SENSITIVITY OF SURFACE TEMPERATURE ANALYSES TO SPECIFICATION OF BACKGROUND AND OBSERVATION ERROR COVARIANCES

by

Daniel Paul Tyndall

A thesis submitted to the faculty of The University of Utah in partial fulfillment of the requirements for the degree of

Master of Science

Department of Meteorology

The University of Utah

December 2008
ABSTRACT

A two-dimensional variational analysis for surface temperature is developed for a limited domain (4° latitude × 4° longitude) in order to evaluate approaches to efficiently examine the sensitivity of similar variational analysis systems to the specification of observation and background errors. This work is intended to help facilitate improvements to the operational Real-Time Mesoscale Analysis (RTMA) developed by the National Centers for Environmental Prediction. The local surface analysis (LSA) uses the same projection and terrain (on a roughly 5-km × 5-km grid), background fields derived from 1-h forecasts of the Rapid Update Cycle (RUC) downscaled to that grid, and observation assets used by the RTMA.

The observation error variance as a function of broad network categories and error variance and covariance of the downscaled 1-h RUC background fields are estimated using a sample of over 7 million surface temperature observations in the continental United States collected during the period 8 May – 7 June 2008. The ratio of observation to background error variance is found to be between 2 and 3, which is higher than that used operationally by the RTMA. This ratio is likely even higher in mountainous regions where the representativeness errors attributed to the observations are large. The background errors also tend to remain more strongly correlated over longer horizontal distances than those specified operationally for the RTMA.
Analysis sensitivity to both the ratio of the observation and background error variance and background error decorrelation length scale is examined for a single case (0900 UTC 22 October 2007) using the LSA centered over the Shenandoah Valley of Virginia. The RTMA surface temperature analysis for that case exhibited several unrealistic features in that region as a result of a pronounced surface-based temperature inversion. Sets of 10 data denial experiments in which 10% of the observations are withheld randomly and uniquely from each analysis are used. The analysis error is estimated by the differences between the withheld observations and the corresponding analyses from which the observations are withheld while the analysis sensitivity to the withheld observations is computed from the differences between control analyses and the analyses from which the observations are withheld. For this case it is possible to improve analysis accuracy in multiple ways, i.e., by making the analysis less (more) detailed by broadening (shortening) the decorrelation length scales of the background error covariance in combination with increasing (decreasing) the observation to background error variance ratio. These results, not surprisingly, confirm the need to examine analysis sensitivity over many types of synoptic situations and the difficulty in specifying those parameters a priori.
# TABLE OF CONTENTS

ABSTRACT......................................................................................................................................................... iv

ACKNOWLEDGEMENTS........................................................................................................................................ v

1 INTRODUCTION............................................................................................................................................... 1
   Background ................................................................................................................................................ 1
   Objectives ................................................................................................................................................ 15

2 SURFACE ANALYSIS SYSTEMS....................................................................................................................... 17
   Introduction ................................................................................................................................................ 17
   Background Field ...................................................................................................................................... 17
   Observations ............................................................................................................................................. 19
   Two-Dimensional Variational Theory ....................................................................................................... 25
   Observation and Background Error Covariances ....................................................................................... 30
   Estimation of Observation and Background Error Covariances ............................................................. 33
   Local 2D-Var Surface Analysis .................................................................................................................. 45

3 OBJECTIVE EVALUATION OF SURFACE ANALYSES ................................................................................. 50
   Withholding Observations ......................................................................................................................... 50
   Estimating Analysis Accuracy and Sensitivity ......................................................................................... 54
   Sensitivity to the Specification of Observation and Background Error Covariances .......................... 69

4 SUMMARY AND DISCUSSION.......................................................................................................................... 86
   Summary .................................................................................................................................................... 86
   Discussion and Recommendations ........................................................................................................... 89

REFERENCES ..................................................................................................................................................... 96
ACKNOWLEDGEMENTS

I would like to thank my advisor, John Horel, for his support and guidance throughout this entire project. I would also like to thank my other two committee members, Jim Steenburgh and Zhaoxia Pu, for the valuable comments they provided for this project.

This project was greatly aided by the support of the developers and evaluators of the Real-Time Mesoscale Analysis (RTMA), especially the help of the lead developer of the RTMA, Manuel de Pondeca. In addition to Manuel, I would also like to thank Jim Purser, David Parrish, John Derber, Wanshu-Wu, Geoff DiMego, Geoff Manikin, Brad Colman, Lee Anderson, Kirby Cook, David Novak, Stan Benjamin, David Myrick, Stephen Jascourt, David Radell, and Steven Lazarus.

I would also like to acknowledge the computer support utilized in this study at the National Centers for Environmental Prediction, including permission to use the Haze supercomputer as well as its computer support personnel. Thanks also goes to Dan Trentman, of the College of Mines and Earth Sciences computer support, as well as the support personnel from the Center for High Performance Computing at the University of Utah.

Finally, I would like to thank my friends here in Utah as well as my family back in Illinois. I wouldn’t have gotten this far without them.
This research was supported by National Oceanic and Atmospheric Administration grant NA07NWS4680003 as part of the CSTAR program.
CHAPTER 1

INTRODUCTION

Background

The National Weather Service (NWS) issues gridded forecasts from 3 h to 7 days of sensible surface weather elements that are aggregated into the National Digital Forecast Database (NDFD; Glahn and Ruth 2003). NWS forecasters create and update the forecast grids using output from forecast models or local surface analyses. The need for corresponding national-scale analyses to help provide consistent initialization and verification of the NDFD forecast grids was recognized by many forecasters and, based on this input, the NWS Office of Science and Technology established an Analysis of Record program to address this need (Horel and Colman 2005).

As an initial proof of concept for the Analysis of Record effort, the National Centers for Environment Prediction (NCEP) has led the development of the Real-Time Mesoscale Analysis (RTMA; de Pondeca et al. 2007). The RTMA is a 5-km surface analysis of temperature, wind, moisture, and pressure with additional precipitation and cloud grids provided from other sources. The analysis grids are produced within 45 min of the nominal top-of-the-hour valid time. The RTMA uses a two-dimensional variational (2D-Var) version of NCEP’s Gridpoint Statistical Interpolation (GSI; Purser et al. 2003a,
software to adjust the downscaled Rapid Update Cycle background field (RUC; Benjamin et al. 2007) using surface observations, with satellite-derived winds used over the oceans.

This research began in 2006 to help evaluate the RTMA, with particular attention placed on the characteristics of the analyses in regions of complex terrain. Web-based graphics were developed for many subregions around the country in order to assess the RTMA routinely. This thesis research builds on the examination of a particular case, 0900 UTC 22 October 2007, when a number of notable features were evident in many of the subdomains routinely examined (i.e., wildfires and strong winds in southern California and cold pools in northern Utah and Virginia). In addition to evaluation of the operational and developmental versions of the RTMA available at that time, this case has been examined further by running the RTMA on the Haze NCEP development computer using the RTMA source code revision implemented on 17 January 2008. Figure 1.1 shows an RTMA analysis of 2-m surface temperature for the continental United States using the January version of the RTMA. The complex mix of synoptic, mesoscale, and local scale features clearly evident around the country demonstrates the need for and utility of the RTMA.

To motivate the research conducted in this study, the analysis for the 0900 UTC 22 October 2007 case is examined in greater detail in one of the subdomains of interest, the Shenandoah Valley/Sharandoah National Park region near Washington, D.C. (Figure 1.2a). The Shenandoah Valley is located to the west of the National Park, which encompasses a swath of the Blue Ridge Mountains (with peaks in excess of 1000 m). The NDFD 5-km terrain used by the RTMA for this region is shown in Figure 1.2b. While the
Figure 1.1. RTMA temperature analysis (°C) at 0900 UTC 22 October 2007 over the continental United States domain. The boxed area indicates the Shenandoah Valley subdomain that is analyzed in this study.
Figure 1.2. Topography of the Shenandoah Valley subdomain. a. High resolution terrain plotted.
Figure 1.2. continued. b. NDFD terrain (m) used by RTMA.
major mesoscale terrain features evident in Figure 1.2a are present in the NDFD terrain, many small-scale, yet often important, local terrain features are lost. The 1200 UTC atmospheric sounding from Sterling, VA illustrates a strong surface-based temperature inversion that morning (Figure 1.3). As discussed by Myrick et al. (2005), these types of synoptic situations are often difficult to analyze in mountainous regions since the surface temperature gradient can be very large with strong cold pools typically located in the valleys adjacent to warmer conditions on surrounding slopes. In this particular situation, the downscaling procedure used to transform the background field to the 5-km resolution of the RTMA develops a number of mesoscale southwest-northeast oriented bands across Virginia as shown in Figure 1.4: higher temperature on the east, lower temperature to the east of the Blue Ridge Mountains, higher temperature over the background’s approximation of the Blue Ridge Mountains, and generally lower temperature to the west of that range. As is typical throughout the country, there are far fewer observations (~300) available than the 1,595 gridpoints of this subdomain (or approximately 700 observations for the 6,384 gridpoints of the entire region depicted in Figure 1.4). In addition, the unbalanced distribution of observations across the region with dense coverage in the Washington D.C. metropolitan area and limited coverage elsewhere is common around the country as well. The observations, however, help to provide greater detail in many locations, e.g., lower temperatures along the valley floor of the Shenandoah Valley and higher temperatures on the nearby slopes, which an experienced forecaster would know to be typical of the conditions in other regions of the subdomain as well.
Figure 1.3. Atmospheric sounding from IAD valid 1200 UTC 22 October 2007. The red and green lines depict the temperature profile and dewpoint profile of the atmosphere, respectively.
Figure 1.4. Background field over analysis domain, using a 1-h downscaled RUC forecast valid at 0900 UTC 22 October 2007. Temperature is plotted in °C; elevation is contoured in meters.
The goal of an objective analysis is to blend the observations and the background in such a way that the final product reflects an optimal mix of the two (McPherson 1975; Talagrand 1997). Defining whether an analysis is optimal inherently involves subjective and objective evaluation. The RTMA analysis shown in Figure 1.5a in this instance exhibits a number of features that do not appear optimal subjectively based on physical reasoning for this type of synoptic situation. One would expect that the inclusion of the observations would help to provide better definition of the valley cold pools and higher temperatures on nearby slopes. However, in this instance the RTMA analysis exhibits broad features that do not conform to the analysis terrain, e.g., the bull’s-eye of high temperature over the highest NDFD terrain of the Blue Ridge Mountains. As will be discussed in Chapter 2, the analysis at a particular gridpoint depends most heavily on nearby innovations (differences between the observations and the corresponding background values). As shown in Figure 1.5b, the adjustments to the background made by the RTMA analysis are appropriately constrained in the vicinity of the innovations, i.e., where the innovations are positive (negative) the analysis temperature is higher (lower) than the background. However, in some data voids (e.g., at the NDFD terrain’s peak of the Blue Ridge Mountains), the analysis increments are larger than nearby innovations, which leads, for example, to the physically unrealistic analysis bull’s-eye in Figure 1.5a.

