9.6 OBJECTIVE ANALYSIS ON TERRAIN-FOLLOWING SURFACES IN COMPLEX TERRAIN

Xingxiu Deng* and Roland Stull
University of British Columbia,
Vancouver, BC, Canada

Abstract

A technique is developed to anisotropically spread surface observations in steep valleys, to create an objective analysis for the lowest, terrain-following mesoscale model levels. The goal is to improve the mesoscale details of lowest level analysis in complex terrain, with an ultimate goal to provide better input-fields for modern numerical weather prediction models (NWP) that can resolve the weather in individual valleys.

The method is a mother-daughter approach where the spreading of information from generation to generation is reduced by the relative and absolute terrain height differences between offspring. This allows information to follow valleys around ridges, while reducing spread over the ridge top.

1. Introduction

Objective analysis transforms information from randomly spaced observing sites into data at regularly spaced grid-points. In the mountainous terrain of western North America, surface observation sites are usually located in the valleys. Thus, the analysis of surface data is made complicated by orographic features in complex terrain.

Any kind of analysis scheme, such as optimum interpolation (OI), successive corrections (Bratseth 1986), or 3D variational analysis (Laroche and Gauthier 1999), requires the specification of a covariance function to model the spatial correlations of background errors. In order to make the analysis scheme feasible, some crude simplifications on the form of these forecast errors are usually done. In particular, isotropy and homogeneity of the background error correlations are most often assumed.

Isotropy, however, is a particularly questionable assumption for variables that are affected by mesoscale phenomena or for analyses in mountainous regions. For this reason, flow-dependent or anisotropic covariance models are receiving more and more attention.

We develop a mother-daughter approach to account for terrain effects in the spread and analysis of surface data. The result is then incorporated into an existing analysis tool (the ARPS data assimilation system - ADAS).

**Corresponding author address:* Xingxiu Deng, Dept. of Earth & Ocean Sciences, University of British Columbia, 6339 Stores Rd. Vancouver, BC, Canada V6T 1Z4. E-mail:xdeng@eos.ubc.ca

2. Mother-Daughter approach for spreading data

Consider a single, isolated surface-weather station in a valley having twists and turns. Given typical boundary-layer processes (Stull 1988), one would expect that the weather observations at this site are fairly representative of the near-surface air further up- and down-stream in this valley in the vicinity of the surface station, regardless of the turns of the valley floor. It is unlikely that this surface observation is representative of near-surface air in a neighboring valley on the other side of a ridge, unless these two valleys are connected. Given modern mesoscale models with fine horizontal resolution, there are often many grid points that are within each valley that need to be properly initialized. So a technique is needed to spread the surface observation to the grid points in a way that favours along-valley spread and suppresses cross-ridge spread. It is assumed that horizontal grid spacing is fine enough to resolve all important ridges and hills.

For simplicity, first consider a surface observation that happens to lie directly on a model grid point. In the mother-daughter approach designed here, the surface observation is treated as an Honored Ancestor (HA), who has eight daughters, one at each of her neighboring grid points. Then every daughter becomes a mother and has her own daughters around her, and so on. The amount of information that a daughter gains from her mother is reduced by terrain height difference between them. The efficiency of transferring information from a mother to a neighboring daughter is called the "sharing factor", which has values between 0 and 1. This approach can be expressed by following iterative equation for sharing factor S:

$$S_{daughter}(\nu) = S_{mother}(\nu - 1) \left\{ 1 - \left[\frac{|Z_{mother}(\nu-1) - Z_{daughter}|}{zref1} \right]^a \right\} \left\{ 1 - \left[\frac{|Z_{HA} - Z_{daughter}|}{zref2} \right]^b \right\} \quad (1)$$

where ν is iteration count, $Z_{mother}(Z_{daughter})$ is terrain elevation at a mother (daughter) grid point, Z_{HA} is terrain elevation at the HA, a, b, $zref1$, $zref2$ are tunable parameters. The equation determines how much the daughters share the Honored Ancestor's information, where the equation is initialized with $S_{mother}(0) = 1$ and $Z_{mother}(0) = Z_{HA}$ (representing the Honored Ancestor). After each application of this equation, the daughter becomes a mother, and her sharing factor (terrain elevation) is used as S_{mother} (Z_{mother}) in the next iteration. The circuitous traveling distance from the Honored Ancestor to every grid-point plays an role on determining the final sharing factors at every grid-point via a Gaussian drop-off (see section 5.).

