Multisensor Estimation of Mixing Heights over a Coastal City

JOHN W. NIELSEN-GAMMON AND CHRISTINA L. POWELL*
Department of Atmospheric Sciences, Texas A&M University, College Station, Texas

M. J. MAHONEY
Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California

WAYNE M. ANGEVINE, CHRISTOPH SENFF, AND ALLEN WHITE
Cooperative Institute for Research in Environmental Sciences Climate Diagnostics Center, University of Colorado, and NOAA/Earth System Research Laboratory, Boulder, Colorado

CARL BERKOWITZ AND CHRISTOPHER DORAN
Pacific Northwest National Laboratory, Richland, Washington

KEVIN KNUPP
Department of Atmospheric Science, University of Alabama in Huntsville, Huntsville, Alabama

(Manuscript received 23 May 2006, in final form 15 March 2007)

ABSTRACT

An airborne microwave temperature profiler (MTP) was deployed during the Texas 2000 Air Quality Study (TexAQS-2000) to make measurements of boundary layer thermal structure. An objective technique was developed and tested for estimating the mixed layer (ML) height from the MTP vertical temperature profiles. The technique identifies the ML height as a threshold increase of potential temperature from its minimum value within the boundary layer. To calibrate the technique and evaluate the usefulness of this approach, coincident estimates from radiosondes, radar wind profilers, an aerosol backscatter lidar, and in situ aircraft measurements were compared with each other and with the MTP. Relative biases among all instruments were generally less than 50 m, and the agreement between MTP ML height estimates and other estimates was at least as good as the agreement among the other estimates. The ML height estimates from the MTP and other instruments are utilized to determine the spatial and temporal evolution of ML height in the Houston, Texas, area on 1 September 2000. An elevated temperature inversion was present, so ML growth was inhibited until early afternoon. In the afternoon, large spatial variations in ML height developed across the Houston area. The highest ML heights, well over 2 km, were observed to the north of Houston, while downwind of Galveston Bay and within the late afternoon sea breeze ML heights were much lower. The spatial variations that were found away from the immediate influence of coastal circulations were unexpected, and multiple independent ML height estimates were essential for documenting this feature.

1. Introduction

The depth of the mixed layer (ML), also known as the ML height or mixing height, determines the volume in which pollution emitted near the ground is primarily concentrated. Various imperfect observational tech-

* Current affiliation: Environmental Sciences Program, Wright State University, Dayton, Ohio.

Corresponding author address: John W. Nielsen-Gammon, Dept. of Atmospheric Sciences, 3150 TAMUS, College Station, TX 77843-3150.
E-mail: n-g@tamu.edu

DOI: 10.1175/2007JAMC1503.1

© 2008 American Meteorological Society
commonly used as a reference for evaluating other techniques, but radiosonde-based definitions of ML height vary from study to study and the particular technique used is not always described. Grimsdell and Angevine (1998), using an unspecified subjective technique for the radiosonde data, found a profiler bias of 39 m for clear conditions and 27 m for cloudy conditions in summertime in Illinois. White et al. (1999) compared profilers and lidar during summertime in Tennessee and found a lidar bias of 69 m and an RMS scatter about the regression of around 150 m. Van Pul et al. (1994) compared sonde (using the method of Holzworth 1964) and lidar from a representative sample of conditions throughout the year at the Netherlands and found a standard error of the regression of about 120 m. S. Cohn (2005, personal communication) has performed additional computations on profiler and lidar data from the field study previously analyzed by Cohn and Angevine (2000): using the radar wind profilers as ground truth, the bias of two ground-based lidars were −3 and 39 m and the standard deviations were 109 and 176 m. The most commonly used metric for comparisons, the correlation coefficient, does not convey information on the relative biases or accuracy of the different estimates.

This study evaluates the absolute and relative accuracy of the first ML height estimates made using an airborne microwave temperature profiler (MTP; Denning et al. 1989). The airborne MTP was flown for the first time for extensive boundary layer temperature sampling during the Texas 2000 Air Quality Study (TexAQS-2000), conducted in and around Houston, Texas, in August and September 2000. Preliminary to this evaluation, the relative performance of other conventional instruments (radiosonde, radar wind profiler, and airborne lidar) in Houston’s subtropical coastal environment is evaluated using a consistent and informative set of metrics. The value of a multi-instrument ML height dataset is then illustrated by a spatial and temporal analysis of ML heights during a high-ozone day in Houston. The combination of instruments provides a uniquely comprehensive view of a spatially inhomogeneous urban mixed layer.

2. Data and methods

The ML heights used in this paper were determined from data collected from five boundary layer profilers, three radiosonde sites, an airborne lidar, an airborne MTP, and in situ temperature and dewpoint measurements collected by instruments aboard the National Center for Atmospheric Research (NCAR) L-188C Electra aircraft. The positions of the wind profilers and radiosonde sites (as well as surface sites to be mentioned in section 6) are displayed in Fig. 1. The profilers operated continuously, and soundings were taken every 3–12 h. The Electra, which carried the MTP, most often flew at about 650 m above ground level (AGL), within the mixed layer, while the aircraft that carried the lidar generally flew above 3000 m AGL so that the downward-pointing lidar could sample the entire PBL. The two aircraft generally operated independently, performing multiple transits across the Houston area or across downwind plumes of pollutants.

The intercomparison dates are given in Table 1. Most days with airborne data were cloud-free or had scattered clouds that could be avoided by the lidar aircraft. In general, conditions in Houston during the field program were somewhat hotter and drier than normal, and the prevailing wind direction was most often southerly or northeasterly. Sea breezes were common on days with light winds.

