

Training warm-rain bulk microphysics schemes using super-droplet simulations

Sajjad Azimi¹, Anna Jaruga¹, Emily de Jong², Sylwester Arabas³, Tapio
Schneider^{1,4}

¹Department of Environmental Science and Engineering, California Institute of Technology, Pasadena,
CA, USA

²Department of Mechanical and Civil Engineering, California Institute of Technology, Pasadena, CA, USA

³Faculty of Physics and Applied Computer Science, AGH University of Krakow, Kraków, Poland

⁴Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA

Key Points:

- A calibration framework for warm-rain bulk microphysics parameterizations is presented.
- The framework relies on a library of super-droplet simulations of a rain shaft.
- Calibrating a single-moment microphysics scheme with the calibration framework substantially reduces the model-data mismatch.

Corresponding author: Sajjad Azimi, azimi@caltech.edu

Abstract

Cloud microphysics is a critical aspect of the Earth’s climate system, which involves processes at the nano- and micrometer scales of droplets and ice particles. In climate modeling, cloud microphysics is commonly represented by bulk models, which contain simplified process rates that require calibration. This study presents a framework for calibrating warm-rain bulk schemes using high-fidelity super-droplet simulations that provide a more accurate and physically based representation of cloud and precipitation processes. The calibration framework employs ensemble Kalman methods including ensemble Kalman inversion (EKI) and unscented Kalman inversion (UKI) to calibrate bulk microphysics schemes with probabilistic super-droplet simulations. We demonstrate the framework’s effectiveness by calibrating a single-moment bulk scheme, resulting in a reduction of data-model mismatch by more than 75% compared to the model with initial parameters. Thus, this study demonstrates a powerful tool for enhancing the accuracy of bulk microphysics schemes in atmospheric models and improving climate modeling.

Plain Language Summary

Cloud microphysics is a complex set of processes that determine the formation and evolution of particles in clouds, which affects the Earth’s climate by regulating precipitation and cloud cover. However, the vast difference in scale between the microphysics and large-scale atmospheric flows makes it impossible to simulate these processes in climate models directly. Instead, climate models use simplified methods to represent cloud microphysics, which can result in inaccuracies. In this study, we focus on calibrating the simplified models with more detailed simulations of cloud microphysics using the super-droplet method. We demonstrate a framework for calibrating the simplified models using high-fidelity simulations, which improves the accuracy of these models.

1 Introduction

Cloud microphysics refers to the microscale processes within clouds that control the formation and evolution of hydrometeors, such as cloud droplets, ice crystals, and raindrops. These processes are essential for regulating many mesoscale properties of clouds, such as precipitation and cloud albedo, which are important factors in the Earth’s climate system. Despite the crucial role of cloud microphysics, climate models cannot resolve these processes mainly due to the vast scale separation between the micro-scale dynamics of hydrometeors and large-scale atmospheric flows. As a result, climate models commonly represent cloud microphysics by representing particle size distributions (PSD) of hydrometeors through bulk methods. Bulk methods track the evolution of aggregate properties of the PSD, such as the total mass or number of particles. While bulk schemes are the dominant numerical approach in climate modeling, they have significant uncertainty in both the structure of the model and the parameters (Khain et al., 2015; Morrison et al., 2019; Igel et al., 2022). However, the uncertainty in the parameters can be reduced through calibration against more detailed methods such as spectral bin methods and particle-based super-droplet methods. In this paper, we will focus on calibrating bulk methods with detailed results of the particle-based super-droplet method to improve the accuracy of climate models.

While bulk methods have the advantage of reducing the computational cost of microphysics modeling, their accuracy is challenged by several factors. First, bulk methods follow the evolution of a few moments of the PSD, while many process rates depend on higher moments. Therefore, the bulk methods require closures that express higher moments in terms of the tracked moments. These closures are typically derived by assuming specific functional forms for the size distribution, such as a gamma or exponential distribution (e.g., Khairoutdinov & Kogan, 2000; Liu & Daum, 2004; Seifert & Beheng, 2006; Morrison & Grabowski, 2007). However, in reality, the size distribution of

66 hydrometeors is multimodal. Consequently, bulk methods consider different particle cat-
67 egories, such as cloud droplets and raindrops, each represented by different unimodal dis-
68 tributions. The conversion rate between these categories is parameterized, leading to in-
69 creased uncertainties in climate modeling. Furthermore, the use of multiple categories
70 is an artificial representation of the real-world physics of hydrometeors, and the conver-
71 sion rates may not be able to capture the collective physics of hydrometeors well. Sec-
72 ond, parameterization models in bulk schemes typically include several process rate pa-
73 rameters that need to be calibrated with reference physics, which can be observational
74 data or high-fidelity numerical simulations. However, despite the abundance of satellite
75 observations available, it remains challenging to leverage them effectively for the devel-
76 opment of microphysics schemes due to the difficulties in accurately mapping from these
77 observations to microphysical variables (Morrison et al., 2020).

78 Despite their limitations, bulk methods are widely used in climate modeling due
79 to their simplicity, motivating researchers to continually develop new parameterizations
80 to improve their accuracy (e.g., Kessler, 1969; Tripoli & Cotton, 1980; Milbrandt & Yau,
81 2005; Morrison & Milbrandt, 2015; Morrison et al., 2019). The complexity of bulk meth-
82 ods varies depending on the number of prognostic moments they track. While most cloud
83 microphysics schemes only describe one or two moments, higher-moment schemes are more
84 accurate, albeit at increased computational costs. Regardless of the complexity of a new
85 bulk parameterization idea, poorly estimated parameters can impact the performance
86 of the entire modeling system. Therefore, careful attention must be given to this aspect
87 of bulk method development to ensure that new parameterization ideas are effective and
88 reliable. Several recent studies highlighted the application of Bayesian techniques in pa-
89 rameter estimation for bulk microphysics schemes. Posselt and Vukicevic (2010) and Posselt
90 (2016) employed a Markov chain Monte Carlo algorithm to investigate the relationship
91 between cloud microphysical parameters and deep moist convection simulations. Morrison
92 et al. (2019) and van Lier-Walqui et al. (2020) introduced the Bayesian observationally
93 constrained statistical-physical scheme, a flexible framework designed to learn microphys-
94 ical parameter distributions through Bayesian inference. Bieli et al. (2022) proposed a
95 bulk microphysics scheme with adjustable complexity, and presented an efficient param-
96 eter learning approach using the calibrate-emulate-sample algorithm (Dunbar et al., 2021).
97 Notably, both of these studies demonstrated learning parameters of their bulk schemes
98 by using perfect-model experiments with data generated by the same models.

99 Access to microphysics observations for calibration and validation of bulk schemes
100 is often limited, making high-fidelity simulations using detailed microphysics represen-
101 tations a critical data source. Researchers have commonly used spectral bin methods to
102 calibrate and evaluate bulk schemes (e.g., Khairoutdinov & Kogan, 2000; Kogan, 2013;
103 Zeng & Li, 2020; Gettelman et al., 2021). However, bin methods can be susceptible to
104 numerical diffusion, and - in the case of modeling coalescence - they inherit the limit-
105 ing assumptions necessary to derive the underlying deterministic Smoluchowski equa-
106 tions, both of which limit their accuracy (Grabowski et al., 2019). Another detailed method
107 that has gained increasing attention in recent years is the particle-based super-droplet
108 method (SDM) (Shima et al., 2009; Andrejczuk et al., 2010; Riechelmann et al., 2012).
109 This method uses a probabilistic particle-based approach to track individual super-droplets
110 explicitly and allows for a more realistic representation of the microphysics involved in
111 cloud and precipitation processes. Each super-droplet is treated as an ensemble of ac-
112 tual particles that share the same attributes, such as size, composition and location. SDM
113 simulations are probabilistic because they involve random sampling of the attribute space
114 at initialization and feature Monte-Carlo-type representation of stochastic processes such
115 as coagulation and breakup. Each SDM simulation yields a single realization of the sys-
116 tem evolution which includes tracking of each super-droplet’s properties through par-
117 ticle processes such as aerosol activation, condensation, evaporation, collision, coalescence,
118 and break-up. Unlike bulk schemes that require parameterizations of conversion rates
119 between artificial categories, the SDM avoids such parameterizations, providing a more

120 accurate and physically based representation of cloud and precipitation processes. As
121 such, the particle-based super-droplet approach has the potential to provide more real-
122 istic and detailed data for improving the accuracy of bulk schemes in simulating cloud
123 and precipitation processes. Noh et al. (2018) employed the particle-based super-droplet
124 approach to evaluate several bulk parameterizations for collisional growth in shallow cu-
125 mulcus clouds. However, their study is limited to few simulations initialized with a sin-
126 gles thermodynamic condition and excludes considerations of raindrop breakup and evap-
127 oration.

128 Here in this study, we present a framework for calibrating warm-rain bulk schemes
129 using high-fidelity super-droplet simulations. We implement the one-dimensional kine-
130 matic driver (KiD-1d) model, proposed by Shipway and Hill (2012), and generate a li-
131 brary of super-droplet simulations in this model. The KiD-1d model is a one-dimensional
132 warm rain shaft model with a prescribed flow field and constant temperature profile. The
133 flow and temperature fields are prescribed to isolate microphysics processes from their
134 feedbacks with dynamics and thermodynamics, enabling us to calibrate and validate mi-
135 crophysics schemes dynamically consistently. This means that any variations in the re-
136 sults can only be attributed to changes in microphysics schemes. We utilize ensemble
137 Kalman methods, including ensemble Kalman inversion (EKI) (Iglesias et al., 2013) and
138 unscented Kalman inversion (UKI) (Huang et al., 2022), to calibrate bulk microphysics
139 schemes with the super-droplet simulations. EKI and UKI are ensemble-based gradient-
140 free methods that have demonstrated remarkable success in a wide variety of calibration
141 studies (e.g., Xiao et al., 2016; Kovachki & Stuart, 2019; Dunbar et al., 2022). EKI is
142 more robust than UKI concerning noise in observations, while UKI provides parameter
143 uncertainties and allows for model error quantification (Lopez-Gomez et al., 2022). We
144 demonstrate the application of the calibration framework by calibrating a single-moment
145 warm-rain bulk scheme, targeting parameters of conversion rates such as condensation,
146 auto-conversion, accretion, sedimentation, and evaporation rates. Remarkably, calibra-
147 tions using EKI and UKI obtain two different sets of optimal parameters, both result-
148 ing in a similar reduction of model-data mismatch. The difference between these two pa-
149 rameter sets is consistent with the parameter correlations obtained from UKI. Through
150 our calibration process, we achieve a significant enhancement in the accuracy of the bulk
151 model by more than 75% compared to the model with initial parameter values.

152 The calibration framework presented here has several notable properties compared
153 to previous studies. First, we employ the SDM as a tool capable of providing a physi-
154 cally based representation of microphysics for generating benchmark simulations. Sec-
155 ond, the framework offers an efficient setup to calibrate and evaluate bulk methods by
156 using a diverse set of rain shaft simulations with a wide variety of precipitation condi-
157 tions. Finally, by using ensemble Kalman methods, which are gradient-free, we ensure
158 both efficient parameter learning and the ability to quantify parameter uncertainties and
159 model error. The calibration framework presented in this study provides a promising tool
160 for enhancing the accuracy of bulk microphysics schemes in atmospheric models, with
161 potential implications for improving climate modeling.

162 The manuscript is organized as follows: Section 2 provides an overview of the KiD-
163 1d model, along with a discussion of the SDM used to generate simulations of the KiD-
164 1d model. The section also describes the calibration methods employed in our framework
165 for calibrating bulk schemes. In Section 3, we present a library of super-droplet simu-
166 lations of the KiD-1d model and report the results of calibrating a single-moment bulk
167 scheme using this library of rain shaft simulations. Finally, Section 4 summarizes our
168 findings and provides an outlook for future research.

Table 1. Data points for interpolating the initial water vapor mixing ratio $r_{v,0}$ and potential temperature θ .

Height (m)	$r_{v,0}$ (kg kg ⁻¹)	θ (K)
0	0.015	297.9
740	0.0138	297.9
3260	0.0024	312.66

2 Methods

This section provides an overview of the methods employed in this study. We introduce the one-dimensional rain-shaft model, which serves as a testbed for calibrating and evaluating warm-rain bulk schemes in relation to high-fidelity particle-based simulations. Subsequently, we discuss the SDM utilized to generate a comprehensive library of simulations for benchmarking bulk schemes. Next, we present a specific example of a single-moment warm-rain bulk scheme used to demonstrate the application of the calibration framework. Lastly, we explain the calibration methods employed to refine and optimize the bulk scheme.

