



# Four Methodological Guidelines to Evaluate the Research Impact of Co-produced Climate Services

Mathilda Englund<sup>1\*</sup>, Karin André<sup>1</sup>, Åsa Gerger Swartling<sup>1</sup> and Jenny Iao-Jørgensen<sup>2</sup>

<sup>1</sup> Stockholm Environment Institute, Stockholm, Sweden, <sup>2</sup> Division of Risk Management and Societal Safety, Faculty of Engineering, Lund University, Lund, Sweden

## OPEN ACCESS

### Edited by:

Scott Bremer,  
University of Bergen, Norway

### Reviewed by:

Rehana Shrestha,  
Leibniz Science Campus Digital Public  
Health Bremen, Germany  
Neha Mittal,  
University of Leeds, United Kingdom

### \*Correspondence:

Mathilda Englund  
mathilda.englund@sei.org

### Specialty section:

This article was submitted to  
Climate Risk Management,  
a section of the journal  
Frontiers in Climate

Received: 31 March 2022

Accepted: 09 June 2022

Published: 22 July 2022

### Citation:

Englund M, André K, Gerger Swartling Å and Iao-Jørgensen J (2022) Four Methodological Guidelines to Evaluate the Research Impact of Co-produced Climate Services. *Front. Clim.* 4:909422. doi: 10.3389/fclim.2022.909422

As climate change impacts unfold across the globe, growing attention is paid toward producing climate services that support adaptation decision-making. Academia, funding agencies, and decision-makers generally agree that stakeholder engagement in co-producing knowledge is key to ensure effective decision support. However, co-production processes remain challenging to evaluate, given their many intangible effects, long time horizons, and inherent complexity. Moreover, how such evaluation should look like is understudied. In this paper, we therefore propose four methodological guidelines designed to evaluate co-produced climate services: (i) engaging in adaptive learning by applying developmental evaluation practices, (ii) building and refining a theory of change, (iii) involving stakeholders using participatory evaluation methods, and (iv) combining different data collection methods that incorporate visual products. These methodological guidelines offset previously identified evaluation challenges and shortcomings, and can be used to help stakeholders rethink research impact evaluation through their complementary properties to identify complex change pathways, external factors, intangible effects, and unexpected outcomes.

**Keywords:** climate adaptation, climate services, decision support, knowledge co-production, transdisciplinary research, participatory research, evaluation method, research impact

## INTRODUCTION

As climate change unfolds across the globe, growing attention is paid toward producing climate services that supports adaptation decision-making (Papathoma-Köhle et al., 2016; Adger et al., 2018). Despite recent advancements in risk and vulnerability assessments, climate impact studies, and adaptation research, the use of such knowledge remains limited in practice (Klein and Juhola, 2014; Bremer and Meisch, 2017; Palutikof et al., 2019). Academia, funding agencies, and decision-makers are increasingly adopting knowledge co-production in order to transcend the divide between academia and practice, and take advantage of potential intangible co-benefits, for example mutual learning, social capital, and institutional capacity (Hansson and Polk, 2018; Bremer et al., 2019; Cvitanovic et al., 2019). This indicates a shift in the role of science in society (Jasanoff, 2004) in which science is held accountable for providing applicable and useful research of societal relevance (Barry et al., 2008; Wiek et al., 2014). The question, however, remains whether co-produced climate services fulfill these accountabilities as evaluations remain rare (Vincent et al., 2018; Daniels et al., 2020). It is still unclear how such co-produced climate services contribute to societal change (Lourenço et al., 2016; Wall et al., 2017). Consequently, funding agencies and

decision-makers lack information to make sound decisions regarding where, or if, to spend their often limited resources to improve co-produced climate services (Vaughan and Dessai, 2014; Lemos et al., 2018; Visman et al., 2022). Evaluation can bridge this gap by contributing to a broader evidence base that can inform future climate service practices to maximize their impact. Hence, evaluations can support and improve climate risk-informed decision support, in the long run increasing the efficiency and effectiveness of climate risk management as a whole (Vaughan and Dessai, 2014; Daniels et al., 2020; Salamanca and Biskupska, 2021).

In this vein, scholars have recently started to outline evaluation practices that are fit for appraising co-produced climate services (see for example Vogel et al., 2017; Wall et al., 2017; Tall et al., 2018; Bremer et al., 2021; Salamanca and Biskupska, 2021; Visman et al., 2022). Many, however, continue to employ traditional evaluation procedures that solely focus on assessing academic outputs, thus failing to capture the many co-benefits that may emerge when co-producing climate services (Sarkki et al., 2015; Schuck et al., 2017). Tracking pathways to the impact of co-produced climate services remain equally limited (Jones et al., 2018), and further research is required to better design evaluation practices to capture long-term impacts and intangible benefits (Daniels et al., 2020). Novel approaches are, therefore, called for. Hence, this paper aims to address this limitation in current research by identifying methodological guidelines that outline approaches fit for evaluating co-produced climate services. We investigate the following research question: What methodological guidelines can be used to evaluate co-produced climate services more effectively? To this end, we review 25 scientific papers in-depth, followed by a survey study targeting actors with experience in co-producing knowledge.

## CONCEPTUAL LANDSCAPE

Climate services first emerged when the World Meteorological Organization (WMO) in collaboration with various UN agencies initiated the World Climate Conference-3 (WCC-3) in 2009 to improve information for decision-making (WMO, 2009). Although still in its infancy, climate services, as a concept, is gaining prominence in the adaptation discourse (Vaughan and Dessai, 2014; Tall et al., 2018; Bremer et al., 2019; Hewitt and Stone, 2021). Climate services are commonly understood as efforts seeking to support climate risk-informed decision-making by providing timely, tailored, and usable knowledge and information (Vincent et al., 2018; Gerger Swartling et al., 2019; Daniels et al., 2020). Although many international organizations present definitions of climate services (see for example WMO, 2009; European Commission, 2015; IPCC, 2018), in practice, climate services tend to be confused with weather forecasts and climate research (Vaughan et al., 2018).

Other constraints further inhibit climate services to fulfill their stated aims, including, for example, a disconnect between stakeholders' expectations, inadequate consideration of stakeholders' differing realities, and data issues (Porter and Dessai, 2017; Ernst et al., 2019; Vogel et al., 2019).

Many scholars attribute these shortcomings to the one-directional delivery of climate services from providers to users, which continue to dominate the field (McNie, 2007; Steynor et al., 2016, 2020). Climate services are supply-driven as to which decision-makers' demand for specific knowledge and information remains lacking (Lourenço et al., 2016), in the end inhibiting decision-makers to take ownership over the climate information and apply it in practice (Dilling and Lemos, 2011; Vaughan and Dessai, 2014). In addition, climate services tend to emphasize tailored products despite that other more intangible outcomes and impacts can be far more important (Daniels et al., 2020; Norström et al., 2020).

For this reason, knowledge co-production is considered a promising approach for making climate services more accessible, relevant, and actionable (Vincent et al., 2018; Bremer et al., 2019; Carter et al., 2019). As a wicked problem, climate adaptation cuts across sectors and disciplines, which calls for a collaborative and interdisciplinary approach that fosters knowledge exchange and action across different stakeholder groups (Cash et al., 2003; Jones et al., 2018; Harvey et al., 2019). Accordingly, academia, decision-makers, and funding agencies suggest that knowledge co-production deserves a central role in the environmental governance discourse (Vincent et al., 2018; Romina and Gerger Swartling, 2019). In broad terms, co-production refers to the process in which researchers and decision-makers collaborate when producing knowledge (Blackstock et al., 2007; Heink et al., 2015; Belcher et al., 2016). Norström et al. (2020) provides a more encompassing definition, in which knowledge co-production implies a collaborative research process involving diverse types of expertise and actors, to solve real-world problems and produce situation-relevant knowledge. In the literature, many benefits are associated with knowledge co-production such as better adaptation decision support, strengthened cross-sectorial networks, improved trust and confidence, increased institutional capacity, and better scientific quality (Bremer et al., 2019; Cvitanovic et al., 2019; Daniels et al., 2020).

There is, however, little existing evidence showing if co-produced climate services deliver on these potential benefits, and whether they are utilized in practice (Swart et al., 2017; VanderMolen et al., 2019). Research impact evaluation can bridge this gap. Numerous definitions of research impact evaluation exist (see Alla et al., 2017 for a review of definitions). For the purpose of this paper, we apply the definition suggested by Reed et al. (2021, p. 3): "the process of assessing the significance and reach of both positive and negative effects of research." Looking at the evaluation literature at large, three main typologies emerge, namely: *summative evaluation* that takes place at the end of an intervention to assess its overall merit; *formative evaluation* that is embedded into the project life cycle to enhance learning with intent to improve project performance; and *developmental evaluation* that offers an ongoing process supporting adaptive management in complex social interventions (Patton, 2006, 2010; Dozois et al., 2010; Mitchell and Lemon, 2020). Research impact evaluation may handle one or more of the following effects: *outputs*, which are the tangible products of the process; *outcomes*, as the less tangible effects and results of the co-production process; and *impacts*, as the long-term effects of

the co-production process (Hassenforder et al., 2015; Wall et al., 2017). In the context of co-produced climate services, scientists and users may have contrasting views on measures of achievement, hence what outcomes and impacts to evaluate. It is, therefore, imperative to consider this multitude of perspectives when evaluating co-produced climate services (Roux et al., 2010; Fazey et al., 2014).

