https://doi.org/10.5194/egusphere-2026-3075
Preprint. Discussion started: 11 June 2026

Systematic Observation-Based Estimate of Effective Radiative Forcing from Aerosol–Cloud Interactions

Omer Roi-Cohen[1], Gaea Hadary[1], Casey J. Wall[2,3], Paulo Ceppi[4], and Guy Dagan[1]

[1]Fredy and Nadine Herrmann Institute of Earth Sciences, Hebrew University, Jerusalem, Israel
[2]Department of Meteorology, Stockholm University, Stockholm, Sweden
[3]Bolin Centre for Climate Research, Stockholm University, Stockholm, Sweden
[4]Department of Physics, Imperial College London, London, United Kingdom

Correspondence: Omer Roi-Cohen (omer.cohen15@mail.huji.ac.il); Guy Dagan (guy.dagan@mail.huji.ac.il)

Abstract. The change in Earth's energy budget caused by anthropogenic aerosols interacting with clouds is the most uncertain contributor to the historical energy budget trend, with important implications for future climate projections. Recent studies estimating the effective radiative forcing from aerosol-cloud interactions (ERF_{aci}) using satellite observations and a Cloud-Controlling Factor (CCF) analysis have produced a large spread of results, ranging from approximately -0.3 to -1.5 W m^{-2}. This spread is comparable to the full IPCC AR6 uncertainty range, reflecting the use of different datasets and methodological choices across studies, often without a systematic basis for selecting among them. Here we develop a unified framework to rigorously evaluate these methodological choices across multiple reanalysis datasets, using both climate model simulations and observed regional aerosol trends as independent validation tests. Applying model based bias-correction to the best configuration yields a best-estimate global ERF_{aci} of -0.84 W m^{-2} (66% confidence interval: -1.21 to -0.47 W m^{-2}) and an implied Equilibrium Climate Sensitivity of 3.33 K (66% confidence interval: 2.65 to 4.22 K), both consistent with IPCC AR6 and WCRP 2019 assessments but different from previous CCF-based estimates.

1 Introduction

Clouds scatter incoming shortwave radiation and reflect it back to space, creating a cooling effect; yet because cloud tops are typically colder than the surface below them, they also trap outgoing longwave radiation, producing a compensating warming effect (Loeb et al., 2018). Among all cloud regimes, low clouds are especially radiatively important due to their wide coverage, high shortwave reflectance and low cloud tops (Wood, 2012; Schneider et al., 2017). Marine boundary layer clouds, such as stratocumulus and trade cumulus, are particularly critical because they form vast decks over dark ocean surfaces, maximizing their radiative cooling effect (Hartmann and Short, 1980; Ramanathan et al., 1989; Bender et al., 2011).

Anthropogenic emissions can impact cloud-radiative properties by introducing aerosols into the atmosphere, which act as cloud condensation nuclei (CCN). A higher concentration of CCN results in a larger number of smaller cloud droplets. These smaller droplets scatter shortwave radiation more efficiently, creating an immediate cooling effect known as the Twomey effect (Twomey, 1974). Beyond this instantaneous response, aerosols also induce rapid macrophysical adjustments, such as suppressing precipitation, altering the liquid water path and modifying cloud fraction (Albrecht, 1989; Rosenfeld et al., 2006; Bretherton

https://doi.org/10.5194/egusphere-2026-3075
Preprint. Discussion started: 11 June 2026

et al., 2007; Quaas et al., 2024). The total effect of anthropogenic aerosol emissions on the global Top Of Atmosphere (TOA) radiation budget via modifying clouds, including both the instantaneous forcing and the subsequent adjustments, is known as the Effective Radiative Forcing from Aerosol-Cloud Interactions (ERF_{aci}; Charlson et al. 1992; Stevens 2015; Bellouin et al. 2020).

ERF_{aci} is one of the largest and most uncertain terms in the historical energy budget trend, and constraining it from observations remains an open challenge (Boucher et al., 2013; Bellouin et al., 2020). The uncertainty in our assessment of ERF_{aci} propagates directly into global projections of our changing climate (Bellouin et al., 2020; Watson-Parris and Smith, 2022). Hence, constraining the ERF_{aci} to a narrower range could greatly reduce the uncertainty of the Equilibrium Climate Sensitivity (ECS; Andreae et al. 2005; Sherwood et al. 2020; Watson-Parris and Smith 2022), which is defined as the equilibrium change in the global-mean temperature resulting from a doubling of atmospheric CO_2 concentrations.

Since shallow clouds are radiatively important, their ERF_{aci} is particularly crucial to constrain (Bellouin et al., 2020). Additionally, because shallow marine clouds frequently form in pristine, aerosol-limited environments, they are the clouds most susceptible to anthropogenic emissions (Koren et al., 2014; Bellouin et al., 2020; Wood, 2021). Understanding how these clouds respond to aerosols is therefore key for constraining historical radiative forcing and future warming.

In an effort to improve our assessment of ERF_{aci}, recent studies used observations and reanalysis datasets to calculate ERF_{aci} from shallow clouds via a Cloud Controlling Factor (CCF) analysis (Wall et al., 2022; Park et al., 2025; Zelinka et al., 2026). The CCF method was used successfully in numerous cloud feedback studies (Klein et al., 2017; Scott et al., 2020; Myers et al., 2021; Ceppi and Nowack, 2021; Ceppi et al., 2024) and was recently adopted for ERF_{aci} estimations (Wall et al., 2022). The CCF method isolates the aerosol signal from meteorological covariation (Gryspeerdt et al., 2019b) by treating environmental variables alongside an aerosol proxy as simultaneous cloud predictors. Specifically, linear regression between these CCFs and Cloud-Radiative Effect (CRE) anomalies, applied at each $5^\circ \times 5^\circ$ grid cell using monthly mean anomalies, yield a unique sensitivity of the CRE to each CCF including the aerosol proxy, with all other predictors held constant.

Calculating the ERF_{aci} with the CCF method requires multiplying the sensitivity of the aerosol proxy obtained from the regression by the change in aerosol proxy from pre-industrial (PI) to present-day (PD) conditions. The PD–PI difference can be taken, for example, from Coupled Model Intercomparison Project Phase 6 (CMIP6) historical simulations (Wall et al., 2022). However, previous studies use different choices of aerosol proxies (Aerosol Index, Sulfate Aerosol Optical Depth, and Sulfate Mass Concentration) and different reanalysis datasets for both the meteorological CCFs and the aerosol proxy (Wall et al., 2022; Park et al., 2025; Zelinka et al., 2026; Ceppi et al., 2026). Additionally, one recent study takes into account the activation rate of CCNs to droplets (Park et al., 2025). These methodological differences produce ERF_{aci} estimates ranging from approximately -0.3 to -1.5 W m^{-2} (Fig. 1), a spread comparable to the full IPCC AR6 (Forster et al., 2021) uncertainty range, driven by analytical choices rather than physical uncertainty.

In this study, we develop a unified framework to systematically evaluate and compare these methodological choices. Through rigorous analysis of existing methods and by incorporating regression techniques from cloud feedback studies (Ceppi and Nowack, 2021) that have not been previously applied to ERF_{aci} estimation, we derive observationally-constrained estimates of both ERF_{aci} and ECS. Our results are in close agreement with the IPCC AR6 and WCRP 2019 consensus assessments,

https://doi.org/10.5194/egusphere-2026-3075
Preprint. Discussion started: 11 June 2026

