

SENSITIVITY OF SURFACE METEOROLOGICAL ANALYSES TO
OBSERVATION NETWORKS

by

Daniel Paul Tyndall

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

Doctor of Philosophy

Department of Atmospheric Sciences

The University of Utah

December 2011

Copyright © Daniel Paul Tyndall 2011

All Rights Reserved

ABSTRACT

A computationally efficient variational analysis system for two-dimensional meteorological fields is developed and described. This analysis approach is most efficient when the number of analysis grid points is much larger than the number of available observations, such as for large domain mesoscale analyses. The analysis system is developed using MATLAB software and can take advantage of multiple processors or processor cores. A version of the analysis system has been exported as a platform independent application (i.e., can be run on Windows, Linux, or Macintosh OS X desktop computers without a MATLAB license) with input/output operations handled by commonly available internet software combined with data archives at the University of Utah.

The impact of observation networks on the meteorological analyses is assessed by utilizing a percentile ranking of individual observation sensitivity and impact, which is computed by using the adjoint of the variational surface assimilation system. This methodology is demonstrated using a case study of the analysis from 1400 UTC 27 October 2010 over the entire contiguous United States domain. The sensitivity of this approach to the dependence of the background error covariance on observation density is examined. Observation sensitivity and impact provide insight on the influence of observations from heterogeneous observing networks as well as serve as objective

metrics for quality control procedures that may help to identify stations with significant siting, reporting, or representativeness issues.

TABLE OF CONTENTS

ABSTRACT.....	iii
ACKNOWLEDGEMENTS.....	vii
1 INTRODUCTION	1
Background and Motivation	1
Objectives and Outline.....	8
2 VARIATIONAL ASSIMILATION THEORY	12
Introduction.....	12
Analysis Space.....	14
Observation Space	17
Implementation of Constraints.....	19
3 IMPLEMENTATION OF VARIATIONAL ASSIMILATION THEORY WITHIN THE UU2DVAR.....	26
Introduction.....	26
Background Error Covariance	26
Computation and Storage of the Background Error Covariance Matrix	28
Background, Terrain, and Land/Water Mask	34
Usage of Observations and Quality Control	35
UU2DVar General Characteristics and Analysis Cycle	39
4 OBSERVATION IMPACTS AND SENSITIVITY	53
Introduction.....	53
Computation.....	53
Observation Networks	56
5 RESULTS AND DISCUSSION	72
Description of Case Study	72
Evaluation of Data Density Constraints.....	75
Sensitivity and Impact to Observation Networks	78

6 CONCLUSION.....	133
Summary	133
Recommendations and Future Work	137
REFERENCES	141

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. John Horel, for his support and guidance throughout this entire project, as well as providing the original NetCDF input/output routines and the asymmetric quality control used for wind observations used by the University of Utah Variational Surface Analysis (UU2DVar). I would also like to thank the other four members of my committee—Drs. Jim Steenburgh, Jan Paegle, Zhaoxia Pu, and Thomas Haiden, for their valuable contributions to this project. Dr. Haiden was especially helpful in providing background material and data from his own analysis system, as well as many helpful discussions on surface analyses in regions of complex terrain.

I would also like to thank Dr. Manuel de Pondeca for his help on understanding the covariance computations within the Real-Time Mesoscale Analysis, as well as providing me with the background fields used for the analyses computed as part of this study. MesoWest support, specifically Mr. Chris Galli and Dr. Judy Pechmann, were tremendously helpful in the development of a routine to download observations out of MesoWest to be used by UU2DVar. WeatherFlow, Inc., and the Oklahoma Mesonet also deserve special recognition in allowing their observing networks to be used free of charge for the case study presented here.

I would also like to thank Dr. Fred Carr for guiding me to adjoint analysis as a technique to be used to compute observation sensitivity and impact. The Center of High

Performance Computing at the University of Utah greatly aided this project by providing server and desktop support for this work.

Finally, I would like to thank my friends and family, as I wouldn't have gotten this far without them.

This research is supported by the National Ocean and Atmospheric Administration under grants NA07NWS468003 and NA10NWS4680005 as part of the CSTAR program.

CHAPTER 1

INTRODUCTION

Background and Motivation

High spatial and temporal objective surface analyses have become increasingly vital during the past decade. Such mesoscale analyses are needed in nowcasting and short-range forecasting for wind power management, transportation safety, wildfire management, dispersion modeling, as well as defense applications (Horel and Colman 2005). Some of these high resolution real-time objective analyses are generated by tools that are not part of a fully integrated analysis/forecast data assimilation cycle, as most numerical models fail to capture adequately many surface weather features due to insufficient spatial resolution as well as incomplete parameterization of boundary layer processes (Uboldi et al. 2008; Glowacki et al. 2011). Instead, surface grids from short-range forecasts are often used as a starting point in the objective analysis process and then adjusted on the basis of high density mesonet observations.

Lazarus et al. (2002) reviewed many of the operational and research mesoscale analysis systems available during the late 20th century. Some of these systems are no longer undergoing further development or have been officially retired. Examples of current operational high resolution objective analyses developed internationally include the Vienna Enhanced Resolution Analysis (VERA; Steinacker et al. 2006) and the

Integrated Nowcasting through Comprehensive Analysis (INCA; Haiden et al. 2010) systems for Austria and the Mesoscale Surface Analysis System (MSAS; Glowacki et al. 2011) run over Australia. All 3 of these analysis systems incorporate high density mesonet observations and generate surface analyses of temperature, moisture, and wind at resolutions of 1-4 km. Mesoscale objective analysis systems available in the United States of particular note include MatchObsAll (Foisy 2003), the Space and Time Multiscale Analysis System (STMAS; Xie et al. 2011), and the Real-Time Mesoscale Analysis System (RTMA; de Pondecia 2011). STMAS and the RTMA are run at regular intervals (15 minutes for STMAS; 1 h for the RTMA) over the contiguous United States (CONUS) domain. MatchObsAll is run at the discretion of National Weather Service (NWS) forecasters over local domains, which typically extends slightly beyond their areas of forecast responsibility. MatchObsAll and the RTMA are used operationally by NWS forecasters to help create and verify high resolution gridded forecasts of near-surface conditions across the United States (Glahn and Ruth 2003).

The methodologies used by these analysis systems can be categorized into two general classes. The first type consists of interpolation techniques (VERA, spline; INCA, inverse distance; and MatchObsAll, serpentine curve) that strive to have the analysis agree very closely with the available observations (Daley 1991). These approaches tend to be computationally efficient and work very well when the observations are spread relatively uniformly across the analysis domain and erroneous observations are identified and rejected as part of preprocessing quality control procedures. These approaches tend to suffer if the density of observations varies widely within the analysis domain as the interpolation techniques may tightly constrain the analysis where the observations are

plentiful leading to overfitting in nearby data void regions (Myrick et al. 2005; Barker et al. 2007). Approaches that fall within the second general class of analysis system (MSAS, optimum interpolation; RTMA, two-dimensional variational, 2DVar; and STMAS, three-dimensional variational, 3DVar) assume that observations may contain errors arising from the representativeness of the observations within their surrounding environment as well as instrumentation errors (Daley 1991; Kalnay 2003). These approaches are particularly appropriate for analysis systems that rely on observations from heterogeneous networks with differing quality control standards that are often distributed unevenly within the analysis domain.

