Scientist-Stakeholder Relationships Drive Carbon Data Product Transfer Effectiveness within NASA Program

Molly E Brown[1]* mbrown52@umd.edu
ORC-ID: 0000-0001-7384-3314
Vanessa M. Escobar[2], vanessa.escobar@noaa.gov
ORC-ID: 0000-0002-6837-8099
Fatima M. Younis[3], fyounis@umd.edu
ORC-ID: 0000-0002-7228-1724
Edil Sepúlveda Carlo4, edil.sepulvedacarlo@nasa.gov
ORC-ID: 0000-0002-7650-7970
Megan McGroddy[4], megan.e.mcgroddy@nasa.gov
ORC-ID: 0000-0002-8481-1550
Sabrina Delgado Arias[4], sabrina.delgadoarias@nasa.gov
ORC-ID: 0000-0003-4411-1717
Peter Griffith[4], peter.c.griffith@nasa.gov
ORC-ID: 0000-0002-4267-7429
George Hurtt[1], gchurtt@umd.edu,
ORC-ID: 0000-0001-7278-202X

1. Department of Geographical Sciences, University of Maryland, College Park MD 20742
*Corresponding Author, mbrown52@umd.edu
Phone: 703-855-6190

2. NOAA/NASA GEO-XO, Radiometric Calibration Laboratory, NASA Goddard Space Flight Center

3. Department of Government and Politics, University of Maryland, College Park, MD 20742

4. Science Systems and Applications, NASA Goddard Space Flight Center, Greenbelt, MD

Abstract

Carbon cycle science is at the heart of research on global climate change and its long-term impacts, as it examines the exchange of carbon between the atmosphere, oceans, land, and the impact of fossil fuel emissions on this cycling. Given the urgency of the climate challenge, NASA's Carbon Monitoring System (CMS) requires all funded investigators to identify and work with stakeholder organizations at project inception to accelerate the transfer of the products developed by funded research into decision making systems. In this study, we contribute to the literature through the implementation of a quantitative analysis of 908 unique survey responses from funded investigators to explore the maturity of the scientist-stakeholder engagement. The paper employs multiple correspondence analysis to provide evidence to support policy options to increase stakeholder integration into research programs. Despite limitations of the dataset used, we demonstrated that multiple funding rounds, long-standing relationships between the stakeholder and scientist, and the scientific productivity of the PI, including the ability to produce datasets and research papers on these datasets, all contribute to carbon products moving from research to operational use. The maturity of relationships between scientists and stakeholders was shown to result improved stakeholder engagement. The use of carbon products should be identified in every stage of the program, and that capacity building is needed to support both existing and newly identified stakeholders better understand and use CMS products. As Federal, State, and local policy on climate adaptation and mitigation matures, the need for information on carbon will expand. Building of stakeholder-scientist relationships in CMS results in an effective generation and use of datasets to support this need and prototype ways that improved information needed for decision making can be created.

Keywords: Monitoring, Evaluation, carbon monitoring system, stakeholder, engagement

1. Introduction

Carbon cycle science examines the exchange of carbon between the atmosphere, oceans, land and the impact of fossil fuel emissions on this cycling. Studying the carbon cycle helps us understand the probable impact of climate change on humanity through rising temperatures and increasing carbon dioxide levels (Allen and others 2018). Climate change also threatens long-term economic development (Liobikiene and Butkus, 2018), food production (Ray *et al* 2013), and will damage urban infrastructure (Wilbanks and Fernandez, 2014) necessary to support a growing human population. However, ensuring that critical information on climate is actually used in day-to-day decisions and policy making by stakeholders such as governments, businesses, and institutions requires engagement and communication between the user and the producer of the information (Cash *et al* 2006).

We define stakeholders as an individuals, groups or organizations that are affected by climate change, who can make policy, investment or activity decisions with carbon data, and who are end-users of the data CMS produces. Carbon data is information, analysis, visualizations and data products that inform decision makers about carbon stocks and fluxes that move throughout the Earth system across a range spatial and temporal scales. A carbon stock, or carbon pool, is a system that has the capacity to store or release carbon. A carbon flux refers to the amount of carbon exchanged between carbon stocks over a specified time. In simple terms, CMS data seeks to model and measure the movement of carbon between land, oceans, atmosphere, and living things (Hurtt *et al* 2022). Although we recognize that scientists are often also stakeholders of scientific information and models, we focus on non-scientist stakeholders in this context for clarity.

NASA’s Carbon Monitoring System (CMS) has worked for the past ten years to prototype capabilities necessary to support stakeholder needs for Monitoring, Reporting, and Verification of carbon stocks and fluxes (Hurtt *et al* 2019). The result of this sustained funding is the development of a community of practice where scientists have learned from each other about how to do meaningful stakeholder engagement, the value of this engagement, and have learned through annual Science Team meetings and stakeholder workshops about applications of CMS products (Brown *et al* 2020). By connecting cutting edge carbon cycle science research to stakeholders beyond the scientific community who may use the data in their decision making, NASA CMS contributes to understanding the needs of the climate data end-user community (Moser and Ekstrom 2010). For the past eight years, the NASA Goddard Space Flight Center (GSFC) science applications team has engaged both the CMS PIs and a diverse set of stakeholders to encourage mutual understanding of data needs and functionality of the current and planned CMS data products for effective use in decision making contexts. The goal of the CMS applications efforts is to link stakeholders to CMS science products and provide a path for feedback and lessons learned for CMS PIs so CMS is more accessible and user friendly. Stakeholders closely engaged with CMS projects at the federal level include the U.S. Environmental Protection Agency (EPA), the USDA Forest Service, and NOAA (Figure S1) (Carlo *et al* 2018).

Challenging the science community to identify, learn from and engage directly with potential users of their science has resulted in improved relevance and uptake of scientific products (Brugger *et al* 2016). The CMS program motivates new basic research while integrating the user community into data product creation and distribution, demonstrating how science and technology can be integrated into decision making (West *et al* 2018).

Here we use a quantitative approach to assess CMS scientist engagement with stakeholders and promote use of carbon cycle science data developed during the project. Our primary hypothesis is that by measuring specific characteristics of the CMS Principal Investigator (PI) scientist and their institution, such as their experience, personal relationships with the stakeholder, frequency of communication with the stakeholder, and the period of support for the project, we can estimate the effectiveness of CMS PIs in creating useful carbon products and transferring them to support decision making. We focus on the CMS PI, the data they produce, and their engagement with stakeholders, not the stakeholders themselves.

We also hypothesize that traditional scientific measures of 'success', such as citation of peer reviewed articles, can be related quantitatively to changes in stakeholder engagement, as measured by our impact metric Difference of Applications Readiness Levels (DiffARL). Our hypothesis regarding co-production is that when an investigator receives multiple rounds of CMS funding, this enhances the likelihood that the CMS PI will build a mature, long-standing relationship with a stakeholder (Brugger *et al* 2016, Jahn *et al* 2012). To test these hypotheses, we develop categorical variables describing each funded CMS product, and use multiple correspondence analysis to explore the efficacy of the stakeholder-scientist interaction.