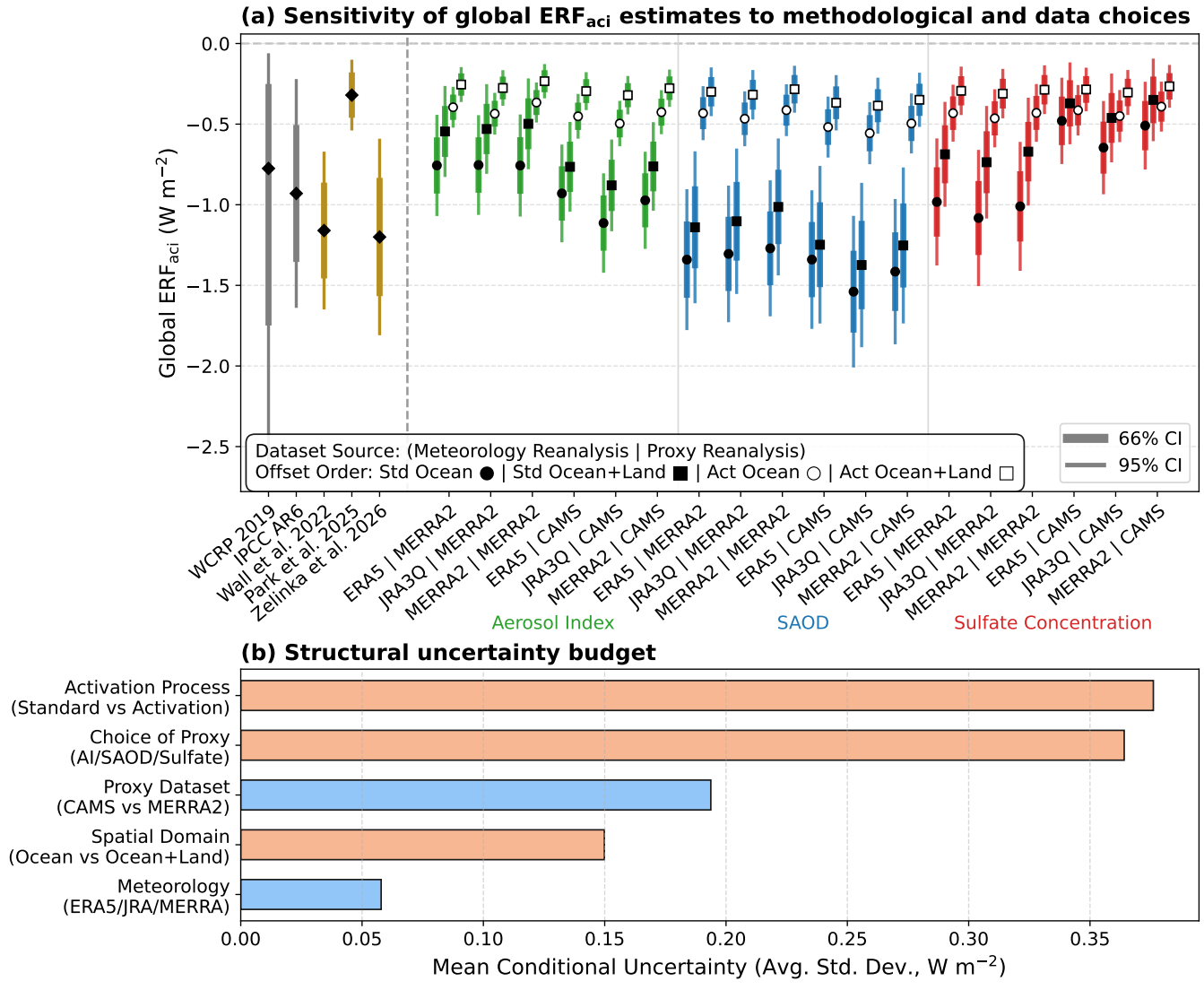

Figure 1. Sensitivity of global observation-based estimates of Effective Radiative Forcing from Aerosol-Cloud Interactions (ERF$_{aci}$) to methodological and data choices. **(a)** Comparison of global mean ERF$_{aci}$ estimates (markers) with their 66% (thick lines) and 95% (thin lines) confidence intervals. The left section displays reference estimates from major assessments (WCRP 2019 (Bellouin et al., 2020) and IPCC AR6 (Forster et al., 2021)) in gray, and published CCF-based estimates (Wall et al., 2022; Park et al., 2025; Zelinka et al., 2026) in yellow. The right section displays estimates calculated in this study using Ordinary Least Squares (OLS) regression with local predictors on a $5^\circ \times 5^\circ$ grid, following the methodology of those previous studies for both the central estimate and uncertainty quantification (Sect. S1 in the Supplement). Results are shown across all dataset combinations, categorized by **aerosol proxy** (green for Aerosol Index, blue for Sulfate Aerosol Optical Depth, red for Sulfate Concentration), **spatial domain** (circles for ocean-only, squares for ocean and land), and **physical pathway** (filled markers for the standard direct proxy route, open markers for the intermediate droplet activation route; see Sect. 3 for details). **(b)** Structural uncertainty budget, quantifying the mean conditional standard deviation (W m^{-2}) introduced to the global ERF$_{aci}$ estimate by changing a single methodological factor while keeping all other choices fixed. The orange bars indicate uncertainty in methodological choice while the blue bars indicate uncertainty in dataset selection.

https://doi.org/10.5194/egusphere-2026-3075
Preprint. Discussion started: 11 June 2026

strengthening confidence in those values and resolving the large discrepancies found among recent CCF-based studies. For the first time, every major methodological choice is evaluated under identical conditions — enabling the uncertainty contributed by each individual choice to the total ERF_{aci} spread to be quantified and making the source of discrepancies among prior studies transparent.

2 Data

Our analysis draws on three categories of data: 1) satellite observations of cloud-radiative properties that serve as the regression target, 2) reanalysis products that provide the CCF predictors (including both meteorological variables and aerosol proxies), and, 3) CMIP6 model output used to estimate the pre-industrial to present-day change in aerosol loading, to evaluate method performance against a known "ground-truth" ERF_{aci}, and to derive a scaling factor that converts our regional estimate to a global forcing (Sect. 3.2).

2.1 Observations and Reanalyses

We use monthly anomalies of satellite observations and reanalysis datasets spanning from 2003 to 2022. To ensure spatial consistency, all variables are first mapped to a common $5° \times 5°$ grid using area-weighted spatial averaging. Once on the common grid, we calculate the anomalies used in our regression by removing the climatological seasonal cycle and the linear trend from all variables.

For the target variable in our regression, we calculate the CRE anomalies of low clouds using the Clouds and the Earth's Radiant Energy System (CERES) FluxByCldTyp Edition 4.1 dataset (Loeb et al., 2018). We include all grid cells between $55°$S to $55°$N in order to minimize retrieval biases due to high solar zenith angle in the polar regions. Low clouds are defined as having a cloud top pressure (CTP) ≥ 680 hPa. As passive satellite instruments view only the highest cloud in a scene, overlying mid- and high-level clouds can artificially obscure the low clouds below. To account for this, we define the non-obscured low-cloud fraction as $L_n = L/(1-U)$, where L is the retrieved low-cloud fraction and U is the retrieved fraction of non-low clouds (CTP < 680 hPa).

Following Wall et al. (2022) and Scott et al. (2020), we compute the unobscured low-cloud radiative effect anomaly to isolate the specific contribution of low clouds to the TOA radiative flux. The anomaly is decomposed into two physically distinct components: changes in total low-cloud amount, which modulate how much of the sky is covered by reflecting clouds; and changes in cloud optical depth and cloud-top pressure, which modulate how strongly a given cloud cover reflects. Together, these two terms account for the full radiative response. Formally, the total low-cloud radiative anomaly R' is:

$$R' = \sum_{p=1}^{P}\sum_{t=1}^{T}(\overline{R_{pt}} - \overline{R_{\mathrm{clr}}})\left[\frac{\overline{f_{pt}}}{\overline{L}}L'_n(1-\overline{U}) + \left(f'_{pt} - \overline{f_{pt}}\frac{L'}{\overline{L}}\right)\right], \tag{1}$$

where p and t index the CTP and optical depth bins respectively, f_{pt} is the retrieved cloud fraction in bin (p,t), R_{pt} is the TOA flux above clouds in that bin, R_{clr} is the clear-sky TOA flux, overbars denote the climatological seasonal cycle, and primes denote monthly anomalies. The first bracketed term captures the radiative contribution from changes in low-cloud amount; the

https://doi.org/10.5194/egusphere-2026-3075
Preprint. Discussion started: 11 June 2026

second captures the contribution from changes in the distribution of optical depth and CTP. We evaluate Eq. (1) separately for shortwave and longwave fluxes to yield the net low-cloud radiative effect anomaly R'_{net}. Finally, prior to performing the regression analysis, we remove the linear trend from the resulting anomaly data.

For the meteorological cloud-controlling factors (CCFs), we separate the grid cells into land and ocean based on their dominant surface type. Over the ocean, following previous studies (Wall et al., 2022; Ceppi et al., 2024; Park et al., 2025; Zelinka et al., 2026), the predictors include Sea Surface Temperature (SST), Estimated Inversion Strength (EIS), relative humidity at 700 hPa (RH_{700}), subsidence at 700 hPa (ω_{700}), surface temperature advection ($-\mathbf{v}\cdot\nabla T_s$, where $\mathbf{v}$ is the near-surface wind vector and T_s is the surface temperature), and surface wind speed. Over land, the predictors include Surface Skin Temperature (T_{s}), Lower Tropospheric Stability (LTS), RH_{700}, and ω_{700} (Ceppi et al., 2024). EIS and LTS are defined as:

$$\mathrm{EIS} = \mathrm{LTS} - \gamma_m\,(Z_{700} - Z_{\mathrm{LCL}})\,; \quad \mathrm{LTS} = \theta_{700} - \theta_{\mathrm{sfc}}, \tag{2}$$

where γ_m is the moist adiabatic lapse rate at 850 hPa, θ is the potential temperature (Klein and Hartmann, 1993), and Z_{LCL} is the lifting condensation level height (Wood and Bretherton, 2006). High-elevation land grid cells are excluded as described below (Fig. S1 in the Supplement). We obtain three sets of these meteorological CCFs using three reanalysis datasets: the fifth generation of the European Center for Medium-Range Weather Forecasts (ECMWF) atmospheric reanalysis of the global climate (ERA5; Hersbach et al. 2020), the Modern-Era Retrospective Analysis for Research and Applications version 2 reanalysis (MERRA-2; Gelaro et al. 2017; Randles et al. 2017), and the Japanese Reanalysis for Three Quarters of a Century (JRA-3Q; Kosaka et al. 2024).

To directly compare our results with prior CCF-based $\mathrm{ERF}_{\mathrm{aci}}$ studies, we consider three aerosol proxies used in the existing literature: Wall et al. (2022) use sulfate mass concentration (SO_4), Park et al. (2025) use SO_4 and Aerosol Index (AI), and Zelinka et al. (2026) use Sulfate Aerosol Optical Depth (SAOD) at 550 nm. We therefore include all three aerosol proxies, each obtained from two reanalysis datasets: MERRA-2 and the Copernicus Atmosphere Monitoring Service (CAMS; Inness et al. 2019). In MERRA-2, AI is calculated using the total aerosol Angstrom parameter (470–870 nm) and the total aerosol extinction at 550 nm. In CAMS, the Angstrom parameter is calculated from the 469 nm and 865 nm wavelengths and multiplied by the AOD at 550 nm. Over ocean grid cells, we extract the sulfate concentration at 910 hPa (Wall et al., 2022). Over land, in the presence of topography, to maintain consistent pressure levels across MERRA-2, CAMS, and the climate models (Sect. 2.2), we interpolate the CAMS sulfate concentrations to the MERRA-2 pressure levels (between 910 hPa and 700 hPa). We then extract the data from the highest pressure level that contains valid data across all models to ensure spatial and temporal consistency. Land grid cells with a surface pressure below 700 hPa (i.e., high mountain ridges) are excluded (Fig. S1 in the Supplement). All aerosol proxies are converted to $\log_{10}$ space for the regression, as in previous studies (Wall et al., 2022; Park et al., 2025; Zelinka et al., 2026).

Finally, for the intermediate activation route analysis (Park et al., 2025), we use cloud droplet number concentration (N_d) retrieved from the Moderate Resolution Imaging Spectroradiometer (MODIS; Gryspeerdt et al. 2019a), following the methodology of Park et al. (2025). Because the N_{d} dataset concludes in 2020, any analysis using N_{d} is restricted to the 2003–2019 period, with anomalies calculated relative to the climatology and trends of that specific timeframe.

https://doi.org/10.5194/egusphere-2026-3075
Preprint. Discussion started: 11 June 2026

2.2 CMIP6 Models

We use gridded data from the Coupled Model Intercomparison Project Phase 6 (CMIP6) historical simulations (Eyring et al., 2016). To calculate the ERF_{aci}, we determine the change in the $\log_{10}$ aerosol proxy between the present day (PD, 2005-2014 average) and the pre-industrial era (PI, 1850-1859 average). We calculate this using all models where the respective proxies are available (11 models for AI, 24 for SAOD, and 21 for SO_4; see Table S1 in the Supplement, for a complete list of models and their specific applications). For models with multiple ensemble members, we average the parameters across all realizations to represent a single model estimate. For AI in CMIP6, the Angstrom parameter is calculated from the 440 nm and 870 nm wavelengths.

In addition, we use the Radiative Forcing Model Intercomparison Project (RFMIP; Pincus et al. 2016) aerosol-only simulations from 20 models (Table S1 in the Supplement) (Smith et al., 2020) as our "ground-truth" ERF_{aci} for two purposes: first, to derive a scaling factor that links the ERF_{aci} of our specific spatial domain ($ERF_{aci,d}$) to the global ERF_{aci} ($ERF_{aci,g}$); and second, to evaluate the predictive skill of our CCF methodology by applying our framework to the CMIP6 historical runs and comparing our predicted ERF_{aci} against the known RFMIP responses.

