NOTES AND CORRESPONDENCE

A Simple Parcel Method for Prediction of Cumulus Onset and Area-Averaged Cloud Amount over Heterogeneous Land Surfaces

PETER J. WETZEL

Laboratory for Atmospheres, NASA/Goddard Space Flight Center, Greenbelt, Maryland

25 April 1989 and 15 January 1990

ABSTRACT

The purpose of this note is to compare several methods for predicting the onset and quantitative amount of cloud cover over heterogeneous land surfaces. Among the methods tested are that of Wilde et al. (1985) and a new, simple parcel approach. Model comparison is accomplished by running each model using a series of six initial conditions from the Wangara experiment. Case days were chosen because they had relatively quiet synoptic conditions, and exhibited the formation of cumulus clouds from an initially mostly clear sky during the period of solar heating. Each model contains two or three free parameters that were systematically varied until the optimum agreement was reached between observed and predicted cloud amount. The single best run for each method was chosen based on the RMSE and coefficient of determination. The best runs are compared and plotted against the observations for the six case days.

Results of these limited tests do not necessarily suggest the absolute degree of accuracy to which low cloud cover may be predicted. This is left for a future study. Rather, the focus is on the relative skill and flexibility of the various models. It is shown that parcel methods, in which surface air is lifted to its equilibrium level while being diluted by a defined amount of mixed layer air, produce substantially superior prediction of cloud amount, particularly during periods of rapid cloud onset when the mean boundary layer top is swiftly rising through a near-neutral layer. Pending verification from independent datasets, it appears that an rms error in instantaneous cloud amount of ±10% may be achievable.

1. Introduction

The natural variability of the land surface, when heated by the sun, produces a spectrum of buoyant parcels. Both the temperature and the water vapor mixing ratio of these parcels can vary across a significant range over relatively short distances. As the daytime planetary boundary layer (PBL) depth increases with time in response to surface heating, and if conditions are right, the more moist buoyant parcels can begin to condense and form clouds as they rise. Specifically, clouds will appear when the first rising parcel reaches its lifting condensation level (LCL) before its upward motion is stopped by the inversion capping the PBL. The fraction of the rising parcels that reach their LCL should be related in some manner to the observed fraction of low (cumulus or stratocumulus) cloud cover.

The purpose of this note is to test and compare various methods of modeling PBL-capping cloud formation, including methods that explicitly track rising buoyant parcels and a method recently reported by Wilde et al. (1985; hereafter WSE). The WSE method used PBL observations to define a statistical distribution function describing the probability of occurrence as a function of height, of both the LCL and the height of the top of the boundary layer. Although they found that no single function satisfied goodness-of-fit tests in all the cases studied, they chose a three-parameter, double-exponential distribution as the best compromise. Applying this distribution to a number of observations, they were able to show a good, quantitative correspondence between diagnosed cumulus amount and the amount of vertical overlap between the LCL zone and entrainment zone. Their approach also showed skill in diagnosing cumulus onset times.

The present study uses a set of relevant PBL data selected from the Wangara experiment (Clarke et al. 1971) to examine the relationship between cumulus amount and the calculated relative humidity of rising parcels. WSE's concept of defining a statistical distribution of LCL heights and PBL top heights is compared with two explicit parcel methods. For the purposes of this study a parcel is defined by its representative or average potential temperature and mixing ratio at the surface and by a mixing factor, $m$, defined as the fraction of surface air contained in the parcel as it reaches

Corresponding author address: Dr. Peter J. Wetzel, Laboratory for Atmospheres, Code 612, NASA/Goddard Space Flight Center, Greenbelt, MD 20771.
erode clouds at the PBL top that are formed by cooler, but equally moist, parcels.