2.0 Literature Review

Previous research has shown that a collaborative approach to knowledge development is more likely to result in usable information than when research is conducted in isolation (Wall *et al* 2017, Fazey *et al* 2014). Co-production of knowledge or transdisciplinary research (Jahn *et al* 2012) lies between basic research into processes, relationships and product development typically funded by NASA's Earth Science Division and applied research focused on defining applications that can be supported with Earth Science products and guide scientific priorities (Moran *et al* 2015).

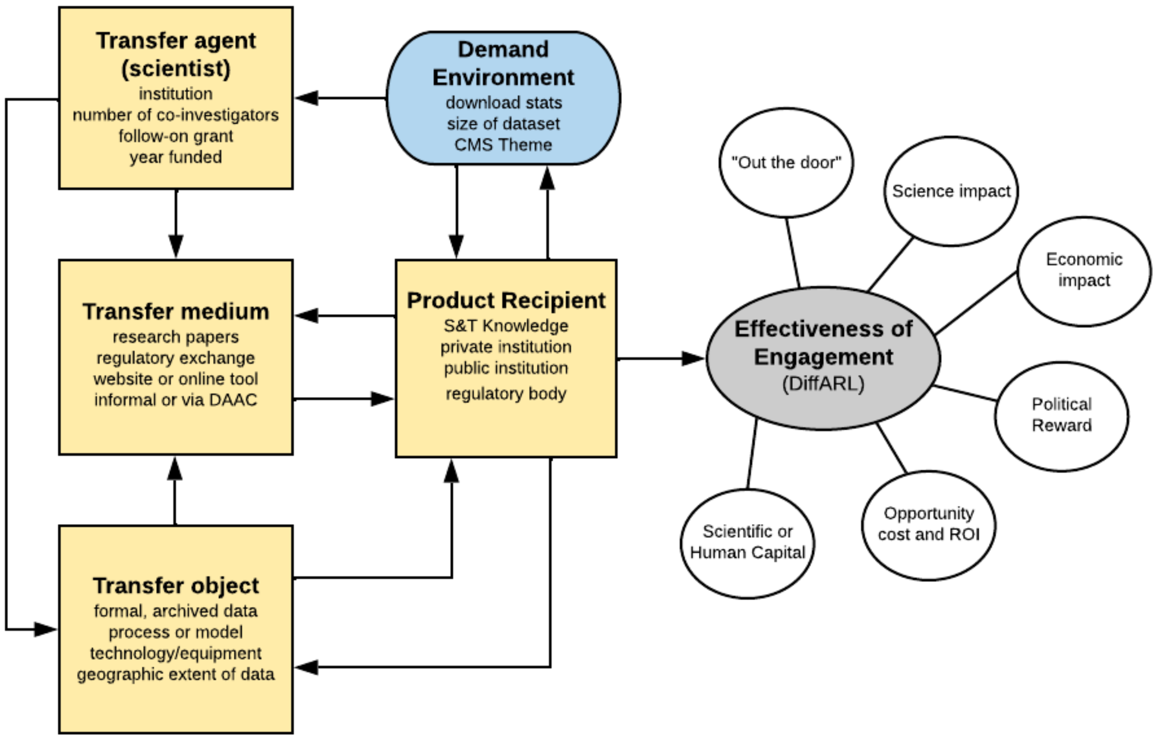
Extensive previous work has been done on understanding the link between applied and basic research and its use in policy and decision making. Sarewitz and Pielke (2007) conceptualized how the supply of information generated through investment in basic research could meet the need of society. Matching the ‘demand’ for science, particularly in support of decision-making in public affairs, to monitor and assess the impact on society that science and technology has created, to the ‘supply’ of basic or applied research requires constant and early interaction between the producers and users of the information (Sarewitz and Pielke 2007). The utility of scientific information does not fall directly out of the knowledge itself, but requires that the knowledge be ‘socially robust’, valid and reliable in the context in which it could be used, which is attained through engagement with experts and stakeholders throughout its development (Gibbons 1999, Cash *et al* 2003).

We recognize that fundamentally, building relationships between scientists, experts and stakeholders who will use carbon data is at the heart of the CMS program. Engaging with stakeholders over time in ways that allow for two way-learning, the development of long-term relationships, and transformation of methods and datasets to meet the needs of stakeholders (Cook and Zurita 2019). Issues related to climate and environment are particularly thorny to resolve because they require both scientific knowledge and political and social values. Addressing these problems requires establishing and maintaining dialogs among interested parties with differing values to bring scientific expertise together with local and environmental concerns to find solutions (Meadow *et al* 2015, Ludwig 2001). Our research therefore focuses on the CMS science community, their ability and interest in engaging in this dialog, and seeks to better understand the context in which their research could be used.

Drawing on research from the field of technology transfer, the Contingent Effectiveness Model draws its name from the assumption that parties to technology transfer have multiple goals and effectiveness criteria that depend on who the user is and how they value the dataset (Bozeman 2000). Similar to Cash et al (2006)'s four critical functions in application of science, which include convening, translation, collaboration, and mediation, the Bozeman model provides five broad dimensions that determine effectiveness: 1) characteristics of the scientist or transfer agent who is guiding the research and product development; 2) characteristics of the method through which the technology is transferred (transfer medium); 3) characteristics of the product itself such as resolution, time step and latency (transfer object); 4) the demand environment or the need for the data in the user community; and finally 5) the characteristics of the product recipient or stakeholder (Figure 1). An assumption of the Contingent Effectiveness Model is that there is no single way to measure effectiveness of technology transfer since effectiveness is defined by each stakeholder individually and in the context in which the data are being used (right side of Figure 1). This results in highly contextualized and fundamentally incomparable 'success' criteria, which although relevant, is also difficult for a funding agency like NASA to use in evaluating the success of its program to communicate, disseminate, and encourage use of its products. This is the primary reason why we use the scientist-provided product Application Readiness Level (ARL) change metric as a way to determine 'success' of CMS's impact on society through describing the products' progression of use of a product within a stakeholders' decision-making activities (NASA 2017).

Connections between research and societal outcomes are affected by a wide array of contingency, complexity and non-linearity factors, but these factors need not prevent the use of data for improved decision-making (Changnon *et al* 2000, Lasswell 1971). Here we use data on product application readiness levels reported by the project scientist to determine how characteristics of the transfer agent, transfer medium, the product, the stakeholder and the demand environment affect the uptake of the product by the stakeholder (Bozeman 2000) (Table 1). Contributing to the literature about co-production of knowledge (Wall *et al* 2017), we present a quantitative analysis that focuses on determining the potential causes for why some CMS funded programs were able to engage effectively, as described in Table 1, and others were not.

