# Small-Scale and Mesoscale Variability in Cloudy Boundary Layers: Joint Probability Density Functions

VINCENT E. LARSON

*Atmospheric Science Group, Department of Mathematical Sciences, University of Wisconsin–Milwaukee, Milwaukee, Wisconsin*

JEAN-CHRISTOPHE GOLAZ AND WILLIAM R. COTTON

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

(Manuscript received 5 November 2001, in final form 6 May 2002)

ABSTRACT

The joint probability density function (PDF) of vertical velocity and conserved scalars is important for at least two reasons. First, the shape of the joint PDF determines the buoyancy flux in partly cloudy layers. Second, the PDF provides a wealth of information about subgrid variability and hence can serve as the foundation of a boundary layer cloud and turbulence parameterization.

This paper analyzes PDFs of stratocumulus, cumulus, and clear boundary layers obtained from both aircraft observations and large eddy simulations. The data are used to fit five families of PDFs: a double delta function, a single Gaussian, and three PDF families based on the sum of two Gaussians.

Overall, the double Gaussian, that is binormal, PDFs perform better than the single Gaussian or double delta function PDFs. In cumulus layers with low cloud fraction, the improvement occurs because typical PDFs are highly skewed, and it is crucial to accurately represent the tail of the distribution, which is where cloud occurs. Since the double delta function has been shown in prior work to be the PDF underlying mass-flux schemes, the data analysis herein hints that mass-flux simulations may be improved upon by using a parameterization built upon a more realistic PDF.

## 1. Introduction

This paper investigates the joint probability density function (PDF) of vertical velocity $w$, liquid water potential temperature $\theta_l$, and total specific water content $q_t$, from cumulus, stratocumulus, and cumulus-rising-into-stratocumulus layers. A joint PDF $P(w, \theta_l, q_t)$ is the probability density of finding a particular value of the triplet $(w, \theta_l, q_t)$ when a measurement is made. We study joint PDFs derived from aircraft observations and large eddy simulations (LES). Our interest in joint PDFs has two related motivations.

A first motivation is that we seek a general representation of the buoyancy flux $(g/\theta_0)\overline{w'\theta_v'}$ in partly cloudy layers. Here, $\theta_v$ is the virtual potential temperature, $g$ is the acceleration due to gravity, and $\theta_0$ is a reference temperature. The buoyancy flux is important because it is the means by which convection generates turbulence. Very different formulas for buoyancy flux in partly cloudy layers can be derived depending on whether one assumes that $w$, $\theta_l$, and $q_t$ are distributed according to a Gaussian joint PDF (Mellor 1977; Sommeria and Deardorff 1977) or a convective mass-flux model (Randall 1987). In general, what is needed to accurately determine the buoyancy flux is an accurate representation of the joint PDF $P(w, \theta_l, q_t)$. Furthermore, if this PDF is known, determining the buoyancy flux requires no further information about spatial structure. This motivates us to focus directly on the PDF. If its shape can be parameterized, the buoyancy flux follows as a by-product. As will be shown below, low-cloudiness cumulus boundary layers exhibit PDFs that resemble neither Gaussians nor double delta functions (see also Wang and Stevens 2000).

A second, related motivation to study PDFs is to provide an observational test for a key assumption behind a new cloudy boundary layer parameterization (Golaz et al. 2002a,b). Traditionally, cloud parameterization has been viewed as a multiplicity of tasks. Such tasks include the prediction of heat flux, moisture flux, cloud fraction, and liquid water content. In contrast, Golaz et al. (2002a) adopt the alternative viewpoint that the goal of parameterization consists largely of a single task: the prediction of the joint PDF $P(w, \theta_l, q_t)$. The PDF viewpoint is more general because, if $P(w, \theta_l, q_t)$ is given,

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

the fluxes, cloud fraction, and liquid water content can all be diagnosed.

Unfortunately, direct calculation of $P(w, \theta_l, q_t)$ is computationally expensive. Therefore, the following methodology is followed. Golaz et al. (2002a,b) assume that the joint PDF $P(w, \theta_l, q_t)$ in each grid box and time step is the sum of two Gaussians. The double Gaussian shape defines a family of joint PDFs that depends on several variable parameters, such as the mean and variance of each *individual* Gaussian. From this family of PDFs, a particular member is selected for each grid box and time step. The member is fixed by requiring it to satisfy various moments of the PDF, such as $\overline{w}$, $\overline{w'^2}$, $\overline{w'^3}$, and $\overline{w'q_t'}$, that are predicted using standard moment equations derived from the Navier–Stokes and advection–diffusion equations. The number of prognosed moments is chosen to equal the number of PDF parameters. The PDF is then used to close various unclosed terms in the prognostic equations, such as the buoyancy flux $(g/\theta_0)\overline{w'\theta_v'}$ and $\overline{w'q_t'^2}$. Finally, the equations are advanced another timestep, and the cycle repeats. The choice of PDF is a key closure assumption. This method, the ''assumed PDF'' method, has a long history in combustion research (e.g., O'Brien 1980; Frankel et al. 1993; Bray and Libby 1994; Cook and Riley 1994) and meteorology (e.g., Manton and Cotton 1977; Sommeria and Deardorff 1977; Bougeault 1981; Chen and Cotton 1987). However, only recently have atmospheric models included $w$ in the PDF (Randall et al. 1992; Lappen and Randall 2001a,b,c).

The assumed PDF method has several advantages. First, because the fluxes, cloud fraction, and liquid water are all derived from the same joint PDF, they are all calculated in an internally consistent manner (Lappen and Randall 2001a). Second, predicting the PDF allows one to remove certain systematic biases that occur in the microphysics of models that ignore subgrid variability (Rotstayn 2000; Pincus and Klein 2000; Larson et al. 2001a,b). Third, although the moment equations require closure assumptions for the pressure and dissipation terms, many terms in these equations are derived directly from accepted theory, namely the Navier–Stokes and advection–diffusion equations. The assumed PDF method also includes empiricism in the form of the family of PDFs. However, the shape of the family of PDFs only determines the higher-order moments, and if these moments are of sufficiently high order that they do not strongly affect quantities of primary interest, such as cloud fraction, then errors due to the shape of the PDFs will be acceptably small. The errors in cloud fraction, liquid water, and liquid water flux incurred by an assumed PDF family can be assessed with observational or numerical data, and we do so in the present paper. Whether or not a double Gaussian PDF leads to tolerable errors when used in an interactive, prognostic parameterization is explored in Golaz et al. (2002a,b).

To predict buoyancy flux and, more generally, to develop a PDF-based parameterization of cloudy boundary layers, it is crucial to choose a satisfactory family of PDFs. We seek a family of PDFs that is simple, because the more parameters a PDF involves, the more moments must be prognosed. We also seek a family of PDFs that is sufficiently flexible and general that it can model both stratocumulus and cumulus regimes. Stratocumulus clouds often have Gaussian-like PDFs, whereas cumulus layers with a cloud fraction below 20% tend to have skewed PDFs (Bougeault 1981, 1982; Cuijpers and Bechtold 1995). If a parameterization package in a large-scale model includes separate schemes for separate regimes, it faces the difficulty of choosing which scheme to trigger.

This paper uses aircraft data and large eddy simulations to test how well observed PDFs are fit by five families of PDFs. The first and second families are double delta and single Gaussian functions. The third family is based on a double Gaussian functional form (Lewellen and Yoh 1993). The fourth and fifth are analytic formulations based on the double Gaussian form. We compare cloud fraction, specific liquid water content, and liquid water flux calculated from observed and parameterized PDFs. Since we desire to develop a parameterization for mesoscale models, we truncate the aircraft legs to lengths ranging from 10 to 50 km. The two large eddy simulations we perform span 6.4 km $\times$ 6.4 km and 6.7 km $\times$ 6.7 km in the horizontal.

The present paper differs from prior studies of one-dimensional PDFs of thermodynamic variables such as total water content (Bougeault 1981, 1982; Lewellen and Yoh 1993; Xu and Randall 1996; Larson et al. 2001b; Price 2001; Tompkins 2002). Such one-dimensional PDFs are useful for developing subgrid cloud schemes. However, to represent the buoyancy flux or develop a PDF-based parameterization of boundary layer turbulence, one needs a joint PDF that includes $w$. Wang and Stevens (2000) do present such joint PDFs, but we additionally fit various families of PDFs, and we test PDF-based diagnoses of cloud fraction, liquid water, and liquid water flux.

## 2. Families of PDFs to be tested

This section briefly summarizes the parameterizations of the joint three-dimensional PDF, $P(w, \theta_l, q_t)$. A fuller description is given in the appendix. For simplicity, our PDFs do not include any hydrostatic pressure variation.

### *a. Double delta function (7 parameters)*

This PDF consists of two Dirac delta functions whose locations and relative amplitude may vary. A double delta function PDF corresponds to a mass-flux scheme consisting of an updraft and downdraft plume, with no subplume variability. This PDF is perhaps the simplest that permits nonzero skewness and bimodality. Following Randall et al. (1992) and Lappen and Randall

(2001a), we fix the relative amplitude of the delta functions and their positions in the $w$ coordinate by choosing them such that the resulting PDF matches the observed values of the moments $\overline{w}$, $\overline{w'^2}$, and $\overline{w'^3}$. The positions of the delta functions in the $\theta_l$ coordinate are determined by $\overline{\theta_l}$ and $\overline{w'\theta_l'}$. Likewise, the positions in the $q_t$ coordinate are determined by $\overline{q_t}$ and $\overline{w'q_t'}$. Therefore, when we use this procedure to fit this PDF to data, we ensure that the PDF exactly satisfies the observed scalar fluxes, namely, $\overline{w'\theta_l'}$ and $\overline{w'q_t'}$.

### *b. Single Gaussian (9 parameters)*

This PDF consists of a single Gaussian that in general has nonzero correlations between the variables. That is, $\overline{w'\theta_l'}$, $\overline{w'q_t'}$, $\overline{q_t'\theta_l'} \neq 0$. The Gaussian PDF does not allow the possibility of skewness or bimodality, but, unlike the double delta PDF, it does exactly satisfy the scalar variances $\overline{\theta_l'^2}$ and $\overline{q_t'^2}$.

### *c. Lewellen–Yoh (12 parameters)*

This PDF is based on a double Gaussian function, that is, the sum of two Gaussians. The positions and relative amplitudes of the two Gaussians may vary, permitting skewed and bimodal shapes. On the other hand, if the two Gaussians overlap, the PDF can reduce to a single Gaussian. Lewellen and Yoh (1993) reduce the number of PDF parameters, so that the PDF is fully determined by the means, variances, covariances, and $\overline{w'^3}$, $\overline{\theta_l'^3}$, and $\overline{q_t'^3}$. Although the Lewellen–Yoh scheme performs the best of those we tested, it contains two complications. First, its PDF parameters cannot be solved analytically. Second, a host model may predict moments that result in unphysical values of the Lewellen–Yoh PDF parameters. A plausible procedure for limiting the PDF parameters to acceptable ranges is listed in the appendix.

### *d. Analytic Double Gaussian 1 (10 parameters)*

Like Lewellen–Yoh, this PDF is based on the double Gaussian shape with a reduced number of parameters. However, in this case the widths of the individual Gaussians in $w$ are assumed equal, and the parameters are found analytically. An analytic PDF scheme is more robust than one that requires a numerical root finder. This is an advantage if the scheme is implemented as a parameterization in a large-scale model. The only third moment required for Analytic Double Gaussian 1 is $\overline{w'^3}$. Letting $Sk_x$ denote the skewness of variable $x$, we set $Sk_{\theta_l} = 0$ and $Sk_{q_t} = 1.2Sk_w$. Analytic Double Gaussian 1 is constructed to satisfy $\overline{q_t'\theta_l'}$ exactly.

### *e. Analytic Double Gaussian 2 (10 parameters)*

This scheme is exactly like Analytic Double Gaussian 1, except that the widths of the individual Gaussians in $w$ are obtained using the formulas of Luhar et al. (1996). The widths are assumed unequal in general, and the PDF reduces to a single Gaussian when the skewness in $w$ vanishes.

## 3. Observational data and large eddy simulations

To test the aforementioned PDFs, we use two sources of data: aircraft observations and large eddy simulations (LESs). Both sources contain only ice-free boundary layer clouds.

The aircraft observations were made during the Atlantic Stratocumulus Transition Experiment (ASTEX; Albrecht et al. 1995) and the First International Satellite Cloud and Climatology Project (ISCCP) Regional Experiment (FIRE; Albrecht et al. 1988). ASTEX investigated stratocumulus layers, cumulus-rising-into-stratocumulus layers, and some cumulus layers, whereas FIRE focused more on stratocumulus layers. We fit the PDF parameterizations to the ASTEX and FIRE datasets separately. These datasets contain many cloud-free legs flown in clear boundary layers or flown above or below cloud layers. Of the 184 50-km ASTEX legs, 116 are cloud free; of the 92 50-km FIRE legs, 48 are cloud free. We keep these because we want to be able to predict the absence of cloud when appropriate. Additionally, we wish to test the PDF parameterizations independently on cumulus legs. To do so, we isolate all ASTEX legs that contain some cloud and that occurred on flights that, according to observers' notes, sampled cumulus layers with no stratocumulus or at most broken stratocumulus above. This yields a third dataset, denoted ''ASTEX cumulus legs,'' that consists of eight legs. Six of these legs had cloud fractions $C \lesssim 0.02$, one had $C = 0.08$, and one had $C = 0.24$. (In general $0 \leq C \leq 1$.)

The aircraft was a C-130 operated by Meteorological Research Flight, a branch of the Met Office. Its instrumentation is described in Rogers et al. (1995). Here we merely mention the following. Liquid water content was measured with a Johnson–Williams hot wire probe, which has a response time of about 1 s. The vertical velocity was measured with an estimated accuracy of $\pm 0.1$ m s$^{-1}$. Temperature was measured using a Rosemount de-iced total temperature sensor and corrected for dynamic heating effects. In ASTEX, total water content was measured with a fast–response Lyman-$\alpha$ hygrometer, logged at 64 Hz, referenced to out-of-cloud data from a General Eastern 1011B dew/frost point hygrometer. In FIRE, total water content was obtained by adding the liquid and vapor content obtained from the Johnson–Williams probe and the dew/frost point hygrometer.