For each model, we compute the domain-to-global scaling factor γ (defined and applied in Sect. 3.2) by dividing the model's known global $ERF_{aci,g}$ by the $ERF_{aci,d}$ of the $55°$S–$55°$N belt, evaluated over either ocean-only or combined ocean-and-land grid cells (a methodological choice).

For the CMIP6 evaluation of our methods, we extract the CCFs and low-cloud CRE of the latest 20 years of the historical simulations (1995–2014) to align closely with the observational period. To ensure spatial consistency with our observational framework, all CMIP6 outputs are first regridded to the common $5° \times 5°$ grid using area-weighted spatial averaging. We use the International Satellite Cloud Climatology Project (ISCCP; Young et al. 2018) simulator output (Bodas-Salcedo et al., 2011) to obtain cloud fraction across optical depth (τ) and CTP bins, ensuring a like-for-like comparison with the satellite observations. Cloud droplet number concentration, N_d, is extracted following Park et al. (2025) as the column maximum within each vertical atmospheric profile.

Following Ceppi et al. (2024), we compute low-cloud radiative anomalies in the models by applying the Zelinka et al. (2012) cloud-radiative kernels (K_{pt}) on a monthly basis. Because the shortwave kernel depends on the underlying surface albedo, we calculate a monthly climatological clear-sky albedo for each grid cell from the ratio of upwelling to downwelling clear-sky shortwave radiation.

The model low-cloud radiative anomaly R' is then computed using the same two-term decomposition as in Eq. (1), replacing the observed cloud-radiative kernel $(\overline{R_{pt}} - \overline{R_{\mathrm{clr}}})$ with the model-derived kernel K_{pt}:

$$R' = \sum_{p=1}^{P}\sum_{t=1}^{T} K_{pt}\left[\frac{\overline{f_{pt}}}{\overline{L}}L'_n(1-\overline{U}) + \left(f'_{pt} - \overline{f_{pt}}\frac{L'}{\overline{L}}\right)\right] \tag{3}$$

As in the observational framework, the first term captures changes in low-cloud amount and the second captures changes in the distribution of optical depth and CTP. We evaluate this separately for shortwave and longwave fluxes, detrend the resulting anomalies, and apply the same latitude cutoffs and grid-cell exclusion criteria as in the observational analysis.

https://doi.org/10.5194/egusphere-2026-3075
Preprint. Discussion started: 11 June 2026

Ten models provide the complete set of ISCCP simulator outputs, surface fluxes, and ERF_{aci} ground-truth data required for this full evaluation pipeline; of these, 8 contain the SO_4 aerosol proxy, 7 contain the SAOD aerosol proxy and 6 contain the AI aerosol proxy (Table S1 in the Supplement).

3 Methods

3.1 Cloud-Controlling Factor Analysis

Each grid cell contains 5 (land cells) or 7 (ocean cells) CCFs, or predictors, along with the target low-cloud CRE anomalies (the predictand; Sect. 2.1). We compare three regression techniques spanning a range from the strictly local approach used in all prior CCF-based ERF_{aci} studies to methods that explicitly incorporate non-local influences on low clouds. The motivation for testing non-local approaches is twofold: non-local ridge regression has been shown to improve prediction skill in cloud feedback studies (Ceppi et al., 2024); and large-scale circulation adjustments driven by aerosol perturbations, which local regression cannot capture, have recently been shown to substantially amplify ERF_{aci} (Dagan et al., 2023; Yamaguchi et al., 2026). Testing all three methods within a common framework allows us to assess which performs best for estimating ERF_{aci}.

The three regression techniques we use are:

1. **Ordinary Least Squares (OLS) Linear Regression:** Applied locally at each grid cell (Wall et al., 2022; Park et al., 2025; Zelinka et al., 2026) with the cost function of:

$$\beta' = \arg\min_{\beta} \left(||y - X\beta||_2^2 \right), \tag{4}$$

where β are the coefficients, X is the CCF predictors and y is the target low cloud radiative anomalies. β' is the coefficients learned by the regression.

2. **Ridge Regression:** Introduces an L_2 regularization penalty to prevent overfitting and handle multicollinearity among the CCFs:

$$\beta' = \arg\min_{\beta} \left(||y - X\beta||_2^2 + \alpha \sum_k \beta_k^2 \right), \tag{5}$$

where α is the ridge regularization parameter, β_k are the coefficients of cell k in the regressed domain. X, y and β' are the same as in the OLS case (Eq. 4). This allows us to use a domain of non-local CCFs from neighboring grid cells (around the local regressed grid cell) as predictors, motivated by previous cloud feedback studies (Ceppi and Nowack, 2021; Ceppi et al., 2024). We test domains from 1×1 to 9×9 grid cells around the target cell.

3. **Distance-Weighted Ridge Regression:** To account for remote, large-scale circulation adjustments to aerosols (Dagan et al., 2023; Yamaguchi et al., 2026), we expand the regression domain globally and apply a spatially decaying penalty. The coefficients β' are determined by minimizing:

$$\beta' = \arg\min_{\beta} \left(||y - X\beta||_2^2 + \alpha \sum_k (d_k + 1)^{\lambda} \beta_k^2 \right), \tag{6}$$

https://doi.org/10.5194/egusphere-2026-3075
Preprint. Discussion started: 11 June 2026

where α is the ridge regularization parameter, d_k is the geographic distance of the k-th non-local grid cell predictor from the target cell (measured in cell widths), and λ controls the spatial decay regularization. X, y and β' are the same as the OLS case (Eq. 4).

In all three regression techniques, the target variable is the low-cloud CRE anomaly at the local grid cell. In the latter two approaches, however, the CCF predictors are drawn from both local and non-local grid cells and may include contributions from latitudes spanning 55°S to 55°N. We can express the low-cloud CRE anomalies (R'_{net}) as a linear combination of the 5 or 7 CCFs (x_i) of each grid cell in the predictors domain:

$$R'_{\mathrm{net}} \approx \sum_{j=1}^{n} \sum_{i=1}^{5 \text{ or } 7} \frac{\partial R_{\mathrm{net}}}{\partial x_{ij}} x'_{ij}, \tag{7}$$

where primes indicate an anomaly of a variable, j indicates the cell index of the predictor grid cell, and n is the number of predictor grid cells. The regression is repeated individually for all grid cells in the chosen spatial domain of analysis (Sect. 3.2).

We determine the hyperparameters for the ridge and distance-weighted ridge regressions using 5-fold cross-validation. Out-of-sample predictive skill is evaluated using a leave-one-year-out approach, training on 19 years and testing on the remaining year. Finally, we apply 5-fold cross-validation over the entire 20-year period to obtain the final coefficients (Ceppi et al., 2024). For a local cell to be analyzed, it must have at least 4 years of non-missing data across all variables. In the non-local ridge models, if a remote cell is missing more than 10% of the time series of the local target cell, it is excluded; if it is missing less than 10%, the missing temporal gaps are filled with zeros prior to standardization. We iterate this methodology across all 6 combinations of reanalyses (3 meteorological × 2 aerosol datasets) for each of the 3 aerosol proxies.

To address structural uncertainties regarding the physical pathway of aerosol activation, we perform a mediation analysis through the cloud droplet number concentration (N_{d}), following Park et al. (2025). Instead of regressing CRE directly on the aerosol proxy, we split the process into two steps. First, we regress the $\log_{10}(N_{\mathrm{d}})$ anomalies onto the CCFs (including the aerosol proxy) to determine the sensitivity of N_d to aerosols. Second, we regress the CRE anomalies onto the meteorological CCFs and the $\log_{10}(N_{\mathrm{d}})$ anomalies to obtain the sensitivity of CRE to N_{d}. The total aerosol sensitivity is calculated as the product of these two intermediate sensitivities. When using a regression technique that includes non-local cells, we first use OLS between the local CCFs and the N_{d} to obtain the local activation rate of cloud droplets to the aerosol proxy, and then use the non-local regression technique with N_{d} as a predictor instead of the aerosol proxy.

In total, this methodology is applied across 11 combinations of regression technique and domain size, 3 aerosol proxies, 6 reanalysis combinations, and 2 physical pathways, yielding over 100 distinct configurations evaluated at each of the ∼1,500 grid cells in the analysis domain. The full pipeline (including cross-validation, $\mathrm{ERF_{aci}}$ calculation, bias correction, and uncertainty propagation) is applied identically to each configuration, ensuring that all comparisons are made on equal footing.

3.2 Calculating $\mathrm{ERF_{aci}}$ and Model Evaluation

To calculate the total $\mathrm{ERF_{aci,d}}$ of the domain of analysis, we multiply the sensitivity of the aerosol proxy derived from the local regression by the difference in aerosol proxy between PD and PI (ΔProxy) taken from the CMIP6 historical runs. This is cal-

https://doi.org/10.5194/egusphere-2026-3075
Preprint. Discussion started: 11 June 2026

culated at each grid cell and subsequently averaged across the domain of analysis. In our two non-local regression techniques, we also multiply the sensitivity of the local low cloud CRE to each of the non-local cells, with the aerosol proxy change in each of these non-local cells:

$$\mathrm{ERF_{aci,d}} \approx \left\langle \sum_{j=1}^{n} \frac{\partial R_{\mathrm{low}}}{\partial \mathrm{Proxy_j}} \Delta \mathrm{Proxy_j} \right\rangle, \tag{8}$$

where j represents the cell index of the predictor grid cell, and n is the number of predictor grid cells. For each grid cell we regress, we sum the contribution of the domain's proxy CCF to its local $\mathrm{ERF_{aci}}$ estimation, and then we average over the whole domain of analysis to get $\mathrm{ERF_{aci,d}}$. Because our analysis is restricted to the 55°S–55°N latitude belt (evaluated over either ocean-only or combined ocean and land grid cells), we must scale this regional estimate to obtain the global forcing. Following Wall et al. (2022), we use CMIP6 aerosol-only simulations to estimate γ. For each CMIP6 model m, we compute it as:

$$\gamma^m = \frac{\mathrm{ERF}^m_{\mathrm{aci,g}}}{\mathrm{ERF}^m_{\mathrm{aci,d}}}, \tag{9}$$

where the superscript m denotes model-derived quantities. The resulting distribution of γ^m across the model ensemble is then propagated through our Monte Carlo uncertainty procedure described below. This per-model γ is applied to scale the observationally-derived domain forcing to a global estimate:

$$\mathrm{ERF_{aci,g}} = \gamma \mathrm{ERF_{aci,d}}. \tag{10}$$

We evaluate method performance using two independent frameworks that together span the full chain from regression sensitivity to radiative forcing. This dual-evaluation approach, which combines perfect-model testing with real-world observational validation, goes beyond the single-metric evaluations used in prior CCF-based $\mathrm{ERF_{aci}}$ studies and is a central methodological contribution of this work. First, we apply our full regression pipeline to the CMIP6 historical models. We multiply the model-derived aerosol sensitivity by the model's ΔProxy, average across the domain, and compare our prediction to the known aerosol-only simulation ground truth. We test this across different spatial domains as the domain of analysis (55°S to 55°N), both only over ocean grid cells (Wall et al., 2022; Park et al., 2025; Zelinka et al., 2026), and including land grid cells.