Figure 1. CMS Contingent Effectiveness Model of technology transfer, amended from Bozeman (2000)



Using this framework as a guiding principle, we will examine the likely impact on our ‘Effectiveness of Engagement’ metric from the scientist perspective, as measured by the difference in ARL level from the start to the end of the development of a product (referred to here as the DiffARL variable). Those products with that report change in ARL level are the result of engagement with the stakeholder or target organization. This change will deliver a variety of CMS ‘success criteria’ (Table 1), such as ‘out the door’, increased scientific and human capital through the engagement between the scientist organization and the stakeholder organization, and science impact through the process of writing and publishing papers on the new data product.

3.0 NASA’s Carbon Monitoring System (CMS)

NASA’s CMS initiative was initially funded in the 2010 Congressional Appropriation, which directed NASA to start working towards a Carbon Monitoring System (CMS) and provided specific guidance on how this could be done. NASA CMS emphasizes exploitation of the satellite remote sensing resources, scientific knowledge, and modeling expertise that are major strengths of the NASA Earth Science program (Hurtt *et al* 2014). The approach focuses on product development and requires close communications and/or partnerships with state, local, tribal and federal government agencies and their technical experts who develop and produce carbon inventory and biomass inventories. Here we assess CMS scientists' perceptions of their stakeholder engagement to provide relevant programmatic lessons learned for NASA Earth Science Division (ESD). Improvements in the use of Earth Science data can have a societal benefit by supporting decision making by stakeholders in their efforts to mitigate or adapt to a changing climate. Improving decision support and use of NASA data products is a key goal of the NASA ESD and of CMS.

CMS requires that all funded PIs have users and stakeholder organizations included when proposing, conducting their research, and documenting their results. The focus of CMS is to iteratively develop data products in collaboration with stakeholder organizations so that the data products better inform monitoring, reporting, and verification of carbon fluxes and stocks across a variety of institutions and decisions. Inclusion of users and stakeholder organizations is now a requirement for NASA missions.

4.0 Measuring CMS Impact

To test our hypotheses, we use information on each product and its use in a stakeholder environment generated through scientist questionnaire. Because these questionnaires are repeated every year and the CMS program has specific and independent stakeholder engagement activities, there are multiple evaluation points for the data to ensure its consistency and quality. Below we set out metrics we use to describe the drivers of the DiffARL metric.

The CMS Principal Investigator (PI) is our transfer agent in this context. Recent research has shown that there is a great deal of learning (Ernst 2019) that occurs within the science team and stakeholder engagement meetings supported by CMS, engendering a community of practice (Wenger 2011). Here we use information on the PI institution, the number of co-investigators they have on the grant, whether the grant is one of a series that was funded by CMS and the year the

Table 1. Technology transfer effectiveness criteria from the Contingent Effectiveness Model, derived from Bozeman (2000)

Effectiveness criterion	Focus	Relation to research and practice	CMS Success Criteria
"Out the door"	Success requires at least one organization to learn about data product, without regard to impact.	Extremely common practice to simply determine if someone used the data product with no regard to impact on decision making.	DiffARL is low or zero, since the project stays in 'Discovery and Feasibility stage', but papers are written to increase product awareness.
Science impact	Citation score documenting that the research or dataset description been used in a scientific literature, demonstrating 'science impact'.	Widespread success criteria for a funded research program, with many research studies and methods available to measure science impact.	DiffARL is low, since the project stays in 'Discovery and Feasibility stage'.
Economic impact	Has the transfer resulted in economic benefit for institution, community or industry through its use?	Important criteria for perception of impact but can be difficult to measure without access to private data and is beyond the scope of nearly all research programs.	*Not measured with DiffARL - not evaluated here*
Political Reward	Based on the expectation of reward or impact flowing from the use of the data product, such as increased importance of fighting forest fires or regulation to reduce pollution sources identified.	Widely used as a success criteria in practice, but poorly quantified or examined in the literature.	DiffARL is 6 or greater, as product moves from 'discovery' to 'application' stages
Opportunity Cost and ROI	Return on investment on the part of the scientist and the stakeholder organization, particularly against other ways of using resources and time.	Concern among practitioners, rarely studied in the literature except in cost-benefit analyses.	*Not measured with DiffARL - not evaluated here*
Scientific or Human Capital	Impact of engagement with stakeholder on the enhanced scientific, technological and	A high priority for CMS in its mission, but poorly measured	DiffARL is 6 or greater, as product moves from

	communication skills, particularly focused on social capital and on students and other team members in both user and producer organizations.	and rarely studied in the literature.	'discovery' to 'application' stages

project was funded as potential drivers of the maturity of the PI-stakeholder relationship (Table 2). Although economic impact is an important part of assessing the value of carbon products, the data that we had available for this review did not include information on potential economic benefits of the data. Mature engagement with a stakeholder, including generating a deep understanding of the organizational context in which the product is used (VanderMolen *et al* 2020), may result in a product moving from a Stage 1 ARL (research) to a Stage 3 ARL (stakeholder use of the product in decision making) (Wall *et al* 2017).

The transfer medium describes the way the CMS carbon data is transferred to the user. Research papers and other publications are the primary way most scientists communicate their findings about the carbon cycle to others, including stakeholders, regulatory bodies and scientists. Generating knowledge in a systematic way and publishing it is widely accepted as a primary output from NASA research funding and can be instrumental in communicating results to a broad community. Other ways CMS PIs transfer their products include direct transfer from PI to stakeholder; presentations in CMS meetings; and the engagement work of CMS applications efforts. Here we use quantitative data on total number of citations on datasets in the Distributed Active Archive Centers (DAAC) as a measure of the transfer medium.

Similar to the transfer medium, the transfer object is the carbon dataset produced by the project, its size and geographic extent. The focus is on the content and form of the dataset, and its characteristics such as spatial and temporal data extent that determines whether the stakeholder can use it or not (Table 2). For example, if the stakeholder is a local user in Reno, Nevada, who is making decisions on investments in urban tree canopy, having a carbon data product on forest biomass in Maryland will not improve the user in Nevada's ability to make decisions. Similarly, if the dataset ends in the year 2000 but the decision maker needs near-real time information, the stakeholder will not be able to use the data. We also use a total data size metric as a single metric to indicate how many files and resolutions are available for use by stakeholders.

The product recipient is the stakeholder or end-user organization. Here we use information provided by the CMS PI on the recipients of their datasets. We have a PI-determined assessment of the strength of the relationship with the stakeholder for each product, as a determinant of 'effectiveness of engagement' outcome variables. If the PI considers the relationship to be strong, then theoretically the ARL change has been large if the PI has a positive interaction with the product recipient. We will test this idea here.

