Diagnostic Equations for the Stable Boundary Layer Height: Evaluation and Dimensional Analysis

G. J. Steeneveld, B. J. H. van de Wiel, and A. A. M. Holtslag

Meteorology and Air Quality Group, Wageningen University, Wageningen, Netherlands

(Manuscript received 22 September 2005, in final form 28 May 2006)

ABSTRACT

The performance of diagnostic equations for the stable boundary layer height \( h \) is evaluated with four observational datasets that represent a broad range of latitudes, land use, and surface roughness. In addition, large-eddy simulation results are used. Special care is given to data-quality selection. The diagnostic equations evaluated are so-called multilimit equations as derived by Zilitinkevich and coworkers in a number of papers. It appears that these equations show a serious negative bias, especially for \( h/H < 100 \) m, and it was found that the parameters involved could not be determined uniquely with calibration. As an alternative, dimensional analysis is used here to derive a formulation for \( h \) that is more robust. The formulation depends on the surface friction velocity \( u_* \), surface buoyancy flux \( B_s \), Coriolis parameter, and the free-flow stability \( N \). The relevance of the Coriolis parameter for the boundary layer height estimation in practice is also discussed. If the Coriolis parameter is ignored, two major regimes are found: \( h \sim u_*/N \) for weakly stable conditions and \( h \sim (|B_s|/N^3)^{1/2} \) for moderate to very stable conditions.

1. Introduction

The stable boundary layer (SBL) height \( h \) is an important quantity to describe the relevant processes that govern the SBL development and its vertical structure (Holtslag and Nieuwstadt 1986). The stable boundary layer height clearly has an impact on the mixing properties of the SBL. Also, model formulations that use an explicit prescription of the vertical profile of an eddy diffusion coefficient \( K \) require an explicit expression for the SBL height (e.g., Troen and Mahrt 1986; Holtslag and Boville 1993).

Furthermore, the dispersion of pollutants during stable stratification is strongly affected by \( h \) (e.g., Salmond and McKendry 2005). Release of pollutants below \( h \) during periods of weak winds and consequently weak vertical mixing may result in very high concentrations of primary and secondary pollutants, causing serious consequences for life. Therefore, \( h \) is a critical quantity to estimate for meteorological preprocessors in air-quality models (Venkatram 1980; Gryning et al. 1987; Lena and Desiato 1999; Seibert et al. 2000; Karpfpen et al. 2001).

In contrast to the daytime convective boundary layer, where its height can be obtained relatively simply from profile observations (e.g., Vogelezang and Holtslag 1996, henceforth VH96), observing the SBL height \( (h_{\text{obs}}) \) is less straightforward because turbulence is gradually suppressed during stable conditions and may also show intermittent behavior (e.g., Holtslag and Nieuwstadt 1986). Vickers and Mahrt (2004, henceforth VM04) found a classical SBL with well-defined surface-based turbulence for only 22% of the time for the Cooperative Atmosphere–Surface Exchange Study (CASES)-99 field campaign. In addition to turbulence, other processes such as radiation divergence (Anfossi et al. 1976; Garratt and Brost 1981), gravity waves, wave breaking (Newsom and Banta 2003), and baroclinicity (Zilitinkevich and Esau 2003) influence the SBL structure for very stable conditions. In that case, problems occur in measuring \( h_{\text{obs}} \) in the absence of a universal relationship among the profiles of temperature, wind speed, and turbulence variables. The interpretation of wind speed and temperature profiles is not straightforward in that case, and it is therefore not surprising that several definitions for \( h \) are in use nowadays (VM04; Beyrich 1994).
Despite the uncertainties mentioned above, much work has been done in modeling (scaling) the SBL height ($h_{mod}$; see section 2) from profile information or from turbulent variables at the surface. However, many of these models have been validated for a limited range of boundary layer heights or for a single dataset. Zilitinkevich and Mironov (1996, henceforth ZM96) presented a framework paper in which they summarize classical material on the subject. They derive two diagnostic multilimit equations for the SBL equilibrium height by inverse interpolation of relevant SBL height scales. The first aim of the current paper is to evaluate the performance of these two multilimit equations against four observational datasets at different latitudes and against large-eddy simulation (LES) results. The second aim is to present an alternative, robust, and practical formulation for $h$ based on the same variables as in the multilimit equations, but based on formal dimensional analysis instead of inverse interpolation. Last, we discuss and assess the relevance of the Coriolis parameter in practical estimates for $h$.

In section 2 we present background information on defining and modeling $h$. Section 3 describes the observational datasets and the LES data; then in section 4 the two multilimit equations are evaluated and recalibrated. In section 5, we use dimensional analysis to obtain an alternative formulation, and concluding remarks are made in section 6.

2. Background

The subject of defining and modeling the stable boundary layer height has a long history. Some well-known definitions for $h$ are (e.g., VM04) 1) a fraction of the layer through which turbulence exists (Lenschow et al. 1988), 2) the top of the downward turbulent sensible heat flux (Caughey et al. 1979), 3) the lowest maximum of the wind speed [often referred to as the low-level jet (LLJ); Melgarejo and Deardorff 1974], and 4) the top of the temperature inversion or the first discontinuity in the potential temperature profile (Yamada 1979). Some other definitions are given in Stull (1988). Seibert et al. (2000) state that no final answer is received to the question, “What is the most suitable height scale to characterize the vertical mixing in the SBL?” The first two definitions focus on the turbulent structure of the SBL near the surface, and the latter two are based on profile information from radiosondes, tethered balloons, or remote sensing methods.

