Objectively Determined Fair-Weather NBL Features in ARW-WRF and Their Comparison to CASES-97 Observations

MARGARET A. LEMONE, MUKUL TEWARI, FEI CHEN, AND JIMY DUDHIA

National Center for Atmospheric Research, Boulder, Colorado

(Manuscript received 11 November 2013, in final form 27 March 2014)

ABSTRACT

Heights of nocturnal boundary layer (NBL) features are determined using vertical profiles from the Advanced Research Weather Research and Forecasting Model (ARW-WRF), and then compared to data for three moderately windy fair-weather nights during the April–May 1997 Kansas-based Cooperative Atmosphere–Surface Exchange Study (CASES-97) to evaluate the success of four PBL schemes in replicating observations. The schemes are Bougeault–LaCarrere (BouLac), Mellor–Yamada–Janji/C19 (MYJ), quasi-normal scale elimination (QNSE), and Yonsei University (YSU) versions 3.2 and 3.4.1. This study’s chosen objectively determined model NBL height \( h \) estimate uses a turbulence kinetic energy (TKE) threshold equal to 5% TKE_{max}, where TKE_{max} is relative to its background (free atmosphere) value. The YSU- and MYJ-determined \( h \) could not be improved upon. Observed heights of the virtual temperature maximum \( h_{T_{v_{max}}} \) and wind speed maximum \( h_{S_{max}} \), and the heights \( h_{1_{wsonde}} \) and \( h_{2_{wsonde}} \), between which the radiosonde slows from turbulent to nonturbulent air, and thus brackets \( h \), were used for comparison to model results. The observations revealed a general pattern: \( h_{T_{v_{max}}} \) increased through the night, and \( h_{T_{v_{max}}} \) and \( h_{S_{max}} \) converged with time, and the two mostly lay between \( h_{1_{wsonde}} \) and \( h_{2_{wsonde}} \) after several hours. Clear failure to adhere to this pattern and large excursions from observations or other PBL schemes revealed excess mixing for BouLac and YSU version 3.2 (but not version 3.4.1) and excess thermal mixing for QNSE under windy conditions. Observed friction velocity \( u^* \) was much smaller than model values, with differences consistent with the observations reflecting local skin drag and the model reflecting regional form drag + skin drag.

1. Introduction

The objectives of this paper are to evaluate the heights of significant nocturnal boundary layer (NBL) features using observed and model profiles and then use them to evaluate four planetary boundary layer (PBL) schemes in high-resolution (1-km innermost grid) simulations using the Advanced Research version of the Weather Research and Forecasting Model (ARW-WRF, hereafter referred to as WRF). As in its companion paper, LeMone et al. (2013), which examined the convective boundary layer (CBL), we focus on the Yonsei University (YSU; Hong et al. 2006), Mellor–Yamada–Janjić (MYJ; Janjić 2001), and Bougeault–LaCarrere (BouLac; Bougeault and LaCarrere 1989), and quasi-normal scale elimination (QNSE; Sukoriansky and Galperin 2008; Sukoriansky et al. 2005) and use data from three fair-weather nights with mostly moderate winds during the April–May 1997 Kansas-based Cooperative Atmosphere–Surface Exchange Study (CASES-97; LeMone et al. 2000) field program. After identifying a good criterion for objectively determining NBL depth \( h \), we explore PBL-scheme strengths and shortcomings by comparing the behavior and magnitudes of \( h \), \( h_{T_{v_{max}}} \), and \( h_{S_{max}} \), respectively the heights of the virtual temperature \( T_v \) and wind speed \( S \) maxima, to observations.

While the CBL has a well-defined mixed-layer top, the NBL \( h \) can be difficult to identify. Numerous methods have been tried (Table 1). Diagnosis of \( h \) using mean or instantaneous vertical profiles has had mixed success. Using Doppler lidar data, Pichugina and Banta (2010) found a strong correspondence of \( h_{S_{max}} \) to \( h \) as defined by profiles of the variance of the radial velocity for a subset of wind profiles (one maximum, wind in lowest 200 m greater than 5 m s\(^{-1}\)), with minimum curvature in the \( S \) profiles...
yielding even better results. However, in a large-eddy simulation (LES) of a weakly stable NBL (Obukhov length $L > 100$ m) by Kosovic and Curry (2000), $h_{Smax}$ and $h_{Tvmax}$ coincided with $h$ only under steady-state conditions, after about one inertial period of simulation. Similarly, observations show that $h_{Tvmax}$ does not necessarily coincide with $h_{Smax}$ or $h$ in moderately to very stable conditions (e.g., Figs. 6 and 7 of Mahrt and Vickers 2006). Indeed, as illustrated in Banta et al. (2007) and Sun et al. (2004) and elsewhere, the NBL often has a complex structure that varies with time.

Bulk Richardson numbers (e.g., Vogelezang and Holtslag 1996) and more complex formulations (e.g., Vickers and Mahrt 2004; Steeneveld et al. 2007) have also been used, with varying degrees of success. A significant shortcoming of such approaches is that radiosonde data need to be smoothed for reliable estimates. Also, magnitudes of criteria using vertical gradients, including Richardson numbers, tend to vary with the vertical spacing used.

When turbulence data are available, the height at which a second-moment variable decreases to a specific fraction of its surface or near-surface maximum provides a useful estimate of $h$. Examples of such parameters are buoyancy flux (Caughey et al. 1979), vertical velocity variance (Vickers and Mahrt 2004), vertical flux of the component of the horizontal momentum along the surface wind direction (Kosovic and Curry 2000), and the turbulence kinetic energy (TKE; Lenschow et al. 1988). Fortunately, such parameters tend to be internally consistent, at least for weakly stable NBLs [e.g., see LES of Kosovic and Curry (2000), Basu and Porte-Agel (2006), and simulations summarized in Beare et al. (2006); compare to http://gabls.metoffice.com/variance_625.html and follow the menu for profiles of fluxes and means].

Since TKE profiles are available from WRF runs, we use them to determine subjective $h$ ($h_{subj}$) for BouLac, MYJ, and QNSE and use the profiles and resulting $h_{subj}$ to judge the model $h$ metrics in Table 1. For YSU, we use the eddy exchange coefficient $K$, noting its relationship to TKE (Shin et al. 2013). The LES-generated TKE profiles of Kosovic and Curry (2000) and Basu and

### Table 1. Selected potential criteria for NBL depth $h$ from vertical profiles.

<table>
<thead>
<tr>
<th>Observations</th>
<th>Model</th>
<th>Comments/source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\partial T/\partial z = 0$</td>
<td>$\partial \Theta_v/\partial z = 10 , \text{K km}^{-1}$</td>
<td>Both nearly equivalent to $\partial T/\partial z = 0$ (Yamada 1979)</td>
</tr>
<tr>
<td>$\partial S/\partial z = 0$ and minimum curvature</td>
<td>$\partial S/\partial z = 0$</td>
<td>Minimum curvature found subjectively (Melgarejo and Deardorff 1974)</td>
</tr>
<tr>
<td>$R_{loc} \sim 1.0$</td>
<td>$R_{loc} = 0.5$</td>
<td>Values based on trial and error for subset of soundings (Kosovic and Curry 2000)</td>
</tr>
<tr>
<td>Drop in radiosonde rise rate</td>
<td>$\text{Ri(level 1–h)} = \text{constant}$</td>
<td>Result of TKE dropoff at top of NBL (Johansson and Bergstrom 2005)</td>
</tr>
<tr>
<td></td>
<td>TKE = $0.101, 0.2 \text{ m}^2 \text{s}^{-2}$</td>
<td>Constant = 0.2 based on subset of cases (Vogelzang and Holtslag 1996)</td>
</tr>
<tr>
<td></td>
<td>$f/TKE_{max} + \text{TKE}_b$</td>
<td>Modifications of MYJ $h$ criterion (YSU; Hong et al. 2006)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Figure 1 plus comparison of TKE profiles to $K$ and flux profiles for LES and PBL schemes (LeMone et al. 2013)</td>
</tr>
</tbody>
</table>