2. The model

The approach to testing the various methods for diagnosing cloud cover involves incorporating them into a simple one-dimensional prognostic PBL model with a full surface energy budget and with parameterized bulk PBL top entrainment dynamics. The model is the same one used by Wetzel and Chang (1988), and therefore will not be described in detail here. The computation of mean PBL height follows Tennekes (1973) as modified by the "Zilitinkevich correction" (Tennekes and Driedonks 1981). The surface energy budget computes a single, bulk average surface temperature within the biomass layer. Surface mixing ratio is computed as a by-product of the detailed evapotranspiration model described in Wetzel and Chang (1988). The mixed-layer mean virtual potential temperature and mixing ratio, which are used to dilute the parcel as it rises, are computed as by-products of the prognostic equation for PBL depth [see Wetzel and Chang (1988) and Wetzel (1983)].

It should be noted, however, that the results presented below using the parcel methods do not directly depend on the chosen formula for mean PBL depth, since parcels are simply lifted from the surface until they encounter air of the same temperature in the stable layer above the PBL. The parcel method is,

the top of the mixed layer. Figure 1 illustrates this simple process. The initial parcel is lifted dry-adiabatically to its equilibrium level (the level where it first becomes neutrally buoyant relative to the atmospheric sounding). By the time it reaches its equilibrium level, it is assumed to have been diluted and cooled by a fraction $(1 - m)$ of mixed-layer air. The relative humidity of the resultant parcel where it comes to rest (usually near its equilibrium point) obviously determines whether or not it forms a cloud.

Since a family of rising parcels establishes its own vertical distribution statistics for each case, there is no need to specify an invariant distribution function. The advantage of this becomes apparent when one considers how the actual vertical distribution might vary in two special cases. First, in the case of a near-neutral lapse rate above the PBL, the warmest parcels may rise far above the mean PBL top height before reaching equilibrium. They may therefore cool and condense into clouds long before the mean PBL warms enough to begin to systematically entrain the capping near-neutral layer. On the other hand, when a strong stable layer is encountered above the PBL, as illustrated in Fig. 2, the warmest, highest-rising parcels can actually act to

![Diagram of potential temperature](image)

**FIG. 1.** Schematic "paths" or rising parcels, with initial surface potential temperature $\theta_s$, to their equilibrium point assuming various amounts of entrained mixed layer air. In this figure lines of constant potential temperature are vertical. The heavy solid line represents the mean atmospheric profile of potential temperature. Values of $m$ greater than 1 represent a mathematical construct useful to describe parcels warmer and wetter than the area-mean surface value. The entrainment factor is assumed to be the same for mixing ratio (profiles not shown) as for potential temperature.

![Diagram of thermodynamic diagram](image)

**FIG. 2.** Schematic thermodynamic diagram (skew $T$-log$p$) illustrating the entrainment zone and the LCL zone. The heavy solid line represents the mean atmospheric temperature sounding. The entrainment zone is the vertical region where parcels rising dry-adiabatically (following the lines sloping upward to the left) encounter the boundary-layer-capping inversion. The LCL zone is the shaded region where the adiabats intersect the mixing ratio lines (dashed and sloping upward to the right) of the family of parcels. Percentages indicate the proportion of parcels having potential temperatures or mixing ratios less than the value represented by the particular labeled line. Note that in this case, although the entrainment zone intersects the LCL zone in the vertical coordinate, no clouds form because no parcel is simultaneously cool and moist enough to allow condensation.
in fact, an entirely independent means of defining a local PBL top height.

Once clouds form, they are assumed to reflect 60% of solar radiation and to be blackbodies to longwave radiation. No buoyant effects due to latent heat release were considered. For very shallow clouds, stratocumulus for example, the radiational effects are likely to be more important to cloud dynamics than the latent heat effects. As the vertical extent of the cloud increases, the buoyancy generated by latent heat release in the rising parcels begins to dominate. Since latent heat release is not considered in any of the models tested here, their skill at diagnosing cloud amount may be expected to deteriorate when the cloud regime is dominated by cumulus congestus or cumulonimbus.

In the predictions of cloud amount presented below, two variations of the parcel method are compared with the method prescribed by WSE and with a simple method based on relative humidity at the mean PBL top. A description of each of the four approaches follows.

