Scale decisions and good practices in socio-environmental systems modelling: guidance and documentation during problem scoping and model formulation

Hsiao-Hsuan Wang[1*], George van Voorn[2], William E. Grant[1], Fateme Zare[3], Carlo Giupponi[4], Patrick Steinmann[2], Birgit Müller[5,6], Sondoss Elsawah[3], Hedwig van Delden[7,8], Ioannis N. Athanasiadis[9], Zhanli Sun[10], Wander Jager[11], John C. Little[12], and Anthony J. Jakeman[13]

[1] Ecological Systems Laboratory, Department of Ecology and Conservation Biology, Texas A&M University, USA
[2] Biometris, Wageningen University & Research, The Netherlands
[3] Capability Systems Centre, School of Engineering and Information Technology, University of New South Wales, Canberra, Australia
[4] Department of Economics, Ca' Foscari University of Venice and Fondazione Eni Enrico Mattei, Italy
[5] Department of Ecological Modelling, Helmholtz Centre for Environmental Research - UFZ, Leipzig, Germany
[6] Chair of Modelling of Human-Environment-Systems, Brandenburg University of Cottbus-Senftenberg, Cottbus, Germany.
[7] Research Institute for Knowledge Systems, Maastricht, The Netherlands
[8] School of Civil, Environmental and Mining Engineering, University of Adelaide, Adelaide, Australia
[9] Data Competence Center, Wageningen University and Research, Wageningen, The Netherlands
[10] Leibniz Institute of Agricultural Development in Transition Economies (IAMO), Halle (Saale), Germany
[11] Groningen Center for Social Complexity Studies, University of Groningen, Groningen, The Netherlands
[12] Department of Civil and Environmental Engineering, Virginia Tech, USA
[13] Institute for Water Futures, Fenner School of Environment and Society, Australian National University, Canberra, Australia

Abstract

Models of socio-environmental or social-ecological systems (SES) commonly address problems requiring interdisciplinary scientific expertise and input from a heterogeneous group of stakeholders. In SES modelling multiple interactions occur on different scales among various phenomena. These scale phenomena include the technical, such as system variables, process detail, inputs and outputs, which most often require spatial, temporal, thematic and organisational choices. From a good practice and project efficiency perspective, the problem scoping and conceptual model formulation phase of modelling is the one to address well from the outset. During this phase, intense and substantive discussions should arise regarding appropriate scales at which to represent the different phenomena. Although the details of these discussions influence the path of model development, they are seldom documented and as a result often forgotten. We draw upon personal experience with existing protocols and communications in recent literature to propose preliminary guidelines for documenting these early discussions about the scale(s) of the studied phenomena. Our guidelines aim to aid modelling group members in building and capturing the richness of their rationale for scoping and scale decisions. The resulting transcripts are intended to promote transparency of modelling decisions and provide essential support for the justification of the final model for its intended use. They also facilitate adaptive

Correspondence:
Contact H.-H. Wang at hsiaohsuan.wang@ag.tamu.edu

Cite this article as:
Wang, H.-H., van Voorn, G., Grant, W.E., Zare, F., Giupponi, C., Steinmann, P., Müller, B., Elsawah, S., van Delden, H., Athanasiadis, I.N., Sun, Z., Jager, W., Little, J.C., & Jakeman, A.J.
Scale decisions and good practices in socio-environmental systems modelling: guidance and documentation during problem scoping and model formulation
Socio-Environmental Systems Modelling, vol. 5, 18563, 2023, doi:10.18174/sesmo.18563

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

modifications of the pathway of model development via retracing decisions and iterative reflection upon alternative scale options.

Keywords
Organisational scale; spatial scale; temporal scale; thematic scale; uncertainty

1. Introduction

Models are indispensable tools in the analysis and any subsequent decision making related to socio-environmental or social-ecological systems (SES) problems and issues. They help describe, represent and analyse complex human-environment interactions for the system of concern, and thereby assist in understanding the system and potential ramifications of specific management decision options and uncontrollable system drivers. Treatment of scale is a crucial consideration in such integrated modelling, along with the identification of issues of concern, management options and governance arrangements, models, and sources and types of uncertainty (Hamilton et al., 2015). Much recent literature on modelling SES has focused on difficulties associated with identifying the appropriate scales at which to represent different system structures and processes, as well as different types of model output (Elsawah et al., 2020; Iwanaga et al., 2021a; Lippe et al., 2019).

1.1 The wide nature of scales

Scale decisions in SES modelling include not only those associated with defining temporal and spatial scales per se, but also those associated with defining thematic and organisational scales, especially in the social dimension (e.g., individual, community, and/or national level) (Elsawah et al., 2020; Gibson et al., 2000; Müller et al., 2013; van Delden et al., 2011). Here, the thematic resolution refers to the granular component of scale, regarding the thematic detail at which categorical maps of system components are made. The higher the thematic detail or resolution, the larger the amount of information we have and the more complex the analysis will be (García-Álvarez et al., 2019). Difficulties in identifying which components of the real system to include in the integrated model, and at what level of detail to represent them (intrinsic scale), are also related directly to issues of scale (Cumming et al., 2006; Gotts et al., 2019; Iwanaga et al., 2022).

The focus of this article is on the decision choices with respect to 'model scale', including not only spatial and temporal scales but also 'observational scale' (scale of measurement and sampling) and 'policy scale' (the scale for policy making) (for more details see Section 3.2.). More concretely, we follow Gibson et al. (2000) and Iwanaga et al. (2021a) in using the term "scale" in an expansive sense to refer to the spatial, temporal, thematic and organisational dimensions used to cover the scope of work conducted during the treatment and representation of system processes. In particular, organisational scales, although no less intuitive than temporal, spatial or thematic scales, are usually difficult to define precisely. Even though we agree that ontology and epistemology insights would be essential to reach a more solid and consistent scale type, we provide brief definitions in Box 1 of the various aspects of scale as we intend for them to be understood within the context of the present paper.

Box 1: Brief definitions of the various aspects of scale as we intend for them to be understood within the context of the present paper.

Scale encompasses the **spatial, temporal, thematic and organisational** dimensions used to describe system processes represented in the model (Gibson et al., 2000; Iwanaga et al., 2021a).

Model scale, in addition to **spatial, temporal, thematic, and organisational scales**, encompasses **observational scale, policy scale and intrinsic scale** (Wu & Li, 2006).

Thematic scale (= thematic resolution) refers to the amount of detail with which thematic categories are defined (García-Álvarez et al., 2019).

Organisational scale refers to aggregates of individuals in different social networks (e.g., individual, community and/or national level) (Elsawah et al., 2020; Gibson et al., 2000; Müller et al., 2013; van Delden et al., 2011).

Observational scale refers to the scale of measurement and sampling (Wu & Li, 2006).

Policy scale refers to the scale relevant for policy making (van Delden et al., 2011).

Intrinsic scale (= process scale) refers to the scale at which a process operates (Cumming et al., 2006; Gotts et al., 2019; Iwanaga et al., 2022).

Organisational scales are often associated with human components, like aggregates of individuals in different informal and formal social networks, for example, households, jobs, farmer networks, internet fora and organisations. While some networks are considered to be "scale-free", there are no universal units of measure when referring to human components of a system. Moreover, there usually are many interactions between agents, other units, and their environment. Hence, preferences by modellers and stakeholders on organisational scales can easily differ. For example, a model may be at the individual level or consider a larger unit (e.g., a village). An important aspect to consider is the scale of decision making by human components, which is equally challenging to specify. For instance, it is commonly assumed that different types of decisions made by agents are on the same timescale, which is not necessarily valid (Wernz & Deshmukh, 2012).

1.2 Why emphasise the communication and documentation of scale decisions?

Here, we emphasise the need for improving communication in modelling exercises or projects as a first step in resolving scale and associated issues in models. Disagreements during problem scoping and conceptual model formulation of SES models are both inevitable and desirable. Without such disagreements, assumptions might be accepted tacitly without sufficient thought given to justifying their acceptance. Thus, disagreements promote transparency, given they are openly communicated within the involved group of stakeholders. Productive debates are more likely when participants agree to clarify their own expectations regarding the outcomes of the model, to allow each other space to disagree and provide alternative viewpoints, and also have a clear understanding of end-user expectations. A useful lens here for such discussions is the principle of fitness-for-purpose. In contending that fitness-for-purpose must go beyond the functional use of a model to include its management, problem and project contexts, it has recently been honed and illustrated in the terms of model usability, reliability and feasibility, along with providing considerations and potential evaluation criteria for those terms (Hamilton et al., 2022).

Detailed explanations supporting decisions that define the scope and objectives of the model - such as (1) confining or extending the scope of questions the model will address, (2) including or excluding the explicit representation of particular system components and processes, and (3) using coarser or finer resolutions to represent constituent parts of the model - reveal the richness of thought embodied in this initial stage of model development (Jakeman et al., 2006). The inability to communicate clearly how the modelling process led to the model structure and subsequently to the model results undermines model credibility and limits model usefulness (Grimm et al., 2020). Lack of documentation during problem scoping and conceptual model formulation also prevents modelling group members from recalling their rationale for choosing the current, as opposed to an alternative, pathway of model development (Zare et al., 2020). And the adequacy of the documentation process will very likely affect the comprehensiveness and seriousness of the range of scale choices considered.

Hence, we also argue here that communication, in turn, can be greatly improved by better documentation of the modelling process. When properly documented, the content of ensuing debates can provide both guidance and justification for the pathway of model development (Zare et al., 2020). Awareness of the need to improve communication is growing and has resulted in the development of a variety of protocols and guidelines that are applicable to SES modelling (Ayllón et al., 2021; Badham et al., 2019; Grimm et al., 2014; Grimm et al., 2006; Grimm et al., 2010; Jakeman et al., 2006; Müller et al., 2013; van Voorn et al., 2016). All prescribe formal descriptions of each step in the modelling process (e.g., statement of model objectives, conceptual model formulation, quantitative model specification, model evaluation, and model application). Although different authors divide the modelling process into more or fewer steps and name the steps differently, the description of the process itself is consistent among the various schemes. Most schemes focus on documenting the details of a single model development path leading from a statement of objectives to the application of the model; few explicitly consider the need for discussing potential alternative paths that have not been taken in the modelling process.