As the population grows and the progeny spread further and further from the HA, it is often the case that many different generations of relatives try to live in the same house (i.e. at the same grid point). At each grid point, only the strongest (i.e. most informed) daughter with biggest sharing factor survives, similar to the evolution theory of Darwin.

Figure 1 illustrates the sharing factors and traveling distance from one HA.

3. Observations and first-guess fields

To test the above approach, we need as many surface observations as possible. An example will be shown for valleys in Southwestern British Columbia (BC) Canada. Emergency Weather Net Canada, operated and maintained by the Geophysical Disaster Computational Fluid Dynamics Centre at UBC, exists to provide timely and comprehensive surface observations. Hourly data sets are combined from several agencies, including the BC Ministry of Transportation and Highways, BC Ministry of Forests, BC Ministry of Water Land and Air Protection, Greater Vancouver Regional District, Environment Canada, BC Hydro, CN rail and the National Centers for Environmental Prediction.

The error statistics associated with the data sources are unknown.. To simplify problem and avoid complexity added by observations at the first stage, "fraternal-twin" experiments are used to evaluate impact of mother-daughter approach on surface data analysis. Namely,

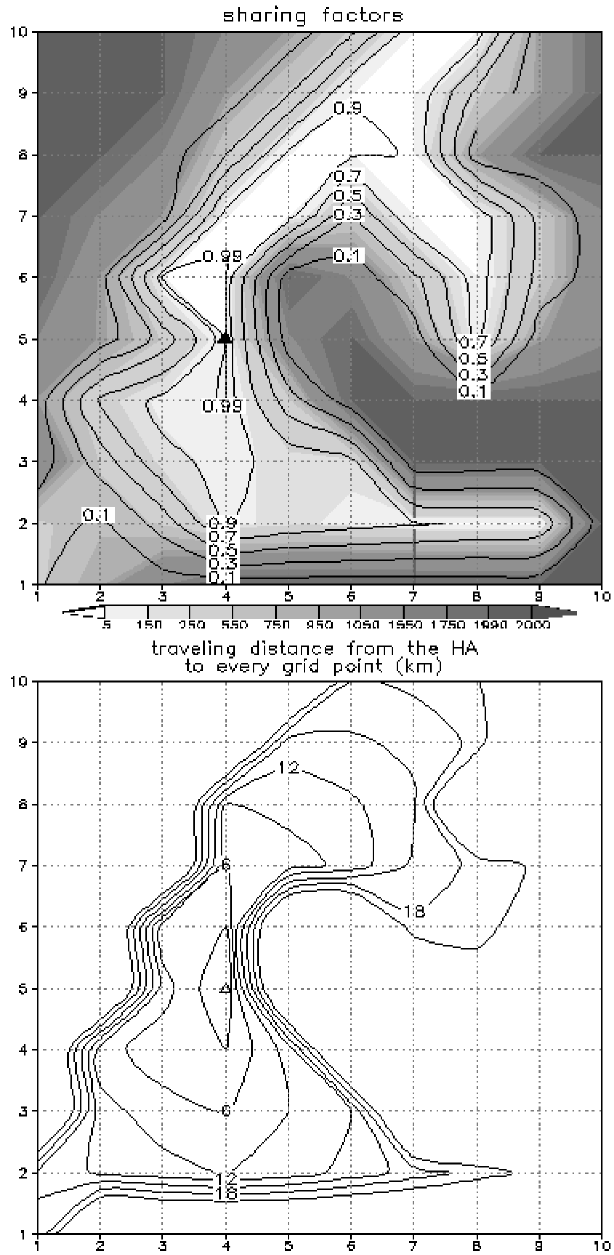

Figure 1: Sharing factors (top, after a Gaussian drop-off with standard deviation of 30 km) and travelling distance (bottom) for an idealized valley within a 2_D domain that spans 300 km x 300 km with a resolution of 3 km. Terrain elevations in meters are shown by shading in the top picture. Darker shading indicates higher elevations, with a maximum terrain height difference of 2000 m. The HA is indicated by a triangle sign ($\triangle$) at $x = 4$ and $y = 5$. The parameters used are: $a = 2$, $b = 2$, $zref1 = 750$ m, $zref2 = 750$ m.

the NWP model (hereafter the "truth model") is first integrated 12 hours to generate a reference atmosphere, which we call the "truth" here. Simulated surface observations are then extracted from the "truth" atmosphere and horizontally interpolated to virtual observation locations. All simulated surface observations are assumed to be perfect, with zero observation error.