It is obvious that different values are found by using different definitions. For example, the LLJ (definition 3) is also influenced by processes in the boundary layer other than the actual turbulence intensity (e.g., inertial and baroclinicity effects; Stull 1988). Therefore, the LLJ is perhaps a better measure for the history of the boundary layer than of the actual SBL structure (Mahrt et al. 1982). On the other hand, shear below and above the LLJ will generate turbulence and thus affects the turbulence intensity over the whole boundary layer (Cuxart and Jiménez 2007). In addition, definition 3 is hard to apply when an LLJ is absent, which occurs in reality. For definition 4 we should realize that, besides the effect of turbulent fluxes, the temperature profile is also strongly influenced by radiative cooling (André and Mahrt 1982) and by heat release of breaking waves at the SBL top. A surprise is that Kosovic and Curry (2000) indicate, based on their LES simulations, that under ideal conditions definitions 1, 3, and also 4 are equivalent. In this study, we will mainly apply the last two definitions (section 3), because turbulent flux profile data are often absent aloft and radiosondes provide information concerning $h$ aloft. In this manner, observations biased toward smaller values of $h$ are circumvented.

In principle, also subsidence and baroclinicity determine the value of $h$ (Zilitinkevich and Bakanov 2002; Zilitinkevich and Essau 2003). However, routine observations of subsidence and baroclinicity are usually unavailable and are therefore not taken into account in this study. Note that the technique of dimensional analysis that we use in the second part of the paper is also applicable to other definitions of $h$.

Apart from defining $h$, Seibert et al. (2000) and Beyrich (1997) review methods to observe the SBL height. Traditional methods use sodar or other remote sensing methods (e.g., van Pul et al. 1994) or employ parcel methods on atmospheric profiles (VH96). Beyrich (1997) states that automated detection of $h$ works unsatisfactorily at this moment, and that visual inspection of vertical profiles by experts is still recommended. Therefore, in this study an approach with visual inspection is followed. In any case, the uncertainty of $h_{obs}$ is typically 30%–40%.

From the modeling perspective, several parameterizations for $h_{mod}$ have been proposed. These can in general be subdivided into prognostic equations (e.g., Nieuwstadt 1980a; Nieuwstadt and Tennekes 1981; Mahrt 1981; Gassmann and Mazzeo 2001) and diagnostic equations based on surface parameters (e.g., Zilitinkevich 1972; Arya 1981; Estournel and Guedalia 1990). Nieuwstadt and Tennekes (1981) conclude that, from a fundamental point of view, prognostic equations should be preferred. However, their performance depends strongly on the initial conditions because the intranight variation of $h$ is larger than the intranight variation (André 1983). In contrast, Yu (1978).
concludes that diagnostic models perform better than prognostic models.

ZM96 proposed a framework and identified rotation, surface buoyancy flux, and free-flow stability to be the key physical processes that govern $h$. A formula for the equilibrium value of $h$ was derived by ZM96, considering the steady-state turbulent kinetic energy equation and parameterizing the vertical mixing in the SBL due to these three processes. The formula uses the surface friction velocity $u_*$, surface buoyancy flux $B_s$ [$=(g/\theta)\nabla \theta$], the earth’s rotation given by the Coriolis parameter $f$, and free-flow stability [represented by the Brunt–Väisälä frequency: $N = [(g/\theta)\nabla \theta]^{1/2}$, where $g$ is the gravity acceleration, $\theta$ is the potential temperature, and $z$ is the height above ground level]. The equilibrium height $h$ is then given by

$$
\left( \frac{fh}{C_i u_*} \right)^2 + \frac{h}{C_i L^*} + \frac{Nh}{C_i u_*} = 1,
$$

where $L^*$ is the (modified) Obukhov length, given by $L^* = -\frac{g}{\theta}B_s$ (as defined without the von Kármán constant, following ZM96) and $C_n$, $C_s$, and $C_i$ are constants of proportionality. As such, this is effectively a diagnostic method. The three terms represent the impact of rotation, the surface buoyancy flux, and free-flow stability, respectively. The main advantage of Eq. (1) is its multilimit behavior, that is, for $f \to 0$ or $N \to 0$ or $L^* \to \infty$ the value of $h$ by Eq. (1) remains defined. Note that many earlier proposals—for example, by Zilitinkevich (1972) and Nieuwstadt (1980b)—are undefined for $f \to 0$.

Based on Zilitinkevich (1972) and Pollard et al. (1973), ZM96 add (using inverse interpolation, i.e., geometric averaging) two additional terms to Eq. (1) to “include the cross interactions” between rotation, the buoyancy flux, and free-flow stability. This approach results finally in a multilimit equation for $h$:

$$
\left( \frac{fh}{C_i u_*} \right)^2 + \frac{h}{C_i L^*} + \frac{Nh}{C_i u_*} + \sqrt{\frac{\|B_s\| h}{C_{ir} u_*^2}} + \sqrt{\frac{\|N\| h}{C_{ir} u_*}} = 1,
$$

where $C_n = 0.5$, $C_s = 10$, $C_i = 20$, $C_{ir} = 1$, and $C_{ir} = 1.7$. Joffre et al. (2001) recently reevaluated these coefficients and found $C_n = 0.2$, $C_s = 2.5$, $C_i = 10$, $C_{ir} = 0.4$, and $C_{ir} = 1.2$ from tower observations at Sodankylä, Finland, and VM04 proposed $C_n = 0.04$, $C_s = 6$, $C_i = 15$, $C_{ir} = 0.7$, and $C_{ir} = 0.8$ based on observations from CASES-99, Coupled Boundary Layers/Air–Sea Transfer (CBLAST), and Fluxes over Snow Surfaces (FLOSS) experiments.