Even though resource limitations can inhibit the exploration of alternative pathways of model development, recently there has been an increasing recognition of the importance of documenting the discussions at each important decision point during model development, in other words, the discussions which led to choosing a particular path forward over others (Ayllón et al., 2021; Grimm et al., 2014; Grimm et al., 2020; Zare et al., 2020). Grimm et al. (2014) introduced the idea of keeping a modelling notebook, analogous to a lab journal in laboratory research, which documents key assumptions and decisions along the path of model development to be traced back if needed. Ayllón et al. (2021) further elaborated on the concept of a modelling notebook by suggesting that it should also document the outcomes of relevant discussions among modelling group members and how these outcomes influenced the future directions of the modelling project. Zare et al. (2020) suggested using a pathway diagramming approach to document more formally the alternative pathways of model development that are discussed and to facilitate iterative reflection upon them. Such diagrams would indicate pathways that led to dead ends, both conceptually and in practice, and alternative potentially successful pathways not taken, as well as the actual pathway of model development. Zare et al. (2020) also proposed a set of symbols representing key pathway diagramming components, such as decision forks, selected options, other possible options, documentation points, and reflection points.

1.3 Why focus on the problem scoping and conceptual model formulation phases?

In this paper, we draw upon the above-mentioned literature and personal experience with existing protocols to propose improvements that facilitate documentation of, and reflection on, the discussions that direct the path of model development. But our main focus is on the discussions of scale issues by modelling group members and with stakeholders during problem scoping and conceptual model formulation, as this is an essential phase in SES model development to address well from the outset, notwithstanding that elements of it may be revisited in later phases. Problem scoping and conceptual model formulation provide the foundation for all subsequent model development, and virtually all choices encountered during problem scoping and conceptual model formulation involve issues of scale. The decisions made in that phase will transfer to support the selection of the integrated model type (Kelly et al., 2013) and its computational model implementation, and enable the potential application of subsequent methods that may sharpen choices further. Quantitative methods that enable discrimination among technical scale choices principally include assessing the identifiability of the computational model such as via sensitivity and uncertainty analyses (Guillaume et al., 2017; Guillaume et al., 2019; Razavi et al., 2021; Saltelli et al., 2021).

We hypothesise that more formal guidance for documenting the flow of these early discussions of scale will aid modelling group members in capturing the richness of their rationale for scoping and scale decisions, and thus help resolve scale issues. The additional time and money required to document these discussions are likely negligible compared to issues encountered later in the modelling process when scale issues have not been resolved, and the resulting transcripts will provide essential support for the justification of the final model for its intended purpose. The transcribing process will also facilitate adaptive modifications of the pathway of model development via iterative reflection upon the alternative scale options discussed.

1.4 Outline of the paper

In the following sections, we first justify the nature and importance of identifying scale issues that may arise during problem scoping and conceptual model formulation (Section 2), and then present a framework of considerations for the associated documentation, a structured process for thinking about and validating scale decisions, and guidance on communication issues (Section 3). We also point readers toward recent review papers. Next, we describe our proposed guidelines (Section 4), relating each step in the guidelines to a particular scale issue. We then offer suggestions for future work on scale concepts and issues in SES and related communities around developing a common understanding, building capacity and training, and widespread application to continually refine the guidelines (Section 5).

2. Scale issues arising during problem scoping and model formulation

Scale issues inevitably arise during the initial phases of the modelling process, and it is during these initial phases that we confront the most fundamental choices affecting model development. Virtually all choices encountered during problem scoping and conceptual model formulation (e.g., specifying model purpose, bounding the model, representing processes connecting model components, and specifying model inputs and outputs) involve issues of scale.

Problem scoping and conceptual model formulation are widely acknowledged and emphasised as crucial rungs upon which subsequent quantitative model development steps depend (Grimm et al., 2014; Jakeman et al., 2006; Refsgaard & Henriksen, 2004). The tenet that model objectives should guide all phases of the modelling process has long been accepted, regardless of the research field (Ford, 1999; Forrester, 1961; Grant, 1986; Patten, 1971; Schmolke et al., 2010). Perhaps less widely emphasised is the importance of detailed consideration of model purpose, especially for socio-environmental models. As Edmonds et al. (2019) point out, the purpose of a socio-environmental model may be to predict, explain, describe, theoretically explore, illustrate, serve as an analogy, or facilitate social interaction. Problem scoping and conceptual model formulation ideally are conducted within the context of a well-defined model purpose. For modelling that addresses problems requiring interdisciplinary scientific expertise and input from a heterogeneous group of stakeholders, identifying appropriate model objectives is particularly challenging (Elsawah et al., 2020; Iwanaga et al., 2021b; Jakeman et al., 2006). Most commonly a general statement of project objectives from a funding source is what brings modelling groups together in the first place. However, converting these general objectives into specific objectives that a model project or exercise actually can accomplish to serve a specifically stated purpose, whilst taking into account timeframe, various resource and other constraints, is often a difficult task. Furthermore, the specific objectives of the modelling, as initially stated, often warrant modification as alternative pathways of model development present themselves (Zare et al., 2020).

Groups involved in modelling socio-environmental systems (SES) generally include scientists from the physical, biological and social sciences, as well as a multi-vocational assemblage of stakeholders from different professions, occupations and pursuits who view the modelling project from different perspectives (Voinov & Bousquet, 2010; Wang & Grant, 2021). A diversity of perspectives on the problem at hand is a strength of interdisciplinary, multi-vocational groups, but can also be a potential source of strife. Difficulties arise when group members underestimate the overhead of interdisciplinary communication. As scientists, we are largely ill-trained to communicate our perspectives in a way that allows the development of a shared understanding of project objectives among stakeholders, who typically have different values and concerns, and suggests how project objectives might be achieved.

The typical research management framework is also ill-suited to facilitate the interdisciplinary communication required to address integrated modelling challenges. Arguably, these difficulties arise in large part because of a lack of transparency in discussions about scale decisions (Nabavi et al., 2017), frequently leading to unproductive, time-consuming arguments among group members, which can affect the trustworthiness and fitness-for-purpose of the final model. Such arguments during the problem scoping and conceptual model formulation phases often are exacerbated by a limited budget and impending deadlines, and limited data availability, as well as the expediency of already existing models ('models on the shelf') that may, possibly falsely, seem suitable for the current problem. Thus, a lack of transparency in group discussions can lead to scale choices made without sufficient awareness or reflection upon their impact on subsequent model development (Zare et al., 2020).

3. Disentangling issues around scales and communication about scales

The literature on scale issues arising during SES modelling is becoming voluminous, as is the literature on the involvement of stakeholders in the modelling process. Thus, for meeting best practice, it has been established what are some of the major scale decisions that should be documented (spatial, temporal, process detail, thematic, organisational) and for whom they should be documented (modellers, stakeholders, and/or the general public) (Figure 1).

As stressed in Section 1, scale issues arise in the early stages of modelling, and so communication on scale issues should take place during those stages of an SES modelling exercise. Modellers include those scientists who are involved in model development and their peers who potentially will review the model. Stakeholders include "non-scientist" stakeholders with an interest in the model and its output (e.g., natural resource managers, policy makers, risk assessors, investors), as well as "final impact" stakeholders who would be affected directly by decisions based on model output (e.g., farmers/fishers, developers, manufacturers, local residents) (Cartwright et al., 2016). Accordingly, the content of the "message" embodied in the conceptual model and style in which that message is transmitted to stakeholders, particularly to "final impact" stakeholders, should be tailored to the decision-making context (see literature review on stakeholder participation for environmental management by Reed (2008)).

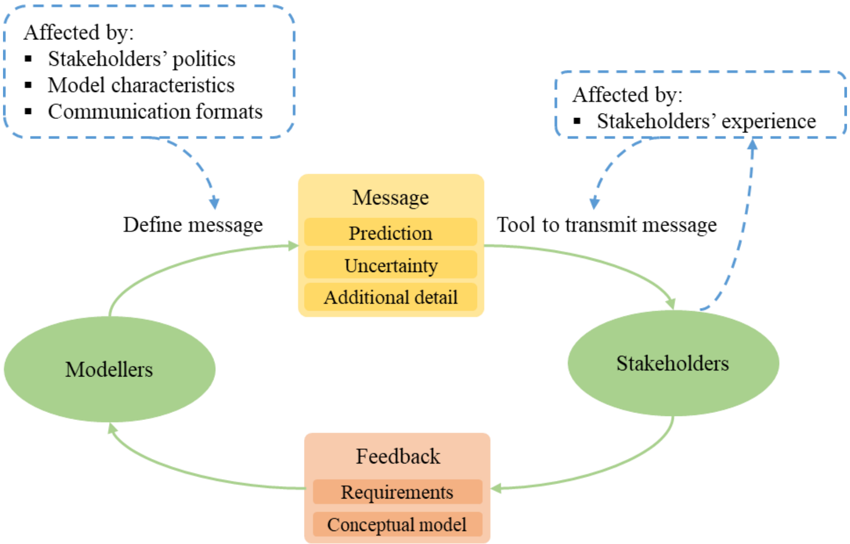

Figure 1: A general framework for the documentation of scale decisions during problem scoping and conceptual model formulation, indicating who should communicate what to whom. Adapted from Cartwright et al. (2016).

The content and style of documentation supporting effective communication of scale decisions, not only among those directly involved in the modelling process but also with the various parties interested in the model and its output, are arguably not well understood by modellers. The way scale decisions are documented should include consideration of both the interdisciplinarity of peers who will evaluate the scientific credibility of the model, and the level of experience of stakeholders who will assess the usefulness of model output within a particular socio-cultural context. In Figure 2 we have organised important scale issues that arise during problem scoping and conceptual model formulation using an Ishikawa fishbone diagram (Ishikawa, 1990). Below we will discuss each "bone" in detail. Our intent is not to provide a literature review, but rather to organise frequently encountered aspects of scale decisions within a framework that facilitates communication of the general set of guidelines that we will suggest in Section 4.