Then we run a different NWP model (hereafter called the "analysis model") for 12 hours with the same initial fields as for the "truth" run, but randomly perturbed to generate a different first-guess. When the truth model and the analysis model are similar but not identical, the experiment is called fraternal-twin. As pointed out by Arnold and Dey (1986), the truth model and the analysis model should have the same relationship to each other as the real atmosphere and the model do. As we all know, a NWP model is always less sophisticated than the real atmosphere. Therefore, the analysis model should be less sophisticated than the truth model. This can be easily achieved by using a coarser grid and less complicated physics parameterizations in the analysis model.

In this study, the Mesoscale Compressible Community (MC2) model (Benoit et al. 1997) at 2 km horizontal grid spacing is used as the truth model and MC2 at 3 km as the analysis model. The initial fields and boundary conditions for both the 2 km and 3 km runs come from the output of an MC2 4 km run (figure 2). Perturbations within the range (-pert0, +pert0), with zero average inside the domain, are added to prognostic variables in the initial fields for 3 km runs. Amplitudes (pert0) of the added perturbations for temperature (u-wind component, v-wind component, logarithm of pressure perturbation, specific humidity) are 5.0 °C (5.0 m s^{-1}, 5.0 m s^{-1}, 5.0E-04, 5.0E-03 kg/kg). Finally, an offset of -1.5 °C (2.0 m s^{-1}, -2.0 m s^{-1}, -1.0E-04,1.0E-03 kg/kg) is added to the above perturbed initial fields with an outcome of non-zero average inside the domain. A dynamic initialization algorithm that is built into the MC2 model is activated to remove spurious gravity waves excited by the added perturbations.

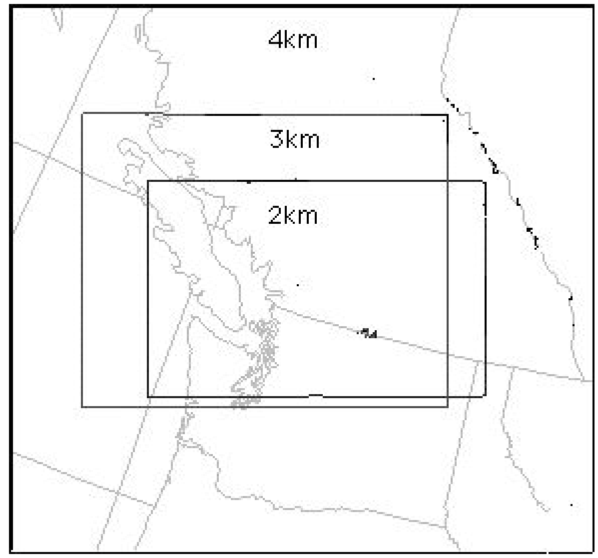

Figure 2: MC2 grid domains for 4 km, 3 km, 2 km. Output from 4 km provides nesting files to initialize 2 km and 3 km runs. 2 km is run to provide a reference atmosphere to generate virtual surface observations for analysis and verification. 3 km is run to provide first-guess fields for analysis.

4. A brief description about ADAS

The Advanced Regional Prediction System (ARPS; Xue et al. 2000) Data Assimilation System (ADAS; CAPS 1995) was developed by the Center for Analysis and Prediction of Storms at the University of Oklahoma. The ADAS is a 3-dimensional mesoscale analysis system that ingests and analyzes meteorological data coming from different observational sources, i.e. single-level observations (typically surface observations) and multiple-level observations such as upper-air soundings.

The ADAS combines observations with a background (first-guess) field by using the Bratseth method, which converges to optimum interpolation. The Bratseth scheme accounts for the relative error between the observations and the first-guess. In ADAS, the spatial correlations of background error are modeled as Gaussian.

