JGR Atmospheres

RESEARCH ARTICLE
10.1029/2022JD038134

Systematic Validation of Ensemble Cloud-Process Simulations Using Polarimetric Radar Observations and Simulator Over the NASA Wallops Flight Facility

Toshi Matsui[1,2], David B. Wolff[3], Stephen Lang[2,4], Karen Mohr[2], Minghua Zhang[5], Shaocheng Xie[6], Shuaiqi Tang[7], Stephen M. Saleeby[8], Derek J. Posselt[9], Scott A. Braun[2], Jiun-Dar Chern[1,2], Brenda Dolan[8], Jason L. Pippitt[3,4], and Adrian M. Loftus[2]

[1]Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA, [2]NASA Goddard Space Flight Center, Greenbelt, MD, USA, [3]NASA Goddard Space Flight Center, Wallops Flight Facility, Wallops Island, VA, USA, [4]Science Systems and Applications, Inc., Lanham, MD, USA, [5]Stony Brook University, Stony Brook, NY, USA, [6]Lawrence Livermore National Laboratory, Livermore, CA, USA, [7]Pacific Northwest National Laboratory, Richland, WA, USA, [8]Colorado State University, Fort Collins, CO, USA, [9]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA

Abstract The BiLateral Operational Storm-Scale Observation and Modeling (BLOSSOM) project was initiated to establish a long-term supersite to improve understanding of cloud physical states and processes as well as to support satellite and climate model programs over the Wallops Flight Facility site via a bilateral approach of storm-scale observations and process modeling. This study highlights a noble systematic validation framework of the BLOSSOM ensemble cloud-process simulations through mixed-phase, light-rain, and deep-convective precipitation cases. The framework consists of creating a domain-shifted ensemble of large-scale forcing data sets, and configuring and performing cloud-process simulations with three different bulk microphysics schemes. Validation uses NASA S-band dual-POLarimetric radar observations in the form of statistical composites and skill scores via a polarimetric radar simulator and newly developed CfRad Data tool (CfRAD). While the simulations capture the overall structures of the reflectivity composites, polarimetric signals are still poorly simulated, mainly due to a lack of representation of ice microphysics diversity in shapes, orientation distributions, and their complex mixtures. Despite the limitation, this new ensemble-based validation framework demonstrates that (a) no particular forcing or microphysics scheme outperforms the rest and (b) the skill scores of coarse- and fine-resolution ensemble simulations with different domain-shifted forcing and microphysics schemes are highly correlated with each other with no clear improvement. On the other hand, this suggests that coarse-resolution ensemble simulations are relevant for selecting the best meteorological forcing and microphysics scheme before conducting computationally demanding large eddy simulations in support of aircraft and satellite instrument development as well as cloud-precipitation-convection parameterizations.

Plain Language Summary Cloud-process simulations are state-of-art numerical simulations to represent realistic cloud and convection development through numerous microphysical processes. These simulations are largely affected by thermodynamic fields. This work generates slightly modified thermodynamic fields to create many different states of cloud process simulations, which are systematically validated by the weather radar available in NASA Wallops Flight Facility.

Key Points:

- Systematic validation of ensemble cloud-process simulations is established over the NASA Wallops Flight Facilities
- NASA polarimetric radar and simulators reveal strength and weakness of cloud-process simulations
- Coarse-resolution cloud-process simulations are sufficient to select best forcing and microphysics options

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

T. Matsui,
Toshihisa.Matsui-1@nasa.gov

Citation:

Matsui, T., Wolff, D. B., Lang, S., Mohr, K., Zhang, M., Xie, S., et al. (2023). Systematic validation of ensemble cloud-process simulations using polarimetric radar observations and simulator over the NASA Wallops Flight Facility. *Journal of Geophysical Research: Atmospheres*, *128*, e2022JD038134. https://doi.org/10.1029/2022JD038134

Received 5 NOV 2022
Accepted 11 AUG 2023

Author Contributions:

Conceptualization: Toshi Matsui
Data curation: Toshi Matsui, David B. Wolff, Jason L. Pippitt
Formal analysis: Toshi Matsui, Stephen Lang
Funding acquisition: Toshi Matsui, Karen Mohr, Jiun-Dar Chern
Investigation: Toshi Matsui, David B. Wolff, Stephen Lang

1. Introduction

Over the past several decades, cloud-resolving models (CRMs) have been advanced and utilized for many applications, such as the study of mesoscale convective systems (Fan et al., 2017; Li et al., 2010; Saleeby et al., 2016; Tao et al., 2013; Zeng et al., 2013), snowstorms (Iguchi et al., 2012a; Saleeby et al., 2013; Shi et al., 2010), cloud and precipitation microphysics (Chern et al., 2020; Iguchi et al., 2012b; Li et al., 2018; Tao et al., 2016), aerosol-cloud interactions (Iguchi et al., 2020; Loftus et al., 2016; Saleeby et al., 2015; Shi et al., 2014, 2021; Tao et al., 2007), cloud radiative forcing (Matsui et al., 2018; Zeng et al., 2011), and many other science applications (Tao et al., 2014). High-quality validated CRMs are currently used in the satellite programs of various space agencies as algorithm testbeds and instrument design (Grasso & Lindsey, 2011; Matsui et al., 2013) and for remote sensing algorithm databases (Kidd et al., 2016; Lang & Tao, 2018; Tao et al., 2019). Yet, one of the

Methodology: Toshi Matsui, David B. Wolff, Stephen Lang, Minghua Zhang, Shuaiqi Tang
Project Administration: Toshi Matsui
Resources: Toshi Matsui, David B. Wolff
Software: Toshi Matsui, David B. Wolff, Minghua Zhang, Shaocheng Xie, Shuaiqi Tang, Stephen M. Saleeby, Derek J. Posselt, Brenda Dolan
Supervision: Toshi Matsui, David B. Wolff, Karen Mohr, Minghua Zhang, Scott A. Braun
Validation: Toshi Matsui, David B. Wolff, Stephen Lang
Visualization: Toshi Matsui, David B. Wolff
Writing – original draft: Toshi Matsui, Stephen Lang, Karen Mohr
Writing – review & editing: Toshi Matsui, David B. Wolff, Minghua Zhang, Shaocheng Xie, Shuaiqi Tang, Stephen M. Saleeby, Derek J. Posselt, Scott A. Braun, Jiun-Dar Chern, Brenda Dolan, Adrian M. Loftus

long-standing uncertainties in CRMs is the spatial-temporal distribution of bulk hydrometeors and their properties, especially mixed-phase and frozen precipitation (Posselt, 2016; Posselt et al., 2019), often contrasted between continental and ocean conditions (Matsui et al., 2016, 2020).

With the development of detailed multi-instrument satellite simulators, multi-channel multi-instrument remotely sensed measurements can be used to provide a unique radiance-based evaluation framework anywhere there is earth-observing satellite coverage (Hashino et al., 2013; Masunaga et al., 2011; Matsui et al., 2009, 2014). However, these space-borne radar and microwave instruments are on low-altitude-orbiting satellite platforms, which are not useful for validating the detailed storm-scale evolution of convection and microphysics information due to limited temporal coverage. Ground-based radar instruments can provide detailed high temporal resolution. Particularly, scanning weather radar can capture important storm information from four-dimensional polarimetric and Doppler velocity signals (Bringi & Chandrasekar, 2001; Kumjian et al., 2022). Hence, in addition to global satellite measurements and simulators, polarimetric radar measurements and simulators complement the development of CRMs (Kollias & Tatarevic, 2017; Matsui et al., 2019; Ryzhkov et al., 2011; van Lier-Walqui et al., 2016).

The NASA S-band dual-POLarimetric radar (NPOL) and Dual-polarization Dual-frequency Doppler Radar (D3R) are NASA's signature ground-based polarimetric radars and have been deployed to many Global Precipitation Measurement (GPM) ground validation (GV) sites, providing vital storm-scale observations for CRM evaluation. The GPM GV program has conducted a series of field campaigns in mid- and high-latitude regions to study various types of precipitation processes and to better understand cloud microphysical properties and validate CRMs (Iguchi et al., 2012a, 2012b, 2014). Instrumentation in these campaigns consisted of a suite of ground-based dual-polarimetric multifrequency radar, rain and snow gauges, disdrometers, and aircraft-based microphysics measurements.

The Precipitation Research Facility (PRF) at Wallops Flight Facility (WFF) in Virginia, USA has been providing the instruments for these field campaigns (https://wallops-prf.gsfc.nasa.gov). Outside of the GV field campaigns, the PRF has maintained these instruments at the WFF site to provide ground precipitation measurements to validate measurements and retrievals from the radar and passive microwave instruments on the GPM core satellite in and around the WFF site. The Wallops PRF provides an excellent opportunity to utilize high-quality precipitation data not only for satellite validation but for physical process studies and model development.

To continue to improve and advance CRMs for satellite programs in unison with the steady progression to finer grid resolutions, a robust long-term evaluation is required that facilitates comprehensive comparisons between high-resolution, detailed cloud-process simulations and systematic routine storm-scale observations at high spatio-temporal scales.

To this end, BiLateral Operational Storm-Scale Observation and Modeling (BLOSSOM) has been initiated in order to establish a long-term super site at WFF that harnesses storm-scale measurements and modeling to support NASA satellite-aircraft missions and climate models. To achieve this goal, BLOSSOM aims to:

1. Provide meteorological large-scale forcing ensembles to drive cloud-process models to quantify errors and uncertainties,
2. Organize routine storm-scale measurements from ground-based polarimetric Doppler radars and in situ data to validate skills from cloud-process models and single-column models (SCMs), and
3. Generate a high-quality large-eddy simulation (LES) and observation database to support satellite and aircraft missions as well as new parameterization development for climate models via SCMs.

BLOSSOM was inspired by a similar DOE project: LES ARM Symbiotic Simulation and Observation (LASSO) over the Southern Great Plains (SGP) site (Gustafson et al., 2020). However, the ultimate objective of BLOSSOM is to support NASA modeling systems (https://map.nasa.gov/models/) as well as satellite programs. Particularly satellite support is one of the most important objectives for NASA Earth Science Division. Well validated high-quality cloud process simulations can feed to a detailed satellite instrumental simulator as input of atmosphere-surface states. In this way, realistic satellite observable signals can be simulated for current or future satellite missions (Matsui et al., 2013), because a detailed simulator can transform fine-resolution LES input to coarse-resolution satellite observables through calculating antenna gain function with respect to sensor beam widths (Pielke, 2013). Simulated satellite signals are further used to examine the remote sensing algorithms whether retrieval can accurately retrieve geophysical parameters or not in comparison with known process-model

input (Masunaga et al., 2011). This framework, often denoted as retrieval Observing System Simulation Experiment (Zeng et al., 2020), is becoming a more important framework in the future, because of the relatively cost-effective framework in comparison with a full suite of field campaigns.

This paper presents the initial case studies, model experiment description and validation methods of BLOSSOM. The science goal is to evaluate the skill (or performance) and uncertainties in the cloud-processes model using ground-based polarimetric radar and assess the impact of model resolution on skill to determine an ensemble framework. Section 2 details the PRF instruments, Section 3 explains the method for deriving large-scale forcing, Section 4 provides the CRM specifications, Section 5 gives the validation methods and results, Section 6 shows the skill scores and impact of model resolution, Section 7 examines sensitivity of resolution, and Section 8 provides conclusions and remarks.

2. Wallops Flight Facility and Case Studies

The WFF PRF observatory is situated on the coast of Virginia's Eastern Shore (Figure 1a). There are a wide range of storm types observed at WFF, including continental organized convection, sea-breeze-driven isolated convection, mid-latitude frontal systems, winter-time nor'easters, and tropical cyclones among others, offering a different ranges of events than the DOE ARM SGP site (Sisterson et al., 2016). This makes the WFF site a unique "supersite" for physical process studies and model development in addition to the existing DOE ARM SGP site. Core measurements from the Wallops PRF include a ground-based radar network, in situ precipitation measurements, and routine collection of GPM satellite measurements. Among these instruments, the most important for the BLOSSOM project are the ground-based polarimetric Doppler weather radars since they can provide the most comprehensive and critical storm-scale measurements of precipitating cloud systems at high temporal and spatial resolutions.

NPOL is NASA's flagship weather radar that provides high-resolution sampling of precipitation over the entire Delmarva Peninsula. NPOL is transportable and has been the backbone for validation efforts during several GPM field campaigns (e.g., Houze et al., 2017; Jensen et al., 2015; Wolff et al., 2015). NPOL is an S-band (10 cm) dual-polarization radar with an 8.5-m parabolic dish that provides a data resolution of 125 m radially and 0.95° azimuthally. Radar measurements include horizontally polarized reflectivity (Z_h), differential reflectivity (Z_{dr}), radial velocity (V_{rad}), spectrum width (W_{sp}), differential phase (F_{dp}), and horizontal-vertical correlation (ρ_{hv}). These measurements are used to calculate specific differential phase (K_{dp}), estimate precipitation rates, median drop diameters, and perform hydrometeor identification (Dolan & Rutledge, 2009; Dolan et al., 2013).

For proving BLOSSOM concepts, limited case studies are selected from the NPOL measurements to prove the experimental design and validation methods in the BLOSSOM project. The first case is a wintery mixed-phase event on 20 March 2015 (Figure 1a), in which hydrometeors are generally characterized as snow aggregates above 2 km in height and wet snow below (Figure 1b). The second case is a spring light rain event on 21 May 2015 (Figure 1c), in which wide-spread drizzle is present below a wet snow layer capped by a shallow layer of snow aggregates (Figure 1d). The last case is summer deep convection on 28 July 2017 (Figure 1e), in which echo top heights exceed 10 km and convective cores are characterized by graupel (Figure 1f). In these cases, precipitation is generally widespread in the domain during the peak times. These cases are diverse in terms of background meteorology and hydrometeor types and provide a robust opportunity to validate the different aspects of cloud process performance and the overall BLOSSOM framework, while widespread precipitation can be more readily captured by the large-scale advective forcing and idealized set of process simulations (discussed in following sections).

3. Large-Scale Forcing

There are two practical ways to conduct process-scale cloud models over the WFF region. The first approach is to use a regional mesoscale model with lateral boundary conditions (LBCs) driven by global forecasts or reanalysis (Pielke, 2013). The first approach can introduce realistic large-scale and heterogeneous mesoscale dynamics, such as orographic convection, sea-breeze circulation, and sharp frontal head as well as large-scale dynamics with given accurate initial and LBCs. However, even if initial and LBCs are accurate, timing and location of precipitation systems can be mis-represented due to physics biases and downscaling through LBCs (H. C. Davies, 1983).

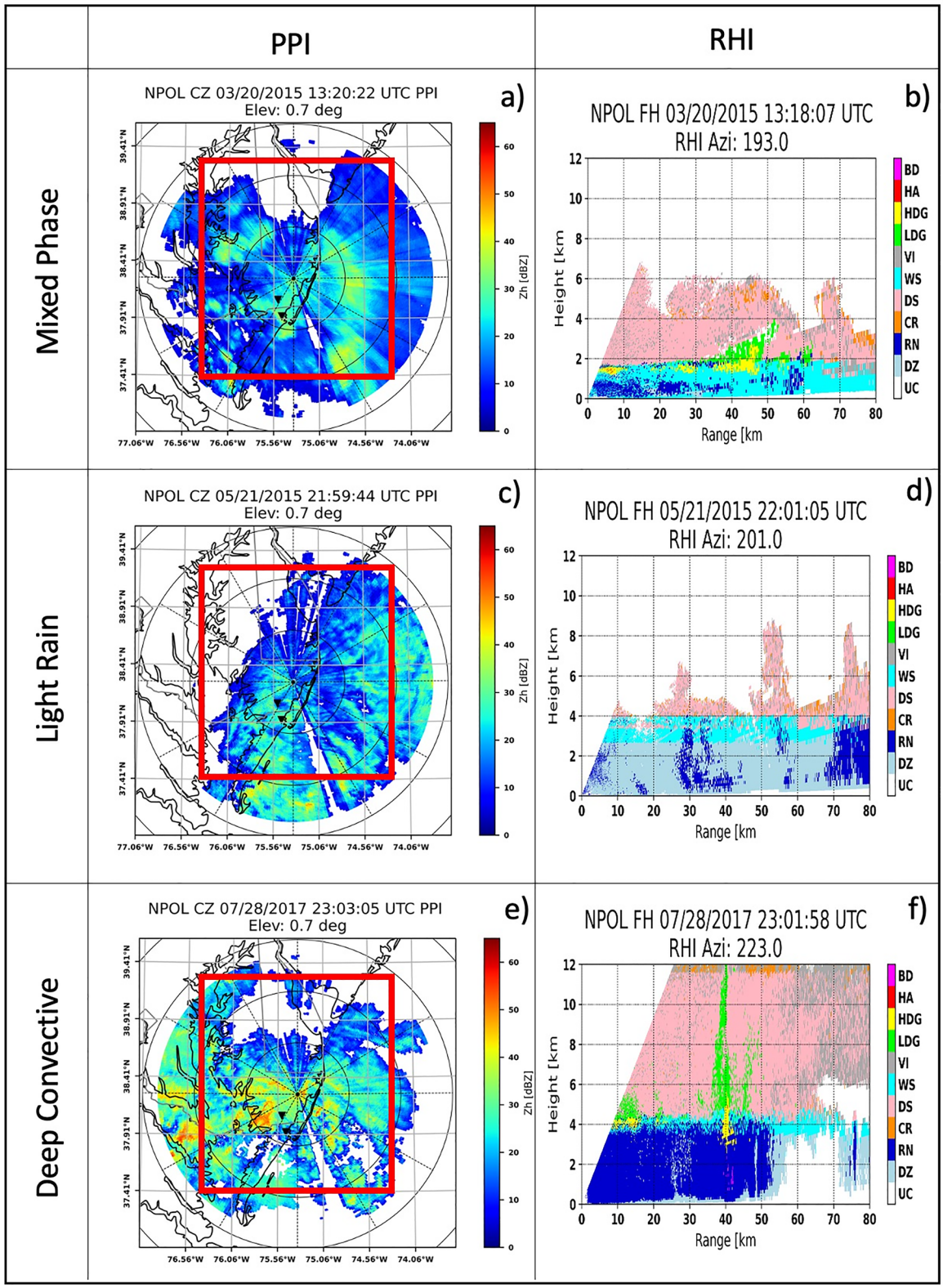

Figure 1. NASA S-band dual-POLarimetric radar images of three case studies: Plan Position Indicator radar reflectivity in the left column (a, c, and e) and Range Height Indicator HID images (BD: big drop, HA: hail, HDG: high-density graupel, LDG: low-density graupel, VI: vertical aligned ice, WS: wet snow, DS: snow aggregate, CR: ice crystal, RN: rain, DZ: drizzle, and UC: unclassified) in the right column (b, d, and f). Red rectangular in the left column represent "control" domain of large-scale meteorological forcing.

