Systematic Biases in the Microphysics and Thermodynamics of Numerical Models That Ignore Subgrid-Scale Variability

VINCENT E. LARSON

Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, Colorado

ROBERT WOOD AND PAUL R. FIELD

Meteorological Research Flight, Farnborough, United Kingdom

JEAN-CHRISTOPHE GOLAZ

Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado

THOMAS H. VONDER HAAR

Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, Colorado

WILLIAM R. COTTON

Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado

(Manuscript received 22 October 1999, in final form 18 September 2000)

ABSTRACT

A grid box in a numerical model that ignores subgrid variability has biases in certain microphysical and thermodynamic quantities relative to the values that would be obtained if subgrid-scale variability were taken into account. The biases are important because they are systematic and hence have cumulative effects. Several types of biases are discussed in this paper. Namely, numerical models that employ convex autoconversion formulas underpredict (or, more precisely, never overpredict) autoconversion rates, and numerical models that use convex functions to diagnose specific liquid water content and temperature underpredict these latter quantities. One may call these biases the ''grid box average autoconversion bias,'' ''grid box average liquid water content bias,'' and ''grid box average temperature bias,'' respectively, because the biases arise when grid box average values are substituted into formulas valid at a point, not over an extended volume. The existence of these biases can be derived from Jensen's inequality.

To assess the magnitude of the biases, the authors analyze observations of boundary layer clouds. Often the biases are small, but the observations demonstrate that the biases can be large in important cases.

In addition, the authors prove that the average liquid water content and temperature of an isolated, partly cloudy, constant-pressure volume of air cannot increase, even temporarily. The proof assumes that liquid water content can be written as a convex function of conserved variables with equal diffusivities. The temperature decrease is due to evaporative cooling as cloudy and clear air mix. More generally, the authors prove that if an isolated volume of fluid contains conserved scalars with equal diffusivities, then the average of any convex, twice-differentiable function of the conserved scalars cannot increase.

1. Introduction

Many atmospheric processes 1) can be represented in a numerical model as nonlinear functions of model-predicted variables and 2) are too small to be resolved by the model. Prior authors have noted that such nonlinear, small-scale processes pose a special difficulty for numerical models (e.g., Kogan 1998; Stevens et al. 1998; Stevens et al. 1996; Fowler et al. 1996; Fowler and Randall 1996). Namely, a model errs if it substitutes grid box average values into the nonlinear function corresponding to the small-scale process. To avoid such ''averaging errors,'' the model needs information about subgrid variability.

Corresponding author address: Vincent E. Larson, Department of Mathematical Sciences, University of Wisconsin–Milwaukee, P.O. Box 413, Milwaukee, WI 53201-0413.
E-mail: vlarson@uwm.edu

Prior authors have pointed out several undesirable consequences of ignoring subgrid variability. For example, Fowler et al. (1996) and Fowler and Randall (1996) parameterized autoconversion of cloud droplets to raindrops in a general circulation model by implementing an autoconversion parameterization developed for models with much smaller grid boxes. To obtain reasonable simulations, they were forced to reduce the threshold in specific liquid water content, q_l, at which autoconversion begins. The problem, they suggested, was that the grid box average of q_l may be much smaller than q_l in the localized areas where precipitation forms. Kogan (1998) performed a large eddy simulation of drizzling stratocumulus and then compared domain-averaged autoconversion, as calculated by the formula of Khairoutdinov and Kogan (2000), with the result of substituting domain-averaged quantities into the Khairoutdinov–Kogan formula. The two results differed markedly. Stevens et al. (1998) hypothesized that neglecting correlation terms in Reynolds-averaged microphysical equations causes some one-dimensional models to spuriously activate cloud droplets at the tops of stratocumulus layers; these authors also used a sophisticated bin-microphysics model to study the influence of averaging errors on computations of drizzle formation. Stevens et al. (1996) suggested that ignoring subgrid variability can lead to spurious supersaturation at cloud edges.

Prior authors have also discussed biases. Here, we define a bias to be an error that always has the same sign. Cahalan et al. (1994) pointed out that ignoring subgrid variability in computations of cloud albedo can lead to the so-called plane-parallel albedo bias. They noted that the bias is a consequence of a theorem known as Jensen's inequality (Jensen 1906). The origin of biases in radiative transfer applications was also discussed by Newman et al. (1995). The independent work of Rotstayn (2000) and Pincus and Klein (2000) discussed biases in autoconversion and their implications for general circulation models. Stevens et al. (1996) stated that ignoring subgrid variability leads to underpredictions of $\overline{q_l}$ and temperature, $\overline{T}$, where an overbar denotes a grid box average. Stevens et al. (1996) provided an illuminating physical interpretation of the underpredictions but did not prove that they are systematic.

The present paper proves that ignoring subgrid-scale variability leads not only to errors in certain microphysical and thermodynamic quantities, but also to biases (that are systematic). The proof follows directly from Jensen's inequality, upon noting that these quantities are represented by convex or concave functions.[1] A bias is more pernicious than an ordinary error because a bias is single-signed, rather than partially self-canceling. Our analysis provides a straightforward way to identify what conditions and parameterizations lead to (systematic) biases.

In section 2 of this paper, we first note that the Kessler autoconversion formula is convex, and review the reason that neglecting subgrid variability leads to a bias when using a convex or concave formula. Then we note that the nonlinear autoconversion parameterization of Khairoutdinov and Kogan (2000) is neither convex nor concave and hence is not associated with a (systematic) bias. To demonstrate that errors associated with the Kessler and Khairoutdinov–Kogan parameterizations can be significant in important cases, in section 3 we examine boundary layer clouds observed during the Atlantic Stratocumulus Transition Experiment (ASTEX) field experiment. In section 4, we prove that if $\overline{q_l}$ is diagnosed via a convex function and subgrid variability is ignored, then $\overline{q_l}$ and also the average temperature $\overline{T}$ are underpredicted in partly cloudy grid boxes. By deriving the autoconversion, liquid water, and temperature biases from Jensen's inequality, we make clear that all these biases have an underlying and common feature: namely, all are associated with convex formulas. In section 5, we prove that if a material fluid volume is isolated and at constant pressure, and if total specific water content q_t and liquid water temperature T_l are conserved within the volume, then $\overline{q_l}$ and $\overline{T}$ decrease or stay constant in time. The decrease is intimately related to the fact that $\overline{q_l}$ and $\overline{T}$ can be approximated as convex functions of conserved variables.

We stress that the present paper is not intended to be a critique of parametric formulas such as the Kessler or Khairoutdinov–Kogan autoconversion formulas. However, we do indirectly critique numerical models that ignore subgrid variability.

2. Grid box average autoconversion bias

We discuss the Kessler autoconversion parameterization first because it is simple, familiar, and widely used. The Kessler autoconversion rate for cloud droplets to raindrops is

$$A_K = K_1(q_l - q_{\text{crit}})H(q_l - q_{\text{crit}}), \tag{1}$$

where q_{crit} is a critical threshold in q_l below which autoconversion vanishes, K_1 is an inverse timescale, and H is the Heaviside step function (Kessler 1969). In the Pennsylvania State University–National Center for Atmospheric Research Mesoscale Model version 5, $q_{\text{crit}} = 0.5$ g kg^{-1} and $K_1 = 10^{-3}$ s^{-1} (Grell et al. 1994). The Kessler formula is displayed as the thick solid line in Fig. 1.

The Kessler formula is convex. Graphically, this means that any line segment (e.g., the dashed line in Fig. 1) that joins two points on the Kessler curve lies entirely on or above the Kessler curve, that is, in the gray shaded region in Fig. 1 (Hiriart-Urruty and Lemaréchal 1993, p. 3). Because the Kessler formula is

[1] Some authors refer to a convex function as ''concave up'' and a concave function as ''concave down.''

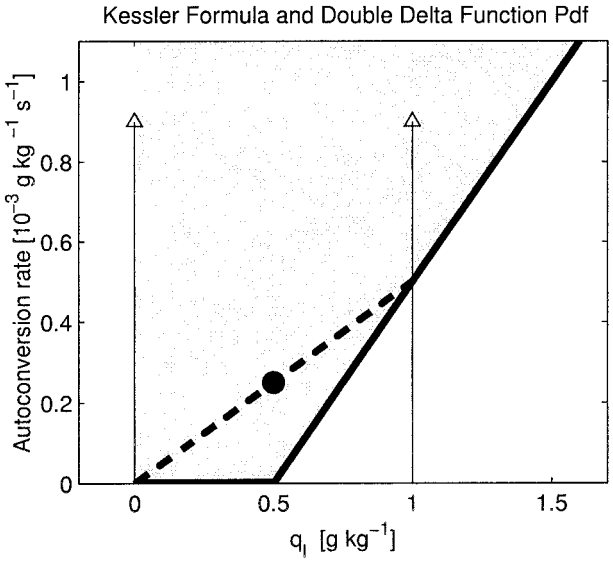

FIG. 1. The Kessler autoconversion rate (thick solid line) and a possible pdf of q_l composed of two delta functions at $q_l = 0$ g kg^{-1} and $q_l = 1$ g kg^{-1}. For this pdf, $[\overline{q_l}, \overline{A_K(q_l)}]$ is represented by the large dot. Because the Kessler formula is convex, $A_K(\overline{q_l}) = 0$ is less than $\overline{A_K(q_l)}$, and any line segment that connects two points on the Kessler formula, such as the thick dashed line shown, must lie within the gray shaded region.

convex, it turns out that if the grid box average specific liquid water content, $\overline{q_l}$, is substituted into the Kessler formula [(1)], the resulting autoconversion rate is always less than or equal to the autoconversion rate obtained by calculating the Kessler autoconversion rate at each point within the grid box and then averaging. That is, $A_K(\overline{q_l}) \leq \overline{A_K(q_l)}$, where an overbar denotes a grid box average.