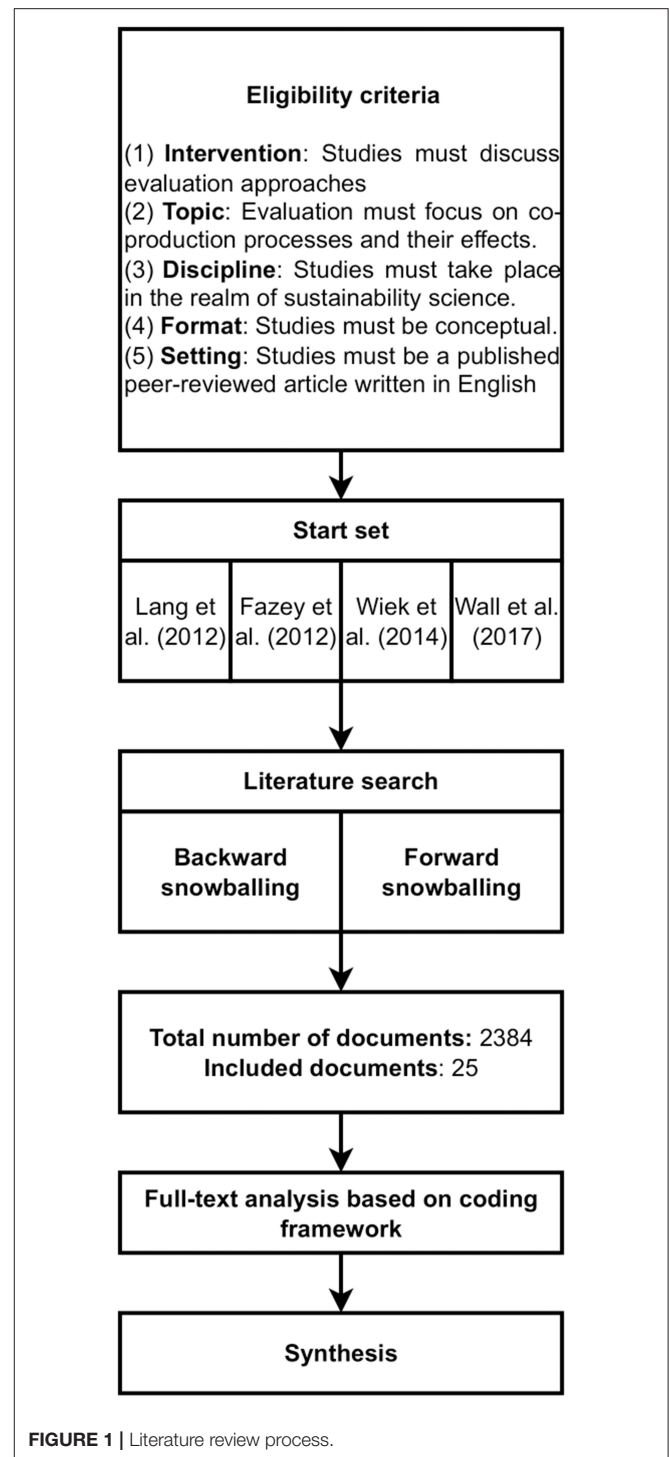
## MATERIALS AND METHODS

To identify methodological guidelines for evaluating co-produced climate services, we first carried out a literature review exploring previous attempts to evaluate co-production processes. We paid special attention to challenges and good practices. We extracted lessons identified, that later were transformed into methodological guidelines. Lastly, through an online survey, we shared among actors with previous experience in co-producing knowledge our proposed methodological guidelines and validated them.

### Literature Review

We first performed a literature review. To do so, we drew inspiration from the systematic snowballing approach outlined by Wohlin (2014) with methodological additions from Haddaway et al. (2015) and Dawkins et al. (2019). Previous research shows that snowballing is equally reliable as the traditional systematic review methods that rely on database searches (Badampudi et al., 2015). Snowballing, however, tends to have higher precision and therefore retrieve much fewer studies to be analyzed, which, therefore, arguably mitigates the risk of human error in comparison to database searches (Felizardo et al., 2016). This adapted approach consisted of five steps, as illustrated in **Figure 1**: (i) determine eligibility criteria, (ii) identify a start set, (iii) literature search applying backward and forward snowballing, (iv) coding and analysis, and (v) synthesis.

As an initial step, we developed a set of eligibility criteria that determined the basic conditions that a document must fulfill for inclusion in the final sample. This represented our attempt to ensure repeatability and reliability (Haddaway et al., 2015; Dawkins et al., 2019). We identified five criteria. First, documents were included if the studies suggested an *evaluation framework, approach, or method*. Second, documents were included if the studies concerned *evaluating knowledge co-production* or adjacent research practices like transdisciplinary research, participatory methods, and science-policy interface. Third, documents were included if the studies were related to *sustainability science*. The sole focus on climate services proved insufficient due to the lack of previous academic literature. Although considered as separate disciplines, climate services and sustainability science share many characteristics due to the many challenges invoked by complexity, uncertainty, and long time horizons. Fourth, documents were included if the studies were *conceptual*. Case studies were at first included, but it turned out that most focused on the co-production initiative itself rather than the evaluation approach. Case studies were, therefore, excluded. Some papers were both conceptual and empirical, as they developed and tested a novel evaluation approach.



These papers were included to gain a better understanding of potential practical challenges and good practices that may arise when evaluating co-production initiatives. Fifth, documents were included if they were published in *peer-reviewed journals* written in *English*.

Next, we selected a start set compliant with the eligibility criteria. We first performed a preliminary literature search to

gain a quick overview of available research. We used a simple search string that reflected the key concepts of interest: TITLE-ABS-KEY (evaluating AND knowledge AND co-production). The preliminary literature search yielded 32 documents, of which four documents were tentatively included. Thereafter we added six documents that were known among the authors but missing from the preliminary literature search. The start set was then reduced in size. Documents with the most citations were selected, in order to provide a larger input for the snowballing. A broad representation of academic journals was also considered. In total, four documents were included. The documents that were initially excluded were later included as they were identified in the literature search. For more details, see the **Supplementary Material**.

We thereafter began the literature search, applying backward and forward snowballing in iterations. Backward snowballing reviewed the reference list of the documents in the start set, whereas citations were considered during the forward snowballing (Wohlin, 2014). Only one author was involved in the screening process. Meetings were held with the remaining co-authors on multiple occasions to ensure consistency. Citations and references lists were identified using the well-regarded database Scopus in October 2020. Documents were included if titles and abstracts met the eligibility criteria. Documents found during the initial iteration were added to the start set, and subject to snowballing during the next iteration. The process continued until no additional documents were found. In total, the literature search generated 2384 documents, of which 70 were screened a second time. In the end, 25 documents were included for full-text analysis (see **Supplementary Material** for more information).

Once selected, documents were coded. A deductive approach was employed, using a pre-defined coding form to ensure consistency and replicability (Haddaway et al., 2015). Three types of codes were considered. First, basic citation information was noted. Second, conceptualizations and approaches to knowledge co-production were registered to avoid any terminological ambiguity. Third, the proposed evaluation design was considered, paying special attention to challenges and good practices. An overview is provided in **Table 1**.

Findings were then synthesized narratively, by taking a textual approach to summarize the findings (Popay et al., 2006). Information was synthesized for each code. Data was clustered into classes of similar objects, which revealed key themes and patterns. We performed some simple statistical analysis for those codes to be easily quantifiable.

Next, we identified methodological guidelines for evaluating co-produced climate services based on the findings from the literature review. We paired all identified challenges with potential solutions outlined in the reviewed literature. Solutions were clustered into groups based on what methods they suggested. Four themes emerged, which were labeled and translated into methodological guidelines. Some additional literature was added at this stage to collect more information about the methods and approaches outlined in the methodological guidelines. Here it is worth noting, in line with Hassel (2010), that there are an infinite number of solutions to a single problem. We, therefore, refrained from making any claims

**TABLE 1** | Coding form.

Code	Type of code
Author(s)	Descriptive text
Title	Descriptive text
Year	Descriptive text
Abstract	Descriptive text
Research design	Descriptive text (case study, literature review, etc.)
Discipline	Descriptive text (sustainable development, natural resource management, climate change adaptation, etc.)
Country/Region/Sub-national	Country/Region/Sub-national area
Empirically tested	Yes, no
Term used	Descriptive text (transdisciplinary, participatory, co-producing, etc.).
Definition	Descriptive text
Theoretical approach	Descriptive text
Effects	Output, outcome, impact
Typology	Summative, formative, developmental
Timing	Pre-assessment, monitoring, retrospective, all
Design	Descriptive text (e.g. qualitative, quantitative, mix-method, participatory, etc.)
Data collection	Descriptive text (workshops, surveys, interviews, knowledge tests, etc.)
Evaluation framework	Descriptive text explaining how the evaluation is approached
Evaluation criteria	Descriptive text presenting and explaining evaluation criteria
Success factors for evaluation	Descriptive text
Challenges for evaluation	Descriptive text
Other useful information	Descriptive text

on presenting an optimal solution. Instead, we aimed to find one possible solution that addresses the methodological challenges that arise when evaluating co-produced knowledge.

## Survey

To increase the reliability of our findings and validate the emerging methodological guidelines, we distributed an online survey to actors with previous experience in co-producing knowledge. Survey responses also provided an in-depth understanding of practical barriers. Before its launch, the survey was tested to identify potential ambiguities. The survey was thereafter launched in February 2021, and remained open for a month. The survey included both qualitative and quantitative questions. Respondents were given the option to answer the questionnaire in Swedish or English, depending on what they felt the most comfortable with.

Respondents were identified through existing networks at the Stockholm Environment Institute (SEI) as well as personal networks through LinkedIn. Three groups were targeted: (1) previous project participants, (2) staff at SEI, and (3) personal networks. Respondents were also asked if they could recommend any other people to respond to the survey. In total, 61 complete responses were collected. 91% of the respondents

self-identified as fulfilling multiple roles when co-producing knowledge – including users, providers, intermediaries, financiers, and evaluators. Respondents represented research institutes or universities (64%), non-governmental organizations (20%), governmental agencies (10%), municipalities (2%), and private consultancies (2%). Among the respondents, different geographical regions were represented – including Europe, Africa, Asia, Latin America, and North America.

We first asked some introductory questions to better understand the respondent's background, role, and experience in co-producing climate services. Subsequently, the survey asked questions that reflected the emerging methodological guidelines, with a focus on understanding potential benefits and barriers in applying the methodological guidelines in practice. For a detailed description, see the **Supplementary Material**.

Once collected, the data was analyzed using Excel. Quantitative data was summarized and visualized in different types of graphs. All numbers were rounded to the nearest integer. Statistical analysis was avoided, as most of the data was structured on an ordinal scale which eliminated most statistical methods (Bryman, 2012). The qualitative data was treated as one cohesive dataset, meaning that significant patterns were identified across the entire dataset rather than for the single questions alone (Braun and Clarke, 2019). We applied an inductive approach in line with Thomas (2006), in which we reduced the data by developing themes based on our interpretations and previous research. We noted the following for each theme: category label, short description, direct quotes, and potential links.

## RESULTS

Results are presented in three parts. First, we present findings from the literature review focusing on current evaluation practices, challenges, and good practices. Based on these findings, we identify four methodological guidelines for evaluating co-produced climate services. Lastly, we present the survey responses to validate the methodological guidelines.

