https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

Towards Resolving Equifinality in Aerosol-Cloud Radiative Forcing Through Process-Level Constraints

McKenna W. Stanford[a], Naser Mahfouz[a], Johannes Mülmenstädt[a], August Mikkelsen[b], and Susannah Burrows[a]

[a]Atmospheric, Climate, and Earth Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA
[b]Department of Atmospheric Science, University of Wyoming, Laramie, WY, USA

Correspondence: McKenna W. Stanford (mckenna.stanford@pnnl.gov) and Johannes Mülmenstädt (johannes.muelmenstaedt@pnnl.gov)

Abstract. Equifinality—where multiple parameter combinations produce indistinguishable climate states—is a fundamental obstacle to calibrating Earth system models (ESMs). Parameter sets that agree well with the observed mean state can produce dramatically different responses to external perturbations, reflecting a mathematical non-uniqueness that standard calibration approaches cannot resolve. This poses a critical challenge for constraining the effective radiative forcing due to aerosol-cloud interactions ($\mathrm{ERF_{ACI}}$), which depends sensitively on cloud responses to aerosol perturbations. We use a perturbed parameter ensemble that, as a minimal working example, varies only two parameters related to warm-rain formation (autoconversion and accretion) and constrain on synthetic observations in a perfect-model setup. Even under ideal conditions, radiation-only constraints produce a fundamentally degenerate, bimodal $\mathrm{ERF_{ACI}}$ posterior. Incorporating the observable cloud state (liquid water path, $\mathcal{L}$) for warm marine stratocumulus clouds can resolve this bimodality, but its utility diminishes with large observational uncertainty. A far more robust constraint is achieved by directly targeting the underlying physical process rates that modulate $\mathcal{L}$. Process-rate constraints effectively eliminate bimodality and dramatically reduce posterior uncertainty even under large assumed observational uncertainties. As process rates are not directly observable, we test their observable proxies as a tractable alternative, finding they offer only limited additional value over cloud state due to their emergent behavior. These results establish a critical principle for ESM calibration: to achieve robust constraints on $\mathrm{ERF_{ACI}}$ and beyond, it is necessary to constrain both the emergent climate state and the physical pathways that govern its formation.

1 Introduction

Equifinality, in which multiple distinct parameter combinations produce equally plausible emergent system states, is a pervasive challenge in calibrating Earth system models (Von Bertalanffy, 1950). First formalized in Earth system modeling by Beven and Binley (1992) in the context of hydrological models, the problem is directly analogous to the parametric degeneracies that plague climate model calibration. The phenomenon has been recognized across a broad range of scientific disciplines, including geophysics (Beven, 2006; Nicholas and Quine, 2010; Mülmenstädt and Feingold, 2018; Cho et al., 2025), biology (Tang and Zhuang, 2008), ecology (Luo et al., 2009), neuroscience (Prinz et al., 2004; Marder and Taylor, 2011), and the social sciences (Gresov and Drazin, 1997). In ESMs, the core problem is that parameter sets optimized against an identical observed mean state

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

can yield substantially different responses when the system is perturbed, a degeneracy that standard calibration strategies are ill-equipped to detect or resolve. Tuning approaches that optimize models against Earth's observed top-of-atmosphere (TOA) radiation balance (e.g., Mauritsen et al., 2012; Golaz et al., 2013; Hourdin et al., 2017; Schmidt et al., 2017) are especially susceptible to this problem. While adding constraints from observed cloud properties provides additional information, these emergent properties result from a complex balance of underlying processes (Quaas et al., 2009; Wang et al., 2012; Klein and Hall, 2015) and are themselves subject to large observational uncertainties (Bing Lin and Rossow, 1994; O'Dell et al., 2008; Seethala and Horváth, 2010; Greenwald et al., 2018), often leaving them insufficient to break physical degeneracies between competing parameterizations.

Equifinality is particularly consequential for constraining the effective radiative forcing due to aerosol-cloud interactions ($\mathrm{ERF_{ACI}}$), which remains a leading source of uncertainty in projections of Earth's total radiative forcing (Carslaw et al., 2013; Boucher et al., 2014; Mülmenstädt and Feingold, 2018; Bellouin et al., 2020; Forster et al., 2021). Aerosol-cloud interactions (ACI) occur where anthropogenic aerosol perturbations modify cloud properties and alter the planetary radiation budget (Rosenfeld et al., 2014; Gettelman, 2015). $\mathrm{ERF_{ACI}}$ integrates the increase in cloud albedo from smaller, more numerous cloud droplets (i.e., the Twomey effect; Twomey, 1974) with subsequent cloud adjustments including modulations to liquid water path ($\mathcal{L}$) and cloud cover (Albrecht, 1989; Ackerman et al., 2004; Bretherton et al., 2007). These adjustments involve competing physical pathways—including precipitation suppression and enhanced cloud-top entrainment—leaving the sign and magnitude of their net effect uncertain (Goren and Rosenfeld, 2014; Mülmenstädt et al., 2024a, b). Equifinality sharpens this challenge considerably: parameter sets that reproduce the observed present-day radiation budget with equal fidelity can yield sharply divergent cloud responses to aerosol perturbation and, consequently, widely varying $\mathrm{ERF_{ACI}}$ estimates. Reducing this uncertainty requires methodologies capable of probing process-level behavior and confronting the structural degeneracies that conventional calibration approaches leave unresolved (Carslaw et al., 2013; Bony et al., 2015; Knutti et al., 2017).

To address this limitation, we propose that process-level constraints can conditionally resolve equifinality as it manifests in $\mathrm{ERF_{ACI}}$ uncertainty. We demonstrate this using a 50-member perturbed parameter ensemble (PPE; e.g., Carslaw et al., 2026) within the U.S. Department of Energy's (DOE) Energy, Exascale, and Earth System Model version 3 (E3SMv3; E3SM Project, 2024). As a minimal working example, we explore parametric uncertainty by perturbing only two warm-rain process parameterizations: autoconversion and accretion. Autoconversion describes the spontaneous collision and coalescence of cloud droplets to form the first drizzle drops, while accretion describes their subsequent growth by sweeping up additional cloud droplets. Together, these two processes control the onset and intensity of warm rain and are the primary microphysical pathways through which aerosol perturbations modify $\mathcal{L}$ and cloud lifetime. We employ a perfect-model setup in which a control simulation provides synthetic truth (e.g., Lee et al., 2016), deliberately isolating the parametric degeneracy that is the focus of this study from the confounding effects of structural model error—deficiencies in the model's process representations that would introduce additional sources of misfit independent of parameter choice—and observational bias. We then apply hierarchical constraints using a Markov Chain Monte Carlo (MCMC) Bayesian inference approach across three tiers of constraints: (1) radiative state (TOA fluxes), (2) cloud state (liquid water path, $\mathcal{L}$), and (3) direct process rates (autoconversion and accretion rates).

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

Radiation-only constraints produce a bimodal $\mathrm{ERF_{ACI}}$ posterior reflecting a fundamental parametric degeneracy, where multiple parameter combinations satisfy the radiative constraint via physically distinct cloud regimes. A single cloud-state constraint ($\mathcal{L}$) resolves this bimodality under realistic observational uncertainties but fails as uncertainty approaches 50%. Process-rate constraints, even under large observational uncertainty, substantially narrow the $\mathrm{ERF_{ACI}}$ posterior and robustly eliminate the bimodality. Observable proxies for process rates—surface precipitation rate and precipitation occurrence frequency—offer limited additional constraining value beyond $\mathcal{L}$ alone and are insufficient to fully resolve this degeneracy. This framework thus serves as a controlled demonstration that equifinality in $\mathrm{ERF_{ACI}}$ is conditionally reducible, but only when constraints are anchored in the physical mechanisms that govern the emergent climate state.

The remainder of this article is structured as follows: Sect. 2 describes the PPE design, emulator construction, and Bayesian calibration framework. Results are presented in Sect. 3, followed by discussion in Sect. 4 and conclusions in Sect. 5.

2 Data and Methods

The overarching goal of this study is to quantify the extent to which different classes of observational constraints can reduce parametric degeneracy in $\mathrm{ERF_{ACI}}$ within a controlled Bayesian calibration framework. This framework, summarized in Fig. 1, consists of five stages: construction of a perturbed parameter ensemble (PPE), training of Gaussian Process (GP) emulators as fast model surrogates, application of hierarchical constraints via Markov Chain Monte Carlo (MCMC) sampling, and propagation of the constrained parameter distributions back through the emulators to recover posterior distributions of key output variables including $\mathrm{ERF_{ACI}}$. The following subsections describe the E3SM model configuration and PPE design (Sect. 2.1), the output variables and derived quantities used as constraints (Sect. 2.2), the GP emulator construction (Sect. 2.3), the hierarchical constraint tiers (Sect. 2.4), and the Bayesian calibration and posterior propagation methodology (Sect. 2.5).

2.1 E3SM Perturbed Parameter Ensemble

The PPE is performed with the U.S. Department of Energy (DOE) Energy Exascale Earth System Model version 3 (E3SM Project, 2024; Xie et al., 2025) at a standard effective horizontal resolution of 100 km with 80 vertical levels. Key physics schemes include the Cloud Layers Unified By Binomials (CLUBB) scheme for macrophysics and vertical mixing (Golaz et al., 2002; Larson et al., 2005; Larson, 2017), the Predicted Particle Properties (P3) microphysics scheme (Morrison and Milbrandt, 2015; Terai et al., 2025) for stratiform cloud microphysics, the Zhang and McFarlane (1995) convection scheme which includes effects from aerosol-modulated convective cloud microphysics (Terai et al., 2025), the Rapid Radiative Transfer for General circulation Models (RRTMG; Mlawer et al., 1997; Clough et al., 2005; Iacono et al., 2008) radiation code, and the Modal Aerosol Model (MAM) scheme based on MAM4 (Wang et al., 2026).

We perturb two process rates known for their strong influence on warm-rain formation and ACI (e.g., Mahfouz et al., 2024; Feng et al., 2026): autoconversion and accretion. Perturbations are applied as multiplicative factors ($\mathcal{C}_{\mathrm{auto}}$ and $\mathcal{C}_{\mathrm{accr}}$, respectively) to the default process rate formulations within P3, which follow Khairoutdinov and Kogan (2000, their Eqs. 33

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

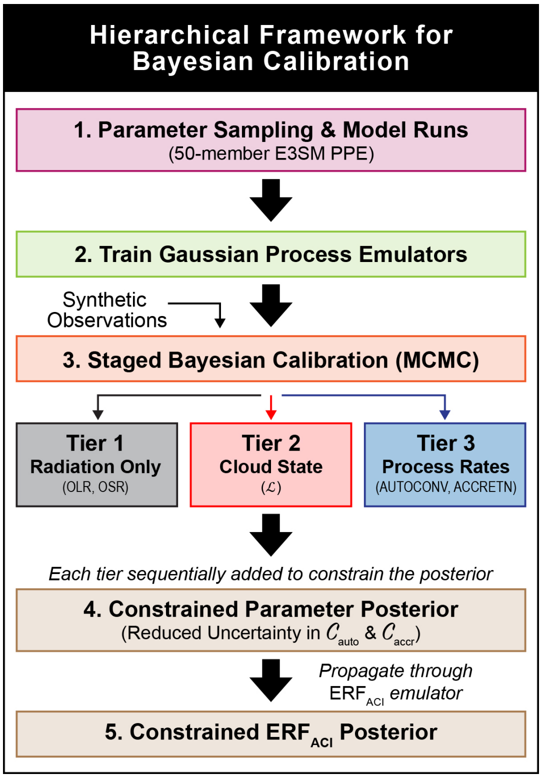

Figure 1. Conceptual diagram of the hierarchical Bayesian calibration framework used in this study. In Step 1, the PPE is constructed and realized. In Step 2, Gaussian Process (GP) emulators are trained as fast statistical surrogates of the model. In Step 3, the emulators are used with synthetic observational targets in MCMC simulations performed across three staged constraint tiers: **Radiation Only**, **Cloud State**, and **Process Rates**. The constrained input parameters ($\mathcal{C}_{\mathrm{auto}}$ and $\mathcal{C}_{\mathrm{accr}}$, Step 4) are propagated back through the emulators to realize constrained posterior distributions of $\mathrm{ERF}_{\mathrm{ACI}}$ and other output variables (Step 5).

and 34):

$$\left(\frac{dq_r}{dt}\right)_{\mathrm{auto}} = \mathcal{C}_{\mathrm{auto}}\mathcal{A}_{\mathrm{auto}}N_c^{\beta}q_c^{\gamma} \tag{1}$$

$$\left(\frac{dq_r}{dt}\right)_{\mathrm{accr}} = \mathcal{C}_{\mathrm{accr}}\mathcal{A}_{\mathrm{accr}}(q_c q_r)^{\lambda} \tag{2}$$

where N_c is the in-cloud cloud droplet number concentration, q_c and q_r are the cloud droplet and rain mass mixing ratios, respectively (Morrison et al., 2008), and $\mathcal{C}_{\mathrm{auto}}$, $\mathcal{C}_{\mathrm{accr}}$ default to unity. The coefficients $\mathcal{A}_{\mathrm{auto}}$, β, and γ take default values of 30500, -1.1, and 3.19, respectively; for accretion, $\mathcal{A}_{\mathrm{accr}} = 117.25$ and $\lambda = 1.15$. A 50-member ensemble was generated by sampling $\mathcal{C}_{\mathrm{auto}}$ and $\mathcal{C}_{\mathrm{accr}}$ via Latin hypercube sampling in $\log_2$ space over the range $[-10, 10]$ (i.e., 2^{-10} to 2^{10}; see Fig. A1).

Each ensemble member, plus a control run (CNTL) with default parameters ($\mathcal{C}_{\mathrm{auto}} = \mathcal{C}_{\mathrm{accr}} = 1$), was integrated for 13 months with prescribed sea surface temperatures and sea ice extent; the first month is discarded as spin-up. For each member,

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

paired present-day (PD; aerosols representative of 2010) and pre-industrial (PI; aerosols representative of circa 1850) simulations were run to permit calculation of $\mathrm{ERF_{ACI}}$ and cloud adjustments. Nudging is used to control the meteorological state to enhance the signal-to-noise ratio when calculating the ERF metrics between the PD and PI aerosol states. Following previous studies (Mahfouz et al., 2024, 2025), we nudge horizontal winds to MERRA2 at all vertical levels every six hours with a relaxation timescale of six hours.

2.2 Output Variables and Derived Quantities

Output variables are processed from 3-hourly simulation output. Radiation variables, including outgoing longwave radiation (OLR) and outgoing shortwave radiation (OSR), are averaged globally. Cloud-state and process-rate variables are analyzed specifically within warm, overcast marine stratocumulus decks following the dynamical regime definition of Medeiros and Stevens (2011) as applied by Mülmenstädt et al. (2024b): grid points where the lower-tropospheric stability (LTS) exceeds 18.55 K, large-scale subsidence (ω) at 700 and 500 hPa exceeds 10 hPa $\mathrm{d^{-1}}$, and cloud fraction exceeds 90% for at least 30% of the time. These criteria isolate the subtropical eastern ocean basins where the stratocumulus regime dominates (Fig. A3). Radiative state variables are computed as temporal and global area-weighted means; cloud-state and process-rate variables are computed as temporal and area-weighted means over columns meeting the stratocumulus regime criteria.

The analyzed cloud-state and process-rate variables include in-cloud liquid water path ($\mathcal{L}$) and vertically integrated autoconversion (AUTOCONV) and accretion (ACCRETN) rates. The $\mathrm{ERF_{ACI}}$ is computed following the decomposition of Ghan (2013). The top-of-atmosphere radiative imbalance (or the RESidual radiation at Top-Of-Model, RESTOM) is computed under three conditions: all-sky ($\mathrm{RESTOM_{all}}$), clean-sky ($\mathrm{RESTOM_{cln}}$), and clean-clear-sky ($\mathrm{RESTOM_{clrcln}}$). Clean-clear-sky fluxes represent the fluxes that would be present if there were neither aerosols nor clouds and are calculated by adding an additional radiation call at the very beginning of the radiative transfer logic before the optics class is endowed with aerosol and cloud properties. Clean-sky fluxes represent the fluxes that would be present if there were no aerosols and are calculated by adding an additional radiation call after instantiating an additional optics class, but not endowing it with aerosol properties. $\mathrm{ERF_{ACI}}$ is then defined as:

$$\mathrm{ERF_{ACI}} = \Delta\mathrm{RESTOM_{cln}} - \Delta\mathrm{RESTOM_{clrcln}} \tag{3}$$

where Δ denotes the PD−PI difference. This isolates the radiative effect of aerosol-induced cloud changes from the direct aerosol radiative effect and residual contributions. The $\mathcal{L}$ adjustment is defined as $\mathcal{L}_{\mathrm{adj}} = \ln(\mathcal{L}_{\mathrm{PD}}/\mathcal{L}_{\mathrm{PI}})$.