The difference among these three sets of coefficients for Eq. (2) is relatively large, which indicates that these (many) coefficients are hard to obtain with a sufficient degree of confidence. For example, $C_n$ ranges from 0.045 to 0.6 in the literature (Benkley and Schulman 1979; Mason and Thomson 1987; Beyrich and Kotroni 1993; Kosovic and Lundquist 2004) and $C_s$ ranges from 1.2 to 100 (ZM96). This subject will be addressed in more detail in section 4. Furthermore, VM04 found a stability dependence of $C_s$ and $C_{ir}$, the existence of which questions the uniqueness of the constants. We have seen that several of these coefficients have a large range in the literature.

Kosovic and Curry (2000) remark that Eq. (2) uses five dimensionless groups, whereas only three can be justified according to Buckingham II theory. Based on the five relevant quantities and only two basic dimensions (viz., meters and seconds), three dimensionless groups should be sufficient to describe the datasets. As such, Eq. (2) has redundancy. Inspired by this comment and by the fact that the coefficients do not seem to be robust, we decided to start with formal application of the II theorem without prescribing some particular shape (such as with inverse interpolation). In this way we will also show that bias for small values of $h$ disappears so that the result is applicable to high-stability cases. Note that the II theorem is semiempirical and is only applicable in the range of available observations (see section 5).

3. Observations

Many model formulations for the SBL height have been proposed based on a single dataset or on datasets biased toward shallow boundary layers (VM04), and thus universality is not a priori guaranteed for these models. On the contrary, the analysis in this paper is based on four observational datasets over different terrain types according to their surface roughness and land use. We also use LES results of the Global Energy and Water Cycle Experiment (GEWEX) Atmospheric Boundary Layer Study (GABLS; Beare et al. 2006) for model verification. The latter reflects a moderately stable case. We do not use LES for more stratified cases because of the many uncertainties that still exist in LES for stratified conditions.

Below we describe the observational sites, the data processing, and the method we used to obtain $h$ from profile observations. Table 1 provides the number of data points per dataset and the observed range of the relevant quantities.
This measurement campaign was held 1–31 October 1999 near Leon, east of Wichita, Kansas (37.6486°N, −96.7351°E, 430 m MSL). The objective was to study the relevant processes in the SBL (Poulos et al. 2002). The area is relatively flat prairie grassland, free of obstacles, with an estimated roughness length $z_0 = 0.03$ m (van de Wiel et al. 2003). During CASES-99, 126 radiosondes were launched from the central site at night (defined here as 0100–1300 UTC). Soundings were launched every 6 h and hourly during intensive observation periods [IOP—see Poulos et al. (2002) for an overview of the IOPs]. The quality of the validation and calibration of SBL height models strongly depends on the data quality, data range (Table 1), and careful data selection and processing. Because many variables are used and we want to have reliable estimates for all of them, we have to reduce the data strictly. Similar to van de Wiel et al. (2003), nights that are subject to rapidly changing synoptic conditions are disregarded (those were classified as “Non” in that paper). We consequently determined the SBL height for the remaining 101 profiles using the method in Joffre et al. (2001). They identified the SBL height subjectively by inspecting together the wind speed, potential temperature, and Richardson number profiles for clear changes below the inversion height that would indicate a change in the structure of the lower atmosphere. This method is illustrated in Fig. 1 for 0700 UTC 23 October 1999. The SBL height was based on 1) the LLJ height $h_{LLJ}$ and 2) the height of the first discontinuity in the potential temperature profile $h_\theta$, from curved to linear in Fig. 1a. Figure 1b illustrates the first peak in the gradient Richardson number at that level. If $h_{LLJ}$ and $h_\theta$ differ by more than 60 m, the data points were rejected, which left 58 data points. In addition, we selected data for $N > 0.015$ s$^{-1}$, because weaker $N$ could not be determined accurately, and then 48 data points remained. The reliability of eddy covariance measurements (10-min averages) for weak turbulence is questionable, and therefore we disregarded data with friction velocity $u_* < 0.04$ m s$^{-1}$ and sensible heat flux $H > -2$ W m$^{-2}$. Because all nights in CASES-99 are described well in

<table>
<thead>
<tr>
<th>Table 1. Overview of observations (per station) during stable stratification, and before and after data reduction.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sodankylä</td>
</tr>
<tr>
<td>No. of data points before reduction</td>
</tr>
<tr>
<td>No. of data points after reduction</td>
</tr>
<tr>
<td>Range $h$ (m)</td>
</tr>
<tr>
<td>Range $H$ (W m$^{-2}$)</td>
</tr>
<tr>
<td>Range $u_*$ (m s$^{-1}$)</td>
</tr>
<tr>
<td>Range $N$ (s$^{-1}$)</td>
</tr>
</tbody>
</table>

