RMetS
Royal Meteorological Society

Analytic upscaling of a local microphysics scheme. Part I: Derivation

Vincent E. Larson* and Brian M. Griffin
Department of Mathematical Sciences, University of Wisconsin–Milwaukee, WI, USA

*Correspondence to: V. E. Larson, Department of Mathematical Sciences, University of Wisconsin–Milwaukee, PO Box 413, Milwaukee, WI 53201, USA. E-mail: vlarson@uwm.edu

In large-scale simulations of the atmosphere, microphysical processes occur on smaller scales than the grid-box size. If the processes are nonlinear, and the variability within a grid box is large, then ignoring the subgrid variability leads to inaccuracy. To avoid this inaccuracy, local microphysical formulas, valid at a point in space, may be upscaled to a large grid box.

This may be done by integrating the local microphysical formula over the probability density function (PDF) that represents spatial subgrid variability. This paper will upscale local formulas proposed by Khairoutdinov and Kogan, which constitute a complete microphysics scheme for drizzle in marine stratocumulus. It is tractable to upscale this scheme because all formulas in it are power laws. The marginal PDFs will be assumed to have either a normal mixture or log-normal functional form.

In Part II of this pair of papers, the upscaled formulas are implemented interactively in a single-column model and tested for a drizzling marine stratocumulus case.

Key Words: microphysics; parametrization; probability density function; upscaling

Received 13 May 2010; Revised 29 February 2012; Accepted 30 March 2012; Published online in Wiley Online Library 7 June 2012

Citation: Larson VE, Griffin BM. 2013. Analytic upscaling of a local microphysics scheme. Part I: Derivation. *Q. J. R. Meteorol. Soc.* **139**: 46–57. DOI:10.1002/qj.1967

1. Introduction

Microphysical processes are local. Consider, for instance, autoconversion, the process whereby cloud droplets grow to drizzle-drop size. The autoconversion rate at a point in space depends on the population of cloud droplets in the immediate vicinity, not on cloud droplets at distant locations.

However, large-scale atmospheric numerical models of atmospheric flows that compute such microphysical processes typically discretize the domain into large grid boxes and compute fields, e.g. moisture and temperature, only at the grid-box scale or larger. Such models require not local microphysical rates but rather average rates over a grid box that can be used to increment the grid-box-scale moisture and temperature fields.

When we desire a grid-box-average rate, substituting a local microphysical rate is often inaccurate. This is especially true when the microphysical processes are both nonlinear and small in scale.

The nonlinearity implies that, for a process rate represented by a nonlinear function f that depends on variable x, usually

$$\langle f(x) \rangle \neq f(\langle x \rangle), \tag{1}$$

where the angle brackets denote a grid-box average (Pincus and Klein, 2000). In other words, the grid-box average (the left-hand side) does not equal the result of feeding the average of x into the local formula f (the right-hand side), except by coincidence. For example, the grid-box-average cloud water is insufficient information to predict grid-box-average autoconversion rate. One reason is that

autoconversion occurs disproportionately in the part of the grid box with more cloud water.

The small-scale nature of such processes allows variability to occur on the subgrid scale. When the subgrid variability is large, the microphysical rates need to be averaged in a way that accounts for this variability. Stated differently, local microphysical formulas, even if perfectly accurate, need to be 'upscaled' to the grid-box scale.

To do so, a model may estimate, for variables like x, the probability density function (PDF) of spatial variability within a grid box, $P(x)$. Then the model may integrate over the PDF to produce a grid-box average of the microphysical process rate:

$$\langle f(x) \rangle = \int f(x) P(x) \mathrm{d}x. \tag{2}$$

In many cases, the PDF provides sufficient information to account for subgrid variability. Given the within-grid-box PDF, a model 'knows' how much cloud water is present in the moistest part of a cloudy grid box and therefore 'knows' how to weight an autoconversion formula in order to account for the moistest part.

Predicting the PDF with sufficient accuracy is a significant task, but if the PDF can be accurately predicted, then, in many cases, the problem of parametrizing subgrid variability reduces to the problem of quadrature, namely, the integration of Eq. (2). The integral can be evaluated via Monte Carlo integration (Pincus *et al.*, 2003; Räisänen *et al.*, 2004; Larson, 2007). That is, sample points can be drawn from the PDF, $P(x)$, and fed into the microphysical formula $f(x)$. The resulting points can be averaged to yield an estimate of $\langle f(x) \rangle$. This technique is flexible and is applicable when $f(x)$ is a complex numerical subroutine. However, the technique also introduces statistical noise into the averages due to the necessarily limited sample size (Räisänen and Barker, 2004; Larson *et al.*, 2005; Räisänen *et al.*, 2005; Pincus *et al.*, 2006).

Rather than Monte Carlo integration, this paper instead integrates Eq. (2) analytically. This yields exact integrals, and in particular avoids the noise associated with Monte Carlo integration. Analytic integration is possible only because we use microphysical formulas $f(x)$ and PDFs $P(x)$ that are simple. Specifically, the microphysical formulas are those of Khairoutdinov and Kogan (2000) (hereafter denoted KK) for drizzling marine stratocumulus clouds. KK parametrize all microphysical sources and sinks in terms of power laws, which greatly simplifies the mathematics. Furthermore, our chosen PDF has a simple functional form. It is a mixture of two multivariate normals–in other words, a sum of two joint Gaussians. (For the precipitation variables, the normal mixture form is transformed from a log-normal form.) Each univariate marginal PDF is a normal mixture. (A univariate marginal PDF is the single-variable PDF that remains when all other variates of the full multivariate PDF are integrated over the entire domain.) This combination of simple functional forms for $f(x)$ and $P(x)$ yields analytically tractable integrals.

This analytic approach was outlined and tested by Larson and Griffin (2006). Further results using this method were presented in Wyant *et al.* (2007). The present paper presents the mathematical details that underlie these results.

A related analytic approach has been used by Cheng and Xu (2009) for a single-moment microphysics scheme and a normal mixture PDF for rain. The approach of Larson and Griffin (2006), Wyant *et al.* (2007), and the present paper differ from the approach of Cheng and Xu (2009) in a couple ways. First, we use a single log-normal marginal PDF for precipitation variables. The log-normal form ensures that small but non-negative variables such as rain indeed are non-negative. It also avoids the need to explicitly predict skewness of these variables. Second, we use a microphysics scheme that includes both droplet number and mixing ratio in its diagnostic formula for autoconversion.

Two other related analytic approaches were outlined by Zhang *et al.* (2002) and Morrison and Gettelman (2008). These approaches used a univariate PDF for cloud water, rather than a multivariate approach that includes rainwater such as we propose here. A multivariate approach is useful for processes that involve two variables, such as collection of cloud water by rainwater.

The remainder of this paper is organized as follows. Section 2 provides an overview of the local KK prognostic equations. Section 3 presents the functional form of the subgrid PDF. Section 4 presents upscaled versions of the KK formulas for mean volume radius and sedimentation velocity of rain. Section 5 presents the upscaled version of the KK autoconversion formula; section 6, the KK accretion formula; and section 7, the KK condensation/evaporation formula. Section 8 describes various parameters that currently need to be specified in order to use the aforementioned set of equations. Section 9 summarizes the upscaling method.

2. Mathematical framework

2.1. The Khairoutdinov and Kogan (2000) prognostic equations for drizzle

The KK formulas parametrize drizzle in marine stratocumulus clouds. The KK scheme is double moment in both rainwater and cloud water. (This paper will use the word 'rain' synonymously with 'drizzle', although the KK scheme is designed for small raindrops, that is, drizzle drops.) That is, the KK scheme prognoses both number concentrations and mixing ratios for both rain and cloud water. In our simulations, we prognose number concentration and mixing ratio of rain, but we diagnose cloud water mixing ratio, and we prescribe the number concentration of cloud droplets.

The prognostic equation that we use for rainwater mixing ratio, r_r, and raindrop concentration per mass of air, N_r, are based on Eqs (8) and (9) of Khairoutdinov and Kogan (2000). After Reynolds averaging (Stull, 1988), representing turbulent advection using eddy diffusivity, and neglecting correlations with perturbation sedimentation velocity or eddy diffusivity, we find an equation for the grid-box average of r_r:

$$\begin{aligned} \frac{\partial \langle r_r \rangle}{\partial t} = &- \langle w \rangle \frac{\partial \langle r_r \rangle}{\partial z} + \frac{\partial \langle V_{r_r} \rangle \langle r_r \rangle}{\partial z} \\ &+ \left\langle \left(\frac{\partial r_r}{\partial t} \right)_{\text{evap}} \right\rangle + \left\langle \left(\frac{\partial r_r}{\partial t} \right)_{\text{auto}} \right\rangle \\ &+ \left\langle \left(\frac{\partial r_r}{\partial t} \right)_{\text{accr}} \right\rangle + \frac{\partial}{\partial z} \langle K \rangle \frac{\partial \langle r_r \rangle}{\partial z}, \end{aligned} \tag{3}$$

and an equation for N_r:

$$\frac{\partial \langle N_r \rangle}{\partial t} = - \langle w \rangle \frac{\partial \langle N_r \rangle}{\partial z} + \frac{\partial \left\langle V_{N_r} \right\rangle \langle N_r \rangle}{\partial z} + \left\langle \left(\frac{\partial N_r}{\partial t} \right)_{\text{evap}} \right\rangle + \left\langle \left(\frac{\partial N_r}{\partial t} \right)_{\text{auto}} \right\rangle + \frac{\partial}{\partial z} \langle K \rangle \frac{\partial \langle N_r \rangle}{\partial z}. \quad (4)$$

Here w is the vertical wind component in $\mathrm{m\,s^{-1}}$, V_{r_r} and V_{N_r} are the sedimentation velocities (defined positive downward) in $\mathrm{m\,s^{-1}}$ of r_r and N_r, respectively, and K is the eddy diffusion coefficient in $\mathrm{m^2\,s^{-1}}$. The subscript 'evap' denotes condensation/evaporation; 'auto' denotes autoconversion, and 'accr' denotes accretion (or collection) of cloud droplets by raindrops. A grid-box mean is denoted by $\langle\ \rangle$, and a perturbation from that mean will be denoted hereafter by a prime ($'$).