The second approach is to use an idealized CRM with periodic boundary conditions driven by uniform large-scale meteorological forcing (LSF) (e.g., Soong & Tao, 1980). This second approach, large-scale advective forcings are imposed to generate an ensemble of convection, often referred as cumulus ensemble (Tao & Simpson, 1993). This configuration causes unrealistic cloud morphologies across periodic boundary conditions. Thus the simulated results must be "statistically" evaluated against the observations (e.g., Waliser et al., 2002). Although this approach cannot incorporate realistic aforementioned mesoscale dynamics or waves, it has been proven working well for summer convection and even non-uniform cloud simulations of convective clouds (L. Davies et al., 2013; Fridlind et al., 2012; Ghan et al., 2000; Tang et al., 2022; Tao et al., 2014; Xie et al., 2002, 2005; Xu et al., 2002, 2005). These studies suggest that different model parameterizations create diverse statistics for convective storm simulations for a given large-scale forcing, but error sources from the large-scale forcing, how it is applied, and the grid configuration have yet to be examined. The major advantages is that this configuration can set up identical modeling set up between the process model and SCM for parameterization development (Randall et al., 1996).

BLOSSOM follows the uniform LSF approach because the same forcing can be applied to SCMs, allowing LES-SCM comparisons for the mutual development of physics parameterizations (Lopez-Gomez et al., 2020) and model intercomparison (Blossey et al., 2016; Xu et al., 2002), although flat and homogeneous terrain and forcing often lack important mesoscale-forcing and associated representation in precipitation systems. Thus, BLOSSOM simulations will focus on more wide-spread precipitation events rather than cases driven by strong mesoscale forcing.

The VARiational ANALysis (VARNAL) approach is a broadly accepted method for generating large-scale forcing wherein data are collected and adjusted based on the vertical integration of the atmospheric mass, moisture, dry static energy, and momentum budgets (M. Zhang & Lin, 1997; M. Zhang et al., 2001). Vertical profiles from a sounding network or reanalysis data can be used to calculate the tendency and advection terms of the budget equations, while the source/sink terms can be constrained by various sources of in situ and remote sensing observations. Xie et al. (2004) found that precipitation observations are the most important for constraining the moist and dry static energy budgets.

In a similar manner, BLOSSOM LSF is constructed from NASA's reanalysis: the Modern-Era Retrospective analysis for Research and Applications Version 2 (MERRA2, Rienecker et al., 2011). MultiRadar, MultiSensor system (MRMS) precipitation data (J. Zhang et al., 2016) are used as the precipitation constraint while other constraints include hourly atmospheric radiation budgets from the Clouds and the Earth's Radiant Energy System (CERES, Wielicki et al., 1996) SYNoptic product (SYN1deg, Rutan et al., 2015) and surface turbulent fluxes from MERRA2. These data have been compiled in three-hourly temporal resolution for the VARNAL approach. Then, the three-hourly large-scale forcing data are derived over the WFF $1.8 \times 2.2°$ LSF domain (latitude 37.4°–39.2°N, longitude 76.4°–74.2°W, Figure 1).

To characterize uncertainties in its model simulations, DOE's LASSO uses multiple forcing ensembles to achieve simulations close to the observations by changing the domain sizes of the reanalysis fields as well as the sources of the observation constraints (Gustafson et al., 2020). Similarly, BLOSSOM creates ensemble LSFs by shifting the MERRA2 domains (both atmosphere and surface fluxes) to the northwest, southwest, southeast, and northeast by 1° (denoted as EN1, EN2, EN3, and EN4 forcing, respectively) while the MRMS and CERES domains remain in the default control domains (Figure 1) This method accounts for uncertainties in the location and timing of precipitation systems and air mass in the reanalysis fields due to errors in data assimilation or biases in the MERRA2 system. In this way, five LSFs are generated through the VARNAL.

Figure 2 shows time series of precipitation rate, the apparent heat source (Q1) profile, and apparent moisture sink (Q2) profile (Yanai et al., 1973) from the control forcing (a, c, and e) as well as the variability (standard deviation) between five ensemble forcings (b, d, and f) for each case. Large heat sources and moisture sinks are associated with precipitation variability to different degrees in the three cases. Large ensemble variability also tends to be associated with large heat/moisture tendencies. Q1 tends to have larger variability above 400 mb levels, whereas Q2 tends to have large variability below 400 mb. This LSF domain-shifted ensemble variability strongly impacts the model macrophysics as well as microphysics validation as discussed in later sections.

4. Ensemble Simulations With the Cloud-Process Model

BLOSSOM employs the Goddard Cumulus Ensemble (GCE) as its default CRM. GCE is a cloud-process model developed and improved at NASA GSFC over several decades (Tao & Simpson, 1993; Tao et al., 2003, 2014). The GCE responds well to imposed LSF producing clouds and precipitation via the microphysics. Cyclic boundary conditions are used, so no additional heat, moisture or momentum enters the domain apart from that imposed

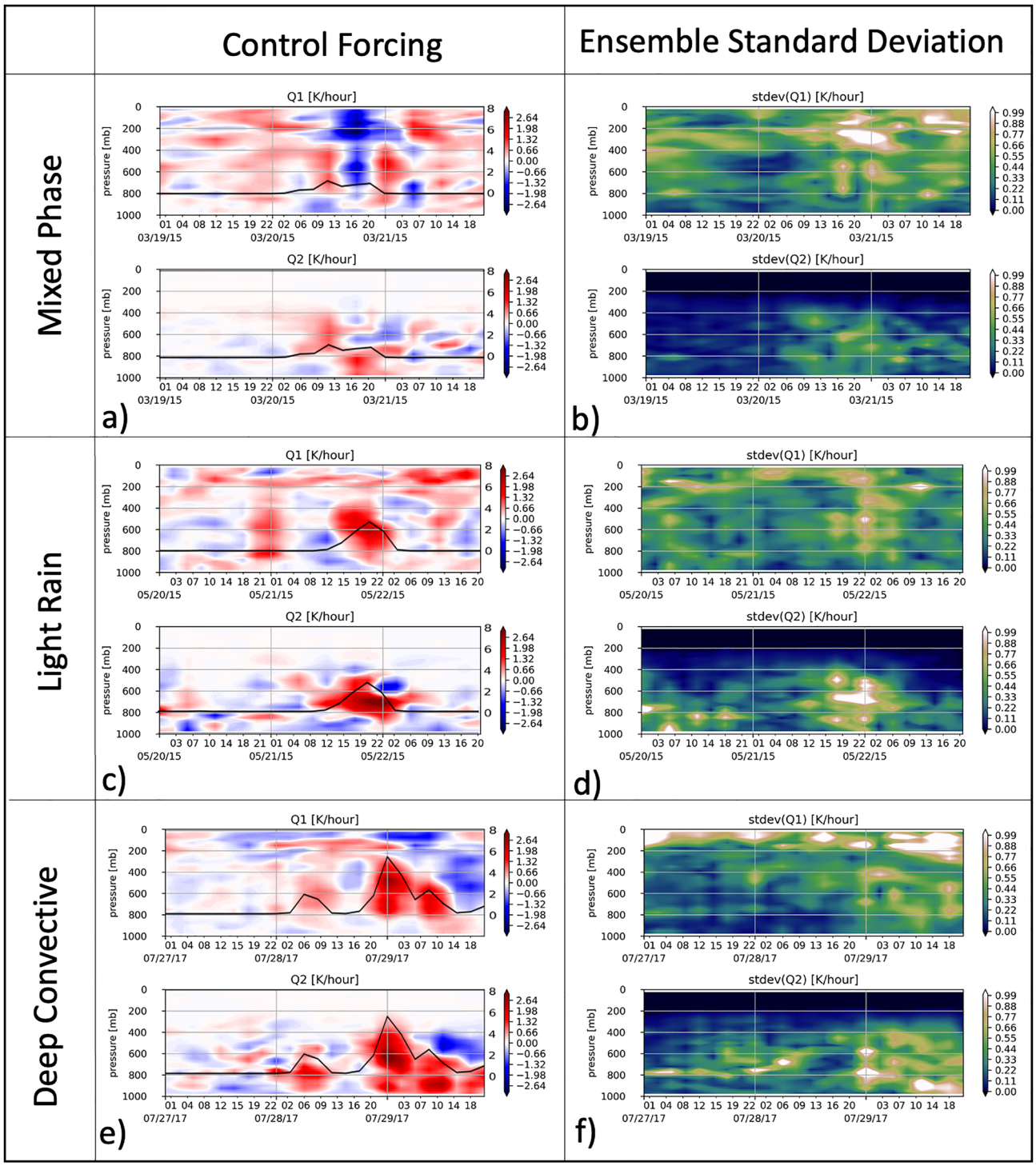

Figure 2. Time series of the apparent heat source (Q1) and moisture sink (Q2) (Yanai et al., 1973) profiles in three cases (a, c, and e) with surface precipitation (black line with green *y*-axis in [mm/hr]) and standard deviations among ensemble large-scale meteorological forcings (b, d, and f). Unit of moisture sink is converted to K/hr in energy point of view.

by the LSF or solar/infrared radiative processes. In addition, GCE's anelastic dynamic core option allows faster integration of finer-resolution runs (1.5–2 times) than its compressible dynamic core option. For this project, the GCE code was extensively upgraded to incorporate (a) sub-second time steps, (b) a stretched grid, (c) generalized forcing options, and (d) three bulk microphysics options.

Table 1
Comparison of Microphysics Scheme Attributes

	GMP	Morrison	RAMS
Cloud droplets	Q	Q, N	Q_{small}, Q_{large}, N_{small}, N_{large}
Rain	Q	Q, N	Q, N
Cloud ice	Q	Q, N	Q_{small}, Q_{large}, N_{small}, N_{large}
Snow aggregates	Q, s-m[a]	Q, N	Q, N
Graupel	Q, s-m[a]	Q, N	Q, N
Hail	Q, s-m[a]	Option for graupel or hail	Q, N
Ice density	Power law	Constant	Power law

Note. Q and N represent mass and number concentrations, respectively.
[a]The GMP scheme includes pre-set size-mappings (s-m) for snow, graupel, and hail that depend on mass and ambient air temperature.

BLOSSOM uses traditional 1-km horizontal grid spacing with 128 × 128 × 96 grids (*x*-*y*-*z* cartesian coordinates) because it requires a number of ensembles to examine the uncertainties of the model performance. Vertical grid spacings are ranged from roughly 44 to 250 m for low atmosphere (below 4 km level), 250 to 400 m for middle atmosphere (between 4 and 10 km level), and 400 to 600 m for upper troposphere (above 10 km level). However, additional finer resolution runs with 250-m grid spacing are conducted to examine the impact of model resolution upon skill scores of model performance. This domain size approximately corresponds to the grid size of climate model (1°).

Precipitating systems over WFF often contain solid- and mixed-phased microphysics processes, so there is a larger degree of uncertainty when simulating these types of systems versus shallow warm-phase clouds in the LASSO project over the ARM SGP site (Gustafson et al., 2020). Thus, BLOSSOM uses three different bulk microphysics schemes (Table 1). The Goddard bulk one-moment 6-class scheme (denoted as GMP hereafter) has four ice classes and uses preset size and density mapping for snow, graupel and hail (Lang et al., 2014; Tao et al., 2016). The Morrison two-moment 5-class scheme (denoted as MORR hereafter) predicts mass and number concentration and has been widely used in the Weather Research and Forecasting community (Morrison et al., 2005). The Regional Atmospheric Modeling System (RAMS) microphysics scheme is a 2-moment 8-class bin-model-emulating scheme (Meyers et al., 1997; Saleeby & Cotton, 2004; Saleeby & van den Heever, 2013; Walko et al., 1995) and is the most complex of the three schemes. Details are given in Table 1. Besides the microphysics scheme, all simulations use the 1.5-order turbulent kinetic energy scheme for turbulent mixing and the Goddard radiation scheme for radiative heating (Matsui et al., 2018).

For the control simulations (denoted as "CTL"), the horizontal and vertical advective forcings, mean wind speed as well as prescribed surface flux are used to drive the GCE. These fields are interpolated and imposed as tendency terms at every model time step. In addition, GCE allows the nudging of atmospheric temperature (T), humidity (QV), or temperature-and-humidity (TQV) profiles for stronger constraints. For the control-domain forcing only (Figure 1), these three additional nudging methods (T, QV, or TQV) are applied. Thus, four different nudging (CTL, T, QV, TQV) and four domain–shifted ensemble LSFs (EN1, EN2, EN3, and EN4 in Section 3) are used to drive the GCE. Thus, an ensemble of eight different simulations are generated for each case with each member run with three different bulk microphysics parameterizations (Table 1), resulting in 24 simulations for each case to characterize the model uncertainties related to forcing and microphysics. Hereafter, each specific run for the ensemble experiments is labeled as "microphysics tag (GMP, MORR, or RAMS)"-"forcing tag (CTL, CTL_T, CTL_QV, CTL_TQV, EN1, EN2, EN3, or EN4)."

5. Model Validation Framework Using Polarimetric Radar

5.1. Polarimetric Radar Simulator and Data Processing Tool

Ground-based polarimetric scanning radar can provide comprehensive 4D radar backscatter over a few hundred km^2 domain and characterize the microphysics in convective and stratiform precipitation (Bringi & Chandrasekar, 2001). Thus, the data is very useful for evaluating and improving microphysical characteristics

and parameterizations (Jung et al., 2010; Matsui et al., 2019; Putnam et al., 2014; Snyder et al., 2017). To bridge the gap between CRMs and polarimetric radar measurements, a POLArimetric Radar Retrieval and Instrument Simulator (POLARRIS) has been developed (Matsui et al., 2019). POLARRIS is composed of a forward model based on a T-matrix and Mueller matrix (Vivekanandan et al., 1991) to generate polarimetric radar observables (Z_h, Z_{dr}, K_{dp}, ρ_{hv}) and radial velocity (V_{rad}). POLARRIS accounts for microphysics specifications (phase, size, and densities) from the bulk microphysics assumptions with hydrometeor shape and orientation distributions from various observations (Matsui et al., 2019). All assumptions and equations of particle axis ratio and orientation angle distributions are described in Table 1 of Matsui et al. (2019).

In contrast to generating output on the model grid (Matsui et al., 2019), the latest version of POLARRIS can re-sample model-grid fields to more realistic radial coordinates consistent with the radar scanning volume. This allows the simulated polarimetric radar signal to be output in the more common CfRadial format (Dixon & Lee, 2016) consistent with the NPOL CfRadial output. In this project, GCE 3D fields are output every 10 min with POLARRIS forward simulations output at the same frequency, which is close to the current NPOL scanning cycle (~14 min) with dense elevation angles. In the future, POLARRIS will be implemented inline within the GCE for more efficient radar sampling versus outputting large numbers of CRM files (Oue et al., 2019).

POLARRIS and NPOL CfRadial data sets are then processed in the CfRadial Data tool (CfRaD), a newly developed Fortran-based program that can calculate various statistical composites as well as convert data coordinates with parallel processing capability. First, CfRaD converts radial coordinates into latitude-longitude coordinates via a matrix rotation method for given radar latitude-longitude locations, elevation angle and range distance. Radar beam width and range resolution are also considered. Second, CfRaD computes composite reflectivity by extracting the column maximum Z_h value. The composite reflectivity values are derived only in the lowest 3 km to avoid including anvil types of clouds in the convective-stratiform (C-S) separation. Z_{dr} and K_{dp} values coincident with the composite reflectivity are also stored for C-S separation. Third, CfRaD calculates statistical composites, including the horizontal fractional coverage of composite reflectivity (denoted as *RadarFrac*), time series of domain-averaged polarimetric radar profiles (denoted as *RadarProf*), and contoured frequency with altitude diagrams (CFADs) of polarimetric radar variables (denoted as *RadarCFAD*).

Note that an experimental C-S separation method is being applied in the BLOSSOM project using the product of Z_{dr} and K_{dp} values. The physics implication is that convective cores tend to generate large-sized, voluminous oblate raindrops as a result of melting aggregates, graupel and hail, and consequently tend to have larger values of Z_{dr} and K_{dp} in both continental and maritime convective systems (e.g., Matsui et al., 2020). Instead of using either value alone, their product physically represents a combination of a large size (Z_{dr}) and volume (K_{dp}) of oblate raindrops. Currently, a threshold of 2.0 is used to identify strong convective cores after examining several midlatitude convective systems over the WFF after some initial investigation. Associated Z_h values in the convective columns derived from this method tend to be very large (shown later in this section), and has been demonstrated well over Houston area during the Tracking Aerosol Convection Interactions Experiment field campaign (https://portal.nccs.nasa.gov/datashare/tracer/).

The approach is similar in concept to the raindrop microphysics-based separation technique in Thurai et al. (2021), but our experimental method uses polarimetric radar signals more directly. Although there are a number of different C-S separation methods with their advantages and disadvantages (Lang et al., 2003), the key aspect is to apply an identical C-S separation method to models and observations using an instrument simulator (Matsui et al., 2019). This experimental method tends to select a narrower "core" of convective portion in comparison with the traditional Z_h-based C-S separation (Steiner et al., 1995).

