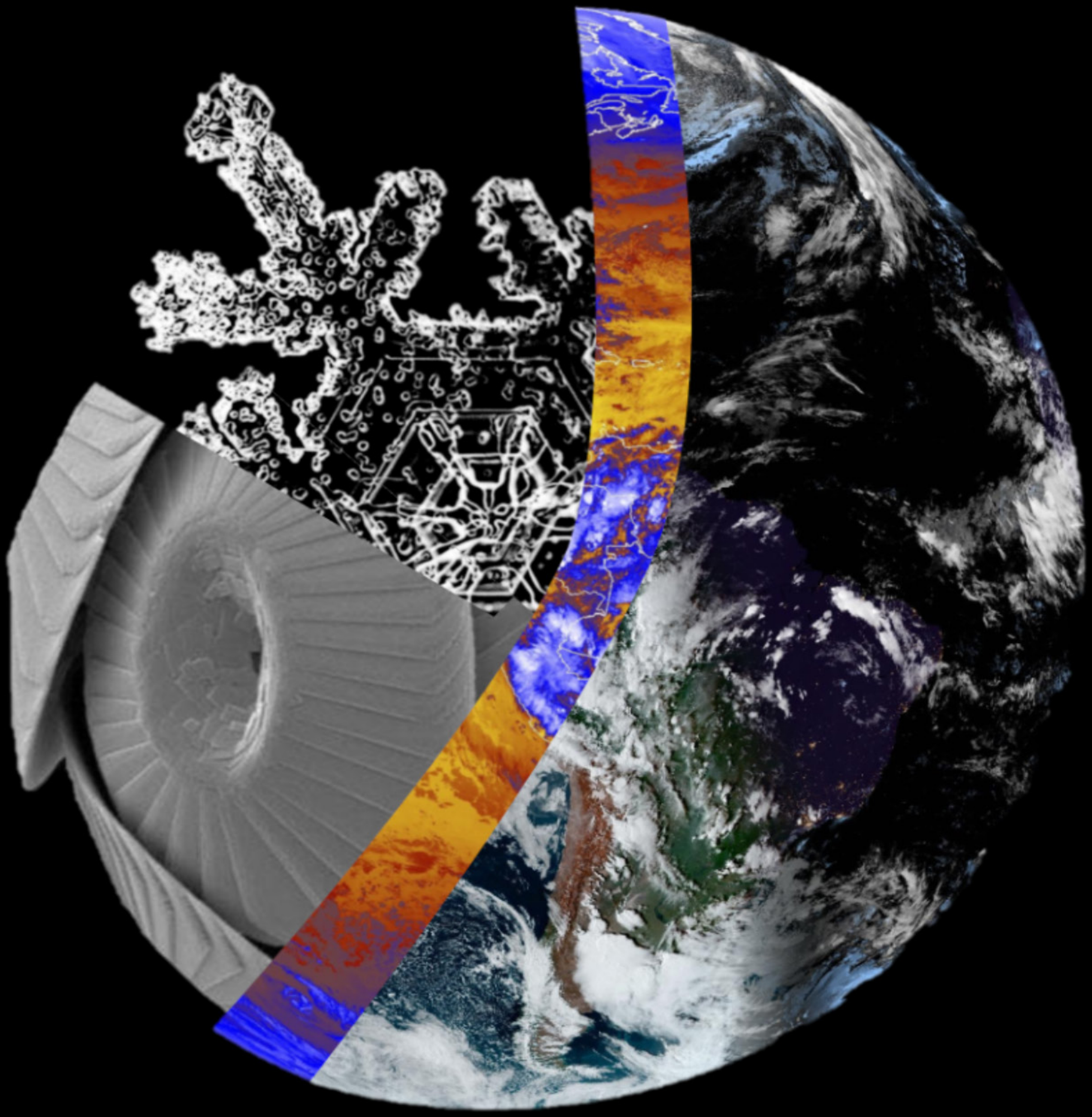

MICRO2MACRO: ORIGINS OF CLIMATE CHANGE UNCERTAINTY

A US CLIVAR Workshop

October 28 – 30, 2024

University of Wyoming | Laramie, WY, and Virtual

MICRO2MACRO: ORIGINS OF CLIMATE CHANGE UNCERTAINTY

Workshop Report

BIBLIOGRAPHIC CITATION

McCoy, D., R. Wood, S. M. Burrows, A. Fridlind, A. Igel, C. Jen, L. Regayre, M. Saito, and D. Watson-Parris, 2025: Micro2Macro: Origins of Climate Change Uncertainty. A US CLIVAR Report. 2025-6, 64 pp, https://doi.org/10.5065/q8f4-yc54.

EDITORS

Alyssa Cannistraci and Mike Patterson (US CLIVAR)

FRONT COVER IMAGE

Composite image showing a snow crystal captured by Gabor Vali (University of Wyoming) on Elk Mountain, WY; a coccolithophore captured by Jeremy Young (University College London); and a full disk image from NOAA GOES East.

BACK COVER IMAGES

Workshop images from plenary and poster sessions, courtesy of University of Wyoming Photo Service.

Table of Contents

EXECUTIVE SUMMARY

Where we are: Global Earth system models (ESMs) are essential tools for seasonal-to-decadal environmental predictions, which decision-makers across government and industry require. However, uncertainties originating at the microphysical scale and more broadly in processes occurring on scales smaller than typical ESM grid boxes, remain a major challenge. While resolution continues to improve in our predictive models, we will need to parameterize microphysical processes that contribute the bulk of prediction uncertainty for the foreseeable future. Microphysical uncertainties contribute broadly to remaining limitations on Earth system predictability on seasonal and decadal time scales.

Progress made:

- Advances have been made in high-resolution modeling, improved parameterization techniques, and growing recognition of the importance of robust uncertainty quantification.
- Fundamental scientific understanding of some key processes (e.g., warm rain formation) and identification of process-relevant observational constraints has improved.
- Simulation of the global-mean atmospheric state and key modes of atmosphere-ocean variability (e.g., ENSO) that impact weather at seasonal-to-decadal time scales has improved significantly.
- Promising applications of machine learning have been employed to objectively establish the impact of parametric uncertainty on model predictions, to reduce process uncertainties through observational constraints, and to enhance and accelerate simulation of key microphysical processes.

Remaining gaps:

- Some key processes remain fundamentally very poorly understood and/or difficult to measure or confidently attribute (e.g., primary and secondary ice production).
- Some crucial observations are not available, and necessary error characterization is frequently not available or inadequate.
- There is a lack of protocols on how to deal with structural and parametric uncertainties in how microphysical processes affect predictions at decadal scales and in regional extremes.
- Frameworks are needed that unite machine learning with improved physical understanding to allow out of sample generalization to states that the system has yet to experience, years-to-decades in the future.

What's needed: A systematic, multi-scale approach is needed to integrating observational data, process studies, and model development, alongside improved three-way communication between lab, observation, and modeling communities across scales to enhance model-data integration.

Anticipated outcomes: Improved Earth system prediction can be achieved, including improvements in regional forecasting of water availability to inform decision-makers about agriculture, energy, and infrastructure needs at seasonal-to-decadal time scales. The impact from past investments can be enhanced to improve fundamental process understanding. Improved guidance can be available for future campaign planning and infrastructure design and implementation.

1 INTRODUCTION

Earth System Models (ESMs) remain the only viable tool for long-term environmental predictions (seasonal-to-decadal), but microphysical and small-scale processes drive key uncertainties that render their predictions uncertain. Progress is being made, with ESMs now treating many more processes in physically-realistic ways than before. This increase in realism is supported by the emergence of many new observations. Tools to quantify microphysical process uncertainty in ESMs are also being more widely used. The inclusion of more processes in ESMs has led to more complexity. However, gaps remain. There is a lack of systematic consideration of the impact of parameterized processes on prediction uncertainty (within a context of increasing model complexity), as well as lack of fundamental understanding or order-of-magnitude uncertainties in representing some key processes (e.g., ice formation processes and rates). Moving forward, better integration of observational constraints to feed model improvement, coordinated evaluation frameworks to reduce process uncertainties, and a shift toward process-driven model evaluation across scales are needed.

Figure 1. Workshop participants, courtesy of University of Wyoming Photo Service.

1.1 Workshop context

Micro2Macro: Origins of Climate Change Uncertainty (M2M) brought together 154 scientists in Laramie, Wyoming, and online for a 2.5-day workshop in October of 2024 (Figure 2). Attendees included experts in Earth System Model (ESM) development; in situ observations; remote sensing; atmospheric chemistry; aerosol, cloud and precipitation physics; high-resolution cloud and aerosol modeling, laboratory experiments; uncertainty quantification; and field campaign design. Two-thirds of participants identified as students or early career scientists.

The motivation for gathering a large body with such diverse expertise was simple: our ability to offer predictions of the Earth system on seasonal-to-decadal scales hinges on processes that are small-scale and fast (Lehner et al. 2020; Watson-Parris and Smith 2022). Uncertainties in these “subgrid” processes and their representation in models (“parameterization”) limit Earth System predictability at regional and global scales (Morrison et al. 2020; Peace et al. 2020; Liu and Kollias 2023), and the process of translating advances in fundamental understanding to improvements in predictability needs to accelerate.

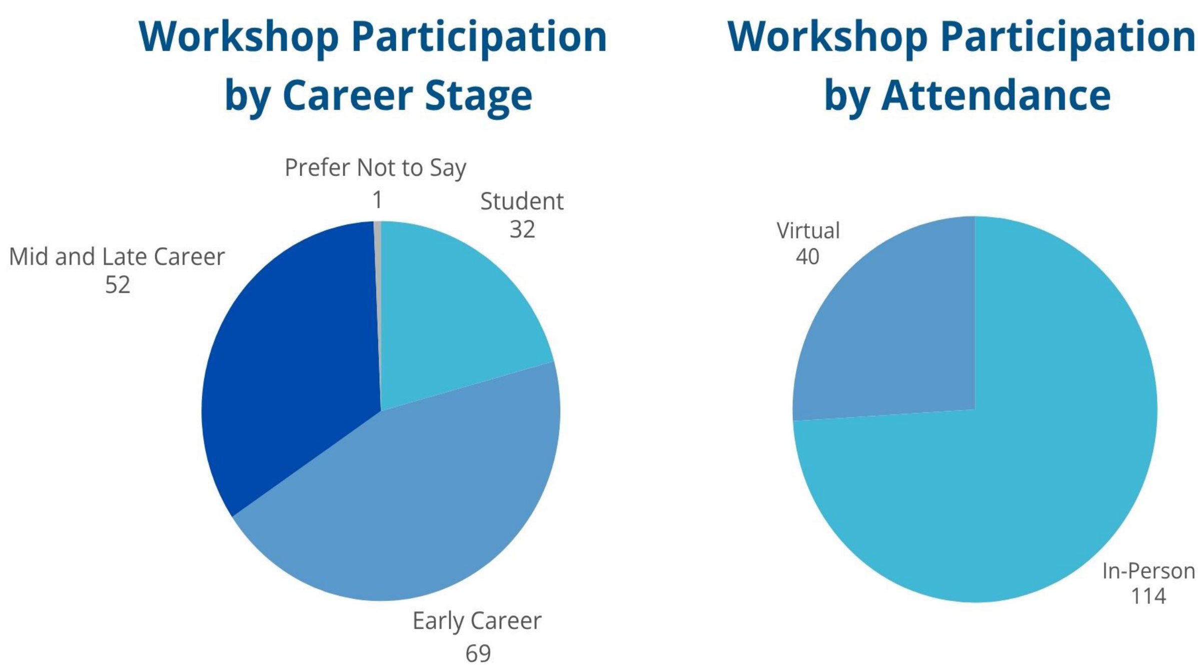

Figure 2. Summary of workshop participants. Courtesy of Alyssa Cannistraci.

It is enticing to imagine that we can resolve our way out of this problem as we advance to exascale computing. However, it is important to appreciate the massive range of scales needed to provide global predictions that explicitly resolve the parameterized processes that are assessed to dominate prediction uncertainty (Sherwood et al. 2020) (Figure 3; orange boxes).

Between the planetary wave spatial scale and the scales of individual hydrometeor interactions there are many orders of magnitude that need to be bridged. Efforts in pushing the bounds of what a global model can resolve have advanced the field (Wang et al. 2011; Terai et al. 2020), but substantial computing power is needed to achieve this, which in turn shortens the total length of prediction available from these models (Figure 3; green bars).

We provide a simple heuristic model to contextualize the huge range of scales of processes. It also illustrates that as resolution increases, the length into the future that is available for global predictions decreases. This can be illustrated with a heuristic calculation. Assuming that the number of computations scales with the number of horizontal (n_x) and vertical grid cells (n_z) as $n_x^2 n_z$, where the n_z>1 and n_x>1 for the number of grid boxes in each dimension (or conversely horizontal resolution scales as $1/n_x$). Putting this in the context of Figure 3, available integration length for global simulations scales roughly as a function of horizontal resolution following this relationship (Figure 3; dashed line). For simplicity in Figure 3 we neglect the vertical dimension

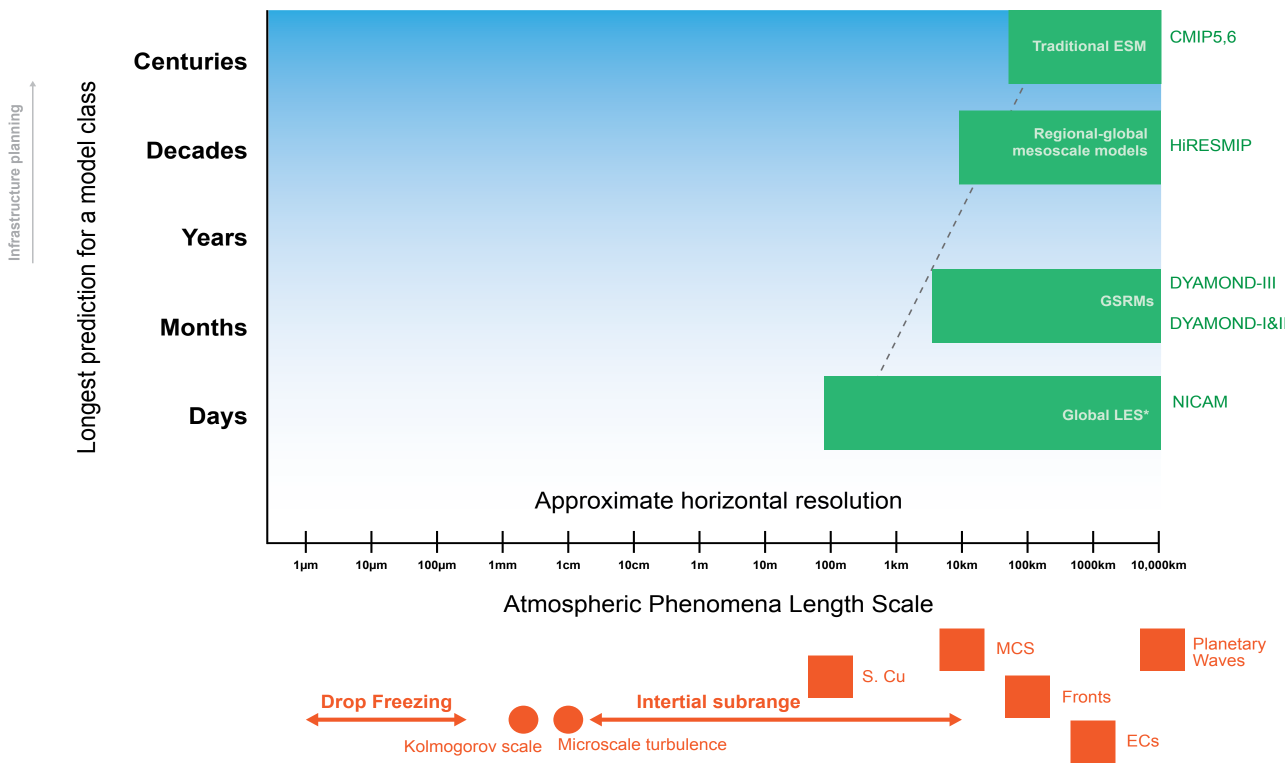

Figure 3. The tremendous scale gap between microphysical processes and global-scale prediction means that we will need to parameterize processes for the foreseeable future. The x-axis shows a representative scale of various atmospheric processes (adapted from Morrison et al. 2020, CC) (orange). The approximate horizontal resolution of different model classes at the time of writing is shown as green bars. The y-axis shows the time scale of global predictions available from each model class. Representative modeling initiatives are listed on the right. The timescale at which model physics dominates prediction uncertainty based on Lehner et al. (2020, CC) is shown as blue shading. The approximate planning time scale of infrastructure is shown on the left.

and assume a constant time step, but we stress that vertical resolution is very important to model performance (Yamaguchi et al. 2017; Lee et al. 2022). While the argument above is simplistic, an emergent property of global-scale models is that there is a tradeoff between our ability to predict for longer time scales and our ability to move to higher resolution and reduce our need for parameterization (Figure 3). Given the huge scale gap between the global scale and key microphysical processes, we are going to need to parameterize for the foreseeable future. One way to think about this is how long before we will be able to offer global integrations for a given resolution. The number of available computations on supercomputers has increased steadily over time with the number of computations on the world's fastest supercomputer increases by an order of magnitude roughly every 3–5 years. Following this scaling of available computation, moving global large eddy simulation (Δx ~100 m) from days to years might be possible in 6–10 years assuming no changes in timestep or vertical resolution. This is a long time to wait in the context of society and even for these very high-resolution classes of model, we must still parameterize microphysical processes that are not resolved by the O~100 m grid resolution.

In addition to the computation-resolution trade off described above, it is also important to note that at present the highest resolution global models do not tend to couple microscale processes such as aerosol (Haarsma et al. 2016), which artificially removes a source of prediction uncertainty. We highlight that high resolution models are a key part of how we develop better predictions of the Earth system, but we will need to abstract microphysical processes with simplified parametric representations for the foreseeable future. Unfortunately, microphysical process uncertainty does not converge like dynamics in the limit of high spatial resolution (Morrison et al. 2020; Liu and Kollias 2023). This mathematical abstraction of processes is referred to as *parameterization*.

Prediction uncertainty due to parameterized processes in ESMs is nothing new. Historically, these processes have been used to try and adjust ESMs to be in agreement with observations such as historical temperature trends and top of atmosphere radiative flux (Bender 2008; Mauritsen et al. 2012). New statistical techniques have allowed parameterized processes to be considered as a legitimate component of the uncertainty in predictions from ESMs (Lee et al. 2011; Tsushima et al. 2020; Sexton et al. 2021; Gettelman et al. 2024). This approach involves simultaneously perturbing the values of multiple uncertain model parameters within process parametrizations. Exploring parameter space in an ESM is a substantial task, requiring specialist knowledge about parametrizations, plausible parameter ranges in a given model configuration and the most efficient method for perturbing values in the model. These model-specific complexities mean efforts to explore parameter space must be tailored to each model version. At present, efforts exist across several modeling centers (Qian et al. 2018; Yoshioka et al. 2019; Regayre et al. 2023; Eidhammer et al. 2024; Elsaesser et al. 2024) but would benefit from greater coordination and inter-model analyses.

Advancements in characterizing prediction uncertainty come on the heels of our improved understanding of what processes operate in the atmosphere through a combination of space-borne, airborne, and surface observations. As our understanding has grown, developing ESMs has become more complex. This requires decisions about where to simplify process representation without compromising predictive skill (Proske et al. 2022).

1 See: Jack Dongarra, Martin Meuer, Horst Simon, Erich Strohmaier. TOP500, Performance Development

Although the past decade has provided a wealth of new process understanding and observations as well as new tools to tackle the implementation of parameterizations in ESMs, the full impact of new observations, modeling, and statistical tools on prediction uncertainty has lagged, as evidenced by the recalcitrant and even widened range of CMIP-predicted climate sensitivity in successive assessment reports (McDonnell et al. 2024).

1.2 Workshop objectives

There are numerous interesting questions related to how we create parameterizations, observe processes in the field, and treat remote sensing retrievals. We consider these questions in the context of our ability to provide accurate seasonal-to-decadal predictions that can go to informing policy makers (Hope 2015). The goal of M2M is to provide a community roadmap in the context of our process-based understanding of microphysical processes and seasonal-to-decadal global predictions. This overarching goal informs how we discuss problems related to ESM development, observations, and their intersection. In service of this, the Scientific Organizing Committee (SOC) split the workshop into sessions focused on a series of sub-questions.

Before embarking on developing a series of questions related to how we approach ESM prediction uncertainty, the workshop summarized the current state of the field and attempted to identify challenges that limit the predictive ability of ESMs at the seasonal-to-decadal scale. This summary focused on ESM performance with the goal of addressing ESM decadal prediction uncertainty as opposed to fidelity in reproducing current climate state. The SOC highlighted a focus on reducing prediction uncertainty as opposed to increasing fidelity in representing the present-day climate, because present-day fidelity is a profoundly multifaceted objective whose relevance can only be established by robustly assessing prediction uncertainty.

While an ability to match present-day observations is necessary, it is insufficient to produce a reliable ESM, and it is also extremely challenging to observationally-constrain the chemical and physical processes occurring in the atmosphere. Observation challenges affect spaceborne and surface remote sensing of microphysical properties; in situ measurements at the surface and from airborne platforms; and in the lab where microphysical quantities of interest may need to be inferred from observables. A key finding emerging from this discussion and summary of how observations are being used to improve models is a pressing need for a clear communication channel between groups making/developing observations and groups developing/evaluating models, and the recognition that both sets of problems are subject to spiraling complexity. We need to develop collaborative pathways forward that offer actionable information about structural improvements that will reduce model prediction uncertainty.

