Assimilating Surface Weather Observations from Complex Terrain into a High-Resolution Numerical Weather Prediction Model

XINGXIU DENG AND ROLAND STULL

University of British Columbia, Vancouver, British Columbia, Canada

(Manuscript received 12 July 2005, in final form 27 June 2006)

ABSTRACT

An anisotropic surface analysis method based on the mother–daughter (MD) approach has been developed to spread valley station observations to grid points in circuitous steep valleys. In this paper, the MD approach is further refined to allow spreading the mountain-top observations to grid points near neighboring high ridges across valleys. Starting with a 3D first guess from a high-resolution mesoscale model forecast, surface weather observations are assimilated into the boundary layer, and pseudo-upper-air data (interpolated from the coarser-resolution analyses from major operational centers) are assimilated into the free atmosphere. Incremental analysis updating is then used to incorporate the final analysis increments (the difference between the final analysis and the first guess) into a high-resolution numerical weather prediction model. The MD approaches (including one with shoreline refinement) are compared with other objective analysis methods using case examples and daily mesoscale real-time forecast runs during November and December 2004. This study further confirms that the MD approaches outperform the other methods, and that the shoreline refinement achieves better analysis quality than the basic MD approach. The improvement of mountain-top refinement over the basic MD approach increases with the percentage of mountain-top stations, which is usually low. Higher skill in predicting near-surface potential temperature is found when surface information is spread upward throughout the boundary layer instead of at only the bottom model level. The results show improved near-surface forecasts of temperature and humidity that are directly assimilated into the model, but poorer forecasts of near-surface winds and precipitation, which are not assimilated into the model.

1. Introduction

This paper presents a method to fully incorporate both dense local surface weather observations from complex terrain and existing coarse-resolution 3D analyses into a cost-efficient higher-resolution objective analyses for local mesoscale modeling. The goal is to reduce near-surface high-resolution numerical weather prediction (NWP) errors in mountainous terrain.

Many parts of the world have complex mountain ranges with valleys that are narrow and long with kinks and twists. To tackle the analysis problems in such complex terrain, Deng and Stull (2005, hereafter DS05) developed a mesoscale analysis method using a mother–daughter (MD) approach to spread valley observations along circuitous valleys. The MD approach is further refined here for mountain-top observations, as will be described in section 2.

Section 3 describes the NWP model and surface weather observations used here. Surface observations are valuable for mesoscale data analysis, assimilation, and forecasting (Yee and Jackson 1988; Stauffer et al. 1991; Miller and Benjamin 1992; Ruggiero et al. 1996). Ruggiero et al. (2000) demonstrated that a combination of continuous assimilation of satellite imager data and intermittent assimilation of hourly surface observations led to a better depiction of circulations caused by cloud-shading contrasts. Alapaty et al. (2001) found that assimilating surface data into NWP models led to significant reduction in atmospheric boundary layer (BL) modeling errors, which is beneficial to subsequent air pollution forecasts.

Methodology for assimilating surface observations together with a coarse analysis into a high-resolution

Corresponding author address: Xingxiu Deng, Numerical Weather Prediction Division, Canadian Meteorological Centre, Environment Canada, 2121 Trans-Canada Highway, Dorval, QC H9P 1J3, Canada.
E-mail: xingxiu.deng@ec.gc.ca

DOI: 10.1175/MWR3332.1

mesoscale model is detailed in section 4. The method is tested with case studies and daily real-time runs in sections 5 and 6, respectively. A summary is provided in section 7.

2. Mountain-top refinement

The MD approach in DS05 works well for surface observations in valleys, but could be problematic for mountain-top observations. The MD approach is not able to spread observation information from one mountain top to a neighboring one, because the deep valley in between attenuates the information spread. The portion of information spread is called sharing factor (SF) ranging from 0 to 1. As explained in DS05, the spreading of observation data from a mountain top to grid points near other mountain tops is desirable for shallow cold-air pooling events when the mountain tops are all protruding into the same free atmosphere air, and for deep BL events when the mountain tops are all in the same BL air. Therefore, the MD approach is further refined here to allow such spreading by treating mountain-top observations differently than valley observations.

A modified SF for a mountain-top observation is proposed as follows:

$$S_{oa}^{\mathrm{MT}} = S_{o}^{\mathrm{MT}}\left[1 - \left(\frac{|Z_o - Z_a|}{\mathrm{zref2}}\right)^b\right] \tag{1}$$

for $|Z_o - Z_a| < \mathrm{zref2}$; otherwise $S_{oa}^{\mathrm{MT}} = 0$. The superscript MT indicates mountain-top observations; the subscripts o and a represent an observation and an analysis grid point (GP), respectively. Here S_{o}^{MT} is the SF at the mountain-top observation location, which is 1.0. Parameter b controls the analysis decorrelation rate. The SF at any surrounding analysis GP (S_{oa}^{MT}) depends upon the elevation (Z) difference between the observation and GP, and upon the "level-top" BL depth parameter zref2 (which was defined in DS05). No iteration is needed.

To use this approach, one must be able to distinguish a mountain-top observation from a valley observation. To do this, we first smooth the model terrain heights by the Barnes (1964) method. The shape factor (R_v) of the empirical Gaussian weights is taken as 90 km, which is the same value as used for the horizontal correlation length scale in the first pass of the Bratseth (1986) successive-correction scheme used in the surface data analysis. Second, the standard deviation (σ_Z) of the difference between the model terrain height and the smoothed model terrain height is calculated within a region of radius R_v around each GP. If Z_s is the smoothed model terrain height, then a zeroth-order estimate of the height of valley BL top (Z_{BL}) can be approximated by

$$Z_{\mathrm{BL}} = Z_s + \max(0., \mathrm{zref2} - \sigma_Z), \tag{2}$$

where zref2 is the level-top BL depth parameter in (1). The valley BL top here acts like a capping inversion top, which is used to distinguish valley locations (below the base of the free atmosphere) from mountain-top locations (within the free atmosphere). The approximated valley BL top follows the smoothed model topography, but with a positive displacement (zref2 − σ_Z). Over complex terrain with varying ridges and valleys, σ_Z is larger and hence the displacement is smaller. Over flat terrain, σ_Z is smaller and therefore the displacement is larger. In the case of zero σ_Z, the displacement equals the level-top BL depth parameter zref2 used in the SF equations [(1) in this paper and (5) in DS05].

If the unsmoothed terrain height for any GP is above Z_{BL}, then that GP is assumed to be in the free atmosphere, and is treated as a mountain-top location. The approximation of Z_{BL} is crude, but provides a simple and effective way to distinguish mountain-top from valley observations. As mentioned in DS05, for any station not collocated with a model GP, the station is first approximated as being collocated at the nearest-neighboring model GP having the least elevation difference between them. In such a case, the elevation of the nearest-neighboring model GP is compared with Z_{BL}. Any surface weather station that is above Z_{BL} is treated as a mountain-top station, whereas any surface station that is below Z_{BL} is treated as a valley station. The SF for a mountain-top station is obtained through (1), whereas the SF for a valley station is calculated via Eq. (5) in DS05.

To illustrate, Fig. 1 shows the model topography of test domains of 3- and 2-km horizontal grid spacing in south British Columbia (BC), Canada, where our assimilation and forecast experiments are conducted.

Figure 2 shows cross sections of the model topography, the highly smoothed model topography (Z_s), the approximated height of valley BL top (Z_{BL}), and σ_Z. The top (bottom) panel in Fig. 2 is along the projection line in the top (bottom) panel of Fig. 1. Any surface station or GP above the dot–dashed Z_{BL} line will be treated as a mountain-top station or GP.

To further illustrate, Fig. 3 shows terrain heights in the Coast Mountains of BC, just north of Vancouver [Lower Fraser Valley (LFV) in Fig. 1]. With the same case example (7–8 March 2003) as in DS05, a virtual mountain-top observation at o3 (Fig. 3) is extracted from the "truth" model surface forecasts of "fraternal twin" experiments (see DS05). The observation incre-

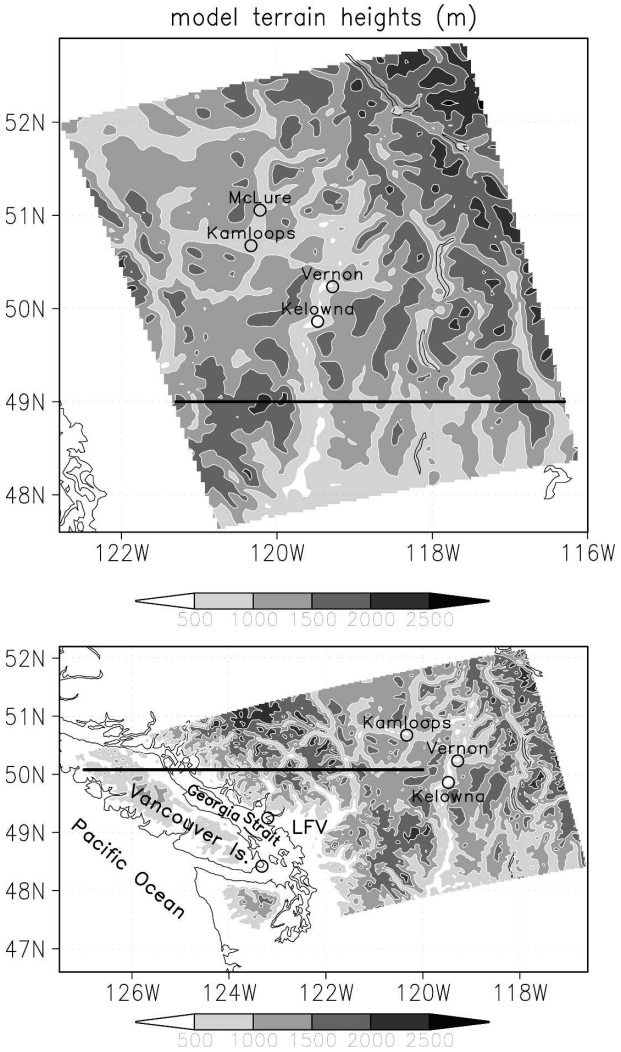

FIG. 1. The model topography (darker shading indicates higher elevations) for (top) 3- and (bottom) 2-km grid spacing. The black lines are the projection lines for the cross sections in Fig. 2. See Fig. 5 for the geographic relationship between these two domains. The cities of Vancouver and Victoria, British Columbia, are also indicated with unlabeled circles, near the LFV and the southern tip of Vancouver Island, respectively.

