

OAE Model Skill Assessment Workshop *Proceedings Document*

Quartier, Garmisch-Partenkirchen, Germany

June 29 – July 3, 2026

Table of Contents

Executive Summary	3
1 Introduction	4
1.1 Organizers and Participants	4
1.2 Scope and Objectives	5
2 Level- and scene-setting	7
2.1 Foundational Texts	7
2.2 Curated Input Presentations	8
3 Structural workshop elements	13
3.1 Oceanographic process groups	13
3.2 Stages of an OAE intervention	14
4 Workshop Exercises	15
4.1 Establish fit-for-purpose criteria	15
4.2 Define model-observation comparisons	15
4.3 Identify uncertainty sources, quantification, and reduction methods	16
5 Emerging R&D Opportunities	19
5.1 R&D Opportunity Exercise	19
5.2 Catalytic Grants	20
6 Next steps	22
7 Acknowledgements	22
Appendix	23
Full Participant List	23
Agenda	23

Authors

Anna Madlener, Jens Daniel Müller, Lauren Moseley, Irene Polnyi, David P. Keller, Laura Stieghorst

Project Lead & Contact

Anna Madlener, Senior Manager, MRV
anna@carbontosea.org

Photo Credits

Front page: Quartier

Workshop photos: Viktoria Pezzei for Carbon to Sea, workshop participants

Publication Date

September 30, 2026

Executive Summary

From June 29 to July 4, 2026, Carbon to Sea hosted the ocean alkalinity enhancement (OAE) Model Skill Assessment Workshop in Garmisch-Partenkirchen, Germany, bringing together 35 researchers, practitioners, and private sector representatives to lay the groundwork for model skill assessment guidelines for monitoring, reporting, and verification (MRV) of OAE. A steering committee composed of six experts helped guide the design of the workshop since November 2025 and led five oceanographic process-oriented groups: dissolution & conversion of alkalinity, biogeochemical cycles, physical transport mechanisms, air-sea gas exchange, and sediment biogeochemistry. Across these groups, participants worked to define fit-for-purpose criteria, model-observation comparison approaches, and uncertainty quantification methods as essential layers of model skill assessment that are reasonably implementable today.

A set of curated input presentations grounded our discussions in the realities of the OAE field today. The Intergovernmental Panel on Climate Change (IPCC) does not yet consider OAE in the development of removal methodologies due to insufficient peer-reviewed literature on OAE's quantification methods. It is crucial to provide this literature, including on effective model assessment, by the time the IPCC revisits methodologies in 2029. This timeline became an important anchor as participants considered what model skill assessment methods are feasible today, versus desirable and realistically achievable in the near future.

Participants drew lessons from adjacent, more mature industries such as watershed monitoring, ecotoxicology, and fisheries, all of which took years or decades to establish accepted fit-for-purpose criteria and skill metrics, underscoring that our own path will be iterative rather than immediate. Working towards acceptable uncertainty levels, moving beyond supplier-specific MRV protocols, and increasing the number of deployments to build cross-site model validation datasets and statistical relevance is critical in spurring further buyer interest and increasing OAE's qualifications for eventual compliance markets.

Beyond the core guidelines work, the workshop surfaced a pool of R&D priorities spanning foundational science, tools and decision-support, and longer-term efforts. Several of these priorities are now under further assessment by participants who received catalytic grants following the workshop.

This document captures the design, discussions, and early insights of the workshop. It is not the guidelines themselves; rather, it's a record of how we got there and an invitation for the broader community to engage as the work moves toward a public comment period later this year.

A lead editors team consisting of Alicia Karspeck, Tyler Rohr, Veronica Tamsitt, and Anna Madlener is harmonizing group outputs into a single set of model skill assessment guidelines for public comment. A separate lead authorship team coordinated by Lauren Moseley is preparing the core aspects of the guidelines and R&D priorities as a peer-reviewed perspectives paper. Questions about this process, the workshop outcomes, or further involvement can be directed to Anna Madlener, anna@carbontosea.org.

1 Introduction

From June 29 to July 4, 2026, Carbon to Sea hosted the Ocean Alkalinity Enhancement (OAE) Model Skill Assessment Workshop (OAE MSAW) in Garmisch-Partenkirchen, Germany. The workshop invited 35 participants who worked in five groups to develop guidelines for assessing the skill of regional models for OAE quantification and identify R&D needs. The goal of the workshop was to develop model selection and skill assessment guidelines immediately applicable and implementable for monitoring, reporting, and verification (MRV) purposes, as well as to distinguish near- versus long-term R&D perspectives in a separate Perspective Paper. In this Proceedings Document, we aim to share the design, insights, and takeaways from the workshop and invite collaboration in shaping its outcomes over the next months.

Marine biogeochemical models are essential for OAE research and for quantifying carbon dioxide removal (CDR) in monitoring, reporting, and verification (MRV) practices. While these models have proven to skillfully simulate ocean circulation and emission-induced CO₂-uptake, key OAE-relevant processes still need better or entirely new representations¹. A targeted model skill assessment is crucial to guide such upgrades and transparently quantify uncertainties of the simulated CDR². In academia, publication of a model-observation comparison in a peer-reviewed journal often serves as de facto "validation," but this approach may not be sufficient for OAE MRV and private sector carbon crediting if it overlooks processes that matter most for OAE, does not follow a consistent evaluation criteria, or produces an unconstrained uncertainty budget.

Recent private sector work on OAE has started to operationalize biogeochemical model simulations in a crediting context. These modeling efforts typically consist of a suite of near-field, regional, and global models that are evaluated against tracer experiments³, and baseline physical and biogeochemical observations. Guidelines for these evaluations are externally defined through a registry and reviewed by independent verifiers. In-trial measurements offer important opportunities to further assess the models' capabilities to reproduce an alkalinity perturbation, the induced CO₂ uptake, and potential biogeochemical feedbacks. This workshop aimed to advance the field of OAE modeling by establishing consensus across academic and private sector participants for the skill assessment and uncertainty quantification of models currently used to simulate OAE.

1.1 Organizers and Participants

Carbon to Sea initiative was the convener, funder and organizer of the workshop. Lead organizers included Anna Madlener and Dr. Jens Daniel Müller. Scientific and strategic oversight and support was provided by Dr. David Keller, Irene Polnyi and Dr. Lauren Moseley.