2.2. *Combining the effects of heat and moisture on cloud water through the extended liquid water mixing ratio, s*

The effects of liquid water potential temperature, θ_l, and total water mixing ratio (vapour plus cloud water), r_t, on cloud water are incorporated through the extended liquid water mixing ratio, s (see Eq. 1 of Lewellen and Yoh, 1993). Extended liquid water mixing ratio s equals cloud-water mixing ratio, r_c, in cases where the air is saturated. However, unlike r_c, s is negative rather than zero when the air is subsaturated. Extended liquid-water mixing ratio can be defined as

$$s = \frac{r_t - r_s\left(T_l, p\right)}{1 + \Lambda\left(T_l\right) r_s\left(T_l, p\right)}, \quad (5)$$

where $r_s\left(T_l, p\right)$ is the saturation mixing ratio at the liquid-water temperature, T_l, which in turn is given by

$$T_l = T - \frac{L_v}{c_{pd}} r_c = \theta_l \left(\frac{p}{p_0} \right)^{\frac{R_d}{c_{pd}}}, \quad (6)$$

where T is absolute temperature, L_v is the latent heat of vaporization, c_{pd} is the specific heat of dry air at constant pressure, p is pressure, p_0 is the reference pressure of 1.0×10^5 Pa, and R_d is the gas constant for dry air. The function $\Lambda\left(T_l\right)$ is defined as

$$\Lambda\left(T_l\right) = \frac{R_d}{R_v} \left(\frac{L_v}{R_d T_l} \right) \left(\frac{L_v}{c_{pd} T_l} \right), \quad (7)$$

where R_v is the gas constant for water vapour.

For later convenience, we now define a dimensionless version of s, $s_* = s/s_0$, where $s_0 = 1\ \mathrm{kg\,kg^{-1}}$. We denote the overall mean value of s_* by μ_{s_*}, and the overall variance of s_* by $\sigma^2_{s_*}$. The mean and standard deviation of s_* for each individual normal component in the normal mixture are defined, respectively, as $\mu_{s_{*(i)}} = \left\langle s_{(i)} \right\rangle / s_0$ and $\sigma_{s_{*(i)}} = \left\langle s'^2_{(i)} \right\rangle^{1/2} / s_0$. A subscript i, where i equals 1 or 2, denotes a mixture component.

3. The PDF functional form

The subgrid-scale variability in all relevant fields is represented by a single, multivariate PDF. The PDF that we use is predicted by the Cloud Layers Unified By Binormals (CLUBB) single-column model (SCM) (Golaz *et al.*, 2002a, 2002b; Larson and Golaz, 2005). CLUBB assumes that the PDF is a weighted mixture, or sum, of two multivariate normal/log-normal functions (see Eq. 17 below). A mixture PDF can take on a variety of skewed and unskewed shapes because its two components can vary independently, as depicted in Figure 1 of Golaz *et al.* (2002a). In Golaz *et al.* (2002a), the PDF, $P(w, \theta_l, r_t)$, was a function only of liquid water potential temperature, θ_l, total water mixing ratio, r_t, and vertical velocity, w. We assume that θ_l, r_t, and w are each distributed as a normal mixture. Because subgrid variations in θ_l and r_t are small, perturbations in s may be approximated as a linear combination of perturbations in θ_l and r_t. Therefore, the distribution of s may also be approximated as a normal mixture.

Unlike Golaz *et al.* (2002a), this paper extends the PDF to include r_r, N_r, and cloud droplet concentration per mass of air, N_c. It can be written $P(w, \theta_l, r_t, r_r, N_r, N_c)$. The marginal distribution of r_c is diagnosed by truncating the distribution of s at saturation and adding a delta function at $r_c = 0$ in order to represent clear air. The marginal distributions of r_r, N_r, and N_c are all assumed to be single log-normal distributions. A log-normal shape is suitable for hydrometeor species because it tends to an unskewed shape when its mean is large and a positively skewed shape when its mean is small. Furthermore, a log-normal distribution is positive definite. This property is unimportant for variables such as *total* water mixing ratio (vapour plus liquid), which are non-negative but whose left tail falls off well above zero mixing ratio. However, variables such as *drizzle* mixing ratio often cluster near zero but can never be negative. For such variables, it is important that the PDF allow strong positive skewness and yet ensure that negative values never occur. Additionally, a single log-normal distribution has the virtue of simplicity. The form of the PDF for a single variable, x, that is distributed log-normally is

$$P(x) = \left[(2\pi)^{1/2} \sigma x \right]^{-1} \exp\left[\frac{-(\ln x - \mu)^2}{2\sigma^2} \right], \quad (8)$$

where μ and σ^2 are the mean and the variance, respectively, of $\ln x$, which is normally distributed.

In order to compute grid-box averages that account for subgrid variability, we need to integrate local microphysical formulas over the PDF, as in Eq. (2). To do so, it proves mathematically convenient to transform log-normal PDFs to normal ones.

We transform the log-normally distributed variables r_r, N_r, and N_c to normally distributed variables r_{rn}, N_{rn}, and N_{cn} as follows:

$$r_{rn} = \ln\left(\frac{r_r}{r_{r0}} \right), \quad (9)$$

$$N_{rn} = \ln\left(\frac{N_r}{N_{r0}} \right), \quad (10)$$

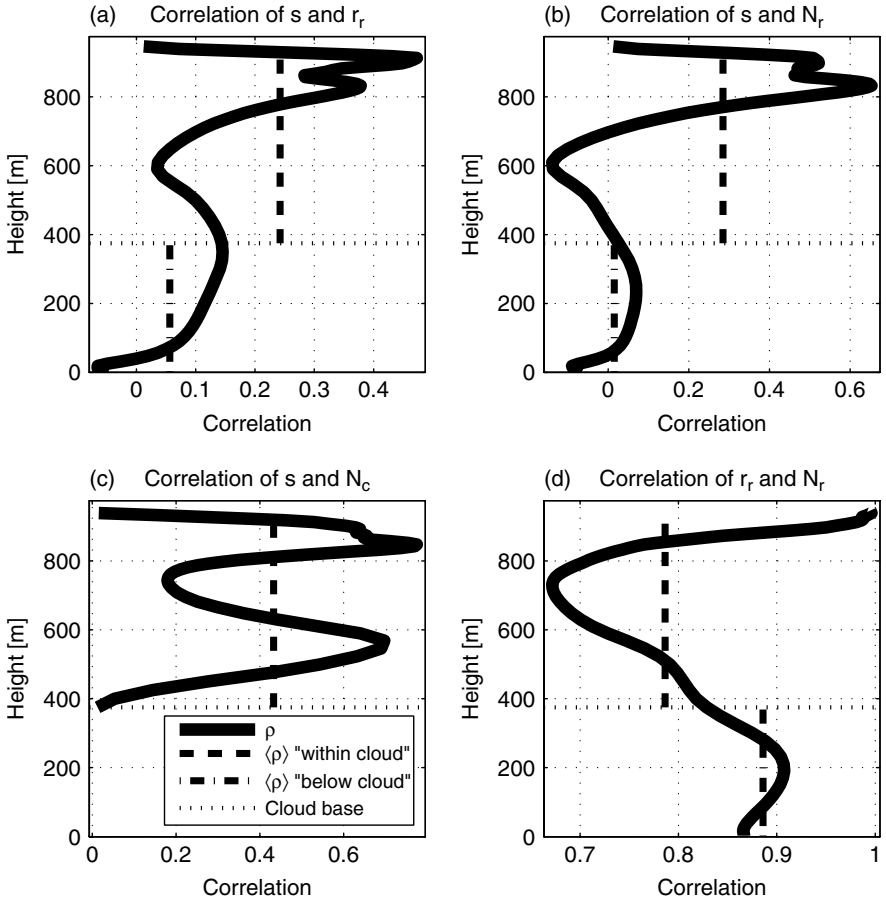

Figure 1. The time-averaged values of $\rho_{(s,r_r)}$, which is the correlation of s and r_r (a), $\rho_{(s,N_r)}$, which is the correlation of s and N_r (b), $\rho_{(s,N_c)}$, which is the correlation of s and N_c (c), and $\rho_{(r_r,N_r)}$, which is the correlation of r_r and N_r (d). The values of the correlations are taken from a SAM LES simulation of the DYCOMS-II RF02 case time-averaged over hours 4–6. For each plot, the thick, solid line is the time-average value of the correlation at each vertical level. The dashed line is the average value of the correlation at all levels with at least some cloud. The dash-dotted line is the average value of the correlation at levels entirely without cloud. The dotted line is the time-averaged altitude of cloud base, which does not change much over the time-averaging period.

and

$$N_{cn} = \ln\left(\frac{N_c}{N_{c0}}\right), \tag{11}$$

where we introduce the constants $r_{r0} = 1\ \mathrm{kg\,kg^{-1}}$, $N_{r0} = 1\ \mathrm{kg^{-1}}$, and $N_{c0} = 1\ \mathrm{kg^{-1}}$ in order to non-dimensionalize the log-normally distributed variables and thereby avoid taking the logarithm of dimensional quantities. The normally distributed variables r_{rn}, N_{rn}, and N_{cn} are dimensionless. Their (dimensionless) means are denoted by $\mu_{r_{rn}}$, $\mu_{N_{rn}}$, and $\mu_{N_{cn}}$, and their (dimensionless) variances are denoted by $\sigma_{r_{rn}}$, $\sigma_{N_{rn}}$, and $\sigma_{N_{cn}}$.