Because direct process rates are not observable, we also evaluate three integrated proxies for warm-rain formation that are retrievable by modern space-borne platforms. The warm rain occurrence frequency (WROF) is the fraction of time a stratocumulus column precipitates at a rate exceeding 10^{-3} mm $\mathrm{h^{-1}}$, a threshold representative of active space-borne sensors such as CloudSat (Lebsock and L'Ecuyer, 2011), and acts as a maximally coarse-grained proxy for rain initiation (Mülmenstädt et al., 2021; Stanford et al., 2023). The precipitation amplification frequency (PAF) is the conditional frequency that a precipitating column exceeds 10^{-1} mm $\mathrm{h^{-1}}$ and acts as a proxy for rain intensity. The large-scale surface precipitation rate ($\mathcal{P}$) is a direct

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

output that represents the integrated outcome of autoconversion, accretion, and evaporation. The constraining performance of these proxies relative to direct process rates is evaluated in Sect. 3.3.

2.3 Gaussian Process Emulators

Fast statistical surrogates of simulated responses to perturbed input parameters are constructed as Gaussian Process (GP) emulators. The 2D input space ($\mathcal{C}_{\mathrm{auto}}$, $\mathcal{C}_{\mathrm{accr}}$) is transformed by taking $\log_{10}$ and scaled to the unit interval $[0,1]$ to optimize emulator performance. We use the `scikit-learn` (Pedregosa et al., 2011) package's `GaussianProcessRegressor` based on Rasmussen and Williams (2005). Separate emulators are trained for each output variable; for variables spanning multiple orders of magnitude (AUTOCONV, ACCRETN), emulators are trained on their natural logarithm. Full details of the kernel specification and hyperparameter optimization are given in Appendix B.

Emulator performance is validated using 5-fold cross-validation, with fidelity quantified by the coefficient of determination (R^2) and the percentage of true values falling within the emulator's predicted 95% confidence interval (full results in Fig. B1). For primary state variables (OLR, OSR, $\mathcal{L}$) and process rates (AUTOCONV, ACCRETN), R^2 exceeds 0.99. One exception is $\mathcal{L}_{\mathrm{adj}}$ ($R^2 = 0.223$), where poor performance is localized to physically pathological regimes (near-total cloud desiccation or hyper-suppressed precipitation) at the extremes of the parameter space. However, the emulator captures the correct physical gradient for the well-behaved majority of the parameter space, making it suitable for the diagnostic purposes of this study (see Appendix B for full discussion). Following validation, production emulators are trained on all 50 members and used throughout the analysis.

2.4 Hierarchical Constraint Framework

The core framework of our study tests three hierarchical tiers of observational constraints, each representing a different level of physical detail:

1. **Radiative State (Tier 1):** TOA OLR and OSR; representing the canonical approach for tuning climate models to achieve global radiative balance (e.g., Mauritsen et al., 2012; Hourdin et al., 2017).

2. **Cloud State (Tier 2):** Adds a constraint on $\mathcal{L}$, the primary bulk cloud property expected to respond to perturbations of warm-rain parameters.

3. **Process Rates (Tier 3):** Targets AUTOCONV and ACCRETN directly, representing the hypothetical best-case scenario in which the underlying warm-rain formation processes are observable.

These tiers represent a progression from canonical model calibration (Tier 1) to crucially but potentially under-utilized emergent cloud state constraints (Tier 2) to an aspirational process-level target (Tier 3).

To isolate the constraining performance of each tier from the confounding effects of structural model error and observational bias, the framework is evaluated in a perfect-model setup (e.g., Lee et al., 2016). In this approach, the CNTL simulation ($\mathcal{C}_{\mathrm{auto}} = \mathcal{C}_{\mathrm{accr}} = 1$) serves as synthetic truth, and all observational targets are drawn directly from it. This design provides an

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

unambiguous upper bound on the constraining power of each tier: any residual uncertainty in the posterior reflects the mathematical structure of the parametric degeneracy itself, not limitations of the observations or the model's physical formulation.

Within this framework, observational uncertainties do not represent measurement errors in the traditional sense but rather the assumed tolerance around the synthetic target value—that is, how precisely each constraint variable must be matched to be considered consistent with the CNTL state. For Tier 1, we follow L'Ecuyer et al. (2015, their Table 3) and apply fixed absolute uncertainties of 2 W m^{-2} to both OLR and OSR, consistent with satellite-derived estimates of TOA energy budget uncertainty. For Tiers 2 and 3 (and integrated process-rate proxies), we test relative uncertainties of 5%, 20%, and 50% of the CNTL value, representing optimistic, realistic, and lenient assumptions, respectively. MCMC experiments sequentially add constraints beginning from Tier 1, allowing precise quantification of the added value of each observational class in reducing ERF_{ACI} uncertainty. The technical implementation of these uncertainties within the likelihood function is described in the following section.

2.5 Bayesian Calibration and Posterior Propagation

Constrained posterior distributions for $\mathcal{C}_{\mathrm{auto}}$ and $\mathcal{C}_{\mathrm{accr}}$ are obtained via Bayesian inference using the affine-invariant ensemble sampler `emcee` (Foreman-Mackey et al., 2013) based on the methods proposed by Goodman and Weare (2010). This algorithm uses an ensemble of "walkers" that propose moves based on the positions of other walkers, allowing the sampler to efficiently adapt to the posterior distribution's geometry even under anisotropic conditions (see Fig. C1) and without the manual tuning required by traditional Metropolis-Hastings methods. Full details of the sampler configuration are given in Appendix C; briefly, we use 100 walkers each taking 5000 steps with the first 20% discarded as burn-in, yielding 400,000 posterior samples.

The log-posterior is the sum of a log-prior and log-likelihood. The log-prior enforces a valid parameter range of 2^{-8} to 2^{8} to account for extreme emulator uncertainty at the boundaries of the input parameter space (see Fig. B2), returning 0 within bounds and $-\infty$ otherwise. The core log-likelihood function, $\ln(\mathbb{L}(\Phi))$, where Φ is a set of $\mathcal{C}_{\mathrm{auto}}$ and $\mathcal{C}_{\mathrm{accr}}$, quantifies the misfit between emulator predictions and the synthetic observations from CNTL. For single-variable constraints, assuming errors are normally distributed across the constraint value, the log-likelihood is given by the squared residual:

$$\ln(\mathbb{L}(\Phi)) = -\frac{1}{2}\frac{(f_i(\Phi)-y_i)^2}{\sigma^2_{\mathrm{obs},i}+\sigma^2_{\mathrm{GP},i}(\Phi)} \tag{4}$$

where $f_i(\Phi)$ is the emulator prediction of variable i, $\sigma^2_{\mathrm{obs},i}$ is the assumed observational variance, and $\sigma^2_{\mathrm{GP},i}(\Phi)$ is the emulator variance. In the perfect-model setup employed here, the observational target y_i is set to the corresponding output value from the CNTL simulation ($\mathcal{C}_{\mathrm{auto}} = \mathcal{C}_{\mathrm{accr}} = 1$), such that the likelihood penalizes parameter combinations whose emulated outputs deviate from the control state. To realistically account for known physical covariances between multiple climate variables (e.g., Fig. A4), we use a multivariate log-likelihood based on the Mahalanobis distance:

$$\ln(\mathbb{L}(\Phi)) = -\frac{1}{2}(f(\vec{\Phi})-\vec{y})^{\top}\vec{\Sigma}^{-1}(f(\vec{\Phi})-\vec{y}) \tag{5}$$

where $f(\vec{\Phi})$ and $\vec{y}$ are vectors of predictions and observations, and $\vec{\Sigma}$ is the total error covariance matrix. To prevent the large structural variance of the PPE (e.g., see prior distributions in Fig. A2) from overwhelming the observational variance and

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

rendering the likelihood surface uselessly flat, we construct a hybrid covariance matrix:

$$\vec{\Sigma}_{\mathrm{hybrid}} = \vec{D}\vec{R}\vec{D} \tag{6}$$

where $\vec{R}$ is the dimensionless correlation matrix computed from the 50-member PPE and $\vec{D}$ is the diagonal matrix of total standard deviations, with each entry $\sqrt{\sigma^2_{\mathrm{obs}} + \sigma^2_{\mathrm{GP}}}$. When $\vec{R}$ is set to the identity matrix (i.e., assuming no correlation between variables), Eq. 5 reduces to the sum of independent log-likelihoods in Eq. 4.

The final step propagates the constrained posterior samples of $\mathcal{C}_{\mathrm{auto}}$ and $\mathcal{C}_{\mathrm{accr}}$ from the MCMC chain back through the trained emulators to generate posterior distributions of output variables (e.g., $\mathrm{ERF_{ACI}}$). This propagation can be performed either deterministically or stochastically. The deterministic method uses only the mean prediction from the emulator. The stochastic approach involves sampling from the emulator's full predictive Gaussian distribution (i.e., mean and variance), therefore including both parametric and emulator uncertainty. We follow the deterministic approach to isolate the effect of parametric uncertainty while accounting for emulator uncertainty in the MCMC likelihood (i.e., the denominator in Eq. 4).

3 Results

3.1 Constrained $\mathrm{ERF_{ACI}}$ Posterior

The unconstrained $\mathrm{ERF_{ACI}}$ prior spans 0 to -1 W m^{-2} (Fig. A2), reflecting the full range of climate states produced by the aggressive parameter perturbations—from essentially cloud-free (immediate precipitation) to persistently overcast (no precipitation). Radiation constraints eliminate these implausible states (Fig. 2, gray distributions) and narrow $\mathrm{ERF_{ACI}}$ to values between approximately -0.75 and -0.95 W m^{-2}. However, the radiation-constrained distribution reveals the fundamental model degeneracy manifesting as a bimodal posterior. The dominant mode (≈ -0.78 W m^{-2}) is aligned with the CNTL value, while the spurious stronger-forcing mode (≈ -0.95 W m^{-2}) arises from compensating autoconversion and accretion parameter pairs. Resolving this degeneracy—and identifying which constraints are capable of doing so—is the central question of the analysis that follows.

The radiation-only tier serves as the baseline; all uncertainty reductions in higher tiers are reported relative to it, with uncertainty quantified as the standard deviation (σ) of the constrained posterior (text boxes in Fig. 2). Adding the cloud-state constraint ($\mathcal{L}$) substantially narrows uncertainty under both optimistic (5%) and realistic (20%; Bing Lin and Rossow, 1994; O'Dell et al., 2008; Seethala and Horváth, 2010; Greenwald et al., 2018) observational errors and successfully resolves the bimodality (Fig. 2a,b, red distributions). However, at large observational uncertainties of 50% (Fig. 2c), the cloud-state constraint fails, producing a bimodal posterior virtually indistinguishable from the radiation-only result and failing to suppress the spurious strong-forcing mode.

The addition of process-rate constraints provides a far more robust path to reducing uncertainty across all tested observational errors (Fig. 2, blue distributions). At all error levels—from 5% to 50%—process-rate constraints effectively eliminate the spurious mode and consistently reduce the width of the dominant mode by approximately half relative to adding only the cloud-state-only constraint. These findings reveal that the emergent cloud state can statistically reject an implausible climate

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

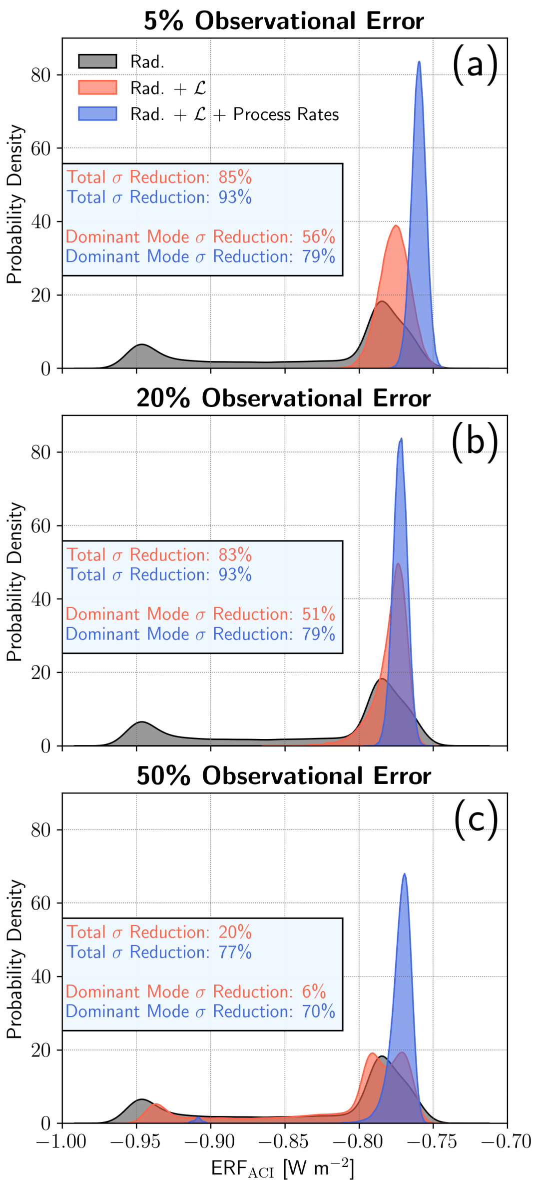

Figure 2. Posterior distributions (as Kernel Density Estimates, KDEs) of $\mathrm{ERF_{ACI}}$ produced by constraining on radiation only (black), radiation + $\mathcal{L}$ (red), and radiation + $\mathcal{L}$ + process rates (AUTOCONV and ACCRETN, blue) for assumed observational uncertainties of (a) 5%, (b) 20%, and (c) 50%. Text boxes show the reduction in uncertainty (1-σ) for each tier relative to the radiation-only distribution, computed as reductions in total σ and in the dominant mode's σ. The dominant mode is defined as the portion of the distribution above the local minimum of the radiation-only KDE between the two peaks (demarcation ≈ -0.85 W m^{-2}).

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

state when observational errors in $\mathcal{L}$ are sufficiently small, while process-level constraints both stabilize on the appropriate climate state and narrow the uncertainty within it regardless of observational error.

3.2 Resolving Equifinality in the Input Parameter Space

Reductions in $\mathrm{ERF_{ACI}}$ uncertainty arise fundamentally from reductions in the constrained input parameter space—ultimately the parameters we wish to calibrate—as demonstrated via joint posterior distributions of $\mathcal{C}_{\mathrm{auto}}$ and $\mathcal{C}_{\mathrm{accr}}$ (Fig. 3). The unconstrained sparse prior within this parameter space is simply the 50-member Latin hypercube samples used to generate the PPE (Figs. A1 and A2), while constraining on any given tier yields a high-density posterior within that 2D space. A perfect constraint would back out the target $\mathcal{C}_{\mathrm{auto}}$,$\mathcal{C}_{\mathrm{accr}}$ values of unity, essentially collapsing the posterior marginal distribution of either input parameter to a near Delta function.

When the target constraint is the TOA radiative state, which essentially reflects a target cloud brightness, then pairs of $\mathcal{C}_{\mathrm{auto}}$ and $\mathcal{C}_{\mathrm{accr}}$ are required that achieve some plausible combination of precipitating and non-precipitating cloud. Under these radiation-only constraints (black distributions in Fig. 3), plausible parameter pairs lie along a compensating ridge from low-$\mathcal{C}_{\mathrm{auto}}$/high-$\mathcal{C}_{\mathrm{accr}}$ (rare but heavy rain) to high-$\mathcal{C}_{\mathrm{auto}}$/low-$\mathcal{C}_{\mathrm{accr}}$ (frequent but light rain). Radiative state constraints therefore eliminate only the most implausible states—cloud-free or cloud-saturated. The compensating ridge therefore demonstrates a degeneracy whereby cloud brightness is correct but the underlying cloud state is not. Adding the $\mathcal{L}$ constraint (red distributions) breaks this degeneracy under small observational errors, concentrating the posterior near unity (Fig. 3a). At 20% error (Fig. 3b), $\mathcal{L}$ eliminates the extremes of the ridge but compensating processes still allow the target cloud state to be achieved via multiple parameter pathways. At 50% error (Fig. 3c), the joint distribution is largely indistinguishable from the radiation-only result. Adding process-rate constraints (blue distributions) finally achieves a climate state in which the clouds are appropriately bright and the cloud state that controls brightness is achieved through the correct physical pathways even under large observational uncertainty.