We have variables that describe the demand for carbon datasets across all the funded research projects and stakeholders. We use here the size of the dataset, the theme in which the project is working (biomass, oceans or atmospheric flux) and the download statistics for datasets archived at the NASA DAACs. Although the demand is very challenging to determine, if a dataset is downloaded more or the paper cited more, then the scientist either has done a good job publicizing it or is working in an area with a real need or both.

5.0 Data and Methods

Table 2 summarizes the dataset used in the analysis. The data was derived from three different CMS PI surveys from 2016 to 2020 (Table S2). Each PI has their own project profile section on the CMS website https://carbon.nasa.gov, where the submitted abstract, participating scientists, project description and datasets produced by each PI's project are available for each year of the CMS solicitation. The Difference ARL variable has a total of 908 ARL observations, which are the change in ARL levels from survey responses from start to end for each data product-stakeholder pair (Table 2). The ARL levels change through time because of the maturity of the product changes, along with the use of the product by the stakeholder, over the period of the grant.

Because here we are connecting data products to ARL levels and the relationship with users, we use the information from each data-stakeholder response. In 2020, we added two questions to the survey that was not previously present (strength of the relationship and first engagement with the stakeholder), which has resulted in over 300 responses regarding the relationship between the scientist and the stakeholder for each product under consideration. Finally, we use information from the Distributed Active Archive Center (DAAC) where final, completed CMS data products are hosted for archiving and final distribution. The DAAC provided citations, data size and number of files in database for 98 data products archived.

Table 2. Data used in the analysis, along with the part of the technology transfer model that they address, the number of observations, the description of the variable and the source of the information.

Variable Name	Use in analysis	Number of observations	Variable description	Source of observation
CMS Theme	Demand Environment	1178	Biomass products = 1 Flux products = 2 Ocean products = 3	CMS Database
Year funded	Transfer agent	1178	2013-2018	CMS Database
PI Institution code	Transfer agent	1147	1 = other, 2=USDA, 3=UMD, 4=JPL, 5=GSFC	CMS Database
Number of Co-Is	Transfer agent	1159	Total number of co-investigators proposed on the project	CMS Database
Follow-on grant or number of precursor projects	Transfer agent	1156	0 to 3 precursor projects	CMS Database
Start ARL[1]	Engagement Effectiveness measure	927	ARL[1] levels 1-9	PI Questionnaire
Current ARL		910		
Target ARL		907		

DiffARL		908[3]	Difference between stated start and final or current ARL for each product-stakeholder pair	Calculated from the PI Questionnaire responses
Number Stakeholders Engaged/Identified for each product	Product recipient	716	1 point for each current or expected stakeholder for each product, values 0-7	PI Questionnaire
Stakeholder Communication Mechanism	Product recipient	716	If communication by proxy = 1, email=2, by phone=3, in-person=4	PI Questionnaire
Frequency of Engagement	Product recipient	716	Never communicated = 0 Communicated once= 1 Yearly = 2 Semi-annually = 3 Quarterly =4 Monthly = 5 Weekly = 6 Daily = 7	PI Questionnaire
First Engagement with Stakeholder	Product recipient	311	Long time ago = 4 When writing the proposal = 3 At start of CMS project = 2 Recently engaged = 1	PI Questionnaire
Strength of Relationship with stakeholder	Product recipient	311	Weak = 1 Somewhat weak = 2 Normal = 3 Somewhat strong = 4 Strong = 5	PI Questionnaire
Download statistics	Demand environment	98	Number of downloads	DAAC database
Citations of journal articles associated with dataset	Demand environment	56	Number of citations	DAAC database
Citations of assigned dataset doi [2]	Transfer medium	73	Number of citations	DAAC database
Data archived by CMS PI	Transfer object	1159	Number of archived datasets, min=0, max = 8	PI Questionnaire
Size of data product	Transfer object	97	Total size in MB Min = 0.1950, Max = 954300	DAAC database
Number of files in database	Transfer object	98	Number of files in database	DAAC database

1. Applications Readiness Levels, see Supplemental Table S1 for description
2. Digital Object Identifier
3. The number of DiffARL observations was limited by the PI response to the questionnaires

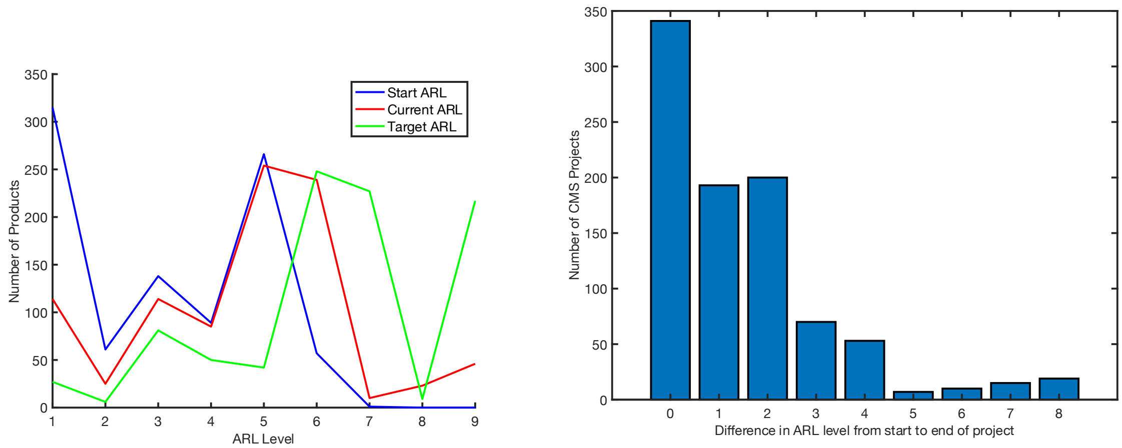
To establish a connection between the dependent variable, DiffARL, and each of the variables described in Table 1, we use a correspondence analysis (CA) technique (Greenacre & Hastie, 1987). This data analysis technique is based on singular value decomposition and is used to detect and represent underlying structures in categorical data. The primary goal of CA is to illustrate important relationships among qualitative variables using a graphical representation without assuming any particular data distribution and can accommodate any type of categorical variable whether binary, ordinal or nominal (Greenacre (1994, 1984). Here we present qualitative variables, such as the response of a CMS PI to a question regarding their relationships with stakeholders and how they interact with them, in quantitative ways. By transforming these responses into quantitative variables, we can test which aspects of the CMS program has the greatest influence on the ability of CMS PIs to increase a product's ARL level.

Each graph presented has percentage of the total variance captured by the two axes for each variable examined. The more variance captured in the second dimension, the less likely that the analysis is missing elements important for understanding how the two variables are related. We also provide the total inertia value, which is defined as the total Pearson Chi-square for the two-way variance table divided by the total sum, and therefore represents the goodness of fit of the two variables to capture all the variance present in the table. In general, the higher the inertia, the better the goodness of fit the second variable has to capture all the variance.

