Global Environmental Change 50 (2018) 1–14

Contents lists available at ScienceDirect

Global Environmental Change

journal homepage: www.elsevier.com/locate/gloenvcha

Closing global knowledge gaps: Producing generalized knowledge from case studies of social-ecological systems

Nicholas R. Magliocca[a,*], Erle C. Ellis[b], Ginger R.H. Allington[c], Ariane de Bremond[d,e], Jampel Dell'Angelo[f], Ole Mertz[g], Peter Messerli[e], Patrick Meyfroidt[h,i], Ralf Seppelt[j,k,l], Peter H. Verburg[m]

[a] Department of Geography, University of Alabama, Tuscaloosa, AL, USA
[b] Department of Geography and Environmental Systems, University of Maryland, Baltimore County, Baltimore, MD, USA
[c] Department of Geography, George Washington University, Washington, DC, USA
[d] Geographical Sciences Department, University of Maryland, College Park, MD, USA
[e] Centre for Development and Environment (CDE) and Institute of Geography, University of Bern, Bern, Switzerland
[f] Department of Environmental Policy Analysis, Institute for Environmental Studies, VU University Amsterdam, Amsterdam, The Netherlands
[g] Department of Geosciences and Natural Resource Management, University of Copenhagen, Øster Voldgade 10, 1350 Copenhagen, Denmark
[h] Georges Lemaître Center for Earth and Climate Research, Earth and Life Institute, Université Catholique de Louvain, Louvain-la-Neuve, Belgium
[i] F.R.S.-FNRS, Brussels, Belgium
[j] UFZ – Helmholtz Centre for Environmental Research, Department of Computational Landscape Ecology, 04318, Leipzig, Germany
[k] Institute of Geoscience & Geography, Martin-Luther-University Halle-Wittenberg, 06099, Halle (Saale), Germany
[l] iDiv – German Centre for Integrative Biodiversity Research, Halle-Jena-Leipzig, 04103, Leipzig, Germany
[m] Department of Environmental Geography at the Institute for Environmental Studies, VU University Amsterdam, Amsterdam, The Netherlands

ARTICLE INFO

Keywords:
Synthesis
Meta-analysis
Causal inference
Land-use change
Biodiversity

ABSTRACT

Concerns over rapid widespread changes in social-ecological systems and their consequences for biodiversity, ecosystem functioning, food security, and human livelihoods are driving demands for globally comprehensive knowledge to support decision-making and policy development. Claims of regional or global knowledge about the patterns, causes, and significance of changes in social-ecological systems, or 'generalized knowledge claims' (GKCs), are generally produced by synthesis of evidence compiled from local and regional case study observations. GKCs now constitute a wide and varied body of research, yet they are also increasingly contested based on disagreements about their geographic, temporal, and/or thematic validity. There are no accepted guidelines for detecting biases or logical gaps between GKC's and the evidence used to produce them. Here, we propose a typology of GKCs based on their evidence base and the process by which they are produced. The typology is structured by three dimensions: i) the prior state of knowledge about the phenomenon of interest; ii) the logic of generalization underlying the claim; and iii) the methodology for generalization. From this typology, we propose a standardized approach to assess the quality and commensurability of these dimensions for any given GKC, and their ability to produce robust and transparent knowledge based on constituent evidence. We then apply this approach to evaluate two contested GKCs – addressing global biodiversity and large-scale land acquisitions – and in doing so demonstrate a coherent approach to assessing and evaluating the scope and validity of GKCs. With this approach, GKCs can be produced and applied with greater transparency and accuracy, advancing the goal of actionable science on social-ecological systems.

1. Introduction

Contemporary rates and scales of social and environmental change are unprecedented (Cleland et al., 2007; Ellis et al., 2016; Steffen et al., 2011; Seppelt et al., 2014). Scientific knowledge about the causes and/or consequences of these changes typically comes from one of two sources: spatially extensive and coarse resolution datasets, or case study or field observations. Large-scale datasets are useful for understanding broad patterns in space and time (e.g. Hansen et al., 2013), but can be insufficient to causally link social and ecological processes to observable conditions (Rindfuss et al., 2004). Observations from case studies enable in-depth understanding of causal relationships among

* Corresponding author at: Department of Geography, University of Alabama, Box 870322, Tuscaloosa, Alabama 35487-0322, USA.
E-mail address: nrmagliocca@ua.edu (N.R. Magliocca).

https://doi.org/10.1016/j.gloenvcha.2018.03.003
Received 24 August 2017; Received in revised form 26 February 2018; Accepted 5 March 2018

social-ecological processes, but these tend to be strongly context-dependent, necessitating large numbers of studies from different places in order to infer broader patterns. Additional challenges to case study synthesis include biases in the selection of sites for case study research, variety of disciplines that contribute to social-ecological systems (SES) research, and mixed or inconsistent methodologies and data sources (Hersperger and Bürgi, 2009; Keys and McConnell, 2005; Magliocca et al., 2015; van Vliet et al., 2015a). Efforts to combine large-scale datasets and case study observations also suffer from epistemological challenges when different levels of explanations (i.e., causal effects versus mechanisms; Meyfroidt, 2016) are used. All of these issues limit our ability to generalize knowledge beyond the specific case, at a time when such extrapolation is increasingly needed to inform policy.

Given the widespread influences of global environmental and economic change, interest is growing in connecting localized social-ecological system changes to related trends, causes, or consequences observed at other locations and/or broader scales (i.e., regional or global) (Magliocca et al., 2015; van Vliet et al., 2015b). Producing such knowledge, or 'generalized knowledge claims' (GKCs), requires expanding observations and inferences from individual case studies beyond the spatial and/or temporal boundaries from which they were originally produced (An et al., 2005; Magliocca et al., 2015; Meyfroidt, 2016; Valbuena et al., 2010; Verburg et al., 2009). Many efforts have been made to synthesize case study observations of social-ecological system change to support assessment and policy, and the GKCs produced in the process vary in type and scope. For example, the intergovernmental panels on climate change (IPCC) and biodiversity and ecosystem services (IPBES) produce GKCs based on a process of expert consultation and thorough expert review. The Collaboration for Environmental Evidence[1] provides rigorous guidelines for conducting and reporting systematic reviews, and a dedicated, peer-reviewed journal for disseminating results. Other analytical efforts typically rely on data synthesis and meta-analytical approaches to extend or scale-up case-based knowledge to broader scales and more general inferences (Magliocca et al., 2015; van Vliet et al., 2015b). Some GKCs are strongly quantitative, such as meta-analysis of effect sizes, while others are a mix or primarily qualitative, such as archetype analysis. Some GKCs attempt to posit causal explanations of general patterns and processes of change as the basis for theory (e.g., Oberlack et al., 2016; Dell'Angelo et al., 2017a), whereas others establish the empirical state of knowledge for a given phenomenon (e.g., Geist and Lambin, 2002; Dell'Angelo and Rulli et al., 2017b). GKCs may also be regionally or narrowly focused on a specific system or set of conditions (e.g., illegal deforestation in Mesoamerica, or only under stable, democratic political systems) or truly global in scope (e.g., forest loss and gain). Any given GKC may have all or some of these characteristics in various combinations.

However, GKCs are also increasingly contested based on disagreements about geographic, temporal, and/or thematic validity (e.g., Bäckstrand and Lövbrand, 2006; Breu et al., 2016; Gonzalez et al., 2016; Rasmussen et al., 2016). While there are shared approaches and accepted guidelines for conducting synthesis (e.g., Gerstner et al., 2017; Koricheva et al., 2013; Liberati et al., 2009), there is no such guidance for how best to communicate: i) the GKC relative to the existing state of knowledge; ii) the logic used to generalized from the scale of case study observations to broader contexts; and iii) the analytical approach used for synthesizing case study observations. Even major institutionalized efforts by IPCC and IPBES, while following thorough procedures to reach consensus among experts (and policy), lack a thorough documentation of the nature of the knowledge claims. Without transparent or consistent reporting of how these components of GKCs align, it is unclear for what conditions different types of GKCs can be expected to be valid, or what possible biases or logical gaps exist between the GKCs and the case study evidence used to produce them (Gonzalez et al., 2016; Martin et al., 2012). Therefore, the goal of this paper is to develop a typology of GKCs and establish a standard approach that aids in assessing and/or producing robust and transparent GKCs based on collections of empirical case studies. This standardization will strengthen future synthesis and generalization efforts (e.g., meta-analyses) and the peer-review process by providing a common vernacular for presenting GKCs, and making explicit often implicit logic of how researchers relate case study observations to generalized knowledge. Ultimately, we hope this will enable reliable assessment and utilization of GKCs to advance research and inform policy.

The typology and evaluation criteria presented in this paper came out of discussions among participants (the authors of this paper) at a working-group convened at the National Socio-Environmental Synthesis Center in Annapolis, MD in June 2016. Workshop participants were assembled specifically to be broad in terms of disciplines and paradigmatic stances represented, specifically including critical/constructivists social scientists to ensure that our discussions captured a broad range of epistemologies and forms of knowledge. The original aim of the working group was to identify the knowledge gaps in three global phenomena research domains: deforestation; large-scale land acquisitions (LSLAs); and biodiversity changes. Specifically, participants were tasked with assessing gaps in the geographic coverage, thematic content, and knowledge production processes (methods, data, and semantics) of recently published and contested knowledge claims in these domains. The intended goal of the workshop was to develop a rigorous and widely applicable meta-study framework for assessing knowledge gaps in this kind of work. We quickly realized that in order to accomplish that goal, we would need to rigorously assess the knowledge claims embedded within current studies, but that at the time there were no guidelines for how to approach such an assessment. Facilitated discussions in the following days resulted in the formulation of the typology and approach we present herein. Below, we provide details on our process and outcomes, including the typology we developed, a recommended protocol for developing a GKC, a rubric for assessing existing GKCs, and application and validation of the rubric to two existing studies. The participants in these discussions are all land-change scientists from academic institutions in the United States and Western Europe with experience conducting meta-studies in one of the three focal research domains.

2. Defining generalized knowledge claims (GKCs)

2.1. Generalization, synthesis, and conditionality

In order to engage in a discussion of how generalized claims are produced, it is important to define our use of three key concepts: generalization, synthesis, and conditionality. In particular, we want to emphasize the difference between generalization, as a research endeavor, and synthesis as the methodological approaches to achieve generalization. Understanding this distinction is a necessary step in order to ensure that the chosen synthesis methods support the logics of generalization used (Magliocca et al., 2015; Steinberg, 2015; van Vliet et al., 2015b).