In a previous field program, White et al. (1999) av-

<table>
<thead>
<tr>
<th>Comparisons</th>
<th>Days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wind profiler and sonde</td>
<td>17 Aug–19 Sep</td>
</tr>
<tr>
<td>Wind profiler and lidar</td>
<td>25 Aug, 28–30 Aug, 1 Sep, 6–7 Sep</td>
</tr>
<tr>
<td>Wind profiler and MTP</td>
<td>25 Aug, 27–28 Aug, 1 Sep, 6 Sep</td>
</tr>
<tr>
<td>MTP and in situ data</td>
<td>25 Aug, 27–28 Aug, 30 Aug, 1 Sep</td>
</tr>
</tbody>
</table>

FIG. 1. Map of the Houston area, showing locations of radar wind profiler, radiosonde, and surface stations mentioned in the text. The box in the center gives the rough outline of the densely populated core of the Houston metropolitan area and is approximately 60 km × 70 km.
eraged the ML height estimates from airborne instruments for each flight segment within a 20-km radius of each profiler site to produce airborne observations that were comparable to the profiler estimates. A similar technique is applied here, except that a radius of 10 km was found to give the greatest agreement between the lidar and profiler ML heights; that radius is adopted for the airborne MTP comparisons as well. Temporal collocation was assumed if the airborne observation was within 15 min of the fixed observation. The ML height agreement between the airborne instruments and the LAM and ELL wind profilers, located near the coast (Fig. 1), was found to be worse than with the other profilers. Because the effect of the sea–land boundary layer transition and the sea breeze itself could not be easily separated from effects due to different instrument characteristics, ML height comparisons with the LAM and ELL wind profilers are not shown.

The differences between techniques of ML height estimation are quantified below using bias and standard deviation. The bias \( b \) is the difference between the means of the paired samples. For each comparison, a reference technique is defined and the bias is positive if the mean estimates from the comparison technique are greater than the mean estimates from the reference technique. The standard deviation \( s \) is the root-mean-square value of the departures of the individual pair sample differences from the mean difference (or bias). The RMSE, or root-mean-square difference between the individual estimates, is directly computable from the bias and standard deviation: \( \text{RMSE} = (b^2 + s^2)^{1/2} \).

The statistical significance of the bias was estimated using a two-tailed, one-sample \( t \) test. The null hypothesis was a statement that the test instrument estimate of ML height is unbiased when compared with the reference estimate. The null hypothesis was not rejected when the calculated \( t \) statistical value was between \(-1.96 \) and \( 1.96 ( -1.96 \leq t \leq 1.96 ) \) and the \( p \) value was greater than 0.05.

### 3. Retrieval of mixing heights from airborne Microwave Temperature Profiler data

#### a. MTP operation and temperature retrieval

The Jet Propulsion Laboratory’s MTP, flown on the NCAR Electra during TexAQS-2000, is a passive microwave radiometer that measures the thermal rotational-line emission from oxygen molecules at three frequencies (55.51, 56.65, and 58.80 GHz). The instrument uses a scanning mirror to view 10 elevation angles from \(-80^\circ\) to \(+80^\circ\) in the flight direction (Denning et al. 1989).

The MTP-measured microwave brightness temperatures are converted to a vertical temperature profile along the flight track by using a statistical retrieval procedure (Strand and Westwater 1968) with Bayesian aspects. To do this, a forward radiative transfer calculation must first be made. Hundreds of soundings representative of the time and location of the actual measurements are used to calculate the expected brightness temperatures for each radiosonde temperature profile. A linear multiple regression is then used to statistically relate the expected brightness temperatures (from the forward radiative transfer calculation) to physical temperature profiles (from the soundings). This results in a set of retrieval coefficients that can be used to convert actual measured brightness temperatures to a physical temperature profile. An information-theory-based metric then compares the measured observables with the average calculated observables for each retrieval coefficient set, and determines which set (or pair of sets) of retrieval coefficients to use. [For further background on the temperature retrieval technique, see Gary (1989, 2006).]

Except for a few short excursions into the PBL, MTP flights have been generally in the upper troposphere and lower stratosphere (UT–LS), which is where the instruments were designed to fly. Because of the short time available to prepare for TexAQS-2000, the existing DC-8 MTP was flown unmodified. In hindsight, better performance might have been obtained by scanning more rapidly to minimize horizontal temperature fluctuations within a scan, by using different (or more) absorption frequencies, by employing a scan strategy optimized for detecting boundary layer structure, and by using a larger mirror to reduce beamwidth and thereby improve vertical resolution.

When flying in the UT–LS, the MTP does not measure emission originating from near the earth’s surface, and therefore measurements are not influenced by either the enhanced daytime surface temperature there (relative to the air temperature), or the substantial variation in emissivity over land (and to a lesser degree over water). The situation is very different for flight in the PBL. To avoid the air temperature retrievals being affected by the combination of surface temperature and emissivity, we downweighted any measurements that did not have at least five \( e \)-folding distances of absorption from the MTP’s line of sight to the surface. Unfortunately, this approach throws away potentially useful information needed to understand surface forcing and can produce artifacts in the retrieval because of the correlation of temperature with altitude in the atmosphere.
For a statistical retrieval to be effective, it is essential that the soundings used to calculate retrieval coefficient sets resemble the actual temperature field being measured. With any soundings, whether observed, modeled, or hypothesized, the information-theory-based metric provides information on whether the soundings are sufficiently consistent with the MTP observations to serve as starting points for the retrievals. Since the Electra generally flew between 1500 and 2300 UTC (0900–1700 LST), soundings from nearby radiosonde sites at 0000 and 1200 UTC did not capture all of the relevant temperature structure, especially ground-level and elevated temperature inversions. The radiosondes also did not sample typical coastal and overwater thermodynamic profiles. To overcome this limitation, synthetic soundings were generated from mesoscale model simulations [fifth-generation Pennsylvania State University–NCAR Mesoscale Model (MM5), 4-km grid spacing, 42 vertical levels with half located below 3 km] for six of the Electra flight days to supplement actual soundings from HDT and LAM, to ensure that sufficiently representative soundings were available to calculate retrieval coefficients.