Steering committee members were selected based on their different expertise and experience with respect to modeling, their deep literacy of the state of the science and associated knowledge gaps in OAE, modeling and MRV questions, as well as their recognition for being able to work collaboratively across disciplines. Steering committee members shaped the scope and design of

¹ [Ward et al, 2025](#)

² Fennel et al., [2022](#), [2023](#)

³ [Subhas et al, 2025](#), [Heede et al, 2026](#) (preprint), [Gerke et al, 2026](#)

the workshop, formalized outcomes, selected participants, and led working groups throughout the workshop. The steering committee included: Dr. Abby Lunstrum (University of Chicago), Dr. Adam Subhas (Woods Hole Oceanographic Institution), Dr. Alicia Karspeck ([C]Worthy), Dr. Elizabeth Yankovsky (Yale University), Dr. Pedro Monteiro (Stellenbosch University) and Dr. Veronica Tamsitt (Submarine Scientific).

Workshop participants comprised a multi-disciplinary group of researchers, practitioners, and private sector representatives, selected for their relevant oceanographic expertise. The workshop conveners sought to establish a mix of disciplines, seniority, sector representation, gender, and geographies and were recruited both directly and through an open [application process in April](#). A full list of participants can be found in the group work section and at the end of this document.



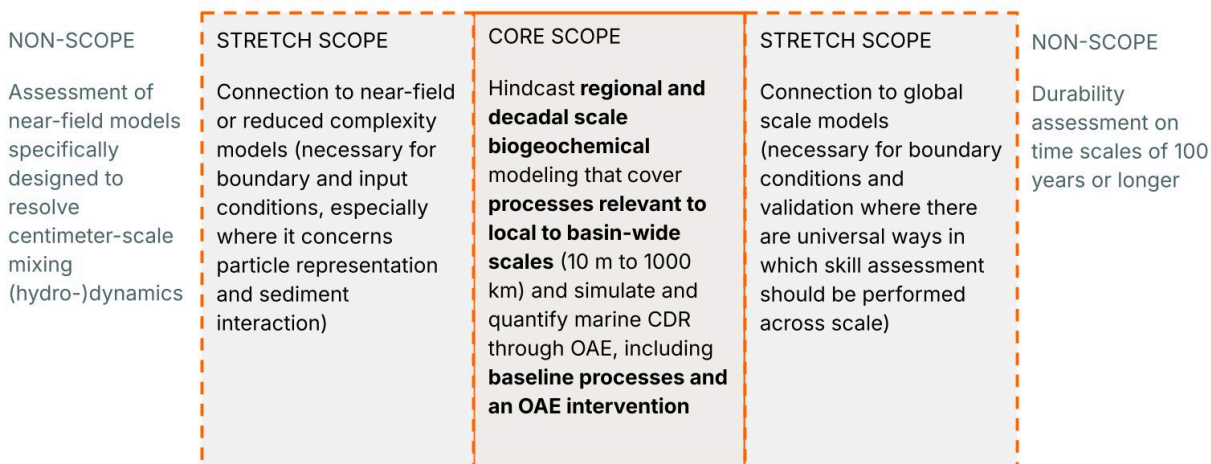
Figure 1: Workshop participants

1.2 Scope and Objectives

The workshop's primary objective was to lay the foundation for model skill assessment guidelines that can be prepared for a public comment period. Rather than focusing on a theoretical future in which ideal conditions for model design and validation exist — for example, inexpensive and ubiquitous observations of carbonate chemistry, computationally unconstrained modeling access, or R&D ideas that have resulted in improved methodologies — participants were specifically asked to think in terms of assessment activities that could reasonably be done right now, using current observations, tooling, and the budgetary constraints that a research or private sector project-deployment team would be operating within.

To this end, the workshop focused on the skill assessment of **OAE-induced additional CO₂ uptake simulated with regional and decadal scale biogeochemical ocean modeling**. Regional ocean biogeochemical models play a key process-integrating role in OAE quantification, because they are capable of coupling the relevant seawater chemistry, biological processes, physical ocean state

and transport, and air-sea gas exchange that govern the cycling of carbon in the ocean and the ultimate removal of additional CO₂ due to an OAE activity.



The secondary focus, “stretch scope,” lay on roadmapping the extensions of the skill assessment methodology to near-field and global models, including a path towards coupling different model scales. Although model evaluation should eventually also cover wider environmental impacts, this was not a focus of this workshop.

Three layers of model skill assessment were defined as a primary organizing principle and required outcome for model skill assessment guidelines:

1. **General fit-for-purpose criteria:** Model specifications (e.g., processes, resolution and scales) required to demonstrate suitability for OAE MRV
2. **Model-observation comparisons:** Approaches and data requirements for model-observation comparisons between baseline processes and OAE intervention
3. **Uncertainty quantification:** Statistical approaches, metrics, confidence bounds, and error propagation required for determining, reporting, and accounting for CDR uncertainty

Additionally, participants contributed ideas to a collective pool of R&D priorities to elevate model skill assessment in the near (~3 year) future. These could range in scale from individual projects to larger-scale initiatives that would help put the entire community on track to address complex sources of uncertainty and needed model development. These priorities will be shared in a peer-reviewed perspectives paper alongside a summary of the alignment on operational and feasible assessment methods published in the model skill assessment guidelines.

2 Level- and scene-setting

A collection of foundational texts and curated input presentations provided common taxonomies, references and relevant context to shape our work.

2.1 Foundational Texts

Participants received introductory reading material in order to prepare for various aspects of the workshop. This included an **uncertainty taxonomy** to facilitate categorization and common language during the time spent in person, a **statistical metrics and methods glossary** that summarizes common metrics of skill assessment, uncertainty reduction, and quantification methods, and a **market crediting overview** of external reference material related to current OAE MRV and modeling protocols and practices.

