# Limitations of separate cloud and rain categories in parameterizing collision-coalescence for bulk microphysics schemes

Adele Igel[1], H Morrison[2], S P Santos[3], and M Van Lier-Walqui[3]

[1]University of California, Davis
[2]National Center for Atmospheric Research
[3]Columbia University

November 26, 2022



### Abstract

Warm rain collision coalescence has been persistently difficult to parameterize in bulk microphysics schemes. Here we use a flexible bulk microphysics scheme with bin scheme process parameterizations, called AMP, to investigate reasons for the difficulty. AMP is configured in a variety of ways to mimic bulk schemes and is compared to simulations with the bin scheme upon which AMP is built. We find that the biggest limitation in traditional bulk schemes is the use of separate cloud and rain categories. When the drop size distribution is instead represented by a continuous distribution with or without an explicit functional form, the simulation of cloud-to-rain conversion is substantially improved. We find that the use of an assumed double-mode gamma distribution and the choice of predicted distribution moments do somewhat influence the ability of AMP to simulate rain production, but much less than using a single liquid category compared to separate cloud and rain categories. Traditional two category configurations of AMP are always too slow in producing rain due to their struggle to capture the emergence of the rain mode. Single category configurations may produce rain either too slowly or too quickly, with too slow production more likely for initially narrow droplet size distributions. However, the average error magnitude is much smaller using a single category than two categories. Optimal moment combinations for the single category approach appear to be linked more to the information content they provide for constraining the size distributions than to their correlation with collision-coalescence rates.

# Limitations of separate cloud and rain categories in parameterizing collision-coalescence for bulk microphysics schemes

**A. L. Igel[1], H. Morrison[2], S. P. Santos[3,4], and M. van Lier-Walqui[3,4]**

[1]Department of Land, Air and Water Resources, University of California Davis
[2]National Center for Atmospheric Research
[3]Center for Climate Systems Research, Columbia University
[4]NASA Goddard Institute for Space Studies

**Key Points:**

- A single category, four moment scheme simulates autoconversion and accretion far better than a two category, two moment scheme.
- The rain mode forms at diameters that are much smaller than are traditionally considered to be rain.
- Using one versus two liquid categories is more important than assumptions about drop size distributions.

Corresponding author: Adele L. Igel, aigel@ucdavis.edu

## Abstract

Warm rain collision coalescence has been persistently difficult to parameterize in bulk microphysics schemes. Here we use a flexible bulk microphysics scheme with bin scheme process parameterizations, called AMP, to investigate reasons for the difficulty. AMP is configured in a variety of ways to mimic bulk schemes and is compared to simulations with the bin scheme upon which AMP is built. We find that the biggest limitation in traditional bulk schemes is the use of separate cloud and rain categories. When the drop size distribution is instead represented by a continuous distribution with or without an explicit functional form, the simulation of cloud-to-rain conversion is substantially improved. We find that the use of an assumed double-mode gamma distribution and the choice of predicted distribution moments do somewhat influence the ability of AMP to simulate rain production, but much less than using a single liquid category compared to separate cloud and rain categories. Traditional two category configurations of AMP are always too slow in producing rain due to their struggle to capture the emergence of the rain mode. Single category configurations may produce rain either too slowly or too quickly, with too slow production more likely for initially narrow droplet size distributions. However, the average error magnitude is much smaller using a single category than two categories. Optimal moment combinations for the single category approach appear to be linked more to the information content they provide for constraining the size distributions than to their correlation with collision-coalescence rates.



## Plain Language Summary

Weather and climate forecast models have always struggled to simulate the production of rain from warm, shallow clouds. As a result, these models often cannot reproduce observed surface rain rates and cloud radiative forcing. Here, we investigate why this rain production is so difficult for the bulk microphysics schemes in these models. We address a number of possibilities: the drop size distribution assumption, the choice of predicted cloud and rain properties, and the decision to treat cloud and rain drops as separate categories. We find the latter is most likely to be the source of difficulty. Most existing models choose to distinguish between cloud and rain drops, which necessitates methods to transfer mass (and other properties) from the cloud category to the rain category during rain production. We find that if we instead use a single liquid drop category that contains both cloud and rain drops, we can substantially improve the prediction of rain formation. This is true even when we use the same total number of predicted properties in each approach. These results imply that we could improve rain production in models without any additional computational cost by moving to a single liquid drop category in bulk microphysics schemes.


## 1 Introduction

The representation of warm phase collision-coalescence in global weather and climate models (GCMs) is notoriously challenging and is often a large source of disagreement between models and observations. Several studies have found that GCMs produce too much light rain and potentially not enough heavy rain (Jing et al., 2017; Kay et al., 2018) and that these errors can lead to a substantial bias in the cloud radiative forcing (Mülmenstädt et al., 2021). Due to these known issues, improving the parameterization of collision-coalescence rates is an active area of current research.

Liquid water in the bulk microphysics schemes used in cloud-resolving and global climate models is typically represented by an artificial division into two categories, one for small cloud droplets and one for larger rain drops. This is based on widespread observations that the liquid water mass distribution is often bimodal and the idea that cloud and rain drops generally grow by different processes (vapor diffusion for the former and collision-coalescence for the latter). With a few exceptions (Szyrmer et al., 2005; Y. L. Kogan & Belochitski, 2012; Morrison et al., 2020), the drop size distribution (DSD) of each category is assumed to follow a theoretical distribution function, most commonly the gamma distribution (e.g. Clark, 1974; Khairoutdinov & Kogan, 2000; Morrison et al., 2005; Seifert & Beheng, 2001; Walko et al., 1995). Bulk schemes then predict one to three moments of the DSD, or integral quantities of these functions. Most commonly these are the 0th moment (M0) of the size distribution, which corresponds to the total number concentration, the 3rd moment (M3), which is proportional to the mass mixing ratio, and possibly the 6th moment (M6), which is proportional to the radar reflectivity factor.

With this basic framework for representing cloud liquid water in bulk schemes, warm phase collision-coalescence is forced to be divided into two main processes, namely autoconversion, the self-collection of cloud droplets to make rain, and accretion, the collection of cloud droplets by raindrops. Some schemes also include self-collection of cloud droplets and/or rain which remain in their respective categories. Autoconversion in particular has been difficult to parameterize. The most common type of autoconversion parameterization is the Kessler-type. These parameterizations allow autoconversion only after some threshold, often in terms of mass mixing ratio or mean droplet size, has been reached. Liu and Daum (2004) provide a summary of many of these parameterizations. Others, such as Seifert and Beheng (2001) and Lee and Baik (2017) make simplifying assumptions to the stochastic collection equation to arrive at analytic equations for autoconversion and accretion rates. Some success has also been found with empirically derived equations or lookup tables based on bin model rates (Berry & Reinhardt, 1974; Feingold et al., 1998; Khairoutdinov & Kogan, 2000; Y. Kogan, 2013) or with a combination of analytic and empirical approaches (Zeng & Li, 2020). Finally, machine learning has also been employed to develop new parameterizations based on bin or Lagrangian model data (Chiu et al., 2021; Seifert & Rasp, 2020).

Seifert and Rasp (2020) and Chiu et al. (2021) both suggest that autoconversion parameterizations may be improved by incorporating information about rain. While rain has no direct impact on autoconversion by definition, its inclusion improves the machine-learned parameterizations and is shown to be strongly related to the cloud droplet size distribution width in idealized conditions (Zeng & Li, 2020). Even with these recent efforts to improve autoconversion, Seifert and Rasp (2020) propose that a fundamental problem with autoconversion parameterizations generally may be that autoconversion is ill-posed for small, narrow cloud droplet size distributions. Prediction of higher-order moments may be helpful as shown in Igel (2019). Careful tuning has alleviated this problem in many parameterizations but often at the cost of overpredicting autoconversion rates early and underpredicting them later in the rain formation process. This tuning is consistent with the known overproduction of light rain in GCMs (Jing et al., 2017; Kay et al., 2018). Another persistent issue is that both analytic and empirical parameterizations must make some assumption about the cutoff size that distinguishes cloud droplets from raindrops. Berry and Reinhardt adopted a radius of 40 $\mu$m as the cutoff size based on simulations and that value has been adopted by most others (e.g. Lee & Baik, 2017; Seifert & Beheng, 2001). Khairoutdinov and Kogan (2000) used a cutoff radius of 25 $\mu$m. Regardless, observations show that the local minimum of the liquid DSD can be variable and as small as about 20 $\mu$m (Austin et al., 1995; Ferek et al., 2000; Sinclair et al., 2021). Such a discrepancy between the parameterizations and observations may be another reason for the difficulty in simulating warm-rain formation using bulk schemes.

In summary, the struggle to predict collision coalescence in bulk schemes has many potential sources. Namely,

1. Poor choice of predicted moments (e.g. 0th, 3rd, and 6th are not the ideal combination)
2. The use of artificially separate cloud and rain modes
3. The use of assumed analytic functions for the DSDs
4. The use of a limited number of predicted moments to describe the DSD in bulk schemes rather than the use of a resolved DSD in bin schemes
5. Fundamental lack of knowledge of the collision-coalescence rates in nature

In this study we aim to assess reasons 1-4. We will do so by employing a flexible, hybrid bulk-bin scheme called the Arbitrary Moment Predictor (AMP; Igel, 2019). AMP and the simulations we run are described in Section 2. A series of tests with AMP in a variety of configurations are presented and discussed in Section 3. Insights about optimal moment combinations are given in Section 4. Conclusions are presented in Section 5.

# 2 AMP Description

## 2.1 AMP Overview

This study makes use of the Arbitrary Moment Predictor (AMP) which was first described in Igel (2019). AMP uses the cloud microphysical parameterizations of the Hebrew University spectral bin model (Khain et al., 2004). However, rather than saving the explicit size distribution between time steps, AMP calculates a limited number of integral moments of the size distribution and saves only these for use in the next time step. At the beginning of a time step, an explicit DSD is obtained such that the integral moments of the explicit DSD function are consistent with the moments predicted by AMP. This explicit DSD is then fed to the microphysical parameterizations of the spectral bin model. Updated integral moments are calculated and the process continues at the next time step. The number of integral moments and the values of the predicted moments are selected by the user. As such, AMP is a bulk microphysics scheme in that it only predicts bulk quantities of the size distribution, but it is a bin microphysics scheme in that it uses bin parameterizations to evolve those bulk quantities.

Due to its design, AMP is a useful tool for understanding the inherent limitations of bulk schemes compared to bin schemes. In this paper, we will compare AMP simulations with simulations run with the bin parameterization on which AMP is built (BIN). Any differences that arise between AMP and BIN are therefore due solely to the representation of the size distribution and not due to differences in the parameterization of the microphysical processes. In this study, we will use three different versions of AMP. These are described in the next three subsections. To easily distinguish among the basic AMP configurations, AMP configured with separate cloud and rain categories will be referred to as AMP-CR; AMP with a single liquid category, an assumed double-mode gamma distribution, and prediction of full moments will be referred to as AMP-F; and AMP with a single category, nonparametric distributions, and prediction of full moments will be referred to as AMP-NP. Here, full moments refer to moments calculated using all bins. Partial moments will refer to moments calculated using only a subset of bins corresponding to either cloud droplets or rain drops.

