Evaluating Formulations of Stable Boundary Layer Height

D. VICKERS AND L. MAHRT
College of Oceanic and Atmospheric Sciences, Oregon State University, Corvallis, Oregon

(Manuscript received 8 October 2003, in final form 23 March 2004)

ABSTRACT

Stable boundary layer height \( h \) is determined from eddy correlation measurements of the vertical profiles of the buoyancy flux and turbulence energy from a tower over grassland in autumn, a tower over rangeland with variable snow cover during winter, and aircraft data in the stable marine boundary layer generated by warm air advection over a cool ocean surface in summer. A well-defined \( h \) within the tower layer at the grass site (lowest 50 m) and the snow site (lowest 30 m) was definable only about 20% of the time. In the remaining stable periods, the buoyancy flux and turbulence energy either (a) remained constant with height, indicating a deep boundary layer, (b) increased with height, or (c) varied erratically with height. Approximately one-half of the tower profiles did not fit the traditional concepts of a boundary layer. The well-defined cases of \( h \) are compared with various formulations for the equilibrium depth of the stably stratified boundary layer based on the Richardson number or surface fluxes. The diagnostic models for \( h \) have limited success in explaining both the variance and mean magnitude of \( h \) at all three sites. The surface bulk Richardson number and gradient Richardson number approaches perform best for the combined data. For the surface bulk Richardson number method, the required critical value varies systematically between sites. The surface bulk Richardson number approach is modified to include a critical value that depends on the surface Rossby number, which incorporates the influence of surface roughness and wind speed on boundary layer depth.

1. Introduction

Numerous numerical models of the atmosphere and atmospheric dispersion require an estimate of boundary layer depth in their parameterization of mixing processes. Existing definitions of the stable boundary layer depth \( h \) include (a) the layer through which surface-based turbulence extends (e.g., Lenschow et al. 1988), (b) the top of the layer with downward heat flux (Caughey et al. 1979), (c) the height of the low-level jet or minimum wind shear (Melgarejo and Deardorff 1974), and (d) the top of the temperature inversion layer or the layer with significant cooling (Yamada 1976). Definition a is probably the most consistent with traditional thinking of a boundary layer as being the layer in contact with the surface with at least intermittent turbulence. However, with partially collapsed turbulence near the surface because of strong stratification, even weak elevated sources of turbulence can cause the turbulence energy to remain constant or even increase with height above the boundary layer. This can occur, for example, when the principle source of turbulence is shear generation associated with elevated nocturnal accelerations and formation of a low-level jet. A boundary layer may not be definable in such cases. Definition b is relatively straightforward to apply when the measurements are available, and it will be used in this study. Even with elevated turbulence that is partially decoupled from the surface, the heat flux often becomes small above the boundary layer because of a decrease in the mean vertical temperature gradient, making the buoyancy flux an easily defined indicator of boundary layer depth. Definitions c and d could be considered indirect methods that rely on characteristics of the mean flow rather than the turbulence.

Previous work has defined \( h \) primarily in terms of mean wind and temperature profiles because adequate vertical resolution of the turbulence energy and fluxes has not been available. Wetzel (1982) noted that the goal of these methods is to identify some feature in the mean profile that unambiguously relates to the vertical extent of the turbulent or thermal effects of the surface. However, the relationship between the mean profiles and the turbulence can be complex. For example, the depth of the layer with significant turbulence and downward heat flux may be less than the depth of the temperature inversion layer, which may be more related to the time history of the turbulence and can include effects of radiational cooling (Brutsaert 1972; Yu 1978; Mahrt et al. 1979; Mahrt 1981). Radiational cooling is relatively more important in weak-wind stable boundary layers (Krishna et al. 2003). The nocturnal boundary layer
depth, as determined by flux profiles, sometimes has no obvious relationship with the mean profiles of wind and temperature. Estimates of h based on mean temperature profiles are often poorly correlated with estimates based on mean wind profiles (Mahrt and Heald 1979).

In this study, we determine h based on profiles of the buoyancy flux when the measurements are available. The main reason is that the heat fluxes go to small values above the surface inversion because the potential temperature gradient becomes small, while momentum flux may not become small. One could argue that this approach underestimates the boundary layer depth, but any layer with turbulence in communication with the ground should experience significant cooling. For this reason, estimates of h based on the momentum flux profile are thought to be less reliable. For the nocturnal tower data over land the boundary layer is defined as the depth of the stratified part of the flow associated with surface cooling that is at least intermittently turbulent with downward heat flux.

Numerous different formulations of h have been proposed for use in models (section 2). Most of these rely on a single dataset for evaluation of the method and the coefficients, and it is not always clear if the results are applicable to other sites where different physics may be important. We evaluate formulations for h using two datasets over land and one over the ocean, with each representing different surface and meteorological conditions. The first type is characterized by strong surface radiative cooling and strong stability over a 60-m tower located in Kansas grassland in autumn during the Cooperative Atmosphere–Surface Exchange Study in 1999 (CASES99). The potential temperature gradient decreases with height close to the surface (negative curvature of the potential temperature profile). The second dataset is from aircraft observations of stable boundary layers associated with summertime warm-air advection from the northeastern United States over the cool Atlantic Ocean during the Coupled Boundary Layers Air–Sea Transfer (CBLAST) experiment. The potential temperature gradient is usually either constant with height or increases with height (positive curvature). A well-defined low-level jet is nearly always located at the top of these stable marine boundary layers. The third stable boundary layer dataset is from a 34-m tower in Colorado rangeland at 2500-m elevation during winter in the Fluxes over Snow Surfaces (FLOSS) experiment where the ground was often covered by snow.

2. Background

Formulations for the equilibrium depth of the stably stratified boundary layer fall into two main categories—Richardson number forms, where the bulk stratification and wind shear are employed, and surface flux–based formulations that require estimating the turbulent heat and momentum fluxes at the surface. These formulations have been recently reviewed by Zilitinkevich and Mikhailov (1996), Vogeleganz and Holstlag (1996), and Zilitinkevich and Baklanov (2002).