Second, we use "opportunistic experiments" (Christensen et al., 2022) following Wall et al. (2022). We isolate two regions with known, strong anthropogenic aerosol trends between 2003 and 2019: East Asia and the East Coast of North America. We apply our CCF analysis methodology to predict the change in low-cloud shortwave radiative effect. The predicted change — calculated by multiplying our regression sensitivities by the observed change in the CCFs (2012–2019 average minus 2003–2010 average) — is compared directly against the observed shortwave radiation changes over the same periods. Uncertainty in the predicted ΔR_{SW} is estimated using a stationary bootstrap (Politis and Romano, 1994), described in Sect. S2 in the Supplement.

Observational CCF methods can be biased because the regression captures not only the aerosol signal but also residual meteorological covariation that the predictors do not fully account for. We correct for this bias by comparing our uncalibrated $\mathrm{ERF_{aci,d}}$ predictions against the known aerosol-only simulations $\mathrm{ERF_{aci}}$ values across a model ensemble and fitting a linear slope to map our raw estimates onto the "true" forcing (Ceppi et al., 2024).

https://doi.org/10.5194/egusphere-2026-3075
Preprint. Discussion started: 11 June 2026

We quantify our uncertainty using Monte Carlo sampling. For each method and domain combination, we draw 1,000,000 samples. In each iteration, we randomly sample one set of aerosol proxy coefficients (from the 6 reanalysis combinations), one CMIP6 model ΔProxy, one slope and intercept from the bias correction distribution, and one domain-to-global scaling parameter (γ). This approach robustly propagates both the structural uncertainty of dataset choice and the statistical uncertainty of the regression. Finally, the resulting probability distribution of the global $ERF_{aci,g}$ is inputted into the World Climate Research Program (WCRP) climate sensitivity framework (Sherwood et al., 2020), that incorporates multiple lines of evidence, to derive the constrained probability distribution of ECS.

4 Results and Discussion

CCF-based ERF_{aci} estimates are highly sensitive to methodological choices. For example, Fig. 1a replicates the approaches of Wall et al. (2022), Park et al. (2025), and Zelinka et al. (2026) across all combinations of aerosol proxy, reanalysis datasets, spatial domain, and physical pathway, using OLS regression with local predictors as in those studies. Minor discrepancies from the original published values reflect differences in land/ocean grid-cell definitions and the order of regridding and anomaly calculation; here all data are regridded prior to anomaly calculation to unify the procedure across observations and CMIP6 models. The resulting global estimates span from approximately -0.3 to over -1.5 W m^{-2}, a spread comparable to the full IPCC AR6 uncertainty range, arising entirely from analytical choices rather than physical uncertainty. Including land cells and incorporating the droplet activation step both substantially reduce the magnitude of the estimated forcing. As shown in Fig. 1b, by holding all other choices fixed while varying one at a time, we decompose the total spread into per-factor contributions, showing that structural uncertainty is not uniformly distributed across methodological decisions but is dominated by a small number of high-impact choices. The activation step introduces the largest single source of structural uncertainty, followed by the choice of aerosol proxy. Both of these methodological choices, alongside the spatial domain and regression technique, are amenable to objective evaluation using our CMIP6 and opportunistic experiment frameworks described below.

We evaluate all combinations of regression technique, activation pathway, aerosol proxy, and spatial domain against the CMIP6 ground-truth $ERF_{aci,d}$, yielding a total of over 100 distinct method configurations tested systematically within a single unified framework. Figure 2a presents the Root Mean Square Error (RMSE) of the predicted $ERF_{aci,d}$ across these configurations evaluated over the combined ocean-and-land domain. The Aerosol Index (AI) consistently outperforms sulfate mass concentration (SO_4) and Sulfate Aerosol Optical Depth (SAOD) across all regression strategies, yielding the lowest RMSE. This advantage holds when RMSE is computed on matched model subsets providing output for both AI and SAOD, or both AI and SO_4, ruling out differences in model subsets as an explanation (Fig. S3 in the Supplement), although the small number of models available for these matched comparisons (as few as 4–5 models per pair; Table S1 in the Supplement) limits the statistical robustness of this conclusion and should be interpreted with appropriate caution.

The standard direct proxy route (solid lines in Fig. 2a) significantly outperforms the intermediate droplet activation route (dashed lines; Park et al. (2025); Sect. 3.1) across all proxies and regression strategies. The activation step produces consistently higher RMSE, with errors varying in direction across models rather than exhibiting a systematic bias, making it a less reliable

https://doi.org/10.5194/egusphere-2026-3075
Preprint. Discussion started: 11 June 2026

pathway for ERF_{aci} estimation, at least based on the evaluation method used here, regardless of the aerosol proxy used. The inconsistency in error direction across models suggests that regression dilution is unlikely to be the primary explanation for the weaker ERF_{aci} estimate of Park et al. (2025), and may instead reflect the compounding of statistical uncertainty across the two regression steps, or the possibility that column-maximum N_d does not fully mediate the aerosol-to-CRE pathway in every model.

Among the regression strategies, three configurations using AI demonstrate the highest predictive skill: the local OLS 1×1 regression, the Ridge 3×3 regression, and the Distance-Weighted Ridge regression (Fig. 2a, highlighted markers). Notably, good out-of-sample skill for cloud-radiative anomalies does not guarantee good ERF_{aci} prediction skill – the two metrics diverge substantially across configurations (Fig. S4 in the Supplement). This likely reflects the fact that out-of-sample skill is dominated by the regression's ability to capture cloud responses to the meteorological CCFs, which collectively outnumber the single aerosol predictor; as a result, differences in cross-validation skill across configurations do not reliably track differences in aerosol sensitivity accuracy. Evaluating methods directly against known forcing responses is therefore essential, a step not taken in prior studies, using the full observational CCF analysis pipeline applied end-to-end within the models.

Including land grid cells substantially reduces uncertainty in the domain-to-global scaling parameter γ. While raw RMSE scores are similar across the two spatial domains (Fig. S5 in the Supplement), the 95% confidence interval on γ narrows from 1.08 ± 0.38 under the ocean-only domain to 0.87 ± 0.06 under the combined ocean-and-land domain, a more than sixfold reduction in half-width. This reflects far greater inter-model agreement on γ when land cells are included, and propagates directly into a consistently lower $ERF_{aci,g}$ uncertainty across all top-performing configurations (Fig. 2b). We therefore evaluate all subsequent results over the combined ocean-and-land domain.

To select a single primary methodology among the three top-performing AI configurations (OLS 1×1, Ridge 3×3, and Distance-Weighted Ridge; all having a similar RMSE), we turn to "opportunistic experiments" – a second, fully independent line of evaluation that tests whether each method can predict observed cloud-radiative trends in regions where aerosol forcing is well-constrained by observations (Fig. 3; Christensen et al. 2022; Wall et al. 2022). Between 2003 and 2019, Eastern North America and Eastern Asia experienced strong, well-documented trends in sulfate aerosol loading driven by regulatory and economic changes in sulfur emissions, producing measurable trends in low-cloud shortwave radiative effects. Since the linear trend is removed from all variables prior to the CCF regression (Sect. 2.1), predicting these trends evaluates the method on information withheld from the regression. The MERRA-2 and CAMS reanalyses capture these trends at slightly different magnitudes (Figs. 3a and 3c), providing an additional test of sensitivity to reanalysis choice.

We apply the regression sensitivities from each of the three top-performing configurations to the observed CCF changes in each region to predict the decadal change in shortwave low-cloud radiative effect (ΔR_{SW}). Across both regions and independent of reanalysis dataset, the OLS 1×1 framework generates predictions with the lowest RMSE and bias relative to the CERES ground truth (Figs. 3b and 3d). We attribute this to the non-local methods over-regularizing the localized aerosol signal, damping the sharp regional gradients that characterize these experiments, a limitation that further supports the selection of OLS 1×1 as our primary methodology. A related factor may be that aerosols vary on shorter timescales than the meteorological CCFs, reducing the benefit of incorporating non-local CCFs from neighboring grid cells for the aerosol coefficient specifically

https://doi.org/10.5194/egusphere-2026-3075
Preprint. Discussion started: 11 June 2026

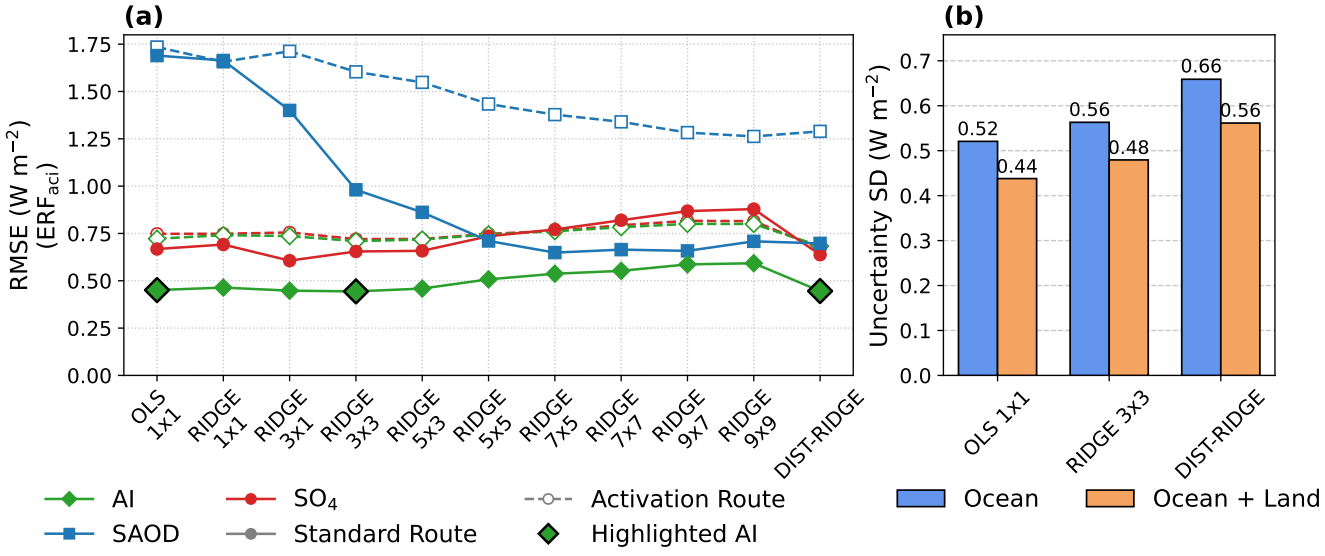

Figure 2. Evaluation of regression methodologies and spatial domains against CMIP6/RFMIP "ground truth". **(a)** Root Mean Square Error (RMSE) of the predicted $ERF_{aci,d}$ across all tested configurations of aerosol proxy (AI, SAOD, SO_4), regression strategy, and physical pathway, evaluated against aerosol-only forcing in CMIP6 models over the combined ocean-and-land domain. Solid lines denote the standard direct proxy route; dashed lines denote the intermediate droplet activation route. Highlighted markers (black outline) indicate the three best-performing configurations, all using AI. The regression CCF grid cell domain size around the target cell is written in each regression technique, except the Distance-Weighted Ridge that takes CCFs from the whole globe. The direct proxy route consistently outperforms the activation route across all proxies and regression strategies. Scatter plots of per-model $ERF_{aci,d}$ prediction skill for OLS, Distance-Weighted Ridge, and the best-performing Ridge configuration, across all proxies and both physical pathways, are provided in Fig. S2 in the Supplement. **(b)** Total structural uncertainty standard deviation (W m^{-2}) of the global $ERF_{aci,g}$ estimate under the ocean-only and ocean-and-land analysis domains, for each of the three top-performing configurations. Including land grid cells consistently reduces the total uncertainty, driven primarily by a sixfold reduction in the confidence interval of the domain-to-global scaling factor γ.

https://doi.org/10.5194/egusphere-2026-3075
Preprint. Discussion started: 11 June 2026

– which could explain why ridge regression improves prediction skill for cloud feedbacks (Ceppi and Nowack, 2021; Ceppi et al., 2024) but not for aerosol forcing estimation.