As an illustration of a technique to evaluate the analysis further, Figure 1.6 shows a data denial experiment in which 10% of the observations are withheld from the analysis using the approach that will be discussed in Chapter 3. The differences between the control analysis (Figure 1.5a) and the analysis using 90% of the observations (Figure
Figure 1.5. RTMA temperature analysis and increments over the Shenandoah Valley subdomain valid 0900 UTC 22 October 2007, with 5-km resolution elevation contoured in meters. a. Temperature analysis shaded in °C. Purple numbers are the observed temperature values (°C) used in the analysis.
Figure 1.5. continued. b. Temperature analysis increments (analysis minus background, °C) plotted in shades of orange and purple. Green numbers denote innovation values (observation minus background in °C) used by the analysis.
1.6) are shown in Figure 1.7. Although these differences tend to be small (generally less than 2°C) at any particular gridpoint, the spatial pattern of those differences help highlight the sensitivity to the withheld observations. Most of the areas in Figure 1.7 with the largest differences (dark red or blue colors) are associated with withheld observations. For example, adding the 4.4°C observation at Fort Valley (marked FVRV2 in Figure 1.5a) in the Shenandoah Valley results in a much colder control analysis. The sensitivities evident in Figure 1.7 to the withheld observations tend to be broader than the terrain features and qualitatively not physical for this synoptic situation. These tendencies reflect that the RTMA analysis underfits the observations in some locations. In addition, overfitting to the cold valley observations is evident nearby (e.g., to the southwest of Fort Valley) since higher temperature is evident in the control analysis even though no positive innovation was withheld from the other analysis.

Over and underfitting are frequent concerns in data assimilation when observation density is low (Daley 1991). Underfitting results from the analysis having too few degrees of freedom compared to the underlying signal and causes the analysis to fail to capture observed maxima and minima. Overfitting artificially creates false maxima and minima. These problems may not be obvious to the user unless the background field and the observations are closely compared to the final analysis. Controlling over- and underfitting in a variational analysis requires appropriate specification of the background and observation errors (Daley 1991).
Figure 1.6. RTMA temperature analysis (shaded, °C) valid at 0900 UTC 22 October 2007 in which 10% of the observations are withheld. Elevation is contoured in meters. Purple numbers are the observed temperature values (°C) used in the analysis.
Figure 1.7. Difference between control RTMA analysis and corresponding analysis with 10% of the observations withheld. Differences between analyses are shown by red or blue shading. Observations withheld from the data withheld analysis are shown in green and the locations of observations used by both analyses are denoted by purple dots. Elevation is contoured in meters.
Objectives

The opportunity to run the RTMA on the Haze development computer as demonstrated in the case study described above has helped to foster improvements to the NCEP analysis code during the past year. However, as that research progressed, the need to also test and evaluate 2D-Var analyses using a simple univariate code for local domains rather than the complex RTMA for the nation as a whole became apparent. In addition, work was needed to assess and document the observation error variances of the surface observations and the RUC background error variances and covariances, which affect the characteristics of 2D-Var analyses.

Hence, the objectives of this study are:

- To examine for the nation as a whole the observation error variance, background error variance, and covariance of the downscaled RUC background used by the RTMA.
- To develop procedures to quantify objectively the analysis error of a 2D-Var system in order to assess the degree to which the background field has been improved by assimilating the available observations.

Characteristics of variational data assimilation systems, including the RTMA and the local 2D-Var surface analysis (LSA) developed here will be described in the next chapter. This study is focused on analyses of surface temperature only. Evaluation of observation and background errors based on a month-long data sample are made as well. Data denial procedures to estimate analysis error are developed in Chapter 3 and applied to the domain centered on the Shenandoah Valley domain for the 0900 UTC 22 October 2007 case. Sets of data denial experiments are performed using the LSA to assess the
sensitivity for this particular case and region to the specification of observation error variance and background error covariance. A summary and conclusions follow in Chapter 4. Implications of this research for the RTMA developers as well as operational users of the RTMA are presented.
CHAPTER 2

SURFACE ANALYSIS SYSTEMS

Introduction

The ingredients required to produce a surface analysis using the RTMA or LSA are summarized in this chapter. These ingredients are: a background field, observations, an analysis system, and specification of the error covariances of the observations and background field. The surface temperature analysis for 0900 UTC 22 October 2007 shown in Chapter 1 using the RTMA is contrasted to that generated for the Shenandoah region using the LSA.

Background Field

A surface analysis, such as the RTMA or LSA, is dependent on an operational data assimilation system to supply the background field. The RTMA depends on the 13 km Rapid Update Cycle (RUC; Benjamin et al. 2004) 1-h forecast to specify its background field. Physical consistency in space and time is provided to the RTMA by using the RUC13, which is a three dimensional atmospheric model. The RUC13 background field is downscaled to the 5-km NDFD grid using the methods described by
Benjamin et al. (2007). The downscaling process for temperature is composed of three parts:

1. The RUC temperature fields at all vertical levels are bilinearly interpolated horizontally from the 13-km resolution to the 5-km grid.

2. Temperature grids are vertically interpolated to the height of the RTMA terrain, in a manner depending on one of two conditions:
   - When the NDFD terrain is lower than the RUC13 terrain, the RUC13’s lapse rate from the lowest 25 mb is multiplied by the distance between the two elevations and is added to the 2-m temperature from the RUC13 to yield the new 2-m temperature. In cases in which the low level lapse rate shows an inversion, the RUC13’s 2-m temperature is used as is.
   - When the NDFD terrain is higher than the RUC13 terrain, the downscaling uses RUC13’s temperature at a height of 2-m above the downscaled terrain for the new background temperature.

3. The land-water mask is downscaled by comparing the new 5-km land-water mask and the RUC13’s land-water mask. When values differ at a point, the coastal temperature gradients from the RUC13 are moved to the new land (water) point if there is a continuous path of land (water) between the two points.

Other aspects of the background fields used by the RTMA are discussed by Benjamin et al. (2007) and Jascourt (2007). When the strong surface based inversions are present in low lying areas (such as for the case presented in Chapter 1), it is not surprising
that the background field will be too warm in the valleys, as the RUC’s terrain will generally be higher than that of the NDFD.

Observations

This study relies on the data normally available and used by the RTMA, which includes a number of data streams relied upon by other NCEP data assimilation systems. These include synoptic and aviation observations, surface mesonet observations from a variety of other networks, Coastal-Marine Automated Network (C-MAN) buoy observations, as well as ship observations that are obtained from the Meteorological Assimilation Data Ingest System (MADIS; Miller et al. 2005) at 30 min past the hour. The RTMA’s wind analyses also include satellite-derived winds from the Special Sensor Microwave/Imager (SSM/I) and Quick Scatterometer (QuikSCAT) instruments. All observations must fall within a ±12 min time window centered about the analysis time, except for observations from QuickSCAT and observations from the Remote Automated Weather Stations (RAWS) mesonet, which must fall within a −30/+12 min time window about the analysis time (M. de Pondeca, personal communication). The latter broader time window for RAWS observations reflects the need to allow for the fixed hourly data collection times outside the ±12 min window for many of those stations that are usually located in critical data sparse regions. The operational constraints of the RTMA to provide a timely analysis within 45 min past the hour implies that many mesonet observations for which data transmission times are relatively slow are not used (B. Olsen, personal communication). For example, Snowpack Telemetry (SNOTEL) observations from the National Resources Conservation Service (NRCS) in mountainous regions are
rarely used in the RTMA, since the meteor-burst data collection approach used by that network introduces typical latencies in data transmission greater than 30 min. While the RTMA has extensive gross error and quality checks developed as part of the analysis system, all observations that pass the quality control of MADIS are used in this study.

Figure 2.1 illustrates the locations for which observations are available for the 0900 UTC 22 October 2007 analysis used in this study. The surface observations are subdivided into four categories: METAR (59 observations), primarily aviation and synoptic reports at airport locations; PUBLIC (575 observations), an aggregation of networks including the Automated Weather System (AWS) and the Citizen Weather Observing Program (CWOP); Remote Automated Weather Stations (RAWS; 11 observations), located typically in remote locations for fire weather applications; and OTHER (75 observations), which includes all remaining surface mesonet observations.