**Fig. 1.** Steady-state LES TKE profiles for the weakly stable ($L > 100$ m) Arctic SBL. Black refers to Kosovic and Curry (2000), red refers to Basu and Porte-Agel (2006), and SGS is the subgrid scale.
Porte-Agel (2006) in Fig. 1 illustrate what we can expect for weakly stable conditions. The two profiles in the figure have weakly concave-up, almost linear shapes, with a maximum at the lowest grid level (10 m). The authors found $h$ by dividing by 0.95 the height at which $u'I_p w' = 0.05(u'I_p w')_{SFC}$, where the wind component $u_p$ is parallel to the surface wind, $w$ is the vertical wind, the overbars indicate an average, and the primes indicate a deviation from that average. In the figure, their $h$ is roughly the height at which $TKE(h) = 0.045TKE(10 \text{ m})$. However, we recognize that the TKE profile can depart significantly from this ideal. For example, Pichugina and

![FIG. 2. CASES-97 observational array. Numbers indicate surface flux sites. At the vertices of the triangle lie 915-MHz RWP/MS sites BEA (elevation 478 m), OXF (360 m), and WHI (430 m), with collocated radiosonde releases. Solid lines indicate flight tracks. Terrain contour interval is 20 m.](image)

![FIG. 3. Illustration of how $h_{\text{Smax}}$, $h_{\text{Tmax}}$, $h_{\text{Riloc}}$, $h_{\text{1wsonde}}$, and $h_{\text{2wsonde}}$ are estimated from radiosonde data.](image)
**Table 2.** Conditions for days examined (time in UTC). B is Beaumont (open grassland), O is Oxford (some trees), and W is Whitewater (grassland). Figure 2 shows site locations. Italics indicate data that are from the radiosonde.

<table>
<thead>
<tr>
<th>Date</th>
<th>Parameter (avg of four 30-min avg)</th>
<th>Sites 1 + 2 (grassland, near BEA)</th>
<th>Sites 5, 6, and 7 (wheat)</th>
<th>Sparse veg (V) + grass (G), sites 3 (V) 4 (V/G) + 8 (G)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Direction/speed (m s⁻¹)</td>
<td>0200</td>
<td>0930</td>
<td></td>
</tr>
<tr>
<td>4–5 May</td>
<td></td>
<td>194/6.6</td>
<td>193/9.1</td>
<td>185/3.6</td>
</tr>
<tr>
<td>B</td>
<td></td>
<td>190/6.6</td>
<td>190/6.6</td>
<td></td>
</tr>
<tr>
<td>O</td>
<td></td>
<td>189/3.6</td>
<td>189/6.1</td>
<td>119/6.6</td>
</tr>
<tr>
<td>W</td>
<td></td>
<td>189/3.6</td>
<td>189/6.1</td>
<td>119/6.6</td>
</tr>
<tr>
<td>10–11 May</td>
<td></td>
<td>210/4.4</td>
<td>213/6.5</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
<td>176/1.8</td>
<td>198/2.5</td>
<td>+9</td>
</tr>
<tr>
<td>O</td>
<td></td>
<td>176/1.8</td>
<td>198/2.5</td>
<td></td>
</tr>
<tr>
<td>W</td>
<td></td>
<td>176/1.8</td>
<td>198/2.5</td>
<td></td>
</tr>
<tr>
<td>20–21 May</td>
<td></td>
<td>82/4.5</td>
<td>72/2.2</td>
<td>+8</td>
</tr>
<tr>
<td>B</td>
<td></td>
<td>85/2.6</td>
<td>69/2/6</td>
<td></td>
</tr>
<tr>
<td>O</td>
<td></td>
<td>85/2.6</td>
<td>69/2/6</td>
<td></td>
</tr>
<tr>
<td>W</td>
<td></td>
<td>85/2.6</td>
<td>69/2/6</td>
<td></td>
</tr>
</tbody>
</table>

- **Note:** 48-m wind direction/speed (°m s⁻¹) minisodar (sonde).
- **Includes stations 1–8, unless otherwise indicated.
- **Does not include site 8.

---

**Figure 2** shows site locations.
Whitewater data totally absent on 29 April, of poor quality on 5 May, good for 10–11 and 20–21 May, and Oxford winds were of marginal quality. Radiosonde data were missing after 0500 UTC at Whitewater on 11 May (see Fig. 7).

b. Determination of NBL depth from observations

Figure 3 illustrates how the heights of the NBL features listed in Table 1 are identified subjectively from observations. Temperature $T$ profiles were smooth enough to make $h_{\text{T, max}}$ easy to determine, with an uncertainty of the order of the radiosonde data spacing, 30–50 m, while scatter in the wind profile can increase uncertainty for $h_{\text{S, max}}$ to $\sim 50$ m. Heights at which the increase of wind with height suddenly slows down without reaching a maximum [a rough correspondence to Pichugina and Banta (2010)’s minimum curvature in the speed profile] were also documented, but this situation was rare (see Fig. 7). Evaluation of $h_{\text{S, max}}$ from the blended RWP + MS data was sometimes complicated by a jump in speed in the transition from RWP to MS measurements at about 150 m, likely a result of ground clutter contamination of the RWP data. Vertical spacing is 60 m for the RWP data at Whitewater and Beaumont and 5 m for the MS data. Scatter and coarse data spacing in the radiosonde data made Richardson numbers unreliable on most days.

In the final technique, it is based on the rapid decrease in sonde vertical velocity $w_{\text{sonde}}$ as the balloon travels from turbulent to nonturbulent air (Johansson and Bergstrom 2005), which results from an increase in drag on the balloon once it enters laminar flow (MacCready 1965; Gallice et al. 2011; Wang et al. 2009). This method is appealing because it directly relates to our TKE-based NBL definition. In contrast to the one PBL depth chosen by Johannson and Bergstrom, we used two depths, $h_{\text{w, sonde}}$ and $h_{2\text{w, sonde}}$, to identify the lower and upper limits of the height interval through which $w_{\text{sonde}}$ falls from its “turbulent” to “nonturbulent” value. A low bias for $h_{\text{w, sonde}}$ of up to 30 m results from the fact that the balloon responds to the turbulence, while the sonde, which collects the data, is attached to the balloon by a 30-m string. The angle of the string to the vertical is unknown, so we do not correct for this bias.

It was not possible to estimate $h$ from $w_{\text{sonde}}$ if the NBL was much less than $\sim 100$ m deep. It takes time for the sonde to unreel from the balloon and for the balloon–sonde system to accelerate to a typical speed of $\sim 5$ m s$^{-1}$, and there were typically only 3–4 points below 100 m at the 10-s data rate. Given the general association of deeper NBLs with larger $u_{\text{*_k}}$ (e.g., Caughey et al. 1979; Steeneveld et al. 2007), we limited our analysis to windier nights. Even so, the balloon rise rate method did not always work, because balloon inflation was not optimum or vertical air motions modified balloon rise rate.

c. NBL classification

The data used for this study were gathered in rolling terrain (Fig. 2) with varying land cover (Table 2); both cause wind and turbulence to vary horizontally. Indeed, according to Acevedo and Fitzjarrald (2001), Fiebrich and Crawford (2001), Van de Wiel et al. (2002), and others, the turbulent near-surface flow sometimes detaches from the surface, especially in lower-lying areas. With this in mind, we characterize the NBL in a regional sense.
TABLE 4. Characteristics of PBL schemes for stable conditions. N is the Brunt–Väisälä frequency. TKE units are m$^2$ s$^{-2}$.