Emergent from discussing how we make observations of microphysical properties and how we can use these observations to reduce prediction uncertainty from ESMs is the need for a unified framework, described in Section 1.3, where the utility of a given observation to improve ESM prediction can be robustly evaluated. This point was discussed in the context of existing tools being used to interrogate the impact of parameterized processes on ESM behavior and how to better integrate them with observations to reduce ESM prediction uncertainty. A benefit of this approach is that it also provides guidance to plan future measurements (Section 5). Such an approach is more mature in the realm of planning satellite missions but is relatively unexplored in terms of planning field deployments and campaigns.

1.3 Bayesian Inference Frameworks for the Earth System (BIFES)

A key tool discussed across M2M sessions is the perturbed process/parameter/physics ensemble (PPE; Carslaw et al. 2025). While PPEs are a relatively new idea in the field (Rougier et al. 2009; Lee et al. 2011), the basic concept is simple and intuitive from a physical sciences and statistics perspective. PPEs sample uncertainties across a wide range of model processes simultaneously, making them well-suited to statistical methods that effectively scale ensembles from hundreds to millions of samples (Lee et al. 2011; Watson-Parris et al. 2021). This scale of model output enables powerful analyses, such as identifying the causes of model uncertainty and applying observational constraints to assess the plausibility of different model[2] variants. We provide a brief, concrete description of how one might use this framework for a simple problem within the field and refer back to this description throughout the remainder of this report.

It is important to realize that the challenges in predicting the Earth system are not unique in the sciences. One of the lingering issues in science is the problem of induction (Hume 1751). In essence, we observe snapshots of states, not processes. Physical science builds models that sit underneath these observations to allow us to predict future events. A simple example is observing the force, acceleration, and mass of a block on a plane and having a physical model of friction that explains our observations. We do not observe friction directly.

The problem of induction was directly addressed during M2M. An example from Graham Feingold's talk during the workshop used an analogy of watching snapshots of a soccer game and inferring the rules of the game from those snapshots (Feingold et al. 2025). If we can build up a model that sits beneath our observations (perhaps the rules of soccer) we can make predictions of the future (whether or not one of the players will be called for being offside). The problem of induction dates back in some form or another to the second century AD and the solution taken by most physical science fields has been to establish a level of causation that we can infer from observations that is 'comfortable' for that field. In many fields that are built on bench-scale science the causation criteria do not need to be very stringent. If one adds acid to a basic solution, the causation as the solution titrates is clear. Bench-scale experiments also enable repeating the same experiments multiple times to establish repeatable causality. Earth science presents a uniquely challenging environment with numerous uncertain, non-linear processes interacting in a chaotic system and a limited ability to repeat experiments. A concept that is frequently discussed is equifinality—the idea that many process combinations can result in the same value of an observable (Beven and Freer 2001).

Discussing inferring soccer rules from snapshots is somewhat abstract, so we provide a concrete example of how this affects the goals of M2M. Imagine that you are attempting to develop a new

2 Here, and below, we specifically refer to models by their utilization (i.e. ESM, large eddy simulation, single-column model) rather than simply 'model.' Here, we specifically keep this general since this applies to any mathematical abstraction of a physical system.

physical model of coalescence scavenging (loss rate of cloud droplets to collision-coalescence) in liquid clouds (Wood 2006) (Figure 3). Obviously, the model is approximate since it is, after all, a model. While M2M focuses on understanding what limits ESM prediction skill (Section 2), it is important to keep in mind that every process parameterization is a form of abstraction, for example, a simplified representation of the true underlying, and typically complex process (Bony et al. 2013; Mülmenstädt and Feingold 2018). Take for example friction between a block and a plane as discussed above, most models of friction do not attempt to represent the meshing of every tiny irregularity between the block and the plane and abstract these as static and dynamic friction coefficients, representing the aggregated macroscale force. In the case of abstracting coalescence, one might take previous observations of state (e.g., cloud droplet number Nd, condensate amount) and use these to develop a physical model that you can apply to make predictions. For instance, one might take multiple observations of state (Nd) and use them to infer how droplets coalesce to affect droplet number and size distributions, which could then be used in the ESM to make future predictions. In Figure 4 previous observations are shown as squares, and a first attempt at a physical model is shown as a gray line. The figure illustrates this in the context of closure in droplet number, where closure is achieved if the droplet number from the physical model and droplet number as observed by probes agree.

Let's imagine that we are envisioning a new set of measurements in a regime of droplet number that hasn't been tackled before. You can imagine three potential outcomes:

(a) Hypothetical observation #1: you observe a droplet number far different than the predictions of your physical model (i.e. your null hypothesis);
(b) Hypothetical observation #2: you observe a droplet number closer, but still not in agreement with your physical model (doesn't intersect the dark gray line); and
(c) Hypothetical observation #3: you observe a droplet number in agreement with your physical model (intersects the dark gray line).

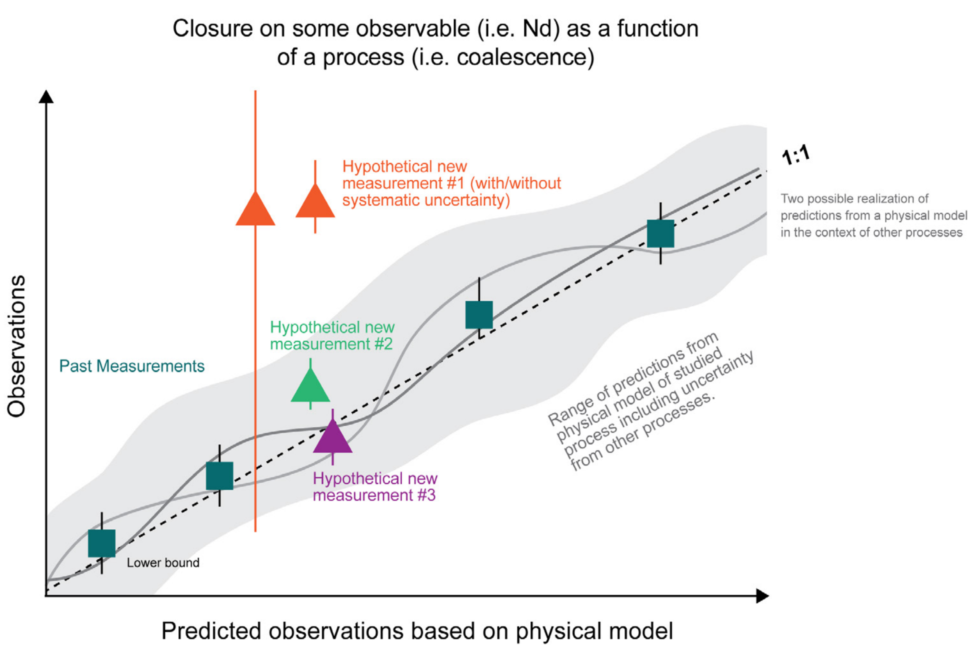

Figure 4. A hypothetical closure between observations of droplet number and a physical model of coalescence. Squares illustrate past measurements used to develop the model of coalescence. A researcher proposes to measure in a new regime where we discuss the consequences for their model of coalescence of new observations that look like the orange, green, or purple triangles. The gray lines illustrate the predictions of the process model ignoring uncertainty from all other processes for two possible configurations of the process model. The shading illustrates when uncertainty in other processes is considered. Uncertainty in observations is shown with vertical lines. Credit: Jennifer Power Art & Design.

Only in the last case can you confirm your proposed physical model. Hypothetical cases #1 and #2 result in rejecting your physical model of the process.

Here, we argue that PPEs coupled to observing system simulation experiments (OSSEs) (Arnold and Dey 1986) and instrument simulators (Swales et al. 2018; Silber et al. 2022) provide a robust alternative to conventional ESM development strategies. We argue that cases (b) and (c) can both be consistent with our proposed model and only case (a) is inconsistent with it. This is because observables, especially in the Earth system, are rarely affected by a single process, and other processes can also carry non-zero uncertainty. In our example, many different processes affect droplet number, and those processes are not derived with zero uncertainty from first principles. For instance, cloud droplet number is also affected by nucleation and growth of aerosol particles (Wood et al. 2012), which are not zero-uncertainty processes. PPEs provide a framework to explore uncertainty in multiple simultaneous and non-linear processes. In our example, if we were to build on our simple process model of coalescence scavenging and also consider uncertainty in other processes that are at work, the uncertainty around our model might look more like the grey shaded region in Figure 4 as opposed to only considering uncertainty from a single process. If we made observations that look like cases (b) or (c), we would not be able to reject our physical model (the null hypothesis) of coalescence scavenging (this is discussed in the context of Type 1 error in Figure 4). The discussion is presented in the context of testing a hypothesized new process representation relative to the current process representation (the null hypothesis or H_0). A set of three possible measurements is given as shown schematically in Figure 4 and the consequence of following those measurements through testing the null hypothesis is shown in Figure 5. If we neglect the uncertainty from processes beyond the process we are evaluating with our hypothesis, there is always a higher chance of incorrectly rejecting the null hypothesis.

Stepping back from this example, this harkens back to a basic idea in statistics: Bayesian versus frequentist statistics. In the frequentist context we can only answer whether observations are consistent with a hypothesis. In the Bayesian context we can ask, is our hypothesis consistent with the observations? This is a subtle, but key, difference. The PPE framework allows us to ask this question in the Earth system by asking whether our theories of processes in the Earth system can be reconciled with observables given that multiple uncertain processes are at work.

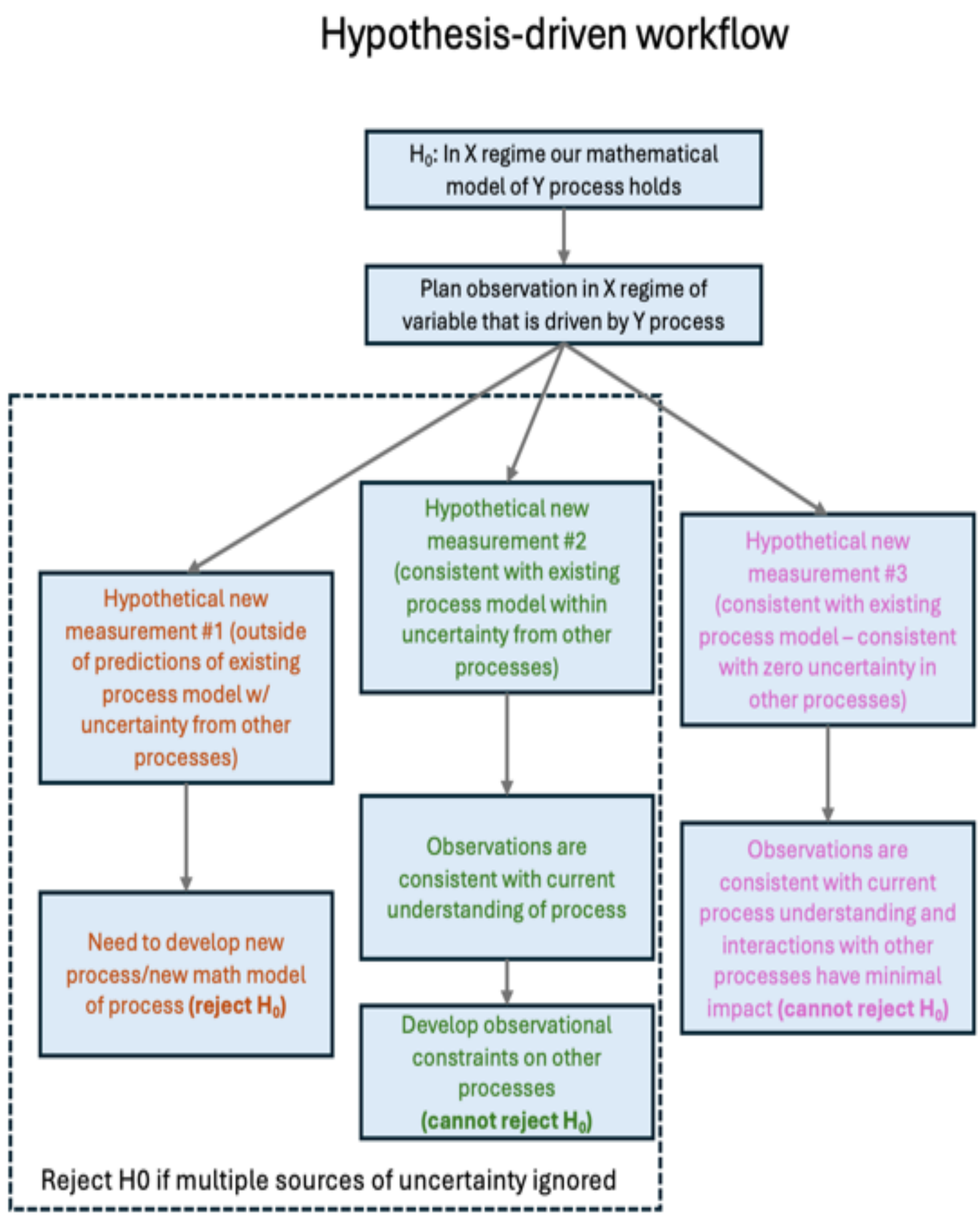

Figure 5. A step-by-step analysis of how different hypothetical observations in Figure are reflected in rejecting the null hypothesis (i.e. underlying physical model). While considering uncertainty in all processes that the observables we are considering results in only rejecting the hypothesis in the case of measurement #1, failing to consider this results in a type 1 error (incorrectly rejects the null hypothesis). The different measurements are shown schematically in Figure 4. Courtesy of M2M SOC.

Another key point is that to reject or accept physical models, we need to have a quantification of observational uncertainty, in addition to the uncertainty stemming from how we represent processes (Section 2). We illustrate this below as vertical lines on the points in Figure 4. If a clear characterization of uncertainty is not available for an observation, then we assume it is effectively infinite and accept all possible models (the long vertical bar on one of the examples for case 1). While this example seems so obvious, systematic error characterization is often not available (Section 3). This was a key theme of M2M that was returned to many times.

Overall, PPEs give us a tool that allows us to pursue hypothesis-driven science in the Earth system where observations and processes are both subject to uncertainty. A discussion of how to reconcile physical models with observations in this framework is provided in Johnson et al. (2020) and in the supplementary to Regayre et al. (2020). A recurrent theme during M2M was the OSSE concept that has been employed effectively in the design of spaceborne remote sensing for weather applications (Cucurull et al. 2024), and to a lesser degree for Earth system predictability applications (see Section 5). We refer to this framework as Bayesian Inference Frameworks for the Earth System (BIFES) throughout the rest of this report. We present BIFES as a unified way to deal with observational and multi process uncertainty. We note that BIFES is heavily inspired by OSSEs. Here, we consider OSSEs in the context of the connector between models and observations (Arnold and Dey 1986; Bodas-Salcedo et al. 2011; Silber et al. 2022; Cucurull et al. 2024). PPEs are a model exploration framework and machine learning (ML) emulation system for perturbing model representation in a systematic way (Lee et al. 2011). The BIFES testbed unifies both concepts to allow a robust, ML-enabled, digital testbed for enhancing Earth system prediction.

1.4 Report structure

Based on discussion and presentations from M2M, we provide a set of actionable suggestions with the goal of improving ESM predictability through consideration of subgrid-scale (parameterized) processes. These suggestions revolve around how we can synergize recent tools developed in the context of ESMs with advances in our ability to observe the Earth system. We outline the workshop goals as:

- Summarize the state of the art in terms of the impact of microphysics on ESM prediction skill at the seasonal-to-decadal scale. [Section 2]
- Characterize how we observe microscale processes and create clear recommendations for interactions between the observational community and the model development community. [Section 3]
- Develop recommendations for observational methods/strategies to improve ESM interannual to decadal prediction skill. [Section 4]
- Identify areas of synergy between observational planning/interpretation and ESMs. [Section 5]
- Advance statistically robust end-to-end methodologies for confronting the multiscale problem of predicting the Earth system (e.g., BIFES). [Section 6]

Each of the following chapters starts out with a summary of current knowledge, recent advances, remaining gaps, future needs, and key findings.

2 WHAT'S WRONG WITH MICROPHYSICS IN EARTH SYSTEM MODELS?

Where we are: Microphysics remains a major source of uncertainty in ESMs, with parameterized processes often treated as tuning knobs for macro-scale features rather than legitimate sources of uncertainty.

Progress made: Advances in microphysical parameterization techniques are being made, and there is a growing recognition of the need for uncertainty characterization. The sophistication of microphysical representation in ESMs has improved in the last decades through work by modeling centers and better leveraging observations.

Remaining gaps: Most models lack robust parametric uncertainty characterization, and structural error evaluation is still in its early stages. Some processes remain fundamentally poorly understood, poorly measured, or both.

What's needed: A shift from simply prioritizing model fidelity to uncertainty-based evaluation, can be achieved through improved collaboration between scientists developing models and those making observations.

Key Findings:
- The characterization of parametric uncertainty is limited in most ESMs.
- There are no established methods for evaluating structural error in ESMs, and few for structural error in smaller-scale models.
- ESM development has historically focused on increasing fidelity rather than quantifying and reducing projection uncertainty.
- Better two-way communication is needed between scientists developing ESMs and making observations.
- Microphysical process parameterization has progressed, but further improvements are needed.

It is important to recognize recent major advances made in implementing parameterizations that represent microphysical processes. In recent generations of ESMs, microphysical processes have risen to be the dominant sources of prediction uncertainty. This is because microphysical processes are at the root of setting cloud feedback (Blossey et al. 2013; Myers and Norris 2013; Tsushima et al. 2020; Zelinka et al. 2020), aerosol radiative forcing (Regayre et al. 2018; Watson-Parris and Smith 2022), as well as many other ESM processes (Liu and Kollias 2023). This prediction uncertainty is a product of the rapid progress made in recent decades in unifying our ability to observe microphysics, infer processes, and parametrize them into ESMs so that we better understand how they impact our predictions of the future. In essence, the uncertainty we see stemming from model physics (Lehner et al. 2020) is a sign that we are starting to come to terms with representing these processes.

Bridging the scale gap between small, fast processes and their effects at the global scale is hard, but we know that these processes matter and neglecting them provides unrealistic predictions to stakeholders. For instance, one could envision a low complexity model where clouds are set to their climatological mean state with static coverage and radiative effect. This would be a low uncertainty model, but output would completely disagree with what we observe happening in the Earth system (Norris et al. 2016; McCoy et al. 2018; Bellouin et al. 2019; Gryspeerdt et al. 2019) because climate feedbacks that make the planet more sensitive to radiative forcing (Soden and Held 2006; Sherwood et al. 2020) would be neglected and forcing trends from microscale processes (Bennartz et al. 2011; McCoy et al. 2018; Myers et al. 2021; Wall et al. 2022; Grosvenor and Carslaw 2023; McCoy et al. 2023) would be removed. It is tempting to embrace a maximalist approach where every process is implemented in ESMs, but increased model fidelity does not ensure improved model predictive skill at the critical seasonal-to-decadal scale. Thus, we need to strike a balance between complexity and interpretability (Bony et al. 2013; Mülmenstädt and Feingold 2018). ESMs cannot mimic every process in the Earth system and remain tractable (or computationally viable). However, not all processes implemented in an ESM have a major influence on future environmental state. This is a key message identified in M2M: we need to consider process representation in ESMs in terms of their payoff in accurately representing projections as opposed to the mean state.

Sources of microphysical uncertainties can be broadly categorized into two groups: sources associated with parametric uncertainty and sources associated with structural error (Sexton et al. 2012; Bony et al. 2013; Regayre et al. 2023). Parametric uncertainty arises when a process is described by the mathematical structure of the parameterizations, but those equations contain empirical or semi-empirical parameters whose values are not sufficiently constrained. Structural error arises when a process rate is not adequately described by the parameterization equations or when a process is entirely missing. It may be associated with missing process dependencies (Lee et al. 2011) and/or inadequate simplifying assumptions.

There was general consensus at M2M that perturbed parameter/process/physics ensembles (PPEs) are promising tools for (a) identifying the most uncertain parameters and, by extension, the most uncertain processes (microphysical or otherwise); (b) constraining uncertain parameters; and (c) identifying sources of structural error (see Section 1.3).

There are many uses of PPEs and their utility will be discussed throughout this report. The use of PPEs for characterizing parametric uncertainty and identifying structural error is currently limited to a few ESMs (Qian et al. 2018; Regayre et al. 2023; Eidhammer et al. 2024; Elsaesser et al. 2024; Nugent 2025). A more widespread adoption of PPEs in ESM development would be beneficial. This is because it provides a more complete assessment of ESM uncertainty and because it can make the process of ESM development much more efficient and robust (Sexton et al. 2021; Yarger et al. 2024; Elsaesser et al. 2025). By systematically sampling uncertainty space in an ESM, we can link processes to prediction uncertainty as opposed to simply linking them one at a time to fidelity. This has the potential to accelerate ESM development because rather than playing whack-a-mole with processes in sequential versions of ESMs, we can survey and identify optimal parameterizations (Elsaesser et al. 2025). We argue in Section 4 that this also allows us to identify observations that provide maximum leverage towards reducing prediction uncertainty and allow us to explore interactions between processes (Song et al. 2024).

Participants in M2M recognized that structural error is present in ESMs in the same way that all models are structurally incomplete. This source of error can sometimes be masked, but not fixed, with parameter tuning. Methods to address structural error in ESMs are still in their infancy (Regayre et al. 2023). Without frameworks to systematically explore parametric uncertainty, it is difficult to operationalize rooting out structural uncertainty. Additional methodology development to account for systematic uncertainty would benefit from a top-down approach to identifying the factors that drive ESM prediction uncertainty.

