https://doi.org/10.5194/egusphere-2026-3606
Preprint. Discussion started: 16 September 2026

ECOMIP: The EarthCARE-ORCESTRA Model Intercomparison Project

Woosub Roh[1], Masaki Satoh[1,2], Romain Fiévet[3], Mark D. Fielding[4], Robin J. Hogan[4], Bjorn Stevens[3], Takuji Kubota[5], and Hajime Okamoto[6]

[1]Atmosphere and Ocean Research Institute, The University of Tokyo, Kashiwa, Chiba 277-8564, Japan
[2]Typhoon Science and Technology Research Center, Institute for Multidisciplinary Sciences, Yokohama National University
[3]Max Planck Institute for Meteorology, Hamburg, Germany
[4]European Centre for Medium-Range Weather Forecasts, Reading, UK
[5]Earth Observation Research Center, Japan Aerospace Exploration Agency, Tsukuba, Ibaraki, Japan
[6]Research Institute for Applied Mechanics, Kyushu University, Kasuga, Fukuoka, Japan

Correspondence to: Woosub Roh (ws-roh@aori.u-tokyo.ac.jp)

Abstract.

This study proposes the EarthCARE-ORCESTRA Model Intercomparison Project (ECOMIP), a framework for evaluating various range of atmospheric models, including global or regional km-scale models using observations from the EarthCARE satellite and the ORCESTRA (Organized Convection and EarthCARE Studies over the Tropical Atlantic) field campaign. The goal of ECOMIP is not to rank model performance but to provide a clear path to improve models by considering how new satellite observations, combined with in situ measurements and satellite simulators, can constrain cloud microphysics, precipitation, and convective dynamics. This paper presents the protocol for Phase 1 of the ECOMIP framework, consisting of 2-day simulations designed for comparison with EarthCARE and ORCESTRA observations. As a first examination of the model intercomparison, NICAM, ICON, and IFS simulations are evaluated with EarthCARE observations during the satellite overpass over the ORCESTRA domain on 3 September 2024. Along the EarthCARE track, the occurrence of a mesoscale convective system and Southern Hemisphere cirrus clouds is captured in both simulations. The three models exhibit different convective organization characteristics. NICAM and IFS tend to produce larger and more coherent convective systems, whereas ICON shows more scattered convective cores. Although conclusions from a single EarthCARE frame cannot be generalized, these results provide an example of model-dependent characteristics and demonstrate the potential of the proposed protocol for improving our understanding through the analysis of a larger number of EarthCARE cases and datasets in future studies. We also present the comprehensive framework of ECOMIP after Phase 1, including statistical free-run simulations and observationally constrained parameter-sensitivity experiments.

1 Introduction

Recent advances in satellite observation provide new opportunities for constraints in atmospheric numerical models, including climate and weather forecasting models. The Earth Cloud, Aerosol and Radiation Explorer (EarthCARE; Illingworth et al., 2015; Wehr et al., 2023), a joint mission of the European Space Agency (ESA) and the Japan Aerospace Exploration Agency (JAXA), launched in May 2024 (Kubota et al. 2026),carries synergistic active and passive sensors including a Cloud Profiling Radar (CPR), the ATmospheric LIDar (ATLID), a Multi-Spectral Imager (MSI), and a Broad-Band Radiometer (BBR). The combined radar–lidar profiling

https://doi.org/10.5194/egusphere-2026-3606
Preprint. Discussion started: 16 September 2026

capability enables detailed characterization of vertical cloud and aerosol structure, while the BBR provides top-of-atmosphere radiative fluxes. Importantly, the EarthCARE CPR includes Doppler capability, allowing retrieval of hydrometeor terminal velocities and information on vertical air motion under appropriate assumptions. This combination of vertical structure and radiative measurements offers unprecedented constraints on cloud microphysics and cloud–radiation interactions at the global scale.

Complementary to satellite observations, the ORCESTRA (Organized Convection and EarthCARE Studies over the Tropical Atlantic; Stevens et al. 2026) campaign, conducted from August to September 2024, provides a comprehensive observational framework for studying mesoscale convective organization in the tropical Atlantic, particularly within the Intertropical Convergence Zone (ITCZ). Through coordinated deployment of research aircraft, a research vessel, and ground-based stations, ORCESTRA produced a multi-platform dataset spanning from the upper ocean to the tropopause. The combination of in situ microphysical measurements, airborne Doppler radar observations, aerosol profiling, and surface flux measurements offers process-level constraints that complement EarthCARE's global coverage. Owing to its spatial coverage and strong synergy with EarthCARE overpasses, the ORCESTRA dataset provides a powerful basis for evaluating satellite retrievals and constraining atmospheric models.

With unprecedentedly rich observational data, atmospheric numerical models can be evaluated and improved through detailed analysis of the most uncertain aspects of cloud and convective processes. In principle, various types of atmospheric models can be targeted for evaluation against these observations. In particular, atmospheric models with the horizontal grid spacing refined to about approximately 0.5–5 km, known as convection permitting models, directly represent cloud and convective systems without using cumulus parameterization, and comparisons with these observations are straightforward (Hashino et al., 2013; Matsui et al., 2016; Satoh et al., 2022). Coarser-resolution atmospheric models with cumulus parameterization are also possible if convective parameterization is properly analyzed (Hashino et al. 2025; Nakamura et al. 2026).

In particular, global km-scale models, or global storm-resolving models (GSRMs), represent a new generation of global atmospheric models with horizontal grid spacing typically finer than ~5 km and dynamical cores based on nonhydrostatic governing equations (Satoh et al. 2019; Stevens et al. 2019). At this resolution, a substantial portion of deep convective motions is explicitly resolved rather than parameterized, in contrast to conventional global circulation models (GCMs) operating at O(10–100 km) grid spacing. By reducing reliance on convective parameterization schemes, GSRMs provide a dynamically consistent framework for representing mesoscale convective systems, tropical cyclones, and organized convection within a global domain. As a consequence, GSRMs have been actively developed and applied by multiple research groups worldwide for both weather prediction and climate research (Satoh et al., 2019; Stevens et al., 2019).

Intercomparison efforts have played a central role in evaluating the performance of these km-scale models. Early coordinated experiments, conducted under Northern Hemisphere winter/summer conditions, employed standardized initial conditions, approximately 40-day simulation periods, and harmonized output protocols, thereby enabling systematic comparison of large-scale circulation, precipitation, cloud organization, and radiative fluxes across multiple nonhydrostatic global models (Stevens et al., 2019). These studies demonstrated that independent GSRMs reproduce realistic synoptic-scale structures and mesoscale convective organization comparable to satellite observations.

https://doi.org/10.5194/egusphere-2026-3606
Preprint. Discussion started: 16 September 2026

Subsequent intercomparison experiments extended this framework to winter conditions and to seasonal or year-long integrations, allowing assessment of intraseasonal variability, monsoon dynamics, and quasi-equilibrium climatological states beyond the initial spin-up period (Takasuka et al., 2024). In addition, idealized aqua-planet experiments have been conducted to isolate fundamental cloud and convective processes in the absence of land–sea contrasts and topographic complexity.

However, most previous intercomparisons have focused primarily on model-to-model differences evaluated using bulk diagnostics such as precipitation statistics, outgoing longwave radiation (OLR), outgoing shortwave radiation (OSR), and cloud fraction. Indeed, agreement in horizontally integrated radiative fields does not necessarily imply consistency in vertical cloud structure.

In this context, active satellite observations from CloudSat and CALIPSO have demonstrated the importance of evaluating the vertical distributions of clouds and hydrometeors in atmospheric models. For example, Delanoë et al. (2011) used CloudSat observations to evaluate the representation of ice water content (IWC) in numerical models, while Bodas-Salcedo et al. (2008) demonstrated the usefulness of radar simulators for generating CloudSat-like observations from models and enabling consistent comparisons between models and satellite observations. Forbes and Ahlgrimm (2014) showed that combined CloudSat and CALIPSO observations can help constrain mixed-phase cloud processes in forecast models.

Similarly, Roh et al. (2021) showed that although multiple GSRMs produce broadly similar horizontal OLR distributions, substantial differences exist in the vertical distribution of liquid and ice hydrometeors. Such discrepancies suggest that compensation among microphysical assumptions, particle size distributions, and fall-speed formulations can yield similar radiative signatures despite divergent internal cloud structures. Because vertical hydrometeor profiles strongly influence precipitation efficiency, convective organization, and radiative feedbacks, constraining these process-level differences represents a critical next step in advancing GSRM development.

To enable consistent comparison between models and observations, forward satellite simulators are required. These tools translate model state variables into synthetic observables such as radar reflectivity, lidar backscatter, Doppler velocity, and radiative fluxes. Prior to the launch of EarthCARE, (Roh et al., 2023) generated EarthCARE-like synthetic observations by coupling a GSRM with the Joint-Simulator for Satellite Sensors (J-Sim: Hashino et al., 2013; Roh et al., 2020). Building on this approach, a recent study has demonstrated that combined analysis of EarthCARE radar reflectivity and Doppler velocity provides valuable constraints on hydrometeor terminal velocities and atmospheric vertical motions in a GSRM (Roh et al., 2025). Similarly, Fielding and Janisková (2020) described the cloud simulator that is used to assimilate EarthCARE radar and lidar data into the ECMWF model.

We introduce EarthCARE-ORCESTRA Model Intercomparison Project (ECOMIP), a coordinated framework for model intercomparison and evaluation that integrates observations from the EarthCARE mission and the ORCESTRA tropical Atlantic field campaign. The scientific objectives of this framework are fourfold.

First, we aim to compare cloud microphysics across models and observations to understand the characteristics of model representations and their diversity. This comparison will clarify the models' deficiencies or biases and provide guidance for improving cloud processes.

Second, we evaluate models by applying satellite simulators to directly compare sensor signals of the EarthCARE and the ORCESTRA instruments. This analysis clarifies the consistency of cloud microphysical

https://doi.org/10.5194/egusphere-2026-3606
Preprint. Discussion started: 16 September 2026

process models and their uncertainties. For example, it will help us understand how differences in particle-size distributions and fall-speed parameterizations shape vertical cloud and precipitation structure.

Third, we examine how differences in cloud structure translate into variations in top-of-atmosphere radiative fluxes (OSR and OLR). This will be explored by consistent use of forward radiative transfer models, thereby linking microphysical assumptions to radiative impacts and identifying the dominant sources of radiative sensitivity. Again, we will clarify how cloud microphysical assumptions and their representation affect radiation.

Fourth, we aim to provide guidance for systematic evaluation across various observations to improve models. This will be pursued by analyzing robust features and systematic differences among models using a suite of observationally consistent metrics, including cloud microphysics, radiation fluxes and cloud and precipitation distribution. In addition, analysis of global models, particularly GSRMs, will examine how changes in cloud microphysics affect large-scale circulation and the climate state. This will lead to a re-evaluation of climate sensitivity, with updated versions of global climate models, following the evaluation and update of cloud microphysics processes using EarthCARE and ORCESTRA observation data.

The present paper initiates this intercomparison by leveraging synergistic observations from EarthCARE and ORCESTRA. We first describe an experimental protocol, including simulation configuration, required output variables, temporal and spatial sampling strategies, and satellite-simulator settings. We then present example analyses using three representative GSRMs, the Nonhydrostatic Icosahedral Atmospheric Model (NICAM; Satoh et al., 2008, 2014), the ICOsahedral Nonhydrostatic (ICON; Wan et al., 2013), Integrated Forecasting System (IFS; Malardel et al. 2016), to demonstrate the applicability of the framework. The protocol is designed to be extensible across different temporal scales, from short-term case studies aligned with intensive observation periods.

2 The intercomparison protocol

The concept of this protocol is to intercompare atmospheric numerical models with observational datasets from the EarthCARE satellite and ORCESTRA field campaigns over the tropical Atlantic Ocean, in order to better understand how clouds, precipitation, and storm systems are represented in current models. By comparing model simulations with various observational datasets, we aim to better understand model behavior and identify which cloud systems and microphysical processes are well reproduced and which differ from the observations.

For comparisons with observations, the model setup will be designed to limit spin-up time and maintain consistency with the synoptic scales of the observational periods. This approach is intended to reduce differences arising from models' departures from observations and to allow more direct comparisons between simulated and observed cloud and precipitation fields on a case-by-case basis. The intercomparison will focus on the vertical distribution of hydrometeors and their relationship to radiation, precipitation, and storm dynamics. Variations in simulated hydrometeor profiles among models may influence radiative heating rates, surface precipitation, and the development and organization of convective systems.

