Iwanaga, T, et al. 2021. Toward a complete interdisciplinary treatment of scale: Reflexive lessons from socioenvironmental systems modeling. *Elem Sci Anth*, 9: 1. DOI: https://doi.org/10.1525/elementa.2020.00182

PRACTICE BRIDGE

Toward a complete interdisciplinary treatment of scale: Reflexive lessons from socioenvironmental systems modeling

Takuya Iwanaga[1,*], Hsiao-Hsuan Wang[2], Tomasz E. Koralewski[2], William E. Grant[2], Anthony J. Jakeman[1], and John C. Little[3]

The pathways taken throughout any model-based process are undoubtedly influenced by the modeling team involved and the decision choices they make. For interconnected socioenvironmental systems (SES), such teams are increasingly interdisciplinary to enable a more expansive and holistic treatment that captures the purpose, the relevant disciplines and sectors, and other contextual settings. In practice, such interdisciplinarity increases the scope of what is considered, thereby increasing choices around model complexity and their effects on uncertainty. Nonetheless, the consideration of scale issues is one critical lens through which to view and question decision choices in the modeling cycle. But separation between team members, both geographically and by discipline, can make the scales involved more arduous to conceptualize, discuss, and treat. In this article, the practices, decisions, and workflow that influence the consideration of scale in SESs modeling are explored through reflexive accounts of two case studies. Through this process and an appreciation of past literature, we draw out several lessons under the following themes: (1) the fostering of collaborative learning and reflection, (2) documenting and justifying the rationale for modeling scale choices, some of which can be equally plausible (a perfect model is not possible), (3) acknowledging that causality is defined subjectively, (4) embracing change and reflection throughout the iterative modeling cycle, and (5) regularly testing the model integration to draw out issues that would otherwise be unnoticeable.

Keywords: Reflexive analysis, Integrated assessment and modeling, System-of-Systems, Socioenvironmental modeling, Interdisciplinary teams, Uncertainty

1. Introduction

Consideration of scale is a common activity in all system-of-systems (SoS) modeling approaches involving the integration of multiple models when representing any complex socioenvironmental system (SES) of interest. Unfortunately, such consideration is all too often conducted tacitly, or at best minimally, and recently has been considered a grand challenge in SES modeling (Elsawah et al., 2020). Scale underlies many modeling concerns including how to address model complexity, conceptual mismatches, and uncertainty. In short, explicit consideration of scale issues provides a valuable, and indeed critical, lens to view the decisions made in any SES modeling activity.

This article follows an earlier publication (Iwanaga et al., 2021b) in which the current practices, issues, and challenges with respect to scale were explored through a sociotechnical lens. Scale can thus be characterized as an expansive term relating not just to the properties of the system under investigation but also the interplay between the social and technical dimensions. These influence what is considered, what is not, and what is eventually included in the modeling. A crucial aspect is the influence of the people involved and the subsequent technical processes and decisions that produce a model for a given purpose. These underlying influences, including scale decisions taken, often remain implicit and are not explicitly discussed. But for reasons of saliency, legitimacy, and transparency, they are best appreciated and considered by team members in as complete a sense as possible, albeit taking resources and time available into account.

Interdisciplinarity is now recognized as a crucial necessity in understanding and dealing with the complexity of socioenvironmental interactions (Hall et al., 2012; Saltelli

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

[2] Ecological Systems Laboratory, Department of Ecology and Conservation Biology, Texas A&M University, College Station, TX, USA

[3] Department of Civil & Environmental Engineering, Virginia Tech, Blacksburg, VA, USA

* Corresponding author:
Email: iwanaga.takuya@anu.edu.au

and Funtowicz, 2017; MacLeod and Nagatsu, 2018; Sterling et al., 2019). Challenges to a successful modeling process and set of outcomes necessitate effective collaboration, teamwork, and cross-disciplinary communication and discussion of a high order among team members (Nancarrow et al., 2013; Hall et al., 2018). Technological solutions cannot resolve mismatches in understanding among people, although they can facilitate and prompt discussions. Thus, there is a need to examine the practices and decision choices we make in any modeling activity, but especially so for one as complex as in SES. Acknowledgment of the human aspects that influence the choice and treatment of scales in modeling, and their implications, is therefore crucial in moving beyond the status quo. A key activity then is identifying the practices and approaches that facilitate and promote effective interdisciplinary cohesion among and within modeling teams.

The treatment of scale in the modeling process is an essential and valuable activity for focusing the attention of modelers on many of their key decisions. But that treatment can be affected by the level of cohesion and reflexivity within the collaborative process, which in turn may have a substantial influence on modeling outcomes, especially with greater interdisciplinarity of the modeling issue being addressed (Jones et al., 2011; Lahtinen et al., 2017). The level of inclusivity in communication that leads to interdisciplinary considerations and participation of all stakeholders, where stakeholders may include modelers themselves (following the definition in Freeman, 2010), are then also issues of scale to be explored. Recent publications espouse similar positions in recognizing the role that researchers play in shaping the scientific and policy discourse (Crouzat et al., 2018; Connolly, 2020; Walsh et al., 2020). We, as researchers, are perhaps coming to the full realization that "the technique is never neutral" (Saltelli et al., 2020) and that we cannot divorce ourselves from the influence we have on processes we take part in (Glynn et al., 2017; Cockerill et al., 2019).

1.1. The reflexive approach

A (more) reflexive approach to interdisciplinarity has been suggested over the years to aid in bridging the gap in understanding between the research that is conducted and the interdisciplinary processes that produce research outcomes (Finlay, 2002; Preston et al., 2015; Lahtinen et al., 2017). As with many terms that cross disciplinary bounds, "reflexivity" has several meanings with different practitioners holding differing views on its definition. The term "reflexive" is adopted here to convey a more transformative intent; the goal is to improve future practice through reflection on the seemingly self-evident choices and influences in the activities undertaken, the underlying assumptions, the role one played in the decisions, and the broader context in which these choices occur (Preston et al., 2015; May and Perry, 2017; Bolton and Delderfield, 2018). Reflexive evaluation is therefore one approach to considering the implication of scale choices, a practice which can aid in identifying the lessons learnt that are of benefit to future research (Krueger et al., 2016; Montana et al., 2020).

In this article, we draw five lessons through the reflexive accounts of the treatment of scale across two interdisciplinary socioenvironmental modeling case studies, also drawing upon diverse literature, where appropriate, to corroborate our experiences. As noted by others, the reflexive approach is highly situation-specific, such that there is no "one" approach to reflexivity (e.g., Montana et al., 2020). The reflexive process applied here was, however, informed by descriptions of reflexivity given in Finlay (2002) and May and Perry (2017), alongside accounts provided by Krueger et al. (2016) and Preston et al. (2015).

The described approaches involve critical self-analysis, which we define in this context as analyzing one's own influence on the modeling process, and a process of joint discussions to form reflexive accounts of our experiences. The choices made in the modeling and their implications were analyzed as part of the reflexive process to elicit the how and why of the modeling and their influence on outcomes. The adopted approach also involved a third party who acted to provide an external viewpoint to elicit further reflection and pushed forward the reflexive process. The approach aided in drawing out the successes and the struggles encountered when working within an interdisciplinary context. It is acknowledged here that the described approach is subject to some uncertainty as not all those involved in the original case studies could participate (due to availability and the necessary time commitment) and so may not include their valuable insights and perspectives (a matter revisited in Section 2).

The reflexive approach encompasses not just the "technical" decisions made (such as what models to use and the scope of stakeholder engagement) but also acknowledges that the modeling teams form a social system in its own right with their own complex interactions which influence the path taken. Model outcomes are therefore heavily influenced by the social context of the modeling process as well as the technical decisions made therein. Future efforts can be improved by concretely acknowledging this interplay (Catalano et al., 2019; Sterling et al., 2019; Montana et al., 2020). A sociotechnical view was taken to elicit these aspects in the reflexive process.

In the next section, we briefly detail the modeling conducted for the two case studies alongside the reflexive accounts of the choices made in the consideration of modeling scale, the team processes involved, and the decisions made. Both studies employed an SoS approach involving the integration of multiple models to represent the SES of interest. The fundamental need to consider these scale aspects has been previously articulated in Elsawah et al. (2020), Little et al. (2019), Badham et al. (2019), and Hamilton et al. (2015), albeit from different perspectives. We then synthesize the five main lessons learnt from the case studies, which we hope might enhance future SES modeling activities.

2. The case studies

The two case studies represent different facets of the issues that SES modelers face within an SoS context. A reflexive account for each case study is provided below and is aligned with the basic steps in the modeling

Table 1. Overview of each case study including the team context, socioenvironmental systems (SESs) involved, and purpose of the modeling. DOI: https://doi.org/10.1525/elementa.2020.00182.t1

Case Study	Team Context	SESs Involved	Time Steps	Purpose
Sugarcane aphids in Great Plains	Interdisciplinary group including experts in areawide pest management, entomology, and ecological modeling located in several states and employed by federal, state, and private institutions The core modeling team consisted of three ecological modelers, an areawide pest manager, an entomologist, and a meteorologist/aeroecologist	Four in total: agroecological systems (sorghum growth, aphid life cycle, and crop management); meteorological system (airborne aphid dispersal)	Once per model run: crop management model Daily: sorghum growth model, aphid life-cycle model Hourly: meteorological dispersal model	Forecasting sugarcane aphid infestations of sorghum fields within an areawide pest management program, providing infestation forecasts to areawide pest managers and sorghum producers
Campaspe	Large group of participants across different disciplines (>10) geographically spread across many institutions (>6). The team included modeling specialists across five systems, and one generalist who developed the farm model, aided in integrated design and development, and led the integration of models	Seven in total: agricultural, hydrological (surface and groundwater), ecological, climatic variability, policy, and recreational suitability	Daily: surface and groundwater, climate Two weekly: agriculture and policy Once per model run at end of scenario: ecology and recreational water suitability index	Knowledge integration and stakeholder discussion of the range of impacts that changing climatic and policy contexts have on water-related farm and environmental concerns

process. The subsections are not organized identically, however, owing to the different experiences encountered and the focus on providing a reflexive account. Key information is briefly summarized here with an overview provided in **Table 1**, and readers who feel sufficiently informed may skip ahead to Section 3 (Lessons learnt).

Both models are of the SoS type as they leverage constituent models that individually represent separate systems wherein each model could, potentially, be applied separately. As is typical of SoS approaches, each case study (1) considered different time frames and spatial/temporal granularities, (2) spanned multiple systems, and (3) involved multiple disciplines and stakeholders. An aspect of scale to be considered is the process of deciding which representations are to be included or excluded and how they are to be represented in terms of the scale of the modeling to be conducted. Modeling scale therefore includes all aspects of the modeling process including the conceptualization of the model, the relationships between constituent models, model structures, boundaries, parameterizations, implementation approach, and the decisions that underpin each of these. These decisions may be influenced by factors external to the modeling concerns, such as the available resources or imposed legacy software, but are also influenced by the disciplinary representation within the team, the interests represented by stakeholders, and the level of interdisciplinary cohesiveness.

The first case study, referred to as the sugarcane aphids in Great Plains (GPSCA) case study, focuses on areawide integrated pest management of aphids infesting grain sorghum fields across a large spatial area, incorporating local- and regional-scale dynamics. More expansive descriptions for each case study are available—Wang et al. (2019) and Koralewski et al. (2019, 2020a) for the GPSCA study and Iwanaga et al. (2018, 2020) for the Campaspe case study. The GPSCA case study emphasizes modeling choices from a technical point of view, while the Campaspe case study offers a description of the team processes, which influenced decisions during model development.

2.1. The GPSCA case study

The sugarcane aphid is an economic pest of sorghum worldwide (Singh et al., 2004), and outbreaks in U.S. sorghum fields have been recurring annually since 2013. Economic losses result from direct feeding, compromised harvesting efficiency, and damage to harvesting equipment and may exceed 50% of the yield (Bowling et al., 2016). Aphids are highly prolific and disperse with wind over long distances within the prairie-steppe region of the North American Great Plains, which is the principal sorghum production area in the United States (van Rensburg, 1973; Singh et al., 2004; USDA-NASS, 2010; Bowling et al., 2016).

Two key tactics within an areawide integrated pest management program for cereal aphids include deployment of aphid-resistant sorghum cultivars and selective use of insecticides (Elliott et al., 2008; Giles et al., 2008; Brewer et al., 2019). The model of Wang et al. (2019) was developed to support wise use of these management

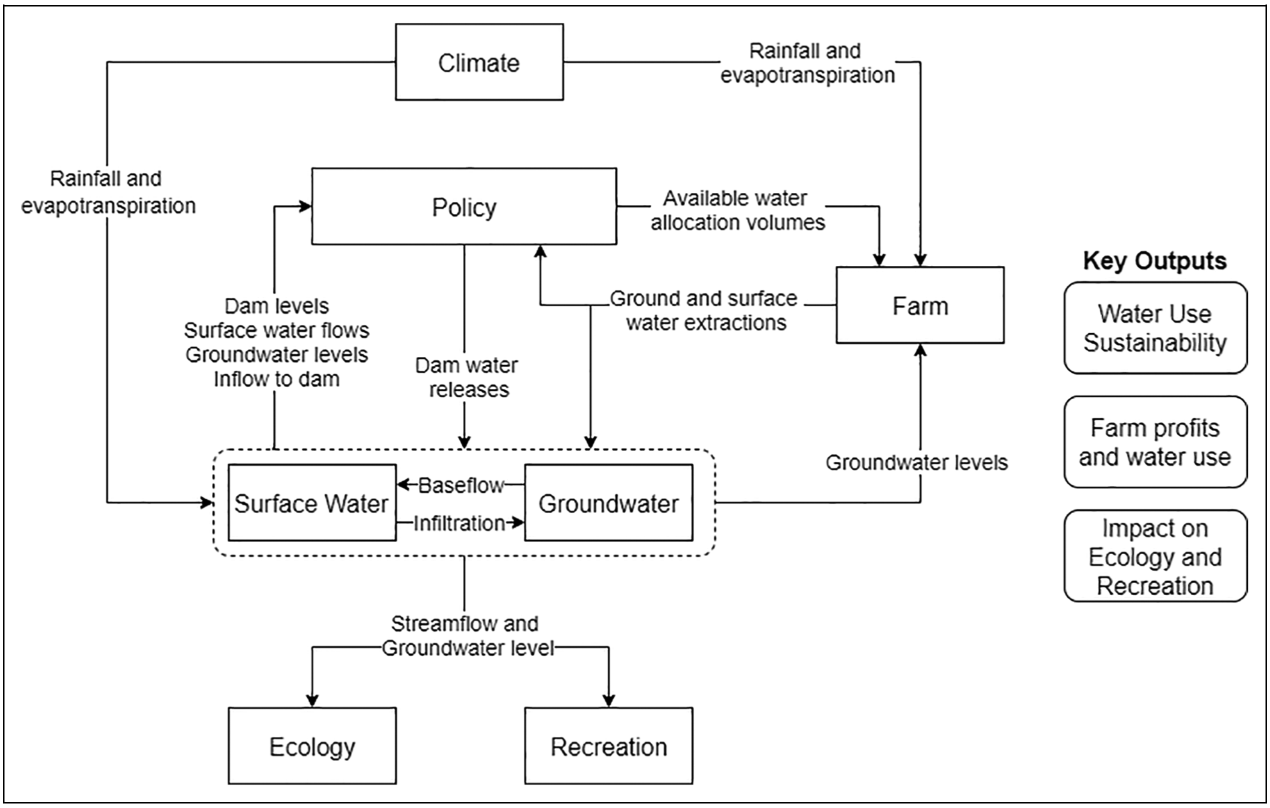

Figure 1. Relationship and interactions between constituent models and the key outputs of the Campaspe integrated model (Iwanaga et al., 2020). Each box represents a constituent model. Dashed line around surface and groundwater models is to simplify the diagram and does not signify a separate model. Arrowed lines indicate the process of data interoperation; the direction of interaction and the data communicated between models. DOI: https://doi.org/10.1525/elementa.2020.00182.f1

tactics for sugarcane aphids. The purpose of the model was to simulate areawide spatiotemporal patterns of sugarcane aphid infestations of sorghum fields, with a focus on timing of initial infestations. Near-real-time model forecasts could inform growing season activities such as timing of field monitoring to detect aphids and optimal insecticide use (Koralewski et al., 2020b). Model outputs also could be useful for region-scale management recommendations such as deployment of aphid-resistant sorghum cultivars (Koralewski et al., 2020a).