## 2.2 AMP-CR

In Igel (2019), the liquid size distribution in AMP is split into two categories corresponding to cloud droplets and rain drops. Integral moments of the two categories are predicted separately. A gamma size distribution ($N(D)$) is assumed for both categories:

$$N(D) = N_0 D^{\nu-1} e^{-\lambda D} \tag{1}$$

where $N_0$, $\nu$ and $\lambda$ are the intercept, shape and slope parameters of the distribution. In the double-moment (2M) configuration, $\nu$ is specified. In the triple-moment (3M) configuration, all three parameters are determined from the prognosed moments. At the start of each time step and for each category, the prognosed moments are used to find the parameters $N_0$, $\nu$ (for 3M only), and $\lambda$ such that the moments of $N(D)$ integrated over the bins corresponding to the category (separated into cloud and rain using a threshold radius) equal the prognosed values of the moments. $N_0$ can be solved for through normalization of the DSD. There are no analytical equations to solve generally for $\nu$ and $\lambda$ when the distributions are incomplete. We use iterative procedures with a first guess based on look up tables. It is possible that no set of distribution parameters is consistent with the predicted moments. In this case, AMP-CR always ensures that the distribution parameters give the correct mass such that mass conservation in the model is guaranteed. AMP-CR next tries to ensure that number concentration is conserved. Once the parameters for the two categories have been found, the resulting DSDs are concatenated to produce a single DSD that is fed to the process parameterizations. After the process rate calculations to evolve the DSD, partial moments are calculated over the bins corresponding to cloud droplets and rain drops to update the values of the prognosed moments. We use a threshold radius of 40 $\mu$m to distinguish between cloud and rain. Full details of AMP-CR are given in Igel (2019).

### 2.3 AMP-F

AMP-F is similar to AMP-CR, but rather than splitting the distribution in two parts, AMP-F uses a single liquid category that is represented by a double-mode gamma DSD:

$$N(D) = N_1 D^{\nu_1 - 1} e^{-\lambda_1 D} + N_2 D^{\nu_2 - 1} e^{-\lambda_2 D} \tag{2}$$

where subscripts "1" and "2" indicate the distribution parameters for each mode. We use either four (4M) or six (6M) prognosed moments. In the 4M configuration, $\nu_1$ and $\nu_2$ are specified; in the 6M configuration, all parameters are diagnosed from the moments. Like $N_0$ in AMP-CR, $N_1$ and $N_2$ can both be solved for through normalization of the DSD. The remaining parameters are again solved for through iterative procedures. As with all versions of AMP, the resulting DSD is then fed to the parameterizations, and updated integral moments are calculated. In AMP-F, the full moments are calculated over all liquid bins. For diagnostic purposes, we also calculate partial moments over the cloud and rain bins separately, again using a 40 $\mu$m threshold radius. However, these calculations are purely diagnostic and do not impact the simulations using AMP-F.

### 2.4 AMP-NP

Finally, rather than using a gamma function or any other analytic function, we developed a single category approach that makes use of nonparametric size distributions; that is, it does not assume any explicit functional form a priori for the DSD. This approach is related to the general problem of reconstructing a distribution from a set of its moments. For AMP, we are interested in reconstructing a discretized DSD comprising $L$ bins. For a mass doubling bin grid (consistent with the discretzied DSDs used in AMP), the first bin contains a number of droplets $n_0$ having mass $m_0$, the next bin contains $n_1$ droplets of mass $2m_0$, the next bin contains $n_2$ droplets of mass $2^2 m_0$, and in general $n_l$ is the number of droplets of mass $2^l m_0$. The 3rd moment, proportional to total mass, can be expressed as $\mathrm{M}3 = [6m_0/(\rho_w \pi)] \sum_{l=0}^{L-1} 2^l n_l$, where $\rho_w$ is the density of water. We will chose our units such that $6m_0/(\rho_w\pi) = 1$ to nondimensionalize this expression. We can then generalize to give the pth moment of the distribution as $\mathrm{M}p = \sum_{l=0}^{L-1} n_l \left(2^{p/3}\right)^l$. To calculate several moments of the distribution $\vec{M} = (\mathrm{M}p_1, \mathrm{M}p_2, \dots)$, we can express this as a matrix multiplication $\vec{M} = V\vec{n}$ where the number of droplets in each bin is denoted as vector $\vec{n} = (n_0, n_1, \dots)$ and

$$V = \begin{pmatrix} 1 & 2^{p_1/3} & \left(2^{p_1/3}\right)^2 & \cdots & \left(2^{p_1/3}\right)^{L-1} \\ 1 & 2^{p_2/3} & \left(2^{p_2/3}\right)^2 & \cdots & \left(2^{p_2/3}\right)^{L-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \end{pmatrix} \tag{3}$$

The matrix in the above expression belongs to a class of matrices known as Vandermonde matrices, which are of the form:

$$\begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{L-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{L-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \end{pmatrix} \tag{4}$$

A Vandermonde matrix is used to evaluate a polynomial at an ordered set of points $(x_1, x_2, \dots)$. Thus, calculating the pth moment of the discretized distribution is equivalent to evaluating the polynomial $n_0 + n_1 x + \cdots + n_{L-1} x^{L-1}$ at the point $x = 2^{p/3}$. A square Vandermonde matrix is invertible if and only if all values $(x_1, x_2, \dots, x_L)$ are distinct. Since this is the case for a discretized distribution with a fixed size or mass grid, it means that if we are given $L$ moments of a size distribution, we can in principle exactly reconstruct the distribution.

Since AMP predicts a limited set of moments with the number of moments $< L$, we need an additional closure assumption to obtain distributions. This is done by using multi-dimensional lookup tables built from a large set of reference binned DSDs, which is further described below. Each dimension of the table corresponds to a moment, and the reference DSDs are averaged over sections of the multi-dimensional space of these moments (using the median instead of mean does not appreciably change results). DSDs are then obtained from input sets of predicted moment values by interpolating the lookup table reference DSDs over this multi-moment space.

In principle, this approach converges to the square Vandermonde matrix problem, and hence to the exact discretized DSD, as the number of predicted moments approaches $L$. This is not true if an explicit DSD functional form is assumed a priori. However, convergence would be difficult to demonstrate in practice because Vandermonde matrices are notoriously ill-conditioned except when $L$ is small (say, $< 15$). This means that small errors in the moments (e.g., owing to machine roundoff) can produce large errors in the reconstructed size distributions. This can be improved by careful choice of moments (including negative moments) and the use of a matrix preconditioner, but even in this case relative errors of $O\left(10^{-8}\right)$ introduced to a single moment can lead to large oscillations in the reconstructed DSD for $L > 30$.

In AMP-NP, the number of predicted moments can be set by the user. In our study we use four moments. The choice of moment orders is also flexible. Here we test three different sets: 1) M0, M3, M6, M9, 2) M0, M3, M4, M5, and 3) M0, M3, M4, M9. To generate the lookup tables, moment values for the above moment sets are calculated for each reference DSD. For all three cases above, the reference DSDs (over 34 mass doubling bins) are first normalized by M0, which effectively reduces the required lookup table dimensionality by one. Thus, the dimensionality of the lookup table for each case is three, corresponding to the other three predicted moments besides M0. The first dimension of the lookup table for all three cases is then chosen as the normalized $\mathrm{M3}^* = \mathrm{M3}/\mathrm{M0}$ ($^*$ denotes moments normalized by M0). For the other lookup table dimensions we employ non-dimensional moments (denoted by #). For this study, we define the following non-dimensional moments for each moment set above:

1) M0, M3, M6, M9: $\mathrm{M6}^{\#} = \mathrm{M3}^*/\mathrm{M6}^{*2}$, $\mathrm{M9}^{\#} = \mathrm{M3}^*\mathrm{M6}^*/\mathrm{M9}^*$

2) M0, M3, M4, M5: $\mathrm{M4}^{\#} = \mathrm{M4}^*/\mathrm{M3}^{*4/3}$, $\mathrm{M5}^{\#} = \mathrm{M5}^*\mathrm{M4}^*/\mathrm{M3}^{*3}$

3) M0, M3, M4, M9: $\mathrm{M4}^{\#} = \mathrm{M4}^*/\mathrm{M3}^{*4/3}$, $\mathrm{M9}^{\#} = \mathrm{M9}^*\mathrm{M3}^*/\mathrm{M4}^{*3}$

Using non-dimensional moments greatly reduces the range of moment values of the reference DSDs (for example, $\mathrm{M6}^{\#}$ varies by about 7 orders of magnitude versus 18 orders of magnitude for $\mathrm{M6}^*$ for the set of reference DSDs described below). This facilitates interpolation over the lookup table for the DSD retrieval. Normalized DSDs stored in the lookup table are the mean of all reference DSDs having moment values falling within a given section of the moment space. The three-dimensional lookup tables consist of $400 \times 200 \times 100$ total sections. Sections are spaced logarithmically given the wide range of moments values even when normalized and non-dimensionalized. Given input values for the set of predicted moments in AMP, DSDs are retrieved from the lookup tables by multi-dimensional linear interpolation in logarithmic space of the moments. The interpolated normalized DSDs are then multiplied by M0 to obtain the full DSDs.

The reference DSDs used to generate the lookup tables include 3,450,230 individual DSDs. These come from previous simulations of shallow and deep convection using the Hebrew University spectral bin model within the Regional Atmospheric Modeling System (Cotton et al., 2003). As such, the DSDs include a variety of distribution shapes that span the full multi-moment space well. Similar to AMP-F, the full moments in AMP-NP are calculated over all liquid bins. However, for purely diagnostic purposes we calculate partial moments over the cloud and rain bins separately, again using a 40 $\mu$m threshold radius.

### 2.5 Simulations

In this study, collision-coalescence is the only microphysical process allowed in AMP and BIN. We ran several test suites of collision-coalescence with a wide variety of initial conditions. The initial conditions are the same as described in Igel (2019), namely, we vary the initial mass mixing ratio from 1 g kg$^{-1}$ to 5 g kg$^{-1}$ in increments of 1 g kg$^{-1}$, we successively double the initial droplet concentration from 50 cm$^{-3}$ to 1600 cm$^{-3}$, and we vary the shape parameter from 1 to 15 in increments of 2. These values are used to initialize a single-mode gamma distribution (Eq. 1) at time zero. In the case of AMP-NP simulations, the initial distribution is created, its moments are calculated, and a moment-matching distribution is found in the look-up table for the initial conditions. BIN and each configuration of AMP are run with these 240 different initial conditions for 30 minutes. Note that in all cases, each AMP or BIN simulation pair begins with identical initial size distributions. Simulations with initial conditions which fail to fully convert the initial cloud water to rainwater in BIN are discarded. Doing so excludes 26 sets of initial conditions and leaves 214 sets for analysis.