The RTMA serves as a reference analysis system for the research undertaken in this study. The National Centers for Environmental Prediction (NCEP) developed the RTMA to support the needs of NWS forecasters (Pondeca et al. 2011). The RTMA is an objective surface analysis system with the ability to assimilate tens of thousands of surface observations collected from many different data providers to yield analysis grids of 2-m temperature, 2-m dewpoint, surface pressure, and 10-m u and v winds over a CONUS domain as well as Alaska, Hawaii, Puerto Rico, and Guam domains. The analysis grids of the RTMA at resolutions of 2.5- and 5-km conform to the National Digital Forecast Database (NDFD) grid described by Glahn and Ruth (2003).

The computational resources required to compute mesoscale surface analyses such as the RTMA (with $\sim 10^7$ gridpoints) are considerable. In addition, techniques to manage appropriately the diversity of observational assets that lead to variations in data density and quality have heretofore remained largely unexplored. Tyndall et al. (2010) examined the sensitivity of the RTMA and a comparable 2DVar system to assumptions

about the observational and background error covariances in part as a function of observational type. Mesoscale data assimilation depends to a great extent on the number of observations available to modify the specified background field. For example, the 5-km resolution RTMA has approximately 15,000 surface observations available to adjust the background fields at over 700,000 gridpoints, while nearly 2.4 million gridpoints are required for the 2.5-km resolution CONUS RTMA. In addition, since these ~15,000 surface observations are not evenly spread throughout the entire analysis grid and are often clustered near urban areas, the number of observations providing independent observations is often substantially less.

The Integrated Data Influence (IDI), as described by Uboldi et al. (2008), can be used as a nondimensional measure of data density. Figure 1.1 depicts the IDI of all surface observations used by the RTMA to compute the 1400 UTC 27 October 2010 temperature analysis, using the RTMA's assumptions regarding the observational and background error covariances, i.e., the factors that affect the influence of observations on the analysis. Subject to the aforementioned assumptions related to the error covariances, regions of the domain with IDI values approaching one have more complete data coverage, while regions with low IDI values have few observations available. The inequitable distribution of observations is of concern everywhere, but the complex underlying terrain of the western United States results in localized microclimates that remain difficult to resolve on the basis of the present observational network. Proposed improvements to the current approach through the development of a Nationwide Network of Networks (NNoN; National Research Council 2009) are unlikely to provide the

number of observations necessary to resolve all of these local weather features around the country.

Variational data assimilation systems suffer from the necessity to specify the spatial scales of the background error covariance. Specifying large spatial scales for those errors appropriate for regions where few observations are available may lead to the inability to capture small-scale structures evident in data-rich areas. For example, Figure 1.2 presents an artificial 2-m temperature analysis and corresponding analysis increments (adjustments to the background field) where a relatively dense observation network is embedded within a data sparse region. All of the observations (outlined circles) in Figure 1.2a generally have good agreement with each other, except at the very center of the observing network, where there are three observations that are warmer than those surrounding it. The temperature analysis in this case fails to capture the higher temperature feature here, as the assimilation scheme is tuned to extend the influences of the observations to the data sparse areas at the edge of the domain. Further, the “washed out” nature of the analysis increments (Figure 1.2b) near the center of the domain is due to the large number of cooler observations surrounding the three warmer observations, which limit the influence of the warmer observations to properly adjust the background field to the observed temperatures in the center of the domain. Tuning the assumptions about the background error covariance to resolve the small scale features would help to define the higher temperatures near the center, but would degrade the analysis in the surrounding data-sparse areas. Hence, adjusting the background error covariance as a function of data density may help to extend the influence of observations in otherwise

data sparse regions while maintaining the ability to resolve smaller-scale features where the observing network is capable of resolving them.

The application of data density dependent observation weights or background error covariances has been studied previously for several different assimilation methods with mixed results. Lorenc et al. (1991) implemented decreased weights for observations located in data dense areas in the United Kingdom Meteorological Office's Analysis Correction data assimilation scheme, to improve the influence of observations in nearby data sparse regions. Later research using the European Centre for Medium-Range Weather Forecasts (ECMWF) 3DVar assimilation system showed that shorter spatial scales used to specify the background error covariance improved forecasts in data dense areas, while longer spatial scales improved the forecast in data sparse areas (Andersson et al. 1998). Unfortunately, the assimilation system used in that research could only utilize a single structure function at a time (which is used to specify the spatial scales and construct the background error covariance), and Andersson et al. were unable to evaluate the impacts to the forecast on using an observation density dependent structure function. However, their research notes that implementation of this feature into the 3DVar system would likely be beneficial.

The ~15,000 observations that were used to generate the IDI analysis in Figure 1.1 come from over 100 different mesonets across the United States. The impact on analyses of the quality of observations resulting from networks with differing reporting practices, instrumentation, maintenance, siting, and representativeness is of great interest, especially for the development of the NNoN (National Research Council 2009). The National Research Council report, motivating the necessity of the NNoN, emphasizes the

need for improved and ongoing documentation of metadata regarding existing mesonets for such applications. For example, observations from the Remote Automated Weather Station (RAWS) mesonet are typically sited on southern slopes with anemometer heights of 6 m, instead of the 10 m height standard utilized by observations from the NWS (Horel and Dong 2010; Tyndall et al. 2010). Observations from the Citizen Weather Observing Program (CWOP) typically come from consumer grade instrumentation and may be sited on the roof of or next to a building, unlike the mandatory field of clearance and professional grade equipment required for NWS observations. Observations from different mesonet providers with differing instrumentation, standards, and siting can be used by the analysis, provided that the assumptions about the observation errors for each network are appropriately evaluated. Defining those assumptions is facilitated by determining the impact of each network on the analysis.

As described by Tyndall et al. (2010) and Horel and Dong (2010), the Local Surface Analysis (LSA), a 2DVar analysis tool written in MATLAB that utilizes an assimilation scheme similar to the RTMA, has been used on local computer nodes maintained by the Center for High Performance Computing (CHPC) at the University of Utah. Although examination of appropriate error covariances for the LSA (and correspondingly the RTMA) as well as analysis sensitivity to selected observation networks was shown to be possible with the LSA, that approach is practical only for limited regional domains (approximately 6° latitude by 6° longitude) due to the computational requirements of the assimilation algorithm. In order to be able to efficiently compute analyses over continental scale domains, the development of a new variational surface analysis tool was initiated as part of this research and herein referred

to as the University of Utah Variational Surface Analysis (UU2DVar). This development included parallelizing the assimilation computation, implementing highly efficient programming practices using modest computer resources, as well as shifting the computation of the analysis from analysis space to observation space. The adjoint of the UU2DVar has also been developed as part of this research so that it may be used in future research to efficiently assess the impact of observation networks as part of efforts related to the NNoN. The UU2DVar, as well as differences between it and the RTMA, are described in a Chapter 3.