Both independent evaluation frameworks of CMIP6 perfect-model testing and observational opportunistic experiments point to the same "best-performing" configuration (AI OLS 1×1 over ocean and land), which substantially strengthens confidence in this choice. These frameworks are complementary: the CMIP6 evaluation tests whether the method correctly recovers a known forced response in a controlled setting, while the opportunistic experiments test whether the derived sensitivities correctly predict observed radiative trends driven by real-world aerosol changes. Their agreement is not guaranteed and reflects genuine predictive skill.

Figure 4 details the global spatial structure of the $\mathrm{ERF_{aci}}$ derived using the selected AI OLS 1×1 methodology. The sensitivity of the unobscured low-cloud radiative effect to the aerosol proxy ($\partial R_{\mathrm{low}} / \partial \log_{10} \mathrm{AI}$) is strongly negative over the global oceans, with the largest magnitudes concentrated in subtropical marine stratocumulus regions off the western coasts of continents (Fig. 4a) – precisely the cloud regimes expected to be most susceptible to CCN perturbations (Wood, 2021). Multiplying this sensitivity by the pre-industrial to present-day change in aerosol proxy ($\Delta \log_{10} \mathrm{AI}$; Fig. 4b), yields the predicted local $\mathrm{ERF_{aci}}$ (Fig. 4c). The resulting pattern shows a pronounced cooling effect predominantly over marine environments, consistent with physical expectations of boundary-layer cloud susceptibility. Hatching indicates regions of high uncertainty where reanalysis combinations or CMIP6 models disagree on the sign of the local response, highlighting the remaining uncertainty at the regional scale even after methodological choices are optimized.

While observational CCF methods provide spatially detailed estimates, they can be biased by residual meteorological covariation that the regression predictors do not fully control. We apply an emergent-constraint-style bias correction (Sect. 3.2) by fitting a linear slope between our framework's uncalibrated predictions and the CMIP6 "ground-truth" $\mathrm{ERF_{aci,d}}$ across the CMIP6 evaluation ensemble (Fig. S7 in the Supplement).

After propagating all sources of uncertainty, including regression coefficient uncertainty across six reanalysis combinations, inter-model spread in $\Delta \log_{10} \mathrm{AI}$ across 11 CMIP6 models, calibration uncertainty from the bias correction slope and intercept, and γ scaling uncertainty, through a one-million sample Monte Carlo procedure, our best estimate of the global $\mathrm{ERF_{aci}}$ is -0.84 W m^{-2} (66% confidence interval (CI): -1.21 to -0.47 W m^{-2}; 90% CI: -1.56 to -0.18 W m^{-2}). Inputting this constrained forcing distribution into the WCRP multi-line-of-evidence climate sensitivity framework (Sherwood et al., 2020) yields an Equilibrium Climate Sensitivity (ECS) of 3.33 K (66% CI: 2.65 to 4.22 K; 90% CI: 2.25 to 5.08 K; Fig. 5).

As shown in Fig. 5, our estimates diverge substantially from prior CCF-based studies in instructive ways. Park et al. (2025), who explicitly incorporate the droplet activation step without model-based calibration, obtain a considerably weaker forcing of -0.32 W m^{-2}; our CMIP6 evaluation shows that the activation pathway produces larger RMSE relative to the direct proxy route, with errors varying in direction across models, making it a less reliable basis for $\mathrm{ERF_{aci}}$ estimation regardless of aerosol proxy. Wall et al. (2022) and Zelinka et al. (2026), who use $\mathrm{SO_4}$ and SAOD aerosol proxies, respectively, over the ocean-only domain, obtain a stronger estimate of -1.11 W m^{-2} and -1.20 W m^{-2}, respectively; our evaluation identifies $\mathrm{SO_4}$ and SAOD as less skillful proxies than AI based on the CMIP6 test, and that the ocean-only domain tends to negatively bias $\mathrm{ERF_{aci}}$ estimations. By systematically identifying and correcting each of these sources of structural uncertainty, and applying model-

https://doi.org/10.5194/egusphere-2026-3075
Preprint. Discussion started: 11 June 2026

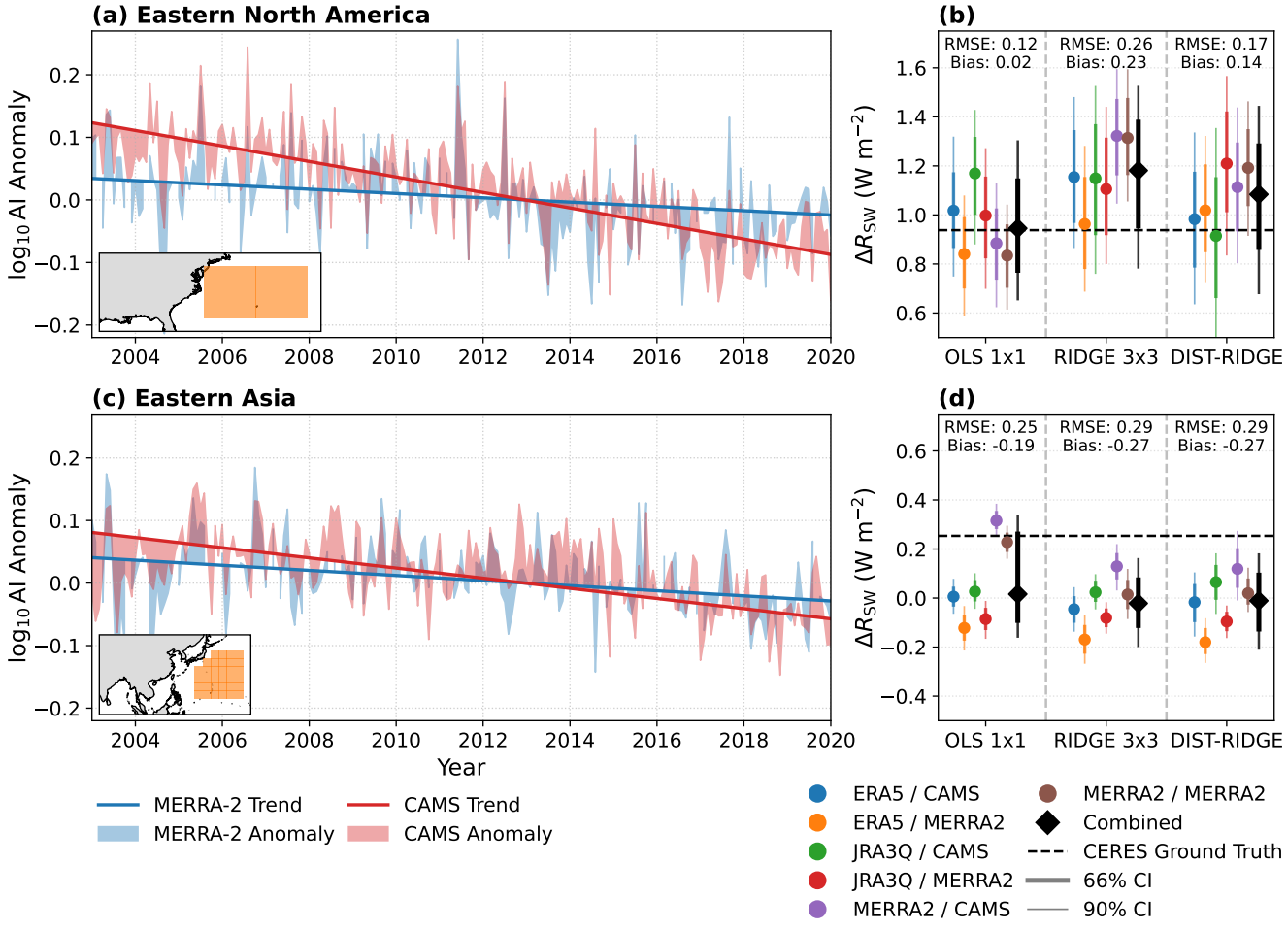

Figure 3. Validation via opportunistic experiments in regions with well-characterized recent aerosol trends. **(a, c)** Time series (2003–2020) of $\log_{10}$(AI) anomalies and linear trends from MERRA-2 and CAMS over Eastern North America and Eastern Asia, respectively. Both regions experienced sustained multi-year trends in aerosol loading driven by documented changes in sulfur emissions, providing a strong observational test of method skill. **(b, d)** Predicted decadal change in shortwave low-cloud radiative effect (ΔR_{SW}, defined as the 2012–2019 average minus the 2003–2010 average) from the three top-performing regression configurations of the CMIP6 evaluation, compared against the CERES observed change (dashed black line). Predictions are shown for both MERRA-2 and CAMS aerosol datasets, with all the meteorological CCFs datasets combinations. Error bars show 66% and 90% confidence intervals derived from a stationary bootstrap (Sect. S2 in the Supplement). The OLS 1×1 methodology yields the lowest RMSE and bias across both regions and reanalysis datasets combinations, supporting its selection as the primary methodology. Results for all configurations are presented in Figure S6 in the Supplement.

https://doi.org/10.5194/egusphere-2026-3075
Preprint. Discussion started: 11 June 2026