### Literature Review

The 25 studies reviewed in our full-text analysis represent a wide variety of disciplinary fields: sustainability research (24%), natural resource management (20%), environmental science (20%), climate adaptation (4%), socio-ecological research (4%), and climate science (4%). The remaining studies (24%) take an interdisciplinary approach focusing on complex societal and environmental problems in general. None of the reviewed literature focus on climate services.

Following the eligibility criteria, all studies relate to knowledge co-production processes, although the conceptualization of knowledge co-production diverges. The reviewed studies refer to knowledge co-production as transdisciplinary, participatory research, communities of practice, knowledge exchange, joint knowledge production, science-policy interfaces, and knowledge integration.

### Evaluating Co-produced Knowledge

Traditionally, research evaluations employ reductionist procedures solely focusing on assessing academic outputs,

thus inadequately capturing the broad range of effects that can emerge when co-producing knowledge (Sarkki et al., 2015; Zscheischler et al., 2018). The reviewed literature seems to acknowledge this shortcoming, and suggests evaluation approaches that appraise all or a combination of outputs, outcomes, and impacts. None proposes a sole focus on outputs.

Most scholars suggest a formative approach when evaluating co-production endeavors (Jones et al., 2009; Lang et al., 2012; Sarkki et al., 2015), which in turn affects the timing of the evaluation. Integrating evaluation practices into the co-production process allows for reflection and learning, thus providing an opportunity for influencing the direction in which the co-production processes are heading (Roux et al., 2010). It also allows for trust to emerge among the involved actors (Wall et al., 2017). However, ex-post evaluations may be necessary to capture those outcomes and impacts that emerge after the end of a co-production process (Walter et al., 2007).

A systematic review performed by Ernst (2019) shows the many methods used in evaluating co-production initiatives, such as questionnaires, interviews, document analysis, and observation. Looking at the literature, most suggest using a Likert scale questionnaire. These studies tend to also employ evaluation criteria (Walter et al., 2007; van der Wal et al., 2014; Zscheischler et al., 2018; Hitziger et al., 2019; Fulgenzi et al., 2020). Others propose a mixed-method approach sequencing data collection methods to serve a specific purpose at different points of time, arguing that the strengths of one method can offset the weaknesses of another (Jones et al., 2009; Wiek et al., 2014; Holzer et al., 2018). Moreover, some studies advocate for participatory evaluation methods, in order to use the evaluation as an opportunity to further strengthening knowledge exchange (Fazey et al., 2014; Norström et al., 2020).

Evaluation criteria are contested subject, as explained by O'Connor et al. (2019, p.2): “developing evaluation criteria for knowledge co-production remains a challenge because of its variety of forms, contexts, and participants who may have differing views of what is valuable”. Moreover, evaluation criteria are inappropriate when appraising complex systems, as it attempts to fit complexity into a few variables and consequently tends to fall into narrow ranges (Jones et al., 2009; Hassenforder et al., 2015). However, 12 studies present evaluation criteria as they can indicate signs of change and allows for comparison across contexts. Looking at the literature, some studies propose evaluation criteria that assess research quality in terms of relevance, credibility, legitimacy, and effectiveness (Sarkki et al., 2015; Belcher et al., 2016; Knickel et al., 2019). Others suggest evaluation criteria representing the co-production process, its effects, and the context in which it operates (Blackstock et al., 2007; Hassenforder et al., 2015; Jahn and Keil, 2015; Wall et al., 2017; Hitziger et al., 2019). Fulgenzi et al. (2020) identify good practices in knowledge co-production and outline evaluation criteria accordingly, whereas Lang et al. (2012) outline evaluation criteria for assessing the co-production process itself. Lastly, Roux et al. (2010) outline evaluation criteria assessing to what extent funders, researchers, and end-users fulfill their accountabilities when co-producing knowledge.

An overview of the evaluation criteria mentioned in the reviewed literature are presented in **Tables 2–4** and organized

**TABLE 2** | Criteria for evaluating co-produced knowledge – the enabling environment.

Criteria	Description	References
Access to resources	The support to participants for them to meet their responsibilities (competence, time, scientific disciplines in research team, budget, adequate infrastructure, practical information, staffing)	Blackstock et al. (2007), Roux et al. (2010), Lang et al. (2012), Sarkki et al. (2015), Belcher et al. (2016), Wall et al. (2017), Hitziger et al. (2019), Knickel et al. (2019), Fulgenzi et al. (2020)
Drivers	Incentives - demand-driven or supply-driven	Blackstock et al. (2007), Hassenforder et al. (2015), Sarkki et al. (2015), Knickel et al. (2019)
External context	Characteristics of the system in which the co-production process operates, with a focus on complexity, boundaries, synergies, and catalyzing events	Blackstock et al. (2007), Roux et al. (2010), Hassenforder et al. (2015), Jahn and Keil (2015), Belcher et al. (2016), Wall et al. (2017), Hitziger et al. (2019)
Expectations	Whether participants are confident that the process will yield positive effects	Hassenforder et al. (2015)
Institutional memory	Safeguarding mechanisms to protect the acquired collective knowledge	Jahn and Keil (2015), Hitziger et al. (2019), Knickel et al. (2019)
Preexisting relationships	Preexisting professional relationships between involved actors	Wall et al. (2017)
Willingness to learn	The capacity and personal motivation to participate in the co-production process	Blackstock et al. (2007), Walter et al. (2007), Roux et al. (2010), Lang et al. (2012), Jahn and Keil (2015), Sarkki et al. (2015), Wall et al. (2017), Hitziger et al. (2019), Knickel et al. (2019)

into the following: (i) criteria assessing the enabling environment, (ii) criteria assessing the process, and (iii) criteria assessing the effects.

## Challenges

A major challenge in evaluating knowledge co-production is the *complexity* of the process itself and of the system in which it operates (Roux et al., 2010; Lang et al., 2012; Fazey et al., 2014). Complex systems are characterized by non-linearity, multi-pathways, emergent properties, dynamic change, and interdependencies (Zscheischler et al., 2018), which in combination with the *long timeframes* in adaptation decision-making makes it *difficult to establish causality* (Jahn and Keil, 2015; Hitziger et al., 2019; Norström et al., 2020). Making mono-causal connections are further inhibited by the ongoing influence of *unforeseeable external factors* (Zscheischler et al., 2018). Trying to fit complexity into a few variables can cause distortion as complex problems are greater than the sum of their parts (Jones et al., 2009; Hassenforder et al., 2015). Tensions arise when trying to apply linear frameworks to capture change that occurs in a messy and complex reality (Walter et al., 2007).

Co-production is subject to *uncertainty* as objectives and practices tend to adapt as the process evolves (Laycock et al., 2019). In addition, uncertainty is inherent to the problem, namely climate change, which is being addressed (Hegger et al., 2012). This poses significant challenges for research impact evaluation, as success is defined in relation to formulated objectives.

In addition, the *intangible* nature of many key elements in knowledge co-production, such as learning, empowerment and trust, further complicates evaluation efforts. These intangible effects are difficult to objectively judge, and therefore tend to rely on subjective estimations (Blackstock et al., 2007; Hassenforder et al., 2016).

Moreover, evaluations are expected to yield different results depending on their timing (Wall et al., 2017; Fulgenzi et al., 2020). Outcomes and impacts emerge at different points in time

(Roux et al., 2010; Ernst, 2019). There are significant *time-lags between causes and effects* as societal impacts evolve over a long period of time (Blackstock et al., 2007; Jahn and Keil, 2015; Wall et al., 2017), making them difficult to capture within the timeframe provided in an externally funded project (Norström et al., 2020).

Furthermore, co-production initiatives involve *stakeholders with different backgrounds* (Jones et al., 2009; Roux et al., 2010; Hitziger et al., 2019), which may complicate evaluation practices due to at times contrasting values, epistemological beliefs, educational background, professional jargons, and objectives. Motivation can also vary among involved stakeholders. Some might consider evaluations as a burden that distracts from the main co-production activities, especially if they are struggling with limited financial resources (Knickel et al., 2019).

## Good Practices

*Flexible practices* are considered key when evaluating co-production processes (Lang et al., 2012; van der Wal et al., 2014; Knickel et al., 2019). Evaluation frameworks must be adapted to the needs of the intended users, considering timing, purpose, scale, and context (Belcher et al., 2016; Knickel et al., 2019). Furthermore, evaluation strategies should adapt and adjust as new insights arise (Blackstock et al., 2007; Carew and Wickson, 2010; Belcher et al., 2016). Evaluation objectives should be revisited and adapted as new information emerges (Norström et al., 2020).

Walter et al. (2007) call for novel evaluation approaches, of which *participatory evaluation* is a promising alternative (Fazey et al., 2014; Norström et al., 2020). Participatory evaluation encourages learning, ultimately transforming the evaluation into a learning activity in itself (Lang et al., 2012). In relation to this, stakeholder engagement is especially important when deciding on evaluation objectives to encourage ownership and buy-in from the involved stakeholders while ensuring contextual relevance (Wiek et al., 2014). It is also suggested to involve stakeholders