## 3 AMP Performance

### 3.1 a. Standard Double-Moment Performance

We first show results for AMP-CR run in a standard two-moment bulk scheme configuration. Specifically, AMP-CR is configured to predict the 0th and 3rd moments of the cloud and rain modes (c03-r03). The AMP-CR simulations are compared to the reference BIN simulations. Note the comparison is done in the same way for other AMP configurations. Consider a single simulation pair for AMP and BIN. First, we normalize the time ($t_n$). Normalized time zero is the simulation start. Normalized time $t_n = 1$ is defined as the time when 99% of the cloud water has been converted to rainwater in the BIN simulation. The evolutions of all moments in both the BIN and AMP simulations are re-gridded to the normalized time. Next, the moment values in both the BIN and AMP simulations are normalized by the maximum value in the BIN simulation occurring between $t_n$ of 0 and 1. This procedure is repeated for each pair of BIN and AMP simulations. Finally, the simulation pairs are grouped into terciles based on the difference in cloud droplet normalized M3 between BIN and AMP when 50% of the water mass has been converted to rain in each BIN simulation. Normalized evolutions within each error tercile are averaged together.

Figure 1a shows the normalized evolutions for each error tercile of the 3rd, 0th, and 6th moments of the cloud droplet distribution, and the 0th and 6th moments of the raindrop distribution. Note that normalized M3 of the raindrop distribution is one minus normalized M3 of the cloud droplet distribution and that the 6th moments are purely diagnostic. There are several features of the AMP-CR performance to notice. In the first tercile, the difference between AMP-CR and BIN is nearly zero for all moments (purple dotted line). In these cases, rain is made relatively quickly. There are often large cloud droplets or small rain drops already present (notice the non-zero values of rain M0 present at the start of the simulations in Fig. 1a4) and little autoconversion is required before accretion becomes the dominant rain formation process. On the other hand, in the third tercile (gold lines), AMP-CR struggles to convert cloud water to rainwater and rain production is severely delayed. Figure 1a4 shows that essentially no raindrops are created by AMP-CR in this tercile (gold solid line).

### 3.2 Triple-Moment Performance

Perhaps the most obvious way to improve the accuracy of a bulk scheme is to predict more moments. Figure 1b shows the normalized moment evolutions for a standard triple-moment bulk AMP-CR configuration in which the 6th moments of both the cloud and rain drop distributions are predicted in addition to the 0th and 3rd moments (c036-r036). The performance of triple-moment AMP-CR compared to BIN is somewhat improved over that with double-moment AMP-CR. The third tercile (gold), in which AMP struggles the most to produce rain quickly, now converts about 25% more cloud water to rainwater as in c03-r03 by $t_n = 1$. That said, 30% of the cloud water still remains on average in c036-r036 when

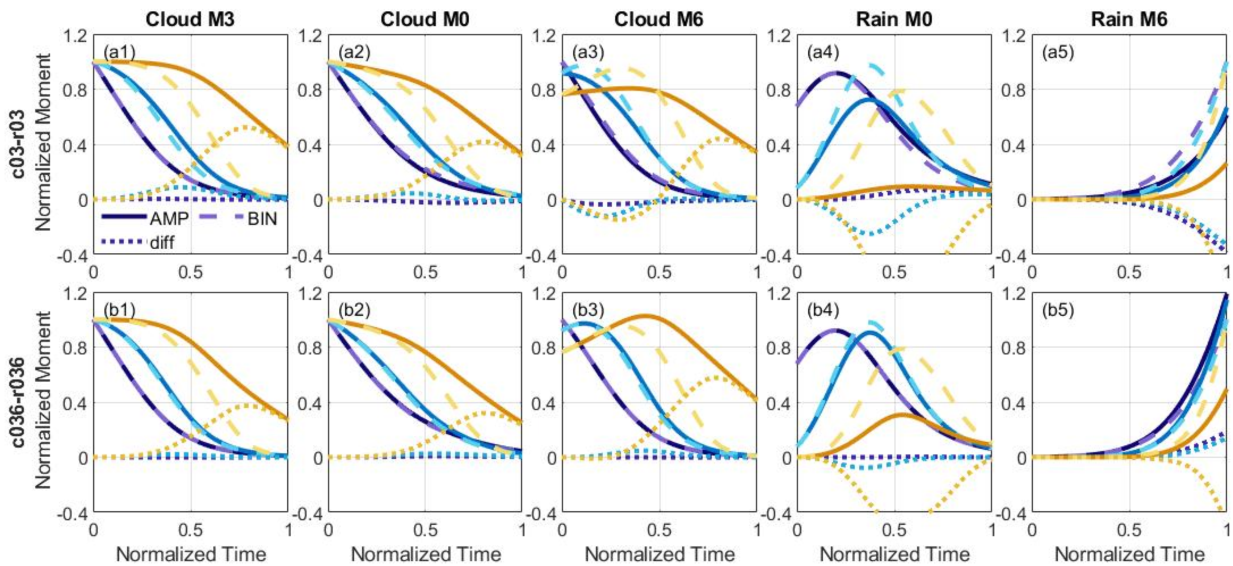


**Figure 1.** Normalized evolutions of distribution moments in AMP and BIN and their differences for (a) AMP-CR c03-r03 and (b) AMP-CR c036-r036. The specific distribution moments are indicated in the column titles. The simulations are sorted into tercile groups based on the difference in cloud mass between AMP and BIN (see the main text) and the average evolution is shown for each group. Note that tercile groups are different for each AMP configuration. Solid lines show results for AMP, long dashed lines for BIN, and dotted lines show the difference. Gold lines show results for the tercile group with the largest errors, blue lines for the middle group, and purple lines for the group with the smallest errors.

BIN has completely converted the cloud water to rain. Rain nM0 shows perhaps the biggest improvement, but the AMP values are still too low by about a factor of 2. Even better performance would be preferred.

An idea that has been suggested recently is that cloud processes may be better represented if different distribution moments were predicted. This idea was explored in Igel (2019) using AMP. They found that the mass evolution during collision-coalescence could be better represented by predicting the 3rd and 8th moments of the cloud droplet distribution rather than the 0th and 3rd. For a triple-moment configuration, predicting the 0th, 3rd, and 8th cloud droplet moments was shown to be best. Different rain moment combinations were not tested. Figure 2 shows results for AMP-CR configured to predict the 0th, 2nd, and 3rd (032) or 0th, 3rd, and 8th (038) moments of the cloud and rain distributions. Consistent with Igel (2019), changing the predicted cloud moments does impact the evolution of collision-coalescence with combinations c038-r032 and c038-r038 (Fig. 2b and 2c) producing a substantial improvement over c032-r032 (Fig. 2a) in terms of cloud M3, cloud M0, and rain M0. The combination of predicted rain moments has very little influence on the moment evolutions except for rain M6. Overall c038-r038 (Fig. 2c) performs marginally better than the standard combination of c036-r036 (Fig. 1b), but predicting different moments does not appear to be a promising way to improve the representation of autoconversion and accretion.

### 3.3 Single Liquid Category Performance

Another idea that has been proposed in the past is to use a single category for cloud and rainwater. Clark (1976) and Clark and Hall (1983) used this approach. They assumed that the liquid size distribution could be described by the sum of two lognormal PDFs. The Clark scheme was revived and modernized with machine-learned moment tendencies by Rodríguez Genó and Alfonso (2022). All three stud-

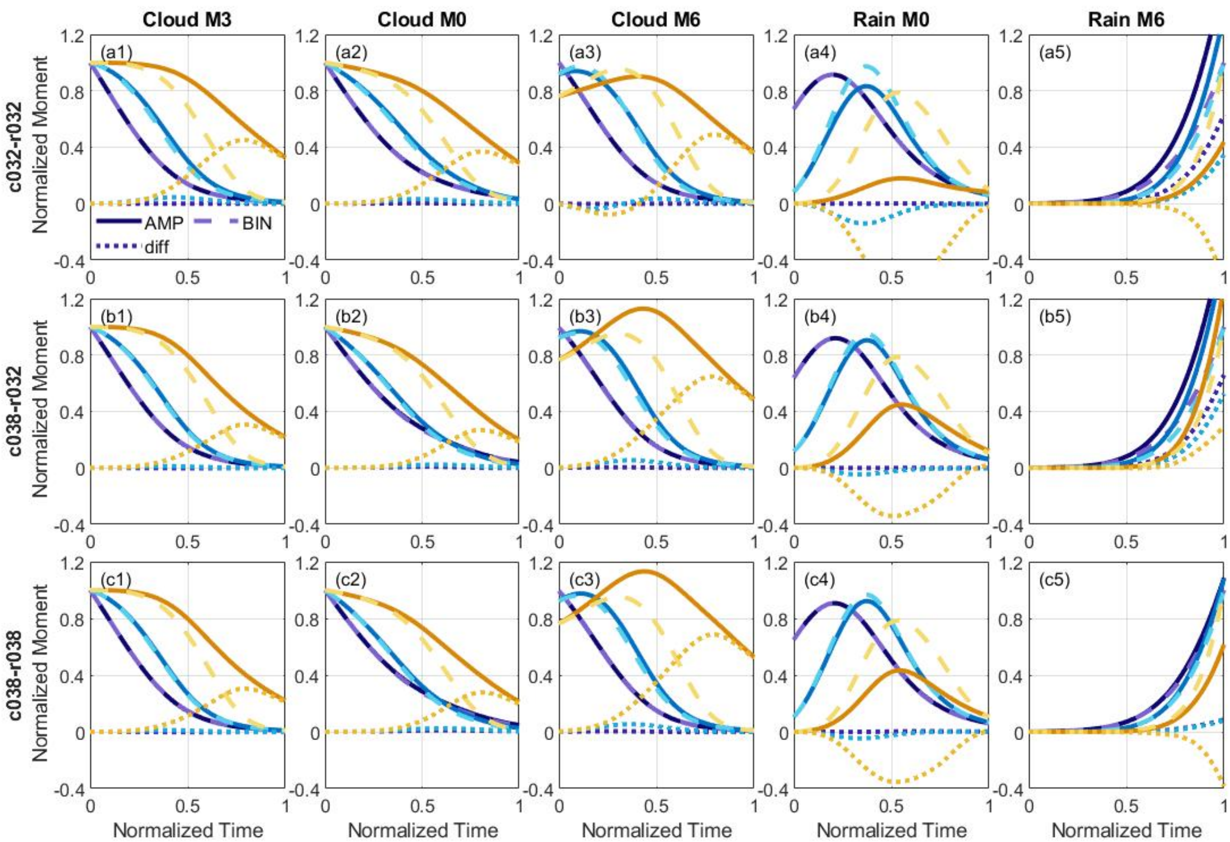


**Figure 2.** As in Figure 1, except for various configurations of 3M AMP-CR: (a) c032-r032, (b) c038-r032, and (c) c038-r038.

ies found the prediction of six total moments could adequately simulate the collision-coalescence process. Y. L. Kogan and Belochitski (2012) developed a full warm-phase bulk microphysics scheme with a single liquid category. They predicted five total moments, made no assumptions about the underlying size distribution, and formulated moment tendency equations through a combination of theory and empirical fitting to bin microphysics model process rates. Their simulations of non-precipitating and drizzling stratocumulus clouds with the total moment scheme were comparable to simulations with a traditional, two-category scheme.