Objectives and Outline

The objectives of this study are:

- To document the algorithms used by the UU2DVar to efficiently produce continental-scale surface analyses.
- To show that specifications of the background error covariance based on observation density allow relatively small-scale features to be resolved in areas of high data density while allowing the limited observations in data-sparse regions to influence analyses on broader scales.
- To apply the adjoint of the UU2DVar to assess analysis sensitivity to and impacts of individual mesonets on analyses.

Chapter 2 of this document discusses and describes variational assimilation theory, which is used by the UU2DVar to generate surface analyses. Variational theory is discussed in both the analysis space framework (utilized by the LSA) and the observation space framework (utilized by UU2DVar). Chapter 3 discusses the computational

implementation of the variational framework used by the UU2DVar, namely the mathematical technique used to simplify the background error covariance matrix, the parallelization technique and usage of sparse matrices to decrease both wall clock time (the time needed to compute a quantity) and memory usage, formulation of the background error correlations, as well as adjustments to and quality control of observations utilized by the data assimilation tool. The adjoint of the UU2DVar, its derivation, and its application to specify background error covariance as a function of data density as well as the analysis sensitivity to differing observation networks is described in Chapter 4. A particular case study is used in Chapter 5 to demonstrate the use of this methodology. Finally, a summary and conclusions follow in Chapter 6. Future work is also presented in that chapter.

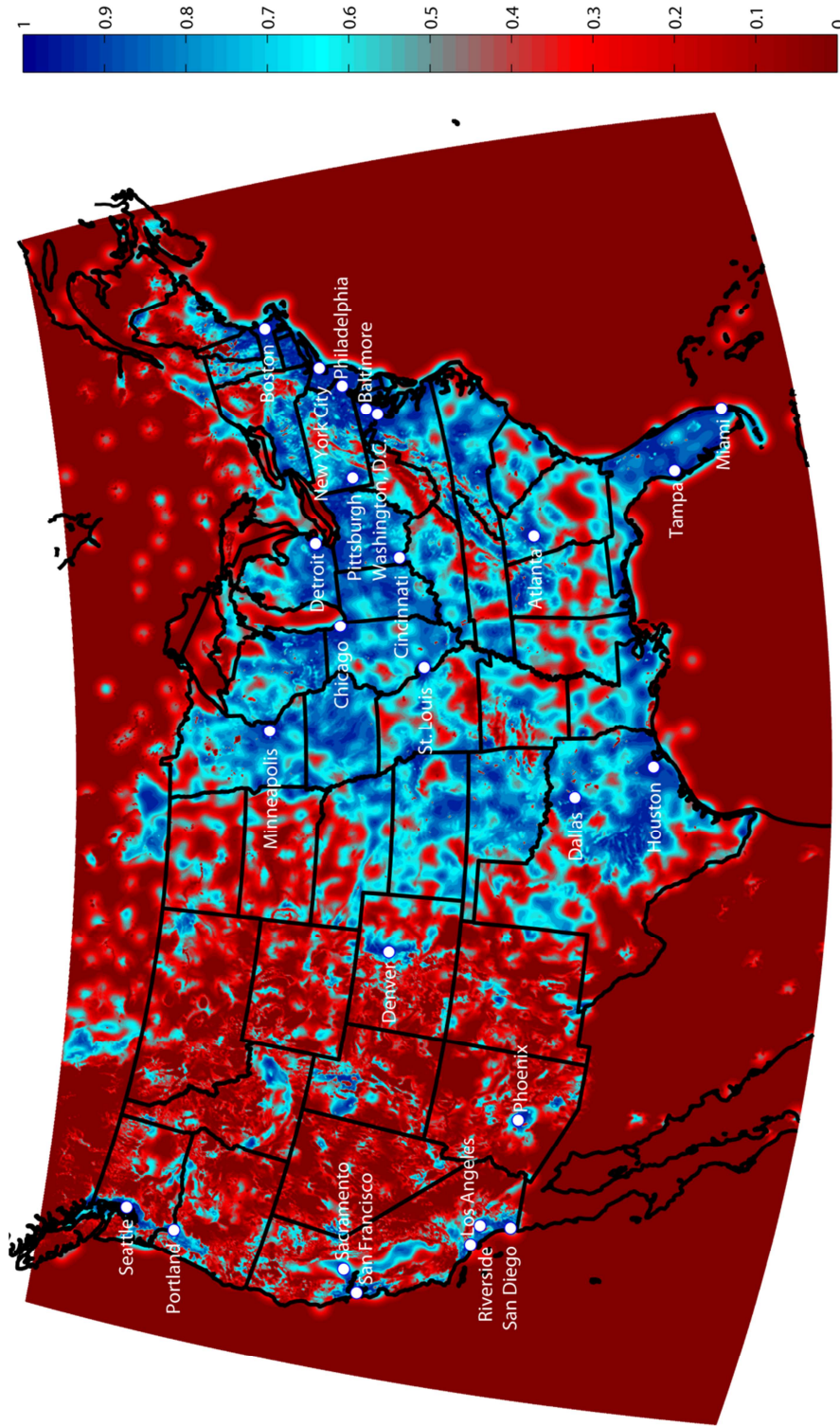


Figure 1.1.1. IDI analysis computed using 2-m temperature observations available to RTMA on 1400 UTC 27 October 2010 along with the RTMA's spatial scales used to specify the background error covariance. The 25 largest metropolitan within CONUS areas (by population) are also noted on the map. High data densities (blue) are clustered around urban areas, while rural areas are often data sparse (light red) or data void (dark red).