2.1 System: One-dimensional kinematic driver model

The calibration framework utilizes an implementation of the one-dimensional kinematic driver (KiD-1d) model as a testbed for calibrating and evaluating warm-rain bulk schemes. The KiD-1d model is specifically designed to facilitate the assessment of microphysics parameterizations by prescribing both the velocity and temperature fields (Shipway & Hill, 2012; Hill et al., 2023). This prescription effectively prevents any feedback from dynamics and thermodynamics on microphysics processes, ensuring that observed variations in the results can be solely attributed to changes in microphysics parameterizations. In the employed implementation of the KiD-1d model, we consider a stratified air density profile, and thus prescribe the flow by using an air momentum profile, unlike (Hill et al., 2023) where a constant density is used.

The KiD-1d model represents shallow convection in a column of moist air within a height range of 3 km from the ground level. The prescribed flow field represents an updraft, which is uniform in height z and sinusoidal in time t , as given by the equation:

$$\rho w(z, t) = (\rho w)_0 \sin(\pi t/t_1), \quad 0 < t < t_1. \quad (1)$$

Here, ρ represents the dry-air density, w denotes the vertical velocity component, and $(\rho w)_0$ is the maximum updraft momentum. The parameter t_1 represents the duration of the updraft. Beyond t_1 , there is no updraft, and ρw remains at 0. This updraft motion lifts moist air to higher, colder levels, facilitating condensation of water vapor and cloud formation. The initial vapor mixing ratio $r_{v,0}$ and the potential temperature θ are represented as piecewise linear profiles interpolated from data points provided in Table 1. The initial temperature profile $T(z)$ at $t = 0$ is computed from the potential temperature $\theta(z)$ and is held constant throughout the simulations to eliminate any potential thermodynamics feedback on microphysics.

2.2 Particle-based simulation method

For generating a library of particle-based simulations of the KiD-1d model, we use the PySDM package (Bartman, Bulenok, et al., 2022; de Jong et al., 2023). PySDM is a Python-based code designed to run particle-based simulations of clouds and precipitation using super-droplets. Each super-droplet corresponds to multiple particles shar-

206 ing the same properties, including location, size and composition. The multiplicity of a
 207 super-droplet indicates the number of actual particles it represents. For further details
 208 on the models employed in PySDM, refer to Bartman, Bulenok, et al. (2022) and de Jong
 209 et al. (2023).

210 Because the particle-based simulations are inherently stochastic, we generate 100
 211 simulations for each configuration to determine the mean and variability of the results
 212 used for calibration purposes. In each simulation, we utilize an average of $N_{sd} = 512$
 213 super-droplets per grid box, with a grid spacing of $dz = 50$ m and a time step of $dt =$
 214 5 s for advection computations. The Python-based code PyMPDATA (Bartman, Banašiewicz,
 215 et al., 2022) is used for solving the advection equation. We study the independence of
 216 the results from the chosen numerical values by performing simulations with doubled N_{sd} ,
 217 halved dz , and halved dt . The results from these simulations show excellent agreement
 218 with the original findings, indicating that the results are not influenced by the specific
 219 numerical values employed. (For more detailed information, see Appendix A.)

220 2.3 Single-moment warm-rain bulk scheme

221 To demonstrate the application of our calibration framework, we focus on calibrat-
 222 ing and evaluating a single-moment warm-rain bulk scheme. Specifically, we examine the
 223 single-moment bulk scheme implemented in CloudMicrophysics.jl, an open-source Ju-
 224 lia package developed and utilized within the CliMA project (clima.caltech.edu). This
 225 bulk scheme is based on the original concept introduced by Kessler (1969). It divides the
 226 total water content into three categories: water vapor, cloud water, and rainwater. The
 227 conversion of water vapor into cloud water occurs through condensation. The conver-
 228 sion of cloud water to rainwater involves two processes: auto-conversion, accounting for
 229 the collision and coalescence of droplets in the cloud phase to form raindrops, and ac-
 230 cretion, representing the collection of cloud droplets by raindrops. The sedimentation
 231 of raindrops causes them to descend to subsaturated regions, leading to the partial con-
 232 version of rainwater back into water vapor through evaporation.

233 The auto-conversion rate is represented as the ratio of the specific content of cloud
 234 water to the auto-conversion time scale. This time scale is determined by a power-law
 235 function of the initial aerosol number density (N_a). The auto-conversion rate is expressed
 236 as follows:

$$\frac{\partial q_r}{\partial t} \Big|_{acnv} = - \frac{\partial q_c}{\partial t} \Big|_{acnv} = \frac{q_c}{\tau_{acnv,0} \left(\frac{N_a}{100 \text{ cm}^{-3}} \right)^{\alpha_{acnv}}}. \quad (2)$$

237 In this equation, q_c and q_r represent the specific content of cloud and rainwater, respec-
 238 tively. The constant $\tau_{acnv,0}$ denotes the reference auto-conversion time scale, and α_{acnv}
 239 represents the power law parameter of the number density.

240 The process rate equations provided in the CloudMicrophysics.jl package are based
 241 on the following assumptions regarding the raindrop size distribution n , mass m , area
 242 a , and terminal velocity v as functions of the particle radius r :

$$n(r) = n_0 \exp(-\lambda r) \quad (3)$$

$$m(r) = m_0 \left(\frac{r}{r_0} \right)^3 \quad (4)$$

$$a(r) = \chi_a a_0 \left(\frac{r}{r_0} \right)^{2+\Delta_a} \quad (5)$$

$$v(r) = \chi_v v_0 \left(\frac{r}{r_0} \right)^{1/2+\Delta_v}, \quad (6)$$

243 where r_0 denotes the reference raindrop radius used for nondimensionalization. The val-
 244 ues of the reference raindrop mass m_0 , area a_0 , and terminal velocity v_0 are calculated

245 as follows: $m_0 = (4/3) \pi \rho_w r_0^3$, $a_0 = \pi r_0^2$, and

$$v_0 = \left(\frac{8(\rho_w/\rho_m - 1)gr_0}{3C_d} \right)^{1/2}. \quad (7)$$

246 Here, ρ_w represents the density of water, ρ_m is the moist-air density, g denotes the ac-
 247 celeration due to gravity, and C_d is a constant drag coefficient. The coefficients χ_a , Δ_a ,
 248 χ_v , and Δ_v are free parameters that can be adjusted during model calibration. The pa-
 249 rameters n_0 and λ serve as distribution parameters. Integrating the mass of particles over
 250 the distribution, we obtain the following equation for λ :

$$\lambda = \left(\frac{4\pi\rho_w n_0 \Gamma(4)}{3q_r \rho_m} \right)^{\frac{1}{4}}, \quad (8)$$

251 where Γ denotes the gamma function. The condensation of water vapor is modeled by
 252 relaxing the excess of water vapor towards the saturation specific humidity over the con-
 253 densation time scale:

$$\left. \frac{dq_c}{dt} \right|_{cond} = \frac{q_v - q_v^*}{\tau_{cond}}, \quad (9)$$

254 where q_v represents the specific humidity, q_v^* is the saturation specific humidity, and τ_{cond}
 255 represents the time scale of condensation. The accretion rate is obtained by integrating
 256 the rate of collection of cloud droplets by raindrops while falling at their terminal ve-
 257 locity over the assumed raindrop size distribution. It is expressed as follows:

$$\left. \frac{dq_r}{dt} \right|_{accr} = - \left. \frac{dq_c}{dt} \right|_{accr} = n_0 \Pi_{a,v} q_c E_{cr} \Gamma(\Sigma_{a,v} + 1) \frac{1}{\lambda} \left(\frac{1}{r_0 \lambda} \right)^{\Sigma_{a,v}}, \quad (10)$$

258 where $\Pi_{a,v} = a_0 v_0 \chi_a \chi_v$, and $\Sigma_{a,v} = 5/2 + \Delta_a + \Delta_v$. Additionally, E_{cr} represents the
 259 collision efficiency between cloud droplets and raindrops. The sedimentation of rain is
 260 accounted for by the following equation, which describes the terminal velocity:

$$v_t = \chi_v v_0 \left(\frac{1}{r_0 \lambda} \right)^{1/2 + \Delta_v} \frac{\Gamma(9/2 + \Delta_v)}{\Gamma(4)}. \quad (11)$$

261 Finally, the rate of rain evaporation is modeled by integrating the evaporation of indi-
 262 vidual particles over the spectrum of raindrops. This leads to the following expression:

$$\begin{aligned} \left. \frac{dq_r}{dt} \right|_{evap} &= \frac{4\pi n_0}{\rho_m} (S - 1) G(T) \lambda^{-2} \\ &\times \left[a_{vent} + b_{vent} \left(\frac{\nu_a}{D_v} \right)^{\frac{1}{3}} \left(\frac{1}{r_0 \lambda} \right)^{\frac{\nu_a + \Delta_v}{2}} \left(\frac{2\chi_v v_0}{\nu_a \lambda} \right)^{\frac{1}{2}} \Gamma \left(\frac{11}{4} + \frac{\Delta_v}{2} \right) \right]. \end{aligned} \quad (12)$$

263 In this equation, $S = q_v/q_v^*$ represents the saturation, T denotes the temperature, D_v
 264 is the diffusivity of water vapor, ν_a is the kinematic viscosity of air, and a_{vent} and b_{vent}
 265 are ventilation parameters. The function $G(T)$ is defined as:

$$G(T) = \left(\frac{L}{kT} \left(\frac{L}{R_v T} - 1 \right) + \frac{R_v T}{p_v^* D_v} \right)^{-1} \quad (13)$$

266 where L is the latent heat of vaporization, k is the thermal conductivity of air, R_v is the
 267 gas constant of water vapor, and p_v^* represents the saturation vapor pressure.

268 The single-moment bulk scheme considered in this study involves several notable
 269 simplifications. First, the functional form of the auto-conversion parameterization is straight-
 270 forward, representing it as the ratio of available cloud water to an auto-conversion time
 271 scale. Second, the scheme assumes that the distribution of raindrops follows an expo-
 272 nential distribution, characterized by a constant scaling parameter n_0 . Third, in the pa-
 273 rameterization of terminal velocity, a constant drag coefficient is employed, which is as-
 274 summed to apply uniformly to all particles, while in reality, the drag coefficient is a func-
 275 tion of raindrop size. Finally, the scheme adopts a constant collision efficiency in the pa-
 276 rameterization of accretion rate. These simplifications, while enhancing computational
 277 efficiency, can affect the model's performance.

Table 2. Parameters of the single-moment bulk scheme. The columns show parameter names, brief descriptions, and prior values with references. The references are KM2003 (Korolev & Mazin, 2003), GS1996 (Grabowski & Smolarkiewicz, 1996), and LD2004 (Liu & Daum, 2004), MP1948 (Marshall & Palmer, 1948), and G1998 (Grabowski, 1998). Note that the values of a_{vent} and b_{vent} are selected to achieve a close agreement with the evaporation rate of GS1996 at a specific humidity of 15 g/kg and $T = 288$ K. Additionally, the value of C_d is selected to closely approximate the terminal velocity of GS1996.

Parameter name	Description	Value
τ_{cond}	Condensation time scale	10 s, KM2003
$\tau_{acnv,0}$	Auto-conversion time scale	1000 s, GS1996
α_{acnv}	Auto-conversion coefficient	1, LD2004
χ_v	Terminal velocity coefficient	1
Δ_v	Terminal velocity coefficient	0
χ_a	Accretion coefficient	1
Δ_a	Accretion coefficient	0
a_{vent}	Evaporation coefficient	1.5, GS1996
b_{vent}	Evaporation coefficient	0.53, GS1996
r_0	Reference raindrop radius	10^{-3} m
n_0	Size distribution parameter	$16 \cdot 10^6$ m $^{-4}$, MP1948
C_d	Raindrop drag coefficient	0.55, GS1996
E_{cr}	Collision efficiency	0.8, G1998

278 Table 2 provides a list of parameters of the single-moment bulk method, along with
 279 their prior values. We select a subset of the parameters for calibration, specifically fo-
 280 cusing on those that do not have easily definable physical values. These choices aim to
 281 comprise a set of parameters that uniquely govern auto-conversion, accretion, the ter-
 282 minal velocity of raindrops, and the rain evaporation rate. To ensure coverage of these
 283 processes, we selected one or two parameters from each process, each capable of signif-
 284 icantly modifying that specific process. Specifically, we select τ_{cond} to represent the con-
 285 densation process, $\tau_{acnv,0}$ and α_{acnv} for auto-conversion, χ_v and Δ_v for raindrop termi-
 286 nal velocity, and χ_a and Δ_a for accretion. Additionally, we include b_{vent} to regulate the
 287 rate of evaporation. Other parameters of the model that are modulated by the calibrated
 288 parameters remain constant during model calibration.