5.2. Model Validation Using Statistical Composites

RadarFrac composites are time series of the areal fraction of composite radar reflectivity with 0, 10, 20, 30, 40, and 50 dBZ thresholds (Figure 3). These composites capture the time evolution of macroscopic storm structures of intensity and extent. The 0–10 dBZ threshold captures the entire storm (or cloud) coverage, 10–30 dBZ thresholds capture light precipitation coverage, 30–40 dBZ thresholds capture moderate precipitation coverage, and 40–50 dBZ thresholds capture heavy to intense precipitation coverage. Figure 3 shows observed *RadarFrac* composites from the mixed-phase, light-rain, and deep-convective cases. Radar fractions are missing from 16 to 21 Z in the light rain case due to missing Plan Position Indicator (PPI) scans. Otherwise, *RadarFrac* shows time transitions of storm coverage for the different case studies. The deep-convective case has a small area of

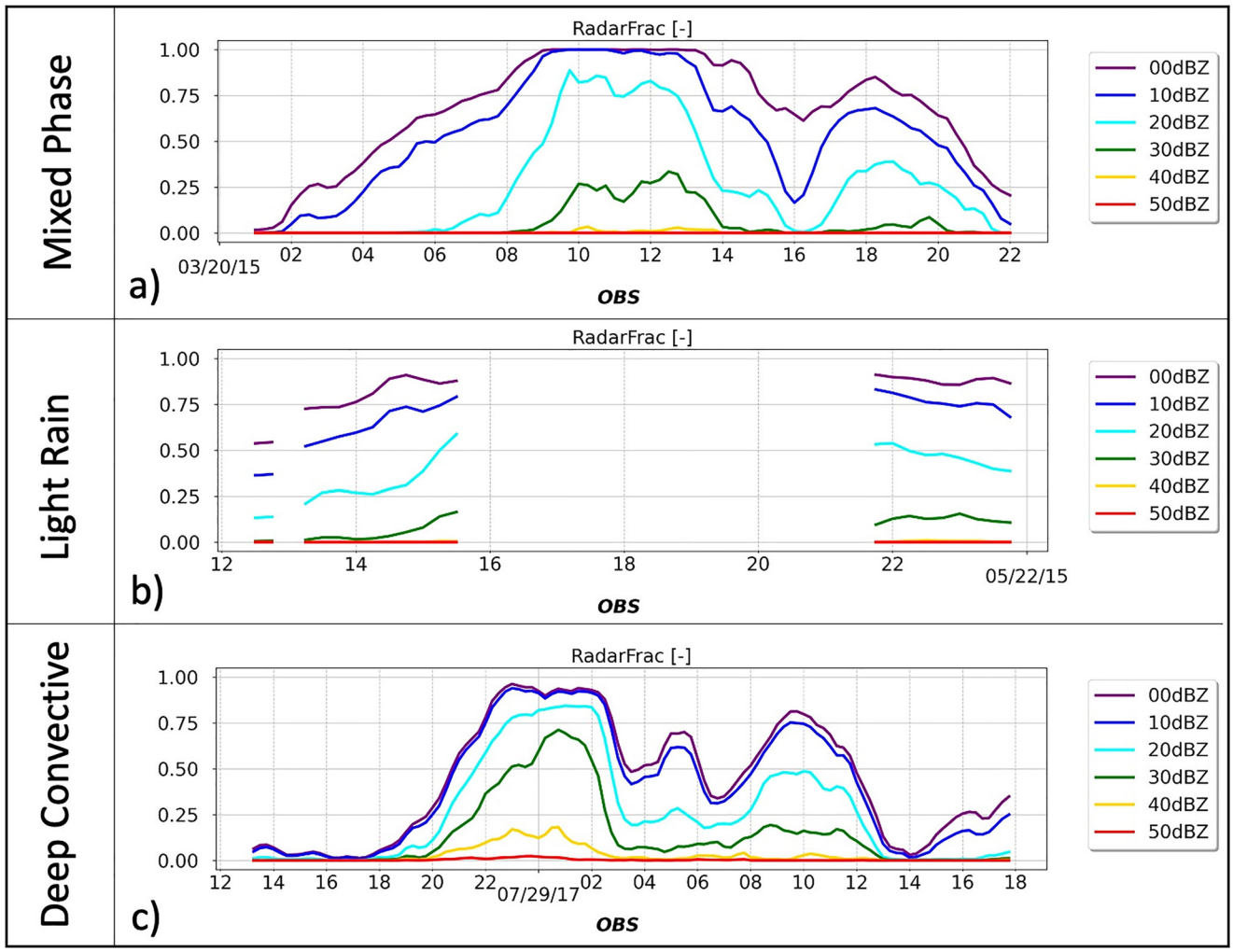

Figure 3. Time series of areal fractions of composite radar reflectivity with 0, 10, 20, 30, 40, and 50 dBZ thresholds for three cases. The light-rain case have gaps in radar Plan Position Indicator scan between 16 and 21 Z.

40–50 dBZ values suggesting the presence of convective cores (Figure 3c), whereas the mixed-phase and light-rain cases only exceed 30 dBZ (Figures 3a and 3b). All three cases have 0 dBZ coverages that exceed 0.8, suggesting the precipitation systems are widespread at their peak in the BLOSSOM domain.

Figure 4 shows *RadarFrac* composites of best skill scores (discussed later) derived from the GCE-POLARRIS simulations. For efficiency, only the best-skilled (discussed in next section) composites are shown in the model figure (applying to Figures 4, 6 and 8). The mixed-phase case is represented by the MORR-EN2 run (Figure 4a). Radar fractions for the 0, 10, and 20 dBZ thresholds show a more rapid increase up to the entire domain around 06 Z, 20 March than the decrease between 18 and 22 Z. 30–40 dBZ fractions are also overestimated versus the observed. The light-rain case is represented by the GMP-EN4 run. This simulation shows reasonable radar fractions of 0–40 dBZ thresholds in comparison with the observations, although the model exceeds 40 dBZ (Figure 4b). The deep-convective case is represented by the RAMS-CTL_T run. Although the temporal evolution is not well captured, the spectrum of radar fractions from 0 to 40 dBZ is realistically simulated (Figure 4c). Again, the radar fractions tend to increase rapidly at the beginning of the storm versus the gradual increase in the observed composites. This is due to the limitation of using horizontally homogeneous LSF with cyclic boundary conditions. The forcing tends to generate clouds simultaneously over the entire domain. These errors/biases could not be fixed unless regional weather simulations are employed with accurate initial and LBCs.

RadarProf composites show a time series of horizontal domain-averaged polarimetric radar profiles from Range Height Indicator (RHI) scans (Figure 5). Unlike quasi-vertical profiles (Ryzhkov et al., 2016), this composite

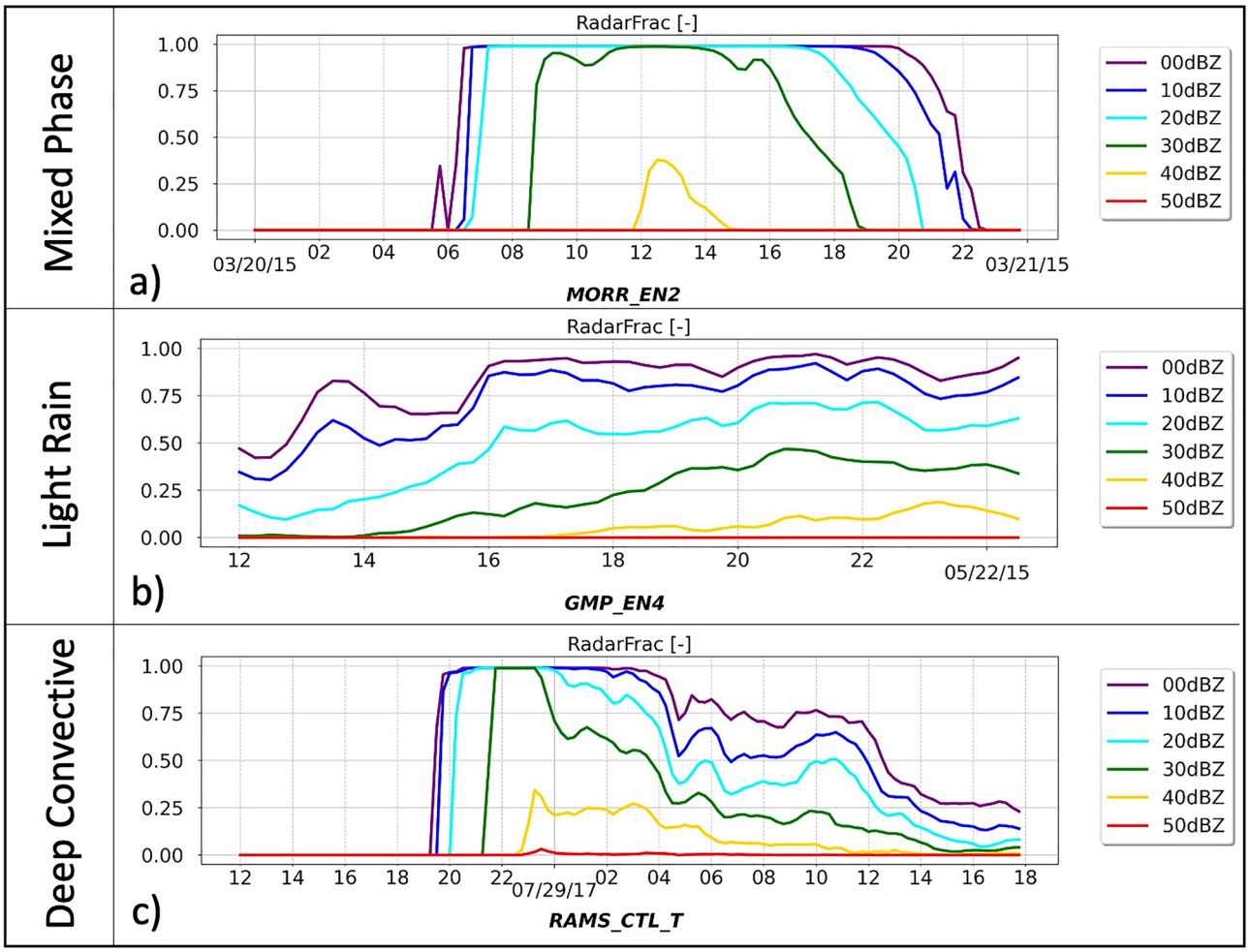

Figure 4. Same as Figure 3, but represented by the best-skill simulations in terms of large-scale meteorological forcing and microphysics scheme.

collects radar observations from all elevation angles to create averaged profiles. For these case studies, NPOL PPI scan is limited to very low elevation angles, thus measurements from RHI scan are used in this analysis. Z_h approximates the sixth moment of particle sizes and so emphasizes the variability of larger particles in the sampling volume. Z_{dr} is a measure of the size-weighted (sixth moment) oblateness of particles in the Rayleigh scattering regime and is also sensitive to particle phase (liquid vs. ice); thus, larger values are often associated with large oblate raindrops in convective cores or horizontally aligned ice crystals above the melting layer. K_{dp} is sensitive to the volume-weighted (third moment) oblateness of particles. ρ_{hv} is sensitive to the diversity in particle shapes and phases in a pulse volume, such as melting layers.

Figure 5a shows that the mixed-phase case depicts low to moderate Z_{dr} values (up to 0.9) and relatively low ρ_{hv} values (down to 0.9) near the surface during the 06–14 Z period. However, Z_{dr} values become nearly zero and ρ_{hv} values become close to unity after 16 Z. The 19–21 Z period shows narrow melting signals at the 2.5 km level characterized with very large Z_h and Z_{dr} values in conjunction with low ρ_{hv} values. This physically represents the transition from mixed-phase precipitation to light rain during this last time period. The light rain case is similar but shows a smoother transition from a mixed phase in the beginning to signals dominated by light rain (Figure 5b). In the deep-convective case, *RadarProf* composites are separately analyzed for convective and stratiform columns. In the convective columns, Z_h, Z_{dr}, and K_{dp} values become very large below 5 km, indicating large, voluminous, oblate raindrops (Figure 5c). However, strong melting signals (large Z_h and small ρ_{hv}) are also captured from 22 to 01 Z, suggesting either possible misclassification in the current experimental C-S separation method or mismatches between the surface and mid-level microphysics along rain shafts. Stratiform columns show a more extensive period of melting layer from 20 Z 28 July to 12 Z 29 July and smaller Z_h, Z_{dr}, and K_{dp}.

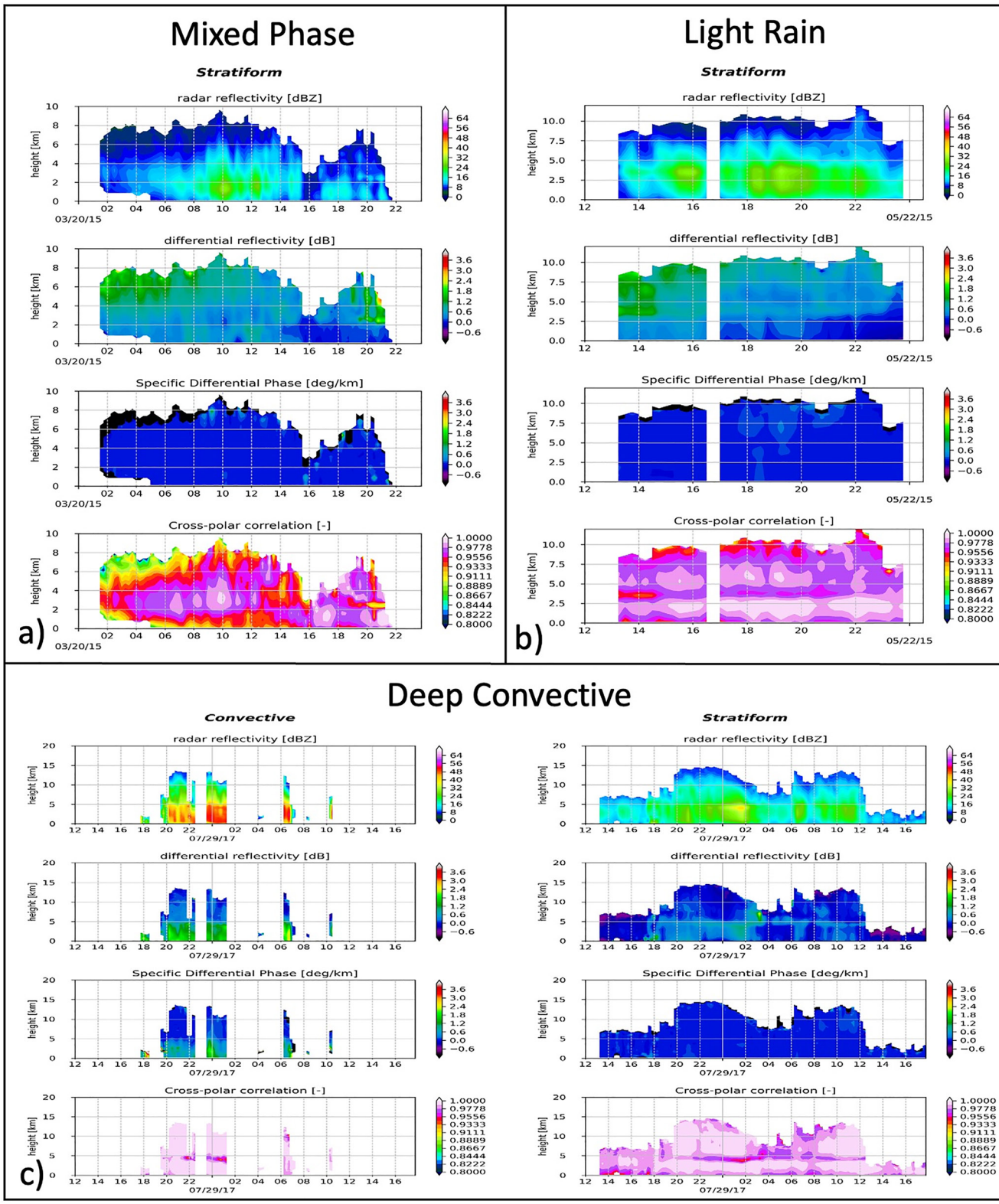

Figure 5. Time series of domain-mean vertical profiles of Z_h, Z_{dr}, K_{dp}, and ρ_{hv}. The mixed-phase and light-rain case shows stratiform columns only, since convective columns have very small sampling numbers, while the deep-convective cases shows both convective and stratiform columns. Note that *y*-axis is scaled differently for each case.

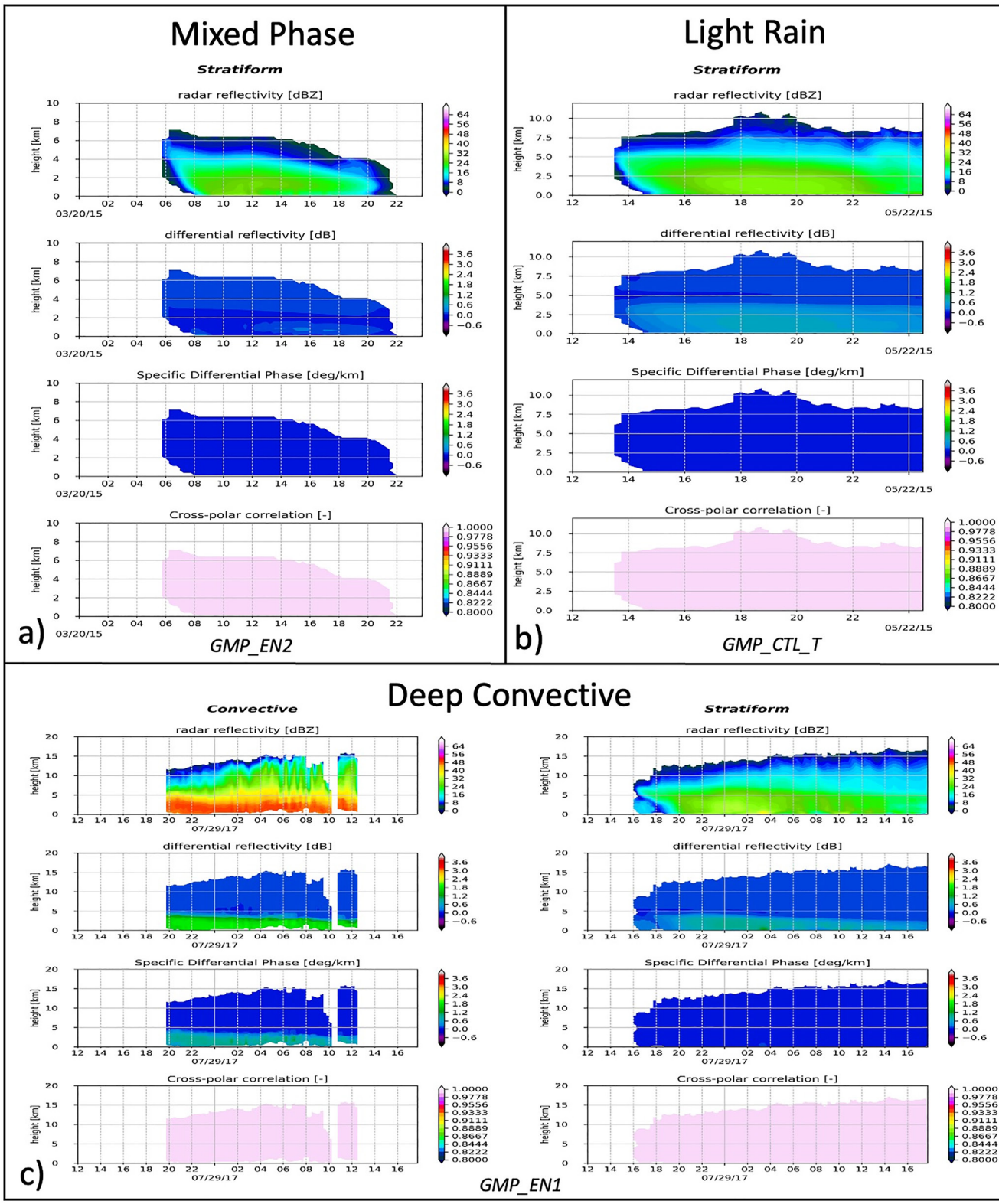

Figure 6. Same as Figure 5, but represented by the best-skill simulations in terms of large-scale meteorological forcing and microphysics scheme.

Figure 6 shows *RadarProf* composites of best skill scores (shown later) derived from the GCE-POLARRIS simulations. The mixed-phase case is represented by the GMP-EN2 run (Figure 6a). Compared to the observations (Figure 5a), the model *RadarProf* composites are more homogeneous versus time, especially Z_{dr} and ρ_{hv}. Upper-level Z_{dr} is near zero (~2 dB is observed), indicating a lack of horizontally aligned ice crystals in the model. Near-zero values of ρ_{hv} also indicate the simulated hydrometeors lack diversity in particle shapes and phases. Similar results are shown in the light rain case, represented by the GMP-CTL run (Figure 6b). In the deep convective case, represented by the GMP-EN1 run (Figure 6c), Z_h, Z_{dr}, and K_{dp} values show consistent large values in the convective regime indicating large raindrops in size and volume. Again, ρ_{hv} is very homogeneous due to a lack of particle diversity in the model. Though not shown, the RAMS microphysics has slightly more variability in ρ_{hv} because of its additional hydrometeors types (Table 1). Repeating the discussions in Matsui et al. (2018), most of the bulk microphysics (including three schemes used in this study) do not provide information on orientation of particles and complex structure of ice particles. Thus, this information must be assigned in the POLARRIS, which results in more homogeneous signals than the observed signals.