Uncertainty Framework

Across the OAE and MRV field, the same underlying uncertainty phenomena are often described using different terms depending on the scientific, commercial, or regulatory contexts. **This fragmentation affects what gets quantified, what gets reported, and makes it difficult to see where additional steps and investments can potentially reduce uncertainty over time.** To address this, the organizers developed an uncertainty framework organized into uncertainty categories (process, measurement, and model) and types (e.g., parametric uncertainty) per category. In the workshop, participants spent a full day pressure-testing and completing the framework by producing a (near-)exhaustive list of uncertainty sources per type (e.g., gas transfer velocity as a parametric uncertainty in models), assigning potential magnitudes and responsibilities, as well as documentation of measurement and modeling methods to quantify and reduce the uncertainty source (see Section 4.3).

Statistical Metrics and Methods Glossary

The Statistics Glossary provided non-exhaustive reference tables covering two complementary dimensions of model skill assessment: (1) common statistical metrics used to quantify modeling uncertainties and evaluate model skill (e.g., bias and mean-state metrics), and (2) methods for quantifying and/or reducing sources of uncertainties (e.g., model-observation comparisons, model-intercomparison projects, observing system simulation experiments (OSSEs)). One of the core goals of producing model skill assessment guidelines in this workshop was to provide concretely acceptable and commonly used skill evaluation metrics and methods as part of the final outcome. This overview was a resource amongst participants and, along with the uncertainty framework, will be adapted based on workshop discussions and shared in the final guidelines as appropriate.

Market Crediting Overview

To establish the baseline for current model skill assessment in MRV, the organizers also highlighted public documentation from previous crediting events. This guidance was designed to help participants understand current practitioner efforts and identify potential areas for improvement. These references and subsequent discussions highlighted that early private sector operations,

when conducted carefully and using structured, stage-gated approaches, offer a valuable testing ground to adopt and improve high-integrity MRV practices.

Table 1: Overview of external reference documents

Reference Documents	
Isometric Protocol	OAE Protocol references: Air-Sea CO₂ Uptake Module Protocol Section 10: Pre-deployment requirements Protocol Section 11: Monitoring Requirements
Planetary PDD, Monitoring, and Quantification Reports	Example Project-level MRV Reporting: Project Page Project Design Document (March 2026) , Section 1.4 (p. 52), Section 1.5 Models (p. 65-75) Monitoring & Operations Report (March 2026) , particularly Section 6 Research & Development (p 24-36) Quantification Report (March 2026) , particularly Section 6 Determination of CDR _{gross} <i>Additional references in previous issuances:</i> Dec '25 , Oct '25 , Jun '25 (date refers to the point of credit issuance)
Credit Background Reading	Carbon to Sea Overview of first OAE credits: The First OAE Credits: An Independent MRV Review

2.2 Curated Input Presentations

The organizers identified a set of perspectives to inspire and shape the working discussions through live and virtual presentations.

Table 2: Overview of input presentations

Presenter	Title
Dr. Frauke Kracke, Frontier <i>REMOTE</i>	<i>What does a CDR buyer need from ocean modelers?</i>
Dr. Alice Alpert, Environmental Defense Fund <i>REMOTE</i>	<i>Biogeochemical Model Skill Assessment for MRV: Key motivations and considerations</i>
Dr. Judith Hauck, Alfred Wegener Institute <i>REMOTE</i>	<i>Ocean biogeochemical model skill assessment in the Global Carbon Budget</i>
Dr. Lennart Gerke, Dr. Robert Izett, Dr. Adam Subhas	<i>Photo show: three operational examples from field studies</i>
Dr. Heather Kim, Woods Hole Oceanographic Institution <i>REMOTE</i>	<i>Modeling the LOC-NESS OAE field trials: Site selection, model skill, and sources of CDR uncertainty</i>
Dr. Ruth Musgrave, Dalhousie University	<i>Quantifying uncertainties in estimates of marine carbon dioxide removal using reduced complexity models</i>

Dr. Kristi Weighman, Hourglass Climate
REMOTE

Modeling Environmental Risk: Charting mechanistic modeling pathways from research to regulation

Dr. Benjamin Saenz, [C]Worthy REMOTE

Operational Watershed and Water Quality Models in California: Requirements, Skill, and Regulation

Dr. Charles Stock, National Oceanic and Atmospheric Association REMOTE

Ocean Predictions For Climate-Informed Ecosystems and Fisheries Management

Role of model skill assessment for ensuring climate ambition

Dr. Alice Alpert, Senior Scientist at the [Environmental Defense Fund](#), provided a general overview of the role of CDR in climate scenarios with particular focus on the Intergovernmental Panel on Climate Change (IPCC) and called for a strategic research approach to close knowledge gaps. She explained the IPCC's role in creating methodologies for countries to quantify removals for Paris Agreement commitments, noting that OAE is not currently included due to a lack of sufficient peer-reviewed scientific literature for quantifying removals. Specifically, Alpert shared MRV priorities that the OAE community should consider in order to provide the needed knowledge for an IPCC methodology development by the time the next expert meeting in 2029 takes place:

- 1) Field experiments that are accompanied by fit-for-purpose, site-specific modeling,
- 2) Validating model skill on perturbed and unperturbed conditions, and
- 3) Expressing overall uncertainties taking into account all potential sources.

Keeping an eye on the 2029 timeline was an important perspective throughout the week to decipher early alignment and key research priorities, solvable in the near term, from broader scientific knowledge gaps within the field of OAE.

Role of models in driving market legibility

Dr. Frauke Kracke, Head of Science at the advanced market commitment [Frontier](#), shared insights into Frontier's purchasing approach for CDR in the voluntary carbon market. Frontier aims to accelerate the development of carbon removal methods through catalytic funding and long-term offtake agreements. Kracke noted that while some long-term OAE offtake contracts have been signed for specific, well-characterized regions, broader buyer confidence for OAE in general requires more robust, defensible, and auditable MRV frameworks. She also emphasized that large-scale success of the CDR field including OAE requires compliance markets. In Kracke's views, critical OAE bottlenecks currently consist of the difficulty to detect signals through measurements especially at small scale, limited standardized frameworks, and slow emergence of new deployments. Further, she named a lack of agreement about acceptable uncertainty levels,