2.1.1. Conceptualization

The core modeling team that developed the conceptual model consisted of three ecological modelers: an areawide pest manager, an entomologist, and a meteorologist/aeroecologist. Although we all conceptualized the SoS as consisting of linked agroecological and meteorological systems, agreement on representation of causal processes operating within, and especially between, these two constituent systems was achieved only after considerable debate. The debate was centered explicitly on our choice of appropriate temporal and spatial scales at which to represent system processes. However, implicitly, we were debating the level of causality to include in the representation of those processes. That is, did we require our model, or parts thereof, to be interpretable as embodying cause–effect relationships or did we require only that model outputs correspond well with available real-world observations. Below, we describe the details of the conceptual model of the integrated SoS that emerged as a shared understanding of the modeling team (see also figure 1 in Wang et al., 2019).

Important processes modeled in the agroecological system included sorghum growth, aphid development, and crop management. Both sorghum growth (through phenological stages) and aphid development (through life-cycle stages) were modeled primarily as a function of environmental temperature. Crop management (i.e., decisions to plant aphid-resistant sorghum cultivars and rules for insecticide use) was modeled as a set of external variables. Important processes modeled in the meteorological system included emigration (time and location of aphid “takeoff”), wind-borne aphid migration, and immigration (time and location of aphid “landing”).

Migration was modeled primarily as a function of wind velocity and direction and flight duration. The processes of emigration and immigration linked the agroecological and meteorological systems, with emigration initiated in the agroecological model (based on sorghum phenological stage and aphid life stage) and immigration initiated in the meteorological model (based on the deposition pattern of aphids). Additional conceptual details on linkage of the meteorological and agroecological components are provided in Koralewski et al. (2019; figure 1 therein).

The spatial extent of the model included the primary sorghum-producing areas in the United States, including Kansas, Oklahoma, and Texas. The spatial resolution was set to 0.5° latitude by 0.5° longitude, which resulted in about 700 georeferenced landscape cells of approximately 2,500 km^2. This coarse-grained scale facilitated linkage of the agroecological system of the SoS to an existing atmospheric particle trajectory model (HYSPLIT; Stein et al., 2015; see Section 2.1.3.2 on model construction). The number of cells may vary annually depending on whether or not sorghum is present in the cell during a given year. The temporal scale was 1 year, which allowed for encompassing one sorghum growing season, and thus the maximum potential extent of aphid infestations, with resolution of a daily step to capture important details related to phenological development of sorghum and population dynamics of aphids. Immigration of aphids from outside of the southern boundary was approximated as occurring from the Rio Grande Valley along the border between Texas and Mexico based on reported field detection of aphids (Bowling et al., 2016). Aphids arrive in the Valley from Mexico "unannounced" because producer reports of infestations are not available from Mexico on a regular basis. We elaborate our rationale for scale choices in the subsection below.

However, before proceeding further, we reflect for a moment on our modeling team for this case study. Our team did not include a social scientist. This was not by design, in the sense of an explicit decision in favor of exclusion. Rather, it resulted from funding priorities within the overall project and the associated restriction on the number of modeling team members. Our team also did not include sorghum producers. Our entomologist maintained close ties with numerous producers via his agricultural extension activities and could explain their perspectives and main interests with relative confidence. Nonetheless, sorghum producers did not participate directly in discussions among the members of the modeling team.

Ideally, a social scientist and at least one sorghum producer would have been members of the core modeling team from the beginning. These "social voices" would have enriched our shared understanding of the SES and, arguably, could have modified the course of model development. For example, one of our livelier debates during model development, which we describe below in Section 2.1.2.2, undoubtedly would have included a more detailed discussion of the guidelines that have been developed for sorghum planting dates and subsequent management activities. We currently are exploring ways to quantify producer decision making within predominantly biophysical models (e.g., Wang et al., 2020b), which may be applicable to the GPSCA model. But perhaps more challenging than the quantitative details involved are the financial and logistical problems that impede direct long-term involvement of stakeholders in the model development process.

2.1.2. Scale choices

The multifarious reasoning that led to our choice of scales is not intuitively obvious. In principle, we based our determination of spatial and temporal scales (outlined in Table 1) on model objectives, the ecology of the organisms involved, the level of detail contained in information available from literature and from stakeholders (sorghum producers), and computational considerations. Sorghum producers did not participate directly in discussions among members of the core modeling team. However, our entomologist maintained close ties with numerous producers via his agricultural extension activities and could represent with confidence their perspectives and main interests. Spatial and temporal scales both spanned several orders of magnitude. The spatial scale of interest ranged from the regional management perspective (approximately 1.75 million km^2 of modeled area) to that of the sorghum producers' and field scientists', focused on a single sorghum leaf which, for practical purposes, encompasses an aphid colony. Temporal scales of interest ranged from an approximately 9-month period of sorghum availability in the region (for regional managers) to a "near-real-time" estimation of aphid density in a sorghum field (for sorghum producers and field scientists).

Scale choices were complicated further because aphids are small (approximately 0.05 mg) and prolific (population doubling time as short as a few days). Non-winged (apterous) morphs are relatively sedentary, whereas wind-borne dispersal can carry winged (alate) morphs over long distances (hundreds of kilometer). Thus, densities of local colonies can exceed 1,000 individuals per sorghum leaf, while emigrants from a single colony can be dispersed over thousands of square kilometers. Important life processes occurring during the terrestrial portion of the aphid life cycle are commonly measured in terms of daily rates, and the most common metric used to record field measurements of aphid densities is individuals per sorghum leaf. Sorghum development through phenological stages also is measured in terms of days (or "degree-days") per stage. However, important dynamics occurring during wind-borne aphid migration result from physical environmental conditions (wind velocities and directions) that are highly variable over the entire U.S. Great Plains.

Reflecting on these various considerations, we needed to "scale up" spatially and temporally from representation of the agroecological processes occurring at the individual aphid/sorghum leaf interface to generate seasonally variable regional patterns of aphid infestations of sorghum of interest at the areawide pest management level. Placing our model objectives within the context of Levins's (1966) classical modeling trade-offs (precision vs. generality vs. realism), it also seemed clear that our priority was realism. That is, we wanted to explicitly consider the agroecological characteristics specific to the south-central U.S. Great Plains.

In addition, we wanted to explicitly consider stochastic effects on these agroecological processes that are dependent on meteorological conditions. Infestation forecasts needed to be probabilistic. Within this context, the inherent stochasticity of the SoS and the parametric uncertainty associated with representations of system processes shaped our scale decisions. To provide some insight into our thought processes, we initially focused on the model output of most interest to end users and worked our way back to sets of modeled processes that might generate that output, noting the relative level of detail included in representation of the various processes (Wang and Grant, 2021).

2.1.2.1. Model output

The model output of most interest to end users (sorghum producers) was a set of calendar dates indicating when aphids were most likely to first infest their sorghum fields. We began by conceptually bounding the level of detail at one end with a deterministic, static, correlative model that estimates a mean date of the first infestation of the south-central U.S. Great Plains based on observed first infestation dates (which date back to 2013). At the other end, we conceptually bounded the level of detail with a dynamic, spatially explicit, individual-based model that represents all of the individual sorghum leaves in the south-central U.S. Great Plains and all of the individual aphids that might infest them.

Given the purpose, a useful model needed to be probabilistic, dynamic, and spatially explicit. Thus, regarding spatial and temporal scales, we divided the south-central U.S. Great Plains into smaller-sized areas and the approximately 9-month period into shorter time steps. Furthermore, we knew that producers were most interested in their sorghum fields and in associated management activities (e.g., planting, monitoring for aphids, pesticide applications), which might be shifted by a few days or weeks. Areawide pest managers were interested in helping individual producers make such decisions, but via more synoptic infestation forecasts, which could be individualized by local agricultural advisors (e.g., in the United States, agricultural extension agents working at the county level). Thus, for end users, the model needed to provide daily forecasts that could be interpreted at farm-level and regional-level spatial scales.

2.1.2.2. Process representation

System processes needing to be explicitly modeled included those at the sorghum/aphid/crop management interface. As we mentioned in Section 2.1.1 on model conceptualization, our main debates about scale choices were primarily debates about the level of causality to include in the representation of SoS processes. In particular, we debated whether our model, or parts thereof, needed to be interpretable as embodying cause–effect relationships. Below, to avoid an overly confusing description of process representation, we first present our final shared understanding of the appropriate scales to use. We then conclude this section with an attempt to provide some insight into the sorts of debates that led to that shared understanding.

Guidelines have been developed for sorghum planting dates and subsequent management activities in terms of latitudinal differences in weather patterns during the growing season. Population dynamics of sugarcane aphids on grain sorghum have been widely studied over the past several years, although our ability to quantify with confidence the effects of aphid-resistant sorghum cultivars, natural aphid enemies, and proximate causes of emigration remains quite limited. The fact that migrating aphids are dispersed by the wind as essentially inert particles above the flight boundary layer (i.e., a few meters above ground level) allows representation of migration via the use of well-developed meteorological particle dispersion models but also results in the uncertainty necessarily associated with weather forecasts.

Thinking about positioning our representations of these processes at the sorghum/aphid/crop management interface with regard to the level of detail included in the representations, it seemed that the modeled processes should meet two criteria. They should generate output directly comparable to personal observations commonly made by end users, and they should be viewed by research scientists as being acceptable mechanistic representations. The most common observational metric used by producers and field biologists was the number of aphids on a sorghum leaf. Usually, several leaves per plant and several plants per field were sampled on a given day, with results accumulated over time and summarized at field-, farm-, county-, and regional-level spatial scales. Regarding mechanistic (cause–effect) representation, we emphasized the term "acceptable" to acknowledge that causality is defined subjectively. The requisite level of detail to claim that a process is represented mechanistically is to a large degree problem-specific.

There was a reasonably narrow range of defensible levels of detail to consider for the model to be viewed as mechanistic. Aphid development, reproduction, mortality, and emigration, as well as processes affecting the quality of sorghum leaf (sorghum phenological development), were represented as functions of environmental temperature modified by aphid density and seemed a defensible "mark" along the level of detail scale for the agroecological model. One step toward the more detailed representation might be marked by a representation of the processes just mentioned explicitly in terms of the physiology involved in sorghum and aphid development and the frequency of physical contact among aphids. One step toward a less detailed representation might be marked by an implicit representation of these processes in terms of sorghum phenological stage and aphid population density as functions of days since planting and days since initial infestation, respectively, and emigration as a function of population density per se.

The level of detail for representation of agroecological processes that met the two criteria just described suggested a sorghum leaf and a day as appropriate spatial and temporal scales. This left us with two final considerations related to scale choice. One involved summarizing numerically the results of mechanistically modeled daily processes occurring on individual sorghum leaves in terms of a set of calendar dates indicating when aphids were most likely to first infest sorghum fields in the south-central U.S. Great Plains. The other involved dealing with potential phase shifts along the level of detail continuum that might be needed when passing information about migrating aphids between the agroecological and meteorological models.

The first step in summarizing results from individual sorghum leaves involved deciding how many leaves we needed to represent explicitly, how they might differ from one another, and how aphids on one leaf might affect aphids on another leaf. There is, however, relatively large observed variation in aphids/leaf on a single plant, aphids/plant within a single field, and aphid densities among neighboring fields, as well as spatial variation in environmental temperatures to which leaves (and the aphids on them) were exposed. We felt comfortable, therefore, letting a single sorghum leaf represents a mean-field approximation of the conditions of sorghum leaves over an area large enough to be of interest from the synoptic perspective of areawide managers.

We felt that forecasts summarized probabilistically from this synoptic perspective also would be interpretable at the farm level by producers. Since we would be executing sets of Monte Carlo simulations to make infestation forecasts, which would encompass the environmental stochasticity inherent in the modeled system, they could be interpreted in a similar manner to local weather forecasts. Producers were accustomed to inferring probable future weather conditions for their specific location based on weather forecasts for areas much larger than their sorghum fields. They also were accustomed to interpreting field-based observations of aphid infestations summarized at the county level in terms of infestation likelihoods for their fields. The final detail involved in summarizing results based on dynamics occurring on single sorghum leaves simply involved making the required unit conversions. For this, we had estimates of mean number of leaves per sorghum plant, mean number of sorghum plants per hectare, and number of hectares of sorghum within various-sized areas of the south-central U.S. Great Plains.

Regarding potential phase shifts along the level of detail continuum that were needed when passing information between agroecological and meteorological models, we identified two. One was conceptual and one was quantitative. Conceptually, aphids were treated as inert particles in the meteorological model as they are weak flyers. Within the meteorological model, particle depositions were updated hourly (during the 12-h migration time), but deposited particles (immigrating aphids) were passed back to the agroecological model as daily cohorts.

Quantitatively, aphids underwent a phase shift within the meteorological model in that we severed the numerical connection between the number of aphids emigrating and the number of aphids immigrating by placing an arbitrarily small number of (super-) aphids on each sorghum leaf receiving immigrants. Although not ideal, we felt this phase shift did not compromise the forecasting ability of the model. Given the variable size of emigration events, the lack of data on mortality rates during migration, and the dependency of successful colonization on the time lag between arrival of immigrants and arrival of natural enemies, we felt colonization could be represented appropriately as a stochastic event occurring within any landscape cell in the agroecological model (Wang et al., 2020a).

Having presented our final shared understanding of appropriate scales, we now attempt to provide some insight into one of the livelier scale debates with regard to the level of detail with which to represent SoS processes. As we described above, our final decision with regard to aphid development, reproduction, mortality, and emigration was to represent these processes as functions of environmental temperature modified by aphid density. Our meteorologist/aeroecologist would have been satisfied with a "causal" representation of aphid population dynamics that represented population density as a function of number of days since initial infestation and emigration as a function of population density. Such a representation was perceived as unacceptably phenomenological by our entomologist. Our entomologist initially proposed a more mechanistic representation of the aphid life cycle, which included, among other things, mortality due to natural enemies (predators and parasites). Arguably, aphid population growth depends on timing and magnitude of mortality imposed by their natural enemies, which depends on species composition of the community of natural enemies, which depends on the characteristics of the landscape surrounding a sorghum field. However, in view of (1) the site-specificity of such relationships, (2) the fact that connection of the terrestrial portion of the SoS model with the aeroecological portion required just a single number of aphids emigrating from each of the approximately 2,500 km^2 landscape cells, and (3) the fact that the purpose of the model was to simulate areawide spatiotemporal patterns of aphid infestations, our entomologist agreed to a simpler "causal" representation of the aphid life cycle. The simpler representation upon which we finally agreed was acknowledged as acceptably "causal" by our entomologist because of the general acceptance among subject-matter experts of the temperature dependency of insect reproduction and development and the density dependency of aphid emigration. Our meteorologist/aeroecologist doubted that model output would be improved by this, from his perspective, more complicated representation but acknowledged the benefits in terms of increasing model credibility.

2.1.3. Development

The integrated SoS model was built for use specifically within the context of the areawide pest management program for sugarcane aphids in the south-central U.S. Great Plains. It was developed by the three ecological modelers, all of whom worked at the same physical location. The modelers maintained frequent direct communication with the areawide pest manager, the entomologist, and the meteorologist/aeroecologist, each of whom facilitated indirect communication with a broad array of specific subject-matter specialists, as well as sorghum producers throughout the south-central U.S. Great Plains.

2.1.3.1. Collecting data, information, and knowledge

Several important processes included in the agroecological model had been studied extensively. Data representing growth of sorghum and development of sugarcane aphids to environmental temperature were available in the scientific literature. Information on crop management (e.g., guidelines for planting and harvesting) for sorghum in the U.S. Great Plains had been summarized and was easily accessible. Other important processes, while generally understood conceptually, could not be quantified based on available data. Proximate causes of aphid mortality and emigration remained topics of debate among subject-matter specialists. We drew upon the knowledge of the core modeling team, supplemented by the array of subject-matter specialists with whom we communicated, to quantify these processes.

Most of the important processes needed in the meteorological model had been incorporated into an existing, readily available, atmospheric particle trajectory model (see next section), which we used to simulate wind-borne aphid migration and subsequent immigration (particle deposition; aphids are weak flyers and, once airborne, are dispersed essentially as inert particles).