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Basic physics</th>
<th>Default $h$ criterion</th>
<th>Background TKE (m$^2$ s$^{-2}$)</th>
<th>Primary References</th>
</tr>
</thead>
<tbody>
<tr>
<td>BouLac Level 1.5 scheme</td>
<td>Solve for $e$ using simplified form of (1): $K_H = K_M = 0.4L_{\text{mix}}e^{1.5}$; $L_{\text{mix}}$ based on how far parcel can travel vertically with initial vertical velocity = $(2e)^{0.5}$</td>
<td>None for NBL</td>
<td>0.0001</td>
<td>Bougeault and LaCarrere (1989) Therry and LaCarrere (1983)</td>
</tr>
<tr>
<td>MYJ Level 2.5 scheme</td>
<td>Solves for $e, \theta$; (latter equation simplified)→2.5; $K_{H,M} = f_{H,M}L_{\text{mix}}e^{0.5}$; $f_{H,M} = f_{H,M} e$; $L_{\text{mix}} = L_{\text{mix},0}$; $L_{\text{mix}},0 = \alpha \left(\int_0^{h_{\text{bl}}} (q, dz)\right)/(\int_0^{h_{\text{bl}}}(q, dz))$; Preliminary</td>
<td>$h_{\text{MYJ}}$, lowest level at which $e &lt; 0.101$ m$^2$ s$^{-2}$ (not interpolated)</td>
<td>0.1000</td>
<td>Janjic (2001) Mellor and Yamada (1982) Mellor and Yamada (1974)</td>
</tr>
<tr>
<td>ONSE In (1) vertical diffusion as in MYJ,</td>
<td>shear and buoyancy terms use $K_{H,M}$ + vertical gradients as in MYJ,</td>
<td>$h_{\text{ONSE}}$, lowest level at which $e &lt; 0.0050$ m$^2$ s$^{-2}$ (not interpolated)</td>
<td>0.0050</td>
<td>Sukoriansky et al. (2006) Detering and Etling (1985) Sukoriansky and Galperin (2008)</td>
</tr>
<tr>
<td>YSU</td>
<td>$K_{H,M} = w_v k z(1 - z/h_{\text{YSU}})^2$; $w_v = u^* k_z$ for version 3.2; $w_v = u^*/(1 + 5z/L)$ for version 3.4.1; $L_{\text{mix}} = L_{\text{mix},0} k z/(k z + L_{\text{mix},0})$, $L_{\text{mix},0} = 30$ m</td>
<td>$R_i(h_{\text{YSU}}) = 0.25$ (interpolated), from (6) with $z_1 = z h$, $U_1 = U = 0$</td>
<td>N/A</td>
<td>Hong et al. (2006) Noh et al. (2003) Louis (1979) Betts et al. (1996)</td>
</tr>
</tbody>
</table>

According to LeMone et al. (2003), NBLs regionally vary from being continuously turbulent and fully coupled to the surface, with air trajectories following the terrain along the synoptic wind direction, to having only weak turbulence driven by drainage winds. In the former case, the 2-m $T$ changes with the elevation, $T_{2m,el} \sim -9.8$ K km$^{-1}$, following the adiabatic lapse rate, while $T_{2m,el} > +40$ K km$^{-1}$ for the latter case (their Fig. 5), with a magnitude that increases with the vertical $T$ gradient; low-lying locations where radiative cooling is not offset by downslope winds or turbulent mixing also increase $T_{2m,el}$. From Table 2, $T_{2m,el} = -11$ K km$^{-1}$ (based on a least squares straight line for $T_{2m}$ as a function of elevation for sites 1–8), and the wind direction varies little spatially, indicating that the synoptic flow on 4–5 May is continuously coupled to the surface at 0930 UTC. On the other hand, intermediate $T_{2m,el}$ and more variation in wind direction on the nights of 10–11 May and 20–21 May suggest some influence by drainage flow, with possible occasional decoupling and associated cooling, especially for the low-lying stations.

A similar picture emerges from the local classification scheme of Van de Wiel et al. (2003), who use net radiation $R_{\text{net}}$ and $u_{*}^2/\theta_h$, where the friction velocity $u_{*} = [(u'w')_{dc}^2 + (v'w')_{dc}^2]^{1/4}$ and $u'$ and $v'$ are the horizontal wind components, to identify three regimes: “continuous turbulent,” “intermittent,” and “radiative.”
To identify the regimes for CASES-97, we plotted on their Fig. 8 our best estimates of $h$ along with three sets of flux averages (for sites 1–8, the two highest grassland sites 1 and 2, and the two floodplain sites 3 and 6). Assuming that the regimes in the figure apply (reasonable since CASES-97 was in the same location), we found that 4–5 May falls in the continuous turbulence regime for all three averages at 0930 UTC and a mix of intermittent and continuous turbulence at 0200 UTC. The other 2 days fall mostly in the intermittent regime, consistent with the evaluation based on $T_{2m,e}$. The fourth night, 28–29 April (not shown), falls in the radiative regime for all three

![Stable Comparison of various PBLH criteria -- MYJ for BEA 4-5 May 1997](image)

**Fig. 5.** For MYJ at Beaumont on 5 May 1997, the evaluation of $h$ criteria based on TKE profiles (shifted 2 m$^2$ s$^{-2}$ each 2 h); “1M” in upper left of the first panel indicates that criterion 1 failed to identify $h$. Sunset was around 0130 UTC (1930 CST).

![Evaluation of eight potential $h$ criteria based on comparisons to series of TKE profiles like those in Fig. 5 for nights of 4–5, 10–11, and 20–21 May. Labels refer to dates in UTC and M signifies May](image)

**Fig. 6.** For Beaumont, evaluation of eight potential $h$ criteria based on comparisons to series of TKE profiles like those in Fig. 5 for nights of 4–5, 10–11, and 20–21 May. Labels refer to dates in UTC and M signifies May.
averages and thus was considered too stable to be included here.

Following Van de Wiel et al. (2012) and Sun et al. (2012), classifying the days according to whether the wind speed at a given height is capable of sustaining turbulence beneath also reveals a similar picture. Based on Cabauw data, Van de Wiel et al. found that continuous turbulence is maintained when the wind at 40 m exceeds $5 \text{ m s}^{-1}$, with the threshold increasing with $|R_{	ext{net}}|$. Using data from the 55-m CASES-99 tower, Sun et al. found that threshold speeds increase with height, with values of $7 \text{ m s}^{-1}$ at 40 m and $8 \text{ m s}^{-1}$ at 50 m. The CASES-99 thresholds are larger than for Cabauw at least partially because of larger $|R_{	ext{net}}|$ [cf. Fig. 4 of Van de Wiel et al. (2012) for Cabauw to Table 2 of Van de Wiel et al. (2003) for CASES-99]. Since $|R_{	ext{net}}|$ in our Table 3 is close to that during CASES-99, 7.5–8 m s$^{-1}$ is a good threshold speed at 48 m for CASES-97 as well as CASES-99. Based on this criterion, turbulence below 48 m can be sustained on 4–5 May, at Beaumont and Whitewater at 0930 UTC 10–11 May, for Beaumont on 20–21 May, and for Oxford at 0930 UTC 21 May.

2. Model setup and analysis

a. WRF runs

The model results analyzed are from WRF version 3.2 runs described in LeMone et al. (2013) for 4–5 May, 10–11 May, and 20–21 May, with an additional set performed using WRF version 3.4 with YSU version 3.4.1. Each
simulation was run for 24 h, starting at 1200 UTC (0600 LST), using four two-way interacting nested grids with spacing of 27, 9, 3, and 1 km, respectively. The 2128 × 2547 km² outer domain (Fig. 4; LeMone et al. 2013) extends across most of the continental United States, and the inner 127 × 107 km² grid is centered on the CASES array. The vertical grid has 44 sigma levels, with the lowest half model level just below 5 m, spacing increasing with height (e.g., Fig. 4), and the top level at about 16 km. Initial and boundary conditions for WRF are from the 3-h North American Regional Reanalysis (NARR; http://rda.ucar.edu/datasets/ds608.0/) data on a 32-km grid.

