in later stages of the modeling process. Such an approach can also assist in assessing uncertainty rooted in model building choices, as the treatment of scale may affect model outputs and outcomes (further discussed in Section 3.2.4). Although conceptual diagrams can be developed without specifying the scales involved, explicit consideration of scale is valuable for avoiding misinterpretation of the conceptualization and ensuring key variables and processes are included. A useful reflexive exercise, not usually reported but aiding transparency, is to identify what alternative approaches were considered, or could have been considered, and how these may have affected results and outcomes, if adopted.

*Table 2. Description of the general approaches in the development of multi-scale models, adapted from Ingram et al. (2004).*

| Approach | Description |
|---|---|
| Top-down | Creation of a coarse generalized model which is then progressively refined to an appropriate mix of scales. |

| | |
|---|---|
| Bottom-up | Models are developed at the smallest resolution initially conceptualized to be necessary and are then expanded to encompass scales as further information becomes available. |
| Middle-out | Development of the SoS model begins at the scale richest in data or information, working "outwards" towards smaller and larger scale models, as necessary. In SoS modeling, what is "richest" is likely to be subjective to each discipline and available understanding. |
| Concurrent | The process of constructing models to represent all hierarchical levels at the same time. |

## 3.2. Development Phase

### 3.2.1. Collecting data, information, and knowledge

Data, information and knowledge for each constituent model may come from the field or through literature, solicited through expert and stakeholder engagement, or collected through analysis. Considerations towards data collection in the integrated setting have been previously explored in Badham et al. (2019). Correctly communicating and interpreting data across heterogeneous systems, however, requires that the data are interoperated between constituent models and that model behavior across scales remains valid and meaningful (Renner, 2001). For this purpose, metadata serves an essential role.

Transparency in the collection process and approval from those involved in the modeling are necessary to ensure that collected data remain conceptually relevant across scales. Furthermore, transparency in the context of data collection and usage is a key factor to develop trust among stakeholders and model users, and future adoption of the constituent models (Barba, 2019; Gray and Marwick, 2019). Data may need to be transformed to be fully relevant for the context of its intended use, such as up-or-downscaling to ensure compatibility with other processes. Ideally, metadata would include information on the data collection, uncertainty and transformation process, which aids in determining the appropriateness of data for the SoS model. Explicit descriptors of both input and output data can assist in identifying the commensurate level of data collection with respect to available resources.

Modeler bias may be compounded as the choice of data collection, as well as the metadata that describes the data, may influence how system interactions are perceived, and thus conceptualized (Bhattacherjee et al., 2008). What may be considered irrelevant in one field may dictate modeling pathways in another. In an SoS setting there are many more participants involved and so there is a high degree of uncertainty stemming from the decisions made as a result.

Data quality and informativeness (e.g. accuracy or precision) provided by constituent models may also be diverse. Diversity of data obtained from a diversity of sources, however, runs the risk of conflicting information (Gray et al., 2012). Modelers from different disciplines may also utilize different scales for the same process, resulting in inconsistencies, and thus errors, the sources of which are difficult to identify. In this regard, non-quantitative sources of information, gathered from literature and/or through stakeholder engagement, may become key assets that resolve such issues (Grant and Swannack, 2007). In cases where data describing a particular linkage in an SoS model are not available, theoretical relationships, generally applicable empirical relationships, or model process and output can be useful representations for the purpose of the SoS model (Rai et al., 2002). The documentation developed in the Scoping phase can be leveraged to ensure applicability and validity with regard to the model purpose.

### 3.2.2. Construction

Construction of computational SoS models requires the marrying of domain expertise from across the various disciplines involved with technical software development knowledge. While the overarching context may be well-defined within the scoping phase, it is in this Construction step that the individual components, and the scales they represent, are developed, and coupled, tested and validated. Here, existing models may be repurposed or new models developed. The specifics of their initialization, interoperation, method of execution and management of the data involved are to be determined and prototyped in this phase (Igamberdiev et al., 2018; Madni and Sievers, 2014).

A balanced approach is needed in SoS model development that takes several factors into account. There is a danger that the models themselves become treated as pieces of software that merely require connection, ignoring the socio-technical context for their intended use (Voinov and Shugart, 2013). Another issue is the overparameterization of constituent and component models (Brun et al., 2001; Nossent and Bauwens, 2012), as simply integrating these models to form an SoS model exacerbates issues of uncertainty and identifiability (considerations of which are explored in the following sections). At the same time, ignoring the technical considerations of integration is also inadvisable (Verweij et al., 2010). Mitigating the issues that consequently arise becomes increasingly difficult as more systems and scales are included (Voinov and Shugart, 2013; Wirtz and Nowak, 2017).

Requisite systems could be represented at the level of detail necessary for the SoS model purpose through a tiered modeling structure (Little et al., 2019). Implementation of such a tiered approach can involve developing metamodels or entirely different system models. Metamodels being simplified representations of more complex models (revisited in Section 3.3). Two pertinent issues in SoS model construction are the focus below: managing the conceptual inter-connection between models, and the process of integration.

#### 3.2.2.1. Conceptual Integration

Conceptual integration of constituent models can benefit from requiring that constituent models be mechanistic as opposed to black boxes. When a model is implemented as a black box, it becomes difficult to evaluate and understand (Lorek and Sonnenschein, 1999). SoS modeling may make use of pre-existing models which constitutes re-purposing, implying the transference of the model assumptions, limitations, and scale to a new context. It is emphasized here that model suitability within its original context is not necessarily applicable to the new context (Ayllón et al., 2018; Belete et al., 2017; Voinov and Shugart, 2013). And availability of code alone, for example, does not imply transparency. What is important is the contextual information that is necessary to assess the suitability of the model purpose and functionality.

A key challenge then is ensuring the box remains open and transparent rather than closed and opaque. Opaque development can be attributed to the modular nature of constituent model development, with the teams working separately - both conceptually and geographically - and often split along disciplinary lines. Such teams can be described as self-organizing (Sletholt et al., 2012) but may lack cross-disciplinary knowledge (cross-functionality, as in Hidalgo, 2019; Hoda et al., 2013). The lack of interdisciplinary communication between teams then results in black, or at best gray, box models to those not involved in their development.

What is important in this interdisciplinary context is clear documentation and an organizational culture that supports the perpetuation of the relevant contextual knowledge. As previously mentioned in Section 3.1.3, describing the model and its conceptual linkages in a single canonical document via the ODD Protocol (introduced in Section 3.1.3) is one approach that could be leveraged. Furthermore, a "nested ODD" approach may be adopted in the case of complex SoS models wherein the constituent models may be another SoS model.

#### 3.2.2.2. Technical Integration

Technical integration refers to the correctness of model interactions, recognizing the distinction between conceptual or abstract representation (e.g. an equation or flow diagram) and its implementation as software. Successful technical integration of computational models requires the necessary engineering expertise to be available (Knapen et al., 2013). Crucial considerations are that constituent models interact and accordingly that errors will propagate (cf. Dunford et al., 2015), and that each constituent model may undergo its own separate development cycle which invariably necessitates continual adjustments to be made.

Flexibility of integration is often desirable as it allows the model to be resilient against changes in the modeling scope. Flexibility facilitates investigations into model structure (of both constituent and component models) and the technical design considerations that lead to flexibility allows for the composition of different combinations of relevant code and data represented through a nested hierarchy (e.g. 'loose coupling'; Elag et al., 2011; Vale et al., 2016; Whelan et al., 2014). Use of integration frameworks are helpful in that they allow the treatment of individual models as loose, composable, modules that provide some flexibility in dealing with the range of scales involved.

Current integration frameworks typically have their roots in specific disciplines and tend to focus on physical processes (cf. Ayllón et al., 2018). The Open Modeling Interface (OpenMI, Moore and Tindall, 2005), for example, has had to evolve from its initial focus in the hydrological sciences to accommodate an interdisciplinary modeling process (Buahin and Horsburgh, 2018). Thus, while the processes and requirements of such frameworks may be generally applicable, there remains some difficulty in their generic implementation and adoption within the interdisciplinary context of SoS modeling.

In some cases, such frameworks may be overly complex or otherwise unsuitable for the purpose and context in which the modeling is being conducted. Such difficulties may be resolved in the future as improvements to these frameworks are ongoing (Voinov and Shugart, 2013). Often modelers adopt a less formalized approach to avoid an inappropriate or constraining framework. In either case, ensuring semantic and conceptual correctness between models is typically left to the modelers themselves (cf. Hutton et al., 2020). Direct, manual, "tight-coupling" of models without the use of integration frameworks is still very much the norm.

More recent efforts include a collaborative web-based platform through which the conceptual, semantic and technical integration occurs (OpenGMS, in Chen et al., 2019). Faster feedback between participants then allows identified issues to be addressed earlier. Other approaches provide a curated ontological set of descriptors for common phenomena of interest (e.g. snowmelt or rainfall). These can be referred to as "system variables" (as in Pacheco-Romero et al., 2020) and efforts to record their quantities (e.g. centimetre, grams, etc.) and relevant operators in a specific metadata format have also been undertaken (e.g., the Standard Names, in Hobley et al., 2017). Having the inputs and outputs described and documented in such a way aids in reducing potential mismatches in later (re)use and could be used to enable later automated model coupling. Frameworks do not yet automate conversions or identify incompatible or inconsistent usage (e.g. litres per second to degrees Celsius) although this is likely to change in the near future.

Both the selected framework and constituent models may change over the course of the modeling cycle along with the scales represented. Such changes may affect its appropriateness with respect to the model purpose. For example, adoption of a particular framework or model may increase the computational requirements or necessitate changes to constituent models to allow interoperation. Inadequate consideration of the concerns and requirements of the modeling as a whole may occur in cases where cognitive constraints are still in place. The modeling process may be smoothed if requirements of the later phases are kept in mind during the design, construction (or selection) of models, and the resources allocated – including the availability of expertise – to each of these activities.

### 3.2.3. Model Calibration

Calibration is the process of tuning parameters or altering the functional forms of equations or relations to achieve desired model behavior (Bennett et al., 2013). In SoS modeling, issues such as non-identifiability and equifinality (Beven and Freer, 2001; Guillaume et al., 2019), curse of dimensionality (Bellman, 2015), computational burden (Razavi et al., 2010), and data representativeness (Beven and Westerberg, 2011; Singh and Bárdossy, 2012) may all be amplified.

Calibration implies the existence of appropriate and sufficient data to calibrate models against. Availability of data relevant for the modeling purpose is a requirement no matter how perfect the model may be. Conversely, a lack of data does not imply subsequent modeling is not useful. A model with high uncertainty may still characterize uncertainty in a way that is meaningful to decision makers, for example indicating the comparative tradeoffs between available management options (Reichert and Borsuk, 2005). Such high-uncertainty models may be helpful in determining the relative "worth" of data to be collected to better characterize uncertainty and inform future modeling or research. Such optimal experiment design approaches may also be leveraged to maximize the use of available data (Bandara et al., 2009; López–Fidalgo and Tommasi, 2018; Vanlier et al., 2014).

Arguably, model calibration within the SoS paradigm can take three general approaches: (1) calibration of each constituent model independently before integration, (2) calibration of all models together after integration, or (3) a combination thereof. The first approach is the simplest and most straightforward as each constituent model would be calibrated within its own domain (Phillips et al., 2001). While pragmatic, it ignores the effect of representing different scales across the represented SoS and system-system interactions, which in turn affects model behavior and performance of the individual constituent model. If a model is considered "calibrated" when both an acceptable level of fit and reasonable parameter values are found (as in Anderson et al., 2015), calibration in the *dis*integrated context does not necessarily transfer to the integrated context. In other words, what is "reasonable" in one context may not be so in another, and the selected parameter values may not be robust to the change in context that integration brings due to the different scales, interactions and data space involved.

The second approach is seemingly the most comprehensive approach to model calibration, as every possible interaction between models could be present in the process of model calibration (Huang et al., 2013). Interdisciplinary knowledge is leveraged to ensure calibrated values are both reasonable for the expanded operationalization. This then enriches the data space for individual constituent models and improves their performance (Jones et al., 2017). The approach, however, has the following major bottlenecks:

- The search space for model calibration will be excessively large (Ling et al., 2012). In addition, new (possibly erroneous) interaction effects might emerge between the parameters of one model with those of another model, especially with different scales of information, which makes the response surface extremely complex for model calibration. The calibration process might then become computationally cumbersome and/or infeasible.

- The available data with different scales may not be enough to properly constrain the model in the process of calibration (Ingwersen et al., 2018), as it is not identifiable from the data (Guillaume et al., 2019). There is a risk of overfitting as well, as the available data might be insufficient to produce a generalized model sufficient to cover the integrated domain.

- Expert knowledge for each model may have scale constraints and may not be easily transferable to the full SoS domain (Howard and Derek, 2016).

In the third approach, models are integrated one-at-a-time, incrementally adding complexity so that the influence of each constituent model can be directly attributed and subsequent issues can be addressed. This approach may include modifying the conceptualization as necessary and sequentially calibrating the resulting integrated configurations (Duchin, 2016; Duchin and Levine, 2019). While this approach may be as pragmatic as the first, and perhaps as comprehensive as the second, the disadvantage is the time and computational cost to perform sequential coupling and calibration. Such an approach would seem more practical in cases where there is little disciplinary friction and a relatively small number of models to be integrated.

In all approaches above, the role of expert knowledge in determining the acceptability of the calibration cannot be understated. In management contexts, for example, change in policy (e.g. the governing rulesets) may impart shifts in system behavior that may be hard to discern by examining quantitative data alone, and even more difficult to represent. Machine learning approaches may assist in identifying and representing non-stationary system behavior (e.g. Rui Wu et al., 2019; Razavi and Tolson, 2013) but still require intensive data for training and validation by experts where possible (Razavi and Tolson, 2013), and scale issues still exist between different single-system models or different levels of model integration. Such information in one system may have implications for how other constituent models are calibrated, and so interdisciplinary communication, awareness and consideration of the intertwining issues is necessary to safeguard against mismatches.