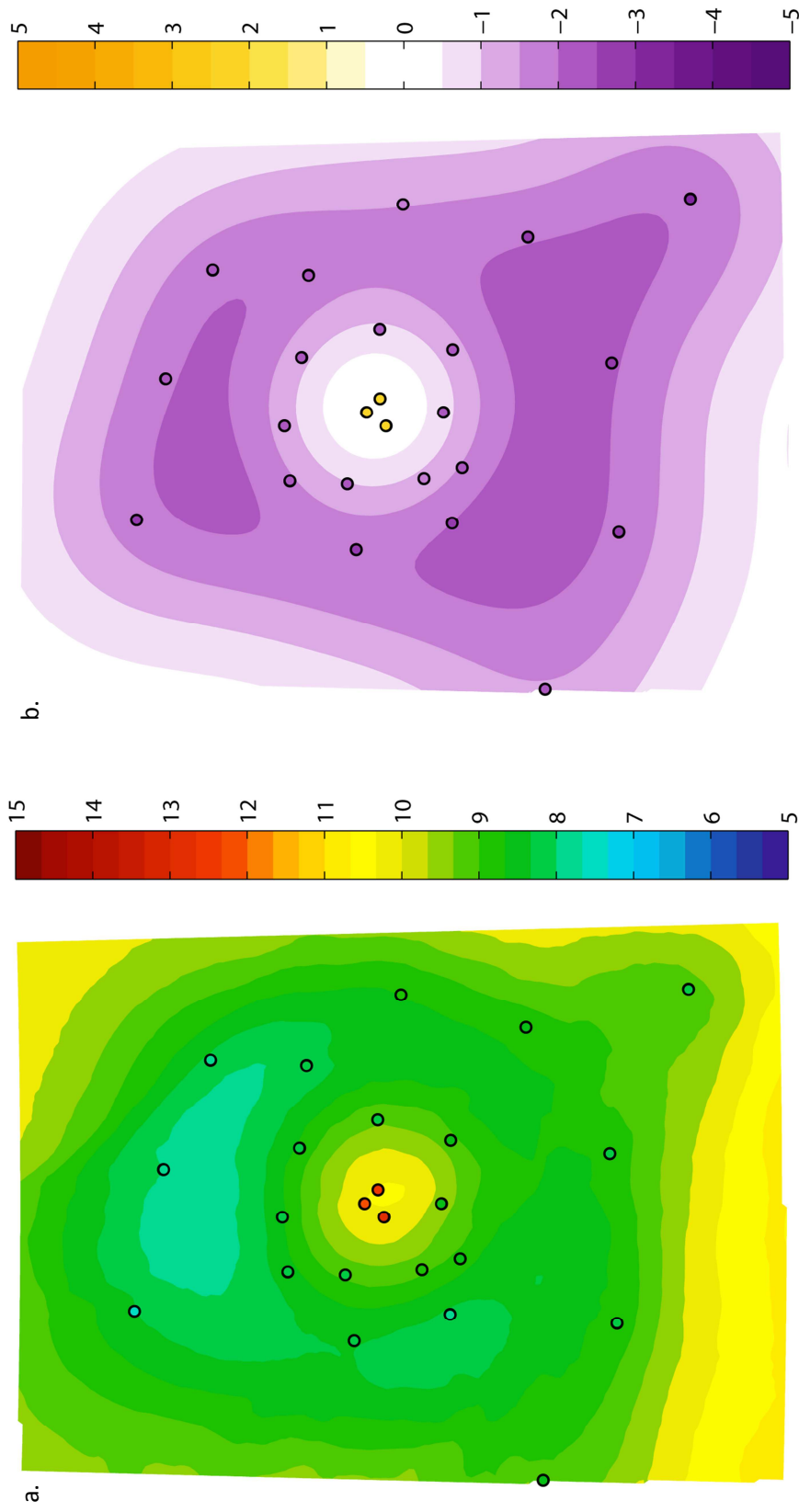


Figure 1.2. Artificial 2-m temperature analysis and observations, along with corresponding analysis increments and observation innovations. a. Analysis temperature, shaded in $^{\circ}\text{C}$, along with temperature observations (circle markers, shaded in $^{\circ}\text{C}$). b. Analysis temperature increments in $^{\circ}\text{C}$, along with temperature observation innovations (circle markers, shaded in $^{\circ}\text{C}$).

CHAPTER 2

VARIATIONAL ASSIMILATION THEORY

Introduction

Gridded objective analyses are generated from typically irregularly distributed observations combined with a background field on a continuous grid subject to statistical assumptions and constraints (McPherson 1975; Talagrand 1997; Kalnay 2003). Such data assimilation algorithms have been necessary since the advent of meteorological modeling in the 1950s and many of those early approaches (e.g., Cressman method [Gilchrist and Cressman 1954], and successive corrections [Bergthórsson and Döös 1955; Cressman 1959]) continue to be used. As computational resources have improved, techniques such as optimal interpolation (Gandin 1963) and time-independent 2DVar and 3DVar assimilation (Sasaki 1958) have been introduced. Most recently, time-dependent (four-dimensional) variational assimilation (4DVar; Sasaki 1970) and ensemble-based data assimilation (e.g., Evensen 1994) are used by some operational centers and by many research groups.

While some research groups are beginning to study 4DVar and ensemble based techniques for high resolution surface assimilation (N. Baker, personal communication; Ancell et al. 2011), these approaches are too computationally expensive to be used for real-time, high resolution, large scale mesoscale analyses. Instead, time-independent

variational approaches remain the most computationally affordable solutions for mesoscale analyses and will be studied here. Following Kalnay (2003), all 2DVar and 3DVar approaches seek to minimize the cost function, J ,

$$2J(\mathbf{x}_a) = J_b + J_o \quad (2.1)$$

where the terms J_b and J_o penalize the analysis \mathbf{x}_a for differences from the background field and observations respectively. As will be discussed later in this chapter, additional weak constraints (J_c) based on the underlying terrain or flow dependencies can be introduced:

$$2J(\mathbf{x}_a) = J_b + J_o + J_c \quad (2.2)$$

To minimize the cost function, Equation 2.1 is expanded:

$$2J(\mathbf{x}_a) = (\mathbf{x}_a - \mathbf{x}_b)^T \mathbf{P}_b^{-1} (\mathbf{x}_a - \mathbf{x}_b) + [\mathbf{H}(\mathbf{x}_a) - \mathbf{y}_o]^T \mathbf{P}_o^{-1} [\mathbf{H}(\mathbf{x}_a) - \mathbf{y}_o] \quad (2.3)$$

where \mathbf{x}_b and \mathbf{y}_o correspond, respectively, to the background field and observation vectors, \mathbf{P}_b and \mathbf{P}_o define, respectively, the background and observation error covariance matrices, and \mathbf{H} is an operator that maps the analysis onto the observations. There are two widely used approaches to minimize Equation 2.3 to yield an analysis: (1) solve iteratively for a solution on the analysis grid (analysis space, Parrish and Derber 1992; Courtier et al. 1998), or (2) solve iteratively for a solution at the observation locations

(observation space, Lorenc 1986; Daley and Barker 2001). These two methods are discussed in the next sections.

Analysis Space

In the analysis space framework, the relationship

$$\mathbf{H}(\mathbf{x}_a) - \mathbf{y}_o = \mathbf{H}(\mathbf{x}_a - \mathbf{x}_b + \mathbf{x}_b) - \mathbf{y}_o = \mathbf{H}(\mathbf{x}_b) + \mathbf{H}(\mathbf{x}_a - \mathbf{x}_b) - \mathbf{y}_o \quad (2.4)$$

is substituted into Equation 2.3 yielding:

$$\begin{aligned} 2J(\mathbf{x}_a) = & (\mathbf{x}_a - \mathbf{x}_b)^T \mathbf{P}_b^{-1} (\mathbf{x}_a - \mathbf{x}_b) \\ & + [\mathbf{H}(\mathbf{x}_a - \mathbf{x}_b) + \mathbf{H}(\mathbf{x}_b) - \mathbf{y}_o]^T \mathbf{P}_o^{-1} [\mathbf{H}(\mathbf{x}_a - \mathbf{x}_b) + \mathbf{H}(\mathbf{x}_b) - \mathbf{y}_o] \end{aligned} \quad (2.5)$$