Specifically, the agroecological component uses data on air temperature at the soil surface and at 2 m above the soil surface, sorghum planting and harvest dates, and percentage of land on which sorghum was grown. Published information was used to model sorghum growth stages (Gerik et al., 2003), sorghum leaf area (Roozeboom and Prasad, 2019), sorghum harvest dates (USDA-NASS, 2010), aphid life stages (Davidson, 1944; Poché et al., 2016), aphid reproduction (Brewer et al., 2017; Hinson, 2017), and density-dependent reduction of aphid population size (Brewer et al., 2017). EDAS 40-km resolution data (National Oceanographic and Atmospheric Administration, 2019) were used as input for the atmospheric dispersion model HYSPLIT (Stein et al., 2015). HYSPLIT also received georeferenced information on emigrating aphids from the ecological component of the model. References for data and other sources of other information used to parameterize the agroecological and meteorological models are available in Wang et al. (2019).

Documentation to support interdisciplinary cohesion followed established standards for documenting individual-based (or agent-based) models in the field of ecological modeling, including the overview, design concepts, and details (ODD) protocol (Grimm et al., 2006, 2010).

2.1.3.2. Construction

The agroecological component of the integrated model was constructed using the individual-based modeling framework NetLogo (Wilensky, 1999). The need to model aphid life-cycle processes at an acceptably "causal" scale (see Section 2.1.2.2) prompted our choice of an individual-based model. Our choice of NetLogo over other types of modeling platforms within which individual-based models can be developed (e.g., see Ch. 8 in Grimm and Railsback, 2005) was based on our familiarity with NetLogo, its wide acceptance for individual-based modeling in ecology (Grimm et al., 2020), and its facilitation of model documentation via the ODD protocol. Our choice of NetLogo imposed computational limitations with regard to the number of individual entities that could be represented explicitly during simulations, as we describe below. The meteorological component was constructed using the established and widely used atmospheric particle trajectory model HYSPLIT (Stein et al., 2015), which computes airborne dispersal of aphids as inert particles. The NetLogo and HYSPLIT components were connected computationally with a custom-developed algorithm "Link" (Koralewski et al., 2019), with data exchange possible at a daily time step. The NetLogo platform is often used for individual-based ecological models (see, e.g., Thiele et al., 2014).

Two HYSPLIT input files EMITIMES and CONTROL are used to pass georeferenced information on emigrants from the agroecological component of the model. HYSPLIT estimates synoptic dispersal of aphids aloft. The georeferenced information on aphid immigrants is passed back to the agroecological component of the model, and subsequent updates of landscape cell states follow. Considering the spatial resolution and the regional scale, and to reduce the overall computational cost, a cohort of aphids is represented by a collective super-aphid (Scheffer et al., 1995).

An individual-based modeling approach allowed explicit representation and customization of the stage- and morph-specific reaction of sugarcane aphids to changing environmental conditions (e.g., sorghum phenological stage and environmental temperature). These reactions, or behavioral responses, of individual aphids were programmed in NetLogo via sets of equations, often embedded within logical statements. The rules were realistic, that is, they were interpretable in terms of sugarcane aphid physiology and ecology on grain sorghum in the south-central U.S. Great Plains. Population-level phenomena of interest (e.g., migration events) then emerged as the cumulative result of understandable cause–effect reactions of individuals rather than as a correlate of an arbitrary index, such as calendar date.

The conceptual basis for our choice to use an existing atmospheric model was the universal applicability of the laws of fluid mechanics upon which such models are founded. Thus, our need for a realistic integrated model, which required a "custom-built" agroecological model to accommodate the unique biological characteristics of the organisms involved, was not compromised by the generality of a model based in the physical sciences; of course, as per Levins (1966), we necessarily sacrificed precision in the sense that any realistic ecological model will contain stochastic effects, which will inevitably reduce precision (Evans, 2012). As noted earlier, aphids were treated as inert particles during the migration phase. Parameterization of the particle dispersion model required specification of the point sources (latitude and longitude) of particle emission (aphid emigration), number of particles (aphids) emitted, altitudes (meters above ground level) at which particles are dispersed (migration altitudes), and duration (hours) of dispersal events (migration duration).

Computational considerations limited the number of entities that could be dealt with numerically during simulations. We reduced the number of entities involved in calculations by simulating the phenological development of only one sorghum plant within each ≈2,500 km^2 landscape cell and the population dynamics of the aphids on only one leaf on each plant. That is, each aphid population consisted of a series of daily cohorts, with each cohort (superindividual) representing a variable number of identical aphids. The number of aphids represented by a superindividual was initialized by a reproduction or immigration event and subsequently reduced by mortality and emigration events. Each simulation, which forecasted spatiotemporal patterns of aphid infestations of sorghum during one growing season, required less than an hour of runtime on a desktop personal computer, and the necessary data input files for the meteorological model fitted comfortably within available data storage space.

Worthy of comment here is the fact that we did not face model construction problems related to concurrent development of the agroecological and meteorological models. The following case study describes communication problems, both human and computational, associated with the integration of models that were being developed concurrently (see Section 2.2.3.3). Although we needed to develop a customized algorithm ("Link") to connect NetLogo and HYSPLIT, the information passed between the two models (aphids treated as inert particles) did not change as a result of coding changes in NetLogo during the development of the agroecological model.

2.1.3.3. Model calibration

Model calibration was twofold. First, sorghum development was calibrated to adjust simulated sorghum harvest dates and number of days from planting to harvest to those reported by USDA-NASS (2010). Second, the regional migration of aphids was calibrated to adjust the simulated spatiotemporal pattern of infestations to field data from sorghum producers in Texas in 2017. This step was accomplished by adjusting colonization probabilities and did not require changes to the meteorological component of the integrated model.

2.1.4. Uncertainty analysis

The primary source of uncertainty in the integrated SoS model arose at the intersection of aphid terrestrial ecology and airborne aphid dispersal. At the time this study was published, we based this assessment on an informal sensitivity analysis that consisted of qualitative analyses of aphid infestation maps (based on expert opinion) produced by simulations with a variety of different iterations of parameters in the agroecological and aeroecological portions of the model (the maps analyzed were analogous to those in figure 8 of Wang et al., 2019). We describe the manner in which we conducted this initial, and a subsequent, sensitivity analysis in the next section on model testing and evaluation. Initiation of emigration from local populations likely depends on (1) host plant growth stage, (2) pest density and (3) developmental stage, and (4) weather or some combination thereof (Parry, 2013 and references therein). There also was uncertainty regarding duration of migration events, mortality while aloft (and thus also vigor upon landing), and aphid responses to meteorological factors in general while aloft (Eagles et al., 2013). Since processes governing initiation of emigration were modeled at the surface of a sorghum leaf, whereas processes governing airborne migration were modeled over the entire south-central U.S. Great Plains, scale issues pervaded uncertainty analysis. Furthermore, end users of the model fell into two groups with different spatiotemporal perspectives on system uncertainty.

Model purpose dictated that uncertainty analysis be focused primarily on forecasts of timing of initial aphid infestations of sorghum fields. Day-of-year of initial infestation is a common metric used by both areawide pest managers and sorghum producers to analyze and discuss infestation dynamics. However, a statement that an infestation may occur sometime during a 10-day period is likely to be interpreted quite differently by an areawide manager compared to a producer. From the spatiotemporal perspective of an areawide manager, a 10-day window of uncertainty associated with the northward advance of an aphid infestation front over the south-central U.S. Great Plains during the sorghum growing season may provide useful planning information. But from the spatiotemporal perspective of a producer, such a window of uncertainty associated with the first appearance of aphids in their sorghum field may be less useful. Likewise, a forecasted infestation front advancing via 2,500 km^2 "footsteps" may provide useful areawide management information but be less useful to a producer with a few thousand hectares of sorghum. Nonetheless, although synoptic areawide forecasts may not contain the specificity desired by producers, they do contain useful information if the forecast uncertainty is interpreted within the appropriate spatiotemporal context. Analogous to regional weather forecasts, uncertainty inevitably increases with decreasing spatial scale. SoS modelers might make more effective use of this analogy when interpreting their uncertainty analyses to end users.

2.1.5. Testing and evaluation

The initial assessment of model structure, linkages between model components, and overall model function was performed to verify overall correspondence with model purpose and to identify potentially missing components. Model behavior was then evaluated regarding the ability to produce the general south-to-north temporal trend in emergence of sorghum and the subsequent infestation of sorghum fields by sugarcane aphids.

Simulated and observed spatiotemporal patterns of aphid infestations were then compared to validate the model. The simulated data were based on 10 replicate stochastic simulations. The field data were collected in Texas, Oklahoma, and Kansas during 2017 and were not used in model development. The average simulated dates of first aphid infestations were within the range of observed dates of first infestations in four of the five sorghum growing regions (Wang et al., 2019; figure 5). The ranges of observed dates were narrower than the corresponding simulated ones, which was attributed to the fact that all simulated infestations were detected whereas field data were limited by temporal and spatial field sampling constraints. Initial testing and evaluation details are available in Wang et al. (2019).

After publication of the work reported above, in which model testing was limited by the ever-present combination of limited funding and impending deadlines, we were fortunate to have the opportunity to extend our testing in two areas of particular interest. The testing was basically a sensitivity analysis that consisted of varying the value of one parameter at a time in either the agroecological model or the aeroecological model and qualitatively assessing the effects on SoS model outputs. Both involved aphid migration, the key process (which includes the processes of immigration and emigration) connecting aphid terrestrial ecology and airborne aphid dispersal. We were interested particularly in evaluating more formally the uncertainty in model outputs describing spatiotemporal infestation trends resulting from uncertainty in the values of key parameters affecting migration. First, we evaluated the effects of altering timing of first appearance of aphids in the southernmost U.S. Great Plains. The first appearance of aphids is an initial condition of the agroecological model representing immigration from an external source (Mexico). Next, we evaluated effects of altering dispersal duration, minimum dispersal height (meters above ground level), and maximum dispersal height. Dispersal duration and heights are parameters controlling airborne dispersal in the aeroecological model. Results of these new tests indicated alteration of the timing of first appearance of aphids in the southernmost U.S. Great Plains affected forecasted spatiotemporal patterns of infestation (as indicated by georeferenced probabilities of first infestations) throughout the entire south-central Great Plains region (Koralewski et al., 2020a, 2020b). However, alteration of the three dispersal parameters, over the 63 combinations of values tested, had little effect on georeferenced probabilities of first infestations.

These new results more clearly identified the timing of first aphid infestations in landscape cells as the primary source of forecasting uncertainty in the integrated SoS model. They also suggested some rescaling of modeled processes that would be interesting to examine from the standpoint of increasing utility of infestation forecasts for sorghum producers, specifically, reducing the level of detail with which we represent processes in the agroecological model and increasing the spatial resolution with which we represent migration in the aeroecological model. We have conducted a series of thought experiments, which suggests accurate forecasting of timing of initial infestations is more important than accurate forecasting of magnitudes of migrations and initial infestations within the context of areawide pest management (Wang et al., 2020a). Given the high fecundity and rapid development of aphids at temperatures characteristic of the sorghum growing season, time lags between initial infestation, and the presence of potential emigrants is only a few days. Aphid colony growth versus local extinction depends on interaction of myriad processes (see Section 2.1.2.2) that can be aggregated into a single stochastic variable without increasing the level of uncertainty associated with colony survival and production of emigrants. However, increasing the spatial resolution of simulated immigration points poses a technical problem. Although there is increasing availability of high-resolution atmospheric data and increasing sophistication of atmospheric particle trajectory models, it is unlikely that data supporting field validation of fine-scale immigration forecasts will be available in the foreseeable future.

2.2. The Campaspe case study

The Campaspe study focused on the long-term management of water resources between agroeconomic and environmental concerns at a regional scale, under a backdrop of uncertain future climate and policy conditions. The study area, the Lower Campaspe subcatchment, is in South-East Australia and part of the Southern Murray–Darling Basin. The area is of ecological, socioeconomic, and agricultural importance. Increasing agricultural and environmental concerns and the impact of recent droughts (e.g., the Millennium Drought, 1996–2010; Kendall, 2013) have spurred a series of hotly contested water policy reforms. Regionally, riverine health is said to be poor (Murray–Darling Basin Authority, 2012; North Central CMA, 2014) and is set to become increasingly challenging, especially under uncertain climate conditions (Dey et al., 2019) that are likely to exacerbate water availability. The Campaspe integrated model (CIM) was developed to facilitate discussion among stakeholders of the long-term implications of water management decisions and potential policy changes, including conjunctive use of surface and groundwater, under a range of uncertain futures.

The interplay between the scale decisions made by the team and the implications regarding modeling scale and treatment thereof is explored here. Some context on the team and the model development approach is first provided (in Sections 2.2.1 and 2.2.2), followed by an exploration of the scale issues, and the decisions in their treatment in Section 2.2.3. The team aspects and decision choices from a scale perspective are the focus of the exploration.

2.2.1. The team context

The team consisted of specialists and research students across the fields of ground and surface water hydrology, the social sciences, software engineering, economics, systems analysis, and uncertainty assessment. Local specialists in water management, agricultural and ecological matters were engaged as part of the project. Organizationally, the team spanned six Australian institutions. Subgroups within the team each focused on an aspect of the SES. The bulk of the team had prior working relationships conducting integrated assessments, but this was the first time their models were so intimately integrated and in a manner that accounted for feedbacks between systems.

The team previously underwent a self-reflection process using a survey-based approach (discussed in Zare et al., 2021). The "Monitoring and Evaluation" process described therein aided in identifying opportunities for improvement of practices that could better structure the modeling processes and enhance team efficiency. The account provided here differs from the first in that the focus here is on the issues of scale that arise throughout rather than demonstrating the value of self-reflection in the modeling process. Common experiences then inform the lessons learnt (discussed in Section 3).

2.2.2. Development and application

The CIM was developed to represent the spatiotemporal forcing and system interactions that changing climatic, market, and policy contexts have on water-related farm decisions and profits, as well as catchment-scale groundwater and ecological concerns. Team members self-organized to develop constituent models for this SoS model and, at least initially, focused on the processes and issues of concern specific to their system of interest. The approach, and the number of people involved, then had interrelated implications regarding the treatment of scale issues and the decisions therein (which are explored in Section 2.2.3). Here, the approach to construction and simulation of the model is described to provide some context.

2.2.2.1. Construction

To address the spatiotemporal forcing and system interactions that changing climatic and policy contexts have on water-related farm and environmental concerns, an integrated model built from a collection of system-specific models was developed. Having experience in integrated assessment, modelers were aware that models would be dependent on data interoperated between models. A practical approach was taken in integrating these models, and so the CIM operates on a linear feed-forward concept where outputs from one model are fed into other models with which it has a direct relationship (see **Figure 1**). Interoperation of data occur at a daily time step for surface and groundwater models and a two-weekly step for policy and agricultural models. Feedback between models occurs once at their respective time steps, except for the two indicator models (i.e., ecology and recreation impact evaluation) that are run at the end of a scenario. Further detail on the models is provided in Appendix A.

It was known and expected early in the modeling process that the constituent models were to be developed in a variety of approaches and programming languages. Different development environments (e.g., laptop vs. supercomputer) would have to be accommodated. Technical integration of the constituent models was achieved through a purpose-built (software) framework developed in Python. The primary reason for Python is that it is cross-platform and is popular within the sciences as a "glue" between models (Muller et al., 2015; Dysarz, 2018).

2.2.2.2. Simulation approach

Exploratory scenario modeling (ESM) was the selected approach in simulating outputs with the CIM as it allows for the consideration of a multitude of plausible futures in conjunction with scenario, model, and decision uncertainty (Maier et al., 2016; Horne et al., 2019). Certainly, the involvement of researchers with a history and expertise in uncertainty assessment brought considerations of uncertainty to the forefront. Another key reason for the adoption of ESM was to better enable the communication of the scale of uncertainty to local stakeholders, which may influence the decisions enacted (Maier et al., 2016; Little et al., 2019).

Exploratory approaches involve many model runs, with each run representing a possible plausible future (i.e., a scenario) under a variety of conditions. With the CIM, these include hypothetical policy changes (e.g., conjunctive use of surface and groundwater resources), changing climate conditions, market prices for commodities and input costs, and on-farm management options to allow assessment of impacts on the agricultural, groundwater, recreational, and ecological systems.