The physical basis for these posterior shapes is revealed by projecting the likelihood surfaces of each constraint tier onto the parameter space (Fig. 4, shown for 20% observational error). To standardize these surfaces, confidence intervals (CIs) are defined using Wilks' theorem (Wilks, 1938): for a given likelihood surface $\mathbb{L}(\Phi)$ with maximum $\mathbb{L}_{\max}$, the quantity $2(\mathbb{L}_{\max} - \mathbb{L}(\Phi))$ is asymptotically distributed as a χ^2 distribution with k degrees of freedom equal to the number of parameters (here, $k = 2$). The critical χ^2 values enclosing 68% (1-σ) and 95% (2-σ) of the probability are approximately 2.30 and 5.99, respectively, so CI boundaries are defined as contours where the log-likelihood has dropped from its maximum by $\Delta\mathbb{L} = \chi^2/2$. In Fig. 4, we plot the 1-σ contour at $\mathbb{L}_{\max} - 1.15$ for each constraint.

The 1-σ CIs for the radiation-only constraint (white contours) follow the same compensating ridge seen in the posterior (Fig. 3), confirming that a wide range of parameter combinations can satisfy tight TOA radiative state requirements. The CIs for $\mathcal{L}$ alone (green contours) trace a similar shape, following the parameter combinations that yield $\mathcal{L}$ near the control value of $\approx$80 g m^{-2} (Fig. 4b). The intersection of these two broad surfaces explains the shape of the radiation + $\mathcal{L}$ posterior in Fig. 3b.

In stark contrast, the process-rate CIs (dark blue contours) exhibit a fundamentally different structure, forming a diagonal likelihood space extending from the center towards the upper right. This seemingly counter-intuitive result—where hyper-

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

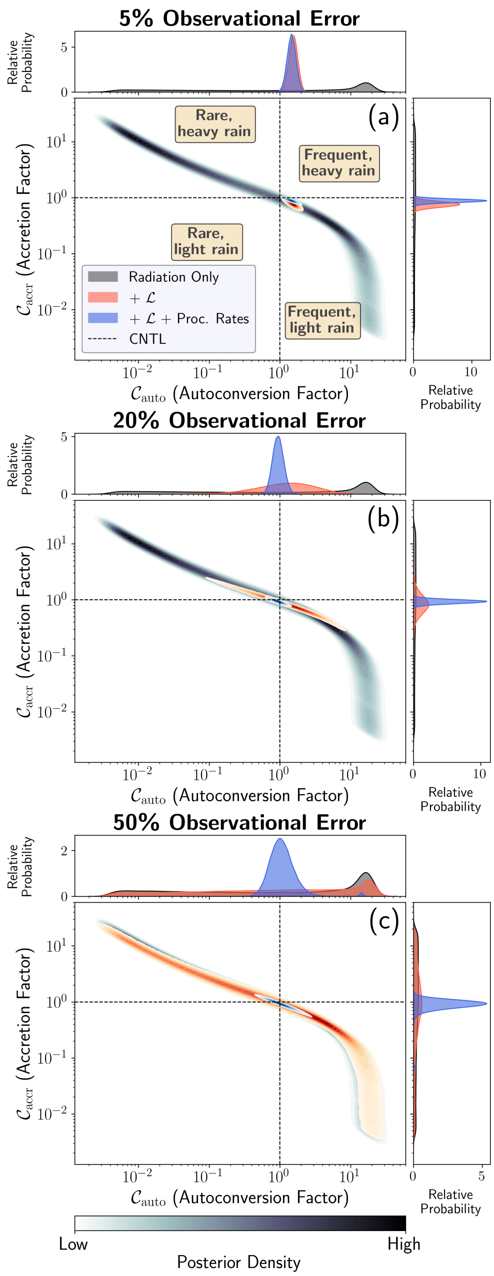

Figure 3. Conditional (large panels) and marginal (edge panels) posterior KDEs of input parameters $\mathcal{C}_{\rm auto}$ (abscissas) and $\mathcal{C}_{\rm accr}$ (ordinates) for radiation-only constraints (black hues), radiation + $\mathcal{L}$ (red hues), and radiation + $\mathcal{L}$ + process rates (blue hues) for assumed observational uncertainties of (a) 5%, (b) 20%, and (c) 50%. Black dashed lines at unity indicate the CNTL value. Filled contours represent probability densities from 5% to 95% (lighter to darker hues) as indicated by the color bar. Quadrant labels in (a) describe the warm-rain precipitation regime defined by the autoconversion (occurrence frequency) and accretion (intensity) factor combinations.

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

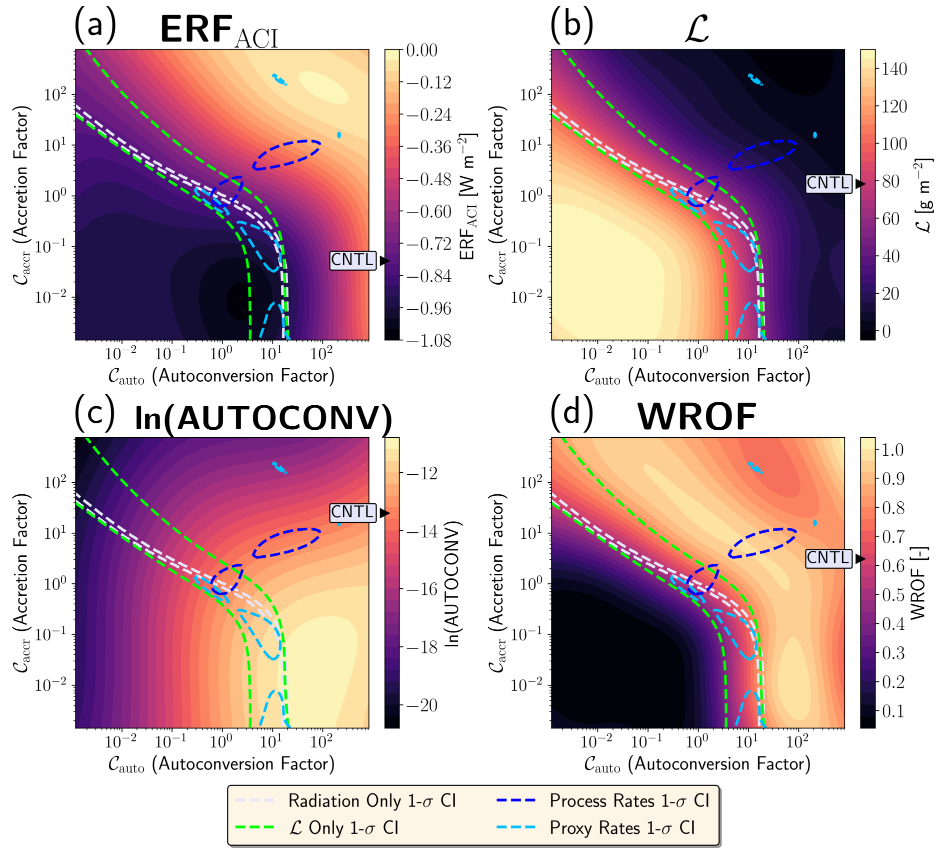

Figure 4. Emulator-predicted surfaces of select variables as a function of $\mathcal{C}_{\text{auto}}$ and $\mathcal{C}_{\text{accr}}$ (filled contours). Dashed contour lines are the 1-σ CIs for radiation only (white), $\mathcal{L}$ only (green), process rates (dark blue), and proxy rates (light blue), computed with 20% observational error. CI contours are identical across all panels. The CNTL value is indicated by a notch on each color bar.

efficient rain parameters are considered plausible—reveals a self-limiting feedback within the model physics. In the upper-right regime, large multiplicative factors drive such efficient rain formation that they deplete the available $\mathcal{L}$ (Fig. 4b). This depletion of $\mathcal{L}$, a primary ingredient for the process rates themselves (Eqs. 1 and 2), acts as a negative feedback, causing the final integrated AUTOCONV and ACCRETN values to fall back towards the CNTL values (e.g., Fig. 4c).

This demonstrates that while process-rate constraints are uniquely powerful at eliminating the compensating ridge degeneracy, they are susceptible to a different form of equifinality if used in isolation. The ultimate success of the framework therefore lies in the synergistic combination of all three tiers. Radiation and $\mathcal{L}$ constraints exclude the nonphysical self-depleted cloud states in the upper right, while the process-rate constraint eliminates the pathological lower-right quadrant. The only area satisfying all three conditions simultaneously is the small, tight lobe where the CIs intersect, forcing the solution towards the true parameter values near unity. Full likelihood surfaces for all constraint variables are given in Fig. D1.

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

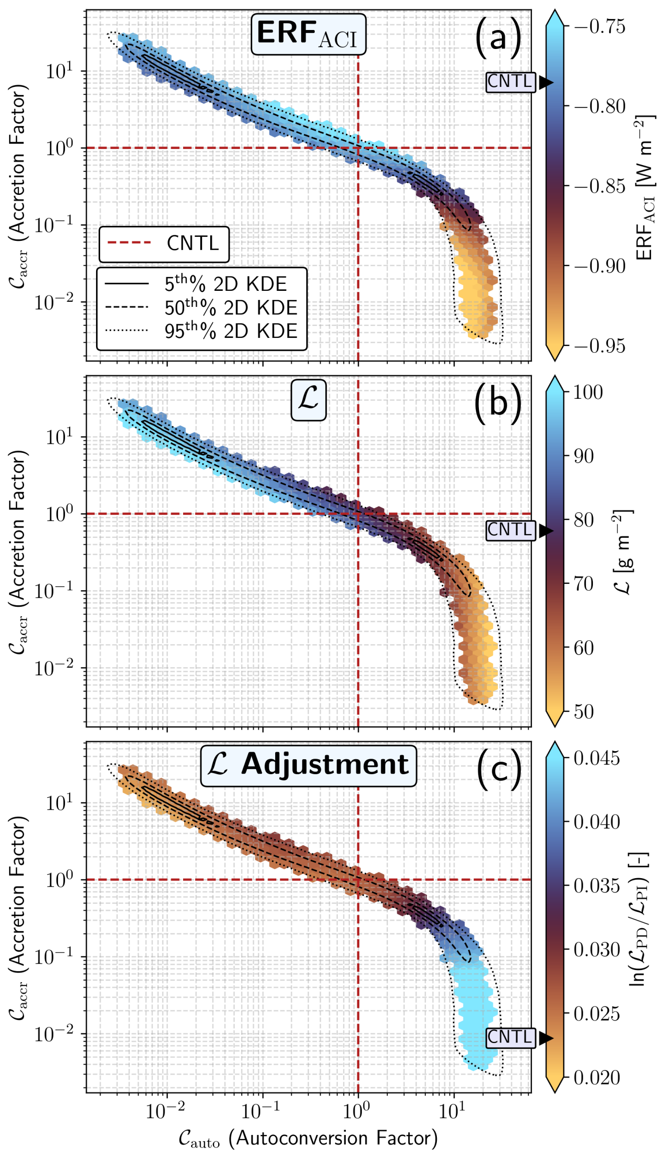

Figure 5. Joint KDEs between $\mathcal{C}_{\mathrm{auto}}$ (abscissas) and $\mathcal{C}_{\mathrm{accr}}$ (ordinates) showing the 5$^{\mathrm{th}}$, 50$^{\mathrm{th}}$, and 95$^{\mathrm{th}}$ percentiles (black lines) under radiation-only constraints. Filled hex-bins show the average (a) ERF$_{\mathrm{ACI}}$, (b) $\mathcal{L}$, and (c) $\mathcal{L}$ adjustment for a given $\mathcal{C}_{\mathrm{auto}}$-$\mathcal{C}_{\mathrm{accr}}$ pair. CNTL values are indicated by notches on the color bars.

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

An important question remains regarding the physical characteristics of the parameter space that produce the spurious strong-forcing mode. This is diagnosed by projecting emulated output variables onto the radiation-constrained posterior (Fig. 5). The stronger $\mathrm{ERF_{ACI}}$ values (≈ -0.95 W m^{-2}) are located exclusively within a "hook" region of high-$\mathcal{C}_{\mathrm{auto}}$/low-$\mathcal{C}_{\mathrm{accr}}$ (lower-right quadrant of Figs. 3–5). This reveals a paradox: the region of strongest radiative cooling (Fig. 5a) corresponds to the region with the lowest $\mathcal{L}$ (Fig. 5b), which is counter-intuitive since thicker clouds are generally more reflective. The paradox is resolved by the $\mathcal{L}$ adjustment from PI to PD (Fig. 5c), which is largest in this same hook region, confirming that the strong $\mathrm{ERF_{ACI}}$ is an artifact of the underlying model physics: the hook parameter combination creates a pre-industrial cloud state that is extremely fragile, producing a very large relative adjustment in $\mathcal{L}$ when anthropogenic aerosols are introduced. This same feedback was demonstrated by Feng et al. (2026) for simulations with increased autoconversion rates in both E3SMv2 (Golaz et al., 2022) and CESM2 (Danabasoglu et al., 2020).

This demonstrates that radiative state constraints alone are insufficient because they are blind to this underlying fragility, validating a plausible present-day state without testing whether that state results from a robust or pathological physical mechanism. These findings illustrate the core conclusion of this work: achieving a realistic emergent state is insufficient. Constraints must be powerful enough to select parameter sets that are physically robust across multiple climate states.

3.3 Process-Rate Proxies as Tractable Constraints

While direct process rates offer the most powerful constraint, their lack of observability motivates the search for tractable alternatives. We evaluate three process-oriented proxies retrievable by modern space-borne platforms: $\mathcal{P}$, WROF, and PAF (defined in Sect. 2.2). WROF acts as a maximally coarse-grained proxy for rain initiation (Mülmenstädt et al., 2021; Stanford et al., 2023) while PAF captures rain intensity by measuring the conditional frequency of heavy rain events; $\mathcal{P}$ represents the integrated outcome of both processes together with evaporation.

An evaluation of the emulated surfaces (Figs. 4, D1) and correlograms (Fig. D2) reveals that these proxies are not simple one-to-one alternatives for the direct process rates. Their behavior is conditioned on the background cloud state; for instance, WROF (Fig. 4d) is positively correlated with autoconversion but is also strongly anti-correlated with $\mathcal{L}$ (Fig. D2), underscoring that the proxies constrain the integrated behavior of the precipitation process rather than the underlying mechanisms themselves.

The combined proxy constraint (Fig. 4, cyan contours) shows maximum likelihoods near unity but also extending into the hook region, failing to break the model's fundamental degeneracy. This manifests in the $\mathrm{ERF_{ACI}}$ posteriors with proxy constraints (Fig. 6): at both 20% and 50% observational errors (panels b,c), the combined proxy constraint still produces a bimodal posterior, failing to eliminate the spurious strong-forcing mode near -0.95 W m^{-2}.

Despite this limitation, the proxies provide some value. At 20% observational error (Fig. 6b), the combined proxy constraint narrows the dominant $\mathrm{ERF_{ACI}}$ mode, approaching the performance of the direct process-rate constraint. A decomposition of individual proxies shows this is a synergistic effect, particularly at higher uncertainties, where the combination of all three provides a tighter constraint than any single proxy alone. These results confirm that while observable process-oriented proxies offer moderate constraining value, they remain insufficient to fully resolve the parametric equifinality that only direct process-rate constraints can eliminate.