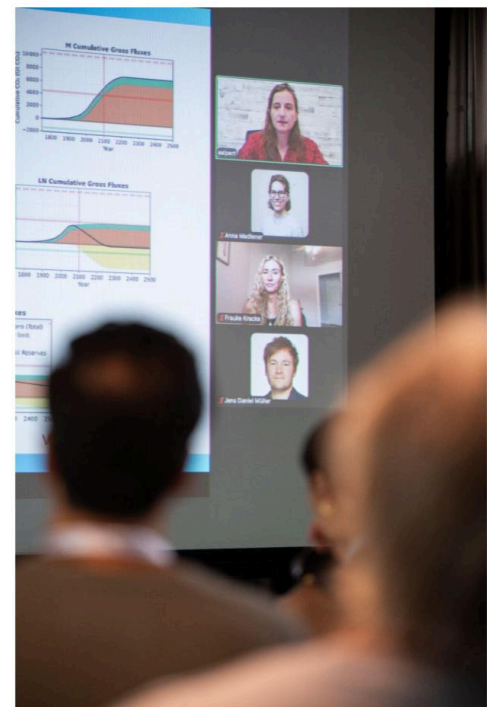


Figure 2: Audience listening to Dr. Frauke Kracke and Dr. Alice Alpert presenting virtually.

the supplier-specific MRV protocol focus, and limited number of deployments to build cross-site model validation datasets and statistical relevance as a key motivating perspective for the workshop. Kracke also highlighted the challenge of MRV to meet scientific rigor while still being operationally scalable and cost-effective, underscoring Frontier's philosophy that early and iterative crediting can fund the research necessary to gradually improve crediting frameworks in the future.

Parallels from the Global Carbon Budget

Dr. Judith Hauck, deputy head of Marine Biogeosciences at the [Alfred-Wegener-Institute](#), presented on uncertainty quantification and model skill assessment approaches implemented in the Global Carbon Budget (GCB)⁴, which operationally quantifies carbon fluxes between the atmosphere, ocean, and terrestrial biosphere on an annual basis. The GCB uses a combination of observation-based machine learning products and dynamic ocean models to estimate the ocean carbon sink. The ensemble of GCB models is systematically evaluated against physical variables and carbonate system metrics with the Integrated Ocean Model Benchmarking (IOMB)⁵ tool. These benchmarks allow for qualitative assessment, identifying which models perform best in specific metrics, though no single model excels in every category. In her presentation, Hauck inferred that for regional OAE assessments, models must accurately represent local spatial and temporal variability, particularly in mixed layer depth, which strongly influences additional carbon flux and alkalinity redistribution. Hauck stressed that while global carbon sink models currently focus on physics and carbonate parameters, regional OAE applications may require the inclusion of biological systems to achieve necessary accuracy. During discussion with participants, Hauck pointed out that there is currently no community-agreed threshold for model scores, but agreed that this could be a viable path forward for MRV. To achieve this, Hauck emphasized the need for community consensus on which metrics are most critical for regional MRV, acknowledging that these criteria may evolve over time.

OAE in the real world: three operational examples from field studies

Model development and skill assessment for OAE requires real-world experiments to provide critical data. As in many oceanographic disciplines, OAE field research is complex and logistically challenging and may affect what is operationally possible. To illustrate this, three experienced OAE field researchers shared practical insights from past and ongoing field campaigns using illustrative photos and videos.

Dr. Lennart Gerke, PostDoc at [\[C\]Worthy](#), shared practical experiences from dual-tracer and dye-release studies in diverse environments, ranging from research vessels to improvised setups on small boats. Dr. Robert Izett, observational lead at [Planetary](#), shared challenges of industrial-scale operations, including working with impure feedstocks, biofouling, and strict regulatory compliance. Dr. Adam Subhas from the [Woods Hole Oceanographic Institution](#) (WHOI) and co-PI of the [LOC-NESS](#) project highlighted the importance of engineering expertise, biological replication, and drone monitoring for near-field plume analysis. Together, the three presentations highlighted the early and challenging nature of OAE field research and the importance of data collection for model development.

⁴ [Friedlingstein et al, 2026](#)

⁵ [Fu et al, 2022](#)



Figure 3: Dr. Lennart Gerke, Dr. Robert Izett, and Dr. Adam Subhas (left to right) sharing insights from OAE research in practice

Learnings from LOC-NESS model skill assessment efforts

Dr. Heather Kim of WHOI further presented the LOC-NESS project's approach towards building model confidence. The modeling framework for the project involved an ocean circulation model coupled with biogeochemistry, with explicit calcification processes added to better represent baseline carbonate chemistry. The presentation emphasized that, while models are critical for planning and estimation, **different choices in physical and biogeochemical model formulation can lead to variations in predicted outcomes, underscoring the need for coordinated model intercomparison.**

Quantifying uncertainty with a simplified 1-dimensional model tool

Dr. Ruth Musgrave, Associate Professor at [Dalhousie University](#), presented an approach towards practical, immediately implementable uncertainty quantification through a 1-dimensional diffusivity model. By varying parameters like interior mixing, deployment depth, and baseline carbonate chemistry, researchers can identify which factors drive the largest uncertainties in CDR calculations. One of Musgrave's key findings is that **the importance of various uncertainty contributions changes between the near-term (~5 years) and long-term (50 years), which can affect uncertainty calculations for MRV.** Further, she shared that alkalinity enhancement beneath the mixed layer results in large, long-lasting uncertainties, highlighting the importance of understanding near-field mixing.



Figure 4: Dr. Ruth Musgrave during her presentation on reduced complexity modeling for uncertainty quantification.

Lessons from adjacent fields: ecotoxicology, watershed, and fisheries management

OAE is not the first oceanographic industry grappling with challenges around model skill assessment. To illustrate this, we invited three virtual presentations to share learnings from adjacent fields in their path towards regulated modeling practices.

Dr. Kristi Weighman, ecotoxicologist at [Hourglass Climate](#), shared how regulators have been applying ecotoxicological models in permitting considerations. These models are used to support high stakes environmental decisions, serving as an example of credibility amongst regulators that while tolerating ambiguity. She highlighted that **specific models generally considered fit-for-purpose in this context have gone through iterations of building one validated process at a time over several decades**, illustrating the likely length of an ongoing path for solidifying regulations in OAE. Evaluating whether a model is ready for regulatory purposes in this field typically relies on a well-defined scope that the model is valid for, well-defined processes of a model that can be examined, and a validation that shows a model meets pre-defined performance criteria.