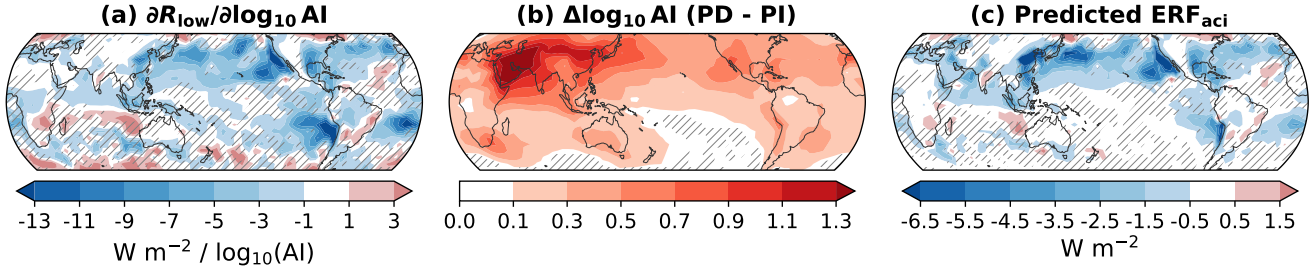

Figure 4. Global distribution of the sensitivity of low-cloud to aerosol and the predicted radiative forcing. All panels are generated using the selected AI OLS 1×1 methodology over the combined ocean-and-land domain, with results shown as the mean across six reanalysis combinations. **(a)** Sensitivity of the unobscured low-cloud radiative effect to the Aerosol Index ($\partial R_{\mathrm{low}}/\partial \log_{10}\mathrm{AI}$; W m^{-2} per unit $\log_{10}$AI). **(b)** Multi-model mean change in the aerosol proxy ($\Delta \log_{10}\mathrm{AI}$) from the pre-industrial to present-day era, derived from CMIP6 historical simulations. **(c)** Predicted local ERF$_{\mathrm{aci}}$ (W m^{-2}), computed as the product of (a) and (b) at each grid cell. The cooling signal is concentrated over marine environments, consistent with the known susceptibility of marine boundary layer clouds to CCN perturbations. Hatching denotes regions of high uncertainty where reanalysis combinations or CMIP6 models disagree on the sign of the local response.

based bias-correction absent from previous studies, our framework converges on a central ERF$_{\mathrm{aci}}$ estimate in close agreement with IPCC AR6. The implied ECS is consistent with the WCRP 2019 consensus, providing independent observational support to both assessments and resolving the wide spread among previous CCF-based studies.

5 Conclusions

We have developed a unified framework to systematically evaluate over 100 methodological configurations for estimating the ERF$_{\mathrm{aci}}$ using the CCF analysis approach. This scope allows us to quantify the uncertainty related to individual methodological choices and robustly identify the optimal configuration for isolating the aerosol forcing signal. Through two independent evaluation frameworks, we identify the OLS 1×1 regression with the AI aerosol proxy over a combined ocean-and-land domain as the most skillful configuration. In contrast, incorporating the intermediate droplet activation pathway (Park et al., 2025) consistently produces higher RMSE with errors that vary in direction across models, making it a less reliable route regardless of the aerosol proxy used – at least under the criteria adopted here. Applying our optimized, bias-corrected framework yields a best-estimate global ERF$_{\mathrm{aci}}$ of -0.84 W m^{-2} (66% CI: -1.21 to -0.47 W m^{-2}) and an implied ECS of 3.33 K (66% CI: 2.65 to 4.22 K). Both estimates are in close agreement with the IPCC AR6 and WCRP 2019 assessments estimates, but differ from previous CCF-based estimates.

The convergence of two independent evaluation frameworks, the CMIP6 perfect-model testing and the "opportunistic experiments" capturing observed aerosol trends in East Asia and Eastern North America, substantially strengthens confidence in this methodology. Because such agreement across two independent evaluation paradigms is not guaranteed, it reflects genuine predictive skill rather than statistical tuning. Furthermore, including land grid cells reduces the 95% CI half-width of γ by more

https://doi.org/10.5194/egusphere-2026-3075
Preprint. Discussion started: 11 June 2026

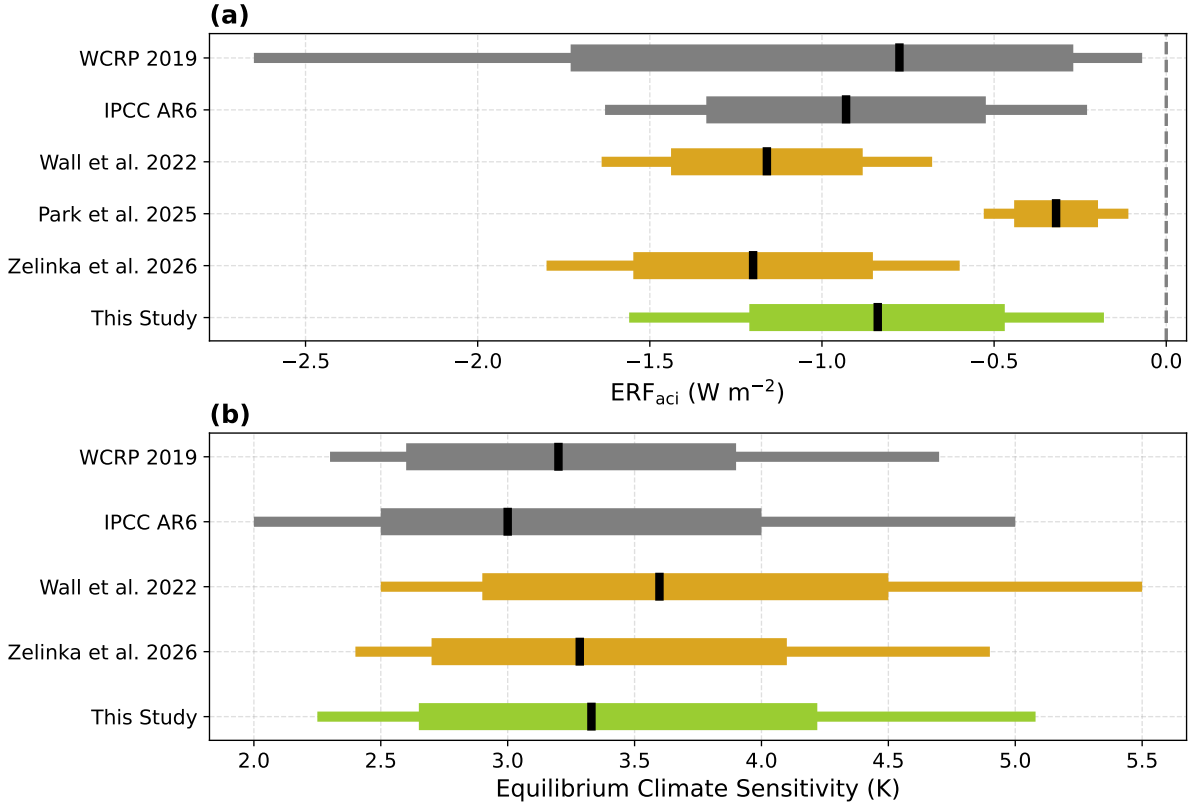

Figure 5. Globally constrained **(a)** ERF_{aci}, and **(b)** Equilibrium Climate Sensitivity (ECS) from this study compared with prior assessments. Markers indicate central estimates; thick lines denote 66% confidence intervals; thin lines denote 90% confidence intervals. Results from this study (derived using the AI OLS 1×1 methodology with CMIP6-based bias correction and full uncertainty propagation) are shown alongside the WCRP 2019 (Bellouin et al., 2020; Sherwood et al., 2020) and IPCC AR6 (Forster et al., 2021) consensus assessments, and three recent CCF-based studies (Wall et al., 2022; Park et al., 2025; Zelinka et al., 2026). Our estimates align closely with the major assessment values and resolve the large spread among prior CCF-based studies, providing independent observational support to the assessed ERF_{aci} and ECS ranges.

https://doi.org/10.5194/egusphere-2026-3075
Preprint. Discussion started: 11 June 2026

than a factor of six, improving both the physical coverage and the final forcing uncertainty (Fig. 2b). Our framework highlights that a bias correction against a ground truth from aerosol-only simulations is essential. Without it, residual meteorological covariation in the regression systematically biases the ERF_{aci} estimates, explaining much of the spread observed among prior studies.

By evaluating all methodological choices within a single controlled pipeline, our results provide important context for recent literature. We demonstrate that the wide spread in prior CCF-based estimates is almost entirely from subjective analytical choices such as the aerosol proxy, spatial domain, and activation pathway, rather than from underlying physical uncertainty. The stronger cooling estimates of Wall et al. (2022) and Zelinka et al. (2026) reflect the use of less skillful aerosol proxies and an ocean-only domain that inflated γ uncertainty. Conversely, the weaker estimate of Park et al. (2025) stems from the intermediate droplet activation pathway, which, when evaluated within our CMIP6 framework, underperforms the direct proxy route in terms of ERF_{aci} prediction skill. A critical methodological finding of our study is that out-of-sample predictive skill for cloud-radiative anomalies does not reliably predict ERF_{aci} estimation accuracy. While prior CCF-based studies relied primarily on regression cross-validation as a quality criterion, we show this to be insufficient; method evaluation must be performed directly against known modeled forcing responses.

Several caveats should be mentioned and point to directions for future work. First, the bias correction and perfect-model evaluation for the AI proxy rely on a limited ensemble of 6 CMIP6 models that provide the necessary output. While this uncertainty is explicitly propagated through our Monte Carlo procedure and is validated by independent opportunistic experiments, expanding the model ensemble in future phases of CMIP would further refine these constraints. Second, CERES-based low-cloud CRE retrievals over land carry greater uncertainty than those over the ocean; future work incorporating multiple satellite datasets for land CRE (following the multi-dataset approach of Ceppi et al. (2024)) could include this uncertainty. Third, there is an inherent asymmetry between the observational and CMIP6 arms of our framework: the CERES-FBCT analysis uses observed cloud-radiative effects directly, while the CMIP6 evaluation employs the Zelinka et al. (2012) kernels applied to ISCCP simulator output, which handle cloud masking differently; replacing the time-varying observational kernel with a monthly climatological mean could bring the two arms into closer alignment. Fourth, while the CCF regression framework effectively captures local aerosol-cloud sensitivities, it may underrepresent large-scale circulation adjustments to aerosol forcing, which can amplify ERF_{aci} (Dagan et al., 2023; Yamaguchi et al., 2026). Although our distance-weighted ridge regression attempted to capture these non-local effects, it did not outperform the local OLS approach in the opportunistic experiments or CMIP6 models evaluation, suggesting that capturing large-scale circulation adjustments in observation-based forcing estimates remains an open problem. Fifth, our chosen aerosol proxies (AI, SAOD, and SO_4) may not fully capture all CCN-active aerosol species that contribute to ERF_{aci}; other aerosol species and specific land-region contributions introduce additional uncertainties not fully quantified here. Finally, our analysis assumes that low clouds dominate the ERF_{aci} and neglects the contribution of non-low clouds; incorporating these higher-level clouds in future analyses could change both the best-performing methodological configuration and the estimated uncertainty range.