There are several ongoing efforts to quantify and/or characterize structural error for individual processes. Examples of such efforts include the characterization of internal vs external mixing assumptions for aerosol particles (Curtis et al. 2017), characterization of errors associated with separate cloud and rain categories (compared to a single category) (Igel et al. 2022; Hu and Igel 2024), and consistent treatment of ice effective radius across microphysics and radiation (Fierce et al. 2024). These efforts often start in very simplified modeling frameworks (e.g., a box or parcel) and are slow to make their way to ESMs. Informed ESM development using this bottom-up approach could be better streamlined within the community.

ESMs are, at their core, a predictive model based on underlying physical models of processes. They can predict both the current state of the planet as well as predicting into the future. It was discussed that evaluation of ESM behavior has been historically focused on fidelity as opposed to prediction uncertainty. This reflects the common aphorism "it's difficult to make predictions, especially about the future."[3] Predicting a state we can observe is a well-defined problem, and makes it difficult to effectively prioritize processes to provide more robust predictions of the future using ESMs. PPEs were discussed as one way to systematically evaluate which processes limit ESM predictive ability.

Workshop participants struggled to agree on exactly which microphysical processes contribute most to the uncertainty in specific projections (e.g., temperature, precipitation, cloud radiative effect) from ESMs due to a lack of comparable evidence. Based on discussions, process studies tend to be conducted one at a time and it is difficult for researchers studying, for instance, secondary ice production to be able to offer a definitive comparison in terms of contribution to prediction uncertainty with, for instance, aerosol mixing state. Overall, this lack of agreement highlighted the need for more systematic studies of the parametric and structural uncertainties in current ESMs to allow efforts to be focused on major limitations. However, there were some consistent calls for specific observations (see Section 3).

---

3 Attributed to Neils Bohr (1885-1962).

Finally, workshop participants also emphasized the need for better communication among modelers, observationalists, and laboratory experimentalists. Such communication is essential to translate fundamental process work into model parameterizations as well as for robustly evaluating model updates. Observationalists need to effectively communicate uncertainties in their datasets (see Section 3). However, communication also needs to go the other way. Modelers need to inform laboratory experimentalists and observationalists about their greatest needs. One pathway identified to streamline the connection between observations, experiments, and ESM advancement is to more regularly implement simulation studies before experimental design or field deployment (Section 5) and leverage the proposed BIFES framework (Section 1.3).

3 CAN WE EVEN OBSERVE MICROPHYSICS?

Where we are: We have a wide array of microphysical observations. However, existing observational data sets are often not well-suited for ESM development and evaluation, known instrument deficits are not being sufficiently addressed by innovation and deployment, and laboratory work remains undervalued.

Progress made: Recent decades have greatly advanced microphysical remote sensing and in situ observational techniques, and there is increased awareness of chronic instrument shortfalls and observational uncertainty.

Remaining gaps: There remains a lack of systematic and representation uncertainty, insufficient availability of community open-source simulators, and deemphasized importance of instrument development and deployment and laboratory research. There is a lack of frameworks for analyzing causality and process rates in observational datasets.

What's needed: Greater emphasis on laboratory experiments and instrument development, and better integration of observations with modeling would enable progress.

Key Findings:

- Existing remote sensing and in situ products are often inadequate for robust model evaluation. Systematic and representation uncertainty quantification is frequently missing or inadequate. There is a need for improved, community-driven open-source simulators.
- Systematic and representation uncertainty quantification is frequently missing or inadequate.
- There is a need for improved, community-driven open-source simulators.
- Controlled laboratory and perturbation experiments are critical for improving process understanding and discovering new processes.
- Targeted closure studies are needed to constrain microphysical properties that are major sources of ESM seasonal-to-decadal prediction uncertainty.
- Observing systems and campaigns should be designed and dedicated to constraining specific processes, with instrumentation adequate for purpose.
- Sustained post-observation analysis is required to accomplish process constraint studies at the community level.

Observations of microphysics include in situ observations, laboratory measurements, and remote sensing, and the synergies between them are being increasingly exploited to develop, confront, and constrain ESMs. “Microphysics” in the context of atmospheric measurements refers to both microphysical properties and microphysical processes. Microphysical properties are the characteristics of atmospheric particles or hydrometeors that result from relevant microphysical processes and can be measured using various techniques. Microphysical processes are time-varying phenomena that causally change microphysical properties. It is critical to acknowledge that microphysical observations provide information about the state of a system, and do not directly provide process information. The links between observations and processes are discussed in other sections of this report.

The last three decades have seen a rapid development of observations of small, fast processes in the atmosphere. In the US, the Atmospheric Radiation Measurement (ARM) program set up its first permanent site in 1992 with a core mission related to clouds and aerosol (Ackerman and Stokes 2003). In 1999 NASA launched Terra, which provided new avenues to examine the global distributions of aerosol and cloud properties (Nakajima et al. 2001; King et al. 2003). This was followed by the launches of spaceborne lidar and radar in 2006 and a satellite constellation (the A-Train) that allowed examination of small, fast processes at a new level (Stephens et al. 2002). Of course, with any set of observations, there is a time lag between first light and when a robust, vetted product is available for science applications, and a sufficiently long data record is available (McComiskey et al. 2009; Gryspeerdt et al. 2022a; Zhang et al. 2023). As motivated in Section 1, observations are the tether to reality for ESMs (Regayre et al. 2023; Elsaesser et al. 2024). They are what allow us to develop ESMs and implement physical processes that are likely to affect system predictability (Twomey 1977; Boucher and Lohmann 1995; Gettelman and Morrison 2015; Khairoutdinov and Kogan 2000; Gettelman et al. 2025), and they are how we evaluate the performance of our ESMs by challenging our models to produce output that matches observations of system state (Gettelman et al. 2020; Regayre et al. 2020; Mikkelsen et al. 2024; Elsaesser et al. 2025).

It is important to recognize that observing microphysical properties is challenging. From space, remote sensing must probe through the depth of the atmosphere and interpret radiometric signals to infer properties that are at small scales relative to the instrument footprint. These scale discrepancies introduce uncertainties in satellite-based quantification of climate-relevant properties, such as the covariance of Nd and aerosol concentration (Quaas et al. 2020). From the surface, sampling is limited due to constraints on time and resources (in that we cannot place surface observations every few km). Surface observations are in closer proximity to the system being studied, but they are also embedded in that system so may be affected by the environment being observed. For instance, liquid water path estimates from microwave radiometers can be biased due to a rain water layer developed on the radome (Kurri and Huuskonen 2008), and delicate aerosol measurements can be contaminated in the vicinity of broken clouds (Kaufman et al. 2005). Even within the laboratory, it is difficult to realistically isolate specific real-world processes in a controlled environment and relate these observations to key atmospheric processes. Despite huge strides forward in measurement and experimental techniques, much work remains to reconcile observations from different instruments, experimental setups, and sampling locations. This presents opportunities in refining remote sensing, instrument development, and laboratory studies to better quantify cloud and aerosol microphysical proper ties. The scope of M2M was to identify where developing observational techniques can be most impactful to support ESM predictability.

A key need identified during M2M and feeding into BIFES (Section 1.3) is frameworks that can fill the gap between observations (our measure of reality) and ESMs (our tool to understand if observations have value for prediction). This framework needs to be able to bridge observational/retrieval uncertainty and representational uncertainty in ESMs. Some efforts are currently underway, supported by NSF[4] and DOE.[5] Related to this key need, observational instrument simulators will enable us to bridge between instruments and models. An example is a satellite data simulator unit (Masunaga et al. 2010), which is useful for ESM evaluations regarding model assumptions, remote sensing algorithm improvements, and systematic retrieval error quantifications. Community simulators (Bodas-Salcedo et al. 2011; Silber et al. 2022) could be given more focus.

3.1 Systematic and representation error

While we have made tremendous strides as a field in our ability to provide observations that constrain microphysical processes, M2M identified that one of the key observational shortfalls is our inability to accurately describe systematic error[6] (Figure 6). Systematic error in observations can be corrected if the main cause behind the systematic error for the instrument is known. However, this is typically difficult to quantify as no standard, or true value, exists to compare for most atmospheric instruments. For example, observations of number concentration of ice nucleating particles (INP) present in a given air sample varies by one to two orders of magnitude depending on the instrument type and the composition of the ice nucleation particle (Knopf et al. 2021; DeMott et al. 2018). No true standard exists for INP that would allow researchers to quantify the systematic uncertainty and separate out this error from random error, through some complex compounds, such as lignin or Snomax, have been prop-

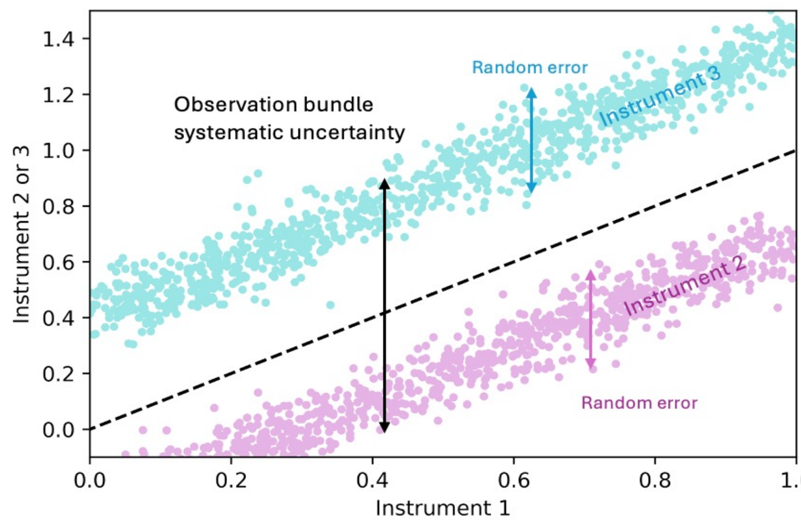

Figure 6. An idealized representation of systematic and random error estimated for a single observable. The observation bundle concept is illustrated for a bundle created by contrasting 3 different measurement techniques. Courtesy of M2M SOC.

4 INtegrating Field Observations and Research Models (INFORM) https://sites.google.com/ucar.edu/mccluskey-research/research#h.p_rufOjxB1oWNe

5 PPE Regression Optimization Center for ESM Evaluation and Development https://github.com/PROCEED-ESM

6 For instance, bias in the measurements due to how an instrument operates which would shift all observations away from the true value.

osed as a calibration standard with caveats on potential degradation over time (Polen et al. 2016; Miller et al. 2021). In contrast, random error characterization is well-posed and can be readily evaluated by simply placing two calibrated detectors, if calibration standards exist, next to each other. However, this is infrequent for many observational studies due to resource limitations in deploying two identical instruments for each measurement parameter. It can also be a major challenge for intercomparison of cloud probes, which must typically be compared aloft (from airborne platforms) rather than at the surface. Limited information on systematic and random errors is a major issue for physical process understanding and predictability because it means that the error bars in Figure 4 should be assumed to be infinite. In turn, this means that all potential models are consistent with observations. Clearly, this is an extreme case and most scientists making observations will argue that there are some limits on the plausible range for a given measurement. Unfortunately, studies that evaluate systematic errors in estimated geophysical variables from measurements are quite limited (Iwabuchi et al. 2014; Saito et al. 2020). If these are not provided as part of a data set, the assumption has to be that there is no real characterization of systematic error.

This paints what is perhaps an excessively dark picture of how observational estimates and their uncertainties are made and it should be emphasized that this is in the context of the hand off between those making the observations and those leveraging the observations for evaluating ESMs (Schutgens et al. 2016; Regayre et al. 2020). In discussion during M2M, observational groups pointed out that systematic uncertainty was potentially available for many different observations. However, this is often in the context of the individual retrieval level for observations where scientists are able to characterize how different assumptions propagate to uncertainty. These uncertainties are typically not readily available in the scaled-up observations needed for comparison to ESMs. One recommendation from M2M was for data sources to tag data with a statement about potential systematic error. This could be a statement with an estimate of error or a statement that it is not available at the scale of the data set . Propagating uncertainty through retrievals to physical quantities is an important step to characterizing systematic uncertainty but does not provide a truly independent measure of systematic uncertainty.

A suggestion made during Ken Carslaw's talk (Carslaw 2025) is that our urgent need for robust future predictions means that we should avoid recreating observational data sets from the ground up and instead should try to leverage what we have and require more concrete descriptions of uncertainty for existing products. Collection and formatting of observations into 'ESM-ready' formats with characterized systematic error, would be ideal. A suggestion made during Gregory Elsaesser's talk (Elsaesser et al. 2025) was to use 'observation/retrieval bundles' to evaluate systematic error (Figure 5).

Observation bundles take several different methods of retrieving a given geophysical variable and interpret the spread between methods as a characterization of systematic error (Elsaesser et al. 2025). This has two substantial benefits: first, it is something that can be done immediately by groups evaluating ESMs to provide a characterization of systematic error to inform ESM constraints; and second, it is a more robust estimation of the systematic error than is provided by just propagating assumptions in a single retrieval/measurement technique. We note that programs like ARM provide an effective pathway forward for this due to the large number of collocated measurements being taken, allowing for the building of retrieval bundles that are being stochastically sampled. One example of this is the recent intercomparison of different retrievals of cloud droplet number concentrations and comparison to aircraft data over the Açores (Zhang et al. 2023); another is the intercomparison of multiple spaceborne retrievals with airborne measurements (Gryspeerdt et al. 2022b). Finally, comparison of estimates of a geophysical variable (e.g., LWP) from two independent measurement techniques (e.g., passive microwave vs vis/NIR retrievals) may also serve as an approach to estimating systematic uncertainty.

One issue raised several times during M2M was the scale mismatch between observations and ESM grid box scales. Frameworks are needed to scale observations with very different footprints up to the scale for direct evaluation against an ESM's much larger grid size (Field and Furtado 2016; Schutgens et al. 2016). This could be addressed if systematic observations were collected (i.e., measure the same parameter over multiple points in a model grid box, or the same parameter over a long period) to quantify this uncertainty (Schutgens et al. 2017). There are several examples of this being leveraged in the design of observation collection (Wood et al. 2011; Albrecht et al. 2019; Redemann et al. 2021), which has proven beneficial to the usability of that data for ESM improvement (Section 6).

Finally, the idea of adapting the PPE framework to methodically explore the causes and impacts of systematic uncertainty was discussed. There are commonalities between how PPEs are applied to understanding ESMs (complex, non-linear models) and how these might be able to help understand systematic error in observations (complex, non-linear retrievals). Priors for different assumptions going into retrievals could be systematically sampled and propagated through the retrieval to create an uncertainty space and best estimate for each observation.

3.2 What microphysical measurements are we missing?

As motivated above, recent decades have yielded major strides in our ability to observe microphysics. However, there remain key frontiers where observations of microphysics have the potential to improve process understanding, improve ESM robustness, and reduce projection uncertainty. It is critical to note, while there is always the potential for more observations to be made, the community needs to coordinate across those developing ESMs, process models, and observations to identify which observations provide leverage on predictability at the annual-decadal scale (and which do not) to be able to prioritize these efforts.

The discussion during the workshop identified numerous new microphysical measurements that would reduce prediction uncertainty in ESMs. M2M determined that vertically-resolved liquid condensate from remote sensing measurements (Leinonen et al. 2016; Saito et al. 2019) is critical for modeling and assessment of cloud processes (e.g., Beall et al. 2024). In addition, inconsistent remote sensing LWP estimates from vis/NIR and passive microwave techniques result in uncertainty in cloud properties (Greenwald et al. 1997). Attempts should be made to reconcile these measurements with closure measurement campaigns.

The workshop also identified the paucity of ice cloud measurements as a major knowledge gap, with a particular focus on knowledge gaps surrounding secondary ice production. Secondary ice production refers to a collection of at least eight currently identified mechanisms that lead to the formation of new ice crystals from existing ice, through processes such as fragmentation and splintering of ice. These processes can occur during the freezing process, or upon collision of ice with another hydrometeor, and lead to increases in ice crystal number. Ice crystals formed via secondary ice production can grow through riming and vapor deposition and have the potential to participate in additional secondary ice production, leading to a cascading process of ice multiplication that can rapidly glaciate a cloud.

Field observations provide strong evidence that secondary ice production is important in the atmosphere under appropriate conditions (Korolev et al. 2022), but the relative importance of the various mechanisms across different cloud regimes is not yet well understood, and laboratory studies are limited (Sullivan et al. 2017; Korolev and Leisner 2020; Burrows et al. 2022). Given the current limited state of knowledge, new laboratory studies that isolate and quantify specific secondary ice production pathways under a range of relevant atmospheric conditions are likely to lead to significant new insights with potential to improve the predictability of precipitation and clouds.

Our ability to measure cloud droplet number concentration (Nd) in liquid-topped clouds was discussed. Previous work has argued that Nd is the variable of state in aerosol-cloud interactions (Wood 2012), which in turn are a major source of global-scale prediction uncertainty (Bellouin et al. 2019). While previous climate process teams have made strides in establishing Nd as a worthwhile observable and building a baseline for how we observe this quantity from space (Grosvenor and Wood 2014; Grosvenor et al. 2018), there is still substantial uncertainty in Nd across different observational methodologies. This includes airborne (Lance et al. 2010), surface remote sensing (Zhang et al. 2023), and remote sensing from space (Gryspeerdt et al. 2022b). It is encouraging that when attempting closure between surface and airborne measurements of Nd (Zhang et al. 2023) and spaceborne and airborne (Gryspeerdt et al. 2022b) there is some level of consistency. Overall, the leverage that Nd exerts on our ability to offer predictions (Carslaw et al. 2013; McCoy, I. L. et al. 2020; Jones et al. 2025) supports the need to continue to refine our measurements of Nd, even if these observations are broadly available. Advances in other measurement techniques such as multi-angular polarimetry and high vertical resolution lidar (Hu et al. 2021; Zhang et al. 2023) are promising avenues that are likely to provide substantial return on time invested (Gao et al. 2023). However, it is important to keep in mind that even what seem to be relatively direct measurements from cloud droplet probes can be subject to substantial systematic uncertainty due to uncertainty in sampling volume (Lance et al. 2010).

The workshop discussion briefly mentioned the importance of observing aerosol and cloud chemistry and the lack of systematic observations of chemical composition in surface measurements. Chemical processes alter cloud properties primarily by affecting aerosol composition (McFiggans et al. 2006). The chemical components of particles and droplets likely vary widely over spatial and temporal scales. Laboratory measurements of chemical properties of aerosol particles and droplets are extremely diverse and have spurred major advancement in instrumentation from Aerosol Mass spectrometers, chemical ionization mass spectrometers, single-particle mass spectrometers, and numerous others (Seinfeld and Pandis 2016). Laboratory measurements have shown that particle composition influences the temperature at which cloud droplets freeze (Jahn et al. 2020; Bieber et al. 2024). The sheer complexity of the chemistry occurring in particles and droplets makes it extremely difficult to parameterize and incorporate into ESMs. As such, the challenge remains to obtain sufficiently detailed measurements, both in the laboratory and field, on key chemical processes and relating these processes to the microphysics expressed in ESMs across the large spatial and temporal scales of ESMs.

4 HOW ARE OBSERVATIONS BEING USED TO IMPROVE ESMS?

Where we are: Observations are increasingly being used to constrain ESMs systematically, but major gaps remain in uncertainty quantification and process-level constraints.

Progress made: Some success has been achieved in using PPEs to quantify uncertainties, identify ESM structural deficiencies, and improve calibration.

Remaining gaps: There is a lack of well-quantified observational uncertainties, inter-instrument differences, and robust model-observation representation error characterization. We also lack the ability to perform the requisite LES experiments required to bridge observations and ESMs.

What's needed: More rigorous and comprehensive observational uncertainty quantification and multi-scale frameworks would enable linking observations to ESMs.

Key Findings:

- Observational constraints on PPEs can reveal ESM structural deficiencies.
- Careful calibration of PPEs as opposed to model tuning are needed to match a wide range of observation metrics, accounting for observational uncertainties, and to avoid overconfidence in constraints.
- PPE calibration requires better-quantified observational uncertainties, characterization of inter-instrument errors/differences, and model-observation representation errors.
- Model constraints on atmospheric state alone are insufficient for constraining aerosol forcing and climate sensitivity. There is a need to additionally constrain key processes, whilst accounting for covariability of aerosol, clouds, and vertical motion.
- LES simulations provide a crucial bridge across scales, but a lack of routine LES, and limited representativeness of existing LES ensembles hinders progress.
- Efforts are needed to operationalize the cycle of model development, PPE creation, and observational constraint.
- Multi-scale frameworks are needed to link observations to ESMs.

Discussions at M2M underscored the pivotal role of observations, including laboratory, in situ, and remote sensing, in informing and improving ESMs. Traditionally, ESMs have been tuned to match a variety of metrics of the present-day state (Mauritsen et al. 2012; Hourdin et al. 2017); however, workshop participants agreed that this approach alone is likely insufficient to constrain aerosol forcing or climate sensitivity and produce robust future projections. Instead, a range of strategies is emerging for harnessing observations to refine ESMs (Regayre et al. 2018; Watson-Parris et al.

2021), from improving individual process representations to integrating multi-scale observational data into comprehensive uncertainty frameworks. The BIFES framework described in Section 1.3 builds on this.

A range of state-of-the-art strategies for reducing ESM prediction uncertainty using observations was showcased at M2M. These include: (a) improving parameterizations of individual microphysical processes; (b) emulating or applying machine-learning to represent microphysical processes; (c) calibrating model ensembles; and (d) using PPEs to identify structural model deficiencies.

Based on guidance emerging from discussion at M2M, (a) would ideally adopt a Bayesian context for any model improvement (Section 1.3). Discussion at M2M recommended considering multiple interacting processes in the development of ESMs (e.g., via (b), (c), and (d)).