[C]Worthy's Head of Applications, Dr. Benjamin Saenz, shared insights from previous work in the State of California where model-based decision making informs watershed management that balances watershed competing uses (endangered species, flood control, water quality, utility/diversion) with constraints (water system, dams, climate) and ensures operating within strict regulation. He shared that **quantitative skill metrics were developed over time when operations violated regulation under new climate stresses that were not predicted by established metrics and models**. These metrics were primarily driven by the legal expectation of "best available science," where agencies use precedent to establish metrics but also are expected to leverage new applicable science. An important insight for workshop participants was considering that in the future, models used for OAE may have to adhere to specific performance criteria, as well as legal precedent, in order to withstand scrutiny.

Lastly, Dr. Charlie Stock of the [National Oceanographic and Atmospheric Association](#) (NOAA) reported from his experience developing the [Changing Ecosystems and Fisheries Initiative](#) (CEFI) Modeling System, an example demonstrating how ocean models can mature into decision-support systems and the importance of incorporating decision criteria early. **Stock highlighted the evolution of the modeling system through strong feedback between fishers, decision makers, and modelers to ensure it is fit-for-purpose over time**. Confidence in model systems was built through engaging relevant stakeholders to determine relevant metrics, as well as combining retrospective hindcasts and retrospective forecast experiments. He stressed **the importance of establishing sufficient skill at the beginning to be useful, while laying the foundation for progressive improvement**. Stock's talk highlighted the importance of iterative progress towards confidence in OAE simulations while regularly pressure testing our ability to properly simulate past interventions and eventually predict quantification in a forward-looking manner.

All three examples showed that achieving an understanding of adequate fit-for-purpose criteria, identifying the right skill metrics, and accepted uncertainty management required a somewhat novel process and years or decades to establish. They are all fundamentally driven by an underlying scientific focus that materially affects the choices made.

3 Structural workshop elements

Two organizing principles were leveraged in the workshop in addition to the three layers of model skill assessment: **oceanographic processes** and **research stages**. These were selected and consistently applied to make our exercises and problem-solving more comprehensive.

3.1 Oceanographic process groups

Five process-focused working groups were established to facilitate in-depth exploration of key OAE-related themes. These specific areas were selected because they represent the primary drivers affecting OAE efficiency. Furthermore, this structure recognizes that scientific maturity and regional model representations vary significantly across these topics. By separating these themes, we enabled participants to engage deeply in targeted scientific discussions while acknowledging that, within MRV frameworks, certain processes are more advanced and straightforward to address than others.

Table 3: Overview of core group questions and all members

Dissolution & Conversion	Biogeochemical cycles and carbon storage	Physical transport mechanisms	Air-sea gas exchange	Sediment biogeochemistry
How are alkaline feedstocks converted into seawater alkalinity? How stable is the additional alkalinity on short time scales?	Where does additional alkalinity and DIC accumulate? How well are background BGC dynamics simulated and are they important for CDR? Which BGC processes might be altered and affect CDR?	Where does the alkalized water go? How well do models account for unresolved physical processes?	How do gas-exchange parameterizations, met conditions, waves and geographic settings control additional carbon uptake?	What happens to added alkalinity in sediments? How is sediment biogeochemical cycling affected by enhanced alkalinity?
Adam Subhas Alexander Andriatis Mohammed Hashim Robert Izett Anne Kiely Charly Moras	Jens Daniel Müller Matthew Humphreys Rui Jin Matthew C. Long Tyler Rohr	Elizabeth Yankovsky Pedro Monteiro Aurora Leeson Chih-Lun Liu Ruth Musgrave Daan Reijnders	Veronica Tamsitt Lennart Gerke Hendrik Grosselindemann Jing He Jesse Vance Brian Ward	Abby Lunstrum Anna-Adriana Anschütz Devon Cole Mojtaba Fakhraee Astrid Hylén Olivier Suplis

In addition to defined groups, six participants facilitated the overall workshop ensuring groups worked cohesively, detected, and synthesized themes on a daily basis, and remained goal-oriented throughout.

Table 4: Overview of subject matter experts and facilitators

Role	Members	Core Task
Subject Matter Experts	Jann Paul Mattern David Keller	Statistical approaches guidance OAE-specific guidance, overall scientific steering

Table 4: Overview of subject matter experts and facilitators

Role	Members	Core Task
Facilitation	Lauren Moseley	OAE-specific guidance and OAEMIP linkages
	Alicia Karspeck	Lead facilitator, building strategic cohesion across groups
	Anna Madlener Irene Polnyi	MRV specific guidance, overall workshop steering Strategic partnerships and R&D opportunities

3.2 Stages of an OAE intervention

In addition to the skill assessment layers and oceanographic processes, a research staging framework was used to further guide discussions and allow prioritization. In 2025 and 2026, Carbon to Sea and Plymouth Marine Laboratories developed an Environmental Impact Monitoring Framework (EIMF) that introduces a stage-gated approach to guide the responsible advancement of OAE field research⁶. The framework was used for this workshop to guide model assessment across these same stages and define which aspects of each assessment layer are required and feasible. By generally asking the question what a model must be capable of at each stage, the stages intended to reflect the reality that each phase of OAE field research offers different opportunities and imposes different limitations on how practical and informative model assessment activities can be, e.g., through observations constrained by signal detectability at early stages.

