Assimilating Coherent Doppler Lidar Measurements into a Model of the Atmospheric Boundary Layer. Part II: Sensitivity Analyses

ROB K. NEWSOM
Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, Colorado

ROBERT M. BANTA
NOAA/Environmental Technology Laboratory, Boulder, Colorado

(Manuscript received 9 February 2004, in final form 21 June 2004)

ABSTRACT

A series of trials are performed to evaluate the sensitivity of a 4DVAR algorithm for retrieval of microscale wind and temperature fields from single-Doppler lidar data. These trials use actual Doppler lidar measurements to examine the sensitivity of the retrievals to changes in 1) the prescribed eddy diffusivity profile, 2) the first-guess or base-state virtual potential temperature profile, 3) the phase and duration of the assimilation period, and 4) the grid resolution.

The retrieved fields are well correlated among trials over a reasonable range of variation in the eddy diffusivity coefficients. However, the retrievals are quite sensitive to changes in the gradients of the first-guess or base-state virtual potential temperature profile, and to changes in the phase (start time) and duration of the assimilation period. Retrievals using different grid resolutions exhibit similar larger-scale structure, but differ considerably in the smaller scales. Increasing the grid resolution significantly improved the fit to the radial velocity measurements, improved the convergence rate, and produced variances and fluxes that were in better agreement with tower-based sonic anemometers.

Horizontally averaged variance and heat flux profiles derived from the final time steps of all the retrievals are similar to typical large-eddy-simulation (LES) results for the convective boundary layer. However, all retrieved statistics show significant nonstationarity because fluctuations in the initial state tend to be confined within the boundaries of the scan.

1. Introduction

In Part I of this study (Newsom and Banta 2004), we described a four-dimensional variational data assimilation (4DVAR) algorithm for retrieval of microscale wind and temperature fields from single-Doppler lidar data. This retrieval algorithm is essentially a sophisticated analysis procedure, whereby 3D, high-resolution, three-component-velocity and thermodynamic fields are retrieved from sequential 3D volumes of the measured radial-velocity component from a scanning Doppler lidar. The retrieved 3D fields can then be further analyzed to provide vertical profiles of turbulent variances and fluxes within the volume, in a manner similar to the way large-eddy-simulation (LES) numerical model volumes are analyzed.

The algorithm described in Part I is based on a method developed for Doppler radar by Sun et al. (1991) and Sun and Crook (1994) and later adapted for Doppler lidar by Lin et al. (2001) and Chai et al. (2004). The underlying concept is to fit the output of a dynamic model to spatially and temporally resolved measurements of radial velocity from a scanning Doppler lidar. The algorithm described in Part I incorporates a novel data ingest scheme and accounts for radial velocity measurement uncertainty.

Progress in understanding turbulent transport between the surface and the free troposphere has long been impeded by an inability to measure both the spatial variation and temporal evolution of eddy structures in the atmospheric boundary layer (ABL). The 4DVAR retrieval technique represents one potentially powerful tool that can be used to address this problem. The method retrieves the velocity, temperature, and pressure fields as functions of time from the lidar volume scan data. The retrieved fields can be analyzed to produce vertical profiles of the mean and turbulence statistics via horizontal averaging in the same manner as these profiles are obtained from output.

The accuracy of the retrievals will be affected by several factors, including model specifications, first-guess fields, and data quality. Since the method de-
scribed in Part I treats only the initial conditions as control parameters, it is important to understand how the retrieved fields change with respect to changes in the grid resolution, treatment of the boundary conditions, and constants that control physical parameterizations. For this study, we specifically examine the sensitivity of the retrievals to changes in 1) the prescribed eddy diffusivity profile, 2) the first-guess or base-state virtual potential temperature profile, 3) the phase and duration of the assimilation period, and 4) the grid resolution. We also examine the characteristics and behavior of turbulence statistics derived from the retrieved fields.

A common approach in many sensitivity studies is to perform identical twin experiments using simulated measurements. The obvious advantage in this approach is that the true (simulated) fields are known and it is possible to quantify the errors. On the other hand, real datasets offer many challenges that are difficult to simulate, and the results of identical twin experiments may indicate better performance than can be realized in practice. The sensitivity studies presented in this paper make use of actual Doppler lidar data collected under convective conditions during the Cooperative Atmosphere–Surface Exchange Study (CASES-99) field campaign. The sensitivity of the retrieval algorithm to changes in various parameters is assessed by examining the consistency, or lack of thereof, between retrieval results.

This paper is organized as follows. Sections 2 and 3 discuss the retrieval algorithm and observations. These sections provide only brief overviews since these topics were covered in more detail in Part I. The results of the retrieval experiments are presented in section 4, and our conclusions are given in section 5.

2. Retrieval method

The retrieval algorithm fits the output of a model of ABL flow to radial velocity data from a coherent Doppler lidar. The algorithm is similar to that described by Sun et al. (1991) and Sun and Crook (1994), except for several modifications, which include the following.

- The cost function is computed by interpolating model quantities to the measurement points instead of vice versa. Thus, measurements are used only where they are available, and the problem of interpolating through data voids is left to the model itself.

- The adjoint equations are developed assuming that the subresolution-scale (SRS) fluxes can be represented as generic functions of the resolved scale velocity and temperature gradients. Specific SRS schemes can then be input. Here, and in Part I, specific forms are derived using gradient transport theory in which the eddy and temperature diffusivity coefficients are prescribed as functions of height.

- Measurement uncertainty is included in the cost function evaluation, reducing the impact of regions of low signal-to-noise ratio (SNR) data.