Generalization is a mostly inductive logical argument for extending an empirical or conceptual relationship deduced from a set of case study observations beyond the specific contexts in which the relationship was first observed (Steinberg, 2015). Researchers generalize by connecting trends in, relationships among, or explanations for observations of change in specific contexts to more general patterns, causes, and/or consequences, thereby moving beyond case-specific explanations and the 'variance of place' (Turner et al. 2007). Although GKCs addressing global-scale phenomena are becoming more common, the generalized pattern or process does not have to be applicable globally in the geographic sense, but rather applicable beyond the specific contexts under which its constituent observations were acquired.

[1] http://www.environmentalevidence.org/

Synthesis involves a broad family of methods for knowledge integration that draw upon multiple sources of data, explanations, and analytical techniques (http://sesync.org/glossary/). Synthesis methods differ widely in their sampling rigor, use of quantitative and qualitative data, and inferential strength. Meta-study (of which meta-analysis is a special type) is a synthesis method for distilling the findings of many narrowly focused analyses (i.e., 'cases') to produce knowledge that is more generally applicable than what can be derived from a single case (Magliocca et al., 2015). Applying a synthesis method necessarily formalizes a GKC and results in a bounded range of conditions, or '*conditionality*', under which the generalization is expected to hold true. The conditionality of a claim comes from the intersection between how the authors intend for the GKC to be applied and any statistical or methodological limitations imposed by how the empirical evidence were gathered, interpreted, and synthesized. But to date, there are no agreed-upon terminology, structured workflows, or rubrics to guide the development or assessment of such endeavors. Thus, the landscape of GKCs is conspicuously heterogeneous with various combinations of sample sizes, synthesis methods, and generalization logics used by social-ecological researchers.

2.2. Dimensions of generalized knowledge claims

The first step in creating a consistent approach to developing or assessing GKCs is to identify their key attributes. We classify GKCs into types based on three essential dimensions:

(1) the prior *state of knowledge* about the phenomenon of interest;
(2) the *logic of generalization* underlying the GKC; and
(3) the *methodology* for producing a GKC.

Once classified into types, the production, reporting, and evaluation of GKCs can be standardized with a structured workflow, similar to standardization efforts in other disciplines (e.g., PRISMA statement in medical research, Liberati et al., 2009; Collaboration for Environmental Evidence).

We developed a standardized approach around these three dimensions to systematically characterize, assess, and compare diverse GKCs (summarized in Table 1). The first dimension relates to the state of knowledge about the study phenomenon. This dimension describes whether a claim is building-on, extending, or refuting current knowledge, because GKCs are made about phenomena for which there are varying levels of existing knowledge. This entails being explicit about the existing knowledge base's level of empirical evidence, its interpretation, and causality established from those interpretations that the GKC is reinforcing or revising, which goes beyond what is typically reported in a literature review. The second dimension describes the generalization logic underlying the GKC, which entails the level of causal inference posited by the GKC and its stated conditionality. This dimension describes how any given case study observation relates to broader conditions under which similar observations are expected, and how broader-scale inference from a specific context relies on or departs from existing knowledge. The third dimension characterizes the methodology for producing a GKC. This includes the quantitative and/or qualitative nature of case study observations, geographic scale and extent of that data, and synthesis method used. This dimension allows assessment of whether the type and scope of the GKC is consistent with the empirical evidence available to build it. Each of these dimensions and their associated sub-dimensions are discussed in detail with examples from the social-ecological literature below (Table 2).

Table 1
Summary of the three dimensions of generalized knowledge claims (GKC) and their subcategories.

1: Existing Knowledge Base
Prior Empirical Studies: (i) Limited; (ii) Comprehensive
Existing Theory: (i) Scant or many competing theories; (ii) One dominant theory
2: Logic of Generalization of the Claim
Causality: (i) Descriptive; (ii) Causal effects; (iii) Causal mechanism
Conditionality: (i) Narrow; (ii) Middle-range; (iii) Universal
3: Methodology for Producing Claim
Empirical Basis: (i) Qualitative data only; (ii) Quantitative data only; (iii) Mixed data types
Number of Observations: (i) Small-N; (ii) Large-N
Spatial Scale of Observation: (i) Local; (ii) Sub-national to national; (iii) Regional to global
Geographic Representativeness: (i) Limited; (ii) Regional; (iii) Global

2.2.1. Dimension 1: existing knowledge base

2.2.1.1. Existing empirical studies. The standard of evidence for making GKCs depends on the state of existing knowledge about the phenomenon of study, which can vary in its comprehensiveness and coherence. Comprehensiveness is simply the volume of empirical studies that have been accumulated and can inform analysis of the given topic. The more studies available to the researcher, the greater the potential empirical evidence that can be brought to bear on a GKC. However, an abundance of empirical studies does not necessarily lead to a coherent knowledge base from which to generalize. While the number of studies used to make a GKC is important (discussed in Section 2.2.3), the extent to which empirical evidence has been systematically organized and interpreted determines the state of the existing knowledge base. This in turn determines the consistency and validity of the GKC relative to existing evidence. Thus, GKCs about emergent phenomena which have been identified through statistical evidence, individual case studies, or monitoring systems may be restrained to phenomenological or frequentist arguments describing the state of knowledge. In contrast, GKCs about well-studied phenomena may advance more strongly on causal or theoretical arguments.

For example, an observation of changing phenology in Wisconsin (Bradley et al., 1999) is readily accepted as evidence of more general causal effects of climate change occurring around the planet (Parmesan and Yohe, 2003). This is because climate science, at least in this regard, has amassed a multitude of such observations, articulated a causal mechanism linking climate change and phenology, and formed and tested predictions of phenology changes in many locations and climatic conditions. The result is a well-developed, shared, and community-refined knowledge base from which a standard logic of generalization between each individual observation and the larger class of phenomena has been established (Steinberg, 2015). Importantly, the strength of evidence needed to make new claims based on this generalization logic is relatively low, compared to that of an emerging phenomenon (e.g., transnational LSLAs) for which fewer observations have been made and/or hypotheses tested. Thus, the scope of GKCs depends on the existing knowledge base and logic of generalization for a given object of study.

2.2.1.2. Existing theory. The theoretical landscape within which a GKC is made also influences its possible goal and the process of formulating the claim (Table 3). When a given topic is dominated by a well consolidated set of theories, a new generalization may serve to expand or restrict the range of conditions under which this theory applies: testing if the theory still holds under a new or broader range of conditions; taking into account variation in a previously not measured variable; or identifying a certain range of conditions under which the validity of the dominant theory is constrained. It is also possible, though less frequent, to construct a new generalization claim that refutes a dominant theory. In this case, the strength of evidence supporting such "overturning" GKCs ought to at least meet, or better, exceed that supporting the prior theory, including a comprehensive empirical knowledge base meeting or extending beyond the contexts (e.g., conditionality) under which the original theory claims to be applicable.

Table 2
Examples of the dimensions of generalized knowledge claims (GKC) from social-ecological systems literature.

Dimension	Existing Knowledge Base				Logic of Generalization of Claim		
Sub-Dimensions	Existing Empirical Studies		Existing Theory		Causality		
Categories	Limited	Comprehensive	Scant or Many Competing Theories	One Dominant Theory[a]	Descriptive	Causal Effects	Causal Mechanisms
Description	Emerging phenomenon	Established or well-documented phenomenon	Theory has yet to emerge or many competing theories have emerged but have not been comprehensively tested	One or a few theories have been empirically verified within posited conditions of applicability	Phenomenological descriptions of the occurrence and/or relevant importance of an observed outcome (Meyfroidt, 2016)	Frequentist analysis of a set of factors linked to an outcome (Meyfroidt, 2016)	Chain of causal mechanisms linking a set of factors to an outcome (Meyfroidt, 2016)
Examples	Large-scale land acquisitions	Tropical deforestation	Formulating hypotheses on the main factors affecting pathways of commodity crop expansion in tropical landscapes (Meyfroidt et al., 2014)	Induced intensification (Turner and Ali, 1996)	Global forest change trends derived from remote sensing time series analysis (Hansen et al., 2013)	Meta-analysis synthesizing the frequency of different underlying drivers of agricultural land use change in Europe (van Vliet et al., 2015b)	Description of the archetypical pathways through which large-scale land acquisitions can impact livelihoods (Oberlack et al., 2016)

Dimension	Logic of Generalization of Claim			Methodology for Producing Claim		
Sub-Dimensions	Conditionality			Empirical Basis		
Categories	Narrow	Medium	Universal	Qualitative Data Only	Quantitative Data Only	Mixed Data Types
Description	Generalization applied to limited contextual conditions (i.e., within-system; Steinberg, 2015)	Generalization applied to conditions shared by similar systems (i.e., cross-system resonance groups; Steinberg, 2015)	Generalization applied to a broad range of conditions across systems	Analysis of descriptive data with comparative methods (Ragin, 1987; Young et al., 2006)	Analysis of numerical data with statistical techniques	Using both quantitative and qualitative data and/or analytical methods
Examples	Study of the conditions under which perceptions of environmental change have contributed to forest transition in four village case studies in northern Vietnam (Meyfroidt, 2013); Modeling local land-use and livelihood responses to global market signals at six case study sites (Magliocca, 2015)	Two types of forest transitions. *Condition*: In Central America (Grau and Aide, 2008); Identification of major tropical deforestation drivers. *Condition*: Tropical forests, over 1960-80s and 1980-2000s (Rudel et al., 2009); Important behavioral variations in economic decisions across cultural contexts. *Condition*: In small-scale societies (Henrich et al., 2005)	No net change in local plant biodiversity globally over time (Vellend et al., 2013)	Qualitative comparative analysis of case studies of land grabbing (Dell'Angelo et al., 2017a); historical process tracing of China and Brazil energy policy (Hochstetler and Kostka, 2015)	Meta-analysis of land use intensity effect sizes on plant species richness (Gerstner et al., 2014)	Global meta-study of agricultural change case studies (Keys and McConnell, 2005)

Dimension	Methodology for Producing Claim							
Sub-Dimensions	Number of Observations		Spatial Scale of Observations			Geographic Representativeness		
Categories	Small-N	Large-N	Local	Sub-National to National	Regional to Global	Limited	Regional	Global
Description	Roughly 2 to 30 'cases' and number of variables is large relative to number of observations (Rihoux and Ragin,2009)	Roughly 30 or more 'cases' and number of variables is small relative to number of observations.	Field plots; villages; community surveys	Administrative Units	Biomes; Continents	Variation in observational conditions is very limited relative to variation observed across the expected geographical extent of the study phenomenon	Variation in observational conditions is incomplete relative to variation observed across the expected geographical extent of the study phenomenon	Variation in observational conditions is equal to variation observed across the expected geographical extent of the study phenomenon

(*continued on next page*)

Table 2 (*continued*)

Dimension	Methodology for Producing Claim							
Examples	Cross-site comparison of land-use change around protected areas (e.g. DeFries et al., 2007)	Meta-analysis of effect sizes of CO_2 treatments on food crops (e.g., Taub et al., 2008	Global meta-analysis of land use effects on local biodiversity (e.g., Newbold et al., 2015)	Cross-national comparison of virtual water use (Tamea et al., 2016)	Global meta-analysis of regional variations in soil phosphorus (MacDonald et al., 2012);	Cross-site comparisons, field plots from a localized area, single remote sensing image, or community survey	Regional synthesis of local case studies (e.g., Grau and Aide, 2008; van Vliet et al., 2015a); Sampling bias towards developed world regions (e.g., Gerstner et al., 2014)	Corrected regional bias in global meta-analysis of swidden agricultural change (van Vliet et al., 2012)

[a] Theories are dominant or 'accepted' in the sense that they are applied as explanatory frameworks and have not yet been falsified under conditions in which they are posited to be valid.