RadarCFAD composites show time-integrated profiles of domain-wide histograms (Figure 7). Similar to RadarProf, mixed-phase and light rain cases only show stratiform components. In contrast to *RadarProf*, *RadarCFAD* provides more details on the polarimetric radar variable distributions. Z_h in the mixed-phase case depicts mode values around 10 dBZ below 5 km; however, Z_h reaches nearly 50 dBZ at 2.5 km in association with melting particles (Figure 7a). Most of the Z_{dr} values (color shading) range between 0 and 2 above 5 km in association with weak Z_h values (0–10 dBZ); the mode shifts toward zero near the surface, while K_{dp} is concentrated around zero value throughout the profile. ρ_{hv} departs from unity above 5 km (Figure 7a). These polarimetric signals imply that the precipitation system is dominated by a mixture of ice crystal habits, especially aloft, including horizontally aligned oblate crystals. The light rain case shows similar trends in polarimetric signals, but storm heights and Z_h are slightly higher, and there are more pronounced melting signals in Z_h, Z_{dr}, and ρ_{hv} at 4 km (Figure 7b).

The deep-convective case highlights the contrast between the convective and stratiform CFADs (Figure 7c). Because of the Z_{dr}-K_{dp}-based C-S separation method, most of Z_{dr} and K_{dp} values range from 1 to 2 dB and from 0 to 1.5 dB, respectively in the convective CFAD below 5 km, while Z_{dr} and K_{dp} are distributed around zero above 5 km. These tendencies differ from the mixed-phase and light rain cases and indicate that ice crystals are more randomly oriented or shaped in the deep convective case. Z_h values dramatically increase from storm-top toward the surface, which is typical of convective core CFADs associated with graupel, hail, and large raindrops (e.g., Matsui et al., 2020), though a small melting signal is present in ρ_{hv} near 5 km. Stratiform Z_h distributions, in contrast, show smaller values with a sharp melting-layer signal at 5 km, associated with peaks in Z_{dr} and ρ_{hv}. Overall, this simple Z_{dr}-K_{dp}-based C-S separation method distinguishes convective and stratiform columns reasonably with well-known C-S features (Houze, 1997; Steiner et al., 1995). More detailed comparison of C-S separation methods will be studied upon the long-term BLOSSOM database in the future manuscript. Note that the quick decline in Z_h values near the surface is most likely due to NPOL's beam blockage at the lowest beam angle.

Figure 8 shows *RadarCFAD* composites of best model skill score derived from the GCE-POLARRIS simulations. The mixed-phase case is represented by the RAMS-EN3 run (Figure 8a). Clearly, the Z_h, Z_{dr}, and ρ_{hv} values have much less variability than the observations, while Z_h tends to be overestimated by the simulation. The light rain case has very similar results despite the different microphysics and forcing options (GMP-EN2, Figure 8b). As noted earlier, the RAMS microphysics has slightly more variable ρ_{hv} values. The deep convective case, represented by the RAMS-CTL_T run, has closer agreement in the mode values of all radar observables in both the convective and stratiform regimes (Figure 8c), though the variability is very limited in model distributions. Matsui et al. (2019, 2020) noted that some of these model biases stem from the microphysics parameterizations but some are from uncertainties in the assumed particle axis ratio and orientation distributions in the forward simulations and will be improved in future studies. More detailed simulated radar composites are available in Supporting Information S1.

6. Skill Score of Ensemble Simulations

While detailed manual inspections are necessary to improve the specific physical processes and assumptions in the forward model, a more automated, systematic validation is needed using these polarimetric radar composites given the number of meteorological forcing and microphysics options available. For this, a validation skill score

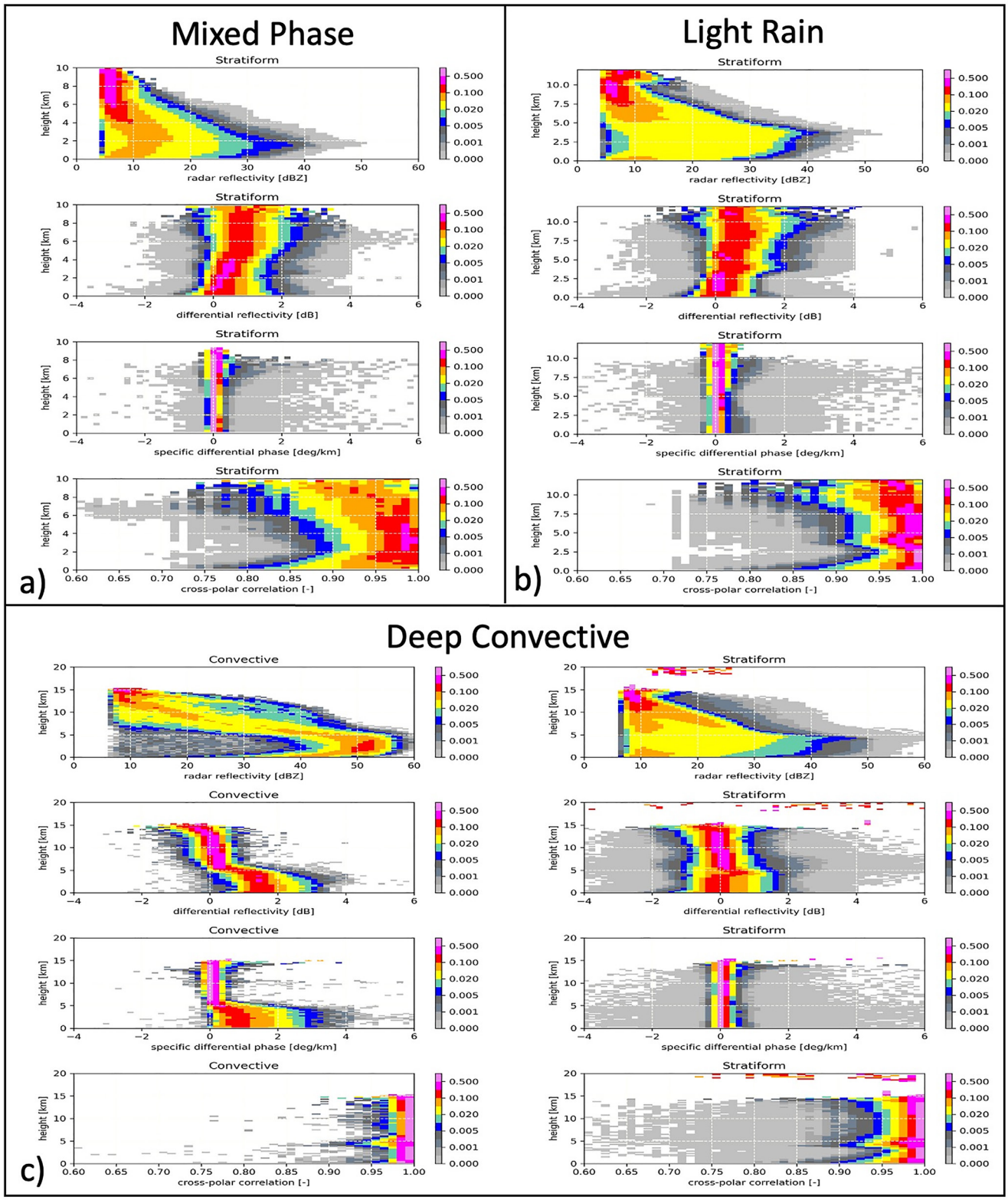

Figure 7. Radar contoured frequency with altitude diagrams of Z_h, Z_{dr}, K_{dp}, and ρ_{hv} integrated over the BiLateral Operational Storm-Scale Observation and Modeling domain. The mixed-phase and light-rain case shows stratiform columns only, since convective columns have very small sampling numbers, while the deep-convective cases shows both convective and stratiform columns.

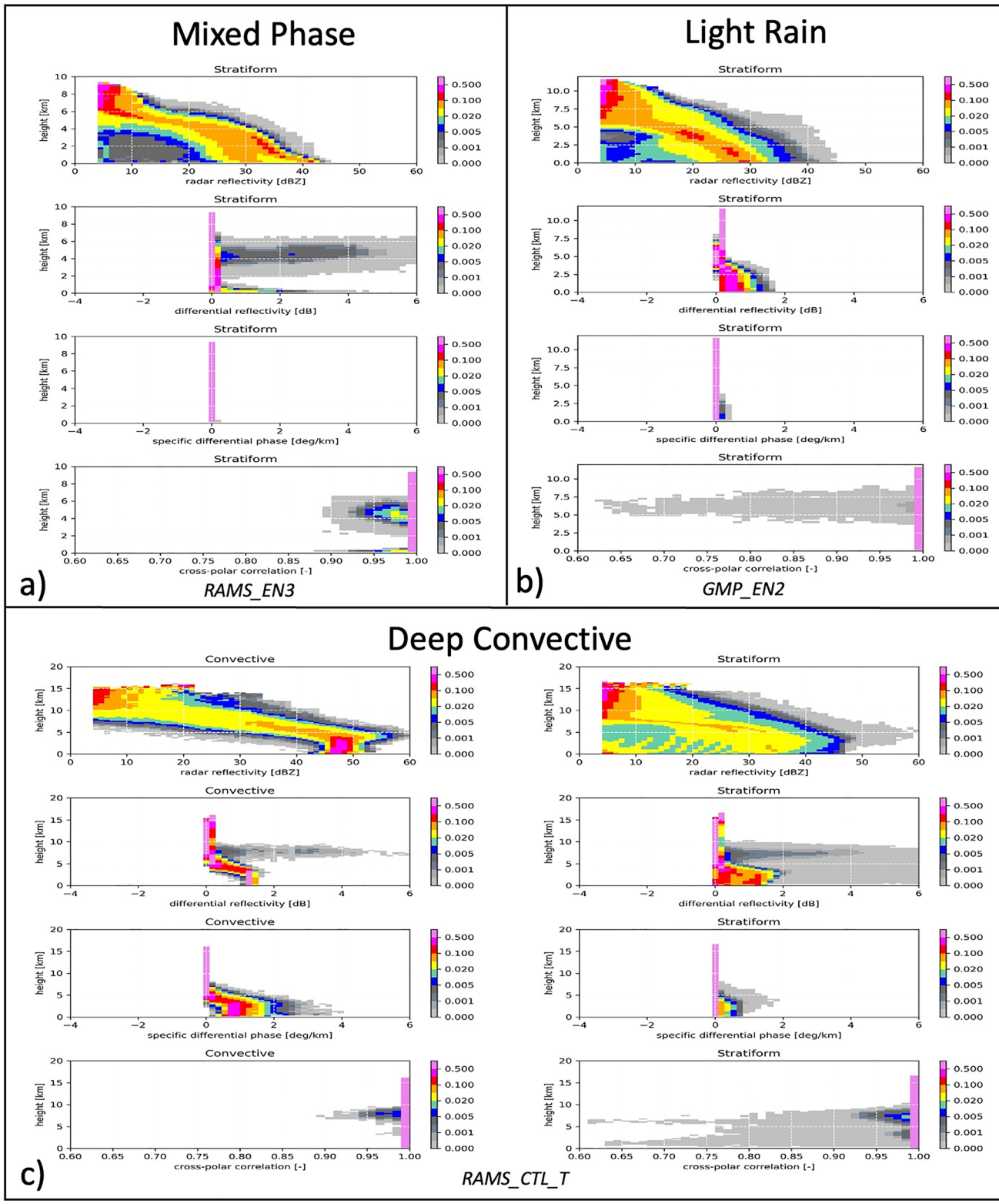

Figure 8. Same as Figure 7, but represented by the best simulations in terms of large-scale meteorological forcing and microphysics scheme.

is established via a combination of simple Spearman correlations (Corr) and root-mean square errors between the observed and simulated radar composites. RMSE is further normalized to a range of 0 to unity, hereafter denoted as normalized root-mean square error (NRMSE), equalizing the range of Corr.

For each radar composite, Corr and NRMSE are averaged with various weights (w) to derive total $\mathrm{NMRSE}_{\mathrm{tot}}$ and $\mathrm{Corr}_{\mathrm{tot}}$:

$$\mathrm{NRMSE}_{\mathrm{tot}} = \frac{\sum_p w_p * \mathrm{NRMSE}_p}{\sum_p w_p}$$

$$\mathrm{Corr}_{\mathrm{tot}} = \frac{\sum_p w_p * \mathrm{Corr}_p}{\sum_p w_p},$$

where p represents different parameters: 00, 10, 20, 30, 40, and 50 dBZ for *RadarFrac*; Z_h, Z_{dr}, K_{dp}, and ρ_{hv} for *RadarProf* and *RadarCFAD*. The weights are all unity for *RadarFrac*, meaning NRMSE and Corr from different p are equally weighted, whereas the weights for *RadarProf* and *RadarCFAD* are not. Instead, the weights for Z_h, Z_{dr}, K_{dp}, and ρ_{hv} are 30:5:5:2, respectively, much more heavily weighted toward Z_h than the polarimetric parameters (Z_{dr}, K_{dp}, and ρ_{hv}). This heterogeneous weighting is due to uncertainties in the assumed particle shapes and orientations rather than the explicit calculations for Z_{dr}, K_{dp}, and ρ_{hv} (Matsui et al., 2019). While the orientation and oblateness of raindrops are well known with high confidence (e.g., Brandes et al., 2002), those for solid particles are often characterized by unique shapes and orientations depending on environmental conditions (Schrom & Kumjian, 2018). Also, the weighting ratio between convective and stratiform is 1:9 due to the much larger horizontal coverage of stratiform precipitation (e.g., Houze, 1997; Steiner et al., 1995). These weights can be adjusted depending on the target physics validation and improvement.

Then, a skill score (SCORE), defined as the Euclidean distance scaled to 100 for a perfect point (zero $\mathrm{NRMSE}_{\mathrm{tot}}$ and $\mathrm{Corr}_{\mathrm{tot}}$ of unity), is calculated from the following equation:

$$\mathrm{SCORE} = 100 \times \frac{\sqrt{(1 - \mathrm{Corr}_{\mathrm{tot}})^2 - \mathrm{NRMSE}_{\mathrm{tot}}{}^2}}{\sqrt{D_{\max}}},$$

where $D_{\max}$ is maximum Euclidean distances, which are two for RadarFrac and RadarProf, while five for RadarCFAD, because slight shifts of CFADs often create negative correlation (i.e., larger Euclidean distance). Figure 9 shows *SCORE* diagrams from *RadarFrac* validation. SCOREs are randomly distributed in general. The best skills for the mixed-phase, light rain, and deep convective cases are the MORR-EN2, GMP-EN4, and RAMS-CTL_T runs, respectively. The worst skills for these cases are the MORR-EN1, RAMS-CTL_TQV, MORR-EN1 runs in the same order. None of the forcing nor microphysics show consistent superior skill. Detailed statistical distributions of *RadarFrac* SCORE are summarized in the boxplots later. SCORE variability is particularly large for MORR microphysics in the mixed-phase case. Beside these anomalies, SCORE variabilities appear to be similar to each other. Since *RadarFrac* validates the domain precipitation intensity spectrum, the best (worst) SCORE indicates the run that performed best (worst) at capturing the macroscopic structure of precipitation systems.

Figure 10 shows skill diagrams from *RadarProf*. The best skills for the mixed-phase, light rain, and deep convective cases are the GMP-EN2, GMP-CTL_T, and GMP-EN1 runs, respectively. The worst skills for these cases are the MORR-EN1, RAMS-CTL_QV, and RAMS-EN4 in the same order. The GMP microphysics tends to perform the best in this composite for all three cases. Detailed statistical distributions in the boxplot (Figure 12, middle columns) also confirm that the GMP microphysics has a higher skill than RAMS and MORR schemes. RAMS tends to have lower skill than the GMP and MORR schemes with respect to the median and interquartile values.

Figure 11 shows skill diagrams from *RadarCFAD*. The best skills for the mixed-phase, light rain, and deep convective cases are the RAMS-EN3, GMP-EN2, and RAMS CTL_T runs, respectively. The worst skills for these cases are the GMP EN1, MORR CTL_QV, and MORR CTL_T runs in the same order. Detailed statistical distributions in the boxplots (Figure 12) show that the RAMS scheme has slightly better *RadarCFAD* skill than the GMP and MORR schemes in the mixed-phase and deep convective cases, but the GMP scheme is better in the light rain case. Compared to the *RadarFrac* and *RadarProf* composites, *SCORE* variability from *RadarCFAD* tends to be small in terms of interquartile values as well as min-max ranges. This suggests that *RadarCFAD* composites are less affected by variability in the forcing ensemble than the *RadarFrac* and *RadarProf* composites.