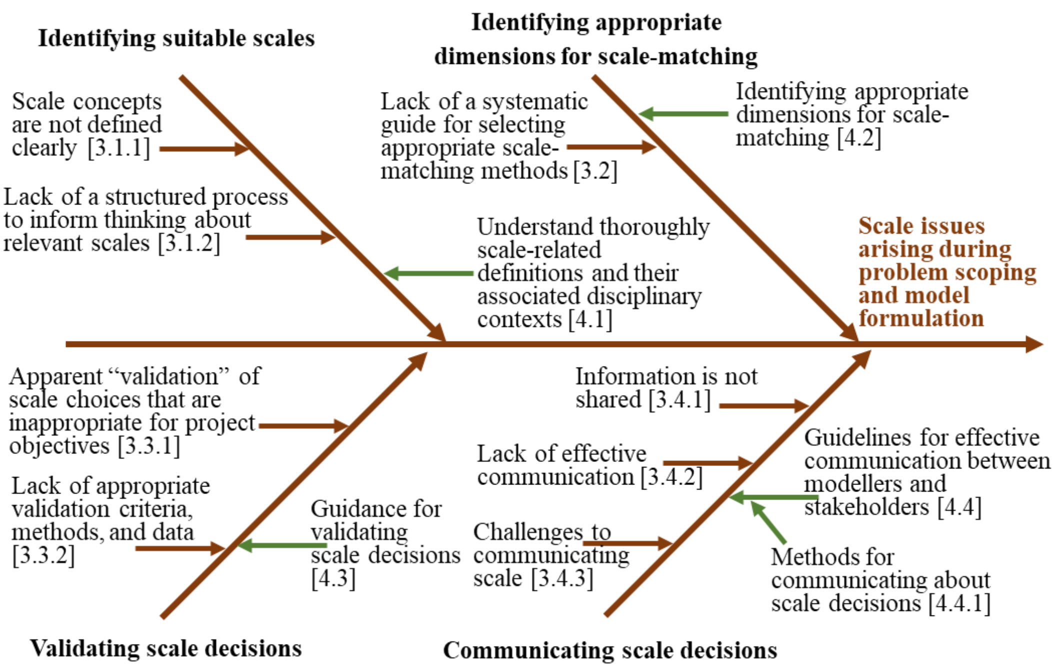

Figure 2. Important scale issues that often arise during SES problem scoping and conceptual model formulation (brown arrows), organised using an Ishikawa fishbone diagram (Ishikawa, 1990). Also shown are proposed guidelines associated with each issue (green arrows, discussed in Section 4).

3.1 Identifying suitable scales

3.1.1 Define scale concepts clearly

Scale and associated issues can originate from many sources, and these sources need to be identified and addressed before scales can be clearly defined. We list here issues arising from:

- Ambiguity in use and meaning of the word 'scale' and words associated with it. For example, the word 'scale' mathematically can refer to a series of marks at regular intervals on a line used for measuring or a rule for determining the distances between these marks. Geographically, 'scale' refers to the ratio of size to indicate proportional dimensions when shrinking objects to represent them on a map. Three attributes of scale are extent (the 'width' of what is possible when measuring), support (the size of the discretised intervals of measuring), and coverage (the ratio between the intervals and the extent) (Bierkens et al., 2000; Gibson et al., 2000). In everyday language, 'scale' often refers to the size of something, e.g., "the large scale of the COVID-19 pandemic". (Note that 'resolution' refers to the level of detail at which an entity or process is represented.)
- Ontological differences. It is critical to clarify the ontology (Gotts et al., 2019). For example, when talking about 'farms', scale could concern the spatial boundaries of farms (measured in spatial units), the judicial farm units (possibly expressed in monetary units), or the position of farmers in a social network.
- Discrepancies among what is desired by the exercise or from a stakeholder perspective, the process understanding of the modeller, and the support available through data (van Delden et al., 2011). These do not necessarily overlap, as available data and the scale at which process(es) can be modelled may not provide the support desired by the exercise (similar to balancing model complexity between model requirements and the available support from data) (Wagener et al., 2001).
- Differences in the objective functions that modellers and stakeholders may have in mind, for which different scales may be applicable, including different appropriate or preferential scales for different processes and related model components. Any effort to simulate the behaviour of an SES will have to make deliberate decisions on scales. There may not be one 'perfect' scale for the SES model under consideration.
- Mismatches between the real-life scale of attributes and their discrete classification and/or numerical representation in models (e.g., the classification of soil types and the implementation of layers in soil hydrological models). Interpolation and extrapolation may (partially) solve issues associated with scale mismatches between model and data, but this usually comes at the expense of increased uncertainty about model predictions.
- The (implicit) assumption that attributes are scaled homogeneously, whereas they may scale heterogeneously. For example, in climate studies it is not uncommon to select only the dominant soil type and cropping system in a fixed radius around individual weather stations and assume these occupy the whole area, disregarding other soil types and cropping systems in that area that may still play a substantial role (Schils et al., 2022). For some areas this homogeneity may be roughly accurate, while in other areas soil types or cropping systems may be very diverse.

A particular feature with modelling SES is that 'scale' can encompass multiple aspects. Many ecological studies consider only time and space (Turner et al., 2001). Although many 'natural' boundaries are related to time and space because they are defined by dominant biophysical, geological or chemical processes (e.g., in the case of a catchment area for groundwater modelling), identification of the scale of these boundaries may not be trivial. Boundaries may also be the result of underlying processes. For instance, several patterns observed in ecosystems can be explained by smaller-scale mechanisms of resource concentration (Kéfi et al., 2010). Moreover, SES also typically involve 'artificial' boundaries. Even though some of these artificial boundaries concern time or space, like country borders or the administrative start or finish of political cycles, many concern other attributes relating to thematic or organisational scales.

The choice for organisational scales may be rather straightforward, for instance, many judicial or governmental units enforce a 'pyramid-like' structure in which one agent controls multiple agents 'downstream'. Ecological food webs may follow a similar structure, in particular when functional groups are considered (Allen & Hoekstra, 1994). Often, however, the organisational scales are debatable. For example, scales may result from social processes, such as networking among actors. Thus, agents can have different roles (people are workers, team members, and parents, for instance), where each role can be accompanied by different action modes, and agents interact with other agents from the same group with the same role (working together, playing as a team, and being with the family) – the so-called 'agent-group-role' paradigm (Brugière et al., 2022). Further entanglement of scales may occur when agents are part of different social and organisational networks. Decisions need to be made about the number and type of actors and the level of differentiation (e.g., number of actor groups, household/farm types, type of irrigation licence, types of technology, types of behaviour), as well as decisions about the thematic resolution in all other parts of the model (e.g., land use classes, crop types, economic sectors).

3.1.2 Towards a structured process to inform thinking about relevant scales

Good modelling practices should include explicit discussion and documentation of the choices of scale that were made, why they were made, and by whom. We argue it is best to not aim for agreement about scale at the start, but rather to allow room for discussion and to explore different scale options, including the possibility of having multiple scales.

For each different viewpoint provided by stakeholders, at least the following questions/items should be considered:

- Are there natural or artificial boundaries to time, space, or otherwise?
- Are there agents in the system? What limitations or demands on the scale do they impose? What are the options for agents to act that are relevant for inclusion in the model? What is the best level of detail at which to represent the different (groups of) agents?
- Are there any emerging scales, such as spatial patterning or social clustering?
- Regarding space, is the best approach a point model, or a one-, two-, or three-dimensional model?
- If maps are used, are they grid-based or vector-based?
- Are there any transformations used, like (dis)aggregations, interpolation or extrapolation?
- What heterogeneities are there, and what is their origin?
- What is the extent, support and coverage of the data? And of the model?
- What is desired from the application side (this may be stakeholder-specific)?

3.2 Identifying appropriate dimensions for scale-matching

Once the sources of scale issues are clarified, and discussions about the desired scale(s) are underway, several new questions must be addressed about the matching of scales. As mentioned above, what is desired as scale and what is supported by data and process knowledge regarding the scale may not overlap. First, what are potential scale-matching methods and what is meant by it? What is being matched to what? Possible answers involve: (1) a match to what the different disciplines involved in the modelling process find important for answering the research question, (2) a match to the resolution and extent of the available data, or (3) a match to the resolution and extent of the envisioned application scale. It is important to realise that selection of model scale commonly involves a trade-off, or tension, between the scales desired for model application and the scales supported by the available data.

As mentioned briefly in the introduction, we should further recognise that there are different types of model scales. Such scales are intrinsic scale (the scale at which a process operates), observational scale (the scale of measurement and sampling), and policy scale (the scale for policy making). The relation among these scales is discussed by Wu & Li (2006) and Bierkens et al. (2000) and has been taken into consideration regarding the 'model scale' decisions that we focus on here.

Sitas et al. (2021) provide a chapter on system scoping in SES research, including a table of methods containing policy scoping and social-ecological inventories, and links to introductory texts describing associated methods. Furthermore, they acknowledge that system scoping is commonly constrained by institutional and biophysical boundaries. Existing conceptual frameworks directed at system scoping include the Intergovernmental Science Policy Platform for Biodiversity and Ecosystem Services (IPBES) framework (Díaz et al., 2015), and Ostrom's framework for analysing the sustainability of SES (Ostrom, 2009). These conceptual frameworks can help to set the boundaries, conditions and variables for a deeper exploratory or analytical exercise based on addressing specific questions or problems.

3.3 Validating scale decisions

3.3.1 Avoid apparent "validation" of scale choices that are inappropriate for project objectives

As questions of scale have a major influence on the design and functionality of the model, we explicitly discuss the validation of scale decisions here. Model validation broadly aims to answer the question "Did we build the *right* model?", thus establishing that the model is a useful and credible representation of reality in the eyes of the problem owner(s). In models where scale plays a significant role, scale validation is a necessary precondition for establishing this credibility by answering the question "Did we choose the *right* scale?" It should be noted that, despite the phrasing of these questions, validation is never "complete" in a binary sense, but always "more or less" done.