Flow in the ABL is modeled using the following system of equations:

\[
\frac{\partial u_i}{\partial t} = -\frac{\partial p}{\partial x_i} - \frac{\partial}{\partial x_j}(u_i u_j) + \frac{g}{\Theta_{ref}} \frac{\partial}{\partial x_j} (\theta - \langle \theta \rangle) + \frac{\partial}{\partial x_j}(2K_w S_{ij}),
\]  

(1)

\[
\frac{\partial \hat{\theta}}{\partial t} = -\frac{\partial}{\partial x_j}(u_i \hat{\theta}) - \frac{\partial \theta_e}{\partial z} + \frac{\partial}{\partial x_j} \left( K_h \frac{\partial \hat{\theta}}{\partial x_j} \right),
\]  

and (2)

\[
\nabla^2 p = -\frac{\partial}{\partial x_j} \left[ \frac{\partial}{\partial x_j}(u_i u_j) - \frac{g}{\Theta_{ref}} \frac{\partial}{\partial x_j} (\theta - \langle \theta \rangle) \right]
\]  

\[-\frac{\partial}{\partial x_j}(2K_w S_{ij}) - \frac{u_i}{\Delta t} \right].
\]  

(3)

This model simulates dry, shallow incompressible flow under the Boussinesq approximation. In Eqs. (1)–(3) \( u \) is the velocity, \( g \) is the acceleration of gravity, \( \Theta_{ref} \) is a reference virtual potential temperature, \( p \) is related to the perturbation pressure, \( S_y = 0.5[\left( \partial u_i / \partial x_j \right) + \left( \partial u_j / \partial x_i \right)] \) is the strain rate, and horizontal averages are denoted by \( \langle \cdot \rangle \). Implicit in the formulation of Eq. (3) is the use of the Harlow–Welch (Harlow and Welch 1965) scheme, where \( \Delta t \) is the time step size.

The potential temperature is \( \theta = \hat{\theta} + \theta_b \) where \( \theta_b \) is the base-state potential temperature, and \( \theta \) is the departure from the base state. The base-state potential temperature depends only on height and serves as a reference profile. The advective term in Eq. (2) has been written in terms of \( \hat{\theta} \) rather than \( \theta \), because this form is more numerically stable. When the advective term in Eq. (2) is expressed in terms of \( \theta \), the adjoint equations contain a term that depends linearly on \( \theta \). This can cause instabilities in the conjugate gradient algorithm because the numerical value of \( \theta \) is so much larger than the other prognostic variables.

The coefficients of eddy diffusivity \( K_w \) and temperature diffusivity \( K_h \) are handled as prescribed functions of height. The vertical variation of \( K_w \) is modeled using an expression similar to the one proposed by Troen and Mahrt (1986). This is given by

\[
K_w = K_{\text{max}}(1 + \alpha) \left( \frac{1 + \alpha}{\alpha} \right)^\alpha \frac{z}{z_{\text{max}}} \left( 1 - \frac{z}{z_{\text{max}}} \right) ^\alpha
\]

for \( z_{\text{max}} \equiv z \),

(4)

where \( z_{\text{max}} \) is the height of the computational domain, \( K_{\text{max}} \) determines the maximum value of \( K_w \), and \( \alpha \) controls the shape of the profile and the height of the maximum. We further assume that

\[
K_h = 2K_w.
\]

(5)
Retrievals are performed by minimizing a cost function, which is a measure of the discrepancy between the model output and the radial velocity observations. The primary term in the cost function is given by (Newsom and Banta 2004)

$$J_{\text{obs}} = \sum_m \left( \frac{\Delta_m}{\sigma_m} \right)^2,$$

where

$$\Delta_m = \mathbf{u}_m \cdot \mathbf{r}_m - u_{m,\text{obs}}$$

is the difference between the modeled and observed radial velocity. The velocity field generated by the model is denoted by $\mathbf{u}_m$, where the overbar implies interpolation to the coordinates of the $m$th radial velocity observation, $u_{m,\text{obs}}$. The unit vector from the lidar to the $m$th observation is $\mathbf{r}_m$, and $\sigma_m$ is the error in this measurement.

The observed radial velocity is a function of the position vector $\mathbf{r}$ from the lidar to a point in the scan volume. This position vector is a function of time, $\mathbf{r} = \mathbf{r}(t)$, since the beam is scanned. In practice, it may take several minutes to scan a full volume, and during that time the eddy structure within the volume changes. The cost function used in this study is evaluated by interpolating the model output to the space–time coordinates of the observations. This approach is a dynamically consistent means of ingesting the data, since the observations are used only when and where they are available.

Each term in the cost function is also weighted by the reciprocal of the square of the measurement error $\sigma$. Radial velocity error was estimated from fixed-beam (staring) data that were acquired periodically during the CASES-99 field deployment and used to assess instrument performance. For moderate to strong return signals the radial velocity error is dominated by random noise. The variance of the radial velocity noise versus return signal strength was estimated from time series analysis of fixed beam data. The relationship between $\sigma$ and a measure of the return signal strength is shown in Fig. 8 of Part I. The same curve was used to provide estimates of $\sigma$ for the volume scan data used in this study.

As in Part I the model and its adjoint are integrated using a space-centered, forward-in-time scheme on a staggered grid with periodic lateral boundary conditions. At the top and bottom of the domain the vertical velocity vanishes, and the potential temperature is set to its base-state value. Horizontal velocities are set to zero at the bottom of the domain. At the top of the domain the horizontal velocities are set to the base-state values derived from a VAD-type (Banta et al. 2002; Chai et al. 2004) analysis of the lidar data. For purposes of the sensitivity tests the model was run at relatively lower resolution to perform the maximum number of tests. This was deemed acceptable because we were attempting to assess similarities and differences among runs rather than the absolute accuracy of the runs. An exception was the comparisons between the low-resolution and high resolution runs, where we were interested in determining which run produced the best agreement with measured fluxes. These were compared with measurements from the nearby 60-m CASES-99 tower.

The cost function is minimized by variational adjustment of the initial prognostic variables using a conjugate gradient algorithm (Press et al. 1988), in which gradient information is derived from the model’s adjoint. This iterative process is terminated when a predefined convergence tolerance is achieved. For this study, the algorithm terminates when the number of iterations reaches 200, or when the relative change in the cost function from one iteration to the next is less than $10^{-8}$.