Because the set of observed cumulus legs is small, we supplement it with LES of two cumulus cases, both simulated by the Global Energy and Water Cylee Experiment (GEWEX) Cloud System Studies (GCSS) boundary layer cloud working group. The first is a con-

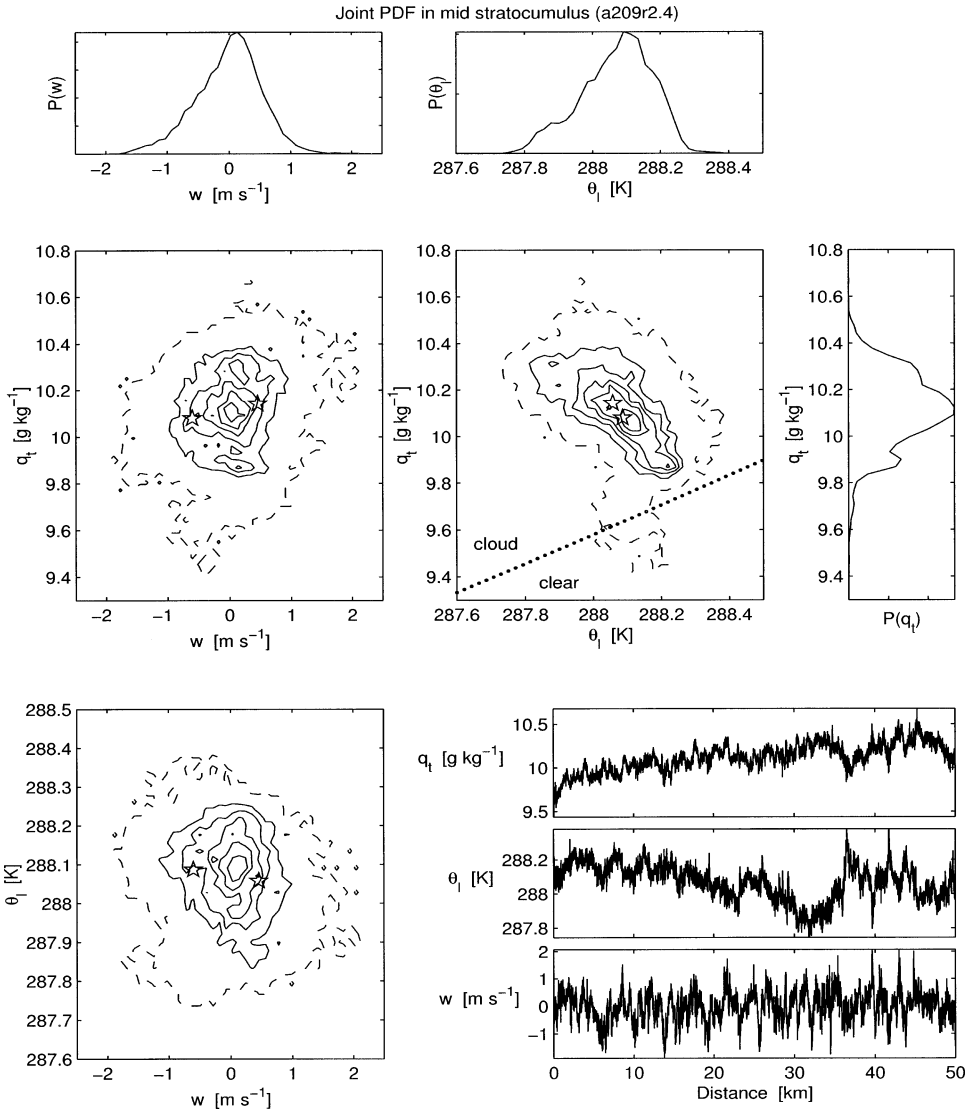


FIG. 1. One- and two-dimensional projections of a joint PDF constructed from an aircraft transect through the middle of a stratocumulus cloud observed during ASTEX. The dashed contour in the two-dimensional PDFs is at 0.005 times the maximum of the PDF. The solid contours are evenly spaced. The dotted line in the $q_t - \theta_l$ panel corresponds to saturation. Above and to the left of this line all points contain liquid. The stars mark the location of the best fit double delta function. Shown in the lower-right corner are the space series used to construct the PDF.

tinental cumulus case based on measurements over the Atmospheric Radiation Measurement (ARM) program Cloud and Radiation Testbed (CART) site in Oklahoma. The second is a trade wind cumulus case based on observations from the Barbados Oceanographic and Meteorological Experiment (BOMEX).

The LES model is the Regional Atmospheric Modeling System (RAMS; Pielke et al. 1992). All ice and precipitation processes in the model are turned off. Any water vapor in excess of saturation is immediately condensed into liquid. In the ARM simulation, the horizontal and vertical grid spacings are 100 m and 40 m, respectively, and the domain is 6.7 km $\times$ 6.7 km $\times$ 4.4 km high. In the BOMEX simulation, the horizontal and vertical grid spacings are 100 m and 40 m, respectively, and the domain is 6.4 km $\times$ 6.4 km $\times$ 3 km high. (Details of the setups can be found online at http://www.atmos.washington.edu/breth/GCSS/GCSS.html.) The ARM simulation we use was submitted as part of the GCSS intercomparison workshop. The BOMEX simulation we use was performed subsequent to the BOMEX intercomparison workshop. Both simulations we use compare well with other models' results from the GCSS workshops.

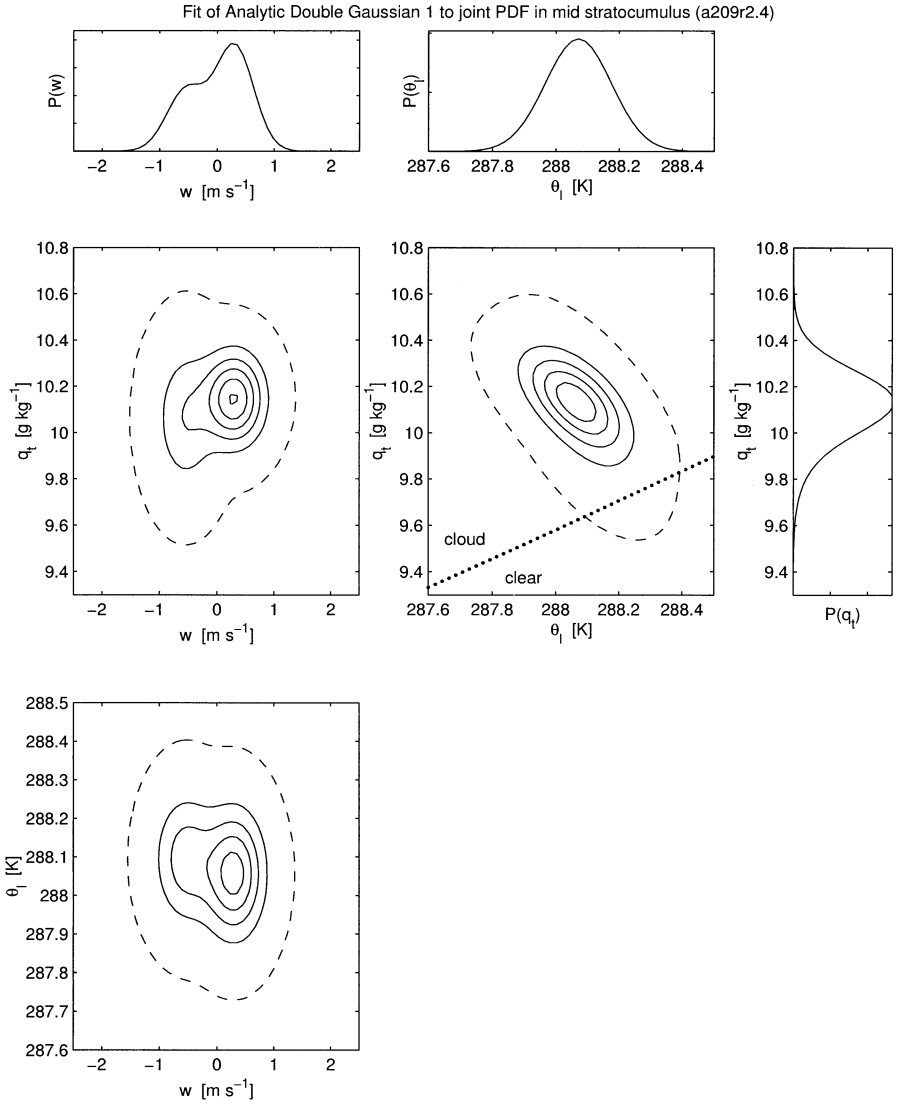


FIG. 2. The best fit of Analytic Double Gaussian 1 for the stratocumulus joint PDF shown in Fig. 1. The axes and contours are the same as in Fig. 1. The dotted line in the $q_t - \theta_l$ panel corresponds to saturation. Above and to the left of this line all points contain liquid. The upward pointing arrow in the top-right panel denotes a Dirac delta function.

## 4. Examples of joint PDFs from aircraft data

Before testing the PDF parameterizations quantitatively, we first plot two-dimensional slices of observed and fitted PDFs to give a sense of their overall character.

The first PDF is derived from an ASTEX aircraft leg through the middle of a stratocumulus layer (Fig. 1). The joint PDF is unimodal and relatively unskewed, with a quasi-Gaussian shape. The PDF is spread out and does not resemble a double delta function. There is some negative correlation between $q_t$ and $\theta_l$. This was commonly observed in the PDFs and may correspond to legs that have little variability in temperature but large variability in liquid water. For comparison, a one-dimensional PDF of liquid water using this leg was plotted in Fig. 2a of Larson et al. (2001b). To give a sense of how well a double Gaussian can fit such a PDF, Fig. 2 displays the best-fit Analytic Double Gaussian 1. This fitting formula approximates the broad quasi-Gaussian shape, yields an appropriately large cloud fraction, and permits negative correlation between $q_t$ and $\theta_l$.

The second PDF is derived from an ASTEX aircraft leg through cumuli (Fig. 3). In contrast to the stratocumulus PDF, the cumulus PDF is highly skewed. In the cumulus case, the cloud portion of the PDF consists of a long tail that does not resemble a delta function. The negative correlation of $q_t$ and $\theta_l$ is stronger here than in the stratocumulus leg. A one-dimensional PDF using this

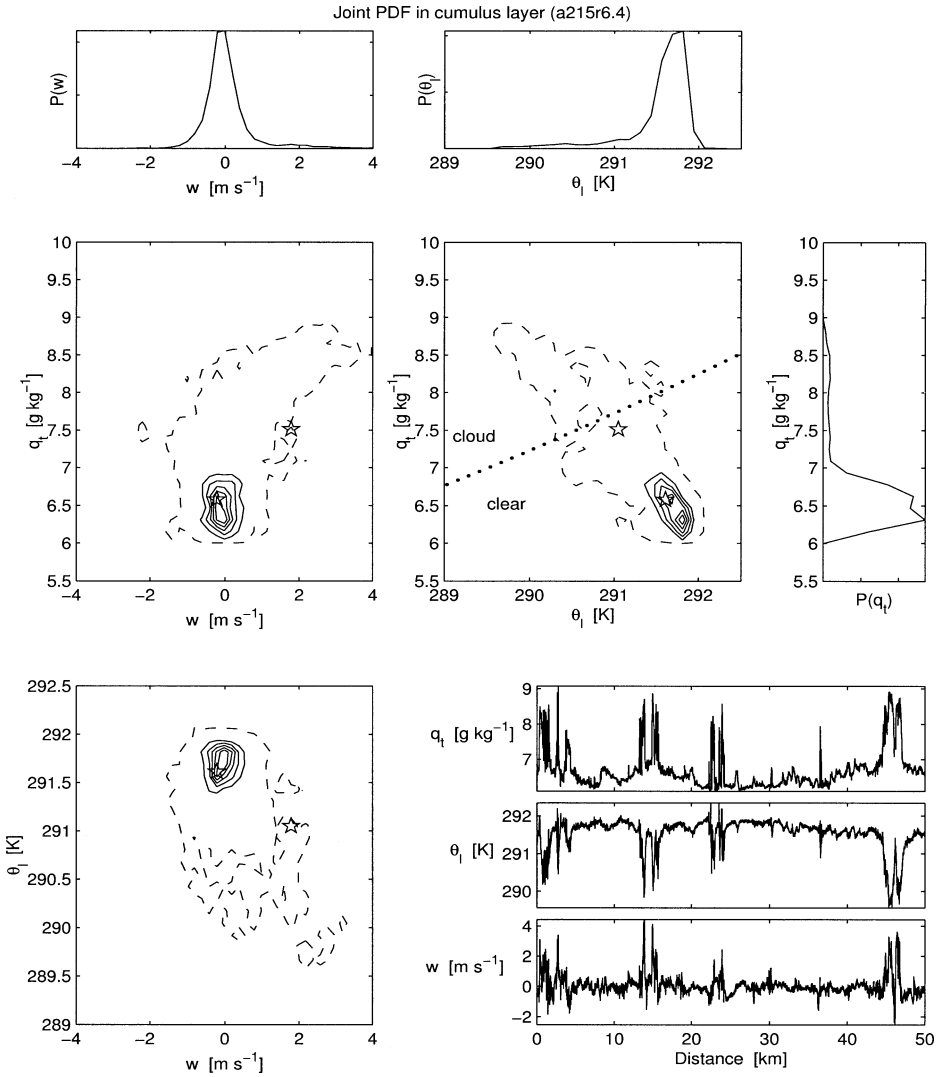


FIG. 3. One- and two-dimensional projections of a joint PDF of an aircraft transect through a cumulus layer observed during ASTEX. The dashed contour in the two-dimensional PDFs is at 0.005 times the maximum of the PDF. The solid contours are evenly spaced. The dotted line in the $q_t - \theta_l$ panel corresponds to saturation. Above and to the left of this line all points contain liquid. The stars mark the location of the best fit double delta function. Shown in the lower-right corner are the space series used to construct the PDF.

leg was plotted in Fig. 2d of Larson et al. (2001b). The best fit provided by Analytic Double Gaussian 1 is shown in Fig. 4. In this particular example, the updraft ''plume'' is diagnosed to have zero width in $\theta_l$, that is, to be a delta function in $\theta_l$, which is unphysical. However, the fitting formula does yield within-updraft variability in $q_t$ and $w$. The fitted PDF has significant skewness and an associated tail containing cloud, but the tail is somewhat too long. In summary, although Analytic Double Gaussian 1 sometimes produces unrealistic features, it is capable of distinguishing qualitatively between stratocumulus PDFs that are relatively unskewed and cumulus PDFs that have long tails.

## 5. Evaluations of PDF parameterizations

Now we quantify the errors induced by approximating an observed PDF with a parameterized PDF. We adopt the following procedure. First, we use an aircraft leg to compute observed values of cloud fraction ($C$), specific liquid water content ($\overline{q_l}$), and liquid water flux ($\overline{w'q_l'}$). Second, we use the moments computed from the aircraft leg and the five families of PDFs to parameterize the leg's PDF, as described in the appendix. Then, using the parameterized PDF, we compute parameterized values of $C$, $\overline{q_l}$, and $\overline{w'q_l'}$ as in the appendix. Finally we compare the parameterized and observed values.