A calibration method which seems not to have been used explicitly for SoS models is pattern-oriented modeling (Grimm and Railsback, 2012; Railsback and Grimm, 2019; Wiegand et al., 2004, 2003). Here, a set of patterns observed at different scales and levels of organization is used to reject, as a set of filters, unsuitable parameter combinations and process representations, and may be closely related to the use of hydrologic signatures for (hydrological) model calibration and testing (Gupta et al., 2008). As for parameters, this approach corresponds to the rejection method in Approximate Bayesian Computing (van der Vaart et al., 2016). The basic idea is that a combination of "weak" patterns, which by themselves do not contain much information and thus would not reject many parameter combinations, can be as efficient as using a "strong" pattern, which is highly distinctive, but might not be available. For models with multiple scales, this approach holds high potential as it would help to keep both the SoS and constituent models within realistic operation spaces.

#### 3.2.4. Uncertainty analysis

SoS models often target large problem domains which necessitate complex models for their assessment and by their nature have a high degree of uncertainty. For the discussion here, we speak to the quantitative and qualitative aspects of uncertainty, which may be further classified based on their source or primary influence. Prior literature, for example, speaks of model structure, technical, parameter, scenario, contextual and predictive uncertainty (for further description, see Beven, 2009; Pianosi et al., 2016; Walker et al., 2003).

Quantitative approaches aim to measure the effect of uncertainty in a specific parameter, input or assumption on an output and allow the numerical characterization of the output distribution and therefore model behavior (Saltelli et al., 2019; Zimmermann, 2000). Qualitative uncertainty, however, cannot be characterized with a value and arises from sources such as the biases and subjective beliefs of human actors (Chen et al., 2007). Qualitative uncertainty can also arise from the modelers' subjective judgment, linguistic imprecision and disagreement across actors involved (Linkov and Burmistrov, 2003; Refsgaard et al., 2007).

One reason for increased model uncertainty in SoS modeling is the complexity that is largely a result of the increased scope of modeling, which comes with a larger number of models and people (and their perspectives) involved. The increase in the number of actors typically results in an increase in the overall number of parameters and their possible interactions (Oreskes, 2003), the number of possible decision pathways in the modeling process (Lahtinen et al., 2017), and the level of stakeholder influence at each decision fork (Ostrom, 2007).

Increasing model complexity allows for a higher-fidelity model, but can also increase the perceived uncertainty in a traditional sense; known as the complexity paradox (Oreskes, 2003). Characterizing "true" uncertainty in an SoS model, however, is impossible as it requires a model that represents everything perfectly including unknown unknowns (Hunt, 2017). Uncertainty may then compound with each interaction across constituent models in the SoS framework, propagating some amount of error (Dunford et al., 2015). Thus, it becomes progressively difficult to gain insights as to what effect and influence the combinations of these have (structural and parameter identifiability as in Bellman and Åström, 1970; Guillaume et al., 2019). High levels of model uncertainty need not be a barrier to effective decision support, however, and is ameliorated by providing estimates or assessments of such uncertainties (Reichert and Borsuk, 2005), both quantitative and qualitative. Different strategies and further considerations for uncertainty assessment are needed in SoS modeling compared to single-system modeling.

One commonly suggested approach to restricting model complexity (and possibly runtime) is to screen for insensitive parameters (Pianosi et al., 2016). Such parameters are said to have negligible influence on model output and so may be "fixed", i.e., made static in subsequent analyses, or otherwise removed from the model. Another is to "tie" related parameters so that they may be represented by a single "hyperparameter" (Raick et al., 2006). Reducing the number of parameters, however, does not necessarily equate to a reduction in uncertainty. Rather, it may simply mean consideration of an uncertainty source is determined to be unimportant for a given context or purpose (Pianosi et al., 2016), and doing so may trade off model fidelity under new unseen conditions.

Use of a constituent model within an SoS model as opposed to its individual operation, or its modification or simplification through parameter screening and tying constitutes a change in context. Therefore, parameters initially found to be influential might become inactive and non-influential (and vice versa), or the relationships that led to parameters being tied may change. The change of context also changes the relevance of the assumptions and objectives, and what constitutes an appropriate uncertainty analysis (Song et al., 2015). Uncertainty analysis conducted in one context is not valid across all scales. Thus, premature model simplification may ultimately affect the appropriateness of the SoS model for its overarching purpose. A comprehensive sensitivity analysis under current and possibly alternative conditions can provide valuable insights into a key question: "*when and how does uncertainty matter?*", as discussed in Razavi et al. (2019). An alternate view is that, given the likelihood of limited computational resources, efforts to characterize and communicate uncertainties to stakeholders may be more beneficial than an exhaustive sensitivity analysis (Reichert, 2020; Anderson et al., 2015).

An additional consideration is that a constituent model may be a legacy or third-party model that cannot be modified (e.g., due to lack of access to the underlying code). This would introduce some hidden or uncharacterized uncertainty into the SoS modelling. In this case, metamodeling (expanded on in the next subsection) might provide some help in simplifying the model.

Explicit documentation of the criteria used for each constituent model can ensure relevance of its application and reduce contextual uncertainty (see Walker et al., 2003) across all the scales involved. Accordingly, in the recent update of the ODD protocol (Grimm et al., 2020), a standard format for describing models, the element "Purpose" has been changed to "Purpose and patterns", with patterns being the multiple criteria for ensuring a model's structural realism, as defined in the "pattern-oriented" modeling strategy (Grimm, 2005; Grimm and Railsback, 2012). The effect and relative importance of model structure uncertainty may be assessed through expert and stakeholder knowledge of alternate models (van der Sluijs, 2007) and Bayesian approaches could be applied to characterize the known unknowns (Clark, 2005). Uncertainty matrices have also been suggested as a tool to qualitatively identify and document the source, type and nature of uncertainty and assess its relative priority in a table-like format (see Refsgaard et al., 2007).

Increased consideration of technical uncertainty (adopting the term from Walker et al., 2003) is another area which warrants further consideration in the SoS modeling context. Choice of what infrastructure and technologies to use is likely to stem from the prior experiences of the team(s) involved. Constituent models may be run on different infrastructure than was originally intended, especially as issues around computational reproducibility are addressed (Barba, 2019; Hutton et al., 2016). Identical code run under different computational environments may produce different results (see for example Bhandari Neupane et al., 2019). Such infrastructure may differ in physical or virtual architecture (e.g., laptop, supercomputer, or operating systems) or method of generating/interpreting code (e.g., different languages, compilers, package versions). Various combinations of these may be used and may also differ in the development and application phases. For these reasons the influences of different and interoperating infrastructure are important considerations (Iwanaga et al., 2020).

Correlation between parameters is another issue that is often ignored in the characterization and attribution of uncertainty (Do and Razavi, 2020). Correlation refers to statistical dependency between parameters. It is different from interaction effects which refer to the presence of non-additive operations among two or more factors embedded in constitutive equations of the model. In SoS modeling the issue is further escalated as possible correlations between the factors of different models needs to be accounted for. Ignoring correlations can falsify any estimation of uncertainty (Do and Razavi, 2020).

### 3.2.5. Testing and Evaluation

Testing and evaluation can assist in the assessment of the ramifications of scale choice. In this step reasonableness of model structure and interpretability of relationships within models are assessed along with the traditional analysis of model behavior. Not all outputs produced by the constituent models may be relevant for the SoS model purpose and the validity of their outputs are affected due to the integrated nature of SoS modeling. For any evaluation to be effective, the specific model outputs of interest that are relevant for the model purpose must be well understood. Outputs may be at a particular spatio-temporal scale, for instance a long-term average of a model output over a large spatial domain or an extreme event at a specific point location. Issues may also stem from the conceptual suitability of constituent models as uncertainty may be propagated throughout and may compound as more models are integrated (Dunford et al., 2015). Thus, the first step in testing and evaluation involves attempting to refute aspects of SoS model structure and functional relationships within the model based on their lack of correspondence with the represented system and the model outputs. Stakeholders could be leveraged to evaluate the conceptual alignment and appropriateness of the SoS representation at the selected scales.

Evaluation of the behavioral relationships at the integrated level is similar to scientific hypothesis testing (Wilson et al., 2017) or "conceptual testing" (Iwanaga et al., 2020) wherein functional relationships within the SoS model are examined. Such tests may be especially useful in cases where the internal workings of a model are inaccessible or otherwise unknown but expected behavior of the constituent model in the integrated context can be characterized (Iwanaga et al., 2020). These approaches can be used to identify impossible or implausible aspects of the SoS model output. If any aspect of model structure or any functional relationship within the model can be shown to be an inadequate representation of the corresponding aspects of the real system, then that particular portion of the model is refuted (Li et al., 2016). Examination of model behavior over a range of inputs will also help to expose additional inadequacies in the model (Bennett et al., 2013).

The interesting aspect in this regard is that successful testing and evaluation of the constituent models does not guarantee correctness of the SoS model and vice versa. Testing and evaluation may happen at different scale levels, and acceptable model behavior depends on the model purpose and consequent measures or indicators of interest. Model behavior of constituent models could be examined quantitatively through assessment of the intermediate data in the models to ensure their behavior is consistent with a priori expectations.

It is necessary to test the software used to interoperate data across the different hierarchical levels using relevant testing approaches. These include checking the mapping of input-outputs between models, conversion of units, use of metadata to perform semantic operations, and translation of spatial temporal dimensions (Ayllón et al., 2018; Belete et al., 2017; Voinov and Shugart, 2013). Testing processes found in software engineering may additionally aid in conducting such checks (see for example, Laukkanen et al., 2017; Verweij et al., 2010; Yoo and Harman, 2012).

It may also be possible that some data gaps or uncertainties from constituent models have a lesser or negligible effect on the SoS model depending on how the constituent model is leveraged at the SoS level. Furthermore, constituent models may present overlapping and/or conflicting data or assumptions that will only be revealed when testing and evaluating their integration. A common example is *double counting* uncertainty due to embedded assumptions in the model or failure to detect correlated variables with a common cause.

The next step focuses more specifically on the correspondence between model projections and observed data. Strictly speaking, data used in model testing and evaluation must be independent of data used to develop the model (Raick et al., 2006). A variety of visual, statistical, and machine learning methods are widely used to evaluate SoS models. The choice of method, however, should be based on the fundamental questions of what scenarios and observations to use in the evaluation. Evaluation of models under the range of conditions similar to those of interest can aid in identifying limitations of the model (Ramaswami et al., 2005).

Sensitivity analysis is now regarded as standard practice in modeling (Norton, 2015; Pianosi et al., 2016; Razavi and Gupta, 2015). The sensitivity of SoS model behavior to changes to its constituents and their interactions is the target of the assessment (Moriasi et al., 2007). An issue stemming from the likely overparameterization of constituent models is equifinality and the lack of identifiability. Equifinality refers to the phenomenon of different implementations or combinations of model structure, parameter values, and their interactions producing equally acceptable results (Wagener et al., 2003; Beven, 2006). Identifiability then refers to the ability to attribute the influence on model outputs to unique model parameters or structure (Muñoz et al., 2014; Guillaume et al., 2019). Therefore, the greater the number of parameters, the less identifiable the model becomes.

Sensitivities are assessed as part of identifiability analysis, typically by ranking parameters based on their influence on outputs which can aid in determining what parameters require focused efforts to reduce uncertainty or improve identifiability (e.g. Factor Prioritization; Nossent and Bauwens, 2012). Information from sensitivity and identifiability analysis can then aid in simplifying the model (as discussed in the previous section). Similar to what was noted in Section 3.2.3, naively applying sensitivity and identifiability analysis without consideration of the SoS context may adversely affect modeling outcomes.

Assessment of sensitivities would ideally rely on global, rather than local analyses for reasons that have been expounded in prior literature (see for example Pianosi et al., 2016; Saltelli and Annoni, 2010). Use of global sensitivity analyses in model assessment has seen increasing use, despite the lack of uptake or reported use of available software tools to conduct such analyses (Douglas-Smith et al., 2020). Still, the importance of such analyses tends to be under-appreciated (Saltelli et al., 2019).

One practical reason for the lack of global sensitivity analyses is that they are typically computationally expensive to perform and the SoS models themselves typically exhibit long runtimes. Dependencies and correlations between parameters across constituent models and their respective scales pose another challenge. Metamodeling (expanded on in the next section) along with recently developed sampling and analysis methods may be more amenable to the SoS context. Examples of such methods that warrant further investigation include moment-independent methods (such as PAWN; Pianosi and Wagener, 2015) which can be applied independent of the sampling scheme used, and variogram-based approaches (e.g. STAR-VARS; Razavi and Gupta, 2015) which can reportedly account for temporal and spatial correlations. Adaptive sampling of the parameter space, through sparse-grids for example, in combination with these analysis techniques, may also aid in reducing the computational costs associated with sensitivity and uncertainty analyses (Buzzard and Xiu, 2011; Xiong et al., 2010).

## 3.3. Application Phase

A critical aspect in the *application* of SoS models is that constituent models evolve independently. Development of each constituent model, by necessity, is led by disciplinary experts and undergoes separate, asynchronous, development cycles. As each model may come from different paradigms and sources of knowledge, the implementation may be adjusted over time or even replaced in response to newly acquired knowledge. Advancing towards trial model applications using the expected type and volume of data as early, quickly and often as possible allows modelers to encounter issues in the model application earlier in the process (Warren, 2014). Experience gained with each iteration subsequently serves to rectify and protect against future application challenges. Application of the model then requires monitoring and scrutinizing to ensure the underlying models (including their metadata, represented knowledge and application context) remain current and appropriate.

When models are integrated, the runtime may prevent practical application for its primary purpose, such as social learning through interactive use with stakeholders, or for global sensitivity analyses. One option to overcome this problem is to simplify the constituent models for the specific purpose. Doing so requires a high degree of knowledge of the constituent models, however, and may not be practical in cases where legacy models are used. Spatially explicit models can especially be a problem in regard to runtime, and a solution for reduction in computational burden may be achieved through aggregating grid cells into similar zones (e.g. groundwater model aggregated into hydraulic conductivity zones; Elsawah et al., 2017).