We can express the means and the variances of the normally distributed variables in terms of their log-normally distributed counterparts (Garvey, 2000, Appendix B):

$$\mu_{r_{rn}} = \ln\left\{\frac{\langle r_r\rangle}{r_{r0}}\left(1+\frac{\langle r_r'^2\rangle}{\langle r_r\rangle^2}\right)^{-1/2}\right\}, \tag{12}$$

$$\sigma_{r_{rn}} = \left\{\ln\left(1+\frac{\langle r_r'^2\rangle}{\langle r_r\rangle^2}\right)\right\}^{1/2}, \tag{13}$$

with analogous equations for N_r and N_c.

When a variable is transformed from log-normal to normal, its correlations with other variables change. In this paper, a linear correlation coefficient between two arbitrary variables, e.g. x and y, will be denoted

$$\rho_{(x,y)} = \frac{\langle x'y'\rangle}{\sqrt{\langle x'^2\rangle}\sqrt{\langle y'^2\rangle}}. \tag{14}$$

Such correlations may be interpreted as non-dimensionalized covariances.

First, we list a formula for the correlation between a normal variable, such as $s_{(i)}$, and a variable that has been transformed from log-normal to normal. Specifically, the correlations between $s_{(i)}$ and any of the normally distributed variables r_{rn}, N_{rn}, and N_{cn} are given in terms of the correlations between $s_{(i)}$ and the corresponding log-normal variables r_r, N_r, and N_c by (Garvey, 2000, Eq. B-1)

$$\rho_{(s_{(i)},r_{rn})} = \rho_{(s_{(i)},r_r)}\frac{\left\{\exp\left(\sigma_{r_{rn}}^2\right)-1\right\}^{1/2}}{\sigma_{r_{rn}}}, \tag{15}$$

with similar equations for N_r and N_c.

Second, we list a correlation formula for two log-normal variables (r_r and N_r) that have been transformed to normal

variables (r_{rn} and N_{rn}) (Garvey, 2000, Eq. C-3):

$$\rho_{(r_{rn},N_{rn})} = \frac{1}{\sigma_{r_{rn}}\sigma_{N_{rn}}} \ln \left\{ 1 + \rho_{(r_r,N_r)} \times \left(\exp\left(\sigma_{r_{rn}}^2\right) - 1\right)^{1/2} \times \left(\exp\left(\sigma_{N_{rn}}^2\right) - 1\right)^{1/2} \right\}. \quad (16)$$

At present, our SCM does not contain prognostic equations for the following variances and correlations: $\left\langle r_r'^2 \right\rangle$, $\left\langle N_r'^2 \right\rangle$, $\left\langle N_c'^2 \right\rangle$, $\rho_{(s_{(i)},r_r)}$, $\rho_{(s_{(i)},N_r)}$, $\rho_{(s_{(i)},N_c)}$, and $\rho_{(r_r,N_r)}$. Instead, in this study the variances are diagnosed, and the correlations are prescribed as constants on the basis of a large-eddy simulation (LES). A table of the values can be found in section 8.

Now that the log-normally distributed variables (r_r, N_r, and N_c) have been transformed into normally distributed variables, the PDF may be written as a weighted mixture of two 6-variate normal distributions:

$$P\left(w, \theta_l, r_t, r_{rn}, N_{rn}, N_{cn}\right) = (a)\, P_{(1)}\left(w, \theta_l, r_t, r_{rn}, N_{rn}, N_{cn}\right) + (1-a)\, P_{(2)}\left(w, \theta_l, r_t, r_{rn}, N_{rn}, N_{cn}\right), \quad (17)$$

where a is the mixture fraction, which is the relative weight of the first normal component, and where

$$P_{(i)}\left(w, \theta_l, r_t, r_{rn}, N_{rn}, N_{cn}\right) = \frac{1}{(2\pi)^3 \left|\Sigma_{(i)}\right|^{1/2}} \times \exp\left\{ -\frac{1}{2} \left(\mathbf{x} - \mu_{(i)}\right)^T \Sigma_{(i)}^{-1} \left(\mathbf{x} - \mu_{(i)}\right) \right\}. \quad (18)$$

Here $\mathbf{x} = (w, \theta_l, r_t, r_{rn}, N_{rn}, N_{cn})$ is a column vector of variables, and $\mu_{(i)}$ and $\Sigma_{(i)}$ denote the mean vector and covariance matrix, respectively, of the ith normal component of those variables, where i can be 1 or 2. An important note is that each subset of $\mathbf{x}$ is distributed according to a multivariate normal. The means and variances of r_{rn}, N_{rn}, and N_{cn} are equal for each normal component. For example, in the case of r_{rn}, $\mu_{r_{rn(1)}} = \mu_{r_{rn(2)}}$ and $\sigma_{r_{rn(1)}}^2 = \sigma_{r_{rn(2)}}^2$. Hence, for these variates, the two normal components overlap exactly, thereby reducing to a single normal. When transformed back, these variates are distributed according to a single log-normal.

4. Sedimentation velocity and mean volume radius

The KK formula for the sedimentation velocity of r_r is (Khairoutdinov and Kogan, 2000, Eq. 37)

$$V_{r_r} = 12000\mathrm{s}^{-1} R_{vr} - 0.2 \text{ m s}^{-1}, \quad (19)$$

where R_{vr} is the mean volume radius of the raindrops. Averaging over a grid box, we find

$$\left\langle V_{r_r} \right\rangle = 12\,000\mathrm{s}^{-1} \left\langle R_{vr} \right\rangle - 0.2 \text{ m s}^{-1}. \quad (20)$$

The mean volume radius of the raindrops is (Khairoutdinov and Kogan, 2000, Eq. 3)

$$R_{vr} = \left(\frac{4\pi\rho_l}{3\rho_a}\right)^{-1/3} r_r^{1/3} N_{rV}^{-1/3}, \quad (21)$$

where ρ_l is the density of liquid water and ρ_a is the air density. N_{rV} is the raindrop concentration per volume of air, in m^{-3}. If we note that N_{rV} is related to the number concentration per mass of air, N_r, by $N_r = N_{rV}/\rho_a$, assume that ρ_l is constant, and substitute Eq. (21) into Eq. (20), we find

$$\left\langle V_{r_r} \right\rangle = 12000 \text{ s}^{-1} \left\{ \left(\frac{4\pi\rho_l}{3}\frac{N_{r0}}{r_{r0}}\right)^{-1/3} \times \left\langle \left(\frac{r_r}{r_{r0}}\right)^{1/3} \left(\frac{N_r}{N_{r0}}\right)^{-1/3} \right\rangle \right\} - 0.2 \text{ m s}^{-1}. \quad (22)$$

To compute $\left\langle V_{r_r} \right\rangle$, we need to find the grid-box average $\left\langle (r_r/r_{r0})^\alpha \, (N_r/N_{r0})^\beta \right\rangle$, where $\alpha = 1/3$ and $\beta = -1/3$. This average is an integral of a power law of two log-normally distributed variables, r_r and N_r. To compute this integral, we transform variables from r_r and N_r, which are log-normally distributed, to r_{rn} and N_{rn}, which are normally distributed, using Eqs (9) and (10). Although r_r and N_r are non-negative, r_{rn} and N_{rn} may range from $-\infty$ to ∞. We find

$$\left\langle \left(\frac{r_r}{r_{r0}}\right)^\alpha \left(\frac{N_r}{N_{r0}}\right)^\beta \right\rangle = \left\langle \exp\left(\alpha r_{rn} + \beta N_{rn}\right) \right\rangle = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left(\alpha r_{rn} + \beta N_{rn}\right) \times P\left(r_{rn}, N_{rn}\right) \, \mathrm{d}N_{rn} \, \mathrm{d}r_{rn}. \quad (23)$$