Five sets of retrieval coefficients were needed to represent the temperature fields encountered by the Electra: an ocean set with no excess skin temperature, two sets representing well-mixed profiles over land, with random excess skin temperatures averaging 20°C more than ground temperatures added at the sounding’s ground level to mimic the excess surface emission, and two sets to represent surface-based inversions and inversions above 150 m.

b. ML height estimation from MTP

The vertical spacing of the MTP retrievals increases with distance from the aircraft flight level: 100 m near aircraft flight level, 300 m at 1 km above or below flight level, 1 km at 4 km above or below flight level, and so on. To retain only those retrievals providing adequate resolution near the expected mixing height, data from flight levels above 2.8 km or below 0.2 km were excluded. Also excluded were data from aircraft ascents, descents, and turns, as attitude changes affect the viewing angle of the MTP and degrade the retrievals.

Data from the remaining scans were linearly interpolated to 100-m vertical intervals. An example of interpolated potential temperatures from a flight leg on 1 September 2000 is shown in Fig. 2. The rapid, vertically correlated potential temperature variations occur primarily over the city and are attributed to contamination from surface emissions, the finite duration of scans, and other sources of instrument noise. To eliminate these artifacts prior to computing ML heights, nine-point horizontal moving averages were computed from successive scans. The endpoints of the nine-scan sequence were constrained to be within approximately 15 km of each other to avoid averaging across large distances.

The direct application of subjective mixing height al-

![Fig. 2. Vertical section of potential temperature, 1730–1800 UTC 1 Sep 2000, as retrieved by airborne MTP. Isentropes are every 0.5 K and increase from blue to red, with thick lines for the 308- and 312-K isentropes. Noise near the middle of the segment is most likely an artifact of rapidly varying surface temperatures and emissivities in the Houston urban core.](image-url)
algorithms based on radiosonde data (e.g., Kaimal et al. 1982) is not appropriate, because the MTP retrievals have much coarser vertical resolution than a typical radiosonde. The “blurry” data from the MTP obscures otherwise clear transitions between neutral and stable stratifications. Thus, we test ML height estimates from various modified versions of two objective radiosonde-based methods (Marsik et al. 1995) that we will call the theta-increase method (Heffter 1980; Marsik et al. 1995) and the T-lapse rate method. Application of these techniques to a sample MTP temperature profile is illustrated in Fig. 3.

The T-lapse rate method, as applied by Garrett (1981), assigns the ML height to the base of the first layer with a lapse rate more stable than 8 K km$^{-1}$, equivalent to a vertical derivative of potential temperature greater than 2 K km$^{-1}$. As with any approach based on a vertical derivative, this approach is sensitive to the vertical resolution of the underlying data. To find appropriate lapse rate criteria for the coarse (100 m or worse) MTP data, a variety of algorithms were tested. Two techniques for defining the lapse rate were tested: the lapse rate over the 100-m-thick layer above a given point, and the least stable lapse rate found by varying the thickness of the layer above a given point. Two techniques for specifying the lapse rate criterion were also tested. The first was a simple numerical value for the vertical derivative of potential temperature (1.5, 1.75, 2.0, 2.25, 2.5, and 3.0 K km$^{-1}$), and the second was a value computed by multiplying the average vertical derivative of potential temperature from 2.0 to 3.5 km by a constant factor (0.3, 0.5, 0.7, 0.8, 0.9, and 1.0). This combination of techniques and thresholds yields a total of 24 different estimates.

The theta-increase method, described by Heffter (1980), Marsik et al. (1995), and Seibert et al. (2000), is based on finding the level of maximum ascent of a low-level air parcel. When applied to radiosonde data, a surface-based parcel is commonly used, but the MTP does not retrieve temperatures well near the ground. As applied here, the ML height is defined as the level at which the observed potential temperature first exceeds the minimum potential temperature within the boundary layer by some threshold amount. Six different values of the threshold potential temperature exceedance are tested (1.0, 1.25, 1.5, 1.75, 2.0, and 2.25 K).

The various ML height estimates from the MTP are tested against profiler ML height estimates to determine the optimal technique and threshold (section 5c).

4. ML height estimates from other instruments

a. Radiosondes and in situ aircraft

The radiosonde-based estimates are assumed here to provide benchmark estimates of the “true ML height” because they incorporate the vertical profiles of multiple thermodynamic quantities at high vertical resolution. While these estimates are assumed to be bias-free, they are subject to sampling error caused by a radiosonde passing through an individual thermal, updraft, or downdraft that is exceptionally strong, yielding higher or lower estimates of the ML height than a different instrument whose estimate is averaged temporally and spatially.

Soundings were plotted on skew T diagrams for subjective identification of the ML height (Fig. 4). An ideal sounding will have an ML with a constant mixing ratio. If the mixing ratio was not well mixed, the ML was taken as the layer where the mixing ratio was nearly constant. The base of the moisture transition layer was then defined as the first level in the dewpoint profile with a prominent decrease in mixing ratio above the ML. The ML height was defined as the point halfway between the base and top of the moisture transition layer. An ML height estimate was assigned a high-quality flag when the dewpoint profile was nearly constant within the ML and the ML height estimate was unambiguous. When the moisture transition layer was
ill defined but a well-defined stable layer was present, the base of the stable layer was taken as the ML height.