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

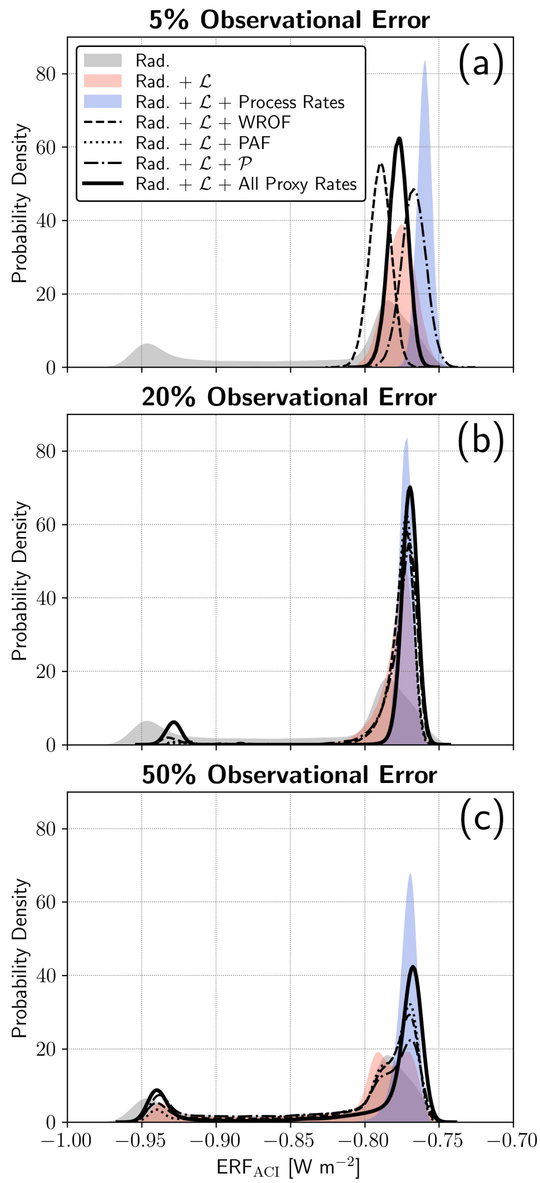

Figure 6. As in Fig. 2, but with the addition of constrained $\mathrm{ERF_{ACI}}$ distributions for individual proxy rates (WROF, PAF, and $\mathcal{P}$—dashed, dotted, and dash-dotted line styles, respectively) and their combination (solid line). Distributions constrained by radiation only (grey), radiation + $\mathcal{L}$ (red), and radiation + $\mathcal{L}$ + process rates (blue) are shown with enhanced transparency for reference.

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

4 Discussion

Our results demonstrate that equifinality in $\mathrm{ERF_{ACI}}$ is conditionally resolvable in a minimal working example, but only when constraints are anchored in the physical mechanisms that govern the emergent climate state. Constraining on a target TOA radiative state alone produces a fundamental degeneracy—a bimodal posterior reflecting two physically distinct but radiatively indistinguishable cloud regimes. Requiring the underlying cloud state to also be physically plausible substantially resolves this degeneracy. Additional cloud-state components beyond $\mathcal{L}$, such as droplet number concentration, may provide further constraint, and our multivariate MCMC framework is designed to leverage the physical correlations between such variables (Sect. 2.5). By combining constraints on radiative fluxes and $\mathcal{L}$, we demonstrate that the target radiative state is achieved with a realistic underlying cloud state—a step that is currently underutilized in standard model calibration but which our results identify as a critical intermediate constraint.

The efficacy of the combined cloud-state and radiative constraint is nonetheless sensitive to observational uncertainty. It remains robust at a realistic 20% uncertainty in $\mathcal{L}$—consistent with known systematic errors in satellite retrievals (O'Dell et al., 2008)—but fails as uncertainty approaches 50%. This highlights that continuing efforts to quantify uncertainty in observed $\mathcal{L}$ (e.g., using variability among retrieval products; Elsaesser et al., 2025) are crucial for robustly constraining models.

The only constraint configuration capable of fully resolving equifinality across all tested observational uncertainties combines all three tiers, with the process rates providing the critical final constraint that no emergent observable can replicate. This is because process rates are not emergent quantities—they reflect the local physical tendencies that drive the system rather than the integrated outcome of competing processes—and therefore carry information about the model's internal mechanisms that no emergent observable can supply. Crucially, this does not render emergent constraints redundant: the full resolution of equifinality requires their synergy, with radiation and cloud-state constraints eliminating the nonphysical self-depleted regime that process-rate constraints alone cannot exclude. The stark difference in constraining power is reflected in the likelihood surfaces (Fig. 4), where the process-rate contours exhibit a completely different orientation from all emergent constraints, reflecting a fundamentally different set of physical trade-offs that enable complete resolution of the bimodal posterior.

Despite this finding, a natural question is whether Tier 3 constitutes a tautological constraint—that is, whether constraining on vertically integrated process rates is trivially related to the multiplicative perturbations applied to those same rates. We argue it does not undermine our central findings for two reasons. First, the perturbed quantities ($\mathcal{C}_{\mathrm{auto}}$, $\mathcal{C}_{\mathrm{accr}}$) are multiplicative factors applied to local, instantaneous process rates, while the constrained quantities (AUTOCONV, ACCRETN) are vertically integrated, temporally averaged, and spatially aggregated over the stratocumulus regime. The mapping between these involves nonlinear feedbacks through cloud fraction, $\mathcal{L}$, and precipitation evaporation, and exhibits saturation at extreme perturbation values (Fig. D2). Second, and more fundamentally, Tier 3 is intended as a hypothetical best-case scenario—an upper bound on what process-level observational constraints could achieve—rather than a practically deployable calibration strategy. The substantial gap between this upper bound and what is achievable with emergent proxies is itself a primary quantitative result of this work, motivating the search for observational strategies that can approach this bound. We further note that the resolution of equifinality demonstrated here is specific to this idealized, low-dimensional configuration; in higher-dimensional parameter

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

spaces or when constraints are less directly related to the perturbed mechanisms, equifinality may prove considerably harder to resolve, and this remains an important open question.

The perfect-model approach is a deliberate feature of this study. By deriving synthetic observations from the same model used for calibration, we eliminate structural model error as a contributor to the equifinality we observe, demonstrating conclusively that it arises from the mathematical structure of the problem itself. Future work will relax this assumption by applying process-oriented constraints to real observations with realistic uncertainties (e.g., Elsaesser et al., 2025; Nugent et al., 2026; Fridlind et al., 2026). In these more realistic settings, both structural model error and observational bias are introduced, which may cause the degradation in constraining power observed between 20% and 50% observational error for the cloud-state constraint to occur at considerably lower error levels—making process-rate constraints potentially even more important than our idealized results suggest. Autoconversion and accretion rates have been estimated from observations in targeted settings—including aircraft-based integration of the stochastic collection equation (Wood, 2005) and synergistic radar-radiometer retrievals of coalescence rates (Stephens and Haynes, 2007)—and observationally constrained parameterizations of these rates have been developed from field campaign measurements and applied to improve climate model representations of warm rain (Chiu et al., 2021; Dong et al., 2021). However, these approaches remain limited to specific field campaigns or idealized retrieval conditions and do not yet provide the routine, large-scale process-rate estimates that a calibration framework like the one demonstrated here would require. This gap motivates the development of scalable retrieval algorithms and raises the fundamental question of whether synergistic combinations of observable, global proxies can ever approach the constraining power of the underlying process rates themselves.

The simplicity of our two-parameter design (an intentional deviation from high-dimensional calibration efforts, e.g., Elsaesser et al., 2025; Eidhammer et al., 2024) was essential for diagnosing the physical mechanisms underlying equifinality, enabling interpretation of the compensating parameter ridge and the diagnosis of the paradoxical regime in which the largest radiative cooling coincided with the least amount of cloud. This level of mechanistic insight is often obscured in higher-dimensional spaces where compensating parameter combinations are harder to visualize and diagnose. As the number of perturbed parameters grows, the equifinality problem becomes more severe and the need for process-level constraints more acute, yet the mechanistic interpretation of compensating parameters becomes correspondingly harder to extract. Iterative ensemble designs such as the Calibrated Physics Ensemble approach (Elsaesser et al., 2025) offer a tractable path forward by using initial constraint posteriors to guide subsequent ensemble sampling toward high-probability regions of parameter space. Additional tools that isolate the marginal influence of individual parameters on model output—such as partial dependence plots and individual conditional expectation curves (Goldstein et al., 2013; Hastie et al., 2009)—provide an analogous diagnostic capability in higher dimensions. The presented framework therefore demonstrates how PPEs can be used not just for model tuning but for diagnosing the parametric degeneracies that govern model uncertainty, a purpose that could be extended to calibrate the underlying physical parameters of the autoconversion and accretion schemes directly (Eqs. 1, 2) to further guide model development (e.g., Mikkelsen et al., 2025).

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

5 Conclusions

Equifinality—where multiple parameter combinations produce indistinguishable climate states—is a fundamental obstacle to constraining $\mathrm{ERF_{ACI}}$. Using a minimal working example with a two-parameter PPE and a hierarchical Bayesian calibration approach, we demonstrate that this degeneracy is conditionally resolvable within an idealized perfect-model framework and identify the observational constraints most capable of resolving it. Our results yield three primary conclusions.

First, constraining on a target TOA radiative state alone is insufficient. This canonical approach exposes a fundamental parametric degeneracy, resulting in a bimodal $\mathrm{ERF_{ACI}}$ posterior that includes a spurious, physically implausible solution arising from compensating parameter combinations that produce the correct cloud brightness via the wrong physical pathway.

Second, adding a constraint on the observable cloud state—here the liquid water path ($\mathcal{L}$)—is a powerful and robust subsequent step. It successfully eliminates the spurious mode under both optimistic (5%) and realistic (20%) observational uncertainty, but its utility collapses at large (50%) uncertainty, revealing a critical sensitivity where the constraining power of emergent state variables can fail.

Third, a direct constraint on the underlying physical processes is the most powerful approach. The full combination of constraints, anchored in mechanism-level information, is the only configuration capable of completely resolving the parametric degeneracy, providing the narrowest $\mathrm{ERF_{ACI}}$ uncertainty even under large observational error. Observable proxies for these processes offer some value but remain emergent properties and are ultimately insufficient to resolve the degeneracy.

Although demonstrated here in the context of warm-rain processes and $\mathrm{ERF_{ACI}}$, this hierarchy of constraint value is unlikely to be unique to this system. Anywhere that emergent climate states are governed by competing physical pathways, state-only constraints may be similarly insufficient. This work establishes a clear and generalizable principle for model evaluation and calibration: to build high-fidelity, unambiguous predictions, it is necessary to move beyond emergent constraints and towards constraining the fundamental physical mechanisms that produce them.

Code and data availability. We provide a Zenodo archival repository with the following: (1) the E3SMv3 code used; (2) instructions to reproduce the runs in the manuscript; (3) the Latin hypercube samples; (4) instructions to retrieve the data files produced by the simulations in this study, which have been publicly long-term archived using NERSC HPSS; (5) the analysis code, plotting code, emulators, and posterior chains required to reproduce the analyses and figures in the manuscript. The archival repository is available at https://doi.org/10.5281/zenodo.20739692 (Stanford and Mahfouz, 2026).

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

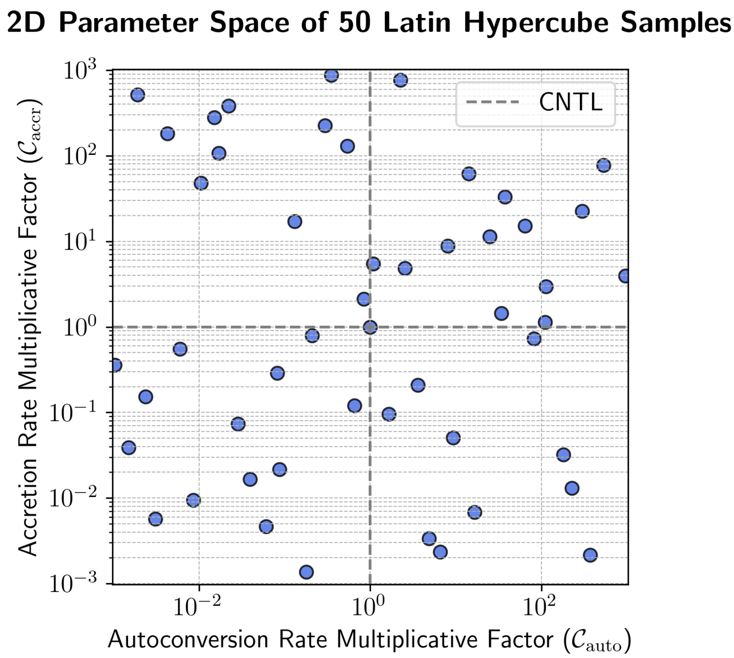

Figure A1. Two-dimensional parameter space for multiplicative factors applied to the autoconversion ($\mathcal{C}_{\mathrm{auto}}$, abscissa) and accretion ($\mathcal{C}_{\mathrm{accr}}$, ordinate) rates obtained via Latin hypercube sampling.

Appendix A: PPE and Experimental Design Diagnostics

The 50-member Latin hypercube sampling of the $\mathcal{C}_{\mathrm{auto}}$-$\mathcal{C}_{\mathrm{accr}}$ parameter space is shown in Fig. A1, illustrating the space-filling design that ensures adequate coverage of the full parameter range used to train the GP emulators (Sect. 2.3 and Appendix B). The output variables from the 50-member PPE form the unconstrained sparse prior, shown in Fig. A2 (orange). The dense prior (light blue) is generated by evaluating the production GP emulators at 10^6 randomly sampled points and provides a high-resolution view of the unconstrained output uncertainty prior to any constraint being applied. The spatial coverage of the stratocumulus regime used to compute cloud-state and process-rate variables is shown in Fig. A3, confirming that the regime criteria isolate the subtropical eastern ocean basins as expected and is consistent with Mülmenstädt et al. (2024b). Finally, the physical covariance structure between all PPE output variables is shown in Fig. A4. This correlation heatmap, computed from the 50-member PPE, reveals the inter-variable relationships that motivate the hybrid covariance matrix construction in the multivariate likelihood framework described in Sect. 2.5.

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

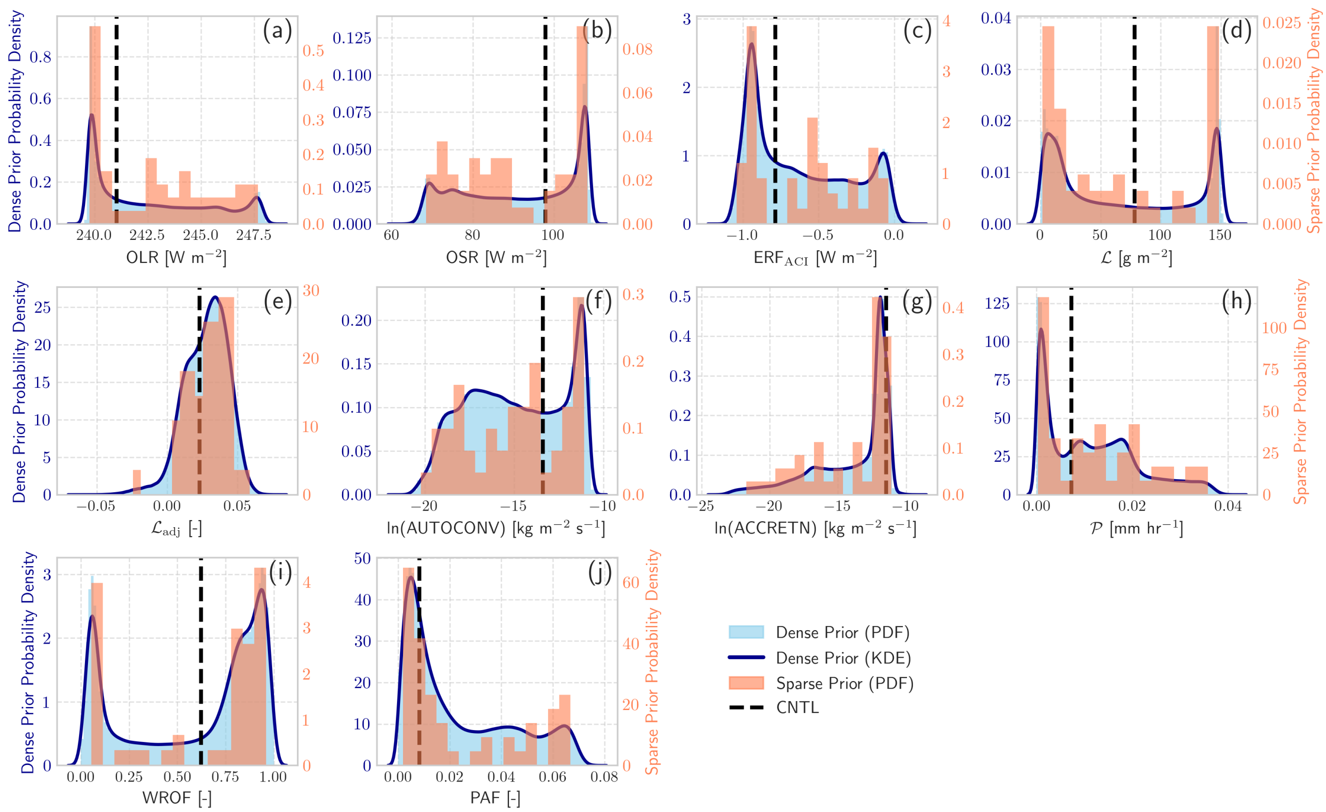

Figure A2. PDFs of the sparse prior distributions (orange, from the 50-member PPE) and 10^6 emulated samples produced using uniform random sampling (light blue) for all output variables. The dark blue line shows the KDE of the dense prior. The dashed black line indicates the CNTL value ($\mathcal{C}_{\text{auto}} = \mathcal{C}_{\text{accr}} = 1$).