**Fig. 1.** Profiles of (a) potential temperature, wind speed, and (b) gradient Richardson number for 0700 UTC 23 Oct during CASES-99. The arrow indicates the stable boundary layer height.
the literature, we checked the literature to see whether the remaining data were probably subject to “events” (e.g., the density current of 18 October). During such events, the boundary layer will not behave like a classical boundary layer, and therefore these data should be disregarded. After this selection, only 32 data points remained.

b. Sodankylä

The dataset (34 data points during stable stratification) is based on an intensive campaign of radiosondes performed during the international Northern Hemisphere Climate-Processes Land-Surface Experiment (NOPEX)/Winter Experiment (WINTEX) program at the observatory of Sodankylä (67.4°N, 26.7°E, 180 m MSL) between 10 and 21 March 1997 (Halladin 1999). The terrain topography around the site is mostly flat, characterized by isolated gently rolling hills (altitude differences of 50–150 m). This area of Lapland is covered by sparse forest (mean tree height of about 8 m around the site). The immediate vicinity of the mast was characterized by semi-open pine forest with a mean height $H_s \approx 8$ m for the trees around the mast. By assuming a displacement height of $d = 0.66H_s = 5.3$ m, neutral profiles yielded $z_0 \approx 1.4$ m (Joffre and Kangas 2002). The data were already carefully selected by Joffre et al. (2001). After applying selection criteria on $u_s$, $H_s$, and $N$ similarly as for CASES-99, 30 data points remained.

c. Cabauw

The current dataset (189 data points) has been obtained in the period 1977–79 at Cabauw, Netherlands (51.971°N, 4.927°E, −0.7 m MSL), by Nieuwstadt (1980b). For boundary layer heights, VH96 made an extensive study with these data, and later the data were also used by Zilitinkevich and Bakanov (2002). The data were carefully selected (e.g., filtered for gravity waves) before in VH96. Some additional restrictions made by us were $u_s < 0.04$ m s$^{-1}$, sensible heat flux $H > -2$ W m$^{-2}$, $N > 0.015$ s$^{-1}$, and the rejection of Type-I boundary layers (dominantly driven by radiation divergence and low winds, as defined in VH96).

The Cabauw area is flat and covered with grass with an overall roughness length of 0.20 m (mesoscale effects are included; de Rooy and Holtslag 1999). A more extensive description of this site can be found in van Ulden and Wieringa (1996). The SBL height was observed with a sodar with an uncertainty of about 40% (VH96). This method differs from the one used at the two previously mentioned sites. However, Arya (1981) states that $h_{LLJ}$ is probably a suitable alternative to represent the observed boundary layer height with a sodar. In addition, Hicks et al. (1977) found no systematic differences between sodar observations and the nocturnal surface inversion, based on temperature profile information. We therefore assume the two methods ($h_{LLJ}$ and sodar) to be “equivalent,” although it is not necessarily true (see section 4).

The free-flow stability was obtained from the 160- and 200-m levels of the tower for cases in which $h$ was less than 200 m. When $h$ was larger than 200 m, we estimated $N$ above the SBL by a statistical relationship obtained from the CASES-99 (section 3a) observations: $N = 0.015 + 0.51N_{BL}$, in which $N_{BL}$ is $N$ from measurements made between the surface and the 200-m level. This method certainly introduces an additional uncertainty for $N$ but meanwhile offers the possibility to increase the number of available data points in this dataset.

d. SHEBA

The Surface Heat Budget of the Arctic (SHEBA) dataset (338 data points with stable stratification) was obtained over the Arctic ice pack north of Alaska between October 1997 and October 1998. During this period, the SHEBA ice station drifted from approximately 75°N, 144°W to 80°N, 166°W. Visual inspection of the profiles left 47 data points; the other data were rejected because of insufficient wind speed observations in the lowest part of the sounding or because of fog conditions (particularly in wintertime). The latter especially caused a large reduction of useful data. For this site, $h_{LLJ}$ was taken as the SBL height as observed from radiosondes (as discussed for the CASES-99 dataset), taking into account that this height should also be supported by a “discontinuity” in the $\theta$ profile. Forty-one data points were not saturated near the surface. After applying the same criteria for surface fluxes and free-flow stability as for CASES99, 20 data points remained.

e. GABLS LES

Next to datasets from measuring campaigns, we also used ensemble mean LES results from the GABLS intercomparison study (Beare et al. 2006) to verify our model. The LES case study was based on simulation of an Arctic SBL over an ice surface during 9 h, with 0.25 K h$^{-1}$ of prescribed surface cooling, and for a geostrophic wind speed of 8 m s$^{-1}$. After 9 h, a moderate SBL of about 180-m height developed. The advantage of LES is that the free-flow stability is prescribed ($N = 0.019$ s$^{-1}$) without any uncertainty, whereas in the observations $N$ is relatively uncertain. The most relevant
disadvantage is that the current LES dataset consists of only one “surface stability point.”

f. Data considerations

We stress that, although these datasets cover a broad range of stabilities, surface roughnesses, and latitudes, the datasets differ in averaging time of the surface fluxes. Different averaging times can affect the surface fluxes and consequently $h_{\text{mod}}$. Furthermore, one should be aware that the friction velocity at a local scale used in this study is not a priori the amount of friction that is felt by the whole boundary layer on a larger horizontal scale (Beyrich and Kotroni 1993). Mesoscale circulations are often reported in SBL studies (e.g., Mahrt and Vickers 2002), and these circulations can contribute considerably to the surface fluxes (VM04). Garratt (1982) found that in studies that compare modeled and observed SBL heights the largest uncertainty is associated with the uncertainty of surface fluxes, which act as input variables for the model calculations.