The physical parameterizations include the Noah land surface model (Chen and Dudhia 2001a,b; Ek et al. 2003), the Rapid Radiative Transfer Model (RRTM) long-wave parameterization scheme (Mlawer et al. 1997), the Dudhia (1989) shortwave radiation scheme, and the Lin et al. (1983) bulk microphysics scheme. Three PBL schemes (described in more detail in the next section) were linked to their default surface layer options (option 1 for YSU, option 2 for MYJ, and option 4 for QNSE); BouLac uses the same option as MYJ. Surface characteristics are based on the Moderate Resolution Imaging Spectroradiometer (MODIS) VEGPARM Table version 3.1.1, with modified surface roughness values z₀ (see Table 3; LeMone et al. 2013). Land use types over the CASES-97 array are mainly crop- and grassland, with the latter increasing eastward. All three grassland sites used for model observation comparisons (Beaumont, site 1, and site 2) correspond to grassland grid cells in WRF.

b. PBL schemes in stable conditions

The four PBL schemes, BouLac, MYJ, QNSE, and YSU, are outlined in Table 4, along with references. The first three, henceforth called TKE schemes, solve various forms of the equation for TKE (represented here by $\varepsilon$), given by

$$\frac{\partial \varepsilon}{\partial t} = \frac{\partial (w' \varepsilon')}{\partial z} - \frac{1}{\rho} \frac{\partial (w' \rho' T'_v)}{\partial z} - \frac{u'w'}{\rho} \frac{\partial U}{\partial z} - \frac{v'w'}{T'_v} \frac{\partial V}{\partial z} + g \frac{\varepsilon}{T'_v} T'_v = -\varepsilon, \tag{1}$$

where the wind components ($u$, $v$, $w$) are in a right-handed coordinate system with $u$ positive east, and each component is the sum of the resolved (upper case) and unresolved/parameterized (primed) component. For $T'_v$ and air density $\rho$ the resolved portion is indicated with an overbar. The first two terms on the right-hand side are the vertical divergence of vertical energy transport and vertical pressure transport by $w'$, the third and fourth terms are shear production, the fifth term is buoyancy production, and the sixth term is dissipation. Note that the three TKE schemes allow neither TKE transport by the resolved flow nor horizontal TKE transport by turbulence.

For all four PBL schemes (including YSU, for which the nonlocal terms are zero for stable conditions), the tendency for a quantity $C$ due to subgrid fluxes is found using an eddy diffusivity $K_c$ from

---

1 Both vertical divergence terms are much smaller than the other terms in (1), with the pressure transport contribution close to zero in the Kosovic and Curry (2000) LES of a weakly stable NBL.
The eddy viscosity $K_M$ for the TKE schemes is given by

$$K_M = F_{M,H} L_{\text{mix}} e^{0.5}$$  \hspace{1cm} (3)$$

In (3), $K$ is the eddy diffusivity, and the subscripts $M$ and $H$ refer to momentum and heat, respectively. The master length scale $L_{\text{mix}}$ and the function $F_{M,H}$ varies with the scheme (Table 4) as does the Prandtl number $Pr$, given by the ratio $K_M/K_H$. Note that $Pr = 1$ for BouLac. Also, $L_{\text{mix}}$ for BouLac is the maximum vertical distance traveled by a frictionless air parcel with initial vertical velocity $(2e)^{1/2}$ against unfavorable thermal stratification. This simplification, which ignores the decrease with stability of the fractional contribution of $w^2$ to $e$, is expected to overestimate $L_{\text{mix}}$ (Therry and LaCarrere 1983).

For YSU, $K_H$, for a stable boundary layer is calculated from the YSU NBL depth $h_{\text{YSU}}$ via

$$K_H = K_M = kw_z \left(1 - \frac{z}{h_{\text{YSU}}} \right)^2,$$  \hspace{1cm} (4)$$

where the von Kármán constant $k = 0.4$ and $w_z$ is essentially equal to the friction velocity $u_*$ in YSU version 3.2 and $w_z = u_*(1 + 5z/L)^{-1}$ in YSU version 3.4.1.

c. Evaluation of model NBL depth and selection of the 5% TKE$_{\text{max}}$ criterion

The model output $h$ values, $h_{\text{MYJ}}$, $h_{\text{QNSE}}$, and $h_{\text{YSU}}$, based on the criteria listed in Table 4 and shown as solid lines in Fig. 4, diverge much more than implied by the actual profiles of TKE or, for YSU, $K$ ($h_{\text{sub}}$ indicated by dashed lines). Also, TKE falls with height to a different “background” value for each scheme. Thus, we defined a candidate NBL depth $h_{f_{\text{TKE}_{\text{max}}}}$ as the height satisfying

$$h_{f_{\text{TKE}_{\text{max}}}} = f(\text{TKE}_{\text{max}} - \text{TKE}_b) + \text{TKE}_b,$$  \hspace{1cm} (5)$$

where TKE$_b$ is the background value and $f = 0.05$ ($\approx$ the value in Fig. 1) or 0.10. For BouLac, TKE$_{\text{max}}$ was always at half grid level $zh_2$ ($\approx$ 5 m). For MYJ and QNSE, TKE$_{\text{max}}$ tended to occur at the surface (full grid level $zf_1$) but not always; so TKE$_{\text{max}}$ was determined for the lowest kilometer. We also considered the same TKE thresholds as in LeMone et al. (2013), namely, 0.101 and 0.2 m$^2$s$^{-2}$.

The candidate $h$ criteria based on the mean profiles and Richardson numbers are listed in Table 1. Note that the altitude at which the vertical gradient of virtual potential temperature $\Theta_{v,z} = 10$ K km$^{-1}$ is within $\approx$ 10 m of $h_{\text{Tvmax}}$. This is not surprising, as can be shown by subtracting the adiabatic lapse rate $9.8$ K km$^{-1} \times (z - z_1)$ from $\Theta_{v,z}$ to obtain a profile with the same shape as $T_v$. Thus, we use $h_{\text{Tvmax}}$ in the text for the sake of brevity, while using $h_{\Theta_{v,z}}$ or $\Theta_{v,z} = 10$ K km$^{-1}$ in the figures for the sake of accuracy.
We do not evaluate minimum curvature objectively but discuss it in section 4.

To describe the bulk Richardson number $R_i$ criteria, we start with the expression

$$R_i = \frac{2g(\Theta_{vt} - \Theta_{vb})(\zeta_t - \zeta_b)}{(\Theta_{vt} + \Theta_{vb})[(U_t - U_b)^2 + (V_t - V_b)^2]},$$

where the subscripts $b$ and $t$ denote the bottom and top of the layer for which $R_i$ is calculated. Thus, for the local Richardson number $R_{iloc}$ in Table 1, the subscripts $b$ and $t$ refer to adjacent grid points. For the Richardson number $R_{ilayer}$, the two subscripts refer to the top grid point of the layer and the lowest grid point above the surface (level 1 at height $z_{h1}$). The $R_{ilayer}$ used to find $h_{YSU}$ ($R_{iYSU}$) is based on (6) with $U_b$ and $V_b$ set to zero. The layer $R_i$s are assigned to the layer top; $R_{iloc}$ is assigned to the layer midpoint.

All thresholds are examined moving upward until their value is bracketed and interpolation can take place to determine the corresponding $h$.

The eight candidate $h$ indicators were plotted on TKE (or for YSU, $K$) profiles for each night and location for subjective assessment, as illustrated in Fig. 5.² The TKE only slightly above the background ($0.1 \text{ m}^2 \text{s}^{-2}$) in the upper part of the profiles at 0000 and 0200 UTC is associated with decaying CBL turbulence. After 0200 UTC, the two $h_{Ri}$ are close to or slightly higher than $h_{subj}$, the top of the enhanced TKE profile. The TKE-based criteria 5, 7, and 8 do well through the night, but $h_{TKE}$ from criterion 6 (TKE $= 0.101 \text{ m}^2 \text{s}^{-2}$) is initially greater than $h_{subj}$ by two grid points, changing to one grid point later on. Finally, $h_{Tvmax}$ and $h_{Smax}$ are different from one another and from $h_{subj}$ at the beginning of the night, but converge by 0500 UTC, after which they correspond to within one grid point.

Figure 6 summarizes the height comparisons for Beaumont. From the figure, criteria 7 or 8 were closest to $h_{subj}$. Comparing results from all three profiler sites, we chose criterion 7, which we will call the 5% TKE $= 0.010$ criteria, as our best estimate of $h$. Constant TKE thresholds did not work as well, with high biases for large TKE changing to low biases for small TKE, and sometimes TKE maxima remained below the TKE threshold, implying “no” NBL when subjective or percentage-based assessments indicated there was one. Nor were Richardson number criteria a good match, with the relationship of $h_{Ri}$ to $h_{subj}$ varying from hour to hour and from day to day. Indeed, QNSE $h_{Rloc}$ (criterion 3) was both greater and less than $h_{subj}$ on 5 May. As in Fig. 5, $h_{Tvmax}$ (criterion 1) and $h_{Smax}$ (criterion 2) for MYJ and QNSE were different from $h_{subj}$ after sunset, but were mostly within one grid point of $h_{subj}$ after several

² Odd hours are omitted for readability.
hours. The YSU $h$ value $h_{YSU}$ was closer to $h_{subj}$ than those based on the criteria in Fig. 6 and hence will be used here. This is not surprising, since $h_{YSU}$ is used to calculate $K$ in (4).