The right side of Equation 2.5 is algebraically expanded yielding Equation 2.6:

$$\begin{aligned} 2J(\mathbf{x}_a) = & (\mathbf{x}_a - \mathbf{x}_b)^T \mathbf{P}_b^{-1} (\mathbf{x}_a - \mathbf{x}_b) + [\mathbf{H}(\mathbf{x}_a - \mathbf{x}_b)]^T \mathbf{P}_o^{-1} [\mathbf{H}(\mathbf{x}_a - \mathbf{x}_b)] \\ & + [\mathbf{H}(\mathbf{x}_a - \mathbf{x}_b)]^T \mathbf{P}_o^{-1} [\mathbf{H}(\mathbf{x}_b) - \mathbf{y}_o] \\ & + [\mathbf{H}(\mathbf{x}_b) - \mathbf{y}_o]^T \mathbf{P}_o^{-1} [\mathbf{H}(\mathbf{x}_a - \mathbf{x}_b)] \\ & + [\mathbf{H}(\mathbf{x}_b) - \mathbf{y}_o]^T \mathbf{P}_o^{-1} [\mathbf{H}(\mathbf{x}_b) - \mathbf{y}_o] \end{aligned} \quad (2.6)$$

To reduce the expense of computing the inverse of \mathbf{P}_b (due to its large size), the cost function is transformed into a function of \mathbf{v} :

$$\mathbf{v} = \mathbf{P}_b^{-1}(\mathbf{x}_a - \mathbf{x}_b) \quad (2.7)$$

yielding:

$$\begin{aligned} 2J(\mathbf{v}) = & \mathbf{v}^T \mathbf{P}_b^T \mathbf{v} + \mathbf{v}^T \mathbf{P}_b^T \mathbf{H}^T \mathbf{P}_o^{-1} \mathbf{H} \mathbf{P}_b \mathbf{v} + \mathbf{v}^T \mathbf{P}_b^T \mathbf{H}^T \mathbf{P}_o^{-1} [\mathbf{H}(\mathbf{x}_b) - \mathbf{y}_o] \\ & + [\mathbf{H}(\mathbf{x}_b) - \mathbf{y}_o]^T \mathbf{P}_o^{-1} \mathbf{H} \mathbf{P}_b \mathbf{v} \\ & + [\mathbf{H}(\mathbf{x}_b) - \mathbf{y}_o]^T \mathbf{P}_o^{-1} [\mathbf{H}(\mathbf{x}_b) - \mathbf{y}_o] \end{aligned} \quad (2.8)$$

The minimum of the cost function from Equation 2.8 is computed by finding where the gradient of the function is 0:

$$0 = \nabla J(\mathbf{v}) = \mathbf{P}_b^T \mathbf{v} + \mathbf{P}_b^T \mathbf{H}^T \mathbf{P}_o^{-1} \mathbf{H} \mathbf{P}_b \mathbf{v} + \mathbf{P}_b^T \mathbf{H}^T \mathbf{P}_o^{-1} [\mathbf{H}(\mathbf{x}_b) - \mathbf{y}_o] \quad (2.8)$$

yielding:

$$-\mathbf{P}_b^T \mathbf{H}^T \mathbf{P}_o^{-1} [\mathbf{H}(\mathbf{x}_b) - \mathbf{y}_o] = [\mathbf{P}_b^T + \mathbf{P}_b^T \mathbf{H}^T \mathbf{P}_o^{-1} \mathbf{H} \mathbf{P}_b] \mathbf{v} \quad (2.9)$$

Equation 2.9 is solved iteratively for \mathbf{v} by the conjugate gradient solution method (CGS; Hestenes and Stiefel 1952) or the generalized minimum residual method (GMRES; Saad and Schultz 1986). The analysis is computed from Equation 2.10, which reflects that the background field is modified by the innovation ($\mathbf{P}_b \mathbf{v}$):

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{P}_b \mathbf{v} \quad (2.10)$$

Although equations 2.9 and 2.10 simplify the analysis by eliminating the computation of the inverse of \mathbf{P}_b , the computation and storage of \mathbf{P}_b itself is no trivial task. \mathbf{P}_b is a matrix of size $[n \times n]$, where n is the number of gridpoints in the analysis. For a continental scale two-dimensional analysis, n can be on the order from 10^5 to 10^6 depending upon the horizontal spacing of the grid. Storage of a double precision matrix of these sizes ranges from 74 GB to 7.4 TB, which can be difficult to store in memory even on supercomputers. Furthermore, computation of the full background error covariance matrix is expensive; e.g., computing \mathbf{P}_b generally takes about 5.5 h on 8 processor cores (unless otherwise specified, all wall clock times were measured using a compute node with 2 Xeon hex-core processors clocked at 2.80 GHz) for the types of cases studied here, which is not suitable for an analysis that might be needed for real-time applications. Although wall clock time can be reduced by using more processors, operational centers often have serious constraints on computing resources due to the large number of numerical products needed to run on the same supercomputer (for example, the 5-km RTMA is only run on 16 processors of NCEP's 4,992 processor computer cluster [M. de Ponca, personal communication]).

The background error covariance matrix is often approximated to circumvent these storage and computation problems (Fisher 2003). There are many different ways to approximate \mathbf{P}_b , such as modeling the matrix in both spectral and spatial coordinates using the wavelet formation (Buehner and Charron 2007), using a diffusion operator (Weaver and Courtier 2001), or using a recursive filter (Lorenc 1986; Purser et al. 2003a,

2003b; de Ponte et al. 2011). Since the background error covariance matrix defines the spatial scales over which observations influence the analysis, it is important to approximate the matrix as accurately as possible (Daley 1991; Fisher 2003). Since much of the covariance modeling research (Weaver and Courtier 2001; Purser et al. 2003a, 2003b; Buehner and Charron 2007) has focused on global analysis products, simplifications of the covariance matrices over such large domains take advantage of synoptic-scale balances, such as geostrophic and hydrostatic balance. Unfortunately, for mesoscale conditions within the planetary boundary layer, such balances are not appropriate (Bannister 2008a, 2008b).

Observation Space

In the observation space framework, the gradient of the cost function, presented in Equation 2.3, is computed immediately:

$$0 = 2\nabla J(\mathbf{x}_a) = \mathbf{P}_b^{-1}(\mathbf{x}_a - \mathbf{x}_b) + \mathbf{H}^T \mathbf{P}_o^{-1}[\mathbf{H}(\mathbf{x}_a) - \mathbf{y}_o] \quad (2.11)$$

Equation 2.11 is multiplied by the background error covariance to avoid computing its inverse yielding:

$$\mathbf{x}_a - \mathbf{x}_b = \mathbf{P}_b \mathbf{H}^T \mathbf{P}_o^{-1}[\mathbf{y}_o - \mathbf{H}(\mathbf{x}_a)] \quad (2.12)$$

Similar to the derivation in physical space, the substitution

$$\boldsymbol{\eta} = \mathbf{P}_o^{-1}[\mathbf{y}_o - \mathbf{H}(x_a)] \quad (2.13)$$

is introduced into Equation 2.12 to yield Equation 2.14:

$$x_a - x_b = \mathbf{P}_b \mathbf{H}^T \boldsymbol{\eta} \quad (2.14)$$

Equation 2.14 is transformed into observation space by multiplying the equation by the forward transform operator, \mathbf{H} :

$$\mathbf{H}(x_a) - \mathbf{H}(x_b) = \mathbf{H} \mathbf{P}_b \mathbf{H}^T \boldsymbol{\eta} \quad (2.15)$$