4. Evaluation and parameter estimation

a. Evaluation

Next we evaluate the performance of Eqs. (1) and (2) with the coefficients given in ZM96 and Zilitinkevich and Baklanov (2002) (see discussion in section 2). Furthermore, we also evaluate the simple estimate by Koracin and Berkowicz (1988), written as $h = 700u_{*}$. Note that a time scale is implicitly included in the numerical constant.

Table 2 summarizes model performance given by several statistical quantities: the mean absolute error (mae), the root-mean-square error (rmse), subdivided into a systematic part (rmse-s) and an unsystematic (rmse-u) part, the median of the absolute errors (meae), the fractional bias (FB), and the index of agreement (IoA; Willmott 1982). Note that the correlation coefficient $r$ is not a reliable parameter for model quality (in the case of strong bias, see below) and is omitted here. Good model performance will result in IoA $\geq 1$, a small mae and meae, and, in the ideal case, rmse-s close to zero and rmse-u equal to the total rmse. To avoid that too much weight is given to the larger datasets, we present the statistics per dataset separately. Apart from these statistical measures, scientific evaluation can be enhanced by examination of graphical data display (Willmott 1982).

The performance of the three-term multilimit formula Eq. (1) is shown in Fig. 2. For the Sodankylä data, $h$ is estimated reasonably well for thick SBLs, although the number of data points is limited in that range. In contrast, for shallow SBL heights, we find a clear offset: $h_{\text{mod}}$ is only several meters where $h_{\text{obs}}/H_{1100} = 50–80$ m. Similar results were found for Cabauw. For SHEBA, $h_{\text{obs}}$ is relatively shallow and is underestimated by Eq. (1) despite the high correlation between $h_{\text{mod}}$ and $h_{\text{obs}}$. The IoA ranges from 0.62 to 0.82 for this equation (Table 2). For CASES-99, $h_{\text{mod}}$ is nearly always smaller than $h_{\text{obs}}$ even for thick SBLs (FB $= -0.91$). This result agrees with the results of VM04, who found a bias of $-68\%$ with Eq. (1) for CASES-99. They conclude that the introduction of the free-flow stability does not improve model performance (in contrast to our results in section 4), even in the case of marine boundary layers (CBLAST campaign) with a relatively strong free-flow stability aloft. For the LES model results, Eq. (1) gives $h_{\text{mod}} = 218$ m while $h_{\text{LES}} = 180$ m. Based on the rmse, Eq. (1) with coefficients from Joffre et al. (2001) results in similar performance as with the ZM96 coefficients.
With the coefficients in VM04 also a clear negative bias is found. This difference is perhaps also due to the different definitions used here and in VM04. Note that in the current analysis the impact of subsidence was not taken into account.

Figure 3 depicts the modeled and observed $h$ for the five-term multilimit model Eq. (2). For Sodankylä, $h_{\text{mod}}$ is systematically smaller than $h_{\text{obs}}$ for both deep and shallow boundary layers. Table 2 demonstrates that a large part of the rmse is of systematic nature in this case, and the FB is less than $0.48$. For the LES case, $h_{\text{mod}} = 90$ m while $h_{\text{LES}} = 180$ m. Except for Sodankylä, the IoA is smaller for Eq. (2) than for Eq. (1). The offset for shallow SBLs, as with Eq. (1), persists in Eq. (2). In general, Eq. (2) underestimates $h_{\text{obs}}$ by a factor of 2, in agreement with VM04 who found a bias of $-78\%$ with Eq. (2). This was also found for other observational sites and for the LES data. In fact, the addition of the two terms that should account for the interaction among $f$, $N$, and $L$ causes the model performance to deteriorate! It seems that too much weight is given to the added (shortest) length scales with the inverse interpolation method. Especially for Eq. (2), the use of parameters with values given in Joffre et al. (2001) gave poor results. For example, the mean meae increased from $94.6$ m with the coefficients in ZM96 to $137.3$ m with the use of coefficients of Joffre et al. (2001).

Our conclusion is that the current parameter set in Eqs. (1) and (2) cannot properly predict the observed boundary layer heights. We will consequently undertake a recalibration of the coefficients in Eq. (1). Note that we also found a negative bias for shallow SBLs, and no unique parameter set could be obtained for Eq. (10) in Zilitinkevich and Esau (2003; not shown), which is an improved version of Eqs. (1) and (2). Before we proceed, we note that the simple empirical estimate of $h = 700u_*$ performs well overall.

b. Calibration

With the available dataset it is tempting to recalibrate the coefficients in Eqs. (1) and (2). Note that Zilitinkevich and Baklanov (2002) determined the numerical values of the coefficients among others on the same Cabauw dataset as in the current study. However, they limited themselves to $h < 200$ m, and therefore the obtained coefficients can only be representative for a small subset of the total dataset. Especially for $h < 200$ m, this Cabauw dataset contains a relatively large scatter in comparison with the scatter for $h > 200$ m. The use of this data subsection may hamper robust parameter estimation.