Motivated by these previous studies, we ran AMP-F with a double-mode gamma distribution with both four and six predicted full moments. In the 4M and 6M configurations, the choice of predicted full moments is not obvious. We ran a large number of predicted full moment combinations; all included both the 0th and 3rd moments. Partial moments of the "cloud" and "rain" distributions were diagnosed by integrating over the appropriate bins (40 $\mu$m radius threshold) at the end of each time step in order to facilitate the same analysis shown in Figures 1 and 2. Note that the evolution of full M0 is nearly identical to that of cloud M0 and likewise full M6 is nearly identical to rain M6.

First, we show results from the 4M AMP-F simulations in Figure 3. Three different predicted moment combinations are shown: in addition to the 0th and 3rd, the 6th and 9th (f0369, Fig. 3a), the 4th and 5th (f0345, Fig. 3b), and the 4th and 9th (f0349, Fig. 3c). (Note that some noise appears in Fig. 3b3-5 and 3c3-5 toward the end in the AMP-F simulations. This occurs when AMP fails to find distribution parameters that are consistent with the predicted moments. In the results shown here, the problem is minor. For other moment combinations, the problem is a major one.) The difference between AMP-F and BIN for all three terciles is substantially smaller for all moment combinations compared to the previously best AMP-CR combination, c038-r038 (Fig. 2c). AMP-CR nearly always produced rain too slowly; the lowest error tercile for AMP-F corresponds to rain production that is too fast (Fig. 3a-c1) and seems to correspond to cases in which rain production is initially slow. These results are particularly remarkable given that the 3M, two category AMP-CR simulations in Fig. 2 predict two additional quantities than the 4M, single category AMP-F simulations in Fig. 3. When Fig. 3 is compared to Fig. 1a, in which case both sets of AMP simulations use the same total number of predicted moments, the improvement with the use of a double-mode distribution becomes even more noteworthy.

Errors in the rain moments (Fig. 3 columns 4-5) are generally larger than errors in the cloud droplet moments in 4M AMP-F (Fig. 3 columns 2-3). Rain moment errors are on average smaller than those for the double-moment AMP-CR configuration (Fig. 1a) and comparable to somewhat worse than those for the triple-moment AMP-CR configurations (Fig. 1b and 2).

The evolution of the mass distribution for a sample initial condition is shown in Figure 4 to better understand the differences between the simulations with various AMP configurations and BIN. At 300 seconds, a small amount of rain has formed in BIN. AMP-CR c03-r03 has totally failed to produce this rain and has a distribution that is nearly identical to the initial distribution. AMP-CR c038-r038 has produced some rain by first increasing the mean size of the cloud droplet mode relative to AMP-CR c03-r03. The increased mean size of the cloud droplet mode is even more apparent at 450 and 600s. So, while the effect of creating some rain is more consistent with the reference BIN distribution, the way in which it has done so is inconsistent with the BIN simulation. The single-category, double-mode AMP-F f0349 simulation produces a distribution that most closely matches BIN at 300s. In all three times shown, AMP-F f0349 maintains the mean size of the cloud droplet mode. It does struggle to capture the shape of the rain mode, which leads to errors in the rain number concentration and 6th moment. However, such a result is unsurprising given that the shape parameters are held constant. Its performance is still greatly improved compared to AMP-CR c03-r03. These results strongly suggest that the conversion of cloud water to rainwater could be substantially better simulated by the use of a single liquid hydrometeor category.

Next, we ran AMP-F with six predicted moments such that no parameters of the double-mode gamma distribution were fixed. Unsurprisingly, we find that the performance is improved further and we find almost perfect agreement in the mass and cloud droplet concentration evolutions between 6M AMP-F and BIN, and to a lesser extent with the rain M6 (Figure 5). Agreement for cloud M6 and rain M0 is also im-

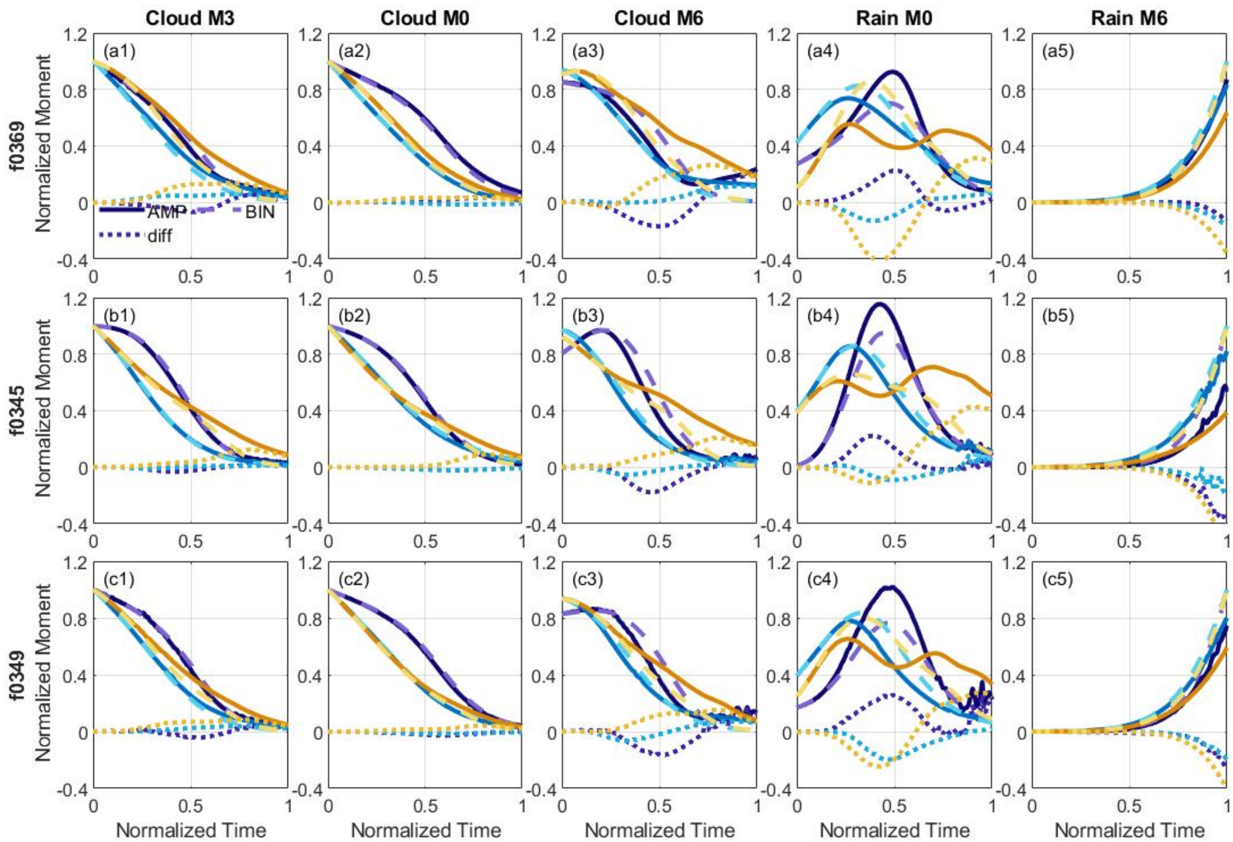


**Figure 3.** As in Figure 1, except for various configuration of 4M AMP-F: (a) f0369, (b) f0345, and (c) f0349.

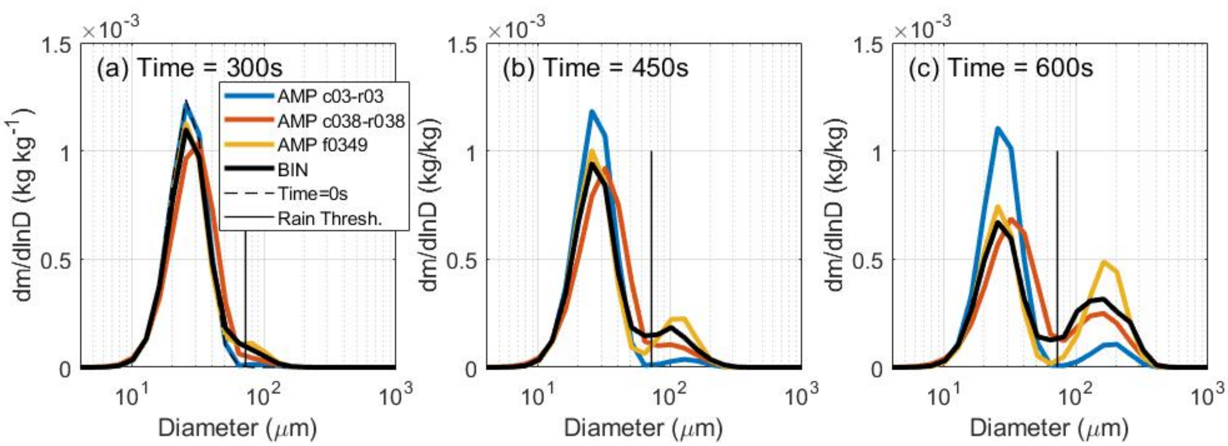


**Figure 4.** Sample evolution of mass distributions with BIN (thick black line) and three different configurations of AMP (colored lines) as indicated in the legend. The thin dashed line shows the initial distribution in (a), which is nearly overlayed by the AMP-CR c03-r03 simulation. The thin solid line indicates the diameter which separates cloud and rain.

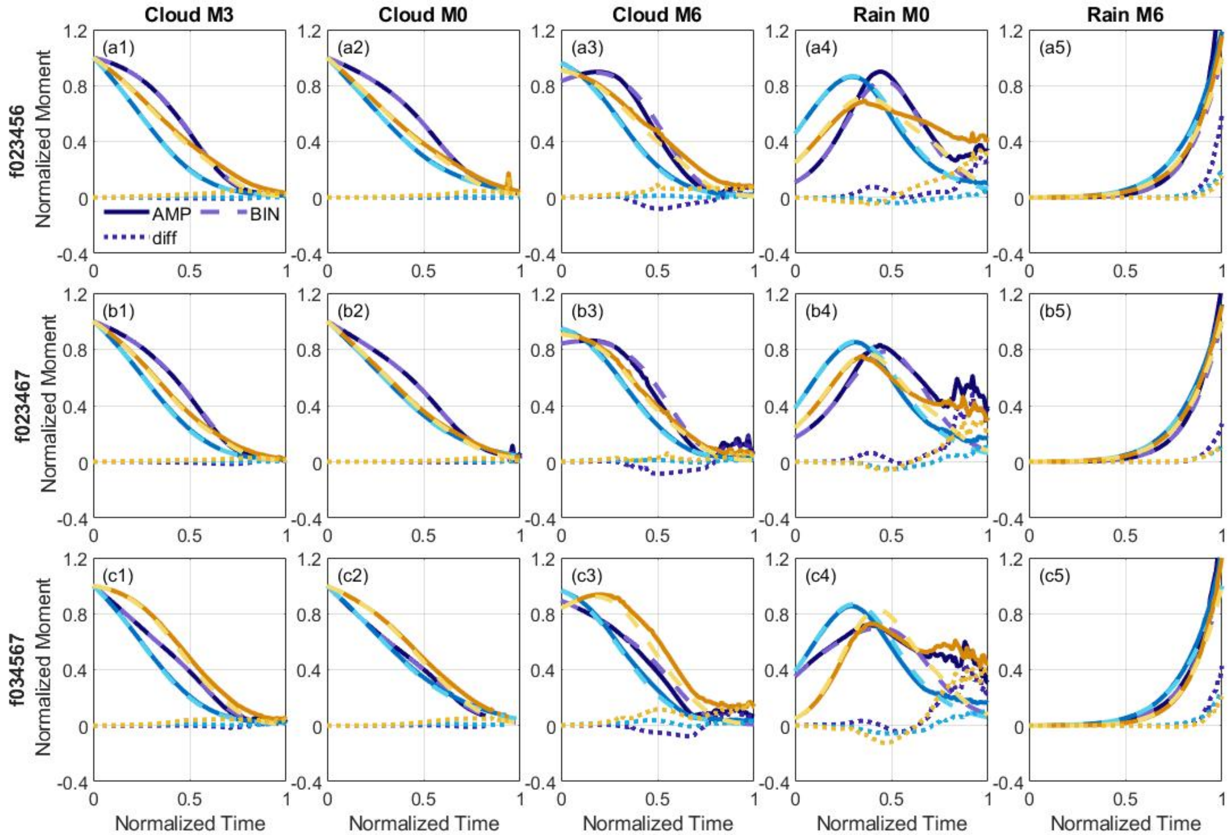


**Figure 5.** As in Figure 1, except for various configuration of 6M AMP-F: (a) f023456, (b) f023467, and (c) f034567. Noise near the end of the evolutions of cloud M0, cloud M6, and rain M0 appears due to the inability to find DSD parameters given the predicted moment values.