ML heights were estimated from in situ aircraft data with the same procedure, using ascent–descent segments of the flights. Measurements included GPS altitude, temperature (Rosemount Type 102, unheated), and dewpoint (General Eastern 1011B thermoelectric hygrometer). The dewpoint sensor uses the chilled mirror technique, so dewpoint changes greater than 1 K s\(^{-1}\) are unresolved. Data were available at 1-s intervals. The typical ascent–descent rate of the Electra was 7 m s\(^{-1}\) and typical airspeed was 120 m s\(^{-1}\), yielding a glide slope of 0.06.

b. Radar wind profilers

The radar reflectivity is computable from the signal-to-noise ratio (SNR) measured by the wind profilers. A peak in reflectivity is found at the entrainment zone. Reflectivity peaks can also be caused by other phenomena (Angevine et al. 1994; White et al. 1999; Grimsdell and Angevine 1998). To date, effectively distinguishing the ML height signature from contaminants has been done by manual methods; automated algorithms are under development.

ML heights were estimated from the wind profilers by manually examining the reflectivity and spectral width patterns of each beam. Based on these patterns, the peak of reflectivity corresponding to the ML top was chosen. The resolution of an individual reflectivity measurement is 60 m. To obtain robust estimates, the reflectivity peaks were averaged by eye over an hour (approximately 90 separate radials), and the resulting estimates are reported at a resolution of 10 m. An example of the ML height estimates from one profiler on a typical day is shown in Fig. 5.

A “good” quality flag was given to those data based on clear, well-defined patterns, and a “marginal” quality flag was given to periods when the value was, in the analyst’s judgment, likely but not certain to be the correct ML height. Patterns showing poor definition of the ML height were flagged as poor quality and not included in these comparisons. The hour-long averaging could introduce inaccuracies if the change of height with time during the averaging period is strongly nonlinear.

c. Airborne aerosol backscatter lidar

Backscatter lidar systems observe the distribution of particulate matter in the ML. A commonly used method to determine ML heights from backscatter lidar profiles is to find a maximum gradient in the lidar backscatter signal associated with the decrease in aerosol backscatter in the transition zone from the ML to the free atmosphere (Davis et al. 2000). This approach is followed here (an example is given in Fig. 6), after first specifying a maximum ML height based on subjective examination of the backscatter data in space and time.

The vertical resolution of the lidar estimates was approximately 15 m, and the horizontal resolution was about 600 m. Figure 6 provides a snapshot of the spatial distribution of mixed layer depth in the Houston area: over the Gulf of Mexico (left side of panel) the boundary layer was about 500 m deep, whereas the mixed layer depth increased rapidly to about 1800 m as the lidar aircraft crossed the coastline and proceeded toward Houston (right side of panel).
High backscatter signals received from cloud tops can be mistaken for the top of the ML, which can create a positive bias in the ML depth if cloud signals are not otherwise detected and removed. Also, as with other techniques involving constituent mixing, errors can result when an internal boundary layer is present, and the lidar measures the top of a residual layer as opposed to the current surface mixed layer. Residual layers occur after the convective PBL collapses in the afternoon or when the continental ML is undercut by a sea breeze, and contain properties of the past ML air that is unrepresentative of the present ML. To minimize this problem, the TexAQS-2000 lidar data were prescreened to constrain the objective algorithm to identify the lowest aerosol transition layer if multiple such layers were seen.

5. Instrument intercomparison

a. Background

Five boundary layer wind profilers, an airborne lidar, and the airborne MTP were used to estimate the ML height. These instruments have different strengths and are often only appropriate under certain conditions. Even under optimal conditions, ML height estimates can differ because each instrument measures a different aspect of the ML-to-free-atmosphere transition.

Because wind profilers are stationary instruments that provide nearly continuous measurements of the ML, their ML height estimates can be compared directly with the estimates of the airborne instruments when the aircraft flew suitably close to profiler sites. To relate these comparisons to radiosonde estimates of the “true ML height,” the profiler estimates are first compared with the sonde benchmark estimates. Unless otherwise stated, the bias is defined relative to the wind profiler ML height estimates. So, for example, if another instrument yields an estimate higher than that of the wind profilers, the bias would be positive.

b. Wind profiler versus radiosonde benchmark

To make comparisons with the radiosonde benchmark method, ML heights were obtained from wind profiler data for the closest time following each radio-
sonde release. Both WPP and LAM had wind profilers collocated with radiosonde release sites.

The results of the statistical comparison are presented in Table 2 and shown graphically in Fig. 7. Overall, the ML heights estimated by the wind profiler data had a bias of at most $0.08$ km relative to the benchmark-method ML heights and the differences had a standard deviation of $0.27$ km or less. Evaluation of the LAM and WPP comparison together illustrated that the results were mainly dependant on the quality flag of the benchmark method heights. When only the high-quality benchmark method heights were used, the bias between the two instruments was $10$ m. In almost all comparisons shown in Table 2, the bias was negative, indicating the ML heights estimated from the wind profiler data were higher on average than the ML heights estimated from the benchmark method. None of these biases are statistically significant, which is consistent with past research that also found good agreement between these two instruments and the wind profiler heights to be slightly higher than those determined from radiosonde data (Grimsdell and Angevine 1998; Angevine et al. 1994). The small value of the bias is sensitive to the choice of radiosonde ML height estimation method and would increase by $50$ m if the profiler heights were compared with ML heights defined as the base of the stable layer.

The comparisons between the wind profilers and the benchmark method were consistent over all measured heights and between profilers. The greatest agreement, the good-quality estimates from LAM, occurred with a sample size of only five. ML height estimates from the coastal location agreed as well as height estimates from the inland location. The biases were not statistically significant. The small bias and apparently reasonable standard deviation justifies use of the profiler estimates for comparison with the airborne instruments.

c. Wind profiler versus airborne instruments

The results of the statistical comparison between the WPP, HSW, and LIB (noncoastal) profilers and the airborne instruments are presented in Table 3 and shown graphically in Fig. 8.