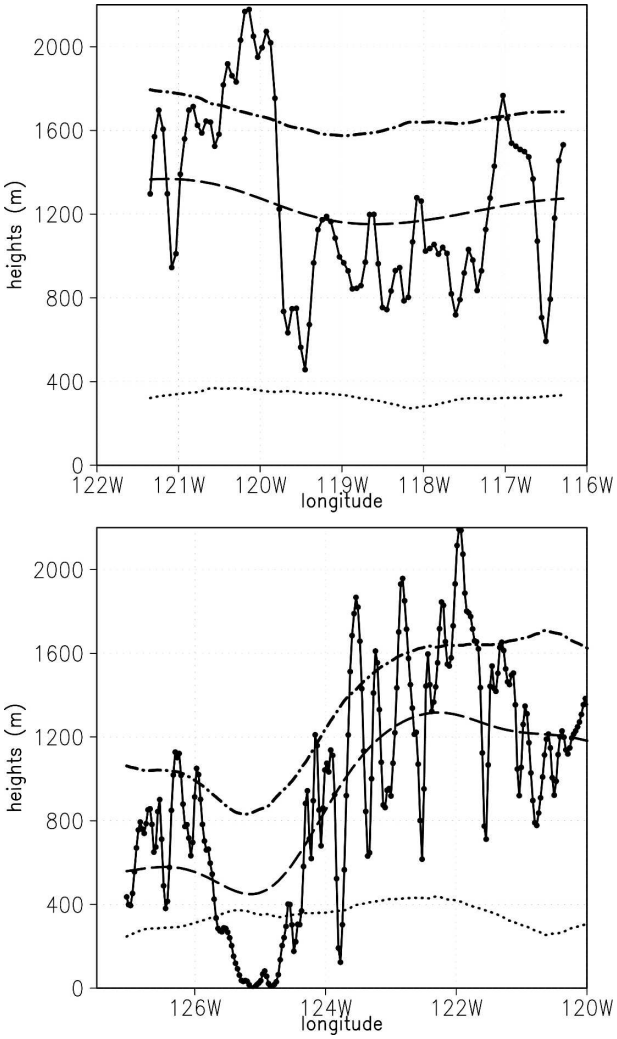

FIG. 2. Cross sections of the model topography (solid line), the smoothed model topography (dashed line), the approximated height of valley BL top (dot–dashed line) and the std dev of the difference between the model topography and the smoothed model topography (dotted line). (top) The cross section is along 49.0°N cutting through the 3-km domain (see the black line in the top of Fig. 1). (bottom) The cross section is along 50.08°N cutting through the 2-km domain (see the black line in the bottom of Fig. 1). Any surface station located above the dot–dashed line is treated as a mountain-top station. The same is true for GPs; namely, the regions where the solid line is above the dot–dashed line are considered to be mountain-top locations.

ment (observation minus the first guess from the "analysis" model forecasts) is −1.56 K. The analysis increments (analysis minus the first guess) produced by the observation o3 using the MD approach after mountain-top refinement (method MD_MT) are shown in Fig. 4. Method MD_MT successfully spreads the mountain-top observation increment to other mountain tops, while giving very small spread in the valleys.

In DS05, the MD approach was compared with two other published methods: GAUSS is the generic Gaussian spread built into the Advanced Regional Prediction System (ARPS; Xue et al. 2000) Data Assimilation System (ADAS; Brewster 1996); and TERR_DIFF is the terrain difference method of Miller and Benjamin (1992). See DS05 for more details about GAUSS and TERR_DIFF. Redoing that fraternal twin experiment from DS05, but now including method MD_MT, yields the analysis verification statistics in Tables 1 and 2. When only the mountain-top observation at o3 is analyzed, Table 1 shows that MD_MT verifies second best

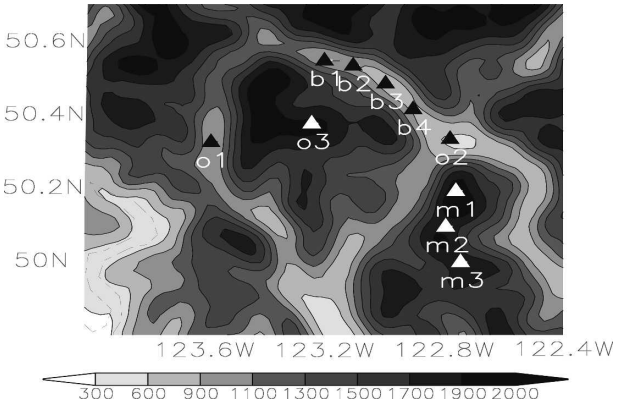

FIG. 3. Virtual surface observation stations (solid triangles) used in the analysis and verification. The stations are positioned at the truth model terrain height in the Coast Mountains north of Vancouver. Station o1 (50.326°N, 123.578°W) is near the mouth of the Elaho River. Station o2 (50.3344°N, 122.767°W) is near the town of Pemberton. Station o3 (50.3773°N, 123.2363°W) is over the mountain top. Terrain elevations (m) are from the truth model (MC2 2 km). Darker shading indicates higher elevations, with a maximum elevation difference of 2055 m in this figure. The dashed line in lower-left corner is a fjord (Jervis Inlet).

(after MD) against valley stations only, and best when verified against mountain stations only. When two valley observations (o1 and o2 from Fig. 3) and one mountain-top observation (o3) are analyzed, Table 2 shows that MD_MT verifies best against all the other valley and mountain stations combined.

To further test MD_MT, we next describe the NWP model and methodology used for one case study (a forest fire event) and for 2 months of daily real-time runs using real data.

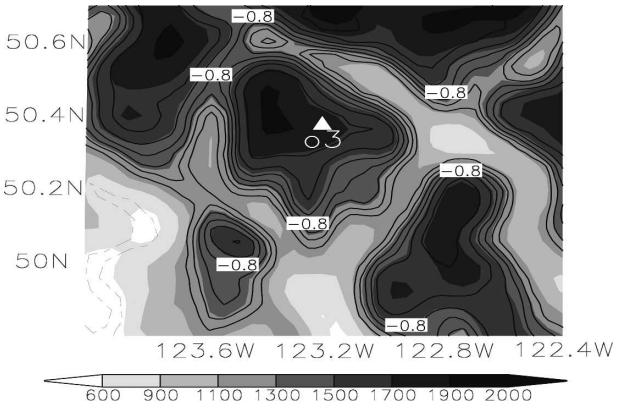

FIG. 4. The AIs (isopleths) from method MD_MT for potential temperature at the lowest (terrain following) model level in response to a single mountain-top observation at o3 (white solid triangle). The contour interval is 0.2 K. Darker shading indicates higher elevations (m). The observation increment at o3 is −1.56 K.

TABLE 1. The NRMSEs between the analyses and observations for different sets of the verification stations. One set is the valley observations (b1, b2, b3, and b4) from Fig. 3; the other is the mountain-top observations (m1, m2, and m3). The analyses were produced when only the single mountain-top observation o3 is used. The observation stations are shown in Fig. 3. NRMSE is RMSE for each method normalized by the RMSE of the FSTG. Smaller NRMSE corresponds to better analyses. NRMSE close to 1.0 indicates that the observations contributed relatively little value relative to the first guess.

Method	Verified against b1, b2, b3, and b4	Verified against m1, m2, and m3
FSTG	1.0	1.0
GAUSS	1.103	0.275
TERR_DIFF	1.406	0.214
MD	1.001	1.0
MD_MT	1.003	0.163

3. The numerical model and data

a. The numerical model

The NWP model into which the surface data are assimilated is the Mesoscale Compressible Community (MC2) model (Laprise et al. 1997; Benoit et al. 1997) version 4.9.1. It uses vertically stretched Gal–Chen terrain-following coordinates to obtain greater resolution near the surface.

MC2 is configured with five one-way self-nested grids with horizontal grid spacings of 108, 36, 12, 4, and 2 or 3 km. The National Centers for Environmental Prediction (NCEP) Eta Model[1] analysis and forecasts from the "104" grid (with a 90.7-km grid spacing true at 60°N) were used as the initial and boundary conditions for the coarsest MC2 grid. The Eta forecasts were available every 3 h up to 84 h. We start from the 104-grid output because this domain extends far enough west over the Pacific to reduce upstream boundary errors, and also because this output was continuously available and timely accessible via the Internet.

The 4-, 2-, and 3-km meshes (Fig. 5) have 35 layers (18 below 1500 m AGL) in the vertical, with the model top at 23 km. The first "thermodynamic" ("momentum") level is located at 5.3 (10.6) m above the model ground.

Objective analyses of surface potential temperature and specific humidity are performed at the first thermodynamic level. A data assimilation (DA) experiment for the case study is performed on the 3-km domain (Fig. 5 and Fig. 1, top), centered at Vernon in south-

[1] Effective 25 January 2005, the Eta Model has been renamed the North American Mesoscale (NAM) model.

TABLE 2. Same as in Table 1, but for all of the verification stations (the valley stations b1, b2, b3, and b4, and the mountain-top stations m1, m2, and m3 from Fig. 3). The analyses were produced using two observations (o1 and o2) in different valleys and one mountain-top observation o3.

Method	Verified against b1, b2, b3, b4, m1, m2, and m3
FSTG	1.0
GAUSS	1.759
TERR_DIFF	1.967
MD	0.972
MD_MT	0.415

central BC. Real-time DA runs are performed for the 2-km mesh (Fig. 5 and Fig. 1, bottom). For all DA runs on 3- or 2-km meshes, five GPs are cut from each side of the analysis domains shown in Fig. 5, to reduce the effects of lateral boundary errors.

b. Data

Hourly surface weather observations are from the real-time "Emergency Weather Net Canada" (EmWxNet) database, which includes surface temperature, relative humidity, wind speed/direction, and mean sea level pressure (or surface pressure at some stations). Temperature and humidity are used for both assimilation and verification, while other variables are used for forecast verification only.

Potential temperature varies smoothly over mountainous terrain when the BL is relatively deep and well mixed (Miller and Benjamin 1992); thus, it is chosen as the temperature variable for analysis. It is also the temperature variable analyzed in ADAS.

Specific humidity is chosen as the moisture variable for analysis and assimilation into MC2 for two reasons: 1) MC2 uses specific humidity; and 2) specific humidity is a continuous variable. Specific humidity is also the moisture variable analyzed in ADAS. To derive specific humidity from relative humidity, the saturation specific humidity is first calculated using the enhanced Teten's formula, as in ARPS. This calculation requires surface temperature and station pressure as input. There are few reports of station pressure from the EmWxNet. In the case of missing station pressure, mean sea level pressure is used to approximate station pressure according to the altimeter equation (standard atmosphere) for stations with elevations below 500 m.

4. Methodology

Assimilating surface weather observations into a high-resolution NWP model in this study involves three major components: horizontal spreading, vertical spreading, and data insertion.