The equation for the PDF for the bivariate normal distribution formed by r_{rn} and N_{rn}, $P\left(r_{rn}, N_{rn}\right)$, is

$$P\left(r_{rn}, N_{rn}\right) = \left(2\pi\sigma_{r_{rn}}\sigma_{N_{rn}}\sqrt{1-\rho_{(r_{rn},N_{rn})}^2}\right)^{-1} \exp\left(-\frac{1}{2}\phi\right), \quad (24)$$

where

$$\phi = \left(1-\rho_{(r_{rn},N_{rn})}^2\right)^{-1} \times \left\{ \frac{1}{\sigma_{r_{rn}}^2}\left(r_{rn}-\mu_{r_{rn}}\right)^2 + \frac{1}{\sigma_{N_{rn}}^2}\left(N_{rn}-\mu_{N_{rn}}\right)^2 - \frac{2\rho_{(r_{rn},N_{rn})}}{\sigma_{r_{rn}}\sigma_{N_{rn}}} \times \left(r_{rn}-\mu_{r_{rn}}\right)\left(N_{rn}-\mu_{N_{rn}}\right) \right\}. \quad (25)$$

To evaluate this double integral, we first integrate over N_{rn} using Gradshteyn and Ryzhik (1980, Eq. 3.323 (2)). Then, the remaining single integral is integrated over r_{rn} using the same formula (Gradshteyn and Ryzhik, 1980, Eq. 3.323 (2)):

$$\left\langle \left(\frac{r_r}{r_{r0}}\right)^\alpha \left(\frac{N_r}{N_{r0}}\right)^\beta \right\rangle = \exp\left(\mu_{r_{rn}}\alpha + \mu_{N_{rn}}\beta + \frac{1}{2}\sigma_{r_{rn}}^2\alpha^2 + \frac{1}{2}\sigma_{N_{rn}}^2\beta^2 + \rho_{(r_{rn},N_{rn})}\sigma_{r_{rn}}\alpha\sigma_{N_{rn}}\beta \right). \quad (26)$$

With this result in hand, we can find $\langle V_{r_r} \rangle$ by substituting Eq. (26) into Eq. (22).

Once a formula for $\langle V_{r_r} \rangle$ is given, then $\langle V_{N_r} \rangle$ can be found readily because the KK equations for V_{N_r} and V_{r_r} differ only by constants.

5. Autoconversion

The KK formula for the time rate of change of r_r due to autoconversion of small cloud drops to large rain (drizzle) drops is (Khairoutdinov and Kogan, 2000, Eq. 29)

$$\left(\frac{\partial r_r}{\partial t}\right)_{\text{auto}} = 7.4188 \times 10^{13}\ \text{kg kg}^{-1}\ \text{s}^{-1} \left(\frac{\rho_a}{\rho_{a0}}\right)^{-1.79} \times \left(\frac{s}{s_0}\right)^{2.47} \left(\frac{N_c}{N_{c0}}\right)^{-1.79}, \tag{27}$$

where $s_0 = 1\,\text{kg kg}^{-1}$, $N_{c0} = 1\,\text{kg}^{-1}$, and $\rho_{a0} = 1\,\text{kg m}^{-3}$. The value and units of our constant pre-factor differ from that of KK because our equation employs the droplet concentration per mass of air, N_c, instead of the concentration per volume of air, N_{cV}. The two are related by $N_c = N_{cV}/\rho_a$, where ρ_a is the air density. Also, we substitute s for r_c, because autoconversion only occurs within cloud, and there $s = r_c$.

If we average both sides of Eq. (27) and regard ρ_a as a constant in the horizontal, then we find

$$\left\langle \left(\frac{\partial r_r}{\partial t}\right)_{\text{auto}} \right\rangle = 7.4188 \times 10^{13}\ \text{kg kg}^{-1}\ \text{s}^{-1} \left(\frac{\rho_a}{\rho_{a0}}\right)^{-1.79} \times \left\langle \left(\frac{s}{s_0}\right)^{2.47} \left(\frac{N_c}{N_{c0}}\right)^{-1.79} \right\rangle. \tag{28}$$

We see that in order to compute $\left\langle \left(\frac{\partial r_r}{\partial t}\right)_{\text{auto}} \right\rangle$, we need to compute the grid-box average $\left\langle (s/s_0)^{\alpha} (N_c/N_{c0})^{\beta} \right\rangle$, where now we set $\alpha = 2.47$ and $\beta = -1.79$. This average differs from the average for sedimentation velocity, Eq. (23), because (1) s is distributed according to a normal mixture, rather than a single log-normal, and (2) the lower limit of integration for s is 0, because the variable that it replaces, r_c, cannot be negative. However, N_c, which is originally distributed as a single log-normal, is transformed via Eq. (11) to N_{cn}, which is distributed as a single normal. The grid-box average that we need is

$$\begin{aligned} \left\langle \left(\frac{s}{s_0}\right)^{\alpha} \left(\frac{N_c}{N_{c0}}\right)^{\beta} \right\rangle &= \left\langle s_*^{\alpha} \exp(\beta N_{cn}) \right\rangle \\ &= \int_0^{\infty} \int_{-\infty}^{\infty} s_*^{\alpha} \exp(\beta N_{cn}) \times \Big\{ (a)\, P_{(1)}(s_*, N_{cn}) + (1-a)\, P_{(2)}(s_*, N_{cn}) \Big\} \mathrm{d}N_{cn}\, \mathrm{d}s_* \\ &\equiv \Big(a\Big)\text{Auto}_{(1)} + \Big(1-a\Big)\text{Auto}_{(2)}. \end{aligned} \tag{29}$$

Here $P_{(i)}(s_*, N_{cn})$ is the ith bivariate normal component, where i equals 1 or 2, and $\text{Auto}_{(i)}$ is proportional to the average autoconversion rate in the ith normal component:

$$\text{Auto}_{(i)} \equiv \int_0^{\infty} \int_{-\infty}^{\infty} s_*^{\alpha} \exp(\beta N_{cn}) \times P_{(i)}(s_*, N_{cn})\, \mathrm{d}N_{cn}\, \mathrm{d}s_*. \tag{30}$$

Each individual bivariate normal component in the mixture has the functional form

$$P_{(i)}(s_*, N_{cn}) = \left(2\pi \sigma_{s_{*(i)}} \sigma_{N_{cn}} \sqrt{1 - \rho^2_{(s_{(i)}, N_{cn})}}\right)^{-1} \exp\left(-\frac{1}{2}\phi\right), \tag{31}$$

where

$$\begin{aligned} \phi = &\left(1 - \rho^2_{(s_{(i)}, N_{cn})}\right)^{-1} \\ &\times \left\{ \frac{1}{\sigma^2_{s_{*(i)}}} \Big(s_* - \mu_{s_{*(i)}}\Big)^2 + \frac{1}{\sigma^2_{N_{cn}}} \Big(N_{cn} - \mu_{N_{cn}}\Big)^2 - \frac{2\rho_{(s_{(i)}, N_{cn})}}{\sigma_{s_{*(i)}} \sigma_{N_{cn}}} \times \Big(s_* - \mu_{s_{*(i)}}\Big)\Big(N_{cn} - \mu_{N_{cn}}\Big) \right\}. \end{aligned} \tag{32}$$

In order to perform the double integral in Eq. (30), we first integrate over N_{cn} using Gradshteyn and Ryzhik (1980, Eq. 3.323 (2). Subsequently, we integrate over s_* using Gradshteyn and Ryzhik (1980, Eq. 3.462 (1).

The result is that for each bivariate component in the weighted mixture of two bivariate normal components there are three different cases, depending on whether there is variability in s_* within the ith normal component and whether there is cloudiness. For cases where $\sigma_{s_{*(i)}} > 0$, then s_* varies within the ith normal component. Regardless of cloudiness, we find

$$\begin{aligned} \text{Auto}_{(i)} = &\frac{1}{\sqrt{2\pi}} \Big(\sigma_{s_{*(i)}}\Big)^{\alpha} \exp\left[\mu_{N_{cn}}\beta + \frac{1}{2}\sigma^2_{N_{cn}}\beta^2 - \frac{1}{4}s^2_{c(i)}\right] \\ &\times \Gamma(\alpha+1)\ \ D_{-(\alpha+1)}\Big(-s_{c(i)}\Big), \end{aligned} \tag{33}$$

where:

$$s_{c(i)} = \frac{\mu_{s_{*(i)}}}{\sigma_{s_{*(i)}}} + \rho_{(s_{(i)}, N_{cn})} \sigma_{N_{cn}} \beta, \tag{34}$$

and where Γ is the Gamma function and D is the parabolic cylinder function.

In Eq. (34), $\sigma_{s_{*(i)}}$ appears in the denominator. Therefore, if $\sigma_{s_{*(i)}} = 0$, then $s_{c(i)}$ is undefined and hence so is $\text{Auto}_{(i)}$. To

avoid this, we handle the case of $\sigma_{s_{*(i)}} = 0$ specially. When $\sigma_{s_{*(i)}} = 0$, then there is no variability in s within mixture component i. That is, s is uniform within component i. The normal distribution collapses to a delta function distribution. In this limit, the parabolic cylinder function can be approximated according to Gradshteyn and Ryzhik (1980, Eq. 9.246.2).