The results improve our understanding of the climate system's response to anthropogenic forcing. By independently reproducing ERF_{aci} and ECS values consistent with IPCC AR6 and WCRP 2019 using an observationally-constrained approach with

https://doi.org/10.5194/egusphere-2026-3075
Preprint. Discussion started: 11 June 2026

rigorous model-based calibration, our study lends strong independent support to these consensus assessments and narrows the uncertainty in historical aerosol forcing. This constrained ERF_{aci} range has direct consequences for ECS uncertainty. A moderate, well-constrained aerosol cooling effect leaves less room for a very low climate sensitivity to explain observed historical warming, supporting a likely ECS above ∼2.6 K. Ultimately, the framework developed here, of combining CMIP6 perfect-model evaluation with real-world opportunistic experiments as complementary validation, provides a transferable template for future observational forcing studies, showing that method choices must be evaluated against known forcing responses.

Data availability. All satellite, reanalysis, and climate model data used in this study are publicly available. CERES FluxByCldTyp Edition 4.1 data were downloaded from the NASA CERES ordering tool (https://ceres.larc.nasa.gov/data/). Cloud droplet number concentration (N_d) derived from MODIS was obtained from the NERC EDS Centre for Environmental Data Analysis (Gryspeerdt et al., 2022). MERRA-2 reanalysis data were downloaded from NASA Goddard Earth Sciences Data and Information Services Center, including the following collections: Global Modeling and Assimilation Office (GMAO) (2015a), Global Modeling and Assimilation Office (GMAO) (2015c), Global Modeling and Assimilation Office (GMAO) (2015e), Global Modeling and Assimilation Office (GMAO) (2015f), Global Modeling and Assimilation Office (GMAO) (2015d), and Global Modeling and Assimilation Office (GMAO) (2015b) (Gelaro et al., 2017; Randles et al., 2017). ERA5 reanalysis data is available from the Copernicus Climate Data Store (Hersbach et al., 2023a, b, 2020). JRA-3Q reanalysis data is available from the Data Integration and Analysis System (DIAS) portal (Japan Meteorological Agency (JMA), 2022; Kosaka et al., 2024). CAMS global reanalysis (EAC4) data were obtained from the Copernicus Atmosphere Data Store (Copernicus Atmosphere Monitoring Service (CAMS), 2020; Inness et al., 2019). CMIP6 historical and RFMIP aerosol-only simulation output, including ISCCP satellite simulator diagnostics, can be downloaded from the Earth System Grid Federation (https://esgf-node.llnl.gov/projects/cmip6/; Earth System Grid Federation; Eyring et al. 2016; Pincus et al. 2016; Smith et al. 2020; Young et al. 2018; Bodas-Salcedo et al. 2011). Cloud-radiative kernels were downloaded from Zelinka (2021).

Author contributions. ORC led the data collection, analysis, and manuscript preparation. GH contributed to data collection and preprocessing. CJW and PC provided data, tools, and scientific guidance. GD initiated and supervised the project. All authors reviewed the manuscript.

Competing interests. At least one of the (co-)authors is a member of the editorial board of Atmospheric Chemistry and Physics. The authors declare no other competing interests.

Acknowledgements. We thank Chaim I. Garfinkel for interesting discussion on the presented results. Large language models (LLMs) such as Gemini (Google) and Claude (Anthropic) were used for limited coding tasks and to refine the structure and scientific tone of the manuscript.

https://doi.org/10.5194/egusphere-2026-3075
Preprint. Discussion started: 11 June 2026

Financial support. This research has been supported by the German Research Foundation (DFG) under grant HO 6588/3-1. PC was supported by UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding Guarantee (grant EP/Y036123/1). PC was additionally supported through UK Natural Environmental Research Council (NERC) grants NE/V012045/1 and NE/T006250/1.

https://doi.org/10.5194/egusphere-2026-3075
Preprint. Discussion started: 11 June 2026

References

Albrecht, B. A.: Aerosols, Cloud Microphysics, and Fractional Cloudiness, Science, 245, 1227–1230, https://doi.org/10.1126/science.245.4923.1227, 1989.

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

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/https://doi.org/10.1029/2019RG000660, e2019RG000660 10.1029/2019RG000660, 2020.

Bender, F. A.-M., Charlson, R. J., Ekman, A. M. L., and Leahy, L. V.: Quantification of Monthly Mean Regional-Scale Albedo of Marine Stratiform Clouds in Satellite Observations and GCMs, Journal of Applied Meteorology and Climatology, 50, 2139 – 2148, https://doi.org/10.1175/JAMC-D-11-049.1, 2011.

Bodas-Salcedo, A., Webb, M. J., Bony, S., Chepfer, H., Dufresne, J.-L., Klein, S. A., Zhang, Y., Marchand, R., Haynes, J. M., Pincus, R., and John, V. O.: COSP: Satellite simulation software for model assessment, Bulletin of the American Meteorological Society, 92, 1023 – 1043, https://doi.org/10.1175/2011BAMS2856.1, 2011.

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, pp. 571–657, Cambridge University Press, Cambridge, UK, https://doi.org/10.1017/CBO9781107415324.016, 2013.

Bretherton, C., Blossey, P. N., and Uchida, J.: Cloud droplet sedimentation, entrainment efficiency, and subtropical stratocumulus albedo, Geophysical research letters, 34, 2007.

Ceppi, P. and Nowack, P.: Observational evidence that cloud feedback amplifies global warming, Proceedings of the National Academy of Sciences, 118, e2026290 118, https://doi.org/10.1073/pnas.2026290118, 2021.

Ceppi, P., Myers, T. A., Nowack, P., Wall, C. J., and Zelinka, M. D.: Implications of a Pervasive Climate Model Bias for Low-Cloud Feedback, Geophysical Research Letters, 51, e2024GL110 525, https://doi.org/https://doi.org/10.1029/2024GL110525, e2024GL110525 2024GL110525, 2024.

Ceppi, P., Wilson Kemsley, S., Andersen, H., Andrews, T., Kramer, R. J., Nowack, P., Wall, C. J., and Zelinka, M. D.: Emerging low-cloud feedback and adjustment in global satellite observations, Atmospheric Chemistry and Physics, 26, 4153–4171, 2026.

Charlson, R. J., Schwartz, S. E., Hales, J. M., Cess, R. D., Coakley, J. A., Hansen, J. E., and Hofmann, D. J.: Climate Forcing by Anthropogenic Aerosols, Science, 255, 423–430, https://doi.org/10.1126/science.255.5043.423, 1992.

Christensen, M. W., Gettelman, A., Cermak, J., Dagan, G., Diamond, M., Douglas, A., Feingold, G., Glassmeier, F., Goren, T., Grosvenor, D. P., Gryspeerdt, E., Kahn, R., Li, Z., Ma, P.-L., Malavelle, F., McCoy, I. L., McCoy, D. T., McFarquhar, G., Mülmenstädt, J., Pal, S., Possner, A., Povey, A., Quaas, J., Rosenfeld, D., Schmidt, A., Schrödner, R., Sorooshian, A., Stier, P., Toll, V., Watson-Parris, D., Wood, R., Yang, M., and Yuan, T.: Opportunistic experiments to constrain aerosol effective radiative forcing, Atmospheric Chemistry and Physics, 22, 641–674, https://doi.org/10.5194/acp-22-641-2022, 2022.

https://doi.org/10.5194/egusphere-2026-3075
Preprint. Discussion started: 11 June 2026

Copernicus Atmosphere Monitoring Service (CAMS): CAMS global reanalysis (EAC4), https://doi.org/10.24381/d58bbf47, accessed [03-11-2025], 2020.

Dagan, G., Yeheskel, N., and Williams, A. I.: Radiative forcing from aerosol–cloud interactions enhanced by large-scale circulation adjustments, Nature Geoscience, 16, 1092–1098, 2023.

Earth System Grid Federation: WCRP Coupled Model Intercomparison Project Phase 6, https://esgf-node.llnl.gov/projects/cmip6/.

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

Forster, P., Storelvmo, T., Armour, K., Collins, W., Dufresne, J.-L., Frame, D., Lunt, D. J., Mauritsen, T., Palmer, M. D., 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 Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, edited by Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B., pp. 923–1054, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, https://doi.org/10.1017/9781009157896.009, 2021.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., et al.: The modern-era retrospective analysis for research and applications, version 2 (MERRA-2), Journal of climate, 30, 5419–5454, 2017.

Global Modeling and Assimilation Office (GMAO): MERRA-2 tavgM_2d_aer_Nx: Monthly Mean, Time-Averaged, Single-Level, Assimilation, Aerosol Diagnostics, version 5.12.4, https://doi.org/10.5067/FH9A0MLJPC7N, 2015a.

Global Modeling and Assimilation Office (GMAO): MERRA-2 inst3_3d_aer_Nv: 3d, 3-Hourly, Instantaneous, Model-Level, Assimilation, Aerosol Mixing Ratio, version 5.12.4, https://doi.org/10.5067/LTVB4GPCOTK2, 2015b.

Global Modeling and Assimilation Office (GMAO): MERRA-2 instM_2d_asm_Nx: Monthly Mean, Instantaneous, Single-Level, Assimilation, Single-Level Diagnostics, version 5.12.4, https://doi.org/10.5067/5ESKGQTZG7FO, 2015c.

Global Modeling and Assimilation Office (GMAO): MERRA-2 instM_3d_asm_Np: Monthly Mean, Instantaneous, Pressure-Level, Assimilation, Assimilated Meteorological Fields, version 5.12.4, https://doi.org/10.5067/2E096JV59PK7, 2015d.

Global Modeling and Assimilation Office (GMAO): MERRA-2 tavgM_2d_flx_Nx: Monthly Mean, Time-Averaged, Single-Level, Assimilation, Surface Flux Diagnostics, version 5.12.4, https://doi.org/10.5067/0JRLVL8YV2Y4, 2015e.

Global Modeling and Assimilation Office (GMAO): MERRA-2 tavgM_2d_slv_Nx: Monthly Mean, Time-Averaged, Single-Level, Assimilation, Single-Level Diagnostics, version 5.12.4, https://doi.org/10.5067/AP1B0BA5PD2K, 2015f.

Gryspeerdt, E., Goren, T., Sourdeval, O., Quaas, J., Mülmenstädt, J., Dipu, S., Unglaub, C., Gettelman, A., and Christensen, M.: Constraining the aerosol influence on cloud liquid water path, Atmospheric Chemistry and Physics, 19, 5331–5347, https://doi.org/10.5194/acp-19-5331-2019, 2019a.