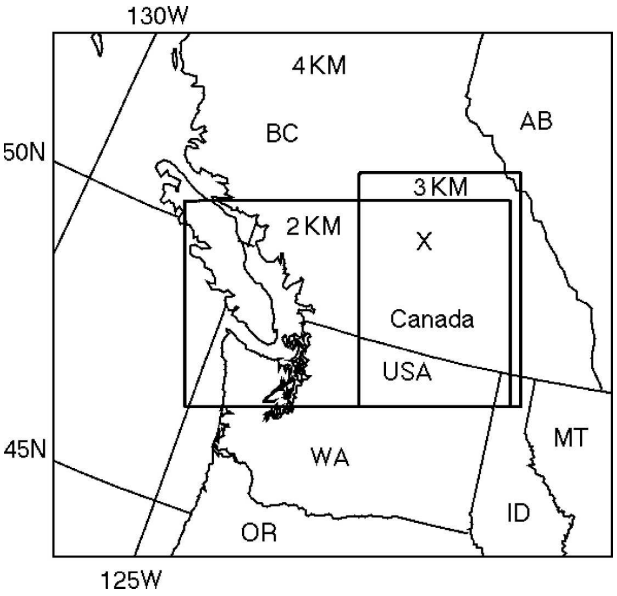

FIG. 5. MC2 grid domains for $\Delta x = 4$, 3, and 2 km. MC2 4-km output provides nesting files to initialize 2- and 3-km runs. MC2 at 3 or 2 km is run to provide first-guess fields for analysis, and also for assimilation runs after analysis. Five grid points are cut from each side of the analysis domain shown here, to reduce lateral boundary effects in the assimilation. The forest fire location near the town of McLure is indicated with the X.

a. Horizontal spreading of surface observations

The isotropic assumption typically used in background error correlation models for horizontally spreading observations is not valid for mountainous terrain. Instead, surface observations are analyzed anisotropically using MD_MT with ADAS to generate a surface data analysis (see DS05). This is done by optimally combining surface observations and a high-resolution gridded first guess from the lowest terrain-following level of a previous MC2 run. The parameters used in each analysis method are the same as in DS05.

b. Vertical spreading of surface observations and combination with pseudo-upper-air data

Surface observations are available at only one terrain-following level. Meanwhile, major operational centers generate 3D meteorological analyses daily by assimilating many types of measurements (rawinsondes, aircraft, wind profilers, satellite and radar, etc). This section describes how to spread surface information upward, and to combine it with pseudo-upper-air data interpolated from a coarser, 3D analysis from the major operational centers.

To obtain pseudo-upper-air data on the terrain-following MC2 3-km (or 2 km) model levels, MC2 is utilized as an interpolation tool here. This is done by integrating MC2 for only 1 h and self-nesting the coarsest grid (108 km) successively down to 36, 12, 4, and finally to the finest one (3 or 2 km). The resulting MC2 3-km (or 2 km) output on terrain-following levels at 0 h are the desired pseudo-upper-air data. By assimilating this interpolated Eta data, we effectively incorporate all the satellite, radar, aircraft, and rawinsonde data that were assimilated by NCEP.

At this stage, we have a surface data analysis, pseudo-upper-air data, and the first guess from a previous MC2 3-km (or 2 km) forecast, all on the MC2 model levels and at every model grid column. In devising a profile (PROF) scheme to combine the three sets of data, we adopt assumptions similar to the ones of Yee and Jackson (1988): surface observations describe atmospheric state in the BL, while coarser, 3D analyses from major operational centers provide information above the BL. In the PROF scheme, potential temperature and specific humidity analyses at the lowest model level are applied to the whole BL. Above the BL top in the free atmosphere, the final analysis F^A is a weighted average of the first guess F^B and the pseudo-upper-air data F^{UA}:

$$F^A(r, k) = [1 - W(r, k)]F^B(r, k) + W(r, k)F^{\mathrm{UA}}(r, k), \tag{3}$$

where indexes r and k correspond to horizontal position and vertical level, respectively. Weights $W(r, k)$ in (3) depends on the ratio of error variances of the pseudo-upper-air data and first guess:

$$W(r, k) = \frac{1}{1 + \left[\dfrac{\sigma^{\mathrm{UA}}(r, k)}{\sigma^B(r, k)}\right]^2}. \tag{4}$$

As the pseudo-upper-air data are obtained by interpolating the Eta analysis to the MC2 3-km (or 2 km) model levels, the error variance of the pseudo-upper-air data is assumed to be the same as that of the first guess: $\sigma^{\mathrm{UA}}(r, k) = \sigma^B(r, k)$.

The BL height is diagnosed from a profile method using a slab idealization of the mixed layer (Stull 2000), where the potential temperature jump in the entrainment zone is taken as $\Delta\theta = 1.5$ K. Namely, the analyzed potential temperature at the lowest model level $\theta^{\mathrm{SA}}(1)$ is compared with the potential temperature profile of the pseudo-upper-air data (θ^{UA}) at successively higher grid points. When $\theta^{\mathrm{SA}}(1) - \theta^{\mathrm{UA}}(1) > -1.5$ K, the BL depth Z_i is the height (above the model ground) of the model level k at which the criterion of $\theta^{\mathrm{UA}}(k) - \theta^{\mathrm{SA}}(1) \geq 1.5$ K is first met. When $\theta^{\mathrm{SA}}(1) - \theta^{\mathrm{UA}}(1) \leq 1.5$ K, Z_i is assumed to be not shallower than 300 m, which represents nocturnal BLs or shallow BLs with cold-air pooling. Figure 6 illustrates the relationship between the BL depth Z_i and $\theta^{\mathrm{SA}}(1) - \theta^{\mathrm{UA}}(1)$.

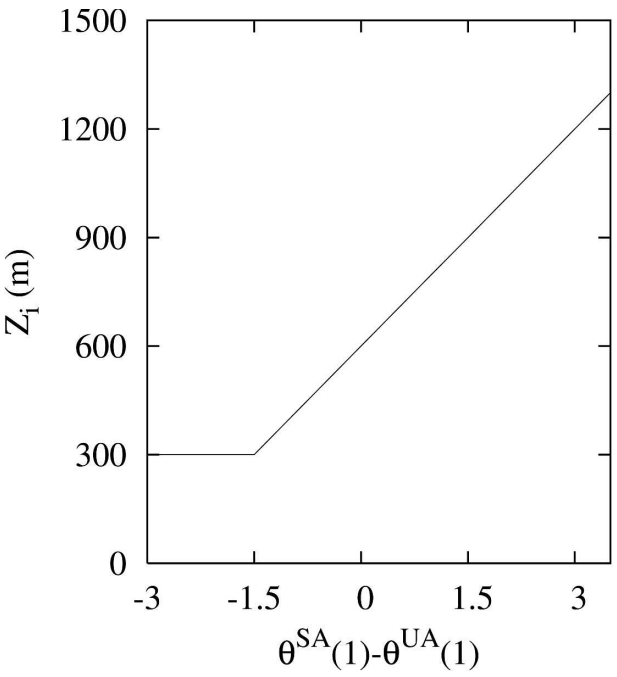

FIG. 6. Schematic diagram illustrating the relationship of the BL depth Z_i and the difference between the analyzed potential temperature and the interpolated Eta analysis at the bottom model level [θ^{SA} (1) − θ^{UA} (1)]. The actual slope of the diagonal line depends on the upper-air profile.

c. Insertion of analysis increments

The objective analysis procedures used in this work (same as in DS05) do not provide mass and motion fields that are in optimal balance to initiate NWP. This study uses the incremental analysis updating (IAU; Bloom et al. 1996), instead of an explicit initialization procedure. In IAU, the analysis increments (AIs) are gradually incorporated into the MC2 model. Because MC2 uses a process-splitting method (Bergeron et al. 1994) for its different prognostic terms, it was convenient to apply the AIs after the physical parameterizations in the vertical, but before the treatment of horizontal diffusion and the Asselin–Robert time filter.

IAU is implemented in MC2 largely following the existing ADAS IAU (Brewster 2003; Bloom et al. 1996). Parallelization of IAU uses the software architecture of MC2. In their applications of ADAS IAU (Brewster 2003; Xue et al. 2002), the analysis incre-

ments for radar data were introduced to a storm-scale numerical model employing a constant time weighting over a 10-min period. In our work, the analysis increments are introduced into the model over a 1-h IAU window using a constant time weighting (sections 5 and 6).

5. Case study results

Numerical experiments are performed for 29–30 July 2003 when a large forest fire occurred near McLure, BC (see Fig. 5). During that time, fair weather associated with a 50-kPa ridge over southern BC allowed BL processes to dominate. The 3-km domain (Fig. 5) is used for this assimilation case study.

The Eta Model analysis and forecasts, started at 0000 UTC 29 July 2003, are used to drive the 108-km grid, which in turn drives the nests of finer grids (see Fig. 7). The MC2 3-km control run (CTRL) is started at 1200 UTC 29 July 2003 from the MC2 4-km output, and provides its 12-h output as the first guess for the analysis at 0000 UTC 30 July 2003. Surface analyses are performed for 0000 UTC 30 July 2003 by blending hourly surface observations with the first guess at the lowest terrain-following model level valid at the same time. Those surface stations with a difference between their actual elevation and model topography greater than 500 m are excluded from analysis and verification. Out of the total 134 stations in the DA domain, 20 stations were excluded for this reason.

Pseudo-upper-air data are obtained from the Eta analysis valid at 0000 UTC 30 July 2003. A final high-resolution analysis at 0000 UTC 30 July 2003 is produced by combining the surface analysis, pseudo-upper-air data, and first guess using the PROF scheme. A DA run at 3-km grid spacing is started at 0000 UTC 30 July 2003 from the final analysis via the IAU technique. Lateral boundary conditions for the DA runs and for the CTRL run are the same, and are from the MC2 4-km output. Verification of subsequent model forecasts at the lowest model level against surface observations is performed during a 24-h forecast period for the DA forecast run, and are compared with verification of the CTRL run during the 12–36-h forecast period from 0000 UTC 30 July to 0000 UTC 31 July 2003. The verification of both the surface analysis and forecast fields are presented next.

Note that the comparison against this CTRL run is not strictly optimal, since this CTRL run starts earlier than the DA runs. This suboptimal choice of CTRL implies a probable favor in DA experiments compared to CTRL. A much more appropriate CTRL run would start from the analysis of an operational model valid at the same initial time as that of all other experiments. The purpose of this paper is not only to demonstrate the usefulness of surface data but also to assess the impact of data assimilation on real-time high-resolution model runs that do not have DA module originally. In this paper, we sacrifice the optimum, in order to verify the assimilation strategy over a long time period in real time without modifying the current daily real-time forecast system. Nevertheless, one should keep this problem of suboptimal CTRL in mind when comparing the result from one DA experiment against that from CTRL. The problem with suboptimal CTRL is minimized when comparing the results from two DA experiments.