Additional simple manipulation leads to:

$$\mathbf{H}(x_a) - \mathbf{y}_o + \mathbf{y}_o - \mathbf{H}(x_b) = \mathbf{H} \mathbf{P}_b \mathbf{H}^T \boldsymbol{\eta} \quad (2.16)$$

$$-\mathbf{P}_o \boldsymbol{\eta} + \mathbf{y}_o - \mathbf{H}(x_b) = \mathbf{H} \mathbf{P}_b \mathbf{H}^T \boldsymbol{\eta} \quad (2.17)$$

Finally, the terms multiplied by $\boldsymbol{\eta}$ in Equation 2.17 are separated to one side of the equation, which allows for $\boldsymbol{\eta}$ to be solved iteratively:

$$\mathbf{y}_o - \mathbf{H}(x_b) = (\mathbf{H} \mathbf{P}_b \mathbf{H}^T + \mathbf{P}_o) \boldsymbol{\eta} \quad (2.18)$$

The analysis is then computed by rearranging Equation 2.14:

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{P}_b \mathbf{H}^T \boldsymbol{\eta} \quad (2.19)$$

The observation space approach eliminates the storage problem of the background error covariance since the transpose of \mathbf{H} filters unneeded information from the background error covariance. Rows of \mathbf{P}_b can be computed individually, multiplied by their respective columns of \mathbf{H}^T , and stored. The product, $\mathbf{P}_b \mathbf{H}^T$ has dimensions $[n \times m]$, where m is the number of observations assimilated by the analysis (see Figure 2.1). Memory efficiency is greatly improved by using this approach, provided that the analysis is under-sampled (i.e., the number of observations is much less than the number of gridpoints [Daley and Barker 2001]). Memory requirements for the computation in observation space for double precision data can range from 7.4 GB to 74.5 GB for analyses of 10^5 to 10^6 gridpoints, which is within the memory capacity of many computer clusters. While these memory requirements are still significant, additional approximations can be made and additional computational methods can be implemented to allow analyses to be generated using modest computing resources. These methods and approximations used by the UU2DVar are covered in Chapter 3.

Implementation of Constraints

Although the undersampling of observations is exploited by the analysis technique presented by Equations 2.18 and 2.19, undersampling remains a significant problem for high resolution analyses. If the observations are sparse in a particular area of the analysis domain, then the data assimilation system depends upon the background field to produce the analysis (Equation 2.1). Many high resolution surface analysis frameworks

and systems (Myrick et al. 2005; de Pondecà et al. 2011; Haiden et al. 2011) downscale coarse resolution background fields to the analysis grid using a variety of methods. Unfortunately, the downscaled background field may not resolve many small-scale weather features (Myrick et al. 2005), and in some cases, may produce erroneous features in these areas through the downscaling process.

The usage of constraints can help improve the analysis by supplying information to the data assimilation system not provided by the background field or observations (Lorenc 1986; Xie et al. 2002). The constraint can either be formulated as a weak constraint or a strong constraint (Zhu and Yan 2006). Strong constraints modify either the background error covariance or the background itself. The strong constraint may add balanced coupling between two different assimilated fields, add background error correlation to a meteorological parameter or topography field, or may impose some other fundamental limit or law to the analysis (Lorenc 1986; Protat and Zawadzki 1999; Xie et al. 2002). As will be discussed in Chapter 3, the UU2DVar (as well as the RTMA) uses differences in elevation as an anisotropic constraint.

The addition of a strong constraint defined by the density of observations is introduced in Chapter 3 and tested in Chapter 4 as part of this research. Because of the direct modification to the background error covariance or background field, the strong constraints are assumed to be perfect and force the subsequent analyses to meet the balance requirements of the specific constraint (Lorenc 1986; Xie et al. 2002).

In contrast, a weak constraint does not force the analysis to meet the constraint exactly, which can be advantageous if the constraint is only an approximation (Xie et al.

2002). One formulation of the weak constraint, J_c , presented in Equation 2.2 is often expanded in the form:

$$J_c = (\mathbf{x}_a - \mathbf{x}_c)^T \mathbf{P}_c^{-1} (\mathbf{x}_a - \mathbf{x}_c) \quad (2.20)$$

where \mathbf{x}_c is the constrained field and \mathbf{P}_c is a term that describes the error covariance of the constrained field (Lorenz 1986). The term \mathbf{P}_c describes the weighting of the minimization of the difference between the analysis and the constraint relative to the difference between the analysis and the observations and the analysis and the background field. The weak constraint may also be implemented as additional artificial observations in \mathbf{y}_o in the cost function.

The solution to the variational analysis equation with a weak constraint as described by Equation 2.2 becomes more complicated than the basic observation space equations presented by Equations 2.18 and 2.19 and also doubles the memory cost of the analysis. Equation 2.21 presents the variational cost function with the explicit weak constraint implementation presented by Equation 2.20:

$$\begin{aligned} 2J(\mathbf{x}_a) = & (\mathbf{x}_a - \mathbf{x}_b)^T \mathbf{P}_b^{-1} (\mathbf{x}_a - \mathbf{x}_b) + [\mathbf{H}(\mathbf{x}_a) - \mathbf{y}_o]^T \mathbf{P}_o^{-1} [\mathbf{H}(\mathbf{x}_a) - \mathbf{y}_o] \\ & + (\mathbf{x}_a - \mathbf{x}_c)^T \mathbf{P}_c^{-1} (\mathbf{x}_a - \mathbf{x}_c) \end{aligned} \quad (2.21)$$

The gradient of Equation 2.21 is computed to find the minimum of the cost function:

$$0 = 2\nabla J(\mathbf{x}_a) = \mathbf{P}_b^{-1}(\mathbf{x}_a - \mathbf{x}_b) + \mathbf{H}^T \mathbf{P}_o^{-1} [\mathbf{H}(\mathbf{x}_a) - \mathbf{y}_o] + \mathbf{P}_c^{-1}(\mathbf{x}_a - \mathbf{x}_c) \quad (2.22)$$

As with the observation space framework presented in the last section, Equation 2.22 is multiplied by the background error covariance to simplify the computation of the analysis:

$$0 = \mathbf{x}_a - \mathbf{x}_b + \mathbf{P}_b \mathbf{H}^T \mathbf{P}_o^{-1} [\mathbf{H}(\mathbf{x}_a) - \mathbf{y}_o] + \mathbf{P}_b \mathbf{P}_c^{-1}(\mathbf{x}_a - \mathbf{x}_c) \quad (2.23)$$

Equation 2.23 is multiplied by the observation operator \mathbf{H} to simplify the terms involving \mathbf{P}_b :

$$0 = \mathbf{H}(\mathbf{x}_a - \mathbf{x}_b) + \mathbf{H} \mathbf{P}_b \mathbf{H}^T \mathbf{P}_o^{-1} [\mathbf{H}(\mathbf{x}_a) - \mathbf{y}_o] + \mathbf{H} \mathbf{P}_b \mathbf{P}_c^{-1}(\mathbf{x}_a - \mathbf{x}_c) \quad (2.24)$$