Gryspeerdt, E., Goren, T., Sourdeval, O., Quaas, J., Mülmenstädt, J., Dipu, S., Unglaub, C., Gettelman, A., and Christensen, M.: Constraining the aerosol influence on cloud liquid water path, Atmospheric Chemistry and Physics, 19, 5331–5347, 2019b.

Gryspeerdt, E., McCoy, D., Crosbie, E., Moore, R. H., Nott, G. J., Painemal, D., Small-Griswold, J., Sorooshian, A., and Ziemba, L.: Cloud droplet number concentration, calculated from the MODIS (Moderate resolution imaging spectroradiometer) cloud optical properties retrieval and gridded using different sampling strategies, https://doi.org/10.5285/864a46cc65054008857ee5bb772a2a2b, 2022.

https://doi.org/10.5194/egusphere-2026-3075
Preprint. Discussion started: 11 June 2026

Hartmann, D. L. and Short, D. A.: On the Use of Earth Radiation Budget Statistics for Studies of Clouds and Climate, Journal of Atmospheric Sciences, 37, 1233 – 1250, https://doi.org/10.1175/1520-0469(1980)037<1233:OTUOER>2.0.CO;2, 1980.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al.: The ERA5 global reanalysis, Quarterly journal of the royal meteorological society, 146, 1999–2049, 2020.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., et al.: ERA5 monthly averaged data on pressure levels from 1940 to present, https://doi.org/10.24381/cds.6860a573, 2023a.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., et al.: ERA5 monthly averaged data on single levels from 1940 to present, https://doi.org/10.24381/cds.f17050d7, 2023b.

Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J., Engelen, R., Eskes, H., Flemming, J., et al.: The CAMS reanalysis of atmospheric composition, Atmospheric Chemistry and Physics, 19, 3515–3556, https://doi.org/10.5194/acp-19-3515-2019, 2019.

Japan Meteorological Agency (JMA): JRA-3Q: Japanese Reanalysis for Three Quarters of a Century, https://doi.org/10.20783/DIAS.645, 2022.

Klein, S. A. and Hartmann, D. L.: The Seasonal Cycle of Low Stratiform Clouds, Journal of Climate, 6, 1587 – 1606, https://doi.org/10.1175/1520-0442(1993)006<1587:TSCOLS>2.0.CO;2, 1993.

Klein, S. A., Hall, A., Norris, J. R., and Pincus, R.: Low-Cloud Feedbacks from Cloud-Controlling Factors: A Review, Surveys in Geophysics, 38, 1307–1329, https://doi.org/10.1007/s10712-017-9433-3, 2017.

Koren, I., Dagan, G., and Altaratz, O.: From aerosol-limited to invigoration of warm convective clouds, science, 344, 1143–1146, 2014.

Kosaka, Y., Kobayashi, S., Harada, Y., Kobayashi, C., Naoe, H., Yoshimoto, K., Harada, M., Goto, N., Chiba, J., Miyaoka, K., Sekiguchi, R., Deushi, M., Kamahori, H., Nakaegawa, T., Tanaka, T. Y., Tokuhiro, T., Sato, Y., Matsushita, Y., and Onogi, K.: The JRA-3Q Reanalysis, Journal of the Meteorological Society of Japan. Ser. II, 102, 49–109, https://doi.org/10.2151/jmsj.2024-004, 2024.

Loeb, N. G., Doelling, D. R., Wang, H., Su, W., Nguyen, C., Corbett, J. G., Liang, L., Mitrescu, C., Rose, F. G., and Kato, S.: Clouds and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) Top-of-Atmosphere (TOA) Edition-4.0 Data Product, Journal of Climate, 31, 895 – 918, https://doi.org/10.1175/JCLI-D-17-0208.1, 2018.

Myers, T. A., Scott, R. C., Zelinka, M. D., Klein, S. A., Norris, J. R., and Caldwell, P. M.: Observational constraints on low cloud feedback reduce uncertainty of climate sensitivity, Nature Climate Change, 11, 501–507, https://doi.org/10.1038/s41558-021-01039-0, 2021.

Park, C., Soden, B. J., Kramer, R. J., L'Ecuyer, T. S., and He, H.: Observational constraints suggest a smaller effective radiative forcing from aerosol–cloud interactions, Atmospheric Chemistry and Physics, 25, 7299–7313, https://doi.org/10.5194/acp-25-7299-2025, 2025.

Pincus, R., Forster, P. M., and Stevens, B.: The Radiative Forcing Model Intercomparison Project (RFMIP): experimental protocol for CMIP6, Geoscientific Model Development, 9, 3447–3460, https://doi.org/10.5194/gmd-9-3447-2016, 2016.

Politis, D. N. and Romano, J. P.: The Stationary Bootstrap, Journal of the American Statistical Association, 89, 1303–1313, https://doi.org/10.1080/01621459.1994.10476870, 1994.

Quaas, J., Andrews, T., Bellouin, N., Block, K., Boucher, O., Ceppi, P., Dagan, G., Doktorowski, S., Eichholz, H. M., Forster, P., et al.: Adjustments to climate perturbations—mechanisms, implications, observational constraints, AGU Advances, 5, e2023AV001 144, 2024.

Ramanathan, V., Cess, R. D., Harrison, E. F., Minnis, P., Barkstrom, B. R., Ahmad, E., and Hartmann, D.: Cloud-Radiative Forcing and Climate: Results from the Earth Radiation Budget Experiment, Science, 243, 57–63, https://doi.org/10.1126/science.243.4887.57, 1989.

https://doi.org/10.5194/egusphere-2026-3075
Preprint. Discussion started: 11 June 2026

Randles, C., Da Silva, A., Buchard, V., Colarco, P., Darmenov, A., Govindaraju, R., Smirnov, A., Holben, B., Ferrare, R., Hair, J., et al.: The MERRA-2 aerosol reanalysis, 1980 onward. Part I: System description and data assimilation evaluation, Journal of climate, 30, 6823–6850, 2017.

Rosenfeld, D., Kaufman, Y., and Koren, I.: Switching cloud cover and dynamical regimes from open to closed Benard cells in response to the suppression of precipitation by aerosols, Atmospheric Chemistry and Physics, 6, 2503–2511, 2006.

Schneider, T., Teixeira, J., Bretherton, C., Brient, F., Pressel, K., Schär, C., and Siebesma, A.: Climate goals and computing the future of clouds, Nature Climate Change, 7, 3–5, https://doi.org/10.1038/nclimate3190, 2017.

Scott, R. C., Myers, T. A., Norris, J. R., Zelinka, M. D., Klein, S. A., Sun, M., and Doelling, D. R.: Observed Sensitivity of Low-Cloud Radiative Effects to Meteorological Perturbations over the Global Oceans, Journal of Climate, 33, 7717 – 7734, https://doi.org/10.1175/JCLI-D-19-1028.1, 2020.

Sherwood, S. C., Webb, M. J., Annan, J. D., Armour, K. C., Forster, P. M., Hargreaves, J. C., Hegerl, G., Klein, S. A., Marvel, K. D., Rohling, E. J., Watanabe, M., Andrews, T., Braconnot, P., Bretherton, C. S., Foster, G. L., Hausfather, Z., von der Heydt, A. S., Knutti, R., Mauritsen, T., Norris, J. R., Proistosescu, C., Rugenstein, M., Schmidt, G. A., Tokarska, K. B., and Zelinka, M. D.: An Assessment of Earth's Climate Sensitivity Using Multiple Lines of Evidence, Reviews of Geophysics, 58, e2019RG000 678, https://doi.org/https://doi.org/10.1029/2019RG000678, e2019RG000678 2019RG000678, 2020.

Smith, C. J., Kramer, R. J., Myhre, G., Alterskjær, K., Collins, W., Sima, A., Boucher, O., Dufresne, J.-L., Nabat, P., Michou, M., et al.: Effective radiative forcing and adjustments in CMIP6 models, Atmospheric Chemistry and Physics, 20, 9591–9618, https://doi.org/10.5194/acp-20-9591-2020, 2020.

Stevens, B.: Rethinking the Lower Bound on Aerosol Radiative Forcing, Journal of Climate, 28, 4794 – 4819, https://doi.org/10.1175/JCLI-D-14-00656.1, 2015.

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

Wall, C. J., Norris, J. R., Possner, A., McCoy, D. T., McCoy, I. L., and Lutsko, N. J.: Assessing effective radiative forcing from aerosol–cloud interactions over the global ocean, Proceedings of the National Academy of Sciences, 119, e2210481 119, https://doi.org/10.1073/pnas.2210481119, 2022.

Watson-Parris, D. and Smith, C. J.: Large uncertainty in future warming due to aerosol forcing, Nature Climate Change, 12, 1111–1113, 2022.

Wood, R.: Stratocumulus clouds, Monthly weather review, 140, 2373–2423, 2012.

Wood, R.: Assessing the potential efficacy of marine cloud brightening for cooling Earth using a simple heuristic model, Atmospheric Chemistry and Physics, 21, 14 507–14 533, https://doi.org/10.5194/acp-21-14507-2021, 2021.

Wood, R. and Bretherton, C. S.: On the Relationship between Stratiform Low Cloud Cover and Lower-Tropospheric Stability, Journal of Climate, 19, 6425 – 6432, https://doi.org/10.1175/JCLI3988.1, 2006.

Yamaguchi, T., Yoshida, R., Chen, Y.-S., McCoy, I. L., and Feingold, G.: Dynamical and radiative influence on the Hadley circulation by aerosol-cloud interactions, Journal of Advances in Modeling Earth Systems, 18, e2025MS005 465, 2026.

Young, A. H., Knapp, K. R., Inamdar, A., Hankins, W., and Rossow, W. B.: The International Satellite Cloud Climatology Project H-Series climate data record product, Earth System Science Data, 10, 583–593, https://doi.org/10.5194/essd-10-583-2018, 2018.

Zelinka, M. D.: Mzelinka/cloud-radiative-kernels: Sep 17, 2021 release, https://doi.org/10.5281/zenodo.5514137, 2021.

https://doi.org/10.5194/egusphere-2026-3075
Preprint. Discussion started: 11 June 2026

Zelinka, M. D., Klein, S. A., and Hartmann, D. L.: Computing and partitioning cloud feedbacks using cloud property histograms. Part I: Cloud radiative kernels, Journal of Climate, 25, 3715–3735, https://doi.org/10.1175/JCLI-D-11-00248.1, 2012.

Zelinka, M. D., Myers, T. A., Qin, Y., Chao, L.-W., Klein, S. A., Po-Chedley, S., Ma, P.-L., Wall, C. J., Ceppi, P., and Gettelman, A.: Recent cloud trends and extremes reaffirm established bounds on cloud feedback and aerosol-cloud interactions, Communications Earth & Environment, https://doi.org/10.1038/s43247-026-03461-8, 2026.