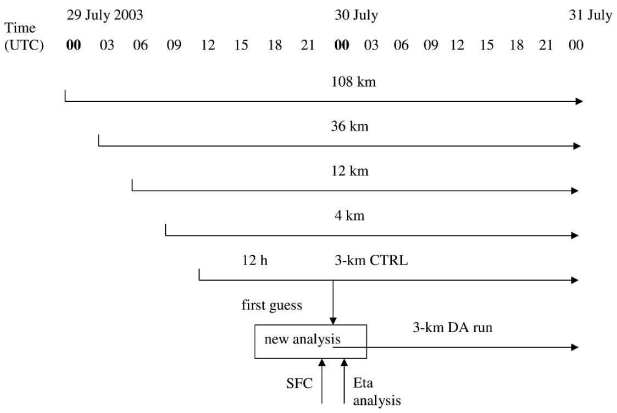

FIG. 7. Schematic diagram illustrating model time lines of the MC2 self-nested grids including the 3-km CTRL and DA runs. MC2 4-km provides boundary conditions for the 3-km CTRL run and the 3-km DA runs. SFC indicates surface data.

a. Analysis verification

Figure 8 shows all 118 surface stations within the analysis domain. The domain for DA runs is smaller than the analysis domain, leaving 114 stations within the DA domain for verification. Among them, 65 stations (indicated by closed triangles) reported temperature observations at 0000 UTC 30 July 2003. Observations separated by 100 m or less in the horizontal and vertical are averaged to create a smaller number of "superobservations." Because only a subset of these stations also reported relative humidity and pressure, there are only 16 stations for which specific humidity could be calculated (Fig. 9).

Table 3 shows that all methods give improved potential temperature analyses compared to the first guess (FSTG) as measured by bias, mean absolute error (MAE), and normalized RMSE (NRMSE). Both MD and MD_MT outperform GAUSS and TERR_DIFF; MD produces the lowest bias and MAE, whereas

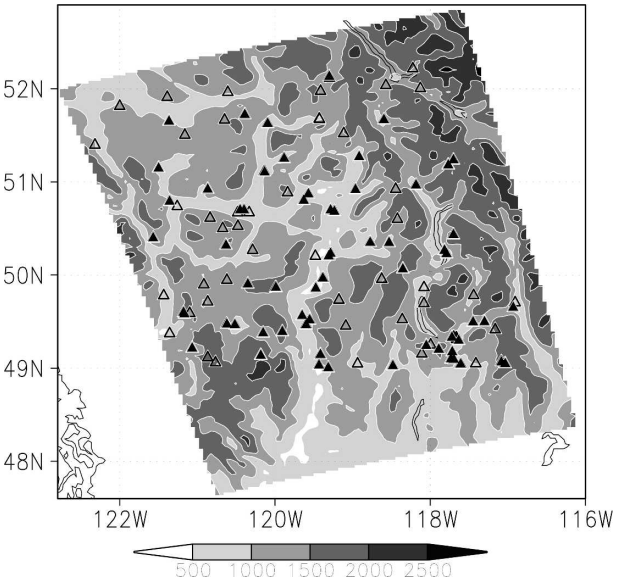

FIG. 8. Surface weather stations (triangles) are superposed on the model topography (m) for the 3-km domain shown in Fig. 5. Darker shading indicates higher elevations. Open triangles represent station locations with missing temperature observations at the analysis time (0000 UTC 30 Jul 2003), whereas closed triangles indicate station locations reporting temperatures at the analysis time.

MD_MT has the lowest NRMSE. The improvement of MD_MT over MD is not as large as that for the virtual observation case in section 2. One reason is that only 5% of the stations (3 out of 65) are identified as mountain-top observations in this case, while 33% (1 out of 3) were from mountain top for the virtual observation case.

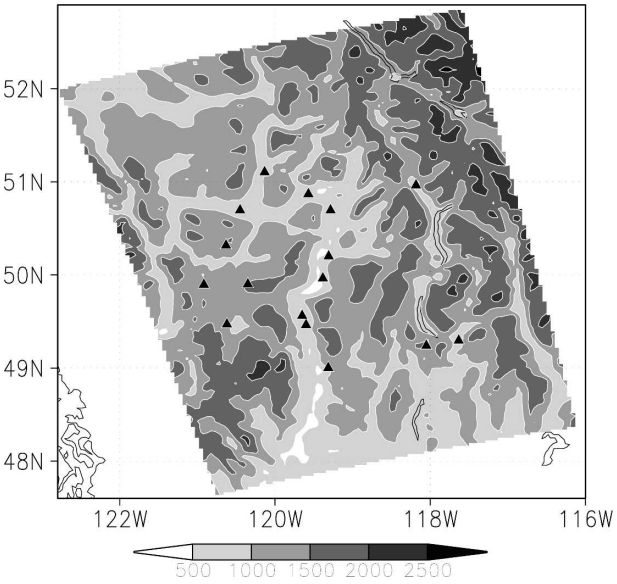

FIG. 9. Surface station locations (closed triangles) with sufficient data to calculate specific humidity at the analysis time (0000 UTC 30 Jul 2003) and the model topography (m). Darker shading indicates higher elevations.

TABLE 3. Verification of analyzed potential temperatures for all reporting verification stations at the analysis time: 0000 UTC 30 Jul 2003. Here *N* equals the total number of observations. For all these statistics, smaller is better.

Method	*N*	Bias (K)	MAE (K)	NRMSE
FSTG	64	−5.258	7.687	1.0
GAUSS	64	0.758	2.099	0.515
TERR_DIFF	64	0.712	2.324	0.546
MD	64	0.089	1.600	0.391
MD_MT	64	0.133	1.610	0.387

For specific humidity, Table 4 shows that MD and MD_MT outperform GAUSS and TERR_DIFF. No specific humidity observations are from the mountain-top stations, therefore MD_MT = MD. The MD approach with shoreline refinement (MD_LSMG, see DS05) is not studied here, because there are no stations over water in this domain.

The results above confirm the conclusion from DS05: the MD approach outperforms GAUSS and TERR_DIFF in producing better surface analyses. The results also show that the mountain-top refinement provides a small improvement over the basic MD approach. In section 5b, the impacts of assimilating surface weather observations using MD_MT on subsequent near-surface forecasts are examined.

b. Verification of MC2 forecasts

1) FORECAST EXPERIMENTS

Eight experiments are conducted: one control experiment without assimilation (CTRL) and seven other experiments, which assimilate various combinations of the surface and pseudo-upper-air data, use different insertion rates, or assimilate different meteorological fields (see Table 5).

Experiment SURFDA assimilates the surface data into only the lowest model level, and uses the first guess at all higher model levels. Experiment SF_PROF uses the surface analysis only, but the AIs at the lowest model level are spread vertically by the PROF scheme to the whole BL. Experiment PROFDA uses the surface and pseudo-upper-air data combined by the PROF scheme. By comparing SF_PROF and PROFDA, one can see if the assimilation of surface data plays a dominant role on subsequent near-surface forecasts. For the above three experiments, the temperature AIs are incorporated all at once within a single time step (30 s) window.

TABLE 4. Same as in Table 3, but for verification of analyzed specific humidity.

Method	N	Bias (1.0×10^{-4} kg kg^{-1})	MAE (1.0×10^{-4} kg kg^{-1})	NRMSE
FSTG	16	14.941	21.921	1.0
GAUSS	16	−3.595	7.878	0.621
TERR_DIFF	16	−4.480	9.440	0.665
MD	16	−3.181	6.770	0.594
MD_MT	16	−3.181	6.770	0.594

Experiments TH1 (i.e., experiment PROFDA) through TH6 assimilate surface and upper-air temperature fields employing different insertion rates over different IAU time windows. TH1 incorporates AIs all at once within a single 30-s time step. Over a 1-h window, experiments TH2, TH3, and TH6 incorporate AIs every 30, 120, and 1200 s, respectively. The 1-h window is selected because the BL responds to surface forcings with a time scale of about 1 h or less (Stull 1988).

The settings for experiments ATQ and AQ are the same as those for experiment TH6, except that ATQ assimilates both temperature and specific humidity, and AQ assimilates only specific humidity.

2) VERIFICATION RESULTS

Statistical assessments during the 1–12-h forecast period in Table 6 suggest that different DA experiments have variable success at predicting near-surface fields. The seven DA experiments can be divided into three groups. Comparisons are made below within each group.

For potential temperature forecasts, experiments SURFDA, SF_PROF, and PROFDA outperform CTRL. By assimilating surface temperature at only the lowest model level, SURFDA produces very little improvement over CTRL. When surface potential temperature analyses are spread throughout the whole BL (viz. experiment SF_PROF), larger improvement for predicting surface potential temperatures are achieved. By combining surface and pseudo-upper-air data, PROFDA gains only slightly larger improvement than SF_PROF. Thus, assimilating surface data that were spread vertically over the BL was valuable for this forest fire weather case.

Verification of mean sea level pressure (SLP) forecasts reveals similar performance. All experiments except SURFDA outperform CTRL. Overall, improvement of SLP forecasts is smaller than that of potential temperature. This is reasonable as temperature, not pressure, observations are directly assimilated into the model. For wind forecasts, all three DA experiments underperform CTRL in terms of normalized root-mean-square vector error (NRMSVE) due to initial imbalances between mass and wind fields caused by a sudden change in temperature fields. SURFDA has the smallest NRMSVE, probably because the assimilated temperature information is lost soon after insertion.

The impact of varying insertion rate on potential temperature forecasts is much larger than on SLP and wind forecasts (cf. experiments TH1, TH2, TH3, and TH6 in Table 6). The errors of potential temperature decrease significantly from experiment TH1 to TH2 (or TH3 and TH6). This implies that introducing the AIs over a 1-h window helps the model to better retain the assimilated information than adding the increments at only the initial time. The differences between the error measures of experiments TH2, TH3, and TH6 are small. Comparatively, TH6 has the smallest errors for potential temperature and vector wind.

The top panel of Fig. 10 shows NRMSE of potential temperature versus forecast hour for experiments CTRL, TH1, TH3, and TH6. The NRMSEs for the DA experiments are much smaller than 1.0 (NRMSE of CTRL) during the first 12 h (except at 4 h, see explanation for the exception later in this section), and later increase gradually. Finally the NRMSE becomes close to 1.0. By applying the AIs over a 1-h window rather than all at once, TH3 and TH6 produce much smaller NRMSE than TH1 for the first 7 h (except 4 h) of forecasts. Reducing the insertion rate at which the AIs

TABLE 5. Experimental design. Variable T is temperature and q is specific humidity.