To facilitate comparisons with EarthCARE observations, satellite simulators will be used to convert model outputs into observation-equivalent quantities, such as radar reflectivity and Doppler velocity. This enables direct comparisons with satellite measurements and provides a framework for assessing how assumptions about particle size distributions and microphysical processes are reflected in simulated radar signals in an observational context.

https://doi.org/10.5194/egusphere-2026-3606
Preprint. Discussion started: 16 September 2026

In this intercomparison protocol, participating models include both global and regional models, the latter focusing on the ORCESTRA observation domain. This combined framework allows both global-scale consistency and regionally targeted analyses to be addressed within a common protocol.

The key aspect of ECOMIP is the use of diverse, complementary observational datasets. For EarthCARE, data from the CPR, including radar reflectivity and Doppler velocity, will be employed to examine the vertical structure of clouds. These measurements provide information on ice particle size and ice and liquid water contents, while Doppler velocity offers insight into hydrometeor terminal velocities and vertical air motions. In addition, observations from ATLID will be used to identify the thermodynamic phase of clouds and to characterize the vertical distribution of aerosols. ORCESTRA observations further complement satellite data: two research aircraft observed cloud vertical structures using radars at different wavelengths, and dropsonde measurements provide information on large-scale vertical atmospheric motions. Moreover, ship-based vertically pointing sensors and polarimetric radar observations contribute to understanding precipitation characteristics and cloud microphysical processes. Together, these datasets form a comprehensive observational basis for the planned intercomparison.

This study will also make use of satellite simulators developed and applied by different modeling groups in order to intercompare forward-modeling approaches. Forward models such as J-Sim (Hashino et al., 2013; Roh et al., 2020), PAMTRA (Mech et al., 2020) and ZmVar (Fielding and Janisková 2020) employ different scattering databases and computational modules, leading to variations in simulated observation-equivalent quantities. Climate modeling groups are familiar with the CFMIP Observation Simulator Package (COSP, Bodas-Salcedo et al. 2012), to which the EarthCARE CPR simulators have recently been introduced (Nakamura et al. 2026). Intercomparison of these forward models provides an opportunity to examine uncertainties associated with satellite simulations and to assess how differences in physical assumptions are reflected in simulated radar and lidar signals. Such comparisons are expected to support a more consistent interpretation of model–observation differences and to contribute to improved understanding of the underlying physical processes.

3 Experimental designs

3.1 Overviews

We propose a three-phase experimental framework (Table 1) that progressively links process-level diagnostics with statistical model evaluation and parameter sensitivity analysis. Phase 1 consists of targeted case studies aligned with EarthCARE overpasses and ORCESTRA field observations. The primary objective of this phase is to enable direct model–observation comparisons in observation space, thereby improving process-level understanding of convection, vertical cloud structure, hydrometeor microphysics, and aerosol–cloud interactions. By minimizing spin-up effects and synchronizing simulations with observational periods, this phase emphasizes detailed interpretation of vertical profiles, radar and lidar signatures, and associated radiative characteristics. In addition to the core case jointly simulated by all participating groups, we encourage contributors to propose additional scientifically interesting cases for coordinated analysis and intercomparison (Table 2). This flexible design allows the framework to capture a broader range of convective regimes and environmental conditions while maintaining consistency in diagnostics and evaluation metrics.

Phase 2 extends the analysis to two-month free-run simulations under realistic boundary conditions. Simulations are initialized at 00:00 UTC on 1 August, and the analysis period spans from 00:00 UTC on 10 August

https://doi.org/10.5194/egusphere-2026-3606
Preprint. Discussion started: 16 September 2026

to 30 September. This phase allows for statistical evaluation of model performance across a range of meteorological regimes and internal variability. Diagnostics include cloud occurrence frequency, hydrometeor vertical distributions, precipitation characteristics, and top-of-atmosphere radiative fluxes. By comparing models over an extended integration period, Phase 2 provides a robust assessment of systematic differences and commonalities among participating models beyond individual case events.

Phase 3 builds on insights from the previous phases by conducting constrained free-run simulations in which selected cloud microphysics and radiation parameters—such as fall-speed formulations, hydrometeor density assumptions, and particle-size distribution parameters—are adjusted based on observational evidence from EarthCARE and ORCESTRA. The objective is to evaluate how observationally informed parameter modifications influence the simulated radiation budget, precipitation statistics, convective organization, and cloud climatology over the two-month integration period. This phase thus provides a structured pathway for quantifying the sensitivity of large-scale model behavior to process-level parameter constraints derived from vertically resolved observations.

The timeline of the three phases is schematically outlined in Figure 1.

Table 1. Overview of the three-phase experimental framework

Phase	Objectives and Description	Target Cases / Duration
Phase 1 – Case Studies	Direct comparison and interpretation of EarthCARE and ORCESTRA observations to improve understanding of convection, cloud microphysics, and aerosol–cloud interactions.	2-day simulations. Mandatory: 3 Sep 2024 targeted for deep convection Optional: Additional events selected based on participant interests.
Phase 2 – Free-run Simulations	Conduct two-month continuous simulations for statistical evaluation of model performance and intercomparison under natural variability.	Around 2 months
Phase 3 – Constrained Free-run Simulations	Perform free-run experiments with observationally constrained parameters (e.g., terminal velocity, hydrometeor density) derived from EarthCARE/ORCESTRA insights, to examine their impact on radiation, precipitation, convection, and cloud statistics. A common configuration will be applied across models.	Around 2 months

Table 2. Phase 1 mandatory and optional cases: analysis period, initial conditions, and description.

No.	Analysis period (target time)	Initial condition	Comments
Case 1 (mandatory)	00:00 UTC 3–00:00 UTC 4 September 2024 (Target: 16:07 UTC 3 September 2024)	00:00 UTC 2 September 2024	Deep convection, 'golden case' with EarthCARE, HALO, Meteor Frame number: 01520E
Case 2 (optional)	00:00 UTC 18–00:00 UTC 19 August 2024 (Target: 16:04 UTC 18 August 2024)	00:00 UTC 17 August 2024	Deep convection Frame number: 01271E

https://doi.org/10.5194/egusphere-2026-3606
Preprint. Discussion started: 16 September 2026

Case 3 (optional)	00:00 UTC 21–00:00 UTC 22 September 2024 (Target: 17:33 UTC 21 September 2024)	00:00 UTC 20 September 2024	Scattered convection Frame number: 01801E
Case 4 (optional)	00:00 UTC 27–00:00 UTC 28 August 2024 (Target: 16:00 UTC 27 August 2024)	00:00 UTC 26 August 2024	Deep convection Frame number: 01411E
Case 5 (optional)	00:00 UTC 19–00:00 UTC 20 September 2024 (Target: 17:43 UTC 19 September 2024)	00:00 UTC 18 September 2024	Anvil clouds Frame number: 01770E

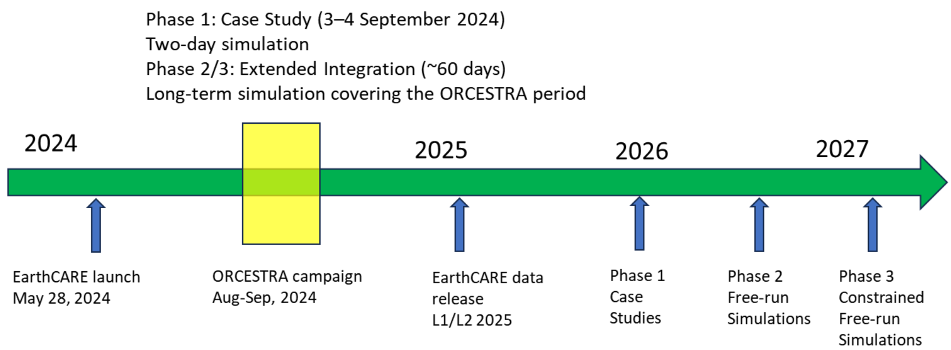

Figure 1: Timeline illustrating the EarthCARE launch, the ORCESTRA field campaign, and the corresponding availability of EarthCARE data used in this study.

https://doi.org/10.5194/egusphere-2026-3606
Preprint. Discussion started: 16 September 2026

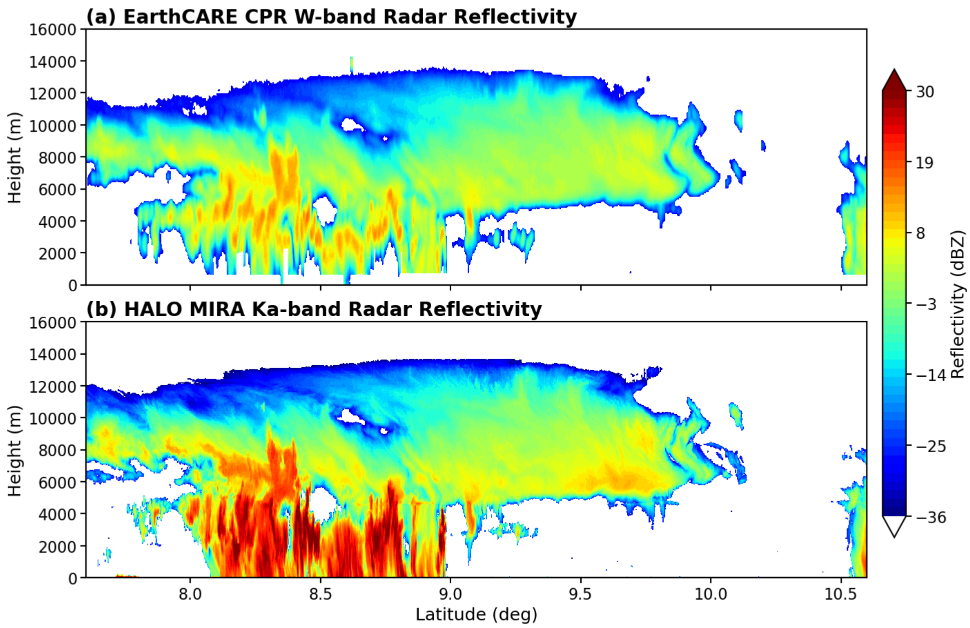

Figure 2: Vertical profiles of radar reflectivity observed by (a) the EarthCARE 94-GHz CPR (frame number: 01520E), and (b) the HALO aircraft 35-GHz radar (b) during the observations around 16:00 UTC on 3 September 2024. Contours indicate radar reflectivity in units of dBZ.

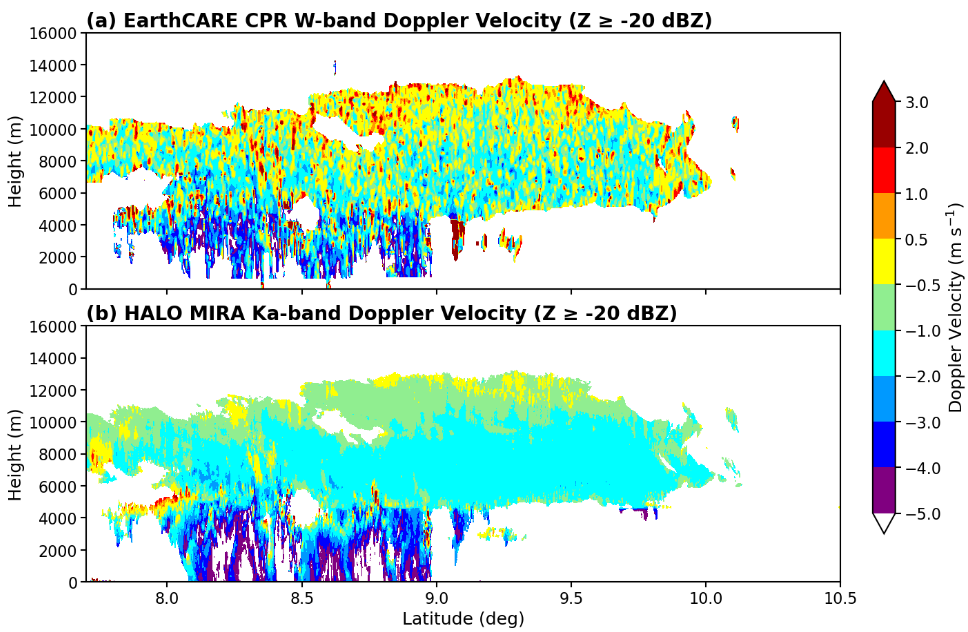

Figure 3: The same case as Fig. 2 but showing vertical profiles of Doppler velocity (positive upwards) observed by (a) EarthCARE CPR and (b) the HALO aircraft radar. Contours indicate radar reflectivity in units of m s^{-1}.

https://doi.org/10.5194/egusphere-2026-3606
Preprint. Discussion started: 16 September 2026

Figure 2 shows vertical profiles of radar reflectivity observed at approximately 16:00 UTC on 3 September 2024 by the EarthCARE CPR and by the HALO aircraft radar over the ORCESTRA domain. The two datasets correspond to nearly the same time and location, with the collocation occurring near 8.6°N. In the EarthCARE profile, a reflectivity maximum near 13 km altitude corresponds to the HALO overflight signal. Both observations exhibit similar overall cloud structures. In particular, a mesoscale convective system (MCS) is evident south of 10° N, while a congestus-type convective structure appears near 11° N.