When $\sigma_{s_{*(i)}} = 0$, then there are two distinct cases. If $\mu_{s_{*(i)}} >= 0$, then the component is entirely cloudy, and we obtain

$$\begin{aligned} &\mathrm{Auto}_{(i)} \\ &= \left(\mu_{s_{*(i)}}\right)^{\alpha} \exp\left(\mu_{N_{cn}}\beta + \frac{1}{2}\sigma^2_{N_{cn}}\beta^2\right). \end{aligned} \tag{35}$$

If $\sigma_{s_{*(i)}} = 0$ and $\mu_{s_{*(i)}} < 0$, then the component is clear, and therefore

$$\mathrm{Auto}_{(i)} = 0. \tag{36}$$

Once we have obtained the autoconversion rate $\mathrm{Auto}_{(i)}$, then we can obtain the average rate of change of r_r due to autoconversion by substituting $\mathrm{Auto}_{(i)}$ into Eq. (29), which in turn can be substituted into Eq. (28).

The KK equation for the rate of change of N_r due to autoconversion is simply proportional to the rate of change of r_r due to autoconversion (Khairoutdinov and Kogan 2000, Eq. 32). Thus, solving for the average rate of change of r_r due to autoconversion provides the necessary information to solve for the average rate of change of N_r due to autoconversion.

6. Accretion

The KK formula for accretion (collection) of cloud droplets onto raindrops may be written, in slightly modified form, as (Khairoutdinov and Kogan, 2000; see Eq. (33))

$$\begin{aligned} &\left(\frac{\partial r_r}{\partial t}\right)_{\mathrm{accr}} \\ &= 67\ \mathrm{kg\,kg^{-1}\ s^{-1}} \left(\frac{s}{s_0}\right)^{1.15} \left(\frac{r_r}{r_{r0}}\right)^{1.15}, \end{aligned} \tag{37}$$

where we have substituted s for r_c because we assume that accretion occurs only where the air is saturated. Also, we have introduced the constants $s_0 = 1\,\mathrm{kg\,kg^{-1}}$ and $r_{r0} = 1\,\mathrm{kg\,kg^{-1}}$ in order to explicitly non-dimensionalize the formula.

In order to incorporate subgrid variability and upscale to an extended grid box, the average is taken of both sides of Eq. (37):

$$\begin{aligned} &\left\langle\left(\frac{\partial r_r}{\partial t}\right)_{\mathrm{accr}}\right\rangle \\ &= 67\ \mathrm{kg\,kg^{-1}\ s^{-1}} \left\langle\left(\frac{s}{s_0}\right)^{1.15} \left(\frac{r_r}{r_{r0}}\right)^{1.15}\right\rangle. \end{aligned} \tag{38}$$

This shows that the key to determining the effect of accretion on drizzle is to compute a grid-box average of the form $\langle (s/s_0)^{\alpha} (r_r/r_{r0})^{\beta}\rangle$, where now we take $\alpha = \beta = 1.15$. This has the same form as the average that we computed above for autoconversion, if we substitute r_r for N_c (see Eq. 29). Both r_r and N_c are assumed to be distributed log-normally, and hence the mathematics of the integration is identical. For convenience of the reader, however, we now list the corresponding formulas for accretion.

In analogy to the integral that we computed for autoconversion (Eq. 29), for accretion we compute

$$\begin{aligned} &\left\langle\left(\frac{s}{s_0}\right)^{\alpha}\left(\frac{r_r}{r_{r0}}\right)^{\beta}\right\rangle = \left\langle s_*^{\alpha}\exp\left(\beta r_{rn}\right)\right\rangle \\ &= \int_0^{\infty}\int_{-\infty}^{\infty} s_*^{\alpha}\exp\left(\beta r_{rn}\right) \\ &\qquad \times \Big\{ (a)\, P_{(1)}\left(s_*, r_{rn}\right) \\ &\qquad\quad + (1-a)\, P_{(2)}\left(s_*, r_{rn}\right) \Big\} \mathrm{d}r_{rn}\, \mathrm{d}s_* \\ &= \Big(a\Big)\mathrm{Accr}_{(1)} + \Big(1-a\Big)\mathrm{Accr}_{(2)}, \end{aligned} \tag{39}$$

where for each normal component, i, we have defined

$$\begin{aligned} &\mathrm{Accr}_{(i)} \equiv \\ &\int_0^{\infty}\int_{-\infty}^{\infty} s_*^{\alpha}\exp\left(\beta r_{rn}\right) \\ &\qquad \times P_{(i)}\left(s_*, r_{rn}\right) \mathrm{d}r_{rn}\, \mathrm{d}s_*. \end{aligned} \tag{40}$$

Each component of the PDF is a bivariate normal with an analogous form to Eq. (31).

We can solve the desired integral for accretion (Eq. (40)) as we did for autoconversion (see the steps leading to Eq. (33)). We obtain

$$\begin{aligned} &\mathrm{Accr}_{(i)} \\ &= \frac{1}{\sqrt{2\pi}}\left(\sigma_{s_{*(i)}}\right)^{\alpha} \exp\left(\mu_{r_{rn}}\beta + \frac{1}{2}\sigma^2_{r_{rn}}\beta^2 \right. \\ &\qquad\qquad \left. - \frac{1}{4}s^2_{c(i)}\right) \\ &\times \Gamma\left(\alpha+1\right)\ \ D_{-(\alpha+1)}\Big(-s_{c(i)}\Big), \end{aligned} \tag{41}$$

where

$$s_{c(i)} = \frac{\mu_{s_{*(i)}}}{\sigma_{s_{*(i)}}} + \rho_{(s_{(i)},r_{rn})}\sigma_{r_{rn}}\beta, \tag{42}$$

and where Γ is the Gamma function and D is the parabolic cylinder function. As before, the case of $\sigma_{s_{*(i)}} = 0$ must be handled specially. If $\sigma_{s_{*(i)}} = 0$ and $\mu_{s_{*(i)}} >= 0$, then the ith component is cloudy, and therefore

$$\begin{aligned} &\mathrm{Accr}_{(i)} \\ &= \left(\mu_{s_{*(i)}}\right)^{\alpha} \exp\left(\mu_{r_{rn}}\beta + \frac{1}{2}\sigma^2_{r_{rn}}\beta^2\right). \end{aligned} \tag{43}$$

In contrast, if $\sigma_{s_{*(i)}} = 0$ and $\mu_{s_{*(i)}} < 0$, then component i is clear, and hence there is no accretion:

$$\mathrm{Accr}_{(i)} = 0. \tag{44}$$

To compute the average rate of change of r_r due to accretion, we may substitute $\mathrm{Accr}_{(i)}$ into Eq. (39), which in turn may be substituted into Eq. (38).

7. Condensation/evaporation

Khairoutdinov and Kogan (2000, Eq. 22) approximates the rate of change of r_r due to condensation or evaporation as

$$\left(\frac{\partial r_r}{\partial t}\right)_{\text{evap}} = 3C_{\text{evap}}G\left(T,p\right)\left(\frac{4\pi\rho_l}{3\rho_a}\right)^{2/3} \times r_r^{1/3}N_{rV}^{2/3}S, \tag{45}$$

where C_{evap} is the (dimensionless) ratio of the raindrop mean geometric radius to the raindrop mean volume radius. Following Khairoutdinov and Kogan (2000, p. 233), we specify $C_{\text{evap}} = 0.86$, which is appropriate for drizzle drops. Supersaturation, S, is dimensionless and is defined as

$$S \equiv \frac{e}{e_s(T)} - 1, \tag{46}$$

where e is the vapour pressure and $e_s(T)$ is the saturation vapour pressure over liquid. The function $G\left(T,p\right)$ is a coefficient with units of $\text{m}^2\ \text{s}^{-1}$ in the drop radius growth equation (Rogers and Yau, 1989, Eq. 7.17):

$$G\left(T,p\right) = \frac{1}{\rho_l\left\{\left(\frac{L_v}{R_vT}-1\right)\frac{L_v}{K_aT} + \frac{R_vT}{D_ve_s(T)}\right\}}, \tag{47}$$

where K_a is the coefficient of thermal conductivity of air, and D_v is the coefficient of diffusion of water vapour in air.

In our SCM, $S > 1$ is prohibited. Instead, vapour in excess of saturation is instantly condensed into cloud water via a saturation adjustment. Therefore, rainwater cannot grow by condensation and is only allowed to evaporate.