The METAR reports have provided the foundation for surface weather observations around the nation, and are incorporated into the RUC data assimilation system. For the RTMA analysis shown in Figure 1.1, 4,938 METAR observations were available to estimate the conditions at 892,375 gridpoints. Obviously, that network by itself provides insufficient coverage to define mesoscale and local scale weather features. As is evident in Fig. 2.1, the PUBLIC source of temperature observations is clearly a major one for the RTMA, especially in metropolitan areas. Over 7,500 PUBLIC observations across the nation were used in the analysis shown in Figure 1.1. Concerns about the quality and representativeness of those observations have been raised, primarily because of the relatively inexpensive sensors, equipment often used, as well as potential siting issues since the equipment is typically mounted on roofs of schools or homes.
Figure 2.1. Observations available for 0900 UTC 22 October 2007, plotted over NDFD terrain (shaded in meters). a. Locations of METAR observations.
Figure 2.1. continued. b. Locations of PUBLIC observations.
Figure 2.1. continued. c. Locations of RAWS observations.
Figure 2.1. continued. d. Locations of OTHER observations.
Although several layers of quality control are utilized by the RTMA to help reduce the impacts from these issues, it has been observed (not shown) that it does not reject all poor observations and it may on occasion incorrectly reject good observations. However, when averaged across the entire CONUS domain over many analyses, the RTMA does effectively reduce the magnitude of the errors associated with these problems, which will be presented later in this chapter. For the subdomain used in this study, RAWS stations are not a major asset (Figure 2.1b). However, nearly 2,200 stations are located nationwide in support of fire weather applications, and are used increasingly for a variety of emergency management applications as well. RAWS observations are particularly valuable throughout the western United States, since they are typically located in remote locations in areas for which no other automated observations are available. The OTHER category (Figure 2.1d) includes a mix of Road Weather Information Systems installed by state transportation departments, air quality agencies, and a variety of additional government, commercial, and educational networks installed for many different types of applications. There are over 1500 such observations typically available to the RTMA nationally to help provide weather information along transportation corridors in a mix of remote and metropolitan areas.

Two-Dimensional Variational Theory

Following Kalnay (2003), one of the simplest forms of an analysis equation can be written as the sum of weighted innovations to the background field,

$$
\bar{x}_a = \bar{x}_b + \bar{W}(\bar{x}_o - \bar{x}_b)
$$

(2.1)
where \( \bar{x}_a \) corresponds to the analysis field, \( \bar{x}_b \) corresponds to the background field, \( \bar{x}_o \) corresponds to the observations transformed to the analysis grid, and \( \bar{W} \) corresponds to an optimal weight that takes into account the probability that the innovation \((x_o - x_b)\) accurately represents the difference between the true state of the atmosphere and the background field.

Variational approaches modify the background field by minimizing the cost function, \( J \), which is a function of the analysis parameter, \( \bar{x}_a \). The cost function originates from the assumed statistical properties of the observations and the background field, i.e., a Gaussian distribution probability \( (P) \) of the form:

\[
P(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}
\]

(2.2)

for an independent variable \( x \), with mean \( \mu \), and a variance \( \sigma^2 \). The Guassian distribution assumption is used to yield the likelihood \( (L) \) of the analysis given a single background (subscript \( b \)) and observation value (subscript \( o \)), respectively:

\[
L_{\sigma_b}(x_a | x_b) = P_{\sigma_b}(x_b | x_a) = \frac{1}{\sigma_b \sqrt{2\pi}} e^{-\frac{(x_b-x_a)^2}{2\sigma_b^2}}
\]

(2.3)

\[
L_{\sigma_o}(x_a | x_o) = P_{\sigma_o}(x_o | x_a) = \frac{1}{\sigma_o \sqrt{2\pi}} e^{-\frac{(x_o-x_a)^2}{2\sigma_o^2}}
\]

(2.4)

The most likely value of the analysis given these two probabilities maximizes the joint probability:
\[
\begin{align*}
\max_{x} L_{\sigma_a,\sigma_b}(x_a|x_o, x_b) &= P_{\sigma_a}(x_o|x_a)P\sigma_b(x_b|x_a) = \frac{1}{2\pi\sigma_a\sigma_b}e^{\frac{(x_o-x_a)^2}{2\sigma_a^2}} \frac{(x_b-x_a)^2}{2\sigma_b^2} \\
\text{(2.5)}
\end{align*}
\]

Taking the natural logarithm of both sides of the equation to simplify yields

\[
\begin{align*}
\max_{x} \ln L_{\sigma_a,\sigma_b}(x_a|x_o, x_b) &= \max_{x} \left[ \ln \frac{1}{2\pi\sigma_a\sigma_b} - \frac{(x_o-x_a)^2}{2\sigma_a^2} - \frac{(x_b-x_a)^2}{2\sigma_b^2} \right] \\
\text{(2.6)}
\end{align*}
\]

Multiplying by negative one, moving the constant from the right hand side of the equation to the left, and assembling the terms in the left side of the equation into a function of \(x_a\) and rearranging yields

\[
2J(x_a) = \frac{(x_a-x_b)^2}{\sigma_b^2} + \frac{(x_a-x_o)^2}{\sigma_o^2} \\
\text{(2.7)}
\]

Finally, applying this cost function to all observations and the entire background field yields:

\[
2J(x_a) = (x_a - x_b)^T P_b^{-1} (x_a - x_b) + (\bar{y}_o - H(x_a))^T P_o^{-1} (\bar{y}_o - H(x_a)^-1) \\
\text{(2.8)}
\]

This equation utilizes the observation values \((\bar{y}_o)\) directly and applies the linear forward operator \(H\) on the analysis values to transform them to the locations of the observations.

The error covariance matrices \(P_b\) and \(P_o\) define how the errors of the background and observations are related from one location to another. The “optimal mix” between the
background field and observations is clearly evident in Equations 2.7 and 2.8; the first term on the right side of both equations corresponds to the penalty associated with differences between the analysis and the background field while the second term corresponds to the penalty associated with differences between the analysis and the observations.

The variational approach attempts to minimize Equation 2.8 by setting \( \nabla_x J(x_a) = 0 \), where the subscript \( x_t \) corresponds to truth. Equation 2.8 is expanded using the relationship:

\[
\overline{y}_o - H(x_a) = y_o - H(x_b + (x_a - x_b)) = \overline{y}_o - H(x_b) - H(x_a) - x_b
\]  

This relationship is substituted in Equation 2.8 to yield:

\[
2J(x_a) = (x_a - x_b)^T P_b^{-1} (x_a - x_b) + (x_a - x_b)^T \bar{H}^T P_o^{-1} \bar{H} (x_a - x_b)
\]

\[
- [\overline{y}_o - \bar{H}(x_b)]^T P_o^{-1} \bar{H} (x_a - x_b)
\]

\[
- (x_a - x_b)^T \bar{H}^T P_o^{-1} [\overline{y}_o - \bar{H}(x_b)]
\]

\[
+ [\overline{y}_o - H(x_b)]^T P_o^{-1} [\overline{y}_o - H(x_b)]
\]  

Rearranging and combining terms in Equation 2.10 yields:
\[ 2J(\bar{x}_a) = (\bar{x}_a - \bar{x}_b)^T P_b^{-1} (\bar{x}_a - \bar{x}_b) + (\bar{x}_a - \bar{x}_b)^T \tilde{H}^T P_o^{-1} \tilde{H} (\bar{x}_a - \bar{x}_b) \]

\[ - \left[ \bar{y}_o - \tilde{H}(\bar{x}_b) \right]^T P_o^{-1} \tilde{H} (\bar{x}_a - \bar{x}_b) \]

\[ - (\bar{x}_a - \bar{x}_b)^T \tilde{H}^T P_o^{-1} \left[ \bar{y}_o - \tilde{H}(\bar{x}_b) \right] \]

\[ + \left[ \bar{y}_o - \tilde{H}(\bar{x}_b) \right]^T P_o^{-1} \left[ \bar{y}_o - \tilde{H}(\bar{x}_b) \right] \]

(2.11)

Because of its large size (i.e., the square of the number of gridpoints), it is not feasible to compute the inverse of the background error covariance $P_b$ for most applications. One approach to avoid calculating the inverse is to substitute

\[ \bar{v} = P_b^{-1} (\bar{x}_a - \bar{x}_b) \]

(2.12)

in the penalty function, which yields

\[ J(\bar{x}_a) = \frac{1}{2} \bar{v}^T P_b^{-1} \bar{v} + \bar{v}^T \tilde{H}^T P_o^{-1} \tilde{H} P_b \bar{v} - \frac{1}{2} \left[ \bar{y}_o - \tilde{H}(\bar{x}_b) \right]^T P_o^{-1} \tilde{H} P_b \bar{v} \]

\[ - \frac{1}{2} \bar{v}^T P_b^{-1} \tilde{H}^T P_o^{-1} \left[ \bar{y}_o - \tilde{H}(\bar{x}_b) \right] \]

\[ + \frac{1}{2} \left[ \bar{y}_o - \tilde{H}(\bar{x}_b) \right]^T P_o^{-1} \left[ \bar{y}_o - \tilde{H}(\bar{x}_b) \right] \]

(2.13)

Minimizing Equation 2.13 produces

\[ \nabla_{\bar{x}} J(\bar{x}_a) = 0 = P_b^T \bar{v} + \tilde{H}^T P_o^{-1} \tilde{H} P_b \bar{v} - \tilde{H}^T P_o^{-1} \bar{y}_o \]

\[ + \left( P_b^T + \tilde{H}^T P_o^{-1} \tilde{H} P_b \right) \bar{v} = \tilde{H}^T P_o^{-1} \left[ \bar{y}_o - \tilde{H}(\bar{x}_b) \right] \]

(2.14)

(2.15)
From Equation 2.15, the term $\tilde{v}$ can be iteratively solved as a function of the innovations and error covariances using a variety of methods, such as the conjugate gradient solution method (Hestenes and Stiefel 1952), the Generalized Minimum Residual method (GMRES; Saad and Schultz 1986), or through the use of recursive filters (Purser et al. 2003a, 2003b). The resulting 2D-Var analysis is then determined from:

$$\bar{x}_a = \bar{x}_b + \bar{P}_b \tilde{v}$$  \hspace{1cm} (2.16)

Equation 2.16 is analogous to Equation 2.1 in that the analysis equation is simply the addition of an increment term ($\bar{P}_b \tilde{v}$ in Equation 2.16 or $\tilde{W}(\bar{x}_o - \bar{x}_b)$ in Equation 2.1) to the background field ($\bar{x}_b$).