Expt name	Surface analysis	UA analysis	Vertical spreading	DA window	Insertion rate	Variables assimilated
CTRL	No					
SURFDA	Yes	No	No	30 s	30 s	T
SF_PROF	Yes	No	Yes	30 s	30 s	T
PROFDA (TH1)	Yes	Yes	Yes	30 s	30 s	T
TH2	Yes	Yes	Yes	1 h	30 s	T
TH3	Yes	Yes	Yes	1 h	120 s	T
TH6 (AT)	Yes	Yes	Yes	1 h	1200 s	T
AQ	Yes	Yes	Yes	1 h	1200 s	q
ATQ	Yes	Yes	Yes	1 h	1200 s	T, q

TABLE 6. Verification of surface potential temperature (θ), vector wind (**V**), and MSLP forecasts in terms of bias during a 12-h forecast period from 0100 to 1200 UTC 30 Jul 2003. NRMSE (NRMSVE) is RMSE (RMSVE) for each experiment normalized by the RMSE (RMSVE) of the control run (CTRL). The DA experiments are defined in Table 5. For all these statistics, smaller (closer to zero) is better.

	1–12-h forecast								
		θ(K) (N = 601)			SLP (kPa) (N = 153)			**V** (m s^{-1}) (N = 506)	
Expt	Bias	MAE	RMSE	NRMSE	Bias	MAE	NRMSE	RMSVE	NRMSVE
CTRL	−4.838	5.666	6.472	1.000	0.376	0.397	1.000	1.745	1.0
SURFDA	−4.853	5.623	6.442	0.996	0.376	0.397	1.000	1.750	1.003
SF_PROF	−4.328	5.288	6.089	0.941	0.321	0.365	0.951	1.968	1.128
PROFDA	−4.163	5.179	5.979	0.924	0.301	0.357	0.935	1.996	1.144
TH2	−3.846	4.885	5.679	0.878	0.269	0.363	0.940	2.024	1.160
TH3	−3.871	4.910	5.700	0.881	0.270	0.362	0.939	2.023	1.160
TH6 (AT)	−3.752	4.792	5.600	0.865	0.277	0.358	0.936	1.991	1.141
AQ	−4.921	5.731	6.554	1.013	0.385	0.408	1.021	1.773	1.016
ATQ	−3.688	4.742	5.548	0.857	0.275	0.357	0.924	2.033	1.166

are introduced decreases NRMSE mainly for the first 1 h of forecasts. Reducing the insertion rate over a 1-h window implies a larger fraction of the AIs is applied at each insertion step and thus experiences less diffusion after 1 h.

The time series of NRMSVE for experiments CTRL, TH1, TH3, and TH6 are shown in the bottom panel of Fig. 10. By assimilating temperatures only, the model produces poorer wind forecasts than CTRL during the first 12 h due to initial imbalances between the mass and wind fields. During the second 12 h, all DA experiments and CTRL produce similar forecasts. This could be because the assimilated information propagates out of the domain. Another reason could be that the mass and wind fields are adjusted to be in balance. By applying the temperature AIs over a 1-h window rather than all at once (thus reducing initial imbalances), TH3 and TH6 decrease NRMSVE for the 1-h forecast compared to TH1, but increase NRMSVE a little during the 2–8 h. Reducing the rate at which the AIs are introduced further decreases NRMSVE for the 1-h forecast while reducing the growth of NRMSVE later.

The summed total precipitation amounts over the entire domain for experiments CTRL, TH1, TH2, TH3, and TH6 are listed in Table 7. Actual observations (not listed) show trace to very light precipitation only at some stations during the simulation period. Compared to CTRL, data assimilation experiments result in much more precipitation during the first 3 forecast hours, and slightly smaller precipitation afterward. By applying the AIs over a 1-h IAU window rather than all at once, TH2, TH3, and TH6 reduce excessive precipitation during the first 3 forecast hours compared to TH1. But the precipitation forecasts from the DA runs remain poorer than from the CTRL run. This may imply that a better initialization step is required when assimilating surface data into the NWP model. Also, in steep mountainous regions much of the precipitation is orographically modified, so poor wind forecasts lead to poor precipitation forecasts.

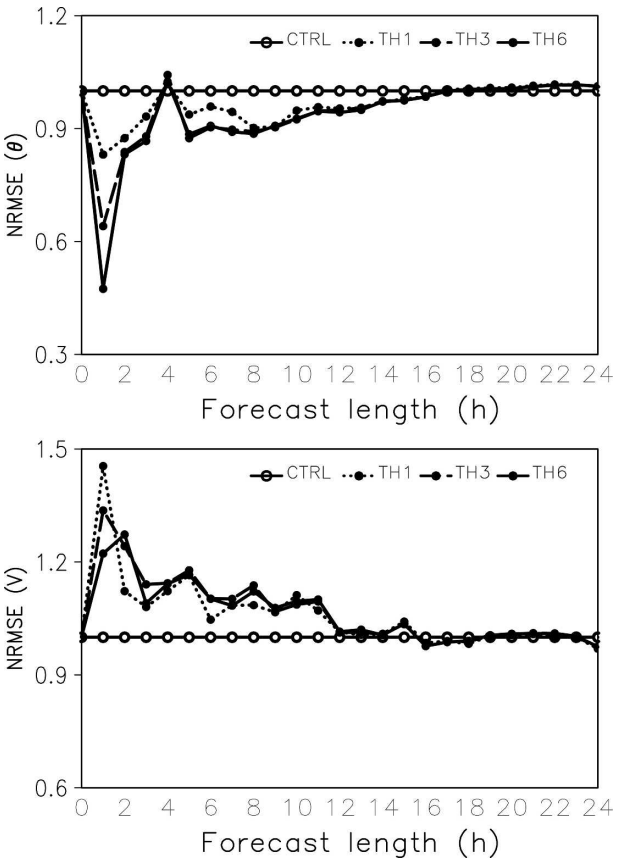

FIG. 10. Time series, for (top) NRMSE of surface potential temperature and (bottom) NRMSVE of surface winds from experiments CTRL, TH1, TH3, and TH6. Smaller NRMSE (or NRMSVE) corresponds to better forecasts.

TABLE 7. The summed total precipitation amounts (mm) over the entire domain from experiments CTRL, TH1, TH2, TH3, and TH6.

Expt	0–3 h	3–6 h	6–9 h	9–12 h
CTRL	96.9	0.0	0.9	7.6
TH1	1095.6	0.0	0.5	5.0
TH2	651.4	0.0	0.5	4.3
TH3	653.4	0.0	0.5	4.3
TH6	592.3	0.0	0.4	4.1

Now we look at the impact of assimilating different meteorological fields on subsequent model forecasts. For potential temperature and SLP forecasts, experiment AQ (assimilating specific humidity only) produces poorer forecasts than CTRL (Table 6). The forecasts from experiments AT and ATQ are greatly improved compared to the CTRL forecast. By assimilating both temperature and specific humidity, ATQ is improved compared to AT that assimilates temperature only.

The DA experiments AT, AQ, and ATQ give poorer wind forecasts than CTRL. In contrast to the error measures of potential temperature and SLP, NRMSVE for wind is the smallest from AQ, and the largest from ATQ. Winds are more sensitive to temperature perturbations than to moisture perturbations.

Verification statistics of specific humidity for experiments AT, AQ, and ATQ are shown in Table 8. Assimilating only specific humidity (experiment AQ) greatly decreases bias, MAE, and NRMSE in the forecasts of specific humidity compared to CTRL. ATQ has smaller MAE and NRMSE than AT.

The impact of assimilating different meteorological fields on the diurnal variation of surface temperature is examined at three stations near McLure and three stations in the Okanagan Valley (Fig. 11). McLure and the Okanagan Valley suffered extensive forest fires in the past few years, for which improved forecasts could have aided fire fighters in saving more homes. The diurnal variations of surface temperature forecasts for experiments AT and AQ are almost the same as experiments ATQ and CTRL, respectively, and therefore not included here.

TABLE 8. Verification of surface specific humidity (q) for the 1–12-h forecast period from 0100 to 1200 UTC 30 Jul 2003. The DA experiments differ in the variables assimilated into the model.

	1–12-h forecast			
	q (1.0×10^{-4} kg kg^{-1}) ($N = 169$)			
Expt	Bias	MAE	RMSE	NRMSE
CTRL	18.883	24.465	29.723	1.0
AT	17.605	24.051	28.801	0.969
AQ	14.082	19.810	24.249	0.816
ATQ	18.410	23.517	28.369	0.954

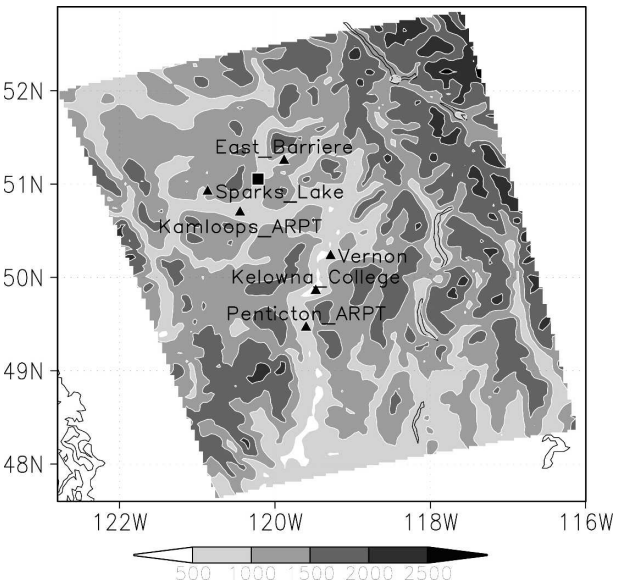

FIG. 11. Three surface stations (closed triangles) near McLure (a forest fire location indicated by a closed square) and three others farther south in the Okanagan Valley. The model terrain heights (m) are shown by shading. Darker shading corresponds to higher elevations. The shaded region corresponds to the 3-km domain in Fig. 5.

Figure 12 compares near-surface temperature forecasts (at the lowest model level, which is 5.3 m above model ground) from experiments CTRL and ATQ to the observed surface temperatures at three stations near McLure every 1 h during the 24-h forecast period. The first guess (from CTRL) underforecasts surface temperature at the analysis time (0000 UTC 30 July 2003). By assimilating local surface temperature observations, experiment ATQ produces improved forecasts. The improvement decreases from 1- to 2-h forecast for these three stations.

Time series of observed and forecasted surface temperatures are also shown for three stations located in the populated Okanagan Valley in Fig. 13, where similar conclusions can be drawn. However, the difference in the magnitude between the observed and forecasted surface temperature is obvious, caused partly by the difference between actual and model elevation (Table 9) in steep terrain.