When a given question corresponds to a scant theoretical landscape (e.g., 'virtual water' when it was first proposed; Allan, 2003), a GKC may serve as a benchmark for future tests. Theories about phenomena for which limited consistent empirical studies exist can only be tested or refined rather than confirmed, because it is likely that the empirical observations do not yet sufficiently cover the range of conditions in which the phenomenon might be expected to occur. Where different competing theories are struggling, a GKC may test one theory against another in order to determine which one is valid or how they can be valid under different ranges of conditions.

2.2.2. Dimension 2: logic of generalization of the claim

2.2.2.1. Causality. The type of inference presented in a GKC depends on the extent to which the analysis can establish causality between the outcomes of interest and hypothesized related factors. Purely descriptive or phenomenological GKCs, such as documenting the rates of forest cover change globally (e.g., Hansen et al., 2013), establish the occurrence and/or relevant importance of an observed outcome, but do not make assertions or inferences about the factors causing such outcomes. In contrast, GKCs aiming to establish causal relationships between broad-scale outcomes and variations in factors observed from case studies will not be robust unless these are based on some form of causal inference. Following Meyfroidt (2016), causal inference may come in one of two forms. A *causal effect* is when a change in outcome Y is brought about by the change in a factor X. The proposed effect of factor X is underpinned by a hypothesis originating from the related science fields, and tested with a single or collection of measurements that describe variations in Y. For example, meta-analyses that quantify the frequency and/or correlation of specific factors (e.g., proximate causes and underlying driving forces) with outcomes of interest (e.g., deforestation) are claiming causal effects (Geist and Lambin, 2002). The strength of the causal effect depends on establishing the necessity and/or sufficiency of causal factors. A *causal mechanism* extends the assertion of causal effects by providing an explanation of *how* a causal factor or combination of causal factors produces an effect. A chain of interacting 'instances' is assembled as a causal mechanism to describe a cause and/or consequence (Meyfroidt, 2016). For example, linking a series of specific factors and processes, such as poor law enforcement, asymmetric access to capital, and elite capture, to specific outcomes, such as shifts in livelihood strategies, asserts a causal mechanism for explaining livelihood changes in the context of LSLAs (Oberlack et al., 2016). Causal inferences can build on different logics, i.e. deductive approaches deriving causal inferences from already established general principles or assumptions that are tested across the range of observations, or inductive generalizations of a set of abductive explanations of the different cases studied (Walton, 2005), or combinations of these.

There are fundamental differences in the explanations established using inference based on causal effects and mechanisms, and thus in the logic of generalization used to make a GKC. GKCs based on case studies that explain the effects of one or more causal factor(s) on outcomes in a specific context can describe general patterns of change across sampled contexts and conditions under which case-specific effects deviate from general patterns. GKCs based on case studies that propose how and why causal factors affected outcomes in a specific context (i.e., causal mechanism) are additionally capable of making predictions about the most likely conditions under which a causal mechanism will operate, outcomes vary, and when and where a GKC will be applicable. Thus, whether GKCs remain descriptive or articulate 'grand theories' of social-ecological changes is contingent on the level of causal explanation achieved in the case study observations on which GKCs are based.

2.2.2.2. Conditionality. The second aspect which defines the logic of generalization underlying a given GKC is its level of conditionality, or the "range" or "contextual factors", under which the assertion is claimed to be valid. In principle, a GKC is reliable and robust only

Table 3
Goals of generalization depending on the existing knowledge base. Intersecting the states of current empirical data and theories can help define the appropriate goal for the GKC one is attempting to make.

		Existing Theory	
		Scant / Unsettled	One dominant
Existing Empirical Studies	Limited	Identifying initial knowledge evidence and knowledge gaps, formulating hypotheses as basis for theories and additional case study collection	Questioning/refuting dominant theory, suggesting alternatives
	Comprehensive	Testing alternative theories, identifying conditions under which they apply	Confirming theory, refining subtheories, sometimes refuting a theory

within the conditions covered by the observations used to produce it. These conditions relate to geographic space, to the extent that things nearer to one another are more similar than those that are far apart. But the scope of applicability of a GKC does not only apply to its geographic characteristics but to any variable that can be used to describe the specific conditions of its constituent case study observations. Environmental variables, such as the types of climate, biomes, and soils, and human variables, such as political or economic systems, social structures, and cultural traits, can all define the range of conditions under which the GKC is expected to be valid. The conditions may also be temporal. In SES and especially in the context of the Anthropocene, human agency and historicity, and other factors such as global environmental change modify the processes and dynamics that are studied. Some GKCs can be valid only over a certain time period or under certain historic conditions, and the manifestation of new phenomena over time changes the universe being studied, as manifested by the emergence of private-led governance of natural resources, or the dynamics of LSLAs, for example (Meyfroidt, 2016). While characteristics of a study site or collection of case study observations are often reported, they are seldom reported with an eye towards synthesis such that a level of precision and/or quantification is provided that enables testing the assertions of a GKC beyond the conditions in which its constituent cases were observed. Extrapolating current or past GKC to the future or using synthesis to inform model development thus requires careful consideration of how these extrapolations rely on assumptions about the stationarity of certain processes in the future.

The scope of spatial and temporal conditionality – relatively narrow or universal – depends on the limitations of empirical data collection and sampling ability and/or (theoretical) understanding of the phenomenon of study. Limitations imposed by data sampling are straightforward – the more comprehensive the sampling, the broader range of conditions to which the GKC can apply. If a study makes a GKC beyond the range of conditions that are present in the observations used to derive it, such claims cannot be considered robust unless evidence is presented demonstrating that the claim is not influenced by variations in the conditions or contextual factors beyond those observed.

Together, the type of causality and conditionality posited by a GKC set the bounds for where the claim falls along the spectrum from descriptive and narrow conditionality to theoretical and universal conditionality.

2.2.3. Dimension 3: methodology for producing the claim

2.2.3.1. Empirical basis. The type of data contained in case study observations will influence the comparability of data across cases and the choice of synthesis method (Magliocca et al., 2015). *Qualitative* data typically describes the directionality and/or relative strength of the link between two variables, the general property of a phenomenon (e.g., positive or negative relation between a factor and an outcome, increasing prevalence of a given phenomenon over time), or the possible ways to structure causal linkages between different variables. Such means of description/measurement may not be consistent across observers and/or contexts, which requires standardization or harmonization in the form of coding before comparisons can be conducted. Generating standardized codes when using qualitative and/or mixed data can increase comparability and statistical power (Rudel, 2008). *Quantitative* data describes the magnitude or degree of a factor's effect on an outcome (e.g., 30% vs. 60% of cases show an increase vs. decrease in forest cover, respectively). Generally, the means of quantification is more consistent across observers (e.g., statistical), or, provided a conversion of measurement units, allows for direct comparison. Standardization of quantitative measurements may enable the pooling of data and increased statistical power (e.g., meta-analysis of effect sizes; Rudel, 2008).

Whichever type of data is used to make a GKC, the approaches and methods used to synthesize qualitative versus quantitative data are different (Magliocca et al., 2015). Both qualitative and quantitative types of data can lead to explanations of causal effects (e.g., descriptive, place-based narratives and regression analyses, respectively) or causal mechanisms (e.g., historical process tracing or configurational causal analysis and time-series analysis; George and Bennett, 2005). However, the selection of synthesis method – meta-analysis of effect sizes versus case-oriented meta-analysis – will follow from the type and standardization challenges of data on which the GKC is based (Magliocca et al., 2015).

2.2.3.2. Number of observations. Synthesis researchers are often faced with a trade-off between causal inference and conditionality. Generalization from single or small-N studies, using methods like historic process tracing (George and Bennett, 2005) or case-oriented meta-analysis (Ebbinghaus, 2005; Ragin, 1987; Steinberg, 2015), reduces the loss of information during synthesis and can enable the researcher to gain a deeper understanding of the context and structural relationships between explanatory factors and outcomes. This mode of generalization lends itself to configurational analytical methods, such as qualitative comparative analysis, which may provide a more direct path to causal inference based on mechanisms. However, small-N generalization may suffer from selection or geographic bias and be analytically constrained by degrees of freedom (Rudel, 2008). Conversely, generalization from large-N studies, such as meta-analysis of effect sizes or meta-regression, can capture more observational variation to address selection bias, and distill findings from a large number of studies in a simple, quantified way. However, the need for a large number of cases to build statistical power often limits analysis to descriptions of central tendencies, ignoring outliers that might be particularly informative (Rudel, 2008). Because of the volume of information that needs to be synthesized, researchers using large-N samples also generally have to focus on simple, reduced-form questions, such as "What is the effect of variable X on outcome Y", preventing analysis of more complex causal relations.

The size of the case study collection also constrains the choice of the synthesis method that can be used. From an analytical perspective, a general 'rule of thumb' is a sample size of 30, below which certain statistical methods lose applicability (e.g., regression analysis) and above which comparative methods become overwhelmed (e.g., configurational comparative methods based on set theory) (Ragin, 1987; Rihoux and Ragin, 2009). Thus, more case studies do not necessarily result in more powerful or precise synthesis analysis. Conceptually, the number of observations required depends on the intended conditionality of the GKC, and the adequacy of sampling throughout the range of variability associated with explanatory factors and the outcome of interest. For example, a GKC that is presented as universal, but with a small number of observations and/or over a limited geographic extent, may not have captured alternative outcomes under similar conditions as the sample (i.e., multifinality) or the same outcomes that occurred under different conditions (i.e., equifinality).