Before proving this inequality, we illustrate it by a simple example. Consider a grid box that is occupied half by clear air and half by cloudy air with $q_l = 1$ g kg^{-1}. The probability density function (pdf) of q_l for this grid box consists of two equally strong Dirac delta functions, one located at $q_l = 0$ and the other at $q_l = 1$ g kg^{-1}, as depicted in Fig. 1. According to the Kessler formula, autoconversion occurs in the cloudy half of the grid box but not the clear half. The autoconversion rate, averaged over the grid box, is then $\overline{A_K(q_l)} = 0.25 \times 10^{-3}$ g kg^{-1} s^{-1}, as shown by the dot in Fig. 1. But if one inserts $\overline{q_l} = 0.5$ g kg^{-1} into the Kessler formula [(1)], then one finds that the autoconversion rate $A_K(\overline{q_l})$ is zero.

Inspection of the Kessler autoconversion curve in Fig. 1 indicates the situations in which the underestimate is large. Since the Kessler curve is linear except at q_{crit}, there is no bias if $q_l \geq q_{\text{crit}}$ everywhere in the grid box or if $q_l \leq q_{\text{crit}}$ everywhere. Also, there is no bias if q_l is uniform throughout the grid box, that is, if the pdf of q_l is a single delta function. Therefore we can surmise that the Kessler bias is most significant when $\overline{q_l}$ is near the threshold for autoconversion (i.e., the cloud is weakly drizzling) and q_l is highly variable.

To prove that a bias is associated with the Kessler formula, we use Jensen's inequality (Jensen 1906). But first, we define convex functions more formally. Consider a function $f(\Lambda_1, \ldots, \Lambda_{N_s})$ of N_s scalars. Suppose that $0 < \lambda < 1$ is a fractional weighting and that $(\Lambda_1, \ldots, \Lambda_{N_s})$ and $(\Omega_1, \ldots, \Omega_{N_s})$ are two vectors. Then f is convex if and only if

$$f[\lambda\Lambda_1 + (1-\lambda)\Omega_1, \ldots, \lambda\Lambda_{N_s} + (1-\lambda)\Omega_{N_s}] \leq \lambda f(\Lambda_1, \ldots, \Lambda_{N_s}) + (1-\lambda) f(\Omega_1, \ldots, \Omega_{N_s}), \quad (2)$$

for all such $(\Lambda_1, \ldots, \Lambda_{N_s})$, $(\Omega_1, \ldots, \Omega_{N_s})$, and λ (Dudley 1989, p. 272). A simple criterion for convexity is available in the case that $N_s = 1$ and f is twice differentiable, namely, that the second derivative of f be everywhere nonnegative: $d^2 f(\Lambda_1)/d\Lambda_1^2 \geq 0$ (Hiriart-Urruty and Lemaréchal 1993, p. 34). A function f is said to be concave when $-f$ is convex (Hiriart-Urruty and Lemaréchal 1993, p. 145). Now define an average over a grid box as

$$\overline{f(\Lambda_1, \ldots, \Lambda_{N_s})} \equiv \int d\Lambda_1 \cdots d\Lambda_{N_s} P(\Lambda_1, \ldots, \Lambda_{N_s}) f(\Lambda_1, \ldots, \Lambda_{N_s}), \quad (3)$$

where $P(\Lambda_1, \ldots, \Lambda_{N_s})$ is the joint probability density of $(\Lambda_1, \ldots, \Lambda_{N_s})$. The integral is taken over all possible values of $(\Lambda_1, \ldots, \Lambda_{N_s})$.

Then Jensen's inequality may be stated as follows (Dudley 1989, p. 273): If $f(\Lambda_1, \ldots, \Lambda_{N_s})$ is a convex function, then

$$f(\overline{\Lambda_1}, \ldots, \overline{\Lambda_{N_s}}) \leq \overline{f(\Lambda_1, \ldots, \Lambda_{N_s})}. \quad (4)$$

Therefore, substituting grid box average values $(\overline{\Lambda_1}, \ldots, \overline{\Lambda_{N_s}})$ into a convex parameterization f sometimes underpredicts and never overpredicts the quantity we usually desire, the grid box average of f, that is, $\overline{f(\Lambda_1, \ldots, \Lambda_{N_s})}$. Likewise, substituting grid box average values into a *concave* function sometimes overpredicts and never underpredicts the quantity we desire. Jensen's inequality has very general implications. Whenever a convex or concave formula of any sort is applied in a numerical model grid box, and subgrid-scale variability is ignored, a (systematic) bias occurs, relative to what would be obtained if subgrid variability were accounted for. (Of course, the bias may have negligible magnitude in practice.)

To prove that a bias is associated with the Kessler formula, we simply note that this formula is convex. Then Jensen's inequality implies that $A_K(\overline{q_l}) \leq \overline{A_K(q_l)}$.

Some autoconversion parameterizations depend on droplet number concentration, N, as well as q_l. Including the effect of N on autoconversion is important for assessing how clouds respond to a change in the concen-

tration of cloud condensation nuclei (Cotton and Anthes 1989, p. 92). When an autoconversion formula depends on two parameters, determining whether the formula is convex or concave becomes subtle. Consider the autoconversion formula of Khairoutdinov and Kogan (2000):

$$A_{\mathrm{KK}} = K_2 q_l^a N^b, \tag{5}$$

where $K_2 = 7.419 \times 10^{13}$ m$^{-5.37}$ s^{-1}, $a = 2.47$, and $b = -1.79$. A sufficient condition for the convexity of a two-dimensional function is (see Boas 1983, 170–171 and Hiriart-Urruty and Lemaréchal 1993, 190–191)

$$\frac{\partial^2 A_{\mathrm{KK}}}{\partial q_l^2} \geq 0 \qquad \frac{\partial^2 A_{\mathrm{KK}}}{\partial N^2} \geq 0$$

$$\frac{\partial^2 A_{\mathrm{KK}}}{\partial q_l^2}\frac{\partial^2 A_{\mathrm{KK}}}{\partial N^2} \geq \left(\frac{\partial^2 A_{\mathrm{KK}}}{\partial q_l \partial N}\right)^2. \tag{6}$$

The first two of these conditions are satisfied for Khairoutdinov and Kogan's values of the exponents a and b. Therefore, if N were treated as a constant parameter in A_{KK}, then ignoring subgrid variability in q_l would lead to systematic underestimates; that is, $A_{\mathrm{KK}}(\overline{q_l}, \overline{N}) \leq A_{\mathrm{KK}}(q_l, \overline{N})$. Likewise, for $b = -1.79$ the second condition implies that $A_{\mathrm{KK}}(\overline{q_l}, \overline{N}) \leq A_{\mathrm{KK}}(\overline{q_l}, N)$. However, if $a > 0$ and $b < 0$, as in the Khairoutdinov–Kogan parameterization, then the third condition in (6) holds if and only if $a + b > 1$. This condition is violated for the values of a and b used by Khairoutdinov and Kogan. Although the Khairoutdinov–Kogan formula would be convex if either N were a constant parameter or q_l were a constant parameter, the formula is not convex when both are variable. For example, consider a grid box for which q_l is proportional to N. Then $A_{\mathrm{KK}} \propto N^{0.68}$, and $N^{0.68}$ is a *concave* function of N.

3. Assessment of autoconversion errors using observational data

We have seen that when the Kessler autoconversion formula is used, a systematic bias can result from neglect of subgrid variability. This fact can be deduced solely by noting that the Kessler formula is convex. But to determine whether or not the bias is significant, one needs to examine data. To show that the Kessler bias can be large in important cases, we calculate the biases associated with a few examples of boundary layer clouds from the ASTEX field experiment. We also examine averaging errors associated with inserting mean values into the Khairoutdinov–Kogan parameterization. Using field data to assess biases has an advantage over using numerical data calculated by large eddy simulation models. Namely, observed variability of scalars in boundary layers is often greatest at long wavelengths (Cotton and Anthes 1989, 373–383), but large eddy simulations are still often restricted to horizontal domains of several kilometers or less in length; therefore, a large eddy simulation may not capture the full variability that should be accounted for within a mesoscale (e.g., 50 km) grid box.