Appendix B: Gaussian Process Emulators

B1 Kernel Specification and Hyperparameter Optimization

The GP emulator's predictive performance is determined by its covariance function, or kernel, $\mathcal{K}(x_i, x_j)$, which defines the similarity between any two points x_i and x_j. We use a composite kernel given by:

$$\mathcal{K}(x_i, x_j) = \sigma_f^2 \exp\left(-\frac{||x_i - x_j||^2}{2l^2}\right) + \sigma_n^2 \delta_{ij} \tag{B1}$$

where δ_{ij} is the Kronecker delta. The first term is a scaled Radial Basis Function (RBF) kernel that models the smooth, nonlinear relationship between the input parameters and model output; its length-scale l governs the smoothness of the emulated surface. The second term is a white noise kernel that accounts for small-scale variations in the training data and improves

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

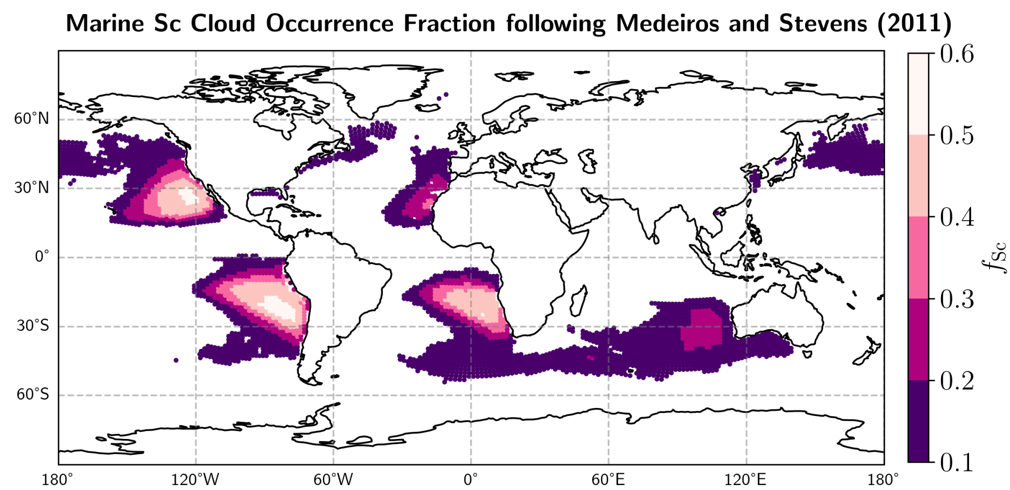

Figure A3. Cloud occurrence frequency map of stratocumulus (f_{Sc}), following the dynamical and thermodynamic regime definition of Medeiros and Stevens (2011). Grid points with $f_{\mathrm{Sc}} > 0.3$ are used for classifying overcast marine stratocumulus in this study following Mülmenstädt et al. (2024b).

numerical stability. The kernel hyperparameters—signal variance σ_f^2, length-scale l, and noise variance σ_n^2—are optimized during training by maximizing the log-marginal likelihood. In the `scikit-learn` implementation, this corresponds to a canonical `ConstantKernel` × `RBF` + `WhiteKernel` structure, with initial guesses of unity and broad, uninformative bounds on the hyperparameters to guide optimization.

B2 Emulator Validation

To ensure the GP emulators are accurate and reliable surrogates of the parent model, we performed a 5-fold cross-validation procedure, with results presented as parity plots in Fig. B1. For primary state variables (OLR, OSR, $\mathcal{L}$), R^2 exceeds 0.99, indicating the emulators explain more than 99% of the variance in the true model output, with out-of-sample predictions falling tightly along the 1:1 line and confirming the absence of systematic bias.

While R^2 values are high, several variables show slightly fewer than 95% of true values falling within the emulator's predicted 95% CI. This is partly a consequence of the small ensemble size—only three or more test points falling outside the CI are required to reduce the coverage below 94%. An examination of their locations reveals no single universal cause; instead,

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

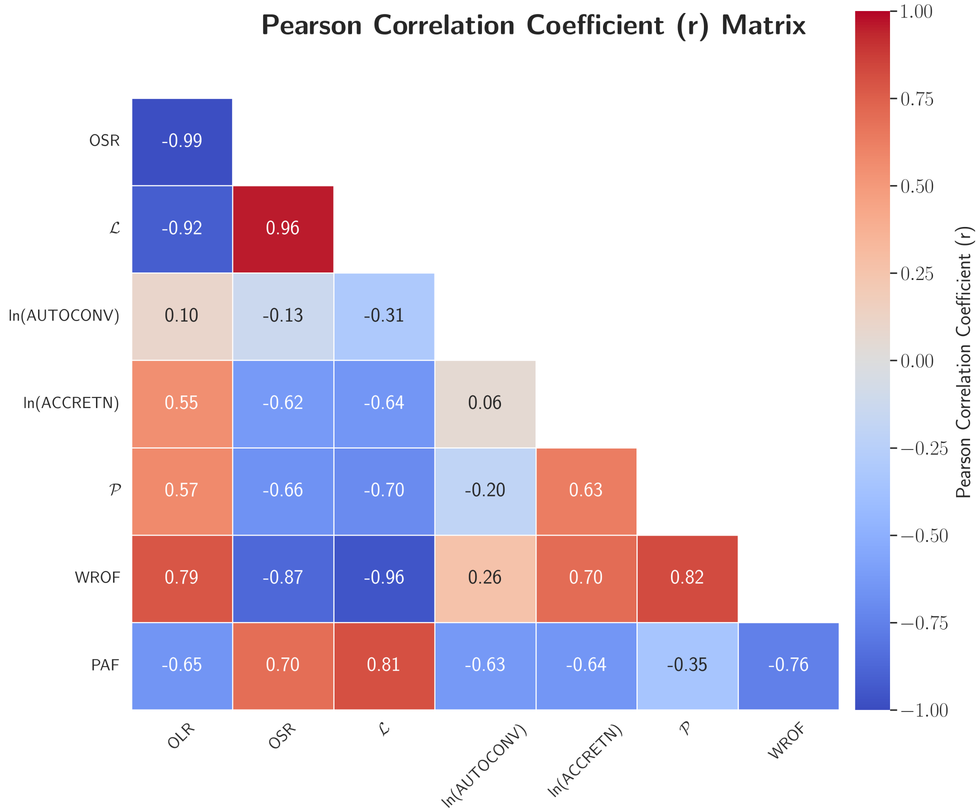

Figure A4. Correlogram of all constraint variables used in this study. Color corresponds to the Pearson correlation coefficient (r), with reds indicating positive and blues indicating negative correlations. Values $> |0.5|$ are displayed in white font and values $\leq |0.5|$ in black font.

they appear to arise from a combination of slight emulator overconfidence in densely sampled regions, sharp nonlinearities in a variable's physical response, or sampling gaps in a given cross-validation fold. Crucially, even for variables with lower CI coverage (e.g., $\mathcal{P}$, 84%), the vast majority of points demonstrate a strong linear relationship along the 1:1 line, indicating the emulator's predictive trend is sound.

A notable exception is $\mathcal{L}_{\mathrm{adj}}$ ($R^2 = 0.223$). The outlier points correspond to two physically pathological regimes: near-total cloud desiccation ($\mathcal{L} < 20$ g m^{-2}) and hyper-suppressed precipitation ($\mathcal{L} > 140$ g m^{-2}). Removing these six samples improves R^2 for the remaining well-behaved data to 0.74, confirming that poor performance is localized to physically unrealistic climate states at the extremes of the parameter space. For the purposes of this study, where $\mathcal{L}_{\mathrm{adj}}$ is used diagnostically to identify the cause of the spurious $\mathrm{ERF}_{\mathrm{ACI}}$ mode (Fig. 5c), absolute accuracy is less important than capturing the correct physical gradient across the parameter space, which the emulator does successfully.

Finally, we emphasize that the Bayesian MCMC framework is naturally robust to minor emulator imperfections. Unlike rejection sampling, the MCMC navigates a smooth likelihood surface such that slight emulator errors result only in a marginally

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

lower likelihood score, gently guiding walkers toward higher-probability regions. The MCMC process inherently focuses sampling effort in these high-probability regions, which are also the most densely populated by training data and where emulators are most accurate, naturally down-weighting the influence of emulator uncertainty from sparse, extreme regions of the parameter space.

GP emulator uncertainty maps for selected output variables are shown in Fig. B2. These maps illustrate the increase in predictive uncertainty towards the boundaries of the parameter space, motivating the prior bounds applied in the MCMC to avoid these regions (see Appendix C1). Fig. B2b also demonstrates the large uncertainty in $\mathcal{L}_{\mathrm{adj}}$ within regions not populated by the sparse prior samples.

Appendix C: MCMC Configuration and Diagnostics

C1 Sampler Configuration

The `emcee` sampler is configured with 100 walkers each taking 5000 steps; the first 20% of steps are discarded as burn-in, yielding 400,000 posterior samples. A stretch move scale factor of 3 is used to encourage larger exploratory steps suitable for the complex posterior surfaces encountered in this work (e.g., Figs. 3, 5).

We assume an informed initial placement of walkers. The valid parameter range is restricted to 2^{-8} to 2^{8} to avoid the high-uncertainty boundaries of the parameter space (Fig. B2). Walkers are initialized from a normal distribution centered on the scaled inputs (i.e., 0.5 in $[0,1]$ space) with a width such that 99.7% of walkers lie within $\pm 3\sigma$ of the mean, giving the MCMC a warm start that accelerates convergence relative to a randomly initialized distribution across the full $[0,1]$ space.

C2 MCMC Diagnostics

Convergence of the walker ensemble is confirmed via trace plots for the three primary MCMC configurations (Radiation Only, Radiation + $\mathcal{L}$, and Radiation + $\mathcal{L}$ + Process Rates) at 20% observational uncertainty, shown in Fig. C1. These confirm that walkers converge onto high-probability regions of the parameter space within the burn-in period and validate the robustness of the final posterior distributions.

Appendix D: Extended Likelihood Surfaces and Correlation Diagnostics

Figure D1 extends Fig. 4 to show emulator-predicted surfaces and 1-σ CI contours for all constraint variables, including the full suite of proxy rates. Figure D2 shows correlograms of all variables across the full radiation-only MCMC posterior, with points colored by their location in the $\mathcal{C}_{\mathrm{auto}}$-$\mathcal{C}_{\mathrm{accr}}$ quadrant. Together these figures support the interpretation of the proxy likelihood surfaces and the regime-dependent nonlinear relationships between proxies and direct process rates discussed in Sect. 3.3.

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

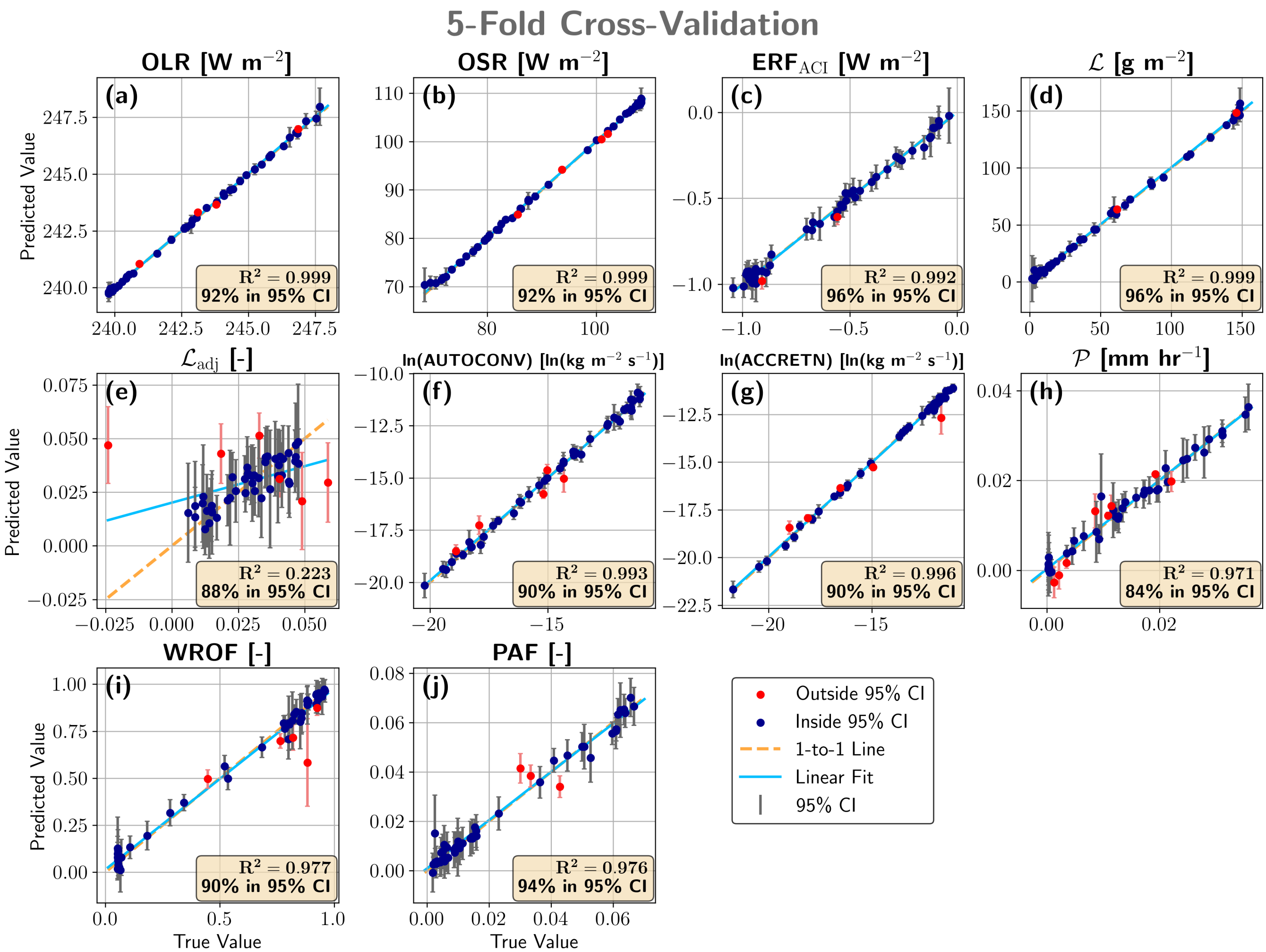

Figure B1. Parity plots of true values (abscissas) versus GP emulator predictions (ordinates) for all variables used in this study. Blue circles are samples within the 95% CI and red circles are those outside, where error bars represent the 95% CI. Text boxes display R^2 and the percentage of samples within the 95% CI. The 1:1 line is shown as an orange dashed line and the linear regression fit as a solid light blue line.