### 3. Observations

The retrieval algorithm assimilates Doppler lidar measurements of radial velocity that are obtained by repeatedly scanning a volume of the ABL. For this study, volume scans were acquired using The National Oceanic and Atmospheric Administration’s (NOAA’s) high-resolution Doppler lidar (HRDL) during the CASES-99 field program in October 1999 (Poulos et al. 2002). Basic performance parameters for HRDL are shown in Table 1 of Part I, and the design is discussed in detail by Grund et al. (2001). Wulfmeyer et al. (2000) describe further refinements. Data from HRDL acquired during CASES-99 have been used in numerous studies (Blumen et al. 2001; Newsom and Banta 2003; Banta et al. 2002, 2003; Sun et al. 2002, 2003; Fritts et al. 2003; Drobinski et al. 2004). The CASES-99 field site was located in an area characterized by open and gently rolling terrain. A map of the field site indicating the lidar location, scan volume, and model domain is shown in Fig. 3 of Part I. That map also shows the location of a heavily instrumented 60-m tower approximately 1.4 km north of the lidar position.

A total of eight volume scans were recorded by the lidar between 2055:20 and 2107:21 UTC during the afternoon of 25 October 1999. The start and end times of each of these volume scans are given in Table 1. Volume data were acquired and processed in the manner described in Part I. The volume scans were updated every 1.5 min.
For this study we use four volume scans recorded during a 6-min period from 2058:20 to 2104:21 UTC.

Figure 1 displays quality-controlled radial velocity data and the corresponding measurement errors from volume scans 3, 4, 5, and 6. The panels in this figure show individual sector scans taken at an elevation angle of 4°. The eastward movement of inhomogeneities in the radial velocity field is clearly indicated. Since the mean winds were light, animations of the lidar data show good continuity in the structure from scan to scan. This is necessary in order to obtain meaningful results from the retrieval algorithm.

The retrieval algorithm requires specification of the base-state wind and potential temperature profiles. The base state wind profile was computed from the lidar volume scan data using a VAD-type processing technique (Banta et al. 2002; Chai et al. 2003). The base-state virtual potential temperature field was obtained from a radiosonde that was released from the main field site at approximately 1900 UTC, or about 2 h prior to the acquisition of the volume data. Figure 2 shows the base-state wind and virtual potential temperature profiles. During the time of the volume scans the sky was clear and the boundary layer was moderately unstable. The profiles in Fig. 2 indicate a shallow convective boundary layer with light westerly winds (~2 m s⁻¹). The base of the capping inversion was located at approximately 600 m AGL. Above this level the winds veered with height and the wind speeds increased. At 800 m AGL the winds were northwesterly at about 5 m s⁻¹.

4. Assimilation experiments and results

All assimilation trials were performed using a domain size of 3 km × 3 km in the horizontal and 800 m in
the vertical. The coordinate axes are oriented in the standard manner such that the x and y axes point toward the east and north, respectively. The domain boundaries extend from −3 to 0 km in x, −1.5 to 1.5 km in y, and 0 to 800 m in z. The lidar is located at the origin of the domain.

Parameters used for each of the assimilation trials are listed in Table 2. The eddy diffusivity parameters were varied in trials 37–42, base-state virtual potential temperature profiles were varied in trials 44 and 45, and the phase and duration of the assimilation period were varied in trials 46–48. Twelve trials used a relatively coarse grid with $24 \times 24 \times 20$ points in x, y, and z, respectively. Trial 49 used a higher resolution grid with $40 \times 40 \times 34$ points. The low-resolution trials used a time step size of $\Delta t = 2$ s, and the higher resolution trial used a time step size of $\Delta t = 1$ s. The last column in Table 2 contains an index that refers to one of three base-state virtual potential temperature profiles. These three profiles are shown in Fig. 2b. Base-state 1 is a measured profile obtained from a 1900 UTC radiosonde release at the main field site. Base-states 2 and 3 are artificial profiles that are used to test the sensitivity of the retrievals to different first-guess temperature fields.

In this study we evaluate the sensitivity of the retrievals to changes in the eddy diffusivity profile [Eq. (4)] by considering a restricted range of $\alpha$ and $K_{\text{max}}$ values. Profiles of $K_{\alpha}/K_{\text{max}}$ for $\alpha = 2$ and $\alpha = 4$ are shown in Fig. 3.

Prior to performing the retrieval trials, we conducted several runs of a test version of the predictive model component of the assimilation scheme to establish a reasonable range of $K_{\text{max}}$ values for a given grid resolution. These test runs were initialized using the wind and temperature profiles shown in Fig. 2 (profile 1 in Fig. 2b) with small random perturbations added to the temperature profile to initiate turbulent motions. The test model was then integrated in time and the dynamics allowed to spin up. Many different test runs were performed with different values of $K_{\text{max}}$, and we observed which values resulted in approximately statistically steady state turbulence after the initial spinup period. For a grid resolution of $24 \times 24 \times 20$, variance and flux profiles did not change significantly over 30 min of simulation time after spin up for $K_{\text{max}}$ values between 5 and 20 m$^2$ s$^{-1}$. For the higher resolution trial ($40 \times 40 \times 34$), we used $K_{\text{max}} = 7$ m$^2$ s$^{-1}$.

Table 3 provides a synopsis of the trial results. The second column in Table 3 represents the percentage of model generated radial velocities, $\mathbf{u} \cdot \mathbf{\hat{f}}_m$, that occur within $\sigma_{\text{rms}} \pm \sigma_{\text{m}}$. This provides a measure of how well the model fits the available radial velocity measurements. The third column in Table 3 is the number of iterations that were performed for each trial. The retrieval algorithm was terminated after 200 iterations or when the convergence tolerance was achieved, whichever came first. Only in three cases (trials 40, 48, and 49) was the convergence tolerance achieved before 200 iterations were completed.
TABLE 3. Trial number, percentage of modeled radial velocities within measurement precision, number of iterations performed, and total number of measurements used in each trial.