As with model validation, scale validation should be performed with specific criteria in mind. One such criterion might be that the scale in the model matches some real-world scale (which is deeply intertwined with the observational scale), and therefore represents real-world entities at a comparable level of detail. Another example of a criterion might be that the chosen scale permits the replication of some observed real-world pattern (Grimm & Railsback, 2012). By underpinning the validation process with explicit criteria, the bounds of this incompletable process can be maintained, and its limitations clearly documented.

However, data about the real world are rarely available at exactly the scale required by problem owners. Model-based analyses are therefore often conducted at scales unsuitable for the problem owners, as analysts choose to adjust their models to scales at which data are available (Meinen & Robinson, 2021). While pragmatic, such a decision invariably invalidates both the scale and the model. After all, if the analytical outcomes are not useful for the problem owners, neither the scale nor the model is *right* in the sense described above.

To alleviate the disconnect between desired scale and data availability, four options can be proposed. First, the model may be validated at a scale known to be incorrect, and then be supplied to problem owners with an explicit warning. Second, exploratory modelling (Bankes, 1993) could be used to evaluate multiple alternative scale choices and the resulting model behaviour. This evaluation could potentially reveal that model behaviour is insensitive to scale choice, that is, regardless of the scale at which the model is implemented, the resulting analytical insights are broadly the same. Third, it may be necessary to use validation methods that are not data-driven, but take a more qualitative approach, such as model reviews with domain experts. Finally, the increasing availability of low-cost data collection methods such as digital survey tools, commercial satellite imagery, and uncrewed aerial vehicles can lower the barrier to collecting new data at the required scales (Meinen & Robinson, 2021).

3.3.2 Partition validation concepts: conceptual, operational and data validity

There seem to be no generic criteria for the validation of a scale. What is credible and useful may depend on different things. Validation tests focused on scale choices should include both qualitative and quantitative measures of system performance, and project-specific tests are important. From a practical point of view, validation of scale may be undertaken together with model validation, while keeping in mind the distinction between the two. For scale validation, we cannot simply replicate all methods that are available for model validation, but some model validation methods do focus on scale. It seems useful to consider a minimal partitioning of validation concepts, namely (1) conceptual validity, (2) operational validity, and (3) data validity (Rykiel, 1996).

Conceptual validation of scale choices requires that theories and assumptions underlying the scale choices are correct, or at least justifiable, and that the scales used to represent the system of interest are reasonable for the intended model use. It also requires a scientifically acceptable explanation of the cause-effect relationships operating at the chosen scales, and/or a justification for the aggregation or disaggregation of known processes to the chosen scales. Such justification may include a rationale for omitting processes known to be involved in the system's dynamics, and for using representations known to be false at a chosen scale. Differences in scale among model components may necessitate aggregations that are conceptually erroneous when viewed at one scale in order to produce operational results that are acceptable.

Conceptual hypothesis testing via pattern-oriented modelling is one approach that integrates both qualitative and quantitative aspects of model validation. The approach involves the identification of questions with a known range of acceptable answers at the scales chosen, and the evaluation of the model's ability to produce these scale-dependent patterns of system behaviour (Grimm et al., 2005). The greater the numbers and ranges of questions and patterns over which the model can be tested, the better. Expert validation through one-on-one dialogues or serious games often may be useful, and can close the loop on conceptual modelling early in the development process (Elsawah et al., 2015). One practical approach would be to ask stakeholders who initially were not involved in the modelling project if they feel the current scale choices are appropriate for the given problem. We can also compare our scale choices with those of other works on the same topic, preferably involving a justification of their choice.

Operational validation of scale choices requires that model output meets the performance standards set for the model, given its purpose. For instance, when one is trying to reproduce patterns using a model, one would like to be able to compare model output with observed real-world data, although this can be tricky with models integrating many components (Louie & Carley, 2008). Operational validation is a pragmatic approach to scale validation, being concerned primarily with how well the model mimics the system of interest at the scale chosen for model outputs. Historical replay is one such method that involves running the model with data independent from the data used to calibrate the model, and then seeing if the model generates realistic historic data when run at the currently chosen scales. One can also see if model outcomes change substantially when the chosen scales are changed, which may involve a different resolution when space or time scales are concerned, or a more fundamental difference such as the level of process detail. Statistical tests comparing simulated and real data are used to evaluate how significantly the model output corresponds with observed data. However, such correspondence does not imply that the scientific basis of the model or the spatial and temporal scales embedded in its internal structure corresponds to the actual processes or cause-effect relationships operating in the real system. Failure to achieve operational validation may reveal underlying conceptual problems, often involving issues of scale.

In regard to *data validation*, it is important to emphasise that data are not an infallible standard for judging model performance. Rather, model output (simulated data) and data from the real system are what have been referred to as "two moving targets" that we try to overlay one upon the other (Rykiel, 1996). The data used to validate the scale and/or the model must themselves be validated (Robinson, 1997). Data validation requires that the data and its processing meet a specified standard (quality assurance) and that the interpretation of the data is demonstrably valid (Sargent, 2008). It follows that the standards set for data quality must preferably be set at the same scales of the data collection. Likewise, the initial demonstration of the validity of data interpretation must occur at the same scale as the data collection. The establishment of the validity of subsequent aggregations or interpolations of the data, which typically are involved in passing information among models or sub-models with different scales, involves necessarily subjective and often controversial decisions. Where these decisions must be made, it is especially important to incorporate a broad range of perspectives on the studied system so as to fully understand the potential drawbacks of a specific choice (Hamilton et al., 2015). One final consideration related to data validation is that even high-quality data represent only one historical trajectory of the system, one manifestation of reality. That is, just as an SES has more than one possible future trajectory, it had more than one possible past trajectory, only one of which was realised and provides us with empirical data. Assessment of data validity requires consideration of the likely magnitude of differences among data from alternative historical system trajectories given the scale at which the data were collected.

3.4 Communicating scale decisions

As a central thesis in this paper, we argue that communication with stakeholders is of key importance both during the discussions on scales, and at the end of these discussions when a decision is being or has been taken regarding scales. Important information to share includes: (1) What is the actual decision, and why? (2) Who made the decision, and how was it made? (3) What are the expected consequences of the decision?

3.4.1 Information to share

Regarding item (1), the actual decision and why, it is relevant to communicate what decision has been reached. Only one scale may have been chosen for some phenomena, or more than one scale, or perhaps no agreement has been reached (there is an agreement to disagree). A distinction must be made between spatial scale(s), temporal scale(s), thematic scale(s), and organisational scale(s), if applicable. In addition, a justification should be given in regard to why the scale decision is what it is, as well as a concise description of the path to the decision. This information makes it possible to review the decision and potentially provide input to alter the decision, if applicable. Reasons for a certain choice can involve the desired scale, for example, the (un)availability of information on stakeholder requirements, or support for certain scales, for example, related to biophysical or socioeconomic limitations or the (un)availability of certain data.

Regarding item (2), who made the decision, and how, it is relevant to explain the process that was followed to reach the decision. Who was involved, and in what capacity? How much time was dedicated to discussing scales? Were dedicated exercises, workshops, or maybe even serious games organised to explore scale issues and possible outcomes of decisions? Was the decision the result of a democratic process, or did a particular person (for example, the project leader or a specific stakeholder) make the final decision? What attempts were made to avoid bias in the decision? Ideally, as already motivated in the Introduction of this paper, any scale decision should have broad support to avoid further issues down the road. A proper written description and justification should be available, including documentation of the roles played by people involved in the various phases of the decision-making process.

Regarding item (3), the consequences of the decision, a written reflection on the ramifications of the decision should be provided. What options were not selected, and why? What could be the result of not selecting those options? Ramifications of certain decisions need not only be numeric or categorical, in terms of model output and uncertainty; there can also be social ramifications, for example, that the minority who were not in favour of the scale will show a reduced willingness to continue participating in the remainder of the project.

To address the three key items mentioned above, those in charge of a project should obtain clarification of the decision objective or research question. This clarification is essential for the justification of the scale decision(s). A project will seldom be directly aimed at scales, so a ‘translation’ of the societal and research objectives to scale decisions is usually needed. An identification of the relevant decision makers and stakeholders is equally important, and people in charge have a responsibility to justify the reasons certain stakeholders were considered for involvement in scale decisions or were excluded from such involvement.

3.4.2 Towards more effective communication

One of the most critical points of effective communication is the ability to understand the audience and being audience-centred. This involves an appreciation of the diverse knowledge about the subject, including scales, that exists among members of the audience (Hall et al., 2014). To reach appropriate scale decisions, we must communicate effectively. Modellers need to realise that behavioural traits can affect group interactions and effective communication as well as lead to the possibility of procedural mistakes and cognitive biases (Hämäläinen et al., 2013). We should anticipate how and why different audiences might fail to understand various aspects of these decisions. We also should acknowledge that in some cases disagreements over scale issues cannot be resolved definitively. In such cases, we should agree to disagree and move forward with the modelling process in search of new knowledge and/or experience that will help resolve the scale issues.

Effective dialogues with stakeholders to reach a shared decision about scales require a two-way flow of communication (Carrada, 2006). Currently, not enough two-way communication about scale issues happens during the modelling process (Zare et al., 2020). The idea that society must understand the ivory tower knowledge of science is not appreciated because it is based on the idea that knowledge is a one-way flow from scientists to society. Modellers need to avoid using technical terms without simple explanations of their meanings. The inability to communicate effectively is often due to the different mental models held by the different participants in the modelling process (Hall et al., 2014). Furthermore, differences in the level of interest in, and understanding of, the role and importance of scale (and other modelling terminology) exist not only among stakeholders but also among modellers. Modelling procedures in different disciplines have differences in type, steps, specifications, assumptions, and the importance of the role of scales. Even experts tend to communicate poorly about the scale aspects of their modelling, and their pre-existing experience can affect their level of interest and knowledge about the importance of scale issues (Hall et al., 2014). Documentation of, and access to, the scale decision-making process in modelling exercises can only ameliorate such situations and enhance good practice in the longer term.