Expansion and rearrangement yields:

$$\begin{aligned} \mathbf{H}(\mathbf{x}_b) + \mathbf{H} \mathbf{P}_b \mathbf{H}^T \mathbf{P}_o^{-1} \mathbf{y}_o + \mathbf{H} \mathbf{P}_b \mathbf{P}_c^{-1} \mathbf{x}_c \\ = \mathbf{H}(\mathbf{x}_a) + \mathbf{H} \mathbf{P}_b \mathbf{H}^T \mathbf{P}_o^{-1} \mathbf{H}(\mathbf{x}_a) + \mathbf{H} \mathbf{P}_b \mathbf{P}_c^{-1} \mathbf{x}_a \end{aligned} \quad (2.25)$$

Finally, the analysis vector is factored out of the right hand side of the equation, yielding:

$$\begin{aligned}
\mathbf{H}(x_b) + \mathbf{H}\mathbf{P}_b\mathbf{H}^T\mathbf{P}_o^{-1}y_o + \mathbf{H}\mathbf{P}_b\mathbf{P}_c^{-1}x_c \\
= \left[\mathbf{H} + \mathbf{H}\mathbf{P}_b\mathbf{H}^T\mathbf{P}_o^{-1}\mathbf{H} + \mathbf{H}\mathbf{P}_b\mathbf{P}_c^{-1} \right] x_a
\end{aligned} \tag{2.26}$$

In this form, the analysis is computed by directly solving for the analysis vector through an iterative solution method (as in Equations 2.8 and 2.18). Equation 2.26 can be generalized for K multiple constraints:

$$\begin{aligned}
\mathbf{H}(x_b) + \mathbf{H}\mathbf{P}_b\mathbf{H}^T\mathbf{P}_o^{-1}y_o + \mathbf{H}\mathbf{P}_b \sum_{k=1}^K \mathbf{P}_{c_k}^{-1}x_{c_k} \\
= \left[\mathbf{H} + \mathbf{H}\mathbf{P}_b\mathbf{H}^T\mathbf{P}_o^{-1}\mathbf{H} + \mathbf{H}\mathbf{P}_b \sum_{k=1}^K \mathbf{P}_{c_k}^{-1} \right] x_a
\end{aligned} \tag{2.27}$$

The preceding derivation assumes that \mathbf{P}_c is a diagonal matrix, as it is not computationally feasible to calculate the inverse of a $[n \times n]$ matrix. As in the basic observation space framework, the combined matrix $\mathbf{P}_b\mathbf{H}^T$ can be stored efficiently. However, the computational memory required doubles using Equation 2.27 because another product must be stored: $\mathbf{H}\mathbf{P}_b$. This product also does not require storing explicitly the entire background error covariance matrix as individual rows of the \mathbf{H} matrix can be multiplied by individual columns of the \mathbf{P}_b matrix to yield the product $\mathbf{H}\mathbf{P}_b$. Although the memory cost for an analysis using Equation 2.27 doubles compared to Equations 2.18 and 2.19, wall clock time increases only slightly. Because \mathbf{P}_b is symmetric, individual rows of \mathbf{P}_b computed during the computation of $\mathbf{P}_b\mathbf{H}^T$ can be transposed to yield the

individual columns of \mathbf{P}_b needed for the computation of \mathbf{HP}_b (computation of these matrix products is further discussed in Chapter 3).

The UU2DVar supports both the usage of strong and weak constraints in the computation of the analysis; however, weak constraints are not investigated as part of this research. The UU2DVar could be utilized to study weak constraints as described by Equation 2.21 such as the additional utilization of a statistical model of orographic flow appropriate to the underlying terrain, as the product \mathbf{HP}_b is computed as part of the assimilation cycle to compute the adjoint (described in Chapter 4). The implementation of the strong constraints used as part of this research, which includes basic terrain anisotropy and a data density term, are discussed in Chapter 3.

$$\begin{array}{c}
 \begin{array}{cccccc}
 \sigma_b & A & F & J & M & O \\
 A & \sigma_b & B & G & K & N \\
 F & B & \sigma_b & C & H & L \\
 J & G & C & \sigma_b & D & I \\
 M & K & H & D & \sigma_b & E \\
 O & N & L & I & E & \sigma_b
 \end{array} \\
 \mathbf{P}_b
 \end{array}
 \times
 \begin{array}{c}
 \begin{array}{cc}
 a & g \\
 b & h \\
 c & i \\
 d & j \\
 e & k \\
 f & l
 \end{array} \\
 \mathbf{H}^T
 \end{array}
 =
 \begin{array}{c}
 \begin{array}{cc}
 \mathbf{P}_b \mathbf{H}^T_{1,1} & \mathbf{P}_b \mathbf{H}^T_{1,2} \\
 \mathbf{P}_b \mathbf{H}^T_{2,1} & \mathbf{P}_b \mathbf{H}^T_{2,2} \\
 \mathbf{P}_b \mathbf{H}^T_{3,1} & \mathbf{P}_b \mathbf{H}^T_{3,2} \\
 \mathbf{P}_b \mathbf{H}^T_{4,1} & \mathbf{P}_b \mathbf{H}^T_{4,2} \\
 \mathbf{P}_b \mathbf{H}^T_{5,1} & \mathbf{P}_b \mathbf{H}^T_{5,2} \\
 \mathbf{P}_b \mathbf{H}^T_{6,1} & \mathbf{P}_b \mathbf{H}^T_{6,2}
 \end{array}
 \end{array}$$

$$\mathbf{P}_b \mathbf{H}^T_{1,1} = \sigma_b a + Ab + Fc + Jd + Me + Of \quad \mathbf{P}_b \mathbf{H}^T_{1,2} = \sigma_b g + Ah + Fi + Jj + Mk + Ol$$

$$\mathbf{P}_b \mathbf{H}^T_{2,1} = Aa + \sigma_b b + Bc + Gd + Ke + Nf \quad \mathbf{P}_b \mathbf{H}^T_{2,2} = Ag + \sigma_b h + Bi + Gj + Kk + Nl$$

$$\mathbf{P}_b \mathbf{H}^T_{3,1} = Fa + Bb + \sigma_b c + Cd + He + Lf \quad \mathbf{P}_b \mathbf{H}^T_{3,2} = Fg + Bh + \sigma_b i + Cj + Hk + Ll$$

$$\mathbf{P}_b \mathbf{H}^T_{4,1} = Ja + Gb + Cc + \sigma_b d + De + If \quad \mathbf{P}_b \mathbf{H}^T_{4,2} = Jg + Gh + Ci + \sigma_b j + Dk + Il$$

$$\mathbf{P}_b \mathbf{H}^T_{5,1} = Ma + Kb + Hc + Dd + \sigma_b e + Ef \quad \mathbf{P}_b \mathbf{H}^T_{5,2} = Mg + Kh + Hi + Dj + \sigma_b k + Ek$$

$$\mathbf{P}_b \mathbf{H}^T_{6,1} = Oa + Nb + Lc + Id + Ee + \sigma_b f \quad \mathbf{P}_b \mathbf{H}^T_{6,2} = Og + Nh + Li + Ij + Ek + \sigma_b l$$

Figure 2.1. Methodology to compute each row of the background error covariance individually (shaded) to yield $\mathbf{P}_b \mathbf{H}^T$. Variables A through O and a through l are dummy variables.