<table>
<thead>
<tr>
<th>Trial</th>
<th>Percent ≤ σ</th>
<th>Iterations</th>
<th>Measurements</th>
</tr>
</thead>
<tbody>
<tr>
<td>37</td>
<td>80.49</td>
<td>200</td>
<td>11 274</td>
</tr>
<tr>
<td>38</td>
<td>79.84</td>
<td>200</td>
<td>11 274</td>
</tr>
<tr>
<td>39</td>
<td>76.37</td>
<td>200</td>
<td>11 274</td>
</tr>
<tr>
<td>40</td>
<td>77.65</td>
<td>149</td>
<td>11 274</td>
</tr>
<tr>
<td>41</td>
<td>77.77</td>
<td>200</td>
<td>11 274</td>
</tr>
<tr>
<td>42</td>
<td>72.30</td>
<td>200</td>
<td>11 274</td>
</tr>
<tr>
<td>43</td>
<td>79.93</td>
<td>200</td>
<td>11 274</td>
</tr>
<tr>
<td>44</td>
<td>79.86</td>
<td>200</td>
<td>11 274</td>
</tr>
<tr>
<td>45</td>
<td>69.32</td>
<td>200</td>
<td>17 002</td>
</tr>
<tr>
<td>46</td>
<td>69.29</td>
<td>200</td>
<td>16 966</td>
</tr>
<tr>
<td>47</td>
<td>62.07</td>
<td>196</td>
<td>22 694</td>
</tr>
<tr>
<td>48</td>
<td>87.97</td>
<td>64</td>
<td>11 274</td>
</tr>
</tbody>
</table>

The sensitivity of the retrievals to changes in parameters is quantified using the linear correlation coefficient between the same prognostic variables from different trials. Here we define the height dependent linear correlation coefficient between a prognostic variable $\xi$ from trials $p$ and $q$ as

$$ R_p(\xi_{pq}, \xi_q) = \frac{\sum_{x,y,t} (\xi_{pq} - \langle \xi_{pq} \rangle)(\xi_q - \langle \xi_q \rangle)}{\sqrt{\sum_{x,y,t} (\xi_{pq} - \langle \xi_{pq} \rangle)^2 \sum_{x,y,t} (\xi_q - \langle \xi_q \rangle)^2}}, $$

(8)

where $\Sigma_{x,y,t}$ denotes summation over time and over the horizontal grid points within the scan volume $\Omega$. The symbol $\xi$ is used to represent any prognostic variable, $u, v, w,$ or $\theta$, and the subscripts $p$ and $q$ refer to trial numbers. The horizontal mean of $\xi$ is

$$ \langle \xi \rangle_{ht} = \frac{1}{N_x N_y} \sum_{x,y} \xi, $$

(9)

where $N_x$ is the number of grid points in $x$ at each vertical level inside $\Omega$. Similarly, $N_y$ is the number of grid points in $y$ at each vertical level inside $\Omega$. We also define the volume–time–averaged correlation as

$$ R_p(\xi_{pq}, \xi_q) = \frac{\sum_{x \in \Omega} (\xi_{pq} - \langle \xi_{pq} \rangle)(\xi_q - \langle \xi_q \rangle)}{\sqrt{\sum_{x \in \Omega} (\xi_{pq} - \langle \xi_{pq} \rangle)^2 \sum_{x \in \Omega} (\xi_q - \langle \xi_q \rangle)^2}}, $$

(10)

where $\Sigma_{x \in \Omega}$ denotes summation over time and over all grid points within $\Omega$. The correlation $R_p(\xi_{pq}, \xi_q)$ depends only on height, and $R_p(\xi_{pq}, \xi_q)$ is a scalar that represents the overall correlation between $\xi_{pq}$ and $\xi_q$.

**a. Sensitivity to the eddy diffusivity profile**

Trials 37–42 were performed to investigate how the retrievals are affected by changes in the prescribed eddy diffusivity profile as given by Eq. (4). Trials 37, 38, and 39 were conducted with $\alpha = 4$ and $K_{\text{max}}$ equal to $5, 10,$ and $20 \text{ m}^2 \text{s}^{-1}$, respectively. Trials 40, 41, and 42 were conducted over the same range of $K_{\text{max}}$ but with $\alpha = 2$.

The effects of varying either $K_{\text{max}}$ or $\alpha$ separately are illustrated in Fig. 4. This figure shows vertical profiles of $R(\xi_{pq}, \xi_q)$ for all prognostic variables in which trial 38 ($p = 38$) is used as a reference. Figure 4 also shows profiles of the squared difference in the eddy diffusivity, $\Delta K^2_p$, where $\Delta K^2_p = K_{\text{ref}}^p - K_{\text{m}}^p$, and $K_{\text{ref}}^p$ is the eddy diffusivity profile for trial 38, which is defined by $K_{\text{max}} = 10 \text{ m}^2 \text{s}^{-1}$ and $\alpha = 4$. The profiles in Figs. 4a and 4c illustrate the effect of changing $K_{\text{max}}$ with $\alpha = 4$, and the profiles in Fig. 4e show the effects of changing $\alpha$ from 2 to 4 for $K_{\text{max}} = 10 \text{ m}^2 \text{s}^{-1}$. In Figs. 4a and 4b the correlations generally decrease toward the surface. Figure 4c shows a distinct inverse relation between $\Delta K^2_p$ and the correlation, that is, in layers where $\Delta K^2_p$ is large, the correlations are smaller, and vice versa. Overall, however, correlations between trials remain above ~0.7 at all levels.

Table 4 lists the volume–time–averaged correlations between trial 38 and trials 37, 39, 40, 41, and 42. This table also shows rms($\Delta K_m$) = [$\sum_p (K_m^p - K_{\text{ref}}^p)^2/N]$ for all of these five comparisons. The volume–time–averaged correlations, which range between 0.73 and 0.93, indicate that the fields remain relatively well correlated over a reasonable range of variation in the eddy diffusivity coefficients. The correlations increase with decreasing rms($\Delta K_m$). The $u$ component exhibits the best correlation between the various retrievals. This is expected because the radial velocity measurements are most sensitive to the $u$ component due to the orientation of the volume scan in this case. In other words, the $u$ component is the most directly observed variable. The $v$ component and the potential temperature exhibit the poorest correlations between the various trials. This is likely due to the fact that the $v$ component is mostly orthogonal to the observed radial velocity, and no direct observations of potential temperature fluctuations are used in the assimilation.