3.4.3 Challenges to communicating scale

There are several challenges in relation to communicating scale concepts to non-technical audiences, including (1) confusion regarding scale concepts, (2) conflicts with pre-existing scales in other (modelling) studies, (3) the use of inappropriate communication methods (first and foremost, using 'slang' or jargon language that is too technical), and (4) the existence of cognitive boundaries and behavioural phenomena that inhibit effective communication (Glynn et al., 2017; Moallemi et al., 2020).

To address confusion regarding scale concepts, we must discuss the unique meaning of the concept. However, as mentioned earlier, the meaning of 'scale' may differ by trade, academic discipline, and geographic locale. Modellers should determine how audiences will interpret specific technical terms such as scale by simply asking for their understanding and definitions of them (Hall et al., 2014)

To address conflicts with pre-existing scales in other (modelling) studies, we should acknowledge that scale choices in different studies on a unique geographic location could be different based on different aspects of the study, which is why it is important to provide a motivation for decisions on the scale. Also, stakeholders and modellers might have experiences with studies using different scales. These conflicts with pre-existing scales can be challenging to overcome. Communication with a transformative explanation that acknowledges the audience's pre-existing knowledge and transforms it with alternatives could help in overcoming this obstacle (Hall et al., 2014; Modell, 2009). Steps in using a transformative explanation to communicate the scale include: (1) stating or asking questions to elicit the existing description of the scale and possible choices, (2) discussing the plausibility of existing scale choices by including supporting observations, (3) discussing and demonstrating with examples why and where the existing scale choices might not be the best choice, (4) presenting alternative explanations, descriptions and choices for the scale, and (5) explaining why and how any new scale is a better choice (Hall et al., 2014; Rowan, 2009).

Presenting dummy results, whereby a simplified and mock-up of the results showing what the eventual outputs might look like by choosing different scale options, can be a useful method to demonstrate the plausibility of existing scale choices. Dummy results can demonstrate the effect of different scale choices on the results and their importance and could foster critical thinking and reflective practice. Additionally, using examples of other case studies could be beneficial in explaining the choices. A simple risk analysis, such as structured risk analysis (McEvoy & Whitcombe, 2002), to highlight the trade-offs involved in selecting the different choices can also be useful. For detailed information about these methods please check Zare et al. (2020).

4. Description of proposed guidelines

In this section, we present our proposed guidelines aimed at addressing scale issues in integrated SES models during the problem scoping and conceptual model formulation phase, linking them to the "fishbone" framework described in Section 3 (Figure 2).

4.1 Understand thoroughly scale-related definitions and their associated disciplinary contexts

When modelling SES, multiple dimensions of both environmental and socioeconomic factors must be considered simultaneously. A prime requirement for the development of an integrated SES model is the identification of suitable scales for the representation and linkage of the system components to be considered for inclusion in the model. As always, the first step is to define the objectives of the modelling exercise: What questions are to be addressed with the model and who are the stakeholders who will be using the model output? This will determine bounds to scales as well as system and subsystem boundaries, and the separation between what elements are considered endogenous and exogenous. Typically, a more detailed representation and consideration of scale is entailed for endogenous elements crucial to meeting the modelling objective(s), than for forcing exogenous variables. The challenge is to identify coherent combinations of spatial, temporal, thematic and organisational scale details across the system representation.

We propose a set of general considerations and principles to help identify suitable scales for SES modelling and its documentation:

- **A1**: Realise there could be a *single extent* of the integrated model, but there is *no single scale* and *no single resolution* for it. Even within the same SES conceptual model of a given problem type or set of issues, interdependently-nested scales typically are involved in which model component scales will change in relation to that interdependence. Furthermore, realise there are different scales to decide upon, namely spatial, temporal, thematic and organisational.
- **A2**: Recognise that in general there is no predefined scale(s) for SES modelling in general because it depends on the objective(s) of the exercise, the knowledge and data available and indeed stakeholder perspectives included.
- **A3**: Look for a joint definition of objectives and system boundaries. This should guide the identification of system elements, exogenous drivers and endogenous variables, and their scales.
- **A4**: Consider proper data typologies and anticipate that there may be mismatches that need to be resolved. Data typologies can also help guide the definition of scales and resolutions. For example, in the case of remotely sensed data, space and time are fixed by the platform and the sensor used. However, scale and resolution can be aggregated or interpolated if needed. The scales to be jointly considered may have natural boundaries, but a common occurrence is that data with different typologies (e.g., environmental data versus economic data) will generate mismatches. An example is the integration of data based upon natural boundaries such as watersheds or soil units versus data referring to administrative units such as social or demographic units.
- **A5**: Expect that there is often a trade-off between desired scales and their support from the data. Data availability constrains scales and resolutions, at least at the level of model inputs. In general, a higher level of detail and accuracy is needed for those variables that are crucial for the dynamics to be simulated. Sensitivity analysis may help to identify these variables and supports the analysis and mapping of uncertainty. No simple rules exist, but it should be pointed out that in some cases these analyses could demonstrate that the declared objectives cannot be met with the available information and thus suggest that the modelling exercise should be abandoned. Mismatches between model scales and resolutions and those of available data can be treated to some extent by adopting a common discretisation approach in space and time, and again by initiating the discussion regarding uncertainty.
- **A6**: Initiate and resolve negotiations within the modelling team and with the relevant stakeholders to specify the definition of scales, resolutions and extents (in the following Subsections we discuss how to organise these negotiations).

In Box 2, we discuss two examples to illustrate this set of rules.

4.2 Identifying appropriate dimensions for scale-matching

Following the above set of principles, the appropriate scales should be identified. As the examples in Box 2 illustrate, there are alternative temporal, spatial, thematic and organisational scales that can be selected. For each intended application, scientists and stakeholders should consider alternative scales. Rather than presenting a formal/statistical approach to the selection of the right scale, we propose the use of a set of questions to aid in scale selection. These questions are based on a synthesis of considerations presented earlier and principles A1 through A6. This approach involves considerable communication between modellers and stakeholders (for the 'how' on communication, see Subsection 4.4). Questions to guide decisions for identifying suitable scales are:

- **B1:** What does scale mean in the project (see Subsection 3.1.1)? Is there a clear and unified use of ontology? Have discrepancies been clarified between stakeholder desires, the modellers' understanding, and data?
- **B2**: What are the desired scale properties for the different disciplines? What are the reasons for the acceptance or rejection of particular scales?
- **B3**: Depending on the above, which temporal, spatial, thematic and organisational scales have been selected for? Have these choices and their justification been documented during the modelling process, including why certain options were not selected?

Box 2: Two examples of identifying suitable scales.

As a first example, consider a modelling exercise aimed at analysing a rural SES, which, at a high level of detail, exhibits extreme complexity in terms of its numerous socio-environmental processes and their interrelationships. Our modelling objectives might be related to the assessment of the environmental impacts of agricultural activities. In this case, we would need to build an SES model with relevant, though simplified, complexity. Our SES model would require the identification of system boundaries in terms of a geographical unit such as a watershed, within which one could observe the main phenomena of pollution generation, transport and discharge leading to environmental impacts such as the eutrophication of water bodies. Examples of relevant elements to be considered with greater granularity, i.e., with higher spatial, temporal and thematic resolutions, would be the unit of agricultural production in which management practices could generate pollution phenomena. Typically, these are fields, i.e., the management units for the farmer. Given the relevance of soils and water balance for those phenomena, we should expect the spatial modelling unit would derive from the intersection of hydrologic units (soil and water combinations), which could be smaller than the size of a cultivated field. In that case, the level of detail of the simulation would have a sub-field spatial scale. A consistent time scale could be discretised on a daily basis, but this would depend on the dynamics of the spatial units. Ideally, the simulation should last for several years, enough to allow for an adequate number of climate events to be described to capture the variability of the system response.

As a second example, the same rural SES in terms of its inherent complexity could be studied for different purposes. For example, aspects of the same SES might be studied with the objective of conducting economic analyses of the market for agricultural commodities, which would result in different conclusions regarding scale, extent and resolution. In that case, the area defined by a watershed would not be so relevant. Of more importance would be the area identified by administrative boundaries managed by the same agricultural policy instruments or having a unique market as the destination of the agricultural products. The spatial size could be similar, but the boundaries would be different, as would the main elements to be identified as simulation units. In this case, farms typically would identify the main decision unit of the SES. Soils and the water cycles would be less important, and they could be aggregated at coarser scales by identifying areas with different average productivity. Climatic variables could be described as annual or multiannual averages, rather than using a daily resolution, as in the example of the previous paragraph. The prices of production factors, however, might need to be described on a relatively fine time scale that captures significant fluctuations in price.

Similarly, exogenous drivers of interest would be different in these two examples. In the first example, exogenous drivers could be regional or national environmental policies and regulations. In the second example, they could be global prices of agricultural commodities and international trade regulations. But in both cases, climate might be an exogenous driver though its temporal resolution, say, may need to be different. In the evaluation of scales, the underlying processes must be considered. Questions to be considered are: What is the minimal set of processes and variables that need to be included? What seem to be important processes, but are too complex to include? And what would be nice to include, but not essential?

In both examples, nested scales are needed. Spatially, local management practices might be represented by more detailed variables defined at the farm level. Regulations affecting agricultural production might be defined at a regional scale, and factors affecting market dynamics might be defined at either local or global scales, or both, depending on modelling objectives. Temporally, socio-economic variables might have relatively long or relatively short time steps. Commodity prices, for example, might be fixed as parameters in some exercises with limited emphasis on economics, or they might be monthly values of exogenous drivers when economic optimisation is one of the objectives. Or they could be endogenous variables in large-scale simulations, such as global climate change general equilibrium modelling, in which an annual time step is typically adopted. An actual case study in Thailand can be explored further in Trébuil et al. (2005).