In order to account for subgrid variability in evaporation of rain, we need to phrase Eq. (45) in terms of variables that are predicted by our SCM. Therefore, we need to write supersaturation, S, in terms of extended liquid water mixing ratio, s. To do so, we approximate

$$S = \frac{e}{e_s(T)} - 1 \approx \frac{r_t}{r_s\left(T,p\right)} - 1, \tag{48}$$

where the use of r_t is valid because cloud water is assumed to be absent in subsaturated areas. To solve for S in terms of s, we combine Eq. (48), the fact that $T = T_l$ when no liquid is present, and Eq. (5). We find

$$S \approx \left(\frac{1+\Lambda\left(T_l\right)r_s\left(T_l,p\right)}{r_s\left(T_l,p\right)}\right)s. \tag{49}$$

If we substitute Eq. (49) into Eq. (45) and convert the raindrop concentration per volume of air, N_{rV}, to raindrop concentration per mass of air, N_r, via $N_r = N_{rV}/\rho_a$, then we find

$$\left(\frac{\partial r_r}{\partial t}\right)_{\text{evap}} = \chi\left(T_l\right)\left(\frac{s}{s_0}\right)\left(\frac{r_r}{r_{r0}}\right)^{1/3}\left(\frac{N_r}{N_{r0}}\right)^{2/3}, \tag{50}$$

where

$$\chi\left(T_l\right) \equiv 3C_{\text{evap}}G\left(T,p\right)\left(\frac{4\pi\rho_l}{3}r_{r0}^{1/2}N_{r0}s_0^{3/2}\right)^{2/3} \times \left(\frac{1+\Lambda\left(T_l\right)r_s\left(T_l,p\right)}{r_s\left(T_l,p\right)}\right). \tag{51}$$

To write an expression for the grid-box average evaporation rate, we substitute Eq. (51) into Eq. (50) and average both sides:

$$\left\langle\left(\frac{\partial r_r}{\partial t}\right)_{\text{evap}}\right\rangle = \left\langle\chi\left(T_l\right)\left(\frac{s}{s_0}\right)\left(\frac{r_r}{r_{r0}}\right)^{1/3}\left(\frac{N_r}{N_{r0}}\right)^{2/3}\right\rangle. \tag{52}$$

Even if pressure is assumed to be constant within a grid box, computing this average requires evaluating a quadruple integral. This is intractable. Fortunately, variations in T_l across a grid box are typically small enough to approximate $\chi\left(T_l\right)$ within the ith normal component using the average of T_l within that component, $T_{l(i)}$. That is, we replace $\chi\left(T_l\right)$ with

$$\chi\left(T_{l(i)}\right) = 3C_{\text{evap}}G\left(T_{l(i)},p\right)\left(\frac{4\pi\rho_l}{3}r_{r0}^{1/2}N_{r0}s_0^{3/2}\right)^{2/3} \times \left(\frac{1+\Lambda\left(T_{l(i)}\right)r_s\left(T_{l(i)},p\right)}{r_s\left(T_{l(i)},p\right)}\right). \tag{53}$$

(Here we have set $T = T_l$ because we assume that no liquid is present.) Using $T_{l(i)}$ accounts for some variability in T_l within a grid box and is therefore (slightly) more accurate than simply replacing T_l with $\langle T_l\rangle$ in χ. With this approximation, χ may be treated as a constant within a normal component and therefore separately pulled outside the average in Eq. (52) for each normal component. In this way, a quadruple integral is reduced to two triple integrals, one for each normal component. If we transform the log-normal variables r_r and N_r to their normally distributed counterparts r_{rn} and N_{rn}, and we let $\alpha = 1$, $\beta = 1/3$, and $\gamma = 2/3$, then we find

$$\left\langle\left(\frac{s}{s_0}\right)^{\alpha}\left(\frac{r_r}{r_{r0}}\right)^{\beta}\left(\frac{N_r}{N_{r0}}\right)^{\gamma}\right\rangle = \left\langle s_*^{\alpha}\exp\left(\beta r_{rn} + \gamma N_{rn}\right)\right\rangle. \tag{54}$$

This integral is tractable because we assume that the PDF is a trivariate (s, r_{rn}, N_{rn}) mixture of two normals:

$$\begin{aligned}\left\langle\left(\frac{\partial r_r}{\partial t}\right)_{\text{evap}}\right\rangle &= \int_{-\infty}^{0}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} s_*^{\alpha}\exp\left(\beta r_{rn} + \gamma N_{rn}\right) \\ &\quad\times\Big[(a)\,\chi\left(T_{l(1)}\right)P_{(1)}\left(s_*, r_{rn}, N_{rn}\right) \\ &\quad + (1-a)\,\chi\left(T_{l(2)}\right)P_{(2)}\left(s_*, r_{rn}, N_{rn}\right)\Big] \\ &\quad\times \mathrm{d}N_{rn}\,\mathrm{d}r_{rn}\,\mathrm{d}s_*.\end{aligned} \tag{55}$$

Here we have integrated s_* from $-\infty$ to 0, which corresponds to subsaturation, because our SCM immediately condenses vapour in excess of saturation.

If we pull $\chi\,(T_l)$ and a outside the integral in Eq. (55), then we may rewrite it as

$$\begin{aligned}&\left\langle\left(\frac{\partial r_r}{\partial t}\right)_{\text{evap}}\right\rangle \\ &= (a)\ \chi\left(T_{l(1)}\right)\ \text{Evap}_{(1)} \\ &\quad + (1-a)\ \chi\left(T_{l(2)}\right)\ \text{Evap}_{(2)},\end{aligned} \tag{56}$$

where we have defined, for each normal component, a quantity proportional to the rate of change of rain mixing ratio due to evaporation:

$$\begin{aligned}&\text{Evap}_{(i)} \equiv \\ &\int_{-\infty}^{0}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} s_*^{\alpha}\exp\left(\beta r_{rn}+\gamma N_{rn}\right) \\ &\qquad\times P_{(i)}\left(s_*, r_{rn}, N_{rn}\right) \\ &\qquad\times \mathrm{d}N_{rn}\,\mathrm{d}r_{rn}\,\mathrm{d}s_*.\end{aligned} \tag{57}$$

Instead of the bivariate PDFs shown earlier for autoconversion and accretion, now each normal component is trivariate in s_*, r_{rn}, and N_{rn}. The trivariate functional form is (Burington and May, 1970, Eq. 11.14.1)

$$\begin{aligned}&P_{(i)}\left(s_*, r_{rn}, N_{rn}\right) = \\ &\Bigg[(2\pi)^{3/2}\,\sigma_{s_{*(i)}}\sigma_{r_{rn}}\sigma_{N_{rn}} \\ &\quad\times\Bigg\{1-\left(\rho^2_{(s_{(i)},r_{rn})}+\rho^2_{(s_{(i)},N_{rn})}+\rho^2_{(N_{rn},r_{rn})}\right) \\ &\qquad+2\rho_{(s_{(i)},r_{rn})}\rho_{(s_{(i)},N_{rn})}\rho_{(r_{rn},N_{rn})}\Bigg\}^{1/2}\Bigg]^{-1} \\ &\times\exp\left(-\frac{1}{2}\phi\right),\end{aligned} \tag{58}$$

where

$$\begin{aligned}\phi = &\Bigg\{1-\left(\rho^2_{(s_{(i)},r_{rn})}+\rho^2_{(s_{(i)},N_{rn})}+\rho^2_{(r_{rn},N_{rn})}\right) \\ &\qquad+2\rho_{(s_{(i)},r_{rn})}\rho_{(s_{(i)},N_{rn})}\rho_{(r_{rn},N_{rn})}\Bigg\}^{-1} \\ &\times\Bigg\{\frac{1-\rho^2_{(r_{rn},N_{rn})}}{\sigma^2_{s_{*(i)}}}\left(s_*-\mu_{s_{*(i)}}\right)^2 \\ &\quad+\frac{1-\rho^2_{(s_{(i)},N_{rn})}}{\sigma^2_{r_{rn}}}\left(r_{rn}-\mu_{r_{rn}}\right)^2 \\ &\quad+\frac{2\left(\rho_{(s_{(i)},N_{rn})}\rho_{(r_{rn},N_{rn})}-\rho_{(s_{(i)},r_{rn})}\right)}{\sigma_{s_{*(i)}}\sigma_{r_{rn}}} \\ &\qquad\times\left(s_*-\mu_{s_{*(i)}}\right)\left(r_{rn}-\mu_{r_{rn}}\right) \\ &\quad+\frac{2\left(\rho_{(s_{(i)},r_{rn})}\rho_{(r_{rn},N_{rn})}-\rho_{(s_{(i)},N_{rn})}\right)}{\sigma_{s_{*(i)}}\sigma_{N_{rn}}} \\ &\qquad\times\left(s_*-\mu_{s_{*(i)}}\right)\left(N_{rn}-\mu_{N_{rn}}\right) \\ &\quad+\frac{2\left(\rho_{(s_{(i)},r_{rn})}\rho_{(s_{(i)},N_{rn})}-\rho_{(N_{rn},r_{rn})}\right)}{\sigma_{r_{rn}}\sigma_{N_{rn}}} \\ &\qquad\times\left(r_{rn}-\mu_{r_{rn}}\right)\left(N_{rn}-\mu_{N_{rn}}\right) \\ &\quad+\frac{1-\rho^2_{(s_{(i)},r_{rn})}}{\sigma^2_{N_{rn}}}\left(N_{rn}-\mu_{N_{rn}}\right)^2\Bigg\}.\end{aligned} \tag{59}$$