The 5% TKE$_{max}$ criterion has its drawbacks. For example, the logical choice of $h_{subj}$ as the height at which the TKE decrease with height abruptly slows down can lead to a value quite different from $h_{5\%TKE_{max}}$. This problem was common for QNSE, since its TKE profile decreases asymptotically with height toward its background value; such profiles also occur in the early evening for the other TKE schemes due to decaying turbulence in the residual layer. Further, the TKE profiles sometimes depart significantly from Fig. 1, especially on weaker wind nights, with TKE$_{max}$ several grid points above the surface. In such circumstances, $h_{5\%TKE_{max}}$ can occur either above or below the height of TKE$_{max}$, depending on where the criterion is first met.

d. Uncertainty in heights of NBL features in WRF output

Our analysis is limited by relatively coarse vertical grid spacing compared to $h$, which varied from $\sim 100$ m (resolved by 5 grid points) to $\sim 500$ m (resolved by 10 grid points). In addition, maps of 1-km domain w at 270 m (a “typical” NBL depth) indicate weak but noticeable resolved wave structures, which could displace $h$, $h_{S_{max}}$, and $h_{T_{v_{max}}}$ vertically. While their impact appears to be minor for the weaker wind nights, the structures reach an amplitude of $\sim 0.1$ m s$^{-1}$ by 0900 UTC 5 May. With a northwest–southeast orientation and a 30-km wavelength along the north–south wind (20 m s$^{-1}$) this translates to a worst-case displacement of features of up to $\sim 24$ m.

3. Comparison to observations

a. Relationship among observed NBL profile features

Since our sample is small, we look for repeatable behavior of $h_{S_{max}}$, $h_{T_{v_{max}}}$, $h_{1w_{sonde}}$, and $h_{2w_{sonde}}$ before comparison to model results. The observations are summarized in Fig. 7. Though there are considerable differences among the three heights for some of the cases, there is a close match between $h_{T_{v_{max}}}$ and $h_{S_{max}}$ before 0800 UTC (0200 LST), just as for the model results in Figs. 5 and 6. Note that Beaumont had the fewest clear estimates of $h_{1w_{sonde}}$ and $h_{2w_{sonde}}$, perhaps due to air currents associated with nearby terrain.
The composited data (Fig. 8) show convergence of $h_{T\text{max}}$ and $h_{S\text{max}}$ with time, with $h_{T\text{max}}$ starting out lower than $h_{S\text{max}}$ but increasing fast enough to catch up with it by around 0800 UTC, at which time both lie within the $h$ range bracketed by minimum $h_{1\text{wsonde}}$ and maximum $h_{2\text{wsonde}}$. Composite time series of each height were estimated using its value at 0930 UTC. For example, for each night and location, (i) the time series $h_{T\text{max}}(t)$ was divided by $h_{T\text{max}}(t=0930\text{ UTC})$, (ii) the $h_{T\text{max}}(t=0930\text{ UTC})$ values were averaged (for all cases in Fig. 7 except for 11 May/Whitewater, when soundings ended before 0930 UTC), and (iii) the normalized heights were multiplied by the average 0930 UTC value to obtain its composite value. The procedure was similar for $h_{S\text{max}}$, $h_{1\text{wsonde}}$, and $h_{2\text{wsonde}}$.

Time convergence of $h_{T\text{max}}$ and $h_{S\text{max}}$ is supported by the closeness of average $h_{T\text{max}}$ at 0930 UTC (279 m) to average $h_{S\text{max}}$ from radiosondes (264 m) and blended RWP + MS data (265 m), using the four cases with good data from both sources (Fig. 7). Taking the seven cases for which 0930 UTC values can be determined without extrapolation, average $h_{T\text{max}} = 377$ m and average $h_{S\text{max}} = 354$ m, close to average $h_{2\text{wsonde}} (333–363\text{ m})$ but greater than average $h_{1\text{wsonde}} (212–242\text{ m})$, where the first number is the sonde height and the second number accounts for the maximum possible correction for the balloon–sonde separation.

b. Relationship among modeled NBL profile features and comparison to observations

Figure 9 shows four types of modeled behavior for $h_{T\text{max}}$, $h_{S\text{max}}$, and $h$. The pattern in the top panel, that is, converging of $h_{T\text{max}}$, $h_{S\text{max}}$, and $h$, with time, is most consistent with observations (Figs. 7, 8). In this case $h_{\text{MYJ}}$ overlaps with the heights of the two maxima, while $h_{5\%\text{TKE}_{\text{max}}}$ is about a grid point lower. The second and third patterns involve very low $h_{T\text{max}}$ and do not correspond to observations. For the final pattern, the more rapidly growing $h_{T\text{max}}$ approaches $h_{S\text{max}}$, but then overtakes it, ending up several grid points higher. In this case, the evolution of $h_{S\text{max}}$ follows observations, but
$h_{TV_{max}}$ grows too rapidly (cf. Figs. 7 and 9), eventually exceeding observed values as well as $h_{S_{max}}$. Since the last three patterns in the figure are not observed, we refer to these patterns as “pathological”$^3$ for this dataset.

To trace the origins of the pathologies in Fig. 9, we plot their frequency as a function of stability (Obukhov length $L$) for three types of $h_{TV_{max}}$ behavior: “increasing,” “too low,” and “too high” in Figs. 10 and 11. Individual points were counted. Thus, all the $h_{TV_{max}}$ points showing an increase toward $h_{S_{max}}$ with time were counted in the increasing pattern; points for which $h_{TV_{max}} < 50$ m were counted as too low, and points for which $h_{TV_{max}}$ exceeds $h_{S_{max}}$ by more than a grid point were counted as too high.

Thus, for example, BouLac 11 May in Fig. 9 has five increasing points and six too-low points. Similarly, only the last three points for the YSU 3.4.1 case on 5 May fall in the too-high category.

From Figs. 10 and 11, the behavior of $h_{TV_{max}}$ varies with PBL scheme, with MYJ reproducing the observed increasing pattern most often (89% of the samples). YSU 3.2 (67%) and BouLac (55%) show too-low $h_{TV_{max}}$ most often and MYJ (7%) the least. The too-high values are the least common, with a few examples for YSU 3.4.1 (7%) and MYJ (4%). Both too-low and too-high $h_{TV_{max}}$ occurred more often for more near-neutral situations (larger $L$), although the association is not perfect. The lack of a sharp distinction is likely related to horizontal variation of wind, temperature, and $L$; so, any relationship reflects upstream as well as local behavior. Also, the TKE schemes from (1) respond to vertical gradients (and for QNSE, the bulk Richardson number; see Table 4) more directly than the surface flux-determined $L$.

$^3$ A persistent (3 h) shallow $S$ maximum occurs at ~50–100 m, which is sometimes linked to a $T_v$ maximum on 5 May at Whitewater (Fig. 7). A check of other observed sounding sequences showed this behavior to be unique.
The origins of too-low $h_{T_{\text{vmax}}}$ become apparent when we examine the results for Beaumont on 5 May, the windiest night, in Fig. 12. In the time series (left side), MYJ and YSU 3.4.1 replicate observed $h_{T_{\text{vmax}}}$ to within a grid interval until at least 0800 UTC, while BouLac, QNSE, and YSU 3.2 show $h_{T_{\text{vmax}}} < 50 \text{ m AGL}$. The vertical profiles (right side) reveal the explanation: too-strong vertical mixing produces deep near-neutral layers for BouLac and YSU 3.2 and a more modestly well-mixed layer for QNSE; all result in low-level $T_{\text{v}}$ maxima. (The lower secondary maximum in the observed $T_{\text{v}}$ profile does not persist.) For the BouLac and YSU 3.2 wind profiles, a height of minimum curvature below 100 m and an $h_{S_{\text{max}}}$ much greater than observed roughly bracket the well-mixed thermal layer, while $h_{S_{\text{max}}}$ for QNSE, MYJ, and YSU 3.4.1 is close to the observed value.