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

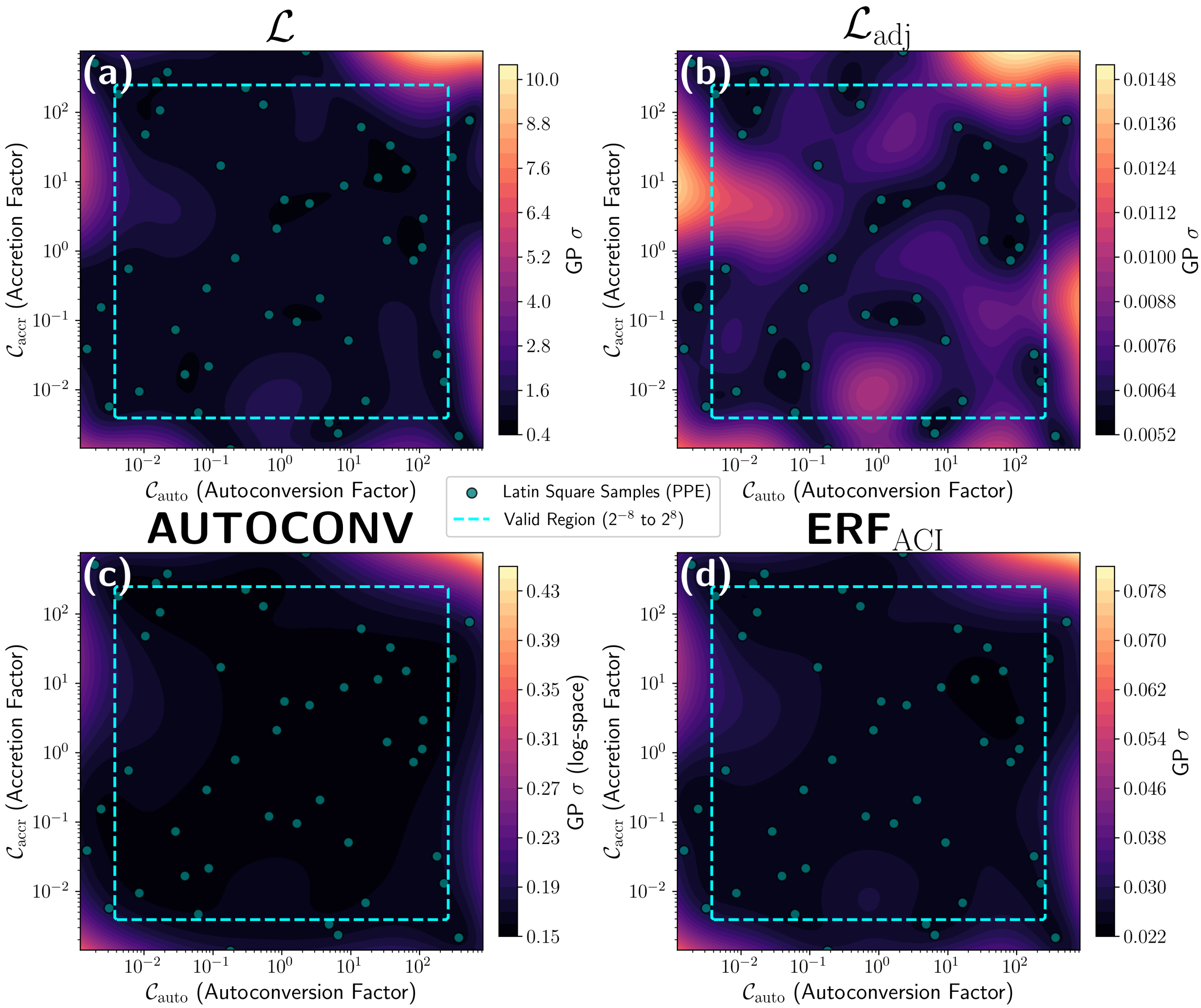

Figure B2. GP-predicted standard deviation (σ) as a function of $\mathcal{C}_{\mathrm{auto}}$ (abscissa) and $\mathcal{C}_{\mathrm{accr}}$ (ordinate) for selected output variables. Teal circles represent the 50-member PPE sampling locations (Fig. A1). Dashed cyan lines indicate the prior bounds supplied to the MCMC log-prior function to avoid the high-uncertainty boundaries of the parameter space (Appendix C1).

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

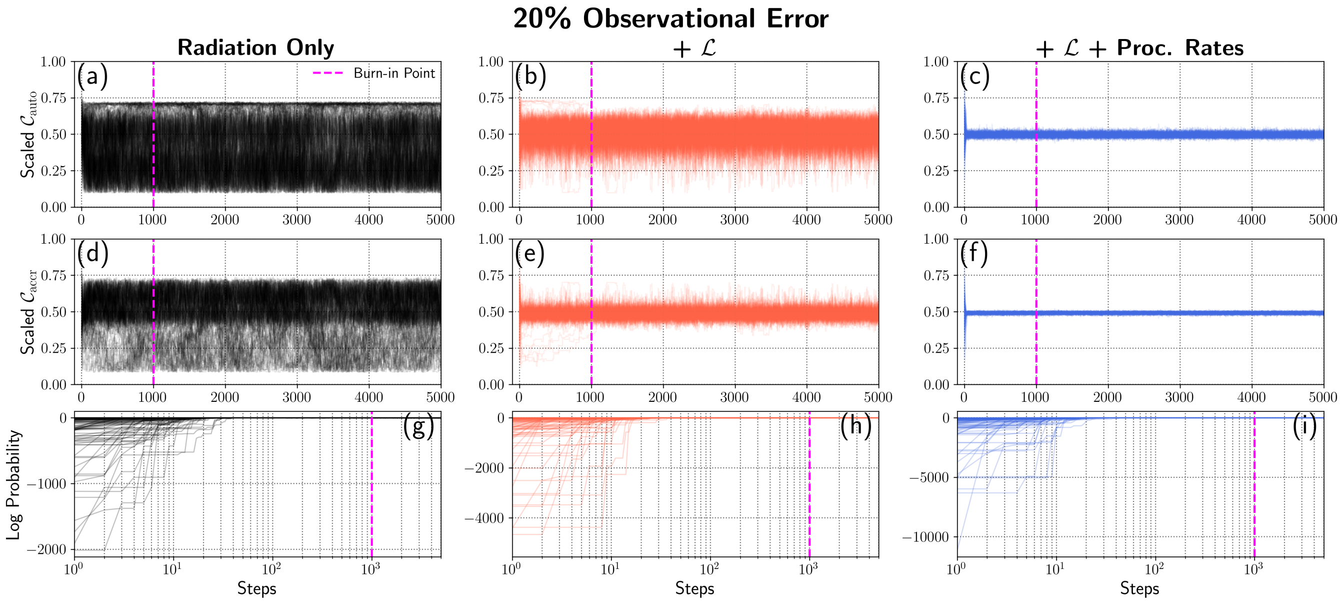

Figure C1. Trace plots for MCMC runs constraining on Radiation Only (left column), Radiation + $\mathcal{L}$ (middle column), and Radiation + $\mathcal{L}$ + Process Rates (right column) at 20% observational uncertainty (2 W m^{-2} for Radiation Only). The top and middle rows show walker exploration of $\mathcal{C}_{\mathrm{auto}}$ and $\mathcal{C}_{\mathrm{accr}}$, respectively; the bottom rows show the log-likelihood (Eq. 5) on a logarithmic step axis to highlight convergence behavior.

Author contributions. M.S., J.M., N.F., and S.B. designed the research; M.S. and J.M. performed the research; N.M. contributed new reagents/analytic tools; A.M. contributed methodological design; M.S. analyzed data; and M.S. wrote the paper.

470 *Competing interests.* The authors declare they have no competing interests.

Acknowledgements. We thank Duncan Watson-Paris and Sophie Wynn for helpful conversations on statistical emulation. Additional thanks are given to Daniel McCoy, Jacqueline Nugent, and Hunter Brown for providing computing allocation and aiding PPE logistics. We would like to acknowledge the use of the Derecho system (doi:10.5065/qx9a-pg09) and the Casper system ((https://ncar.pub/casper) supported by the NSF National Center for Atmospheric Research (NCAR) at the NSF NCAR-Wyoming Supercomputing Center, sponsored by the National 475 Science Foundation and the State of Wyoming. This was completed under allocation number WYOM-0191. This work was funded by the Laboratory Directed Research and Development (LDRD) program at Pacific Northwest National Laboratory (PNNL). PNNL is operated by Battelle for the DOE under Contract DE-AC05-76RLO1830. Additional computational resources were provided by the National Energy Research Scientific Computing Center (NERSC) user facility supported by the Office of Science of the U.S. Department of Energy under Contract DE798 AC02-05CH11231. AM's efforts were supported by Regional & Global Model Analysis (RGMA) DE-SC0025208. For

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

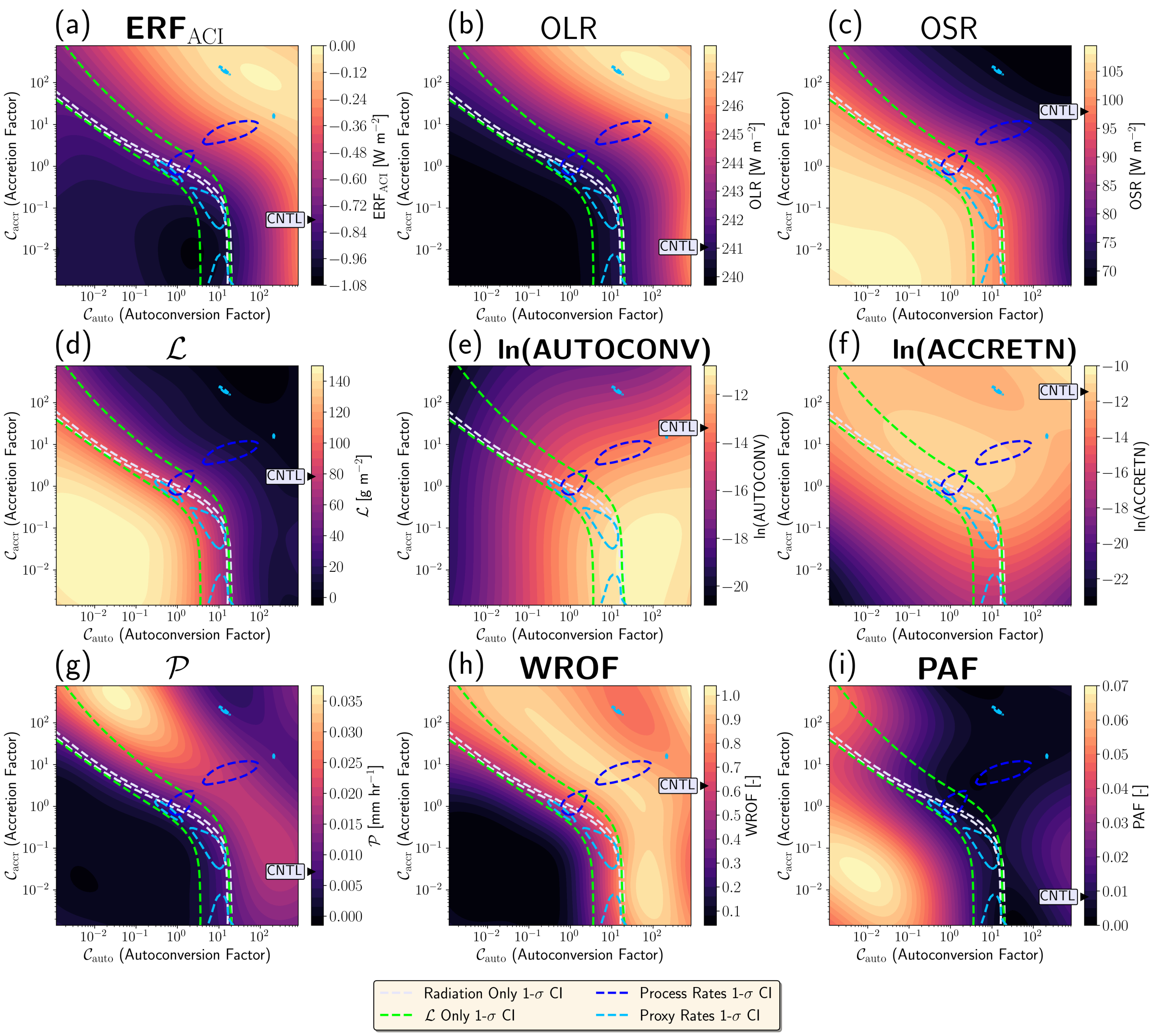

Figure D1. As in Fig. 4 but for all variables used as constraints in this study.

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

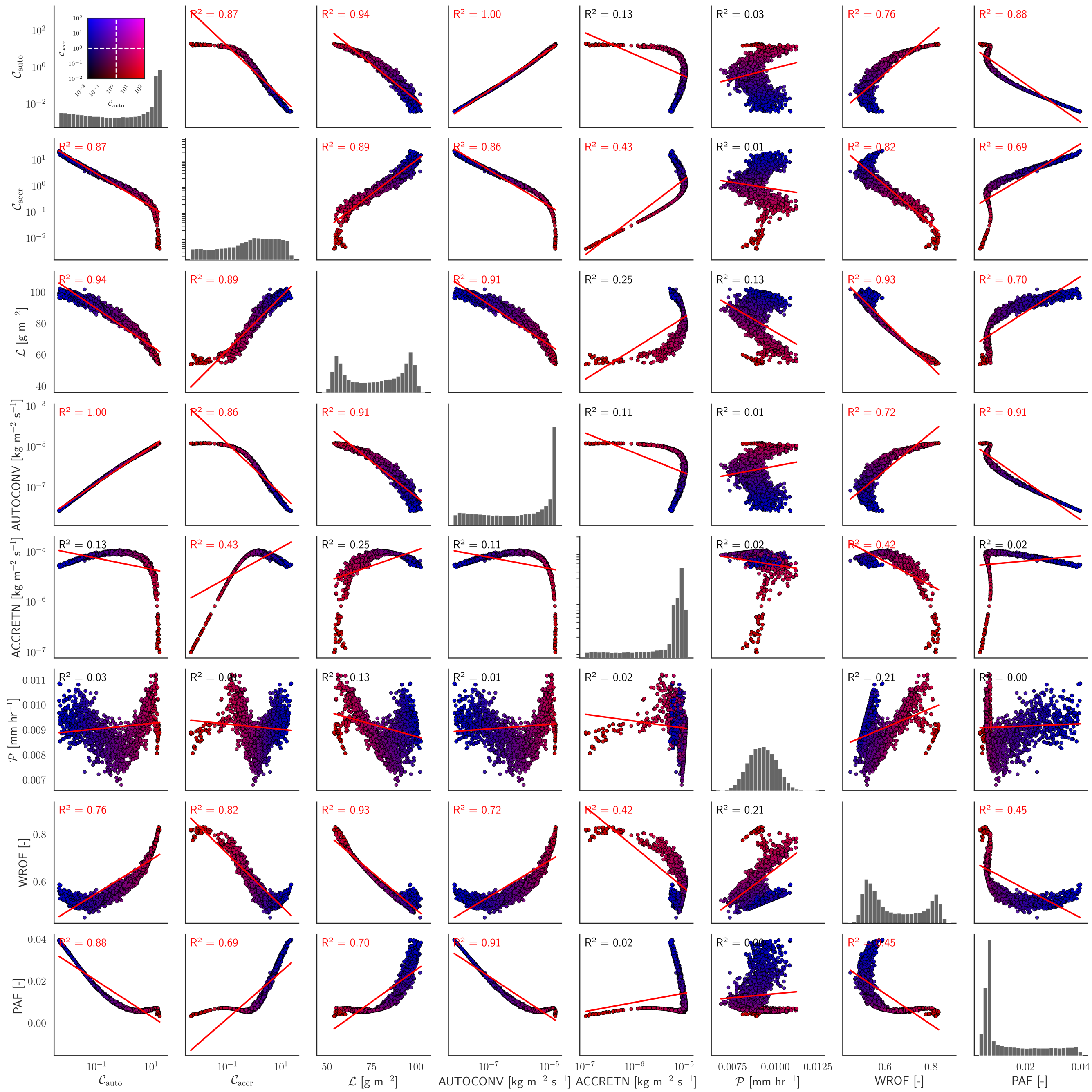

Figure D2. Correlograms of all variables used in this study, including inputs ($\mathcal{C}_{\mathrm{auto}}$, $\mathcal{C}_{\mathrm{accr}}$), for the 400,000 radiation-only MCMC posterior samples. Point color indicates the quadrant of the $\mathcal{C}_{\mathrm{auto}}$-$\mathcal{C}_{\mathrm{accr}}$ space as defined in the inset of the uppermost left panel. Lines and annotations show R^2 values, with red text indicating $R^2 > 0.3$.

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

transparency and following the position statement of the Committee on Publication Ethics (COPE), we acknowledge the assistance of large language models in aiding construction of processing code, interpretation of software, and providing grammatical and contextual refinements.

© Author(s) 2026. CC BY 4.0 License.

References

Ackerman, A. S., Kirkpatrick, M. P., Stevens, D. E., and Toon, O. B.: The impact of humidity above stratiform clouds on indirect aerosol climate forcing, Nature, 432, 1014–1017, https://doi.org/10.1038/NATURE03174, 2004.