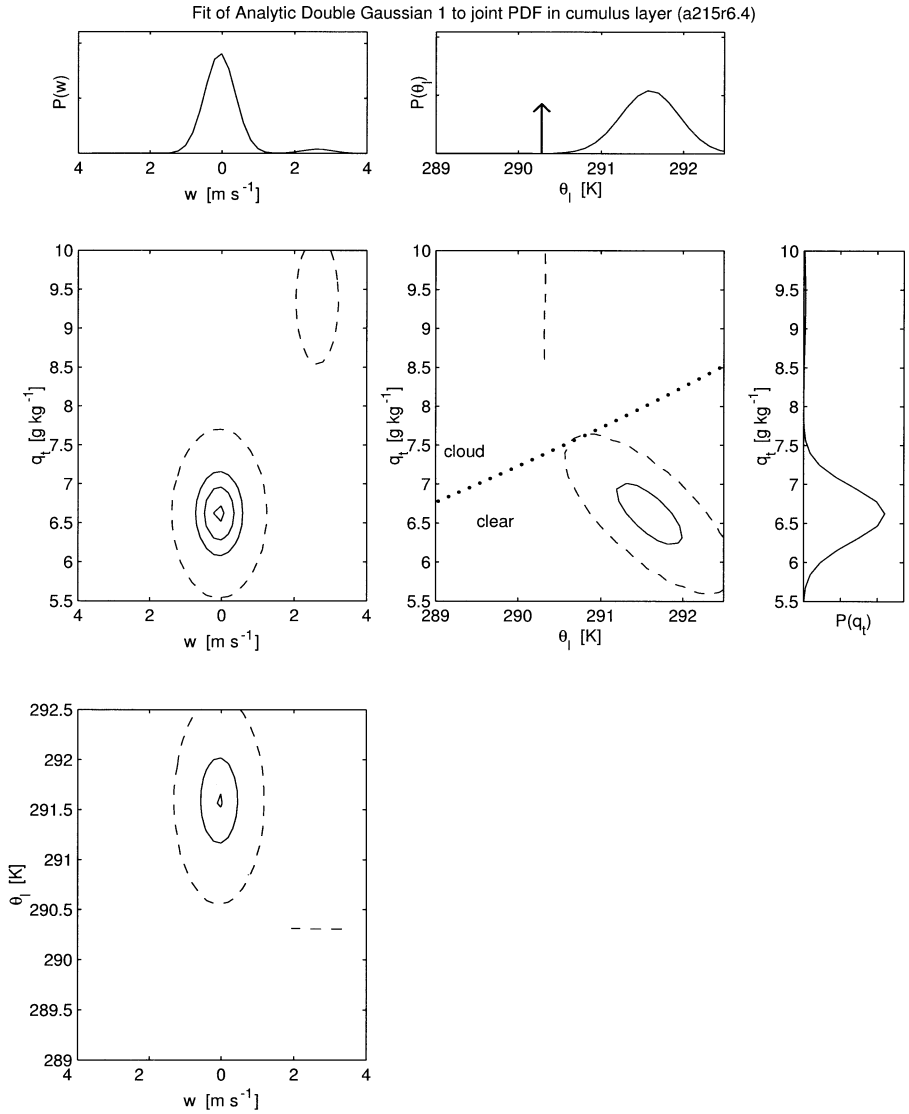


FIG. 4. The best fit of Analytic Double Gaussian 1 for the cumulus joint PDF in Fig. 3. The axes and contours are the same as in Fig. 3. The dotted line in the $q_t - \theta_l$ panel corresponds to saturation. Above and to the left of this line all points contain liquid.

Because we use observed moments to compute parameterized values of $C$, $\overline{q_l}$, and $\overline{w'q_l'}$, we do not test errors in the host model's prediction of these moments. Instead, we isolate errors arising from misrepresentations in the shape of the PDF. Therefore, the errors we plot are the best that a model could hope to achieve if it used these families of PDFs and if it predicted the moments perfectly. The question of how accurately the moments can be predicted is deferred to Golaz et al. (2002b).

We choose to investigate $C$ and $\overline{q_l}$ because they are indicators of the cloud structure, which is important for calculating radiative transfer and precipitation. We choose to compute $\overline{w'q_l'}$ because it is a key component of the buoyancy flux $(g/\theta_0)\overline{w'\theta_v'}$, which in turn generates turbulent kinetic energy. We may write

$$\overline{w'\theta_v'} = \overline{w'\theta_l'} + \frac{1 - \epsilon_0}{\epsilon_0}\theta_0\overline{w'q_t'} + \left[\frac{L_v}{c_p}\left(\frac{p_0}{p}\right)^{R_d/c_p} - \frac{1}{\epsilon_0}\theta_0\right]\overline{w'q_l'}. \quad (1)$$

Here, $\epsilon_0 = R_d/R_v$, $R_d$ is the gas constant of dry air, $R_v$ is the gas constant of water vapor, $L_v$ is the latent heat of vaporization, $c_p$ is the heat capacity of air, $\theta_0$ is a reference temperature, $p$ is the pressure field, and $p_0$ is a constant reference pressure. This expression shows

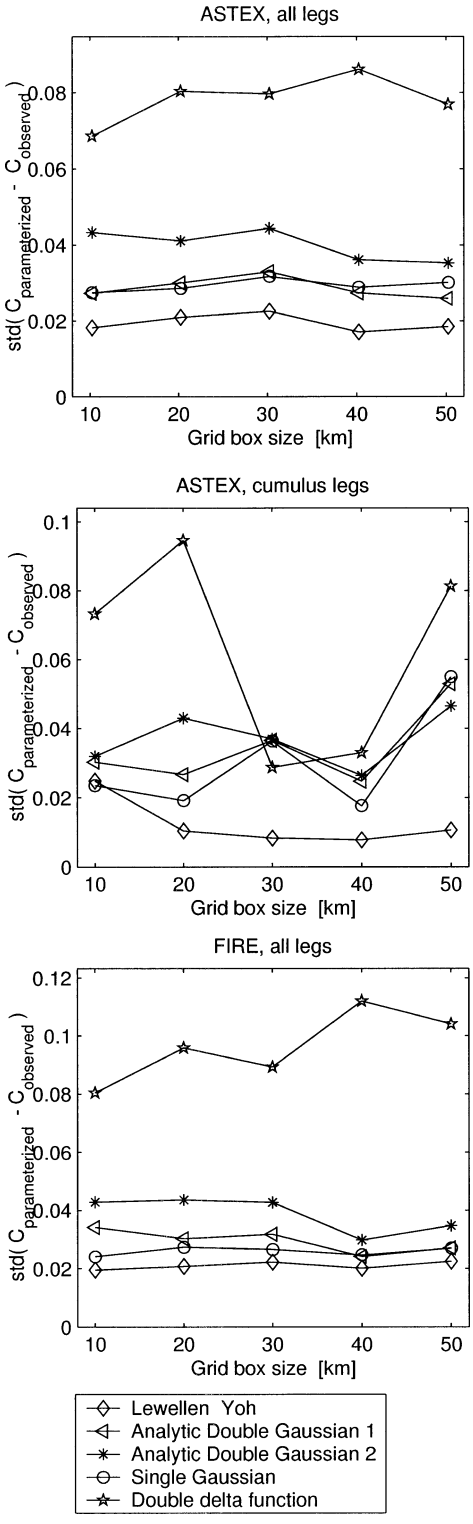


FIG. 5. Errors in parameterized cloud fraction. A higher value along the ordinate corresponds to greater error. To construct this plot, the three datasets were divided into legs of a given length. Then observed and parameterized cloud fraction, $C_{\text{observed}}$ and $C_{\text{parameterized}}$, were computed for the parameterizations listed in the legend. Next we computed the standard deviation of ($C_{\text{parameterized}} - C_{\text{observed}}$) over all the legs. Then we repeated the procedure for different leg lengths.

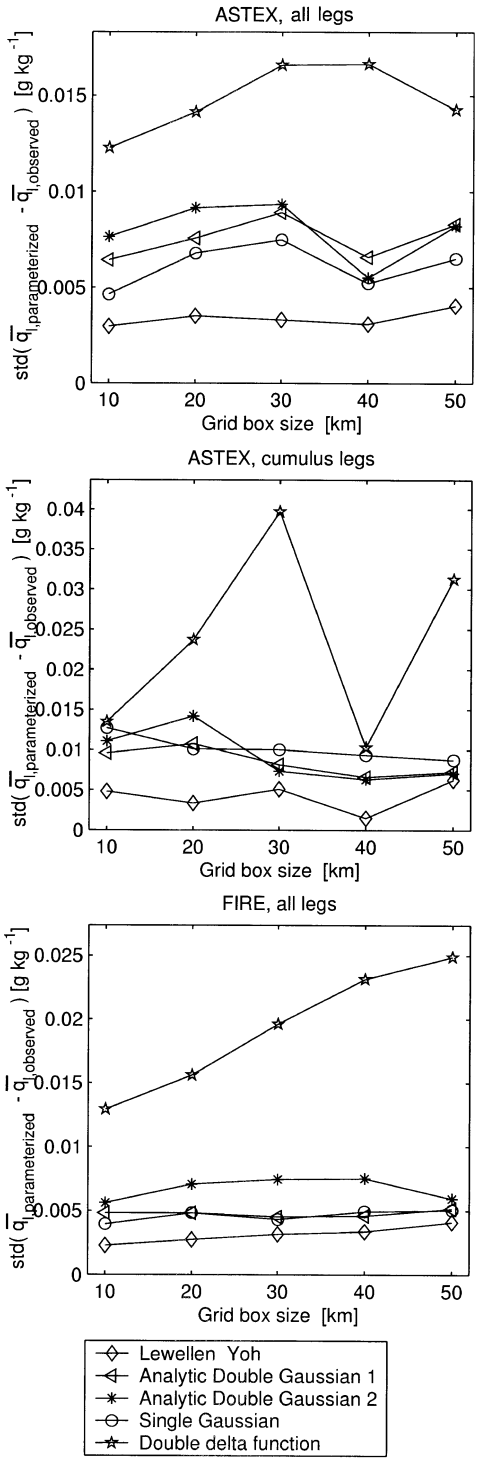


FIG. 6. Same as in Fig. 5, but here we compute errors in parameterized specific liquid water content.

that the buoyancy flux may be approximated as a sum of three fluxes, two of which are multiplied by slowly varying coefficients. If $\overline{w'\theta_l'}$ and $\overline{w'q_t'}$ are prognosed or otherwise determined by a model, then the remaining

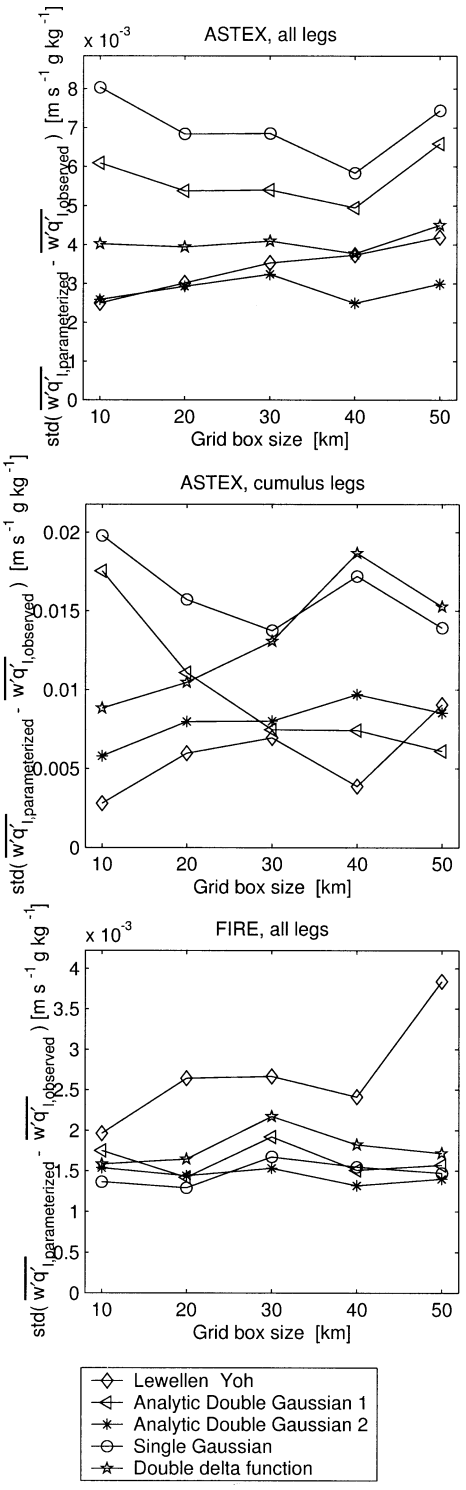


FIG. 7. Same as in Fig. 5, but here we compute errors in parameterized turbulent flux of specific liquid water content.

parameter to determine is $\overline{w'q_l'}$. This depends sensitively on the joint PDF $P(w, \theta_l, q_t)$ (Bougeault 1981). Hence $\overline{w'q_l'}$ strongly influences the dynamics of a cloudy boundary layer and also indicates the quality of a parameterized PDF.

Before performing the above calculations, we truncate our aircraft legs to lengths ranging from 10 to 50 km. Relative to a one-dimensional PDF, a three-dimensional joint PDF has a relatively large volume of parameter space for a given number of sample points. To maintain an adequate number of sampling points, we truncate legs at no less than 10-km intervals. Despite this, a one-dimensional leg samples only a small portion of the three-dimensional volume spanned by a grid box of corresponding size. This problem is difficult to avoid with typical aircraft datasets. We limit legs to 50 km because few aircraft legs in our datasets are longer than 50 km.

We do not ensemble average or remove linear trends from the aircraft legs. This is because we want to mimic what is done in the filtering approach to numerical modeling of fluid flows. In this approach, the resolved fields of a numerical model are regarded as flow fields filtered by a running spatial average of finite width, rather than an ensemble or Reynolds average (Germano 1992). Then what needs to be parameterized is a localized spatial average of moments such as liquid water flux, not an ensemble average. The probability distribution that corresponds to the filtering approach is the probability of finding $w$, $\theta_l$, and $q_t$ within a *spatially filtered* region. Therefore the PDFs that this paper discusses are, strictly speaking, more properly denoted ''filtered density functions'' (see Colucci et al. 1998).

Filtered fields do not obey Reynolds rules of averaging, such as $\overline{\overline{f}} = \overline{f}$. This turns out not to hinder derivation of the moment equations if central moments such as $\overline{w'q_l'}$ are replaced by $\overline{wq_l} - \overline{w}\,\overline{q_l}$, and similarly for other moments (Germano 1992). This avoids the extra terms that would otherwise appear because spatial filters over finite volumes do not obey Reynolds rules of averaging (Leonard 1974). To shorten notation, this paper will write $\overline{w'q_l'}$, but this is meant to be interpreted as $\overline{wq_l} - \overline{w}\,\overline{q_l}$, and analogously for all other higher-order moments. All calculations we perform on the data are consistent with the latter form. That is, a prime in our calculations denotes a deviation from a grid box average field, not a deviation from a running average field that varies continuously. One particularly simple filter is a spatial average over a rectangular parallelepiped centered on a grid box. To mimic this filter, we truncate each aircraft leg to a given length and calculate its moments by performing the appropriate spatial averaging over the truncated leg.