The comparisons between the wind profilers and the airborne lidar were consistent over the full range of ML heights. The ML heights measured by both instruments are consistent with past research that also found good agreement between these two instruments and the wind profiler heights to be slightly higher than those determined from radiosonde data (Grimsdell and Angevine 1998; Angevine et al. 1994). The small value of the bias is sensitive to the choice of radiosonde ML height estimation method and would increase by $50$ m if the profiler heights were compared with ML heights defined as the base of the stable layer.

The comparisons between the wind profilers and the benchmark method were consistent over all measured heights and between profilers. The greatest agreement, the good-quality estimates from LAM, occurred with a sample size of only five. ML height estimates from the coastal location agreed as well as height estimates from the inland location. The biases were not statistically significant. The small bias and apparently reasonable standard deviation justifies use of the profiler estimates for comparison with the airborne instruments.

### Table 2: The std dev, bias, and number of data points for comparison of benchmark method with profiler estimates. Bias is defined as sonde estimate minus profiler estimate. Data are divided into high or good quality estimates only and all available estimates.

<table>
<thead>
<tr>
<th>Sonde benchmark</th>
<th>Quality</th>
<th>Std dev (km)</th>
<th>Bias (km)</th>
<th>No.</th>
<th>Std dev (km)</th>
<th>Bias (km)</th>
<th>No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>LAM</td>
<td>High</td>
<td>0.11</td>
<td>0.00</td>
<td>5</td>
<td>0.20</td>
<td>0.08</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>0.26</td>
<td>-0.08</td>
<td>17</td>
<td>0.27</td>
<td>-0.07</td>
<td>22</td>
</tr>
<tr>
<td>WPP</td>
<td>High</td>
<td>0.23</td>
<td>-0.02</td>
<td>13</td>
<td>0.22</td>
<td>-0.04</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>0.21</td>
<td>-0.01</td>
<td>15</td>
<td>0.24</td>
<td>-0.01</td>
<td>18</td>
</tr>
</tbody>
</table>

### Table 3: The std dev, bias, and number of data points for comparison of airborne instrument estimates with profiler estimates. Bias is defined as airborne estimate minus profiler estimate.

<table>
<thead>
<tr>
<th>Airborne instrument</th>
<th>High quality</th>
<th>All quality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Std dev (km)</td>
<td>Bias (km)</td>
</tr>
<tr>
<td>Lidar</td>
<td>0.20</td>
<td>-0.06</td>
</tr>
<tr>
<td>MTP</td>
<td>0.10</td>
<td>-0.07</td>
</tr>
</tbody>
</table>
were within a few meters of each other with a slight bias, not statistically significant, toward higher ML heights estimated from the wind profiler data and a standard deviation of about 0.2 km. One exceptional data point has a lidar ML height 0.6 km higher than the corresponding profiler ML height. A few other similarly biased outliers were found at the coastal profiler sites (not shown). A few of these differences appear to be due to the spatial variations in ML heights associated with the sea-breeze front, while others may be due to the lidar detecting the top of the residual layer rather than the surface-based ML, or they may be due to ambiguities in boundary layer height retrieval caused by poor contrast in aerosol loading between the ML and the lower free troposphere. One example of ambiguity is shown in Fig. 9, where there is considerable aerosol loading aloft as well as within the ML, and air with high aerosol content between 0.8 and 1.3 km appears to be above the ML between 1736 UTC and 1739 UTC and within the ML after 1743 UTC.

The MTP method producing ML height estimates most similar to profiler ML height estimates was the theta-increase method. The smallest standard deviation (0.12 km) was found with a theta-increase value of 1.75 K. A standard deviation nearly as small (0.13 km) and a much smaller bias (−0.01 km versus −0.07 km) was found with a theta-increase value of 1.5 K; for simplicity (given the inaccuracies associated with a small sample size) we choose 1.5 K for the theta-increase value. The next-best-performing algorithm was the $T$-lapse rate method using an arbitrarily specified lapse rate criterion and a variable layer over which the lapse rate is measured; the best criterion was a potential temperature vertical derivative more stable than 3 K km$^{-1}$. This algorithm produced a standard deviation of 0.21 km and a bias of −0.03 km. The $T$-lapse rate techniques did not work as well because they were very sensitive to details of the stability profiles, and apparently, the coarse vertical resolution of the MTP (relative to radiosondes) did not resolve stability variations to sufficient accuracy.

The MTP results with a theta-increase criterion of 1.5 K are shown in Fig. 8 and Table 3. The standard deviation and RMSE values found for the comparisons between the MTP and wind profilers were lower than the values found from comparisons between the wind profilers and airborne lidar and the wind profilers and radiosonde benchmark. Because of small sample sizes, this difference is not significant and may not be reproducible; different potential temperature criteria may perform better in other circumstances. The differences between the wind profilers and the MTP were consistent over the full range of ML heights.

d. Airborne MTP versus in situ observations

Along with the inland wind profiler heights, the MTP heights were also compared with ML heights determined by in situ Electra aircraft data using temperature and dewpoint profiles. The MTP was located aboard the Electra, so collocated observations were available when the aircraft ascended or descended through the ML top before or after level flight. The purpose of this comparison was to perform an independent check on the MTP heights, since the airborne MTP ML height algorithm was tuned to agree with profiler estimates. The portions of aircraft flight over Galveston Bay and the Gulf of Mexico were excluded.