The EarthCARE CPR operates at W-band (94 GHz), whereas the HALO radar operates at Ka-band (35 GHz). Due to lower attenuation at Ka-band, the HALO radar generally shows stronger reflectivity values and a more clearly defined melting layer than W-band observations. The complementary sensitivities of the two radar frequencies suggest that combined multi-frequency observations provide enhanced constraints on cloud microphysical properties and riming progression.

Figure 3 shows the vertical distribution of Doppler velocity observed at the same location as in Fig. 2. Unlike the earlier CloudSat mission, the EarthCARE Cloud Profiling Radar (CPR) provides Doppler velocity measurements, enabling direct observation of hydrometeor terminal velocities and information on vertical air motions. To reduce uncertainties associated with low signal-to-noise ratios, the analysis presented here is limited to radar reflectivity values greater than −20 dBZ.

The EarthCARE observations clearly capture the contrast between relatively slow-falling ice particles above the melting layer and the faster-falling rain particles below it. A similar Doppler velocity structure is also observed in the HALO aircraft radar measurements, indicating general consistency between the satellite and airborne observations. However, small-scale variations in the EarthCARE Doppler velocity field are apparent. These fluctuations may arise from several sources, including the high orbital velocity of the satellite (approximately 7 km s^{-1}), finite number of pulses in each Doppler estimate, and horizontal heterogeneity within the cloud field. Such factors can introduce additional variability and uncertainty in the retrieved Doppler velocity. The combined analysis of satellite and airborne Doppler observations therefore offers a valuable opportunity to better understand measurement uncertainties and to improve the interpretation of vertical air motion and hydrometeor fall velocities in convective systems.

3.2 Experimental design

The basic experimental protocol is summarized in Table 3. ECOMIP is open to any type of atmospheric model. Here, we focus on global and regional km-scale models for simplicity. Coarser resolution models can also follow this protocol, but additional treatments of convective parameterization is generally required for interpretation (Hashino et al., 2025.; Nakamura et al., 2026.).

Consistent with previous GSRM intercomparison studies, participating km-scale models are required to use a horizontal grid spacing finer than 5 km to ensure that deep convective motions are resolved rather than parameterized. The vertical resolution may be freely chosen by each modeling group, provided that it is adequate to resolve the vertical structure of clouds, precipitation systems, and boundary-layer processes. In general, for km-scale models, the use of cumulus parameterization and gravity wave drag schemes is discouraged. However, it is not required; one must be careful when interpreting models with subgrid parameterization.

https://doi.org/10.5194/egusphere-2026-3606
Preprint. Discussion started: 16 September 2026

Regarding initialization, we suggest using ERA5 reanalysis data or the IFS 9 km operational analysis as initial and boundary conditions; the latter is provided to participants. These datasets provide sufficiently high spatial and temporal resolution to support km-scale simulations and ensure consistency with EarthCARE and ORCESTRA observation periods. As options, participating groups may employ other reanalysis products or high-resolution operational analyses if it is straightforward. However, the dependency on initial conditions may be one of the target analysis of sensitivity of the results.

Table 3. Overview of the basic ECOMIP intercomparison protocol. (*) Note that the horizontal resolution of less than 5 km is requested for global or regional km-scale models. Coarser-resolution models can also participate in ECOMIP.

Setting	Descriptions
Horizontal resolution	Less than 5 km(*)
Vertical resolution	Not specified
Model	Global or regional atmosphere model
Output interval	3D: 30 mins, 2D: 10 mins for Phase 1 3: 3 hours, 2D: 30 mins for Phase 2 and 3
Initialization of land and ocean	Arbitrary up to each group
Initial data	00UTC the 1day before the target case time
Domain size for the regional model	2°S – 22°N, 10°W – 62°W
Initial Condition	ERA5 or IFS 9 km

To enable comprehensive process-level evaluation and satellite-simulator applications, a standardized set of output variables is required. The list of mandatory three-dimensional output variables is provided in Table 4. These include atmospheric state variables (e.g., temperature, pressure, water vapor mixing ratio), hydrometeor mixing ratios or mass concentrations for all prognostic categories (e.g., cloud water, rain, cloud ice, snow, graupel), vertical velocity, and additional microphysical diagnostics where available. The list of required two-dimensional output variables is summarized in Table 5. This includes surface precipitation rates, top-of-atmosphere radiative fluxes (shortwave and longwave), surface fluxes, and column-integrated hydrometeor quantities such as liquid water path (LWP) and ice water path (IWP).

For all output variables, clear documentation of variable names, units, definitions, and diagnostic methods is mandatory. Particular attention must be paid to consistency in unit conventions (e.g., kg kg^{-1} vs. kg m^{-3} for hydrometeors; Pa vs. hPa for pressure; m s^{-1} for velocities) and to explicit specification of whether quantities represent grid-mean, in-cloud, or mass-weighted averages. Any diagnostic transformations applied within the model (e.g., conversion from mixing ratio to specific mass) should be documented to ensure reproducibility and consistent interpretation.

Table 4. List of required three-dimensional output variables for the intercomparison protocol.

No.	**Variable Name**	**Unit**	**No.**	**Variable Name**	**Unit**

https://doi.org/10.5194/egusphere-2026-3606
Preprint. Discussion started: 16 September 2026

1	Model height	m	6	Pressure	Pa
2	Zonal wind velocity	m s⁻¹	7	Specific humidity	kg kg⁻¹
3	Meridional wind velocity	m s⁻¹	8	Mixing ratio for each hydrometer	kg kg⁻¹
4	Vertical wind velocity	m s⁻¹	9	Number concentration for a double moment scheme	kg kg⁻¹
5	Temperature	K	10	Cloud fraction	

Table 5. List of required two-dimensional output variables for the intercomparison protocol.

No.	Variable Name	Unit	No.	Variable Name	Unit
1	Total column ice water	kg m⁻²	16	Upward SW radiation at the surface (Clear sky)	W m⁻²
2	Total column liquid water	kg m⁻²	17	Downward SW radiation at the surface (All sky)	W m⁻²
3	Surface latent heat flux	W m⁻²	18	Downward SW radiation at the surface (Clear sky)	W m⁻²
4	Surface sensible heat flux	W m⁻²	19	Precipitation rate	kg m⁻² s⁻¹
5	Upward LW radiation at the TOA (All sky)	W m⁻²	20	Total column water vapor	kg m⁻²
6	Upward LW radiation at the TOA (Clear sky)	W m⁻²	21	Surface pressure	Pa
7	Downward LW radiation at the TOA	W m⁻²	22	Sea level mean pressure	Pa
8	Upward LW radiation at the surface (All sky)	W m⁻²	23	2 m specific humidity	kg kg⁻¹
9	Upward LW radiation at the surface (Clear sky)	W m⁻²	24	2 m temperature	K
10	Downward LW radiation at the surface (All sky)	W m⁻²	25	10 m zonal wind velocity	m s⁻¹
11	Downward LW radiation at the surface (Clear sky)	W m⁻²	26	10 m meridional wind velocity	m s⁻¹
12	Upward SW radiation at the TOA (All sky)	W m⁻²	27	Surface temperature	K
13	Upward SW radiation at the TOA (Clear sky)	W m⁻²	28	Zonal momentum flux at the surface	N m⁻²
14	Downward SW radiation at the TOA	W m⁻²	29	Meridional momentum flux at the surface	N m⁻²

https://doi.org/10.5194/egusphere-2026-3606
Preprint. Discussion started: 16 September 2026

15	Upward SW radiation at the surface (All sky)	W m⁻²	30	Vertically projected cloud cover	[unitless]

For calculations using a satellite simulator, additional physical and microphysical information is required to represent the scattering and radiative properties of hydrometeors. In particular, it is necessary to specify the size distribution parameters, such as those for gamma or log-normal distributions, or to provide the statistical moments derived from these distributions for each hydrometeor category. Furthermore, the terminal velocity of particles as a function of their diameter must be defined, along with the bulk density of ice particles, including snow and graupel. If applicable, assumptions regarding the nonspherical shapes of ice particles should also be included, as they significantly influence the scattering characteristics and, in turn, the simulated satellite observations.

When using subgrid cloud assumptions, it is essential to describe how these subgrid-scale variations are treated within the satellite simulator framework. Since satellite simulators typically operate on grid-mean model variables, the representation of subgrid cloud heterogeneity—such as fractional cloud cover, variability in hydrometeor content, or unresolved vertical overlap—can substantially affect the simulated observables (e.g., radar reflectivity, lidar backscatter, radiative fluxes). Therefore, explicit assumptions regarding the subgrid distribution of clouds and hydrometeors (e.g., probability density functions of LWC/IWC or optical properties) should be clearly stated, as they influence both the simulated radiances and the interpretation of model–observation comparisons.

Regarding the output format, the planned data products will be provided either on a latitude–longitude grid converted from the model's native grid or in the HEALPix grid format. If the grid conversion process proves technically challenging or computationally demanding, the data will instead be distributed on the native model grid, accompanied by conversion tools that enable users to reproject the data into either a latitude–longitude grid or the HEALPix grid format.

4 Examples of the analysis

4.1 Global view of simulations

To demonstrate the feasibility of the ECOMIP framework, we first show the analysis using three representative GSRMs, NICAM, ICON, and IFS. For Phase 1, both models performed simulations as described for Case 1 in Table 2, corresponding to the EarthCARE overpass on 3 September 2024. Following the ECOMIP protocol, the simulations were initialized at 00:00 UTC on 2 September 2024. For the initial conditions, NICAM used ERA5 reanalysis data, while ICON used ECMWF IFS 9 km operational forecast data due to a technical constraint. The dependency on the initial condition is one of the issues of the intercomparison. The NICAM simulation was performed at a horizontal resolution of 3.5 km, following the experimental configuration of Takasuka et al. (2024). The ICON simulation was conducted according to the framework described in (Fiévet et al., n.d.; Hohenegger et al., 2023). The IFS simulations was performed at a horizontal resolution of 4. 4km.

Figure 4 compares the horizontal distribution of precipitation from the GSMaP observations (Kubota et al. 2020, 2026), NICAM and ICON simulations. For the present study, the IFS output was only available over the Atlantic Ocean domain, and therefore no global precipitation distribution was provided. The NICAM and ICON simulations reproduce the overall precipitation patterns reasonably well when compared with the GSMaP observational data. In particular, the structures of the ITCZ and midlatitude frontal systems are well captured by

https://doi.org/10.5194/egusphere-2026-3606
Preprint. Discussion started: 16 September 2026

both models, with spatial features consistent with observations. However, differences are evident in total precipitation and in regions characterized by weak precipitation, where the two models and observations show noticeable discrepancies. For weak precipitation, the GSMaP observations that rely primarily on passive microwave channels are subject to relatively large uncertainties. The Ka/Ku bands of DPR on board GPM are used for calibration, but they are not sensitive to weak precipitation (Hamada and Takayabu 2016; Masaki et al. 2022). In this context, high-sensitivity radar observations from EarthCARE are expected to provide valuable information for improving the understanding and evaluation of weak precipitation processes in both observations and models.