There are significant trends in some aircraft legs, such as the $q_t$ leg in Fig. 1. In the filtering approach, these trends should be and are included in the calculation of deviations and the filtered density function. Prognosing moments of fluid fields that contain trends is a difficult

TABLE 1. Summary of bias and errors in cloud fraction $C$. This table lists the bias and errors averaged over all points shown in Fig. 5, that is, averaged over all datasets and grid box sizes. The last row lists the standard deviation of $C$ among all legs; it is equivalent to the error incurred by a ''parameterization'' that predicts clear skies always.

| | Mean $(C_{\text{parameterized}} - C_{\text{observed}})$ | Standard deviation of $(C_{\text{parameterized}} - C_{\text{observed}})$ |
|---|---|---|
| Lewellen–Yoh | −0.0016 | 0.018 |
| Analytic Double Gaussian 1 | −0.0040 | 0.031 |
| Analytic Double Gaussian 2 | 0.000 68 | 0.039 |
| Single Gaussian | −0.0046 | 0.029 |
| Double delta function | −0.016 | 0.079 |
| | | Standard deviation of $C_{\text{observed}}$ |
| | | 0.25 |

problem in the theory of fluid mechanics. Traditional turbulence closures do not model mesoscale trends. The present paper eschews this problem and focuses instead on relating the moments to PDF parameters, given accurate moments.

The errors in $C$, $\overline{q_l}$, and $\overline{w'q_l'}$ at various leg lengths are shown in Figs. 5, 6, and 7. Consider first the errors in $C$ and $\overline{q_l}$. The results from ASTEX cumulus legs are somewhat noisy because there are only eight legs. Overall, however, the largest errors are associated with the double delta function (7 parameters). Superior to this are the single Gaussian (9 parameters) and analytic double Gaussian (10 parameters) parameterizations, which perform comparably to each other. The best predictions are provided by the Lewellen–Yoh scheme (12 parameters). As expected, the parameterizations with more fitting parameters produce better fits. Now consider errors in $\overline{w'q_l'}$, shown in Fig. 7. Overall, this plot reveals no obviously superior scheme. The ASTEX cumulus legs hint, however, that the double Gaussian schemes may perform better for long legs in cumulus layers. The double Gaussian schemes have the capability of producing long tails when skewness is large. This is an advantage when parameterizing cumulus PDFs. However, our experience was that these tails can occasionally lead to poor predictions of $C$ and $q_l$. Such outliers dominate the errors in the double Gaussian schemes. The double delta function and single Gaussian PDFs are more subdued, but they tend to underpredict $\overline{w'q_l'}$. The (nonsystematic) biases and errors for all the parameterizations and all three datasets are summarized in Tables 1, 2, and 3. For the most part, parameterizations with larger errors also tend to have larger biases.

Figure 8 displays the errors in ASTEX and FIRE legs that have midrange cloud fractions, $0.2 \leq C \leq 0.8$. This restriction leaves fifteen 50-km ASTEX legs and twelve 50-km FIRE legs. The errors are several times larger than when all cloud fractions are permitted. This is because fits to clear or overcast legs have small errors for all PDF parameterizations that we tested. However, the performance of the parameterizations relative to each other is similar for midrange $C$ and all $C$.

A more detailed picture of the errors in long (50 km) ASTEX cumulus legs is shown in Fig. 9. It turns out that for each of these legs, the double delta function diagnoses zero $C$, $\overline{q_l}$, and $\overline{w'q_l'}$. Why is this? Consider Fig. 3, which fits a double delta function to a particular observed PDF. Each delta function, denoted by a star, lies within unsaturated air (although one lies close to the saturation line). Recall that the locations of the delta functions are chosen such that the observed and parameterized fluxes $\overline{w'q_t'}$ are equal, and similarly for $\overline{w'\theta_l'}$. To prevent these fluxes from being overestimated, the delta functions must be rather closely spaced. Then neither delta function lies within the cloudy region. In Fig. 9, the single Gaussian underestimates $C$ modestly, $\overline{q_l}$ moderately, and $\overline{w'q_l'}$ severely. Similar results were obtained by Bougeault (1981). This is related to the fact that the single Gaussian is unskewed, whereas cumulus layers are positively skewed. The percentage underestimates are more severe for cases with less cloud. The double Gaussian parameterizations produce scatter, but

TABLE 2. Summary of bias and errors in specific liquid water content $\overline{q_l}$ averaged over all points in Fig. 6, that is, averaged over all datasets and grid box sizes. The last row lists the standard deviation of $\overline{q_l}$ among all legs; it is equivalent to the error incurred by a ''parameterization'' that predicts clear skies always.

| | Mean $(q_{l,\text{parameterized}} - q_{l,\text{observed}})$ $(10^{-3}$ g kg$^{-1})$ | Standard deviation of $(q_{l,\text{parameterized}} - q_{l,\text{observed}})$ $(10^{-3}$ g kg$^{-1})$ |
|---|---|---|
| Lewellen–Yoh | −0.34 | 3.6 |
| Analytic Double Gaussian 1 | −0.86 | 6.9 |
| Analytic Double Gaussian 2 | −0.58 | 8.0 |
| Single Gaussian | −2.5 | 7.0 |
| Double delta function | −7.3 | 19 |
| | | Standard deviation of $q_{l,\text{observed}}$ $(10^{-3}$ g kg$^{-1})$ |
| | | 98 |

TABLE 3. Summary of bias and errors in liquid water flux $\overline{w'q_l'}$. This table lists the bias and errors averaged over all points shown in Fig. 7, that is, averaged over all datasets and grid box sizes. The last row lists the standard deviation of $\overline{w'q_l'}$ among all legs; it is equivalent to the error incurred by a ''parameterization'' that predicts clear skies always.

| | Mean ($\overline{w'q'_{l,\text{parameterized}}} - \overline{w'q'_{l,\text{observed}}}$) ($10^{-3}$ m s$^{-1}$ g kg$^{-1}$) | Standard deviation of ($\overline{w'q'_{l,\text{parameterized}}} - \overline{w'q'_{l,\text{observed}}}$) ($10^{-3}$ m s$^{-1}$ g kg$^{-1}$) |
|---|---|---|
| Lewellen–Yoh | −0.43 | 3.9 |
| Analytic Double Gaussian 1 | 1.4 | 5.7 |
| Analytic Double Gaussian 2 | −1.6 | 4.1 |
| Single Gaussian | −3.3 | 8.2 |
| Double delta function | −2.8 | 6.4 |
| | | Standard deviation of $\overline{w'q'_{l,\text{observed}}}$ ($10^{-3}$ m s$^{-1}$ g kg$^{-1}$) |
| | | 12.7 |

do not have general underestimates, except in the diagnosis of $\overline{w'q_l'}$ by Analytic Double Gaussian 2 (asterisks).

To investigate cumulus PDFs more closely, Figs. 10 and 11 present results from certain time periods of LESs that were set up according to the GCSS specifications. Figure 10 displays the ARM continental cumulus case, and Fig. 11 displays the BOMEX trade wind cumulus case. Values of $C$, $\overline{q_l}$, and $\overline{w'q_l'}$ for the various parameterizations are calculated as we did for the aircraft legs, but as input we use moments computed from the LES fields. For the most part, the results corroborate those of our observational data. As before, the double delta function PDF produces no cloud for these time periods. As before, the single Gaussian underestimates $C$ moderately, $\overline{q_l}$ more severely, and $\overline{w'q_l'}$ even more

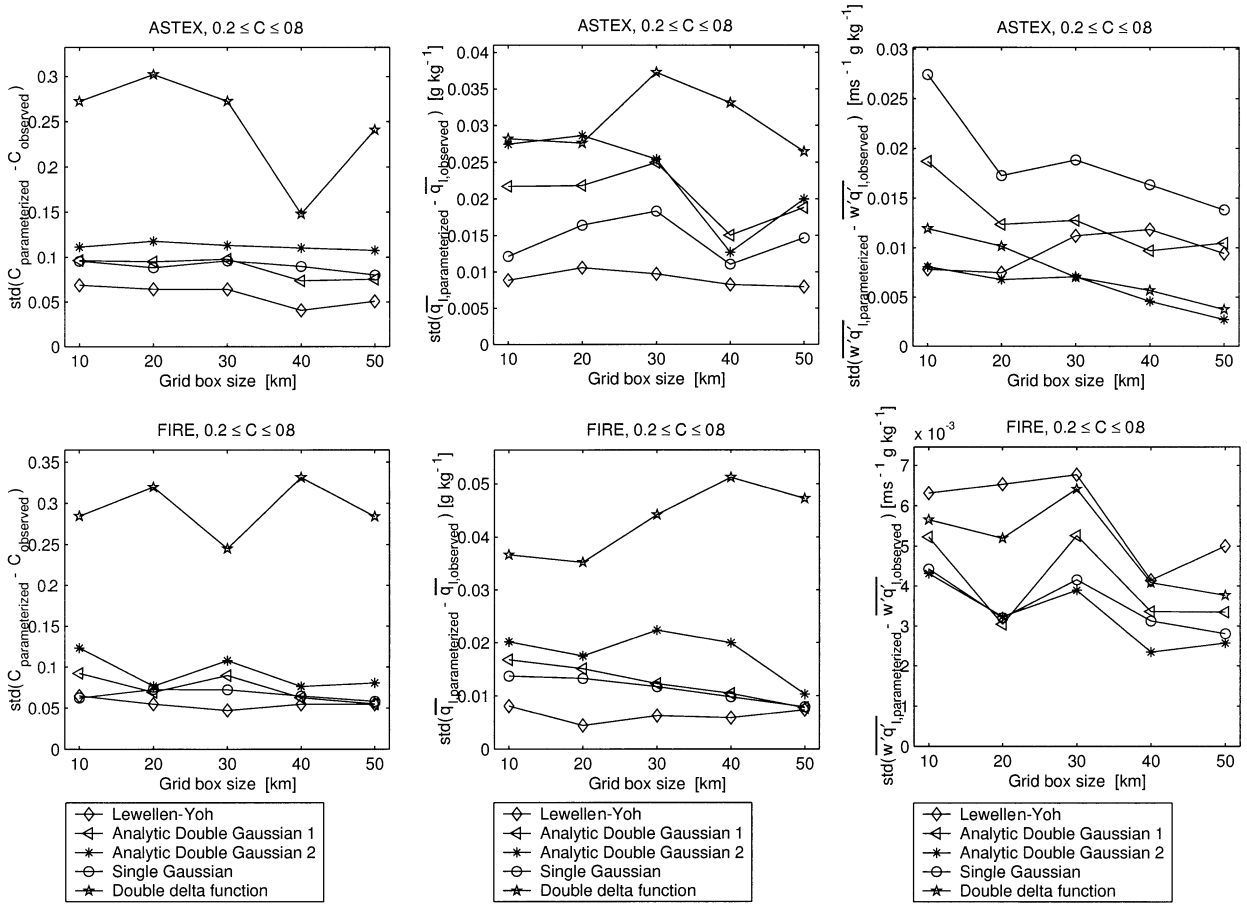


FIG. 8. Same as in Fig. 5, but here we include only those legs from ASTEX and FIRE that have cloud fractions $C$ between 0.2 and 0.8. This provides an estimate of error for legs that are neither clear nor overcast. The figure also displays errors in parameterized specific liquid water content and its flux.

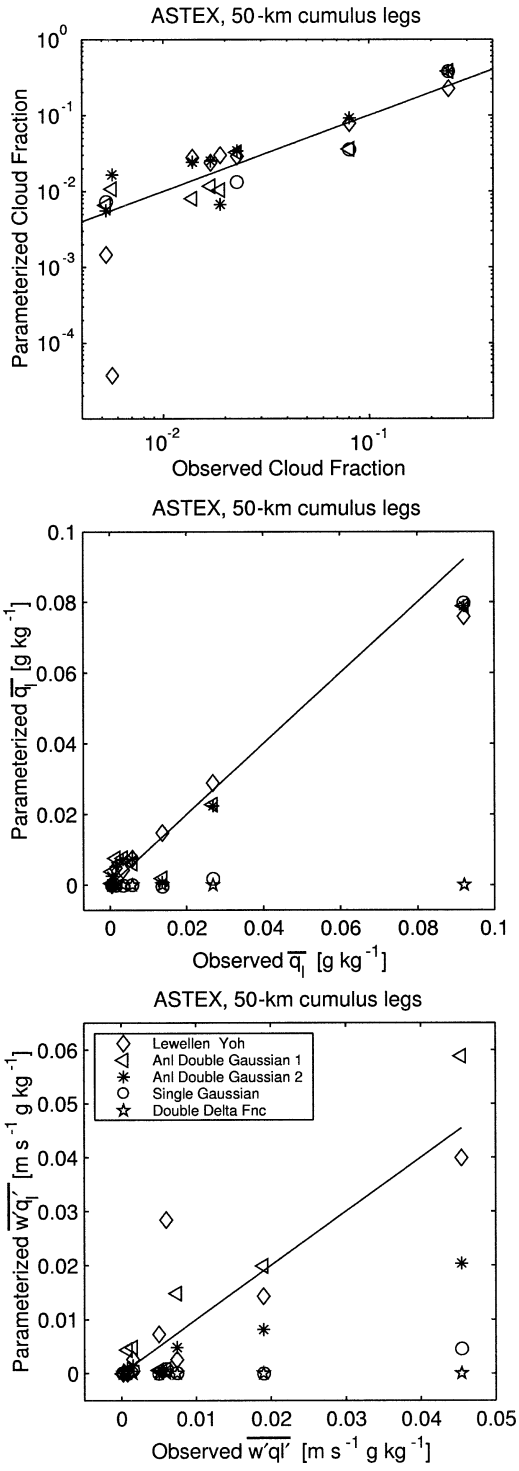


FIG. 9. A scatterplot of parameterized vs observed cloud fraction ($C$), average specific liquid water content ($\overline{q_l}$), and turbulent flux of specific liquid water content ($\overline{w'q_l'}$). The various parameterizations tested are listed in the legend. Each point corresponds to one long (50 km) ASTEX cumulus leg. A perfect fit would correspond to all points lying along the solid line. The Gaussian PDF (circles) leads to underpredictions. The double delta function (stars) diagnoses no cloud or liquid for these cumulus legs, and therefore does not appear in the top panel.

severely. The double Gaussian schemes perform significantly better, but they still underestimate liquid water flux by a factor of two to three. Although the percentage errors in fitting our cumulus simulations are large, the absolute errors are comparable with the errors in fitting the aircraft data. For instance, the error in fitting $\overline{w'q_l'}$ diagnosed by the double delta PDF over the cloud layer is $6.3 \times 10^{-3} \pm 4.3 \times 10^{-3}$ m s$^{-1}$ g kg$^{-1}$ for BOMEX and $20 \times 10^{-3} \pm 17 \times 10^{-3}$ m s$^{-1}$ g kg$^{-1}$ for ARM. This is comparable to the errors shown in Table 3 and Figs. 7 and 8. Whether either the Gaussian or double delta function performs well for cumulus layers with higher cloudiness is an open question that we have not explored.