Importantly, all considerations in the choice of processes and scales should be documented and explained for future reference as input for scale validation (see next Subsection). Reasons for the selection or rejection of some scales may be due to direct technical modelling considerations, stakeholder desires, lack of data to support a particular scale, or the problem at hand may not require representation at a particular scale. For example, even if remote sensing data allow for a small-grained spatial resolution, it may not be required for the application. Another example of rejecting a scale may be the uncertainty associated with the (dis)aggregation of available data to a particular scale. For example, even if disaggregation methods are available, the data may be too coarse for the desired disaggregation level.

4.3 Validating scale decisions

When deciding on scale, the choice of scale should also be validated. We propose the following guidance regarding the validation of scale decisions:

- **C1**: Recognise the difference between model validation and scale validation. Also, be aware that what is invalid for one type of application may be valid for another.

- **C2**: Re-validate scale decisions if the context changes. Scale choices are assessed within the domain of model applicability (*sensu* Rykiel, 1996) and are validated under the specified conditions that the model is designed to simulate.
- **C3**: Expect and embrace change in the selection of scale choices as inevitable due to the complexity of the system being studied and the speed at which new information may become available. During problem scoping and conceptual model formulation, it is common to go through several iterations of a set of concurrent steps/activities. The flexibility to formally assess alternative scale choices must be maintained, as this is critical to problem scoping and conceptual model formulation. Scale choices should not be invalidated prematurely, and justifications should be provided for decisions to declare a scale choice (in)valid. Such justifications add to the possibilities of exploring alternative pathways.
- **C4**: Clarify which scale and model validation criteria and approaches are used, and why. Consider the distinction between conceptual, operational and data validity (see Subsection 3.3.2 for suggestions).
- **C5**: Realise that data do not represent an infallible standard in the sense that their scale should not solely dictate choice. Explain the origin of the involved data and how they are used in the scale validation. This especially includes data that have been interpolated or extrapolated.
- **C6**: Try to include both qualitative and quantitative measures of system performance. Project-specific tests are important. Identify questions with a known range of acceptable answers at the scales chosen and scale-dependent patterns of system behaviour that the model is expected to produce.

4.4 Communicating scale decisions

Communication and justification are the central issues around scale choices. Enhancing the quality of communication and documentation is crucial for improving the transparency of the modelling process and stakeholder trust in the model. Subsection 4.1 discussed considerations and principles to help identify suitable scales, Subsection 4.2 presented questions to consider in choosing a scale, and Subsection 4.3 dealt with scale validation. This Subsection considers some good communication practices during all three steps and is intended to aid model reproducibility and communication effectiveness and thereby help both modellers and stakeholders. We propose the following guidelines for effective communication between modellers and stakeholders regarding the decision making related to scales:

- **D1**: Get clarity on the decision objective(s) or research question(s). This precision is needed for justification of the scale decisions adopted. A project will seldom be directly aimed at scales, so a ‘translation’ of the societal and research objectives to scale decisions is usually needed.
- **D2**: Analyse the audience of stakeholders for possible inclusion or exclusion in the communication process (see Box 3 below). Be aware of differences in the level of interest in, and understanding of, the role and importance of scale among stakeholders and modellers. Do not consider the ‘scientific view’ to automatically be the right one. Recognise that behavioural traits can affect group interactions and effective communication, and that they may lead to the possibility of procedural mistakes and cognitive biases; for instance, focusing on the time scale and ignoring important organisational scale aspects, or emphasising fast practical model use while overlooking concerns regarding scale validation.

Box 3: A list of questions for project leaders and teams to consider in analysing an audience of stakeholders for possible inclusion or exclusion.

1) Who are the decision makers and stakeholders?
2) What are their interests, desires and motivations for participating in project planning?
3) How do audience members understand the system of interest?
   (i) How do they interact with, or experience, the system?
   (ii) What features of the system do they pay attention to?
   (iii) How do they monitor or observe the system?
   (iv) How do practitioners make sense of changing conditions?
   (v) What local vocabularies of terms, images and concepts do they use to explain various system aspects and characterise problems within the system?
4) What other driving factors, significant events, and past conflicts with the decision-making context should be considered?
5) What prior experiences have audiences had with models?

- **D3**: Pay attention to communication between team members as well as stakeholders. In interdisciplinary research exercises such as occur in SES modelling, communication among team members is as important as communication between stakeholders and the research team. Improved communication among team members enhances the transfer of knowledge within the team and helps to achieve a strong connection among different disciplines, especially regarding critical decisions such as scale.
- **D4**: Organise communication in a suitable way; in the Subsection below, we discuss some approaches that can be followed. Important principles are:
    1) Involve sufficient two-way communication between modellers and stakeholders.
    2) Consider the earlier-mentioned challenges with communicating scale concepts to non-technical audiences, including mitigating (i) confusion regarding scale concepts, (ii) conflicts with pre-existing scales in other studies, (iii) the use of inappropriate communication methods, and (iv) the existence of cognitive boundaries and behavioural phenomena that inhibit effective communication.
    3) Avoid technical language that is unnecessary or without proper explanation.
- **D5**: Provide concise and comprehensible documentation with respect to the important information regarding decisions and the rationale supporting those decisions. When properly performed and completed, this documentation makes it possible for people from within and outside the project to review the decisions and, if applicable, potentially provide input to alter the decisions. It should contain essential information about the following:
    1) The actual decision, why it was made, who made the decision, and how it was made. For example, were dedicated exercises, workshops, or even serious games organised to explore scale issues and possible outcomes of decisions? Was the decision the result of a group discussion, or did a particular person (for example, the project leader or a specific stakeholder) make the final decision? If it were the results of a group discussion, who was involved in the process (Moallemi et al., 2022)?
    2) Documented considerations regarding whether we use one or more scales for the study depending on the different phenomena, and why (see A1), what is desired for a particular simulation or project (see A2), and the possible trade-offs between desired scales and support provided by data (see A5). Also include options for scales that were *not* followed up.
    3) A clear description of the joint definition of objectives and system boundaries (see A3) and the used data typologies (see A4).
    4) A description of the validation (see the guidelines under C), and a critical reflection of the expected consequences of the decisions regarding scale, including what scale means in the project (e.g., is it temporal or spatial?) (see B1), whether it should be the same for all disciplines or not (see B2), and what the desired temporal, spatial, thematic and organisational scales are for the social and for the environmental view (B3). Consider what could be the result of not selecting particular scale options? The ramifications of certain decisions need not only be in terms of model output and uncertainty; there can also be social ramifications, for example, that the minority who was not in favour of the chosen scale will show a reduced willingness to continue participating in the remainder of the project.
- **D6**: Use proper language. In their communication with stakeholders and fellow scientists, people should opt for the use of non-technical language and means of communication that are broadly supported and traceable. Avoid the use of scientific terms as well as popular 'slang' and establish a common ontology to avoid confusion. For example, a 'matrix' can be a mathematical object from linear algebra, the physical attribute of the soil, or the landscape over which a species is dispersed. This is confusing, and many stakeholders will not even know what this word means. Technical terms and implicit assumptions must be clarified.
- **D7**: Communication and documentation should be an ongoing, regularly occurring process, with sufficient information about the possible scale options and scale decisions presented in such a way as to be as transparent as possible.
- **D8**: Activities aimed at communication should be planned from the start of the project to ensure incentivised communication. The occurrence of many, short interactions among a few team members

has been found to improve the transfer of relevant project information between scientists and stakeholders (Podestá et al., 2013). Hence, ample communication moments should be planned and followed.

4.4.1. Methods for communicating about scale decisions

Different ways of communication should be used to maximise the probability of getting the information across. Communicators should use simple exercises and tools to reduce confusion about the scale concept, identify the available options, and demonstrate the process. We list some available tools and suggestions for selecting tools:

- **E1**: Using ID cards (Zare et al., 2019). An ID card is an organised collection of explicit statements about each step of the modelling process which can be used to communicate a common story about the scale selection step and a better understanding of the scale selection process between modellers and stakeholders. Such a card shows a quick summary of what has been said about this step across multiple guidelines using a common language and provides information about different aspects of the step such as aim, actors involved, and main decisions. Team members should fill out such cards.
- **E2**: The use of narratives with examples and a scientific storytelling approach has been found to be useful in the communication of complex information to non-specialists (Howarth & Anderson, 2019). This could be, for example, defining the spatial and temporal scales in narratives about the case study in a different time and physical boundary and conveying the differences within the story.
- **E3**: The elucidating explanation of the concept (Hall et al., 2014). This uses an example of the concept, here scale, that lists the concept's essential features and also offers a list of examples and non-examples. Therefore, the audience has the opportunity to understand the difference between examples and non-examples by looking for their essential features. This tool is helpful if the concept of scale is mistaken with other concepts such as resolution or if there are several similar options for scale.
- **E4**: The practice of regular reflection on the modelling pathway during a project, particularly with respect to scale choices, can increase transparency and trust among stakeholders and help them reach better decisions as well as promote efficiency and effectiveness of the modelling process (Zare et al., 2021).
- **E5**: Documentations should include the following information, noting that stakeholder reflections on the information should be solicited to avoid one-way communication:
    1) the scale decisions that were made,
    2) the other scale options that were discussed but not chosen,
    3) what each alternative option might mean in terms of possibly affecting project results, costs and next steps in the project,
    4) what each group of stakeholders thought about the scale decisions, and
    5) the final decisions and selected scales.
- **E6**: Tools for documentation should be more than just deliverables and reports at the final stage of the project. Metadata, summaries and easy-to-understand fact sheets for each document could convey the meaning and aim of each document to people who will not read the details in a final report.
- **E7**: Reflective communication: Reflections could be solicited within a workshop or by sending a list of reflective questions with any document to facilitate engagement and draw attention to the points of concern. Reflection to validate, confirm, or debunk the result of a decision (such as scale selection) should happen frequently, and thereby assist in efficiently reaching objectives by identifying the need for iteration as early as possible (Zare et al., 2020). Steps in the reflective cycle include:
    1) describing the event and situation,
    2) expressing the sentiments and thinking about them,
    3) assessing and analysing the discussion, and finally,
    4) concluding and setting an action plan (Gibbs, 1988).