### Observation and Background Error Covariances

As shown in Equation 2.15, the analysis depends strongly on the specification of the observation and background error covariances. To illustrate this further, for the case of a single observation and background value (Equation 2.7), the variational analysis would be:

$$T_a = \frac{\sigma_o^2}{\sigma_o^2 + \sigma_b^2} T_b + \frac{\sigma_b^2}{\sigma_o^2 + \sigma_b^2} T_o$$  \hspace{1cm} (2.17)

Hence, if the observations are assumed to have small errors ($\sigma_o < \sigma_b$), then the analysis would be tightly constrained by the observations. If the observation and the background errors are assumed to be equal, then the analysis would simply be the average of the two
values. Hence, the prescribed ratio of the observation and background error variances controls to a large degree the extent to which a variational analysis will resemble the observations or the background.

For the more general case of many observations (Equation 2.15), it is assumed that the observation errors at one location are uncorrelated with those at another, and hence, $\overline{P}_o$ (the observation error covariance) has diagonal elements determined by the magnitude of the observation error variance assigned for that type of observation (i.e., it is possible to specify a larger or smaller error variance for observations as a function of network type, RAWS or PUBLIC, for example). For most variational analysis applications, it is not practical to specify $\overline{P}_b$ (the background error covariance) uniquely for every pair of gridpoint locations since its array size is so large. Rather, assumptions are used which lead to $\overline{P}_b$ being a sparse matrix in which only the background error covariances between pairs of nearby gridpoints are assumed to be related to one another and the diagonal elements are determined by the magnitude of the background error variance assigned for that variable.

The present RTMA relies on the specified estimates of observation and background error variances developed from regional-scale models (Wu et al. 2002) and adjusted subjectively based on experience and sensitivity experiments using the RTMA during the past couple of years (M. de Pondeca, personal communication). The observation/background error variance ratio is assumed to be lower ($\sigma_o^2/\sigma_b^2 = 1.0$) for METAR observations than those for other mesonets (RAWS, PUBLIC, OTHER; $\sigma_o^2/\sigma_b^2 = 1.2$). The background error covariances used by the RTMA are of quasi-Gaussian form and a function of horizontal and vertical separation distance between any
two grid locations (de Pondeca et al. 2007). The covariance shapes used in the RTMA are
efficiently defined with the help of spatial recursive filters used in conjunction with the
sequential line-filtering “triad” algorithm (Purser et al. 2003a, 2003b; Purser 2005). The
structure functions are prescribed to display a controlled degree of correlation with the
underlying smooth representation of the NDFD terrain based on a variant of the
Riishøjgaard method (Riishøjgaard 1998).

During the evaluation of the RTMA completed as part of this research, the
sensitivity of the analysis to the smoothing used to define the background error
covariances was examined and was found to have negative impacts on the analysis as
shown in Chapter 1. The most recent implementation of the developmental version of the
RTMA relies on reduced terrain smoothing for the structure functions as a result of this
study’s early results.

Although the specification of the RTMA’s background error covariance structure
functions can be generalized to any field defined by the underlying surface or background
field, the present implementation essentially approximates the following:

\[
\rho_{ij} = \sigma_{E}^2 \exp\left( -\frac{r_{ij}^2}{R^2} \right) \exp\left( -\frac{z_{ij}^2}{Z^2} \right) 
\]  

where \( \rho_{ij} \) is the background error correlation between each gridpoint and observation, \( r_{ij} \)
and \( z_{ij} \) are the horizontal and vertical distances between each gridpoint and observation
respectively, and \( R \) and \( Z \) are horizontal and vertical scaling factors that determine the
decorrelation length scale. The first exponential defines an isotropic dependence on
separation distance, with the controlling scale \( R \) set presently to roughly 40 km in the
RTMA. The covariance structure functions used by the RTMA can be interpreted as an anisotropic adjustment such that the background errors decorrelate faster with horizontal distance if the two gridpoints differ in elevation. The vertical scale height $Z$ used in the RTMA is presently set to 100 m for temperature. Additional terms can be added to Equation 2.18 to constrain the background errors in terms other than horizontal and spatial distance. For example, Myrick et al. (2005) introduced a third term to this equation to reflect that background error covariances are less likely to be correlated across mountain ranges. Specification of the background error covariances in terms of spatial distance alone may not be optimal for the RTMA, which will be discussed further in Chapter 4.

**Estimation of Observation and Background Error Covariances**

Using statistics obtained as part of the RTMA production cycle, it is possible to assess whether the assumptions used for the observation and background error variances and error covariances are appropriate. As discussed by Myrick and Horel (2006), the magnitudes of the observation and background error variances as well as the rate at which these errors decorrelate with distance can be estimated using an approach used by previous investigators (Lönnberg and Hollingsworth 1986; Xu et al. 2001). That approach is followed here. Observations and corresponding nearest values of the background fields during the 30-day period from 8 May-7 June 2008 were used to assess characteristics of the observation and background errors for the continental United States as a whole. Over 7 million observation-background pairs were used.
First, consider the bias (mean difference between the background field and the observations) as a function of time of day relative to solar noon (Figure 2.2). The bias of the background is very small when these hourly biases are averaged over the entire day and over all network types (the AVERAGE line). However, that reflects the offsetting tendencies for the downscaled RUC to be too cold at night and too warm during the day. One of the goals of this phase of the research is to estimate the observation errors as a function of network type. As mentioned earlier, concerns have been raised about the siting and reporting practices of some mesonets compared to others. The biases as a function of time of day tend to be quite similar for all observation types, with the exception of a lower bias for RAWS observations during the evening, which is likely due to the south-facing exposure of RAWS stations leading to higher late afternoon and evening temperatures compared to other stations. The sign of the bias during the night is opposite to that found for 1200 UTC by Myrick and Horel (2006). The study completed by Myrick and Horel evaluated the RUC only over the western United States during the winter season, in contrast to this study which evaluated the RUC over the entire CONUS domain during a month long period at the end of the spring season. The downscaling techniques between the two versions of the RUC evaluated were different as well. In the Myrick and Horel study, the RUC was downscaled by a simple horizontal bilinear interpolation. The RUC evaluated as part of this study used the RUC’s low level lapse rate to vertically interpolate the RUC’s surface temperature to the elevation of the new 5-km terrain. Discussions about other possible causes for the biases found here are currently ongoing with the RUC developers at the Earth Systems Research Laboratory.
Figure 2.2. Difference between background field and surface temperature observations (°C) averaged over the period 8 May 2008 – 7 June 2008 as a function of the local observation time and network type.
The present developmental version of the RTMA (at the time of this study’s evaluation) uses a bias correction method suggested by Dee and da Silva (1998) to mitigate the impact of systematic errors of the background field. This approach accumulates bias errors continuously and weights the most recent biases more heavily than older ones. The strong diurnal cycle in systematic errors evident in Figure 2.2 is not resolved by that procedure and not eliminated at the present time.

As discussed by Myrick and Horel (2006), the structure of the background error covariance can be estimated by examining the correlations between the innovations at one station with those at all nearby stations, which approximates a single vector of the array \( \bar{P}_b \) at the \( i \)th station and computed as follows:

\[
\rho_{ij} = \frac{(o_i - b_i')(o_j - b_j')}{s_{o_i-b_i}s_{o_j-b_j}} \tag{2.19}
\]

where \( i \) (\( j \)) denotes the \( i \)th (\( j \)th) observation location, the prime denotes the departure from the sample mean, and \( s \) denotes the standard deviation of the innovations about the sample mean. For example, Figure 2.3a shows the correlations between the innovations at Charlottesville, VA (KCHO) and those within the Shenandoah subdomain computed from all available observations in the 30-day period. The correlations tend to drop off sharply with distance and then remain above 0.3 for roughly 75 km.

The spatial pattern of the background error correlation in the vicinity of KCHO defined by Equation 2.19 using the decorrelation horizontal and vertical length scales of the RTMA is indicated by the by the shading in Figure 2.3a. Generally, the estimates of the observed correlations tend to be smaller nearby and larger over longer separation
distances than that used in the RTMA. They also do not show evidence of the strong
decorrelation with elevation implied by Equation 2.19, although few observations are
available at the higher elevations to help define that structure. The estimate and
specification of the correlation spatial pattern for Winchester, VA (KOKV) is shown in
Figure 2.3b. Again, the observed correlations between the background errors remain
above 0.3 for distances upwards of 75 km, while the specified background error
correlations are more tightly constrained both horizontally and vertically than those
estimated from the observations.

The examples in Figure 2.3 are two of the more than 11000 estimates of the
background error covariance that can be computed from the month-long sample of
observations and background fields. As discussed by Myrick and Horel (2006), the
covariance between observation innovations at two points separated by distance $r$ can be
accumulated over all the observation-gridpoint pairs for the monthly sample:

$$\text{cov}(r) = (o_1 - b_1)(o_2 - b_2)$$  \hspace{1cm} (2.20)

Then, using the same assumptions as Myrick and Horel (2006),

$$\text{cov}(r = 0) = \sigma_o^2 + \sigma_b^2$$  \hspace{1cm} (2.21)

and
Figure 2.3. Correlations (red numbers) between temperature innovations at a chosen example observation location and those at other locations within the Shenandoah subdomain computed over the 8 May - 7 June 2008 period. Shading indicates the shape of the background error correlation calculated by Equation 2.18 with selected decorrelation length scales. Distance (25 km intervals) from the selected observation location is indicated by the range rings. Elevation is contoured in meters. a. Correlations between Charlottesville, VA (KCHO) and all other observations using a horizontal (vertical) length scale of $R = 40$ km ($Z = 100$ m).
Figure 2.3. continued. b. Correlations between Winchester, VA (KOKV) and all other observations using a horizontal (vertical) length scale of $R = 40$ km ($Z = 100$ m).
Figure 2.3. continued. c. Correlations between Charlottesville, VA (KCHO) and all other observations using a horizontal (vertical) length scale of $R = 80$ km ($Z = 200$ m).
Figure 2.3. continued. d. Correlations between Winchester, VA (KOKV) and all other observations using a horizontal (vertical) length scale of $R = 80$ km ($Z = 200$ m).
where $\rho_{ij}$ is defined typically isotropically as the first exponential term of the right hand side of Equation 2.18.