We note that, in practice, there are difficulties measuring surface fluxes in strongly stable conditions. The intermittent nature of the turbulence requires a long flux-averaging time to obtain a representative sample of the transporting eddies and to reduce random errors. Mesoscale motions can contaminate the calculated turbulent fluxes when using standard analysis methods that use an averaging time scale of 10 min or longer to define the turbulent fluctuations (Vickers and Mahrt 2003). Strong relative flux divergence (thin boundary layers) can cause fluxes measured at standard heights to be significantly smaller than the surface flux. When the sonic anemometer is moved down closer to the surface, flux loss caused by pathlength averaging can become significant.

a. Surface flux–based methods

Rossby and Montgomery (1935) considered the influences of friction and the earth’s rotation on boundary layer depth and proposed

\[ h = C_s \frac{u^*_h}{f}, \]

where h is the stable boundary layer height, \( u^*_h \) is the surface friction velocity or square root of the kinematic momentum flux at the surface, \( f \) is the Coriolis parameter, and \( C_s \) is a nondimensional coefficient with values ranging from 0.1 to 0.5, where larger values are associated with neutral conditions. While Ekman dynamics predict an inverse dependence of \( h \) on the Coriolis parameter, no meaningful dependence between \( h \) and \( f \) has been demonstrated with atmospheric observations. This formulation may still explain variations in \( h \) because of the correlation between \( u^*_h \) and \( h \).

Kitaigorodskii (1960) concluded that the Obukhov length \( L \) is the appropriate length scale for the depth of the layer when dominated by the surface fluxes of heat and momentum,

\[ h = C_L L, \]

where \( L = -w^*_B B_s \) and \( B_s = (g/\theta)w^*/\theta^* \) is the scaled surface buoyancy flux. The von Kármán constant is not included in his definition of \( L \). Here, \( C_L \) is a nondimensional coefficient ranging from 100 as reported by Kitaigorodskii and Joffre (1988) using data from the Wangoora and “ICE-77” experiments (Joffre 1981) to 1 for an oceanic case (Stigebrandt 1985). Zilitinkevich (1972) incorporated the influences of the earth’s rotation and the surface fluxes to obtain

\[ h = C_s \frac{u^*_h}{(-f B_s)^{1/2}}, \]

where \( C_s \) is a nondimensional coefficient of order 1 (Zilitinkevich 1989). This expression [Eq. (3)] has been
widespread use in observational and modeling studies, including air pollution dispersion modeling (see references in Zilitinkevich and Mironov 1996). Pollard et al. (1973) proposed using friction, rotation, and the free-flow stratification such that

\[ h = C_s \frac{u^*}{N}, \]  

where \( N \) is the buoyancy frequency \( \left[ N^2 = \frac{g}{\rho} \frac{\partial \theta}{\partial z} \right] \) in the free atmosphere above the boundary layer and \( C_s \) is a nondimensional coefficient equal to 1.7. This form requires surface momentum flux information and temperature information at two levels above the boundary layer to evaluate the temperature gradient. Considering surface friction and background stratification alone leads to

\[ h = C_s \frac{u^*}{N}, \]  

as proposed by Kitaigorodskii and Joffre (1988) and others, where \( N \) is intended to describe the influence of stratification on entrainment at the top of the boundary layer and \( C_s \) is a nondimensional coefficient ranging from 4 to 20.

The stable nocturnal boundary layers studied here develop within a well-mixed daytime residual layer that remains neutrally or weakly stratified during the night. The stratification is stronger within the nocturnal boundary layer than above it, and the free-flow stratification \( (N) \) is not expected to be an important constraint on \( h \). For the purpose of applying the formulations, we evaluate \( N \) using the highest measurement levels available on the tower. The marine boundary layers studied here are typically capped by a temperature inversion, and the free-flow static stability could be an important influence on \( h \). For these data we evaluate \( N \) at a height just above the boundary layer using the aircraft data.

A limiting form for the stable boundary layer height derived by Zilitinkevich and Mironov [1996, their Eq. (26)] is

\[ \left( \frac{f h}{C_s u^*} \right)^2 + \frac{h}{C_s L} + \frac{N h}{C_s u^*} = 1. \]  

A second, more complex form with two additional terms [their Eq. (30)] is

\[ \left( \frac{f h}{C_s u^*} \right)^2 + \frac{h}{C_s L} + \frac{N h}{C_s u^*} + \frac{h f L}{u^*} + \frac{h (f L)^{1/2}}{C_s u^*} = 1, \]  

where they proposed that the latter better incorporates intermediate regimes associated with boundary layers dominated by buoyancy. The expression [Eq. (6)] is intended to approximate the three asymptotic cases of truly neutral, surface heat flux–dominated, and inversion-capped boundary layers. Using a vertical profile for eddy viscosity based on numerical simulations of the stable boundary layer by Brost and Wyngaard (1978), Nieuwstadt (1981) proposed

\[ h = \frac{0.3 u^* f L}{1 + 1.9 h L}. \]  

Yu (1978) evaluated expressions for the stable boundary layer height from Deardorff (1972):

\[ h = \left( \frac{1}{30 L} + \frac{f}{0.35 u^*} \right)^{-1}, \]

and from Businger and Arya (1974):

\[ h = \left( \kappa u^* L/f \right)^{1/2}. \]

Zilitinkevich et al. (2002) derived a new formulation from the momentum equations that implicitly accounted for the turbulence kinetic energy (TKE) budget,

\[ h = \frac{C_s u^*}{f} \left[ 1 + \frac{C_s u^*}{\kappa u^* L} \left( 1 + \frac{C_s u^* L}{\kappa u^*} \right) \right]^{1/2}. \]  

Zilitinkevich and Baklanov (2002) report that Eq. (11) is a reasonable diagnostic formulation for \( h \), and they recommend its use in one-dimensional models with coefficients \( C_s = 0.4, C_s = 0.75, \) and \( C_s = 0.25 \). The same study reports that Eq. (11) was superior to Eq. (7), when evaluated from the Cabauw, Netherlands, data. Additional stable boundary layer depth formulations are listed in Zilitinkevich and Baklanov (2002, their Table IV).

b. Richardson number–based methods

An alternative approach to the surface flux formulations is to use the bulk stratification and wind shear. This is based on the idea that the depth and strength of mixing in a stratified flow are observed to either increase or decrease depending on whether the Richardson number is less than or greater than some critical value. Hanna (1969), Wetzel (1982), Troen and Mahrt (1986), Vogelegzang and Holtslag (1996), and others have proposed formulating the stable boundary layer height as the lowest level at which the bulk Richardson number exceeds a critical value. A practical drawback to this approach, in comparison with surface flux–based forms, is that the resolved \( h \) is only as good as the vertical resolution of the observations or the vertical resolution in the model. An advantage is that the difficult-to-measure surface fluxes are not required.