6.0 Results

The difference between the start and end Applications Readiness Level (DiffARL) for all 908 dataset-stakeholder pairs is shown in Figure 2A. There are relatively few products with large changes in ARL levels, with only 104 products, or 11% of the total having ARL change greater than 4 (see Table S1 for ARL definitions). No CMS products began at ARL levels at 7, 8 or 9, which denotes operational readiness. In total, 19 products have gone from conceptual ARL1 to an operational ARL9 during the project period.

Figure 2. Distribution of the DiffARL variable for 908 dataset-stakeholder pairs. **2A.** Histogram of start, current, and target project ARL levels, as reported by the PI. **2B**. Histogram of the difference between start and current ARL (DiffARL).



A. B.

Using the correspondence analysis, we present in Table 3 a summary of the ability of each independent variable to capture the variance in DiffARL. The results show that the ability of the scientist to address stakeholder demand for carbon products, including the topic that the product addresses, the interest and subsequent citation of the papers written by the project about the data, and the number of files archived in permanent storage by each product are important factors in explaining the maturity of the CMS data product for stakeholder use. Characteristics of the agent and the recipient are also important, however.

Table 3. Results of correspondence analysis, presented by the proportion of variance captured by the variable (the "inertia" parameter), along with the variance in each dimension explained from the decomposition of the categorical matrix and DiffARL. We provide the number of observations in each CA analysis.

Variable vs Diff ARL	**Use**	Inertia	% Variance in Dimension 1	% Variance in Dimension 2	Number of observations
Citation of journal articles associated with dataset	Demand	0.893	54.84%	85.63%	35
Archived data	Transfer object	0.831	33.29%	59.47%	908
Number of Files	Transfer object	0.799	38.09%	72.09%	61
Number of Stakeholders Engaged/Identified	Recipient	0.737	45.59%	70.54%	596
Citations of Dataset DOI	Medium	0.750	38.98%	66.28%	48
Year	Agent	0.542	43.25%	78.93%	908
Number of downloads	Demand	0.503	47.91%	77.59%	61
First engaged stakeholder	Recipient	0.501	53.71%	83.24%	240
Size of Data product	Object	0.489	49.87%	81.81%	60
Frequency of Engagement	Recipient	0.488	37.36%	65.61%	583

Institution code	Agent	0.487	42.53%	78.39%	896
Stakeholder Communication Mechanism	Recipient	0.445	32.17%	56.87%	595
Strength of Relationship	Recipient	0.418	58.26%	93.54%	240
Resolution of dataset (pixel size)	Demand	0.426	46.61%	70.41%	302
Spatial extent of dataset	Demand	0.363	42.69%	73.66%	365
Number of Precursor Projects	Agent	0.281	52.39%	78.57%	905
Number of Co-Is	Agent	0.264	61.65%	99.24%	908
CMS Theme	Demand	0.163	60.36%	100%	908

6.1 Transfer Agent Characteristics

The results show that characteristics of the scientist or PI developing the CMS product are important to explaining product maturity, particularly the period over which the PI was funded and the year the PI submitted the proposal. The categorical data created from the DiffARL and year information shown in Table 4 documents an increase in ARL levels each year. Figure 3 shows the correspondence analysis diagram for this same data. Projects funded in 2015 and 2016 were those that produced the largest increases in 7 and 8 ARL levels. Previous research showed that in-depth understanding of stakeholders' information needs was important to data use, but that this takes time and requires continuity in relationships (VanderMolen *et al* 2020). In CMS, NASA-supported applications personnel has helped to increase communication and engagement with stakeholders, particularly for projects that funded time within their own grants for stakeholder engagement.

Table 4. Categorical table used to create the Figure 3 correspondence analysis figure, showing the number of project-stakeholder ARL changes were documented for projects that were funded in each year.

	No change	**1 DiffARL**	**2 DiffARL**	**3 DiffARL**	**4 DiffARL**	**5 DiffARL**	**6 DiffARL**	**7 DiffARL**	**8 DiffARL**
2011	291	10	10	15	26	0	0	0	0
2012	0	0	0	0	0	0	0	0	0
2013	26	9	18	6	3	0	0	0	0
2014	199	65	68	22	7	1	9	0	6
2015	24	10	62	25	4	0	1	15	0

2016	38	98	42	2	11	6	0	0	13
2017	0	0	0	0	0	0	0	0	0
2018	33	1	0	0	2	0	0	0	0

Aspects of the transfer agent that seem to capture less of the variance of the DiffARL metric include the institution from where the PI is based, the number of precursor projects and the number of co-investigators funded under the program. In a previous paper, Brown et al (2020) found that the CMS program's ability to provide consistent funding year after year, and to provide engagement and learning of both the agent and the recipient of the data (here the stakeholder) were essential elements of the program.