2.2.3.3. Scale of observations. The third aspect of GKC methodology to be considered is the spatial scale of the observations under consideration. Issues arise when observations made at different scales are treated as equivalent during synthesis (Margulies et al., 2016). Observational scale has two constituent parts: resolution and extent. For example, field plot measurements or household surveys are high resolution observations, and are most often collected across a small extent, such as a single landscape or series of villages (e.g., Rudel et al., 2000). National-scale observations are of coarser spatial resolution than household data, but better support regional or global GKCs because of their wide availability (e.g., Kaimowitz and Angelsen, 1998). Data of various scales can be used to support the formation of GKCs as long as case study observations are harmonized to be of comparable resolution and extent, and both are consistent with the intended logic of generalization and geographic representativeness of the GKC (Karl et al., 2013; Magliocca et al., 2015).

Uncertainties introduced by comparing observations with ambiguous and/or inconsistent geographic scales have not been adequately considered by global change researchers (Turner et al., 1990, Karl et al., 2013; Magliocca et al., 2015; Margulies et al., 2016). Issues of aggregation and the modifiable areal unit problem (MAUP) in spatial analyses have received significant consideration within the broader field of geography (for a review of some key works, see Marston, 2000). It is well known that inconsistency between the resolution of the data and the level at which claims are derived can lead to “ecological fallacy” (Robinson, 1950). However, these lessons have not necessarily been carried forward when considering how to synthesis multiple, heterogeneous case study observations of social and ecological, factors in order to make broader inferences. Scale mismatches are of particular concern when studying SES, because many social and economic factors are scale-dependent and discontinuous over space (e.g., population density and market access [Verburg et al., 2011a]; agricultural intensity [Laney, 2002]). For example, Vandergeten et al. (2016) conducted a meta-analysis of ‘land grabbing’ claiming local factors, such as land tenure and livelihood strategies, were the main drivers of the social acceptability of land deals. However, the scale of observations contained within the studies synthesized for the meta-analysis ranged from village-level case studies to provincial reviews to national commentaries, each of which presented different forms of scale-dependent explanations or descriptions of change for a local process. Robust generalized claims relating to social-ecological interactions at specific scales rely on observations consistent with these spatial scales.

2.2.3.4. Geographic representativeness. Observations of SES change are almost never made at random across Earth’s surface. For example, Martin et al. (2012) examined the global distribution of terrestrial ecology study site locations and found geographic bias in site selection towards ecological conditions that resembled ‘undisturbed by humans’, but which were also the most accessible from urban areas (e.g., urban parks, protected areas). Research trends can also introduce bias towards locations where particularly salient change is occurring (e.g., desertification, afforestation [Magliocca et al., 2015], and/or locations prioritized for research funding [Martin et al., 2012; McMichael et al., 2017]). As a result, sets of existing case studies selected for making GKCs can be highly biased, over-representing or under-representing more accessible areas, wealthy areas, high population areas, areas of disciplinary research focus, the temperate zone, etc. (Martin et al., 2012; Schmill et al., 2014).

The concept of geographic representativeness is intuitive but can be difficult to assess. Given a population of potential occurrences of a phenomenon for which a GKC will be generated, does the set of case study observations selected represent an unbiased sample from the population? Take the following simplified, hypothetical example for illustration. If the potential population of observations is “the global extent of forests”, and the GKC is “most deforestation is caused by road construction”, the ideal set of observations would be randomly distributed across Earth’s forest biomes (areas that would be forest-covered if not for human use of land). Furthermore, issues of non-stationarity must also be considered such that observations are contemporaneous with hypothesized causes (e.g., changing drivers of deforestation over decades, Rudel et al., 2009). If it were found that 90% of the available forest change observations were instead located near urban areas in the temperate zone, and/or observations where from a limited time span and single mode of deforestation, this would clearly bias any claims made for Earth’s forests in general.

Increasingly, researchers are accompanying their claims with statements regarding the geographic representativeness of their sets of case study observations. For example, they might note the number of sites in different world regions or biomes as evidence that observations are unbiased. These kinds of statements can be more powerful and transparent by taking advantage of more robust statistical tools that now exist for assessing the geographic representativeness (e.g., Magliocca, 2015; Meyfroidt et al., 2014). The GLOBE project introduced online tools that provide a targeted set of statistical indicators for “global representativeness analysis”, enabling researchers to rapidly assess the geographic representativeness of sets of case study observations in relation to global patterns across a broad array of social and environmental variables using a three-step process (Schmill et al., 2014). After mapping the site of each case study observation, ideally as a polygon(s) covering its full extent, the global region to be represented (the “population”), e.g. the global extent of forest biomes, and the global pattern of variation to be assessed, e.g. forest biomes or distance from urban areas, are defined. The degree to which the set of sites differs from a random sample across the global region can then be computed using robust statistical methods that can also map and quantify biases in the sample set caused by over- and under-sampling. As with bias correction in survey statistics, samples may also be reweighted to correct for geographic bias. All of these assessments and corrections can be shared online for citation using a DOI or incorporated as appendices in published work. An example of such analysis performed in the GLOBE system on a contested GKC of global biodiversity changes put forth by Vellend et al. (2013) is publically available online (doi:10.7933/K19S1P7K). Further details are provided in Appendix C. Similar computations have also been conducted outside of GLOBE (e.g., Gonzalez et al., 2016), and Vaclavik et al. (2016) suggested a similar procedure with different statistical indicators for assessing multifactorial representativeness of sample sites within a parameter space of 32 social and environmental variables.

2.3. Typology of GKCs from case study evidence

Based on these dimensions, we have developed a typology of GKCs in which claims are organized by the level of inference, conditionality, and production of testable predictions (Table 4 and Fig. 1). Advancing scientific knowledge about regional or global phenomenon (i.e., moving up and/or to the right in Fig. 1) can be done by independently extending either the level of inference or conditionality of previous GKCs. Additionally establishing or building theory requires inference of causal mechanisms to be aligned with the specific empirical conditions under which predictions can be tested. Where a GKC ultimately falls within this typology is determined by its intended level of inference along a spectrum of descriptive to mechanistic explanations, and the extent to which methodologies used to acquire, analyze, and interpret

Table 4
Generalized knowledge claim (GKC) typology organized primarily by level of inference with variations in claim conditionality within each level.

1) Phenomenological/Descriptive
State, variability, and/or correlations described among patterns of change and influential factors across medium to universal range of conditions with no clear causal inference made.
e.g., "Deforestation occurs across a range of contexts with a variety of outcomes", or " 40% of cropland expansion occurred into forest, and 60% expanded into pasturelands".
2) Inference of Causal Effects
Causal links posited between observed patterns of change and critical variables responsible for the change across medium to universal range of conditions.
e.g., "40% of the variability in deforestation rates is due to economic factors".
3) Inference of Causal Mechanisms
Causal mechanism (i.e., process) posited for how critical variables are responsible for patterns of change across medium to universal range of conditions.
e.g., "Deforestation increased 40% under conditions of good governance because it provided a more favorable environment for business investments of large-scale companies, which are a dominant agent of deforestation".
3a) Middle-Range Theory
Causal mechanisms for explaining variability in outcomes with change in critical variables specified with testable predictions made for all observable conditions with a specified range.
3b) Grand Theory
Causal mechanisms for explaining variability in outcomes with change in critical variables specified with testable predictions made for all observable conditions.

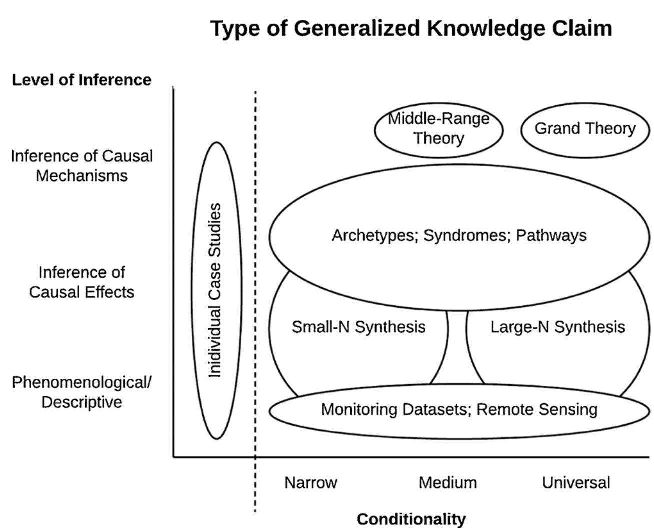

Fig. 1. Types of generalized knowledge claims (GKCs) with examples that vary in dimensions of causal inference, characteristics of local observations used, and synthesis method, which result in differing levels of conditionality. Individual case studies are not GKCs, but are included to the left of the dashed line for comparison and to illustrate alignment between inference level between case studies and the GKC they support. For example, "Monitoring Datasets; Remote Sensing" refers to GKCs made from descriptive analyses of such datasets.

empirical evidence constrain conditionality. It is also important to note that while this typology characterizes only GKCs, the level of inference posited by the GKCs can only be valid to the degree that this aligns with that present in the case studies supporting the GKC, which is why case studies have been included in Fig. 1.

Thus far, we have provided a coherent approach for describing the dimensions of GKCs, and how claims are categorized based on the relationships among those dimensions. In the next section, we introduce a standardized protocol for reporting each dimension of a GKC, which provides the basis for evaluating the commensurability – and thus overall strength – of each dimension given the type of claim being made.

3. A standardized protocol for reporting and assessing GKCs

The proposed protocol reports the essential details of producing GKCs from case study evidence, comparable to the ODD protocol (Overview, Design Concepts, and Details) used for describing agent-based models (Grimm et al., 2005). Our main objective is to provide guidance for increasing rigor and transparency in and improving communication of the GKC generation process. Each step in the protocol describes the research approach used to address one or more dimensions of GKCs as part of the workflow illustrated in Fig. 2. Nine heuristic questions and explanations are provided to elicit sufficient detail about the logic and methodology used at each step to generate or evaluate the validity of a GKC. Current GKCs already address some or most of the suggested reporting details, but a combination of all these reporting elements together are essential to fully assess or communicate the internal consistency of a GKC. We hope that this protocol will evolve once it is used by a sufficiently large number of SES researchers from diverse disciplines.

3.1. Existing knowledge base

3.1.1. Assessing current knowledge and/or GKCs

Question: What is the purpose of the investigation or the GKC? In what ways does it challenge current knowledge and/or existing GKCs?

Explanation: To be transparent, comparable, and falsifiable, GKC's require a precise *system definition* and *state of existing knowledge* about the study system. The *system definition* describes what is and is not being studied, and what aspects of system change (e.g., causes, observed changes, and/or consequences) to which the knowledge claim applies. Specific research questions and hypotheses flow naturally from this system definition. This is followed by a description of how the GKC relates to the *state of existing knowledge* about the study system, and particularly how the claim differs from existing GKCs and the specific limitations of existing knowledge or GKCs that are being addressed. With these characteristics clearly laid out, Fig. 1 and the following questions guide where a claim lies along the axes of conditionality and inference in the GKC typology.