Current ESM evaluation methods are diverse, with different modeling centers utilizing their own preferred subset of available observations. This has led to a scenario where individual ESMs exhibit skill at simulating certain aspects of the present-day climate yet, when compared across a broad set of observational metrics, none outperforms the (unphysical) multi-model mean (Gleckler et al. 2008; Knutti et al. 2010; Parsons et al. 2021). Such diversity suggests that despite advances in ESM fidelity informed by observational constraints, ESMs still produce inherently uncertain future projections in part because each ESM is getting different aspects of the Earth system 'right.'

Contemporary ESM development efforts focus on ensuring that simulated atmospheric states agree with observational benchmarks. This is a zeroth-order metric for ESM believability that makes good use of abundant observations of the current atmospheric state. However, matching present-day conditions does not automatically translate to reliable future projections (Figure 7).

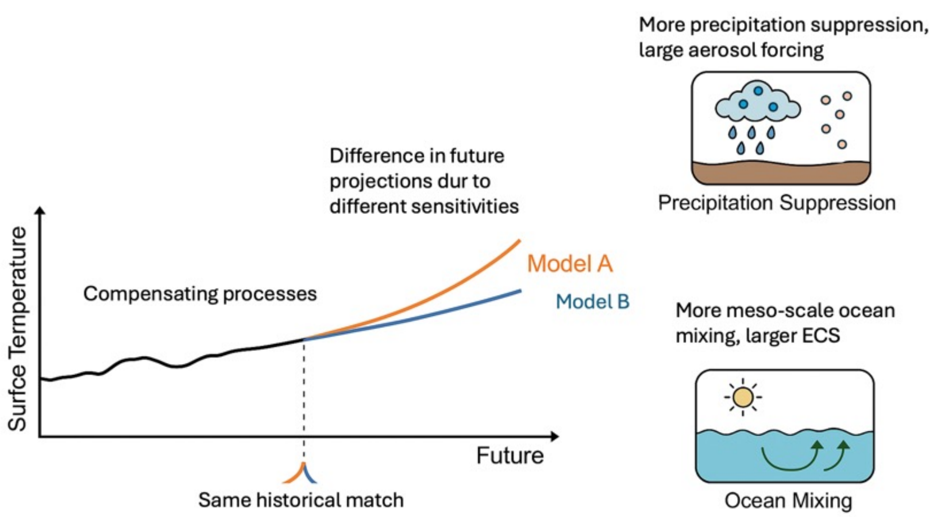

Figure 7. An illustration of how the historical record of observations can be matched by more than one set of process representations, but with dramatically different future predictions (higher or lower Equilibrium Climate Sensitivity; ECS). Courtesy of M2M SOC.

There are many observables that have little to no bearing on the processes that we know are important drivers of prediction uncertainty. Participants emphasized the need for further process-level observations that explicitly account for meteorological covariability across multiple atmospheric regimes to target the persistent drivers of ESM uncertainty. Lagrangian observations (e.g., via unmanned aerial vehicles) and enhanced geostationary satellite systems could also be used to target specific model processes. With further advancements in instrumentation, these systems could more effectively constrain process-level model uncertainties.

Increasing the complexity of individual parameterizations can lead to higher ESM fidelity by increasing the number of ways that a given ESM can move in to match observations. These increased degrees of freedom come at the expense of increased computational cost. One option to increase degrees of freedom with minimal additional cost is to replace some ESM components with parameterizations generated by machine learning either through empirically matching observations or attempting to emulate existing high-cost parameterizations (Beucler et al. 2021; Harder et al. 2022; Perkins et al. 2024).

As noted by participants, increased parameterization complexity can increase model fidelity and may potentially improve processes that are relevant to seasonal-to-decadal prediction. However, parameterizations cannot fully capture the complex interactions that drive predictability even if they are improved to more accurately represent a given process or observable. The analogy of "perfectly crafted puzzle pieces that don't fit together" illustrates the limitations of isolated process improvements and reinforces the need to consider how processes interact and differ between environments (Section 1.3). Challenges remain in producing parameterizations that capture essential principles governing microphysical systems, rather than just statistical correlations in training data. A key concern is that overfitting or attempting to replace physics with machine learning could be inherently risky for the core mission of ESMs. ESMs need to provide stakeholders with seasonal-to-decadal predictions, which often requires making out-of-sample predictions and so require careful validation and testing procedures. This was brought up both at M2M and at the follow-on breakout at the ARM/ASR meeting.

Another option to improve ESM fidelity is by more explicitly leveraging LES, which provide a crucial bridge between explicitly resolved processes and the subgrid scale parametrizations used in ESMs. LES allow for controlled experimentation and detailed process-level insight, making them a powerful tool for constraining and developing parameterizations. However, routine use of LES for this purpose is in its infancy, and existing LES ensembles are often highly idealized and narrow in scope. For instance, large-scale LES ensembles have explored key axes of cloud variability, such as cloud droplet number (Glassmeier et al. 2019) or surface and environmental conditions (Jansson et al. 2023), but rarely both simultaneously (Sansom et al. 2024). While these datasets have enabled progress in quantifying physical sensitivities and parameter uncertainty (Hoffmann et al. 2020), their representativeness remains limited. This restricts the generality of any parameterizations derived from them, particularly when the underlying processes are sensitive to initialization timescales or exhibit transient behavior (Glassmeier et al. 2021). Broader and more systematically designed LES ensembles, including those targeting observationally relevant regimes (Erfani et al. 2025), are therefore essential for realizing their full potential in ESM development.

A significant advancement in how we develop ESMs, relative to one-at-a-time/artisanal tuning approaches has been the use of PPEs to systematically quantify model uncertainty, then calibrate against a wide range of observations to constrain that uncertainty (Yarger et al. 2024). These approaches have been effective in exposing structural model deficiencies, such as when a model may excel at simulating one aspect of the Earth system at the expense of another (Regayre et al. 2023). Despite these successes, the potential to calibrate PPEs against observations is currently limited by uncertainties in the observational data (Section 3.1) (Elsaesser et al. 2025) and the lack of robust frameworks to quantify model–observation representation errors (Schutgens et al. 2017) and efficiently identify and address structural model deficiencies.

For observations to more effectively inform ESM development, robust observational uncertainty quantification is essential. This means not only recording instrument-specific uncertainties but also understanding and accounting for inter-instrument discrepancies and spatial-temporal representation errors (Section 3.1). M2M participants envisaged the creation of a comprehensive observational database that incorporates these uncertainties and enables consistent benchmarking across ESMs. In addition, a clear need was identified to conduct more laboratory measurements to better estimate measurement uncertainties, and indirectly improve process-level understanding, as these uncertainties are often unknown beyond the intuition of the instrument developer. The observational database would include multiple retrievals of key observables such as cloud droplet number, along with pixel level uncertainties for each product. This supports the retrieval bundle idea discussed above. This database would enable modelers and observationalists to establish agreed tolerances for each observable variable, thereby determining which ESM variants are most likely to produce observationally-plausible future projections.

In turn, well-calibrated PPEs could provide actionable insights on where new observations are most needed (Section 5), creating a feedback loop that can be used to refine both ESM development and observational practices. Furthermore, because certain ESM processes and properties, such as ice microphysics, ice crystal radii, and aerosol-cloud interactions, are highly sensitive to structural choices made during parametrization, parameter perturbation effects will vary between ESMs. One way around this that was discussed at M2M was the idea of multi-model PPEs that sample both parametric and structural uncertainties (Shiogama et al. 2014). While not a new idea, the approach has not yet been attempted for several reasons: only a small number of modeling centers have historically had the capability to perform PPEs; there are large structural differences between ESMs and a lack of agreed-upon observational metrics. Nevertheless, these multi-structure ensembles will be a crucial tool for characterizing the shared causes of prediction uncertainty across the next generation of ESMs.

A recurring theme was the need for multi-scale frameworks that can bridge the gap from lab experiments and microphysical observations to particle-resolved and large-eddy simulations, and ultimately to ESMs. Some recent work was presented along these lines connecting a process model-based PPE with a simple energy-balance model to constrain the processes based on historical temperature trends and leading to a reduced posterior uncertainty for this model (Watson-Parris 2025). Further work is essential to quantify how uncertainties propagate across these scales so that climate-scale parametrization uncertainties can be robustly quantified.

Overall, while M2M highlighted progress in using observations to improve ESMs, the transition from promising pilot studies to an operational framework applicable across multiple modeling centers will require a more rigorous, multi-scale, and uncertainty-aware approach. By integrating enhanced process-level observations with comprehensive uncertainty quantification and multi-model calibration techniques, the community can make significant strides toward reducing the uncertainty in Earth system projections and improving the reliability of ESMs.

5 HOW CAN WE PLAN FUTURE PROCESSES OBSERVATIONS TO REDUCE PREDICTION UNCERTAINTY?

Where we are: As the field of atmospheric science matures, field campaigns are increasingly shifting from exploratory to hypothesis-driven science, leading to a need for strategic campaign planning that can robustly test specific hypotheses. Simultaneously, observation-driven improvements to ESMs remain limited by a lack of available frameworks to differentiate between parametric and structural uncertainties in complex systems.

Progress made: There is growing recognition of OSSEs, digital testbed approaches, and model-guided experimental design for planning and prioritizing observations.

Remaining gaps: Observations often lack systematic uncertainty quantification, and observational campaign designs are not always well-integrated with ESM development needs.

What's needed: Innovative, robust machine learning and software frameworks are needed to support parametric and structural uncertainty quantification, accelerate the model development life cycle, and support strategic field-campaign design. A structured approach is needed to designing and prioritizing field campaigns for process-level constraints, using a hierarchy of models.

Key Findings:

- Model-Guided Experimental Design: observing system simulation experiments (OSSEs), virtual deployments, and model hierarchies are needed to optimize and prioritize high-cost fieldwork aiming to improve understanding of the processes relevant to seasonal-to-decadal environmental prediction.
- Shorter-term predictability of weather can be evaluated during the period of most field campaigns. For seasonal-to-decadal environmental prediction there needs to be a modeling step that lets short-term observations link to seasonal-to-decadal questions.
- Advances in process understanding are possible from opportunistic experiments and perturbation studies in outdoor environments.
- Field campaign proposals targeting specific processes through "closure studies" are often deemed too pedestrian, but such targeted measurement strategies are critical for constraining processes.

Efforts to better understand microscale process observations and their larger-scale effects now span several decades. In the period after World War II, as observations of microphysics underwent a surge of interest (Vonnegut 1947), there was so much new scientific territory to explore that any foray into the atmosphere yielded observational riches and improved physical understanding. This can be thought of as a period of setting our priors. Over time, understanding of microphysical processes and their interactions has improved, and models have been developed and refined to represent them. Scientists have developed increasingly complex models of cloud microphysics processes, drawing on laboratory experiments, observations of clouds and precipitation in the natural environment, and theory (Morrison et al. 2020). Aerosol sources and distributions have been better characterized, and major global sources of aerosol and their primary known feedbacks within the Earth System are increasingly represented in models (Carslaw et al. 2010).

As the field of atmospheric science matures, we have witnessed a shift from exploratory field campaigns to more hypothesis-driven and integrative science, including efforts to characterize and reduce uncertainties in key large-scale ESM-projected outcomes that depend critically on microscale processes. Programs such as the Climate Process Teams (CPTs) have succeeded in bringing together modelers and observational scientists to target specific process improvements in ESMs, using existing observations as constraints.

Simultaneously, advances in computational power are enabling scientists to simulate larger domains at higher resolutions than ever before. This includes both the recent emergence of a new class of kilometer-scale "convection-permitting" models (Donahue et al. 2024), and the application of limited-domain models, including large eddy simulation (LES) models, to increasingly larger domains. From both sides, the gap in scales continues to narrow between microscale process observations and large-scale simulations of cloud, aerosol, and boundary-layer processes.

These transitions have led to both increased opportunities to improve the process fidelity of atmospheric models, and an increasing need for strategic campaign planning to robustly test specific hypotheses. Simultaneously, the pace of observation-driven improvements to ESMs is still limited by a lack of robust, hierarchical frameworks that facilitate the integration of process-oriented observations into the development workflow.

There is a growing recognition that simulations can be used to identify where and how potential new observations will have the greatest ability to constrain specific physical processes, or address process-oriented science questions. For example, high-resolution atmospheric models such as LES are now sufficiently advanced that they can produce skillful representations of field experiments (McCoy et al. 2024), suggesting that they could and should be used more extensively in field experiment design. Models are increasingly being deployed to compare alternative designs for proposed laboratory facilities that aim to isolate and study specific cloud microphysical processes (Wang et al. 2024). And models are being used at the ESM scale to identify the locations and types of observations that are most informative for specific science questions (Regayre et al. 2014; Gettelman et al. 2020; Song et al. 2025).

One approach for systematically employing models in such a fashion is an observing system simulation experiment, or OSSE. The term "OSSE" originates in the remote sensing community, where OSSEs have long been used, for example, to simulate the impacts of proposed designs of new satellite systems (Hoffman and Atlas 2016).

More broadly, an OSSE can be defined as a simulation-based quantification of the expected capability of any future observing system to meet its stated design purpose. In their review of current OSSE use by US agencies that make Earth observations, Zeng et al. (2020) differentiate three different types OSSEs, which can be employed to achieve different aims:

1. to optimize observational data for forecast improvements (forecasting OSSEs),
2. to determine the spatial and temporal sampling requirements to address a specific science question (sampling OSSEs), and
3. to define technical requirements for new instrumentation to address a specific science question, before that instrumentation is built (retrieval OSSEs).

An OSSE framework that directly connects process-oriented observing systems to key aspects of Earth System predictability would provide a powerful tool to optimize the design of future field campaigns, ground-based observing networks, and satellite observations, for the purpose of addressing specific science questions and optimizing specific predictions of interest.

One major challenge in building such a framework is the need to connect fast microscale processes to their macroscale effects at the seasonal-to-decadal time frames. Tackling this challenge requires a hierarchy of models that can resolve processes occurring at different scales (Figure 3).

Bridging across these multiple scales is a longstanding challenge, but modern machine learning methods are opening new pathways to tackle this challenge. The application of machine learning to problems of Earth system predictability is currently an active research frontier. The simultaneous exploration of a variety of approaches by multiple research teams will drive innovation in this area in the coming decade.

We argue that we are now in the period of needing to craft studies that strategically target high-value processes that we know impact predictability (Kahn et al. 2017). As Earth system science evolves and matures, there should be an expectation that large, resource-intensive field studies and observing systems need to objectively demonstrate how they will target specific model process improvements, either by providing observations that are sufficient to differentiate between specific hypotheses, or by obtaining quantitative reductions in known uncertainties through observational constraints. Simply put, observational campaign/deployment designs need to be better integrated with OSSEs that assist in planning of technical requirements and sampling needs and driven by ESM predictive needs. This requires practitioners across multiple subdisciplines—from process-oriented observations to modeling at the Earth System scale—to develop new ways of coordinating.

5.1 Causality in the Earth system

Studying microscale processes in the context of the Earth system is not only a challenging multiscale problem (Figure 2), but also requires grappling with a system intensely subject to equifinality[7] (Beven and Freer 2001), where causal inference is difficult (Section 1.3). Being able

7 Many combinations of processes can produce the same outcome.

to isolate flows of causality between observables is challenging in the atmosphere (Stevens and Feingold 2009; McCoy et al., D. T. 2020; Mikkelsen et al. 2025) and ultimately necessitates observing some sort of perturbation to the system (McCoy and Hartmann 2015; Mace and Abernathy 2016; French et al. 2018), modeling to build causal understanding (McCoy, D. T. et al. 2020; Mikkelsen et al. 2025), or both (Malavelle et al. 2017; Toll et al. 2019). During M2M, several approaches were discussed as promising avenues to tackle causality inference using observations and connect it to decadal predictability. We briefly outline them below.

Opportunistic or “natural” experiments (Christensen et al. 2022) enable studies of a system’s response to external perturbations, for example cloud responses to volcano plumes. A number of recent ‘opportunistic experiments’ have studied cloud microphysical responses to transient aerosol changes such as those from volcanoes (McCoy and Hartmann 2015; Mace and Abernathy 2016; Malavelle et al. 2017), shipping (Gryspeerdt et al. 2019; Diamond et al. 2020), and changes in emission controls (McCoy et al. 2018; Wall et al. 2022). Other phenomena can also be leveraged to understand how clouds respond to warming; for example, through studies of marine heat waves (Myers et al. 2021; McCoy et al. 2023).

Studies where observations are coordinated with intentional perturbations to the atmospheric state also have a long history in the field (Vonnegut 1958). Historically, these studies were not designed to achieve the goal of advancing fundamental process understanding. However, new approaches in recent work where the experimental design of perturbation studies has been scoped to improve basic physical process understanding have produced valuable insights (French et al. 2018; Henneberger et al. 2023). Perturbative studies in which microphysical cloud or aerosol changes are induced by deliberate injection of aerosol particles provide a clear way around the causality problem. These include projects such as Seeded and Natural Orographic Wintertime clouds: the Idaho Experiment (SNOWIE) (Tessendorf et al. 2019); Eastern Pacific Emitted Aerosol Cloud Experiment (E-PEACE) (Russel et al. 2013); and marine cloud brightening over the Great Barrier Reef (Latham et al. 2013; Hernandez-Jaramillo 2025).

5.2 From probe to planet

As pointed out by several speakers at M2M, we need to identify ways to make our observations better constraints for ESMs, because ESMs will continue to be a critical tool needed for predictions at the key seasonal-to-decadal scale. This needs to be done with alacrity. We need to develop efficient, tractable solutions for how we unite field campaigns and the laboratory with ESM development to be able to provide stakeholders with robust, actionable projections of years and decades into the future. Field experiments have typically been designed to provide new knowledge and understanding of basic processes but are increasingly being used to address deficiencies in the representation of these processes in large scale models.

As an example, the 2008 Variability of the American Monsoon Systems (VAMOS) Ocean-Cloud-Atmosphere-Land Study (VOCALS) program (Mechoso et al. 2014) involved two large-scale model assessment projects. The first (Wyant et al. 2010) was designed to take place prior to the VOCALS field campaign and helped set the observational sampling protocol for the campaign. Based on this, VOCALS adopted a “routine” sampling approach that allows the building of statistical sampling along one geographical transect that is repeated numerous times during the campaign (Bretherton et al. 2010; Wyant et al. 2015). The routine sampling approach is more common now (Wofsy 2011; Albrecht et al. 2019; Redemann et al. 2021), establishing a framework for meaningful direct comparison of campaign-mean properties and process-driven relationships between variables with ESMs. However, we are still falling short on using ESMs to systematically aid the design of most field studies. Model simulations should be sampled to evaluate a variety of different possible sampling approaches, for example whether continuous sampling at a stationary surface site or sampling at limited times but over a larger spatial area provides more stringent constraints on processes of interest in ESMs.

An extensive literature for spaceborne observations exists through OSSE (Arnold and Dey 1986; Cucurull et al. 2024) approaches. Field observations can be modeled using an OSSE approach, but currently this is rarely carried out prior to the observational campaign. This would be extremely valuable for field experiment design and necessary to tell us the impact of a given field campaign on our ability to offer robust decadal predictions. Satellite OSSEs use models to test different satellite system designs before instruments are built or deployed. A critical step is to compare the proposed new satellite performance with what is available from the current observing platforms to establish the benefits of new measurements.

A field campaign approach to the OSSE framework could be used to optimize where and when to deploy field measurements, determine what geophysical variables must be measured and how accurately, and determine if the experimental hypotheses to be tested can be refuted given the expected sampling uncertainties and covariances between variables that may make hypothesis testing difficult. An example is a potential closure study for warm rain falling from marine stratocumulus clouds, in which the hypothesis to be tested is that, after controlling for meteorologically driven variability, we are able to constrain, with sufficient statistical power, the sensitivity of warm rain to the concentration of cloud droplets in the cloud. This is the essence of the precipitation susceptibility metric (Feingold and Siebert 2009). We know that the amount of condensate a cloud contains, sometimes expressed in the form of the liquid water path (LWP), is the leading predictor of a cloud's propensity to precipitate (Byers and Hall 1955; Pawlowska and Brenguier 2003). To account for this thermodynamic and meteorological control on precipitation, the field study would need to be able to take sufficient coincident measurements of LWP and precipitation rate to be able to regress out the impact of LWP. But there must also be a sufficient variability in the cloud droplet concentration to exert a detectable influence on precipitation above that driven by variability in LWP. Using the field OSSE approach, we would be able to essentially perform a precipitation susceptibility analysis with the ESM output to establish the potential for successfully testing the experimental hypothesis.

There are consequential choices involved in how to sample the atmosphere in field studies [e.g., survey type flying designed to sample large-scale gradients in cloud or aerosol properties (Wood et al. 2011) vs highly localized flights designed to capture specific processes (e.g., cloud droplet activation)]. These different types of sampling are often in tension. A result of this tension is that field experiments frequently have too many objectives, make too many sampling compromises, and do not focus sufficiently on pressing unsolved issues. How would the experimental planning process differ if we were to use OSSE data (e.g., from a month-long simulation of a specific target region) to design sampling strategies in an objective way? Spaceborne missions have to make hard choices on the ground that cannot be fixed once the satellite is in orbit, while there may be the perception with field campaigns that 'we can fix it once we are there.' However, once in motion field campaigns have substantial inertia and decisions like what instrumentation is being deployed are rarely revised in any substantial way. It would make sense to use OSSE techniques for field deployments and campaigns in the same way that we use these for spaceborne missions.