289 2.4 Algorithms for learning parameters

290 The problem of learning parameters for the bulk method is formulated as an in-
 291 verse problem, represented by the equation

$$y = \mathcal{H} \circ \Psi \circ \mathcal{T}^{-1}(\theta) + \delta + \eta. \quad (14)$$

292 Here, y represents the vector of observations, and θ represents the vector of learnable
 293 parameters, which are transformed into an unconstrained space $\theta \in \mathbb{R}^p$. The operator
 294 \mathcal{T} is a transformation map that converts parameters ϕ from their constrained subspace
 295 (where they satisfy constraints such as positivity) to the unconstrained space, such that
 296 $\theta = \mathcal{T}(\phi)$. The mapping Ψ represents the dynamical model, while \mathcal{H} denotes the ob-
 297 servational map incorporating necessary post-processing operations to generate model
 298 predictions aligned with the observations. For example, y may represent averaged spe-
 299 cific water content data from particle-based simulations, Ψ represents bulk scheme sim-
 300 ulation results, and \mathcal{H} could involve spatial and temporal averaging. The observational

301 noise associated with the observations y is indicated by η , and the model error by δ . Both
 302 η and δ are assumed to follow a Gaussian distribution with zero mean.

303 To ensure the generalizability of the calibrated model, we train it using multiple
 304 system configurations. We refer to the set of system configurations used for model train-
 305 ing as C . In this study, $|C| = 49$ configurations are used for the calibrations. The ob-
 306 servation vector y consists of observations obtained from all system configurations: $y =$
 307 $[y_1, y_2, \dots, y_{|C|}]^T$. For each system configuration, 100 SDM simulations are conducted, and
 308 the mean values of specific contents of cloud water, rainwater, and water vapor over in-
 309 tervals of 100 m and 10 min are extracted. The data are then normalized by dividing each
 310 field by the maximum of its standard deviation across the 100 simulations. Subsequently,
 311 the observation vector y_c and the noise covariance Γ_c are computed for each configura-
 312 tion c using the normalized data obtained from the 100 SDM simulations.

313 To calibrate the parameters of the bulk scheme using particle-based simulations,
 314 we employ two gradient-free algorithms available in the EnsembleKalmanProcesses.jl pack-
 315 age: ensemble Kalman inversion (EKI) (Iglesias et al., 2013) and unscented Kalman in-
 316 version (UKI) (Huang et al., 2022). These algorithms are derived from the extended Kalman
 317 filter and heavily rely on Gaussian conditioning. EKI utilizes an iterative procedure to
 318 search for the optimal parameter set (maximum likelihood estimator, MLE) by updat-
 319 ing an ensemble of J parameter sets with $J \sim p$. For our calibrations, we choose $J =$
 320 20 . The initial ensemble is formed by randomly sampling parameters from a Gaussian
 321 distribution. On the other hand, UKI adopts a deterministic approach to update an ini-
 322 tial Gaussian estimate represented by an ensemble of $J = 2p + 1$ parameter sets, aim-
 323 ing to approximate the likelihood centered around the MLE. EKI shows greater robust-
 324 ness against observation noise than UKI, while UKI quantifies model error and estimates
 325 parameter uncertainties. For a detailed discussion on both algorithms, refer to Lopez-
 326 Gomez et al. (2022).

327 Training the model involves minimizing the average configuration loss function that
 328 penalizes the mismatch between observations and model outputs. The average config-
 329 uration loss is given by

$$L(\theta; y) = \frac{1}{2|C|} \sum_{c=1}^{|C|} \|y_c - \mathcal{H}_c \circ \Psi_c \circ \mathcal{T}^{-1}(\theta)\|_{\Gamma_c}^2, \quad (15)$$

330 where $\|\cdot\|_{\Gamma_c}$ represents the Mahalanobis norm, with $\|\cdot\|_{\Gamma_c}^2 = \langle \cdot, \Gamma_c^{-1} \cdot \rangle$. Both EKI and
 331 UKI require evaluating the loss at each iteration, which involves running the model for
 332 all configurations. However, this can be computationally expensive. To address this, we
 333 employ mini-batches of configurations denoted as $B \subset C$ to approximate the average
 334 configuration loss:

$$L(\theta; y_B) = \frac{1}{2|B|} \sum_{c \in B} \|y_c - \mathcal{H}_c \circ \Psi_c \circ \mathcal{T}^{-1}(\theta)\|_{\Gamma_c}^2. \quad (16)$$

335 Batching is a commonly used technique that helps prevent convergence to local minima
 336 and thus improves generalization (Li et al., 2014). For our study, we choose a batch size
 337 of $|B| = 6$ for running the calibrations. During model training, EKI and UKI receive
 338 data from a mini-batch of $|B|$ configurations at each iteration. The mini-batches are ran-
 339 domly drawn without replacement from the set of training configurations C . An epoch
 340 corresponds to a complete cycle through all available configurations such that no other
 341 mini-batch can be composed of the remaining configurations. At the end of each epoch,
 342 the configurations are reshuffled. With $|C| = 49$ and $|B| = 6$, each epoch consists of
 343 8 iterations.

3 Results and discussion

In this section, first, we discuss a library of particle-based simulations of the KiD-1d model for different system configurations. Then, we continue by demonstrating the calibration of the single-moment bulk scheme using the library of particle-based simulations as a benchmark.

3.1 Library of rain shaft simulations

We have generated a library of KiD-1d model simulations using the super-droplet method (SDM). This library includes simulations with varying values of the updraft amplitude ($(\rho w)_0$), initial aerosol number density (N_a), and ground-level pressure (p_0). The updraft amplitude ranges from $1.0 \text{ kg m}^{-2} \text{ s}^{-1}$ to $4.0 \text{ kg m}^{-2} \text{ s}^{-1}$ in increments of $0.5 \text{ kg m}^{-2} \text{ s}^{-1}$. The initial aerosol number density takes seven values ranging from $N_a = 10 \text{ cm}^{-3}$ to $N_a = 1000 \text{ cm}^{-3}$, all corresponding to concentration at standard temperature and pressure conditions for dry air. Simulations are conducted for five different surface air pressures, ranging from $p_0 = 988 \text{ hPa}$ to $p_0 = 1012 \text{ hPa}$ in increments of 6 hPa . Each increment in air pressure corresponds to an approximate increase of 0.5 K in the prescribed temperature profile, which impacts the cloud condensate profile. For each combination of variables, we produce 100 simulations to compute the average and variability of the results.

By varying the values of the updraft speed and surface pressure, we can influence the amount of condensed cloud water and, consequently, the precipitation. Additionally, changing the initial aerosol number density influences the collision and coalescence of droplets, thereby influencing the formation of rain (Tao et al., 2012). The selection of different values for these control parameters allows us to generate various rain formation conditions. This variety is crucial for providing the calibration process with diverse training data, thus enhancing the generalizability of the trained model.

In addition to the simulations of the KiD-1d model involving all processes, we conducted additional simulations where the collision and coalescence processes were excluded. These simulations, referred to as condensation-only cases, do not result in rain formation as droplets do not grow large enough to sediment through condensation alone. We performed these simulations with the intention of using them as a reference to evaluate the numerical advection of ambient moisture and the condensation scheme of the bulk model separately from other process parameterizations. Figure 1 (left panels) illustrates an example simulation of a condensation-only case. The figure shows the height-time contours of the cloud water and rainwater specific content, as well as the cloud water path (CWP), rainwater path (RWP) and surface rain rate (RR) over time. The cloud water path and rainwater path represent the total amount of cloud water and rainwater in a column of moist air per unit area, respectively. As is evident in Figure 1 (left panels), condensation primarily occurs within the first ten minutes of the simulation ($t < t_1$) when the updraft speed is non-zero. After $t_1 = 10 \text{ min}$, no rainwater forms as collision and coalescence processes are not considered, and the cloud water is preserved.

When collision and coalescence processes are involved, formation of raindrops is observed. We use a fixed radius threshold of $50 \mu\text{m}$ to differentiate raindrops from cloud droplets for simulation output analysis. We found the sensitivity of the results to the exact value of this threshold to be insignificant. Figure 1 (right panels) illustrates the generation of rain in the simulation of the KiD-1d model with the inclusion of rain production through particle collision and coalescence. Following the coalescence of particles and the formation of raindrops, the raindrops descend due to sedimentation, moving below the cloud base where water vapor is not saturated. Consequently, rain evaporation occurs, resulting in only some of the initial rain water reaching the surface.

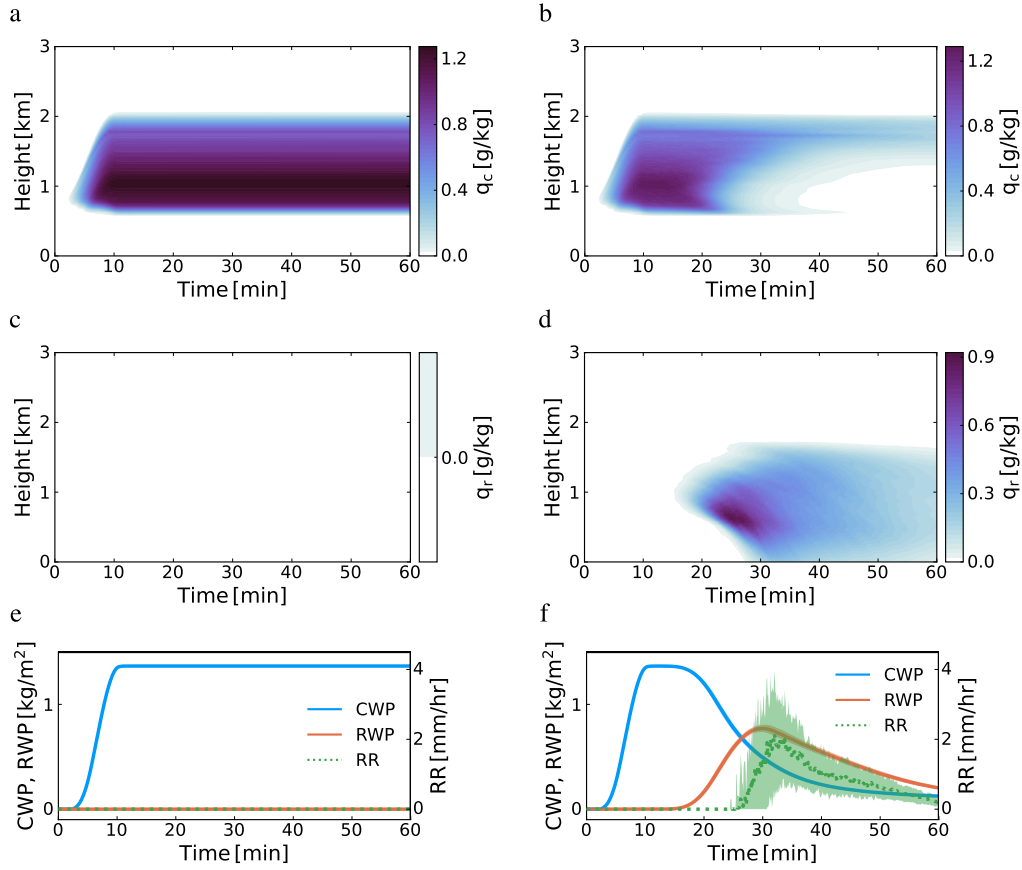


Figure 1. Simulations of the KiD-1d model using the SDM, both without (left panels) and with (right panels) the inclusion of rain production through particle collision and coalescence. The simulations employ an updraft momentum amplitude of $3 \text{ kg m}^{-2} \text{ s}^{-1}$ and an initial aerosol number density of 100 cm^{-3} . Height-time contours for the average specific cloud water content, q_c (a and b), as well as the average specific rainwater content, q_r (c and d) are shown. Panels e and f illustrate the evolution over time of cloud water path, rainwater path, and surface rain rate. In panel f, variations in the graphs are represented by shading, indicating one standard deviation above and below the mean.