From Fig. 13, the excessive mixing for YSU 3.2 and BouLac can be traced to their $K (= K_H = K_M)$ being much larger than for the other three PBL schemes. In the case of YSU 3.2, large $K_H$ results from setting the scaling velocity $w_s$ to its neutral stratification value $u^*$; the more reasonable $K$ in YSU 3.4.1 results from accounting for stratification in the $w_s$ formulation (Table 4). In the case of BouLac, the large $K$ results primarily from the too-large $L_{\text{mix}}$. The QNSE $h_{T_{\text{vmax}}}$ collapse, like that for BouLac and YSU 3.2, is related to relatively large $K_H$ in the lowest 150 m. Unlike BouLac and YSU 3.2, QNSE produces a $K_M$ close to that of MYJ, leading to reasonable agreement with the observed $S$ profile in Fig. 12. The final pathology, for which $h_{T_{\text{vmax}}} > h_{S_{\text{max}}}$, is also associated with too-strong mixing that just happens not to be strong enough to form a low-level $T_{\text{v}}$ maximum.

The differences between QNSE and MYJ are related to differences in their $Pr$ values. Figure 14 (top) compares $Pr^{-1} = K_H/K_M$ for hourly profiles at Beaumont for the same night as Figs. 12 and 13. While $K_H/K_M = 1.4$ for

---

**Fig. 14.** (top) For 5 May Beaumont, inverse Prandtl number $K_H/K_M$ as a function of height for hourly profiles from 0200 to 1000 UTC. (bottom) Relationship between $K_H/K_M$ and local Richardson number, superposed on plot from Monti et al. (2002). For the bottom plot, the field data were collected in nocturnal downslope flows during the Vertical Transport and Mixing Experiment. Laboratory data are from Strang and Fernando (2001). The dashed line extending QNSE to higher $Ri_{\text{loc}}$ is based on Fig. 4 of Sukoriansky et al. (2006).
c. Collapse of the NBL

A final behavior, “collapse” of $h$ to near-zero values, occurs only for the most stable NBL encountered, on 21 May at Whitewater (Fig. 15), with $h_{\text{MYJ}} < 10 \text{ m}$ for 11 h. For QNSE, $h_{\text{TKE},\text{max}}$ is undefined at 0600 UTC (TKE = its background value) and small at 0800 UTC. Based on WRF simulations, the weak winds result from the deceleration of the easterly synoptic-scale wind by downslope forces on the west side of the watershed, consistent with observations (Table 2). In spite of small TKE, $h_{\text{wsonde}}$ and $h_{\text{wsonde}}$ are mostly close to observed $h_{\text{wsonde}}$ and $h_{\text{TKE},\text{max}}$, as well as $h_{\text{TKE},\text{max}}^{5\%}$ for MYJ, QNSE, and YSU 3.4.1. Model success in replicating observations likely results from the momentum and temperature profiles (and thus the TKE they produce) bearing the effects of forces and stronger turbulence upstream.

Figure 16 indicates that small $h_{\text{MYJ}}$ is associated with TKE hitting its lowest-allowed value, $0.1 \text{ m}^2 \text{ s}^{-2}$ at the surface; $h_{\text{TKE},\text{max}}^{5\%}$ is based on profiles of $\text{TKE}^\prime < 0.05 \text{ m}^2 \text{ s}^{-2}$, sometimes with double maxima, as illustrated in the figure. Given stronger modeled (not shown) and observed (Table 2) easterlies, greater model TKE upstream (near-surface TKE at Beaumont 0.3–0.5 $\text{ m}^2 \text{ s}^{-2}$ at 0500–0800 UTC), and the well-defined change in $w_{\text{sonde}}$ in Fig. 16, we suspect that model TKE at Whitewater was too small, something that could be remedied by allowing horizontal transport of TKE. As for YSU 3.4.1, $K$ reverts to a function of vertical grid spacing since $K$ from (4) is less than that value.

d. Daily model bias in depth of NBL profile features

Figure 17 compares modeled and observed $h_{\text{TKE},\text{max}}$, $h_{\text{Smax}}$, and $h$. The green-shaded cells, which indicate model heights within about one grid point of observed heights, show that the number of successful predictions so defined is about the same for MYJ (using $h_{\text{TKE},\text{max}}^{5\%}$), QNSE, and YSU 3.4.1. The most obvious problem in the figure is the relatively large positive $h$ and $h_{\text{Smax}}$ biases for all the PBL schemes on 5 May, the windiest day. Also evident are the previously discussed impacts of excess mixing on $h_{\text{TKE},\text{max}}$, with unrealistically low values, especially on 5 May for BouLac and QNSE (light blue cells) and both too-high and too-low values for BouLac on the less windy days (red cells). Excess mixing on 5 May is also reflected in too-high $h_{\text{Smax}}$ for MYJ and to a lesser degree for QNSE and for all days for BouLac. The too-high $h_{\text{Smax}}$ for BouLac is commonly paired with a height of minimum curvature close to $h_{\text{TKE},\text{max}}$ (e.g., Fig. 12) or the minimum curvature height in the $T_v$ profile. The MYJ, QNSE, and YSU 3.4.1 $S$ profiles rarely show similar behavior. Spikes in $h_{\text{Smax}}$ most common for MYJ, are mostly associated with real shifts in $S_{\text{max}}$. 

---

4 This is consistent with Janjic (2001, p. 13) as well as our choice of $R_{\text{i,loc}} = 0.5$ as a potential NBL depth criterion (Table 1). On nights when it failed for MYJ, $h_{\text{Ri,loc}} > h$. 

---

**Fig. 15.** For 21 May Whitewater, observed and modeled time series of NBL profile features. Observations: Red with $h_{\text{Smax}}$ from both radiosondes (circles) and RWP + MS (squares), and $h$ zone based on $h_{\text{wsonde}}$ (upside down triangles) and $h_{\text{wsonde}}$ (triangles). For PBL schemes: BouLac (green), MYJ (turquoise), QNSE (blue), and YSU 3.4.1 (purple). For BouLac, YNSE, and MYJ, $h = h_{\text{TKE},\text{max}}^{5\%}$; for YSU, $h = V_{\text{MYJ}}$. Also shown is $h = h_{\text{MYJ}}$ (light turquoise). (right) Indicates grid heights for (top), (middle) $zh$ and (bottom) $zf$. 

**Fig. 16.** Whitewater 21 May 1997

QNSe [consistent with Sukoriansky et al. (2006)] near the surface where $R_{\text{i,loc}}$ is closest to neutral, $K_{\text{H}}/K_{\text{M}} = 1.0$ for MYJ, a consequence of (A8) in Janjic (2001). If one accepts $Pr = 1$ for the surface layer (e.g., Kaimal and Finnigan 1994, their Fig. 1.8), the MYJ Prandtl number is closer to correct near the surface. Further, when $Pr^{-1}$ is plotted against $R_{\text{i,loc}}$ (Fig. 14, bottom), the MYJ points fit the Monti et al. (2002) data at least as well as the QNSE points [though Grachev et al. (2007) suggest that the relationship in such plots is contaminated by self-correlation and that $Pr$ might actually increase with $R_{\text{i,loc}}$ if self-correlation is eliminated]. Note that $R_{\text{i,loc}} < 0.5$ for all MYJ NBL output examined. However, $Pr = 1$ when TKE = $0.1 \text{ m}^2 \text{ s}^{-2}$, its background value.
The assumption that $K_H = K_M$ for $z < h_{YSU}$ does not seem to have had a negative impact on YSU 3.4.1. Why is this? As illustrated by Fig. 18, $R_{\text{loc}}$ (and thus $Pr$, see Fig. 14, bottom) reaches a maximum above the height of the $K$ maximum, keeping $Pr$ closer to 1 where $K$ is the largest. Furthermore, NBL $R_{\text{loc}}$ reaches only $\sim 0.5$–1 in windy conditions. However, it should be noted that $Pr > 1$ for $z > h_{YSU}$.