In cases of high runtime, replacing the most computationally expensive constituent models with metamodels may be a viable option. Metamodels approximate the input-output behavior of the original model (Castelletti et al., 2012; Christelis and Hughes, 2018; Pietzsch et al., 2020) and therefore provide simplified representation(s) of more complex models (Asher et al., 2015; Razavi et al., 2012). Metamodels leverage the emergent simplicity of complex systems and although there are a variety of methods available to accomplish this, generally metamodels require the complex models (i.e. the original constituent models) to be available beforehand. Meta models, being approximations of an original model's response surface, are most relevant to the conditions existing in the datasets upon which they are tuned, so care needs to be taken if using them under conditions that transcend those extant in the data. System forcing data beyond that currently experienced, such as climate change or groundwater extractions, are of particular concern in this regard. If possible, simply allocating more computational resources (e.g. supercomputers) may be the most pragmatic and resource efficient alternative, especially considering the time taken to investigate and implement the options listed above. It is acknowledged, however, that more computational capacity may not be available.

### 3.3.1. Analysis and Visualization

In the management context, where SoS models are typically applied, there is a need to adequately describe the level of uncertainties in the SoS model and its predictions. Individual stakeholders may react differently to uncertainties and levels of uncertainty (Cockerill et al., 2019). Presenting scenario results relative to the modeled baseline neatly reduces the inherent biases that come with relying on stakeholder preferences to inform desirable thresholds, as would usually occur in multi-criteria, or multi-objective, analysis approaches (Maier et al., 2016; Martin et al., 2017; Reichert and Borsuk, 2005). With such an approach, the acceptability of a (possible) maximum or minimum relative change becomes the focus of stakeholder discussion.

Software tooling for supporting analyses of model results (including sensitivity and uncertainty analyses) typically necessitates interaction between the analysis software and the model(s), which may require the development of additional interfaces (i.e. code or supporting software). Due to the number of models involved, the associated parameters, and the possibly dynamic model structure (Wirtz and Nowak, 2017), maintaining these interfaces in the SoS context may quickly become unwieldy. Additionally, it may be desirable to replace entire models to analyze the influence of model structure and the scales they represent (Ewert et al., 2011), thus potentially rendering existing interfaces obsolete. Recent efforts circumvent this issue by supporting the near-seamless transition between the nested hierarchical representation common in SoS design to the conceptually simpler "flat" structure expected in typical analyses (e.g. Schouten and Deits, 2020). An example of nested and flattened representations of a node network is provided in Appendix 1.

A common requirement shared with tooling for conducting analyses (e.g. for sensitivity and uncertainty analysis, and exploratory modeling) is the provision and definition of parameter values. These may consist of a "default" value, a range within which values may vary, whether these values are categorical, scalar, or regarded as constants (examples may be found in Adams et al., 2014; Kwakkel, 2017; Pianosi et al., 2015; Razavi et al., 2019). Categorical values may indicate substitution with other data types or a collection of data types (e.g. rasters, climate sequences, etc.) Such information may be the minimum necessary to conduct such analyses, to reproduce and replicate results, and to support later automation of these activities. Parameter values in effect represent dimensions of scale and the inappropriate selection of their values and ranges may result in misleading results (Shin et al., 2013; Wagener and Pianosi, 2019).

## 3.4 Perpetuation Phase

As in Badham et al., (2019), *perpetuation* is about the intended influence the modeling is to have into the future. The focus here is on the scale of documentation and process evaluation in SoS modeling which is informed by the level of consensus among stakeholders and modelers as to its purpose. In the research context, for example, there is a newfound expectation that the model be developed and provided in a manner that supports reproducibility and replicability. Reproducibility is the ability to recreate results, whereas replicability captures the ability of the model to generate new but consistent data in other applications (Patil et al., 2016).

Where SoS models are used by external stakeholders, some amount of technical support is likely expected. Without this, use of the model and thus its impact is likely to be minimal. Computational models are software in that they are made of code, and so continued use comes with a baseline cost to cover maintenance, improvements, and updating of documentation. Such capacity is crucial in contexts where long-term management and decision support is an acknowledged requirement. In such cases the design, implementation and documentation of the model should plan for these long-term activities from the beginning. In the SoS context this implies retaining the interdisciplinary knowledge within a team or organization (e.g. Cockerill et al., 2019; Kragt et al., 2013).

### 3.4.1 Documentation

Whereas earlier sections spoke to the content of documentation, this section focuses on the role of documentation in an interdisciplinary setting such as SoS modeling. Documentation is a conduit through which information and knowledge are propagated and provides the necessary context for model evaluation (Cockerill et al., 2019). Without sufficient documentation, it is difficult to understand the context that led to any specific issue, including mismatches between constituent models. Lack of context then affects the perceived validity of the model conceptualization, restricts model use, rendering the model inappropriate or invalid for its purpose.

The act of documenting itself allows for reflexive and transparent communication and for new insights to be gained. Undocumented assumptions regarding scale and their influence may compromise other constituent models, thus holistic awareness of the SoS issues can be obstructed by a lack of documentation. Long-term maintenance and use of the model may also be impeded (Ahalt et al., 2014). No individual holds the knowledge and awareness of the modeling details in their entirety, let alone the effects of interactions between models. It is therefore important to recognize that writing and maintaining documentation should be a team effort, and a culture to support this should be fostered.

In practice there are few incentives for documenting models to such an extent. A key problem in SoS model documentation is that details of the constituent models important for the SoS team may be considered unnecessary for the teams developing the constituent models. Once again, this stems from potential disconnects between the purpose of the SoS model and the individual (or original) objectives of each constituent model. In the sciences the focus is often on the publication of papers at the expense of ensuring model reuse or reproducibility and replicability (Easterbrook, 2014; Joppa et al., 2013; Peng, 2011; Schnell, 2018). There is an increasing push to change the culture surrounding the publication process, however, to better recognize, credit and incentivize model code publication. For example, a number of organizations have begun supporting "Open Code Badges" to highlight reproducible work (https://www.comses.net/resources/open-code-badge/).

### 3.4.2 Process evaluation

The extent to which the modeling has achieved its overarching purpose is evaluated in this step (Badham et al., 2019). This evaluation extends beyond the technical performance of the SoS model (Bennett et al., 2013) to consider outcomes of modeling as a social process. Success of a model depends on the beliefs and expectations of the intended users and in their satisfaction with the model and its results (Hamilton et al., 2019). It may also depend on the biases and beliefs of the model creators (Glynn et al., 2017) and in an alignment of expectations between creators and users (Sterling et al., 2019). The suitability of the success criteria is dependent on the context of the project, including not only the model purpose, but also the characteristics of the problem, such as its complexity and the resources that were available (Hamilton et al., 2019).

Process evaluation in SoS focuses on two facets: achievement of goals and longevity of the models. In terms of goal achievement, process evaluation considers whether the goals of the SoS model were supported by its constituent models and, where applicable, whether constituent models achieved their own goals. Although satisfying the goals of the constituent models may seem an indirect path to satisfying the goals of the SoS model, this interpretation is misleading. An SoS approach to modeling, instead of simply a multi-modeling approach, leverages the autonomy and independence of the constituent models. Constituent models still need to be capable of yielding their own outcomes, regardless of how those models are used in the context of the SoS model (Salado, 2015).

Evaluation of the longevity of the SoS model, referring to the ability to leverage or reuse the SoS model over time, requires the development and assessment of a targeted plan for its sustainment that includes: (1) monitoring the evolution of the constituent models; (2) identifying alternatives for models that may cease their validity, availability or accessibility during the lifetime of the SoS model; (3) establishing a strategy for the continued evolution of the SoS model, including the development of potential transformation frameworks and implementations; and (4) identifying opportunities to facilitate the sustainment of constituent systems aligned with the sustainment of the SoS model.

Process evaluation for SoS models may consider adopting a reflexive process in which questions are asked of those involved in the modeling, such as 'did the modeling process help to improve understanding of the system/problem?' or 'did the modeling process help facilitate communication between stakeholders?' (Hamilton et al., 2019). The line of questioning can then leverage input from the various perspectives available, including those of experts and stakeholders for the different constituent systems of an SoS. Bias in the model, such as whether their respective positions were adequately represented, may then be assessed. Alternative conceptions and processes of the system and their scales could also be assessed at this stage (Voinov et al., 2016).



# 4. The Paths Forward

## 4.1. A grander vision and commensurate funding

Addressing all the scale-related issues outlined in the paper requires a level of cooperation and concerted integrative effort that is by and large not possible given the usual short-term funding of the sciences (e.g. Saltelli, 2018). Recent publications have also brought attention to deficiencies in the current science resourcing structure, characterized in part by competition over limited funding and an emphasis on (number and citation counts of) publications. Existing funding mechanisms may well be detrimental to the quality of science produced (Binswanger, 2013; Sandström and Besselaar, 2018).

Limited resourcing is one reason for the multiple, albeit siloed, efforts with a focus on single case studies (Pulver et al., 2018; Hoekstra et al., 2014), and the necessity of excluding salient aspects of the modeling (such as adequate participatory processes; Eker et al., 2018) or making less than ideal choices about the model or data (e.g. using existing coarser scale data rather than collecting new data at a finer scale). Commentary by researchers highlight the importance of interdisciplinary work (Kretser et al., 2019; Meirmans et al., 2019), which is typically not funded to the same extent as monodisciplinary efforts (Kwon et al., 2017; Bromham et al., 2016). Regardless of the importance of such holistic assessments these real-world constraints essentially make holistic SoS modeling and analyses unrealistic.

On the other hand, examples of large concerted efforts can be found, such as in astronomy and physics which have produced groundbreaking work with the Event Horizon Telescope (e.g. first photograph of a blackhole, Akiyama et al., 2019) and the Large Hadron Collider (e.g. discovery of the Higgs boson, Aad et al., 2012). These resource intensive projects are important and could substantially influence future societal development. At the same time, lesser importance is placed by funding organizations on interdisciplinary socio-environmental works which arguably have a more immediate impact and benefit to society.

A grander vision for SoS research, in line with large-scale collaborations in other fields, is vital to achieve a truly holistic consideration of SoS modeling for resolving socio-environmental issues. Realizing this vision itself requires fundamental shifts in how such interdisciplinary work, and associated expertise, are viewed and funded (Elsawah et al., 2020). Greater funding focused on education and training of interdisciplinary system practitioners is fundamental for greater cohesion and consensus in the socio-environmental sciences (Little et al., 2019). While alternative funding models have been suggested for the sciences (see for example Meirmans et al., 2019; Higginson and Munafò, 2016), the current state of affairs is unlikely to change in the near future. Thus, any benefits from a systemic change, if they occur at all, will be experienced only in the long-term.

Although disciplinary experts may collaborate, pool resources, engage with stakeholders and gain experience in interdisciplinary work in the process of investigating a socio-environmental issue, this is not an effective way forward. In the medium-term, existing case studies could be leveraged to perform a comparative meta-analysis to determine the level of influence system connections have, and the scales at which such connections matter (Pulver et al., 2018). Such meta-analyses could extend to the practices used to manage the socio-technical influences in the modeling process. Shifts towards leveraging collections of studies for meta-analyses are emerging in fields such as psychology to allow for what is known as “statistical objectivity” towards reported findings in the literature (Freese and Peterson, 2018). Although the focus there is in resolving issues of replicability, the same approach can be leveraged to characterize scale commonalities.

We conclude here by re-emphasizing three key considerations which can reinforce current SoS modeling efforts in a move towards the larger consensus needed for this grander vision.

## 4.2. Strengthen interdisciplinary communication

Here lies the crux of the challenge in developing a tiered SoS model. It is not only necessary for the science and engineering to mesh together appropriately, but it is fundamental that the modeling process also consider and embed the socio-technical considerations. While we as modelers struggle with the former, the latter is too often ignored. As there are a variety of participants, and therefore disciplinary perspectives involved, a key set of considerations are in the social dimensions that provide the interface between modeling efforts.

Integrating multiple perspectives requires an integrative approach which is necessary, ultimately, to navigate towards a beneficial system change (why else do we model?). Choices made in the treatment of scale are unavoidable and may result in conflicting decisions with separate implications. Just to name one, members of teams may have a path pre-selected without full consideration of the implications on the system representations, leading to further issues when such decisions are not communicated.

The next generation of systems modelers would ideally embody a culture that is cognizant of the socio-technical issues, considerations, and their influences throughout the modeling process (e.g. Little et al., 2019). Such a systemic cultural shift can only be developed in the longer term, however, and so in the meantime clearer communication requires adequate resourcing for documenting decisions made, and code and data used, including their maintenance. Practices for the co-production of knowledge to fulfill the needs and requirements of the modeling is necessary for advances to be made (Norström et al., 2020).

There is often a preference for face-to-face meetings to facilitate the necessary level of communication but that may not always be possible. Geographic distance, scheduling conflicts, travel restrictions and other factors may preclude such activities. Communication technologies play a critical role in mitigating some aspects of the issue. For example, travel and social distancing restrictions during the COVID-19 pandemic has prohibited many teams from meeting in person, forcing reliance on technologies such as video conferencing. Regardless of the mode of communication, a team and organizational culture of consistent and continual communication is one necessity repeatedly highlighted to resolve a variety of scale issues and the conflict that may arise between actors throughout the modeling process. Incorporating knowledge beyond the bounds of one's own disciplinary training is crucial to the holistic attention to and incorporation of scales and to avoid the siloing of information and knowledge, and to break down cognitive constraints.

## 4.3. Improve documentation processes

The importance of documentation is another aspect that was repeatedly raised throughout this paper. Documentation of the modeling process communicates, and makes accessible, the decisions, actions, the context of those decisions and actions, and reflection on those choices to those who may or may not have been active participants in their making. Insufficient documentation affects many aspects from the pace of model development throughout the modeling cycle, quality of model integration especially across disciplinary boundaries, and the perceived quality of the modeling conducted. A lack of documentation accessibility additionally affects the (re)use and maintenance of the SoS model (or its constituents) and so could lead to duplication of effort across those involved in modeling SESs.

One approach to ensure that documentation is made a priority is to adopt a documentation-driven development and design approach (Heeager, 2012). Such approaches are exemplified by the ODD Protocol (Grimm et al., 2020, 2014, 2010). In this paradigm, documentation is developed first, serving as a vehicle for discussion, ideally prior to any model development (Heeager, 2012). Ambiguities in the documentation (and thus the modeling) may be addressed earlier in the process as a result, and documentation could be iteratively revised, commensurate with any changes to modeling scale. Furthermore, maintaining Records of Engagement and Decision-making (RoED, Cockerill et al., 2019) to document the process and pathway decisions were made in a context-appropriate manner may be crucial to ensuring conceptual and technical validity throughout the modeling cycle. Sufficient, rather than exhaustive, documentation to describe model context would be preferred (Ambler, 2002; Cockerill et al., 2019).