At 0400 UTC, Figs. 12 and 13 show either a large drop in surface temperature at some stations or missing reports at others. This helps explain the sudden jump of NRMSE for surface potential temperature at 4 h in Fig. 10.

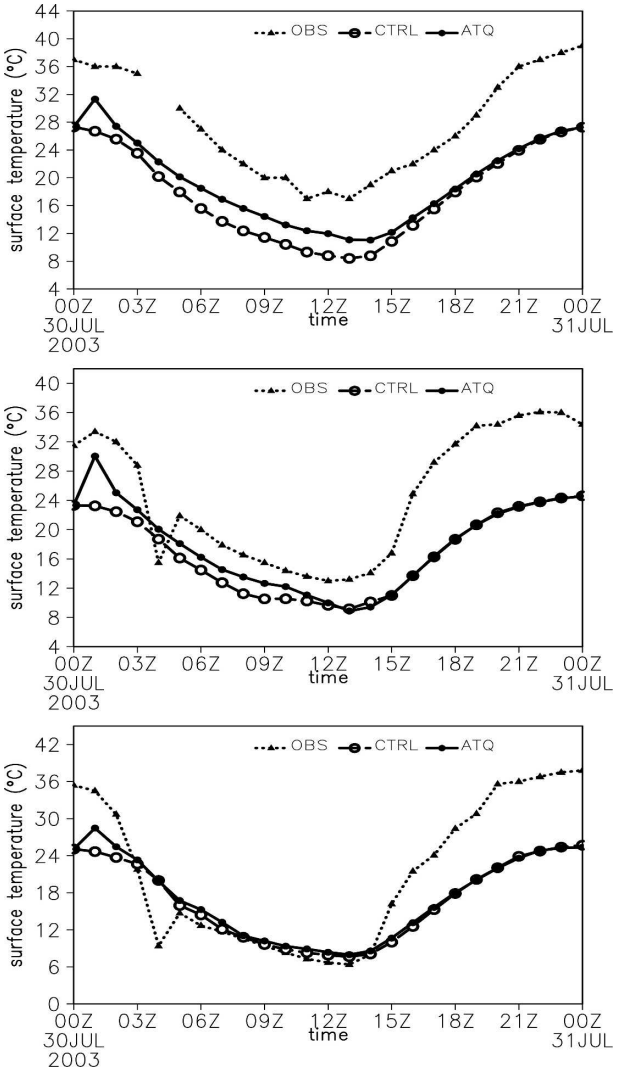

FIG. 12. Time series of observed surface temperature (dotted line with closed triangles) and the forecast temperature at the lowest model level from the control run (CTRL; dashed line with open circles) and from experiment ATQ (solid line with closed circles) for stations near McLure: (top) Kamloops ARPT, (middle) Sparks Lake, and (bottom) East Barriere.

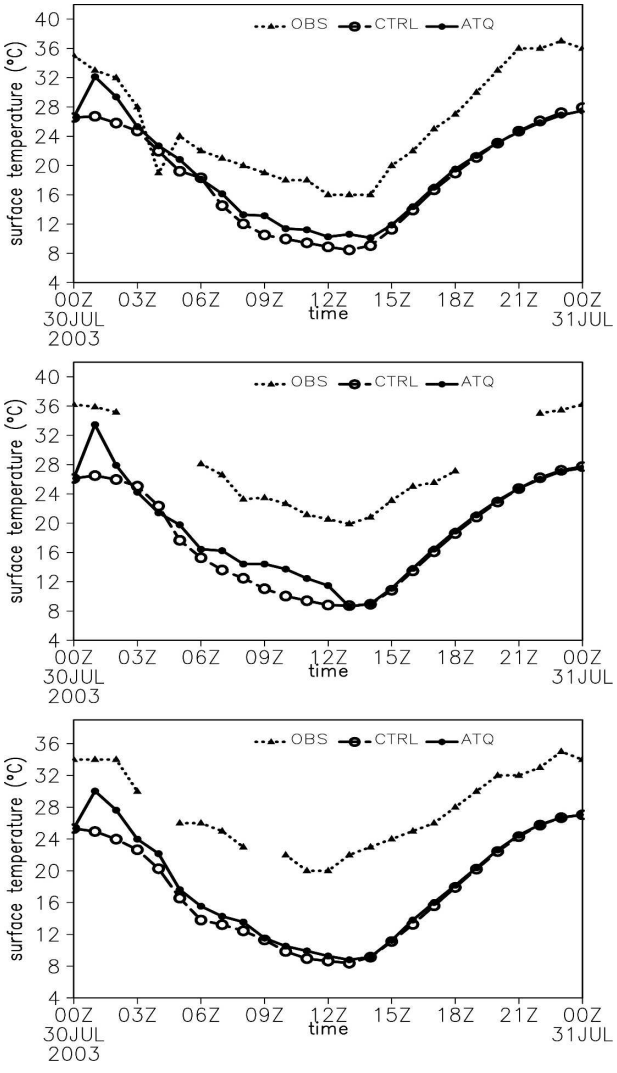

FIG. 13. Same as in Fig. 12, but for stations in Okanagan Valley: (top) Vernon, (middle) Kelowna College, and (bottom) Penticton ARPT.

6. Daily real-time runs

This section further tests MD_MT and the assimilation methodology for robustness over a longer period by applying them in daily real-time forecast runs for 2 months.

a. Experimental setting

The real-time MC2 model at UBC has five one-way self-nested grids with horizontal grid spacings of 108, 36, 12, 4, and 2 km. The DA run is performed daily for only the 2-km grid. The model time lines are shown in Fig. 14. In a real-time setting, the grids at 108, 36, 12, and 4 km are actually run into the future at 1200 UTC on day 3. Limited by computational resources, the primary real-time 2-km grid (hereafter the CTRL) driven

TABLE 9. The actual station elevations and the modeled station elevations for the surface stations shown in Fig. 11.

Station	Actual elev (m)	Modeled elev (m)	Elev error of model (m)
Kamloops ARPT	346.0	541.0	195.0
Sparks Lake	1036.0	1030.0	−6.0
East Barriere	671.0	912.0	241.0
Vernon	556.0	487.0	−69.0
Kelowna College	300.0	399.0	99.0
Penticton ARPT	344.0	569.0	225.0

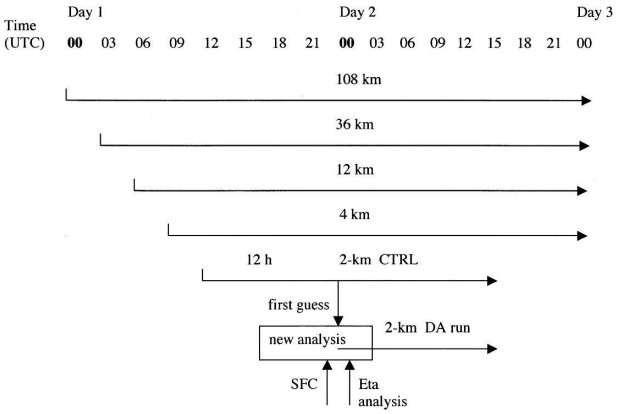

FIG. 14. Schematic diagram illustrating model time lines of the real-time MC2 self-nested grids including the 2-km DA run. The 2-km CTRL run is the regular real-time 2-km run. MC2 4-km provides boundary conditions for the 2-km CTRL run and the parallel 2-km assimilation run. SFC indicates surface data.

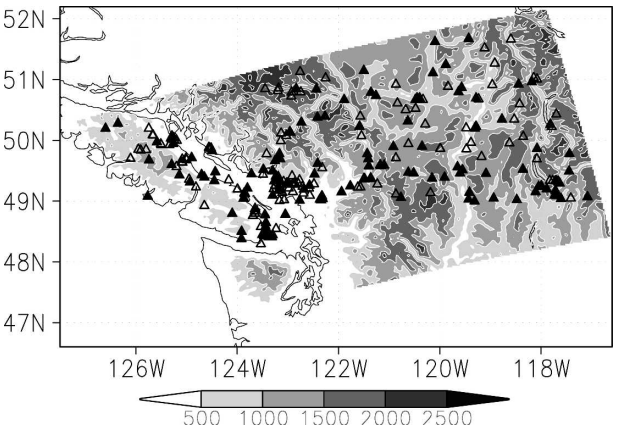

FIG. 15. Surface weather stations (triangles) superposed on the 2-km MC2 model topography (m). Darker shading indicates higher elevations. Open triangles represent station locations with missing temperature observations at this one sample analysis time 0000 UTC 2 Nov 2004, while closed triangles indicate station locations with available temperature observations at that analysis time.

by the 4-km output is run for only 27 h. This CTRL run is started at model-time 1200 UTC on day 1 (Fig. 14), and provides its 12-h output as a first guess for analyses at 0000 UTC on day 2 for the parallel DA run.

The surface observations of potential temperature and specific humidity at 0000 UTC on day 2 are analyzed using the ADAS Bratseth (1986) scheme (Brewster 1996) modified to include the MD approach after mountain-top refinement (MD_MT). The final analyses are obtained by combining the first guess, surface, and pseudo-upper-air data by using the PROF scheme. Out of the total 305 surface stations in the DA domain, 22 stations were excluded from analysis and verification because their actual and model elevation differed by more than 500 m. Stencils of sharing factors at the model grid points are generated for each of the remaining 283 surface stations using MD_MT, so that there is no need to recalculate the sharing factors during each day's analysis. These analyses are verified in section 6b.

The 2-km DA parallel run is started at model time 0000 UTC on day 2 and is integrated forward 15 h. Observation information are gradually incorporated into MC2 every 1200 s over a 1-h window using the IAU technique. Boundary conditions for both CTRL and DA runs were from MC2 4-km output. Verifications of subsequent forecasts at the lowest model level against surface observations are performed during the 15-h forecast period for the DA runs, and are compared with verifications of the CTRL runs during the same valid time. As discussed earlier in section 5, the CTRL runs are not strictly optimal. The improvement of the DA runs over the CTRL runs might be slightly overestimated.

Each day, the analysis obtains its first guess from the 12-h output of the 2-km CTRL run, not from the 2-km DA run. This implies that the local surface observations assimilated are not incorporated into the forecast cycle; thus, each analysis and DA forecast are independent of past local surface observations. This daily cold start is good for independent tests of the robustness of the DA module. Verification results for November and December 2004 forecasts are given in section 6c. Due to a hardware failure of the supercomputer in December 2004, the sample size for December is smaller than for November.

b. Analysis results

Figure 15 shows all 283 surface stations within the Canadian portion of the analysis and DA domain. The number of observations varied from day to day depending on the number of stations actually reporting. Because verification was performed 1 day later than analysis (and some stations report late), the number of available reports for verification in November 2004 was larger than that for analysis (Fig. 16). December was similar (not shown).