The surface bulk Richardson number is calculated as

\[ R_{in} = \frac{\kappa u^*}{\bar{\theta}^*} \left[ \frac{\theta(z) - \theta_s}{U(z)^2} \right], \]

where the gradients are evaluated between the surface and some level \( z \), \( \theta^* \) is the layer average potential temperature, \( \theta(z) \) is the potential temperature at \( z \), \( \theta_s \) is the surface temperature, and \( U(z) \) is the mean wind speed at \( z \). Most models carry a surface radiative temperature.
is the bottom. For \( R_i \) values indicate the top of the layer and the surface is the bottom of the layer. For \( R_s \) values are the top of the layer and 10 m is the bottom. For \( R_u \) values are the midpoint of a 10-m-thick layer.

<table>
<thead>
<tr>
<th>Exp</th>
<th>Parameter</th>
<th>Levels (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CASES99</td>
<td>( R_n )</td>
<td>5, 10, 15, 25, 35, 45, 55</td>
</tr>
<tr>
<td></td>
<td>( R_b )</td>
<td>15, 25, 35, 45, 55</td>
</tr>
<tr>
<td></td>
<td>( R_l )</td>
<td>10, 20, 30, 40, 50</td>
</tr>
<tr>
<td>FLOSS</td>
<td>( R_u )</td>
<td>5, 10, 15, 20, 30</td>
</tr>
<tr>
<td></td>
<td>( R_b )</td>
<td>15, 20, 30</td>
</tr>
<tr>
<td></td>
<td>( R_l )</td>
<td>10, 15, 25</td>
</tr>
<tr>
<td>CBLAST</td>
<td>( R_u )</td>
<td>15, 25, 35, 175, 185</td>
</tr>
<tr>
<td></td>
<td>( R_b )</td>
<td>25, 35, 175, 185</td>
</tr>
<tr>
<td></td>
<td>( R_l )</td>
<td>25, 35, 175, 185</td>
</tr>
</tbody>
</table>

for the surface energy budget, but in many observational studies the surface temperature is not available or is spatially heterogeneous.

To initiate evaluation of Eq. (12), \( R_n \) is calculated from the lowest model or observational level above the surface and the surface temperature and is compared with a critical value \( R_c \). If \( R_n > R_c \), then \( h < z \); otherwise, \( R_n \) is calculated from data at the next highest level and the surface, and so on. Wetzel (1982) replaced the surface temperature \( \theta_s \) with the modeled aerodynamic temperature \( \theta_a \) and found the critical Richardson number to be one-third based on the Wangara data. Troen and Mahrt (1986) suggest a value of one-half for \( R_s \). Vogelezang and Holtslag (1996) found \( R_s = 0.22 \) based on the Cabauw tower data; however, they used temperature and wind at 2 m instead of at the surface. Other published estimates of the critical Richardson number are summarized in Zilitinkevich and Baklanov (2002, their Table II).

Vogelezang and Holtslag (1996) proposed a bulk Richardson number approach with the air temperature at some level near the surface and the wind speed replaced by the wind shear between the two atmospheric levels,

\[
R_b = (z - z_i) \left( \frac{g}{\theta_s} \right) \frac{\theta(z) - \theta(z_i)}{[U(z) - U(z_i)]^2},
\]

where \( z \) is the top of the layer being considered and \( z_i \) is the bottom of the layer. Their study and others have found that use of the wind speed rather than the wind shear in the Richardson number, as in Eq. (12), can overestimate shear production of turbulence in strong winds. Unfortunately, Eq. (13) may be sensitive to the selection of the lower level for evaluating the gradients.

Vogelezang and Holtslag (1996) proposed that the lower level should be above the surface layer and tested various \( z_i \) values between 20 and 80 m. They did not find significant sensitivity of the calculated boundary layer height on the precise value of \( z_i \) tested using the Cabauw data. They found a value of 0.3 for the critical Richardson number \( R_c \) applicable to Eq. (13) for \( z_i \) between 20 and 80 m, where their stable boundary layer height varied from approximately 80 to 200 m. For the shallow boundary layers considered here, the surface layer is typically less than 10 m deep and we test this bulk Richardson number method using \( z_i = 10 \) m.

We also test a gradient Richardson number approach as proposed by Busch et al. (1976) and others, where

\[
R_i = \frac{\Delta z}{z} \frac{\theta(z + \Delta z/2) - \theta(z - \Delta z/2)}{U(z + \Delta z/2) - U(z - \Delta z/2)},
\]

and the gradients are calculated over a fixed layer thickness of \( \Delta z \) equal to 10 m. This approach is intended to be more sensitive to local variations in the gradients. The profile of \( R_i \) is compared with a critical value beginning with the level closest to the surface and proceeding upward, and the boundary layer height is that height where \( R_i \) exceeds some critical value \( R_{ic} \). We initially test this scheme with \( R_{ic} = 0.5 \) and 1/2.

Miles (1961) and Howard (1961) suggested that the critical Richardson number was 0.25 based on linear, invicid stability analysis. A few possible reasons may be considered for higher critical Richardson numbers observed. First, the Richardson number could be larger if the turbulence is generated over a layer that is thinner than the layer of calculation (Woods 1969; Kunkel and Walters 1982; Kim and Mahrt 1992). Second, finite-amplitude instability occurs at larger Richardson number than 0.25 (Abarbanel et al. 1984). Preexisting weak turbulence and mesoscale velocity variations seem to be always present. Last, there may be cases of large Richardson numbers in which turbulence that is advected past the tower originates upwind where the Richardson number was smaller.