**TABLE 3** | Criteria for evaluating co-produced knowledge – the process.

Criteria	Description	References
Awareness	The extent to which participants can identify available resources and possible gaps	Fulgenzi et al. (2020)
Capacity development	The process of developing skills, knowledge, and awareness	Blackstock et al. (2007), Roux et al. (2010), Sarkki et al. (2015), Wall et al. (2017), Fulgenzi et al. (2020)
Co-location	Whether involved actors are willing to host junior researchers	Roux et al. (2010)
Conflict resolution	The degree of conflicts between participants, and the ability to manage such conflicts	Blackstock et al. (2007), Roux et al. (2010), Lang et al. (2012), Wall et al. (2017), Hitziger et al. (2019), Knickel et al. (2019)
Continuity	The consistency in participation, in terms of recurring participants	Wall et al. (2017)
Cost effectiveness	Cost of achieving identified objectives	Blackstock et al. (2007), Jahn and Keil (2015)
Effective collaboration	Mechanisms promoting collaboration (research plan, documentation, agenda, roles and responsibilities)	Belcher et al. (2016), Wall et al. (2017), Knickel et al. (2019), Fulgenzi et al. (2020)
Effective communication	Appropriateness, relevance, clarity, and accessibility of communication efforts	Walter et al. (2007), Roux et al. (2010), Jahn and Keil (2015), Sarkki et al. (2015), Belcher et al. (2016), Wall et al. (2017), Hitziger et al. (2019), Knickel et al. (2019)
Ethical aspects	Whether research adheres to ethical standards	Belcher et al. (2016)
Inclusion of all relevant perspectives	The creation of a safe space for participants to voice their opinions and influence the decision-making process	Blackstock et al. (2007), Walter et al. (2007), Hassenforder et al. (2015), Hitziger et al. (2019), Knickel et al. (2019), Fulgenzi et al. (2020)
Involvement	At what stage of the process different participants are engaged	Knickel et al. (2019)
Leadership	A leadership figure facilitating the process	Blackstock et al. (2007), Walter et al. (2007), Roux et al. (2010), Lang et al. (2012), Jahn and Keil (2015), Sarkki et al. (2015), Wall et al. (2017), Hitziger et al. (2019), Knickel et al. (2019)
Methods	Appropriateness of the selected disciplines, epistemology, methods, approaches, and theories	Blackstock et al. (2007), Lang et al. (2012), Hassenforder et al. (2015), Jahn and Keil (2015), Belcher et al. (2016), Wall et al. (2017), Hitziger et al. (2019), Knickel et al. (2019)
Objectives	Whether goals are established for the co-production process	Blackstock et al. (2007), Walter et al. (2007), Lang et al. (2012), Hassenforder et al. (2015), Jahn and Keil (2015), Sarkki et al. (2015), Belcher et al. (2016), Wall et al. (2017), Hitziger et al. (2019), Knickel et al. (2019)
Practicalities	Frequency of events, number of participants, and setting of exchange	Hassenforder et al. (2015), Hitziger et al. (2019)
Reflection	Opportunities to reflect upon the collective experience and adjust the plan accordingly	Lang et al. (2012), Jahn and Keil (2015), Sarkki et al. (2015), Belcher et al. (2016), Wall et al. (2017), Hitziger et al. (2019), Knickel et al. (2019), Fulgenzi et al. (2020)
Relationships	The improvements in social capital, and the development of new social networks	Blackstock et al. (2007), Walter et al. (2007), Lang et al. (2012), Sarkki et al. (2015), Knickel et al. (2019), Fulgenzi et al. (2020)
Representation	The genuine inclusion of a diverse set of actors	Blackstock et al. (2007), Lang et al. (2012), Hassenforder et al. (2015), Jahn and Keil (2015), Sarkki et al. (2015), Belcher et al. (2016), Wall et al. (2017), Hitziger et al. (2019), Knickel et al. (2019), Fulgenzi et al. (2020)
Theory of change	The development of a theory of change that matches the objectives and clarifies any underlying assumptions	Jahn and Keil (2015), Belcher et al. (2016), Hitziger et al. (2019)
Transparency	The extent to which participants and observers can understand the process	Blackstock et al. (2007), Knickel et al. (2019)

when developing evaluation criteria to ensure that effects perceived as important are being considered (Fazey et al., 2014).

Additionally, scholars suggest considering the *evaluation as a process*. It is proposed to integrate the evaluation from the start to allow for social learning and trust to emerge (Roux et al., 2010; Wall et al., 2017). The evaluation should aim to be *comprehensive* to capture both the co-production process itself, and its expected and unexpected outputs, outcomes, and impact (Fazey et al., 2014; Belcher et al., 2016; Wall et al., 2017). Intangible aspects should also be assessed, despite being difficult to measure (Norström et al., 2020).

It is recommended to develop a *theory of change*, which is a logic model supporting project management, stakeholder engagement, and evaluation practices which seeks to describe how change is expected to occur. Theory of change offers greater flexibility in comparison to other logic models and can capture complexity, clarify causal linkages, and bridge conflicting interests (Fazey et al., 2014; Wall et al., 2017; Knickel et al., 2019; Norström et al., 2020). Others suggest using visual products to encourage meaningful discussions among involved stakeholders, as it can help to overcome barriers like differences in educational backgrounds and language preferences

**TABLE 4** | Criteria for evaluating co-produced knowledge – the effects.

Criteria	Description	References
Accountability	The extent to which participants have satisfied their personal core responsibilities	Blackstock et al. (2007),
Inspiration	The motivation to pursue follow-up projects	Wall et al. (2017), Fulgenzi et al. (2020)
Outcomes	Whether desired change is achieved	Lang et al. (2012), Wall et al. (2017)
Outputs	The timely delivery of the tangible products (peer-reviewed articles, workshops, meetings, reports)	Roux et al. (2010), Hassenforder et al. (2015), Wall et al. (2017), Knickel et al. (2019)
Quality of decision-making	The implementation, integration, and maintenance of findings	Blackstock et al. (2007), Roux et al. (2010), Jahn and Keil (2015), Knickel et al. (2019)
Quality of research product	Validity of the final research product, considering its legitimacy, transferability, credibility, comprehensiveness, and robustness	Blackstock et al. (2007), Sarkki et al. (2015), Wall et al. (2017)
Recognized impacts	Perceived changes associated with the co-production process (unintended effects, changes in perspectives, and improved organizational performance)	Blackstock et al. (2007), Hassenforder et al. (2015), Jahn and Keil (2015), Fulgenzi et al. (2020)
Relevance to society	Whether research findings are used in practice to solve the targeted problem	Roux et al. (2010), Lang et al. (2012), Jahn and Keil (2015), Sarkki et al. (2015), Belcher et al. (2016), Wall et al. (2017), Knickel et al. (2019)
Social learning	Changes collective culture and values	Blackstock et al. (2007), Jahn and Keil (2015), Wall et al. (2017), Hitziger et al. (2019), Knickel et al. (2019)

(Lang et al., 2012). Evaluation practices should also allow for *maximum participation*, and *adjust for potential memory distortion* (Wiek et al., 2014).

## Methodological Guidelines

Reviewing the literature, many insights strike as relevant although none addresses climate services. Challenges associated with complexity, long time horizons, uncertainty, and stakeholder diversity are inherent to many co-production processes, and thus cut across disciplinary boundaries. Drawing from the literature review, we identify four methodological guidelines fit for evaluating co-produced climate services. An overview is provided in **Table 5**.

### Engaging in Adaptive Learning by Applying Developmental Evaluation Practices

As its name suggests, developmental evaluation puts emphasis on development rather than accountability or improvement (Mitchell and Lemon, 2020). Developmental evaluation seeks to support adaptive management, allowing practices to adapt as new insights emerge or circumstances change (Patton, 2010). Developmental evaluation rests on the same assumptions that underpin knowledge co-production initiatives. Co-production builds on the assumption that change is complex, non-linear, and emergent (Norström et al., 2020), and developmental evaluation is designed to understand such complexity. Drawing inspiration from complexity theory, developmental evaluation sets to support adaptive management in social innovation initiatives, like when co-producing climate services, characterized by complexity, emergence, stakeholder diversity, long time horizons, and uncertainty (Dozois et al., 2010; van Tulder and Keen, 2018).

Developmental evaluation is per design flexible, thus offsetting challenges encountered when applying summative and formative assessment approaches. Summative evaluations

**TABLE 5** | Overview of the methodological guidelines.

Methodological guidelines	Justification
Engaging in adaptive learning by applying developmental evaluation practices	<ul style="list-style-type: none"> <li>• To handle complexity and uncertainty by engaging stakeholders in adaptive management.</li> <li>• Continuous evaluation process to capture change as it emerges.</li> </ul>
Building and refining a theory of change	<ul style="list-style-type: none"> <li>• To capture complexity.</li> <li>• Easy to update as outcomes unfold over time.</li> <li>• To take a system perspective to capture external factors.</li> <li>• To allow stakeholders to reflect upon causal linkages.</li> </ul>
Involving stakeholders using participatory evaluation methods	<ul style="list-style-type: none"> <li>• To develop a shared problem understanding.</li> <li>• To ensure that all perspectives are equally considered and represented.</li> <li>• To allow stakeholders to draw attention to unexpected outcomes.</li> <li>• To develop a shared vocabulary to overcome any professional jargon.</li> </ul>
Combining different data collection methods that incorporate visual products	<ul style="list-style-type: none"> <li>• To create a robust data set for analysis.</li> <li>• To provide a more comprehensive understanding of the problem at hand.</li> <li>• To capture both tangible and intangible effects.</li> <li>• To make it easier to grasp complexity.</li> </ul>

aim for predictability by using a linear cause-effect model and rigid methods (Fazey et al., 2014), which is considered unfit for evaluating co-produced climate services as they require flexible practices that can adapt as uncertainties and complexity unfold (Salamanca and Biskupska, 2021). Similarly, formative evaluations also prove inadequate in terms of flexibility, as they seek to support improvements toward a pre-defined objective



(Patton, 2010). Co-production initiatives tend to change their objectives and practices as the process evolves (Blackstock et al., 2007), thus making it unfeasible to try to measure success against a set of pre-defined objectives. Instead, developmental evaluation promotes adaptive management in order for evaluation practices to adapt to changes in objectives, research design, or stakeholder constellation. In practice, developmental evaluation support adaptive management by engaging stakeholders in an ongoing evaluation process in which an embedded evaluator provides actionable feedback to facilitate continuous learning (Patton, 2006, 2010).

Adaptive management can serve as a vehicle for joint action, in which stakeholders can bring their experience and feedback into action and adjust evaluation practices accordingly (Reynolds, 2014; Gerlak et al., 2018). Developmental evaluation supports a shift from linear single-loop thinking toward transformative double-loop learning in which adaptive management allows evaluation practices, objectives, or metrics of success to change in response to experience. Double-loop learning is fit if adapting to uncertainty or complexity, as it supports transformation rather than retainment (Shea and Taylor, 2017). Arguably, climate services can benefit from adaptive management and double-loop learning by helping the involved stakeholders to navigate the inherent uncertainty, complexity, and long time horizons associated with adaptation decision-making, as it allows stakeholders to adjust their evaluation practices to emerging and changing contexts.

### Building and Refining a Theory of Change

Theory of change is increasingly used to inform baseline studies, organizational design in complex and multi-stakeholder settings, and to facilitate adaptive learning from a systems perspective throughout a project life cycle (van Es et al., 2015). Theory of change is designed to support interventions subject to complexity and uncertainty, which makes it fit for co-production processes that address climate risks. Furthermore, climate services cannot be considered in isolation from the context they operate in as decision-makers combine different sources of information when planning for adaptation (Zscheischler et al., 2018; André et al., 2021). Theory of change acknowledges these external influences by identifying and monitoring them, ultimately strengthening any causal claims (van Es et al., 2015).