**b. Sensitivity to the base-state $\theta$ profile**

The retrieval algorithm was initiated by setting the first-guess fields to base-state profiles of wind and potential temperature. The base-state profiles are assumed to be representative of the horizontal mean. The expectation is that the retrieved perturbation fields will be close to the base state. For this study, the base-state wind profiles are computed directly from the lidar volume scan data using a VAD-type processing technique (Banta et al. 2002; Chai et al. 2004). As a result, we are confident that the base-state wind profiles shown in Fig. 2a are representative of the true horizontally averaged mean flow during the assimilation period. On the other hand, the base-state potential temperature profile was obtained from a radiosonde released at the main
Fig. 4. Profiles of linear correlations $R(z_p, z_q)$ between trial $p = 38$ and trials (a) $q = 37$, (c) $q = 39$, and (e) $q = 41$. Profiles of the squared difference in the eddy diffusivities between trial $p = 38$ and trials (b) $q = 37$, (d) $q = 39$, and (f) $q = 41$. The legend appearing in (a) applies to (c) and (e).

Table 4. Results of the eddy diffusivity sensitivity study.

<table>
<thead>
<tr>
<th>Trials</th>
<th>Rms ($\Delta K_m$) ($\text{m}^2\text{s}^{-1}$)</th>
<th>Correlations, $R(z_p, z_q)$</th>
<th>$\xi = u$</th>
<th>$\xi = v$</th>
<th>$\xi = w$</th>
<th>$\xi = \theta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>38, 37</td>
<td>2.769</td>
<td></td>
<td>0.927</td>
<td>0.867</td>
<td>0.899</td>
<td>0.881</td>
</tr>
<tr>
<td>38, 39</td>
<td>5.538</td>
<td></td>
<td>0.879</td>
<td>0.819</td>
<td>0.844</td>
<td>0.825</td>
</tr>
<tr>
<td>38, 40</td>
<td>2.917</td>
<td></td>
<td>0.882</td>
<td>0.800</td>
<td>0.859</td>
<td>0.816</td>
</tr>
<tr>
<td>38, 41</td>
<td>2.840</td>
<td></td>
<td>0.931</td>
<td>0.888</td>
<td>0.928</td>
<td>0.896</td>
</tr>
<tr>
<td>38, 42</td>
<td>8.505</td>
<td></td>
<td>0.833</td>
<td>0.744</td>
<td>0.793</td>
<td>0.732</td>
</tr>
</tbody>
</table>

Field site about 2 h prior to the acquisition of the lidar volume scan data. Due to this delay, it is possible that the data from this sounding may not be representative of the conditions that prevailed when the lidar volume scan data was recorded; for example, the boundary-layer depth may have increased during that time. We are also concerned that the radiosonde measurements may not be representative of the true horizontal mean because the atmosphere is sampled only along the ascent path. In this section we examine the extent to which the retrievals are affected by changes in the prescribed base-state potential temperature profile.

To test the sensitivity of the retrievals to the prescribed potential temperature base state we ran the re-
trieval algorithm using two artificial profiles. These profiles are labeled 2 and 3 in Fig. 2b. The wind profiles in Fig. 2a indicate veering with height starting at about 650 m AGL. However, the base of the capping inversion layer in profile 1 (Fig. 2b) is approximately 100 m below this level. This suggests that the boundary-layer depth increased during the intervening period between the assimilation time and the sounding time. In order to test the impact of the uncertainty in the boundary-layer depth on the retrieved fields, we constructed profile 2 with an inversion layer starting at about 650 m AGL. Additionally, the retrieval algorithm was run using profile 3. This profile is constant through the depth of the domain and was obtained by vertically averaging profile 1. Clearly, profile 3 is not representative of the stability structure of typical convective boundary layers (CBLs).

Two trials were conducted that used the artificial \( \theta_b \) profiles shown in Fig. 2b. Trials 44 and 45 use the base-state profiles labeled 2 and 3, respectively. Table 5 lists the volume–time-averaged linear correlation coefficients [Eq. (10)] between trials 38 and 44 and between trials 38 and 45. Once again, \( u \) exhibits the best correlation and \( \theta \) exhibits the poorest correlation. Table 5 indicates reasonably good overall correlations with values ranging from 0.75 to 0.93. However, these volume–time-averaged correlations convey no information about the vertical structure.

Figure 5 shows profiles of the height-dependent correlation [Eq. (8)] between trials 38 and 44 (Fig. 5a) and trials 38 and 45 (Fig. 5c). All fields remain relatively well correlated below 400 m, although the absence of a superadiabatic surface layer in profile 3 slightly degrades the correlation between trials 38 and 45 in this region. Above about 500 m, the correlations become much worse. Figure 5c shows that the \( \theta \) fields from trials 38 and 45 become very poorly correlated near 600 m AGL.

The purpose of using profile 3 was to determine if the retrieval algorithm is capable of adjusting the first-guess temperature field sufficiently to obtain a mean vertical structure that is consistent with the CBL.

![Figure 5](image_url)

**Figure 5.** Profiles of (a) linear correlation and (b) the absolute deviation in the stability between trials 38 and 44. Also shown are (c) the profiles of the correlation and (d) the absolute deviation in the stability between trials 38 and 45. The legend appearing in (a) also applies to (c).
Figure 5 also displays profiles of the absolute deviation in the base-state stability, \(|\partial \Delta \theta / \partial z|\), where \(\Delta \theta_b = \theta_b - \theta^e_0\), and \(\theta^e_0\) refers to profile 1 in Fig. 2b. This shows the effects of the difference in the \(\theta\) profile on the vertical \(\theta\) gradient at each level. A clear relationship between \(\partial \Delta \theta / \partial z\) and the corresponding correlation profile is evident. In layers where the base-state stability deviation is large, the correlation between the variables is poor, indicating that the retrieved fields are quite sensitive to the first-guess (base state) stability. We also note that for all trials the retrieved mean potential temperature profile deviates little from the prescribed base state. This indicates that there is simply insufficient information in the radial velocity field alone for the retrieval method to determine the correct mean stability profile. Thus, without an accurate representation of the base-state stability, the retrieval algorithm can be expected to produce dubious results, particularly in the temperature and vertical velocity fields.