The importance of and requirement for OSSEs as part of field campaigns were key points of community consensus at M2M. This was revisited at the DOE ASR/ARM meeting in 2025 as part of a M2M follow-on breakout on OSSEs and uncertainty quantification.[8] There was agreement in the breakout that these tools would have been useful for recent major efforts such as the Bankhead National Forest deployment for topics as simple as the site deployment and has been critical to result from the Pi Chamber at Michigan Tech where processes need to be inferred from observables—even when the environment (a cloud chamber) is strongly controlled.

We view the OSSE concept as a key component of the overall BIFES framework proposed in Section 1.3. OSSEs form the link between observations and models that can be used to explore multi-process uncertainty. One key recommendation from M2M was to continue to build a framework that allows us to plan future field campaigns and make clear, supported statements about how the cost of deployment is valued in reduced prediction uncertainty.

While the OSSE and encompassing BIFES framework may not be needed for every sort of field deployment, it is needed for studies that seek to improve our seasonal-to-decadal predictions. Studies examining short-term phenomena that can be characterized within the days-to-weeks period of a typical field campaign may not benefit from this step, but as far as we know there is no way to assess the impact of new observations on longer-term predictions without incorporating an ESM (Figure 3).

One consistent feature of discussion at M2M was the idea that closure studies were a method that allowed observations and models to close the gap between observables, causality, and process understanding. Closure can take many forms and effectively means the attempt to predict an observable based on a model of the processes linking different characteristics of the system (Wood et al. 2012; Kang et al. 2022). Closure studies represent one of our best tools that we can use to really build process understanding and link between microphysics and what we can implement in an ESM. It was pointed out that it is difficult to propose campaigns that focus on getting a single process study right as being too narrow and pedestrian. One possibility is that by being able to better contextualize the impacts of getting a narrow set of processes right with a focused campaign would better support the need for more targeted field work. For instance, a closure study focused on one specific convective rain process might not find support using typical campaign criteria, but if the infrastructure existed to show that a campaign focused on closure in this quantity would improve, say, our prediction of heavy rain over Alabama in the

8 See: Uncertainty quantification and Observing System Simulation Experiments (OSSEs) for ARM data

next decade, support for the campaign could be motivated by economic necessity. One example of this is the SAM-CAAM systematic measurement concept (Kahn et al. 2017).

One impact of operationalizing and requiring this step alongside traditional pre-campaign activities like flight planning would be to lead to a greater return on investment in field experiments and the targeting specific processes. Discussion at M2M began to sketch out what this might look like. We have provided a rough flow chart to illustrate this in Figure 8 and we discuss this concept below.

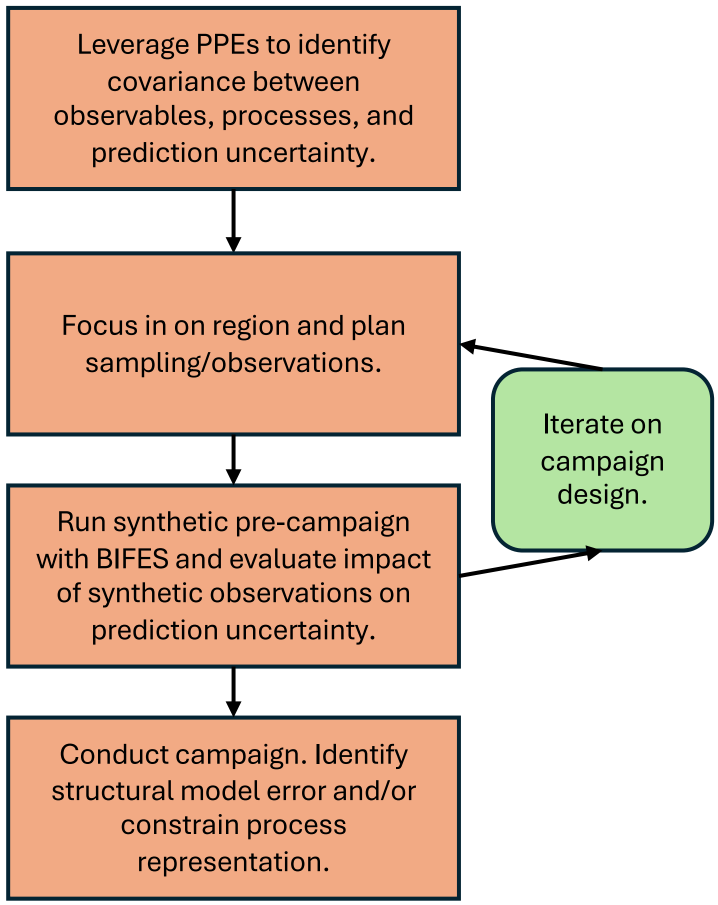

Figure 8. A potential approach to deploying observations to improve Earth system predictability. Courtesy of M2M SOC.

This is illustrated with the idea that the proposed campaign should improve seasonal-to-decadal predictability in the Earth system. First, we need to evaluate where we think uncertainty is coming from. One way to do this systematically could be to use one of the many existing PPEs, or some aggregate of them, to look at how observables from a campaign correspond to ESM predictions. One example of this might be work showing substantial constraint of historical radiative forcing by summertime cloud droplet number in pristine, high-latitude environments (Carslaw et al. 2013; McCoy, I. L. et al. 2020; Song et al. 2025). Having identified an observable (or set of observables) and a region, and given ourselves some priors regarding what sort of constraints on predictability we can get from measurements there, we can start to burrow down to the details of campaign design. This could be achieved through a mix of single column modeling and relatively high-resolution simulations to plan details like whether to sample systematically or whether to target particular weather states. Once we have an idea of the sampling strategy, we can plug our synthetic campaign into the BIFES framework. Tools to contrast ESMs and field observations exist (Gettelman et al. 2020) and can be implemented in this step with the background of a PPE-type design in BIFES. At this point we may find that our decisions about what sort of measurements to make and what sort of sampling to employ may produce insufficient changes on predictability at the global scale. Different combinations of measurements and sampling approaches can be rapidly attempted to arrive at an optimal deployment. Key observational metrics like how big a systematic uncertainty are tolerable can be tried out here. For instance, we might find that we need to fly several instruments measuring the same quantity to create an observation bundle (Elsaesser et al. 2025; Jones et al. 2025; Zhang et al. 2023). Finally, once the campaign is completed, the results can be plugged back into the existing framework relatively quickly and can evaluate whether we are able to reduce uncertainty in a purely parametric manner or whether we also find a structural disagreement with the ESMs we rely on to produce long-term predictability.

6 WHAT DO WE DO NEXT?

Where we are: ESM developers increasingly recognize the importance of prediction uncertainty quantification, but practical implementation is lagging.

Progress made: Steps have been taken toward integrating PPEs into ESM development cycles and multi-ESM comparisons. Major steps forward have occurred in observing microphysics and adding new processes to ESMs.

Remaining gaps: There is need for a community of practice, standardized best practices for ESM-observation comparisons, and more coordinated multi-ESM studies.

What is needed: New approaches to implement include multi-ESM PPE studies, multi-scale frameworks to link from observations to ESMs, hypothesis-driven multi-model attributions, and uncertainty-aware seasonal-to-decadal projections.

Key Findings:

- Emphasis is needed on uncertainty in ESM seasonal-to-decadal predictions, departing from a historical focus on state fidelity.
- Advance tractable and robust uncertainty quantification in observational products, paired with instrument simulators and Bayesian uncertainty (Section 1.3) are needed.
- There is a need to plan observations that are coupled to ESM development to identify high-value observations.
- Coordinate Multi-ESM PPE studies to characterize both structural and parametric uncertainty.

After two days of presentations and discussion the SOC went into retreat to develop a top-level set of recommendations based on M2M. These top-level recommendations were presented in the final half day to all in attendance to assess broad consensus. We describe these recommendations in this section.

First, it is important to recognize the massive strides that the community has made in moving towards tackling one of today's most challenging and societally-impactful multi-scale scientific problems. This problem has been with us since the very earliest days of developing ESMs. Syukuro Manabe's recent Nobel prize win emphasizes the need to consider small, fast convective processes to represent and predict the atmospheric structure (Manabe and Wetherald 1967). ESMs that can deliver predictability across the seasonal-to-decadal boundary have grown in complexity and skill. However, simply because of the massive scale gap between the planet and the microscale (Figure 2), we are still confronted with the grand challenge of abstracting those processes and harnessing them to offer robust, reliable predictions.

Bridging the scale gap between planetary and microscales is hard and by nature will introduce uncertainty. We emphasize that the Earth system is coupled, and we need to have global models (ESMs) that allow for this coupling and represent teleconnections across the planet. We know that small, fast processes directly impact ESM predictions (Lehner et al. 2020; Sherwood et al. 2020; Watson-Parris and Smith 2022; Gettelman et al. 2025). There has historically been a great deal of emphasis on the importance of parameterized processes related to the top-of-atmosphere energy budget (Andreae et al. 2005; Zelinka et al. 2020). However, regional changes are the key consideration as we provide guidance to stakeholders. While microscale processes are clearly important for the global energy budget, these processes are also important for our ability to provide regional forecasts at the seasonal-to-decadal scale. Parameterized processes set patterns of heating and cooling in the atmosphere that in turn set circulations. We can leverage high resolution models to reduce our reliance on some parameterizations, but it is important to emphasize that the scale gap is still enormous, and regional models do not explicitly resolve phenomena like convective clouds and smaller turbulent motions that are key to mesoscale cloud organization and cloud radiative properties (Figure 2). In addition to not resolving relevant microscale processes, regional models are still dependent on the boundary conditions from ESMs to offer future predictions.

The discussion above is not meant to disparage any class of model. It is important to recognize that confronting the complex, multiscale problem of predicting the Earth system requires a hierarchy of model classes that each resolve different features and have different scales of geographic and temporal coverage. However, if we seek to offer stakeholders predictions on the scale needed to make societal decisions for the future, we are going to need ESMs, and to have reliable ESMs, we need to better understand and constrain ESM prediction uncertainty.

6.1 Prediction uncertainty instead of fidelity

Given that we are stuck with parameterization, what steps can we take to reduce prediction uncertainty? As a scientific community, we have been making ever more complex ESMs over the last few decades, yet key metrics related to historical forcing uncertainty and climate sensitivity have not narrowed (Bellouin et al. 2019) and have in some cases broadened (Zelinka et al. 2020) as we move to ESMs that represent more processes and allow higher prediction accuracy.

Broadly, the key deliverable of an ESM is a reliable, interpretable prediction of future environmental state with the ability to couple to emissions decisions. The seasonal-to-decadal time scale is key to most forms of infrastructure planning from bridges to nuclear power plants. However, the way that we develop ESMs is often not focused on uncertainty in future predictions as much as it is focused on matching the mean-state of various observables. Essentially, any ESM can be built to match observables well by adding more and more parameterization and model structure, but this may not necessarily address underlying causes of prediction uncertainty. Here, we make suggestions for dealing with more sophisticated ESMs to allow higher accuracy predictions of the future.

Perturbed parameter/process/physics ensembles (PPEs) were discussed at length during M2M. These are powerful tools that allow a systematic evaluation of parametric uncertainty. They also reveal how ESM equifinality limits observational constraint (Lee et al. 2016), which needs to be addressed to provide stakeholders with robust predictions. A clear example of the tension between calibrating ESMs in this framework was presented by Greg Elsaesser (Elsaesser et al. 2025). The NASA GISS calibrated PPE (CPE) is built on the GISS E3 model where many parameters that we know affect ESM behavior are systematically perturbed. The resultant ensemble is confronted with an array of spaceborne observations. This results in a set of ESMs where our observational record and the ESMs are in the best agreement.

The results presented at M2M by Greg Elsaesser and Ken Carslaw, combined with the results issuing from other systematic explorations of how mean state maps to system response in the UK ESM (Sexton et al. 2021; Yamazaki et al. 2021), underlines the disconnect between observables and predictions. At the same time, they illustrate how we can use new approaches to ESM development to better close the gap between observations and predictions. These approaches show us how constraining an ESM with observations affects prediction uncertainty. Because of this strength, it is important that PPEs are more integrated into the ESM development cycle and that ESM development needs to focus on agreement with observations that exert the most leverage on future prediction uncertainty. Essentially, ESM development needs to move from being concerned with fidelity in replicating the observed state to concern with uncertainty in predicting the future. We have a set of computational tools to make this connection, but it requires sustained commitment to make this a concrete part of the ESM development cycle.

6.2 Planning observations and avoiding false positives

The community should push towards focused observational campaigns and deployments that prioritize reducing projection uncertainty from ESMs. This suggestion is more relevant to suborbital observations where there is a large degree of freedom in how to sample key phenomena, flexibility in which instruments are included, and the deployment cost is relatively low.

A key component of determining the utility of any observations is the need for clear systematic error assessments. For most model evaluation purposes, the key uncertain quantity is the systematic uncertainty rather than the random uncertainty. This uncertainty is conceptually difficult to quantify, but is critically important for ESM development goals. Even for relatively basic observables, systemic uncertainty can be quite large (Lance et al. 2010). Discussion at M2M exposed a disconnect between those developing models and those developing observations. Observationalists are loath to provide a reductive view of uncertainty in their observations. Systematic uncertainty assessments are challenging and are not available for many instruments. ESM developers need assessments of systematic uncertainty that consider all sources of error and scale to the ESM sampling, as opposed to a very complex, but potentially lower estimate of systematic uncertainty that cannot be readily used in the context of an ESM.

One recommendation for near-term ESM development is to create observation bundles where multiple techniques to measure an observable are grouped together as a simple assessment of systematic uncertainty. Such activities could build on the successful NASA Making Earth System

Data Records for Use in Research Environments (MEaSUREs) Program, which focuses on the generation, availability, and utility of Earth system data records. There are also major opportunities through the ARM program, which hosts instrument dense deployments that can develop observation bundles and the NASA A-Train. A longer-term suggestion directed at scientists making and providing observations is to provide an estimate of systematic uncertainty that can be used readily at the ESM level and then, if possible, higher-detail assessments. If no systematic uncertainty is available, this needs to be clearly stated in the documentation of an observation. One of the key frameworks emerging from discussion in the SOC following M2M is BIFES (Section 1.3). The BIFES framework uses PPEs combined with OSSEs/instrument simulators to provide a connection between (a) process and observations and (b) observations and prediction uncertainty. We provide BIFES in the context of a closure between ESMs and observations. Observations can provide insights into but are not themselves processes. We envision closure as the most straightforward way to make observations process-aware (Section 5). The BIFES framework illustrates that without some assessment of systematic uncertainty in observations, no process can be formally ruled out (Section 1.3).

A broader impact from the BIFES framework is that it is by nature Bayesian and can handle the equifinality in the Earth system (Section 1.3; Figure 4). Most observables are dependent on multiple processes, and by understanding this, we can update ESMs and, by extension, their predictions in a robust way. To allow for the equifinality of the Earth system, we need to consider the impacts of multiple processes and associated uncertainties into our evaluation of ESMs with observations. By not considering this, new observations have a higher likelihood of rejecting the current processes representation incorrectly (Figure 5). Rejecting previous process understanding is an alluring prospect. It is incentivized by publication and grant-making that focuses on novelty. However, this rejection may not be robust without including other sources of uncertainty or conflicting observation data. Considering uncertainty from all relevant processes to a set of observations or considering systematic error across multiple observations of a quantity widens the space of observations that are consistent with a given mathematical representation of process.

A recurrent theme of M2M was that we need to develop capacity to plan future observations to maximize leverage from those observations in reducing ESM prediction uncertainty. This process is more formalized for spaceborne deployments but is less common for suborbital observations. We view this as one of the side benefits of the BIFES framework where the effect of updating our process uncertainty with new observations can be propagated forward to prediction uncertainty. This will require sustained effort and commitment to develop an infrastructure for integration between ESMs, satellite observations, and in-situ/lab measurements. By building digital testbeds like BIFES, we envision creation of robust model-guided experimental design and better understanding of how observations feed through to increased predictive skill.

6.3 The incomplete model and bridging across scales

The aphorism that 'all models are wrong, but some are useful'[9] compactly sums up some of the challenges with generating predictions of the future with ESMs. Broadly, M2M discussed the wrongness of ESMs in the context of parametric error and structural error. There is broad agreement that PPEs provide a useful way to evaluate parametric error, and these tools are becoming increasingly common for ESM developers (Sexton et al. 2021; Regayre et al. 2023; Eidhammer et al. 2024; Elsaesser et al. 2025). However, it is likely that we may simply miss processes or have entirely the wrong mathematical representation of them—what then? A common theme across M2M was the idea that we can at least begin to approach the idea of structural uncertainty by contrasting PPEs hosted in different ESMs. There are other ways to incorporate structural uncertainties, including the use of more generalized functions that can be adjusted parametrically to allow a single ESM to allow a range of structurally different parameterizations. Contrasting observations with predictions from multi-ESM ensembles that have both varying both model structure and parametric uncertainties will give us our first real chance of identifying and addressing sources of structural uncertainty.

Exercises like the coupled model intercomparison project (CMIP) (Taylor et al. 2013; Eyring et al. 2016) already exist to compare model structure, but it is opaque whether differences across ESMs are due to their parametric choices or due to structural differences, which are conflated in each ESM submission. A suggestion issuing from M2M was the coordination of multi-center efforts to contrast existing PPEs to inform our understanding of structural errors in ESMs. This approach would feed back onto how we effectively design future observations in a hypothesis-driven framework and tell us which processes we need to focus on to address structural gaps in models.

6.4 A new framework to reduce prediction uncertainty

To summarize, ESMs are a key tool allowing us to provide physical, interpretable predictions of future environmental state at the critical seasonal-to-decadal time scale for infrastructure and society.

We provide a strawman framework to illustrate the big picture guidance for an improved workflow for reducing uncertainty in predictions of the Earth system (Figure 9). The process of reducing ESM prediction uncertainty is inherently cyclical. Step 1 is to characterize which processes drive prediction uncertainty, which we can achieve through perturbing parameters and comparing across ESM structures. Step 2 requires us to formulate a testable hypothesis, which could be as simple as updating the current mathematical abstraction of a process or proposing an entirely new structure. Step 3 requires us to plan out a closure on some set of observation in the context of the aggregate of all process uncertainties. Steps 4 and 5 make observations in the environment, in a controlled lab setting, or from basic theory. Following our updating of the underlying process model, we can insert this back into an ESM in step 6 to evaluate how we have reduced prediction uncertainty, at which point we re-evaluate the dominant causes of prediction uncertainty (step 7).

9 Attributed to George Box (1919-2013).

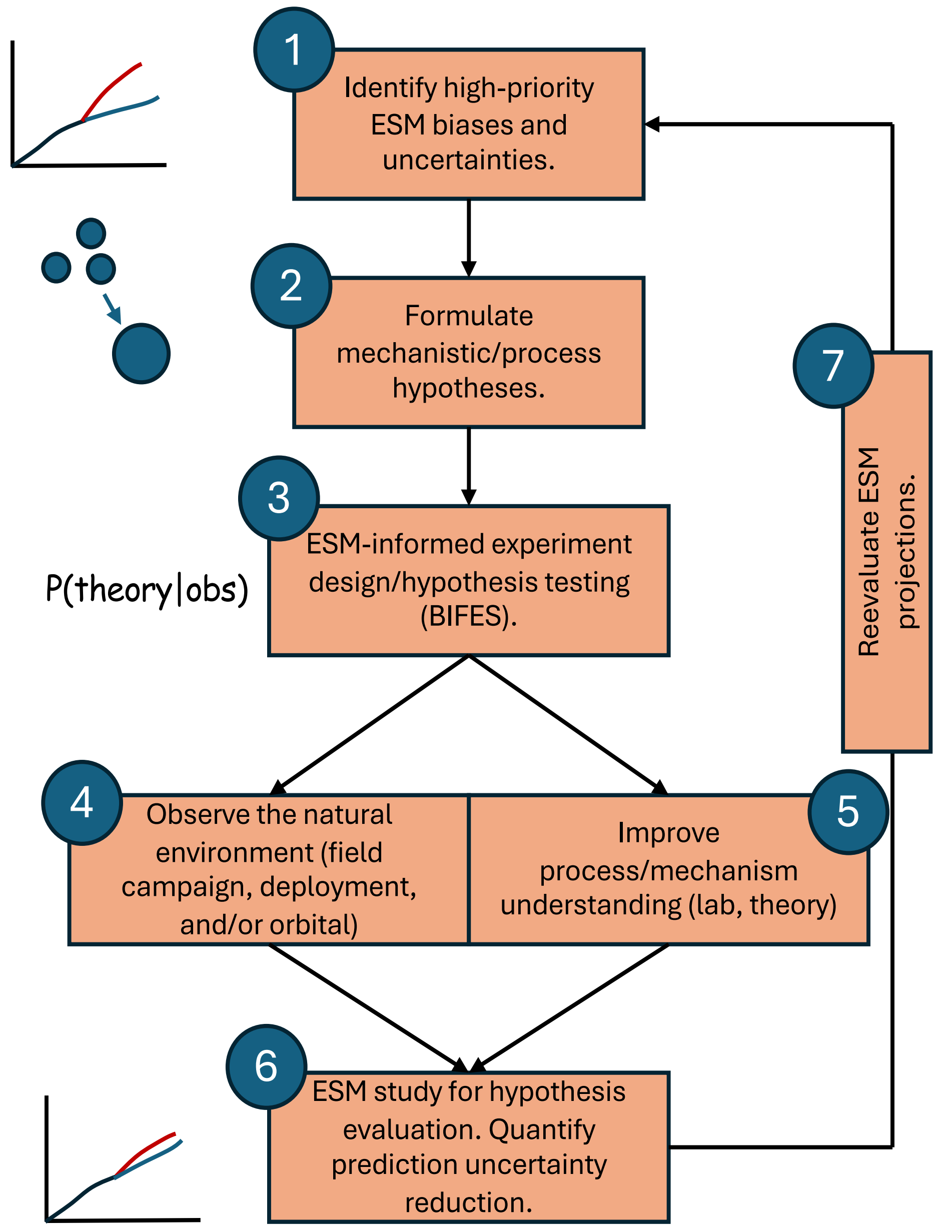

Figure 9. A proposed strawman framework for reducing prediction uncertainty in the Earth system through a synergy of model-based uncertainty quantification and observation planning. Credit: M2M SOC. Courtesy of M2M SOC.