2.2.3. Scaling issues

The scales to be represented in the CIM were identified through analyzing the needs and purpose of the individual systems of interest as well as the intersystem relationships that needed to be represented. These included the agricultural, hydrological (surface and groundwater), ecological, climatic variability, policy, and recreational systems. Specific aspects of these to be represented by models were informed by the range of local stakeholder interests and concerns. Interactions between the seven systems then enhance or degrade the ability to meet the needs of all water users over time. From these, the spatial and temporal scales (including extent and granularity) that were amenable to the context and purpose of the model were identified.

Nominally, each model was informed by both the natural and anthropogenic properties of the catchment. These included water management zones (i.e., areas subject to differing policies), aquifer boundaries, the hydrologic subbasins in the study area, and the available data. The spatial area represented by the surface water, groundwater, and farm models is depicted in **Figure 2**, and a summary of the spatial and temporal scales internal to each model is provided in Appendix A. Further details on the modeling context and findings are available in Iwanaga et al. (2018, 2020).

Because of the number of models and disciplinary experts involved and some geographic dispersion between the team members, maintaining a high degree of cohesion throughout the modeling process was challenging. In the subsections below, a reflexive account is given of the considerations of the SoS approach on the level of detail, participation, interdisciplinarity and team cohesion, and subsequent implications encountered in practice.

2.2.3.1. Scale of detail

Identifying and representing the systems of interest at a level of detail commensurate with the modeling purpose is one challenging aspect that leads to multiple, equally plausible system representations. The farming system, for

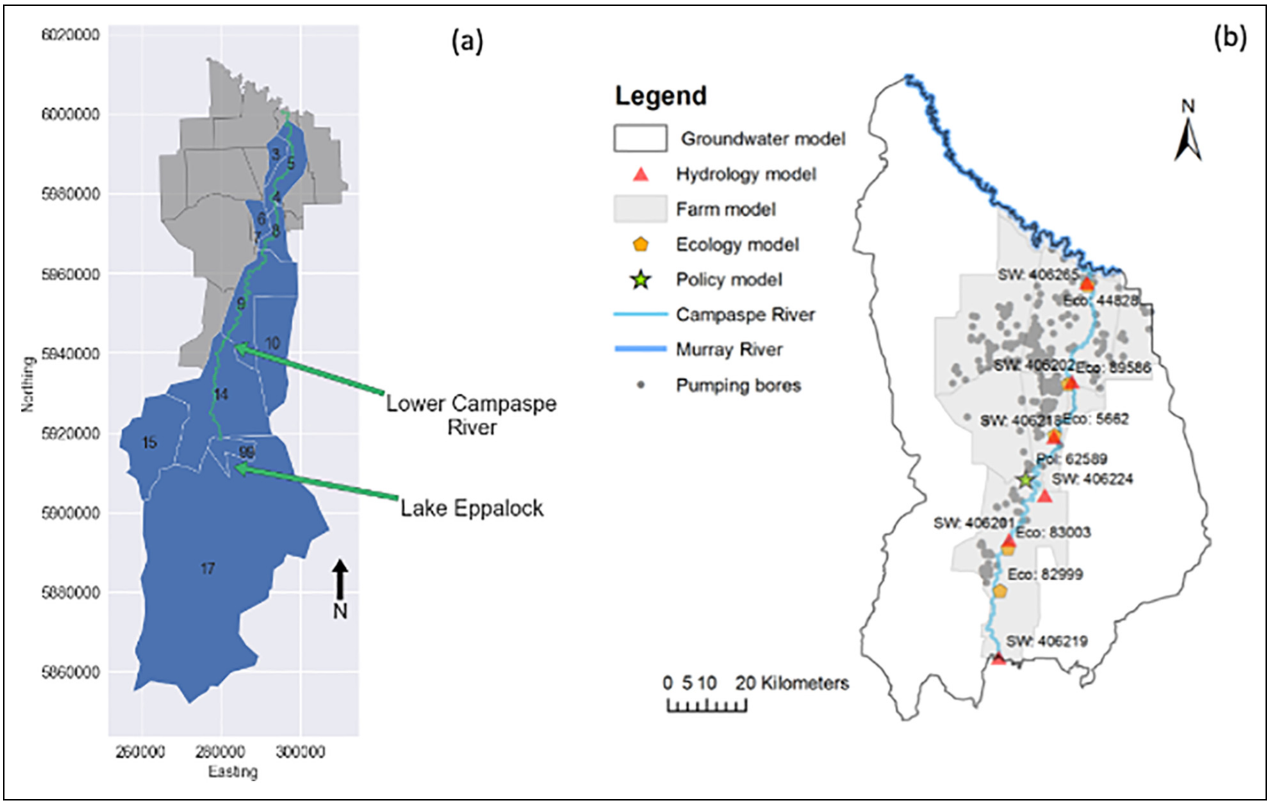

Figure 2. Surface water area (Panel a, left-hand side with subcatchment identifiers annotated) and groundwater area (Panel b, right-hand side) with farm management zones (semitransparent gray areas in both panels). Surface water area extends further south compared to the other models, whereas the represented groundwater area extends further east and west. Figure adapted from Iwanaga et al. (2018). DOI: https://doi.org/10.1525/elementa.2020.00182.f2

example, was represented as a collection of 12 spatially lumped zones primarily determined by local planning areas (known as the Goulburn–Murray Water Supply Protection Areas). Its model additionally operates on a 2-week time step to match the typical irrigation time frame considered by farmers. In other words, anthropogenic considerations (governance boundaries and water use behavior) influenced the representation more so than biophysical concerns (e.g., soil attributes).

As the quantities of interest were to be predicted primarily at the catchment level, it may have been possible to aggregate some representations to a coarser level without compromising the modeled outcomes. Toward one extreme, the catchment could be represented as a single spatial zone in the farm model rather than the adopted 12 zones. On the other end, the ecological indicator model provides a long-term indication of the average suitability of streamflow for ecological purposes (e.g., averaged value over decades). Expanding scale considerations to holistically capture the temporal dynamics, its influence on the constituent systems, and how these may adapt and evolve (e.g., adaptive management of stochastic environmental flow) may influence modeled outcomes (Horne et al., 2019; John et al., 2020). Further research is necessary to determine whether increased or decreased detail is in fact appropriate for the context in which the systems are represented. A move toward a finer level of detail than that chosen, however, would require more data at the farm and field level (e.g., long-term groundwater pumping and irrigation usage) that were not available.

In an ideal setting with more time and resources, one would undertake some analysis of the possible alternative scale assumptions to explore their effects on model outputs. In this way, one could decide on the trade-offs among different scale choices regarding improved model performance versus resources required to implement them. At the very least, for transparency, the scale choices would be documented, and the ones selected for the modeling justified with a narrative that captures the decision context, the decision, and the known implications and consequences of those decisions. There are many "good practices" for documentation in both software and model development. Software practices include the Architectural Decision Records (Emery and Hilliard, 2008; Zdun et al., 2014), which advocate storing such documentation alongside code in version control. Likewise, the TRACE documentation framework suggests keeping "computational notebooks" in version control as a complement to traditional "pen-and-paper" notebooks with similar aims of documenting decisions made throughout the modeling (Ayllón et al., 2021).

Processes and phenomena couched in ambiguous or disciplinary-specific (or context-specific) terms may drive misconceptualizations of the constituent models. For example, surface and farm models both applied separate representations of "effective rainfall." Although the surface water model provides a physically based estimation of effective rainfall at a subcatchment level (see Croke and Jakeman, 2004; Ivkovic et al., 2014), the farm model applies a soil moisture accounting method that is recommended to farmers in the region for each of the 12 farming zones represented (Iwanaga et al., 2020). The moisture accounting approach informs irrigation schedules, helping farmers determine the timing and volume of irrigation, but is not a physically based estimation, to the surprise of some. Considerations around the scale of interdisciplinary communication are explored in Section 2.2.3.3.

Often in SoS modeling, the appropriate level of detail is not readily apparent, particularly during the earlier modeling phases when model development tends to focus on higher level considerations. Choices of scale are often framed by one's disciplinary focus, and individual preferences may result in decisions that lead down unintended pathways (Lahtinen et al., 2017). Modeled scales, and their most appropriate level of representation, are often not readily apparent and could be construed to be somewhat arbitrary, but not senseless, for example, when being constrained by "real-world" considerations. Insufficient consideration of the interdisciplinary aspect and challenges in cross-disciplinary communication may then have implications in the testing, evaluation, and application of the model (i.e., different paths are taken, as in Lahtinen et al., 2017), particularly in the (disaggregated) model development phases (revisited in Sections 2.2.3.4, 2.2.3.5, and 2.2.3.6).

2.2.3.2. Scale of participation

A catchment-wide survey of farmers, a series of workshops with local experts, and targeted engagement with ecologists and those representing recreational interests were among the participatory processes used to collect expert knowledge and perspectives. Furthermore, scenarios of interest were identified and co-developed through stakeholder engagement. In effect, system experts and stakeholders act as representatives of the systems under consideration including the issues and concerns that are most pertinent with respect to the modeling. The participatory process aided in constraining the overarching scenarios to those that were deemed both technically plausible and socially acceptable regarding agricultural water use (Ticehurst and Curtis, 2016, 2017).

Aside from the usual budgetary considerations (of time, money, and personnel), timing was a crucial factor in terms of the social (local stakeholder) engagement process. Not all system experts and stakeholders could be expected to attend face-to-face meetings due to timing and scheduling conflicts and the limited resourcing available. Those involved ultimately had the available time, inclination, and goodwill to participate in the time frame selected and required by researchers. This is also true in the context of writing this reflexive account as not all involved in the original case study could contribute (as noted in the Introduction).

A strong focus on the agricultural system and related water management (albeit underpinned by surface and groundwater modeling) is therefore evident in the model conceptualization as most stakeholders were linked to the agricultural and water sector. A consequence is that the model does not consider certain sociocultural values such as those held by local indigenous peoples. Potential adaptive management processes wherein water use policies change in response to improving or deteriorating ecological flow suitability were also not considered to be in scope (see description of the ecology model in Section 2.2.3.1).

Although not an active or conscious decision, the consequential filtering of participants in this manner may have introduced a self-selection bias in the sample of local stakeholders that took part in discussions. Commensurate with the specified scope—one of investigating and discussing water management and policy changes under uncertainty—future work building on this case study will likely feature a greater emphasis on the social dynamics. Incorporating reflexivity as part of the modeling can aid in managing the scale of participation and recognizing when/where the bounds may not suit objectives. In the grander scheme of things, however, enabling such work requires that commensurate funding be available to enable greater levels of participation (Iwanaga et al., 2021b) and to capture lessons learnt through reflexivity (Montana et al., 2020).

2.2.3.3. Scale of interdisciplinarity and communication

Interdisciplinary work at the heart of SoS modeling comes with unique challenges not found in single-system contexts. Many of these are detailed by Iwanaga et al. (2021b), but key to the discussion here is that in SoS, there are several sectors and disciplines involved with associated systems and models being concurrently developed and ultimately integrated. Changes to one model, because of new information or simply because of continual improvements, may necessitate changes to another model. A continual challenge throughout the project lifecycle was effectively scaling communication and participation to an appropriate level to facilitate a deeper understanding of the SES being modeled. Modelers self-organized into subteams to accomplish goals but were, for the most part, focused on their sectoral concerns. Separate and mismatched conceptualizations of the modeling arose throughout the modeling cycle, in part due to this partitioning.

Members of the team can take the role of a mediator, resolving or otherwise addressing inconsistencies and mismatches. Methodological conflict can be addressed at the technical level via model interfaces, which translate one conceptualization to another. In the CIM, for example, lumped 2-weekly farm water extractions were translated into daily averages for the ground and surface water models. Mediators may also handle task-related and interpersonal conflict (De Dreu, 2008) but may only be effective in cases where the role is assumed by someone with sufficient standing within the team and/or a cooperative team culture exists (De Dreu, 2008; Gren and Lenberg, 2018; Hidalgo, 2019).

Certainly, those managing self-organizing teams can guide interdisciplinary communication by holding regular meetings or team bonding activities (as suggested in Zare et al., 2021). Prior research suggests goal interdependence—where the success of one is contingent on the success of another—can improve team performance by setting the stage for effective collaboration (Knight et al., 2001; Tjosvold and Yu, 2004; Lee et al., 2015), particularly where flexibility and rapid response to complex and emergent issues are important (cf. Hansen et al., 2020). Effectiveness of such management strategies is likely to be highly dependent on team context, however. Depending on the larger cultural context, it may be preferable to allow (or guide) team cultures to evolve organically without direct intervention on the frequency and scale of team interactions (e.g., by mandate from management).

In the case of the CIM, each system of interest had different—but at times overlapping—concerns and issues (with some examples provided in Section 2.2.3.1). Close coordination between collaborators was needed to avoid conceptual mismatches in the models and their coupling, given the variety of scales involved and the separate, but interdependent, development paths for each model. Maintaining a high frequency of face-to-face meetings between team members was problematic because of the geographic spread of participants and financial constraints limiting travel, with the default mechanisms being emails and phone calls between individuals and within subgroups. In retrospect, more regular virtual meetings with the whole team may have helped in the longer run, particularly around technical scaling issues.

It is now seen by the team that the use of technologies and practices available to ease the burden of maintaining communication and documentation of decisions would be valuable (Zare et al., 2021). Certainly, there was a preference toward established, often disciplinary-specific, workflows rather than approaches that are perhaps more suitable for the interdisciplinary SoS context wherein team members are also geographically dispersed. For example, most modelers involved in writing code did not actively use version control, making difficult the review and dissemination of code, changing code, and documenting the reasons behind those changes. Code was instead often shared via email. Given the evolving needs and requirements of both the modeling and interdisciplinary context, it is expected that new skills and approaches should be progressively tried and, where found applicable, incorporated into the modeling workflow (Knapen et al., 2013; Hidalgo, 2019).

Separate and mismatched conceptualizations and expectations (forewarned in Knapen et al., 2013; Kragt et al., 2013; Verweij et al., 2010) of model components arose through insufficient communication. The issues that consequently arose were challenging, not to mention time consuming, to identify and correct. To give examples, in one case, numerical values were hardcoded into a model with the expectation that they would be changed manually for every run; an approach that is inappropriate given that the exploratory approach requires hundreds to thousands of model runs. In another case, input fed in from another model was found not to affect any calculations, as the integrated context was not considered.

It is worth noting that commonly suggested solutions to the above, such as adopting "advanced" communication platforms or increasing the frequency of communication, are tools and strategies that can help maintain existing interdisciplinary foundations (to paraphrase Heffernan, 2011). Care should be taken as use of such communication technologies should not be conflated with, nor a replacement for, interdisciplinarity itself. Recent research suggests continual monitoring, regulation, and a collaborative team culture are ideal, lest discrepancies affect overall team efficacy and performance (Driskell et al., 2020). Supporting lines of evidence show that a level of empathy and receptiveness to the experiences and knowledge outside of one's own ("social intelligence" in Woolley and Malone, 2011) is also needed to effectively leverage the diverse abilities found within interdisciplinary teams (Thomas, 2012; Thomas and McDonagh, 2013). This suggests that it is the culture of empathetic openmindedness, inclusivity, and a motivation to achieve team goals that likely drives communication and the cross-pollination of interdisciplinary ideas, more so than the method and scale of communication.

2.2.3.4. Computational scalability

The computational approach is a pertinent scale consideration, especially when uncertainty primarily involves running many scenarios. In this respect, the computational scalability of the CIM became a concern to manage, mainly due to the combined runtime of the constituent models and overhead associated with their interactions. A major decision taken was to run the SoS model on a 5-km square grid rather than the initially chosen 1-km grid. Even then, a single run of the CIM could take 30 min or more, with initial implementations prior to optimizations exceeding an hour. Runtime was not an obvious issue during the disaggregated development of constituent models, even when partially integrated, especially early in the development process when the full scale and number of interactions was neither apparent nor known.

One technical barrier to increased computational performance was the use of files as an intermediary format to interoperate between models. This decision was somewhat imposed rather than selected due to the use of legacy models. Using the MODFLOW implementation for the groundwater model component as an example, computer memory (i.e., RAM), was far more limited and expensive at the time of MODFLOW's development in the 1970s (McDonald and Harbaugh, 2003). Consequently, intermediate results and parameter values between time steps (for the purpose of the CIM) were required to be written out to files rather than kept in memory. Although this process was automated through the FloPy package (Bakker et al., 2016), the comparatively high cost of file read/write activity was unavoidable and constrained the possible avenues for optimizing runtime performance. The issue was sidestepped by using a high-performance (at least at the time of writing) workstation with 32 cores, running thousands of simulations over a period of days to obtain results. This, however, is not ideal and may not be a viable solution for many.