Five analysis methods are compared: GAUSS, TERR_DIFF, MD, MD_LSMG, and MD_MT. For method MD_LSMG, which included shoreline refinement (see DS05), 18 out of 283 surface weather stations are identified to be stations over water, based on the land–sea mask input data for the MC2 model at 2-km grid spacing.

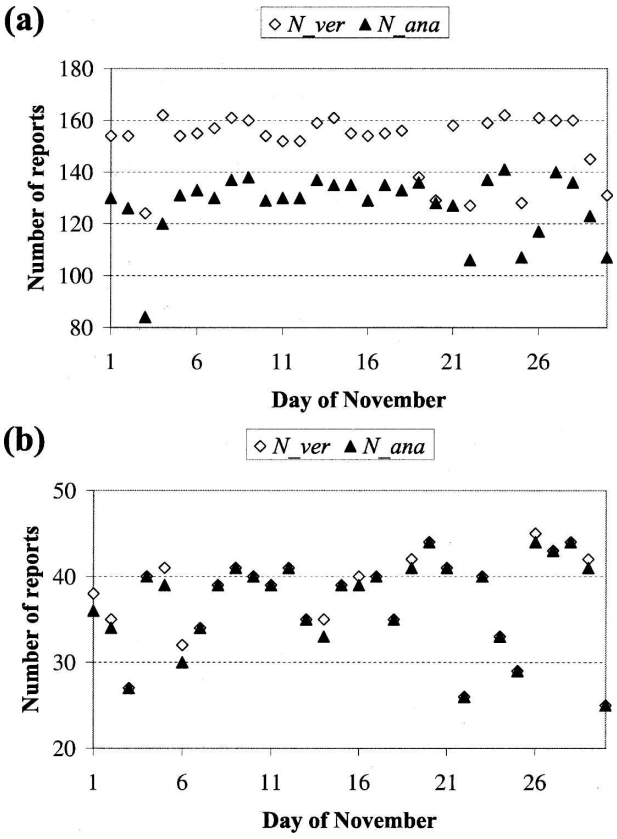

FIG. 16. Number of stations that had reports of (a) potential temperature and (b) specific humidity for each day of November 2004. Each day in the *x* axis corresponds to "day 1" in Fig. 14. Hence, the number of available reports are actually for a date that is 1 day later than the date shown in *x* axis. Those indicated by triangles are the numbers of reports used in analyses; whereas those marked by diamonds are the numbers of reports used in verification.

1) POTENTIAL TEMPERATURE ANALYSES

Table 10 summarizes the verification statistics for November. The analyses from all methods have smaller positive biases than the first guess (FSTG). The MAE and NRMSE of each method are largely decreased from those of the first guess. Even though TERR_DIFF produces smaller bias than all other methods except MD_LSMG, it has the largest MAE and NRMSE. The MD approaches (MD, MD_MT, and MD_LSMG) have smaller MAE and NRMSE than GAUSS and TERR_DIFF, similar to the July 2003 case. The MD_MT approach shows a minor increase of MAE and NRMSE. This is different from the July 2003 case. Comparatively, MD_LSMG has the smallest bias, MAE, and NRMSE, which is similar to what was found for the February 2003 case in DS05. But the improvement of MD_LSMG over MD is smaller, compared to the February 2003 case. The verification statistics for December (not shown) show similar results to November, but with the first guess having small negative bias and all analysis methods having smaller positive bias.

TABLE 10. Verification of analyzed potential temperatures in November 2004. For all these statistics, smaller magnitude is better.

Method	*N*	Bias (K)	MAE (K)	NRMSE
FSTG	4537	0.293	2.238	1.0
GAUSS	4537	0.286	1.158	0.611
TERR_DIFF	4537	0.261	1.278	0.644
MD	4537	0.284	1.024	0.580
MD_MT	4537	0.284	1.031	0.582
MD_LSMG	4537	0.219	0.968	0.563

2) SPECIFIC HUMIDITY ANALYSES

The bias, MAE, and NRMSE of the specific humidity analyses and first guess for November are presented in Table 11. Each method gives largely reduced magnitude of bias, MAE, and NRMSE compared to the first guess. The MD approaches (MD, MD_MT, and MD_LSMG) produce better analyses with the smaller negative bias magnitude, MAE, and NRMSE than GAUSS and TERR_DIFF, similar to the July 2003 case (Table 4). As there are no specific humidity observations from the few mountain-top stations, MD_MT is equivalent to MD. MD_LSMG gives the smallest MAE and NRMSE of the specific humidity analyses. December results are similar, but the magnitudes of bias, MAE, and NRMSE of FSTG are smaller than in November. MD_LSMG slightly underperforms MD in December.

c. Forecast results for MD_MT

1) VERIFICATION OVER THE FULL FORECAST LENGTH

First, look at the verification summary for different near-surface parameters over the full forecast length from the initial 0–15-h forecast period, over all surface

TABLE 11. Same as in Table 10, but for verification of analyzed specific humidity in Nov 2004.

Method	*N*	Bias (1.0×10^{-4} kg kg^{-1})	MAE ($1.0 - 10^{-4}$ kg kg^{-1})	NRMSE
FSTG	1125	−5.987	7.800	1.0
GAUSS	1125	−0.451	1.892	0.292
TERR_DIFF	1125	−0.418	2.443	0.356
MD	1125	−0.169	1.687	0.270
MD_MT	1125	−0.169	1.687	0.270
MD_LSMG	1125	−0.464	1.553	0.247

TABLE 12. Verification statistics of near-surface parameters produced by the MC2 control (CTRL) and MD_MT DA runs. The statistics are calculated for all the observation–forecast pairs over all stations that reported, and over the 15-h forecast period for the month of Nov. The units for each variable are included for bias and MAE (or RMSVE). The NRMSE or NRMSVE is unitless.

		Bias		MAE		NRMSE	
Variable	*N*	CTRL	DA	CTRL	DA	CTRL	DA
θ (K)	55 230	−0.426	−0.524	2.286	2.160	1.000	0.959
RH(%)	31 954	−3.269	−0.718	10.827	9.645	1.000	0.912
q (1.0×10^{-4} kg kg^{-1})	13 290	−5.148	−3.475	7.460	6.390	1.000	0.886
SLP (kPa)	14 172	0.124	0.132	0.224	0.226	1.000	1.004
				RMSVE		NRMSVE	
Variable	*N*			CTRL	DA	CTRL	DA
V (m s^{-1})	43 823			2.696	2.693	1.000	0.999

stations that reported observations, and over the study period of November and December 2004, respectively.

Table 12 summarizes the verification statistics for November. For potential temperature forecasts, the DA run presents slightly larger negative bias than the CTRL run. The magnitude of the biases for both runs is small. The MAE and NRMSE of the DA run are correspondingly smaller than those of the CTRL run. The monthly averaged improvement in potential temperature forecasts, as indicated by the value of NRMSE, is smaller compared to the July 2003 case study (see experiment ATQ in Table 6). A reduction in the improvement of near-surface potential temperature forecasts due to assimilation of surface observations implies that surface thermal forcing in late fall may have been smaller than in the summer.

Humidity forecasts from the DA run have improved bias, MAE, and NRMSE compared to the CTRL run. The improvement in humidity forecasts in November is larger when compared to experiment ATQ for the July 2003 case study. The relatively dry July and wet November help explain the difference in humidity forecast performance.

In contrast to the July 2003 case (see experiment ATQ in Table 6), assimilating temperature and specific humidity does not significantly improve the forecast quality of mean SLP for November.

Assimilating temperature and specific humidity does not disturb the wind fields very much. This finding is different from the July 2003 case study, perhaps because performance is related to surface potential temperature improvement. When the correction to surface potential temperature is large (i.e., July 2003), the initial imbalances between mass and wind fields due to a sudden change in temperature fields are also large and cannot be removed by the IAU technique. However, when the correction to surface potential temperature is small, the initial imbalances are also small and can be suppressed when the AIs are gradually introduced into the model by the IAU technique.

Verification statistics for December (Table 13) show that potential temperature improvement is slightly larger than for November, but is smaller compared to July 2003. The relative humidity forecasts from the DA run are less improved in December than in November. SLP and vector wind forecasts are neither improved nor degraded much in the DA run compared to the CTRL run.

TABLE 13. Same as in Table 12, but for the month of Dec.

		Bias		MAE		NRMSE	
Variable	*N*	CTRL	DA	CTRL	DA	CTRL	DA
θ (K)	37 532	−1.230	−1.227	2.709	2.496	1.000	0.935
RH(%)	21 198	2.685	3.127	10.612	10.272	1.000	0.983
q (1.0×10^{-4} kg kg^{-1})	10 056	−4.053	−3.581	7.436	6.413	1.000	0.888
SLP (kPa)	10 878	−0.055	−0.055	0.336	0.334	1.000	1.002
				RMSVE		NRMSVE	
Variable	*N*			CTRL	DA	CTRL	DA
V (m s^{-1})	29 961			2.935	2.956	1.000	1.007

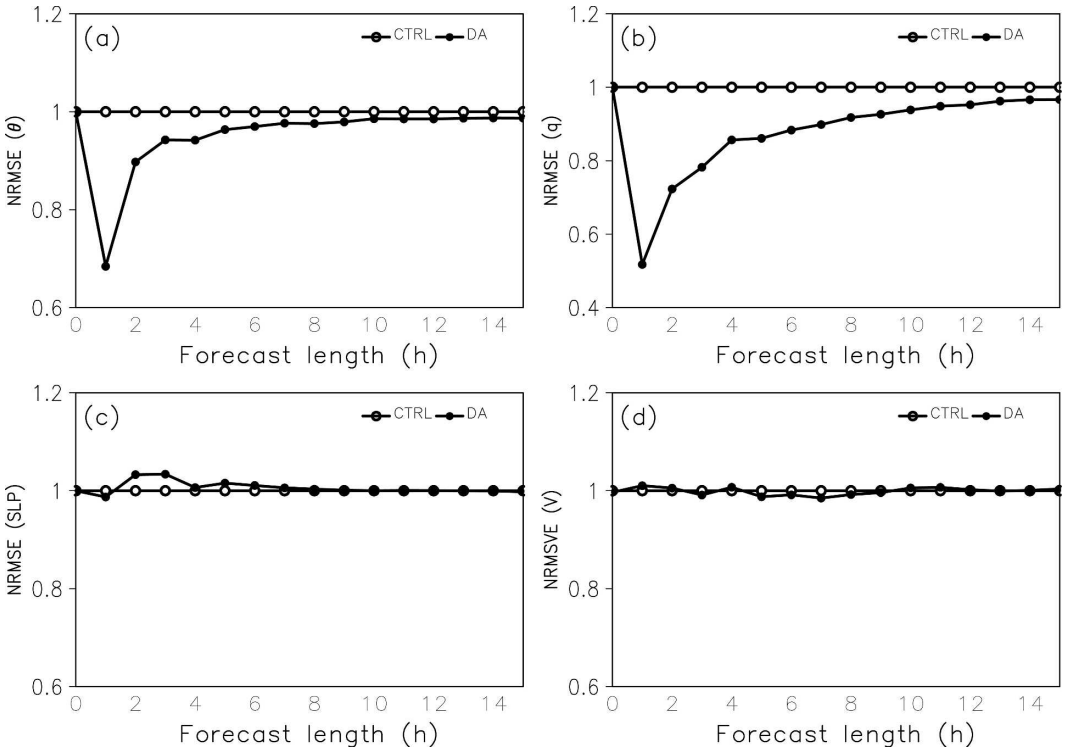

FIG. 17. Time series of monthly averaged NRMSE of surface parameter output from the CTRL and DA runs in November 2004: (a) potential temperature, (b) specific humidity, (c) mean SLP, and (d) vector wind.