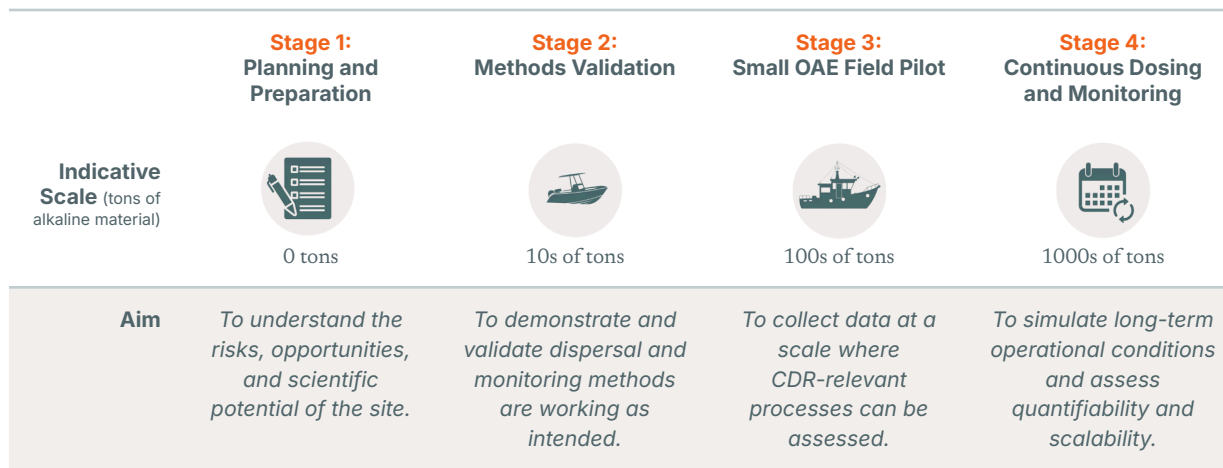


Figure 5: Excerpt of the High Level Framework for Responsible Research Advancement, OAE Environmental Impact Monitoring Framework 2026.

⁶ For further background on the origin of the stage gates, you can download the Environmental Impact Monitoring Framework here: <https://ctsea.info/OAE-EIMF>

4 Workshop Exercises

Throughout the workshop, the groups worked independently through a set of progressive exercises and drafted an initial set of guidelines. Each day, facilitated discussions brought groups together to surface commonalities or tensions that are best resolved as a full group. Group work was structured around three layers of model skill assessment: fit-for-purpose criteria, model-observation comparisons, and uncertainty quantification.

4.1 Establish fit-for-purpose criteria

The goal of this first exercise was to establish model specifications, including required processes, resolution, and scales necessary for robust quantification of OAE. Participants were tasked with defining the minimum process representation and model configuration a model must demonstrate at each stage of an OAE intervention. This served as both a prioritization and accountability exercise, requiring groups to categorize elements as "necessary" versus "safe-to-ignore" at different stages.



Figure 6: Groups A and B as they started their fit-for-purpose criteria discussion.

4.2 Define model-observation comparisons

In this session, participants identified the necessary model-observation comparisons in coherence with their proposed fit-for-purpose criteria as derived in the previous session. The exercise challenged groups to identify both currently available and future-required observations while evaluating signal detectability against specific signal-to-noise ratios at the relevant stages. Participants were tasked with distinguishing between model-observation comparisons of baseline processes versus the OAE intervention signal itself.



Figure 7: Groups D and E during the model-observation comparisons discussion

4.3 Identify uncertainty sources, quantification, and reduction methods

To facilitate the uncertainty exercise, each group utilized pre-printed index cards designed to systematically identify and categorize various sources of uncertainty according to the proposed uncertainty framework. Groups discussed how uncertainties varied by research stage, what their primary drivers are, how to quantify them through measurement and modeling methods, and potential opportunities to reduce them.

Type: Uncertainty source: Stage: 1 · 2 · 3 · 4 Group: A · B · C · D · E Driver: Sampling · Empirical · Ontological · Intrinsic Responsibility: Project · Multi-Institutional · Government	Measurement or Modeling strategies ... <table border="0"> <tr> <td>... to quantify:</td> <td>... to reduce:</td> </tr> <tr><td>.....</td><td>.....</td></tr> <tr><td>.....</td><td>.....</td></tr> <tr><td>.....</td><td>.....</td></tr> <tr><td>.....</td><td>.....</td></tr> <tr><td>.....</td><td>.....</td></tr> <tr><td>.....</td><td>.....</td></tr> <tr><td>.....</td><td>.....</td></tr> </table>	... to quantify:	... to reduce:														
... to quantify:	... to reduce:																
.....																	
.....																	
.....																	
.....																	
.....																	
.....																	
.....																	

Figure 8: Uncertainty index card design

Lastly, assigning a responsibility level to an uncertainty source is critical as not every uncertainty source can reasonably be answered by individual projects. Groups designated project, multi-institutional, or system-level responsibility to each source, per stage:

- Project: uncertainty can reasonably be quantified and/or reduced by standalone projects, and typically involves a high-degree of site-specific information (e.g., Planetary, LOC-NESS)

- Multi-institutional: uncertainty that requires a longer-term, coordinated and dedicated research and/or institutional effort to quantify and reduce, and can typically be generalized across sites (e.g., [Ocean Alk-Align](#), [OAE-PIIP](#), [OAE MIP](#), [NOAA-NOPP mCDR awards](#))
- System-level (called “government” on the cards but revised in hindsight): uncertainty requires system-level funding such as from governments to support long-term quantification and reduction efforts (e.g., [GOOS](#), [SOCCOM](#)).

Taken together across groups, this exercise provided a rich mapping of OAE uncertainties and available methods to address them.