Use of Python itself became an issue as the number, and complexity, of the models that were coupled increased. Python cannot achieve the same level of computational performance as lower-level languages (e.g., Julia, C, Fortran). The same is true for any high-level dynamic and interpreted programming language. Under usual circumstances, this is not a big issue as Python is used to leverage libraries and methods written in lower-level languages (see, e.g., NumPy; Harris et al., 2020), or otherwise "slow" parts of a Python program can be abstracted away into a lower-level language (usually Cython or C). Both strategies were taken with the farm model to improve computational performance. In the case of the CIM, Python handled the interoperation of data between models and so computational performance could not be improved without significant overhaul of the design and structure of the interfacing code, which was not possible in the available time.

As noted earlier (in Section 2.2.2.1), Python was selected for its common use as a "glue" language in the expectation that a variety of languages and approaches would be adopted by the team. It is also well-regarded as a platform for rapid prototyping. In future, an alternative language that is as flexible as Python but is more efficient computationally could well be sought as a replacement. High-performance integration necessitates a high-performance language. The Julia language (Bezanson et al., 2017), a recent addition to the scientific programming landscape, is one promising avenue in this regard.

2.2.3.5. Testing and evaluation

One salient issue that arose in the development of the CIM was the difficulty in assessing the behavior and performance of the integrated model. Calibration of models all together throughout their development was not possible as each model component was at a separate stage in the model cycle. It is acknowledged that models that are calibrated separately may exhibit unexpected behavior when integrated. Model behavior, both in the integrated and *dis*integrated context, was therefore evaluated against available observations and through stakeholder engagement.

Additional concerns revolved around uncertainties that will propagate and compound. Conceptual (or hypothesis) testing was one approach applied to address such concerns. This testing approach involved the identification of questions with a known range of acceptable answers and the subsequent testing of these against the model. The conceptual testing approach is adjustable to the available data and is especially useful in data poor contexts. Framing the context surrounding expected model behavior provides a high-level check of conditions, which can indicate the model is not fit for purpose and that changes are required. The greater the comprehensiveness of such tests, the higher the confidence that the integrated model is fit for purpose (Davidson-Pilon, 2016).

One form of conceptual testing applied was property-based sensitivity analysis. The property-based approach attempts to falsify the conceptual integrity of the integrated model by the sensitivity of model parameters within a restricted area of parameter space (Iwanaga et al., 2021a). Unexpected sensitivity results (e.g., too high, too low, or no sensitivity) then indicate an issue with the model implementation or integration, such as the inadvertent absence of model coupling. Failure of a model to conform to expected/known behavior can then falsify the assumption that the model is functioning correctly or alert to a change of context that invalidates previous understanding of the model (Claessen and Hughes, 2000). Failure of a test then avoids the computational expense of conducting a larger scale global analyses, which, due to the presence of errors, would return misleading and unreliable results.

2.2.3.6. Complexity and model uncertainty

A central challenge in the development of the CIM was determining an appropriate level of complexity while also considering its influence on (model) uncertainty. Complexity of the CIM arose from the variety of workflows, terminologies, expected spatial/temporal scales, and requirements both to individual constituent models and those pertaining to the SoS model and context. Modeler experience (and thus preferences) and available data informed several considerations throughout the modeling process.

As an example, the ground and surface water models were implemented through modifications of existing models, a decision based on prior modeler experience. These were MODFLOW-NWT with FloPy (Bakker et al., 2016) and IHACRES_GW (Ivkovic et al., 2014), respectively. Modified implementations of the groundwater model were additionally applied for other studies that were occurring concurrently (e.g., Partington et al., 2020). Available climate data were at a 5-km grid resolution, which was then the minimum granularity possible, without using interpolation, for the operation of the groundwater model. These models provided inputs for the policy, farm, and ecological indicator model with data upscaled or downscaled as appropriate for their respective purposes (see Appendix A).

The number of models involved and their structure, parameters, resolution/granularity, and data (and sources of data) were all sources of complexity. Increased complexity through the inclusion of additional systems, their interactions, and computational infrastructure generally results in compounding uncertainty (Dunford et al., 2015). This is the uncertainty that arises from the interactions between constituent models with the possibility of each interaction introducing, and propagating, some error (Refsgaard et al., 2007; Dunford et al., 2015). The error propagated may differ depending on what computational platform is in use (Iwanaga et al., 2020).

Additional model complexity allows for further investigation into the possible sources of uncertainty to be considered. Reduction of model complexity and uncertainty is often conflated with reducing its parameterization (or dimensionality), which facilitates the apportioning of parameter uncertainty to a smaller number of (considered) uncertainty sources. Reducing complexity via constraining the number of parameters does not, however, reduce uncertainty in the sense that the effect of random influences or incomplete knowledge is reduced (aleatory or epistemic uncertainty, respectively, as defined in Beven, 2009). On the other hand, model parameterization can be reduced where sensitivity and/or other analysis show that quantities of predictive interest are not influenced by certain choices. These sources of uncertainty can be explicitly documented following processes and considerations as described in Refsgaard et al. (2006, 2007), van der Sluijs (2007), and Reichert (2020).

The decision to adopt established disciplinary-specific models (e.g., MODFLOW-NWT) did quicken model development compared to starting from scratch but introduced additional complexity and considerations. For one, the MODFLOW-based model was to serve multiple purposes (across multiple studies), and so infrastructure to support the generic application and data processing was developed. Use of MODFLOW in this context is one example of a constituent model that is amenable to the overarching modeling purpose, but not necessarily complementary to it. Other constituent models of the CIM required indications of average depth to groundwater for both general and specific locations, whereas MODFLOW operates on a grid-cell (or mesh). Given MODFLOW's computational expense and additional complexity involved, it may have been worthwhile to develop a bespoke model specific to the Campaspe context of lesser complexity. Both approaches are arguably acceptable.

The question then is what level of complexity is warranted for the purpose and context of the model, recognizing constraints due to resources and legacy issues. In the context of the CIM, different scenarios to be explored required different model structures and formulations. Constituent models that could generically represent system behavior across the range of scenarios were considered a necessity. This contrasts with the development of several models specialized for each scenario context, for example, separate models for wet climate conditions, enactment of conjunctive water use policies, and so on. Considerations external to the SoS modeling exercise, as well as prior modeler experience, were additional factors that influenced the choice of constituent models, their implementation, and the process of modeling. Choice of preexisting models arguably allowed models to be developed more quickly, but at the cost of adding model complexity.

A point of interest here is that such considerations regarding the model complexity and uncertainty and their effect on quantities of interest cannot be known in advance, at least not without significant experience with the specific set of constituent models that make up the SoS model. In the context of model development, changes to constituent models invariably happen, which may sufficiently change the context of their application. Prematurely attempting to reduce model complexity and uncertainty before the full context is known (e.g., prior to model integration) is therefore inadvisable (as alluded to in Section 2.2.3.3).

3. Lessons learnt

We conclude our reflexive exercise on two SoS case studies with a synthesis of lessons across five fundamental themes elicited through reflexive self-analysis and discussions between and across the teams involved and supported by corroborating experiences drawn from existing literature. We, at least, would take these lessons forward and incorporate into future SES modeling activities. Although these lessons are also somewhat applicable to single system modeling, we believe they become especially important in the interdisciplinary SoS modeling context. It is acknowledged again here that although efforts toward discussions with team members were made, not all were able to contribute to the reflexive accounts presented. Certainly, availability and the necessary time commitment placed a limit on the scale of participation (as in Section 2.2.3.2).

3.1. Foster constant collaborative learning and reflection

The two case studies detailed in this article both featured a wide variety of disciplinary experts working together. One challenge is a risk that interdisciplinarity can be eroded as researchers gravitate toward the systems that they are familiar with. In the GPSCA case study, even though team members may have initially viewed the problem at hand through different disciplinary lenses, team members shared certain fundamental concepts. In the Campaspe case study, some conceptual mismatches arose that led to problematic issues in the integration of models, lengthening the development/modeling cycle. In our experience, the most efficient way to move an interdisciplinary conversation forward is to look backward in search of those shared concepts (Banerjee et al., 2019). Once we have found common ground, we can move the conversation forward along diverse paths under the guidance of experts who then can explain where they are leading us and why via timely additions of new concepts to our common knowledge base. Common roots were found in the concepts of system dynamics (e.g., Forrester, 1961, is a seminal work in industrial dynamics and is well-known to systems ecologists) and general systems theory. An open attitude and commitment to continual learning, both individually and as a group, is necessary for these guided paths between disciplinary domains to appear (empathetic horizons in Thomas and McDonagh, 2013) and break down disciplinary barriers (MacLeod and Nagatsu, 2018). At the very least, shared concepts avoid potential mismatches in modeler understanding.

Communication among interdisciplinary team members is crucial toward the development of a cohesive systems representation, and its importance cannot be understated. One strategy is to adopt documentation practices to ensure the existence of a collective, and cohesive, body of knowledge (Cockburn and Highsmith, 2001; Kragt et al., 2013). Specific to scale choices, the level of shared understanding and other major considerations could be explicitly catalogued in a "core" table. This table would detail the spatial and temporal scales (Koo et al., 2020), knowledge sources (Kragt et al., 2013), expected computational requirements, major uncertainty sources (Refsgaard et al., 2007; van der Sluijs, 2007; Reichert, 2020), the relevant system(s) affected, and the modeling process (Hutton et al., 2016; Ayllón et al., 2021).

The ODD protocol (Grimm et al., 2006, 2010) was used to capture these considerations in the GPSCA study, adoption of which mandates that pertinent aspects of scale and their representations are documented. The common team goals and the minimum skills/knowledge needed to achieve those goals (e.g., specific expertise in aspects of software and model development) could advantageously be made explicit as part of this process as well. Moreover, such a table is recommended here to be continually updated to consider new information and lessons learnt throughout the modeling cycle.

Others have suggested increasing the number of meetings on the progress of the modeling and to incorporate reflexive evaluation of the team (Preston et al., 2015; Dongen et al., 2018; Delice et al., 2019; Gool et al., 2019). Increased frequency and number of meetings (whether face-to-face or virtual) in effect raises the minimum number of interactions between team members so that knowledge sharing can occur. Contextual examples of how these may be helpful with regard to teams are discussed elsewhere (see Kragt et al., 2013; Cockerill et al., 2019; Zare et al., 2021); however, support for reflexive activities must be available at the organizational level (Salas et al., 2018).

What is perhaps more important than meetings, however, is a team (and organizational) *culture* that allows for empathetic and inclusive communication to occur. Team members may speak different languages or at least adopt heavy disciplinary accents. Preferring one language or dialect at the expense of a "shared language" (Thomas and McDonagh, 2013) could lead to a disregard of relevant knowledge no matter the number, length, format or medium of meetings, or how expansive the documentation (as alluded in Section 2.2.3.3). An overreliance on technological solutions to communication without acknowledging the role of team and organizational culture may lead to more, rather than fewer, misunderstandings (cf. Andres, 2012; Benishek and Lazzara, 2019).

In addition to the reflexive monitoring and evaluation of team processes (as in Driskell et al., 2020; Zare et al., 2021), we recommend that such processes additionally account for the culture that underpins knowledge sharing and communication. Ignoring the role of team and organizational culture risks naturalizing the intuitions of its most privileged members (cf. James, 2014). An open attitude and commitment to continual and collaborative learning, both individually and as a group, is necessary for disciplinary barriers to be broken down and perspectives to be embraced (Woolley and Malone, 2011; Thomas and McDonagh, 2013; MacLeod and Nagatsu, 2018). In essence, teams would ideally culturally evolve throughout the modeling cycle toward more effective models of (interdisciplinary) cooperation (cf. Wilson and Wilson, 2007).

3.2. Document the rationale and reasons for scale choices

Debates about appropriate scales at which to represent structures and processes in multidisciplinary models should pervade discussions among modeling team members, particularly during conceptual model formulation and initial attempts to quantify linkages among model components. Most commonly, however, we begin model formulation with preconceived notions about the appropriate scales with which to represent the structures and processes in those parts of the system with which we are familiar, framed by workflows with which we are accustomed to. These preconceived notions typically are based on the way we have found most useful to think about such structures and processes in the past. Thus, the conceptualizations are coherent from a disciplinary perspective, but the cohesion breaks down when encountering other disciplines.

Our perceived usefulness of system representations is biased by our disciplinary training and experience (Huutoniemi et al., 2010). Such preconceived notions may blind us to alternate, yet still valid, representations or otherwise cause their dismissal as being of little use or simply incorrect. For example, the choice of a daily time step in the GPSCA study was informed by a shared familiarity with daily weather reports and the concept of degree-days of development of plants and insects. A 2-day time step may have been considered, thus cutting computing time in half, arguably without sacrificing usefulness of model output to end users. But a 2-day time step never crossed our minds. With the Campaspe case study, the primary focus on water-related agricultural concerns is partly a result of the level of engagement with agricultural experts (see Section 2.2.3.2), but also that the agency requirements for assessing the instream and riparian ecological impacts were quite modest. Consequently, the possibility of representing adaptive management processes of ecological issues was not actively considered (described in Section 2.2.3.1).

In building a shared understanding to develop a cohesive and complete treatment of scale, it may be more productive to agree to disagree on certain scaling issues that are particularly problematic during conceptual model formulation. Issues that are virtually impossible to resolve conceptually were almost always, in the case of the GPSCA study, clarified via quantification of the factors involved. The issues were clarified in the sense that differences in model output resulting from the use of different scales are made precise. Another reason to move on is that scale transitions that seem easy to accomplish when described in narrative form may be surprisingly difficult to accomplish computationally and which may require modifications that obviate the initially identified scale problems.

The choices made regarding scale were therefore influenced by the people involved and of course their perspectives and judgments. A "perfect" model is not possible, so we choose scales, which we believe best represent the system given "real-world" constraints. These choices are a series of subjective decisions involving consideration of model objectives and available information and resources at the time. A different group of people may arrive at a completely different, and perhaps equally plausible, valuable, and useful, model. The considerations and choices in the treatment of scale should be documented and made transparent for this reason. Such documentation allows researchers external to the process (and their future selves) to better understand the sociotechnical context in which the modeling decisions were made, the reasoning behind the decisions, and any implications or consequences from those decisions. Thus, documentation of the process helps illuminate model limitations and uncertainty (Refsgaard et al., 2006; van der Sluijs, 2007; Reichert, 2020).

We offer a final comment regarding the paucity of documentation available describing the debates preceding final SoS scaling decisions. For example, the ODD protocol, which is widely used to document agent-based models in ecology and which we used to document the GPSCA model, begins with a statement of model purpose followed by a second section that defines model entities (agents), state variables (attributes of agents), and (temporal and spatial) scales. Although this second section requires a justification of the final scale choices for each model component, it does not require documentation of the pros and cons of the alternative scales that were debated over time. Thus, a rich source of information defining the larger context of the modeling decisions, which would be particularly useful when contemplating reuse of the model, often is lost.

3.3. Acknowledge that causality is defined subjectively

When we described process representation in our GPSCA study (Section 2.1.2.2), we referred to the concept of a continuum of levels of perceived causality, of "subject-matter interpretability," extending in a theoretical sense from purely phenomenological/correlative to entirely mechanistic/explanatory. In practice, how different people perceive the representation of any given process in an SoS model will almost surely differ. In terms of model credibility, the important point is that all stakeholders, and here we include members of the modeling team as well as end-users of the model, perceive that the model behavior of most interest to them results from processes represented at an acceptable level of causality, at an acceptable level of subject-matter interpretability. Of overriding importance is that end users can explain, and hence understand, model output in cause–effect terms meaningful to them. But it also is important that members of the modeling team perceive the representations of processes in their areas of expertise as scientifically credible, given the objectives of the integrated SoS model. The cause–effect relationships responsible for output of the integrated model may be explained acceptably to end users in highly aggregated terms, whereas subject-matter specialists may require relatively detailed representations of some modeled processes in order for them to acknowledge those representations as causal.