The WSE method defines an “LCL zone” as the vertical layer within which the large majority of LCLs of individual surface layer parcels occur. Clouds are diagnosed to exist when all or part of the LCL zone falls beneath all or part of the PBL “entrainment zone.” The “entrainment zone” is defined as the layer within which the highest 90% of the local PBL top heights occur. A schematic of this approach as applied to a typical diurnal cycle, under initially clear skies and undisturbed conditions, is presented in Fig. 3. As the boundary layer rises during the day, the zone in which rising heated parcels encounter and mix with overlying undisturbed air rises with it. This zone, the entrainment zone, encompasses a range of values because individual rising parcels are of different sizes and have different degrees of buoyancy. The stippled lines in Fig. 3 represent the heights below which 10%, 50%, and 90% of the rising parcels lose their upward momentum and cease to rise. If these parcels are sufficiently moist, they may reach their lifting condensation level and form clouds. The solid lines in Fig. 3 represent the range of LCLs for this case. The LCL zone also encompasses a range of values because all parcels do not have identical water vapor content or temperature. Only 10% of the parcels have the combination of high water vapor content and low potential temperature that will enable them to condense at the lowest of the solid lines. The remaining parcels must rise further before they condense. The shaded area in Fig. 3 represents the zone in which cloud cover of varying degrees may potentially occur. Presumably the greater the overlap between the LCL zone and the entrainment zone, the greater will be the probability of cloud cover and/or the more areal extent the clouds could be expected to have.

The vertical overlap of the LCL zone and the entrainment zone as shown in Fig. 3 is a necessary condition for boundary-layer-capping cover; however, it is not a sufficient condition in all cases. Particularly where strong static stability exists above the boundary layer, the two zones may overlap substantially without any cloud formation. This situation is depicted in Fig. 2, which is a schematic skew T-logP diagram. In this diagram lines of constant potential temperature are represented by the solid and dotted lines that slope upward to the left, while lines of constant water vapor mixing ratio are represented by the dashed lines that slope upward to the right. The heavy solid line segments represent the temperature sounding of the mean atmosphere. The shaded area indicates the zone on this thermodynamic diagram where the distribution of rising parcels, with their range of potential temperatures, \( \theta \), and mixing ratios, \( q \), reach condensation. Note that this LCL zone substantially overlaps the entrainment zone (here depicted without any parcel overshoot for simplicity), but that all rising parcels will be stopped by the capping inversion before they can reach condensation. Also note that if all rising parcels were allowed to rise a small increment further, it would be only the coolest and wettest parcels that would condense. In that case the warmer parcels, which can be expected to have more vigorous vertical motion, will act to erode any cloud cover that occurs.

Our application of the WSE method employs the statistical distribution functions exactly as prescribed in their paper. Two separate two-tailed exponential functions characterize the probability distribution with height of the LCL zone and the entrainment zone. Each

![Figure 3](image-url)
requires selection of three free parameters, for a total of six free parameters. In order to allow the WSE method to optimally fit the Wangara dataset, we performed preliminary tests to define the two dominant parameters: the mean LCL height and the maximum (90%) height of the LCL.

In order to select the optimum values of these parameters for our dataset, the properties of the air used to calculate the LCL were systematically varied, by means of the mixing factor \( m \) operating on the surface and mixed layer values of temperature and mixing ratio as in the parcel method. Values of \( m \) between 0 and 1.6 were tested (note that values of \( m > 1 \) are simply a mathematical construct, implying that the resultant parcel is warmer than the area-mean surface temperature; see Fig. 1). Based on the results of 216 model runs with the Wangara dataset, it was found that the lowest RMS error between observed and computed cloud amount was achieved with the WSE method using a mixing factor of 0.8 for the mean LCL height and 1.3 for the maximum LCL height. The base of the LCL zone is set, as in WSE, to be the same vertical distance below the mean LCL as the maximum value is above it. The third free parameter in the two-tailed exponential distribution, the shape factor, was assigned a fixed value of 0.01 after it was found that its specification had only a small bearing on the results.