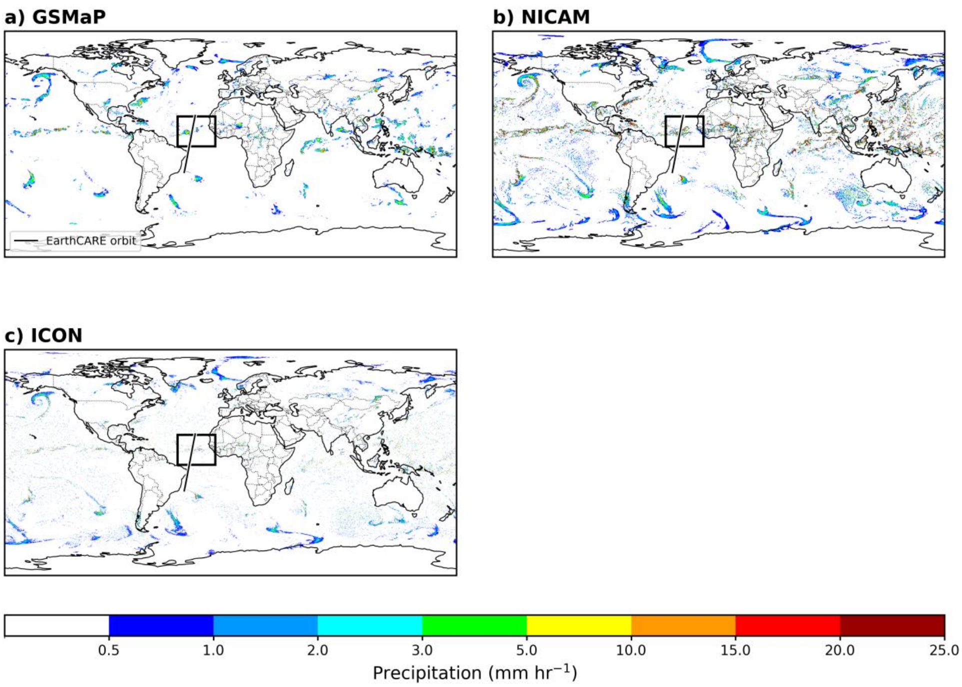

Figure 4: The horizontal distributions of one-hour accumulated precipitation at 16:00 UTC on 3 September 2024 from (a) GSMaP, (b) NICAM, and (c) ICON. The black box indicates the spatial domain of the ORCESTRA campaign. The black line represents the EarthCARE orbit.

https://doi.org/10.5194/egusphere-2026-3606
Preprint. Discussion started: 16 September 2026

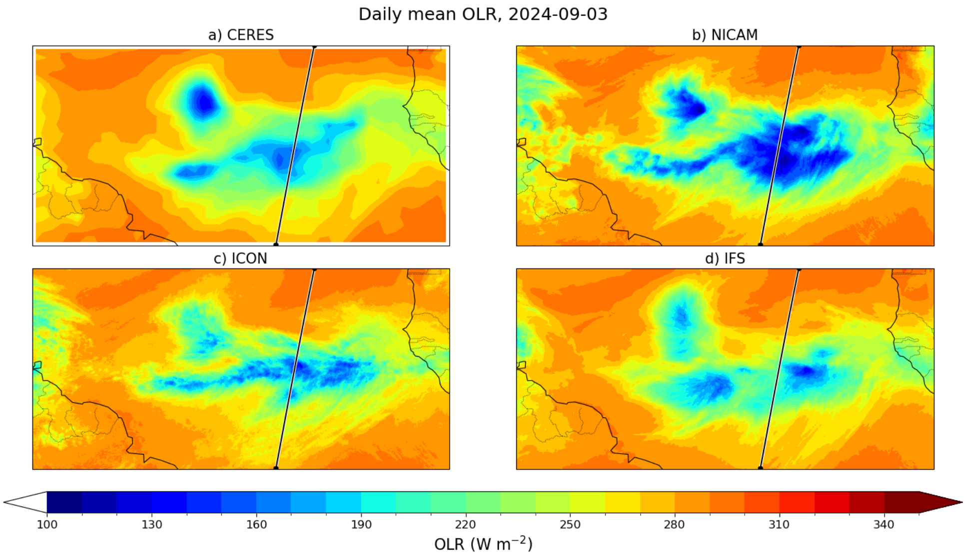

Figure 5: The horizontal distributions of daily averaged Outgoing Longwave Radiation (OLR) on 3 September 2024 over the tropical Atlantic Ocean from (a) CERES, (b) NICAM, (c) ICON, and (d) IFS. The black line represents the EarthCARE orbit.

Figure 5 shows the horizontal distributions of daily averaged OLR over the tropical Atlantic Ocean on 3 September 2024 derived from CERES observations, NICAM, ICON, and IFS simulations. Overall, all models reproduce the large-scale spatial distribution of OLR reasonably well. Regions of reduced OLR associated with deep convective cloud systems are clearly captured in both simulations, indicating that the models successfully represent the large-scale organization of deep clouds. However, differences in OLR magnitude are evident within these convective regions. While the spatial distribution of deep clouds is broadly consistent across the datasets, NICAM tends to underestimate OLR compared with the observations. While the locations of minimum OLR associated with deep convective systems differ slightly among the models, this indicates differences in the simulated structure and organization of deep convection. These differences are related to variations in simulated cloud-top height, anvil cloud extent, or cloud optical properties.

https://doi.org/10.5194/egusphere-2026-3606
Preprint. Discussion started: 16 September 2026

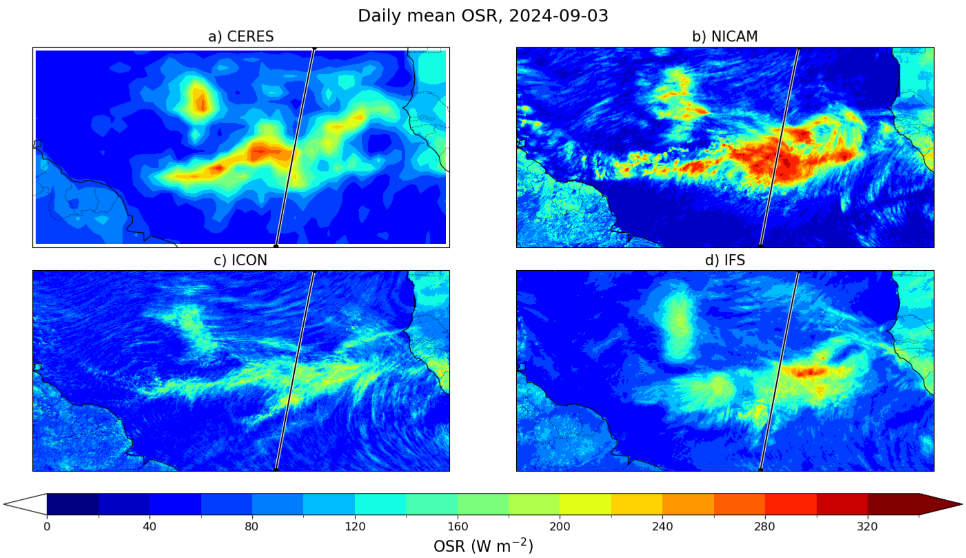

Figure 6: The Horizontal distributions of daily averaged Outgoing Shortwave Radiation (OSR) on 3 September 2024 over the Atlantic Ocean from (a) CERES, (b) NICAM, (c) ICON, and (d) IFS. The black line represents the EarthCARE orbit.

Figure 6 presents the corresponding horizontal distribution of outgoing shortwave radiation (OSR). Regions of deep convection are characterized by strong reflection of incoming solar radiation at the top of the atmosphere, leading to higher OSR values. Three models reproduces these high-reflectance regions reasonably well when compared with CERES observations, indicating that the radiative signatures of deep convective clouds are generally well simulated. Compared with the OLR distribution, the OSR fields show larger differences among the models, indicating a stronger sensitivity to cloud optical properties and cloud vertical structure. In particular, NICAM produces stronger OSR signals associated with deep convective systems than those observed by CERES, suggesting enhanced reflection from optically thicker or more extensive convective clouds. In contrast, ICON shows relatively weaker OSR signals in the same convective regions. The IFS simulation also exhibits enhanced OSR in the convective system located to the right of the EarthCARE orbit. The OSR simulated by NICAM is lower than that observed by CERES and simulated by ICON in regions away from deep convection. These differences indicate the sensitivity of shortwave radiation to the representation of low-level clouds and their optical properties; a characteristic to be explored in the intercomparison of global storm-resolving models.

4.2 Comparisons with the EarthCARE CPR

To enable a direct comparison with EarthCARE observations, synthetic EarthCARE sensor signals can be generated from the model data using J-Sim (Roh et al. 2023). Specifically, radar reflectivity and Doppler velocity fields were calculated from the model output to be consistent with the EarthCARE CPR observation space. For the IFS simulation, radar reflectivity and Doppler velocity were calculated using ZmVar (Fielding and Janisková, 2020), which was developed for the simulation and assimilation of EarthCARE radar and lidar observations in the ECMWF framework.

For the ICON simulation, it is found that detailed information on ice-phase hydrometeors required for radar signal calculations was not fully described. It was due to the design of ICON, in which cloud microphysical process assumptions are not exactly consistent among different cloud microphysics modules. Therefore, in this study, the particle-size-distribution assumptions used in NICAM were applied to estimate radar signals associated with cloud ice, snow, and graupel in the ICON simulation.

We should note that differences in cloud microphysical assumptions between a model and a satellite simulator could generally influence the simulated satellite signals. For the model intercomparison in the satellite signal space, we should carefully set the same assumptions for cloud particle size distributions and other microphysical properties of hydrometeors across the model and satellite simulator. However, this information is not always explicitly constrained or consistently represented in atmospheric numerical models. Through this intercomparison project, we will discuss the desirable configuration of cloud microphysical modules for consistent evaluation by the EarthCARE to gain deeper insights into the cloud microphysical characteristics of hydrometeors and their impact on satellite-observed signals.

The scripts and compact datasets used to reproduce the principal figures and diagnostics presented in this study are publicly available through Zenodo (Roh et al., 2026).

https://doi.org/10.5194/egusphere-2026-3606
Preprint. Discussion started: 16 September 2026

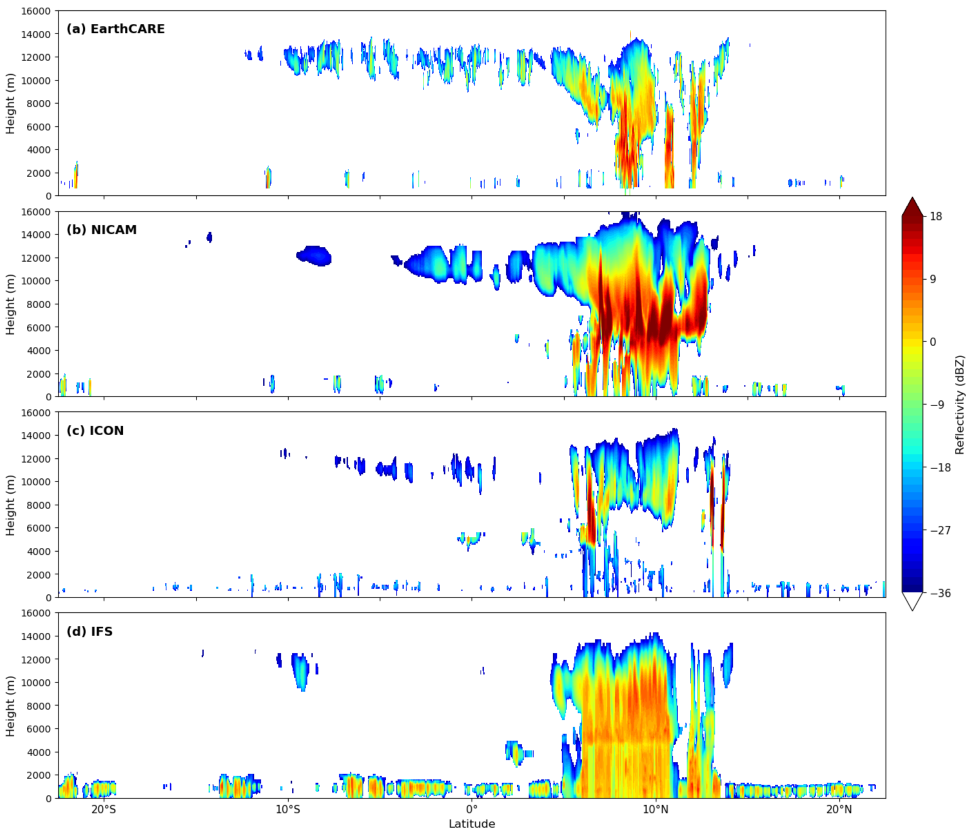

Figure 7: The vertical profiles of radar reflectivity from (a) EarthCARE, (b) NICAM, (c) ICON and (d) IFS.