Debates related to scale decisions in integrated SoS modeling are inextricably related to perceptions of causality. Scale decisions include not only those associated with defining temporal and spatial scales per se but also decisions associated with identifying which components and processes in the real system to include in the model and deciding at what level of detail to represent them. In our GPSCA case study, such debates arose regarding the level of detail with which to represent processes related to the aphid life cycle and the phenological development of sorghum. As described in Section 2.1.2.2, the final decision, which resulted from a lively debate among modeling team members, was to represent these processes as a function of environmental temperature modified by aphid density. Our meteorologist/aeroecologist would have been satisfied with a "causal" representation of aphid population dynamics that represented population density as a function of number of days since initial infestation and emigration as a function of density. Such a representation was perceived as unacceptably phenomenological by our entomologist. Our entomologist initially proposed a more mechanistic representation of the aphid life cycle, which included, among other things, mortality due to natural enemies (predators and parasites). However, in view of the site specificity of such relationships and the fact that the purpose of the integrated SoS model was to simulate areawide spatiotemporal patterns of aphid infestations, our entomologist agreed to a simpler "causal" representation of the aphid life cycle.

As mentioned in Section 2.2.3.1, there were several approaches to represent the spatial areas for the various models in the Campaspe case study. Each were arguably plausible, and objections could be raised depending on modeler perspectives and understanding of the modeling context. Here, we remind modelers that representing greater detail may not be appropriate given the model purpose and context. The "bigger picture" should be kept in mind.

The lesson learnt is that it would serve modeling teams well if their members explicitly acknowledged the subjective nature of their perception of causality at the very beginning of the modeling process. A discussion focused on the concept of a continuum of levels of perceived causality would be time well spent. The initial response to such a discussion most likely would be "everyone already knows that," which probably is true enough if viewed as an abstract concept. But based on our experience, we are quite sure that if early discussions among modeling team members were documented and reexamined, it would be obvious that the subjective nature of defining causality is seldom recognized in practice.

3.4. Embrace change and reflect throughout the iterative modeling cycle

The modeling process is commonly described as undergoing a "cycle" of iterations of a set of (concurrent) phases and steps. Although the number of steps and activities conducted may differ depending on purpose and conceptualization of the cycle (Boehm, 1986; Jakeman et al., 2006; Pianosi et al., 2016; Badham et al., 2019; Arnold et al., 2020; Zare et al., 2021), each step is intended to be revisited as often as needed to incorporate newly discovered or available knowledge, or ideas generated on deep reflection, as "[t]he first model is rarely the best model" (Sterling et al., 2019). It may at times be necessary to abandon an iteration and start over.

Arguably, recognizing and embracing the need for change is fundamental to the flexibility that iterative approaches afford (Dingsøyr et al., 2012; Strode et al., 2012). In the SoS context, the modeling process may have to be restarted due to discovery or incorporation of new knowledge *for another constituent system*, necessitating changes to one's own constituent model or even the modeling process. A shift in scales may be a (pragmatic) necessity to accommodate the integration of constituent models and such a decision may be governed, or have implications toward, data availability/requirement, computational capacity, and model purpose.

Change is inevitable due to the complexity of the systems being studied and the speed at which new information may come to light. Where team members are more accustomed to single-system investigations, a *cultural* shift in thinking may be required to enable flexible response to the (continuous) adjustment of scale, in all its forms. New information may necessitate skills to be acquired or adapted to an unfamiliar modeling context (Knapen et al., 2013; Voznesenskaya et al., 2019). As noted in Section 3.1, being overly tied to a single disciplinary perspective results in an inflexible system conceptualization that is resistant to "new" knowledge or perspectives. The adoption of new practices, technologies, and workflows more amenable to the new modeling context is therefore restricted and hampers team productivity (Cockburn and Highsmith, 2001; Hoda et al., 2013). The lack of version control of model code and data and the consequent effect in the development of the CIM was given as an example in Section 2.2.3.3. Change should be embraced for the lessons learnt to be effectively carried over between iterations and for knowledge to be cross-pollinated between team members (Knight et al., 2001; Lee et al., 2015).

3.5. Regularly test the integration

The reality of iterative development means that (1) constituent models may be of varying complexity and developed against different schedules, (2) changes made in one model may necessitate changes in another, (3) the necessary computational requirements and available computational infrastructure may preclude the possibility of calibrating all models at once, and (4) issues may only become a highlighted concern in the integrated context as the implications of the scale and volume of interactions may not be apparent until all models are coupled.

A somewhat naive view is that any topically relevant sectoral model can be coupled and applied to represent an SoS. This may be true at a technical level but without regard for its conceptual, and contextual, appropriateness, the resulting model is likely to be unwieldy, overly complex, and unsuited for a given purpose (e.g., Voinov and Shugart, 2013). In addressing water resource management problems, for example, Croke et al. (2014) argue that hydrological model choice requires engagement with appropriate concepts, model structures, scales of analyses, performance evaluation, and communication. Again, such issues may not be evident until the scale of the modeling becomes sufficiently expansive. Thus, the relevance of any constituent model to the integrated model's purpose and the propagation of uncertainty needs serious evaluation.

Specific to model coupling, future work could investigate a typology of design elements, which make models more amenable for their use in SoS modeling contexts and classify system models along those lines. In the short-to-medium term, strategies and plans to address or mitigate the impact of a constituent model that turns out to be not wholly suited to the SoS modeling context, such as when the scales of the problem involved increase could be explored. In the case of the CIM, computational performance became a concern as the scale of the modeling increased. One approach would have been to develop a model specifically for the integrated context, as opposed to the (continued) use of a legacy model. In the end, the issue was sidestepped by leveraging high-performance computing infrastructure. Ideally, such considerations would be considered and planned for early in the modeling process.

Methodologically, conceptual-or-hypothesis testing is one (but not the only) approach that may be applied to address concerns around the structure of the SoS model (Wilson et al., 2017; Iwanaga et al., 2020). Such testing approaches involve the identification of questions with a known range of acceptable answers that the SoS model can produce. The greater the number of such tests that can cover the range of possible realities being simulated by the model, the more confident modelers can be that the integrated model is functioning correctly, both technically and conceptually. Property-based sensitivity analysis is one approach (of many) leveraged in the development of the CIM to alert modelers of technical and conceptual issues in model integration (Iwanaga et al., 2021a). Continual testing and integration throughout the modeling process could then highlight context change (e.g., cases wherein previous understandings are falsified) and facilitate understanding of model structure and behavior (Iwanaga et al., 2021a).

In this manner, conceptual tests frame the context for incorrect model behavior. Frequent integration and testing, even at this highly aggregated level, is likely to highlight conceptual mismatches between the knowledge of disciplinary experts and model implementations. Testing of the models and their integration throughout the development cycle then plays an important role in ensuring issues are identified earlier in development (Warren, 2014). Earlier correction of issues helps to avoid "wasteful" model runs and quickens the pace through the modeling cycle. It would be beneficial if all modelers involved strive to enable repeated, and frequent, integration and testing.

Appendix A

Table A1. Individual systems represented in the Campaspe integrated model and their spatial, temporal, and data aspects. DOI: https://doi.org/10.1525/elementa.2020.00182.t2

Constituent System	Spatial	Temporal	Metrics/Data
Climate	5 km grid (0.05°, interpolated) matching the groundwater area	Daily time step. Available data constrained the time frames considered	Data represented differing levels of aridity ranging from extreme dry to "wet" over a 30-year time frame. Data sourced via Climate Change in Australia (CSIRO, 2020)
Groundwater, implemented with MODFLOW-NWT with FloPy interface (Bakker et al., 2016)	5 km grid, seven layers of variable thickness based on hydrogeologic units Higher spatial resolutions were impractical due to the long runtime of MODFLOW-NWT Largest spatial extent, extending further west than other models. Covers 4,896 km^2. Assumes irrigation events are uniformly applied across farm zone areas	Daily time step Assumes irrigation input from the farm model is to be uniformly disaggregated across 14 days	Estimates distance to water table, which influences farm groundwater pumping costs (farmer decisions) and groundwater allocations (policy) Provides estimations of surface–groundwater exchange along the river
Surface water, implemented with IHACRES_GW (Ivkovic et al., 2014)	Lumped, node-based routing model. Nodes represent subcatchments. Covers 3,518 km^2 Extends further south compared to the groundwater model to estimate inflows to the dam Assumes irrigation events are uniformly applied across farm zone areas	Daily time step Assumes irrigation input from the farm model is to be uniformly disaggregated across 14 days	Calculates dam levels, influencing water allocations for both environmental and agricultural users and perceived recreational value Stage height along the river is also provided for policy and ecology models
Farm model	Lumped, zone-based. Each zone represents farming areas of variable size. Covers 2,154 km^2	2 week time step, indicated to be the usual time frame in which irrigation decisions are made (Xie et al., 2019) Total volume of rainfall over the previous 14 days is used to determine irrigation schedule	Crop yield, farm profit estimations, water use (in ml) Incorporated data from farmer surveys
Policy model	Regional/catchment-wide	2 week time step Temporally, the model operates on a 14-day time step matching that of the farm model. In reality, such allocations are announced every 6 weeks and so constitutes	Surface water allocations, determined by dam levels. Groundwater allocations determined by groundwater level at two bores (one in the south, one in the north)

(continued)

TABLE A1. *(continued)*

Constituent System	Spatial	Temporal	Metrics/Data
		a finer grain regulation of water availability	Hypothetical conjunctive use policies allow further extraction of groundwater under periods of low surface water allocations and restrict these in times where surface water allocations are met
Recreational suitability index model	Dam	Lumped—Result is the average index value for the modeled time frame	Suitability of dam for recreational use, tied to the water level in the dam (Lake Eppalock). Generally speaking, higher dam levels allow a higher level of enjoyment to be experienced by recreational users
Ecological suitability index model	River and groundwater levels	Lumped—Result is the average index value for the modeled time frame	Suitability of flow at key gauges to support ecological activities for platypus, fish, and river red gums (iconic trees found in the area). These are lumped together into a single metric

Data accessibility statement

Data sharing is not applicable to this article as the data analyzed were our reflexive experiences, reported within this article.

Acknowledgments

The U.S. National Oceanic and Atmospheric Administration Air Resources Laboratory (ARL) provided access to the HYSPLIT transport and dispersion model. The authors additionally acknowledge past contributions of colleagues and local knowledge holders toward the sugarcane aphids in Great Plains (area wide pest management) and Campaspe case studies without which this work would not have been possible. The authors finally would like to thank the two anonymous reviewers and the Associate Editor, Prof. Zavaleta (University of California, Santa Cruz), for their time and effort. This article was greatly improved as a result of their comments.

Funding

The primary author (Takuya Iwanaga) is supported through an Australian Government Research Training Program Scholarship and a top-up scholarship from the Australian National University Hilda-John Endowment Fund. Hsiao-Hsuan Wang and Tomasz E. Koralewski acknowledge partial support from the USDA Agricultural Research Service provided through the Areawide Pest Management Program, "Areawide Pest Management of the Invasive Sugarcane Aphid in Grain Sorghum," project number 3072-22000-017-07-S. The Campaspe Integrated Model was developed as part of the Murray–Darling Basin Authority's partnership with the National Centre for Groundwater Research and Training under Contract No. MD2594. John Little acknowledges support from National Science Foundation (NSF) Award EEC 1937012. This work was also supported by the National Socio-Environmental Synthesis Center under funding received from the NSF DBI-1639145.

Competing interests

There were no competing interests with regard to this article.

Author contributions

Contributed to the conception and approach: TI, H-HW.

Contributed to the reflexive analysis: TI, H-HW, TEK, WEG, AJJ.

Additional perspectives and commentary to enhance reflexive analysis: JCL.

Drafted and/or revised this article: TI, H-HW.

Additional contributions to manuscript and revisions: TEK, WEG, AJJ, JCL.

References

Andres, HP. 2012. Technology-mediated collaboration, shared mental model and task performance. *Journal of Organizational and End User Computing* **24**: 18. DOI: http://dx.doi.org/10.4018/joeuc.2012010104.

Arnold, T, Guillaume, JHA, Lahtinen, TJ, Vervoort, RW, 2020. From ad-hoc modelling to strategic infrastructure: A manifesto for model management. *Environmental Modelling & Software* **123**: 104563. DOI: http://dx.doi.org/10.1016/j.envsoft.2019.104563.

Ayllón, D, Railsback, SF, Gallagher, C, Augusiak, J, Baveco, H, Berger, U, Charles, S, Martin, R, Focks, A, Galic, N, Liu, C, van Loon, EE, Nabe-Nielsen, J, Piou, C, Polhill, JG, Preuss, TG, Radchuk, V, Schmolke, A, Stadnicka-Michalak, J, Thorbek, P, Grimm, V. 2021. Keeping modelling notebooks with TRACE: Good for you and good for environmental research and management support. *Environmental Modelling & Software* **136**. DOI: http://dx.doi.org/10.1016/j.envsoft.2020.104932.

Badham, J, Elsawah, S, Guillaume, JHA, Hamilton, SH, Hunt, RJ, Jakeman, AJ, Pierce, SA., Snow, VO, Babbar-Sebens, M, Fu, B, Gober, P, Hill, MC, Iwanaga, T, Loucks, DP, Merritt, WS, Peckham, SD, Richmond, AK, 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. DOI: http://dx.doi.org/10.1016/j.envsoft.2019.02.013.

Bakker, M, Post, V, Langevin, CD, Hughes, JD, White, JT, Starn, JJ, Fienen, MN. 2016. Scripting MODFLOW model development using Python and FloPy. *Groundwater* **54**: 733–739. DOI: http://dx.doi.org/10.1111/gwat.12413.

Banerjee, P, Wang, H-H, Peterson, MJ, Grant, WE, Peterson, TR. 2019. Collaborative modeling and social learning in the context of joint forest management in east Sikkim, India. *Frontiers of Environmental Science* **7**. DOI: http://dx.doi.org/10.3389/fenvs.2019.00154.

Benishek, LE., Lazzara, EH. 2019. Teams in a new era: Some considerations and implications. *Frontiers in Psychology* **10**. DOI: http://dx.doi.org/10.3389/fpsyg.2019.01006.

Beven, KJ. 2009. *Environmental modelling: An uncertain future? An introduction to techniques for uncertainty estimation in environmental prediction*. London, UK: Routledge.

Bezanson, J, Edelman, A, Karpinski, S, Shah, VB. 2017. Julia: A fresh approach to numerical computing. *SIAM Review* **59**: 65–98. DOI: http://dx.doi.org/10.1137/141000671.

Boehm, B. 1986. A spiral model of software development and enhancement. *SIGSOFT Software Engineering Notes* **11**: 14–24. DOI: http://dx.doi.org/10.1145/12944.12948.

Bolton, G., Delderfield, R. 2018. *Reflective practice: Writing and professional development* (5th ed.). Los Angeles, CA: SAGE.

Bowling, RD, Brewer, MJ, Kerns, DL, Gordy, J, Seiter, N, Elliott, NE, Buntin, GD, Way, MO, Royer, TA, Biles, S, Maxson, E. 2016. Sugarcane Aphid (Hemiptera: Aphididae): A new pest on sorghum in North America. *Journal of Integrated Pest Management* **7**. DOI: http://dx.doi.org/10.1093/jipm/pmw011.

Brewer, MJ, Gordy, JW, Kerns, DL, Woolley, JB, Rooney, WL, Bowling, RD. 2017. Sugarcane aphid population growth, plant injury, and natural enemies on selected grain sorghum hybrids in Texas and Louisiana. *Journal of Economic Entomology* **110**: 2109–2118. DOI: http://dx.doi.org/10.1093/jee/tox204.

Brewer, MJ, Peairs, FB, Elliott, NC. 2019. Invasive cereal aphids of North America: Ecology and pest management. *Annual Review of Entomology* **64**: 73–93. DOI: http://dx.doi.org/10.1146/annurev-ento-011118-111838.

Catalano, AS, Lyons-White, J, Mills, MM, Knight, AT. 2019. Learning from published project failures in conservation. *Biological Conservation* **238**. DOI: http://dx.doi.org/10.1016/j.biocon.2019.108223.

Claessen, K, Hughes, J. 2000. *QuickCheck: A lightweight tool for random testing of Haskell programs*, in *Proceedings of the Fifth ACM SIGPLAN International Conference on Functional Programming*, ICFP ‘00. New York, NY: Association for Computing Machinery: 268–279. DOI: http://dx.doi.org/10.1145/351240.351266.