We highlight suggestions in summary:

1. **A sustained, coordinated effort in assessing and addressing uncertainties across ESMs is needed.** High-resolution global modeling will not resolve underlying parametric/structural uncertainties in annual-decadal global predictions. While high resolution models (e.g., GSRMs and global LES) are powerful, it is not feasible to replace ESMs with them to offer global predictions at seasonal-to-decadal scale. [Figure 9; step 1]

2. **End-user guidance is needed to inform prediction uncertainty deliverables.** ESM development has focused on fidelity rather than uncertainty—in part because this is a well-posed problem. Uncertainty benchmarks for ESM developers from stakeholders would focus and constrain the ESM development cycle. [Figure 9; step 1]

3. **A need for studies that integrate global modeling with process modeling, satellite observations, and in-situ/lab measurements.** A consequence of a focus on uncertainty rather than fidelity is that it also means we need to use LES and ESMs in a way that integrates closely with the observations that will reduce uncertainty. Perturbed parameter/process/physics ensembles (PPEs) could provide a systematic characterization of uncertainty as well as a way to bridge observations, process, and prediction. Understanding observations in the context of multi-process uncertainties is a key need (Section 1.3). [Figure 9; step 3]

4. **Model-Guided Experiment Design (digital testbeds, virtual deployments, OSSEs) are a key to bridge between models and observations and plan observations intelligently.** One benefit of closer ESM-observation integration is that we can use ESMs as a digital testbed to plan for future observations. This is a shift in how some future observations are planned for and should become a recurrent and ultimately a standard component of observation planning. [Figure 9; step 4 and 5]

5. **Multi-ESM efforts are key to revealing structural uncertainty (across centers/schemes).** A recurrent concern is that while we have developed tools to quantify parametric uncertainty in ESMs, we have not developed the same set of tools to quantify structural uncertainty. A path towards understanding structural uncertainty is to foster collaboration where different ESMs can be contrasted to reveal structural uncertainty. [Figure 9; step 1]

7 ACKNOWLEDGMENTS

The Scientific Organizing Committee (SOC) thank the US CLIVAR Project Office staff and the UCAR CPAESS Events & Travel Team for efforts through all phases of the workshop. The SOC thank the following individuals for their help in organizing the reception and coordinating logistics: Adam Majewski, Charlotte While, Travis Aerenson, Hunter Brown, Ci Song, and Geethma Werapitiya; the NSF University of Wyoming King Air Flight Facility: Matt Burkhart, Anna Robertson, Natalie Kille, Eric Beamesderfer, Coltin Grasmick, Owen Cruikshank, Ed Sigel, and Tom Pierce. We also thank Colleen Kaul for constructive comments that led to improvements in the text of the report.

Additionally, the SOC gratefully acknowledges the sponsoring funding agency programs: DOE Earth and Environmental Systems Modeling (EESM); NOAA Climate Variability and Predictability (CVP); NOAA Modeling, Analysis, Predictions, and Projections (MAPP); and NSF Climate and Large-Scale Dynamics (CLD). We also thank University of Wyoming (UW) Research and Economic Development, UW Atmospheric Science, the UW Science Initiative: EarthBench/Elk Mountain Observatory seed grant, Wyoming Space Grant Consortium, and the UW School of Computing for contributing funds to support M2M. We thank UW IDeA Networks for Biomedical Research Excellence (INBRE) for use of poster boards. Additionally, DTM acknowledges funding from DOE Established Program to Stimulate Competitive Research (EPSCoR), Earth and Environmental Systems Monitoring (EESM), and Atmospheric Systems Research (ASR) through DE-SC0024161. SMB acknowledges funding from the DOE, Office of Science— BER, as part of the Energy Exascale Earth System Model project.

8 REFERENCES

Ackerman, T. P., and G. M. Stokes, 2003: The atmospheric radiation measurement program. *Phys. Today*, **56**, 38–44, https://doi.org/10.1063/1.1554135.

Albrecht, B., and Coauthors, 2019: Cloud System Evolution in the Trades (CSET): Following the evolution of boundary layer cloud systems with the NSF–NCAR GV. *Bull. Amer. Meteor. Soc.*, **100**, 93–121, https://doi.org/10.1175/BAMS-D-17-0180.1.

Andreae, M. O., C. D. Jones, and P. M. Cox, 2005: Strong present-day aerosol cooling implies a hot future. *Nature*, **435**, 1187–1190, https://doi.org/10.1038/nature03671.

Arnold, C. P., and C. H. Dey, 1986: Observing-Systems Simulation Experiments: Past, present, and future. *Bull. Amer. Meteor. Soc*., **67**, 687–695, https://doi.org/10.1175/1520-0477(1986)067<0687:OSSEPP>2.0.CO;2.

Beall, C. M., P.-L. Ma, M. W. Christensen, J. Mülmenstädt, A. Varble, K. Suzuki, and T. Michibata, 2024: Droplet collection efficiencies inferred from satellite retrievals constrain effective radiative forcing of aerosol–cloud interactions. *Atmos. Chem. Phys*., **24**, 5287–5302, https://doi.org/10.5194/acp-24-5287-2024.

Bellouin, N., and Coauthors, 2019: Bounding global aerosol radiative forcing of climate change. *Rev. Geophys*., **58**, https://doi.org/10.1029/2019rg000660.

Bender, F. A.-M., 2008: A note on the effect of GCM tuning on climate sensitivity. *Environ. Res. Lett*., **3**, 014001, https://doi.org/10.1088/1748-9326/3/1/014001.

Bennartz, R., J. Fan, J. Rausch, L. R. Leung, and A. K. Heidinger, 2011: Pollution from China increases cloud droplet number, suppresses rain over the East China Sea. *Geophys. Res. Lett*., **38**, https://doi.org/10.1029/2011gl047235.

Beucler, T., M. Pritchard, S. Rasp, J. Ott, P. Baldi, and P. Gentine, 2021: Enforcing analytic constraints in neural networks emulating physical systems. *Phys. Rev. Lett*., **126**, 098302, https://doi.org/10.1103/PhysRevLett.126.098302.

Beven, K., and J. Freer, 2001: Equifinality, data assimilation, and uncertainty estimation in mechanistic modeling of complex environmental systems using the GLUE methodology. *J. Hydrol*., **249**, 11–29, https://doi.org/10.1016/S0022-1694(01)00421-8.

Bieber, P., G. H. Darwish, W. R. Algar, and N. Borduas-Dedekind, 2024: The presence of nanoparticles in aqueous droplets containing plant-derived biopolymers plays a role inheterogeneous ice nucleation. *J. Chem. Phys*., **161**, 094304, https://doi.org/10.1063/5.0213171.

Blossey, P. N., and Coauthors, 2013: Marine low cloud sensitivity to an idealized climate change: The CGILS LES intercomparison. *J. Adv. Model Earth Sy*., **5**, 234–258, https://doi.org/10.1002/jame.20025.

Bodas-Salcedo, A., and Coauthors, 2011: COSP: Satellite simulation software for model assessment. *Bull. Amer. Meteor. Soc*., **92**, 1023–1043, https://doi.org/10.1175/2011BAMS2856.1.

Bodas-Salcedo, A., J. P. Mulcahy, T. Andrews, K. D. Williams, M. A. Ringer, P. R. Field, and G. S. Elsaesser, 2019: Strong dependence of atmospheric feedbacks on mixed-phase micro physics and aerosol-cloud interactions in HadGEM3. *J. Adv. Model. Earth Syst*., **11**, 1735–1758, https://doi.org/10.1029/2019MS001688.

Bony, S., and Coauthors, 2013: Carbon dioxide and climate: Perspectives on a scientific assessment. *Climate Science for Serving Society*, G. Asrar and J. Hurrell, Eds, Springer Netherlands, 391–413, https://doi.org/10.1007/978-94-007-6692-1_14.

Boucher, O., and U. Lohmann, 1995: The sulfate-CCN-cloud albedo effect. *Tellus B*, **47**, 281–300, https://doi.org/10.1034/j.1600-0889.47.issue3.1.x.

Bretherton, C. S., R. Wood, R. C. George, D. Leon, G. Allen, and X. Zheng, 2010: Southeast Pacific stratocumulus clouds, precipitation and boundary layer structure sampled along 20° S during VOCALS-REx. *Atmos. Chem. Phys*., **10**, 10639–10654, https://doi.org/10.5194/acp-10-10639-2010.

Burrows, S. M., and Coauthors, 2022: Ice-nucleating particles that impact clouds and climate: Observational and modeling research needs. *Rev. Geophys*., **60**, e2021RG000745, https://doi.org/10.1029/2021RG000745.

Byers, H. R., and R. K. Hall, 1955: A census of cumulus-cloud height versus precipitation in the vicinity of Puerto Rico during the Winter and Spring of 1953-1954. *J. Atmos. Sci.*, **12**, 176–178, https://doi.org/10.1175/1520-0469(1955)012<0176:ACOCCH>2.0.CO;2.

Carslaw, K. S., O. Boucher, D. V. Spracklen, G. W. Mann, J. G. L. Rae, S. Woodward, and M. Kulmala, 2010: A review of natural aerosol interactions and feedbacks within the Earth system. *Atmos. Chem. Phys*., **10**, 1701–1737, https://doi.org/10.5194/acp-10-1701-2010.

Carslaw, K. S., and Coauthors, 2013: Large contribution of natural aerosols to uncertainty in indirect forcing. *Nature*, **503**, 67–71, https://doi.org/10.1038/nature12674.

Carslaw, K. S., and Coauthors, 2025: Opinion: The importance and future development of perturbed parameter ensembles in climate and atmospheric science. *EGUsphere* [preprint], https://doi.org/10.5194/egusphere-2025-4341.

Christensen, M. W., and Coauthors, 2022: Opportunistic experiments to constrain aerosol effective radiative forcing. *Atmos. Chem. Phys*., **22**, 641–674, https://doi.org/10.5194/acp-22-641-2022.

Cucurull, L., R. A. Anthes, S. P. F. Casey, M. J. Mueller, and A. Vidal, 2024: Observing System Simulation Experiments (OSSEs) in support of next-generation NOAA satellite constellation. *Bull. Amer. Meteor. Soc*., **105**, E884–E904, https://doi.org/10.1175/BAMS-D-23-0060.1.

Curtis, J. H., N. Riemer, and M. West, 2017: A single-column particle-resolved model for simulating the vertical distribution of aerosol mixing state: WRF-PartMC-MOSAIC-SCM v1.0. *Geosci. Model Dev*., **10**, 4057–4079, https://doi.org/10.5194/gmd-10-4057-2017.

DeMott, P. J., and Coauthors, 2018: The Fifth International Workshop on Ice Nucleation phase 2 (FIN-02): Laboratory intercomparison of ice nucleation measurements. *Atmos. Meas. Tech*., **11**, 6231–6257, https://doi.org/10.5194/amt-11-6231-2018.

Diamond, M. S., H. M. Director, R. Eastman, A. Possner, and R. Wood, 2020: Substantial cloud brightening from shipping in subtropical low clouds. *AGU Adv.*, **1**, e2019AV000111, https://doi.org/10.1029/2019AV000111.

Donahue, A. S., and Coauthors, 2024: To exascale and beyond—The Simple Cloud-Resolving E3SM Atmosphere Model (SCREAM), a performance portable global atmosphere model for cloud-resolving scales. *J. Adv. Model. Earth Syst.*, **16**, e2024MS004314, https://doi.org/10.22541/essoar.171136746.65715825/v1.

Eidhammer, T., and Coauthors, 2024: An extensible perturbed parameter ensemble for the Community Atmosphere Model version 6. *Geosci. Model Dev.*, **17**, 7835–7853, https://doi.org/10.5194/gmd-17-7835-2024.

Elsaesser, G. S., and Coauthors, 2025: Using machine learning to generate a GISS ModelE Calibrated Physics Ensemble (CPE). *J. Adv. Model. Earth Syst.*, **17**, e2024MS004713, https://doi.org/10.1029/2024MS004713.

Erfani, E., R. Wood, P. Blossey, S. J. Doherty, and R. Eastman, 2025: Building a comprehensive library of observed Lagrangian trajectories for testing modeled cloud evolution, aerosol-cloud interactions, and marine cloud brightening. *Atmos. Chem. Phys.*, **25**, 8743–8768, https://doi.org/10.5194/acp-25-8743-2025.

Eyring, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor, 2016: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.*, **9**, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016.

Feingold, G., and H. Siebert, 2009: Cloud-aerosol interactions from the micro to the cloud scale. *Clouds in the Perturbed Climate System: Their Relationship to Energy Balance, Atmospheric Dynamics, and Precipitation*, Strüngmann Forum Report, J. Heintzenberg and R. J. Charlson, Eds, MIT Press, 319–338, https://csl.noaa.gov/staff/graham.feingold/graham.feingold.presentations/14Feingold&Siebert.pdf.

Feingold, G., F. Glassmeier, J. Zhang, and F. Hoffmann, 2025: Opinion: Inferring process from snapshots of cloud systems. *Atmos. Chem. Phys.*, **25**, 1–28, https://doi.org/10.5194/acp-25-10869-2025.

Field, P. R., and K. Furtado, 2016: How biased is aircraft cloud sampling? *J. Atmos. Oceanic Technol.*, **33**, 185–189, https://doi.org/10.1175/JTECH-D-15-0148.1.

Fierce, L., Y. Yao, R. Easter, P.-L. Ma, J. Sun, H. Wan, and K. Zhang, 2024: Quantifying structural errors in cloud condensation nuclei activity from reduced representation of aerosol size distributions. *J. Aerosol Sci.*, **181**, 106388, https://doi.org/10.1016/j.jaerosci.2024.106388.

French, J. R., and Coauthors, 2018: Precipitation formation from orographic cloud seeding. *Proc. Natl. Acad. Sci.*, **115**, 1168–1173, https://doi.org/10.1073/pnas.1716995115.

Gao, M., and Coauthors, 2023: Simultaneous retrieval of aerosol and ocean properties from PACE HARP2 with uncertainty assessment using cascading neural network radiative transfer models. *Atmos. Meas. Tech.*, **16**, 5863–5881, https://doi.org/10.5194/amt-16-5863-2023.

Gettelman, A., and H. Morrison, 2015: Advanced two-moment bulk microphysics for global models. Part I: Off-line tests and comparison with other schemes. *J. Climate*, **28**, 1268–1287, https://doi.org/10.1175/JCLI-D-14-00102.1.

Gettelman, A., and Coauthors, 2020: Simulating observations of Southern Ocean clouds and implications for climate. *J. Geophys. Res. Atmos.*, **125**, e2020JD032619, https://doi.org/10.1029/2020JD032619.

Gettelman, A., T. Eidhammer, M. L. Duffy, D. T. McCoy, C. Song, and D. Watson-Parris, 2024: The interaction between climate forcing and feedbacks. *J. Geophys. Res.: Atmos.*, **129**, e2024JD040857, https://doi.org/10.1029/2024JD040857.

Gettelman, A., C. Tebaldi, and L. R. Leung, 2025: Climate nowcasting. *Environ. Res.: Climate*, **4**, 013002, https://doi.org/10.1088/2752-5295/adc327.

Gleckler, P. J., K. E. Taylor, and C. Doutriaux, 2008: Performance metrics for climate models. *J. Geophys. Res.: Atmos*., **113**, https://doi.org/10.1029/2007JD008972.

Greenwald, T. J., S. A. Christopher, and J. Chou, 1997: Cloud liquid water path comparisons from passive microwave and solar reflectance satellite measurements: Assessment of sub-field-of-view cloud effects in microwave retrievals. *J. Geophys. Res.: Atmos*., **102**, 19585–19596, https://doi.org/10.1029/97JD01257.

Grosvenor, D. P., and R. Wood, 2014: The effect of solar zenith angle on MODIS cloud optical and microphysical retrievals within marine liquid water clouds. *Atmos. Chem. Phys*., **14**, 7291–7321, https://doi.org/10.5194/acp-14-7291-2014.

Grosvenor, D. P., and K. S. Carslaw, 2023: Change from aerosol-driven to cloud-feedback-driven trend in short-wave radiative flux over the North Atlantic. *Atmos. Chem. Phys*., **23**, 6743–6773, https://doi.org/10.5194/acp-23-6743-2023.

Grosvenor, D. P., and Coauthors, 2018: Remote sensing of droplet number concentration in warm clouds: A review of the current state of knowledge and perspectives. *Rev. Geophys*., **56**, 409–453, https://doi.org/10.1029/2017rg000593.

Gryspeerdt, E., T. W. P. Smith, E. O'Keeffe, M. W. Christensen, and F. W. Goldsworth, 2019: The impact of ship emission controls recorded by cloud properties. *Geophys. Res. Lett*., **46**, 12547– 12555, https://doi.org/10.1029/2019GL084700.

Gryspeerdt, E., and Coauthors, 2022a: The impact of sampling strategy on the cloud droplet number concentration estimated from satellite data. *Atmos. Meas. Tech*., **15**, 3875–3892, https://doi.org/10.5194/amt-15-3875-2022.

Gryspeerdt, E., and Coauthors, 2022b: The impact of sampling strategy on the cloud droplet number concentration estimated from satellite data. *Atmos. Meas. Tech*., **15**, 3875–3892, https://doi.org/10.5194/amt-15-3875-2022.

Haarsma, R. J., and Coauthors, 2016: High resolution model intercomparison project (HighResMIP v1.0) for CMIP6. *Geosci. Model Dev*., **9**, 4185–4208, https://doi.org/10.5194/gmd-9-4185-2016.

Harder, P., D. Watson-Parris, P. Stier, D. Strassel, N. R. Gauger, and J. Keuper, 2022: Physics-informed learning of aerosol microphysics. *Environ. Data Sci*., **1**, e20, https://doi.org/10.1017/eds.2022.22.

Henneberger, J., and Coauthors, 2023: Seeding of supercooled low stratus clouds with a UAV to study microphysical ice processes: An introduction to the CLOUDLAB project. *Bull. Amer. Meteor.l Soc*., **104**, E1962–E1979, https://doi.org/10.1175/BAMS-D-22-0178.1.

Hernandez-Jaramillo, D. C., and Coauthors, 2025: First generation outdoor marine cloud brightening trial increases aerosol concentration at cloud base height. *Environ. Res. Lett.*, **20**, https://doi.org/10.1088/1748-9326/adccd7.

Hoffman, R. N., and R. Atlas, 2016: Future Observing System Simulation Experiments. *Bull. Amer. Meteor. Soc*., **97**, 1601–1616, https://doi.org/10.1175/BAMS-D-15-00200.1.

Hope, C., 2015: The $10 trillion value of better information about the transient climate response. *Philos. Trans. R. Soc., A*, **373**, 20140429, https://doi.org/10.1098/rsta.2014.0429.

Hourdin, F., and Coauthors, 2017: The art and science of climate model tuning. *Bull. Amer. Meteor. Soc*., **98**, 589–602, https://doi.org/10.1175/BAMS-D-15-00135.1.

Hu, A. Z., and A. L. Igel, 2024: Bulk microphysics schemes may perform better with a unified cloud-rain category. *J. Adv. Model. Earth Syst*., **16**, e2023MS004068, https://doi.org/10.1029/2023MS004068.

Hu, Y., and Coauthors, 2021: Liquid phase cloud microphysical property estimates from CALIPSO measurements. *Front. Remote Sens*., **2**, 25, https://doi.org/10.3389/frsen.2021.724615.

Hume, D., 1751: *Philosophical Essays Concerning Human Understanding*, 2nd ed., with additions and corrections. Printed for M. Cooper, 259 pp.

Igel, A. L., H. Morrison, S. P. Santos, and M. van Lier-Walqui, 2022: Limitations of separate cloud and rain categories in parameterizing collision-coalescence for bulk microphysics schemes. *J. Adv. Model. Earth Syst*., **14**, e2022MS003039, https://doi.org/10.1029/2022MS003039.

Iwabuchi, H., S. Yamada, S. Katagiri, P. Yang, and H. Okamoto, 2014: Radiative and microphysical properties of cirrus cloud inferred from infrared measurements made by the Moderate Resolution Imaging Spectroradiometer (MODIS). Part I: Retrieval method. *J. Appl. Meteor. Climatol*., **53**, 1297–1316, https://doi.org/10.1175/JAMC-D-13-0215.1.

Jahn, L. G., M. J. Polen, L. G. Jahl, T. A. Brubaker, J. Somers, and R. C. Sullivan, 2020: Biomass com bustion produces ice-active minerals in biomass-burning aerosol and bottom ash. *Proc. Natl. Acad. Sci*., **117**, 21928–21937, https://doi.org/10.1073/pnas.1922128117.

Johnson, J. S., and Coauthors, 2020: Robust observational constraint of uncertain aerosol processes and emissions in a climate model and the effect on aerosol radiative forcing. *Atmos. Chem. Phys*., **20**, 9491–9524, https://doi.org/10.5194/acp-20-9491-2020.

Jones, D., and Coauthors, 2025: Surface observations from atmospheric radiation measurement sites constrain the anthropogenic contribution to cloud droplet number. *ESS Open Archive,* https://doi.org/10.22541/essoar.175087280.06595777/v1.