The EarthCARE satellite passed over the ORCESTRA domain at approximately 16 UTC on 3 September 2024. Using NICAM and ICON simulations, vertical profiles of radar reflectivity were produced in a manner consistent with the EarthCARE observations. In this case, an MCS is observed between 5°N and 15°N within the ORCESTRA campaign region, while cirrus clouds are present in the Southern Hemisphere (Fig. 7a). All three models reproduce the occurrence of the MCS and the Southern Hemisphere cirrus cloud layer (Fig. 7b–d). However, notable differences are evident in the detailed cloud structures and radar reflectivity distributions. In particular, the radar reflectivity associated with cirrus clouds is underestimated in all three models compared with the EarthCARE observations. The NICAM simulation produces a larger and more coherent MCS structure than observed, suggesting stronger convective organization and more extensive anvil development. In contrast, the ICON simulation exhibits a more scattered convective structure with less spatial coherence. The IFS simulation shows relatively low fraction of anvil cirrus clouds and a larger fraction of low-level clouds compared with the observations and the other models. These differences mean model-dependent characteristics in representing convective organization and cloud microphysical properties.

https://doi.org/10.5194/egusphere-2026-3606
Preprint. Discussion started: 16 September 2026

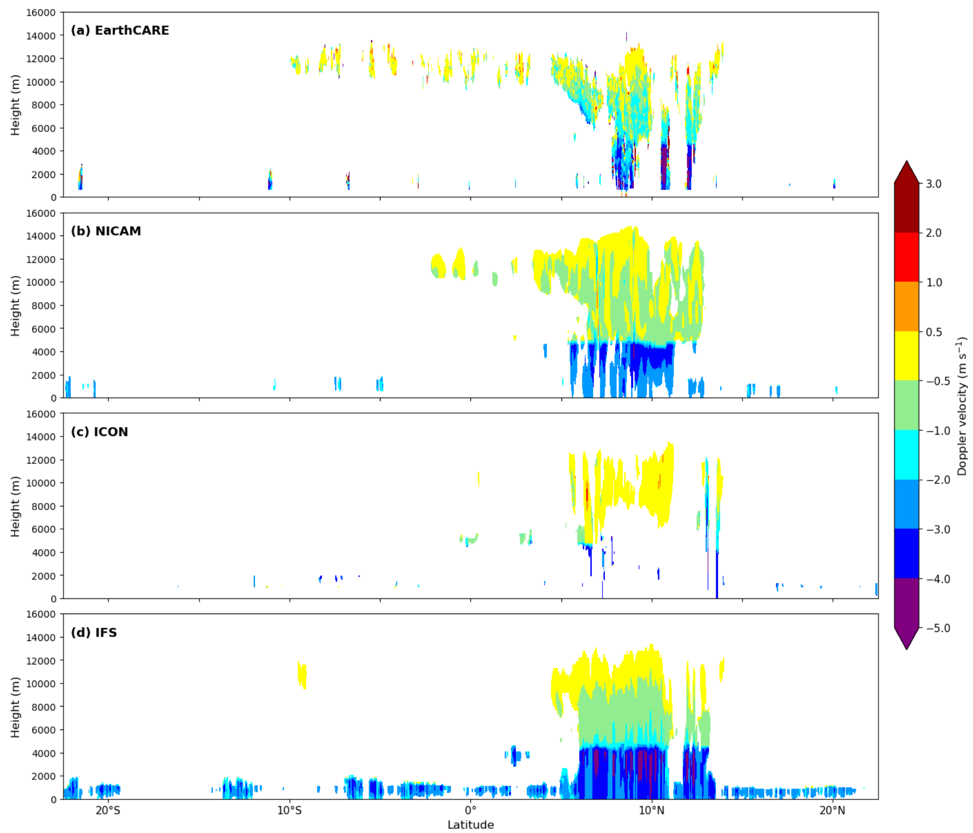

Figure 8: The vertical profiles of Doppler velocity (positive up) from (a) EarthCARE, (b) NICAM, (c) ICON, (d) IFS

A key advantage of the EarthCARE CPR is its capability to observe Doppler velocity by an active radar from space for the first time. This measurement provides valuable information for inferring hydrometeor terminal velocities as well as vertical air motions. In this study, Doppler velocity observations were analysed by applying a signal-to-noise ratio criterion and by restricting the analysis to radar reflectivity values greater than −20 dBZ. The EarthCARE observations show a clear contrast in terminal velocities associated with rain and ice particles near the melting layer, which is located at approximately 5 km altitude (Fig. 8a). Cirrus clouds, in contrast, exhibit relatively low terminal velocities. all three models reproduce similar overall patterns to the observations (Fig. 8b, c, d). For cirrus clouds in the Southern Hemisphere, the simulated radar reflectivity is weaker than observed, particularly in the upper troposphere. In addition, the representation of upward air motions, indicated by orange and red colors, differs substantially among the models. NICAM produces a single convective core, whereas ICON exhibits multiple weaker convective cores with a more fragmented structure. The IFS simulation indicates faster terminal velocities for both ice and liquid-phase hydrometeors compared with the EarthCARE observations, suggesting differences in the representation of particle size distributions or microphysical fall-speed assumptions. These differences indicate model-dependent representations of vertical motion and precipitation processes.

https://doi.org/10.5194/egusphere-2026-3606
Preprint. Discussion started: 16 September 2026

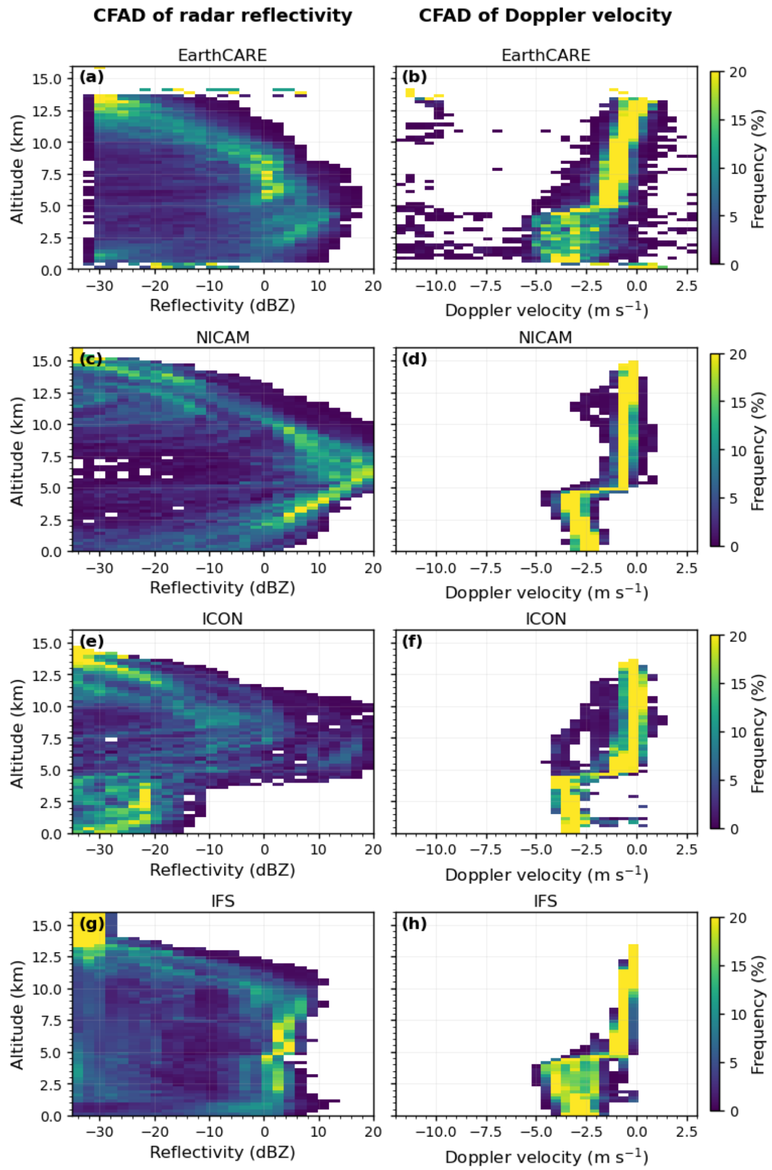

Figure 9: Contoured frequency by altitude diagrams (CFADs) of radar reflectivity (left) and Doppler velocity (right) from EarthCARE, NICAM, ICON, and IFS.. Contours represent normalized frequency (%)

Contoured frequency-by-altitude diagrams (CFADs) were constructed using the datasets presented in Figs. 7 and 8 (Fig. 9). The CFADs of radar reflectivity reveal distinct differences among the datasets. In NICAM, ice-phase radar reflectivity is systematically overestimated compared to EarthCARE observations. In contrast, ICON underestimates radar reflectivity below the melting layer, resulting in weaker simulated rain signatures than

https://doi.org/10.5194/egusphere-2026-3606
Preprint. Discussion started: 16 September 2026

observed. The IFS simulation shows enhanced radar reflectivity in the ice phase, while exhibiting a distinct reduction of maximum radar reflectivity near the melting layer.

For Doppler velocity, all models reproduce the characteristic vertical structure, with an abrupt change near the melting layer at about 5 km. This indicates the faster rain particles below 5 km and the slower ice particles above 5 km. However, notable differences emerge below the melting layer. NICAM underestimates Doppler velocities associated with liquid precipitation and shows a relatively narrow variance, suggesting limited variability in fall speeds or vertical motions. In contrast, ICON shows a broader distribution of Doppler velocities near the boundary layer, indicating a wider range of fall velocities and stronger variability in convective and turbulent motions associated with rain and low-level clouds.

The CFADs of the Doppler velocity are shown in Figs. 9b, d, f, h possess information of upward motion. The Doppler velocity is the summation of vertical velocity and the fall velocity of hydrometeor particles. For NICAM and ICON, the positive values of Doppler velocity indicate the presence of upward motion. Because of the small sample size, the CFADs of NICAM and ICON show upward motion profiles corresponding to specific convective cores, as indicated by the maximum values of the probability distributions. In contrast, the IFS simulation exhibits a lower frequency of upward-motion signatures compared with NICAM and ICON. This separation between air vertical motion and cloud particle fall speed in the EarthCARE data is available in the EarthCARE research product, which is shown below. For models, these values are directly obtained. Roh et al. (2025) discussed a methodology to better understand meso-scale circulation by the synergetic comparison of the EarthCARE data and the NICAM simulation.

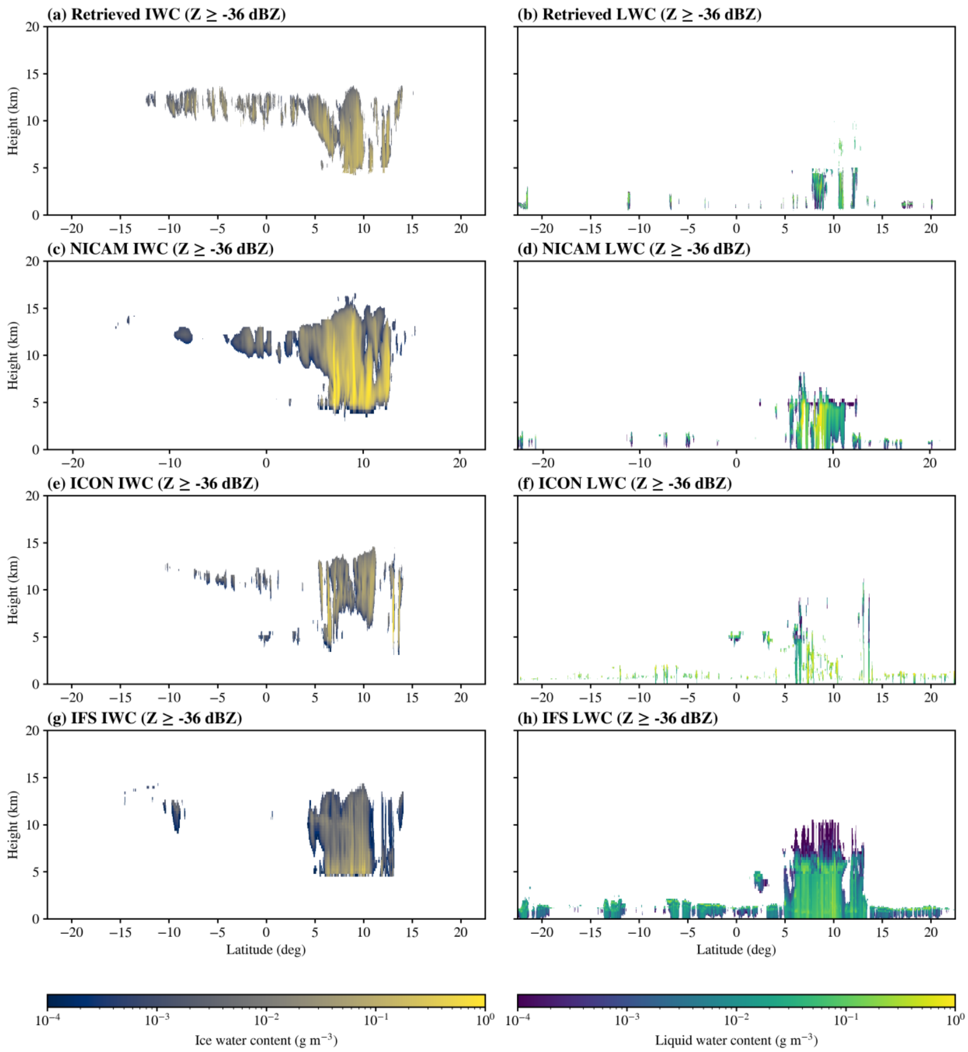

Figure 10: The vertical cross sections of IWC (left) and LWC (right) from (a,d) EarthCARE, (b,e) NICAM, (c,f) ICON, and (g,h) IFS.