To compute the triple integral in Eq. (57), we first integrate over N_{rn} and then r_{rn} using Gradshteyn and Ryzhik (1980, Eq. 3.323 (2)). Finally, the remaining single integral in s_* is evaluated using Gradshteyn and Ryzhik (1980, Eq. 3.462 (1)). We obtain

$$\begin{aligned}&\text{Evap}_{(i)} \\ &= \frac{1}{\sqrt{2\pi}}\left(-\sigma_{s_{*(i)}}\right)^{\alpha}\exp\left(\mu_{r_{rn}}\beta+\mu_{N_{rn}}\gamma\right) \\ &\times\exp\Bigg[\frac{1}{2}\Bigg\{\left(1-\rho^2_{(s_{(i)},r_{rn})}\right)\sigma^2_{r_{rn}}\beta^2 \\ &\qquad+\left(1-\rho^2_{(s_{(i)},N_{rn})}\right)\sigma^2_{N_{rn}}\gamma^2 \\ &\qquad+2\Big(\rho_{(r_{rn},N_{rn})} \\ &\qquad\quad-\rho_{(s_{(i)},r_{rn})}\rho_{(s_{(i)},N_{rn})}\Big) \\ &\qquad\quad\times\sigma_{r_{rn}}\beta\sigma_{N_{rn}}\gamma\Bigg\}\Bigg] \\ &\times\exp\left(\frac{1}{4}s^2_{cc(i)}-\frac{\mu_{s_{*(i)}}}{\sigma_{s_{*(i)}}}s_{cc(i)}+\frac{1}{2}\frac{\mu^2_{s_{*(i)}}}{\sigma^2_{s_{*(i)}}}\right) \\ &\times\Gamma\left(\alpha+1\right)\ \ D_{-(\alpha+1)}\left(s_{cc(i)}\right),\end{aligned} \tag{60}$$

where

$$\begin{aligned}s_{cc(i)} = &\frac{\mu_{s_{*(i)}}}{\sigma_{s_{*(i)}}}+\rho_{(s_{(i)},r_{rn})}\sigma_{r_{rn}}\beta \\ &+\rho_{(s_{(i)},N_{rn})}\sigma_{N_{rn}}\gamma,\end{aligned} \tag{61}$$

and where Γ is the Gamma function and D is the Parabolic Cylinder function. The conditions on α are such that any negative number taken to the α power must result in a real number. For example, $\alpha = 1/2$ would not be acceptable, whereas $\alpha = 1/3$ or $\alpha = 2$ would be acceptable.

As before, we need to treat the case of $\sigma_{s_{*(i)}} = 0$ specially, because in that case Eq. (61) is undefined. As before, we have two sub-cases. If $\sigma_{s_{*(i)}} = 0$ and $\mu_{s_{*(i)}} \leq 0$, then the air corresponding to the ith normal component is clear. In this case, we may approximate the parabolic cylinder function according to Gradshteyn and Ryzhik (1980, Eq. 9.246.2).

Then we obtain

$$\begin{aligned}&\mathrm{Evap}_{(i)}\\&= \left(\mu_{s_{*(i)}}\right)^{\alpha} \exp\left(\mu_{r_{rn}}\beta + \mu_{N_{rn}}\gamma\right)\\&\times \exp\left[\frac{1}{2}\left\{\sigma_{r_{rn}}^2\beta^2 + \sigma_{N_{rn}}^2\gamma^2 + 2\rho_{(r_{rn},N_{rn})}\sigma_{r_{rn}}\beta\sigma_{N_{rn}}\gamma\right\}\right].\end{aligned} \tag{62}$$

In the second sub-case, $\sigma_{s_{*(i)}} = 0$ and $\mu_{s_{*(i)}} > 0$, and the ith normal component is saturated (entirely cloudy). Then we find

$$\mathrm{Evap}_{(i)} = 0. \tag{63}$$

The average rate of change of r_r due to evaporation can be found by substituting both $\mathrm{Evap}_{(i)}$ and $\chi\left(T_{l(i)}\right)$ from Eq. (53) into Eq. (56).

The KK equation for the rate of change of N_r due to evaporation is proportional to the rate of change of r_r due to evaporation (Khairoutdinov and Kogan, 2000, Eq. 23). Thus solving for the average rate of change of r_r due to evaporation provides the necessary information to solve for the average rate of change of N_r due to evaporation.

8. Prescribed parameter values

The current version of our SCM does not prognose the following variances and correlations: $\langle r_r'^2\rangle$, $\langle N_r'^2\rangle$, $\langle N_c'^2\rangle$, $\rho_{(s_{(i)},r_r)}$, $\rho_{(s_{(i)},N_r)}$, $\rho_{(s_{(i)},N_c)}$, or $\rho_{(r_r,N_r)}$. Rather, in the present simulations, the variances are diagnosed by assuming proportionality with the mean squared, and the covariances are calculated by assuming that the correlations are prescribed.

To obtain the relevant parameter values, we use a LES of the 2nd research flight (RF02) of the Second Dynamics and Chemistry of Marine Stratocumulus (DYCOMS-II) field study (Stevens *et al.*, 2003). The LES model that we use is the System for Atmospheric Modeling (SAM) (Khairoutdinov and Randall, 2003).

The parameter values that we use in our RF02 SCM simulations in Part II are displayed in Table 1, which lists variances normalized by squared mean values, and correlations. The normalized variances are extracted from an LES with a horizontal domain of 6400 m × 6400 m, but as the domain increases, the normalized variances may increase as well. Therefore, if one wishes to simulate a larger area with SCM, then the appropriate values of the variances may differ from that shown in Table 1. Choosing appropriate values in such a case is beyond the scope of the present pair of papers.

The table does not list separate correlations for each normal component because we assume that values of correlations are equal for both components. For example, we assume

$$\rho_{(s_{(1)},N_r)} = \rho_{(s_{(2)},N_r)} \equiv \rho_{(s,N_r)}. \tag{64}$$

We make an analogous assumption for the other correlations.

Profiles of the LES correlations are illustrated in Figure 1. This figure shows that the parameter values vary with level (altitude). For simplicity, our SCM assumes that the most significant variation in height occurs between the 'within cloud' levels and the 'below cloud' levels. The 'within cloud' levels are defined as any levels that contain cloud water. The 'below cloud' levels are defined as any levels that do not. To obtain a single 'within cloud' value for use in the SCM, we average the LES-calculated 'within cloud' levels with height and with time (from hours 4 to 6). The values of the 'below cloud' levels are also averaged over height and time. If the SCM determines that there is cloud at a certain vertical level, then it uses the 'within cloud' parameter values in Table 1. If not, then it uses the 'below cloud' parameter set.

Table 1. Parameter values as determined by SAM LES analysis of the DYCOMS-II RF02 case study.

Parameter	Average value within cloud	Average value below cloud
$\rho_{(s,N_r)}$	0.285	0.015
$\rho_{(s,r_r)}$	0.242	0.056
$\rho_{(s,N_c)}$	0.433	NA
$\rho_{(r_r,N_r)}$	0.786	0.886
$\frac{\langle r_r'^2\rangle}{\langle r_r\rangle^2}$	0.766	8.97
$\frac{\langle N_r'^2\rangle}{\langle N_r\rangle^2}$	0.429	12.03
$\frac{\langle N_c'^2\rangle}{\langle N_c\rangle^2}$	0.003	NA

NA, not applicable. This applies to entries involving N_c below cloud.

Figure 1 shows each of the four correlations that are needed: $\rho_{(s,r_r)}$ (Figure 1(a)), $\rho_{(s,N_r)}$ (Figure 1(b)), $\rho_{(s,N_c)}$ (Figure 1(c)), and $\rho_{(r_r,N_r)}$ (Figure 1(d)). We plot the time-averaged simulated values at each vertical level (solid lines), the average 'within cloud' values (dashed lines), and the average 'below cloud' values (dashed-dotted lines). All correlations are positive at most altitudes. The correlation between s and r_r varies with altitude, but the correlations tend to be higher within cloud than below. A similar tendency is exhibited by the correlation between s and N_r. The correlation between s and N_c is only relevant at levels with at least some cloud. The variables r_r and N_r are more highly correlated than the other variables.

9. Conclusions

This paper upscales a local microphysical scheme for drizzle marine stratocumulus (Khairoutdinov and Kogan, 2000) to a larger, grid-box scale. The upscaling is achieved by analytically integrating the Khairoutdinov–Kogan formulas over a multivariate PDF that represents subgrid variability. The main results are grid-box averaged formulas for drizzle sedimentation velocity (Eq. (22)), autoconversion rate (Eq. (28)), accretion rate (Eq. (38)), and condensation/evaporation rate (Eq. (56)). In a model implementation, these are the quantities and terms that would need to be included in the prognostic equation (3) for rainwater mixing ratio, r_r. Analogous terms would need to be implemented in the equation for raindrop number concentration per mass of air, N_r (Eq. 4).

To obtain these formulas, we use PDFs that are assumed mixtures of normal distributions or assumed log-normal distributions, and we use local microphysical formulas that are in the form of power laws. This allows analytic integration

in the case of two- and even three-variable power laws with positive but arbitrary exponents. The formulas are fairly general. They apply to any power laws. In principle, such power laws could be used to approximate a wide variety of local processes, not merely liquid-phase microphysics, but also ice- or mixed-phase microphysics, or other processes.

Upscaling microphysical processes via analytic integration offers an alternative to Monte Carlo integration (e.g. Larson *et al.*, 2005). Analytic integration has the advantage of providing exact grid-box averages. In particular, these averages do not contain the sampling noise inherent in Monte Carlo methods. However, analytic integration has the disadvantage that it can only be applied to microphysical formulas with simple functional forms, such as power laws, as in this paper, or other simple functions such as exponentials. However, if a microphysical functional form is too complex to be integrated analytically, then the functional form can be fit to simpler functional forms. The fitted functions, in turn, can be integrated analytically, providing an upscaled microphysics whose local formulas approximate the original non-integrable microphysics. Accurate fitting is often possible because simple functional forms, such as power laws, provide adjustable pre-factors and exponents, and hence they have considerable flexibility over the ranges of interest. The viability of fitting simple functional forms to microphysical process rates has been demonstrated by Khairoutdinov and Kogan (2000), who developed an entire warm-rain microphysics scheme by fitting only power laws to large-eddy simulation output. The autoconversion and accretion formulas of Khairoutdinov and Kogan (2000) are widely used, including in the Community Atmosphere Model (Morrison and Gettelman, 2008). We speculate that power laws, perhaps coupled with other simple functions, could be equally well used to fit ice microphysical formulas. In this way, a comprehensive microphysical scheme, suitable for use in a general circulation model, could be fitted and upscaled.