5. Summary and suggestions for future work

The choice of scales is an inevitable and critical decision in SES modelling. But the interdisciplinary nature of SES modelling may easily create confusion and frustration when discussing and communicating scale issues during problem scoping and conceptual model formulation. Compromises or sometimes simply convenient choices may be made to accommodate appeals from researchers, stakeholders, and policy makers with specific interests in a particular discipline, question, field or scale. Thus, the final scale choice for any phenomenon is not necessarily the optimal or most scientifically solid and justifiable for all envisioned applications. Accordingly, there are significant benefits in developing a better understanding of issues surrounding the selection of scales in SES modelling.

This paper is intended to stimulate further discussions and inspire future work surrounding the issue of documenting scale choice in SES modelling. We foresee future work including, but not limited to:

- The development of an improved understanding and alignment of scale and associated concepts within the SES and other relevant modelling communities. Despite the fact that many articles focus on the scale issues, confusion regarding scale concepts and lack of understanding of the choice of scale and its implications are still prevalent among modellers and stakeholders.
- The development of education and training materials with easy-to-understand (non-technical, discipline-neutral) terminology in order to advance the handling of scale issues and build associated capacity within the scientific community. These are vital to establishing a common understanding of the concept of scale and relevant terminologies between researchers and non-scientific actors (stakeholders, policymakers and citizens) as well as modellers from different disciplines (the natural sciences and social sciences). We suggest that special workshops could be organised on topics related to scale in SES modelling. For example, topics could include the importance of scale issues, the implications of scale choices, the limitations of models, and the ramifications for modelling results following a particular choice of scales.
- The application of scale guidelines in SES projects and their subsequent refinement to make the discussions and decisions around scale issues transparent and traceable. The potential benefits can only materialise when suggestions for communication and documentation are adopted and applied in SES modelling practice. In the present work, we have presented a preliminary set of guidelines, which now need to be tested 'in the field'. Through real case applications, we hope to further improve the guidelines and make them operational. A clear protocol or template illustrated with examples would be extremely helpful to facilitate the usage of the guidelines. Eventually, an automatic protocol like a web-based tool could facilitate the scale decision process with step-by-step questions.

In this position paper, we have attempted to identify the potential issues, traps and pitfalls in the process of decision making on scale issues and have presented a set of prescriptive guidelines on scale issues in four categories. We argue that following these guidelines can facilitate discussions between researchers and stakeholders, support the justification of scale choice, and help assess the strengths and limitations of models and their outputs.

Acknowledgments

This paper builds on discussions from the Lorentz workshop "Participatory and Cross-scale Modelling of Social Ecological System" at Lorentz Centre, Leiden, the Netherlands in June 2022. The authors acknowledge Lorentz Center, Leiden University and NWO for hosting and providing financial support.

References

Allen, T., & Hoekstra, T. (1994). Toward a definition of sustainability. In Sustainable ecological ecosystems: Implementing an ecological approach to land management, General Technical Report, RM-247 (pp. 98–108). USDA Forest Service.

Ayllón, D., Railsback, S. F., Gallagher, C., Augusiak, J., Baveco, H., Berger, U., Charles, S., Martin, R., Focks, A., & Galic, N. (2021). Keeping modelling notebooks with TRACE: Good for you and good for environmental research and management support. Environmental Modelling & Software, 136, 104932. https://doi.org/10.1016/j.envsoft.2020.104932

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

Bankes, S. (1993). Exploratory modeling for policy analysis. Operations Research, 41(3), 435-449. https://doi.org/10.1287/opre.41.3.435

Bierkens, M. F. P., Finke, P. A., & de Willigen, P. (2000). Upscaling and downscaling methods for environmental research (Vol. 88). Springer.

Brugière, A., Nguyen-Ngoc, D., & Drogoul, A. (2022, 2022-December-01). Handling multiple levels in agent-based models of complex socio-environmental systems: A comprehensive review [Review]. Frontiers in Applied Mathematics and Statistics, 8. https://doi.org/10.3389/fams.2022.1020353

Carrada, G. (2006). Communicating science: Infrastructure, hman resources, international cooperation, research and development, environment and health, societal Issues, industrial innovation - A scientist's survival kit. Office for Official Publications of the European Communities.

Cartwright, S. J., Bowgen, K. M., Collop, C., Hyder, K., Nabe-Nielsen, J., Stafford, R., Stillman, R. A., Thorpe, R. B., & Sibly, R. M. (2016). Communicating complex ecological models to non-scientist end users. Ecological Modelling, 338, 51-59. https://doi.org/10.1016/j.ecolmodel.2016.07.012

Cumming, G. S., Cumming, D. H. M., & Redman, C. L. (2006). Scale mismatches in social-ecological systems: Causes, consequences, and solutions. Ecology and Society, 11(1), 14. http://www.jstor.org/stable/26267802

Díaz, S., Demissew, S., Joly, C., Lonsdale, W. M., & Larigauderie, A. (2015). A Rosetta Stone for nature's benefits to people. PLOS Biology, 13(1), e1002040. https://doi.org/10.1371/journal.pbio.1002040

Edmonds, B., Le Page, C., Bithell, M., Chattoe-Brown, E., Grimm, V., Meyer, R., Montañola-Sales, C., Ormerod, P., Root, H., & Squazzoni, F. (2019). Different modelling purposes. Journal of Artificial Societies and Social Simulation, 22(3), 6. https://doi.org/10.18564/jasss.3993

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. Socio-Environmental Systems Modelling, 2, 16226. https://doi.org/10.18174/sesmo.2020a16226

Elsawah, S., Guillaume, J. H. A., Filatova, T., Rook, J., & Jakeman, A. J. (2015). A methodology for eliciting, representing, and analysing stakeholder knowledge for decision making on complex socio-ecological systems: From cognitive maps to agent-based models. Journal of Environmental Management, 151, 500-516. https://doi.org/10.1016/j.jenvman.2014.11.028

Ford, A. (1999). Modeling the environment: An introduction to systems dynamics modeling of environmental systems. Island Press.

Forrester, J. W. (1961). Industrial dynamics. Cambridge: MIT Press.

García-Álvarez, D., Lloyd, C. D., Van Delden, H., & Camacho Olmedo, M. T. (2019). Thematic resolution influence in spatial analysis. An application to Land Use Cover Change (LUCC) modelling calibration. Computers, Environment and Urban Systems, 78, 101375. https://doi.org/10.1016/j.compenvurbsys.2019.101375

Gibbs, G. (1988). Learning by doing: A guide to teaching and learning methods. Oxford: Further Education Unit.

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(2), 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(4), 356-378. https://doi.org/10.1002/2016EF000487

Gotts, N. M., van Voorn, G. A. K., Polhill, J. G., Jong, E. d., Edmonds, B., Hofstede, G. J., & Meyer, R. (2019). Agent-based modelling of socio-ecological systems: Models, projects and ontologies. Ecological Complexity, 40, 100728. https://doi.org/10.1016/j.ecocom.2018.07.007

Grant, W. E. (1986). Systems analysis and simulation in wildlife and fisheries sciences. New York: Wiley.

Grimm, V., & Railsback, S. F. (2012). Pattern-oriented modelling: a 'multi-scope' for predictive systems ecology. Philosophical Transactions of the Royal Society B: Biological Sciences, 367(1586), 298-310. https://doi.org/10.1098/rstb.2011.0180

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, 280, 129-139. https://doi.org/10.1016/j.ecolmodel.2014.01.018

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(1), 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(23), 2760-2768. https://doi.org/10.1016/j.ecolmodel.2010.08.019

Grimm, V., Johnston, A. S. A., Thulke, H. H., Forbes, V. E., & Thorbek, P. (2020). Three questions to ask before using model outputs for decision support. Nature Communications, 11(1), 4959. https://doi.org/10.1038/s41467-020-17785-2

Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W. M., Railsback, S. F., Thulke, H.-H., Weiner, J., Wiegand, T., & DeAngelis, D. L. (2005). Pattern-oriented modeling of agent-based complex systems: Lessons from ecology. Science, 310(5750), 987-991. https://doi.org/10.1126/science.1116681

Guillaume, J. H. A., Helgeson, C., Elsawah, S., Jakeman, A. J., & Kummu, M. (2017). Toward best practice framing of uncertainty in scientific publications: A review of Water Resources Research abstracts. Water Resources Research, 53(8), 6744-6762. https://doi.org/10.1002/2017WR020609

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 and Software, 119: 418-432. https://doi.org/10.1016/j.envsoft.2019.07.007

Hall, D. M., Lazarus, E. D., & Swannack, T. M. (2014). Strategies for communicating systems models. Environmental Modelling & Software, 55, 70-76. https://doi.org/10.1016/j.envsoft.2014.01.007

Hämäläinen, R. P., Luoma, J., & Saarinen, E. (2013). On the importance of behavioral operational research: The case of understanding and communicating about dynamic systems. European Journal of Operational Research, 228(3), 623-634. https://doi.org/10.1016/j.ejor.2013.02.001

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., Pollino, C. A., Stratford, D. S., Fu, B., & Jakeman, A. J. (2022). Fit-for-purpose environmental modeling: Targeting the intersection of usability, reliability and feasibility. Environmental Modelling & Software, 148, 105278. https://doi.org/10.1016/j.envsoft.2021.105278

Howarth, C., & Anderson, A. (2019). Increasing local salience of climate change: The un-tapped impact of the media-science interface. Environmental Communication, 13(6), 713-722. https://doi.org/10.1080/17524032.2019.1611615

Ishikawa, K. (1990). Introduction to Quality Control (Translator: J.H. Loftus). Productivity Press, New York, USA.