The covariance of observation innovation as a function of distance $r$ during the month period computed for every location in the continental United States is shown in Figure 2.4 for all observation types as well as separated into the four primary network categories. Key statistics are also summarized in Table 2.1. The covariance drops slowly as a function of horizontal distance and does not asymptote to 0, which suggests that the downscaled RUC background fields contain errors that remain correlated over distances of hundreds of kilometers. Although the general behavior of the error covariance as a function radius is similar for the METAR, PUBLIC, and OTHER categories, the RAWS stations exhibit a roughly linear dependence with distance. This result suggests that the characteristics of the background errors in regions of complex terrain differ from those in other regions of the country.

Fitting a fifth-order polynomial to the innovation covariance values for horizontal distance greater than 5 km and extrapolating the curve back to $r = 0$ km allows an estimation of $\sigma_b$ determined for all network types to be 1.2°C. That value is half the estimate for temperature in the western United States during winter (2.5°C) determined by Myrick and Horel (2006). The lower value computed here suggests that the RUC background field is quite good when considered for the locations where observations are available in the continental United States (roughly 1% of the grid). The lower value may depend on a number of factors: (1) the downscaling procedure used for the RTMA is

$$\text{cov}(r) = \overline{b_1 b_2} = \sigma_b^2 \rho_{ij} \quad (2.22)$$
Figure 2.4. Binned innovation covariance (symbols) computed for the downscaled RUC background fields for the period 8 May 2008 to 7 June 2008 as a function of network type. Curve fits to the covariance are shown as a function of network type. The filled in symbols at $r = 0$ km indicate extrapolated estimates of background error variance as a function of network type. The open symbols at $r = 0$ km denote estimates of the sum of the observation and background error variance computed as a function of network type. Background error covariance specified by Equation 2.18 as a function of horizontal distance is also shown assuming lengths scales of 40 and 80 km (dotted lines).
superior to that used by Myrick and Horel (2006); (2) the background errors are less in regions of the country with less topographic variation than the western United States; and (3) the greater density of observations in metropolitan areas may bias the results.

Estimates of the RUC background error variance computed separately as a function of network type are also summarized in Table 2.1 and shown as filled symbols in the left margin of Figure 2.4. The dependence of the results on data density is suggested by the lower estimates of background error variance for the PUBLIC and OTHER categories relative to the METAR observations. The higher estimate of background error variance for RAWS observations reflects the locations of those observations are generally in regions of complex terrain.

Using Equation 2.21, the observation error variance can be estimated from the difference between the innovation covariance value at distance zero (filled symbols in Figure 2.4 and the $\sigma_b^2 + \sigma_o^2$ column in Table 2.1) and the estimates of $\sigma_b^2$. Thus, $\sigma_o$ for all stations in the continental United States is roughly 2.5°C, which is comparable to that estimated by Myrick and Horel (2006). As might be expected, $\sigma_o$ for METAR stations is estimated to be lower (2.1°C) than that for other network types. The larger observation

<table>
<thead>
<tr>
<th>Network</th>
<th>$\sigma_b^2$ (°C$^2$)</th>
<th>$\sigma_b^2 + \sigma_o^2$ (°C$^2$)</th>
<th>$\sigma_o^2$ (°C$^2$)</th>
<th>Avg. num./hr</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALL</td>
<td>1.4</td>
<td>7.5</td>
<td>6.1</td>
<td>11,464</td>
</tr>
<tr>
<td>METAR</td>
<td>2.0</td>
<td>6.2</td>
<td>4.2</td>
<td>1,744</td>
</tr>
<tr>
<td>PUBLIC</td>
<td>1.4</td>
<td>6.6</td>
<td>5.3</td>
<td>6,486</td>
</tr>
<tr>
<td>RAWS</td>
<td>2.6</td>
<td>12.6</td>
<td>10.0</td>
<td>1,301</td>
</tr>
<tr>
<td>OTHER</td>
<td>1.9</td>
<td>8.1</td>
<td>6.2</td>
<td>1,961</td>
</tr>
</tbody>
</table>
error (3.2°C) for RAWS stations is also expected, since the observation error arises from both instrumental and representativeness errors. Even an observation with minimal instrumental error may not be representative of the unknown true value on the scale of the 5×5 km² analysis grid.

Figure 2.4 and Table 2.1 provide support for using ratios of $\sigma_o^2 / \sigma_b^2$ of 2:1 rather than 1:1. Hence, the background temperature field should generally be “trusted” more than the observations. Further, the relatively slow decorrelation of background errors suggest that increasing $R$ and $Z$ is appropriate. For example, Figures 2.3c and 2.3d show the specification of the background error correlation when $R$ and $Z$ are doubled in Equation 2.18. That tends to broaden the error correlation in a manner more consistent with that estimated from the observations in this subdomain. Sensitivity of the LSA to the specification of the observation to background error variance and decorrelation length scales will be examined in the following chapter.

Local 2D-Var Surface Analysis

As mentioned in Chapter 1, considerable insight has been gained by performing sensitivity experiments for the nation as a whole with the RTMA code on the NCEP Haze development computer. However, it was found to be impractical to perform large numbers of sensitivity experiments remotely using that computer. Since some of the most complex aspects of any data assimilation system are those associated with preprocessing and quality control of the data, it is very economical to download the input files used by the RTMA (both the observations and the background fields) and then compute a 2D-Var analysis following Equations 2.15 and 2.16 for a limited domain using the GMRES
method (Saad and Schultz 1986). The GMRES method is an iterative method consisting of an inner and outer iteration that is used for solving large, sparse matrices. GMRES is more computationally efficient in speed and memory requirements than the conjugate gradient solution method, and does not suffer numeric instability from division by zero errors. Implementation of the GMRES method was done in a relatively small number of lines of code using Mathworks™ MATLAB® software.

Since the goal of this research is to evaluate approaches that might lead to improvements of the RTMA, the characteristics of the control experiments using the LSA are designed to be similar to those of the RTMA. For example, the control LSA experiments use Equation 2.18 to specify the temperature background error covariance sharing the same horizontal (40 km) and vertical (100 m) decorrelation length scales, respectively. The observation to background error variance ratio used by the analysis system is 1:1, which is the same ratio used for the METAR observations by the RTMA. The LSA is calculated over a 4° latitude by 4° longitude domain shown in Figure 1.2b centered on the Shenandoah region of Virginia using innovations (differences between observations and the background field) within the 4° by 4° domain for the 0900 UTC 22 October 2007 analysis.

Figure 2.5a shows the LSA temperature analysis valid at 0900 UTC 22 October 2007 for the Shenandoah subdomain. Immediately apparent is that some of the unrealistic features of the RTMA (Figure 1.5a) are not present here, e.g., the very high temperature analyzed by the RTMA over the highest terrain of the Blue Ridge Mountains. The LSA temperature increments shown in Figure 2.5b indicate that the positive innovations to the west of the Blue Ridge Mountains (e.g., nearly 5°C near 38.5°N, 78.6°W) do not
Figure 2.5. LSA temperature analysis and increments valid 0900 UTC 22 October 2007 over the Shenandoah Valley subdomain. Elevation is contoured in meters. a. Temperature analysis plotted in °C. Purple numbers are the observed temperature values (°C) used in the analysis.
Figure 2.5. continued. b. Temperature analysis increments (analysis minus background, °C) shaded in orange and purple. Green numbers denote innovation values (observation minus background in °C).
influence the analysis as much over the nearby higher terrain as seen in the RTMA (Figure 1.5b). Hence, the LSA tends to remain close to the downscaled RUC background in that region.

The differences between Figures 1.5 and 2.5 suggest that the application of the recursive filters and terrain smoothing in the RTMA to efficiently estimate Equation 2.18 may contribute to broadening the effective vertical decorrelation length scale. Given the results of the previous section in which the dependence on vertical separation appears small (see Figure 2.3), this approach is not necessarily a negative one. However, in synoptic situations such as that examined here, where large horizontal temperature gradients arise due to strong surface inversions, the LSA temperature analysis with the increments confined more tightly as a function of vertical separations appears subjectively better than the RTMA temperature analysis. Procedures are described in the next chapter to provide a more objective estimate of analysis sensitivity to the background error decorrelation length scales.
CHAPTER 3

OBJECTIVE EVALUATION OF SURFACE ANALYSES

Withholding Observations

Data denial experiments have been routinely used to quantitatively evaluate objective analyses (Seaman and Hutchinson 1985; Zapotocny et al. 2000; Hiemstra et al. 2006; Myrick and Horel 2008). The analyses computed with the restricted set of observations are usually compared to the withheld observations, the background fields, or the control analyses using all observations to define measures of accuracy and uncertainty of the analysis system (Seaman and Hutchinson 1985). The RTMA is designed to be able to withhold randomly 10% of observations from the analysis until the final iterative steps of the analysis as a means to provide an estimate of analysis uncertainty (de Pondeca et al. 2006).