Changes are made to allow ADAS to directly use first-guess fields from MC2 output, e.g., horizontal and vertical grid definitions are modified to match those in MC2 model; and x and y coordinates of the model grid are modified from 1D to 2D. The results (sharing factors and travelling distance) from mother-daughter approach, described in section 2., are incorporated into ADAS.

5. Experiment design

Three assimilation experiments (see table 1) are designed. Experiment GAUSS accounts for terrain effects by introducing an additional Gaussian function of terrain height difference between the grid point and the observation location. The algorithm for experiment TERR_DIFF is added to the analysis tool following the method proposed by Miller and Benjamin (1992).

Experiment MD employs the sharing factors from mother-daughter approach (see section 2.) to account for terrain effects on analysis increments. This method tracks the analysis increments (anomalies) along the valleys. Experiment MD further defers from experiment GAUSS and TERR_DIFF in that the circuitous traveling

Table 1: Experiments performed. The R_h and R_z are correlation length scales in the horizontal and vertical, respectively, which define the regions of influence. d_{ij} and Δz_{ij} are direct horizontal straight-line distance and terrain-height difference between an analysis grid point and the observation station, respectively. s_{ij} is the traveling distance from the observation to an analysis grid point determined in the mother-daughter program (i.e., usually following the valley). $S_{daughter}$ is the sharing factor at an analysis grid point, which represents how much the analysis grid point shares the observation information. It is where terrain vs. observation height differences are included, as previously described. In this study, $R_h = 90$ km for the first and second iteration, $R_h = 60$ km for the third iteration, $R_h = 30$ km for the fourth and fifth iteration, during the ADAS assimilation cycle.

Experiment	Correlation function
GAUSS	$\exp(-\frac{1}{2}\frac{\lvert d_{ij}\rvert^2}{R_h^2}) \cdot \exp(-\frac{1}{2}\frac{\lvert \Delta z_{ij}\rvert^2}{R_h^2})$
TERR_DIFF	$\exp(-\frac{1}{2}\frac{\lvert d_{ij}\rvert^2}{R_h^2})/(1+K_z \cdot \lvert \Delta z_{ij}\rvert^2)$
MD	$\exp(-\frac{1}{2}\frac{s_{ij}^2}{R_h^2}) \cdot S_{daughter}$

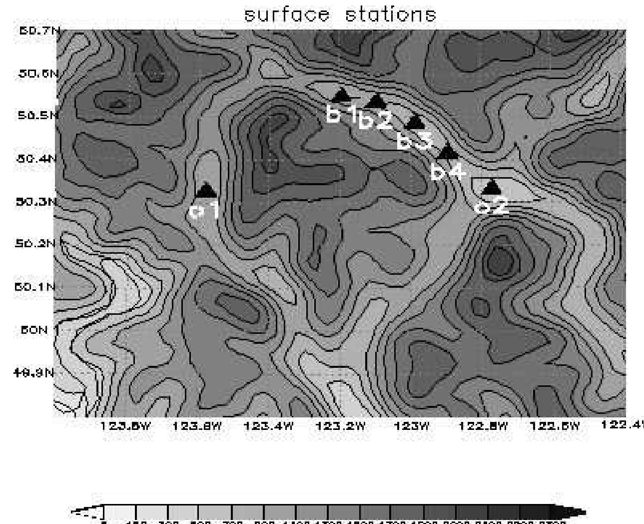

Figure 3: Virtual surface observation stations, indicated by a closed triangle sign (▲), used in the analysis and verification. Station o1 (50.326 °N, -123.578 °W) is near the mouth of the Elaho River. Station o2 (50.3344 °N, -122.767 °W) is near Pemberton. Terrain elevations are from the truth model in meters. Darker shading indicates higher elevations, with a maximum elevation difference of 2054.6557 m in this figure.

distance, rather than conventional horizontal distance, is used in the horizontal part of the correlation function.

6. Results

A case-study comparison of the mother-daughter approach with the other two schemes is presented here for the analysis of potential temperature in mountainous British Columbia. The fraternal-twin experiment is performed for the 7-8 March 2003 case, characterized by an Arctic air outbreak. This case was selected to determine if episodes of cold-air pooling in valleys could be properly analyzed.