## 4.4. Explicit consideration of scale and uncertainty

There is an increasing expectation that SoS models can more completely represent processes within an SES, however, it is impossible to model everything for all purposes. Further explicit consideration of the inter-relationships between scales, choices made in representing scale, and their influence on uncertainty is paramount in the SoS context. Identifying, managing and reconciling the disparate treatment of scale is a key step towards a holistic approach, as opposed to the concurrent, but separate, processes currently applied (Cheong et al., 2012; Elsawah et al., 2020).

As noted several times throughout this paper, the socio-technical context has an inordinate influence on uncertainty. In addition to the communication and documentation considerations outlined above, an avenue for a more holistic assessment of uncertainty includes the use of robustness analysis (Grimm and Berger, 2016). In such analysis, a model with multiple systems is systematically deconstructed through forceful changes to the model parameters, structure, and process representations within each system to assess uncertainty. Use of these approaches with pattern-oriented modeling processes, which filter unsuitable representations across scales, may also be helpful in this regard (Grimm and Railsback, 2012; Gupta et al., 2008).

Additionally, qualitative and quantitative uncertainties could be jointly assessed through the representation of multiple plausible futures that stem from different sets of assumptions through exploratory approaches (Maier et al., 2016; Roberts et al., 2018; Rounsevell and Metzger, 2010). A related approach is a multi-model approach wherein an ensemble of equally plausible models are applied to identify the influence of structural and qualitative uncertainty (Matott et al., 2009; Tebaldi and Knutti, 2007; Uusitalo et al., 2015). Using an ensemble of estimates (such as the average or median of model outputs) may have the benefit of providing more robust and accurate forecasts (Willcock et al., 2020). Applying these on different computational platforms may additionally assist in identifying technical uncertainties (Iwanaga et al., 2020).

It was noted throughout this paper that the scale of the modeling itself should be commensurate with the available resources and purpose. A holistic SoS model may not be entirely possible given resource constraints, however relationships between systems can still be acknowledged and represented (albeit simplistically). Doing so allows some assessment of the uncertainties at least, and constitutes a step towards holistic SoS modeling so long as the underlying assumptions are explicitly documented (e.g. Kloprogge et al., 2011).

# Acknowledgements

This work was supported by the National Socio-Environmental Synthesis Center (SESYNC) under funding received from the National Science Foundation DBI-1639145. The primary author (Takuya Iwanaga) is supported through an Australian Government Research Training Program (AGRTP) Scholarship and a top-up scholarship from the ANU Hilda-John Endowment Fund. Hsiao-Hsuan Wang and Tomasz E. Koralewski acknowledge partial support from USDA, ARS Agreement No. 58-3091-6-035 with Texas A&M AgriLife Research, titled ‘Areawide pest management of the invasive sugarcane aphid in grain sorghum, regional population monitoring and forecasting.’ Min Chen is supported by the Key Program of NSF of China (No.41930648). John Little acknowledges partial support from NSF Award EEC 1937012. The authors would like to thank the three anonymous reviewers and Prof. Randall Hunt (USGS) for their constructive feedback and comments. The authors additionally thank Faye Duchin and Adrian Hindes for comments provided on an earlier draft.



# Appendix 1

*Example of hypothetical model inputs for a hydrological routing model provided in a nested data structure (left column) compared to a more traditional "flat" format (right column). Nested structures are arguably better suited for representing collections of data structures and their relationships (e.g. a network or graph structure) and, pragmatically, are typically more amenable to the inclusion of comments and multiple values associated with specific parameters, reducing cognitive overhead. While perhaps more readable, a disadvantage of nested representations is the additional complexity that may be perceived.*

| Nested | Flat, table-like |
|---|---|
| 406201:<br>node_type: "StreamNode"<br># Interpret as links to other nodes<br>prev_node:<br>- 406214<br>- 406000<br>next_node: 406218<br>formula_type: 1<br>node_params:<br># Default, Min, Max values<br># (assume constant if scalar)<br>d: 200.0, 150.0, 225.0<br>d2: 2.0, 1.5, 2.2<br>e: 1.0<br>f: 1.4, 1.2, 1.5<br>alpha: 0.95<br>a: 0.9<br>b: 0.1<br>initial_storage: 0.0<br>storage_coef: 2.9<br>area: 452.22 | ID, node_type, prev_node, next_node, d, d_min, d_max, d2, d2_min, d2_max, e, e_min, e_max, f, f_min, f_max, alpha, alpha_min, alpha_max, a, a_min, a_max, b, b_min, b_max, initial_storage, storage_coef, area<br>406201, "StreamNode", [406214, 406000], 406218, 200.0, 150.0, 225.0, 2.0, 1.5, 2.2, 1.0, 1.0, 1.0, 1.4, 1.2, 1.5, 0.95, 0.9, 0.9, 0.9, 0.1, 0.1, 0.1, 0.0, 2.9, 452.22 |

# References


Aad, G. et al., 2012. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. Physics Letters B 716, 1–29. https://doi.org/10.1016/j.physletb.2012.08.020

Adams, B.M., Ebeida, M.S., Eldred, M.S., Jakeman, J.D., Swiler, L.P., Stephens, J.A., Vigil, D.M., Wildey, T.M., Bohnhoff, W.J., Dalbey, K.R., 2014. Dakota, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis: Version 6.0 Developers Manual.

Ahalt, S., Band, L., Christopherson, L., Idaszak, R., Lenhardt, C., Minsker, B., Palmer, M., Shelley, M., Tiemann, M., Zimmerman, A., 2014. Water Science Software Institute: Agile and Open Source Scientific Software Development. Computing in Science & Engineering 16, 18–26. https://doi.org/10.1109/MCSE.2014.5

Akiyama, K., Alberdi, A., Alef, W., Asada, K., Azulay, R., Baczko, A.-K., Ball, D., Baloković, M., Barrett, J., Bintley, D., Blackburn, L., Boland, W., Bouman, K.L., Bower, G.C., Bremer, M., Brinkerink, C.D., Brissenden, R., Britzen, S., Broderick, A.E., Broguiere, D., Bronzwaer, T., Byun, D.-Y., Carlstrom, J.E., Chael, A., Chan, C., Chatterjee, S., Chatterjee, K., Chen, M.-T., Chen, Y., Cho, I., Christian, P., Conway, J.E., Cordes, J.M., Crew, G.B., Cui, Y., Davelaar, J., Laurentis, M.D., Deane, R., Dempsey, J., Desvignes, G., Dexter, J., Doeleman, S.S., Eatough, R.P., Falcke, H., Fish, V.L., Fomalont, E., Fraga-Encinas, R., Freeman, W.T., Friberg, P., Fromm, C.M., Gómez, J.L., Galison, P., Gammie, C.F., García, R., Gentaz, O., Georgiev, B., Goddi, C., Gold, R., Gu, M., Gurwell, M., Hada, K., Hecht, M.H., Hesper, R., Ho, L.C., Ho, P., Honma, M., Huang, C.-W.L., Huang, L., Hughes, D.H., Ikeda, S., Inoue, M., Issaoun, S., James, D.J., Jannuzi, B.T., Janssen, M., Jeter, B., Jiang, W., Johnson, M.D., Jorstad, S., Jung, T., Karami, M., Karuppusamy, R., Kawashima, T., Keating, G.K., Kettenis, M., Kim, J.-Y., Kim, Junhan, Kim, Jongsoo, Kino, M., Koay, J.Y., Koch, P.M., Koyama, S., Kramer, M., Kramer, C., Krichbaum, T.P., Kuo, C.-Y., Lauer, T.R., Lee, S.-S., Li, Y.-R., Li, Z., Lindqvist, M., Liu, K., Liuzzo, E., Lo, W.-P., Lobanov, A.P., Loinard, L., Lonsdale, C., Lu, R.-S., MacDonald, N.R., Mao, J., Markoff, S., Marrone, D.P., Marscher, A.P., Martí-Vidal, I., Matsushita, S., Matthews, L.D., Medeiros, L., Menten, K.M., Mizuno, Y., Mizuno, I., Moran, J.M., Moriyama, K., Moscibrodzka, M., Müller, C., Nagai, H., Nagar, N.M., Nakamura, M., Narayan, R., Narayanan, G., Natarajan, I., Neri, R., Ni, C., Noutsos, A., Okino, H., Olivares, H., Oyama, T., Özel, F., Palumbo, D.C.M., Patel, N., Pen, U.-L., Pesce, D.W., Piétu, V., Plambeck, R., PopStefanija, A., Porth, O., Prather, B., Preciado-López, J.A., Psaltis, D., Pu, H.-Y., Ramakrishnan, V., Rao, R., Rawlings, M.G., Raymond, A.W., Rezzolla, L., Ripperda, B., Roelofs, F., Rogers, A., Ros, E., Rose, M., Roshanineshat, A., Rottmann, H., Roy, A.L., Ruszczyk, C., Ryan, B.R., Rygl, K.L.J., Sánchez, S., Sánchez-Arguelles, D., Sasada, M., Savolainen, T., Schloerb, F.P., Schuster, K.-F., Shao, L., Shen, Z., Small, D., Sohn, B.W., SooHoo, J., Tazaki, F., Tiede, P., Tilanus, R.P.J., Titus, M., Toma, K., Torne, P., Trent, T., Trippe, S., Tsuda, S., Bemmel, I. van, Langevelde, H.J. van, Rossum, D.R. van, Wagner, J., Wardle, J., Weintroub, J., Wex, N., Wharton, R., Wielgus, M., Wong, G.N., Wu, Q., Young, A., Young, K., Younsi, Z., Yuan, F., Yuan, Y.-F., Zensus, J.A., Zhao, G., Zhao, S.-S., Zhu, Z., Farah, J.R., Meyer-Zhao, Z., Michalik, D., Nadolski, A., Nishioka, H., Pradel, N., Primiani, R.A., Souccar, K., Vertatschitsch, L., Yamaguchi, P., 2019. First M87 Event Horizon Telescope Results. IV. Imaging the Central Supermassive Black Hole. ApJL 875, L4. https://doi.org/10.3847/2041-8213/ab0e85

Allen, T.J., Henn, G.W., 2006. The Organization and Architecture of Innovation: Managing the Flow of Technology. Routledge.

Ambler, S., 2002. Agile Modeling: Effective Practices for eXtreme Programming and the Unified Process. John Wiley & Sons, Inc., New York.

Anderson, M.P., Woessner, W.W., Hunt, R.J., 2015. Chapter 9 - Model Calibration: Assessing Performance, in: Anderson, M.P., Woessner, W.W., Hunt, R.J. (Eds.), Applied Groundwater Modeling (Second Edition). Academic Press, San Diego, pp. 375–441. https://doi.org/10.1016/B978-0-08-091638-5.00009-2

Asher, M.J., Croke, B.F.W., Jakeman, A.J., Peeters, L.J.M., 2015. A review of surrogate models and their application to groundwater modeling. Water Resources Research 51, 5957–5973. https://doi.org/10.1002/2015WR016967

Ayllón, D., Grimm, V., Attinger, S., Hauhs, M., Simmer, C., Vereecken, H., Lischeid, G., 2018. Cross-disciplinary links in environmental systems science: Current state and claimed needs identified in a meta-review of process models. Science of The Total Environment 622–623, 954–973. https://doi.org/10.1016/j.scitotenv.2017.12.007

Badham, J., Elsawah, S., Guillaume, J.H.A., Hamilton, S.H., Hunt, R.J., Jakeman, A.J., Pierce, S.A., Snow, V.O., Babbar-Sebens, M., Fu, B., Gober, P., Hill, M.C., Iwanaga, T., Loucks, D.P., Merritt, W.S., Peckham, S.D., Richmond, A.K., Zare, F., Ames, D., Bammer, G., 2019. Effective modeling for Integrated Water Resource Management: A guide to contextual practices by phases and steps and future opportunities. Environmental Modelling & Software 116, 40–56. https://doi.org/10.1016/j.envsoft.2019.02.013

Bandara, S., Schlöder, J.P., Eils, R., Bock, H.G., Meyer, T., 2009. Optimal Experimental Design for Parameter Estimation of a Cell Signaling Model. PLOS Computational Biology 5, e1000558. https://doi.org/10.1371/journal.pcbi.1000558

Banerjee, P., Wang, H.-H., Peterson, M.J., Grant, W.E., Peterson, T.R., 2019. Collaborative Modeling and Social Learning in the Context of Joint Forest Management in East Sikkim, India. Front. Environ. Sci. 7. https://doi.org/10.3389/fenvs.2019.00154

Barba, L.A., 2019. Praxis of Reproducible Computational Science. Computing in Science Engineering 21, 73–78. https://doi.org/10.1109/MCSE.2018.2881905

Barton, W., Shan, X., 2017. Environmental, health, and safety department reorganization and prioritization: Using analytical hierarchy process to mitigate scope creep. Environmental Quality Management 26, 31–45. https://doi.org/10.1002/tqem.21505

Bar-Yam, Y., 1997. Dynamics of complex systems. Perseus Books, USA.

Belete, G.F., Voinov, A., Laniak, G.F., 2017. An overview of the model integration process: From pre-integration assessment to testing. Environmental Modelling & Software 87, 49–63. https://doi.org/10.1016/j.envsoft.2016.10.013

Bellman, R., Åström, K.J., 1970. On structural identifiability. Mathematical Biosciences 7, 329–339. https://doi.org/10.1016/0025-5564(70)90132-X

Bellman, R.E., 2015. Adaptive Control Processes. Princeton University Press, Princeton.

Bennett, N.D., Croke, B.F.W., Guariso, G., Guillaume, J.H.A., Hamilton, S.H., Jakeman, A.J., Marsili-Libelli, S., Newham, L.T.H., Norton, J.P., Perrin, C., Pierce, S.A., Robson, B., Seppelt, R., Voinov, A.A., Fath, B.D., Andreassian, V., 2013. Characterising performance of environmental models. Environmental Modelling & Software 40, 1–20.