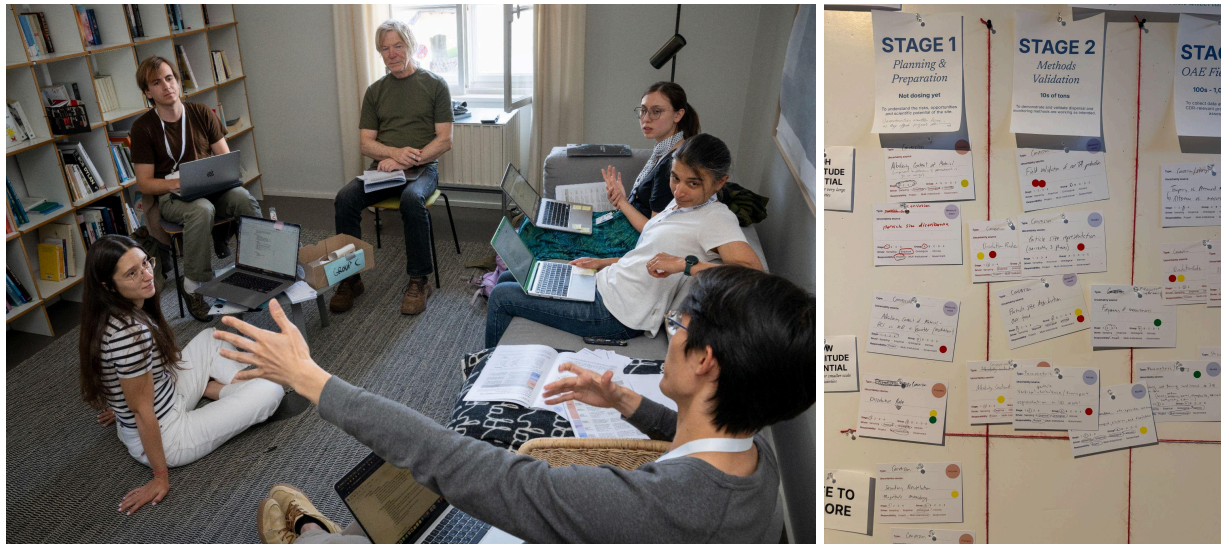


Figure 9: Group C discussing uncertainties and Group A's populated board with uncertainty cards

Cross-group alignment

Across the week, groups shared progress and discussed overlap during readout sessions in the morning and afternoon. This was supported by the lead facilitator, Alicia Karspeck, who regularly rotated through all five working groups to observe discussions in real-time alongside additional subject matter experts. By synthesizing findings during morning readouts and afternoon debriefs, Alicia continuously surfaced high-level themes, identified critical divergences, and connected threads across the process-level groups.



Figure 10: Lead Facilitator Alicia Karspeck during Thursday's morning readout.

Further, a full cross-group alignment day allowed dedicated time for all groups to meet individually, surface shared ownership, align scopes, and flag overlaps on fit-for-purpose criteria and model observations across stages. A few group pairs were converging naturally either thanks to very similar discussions (physical transport and air-sea gas exchange; dissolution and sediments) or because they lay at more opposing process

boundaries (air-sea gas exchange and sediments). In contrast, conversations between biogeochemistry and several other groups (physical transport, dissolution, sediments) surfaced genuine scientific disagreement that the alignment day was meant to expose rather than resolve. Open questions mainly existed around process directionality, near-field versus regional boundaries, the role of sediments, and how uncertainty evolves over time. Questions around the magnitude of the biogeochemical signal were recurring across groups and were flagged for further research. More detail on these overlaps and open needs will be included in the final outcomes of the workshop.

5 Emerging R&D Opportunities

The workshop naturally raised opportunities to advance model skill assessment, such as collecting additional data, running new analysis, and previewing larger projects to improve MRV across the board.

5.1 R&D Opportunity Exercise

A dedicated R&D Opportunity exercise helped crystallize these concepts, categorized broadly into three types:

Foundational science: Fundamental unknowns about mechanistic process and system behavior, necessary to resolve in order to determine their relevance as fit-for-purpose criteria of models, that can be resolved with small, targeted scientific efforts, for example:

- How turbulence, feedstock characteristics, and biogeochemistry jointly affect particle dissolution, settling velocity, and the fate of the added alkalinity
- How sediment-carbonate chemistry dynamics respond to enhanced alkalinity, including alkalinity fluxes, secondary mineral formation, and ion exchange that is currently lacking in essential data
- Whether carbonate equilibrium constants enable internally consistent CO₂ system computations at elevated pH
- What is the direction and likely magnitude of model bias that originates from the omission of organic alkalinity

Tools & decision-support: Applying models, standards, and existing frameworks to help us better prioritize and understand which uncertainty matters most and how to design smarter, more cost-effective field campaigns, for example:

- Shared standardized protocols of feedstock dissolution coefficients and alkalinity enhancement potential
- Sensitivity studies on gas transfer velocity, wind, and atmospheric forcing resolution to help better understand their relative influence on uncertainty
- Idealized process or reduced complexity (1-D) models to help distinguish the importance of dissolution vs sinking vs precipitation
- Observing System Simulation Experiments (OSSEs) to inform and improve observational campaigns to best assess models

Longer-term efforts: Larger projects that rely on more robust empirical datasets, require sustained, multi-year efforts, or must be further validated and scoped, for example:

- Near field to regional ocean model coupling
- Sediment-enabled regional models that incorporate sediment-derived alkalinity, dissolved inorganic carbon and trace elements into regional biogeochemistry models

- Coupled sediment feedback modeling that can represent interactions between sediments, water column chemistry, and biogeochemistry processes over longer timescales
- Evaluation of measurement methods to study CO₂ fluxes and alkalinity delivery at addition site



Figure 11: Participants collaborate to surface R&D opportunities emerging from group discussion

5.2 Catalytic Grants

To build on the momentum and collaborations forged during the workshop, participants had the opportunity to apply for small catalytic grants that can provide early insight or steering on the guidelines, test concrete hypotheses or assumptions that emerged from the workshop, or clarify analyses to inform key areas of uncertainties. Proposals needed to be focused, actionable in the near term, and designed to deliver results that clearly improve workshop outcomes. This could be achieved, for example, by analyzing existing datasets against model outputs to probe a specific workshop question on model-observation comparisons; exploring parameter sensitivities that may require additional model runs; or literature synthesis on a defined topic. Outcomes of these grants will be integrated into the model skill assessment guidelines, the perspective paper, and further disseminated by the individuals through blog posts, talks, and manuscripts.