The MC2 runs consist of five self-nested grids with a resolution of 108 km, 36 km, 12 km, 4 km and 2 or 3 km. The Eta model analysis and forecasts from National Centers for Environmental Prediction (NCEP), valid at 00 UTC 7 March 2003, are used as initial and boundary conditions for the coarsest grid. The truth model at 2 km is started at 12UTC 7 March from output of MC2 4 km run, and integrated 12 hours forward to generate the "truth" atmosphere, from which simulated surface observations are extracted. The analysis model at 3 km is also started at 12UTC 7 March but from randomly perturbed output of MC2 4 km run, and integrated 12 hours forward to provide a first-guess. The analysis is performed at the lowest terrain-following surface of the analysis model at 00UTC 8 March 2003.

Figure 3 shows six virtual surface stations. Among them, station o1 (elevation: 920 m) and o2 (elevation: 536 m), located in different valleys, are used for analysis, whereas four stations (b1, b2, b3, and b4), located in the same valley as station o2, are used for verification. Elevations for station b1, b2, b3 and b4 are 894 m, 881 m, 765 m, and 830 m, respectively. Hereafter, we will identify the valley where station o1 is located as the Elaho River Valley, and the valley where station o2 and the other four stations are located as the Lillooet River Valley.

6.1 Impact of a single surface observation

Objective analysis involves incorporating weather observations available at irregularly distributed locations onto a uniform grid, and merging the results with a background (first-guess) from a numerical model forecast. Usually, the observations are not analyzed directly. Rather, a first-guess from a NWP model forecast is subtracted from each observation to produce observation increments. These observation increments are then analyzed to produce analysis increments (analysis deviations), which are then added onto the first-guess to produce the final analysis.

The analysis increments produced by a single observation reveal the characteristics of the analysis scheme. Therefore, by examining analysis increments from station o1 (in figure 3), one can conveniently evaluate how information from surface observations will affect grid points around it.

Figure 4 shows response to a single potential temperature observation at station o1 (in figure 3) for experiment GAUSS, TERR_DIFF and MD. All of them are anisotropic. The observation mostly influences grid points in the valleys, rather than those over mountains.

Experiment GAUSS and TERR_DIFF exhibit similar features. Namely, large analysis increments can be seen in valleys that are disconnected from the valley where the observation station is located. The Lillooet River Valley where station b1 is located (see figure 3) is such an example. These two schemes show their drawbacks when an analysis grid point is located at the opposite side (i.e. the Lillooet River Valley in this case) of one mountain to the observation location. But in experiment MD, the observation increments largely modify the first-guess fields for those grid points in the Elaho River Valley while only slightly modifying those in the Lillooet River Valley. This is a realistic attribute, particularly for cold-air pooling. The characteristics for the three schemes are expected from their corresponding correlation functions presented in table 1.

6.2 Impact of two surface observations

A difficult problem in complex terrain is the analysis at one grid point using observations from two adjacent but disconnected valleys, where each valley may contain a different meteorological regime. To further examine the behaviors of the three schemes, we consider two observations that are located at different valleys (station o1 and o2 in figure 3). The observation at site o1 is 3.35 °K colder than the first-guess, while the observation at site o2 is 3.36 °K warmer than the first-guess. Four surface stations (station b1, b2, b3, and b4) in the Lillooet River Valley are used to verify analysis results.

First, only the observation at site o2 is used in the analysis. To evaluate the three schemes, the analysis results are interpolated to the four verifying stations in the Lillooet River Valley. The root mean square error (RMSE) between the observations and analyses at those sites are calculated. Tunable parameters for experiment GAUSS, TERR_DIFF and MD are chosen so that the three schemes produce similar analysis results in the sense of RMSE. R_z is taken as 500 m in experiment GAUSS. K_z is taken as 7.0 x 10^{-6} m^{-2} in experiment TERR_DIFF. In experiment MD, parameters a, b, $zref1$, $zref2$ are 2, 2, 750 m, 750 m, respectively.