Climate risks operate on long timescales and so is adaptation decision-making, meaning that benefits emerging from climate services might appear far in the future. We, therefore, argue for considering the theory of change as a living entity that can track progress at different temporal scales. It is iterative, thus expected to be revisited and refined on a regular basis as new information emerges. In this way, the theory of change become more informed over time, as it enjoys continuous refinement (van Tulder and Keen, 2018). When employed or applied iteratively, the theory of change can capture development that occurs over long periods of time and help involved stakeholders, if resources allow, to continue their evaluation efforts after the end of the co-production process. The theory of change is widely, although not exclusively, used to support ex-post evaluations in explaining how change has happened, as it puts a structure in

place for stakeholders to continue evaluating impacts as they unfold (Vogel, 2012; van Es et al., 2015; Mayne, 2017).

### Involving Stakeholders Using Participatory Evaluation Methods

In short, participatory evaluation is an approach for involving stakeholders in the evaluation process (Trimble and Plummer, 2019). Stakeholders can be involved at any stage of the evaluation (Guijt, 2014; Reed et al., 2021). Participatory evaluation and knowledge co-production have the same theoretical and epistemological underpinnings. Our study reveals many overlaps, where participatory evaluation can reinforce many of the positive outcomes and impacts that emerge when co-producing knowledge. Benefits include helping diverse stakeholder groups to form a shared vision and vocabulary (Plottu and Plottu, 2011; Fazey et al., 2014); enhancing motivation and buy-in among involved stakeholders (Fazey et al., 2014); drawing attention to unexpected outcomes and impacts (Norström et al., 2020); and validating evaluation findings among involved stakeholders (Guijt, 2014). In addition, stakeholder participation can improve overall robustness by incorporating multiple sources of knowledge and realities (van Es et al., 2015).

Evaluation findings can have a transformational capacity if being integrated in iteration. Participatory evaluation methods can be instrumental in strengthening the evaluation's utilization. Participatory methods encourages ownership by stakeholders involved in the generation and use of climate services. This ownership contributes to sustainability beyond the limited time span of climate service projects (Patton and Horton, 2009; Fazey et al., 2014; van Es et al., 2015).

### Combining Different Data Collection Methods That Incorporate Visual Products

Mixed-method approaches combine qualitative and quantitative methods, in order to take advantage of their respective strengths while counterbalancing any potential weaknesses (Ernst, 2019). Methods can be sequenced to serve a specific purpose at different points in time (Jones et al., 2009; Holzer et al., 2018), thus forming a comprehensive understanding of the process itself and its outputs, outcomes, and impact. On the one hand, qualitative methods are well suited to explore the many intangible effects that emerge when co-producing knowledge, such as social learning, empowerment, and trust (Fazey et al., 2014). Qualitative methods expect the unexpected, and allow the involved stakeholders to draw attention to any unexpected positive or negative effects (Bryman, 2012). On the other hand, quantitative methods can assess how change unfolds over time by employing longitudinal data (Fazey et al., 2014), which is especially appropriate considering the long time horizon that characterizes adaptation decision-making. Quantitative methods can also improve the generalizability of the evaluation findings, and thus identify transferable lessons. Additional benefits of using mix-methods include allowing for triangulation; increasing robustness; enhancing comprehensiveness; improving credibility and validity of findings; and generating unexpected insights (Reed et al., 2021).

In addition, art-based methods generate tangible products for expression and analysis, which can enhance mutual learning among the involved stakeholders (Chambers, 2008). Visualization complements text and dialogue (van Es et al., 2015). Art-based methods can generate products that act as boundary objects, thus helping to bridge diverging stakeholders' interests, goals, epistemologies, expertise, and languages (Wyborn, 2015). Boundary objects, such as visual products, can enhance meaningful participation (Nel et al., 2016; Reed et al., 2021). Discussing while drawing can create an informal and inclusive setting for knowledge exchange (van Es et al., 2015). As phrased by Chambers (2008, p. 100), "Hands are freer to move tangibles than mouths are to speak words." Visual products can stimulate discussions on the topic of interest, ultimately improving both the quantity and quality of the collected data (Petheram et al., 2012). In addition, visual products can disentangle and represent the complexity present when co-producing climate services, and thus provide a better understanding of causal linkages and change pathways (Chambers, 2008; van Es et al., 2015; Reed et al., 2021). Lastly, visual products can help communicating evaluation findings to a broader audience including new project members (Petheram et al., 2012).

## Validation – Survey Results

To a great extent, the survey responses confirmed the methodological guidelines. However, the survey study revealed a number of benefits and challenges if the respondents were to apply the methodological guidelines. An overview is provided in **Table 6**.

In the open-ended questions, respondents refer to good practices in line with developmental evaluation. Frequently mentioned examples include:

- Utilization-focused approaches to ensure usefulness for intended users;
- Adaptive management to support continuous improvement and social learning; and,
- Importance of reflexive practices.

As such, developmental evaluation presents many benefits when evaluating co-produced knowledge. Barriers related to time allocation and funding are, however, noted.

Overall, 66% are familiar with building a theory of change of which around half recommend it in the case of co-production endeavors. Many benefits are identified, including clarifying underlying assumptions, mapping cause and effect pathways, disentangling complexity and context, and defining objectives. Challenges do, however, exist. Many respondents are unfamiliar with the concept. Others argue that the theory of change is "too abstract," "too academic," "bulky," and even "pointless." Others compare the theory of change with the logical framework approach, criticizing it for being donor-driven and reductionist.

In total, 97% of the respondents recommend using participatory evaluation methods. Participatory evaluation can yield many benefits, including forming a common understanding and vision, building trust, validating evaluation findings, and increasing buy-in and ownership among involved stakeholders. Survey responses indicate that stakeholder engagement is

**TABLE 6** | Overview of survey responses.

Methodological guidelines	Main points from survey responses
Engaging in adaptive learning by applying developmental evaluation practices	<p><b>Benefits</b></p> <ul style="list-style-type: none"> <li>• Utilization-focused approach can increase usefulness</li> <li>• Adaptive management can promote learning</li> <li>• Reflexive practices are important</li> </ul> <p><b>Challenges</b></p> <ul style="list-style-type: none"> <li>• Time allocation and funding</li> </ul>
Building a theory of change	<p><b>Benefits</b></p> <ul style="list-style-type: none"> <li>• Clarify assumptions</li> <li>• Understand causal linkages</li> <li>• Disentangle complexity</li> <li>• Define objectives</li> </ul> <p><b>Challenges</b></p> <ul style="list-style-type: none"> <li>• Difficult to use</li> <li>• Donor-driven</li> <li>• Reductionist</li> </ul>
Involving stakeholders using participatory evaluation methods	<p><b>Benefits</b></p> <ul style="list-style-type: none"> <li>• Forming a shared understanding and vision</li> <li>• Build trust</li> <li>• Validating findings</li> <li>• Increase buy-in and ownership among involved stakeholders</li> </ul> <p><b>Challenges</b></p> <ul style="list-style-type: none"> <li>• Biases</li> <li>• Time-consuming</li> <li>• Trade-offs between validating findings and building ownership</li> </ul>
Combining different data collection methods that incorporate visual products	<p><b>Benefits</b></p> <ul style="list-style-type: none"> <li>• Clarify complex issues</li> </ul>

possible at all stages of the evaluation process, in particular when defining the objectives, developing indicators, and reporting the findings. Nonetheless, some challenges are mentioned. One respondent claim that personal involvement can create biases. Others note that stakeholder involvement is time-consuming, and that extensive participation paradoxically can lower engagement. There are also trade-offs between validating findings on one hand, and building ownership and buy-in on the other.

Many methods are considered useful when evaluating co-production initiatives, including interviews, mixed-methods, group discussions, questionnaires, written reflections, indicators, and document review. 98% agreed that visual products can clarify complex issues.

## DISCUSSION

### Research Implications

Despite recent advances in climate services, research is thus far paying little attention to the evaluation of such services. Many methods exist for evaluating research impact. However, few consider climate services and their impact on adaptation policy and action. Usability is rarely assessed. We address this gap by introducing four methodological guidelines that may serve as stimuli for further discussions on how to evaluate

co-produced climate services. In line with previous research (Sarkki et al., 2015; Belcher et al., 2016; Zscheischler et al., 2018), we argue that novel evaluation practices are needed to capture the broad array of effects that emerge when co-producing knowledge. The proposed methodological guidelines support a shift of evaluation approach from traditional practices emphasizing academic outputs to one that capture the many, often intangible or unexpected, effects that emerge, when co-producing climate services.

Our methodological guidelines add to the body of research that seeks to evaluate research impact and co-produced climate services, and shed light upon the need to rethink evaluation practices. Most previous research has focused on suggesting criteria for evaluating co-produced climate services and adaptation (Wall et al., 2017; Visman et al., 2022) as well as their quality (Bremer et al., 2022). Methodological choices remain understudied. In line with previous research (Walter et al., 2007; Jones et al., 2009; Hassenforder et al., 2015), we acknowledge that metrics and criteria themselves are insufficient when evaluating co-produced climate services. Objectives and strategies tend to change as the co-production process evolves (Laycock et al., 2019), and stakeholders may have contrasting views on what constitutes “success” depending the context in which they operate (Vincent et al., 2020). In this vein, the proposed methodological guidelines support flexible practices and address the challenges that arise when using predefined metrics and criteria in value laden and complex co-production processes.

We believe that the methodological guidelines are applicable in co-production processes that are developed for different purposes from climate services. The methodological guidelines draw on evidence from the broader sustainability literature (Blackstock et al., 2007; Carew and Wickson, 2010), suggesting that they also may prove applicable in such contexts. Sustainability science faces similar challenges as climate services when being evaluated, including complexity, uncertainty, and long time horizons. The methodological guidelines can offset these challenges, and thus support the many science-policy interfaces taking place amid complex socioenvironmental systems.

## Applicability of the Four Methodological Guidelines

The four identified methodological guidelines are designed to fit a broad array of contexts, which enable effective application in a variety of climate service initiatives regardless of their scope, topic, and resources. As demonstrated in the survey responses, applying the methodological guidelines could improve evaluation practices by yielding multiple benefits, such as capturing both tangible and intangible effects; managing complexities and uncertainties; monitoring external factors; bridging stakeholder interests; and better representing causal linkages.