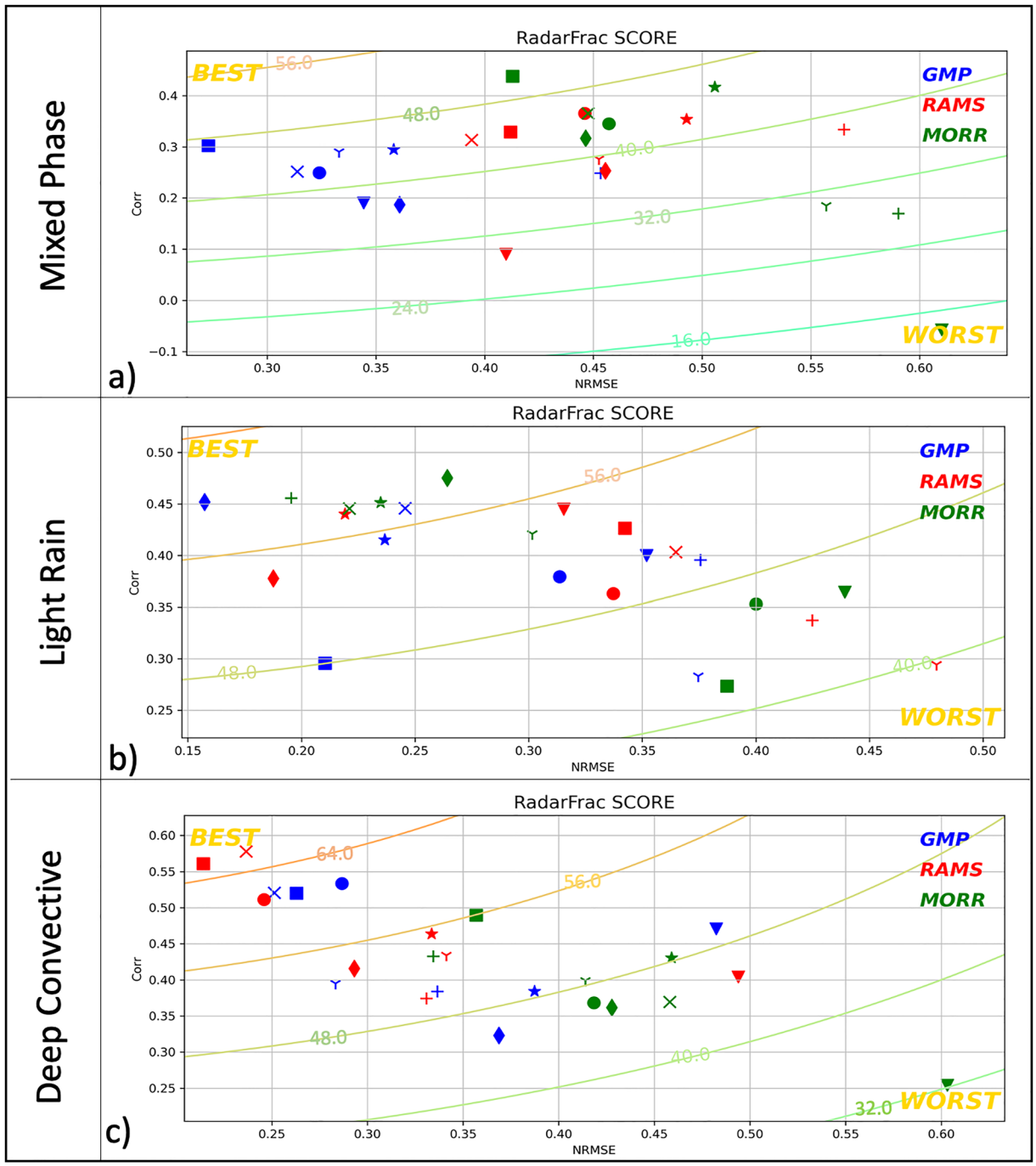

Figure 9. RadarFrac skill score diagrams as a function of Spearman correlation (Cor) and normalized root-mean square error. Blue, red, and green color represent GMP, Regional Atmospheric Modeling System, and MORR microphysics scheme. Different symbols represent types of forcing (CTL(●), CTL_T(×), CTL_QV(+), CTL_TQV(Y), EN1(▼), EN2(■), EN3(★), and EN4(◆)). Contour lines represent skill score.

Overall, none of meteorological forcing and microphysics options consistently outperformed the others in the three case studies. This suggests the design of the LSF ensemble randomly spreads the model skill and so better characterizes uncertainties associated with the construction of the LSF rather than relying on a single member. Also, the advantages of specific microphysics schemes are limited to specific precipitation types. Thus, microphysics improvement must be conducted across a variety of case studies rather than relying on a single case study.

SCOREs appear to be uncorrelated between *RadarFrac*, *RadarProf*, and *RadarCFAD* composites, because SCOREs between different composites do not show any significant correlation coefficients with p-values less

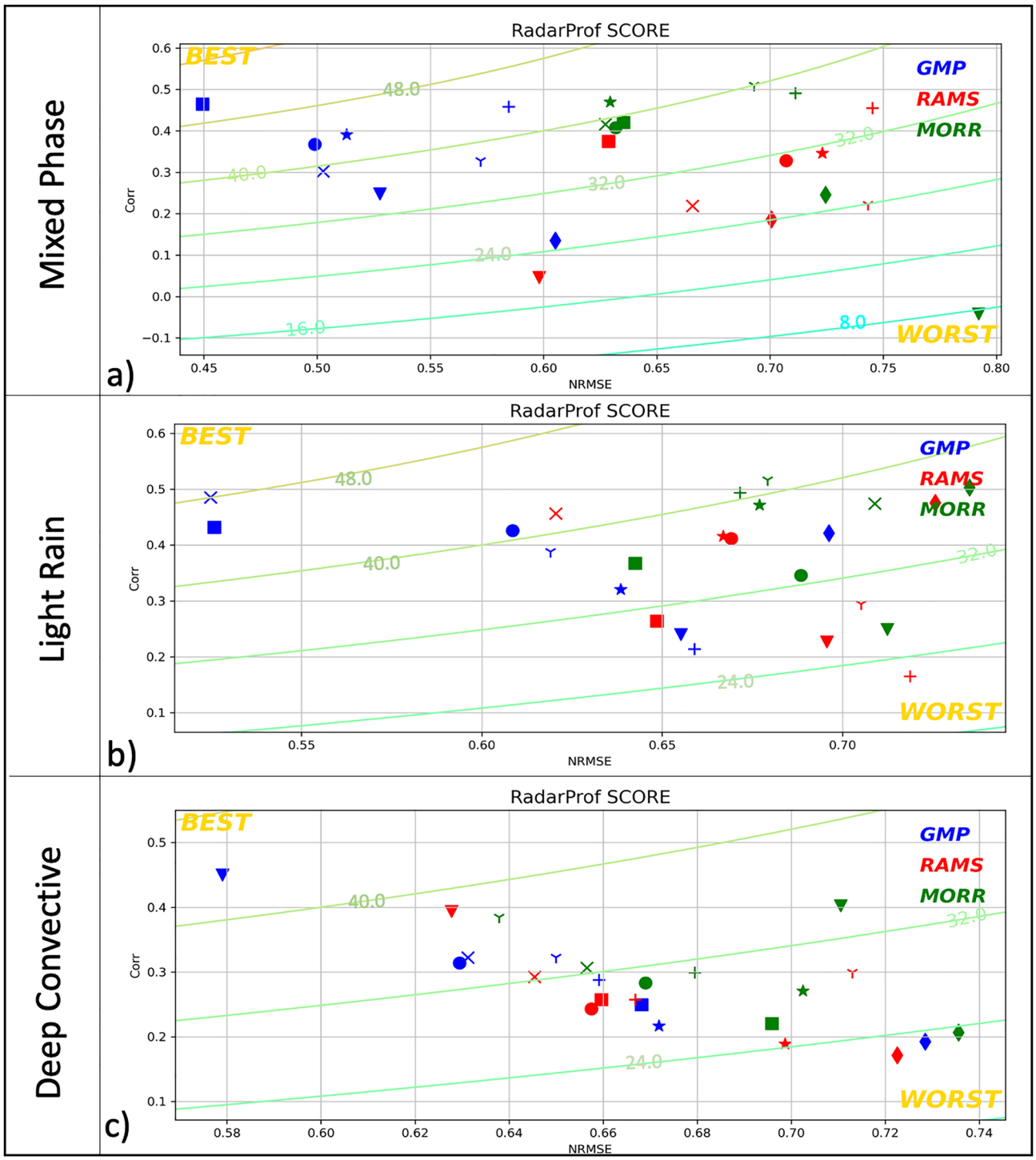

Figure 10. Same as Figure 9, but for RadarProf. Different symbols represent types of forcing (CTL(●), CTL_T(×), CTL_QV(+), CTL_TQV(Y), EN1(▼), EN2(■), EN3(★), and EN4(◆)).

than 0.05 (not shown here). It suggests that these three evaluation composites each validate a unique aspect of the model skill. Which skill score best identifies the best simulation among the ensemble simulations for each case depends on the purpose. To best match a simulation with observations for a specific time from satellite microwave and radar measurements, *RadarFrac* would be the most important skill to choose the best simulation.

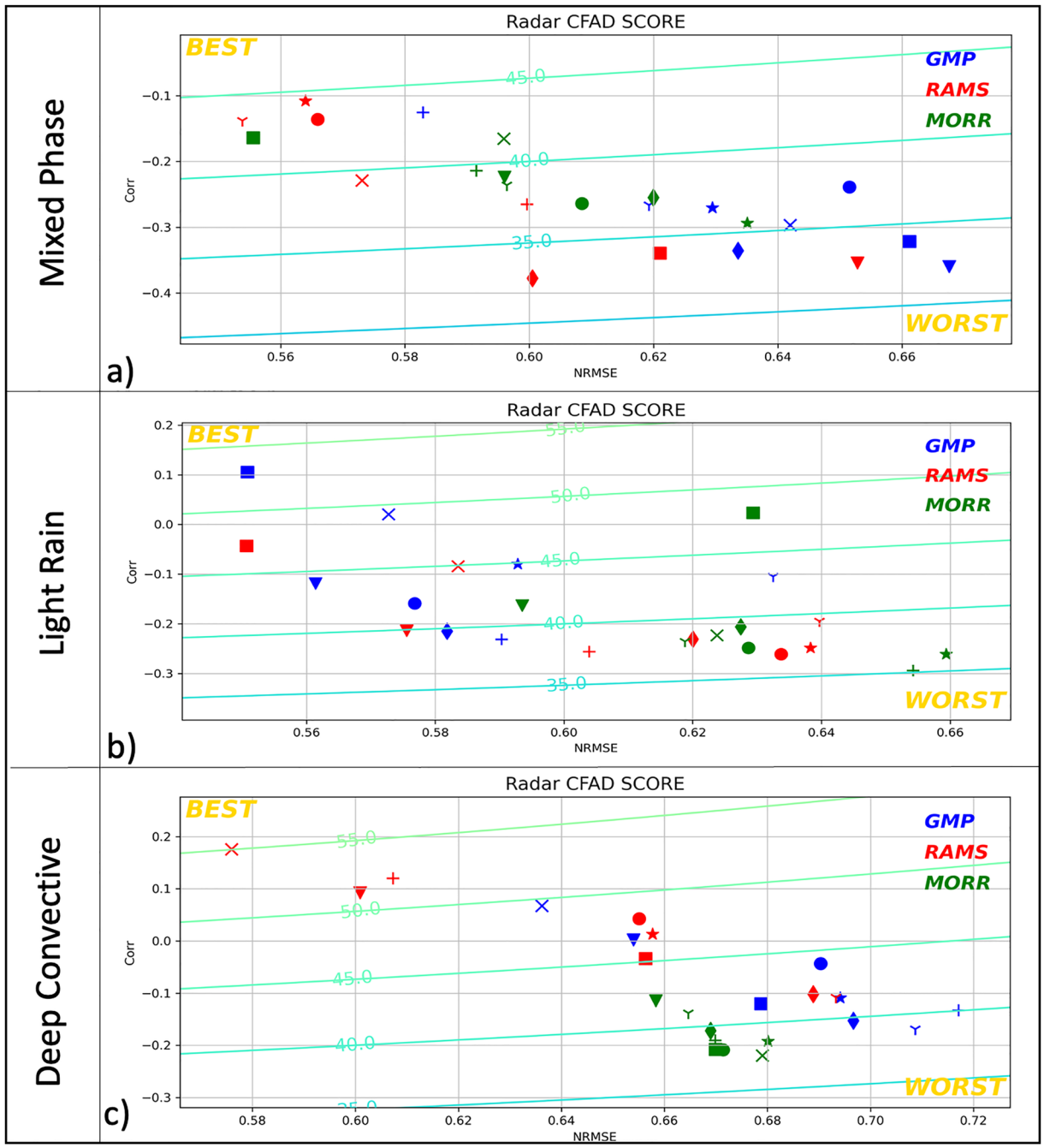

Figure 11. Same as Figure 9, but for RadarCFAD. Different symbols represent types of forcing (CTL(●), CTL_T(×), CTL_QV(+), CTL_TQV(Y), EN1(▼), EN2(■), EN3(★), and EN4(◆)).

7. Sensitivity of High-Resolution Simulations

Previous studies have argued that finer-resolution simulations are better for representing cloud processes, namely moving from traditional CRM resolution (a few km grid spacing) to LES resolution (200 m or less), when simulating deep convection (Bryan & Morrison, 2012; Bryan et al., 2003; Lang et al., 2007). Certainly, LES resolution, especially less than 200 m grid spacing, can better resolve large eddies, cumulus thermals, spectral kinetic energy maxima and boundary layer evolution without subgrid parameterizations (Bryan et al., 2003). It can also better reproduce radar reflectivity, cold pool structure and humidity versus observations (Bryan & Morrison, 2012). However, recent studies also show that the finer-resolution simulations tend to converge

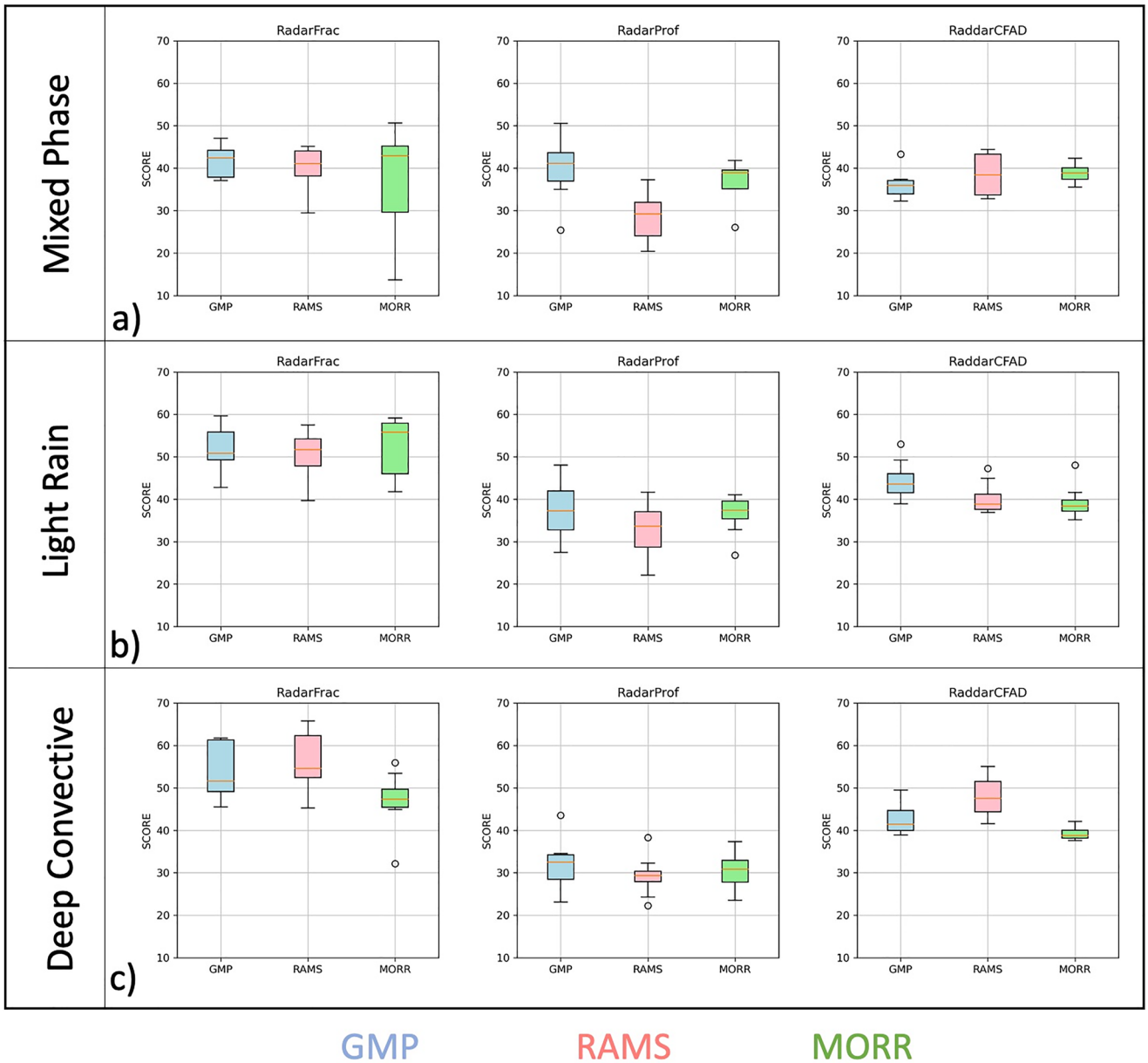

Figure 12. Boxplots of model skill score in each microphysics (GMP: blue, Regional Atmospheric Modeling System: red, and MORR: green) separately for each statistical composite and study case. A red line represents median; a box represents inter quartile (25%), whiskers represent outer quartile (50%), and open-circle dots represent outliers.

simulated states and processes toward the minimum grid spacing but do not necessarily improve the statistical representation of observed vertical velocity (Wang et al., 2022). In fact, none of these studies were systematically validated against the polarimetric radar observations.

For this, additional high-resolution ensemble simulations are used to investigate the impact of model resolution on the radar-based model skill scores. The default GCE ensemble simulations in BLOSSOM use 1 km grid spacing (Section 4), hereafter denoted as the CRM ensemble, as typical traditional CRMs use ∼1 km grid spacing. One simulation takes 30 min to 2.5 hr using 128 processors, depending on the microphysics. The higher-resolution simulations use 250 m grid spacing and require finer time steps (2 s) and more grid points (512 × 512 × 128 in *x*-*y*-*z*). Vertical grid spacing increases from roughly 44 up to 250 m in the lower troposphere and is then held constant at 250 m in the middle and upper troposphere. These finer-resolution simulations are denoted as very large eddy simulations (VLESs) hereafter as the spacing is slightly larger than the traditional LES threshold

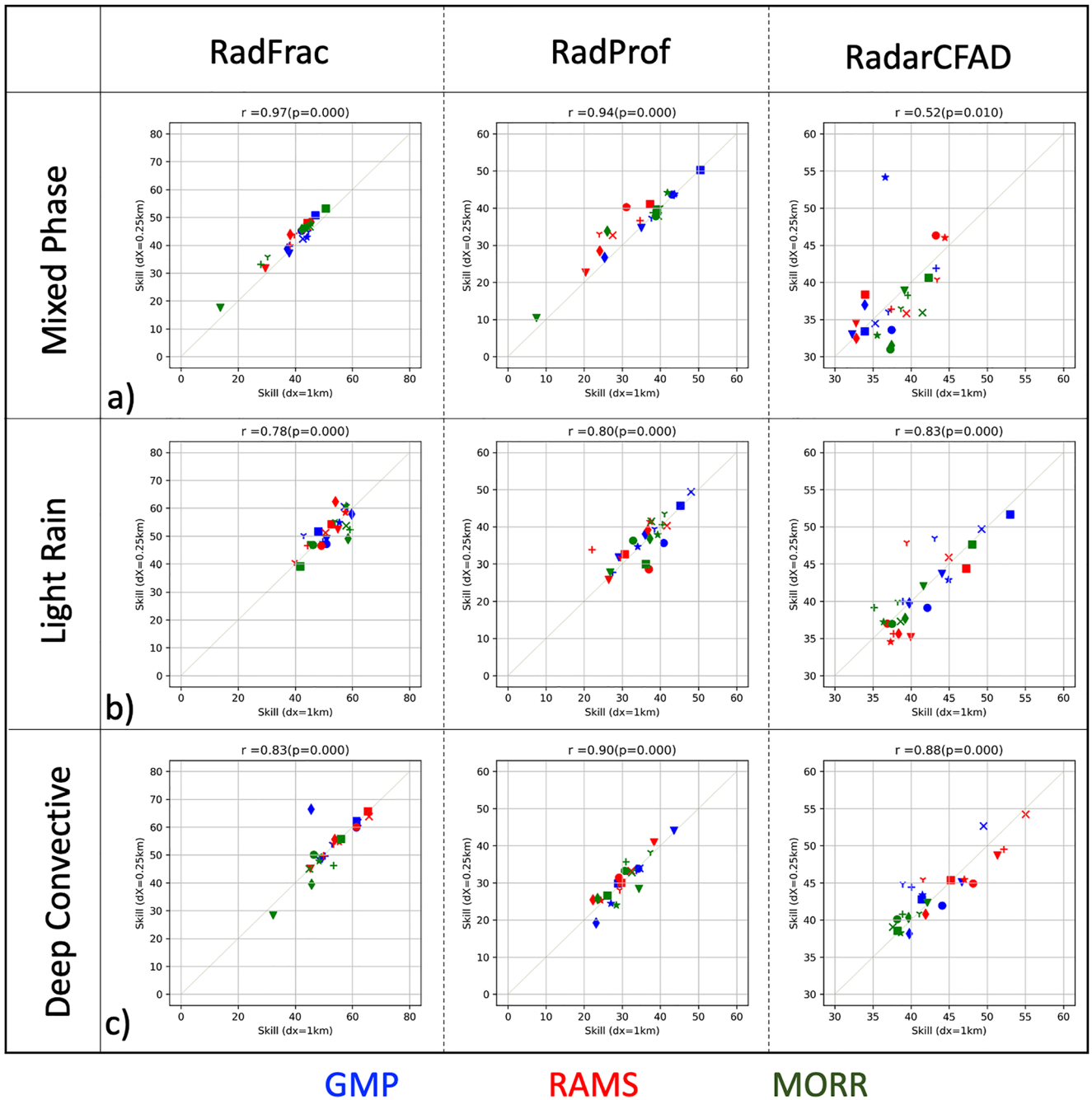

Figure 13. Scatter plots of skill scores between the cloud-resolving model and very large eddy simulation ensembles for each statistical composite and study case. Blue, red, and green dots represent GMP, Regional Atmospheric Modeling System, and MORR microphysics scheme. *r* is Spearman correlation coefficient, and *p* is *p*-value of statistical significance. Different symbols represent types of forcing (CTL(●), CTL_T(×), CTL_QV(+), CTL_TQV(Y), EN1(▼), EN2(■), EN3(★), and EN4(◆)).