The RMSE for experiment GAUSS, TERR_DIFF, and

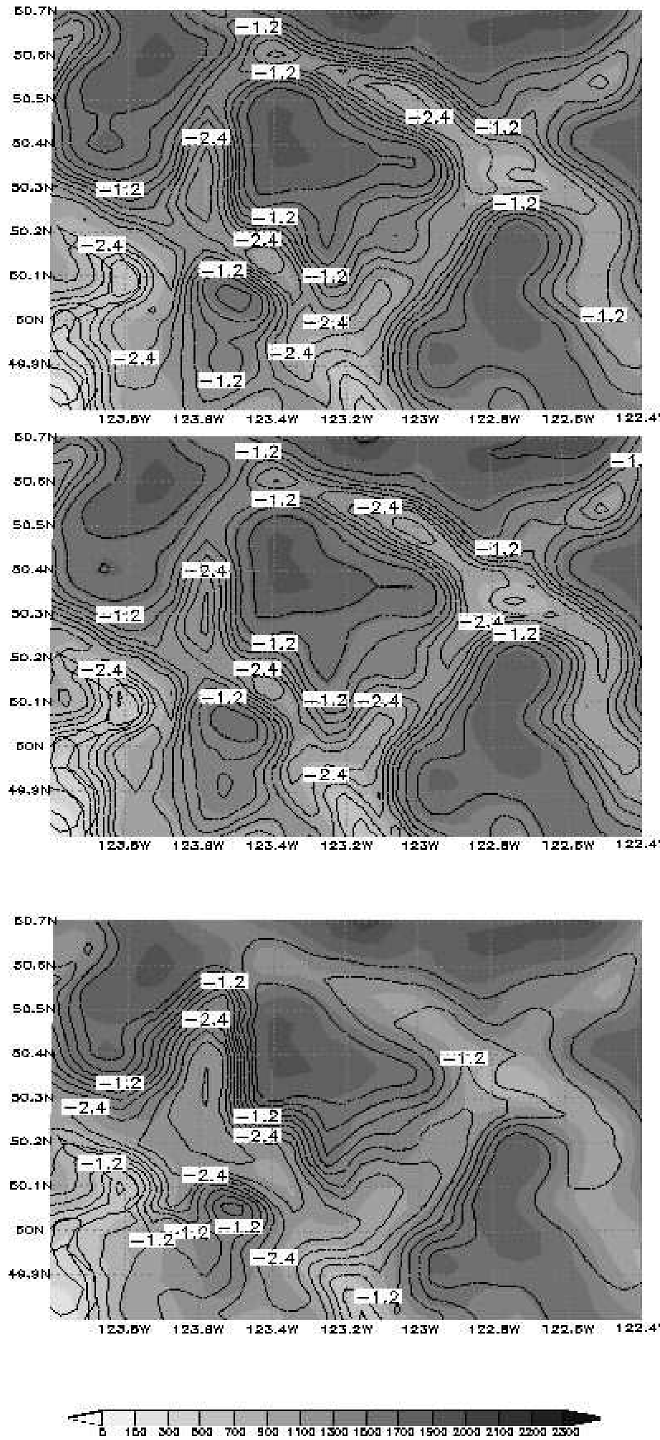

Figure 4: Analysis increments for potential temperature at the lowest model level in response to a single surface potential temperature observation at station o1. Darker shading indicates higher elevations, with a maximum elevation difference of 1939.85 m, which is less than that in figure 3 because of the smoothed terrain used in the NWP model. (top) GAUSS, (middle) TERR_DIFF, (bottom) MD.