Figure 10 shows the vertical cross sections of IWC and LWC derived from EarthCARE observations and from the NICAM, ICON, and IFS simulations. For EarthCARE, IWC and LWC are retrieved using radar reflectivity and Doppler velocity measurements from the EarthCARE CPR (Sato et al., 2025). The retrieval data used in this study correspond to the JAXA Level-2 CPR_CLP product (version vBb), which is based solely on observations from the EarthCARE CPR sensor. For NICAM and ICON, the corresponding model output variables are used. To ensure consistent sampling between observations and simulations, only data with radar reflectivity greater than −36 dBZ are included in the comparison.

https://doi.org/10.5194/egusphere-2026-3606
Preprint. Discussion started: 16 September 2026

Overall, the retrieved and simulated datasets exhibit similar vertical structures and hydrometeor collocation. Both observations and model simulations show comparable vertical distributions of ice and liquid water, indicating that the cloud microphysical structure of the systems is well represented by the models. However, the vertical distribution and magnitude of IWC differ among the models. NICAM tends to overestimate IWC relative to the retrieval data, whereas IFS tends to underestimate it.

The vertical profiles of LWC exhibit broadly similar patterns across the datasets. In particular, supercooled liquid water is present above the melting layer in both the observations and the simulations. However, the amount and vertical distribution of supercooled liquid water vary substantially among the datasets, suggesting differences in the representation of mixed-phase cloud microphysics and associated processes such as riming and condensational growth.

It should be noted that observations of supercooled liquid water within deep convective clouds are very challenging and that the retrieval products contain substantial uncertainties. Nevertheless, comparisons among different model simulations and the retrieval data can provide new insights into the representation of supercooled liquid water and help improve our understanding of mixed-phase processes. In particular, supercooled liquid water can influence Doppler velocity through riming processes and may also affect path-integrated attenuation (PIA). Using the model simulations, it is possible to investigate how strongly these processes influence the simulated radar signals and retrieval characteristics.

https://doi.org/10.5194/egusphere-2026-3606
Preprint. Discussion started: 16 September 2026

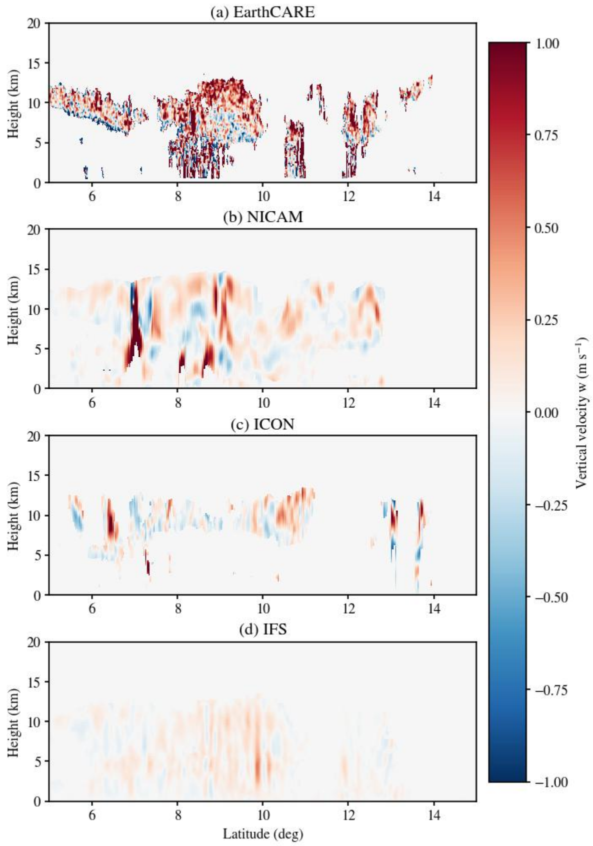

Figure 11: The vertical cross sections of vertical air velocity between 5°N and 15°N from (a) EarthCARE, (b) NICAM, (c) ICON, and (d) IFS.

Figure 11 shows the vertical profiles of vertical air velocity derived from EarthCARE observations and from the NICAM, ICON, and IFS simulations between 5°N and 15°N. For EarthCARE, vertical air motion can be inferred from Doppler velocity measurements by separating the contribution of hydrometeor terminal velocity from the observed Doppler signal (Sato and Okamoto, 2025). This retrieval is sensitive to a variety of factors, including assumptions regarding particle fall speeds and cloud microphysical properties, Doppler velocity observation errors, and aliasing effects, making it a particularly challenging product to interpret and evaluate.

In contrast, model simulations provide direct estimates of vertical air velocity from the dynamical variables, allowing the vertical motion component to be separated more straightforwardly from the simulated Doppler velocity. As a result, the model-derived vertical velocity profiles generally exhibit smoother structures compared with those derived from observations.

The EarthCARE retrieval shows greater variability in vertical velocity than the model results. This enhanced variability is partly attributable to measurement uncertainties in satellite-based Doppler velocity observations, including effects related to platform motion, instrumental noise, and small-scale heterogeneity in the cloud field.

https://doi.org/10.5194/egusphere-2026-3606
Preprint. Discussion started: 16 September 2026

Despite these uncertainties, the observational estimates provide valuable constraints on the magnitude and vertical distribution of atmospheric motions within convective systems (Roh et al. 2025).

All simulations exhibit notable differences in both the distribution and intensity of vertical air motions. These differences reflect model-dependent representations of convective dynamics, cloud microphysical processes, and the mechanisms of precipitation formation. NICAM shows relatively stronger upward motions than the other models, whereas ICON represents smaller-scale upward and downward motions more clearly, likely because of its higher horizontal resolution. In contrast, the IFS simulation exhibits smaller amplitudes of vertical velocity compared with NICAM and ICON.

At the same time, uncertainties also exist in the vertical air velocity retrievals, and it is also important to note that conclusions based on a single EarthCARE frame should be interpreted with caution. In addition, the mean vertical velocity characteristics are expected to vary depending on the type and organization of cloud systems. Nevertheless, understanding model-dependent differences in vertical velocity remains an important research topic. Synergistic comparisons between models and Doppler radar observations can provide valuable insight not only into model behavior but also into the interpretation and complementary understanding of the observational data themselves.

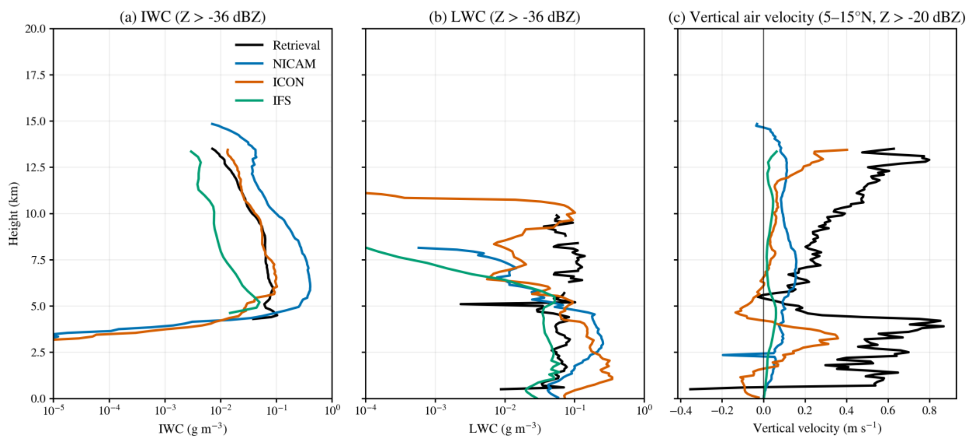

Figure 12: The vertical profiles of IWC, LWC, and vertical air velocity (w) from JAXA retrieval data, NICAM, ICON, and IFS.

Figure 12 shows the horizontally averaged vertical profiles of IWC, LWC, and vertical air velocity for the cross sections shown in Figs. 10 and 11. The averaged IWC in the ICON simulation is generally consistent with the retrieval data, whereas NICAM shows an overestimation and IFS shows an underestimation. The underestimation in IFS is likely related to the smaller amount of anvil cirrus clouds compared with the EarthCARE observations. In contrast, the overestimation in NICAM is consistent with the tendency of the model to produce stronger radar reflectivity than observed by EarthCARE.

For LWC, the differences among the models are generally smaller than those found for IWC below the melting layer. However, notable differences remain in the vertical distribution and cloud fraction of liquid water. ICON

https://doi.org/10.5194/egusphere-2026-3606
Preprint. Discussion started: 16 September 2026

simulates larger LWC values than NICAM and IFS, particularly above the melting layer, where the amount of supercooled liquid water is also enhanced. This result suggests stronger convective activity and more active mixed-phase cloud processes in ICON.

The IFS simulation shows horizontally averaged LWC values comparable to those of NICAM, but exhibits a larger liquid cloud fraction. This indicates that IFS produces more widespread clouds containing relatively small amounts of liquid water, whereas NICAM tends to simulate fewer clouds with comparatively larger LWC values.

The averaged vertical air velocity shows differences between the observations and the models. In this case, the retrieved vertical air velocity from EarthCARE indicates stronger upward motions than those simulated by the models. The distributions of vertical motion also vary among the three models. NICAM shows relatively strong upward motions above the melting layer, whereas IFS exhibits a more hydrostatic structure than the other models. ICON, in contrast, shows relatively stronger upward motions below the melting layer. Because convective cores are highly sensitive to the location of the observation orbit, it is difficult to generalize the results from a single frame. Nevertheless, the comparison helps reveal characteristic differences among the models.

These findings show the value of comparing EarthCARE observations with multiple km-scale models, providing a pathway to improve our understanding of cloud microphysical processes and their representation in numerical models. Since supercooled liquid water plays an important role in riming processes and convective development, further investigation of the microphysical processes associated with convection in different models would be of interest.

https://doi.org/10.5194/egusphere-2026-3606
Preprint. Discussion started: 16 September 2026

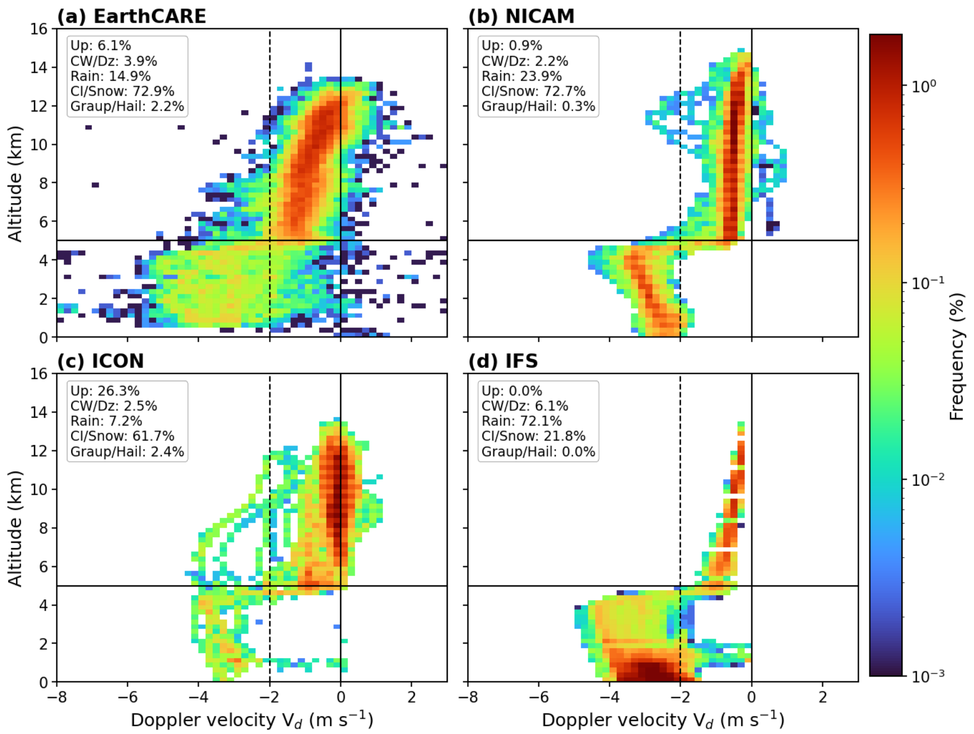

Figure 13: The joint histograms of height and Doppler velocity for a single frame from (a) EarthCARE, (b) NICAM, (c) ICON, and (d) IFS. Contours represent the logarithm of the normalized frequency (%).