Cockburn, A, Highsmith, J, 2001. Agile software development, the people factor. *Computer* **34**: 131–133. DOI: http://dx.doi.org/10.1109/2.963450.

Cockerill, K, Glynn, P, Chabay, I, Farooque, M, Hämäläinen, RP, Miyamoto, B, McKay, P. 2019. Records of engagement and decision making for environmental and socio-ecological challenges. *Euro Journal on Decision Processes*. DOI: http://dx.doi.org/10.1007/s40070-019-00104-6.

Connolly, R., 2020. Why computing belongs within the social sciences. *Communications of the ACM* **63**: 54–59. DOI: http://dx.doi.org/10.1145/3383444.

Croke, BFW, Blakers, RS, El Sawah, S, Fu, B, Guillaume, JHA, Kelly, RA, Patrick, MJ, Ross, A, Ticehurst, J, Barthel, R, Jakeman, AJ. 2014 June 4–6. Marrying hydrological modelling and integrated assessment for the needs of water resource management, in *Proceedings of the International Association of Hydrological Sciences. Presented at the Proceedings of ICWRS2014—6th IAHS-EGU International Symposium on Integrated Water Resources Management*. Bologna, Italy: Copernicus GmbH: 351–356. DOI: http://dx.doi.org/10.5194/piahs-364-351-2014.

Croke, BFW, Jakeman, AJ. 2004. A catchment moisture deficit module for the IHACRES rainfall-runoff model. *Environmental Modelling & Software* **19**: 1–5. DOI: http://dx.doi.org/10.1016/j.envsoft.2003.09.001.

Crouzat, E, Arpin, I, Brunet, L, Colloff, MJ, Turkelboom, F, Lavorel, S. 2018. Researchers must be aware of their roles at the interface of ecosystem services science and policy. *Ambio* **47**: 97–105. DOI: http://dx.doi.org/10.1007/s13280-017-0939-1.

CSIRO. 2020. Climate Change in Australia. Climate Change in Australia. Available at https://climatechangeinaustralia.gov.au/en/. Accessed 24 February 2020.

Davidson, J. 1944. On the relationship between temperature and rate of development of insects at constant temperatures. *Journal of Animal Ecology* **13**: 26–38. DOI: http://dx.doi.org/10.2307/1326.

Davidson-Pilon, C. 2016. *Bayesian methods for hackers: Probabilistic programming and Bayesian inference*. New York: Addison-Wesley.

De Dreu, CKW. 2008. The virtue and vice of workplace conflict: Food for (pessimistic) thought. *Journal of Organizational Behavior* **29**: 5–18. DOI: http://dx.doi.org/10.1002/job.474.

Delice, F, Rousseau, M, Feitosa, J, 2019. Advancing teams research: What, when, and how to measure team dynamics over time. *Frontiers in Psychology*

10. DOI: http://dx.doi.org/10.3389/fpsyg.2019.01324.

Dey, R, Lewis, SC, Arblaster, JM, Abram, NJ. 2019. A review of past and projected changes in Australia's rainfall. *WIREs Climate Change* **10**: e577. DOI: http://dx.doi.org/10.1002/wcc.577.

Dingsøyr, T, Nerur, S, Balijepally, V, Moe, NB. 2012. A decade of agile methodologies: Towards explaining agile software development. *Journal of Systems and Software, Special Issue: Agile Development* **85**: 1213–1221. DOI: http://dx.doi.org/10.1016/j.jss.2012.02.033.

Driskell, JE, Salas, E, Driskell, T. 2020. Foundations of teamwork and collaboration. *American Psychologist* **73**: 334–348.

Dunford, R, Harrison, PA, Rounsevell, MDA. 2015. Exploring scenario and model uncertainty in cross-sectoral integrated assessment approaches to climate change impacts. *Climatic Change* **132**: 417–432. DOI: http://dx.doi.org/10.1007/s10584-014-1211-3.

Dysarz, T. 2018. Application of Python scripting techniques for control and automation of HEC-RAS simulations. *Water* **10**: 1382. DOI: http://dx.doi.org/10.3390/w10101382.

Eagles, D, Walker, PJ, Zalucki, MP, Durr, PA. 2013. Modelling spatio-temporal patterns of long-distance Culicoides dispersal into northern Australia. *Preventive Veterinary Medicine* **110**: 312–322. DOI: http://dx.doi.org/10.1016/j.prevetmed.2013.02.022.

Elliott, NC, Onstad, DW, Brewer, MJ. 2008. History and ecological basis for areawide pest management, in Koul, O, Cuperus, GW, Elliott, N eds., *Areawide pest management: Theory and implementation*. Wallingford, UK: CABI: 15–33.

Elsawah, S, Filatova, T, Jakeman, AJ, Kettner, AJ, Zellner, ML, Athanasiadis, IN, Hamilton, SH, Axtell, RL, Brown, DG, Gilligan, JM, Janssen, MA, Robinson, DT, Rozenberg, J, Ullah, IIT, Lade, SJ. 2020. Eight grand challenges in socio-environmental systems modeling. *Socio-Environmental Systems Modeling* **2**: 16226. DOI: http://dx.doi.org/10.18174/sesmo.2020a16226.

Emery, D, Hilliard, R. 2008. Updating IEEE 1471: Architecture Frameworks and Other Topics, in *Presented at the seventh working IEEE/IFIP conference on Software Architecture (WICSA 2008)*: 303–306. DOI: http://dx.doi.org/10.1109/WICSA.2008.32.

Evans, MR. 2012. Modelling ecological systems in a changing world. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367**: 181–190. DOI: http://dx.doi.org/10.1098/rstb.2011.0172.

Finlay, L. 2002. "Outing" the researcher: The provenance, process, and practice of reflexivity. *Qualitative Health Research* **12**. DOI: http://dx.doi.org/10.1177/104973202129120052.

Forrester, JW. 1961. *Industrial dynamics*. First edition. Cambridge: MIT Press.

Freeman, RE. 2010. *Strategic management: A stakeholder approach, Reissue*. Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore: Cambridge University Press.

Gerik, T, Bean, BW, Vanderlip, R. 2003. *Sorghum growth and development, Texas FARMER Collection*. College Station, TX: Texas AgriLife Extension Service.

Giles, K., Hein, G., Peairs, F., 2008. Areawide pest management of cereal aphids in dryland wheat systems of the Great Plains, USA, in Koul, O, Cuperus, GW, Elliott, N eds., *Areawide pest management: Theory and implementation*. Wallingford, UK: CABI: 441–466.

Glynn, PD, Voinov, AA, Shapiro, CD, White, PA. 2017. From data to decisions: Processing information, biases, and beliefs for improved management of natural resources and environments. *Earth's Future* **5**: 356–378. DOI: http://dx.doi.org/10.1002/2016EF000487.

Gren, L, Lenberg, P. 2018. The importance of conflict resolution techniques in autonomous agile teams, in *Proceedings of the 19th International Conference on Agile Software Development: Companion*, XP '18. New York, NY: Association for Computing Machinery: 1–4. DOI: http://dx.doi.org/10.1145/3234152.3234185.

Grimm, V, Berger, U, Bastiansen, F, Eliassen, S, Ginot, V, Giske, J, Goss-Custard, J, Grand, T, Heinz, SK, Huse, G, Huth, A, Jepsen, JU, Jørgensen, C, Mooij, WM, Müller, B, Pe'er, G, Piou, C, Railsback, SF, Robbins, AM, Robbins, MM, Rossmanith, E, Rüger, N, Strand, E, Souissi, S, Stillman, RA, Vabø, R, Visser, U, DeAngelis, DL. 2006. A standard protocol for describing individual-based and agent-based models. *Ecological Modelling* **198**: 115–126. DOI: http://dx.doi.org/10.1016/j.ecolmodel.2006.04.023.

Grimm, V, Berger, U, DeAngelis, DL, Polhill, JG, Giske, J, Railsback, SF, 2010. The ODD protocol: A review and first update. *Ecological Modelling* **221**: 2760–2768. DOI: http://dx.doi.org/10.1016/j.ecolmodel.2010.08.019.

Grimm, V, Railsback, SF. 2005. *Individual-based modeling and ecology*. STU-Student. Princeton, NJ: Princeton University Press.

Grimm, V, Railsback, SF, Vincenot, CE, Berger, U, Gallagher, C, DeAngelis, DL, Edmonds, B, Ge, J, Giske, J, Groeneveld, J, Johnston, ASA, Milles, A, Nabe-Nielsen, J, Polhill, JG, Radchuk, V, Rohwäder, M-S, Stillman, RA, Thiele, JC, Ayllón, D. 2020. The ODD protocol for describing agent-based and other simulation models: A second update to improve clarity, replication, and structural realism. *Journal of Artificial Societies and Social Simulation* **23**: 7. DOI: http://dx.doi.org/10.18564/jasss.4259.

Hall, KL, Vogel, AL, Huang, GC, Serrano, KJ, Rice, EL, Tsakraklides, SP, Fiore, SM. 2018. The science of team science: A review of the empirical evidence and research gaps on collaboration in science. *American Psychologist* **73**: 532–548. DOI: http://dx.doi.org/10.1037/amp0000319.

Hall, KL, Vogel, AL, Stipelman, BA, Stokols, D, Morgan, G, Gehlert, S. 2012. A four-phase model of transdisciplinary team-based research: Goals, team processes, and strategies. *Translational Behavioral Medicine* **2**: 415–430. DOI: http://dx.doi.org/10.1007/s13142-012-0167-y.

Hamilton, SH, ElSawah, S, Guillaume, JHA, Jakeman, AJ, Pierce, SA. 2015. Integrated assessment and modelling: Overview and synthesis of salient dimensions. *Environmental Modelling & Software* **64**: 215–229. DOI: http://dx.doi.org/10.1016/j.envsoft.2014.12.005.

Hansen, MJ, Vaagen, H, van Oorschot, K. 2020. Team collective intelligence in dynamically complex projects—A shipbuilding case. *Project Management Journal*. DOI: http://dx.doi.org/10.1177/8756972820928695.

Harris, CR, Millman, KJ, van der Walt, SJ, Gommers, R, Virtanen, P, Cournapeau, D, Wieser, E, Taylor, J, Berg, S, Smith, NJ, Kern, R, Picus, M, Hoyer, S, van Kerkwijk, MH, Brett, M, Haldane, A, del Río, JF, Wiebe, M, Peterson, P, Gérard-Marchant, P, Sheppard, K, Reddy, T, Weckesser, W, Abbasi, H, Gohlke, C, Oliphant, TE. 2020. Array programming with NumPy. *Nature* **585**: 357–362. DOI: http://dx.doi.org/10.1038/s41586-020-2649-2.

Heffernan, M. 2011. *Willful blindness: Why we ignore the obvious*. First U.S. edition. New York, NY: Walker & Co.

Hidalgo, ES. 2019. Adapting the scrum framework for agile project management in science: Case study of a distributed research initiative. *Heliyon* **5**: e01447. DOI: http://dx.doi.org/10.1016/j.heliyon.2019.e01447.

Hinson, P. 2017. Effect of temperature on the development of sugarcane aphid, *Melanaphis sacchari*, on Sorghum [Thesis]. West Texas A&M University. Available at http://hdl.handle.net/11310/140. Accessed 2 June 2021.

Hoda, R, Noble, J, Marshall, S. 2013. Self-organizing roles on agile software development teams. *IEEE Transactions on Software Engineering* **39**: 422–444. DOI: http://dx.doi.org/10.1109/TSE.2012.30.

Horne, AC, Nathan, R, Poff, NL, Bond, NR, Webb, JA, Wang, J, John, A. 2019. Modeling flow-ecology responses in the Anthropocene: Challenges for sustainable riverine management. *BioScience* **69**: 789–799. DOI: http://dx.doi.org/10.1093/biosci/biz087.

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

Huutoniemi, K, Klein, JT, Bruun, H. Hukkinen, J. 2010. Analyzing interdisciplinarity: Typology and indicators. *Research Policy* **39**: 79–88. DOI: http://dx.doi.org/10.1016/j.respol.2009.09.011.

Ivkovic, KM, Croke, BFW, Kelly, RA. 2014. Overcoming the challenges of using a rainfall–runoff model to estimate the impacts of groundwater extraction on low flows in an ephemeral stream. *Hydrology Research* **45**: 58–72. DOI: http://dx.doi.org/10.2166/nh.2013.204.

Iwanaga, T, Partington, D, Ticehurst, J, Croke, BFW, Jakeman, AJ. 2020. A socio-environmental model for exploring sustainable water management futures: Participatory and collaborative modelling in the lower Campaspe catchment. *Journal of Hydrology: Regional Studies* **28**. DOI: http://dx.doi.org/10.1016/j.ejrh.2020.100669.

Iwanaga, T, Sun, X, Wang, Q, Guillaume, JHA, Croke, BFW, Rahman, J, Jakeman, AJ. 2021a. Property-based sensitivity analysis: An approach to identify model implementation and integration errors. *Environmental Modelling & Software* **139**. DOI: http://dx.doi.org/10.1016/j.envsoft.2021.105013.

Iwanaga, T, Wang, H-H, Hamilton, SH, Grimm, V, Koralewski, TE, Salado, A, Elsawah, S, Razavi, S, Yang, J, Glynn, P, Badham, J, Voinov, A, Chen, M, Grant, WE, Peterson, TR, Frank, K, Shenk, G, Barton, CM, Jakeman, AJ, Little, JC. 2021b. Socio-technical scales in socio-environmental modeling: Managing a system-of-systems modeling approach. *Environmental Modelling & Software*. DOI: http://dx.doi.org/10.1016/j.envsoft.2020.104885.

Iwanaga, T, Zare, F, Croke, B, Fu, B, Merritt, W, Partington, D, Ticehurst, J, Jakeman, A, 2018 June 13–15. Development of an integrated model for the Campaspe catchment: A tool to help improve understanding of the interaction between society, policy, farming decision, ecology, hydrology and climate, in *Proceedings of the International Association of Hydrological Sciences. Presented at the Innovative Water Resources Management—Understanding And Balancing Interactions between Humankind And Nature—8th International Water Resources Management Conference of ICWRS*. Beijing, China: Copernicus GmbH: 1–12. DOI: http://dx.doi.org/10.5194/piahs-379-1-2018.

Jakeman, AJ, Letcher, RA, Norton, JP. 2006. Ten iterative steps in development and evaluation of environmental models. *Environmental Modelling and Software* **21**: 602–614. DOI: http://dx.doi.org/10.1016/j.envsoft.2006.01.004.

James, R. 2014. What we can learn about Philosophy's diversity problems by comparing ourselves to Music Theory. Available at https://www.its-her-factory.com/2014/10/what-we-can-learn-about-philosophys-diversity-problems-by-comparing-ourselves-to-music-theory/. Accessed 29 October 2020.

John, A, Horne, A, Nathan, R, Stewardson, M, Webb, JA, Wang, J, Poff, NL. 2020. Climate change and freshwater ecology: Hydrological and ecological methods of comparable complexity are needed to predict risk. *WIREs Climate Change*: e692. DOI: http://dx.doi.org/10.1002/wcc.692.

Jones, NA, Ross, H, Lynam, T, Perez, P, Leitch, A. 2011. Mental models: An interdisciplinary synthesis of theory and methods. *Ecology and Society* **16**(1): 46.

Kendall, M. 2013. Drought and its role in shaping water policy in Australia, in Schwabe, K, Albiac, J, Connor, JD, Hassan, RM, Meza González, L eds., *Drought in arid and semi-arid regions: A multi-disciplinary and cross-country perspective*. Dordrecht, the Netherlands: Springer: 451–467. DOI: http://dx.doi.org/10.1007/978-94-007-6636-5_26.

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

Knight, D, Durham, CC, Locke, EA. 2001. The relationship of team goals, incentives, and efficacy to strategic risk, tactical implementation, and performance. *Academy of Management Journal* **44**: 326–338. DOI: http://dx.doi.org/10.5465/3069459.

Koo, H, Iwanaga, T, Croke, BFW, Jakeman, AJ, Yang, J, Wang, H-H, Sun, X, Lü, G, Li, X, Yue, T, Yuan, W, Liu, X, Chen, M. 2020. Position paper: Sensitivity analysis of spatially distributed environmental models- a pragmatic framework for the exploration of uncertainty sources. *Environmental Modelling & Software*. DOI: http://dx.doi.org/10.1016/j.envsoft.2020.104857.