Kahn, R. A., and Coauthors, 2017: SAM-CAAM: a concept for acquiring Systematic Aircraft Measurements to Characterize Aerosol Air Masses. *Bull. Amer. Meteor. Soc*., **98**, 2215–2228, https://doi.org/10.1175/BAMS-D-16-0003.1.

Kang, L., R. T. Marchand, R. Wood, and I. L. McCoy, 2022: Coalescence scavenging drives droplet number concentration in Southern Ocean low clouds. *Geophys. Res. Lett*., **49**, e2022GL097819, https://doi.org/10.1029/2022GL097819.

Khairoutdinov, M., and Y. Kogan, 2000: A new cloud physics parameterization in a large-eddy simulation model of marine stratocumulus. *Mon. Weather Rev*., **128**, 229–243, https://doi.org/10.1175/1520-0493(2000)128%3C0229:ANCPPI%3E2.0.CO;2.

King, M. D., and Coauthors, 2003: Cloud and aerosol properties, precipitable water, and profiles of temperature and water vapor from MODIS. *Geosci. Remote Sens*., IEEE Transactions on, **41**, 442–458, https://doi.org/10.1109/TGRS.2002.808226.

Knopf, D. A., and Coauthors, 2021: Aerosol–ice formation closure: A southern Great Plains field campaign. *Bull. Amer. Meteor. Soc*., **102**, E1952–E1971, https://doi.org/10.1175/BAMS-D-20-0151.1.

Knutti, R., R. Furrer, C. Tebaldi, J. Cermak, and G. A. Meehl, 2010: Challenges in combining projections from multiple climate models. *J. Climate*, **23**, 2739–2758, https://doi.org/10.1175/2009JCLI3361.1.

Korolev, A., and T. Leisner, 2020: Review of experimental studies of secondary ice production. *Atmos. Chem. Phys*., **20**, 11767–11797, https://doi.org/10.5194/acp-20-11767-2020.

Korolev, A ., P. J. DeMott, I. Heckman, M. Wolde, E. Williams, D. J. Smalley, and M. F. Donovan, 2022: Observation of secondary ice production in clouds at low temperatures. *Atmos. Chem.Phys*., **22**, 13103–13113, https://doi.org/10.5194/acp-22-13103-2022.

Kurri, M., and A. Huuskonen, 2008: Measurements of the transmission loss of a radome at different rain intensities. *J. Atmos. Ocean. Technol*., **25**, 1590–1599, https://doi.org/10.1175/2008JTECHA1056.1.

Lance, S., C. A. Brock, D. Rogers, and J. A. Gordon, 2010: Water droplet calibration of the Cloud Droplet Probe (CDP) and in-flight performance in liquid, ice and mixed-phase clouds during ARCPAC. *Atmos. Meas. Tech*., **3**, 1683–1706, https://doi.org/10.5194/amt-3-1683-2010.

Latham, J., J. Kleypas, R. Hauser, B. Parkes, and A. Gadian, 2013: Can marine cloud brightening reduce coral bleaching? *Atmos. Sci. Lett*., **14**, 214–219, https://doi.org/10.1002%2Fasl2.442.

Lee, H.-H., P. Bogenschutz, and T. Yamaguchi, 2022: Resolving away stratocumulus bias es in modern global climate models. *Geophys. Res. Lett*., **49**, e2022GL099422, https://doi.org/10.1029/2022GL099422.

Lee, L. A., K. S. Carslaw, K. J. Pringle, G. W. Mann, and D. V. Spracklen, 2011: Emulation of a complex global aerosol model to quantify sensitivity to uncertain parameters. *Atmos. Chem. Phys*., **11**, 12253–12273, https://doi.org/10.5194/acp-11-12253-2011.

Lee, L. A., C. L. Reddington, and K. S. Carslaw, 2016: On the relationship between aerosol model uncertainty and radiative forcing uncertainty. *Proceed. Natl. Acad. Sci*., **113**, 5820–5827, https://doi.org/10.1073/pnas.1507050113.

Lehner, F., C. Deser, N. Maher, J. Marotzke, E. M. Fischer, L. Brunner, R. Knutti, and E. Hawkins, 2020: Partitioning climate projection uncertainty with multiple large ensembles and CMIP5/6. *Earth Syst. Dynam*., **11**, 491–508, https://doi.org/10.5194/esd-11-491-2020.

Leinonen, J., M. D. Lebsock, G. L. Stephens, and K. Suzuki, 2016: Improved retrieval of cloud liquid water from CloudSat and MODIS. *J. Appl. Meteor. Climatol*., **55**, 1831–1844, https://doi.org/10.1175/JAMC-D-16-0077.1.

Liu, Y., and P. Kollias, 2023: Progress in understanding and parameterizing fast physics in large-scale atmospheric models. *Fast Processes in Large-Scale Atmospheric Models*, Geophysical Monograph Series, 1–10, https://doi.org/10.1002/9781119529019.ch1.

Mace, G. G., and A. C. Abernathy, 2016: Observational evidence for aerosol invigoration in shallow cumulus downstream of Mount Kilauea. *Geophys. Res. Lett*., **43**, 2981–2988, https://doi.org/10.1002/2016GL067830.

Malavelle, F. F., and Coauthors, 2017: Strong constraints on aerosol–cloud interactions from volcanic eruptions. *Nature*, **546**, 485–491, https://doi.org/10.1038/nature22974.

Malavelle, F. F., 2025: Supplemental material for "Strong constraints on aerosol–cloud interactions from volcanic eruptions." *Nature*, https://static-content.springer.com/esm/art%3A10.1038%2Fnature22974/MediaObjects/41586_2017_BFnature22974_MOESM1_ESM.pdf.

Manabe, S., and R. T. Wetherald, 1967: Thermal equilibrium of the atmosphere with a given distribution of relative humidity. *J. Atmos. Sci*., **24**, 241–259, https://doi.org/10.1175/1520-0469(1967)024<0241:TEOTAW>2.0.CO;2.

Masunaga, H., and Coauthors, 2010: Satellite data simulator unit: A multisensor, multi spectral satellite simulator package. *Bull. Amer. Meteorol. Soc*., **91**, 1625–1632, https://doi.org/10.1175/2010BAMS2809.1.

Mauritsen, T., and Coauthors, 2012: Tuning the climate of a global model. *J. Adv. Model Earth Sy.*, **4**, https://doi.org/10.1029/2012MS000154.

McComiskey, A., G. Feingold, A. S. Frisch, D. D. Turner, M. A. Miller, J. C. Chiu, Q. Min, and J. A. Ogren, 2009: An assessment of aerosol-cloud interactions in marine stratus clouds based on surface remote sensing. *J. Geophys. Res.: Atmos.*, **114**, https://doi.org/10.1029/2008JD011006.

McCoy, D. T., and D. L. Hartmann, 2015: Observations of a substantial cloud-aerosol indirect effect during the 2014-2015 Bárðarbunga-Veiðivötn fissure eruption in Iceland. *Geophys. Res. Lett.*, **42**, https://doi.org/10.1002/2015gl067070.

McCoy, D. T., F. A. M. Bender, D. P. Grosvenor, J. K. Mohrmann, D. L. Hartmann, R. Wood, and P. R. Field, 2018: Predicting decadal trends in cloud droplet number concentration using reanalysis and satellite data. *Atmos. Chem. Phys.*, **18**, 2035–2047, https://doi.org/10.5194/acp-18-2035-2018.

McCoy, D. T., P. Field, H. Gordon, G. S. Elsaesser, and D. P. Grosvenor, 2020: Untangling causality in midlatitude aerosol–cloud adjustments. *Atmos. Chem. Phys.*, **20**, 4085–4103, https://doi.org/10.5194/acp-20-4085-2020.

McCoy, I. L., and Coauthors, 2020: The hemispheric contrast in cloud microphysical properties constrains aerosol forcing. *Proceed. Natl. Acad. Sci.*, 201922502, https://doi.org/10.1073/pnas.1922502117.

McCoy, I. L., D. T. McCoy, R. Wood, P. Zuidema, and F. A.-M. Bender, 2023: The role of mesoscale cloud morphology in the shortwave cloud feedback. *Geophys. Res. Lett.*, **50**, e2022GL101042, https://doi.org/10.1029/2022GL101042.

McCoy, I. L., M. C. Wyant, P. N. Blossey, C. S. Bretherton, and R. Wood, 2024: Aitken mode aerosols buffer decoupled mid-latitude boundary layer clouds against precipitation depletion. *J. Geophys. Res.: Atmos.*, **129**, e2023JD039572, https://doi.org/10.1029/2023JD039572.

McDonnell, A., A. M. Bauer, and C. Proistosescu, 2024: To what extent does discounting 'Hot' climate models improve the predictive skill of climate model ensembles? *Earth's Future*, **12**, e2024EF004844, https://doi.org/10.1029/2024EF004844.

McFiggans, G., and Coauthors, 2006: The effect of physical and chemical aerosol properties on warm cloud droplet activation. *Atmos. Chem. Phys.*, **6**, 2593–2649, https://doi.org/10.5194/acp-6-2593-2006.

Mechoso, C. R., and Coauthors, 2014: Ocean–Cloud–Atmosphere–Land interactions in the Southeastern Pacific: The VOCALS Program. *Bull. Amer. Meteor. Soc.*, **95**, 357–375, https://doi.org/10.1175/BAMS-D-11-00246.1.

Mikkelsen, A., D. T. McCoy, T. Eidhammer, A. Gettelman, C. Song, H. Gordon, and I. L. McCoy, 2024: Constraining aerosol-cloud adjustments by uniting surface observations with a perturbed parameter ensemble. *Atmos. Chem. Phys.*, **25**, 4547–4570, https://doi.org/10.5194/acp-25-4547-2025.

Miller, A. J., K. P. Brennan, C. Mignani, J. Wieder, R. O. David, and N. Borduas-Dedekind, 2021: Development of the drop Freezing Ice Nuclei Counter (FINC), intercomparison of droplet freezing techniques, and use of soluble lignin as an atmospheric ice nucleation standard. *Atmos. Meas. Tech.*, **14**, 3131–3151, https://doi.org/10.5194/amt-14-3131-2021.

Morrison, H., and Coauthors, 2020: Confronting the challenge of modeling cloud and precipitation microphysics. *J. Advances in Model. Earth Syst.*, **12**, e2019MS001689, https://doi.org/10.1029/2019MS001689.

Mülmenstädt, J., and G. Feingold, 2018: The radiative forcing of aerosol–cloud interactions in liquid clouds: Wrestling and embracing uncertainty. *Curr. Clim. Change Rep*., **4**, 23–40, https://doi.org/10.1007/s40641-018-0089-y.

Myers, T. A., and J. R. Norris, 2013: Observational evidence that enhanced subsidence reduces subtropical marine boundary layer cloudiness. *J. Climate*, **26**, 7507–7524, https://doi.org/10.1175/JCLI-D-12-00736.1.

Myers, T. A., R. C. Scott, M. D. Zelinka, S. A. Klein, J. R. Norris, and P. M. Caldwell, 2021: Observational constraints on low cloud feedback reduce uncertainty of climate sensitivity. *Nat. Clim. Change*, https://doi.org/10.1038/s41558-021-01039-0.

Nakajima, T., A. Higurashi, K. Kawamoto, and J. E. Penner, 2001: A possible correlation between satellite-derived cloud and aerosol microphysical parameters. *Geophys. Res. Lett*., **28**, 1171–1174, https://doi.org/10.1029/2000GL012186.

Norris, J. R., R. J. Allen, A. T. Evan, M. D. Zelinka, C. W. O'Dell, and S. A. Klein, 2016: Evidence for climate change in the satellite cloud record. *Nature*, **536**, 72–75, https://doi.org/10.1038/nature18273.

Nugent, J. M., 2024: PROCEED-ESM/PROCEED-PPE: Version 0 (v.0.1.0). Zenodo, https://doi.org/10.5281/zenodo.13887194.

Nugent, J. M., and Coauthors, 2025: Overview of the Nephele Perturbed Parameter Ensemble for Aerosol-Cloud Interactions Hosted in E3SMv3. *J. Adv. Model. Earth Syst*. https://doi.org/10.22541/essoar.174907165.57104591/v1.

Parsons, L. A., D. E. Amrhein, S. C. Sanchez, R. Tardif, M. K. Brennan, and G. J. Hakim, 2021: Do multi-model ensembles improve reconstruction skill in paleoclimate data assimilation? *Earth Space Sci*., **8**, e2020EA001467, https://doi.org/10.1029/2020EA001467.

Pawlowska, H., and J.-L. Brenguier, 2003: An observational study of drizzle formation in stratocumulus clouds for general circulation model (GCM) parameterizations. *J. Geophy. Res.: Atmos*., **108**, https://doi.org/10.1029/2002JD002679.

Perkins, W. A., N. D. Brenowitz, C. S. Bretherton, and J. M. Nugent, 2024: Emulation of cloud microphysics in a climate model. *J. Adv. Model. Earth Syst*., **16**, e2023MS003851, https://doi.org/10.1029/2023MS003851.

Polen, M., E. Lawlis, and R. C. Sullivan, 2016: The unstable ice nucleation properties of Snomax® bacterial particles. *J. Geophys. Res.: Atmos*., **121**, https://doi.org/10.1002/2016JD025251.

Proske, U., S. Ferrachat, D. Neubauer, M. Staab, and U. Lohmann, 2022: Assessing the potential for simplification in global climate model cloud microphysics. *Atmos. Chem. Phys*., **22**, 4737–4762, https://doi.org/10.5194/acp-22-4737-2022.

Qian, Y., and Coauthors, 2018: Parametric sensitivity and uncertainty quantification in the Version 1 of E3SM atmosphere model based on short perturbed parameter ensemble simulations. *J. Geophys. Res.: Atmos*., **123**, https://doi.org/10.1029/2018JD028927.

Quaas, J., and Coauthors, 2020: Constraining the Twomey effect from satellite observations: issues and perspectives. *Atmos. Chem. Phys*., **20**, 15079–15099, https://doi.org/10.5194/acp-20-15079-2020.

Redemann, J., and Coauthors, 2021: An overview of the ORACLES (ObseRvations of Aerosols above CLouds and their intEractionS) project: Aerosol–cloud–radiation interactions in the southeast Atlantic basin. *Atmos. Chem. Phys*., **21**, 1507–1563, https://doi.org/10.5194/acp-21-1507-2021.

Regayre, L. A., and Coauthors, 2014: Uncertainty in the magnitude of aerosol-cloud radiative forcing over recent decades. *Geophys. Res. Lett*., **41**, 9040–9049, https://doi.org/10.1002/2014GL062029.

Regayre, L. A., and Coauthors, 2018: Aerosol and physical atmosphere model parameters are both important sources of uncertainty in aerosol ERF. *Atmos. Chem. Phys*., **18**, 9975–10006, https://doi.org/10.5194/acp-18-9975-2018.

Regayre, L. A., and Coauthors, 2020: The value of remote marine aerosol measurements for constraining radiative forcing uncertainty. *Atmos. Chem. Phys*., **20**, 10063–10072, https://doi.org/10.5194/acp-20-10063-2020.

Regayre, L. A., and Coauthors, 2023: Identifying climate model structural inconsistencies allows for tight constraint of aerosol radiative forcing. *Atmos. Chem. Phys*., **23**, 8749–8768, https://doi.org/10.5194/acp-23-8749-2023.

Rougier, J., D. M. H. Sexton, J. M. Murphy, and D. Stainforth, 2009: Analyzing the climate sensitivity of the HadSM3 climate model using ensembles from different but related experiments. *J. Climate*, **22**, 3540–3557, https://doi.org/10.1175/2008JCLI2533.1.

Russel, L. M., and Coauthors, 2013: Eastern Pacific Emitted Aerosol Cloud Experiment. *Bull. Amer. Meteor. Soc.*, 94, 709–729, https://doi.org/10.1175/BAMS-D-12-00015.1

Saito, M., P. Yang, Y. Hu, X. Liu, N. Loeb, W. L. Smith Jr, and P. Minnis, 2019: An efficient method for microphysical property retrievals in vertically inhomogeneous marine water clouds using MODIS-CloudSat measurements. *J. Geophys. Res.: Atmos*., **124**, 2174–2193, https://doi.org/10.1029/2018JD029659.

Saito, M., P. Yang, X. Huang, H. E. Brindley, M. G. Mlynczak, and B. H. Kahn, 2020: Spaceborne middle-and far-infrared observations improving nighttime ice cloud property retrievals. *Geophys. Res. Lett*., **47**, e2020GL087491, https://doi.org/10.1029/2020GL087491.

Sansom, R. W. N., K. S. Carslaw, J. S. Johnson, and L. Lee, 2024: An emulator of stratocumulus cloud response to two cloud-controlling factors accounting for internal variability. *J. Adv. Model. Earth Syst*., **16**, e2023MS004179, https://doi.org/10.1029/2023MS004179.

Schutgens, N., S. Tsyro, E. Gryspeerdt, D. Goto, N. Weigum, M. Schulz, and P. Stier, 2017: On the spatio-temporal representativeness of observations. *Atmos. Chem. Phys*., **17**, 9761–9780, https://doi.org/10.5194/acp-17-9761-2017.

Schutgens, N. A. J., E. Gryspeerdt, N. Weigum, S. Tsyro, D. Goto, M. Schulz, and P. Stier, 2016: Will a perfect model agree with perfect observations? The impact of spatial sampling. *Atmos. Chem. Phys*., **16**, 6335–6353, https://doi.org/10.5194/acp-16-6335-2016.

Seinfeld, J. H., and S. N. Pandis, 2016: *Atmospheric chemistry and physics: from air pollution to climate change*, 3rd ed. John Wiley & Sons, 1152 pp.

Sexton, D. M. H., J. M. Murphy, M. Collins, and M. J. Webb, 2012: Multivariate probabilistic projections using imperfect climate models part I: outline of methodology. *Clim. Dyn*., **38**, 2513–2542, https://doi.org/10.1007/s00382-011-1208-9.

Sexton, D. M. H., and Coauthors, 2021: A perturbed parameter ensemble of HadGEM3-GC3.05 coupled model projections: Part 1: Selecting the parameter combinations. *Clim. Dyn*., **56**, 3395–3436, https://doi.org/10.1007/s00382-021-05709-9.

Sherwood, S. C., and Coauthors, 2020: An assessment of Earth's climate sensitivity using multiple lines of evidence. *Rev. Geophys*., **58**, e2019RG000678, https://doi.org/10.1029/2019RG000678.

Shiogama, H., M. Watanabe, T. Ogura, T. Yokohata, and M. Kimoto, 2014: Multi-parameter multi-physics ensemble (MPMPE): A new approach exploring the uncertainties of climate sensitivity. *Atmos. Sci. Lett*., **15**, 97–102, https://doi.org/10.1002/asl2.472.
Silber, I., R. C. Jackson, A. M. Fridlind, A. S. Ackerman, S. Collis, J. Verlinde, and J. Ding, 2022: The Earth model column collaboratory (EMC2) v1.1: An open-source ground-based lidar and radar instrument simulator and subcolumn generator for large-scale models. *Geosci. Model Dev*., **15**, 901–927, https://doi.org/10.5194/gmd-15-901-2022.
Soden, B. J., and I. M. Held, 2006: An assessment of climate feedbacks in coupled ocean-atmsphere models. *J. Climat*e, **19**, 3354–3360, https://doi.org/10.1175/jcli3799.1.
Song, C., and Coauthors, 2024: Buffering of aerosol-cloud adjustments by coupling between radiative susceptibility and precipitation efficiency. *Geophys. Res. Lett.*, **51**, e2024GL108663, https://doi.org/10.1029/2024GL108663.
Song, C., D. T. McCoy, I. L. McCoy, H. Brown, A. Gettelman, T. Eidhammer, and D. Barahona, 2025: Aircraft in-situ measurements from SOCRATES constrain the anthropogenic perturbations of cloud droplet number. *EGUsphere* [preprint], https://doi.org/10.5194/egusphere-2025-2009.
Stephens, G. L., and Coauthors, 2002: The CloudSat mission and the A-Train — A new dimension of space-based observations of clouds and precipitation. *Bull. Amer. Meteorol. Soc*., **83**, 1771–1790, https://doi.org/10.1175/bams-83-12-1771.
Stevens, B., and G. Feingold, 2009: Untangling aerosol effects on clouds and precipitation in a buffered system. *Nature*, **461**, 607–613, https://doi.org/10.1038/nature08281.
Sullivan, S. C., C. Hoose, and A. Nenes, 2017: Investigating the contribution of secondary ice production to in-cloud ice crystal numbers. *J. Geophys. Res.: Atmos*., **122**, 9391–9412, https://doi.org/10.1002/2017JD026546.
Swales, D. J., R. Pincus, and A. Bodas-Salcedo, 2018: The Cloud Feedback Model Intercomparison Project Observational Simulator Package: Version 2. *Geosci. Model Dev*., **11**, 77–81, https://doi.org/10.5194/gmd-11-77-2018.
Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2013: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteorol. Soc*., **93**, 485–498, https://doi.org/10.1175/bams-d-11-00094.1.
Terai, C. R., M. S. Pritchard, P. Blossey, and C. S. Bretherton, 2020: The impact of resolving subkilometer processes on aerosol-cloud interactions of low-level clouds in global model simulations. *J. Adv. Model. Earth Sys*., **12**, e2020MS002274, https://doi.org/10.1029/2020MS002274.
Tessendorf, S. A., and Coauthors, 2019: A transformational approach to winter orographic weather modification research: The SNOWIE project. *Bull. Amer. Meteor. Soc*., **100**, 71–92, https://doi.org/10.1175/BAMS-D-17-0152.1.
Toll, V., M. Christensen, J. Quaas, and N. Bellouin, 2019: Weak average liquid-cloud-water response to anthropogenic aerosols. *Nature*, **572**, 51–55, https://doi.org/10.1038/s41586-019-1423-9.
Tsushima, Y., M. A. Ringer, G. M. Martin, J. W. Rostron, and D. M. H. Sexton, 2020: Investigating physical constraints on climate feedbacks using a perturbed parameter ensemble. *Clim. Dynam*., **55**, https://doi.org/10.1007/s00382-020-05318-y.
Twomey, S., 1977: Influence of pollution on shortwave albedo of clouds. *J. Atmos. Sci*., **34**, 1149–1152, https://doi.org/10.1175/1520-0469(1977)034<1149:tiopot>2.0.co;2.
Vonnegut, B.,1947: The nucleation of ice formation by silver iodide. *J. Appl. Phys*., **18**, 593–595.