The estimates of analysis accuracy and uncertainty can be sensitive to the approach used to randomly remove the observations from the analysis. Simply removing randomly every tenth observation is not optimal unless the observations are uniformly distributed spatially and exhibit equal observation error. Since observation networks in the continental United States tend to be clustered near urban areas (Figure 2.1), a withheld observation in an urban area may have less of an impact on an analysis than if
that observation was located in an area of low observation density (Seaman and Hutchinson 1985; Myrick and Horel 2008). Previous researchers have avoided this problem by taking into account the spatial distribution of the observations (de Pondeca et al. 2006) or by removing observations only from more randomly distributed networks (Myrick and Horel 2008).

The RTMA’s built-in cross validation and data denial tool attempts to minimize the impact of nonrandom observation networks by using the Hilbert curve (de Pondeca et al. 2006). The Hilbert curve is a space filling curve that occupies its entire domain, maintains spatial uniformity, and never overlaps upon itself (Sagan 1994). As an illustration of the steps required in generating the Hilbert curve, consider Figure 3.1. The first step is simply to define the sample of observations to be used, in this case, an artificial sample of 37 locations scattered across the continental United States (panel 1). Next, the entire domain is converted into a unit domain, and then subdivided into smaller quadrants until there is a maximum of one observation in each of the smaller subsections of the domain (panel 2). As evident in the second panel, many subsections may not contain an observation. The Hilbert curve (green) is drawn through each subsection in the entire domain. Subsections of the domain that do not contain an observation (grey dots) are ignored (panel 3). Next, the order in which the observations are located along the Hilbert curve is used to determine whether or not they are withheld from a specific analysis (panel 4). Here, every fifth observation (blue dots) (starting from the lower left hand corner of the domain and skipping empty subsections) is removed from the observation data set and considered to be the withheld sample (panel 5). This approach
Figure 3.1. Example of the Hilbert curve binning observations into data groups.
leads to five unique verification data sets that are spatially uniform in the context of the observation density.

While Figure 3.1 illustrates how a Hilbert curve can be computed for a small data set, that approach is too computationally expensive to be used for the over 10,000 observations available in the continental United States for the RTMA. The Fortran code developed for the RTMA computes a base-4 Hilbert coordinate based on each observation’s latitude and longitude. Observations are then binned into withholding groups based on their sequential order along the Hilbert curve, ignoring vertices for which no observations are available. This Fortran algorithm was ported to a standalone Fortran program independent of the IBM compiler subroutines used by the RTMA for the data denial experiments of this study.

One approach to using the Hilbert curve is to subdivide the observation data set into 10 subsets in which 10% of the observations are withheld from each analysis such that every observation uniquely belongs to one withholding group (Figure 3.2a). Although the Hilbert curve approach reduces the sensitivity of subsequent data denial experiments to observation density if that approach is used, care is also needed to lessen dependencies due to observation network type. Since the majority of observations available to the RTMA nationwide belong to the PUBLIC category (and represent nearly 80% of the observations in the Shenandoah subdomain), the impact of other network types can be difficult to assess if simply every tenth observation is withheld.

The presumption at the outset of this study was that the inexpensive instrumentation and siting issues typically associated with the PUBLIC network data may adversely affect analysis quality. However, the results presented in Chapter 2 suggest
that, at least for temperature, there is little difference in overall data quality between PUBLIC and other predominantly urban networks. Nevertheless, we control for network type by defining separate Hilbert curves for each of the four primary network categories within the $4^\circ \times 4^\circ$ latitude-longitude domain shown in Figure 2.2. Hence, roughly 6 METAR, 58 PUBLIC, 1 RAWS, and 7 OTHER stations are withheld in each of the 10 subsets. As shown in Figure 3.2b, this approach reclassifies the withholding data subsets from that evident in Figure 3.2a and occasionally leads to removing nearby observations if they belong to different network categories. Hence, the spatial uniformity aspect of the Hilbert curve is occasionally compromised. It should also be noted that every set of observations requires defining a new Hilbert curve, as adding a single location will change the binning for all subsequent observations along the curve.

Estimating Analysis Accuracy and Sensitivity

The LSA analysis shown in Figure 2.5 will be used to illustrate estimation of analysis accuracy and sensitivity, which is dependent on the Hilbert curve data denial technique. The differences between the control analysis and each of the 10 analyses for which 10% of the observations are withheld are shown in Figure 3.3. Including several observations located in the narrow valleys in the Appalachians leads to a colder control analysis than the analysis in which the first set of observations are withheld (upper left corner of Figure 3.3a). Generally, the impact of large innovations is small in the Washington, D.C. area, since the availability of so many other observations in that area diminishes the effect of omitting a few of them (Figures 3.3b, 3.3d, 3.3g, 3.3h, 3.3j). The impact of the relatively strong decorrelation with elevation in this instance ($Z = 100$ m in
Figure 3.2. Withholding observations using the Hilbert curve (bins 1-10). a. Applied to all observation networks.
Figure 3.2. continued. b. Applied separately for each network.
Figure 3.3. Difference (shading, °C) between control LSA temperature analysis at 0900 UTC 22 October 2007 and corresponding analysis using 90% of the observations (10% of observations withheld). Observation innovations (°C) withheld from the data withheld analysis are shown in green and the locations of observations used by both analyses are denoted by purple dots. Elevation is contoured in meters. a. Difference between control analysis and corresponding analysis withholding observation group 1.
Figure 3.3. continued. b. Difference between control analysis and corresponding analysis withholding observation group 2.
Figure 3.3. continued. c. Difference between control analysis and corresponding analysis withholding observation group 3.
Figure 3.3. continued. d. Difference between control analysis and corresponding analysis withholding observation group 4.
Figure 3.3. continued. e. Difference between control analysis and corresponding analysis withholding observation group 5.
Figure 3.3. continued. f. Difference between control analysis and corresponding analysis withholding observation group 6.
Figure 3.3. continued. g. Difference between control analysis and corresponding analysis withholding observation group 7.
Figure 3.3. continued. h. Difference between control analysis and corresponding analysis withholding observation group 8.
Figure 3.3. continued. i. Difference between control analysis and corresponding analysis withholding observation group 9.
Figure 3.3. continued. j. Difference between control analysis and corresponding analysis withholding observation group 10.
Equation 2.18) leads to large differences between the control analysis and the analyses with observations withheld in mountainous regions compared to those on the coastal plain (Figures 3.3c, 3.3e). The impact of large innovations in the low density areas often leads to the greatest sensitivities to the denied observations (Figure 3.3f, 3.3i).

Following Myrick and Horel (2008), the accuracy of the LSA analyses is estimated by calculating the root-mean-square error (RMSE):

\[
RMSE = \sqrt{\frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} (a_{ij} - o_{ij})^2}
\]  

(3.1)

where \(o_{ij}\) are the withheld observations, \(a_{ij}\) are the analysis values at the nearest grid point to the denied observations, \(N\) (roughly 70) is the number of observations withheld in each of the \(M = 10\) data denial experiments. Since each observation can belong to only one member \(j\), the RMSE is calculated using each station once. This RMSE estimate of analysis accuracy ignores observation error, and hence, should be viewed as a relative, not absolute, measure of analysis accuracy (Myrick and Horel 2008).

As shown in Table 3.1, the RMSE between the background values and all of the observations is 2.15°C. For comparison, for the control analysis in which all of the values are used, the RMSE is lowered to 1.62°C. These values bracket that for the data denial experiments—as more (fewer) observations are withheld, the resulting RMSE will approach that of the background (control analysis).

The RMSE accumulated over all 10 analyses in which 10% of the observations are uniquely withheld from the control analysis is estimated to be 1.93°C. Hence,
withholding 10% of the observations causes a degradation in the analysis quality of 0.31°C relative to the control analysis. Since the estimate of analysis quality is determined at only 720 observation and gridpoint pairs, the quality of the analysis at the much greater number of gridpoints away from observations is not evaluated in terms of the RMSE statistic.

Following Zapotocny et al. (2000) and Myrick and Horel (2008), the sensitivity of the analyses to the set of observations used is assessed by computing the root-mean-square sensitivity:

\[
S = \sqrt{\frac{\sum_{i=1}^{M} \sum_{l=1}^{L} (d_{ij} - c_{ij})^2}{ML}}
\]  

(3.2)
where \( c_{ij} \) is the control analysis containing all available observations, \( d_{ij} \) is the \( j \)th analysis withholding one of the \( M = 10 \) sets of observations, and \( L = 6384 \) is the total number of analysis gridpoints. As discussed by Zapotocny et al. (2000) and Myrick and Horel (2008), the sensitivity parameter is a measure of the magnitude of analysis change resulting from withholding data; a small value for \( S \) implies that the analysis is largely unaffected by the removal of the observations. When averaged over the entire LSA domain, \( S \) in this case is 0.26°C, which will be used as a baseline value in the next section.

**Sensitivity to the Specification of Observation and Background Error Covariance**

Using the procedures developed in the previous two sections, a series of experiments were carried out in which the dependence of the analysis at 0900 UTC 22 October 2007 to the specification of the ratio of observation to background error variance and background error covariance is examined. The results of Chapter 2 suggest that increasing the observation to background error variance ratio as well as lengthening the horizontal decorrelation length scale may be appropriate.

Table 3.1 summarizes the 6 sets of data denial experiments completed for this case in addition to the control set described in the previous section. Figure 3.4a shows the analysis for which the horizontal (vertical) decorrelation length scale in Equation 2.18 is doubled to 80 km (200 m). This corresponds to the decorrelation length scales shown in Figures 2.3c and 2.3d as well as that illustrated in Figure 2.4. As expected, expanding the lateral and vertical impact of the innovations results in a smoother analysis, and is evident
from comparing Figures 3.4a and 2.5a. The analysis increments shown in Figure 3.4b clearly have many smoother features compared to Figure 2.5b. Halving the decorrelation horizontal and vertical length scales results in an analysis that corresponds more closely to nearby observations while relaxing towards the background in areas of low observation density (Figure 3.5). While the reduction in RMSE at withheld observations for the halved decorrelation length scale experiment supports reduction of the decorrelation length scales (see Table 3.1), one must remember that (1) this is estimated for only a single analysis and focuses on only one type of synoptic/mesoscale situation; and (2) the RMSE is computed only over a sample size of 700 observation-gridpoint pairs. The RMSE summarized in Table 3.1 at the withheld locations (left column) can be compared to the RMSE computed at all locations (right column) as a measure of over- or under-fitting. For example, broadening (shrinking) the decorrelation length scales leads to higher (lower) RMSE computed at all locations relative to the control analysis that reflects underfitting (overfitting). The sensitivity (Table 3.2) can be used as a measure of the influence of withheld observations in data void regions, but unfortunately, cannot be used as an error statistic because there is no absolute baseline. As expected, sensitivity decreases as decorrelation length scale decreases, which reflects the reduced influence an observation has on the analysis.