Applying the Richardson number approaches [Eqs. (12)–(14)] to data fails to obtain an estimate for \( h \) when the critical value is not exceeded at the highest measurement level. In cases in which the Richardson number increases steadily with increasing height, we extrapolate the profile upward to the critical level to obtain a prediction for \( h \). Estimates of \( h \) obtained in this manner that exceed 2 times the highest data level are discarded because of uncertainty in the estimate. The list of vertical levels where the Richardson numbers were able to be evaluated for the different datasets is shown in Table 1.

The dimensionless coefficients in Table 2 are taken from Zilitinkevich and Mironov (1996), with the exception of \( C_n \); they used \( C_n = 0.5 \) based primarily on

<table>
<thead>
<tr>
<th>( C_n )</th>
<th>( C_i )</th>
<th>( C_s )</th>
<th>( C_w )</th>
<th>( C_u )</th>
<th>( C_g )</th>
<th>( C_s )</th>
<th>( C_{UN} )</th>
<th>( R_s )</th>
<th>( R_{ic} )</th>
<th>( R_{ic} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>10</td>
<td>20</td>
<td>1</td>
<td>1.7</td>
<td>0.4</td>
<td>0.75</td>
<td>0.25</td>
<td>0.3</td>
<td>0.3</td>
<td>0.5</td>
</tr>
</tbody>
</table>
large-eddy simulation (LES) results. We choose a value for \( C_w \) near the low range of suggested values that may be more appropriate for the stable boundary layer. The values of \( C_w \) and \( C_v \) are based primarily on LES modeling with consultation from observations, while \( C_w \) is based on atmospheric data and \( C_v \) is a theoretical estimate (see review in Zilitinkevich and Mironov 1996).

3. CASES99 tower data

Data from the CASES99 grassland site in Kansas during October 1999 are used (Poulos et al. 2002). Sonic anemometers were deployed on the National Center for Atmospheric Research (NCAR) 60-m tower and on a smaller tower located 10 m to the side of the main tower. Fluxes are available at 1, 5, 10, 20, 30, 40, 50, and 55 m. The mean temperature profile is from aspirated and shielded NCAR sensors at 15, 25, 35, 45, and 55 m and from thermocouples at 5 and 10 m. The mean wind profiles are from propeller and vane instruments at 15, 25, 35, 45, and 55 m and from sonic anemometers at 5 and 10 m. Surface radiative temperature was estimated using NCAR’s four-component system of pyranometers for shortwave (Kipp and Zonen) and pyrgeometers for longwave (Epply) radiation. Surface temperature is an average from radiometers at six different sites in close proximity to the tower. Flow through the tower and the transition period just after sunset were excluded. A small amount of data was excluded based on the flux nonstationarity parameter following Mahrt (1998). A larger amount of data was discarded based on quality control testing (Vickers and Mahrt 1997). Eddy correlation fluxes were calculated using a height- and stability-dependent variable averaging time, which minimizes the contamination by mesoscale motions on the calculated fluxes (Vickers and Mahrt 2003). For application of the surface flux–based formulations of \( h \), the surface fluxes were estimated by averaging the fluxes at the 1- and 5-m levels. All quantities, including fluxes, were time averaged for 1 h.

Boundary layer depth was determined from profiles of the buoyancy flux with consultation of the momentum flux and turbulence energy profiles. We use a slightly modified version of the stable boundary layer depths used by Mahrt and Vickers (2003). Boundary layer depth is more easily defined in terms of the buoyancy flux for these data because the small stratification above the surface inversion layer often forces the buoyancy flux to small values, even when the turbulence energy and momentum flux do not decrease with height. The downward buoyancy flux sometimes increases with height for part or all of the tower layer because of nearly collapsed turbulence close to the surface and significant turbulence at higher levels with some stratification. The turbulence at higher levels appears to be generated by breaking waves and shear generation associated with a low-level jet that was typically located at two or more tower heights (Banta et al. 2002).

The determination of stable boundary layer height based on 105 profiles yields the following situations: (a) 22% of the time the buoyancy flux decreases with height to small values near zero and remains small, allowing for a relatively clear definition of the boundary layer height in terms of the buoyancy flux (Fig. 1); (b) 15% of the time the buoyancy flux decreases with height over the lower half of the tower layer but then either increases with height or remains significant and independent of height; (c) 18% of the time the buoyancy flux was relatively constant with height, implying a deep boundary layer relative to the tower height; (d) 29% of the time the turbulence generally increases with height; and (e) 16% of the time the fluxes and turbulent energy vary erratically with height, corresponding to layering in a few of the cases. The first category also includes some cases in which the buoyancy flux increased with height in the layer below 10 m, possibly from flux loss by sonic anemometer pathlength averaging. For comparison with formulations, we consider only cases in which the boundary layer height concept seemed most applicable and the boundary layer depth was resolved by the tower measurements (first category above). We note that about one-half of the profiles do not fit the traditional concepts of a boundary layer (categories b, d, and e). A total of 23 estimates of \( h \) on nine different nights was obtained.

4. CBLAST aircraft data

Aircraft data collected during the pilot program of the CBLAST experiment conducted over the Atlantic Ocean south of Martha’s Vineyard, Massachusetts, during July and August of 2001 are used. The NOAA LongEZ (N3R) aircraft measured the three components of the wind, air temperature, and humidity for calculating eddy correlation fluxes of momentum, heat, and moisture (Crawford et al. 2001; Crescenti et al. 2002). We focus on the flight pattern consisting of low-level flux legs at 10-m altitude followed by up- and downslant soundings. In most every case, the vertical structure revealed by the paired up- and downslant soundings was practically identical. We applied offsets to the surface radiative temperature measurements to correct for an apparent ambient temperature dependence in the Everest radiometer that was found after a comparison with in situ sea surface temperature measurements made by the Air–Sea Interaction Meteorology (ASIMET) buoy. Eddy correlation fluxes were calculated using a 2-km spatial averaging window for the low-level 10-m flight segments where aircraft altitude, roll, pitch, and heading fluctuations remained within prescribed limits.