Iwanaga, T., Steinmann, P., Sadoddin, A., Robinson, D., Snow, V., Grimm, V., & Wang, H.-H. (2022). Perspectives on confronting issues of scale in systems modeling. Socio-Environmental Systems Modelling, 4, 18156. https://doi.org/10.18174/sesmo.18156

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. (2021a). Socio-technical scales in socio-environmental modeling: Managing a system-of-systems modeling approach. Environmental Modelling & Software, 135, 104885. https://doi.org/10.1016/j.envsoft.2020.104885

Iwanaga, T., Wang, H.-H., Koralewski, T., Grant, W., Jakeman, A., & Little, J. (2021b). Toward a complete interdisciplinary treatment of scale: Reflexive lessons from socioenvironmental systems modeling. Elementa: Science of the Anthropocene, 9(1), 00182. https://doi.org/10.1525/elementa.2020.00182

Jakeman, A. J., Letcher, R. A., & Norton, J. P. (2006). Ten iterative steps in development and evaluation of environmental models. Environmental Modelling & Software, 21(5), 602-614. https://doi.org/10.1016/j.envsoft.2006.01.004

Kéfi, S., Eppinga, M. B., de Ruiter, P. C., & Rietkerk, M. (2010). Bistability and regular spatial patterns in arid ecosystems. Theoretical Ecology, 3(4), 257-269. https://doi.org/10.1007/s12080-009-0067-z

Kelly, 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

Lippe, M., Bithell, M., Gotts, N., Natalini, D., Barbrook-Johnson, P., Giupponi, C., Hallier, M., Hofstede, G. J., Le Page, C., Matthews, R. B., Schlüter, M., Smith, P., Teglio, A., & Thellmann, K. (2019). Using agent-based modelling to simulate social-ecological systems across scales. GeoInformatica, 23(2), 269-298. https://doi.org/10.1007/s10707-018-00337-8

Louie, M. A., & Carley, K. M. (2008). Balancing the criticisms: Validating multi-agent models of social systems. Simulation Modelling Practice and Theory, 16(2), 242-256. https://doi.org/10.1016/j.simpat.2007.11.011

McEvoy, N., & Whitcombe, A. (2002). Structured risk analysis. Infrastructure Security, InfraSec 2002,

Meinen, B. U., & Robinson, D. T. (2021). Agricultural erosion modelling: Evaluating USLE and WEPP field-scale erosion estimates using UAV time-series data. Environmental Modelling & Software, 137, 104962. https://doi.org/10.1016/j.envsoft.2021.104962

Moallemi, E. A., Zare, F., Hebinck, A., Szetey, K., Molina-Perez, E., Zyngier, R., Hadjikakou, M., Kwakkel, J., Haasnoot, M., & Miller, K. K. (2022). Knowledge co-production for decision-making in human-natural systems under uncertainty. https://doi.org/10.5281/zenodo.7426011

Moallemi, E. A., Zare, F., Reed, P. M., Elsawah, S., Ryan, M. J., & Bryan, B. A. (2020). Structuring and evaluating decision support processes to enhance the robustness of complex human–natural systems. Environmental Modelling & Software, 123, 104551. https://doi.org/10.1016/j.envsoft.2019.104551

Modell, S. (2009). Institutional research on performance measurement and management in the public sector accounting literature: A review and assessment. Financial Accountability & Management, 25(3), 277-303. https://doi.org/10.1111/j.1468-0408.2009.00477.x

Müller, B., Bohn, F., Dreßler, G., Groeneveld, J., Klassert, C., Martin, R., Schlüter, M., Schulze, J., Weise, H., & Schwarz, N. (2013). Describing human decisions in agent-based models – ODD + D, an extension of the ODD protocol. Environmental Modelling & Software, 48, 37-48. https://doi.org/10.1016/j.envsoft.2013.06.003

Nabavi, E., Daniell, K. A., & Najafi, H. (2017). Boundary matters: the potential of system dynamics to support sustainability? Journal of Cleaner Production, 140, 312-323. https://doi.org/10.1016/j.jclepro.2016.03.032

Ostrom, E. (2009). A general framework for analyzing sustainability of social-ecological systems. Science, 325(5939), 419-422. https://doi.org/doi:10.1126/science.1172133

Patten, B. C. (1971). A primer for ecological modeling and simulation with analog and digital computers (Vol. 1). Academic Press.

Podestá, G. P., Natenzon, C. E., Hidalgo, C., & Ruiz Toranzo, F. (2013). Interdisciplinary production of knowledge with participation of stakeholders: A case study of a collaborative project on climate variability, human decisions and agricultural ecosystems in the Argentine Pampas. Environmental Science & Policy, 26, 40-48. https://doi.org/10.1016/j.envsci.2012.07.008

Razavi, S., Jakeman, A., Saltelli, A., Prieur, C., Iooss, B., Borgonovo, E., Plischke, E., Lo Piano, S., Iwanaga, T., Becker, W., Tarantola, S., Guillaume, J. H. A., Jakeman, J., Gupta, H., Melillo, N., Rabitti, G., Chabridon, V., Duan, Q., Sun, X., Smith, S., Sheikholeslami, R., Hosseini, N., Asadzadeh, M., Puy, A., Kucherenko, S., & Maier, H. R. (2021). The future of sensitivity analysis: An essential discipline for systems modeling and policy support. Environmental Modelling & Software, 137, 104954. https://doi.org/10.1016/j.envsoft.2020.104954

Reed, M. S. (2008). Stakeholder participation for environmental management: A literature review. Biological Conservation, 141, 2417-2431. https://doi.org/10.1016/j.biocon.2008.07.014

Refsgaard, J. C., & Henriksen, H. J. (2004). Modelling guidelines—terminology and guiding principles. Advances in Water Resources, 27(1), 71-82. https://doi.org/10.1016/j.advwatres.2003.08.006

Robinson, S. (1997, 7-10 December 1997). Simulation model verification and validation: Increasing the users' confidence. Proceedings of the 1997 Winter Simulation Conference, Atlanta, GA.

Rykiel, E. J. (1996). Testing ecological models: the meaning of validation. Ecological Modelling, 90(3), 229-244. https://doi.org/10.1016/0304-3800(95)00152-2

Saltelli, A., Jakeman, A., Razavi, S., & Wu, Q. (2021). Sensitivity analysis: A discipline coming of age. Environmental Modelling & Software, 146, 105226. https://doi.org/10.1016/j.envsoft.2021.105226

Sargent, R. (2008). Verification and validation of simulation modelling. Proceedings of the 2008 Winter Simulation Conference, New York, NY.

Schils, R. L. M., van Voorn, G. A. K., Grassini, P., & van Ittersum, M. K. (2022, 2022/01/01/). Uncertainty is more than a number or colour: Involving experts in uncertainty assessments of yield gaps. Agricultural Systems, 195, 103311. https://doi.org/10.1016/j.agsy.2021.103311

Schmolke, A., Thorbek, P., DeAngelis, D. L., & Grimm, V. (2010). Ecological models supporting environmental decision making: a strategy for the future. Trends in Ecology & Evolution, 25(8), 479-486. https://doi.org/10.1016/j.tree.2010.05.001

Sitas, N., Ryan, P., & Schultz, L. (2021). Systems scoping. In R. Biggs, A. de Vos, R. Preiser, H. Clements, K. Maciejewski, & M. Schlüter (Eds.), The Routledge Handbook of Research Methods for Social-Ecological Systems (pp. 83-94). Routledge. https://doi.org/10.4324/9781003021339

Trébuil, G., Bousquet, F., Ekasingh, B., Baron, C., & Le Page, C. (2005). A multi-agent model linked to a GIS to explore the relationship between crop diversification and the risk of land degradation in northern Thailand highlands. In F. Bousquet, G. Trébuil, & B. Hardy (Eds.), Companion modeling and multi-agent systems for integrated natural resource management in Asia (pp. 167-190). International Rice Research Institute.

Turner, M., Gardner, R., O'neill, R., Gardner, R., & O'Neill, R. (2001). Landscape ecology in theory and practice (Vol. 401). Springer.

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, 142(1), 18-28. https://doi.org/10.1016/j.agee.2011.03.005

van Voorn, G. A. K., Verburg, R. W., Kunseler, E. M., Vader, J., & Janssen, P. H. M. (2016). A checklist for model credibility, salience, and legitimacy to improve information transfer in environmental policy assessments. Environmental Modelling & Software, 83, 224-236. https://doi.org/10.1016/j.envsoft.2016.06.003

Voinov, A., & Bousquet, F. (2010). Modelling with stakeholders. Environmental Modelling & Software, 25(11), 1268-1281. https://doi.org/10.1016/j.envsoft.2010.03.007

Wagener, T., Boyle, D. P., Lees, M. J., Wheater, H. S., Gupta, H. V., & Sorooshian, S. (2001). A framework for development and application of hydrological models. Hydrology and Earth System Sciences, 5(1), 13-26. https://doi.org/10.5194/hess-5-13-2001

Wang, H.-H., & Grant, W. E. (2021). Reflections of two systems ecologists on modelling coupled human and natural (socio-ecological, socio-environmental) systems. Ecological Modelling, 440, 109403. https://doi.org/10.1016/j.ecolmodel.2020.109403

Wernz, C., & Deshmukh, A. (2012). Unifying temporal and organizational scales in multiscale decision-making. European Journal of Operational Research, 223(3), 739-751. https://doi.org/10.1016/j.ejor.2012.06.038

Wu, J., & Li, H. (2006). Concepts of scale and scaling. In J. Wu, K. B. Jones, H. Li, & O. L. Loucks (Eds.), Scaling and uncertainty analysis in ecology (pp. 3-15). Springer.

Zare, F., Guillaume, J. H. A., & Jakeman, A. J. (2019). Constructing customized modelling guidelines: A participatory integrated assessment and modelling example. 23rd International Congress on Modelling and Simulation, Canberra, Australia.

Zare, F., Guillaume, J. H. A., ElSawah, S., Croke, B., Fu, B., Iwanaga, T., Merritt, W., Partington, D., Ticehurst, J., & Jakeman, A. J. (2021). A formative and self-reflective approach to monitoring and evaluation of interdisciplinary team research: An integrated water resource modelling application in Australia. Journal of Hydrology, 596, 126070. https://doi.org/10.1016/j.jhydrol.2021.126070

Zare, F., Guillaume, J. H. A., Jakeman, A. J., & Torabi, O. (2020). Reflective communication to improve problem-solving pathways: Key issues illustrated for an integrated environmental modelling case study. Environmental Modelling & Software, 126, 104645. https://doi.org/10.1016/j.envsoft.2020.104645