proved, although these most clearly show the noise that develops toward the end of some simulations due to the inability of AMP-F to find DSD parameters from the prognosed moments.

These simulations are particularly useful for understanding why AMP-F can perform better than AMP-CR. Figure 6 shows the evolution of the shape parameters and mean diameters of the two modes in 6M AMP-F (f023467) and the cloud and rain categories in 3M AMP-CR (c038-r038) from $t_n = 0$ to $t_n = 0.8$ for 25 of the 100 worst performing AMP-CR simulations (the remaining time and simulations are omitted for clarity). Distribution parameters evolve from the "o" to the "x". The two modes in 6M AMP-F (Fig. 6a-b) clearly correspond well to the cloud and rain categories in 3M AMP-CR (Fig. 6c-d), but there are some noticeable differences between the two AMP configurations.

The first (cloud droplet) mode develops quite differently in 3M AMP-CR and 6M AMP-F. In 6M AMP-F, the droplet distributions (Fig. 6a) have a decrease in shape parameter (meaning DSDs become wider), but often later have a substantial increase in shape parameter. The mean diameter consistently decreases, but the total decrease may not be especially large. In contrast, 3M AMP-CR monotonically decreases cloud droplet shape parameter in all simulations and usually predicts a larger change in the mean diameter (Fig. 6c). As also seen in Figure 4, these evolutions suggest that 3M AMP-CR artificially widens the cloud droplet mode because self-collection of droplets produces larger cloud droplets; this increase of mass near the autoconversion threshold results in an increase of the 8th moment and therefore larger diagnosed distribution widths. 6M AMP-F avoids this artificial widening by using the second mode to

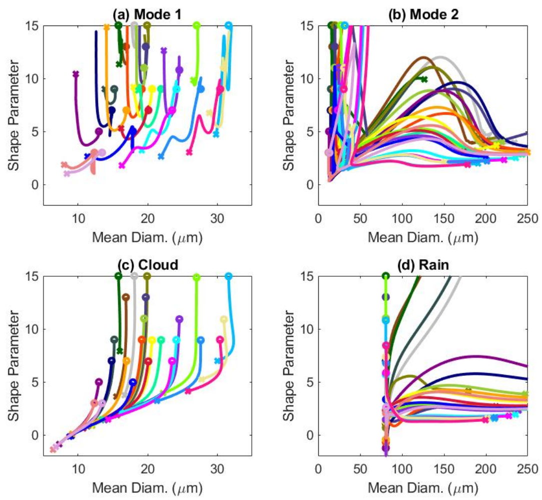


**Figure 6.** Simultaneous evolution of the shape parameter and mean diameter in (a) the smaller mode and (b) the larger mode in AMP-F f023467; and (c) the cloud and (d) rain modes in AMP-CR c038-r038. Time progresses from the "o" to the "x". Each colored line is a separate simulation in the initial condition ensemble.

capture the earliest collisions. This is evidenced by the small initial mean diameters of its second mode (Fig. 6b), much smaller than would usually be considered rain. Once the 6M AMP-F second mode (Fig. 6b) diameters do reach traditional rain drop sizes, the shape parameter tends to increase and then decrease as the mode develops, whereas 3M AMP-CR typically maintains a much more constant shape parameter (Fig. 6d). Overall, this analysis suggests that a key reason that traditional bulk schemes struggle with autoconversion is that the early stages of rain production are not well represented by predefined cloud and rain categories.

### 3.4 Performance with Nonparametric Distributions

Although 4M AMP-F can simulate the rain production well for nearly all tested initial conditions, its use of fixed shape parameters limits its flexibility in representing natural distribution shapes. For this reason, we developed AMP-NP which uses nonparametric size distributions as described in Section 2d. We ran AMP-NP with the same three combinations of predicted moments as 4M AMP-F and the results are shown in Figure 7. Qualitatively, the results are similar to 4M AMP-F (Figure 3) and 3M AMP-CR (Figure 2) and are markedly better than 2M AMP-CR (Figure 1a) despite having the same number of total predicted variables (four).

One notable difference between AMP-NP and AMP-F or AMP-CR is that AMP-NP is much more sensitive to the choice of predicted moments, likely because AMP-NP is not constrained by a functional form for the DSD. AMP-NP is therefore most useful for discussing the optimal combination of full prognostic moments. Configuration np0369 is clearly better than either np0345 or np0349. A likely reason is that orders of the predicted moments, 0369, are more separated than 0345 and 0349. Moments closer to one another become more strongly correlated and thus do not provide as much independent information to reconstruct the DSDs, as discussed in Morrison et al. (2019). Low order moments will give more information about the cloud droplet distribution and higher order moments will give more information about the raindrop size distribution. By having both low and high order moments that are sufficiently separated, np0369 is arguably the best AMP-NP configuration for predicting both cloud and rain partial moments. Another possibility is that certain moments correlate better with the process rates. The reasons for better performance of some moment combinations will be explored further in Section 4.

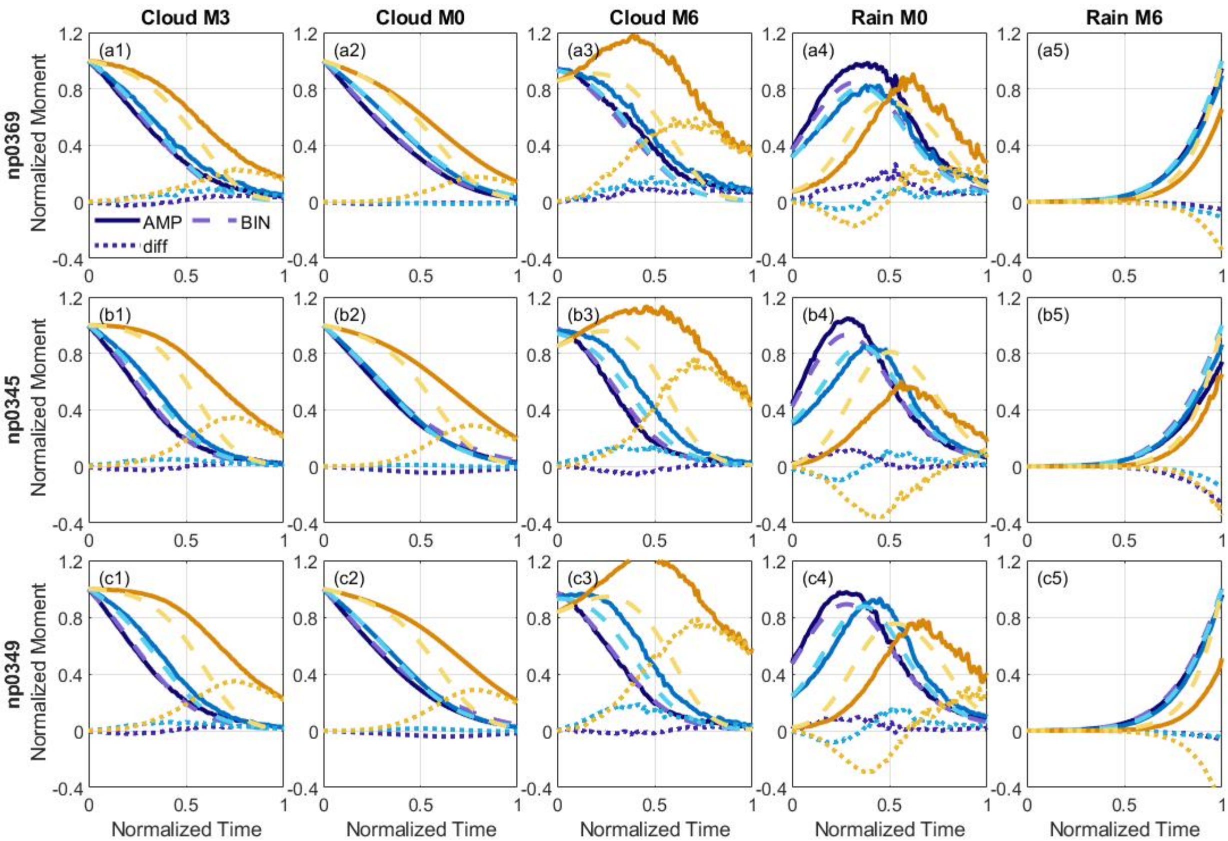


**Figure 7.** As in Figure 1 except for various configurations of AMP-NP: (a) np0369, (b) np0345, and (c) np0349.

Note that the normalized evolutions of the moments in AMP-NP, particularly of cloud M6 and rain M0, are rather noisy. This is a consequence of using size distributions obtained interpolated the lookup tables (see section 2d). While the predicted full DSD moments evolve smoothly, there is no guarantee in AMP-NP that partial moments will evolve smoothly when artificially split between cloud and rain categories. In principle this guarantee is also absent in AMP-F, but the use of a prescribed DSD functional form limits noise when the DSD is diagnostically partitioned into cloud and rain. Regardless, the overall similar performance of AMP-NP (particularly np0369) and 4M AMP-F further points to the conclusion that the major reason traditional 2M bulk schemes struggle with collision-coalescence is their use of separate cloud and rain categories, rather than their assumption of analytic functional forms for the DSD.