3.2. Logic of generalization

3.2.1. Theoretical or conceptual framework for the claim

Question: What theoretical or conceptual framework is used to describe the study system and/or its dynamics? Is a causal mechanism proposed?

Explanation: In non-experimental contexts, the researcher does not know all the conditions under which the studied system's possible states can be observed. In these instances, a theoretical or conceptual framework is necessary to bound the system components, their interactions, and expected variability thought to be important for observing, explaining, and/or predicting changes in the study system. More importantly for making GKCs, a theoretical or conceptual framework identifies the epistemic communities to which the claim is targeted, existing knowledge base on which the GKC is based, critical variables for describing system state and change, and which hypotheses are refutable. While most, if not all, global change research uses some type of theoretical or conceptual framework to structure investigations, such frameworks are not always made explicit to the knowledge user. Being explicit about the conceptual framework increases transparency and reproducibility across knowledge producers, facilitating assessment of GKCs by knowledge users.

3.2.2. Intended claim conditionality

Question: Under what empirical conditions does the chosen theoretical or conceptual framework apply?

Explanation: While *actual* conditionality is determined by methodological constraints, the *intended* conditionality of a GKC depends on

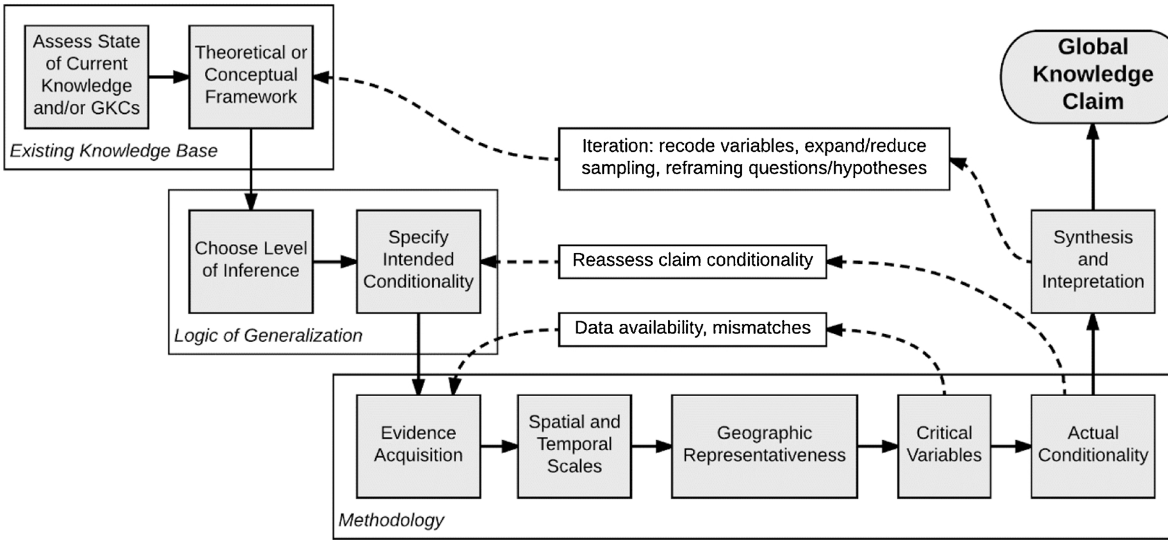

Fig. 2. Workflow for producing generalized knowledge claims.

the chosen theoretical or conceptual framework. Some frameworks are intended for application under specific conditions, such as Ostrom's SES framework (Ostrom, 2007), while others are intended to be applicable across a wider range of conditions, such as the forest transition theory (Mather, 1992; Rudel et al., 2005). When case study observations are all collected and interpreted through a particular framework, the conclusions resulting from a synthesis of those case studies will likely reflect the conceptual or epistemological biases inherent to that framework, potentially excluding important factors or counterfactual outcomes. For example, a political economy analysis of LSLAs may provide substantial depth into the institutional arrangements and power imbalances resulting in conflict, but it may also underrepresent cases where no conflict occurred or institutional factors were not influential in the occurrence of conflicts (e.g., due to biophysical conditions). In the interest of clarity and transparency, we suggest that authors consider providing an explicit discussion of the theoretical or conceptual framework(s) used in the cases at hand, and how it (they) might influence the intended conditionality of a GKC.

3.3. Methodology

3.3.1. Evidence acquisition process

Question: How was case study data compiled and selected?

Explanation: The *actual* claim conditionality is constrained by choices of synthesis methodology. Established guidelines for meta-study and meta-analysis provide a checklist of reporting items to ensure data acquisition and manipulation are transparent and reproducible (e.g., Gerstner et al., 2017; Liberati et al., 2009; Magliocca et al., 2015; Rudel, 2008). Standard reporting items include search terms and phrases, the search engine and/or database used, and explicit criteria for excluding irrelevant results. Additionally, statistics are usually provided for how many initial results were returned by the search strategy, how many sources remained after applying exclusion criteria, and the relationships between number of sources and analytical cases produced from those sources.

Additional reporting needs for GKCs include the steps taken to define and extract comparable data from each case should be documented, including an accounting of the types of data extracted (i.e., quantitative, qualitative, or mixed) and whether intermediate standardization steps, such as qualitative coding or quantitative normalization, were needed. In the case of qualitative coding, inter-coder reliability statistics should also be reported (Magliocca et al., 2015). In addition, if case data have a geographical component, explicit instructions explaining how case geometries (e.g., village boundaries, watershed) were defined within and across sources are needed to provide precise descriptions of the areal extent of the analysis (Margulies et al., 2016).

3.3.2. Spatial and temporal scales

Question: What are the spatial and temporal resolutions and extents (i.e., scale of observations) over which the GKC was developed? Does the scale of the claim match the scale of the case data?

Explanation: Spatial and temporal resolution are important for determining if the data used to make a GKC is sufficiently sensitive to the phenomenon of interest. For example, if spatial configuration is important for the topic of study, spatial data should be sufficiently fine-grained to capture relevant differences across space. Coarse resolution data will tend to exclude high and low values and introduce bias into analyses based on central tendency measures. Similarly, the temporal resolution of data collected requires sufficient frequency to detect changes resulting from the dynamics of the study system. For example, land-use change tends to occur on seasonal to annual time scales, whereas less frequently collected data may be sufficient to observe natural processes of land cover change (e.g., forest regrowth). If dynamics occur at finer intervals than data collection, the exact timing of change may not be detected in ways that can be reliably related to causal factors (i.e., 'censoring' [An et al., 2011]). Also, if a particular claim implies occurrence of a structural change in a system state (e.g., forest land cover, species richness), a stronger claim can be made if case study observations cover comparable time spans.

Choices of spatial and temporal extent can also introduce biases. For example, global data coverage is consistently poor in certain areas, like islands. Excluding such areas may reduce noise in analysis, but may also lose important outlier/extreme cases (Verburg et al., 2011b). Similarly, for temporal extent, including/excluding time periods of unusual volatility or stability can bias analyses. For example, if one is interested in estimating long-term rates of cropland expansion driven by global commodity demands, the decision to include or exclude data from the 2007/8 and 2010/11 food crises will substantially alter estimated rates (e.g., Suweis et al., 2015). If there are known biases created by constraining the spatial and temporal extent for the sake of analytical tractability, these should be made explicit and be reflected in the stated conditionality of the GKC.

3.3.3. Geographic representativeness

Questions: How representative is the collection of case study observations of the conditions in which the phenomenon of interest is expected? How was this assessed? If geographic bias was detected, how was it addressed?

Explanation: The first step in assessing the geographic representativeness of a set of cases is to specify the geographic extent over which the phenomenon of interest is expected. Typically, GKCs are not intended to be truly general in the sense that they apply to all locations on Earth. Rather, claims are meant to apply to places with certain characteristics, such as areas with specific agricultural suitability or within tropical forested biomes. A precise definition is needed of the intended geographic extent by indicating the variables and their values used to delimit the geographic range to which the GKC pertains (Schmill et al., 2014).

Once the applicable spatial extent of the GKC has been defined, a comparison of selected case geographies determines how well they represent the intended geographic space (see Section *2.2.3 Geographic Representativeness* for details). When geographic biases in existing case study observations preclude an analysis across the intended conditionality of the GKC, a researcher can acquire additional observations or reweight the existing sample to reduce bias. Alternatively, the geographic extent of the GKC can be reduced or adjusted, so that the claim matches the extent that can be validly represented (with these procedures reported as statistical methods). Either way, given the wide availability of global data, mapping tools, and robust statistical methods for assessing geographic biases in GKCs derived from sets of case study observations, statistical measures of geographic bias and a precisely defined geographic extent ought to become the standard required of all GKCs published in peer-reviewed literature, just as statistical measures of variance are now required to publish numerical estimates.

3.3.4. Critical variables

Questions: How do the theoretical or conceptual framework and spatial and temporal scales of analysis intersect to determine critical variables for making the GKC? What variables influence the validity of the claim, and thus bound the region of claim validity? What were the measures taken to mitigate these critical variables?

Explanation: Due to the challenges of integrating data of different types and/or with various spatial and temporal characteristics, a theoretically-driven scope of an investigation may not always align with the scope of available data sources. Thus, one or more variables may emerge as ‘critical variables’ that limit the extent and/or resolution of analysis, which force a compromise between theoretical or conceptual comprehensiveness and analytical tractability. For example, land governance is often an important factor in land-use change. However, it is often reported only at national levels, or at sub-national resolution in limited locations, which is at a different resolution or extent than the land conversion process of interest. It is important to explicitly specify any measures taken to address such gaps or limitations. This fosters transparency about the intersection between ‘what should be done’ and ‘what could be done’, so that potential biases or blind spots in the process of making a GKC are explicit. Together, the intersection between the applied theoretical or conceptual framework and scope of the analysis demarcates the boundary of the study system and provides a tangible problem definition (Palmer et al., 2016).

3.3.5. Actual claim conditionality

Questions: Under what empirical conditions is the generalized knowledge claim applicable? How narrow or broad are the range of conditions under which the claim is specified?

Explanation: Because there are multiple methodological choices that determine actual claim conditionality, trade-offs are inevitable. Imposing a stricter range of conditions of validity is likely to result in a more robust claim, while a stricter range of conditions limits the scope of the claim. Failing to specify the range of conditions correctly may result in the claim being shown invalid under some conditions, which may undermine the confidence in the whole claim itself. Much miscommunication and contestation surrounding current GKCs might have been avoided with more precise and explicit definitions of conditionality.