Next we specify the three free parameters in the entrainment zone distribution. We use the mean PBL height as calculated by our model to define the distribution mean. The equilibrium level of the undiluted \( (m = 1) \) average surface parcel is specified as the top of the zone, and the shape parameter is assigned a constant value of 0.05, as in WSE. Because of the small value of the shape parameter, the WSE model is not very sensitive to the exact specification of the entrainment zone top. However, it is obviously sensitive to any errors in prediction of the PBL depth by our one-dimensional model.

The second method tested is a simple one, similar to that used in some current-generation three-dimensional models. The cloud amount at any time is assigned based on the relative humidity (RH) at the mean PBL top. To provide for a gradation of cloud amount rather than a simple yes/no cloud diagnosis, we assumed that the relative humidity at which clouds appear is not a fixed number, but rather is described by a normal distribution function with a mean, \( RH \), and with a standard deviation \( \sigma_{RH} \). The fractional cloud cover is then defined as the probability that the computed RH at the PBL top exceeds \( RH \), given the value of \( \sigma_{RH} \). In a series of model runs, various values of the two free parameters of the normal distribution were tested to find the least RMS error between observed and predicted cloud cover. As a result of 126 runs, the values \( RH = 0.95 \) and \( \sigma_{RH} = 0.10 \) were chosen as optimum for predicting cloud amount from the selected Wangara dataset.

Third, a simple one parcel method was tested, in which the RH of a parcel at its equilibrium point is computed using the mixing parameter \( m \). As in the previous method, a gradation in cloud amount is established by calculating the probability that the parcel RH at the equilibrium point exceeds the normally distributed saturation value defined by RH and \( \sigma_{RH} \). In addition to the two free parameters of the normal distribution, the mixing fraction \( m \) of the parcel is allowed to vary freely. Again, all three parameters were systematically varied in a matrix of model runs to establish the optimum values. In 540 model runs, the lowest RMS error between observed and predicted cloud amount for the Wangara dataset occurred for \( RH = 1.00, \sigma_{RH} = 0.075 \), and \( m = 0.1 \). This implies that for the dataset tested, the average parcel that produced a cloud consisted of 10% surface air and 90% mixed-layer air.

Lastly, some exploratory tests were performed assuming a preselected distribution of nine simultaneously occurring parcels with different potential temperatures and mixing ratios. The nine parcels were defined by a 3 \( \times \) 3 matrix of surface potential temperature and mixing ratio values, separated from one another by a fraction, \( n \), of the difference between the mixed layer mean value (subscript \( M \)) and the surface value (subscript \( s \); see Table 1 for details). Recognizing that parcels with different initial states are likely to occur with unequal frequency over a real surface, but lacking guidance from observation, we weighted the relative abundance of the nine parcels in an ad hoc manner. Table 1 shows the chosen weights. Common sense dictates that it is appropriate to give lower weight to the greater deviations from the mixed layer means of potential temperature and mixing ratio. Beyond that statement, however, the values in Table 1 are chosen without justification. Therefore results of the nine parcel experiment should only be considered as an academic exercise. In this series of experiments the mixing parameter \( m \) is replaced by the fractional separation, \( n \), between the initial states of the various parcels (see Table 1). These tests could by no means be exhaustive, but after 276 runs, values of \( RH = 0.975 \), \( \sigma_{RH} = 0.05 \) and \( n = 0.3 \) were found to have the lowest RMS error between observed and calculated cloud amount.

| Table 1. Relative weighting factors for parcel properties defined by the mixed layer mean potential temperature and mixing ratio (subscript \( M \)), the surface value (subscript \( s \)), and a free parameter \( n \). |
|-----------------|-----------------|-----------------|
| \( \theta_M \)  | \((1 - n)\theta_M + n\theta_s\) | \((1 - 2n)\theta_M + 2n\theta_s\) |
| \( q_M \)      | .515            | .162            | .042            |
| \( (1 - n)q_M + nq_s \) | .162 | .042 | .015 |
| \( (1 - 2n)q_M + 2nq_s \) | .042 | .015 | .005 |
from the experiment site. This evidence includes the time series of temperature, pressure, wind speed, and direction, and the boundary layer profile data, all of which show no sign of disturbance from cumulonimbus outflow or updrafts. No thunder, lightning or precipitation were reported at any time. It must be noted that in the dry, anticyclonic, wintertime conditions found on these days, a cumulonimbus cloud could be sighted from as far as 200 km away. Based on these considerations, it was decided to retain days 4 and 28 as part of the experimental base. As will be shown, results on these two days are very similar in all respects to the other four days, and therefore may prove valuable in defining the limits of applicability of the models. Finally, however, we note that the peculiarities of the weather conditions are confronted equally by all models. The purpose of this paper is to test comparative performance of various cloud cover models under the kind of real-world conditions in which it is hoped they can perform.