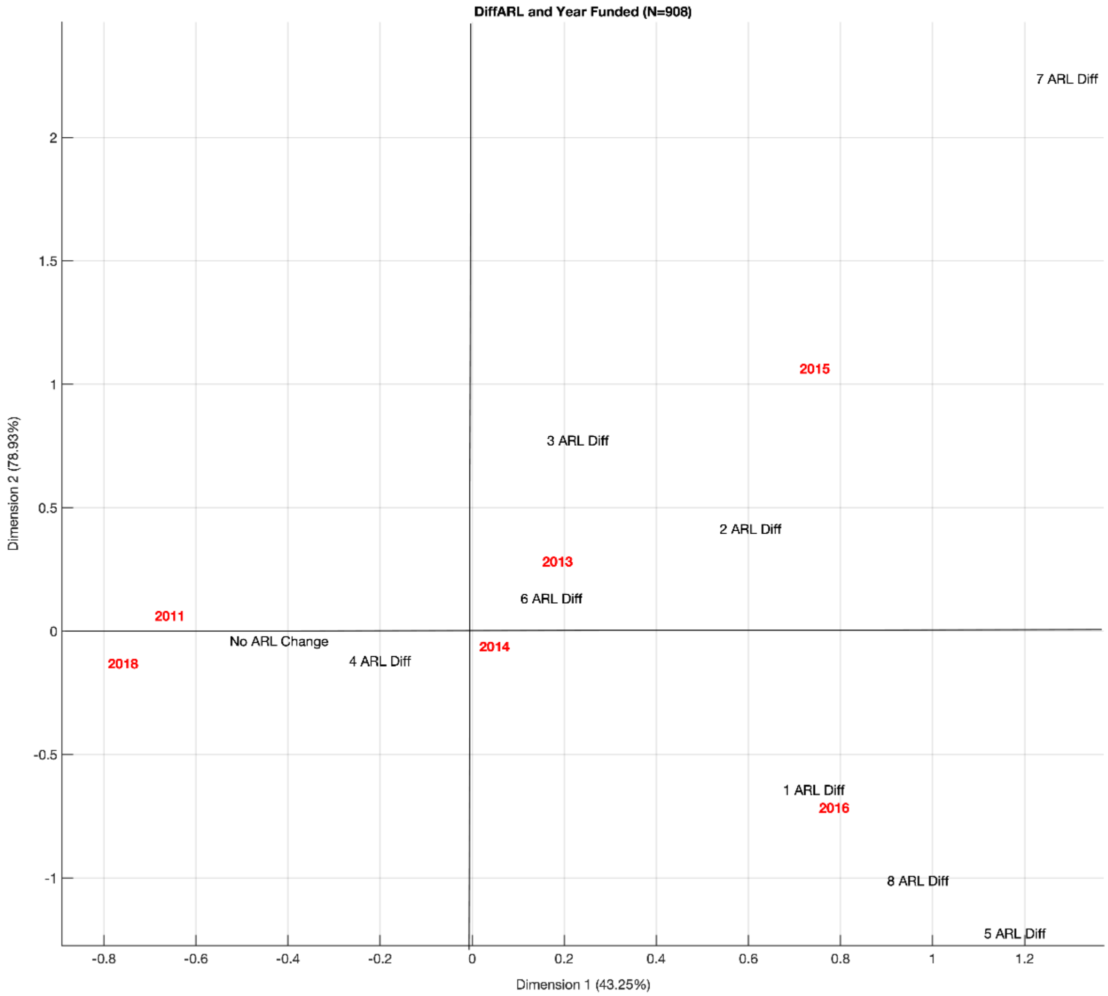


Figure 3. Correspondence analysis between the DiffARL change metric and the year the project was funded. The figure can be interpreted by the closer two elements are to each other, the more similar they are. The further an element is from the 0,0 origin, the more distinctive or different it is from the other elements in the analysis.

6.2 Importance of the Transfer Medium

Few of the variables examined here capture the variability in the transfer medium because we were only able to create variables that captured datasets distributed via the DAAC and not via the CMS PIs to their stakeholders directly, such as web media, videos, and decision support systems (Figure S2). A good example of PI-led data distribution is the NOAA Global Monitoring Laboratory CarbonTracker website that displays and analyzes sources and sinks of carbon dioxide around the world (Butler 2021). However, we did find that the number of dataset citations for datasets distributed via the DAAC explains approximately 75% of the variability of the DiffARL metric, as they relate to the download, use and publication about the use in the peer reviewed literature. Of the projects with data citations of less than 40, 51% were funded in 2011 or 2013, before the CMS project began investing in a broader stakeholder engagement program to support scientists working in the program.

6.3 Transfer Object Variables

Our transfer object variables include the number of products archived, the number of files in the DAAC and the size of the data product. We show that the number of archived datasets and the number files posted at the DAAC by each PI is quite important in explaining the variance of the DiffARL (Figure S3). We found that 56% of the projects with 0-3 archived datasets were funded before 2014. This result may reflect that some projects have not yet finalized their datasets but are still engaging with stakeholders. We found that three datasets are associated with the products that have increased ARL level substantially.

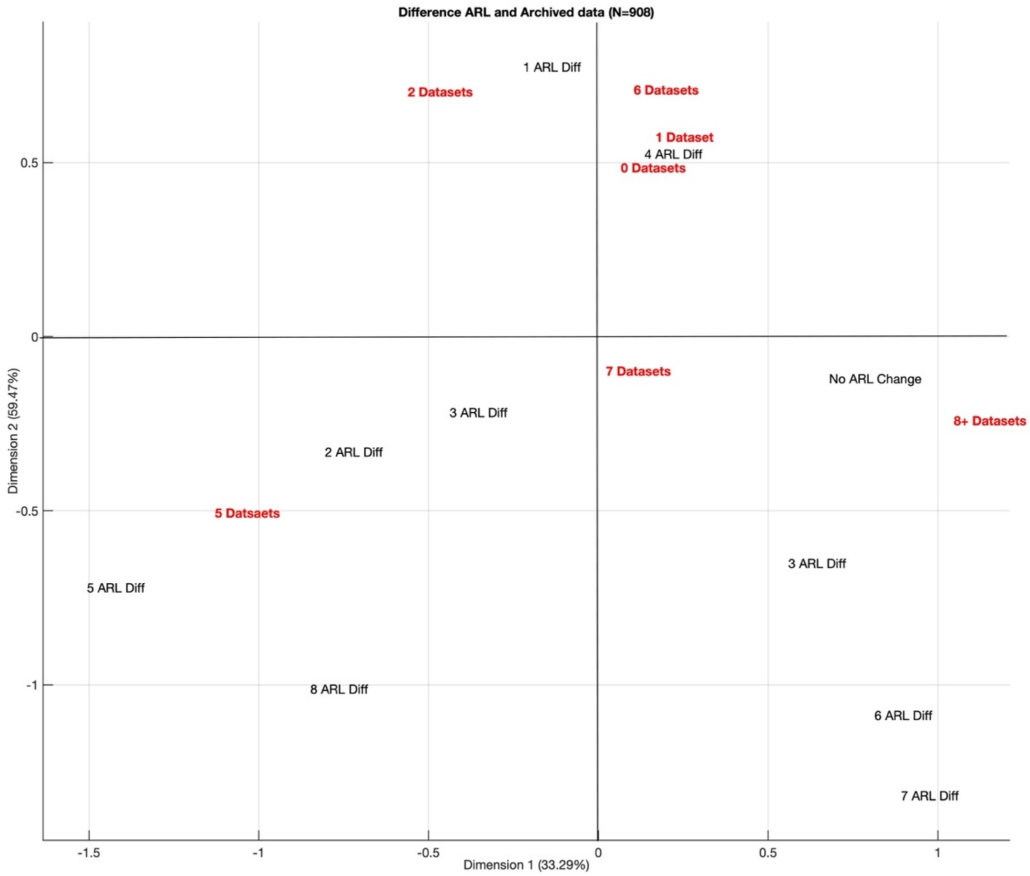


Figure 4. Correspondence analysis map of survey question regarding the number of archived datasets and DiffARL.

6.4 Product Recipient variables

There are more variables describing the product recipient or stakeholder engagement, including the frequency that the PI engages with them, the mechanism through which the communication occurs, when they were first engaged by the PI, the strength of the relationship as described by the PI, and the number of stakeholders were engaged for each product. Of these, the last is most able to capture variability of the DiffARL (Figure S4). We found that 37% of all PIs report that they have between one and two stakeholders, whereas 10% report that they have over 9 stakeholders. The analysis shows that engaging with more stakeholders is not necessarily better for increasing the maturity of each product. Of the 55% who stated that they had engaged the stakeholder ‘a long time ago’, over 80% had an ARL change of 1-3 ARL levels, meaning that although they might have known the stakeholder for a substantial amount of time, the relationship may not be very mature.

6.5 Demand Environment

Finally, the demand environment is critical for understanding how well the CMS PI is to engage with users and increase their ARL levels during the project. One of the most important variables, as indicated by the inertia factor, is the citations of the papers associated with the dataset, which shows the number of other scientists working on the subject being described and the ability of the broader community to hear about and cite the research being conducted to produce that dataset (Figure S4).