3.3.6. Iterations

Question: How did evidence acquisition, interpretation, and/or analysis change as the investigation progressed?

Explanation: The generalization process is often iterative, because each additional piece of case study evidence might demand a new interpretation of general patterns and/or require a different specification of conditionality. A kind of ‘progressive contextualization’ (Vayda, 1983) occurs in which the analyst explores and then explicates the links between patterns in explanatory variables and outcomes and the larger contexts of the change in which case study observations were made. During this process, the researcher may become aware of new patterns in the data, often having to do with new contexts of additional case studies. The researcher may then revise the research question, alter the sampling strategy to include/exclude additional studies, restructure the coding scheme to account for new relationships, and/or quantitatively reweight each observation to address biases. Reporting as a narrative any iterations resulting in a change of research question, search and/or inclusion/exclusion of case study observations, or synthesis methodology makes clear the logic of generalization and its limitations.

4. Evaluating GKCs

The validity of GKCs, like other forms of synthetic knowledge, depends on the *alignment* and *consistency* between the type and scope of the claim and the evidence used to support it. Alignment and consistency are emphasized because specific dimensions of a given GKC might be individually appropriate or robust, such as large-N synthesis methods for medium to universal claims, but the overall strength of a GKC depends on how well these individual dimensions align with each other to build a robust and coherent argument. Unless all of the components discussed above (Section 2) align, the strength of a given GKC is questionable. Similar to the systematic review appraisal approach used by the Collaboration for Environmental Evidence, a GKC is evaluated not only by the scientific merit of the individual conceptual and methodological steps in the process of making a claim, but also how those steps align and are reported in the context of the question addressed by the GKC.

Here we propose a two-step approach to evaluating the relative strength of a GKC: 1) alignment between its scale, observational evidence, and conditionality (*evidence* evaluation); and 2) agreement between the strength of the claim and the strength of prior claims or other existing empirical evidence (*claim* evaluation). The proposed evaluation can be used by authors of GKCs during the research process, or by consumers of GKCs to assess the reliability and transparency of claims. This evaluation uses the data quality pedigree approach of post-normal science to facilitate the communication of uncertainties in scientific evidence outside the expert communities within which these evidence were created (Funtowicz and Ravetz, 1991). This is accomplished here by ranking elements of data quality into relative grades (‘weak’, ‘moderate’, and ‘strong’) using a simple, explicit rubric (Table 5). Grades are then combined using a qualitative logic, generally by downgrading to the lowest grades (Costanza et al., 1992; van der Sluijs et al., 2005). Rubric and grades presented in Table 5 are based on the levels specified for each GKC dimension presented in Table 1. It should be noted here, and wherever this data quality rubric is applied and its results are utilized, that data quality grades are intended only to enable rough, qualitative comparisons among GKCs, based on the principle that an arbitrary but transparent and consistent rubric for assessing the relative quality of GKC's is a useful advance over the arbitrary but non-transparent and inconsistent state of GKC quality evaluation that now exists. The goal here is not to provide absolute objective rankings, but rather, to make it possible for non-experts to assess the relative strength of GKCs at the most basic level- a level both currently lacking and much in need of support.

Based on the principle that the level of analysis should match the scale at which the phenomenon of interest is produced (Munroe et al.,

Table 5
Qualitative rubric for grading the commensurability of scale alignment and observational evidence to support a specific type of GKC, graded as 'strong' (3), 'moderate' (2), or 'weak' (1). Overall grades across categories are downgraded to the lowest grade.

		Scale Alignment						Observational Evidence	
		Scale of Phenomenon			*Geographic Representativeness*			*Number of Observations*	
		Local	Sub-Nat'l to Nat'l	Regional to Global	Limited or Unknown	Regional	Global	Small-N	Large-N
Scale of Observations	Local	**3**	2	1	1	2	**3**	**3**	**3**
	Sub-Nat'l to Nat'l	2	**3**	2	1	2	**3**	2	**3**
	Regional to Global	1	2	**3**	1	1	2	1	**3**

Table 6
Rubric for grading the commensurability of conditionality and level of inference used to make a GKC, which is graded as 'strong' (3), 'moderate' (2), or 'weak' (1). Claims with moderate observational evidence and scale alignment scores are graded weak and moderate for universal and medium conditionality claims, respectively. Scale alignment and observational evidence scores are calculated from Table 5. Overall grades across categories are downgraded to the lowest grade.

Universal or Medium Conditionality				
		Scale Alignment Score		
		Strong	Moderate	Weak
Obs. Evidence Score	Strong	**3**	2	1
	Moderate	2	1 - 2	1
	Weak	1	1	1

		Causality	
		Type of GKC	
		Theory	Phenom. or Descriptive
Level of Inference	Causal Mechanisms	**3**	1
	Causal Effects	2	2
	Descriptive	1	**3**

2014), the first stage of *evidence* evaluation is to test for *scale alignment*. GKCs are most robust when the spatial scale of observations (i.e., resolution and extent) aligns with that of the phenomenon of interest[2] (Table 5). For example, global sets of observations make more extensive claims, but global data tends to be coarse, potentially missing important details detectable from observations made at local scales (Verburg et al., 2011b). When the spatial scales of case study observations and phenomenon of interest are consistent, GKC then *scale alignment* is 'strong'. If spatial resolution *or* extent are one level different from the observed phenomenon, then *scale alignment* is 'moderate' (involves a moderate degree of interpolation, or downscaling). If both the resolution and extent are one level different, or one is more than one level different from the observed phenomenon, then *scale alignment* is 'weak'.

The *scale alignment* grade is then corrected for geographic bias. The grade for spatial scale of observations is compared with the geographic representativeness grade, and the overall grade equals the lowest between the two. For example, 'strong' *alignment* combined with an unknown geographic bias produces a 'weak' GKC, and 'strong' *alignment* combined with an assessed and corrected geographic bias produces a GKC of 'moderate' strength. Of course, a geographically unbiased collection of observations combined with 'strong' *scale alignment* produces the strongest GKC.

The second stage of *evidence* evaluation is an assessment of *observational evidence* by comparing the number of observations with the scale alignment assessment, based on the principle that larger scale claims demand larger sample sizes (Table 5). Small-N studies tend to be best suited for localized phenomenon and synthesis methods aimed at establishing causal effects or mechanisms (e.g., configurational comparative analysis), because context-specific details from each case can be retained more easily during synthesis than for larger samples. In contrast, large-N studies are best suited for analyzing broader scale or localized but widespread phenomenon, and are easily synthesized with methods designed to establish a descriptive or causal effects understanding (e.g., meta-regression analysis). At this point in the evaluation, GKCs based on large-N samples of local scale observations are not penalized, because there is inherently more case study information to analyze with large-N studies. Although, it is likely such GKCs will be downgraded one or two levels in the second stage of evidence evaluation due to difficulties in establishing causal mechanisms with large-N synthesis methods. When the grade of sample size is greater than or equal to the grade *scale alignment*, this is 'strong' *observational evidence*. If the grade of sample size is one step lower than for *scale alignment*, then *observational evidence* is 'moderate'. If sample size is more than one grade lower than for *scale alignment*, then *observational evidence* is 'weak'.

The third stage of *evidence* evaluation is to compare the conditionality and level of causal inference (i.e., causality) of the GKC (Table 6) against the strength of *scale alignment* and *observational evidence* grades (Table 5). Universal GKCs are inherently larger scale claims (large region or Earth) and therefore demand support by the strongest *observational evidence* and *scale alignment*. These are 'strong' only when both *observational evidence* and *scale alignment* are 'strong', and 'moderate' when one of these are 'moderate' (but not both); under any other condition, these are 'weak'. Medium conditionality GKCs are made at all levels below the entire Earth, are 'strong' when both *observational evidence* and *scale alignment* are 'strong', 'moderate' when one or both of these are 'moderate', and 'weak' under all other conditions. Narrow knowledge claims are not GKCs.

[2] A well-established literature and set of methods exists for evaluating the reliability of remote sensing analyses. For example, see recent reviews by Lillesand et al. (2014) and Colomina and Molina (2014).

Grading the appropriate level of causal inference depends on the type of GKC being made (Fig. 1). GKCs for the purpose of theory are ‘strong’ only if inference is based on causal mechanisms, ‘moderate’ is inference based on causal effects is used, and ‘weak’ if made with descriptive analysis. Similarly, phenomenological or descriptive GKCs are best made with and graded ‘strong’ when using large-scale observational tools, such as remote sensing. Such GKCs are graded ‘moderate’ or ‘weak’ when other synthesis methods or theoretical frameworks are used, because methods that rely on higher levels of causal inference may not be appropriate for large-N, descriptive or phenomenological analysis.

Once the *evidence* for a GKC has been graded, the overall strength of the GKC is evaluated relative to its agreement with prior studies and existing theory (*claim evaluation*) using a stylized Bayesian approach that incorporates the strength of prior evidence. Prior evidence is graded ‘strong’ when more than one strong study (graded using the evidence evaluation above) has been conducted and the GKCs of these studies agree. Prior evidence is graded ‘moderate’ when at least one strong or more than one moderate GKC(s) exists. All other prior evidence is graded “weak”. If the new GKC is strong and agrees with prior evidence, the current state of the GKC is graded “strong”. If the new GKC is strong and disagrees with prior evidence[3] or is moderate and agrees with prior evidence, the current state of the GKC is graded “moderate”. All other conditions are graded “weak”.

To demonstrate and validate this grading system, we have evaluated two examples of recent GKCs. Detailed reporting using the proposed standard protocol and grading system are available for each study in Appendices A and B. Differences in the amount of detail provided in each evaluation reflect expert judgements in applying the evaluation rubrics. However, these differences also demonstrate variations in the methodologies and reporting standards used to produce the GKCs.

The first is a global meta-analysis by Gerstner et al. (2014) of the effects of land use on plant biodiversity. The authors concluded that there is a general trend for a loss in plant species richness due to land-use intensification. They used a meta-analysis of effect sizes with 375 case studies with an intended global conditionality, but a geographic bias towards developed world countries was detected. A mix of spatial scales was also observed across the case study collection ranging from 0.1 to 1000 km^2. Interestingly, the authors conducted a sensitivity analysis for the scale of observations and did not detect an effect of spatial scale. Overall, the study’s GKC achieved a ‘moderate’ grade.

The second GKC evaluated was produced by Oberlack et al. (2016), which consisted of an archetypes analysis of the livelihood effects of LSLAs. The authors assert that adverse livelihood conditions arise most frequently from seven processes which are activated in specific configurations of social and ecological factors. The archetypes analysis, which is a form of configurational comparative analysis, used 66 cases of LSLA effects on livelihoods from diverse geographical settings. However, the geographic representativeness of the sample is ambiguous, because the LSLA phenomenon is relatively new and thus its actual conditionality is unknown. Case study observations were based on studies at local, sub-national, and national scales, which presented some comparability issues. Although, these limitations did not prevent the authors from establishing a strong level of inference based on causal mechanisms. Overall, the study’s GKC achieved a ‘moderate’ grade.