Koralewski, TE, Wang, H-H, Grant, WE, Brewer, MJ, Elliott, NC, Westbrook, JK, Szczepaniec, A, Knutson, A, Giles, KL, Michaud, JP. 2020a. Integrating models of atmospheric dispersion and crop-pest dynamics: Linking detection of local aphid infestations to forecasts of region-wide invasion of cereal crops. *Annals of the Entomological Society of America* **113**: 79–87. DOI: http://dx.doi.org/10.1093/aesa/saz047.

Koralewski, TE, Wang, H-H, Grant, WE, LaForest, JH, Brewer, MJ, Elliott, NC, Westbrook, JK. 2020b. Toward near-real-time forecasts of airborne crop pests: Aphid invasions of cereal grains in North America. *Computers and Electronics in Agriculture* **179**: 105861. DOI: http://dx.doi.org/10.1016/j.compag.2020.105861.

Koralewski, TE, Westbrook, JK, Grant, WE, Wang, H-H. 2019. Coupling general physical environmental process models with specific question-driven ecological simulation models. *Ecological Modelling* **405**: 102–105. DOI: http://dx.doi.org/10.1016/j.ecolmodel.2019.02.004.

Kragt, ME, Robson, BJ, Macleod, CJA. 2013. Modellers' roles in structuring integrative research projects. *Environmental Modelling & Software* **39**: 322–330. DOI: http://dx.doi.org/10.1016/j.envsoft.2012.06.015.

Krueger, T, Maynard, C, Carr, G, Bruns, A, Mueller, EN, Lane, S. 2016. A transdisciplinary account of water research. *WIREs Water* **3**: 369–389. DOI: http://dx.doi.org/10.1002/wat2.1132.

Lahtinen, TJ, Guillaume, JHA, Hmlinen, RP. 2017. Why pay attention to paths in the practice of environmental modelling? *Environmental Modelling and Software* **92**: 74–81. DOI: http://dx.doi.org/10.1016/j.envsoft.2017.02.019.

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

Levins, R. 1966. The strategy of model building in population biology. *American Scientist* **54**: 421–430.

Little, JC, Hester, ET, Elsawah, S, Filz, GM, Sandu, A, Carey, CC, Iwanaga, T, Jakeman, AJ. 2019. A tiered, system-of-systems modeling framework for resolving complex socio-environmental policy issues. *Environmental Modelling & Software* **112**: 82–94. DOI: http://dx.doi.org/10.1016/j.envsoft.2018.11.011.

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

Maier, HR, Guillaume, JHA, van Delden, H, Riddell, GA, Haasnoot, M, Kwakkel, JH. 2016. An uncertain future, deep uncertainty, scenarios, robustness and adaptation: How do they fit together? *Environmental Modelling & Software* **81**: 154–164. DOI: http://dx.doi.org/10.1016/j.envsoft.2016.03.014.

May, T, Perry, B. 2017. *Reflexivity: The essential guide* (1st ed.). Thousand Oaks, CA: SAGE.

McDonald, MG, Harbaugh, AW. 2003. The history of MODFLOW. *Groundwater* **41**: 280–283. DOI: https://doi.org/10.1111/j.1745-6584.2003.tb02591.x.

Montana, J, Elliott, L, Ryan, M, Wyborn, C. 2020. The need for improved reflexivity in conservation science. *Environmental Conservation* 1–3. DOI: http://dx.doi.org/10.1017/S0376892920000326.

Muller, E, Bednar, JA, Diesmann, M, Gewaltig, M-O, Hines, M, Davison, AP. 2015. Python in neuroscience. *Frontiers in Neuroinformatics* **9**. DOI: http://dx.doi.org/10.3389/fninf.2015.00011.

Murray–Darling Basin Authority. 2012. *Sustainable Rivers Audit 2: The ecological health of rivers in the Murray–Darling Basin at the end of the millennium drought (2008–2010).* Summary. (No. 75/12). Canberra, ACT, Australia: MDBA.

Nancarrow, SA, Booth, A, Ariss, S, Smith, T, Enderby, P, Roots, A. 2013. Ten principles of good interdisciplinary team work. *Human Resources for Health* **11**: 19. DOI: http://dx.doi.org/10.1186/1478-4491-11-19.

National Oceanographic and Atmospheric Administration. 2019. Air Resources Laboratory—EDAS 40 km Data Archive. Available at https://ready.arl.noaa.gov/edas40.php. Accessed 21 December 2020.

North Central CMA. 2014. *Campaspe river resource guide.* Huntly, UK: North Central Catchment Management Authority.

Parry, HR. 2013. Cereal aphid movement: General principles and simulation modelling. *Movement Ecology* **1**: 14. DOI: http://dx.doi.org/10.1186/2051-3933-1-14.

Partington, D, Knowling, MJ, Simmons, CT, Cook, PG, Xie, Y, Iwanaga, T, Bouchez, C, 2020. Worth of hydraulic and water chemistry observation data in terms of the reliability of surface water-groundwater exchange flux predictions under varied flow conditions. *Journal of Hydrology* **590**: 125441. DOI: http://dx.doi.org/10.1016/j.jhydrol.2020.125441.

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

Poché, DM, Grant, WE, Wang, H-H. 2016. Visceral Leishmaniasis on the Indian subcontinent: Modelling the dynamic relationship between vector control schemes and vector life cycles. *PLOS Neglected Tropical Diseases* **10**: e0004868. DOI: http://dx.doi.org/10.1371/journal.pntd.0004868.

Preston, BL, Rickards, L, Fünfgeld, H, Keenan, RJ. 2015. Toward reflexive climate adaptation research. *Current Opinion in Environmental Sustainability, Open Issue* **14**: 127–135. DOI: http://dx.doi.org/10.1016/j.cosust.2015.05.002.

Refsgaard, JC, van der Sluijs, JP, Brown, J, van der Keur, P. 2006. A framework for dealing with uncertainty due to model structure error. *Advances in Water Resources* **29**: 1586–1597. DOI: http://dx.doi.org/10.1016/j.advwatres.2005.11.013.

Refsgaard, JC, van der Sluijs, JP, Højberg, AL, Vanrolleghem, PA. 2007. Uncertainty in the environmental modelling process—A framework and guidance. *Environmental Modelling & Software* **22**: 1543–1556. DOI: http://dx.doi.org/10.1016/j.envsoft.2007.02.004.

Reichert, P. 2020. Towards a comprehensive uncertainty assessment in environmental research and decision support. *Water Science and Technology* **81**: 1588–1596. DOI: http://dx.doi.org/10.2166/wst.2020.032

Roozeboom, KL, Prasad, PVV. 2019. Sorghum growth and development, in *Sorghum*. Wiley & Sons: 155–172. DOI: http://dx.doi.org/10.2134/agronmonogr58.c8.

Salas, E, Reyes, DL, McDaniel, SH. 2018. The science of teamwork: Progress, reflections, and the road ahead. *American Psychologist* **73**: 593–600. DOI: http://dx.doi.org/10.1037/amp0000334.

Saltelli, A, Benini, L, Funtowicz, S, Giampietro, M, Kaiser, M, Reinert, E, van der Sluijs, JP. 2020. The technique is never neutral. How methodological choices condition the generation of narratives for sustainability. *Environmental Science & Policy* **106**: 87–98. DOI: http://dx.doi.org/10.1016/j.envsci.2020.01.008.

Saltelli, A, Funtowicz, S. 2017. What is science's crisis really about? *Futures, Post-Normal Science in Practice* **91**: 5–11. DOI: http://dx.doi.org/10.1016/j.futures.2017.05.010.

Scheffer, M, Baveco, JM, DeAngelis, DL, Rose, KA, van Nes, EH. 1995. Super-individuals, a simple solution for modelling large populations on an individual basis. *Ecological Modelling* **80**: 161–170. DOI: http://dx.doi.org/10.1016/0304-3800(94)00055-M.

Singh, BU, Padmaja, PG, Seetharama, N. 2004. Biology and management of the sugarcane aphid, *Melanaphis sacchari* (Zehntner) (Homoptera: Aphididae), in sorghum: A review. *Crop Protection* **23**: 739–755. DOI: http://dx.doi.org/10.1016/j.cropro.2004.01.004.

Stein, AF, Draxler, RR, Rolph, GD, Stunder, BJB, Cohen, MD, Ngan, F. 2015. NOAA's HYSPLIT atmospheric transport and dispersion modeling system. *Bulletin of the American Meteorological Society* **96**: 2059–2077. DOI: http://dx.doi.org/10.1175/BAMS-D-14-00110.1.

Sterling, EJ, Zellner, M, Jenni, KE, Leong, K, Glynn, PD, BenDor, TK, Bommel, P, Hubacek, K, Jetter, AJ, Jordan, R, Olabisi, LS, Paolisso, M, Gray, S. 2019. Try, try again: Lessons learned from success and failure in participatory modeling. *Elementa: Science of the Anthropocene* **7**: 9. DOI: http://dx.doi.org/10.1525/elementa.347.

Strode, DE, Huff, SL, Hope, B, Link, S. 2012. Coordination in co-located agile software development projects. *Journal of Systems and Software, Special Issue: Agile Development* **85**: 1222–1238. DOI: http://dx.doi.org/10.1016/j.jss.2012.02.017.

Thiele, JC, Kurth, W, Grimm, V. 2014. Facilitating Parameter estimation and sensitivity analysis of agent-based models: A cookbook using NetLogo and R. *JASSS* **17**: 11.

Thomas, J. 2012. Power and illusion of self-organizing teams. Presented at the PMI® Global Congress 2012, PA: Project Management Institute, North America, Vancouver, BC, Canada. Newtown Square.

Thomas, J. McDonagh, D. 2013. Shared language: Towards more effective communication. *Australasian Medical Journal* **6**: 46–54. DOI: http://dx.doi.org/10.4066/AMJ.2013.1596.

Ticehurst, J, Curtis, A. 2016. *Improving on-farm water use efficiency amongst irrigators in the Campaspe River system and the Rochester Irrigation Area*. Wagga Wagga, Australia: EH Graham Centre for Agricultural Innovation, Charles Sturt University.

Ticehurst, J, Curtis, A. 2017. *Preliminary assessment of conjunctive use opportunities in the Murray-Darling Basin: A case study in the lower Campaspe catchment*. Canberra, Australia: Murray-Darling Basin Authority; Adelaide, Australia: The National Centre for Groundwater Research and Training.

Tjosvold, D, Yu, Z. 2004. Goal Interdependence and applying abilities for team in-role and extra-role performance in China. *Group Dynamics: Theory, Research, and Practice* **8**: 98–111. DOI: http://dx.doi.org/10.1037/1089-2699.8.2.98.

USDA-NASS. 2010. *Field crops usual planting and harvesting dates, Agricultural Handbook Number 628*. Washington, DC: USDA.

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

van Dongen, MJ, van Bokhoven, MA, Goossens, WN, Daniëls, R, van der Weijden, T, Beurskens, A. 2018. Development of a customizable programme for improving interprofessional team meetings: An action research approach. *International Journal of Integrated Care* **18**: 8. DOI: http://dx.doi.org/10.5334/ijic.3076.

van Gool, FW, Bierbooms, JJ, Bongers, IM, Janssen, RT. 2019. Co-creating a program for teams to maintain and reflect on their flexibility. *The International Journal of Health Planning and Management* **34**: e1937–e1947. DOI: http://dx.doi.org/10.1002/hpm.2855.

van Rensburg, NJ. 1973. Notes on the occurrence and biology of the sorghum aphid in South Africa. *Journal of the Entomological Society of Southern Africa* **36**: 293–298.

Verweij, PJFM, Knapen, MJR, de Winter, WP, Wien, JJF, te Roller, JA, Sieber, S, Jansen, JML. 2010. An IT perspective on integrated environmental modelling: The SIAT case. *Ecological Modelling* **221:** 2167–2176. DOI: https://doi.org/10.1016/j.ecolmodel.2010.01.006.

Voinov, A, Shugart, HH. 2013. "Integronsters," integral and integrated modeling. *Environmental Modelling and Software* **39**: 149–158. DOI: http://dx.doi.org/10.1016/j.envsoft.2012.05.014.

Voznesenskaya, T, Krasnov, F, Yavorsky, R, Chesnokova, P. 2019. Modeling self-organizing teams in a research environment. *Business Informatics* **13**: 7–17. DOI: http://dx.doi.org/10.17323/1998-0663.2019.2.7.17.

Walsh, Z, Böhme, J, Wamsler, C. 2020. Towards a relational paradigm in sustainability research, practice, and education. *Ambio*. DOI: http://dx.doi.org/10.1007/s13280-020-01322-y.

Wang, H-H, Grant, WE. 2021. Reflections of two systems ecologists on modelling coupled human and natural (socio-ecological, socio-environmental) systems. *Ecological Modelling* **440**. DOI: http://dx.doi.org/10.1016/j.ecolmodel.2020.109403.

Wang, H-H, Grant, WE, Elliott, NC, Brewer, MJ, Koralewski, TE, Westbrook, JK, Alves, TM, Sword, GA. 2019. Integrated modelling of the life cycle and aeroecology of wind-borne pests in temporally-variable spatially-heterogeneous environment. *Ecological Modelling* **399**: 23–38. DOI: http://dx.doi.org/10.1016/j.ecolmodel.2019.02.014.

Wang, H-H, Grant, WE, Koralewski, TE, Elliott, NC, Brewer, MJ, Westbrook, JK. 2020a. *Where do all the aphids go? A series of thought experiments within the context of area-wide pest management*. Agricultural Systems.

Wang, H-H, Grant, WE, Teague, R. 2020b. Modeling rangelands as spatially-explicit complex adaptive systems. *Journal of Environmental Management* **269**: 110762. DOI: http://dx.doi.org/10.1016/j.jenvman.2020.110762.

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

Wilensky, U. 1999. *NetLogo (and NetLogo user manual)*. Evanston, IL: Center for Connected Learning and Computer-Based Modeling, Northwestern University.

Wilson, C, Kriegler, E, van Vuuren, DP, Guivarch, C, Frame, D, Krey, V, Osborn, TJ, Schwanitz, VJ, Thompson, EL. 2017. Evaluating process-based integrated assessment models of climate change mitigation. Available at http://pure.iiasa.ac.at/id/eprint/14502/. Accessed 16 December 2019.

Wilson, DS, Wilson, EO. 2007. Rethinking the theoretical foundation of sociobiology. *The Quarterly Review of Biology* **82**: 327–348. DOI: http://dx.doi.org/10.1086/522809.

Woolley, A, Malone, T, 2011. What makes a team smarter? More women. *Harvard Business Review* **89**: 32–33.

Xie, Y, Crosbie, R, Simmons, CT, Cook, PG, Zhang, L. 2019. Uncertainty assessment of spatial-scale groundwater recharge estimated from unsaturated flow modelling. *Hydrogeology Journal* **27**: 379–393. DOI: http://dx.doi.org/10.1007/s10040-018-1840-0.

Zare, F, Guillaume, JHA, ElSawah, S, Croke, B, Fu, B, Iwanaga, T, Merritt, W, Partington, D, Ticehurst, J, Jakeman, AJ. 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. DOI: http://dx.doi.org/10.1016/j.jhydrol.2021.126070.

Zdun, U, Tran, H, Capilla, R, Zimmerman, O. 2014. Sustainable architectural design decisions. *InfoQ*. Available at https://www.infoq.com/articles/sustainable-architectural-design-decisions/. Accessed 14 December 2020.

How to cite this article: Iwanaga, T, Wang, H-H, Koralewski, TE, Grant, WE, Jakeman, AJ, Little, JC. 2021. Toward a complete interdisciplinary treatment of scale: Reflexive lessons from socioenvironmental systems modeling. *Elementa: Science of the Anthropocene* 9(1). DOI: https://doi.org/10.1525/elementa.2020.00182

Domain Editor-in-Chief: Steven Allison, University of California, Irvine, CA, USA

Associate Editor: Erika Zavaleta, University of California, Santa Cruz, CA, USA

Knowledge Domain: Ecology and Earth Systems

Published: June 11, 2021 **Accepted:** April 30, 2021 **Submitted:** December 22, 2020

ELEMENTA Science of the Anthropocene

Elem Sci Anth is a peer-reviewed open access journal published by University of California Press.

OPEN ACCESS