Vonnegut, B., 1958: Early work on silver iodide smokes for cloud seeding. Final Rept., *Advisory Committee on Weather Control*, **2**, 283–285.

Wall, C. J., J. R. Norris, A. Possner, D. T. McCoy, I. L. McCoy, and N. J. Lutsko, 2022: Assessing effective radiative forcing from aerosol–cloud interactions over the global ocean. *Proceed. Natl.Acad. Sci*., **119**, e2210481119, https://doi.org/10.1073/pnas.2210481119.

Wang, A., M. Ovchinnikov, F. Yang, S. Schmalfuss, and R. A. Shaw, 2024: Designing a convection-cloud chamber for collision-coalescence using large-eddy simulation with bin microphysics. *J. Adv. Model. Earth Syst*., **16**, https://doi.org/10.1029/2023MS003734.

Wang, M., S. Ghan, M. Ovchinnikov, X. Liu, R. Easter, E. Kassianov, Y. Qian, and H. Morrison, 2011: Aerosol indirect effects in a multi-scale aerosol-climate model PNNL-MMF. *Atmos. Chem. Phys*., **11**, 5431–5455, https://doi.org/10.5194/acp-11-5431-2011.

Watson-Parris, D., 2025: Integrating top-down energetic constraints with bottom-up process-based constraints for more accurate projections of future warming. *Geophys. Res. Lett*., **52**, https://doi.org/10.1029/2024GL114269.

Watson-Parris, D., and C. J. Smith, 2022: Large uncertainty in future warming due to aerosol forcing. *Nature Climate Change*, **12**, https://doi.org/10.1038/s41558-022-01516-0.

Watson-Parris, D., A. Williams, L. Deaconu, and P. Stier, 2021: Model calibration using ESEm v1.1.0 – an open, scalable Earth system emulator. *Geosci. Model Dev*., **14**, 7659–7672, https://doi.org/10.5194/gmd-14-7659-2021.

Wofsy, S. C., 2011: HIAPER Pole-to-Pole Observations (HIPPO): fine-grained, global-scale measurements of climatically important atmospheric gases and aerosols. *Phil. Trans. R. Soc. A.*, **369**, 2073–2086, https://doi.org/10.1098/rsta.2010.0313.

Wood, R., 2006: Rate of loss of cloud droplets by coalescence in warm clouds. *J. Geophys. Res.: Atmos*, **111**, https://doi.org/10.1029/2006JD007553.

Wood, R., 2012: Stratocumulus clouds. *Mon. Weather Rev*., **140**, 2373–2423, https://doi.org/10.1175/MWR-D-11-00121.1.

Wood, R., and Coauthors, 2011: The VAMOS Ocean-Cloud-Atmosphere-Land Study Regional Experiment (VOCALS-REx): goals, platforms, and field operations. *Atmos. Chem. Phys*., **11**, 627–654, https://doi.org/10.5194/acp-11-627-2011.

Wood, R., D. Leon, M. Lebsock, J. Snider, and A. D. Clarke, 2012: Precipitation driving of droplet concentration variability in marine low clouds. *J. Geophys. Res.: Atmos*, **117**, https://doi.org/10.1029/2012jd018305.

Wyant, M. C., and Coauthors, 2010: The PreVOCA experiment: modeling the lower troposphere in the Southeast Pacific. *Atmos. Chem. Phys*., **10**, 4757–4774, https://doi.org/10.5194/acp-10-4757-2010.

Wyant, M. C, and Coauthors, 2015: Global and regional modeling of clouds and aerosols in the marine boundary layer during VOCALS: the VOCA intercomparison. *Atmos. Chem. Phys*., **15**, 153–172, https://doi.org/10.5194/acp-15-153-2015.

Y. J. Kaufman, and Coauthors, 2005: A critical examination of the residual cloud contamination and diurnal sampling effects on MODIS estimates of aerosol over ocean. *IEEE Trans. Geosci. Remote Sens.*, **43**, 2886–2897, https://doi.org/10.1109/TGRS.2005.858430.

Yamaguchi, T., G. Feingold, and V. E. Larson, 2017: Framework for improvement by vertical enhancement: A simple approach to improve representation of low and high-level clouds in large-scale models. *J. Adv. Model. Earth Sys*., **9**, 627–646, https://doi.org/10.1002/2016MS000815.

Yamazaki, K., D. M. H. Sexton, J. W. Rostron, C. F. McSweeney, J. M. Murphy, and G. R. Harris, 2021: A perturbed parameter ensemble of HadGEM3-GC3.05 coupled model projections: Part 2: global performance and future changes. *Clim. Dyn*., **56**, 3437–3471, https://doi.org/10.1007/s00382-020-05608-5.

Yarger, D., B. M. Wagman, K. Chowdhary, and L. Shand, 2024: Autocalibration of the E3SM version 2 atmosphere model using a PCA-based surrogate for spatial fields. *J. Adv. Model Earth Sys*., **16**, https://doi.org/10.1029/2023MS003961.

Yoshioka, M., and Coauthors, 2019: Ensembles of global climate model variants designed for the quantification and constraint of uncertainty in aerosols and their radiative forcing. *J. Adv. Model. Earth Syst.*, **11**, https://doi.org/10.1029/2019ms001628.

Zelinka, M. D., T. A. Myers, D. T. McCoy, S. Po-Chedley, P. M. Caldwell, P. Ceppi, S. A. Klein, and K. E. Taylor, 2020: Causes of higher climate sensitivity in CMIP6 models. *Geophys. Res. Lett*., **47**, https://doi.org/10.1029/2019GL085782.

Zeng, X., and Coauthors, 2020: Use of observing system simulation experiments in the United States. *Bull. Amer. Meteorol. Soc*., **101**, https://doi.org/10.1175/BAMS-D-19-0155.1.

Zhang, D., and Coauthors, 2023: Evaluation of four ground-based retrievals of cloud droplet number concentration in marine stratocumulus with aircraft in situ measurements. *Atmos. Meas. Tech*., **16**, 5827–5846, https://doi.org/10.5194/amt-16-5827-2023.

APPENDIX A: ORGANIZERS

Scientific Organizing Committee

Daniel McCoy, University of Wyoming (co-chair)
Rob Wood, University of Washington (co-chair)
Susannah M. Burrows, Pacific Northwest National Laboratory
Ann Fridlind, NASA Goddard Institute for Space Studies
Adele Igel, University of California Davis
Coty Jen, Carnegie Mellon University
Leighton Regayre, University of Leeds / UK Met Office
Masanori Saito, University of Wyoming
Duncan Watson-Parris, University of California San Diego

Program Organizing Committee

Mike Patterson, US CLIVAR
Alyssa Cannistraci, US CLIVAR
Jessica Martinez, UCAR CPAESS

APPENDIX B: PARTICIPANTS

Katie Ackerman	University of Hawai'i at Mānoa
Travis Aerenson	University of Wyoming
Olivia Alcabes	California Institute of Technology
Marybeth Arcodia	Colorado State University
Adriana Bailey	University of Michigan
Donifan Barahona	NASA Goddard Space Flight Center
Jason Barr	University of Washington
Darrel Baumgardner	Droplet Measurement Technologies
David Benson	NOAA Climate Program Office
Leah Bertrand	University of Colorado
Hassan Beydoun	Lawrence Livermore National Laboratory
Alan Blyth	National Centre for Atmospheric Science / University of Leeds
Corey Bois	University of Utah
Nadine Borduas-Dedekind	University of British Columbia
Hunter Brown	University of Wyoming
Jatan Buch	Columbia University
Susannah Burrows	Pacific Northwest National Laboratory
Alyssa Cannistraci	US CLIVAR
Ken Carslaw	University of Leeds
Yaosheng Chen	University of Colorado / Cooperative Institute for Research in Environmental Sciences / NOAA Chemical Sciences Laboratory
Nicole Chisholm	Dalhousie University
Christine Chiu	Colorado State University
Patrick Chuang	UC Santa Cruz
Jonathan Crosier	University of Manchester
Bithi De	Columbia University / NASA Goddard Institute for Space Studies
Emily de Jong	Lawrence Livermore National Laboratory
Michie De Vera	University of Illinois Urbana-Champaign
Kimberly Dill	Droplet Measurement Technologies
Aryeh Drager	Brookhaven National Laboratory
Tom Dror	University of Colorado / Cooperative Institute for Research in Environmental Sciences / NOAA Chemical Sciences Laboratory
Brandon Duran	UC San Diego / Scripps Institution of Oceanography
Trude Eidhammer	NSF National Center for Atmospheric Research
Gregory Elsaesser	NASA Goddard Institute for Space Studies / Columbia University
Ehsan Erfani	Desert Research Institute

Eshkol Eytan	University of Colorado / Cooperative Institute for Research in Environmental Sciences / NOAA Chemical Sciences Laboratory
Federico Fabiano	Consiglio Nazionale delle Ricerche Istituto di Scienze dell'Atmosfera e del Clima
Jiwen Fan	Argonne National Laboratory
Graham Feingold	NOAA Chemical Sciences Laboratory
Laura Fierce	Pacific Northwest National Laboratory
Declan Finney	University of Leeds
Luke Fraser-Leach	University of Toronto
Jeff French	University of Wyoming
Ann Fridlind	NASA Goddard Institute for Space Studies
Meng Gao	Science Systems and Applications Inc. / NASA Goddard Space Flight Center
Bart Geerts	University of Wyoming
Andrew Gettelman	Pacific Northwest National Laboratory
Hamish Gordon	Carnegie Mellon University
Daniel Grosvenor	UK Met Office / University of Leeds
Siddhant Gupta	Argonne National Laboratory
Walter Hannah	Lawrence Livermore National Laboratory
Ross Herbert	University of Leeds
Ute Herzfeld	University of Colorado
Fabian Hoffmann	Ludwig-Maximilians-Universität München
Arthur Z. Hu	Columbia University
Adele Igel	UC Davis
Robert Jacob	Argonne National Lab
Anna Jaruga	Climate Modeling Alliance / California Institute of Technology
Nadir Jeevanjee	NOAA Geophysical Fluid Dynamics Laboratory
Coty Jen	Carnegie Mellon University
Lihui Ji	University of Illinois Urbana-Champaign
Zhongjing Jiang	University of Illinois Urbana-Champaign
Timothy Juliano	NSF National Center for Atmospheric Research
Nicole June	Colorado State University
Litai Kang	University of Washington
Jan Kazil	University of Colorado / NOAA Chemical Sciences Laboratory
Alexei A. Kiselev	Karlsruher Institut für Technologie Institute für Meteorologie und Klimaforschung
Joseph Ko	Columbia University
Steve Krueger	University of Utah
Chongai Kuang	Brookhaven National Laboratory
Deepak Kumar	State University of New York at Albany
Luis A. Ladino	Universidad Nacional Autónoma de México
Kara D. Lamb	Columbia University
Zachary Lebo	University of Oklahoma

Matt Lebsock	NASA Jet Propulsion Laboratory
En Li	Colorado State University
Jung-Sub Lim	Ludwig-Maximilians-Universität München
Lin Lin	Lawrence Livermore National Laboratory
Yangang Liu	Brookhaven National Laboratory
Kaitlyn Loftus	Columbia University
Jesse Loveridge	Colorado State University
Amy Lu	California Institute of Technology
John Lundstrom	University of Illinois Urbana-Champaign
Salil Mahajan	Oak Ridge National Laboratory
Adam Majewski	University of Wyoming
Hitoshi Matsui	Nagoya University
Christina McCluskey	NSF National Center for Atmospheric Research
Daniel McCoy	University of Wyoming
Isabel McCoy	University of Colorado / Cooperative Institute for Research in Environmental Sciences / NOAA Chemical Sciences Laboratory
Brian Medeiros	NSF National Center for Atmospheric Research
Sebastián Mendoza-Téllez	Universidad Nacional Autónoma de México / Instituto de Ciencias de la Atmósfera y Cambio Climático
Katherine Mifsud	University of Washington
August Mikkelsen	University of Wyoming
Hugh Morrison	NSF National Center for Atmospheric Research
Johannes Mülmenstädt	Pacific Northwest National Laboratory
Samuel O'Donnell	Colorado State University
Camden Opfer	University of Colorado
Lekha Patel	Sandia National Laboratories
Mike Patterson	US CLIVAR
Pappu Paul	University of Illinois Urbana-Champaign
Gwenore Pokrifka	Princeton University
Prasanth Prabhakaran	University of Colorado / Cooperative Institute for Research in Environmental Sciences / NOAA Chemical Sciences Laboratory
Léa Prévost	University of Leeds
Yi Qin	Pacific Northwest National Laboratory
Kwesi T. Quagraine	Indiana University Bloomington
Leighton Regayre	University of Leeds / UK Met Office
Nicole Riemer	University of Illinois Urbana-Champaign
Elise Rosky	University of Michigan
Puja Roy	NSF National Center for Atmospheric Research
Cole Sagan	Princeton University
Masanori Saito	University of Wyoming
Sean Patrick Santos	Pacific Northwest National Laboratory

Luise Schulte	European Centre for Medium-Range Weather Forecasts / University of Cologne
Rick Schulte	Colorado State University
Alice Shen	University of Wyoming
Israel Silber	Pacific Northwest National Laboratory
Clare Singer	NOAA Geophysical Fluid Dynamics Laboratory
Kevin Smalley	Lawrence Livermore National Laboratory
Ci Song	University of Wyoming
McKenna Stanford	Pacific Northwest National Laboratory
Sylvia Sullivan	University of Arizona
Hongwei Sun	University of Washington
Mengyu Sun	University of Manchester
Kenta Suzuki	University of Tokyo
Madhusmita Swain	State University of New York at Albany
Patrick Taylor	NASA Langley Research Center
Shanru Tian	University of Texas at Arlington
Brian Tinsley	University of Texas at Dallas
Velle Toll	Tartu Ülikool
Florian Tornow	Columbia University / NASA Goddard Institute for Space Studies
Marcus van Lier-Walqui	Columbia University / NASA Goddard Institute for Space Studies
Chiara Ventrucci	University of Bologna / Consiglio Nazionale delle Ricerche Istituto di Scienze dell’Atmosfera e del Clima
Mingyi Wang	University of Chicago
Duncan Watson-Parris	UC San Diego / Scripps Institution of Oceanography
Geethma Werapitiya	University of Wyoming
Kayla White	University of Texas at Austin
Luke Whitehead	University of Oslo
Rob Wood	University of Washington
Takanobu Yamaguchi	University of Colorado / Cooperative Institute for Research in Environmental Sciences / NOAA Chemical Sciences Laboratory
Kohei Yamasaki	University of Tokyo / Atmosphere and Ocean Research Institute
Qingyuan Yang	Columbia University / Learning the Earth with Artificial Intelligence and Physics
Mark Zelinka	Lawrence Livermore National Laboratory
Jianhao Zhang	University of Colorado / Cooperative Institute for Research in Environmental Sciences / NOAA Chemical Sciences Laboratory
Yunyan Zhang	Lawrence Livermore National Laboratory
Damao Zhang	Pacific Northwest National Laboratory
Xueying Zhao	NSF National Center for Atmospheric Research / University of Texas at Austin
Xiaoli Zhou	Dalhousie University
Zeen Zhu	Brookhaven National Laboratory
Paquita Zuidema	University of Miami / Rosenstiel School of Marine, Atmospheric, and Earth Science

APPENDIX C: AGENDA

Monday, October 28, 2024		
Time (MDT)	**Agenda**	**Presenter**
7:00	Workshop registration and breakfast	
7:50	Welcome, opening remarks, and workshop goals	
8:15	Session 1: What's Wrong with Microphysics in Climate Models? Chairs: Adele Igel and Daniel McCoy	
8:15	(Invited) What's wrong with microphysics in climate models, and how can we fix it?	Johannes Mülmenstädt, Pacific Northwest National Laboratory
8:40	(Invited) Addressing outstanding uncertainties associated with high clouds	Sylvia Sullivan, University of Arizona
9:05	Detecting and identifying the impact of parameter interaction on climate model outputs based on two Perturbed Parameter Ensembles (PPEs)	Qingyuan Yang, Columbia University
9:20	Tracking structural uncertainty in aerosol model representation	Nicole Riemer, University of Illinois Urbana-Champaign
9:35	The relationship between condensate lifetime and precipitating efficiency and their response to sea surface warming	Hassan Beydoun, Lawrence Livermore National Laboratory
9:50	Open discussion	
10:30	Break	
10:55	Lightning talks for virtual posters	
11:30	Poster Session 1	
12:30	Lunch	
1:30	Session 2: Can We Even Observe Microphysics? Chairs: Coty Jen and Masa Saito	
1:30	(Invited, Virtual) Watching super-cooled water droplets nucleate heterogeneously with high speed cryo-microscopy	Nadine Borduas-Dedekind, University of British Columbia
1:55	(Invited) Remote sensing for cloud and precipitation measurement and science—Promises and challenges	Christine Chiu, Colorado State University
2:20	Can we infer microphysical process information from (infrequent) snapshots of data?	Graham Feingold, NOAA Chemical Services Laboratory
2:35	(Virtual) Aerosol products from PACE multi-angle polarimetric observations	Meng Gao, NASA Goddard Space Flight Center
2:50	Molecular simulations provide evidence in support of ice nucleation upon collision or breakup of supercooled cloud droplets	Elise Rosky, University of Michigan
3:05	Open discussion	
3:30	Break	

4:00	Breakout Session 1	
5:15	End of day 1 for virtual participants; Networking event	
8:00	End of day 1	

Tuesday, October 29, 2024		
Time (MDT)	**Agenda**	**Presenter**
7:00	Breakfast	
8:00	Recap day 1	
8:30	Session 3: How are Observations Being Used to Improve Models? Chairs: Leighton Regayre and Duncan Watson-Parris	
8:30	(Invited) Towards maximum feasible reduction of aerosol forcing uncertainty	Ken Carslaw, University of Leeds
8:55	(Invited) Using observations to design Earth System Model Perturbed Parameter Ensembles	Gregory Elsaesser, NASA Goddard Institute for Space Studies
9:20	Constraining warm cloud precipitation initiation using aircraft measurements	Patrick Chuang, UC Santa Cruz
9:35	Satellite-based model constraints on cloud microphysical processes and its link to radiative forcing of aerosol-cloud interaction	Kenta Suzuki, University of Tokyo
9:50	Unraveling the striking difference in casual low-cloud susceptibility to aerosol between GCM and observation	Jianhao Zhang, NOAA Chemical Sciences Laboratory
10:05	Open discussion	
10:45	Break	
11:00	Poster Session 2	
12:30	Lunch	
1:30	Session 4: How Can We Plan Future Process Observations to Reduce Climate Change Uncertainty? Chairs: Susannah Burrows and Rob Wood	
1:00	(Invited) Ice multiplication as a long-standing question in cloud microphysics	Alexei Kiselev, Karlsruhe Institute of Technology
1:25	(Invited) A process system approach for addressing climate change uncertainties	Christina McCluskey, NSF NCAR
1:50	A novel computational framework for optimal experimental design to improve climate prediction	Zhongjing Jiang, Brookhaven National Laboratory
2:05	Investigating the impacts of aerosol perturbations with a denoising diffusion model	Jatan Buch, Columbia University
2:20	Observing cloud microphysics using (ultra) high resolution radar and lidar system	Zeen Zhu, Brookhaven National Laboratory
2:35	Open discussion	

3:00	Break	
3:15	Poster Session 3	
4:15	Breakout Session 2	
5:30	End of day 2	

Wednesday, October 30, 2024		
Time (MDT)	**Agenda**	**Presenter**
7:00	Breakfast	
8:00	Recap day 2	
8:30	Session 5: What Do We Do Next? Chairs: Ann Fridlind and Susannah Burrows	
8:30	(Invited) Simulating cloud microphysics across scales for predicting climate extremes	Andrew Gettelman, Pacific Northwest National Laboratory
8:55	Significant trend in vertical wind velocity variability revealed by machine learning	Donifan Barahona, NASA Goddard Space Flight Center
9:10	Confronting structural uncertainty in aerosol-cloud interactions through process-level benchmarking	Laura Fierce, Pacific Northwest National Laboratory
9:25	The death of autoconversion?	Kaitlyn Loftus, Columbia University
9:40	Open discussion	
10:00	Break	
10:30	Final discussion and next steps	
11:45	End of workshop	

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from
the Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062
www.osti.gov
ph: (865) 576-8401
fox: (865) 576-5728
email: reports@osti.gov

Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
or (703) 605-6000
email: info@ntis.gov
Online ordering: http://www.ntis.gov

Credit: University of Wyoming Photo Service

usclivar.org/meetings/micro2macro

US Climate Variability and Predictability Program

www.usclivar.org

bsky.app/profile/usclivar.bsky.

PNNL-38643

US CLIVAR acknowledges support from these US agencies and

The material was developed with federal support of NOAA (NA11OAR4310473), NSF (AGS-1502208), and DOE (DE-SC0019366). Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the sponsoring agencies.