The bias and standard deviation values were defined as the MTP estimates minus the Electra estimates. Overall, the comparisons showed good agreement (Table 4; Fig. 10), with minimal bias and standard deviations less than 0.25 km. The intercomparison was generally better when the MTP estimate was more nearly collocated with the in situ aircraft estimate, but the smaller sample size under a stringent collocation criterion inhibits confidence in the results. The MTP-aircraft comparisons were consistent over a range of heights, further supporting the use of MTP data to infer ML heights.

e. Comparison with prior results

In general, the results here are consistent with comparable results reported elsewhere. The profiler–sonde
bias is similar to that found by Grimsdell and Angevine (1998). The lidar biases and standard errors are similar in magnitude to those found by White et al. (1999) and van Pul et al. (1994), taking into account that the standard error of the regression should be somewhat smaller than the standard deviation. White et al. found a positive bias; a negative bias was found here after a few positive outliers were excluded. The ground-based lidars studied by Cohn and Angevine (2000) agreed better with profilers (S. Cohn 2005, personal communication) than did the airborne lidar studied here.

The standard deviation of the benchmark ML heights relative to the profiler ML heights was larger than the standard deviation with respect to the airborne instruments. When combined with the benefit found here of spatial averaging of the airborne data, the evidence suggests that at least part of the large standard deviation is caused by the radiosonde ML heights being point mea-

**Table 4.** The std dev, bias, and number of data points for comparison of MTP estimates with in situ aircraft estimates. Bias is defined as MTP estimate minus in situ estimate.

<table>
<thead>
<tr>
<th>MTP-in situ collocation distance (km)</th>
<th>Std dev (km)</th>
<th>Bias (km)</th>
<th>No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>0.19</td>
<td>0.01</td>
<td>19</td>
</tr>
<tr>
<td>16.8</td>
<td>0.16</td>
<td>0.01</td>
<td>14</td>
</tr>
<tr>
<td>10</td>
<td>0.13</td>
<td>0.06</td>
<td>7</td>
</tr>
</tbody>
</table>
measurements, while the other measurements involve averages in space or time. It seems that part of the standard deviation, perhaps 0.1–0.2 km, reflects the turbulent variability of the instantaneous ML height. This inference is supported by in situ measurements of temperature, humidity, and wind variability by the Electra aircraft when flying just above the top of the mean ML height (not shown).

On average, the ML heights estimated using the airborne lidar data were lower than the ML heights estimated using the wind profiler data by 50 or 60 m. This bias was similar to the bias of the benchmark method relative to the wind profilers, suggesting that the airborne lidar and benchmark methods are nearly unbiased relative to each other. The MTP ML height estimation technique was tuned to be unbiased with respect to the wind profiler estimates, but the standard deviation with respect to the profilers and the bias and standard deviation with respect to the in situ aircraft measurements imply that the MTP ML height accuracy was at least as good as that of the other tested instruments.

6. Spatiotemporal analysis of ML heights

With boundary layer airborne MTP measurements having been established as a viable and accurate technique for estimation of ML heights, an integrated analysis of ML heights on a particular day during the TexAQS-2000 field program is now conducted.

a. Event overview

The period 25 August 2000 through 1 September 2000 was a period of frequent exceedances of the federal ozone standard in the Houston metropolitan area. Intensive field activities were conducted during that period, and intensive mesoscale photochemical modeling efforts have been conducted for regulatory purposes. The date of 1 September 2000 is chosen for ML height analysis here because it included simultaneous comprehensive PBL measurements by both the airborne MTP and the airborne lidar. The five profilers were also operational, and soundings were taken at 3-h intervals at LAM and 6-h intervals at WPP.

September 1 featured light to moderate westerly winds throughout the day, with a weak sea breeze developing near the coast during the afternoon (Fig. 11). Westerly winds are unusual in southeast Texas in the summertime, and the continental airflow allowed daytime temperatures to approach or exceed 40°C (Fig. 12). The westerly winds advected night and morning pollution to the east from the Houston metropolitan area, including from the concentrated petrochemical and industrial emission sources in the Ship Channel area (Ryerson et al. 2003). Hourly averaged ozone values exceeded 140 ppb at two locations on the eastern edge of the Houston metropolitan area and three locations in the Beaumont–Port Arthur area 80–150 km downwind of Houston. Photochemical simulations generally carried the Houston ozone plume too far to the southeast.

b. Temporal ML height variations and adjustments to asynchronous observations

Figure 13 shows the diurnal variation of ML heights according to the profiler observations on 1 September 2000. ML heights increased slowly until about 1800 UTC and rapidly thereafter until leveling off around 1.5–2.0 km, although the upper bound of the ML height is based largely on “marginal” estimates. The evolution of ML heights contrasts with the more normal evolution shown in Fig. 5. Ordinarily, the ML height growth is most rapid in late morning, when both strong insolation and a shallow ML favor rapid growth. Instead, several profilers indicate that the ML height hardly changed at all in late morning on 1 September.

The cause of the atypical ML height behavior is seen in the morning sounding from WPP (Fig. 14). Instead of the typical surface-based nocturnal inversion, an isothermal bottom layer was overlaid by a strong elevated inversion between 980 and 940 hPa (300–670 m). Above the inversion, the lapse rate was nearly dry adiabatic. Consequently, growth of the ML was slow.
throughout the morning until the ML broke through the inversion, and rapid thereafter.

In section 6c, the ML height observations taken by the MTP and lidar during various flight segments will be displayed. However, because the ML heights are rapidly evolving during much of the day, temporal ML height variations would mask the spatial pattern of ML heights that might have been observed by instantaneous observations. To make ML height observations from different times comparable to each other, it is necessary to correct for the changes in ML height that occurred between the times of the observations and the nominal time of the plot.

The wind profiler measurements are useful for this purpose because they provide essentially continuous measurements of ML height variations through most of the day. A scatterplot of ML height versus subsequent half-hour ML height change (Fig. 15a) shows the general tendency for ML height changes later in the day to be larger than the earlier ML height changes, but with
considerable scatter. Part of the cause of the scatter is that the ML broke through the inversion at different times in different locations and attained its maximum depth at different times as well.