Figure 13 presents the joint histograms of height and Doppler velocity using the same dataset as in Fig. 7. (Roh et al., 2024) classified ground-based radar observations into five regimes based on Doppler velocity, distinguishing different hydrometeor types and a positive upward-motion regime associated with atmospheric ascent. For the EarthCARE observations in this frame, the largest fraction corresponds to high clouds associated with cloud ice and snow (72.9%), followed by rain (14.9%). Both NICAM and ICON also show the largest fraction in the cloud ice/snow regime In contrast, the IFS simulation shows the largest fraction in the rain regime, with particularly high frequencies below approximately 2 km altitude within the boundary layer. However, differences are evident in the secondary regimes: in NICAM, rain represents the second-largest fraction (23.9%), whereas in ICON, the upward-motion regime accounts for a larger fraction (26.3 %). In the IFS simulation, the cloud ice/snow regime represents the second-largest fraction (21.8%). These results indicate model-dependent differences in the representation of Doppler velocity regimes. In particular, NICAM appears to underestimate the cloud water and drizzle regimes, while ICON tends to underestimate the rain regime The IFS microphysics scheme does not include graupel, and therefore the graupel/hail regime is absent in the IFS results. Such differences in the distribution of Doppler velocity regimes may influence cloud lifetime and storm dynamics, introducing the importance of regime-based evaluation using Doppler radar observations.

https://doi.org/10.5194/egusphere-2026-3606
Preprint. Discussion started: 16 September 2026

4.3 Comparisons with the ORCESTRA data

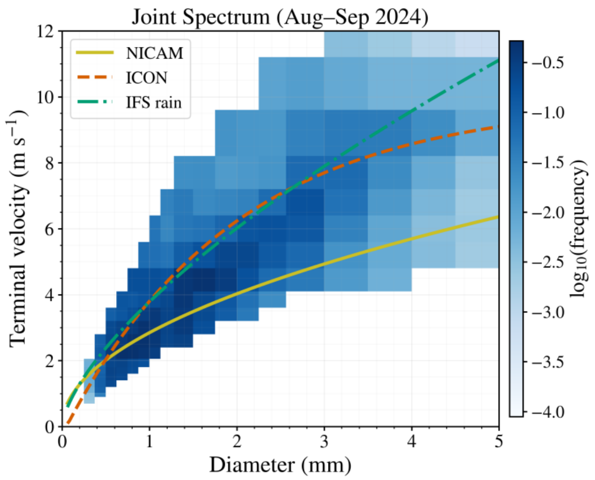

Figure 14: Joint spectrum of particle diameter and fall velocity measured by the disdrometer onboard the research vessel Meteor during August–September 2024. Contours represent the logarithm of the normalized frequency.

The ORCESTRA dataset provides a high degree of observational synergy by combining multiple active sensors, including airborne Ka-band and W-band Doppler radar observations from research aircraft, as well as extensive measurements from the research vessel *Meteor*. Ship-based observations include numerous radiosonde and dropsonde launches, polarimetric radar measurements, and surface-based instruments, offering complementary perspectives on cloud and precipitation processes. Figure 14 presents disdrometer observations collected on board *Meteor* that provide valuable information on the relationship between raindrop size and terminal velocity, a parameterization commonly used in numerical models. These observations are expected to support the evaluation and refinement of rain microphysical parameterizations. The formulation used in ICON and IFS appear to be closer to the observed relationship than that employed in NICAM, although the NICAM formulation still falls within the observed spectral range. Taken together, the diversity and redundancy of the ORCESTRA observations increase the degrees of freedom available for model evaluation and are expected to improve understanding of model parameter choices and associated uncertainties.

5 Possible extensions

The present intercomparison framework mainly focuses on cloud microphysics, convection, and precipitation processes, but it can be extended to additional research areas. One important extension is the inclusion of aerosol transport models. EarthCARE ATLID and airborne lidar observations from ORCESTRA provide vertically resolved aerosol information that can help evaluate aerosol transport, aerosol–cloud interactions, and aerosol radiative effects in atmospheric models.

Another important extension involves the use of multiple satellite simulators. Because different simulators use different assumptions for particle scattering, particle shape, and numerical methods, simulated radar and lidar signals may differ even for the same model output. Intercomparison among satellite simulators will therefore help separate uncertainties associated with model physics from those introduced by the forward operators themselves.

The framework can also be applied to other meteorological phenomena, including tropical cyclones. EarthCARE CPR observations provide vertically resolved radar reflectivity and Doppler velocity measurements that may help evaluate eyewall structure, rainbands, and vertical motions in km-scale simulations of tropical cyclones. Future analyses using multiple EarthCARE cases may improve understanding of model-dependent representations of tropical cyclone dynamics and cloud microphysical processes.

6 Summary

This study introduces the EarthCARE-ORCESTRA Model Intercomparison Project (ECOMIP) and presents its experimental protocol for observation-constrained evaluation of atmospheric numerical models. The framework integrates synergistic datasets from the EarthCARE mission and the ORCESTRA campaign, together with forward satellite simulators. As an initial demonstration, we apply the protocol to two representative global storm-resolving models (GSRMs), NICAM, ICON, and IFS.

Comparison of horizontal precipitation fields against GSMaP observations indicates that both NICAM and ICON reproduce the large-scale spatial distribution of precipitation, including the structure of midlatitude frontal systems. Nevertheless, differences in total precipitation amount and in regions of weak precipitation remain evident. It is also inferred that the precipitation dataset based on passive microwave retrievals is subject to relatively large uncertainties in light-precipitation regimes, and that the high-sensitivity radar observations from EarthCARE offer an important complementary perspective for evaluating weak precipitation processes in both observations and models.

Vertical radar reflectivity comparisons along the EarthCARE overpass reveal that all models capture the presence of a mesoscale convective system (MCS) within the ORCESTRA domain and reproduce Southern Hemisphere cirrus clouds. However, cirrus reflectivity is underestimated in both simulations, and structural differences are apparent in the MCS organization. NICAM and IFS produce a larger and more coherent convective system, whereas ICON exhibits a more scattered structure. These differences point to model-dependent representations of convective organization and cloud microphysical properties.

Doppler velocity observations from the EarthCARE Cloud Profiling Radar (CPR) provide additional constraints on vertical motion and hydrometeor sedimentation. Both models reproduce the overall contrast in terminal velocities between rain and ice particles near the melting layer. However, differences in upward-motion signatures and in the distribution of Doppler velocity regimes are evident. Regime-based analysis indicates clear differences in the dominant Doppler velocity regimes among the observations and models. The EarthCARE observations, NICAM, and ICON are mainly dominated by the high-cloud ice/snow regime, whereas the IFS simulation is dominated by the rain regime. In addition, NICAM shows a relatively larger rain fraction, while ICON exhibits a stronger upward-motion regime. The IFS simulation does not show a distinct graupel/hail regime, which is consistent with the absence of graupel and hail categories in its microphysics scheme. These differences

https://doi.org/10.5194/egusphere-2026-3606
Preprint. Discussion started: 16 September 2026

suggest variations in the representation of hydrometeor types, precipitation formation, convective intensity, and vertical transport processes among the models.

A comparison of retrieved ice water content (IWC) and liquid water content (LWC) further demonstrates microphysical contrasts among the models. NICAM tends to overestimate IWC relative to the retrieval data, particularly in the upper troposphere, whereas ICON shows better agreement with the retrieved IWC profiles. In contrast, IFS tends to underestimate IWC, which is consistent with its weaker representation of anvil cirrus clouds. ICON also produces relatively larger amounts of supercooled liquid water compared with the other models. Although uncertainties exist in the observational retrievals and conclusions based on a single EarthCARE frame should be interpreted with caution, these results demonstrate the importance of extending the analysis to a larger number of EarthCARE cases in order to better understand model-dependent cloud microphysical characteristics. ORCESTRA in situ observations provide an additional degree of constraint. Disdrometer measurements of the raindrop size–terminal velocity relationship indicate that the ICON and IFS fall-speed formulations is closer to the observed relationship than that used in NICAM, although all lie within the observed spectral range. The combination of airborne radar, ship-based polarimetric radar, radiosondes, and surface observations, together with EarthCARE observations, will help improve our understanding of model characteristics.

The present analysis represents Phase 1 of the proposed three-phase framework. Future work will extend the evaluation to two-month free-run simulations (Phase 2) in order to assess the statistical analysis. In Phase 3, observationally constrained nudging experiments will be tested to quantify how microphysical and dynamical sensitivities propagate into radiation budgets, precipitation statistics, and convective organization. Through this structured progression from case-based diagnostics to constrained simulations, the framework provides a pathway toward observation-informed development of atmospheric numerical models.

Code and data availability

The code and compact datasets used to reproduce the figures and analyses presented in this study are publicly available through Zenodo: Roh et al. (2026), Code and data for "ECOMIP: The EarthCARE-ORCESTRA Model Intercomparison Project" (https://doi.org/10.5281/zenodo.21032861). The Zenodo archive contains the scripts and compact datasets required to reproduce the principal figures and diagnostics presented in this paper, including the radar reflectivity, Doppler velocity, ice water content (IWC), liquid water content (LWC), and vertical air velocity comparisons.

The original EarthCARE observation products used in this study are publicly available through the JAXA G-Portal and the ESA MAAP platforms, subject to their respective data access policies. The Joint-Simulator for Satellite Simulators is available from https://www.eorc.jaxa.jp/theme/Joint-Simulator/userform/js_userform.html.

Author contributions

WR conducted the NICAM simulation and analysis using the Joint-Simulator for Satellite Sensors. RF conducted the ICON simulation. MF conducted the IFS simulation. TK provides the EarthCARE data. HO helps with the information on vertical velocity retrieval. MS, RH, and BS coordinated the ECOMIP framework.

Competing interests

Two of the co-authors, Robin Hogan and Hajime Okamoto, serve as Topic Editors for the inter-journal Special Issue "Early results from EarthCARE" (AMT/ACP/GMD). The other authors declare that they have no competing interests.

Acknowledgements

The authors thank members of the JAXA EarthCARE Science Team and the J-Sim project.

Financial support

This work was supported by the EarthCARE satellite study commissioned by the Japan Aerospace Exploration Agency. This work was supported by the JSPS Core-to-Core Program (grant number: JPJSCCA20220001) and the Warm World at the Max Planck Institute for Meteorology. This work was supported by JSPS KAKENHI B (Grant Number 24K00703) and KAKENHI A (Grant Number JP24H00275). Simulations of NICAM in this study were performed on the supercomputer Fugaku (proposal numbers hp250068, hp260066). We also thank the joint research program of the Institute for Space-Earth Environmental Research (ISEE), Nagoya University.

References

Bodas-Salcedo, A., M. J.Webb, M. E.Brooks, M. A.Ringer, K. D.William, S. F.Milton, and D. R.Wilson: Evaluating cloud systems in the Met Office global forecast model using simulated CloudSat radar reflectivities, *J. Geophys. Res*., **113**, D00A13, doi:10.1029/2007JD009620, 2008.

Bodas-Salcedo, A. , Webb, M. J. , Bony, S. , Chepfer, H. , Dufresne, J., Klein, S. A. , Zhang, Y. , Marchand, R. , Haynes, J. M. , Pincus, R. , & John, V. O,: COSP: Satellite simulation software for model assessment. *Bull. Amer. Meteor. Soc.*, **92**, 1023–1043, https://doi.org/10.1175/2011BAMS2856.1, 2011.

Delanoë, J., Hogan, R. J., Forbes, R. M., Bodas-Salcedo, A., & Stein, T. H.: Evaluation of ice cloud representation in the ECMWF and UK Met Office models using CloudSat and CALIPSO data. *Quarterly Journal of the Royal Meteorological Society*, *137*(661), 2064-2078, 2011.

Fielding MD, Janisková M. Direct 4D-Var assimilation of space-borne cloud radar reflectivity and lidar backscatter. Part I: Observation operator and implementation. Q J R Meteorol Soc., 146, 3877–3899, https://doi.org/10.1002/qj.3878, 2020.