(<200 m). VLES ensembles are forced identically to the default CRM ensembles and equivalently processed in POLARRIS and CfRAD to generate model skill scores. VLES ensembles require over a hundred times more computing power (i.e., processor number and integration time).

Figure 13 shows scatter plots of CRM versus VLES ensemble skill scores for all three cases and all three statistical composites. Corr (Spearman correlation) ranges from 0.97 (*RadFrac* for the mixed-phase case) to 0.78 (*RadFrac* for the light rain case), all of which have statistical significance in terms of *p*-value (less than 0.005). These results are concentrated along the one-on-one ratio line, meaning the VLES ensembles do not necessarily

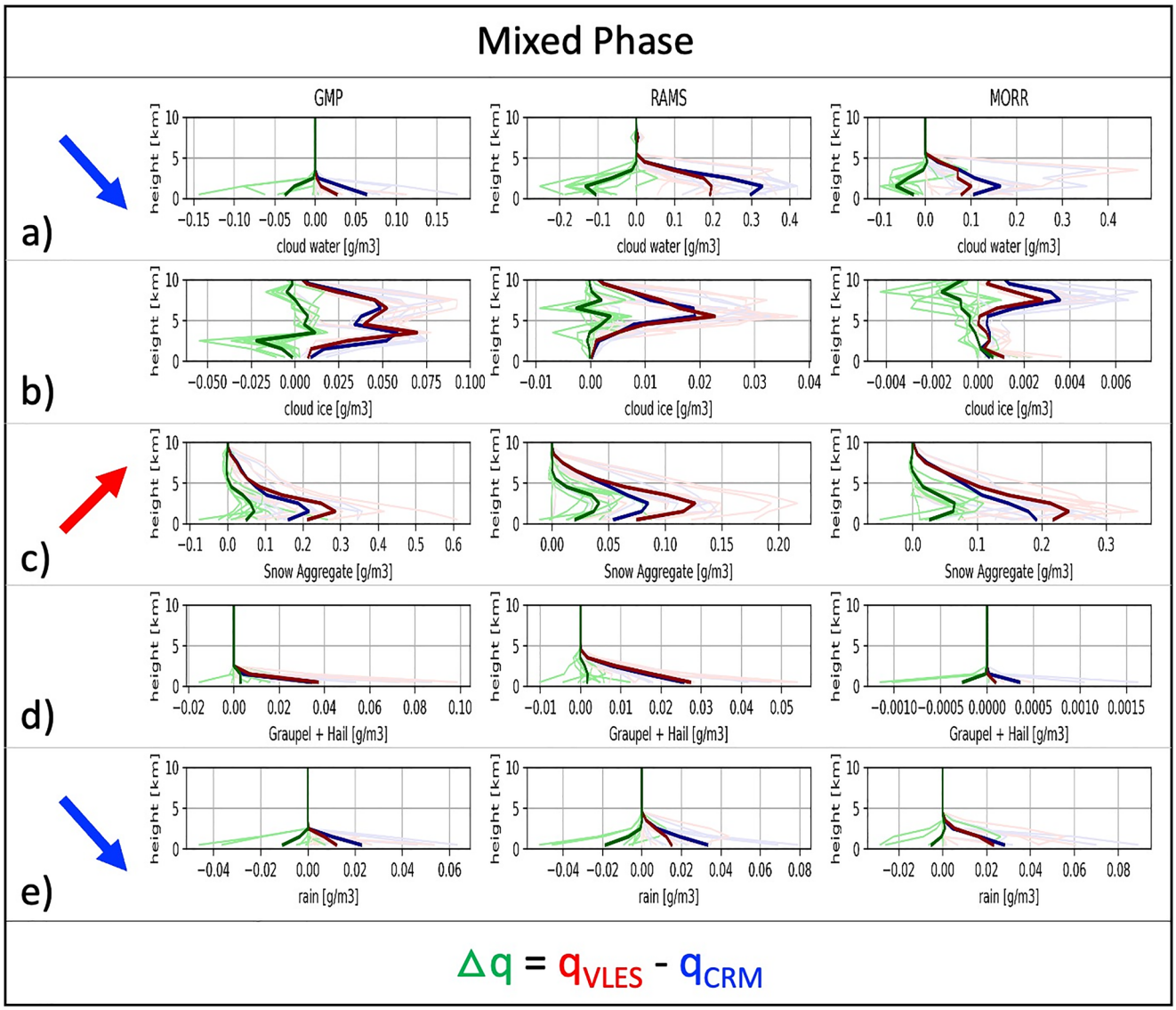

Figure 14. The domain-mean time-integrated hydrometeor profiles of (a) cloud water, (b) cloud ice, (c) snow aggregate, (d) graupel and hail, and (e) rain mass concentrations from the cloud-resolving model (blue) and very large eddy simulation (red) ensembles and their differences (green) for each microphysics scheme in the mixed-phase case. Light lines represent ensemble members, while thick line represent ensemble mean. Red (blue) arrows represent increase (decrease) transitions for all microphysics schemes.

improve the skill scores of the radar composites over the CRM ensembles. There is an anomalous point from GMP-EN3 in the mixed-phase case for *RadarCFAD*. It shows VLES drastically improved the skill score (53.2) over the CRM (36.6). This particular VLES was terminated before completion due to instability, which accidently improved the skill score for *RadarCFAD*. Otherwise, the scatter plots are significantly correlated with each other, meaning the best CRM skill scores for a particular composite also tend to have the best (or near the best) corresponding VLES skill scores.

Since the *S*-band polarimetric radar is more sensitive to larger particles in a common sampling volume, cloud droplets present with large or solid hydrometeors cannot be evaluated. Thus, hydrometeor profiles are used to see whether grid resolution has a consistent impact on a particular hydrometeor or not. Figures 14–16 show the mean hydrometeor profiles from the CRM (blue) and VLES (red) ensembles and their differences (green), separately plotted by the microphysics scheme, for all cases. First, there is a large spread in the mean profiles of the ensembles and the VLES-CRM differences; mean hydrometeor profiles (solid lines) also strongly vary among the three microphysics schemes. Cloud ice mass from the MORR scheme is one-order smaller than the other

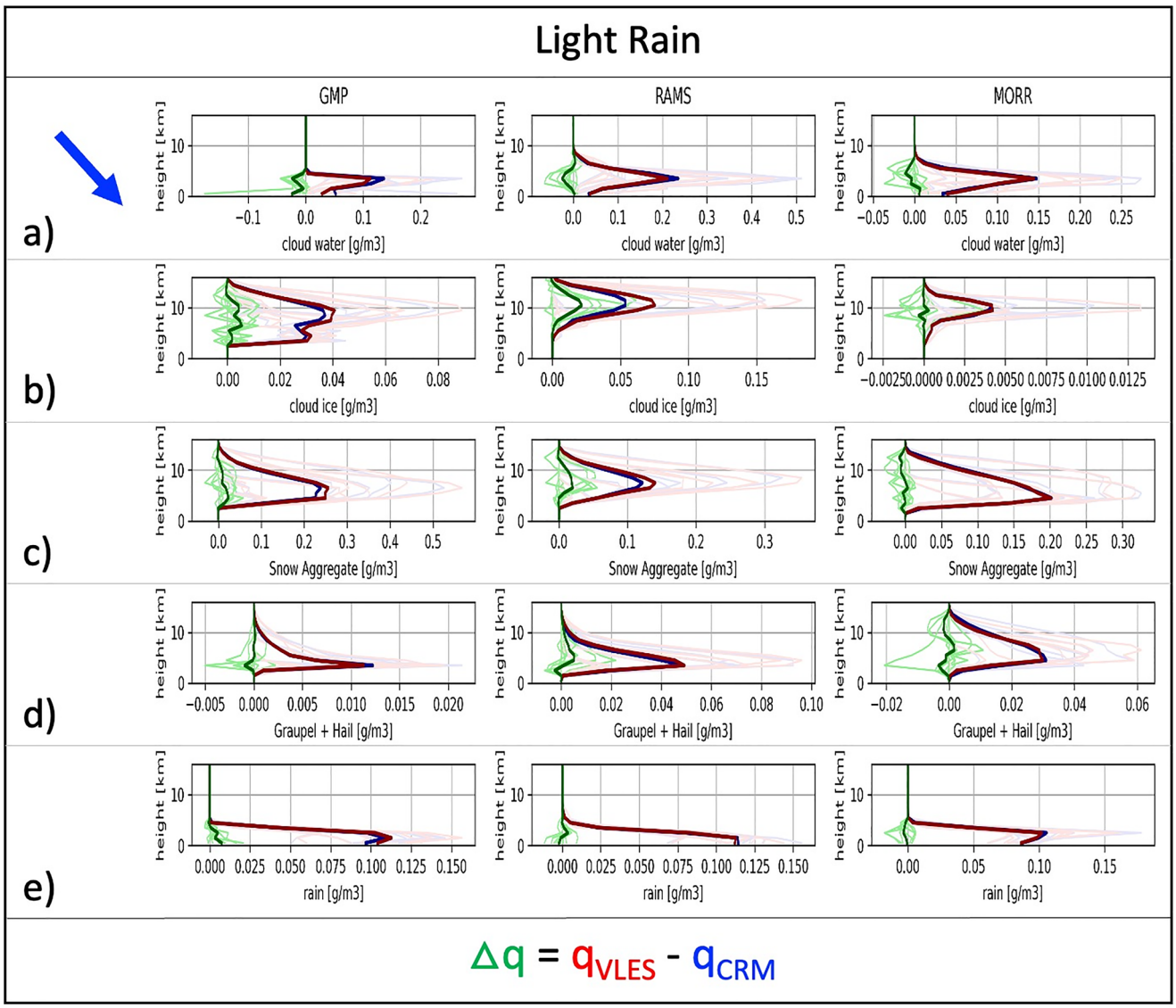

Figure 15. Same as Figure 14, but in the light rain case.

schemes, while cloud water mass from the GMP scheme tends to be concentrated at lower altitudes versus the other schemes.

Despite these differences, there are some similar trends in the hydrometeor profiles among three microphysics schemes. In the mixed-phase case, VLES reduces cloud and rain mass (Figures 14a and 14e), while increasing snow mass (Figure 14c). In the light rain case, VLES reduces cloud water mass only (Figure 15a). In the deep-convective case, VLES increases cloud ice mass (Figure 16b), while reducing graupel/hail mass (Figure 16d). Other hydrometeors in the different case studies had inconsistent sensitivities to VLES among the microphysics schemes.

These results are similar to Bryan and Morrison (2012) wherein idealized simulations of squall lines are sensitive to both microphysics setup and horizontal resolution, but the sensitivity varies among one-moment and two-moment microphysics due to different microphysics processes and feedbacks. Though the cloud systems and model setup differ, a similar conclusion is found here. Finer-resolution simulations allow for more explicit, detailed simulations of turbulence and cumulus thermals by the dynamic cores that can then interact with the various physical processes.

Despite these nonlinear responses to the model spatial resolutions, a robust finding is that VLES ensembles are not required to determine the best meteorological forcing and microphysics options in the three case studies, since skill scores are very close to the default CRM ensemble. In other words, computationally inexpensive

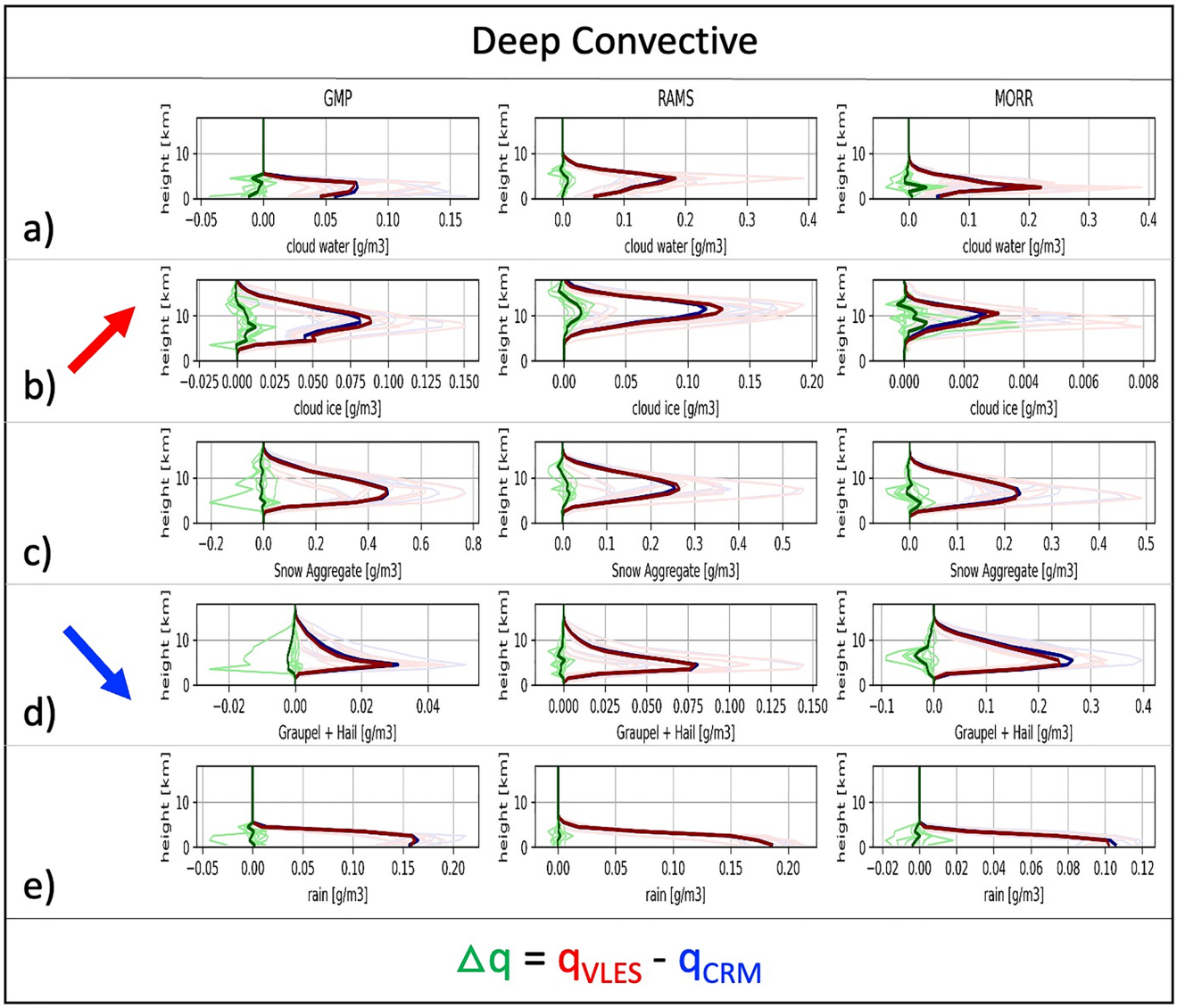

Figure 16. Same as Figure 14, but in the deep convective case.

CRM ensembles are sufficient for identifying the best (or near best) combination of meteorological forcing and microphysics before conducting more detailed LESs in support of remote sensing instruments and parameterization development as VLES ensembles require two orders of magnitude more computational power.

Note that our conclusion here also depends on how vertical resolutions are set with respect to horizontal grid spacing. Vertical resolution of the CRM simulations (96 levels) is set rather finer with respect for 1 km horizontal grid spacing. If we set up much coarse vertical resolution of traditional CRM (e.g., 30–50 levels), our conclusions could have been changed. In fact, finer resolution indeed can improve shallow convective clouds (e.g., Cheng et al., 2010). Nevertheless, our finding for a given spatial grid spacing set up in GCE simulations, we can minimize computational requirements to determine the best combination of microphysics and large-scale forcing.

8. Conclusions

The BLOSSOM project was initiated to use existing precipitation-measuring instruments to continuously improve and apply CRMs and simulations in support of NASA satellite/airborne programs and climate model improvement over WFF, where a variety of storm types are observed throughout the year over the coast of Virginia's Eastern

Shore. For this, LSF ensembles were developed to drive eight CRM member simulations with three different bulk microphysics schemes using the GCE model. Each simulation was processed by the POLARRIS polarimetric radar simulator to generate CfRadial radar output equivalent to NPOL observations in terms of parameters, dimensions, and sampling volumes. These simulated and observed CfRadial data were then processed through the newly developed CfRaD tool to generate three statistical composites: (a) time series of radar fractional coverages, (b) time series of domain-mean vertical polarimetric radar profiles, and (c) time-integrated polarimetric radar CFADs separated into convective and stratiform columns.