Albrecht, B. A.: Aerosols, cloud microphysics, and fractional cloudiness, Science, 245, 1227–1230, https://doi.org/10.1126/SCIENCE.245.4923.1227, 1989.

Bellouin, N., Quaas, J., Gryspeerdt, E., Kinne, S., Stier, P., Watson-Parris, D., Boucher, O., Carslaw, K. S., Christensen, M., Daniau, A. L., Dufresne, J. L., Feingold, G., Fiedler, S., Forster, P., Gettelman, A., Haywood, J. M., Lohmann, U., Malavelle, F., Mauritsen, T., McCoy, D. T., Myhre, G., Mülmenstädt, J., Neubauer, D., Possner, A., Rugenstein, M., Sato, Y., Schulz, M., Schwartz, S. E., Sourdeval, O., Storelvmo, T., Toll, V., Winker, D., and Stevens, B.: Bounding Global Aerosol Radiative Forcing of Climate Change, Reviews of Geophysics, 58, e2019RG000 660, https://doi.org/10.1029/2019RG000660;REQUESTEDJOURNAL:JOURNAL:19449208;ISSUE:ISSUE:DOI, 2020.

Beven, K.: A manifesto for the equifinality thesis, Journal of Hydrology, 320, 18–36, https://doi.org/10.1016/J.JHYDROL.2005.07.007, 2006.

Beven, K. and Binley, A.: The future of distributed models: Model calibration and uncertainty prediction, Hydrological Processes, 6, 279–298, https://doi.org/10.1002/HYP.3360060305;REQUESTEDJOURNAL:JOURNAL:10991085;WGROUP:STRING:PUBLICATION, 1992.

Bing Lin and Rossow, W. B.: Observations of cloud liquid water path over oceans: optical and microwave remote sensing methods, Journal of Geophysical Research, 99, https://doi.org/10.1029/94JD01831, 1994.

Bony, S., Stevens, B., Frierson, D. M., Jakob, C., Kageyama, M., Pincus, R., Shepherd, T. G., Sherwood, S. C., Siebesma, A. P., Sobel, A. H., Watanabe, M., and Webb, M. J.: Clouds, circulation and climate sensitivity, Nature Geoscience 2015 8:4, 8, 261–268, https://doi.org/10.1038/ngeo2398, 2015.

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V.-M., Kondo, Y., Liao, H., Lohmann, U., Rasch, P., Satheesh, S. K., Sherwood, S., Stevens, B., and Zhang, X. Y.: Clouds and Aerosols, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, chap. 7, pp. 571–658, Cambridge Univ. Press, Cambridge, United Kingdom and New York, NY, USA, 2014.

Bretherton, C. S., Blossey, P. N., and Uchida, J.: Cloud droplet sedimentation, entrainment efficiency, and subtropical stratocumulus albedo, Geophysical Research Letters, 34, 3813, https://doi.org/10.1029/2006GL027648, 2007.

Carslaw, K. S., Lee, L. A., Reddington, C. L., Pringle, K. J., Rap, A., Forster, P. M., Mann, G. W., Spracklen, D. V., Woodhouse, M. T., Regayre, L. A., and Pierce, J. R.: Large contribution of natural aerosols to uncertainty in indirect forcing, Nature, 503, 67–71, https://doi.org/10.1038/nature12674, 2013.

Carslaw, K. S., Regayre, L. A., Proske, U., Gettelman, A., Sexton, D. M. H., Qian, Y., Marshall, L. R., Wild, O., van Lier-Walqui, M., Oertel, A., Peatier, S., Yang, B., Johnson, J. S., Li, S., McCoy, D. T., Sanderson, B. M., Williamson, C. J., Elsaesser, G. S., Yamazaki, K., and Booth, B. B. B.: Opinion: The importance and future development of perturbed parameter ensembles in climate and atmospheric science, Atmospheric Chemistry and Physics, 26, 4651–4667, https://doi.org/10.5194/ACP-26-4651-2026, 2026.

Chiu, J. C., Yang, C. K., van Leeuwen, P. J., Feingold, G., Wood, R., Blanchard, Y., Mei, F., and Wang, J.: Observational Constraints on Warm Cloud Microphysical Processes Using Machine Learning and Optimization Techniques, Geophysical Research Letters, 48, e2020GL091 236, https://doi.org/10.1029/2020GL091236;REQUESTEDJOURNAL:JOURNAL:19448007;ISSUE:ISSUE:DOI, 2021.

© Author(s) 2026. CC BY 4.0 License.

Cho, S. J., Lund, J. W., Ball, G., Brown, J., Gellis, A., Gurley, L., Hamshaw, S., Kwang, J., Laws, A., Noe, G., Oelsner, G., Parchaso, F., Phipps, C., Skalak, K., and Sutfin, N.: Disentangling geomorphic equifinality in sediment and hydrologic connectivity through the analyses of landscape drivers of hysteresis, Earth Surface Processes and Landforms, 50, e70 176, https://doi.org/10.1002/ESP.70176;ISSUE:ISSUE:DOI, 2025.

Clough, S. A., Shephard, M. W., Mlawer, E. J., Delamere, J. S., Iacono, M. J., Cady-Pereira, K., Boukabara, S., and Brown, P. D.: Atmospheric radiative transfer modeling: a summary of the AER codes, Journal of Quantitative Spectroscopy and Radiative Transfer, 91, 233–244, https://doi.org/10.1016/J.JQSRT.2004.05.058, 2005.

Danabasoglu, G., Lamarque, J. F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., Emmons, L. K., Fasullo, J., Garcia, R., Gettelman, A., Hannay, C., Holland, M. M., Large, W. G., Lauritzen, P. H., Lawrence, D. M., Lenaerts, J. T., Lindsay, K., Lipscomb, W. H., Mills, M. J., Neale, R., Oleson, K. W., Otto-Bliesner, B., Phillips, A. S., Sacks, W., Tilmes, S., van Kampenhout, L., Vertenstein, M., Bertini, A., Dennis, J., Deser, C., Fischer, C., Fox-Kemper, B., Kay, J. E., Kinnison, D., Kushner, P. J., Larson, V. E., Long, M. C., Mickelson, S., Moore, J. K., Nienhouse, E., Polvani, L., Rasch, P. J., and Strand, W. G.: The Community Earth System Model Version 2 (CESM2), Journal of Advances in Modeling Earth Systems, 12, e2019MS001 916, https://doi.org/10.1029/2019MS001916;WGROUP:STRING:PUBLICATION, 2020.

Dong, X., Wu, P., Wang, Y., Xi, B., and Huang, Y.: New Observational Constraints on Warm Rain Processes and Their Climate Implications, Geophysical Research Letters, 48, e2020GL091 836, https://doi.org/10.1029/2020GL091836;REQUESTEDJOURNAL:JOURNAL:19448007;WGROUP:STRING:PUBLICATION, 2021.

E3SM Project, D.: Energy Exascale Earth System Model v3.0.0. Computer Software., https://doi.org/doi:10.11578/E3SM/dc.20240301.3, 2024.

Eidhammer, T., Gettelman, A., Thayer-Calder, K., Watson-Parris, D., Elsaesser, G., Morrison, H., Van Lier-Walqui, M., Song, C., and McCoy, D.: An extensible perturbed parameter ensemble for the Community Atmosphere Model version 6, Geoscientific Model Development, 17, 7835–7853, https://doi.org/10.5194/GMD-17-7835-2024,, 2024.

Elsaesser, G. S., van Lier-Walqui, M., Yang, Q., Kelley, M., Ackerman, A. S., Fridlind, A. M., Cesana, G. V., Schmidt, G. A., Wu, J., Behrangi, A., Camargo, S. J., De, B., Inoue, K., Leitmann-Niimi, N. M., and Strong, J. D.: Using Machine Learning to Generate a GISS ModelE Calibrated Physics Ensemble (CPE), Journal of Advances in Modeling Earth Systems, 17, e2024MS004 713, https://doi.org/10.1029/2024MS004713;WGROUP:STRING:PUBLICATION, 2025.

Feng, C., Liu, X., Zhao, X., Lu, Z., Lin, L., Zelinka, M. D., Qin, Y., Shan, Y., Zheng, Y., Wu, X., and Saravanan, R.: Interconnection of Aerosol-Cloud Interactions and Cloud Feedback Through Warm Rain Process, Journal of Geophysical Research: Atmospheres, 131, e2025JD044 854, https://doi.org/10.1029/2025JD044854;SUBPAGE:STRING:FULL, 2026.

Foreman-Mackey, D., Hogg, D. W., Lang, D., and Goodman, J.: emcee: The MCMC Hammer, Publications of the Astronomical Society of the Pacific, 125, 306–312, https://doi.org/10.1086/670067, 2013.

Forster, P., Storelvmo, K., Armour, K., Collins, W., Dufresne, J.-L., Frame, D., Lunt, D., Mauritsen, T., Palmer, M., Watanabe, M., Wild, M., and Zhang, H.: The Earth's Energy Budget, Climate Feedbacks and Climate Sensitivity, in: Climate Change 2021: The Physical Science Basis. Contribution of Work Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, chap. 7, pp. 923–1054, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, https://doi.org/https://doi.org/10.1017/9781009157896.009, 2021.

Fridlind, A. M., Elsaesser, G. S., Lier-Walqui, M. v., Cesana, G. V., Weatherhead, E., Tselioudis, G., Schmidt, G., Barahona, D., Cairns, B., Collins, W. D., Considine, D. B., Cucurull, L., DiGirolamo, L., Emory, A., Hasekamp, O., He, S., Kramer, R., Lebsock, M., Lee, T., Leroy,

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

S., Lin, W., Lugauer, S., Miller, D., Mülmenstädt, J., Oreopoulos, L., Posselt, D. J., Zelinka, M. D., Fridlind, A. M., Elsaesser, G. S., Lier-Walqui, M. v., Cesana, G. V., Weatherhead, E., Tselioudis, G., Schmidt, G., Barahona, D., Cairns, B., Collins, W. D., Considine, D. B., Cucurull, L., DiGirolamo, L., Emory, A., Hasekamp, O., He, S., Kramer, R., Lebsock, M., Lee, T., Leroy, S., Lin, W., Lugauer, S., Miller, D., Mülmenstädt, J., Oreopoulos, L., Posselt, D. J., and Zelinka, M. D.: Toward a Climate OSSE Framework for Satellite Mission Design, Bulletin of the American Meteorological Society, 107, E431–E445, https://doi.org/10.1175/BAMS-D-24-0242.1, 2026.

Gettelman, A.: Putting the clouds back in aerosol-cloud interactions, Atmospheric Chemistry and Physics, 15, 12 397–12 411, https://doi.org/10.5194/ACP-15-12397-2015,, 2015.

Ghan, S. J.: Technical note: Estimating aerosol effects on cloud radiative forcing, Atmospheric Chemistry and Physics, 13, 9971–9974, https://doi.org/10.5194/ACP-13-9971-2013, 2013.

Golaz, J.-C., Larson, V. E., and Cotton, W. R.: A PDF-Based Model for Boundary Layer Clouds. Part I: Method and Model Description, Journal of the Atmospheric Sciences, 59, 3540 – 3551, https://doi.org/10.1175/1520-0469(2002)059<3540:APBMFB>2.0.CO;2, 2002.

Golaz, J. C., Golaz, J. C., and Levy, H.: Cloud tuning in a coupled climate model: Impact on 20th century warming, Geophysical Research Letters, 40, 2246–2251, https://doi.org/10.1002/GRL.50232;WEBSITE:WEBSITE:AGUPUBS;JOURNAL:JOURNAL:19448007;ISSUE:ISSUE:DOI, 2013.

Golaz, J. C., Van Roekel, L. P., Zheng, X., Roberts, A. F., Wolfe, J. D., Lin, W., Bradley, A. M., Tang, Q., Maltrud, M. E., Forsyth, R. M., Zhang, C., Zhou, T., Zhang, K., Zender, C. S., Wu, M., Wang, H., Turner, A. K., Singh, B., Richter, J. H., Qin, Y., Petersen, M. R., Mametjanov, A., Ma, P. L., Larson, V. E., Krishna, J., Keen, N. D., Jeffery, N., Hunke, E. C., Hannah, W. M., Guba, O., Griffin, B. M., Feng, Y., Engwirda, D., Di Vittorio, A. V., Dang, C., Conlon, L. A. M., Chen, C. C. J., Brunke, M. A., Bisht, G., Benedict, J. J., Asay-Davis, X. S., Zhang, Y., Zhang, M., Zeng, X., Xie, S., Wolfram, P. J., Vo, T., Veneziani, M., Tesfa, T. K., Sreepathi, S., Salinger, A. G., Reeves Eyre, J. E., Prather, M. J., Mahajan, S., Li, Q., Jones, P. W., Jacob, R. L., Huebler, G. W., Huang, X., Hillman, B. R., Harrop, B. E., Foucar, J. G., Fang, Y., Comeau, D. S., Caldwell, P. M., Bartoletti, T., Balaguru, K., Taylor, M. A., McCoy, R. B., Leung, L. R., and Bader, D. C.: The DOE E3SM Model Version 2: Overview of the Physical Model and Initial Model Evaluation, Journal of Advances in Modeling Earth Systems, 14, e2022MS003 156, https://doi.org/10.1029/2022MS003156;WGROUP:STRING:PUBLICATION, 2022.

Goldstein, A., Kapelner, A., Bleich, J., and Pitkin, E.: Peeking Inside the Black Box: Visualizing Statistical Learning with Plots of Individual Conditional Expectation, Journal of Computational and Graphical Statistics, 24, 44–65, https://doi.org/10.1080/10618600.2014.907095, 2013.

Goodman, J. and Weare, J.: ENSEMBLE SAMPLERS WITH AFFINE INVARIANCE, COMM. APP. MATH. AND COMP. SCI, 5, 2010.

Goren, T. and Rosenfeld, D.: Decomposing aerosol cloud radiative effects into cloud cover, liquid water path and Twomey components in marine stratocumulus, Atmospheric Research, 138, 378–393, https://doi.org/10.1016/J.ATMOSRES.2013.12.008, 2014.

Greenwald, T. J., Bennartz, R., Lebsock, M., and Teixeira, J.: An Uncertainty Data Set for Passive Microwave Satellite Observations of Warm Cloud Liquid Water Path, Journal of Geophysical Research: Atmospheres, 123, 3668–3687, https://doi.org/10.1002/2017JD027638;SUBPAGE:STRING:FULL, 2018.

Gresov, C. and Drazin, R.: Equifinality: Functional Equivalence in Organization Design, The Academy of Management Review, 22, 403, https://doi.org/10.2307/259328, 1997.

Hastie, T., Tibshirani, R., and Friedman, J.: Section 10.13.2, in: The Elements of Statistical Learning, Springer Series in Statistics, chap. Section 10.13.2, pp. XXII–745, Springer, New York, NY, 2nd edn., ISBN 978-0-387-84857-0, https://doi.org/10.1007/978-0-387-84858-7, 2009.

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

Hourdin, F., Mauritsen, T., Gettelman, A., Golaz, J. C., Balaji, V., Duan, Q., Folini, D., Ji, D., Klocke, D., Qian, Y., Rauser, F., Rio, C., Tomassini, L., Watanabe, M., and Williamson, D.: The Art and Science of Climate Model Tuning, Bulletin of the American Meteorological Society, 98, 589–602, https://doi.org/10.1175/BAMS-D-15-00135.1, 2017.

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D.: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, Journal of Geophysical Research, 113, D13 103, https://doi.org/10.1029/2008JD009944, 2008.

Khairoutdinov, M. and Kogan, Y.: A New Cloud Physics Parameterization in a Large-Eddy Simulation Model of Marine Stratocumulus, Monthly Weather Review, 128, 229 – 243, https://doi.org/10.1175/1520-0493(2000)128<0229:ANCPPI>2.0.CO;2, 2000.

Klein, S. A. and Hall, A.: Emergent Constraints for Cloud Feedbacks, Current Climate Change Reports 2015 1:4, 1, 276–287, https://doi.org/10.1007/S40641-015-0027-1, 2015.

Knutti, R., Sedláček, J., Sanderson, B. M., Lorenz, R., Fischer, E. M., and Eyring, V.: A climate model projection weighting scheme accounting for performance and interdependence, Geophysical Research Letters, 44, 1909–1918, https://doi.org/10.1002/2016GL072012;WGROUP:STRING:PUBLICATION, 2017.