While the methodological guidelines can be applied in isolation, we suggest combining them as they are designed to complement each other. Together, they address all identified challenges that emerge when evaluating co-produced knowledge

(see **Table 5**). Moreover, significant overlaps exist between the guidelines, suggesting they can reinforce and perpetuate one another’s positive impacts. For example, developmental evaluation can be introduced to support an adaptive use of the theory of change, allowing it to be refined as change unfolds. A theory of change is better constructed when taking a participatory approach as it allows stakeholders to form a consensus representing the multitude of perspectives involved. Furthermore, a theory of change is best presented as a visual product together with qualitative or quantitative indicators. Visual products tend to be participatory by nature, allowing stakeholders to engage around a boundary object.

## Challenges Applying the Methodological Guidelines

The survey responses shed light on some new challenges not being addressed in the reviewed literature. There appears a significant gap between theory and practice, indicating that current evaluation practices tend to neglect the contextual realities faced by involved stakeholders.

Looking at the survey responses, many are unfamiliar with the theory of change evaluation approach while others regard it as being difficult, academic, reductionist, or donor-driven. The theory of change approach seems to encounter the same shortcomings as other logic models in its practical application, although the reviewed literature makes a clear distinction between the two (Fazey et al., 2014; van Es et al., 2015). A probable reason for this, supported by findings from the survey, is the limited time and budget allocated for reflection and learning. van Es et al. (2015) argue that reflection is key when building a theory of change. Nevertheless, in practice, stakeholders face budgetary and time constraints that inhibit such critical reflection. Arguably, in line with developmental evaluation, there is a need to embed the reflection process into the evaluation cycle to encourage reflexive learning, ultimately stimulating the many benefits associated with building a theory of change.

It is evident in our study that challenges also arise in relation to stakeholder engagement. Participatory evaluations are no silver bullet, and must be adapted to the context at hand. As noted in the survey responses, participatory evaluations are time- and resource-intensive. Extensive participation can cause fatigue and lower engagement. Our findings indicate that funders sometimes require extensive stakeholder participation without fully grasping the research context and conditions, while researchers and practitioners express a lack of budget and time to engage in such activities. There seems to be a disconnect between funders’ expectations and practical realities, highlighting the importance of flexible funding conditions that stimulate adaptive management.

## CONCLUSION

As climate change continues to alter weather patterns, there is a growing need for climate services to support adaptation policy and action. Climate services are, however, rarely

evaluated. This paper addresses current evaluation challenges and opportunities, by identifying methodological guidelines that outline methods and approaches fit for evaluating co-produced climate services. Based on a literature review and survey responses, the following methodological guidelines are identified: (i) engaging in adaptive learning by applying developmental evaluation practices, (ii) building and refining a theory of change, (iii) involving stakeholders using participatory evaluation methods, and (iv) combining different data collection methods that incorporate visual products. Our study indicates that the proposed methodological guidelines can offer significant benefits when evaluating co-produced climate services, such as helping stakeholders to map complex change pathways; capturing external influences; measuring the intangible; bridging conflicting interests; identifying unexpected effects; enhancing usefulness and learning; clarifying underlying assumptions; increasing ownership and buy-in; understanding causal linkages; and building trust.

Our study makes a significant contribution to a better understanding of what methods can be used when evaluating co-produced climate services, hence, marks a step toward improved research impact evaluation. Future empirical testing is, however, required to assure that the proposed methodological guidelines are feasible in practice. We recommend applying these guidelines in an array of empirical contexts to test their applicability in various stakeholder constellations and situations, and thus stimulate further refinement. Future research can engage with a growing body of developmental evaluation literature for cross-learning of methodological challenges and good practices.

Our study shows that evaluation is essential to enhance research impact of climate services, as it can reveal strengths and weaknesses of the current approaches and pave the way for more effective, user-oriented, and demand-driven climate services. Improved evaluation practices can ultimately increase the effectiveness and efficiency of climate services, thus equipping decision-makers with improved climate risk information and assessments. Most importantly, this can better inform the adaptation efforts urgently needed to combat climate change.

## REFERENCES

- Adger, W. N., Brown, I., and Surminski, S. (2018). Advances in risk assessment for climate change adaptation policy. *Philos. Trans. A Math. Phys. Eng. Sci.* 376, 1–13. doi: 10.1098/rsta.2018.0106
- Alla, K., Hall, W. D., Whiteford, H. A., Head, B. W., and Meurk, C. S. (2017). How do we define the policy impact of public health research? A systematic review. *Health Res. Policy Syst.* 15, 84. doi: 10.1186/s12961-017-0247-z
- André, K., Järnberg, L., Gerger Swartling, Å., Berg, P., Segersson, D., Amorim, J. H., et al. (2021). Assessing the quality of knowledge for adaptation—experiences from co-designing climate services in Sweden. *Front. Clim.* 3, 1–12. doi: 10.3389/fclim.2021.636069
- Badampudi, D., Wohlin, C., and Petersen, K. (2015). *Experiences from Using Snowballing and Database Searches in Systematic Literature Studies. Article No. 17*. Available online at: <http://urn.kb.se/resolve?urn=urn:nbn:se:bth-11371> (accessed June 27, 2022).
- Barry, A., Born, G., and Weszkalnys, G. (2008). Logics of interdisciplinarity. *Econ. Soc.* 37, 20–49. doi: 10.1080/03085140701760841

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## AUTHOR CONTRIBUTIONS

ME: conceptualization, methodology, investigation, and writing—original draft preparation. KA, ÅGS, and JI-J: conceptualization, writing—review and editing, and supervision. All authors contributed to the article and approved the submitted version.

## FUNDING

This work builds on insights and data previously published in Englund (2021). This research deriving from the UNCHAIN project is funded through a collaboration between the EU funding mechanisms Joint Programming Initiative (JPI) and Assessment of cross (X)-sectoral climate impacts and pathways for Sustainable transformation (AXIS). All partners are granted financial support through their national funding agency, of which the Stockholm Environment Institute received support from FORMAS (SE) (2018-02737) and the EU (Grant number: 776608).

## ACKNOWLEDGMENTS

We thank the survey respondents for their inputs to this paper. We are grateful to Katrin Danerlöv from the Stockholm Environment Institute for contributing with her knowledge to the survey design. Lastly, we thank the reviewers for their time in providing comments to this paper.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fclim.2022.909422/full#supplementary-material>

- Belcher, B. M., Rasmussen, K. E., Kemshaw, M. R., and Zornes, D. A. (2016). Defining and assessing research quality in a transdisciplinary context. *Res. Eval.* 25, 1–17. doi: 10.1093/reseval/rvv025
- Blackstock, K. L., Kelly, G. J., and Horsey, B. L. (2007). Developing and applying a framework to evaluate participatory research for sustainability. *Ecol. Econ.* 60, 726–742. doi: 10.1016/j.ecolecon.2006.05.014
- Braun, V., and Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qual. Res. Sport Exerc. Health* 11, 589–597. doi: 10.1080/2159676X.2019.1628806
- Bremer, S., and Meisch, S. (2017). Co-production in climate change research: Reviewing different perspectives. *Wiley Interdiscip. Rev. Clim. Change*, 8, 1–22. edswsc. doi: 10.1002/wcc.482
- Bremer, S., Wardekker, A., Baldissera Pacchetti, M., Bruno Soares, M., and van der Sluijs, J. (2022). Editorial: High-Quality Knowledge for Climate Adaptation: Revisiting Criteria of Credibility, Legitimacy, Salience, and Usability. *Front. Clim.* 4, 1–3. doi: 10.3389/fclim.2022.905786
- Bremer, S., Wardekker, A., Dessai, S., Sobolowski, S., Slaattelid, R., and van der Sluijs, J. (2019). Toward a multi-faceted conception of co-production of climate services. *Clim. Serv.* 13, 42–50. doi: 10.1016/j.cliser.2019.01.003