The calibration of coefficient $C_n$, which represents the constant of proportionality for a truly neutral boundary layer ($B_s = 0$ and $N = 0$), from atmospheric observations is difficult because, as a result of radiation divergence, subsidence, and so on, truly neutral boundary layers are generally absent in the atmosphere. Therefore, we prefer to determine $C_n$ from LES studies performed earlier by, for example, Mason and Thomson (1987). The same strategy was followed by ZM96. They find $C_n = 0.6$, and we will use this value from now on. Note that the precise value of $C_n$ does not seem to be very important in the atmosphere, where truly neu-
tral boundary layers hardly exist. From a scale analysis of the different terms of Eq. (1), using typical values of 
\( f = 1 \times 10^{-4} \text{s}^{-1}, \vartheta_{\|} = -0.024 \text{ K m s}^{-1}, u_\infty = 0.3 \text{ m s}^{-1}, \) 
and \( N = 0.03 \text{ s}^{-1}, \) we find that the three terms from left to right contribute for 3%, 25%, and 72% of the sum of the three terms, respectively.

The remaining coefficients in Eq. (1) (\( C_s \) and \( C_i \)) are calibrated with the Sodankylä dataset with a Monte Carlo approach (e.g., Franks et al. 1997). Several statistical quantities can be used as a measure for model quality. Because of the relatively large uncertainty in the observations, we prefer the median of the absolute error (meae) to prevent outliers from dominating the statistical quality parameter too much. Figure 4 shows a surface plot of the meae for parameters ranges \( C_i \in [5, 20] \) and \( C_s \in [0, 100]. \) The meae has a minimum for \( C_i = 11. \) These results are supported by calibration on the whole dataset. Our findings agree with the results of van Pul et al. (1994), Kitaigorodskii and Joffre (1988), VH96 (who found \( C_i = 7–13), \) and Joffre et al. (2001), who found \( C_i = 10.) \) VM04 proposed \( C_i = 15. \) Note that the variability in \( C_i \) may strongly depend on the wind shear across the SBL, as discussed in VH96.

However, no clear minimum is found in the contour plot along the \( C_i \) axis. That means that the model performance is insensitive for \( C_i \) for \( C_i > 40. \) This result also explains the large range of proposed values for \( C_i. \) Similar results were found using other statistical quality measures. Note that also no unique parameter combination for \( C_i \) and \( C_s \) was found with \( 0.1 < C_n < 0.7 \) (not shown).

To summarize, the multilimit equations show a clear bias against observations and their parameters cannot be determined uniquely. We will derive an alternative formulation using formal dimensional analysis instead of using inverse interpolation.

5. Alternative formulation using dimensional analysis

a. Three dimensionless groups

On the basis of the earlier works, we identify that the relevant quantities to describe \( h \) are \( u_\infty, f, B_s, \) and \( N. \) Using the Buckingham II theory (e.g., Langhaar 1951), we find three dimensionless groups:

\[
\Pi_1 = \frac{|B_s|}{hfu_\infty N} \quad \Pi_2 = \frac{k|B_s|}{u_\infty^3} = \frac{h}{L} \quad \text{and} \quad \Pi_3 = \frac{N}{f}.
\]

We may consequently determine the functional form of the surface that describes the relationship among \( \Pi_1, \Pi_2, \) and \( \Pi_3 \) from observations. It should be a universal relationship if all relevant quantities are included.

Figure 5a shows \( \Pi_1 \) versus \( \Pi_2 \) on a linear scale for different classes of \( \Pi_3 \) using Sodankylä observations. Despite the small number of data points per class, \( \Pi_2 \) clearly increases with \( \Pi_1 \) but levels off at different values for different classes of \( N/f. \) This relevance of \( N/f \) was already mentioned by Kitaigorodskii and Joffre (1988). Note that no data are available for \( N/f < 50 \) and \( N/f > 300. \) Furthermore, we remark that the plotted dimensionless groups in Fig. 5 have common terms and the risk of spurious correlation exists (e.g., Baas et al. 2006). However, it turns out that by randomizing the current datasets the scatter increases. Moreover, below we also use dimensional plots to confirm the performance.

On a log-log scale (Fig. 5b), we can determine the different slopes for different classes of \( N/f. \) We propose...
to fit the data according to $\Pi_1 = \alpha(\Pi_2)^{B_1}$. Applying this result and after some rearrangement we find for $h$:

$$h = L \left( \frac{g}{\rho \cdot w_*} \right)^{\lambda}$$

(3)

with $\alpha = 3, \lambda = \left[ C_1 - 0.001(N/f) \right]^{-1}$, and $C_1 = 1.8$. The calibration for $\alpha$ and $C_1$ is discussed in detail below; $L$ is the classical Obukhov length (thus including the von Kármán constant). Note that the innovative aspect of this equation is that the exponent is not constant, but it depends on one of the dimensionless groups. The numerical value 0.001 in $\lambda$ was found by plotting the slopes in Fig. 5b for different classes of $N/f$ (not shown). In addition, the obtained coefficients (based on Sodankylä observations) were confirmed by using half of every dataset for calibration and the other halves for validation.

Considering the applicability range of Eq. (3), we have to realize that the denominator in the exponent should be positive; hence, $N/f < 1800$. With $N = 0.076$ s$^{-1}$, the maximum free-flow stability in the dataset, this corresponds to a latitude $|\phi| > 16^\circ$. In addition, both $N$ and $L$ need to be larger than zero.