Figure 13 suggests that the elevated inversion, while possibly variable in intensity, was at roughly the same height at all profiler sites. Thus, a better approach is to regard changes in ML height as a function of the ML height itself rather than time of day: if the ML top is within the inversion layer, growth should be slow, while if the ML top is above the inversion layer, growth should be rapid. As seen in Fig. 15b, there is somewhat less scatter when ML height change is plotted against ML height itself. Based on the data displayed in Fig. 15b, the following ML height growth is used to adjust asynchronous observations to a nominal observation time: 85 m per 0.5 h when the ML height is less than 500 m; 50 m per 0.5 h when the ML height is between 500 m and 700 m, and 150 m per 0.5 h when the ML height is greater than 700 m. These corrections are applied (iteratively if necessary) for every MTP or lidar observation taken before the nominal map time but within
the specified time window of ±1.5 h. Similar corrections are developed and applied to MTP or lidar observations taken after the nominal map time but within the time window.

A lower bound on the errors of the time-adjusted data may be estimated by applying the adjustments to the profiler observations themselves. The resulting RMS errors are 0.07 km for 0.5-h adjustments, 0.16 km for 1.0-h adjustments, and 0.23 km for 1.5-h adjustments. (RMS errors without time adjustments are 0.12, 0.24, and 0.34 km, respectively.) Thus, somewhere around 1.0–1.5 h, the errors associated with using asynchronous observations should become as large as the errors associated with the individual measurements themselves. Errors will be significantly larger near large bodies of water, where the passage of the sea-breeze front would produce a drop in ML height instead of the assumed steady ML height increase. Within the profiler network, time-adjusted ML height errors at LAM associated with sea breeze front passage are measured at 0.9–1.1 km. Away from coastal effects, the RMS adjustment errors drop to 0.11 km at 1 h and 0.13 at 1.5 h.

c. Spatial variations of ML height

Airborne MTP observations are available between 1532 and 2052 UTC, and airborne lidar observations are available between 1513 and 2151 UTC. Hourly plots were generated and interpreted; shown here (Figs. 16a–c) are plots with data adjusted to the nominal times of 1600, 1800, and 2000 UTC. The size of each data point corresponds to how close the observation is to the nominal time. The two quasi-continuous streaks of dots are the airborne measurements; the lidar measurements are distinguished graphically by the addition of a continuous thin black line and the lack of temporal averaging. The wind profiler and in situ aircraft sounding estimates are shown as isolated dots within black circles; the five large dots in Fig. 16a correspond to the five profilers. The approximate location of the Houston metropolitan area is represented as a brown box with dimensions of 55 km × 45 km.

![Fig. 16. Spatial plots of ML height estimates from profilers, airborne lidar, airborne MTP, and in situ aircraft, for (a) 1600, (b) 1800, and (c) 2000 UTC 1 Sep 2000. Only observations within 1.5 h of the nominal time are shown. The size of the dot or width of the ribbon is proportional to the closeness of the observation time and the nominal map time. The green lines show rivers and coastal features, while the brown box outlines the Houston metropolitan area. A thin line is overlaid on lidar observations, while black circles surround profiler and in situ aircraft observations.](image)
The raw lidar estimates permit identification of the instantaneous ML height through the detection of individual up- and downdrafts. Therefore, the lidar estimates may fluctuate by a few hundred m over a distance of several kilometers, causing the plots of the lidar estimates to appear noisier than those of the MTP. The lidar estimates are not spatially averaged here because the raw estimates provide a useful sense of the small-scale spatial variability of the ML height.

At 1600 UTC (Fig. 16a), most ML heights are between 300 and 700 m. The lower ML heights are found over Galveston Bay and the nearshore Gulf of Mexico. The MTP ML heights indicate a continuous variation with distance from landmasses, being about 600 m over land and near the coast and 300 m over open water. The lidar ML heights offshore are in general agreement but tend to jump between 300 and 600 m. With moderate winds from the west, there would be little to distinguish the young marine layer from the residual aerosol-rich internal boundary layer aloft.

Agreement between the airborne instruments and both profilers and in situ aircraft soundings is excellent at this time, with difference generally being much less than 100 m. North of Houston, one MTP pass indicates a small patch of ML heights in excess of 1200 m. Absent supporting information, this patch might be regarded as an aberration, but analysis of subsequent times indicates that this patch is real and corresponds to the initial puncturing of the elevated inversion by the growing ML.

At 1800 UTC (Fig. 16b), the multiple instruments provide excellent coverage of Houston and its surroundings. ML heights have increased everywhere except over water; the lack of change over water is seen by the diagonal MTP pass at 1600 UTC and the north–south MTP pass at 1800 UTC. (The airborne lidar made a single set of passes over Galveston Bay at around 1700 UTC; an artificial change is introduced by spatially uniform application of the time corrections.)

The most prominent feature in the ML height pattern is the strong gradient between a very deep ML north of Houston and a shallower ML within and south of Houston. According to the MTP measurements closest to the nominal observing time, the ML height exceeded 2500 m to the north while only reaching 650 m over Houston itself. No profilers are sufficiently far north to provide corroboration, but an airborne lidar pass close to an hour later made time-corrected ML height measurements of over 2000 m. Two profilers (WPP and LJB) are within the gradient of ML heights observed by the MTP, and both are locally consistent with the MTP measurements. The temperature structure in the soundings (such as Fig. 14) and the evolution of the profiler ML heights (Fig. 13) are consistent with rapid variations of ML height once the height exceeds 700 m.

Figure 2 shows the MTP potential temperature profiles for the northward leg that passes directly over the ELL profiler site in eastern Houston. The potential temperature structure has the overall appearance of a sea breeze advancing from the left and reaching beyond the center of the vertical section. To the north of the apparent sea breeze, the ML is very deep, as indicated by the gradual vertical variations of retrieved potential temperature. Surface temperature observations at this time (Fig. 12) are consistent with the near-ground temperature variations estimated by the MTP. However, surface wind and temperature observations confirm that a sea breeze had not yet developed.