In our cumulus simulations the cloud fraction maximizes near cloud base and then decreases with increasing altitude. This is because only a small percentage of clouds reach the top of the cloud layer. The single Gaussian and double Gaussian PDFs are all able to reproduce the shape of the cloud fraction profiles. Both analytic double Gaussian schemes underestimate the altitudes of cloud top and cloud base. The Lewellen–Yoh scheme performs better in this respect, probably because it assumes that the skewnesses of $q_t$ and $\theta_l$ are prognosed, not diagnosed.

It is important that a boundary layer parameterization remain valid for a range of grid box sizes, especially if the host model uses nested gridding. We have tested PDF families for horizontal domains ranging from 6.4 km $\times$ 6.4 km for the BOMEX LES to 50 km for the aircraft legs. Whether or not the PDF families will perform well for larger or smaller grid box sizes remains an open question. One would expect the scalar variances to increase with increasing length scale, but all PDF families we tested can accommodate such a change with scale.

The fits produced by the double delta function PDF merit further discussion. The fact that the double delta function diagnoses no cloud in the cumulus cases might lead one to suppose that the double delta function is a more appropriate PDF for stratocumulus cases, where at least it diagnoses some cloud. This may seem counterintuitive, because one might have expected that the double delta function would represent the updraft–downdraft structure better in cumulus layers than in stratocumulus layers. However, the delta function has difficulties with the stratocumulus layers as well, producing the least satisfactory estimates of $C$ and $\overline{q_l}$ among all the schemes tested. The problems in cumulus and stratocumulus layers are both significant, but they are distinct from each other.

In cumulus layers with low cloud fraction, the essential difficulty is that often the observed distribution contains cloud only within the end of a long tail, but a double delta function contains no tail and therefore often produces no cloud. We must note a caveat: although in our examples the double delta function does not produce

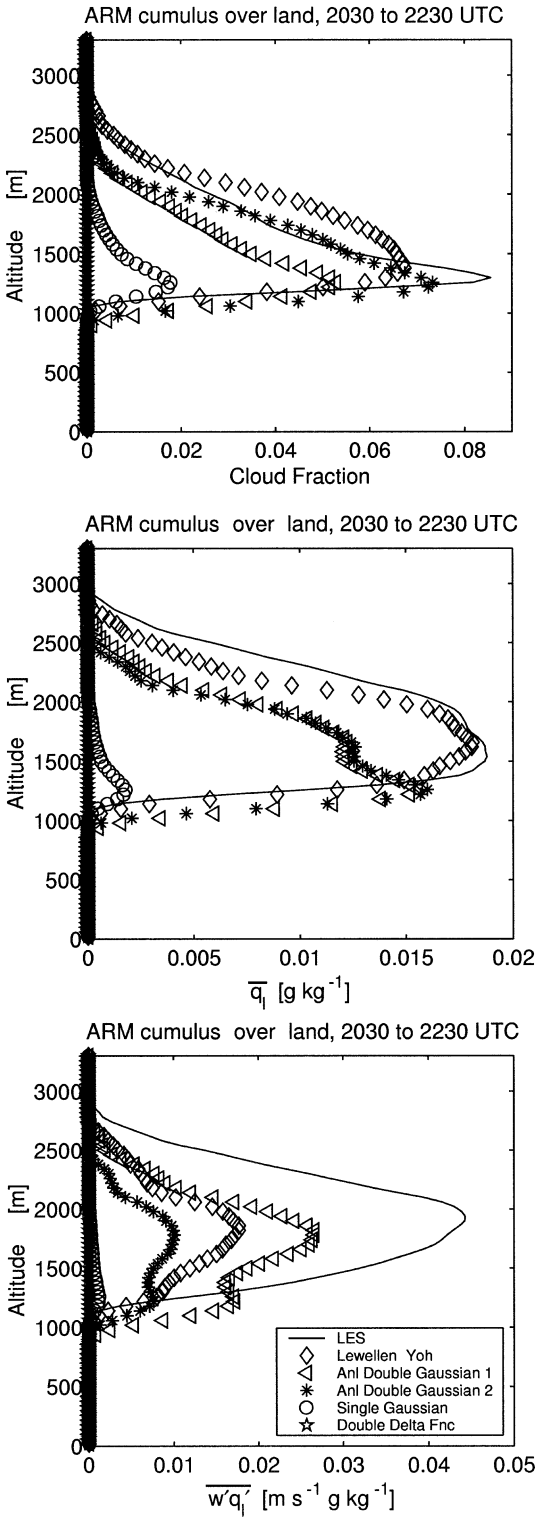


FIG. 10. Diagnosed profiles of cloud fraction ($C$), average specific liquid water content ($\overline{q_l}$), and turbulent flux of liquid water ($\overline{w'q_l'}$), along with corresponding profiles computed by LES. Plotted is the 10th and 11th hours of the GCSS ARM simulation involving cumulus over land. The Gaussian PDF (circles) leads to underpredictions. The double delta function PDF (stars) diagnoses no clouds for this time period.

cloud, it nearly does so in some of our cases. For example, one of the delta functions in Fig. 3 is near the saturation line. To move the delta function within the saturated region, one could choose an algorithm other than Randall et al. (1992) to position the delta functions, but then the fluxes $\overline{w'q_t'}$ and $\overline{w'\theta_l'}$ would no longer be satisfied exactly. However, the error induced in the fluxes may be small if a delta function needs to be moved only slightly.

In stratocumulus layers, the essential difficulty is that many of the PDFs are spread out and not highly correlated. For example, consider the stratocumulus layer depicted in Fig. 1. The correlation coefficient between $w$ and $q_t$ is 0.20, that between $w$ and $\theta_l$ is $-0.14$, and that between $q_t$ and $\theta_l$ is $-0.59$. One can illustrate the difficulty of fitting such PDFs with a double delta function by considering the limiting case of a single Gaussian that is uncorrelated in $w$ and $q_t$, as shown in Fig. 12. Suppose one chooses to match the flux ($\overline{w'q_t'} = 0$) and the variance in $w$, as does Randall et al. (1992). Then for an uncorrelated PDF, the delta functions lie along the $w$ axis (pluses). Unfortunately, then the fitted double delta PDF has $\overline{q_t'^2} = 0$, contrary to the PDF we desire to fit. The double delta fit to the observed stratocumulus PDF in Fig. 1 shows a hint of this behavior. Similarly, one could match the flux and the variance in $q_t$ (triangles), but then unfortunately the variance in $w$ would vanish. Finally, one could match the variances in both directions by placing the delta functions along the diagonal (dots), but then the flux would be grossly overestimated. Regardless of where the delta functions are positioned, either the flux or variances or both must be misrepresented. The problem of matching both flux and variances using the double delta function PDF has been studied in more detail by Wang and Stevens (2000). Our data show additionally that the double delta PDF misrepresents liquid water, and we conjecture that source of the problem in stratocumulus layers is the dissimilarity between double delta PDFs and stratocumulus PDFs, which are spread out and comparatively uncorrelated.

The double delta PDF is of practical importance because several authors have shown that the assumption that a double delta function can adequately represent observed PDFs is the essential assumption behind simple updraft–downdraft mass-flux schemes with no subplume variability (Randall et al. 1992; de Roode et al. 2000; Lappen and Randall 2001a). Therefore, deficiencies of the double delta function PDF are also deficiencies of the simplest mass-flux schemes. Our fits to data evaluate only one variant of the double delta PDF, the Randall et al. (1992) algorithm, in which the fluxes are used to fix the locations of the delta functions. Other possible techniques to close the mass-flux approach, such as using variances to locate the delta functions, may produce better results. However, it seems unlikely to us that a double delta PDF can simultaneously di-

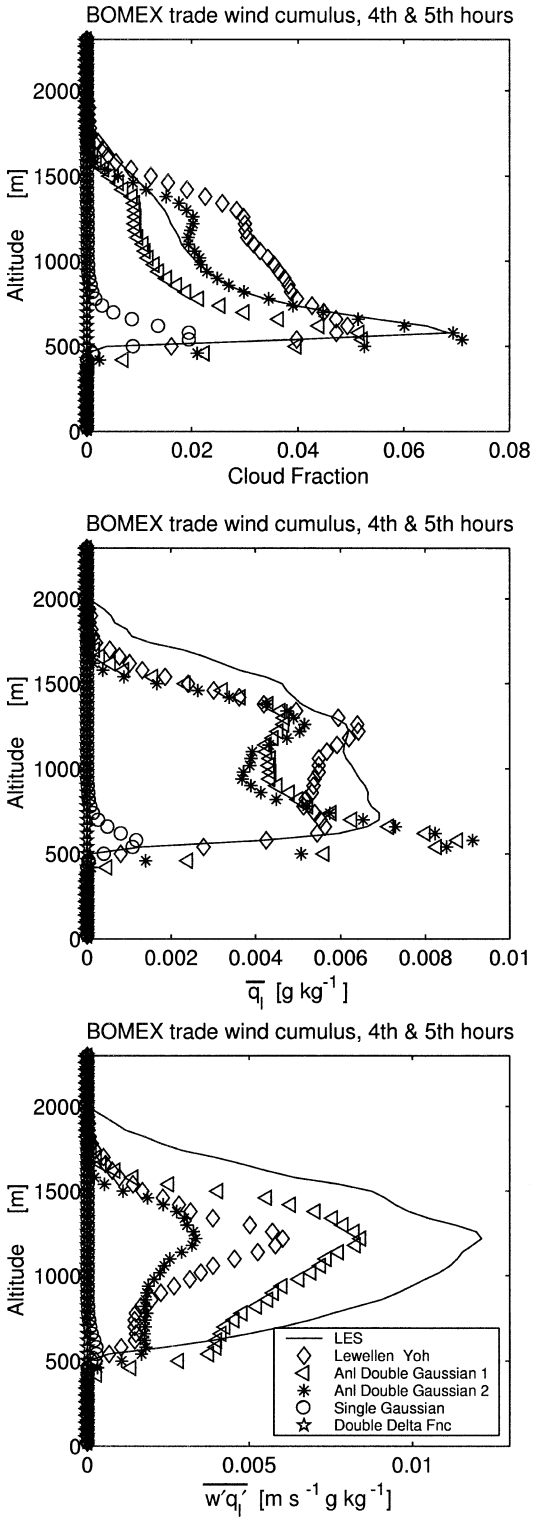


FIG. 11. Diagnosed profiles of cloud fraction ($C$), average specific liquid water content ($\overline{q_l}$), and turbulent flux of liquid water ($\overline{w'q_l'}$), along with corresponding profiles computed by LES. Plotted is the 4th and 5th hours of the GCSS BOMEX simulation involving trade wind cumulus. The Gaussian PDF (circles) leads to underpredictions. The double delta function PDF (stars) diagnoses no clouds for this time period.

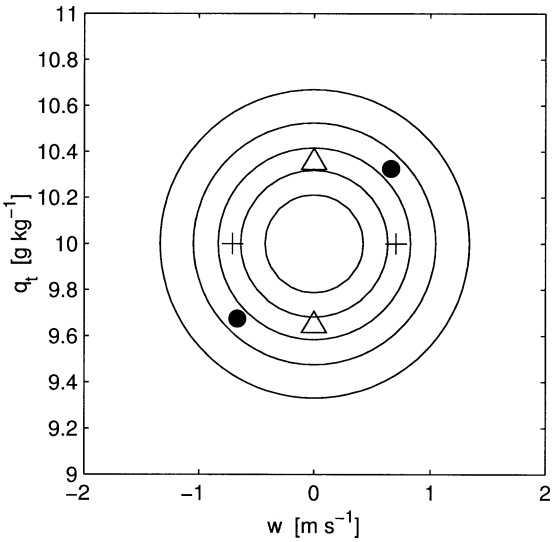


FIG. 12. An idealized joint PDF of vertical velocity $w$ and total specific water content $q_t$ (contours) overplotted with three possible double delta PDF fits. The idealized PDF is a single Gaussian with no correlation. The three fitted PDFs are double delta functions that match $\overline{w'q_t'}$ and $\overline{w'^2}$ (pluses); $\overline{w'q_t'}$ and $\overline{q_t'^2}$ (triangles); and $\overline{w'^2}$ and $\overline{q_t'^2}$ (dots). No double delta PDF can simultaneously match the flux and both variances.

agnose accurate fluxes, variances, and cloud properties, because boundary layer PDFs do not closely resemble double delta functions. We cannot make conclusions about the usefulness of mass flux schemes in deep convecting layers because our data include only boundary layers.

To overcome the deficiencies of mass-flux models, prior authors have proposed several remedies, listed below.

1) Some authors have suggested coupling a mass-flux scheme to a cloud PDF that is not a double delta function, for example, a triangular-shaped PDF (Lock et al. 2000). Such a scheme should then ensure that it does not use two separate and mutually inconsistent PDFs (the double delta function and the triangular-shaped PDF) in different calculations, rather than using a single internally consistent joint PDF (consisting of a delta function *plus* a triangular-shaped function) for all calculations.
2) Other authors have explored the possibility of choosing the updraft and downdraft properties via a sophisticated sampling strategy (Wang and Stevens 2000). For instance, one could define the updraft plume to consist of those parcels that have $w' > 0$, $q_l > 0$, and $\theta_v' > 0$. However, Wang and Stevens (2000) were unsatisfied with the sampling strategies they discussed. Furthermore, it is not obvious how such a redefinition of sampling strategy should be implemented numerically.

3) Several authors have added subplume variability to mass-flux schemes in an attempt to improve the fluxes. For instance, some have increased the fluxes by multiplying them by a constant prefactor whose value could be derived, for example, under the assumption that the underlying PDF is a single joint Gaussian (Randall et al. 1992; Petersen et al. 1999). However, the present paper has demonstrated that cloud fields can be misrepresented by double delta PDFs even if the fluxes are perfect.

We do not evaluate the above remedies because the main topic of this paper is PDFs, not mass-flux schemes. Also, we believe that if one's goal is to parameterize cloudy boundary layers, it is more natural to attempt to model the PDF as directly and faithfully as possible, rather than modifying a mass flux scheme. A possible procedure to directly model PDFs, the assumed PDF method, was outlined in the introduction and is discussed in more detail in Golaz et al. (2002a,b).

## 6. Conclusions

This paper has used aircraft observations and large eddy simulation (LES) of boundary layer clouds to test five families of joint probability density functions (PDFs).