- Bremer, S., Wardekker, A., Jensen, E. S., and van der Sluijs, J. P. (2021). Quality Assessment in Co-developing Climate Services in Norway and the Netherlands. *Front. Clim.* 3, 627665. doi: 10.3389/fclim.2021.627665
- Bryman, A. (2012). *Social Research Methods, 4th ed.* Oxford: Oxford University Press.
- Carew, A. L., and Wickson, F. (2010). The TD wheel: a heuristic to shape, support and evaluate transdisciplinary research. *Futures* 42, 1146–1155. doi: 10.1016/j.futures.2010.04.025
- Carter, S., Steynor, A., Vincent, K., Visman, E., and Lund Waagsaether, K. (2019). *Co-production in African weather and climate services.* Future Climate for Africa and Weather and Climate Information Services for Africa. Available online at: <https://futureclimateafrica.org/coproduction-manual/downloads/WISER-FCFA-coproduction-manual.pdf> (accessed June 27, 2022).
- Cash, D., Clark, W., Alcock, F., Dickson, N. M., Eckley, N., and Jäger, J. (2003). Saliency, Credibility, Legitimacy and Boundaries: Linking Research, Assessment and Decision Making. *KSG Working Papers Series RWP*.
- Chambers, R. (2008). *Revolutions in Development Inquiry.* London: Earthscan.
- Cvitanić, C., Howden, M., Colvin, R. M., Norström, A., Meadow, A. M., and Addison, P. F. E. (2019). Maximising the benefits of participatory climate adaptation research by understanding and managing the associated challenges and risks. *Environ. Sci. Policy* 94, 20–31. doi: 10.1016/j.envsci.2018.12.028
- Daniels, E., Bharwani, S., Gerger Swartling, Å., Vulturius, G., and Brandon, K. (2020). Refocusing the climate services lens: Introducing a framework for co-designing “transdisciplinary knowledge integration processes” to build climate resilience. *Clim. Serv.* 19, 1–15. doi: 10.1016/j.cliser.2020.100181
- Dawkins, E., André, K., Axelsson, K., Benoist, L., Swartling, Å. G., and Persson, Å. (2019). Advancing sustainable consumption at the local government level: a literature review. *J. Clean. Prod.* 231, 1450–1462. doi: 10.1016/j.jclepro.2019.05.176
- Dilling, L., and Lemos, M. C. (2011). Creating usable science: opportunities and constraints for climate knowledge use and their implications for science policy. *Glob. Environ. Change* 21, 680–689. doi: 10.1016/j.gloenvcha.2010.11.006
- Dozois, E., Langlois, M., and Blanchet-Cohen, N. (2010). University of Victoria (B.C.), and International Institute for Child Rights and Development. *DE 201: A Practitioner’s Guide to Developmental Evaluation.* Victoria, BC; Montréal, QC: International Institute for Child Rights and Development, The J.W. McConnell Family Foundation.
- Englund, M. (2021). *Three principles for evaluating co-produced climate services* (Master’s Thesis, Lund University). <https://lup.lub.lu.se/student-papers/search/publication/9062503> (accessed June 27, 2022).
- Ernst, A. (2019). Research techniques and methodologies to assess social learning in participatory environmental governance. *Learn. Cult. Soc. Interact.* 23, 1–17. doi: 10.1016/j.lcsi.2019.100331
- Ernst, K. M., Swartling, Å. G., André, K., Preston, B. L., and Klein, R. J. T. (2019). Identifying climate service production constraints to adaptation decision-making in Sweden. *Environ. Sci. Policy* 93, 83–91. doi: 10.1016/j.envsci.2018.11.023
- European Commission. (2015). *A European Research and Innovation Roadmap for Climate Services.* Publications Office of the European Union. Available online at: <http://op.europa.eu/en/publication-detail/-/publication/73d73b26-4a3c-4c55-bd50-54fd22752a39> (accessed June 27, 2022).
- Fazey, I., Bunse, L., Msika, J., Pinke, M., Preedy, K., Evely, A. C., et al. (2014). Evaluating knowledge exchange in interdisciplinary and multi-stakeholder research. *Glob. Environ. Change* 25, 204–220. doi: 10.1016/j.gloenvcha.2013.12.012
- Felizardo, K. R., Mendes, E., Kalinowski, M., Souza, É. F., and Vijaykumar, N. L. (2016). “Using Forward Snowballing to update Systematic Reviews in Software Engineering,” *Proceedings of the 10th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement*, 1–6.
- Fulgenzi, A., Brouwer, S., Baker, K., and Frijns, J. (2020). Communities of practice at the center of circular water solutions. *WIREs Water* 7, 1–15. doi: 10.1002/wat2.1450
- Gerger Swartling, Å., Tenggren, S., André, K., and Olsson, O. (2019). Joint knowledge production for improved climate services: Insights from the Swedish forestry sector. *Environ. Policy Gov.* 29, 97–106. doi: 10.1002/eet.1833
- Gerlak, A. K., Guido, Z., Vaughan, C., Rountree, V., Greene, C., Liverman, D., et al. (2018). Building a framework for process-oriented evaluation of regional climate outlook forums. *Weather Clim. Soc.* 10, 225–239. doi: 10.1175/WCAS-D-17-0029.1
- Guijt, I. (2014). *Participatory Approaches* (Methodological Briefs: Impact Evaluation 5). UNICEF Office of Research. Available online at: [https://www.betterevaluation.org/en/resources/guide/participatory\\_approaches](https://www.betterevaluation.org/en/resources/guide/participatory_approaches) (accessed June 27, 2022).
- Haddaway, N. R., Woodcock, P., Macura, B., and Collins, A. (2015). Making literature reviews more reliable through application of lessons from systematic reviews. *Conserv. Biol.* 29, 1596–1605. doi: 10.1111/cobi.12541
- Hansson, S., and Polk, M. (2018). Assessing the impact of transdisciplinary research: The usefulness of relevance, credibility, and legitimacy for understanding the link between process and impact. *Res. Eval.* 27, 132–144. doi: 10.1093/reseval/rvy004
- Harvey, B., Cochran, L., and Epp, M. V. (2019). Charting knowledge co-production pathways in climate and development. *Environ. Policy Gov.* 29, 107–117. doi: 10.1002/eet.1834
- Hassel, H. (2010). *Risk and Vulnerability Analysis in Society’s Proactive Emergency Management.* Lund University.
- Hassenforder, E., Pittock, J., Barreateau, O., Daniell, K. A., and Ferrand, N. (2016). The MEPPP Framework: A Framework for Monitoring and Evaluating Participatory Planning Processes. *Environ. Manag.* 57, 79–96. doi: 10.1007/s00267-015-0599-5
- Hassenforder, E., Smajgl, A., and Ward, J. (2015). Towards understanding participatory processes: Framework, application and results. *J. Environ. Manag.* 157, 84–95. doi: 10.1016/j.jenvman.2015.04.012
- Hegger, D., Lamers, M., Van Zeijl-Rozema, A., and Dieperink, C. (2012). Conceptualising joint knowledge production in regional climate change adaptation projects: success conditions and levers for action. *Environ. Sci. Policy* 18, 52–65. doi: 10.1016/j.envsci.2012.01.002
- Heink, U., Marquard, E., Heubach, K., Jax, K., Kugel, C., Neßhöver, C., et al. (2015). Conceptualizing credibility, relevance and legitimacy for evaluating the effectiveness of science–policy interfaces: challenges and opportunities. *Sci. Public Policy* 42, 676–689. doi: 10.1093/scipol/scu082
- Hewitt, C. D., and Stone, R. (2021). Climate services for managing societal risks and opportunities. *Clim. Serv.* 23, 1–7. doi: 10.1016/j.cliser.2021.100240
- Hitziger, M., Aragrande, M., Berezowski, J., Canali, M., Del Rio Vilas, V., Hoffmann, S., et al. (2019). EVOLvINC: EVALuating knOwLedge INtegration Capacity in multistakeholder governance. *Ecol. Soc.* 24, 1–16. doi: 10.5751/ES-10935-240236
- Holzer, J. M., Carmon, N., and Orenstein, D. E. (2018). A methodology for evaluating transdisciplinary research on coupled socio-ecological systems. *Ecol. Indic.* 85, 808–819. doi: 10.1016/j.ecolind.2017.10.074
- IPCC (2018). “Annex 1: Glossary,” in *Global Warming of 1.5°C. An IPCC Special Report on the Impacts of Global Warming of 1.5°C Above Pre-industrial Levels and Related Global Greenhouse Gas Emission Pathways, in the Context of Strengthening the Global Response to the Threat of Climate Change, Sustainable Development, and Efforts to Eradicate Poverty.* Available online at: <https://www.ipcc.ch/sr15/chapter/glossary/> (accessed June 27, 2022).
- Jahn, T., and Keil, F. (2015). An actor-specific guideline for quality assurance in transdisciplinary research. *Futures* 65, 195–208. doi: 10.1016/j.futures.2014.10.015
- Janoff, S. (2004). *States of Knowledge: The Co-Production of Science and the Social Order.* London: Routledge.
- Jones, L., Harvey, B., Cochran, L., Cantin, B., Conway, D., Cornforth, R. J., et al. (2018). Designing the next generation of climate adaptation research for development. *Reg. Environ. Change* 18, 297–304. doi: 10.1007/s10113-017-1254-x
- Jones, N. A., Perez, P., Measham, T. G., Kelly, G. J., d’Aquino, P., Daniell, K. A., et al. (2009). Evaluating participatory modeling: developing a framework for cross-case analysis. *Environ. Manag.* 44, 1180–1195. doi: 10.1007/s00267-009-9391-8
- Klein, R. J. T., and Juhola, S. (2014). A framework for Nordic actor-oriented climate adaptation research. *Environ. Sci. Policy* 40, 101–115. doi: 10.1016/j.envsci.2014.01.011
- Knickel, M., Knickel, K., Galli, F., Maye, D., and Wiskerke, J. S. C. (2019). Towards a reflexive framework for fostering co-learning and improvement of transdisciplinary collaboration. *Sustainability* 11, 1–22. doi: 10.3390/su11236602