**c. Sensitivity to the phase and duration of the assimilation period**

To minimize memory requirements and CPU time it is advantageous to assimilate as few volume scans as necessary. It would seem reasonable that a minimum of two consecutive volume scans would be required in order to capture the evolution of the radial velocity field over the entire scan volume. However, it is not known to what extent the retrievals change as the duration of the assimilation period is increased. Staggering the start time or the phase of the assimilation period may also significantly affect the results.

To address these questions we performed several runs using different assimilation periods and start times, and then examined the correlation between these trials during periods of temporal overlap. Figure 6 is a time line showing the phase and duration of the assimilation periods for four different retrieval runs. The assimilation periods for trials 46 and 47 are equal in length, but their start times are staggered by one volume scan. Trials 38, 46, and 47 overlap during volumes 4 and 5. Trials 46 and 48 overlap during volumes 3, 4, and 5, whereas trials 47 and 48 overlap during volumes 4, 5, and 6.

As the assimilation period increases and more observations are included, the overall fit of the model solution to the radial velocity data degrades. Distributions of \(\Delta_m / \sigma_m\) are shown in Fig. 7 for trials 38, 46, and 48. These distributions become broader as the assimilation period is increased because the number of degrees of freedom or grid resolution remains fixed.

Table 6 lists the volume–time-averaged correlations [Eq. (10)] between the various trials during periods of temporal overlap. For trials with common start times, we observe that increasing the assimilation period from two to three volume scans results in a significant decorrelation between the retrievals. The effect is less pronounced in going from three to four volume scans because this change represents a smaller percentage of the assimilation period. Retrievals with the same start times exhibit better correlations than retrievals with start times that are staggered by one volume scan. For example,
trial 38 is better correlated with trial 47 than with trial 46. Similarly, trial 48 is better correlated with trial 46 than with trial 47. Overall, the correlations in Table 6 range from 0.3 to 0.8, indicating that the retrievals can be fairly sensitive to changes in the phase and duration of the assimilation period.

d. Sensitivity to model resolution

Trial 49 was performed to assess the effects of model grid resolution on the retrieved fields. As indicated in Table 3, trial 49 assimilates data from volume scans 4 and 5, and uses a \( 40 \times 40 \times 34 \) grid. This equates to a grid resolution of \( \Delta x = \Delta y = 75 \text{ m} \) and \( \Delta z = 23.5 \text{ m} \), or nearly twice the resolution of the other trials considered in this study. For this higher-resolution trial, \( K_{\text{max}} \) was set to 7 m/s and the time step \( \Delta t \) was set to 1 s.

Trial 49 resulted in the best fit to the radial velocity measurements and required the fewest iterations to achieve the convergence tolerance. Of course, this improvement in convergence came at the expense of significantly increased CPU time per iteration. The distribution of \( \Delta x / \sigma_m \) for trial 49 is shown in Fig. 10b in Part I. This distribution is quite narrow, with 87% of the radial velocity measurements occurring within \( \pm \sigma_m \) of the retrieved radial velocity field (also see Table 3). This result should not be that surprising, since increasing the resolution also increases the number of degrees of freedom available to the retrieval algorithm.

Table 7 lists volume–time-averaged correlations [Eq. (10)] between \( u, v, w, \) and \( \theta \) for trial 49 and three lower-resolution trials that use the same assimilation period and base-state temperature profile. Because of the differences in spatial and temporal resolution, correlation coefficients had to be computed by linearly interpolating the lower resolution trials to the space–time coordinates of the higher-resolution grid. Once again, Table 7 shows that \( u \) exhibits the best correlations with values between 0.7 and 0.75. Overall, the correlations in Table 7 range from about 0.48 to 0.75, indicating a fair degree of sensitivity to grid resolution.

To illustrate the effect of changing the grid resolution, Fig. 8 shows a comparison between trials 39 and 49. This figure displays horizontal cross sections of the perturbation horizontal velocity, vertical velocity, and perturbation potential temperature fields at 200 m AGL.

As expected, trial 49 exhibits more high-frequency structure. Nevertheless, both the high- and low-resolution trials exhibit many of the same gross features. Both trials agree on general regions of positive and negative vertical velocities. The same is true for regions of positive and negative perturbation potential temperatures.

e. Characteristics of the retrieved variance and flux profiles

This section examines the characteristics of velocity and temperature variances and fluxes of heat and momentum for a given trial. Time-dependent profiles of horizontally averaged, resolved-scale variances and covariances (fluxes) between two prognostic variables, \( \xi \) and \( \eta \), were computed using

\[
\langle \xi' \eta' \rangle_0 = \langle \xi' \rangle_0 \langle \eta' \rangle_0 - \langle \xi \eta' \rangle_0,
\]

where \( \langle \cdot \rangle_0 \) denotes a horizontal average over grid points within the scan volume \( \Omega \). It is necessary to limit the averaging to grid points within \( \Omega \) because variances and magnitudes of covariances are biased toward smaller values when grid points outside \( \Omega \) are included in the averages.