### 3.5 Error Dependencies

We next look to see how AMP errors in the conversion from cloud to rain depend on the initial conditions. Figure 8 shows the dependence of the normalized cloud M3 error at the time that half of the cloud mass has been converted to rain in BIN on the initial shape parameter and initial mean diameter for four configurations of AMP. As such, the maximum possible normalized error is 0.5 and indicates that BIN has converted half of the cloud mass to rain while AMP has converted no cloud mass to rain. Errors tend to be highest for high shape parameters (narrow distributions) and small initial mean diameters for AMP-CR and AMP-NP, which seems consistent with the hypothesis of Seifert and Rasp (2020) that autoconversion is ill-posed for small, narrow cloud droplet size distributions. However, aside from a handful of simulations with the highest errors for initial mean diameters around 15 $\mu$m, errors in AMP-F are high-

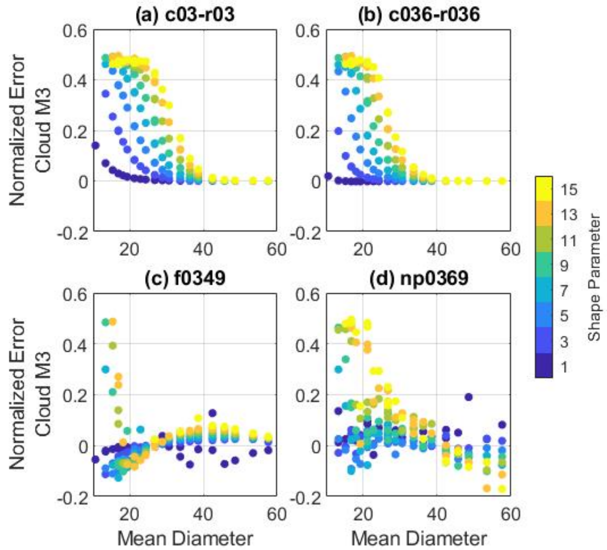


**Figure 8.** Cloud mass normalized errors for various AMP configurations as a function of initial mean diameter and shape parameter (as indicated by the point color).

est for middling values of initial mean diameter and surprisingly are typically higher than for AMP-CR for mean diameters greater than about 30 $\mu$m. Perhaps most notably, AMP configurations with separate cloud and rain categories (Fig. 8a-b) never produce negative errors, that is, AMP-CR always produces rain more slowly than BIN.

Conversely, both AMP configurations using a single category of liquid (Fig. 8c-d), AMP-F and AMP-NP, may produce rain too quickly or too slowly compared to BIN. This analysis again suggests that the traditional separate category approach is limited due to the inability to simulate an initial rain mode that may be much smaller in mean diameter than the typical threshold diameter to distinguish cloud and rain (Fig. 6). Note that AMP-NP does struggle in a similar way to AMP-CR with the smallest and narrowest initial distributions, but the problem is not as severe in AMP-NP. Overall, these results suggest that the use of separate cloud and rain modes does not allow enough flexibility in the DSD shape to capture warm rain production.

## 4 Choice of Predicted Moments

As mentioned in Section 3d, there are two possible reasons for some predictor sets to perform better than others. First, it is possible that some full moments correlate better with the process rates than others and so are more useful for accurately predicting the size distribution evolution. Second, some moment sets may contain more independent information and so better constrain the size distribution. We have explored both of these possibilities.

### 4.1 Process Rate Correlation with Moments

To investigate the possibility of moment correlation with process rates, we calculated the time tendency of each full moment for each DSD in the DSD library (which is used to construct the AMP-NP look up tables, see section 2d) with a mass mixing ratio greater than 1 g kg$^{-1}$ (which is the minimum mixing ratio tested in the simulations above). Multiple linear regression was used to predict the logarithm of these tendencies as a function of the 0th moment, the normalized 3rd moment, and all combinations of two additional normalized moments in the range 1-9, excluding 3 of course. The additional moments

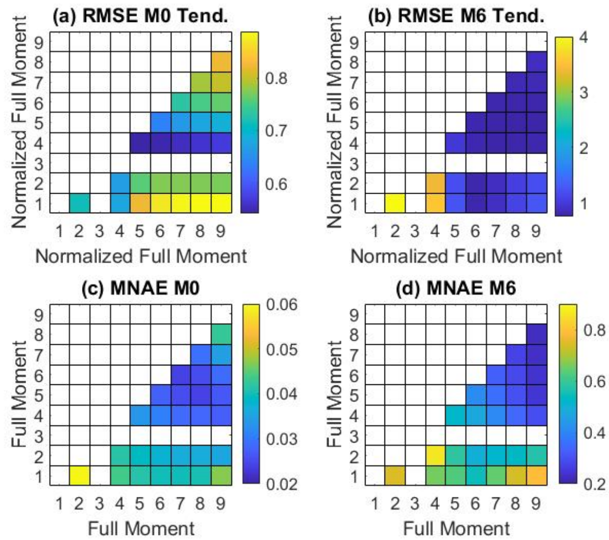


**Figure 9.** RMSE of the regression for moment tendencies for (a) M0 and (b) M6. Mean normalized absolute error for 4M AMP-F simulations for (c) M0 and (d) M6 when half of the cloud mass has been converted to rain in BIN.

are doubly- and triply-normalized, respectively, following Morrison et al. (2019). The root mean square error (RMSE) of the regression was calculated for each combination and the results are shown for the tendencies of the 0th and 6th full moments in Figure 9a-b. Additionally, 4M AMP-F was run for all moment combinations; the mean normalized absolute errors (MNAE) of full moments M0 and M6 (not tendencies) were calculated when half of the cloud mass has been converted to rain in BIN (as in Figure 8), and are shown in Figure 9c-d for comparison.

Figure 9a-b shows that the inclusion of the 4th moment results in the lowest RMSE values for the M0 tendencies; combinations including the 1st moment results in the highest RMSE values. For M6 tendencies, any moment combination that includes the 5th moment or higher substantially reduces the RMSE. But perhaps most noteworthy is that the patterns seen for the tendencies in Fig. 9a-b are not clearly reflected in the AMP-F MNAE values for M0 and M6 seen in Fig. 9c-d. For example, the best moment combination for AMP-F, at least by the metric of MNAE, is 0358 for M0 and 0389 for M6, but neither combination was expected to be best based on the tendency errors in Fig. 9a-b. Conversely, AMP-F configurations that we would have expected to be poor based on RMSE values, such as 0135 for the M0 tendency, are instead mediocre according to MNAE. So, while including moments that are predictive of collision-coalescence rates might be helpful, it does not seem to fully explain the pattern of errors in the moment values seen in Fig. 9c-d.

### 4.2 Information Content

We next investigate which combination of moments provides the most information content for the double-mode gamma DSDs in AMP-F. Because the DSDs in BIN are not double-mode gamma, error in the moment tendencies is unavoidable leading to error in the moments themselves when AMP steps forward in time. We want to determine the optimal combination of moments that minimizes the propagation of this moment error forward to the derived double-gamma DSDs. To quantify this, we will use a standard linear approximation to calculate the propagation of uncertainty in the prognostic moments to the derived double-gamma DSDs.

First, consider a pair of vectors $\vec{y}$ and $\vec{x}$ related by:

$$\vec{y} = f(\vec{x}) \tag{5}$$

If $J_f(\vec{x})$ is the Jacobian of $f$ evaluated at $\vec{x}$, then this relationship can be linearized around $(x = \vec{x}_0)$ to get:

$$\vec{y} \approx f(\vec{x}_0) + J_f(\vec{x}_0)(\vec{x} - \vec{x}_0) \tag{6}$$

If we define $\vec{y}_0 \equiv f(\vec{x}_0)$ and $J_f(\vec{x}_0)$ is invertible:

$$\vec{x} - \vec{x}_0 \approx [J_f(\vec{x}_0)]^{-1} (\vec{y} - \vec{y}_0) \tag{7}$$

To apply this linearization to propagation of uncertainty, assume that $\vec{x}$ is drawn from a distribution with expected value $\vec{\mu}_x$ and covariance matrix $\Sigma_x$, and that the corresponding distribution for $\vec{y}$ has covariance $\Sigma_y$. Then:

$$\Sigma_y = J_f(\vec{\mu}_x)\Sigma_x [J_f(\vec{\mu}_x)]^T \qquad \Sigma_x = [J_f(\vec{\mu}_x)]^{-1} \Sigma_y [J_f(\vec{\mu}_x)]^{-T} \tag{8}$$

This means that given a set of parameters for a double-mode gamma distribution, we can translate between uncertainty of those parameters and uncertainty of any (differentiable) property that can be calculated from those parameters. Furthermore, if a set of prognostic moments is enough to uniquely specify a double-mode gamma distribution, then we can translate the uncertainty of those moments into the uncertainty of the gamma distribution parameters. To do this, we use the formula for the $n$-th moment of a gamma distribution:

$$\mathrm{M}_n = \mathrm{M0}\frac{\Gamma(\nu + n)}{\Gamma(\nu)\lambda^n} \tag{9}$$

In order to nondimensionalize the moment values, we will work with their natural logarithms $L_n = \log(\mathrm{M}_n)$, and define the parameter $\phi = \log(\lambda)$. Then:

$$L_n = L_0 - n\phi + \log(\Gamma(\nu + n)) - \log(\Gamma(\nu)) \tag{10}$$

Taking the derivative of $L_n$ with respect to $L_0$ or $\phi$ is trivial here for a single mode gamma distribution (and possible for $\nu$), but for a double-mode distribution it becomes more complex. If the parameters for mode 1 are $(L_{0,1}, \phi_1, \nu_1)$, and similarly for mode 2 are $(L_{0,2}, \phi_2, \nu_2)$, then the relevant derivatives are

$$\begin{aligned}
\frac{\partial L_n}{\partial L_{0,1}} &= \frac{1}{1 + R_n} & \frac{\partial L_n}{\partial L_{0,2}} &= \frac{R_n}{1 + R_n} \\
\frac{\partial L_n}{\partial \phi_1} &= -\frac{n}{1 + R_n} & \frac{\partial L_n}{\partial \phi_2} &= -\frac{nR_n}{1 + R_n} \\
\frac{\partial L_n}{\partial \nu_1} &= \frac{1}{1 + R_n}\sum_{i=0}^{n-1}\frac{1}{\nu_1 + i} & \frac{\partial L_n}{\partial \nu_2} &= \frac{R_n}{1 + R_n}\sum_{i=0}^{n-1}\frac{1}{\nu_2 + i}
\end{aligned} \tag{11}$$

where $R_n$ is the ratio of the amount of $n$-th moment in the second mode ($\mathrm{M}_{n,2}$) to the amount in the first mode ($\mathrm{M}_{n,1}$)

$$R_n \equiv \frac{\mathrm{M}_{n,2}}{\mathrm{M}_{n,1}} \tag{12}$$

$$= R_3 r_\mu^{n-3} \frac{\Gamma(\nu_2+n)\Gamma(\nu_1+3)\nu_1^{n-3}}{\Gamma(\nu_1+n)\Gamma(\nu_2+3)\nu_2^{n-3}} \tag{13}$$

and $r_\mu$ is the ratio of the two modes' mean diameters

$$r_\mu = \frac{\nu_2\lambda_1}{\nu_1\lambda_2}. \tag{14}$$

To summarize, for a given set of prognostic moments (six if $\nu$ is allowed to vary, or four for fixed $\nu$), we can use a linear approximation to calculate how a small amount of uncertainty in those prognostic moments affects parameters of the double-mode gamma DSD. There are four non-dimensional parameters of the distribution that affect this calculation: the ratio of the two modes' masses $R_3$, the ratio of the two modes' mean diameters $r_\mu$, and the two shape parameters $\nu_1$ and $\nu_2$.