The double delta function satisfactorily diagnoses liquid water flux $\overline{w'q_l'}$ for the ASTEX and FIRE aircraft legs overall. However, compared to the other four PDFs, the double delta function tends to produce the least satisfactory fits of cloud fraction and specific liquid water content. Additionally, in the cumulus data segments we examined, which had low cloud fraction, if one places the delta functions such that the fluxes are matched exactly, then no clouds are diagnosed. (We have not, however, investigated the degree to which the fluxes would have to be compromised in order to produce reasonable cloud values. Other variants of mass-flux schemes may perform better.) The reason that no clouds are diagnosed is that cumulus clouds for our low-cloudiness cases tend to reside on the tail of the distribution, and the double delta function misrepresents the tails. Because there is a close link between the double delta function and mass-flux schemes, this raises doubts about a main assumption underlying the simplest mass-flux schemes. This has inspired Golaz et al. (2002a,b) to develop an alternative to mass-flux schemes that is based on a more realistic PDF. Although it remains an open question whether the higher-order moments, particularly the third-order moments, can be predicted with sufficient accuracy to warrant the use of the more sophisticated PDF families, Golaz et al. (2002a,b) were able to prognose $\overline{w'^3}$ accurately enough to develop a satisfactory PDF-based model for cloudy boundary layers. The PDF family they used was the Analytic Double Gaussian 1.

In partly cloudy layers, the liquid water flux, $\overline{w'q_l'}$, and hence the buoyancy flux $(g/\theta_0)\overline{w'\theta_v'}$, depend on the joint PDF $P(w, \theta_l, q_t)$ of $w$, $\theta_l$, and $q_t$. One finds different expressions for $\overline{w'q_l'}$ depending on whether one assumes that the PDF is a single Gaussian (Mellor 1977) or a double delta function (Randall 1987). One may ask, "Which PDF produces the better estimate of $\overline{w'q_l'}$?" The single Gaussian appears to work slightly better, but both PDFs severely underestimate $\overline{w'q_l'}$ in low-cloudiness cumulus layers. However, we believe the question poses a false dichotomy. Since $\overline{w'q_l'}$ depends on the PDF, there can be as many approximations to $\overline{w'q_l'}$ as there are PDFs. For instance, the formula (A13) for $\overline{w'q_l'}$ is based on the double Gaussian PDF. Therefore, we advocate striving to find the most representative and simplest family of PDFs possible.

In our survey of five PDFs, we found that PDFs that depend on more parameters tend to produce better fits, as expected. The Lewellen–Yoh family of PDFs, which is based on a double Gaussian form, performs best. If an analytic scheme is desired, then the analytic double Gaussians presented here perform satisfactorily. In particular, they fit low-cloudiness cumulus layers somewhat better than the single Gaussian form.

The merits of the various PDF parameterizations must be weighed against their computational costs. Implementing any of the PDF parameterizations we tested would be expensive compared to using a first-order, down-gradient diffusion scheme. It is difficult to accurately compare the costs of the various schemes without implementing each of them. However, assuming that the host model prognoses the mean quantities, the number of additional moments that must be prognosed in the PDF parameterizations ranges from nine (Lewellen–Yoh) to four (double delta). If all moment equations were equally expensive—which is not strictly true, because the higher-moment equations contain extra terms—then the Lewellen–Yoh parameterization would be roughly 9/4 as expensive as a double delta function parameterization.

*Acknowledgments.* We are grateful for helpful comments from Cara-Lyn Lappen, David A. Randall, Adrian M. Tompkins, Robert Wood, Paul R. Field, W. S. Lewellen, and two anonymous reviewers. We also acknowledge the RAF aircrew and MRF staff involved in the ASTEX and FIRE field campaigns. W. R. Cotton and J.-C. Golaz were supported by the National Science Foundation under Grants ATM-9529321 and ATM-9904218. V. E. Larson was supported by the National Oceanic and Atmospheric Administration, Contract NA67RJ0152. For further information about the research in this paper, see http://www.uwm.edu/~vlarson.

## APPENDIX

### Fitting PDFs to Data

This appendix describes how we solve for the parameters that specify the five PDFs, given moments cal-

culated from the aircraft legs or LES output. We also list formulas for cloud fraction ($C$), average specific liquid water content ($\overline{q_l}$), and turbulent flux of liquid water ($\overline{w'q_l'}$).

*a. Double delta function (7 parameters; input moments $\overline{w}$, $\overline{w'^2}$, $\overline{w'^3}$, $\overline{\theta_l}$, $\overline{w'\theta_l'}$, $\overline{q_t}$, $\overline{w'q_t'}$)*

This function is given by

$$P_{\mathrm{dd}} = a\delta(w - w_1)\delta(\theta_l - \theta_{l1})\delta(q_t - q_{t1}) + (1 - a)\delta(w - w_2)\delta(\theta_l - \theta_{l2})\delta(q_t - q_{t2}).$$

This formula lists all the PDF parameters. Namely, $0 \le a \le 1$ is the relative amplitude of the first delta function, and ($w_1$, $\theta_{l1}$, $q_{t1}$) and ($w_2$, $\theta_{l2}$, $q_{t2}$) are the positions of the first and second delta functions, respectively, in ($w$, $\theta_l$, $q_t$) space. To determine these parameters for a particular aircraft leg, we follow the procedure of Randall et al. (1992) and Lappen and Randall (2001a). The strategy is to relate these known moments to the PDF parameters using the definitions of the moments; for example,

$$\overline{w'^2} = \int_{-\infty}^{\infty} dw \int_{-\infty}^{\infty} d\theta_l \int_{-\infty}^{\infty} dq_t P_{\mathrm{dd}}(w, \theta_l, q_t)(w - \overline{w})^2 = a(w_1 - \overline{w})^2 + (1 - a)(w_2 - \overline{w})^2.$$

We first compute $a$, $w_1$, and $w_2$ from $\overline{w}$, $\overline{w'^2}$, and $\overline{w'^3}$ via the following formulas (Randall et al. 1992; Lappen and Randall 2001a):

$$a = \frac{1}{2}\left(1 - \frac{\mathrm{Sk}_w}{\sqrt{4 + \mathrm{Sk}_w^2}}\right),$$

$$w_1 = \overline{w} + \sqrt{\overline{w'^2}}\sqrt{\frac{1 - a}{a}}, \quad \text{and}$$

$$w_2 = \overline{w} - \sqrt{\overline{w'^2}}\sqrt{\frac{a}{1 - a}}.$$

Here, $\mathrm{Sk}_w = \overline{w'^3}/(\overline{w'^2}^{3/2})$ is the skewness of $w$. Next, we compute $\theta_{l1}$ and $\theta_{l2}$ from $\overline{\theta_l}$ and $\overline{w'\theta_l'}$:

$$\theta_{l1} = \overline{\theta_l} - \frac{\overline{w'\theta_l'}}{w_2 - \overline{w}}, \qquad \theta_{l2} = \overline{\theta_l} - \frac{\overline{w'\theta_l'}}{w_1 - \overline{w}}.$$

We use analogous formulas to compute $q_{t1}$ and $q_{t2}$ from $\overline{q_t}$ and $\overline{w'q_t'}$. The procedure guarantees that the delta function PDF satisfies the fluxes $\overline{w'\theta_l'}$ and $\overline{w'q_t'}$, but does not guarantee accurate diagnosis of the scalar variances, $\overline{\theta_l'^2}$ and $\overline{q_t'^2}$. Despite the simplicity of the double delta function, use of this PDF in a numerical model still requires the prediction of $\overline{w'^3}$, which is nontrivial.

Given the PDF parameters, we need to compute cloud fraction $C$, mean specific liquid water content $\overline{q_l}$, and mean liquid water flux $\overline{w'q_l'}$. To do so, we first need to compute a quantity $s$, which equals $q_l$ when $s > 0$ but can also be negative and is conserved under condensation. We have (Lewellen and Yoh 1993)

$$s = q_t - q_s(T_l, p)\frac{(1 + \beta q_t)}{[1 + \beta q_s(T_l, p)]}, \tag{A1}$$

where

$$q_s(T_l, p) = \frac{R_d}{R_v}\frac{e_s(T_l)}{p - [1 - (R_d/R_v)]e_s(T_l)}, \tag{A2}$$

$$\beta = \beta(T_l) = \frac{R_d}{R_v}\left(\frac{L_v}{R_d T_l}\right)\left(\frac{L_v}{c_p T_l}\right), \quad \text{and} \tag{A3}$$

$$T_l \equiv T - \frac{L_v}{c_p}q_l. \tag{A4}$$

Here, $q_s$ is the saturation specific humidity, $e_s$ is the saturation vapor pressure over liquid, $T_l$ is the liquid water temperature, $T$ is temperature, $p$ is pressure, $L_v$ is the latent heat of vaporization, $c_p$ is the specific heat at constant pressure, and $R_d$ and $R_v$ are the gas constants for dry air and water vapor.

We let $s_1$ be the value of $s$ evaluated at $\theta_{l1}$ and $q_{t1}$, and $s_2$ be the value of $s$ evaluated at $\theta_{l2}$ and $q_{t2}$. Then

$$C = aH(s_1) + (1 - a)H(s_2),$$

$$\overline{q_l} = as_1H(s_1) + (1 - a)s_2H(s_2), \quad \text{and}$$

$$\overline{w'q_l'} = \overline{w'q_l} = a(w_1 - \overline{w})s_1H(s_1) + (1 - a)(w_2 - \overline{w})s_2H(s_2).$$

Here, $H$ denotes the Heaviside step function.

*b. Single Gaussian (9 parameters; input moments $\overline{w}$, $\overline{w'^2}$, $\overline{\theta_l}$, $\overline{w'\theta_l'}$, $\overline{q_t}$, $\overline{w'q_t'}$, $\overline{q_t'\theta_l'}$, $\overline{\theta_l'^2}$, $\overline{q_t'^2}$)*

This function is given by (Stuart and Ord 1994, p. 511):

$$P_{\mathrm{sg}}(w, \theta_l, q_t) = (2\pi)^{-3/2}\sqrt{|\mathbf{A}|}\,\exp\left[-\frac{1}{2}(x_i - \overline{x_i})A_{ij}(x_j - \overline{x_j})\right]. \tag{A5}$$

The indices $i$ and $j$ range from 1 to 3, and $x_1 = w$, $x_2 = \theta_l$, and $x_3 = q_t$. $\mathbf{A}$ is the inverse of the covariance matrix $\mathbf{A} = \mathbf{C}^{-1}$, where

$$C_{ij} = \overline{(x_i - \overline{x_i})(x_j - \overline{x_j})}.$$

For a single Gaussian, we are spared the inconvenience of translating from moments to PDF parameters, since the two are the same.

For a Gaussian PDF, cloud fraction $C$ and mean specific liquid water content $\overline{q_l}$ are calculated by linearizing variability in $\theta_l$ and $q_t$ (see Lewellen and Yoh 1993; also Sommeria and Deardorff 1977; Mellor 1977; Bougeault 1982; Chen 1991; Larson et al. 2001b):

$$C = \frac{1}{2}\left[1 + \text{erf}\left(\frac{\overline{s}}{\sqrt{2}\sigma_s}\right)\right], \quad \text{and} \tag{A6}$$

$$\overline{q_l} = \overline{s}C + \frac{\sigma_s}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\overline{s}}{\sigma_s}\right)^2\right]. \tag{A7}$$

Here, erf is the error function, and $\sigma_s$ is the standard deviation of $s$:

$$\sigma_s^2 = c_{\theta_l}^2\sigma_{\theta_l}^2 + c_{q_t}^2\sigma_{q_t}^2 - 2c_{\theta_l}\sigma_{\theta_l}c_{q_t}\sigma_{q_t}r_{q_t\theta_l},$$

where we have defined

$$c_{q_t} = \frac{1}{1 + \beta(\overline{T_l})q_s(\overline{T_l}, p)}, \quad \text{and}$$

$$c_{\theta_l} = \frac{1 + \beta(\overline{T_l})\overline{q_t}}{[1 + \beta(\overline{T_l})q_s(\overline{T_l}, p)]^2}\frac{c_p}{L_v}\beta(\overline{T_l})q_s(\overline{T_l}, p)\left(\frac{p}{p_0}\right)^{R_d/c_p},$$

where $p_0$ is a reference pressure. Here $\sigma_{\theta_l}$ and $\sigma_{q_t}$ are the respective standard deviations of $\theta_l$ and $q_t$. Also, $r_{q_t\theta_l}$ is the correlation coefficient between $q_t$ and $\theta_l$.

The flux of liquid water is given by

$$\overline{w'q_l'} = C\overline{w's'}, \tag{A8}$$

where

$$\overline{w's'} = c_{q_t}\sigma_w\sigma_{q_t}r_{wq_t} - c_{\theta_l}\sigma_w\sigma_{\theta_l}r_{w\theta_l}.$$

Here, $\sigma_w$ is the standard deviation of $w$.

To close the prognostic moment equations of Golaz et al. (2002a), which uses the Analytic Double Gaussian 1, the three formulas below are essential building blocks. We list them here for reference:

$$\overline{w'^2q_l'} = \frac{1}{\sqrt{2\pi}\sigma_s}(\overline{w's'})^2 \exp\left[-\frac{1}{2}\left(\frac{\overline{s}}{\sigma_s}\right)^2\right],$$

$$\overline{\theta_l'q_l'} = \sigma_{\theta_l}(c_{q_t}\sigma_{q_t}r_{q_t\theta_l} - c_{\theta_l}\sigma_{\theta_l})C, \quad \text{and}$$

$$\overline{q_l'q_l'} = \sigma_{q_t}(c_{q_t}\sigma_{q_t} - c_{\theta_l}\sigma_{\theta_l}r_{q_t\theta_l})C.$$

*c. Lewellen–Yoh (12 parameters; input moments $\overline{w}$, $\overline{w'^2}$, $\overline{w'^3}$, $\overline{\theta_l}$, $\overline{w'\theta_l'}$, $\overline{q_t}$, $\overline{w'q_t'}$, $\overline{\theta_l'^2}$, $\overline{q_t'^2}$, $\overline{q_t'\theta_l'}$, $\overline{\theta_l'^3}$, $\overline{q_t'^3}$)*

This PDF is based on a double Gaussian form

$$P_{\text{LY}}(w, \theta_l, q_t) = aG_1(w, \theta_l, q_t) + (1 - a)G_2(w, \theta_l, q_t),$$

where $G_1$ and $G_2$ are three-dimensional Gaussian PDFs, each representing one 'plume':

$$G_1(w, \theta_l, q_t) = (2\pi)^{-3/2}\sqrt{|\mathbf{A}_1|} \exp\left[-\frac{1}{2}(x_i - x_{i1})A_{ij1}(x_j - x_{j1})\right]. \tag{A9}$$

The second plume $G_2$ has an analogous form. Here $\mathbf{A}_1 = \mathbf{C}_1^{-1}$ is the inverse of the covariance matrix of the within-plume correlations, where

$$C_{ij,1} = \overline{(x_i - x_{i1})(x_j - x_{j1})}.$$

Here, $i$ and $j$ vary between 1 and 3, with $x_1 = w$, $x_2 = \theta_l$, and $x_3 = q_t$. The mean values of plume 1 are $x_{11} = w_1$, $x_{21} = \theta_{l1}$, and $x_{31} = q_{t1}$.