ASTEX was an investigation of marine boundary layer clouds near the Azores over the North Atlantic Ocean. An overview of ASTEX can be found in Albrecht et al. (1995) and in a special issue (volume 52, issue 16) of the *Journal of the Atmospheric Sciences*. We examine aircraft data from long, constant-altitude flight legs made by The Met. Office C-130 aircraft in two boundary layer regimes. The first regime is a drizzling stratocumulus layer observed on the night of 12–13 June 1992 during the first Lagrangian intensive operations period (flight A209, denoted DRZ). The second regime is a cumulus-rising-into-intermittent-stratocumulus layer observed on 20 June 1992 during the second Lagrangian intensive operations period (flight A213, denoted CUSCU). We have truncated the legs to 500 s worth of data, so that all legs span approximately the same distance, about 50 km. This is a typical grid box size in a mesoscale model. Linear trends were not removed from the legs.

To quantify autoconversion errors, we use measurements of q_l and droplet number concentration N. Here q_l is obtained from a Johnson–Williams hot-wire probe and logged at 4 Hz. For the autoconversion calculations, q_l is averaged to 1 Hz. Droplet concentration N was measured with a forward scattering spectrometer probe and averaged to 1 Hz. Following Wood and Field (2000), we assume that when $N < 5$ cm^{-3}, large aerosol may be present but cloud droplets are not. In these clear areas, the Johnson–Williams probe may still record small nonzero values of q_l because of instrument noise about the zero threshold. Since the Khairoutdinov–Kogan autoconversion formula, A_{KK}, approaches infinity for vanishing N and nonzero q_l, we set q_l and A_{KK} to zero when $N < 5$ cm^{-3}, in order to avoid spurious contributions to the autoconversion rate. Cloud fraction is computed as the fraction of time (distance) for which $N > 5$ cm^{-3}.

Below we discuss biases in q_l and T. We take this opportunity to mention the additional measurements required to assess these biases. Temperature is obtained using a Rosemount deiced total temperature sensor logged at 32 Hz. To obtain total specific water content q_t, we use data from a Lyman-α hygrometer logged at 64 Hz and averaged to 32 Hz. For the q_l and T biases, q_l is interpolated to 32 Hz. However, the Johnson–Williams instrument response time is only about 1 Hz (which corresponds to roughly 100 m). This prevents an estimate of small-scale variability. Further details of the instrumentation are contained in Rogers et al. (1995).

Time series of the six legs we shall discuss are shown in Fig. 2. Three of the legs—DRZ-OV1, DRZ-OV2, and CUSCU-OV1—are entirely cloudy, that is, overcast. Three others—CUSCU-PC1, CUSCU-PC2, and CUSCU-PC3—are partly cloudy. The pdfs of q_l for these legs (along with the Kessler formula) are plotted in Fig.

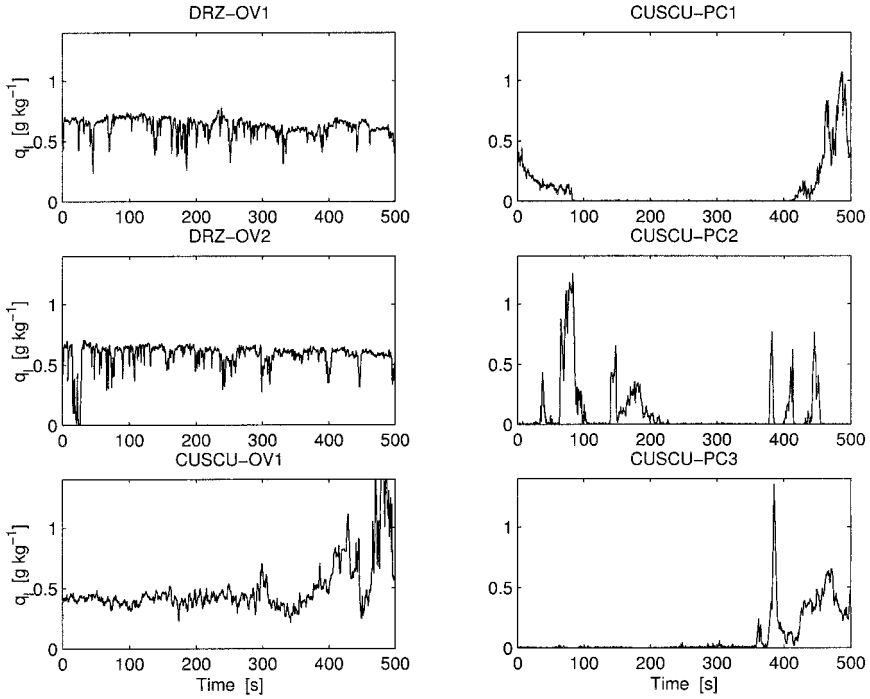

FIG. 2. Time series of specific liquid water content, q_l, for six ASTEX legs.

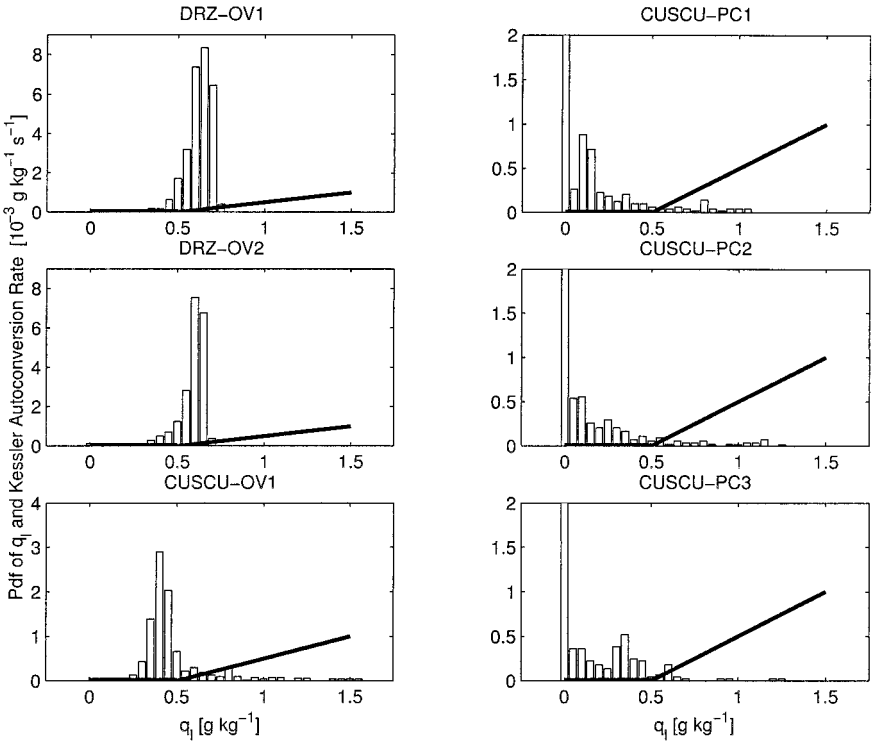

FIG. 3. The pdf of q_l and the Kessler autoconversion rate for six ASTEX legs. In leg DRZ-OV1, the pdf lies mostly to the right of the autoconversion threshold, $q_{crit} = 0.5$ g kg^{-1}, and therefore the grid box average autoconversion bias is small. In the three right-hand panels, the bin at $q_l = 0$ has been truncated by the top border of the plot.

TABLE 1. Six 50-km ASTEX legs.