Simple skill scores were established by combining correlations and NRMSEs to objectively evaluate the ensemble simulations. Results were evaluated for three unique case studies: mixed-phase, light rain, and deep-convective. NPOL observations and GCE simulations reveal unique aspects of the microphysics via three statistical composites. While the GCE-POLARRIS simulations capture the overall structures of the radar composites, polarimetric signals associated with ice microphysics are still poorly represented due to the lack of ice microphysics diversity via uncertainties in crystal shapes, orientations, and their complex mixtures within microphysics scheme and forward model assumptions. No consistent advantages were found among the GMP, MORR, and RAMS microphysics schemes. The same was true for the LSF with skill scores spread randomly over the ensemble members. Although the ensemble LSFs used here do not cover all possible uncertainty sources such as uncertainties from background data and constraint variables (Tang et al., 2016) or variation in the parameters in cloud microphysical schemes (Posselt, 2016; Posselt et al., 2019), the randomly spread results shown in this study indicate that ensemble simulations using the different LSF and microphysics can well characterize the spread (uncertainty) in model performance, at least in this validation framework.

This new ensemble-based validation framework is used to demonstrate the impact of spatial resolution on the cloud-process simulations and model skill scores. Special finer-resolution (250 m grid spacing) ensemble simulations were additionally conducted and compared against the skill scores of the default coarser-resolution (1 km grid spacing) ensemble simulations. Skill scores between the fine- and coarse-resolution ensembles are significantly correlated and show no particular advantage for the fine-resolution simulations, suggesting computationally inexpensive coarse-resolution ensembles are sufficient for finding the best sets of LSF and microphysics without conducting fine-resolution ensembles, which require 100 times more computational resources. Thus, routine BLOSSOM simulations will be at coarser resolution with only the best LSF and microphysics combinations used for fine-resolution LESs. Observation-constrained LESs along with detailed satellite instrument simulators are well-suited to support future satellite and airborne remote sensing, serving as virtual field campaigns (Matsui et al., 2016), including the Atmosphere Observing System (https://aos.gsfc.nasa.gov).

To conclude, the systematic validation framework of ensemble cloud-process simulation using polarimetric radar observant and simulator is a new robust system to validate uncertainties of process-based simulations due to LSFs and microphysics schemes. This framework will be the core component of the BLOSSOM project, and planned to extend the LSF and simulations over several years during the NPOL operation. This framework can be also utilized for improvement of bulk microphysics schemes and various sensitivity studies for better understanding cloud and precipitation interactions and feedback, rather than relying on a single control simulation. It should be noted again that some precipitation systems driven by strong mesoscale forcing over WFF won't be well represented by the GCE simulations due to limitation of using idealized simulation configuration with large-scale forcing. Low skill scores from ensemble simulations could infer such cases. In that case, regional nested modeling approach could be applied.

In future, to address the uncertainty introduced by variability in microphysical parameters (e.g., particle size distribution parameters, densities of snow and graupel, assumed shapes of snow and graupel particles), additional parameter-based ensemble simulations will be demonstrated for select convective cases in future. These ensembles will be generated by perturbing cloud microphysical parameter values within each scheme. We will also examine the potential for introducing stochastic perturbations within cloud microphysics, following recent work in this area using the GCE model (Vukicevic et al., 2022). Similarly, retrievals of hydrometeor properties from polarimetric radar are also subject to uncertainty due to unknown hydrometeor properties, especially in frozen or mixed phase conditions (Posselt et al., 2015). These uncertainties will be explored by introducing perturbations into the POLARRIS model through examining the sensitivity to polarimetric radar variables.

Beyond the NPOL radar capability, additional detailed microphysics aspects will be examined using other measurements available at WFF, such as disdrometers, micro-rain radar, and D3R. Another asset, the NASA-GSFC

Aerosol, Cloud, Humidity Interactions Exploring and Validation Enterprise mobile laboratory features a scanning W-band (93.93 GHz) dual-polarization Doppler cloud radar, a vertically pointing Micro Rain Radar, and a 910 nm ceilometer. Integration of the X-Band (9.6 GHz) Atmospheric Doppler Ground-based Radar (X-BADGER, Emory et al., 2014) dual-polarization Doppler radar alongside the W-band is underway and will offer additional measurements of light to moderate precipitation. These instruments provide even more opportunities to understand storm microphysics and dynamics in the future at WFF.

Data Availability Statement

GCE code used in this manuscript, CfRad tool, Radar observation and simulation statistical data, Python plotting codes, and all figures are freely available in the NCCS Data Portal without registration (https://portal.nccs.nasa.gov/datashare/cloudlibrary/PUB_DATA/BLOSSOM_2022/). NPOL CfRadial data is available via Global Precipitation Mission (GPM) radar archive (https://pmm-gv.gsfc.nasa.gov/pub/gpmarchive/Radar/NPOL/Newark/).

Acknowledgments
This project is funded by the NASA Precipitation Measurement Mission (PMM) program (Program manager: Dr. William McCarty, Grant 80NSSC19K0724) and Modeling Analysis and Prediction (MAP) program (program manager: Dr. David Considine, Grant 80NSSC21K1135). Development of GCE is funded by the NASA GSFC Strategic Science resources. Development of POLARRIS is supported by the Department of Energy (DOE) Atmospheric System Research (ASR) program (Program Managers: Drs. Shaima Nasiri and Ashley Williamson, Grant DE-SC0021247). We also thank the NASA Advanced Supercomputing (NAS) Division and Center for Climate Simulation (NCCS) (Project Manager T. Lee at NASA HQ) for providing the computational resources to conduct the GCE, POLARRIS, and CfRad simulations and stored radar observations and model outputs. The Pacific Northwest National Laboratory (PNNL) is operated for the U.S. Department of Energy by Battelle Memorial Institute under Contract DE-AC05-76RLO1830. This paper is dedicated to Dr. Gail S. Jackson, a former PMM Program Manager, a former Chief of the NASA Goddard Mesoscale Atmospheric Processes Laboratory, a former Project Scientist for the NASA Global Precipitation Measurement (GPM) mission, and a former colleague.

References

Blossey, P. N., Bretherton, C. S., Cheng, A., Endo, S., Heus, T., Lock, A. P., & van der Dussen, J. J. (2016). CGILS Phase 2 LES intercomparison of response of subtropical marine low cloud regimes to CO_2 quadrupling and a CMIP3 composite forcing change. *Journal of Advances in Modeling Earth Systems*, *8*(4), 1714–1726. https://doi.org/10.1002/2016MS000765

Brandes, E. A., Zhang, G., & Vivekanandan, J. (2002). Experiments in rainfall estimation with a polarimetric radar in a subtropical environment. *Journal of Applied Meteorology*, *41*(6), 674–685. https://doi.org/10.1175/1520-0450(2002)041,0674:EIREWA.2.0.CO;2

Bringi, V. N., & Chandrasekar, V. (2001). *Polarimetric Doppler weather radar: Principles and applications* (p. 636). Cambridge University Press.

Bryan, G. H., & Morrison, H. (2012). Sensitivity of a simulated squall line to horizontal resolution and parameterization of microphysics. *Monthly Weather Review*, *140*(1), 202–225. https://doi.org/10.1175/mwr-d-11-00046.1

Bryan, G. H., Wyngaard, J. C., & Fritsch, J. M. (2003). Resolution requirements for the simulation of deep moist convection. *Monthly Weather Review*, *131*(10), 2394–2416. https://doi.org/10.1175/1520-0493(2003)131<2394:rrftso>2.0.co;2

Cheng, A., Xu, K.-M., & Steven, B. (2010). Effects of resolution on the simulation of boundary-layer clouds and partition of kinetic energy to subgrid scales. *Journal of Advances in Modeling Earth Systems*, *2*, 3. https://doi.org/10.3894/JAMES.2010.2.3

Chern, J.-D., Tao, W.-K., Lang, S. E., Li, X., & Matsui, T. (2020). Evaluating precipitation features and rainfall characteristics in a multi-scale modeling framework. *Journal of Advances in Modeling Earth Systems*, *12*(8), e2019MS002007. https://doi.org/10.1029/2019MS002007

Davies, H. C. (1983). Limitations of some common lateral boundary schemes used in regional NWP models. *Monthly Weather Review*, *111*(5), 1002–1012. https://doi.org/10.1175/1520-0493(1983)111<1002:losclb>2.0.co;2

Davies, L., Jakob, C., Cheung, K., Genio, A. D., Hill, A., Hume, T., et al. (2013). A single-column model ensemble approach applied to the TWP-ICE experiment. *Journal of Geophysical Research: Atmospheres*, *118*(12), 6544–6563. https://doi.org/10.1002/jgrd.50450

Dixon, M. J., & Lee, W.-C. (2016). NCAR/UNIDATA CfRadial data file format: Proposed CF-compliant netCDF format for moments data for RADAR and LIDAR in radial coordinates - v1.4. https://doi.org/10.5065/k0md-1642

Dolan, B., & Rutledge, S. A. (2009). A theory-based hydrometeor identification algorithm for X-band polarimetric radars. *Journal of Atmospheric and Oceanic Technology*, *26*(10), 2071–2088. https://doi.org/10.1175/2009JTECHA1208.1

Dolan, B., Rutledge, S. A., Lim, S., Chandrasekar, V., & Thurai, M. (2013). A robust C-band hydrometeor identification algorithm and application to a long-term polarimetric radar dataset. *Journal of Applied Meteorology and Climatology*, *52*(9), 2162–2186. https://doi.org/10.1175/JAMC-D-12-0275.1

Emory, A. E., Demoz, B., Vermeesch, K., & Hicks, M. (2014). Double bright band observations with high-resolution vertically pointing radar, lidar, and profilers. *Journal of Geophysical Research: Atmospheres*, *119*(13), 8201–8211. https://doi.org/10.1002/2013JD020063

Fan, J., Han, B., Varble, A., Morrison, H., North, K., Kollias, P., et al. (2017). Cloud-resolving model intercomparison of an MC3E squall line case: Part I—Convective updrafts. *Journal of Geophysical Research: Atmospheres*, *122*(17), 9351–9378. https://doi.org/10.1002/2017JD026622

Fridlind, A. M., Ackerman, A. S., Chaboureau, J. P., Fan, J., Grabowski, W. W., Hill, A. A., et al. (2012). A comparison of TWP-ICE observational data with cloud-resolving model results. *Journal of Geophysical Research*, *117*(D5), D05204. https://doi.org/10.1029/2011JD016595

Ghan, S., Randall, D., Xu, K. M., Cederwall, R., Cripe, D., Hack, J., et al. (2000). An intercomparison of single column model simulations of summertime midlatitude continental convection. *Journal of Geophysical Research*, *105*(D2), 2091–2121. https://doi.org/10.1029/1999jd900971

Grasso, L. D., & Lindsey, D. T. (2011). An example of the use of synthetic 3.9 μm GOES-12 imagery for two-moment microphysical evaluation. *International Journal of Remote Sensing*, *32*(8), 2337–2350. https://doi.org/10.1080/01431161003698294

Gustafson, W. I., Jr., Vogelmann, A. M., Li, Z., Cheng, X., Dumas, K. K., Endo, S., et al. (2020). The large-eddy simulation (LES) atmospheric radiation measurement (ARM) symbiotic simulation and observation (LASSO) activity for continental shallow convection. *Bulletin of the American Meteorological Society*, *101*(4), E462–E479. https://doi.org/10.1175/BAMS-D-19-0065.1

Hashino, T., Satoh, M., Hagihara, Y., Kubota, T., Matsui, T., Nasuno, T., & Okamoto, H. (2013). Evaluating cloud microphysics from NICAM against CloudSat and CALIPSO. *Journal of Geophysical Research: Atmospheres*, *118*(13), 1–20. https://doi.org/10.1002/jgrd.50564

Houze, R. A., Jr. (1997). Stratiform precipitation in regions of convection: A meteorological paradox? *Bulletin America Meteorology Social*, *78*(10), 2179–2196. https://doi.org/10.1175/1520-0477(1997)078<2179:spiroc>2.0.co;2

Houze, R. A., Jr., McMurdie, L. A., Petersen, W. A., Schwaller, M. R., Baccus, W., Lundquist, J. D., et al. (2017). The olympic mountains experiment (OLYMPEX). *Bulletin of the American Meteorological Society*, *98*(10), 2167–2188. https://doi.org/10.1175/BAMS-D-16-0182.1

Iguchi, T., Matsui, T., Shi, J. J., Tao, W.-K., Khain, A. P., Hou, A., et al. (2012a). Numerical analysis using WRF-SBM for the cloud microphysical structures in the C3VP field campaign: Impacts of supercooled droplets and resultant riming on snow microphysics. *Journal of Geophysical Research*, *117*(D23), D23206. https://doi.org/10.1029/2012JD018101

Iguchi, T., Matsui, T., Tao, W., Khain, A., Phillips, V., Kidd, C., et al. (2014). WRF-SBM simulations of melting layer structure in mixed-phase precipitation events observed during LPVEx. *Journal of Applied Meteorology and Climatology*, *53*(12), 2710–2731. https://doi.org/10.1175/JAMC-D-13-0334.1

Iguchi, T., Matsui, T., Tokay, A., Kollias, P., & Tao, W.-K. (2012b). Two distinct modes in one-day rainfall event during MC3E field campaign: Analyses of disdrometer observations and WRF-SBM simulation. *Geophysical Research Letters*, *39*(24), L24805. https://doi.org/10.1029/2012GL053329

Iguchi, T., Rutledge, S. A., Tao, W.-K., Matsui, T., Dolan, B., Lang, S. E., & Barnum, J. (2020). Impacts of aerosol and environmental conditions on maritime and continental deep convective systems using a bin microphysical model. *Journal of Geophysical Research: Atmospheres*, *125*(12), e2019JD030952. https://doi.org/10.1029/2019JD030952

Jensen, M. P., Toto, T., Troyan, D., Ciesielski, P. E., Holdridge, D., Kyrouac, J., et al. (2015). The midlatitude continental convective clouds experiment (MC3E) sounding network: Operations, processing and analysis. *Atmospheric Measurement Techniques*, *8*(1), 421–434. https://doi.org/10.5194/amt-8-421-2015

Jung, Y., Xue, M., & Zhang, G. (2010). Simulations of polarimetric radar signatures of a supercell storm using a two-moment bulk microphysics scheme. *Journal of Applied Meteorology and Climatology*, *49*(1), 146–163. https://doi.org/10.1175/2009JAMC2178.1

Kidd, C., Matsui, T., Chern, J., Mohr, K., Kummerow, C., & Randel, D. (2016). Global precipitation estimates from cross-track passive microwave observations using a physically-based retrieval scheme. *Journal of Hydrometeorology*, *17*(1), 383–400. https://doi.org/10.1175/JHM-D-15-0051.1

Kollias, P., & Tatarevic, A. (2017). User's guide CR-SIM SOFTWARE v 2.2. Retrieved from http://www.meteo.mcgill.ca/~aleksandra/CRSIM/crsim-UserGuide-v2.0.pdf

Kumjian, M. R., Prat, O. P., Reimel, K. J., van Lier-Walqui, M., & Morrison, H. (2022). Dual-polarization radar fingerprints of precipitation physics: A review. *Remote Sensing*, *14*(15), 3706. https://doi.org/10.3390/rs14153706

Lang, S., & Tao, W.-K. (2018). The next-generation Goddard convective-stratiform heating algorithm: New tropical and warm season retrievals for GPM. *Journal of Climate*, *31*(15), 5997–6026. https://doi.org/10.1175/JCLI-D-17-0224.1

Lang, S., Tao, W. K., Chern, J. D., Wu, D., & Li, X. (2014). Benefits of a 4th ice class in the simulated radar reflectivities of convective systems using a bulk microphysics scheme. *Journal of the Atmospheric Sciences*, *71*(10), 3583–3612. https://doi.org/10.1175/JAS-D-13-0330.1

Lang, S., Tao, W.-K., Cifelli, R., Olson, W., Halverson, J., Rutledge, S., & Simpson, J. (2007). Improving simulations of convective system from TRMM LBA: Easterly and westerly regimes. *Journal of the Atmospheric Sciences*, *64*(4), 1141–1164. https://doi.org/10.1175/JAS3879.1

Lang, S., Tao, W.-K., Simpson, J., & Ferrier, B. (2003). Modeling of convective–stratiform precipitation processes: Sensitivity to partitioning methods. *Journal of Applied Meteorology*, *42*(4), 505–527. https://doi.org/10.1175/1520-0450(2003)042%3C0505:MOCSPP%3E2.0.CO;2

Li, X., Janiga, M. A., Wang, S., Tao, W.-K., Rowe, A., Xu, W., et al. (2018). Evolution of precipitation structure during the November DYNAMO MJO event: Cloud-resolving model intercomparison and cross validation using radar observations. *Journal of Geophysical Research: Atmospheres*, *123*(7), 3530–3555. https://doi.org/10.1002/2017JD027775

Li, X., Tao, W.-K., Matsui, T., Liu, C., & Masunaga, H. (2010). Improving a spectral bin microphysical scheme using long-term TRMM satellite observations. *Quarterly Journal of Royal Metrological Society*, *136*(647), 382–399. https://doi.org/10.1002/qj.569

Loftus, A. M., Tsay, S.-C., Pantina, P., Nguyen, C., Gabriel, P. M., Nguyen, X. A., et al. (2016). Coupled aerosol-cloud systems over northern Vietnam during 7-SEAS/BASELInE: A radar and modeling perspective. *Aerosol and Air Quality Research*, *16*(11), 2768–2785. https://doi.org/10.4209/aaqr.2015.11.0631

Lopez-Gomez, I., Cohen, Y., He, J., Jaruga, A., & Schneider, T. (2020). A generalized mixing length closure for eddy-diffusivity mass-flux schemes of turbulence and convection. *Journal of Advances in Modeling Earth Systems*, *12*(11), e2020MS002161. https://doi.org/10.1029/2020MS002161

Masunaga, H., Matsui, T., Tao, W.-K., Hou, A. Y., Kummerow, C., Nakajima, T., et al. (2011). Satellite data simulator unit: Multi-sensor and multi-frequency satellite simulator package. *Bulletin of the American Meteorological Society*, *91*(12), 1625–1632. https://doi.org/10.1175/2010BAMS2809.1

Matsui, T., Chern, J., Tao, W.-K., Lang, S., Satoh, M., Hashino, T., & Kubota, T. (2016). On the land-ocean contrast of tropical convection and microphysics statistics derived from TRMM satellite signals and global storm-resolving models. *Journal of Hydrometeorology*, *17*(5), 1425–1445. https://doi.org/10.1175/JHM-D-15-0111.1

Matsui, T., Dolan, B., Iguchi, T., Rutledge, S. A., Tao, W., & Lang, S. (2020). Polarimetric radar characteristics of simulated and observed intense convective cores for a midlatitude continental and tropical maritime environment. *Journal of Hydrometeorology*, *21*(3), 501–517. https://doi.org/10.1175/JHM-D-19-0185.1

Matsui, T., Dolan, B., Rutledge, S. A., Tao, W.-K., Iguchi, T., Barnum, J., & Lang, S. E. (2019). Polarris: A POLArimetric radar retrieval and instrument simulator. *Journal of Geophysical Research: Atmospheres*, *124*(8), 4634–4657. https://doi.org/10.1029/2018JD028317

Matsui, T., Iguchi, T., Li, X., Han, M., Tao, W. K., Petersen, W., et al. (2013). GPM satellite simulator over ground validation sites. *Bulletin America Meteorology Social*, *94*(11), 1653–1660. https://doi.org/10.1175/BAMS-D-12-00160.1

Matsui, T., Santanello, J., Shi, J. J., Tao, W.-K., Wu, D., Peters-Lidard, C., et al. (2014). Introducing multisensor satellite radiance-based evaluation for regional Earth system modeling. *Journal of Geophysical Research: Atmospheres*, *119*(13), 8450–8475. https://doi.org/10.1002/2013JD021424

Matsui, T., Zeng, X., Tao, W.-K., Masunaga, H., Olson, W., & Lang, S. (2009). Evaluation of long-term cloud-resolving model simulations using satellite radiance observations and multifrequency satellite simulators. *Journal of Atmospheric and Oceanic Technology*, *26*(7), 1261–1274. https://doi.org/10.1175/2008jtecha1168.1

Matsui, T., Zhang, S. Q., Tao, W.-K., Lang, S., Ichoku, C., & Peters-Lidard, C. (2018). Impact of radiation frequency, precipitation radiative forcing, and radiation column aggregation on convection-permitting west African monsoon simulations. *Climate Dynamics*, *55*(1–2), 1–21. https://doi.org/10.1007/s00382-018-4187-2

Meyers, M. P., Walko, R. L., Harrington, J. Y., & Cotton, W. R. (1997). New RAMS cloud microphysics parameterization: Part II. The two-moment scheme. *Atmospheric Research*, *45*(1), 3–39. https://doi.org/10.1016/s0169-8095(97)00018-5

Morrison, H., Curry, J. A., & Khvorostyanov, V. I. (2005). A new double-moment microphysics parameterization for application in cloud and climate models. Part I: Description. *Journal of the Atmospheric Sciences*, *62*(6), 1665–1677. https://doi.org/10.1175/JAS3446

Oue, M., Kollias, P., Shapiro, A., Tatarevic, A., & Matsui, T. (2019). Investigation of observational error sources in multi-Doppler-radar three-dimensional variational vertical air motion retrievals. *Atmospheric Measurement Techniques*, *12*(3), 1999–2018. https://doi.org/10.5194/amt-12-1999-2019

Pielke, R. A., Sr. (2013). *Mesoscale meteorological modeling* (3rd ed., p. 760). Academic Press.