- Lang, D. J., Wiek, A., Bergmann, M., Stauffacher, M., Martens, P., Moll, P., et al. (2012). Transdisciplinary research in sustainability science: Practice, principles, and challenges. *Sustain. Sci.* 7, 25. edssjs. doi: 10.1007/s11625-011-0149-x
- Laycock, A., Bailie, J., Matthews, V., and Bailie, R. (2019). Using developmental evaluation to support knowledge translation: Reflections from a large-scale quality improvement project in Indigenous primary healthcare. *Health Res. Policy Syst.* 17, 1–11. doi: 10.1186/s12961-019-0474-6
- Lemos, M. C., Arnott, J. C., Ardoin, N. M., Baja, K., Bednarek, A. T., Dewulf, A., et al. (2018). To co-produce or not to co-produce. *Nat. Sustain.* 1, 722–724. doi: 10.1038/s41893-018-0191-0
- Lourenço, T. C., Swart, R., Goosen, H., and Street, R. (2016). The rise of demand-driven climate services. *Nat. Clim. Change* 6, 13–14. doi: 10.1038/nclimate2836
- Mayne, J. (2017). Theory of change analysis: building robust theories of change. *Can. J. Prog. Eval.* 32, 155–173. doi: 10.3138/cjpe.31122
- McNie, E. C. (2007). Reconciling the supply of scientific information with user demands: an analysis of the problem and review of the literature. *Environ. Sci. Policy* 10, 17–38. doi: 10.1016/j.envsci.2006.10.004
- Mitchell, A., and Lemon, M. (2020). Learning how to learn in sustainability transitions projects: the potential contribution of developmental evaluation. *J. Multidiscip. Eval.* 16, 91–103.
- Nel, J. L., Roux, D. J., Driver, A., Hill, L., Maherry, A. C., Snaddon, K., et al. (2016). Knowledge co-production and boundary work to promote implementation of conservation plans. *Conserv. Biol.* 30, 176–188. doi: 10.1111/cobi.12560
- Norström, A. V., Cvitanovic, C., Löf, M. F., West, S., Wyborn, C., Balvanera, P., et al. (2020). Principles for knowledge co-production in sustainability research. *Nat. Sustain.* 3, 182–190. doi: 10.1038/s41893-019-0448-2
- O'Connor, R. A., Nel, J. L., Roux, D. J., Lim-Camacho, L., van Kerkhoff, L., and Leach, J. (2019). Principles for evaluating knowledge co-production in natural resource management: Incorporating decision-maker values. *J. Environ. Manag.* 249, 109392. doi: 10.1016/j.jenvman.2019.109392
- Palutikof, J. P., Street, R. B., and Gardiner, E. P. (2019). Decision support platforms for climate change adaptation: an overview and introduction. *Clim. Change* 153, 459–476. doi: 10.1007/s10584-019-02445-2
- Papathoma-Köhle, M., Promper, C., and Glade, T. (2016). A common methodology for risk assessment and mapping of climate change related hazards—implications for climate change adaptation policies. *Climate* 4, 1–23. doi: 10.3390/cli4010008
- Patton, M. Q. (2006). Evaluation for the way we work. *Nonprofit Q.* 13, 28–33.
- Patton, M. Q. (2010). *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. New York, NY: Guilford Press.
- Patton, M. Q., and Horton, D. (2009). *Utilization-focused evaluation for agricultural innovation* (No. 22; ILAC Brief). The Institutional Learning and Change Initiative. Available online at: [https://www.betterevaluation.org/sites/default/files/ILAC\\_Brief22\\_Utilization\\_Focus\\_Evaluation.pdf](https://www.betterevaluation.org/sites/default/files/ILAC_Brief22_Utilization_Focus_Evaluation.pdf) (accessed June 27, 2022).
- Petheram, L., Stacey, N., Campbell, B. M., and High, C. (2012). Using visual products derived from community research to inform natural resource management policy. *Land Use Policy* 29, 1–10. doi: 10.1016/j.landusepol.2011.04.002
- Plottu, B., and Plottu, E. (2011). Participatory evaluation: the virtues for public governance, the constraints on implementation. *Group Dec. Negot.* 20, 805–824. doi: 10.1007/s10726-010-9212-8
- Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., et al. (2006). *Guidance on the Conduct of Narrative Synthesis in Systematic Reviews*. ESRC Methods Programme. Available online at: <https://www.lancaster.ac.uk/media/lancaster-university/content-assets/documents/fhm/dhr/chir/NSynthesisguidanceVersion1-April2006.pdf> (accessed June 27, 2022).
- Porter, J. J., and Dessai, S. (2017). Mini-me: Why do climate scientists' misunderstand users and their needs? *Environ. Sci. Policy* 77, 9–14. doi: 10.1016/j.envsci.2017.07.004
- Reed, M. S., Ferré, M., Martin-Ortega, J., Blanche, R., Lawford-Rolfe, R., Dallimer, M., et al. (2021). Evaluating impact from research: a methodological framework. *Res. Policy* 50, 1–17. doi: 10.1016/j.respol.2020.104147
- Reynolds, M. (2014). Equity-focused developmental evaluation using critical systems thinking. *Evaluation* 20, 75–95. doi: 10.1177/1356389013516054
- Romina, R., and Gerger Swartling, Å. (2019). Special issue: environmental governance in an increasingly complex world: reflections on transdisciplinary collaborations for knowledge co-production and learning. *Environ. Policy Gov.* 29, 81–82. doi: 10.1002/eet.1842
- Roux, D. J., Storzaker, R. J., Breen, C. M., Lefroy, E. C., and Cresswell, H. P. (2010). Framework for participative reflection on the accomplishment of transdisciplinary research programs. *Environ. Sci. Policy* 13, 733–741. doi: 10.1016/j.envsci.2010.08.002
- Salamanca, A., and Biskupska, N. (2021). *Monitoring, Evaluation and Learning to Build Better Climate Services: A Framework for Inclusion, Accountability and Iterative Improvement in Tandem*. SEI Discussion Brief. Stockholm Environment Institute. [https://www.sei.org/publications/monitoring-evaluation\\_climate-services-in-tandem/](https://www.sei.org/publications/monitoring-evaluation_climate-services-in-tandem/) (accessed June 27, 2022).
- Sarkki, S., Tinch, R., Niemelä, J., Heink, U., Waylen, K., Timaeus, J., et al. (2015). Adding 'iterativity' to the credibility, relevance, legitimacy: a novel scheme to highlight dynamic aspects of science-policy interfaces. *Environ. Sci. Policy* 54, 505–512. doi: 10.1016/j.envsci.2015.02.016
- Schuck, S., Cortekar, J., and Jacob, D. (2017). Evaluating co-creation of knowledge: from quality criteria and indicators to methods. *Adv. Sci. Res.* 14, 305–312. doi: 10.5194/asr-14-305-2017
- Shea, J., and Taylor, T. (2017). Using developmental evaluation as a system of organizational learning: an example from San Francisco. *Eval. Prog. Plan.* 65, 84–93. doi: 10.1016/j.evalprogplan.2017.07.001
- Steynor, A., Lee, J., and Davison, A. (2020). Transdisciplinary co-production of climate services: a focus on process. *Soc. Dyn.* 46, 414–433. doi: 10.1080/02533952.2020.1853961
- Steynor, A., Padgham, J., Jack, C., Hewitson, B., and Lennard, C. (2016). Co-exploratory climate risk workshops: Experiences from urban Africa. *Clim. Risk Manag.* 13, 95–102. doi: 10.1016/j.crm.2016.03.001
- Swart, R. J., de Bruin, K., Dhenain, S., Dubois, G., Groot, A., and von der Forst, E. (2017). Developing climate information portals with users: Promises and pitfalls. *Clim. Serv.* 6, 12–22. doi: 10.1016/j.cliser.2017.06.008
- Tall, A., Coulibaly, J. Y., and Diop, M. (2018). Do climate services make a difference? A review of evaluation methodologies and practices to assess the value of climate information services for farmers: Implications for Africa. *Clim. Serv.* 11, 1–12. doi: 10.1016/j.cliser.2018.06.001
- Thomas, D. R. (2006). A General Inductive Approach for Analyzing Qualitative Evaluation Data. *Am. J. Eval.* 27, 237–246. doi: 10.1177/1098214005283748
- Trimble, M., and Plummer, R. (2019). Participatory evaluation for adaptive co-management of social-ecological systems: a transdisciplinary research approach. *Sustain. Sci.* 14, 1091–1103. doi: 10.1007/s11625-018-0602-1
- van der Wal, M., De Kraker, J., Offermans, A., Kroeze, C., Kirschner, P. A., and van Ittersum, M. (2014). Measuring social learning in participatory approaches to natural resource management: measuring social learning in participatory approaches. *Environ. Policy Gov.* 24, 1–15. doi: 10.1002/eet.1627
- van Es, M., Guijt, I., and Vogel, I. (2015). *Hivos ToC Guidelines: Theory of Change Thinking in Practice. A Stepwise Approach*. Hivos. Available online at: <https://hivos.org/document/hivos-theory-of-change/> (accessed June 27, 2022).
- van Tulder, R., and Keen, N. (2018). Capturing collaborative challenges: designing complexity-sensitive theories of change for cross-sector partnerships. *J. Bus. Ethics* 150, 315–332. doi: 10.1007/s10551-018-3857-7
- VanderMolen, K., Wall, T. U., and Daudert, B. (2019). A call for the evaluation of web-based climate data and analysis tools. *Bull. Am. Meteorol. Soc.* 100, 257–268. doi: 10.1175/BAMS-D-18-0006.1
- Vaughan, C., and Dessai, S. (2014). Climate services for society: Origins, institutional arrangements, and design elements for an evaluation framework. *Wiley Interdiscip. Rev. Clim. Change* 5, 587–603. doi: 10.1002/wcc.290
- Vaughan, C., Dessai, S., and Hewitt, C. (2018). Surveying Climate Services: What Can We Learn from a Bird's-Eye View? *Weather Clim. Soc.* 10, 373–395. doi: 10.1175/WCAS-D-17-0030.1
- Vincent, K., Carter, S., Steynor, A., Visman, E., and Lund Wägsæther, K. (2020). Addressing power imbalances in co-production. *Nat. Clim. Change* 10, 877–878. doi: 10.1038/s41558-020-00910-w
- Vincent, K., Daly, M., Scannell, C., and Leathes, B. (2018). What can climate services learn from theory and practice of co-production? *Clim. Serv.* 12, 48–58. doi: 10.1016/j.cliser.2018.11.001

- Visman, E., Vincent, K., Steynor, A., Karani, I., and Mwangi, E. (2022). Defining metrics for monitoring and evaluating the impact of co-production in climate services. *Clim. Serv.* 26, 100297. doi: 10.1016/j.cliser.2022.100297
- Vogel, C., Steynor, A., and Manyuchi, A. (2019). Climate services in Africa: re-imagining an inclusive, robust and sustainable service. *Clim. Serv.* 15, 100107. doi: 10.1016/j.cliser.2019.100107
- Vogel, I. (2012). *Review of the use of 'Theory of Change' in international development* [Review report]. UK Department for International Development (DFID). Available online at: [https://www.theoryofchange.org/pdf/DFID\\_ToC\\_Review\\_VogelV7.pdf](https://www.theoryofchange.org/pdf/DFID_ToC_Review_VogelV7.pdf) (accessed June 27, 2022).
- Vogel, J., Letson, D., and Herrick, C. (2017). A framework for climate services evaluation and its application to the Caribbean Agrometeorological Initiative. *Clim. Serv.* 6, 65–76. doi: 10.1016/j.cliser.2017.07.003
- Wall, T. U., Meadow, A. M., and Horganic, A. (2017). Developing evaluation indicators to improve the process of coproducing usable climate science. *Weather Clim. Soc.* 9, 95–107. doi: 10.1175/WCAS-D-16-0008.1
- Walter, A. I., Helgenberger, S., Wiek, A., and Scholz, R. W. (2007). Measuring societal effects of transdisciplinary research projects: Design and application of an evaluation method. *Eval. Prog. Plan.* 30, 325–338. doi: 10.1016/j.evalprogplan.2007.08.002
- Wiek, A., Talwar, S., O'Shea, M., and Robinson, J. (2014). Toward a methodological scheme for capturing societal effects of participatory sustainability research. *Res. Eval.* 23, 117–132. doi: 10.1093/reseval/rvt031
- WMO. (2009). *World Climate Conference-3 (WCC-3) | GFCS*. Available online at: [http://www.gfcs-climate.org/wcc\\_3/](http://www.gfcs-climate.org/wcc_3/) (accessed June 27, 2022).
- Wohlin, C. (2014). "Guidelines for snowballing in systematic literature studies and a replication in software engineering," in *8th International Conference on Evaluation and Assessment in Software Engineering (EASE'14)*, 321–330. Available online at: <https://www.wohlin.eu/ease14.pdf> (accessed June 27, 2022).
- Wyborn, C. (2015). Connectivity conservation: Boundary objects, science narratives and the co-production of science and practice. *Environ. Sci. Policy* 51, 292–303. doi: 10.1016/j.envsci.2015.04.019
- Zscheischler, J., Rogga, S., and Lange, A. (2018). The success of transdisciplinary research for sustainable land use: Individual perceptions and assessments. *Sustain. Sci.* 13, 1061–1074. doi: 10.1007/s11625-018-0556-3

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

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Englund, André, Gerger Swartling and Iao-Jørgensen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.