	DRZ-OV1	DRZ-OV2	CUSCU-OV1	CUSCU-PC1	CUSCU-PC2	CUSCU-PC3
Date	13 Jun 1992	13 Jun 1992	20 Jun 1992	20 Jun 1992	20 Jun 1992	20 Jun 1992
Start time (UTC)	0119:11	0253:13	0723:13	0655:42	0638:09	0909:04
Cloud fraction	1.0	0.99	1.0	0.34	0.39	0.32
$\overline{q_l}$ ± standard deviation of q_l (g kg^{-1})	0.62 ± 0.070	0.58 ± 0.097	0.49 ± 0.21	0.093 ± 0.20	0.10 ± 0.22	0.088 ± 0.18
$\overline{N}$ (cm^{-3})	134	138	241	55	97	77
$\overline{A_K(q_l)}$ (g kg^{-1} s^{-1})	1.247×10^{-4}	9.541×10^{-5}	6.588×10^{-5}	1.515×10^{-5}	1.927×10^{-5}	7.562×10^{-6}
$A_K(\overline{q_l})$ (g kg^{-1} s^{-1})	1.213×10^{-4}	8.049×10^{-5}	0	0	0	0
$[\overline{A_K(q_l)} - A_K(\overline{q_l})]/\overline{A_K(q_l)}$	0.027	0.16	1	1	1	1
$\overline{A_{KK}(q_l, N)}$ (g kg^{-1} s^{-1})	2.820×10^{-6}	2.250×10^{-6}	5.424×10^{-7}	1.181×10^{-7}	7.816×10^{-8}	4.861×10^{-8}
$A_{KK}(\overline{q_l}, \overline{N})$ (g kg^{-1} s^{-1})	2.517×10^{-6}	2.010×10^{-6}	4.888×10^{-7}	1.113×10^{-7}	4.560×10^{-8}	5.220×10^{-8}
$[\overline{A_{KK}(q_l, N)} - A_{KK}(\overline{q_l}, \overline{N})]/\overline{A_{KK}(q_l, N)}$	0.11	0.11	0.099	0.058	0.42	−0.074

3. A summary of the characteristics of these legs is listed in Table 1. The biased autoconversion rates in Table 1 were computed by substituting leg-averaged values into the autoconversion formulas. The unbiased autoconversion rates were obtained by computing the autoconversion rate at each sample point along the leg and then averaging. Within the flight legs, the standard deviation in pressure is small, indicating that the aircraft was flying at almost constant altitude.

Figure 4 displays the unbiased $\overline{A_K(q_l)}$ and biased $A_K(\overline{q_l})$ Kessler autoconversion rates. The bias is significant in all cases except DRZ-OV1. This layer is fairly moist and homogeneous ($q_l = 0.62 \pm 0.070$ g kg^{-1}) and therefore the pdf of q_l lies mostly to the right of the threshold in the Kessler curve $q_{\text{crit}} = 0.5$ g kg^{-1} (see Fig. 3). For the most part, the pdf samples a linear region of the Kessler curve, and hence the bias is small. This layer illustrates the point that a strongly drizzling, fairly homogeneous layer has a small or zero Kessler autoconversion bias.

To correct errors in autoconversion due to subgrid variability, Fowler et al. (1996) and Fowler and Randall (1996) suggest predicting cloud fraction and inserting into the Kessler formula the liquid water content averaged over just the cloudy regions, rather than the whole grid box. This remedy can diminish the bias in some partly cloudy layers, but not the ones studied here. For our partly cloudy layers, even the in-cloud average liquid water content is less than the Kessler threshold, so that zero autoconversion is still predicted even when cloud fraction is taken into account. Furthermore, accounting for partial cloudiness cannot aid autoconversion prediction in fully overcast layers, and the bias can be large even when the grid box is entirely cloudy, as illustrated by leg CUSCU-OV1.

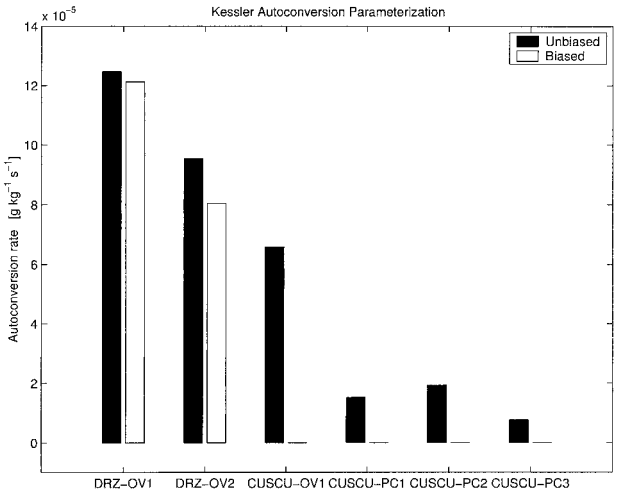

FIG. 4. Kessler autoconversion rates from six ASTEX legs. Unbiased rates, $\overline{A_K(q_l)}$, are shown in black, and biased rates, $A_K(\overline{q_l})$, in white. Because the Kessler formula is convex, $A_K(\overline{q_l}) \leq \overline{A_K(q_l)}$.

In fact, one can prove that if an autoconversion parameterization $A(q_l, N)$ is convex, then accounting for cloud fraction removes some of the autoconversion bias but still leaves a residual, systematic bias. To state this mathematically, we define $\overline{q_l}^{\text{cld}}$ and $\overline{q_l}^{\text{clr}}$ as the liquid water averaged over the cloudy and clear parts of the grid box, respectively. We adopt analogous notation for N, and we also let C denote cloud fraction. Then $\overline{q_l} = C\overline{q_l}^{\text{cld}} + (1 - C)\overline{q_l}^{\text{clr}}$ and similarly for $\overline{N}$. By definition, q_l, N, and A are zero in the clear areas. Then one may show that

$$A(\overline{q_l}, \overline{N}) \leq CA(\overline{q_l}^{\text{cld}}, \overline{N}^{\text{cld}}) \leq C\overline{A(q_l, N)}^{\text{cld}} = \overline{A(q_l, N)}. \quad (7)$$

The middle quantity, $CA(\overline{q_l}^{\text{cld}}, \overline{N}^{\text{cld}})$, is the autoconversion rate that a model would obtain if it accounted for cloud fraction. It exceeds the left-hand expression, $A(\overline{q_l}, \overline{N})$, which is what a model would obtain if it ignored all subgrid variability. But $CA(\overline{q_l}^{\text{cld}}, \overline{N}^{\text{cld}})$ still systematically underpredicts the far right-hand expression, $\overline{A(q_l, N)}$, which is the unbiased autoconversion rate. The left-hand inequality follows directly from the definition of convexity, (2). The right-hand inequality follows from Jensen's inequality, (4), applied with respect to the within-cloud average.

The Khairoutdinov–Kogan autoconversion rates, $\overline{A_{\text{KK}}(q_l, N)}$ and $A_{\text{KK}}(\overline{q_l}, \overline{N})$, are shown in Fig. 5. The Khairoutdinov–Kogan errors are smaller than the Kessler biases, but they are significant in some cases. For all cases examined except CUSCU-PC3, $A_{\text{KK}}(\overline{q_l}, \overline{N})$ underpredicts the quantity we desire, $\overline{A_{\text{KK}}(q_l, N)}$, as did

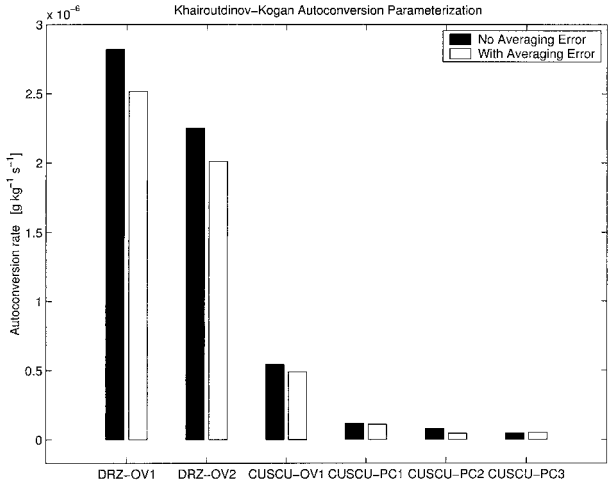

FIG. 5. Khairoutdinov–Kogan autoconversion rates for six ASTEX legs. Rates with no averaging error, $\overline{A_{KK}(q_l, N)}$, are shown in black, and rates that contain averaging errors, $A_{KK}(\overline{q_l}, \overline{N})$, in white. Here $\overline{A_{KK}(q_l, N)}$ exceeds $A_{KK}(\overline{q_l}, \overline{N})$ for all legs except CUSCU-PC3.

the Kessler formula. For CUSCU-PC3, however, $A_{KK}(\overline{q_l}, \overline{N})$ overpredicts $\overline{A_{KK}(q_l, N)}$. The overprediction is possible only because we have treated both N and q_l as variables, rather than constant parameters, in the calculation of $\overline{A_{KK}(q_l, N)}$. In this case, an error of either sign may occur, depending on the pdf associated with a leg.

Table 1 shows that the Kessler and Khairoutdinov–Kogan autoconversion rates differ greatly, a conclusion also reached in the more detailed comparison of Wood (2000). The discrepancy between the Kessler and Khairoutdinov–Kogan rates for leg DRZ-OV1 greatly exceeds the averaging error in the Khairoutdinov–Kogan formula. Over the coming years, autoconversion formulas are likely to improve and hence differences between them are likely to narrow. Such improvements will not, however, reduce averaging errors or biases, for a given grid box size. Therefore, modelers should strive to account for subgrid variability and thereby reduce averaging errors, in addition to developing physically representative autoconversion formulas.