Beven, K., 2006. A manifesto for the equifinality thesis. Journal of Hydrology, The model parameter estimation experiment 320, 18–36. https://doi.org/10.1016/j.jhydrol.2005.07.007

Beven, K., Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. Journal of Hydrology 249, 11–29. https://doi.org/10.1016/S0022-1694(01)00421-8

Beven, K., Westerberg, I., 2011. On red herrings and real herrings: disinformation and information in hydrological inference. Hydrological Processes 25, 1676–1680. https://doi.org/10.1002/hyp.7963

Beven, K.J., 2009. Environmental modelling: an uncertain future?: an introduction to techniques for uncertainty estimation in environmental prediction. Routledge, London ; New York.

Bhandari Neupane, J., Neupane, R.P., Luo, Y., Yoshida, W.Y., Sun, R., Williams, P.G., 2019. Characterization of Leptazolines A–D, Polar Oxazolines from the Cyanobacterium Leptolyngbya sp., Reveals a Glitch with the “Willoughby–Hoye” Scripts for Calculating NMR Chemical Shifts. Org. Lett. 21, 8449–8453. https://doi.org/10.1021/acs.orglett.9b03216

Bhattacherjee, A., Perols, J., Sanford, C., 2008. Information Technology Continuance: A Theoretic Extension and Empirical Test. Journal of Computer Information Systems 49, 17–26. https://doi.org/10.1080/08874417.2008.11645302

Binswanger, M., 2014. Excellence by Nonsense: The Competition for Publications in Modern Science, in: Bartling, S., Friesike, S. (Eds.), Opening Science: The Evolving Guide on How the Internet Is Changing Research, Collaboration and Scholarly Publishing. Springer International Publishing, Cham, pp. 49–72. https://doi.org/10.1007/978-3-319-00026-8_3

Bridle, H., Vrieling, A., Cardillo, M., Araya, Y., Hinojosa, L., 2013. Preparing for an interdisciplinary future: A perspective from early-career researchers. Futures 53, 22–32. https://doi.org/10.1016/j.futures.2013.09.003

Bromham, L., Dinnage, R., Hua, X., 2016. Interdisciplinary research has consistently lower funding success. Nature 534, 684–687. https://doi.org/10.1038/nature18315

Brun, R., Reichert, P., Künsch, H.R., 2001. Practical identifiability analysis of large environmental simulation models. Water Resources Research 37, 1015–1030. https://doi.org/10.1029/2000WR900350

Buahin, C.A., Horsburgh, J.S., 2018. Advancing the Open Modeling Interface (OpenMI) for integrated water resources modeling. Environmental Modelling & Software 108, 133–153. https://doi.org/10.1016/j.envsoft.2018.07.015

Buzzard, G.T., Xiu, D., 2011. Variance-Based Global Sensitivity Analysis via Sparse-Grid Interpolation and Cubature. Communications in Computational Physics 9, 542–567. https://doi.org/10.4208/cicp.230909.160310s

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. Ecology and Society 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. PNAS 100, 8086–8091. https://doi.org/10.1073/pnas.1231332100

Castelletti, A., Galelli, S., Ratto, M., Soncini-Sessa, R., Young, P.C., 2012. A general framework for Dynamic Emulation Modelling in environmental problems. Environmental Modelling & Software, Emulation techniques for the reduction and sensitivity analysis of complex environmental models 34, 5–18. https://doi.org/10.1016/j.envsoft.2012.01.002

Chen, C.-F., Ma, H., Reckhow, K.H., 2007. Assessment of water quality management with a systematic qualitative uncertainty analysis. Science of The Total Environment 374, 13–25. https://doi.org/10.1016/j.scitotenv.2006.12.027

Chen, M., Yue, S., Lü, G., Lin, H., Yang, C., Wen, Y., Hou, T., Xiao, D., Jiang, H., 2019. Teamwork-oriented integrated modeling method for geo-problem solving. Environmental Modelling & Software 119, 111–123. https://doi.org/10.1016/j.envsoft.2019.05.015

Cheong, S.-M., Brown, D.G., Kok, K., Lopez-Carr, D., 2012. Mixed methods in land change research: towards integration. Transactions of the Institute of British Geographers 37, 8–12. https://doi.org/10.1111/j.1475-5661.2011.00482.x

Christelis, V., Hughes, A.G., 2018. Metamodel-assisted analysis of an integrated model composition: An example using linked surface water – groundwater models. Environmental Modelling & Software 107, 298–306. https://doi.org/10.1016/j.envsoft.2018.05.004

Clark, J.S., 2005. Why environmental scientists are becoming Bayesians. Ecology Letters, Vol. 8: 2-14.

Cockerill, K., Glynn, P., Chabay, I., Farooque, M., Hämäläinen, R.P., Miyamoto, B., McKay, P., 2019. Records of engagement and decision making for environmental and socio-ecological challenges. EURO J Decis Process. https://doi.org/10.1007/s40070-019-00104-6

Connor, T.D., Vin'a, A.R., Winkler, J.A., Hull, V., Tang, Y., Shortridge, A.M., Yang, H., Zhao, Z., Wang, F., Zhang, J., Zhang, Z., Zhou, C.-Q., Bai, W., Liu, J., 2019. Interactive spatial scale effects on species distribution modeling: The case of the giant panda. Scientific Reports 9. https://doi.org/10.1038/s41598-019-50953-z

Cooke, N.J., Hilton, M.L., Science, C. on the S. of T., Board on Behavioral, C., Education, D. of B. and S.S. and, Council, N.R., 2015. Supporting Virtual Collaboration, Enhancing the Effectiveness of Team Science. National Academies Press (US).

Cresswell, K.M., Worth, A., Sheikh, A., 2010. Actor-Network Theory and its role in understanding the implementation of information technology developments in healthcare. BMC Medical Informatics and Decision Making 10, 67. https://doi.org/10.1186/1472-6947-10-67

Dahmann, J.S., Baldwin, K.J., 2008. Understanding the Current State of US Defense Systems of Systems and the Implications for Systems Engineering, in: 2008 2nd Annual IEEE Systems Conference. Presented at the 2008 2nd Annual IEEE Systems Conference, pp. 1–7. https://doi.org/10.1109/SYSTEMS.2008.4518994

Daniell, K.A., Barreteau, O., 2014. Water governance across competing scales: Coupling land and water management. Journal of Hydrology, Water governance across competing scales: Coupling land and water management 519, 2367–2380. https://doi.org/10.1016/j.jhydrol.2014.10.055

De Dreu, C.K.W., 2008. The virtue and vice of workplace conflict: food for (pessimistic) thought. Journal of Organizational Behavior 29, 5–18. https://doi.org/10.1002/job.474

Do, N.C., Razavi, S., 2020. Correlation Effects? A Major but Often Neglected Component in Sensitivity and Uncertainty Analysis. Water Resources Research 56, e2019WR025436. https://doi.org/10.1029/2019WR025436

Douglas-Smith, D., Iwanaga, T., Croke, B.F.W., Jakeman, A.J., 2020. Certain trends in uncertainty and sensitivity analysis: An overview of software tools and techniques. Environmental Modelling & Software 124, 104588. https://doi.org/10.1016/j.envsoft.2019.104588

Duchin, F., 2016. A Global Case-Study Framework Applied to Water Supply and Sanitation. Journal of Industrial Ecology 20, 387–395. https://doi.org/10.1111/jiec.12462

Duchin, F., Levine, S.H., 2019. The recovery of products and materials for reuse: The global context of resource management. Resources, Conservation and Recycling 145, 422–447. https://doi.org/10.1016/j.resconrec.2018.10.028

Dunford, R., Harrison, P.A., Rounsevell, M.D.A., 2015. Exploring scenario and model uncertainty in cross-sectoral integrated assessment approaches to climate change impacts. Climatic Change 132, 417–432. https://doi.org/10.1007/s10584-014-1211-3

Easterbrook, S.M., 2014. Open code for open science? Nature Geosci 7, 779–781. https://doi.org/10.1038/ngeo2283

Eker, S., Rovenskaya, E., Obersteiner, M., Langan, S., 2018. Practice and perspectives in the validation of resource management models. Nature Communications 9, 5359. https://doi.org/10.1038/s41467-018-07811-9

Elag, M.M., Goodall, J.L., Castronova, A.M., 2011. Feedback loops and temporal misalignment in component-based hydrologic modeling. Water Resources Research 47. https://doi.org/10.1029/2011WR010792 10.1016/j.envsoft.2010.01.003; Clements, P.C., From subroutines to subsystems: Component-based software development (1995) 17th International Conference on Software Engineering, , paper presented IEEE Comp. Soc., Seattle, Wash; Dozier, J., Opportunities to improve hydrologic data (1992) Reviews of Geophysics, 30 (4), pp. 315-331; Ewert, F., A methodology for enhanced flexibility of integrated assessment in agriculture (2009) Environ. Sci. Policy, 12 (5), pp. 546-561; Fotopou

Elsawah, S., Filatova, T., Jakeman, A.J., Kettner, A.J., Zellner, M.L., Athanasiadis, I.N., Hamilton, S.H., Axtell, R.L., Brown, D.G., Gilligan, J.M., Janssen, M.A., Robinson, D.T., Rozenberg, J., Ullah, I.I.T., Lade, S.J., 2020. Eight grand challenges in socio-environmental systems modeling. 1 2, 16226–16226. https://doi.org/10.18174/sesmo.2020a16226

Elsawah, S., Pierce, S.A., Hamilton, S.H., van Delden, H., Haase, D., Elmahdi, A., Jakeman, A.J., 2017. An overview of the system dynamics process for integrated modelling of socio-ecological systems:

Lessons on good modelling practice from five case studies. Environmental Modelling & Software 93, 127–145. https://doi.org/10.1016/j.envsoft.2017.03.001

Elshafei, Y., Sivapalan, M., Tonts, M., Hipsey, M.R., 2014. A prototype framework for models of socio-hydrology: identification of key feedback loops and parameterisation approach. Hydrology and Earth System Sciences 18, 2141–2166. https://doi.org/10.5194/hess-18-2141-2014

Eusgeld, I., Nan, C., Dietz, S., 2011. "System-of-systems" approach for interdependent critical infrastructures. Reliability Engineering & System Safety, ESREL 2009 Special Issue 96, 679–686. https://doi.org/10.1016/j.ress.2010.12.010

Ewert, F., van Ittersum, M.K., Heckelei, T., Therond, O., Bezlepkina, I., Andersen, E., 2011. Scale changes and model linking methods for integrated assessment of agri-environmental systems. Agriculture, Ecosystems & Environment, Scaling methods in integrated assessment of agricultural systems 142, 6–17. https://doi.org/10.1016/j.agee.2011.05.016

Falcone, J.-L., Chopard, B., Hoekstra, A., 2010. MML: towards a Multiscale Modeling Language. Procedia Computer Science, ICCS 2010 1, 819–826. https://doi.org/10.1016/j.procs.2010.04.089

Freeman, R.E., 2010. Strategic management: a stakeholder approach, Reissue. ed. Cambridge University Press, Cambridge New York Melbourne Madrid Cape Town Singapore.

Freese, J., Peterson, D., 2018. The Emergence of Statistical Objectivity: Changing Ideas of Epistemic Vice and Virtue in Science: Sociological Theory. https://doi.org/10.1177/0735275118794987

Fridman, D., Kissinger, M., 2019. A multi-scale analysis of interregional sustainability: Applied to Israel's food supply. Science of The Total Environment 676, 524–534. https://doi.org/10.1016/j.scitotenv.2019.04.054

Gibson, C.C., Ostrom, E., Ahn, T.K., 2000. The concept of scale and the human dimensions of global change: a survey. Ecological Economics 32, 217–239. https://doi.org/10.1016/S0921-8009(99)00092-0

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. Earth's Future 5, 356–378. https://doi.org/10.1002/2016EF000487

Gokhale, A.A., 2001. Environmental Initiative Prioritization with a Delphi Approach: A Case Study. Environmental Management 28, 187–193. https://doi.org/10.1007/s002670010217

Gorddard, R., Colloff, M.J., Wise, R.M., Ware, D., Dunlop, M., 2016. Values, rules and knowledge: Adaptation as change in the decision context. Environmental Science & Policy 57, 60–69. https://doi.org/10.1016/j.envsci.2015.12.004

Grant, W.E., Swannack, T.M., 2007. Ecological Modeling: A Common-Sense Approach to Theory and Practice. Wiley-Blackwell, Malden, Massachusetts.