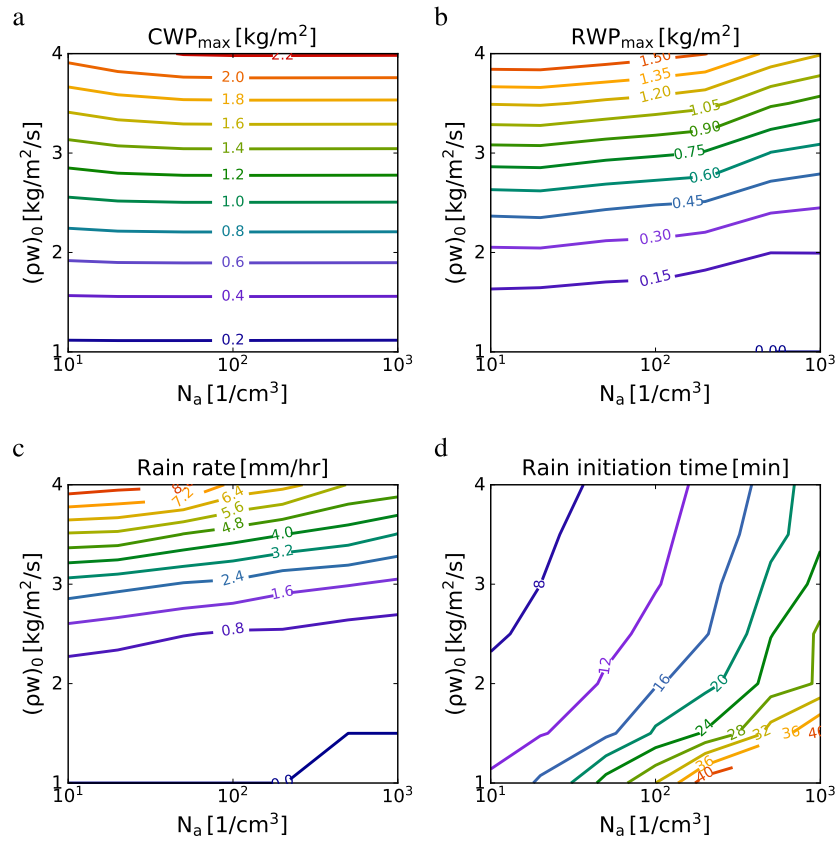


Figure 2. Sensitivity of outputs from the KiD-1d model using the SDM to varying updraft momentum amplitude and initial aerosol number density. The panels display contours of (a) the maximum cloud water path CWP_{max} , (b) the maximum rainwater path RWP_{max} , (c) the maximum surface rain rate, and (d) the rain initiation time. The results are averaged over 100 simulations.

393 Varying the updraft speed and aerosol number density impacts the simulation re-
 394 sults by influencing the availability of water vapor for condensation and the number of
 395 particles contributing to rain formation through collision and coalescence. Figure 2 pro-
 396 vides a visual representation of how updraft speed and aerosol concentration changes af-
 397 fect various properties in the KiD-1d simulations with a fixed surface pressure of $p_0 =$
 398 1000 hPa. Increasing the updraft speed enhances supersaturation at any given altitude
 399 by advecting more water vapor content upward. This heightened supersaturation leads
 400 to the condensation of more cloud water. Consequently, the increased availability of con-
 401 densed cloud water results in greater amount of rain. This is evidenced in Figures 2a and
 402 2b, where both the maximum cloud water path and rainwater path increase with higher
 403 updraft amplitudes. Additionally, Figure 2b demonstrates that an increase in aerosol con-
 404 centration leads to a decrease in the maximum rainwater path. This is due to a higher
 405 number of particles available to carry the same amount of water, resulting in the forma-
 406 tion of smaller droplets. The formation of smaller droplets reduces the likelihood of col-
 407 lision and coalescence, consequently decreasing rain production and surface rain rate (Fig-
 408 ure 2c). The cloud water path (Figure 2a) remains relatively unaffected by aerosol num-
 409 ber density, except for low values of N_a , where simulations with higher N_a yield more
 410 cloud water at high updraft amplitudes. This observation suggests that insufficient aerosols
 411 in the system may delay condensation due to the limited capacity to carry a high vol-
 412 ume of cloud water.

413 Figure 2d illustrates the variations in rain initiation time with changing updraft
 414 amplitude and aerosol number density. The rain initiation time is defined as the time
 415 at which the specific rainwater content surpasses a chosen small threshold ($q_r = 10^{-8}$
 416 g kg^{-1}). Both increasing the updraft amplitude and decreasing the aerosol number den-
 417 sity result in an earlier rain initiation time. Generally, higher updraft amplitudes and
 418 lower aerosol number densities lead to earlier and more substantial rain formation. Note
 419 that similar behavior can be observed at other surface pressures, with less rain observed
 420 for higher surface pressures. These observations highlight the sensitivity of rain forma-
 421 tion to the values of updraft speed and aerosol number density, suggesting that the mi-
 422 crophysical processes governing rain formation are susceptible to certain parameters. These
 423 findings are consistent with the results of Hill et al. (2023), where they demonstrate the
 424 high sensitivity of rain initiation time and amount to specific parameters and different
 425 super-droplet implementations.

426 The simulations conducted with the KiD-1d model using the SDM serve as a bench-
 427 mark for calibrating warm-rain bulk microphysics schemes. This dataset encompasses
 428 a wide range of precipitation conditions, from instances with no rain formation to those
 429 with substantial rainfall, with a maximum rainwater path exceeding 1.6 kg m^{-2} for $p_0 =$
 430 1000 hPa. The observed sensitivities of cloud water content, rain initiation time, and rain-
 431 water content suggest that the dataset represents diverse rates for microphysics processes,
 432 including condensation, auto-conversion, accretion, and rain evaporation. We anticipate
 433 that these sensitivities greatly contribute to the generalizability and effectiveness of the
 434 calibrated bulk microphysics schemes. However, it's important to note that the decou-
 435 pling of microphysics from dynamics, particularly ignoring turbulence effects on collision-
 436 coalescence processes, is a limitation of this study. This limitation may introduce biases
 437 in the calibration results and negatively impact the performance of calibrated schemes
 438 in more complex setups like large eddy simulations or earth system models.

439 It is worth noting that the simulations in the KiD-1D model are not aimed at ac-
 440 curately representing the complex physics of a real precipitating cloud. Specifically, the
 441 KiD-1d model does not take into account turbulence or temperature fluctuations. Its de-
 442 sign isolates microphysics from dynamics and thermodynamics, allowing for a focused
 443 study of microphysics phenomena. This isolation is crucial to ensure that any variations
 444 observed in the results can be attributed solely to changes in the microphysics schemes

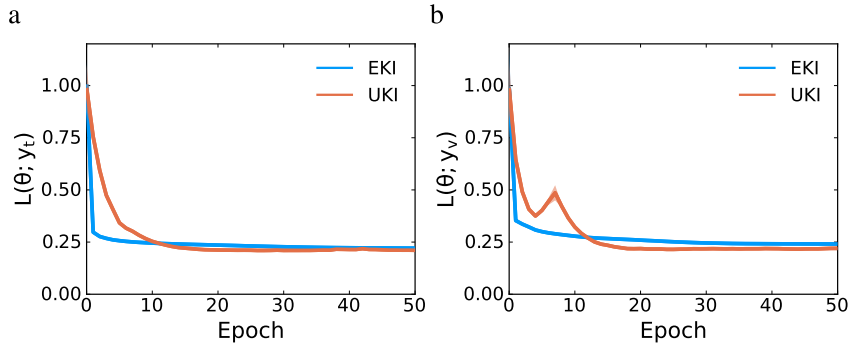


Figure 3. Variations of the loss function during calibration for the training set (a) and validation set (b). Graphs in both panels are normalized by the loss of the model with the initial parameters to allow comparison.

445 being investigated. In the following subsection, we discuss how we employ the results ob-
 446 tained from the SDM simulations to calibrate the single-moment bulk scheme.

447 **3.2 Calibration of a bulk scheme with the library of super-droplet sim-** 448 **ulations**

449 In this subsection, we present the calibration results of the single-moment bulk mi-
 450 crophysics scheme using the library of the SDM simulation results. For training the model,
 451 we use all SDM simulations with varying updraft amplitude and aerosol number den-
 452 sity at the fixed surface pressure of $p_0 = 1000$ hPa. In total, the training set contains
 453 $|C| = 49$ cases. From each case, we extract mean values of specific cloud water content
 454 \bar{q}_c , rainwater content \bar{q}_r , and humidity \bar{q}_v over intervals of 100 m and 10 min to use in
 455 the calibration process.

456 The validation set, on the other hand, is intentionally selected from configurations
 457 at a different ground-level pressure than the training set. This intentional selection al-
 458 lows us to assess whether the calibrated model can effectively capture simulations from
 459 a dataset that is not used for training. Specifically, the validation set consists of simu-
 460 lations performed with the surface pressure $p_0 = 994$ hPa with updraft amplitudes of
 461 $(\rho w)_0 = [2, 3, 4]$ $\text{kg m}^{-2} \text{s}^{-1}$ and aerosol number density $N_a = [50, 200]$ cm^{-3} . It is
 462 worth noting that the lower ground-level pressure of the validation set corresponds to
 463 approximately 0.5 K lower temperature. This leads to higher supersaturation and increased
 464 rain, providing a distinct dataset for validation compared to the training data. As a re-
 465 sult, it is unnecessary to modify the value of the updraft amplitudes and initial aerosol
 466 number densities in the validation set from those used in training.

467 Figure 3 shows the evolution of the configuration-averaged loss during calibrations
 468 for both EKI and UKI, using the training and validation sets. Both EKI and UKI achieve
 469 a reduction of more than 75% in the loss for both the training and validation sets. Al-
 470 though calibration is continued for 50 epochs, loss reduction for both the training and
 471 validation sets mainly occurs within 15 epochs, with EKI reducing the error more rapidly.
 472 Remarkably, the loss reduction for the validation set is almost equal to the reduction for
 473 the training set, indicating that the calibrated model generalizes well to the precipita-
 474 tion conditions in the validation set.

475 Depending on the stochastic initialization of the parameter ensemble for EKI, EKI
 476 and UKI may converge to different sets of parameter values that minimize the mismatch
 477 between bulk method results and SDM simulations. This is demonstrated in Figure 4,

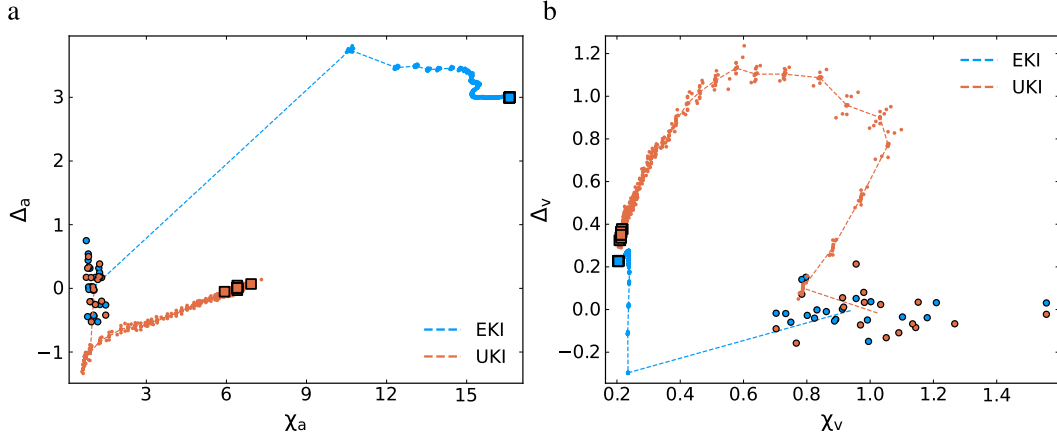


Figure 4. Evolution of (a) the two accretion coefficients, χ_a and Δ_a , and (b) the two terminal velocity coefficients, χ_v and Δ_v , during calibrations for both EKI (dashed blue) and UKI (dashed red). The initial ensembles of parameters are represented by blue circles (EKI) and red circles (UKI), while the final ensembles of parameters are indicated by blue squares (EKI) and red squares (UKI). The final ensemble means for all parameters are given in Table 3.

Table 3. Results of the calibration of the single-moment bulk scheme by EKI and UKI. Columns represent parameter names, the prior parameter values, and the optimal parameter values from EKI and UKI calibrations. The optimal values are obtained by averaging the final ensembles of parameters.

Parameter name	Prior value	EKI optimal value	UKI optimal value
τ_{cond}	10.0 s	39.7 s	35.0 s
$\tau_{acnv,0}$	1.0×10^3 s	13.4×10^3 s	549.1×10^3 s
α_{acnv}	1.0	0.52	2.09
χ_v	1.0	0.205	0.213
Δ_v	0.0	0.228	0.351
χ_a	1.0	16.61	6.41
Δ_a	0.0	3.00	0.01
b_{vent}	0.53	0.98	1.48

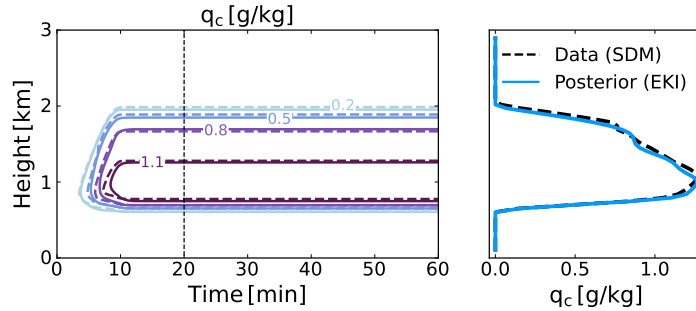


Figure 5. Comparison of the simulations of the KiD-1d model without collision processes using the SDM and the calibrated bulk method by EKI. Updraft amplitude is set to $3 \text{ kg m}^{-2} \text{ s}^{-1}$, and the initial aerosol number density is 100 cm^{-3} . The left panel shows height-time contours of specific cloud water content, while the right panel displays the specific cloud water content at $T = 20 \text{ min}$. The results from both methods, the SDM and the calibrated bulk method, are in excellent agreement showing that the bulk scheme well captures the condensation process.