The impact of the observation to background error ratio is examined in Figures 3.6 and 3.7 and summarized in Tables 3.1 and 3.2. Decreasing (increasing) this ratio, as shown by Figure 3.6 (Figure 3.7), increases (decreases) the influence of the observations by increasing (decreasing) the magnitudes of the increments. These effects lead to essentially no changes in RMSE at withheld locations (Table 3.1) while the sensitivity
Figure 3.4. LSA temperature analysis and increments valid at 0900 UTC 22 October 2007 using a horizontal (vertical) length scale of $R = 80$ km ($Z = 200$ m) and the ratio of observation to background error variance ($\sigma_o^2/\sigma_b^2$) set to 1. Analysis computed over the Shenandoah Valley subdomain. Elevation is contoured in meters. a. Analysis temperature plotted in °C. Purple numbers are the observed temperature values (°C) used in the analysis.
Figure 3.4. continued. b. Analysis temperature increments (analysis minus background) plotted in shades of purple and orange. Green numbers denote innovation values (observation minus background in °C) used in the analysis.
Figure 3.5. LSA temperature analysis and increments valid at 0900 UTC 22 October 2007 using a horizontal (vertical) length scale of $R = 20 \text{ km}$ ($Z = 50 \text{ m}$) and the ratio of observation to background error variance ($\sigma^2_o / \sigma^2_b$) set to 1. Analysis computed over the Shenandoah Valley subdomain. Elevation is contoured in meters. a. Analysis temperature plotted in °C. Purple numbers are the observed temperature values (°C) used in the analysis.
Figure 3.5. continued. b. Analysis temperature increments (analysis minus background) plotted in shades of purple and orange. Green numbers denote innovation values (observation minus background in °C) used in the analysis.
Table 3.2
Sensitivity to Withheld Observations Across Domain

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Sensitivity (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R = 40 \text{ km}, Z = 100 \text{ m}, \sigma_o^2 / \sigma_b^2 = 1$</td>
<td>0.26</td>
</tr>
<tr>
<td>$R = 80 \text{ km}, Z = 200 \text{ m}, \sigma_o^2 / \sigma_b^2 = 1$</td>
<td>0.29</td>
</tr>
<tr>
<td>$R = 20 \text{ km}, Z = 50 \text{ m}, \sigma_o^2 / \sigma_b^2 = 1$</td>
<td>0.20</td>
</tr>
<tr>
<td>$R = 40 \text{ km}, Z = 100 \text{ m}, \sigma_o^2 / \sigma_b^2 = 0.5$</td>
<td>0.34</td>
</tr>
<tr>
<td>$R = 40 \text{ km}, Z = 100 \text{ m}, \sigma_o^2 / \sigma_b^2 = 2$</td>
<td>0.19</td>
</tr>
<tr>
<td>$R = 80 \text{ km}, Z = 200 \text{ m}, \sigma_o^2 / \sigma_b^2 = 2$</td>
<td>0.22</td>
</tr>
<tr>
<td>$R = 20 \text{ km}, Z = 50 \text{ m}, \sigma_o^2 / \sigma_b^2 = 0.5$</td>
<td>0.26</td>
</tr>
</tbody>
</table>
Figure 3.6. LSA temperature analysis and increments valid at 0900 UTC 22 October 2007 using a horizontal (vertical) length scale of $R = 40$ km ($Z = 100$ m) and the ratio of observation to background error variance ($\sigma_o^2 / \sigma_b^2$) set to 0.5. Analysis computed over the Shenandoah Valley subdomain. Elevation is contoured in meters. a. Analysis temperature plotted in °C. Purple numbers are the observed temperature values (°C) used in the analysis.
Figure 3.6. continued. b. Analysis temperature increments (analysis minus background) plotted in shades of purple and orange. Green numbers denote innovation values (observation minus background in °C) used in the analysis.
Figure 3.7. LSA temperature analysis and increments valid at 0900 UTC 22 October 2007 using a horizontal (vertical) length scale of $R = 40$ km ($Z = 100$ m) and the ratio of observation to background error variance ($\sigma_o^2/\sigma_b^2$) set to 2. Analysis computed over the Shenandoah Valley subdomain. Elevation is contoured in meters. a. Analysis temperature plotted in °C. Purple numbers are the observed temperature values (°C) used in the analysis.
Figure 3.7. continued. b. Analysis temperature increments (analysis minus background) plotted in shades of purple and orange. Green numbers denote innovation values (observation minus background in °C) used in the analysis.
decreases by 0.15°C when the observation to background error ratio is increased from 0.5 to 2 (Table 3.2). Over- (under-) fitting is suggested by the combination of no improvement in RMSE at withheld locations and the corresponding changes in sensitivity at all gridpoint locations.

The last two experiments combine the complementary effects of increasing (decreasing) the decorrelation length scales while increasing (decreasing) the observation to background error ratio, as shown in Figure 3.8 (Figure 3.9). The decorrelation length scales and observation to background ratio used in the analysis depicted in Figure 3.8 are suggested by the results of Chapter 2. As shown by the analysis increments in Figure 3.8b (Figure 3.9b), broadening (shortening) the decorrelation length scales and increasing (decreasing) the observation to background error ratio increases (decreases) the area of influence of a particular observation while reducing (intensifying) the magnitude of observation increments. The similarity of the RMSE value at the withheld locations relative to that for all locations for the case of Figure 3.8 as shown in Table 3.1 is a positive result. A user could have some confidence that the analysis in data void regions has comparable quality to that in data rich areas. On the other hand, the large spread for these values for the case of Figure 3.9 continues to reflect overfitting by the analysis. Since the sensitivities for these two cases are not that different from one another (Table 3.2), driving the analysis towards the background (Figure 3.8) yields similar sensitivities to that when the analysis is driven towards the observations (Figure 3.9).

As a final illustration, similar statistics were computed for the single data denial experiment for the RTMA shown in Figure 1.6. The RTMA was found to have an RMSE relative to the withheld observations of 2.12°C and relative to all observations within
Figure 3.8. LSA temperature analysis and increments valid at 0900 UTC 22 October 2007 using a horizontal (vertical) length scale of $R = 80$ km ($Z = 200$ m) and the ratio of observation to background error variance ($\sigma_o^2/\sigma_b^2$) set to 2. Analysis computed over the Shenandoah Valley subdomain. Elevation is contoured in meters. a. Analysis Temperature plotted in °C. Purple numbers are the observed temperature values (°C) used in the analysis.
Figure 3.8. continued. b. Analysis temperature increments (analysis minus background) plotted in shades of purple and orange. Green numbers denote innovation values (observation minus background in °C) used in the analysis.
Figure 3.9. LSA temperature analysis and increments valid at 0900 UTC 22 October 2007 using a horizontal (vertical) length scale of $R = 20$ km ($Z = 50$ m) and the ratio of observation to background error variance ($\sigma_o^2 / \sigma_b^2$) set to 0.5. Analysis computed over the Shenandoah Valley subdomain. Elevation is contoured in meters. b. Analysis temperature plotted in °C. Purple numbers are the observed temperature values (°C) used in the analysis.
Figure 3.9. continued. b. Analysis temperature increments (analysis minus background) plotted in shades of purple and orange. Green numbers denote innovation values (observation minus background in °C) used in the analysis.
the limited domain of 1.64°C. For reference, the RMSE of the background field at the same withheld locations is 2.32°C. It is important to note that these statistics use a Hilbert curve computed over all observations in the United States and are not computed separately for each network type. The sensitivity of the RTMA for this case was 0.47°C, which is nearly double that found for all of the experiments performed using the LSA. This larger sensitivity and the spread between the RMSE computed at the withheld locations versus all locations likely reflects the overfitting assessed subjectively in Chapter 1.
CHAPTER 4

SUMMARY AND DISCUSSION

Summary

This study began as an evaluation of the operational and developmental versions of the RTMA by downloading the analysis grids from NCEP, generating graphics for subdomains of interest, and subjectively evaluating those graphics. Providing feedback to the NCEP developers based on this work led to some improvements of the code. However, permission to use NCEP’s Haze computer to run the RTMA remotely allowed this research to expand in scope and assess the impact of code modifications directly. Unfortunately, the high level of utilization of the Haze computer by NCEP developers for the entire suite of NCEP models led to such slow turnaround in the processing queues that a local analysis system was eventually determined to be a more efficient approach to complete this study. As with most modeling systems, the preprocessing steps required for the RTMA are the most complex to reproduce externally. This research was able to take advantage of all of the preprocessing done for the RTMA and simply download the background fields and observation files required to generate the local variational analyses for any subdomain within the continental United States.
The LSA, as with any 2D or 3D variational analysis system, relies on the specifications of the observation to background error variance ratio and the background error covariance. The RTMA relies on prior estimates of those quantities based on other NCEP data assimilation systems and subsequent modification based on sensitivity experiments and subjective evaluation of the analyses. This study is the first to examine innovations (difference between the observations and the downscaled 1-h RUC forecasts) over the continental United States for a month long period. Following Myrick and Horel (2006), statistics based on the sample innovations were used to crudely estimate the observation to background error variance and the dominant horizontal spatial scales of the background error covariance. The background error for temperature is estimated to be relatively small, roughly 1.4°C. This estimate from the downscaled 13-km 1-h RUC forecasts is a few tenths °C lower than that reported by Benjamin et al. (2004) for 20 km and 40 km 1-h RUC forecasts. In addition, Mesinger et al. (2004) found the RMS analysis error of the North American Regional Reanalysis to be between 2 and 2.5°C. Based on these comparisons, the quality of the RUC 1-h temperature forecasts is judged to be quite good when examined on the scale of the entire continental United States over an entire month.