Although $h_{5\%\text{TKE, max}}$ often appears to be a better measure of $h$ than $h_{\text{MYJ}}$, the differences are not that large, except for around sunset (Fig. 5, noting that $h_{\text{MYJ}} > h$ from criterion 6). From Figs. 5, 9, and 17, the differences between the two estimates are mostly of the order of 1–1.5 grid points. Furthermore, $h_{\text{MYJ}}$ coincides with $h_{\text{vmax}}$ and $h_{\text{smax}}$ by 0700 UTC in Fig. 9, while $h_{5\%\text{TKE, max}}$ lies a grid point lower. Finally, the collapse of $h_{\text{MYJ}}$ to near-zero values is a better indicator of the modeled very stable NBL (MYJ TKE excess over background is less than 0.01 m$^2$ s$^{-2}$). As noted previously, it is likely that the observed $h$ bounded by $h_{\text{1wsonde}}$ and $h_{\text{2wsonde}}$ is strongly influenced by the advection of TKE from upstream. We speculate that the MYJ mean profiles, being partially shaped by the upstream effects of TKE, produce TKE profiles of much smaller magnitude relative to the background but with depths ($h_{\text{subj}}$ and $h_{5\%\text{TKE, max}}$) similar to observations.

d. Surface fluxes

Since surface fluxes influence NBL evolution, we compare model $u_\theta$ and $(\overline{wT^\prime})_{\text{sfc}}$ to the observed values. We use Beaumont and grassland flux sites 1 and 2, since all three sites are only $\sim 10$ km from each other (Fig. 2) and correctly classified as grassland in WRF.

Figure 19 shows variation among model $u_\theta$ values at Beaumont to be small compared to their excess over observed values at the flux sites, while model observation differences in $(\overline{wT^\prime})_{\text{sfc}}$ are only slightly larger than the spread among model values, with model values more negative. To remove the impact of horizontal variability, we plot observed and MYJ surface fluxes at sites 1 and 2 along with MYJ surface fluxes at Beaumont in Fig. 20. As expected, the modeled fluxes at sites 1 and 2 are mostly closer to observations than those for Beaumont, especially for $(\overline{wT^\prime})_{\text{sfc}}$ on 11 and 21 May, but the $u_\theta$ discrepancy is still large. The small spread for model $u_\theta$ compared to observed values suggests the model observation differences are not due to differences in the PBL or associated surface schemes.

Nor does the $u_\theta$ discrepancy appear to be due to biases in the observed values, which are half-hour averages. Averaging fluxes over smaller time intervals has been suggested for stable conditions (e.g., Vickers and Mahrt 2006) in order to obtain internal consistency between turbulence and mean profile measurements and to avoid scatter associated with “mesoscale” motions. However, $u_\theta$ varies smoothly with time. Furthermore, given that averaging times should increase as stability decreases, the 30-min averaging times should work the best for the near-neutral 5 May case, when the discrepancy is largest. Finally, our objective is not to find consistency between measured fluxes and profiles, but rather to account for all unresolved fluxes.
Though the largest \( u* \) discrepancy for 5 May in Fig. 17 is consistent with the greatest model departures from observations, MYJ, QNSE, and YSU 3.4.1 perform rather well on the weaker wind days, despite significant \( u* \) discrepancies. Given this behavior, the poor correspondence of observed and modeled \( u* \) is somewhat surprising, particularly since there is some evidence (e.g., Table 1 of Caughey et al. 1979, their Eq. (4); Steeneveld et al. 2007, p. 222) for a linear relationship between \( h \) and \( u* \), other things being equal.

In addition, MYJ 10-m winds are consistent with observations on all 3 days (Fig. 21). Based on averages of sites 1 and 2 between 0200 and 1100 UTC, the observed speed exceeds the linearly interpolated model results by 0.48, 0.14, and 0.58 m s\(^{-1}\) for 5, 11, and 21 May, respectively. Interpolation assuming a logarithmic profile reduces the differences by up to 0.2 m s\(^{-1}\), based on calculations for the day with the strongest winds (5 May). Thus, model underestimates are \(-2\%–3\%\) for this day.

Moreover, a similar \( u* \) discrepancy emerges when we apply the \( h \) criterion of Richardson et al. (2013),

\[
h = \frac{Ri_c L}{\alpha},
\]

where \( Ri_c \) is the layer Richardson number evaluated at \( h \) and \( \alpha \) is a constant, to modeled and observed data. Both roughly satisfy (7). However, while observed \( \alpha \) is within 40\% of their \( \alpha \) (0.045), the model yields \( \alpha \) values a factor of up to 5 larger (Table 5), depending on how \( Ri_c \) is calculated. The large \( \alpha \) discrepancy results primarily from

\[5\]
from differences between observed and modeled $u_\infty$. For the observations, $\alpha$ was estimated from averaged data using times for which we had high confidence in $h$, while model $\alpha$ was calculated from the corresponding hourly $L$, $h_{\text{subj}}$, and $R_{\text{hyd}}(h_{\text{subj}})$ and then averaged.

Drawing from Kustas et al. (2005) and Strassberg et al. (2008), we hypothesize that both $u_\infty$ values are correct, with the observed values representing $O(100)$ m fetches and the model value representing a much larger region. To demonstrate this, we assume that the model and observed wind are equal, but that the surface roughness length $z_0$ values differ. Thus, assuming neutral stability, the wind speed $S$ is related to model ($M$) and observed ($O$) $u_\infty$ and momentum roughness length $z_0$, via

\[
S_{10} = \frac{u_{\infty M}}{k} \ln \frac{z}{z_{0 M}} = \frac{u_{\infty O}}{k} \ln \frac{z}{z_{0 O}}. \tag{8}
\]

Solving for $z_{0,O}$ by setting $u_{\infty M} = R u_{\infty O}$, we obtain

\[
\ln z_{0,O} = R(\ln z_{0,M} - \ln z) + \ln z. \tag{9}
\]

The stability is closest to neutral on 5 May, when $R = 1.5$ (Fig. 20). Using $z = 10$ m and $z_{0,M} = 0.05$ m (intermediate value for grass in our simulations; see Table 3 of LeMone et al. 2013), we obtain $z_{0,O} = 0.0035$ m. This value is close to the approximate values tabulated online (at http://www.eol.ucar.edu/isf/projects/cases97/), namely, $z_{0,O}$ (site 1) $\sim 0.006$ m and $z_{0,O}$ (site 2) $\sim 0.002$ m.

4. Conclusions

Radiosonde, minisodar, radar wind profiler, and surface observations and WRF simulations for three moderately windy fair-weather nights during the CASES-97 field program are used to identify NBL depth $h$ and the heights of maxima in wind speed $S_{\text{max}}$ and virtual temperature $T_{\text{v,max}}$, which are then used to evaluate four PBL schemes: BouLac, MYJ, QNSE, and YSU. Rather than simply focus on biases, we determine the observed co-evolution pattern of $h$, $h_{\text{Tmax}} \sim h_{\text{v,max}}$, and $S_{\text{max}}$ and then evaluate the success of the four schemes in reproducing that pattern as a function of environmental conditions, as defined by the Obukhov length $L$.

To find $h$ for BouLac, MYJ, and QNSE, we compared eight objectively determined $h$ criteria (four TKE-based criteria, two Richardson number criteria, $h_{\text{Tmax}}$, and $h_{\text{Smax}}$; Table 1) to subjectively based values ($h_{\text{subj}}$) on plots of the model TKE profiles. Based on this comparison, we chose a threshold equal to 5% of the maximum TKE excess from its background value, where the maximum was found for the lowest kilometer. Fortuitously, the height so derived, $h_{\text{5\%TKE}}$, was consistent with Fig. 1. However, $h_{\text{5\%TKE}}$ was not a significant improvement over the MYJ-derived value $h_{\text{MYJ}}$. For YSU, we used its Richardson number–based $h_{\text{YSU}}$, which was a close equivalent to the subjective NBL depth based on $K$, except under the most stable conditions, when $K$ was proportional to the vertical grid spacing.

The observed TKE-based $h$ was based on a decrease in balloon rise rate from $\sim 5$ to $\sim 3$ m s$^{-1}$ going from turbulent to nonturbulent air (Johansson and Bergstrom 2005). Taking the heights at which the deceleration started ($h_{\text{1wsonde}}$) and ended ($h_{\text{2wsonde}}$) yielded reasonable bounds for $h$, given proper balloon launch procedure, the absence of large air vertical velocities, and $h > \sim 100$ m. This method worked least reliably at Beaumont, perhaps because of the currents associated with nearby terrain.