4. Grid box average liquid water content and temperature biases

We now discuss biases that arise in partly cloudy grid boxes when average specific liquid water content, $\overline{q_l}$, and average temperature, $\overline{T}$, are diagnosed from conserved variables without accounting for subgrid-scale variability. Different numerical models perform this diagnosis using different formulas of varying accuracy. We will choose to use a diagnostic formula, (15), that requires us to introduce a variable s. We use this formula because its convexity can be assessed by inspection. Although most models do not use this particular formula (an exception is Bougeault 1981), they use formulas with similar properties.

Consider a grid-box-sized volume of air at constant pressure. Assume that all vapor in excess of saturation is instantly condensed into liquid, and that no ice forms. To facilitate discussion, we first define the liquid water temperature T_l, which is akin to the linearized liquid water potential temperature:

$$T_l = T - \frac{L}{c_p} q_l. \tag{8}$$

The variable L is the latent heat of vaporization, and c_p is the specific heat at constant pressure.

Next we need to define the variable $s = s(q_t, T_l, p) = s(\overline{q_t}, \overline{T_l}, p) + s'$, where q_t is the total specific water content (including vapor and liquid), and p is the pressure. Here s is conceptually appealing because it can be written as a function of conserved variables (and pressure); furthermore, when $q_t > q_s(T, p)$, where $q_s(T, p)$ is the saturation specific humidity, s approximates $q_l = q_t - q_s(T, p)$. However, when $q_t < q_s(T, p)$, $q_l = 0$, but s is negative. Also, in this case $s \neq q_t - q_s(T, p)$, and so s cannot be simply interpreted as the deviation of q_t from saturation.

The variable s can be written (Lewellen and Yoh 1993; see also Sommeria and Deardorff 1977; Mellor 1977):

$$s = q_t - q_s(T_l, p)\frac{(1 + \beta_1 q_t)}{[1 + \beta_1 q_s(T_l, p)]}, \tag{9}$$

where

$$q_s(T_l, p) = \frac{R_d}{R_v}\frac{e_s(T_l)}{p - [1 - (R_d/R_v)]e_s(T_l)}, \tag{10}$$

$$\beta_1 = \beta_1(T_l) = \frac{R_d}{R_v}\left(\frac{L}{R_d T_l}\right)\left(\frac{L}{c_p T_l}\right). \tag{11}$$

Here e_s is the saturation vapor pressure over liquid, and R_d and R_v are the gas constants for dry air and water vapor. When positive, s is an approximation of q_l only because the derivation of s requires that $q_s(T, p)$ be expanded about $T = T_l$. But since $(T - T_l)/T_l \ll 1$, the approximation is accurate.

Taylor expanding s about $s(\overline{q_t}, \overline{T_l}, p)$ to first order, and assuming

$$2\frac{L}{c_p \overline{T_l}}\frac{|\overline{q_t} - q_s(\overline{T_l}, p)|}{1 + \beta_1(\overline{T_l})\overline{q_t}} \ll 1, \tag{12}$$

we find the following approximation for s':

$$s' = \frac{q_t'}{1 + \beta_1(\overline{T_l})q_s(\overline{T_l}, p)} - \frac{1 + \beta_1(\overline{T_l})\overline{q_t}}{[1 + \beta_1(\overline{T_l})q_s(\overline{T_l}, p)]^2}\left.\frac{\partial q_s(T_l, p)}{\partial T_l}\right|_{T_l = \overline{T_l}} T_l'. \tag{13}$$

This formula is somewhat more accurate than the one

TABLE 2. Biases in $\overline{q_l}$ and $\overline{T}$ for three 50-km ASTEX legs.

	CUSCU-PC1	CUSCU-PC2	CUSCU-PC3
$\overline{q_l(q_t, T_l, p)}$ (g kg^{-1})	9.70×10^{-2}	1.31×10^{-1}	1.18×10^{-1}
$q_l(\overline{q_t}, \overline{T_l}, \overline{p})$ (g kg^{-1})	0	0	0
$\overline{q_l(q_t, T_l, p)} - q_l(\overline{q_t}, \overline{T_l}, \overline{p})$ (g kg^{-1})	9.70×10^{-2}	1.31×10^{-1}	1.18×10^{-1}
$\overline{T(q_t, T_l, p)}$ (K)	278.72	282.02	280.08
$T(\overline{q_t}, \overline{T_l}, \overline{p})$ (K)	278.48	281.70	279.79
$\overline{T(q_t, T_l, p)} - T(\overline{q_t}, \overline{T_l}, \overline{p})$ (K)	0.24	0.32	0.29

given by Lewellen and Yoh (1993) because Eq. (13) contains a factor in the second term, $(1 + \beta_1\overline{q_l})/(1 + \beta_1 q_s)$, that is neglected by Lewellen and Yoh. In (13), we have

$$\frac{\partial q_s(T_l, p)}{\partial T_l} = \frac{c_p}{L}\beta_1(T_l)q_s(T_l, p) \tag{14}$$

by the Clausius–Clapeyron equation (see Sommeria and Deardorff 1977).

Given $s(\overline{q_t}, \overline{T_l}, p)$ and s', q_l at a point within a volume can be approximated as follows (Bougeault 1981):

$$q_l = [s(\overline{q_t}, \overline{T_l}, p) + s']H[s(\overline{q_t}, \overline{T_l}, p) + s']. \tag{15}$$

Here H denotes the Heaviside step function. To summarize, (15) relies on two approximations: that $(T - T_l)/T_l \ll 1$, and that fluctuations in T_l and q_t are small.

Now we are ready to prove that ignoring fluctuations in s leads to an underestimation of $\overline{q_l}$. To do so, we note that in the approximate formula (15), q_l is a convex function of s'. The result then follows directly from Jensen's inequality. Most numerical models do not use the particular formula (15), but all such diagnostic formulas for q_l are approximately convex in shape.

Equation (15) resembles the Kessler autoconversion formula (1). Hence the bias in q_l has similar properties to the Kessler bias. For instance, the bias in q_l disappears if the grid box is either entirely overcast or entirely clear. Also, the bias is largest if the grid box contains clear areas that are very dry and cloudy areas that are very moist, with the grid box average near saturation.

We can also prove that ignoring fluctuations in s leads to an underestimate of $\overline{T}$. Suppose the liquid water temperature $\overline{T_l}$ is predicted without bias. Then (8) shows that if $\overline{q_l}$ is underestimated, so is $\overline{T}$.

The approximation (15) is valid only when fluctuations in q_t and T_l over a grid box are small. When the fluctuations are not negligible, but p is approximately constant, then the existence of a bias depends on the convexity of a two-dimensional function $q_l(q_t, T_l, p)$. If we make the approximation that $q_l \cong s(q_t, T_l, p)H[s(q_t, T_l, p)]$, then it turns out that q_l is neither a convex nor concave function of q_t and T_l. Therefore, one can construct cases in which neglecting subgrid variability actually overestimates $\overline{q_l}$. However, the overestimates are typically minuscule. For instance, consider an entirely cloudy volume of air with $p = 600$ hPa and $q_t = 2.2$ g kg^{-1} everywhere, but with $T_l = 254$ K in half the volume and $T_l = 255$ K in the other half. Then neglecting subgrid variability overestimates $\overline{q_l}$ by about 5×10^{-4} g kg^{-1}.

To determine the magnitude of the biases in data, we return to the three partly cloudy ASTEX legs from 20 June 1992 (flight A213, denoted CUSCU), each of which is approximately 50 km in length. Table 2 summarizes the biases for these legs. In this table, q_l is computed via the formula $q_l = s(q_t, T_l, p)H[s(q_t, T_l, p)]$ and T is computed via (8). Averaging is performed as for the autoconversion calculations. The biases are modest but not negligible.

The biases exist for shorter legs as well. To illustrate this, we divide leg CUSCU-PC2 into ~1 km segments. Most segments are either entirely cloudy or entirely clear and therefore have no bias. But some segments near cloud edges do have large biases. The segments with the four largest biases are listed in Table 3. A buoyancy deficit of the magnitude observed here (~0.4 K) could significantly lower the maximum cloud-top height attained by individual clouds in a model.

5. Decrease of $\overline{q_l}$ in an isolated system

As discussed above, the fact that q_l can be approximated as a convex function, (15), has implications for numerical modeling of clouds because convexity of q_l leads to biases in $\overline{q_l}$ and $\overline{T}$. In this section, we set aside

TABLE 3. Biases in $\overline{q_l}$ and $\overline{T}$ for four 1-km segments from ASTEX leg CUSCU-PC2.