A Monte Carlo strategy, similar to the one in the previous section, was followed to estimate $\alpha$ and $C_1$ [on Sodankylä (Fig. 6a) and on the whole dataset (Fig. 6b)]. In this case, a clear minimum in the meae is found in the contour plots with $C_1 = 1.8$ and $\alpha = 3$ as optimal estimates and is confirmed using the other statistical quality measures (not shown). So, in contrast to Eqs. (1) and (2), the parameters in Eq. (3) can be determined with good confidence.

b. Verification

In this section we will verify the performance of Eq. (3) against the independent data from Cabauw, CASES-99, SHEBA (Figs. 7b–d; Table 3), and a cross validation for Sodankylä (Fig. 7a). The good performance for Sodankylä is obvious, because these are the same data as used for the calibration. Nevertheless, it seems that the data collapse onto a single curve. This is not trivial, and it gives confidence in the method and the variables that we selected. The model agrees well with the CASES-99 (rmse = 53.6 m but is largely unsystematic) and Cabauw (rmse = 80.3 m with rmse-u = 62.6 m) observations, although the scatter is larger for the Cabauw dataset than for the other datasets. This relatively large scatter is probably inherent to the sodar-based observations for Cabauw instead of radiosonde profiles for the other datasets (section 3). For SHEBA, the model performance is good (rmse = 40.3 m, of which 37.6 m is unsystematic), although the model underestimates the observations slightly (Fig. 5d). Overall, the magnitude of FB is less than 0.15, which is not achieved for any of the other proposals, and IoA is larger than for Eqs. (1) and (2). In general, the alternative proposal is superior to Eqs. (1) and (2) (see Table 2), where the main improvement by Eq. (3) is achieved for shallow boundary layers. For the LES model, Eq. (3) predicts $h = 173$ m, which is close to $h_{LES} = 180$ m. Given the typical uncertainties of $h_{obs}$ (30%), the model performs well, which gives us good confidence in the applied method.
c. On the relevance of the Coriolis parameter f: Two dimensionless groups only

The recent literature discusses the variables that govern \( h \) (Kosovic and Curry 2000; Kosovic and Lundquist 2004). Although \( f \) should theoretically play a role in governing \( h \) based on its presence in the Ekman equations (at least for neutral boundary layers), we can discuss the relevance of \( f \) as compared with \( N \) in practical application for the SBL as mentioned in VH96 and Zilitinkevich and Baklanov (2002). Because free-flow stability is always present in the atmosphere and is \( O(10^{-2}) \) while \( f \) is typically \( O(10^{-4}) \), VH96 suggest that the impact of \( f \) can be neglected in practice. Mahrt and Heald (1979) and VM04 also argue that the Coriolis parameter is of minor importance because the SBL development is governed by an inertial oscillation. In that case, a pure Ekman boundary layer does not exist and thus the use of \( f \) is doubtful.

In addition, if we analyze sodar observations during so-called intermittent nights (e.g., during CASES-99; van de Wiel et al. 2003), we find that the boundary layer turbulence and also the boundary layer height respond quickly (<10 min) to a change of the near-surface turbulence intensity if decoupling from the surface occurs. This result suggests that the SBL can react on a timescale much shorter than \( f^{-1} \), which is believed to be the governing timescale for the SBL (Nieuwstadt and Duynkerke 1996). This finding implies that \( f^{-1} \) is not a priori the most dominant time scale for the SBL growth but that time scales that originate from the interaction with the surface may be more important. Estournel and Guedalia (1990) and VM04 suggest that the roughness length for momentum \( z_0 \) is a relevant quantity. As a consequence, the relevance of \( f \) and \( z_0 \) will be examined below.

Because of the variable nature of \( h \), it is useful to adopt statistical techniques to gain insight into the relevant quantities that govern \( h \). In this section, we perform a principal component analysis (PCA) on \( h_{\text{obs}} \) from all datasets to obtain information about the relative impact of the different variables on \( h_{\text{obs}} \). Recall that PCA is a statistical technique in which the total variance of a dataset is decomposed along orthogonal vectors by determining the eigenvalues and eigenvectors of the covariance matrix among the variables. These eigenvectors are sorted in descending order according to the eigenvalues. The eigenvector associated with the largest eigenvalue is called the first principal component (FPC), the second largest is called the second principal component, and so on. Then the data are transformed back into real space and are correlated to the original variables.

Table 4 shows the absolute values of the correlation coefficients between the observed \( u_*, N, \overline{w \theta}_s, f, \) and \( z_0 \) and the FPC. The FPC explained 99.9% of the variance; therefore the higher principal components can be neglected safely. It appears that the correlation coefficient between \( u_* \) and the FPC is large relative to the other coefficients. The quantities \( \overline{w \theta}_s \) and \( N \) show a considerable smaller correlation but are still relevant. The Coriolis parameter and \( z_0 \) have a correlation coefficient of only 0.15 and 0.16 with the FPC, respectively. Therefore, the latter quantities are in practice relatively unimportant for \( h \) estimation (at least for these datasets). Note that the dominance of \( u_* \) gives support to the simple estimate by Koracin and Berkowicz (1988).