Gray, C.T., Marwick, B., 2019. Truth, Proof, and Reproducibility: There's No Counter-Attack for the Codeless, in: Nguyen, H. (Ed.), Statistics and Data Science, Communications in Computer and Information Science. Springer, Singapore, pp. 111–129. https://doi.org/10.1007/978-981-15-1960-4_8

Gray, S., Chan, A., Clark, D., Jordan, R., 2012. Modeling the integration of stakeholder knowledge in social–ecological decision-making: Benefits and limitations to knowledge diversity. Ecological Modelling, Modeling Human Decisions 229, 88–96. https://doi.org/10.1016/j.ecolmodel.2011.09.011

Grimm, V., Augusiak, J., Focks, A., Frank, B.M., Gabsi, F., Johnston, A.S.A., Liu, C., Martin, B.T., Meli, M., Radchuk, V., Thorbek, P., Railsback, S.F., 2014. Towards better modelling and decision support: Documenting model development, testing, and analysis using TRACE. Ecological Modelling, Population Models for Ecological Risk Assessment of Chemicals 280, 129–139. https://doi.org/10.1016/j.ecolmodel.2014.01.018

Grimm, V., Berger, U., 2016. Robustness analysis: Deconstructing computational models for ecological theory and applications. Ecological Modelling, Next generation ecological modelling, concepts, and theory: structural realism, emergence, and predictions 326, 162–167. https://doi.org/10.1016/j.ecolmodel.2015.07.018

Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., Grand, T., Heinz, S.K., Huse, G., Huth, A., Jepsen, J.U., Jørgensen, C., Mooij, W.M., Müller, B., Pe'er, G., Piou, C., Railsback, S.F., Robbins, A.M., Robbins, M.M., Rossmanith, E., Rüger, N., Strand, E., Souissi, S., Stillman, R.A., Vabø, R., Visser, U., DeAngelis, D.L., 2006. A standard protocol for describing individual-based and agent-based models. Ecological Modelling 198, 115–126. https://doi.org/10.1016/j.ecolmodel.2006.04.023

Grimm, V., Berger, U., DeAngelis, D.L., Polhill, J.G., Giske, J., Railsback, S.F., 2010. The ODD protocol: A review and first update. Ecological Modelling 221, 2760–2768. https://doi.org/10.1016/j.ecolmodel.2010.08.019

Grimm, V., Railsback, S.F., 2012. Pattern-oriented modelling: a 'multi-scope' for predictive systems ecology. Phil. Trans. R. Soc. B 367, 298–310. https://doi.org/10.1098/rstb.2011.0180

Grimm, V., Railsback, S.F., Vincenot, C.E., Berger, U., Gallagher, C., DeAngelis, D.L., Edmonds, B., Ge, J., Giske, J., Groeneveld, J., Johnston, A.S.A., Milles, A., Nabe-Nielsen, J., Polhill, J.G., Radchuk, V., Rohwäder, M.-S., Stillman, R.A., Thiele, J.C., Ayllón, D., 2020. The ODD protocol for describing agent-based and other simulation models: a second update to improve clarity, replication, and structural realism. JASSS 23, 7. https://doi.org/10.18564/jasss.4259

Groen, D., Knap, J., Neumann, P., Suleimenova, D., Lourens, V., Kenneth, L., 2019. Mastering the scales: a survey on the benefits of multiscale computing software. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 377, 20180147. https://doi.org/10.1098/rsta.2018.0147

Guillaume, J.H.A., Jakeman, J.D., Marsili-Libelli, S., Asher, M., Brunner, P., Croke, B., Hill, M.C., Jakeman, A.J., Keesman, K.J., Razavi, S., Stigter, J.D., 2019. Introductory overview of identifiability analysis: A guide to evaluating whether you have the right type of data for your modeling purpose. Environmental Modelling & Software 119, 418–432. https://doi.org/10.1016/j.envsoft.2019.07.007

Gupta, H.V., Wagener, T., Liu, Y., 2008. Reconciling theory with observations: elements of a diagnostic approach to model evaluation. Hydrological Processes 22, 3802–3813. https://doi.org/10.1002/hyp.6989

Hamilton, S.H., ElSawah, S., Guillaume, J.H.A., Jakeman, A.J., Pierce, S.A., 2015. Integrated assessment and modelling: Overview and synthesis of salient dimensions. Environmental Modelling & Software 64, 215–229. https://doi.org/10.1016/j.envsoft.2014.12.005

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. Environmental Modelling & Software 118, 83–98. https://doi.org/10.1016/j.envsoft.2019.04.008

Hannah, D.M., Wood, P.J., Sadler, J.P., 2004. Ecohydrology and hydroecology: A 'new paradigm'? Hydrological Processes 18, 3439–3445. https://doi.org/10.1002/hyp.5761

Häyhä, T., Lucas, P.L., van Vuuren, D.P., Cornell, S.E., Hoff, H., 2016. From Planetary Boundaries to national fair shares of the global safe operating space — How can the scales be bridged? Global Environmental Change 40, 60–72. https://doi.org/10.1016/j.gloenvcha.2016.06.008

Heeager, L.T., 2012. Introducing Agile Practices in a Documentation-Driven Software Development Practice: A Case Study. Journal of Information Technology Case and Application Research 14, 3–24. https://doi.org/10.1080/15228053.2012.10845693

Hidalgo, E.S., 2019. Adapting the scrum framework for agile project management in science: case study of a distributed research initiative. Heliyon 5, e01447. https://doi.org/10.1016/j.heliyon.2019.e01447

Higginson, A.D., Munafò, M.R., 2016. Current Incentives for Scientists Lead to Underpowered Studies with Erroneous Conclusions. PLOS Biology 14, e2000995. https://doi.org/10.1371/journal.pbio.2000995

Hobday, A.J., Doerr, V.A.J., Marshall, N.A., Cvitanovic, C., Lim-Camacho, L., 2018. Adapting to climate change: the role of organisational personalities in natural resource management. Reg Environ Change 18, 561–571. https://doi.org/10.1007/s10113-017-1227-0

Hobley, D.E.J., Adams, J.M., Nudurupati, S.S., Hutton, E.W.H., Gasparini, N.M., Istanbulluoglu, E., Tucker, G.E., 2017. Creative computing with Landlab: an open-source toolkit for building, coupling, and exploring two-dimensional numerical models of Earth-surface dynamics. Earth Surface Dynamics 5, 21–46. https://doi.org/10.5194/esurf-5-21-2017

Hoda, R., Noble, J., Marshall, S., 2010. Organizing self-organizing teams, in: Proceedings of the 32nd ACM/IEEE International Conference on Software Engineering - ICSE ’10. Presented at the the 32nd ACM/IEEE International Conference, ACM Press, Cape Town, South Africa, p. 285. https://doi.org/10.1145/1806799.1806843

Hoekstra, A., Chopard Bastien, Coveney Peter, 2014. Multiscale modelling and simulation: a position paper. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 372, 20130377. https://doi.org/10.1098/rsta.2013.0377

Hoekstra, A.G., Lorenz, E., Falcone, J.-L., Chopard, B., 2007. Towards a Complex Automata Framework for Multi-scale Modeling: Formalism and the Scale Separation Map, in: Shi, Y., van Albada, G.D., Dongarra, J., Sloot, P.M.A. (Eds.), Computational Science – ICCS 2007, Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 922–930.

Howard, P.H., Derek, S., 2016. The Wisdom of the Economic Crowd: Calibrating Integrated Assessment Models Using Consensus (2016 Annual Meeting, July 31-August 2, Boston, Massachusetts No. 235639). Agricultural and Applied Economics Association. https://doi.org/10.22004/ag.econ.235639

Huang, Y., Yu, D., Gong, Y., Liu, C., 2013. Semi-supervised GMM and DNN acoustic model training with multi-system combination and confidence re-calibration, in: INTERSPEECH.

Hunt, R.J., 2017. Applied Uncertainty. Groundwater 55, 771–772. https://doi.org/10.1111/gwat.12604

Hutton, C., Wagener, T., Freer, J., Han, D., Duffy, C., Arheimer, B., 2016. Most computational hydrology is not reproducible, so is it really science? Water Resources Research 52, 7548–7555. https://doi.org/10.1002/2016WR019285

Hutton, E.W.H., Piper, M.D., Tucker, G.E., 2020. The Basic Model Interface 2.0: A standard interface for coupling numerical models in the geosciences. Journal of Open Source Software 1, 2317. https://doi.org/10.21105/joss.02317

Igamberdiev, M., Grossmann, G., Selway, M., Stumptner, M., 2018. An integrated multi-level modeling approach for industrial-scale data interoperability. Softw Syst Model 17, 269–294. https://doi.org/10.1007/s10270-016-0520-6

Ingram, G.D., Cameron, I.T., Hangos, K.M., 2004. Classification and analysis of integrating frameworks in multiscale modelling. Chemical Engineering Science 59, 2171–2187. https://doi.org/10.1016/j.ces.2004.02.010

Ingwersen, J., Högy, P., Wizemann, H.D., Warrach-Sagi, K., Streck, T., 2018. Coupling the land surface model Noah-MP with the generic crop growth model Gecros: Model description, calibration and validation. Agricultural and forest meteorology v. 262, 322–339. https://doi.org/10.1016/j.agrformet.2018.06.023

Iwanaga, T., Partington, D., Ticehurst, J., Croke, B.F.W., Jakeman, A.J., 2020. A socio-environmental model for exploring sustainable water management futures: Participatory and collaborative modelling in the Lower Campaspe catchment. Journal of Hydrology: Regional Studies 28, 100669. https://doi.org/10.1016/j.ejrh.2020.100669

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

Jones, J.W., Antle, J.M., Basso, B., Boote, K.J., Conant, R.T., Foster, I., Godfray, H.C.J., Herrero, M., Howitt, R.E., Janssen, S., Keating, B.A., Munoz-Carpena, R., Porter, C.H., Rosenzweig, C., Wheeler, T.R., 2017. Toward a new generation of agricultural system data, models, and knowledge products: State of agricultural systems science. Agricultural Systems 155, 269–288. https://doi.org/10.1016/j.agsy.2016.09.021

Joppa, L.N., McInerny, G., Harper, R., Salido, L., Takeda, K., O'Hara, K., Gavaghan, D., Emmott, S., 2013. Troubling Trends in Scientific Software Use. Science 340, 814–815. https://doi.org/10.1126/science.1231535

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., 2018. Twelve Questions for the Participatory Modeling Community. Earth's Future 6, 1046–1057. https://doi.org/10.1029/2018EF000841

Kelly (Letcher), R.A., Jakeman, A.J., Barreteau, O., Borsuk, M.E., ElSawah, S., Hamilton, S.H., Henriksen, H.J., Kuikka, S., Maier, H.R., Rizzoli, A.E., van Delden, H., Voinov, A.A., 2013. Selecting among five common modelling approaches for integrated environmental assessment and management. Environmental Modelling & Software 47, 159–181. https://doi.org/10.1016/j.envsoft.2013.05.005

Kloprogge, P., van der Sluijs, J.P., Petersen, A.C., 2011. A method for the analysis of assumptions in model-based environmental assessments. Environmental Modelling & Software, Thematic issue on the assessment and evaluation of environmental models and software 26, 289–301. https://doi.org/10.1016/j.envsoft.2009.06.009

Knapen, R., Janssen, S., Roosenschoon, O., Verweij, P., de Winter, W., Uiterwijk, M., Wien, J.-E., 2013. Evaluating OpenMI as a model integration platform across disciplines. Environmental Modelling & Software, Thematic Issue on the Future of Integrated Modeling Science and Technology 39, 274–282. https://doi.org/10.1016/j.envsoft.2012.06.011

Knight, D., Durham, C.C., Locke, E.A., 2001. The Relationship of Team Goals, Incentives, and Efficacy to Strategic Risk, Tactical Implementation, and Performance. Academy of Management Journal 44, 326–338. https://doi.org/10.2307/3069459

Kragt, M.E., Robson, B.J., Macleod, C.J.A., 2013. Modellers' roles in structuring integrative research projects. Environmental Modelling & Software, Thematic Issue on the Future of Integrated Modeling Science and Technology 39, 322–330. https://doi.org/10.1016/j.envsoft.2012.06.015

Kretser, A., Murphy, D., Bertuzzi, S., Abraham, T., Allison, D.B., Boor, K.J., Dwyer, J., Grantham, A., Harris, L.J., Hollander, R., Jacobs-Young, C., Rovito, S., Vafiadis, D., Woteki, C., Wyndham, J., Yada, R., 2019. Scientific Integrity Principles and Best Practices: Recommendations from a Scientific Integrity Consortium. Sci Eng Ethics 25, 327–355. https://doi.org/10.1007/s11948-019-00094-3

Kwakkel, J.H., 2017. The Exploratory Modeling Workbench: An open source toolkit for exploratory modeling, scenario discovery, and (multi-objective) robust decision making. Environmental Modelling & Software 96, 239–250. https://doi.org/10.1016/j.envsoft.2017.06.054

Kwon, S., Solomon, G.E.A., Youtie, J., Porter, A.L., 2017. A measure of knowledge flow between specific fields: Implications of interdisciplinarity for impact and funding. PLOS ONE 12, e0185583. https://doi.org/10.1371/journal.pone.0185583

Lahtinen, T.J., Guillaume, J.H.A., Hmlinen, 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

Laukkanen, E., Itkonen, J., Lassenius, C., 2017. Problems, causes and solutions when adopting continuous delivery—A systematic literature review. Information and Software Technology 82, 55–79. https://doi.org/10.1016/j.infsof.2016.10.001

Lee, C., Lin, Y., Huan, H., Huang, W., Teng, H., 2015. The effects of task interdependence, team cooperation, and team conflict on job performance. Social Behavior and Personality: An International Journal 43, 529–536. https://doi.org/10.2224/sbp.2015.43.4.529

Li, G., Bie, Z., Kou, Y., Jiang, J., Bettinelli, M., 2016. Reliability evaluation of integrated energy systems based on smart agent communication. Applied Energy 167, 397–406. https://doi.org/10.1016/j.apenergy.2015.11.033

Ling, Y., Mullins, J., Mahadevan, S., 2012. Calibration of multi-physics computational models using Bayesian networks.

Linkov, I., Burmistrov, D., 2003. Model Uncertainty and Choices Made by Modelers: Lessons Learned from the International Atomic Energy Agency Model Intercomparisons†. Risk Analysis 23, 1297–1308. https://doi.org/10.1111/j.0272-4332.2003.00402.x

Little, J.C., Hester, E.T., Elsawah, S., Filz, G.M., Sandu, A., Carey, C.C., Iwanaga, T., Jakeman, A.J., 2019. A tiered, system-of-systems modeling framework for resolving complex socio-environmental policy issues. Environmental Modelling & Software 112, 82–94. https://doi.org/10.1016/j.envsoft.2018.11.011

López–Fidalgo, J., Tommasi, C., 2018. Optimal Experimental Design for Model Selection: A Partial Review, in: Gil, E., Gil, Eva, Gil, J., Gil, M.Á. (Eds.), The Mathematics of the Uncertain: A Tribute to Pedro Gil, Studies in Systems, Decision and Control. Springer International Publishing, Cham, pp. 253–263. https://doi.org/10.1007/978-3-319-73848-2_24

Lorek, H., Sonnenschein, M., 1999. Modelling and simulation software to support individual-based ecological modelling. Ecological Modelling 115, 199–216. https://doi.org/10.1016/S0304-3800(98)00193-8

MacLeod, M., Nagatsu, M., 2018. What does interdisciplinarity look like in practice: Mapping interdisciplinarity and its limits in the environmental sciences. Studies in History and Philosophy of Science Part A 67, 74–84. https://doi.org/10.1016/j.shpsa.2018.01.001

Macy, M.W., Willer, R., 2002. From Factors to Actors: Computational Sociology and Agent-Based Modeling. Annual Review of Sociology 28, 143–166. https://doi.org/10.1146/annurev.soc.28.110601.141117

Madni, A.M., Sievers, M., 2014. Systems Integration: Key Perspectives, Experiences, and Challenges. Systems Engineering 17, 37–51. https://doi.org/10.1002/sys.21249

Maier, H.R., Guillaume, J.H.A., van Delden, H., Riddell, G.A., Haasnoot, M., Kwakkel, J.H., 2016. An uncertain future, deep uncertainty, scenarios, robustness and adaptation: How do they fit together? Environmental Modelling & Software 81, 154–164. https://doi.org/10.1016/j.envsoft.2016.03.014

Malard, J.J., Inam, A., Hassanzadeh, E., Adamowski, J., Tuy, H.A., Melgar-Quiñonez, H., 2017. Development of a software tool for rapid, reproducible, and stakeholder-friendly dynamic coupling of system dynamics and physically-based models. Environmental Modelling & Software 96, 410–420. https://doi.org/10.1016/J.ENVSOFT.2017.06.053

Martin, D.M., Powell, S.J., Webb, J.A., Nichols, S.J., Poff, N.L., 2017. An Objective Method to Prioritize Socio-Environmental Water Management Tradeoffs Using Multi-Criteria Decision Analysis. River Research and Applications 33, 586–596. https://doi.org/10.1002/rra.3103

Matott, L.S., Babendreier, J.E., Purucker, S.T., 2009. Evaluating uncertainty in integrated environmental models: A review of concepts and tools. Water Resources Research 45. https://doi.org/10.1029/2008WR007301

Meadows, D.H., 2008. Thinking in Systems: A Primer. Chelsea Green Publishing.