	Segment 1	Segment 2	Segment 3	Segment 4
Start time (UTC)	0639:09	0644:29	0644:59	0645:39
$\overline{q_l(q_t, T_l, p)}$ (g kg^{-1})	4.20×10^{-1}	2.26×10^{-1}	1.55×10^{-1}	1.57×10^{-1}
$q_l(\overline{q_t}, \overline{T_l}, \overline{p})$ (g kg^{-1})	2.64×10^{-1}	0	8.20×10^{-3}	5.92×10^{-3}
$\overline{q_l(q_t, T_l, p)} - q_l(\overline{q_t}, \overline{T_l}, \overline{p})$ (g kg^{-1})	1.56×10^{-1}	2.26×10^{-1}	1.47×10^{-1}	1.51×10^{-1}
$\overline{T(q_t, T_l, p)}$ (K)	282.61	281.71	281.81	281.72
$T(\overline{q_t}, \overline{T_l}, \overline{p})$ (K)	282.22	281.15	281.45	281.34
$\overline{T(q_t, T_l, p)} - T(\overline{q_t}, \overline{T_l}, \overline{p})$ (K)	0.39	0.56	0.36	0.38

these modeling implications and turn to theoretical consequences of convexity of q_l on the physics of mixing. Namely, we prove that as a partly cloudy volume of air homogenizes through mixing, $\overline{q_l}$ decreases.

Consider an isolated, partly cloudy volume of air at constant pressure. Suppose that q_l can be adequately approximated by (15), that is, as a convex function of s'. Furthermore, suppose that T_l and q_t are conserved.[2] The time evolution of the pdf of s' determines the time evolution of $\overline{q_l}$. How does the pdf of s' evolve? Advection merely rearranges fluid particles and therefore cannot directly alter the pdf of s' within the volume. Diffusion, however, tends to homogenize s', that is, bring the pdf of s' toward a delta function. In this uniform final state,

$$\overline{q_l[s(\overline{q_t}, \overline{T_l}, p) + s']}\Big|_{\text{time}\to\infty} = q_l[s(\overline{q_t}, \overline{T_l}, p)]. \quad (16)$$

But according to Jensen's inequality [(4)], if q_l is a convex function, then the minimum possible value of $\overline{q_l}$, among all pdfs of s' that are consistent with $\overline{q_t}$ and $\overline{T_l}$, is $q_l[s(\overline{q_t}, \overline{T_l}, p)]$. Hence as time $\to \infty$, $\overline{q_l}$ achieves its minimum value.

One can go further. If pressure is constant, and $\overline{q_l}$ can be approximated as a convex function of conserved variables with equal diffusivities, one can prove that $\overline{q_l}$ in an isolated, partly cloudy system decreases or stays constant in time. Neither advection nor diffusion can cause $\overline{q_l}$ to increase, even temporarily. Using the definition of T_l, Eq. (8), we may also conclude that $\overline{T}$ decreases or remains constant in time. More generally, one can prove that the average of any convex, twice-differentiable function f of conserved variables with equal diffusivities in an isolated fluid system decreases or remains constant in time.

The physical content of the proof may be summarized as follows. Consider a conserved scalar ϕ in an isolated fluid volume and a convex function $f(\phi)$. Assume that initially the pdf of ϕ is broad because ϕ varies strongly within the fluid volume. The broad pdf samples much of the nonlinearity of f. Because f is convex, regions of anomalously large ϕ contribute proportionally more to $\overline{f}$ than regions of small ϕ. As time passes, however, diffusion leads to homogenization of ϕ and hence a narrowing of the pdf of ϕ. Then the pdf samples less of the nonlinearity in f. The contributions to $\overline{f}$ of large ϕ still outweigh those of small ϕ, but by a smaller amount. Hence $\overline{f}$ decreases.

A rather direct proof that the average of a convex function f of conserved variables decreases in time can be constructed via a simple extension of a derivation in Bilger (1989).[3] Below we present an alternative proof that emphasizes the fact that, because of diffusion, the pdf of the conserved variables evolves toward a delta function, and that the decrease in $\overline{f}$ is intimately connected to this evolution of the pdf.

First, we assume that each of N_s scalars, $\Phi(\mathbf{x}, t) = (\phi_1, \ldots, \phi_{N_s})$, is conserved. That is, each scalar ϕ_β satisfies the advection–diffusion equation,

$$\rho\frac{\partial\phi_\beta}{\partial t} + \rho u_i\frac{\partial\phi_\beta}{\partial x_i} = \frac{\partial}{\partial x_i}\left(\rho\kappa\frac{\partial}{\partial x_i}\phi_\beta\right), \quad (17)$$

where u_i is a component of velocity, ρ is fluid density, x_i is a spatial component, and t is time. We use the Einstein summation convention. We have assumed that the diffusivities κ of all scalars are equal. Second, we assume the fluid system is an isolated material fluid volume V with (constant) mass M. There is no flow or molecular diffusion of scalars through the surface enclosing the element. The bounding surface may be removed to infinity.

Now we write down the evolution equation for the joint pdf $P(\Psi; t)$ over a material fluid volume. Here $\Psi = (\psi_1, \ldots, \psi_{N_s})$ denotes a possible set of values of the conserved scalars $\Phi(\mathbf{x}, t)$. Then the governing equation of the pdf can be derived following Colucci et al. (1998):

$$\frac{\partial P(\Psi; t)}{\partial t} = -\frac{\partial}{\partial\psi_\gamma}\frac{\partial}{\partial\psi_\beta}\left(\left\langle\kappa\frac{\partial\phi_\gamma}{\partial x_i}\frac{\partial\phi_\beta}{\partial x_i}\,\middle|\,\Psi\right\rangle P(\Psi; t)\right). \quad (18)$$

The conditional expectation $\langle\kappa(\partial\phi_\gamma/\partial x_i)(\partial\phi_\beta/\partial x_i) \mid \Psi\rangle$ is the value of $\kappa(\partial\phi_\gamma/\partial x_i)(\partial\phi_\beta/\partial x_i)$ averaged over those locations where, for all α, $\phi_\alpha(\mathbf{x}, t) = \psi_\alpha$. When only one scalar is present, it is easy to recognize that the governing pdf equation resembles the diffusion equation in scalar space but with a negative ''diffusivity,'' $-\langle\kappa(\partial\phi/\partial x_i)(\partial\phi/\partial x_i) \mid \Psi\rangle$ (Chen and Kollmann 1994; Dopazo 1994). Hence the pdf evolves from an arbitrary shape to a delta function centered about the mean.

We now demonstrate that $\partial\overline{q_l}/\partial t \le 0$. When q_t and T_l vary little across the fluid volume and p is constant, the specific liquid water content averaged over the volume, $\overline{q_l}$, is given by

$$\overline{q_l} = \int_{-\infty}^{\infty} ds'\, P_{s'}(s'; \overline{q_t}, \overline{T_l}, p, t)[s(\overline{q_t}, \overline{T_l}, p) + s'] \times H[s(\overline{q_t}, \overline{T_l}, p) + s']. \quad (19)$$

Renaming the dummy variable of integration in the $\overline{q_l}$ equation [(19)] and using the Heaviside step function, we find

$$\overline{q_l} = \int_{\psi=-s(\overline{q_t},\overline{T_l},p)}^{\infty} d\psi\, P_{s'}(\psi; \overline{q_t}, \overline{T_l}, p, t)[s(\overline{q_t}, \overline{T_l}, p) + \psi]. \quad (20)$$

Since we assume that $\overline{q_t}$, $\overline{T_l}$, and p are constants over the fluid volume and in time, so is $s(\overline{q_t}, \overline{T_l}, p)$. We now assume that s' obeys an advection–diffusion equation

[2] In this paper, a variable is said to be conserved if it satisfies the advection–diffusion equation with no source term.

[3] The proof can be obtained by substituting Eq. (9) of Bilger (1989) into Eq. (1) of Bilger (1989), integrating over the fluid volume, and then noting that convexity of f ensures a decrease in $\overline{f}$. This proof can be generalized to the case in which f is a function of several conserved variables.

of the form (17). To do so, we assume that q_t and T_l are conserved and neglect the difference between their diffusivities. Then, taking the partial time derivative of (20) and substituting in the expression (18) for $\partial P(\psi; t)/\partial t$, we find

$$\frac{\partial \overline{q_l}}{\partial t} = -\int_{\psi=-s(\overline{q_t},\overline{T_l},p)}^{\infty} d\psi\ [s(\overline{q_t}, \overline{T_l}, p) + \psi] \times \frac{\partial^2}{\partial \psi^2}\left(\left\langle \kappa \frac{\partial s'}{\partial x_i}\frac{\partial s'}{\partial x_i}\middle| \psi \right\rangle P_{s'}(\psi; \overline{q_t}, \overline{T_l}, p, t)\right). \quad (21)$$

Integrating by parts yields

$$\frac{\partial \overline{q_l}}{\partial t} = -[s(\overline{q_t}, \overline{T_l}, p) + \psi] \times \frac{\partial}{\partial \psi}\left(\left\langle \kappa \frac{\partial s'}{\partial x_i}\frac{\partial s'}{\partial x_i}\middle| \psi \right\rangle P_{s'}(\psi; \overline{q_t}, \overline{T_l}, p, t)\right)\Bigg|_{\psi=-s(\overline{q_t},\overline{T_l},p)}^{\psi=\infty} + \left\langle \kappa \frac{\partial s'}{\partial x_i}\frac{\partial s'}{\partial x_i}\middle| \psi \right\rangle P_{s'}(\psi; \overline{q_t}, \overline{T_l}, p, t)\Bigg|_{\psi=-s(\overline{q_t},\overline{T_l},p)}^{\psi=\infty}. \quad (22)$$