Determining the cause of this ML structure is beyond the scope of this paper. Among the possibilities are 1) a residual sea-breeze inversion from the previous day’s sea breeze, 2) spatial variations in the intensity of an inversion of remote origin, 3) a surface heating contrast caused by the observed presence of scattered to broken middle- and upper-tropospheric clouds over the Houston area and clear skies to the north, and 4) a surface heating contrast caused by a pine forest with sandy soils to the north and urban and grassland ecologies with clay soils to the south. Regardless of its cause, the complex ML height structure at 1800 UTC illustrates the importance of ML height observations that span the area of interest, as opposed to reliance on one or two fixed measuring sites.

Differences among various ML height estimates within the Houston area probably do not reflect real ML height variations. The lidar and MTP estimates in Houston are systematically contradictory, with lidar heights higher than MTP heights except for a single MTP pass through west-central Houston. Two of the profilers are in better agreement with the MTP, while one is in better agreement with the lidar. According to the profiler time series (Fig. 13), 1800 UTC was a time of sudden growth in the depth of the ML, and this rapid growth may have produced differing signals in the lidar and MTP measurements. Another possibility is that the lidar estimates may have been complicated by the possible presence of a near-stagnant plume from a nearby power plant.

At 2000 UTC (Fig. 16c), most airborne measurements were east (downwind) of Houston. Both airborne instruments continue to show extremely high ML heights at latitudes north of the northern margin of Houston, with some variability among the north–south legs. The in situ aircraft sounding in the northeast corner of the plot actually represents a minimum ML height, because the sounding terminated before the air-
craft reached the top of the ML. That, plus a 2300-UTC sounding from WPP that shows a well-mixed boundary layer to 3000 m, provides independent support of the lidar and MTP ML height estimates to the north.

As the westerly winds carry air eastward across Galveston Bay, the bottom of the PBL should be cooled because a 5°C land–sea temperature contrast is present. Downwind of Galveston Bay, the growth of the ML should be inhibited relative to locations unaffected by Galveston Bay. Both the MTP and lidar show a strong north–south gradient in ML height just east of the northeast corner of Galveston Bay, consistent with expectations. The gradient is less distinct farther downwind. A single east to west pass of the Electra along the north shore of Galveston Bay indicates somewhat higher ML heights than the nearby ELL profiler, but does directly support a sudden change in ML height (from 2000 to 700 m) as a marine PBL forms and develops within the offshore flow.

By this time of day, a sea breeze has developed along the Gulf of Mexico coastline. Both the lidar and MTP indicate ML heights less than 800 m along the Gulf coast east of Galveston Bay. These estimates are confirmed by the LAM radar profiler, which experienced a break in the height of the ML (Fig. 13) when the apparent sea breeze arrived. The airborne lidar observations close to the LAM site were taken prior to 2000 UTC, so the higher ML heights observed by the lidar are not necessarily inconsistent with the LAM profiler observation.

7. Conclusions

In this paper, the first use of an airborne microwave temperature profiler (MTP) to estimate ML heights is described. The accuracy of the ML height estimates is determined by intercomparison among several more conventional estimation techniques, using radiosondes, profilers, airborne lidar, and in situ aircraft measurements.

Overall, there was a relatively good relationship between the ML determined by the separate conventional instruments with respect to the bias, standard deviations, and RMSE of the difference in the heights. The good agreement between the airborne aerosol backscatter lidar and wind profilers suggest that the airborne lidar can be used to provide fairly accurate estimates of the ML heights even in a coastal city such as Houston where there is likely to be strong spatial variability in the aerosol content. The worse agreement between the wind profilers and sonde benchmark suggests that there was error in the radiosonde estimates due to the lack of spatial representativeness, since the radiosondes were subject to turbulent variations in the local height of the PBL that were averaged out in the profiler estimates.

The MTP ML height estimation technique was selected and calibrated by comparison with profiler estimates. The statistical agreement between the MTP and in situ aircraft estimates and the spatial intercomparison with airborne lidar estimates show that the MTP can be used for determining the spatial distribution of ML heights and can give fairly accurate estimates.

The conclusions regarding comparative instrument performance are strictly applicable to the Houston area in summertime but are likely to be generally applicable to tropical or warm-season subtropical coastal environments.

A spatiotemporal analysis of ML heights was performed on a day in which both airborne MTP and airborne lidar were available. Remarkably large spatial variations of ML height were found in and around Houston and confirmed by multiple instruments and measurements. At 1800 UTC, the ML height in the Houston area, influenced by an elevated inversion, was 700–1100 m, while a few tens of kilometers to the north, the ML height was 2000–2500 m. The cause of this extreme variation is not known, but such an ML height variation would have substantial effects on pollutant concentrations. Other large ML height variations were found downwind of Galveston Bay and in the vicinity of the Gulf of Mexico sea breeze. Any potential enhancement or suppression of the ML growth by the urban area was undetectable in comparison with other ML height variations.

Acknowledgments. This project has been funded in part with funds from the State of Texas as part of the program of the Texas Air Research Center. The contents do not necessarily reflect the views and policies of the State of Texas nor does the mention of trade names or commercial products constitute endorsement or recommendation for use. Work performed by MJM at the Jet Propulsion Laboratory, California Institute of Technology, was carried out under a contract with the National Aeronautics and Space Administration. MJM also received support from the National Oceanic and Atmospheric Administration (NOAA) and the Texas Natural Resource Conservation Commission (TNRCC), now the Texas Commission on Environmental Quality (TCEQ).

REFERENCES