478 which shows the evolution of the two accretion coefficients, χ_a and Δ_a , as well as the two
 479 terminal velocity coefficients, χ_v and Δ_v , by both EKI and UKI. While EKI and UKI
 480 converge to similar results for χ_v and Δ_v , the evolution of the two accretion coefficients
 481 χ_a and Δ_a during the EKI and UKI calibrations shows significant differences, which indic-
 482 ates the convergence of EKI and UKI towards two distinct sets of parameters. The
 483 evolution of all parameters during the EKI and UKI calibrations is provided in Appendix
 484 B. The final values of all parameters obtained by EKI and UKI are provided in Table 3.

485 The main difference between the two parameter sets obtained by EKI and UKI is
 486 in the auto-conversion and accretion parameters. In the UKI set, the auto-conversion
 487 parameters, that control the auto-conversion time scale, are significantly larger than those
 488 in the EKI set. Consequently, the UKI set predicts larger auto-conversion time scales,
 489 leading to lower auto-conversion rates. However, this change is counterbalanced by the
 490 smaller values of the accretion coefficients, including Δ_a , which governs the exponent of
 491 q_r in the accretion process rate. With a smaller exponent, the accretion process yields
 492 larger rain production rates for small q_r values, particularly in the early stages of rain
 493 production. Thus, larger accretion rates compensate for the smaller auto-conversion rates,
 494 resulting in comparable rain formations in the simulations.

495 Notably, the optimal auto-conversion time scale obtained by both EKI and UKI
 496 are larger than the auto-conversion time scale of 1000 s documented in Grabowski and
 497 Smolarkiewicz (1996). This difference may be attributed to the fact that, unlike Grabowski
 498 and Smolarkiewicz (1996), we do not consider any auto-conversion threshold. Moreover,
 499 the exponent of q_r in the accretion parameterization, equation (10), is close to one for
 500 UKI optimal parameters. This is consistent with bulk schemes of Tripoli and Cotton (1980);
 501 Beheng (1994); Seifert and Beheng (2006). In contrast, the exponent for EKI optimal
 502 parameters is relatively larger. Since both EKI and UKI achieve an approximate 75%
 503 reduction in loss, we accept both sets of parameters as valid calibrations for the bulk method.
 504 Incorporating detailed auto-conversion and accretion rate information in the training data
 505 could provide further insights and help obtain a unique set of optimal parameters.

506 In the simulation of the KiD-1d model, when precipitation processes are not included
 507 (condensation-only case), the only parameterized process is the condensation of water
 508 vapor into cloud water by equation 9. Figure 5 shows results of the simulation of the KiD-
 509 1d model with $(\rho w)_0 = 3 \text{ kg m}^{-2} \text{ s}^{-1}$, $N_a = 100 \text{ cm}^{-3}$ and $p_0 = 1000 \text{ hPa}$ in the
 510 condensation-only case by using the calibrated bulk method and SDM. Height-time con-

511 tours of specific cloud water content q_c and the profiles of q_c at $t = 20$ min are com-
 512 pared. Note that the bulk method simulation with the result of UKI is very similar to
 513 that of EKI and is therefore not shown. The simulation results by the calibrated bulk
 514 method are in excellent agreement with the results of the SDM. This excellent agreement
 515 confirms the satisfactory performance of the implementation of condensation and wa-
 516 ter vapor advection in the bulk method simulations.

517 Figures 6 and 7 compare simulations using the SDM, the bulk method before train-
 518 ing, and the calibrated bulk method by EKI and UKI. Figure 6 visualizes contours of
 519 specific cloud water content q_c and specific rainwater content q_r in height and time, while
 520 figure 7 shows profiles of q_c and q_r at $t = 10$ min, $t = 20$ min, $t = 30$ min, and $t = 50$
 521 min. As evidenced in these figures, the bulk method with the initial parameters under-
 522 estimates the specific cloud water content and incorrectly predicts an early peak in spec-
 523 ific rainwater content. These deviations suggest an overestimation of rain production
 524 and sedimentation rates in the bulk method before training. However, both EKI and UKI
 525 optimal parameters significantly improve the bulk method simulations with respect to
 526 the SDM results. After calibrations, the auto-conversion parameters $\tau_{acnv,0}$ and α_{acnv}
 527 increase, resulting in reduced auto-conversion rates. Additionally, the terminal velocity
 528 parameter χ_v decreases, leading to reduced sedimentation. On the other hand, the ac-
 529 cretion parameter χ_a increases in both EKI and UKI calibrations. However, it is impor-
 530 tant to note that the accretion rate, which is influenced by sedimentation, is governed
 531 by the product $\chi_a\chi_v$. In the calibrated bulk method, this product slightly increases com-
 532 pared to that with the initial parameters. These parameter adjustments contribute to
 533 the overall decrease of rain formation and sedimentation, and the reasonable agreement
 534 of the calibrated bulk method, by both EKI and UKI, with the SDM results.

535 While the simulations using the calibrated bulk method by EKI and UKI yield sim-
 536 ilar overall results, there are differences in specific details. For example, the maximum
 537 specific rainwater content for the EKI calibrated bulk method exceeds that for the UKI
 538 calibrated bulk method by more than 30%. Also, when q_r for the SDM peaks ($t \sim 20$
 539 min), the EKI calibrated bulk method underestimates q_c close to the cloud base while
 540 the UKI calibrated bulk method overestimates it compared to SDM results. This obser-
 541 vation suggests that the rain production rate for the EKI calibrated bulk method is over-
 542 estimated while that for the UKI calibrated bulk method is underestimated. This is con-
 543 firmed in figure 8 where cloud and rain water path and surface rain rate are visualized
 544 over time. The rainwater path for the SDM peaks slightly after that for the EKI cali-
 545 brated bulk method and shortly before that for the UKI calibrated bulk method, indi-
 546 cating the overestimation of the rain production rate by the EKI calibrated bulk method
 547 and the underestimation of the rate by the UKI calibrated bulk method. The higher rain
 548 production rates predicted by the EKI calibrated bulk method occur around the peak
 549 of q_r , which corresponds to the period when accretion is the dominant rain formation
 550 process. This observation suggests that the EKI calibrated method predicts higher ac-
 551 cretion rates for large values of q_r compared to the UKI calibrated method. This differ-
 552 ence in accretion rates can be attributed to the higher value of the accretion parame-
 553 ter χ_a in the EKI parameter set. Additionally, it is notable that the surface rain rate
 554 for the EKI calibrated bulk method is more than 30% higher than that for the UKI cal-
 555 ibrated bulk method. The higher surface rain rate is due to the lower evaporation rate
 556 of the EKI calibrated bulk method (caused by smaller b_{vent}) than that of the UKI cal-
 557 ibrated bulk method.

558 The bulk method before training incorrectly predicts an early surface rain rate due
 559 to the incorrect prediction of early rain production. The calibrated bulk methods by both
 560 EKI and UKI predict the timing of the surface rain rate very well. However, they fail
 561 to correctly predict the magnitude of the maximum rain rate. The significant error in
 562 the prediction of the maximum surface rain rate despite capturing q_r well can be attributed
 563 to the inability of the single-moment bulk method to adequately predict the terminal ve-

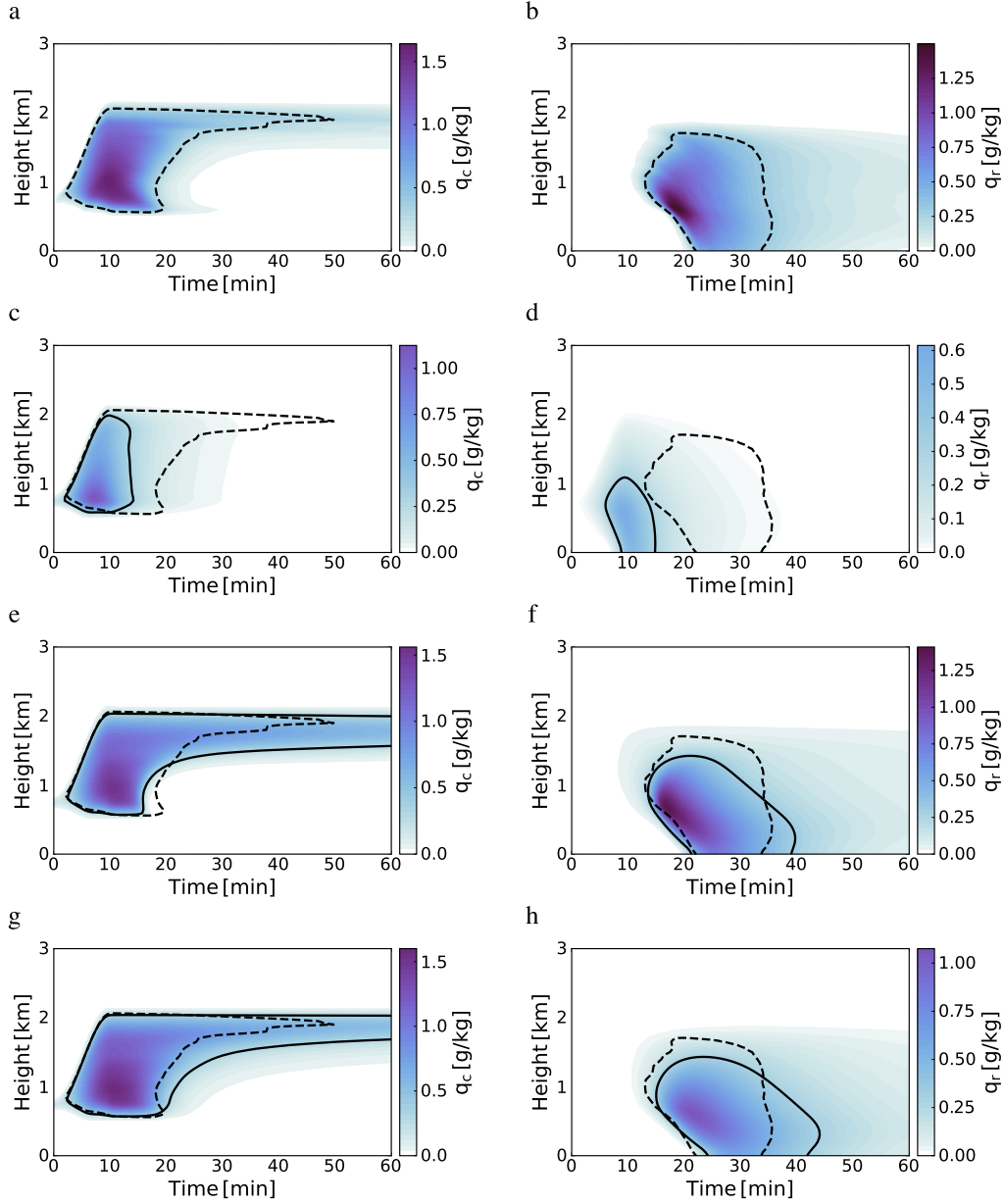


Figure 6. Comparison of the KiD-1d model simulations using the SDM and the bulk method. Height-time contours of specific cloud water content q_c (left panels) and specific rainwater content q_r (right panels) are compared for the simulations using the SDM (a and b), the bulk method with the initial parameters (c and d), and the calibrated method by EKI (e and f) and UKI (g and h). Black solid lines indicate $q_c = 0.3 \text{ g kg}^{-1}$ (left panels) and $q_r = 0.3 \text{ g kg}^{-1}$ (right panels), while black dashed lines represent the same contour levels for the SDM results, overlaid on all panels for comparison. The simulations use $(\rho w)_0 = 3 \text{ kg m}^{-2} \text{ s}^{-1}$, $N_a = 50 \text{ cm}^{-3}$, and $p_0 = 994 \text{ hPa}$. The SDM results are the average of 100 simulations.