There are two other challenges that face both analytic and Monte Carlo methods. A first challenge is diagnosing the needed correlations. Although the correlations are prescribed in the present work, Larson *et al.* (2011) provide a general method to diagnose the correlations given the corresponding means, variances, and vertical turbulent fluxes. The method diagnoses correlations based on a semi-empirical formula that is inspired by rigorous upper and lower bounds. The method is computationally inexpensive, which is important because the number of correlations scales as the square of the number of variates. Furthermore, the method guarantees by construction that diagnosed correlations are mutually consistent with each other, that is, realizable. The method has been tested for mixed-phase Arctic clouds (Larson *et al.*, 2011), but the formulation is quite general. However, it is beyond the scope of the present manuscript to incorporate diagnostic correlations.

A second challenge of both analytic and Monte Carlo methods is computational cost. Analytic integration incurs cost primarily in the numerical computation of special functions. Monte Carlo integration incurs cost in the generation of random numbers, calculation of special functions, and possibly multiple calls to the microphysics parametrization per grid box and time step. We performed timing tests and found that the computational cost of analytic integration of the Khairoutdinov–Kogan microphysics for the DYCOMS-II RF02 case is similar to the cost of Monte Carlo integration with eight sample points per grid box and time step. Eight sample points removes most but not all of the statistical noise inherent in such Monte Carlo methods for a relatively homogeneous case such as DYCOMS-II RF02 (not shown). Analytic integration is advantageous when statistical noise is a concern. The timing tests were performed using the Sun Fortran 95 compiler version 8.6 on a single 3.2 GHz W3565 Intel Xeon processor. The Monte Carlo integration method we test is a sub-column generator called the Subgrid Importance Latin Hypercube Sampler (SILHS). It is an outgrowth of the Latin hypercube method of Larson *et al.* (2005). SILHS reduces the sampling noise by use of stratified sampling, in which sample points are chosen quasi-randomly in a way that mitigates clumping of points. SILHS also reduces noise by use of importance sampling, in which points are chosen preferentially in cloud, but this confers no benefit in fully overcast cases such as DYCOMS-II RF02. SILHS is fully multivariate and permits an arbitrary degree of overlap in the vertical.

Given that complex microphysical formulas can be fitted by simpler ones, and given that the computational cost of analytic integration is competitive, we believe that it is feasible to include multivariate analytic integration in large-scale models of the atmosphere. Indeed, the univariate integration of Morrison and Gettelman (2008) has already been included as a default setting in version 5 of the Community Atmosphere Model.

The usefulness of these formulas when implemented in a fully interactive single-column model is assessed in Part II (Griffin and Larson, 2012). There we simulate a drizzling marine stratocumulus cloud and find that the drizzle at the ocean surface is a factor of four too low when the microphysical formulas are not upscaled.

Acknowledgements

The authors are grateful for financial support provided by Grant 04-062 from the UWM Research Growth Initiative and by Grants ATM-0618818 and ATM-0936186 from the National Science Foundation.

References

Burington RS, May DC Jr. 1970. *Handbook of Probability and Statistics with Tables* (2nd edn). McGraw-Hill: New York.

Cheng A, Xu K-M. 2009. A pdf-based microphysics parameterization for simulation of drizzling boundary layer clouds. *J. Atmos. Sci.* **66**: 2317–2334.

Garvey PR. 2000. *Probability Methods for Cost Uncertainty Analysis*. Marcel Dekker: New York, NY.

Golaz J-C, Larson VE, Cotton WR. 2002a. A PDF-based model for boundary layer clouds. Part I: Method and model description. *J. Atmos. Sci.* **59**: 3540–3551.

Golaz J-C, Larson VE, Cotton WR. 2002b. A PDF-based model for boundary layer clouds. Part II: Model results. *J. Atmos. Sci.* **59**: 3552–3571.

Gradshteyn IS, Ryzhik IM. 1980. *Table of Integrals, Series, and Products*. Academic Press: New York, NY.

Griffin BM, Larson VE. 2012. Analytic upscaling of local microphysics parameterizations. Part II: Simulations. *Q. J. Royal Meteorol. Soc.* DOI: 10.1002/qj.1966.

Khairoutdinov M, Kogan Y. 2000. A new cloud physics parameterization in a large-eddy simulation model of marine stratocumulus. *Mon. Weather Rev.* **128**: 229–243.

Khairoutdinov M, Randall DA. 2003. Cloud resolving modeling of the ARM Summer 1997 IOP: model formulation, results, uncertainties, and sensitivities. *J. Atmos. Sci.* **60**: 607–624.

Larson VE. 2007. From cloud overlap to PDF overlap. *Q. J. R. Meteorol. Soc.* **133**: 1877–1891.

Larson VE, Golaz J-C. 2005. Using probability density functions to derive consistent closure relationships among higher-order moments. *Mon. Weather Rev.* **133**: 1023–1042.
Larson VE, Griffin BM. 2006. Coupling microphysics parameterizations to cloud parameterizations. *Preprints: 12th Conference on Cloud Physics*, Madison, WI, American Meteorological Society.
Larson VE, Golaz J-C, Jiang H, Cotton WR. 2005. Supplying local microphysics parameterizations with information about subgrid variability: Latin hypercube sampling. *J. Atmos. Sci.* **62**: 4010–4026.
Larson VE, Nielsen BJ, Fan J, Ovchinnikov M. 2011. Parameterizing correlations between hydrometeors in mixed-phase Arctic clouds. *J. Geophys. Res.* **116**: D00T02, DOI: 10.1029/2010JD015570.
Lewellen WS, Yoh S. 1993. Binormal model of ensemble partial cloudiness. *J. Atmos. Sci.* **50**: 1228–1237.
Morrison H, Gettelman A. 2008. A new two-moment bulk stratiform cloud microphysics scheme in the Community Atmosphere Model, Version 3 (CAM3). Part I: Description and numerical tests. *J. Climate* **21**: 3642–3659.
Pincus R, Klein SA. 2000. Unresolved spatial variability and microphysical process rates in large-scale models. *J. Geophys. Res.* **105**: 27059–27065.
Pincus R, Barker HW, Morcrette J-J. 2003. A fast, flexible, approximate technique for computing radiative transfer in inhomogeneous cloud fields. *J. Geophys. Res.* **108**: Art. No. 4376.
Pincus R, Hemler R, Klein SA. 2006. Using stochastically-generated subcolumns to represent cloud structure in a large-scale model. *Mon. Weather Rev.* **134**: 3644–3656.
Räisänen P, Barker HW. 2004. Evaluation and optimization of sampling errors for the Monte Carlo Independent Column Approximation. *Q. J. R. Meteorol. Soc.* **130**: 2069–2085.
Räisänen P, Barker HW, Khairoutdinov MF, Li J, Randall DA. 2004. Stochastic generation of subgrid-scale cloudy columns for large-scale models. *Q. J. R. Meteorol. Soc.* **130**: 2047–2067.
Räisänen P, Barker HW, Cole JNS. 2005. The Monte Carlo Independent Column Approximation's conditional random noise: impact on simulated climate. *J. Climate* **18**: 4715–4730.
Rogers RR, Yau MK. 1989. *A Short Course in Cloud Physics* (3rd edn). Butterworth-Heinemann: Woburn, MA.
Stevens B, Lenschow DH, Vali G, Gerber H, Bandy A, Blomquist B, Brenguier J-L, Bretherton CS, Burnet F, Campos T, Chai S, Faloona I, Friesen D, Haimov S, Laursen K, Lilly DK, Loehrer SM, Malinowski SP, Morley B, Petters MD, Rogers DC, Russell L, Savic-Jovcic V, Snider JR, Straub D, Szumowski MJ, Takagi H, Thornton DC, Tschudi M, Twohy C, Wetzel M, Van Zanten MC. 2003. Dynamics and chemistry of marine stratocumulus: DYCOMS-II. *Bull. Am. Meteorol. Soc.* **84**: 579–593.
Stull RB. 1988. *An Introduction to Boundary Layer Meteorology*. Kluwer: Dordrecht.
Wyant MC, Bretherton CS, Chlond A, Griffin BM, Kitagawa H, Lappen C-L, Larson VE, Lock A, Park S, de Roode SR, Uchida J, Zhao M, Ackerman AS. 2007. A single-column-model intercomparison of a heavily drizzling stratocumulus topped boundary layer. *J. Geophys. Res.* **112**: D24204, DOI: 10.1029/2007JD008536.
Zhang J, Lohmann U, Lin B. 2002. A new statistically based autoconversion rate parameterization for use in large-scale models. *J. Geophys. Res.* **107**: 4750, DOI: 10.1029/2001JD001484.