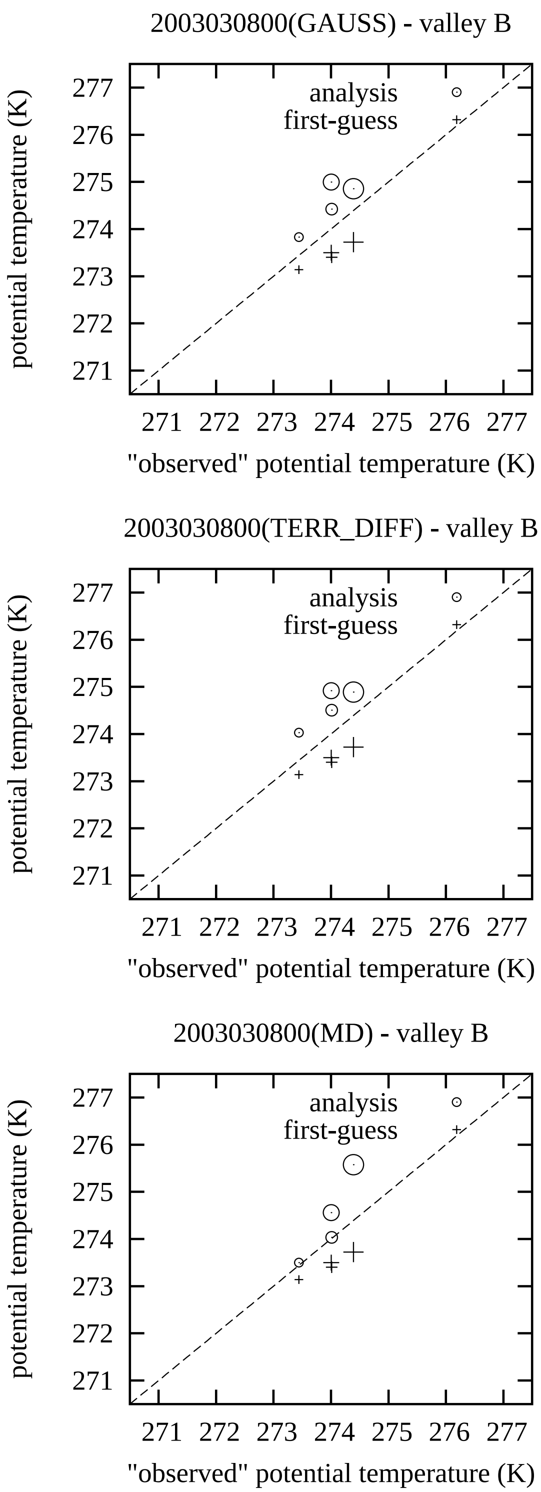

Figure 5: Analysis and first-guess vs. observation, when observation at station o2 is used in the analysis. The analysis is indicated by a circle sign (⊙), whereas the first-guess is indicated by a plus sign (+). Perfect analysis is along the diagonal. (top) GAUSS; (middle) TERR_DIFF; (bottom) MD.