Posselt, D. J. (2016). A Bayesian examination of deep convective squall line sensitivity to changes in cloud microphysical parameters. *Journal of the Atmospheric Sciences*, *73*(2), 637–665. https://doi.org/10.1175/jas-d-15-0159.1

Posselt, D. J., He, F., Bukowski, J., & Reid, J. S. (2019). On the relative sensitivity of a tropical deep convective storm to changes in environment and cloud microphysical parameters. *Journal of the Atmospheric Sciences*, *76*(4), 1163–1185. https://doi.org/10.1175/JAS-D-18-0181.1

Posselt, D. J., Li, X., Tushaus, S. A., & Mecikalski, J. R. (2015). Assimilation of dual-polarization radar observations in mixed- and ice- phase regions of convective storms: Information content and forward model errors. *Monthly Weather Review*, *143*(7), 2611–2636. https://doi.org/10.1175/mwr-d-14-00347.1

Putnam, B. J., Xue, M., Jung, Y., Snook, N., & Zhang, G. (2014). The analysis and prediction of microphysical states and polarimetric radar variables in a mesoscale convective system using double-moment microphysics, multinetwork radar data, and the ensemble Kalman Filter. *Monthly Weather Review*, *142*(1), 141–162. https://doi.org/10.1175/MWR-D-13-00042.1

Randall, D. A., Xu, K., Somerville, R. J., & Iacobellis, S. (1996). Single-column models and cloud ensemble models as links between observations and climate models. *Journal of Climate*, *9*(8), 1683–1697. https://doi.org/10.1175/1520-0442(1996)009<1683:SCMACE>2.0.CO;2

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., et al. (2011). MERRA: NASA's modern-era Retrospective analysis for research and applications. *Journal of Climate*, *24*(14), 3624–3648. https://doi.org/10.1175/jcli-d-11-00015.1

Rutan, D. A., Kato, S., Doelling, D. R., Rose, F. G., Nguyen, L. T., Caldwell, T. E., & Loeb, N. G. (2015). CERES synoptic product: Methodology and validation of surface radiant flux. *Journal of Atmospheric and Oceanic Technology*, *32*(6), 1121–1143. https://doi.org/10.1175/JTECH-D-14-00165.1

Ryzhkov, A., Pinsky, M., Pokrovsky, A., & Khain, A. (2011). Polarimetric radar observation operator for a cloud model with spectral microphysics. *Journal of Applied Meteorology and Climatology*, *50*(4), 873–894. https://doi.org/10.1175/2010JAMC2363.1

Ryzhkov, A., Zhang, P., Reeves, H., Kumjian, M., Tschallener, T., Trömel, S., & Simmer, C. (2016). Quasi-vertical profiles—A new way to look at polarimetric radar data. *Journal of Atmospheric and Oceanic Technology*, *33*(3), 551–562. https://doi.org/10.1175/jtech-d-15-0020.1

Saleeby, S. M., & Cotton, W. R. (2004). A large-droplet mode and prognostic number concentration of cloud droplets in the Colorado State University Regional Atmospheric Modeling System (RAMS). Part I: Module descriptions and supercell test simulations. *Journal of Applied Meteorology*, *43*(1), 182–195. https://doi.org/10.1175/1520-0450(2004)043<0182:almapn>2.0.co;2

Saleeby, S. M., Cotton, W. R., Lowenthal, D., & Messina, J. (2013). Aerosol impacts on the microphysical growth processes of orographic snowfall. *Journal of Applied Meteorology and Climatology*, *52*(4), 834–852. https://doi.org/10.1175/jamc-d-12-0193.1

Saleeby, S. M., Herbener, S. R., van den Heever, S. C., & L'Ecuyer, T. (2015). Impacts of cloud droplet nucleating aerosols on shallow tropical convection. *Journal of the Atmospheric Sciences*, *72*(4), 1369–1385. https://doi.org/10.1175/jas-d-14-0153.1

Saleeby, S. M., & van den Heever, S. C. (2013). Developments in the CSU-RAMS aerosol model: Emissions, nucleation, regeneration, deposition, and radiation. *Journal of Applied Meteorology and Climatology*, *52*(12), 2601–2622. https://doi.org/10.1175/JAMC-D-12-0312.1

Saleeby, S. M., van den Heever, S. C., Marinescu, P. J., Kreidenweis, S. M., & DeMott, P. J. (2016). Aerosol effects on the anvil characteristics of mesoscale convective systems. *Journal of Geophysical Research: Atmospheres*, *121*(18), 10880–10901. https://doi.org/10.1002/2016JD025082

Schrom, R. S., & Kumjian, M. R. (2018). Bulk-density representations of branched planar ice crystals: Errors in the polarimetric radar variables. *Journal of Applied Meteorology and Climatology*, *57*(2), 333–346. https://doi.org/10.1175/JAMC-D-17-0114.1

Shi, J. J., Braun, S. A., Tao, Z., & Matsui, T. (2021). Influence of the saharan air layer on Hurricane Nadine (2021). Part I: Observations from the Hurricane and severe storm sentinel (HS3) investigation and modeling result. *Monthly Weather Review*, *149*(10), 3541–3562. Retrieved from https://journals.ametsoc.org/view/journals/mwre/149/10/MWR-D-20-0344.1.xml

Shi, J. J., Matsui, T., Tao, W. K., Tan, Q., Peters-Lidard, C., Chin, M., et al. (2014). Implementation of an aerosol-cloud microphysics-radiation coupling into the NASA Unified WRF: Simulation results for the 6-7 August 2006 AMMA special observing period. *Quarterly Journal of the Royal Meteorological Society*, *140*(684), 2158–2175. https://doi.org/10.1002/qj.2286

Shi, J. J., Tao, W.-K., Matsui, T., Hou, A., Lang, S., Peters-Lidard, C., et al. (2010). Microphysical properties of the January 20–22 2007 snow events over Canada: Comparison with in-situ and satellite observations. *Journal of Applied Meteorology and Climatology*, *49*(11), 2246–2266. https://doi.org/10.1175/2010jamc2282.1

Sisterson, D. L., Peppler, R. A., Cress, T. S., Lamb, P. J., & Turner, D. D. (2016). The ARM Southern Great Plains (SGP) site. *Meteorological Monographs*, *57*, 6.1–6.14. https://doi.org/10.1175/AMSMONOGRAPHS-D-16-0004.1

Snyder, J. C., Bluestein, H. B., Dawson, D. T., II, & Jung, Y. (2017). Simulations of polarimetric, X-band radar signatures in supercells. Part I: Description of experiment and simulated ρhv rings. *Journal of Applied Meteorology and Climatology*, *56*(7), 1977–1999. https://doi.org/10.1175/JAMC-D-16-0138.1

Soong, S.-T., & Tao, W.-K. (1980). Response of deep tropical clouds to mesoscale processes. *Journal of the Atmospheric Sciences*, *37*(9), 2016–2036. https://doi.org/10.1175/1520-0469(1980)037<2016:rodtcc>2.0.co;2

Steiner, M., Houze, R. A., Jr., & Yuter, S. E. (1995). Climatological characterization of three-dimensional storm structure from operational radar and rain gauge data. *Journal of Applied Meteorology*, *34*(9), 1978–2007. https://doi.org/10.1175/1520-0450(1995)034<1978:ccotds>2.0.co;2

Tang, S., Xie, S., Guo, Z., Hong, S.-Y., Khouider, B., Klocke, D., et al. (2022). Long-term single-column model intercomparison of diurnal cycle of precipitation over midlatitude and tropical land. *Quarterly Journal of the Royal Meteorological Society*, *641*(743), 669. https://doi.org/10.1002/qj.4222

Tang, S., Zhang, M., & Xie, S. (2016). An ensemble constrained variational analysis of atmospheric forcing data and its application to evaluate clouds in CAM5. *Journal of Geophysical Research: Atmospheres*, *121*(1), 33–48. https://doi.org/10.1002/2015JD024167

Tao, W.-K., Iguchi, T., & Lang, S. (2019). Expanding the Goddard CSH algorithm for GPM: New extratropical retrievals. *Journal of Applied Meteorology and Climatology*, *58*(5), 921–946. https://doi.org/10.1175/JAMC-D-18-0215.1

Tao, W.-K., Lang, S., Zeng, X., Li, X., Matsui, T., Mohr, K., et al. (2014). The Goddard Cumulus Ensemble model (GCE): Improvements and applications for studying precipitation processes. *Atmospheric Research*, *143*, 392–424. https://doi.org/10.1016/j.atmosres.2014.03.005

Tao, W.-K., Li, X., Khain, A., Matsui, T., Lang, S., & Simpson, J. (2007). Role of atmospheric aerosol concentration on deep convective precipitation: Cloud-resolving model simulations. *Journal of Geophysical Research*, *112*(D24), D24S18. https://doi.org/10.1029/2007JD008728

Tao, W.-K., & Simpson, J. (1993). The Goddard cumulus ensemble model. Part I: Model description. *Terrestrial, Atmospheric and Oceanic Sciences*, *4*(1), 19–54. https://doi.org/10.3319/TAO.1993.4.1.35(A)

Tao, W.-K., Simpson, J., Baker, D., Braun, S., Chou, M.-D., Ferrier, B., et al. (2003). Microphysics, radiation and surface processes in the Goddard Cumulus Ensemble (GEC) model. *Meteorology and Atmospheric Physics*, *82*(1–4), 97–137. https://doi.org/10.1007/s00703-001-0594-7

Tao, W.-K., Wu, D., Lang, S., Chern, J.-D., Peters-Lidard, C., Fridlind, A., & Matsui, T. (2016). High-resolution NU-WRF simulations of a deep convective-precipitation system during MC3E: Further improvements and comparisons between Goddard microphysics schemes and observations. *Journal of Geophysical Research: Atmospheres*, *121*(3), 1278–1305. https://doi.org/10.1002/2015JD023986

Tao, W.-K., Wu, D., Matsui, T., Peters-Lidard, C., Lang, S., Hou, A., et al. (2013). Precipitation intensity and variation during MC3E: A numerical modeling study. *Journal of Geophysical Research: Atmosphere*, *118*(13), 7199–7218. https://doi.org/10.1002/jgrd.50410

Thurai, M., Bringi, V., Wolff, D., Marks, D., & Pabla, C. (2021). Testing the drop-size distribution-based separation of stratiform and convective rain using radar and disdrometer data from a mid-latitude coastal region. *Atmosphere*, *12*(3), 392. https://doi.org/10.3390/atmos12030392

van Lier-Walqui, M., Fridland, A. M., Ackerman, A. S., Collis, S., Helmus, J., MacGorman, D. R., et al. (2016). On polarimetric radar signatures of deep convection for model evaluation: Columns of specific differential phase observed during MC3E. *Monthly Weather Review*, *144*(2), 737–758. https://doi.org/10.1175/mwr-d-15-0100.1

Vivekanandan, J., Adams, W. M., & Bringi, V. N. (1991). Rigorous approach to polarimetric radar modeling of hydrometeor orientation distributions. *Journal of Applied Meteorology*, *30*(8), 1053–1063. https://doi.org/10.1175/1520-0450(1991)030<1053:RATPRM>2.0.CO;2

Vukicevic, T., Posselt, D. J., & Stankovich, A. (2022). Sensitivity of modeled microphysics to stochastically perturbed parameters. *Journal of Advances in Modeling Earth Systems*, *14*, 20. https://doi.org/10.1029/2021MS002933

Waliser, D. E., Ridout, J. A., Xie, S., & Zhang, M. (2002). Variational objective analysis for atmospheric field programs: A model assessment. *Journal of the Atmospheric Sciences*, *59*(24), 3436–3456. https://doi.org/10.1175/1520-0469(2002)059<3436:VOAFAF>2.0.CO;2

Walko, R. L., Cotton, W. R., Meyers, M. P., & Harrington, J. Y. (1995). New RAMS cloud microphysics parameterization: Part I. The single-moment scheme. *Atmospheric Research*, *38*(1–4), 29–62. https://doi.org/10.1016/0169-8095(94)00087-t

Wang, D., Prein, A. F., Giangrande, S. E., Ramos-Valle, A., Ge, M., & Jensen, M. P. (2022). Convective updraft and downdraft characteristics of continental mesoscale convective systems in the model gray zone. *Journal of Geophysical Research: Atmospheres*, *127*(16), e2022JD036746. https://doi.org/10.1029/2022JD036746

Wielicki, B. A., Barkstrom, B. R., Harrison, E. F., Lee, R. B., Smith, G. L., & Cooper, J. E. (1996). Clouds and the Earth's radiant energy system (CERES): An earth observing system experiment. *Bulletin America Meteorology Social*, *77*(5), 853–868. https://doi.org/10.1175/1520-0477(1996)077<0853:CATERE>2.0.CO;2

Wolff, D. B., Marks, D. A., & Petersen, W. A. (2015). General application of the relative calibration adjustment (RCA) technique for monitoring and correcting radar reflectivity calibration. *Journal of Atmospheric and Oceanic Technology*, *4*(3), 365–380. https://doi.org/10.1175/jtech-d-13-00185.1

Xie, S., Cederwall, R. T., & Zhang, M. (2004). Developing long-term single-column model/cloud system–resolving model forcing data using numerical weather prediction products constrained by surface and top of the atmosphere observations. *Journal of Geophysical Research*, *109*(D1), D01104. https://doi.org/10.1029/2003JD004045

Xie, S., Xu, K. M., Cederwall, R. T., Bechtold, P., Genio, A. D. D., Klein, S. A., et al. (2002). Intercomparison and evaluation of cumulus parameterization under summertime mid-latitude continental conditions. *Quarterly Journal of the Royal Meteorological Society*, *128*(582), 1095–1136. https://doi.org/10.1256/003590002320373229

Xie, S., Zhang, M., Branson, M., Cederwall, R. T., Del Genio, A. D., Eitzen, Z. A., et al. (2005). Simulations of midlatitude frontal clouds by single-column and cloud-resolving models during the Atmospheric Radiation Measurement March 2000 cloud intensive operational period. *Journal of Geophysical Research*, *110*(D15), D15S03. https://doi.org/10.1029/2004JD005119

Xu, K.-M., Cederwall, R. T., Donner, L. J., Grabowski, W. W., Guichard, F., Johnson, D. E., et al. (2002). An intercomparison of cloud-resolving models with the atmospheric radiation measurement summer 1997 intensive observation period data. *Quarterly Journal of the Royal Meteorological Society*, *128*(580), 593–624. https://doi.org/10.1256/003590002321042117

Xu, K.-M., Zhang, M., Eitzen, Z. A., Ghan, S. J., Klein, S. A., Wu, X., et al. (2005). Modeling springtime shallow frontal clouds with cloud-resolving and single-column models. *Journal of Geophysical Research*, *110*(D15), D15S04. https://doi.org/10.1029/2004JD005153

Yanai, M., Esbensen, S., & Chu, J.-H. (1973). Determination of bulk properties of tropical cloud clusters from large-scale heat and moisture budgets. *Journal of the Atmospheric Sciences*, *30*(4), 611–627. https://doi.org/10.1175/1520-0469(1973)030<0611:DOBPOT>2.0.CO;2

Zeng, X., Atlas, R., Birk, R. J., Carr, F. H., Carrier, M. J., Cucurull, L., et al. (2020). Use of observing system simulation experiments in the U.S. *Bulletin America Meteorology Social*, *101*(8), E1427–E1438. https://doi.org/10.1175/BAMS-D-19-0155.1

Zeng, X., Tao, W.-K., Matsui, T., Xie, S., Lang, S., Zhang, M., et al. (2011). Estimating the ice crystal enhancement factor in the tropics. *Journal of the Atmospheric Sciences*, *68*(7), 1424–1434. https://doi.org/10.1175/2011JAS3550.1

Zeng, X., Tao, W.-K., Powell, S. W., Houze, R. A., Jr., Ciesielski, P., Guy, N., et al. (2013). A comparison of the water budgets between clouds from AMMA and TWP-ICE. *Journal of the Atmospheric Sciences*, *70*(2), 487–503. https://doi.org/10.1175/jas-d-12-050.1

Zhang, J., Howard, K., Langston, C., Kaney, B., Qi, Y., Tang, L., et al. (2016). Multi-radar multi-sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities. *Bulletin America Meteorology Social*, *97*(4), 621–638. https://doi.org/10.1175/BAMS-D-14-00174.1

Zhang, M., & Lin, J. (1997). Constrained variational analysis of sounding data based on column-integrated budgets of mass, heat, moisture, and momentum: Approach and application to ARM measurements. *Journal of the Atmospheric Sciences*, *54*(11), 1503–1524. https://doi.org/10.1175/1520-0469(1997)054<1503:CVAOSD>2.0.CO;2

Zhang, M., Lin, J., Cederwall, R. T., Yio, J. J., & Xie, S. C. (2001). Objective analysis of ARM IOP data: Method and sensitivity. *Monthly Weather Review*, *129*(2), 295–311. https://doi.org/10.1175/1520-0493(2001)129<0295:OAOAID>2.0.CO;2