Fiévet, R., Linardakis, L., Kornblueh, L., Daniel-Lacombe, M., Gloeckner, H. M., Fiolleau, T., Kluft, L., Hohenegger, C., and Stevens, B.: Tropical convection and storm-resolving modeling: a case study based on the ORCESTRA campaign, ESS Open Arch., , https://doi.org/10.22541/essoar.177315810.09022965/v1, 2026.

Forbes, R. M., & Ahlgrimm, M.: On the representation of high-latitude boundary layer mixed-phase cloud in the ECMWF global model. *Monthly Weather Review*, *142*(9), 3425-3445, 2014.

Hamada, A. and Takayabu, Y. N.: Improvements in detection of light precipitation with the Global Precipitation Measurement dual-frequency precipitation radar (GPM DPR), J. Atmos. Oceanic Technol., 33, 653–667, https://doi.org/10.1175/JTECH-D-15-0097.1, 2016.

Hashino, T., Satoh, M., Hagihara, Y., Kubota, T., Matsui, T., Nasuno, T., and Okamoto, H.: Evaluating cloud microphysics from NICAM against CloudSat and CALIPSO, J. Geophys. Res. Atmospheres, 118, 7273–7292, https://doi.org/10.1002/jgrd.50564, 2013.

Hashino, T., Satoh, M., Kubota, T., Koshiro, T., Okamoto, K., Hagihara, Y., Okamoto, H., and Seiki, T.: A Sub-Grid Precipitation Generator Based on NICAM for Simulating Cloud Radar Signals With GCMs, https://doi.org/10.1029/2024JD042597, 2025.

Hohenegger, C., Korn, P., Linardakis, L., Redler, R., Schnur, R., Adamidis, P., Bao, J., Bastin, S., Behravesh, M., Bergemann, M., Biercamp, J., Bockelmann, H., Brokopf, R., Brüggemann, N., Casaroli, L., Chegini, F., Datseris, G., Esch, M., George, G., Giorgetta, M., Gutjahr, O., Haak, H., Hanke, M., Ilyina, T., Jahns, T., Jungclaus, J., Kern, M., Klocke, D., Kluft, L., Kölling, T., Kornblueh, L., Kosukhin, S., Kroll, C., Lee, J., Mauritsen, T., Mehlmann, C., Mieslinger, T., Naumann, A. K., Paccini, L., Peinado, A., Praturi, D. S., Putrasahan, D., Rast, S., Riddick, T., Roeber, N., Schmidt, H., Schulzweida, U., Schütte, F., Segura, H., Shevchenko, R., Singh, V., Specht, M., Stephan, C. C., von Storch, J.-S., Vogel, R., Wengel, C., Winkler, M., Ziemen, F., Marotzke, J., and Stevens, B.: ICON-Sapphire: simulating the components of the Earth system and their interactions at kilometer and subkilometer scales, Geosci. Model Dev., 16, 779–811, https://doi.org/10.5194/gmd-16-779-2023, 2023.

Illingworth, A. J., Barker, H. W., Beljaars, A., Ceccaldi, M., Chepfer, H., Clerbaux, N., Cole, J., Delanoë, J., Domenech, C., and Donovan, D. P.: The EarthCARE satellite: The next step forward in global measurements of clouds, aerosols, precipitation, and radiation, Bull. Am. Meteorol. Soc., 96, 1311–1332, 2015.

Kubota, T., Aonashi, K., Ushio, T., Shige, S., Takayabu, Y. N., Kachi, M., Arai, Y., Tashima, T., Masaki, T., Kawamoto, N., Mega, T., Yamamoto, M. K., Hamada, A., Yamaji, M., Liu, G., and Oki, R.: Global satellite mapping of precipitation (GSMaP) products in the GPM era, in: Satellite Precipitation Measurement, vol. 67, edited by: Springer, pp. 355–373, https://doi.org/10.1007/978-3-030-24568-9_20, 2020.

Kubota, T., Yamamoto, M. K., Ito, M., Tashima, T., Hirose, H., Ushio, T., Aonashi, K., Shige, S., Hamada, A., Yamaji, M., Yoshida, N., and Kachi, M.: Construction of a longer-term and more homogeneous GSMaP precipitation dataset, in: Radar Meteorology and its Applications, edited by: Chandrasekar, V., Das, S., and Datta, A., Springer Atmospheric Sciences, Springer, Singapore, https://doi.org/10.1007/978-981-95-5409-6_9, 2026.

Kubota, T., Kikuchi, M., Muto, M., Hashimoto, M., Imura, Y., Maruyama, K., Hoffmann, A., Hummel, T., Koopman, R., Malina, E., Rusli, S., Tzallas, V., von Bismarck, J., and Frommknecht, B.: Summary of the EarthCARE Science and Validation Workshop 2025 ? Early achievements and future perspectives for the Earth Cloud Aerosol and Radiation Explorer (EarthCARE) satellite mission, J. Eur. Meteorol. Soc., 100037, https://doi.org/10.1016/j.jemets.2026.100037, 2026

Masaki, T., Iguchi, T., Kanemaru, K., Furukawa, K., Yoshida, N., Kubota, T., and Oki, R.: Calibration of the dual-frequency precipitation radar (DPR) onboard the Global Precipitation Measurement (GPM) core observatory, IEEE Trans. Geosci. Remote Sens., https://doi.org/10.1109/TGRS.2020.3039978, 2022.

Malardel, S, Wedi N, Deconinck W, Diamantakis M, Kuehnlein C, Mozdzynski G, Hamrud M, Smolarkiewicz P: A new grid for the IFS. ECMWF Newsl 146:23–28. https://doi.org/doi:10.21957/zwdu9u5i, 2016.

Matsui, T., Chern, J.-D., Tao, W.-K., Lang, S., Satoh, M., Hashino, T., and Kubota, T.: On the Land–Ocean Contrast of Tropical Convection and Microphysics Statistics Derived from TRMM Satellite Signals and Global Storm-Resolving Models, https://doi.org/10.1175/JHM-D-15-0111.1, 2016.

https://doi.org/10.5194/egusphere-2026-3606
Preprint. Discussion started: 16 September 2026

Mech, M., Maahn, M., Kneifel, S., Ori, D., Orlandi, E., Kollias, P., Schemann, V., and Crewell, S.: PAMTRA 1.0: the Passive and Active Microwave radiative TRAnsfer tool for simulating radiometer and radar measurements of the cloudy atmosphere, Geosci. Model Dev., 13, 4229–4251, https://doi.org/10.5194/gmd-13-4229-2020, 2020.

Nakamura, Y., Suzuki, K., and Horie, H.: Incorporating W-Band Doppler Velocity Signal Simulator Into COSP2: Model Evaluation Against Ground-Based Radar Measurement, https://doi.org/10.1029/2025MS004958, 2026.

Roh, W., Satoh, M., Hashino, T., Okamoto, H., and Seiki, T.: Evaluations of the thermodynamic phases of clouds in a cloud-system-resolving model using CALIPSO and a satellite simulator over the Southern Ocean, J. Atmospheric Sci., 77, 3781–3801, 2020.

Roh, W., Satoh, M., Hashino, T., Matsugishi, S., Nasuno, T., and Kubota, T.: Introduction to EarthCARE synthetic data using a global storm-resolving simulation, Atmospheric Meas. Tech., 16, 3331–3344, 2023.

Roh, W., Satoh, M., Hagihara, Y., Horie, H., Ohno, Y., and Kubota, T.: An evaluation of microphysics in a numerical model using Doppler velocity measured by ground-based radar for application to the EarthCARE satellite, Atmospheric Meas. Tech., 17, 3455–3466, https://doi.org/10.5194/amt-17-3455-2024, 2024.

Roh, W., Satoh, M., Matsugishi, S., Aoki, S., Kubota, T., and Okamoto, H.: Vertical motions in clouds from EarthCare satellite and a global storm-resolving modeling, Sci. Rep., 16, 2622, https://doi.org/10.1038/s41598-025-32256-8, 2025.

Roh, W., Satoh, M., Fiévet, R., Fielding, M., Hogan, R., Stevens, B., Kubota, T., & Okamoto, H.: Code and data for "ECOMIP: The EarthCARE-ORCESTRA Model Intercomparison Project" [Data set]. Zenodo. https://doi.org/10.5281/zenodo.21032861, 2026.

Sato, K. and Okamoto, H.: EarthCARE L2 cloud-precipitation and vertical velocity products, in: Remote Sensing of Clouds and the Atmosphere XXX, Remote Sensing of Clouds and the Atmosphere XXX, 313–316, https://doi.org/10.1117/12.3070102, 2025.

Sato, K., Okamoto, H., Nishizawa, T., Jin, Y., Nakajima, T. Y., Wang, M., Satoh, M., Roh, W., Ishimoto, H., and Kudo, R.: JAXA Level 2 cloud and precipitation microphysics retrievals based on EarthCARE radar, lidar, and imager: the CPR_CLP, AC_CLP, and ACM_CLP products, Atmospheric Meas. Tech., 18, 1325–1338, 2025.

Satoh, M., Matsuno, T., Tomita, H., Miura, H., Nasuno, T., and Iga, S.-I.: Nonhydrostatic icosahedral atmospheric model (NICAM) for global cloud resolving simulations, J. Comput. Phys., 227, 3486–3514, 2008.

Satoh, M., Tomita, H., Yashiro, H., Miura, H., Kodama, C., Seiki, T., Noda, A. T., Yamada, Y., Goto, D., Sawada, M., Miyoshi, T., Niwa, Y., Hara, M., Ohno, T., Iga, S., Arakawa, T., Inoue, T., and Kubokawa, H.: The Non-hydrostatic Icosahedral Atmospheric Model: description and development, Prog. Earth Planet. Sci., 1, 18, https://doi.org/10.1186/s40645-014-0018-1, 2014.

Satoh, M., Stevens, B., Judt, F., Khairoutdinov, M., Lin, S. J., Putman, W. M., and Düben, P.: Global cloud-resolving models. Curr Clim Change Rep 5: 172–184, 2019.

Satoh, M., Matsugishi, S., Roh, W., Ikuta, Y., Kuba, N., Seiki, T., Hashino, T., and Okamoto, H.: Evaluation of cloud and precipitation processes in regional and global models with ULTIMATE (ULTra-sIte for Measuring Atmosphere of Tokyo metropolitan Environment): a case study using the dual-polarization Doppler weather radars, Prog. Earth Planet. Sci., 9, 51, https://doi.org/10.1186/s40645-022-00511-5, 2022.

Stevens, B., Satoh, M., Auger, L., Biercamp, J., Bretherton, C. S., Chen, X., Düben, P., Judt, F., Khairoutdinov, M., Klocke, D., Kodama, C., Kornblueh, L., Lin, S.-J., Neumann, P., Putman, W. M., Röber, N., Shibuya, R., Vanniere, B., Vidale, P. L., Wedi, N., and Zhou, L.: DYAMOND: the DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains, Prog. Earth Planet. Sci., 6, 61, https://doi.org/10.1186/s40645-019-0304-z, 2019.

Stevens, B., Bony, S., Gross, S., Klocke, D., Windmiller, J. M., Wing, A. A., ... & Wu, Y.: Orcestra: Organized convection and earthcare studies over the tropical Atlantic. *Tellus*, *79*(1), 2026

https://doi.org/10.5194/egusphere-2026-3606
Preprint. Discussion started: 16 September 2026

Takasuka, D., Satoh, M., Miyakawa, T., Kodama, C., Klocke, D., Stevens, B., Vidale, P. L., and Terai, C. R.: A protocol and analysis of year-long simulations of global storm-resolving models and beyond, Prog. Earth Planet. Sci., 11, 66, https://doi.org/10.1186/s40645-024-00668-1, 2024.

Wan, H., Giorgetta, M. A., Zängl, G., Restelli, M., Majewski, D., Bonaventura, L., Fröhlich, K., Reinert, D., Rípodas, P., Kornblueh, L., and Förstner, J.: The ICON-1.2 hydrostatic atmospheric dynamical core on triangular grids – Part 1: Formulation and performance of the baseline version, Geosci. Model Dev., 6, 735–763, https://doi.org/10.5194/gmd-6-735-2013, 2013.

Wehr, T., Kubota, T., Tzeremes, G., Wallace, K., Nakatsuka, H., Ohno, Y., Koopman, R., Rusli, S., Kikuchi, M., and Eisinger, M.: The EarthCARE mission–science and system overview, Atmospheric Meas. Tech., 16, 3581–3608, 2023.