Summary plots of composite $h_{\text{1wsonde}}$, $h_{\text{2wsonde}}$, $h_{\text{Smax}}$, and $h_{\text{Tmax}}$ revealed a general pattern: $h_{\text{Tmax}}$ increases gradually through the night, $h_{\text{Smax}}$ and $h_{\text{Tmax}}$ converge, and the two approach the $h$ zone based on $w_{\text{sonde}}$ after
several hours, after which all three occupy roughly the same altitude range until surface heating starts to form the CBL. On many nights $h_{S\text{max}}$ followed $h$ through most of the night to such a degree that $h_{S\text{max}}$ was a good secondary measure of $h$, in agreement with the work of Banta et al. (2003) and Pichugina and Banta (2010).

Kosovic and Curry (2000) produced such an evolution using LES, although they cautioned that it took an inertial period (about 20 h at this latitude) for the three heights to correspond.

The observed coevolution pattern provided metrics against which the NBL schemes could be judged. Of the
PBL schemes examined, MYJ, QNSE, and YSU version 3.4.1 mostly reproduced the observed converging of $h_{Tvmax}$, $h_{Smax}$, and $h$ with time. However, BouLac (55% of the time) and YSU version 3.2 (67% of the time) produced unrealistically low $h_{Tvmax}$, a sign of too much vertical mixing. In both cases, $h_{Smax}$ tended to be too high compared to observations. The low $h_{Tvmax}$ behavior occurred with intermediate frequency for QNSE (33%) and seemed to be associated with too low a Prandtl number (too large a $K_H$) since QNSE wind profiles were simulated far better ($K_M$ about right). The low $h_{Tvmax}$ behavior occurred least frequently for MYJ (7%) and YSU 3.4.1 (8%). In most cases, the excessive vertical mixing was associated with larger values of $L$ (windier nights). However, the $L$ dependence was not a clean one: a bulk Richardson number could be a better parameter, and upstream as well as local forcing determines the profiles of resolved parameters.

A final behavior, the collapse of the NBL ($h_{MYJ} \sim 0$ for several hours; $h_{5\%TKE_{max}}$ undefined 1 h each for MYJ and QNSE), occurred on 21 May (Whitewater), with TKEs for MYJ and QNSE sometimes pegged at background values. In this case, $h_{Tvmax}$, $h_{Smax}$, and (surprisingly) $h_{5\%TKE_{max}}$ were close to observed values, the presumed effect of forces and stronger turbulence upstream on the wind and temperature profiles and thus the model TKE profiles. Consistently well-defined $h_{1wsonde}$ and $h_{2wsonde}$ suggest model underestimates of TKE, a possible result of neglecting horizontal TKE advection.

There was a large mismatch between observed and modeled friction velocity $u_*$ near Beaumont even though the PBL schemes replicated NBL features at Beaumont reasonably well. The discrepancy appears to reflect a mismatch in model and observed roughness length $z_0$. Indeed, using the modeled and measured $u_*$ and $S$ at 10 m, and the model $z_0$, we were able to reproduce the measured $z_0$. Similarly, the observed and model results each produce internally consistent values of the constant $\alpha$ in Richardson et al. (2013) that relates $h$ to $\text{Ri}(h)$ and $L$ [see (7)], but $\alpha$ for the model results is up to 5 times $\alpha$ for the observations, again a consequence of the $u_*$ discrepancy. We suggest that the observed $z_0$, $u_*$, and 10-m wind speed reflect skin drag over an $O(100)$ m grassland fetch, while the corresponding model values are consistent with a larger flux footprint that produces form drag as well as skin drag. Likewise, the modeled and observed NBL wind profiles likely reflect regional $z_0$ values that include form drag. Similar conclusions were reached by Kustas et al. (2005) and Strassberg et al. (2008) for the CBL. Such discrepancies may also contribute to the surface flux inconsistencies noted by Hacker and Angevine (2013).

The conclusions should be generalized with caution. Because of data limitations, we limited ourselves to windier nights, during which we would expect deeper NBLs and more continuous coupling of the atmosphere to the surface. The 108 h and three locations analyzed thus mostly represent less stable conditions [intermittent to continuous turbulence regimes of Van de Wiel et al. (2003)]. Further, the TKE did not always simply decrease with height, sometimes reaching a maximum (or two maxima) above the surface. Finally, resolved mesoscale NBL structures could influence the results slightly.

The small number of nights sampled is compensated for by the completeness of the dataset, which along with
the WRF runs, allows the examination of the impacts of upstream and surface conditions on the evolution of the NBL. This case study approach draws on the strengths of the simulations and observations to examine the co-evolution of observed and modeled NBL profile features and TKE values, terrain effects on the flow, and vertical surface fluxes. Though we cannot quantify model biases, we can explore their sources.

As to PBL scheme improvements, the results suggest the following:

- for BouLac, allowing for stability in converting TKE to vertical velocity in the expression for estimating $L_{\text{mix}}$, which could mitigate too much mixing at night [already recognized as a shortcoming by Therry and LaCarrere (1983)];
- modifying QNSE to mitigate apparent excess mixing of the temperature profile;
- using the stability-dependent form for $w_s$ in YSU (i.e., use YSU 3.4.1 rather than 3.2); and
- including the advection of TKE by resolved winds.

As to the observations, the results suggest designing field measurements to better measure the relevant parameters in NBL evolution or using datasets that include a useful subset (e.g., CASES-99; Poulos et al. 2002):

- For NBL depth, this would include radiosonde releases optimized to find $h$ using balloon rise rate. Collocated tethersonde, tower, and lidar data could enable sampling a broader range of NBLs as well as comparison of techniques for determining $h$.
- This would also involve designing flux measurements to include “regional” as well as “local” fluxes by adding measurements from a taller tower and/or aircraft or unmanned aerial vehicles to a traditional flux tower network. Analysis would include examination of averaging times.

**Acknowledgments.** The surface meteorology, surface flux, and radiosonde data were collected, processed, and archived by NCAR’s Earth Observations Laboratory and the RWP/MS data were collected, processed, and archived by the Argonne National Laboratory’s former Argonne Boundary Layer Experiments Facility. The authors are indebted to Tom Horst and Steven Onclcy of NCAR and Richard Coulter of Argonne National Laboratory for help in data interpretation as needed. Valuable help in interpreting the PBL schemes was provided by Alberto Martilli (BouLac), Zavisa Janjić (MYJ), Esa-Matti Tastula (QNSE), and Songyu Hong (YSU). Xubin Zeng, Jeff Weil, Sukanta Basu, Jielen Sun, and Wayne Angevine provided useful insights, with the last two also helping to clarify figures. Finally, suggestions by Jun Zhang and an anonymous reviewer helped us clarify several points and improve the figures. This work was supported by the Air Force Weather Agency, the NCAR Water System Program, and NCAR BASE funding from the National Science Foundation.

**REFERENCES**


<table>
<thead>
<tr>
<th>Date</th>
<th>Time (UTC)</th>
<th>$h$</th>
<th>$Ri_k$</th>
<th>$R_i(0-h)$</th>
<th>$R_i(z_1-h)$</th>
<th>$L$</th>
<th>$\alpha$</th>
<th>$\alpha_{0..h}$</th>
<th>$\alpha_{1..h}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 May</td>
<td>0200–1100</td>
<td>277</td>
<td>0.165</td>
<td>—</td>
<td>—</td>
<td>107</td>
<td>0.064</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>11 May</td>
<td>0750–1000</td>
<td>260</td>
<td>0.37</td>
<td>—</td>
<td>—</td>
<td>38.8</td>
<td>0.054</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>21 May</td>
<td>0100–1000</td>
<td>140</td>
<td>0.24</td>
<td>—</td>
<td>—</td>
<td>31.2</td>
<td>0.054</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

For NBL depth, this would include radiosonde releases (e.g., CASES-99; Poulos et al. 2002):

**Table 5. Estimates of Richardson et al. (2013) constant $\alpha$ in (7) at Beaumont. For observations, $Ri$, from (6) with $h$ and lowest level above surface; $L$ based on sites 1 and 2.**