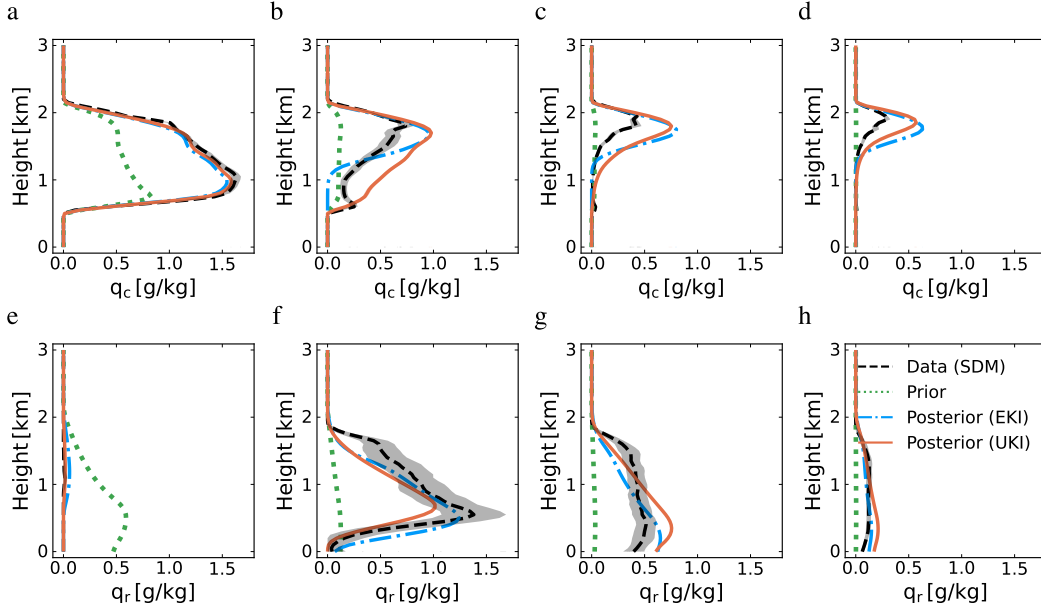


Figure 7. Solutions of the SDM and the bulk method with the initial and the calibrated parameters are compared. The specific cloud water content profiles (top panels) and specific rain-water content profiles (bottom panels) are shown at times $t = 10$ min (panels a and e), $t = 20$ min (panels b and f), $t = 30$ min (panels c and g) and $t = 50$ min (panels d and h). The calibrated bulk method results are obtained by evaluating the model using the ensemble means. For the SDM results, the dashed lines represent the average of 100 simulations, while the shadings visualize the variability, showing plus and minus one standard deviation.

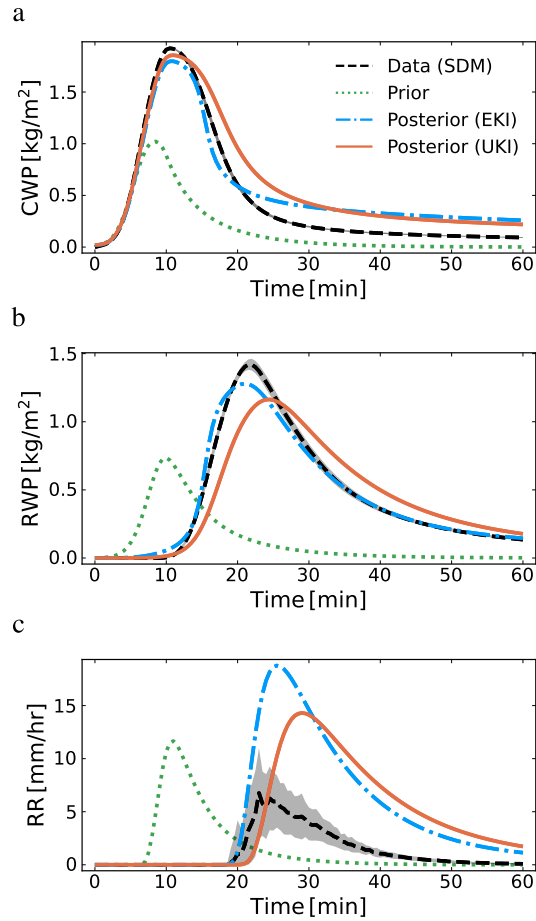


Figure 8. Comparison of the cloud water path CWP (a), rainwater path RWP (b), and surface rain rate RR (c) for simulations using the SDM (black dashed), the bulk method with the initial parameters (green dotted), and the calibrated bulk method by EKI (blue dash-dot) and UKI (red solid). The results of the calibrated bulk method are obtained by evaluating the bulk method with the ensemble mean. The SDM results represent the average of 100 simulations, and the profile variability is indicated by shading plus and minus one standard deviation.

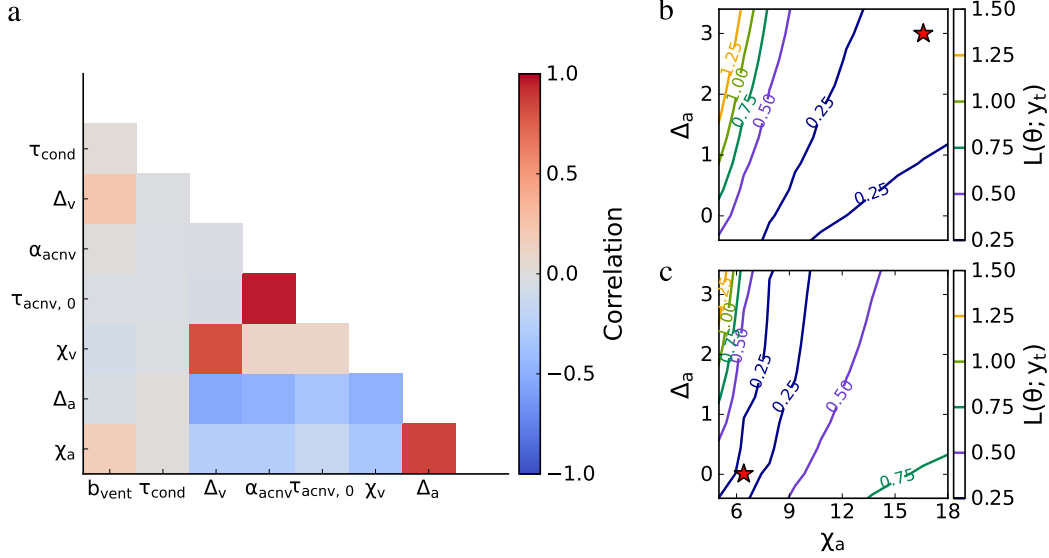


Figure 9. Parameter correlations estimated using the UKI method (a), and contours of the loss function $L(\theta; y_t)$ for varying accretion parameters χ_a and Δ_a , while keeping other parameters fixed at the EKI (b) and UKI (c) optimal values. The markers indicate the optimal values of the accretion parameters obtained by EKI (b) and UKI (c). The loss values are normalized by the value of the loss evaluated for the bulk model with the initial parameters.

564 locity of particles. The poor representation of terminal velocity by the single-moment
 565 bulk scheme is inevitable as terminal velocity is simply a single-valued function of q_r and
 566 the gravitational size sorting is not captured (Milbrandt & McTaggart-Cowan, 2010).
 567 The prediction of the maximum surface rain rate can be improved by using multi-moment
 568 bulk schemes with sedimentation rates that can capture gravitational size sorting.

569 In addition to the maximum likelihood estimator, UKI provides correlations be-
 570 tween model parameters. Figure 9(a) visualizes the correlation map between param-
 571 eters of the single-moment bulk scheme obtained by the UKI calibration. For the employed
 572 training dataset, the calibrated bulk scheme shows strong correlations between the two
 573 auto-conversion parameters $\tau_{\text{acnv},0}$ and α_{acnv} , between the two accretion coefficients χ_a
 574 and Δ_a , as well as between the two terminal velocity coefficients χ_v and Δ_v . Also, both
 575 accretion coefficients are moderately anti-correlated with auto-conversion and terminal
 576 velocity coefficients.

577 The correlations between the two accretion coefficients and between the two ter-
 578 minal velocity coefficients can be attributed to the compensatory nature of these param-
 579 eters in their corresponding process rate equations. Specifically, an increase in the scal-
 580 ing factor (e.g., χ_a or χ_v) is accompanied by a corresponding increase in the exponent
 581 of q_r (e.g., Δ_a or Δ_v). The anti-correlations between the accretion and terminal ve-
 582 locity parameters arise from the direct effect of sedimentation on the accretion rate. The
 583 anti-correlation between the accretion coefficients and the auto-conversion parameters
 584 is due to the counterbalance between these two processes in the early stages of rain for-
 585 mation. The strong correlation between the two auto-conversion parameters suggests that
 586 as the initial number density N_a increases, a greater adjustment in the auto-conversion
 587 process is required to maintain a balanced rain formation process.

588 Utilizing the correlation information provided by UKI can contribute to refining
 589 the parameterizations of the bulk method by identifying a smaller set of uncorrelated

590 parameters for calibrations. For instance, the strong correlation between auto-conversion
 591 parameters suggests that training the model for only one of the two parameters might
 592 result in a similar reduction of the model-data mismatch as training for both param-
 593 eters.

594 The parameter correlations derived from UKI are consistent with the differences
 595 between the optimal parameter sets obtained by EKI and UKI. In the set of UKI op-
 596 timal parameters, both auto-conversion parameters are higher than those in the EKI set,
 597 while both accretion coefficients are lower and both terminal velocity coefficients are slightly
 598 higher. The parameter correlations and the consistent differences between the EKI and
 599 UKI optimal parameter sets suggest the existence of a range of parameters for which the
 600 model-data mismatch remains acceptably small. This is illustrated in Figures 9b and 9c,
 601 where contours of the configuration-averaged loss function are visualized for varying ac-
 602 cretion parameters χ_a and Δ_a , while other parameters are fixed at the EKI or UKI op-
 603 timal values. As evidenced in this figure, the loss function value remains below 25% within
 604 a notably wide region in the space of χ_a and Δ_a . These results demonstrate the exist-
 605 tence of a continuous range of parameter combinations that yield satisfactory model per-
 606 formance, allowing for flexibility in selecting parameter values. Obtaining a unique set
 607 of parameters can be achieved by providing additional constraints for parameter estima-
 608 tion through incorporating detailed information about auto-conversion and accretion pro-
 609 cesses in the training data. By leveraging such information, it may become possible to
 610 refine the parameterizations of the bulk method and enhance the model’s capability to
 611 capture the underlying dynamics. The investigation into incorporating auto-conversion
 612 and accretion process rates into the parameterization of the bulk model is left for future
 613 research.

614 4 Summary and conclusion

615 The aim of this study was to improve the accuracy of the representation of cloud
 616 and precipitation processes within bulk schemes. We presented a calibration framework
 617 for training warm-rain bulk microphysics schemes by using high-fidelity super-droplet
 618 simulations. The calibration framework uses ensemble Kalman methods for training the
 619 models, including ensemble Kalman inversion (EKI) and unscented Kalman inversion
 620 (UKI). Calibrations are carried out by leveraging simulations of the KiD-1d model, a one-
 621 dimensional rain-shaft model that has been widely used for studying microphysics schemes.
 622 In this model, the updraft and the temperature profile are prescribed so that any vari-
 623 ation in the results can only be attributed to changes in the employed microphysics scheme.

624 To benchmark the performance of the bulk methods, we generated a library of super-
 625 droplet simulations of a rain shaft model. Simulations were carried out for different up-
 626 draft amplitudes, initial aerosol number density and surface air pressure to provide a wide
 627 range of precipitation conditions for comparing and evaluating bulk microphysics schemes.

628 Our results demonstrate the effectiveness of the calibration framework by apply-
 629 ing it to a single-moment microphysics model. While calibrations by EKI and UKI re-
 630 sult in two different sets of parameters, the calibrated bulk method by both EKI and UKI
 631 shows a significant reduction in model error with respect to the super-droplet simula-
 632 tions. Specifically, the prediction of cloud and rain profiles showed excellent agreement
 633 with the reference simulations. However, while the timing of the surface precipitation
 634 rate showed improvement, the magnitude of the maximum rain rate was overpredicted
 635 by the single-moment bulk scheme. This finding emphasizes the need for further research
 636 to capture the surface precipitation rate more accurately, particularly by exploring the
 637 potential of higher-moment schemes that can represent the gravitational size sorting of
 638 particles.