5. Conclusions

By classifying, documenting, and evaluating GKCs, we aim to advance the development of a rigorous approach to producing generalized knowledge of SES in a form useful for non-experts. Based on the approach we describe, the production and use of generalized knowledge claims derived from local and regional observations can be made more transparent, reproducible, robust, and reliable while also clarifying knowledge gaps to be filled through further research efforts.

In developing this approach, we found that the types and purposes of GKCs are often not made clear, and consequently the conditions under which these claims are valid (claim conditionality) remain vague. The appropriate conditionality of GKCs differs across problem domains, SES change phenomena, and the existing state of knowledge. Situations exist in which the need for global knowledge is imperative (e.g., transnational LSLAs), while the accumulated local evidence may be insufficient to produce robust synthetic knowledge at the desired levels of generalization. Further, theory may not yet be articulated or tested at levels adequate to guide data collection and inference at desired levels (e.g., general strategies for adaptation to climate change).

The development of GKCs is extremely important, as there is an urgent need to coordinate efforts to respond to environmental and social challenges that operate beyond specific places and have global relevance. Global processes, such as economic globalization, environmental change, and species invasions, are increasingly connecting and transforming local SESs, and there are lessons to be learned from how different locations might respond in similar or different ways to the same global forces. For example, are agricultural development approaches that were successful in Brazil transferrable to East Africa (e.g., Cabral and Shankland, 2013)? How do changes in species richness in human-dominated landscape vary according to land-use intensity (e.g., Gerstner et al., 2014; Gonzalez et al., 2016)? Or, how similar or different are pathways of land conversion for oil palm in Central America and Southeast Asia (e.g., Meyfroidt et al., 2014)?

GKCs can provide the insights and evidence needed to coordinate policy responses at regional and global scales (Messerli et al., 2013). Implementation of the approach we propose will help align the conceptual and methodological elements of generalized knowledge production about SES. In documenting alignments between the scope and evidence of claims, GKC users are made aware that the strength and conditionality of GKCs depends both on methodological choices made by the producer of the claim and also the purposes to which the claim is applied, *a posteriori*. When properly applied, the reporting and assessment approach we present can help to ensure that GKCs are well grounded in evidence and have a clearly defined scope of validity, further increasing the likelihood that generalized knowledge of SESs can be made both reliable and *useable* for sustainable development.

Acknowledgements

This paper was based on a workshop supported jointly by the GLOBE project (US National Science Foundation grant #1125210) and the National Socio-Environmental Synthesis Center (SESYNC; NSF award DBI-1052875) titled “Globalizing Our Understanding of Land-Use Change”, and is one of the synthesis activities of the Global Land Programme (http://glp.earth).

References

An, L., Linderman, M., Qi, J., Shortridge, A., Liu, J., 2005. Exploring complexity in a human–environment system: an agent-based spatial model for multidisciplinary and multiscale integration. Ann. Assoc. Am. Geogr. 95 (1), 54–79.
An, L., Brown, D.G., Nassauer, J.I., Low, B., 2011. Variations in development of exurban residential landscapes: timing, location, and driving forces. J. Land Use Sci. 6 (1), 13–32.
Allan, J.A., 2003. Virtual water-the water, food, and trade nexus. useful concept or misleading metaphor? Water Int. 28 (1), 106–113.
Bäckstrand, K., Lövbrand, E., 2006. Planting trees to mitigate climate change: contested discourses of ecological modernization, green governmentality and civic environmentalism. Global Environ. Politics 6 (1), 50–75.
Bradley, N.L., Leopold, A.C., Ross, J., Huffaker, W., 1999. Phenological changes reflect climate change in Wisconsin. Proceedings of the National Academy of Sciences 96 (17), 9701–9704.
Breu, T., Bader, C., Messerli, P., Heinimann, A., Rist, S., Eckert, S., 2016. Large-scale land

[3] Such a GKC could be upgraded as additional, corroborating evidence challenging prior theory or GKCs emerges. In contrast, the standards for achieving a ‘strong’ GKC are lower when it agrees with prior GKCs.