To examine how the optimum choice of predicted moments depends on these parameters, we consider the optimal set of moments with fixed $\nu_1$ and $\nu_2$, i.e. with four prognostic moments. As in the analysis in section 4.1, we require M0 and M3 to be included. We then find which other pair of moments over the range M1 to M9 (excluding M3) can be added to minimize uncertainty in $\log(R_3)$, which quantifies the uncertainty in the ratio of mass between the left and right modes. We also assume that the covariance matrix for the log-moments ($L_n$) is the identity, i.e. the log-moments are uncorrelated and all have the same variance. In other words, the magnitude of relative uncertainty is identical and uncorrelated between the moments. Parameter values considered for $R_3$ range from $10^{-2}$ to $10^2$, and for $r_\mu$ from 1 to 100. We tested all choices of $\nu \in \{0, 3, 10\}$ for each mode, but found that results were not strongly affected by the $\nu$ values. We therefore only show results where $\nu_1 = 10$ and $\nu_2 = 3$, values that are typical early in the 6M AMP-F simulations, as seen in Figure 6.

Results are shown in Figure 10. We notice first that M9 is always one of the optimal moment choices (there are rare exceptions to this for other values of $\nu$, which is not shown). With M0, M3, and M9 as prognostic moments, the remaining optimal moment depends on the details of the droplet size distribution. If $R_3 \leq 1$, i.e. if most of the mass is in the smaller mode, then the optimal fourth moment will be M4 or higher. Otherwise the optimal moment will be M1 or M2. We can also see that moments closer to M3 are preferred when the two modes are well separated ($r_\mu \gg 1$).

While there is not a one-to-one correspondence of the optimal moment combinations in Fig. 10 to the smallest M0 and M6 MNAE in Fig. 9c-d, there are similar trends. For instance, including M9 as a predicted moment leads to the smallest M6 MNAE when the other predicted moment lies between M4 and M8 (Fig. 9d), consistent with the information content analysis here showing M9 is (nearly) always optimal; M9 only slightly increases error compared to M7 and M8 when the other moment is between M4 and M8 for the M0 MNAE (Fig. 9c). The optimal moment pairs here are more consistent with the MNAE results (Fig. 9c-d) than the RMSE tendency (Fig. 9a-b). However, they cannot explain all trends in MNAE. We highlight one interesting difference between the optimal moment pairs in Fig. 10 and the MNAE analysis. MNAE is generally larger (particular for M6) when one of the predicted moments is M1 or M2, compared to when both moments are between M4 and M9. In contrast, M1 or M2 together with M9 are optimal according to the analysis here when $R_3 > 3$, that is, when the right (large) mode dominates the DSD. A plausible explanation is that, when integrated in time, errors need to be minimized early in the simulations during the rain initiation stage when $R_3 < 1$ (meaning the left mode dominates), in order to minimize overall error. As shown in Fig. 10, this would imply an optimal moment combina-

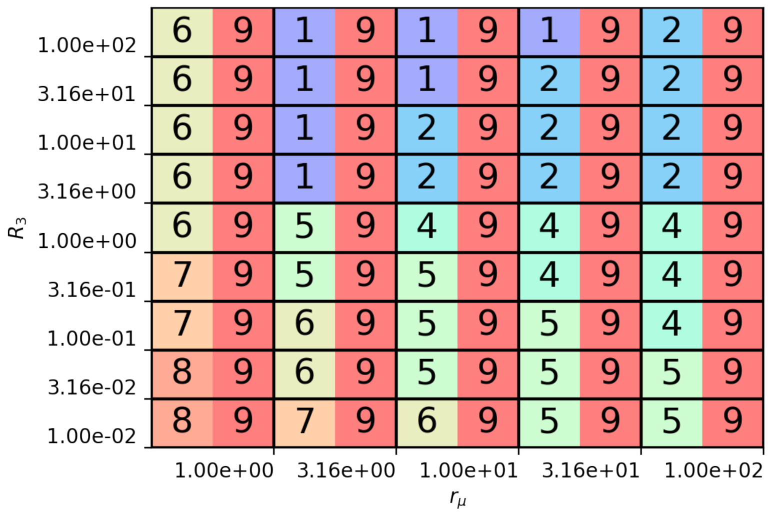


**Figure 10.** Orders of optimal pairs of predicted moments (in conjunction with M0 and M3) leading to the smallest error in $\log(R_3)$, where $R_3$ is ratio of mass between the two gamma distribution modes. Results for the optimal moment pairs are shown across the two-dimensional space of $R_3$ (y-axis) and the ratio between the modes' mean diameters $r_\mu$ (x-axis), with $\nu_1 = 10$ and $\nu_2 = 3$. Each moment is color coded for clarity, with lower order moments in cool colors (blue) and higher order moments in warm colors (orange to red). As highlighted in red, M9 is one of the optimal moments in the pair for all values of $R_3$ and $r_\mu$.

tion generally between M4 and M7, together with M9, which is consistent with the MNAE results. Additional analysis described below supports this idea.

Figure 10 provides information on the optimal combination of predicted moments for partitioning mass between the modes, but not on how much better the optimal combination is compared to other combinations. Thus, we include Figure 11 which shows the ratio of the uncertainty in $\log(R_3)$ to the uncertainty in the input moments for various combinations of predicted moments (which we will call the "uncertainty multiplier"). For instance, if the uncertainty multiplier is 20 (the maximum shown) and all moments are subjected to an uncorrelated error of 0.5 dB, then $R_3$ will be affected by a 10 dB error, i.e. only the rough order of magnitude can be correctly estimated. M0 and M3 are again included as two of the four moments, while all other combinations of moment pairs between M1 and M9 (excluding M3) are analyzed.

At the initial time the droplet size distribution only has one small mode, and the second mode gradually forms from its right tail. Thus, early in the simulations both $R_3$ and $r_\mu$ will be small. Over time, the second mode both separates from the first mode (increasing $r_\mu$) and grows in amplitude (increasing $R_3$), which can be seen in both the BIN and AMP runs in Figure 4. This evolution is followed by the sequence of plots in Fig. 11a-d. In particular, Fig. 11a shows that when the second mode is still relatively undeveloped (i.e., small $R_3$ and $r_\mu$), using M1 or M2 as predicted moments (the bottom two rows), regardless of the other moment, is unable to "resolve" the distinction between the first and second modes at all. Using M4 (particularly with M5 as the other moment) leads to a similar problem, though to a lesser extent. On the other hand, Fig. 11b shows that if the second mode is more separated from the first but the first mode still dominates (i.e., small $R_3$ but large $r_\mu$), M4 produces comparable results to M5-M7, regardless of the other moment, while M1 and M2 still give large uncertainty. This may explain why AMP-F f0349 can do better with smaller initial diameters than middling initial diameters (Fig. 8c).

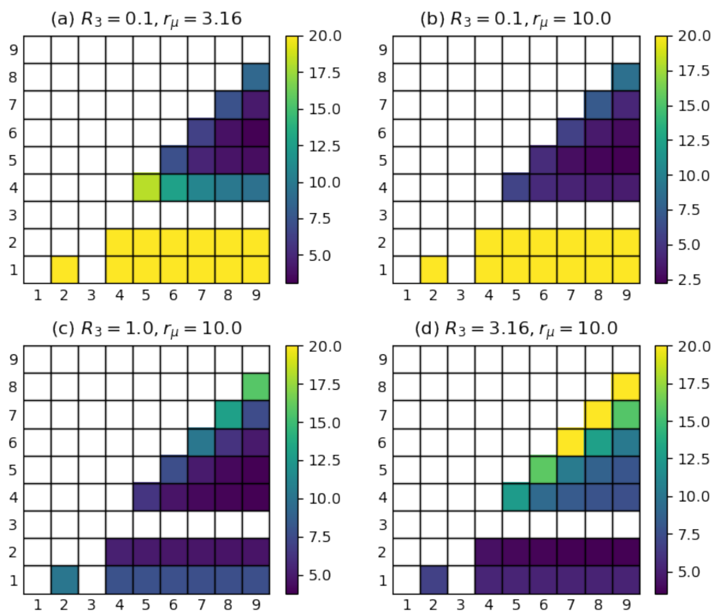


**Figure 11.** Uncertainty multiplier (ratio of the uncertainty in $\log(R_3)$ to the uncertainty in the input moments, where uncertainty is defined as the square root of variance) for various moment pairs (in conjunction with M0 and M3), for different values of $r_\mu$ and $R_3$ as labeled above the four plots. The x- and y-axes are the orders of the moment pairs. For all plots, $\nu_1 = 10$ and $\nu_2 = 3$. Note that the color range only extends to 20, but values can be much larger, e.g. $> 100$ for the (M1,M2) pair in plot (a).

Figure 11c-d shows how the growth of the second mode (meaning larger $R_3$) changes the optimal choice of moments, as combinations that include M1 or M2 become more effective while combinations using higher moments lead to greater DSD uncertainty. This may explain why moment choices that do well early in the simulations, such as (M0,M3,M6,M9), see some loss of accuracy once the majority of cloud has been converted to rain, but why other combinations including M1 or M2 do less well overall as quantified by MNAE (Fig. 9c-d). In other words, even if including M1 or M2 as a predicted moment is more effective at later times, it may not be able to recover from large errors earlier in the simulation.

That said, this analysis still fails to explain why AMP-F works well when using moment combinations without any moments higher than M6. For example, AMP-F f0345 gives comparable or perhaps even slightly better results relative to the benchmark compared to f0349 and f0369 (Fig. 3). We have also examined DSD uncertainties in other quantities apart from $\log(R_3)$, such as the mean particle sizes of the two modes and M0 and M6 partitioned between the modes. Uncertainty using (M0,M3,M6,M9) or (M0,M3,M4,M9) is virtually always far lower than using (M0,M3,M4,M5) for all quantities, even though AMP-F f0345 produces overall similar or slightly better results compared to AMP-F f0369 and f0349. It is possible that higher moments do not work as well in practice due to numerical considerations (e.g. the limited range and resolution used to represent the DSD in the bin model). It is also possible that the errors that result from assuming a double gamma distribution are more pronounced when using moments greater than $M_6$, due to the fact that larger moments depend heavily on the tails of the distribution.

We emphasize that the uncertainty analysis in this section applies strictly to the two-mode gamma DSDs in AMP-F. While AMP-NP np0369 performs similarly to AMP-F, other moment combinations for

AMP-NP produce much poorer results (Fig. 7). Thus, uncertainty characteristics as a function of the choice of prognostic moments are much different in AMP-NP than the two-mode gamma DSDs in AMP-F. As we already noted, the non-parametric reconstruction of DSDs in AMP-NP works best when the orders of the predicted moments are spread apart. In this case, as the difference in the moment orders increases their correlation decreases, meaning the moments are better at providing independent information about the DSD (Morrison et al., 2019). It is clear this situation does not simply translate to two-mode gamma DSDs.