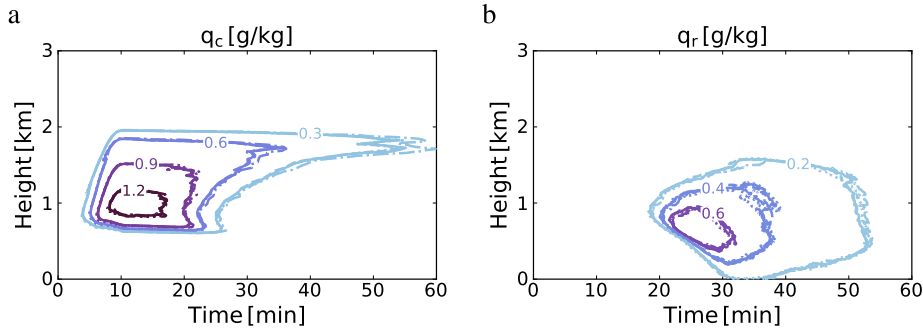


Figure A1. Comparison of KiD-1d model simulations using the SDM with different numerical settings. Height-time contours of specific content for cloud water (panel a) and rain (panel b) are shown. The solid contour lines represent the simulation with $N_{sd} = 512$ super-droplets, grid spacing of $dz = 50$ m, and time steps of $dt = 5$ s. This simulation is compared with simulations using doubled number of super-droplets (dashed), halved grid spacing (dashdot) and halved time steps (dot). The averages of 100 simulations for each set of numerical settings are shown. The excellent agreement of results indicates that the KiD-1d model simulations are insensitive to the numerical settings used.

639 Our study highlights the potential of calibrating classic parameterizations of mi-
 640 crophysics using high-fidelity super-droplet simulations. Although super-droplet tech-
 641 niques are still in their early stages and pose potential limitations in capturing the en-
 642 tirety of the underlying physical phenomena (Morrison et al., 2020; Hill et al., 2023), lever-
 643 aging the valuable insights obtained from these simulations can enhance classic mi-
 644 crophysics parameterizations. Unlike observational data, these simulations allow us to dis-
 645 entangle microphysics from other dynamics and calibrate microphysics processes in iso-
 646 lation from their feedbacks with atmospheric flows. This is a significant advantage, as
 647 it enables us to explore and refine microphysics parameterizations in a controlled man-
 648 ner, which would be challenging even with abundant laboratory or observational data.
 649 Utilizing super-droplet simulations is a promising approach to improve microphysics pa-
 650 rameterizations, particularly in regions where clouds show strong sensitivity to mi-
 651 crophysics parameters. Further research in this direction is needed to explore the full po-
 652 tential and capability of the super-droplet simulations in improving the accuracy of clas-
 653 sic parameterizations of cloud microphysics.

654 Appendix A Result independence from numerical values

655 Figure A1 compares SDM simulations of the KiD-1d model with $(\rho w)_0 = 3 \text{ kg}$
 656 $\text{m}^{-2} \text{ s}^{-1}$, $N_a = 100 \text{ cm}^{-3}$ and $p_0 = 1000 \text{ hPa}$ for different numerical setups. The refer-
 657 ence simulation with an average of $N_{sd} = 512$ super-droplets per grid box, $dz = 50$
 658 m and $dt = 5$ s is compared against simulations with doubled number of super-droplets,
 659 halved grid spacing, and halved time step. The results are in excellent agreement, in-
 660 dicated the independence of the reference simulation from specific numerical values.

661 Appendix B Parameter evolution in EKI and UKI calibrations

662 Figure B1 displays the evolution of all calibrated parameters during the calibra-
 663 tion of the single-moment bulk scheme using EKI and UKI methods. The calibrated pa-
 664 rameters include accretion coefficients, terminal velocity coefficients, auto-conversion co-
 665 efficients, condensation time scale, and evaporation coefficient. While EKI and UKI show

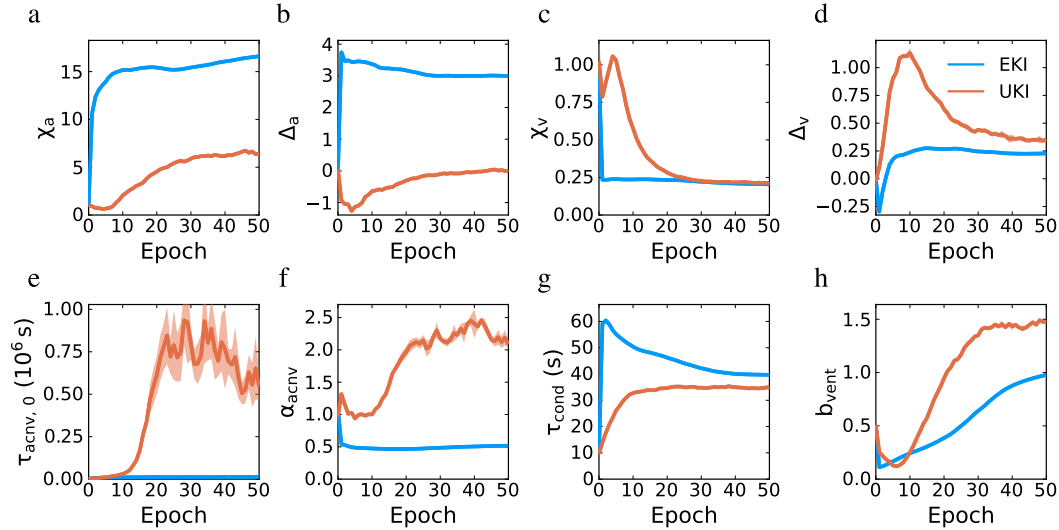


Figure B1. Parameter evolutions of the accretion coefficients χ_a (a) and Δ_a (b), terminal velocity coefficients χ_v (c) and Δ_v (d), auto-conversion coefficients $\tau_{acnv,0}$ (e) and α_{acnv} (f), condensation time scale τ_{cond} (g), and evaporation coefficient b_{vent} (h) during calibrations using EKI (blue) and UKI (red). The parameter uncertainty obtained from UKI is illustrated by shadings, indicating plus and minus one standard deviation of the parameter ensemble.

666 comparable final converged values for the terminal velocity coefficients (χ_v and Δ_v) and
 667 the condensation time scale (τ_{cond}), the final converged values of the remaining param-
 668 eters by EKI and UKI are significantly different.

669 Data Availability Statement

670 The library of super-droplet simulations is available at [https://doi.org/10.5281/](https://doi.org/10.5281/zenodo.8336442)
 671 [zenodo.8336442](https://doi.org/10.5281/zenodo.8336442). We used PySDM v2.15 (<https://github.com/open-atmos/PySDM>)
 672 to generate the super-droplet simulations. The code for the calibration pipeline can be
 673 found at <https://doi.org/10.5281/zenodo.8362305> and [https://github.com/CliMA/](https://github.com/CliMA/Kinematic1D.jl)
 674 [Kinematic1D.jl](https://github.com/CliMA/Kinematic1D.jl). For the calibrations, we used the Julia packages [CloudMicrophysics.jl](https://github.com/CliMA/CloudMicrophysics.jl)
 675 v0.13.3 (<https://github.com/CliMA/CloudMicrophysics.jl>) and [EnsembleKalman-](https://github.com/CliMA/EnsembleKalmanProcesses.jl)
 676 [Processes.jl](https://github.com/CliMA/EnsembleKalmanProcesses.jl) v1.1 (<https://github.com/CliMA/EnsembleKalmanProcesses.jl>).

677 Acknowledgments

678 Sajjad Azimi acknowledges support by the Swiss National Science Foundation (SNSF,
 679 Grant No. P500PN_202876). Sylwester Arabas acknowledges support by the Polish Na-
 680 tional Science Centre (NCN, Grant No. 2020/39/D/ST10/01220). This research was ad-
 681 ditionally supported by the generosity of Eric and Wendy Schmidt by recommenda-
 682 tion of the Schmidt Futures program and by the U.S. National Science Foundation (Grant
 683 No. AGS-1835860). The authors report no conflicts of interest.

684 References

685 Andrejczuk, M., Grabowski, W. W., Reisner, J., & Gadian, A. (2010, 11). Cloud-
 686 aerosol interactions for boundary layer stratocumulus in the Lagrangian
 687 Cloud Model. *Journal of Geophysical Research*, 115(D22), D22214. Re-
 688 trieved from <http://doi.wiley.com/10.1029/2010JD014248> doi:

- 689 10.1029/2010JD014248
- 690 Bartman, P., Banaśkiewicz, J., Drenda, S., Manna, M., Olesik, M. A., Rozwoda, P.,
691 ... Arabas, S. (2022, 9). PyMPDATA v1: Numba-accelerated implementation
692 of MPDATA with examples in Python, Julia and Matlab. *Journal of Open*
693 *Source Software*, 7(77), 3896. doi: 10.21105/joss.03896
- 694 Bartman, P., Bulenok, O., Górski, K., Jaruga, A., Lazarski, G., Olesik, M. A., ...
695 Arabas, S. (2022, 4). PySDM v1: particle-based cloud modeling package
696 for warm-rain microphysics and aqueous chemistry. *Journal of Open Source*
697 *Software*, 7(72), 3219. doi: 10.21105/joss.03219
- 698 Beheng, K. (1994, 6). A parameterization of warm cloud microphysical conver-
699 sion processes. *Atmospheric Research*, 33(1-4), 193–206. doi: 10.1016/0169-
700 -8095(94)90020-5
- 701 Bieli, M., Dunbar, O. R. A., de Jong, E. K., Jaruga, A., Schneider, T., & Bischoff,
702 T. (2022, 8). An Efficient Bayesian Approach to Learning Droplet Collision
703 Kernels: Proof of Concept Using “Cloudy,” a New μ_0 -Moment Bulk Mi-
704 crophysics Scheme. *Journal of Advances in Modeling Earth Systems*, 14(8).
705 doi: 10.1029/2022MS002994
- 706 de Jong, E. K., Singer, C. E., Azimi, S., Bartman, P., Bulenok, O., Derlatka, K., ...
707 Arabas, S. (2023, 4). New developments in PySDM and PySDM-examples
708 v2: collisional breakup, immersion freezing, dry aerosol initialization, and
709 adaptive time-stepping. *Journal of Open Source Software*, 8(84), 4968. Re-
710 trieved from <https://joss.theoj.org/papers/10.21105/joss.04968> doi:
711 10.21105/joss.04968
- 712 Dunbar, O. R. A., Garbuno-Inigo, A., Schneider, T., & Stuart, A. M. (2021, 9).
713 Calibration and Uncertainty Quantification of Convective Parameters in an
714 Idealized GCM. *Journal of Advances in Modeling Earth Systems*, 13(9). doi:
715 10.1029/2020MS002454
- 716 Dunbar, O. R. A., Howland, M. F., Schneider, T., & Stuart, A. M. (2022, 9).
717 Ensemble-Based Experimental Design for Targeting Data Acquisition to In-
718 form Climate Models. *Journal of Advances in Modeling Earth Systems*, 14(9).
719 doi: 10.1029/2022MS002997
- 720 Gettelman, A., Gagne, D. J., Chen, C., Christensen, M. W., Lebo, Z. J., Mor-
721 rison, H., & Gantos, G. (2021, 2). Machine Learning the Warm Rain
722 Process. *Journal of Advances in Modeling Earth Systems*, 13(2). doi:
723 10.1029/2020MS002268
- 724 Grabowski, W. W. (1998, 11). Toward Cloud Resolving Modeling of Large-Scale
725 Tropical Circulations: A Simple Cloud Microphysics Parameterization. *Jour-
726 nal of the Atmospheric Sciences*, 55(21), 3283–3298. Retrieved from [http://](http://journals.ametsoc.org/doi/10.1175/1520-0469(1998)055<3283:TCRMOL>2.0.CO;2)
727 [journals.ametsoc.org/doi/10.1175/1520-0469\(1998\)055<3283:TCRMOL>2](http://journals.ametsoc.org/doi/10.1175/1520-0469(1998)055<3283:TCRMOL>2.0.CO;2)
728 [.0.CO;2](http://journals.ametsoc.org/doi/10.1175/1520-0469(1998)055<3283:TCRMOL>2.0.CO;2) doi: 10.1175/1520-0469(1998)055<3283:TCRMOL>2.0.CO;2
- 729 Grabowski, W. W., Morrison, H., Shima, S.-I., Abade, G. C., Dziekan, P., &
730 Pawlowska, H. (2019, 4). Modeling of Cloud Microphysics: Can We Do
731 Better? *Bulletin of the American Meteorological Society*, 100(4), 655–672. doi:
732 10.1175/BAMS-D-18-0005.1
- 733 Grabowski, W. W., & Smolarkiewicz, P. K. (1996, 3). Two-Time-Level Semi-
734 Lagrangian Modeling of Precipitating Clouds. *Monthly Weather Review*,
735 124(3), 487–497. Retrieved from [http://journals.ametsoc.org/doi/](http://journals.ametsoc.org/doi/10.1175/1520-0493(1996)124<0487:TTLSLM>2.0.CO;2)
736 [10.1175/1520-0493\(1996\)124<0487:TTLSLM>2.0.CO;2](http://journals.ametsoc.org/doi/10.1175/1520-0493(1996)124<0487:TTLSLM>2.0.CO;2) doi: 10.1175/
737 1520-0493(1996)124<0487:TTLSLM>2.0.CO;2
- 738 Hill, A. A., Lebo, Z. J., Andrejczuk, M., Arabas, S., Dziekan, P., Field, P., ... Vié,
739 B. (2023, 1). Toward a Numerical Benchmark for Warm Rain Processes. *Jour-
740 nal of the Atmospheric Sciences*. Retrieved from [https://journals.ametsoc](https://journals.ametsoc.org/view/journals/atsc/aop/JAS-D-21-0275.1/JAS-D-21-0275.1.xml)
741 [.org/view/journals/atsc/aop/JAS-D-21-0275.1/JAS-D-21-0275.1.xml](https://journals.ametsoc.org/view/journals/atsc/aop/JAS-D-21-0275.1/JAS-D-21-0275.1.xml)
742 doi: 10.1175/JAS-D-21-0275.1
- 743 Huang, D. Z., Schneider, T., & Stuart, A. M. (2022, 8). Iterated Kalman methodol-