The first term on the right-hand side vanishes if, for all times of interest, $P_{s'}|_{\psi=\infty}$ and $\partial P_{s'}/\partial\psi|_{\psi=\infty}$ tend to zero faster than ψ^{-1}. The second term vanishes at the upper limit if $P_{s'}|_{\psi=\infty} = 0$. When these conditions hold,

$$\frac{\partial \overline{q_l}}{\partial t} = -\left\langle \kappa \frac{\partial s'}{\partial x_i}\frac{\partial s'}{\partial x_i}\middle| \psi = -s(\overline{q_t}, \overline{T_l}, p) \right\rangle \times P_{s'}[\psi = -s(\overline{q_t}, \overline{T_l}, p); \overline{q_t}, \overline{T_l}, p, t] \le 0. \quad (23)$$

This concludes the proof. Note that the form of the pdf has been left quite general. Here $\overline{q_l}$ cannot increase, regardless of the pdf, as long as the pdf and its derivative vanish fast enough as s' becomes large.

Proofs such as the above are useful because if the conclusion of the proof is violated, then we know at least one assumption must also have been violated. In addition, the proof is useful because Eq. (23) provides a formula for the rate of decrease of $\overline{q_l}$. This formula depends explicitly only on quantities at cloud boundaries, that is, at $\psi = -s(\overline{q_t}, \overline{T_l}, p)$. Specifically, it depends on the dissipation rate of s' and the value of the pdf of s' at cloud edge. It is clear that $\overline{q_l}$ decreases only if $P_{s'}[\psi = -s(\overline{q_t}, \overline{T_l}, p)] \neq 0$, that is, if the volume is partly cloudy. Given a particular dissipation rate, $\overline{q_l}$ decreases rapidly if $P_{s'}[\psi = -s(\overline{q_t}, \overline{T_l}, p)]$ is large, that is, if many elements within the volume are at saturation.

The mechanism of the temperature decrease is evaporative cooling. Advection within a partly cloudy parcel brings clear and cloudy fluid particles to adjacent positions. The subsequent diffusive mixing evaporates liquid water, leading to cooling. Because advection requires time, the cooling is not realized instantaneously (Krueger 1993). The amount of cooling at a particular moment in time depends on how much mixing has occurred in the past. Even if $\overline{q_t}$ and $\overline{T_l}$ remain constant during the mixing process, the parcel ''knows'' how much mixing has occurred via its $\overline{q_l}$ or $\overline{T}$. Therefore a parcel can be said to have a ''memory'' of past mixing. This memory resides in the degree of mixedness of the parcel, that is, the finescale distribution of q_t and T_l. A numerical model that diagnoses $\overline{T}$ from $\overline{q_t}$ and $\overline{T_l}$ under the assumption of grid box homogeneity assumes, effectively, that evaporative cooling occurs instantaneously. Such models exclude memory effects and thereby omit a potentially important process in the entrainment of environmental air by clouds. Krueger (1993) has speculated that artificially rapid evaporative cooling may be the reason that some numerical simulations exhibit cloud-top entrainment instability under conditions in which it is not observed in nature.

The reasoning in the aforementioned proof need not be restricted to the particular quantity $\overline{q_l}$. The average of any twice-differentiable, convex function of conserved scalars with equal diffusivities in an isolated system can be shown to decrease or stay constant in time. Suppose that a function $f(\Psi)$ of N_s variables is such a function. Then define the average of f as

$$A_f(\Psi; t) = \int_{V_\psi} d\psi_1 \cdots d\psi_{N_s}\ P(\Psi; t) f(\Psi), \quad (24)$$

where the integral is taken over the full range of Ψ in the N_s dimensional space of scalars.

To determine whether or not A_f decays, we take the partial time derivative of (24), use the governing equation of the pdf [(18)] to substitute for $\partial P(\Psi; t)/\partial t$, and integrate by parts twice. Then we find

$$\frac{\partial A_f(\Psi; t)}{\partial t} = -\int_{A_\psi} dA_\psi\ n_\beta \left\{ f(\Psi) \frac{\partial\left[\left\langle \kappa \frac{\partial \phi_\gamma}{\partial x_i}\frac{\partial \phi_\beta}{\partial x_i}\middle| \Psi \right\rangle P(\Psi; t)\right]}{\partial \psi_\gamma} - \frac{\partial f(\Psi)}{\partial \psi_\gamma}\left\langle \kappa \frac{\partial \phi_\gamma}{\partial x_i}\frac{\partial \phi_\beta}{\partial x_i}\middle| \Psi \right\rangle P(\Psi; t)\right\} - \int_{V_\psi} d\psi_1 \cdots d\psi_{N_s} \frac{\partial^2 f(\Psi)}{\partial \psi_\gamma \partial \psi_\beta}\left\langle \kappa \frac{\partial \phi_\gamma}{\partial x_i}\frac{\partial \phi_\beta}{\partial x_i}\middle| \Psi \right\rangle P(\Psi; t).$$

The surface integral vanishes if the pdf and its derivatives vanish at the outer surface A_ψ of the N_s dimensional space of scalars Ψ. In this case, we can use the definition of a conditional average (see Colucci et al. 1998) to obtain

$$\frac{\partial A_f(\Psi; t)}{\partial t} = -\frac{1}{M}\int_V d^3x\, \rho\kappa \frac{\partial \phi_\gamma}{\partial x_i}\frac{\partial \phi_\beta}{\partial x_i}\frac{\partial^2 f(\Phi)}{\partial \phi_\gamma \partial \phi_\beta}. \quad (25)$$

Thus A_f decreases or stays constant with time if

$$\frac{\partial^2 f(\Phi)}{\partial \phi_\gamma \partial \phi_\beta}\frac{\partial \phi_\gamma}{\partial x_i}\frac{\partial \phi_\beta}{\partial x_i} \geq 0. \quad (26)$$

Assuming f is twice differentiable, this condition holds for all Φ if and only if f is convex (Hiriart-Urruty and Lemaréchal 1993, p. 190). Therefore, if f is convex, A_f decreases or remains constant in time.

The result (25) has relevance to some fast-chemistry reactions. Suppose a chemical reaction occurs much more rapidly than the timescale for molecular mixing, so that chemical equilibrium prevails at each point in the fluid. Examples of such fast-chemistry reactions in the field of turbulent combustion are discussed by Bilger (1976, 1989) and Libby and Williams (1994). When the chemical and thermal diffusivities can be approximately set equal, and initially the fuel and oxidizer are homogeneous and separated, then the concentrations of reactants and products can be written as functions of a conserved scalar. Sometimes in such reactions the reactants are everywhere convex and the products everywhere concave [as in Fig. 4a of Bilger (1976), Fig. 1 of Bilger (1989), and Fig. 3 of Libby and Williams (1994)]. Then (25) states that the average product concentration cannot decrease nor the average reactant concentration increase, even temporarily. Sometimes a reactant can be approximated, as can q_l, as a convex, continuous, piecewise linear function $f(\psi_1)$ with a slope discontinuity at a single point (Bilger 1976). In this case, A_f still decreases or remains constant with time even though f does not possess a first or second derivative everywhere. This may be seen by proceeding as in the proof for the decay of q_l, Eqs. (19)–(23). Likewise, a product of a reaction can sometimes be approximated as a concave function with a slope discontinuity at a single point; if so, the average product concentration increases or remains constant with time.

6. Conclusions

Jensen's inequality allows one to prove that (systematic) biases can arise from neglect of subgrid variability and the use of convex or concave functions (Cahalan et al. 1994). The present paper points out two examples of convex functions: the Kessler autoconversion formula, and an approximate formula for specific liquid water content, q_l, that is valid if fluctuations in thermodynamic quantities are small. If these functions are used in a numerical model, then neglecting subgrid variability leads to biases in autoconversion rate (particularly for highly variable but weakly drizzling clouds), and biases in $\overline{q_l}$ and average temperature (for partly cloudy regions). If a model contains one of these biases, then to make unbiased calculations, the model must also contain an additional, compensating bias.

The biases in $\overline{q_l}$ and mean temperature pose a difficulty for numerical models that diagnose $\overline{q_l}$. A numerical model could attempt to circumvent the difficulty by *prognosing* $\overline{q_l}$. But accurate prognosis of $\overline{q_l}$ requires information about the probability density function (pdf) of thermodynamic quantities. For instance, the rate at which a radiatively cooled parcel condenses liquid water depends on the pdf of total water content. If the parcel is highly subsaturated everywhere, no q_l forms; if the parcel is slightly subsaturated somewhere, q_l may form rapidly. To calculate the increase in $\overline{q_l}$ associated with cooling, the prognostic $\overline{q_l}$ equation of Tiedtke (1993) relies on a prognostic equation for cloud cover, which in turn assumes a reasonable but ad hoc pdf.