2) VERIFICATION BY FORECAST HOUR

Time series of monthly averaged NRMSE of surface potential temperature for November are shown in Fig. 17a. An immediate improvement in NRMSE of the DA run is apparent at 1-h forecast, as new observation information is introduced into the model within the 1-h window. Then, the improvement diminishes with forecast hour, but is still visible at 15 h. December (not shown) shows the same pattern as for November, but with smaller NRMSE. The gradual decrease in the improvement of the DA forecasts can be partly attributed to the effects of lateral boundary conditions advecting into the domain, and partly to model errors that return after assimilating the observations.

The DA improvement of specific humidity in both November (see Fig. 17b) and December (not shown) behaves very similar to that of potential temperature. However, by assimilating temperature and specific humidity only, the model gives slightly degraded SLP and wind forecasts during the first several forecast hours for both November (see Figs. 17c,d) and December (not shown). The deviation of the DA wind forecasts from CTRL for November and December 2004 is much smaller than for July 2003. Near the middle and end of the forecast period, SLP and wind forecasts from the DA run are very similar to those from CTRL, again possibly due to the assimilated information propagating out of the domain, and due to mass and wind fields adjusted toward balance.

7. Summary

The MD method for analyzing complex terrain surface weather observations into a high-resolution NWP model (MC2) is tested with case studies and with 2 months of daily real-time runs. The MC2 first guess, MD surface analysis, and pseudo-upper-air data (interpolated from coarser-resolution analyses) are combined into a final 3D high-resolution analysis using a profile method, which vertically spreads the surface data throughout the BL. The final analysis increments are incorporated into MC2 every 1200 s over a 1-h window. This incremental analysis updating (Bloom et al. 1996) exhibits better skill than instantaneous analysis increment insertion.

Also, a new mountain-top refinement (MD_MT) is proposed to the basic MD approach, and is tested along with a previously suggested shoreline refinement (MD_LSMG). These approaches are also compared with Gaussian (GAUSS) and terrain difference

(TERR_DIFF) methods for surface data analysis in complex terrain. Verification against surface observations is done both for the final analyses, and for the high-resolution forecasts started from MD_MT analyses.

It is found that the MD approaches outperform GAUSS and TERR_DIFF when tested in steep mountainous regions of southern British Columbia, Canada. The mountain-top refinement adds increasing value over MD as more mountain-top stations are included. The MD_LSMG shoreline refinement excels in coastal mountainous regions.

Surface information assimilated at only the lowest model level is soon lost at the beginning of the forecast period. Larger improvement over the control (CTRL) run is achieved when surface information is spread upward throughout the BL. Combining the surface and pseudo-upper-air data gives slightly larger improvement over the CTRL forecast than assimilating only surface data throughout the whole BL. This implies that assimilation of surface data plays an important role in reducing the model errors of near-surface weather parameters, supporting the findings of Hacker and Snyder (2005) who used an ensemble Kalman filter method.

By assimilating surface temperature and specific humidity, the forecast quality of those parameters are better than CTRL. The DA improvement is the largest at 1 h, and gradually decreases out to 15 h. The CTRL run in this paper is initialized by an analysis including older observations than the DA runs. This implies that the DA improvement over CTRL might be slightly overestimated.

The DA run gives poorer forecasts of near-surface winds and precipitation, which are not assimilated into the model. This research addresses just a small part of a much larger data assimilation need. Namely, it attempts to improve only the potential temperature and humidity analyses and forecasts by assimilating these data. No attempt was made to improve precipitation or wind forecasts in the mountains. While there is indeed a critical need to improve forecasts of precipitation and winds, those weather elements are much more difficult to assimilate successfully at high resolution in the mountains, and no attempt is made to do it here.

For these reasons, minimal discussion is given to precipitation. The only reason precipitation is mentioned at all is so that the reader is aware what fields are and are not improved by the assimilation of only potential temperature and humidity.

Future work could be done to assimilate sequences of hourly surface observations in order to achieve better quality for a longer forecast. Future investigations could include the analysis of surface winds (probably through streamfunction and velocity potential) in complex terrain and the incorporation of winds, together with temperature and specific humidity, into an NWP model. To get better precipitation forecasts while assimilating surface data, the assimilation method could be incorporated into a mesoscale three-dimensional variational data assimilation (3DVAR) system, which could also ease using an optimal CTRL.

Acknowledgments. We thank J. Hacker from NCAR for his helpful recommendations and advice about this work, and S. Chamberland from RPN for MC2 user support. We also thank H. Modzelewski, G. Hicks II, and T. Cannon at UBC for their adept computer administration. Thanks also to two anonymous reviewers. Grant support came from the Canadian Natural Science and Engineering Research Council (NSERC), the BC Forest Investment Account, Environment Canada (EC & LUTE), and the Canadian Foundation for Climate and Atmospheric Science (CFCAS). Geophysical Disaster CFD Center computers were funded by the Canadian Foundation for Innovation, the BC Knowledge Development Fund, and UBC.

REFERENCES

Alapaty, K., N. L. Seaman, D. S. Niyogi, and A. F. Hanna, 2001: Assimilating surface data to improve the accuracy of atmospheric boundary layer simulations. *J. Appl. Meteor.,* **40,** 2068–2082.

Barnes, S. L., 1964: A technique for maximizing details in numerical weather map analysis. *J. Appl. Meteor.,* **3,** 396–409.

Benoit, R., M. Desgagne, P. Pellerin, S. Pellerin, Y. Chartier, and S. Desjardins, 1997: The Canadian MC2: A semi-Lagrangian, semi-implicit wideband atmospheric model suited for fine-scale process studies and simulation. *Mon. Wea. Rev.,* **125,** 2382–2415.

Bergeron, G., R. Laprise, D. Caya, A. Robert, M. Giguere, R. Benoit, and Y. Chartier, 1994: Formulation of Mesoscale Compressible Community (MC2) model. Internal Rep., Cooperative Centre for Research in Mesometeorology, 165 pp.

Bloom, S. C., L. L. Takacs, A. M. da Silva, and D. Ledvina, 1996: Data assimilation using incremental analysis updates. *Mon. Wea. Rev.,* **124,** 1256–1271.

Bratseth, A. M., 1986: Statistical interpolation by means of successive corrections. *Tellus,* **38A,** 439–447.

Brewster, K. A., 1996: Application of a Bratseth analysis scheme including Doppler radar data. Preprints, *15th Conf. on Weather Forecasting and Analysis,* Norfolk, VA, Amer. Meteor. Soc., 92–95.

——, 2003: Phase-correcting data assimilation and application to storm-scale numerical weather prediction. Part II: Application to a severe storm outbreak. *Mon. Wea. Rev.,* **131,** 493–507.

Deng, X., and R. Stull, 2005: A mesoscale analysis method for surface potential temperature in mountainous and coastal terrain. *Mon. Wea. Rev.,* **133,** 389–408.

Hacker, J., and C. Snyder, 2005: Ensemble Kalman filter assimi-

lation of fixed screen-height observations in a parameterized PBL. *Mon. Wea. Rev.,* **133,** 3260–3275.

Laprise, R., D. Caya, G. Bergeron, and M. Giguère, 1997: The formulation of the André Robert MC2 (mesoscale compressible community) model. *Numerical Methods in Atmospheric and Oceanic Modelling: The André J. Robert Memorial Volume,* C. A. Lin, R. Laprise, and H. Ritchie, Eds., Atmosphere–Ocean Series, Canadian Meteorological and Oceanographic Society/NRC Research Press, 195–220.

Miller, P. A., and S. G. Benjamin, 1992: A system for the hourly assimilation of surface observations in mountainous and flat terrain. *Mon. Wea. Rev.,* **120,** 2342–2359.

Ruggiero, F. H., K. D. Sashegyi, R. V. Madala, and S. Raman, 1996: The use of surface observations in four-dimensional data assimilation using a mesoscale model. *Mon. Wea. Rev.,* **124,** 1018–1033.

——, G. D. Modica, and A. E. Lipton, 2000: Assimilation of satellite imager data and surface observations to improve analysis of circulations forced by cloud shading contrasts. *Mon. Wea. Rev.,* **128,** 434–448.

Stauffer, D. R., N. L. Seaman, and F. S. Binkowski, 1991: Use of four-dimensional data assimilation in a limited-area mesoscale model. Part II: Effects of data assimilation within the planetary boundary layer. *Mon. Wea. Rev.,* **119,** 734–754.

Stull, R. B., 1988: *An Introduction to Boundary Layer Meteorology.* Kluwer Academic, 666 pp.

——, 2000: *Meteorology for Scientists and Engineers.* 2d ed. Brooks/Cole, 502 pp.

Xue, M., K. K. Droegemeier, and V. Wong, 2000: The Advanced Regional Prediction System (ARPS)—A multi-scale nonhydrostatic atmospheric simulation and prediction model. Part I: Model dynamics and verification. *Meteor. Atmos. Phys.,* **75,** 161–193.

——, K. Brewster, D. Weber, K. W. Thomas, F. Kong, and E. Kemp, 2002: Real-time storm-scale forecast support for IHOP 2002 at CAPS. Preprints, *15th Conf. on Numerical Weather Prediction and 19th Conf. on Weather Analysis and Forecasting,* San Antonio, TX, Amer. Meteor. Soc., CD-ROM, 4B.3.

Yee, S. Y. K., and A. J. Jackson, 1988: Blending of surface and rawinsonde data in mesoscale objective analysis. AFGL Tech. Rep. 88-0144, Air Force Geophysics Laboratory, Hanscom AFB, MA, 31 pp. [NTIS ADA203984.]