7.0 Discussion

Access, awareness and availability are key to the use and uptake of products by stakeholders. We found that the hypotheses that the scientist's ability to communicate about their product via publications, and the length of time engaged with the stakeholder were key factors in their effectiveness in creating useful carbon products and transferring them to support decision making. Our finding support previous research from Jahn et al. (2012) and Brugger et al. (2016) that demonstrate the ability of a scientist to understand the stakeholder context is critical for uptake. Our quantitative approach revealed the importance of the production scientific articles and datasets as the foundation upon which subsequent use of the data product by stakeholders.

More frequent and decision-targeted engagement with the user during the development of the CMS product increases the awareness of how the product will best integrate into the user framework and directly connects to the stakeholders' needs and decisions. Increased awareness of the product development details has a direct impact on product access and availability to the user and helps the CMS scientist connect with the most relevant organizations. The feedback from the user can help drive the access and availability of the CMS products, directly increasing the use and familiarity, and ultimately increasing the products' ARL through the life of the project.

This paper provides a method that allows for quantitative analysis of scientist surveys to explore drivers of increased product engagement. There has been substantial amount of research showing that meaningful interaction between a scientist and a stakeholder during product development should increase the use of scientific information (Lemos and Morehouse 2005, Arnott *et al* 2020b), with others finding that even with relevant information and an engaged stakeholder, there are significant barriers for scientists to engage effectively with potential users of information (VanderMolen *et al* 2020) (Figure S5). Here we find that a quantitative approach can help identify characteristics of a funding program and actions that a scientist can take to increase their success in moving from basic research to application (Whitney and Leshner 2004).

As previous research has found, characteristics of the stakeholder or recipient of the CMS data are important. Our research shows that the maturity of the user relationship with the PI at the proposal stage of the project is related to how much the ARL evolves during the period of performance (Figure S6). Maturity of relationship, which can be measured through letters of interest and other documentation submitted with the proposal, can be encouraged by clearly and consistently funding CMS projects that build on existing relationships. As Arnott (2020a) points out, funders of science are receptive to new ways of revisiting the 'social contract' for science so that co-production of knowledge can be prioritized. Ensuring CMS scientists prioritize relationships as well as producing products and writing papers is essential.

Engaging with stakeholders frequently, providing transparency on product capabilities and limitations, and integrating feedback while creating a strong relationship with them was also found enhance change in applications readiness. Being transparent about capabilities through frequent communication reduced confusion related to access, awareness and availability, and further strengthens the user/PI trust and relationship. CMS products that were able to achieve this were also more likely to be funded in sequential years and continue to evolve their ARL. Of all products, 43% had no precursor projects and were new to the CMS program. We also find that 52% of projects with one and two precursor projects, were more likely to report an increase in ARL level increase that was higher than those with no precursors.

7.1 Limitations

An important limitation of this research is the focus on using scientist survey results as a proxy for stakeholder use of data products. We are limited by the active participation of the CMS PIs in the survey, and their perceptions as they answer questions on their relationships and engagement with stakeholders. The rigor with which they apply the ARL framework to the stakeholder's use of their product is also a critical limitation. After working in CMS for several years, most PIs are extremely aware of the importance of engagement, and therefore may report a better relationship with stakeholders than is the case. To compare the scientist provided ARLs to those provided by a stakeholder, we interviewed 12 CMS stakeholders in 2021. Of the products reviewed, we found that only 36% of the stakeholders disagreed with the scientist-provided ratings by more than one ARL level, but these were evenly split between the stakeholders who believed the product was more mature than the scientist provided (a higher ARL), and those that said it was less mature (a lower ARL). We recognize the complexity of assigning ARLs, which both scientists and stakeholders find challenging, and the different perspectives that a policy maker has from the developer of the product. Further work is needed on evaluating the consistency of ARL ratings across different communities.

7.2 Significance for Policy and Funding of Carbon Datasets

The CMS project provides a consistent funding stream for scientists and stakeholders who engage with them. The result has been the development of a community of practice that has a coherent engagement of carbon and decision support topics (Brown et al 2020). Annual CMS Science Team meetings, required for CMS funded project scientists, include a one-day applications workshops, where stakeholders identified to be working with projects are invited to present their projects either in a talk or in a poster. At the 2020 meeting, 12 active stakeholders presented, and noted the importance of CMS products across a range of applications including: the role of forests in climate mitigation planning, implementing urban canopy targets, wetland and mangroves carbon monitoring, and monitoring of aquatic and marine primary productivity. Stakeholders also noted remaining data needs and gaps, obstacles, or barriers to use, and other programmatic activities CMS could do better. As the CMS project continues, additional investment in stakeholder engagement has been made, including providing more opportunities for stakeholders to attend the CMS science meetings virtually, participate in surveys and interviews from the CMS Applications team to determine their challenges and needs.

8.0 Conclusions

The maturity of relationships between scientists and stakeholders can be encouraged through both relationship building before the grant is submitted and through more rigorous review of letters of support and clear expression of how the CMS scientist intends to engage with the stakeholder. There are numerous important additional applications that could be supported with CMS products as the need for carbon information grows. The ongoing user engagement continues to inform ways in which CMS data can be applied stakeholder needs.

We found that assessing product maturity with PI-applied Applications Readiness Levels was able to capture investments in stakeholder relationships by CMS PIs. We were able to document changes in product maturity through PI-reported ARL levels, offering a potential management tool that could be used in applications programs seeking to develop datasets usable by stakeholders. The method has the potential to determine the success of the CMS program in achieving its goals of putting data into the hands of decision makers.

New ways to use carbon products should be identified in every stage of the program, and that capacity building is needed to help both existing and newly identified stakeholders better understand and use CMS products. As Federal, state, and local policies on climate accelerate, the need for information on carbon will expand, as will the need for feedback from decision makers at all scales. CMS is an appropriate prototype for generating and using datasets to support this need and to continue assessing the community needs for carbon science in society.

Funding

This research was funded by a 2018 NASA grant # 80NSSC20K0853 from the NASA Carbon Monitoring System.

Conflicts of interest/Competing interests – The authors declare no conflict of interest.

Availability of data and material – All data used in this study are available at https://carbon.nasa.gov

Code availability – Custom code using Matlab will be shared upon request via email from the corresponding author.

Authors' contributions

MB and VE conceptualized the approach and datasets, FY, ESC and MM generated the data, VE, PG and GH helped write the paper and MB did the analysis and wrote the paper.