The PDF parameters are the amplitude of plume 1 ($a$); the means of plume 1 ($w_1$, $q_{t1}$, $\theta_{l1}$); the standard deviations of plume 1 ($\sigma_{w1}$, $\sigma_{q_t1}$, $\sigma_{\theta_l1}$); and the within-plume correlations of plume 1 ($r_{w\theta_l1}$, $r_{wq_t1}$, $r_{q_t\theta_l1}$). There are analogous PDF parameters for the second plume. This leads to a total of 19 PDF parameters. Lewellen and Yoh (1993) use diagnostic assumptions to obtain these 19 parameters from the 12 prognosed moments. A detailed description of the procedure is provided in Lewellen and Yoh (1993) and is summarized here.

Lewellen and Yoh's PDF parameters satisfy the equations obtained by integrating the 12 relevant moments over the double Gaussian PDF. We list four of these equations; the others are analogous:

$$\begin{aligned}
\overline{w} &= aw_1 + (1 - a)w_2,\\
\overline{w'^2} &= a[(w_1 - \overline{w})^2 + \sigma_{w1}^2]\\
&\quad + (1 - a)[(w_2 - \overline{w})^2 + \sigma_{w2}^2],\\
\overline{w'^3} &= a[(w_1 - \overline{w})^3 + 3(w_1 - \overline{w})\sigma_{w1}^2]\\
&\quad + (1 - a)[(w_2 - \overline{w})^3 + 3(w_2 - \overline{w})\sigma_{w2}^2],\\
\overline{w'q_t'} &= a[(w_1 - \overline{w})(q_{t1} - \overline{q_t}) + r_{wq_t1}\sigma_{w1}\sigma_{q_t1}]\\
&\quad + (1 - a)[(w_2 - \overline{w})(q_{t2} - \overline{q_t}) + r_{wq_t2}\sigma_{w2}\sigma_{q_t2}].
\end{aligned}$$

To satisfy these equations for the moments, we first select the skewness with the largest magnitude:

$$\text{Sk}_{\text{max}} = \max(|\text{Sk}_w|, |\text{Sk}_{\theta_l}|, |\text{Sk}_{q_t}|).$$

Then we compute $a$ via

$$\begin{aligned}
a &= 0.75, \quad \text{if } |\text{Sk}_{\text{max}}| \le 0.84, \quad \text{and}\\
a^6 - \text{Sk}_{\text{max}}^2(1 - a) &= 0, \quad \text{if } |\text{Sk}_{\text{max}}| > 0.84.
\end{aligned}$$

The transcendental equation for $a$ must be solved numerically.

Next, we compute the locations $w_1$ and $w_2$, and standard deviations $\sigma_{w1}$ and $\sigma_{w2}$, of the two Gaussians in the $w$ coordinate. First, we compute $B_w \equiv w_2 - w_1$ according to

$$B_w = \text{sign}(\text{Sk}_w)\sqrt{\overline{w'^2}}\left(\frac{|\text{Sk}_w|}{1 - a}\right)^{1/3}.$$

Then we write

$$w_1 = \overline{w} - B_w(1 - a),$$
$$w_2 = \overline{w} + B_w a,$$
$$\sigma_{w1}^2 = \overline{w'^2} - B_w^2(1 - a)(1 + a + a^2)/(3a),$$
$$\sigma_{w2}^2 = \overline{w'^2} + B_w^2(1 - a)^2/3.$$

Analogous formulas hold for the means and widths of the Gaussians in $\theta_l$ and $q_t$.

The within-plume correlation between $w$ and $\theta_l$ is given by

$$r_{w\theta_l 1} = \frac{\overline{w'\theta_l'} - B_w B_{\theta_l} a(1 - a)}{a\sigma_{w1}\sigma_{\theta_{l1}} + (1 - a)\sigma_{w2}\sigma_{\theta_l 2}}. \quad \text{(A10)}$$

Physically, we must have $-1 \le r_{w\theta_l 1} \le 1$. However, the above formula yields a value outside this range for certain values of the moments. To avoid this, we set $-1 + r_{\text{thresh}} < r_{w\theta_l 1} < 1 - r_{\text{thresh}}$, where $r_{\text{thresh}} = 0.05$. We find $r_{wq_t 1}$ and $r_{q_t\theta_l 1}$ using formulas analogous to (A10). We use the same value of $r_{\text{thresh}}$ to restrict $r_{wq_t 1}$ and $r_{q_t\theta_l 1}$ to physically possible values. Furthermore, we set the within-plume correlations in the two plumes equal: $r_{w\theta_l 2} = r_{w\theta_l 1}$ and $r_{wq_t 2} = r_{wq_t 1}$.

Despite the adjustments to the within-plume correlations, for some PDFs, the determinant of $\mathbf{A}_1$, $|\mathbf{A}_1|$, can still turn out to be negative. In this case, the factor $\sqrt{|\mathbf{A}_1|}$ in (A9) renders the PDF imaginary. To prevent this, we ensure

$$1/|\mathbf{A}_1| \propto 1 - r_{w\theta_l 1}^2 - r_{wq_t 1}^2 - r_{q_t\theta_l 1}^2 + 2r_{w\theta_l 1}r_{wq_t 1}r_{q_t\theta_l 1} > 0$$

by adjusting $r_{q_t\theta_l 1}$ such that

$$r_{w\theta_l 1}r_{wq_t 1} - \sqrt{(1 - r_{w\theta_l 1}^2)(1 - r_{wq_t 1}^2)} < r_{q_t\theta_l 1} < r_{w\theta_l 1}r_{wq_t 1} + \sqrt{(1 - r_{w\theta_l 1}^2)(1 - r_{wq_t 1}^2)}.$$

Finally, we set $r_{q_t\theta_l 2} = r_{q_t\theta_l 1}$.

Cloud fraction, $\overline{q_l}$, and $\overline{w'q_l'}$ for the Lewellen–Yoh scheme are

$$C = a(C)_1 + (1 - a)(C)_2, \quad \text{(A11)}$$

$$\overline{q_l} = a(\overline{q_l})_1 + (1 - a)(\overline{q_l})_2, \quad \text{and} \quad \text{(A12)}$$

$$\overline{w'q_l'} = a[(w_1 - \overline{w})(\overline{q_l})_1 + (\overline{w'q_l'})_1] + (1 - a)[w_2 - \overline{w})(\overline{q_l})_2 + (\overline{w'q_l'})_2]. \quad \text{(A13)}$$

Here $(C)_1$, $(\overline{q_l})_1$, and $(\overline{w'q_l'})_1$ are given respectively by the single Gaussian equations (A6), (A7), and (A8), except that they are evaluated with respect to the mean and variances and covariances of Gaussian 1, which has amplitude $a$. Likewise, $(C)_2$, $(\overline{q_l})_2$, and $(\overline{w'q_l'})_2$ are associated with Gaussian 2.

In addition, we have the formulas

$$\begin{aligned}
\overline{w'^2q_l'} &= a\{[(w_1 - \overline{w})^2 + \sigma_{w1}^2][(\overline{q_l})_1 - \overline{q_l}] + 2(w_1 - \overline{w})(\overline{w'q_l'})_1 + (\overline{w'^2q_l'})_1\} \\
&\quad + (1 - a)\{[(w_2 - \overline{w})^2 + \sigma_{w2}^2][(\overline{q_l})_2 - \overline{q_l}] + 2(w_2 - \overline{w})(\overline{w'q_l'})_2 + (\overline{w'^2q_l'})_2\}, \\
\overline{\theta_l'q_l'} &= a[(\theta_{l1} - \overline{\theta}_l)(\overline{q_l})_1 + (\overline{\theta_l'q_l'})_1] + (1 - a)[(\theta_{l2} - \overline{\theta}_l)(\overline{q_l})_2 + (\overline{\theta_l'q_l'})_2], \quad \text{and} \\
\overline{q_t'q_l'} &= a[(q_{t1} - \overline{q_t})(\overline{q_l})_1 + (\overline{q_t'q_l'})_1] + (1 - a)[(q_{t2} - \overline{q_t})(\overline{q_l})_2 + (\overline{q_t'q_l'})_2]. \quad \text{(A14)}
\end{aligned}$$

*d. Analytic double Gaussian 1 (10 parameters, input moments $\overline{w}$, $\overline{w'^2}$, $\overline{w'^3}$, $\overline{\theta_l}$, $\overline{w'\theta_l'}$, $\overline{q_t}$, $\overline{w'q_t'}$, $\overline{\theta_l'^2}$, $\overline{q_t'^2}$, $\overline{q_t'\theta_l'}$)*

Like the Lewellen–Yoh formulation, the analytic double Gaussian families of PDFs are based on the double Gaussian

$$P_{\text{adg1}}(w', \theta_l', q_t') = aG_1(w', \theta_l', q_t') + (1 - a)G_2(w', \theta_l', q_t'), \quad \text{(A15)}$$

where $G_1$ and $G_2$ are Gaussians. For the analytic double Gaussians, however, the parameters can be found analytically. To achieve this simplicity, we invoke several assumptions. First, we assume that *subplume* variations in $w$ are uncorrelated with those in $q_t$ or $\theta_l$. Then, if we let $i = 1$ or 2, the individual Gaussians are given by

$$\begin{aligned}
G_i(w', \theta_l', q_t') &= \frac{1}{(2\pi)^{3/2}\sigma_{wi}\sigma_{q_t i}\sigma_{\theta_l i}(1 - r_{q_t\theta_l i}^2)^{1/2}} \exp\left[-\frac{1}{2}\left(\frac{w' - (w_i - \overline{w})}{\sigma_{wi}}\right)^2\right] \\
&\quad \times \exp\left(-\frac{1}{2(1 - r_{q_t\theta_l i}^2)}\left\{\left[\frac{q_t' - (q_{ti} - \overline{q}_t)}{\sigma_{q_t i}}\right]^2 + \left[\frac{\theta_l' - (\theta_{li} - \overline{\theta}_l)}{\sigma_{\theta_l i}}\right]^2 \right.\right. \\
&\quad \left.\left. - 2r_{q_t\theta_l i}\left[\frac{q_t' - (q_{ti} - \overline{q}_t)}{\sigma_{q_t i}}\right]\left[\frac{\theta_l' - (\theta_{li} - \overline{\theta}_l)}{\sigma_{\theta_l i}}\right]\right\}\right).
\end{aligned}$$

The PDF parameters of $P_{\text{adg1}}$ are the same as for Lewellen and Yoh, except that we set $r_{wq_t1,2} = r_{w\theta_l1,2} = 0$. The PDF parameters are determined by the following procedure. If there is no variability in $w$—that is, $\overline{w'^2} = 0$—then the skewness of $w$, $\text{Sk}_w \equiv \overline{w'^3}/(\overline{w'^2}^{3/2})$, is undefined and we cannot determine the PDF parameters. In the data, $\overline{w'^2} = 0$ never occurs. In a numerical model, however, $\overline{w'^2} = 0$ is an important special case. Golaz et al. (2002a,b) assume that in this case the PDF reduces to a single delta function. If $\overline{w'^2} \neq 0$, then we determine the relative amplitude of the Gaussian $a$ and the centers of the Gaussians in $w$ space, $w_1$ and $w_2$, by integrating over the PDF to obtain moment equations for $\overline{w}$, $\overline{w'^2}$, and $\overline{w'^3}$. We assume that the widths of the two Gaussians in $w$ are equal; that is, $\sigma_{w1} = \sigma_{w2}$. Then we find

$$a = \frac{1}{2}\left\{1 - \text{Sk}_w\left[\frac{1}{4(1-\tilde{\sigma}_w^2)^3 + \text{Sk}_w^2}\right]^{1/2}\right\}, \tag{A16}$$

$$\tilde{w}_1 \equiv \frac{w_1 - \overline{w}}{\sqrt{\overline{w'^2}}} = \left(\frac{1-a}{a}\right)^{1/2}(1-\tilde{\sigma}_w^2)^{1/2}, \quad \text{and} \tag{A17}$$

$$\tilde{w}_2 \equiv \frac{w_2 - \overline{w}}{\sqrt{\overline{w'^2}}} = -\left(\frac{a}{1-a}\right)^{1/2}(1-\tilde{\sigma}_w^2)^{1/2}. \tag{A18}$$

To avoid dividing by a vanishingly small number in equations (A17) and (A18), we insist that $0.01 \leq a \leq 0.99$. We have defined $\tilde{\sigma}_w \equiv \sigma_{w1}/\sqrt{\overline{w'^2}} = \sigma_{w2}/\sqrt{\overline{w'^2}}$. We choose $\tilde{\sigma}_w^2 = 0.4$. As shown below, the widths of the individual Gaussians in $\theta_l$ are permitted to differ in Analytic Double Gaussian 1, and likewise for $q_t$. The assumption that $\sigma_{w1} = \sigma_{w2}$ is relaxed in Analytic Double Gaussian 2.

Now we solve for $\theta_{l1}$ and $\theta_{l2}$ from the equations for $\overline{\theta_l}$ and $\overline{w'\theta_l'}$. If there is no variability in $\theta_l$—that is, $\overline{\theta_l'^2} = 0$—then we set the means of the Gaussians equal, $\theta_{l1} = \theta_{l2} = \overline{\theta_l}$, and the widths of the Gaussians in the $\theta_l$ direction to zero, $\sigma_{\theta_l1} = \sigma_{\theta_l2} = 0$. Otherwise, we set

$$\tilde{\theta}_{l1} \equiv \frac{\theta_{l1} - \overline{\theta_l}}{\sqrt{\overline{\theta_l'^2}}} = -\frac{\overline{w'\theta_l'}/(\sqrt{\overline{w'^2}}\sqrt{\overline{\theta_l'^2}})}{\tilde{w}_2}, \tag{A19}$$

$$\tilde{\theta}_{l_2} \equiv \frac{\theta_{l2} - \overline{\theta_l}}{\sqrt{\overline{\theta_l'^2}}} = -\frac{\overline{w'\theta_l'}/(\sqrt{\overline{w'^2}}\sqrt{\overline{\theta_l'^2}})}{\tilde{w}_1}. \tag{A20}$$

Although the widths of the Gaussians in the $w$ direction are set equal, $\sigma_{w1} = \sigma_{w2}$, we allow the widths of the Gaussians in the $\theta_l$ direction, $\sigma_{\theta_l1}$ and $\sigma_{\theta_l2}$, to differ. Specifically, we integrate over the PDF to relate $\overline{\theta_l'^2}$ and $\overline{\theta_l'^3}$ to $\tilde{\sigma}_{\theta_l1,2}$. We find

$$\frac{\sigma_{\theta_l1}^2}{\overline{\theta_l'^2}} = \frac{3\tilde{\theta}_{l2}[1 - a\tilde{\theta}_{l1}^2 - (1-a)\tilde{\theta}_{l2}^2] - [\text{Sk}_{\theta_l} - a\tilde{\theta}_{l1}^3 - (1-a)\tilde{\theta}_{l2}^3]}{3a(\tilde{\theta}_{l2} - \tilde{\theta}_{l1})}, \tag{A21}$$

$$\frac{\sigma_{\theta_l2}^2}{\overline{\theta_l'^2}} = \frac{-3\tilde{\theta}_{l1}[1 - a\tilde{\theta}_{l1}^2 - (1-a)\tilde{\theta}_{l2}^2] + [\text{Sk}_{\theta_l} - a\tilde{\theta}_{l1}^3 - (1-a)\tilde{\theta}_{l2}^3]}{3(1-a)(\tilde{\theta}_{l2} - \tilde{\theta}_{l1})}. \tag{A22}$$

One difficulty of the expressions (A21) and (A22) is that the denominators vanish if $\tilde{\theta}_{l2} = \tilde{\theta}_{l1}$. Furthermore, a Gaussian cannot have negative width. Therefore we impose the condition

$$0 \leq \frac{\sigma_{\theta_l1,2}^2}{\overline{\theta_l'^2}} \leq 100. \tag{A23}$$

To find $\tilde{q}_{t1}$, $\tilde{q}_{t2}$, $\sigma_{q_t1}^2$, and $\sigma_{q_t2}^2$, we use equations that are exactly analogous to (A19), (A20), (A21), (A22), and (A23), with $q_t$ replacing $\theta_l$ everywhere.