Mahrt et al. (2001) compared eddy correlation fluxes collected by the LongEZ with fluxes from the Naval Postgraduate School buoy (Frederickson and Davidson 2000), a Campbell Scientific, Inc., CSAT3 sonic anemometer at the end of a 570-m pier (U.S. Army Corps of Engineers Field Research Facility, Duck, North Car-
olina), and a buoy (W. Drennan, 1999, personal communication) during the Shoaling Waves Experiment (SHOWEX) in March and November of 1999. The aircraft fluxes agreed well with the fluxes from the other platforms, including qualitative agreement during very stable periods with semicollapsed turbulence where the friction velocities were less than 0.1 m s$^{-1}$.

Our estimate of stable boundary layer depth for this dataset was based on the profile of vertical velocity fluctuations for the periods with downward surface heat flux. Profiles of the fluxes were not available. Vertical velocity is preferred over TKE because the variance of the horizontal wind components can be sensitive to the method used to calculate them. That is, the spectra of the horizontal winds do not always have a well-defined peak.

A preliminary $h$ was assigned where the standard deviation of the vertical velocity ($\sigma_w$) decreased to less than 10% of its value at 10 m. The automated procedure based on the vertical velocity variance sometimes prematurely identified the top of the boundary layer because of a very thin layer with weak or possibly intermittent turbulence, and stronger turbulence above this layer. In these cases, the preliminary estimates for $h$ were adjusted after visually inspecting the profile. In the well-defined cases, a rapid decrease in $\sigma_w$ with height occurred over a thin layer associated with the top of the boundary layer and the turbulence remained weak above the boundary layer (Fig. 2). Sometimes layers of elevated turbulence, which are probably associated with residual boundary layers advected into the study region, complicated the determination of $h$, and such profiles were excluded. A well-defined low-level jet was typically observed near the top of the boundary layer at the same height as the sharp decrease in turbulence level ($\sigma_w$). The height of maximum wind speed was fairly homogeneous in space for a given flight day. Unlike the two tower datasets over land, profiles in CBLAST often showed a temperature inversion at the top of the stable boundary layer.

5. FLOSS tower data

We also analyze data from the FLOSS experiment studying the surface meteorological conditions of snow-covered rangeland in the North Park region of Colorado, near Walden, during the winter of 2002/03. FLOSS is part of the Cold Land Processes Field Experiment (CLPX). The region around the tower was generally snow-covered rangeland with scattered patches of half-meter-high sagebrush protruding above the snow. A 34-m tower was instrumented by NCAR to collect profiles of the mean air temperature, humidity, and wind, as well as profiles of the turbulent fluxes of momentum, heat, and water vapor. Fast-response flux data were collected by seven levels of Campbell Scientific CSAT3 sonic anemometers at 1, 2, 5, 10, 15, 20, and 30 m. The sonic anemometers were also used for the mean wind profile.
The temperature at 5, 10, 15, 20, and 30 m was measured using the same instruments as were used in CASES99. Surface radiative temperature was measured with NCAR’s four-component radiation system. The nocturnal eddy correlation fluxes were calculated using a constant averaging time of 5 min to define the fluctuations. Fluxes and mean quantities were time averaged for 1 h. For the purpose of applying the formulations for $h$, the surface fluxes were taken as the average fluxes at the 2- and 5-m levels. Fluxes at 1 m above ground sometimes appeared inconsistent with those at other levels and are excluded here. Flow through the tower corresponding to a 180° sector centered on the northeast was excluded. Stable boundary layer cases during the transition period near sunset were also excluded.

Stable boundary layer height was determined by examining the buoyancy flux profile in the same manner as described above for CASES99. In contrast to the CASES99 flux profiles, the momentum flux in FLOSS typically decreased with height in a manner consistent with the buoyancy flux. Evaluation of the nocturnal stable boundary layer height based on 525 periods, for which all data were available at all levels, yields the following: (a) 22% of the time the buoyancy flux decreased with height to small values within the tower layer, clearly indicating the height of the stable boundary layer (Fig. 3), or the buoyancy flux decreased linearly with height across three or more of the upper measurement levels, in which case we extrapolated the linear decrease upward to estimate $h$, which is not to exceed two tower heights; (b) 11% of the time the buoyancy flux decreased with height over the lower half of the tower but then increased with height over the upper half; (c) 46% of the time the buoyancy and momentum fluxes were relatively constant with height, indicating a deep boundary layer relative to the height of the tower; (d) 15% of the time the turbulence generally increased with height; and (e) 6% of the time the profiles varied erratically with height. The turbulence ($\sigma_w$) generally increased with height for the majority of the profiles in the deep boundary layer category. All of our estimates of $h$ are derived from profiles in the first category, while the other categories are excluded.

6. Comparisons

We first discuss the degree of skill for each of the different formulations of $h$ for each dataset individually and then for the combined data. This is followed by a discussion of the critical Richardson number and problems associated with the surface flux–based formulations.

For strong surface radiative cooling and shallow boundary layers in CASES99, the Richardson number methods are more correlated to the observed $h$ than the surface flux–based methods, although the Richardson number methods explain only about 18% of the variance of $h$ (Table 3). In terms of variance explained, the 11
different surface flux–based formulations all perform poorly, explaining less than 10% of the variance in CASES99. The two methods that rely most directly on the free-flow stratification [Eqs. (4) and (5)] perform the worst, possibly confirming that $N$ is not relevant for these data. The comparison between the magnitudes of the observed and predicted $h$ in CASES99 (Table 4) tends to confirm the values of the coefficients $C_s = 1$ in Eq. (3), the combination of coefficients in Eq. (11), and $R_s = 0.3$ in the surface bulk Richardson number method. The coefficients or critical values in the remaining forms are apparently not appropriate for these data in that the boundary layer depth is significantly under- or overpredicted depending on the form considered. Variations in the critical Richardson number primarily effect the magnitude of the calculated $h$ and not the variance explained by the formulation.