9.0 References

Allen M and others 2018 IPCC, 2018: Summary for Policymakers *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,* ed V Masson-Delmotte, Z Zhai, H-O Pörtner, D Roberts, S Skea, P R Shukla, A Pirani, W Moufouma-Okia, C Péan, R Pidcock, S Connors, J B R Matthews, Y Chen, X Zhou, M I Gomis, E Lonnoy, M Maycock, M Tignor and T Waterfield (Geneva, Switzerland: World Meteorlogical Organization) p 32

Arnott J C, Kirchhoff C J, Meyer R M, Meadow A M and Bednarek A T 2020a Sponsoring actionable science: what public science funders can do to advance sustainability and the social contract for science *Curr. Opin. Environ. Sustain.* **42** 38–44

Arnott J C, Neuenfeldt R J and Lemos M C 2020b Co-producing science for sustainability: can funding change knowledge use? *Glob. Environ. Chang.* **60** 101979

Bozeman B 2000 Technology transfer and public policy: a review of research and theory *Res. Policy* **29** 627–55

Brown M E, Cooper M W and Griffith P C 2020 NASA's Carbon Monitoring System and Arctic-Boreal Vulnerability Experiment (ABoVE) social network and community of practice *Environ. Res. Lett.*

Brugger J, Meadow A and Horangic A 2016 Lessons from first-generation climate science integrators *Bull. Am. Meteorol. Soc.* **97** 355–65

Butler J 2021 NOAA Global Monitoring Laboratory Online: https://www.esrl.noaa.gov/gmd/

Carlo E S, Delgado Arias S, Forgotson C, Hurtt G, Griffith P and Escobar V M 2018 *Identifying Data Needs and Gaps, and Ways to Improve Decision-Making of Stakeholders via the NASA Carbon Monitoring System Applications Project* (Greenbelt, MD)

Cash D W, Borck J C and Patt A G 2006 Countering the loading-dock approach to linking science and decision making: comparative analysis of El Niño/Southern Oscillation (ENSO) forecasting systems *Sci. Technol. Hum. values* **31** 465–94

Cash D W, Clark W C, Alcock F, Dickson N M, Eckley N, Guston D H, Jäger J and Mitchell R B 2003 Knowledge systems for sustainable development *Proc. Natl. Acad. Sci.* **100** 8086–91

Changnon S A, Ravenscroft R, Pilkey O H, Mattingly S, Walaker D, Fellows J, Pendleton J M, Brunner R, Stewart T R, Gauteir D and others 2000 *Prediction: science, decision making, and the future of nature* (Island Press)

Cook B R and Zurita M de L M 2019 Fulfilling the promise of participation by not resuscitating the deficit model *Glob. Environ. Chang.* **56** 56–65

Ernst A 2019 Review of factors influencing social learning within participatory environmental governance *Ecol. Soc.* **24**

Fazey I, Bunse L, Msika J, Pinke M, Preedy K, Evely A C, Lambert E, Hastings E, Morris S and Reed M S 2014 Evaluating knowledge exchange in interdisciplinary and multi-stakeholder research *Glob. Environ. Chang.* **25** 204–20

Gibbons M 1999 Science's new social contract with society *Nature* **402** C81--C84

Greenacre M 1994 Multiple and joint correspondence analysis *Corresp. Anal. Soc. Sci.*

Greenacre M and Hastie T 1987 The geometric interpretation of correspondence analysis *J. Am. Stat. Assoc.* **82** 437–47

Greenacre M J 1984 *Theory and Applications of Correspondence Analysis.* (London: Academic Press)

Hurtt G C, Andrews A, Bowman K, Brown M E, Chatterjee A, Escobar V, Fatoyinbo L, Griffith P, Guy M, Healey S P and others 2022 The NASA carbon monitoring system phase 2 synthesis: scope, findings, gaps and recommended next steps *Environ. Res. Lett.*
Hurtt G, Wickland D, Jucks K, Bowman K, Brown M E, Duren R M, Hagen S and Verdy A 2014 NASA Carbon Monitoring System: Prototype Monitoring, Reporting, and Verification
Hurtt G, Zhao M, Sahajpal R, Armstrong A, Birdsey R, Campbell E, Dolan K, Dubayah R, Fisk J P, Flanagan S and others 2019 Beyond MRV: high-resolution forest carbon modeling for climate mitigation planning over Maryland, USA *Environ. Res. Lett.* **14** 45013
Jahn T, Bergmann M and Keil F 2012 Transdisciplinarity: Between mainstreaming and marginalization *Ecol. Econ.* **79** 1–10
Lasswell H D 1971 *A pre-view of policy sciences* (Elsevier publishing company)
Lemos M C and Morehouse B J 2005 The co-production of science and policy in integrated climate assessments *Glob. Environ. Chang.* **15** 57–68
Liobikien\.e G . and Butkus M 2018 The challenges and opportunities of climate change policy under different stages of economic development *Sci. Total Environ.* **642** 999–1007
Ludwig D 2001 The era of management is over *Ecosystems* **4** 758–64
Meadow A M, Ferguson D B, Guido Z, Horangic A, Owen G and Wall T 2015 Moving toward the deliberate coproduction of climate science knowledge *Weather. Clim. Soc.* **7** 179–91
Moran M S, Doorn B, Escobar V and Brown M E 2015 Connecting NASA science and engineering with Earth science applications *J. Hydrometeorol.* **16**
Moser S C and Ekstrom J A 2010 A framework to diagnose barriers to climate change adaptation *Proc. Natl. Acad. Sci.* **107** 22026–31
NASA 2017 *The Application Readiness Level Metric* Online: https://www.nasa.gov/sites/default/files/files/ExpandedARLDefinitions4813.pdf
Ray D K, Mueller N D, West P C and Foley J A 2013 Yield Trends Are Insufficient to Double Global Crop Production by 2050 *PLoS One* **8** e66428
Sarewitz D and Pielke R A 2007 The neglected heart of science policy: reconciling supply of and demand for science *Environ. Sci. Policy* **10** 5–16
VanderMolen K, Meadow A M, Horangic A and Wall T U 2020 Typologizing Stakeholder Information Use to Better Understand the Impacts of Collaborative Climate Science *Environ. Manage.* **65** 178–89
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
Wenger E 2011 Communities of practice: A brief introduction
West T O, Gurwick N, Brown M E, Duren R, Mooney S, Paustian K, McGlynn E, Malone E, N. Hultman N, Ocko I B and Rosenblatt A 2018 Chapter 18: Carbon cycle science in support of decision making *Second State of the Carbon Cycle Report (SOCCR2): A Sustained Assessment Report* ed N Cavallaro, G Shrestha, R Birdsey, M A Mayes, R Najjar, S Reed, P Romero-Lankao and Z Zhu (Washington, DC, USA: U.S. Global Change Research Program)
Whitney P L and Leshner R B 2004 The transition from research to operations in Earth observation: the case of NASA and NOAA in the US *Space Policy* **20** 207–15
Wilbanks T J and Fernandez S 2014 *Climate change and infrastructure, urban systems, and vulnerabilities: Technical report for the US Department of Energy in support of the national climate assessment* (Island Press)