Given the discussion above, we may exclude \( f \) and \( z_0 \) from the list of relevant variables, at least for estimation of \( h \) in practical applications. Then only two dimensionless groups remain: \( hN/u_* \) and \( h/L^* \) (Fig. 8a, all data are used). Two regimes can be clearly distinguished. For
For $h/L^* < 1$ (toward the near-neutral limit), $h \propto u_*/N$, in accordance with Kitaigorodskii and Joffre (1988), van Pul et al. (1994), and VH96.

For $h/L^* > 1$ (toward the very stable limit), the two groups are linearly related on the log–log scale. This means that $h \propto (|B_s|/N^3)^{1/2}$. Although it seems that $h$ is independent of $u_*$ in this regime, this is not really the case. The dependence of $u_*$ comes in through $B_s$, because in the very stable regime the turbulent temperature scale $\theta_s$ is linearly dependent on $u_*$ (Holtslag and de Bruin 1988; van de Wiel 2002). Thus, making $B_s$ quadratically dependent on $u_*$ means that $h$ is again proportional to $u_*$, but now with a different factor depending on $N$. Thus, our diagnostic equation for $h$ based on the current analysis reads as

$$ h = \begin{cases} 
10u_*/N & \text{for } u_*/N|B_s| > 10 \\
32(|B_s|/N^3)^{1/2} & \text{for } u_*/N|B_s| < 10 
\end{cases}.$$

It is obvious that the application of Eqs. (3) and (4) is limited to cases in which $N > 0$. Equation (4) is well supported by Fig. 8a. Figure 8a also shows Eq. (1), neglecting the small contribution from the term $(f h)/(C_n u_*)$, for the parameter set proposed by ZM96 and Joffre et al. (2001). For strong stability (i.e., large $h/L^*$)
Eq. (1) does not correspond to the observations. Of interest is that the dimensional analysis (ignoring $f$) corresponds to the same dimensionless groups as in Eq. (1) if the $f$ group is neglected in the latter case! However, as before the important difference lies in the particular functional relationship between the group (Fig. 8a). In our case it has a power law (instead of a geometric average).

For stable conditions, the length scale $(|B_f|/N^3)^{1/2}$ in Eq. (4) has an equivalent form as the classical Ozmidov length scale $l_{oz} = (e/N^3)^{1/2}$, with $e$ being the TKE dissipation rate, above which eddies are deformed by stratification. Both scales represent the destruction of TKE during stable conditions. The final result in terms of Eq. (4) is given in Fig. 8b. Table 3 demonstrates that Eq. (4) performs well for CASES-99 and Cabauw (FB $< 0.15$ and IoA $> 0.75$). For Sodankylä, the rmse increases (although rmse-s $<$ rmse-u) and for SHEBA the performance degrades even more (IoA $= 0.13$). Thus, Eq. (3) performs better and seems to be more general, although Eq. (4) is physically more appealing.

### 6. Conclusions

This paper evaluates two so-called multilimit (diagnostic) equations for the stable boundary layer height $h$ proposed by Zilitinkevich and Mironov (1996) against four observational datasets and large-eddy model simulation. These equations are based on inverse interpolation of characteristic length scales that account for rotation, surface buoyancy flux, and free-flow stability, respectively. We found that the multilimit equations underestimate the stable boundary layer height, especially for shallow boundary layers. This condition is undesirable because shallow boundary layers will result in high pollutant concentration during episodes and a proper estimate is required. In an attempt to recalibrate
the parameters, we found that no unique parameter set for these multilimit equations could be determined.

To circumvent the difficulties with the multilimit equations, an alternative formulation is developed with formal dimensional analysis with the same quantities as in the multilimit equations. The proposed formulation is robust and appears to reduce a significant model bias for shallow boundary layer heights in comparison with the existing formulations. As such, the proposed formulation appears applicable also for high-stability conditions in contrast to existing formulations that primarily are derived for weakly stable cases. Contrary to the multilimit equations, a unique parameter set was found for this new formulation.

Last, we discuss whether the Coriolis force and the surface roughness length \( z_{0,m} \) are relevant parameters for estimating the stable boundary layer height in practice. When the Coriolis parameter \( f \) is omitted, the data reveal a division in two regimes. For moderately stable conditions the boundary layer height is proportional to \( u_s/N \); and for stronger stable conditions it appears that the boundary layer height is proportional to the length scale \( (B/vN^3)^{1/2} \). Overall, our analysis shows that the strongest relation of \( h \) exists with \( u_s \) (and \( B \) in the very stable limit), in addition to \( N \) while \( f \) and \( z_{0,m} \) are less relevant for the data utilized! (Here \( u_s \) is friction velocity, \( B \) is the surface buoyancy flux, and \( N \) is the free flow stability.)

Acknowledgments. The authors thank Sylvain Joffre and Markku Kangas for providing the Sodankylä dataset and for their careful selection of the dataset. We thank the SHEBA Atmospheric Surface Flux Group, Ed Andreas, Chris Fairall, Peter Guest, and Ola Person for help in collecting and processing the data. The National Science Foundation supported this research with grants to the U.S. Army Cold Regions Research and Engineering Laboratory, NOAA’s Environmental Technology Laboratory, and the Naval Postgraduate School. Furthermore, we thank our colleague Oscar Hartogensis for gathering the surface observations in CASES-99. We also acknowledge the GABLS community, whose LES model results have been used in this study. We also thank our colleague Dr. Leo Kroon and two anonymous reviewers for their valuable comments on our work.

REFERENCES


——, and H. A. R. de Bruin, 1988: Applied modeling of the night-