In FLOSS, where the stability is generally weaker than in CASES99, the formulations for $h$ explain considerably more variance than they do in CASES99. This is perhaps due to larger observation errors and partial

---

**Table 3.** Correlation coefficients between observed and modeled stable boundary layer height for data from CASES99, CBLAST, FLOSS, and all datasets combined. Number of 1-h-averaged samples is 23 for CASES99, 42 for CBLAST, and 107 for FLOSS.

<table>
<thead>
<tr>
<th>Model</th>
<th>CASES99</th>
<th>CBLAST</th>
<th>FLOSS</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_{u^*/f}$</td>
<td>0.21</td>
<td>0.65</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>$C_L$</td>
<td>0.30</td>
<td>0.64</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>$C_{u^*/(-fB)^{0.5}}$</td>
<td>0.26</td>
<td>0.66</td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td>$C_{u^*/(fN)^{0.5}}$</td>
<td>0.06</td>
<td>0.72</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>$C_{u^*/N}$</td>
<td>-0.08</td>
<td>0.67</td>
<td>0.57</td>
<td></td>
</tr>
<tr>
<td>Eq. (6)</td>
<td>0.21</td>
<td>0.72</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>Eq. (7)</td>
<td>0.21</td>
<td>0.72</td>
<td>0.59</td>
<td></td>
</tr>
<tr>
<td>$L[(0.3u^*/fL)/(1+1.9h/L)]$</td>
<td>0.26</td>
<td>0.66</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>$[1/(30L) + f/(0.35u^*)]^{-1}$</td>
<td>0.28</td>
<td>0.66</td>
<td>0.68</td>
<td></td>
</tr>
<tr>
<td>$(h^<em>_wL^</em>)^{0.5}$</td>
<td>0.26</td>
<td>0.66</td>
<td>0.66</td>
<td></td>
</tr>
<tr>
<td>Eq. (11)</td>
<td>0.20</td>
<td>0.70</td>
<td>0.46</td>
<td></td>
</tr>
<tr>
<td>$R_v$</td>
<td>0.42</td>
<td>0.71</td>
<td>0.73</td>
<td></td>
</tr>
<tr>
<td>$R_s$</td>
<td>0.38</td>
<td>0.50</td>
<td>0.58</td>
<td></td>
</tr>
<tr>
<td>$R_i$</td>
<td>0.42</td>
<td>0.62</td>
<td>0.46</td>
<td></td>
</tr>
</tbody>
</table>

**Table 4.** Percent bias in the modeled stable boundary layer height. Values are $100(h_{mod} - h_{obs})/h_{obs}$, where $h_{mod}$ is the model prediction and $h_{obs}$ is the observed height.

<table>
<thead>
<tr>
<th>Model</th>
<th>CASES99</th>
<th>CBLAST</th>
<th>FLOSS</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_{u^*/f}$</td>
<td>128</td>
<td>153</td>
<td>619</td>
<td>440</td>
</tr>
<tr>
<td>$C_L$</td>
<td>-52</td>
<td>472</td>
<td>246</td>
<td>261</td>
</tr>
<tr>
<td>$C_{u^*/(-fB)^{0.5}}$</td>
<td>3</td>
<td>297</td>
<td>391</td>
<td>316</td>
</tr>
<tr>
<td>$C_{u^*/(fN)^{0.5}}$</td>
<td>95</td>
<td>474</td>
<td>646</td>
<td>332</td>
</tr>
<tr>
<td>$C_{u^*/N}$</td>
<td>18</td>
<td>39</td>
<td>212</td>
<td>144</td>
</tr>
<tr>
<td>Eq. (6)</td>
<td>-68</td>
<td>-7</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>Eq. (7)</td>
<td>-78</td>
<td>-42</td>
<td>-4</td>
<td>-23</td>
</tr>
<tr>
<td>$L[(0.3u^*/fL)/(1+1.9h/L)]$</td>
<td>-60</td>
<td>34</td>
<td>86</td>
<td>54</td>
</tr>
<tr>
<td>$[1/(30L) + f/(0.35u^*)]^{-1}$</td>
<td>20</td>
<td>463</td>
<td>617</td>
<td>500</td>
</tr>
<tr>
<td>$(h^<em>_wL^</em>)^{0.5}$</td>
<td>-35</td>
<td>155</td>
<td>211</td>
<td>160</td>
</tr>
<tr>
<td>Eq. (11)</td>
<td>-3</td>
<td>76</td>
<td>267</td>
<td>183</td>
</tr>
<tr>
<td>$R_v$</td>
<td>2</td>
<td>168</td>
<td>154</td>
<td>137</td>
</tr>
<tr>
<td>$R_s$</td>
<td>-26</td>
<td>11</td>
<td>-8</td>
<td>-6</td>
</tr>
<tr>
<td>$R_i$</td>
<td>75</td>
<td>-17</td>
<td>-29</td>
<td>-12</td>
</tr>
</tbody>
</table>
decoupling (Mahrt and Vickers 2002) in the stronger stability conditions found in CASES99. For FLOSS, all of the models perform about equally well in terms of correlation, and the best models explain about one-half of the variance. However, with the exception of Eq. (7), the bulk Richardson number $R_b$, and the gradient Richardson number $R_s$, the formulations grossly overpredict the magnitude of $h$, indicating that the coefficients are not suitable. For example, the commonly used form given by Eq. (3) with $C_a = 1$ overpredicts $h$ in FLOSS by nearly 400%. Historically, the coefficients in some formulations of $h$ have been calibrated according to the depth of the surface inversion layer, which may be substantially thicker than the depth based on the vertical structure of the fluxes and turbulence.

For CBLAST, where the stability is weaker and the boundary layers are deeper in comparison with the two land-based datasets, the local gradient Richardson number $R_s$ is the best predictor in terms of variance explained. The local gradient approach is optimal here because it captures the location of minimum wind shear associated with the low-level jet. Adding the free-flow stratification to the formulation does not improve model performance [cf. Eqs. (1), (4) and (5) in Table 3] despite the fact that these marine boundary layers are typically capped by an inversion layer. All of the models but three [Eqs. (6) and (7), and $R_s$] overpredict $h$ for CBLAST.