Table 5: Overview of catalytic grants awarded and proposed study

Catalytic Grant Awards	
Recipients	Description
Tyler Rohr	Model analysis to disentangle preconditioning and feedback effects when quantifying offline OAE discounts
Adam Subhas and Charly Moras <i>With collaboration from Planetary and Anne Kiely</i>	Laboratory study to explore the role of water matrices for alkaline material dissolution and how to reliably estimate EqOH for OAE applications
Anne Kiely	Laboratory analysis studying turbulence effects on dissolution rates via tank design
Alexander Andriatis and Chih-Lun Liu	Model study examining Langmuir turbulence effects on sinking OAE particles in LES
Matthew Humphreys and Matthew C. Long	Model study examining sensitivity of OAE yield to marine carbonate chemistry uncertainties
Rui Jin	Literature study producing an evidence-based synthesis of carbonate-system biogeochemical feedbacks to OAE
Elizabeth Yankovsky	Model study to quantify uncertainty from atmospheric forcing resolution in regional OAE Models
Aurora Leeson	Model study to validate and improve a pre-existing regional model for effective OAE experimentation
Jesse Vance <i>With collaboration from Veronica Tamsitt, Brian Ward, Hendrik Grosseindemann, Lennart Gerke</i>	Model-observation analysis to examine sensitivity of CDR to parameterization of air-sea CO ₂ exchange in coupled biogeochemical hydrodynamic regional ocean models
Abby Lunstrum and Tyler Rohr	Model-observation analysis combining experimental data and modeling to quantify alkalinity losses in sediments
Moji Fakhraee and Olivier Sulpis <i>With collaboration from Devon Cole and Elizabeth Yankovsky</i>	In-person workshop developing a coordinated diagenetic modeling study to reduce sediment uncertainty in ocean alkalinity enhancement MRV

6 Next steps

Following the workshop, groups spent a few weeks finalizing their individual group synthesis chapters that documented their in-person discussions. A lead editors team composed of Alicia Karspeck, Tyler Rohr, Veronica Tamsitt and Anna Madlener is now preparing a harmonized set of model skill assessment guidelines that adequately captures each group's recommendations across stages and layers. This will be prepared for a public comment period later in the fall to offer the broader community the opportunity to engage and review the proposed guidelines. Further, a lead authorship team coordinated by Lauren Moseley will be preparing workshop outcomes, in particular identified R&D needs, for a peer-reviewed perspective paper. If you have any questions about this process and would like to learn more, please reach out to Anna Madlener (anna@carbontosea.org).

7 Acknowledgements

Carbon to Sea extends its sincere gratitude to everyone who made this workshop possible. Special thanks goes to our Steering Committee for their strategic guidance, and to all participants for their dedicated contributions and collaborative spirit.

We are also deeply appreciative of the staff at [Quartier](#) for their hospitality. The Carbon to Sea team sought a location that was central and convenient while also offering proximity to nature. This balance helped participants spend focused time together in a remote location that allowed deep dives, valuable group conversations, sufficient time to revisit ideas, and produce writing documentation. Quartier is a seminar space and hotel in Garmisch-Partenkirchen, about 2 hours south of Munich on Germany's border with Austria. The space is built around a philosophy for sustainability and offered beautiful accommodation, catering, and hands-on logistics support alongside a genuinely kind and welcoming staff.

Further thanks goes to Klaus Schanda, CMO of the Zugspitzbahn, for the invitation and organization of our visit to the Zuspitze, and Anna Widholz, Manager for Climate and Environment at the municipality of Garmisch-Partenkirchen, for the interest in our workshop and sharing a deeply-appreciated list of outdoor recommendations.

Appendix

Full Participant List

Name	Organization	Name	Organization
Abby Lunstrum	University of Pennsylvania	Jann Paul Mattern	University of California Santa Cruz
Adam Subhas	Woods Hole Oceanographic Institution	Jens Daniel Müller	Carbon to Sea Initiative
Alexander Andriatis	Massachusetts Institute of Technology	Jesse Vance	Ebb Carbon
Alicia Karspeck	[C]Worthy	Jing He	Isometric
Anna Madlener	Carbon to Sea Initiative	Lauren Moseley	Carbon to Sea Initiative
Anna-Adriana Anschütz	Leibniz Institute For Baltic Sea Research Warnemünde	Lennart Gerke	[C]Worthy
Anne Kiely	Stanford University	Matt C. Long	[C]Worthy
Astrid Hylan	CEREGE	Matthew Humphreys	NIOZ Royal Netherlands Institute for Sea Research
Aurora Leeson	University of Washington	Mohammed Hashim	University of Hawaii
Brian Ward	University of Galway	Moji Fakhraee	University of Connecticut
Charly Moras	University of Hamburg	Olivier Sulpis	CEREGE
Chih-Lun Liu	National Taiwan University	Pedro Monteiro	Stellenbosch University
Daan Reijnders	SeaO2	Robert Izett	Planetary
David Keller	Carbon to Sea Initiative	Rui Jin	University of Washington
Devon Cole	Hourglass Climate	Ruth Musgrave	Dalhousie University
Elizabeth Yankovsky	Yale University	Tyler Rohr	University of Tasmania
Hendrik Grosselindemann	University of Bern	Veronica Tamsitt	Submarine Scientific
Irene Polnyi	Carbon to Sea Initiative		

Agenda

The workshop week featured a welcome reception, group work, cross-group alignment, presentations, and drafting time. Each day began with a readout by Lead Facilitator Alicia Karspeck and concluded with a debrief between organizers, Steering Committee, and the floating participants. On Tuesday, groups discussed fit-for-purpose criteria and model-observation comparisons. Wednesday's schedule was originally set for uncertainty quantification and a Round Robin session, where each group would have met

with all other groups; however, the agenda was flexibly adapted into a full Round Robin day after groups requested dedicated time to align on scopes, identify process overlaps, and map out responsibilities. Consequently, Thursday focused entirely on uncertainty quantification, followed by the R&D brainstorming session in the afternoon. The workshop concluded on Friday morning as groups digitized their physical documentation, establishing the groundwork for the subsequent MRV Guideline and Perspective Paper writing.

	MONDAY	TUESDAY	WEDNESDAY	THURSDAY	FRIDAY	
09:00		Introduction	Morning Check-in			
10:00			Presentations	Presentations	Group Work (Documenting & Drafting)	
11:00		Group work (Fit-for-purpose)	Round Robin	Group Work (Uncertainty)		
12:00		Lunch 12:45-13:45 daily				
13:00						
14:00		Group work (Model-obs)	Round Robin	Cont'd uncertainty	Zugspitze tour	
15:00			Group Debrief			
16:00		Presentations	Halftime Q&A	R&D Exercise		
17:00	Monday Scene Setting	Free time	Free time	Free time	Free time	
18:00						
19:00	Dinner					
22:00						

Figure 12: Overview of the week's agenda blocks as it was adapted to throughout the week.

Impressions