Figure 9 shows vertical profiles of TKE, potential temperature variance, and the total vertical kinematic heat flux for trials 37, 38, 47, 48, and 49. The velocity variances used to compute TKE and the temperature variance contain only contributions from resolved scale fluctuations. On the other hand, the total vertical kinematic heat flux shown in Fig. 9 is equal to the resolved plus the SRS heat flux, where the SRS contribution is given by

\[
-K_j (\partial \theta / \partial z)_0.
\]

Figure 9 illustrates the effects of changes in the eddy diffusivity, assimilation period, and model resolution on the profiles of TKE, potential temperature variance, and heat flux. Despite subtle differences, the general shape characteristics of each of these turbulence profiles are similar from trial to trial. These profiles are consistent with results for the CBL based on measurements (Kaimal et al. 1976; Lenschow et al. 1980) and LES output (Deardorff 1972; Moeng 1984; Moeng and Sullivan 1994). The most significant difference is in the magnitudes of the profiles associated with the high-resolution run, trial 49. This trial produced significantly smaller values for the variances and fluxes. We emphasize that all profiles shown in Fig. 9 were computed from

| Table 6. Volume–time-averaged correlations due to changes in the phase and duration of the assimilation period. |
|---|---|---|---|---|---|
| Trials | Overlap volumes | Correlations, \( R(\xi_1, \xi_2) \) |
| \( p, q \) | \( 4 \rightarrow 5 \) | \( \xi = u \) | \( \xi = v \) | \( \xi = w \) | \( \xi = \theta \) |
| 38, 46 | 0.655 | 0.462 | 0.549 | 0.447 |
| 38, 47 | 0.698 | 0.498 | 0.504 | 0.493 |
| 38, 48 | 0.562 | 0.341 | 0.441 | 0.326 |
| 46, 47 | 0.657 | 0.439 | 0.466 | 0.359 |
| 46, 48 | 0.802 | 0.671 | 0.644 | 0.641 |
| 47, 48 | 0.764 | 0.642 | 0.690 | 0.592 |

| Table 7. Volume–time-averaged correlations due to changes in the grid resolution. |
|---|---|---|---|---|---|---|
| Trials | Correlations, \( R(\xi_1, \xi_2) \) |
| \( p, q \) | \( \xi = u \) | \( \xi = v \) | \( \xi = w \) | \( \xi = \theta \) |
| 37, 49 | 0.708 | 0.476 | 0.559 | 0.509 |
| 38, 49 | 0.707 | 0.480 | 0.573 | 0.520 |
| 39, 49 | 0.747 | 0.519 | 0.608 | 0.567 |

Unauthenticated | Downloaded 11/20/21 08:35 AM UTC
the last time step of each assimilation period. This is a rather important point since turbulence statistics derived from all trials exhibited significant inhomogeneity and nonstationarity. The nonstationarity shows up as an adjustment in time from the first to the last retrieved volume, as shown in Fig. 10. Because the CBL was approximately statistically stationary during the acquisition of the volume scan data, it is reasonable to expect the retrieved statistics to be stationary during that period. This proves not to be the case, however.

Figure 10 shows the time series of the volume-averaged variances of $u$, $v$, $w$ and $\theta$ for trials 38, 47, 48,
Fig. 9. Profiles of $\langle u^2 + v^2 + w^2 \rangle$ (solid), potential temperature variance $\langle (\theta - \bar{\theta})^2 \rangle$ (dotted), and total kinematic heat flux $\langle w' \theta' \rangle - K \langle \theta \partial \theta / \partial z \rangle$ (dashed) at the end of the assimilation period for trials (a) 37, (b) 38, (c) 39, (d) 47, (e) 48, and (f) 49.

We have defined the time-dependent volume-averaged (resolved scale) variance of a prognostic variable $\xi$ as

$$\langle (\xi - \bar{\xi})^2 \rangle = \frac{1}{N} \sum_{x=1}^{N} (\xi - \langle \xi \rangle)^2,$$  \hspace{1cm} (12)

where $N$ is the number of grid points in $\Omega$. We observe that the variances in general tend to decrease as the assimilation period is increased or, most notably, as the grid resolution is increased. Another feature that is consistent among all of the runs is the nonsteadiness of the variance in all variables. The variance of $w$ and $\theta$, in particular, experience significant transient behavior during the assimilation period. The variance of the potential temperature field is large initially and then decays rapidly. The vertical velocity field exhibits the opposite trend. This behavior is more pronounced for the low-resolution, short-period assimilation runs.

The nonsteadiness of the vertical velocity and potential temperature statistics are clearly illustrated in Fig.
11. This figure shows horizontal cross sections of the perturbation potential temperature and vertical velocity fields at \( z = 200 \) m AGL during the beginning, middle, and end of the assimilation period for trial 48. Another striking feature of Fig. 11, that is also evident in Fig. 8, is the localization of the fluctuations within the boundaries of the scan. This occurs because the adjoint equations are forced primarily in the data-rich regions. As a result, the initial model state remains relatively unchanged from the first-guess field in regions containing no data. This imposes a nonphysical confinement of fluctuations in the initial state. Thus all variables, particularly \( w \) and \( \theta \), are initially out of adjustment, but they approach a more statistical steady state later in the assimilation period.

Figure 12 displays the evolution of vertical profiles of the potential temperature variance, vertical velocity variance, and total kinematic heat flux (resolved plus subresolution scale) during the beginning, middle, and end of the assimilation period for the high-resolution trial (trial 49). These profiles were computed using only those grid points that reside inside \( \Omega \). The initial profiles bear little resemblance to typical LES results for the CBL (Deardorff 1972; Moeng 1984; Moeng and Sullivan 1994).

Finally, we compare variance and flux estimates from the retrievals near the surface to those measured from sonic anemometers on the 60-m tower at the main CAS-ES-99 field site. Table 8 lists the mean “tower-layer” velocity variances, friction velocity, and kinematic heat flux computed from sonic anemometer data. Eight sonic anemometers were deployed on the main 60-m tower at heights of 1.5, 5, 10, 20, 30, 40, 50, and 55 m. Temperature data from the sonics were converted to virtual potential temperature using data from pressure sensors located at heights of 1.5, 30, and 50 m, and relative humidity data from sensors located at heights of 5, 15, 25, 35, and 45 m. The pressure and relative humidity data were interpolated to the heights of the sonic anemometers at 5, 10, 20, 30, and 40 m. Variances and fluxes were then computed at each of these six levels by averaging over a 30-min period from 2045 to 2115 UTC. The mean tower-layer estimates shown in Table 8 were obtained by averaging these fluxes and variances over the six levels.