An additional test of model performance is how well they perform during periods when the fluxes are relatively constant with height, or when the turbulence energy generally increased with height. These two conditions often coincided. These are the periods with deeper boundary layers when $h$ exceeds our measurement height. Most of the models, especially Eq. (7), perform poorly in these situations for CASES99 in that they frequently predict a shallow boundary layer when the observations indicate a deep boundary layer (Table 5). The models generally perform better in recognizing deeper stable boundary layers for FLOSS than for CASES99.

A more rigorous test of the different formulations for $h$ is their overall performance for a variety of stable boundary layers found by combining the three datasets (e.g., Fig. 4). The surface bulk Richardson number performs the best in terms of variance explained and on average predicts the correct magnitude of $h$ in CASES99, but it overpredicts the magnitude for CBLAST and FLOSS by more than 100% (Table 4). The non-uniform behavior of the above formulations suggests that additional physics must be important, which is discussed further below.

### a. Critical Richardson number

The critical surface bulk Richardson number can be evaluated directly from the observations as

$$R_s = h \left( \frac{g}{\theta} \right) \left( \frac{[\theta(h) - \theta_i]}{U(h)^2} \right),$$

(15)

where $h$ is the observed stable boundary layer height. There is considerable scatter in the critical value within datasets, indicating that the critical value may fluctuate depending on other factors not included in the formulation, even at the same location. More significant, the critical Richardson number varies systematically between datasets (Fig. 5). This indicates that the surface bulk Richardson number approach with a constant critical value is not applicable to all sites.

The systematic variations in $R_s$ between sites could be due to differences in the surface aerodynamic roughness length ($z_o$), which varies by three orders of magnitude between the sites. One could argue that for a given bulk stratification and wind shear, deeper stable boundary layers develop over rougher surfaces. Zilitinkevich and Baklanov (2002) suggested that the critical Richardson number should increase with increasing roughness, or decrease with increasing surface Rossby number,

$$R_s = \frac{U_{10}}{f z_o},$$

(16)

although they could not test the roughness length dependence using data from only one site. The Coriolis parameter is included in Eq. (16) only for dimensional reasons, and there is no evidence that the depth of the stable boundary layer varies with $f$. Two other dimensionless combinations of parameters that one could consider that do not include the Coriolis parameter are $U_{10}/(u_{*}z_o)$ or $h/z_o$; however, both of these forms require an iterative method because they contain $h$.

Using the Monin–Obukhov similarity theory and the observed fluxes for all of the data (not just the stable boundary layer cases), we calculate the average roughness lengths to be $2.7 \times 10^{-2}$, $4.0 \times 10^{-3}$, and $2.0 \times 10^{-4}$ m for CASES99, FLOSS, and CBLAST, respectively. The roughness length is actually time dependent for FLOSS because of changing snow cover and snow.
FIG. 4. Predicted $h$ vs observed $h$ for stable boundary layer depth formulations described in the text with parameters in Table 2 for the CASES99 (circles), CBLAST (asterisk), and FLOSS (plus) datasets. The 1:1 line is dashed.

Fig. 4. Predicted $h$ vs observed $h$ for stable boundary layer depth formulations described in the text with parameters in Table 2 for the CASES99 (circles), CBLAST (asterisk), and FLOSS (plus) datasets. The 1:1 line is dashed.

age, and it varies from $7 \times 10^{-4}$ to $1 \times 10^{-2}$ m, depending on the time period considered. The roughness length is also time dependent for the marine case because of varying wind–wave–swell interactions, although the variation found in $z_o$ was not significant for the periods studied here and we use one constant value of $z_o$ for the CBLAST data. Because of flux-sampling errors and potential problems with applying the Monin–Obukhov similarity theory in some situations, the uncertainty in the roughness length estimates may be large.

The critical surface bulk Richardson number $R_s$ does indeed decrease with increasing surface Rossby number (Fig. 6). The two CASES99 points and the two CBLAST points in Fig. 6 denote partitioning each dataset into two surface Rossby number categories with an equal number of observations in each. We partition the FLOSS data, which have more observations and a time-dependent roughness length, into four categories based on the value of the surface Rossby number. The wide separation in Rossby number space for the three datasets is primarily due to the large roughness length differences between the sites. For a given roughness length, a stronger wind speed and, thus, a larger $R_o$, is associated with a smaller critical Richardson number as is demonstrated by the CASES99 and CBLAST data. This was noted in section 2 where the shear term tends to be overestimated in strong winds.

Recalculating $h$ using a fit of the critical Richardson number to the surface Rossby number,

$$ R_i = 0.16(10^{-7}R_o)^{-0.18}, $$

instead of a constant value of $R_i = 0.3$, reduces the bias in the modeled $h$ from 154% to 25% for FLOSS and from 168% to 7% for CBLAST. Because of the inadequacy of the power curve fit at small $R_o$, the bias actually increases to $-33\%$ for CASES99. The correlation
coefficient between the observed and modeled $h$ is unchanged using the Rossby number–dependent critical value in comparison with using a constant critical value. We conclude that the surface bulk Richardson number approach for $h$ is improved by specifying a Rossby number–dependent critical value.

The critical bulk Richardson number based on two atmospheric levels can be evaluated as

$$R_c^5 = \left( h - z_i \right) \frac{g}{\theta} \left[ \frac{\theta(h) - \theta_i}{\left| U(h) - U_i \right|} \right]^2,$$

(18)

where $z_i$ is fixed at 10 m. The critical value $R_c$ is less variable between datasets when compared with the critical value for the surface bulk Richardson number (Fig. 7). This is reflected in the small bias error for $h$ using this approach with a constant critical value (Table 4). However, there is significantly larger scatter in $R_c$ at each site in comparison with $R_{bs}$, especially for the marine CBLAST data. This is also indicated by the small correlations for the $R_b$ method as compared with the $R_{bs}$ method (Table 3). The differences in the best-fit critical value of $R_c$ between datasets are not consistent with the surface roughness influence. We conclude that if a constant critical value is used, the bulk Richardson number method based on two atmospheric levels is superior to the surface bulk Richardson number method. However, the better approach is to use the surface bulk Richardson number method with a Rossby number–dependent critical value.