Meirmans, S., Butlin, R.K., Charmantier, A., Engelstädter, J., Groot, A.T., King, K.C., Kokko, H., Reid, J.M., Neiman, M., 2019. Science policies: How should science funding be allocated? An evolutionary biologists' perspective. J Evol Biol 32, 754–768. https://doi.org/10.1111/jeb.13497

Midgley, G., Pinzón, L.A., 2011. Boundary critique and its implications for conflict prevention. Journal of the Operational Research Society 62, 1543–1554. https://doi.org/10.1057/jors.2010.76

Miyasaka, T., Le, Q.B., Okuro, T., Zhao, X., Takeuchi, K., 2017. Agent-based modeling of complex social–ecological feedback loops to assess multi-dimensional trade-offs in dryland ecosystem services. Landscape Ecol 32, 707–727. https://doi.org/10.1007/s10980-017-0495-x

Moore, R.V., Tindall, C.I., 2005. An overview of the open modelling interface and environment (the OpenMI). Environmental Science & Policy, Research & Technology Integration in Support of the European Union Water Framework Directive 8, 279–286. https://doi.org/10.1016/j.envsci.2005.03.009

Moriasi, D.N., Arnold, J.G., Liew, M.W.V., Bingner, R.L., Harmel, R.D., Veith, T.L., 2007. Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. Transactions of the ASABE 50, 885–900. https://doi.org/10.13031/2013.23153

Muñoz, E., Rivera, D., Vergara, F., Tume, P., Arumí, J.L., 2014. Identifiability analysis: towards constrained equifinality and reduced uncertainty in a conceptual model. Hydrological Sciences Journal 59, 1690–1703. https://doi.org/10.1080/02626667.2014.892205

National Research Council, 2013. Sustainability for the Nation: Resource Connections and Governance Linkages. https://doi.org/10.17226/13471

Neumann, P., Düben, P., Adamidis, P., Bauer, P., Brück, Kornblueh, L., Klocke, D., Stevens, B., Wedi, N., Biercamp, J., 2019. Assessing the scales in numerical weather and climate predictions: will exascale be the rescue? Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 377, 20180148. https://doi.org/10.1098/rsta.2018.0148

Nielsen, C.B., Larsen, P.G., Fitzgerald, J., Woodcock, J., Peleska, J., 2015. Systems of Systems Engineering: Basic Concepts, Model-Based Techniques, and Research Directions. ACM Comput. Surv. 48, 18:1–18:41. https://doi.org/10.1145/2794381

Norström, A.V., Cvitanovic, C., Löf, M.F., West, S., Wyborn, C., Balvanera, P., Bednarek, A.T., Bennett, E.M., Biggs, R., de Bremond, A., Campbell, B.M., Canadell, J.G., Carpenter, S.R., Folke, C., Fulton, E.A., Gaffney, O., Gelcich, S., Jouffray, J.-B., Leach, M., Le Tissier, M., Martín-López, B., Louder, E., Loutre, M.-F., Meadow, A.M., Nagendra, H., Payne, D., Peterson, G.D., Reyers, B., Scholes, R., Speranza, C.I., Spierenburg, M., Stafford-Smith, M., Tengö, M., van der Hel, S., van Putten, I., Österblom, H., 2020. Principles for knowledge co-production in sustainability research. Nature Sustainability 3, 182–190. https://doi.org/10.1038/s41893-019-0448-2

Norton, J., 2015. An introduction to sensitivity assessment of simulation models. Environmental Modelling & Software 69, 166–174. https://doi.org/10.1016/j.envsoft.2015.03.020

Nossent, J., Bauwens, W., 2012. Multi-variable sensitivity and identifiability analysis for a complex environmental model in view of integrated water quantity and water quality modeling. Water Sci. Technol. 65, 539–549. https://doi.org/10.2166/wst.2012.884

O'Connell, P.E., Todini, E., 1996. Modelling of rainfall, flow and mass transport in hydrological systems: an overview. Journal of Hydrology 175, 3–16. https://doi.org/10.1016/S0022-1694(96)80003-5

Oreskes, N., 2003. The Role of Quantitative Models in Science, in: Canham, C.D., Cole, J.J., Lauenroth, W.K. (Eds.), Models in Ecosystem Science. Princeton University Press, pp. 13–31.

Ostrom, E., 2007. A diagnostic approach for going beyond panaceas. Proc Natl Acad Sci USA 104, 15181. https://doi.org/10.1073/pnas.0702288104

Oxley, T., ApSimon, H.M., 2007. Space, time and nesting Integrated Assessment Models. Environmental Modelling & Software 22, 1732–1749. https://doi.org/10.1016/j.envsoft.2007.02.002

Pacheco-Romero, M., Alcaraz-Segura, D., Vallejos, M., Cabello, J., 2020. An expert-based reference list of variables for characterizing and monitoring social-ecological systems. Ecology and Society 25. https://doi.org/10.5751/ES-11676-250301

Pahl-Wostl, C., 2007. Transitions towards adaptive management of water facing climate and global change. Water Resour Manage 21, 49–62. https://doi.org/10.1007/s11269-006-9040-4

Patil, P., Peng, R.D., Leek, J.T., 2016. A statistical definition for reproducibility and replicability. bioRxiv 066803. https://doi.org/10.1101/066803

Peng, R.D., 2011. Reproducible Research in Computational Science. Science 334, 1226–1227. https://doi.org/10.1126/science.1213847

Phillips, S.D., Estler, W.T., Doiron, T., Eberhardt, K.R., Levenson, M.S., 2001. A careful consideration of the calibration concept. Journal of research of the National Institute of Standards and Technology 106. https://doi.org/10.6028/jres.106.014

Pianosi, F., Beven, K., Freer, J., Hall, J.W., Rougier, J., Stephenson, D.B., Wagener, T., 2016. Sensitivity analysis of environmental models: A systematic review with practical workflow. Environmental Modelling & Software 79, 214–232. https://doi.org/10.1016/j.envsoft.2016.02.008

Pianosi, F., Sarrazin, F., Wagener, T., 2015. A Matlab toolbox for Global Sensitivity Analysis. Environmental Modelling & Software 70, 80–85. https://doi.org/10.1016/j.envsoft.2015.04.009

Pianosi, F., Wagener, T., 2015. A simple and efficient method for global sensitivity analysis based on cumulative distribution functions. Environmental Modelling & Software 67, 1–11. https://doi.org/10.1016/j.envsoft.2015.01.004

Pietzsch, B., Fiedler, S., Mertens, K.G., Richter, M., Scherer, C., Widyastuti, K., Wimmler, M.-C., Zakharova, L., Berger, U., 2020. Metamodels for Evaluating, Calibrating and Applying Agent-Based Models: A Review. JASSS 23, 9. https://doi.org/10.18564/jasss.4274

Porporato, A., Rodriguez-Iturbe, I., 2002. Ecohydrology-a challenging multidisciplinary research perspective / Ecohydrologie: une perspective stimulante de recherche multidisciplinaire. Hydrological Sciences Journal 47, 811–821. https://doi.org/10.1080/02626660209492985

Potter, S., Doran, B., Mathews, D., 2016. Modelling collective Yawuru values along the foreshore of Roebuck Bay, Western Australia using fuzzy logic. Applied Geography 77, 8–19. https://doi.org/10.1016/j.apgeog.2016.09.016

Pulver, S., Ulibarri, N., Sobocinski, K., Alexander, S., Johnson, M., McCord, P., Dell'Angelo, J., 2018. Frontiers in socio-environmental research: components, connections, scale, and context. Ecology and Society 23. https://doi.org/10.5751/ES-10280-230323

Rai, A., Lang, S.S., Welker, R.B., 2002. Assessing the Validity of IS Success Models: An Empirical Test and Theoretical Analysis. Information Systems Research 13, 50–69. https://doi.org/10.1287/isre.13.1.50.96

Raick, C., Soetaert, K., Grégoire, M., 2006. Model complexity and performance: How far can we simplify? Progress in Oceanography 70, 27–57. https://doi.org/10.1016/j.pocean.2006.03.001

Ramaswami, A., Milford, J.B., Small, M.J., 2005. Integrated Environmental Modeling: Pollutant Transport, Fate, and Risk in the Environment. John Wiley & Sons Inc., Hoboken, NJ.

Razavi, S., Gober, P., Maier, H.R., Brouwer, R., Wheater, H., 2020. Anthropocene flooding: Challenges for science and society. Hydrological Processes 34, 1996–2000. https://doi.org/10.1002/hyp.13723

Razavi, S., Gupta, H.V., 2015. What do we mean by sensitivity analysis? The need for comprehensive characterization of "global" sensitivity in Earth and Environmental systems models. Water Resources Research 51, 3070–3092. https://doi.org/10.1002/2014WR016527

Razavi, S., Sheikholeslami, R., Gupta, H.V., Haghnegahdar, A., 2019. VARS-TOOL: A toolbox for comprehensive, efficient, and robust sensitivity and uncertainty analysis. Environmental Modelling & Software 112, 95–107. https://doi.org/10.1016/j.envsoft.2018.10.005

Razavi, S., Tolson, B.A., 2013. An efficient framework for hydrologic model calibration on long data periods. Water Resources Research 49, 8418–8431. https://doi.org/10.1002/2012WR013442

Razavi, S., Tolson, B.A., Burn, D.H., 2012. Review of surrogate modeling in water resources. Water Resources Research 48. https://doi.org/10.1029/2011WR011527

Razavi, S., Tolson, B.A., Matott, L.S., Thomson, N.R., MacLean, A., Seglenieks, F.R., 2010. Reducing the computational cost of automatic calibration through model preemption. Water Resources Research 46. https://doi.org/10.1029/2009WR008957

Refsgaard, J.C., van der Sluijs, J.P., Højberg, A.L., Vanrolleghem, P.A., 2007. Uncertainty in the environmental modelling process – A framework and guidance. Environmental Modelling & Software 22, 1543–1556. https://doi.org/10.1016/j.envsoft.2007.02.004

Reichert, P., 2020. Towards a comprehensive uncertainty assessment in environmental research and decision support. Water Sci Technol. https://doi.org/10.2166/wst.2020.032

Reichert, P., Borsuk, M.E., 2005. Does high forecast uncertainty preclude effective decision support? Environmental Modelling & Software, Methods of Uncertainty Treatment in Environmental Models 20, 991–1001. https://doi.org/10.1016/j.envsoft.2004.10.005

Renner, S.A., 2001. A "Community of Interest" Approach to Data Interoperability.

Roberts, J.J., Marotta Cassula, A., Silveira, J.L., da Costa Bortoni, E., Mendiburu, A.Z., 2018. Robust multi-objective optimization of a renewable based hybrid power system. Applied Energy 223, 52–68. https://doi.org/10.1016/j.apenergy.2018.04.032

Rounsevell, M.D.A., Metzger, M.J., 2010. Developing qualitative scenario storylines for environmental change assessment. WIREs Climate Change 1, 606–619. https://doi.org/10.1002/wcc.63

Rubin, R.B., Rubin, A.M., Graham, E., Perse, E.M., Seibold, D., 2010. Communication Research Measures II: A Sourcebook. Routledge.

Salado, A., 2015. Abandonment: A natural consequence of autonomy and belonging in systems-of-systems, in: 2015 10th System of Systems Engineering Conference (SoSE). Presented at the 2015 10th System of Systems Engineering Conference (SoSE), pp. 352–357. https://doi.org/10.1109/SYSOSE.2015.7151936

Saltelli, A., 2018. Why science's crisis should not become a political battling ground. Futures 104, 85–90. https://doi.org/10.1016/j.futures.2018.07.006

Saltelli, A., Aleksankina, K., Becker, W., Fennell, P., Ferretti, F., Holst, N., Li, S., Wu, Q., 2019. Why So Many Published Sensitivity Analyses Are False: A Systematic Review of Sensitivity Analysis Practices. Environmental Modelling & Software 114, 29–39.