- ogy for inverse problems. *Journal of Computational Physics*, 463, 111262. doi: 10.1016/j.jcp.2022.111262
- Igel, A. L., Morrison, H., Santos, S. P., & van Lier-Walqui, M. (2022, 6). Limitations of Separate Cloud and Rain Categories in Parameterizing Collision-Coalescence for Bulk Microphysics Schemes. *Journal of Advances in Modeling Earth Systems*, 14(6). doi: 10.1029/2022MS003039
- Iglesias, M. A., Law, K. J. H., & Stuart, A. M. (2013, 4). Ensemble Kalman methods for inverse problems. *Inverse Problems*, 29(4), 045001. doi: 10.1088/0266-5611/29/4/045001
- Kessler, E. (1969). On the Distribution and Continuity of Water Substance in Atmospheric Circulations. In *On the distribution and continuity of water substance in atmospheric circulations* (pp. 1–84). Boston, MA: American Meteorological Society. Retrieved from http://link.springer.com/10.1007/978-1-935704-36-2_1 doi: 10.1007/978-1-935704-36-2_{_}1
- Khain, A. P., Beheng, K. D., Heymsfield, A., Korolev, A., Krichak, S. O., Levin, Z., ... Yano, J.-I. (2015, 6). Representation of microphysical processes in cloud-resolving models: Spectral (bin) microphysics versus bulk parameterization. *Reviews of Geophysics*, 53(2), 247–322. doi: 10.1002/2014RG000468
- Khairoutdinov, M., & Kogan, Y. (2000, 1). A New Cloud Physics Parameterization in a Large-Eddy Simulation Model of Marine Stratocumulus. *Monthly Weather Review*, 128(1), 229–243. Retrieved from [http://journals.ametsoc.org/doi/10.1175/1520-0493\(2000\)128<0229:ANCPPI>2.0.CO;2](http://journals.ametsoc.org/doi/10.1175/1520-0493(2000)128<0229:ANCPPI>2.0.CO;2) doi: 10.1175/1520-0493(2000)128<0229:ANCPPI>2.0.CO;2
- Kogan, Y. (2013, 5). A Cumulus Cloud Microphysics Parameterization for Cloud-Resolving Models. *Journal of the Atmospheric Sciences*, 70(5), 1423–1436. doi: 10.1175/JAS-D-12-0183.1
- Korolev, A. V., & Mazin, I. P. (2003, 12). Supersaturation of Water Vapor in Clouds. *Journal of the Atmospheric Sciences*, 60(24), 2957–2974. Retrieved from [http://journals.ametsoc.org/doi/10.1175/1520-0469\(2003\)060<2957:SOWVIC>2.0.CO;2](http://journals.ametsoc.org/doi/10.1175/1520-0469(2003)060<2957:SOWVIC>2.0.CO;2) doi: 10.1175/1520-0469(2003)060<2957:SOWVIC>2.0.CO;2
- Kovachki, N. B., & Stuart, A. M. (2019, 9). Ensemble Kalman inversion: a derivative-free technique for machine learning tasks. *Inverse Problems*, 35(9), 095005. doi: 10.1088/1361-6420/ab1c3a
- Li, M., Zhang, T., Chen, Y., & Smola, A. J. (2014, 8). Efficient mini-batch training for stochastic optimization. In *Proceedings of the 20th acm sigkdd international conference on knowledge discovery and data mining* (pp. 661–670). New York, NY, USA: ACM. doi: 10.1145/2623330.2623612
- Liu, Y., & Daum, P. H. (2004, 7). Parameterization of the Autoconversion Process. Part I: Analytical Formulation of the Kessler-Type Parameterizations. *Journal of the Atmospheric Sciences*, 61(13), 1539–1548. doi: 10.1175/1520-0469(2004)061<1539:POTAPI>2.0.CO;2
- Lopez-Gomez, I., Christopoulos, C., Langeland Ervik, H. L., Dunbar, O. R. A., Cohen, Y., & Schneider, T. (2022, 8). Training Physics-Based Machine-Learning Parameterizations With Gradient-Free Ensemble Kalman Methods. *Journal of Advances in Modeling Earth Systems*, 14(8). Retrieved from <https://onlinelibrary.wiley.com/doi/10.1029/2022MS003105> doi: 10.1029/2022MS003105
- Marshall, J. S., & Palmer, W. M. K. (1948, 8). THE DISTRIBUTION OF RAIN-DROPS WITH SIZE. *Journal of Meteorology*, 5(4), 165–166. Retrieved from [http://journals.ametsoc.org/doi/10.1175/1520-0469\(1948\)005<0165:TDORWS>2.0.CO;2](http://journals.ametsoc.org/doi/10.1175/1520-0469(1948)005<0165:TDORWS>2.0.CO;2) doi: 10.1175/1520-0469(1948)005<0165:TDORWS>2.0.CO;2
- Milbrandt, J. A., & McTaggart-Cowan, R. (2010, 12). Sedimentation-Induced Errors in Bulk Microphysics Schemes. *Journal of the Atmospheric Sciences*, 67(12),

- 799 3931–3948. doi: 10.1175/2010JAS3541.1
- 800 Milbrandt, J. A., & Yau, M. K. (2005, 9). A Multimoment Bulk Microphysics
801 Parameterization. Part I: Analysis of the Role of the Spectral Shape Pa-
802 rameter. *Journal of the Atmospheric Sciences*, 62(9), 3051–3064. doi:
803 10.1175/JAS3534.1
- 804 Morrison, H., & Grabowski, W. W. (2007, 8). Comparison of Bulk and Bin Warm-
805 Rain Microphysics Models Using a Kinematic Framework. *Journal of the At-
806 mospheric Sciences*, 64(8), 2839–2861. doi: 10.1175/JAS3980
- 807 Morrison, H., & Milbrandt, J. A. (2015, 1). Parameterization of Cloud Microphysics
808 Based on the Prediction of Bulk Ice Particle Properties. Part I: Scheme De-
809 scription and Idealized Tests. *Journal of the Atmospheric Sciences*, 72(1),
810 287–311. doi: 10.1175/JAS-D-14-0065.1
- 811 Morrison, H., van Lier-Walqui, M., Kumjian, M. R., & Prat, O. P. (2019, 3). A
812 Bayesian Approach for Statistical–Physical Bulk Parameterization of Rain
813 Microphysics. Part I: Scheme Description. *Journal of the Atmospheric Sci-
814 ences*, 77(3), 1019–1041. Retrieved from [http://journals.ametsoc.org/
815 doi/10.1175/JAS-D-19-0070.1](http://journals.ametsoc.org/doi/10.1175/JAS-D-19-0070.1) doi: 10.1175/JAS-D-19-0070.1
- 816 Morrison, H., van Lier-Walqui, M., Fridlind, A. M., Grabowski, W. W., Harrington,
817 J. Y., Hoose, C., ... Xue, L. (2020, 8). Confronting the Challenge of Modeling
818 Cloud and Precipitation Microphysics. *Journal of Advances in Modeling Earth
819 Systems*, 12(8). Retrieved from [https://onlinelibrary.wiley.com/doi/
820 10.1029/2019MS001689](https://onlinelibrary.wiley.com/doi/10.1029/2019MS001689) doi: 10.1029/2019MS001689
- 821 Noh, Y., Oh, D., Hoffmann, F., & Raasch, S. (2018, 11). A Cloud Microphysics Pa-
822 rameterization for Shallow Cumulus Clouds Based on Lagrangian Cloud Model
823 Simulations. *Journal of the Atmospheric Sciences*, 75(11), 4031–4047. doi:
824 10.1175/JAS-D-18-0080.1
- 825 Posselt, D. J. (2016). A Bayesian examination of deep convective squall-line sensitiv-
826 ity to changes in cloud microphysical parameters. *Journal of the Atmospheric
827 Sciences*, 73(2), 637–665. doi: 10.1175/JAS-D-15-0159.1
- 828 Posselt, D. J., & Vukicevic, T. (2010, 5). Robust characterization of model physics
829 uncertainty for simulations of deep moist convection. *Monthly Weather Re-
830 view*, 138(5), 1513–1535. doi: 10.1175/2009MWR3094.1
- 831 Riechermann, T., Noh, Y., & Raasch, S. (2012, 6). A new method for large-
832 eddy simulations of clouds with Lagrangian droplets including the ef-
833 fects of turbulent collision. *New Journal of Physics*, 14(6), 065008. doi:
834 10.1088/1367-2630/14/6/065008
- 835 Seifert, A., & Beheng, K. D. (2006, 2). A two-moment cloud microphysics parame-
836 terization for mixed-phase clouds. Part 1: Model description. *Meteorology and
837 Atmospheric Physics*, 92(1-2), 45–66. Retrieved from [http://link.springer
838 .com/10.1007/s00703-005-0112-4](http://link.springer.com/10.1007/s00703-005-0112-4) doi: 10.1007/s00703-005-0112-4
- 839 Shima, S., Kusano, K., Kawano, A., Sugiyama, T., & Kawahara, S. (2009, 7). The
840 super-droplet method for the numerical simulation of clouds and precipita-
841 tion: a particle-based and probabilistic microphysics model coupled with a
842 non-hydrostatic model. *Quarterly Journal of the Royal Meteorological Society*,
843 135(642), 1307–1320. Retrieved from [https://onlinelibrary.wiley.com/
844 doi/10.1002/qj.441](https://onlinelibrary.wiley.com/doi/10.1002/qj.441) doi: 10.1002/qj.441
- 845 Shipway, B. J., & Hill, A. A. (2012, 10). Diagnosis of systematic differences between
846 multiple parametrizations of warm rain microphysics using a kinematic frame-
847 work. *Quarterly Journal of the Royal Meteorological Society*, 138(669), 2196–
848 2211. Retrieved from [https://onlinelibrary.wiley.com/doi/10.1002/
849 qj.1913](https://onlinelibrary.wiley.com/doi/10.1002/qj.1913) doi: 10.1002/qj.1913
- 850 Tao, W. K., Chen, J. P., Li, Z., Wang, C., & Zhang, C. (2012, 6). Impact of aerosols
851 on convective clouds and precipitation. *Reviews of Geophysics*, 50(2). doi: 10
852 .1029/2011RG000369
- 853 Tripoli, G. J., & Cotton, W. R. (1980, 9). A Numerical Investigation of Several

- 854 Factors Contributing to the Observed Variable Intensity of Deep Convection
855 over South Florida. *Journal of Applied Meteorology*, 19(9), 1037–1063. doi:
856 10.1175/1520-0450(1980)019<1037:ANIOSF>2.0.CO;2
- 857 van Lier-Walqui, M., Morrison, H., Kumjian, M. R., Reimel, K. J., Prat, O. P.,
858 Lunderman, S., & Morzfeld, M. (2020, 3). A Bayesian approach for statisti-
859 cal–physical bulk parameterization of rain microphysics. Part II: Idealized
860 Markov chain Monte Carlo experiments. *Journal of the Atmospheric Sciences*,
861 77(3), 1043–1064. doi: 10.1175/JAS-D-19-0071.1
- 862 Xiao, H., Wu, J.-L., Wang, J.-X., Sun, R., & Roy, C. (2016, 11). Quantifying and
863 reducing model-form uncertainties in Reynolds-averaged Navier–Stokes sim-
864 ulations: A data-driven, physics-informed Bayesian approach. *Journal of*
865 *Computational Physics*, 324, 115–136. doi: 10.1016/j.jcp.2016.07.038
- 866 Zeng, X., & Li, X. (2020, 3). A Two-Moment Bulk Parameterization of the Drop
867 Collection Growth in Warm Clouds. *Journal of the Atmospheric Sciences*,
868 77(3), 797–811. doi: 10.1175/JAS-D-19-0015.1