The prediction for $h$ from the gradient Richardson number approach [Eq. (14)] applied to CBLAST is not especially sensitive to small changes in the critical value. At the jet core the mean wind shear approaches zero, forcing the gradient Richardson number to very large values. The optimal critical value of $R_{ leveled}$ based on CBLAST data is 0.35.

b. Surface flux formulations

The surface flux–based formulations have only limited success in predicting the variability and magnitude of $h$. The observed values of the coefficients calculated from the observed $h$ for the one-parameter formulations [Eqs. (1)–(5)] display considerable variation as demonstrated in Fig. 8 where the ordinate spans three orders of magnitude. Large variation in the coefficients indicate a problem with the approach. For example, for $z/L$ near 0.1, where FLOSS and CBLAST data overlap in $z/L$ space, the clear and large difference in $C_n$ cannot be
attributed to stability or to the difference in $f$, which varies only by about 1%. Here, $C_s$ is a strong function of $z/L$, which demonstrates the failure of the $h = C_s L$ approach. The steady increase of $C_s$ with increasing $z/L$ is partly due to self-correlation. The large variation in the coefficients across different datasets indicates that these surface flux-based approaches may not be universal or may not include all of the important physics. It is not clear how one could optimally calculate the coefficients for Eqs. (6) or (11), which have three free parameters, or for Eq. (7), which has five free parameters.

The coefficients calculated from the observed boundary layer depths for the data considered here (Table 6) are smaller than the reference coefficients in the literature (Table 2). The mean values for the combined data in Table 6 were calculated by averaging the mean values for the individual datasets, such that each dataset receives equal weight. The small magnitude of the coefficients found here could relate to the difference between stable boundary layer height estimates based on the depth of the surface-based mixing, as attempted here, and previous work based primarily on the depth of the temperature inversion layer or some characteristic of the mean temperature or wind profile based on models. Our comparisons of observed and modeled $h$ based on tower data are biased toward the shallowest boundary layer depth cases because of the limited height range of the measurements.

**Table 6. Average (std dev) of the dimensionless coefficients in the surface flux-based formulations derived from data.**

<table>
<thead>
<tr>
<th>Coef</th>
<th>CASES99</th>
<th>CBLAST</th>
<th>FLOSS</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_n$</td>
<td>0.06 (0.04)</td>
<td>0.04 (0.01)</td>
<td>0.02 (0.004)</td>
<td>0.04</td>
</tr>
<tr>
<td>$C_s$</td>
<td>16 (15)</td>
<td>0.9 (0.4)</td>
<td>1.4 (0.6)</td>
<td>6</td>
</tr>
<tr>
<td>$C_{sr}$</td>
<td>1.5 (1.2)</td>
<td>0.3 (0.09)</td>
<td>0.2 (0.07)</td>
<td>0.7</td>
</tr>
<tr>
<td>$C_{ir}$</td>
<td>1.2 (0.8)</td>
<td>0.8 (0.2)</td>
<td>0.3 (0.09)</td>
<td>0.8</td>
</tr>
<tr>
<td>$C_i$</td>
<td>24 (16)</td>
<td>15 (4)</td>
<td>7 (2)</td>
<td>15</td>
</tr>
</tbody>
</table>
7. Summary

The stable boundary layer height is determined from measured vertical profiles of the buoyancy flux for a tower dataset over grassland in autumn and a tower dataset over rangeland with variable snow cover in winter, and from the turbulence energy profile for a summertime aircraft dataset over a cooler ocean surface. For the nocturnal tower datasets over land, a well-defined stable boundary layer height within the height range of the measurements was determined for only 22% of the available data. No estimate of the boundary layer depth was made for the remaining periods when the buoyancy flux either (a) was relatively constant with height, indicating a deep boundary layer relative to the height of the tower, (b) increased with height, indicating an elevated source of turbulence, or (c) varied erratically with height, sometimes due to layered turbulence.

The cases of well-defined boundary layer depth were compared with traditional predictions of the equilibrium boundary layer depth based on the Richardson number and a variety of approaches based on surface fluxes. In general, the existing formulations perform poorly and often grossly overestimate the depth of the boundary layer. The latter may be due to previous verification in terms of the depth of the surface inversion, which can be substantially deeper than the layer of turbulence. Also recall that the cases of deep boundary layers are eliminated from two of the three datasets because of finite tower height. Attempting to readjust the values of the coefficients in the boundary layer depth formulations to reduce biases sometimes improves performance, but major variance remains unexplained.

Approaches based on the bulk and gradient Richardson numbers best approximate the observed boundary layer depths. Further improvements are obtained for the surface bulk Richardson number method by relating the required critical value to the aerodynamic roughness length and wind speed. This is carried out here in terms of the surface Rossby number, even though a dependence on the Coriolis parameter is not anticipated. More physically motivated formulations are implicit and require iteration and are, therefore, excluded from this study. Prognostic models of the boundary layer depth were also excluded from this study because errors in the estimation of the rate of change of the boundary layer depth can be large. We have avoided more complicated estimates of the boundary layer depth, because the frequent failure of formation of a definable boundary layer is a more serious problem. The above study also indicates that using data from only a single site can provide an incomplete picture.

Acknowledgments. We express our gratitude to Tim Crawford for his dedication in collecting the aircraft data in the CBLAST Weak-Wind Experiment. The helpful comments of the reviewers are greatly appreciated. This material is based upon work supported by Grant DAAD19-9910249 from the U.S. Army Research Office, Grant 0107617-ATM from the Physical Meteorology Program of the National Sciences Foundation, Grant NAGS-11586 from the NASA Hydrology Program, and Grant N00014-01-1-0084 from the Marine Meteorology Program of the Office of Naval Research.