Results for the 6 days appear in Figs. 4–9. Comparison between the observed and modeled development of low cloud cover on day 4 (Fig. 4) is as good as any day studied despite the late afternoon appearance of cumulonimbus. Cloud onset is simulated well by all methods except the "RH at PBL top." That and the Wilde et al. (1985) method both tend to delay the rapid morning buildup of cloudiness 1 to 2 hours later than it was observed to occur.

On day 16 (Fig. 5), a possible deck of non-surface-based low clouds may have interfered with the observed low cloud total. Clearly the 20%–30% presunrise cloudiness was not rooted in surface heating and it is suspected that this component of the cloudiness may have persisted through the day. Relative to one another, the four methods perform in a manner similar to their behavior on day 4.

The morning increase in cloud cover is remarkably well duplicated by the nine-parcel method on day 19 (Fig. 6), while the single-parcel method underestimates morning cloudiness and overestimates afternoon amounts.

On day 28 (Fig. 7), the sequence of soundings suggests some possible upward motion. Cumulonimbus are observed in three of the afternoon reports, and a few scattered showers occurred in the evening, reinforcing this suspicion. This day is clearly the most disturbed of the six chosen. Nevertheless the modeled trend of cloud onset and the relative performance of the various models are similar to the other five days. As in day 16, some predawn cloud cover may interfere with the observations, although in this case it appears to have burned off by 1000 local time. Cloud cover is duplicated quite well by the two parcel methods on this day.

On day 32 (Fig. 8), the method of Wilde et al. does the best job of simulating early cumulus onset, but
later in the morning as cloud cover continues to increase it lags the observations. In the WSE paper, the case days chosen for verification never developed more than 20% cloud cover. It is possible therefore that these are the conditions for which their method can be expected to perform best.

Finally, on day 39 (Fig. 9), the soundings indicate some dry advection aloft. Mixing ratios drop from 3.0 to 2.2 g kg$^{-1}$ at 800 mb between sunrise and 1500. This trend may be reflected by the observed decrease in cloudiness during the afternoon. In any event, the afternoon cloud cover is poorly simulated by the parcel methods on this day.

Although the detailed results of Figs. 4–9 varied significantly from day to day, it is clear that on every day the parcel methods outperform the other methods by a rather wide margin. Here is the statistical summary of results for the 6 days as a whole: For the "relative humidity at mean PBL top" method, the RMS error in units of fractional cloud cover is 0.252 with a coefficient of determination of $R^2 = 0.52$ between observed and predicted values of cloud cover. The RMSE is 0.218 for the WSE method with $R^2 = 0.72$, 0.133 for the single-parcel method with $R^2 = 0.86$, and 0.102 for the nine-parcel method with $R^2 = 0.90$.

The results presented here have compared model skill in simulating cloud amount when clouds were observed. The question may arise: Do the models successfully duplicate the absence of cloud cover when none occurred? To answer this, tests were performed using data from day 33, a quiescent, cloudless day. All models successfully predicted the complete lack of cloudiness.

The largest errors for the "RH at PBL top" and WSE methods occur during the rapid buildup of cloud cover in late morning. During this period on nearly all days, the PBL was observed to be growing through a layer of relatively weak stability. This allowed individual buoyant parcels to greatly exceed the mean PBL height, as discussed earlier.