Saltelli, A., Annoni, P., 2010. How to avoid a perfunctory sensitivity analysis. Environmental Modelling and Software 25, 1508–1517. https://doi.org/10.1016/j.envsoft.2010.04.012

Sandström, U., Van den Besselaar, P., 2018. Funding, evaluation, and the performance of national research systems. Journal of Informetrics 12, 365–384. https://doi.org/10.1016/j.joi.2018.01.007

Sarosa, S., Tatnall, A., 2015. Failure to Launch: Scope Creep and Other Causes of Failure from an Actor-Network Theory Perspective. International Journal of Actor-Network Theory and Technological Innovation (IJANTTI) 7, 1–13. https://doi.org/10.4018/ijantti.2015100101

Schlüter, M., Haider, L., Lade, S., Lindkvist, E., Martin, R., Orach, K., Wijermans, N., Folke, C., 2019a. Capturing emergent phenomena in social-ecological systems: an analytical framework. Ecology and Society 24. https://doi.org/10.5751/ES-11012-240311

Schlüter, M., Müller, B., Frank, K., 2019b. The potential of models and modeling for social-ecological systems research: the reference frame ModSES. Ecology and Society 24. https://doi.org/10.5751/ES-10716-240131

Schneider, B., Ehrhart, M.G., Macey, W.H., 2013. Organizational Climate and Culture. Annu. Rev. Psychol. 64, 361–388. https://doi.org/10.1146/annurev-psych-113011-143809

Schnell, S., 2018. "Reproducible" Research in Mathematical Sciences Requires Changes in our Peer Review Culture and Modernization of our Current Publication Approach. Bull. Math. Biol. 80, 3095–3105. https://doi.org/10.1007/s11538-018-0500-9

Scholes, R.J., Reyers, B., Biggs, R., Spierenburg, M.J., Duriappah, A., 2013. Multi-scale and cross-scale assessments of social–ecological systems and their ecosystem services. Current Opinion in Environmental Sustainability 5, 16–25. https://doi.org/10.1016/j.cosust.2013.01.004

Schouten, R., Deits, R., 2020. Flatten.jl.

Schweiger, G., Nilsson, H., Schoeggl, J., Birk, W., Posch, A., 2020. Modeling and simulation of large-scale systems: A systematic comparison of modeling paradigms. Applied Mathematics and Computation 365, 124713. https://doi.org/10.1016/j.amc.2019.124713

Shin, M.-J., Guillaume, J.H.A., Croke, B.F.W., Jakeman, A.J., 2013. Addressing ten questions about conceptual rainfall–runoff models with global sensitivity analyses in R. Journal of Hydrology 503, 135–152. https://doi.org/10.1016/j.jhydrol.2013.08.047

Singh, S.K., Bárdossy, A., 2012. Calibration of hydrological models on hydrologically unusual events. Advances in Water Resources 38, 81–91. https://doi.org/10.1016/j.advwatres.2011.12.006

Sivapalan, M., Savenije, H.H.G., Blöschl, G., 2012. Socio-hydrology: A new science of people and water. Hydrological Processes 26, 1270–1276. https://doi.org/10.1002/hyp.8426

Sletholt, M.T., Hannay, J.E., Pfahl, D., Langtangen, H.P., 2012. What do we know about scientific software development's agile practices? 24–36. https://doi.org/10.1109/MCSE.2011.113

Song, X., Zhang, J., Zhan, C., Xuan, Y., Ye, M., Xu, C., 2015. Global sensitivity analysis in hydrological modeling: Review of concepts, methods, theoretical framework, and applications. Journal of Hydrology 523, 739–757. https://doi.org/10.1016/j.jhydrol.2015.02.013

Spitzberg, B.H., Cupach, W.R., 1989. Handbook of interpersonal competence research. Springer-Verlag.

Steinhardt, U., Volk, M., 2001. Scales and spatio-temporal dimensions in landscape research, in: Krönert, R., Steinhardt, U., Volk, M. (Eds.), Landscape Balance and Landscape Assessment. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 137–162. https://doi.org/10.1007/978-3-662-04532-9_6

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 Anth 7, 9. https://doi.org/10.1525/elementa.347

Tate, L.E., 2013. Growth-Management Implementation in Metropolitan Vancouver: Lessons from Actor-Network Theory. Environ Plann B Plann Des 40, 783–800. https://doi.org/10.1068/b37170

Tebaldi, C., Knutti, R., 2007. The use of the multi-model ensemble in probabilistic climate projections. Philos Trans A Math Phys Eng Sci 365, 2053–2075. https://doi.org/10.1098/rsta.2007.2076

Thomas, J., McDonagh, D., 2013. Shared language:Towards more effective communication. Australas Med J 6, 46–54. https://doi.org/10.4066/AMJ.2013.1596

Thompson, J.L., 2009. Building Collective Communication Competence in Interdisciplinary Research Teams. Journal of Applied Communication Research 37, 278–297. https://doi.org/10.1080/00909880903025911

Tjosvold, D., Hui, C., Ding, D.Z., Hu, J., 2003. Conflict values and team relationships: conflict's contribution to team effectiveness and citizenship in China. Journal of Organizational Behavior 24, 69–88. https://doi.org/10.1002/job.180

Tranquillo, J., 2019. An Introduction to Complex Systems: Making Sense of a Changing World. Springer Nature.

Tscheikner-Gratl, F., Bellos, V., Schellart, A., Moreno-Rodenas, A., Muthusamy, M., Langeveld, J., Clemens, F., Benedetti, L., Rico-Ramirez, M.A., de Carvalho, R.F., Breuer, L., Shucksmith, J., Heuvelink, G.B.M., Tait, S., 2019. Recent insights on uncertainties present in integrated catchment water quality modelling. Water Research 150, 368–379. https://doi.org/10.1016/j.watres.2018.11.079

Uusitalo, L., Lehikoinen, A., Helle, I., Myrberg, K., 2015. An overview of methods to evaluate uncertainty of deterministic models in decision support. Environmental Modelling & Software 63, 24–31. https://doi.org/10.1016/j.envsoft.2014.09.017

Vale, T., Crnkovic, I., De Almeida, E.S., Silveira Neto, P.A.D.M., Cavalcanti, Y.C., Meira, S.R.D.L., 2016. Twenty-eight years of component-based software engineering. Journal of Systems and Software 111, 128–148. https://doi.org/10.1016/j.jss.2015.09.019

van Delden, H., van Vliet, J., Rutledge, D.T., Kirkby, M.J., 2011. Comparison of scale and scaling issues in integrated land-use models for policy support. Agriculture, Ecosystems & Environment, Scaling methods in integrated assessment of agricultural systems 142, 18–28. https://doi.org/10.1016/j.agee.2011.03.005

van der Sluijs, J.P., 2007. Uncertainty and precaution in environmental management: Insights from the UPEM conference. Environmental Modelling & Software, The Implications of Complexity for Integrated Resources 22, 590–598. https://doi.org/10.1016/j.envsoft.2005.12.020

van der Vaart, E., Johnston, A.S.A., Sibly, R.M., 2016. Predicting how many animals will be where: How to build, calibrate and evaluate individual-based models. Ecological Modelling 326, 113–123. https://doi.org/10.1016/j.ecolmodel.2015.08.012

van Ittersum, M.K., Ewert, F., Heckelei, T., Wery, J., Alkan Olsson, J., Andersen, E., Bezlepkina, I., Brouwer, F., Donatelli, M., Flichman, G., Olsson, L., Rizzoli, A.E., van der Wal, T., Wien, J.E., Wolf, J., 2008. Integrated assessment of agricultural systems – A component-based framework for the European Union (SEAMLESS). Agricultural Systems 96, 150–165. https://doi.org/10.1016/j.agsy.2007.07.009

Vanlier, J., Tiemann, C.A., Hilbers, P.A., van Riel, N.A., 2014. Optimal experiment design for model selection in biochemical networks. BMC Systems Biology 8, 20. https://doi.org/10.1186/1752-0509-8-20

Verweij, P.J.F.M., Knapen, M.J.R., de Winter, W.P., Wien, J.J.F., te Roller, J.A., Sieber, S., Jansen, J.M.L., 2010. An IT perspective on integrated environmental modelling: The SIAT case. Ecological Modelling, Model-based Systems to Support Impact Assessment - Methods, Tools and Applications 221, 2167–2176. https://doi.org/10.1016/j.ecolmodel.2010.01.006

Voinov, A., Bousquet, F., 2010. Modelling with stakeholders. Environmental Modelling & Software, 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. Environmental Modelling & Software 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. Environmental Modelling & Software 77, 196–220. https://doi.org/10.1016/j.envsoft.2015.11.016

Voinov, A., Shugart, H.H., 2013. "Integronsters", integral and integrated modeling. Environmental Modelling and Software 39, 149–158. https://doi.org/10.1016/j.envsoft.2012.05.014

Wagener, T., McIntyre, N., Lees, M.J., Wheater, H.S., Gupta, H.V., 2003. Towards reduced uncertainty in conceptual rainfall-runoff modelling: dynamic identifiability analysis. Hydrological Processes 17, 455–476. https://doi.org/10.1002/hyp.1135

Walker, W.E., Harremoës, P., Rotmans, J., Sluijs, J.P. van der, Asselt, M.B.A. van, Janssen, P., Krauss, M.P.K. von, 2003. Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. Integrated Assessment 4, 5–17. https://doi.org/10.1076/iaij.4.1.5.16466

Wang, K., Davies, E.G.R., Liu, J., 2019. Integrated water resources management and modeling: A case study of Bow river basin, Canada. Journal of Cleaner Production 240, 118242. https://doi.org/10.1016/j.jclepro.2019.118242

Warren, K., 2014. Agile SD: Fast, Effective, Reliable, in: Proceedings of the 32nd International Conference of the System Dynamics Society. Presented at the Conference of the System Dynamics Society, Systems Dynamics Society, Delft, Netherlands.

Wessells, A.T., 2007. A Review of: "Reassembling the Social: An Introduction to Actor-Network-Theory". International Public Management Journal 10, 351–356. https://doi.org/10.1080/10967490701515606

Whelan, G., Kim, K., Pelton, M.A., Castleton, K.J., Laniak, G.F., Wolfe, K., Parmar, R., Babendreier, J., Galvin, M., 2014. Design of a component-based integrated environmental modeling framework. Environmental Modelling & Software 55, 1–24. https://doi.org/10.1016/j.envsoft.2014.01.016

White, D.D., Lawless, K.L., Vivoni, E.R., Mascaro, G., Pahle, R., Kumar, I., Coli, P., Castillo, R.M., Moreda, F., Asfora, M., 2019. Co-Producing Interdisciplinary Knowledge and Action for Sustainable Water Governance: Lessons from the Development of a Water Resources Decision Support System in Pernambuco, Brazil. Global Challenges 3, 1800012. https://doi.org/10.1002/gch2.201800012

Wiegand, T., Jeltsch, F., Hanski, I., Grimm, V., 2003. Using pattern-oriented modeling for revealing hidden information: a key for reconciling ecological theory and application. Oikos 100, 209–222. https://doi.org/10.1034/j.1600-0706.2003.12027.x

Wiegand, T., Revilla, E., Knauer, F., 2004. Dealing with Uncertainty in Spatially Explicit Population Models. Biodiversity and Conservation 13, 53–78. https://doi.org/10.1023/B:BIOC.0000004313.86836.ab

Willcock, S., Hooftman, D.A.P., Blanchard, R., Dawson, T.P., Hickler, T., Lindeskog, M., Martinez-Lopez, J., Reyers, B., Watts, S.M., Eigenbrod, F., Bullock, J.M., 2020. Ensembles of ecosystem service models can improve accuracy and indicate uncertainty. Science of The Total Environment 141006. https://doi.org/10.1016/j.scitotenv.2020.141006

Wilson, C., Kriegler, E., van Vuuren, D.P., Guivarch, C., Frame, D., Krey, V., Osborn, T.J., Schwanitz, V.J., Thompson, E.L., 2017. Evaluating Process-Based Integrated Assessment Models of Climate Change Mitigation [WWW Document]. URL http://pure.iiasa.ac.at/id/eprint/14502/ (accessed 12.16.19).

Wirtz, D., Nowak, W., 2017. The rocky road to extended simulation frameworks covering uncertainty, inversion, optimization and control. Environmental Modelling and Software 93, 180–192. https://doi.org/10.1016/j.envsoft.2016.10.003

Wu, R., Yang, L., Chen, C., Ahmad, S., Dascalu, S.M., Harris Jr., F.C., 2019. MELPF version 1: Modeling Error Learning based Post-Processor Framework for Hydrologic Models Accuracy Improvement. Geoscientific Model Development 12, 4115–4131. https://doi.org/10.5194/gmd-12-4115-2019

Xiong, F., Xiong, Y., Greene, S., Chen, W., Yang, S., 2010. A New Sparse Grid Based Method for Uncertainty Propagation. Presented at the ASME 2009 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, American Society of Mechanical Engineers Digital Collection, pp. 1205–1215. https://doi.org/10.1115/DETC2009-87430

Yearworth, M., White, L., 2018. Spontaneous emergence of Community OR: Self-initiating, self-organising problem structuring mediated by social media. European Journal of Operational Research, Community Operational Research: Innovations, internationalization and agenda-setting applications 268, 809–824. https://doi.org/10.1016/j.ejor.2018.01.024

Yoo, S., Harman, M., 2012. Regression testing minimization, selection and prioritization: a survey. Software Testing, Verification and Reliability 22, 67–120. https://doi.org/10.1002/stvr.430

Zare, F., Elsawah, S., Iwanaga, T., Jakeman, A.J., Pierce, S.A., 2017. Integrated water assessment and modelling: A bibliometric analysis of trends in the water resource sector. Journal of Hydrology 552, 765–778. https://doi.org/10.1016/j.jhydrol.2017.07.031

Zhang, C., Lin, H., Chen, M., Li, R., Zeng, Z., 2014a. Scale compatibility analysis in geographic process research: A case study of a meteorological simulation in Hong Kong. Applied Geography 52, 135–143. https://doi.org/10.1016/j.apgeog.2014.05.005

Zhang, C., Lin, H., Chen, M., Yang, L., 2014b. Scale matching of multiscale digital elevation model (DEM) data and the Weather Research and Forecasting (WRF) model: a case study of meteorological simulation in Hong Kong. Arab J Geosci 7, 2215–2223. https://doi.org/10.1007/s12517-014-1273-6

Zimmermann, H.-J., 2000. An application-oriented view of modeling uncertainty. European Journal of Operational Research 122, 190–198. https://doi.org/10.1016/S0377-2217(99)00228-3

# Highlights

- Scale incompatibilities among system representations of constituent systems are identified as a key challenge in socio-environmental systems modeling
- Issues of scale arise from the complexity, size and heterogeneity of the constituent systems and their interactions
- A more holistic systems-of-systems modeling framework is needed within which to integrate current approaches and tools
- A range of system modeling considerations within the socio-technical context of system-of-systems modeling is presented based on input from a variety of disciplinary and interdisciplinary experts

# Competing Interests

None to declare.