## 5 Conclusions

In this study we have used AMP, a flexible bulk scheme with bin scheme process parameterizations, to investigate why warm rain production is so difficult generally to represent in bulk schemes. We configured AMP to run in three ways: with traditional, separate cloud and rain categories using either two or three predicted moments for each category, with a single liquid category described with a double-mode DSD using four or six predicted moments, and with a single liquid category with a nonparametric DSD using four predicted moments. AMP was run as a box model in all configurations with collision-coalescence as the only microphysical process and initialized with a variety of unimodal DSDs. Output was compared to reference simulations using the bin scheme upon which AMP is built.

Based on our analysis, we find that the use of separate cloud and rain modes is the primary reason why bulk schemes struggle with warm rain formation. The primary reason is not the choice of predicted moments nor the use of assumed gamma distributions. When a continuous double-mode distribution is used, we find that the evolutions of initially small and narrow cloud droplet distributions, for which autoconversion has historically been challenging, become much more predictable. We find that the second mode, corresponding to rain, has an initially very small diameter, much smaller than is typically considered to be rain. With separate liquid categories, these nascent "rain" drops remain in the cloud category where they cannot be properly represented with an assumed unimodal cloud DSD.

Traditional bulk schemes may possibly be improved by transferring all droplets involved in collisions to the rain category, even if the resultant drop does not meet some size threshold such as a 40 $\mu$m radius. Alternatively, we would encourage development of single liquid category bulk microphysics schemes. This study suggests that a single liquid category could lead to improvements in our ability to simulate warm rain processes. We have shown here that a four moment single category scheme should likely include prediction of the 0th, 3rd, and 9th full moments of the distribution. The optimal choice of a fourth predicted moment is currently unclear since the optimal combination, from an information content perspective, depends on the relative importance of the modes, but it is likely the 4th-6th. Regardless, exploration of the design and advantages or disadvantages of single category schemes is an avenue for future research.

## 6 Open Research

All AMP simulation data and scripts used to analyze the data are publicly available and are archived at `https://datadryad.org/` (Igel, 2022). If the archive is not yet public, the data can also be accessed at `http://farm.cse.ucdavis.edu/~aigel/AMP` during review.

**Acknowledgments**
A. L. Igel was supported by National Science Foundation awards 1940035 and 2025103. H. Morrison, S. P. Santos, and M. van Lier-Walqui were supported by the DOE Earth System Model Development grant C0021048. We would like to acknowledge high-performance computing support from Cheyenne (doi:10.5065/D6RX99 provided by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation. The National Center for Atmospheric Research is sponsored by the National Science Foundation.

## References

Austin, P., Wang, Y., Kujala, V., & Pincus, R. (1995, 7 1). Precipitation in stratocumulus clouds: Observational and modeling results. *Journal of the Atmospheric Sciences*, *52*(13), 2329–2352. (publisher: American Meteorological Society section: Journal of the Atmospheric Sciences) doi: 10.1175/1520-0469(1995)052⟨2329:PISCOA⟩2.0.CO;2

Berry, E. X., & Reinhardt, R. L. (1974, 10 1). An analysis of cloud drop growth by collection part ii. single initial distributions. *Journal of the Atmospheric Sciences*, *31*(7), 1825–1831. (publisher: American Meteorological Society section: Journal of the Atmospheric Sciences) doi: 10.1175/1520-0469(1974)031⟨1825:AAOCDG⟩2.0.CO;2

Chiu, J. C., Yang, C. K., van Leeuwen, P. J., Feingold, G., Wood, R., Blanchard, Y., … Wang, J. (2021). Observational constraints on warm cloud microphysical processes using machine learning and optimization techniques. *Geophysical Research Letters*, *48*(2). doi: 10.1029/2020gl091236

Clark, T. L. (1974, 1 1). A study in cloud phase parameterization using the gamma distribution. *Journal of the Atmospheric Sciences*, *31*(1), 142–155. (publisher: American Meteorological Society section: Journal of the Atmospheric Sciences) doi: 10.1175/1520-0469(1974)031⟨0142:ASICPP⟩2.0.CO;2

Clark, T. L. (1976). Use of Log-Normal Distributions for Numerical Calculations of Condensation and Collection. *Journal of the Atmospheric Sciences*, *33*(5), 810–821. doi: 10.1175/1520-0469(1976)033⟨0810:uolndf⟩2.0.co;2

Clark, T. L., & Hall, W. D. (1983). A Cloud Physical Parameterization Method Using Movable Basis Functions: Stochastic Coalescence Parcel Calculations. *Journal of the Atmospheric Sciences*, *40*(7), 1709–1728. doi: 10.1175/1520-0469(1983)040⟨1709:acppmu⟩2.0.co;2

Cotton, W. R., Pielke Sr, R. A., Walko, R. L., Liston, G. E., Tremback, C. J., Jiang, H., … McFadden, J. P. (2003). RAMS 2001: Current status and future directions. *Meteorology and Atmospheric Physics*, *82*(1-4), 5–29. doi: 10.1007/s00703-001-0584-9

Feingold, G., Walko, R. L., Stevens, B., & Cotton, W. R. (1998). Simulations of marine stratocumulus using a new microphysical parameterization scheme. *Atmospheric Research*, *47-48*, 505–528. doi: 10.1016/S0169-8095(98)00058-1

Ferek, R. J., Garrett, T., Hobbs, P. V., Strader, S., Johnson, D., Taylor, J. P., … Babb, D. (2000, 8 1). Drizzle suppression in ship tracks. *Journal of the Atmospheric Sciences*, *57*(16), 2707–2728. (publisher: American Meteorological Society section: Journal of the Atmospheric Sciences) doi: 10.1175/1520-0469(2000)057⟨2707:DSIST⟩2.0.CO;2

Igel, A. L. (2019). Using an arbitrary moment predictor to investigate the optimal choice of prognostic moments in bulk cloud microphysics schemes. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4559–4575. doi: 10.1029/2019ms001733

Igel, A. L. (2022). *AMP collision coalescence simulations* [dataset]. Dryad. Retrieved from `https://datadryad.org/stash/dataset/doi:10.25338/B8SK8H` doi: 10.25338/B8SK8H

Jing, X., Suzuki, K., Guo, H., Goto, D., Ogura, T., Koshiro, T., & Mülmenstädt, J. (2017). A multimodel study on warm precipitation biases in global models compared to satellite observations. *Journal of Geophysical Research: Atmospheres*, *122*(21), 11,806–11,824. (eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2017JD027310) doi: 10.1002/2017JD027310

Kay, J. E., L'Ecuyer, T., Pendergrass, A., Chepfer, H., Guzman, R., & Yettella, V. (2018). Scale-aware and definition-aware evaluation of modeled near-surface precipitation frequency using cloudsat observations. *Journal of Geophysical Research: Atmospheres*, *123*(8), 4294–4309. (eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2017JD028213) doi: 10.1002/2017JD028213

Khain, A., Pokrovsky, A., Pinsky, M., Seifert, A., & Phillips, V. (2004). Simulation of Effects of Atmospheric Aerosols on Deep Turbulent Convective Clouds Using a Spectral Microphysics Mixed-Phase Cumulus Cloud Model. Part I: Model Description and Possible Applications. *Journal of the Atmospheric Sciences*, *61*(24), 2963–2982. Retrieved from `http://journals.ametsoc.org/doi/abs/10.1175/JAS-3350.1` doi: 10.1175/JAS-3350.1

Khairoutdinov, M., & Kogan, Y. (2000). A new cloud physics parameterization in a large-eddy simulation model of marine stratocumulus. *Monthly Weather Review*, *128*(1), 229–243. 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

Kogan, Y. L., & Belochitski, A. (2012). Parameterization of Cloud Microphysics Based on Full Integral Moments. *Journal of the Atmospheric Sciences*, *69*(7), 2229–2242. doi: 10.1175/jas-d-11-0268.1

Lee, H., & Baik, J.-J. (2017). A physically based autoconversion parameterization. *Journal of the Atmospheric Sciences*, *74*(5), 1599–1616. doi: 10.1175/jas-d-16-0207.1

Liu, Y., & Daum, P. H. (2004). 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

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. doi: 10.1175/JAS3446.1

Morrison, H., Kumjian, M. R., Martinkus, C. P., Prat, O. P., & van Lier-Walqui, M. (2019). A General N-Moment Normalization Method for Deriving Raindrop Size Distribution Scaling Relationships. *Journal of Applied Meteorology and Climatology*, *58*(2), 247–267. doi: 10.1175/jamc-d-18-0060.1

Morrison, H., Lier-Walqui, M. v., Kumjian, M. R., & Prat, O. P. (2020). A Bayesian approach for statistical-physical bulk parameterization of rain microphysics, Part I: Scheme description. *Journal of the Atmospheric Sciences*, *77*, 1019-1041. Retrieved from `https://journals.ametsoc.org/doi/abs/10.1175/JAS-D-19-0070.1` doi: 10.1175/jas-d-19-0070.1

Mülmenstädt, J., Salzmann, M., Kay, J. E., Zelinka, M. D., Ma, P.-L., Nam, C., . . . Quaas, J. (2021, 6). An underestimated negative cloud feedback from cloud lifetime changes. *Nature Climate Change*, *11*(6), 508–513. doi: 10.1038/s41558-021-01038-1

Rodríguez Genó, C. F., & Alfonso, L. (2022). Parameterization of the collision–coalescence process using series of basis functions: Colnetv1.0.0 model development using a machine learning approach. *Geoscientific Model Development*, *15*(2), 493–507. Retrieved from `https://gmd.copernicus.org/articles/15/493/2022/` doi: 10.5194/gmd-15-493-2022

Seifert, A., & Beheng, K. D. (2001). A double-moment parameterization for simulating autoconversion, accretion and selfcollection. *Atmospheric Research*, *59-60*, 265–281. doi: 10.1016/S0169-8095(01)00126-0

Seifert, A., & Rasp, S. (2020). Potential and limitations of machine learning for modeling warm-rain cloud microphysical processes. *Journal of Advances in Modeling Earth Systems*, *12*(12), e2020MS002301. (eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2020MS002301) doi: 10.1029/2020MS002301

Sinclair, K., van Diedenhoven, B., Cairns, B., Alexandrov, M., Dzambo, A. M., & L'Ecuyer, T. (2021). Inference of precipitation in warm stratiform clouds using remotely sensed observations of the cloud top droplet size distribution. *Geophysical Research Letters*, *48*(10), e2021GL092547. (eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2021GL092547) doi: 10.1029/2021GL092547

Szyrmer, W., Laroche, S., & Zawadzki, I. (2005). A Microphysical Bulk Formulation Based on Scaling Normalization of the Particle Size Distribution. Part I: Description. *Journal of the Atmospheric Sciences*, *62*, 4206–4221. doi: 10.1175/jas3620.1

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. doi: 10.1016/0169-8095(94)00087-T

Zeng, X., & Li, X. (2020, 3 1). A two-moment bulk parameterization of the drop collection growth in warm clouds. *Journal of the Atmospheric Sciences*, *77*(3), 797–811. (publisher: American Meteorological Society section: Journal of the Atmospheric Sciences) doi: 10.1175/JAS-D-19-0015.1