Larson, V. E.: CLUBB-SILHS: A parameterization of subgrid variability in the atmosphere, https://arxiv.org/pdf/1711.03675, 2017.

Larson, V. E., Golaz, J.-C., Larson, V. E., and Golaz, J.-C.: Using Probability Density Functions to Derive Consistent Closure Relationships among Higher-Order Moments, Monthly Weather Review, 133, 1023–1042, https://doi.org/10.1175/MWR2902.1, 2005.

Lebsock, M. D. and L'Ecuyer, T. S.: The retrieval of warm rain from CloudSat, Journal of Geophysical Research: Atmospheres, 116, 20 209, https://doi.org/10.1029/2011JD016076, 2011.

L'Ecuyer, T. S., Beaudoing, H. K., Rodell, M., Olson, W., Lin, B., Kato, S., Clayson, C. A., Wood, E., Sheffield, J., Adler, R., Huffman, G., Bosilovich, M., Gu, G., Robertson, F., Houser, P. R., Chambers, D., Famiglietti, J. S., Fetzer, E., Liu, W. T., Gao, X., Schlosser, C. A., Clark, E., Lettenmaier, D. P., and Hilburn, K.: The Observed State of the Energy Budget in the Early Twenty-First Century, Journal of Climate, 28, 8319–8346, https://doi.org/10.1175/JCLI-D-14-00556.1, 2015.

Lee, L. A., Reddington, C. L., and Carslaw, K. S.: On the relationship between aerosol model uncertainty and radiative forcing uncertainty, Proceedings of the National Academy of Sciences of the United States of America, 113, 5820–5827, https://doi.org/10.1073/PNAS.1507050113/SUPPL_FILE/PNAS.201507050SI.PDF, 2016.

Luo, Y., Weng, E., Wu, X., Gao, C., Zhou, X., and Zhang, L.: Parameter identifiability, constraint, and equifinality in data assimilation with ecosystem models, Ecological Applications, 19, 571–574, https://doi.org/10.1890/08-0561.1, 2009.

Mahfouz, N., Mülmenstädt, J., and Burrows, S.: Present-day correlations are insufficient to predict cloud albedo change by anthropogenic aerosols in E3SM v2, Atmospheric Chemistry and Physics, 24, 7253–7260, https://doi.org/10.5194/ACP-24-7253-2024,, 2024.

Mahfouz, N., Beydoun, H., Mülmenstädt, J., Keen, N., Varble, A. C., Bertagna, L., Bogenschutz, P., Bradley, A., Christensen, M. W., Clevenger, T. C., Donahue, A., Fast, J., Foucar, J., Golaz, J.-C., Guba, O., Hannah, W., Hillman, B., Jacob, R., Lin, W., Ma, P.-L., Qian, Y., Singh, B., Terai, C., Wang, H., Wu, M., Zhang, K., Gettelman, A., Taylor, M., Leung, L. R., Caldwell, P., and Burrows, S.: Prescribing the aerosol effective radiative forcing in the Simple Cloud-Resolving E3SM Atmosphere Model v1, https://doi.org/10.5194/EGUSPHERE-2025-1868, 2025.

Marder, E. and Taylor, A. L.: Multiple models to capture the variability in biological neurons and networks, Nature neuroscience, 14, 133, https://doi.org/10.1038/NN.2735, 2011.

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

Mauritsen, T., Stevens, B., Roeckner, E., Crueger, T., Esch, M., Giorgetta, M., Haak, H., Jungclaus, J., Klocke, D., Matei, D., Mikolajewicz, U., Notz, D., Pincus, R., Schmidt, H., and Tomassini, L.: Tuning the climate of a global model, Journal of Advances in Modeling Earth Systems, 4, 0–01, https://doi.org/10.1029/2012MS000154;REQUESTEDJOURNAL:JOURNAL:19422466, 2012.

Medeiros, B. and Stevens, B.: Revealing differences in GCM representations of low clouds, Climate Dynamics 2009 36:1, 36, 385–399, https://doi.org/10.1007/s00382-009-0694-5, 2011.

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

Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., and Clough, S. A.: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave, Journal of Geophysical Research Atmospheres, 102, 16 663–16 682, https://doi.org/10.1029/97JD00237;WGROUP:STRING:PUBLICATION, 1997.

Morrison, H. and Milbrandt, J. A.: Parameterization of Cloud Microphysics Based on the Prediction of Bulk Ice Particle Properties. Part I: Scheme Description and Idealized Tests, Journal of the Atmospheric Sciences, 72, 287–311, https://doi.org/10.1175/JAS-D-14-0065.1, 2015.

Morrison, H., Gettelman, A., Morrison, H., and Gettelman, A.: A New Two-Moment Bulk Stratiform Cloud Microphysics Scheme in the Community Atmosphere Model, Version 3 (CAM3). Part I: Description and Numerical Tests, Journal of Climate, 21, 3642–3659, https://doi.org/10.1175/2008JCLI2105.1, 2008.

Mülmenstädt, J. and Feingold, G.: The Radiative Forcing of Aerosol–Cloud Interactions in Liquid Clouds: Wrestling and Embracing Uncertainty, Current Climate Change Reports 2018 4:1, 4, 23–40, https://doi.org/10.1007/S40641-018-0089-Y, 2018.

Mülmenstädt, J., Salzmann, M., Kay, J. E., Zelinka, M. D., Ma, P. L., Nam, C., Kretzschmar, J., Hörnig, S., and Quaas, J.: An underestimated negative cloud feedback from cloud lifetime changes, Nature Climate Change, 11, 508–513, https://doi.org/10.1038/S41558-021-01038-1;SUBJMETA=106,1108,2786,35,694,704;KWRD=ATMOSPHERIC+SCIENCE,CLIMATE+AND+EARTH+SYSTEM+MODELLING, PROJECTION+AND+PREDICTION, 2021.

Mülmenstädt, J., Ackerman, A. S., Fridlind, A. M., Huang, M., Ma, P. L., Mahfouz, N., Bauer, S. E., Burrows, S. M., Christensen, M. W., Dipu, S., Gettelman, A., Leung, L. R., Tornow, F., Quaas, J., Varble, A. C., Wang, H., Zhang, K., and Zheng, Y.: Can general circulation models (GCMs) represent cloud liquid water path adjustments to aerosol–cloud interactions?, Atmospheric Chemistry and Physics, 24, 13 633–13 652, https://doi.org/10.5194/ACP-24-13633-2024, 2024a.

Mülmenstädt, J., Gryspeerdt, E., Dipu, S., Quaas, J., Ackerman, A. S., Fridlind, A. M., Tornow, F., Bauer, S. E., Gettelman, A., Ming, Y., Zheng, Y., Ma, P. L., Wang, H., Zhang, K., Christensen, M. W., Varble, A. C., Ruby Leung, L., Liu, X., Neubauer, D., Partridge, D. G., Stier, P., and Takemura, T.: General circulation models simulate negative liquid water path-droplet number correlations, but anthropogenic aerosols still increase simulated liquid water path, Atmospheric Chemistry and Physics, 24, 7331–7345, https://doi.org/10.5194/ACP-24-7331-2024,, 2024b.

Nicholas, A. P. and Quine, T. A.: Quantitative assessment of landform equifinality and palaeoenvironmental reconstruction using geomorphic models, Geomorphology, 121, 167–183, https://doi.org/10.1016/J.GEOMORPH.2010.04.004, 2010.

Nugent, J. M., Brown, H., Kirby, A., McCoy, D. T., Allen, G., Aerenson, T., Burrows, S. M., Caulton, D., Fan, J., Feng, Y., Gettelman, A., Griswold, J., Jones, D. B., Leung, L. R., Mahfouz, N., Mikkelsen, A., Mülmenstädt, J., Qian, Y., Shan, Y., Shpund, J., Silber, I., Song, C., Song, X., Wang, H., Wu, M., Xie, S., Zelinka, M. D., Zhang, D., Zhang, G. J., and Zhang, K.: Overview of the Nephele Perturbed

https://doi.org/10.5194/egusphere-2026-4488
Preprint. Discussion started: 12 August 2026
© Author(s) 2026. CC BY 4.0 License.

Parameter Ensemble for Aerosol-Cloud Interactions in E3SMv3, Journal of Advances in Modeling Earth Systems, 18, e2025MS004 989, https://doi.org/10.1029/2025MS004989, 2026.

O'Dell, C. W., Wentz, F. J., and Bennartz, R.: Cloud Liquid Water Path from Satellite-Based Passive Microwave Observations: A New Climatology over the Global Oceans, Journal of Climate, 21, 1721–1739, https://doi.org/10.1175/2007JCLI1958.1, 2008.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, B., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E.: Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research, 12, 2825–2830, 2011.

Prinz, A. A., Bucher, D., and Marder, E.: Similar network activity from disparate circuit parameters, Nature Neuroscience 2004 7:12, 7, 1345–1352, https://doi.org/10.1038/nn1352, 2004.

Quaas, J., Ming, Y., Menon, S., Takemura, T., Wang, M., Penner, J. E., Gettelman, A., Lohmann, U., Bellouin, N., Boucher, O., Sayer, A. M., Thomas, G. E., McComiskey, A., Feingold, G., Hoose, C., Kristj´nsson, J. E., Liu, X., Balkanski, Y., Donner, L. J., Ginoux, P. A., Stier, P., Grandey, B., Feichter, J., Sednev, I., Bauer, S. E., Koch, D., Grainger, R. G., Kirkevaring, A., Iversen, T., Seland, O., Easter, R., Ghan, S. J., Rasch, P. J., Morrison, H., Lamarque, J. F., Iacono, M. J., Kinne, S., and Schulz, M.: Aerosol indirect effects ĝ" general circulation model intercomparison and evaluation with satellite data, Atmospheric Chemistry and Physics, 9, 8697–8717, https://doi.org/10.5194/ACP-9-8697-2009,, 2009.

Rasmussen, C. E. and Williams, C. K. I.: Gaussian Processes for Machine Learning, Gaussian Processes for Machine Learning, https://doi.org/10.7551/MITPRESS/3206.001.0001, 2005.

Rosenfeld, D., Sherwood, S., Wood, R., and Donner, L.: Climate effects of aerosol-cloud interactions, Science, 343, 379–380, https://doi.org/10.1126/SCIENCE.1247490;ISSUE:ISSUE:DOI, 2014.

Schmidt, G. A., Bader, D., Donner, L. J., Elsaesser, G. S., Golaz, J. C., Hannay, C., Molod, A., Neale, R. B., and Saha, S.: Practice and philosophy of climate model tuning across six US modeling centers, Geoscientific Model Development, 10, 3207–3223, https://doi.org/10.5194/GMD-10-3207-2017,, 2017.

Seethala, C. and Horváth, : Global assessment of AMSR-E and MODIS cloud liquid water path retrievals in warm oceanic clouds, Journal of Geophysical Research Atmospheres, 115, 13 202, https://doi.org/10.1029/2009JD012662;REQUESTEDJOURNAL:JOURNAL:21562202D;PAGE:STRING:ARTICLE/CHAPTER, 2010.

Stanford, M. and Mahfouz, N.: Archive for MSP2, Zenodo [Code], https://doi.org/10.5281/zenodo.20739692, 2026.

Stanford, M. W., Fridlind, A. M., Silber, I., Ackerman, A. S., Cesana, G., Mülmenstädt, J., Protat, A., Alexander, S., and Mcdonald, A.: Earth-system-model evaluation of cloud and precipitation occurrence for supercooled and warm clouds over the Southern Ocean's Macquarie Island, Atmospheric Chemistry and Physics, 23, 9037–9069, https://doi.org/10.5194/ACP-23-9037-2023, 2023.

Stephens, G. L. and Haynes, J. M.: Near global observations of the warm rain coalescence process, Geophysical Research Letters, 34, 20 805, https://doi.org/10.1029/2007GL030259;PAGE:STRING:ARTICLE/CHAPTER, 2007.

Tang, J. and Zhuang, Q.: Equifinality in parameterization of process-based biogeochemistry models: A significant uncertainty source to the estimation of regional carbon dynamics, Journal of Geophysical Research: Biogeosciences, 113, 4010, https://doi.org/10.1029/2008JG000757;REQUESTEDJOURNAL:JOURNAL:21562202G;WGROUP:STRING:PUBLICATION, 2008.

Terai, C. R., Xie, S., Song, X., Chen, C. C., Fan, J., Zhang, G. J., Richter, J. H., Shpund, J., Lin, W., Golaz, J. C., Larson, V. E., Moncrieff, M. W., Shan, Y., Zhang, C., Zhang, K., and Zhang, Y.: Impact of Microphysics and Convection Schemes on the Mean-State and Variability

© Author(s) 2026. CC BY 4.0 License.

of Clouds and Precipitation in the E3SM Atmosphere Model, Journal of Advances in Modeling Earth Systems, 17, e2024MS004 569, https://doi.org/10.1029/2024MS004569, 2025.

Twomey, S.: Pollution and the planetary albedo, Atmospheric Environment (1967), 8, 1251–1256, https://doi.org/10.1016/0004-6981(74)90004-3, 1974.

Von Bertalanffy, L.: The Theory of Open Systems in Physics and Biology, Science, 111(2872), 23–29, https://doi.org/10.1126/science.111.2872.23, 1950.

Wang, H., Wu, M., Ke, Z., Feng, Y., Shrivastava, M., Tang, Q., Wan, H., Liu, X., Easter, R. C., Lin, W., Singh, B., Terai, C. R., Burrows, S. M., Fan, J., Ma, P.-L., Mahfouz, N., Mülmenstädt, J., Qian, Y., Shan, Y., Zhang, K., Zhang, Y., Leung, L. R., Xie, S., Rasch, P. J., Shrivastava, M., Easter, R., Terai, C., Burrows, S., and Muelmenstaedt, J.: Overview of Aerosols in the E3SM Version 3: New Model Features and their Impacts, https://doi.org/10.22541/essoar.176804797.79956486/v1, 2026.

Wang, M., Ghan, S., Liu, X., L'Ecuyer, T. S., Zhang, K., Morrison, H., Ovchinnikov, M., Easter, R., Marchand, R., Chand, D., Qian, Y., and Penner, J. E.: Constraining cloud lifetime effects of aerosols using A-Train satellite observations, Geophysical Research Letters, 39, 15 709, https://doi.org/10.1029/2012GL052204;WGROUP:STRING:PUBLICATION, 2012.

Wilks, S. S.: The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses, https://doi.org/10.1214/aoms/1177732360, 9, 60–62, https://doi.org/10.1214/AOMS/1177732360, 1938.

Wood, R.: Drizzle in Stratiform Boundary Layer Clouds. Part II: Microphysical Aspects, Journal of the Atmospheric Sciences, 62, 3034–3050, https://doi.org/10.1175/JAS3530.1, 2005.

Xie, S., Terai, C. R., Wang, H., Tang, Q., Fan, J., Burrows, S. M., Lin, W., Wu, M., Song, X., Zhang, Y., Taylor, M. A., Golaz, J.-C., Benedict, J. J., Chen, C.-C., Feng, Y., Hannah, W. M., Ke, Z., Shan, Y., Larson, V. E., Liu, X., Prather, M. J., Richter, J. H., Shrivastava, M., Wan, H., Zhang, G. J., Zhang, K., Bradley, A., Cameron-Smith, P. J., Damiano, L., Debusschere, B., Donahue, A. S., Easter, R. C., Eldred, M. S., Griffin, B. M., Guba, O., Guo, Z., Huang, X., Lee, J., Lee, H.-H., Lou, S., Mahfouz, N., Moncrieff, M. W., Mülmenstädt, J., Qian, Y., Rasool, Q. Z., Roberts, A., Santos, S. P., Sargsyan, K., Shpund, J., Singh, B., Tao, C., Xie, J., Yang, Y., Zeng, X., Zhang, C., Zhang, M., Zhang, S., Zhang, T., Zheng, X., Jacob, R. L., Leung, L. R., McCoy, R., and Bader, D. C.: The Energy Exascale Earth System Model Version 3. Part I: Overview of the Atmospheric Component, Authorea Preprints, https://doi.org/10.22541/ESSOAR.174456922.21825772/V1, 2025.

Zhang, G. J. and McFarlane, N. A.: Sensitivity of climate simulations to the parameterization of cumulus convection in the canadian climate centre general circulation model, Atmosphere - Ocean, 33, 407–446, https://doi.org/10.1080/07055900.1995.9649539, 1995.