Networks that tend to be located in regions of relatively high observation density (e.g., METAR and PUBLIC) tend to have lower observation errors than those located predominantly in low density areas (e.g., RAWS and OTHER). The latter results from the larger representativeness errors (discrepancies between the spatial and temporal scales of the observations relative to that of the analysis grid) found especially in mountainous
areas. Hence, appropriate ratios of the observation to background error variance appear to be higher than unity, perhaps as much as 2-3.

Estimating observation variance as a function of network type also suggests that, at least for temperature, the errors of the PUBLIC observations are roughly the same as METAR observations. Concerns regarding the quality of PUBLIC observations have been raised in the past, often due to the challenges of siting instrumentation in urban areas. Based on the results of this work, the quality of PUBLIC temperature observations passing the quality control procedures in place for the RTMA appears comparable to that obtained from the surface aviation network.

As also found by Myrick and Horel (2006) for the western United States only, the background errors remain correlated over longer distances than often assumed when considered over the entire continental United States. The decorrelation length scales of the background errors appear similar for all of the network types except for the RAWS network, which exhibits even higher correlations between the background errors for distances up to 200 km. The estimates of the background error covariance as a function of horizontal distance (Figure 2.4) suggest that the exponential structure of Equation 2.18 is not particularly representative of the downscaled RUC background temperature fields.

Objective evaluation of surface analyses can be accomplished by using data denial experiments. Using a Hilbert curve helps to reduce the influence of variations in data density across the analysis domain on the results of the withholding experiments. In this study, the Hilbert curve is used to bin all observations in the domain into 10 separate withholding groups while also attempting to minimize the impact of the uneven number of observations available from four distinct network types. A relative measure of analysis
accuracy can then be computed using all available observation-gridpoint pairs. However, since the number of observations is much less than the number of gridpoints, a sensitivity parameter was also used here to evaluate the analysis grid in its entirety.

Data denial experiments were performed using the 0900 UTC 22 October 2007 Shenandoah Valley case to evaluate different decorrelation length scales and observation to background error ratios. Although the results from estimation of the background and observation errors suggested that the decorrelation length scales needed to be broadened and the observation to background error ratio increased, the data denial experiments applied to this single case could not provide conclusive results for such adjustments.

Discussion and Recommendations

The focus of this study has been to develop ways to efficiently diagnose the accuracy and uncertainty associated with 2D-Var mesoscale analyses such as the RTMA. This work leads to a number of recommendations for analysis developers and end users as will be expanded upon here.

1. **Routine documentation of estimates of the downscaled RUC background bias and error variance** (similar to that shown in Figure 2.2 and Table 3.1) would be useful to RTMA analysis users to help assess the quality of the analyses. The background bias and error variance would provide useful diagnostic information about the quality of the background field. Due to the limited number of observations available to the analysis, the observations cannot turn a poor background field into a good analysis.
2. **Further research to specify the observation error variance and background error covariance on the basis of parameters other than horizontal and vertical distance is recommended.** While a goal at the outset of this work was to help identify the appropriate decorrelation length scales for observation error variance and background error covariance appropriate for the RTMA, the evaluation of this single case is insufficient to do so. Simply doing similar investigations for more cases for other regions or the entire continental-scale domain of the RTMA may not be productive. Not surprisingly, this work suggests that it may not be practical to tune those parameters as a function of horizontal and vertical separation in such a way that the analysis will be optimal in all synoptic situations and regions of the country. One of the advantages of the RTMA is that the background error covariance can be specified generally in terms of one or more characteristics of the background. Experiments were made with the RTMA specifying the background error covariance in part by the horizontal potential temperature gradient, but the results were not improved substantively relative to a control analysis. One approach may be to use a measure of boundary layer stability from the background field, i.e., assuming that locations and synoptic situations with similar stability will have similar background errors. That hypothesis would need to be tested through examination of background error statistics stratified by boundary-layer stability over a large sample of synoptic situations.
3. **The impact of the observation time window used by the RTMA should be evaluated.** While the count of surface observations typically available for the RTMA (~12000) is much greater than conventional METAR observations (Figure 2.1), a considerable number of available mesonet observations are not used by the RTMA. This limitation results from the tight operational constraint for the RTMA to receive, process, analyze, and deliver to end users the analyses by roughly 50 min past the hour. While plans exist to perform another analysis delayed by several hours or as much as a day later, that work remains unfunded (L. Anderson, personal communication). Evaluation of the benefits of extending the present time window or the creation of a delayed analysis that uses all observations is recommended.

4. **Efforts should continue to identify, access, and transmit additional mesonet data to NCEP.** The total count of observations available now (~12000) is somewhat misleading as so many of the observations used by the RTMA are located in urban areas where the data density is often greater than that necessary to adjust the background field adequately (see Figure 3.3). Since the surface analysis is only as good as the observations used in it, much greater attention needs to be placed on identifying and obtaining access to additional data, particularly in regions of otherwise low data density. Staff at NWS offices around the country have been very successful at identifying local data assets and should be encouraged to continue to assist in these efforts.

5. **The manual and automatic quality control procedures used by the RTMA analysis should be evaluated to determine their effectiveness at**
removing poor quality surface observations and minimizing the loss of high quality observations. It has been noticed subjectively that some high quality mesonet observations are discarded on occasion as part of manual and automated quality control checks prior to and during the RTMA analysis. For example, an entire observation may be omitted because one variable has been judged to fail a quality control check or a station may be assumed to be of poor quality based on network type.

6. **Although the use of recursive filters is necessary to operationally complete a variational mesoscale analysis (Purser 2005), the current implementation of the recursive filters used by the RTMA should continued to be examined.** Running a variational analysis of the sort developed for this study is not feasible even on the most modern supercomputers due to the significant memory requirements required (simply storing the entire background error covariance array for the continental-scale NDFD domain requires approximately 2,900 GB of memory). The results of this study suggest that the smoothing inherent in the recursive filters may be advantageous, if differing length scales for the background error covariance can be defined as a function of iteration. Hence, the initial iterations could rely on broad decorrelation length scales and then followed by later iterations with successively smaller decorrelation length scales. A similar approach was used by Myrick et al. (2005).

7. **The required accuracy for the RTMA for operational NWS applications should be established.** A fundamental question for end users of the RTMA,
particularly one of NWS forecasters, is whether the RTMA is “good enough” or needs to be improved before used more widely in operations. This research does not address this question, but does provide some guidance. First the question could be rephrased to be: Is the downscaled RUC background field “good enough”? The statistics developed as part of this study suggest that, at least for temperature, the background field is quite good in terms of some objective measures when considered on the continental scale. Whether being within 1-2°C when averaged over large spatial extents and large sample is “good enough,” depends on the application. Second, the question could then be asked: Does the variational analysis provide significant improvement to the background field? These results, for a particularly difficult analysis situation, are mixed. The developmental version of the RTMA used in this study did not appear to substantively improve on the background field and in some locations degraded it. Subsequent testing with the LSA suggests that it is possible to improve upon the background field incrementally by several tenths °C. Since the developmental version of the RTMA undergoes periodic improvements and enhancements, ongoing evaluation for operational applications of the downscaled background and RTMA fields is required to address this question.

8. **The strengths and weaknesses of continental-scale versus county warning area scale mesoscale analyses should continue to be examined.** The impetus for the RTMA has been to develop a timely mesoscale analysis on the continental scale, rather than to rely upon local analyses available to NWS
field offices using the Local Analysis and Prediction System (LAPS; Albers et al. 1996) or MatchObsAll (Foisy 2003). The benefits of centralized data assimilation systems outweigh the drawbacks as the preprocessing steps for comprehensive data assimilation are daunting. However, a hybrid centralized/local approach may relieve some of the limitations of the operational constraints of the present approach used for NWS applications. The capability of running a “local” RTMA of the type developed for this study would potentially be an improvement relative to the MatchObsAll analyses, which are constrained tightly by observations. The analysis code required for the prototype LSA is very small and fast using readily available software packages and requires computational resources (memory and processors) readily available in desktop computers and small servers. A highly detailed variational local analysis used locally by NWS weather forecasting offices is achievable for domain sizes up to roughly four times that used by this study (8° latitude by 8° longitude). Disseminating the downscaled RUC background grids has to be done once. NWS field offices could then locally run additional analyses using data assets that arrive after the RTMA cutoff and use specifications of observation and background errors more appropriate for their region of interest to aid in forecast preparation and evaluation.

9. **Continued research and development of high resolution mesoscale data assimilation systems are necessary.** The variational scheme chosen for the RTMA followed from other NCEP data assimilation systems. However, the limitations of a standalone 2D-Var analysis for the analysis of sensible surface
weather parameters on the mesoscale is widely acknowledged (Horel and Colman 2005). Progress towards a more comprehensive 4-dimensional data assimilation system on the mesoscale requires effort beyond that presently supported by operational and research agencies. Identifying the limitations and benefits of present operational techniques (Rabier 2005) versus more experimental approaches (e.g., Torn and Hakim 2008) is a critical necessity.
REFERENCES


Jascourt, S., cited 2008: Real-Time Mesoscale Analysis: What is the NCEP RTMA and how can it be used? [Available online at https://www.meted.ucar.edu/loginForm.php?urlPath=nwp/RTMA.]