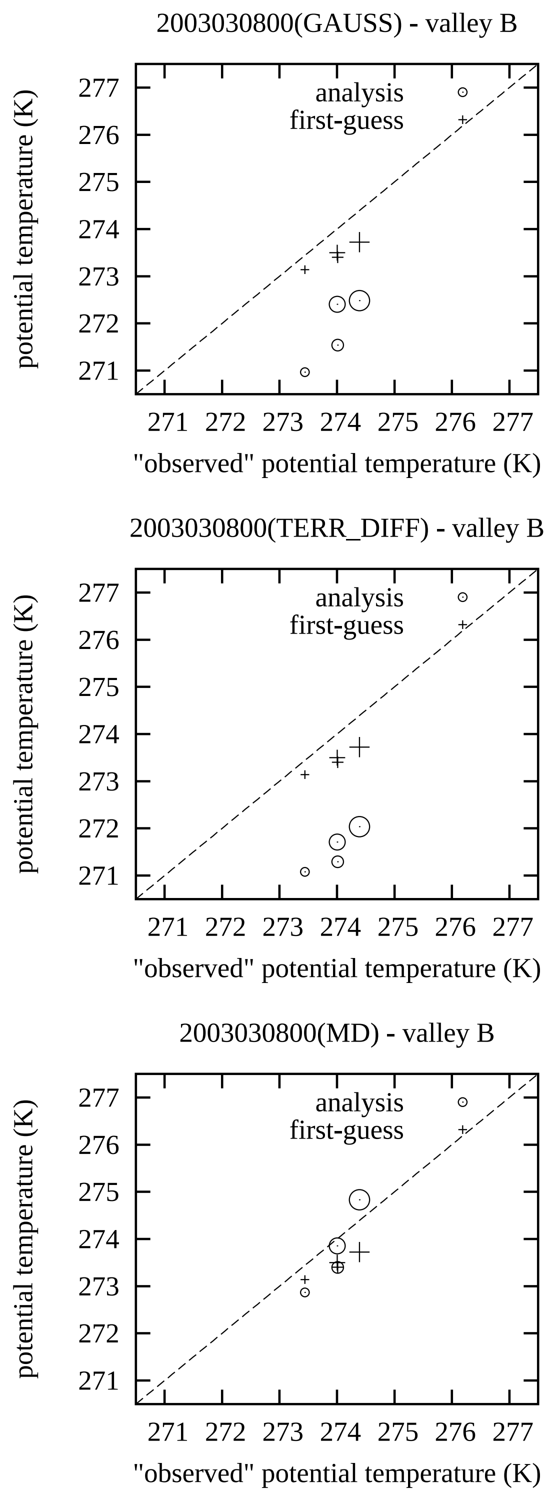

Figure 6: Analysis and first-guess vs. observation, when observations at station o1 and o2 are used in the analysis. The analysis is indicated by a circle sign (⊙), whereas the first-guess is indicated by a plus sign (+). Perfect analysis is along the diagonal. (top) GAUSS; (middle) TERR_DIFF; (bottom) MD.

MD are 0.6179 °K, 0.6468 °K, 0.6559 °K, respectively. The RMSE for the three schemes is slightly greater than the RMSE between observations and the first-guess, which is 0.5395 °K. Analyses and first-guess versus observations for the three schemes are shown in figure 5. It is obvious that they produce similar analysis results for the four verifying stations in the Lillooet River Valley.

Then, observations from both sites (o1 and o2 in figure 3) are used in the analysis. The same set of observations at site b1, b2, b3 and b4 are used again for verification. As shown in figure 6, the added observation at site o1 greatly deteriorates the analysis results from experiment GAUSS and TERR_DIFF. The RMSE increases to 2.1464 °K for GAUSS and 2.4380 °K for TERR_DIFF. However, the added observation at site o1 has minor influence on the analysis of the four verifying stations from experiment MD, as desired because the two valleys with observations are not strongly coupled. The slight influence from the added observation reduces RMSE to 0.4807 °K.

7. Summary and conclusions

A mother-daughter approach is described to tackle the problem of surface data analysis on terrain-following surfaces in complex terrain. The results (sharing factors and travelling distances) of mother-daughter approach are then incorporated in the existing analysis tool (the ARPS data assimilation system - ADAS).

For the one case examined here, the new scheme is compared with original scheme in ADAS, and the scheme developed by Miller and Benjamin (1992). The utility of this approach is found in that the mother-daughter approach can account for both terrain-height difference and valley differences in the analysis of observation increments over mountainous regions. Therefore, this newly developed scheme is preferred to other existing schemes in mountainous regions such as British Columbia, although more case studies and real-time tests are needed to refine the method.

Although calculation of the mother-daughter sharing factors is computationally expensive, it need be done only once for each observation location. The resulting map of sharing factors is then saved as a fixed file, and is applied unchanged as a fast-running stencil for each day's analysis.

REFERENCES

Arnold, C. P. and C. H. Dey, 1986: Observing systems simulation experiments: past, present, and future. *Bull. Amer. Meteor. Soc.*, **67**, 687–695.

Benoit, R., M. Desgagne, P. Pellerin, S. Pellerin, and Y. Chartier, 1997: The Canadian mc2: A semi-lagrangian, semi-implicit wide band atmospheric model suited for fine-scale process studies and simulation. *Mon. Wea. Rev.*, **125**, 2382–2415.

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

CAPS: 1995, ARPS version 4.0 user's guide. Supplement 1: 3-D analysis with ARPS data assimilation system: ADAS version 2.3. Technical report, Center for Analysis and Prediction of Storms, University of Oklahoma, 100 E. Boyd, Suite 1110, Norman OK 73701.

Laroche, S. and P. Gauthier, 1999: Implementation of a 3D variational data assimilation system at the Canadian Meteorological Center. Part II: The regional analysis. *Atmos.-Ocean*, **XXXVII**(**3**), 281–307.

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.

Stull, R. B., 1988: *An Introduction to Boundary Layer Meteorology*. Kluwer Academic Publishers, P.O. Box 17, 3300 AA Dordrecht, The Netherlands, 666 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**(**3** – **4**), 161–193.