acquisition and its effects on the water balance in investor and host countries. PloS one 11 (3), e0150901.
Cabral, L., Shankland, A., 2013. Narratives of Brazil-Africa cooperation for agricultural development: new paradigms? China Braz. Afr. Agric. (CBAA) Project Working Pap. 51, 1–27.
Cleland, E.E., Chuine, I., Menzel, A., Mooney, H.A., Schwartz, M.D., 2007. Shifting plant phenology in response to global change. Trends Ecol. Evol. 22 (7), 357–365.
Colomina, I., Molina, P., 2014. Unmanned aerial systems for photogrammetry and remote sensing: a review. ISPRS J. Photogramm. Remote Sens. 92, 79–97.
Costanza, R., Funtowicz, S.O., Ravetz, J.R., 1992. Assessing and communicating data quality in policy-relevant research. Environ. Manage. 16 (1), 121–131.
Dell'Angelo, J., D'Odorico, P., Rulli, M.C., Marchand, P., 2017a. The tragedy of the grabbed commons: coercion and dispossession in the global land rush. World Dev. http://dx.doi.org/10.1016/j.worlddev.2016.11.005.
Dell'Angelo, J., D'Odorico, P.Rulli, 2017b. The global water grabbing syndrome. Ecol. Econ in press.
Ebbinghaus, B., 2005. When less is more selection problems in large-N and small-N cross-national comparisons. Int. Sociol. 20 (2), 133–152.
Ellis, E.C., Maslin, M., Boivin, Nicole, Bauer, A., 2016. Involving social scientists in defining the anthropocene. Nature 540, 192–193.
DeFries, R., Hansen, A., Turner, B.L., Reid, R., Liu, J., 2007. Land use change around protected areas: management to balance human needs and ecological function. Ecol. Appl. 17 (4), 1031–1038.
Funtowicz, S.O., Ravetz, J.R., 1991. Uncertainty and Quality in Science for Policy. Kluwer Academic Publishers, Dordrecht, The Netherlands.
Geist, H.J., Lambin, E.F., 2002. Proximate causes and underlying driving forces of tropical deforestation tropical forests are disappearing as the result of many pressures, both local and regional, acting in various combinations in different geographical locations. BioScience 52 (2), 143–150.
George, A.L., Bennett, A., 2005. Case Studies and Theory Development in the Social Sciences. MIT Press, Cambridge, MA.
Gerstner, K., Dormann, C.F., Stein, A., Manceur, A.M., Seppelt, R., 2014. Editor's choice: review: effects of land use on plant diversity–a global meta-analysis. J. Appl. Ecol. 51 (6), 1690–1700.
Gerstner, K., Moreno-Mateos, D., Gurevitch, J., Beckmann, M., Kambach, S., Jones, H.P., Seppelt, R., 2017. Will your paper be used in a meta-analysis? make the reach of your research broader and longer lasting. Methods in Ecology and Evolution.
Gonzalez, A., Cardinale, B.J., Allington, G.R., Byrnes, J., Arthur Endsley, K., Brown, D.G., Hooper, D.U., Isbell, F., O'Connor, M.I., Loreau, M., 2016. Estimating local biodiversity change: a critique of papers claiming no net loss of local diversity. Ecology 97 (8), 1949–1960. http://dx.doi.org/10.1890/15-1759.1.
Grau, H.R., Aide, M., 2008. Globalization and land-use transitions in latin America. Ecol. Soc. 13 (2), 16.
Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W.M., Railsback, S.F., Thulke, H.-H., Weiner, J., Wiegand, T., DeAngelis, D.L., 2005. Pattern-oriented modeling of agent-based complex systems: lessons from ecology. Science 310, 987–991.
Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., 2013. High-resolution global maps of 21st-century forest cover change. Science 342 (6160), 850–853.
Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., McElreath, R., Alvard, M., Barr, A., Ensminger, J., Henrich, N.S., 2005. "Economic man" in cross-cultural perspective: behavioral experiments in 15 small-scale societies. Behav. Brain Sci. 28 (6), 795–815.
Hersperger, A.M., Bürgi, M., 2009. Going beyond landscape change description: quantifying the importance of driving forces of landscape change in a Central Europe case study. Land Use Policy 26, 640–648. http://dx.doi.org/10.1016/j.landusepol.2008.08.015.
Hochstetler, K., Kostka, G., 2015. Wind and solar power in Brazil and China: interests, state–business relations, and policy outcomes. Global Environ. Politics 15 (3), 74–94.
Kaimowitz, D., Angelsen, A., 1998. Economic Models of Tropical Deforestation: A Review. Center for International Forestry Research, Bogor, Indonesia.
Karl, J.W., Herrick, J.E., Unnasch, R.S., Gillan, J.K., Ellis, E.C., Lutters, W.G., Martin, L.J., 2013. Discovering ecologically-relevant knowledge from published studies through geo-semantic searching. BioScience 63, 674–682.
Keys, E., McConnell, W.J., 2005. Global change and the intensification of agriculture in the tropics. Global Environ. Change 15 (4), 320–337.
Koricheva, J., Gurevitch, J., Mengersen, K., 2013. Handbook of Meta-analysis in Ecology and Evolution. Princeton University Press.
Laney, R.M., 2002. Disaggregating induced intensification for land-change analysis: a case study from Madagascar. Ann. Assoc. Am. Geogr. 92 (4), 702–726.
Liberati, A., Altman, D.G., Tetzlaff, J., Mulrow, C., Gøtzsche, P.C., Ioannidis, J.P., Clarke, M., Devereaux, P.J., Kleijnen, J., Moher, D., 2009. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. Ann. Intern. Med. 151 (4). http://dx.doi.org/10.7326/0003-4819-151-4-200908180-00136. W-65-W-94.
Lillesand, T., Kiefer, R.W., Chipman, J., 2014. Remote Sensing and Image Interpretation. John Wiley & Sons.
MacDonald, G.K., Bennett, E.M., Taranu, Z.E., 2012. The influence of time, soil characteristics, and land-use history on soil phosphorus legacies: a global meta-analysis. Glob. Change Biol. 18 (6), 1904–1917.
Magliocca, N.R., 2015. Model-based synthesis of locally contingent responses to global market signals. Land 4 (3), 807. http://dx.doi.org/10.3390/land4030807.
Magliocca, N.R., Rudel, T.K., Verburg, P.H., McConnell, W.J., Mertz, O., Gerstner, K., Heinimann, A., Ellis, E.C., 2015. Synthesis in land change science: methodological patterns, challenges, and guidelines. Reg. Environ. Change 15 (2), 211–226.
Margulies, J.D., Magliocca, N.R., Schmill, M.D., Ellis, E.C., 2016. Ambiguous geographies: connecting case study knowledge with global change science. Ann. the Am. Assoc. Geogr. 106 (3), 572–596.
Marston, S.A., 2000. The social construction of scale. Prog. Hum. Geogr. 24 (2), 219–242.
Martin, L.J., Blossey, B., Ellis, E.C., 2012. Mapping where ecologists work: biases in the global distribution of terrestrial ecological observations. Front. Ecol. Environ. 10 (4), 195–201.
Mather, A.S., 1992. The forest transition. Area 24, 367–379.
McMichael, C.N.H., Matthews-Bird, F., Farfan-Rios, W., Feeley, K.J., 2017. Ancient human disturbances may be skewing our understanding of Amazonian forests. Proceedings of the National Academy of Sciences 114, 522–527.
Messerli, P., Heinimann, A., Giger, M., Breu, T., Schönweger, O., 2013. From 'land grabbing' to sustainable investments in land: potential contributions by land change science. Curr. Opin. Environ. Sustain. 5, 528–534.
Meyfroidt, P., 2013. Environmental cognitions, land change and social-ecological feedbacks: local case studies of forest transition in Vietnam. Hum. Ecol. 41 (3), 367–392. http://dx.doi.org/10.1007/s10745-012-9560-x.
Meyfroidt, P., 2016. Approaches and terminology for causal analysis in land systems science. J. Land Use Sci. 1–27.
Meyfroidt, P., Carlson, K.M., Fagan, M.E., Gutiérrez-Vélez, V.H., Macedo, M.N., Curran, L.M., DeFries, R.S., Dyer, G.A., Gibbs, H.K., Lambin, E.F., Morton, D.C., Robiglio, V., 2014. Multiple pathways of commodity crop expansion in tropical forest landscapes. Environ. Res. Lett. 9, 074012. http://dx.doi.org/10.1088/1748-9326/9/7/074012.
Munroe, D.K., McSweeney, K., Olson, J.L., Mansfield, B., 2014. Using economic geography to reinvigorate land-change science. Geoforum 52, 12–21.
Newbold, T., Hudson, L.N., Hill, S.L., Contu, S., Lysenko, I., Senior, R.A., Börger, L., Bennett, D.J., Choimes, A., Collen, B., Day, J., et al., 2015. Global effects of land use on local terrestrial biodiversity. Nature 520 (7545), 45–50.
Oberlack, C., Tejada, L., Messerli, P., Rist, S., Giger, M., 2016. Sustainable livelihoods in the global land rush? archetypes of livelihood vulnerability and sustainability potentials. Global Environ. Change 41, 153–171.
Ostrom, E., 2007. A diagnostic approach for going beyond panaceas. Proceeding of the National Academy of Sciences 104 (39), 15181–15187. http://dx.doi.org/10.1073/pnas.0702288104. USA.
Palmer, M.A., Kramer, J.G., Boyd, J., Hawthorne, D., 2016. Practices for facilitating interdisciplinary synthetic research: the national socio-environmental synthesis center (SESYNC). Curr. Opin. Environ. Sustain. 19, 111–122.
Parmesan, C., Yohe, G., 2003. A globally coherent fingerprint of climate change impacts across natural systems. Nature 421 (6918), 37–42.
Ragin, C.C., 1987. The Comparative Method. University of California Press, Berkeley.
Rasmussen, K., D'haen, S., Fensholt, R., Fog, B., Horion, S., Nielsen, J.O., Rasmussen, L.V., Reenberg, A., 2016. Environmental change in the Sahel: reconciling contrasting evidence and interpretations. Reg. Environ. Change 16 (3), 673–680.
Rihoux, B., Ragin, C.C., 2009. Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques. Sage: Chicago.
Rindfuss, R.R., Walsh, S.J., Turner II, B.L., Fox, J., Mishra, V., 2004. Developing a science of land change: challenges and methodological issues. Proceedings of the National Academy of Sciences 101, 13976–13981.
Robinson, W.S., 1950. Ecological correlations and the behavior of individuals. Am. Sociol. Rev. 15 (3), 351–357.
Rudel, T.K., 2008. Meta-analyses of case studies: a method for studying regional and global environmental change. Global Environ. Change 18, 18–25.
Rudel, T., Flesher, K., Bates, D., Baptista, S., Holmgren, P., 2000. The tropical deforestation literature: geographical and historical patterns of information. Unasylva 51, 46–53.
Rudel, T.K., Coomes, O.T., Moran, E., Achard, F., Angelsen, A., Xu, J., Lambin, E., 2005. Forest transitions: towards a global understanding of the land use change. Global Environ. Change 15, 23–31.
Rudel, T.K., Defries, R., Asner, G.P., Laurance, W.F., 2009. Changing drivers of deforestation and new opportunities for conservation. Conserv. Biol. 23 (6), 1396–1405.
Schmill, M.D., Gordon, L.M., Magliocca, N.R., Ellis, E.C., Oates, T., 2014. GLOBE: analytics for assessing global representativeness. Computing for Geospatial Research and Application (COM. Geo), 2014 Fifth International Conference. IEEE, pp. 25–32.
Seppelt, R., Manceur, A.M., Liu, J., Fenichel, E.P., Klotz, S., 2014. Synchronized peak-rate years of global resources use. Ecol. Soci. 19http://dx.doi.org/10.5751/ES-07039-190450. art50.
Steffen, W., Persson, Å., Deutsch, L., Zalasiewicz, J., Williams, M., Richardson, K., Crumley, C., Crutzen, P., Folke, C., Gordon, L., Molina, M., Ramanathan, V., Rockström, J., Scheffer, M., Schellnhuber, H.J., Svedin, U., 2011. The Anthropocene: from global change to planetary stewardship. Ambio 40 (7), 739–761.
Steinberg, P.F., 2015. Can we generalize from case studies? Global Environ. Politics 15 (3), 152–175. http://dx.doi.org/10.1162/GLEP_a_00316.
Suweis, S., Carr, J.A., Maritan, A., Rinaldo, A., D'Odorico, P., 2015. Resilience and reactivity of global food security. Proc. Natl. Acad. Sci. 112 (22), 6902–6907.
Tamea, S., Laio, F., Ridolfi, L., 2016. Global effects of local food-production crises: a virtual water perspective. Sci. Rep. 6.
Taub, D.R., Miller, B., Allen, H., 2008. Effects of elevated CO2 on the protein concentration of food crops: a meta-analysis. Glob. Change Biol. 14 (3), 565–575.
Turner, B.L., Ali, A.S., 1996. Induced intensification: agricultural change in Bangladesh with implications for Malthus and Boserup. Proceedings of the National Academy of Sciences 93 (25), 14984–14991.
Turner, B.L., Kasperson, R.E., Meyer, W.B., Dow, K.M., Golding, D., Kasperson, J.X., Mitchell, R.C., Ratick, S.J., 1990. Two types of global environmental change: definitional and spatial-scale issues in their human dimensions. Global Environ. Change 1 (1), 14–22.
Turner, B.L., Lambin, E.F., Reenberg, A., 2007. The emergence of land change science for

global environmental change and sustainability. PNAS 104 (52), 20666–20671.

Václavík, T., Langerwisch, F., Cotter, M., Fick, J., Häuser, I., Hotes, S., et al., 2016. Investigating potential transferability of place-based research in land system science. Environ. Res. Lett. 11, 95002. http://dx.doi.org/10.1088/1748-9326/11/9/095002.

Valbuena, D., Verburg, P.H., Bregt, A.K., Ligtenberg, A., 2010. An agent-based approach to model land-use change at a regional scale. Landsc. Ecol. 25 (2), 185–199.

van der Sluijs, J.P., Craye, M., Funtowicz, S., Kloprogge, P., Ravetz, J., Risbey, J., 2005. Combining quantitative and qualitative measures of uncertainty in model-based environmental assessment: the NUSAP system. Risk Anal. 25 (2), 481–492.

van Vliet, J., Magliocca, N.R., Büchner, B., Cook, E., Benayas, J.M.R., Ellis, E.C., Heinimann, A., Keys, E., Lee, T.M., Mertz, O., 2015a. Meta-studies in land use science: current coverage and prospects. Ambio 45 (1), 15–28.

van Vliet, J., de Groot, H.L., Rietveld, P., Verburg, P.H., 2015b. Manifestations and underlying drivers of agricultural land use change in Europe. Landsc. Urban Plan. 133, 24–36.

van Vliet, N., Mertz, O., Heinimann, A., 2012. Trends, drivers and impacts of changes in swidden cultivation in tropical forest-agriculture frontiers: a global assessment. Global Environ. Change 22 (2), 418–429.

Vandergeten, E., Azadi, H., Teklemariam, D., Nyssen, J., Witlox, F., Vanhaute, E., 2016. Agricultural outsourcing or land grabbing: a meta-analysis. Landsc. Ecol. 31 (7), 1395–1417.

Vayda, A.P., 1983. Progressive contextualization: methods for research in human ecology. Hum. Ecol. 11, 265–281.

Vellend, M., Baeten, L., Myers-Smith, I.H., Elmendorf, S.C., Beauséjour, R., Brown, C.D., De Frenne, P., Verheyen, K., Wipf, S., 2013. Global meta-analysis reveals no net change in local-scale plant biodiversity over time. Proceedings of the National Academy of Sciences 110 (48), 19456–19459.

Verburg, P.H., Van De Steeg, J., Veldkamp, A., Willemen, L., 2009. From land cover change to land function dynamics: a major challenge to improve land characterization. J. Environ. Manage. 90 (3), 1327–1335.

Verburg, P.H., Ellis, E.C., Letourneau, A., 2011a. A global assessment of market accessibility and market influence for global environmental change studies. Environ. Res. Lett. 6 (3), 034019.

Verburg, P.H., Neumann, K., Nol, L., 2011b. Challenges in using land use and land cover data for global change studies. Global Change Biol. 17 (2), 974–989.

Walton, D., 2005. Abductive Reasoning. University of Alabama Press.

Young, O.R., Lambin, E.F., Alcock, F., Haberl, H., Karlsson, S.I., McConnell, W.J., Myint, T., Pahl-Wostl, C., Polsky, C., Ramakrishnan, P., Schroeder, H., Scouvart, M., Verburg, P.H., 2006. A portfolio approach to analyzing complex human-environment interactions: institutions and land change. Ecol. Soc. 11 (2), 31. [online] URL. http://www.ecologyandsociety.org/vol11/iss2/art31/.