There is one potential source of systematic error for the "RH at PBL top" and WSE methods—the param-
eterization used to calculate the mean PBL depth in the one-dimensional model. Both of these methods depend heavily on this number, which is taken from a model rather than from the observations (the latter being wholly inadequate because of their temporal infrequency). Figure 10 compares the model-calculated mean PBL depth with the observed values. The observed PBL depth was obtained by careful subjective analysis of the temperature, wind, and mixing ratio profiles for all daytime observations on the six days of the case study. Although there is a small bias toward underprediction of the depth by the model, especially in late afternoon (i.e., the slope of the best linear fit is greater than 1), the agreement with observations is generally quite good. The late afternoon underestimate is most probably caused by the lack of latent heat release by clouds in the model, and does not correspond with the time of worst performance by the WSE and “RH at PBL top” methods. It is therefore concluded that the model PBL depth parameterization is not the primary cause of the errors of these two methods.

4. Conclusions and discussion

Methods have been proposed by Wilde et al. (1985) and herein, whereby the spatial variability of the natural land surface may be considered when predicting the onset and amount of surface-rooted cumulus cloud cover. During the coming era of concern for, and modeling of, the area-averaged response of the atmosphere to altered land surface states, this form of methodology is likely to be called forth. Accurate prediction of cloud cover is paramount to understanding the energy balance of the earth’s surface. It is important for the modeling of all time and space scales of atmospheric motion from the mesoscale to the global climate scale. The parcel methods discussed here are shown to produce a significant improvement in the reliability of such predictions relative to the other methods tested. Although independent verification is required, it appears to be possible to predict low cloud cover to within an RMS error of ±10%, at least in a research mode under conditions of minimal advection and baroclinic activity.

The Wangara data do not indicate the vertical extent of cloudiness, nor do they seem to allow for reporting of cumulus congestus or similar transition states between cumulus and cumulonimbus. Lacking satellite evidence, it is not possible to clearly categorize the cloud regime on the chosen case days. However, as discussed earlier, all possible care was taken to avoid cases where effects of deep cumulus are apparent in the surface pressure, temperature, moisture or wind data, or in the profile data.

Whenever clouds are present the effects of latent heat release are obviously nonzero. Since none of the tested models consider the effects of latent heat release (or other deep cloud phenomena such as downwash density currents), conclusions reported here should be applied in the strictest sense only to relatively shallow cloud cases. The question of to what depth the clouds may grow before these models become invalid was not explored in this study. It is encouraging that the models seem to function adequately in less than ideal real-world environments such as day 28. This suggests that a simple parcel model might be successfully applied to more disturbed cases, such as to predict the precursor cloudiness in severe storm conditions. More testing will be required to confirm this indication.

Cases with significant advection and large scale baroclinic effects may be simulated by incorporating the simple cloud model into a three-dimensional weather prediction model. Any of the cloud models described here could be implemented without difficulty.
Once the values of the free parameters have been established from a satisfactorily broad database, a rapid, noninteractive computation (300 or fewer lines of code for all models) provides a prediction of cloud amount for a particular grid point at a particular time. Under "fair weather" conditions it should be satisfactory to repeat the computation only once every quarter hour of real time. However, care would obviously be required when considering interactions of such a model with cumulus parameterization schemes, and under conditions of large scale (resolvable) latent heat release, such as in midlatitude cyclones.

By no means do the results presented here represent an exhaustive study, as data were selected from only one site and from one time of year. Results should be strictly interpreted as applying only to this dataset. However, tests with the nine parcel model, applied to a summertime Great Plains case using the parameter values established for Wangara ($\overline{RH} = 0.975$, $\sigma_{RH} = 0.05$ and $n = 0.3$) have shown that the model is both transportable and quantitatively useful to predict differential cumulus onset times over a landscape of variable surface characteristics (Rabin et al. 1990).

Many possible alterations to all methods tested can be conceived that would improve their respective results. The WSE method, in particular, might be shown to produce significantly improved results if the statistical distribution functions were defined in a vertical coordinate dependent on potential temperature rather than on pressure or height. That experiment is, however, left for a future study.

Acknowledgments. Special thanks are extended to Kelly Pecnick for her usual prompt and flawless typing of the manuscript.

REFERENCES