Equations (A21) and (A22) depend on the skewness $\text{Sk}_{\theta_l}$, and the analogous equations for $q_t$ depend on $\text{Sk}_{q_t}$. However, we want the analytic double Gaussian closures to be usuable in models that eschew the computational burden of predicting $\text{Sk}_{\theta_l}$ or $\text{Sk}_{q_t}$. Therefore, we make a diagnostic assumption for these quantities, as does the single Gaussian PDF, which assumes that the skewnesses are zero, and as does the double delta PDF discussed above, which implicitly assumes relationships for $\text{Sk}_{\theta_l}$ and $\text{Sk}_{q_t}$. For the analytic double Gaussians, we simply set $\text{Sk}_{\theta_l} = 0$. It may be more realistic to set $\text{Sk}_{\theta_l} = -\text{Sk}_w$, but this assumption led to numerical instability in the single-column model of Golaz et al. (2002a,b). To better represent the skewness in cumulus layers, we set $\text{Sk}_{q_t} = 1.2\ \text{Sk}_w$. However, this leads to a noisy solution when $|\tilde{q}_{t2} - \tilde{q}_{t1}|$ is small, and hence the denominators of the expressions for $\sigma_{q_t1}^2$ and $\sigma_{q_t2}^2$ are small. To reduce the noisiness, we adjust $\text{Sk}_{q_t}$ so that the numerator becomes small when the denominator becomes small. Namely, we set $\text{Sk}_{q_t} = 1.2\ \text{Sk}_w$ when $|\tilde{q}_{t2} - \tilde{q}_{t1}| > 0.4$, we set $\text{Sk}_{q_t} = 0$ when $|\tilde{q}_{t2} - \tilde{q}_{t1}| \leq 0.2$, and we linearly interpolate between these extremes when $0.2 < |\tilde{q}_{t2} - \tilde{q}_{t1}| \leq 0.4$. These formulas for $\text{Sk}_{\theta_l}$ and $\text{Sk}_{q_t}$ are expedient, but it would be worthwhile to attempt to construct more sophisticated and realistic formulas in the future. An estimate of the possible improvement can be gleaned from the performance of the Lewellen–Yoh scheme, which assumes that all skewnesses are prognosed.

Finally, to compute the within-plume correlations, $r_{q_t\theta_l1,2}$, we set $r_{q_t\theta_l1} = r_{q_t\theta_l2}$ and integrate over the PDF to obtain an equation for $\overline{q_t'\theta_l'}$. We find

$$r_{q_t\theta_l 1,2} = [\overline{q_t'\theta_l'} - a(q_{t1} - \overline{q_t})(\theta_{l1} - \overline{\theta_l}) - (1 - a)(q_{t2} - \overline{q_t})(\theta_{l2} - \overline{\theta_l})] \div [a\sigma_{q_t 1}\sigma_{\theta_l 1} + (1 - a)\sigma_{q_t 2}\sigma_{\theta_l 2}]. \quad \text{(A24)}$$

If $\sigma_{q_t 1}\sigma_{\theta_l 1}$ and $\sigma_{q_t 2}\sigma_{\theta_l 2}$ have been set to zero, then the denominator of Eq. (A24) vanishes. In this case, we set $r_{q_t\theta_l 1,2} = 0$. Since a correlation must lie between $-1$ and 1, we also insist that

$$-1 \le r_{q_t\theta_l 1,2} \le 1.$$

The formulas for $C$, $\overline{q_l}$, $\overline{w'q_q'}$, and $\overline{w'^2 q_l'}$ are exactly as for the Lewellen–Yoh scheme [Eqs. (A11), (A12), (A13), and (A14)], except that the within-Gaussian correlation between $w$ and the conserved scalars vanishes: $r_{wq_t 1,2} = r_{w\theta_l 1,2} = 0$. Therefore, $(\overline{w's'})_{1,2}$, $(\overline{w'q_l'})_{1,2}$, and $(\overline{w'^2 q_l'})_{1,2}$ vanish.

### *e. Analytic Double Gaussian 2 (10 parameters; input moments $\overline{w}$, $\overline{w'^2}$, $\overline{w'^3}$, $\overline{\theta_l}$, $\overline{w'\theta_l'}$, $\overline{q_t}$, $\overline{w'q_t'}$, $\overline{\theta_l'^2}$, $\overline{q_t'^2}$, $\overline{q_t'\theta_l'}$)*

This scheme is exactly the same as Analytic Double Gaussian 1 except that we follow the procedure of Luhar et al. (1996) to find the widths and positions of the Gaussians in $w$. Luhar et al. (1996) proposed an analytic one-dimensional PDF that allows the widths of the individual Gaussians to differ and that reduces to a single Gaussian when $\mathrm{Sk}_w = 0$. We modify their closure trivially to permit both negative and positive $\mathrm{Sk}_w$. We also extend their closure to joint three-dimensional PDFs by following the procedure for Analytic Double Gaussian 1. The only formulas that change are those for the relative amplitude $a$ (A16), $\tilde{w}_1$ (A17), and $\tilde{w}_2$ (A18). They become, respectively,

$$a = \frac{1}{2}\left[1 - \mathrm{Sk}_w\left(\frac{1}{4/M + \mathrm{Sk}_w^2}\right)^{1/2}\right],$$

$$\tilde{w}_1 = m\tilde{\sigma}_{w1}, \quad \text{and}$$

$$\tilde{w}_2 = -m\tilde{\sigma}_{w2}.$$

Here,

$$M = \frac{(1 + m^2)^3}{(3 + m^2)^2 m^2},$$

$$\tilde{\sigma}_{w1} = \frac{\sigma_{w1}}{\sigma_w} = \left[\frac{(1 - a)}{a(1 + m^2)}\right]^{1/2},$$

$$\tilde{\sigma}_{w2} = \frac{\sigma_{w2}}{\sigma_w} = \left[\frac{a}{(1 - a)(1 + m^2)}\right]^{1/2}, \quad \text{and}$$

$$m = \frac{2}{3}|\mathrm{Sk}_w|^{1/3}.$$

These equations have the reasonable property that $\sigma_{w1} > \sigma_{w2}$ when $\mathrm{Sk}_w$ is positive, thereby producing a long tail in $w$. A long tail on the opposite side of the PDF is produced when $\mathrm{Sk}_w$ is negative. When $\mathrm{Sk}_w = 0$, this scheme reduces to a single Gaussian, and hence $\tilde{w}_1$ and $\tilde{w}_2$ vanish. This leads to infinite values of $\tilde{\theta}_{l1}$ (A19), $\tilde{\theta}_{l2}$ (A20), $\tilde{q}_{t1}$, and $\tilde{q}_{t2}$. To prevent this, we enforce $|\tilde{w}_1| > 0.05$ and $|\tilde{w}_2| > 0.05$.

## REFERENCES


Albrecht, B. A., D. A. Randall, and S. Nicholls, 1988: Observations of marine stratocumulus clouds during FIRE. *Bull. Amer. Meteor. Soc.,* **69,** 618–626.

——, 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.

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

——, 1982: Cloud-ensemble relations based on the gamma probability distribution for the higher-order models of the planetary boundary layer. *J. Atmos. Sci.,* **39,** 2691–2700.

Bray, K. N. C., and P. A. Libby, 1994: Recent developments in the BML model of premixed turbulent combustion. *Turbulent Reacting Flows.,* P. A. Libby and F. A. Williams, Eds., Academic Press, 115–151.

Chen, C., and W. R. Cotton, 1987: The physics of the marine stratocumulus-capped mixed layer. *J. Atmos. Sci.,* **44,** 2951–2977.

Chen, J.-M., 1991: Turbulence-scale condensation parameterization. *J. Atmos. Sci.,* **48,** 1510–1512.

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,* **10A,** 499–515.

Cook, A. W., and J. J. Riley, 1994: A subgrid model for equilibrium chemistry in turbulent flows. *Phys. Fluids,* **6A,** 2868–2870.

Cuijpers, J. W. M., and P. Bechtold, 1995: A simple parameterization of cloud water related variables for use in boundary layer models. *J. Atmos. Sci.,* **52,** 2486–2490.

de Roode, S. R., P. G. Duynkerke, and A. P. Siebesma, 2000: Analogies between mass-flux and Reynolds-averaged equations. *J. Atmos. Sci.,* **57,** 1585–1598.

Frankel, S. H., V. Adumitroaie, C. K. Madnia, and P. Givi, 1993: Large eddy simulation of turbulent reacting flow by assumed PDF methods. *Engineering Applications of Large Eddy Simulations,* S. A. Ragab and U. Piomelli, Eds., Vol. 162, ASME, 81–101.

Germano, M., 1992: Turbulence: the filtering approach. *J. Fluid Mech.,* **238,** 325–336.

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

——, ——, and ——, 2002b: A PDF-based model for boundary layer clouds. Part II: Model results. *J. Atmos. Sci.,* **59,** 3552–3571.

Lappen, C.-L., and D. A. Randall, 2001a: Towards a unified parameterization of the boundary layer and moist convection. Part I: A new type of mass-flux model. *J. Atmos. Sci.,* **58,** 2021–2036.

——, and ——, 2001b: Towards a unified parameterization of the boundary layer and moist convection. Part II: Lateral mass exchanges and subplume-scale fluxes. *J. Atmos. Sci.,* **58,** 2037–2051.

——, and ——, 2001c: Towards a unified parameterization of the boundary layer and moist convection. Part III: Simulations of clear and cloudy convection. *J. Atmos. Sci.,* **58,** 2052–2072.

Larson, V. E., R. Wood, P. R. Field, J.-C. Golaz, T. H. Vonder Haar, and W. R. Cotton, 2001a: Systematic biases in the microphysics and thermodynamics of numerical models that ignore subgrid-scale variability. *J. Atmos. Sci.,* **58,** 1117–1128.

——, ——, ——, ——, ——, and ——, 2001b: Small-scale and mesoscale variability of scalars in cloudy boundary layers: One-dimensional probability density functions. *J. Atmos. Sci.,* **58,** 1978–1994.

Leonard, A., 1974: Energy cascade in large-eddy simulations of turbulent fluid flows. *Adv. Geophys.,* **18A,** 237–248.
Lewellen, W. S., and S. Yoh, 1993: Binormal model of ensemble partial cloudiness. *J. Atmos. Sci.,* **50,** 1228–1237.
Lock, A. P., A. R. Brown, M. R. Bush, G. M. Martin, and R. N. B. Smith, 2000: A new boundary layer mixing scheme. Part I: Scheme description and single-column model tests. *Mon. Wea. Rev.,* **128,** 3187–3199.
Luhar, A. K., M. F. Hibberd, and P. J. Hurley, 1996: Comparison of closure schemes used to specify the velocity PDF in Lagrangian stochastic dispersion models for convective conditions. *Atmos. Environ.,* **30,** 1407–1418.
Manton, M. J., and W. R. Cotton, 1977: Formulation of approximate equations for modeling moist deep convection on the mesoscale. Atmos. Sci. Paper 266, Colorado State University, 62 pp.
Mellor, G. L., 1977: The Gaussian cloud model relations. *J. Atmos. Sci.,* **34,** 356–358.
O'Brien, E. E., 1980: The probability density function (pdf) approach to reacting turbulent flows. *Turbulent Reacting Flows.,* P. A. Libby and F. A. Williams, Eds., Springer-Verlag, 185–218.
Petersen, A. C., C. Beets, H. van Dop, P. G. Duynkerke, and A. P. Siebesma, 1999: Mass-flux characteristics of reactive scalars in the convective boundary layer. *J. Atmos. Sci.,* **56,** 37–56.
Pielke, R. A., and Coauthors, 1992: A comprehensive meteorological modeling system RAMS. *Meteor. Atmos. Phys.,* **49,** 69–91.
Pincus, R., and S. A. Klein, 2000: Unresolved spatial variability and microphysical process rates in large scale models. *J. Geophys. Res.,* **105,** 27 059–27 065.
Price, J. D., 2001: A study of probability distributions of boundary-layer humidity and associated errors in parameterized cloud-fraction. *Quart. J. Roy. Meteor. Soc.,* **127,** 739–758.
Randall, D. A., 1987: Turbulent fluxes of liquid water and buoyancy in partly cloudy layers. *J. Atmos. Sci.,* **44,** 850–858.
——, Q. Shao, and C.-H. Moeng, 1992: A second-order bulk boundary-layer model. *J. Atmos. Sci.,* **49,** 1903–1923.
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 nonprecipitating clouds. *J. Atmos. Sci.,* **34,** 344–355.
Stuart, A., and J. K. Ord, 1994: *Kendall's Advanced Theory of Statistics.* Vol. 1, 6th ed. Edward Arnold, 676 pp.
Tompkins, A. M., 2002: A prognostic parameterization for the subgrid-scale variability of water vapor and clouds in large-scale models and its use to diagnose cloud cover. *J. Atmos. Sci.,* **59,** 1917–1942.
Wang, S., and B. Stevens, 2000: Top-hat representation of turbulence statistics in cloud-topped boundary layers: A large eddy simulation study. *J. Atmos. Sci.,* **57,** 423–441.
Xu, K.-M., and D. A. Randall, 1996: Evaluation of statistically based cloudiness parameterizations used in climate models. *J. Atmos. Sci.,* **53,** 3103–3119.