Do biases associated with autoconversion and $\overline{q_l}$ significantly affect simulations? Pinpointing the source of model errors is difficult because models contain myriad physical processes and feedbacks. Nevertheless, we now cite two prior papers that noted significant errors in their simulations, and we speculate that the errors may be related, either directly or in conjunction with positive feedbacks,[4] to the biases we have discussed. First, Fowler et al. (1996) and Fowler and Randall (1996) found that autoconversion is reduced in (large) general circulation model grid boxes relative to (small) mesoscale model grid boxes. If large grid boxes encompass more variability than small grid boxes, as one would expect, then according to the arguments in the present paper, underprediction of autoconversion should be more severe in large grid boxes. This is essentially how Fowler et al. (1996) and Fowler and Randall (1996) interpreted their results. Second, Sommeria and Deardorff (1977) performed two simulations of a cumulus layer at fine grid spacing (50 m): one simulation neglected subgrid variability, the other did not. The run that neglected subgrid variability tended to produce less liquid water than the other. Again, this is what one would expect from Jensen's inequality. Although these two sets of authors noticed underpredictions in their simulations, the present paper deepens our understanding by proving that certain microphysical and thermodynamic biases are systematic and placing them in a general theoretical framework that links them to the notion of convexity.

How then should modelers reduce the biases? Equation (3) shows that the average quantities we desire can be obtained without bias if the pdfs of the relevant quan-

[4] For example, a possible feedback was noted by Kristjánsson (1991), who postulated that increased latent heating may lead to increased vertical velocity in cloudy updrafts and hence further latent heating.

tities are known. Therefore we advocate predicting information about the relevant pdfs using the numerical model, as suggested by Manton and Cotton (1977), Sommeria and Deardorff (1977), and Mellor (1977), for example.

Acknowledgments. The authors would like to thank the RAF aircrew and MRF staff involved in the planning and execution of the ASTEX field campaign. The authors also thank three anonymous reviewers for their suggestions. V. E. Larson acknowledges a helpful conversation with R. L. Walko and financial support from the National Oceanic and Atmospheric Administration, Contract NA67RJ0152. W. R. Cotton and J.-Ch. Golaz acknowledge support by the National Science Foundation under NSF-WEAVE Contract ATM-9904128.

REFERENCES

Albrecht, B. A., C. S. Bretherton, D. Johnson, W. H. Schubert, and A. S. Frisch, 1995: The Atlantic stratocumulus transition experiment—ASTEX. *Bull. Amer. Meteor. Soc.,* **76,** 889–904.

Bilger, R. W., 1976: Turbulent jet diffusion flames. *Prog. Energy Combust. Sci.,* **1,** 87–109.

——, 1989: Turbulent diffusion flames. *Annu. Rev. Fluid Mech.,* **21,** 101–135.

Boas, M. L., 1983: *Mathematical Methods in the Physical Sciences.* 2d ed. John Wiley and Sons, 793 pp.

Bougeault, P., 1981: Modeling the trade-wind cumulus boundary layer. Part I: Testing the ensemble cloud relations against numerical data. *J. Atmos. Sci.,* **38,** 2414–2428.

Cahalan, R. F., W. Ridgway, W. J. Wiscombe, T. L. Bell, and J. B. Snider, 1994: The albedo of fractal stratocumulus clouds. *J. Atmos. Sci.,* **51,** 2434–2455.

Chen, J.-Y., and W. Kollmann, 1994: Comparison of prediction and measurement in non-premixed turbulent flames. *Turbulent Reacting Flows,* P. A. Libby and F. A. Williams, Eds., Academic Press, 211–308.

Colucci, P. J., F. A. Jaberi, P. Givi, and S. B. Pope, 1998: Filtered density function for large eddy simulation of turbulent reacting flows. *Phys. Fluids A,* **10,** 499–515.

Cotton, W. R., and R. A. Anthes, 1989: *Storm and Cloud Dynamics.* Academic Press, 883 pp.

Dopazo, C., 1994: Recent developments in pdf methods. *Turbulent Reacting Flows.,* P. A. Libby and F. A. Williams, Eds., Academic Press, 375–474.

Dudley, R. M., 1989: *Real Analysis and Probability.* Wadsworth and Brooks/Cole, 436 pp.

Fowler, L. D., and D. A. Randall, 1996: Liquid and ice cloud microphysics in the CSU general circulation model. Part III: Sensitivity to modeling assumptions. *J. Climate,* **9,** 561–586.

——, ——, and S. A. Rutledge, 1996: Liquid and ice cloud microphysics in the CSU general circulation model. Part I: Model description and simulated microphysical processes. *J. Climate,* **9,** 489–529.

Grell, G. A., J. Dudhia, and D. R. Stauffer, 1994: A description of the fifth-generation Penn State/NCAR Mesoscale Model (MM5). NCAR Tech. Note NCAR/TN-398 + STR, 138 pp.

Hiriart-Urruty, J.-B., and C. Lemaréchal, 1993: *Convex Analysis and Minimization Algorithms I.* Springer-Verlag, 417 pp.

Jensen, J. L. W. V., 1906: Sur les fonctions convexes et les inégalités entre les valeurs moyennes (in French). *Acta Math.,* **30,** 175–193.

Kessler, E., 1969: *On the Distribution and Continuity of Water Substance in Atmospheric Circulation. Meteor. Monogr.,* No. 10, Amer. Meteor. Soc., 84 pp.

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

Kogan, Y. L., 1998: On parameterization of cloud physics processes in mesoscale models. Preprints, *Conf. on Cloud Physics,* Everett, WA, Amer. Meteor. Soc., 348–351.

Kristjánsson, J. E., 1991: Cloud parametrization at different horizontal resolutions. *Quart. J. Roy. Meteor. Soc.,* **117,** 1255–1280.

Krueger, S. K., 1993: Linear eddy modeling of entrainment and mixing in stratus clouds. *J. Atmos. Sci.,* **50,** 3078–3090.

Lewellen, W. S., and S. Yoh, 1993: Binormal model of ensemble partial cloudiness. *J. Atmos. Sci.,* **50,** 1228–1237.

Libby, P. A., and F. A. Williams, 1994: Fundamental aspects and a review. *Turbulent Reacting Flows,* P. A. Libby and F. A. Williams, Ed., Academic Press, 1–61.

Manton, M. J., and W. R. Cotton, 1977: Formulation of approximate equations for modeling moist deep convection on the mesoscale. Atmospheric Science Paper Number 266, Colorado State University, Fort Collins, CO, 62 pp.

Mellor, G. L., 1977: The Gaussian cloud model relations. *J. Atmos. Sci.,* **34,** 356–358.

Newman, W. I., J. K. Lew, G. L. Siscoe, and R. G. Fovell, 1995: Systematic effects of randomness in radiative transfer. *J. Atmos. Sci.,* **52,** 427–435.

Pincus, R., and S. A. Klein, 2000: Unresolved spatial variability and microphysical process rates in large scale models. *J. Geophys. Res.,* **105,** 27 059–27 066.

Rogers, D. P., D. W. Johnson, and C. A. Friehe, 1995: The stable internal boundary layer over a coastal sea. Part I: Airborne measurements of the mean and turbulence structure. *J. Atmos. Sci.,* **52,** 667–683.

Rotstayn, L. D., 2000: On the ''tuning'' of autoconversion parameterizations in climate models. *J. Geophys. Res.,* **105,** 15 495–15 507.

Sommeria, G., and J. W. Deardorff, 1977: Subgrid-scale condensation in models of non-precipitating clouds. *J. Atmos. Sci.,* **34,** 344–355.

Stevens, B., R. L. Walko, W. R. Cotton, and G. Feingold, 1996: The spurious production of cloud-edge supersaturations by Eulerian models. *Mon. Wea. Rev.,* **124,** 1034–1041.

——, W. R. Cotton, and G. Feingold, 1998: A critique of one- and two-dimensional models of boundary layer clouds with a binned representation of drop microphysics. *Atmos. Res.,* **47–48,** 529–553.

Tiedtke, M., 1993: Representation of clouds in large-scale models. *Mon. Wea. Rev.,* **121,** 3040–3061.

Wood, R., 2000: The validation of drizzle parametrizations using aircraft data. Preprints, *13th Int. Conf. on Clouds and Precipitation,* Reno, NV, International Commission on Clouds and Precipitation.

——, and P. R. Field, 2000: Relationships between total water, condensed water and cloud fraction examined using aircraft data. *J. Atmos. Sci.,* **57,** 1888–1905.