Table 8 also displays velocity variances, friction velocity, and the total kinematic heat flux (resolved plus SRS) averaged over the first two grid levels above the surface from the higher-resolution trial (trial 49).
Fig. 11. Evolution of horizontal cross sections of (top) vertical velocity and (bottom) potential temperature for trial 48. The cross sections are taken at $z = 200$ m. Shaded areas indicate negative values.

Fig. 12. Evolution of profiles of vertical velocity variance $\langle w'^2 \rangle$, potential temperature variance $\langle \theta'^2 \rangle$, and total heat flux $\langle w' \theta' \rangle - K_\ell \langle \theta' \theta' \rangle$ for trial 49. Panels (a), (b), and (c) show the profiles at the beginning, middle, and end of the assimilation period, respectively.
friction velocity was computed using total momentum fluxes (resolved plus SRS). Trial 49 produced significantly smaller variances and fluxes than the lower resolution trials. For example, for trial 49 the total kinematic heat flux near the surface is \( \sim 0.2 \) K m s\(^{-1}\). This compares to 0.09 K m s\(^{-1}\) from the tower. By contrast, referring to Fig. 9 we see that the total kinematic heat flux from trial 38 is \( \sim 0.4 \) K m s\(^{-1}\) near the surface, or roughly 4 times larger than the tower estimate. Although the higher resolution trial produced fluxes and variances that are still too large, these results suggest that finer resolution in the model improves the agreement with the tower measurements.

### 5. Summary and conclusions

This study presented the results of several trials designed to investigate the sensitivity of the retrieval algorithm described in Part I. These trials use actual Doppler lidar data collected under convective conditions during the CASES-99 field program. The variability in the retrievals due to changes in various parameters was assessed by examining the correlation between trials. The trials were conducted to specifically examine the sensitivity of the retrievals to changes in 1) the prescribed eddy diffusivity profile, 2) the first-guess or base-state virtual potential temperature profile, 3) the phase and duration of the assimilation period, and 4) the grid resolution.

In general, the \( u \) component is the least sensitive to changes in any of these quantities. This is not surprising because the \( u \) component is the most directly observed variable due to the orientation of the scan volume. The potential temperature tends to exhibit the poorest correlations between the various trials due to the lack of direct measurements of \( \theta \).

Changes in the prescribed eddy diffusivity profile had a relatively minor effect on the retrievals. However, the retrievals were found to be quite sensitive to the gradient of the base-state virtual potential temperature profile. Furthermore, retrieved mean potential temperature profiles deviate little from first-guess temperature profiles, indicating that the radial velocity observations alone are not adequate for the algorithm to determine the correct mean temperature profile. Thus, it is important to use an accurate representation of the base-state temperature profile.

Retrievals were also found to be quite sensitive to the phase (start time) of the assimilation period. For this study the retrieval algorithm was set up such that each trial was essentially started from scratch, using only the base-state wind and temperature profiles as first-guess fields. The effect of phase could be mitigated by incorporating a background term into the cost function that acts to keep the initial model state close to fields derived from a previous assimilation period or forecast. This is a common approach for many forecast models using 4DVAR (Chao and Chang 1992).

As the assimilation period increases and more observations are added, the overall fit of the model to the radial velocity data degrades. The results presented in this study showed that when the duration of the assimilation period is increased by a factor of 2 the correlation during the period of temporal overlap drops to \( \sim 0.4 \). In that case only the gross flow structures exhibit similarity. However, the correlation improves as the grid resolution is increased for short duration trials.

A similar trend was observed by increasing the grid resolution from 24 \( \times \) 24 \( \times \) 20 to 40 \( \times \) 40 \( \times \) 34. Obviously, the higher-resolution retrieval produces more fine detail in the flow structure. The correlations between high- and low-resolution retrievals yielded values near 0.6 overall. A comparison of the horizontal flow structure between the high- and low-resolution retrievals show that they differ considerably in the fine structure but exhibit the same gross structure. Again, this is not surprising, because the dimensions of the largest-scale eddy structures in the volume are dictated by the data (see Fig. 1).

Statistics derived from all retrievals display significant nonstationarity. The variance of \( w \) and \( \theta \), in particular, experience significant transients and are approximately anticorrelated during the assimilation period. The variance of the temperature field is large initially and then decays, whereas the vertical velocity variance exhibits the opposite trend. These transients occur as the flow undergoes adjustment from an initial state in which fluctuations are localized within the boundaries of the scan. This occurs because the adjoint equations are forced primarily in the data-rich regions. As a result, the initial model state remains relatively unchanged from the first-guess fields in regions containing no data. This imposes a nonphysical confinement of fluctuations in the initial state, which in turn results in nonstationary turbulence statistics as the dynamics evolve.

Increasing the grid resolution significantly improved the fit to the radial velocity measurements, improved the convergence rate, and produced smaller variances and fluxes that were in better agreement with tower-based sonic anemometers. The higher-resolution retrie-
al converged in fewer iterations. However, the CPU per iteration increased significantly. The shape of the variances and flux profiles derived from the retrievals late in the assimilation period agree with expected profiles in the CBL. We believe that these results are encouraging and demonstrate that the retrieval method can be a valuable tool for the analysis of boundary layer turbulence.

Acknowledgments. This work was supported by the National Science Foundation (Grant ATM-9908453). Funding for the field measurements, analysis, and manuscript preparation was also provided in part by the Army Research Office under Proposals No. 37522-GS and 43711-EV, and the Cooperative Institute for Research in the Atmosphere (CIRA). We wish to thank Dr. Juanzhen (Jenny) Sun for her consultation on this work and for providing a copy of her code for our reference. We also thank Dr. Yelena Pichugina for assistance with data preparation and display, and members of the Optical Remote Sensing Division at the National Oceanic and Atmospheric Administration, Environmental Technology Laboratory.

REFERENCES