CHAPTER 3

IMPLEMENTATION OF VARIATIONAL ASSIMILATION

THEORY WITHIN THE UU2DVAR

Introduction

As mentioned previously, the UU2DVar solves the variational cost function in observation space. Chapter 2 presented a derivation of the observation space solution to the cost function, and the UU2DVar's implementation of those equations (e.g., the specification of the error covariances, quality control of observations, analysis computation cycle) is covered here.

Background Error Covariance

Individual elements of the background error covariance used by the UU2DVar are computed by:

$$\mathbf{P}_{bij} = \sigma_b \exp\left(-\frac{r_{ij}^2}{\mathcal{R}_k[I_i]^2}\right) \exp\left(-\frac{z_{ij}^2}{\mathcal{Z}_k[I_i]^2}\right) \quad (3.1)$$

where r and z are the horizontal and vertical great circle distances between gridpoints i and j . The horizontal and vertical decorrelation length scale terms (\mathcal{R}_k and \mathcal{Z}_k ,

respectively) in the denominators of the exponential terms in Equation 3.1 are not constants as used in prior studies (Myrick et al. 2005; Tyndall 2008; Horel and Dong 2010; Tyndall et al. 2010), but are instead are assumed to be functions of the data density at gridpoint i as measured by the dimensionless IDI (I). \mathcal{R}_k and \mathcal{Z}_k are defined for the k^{th} sixth order polynomials of the form:

$$\mathcal{R}_k(I_i) = R \cdot (p_{k,1}I_i^6 + p_{k,2}I_i^5 + p_{k,3}I_i^4 + p_{k,4}I_i^3 + p_{k,5}I_i^2 + p_{k,6}I_i + p_{k,7}) \quad (3.2)$$

$$\mathcal{Z}_k(I_i) = Z \cdot (p_{k,1}I_i^6 + p_{k,2}I_i^5 + p_{k,3}I_i^4 + p_{k,4}I_i^3 + p_{k,5}I_i^2 + p_{k,6}I_i + p_{k,7}) \quad (3.3)$$

where $p_{k,1}$ through $p_{k,7}$ are coefficients of the k^{th} polynomial that determine its shape, and R and Z , respectively, are horizontal and vertical decorrelation length scales of the type used in the previously cited studies. The polynomial functions presented in Equations 3.2 and 3.3 were selected due to their ease in modifying their shapes by simply changing the polynomial coefficients. Figure 3.1 depicts the various forms of \mathcal{R}_k and \mathcal{Z}_k as a function of the IDI that were studied as part of this research. In this study, R and Z are set to 80 km and 200 m, respectively; these values were determined by Tyndall et al. (2010) for the CONUS domain and were tested in a case study over the area surrounding the Shenandoah Valley, VA. Similarly, the background error variance, σ_b , is set to 1°C for 2-m temperature and 2-m dewpoint, and 1 m/s for u and v winds.

The IDI, as defined by Ubaldi et al. (2008), is computed by generating an analysis where all of the background values are assumed to be zero and all observations are assumed to be one. The IDI is completely dependent upon the assumptions made regarding the observation and background error covariances. In this study, the

background error covariance used by the IDI is always specified by Equation 3.1 using the set of polynomial coefficients corresponding to $k = 0$ (see Figure 3.1). The ratio of the observation error variance to background error variance (σ_o^2/σ_b^2) is also always 1 for all IDI computations.

The IDI is a measure of the influence on the analysis by the observations; however, it also a measure of observation density (Horel and Dong 2010). Regions of the analysis where the IDI is near 1 indicate data rich areas with multiple stations in close proximity, while values near 0 indicate data void regions. With σ_o^2/σ_b^2 set to 1, the value of the IDI for an analysis with a single observation near that particular observation's gridpoint is 0.5 (i.e., the observation and background contribute equally to the final analysis). As a test of the use of the IDI, 0.5 is used as a point of inflection for the $k = 1$ and $k = 2$ polynomials. These polynomials force the decorrelation length scales to decrease significantly as the IDI approaches 1 and thereby allow finer-scale structures in the analysis than when $k = 0$. In the case of the $k = 2$ polynomial, the decorrelation length scales are substantially increased as the IDI approaches 0 for completely data void areas and thereby allow deviations between isolated observations and the background to influence a broader region. Specifying the background error covariance as a function of data density will be examined in Chapter 5.

Computation and Storage of the Background

Error Covariance Matrix

As mentioned in Chapter 2, the computation and storage of the background error covariance matrix is one of the most significant challenges in variational data

assimilation. Even with the usage of the observation space framework, additional approximations must be made and advanced computation methods must be implemented to compute and store the background error covariance over a continental scale domain within real-time analysis constraints. The UU2DVar utilizes the following four methods to improve the computation and storage of the background error covariance matrix (in the form of $\mathbf{P}_b\mathbf{H}^T$):

1. Usage of sparse matrix mathematics
2. Variational localization
3. Computation of only needed elements of \mathbf{P}_b
4. Parallel computing

Although the largest matrix stored by the UU2DVar is of size $[n \times m]$ instead of $[n \times n]$, a significant amount of memory is required to store this matrix for continental scale variational data assimilation problems. For the CONUS 5-km resolution domain used in this research and the 15,000 observations assimilated each hour, storage of the full $\mathbf{P}_b\mathbf{H}^T$ matrix requires approximately 75 GB of memory. Although this is feasible for large supercomputers, it is not necessary to store the full $\mathbf{P}_b\mathbf{H}^T$ matrix, as sparse matrices can be utilized to reduce memory requirements as well as wall clock time. Unlike the full matrix, which explicitly stores every element of a matrix, the sparse matrix only stores nonzero elements of the matrix, along with the index locations of those nonzero elements. Using sparse matrices only saves significant memory if the matrix to be stored has enough nonzero elements. For example, $\mathbf{P}_b\mathbf{H}^T$ is a two-dimensional matrix; therefore the sparse form of $\mathbf{P}_b\mathbf{H}^T$ must store both row and column indices as well as the values of the

nonzero elements within the matrix. For memory savings to be realized, $\mathbf{P}_b\mathbf{H}^T$ must be at least 66.7% element sparse (i.e., at least 66.7% of its elements must be 0).

Unfortunately, $\mathbf{P}_b\mathbf{H}^T$ does not meet this requirement, even though \mathbf{H}^T is generally an extremely sparse matrix (as discussed below, only one value is nonzero in each row for this study). To force element sparseness, variational localization can be used to add additional zero elements to the matrix product. Depending on how the background error covariance is specified, an observation assimilated using variational methods can influence analysis gridpoint values thousands of kilometers away. These extremely large scale correlations may not be accurate (Hamill et al. 2001), especially in the case of undersampled assimilation problems, which is typical with surface observations. Variational localization refers to the elimination of extremely small error correlations. In the UU2DVar, small error correlations are not even computed as part of the specification of the covariance matrix, which not only reduces memory requirements (from sparse matrix implementations), but also improves computational time. The UU2DVar implements variational localization through a maximum radius of influence, which in this study is set to 3.75 times the maximum horizontal decorrelation scale, i.e., 300 km for polynomial coefficients $k = 0$ and $k = 1$, and 600 km for $k = 2$. This corresponds to removing all correlations that are smaller than 7.8×10^{-7} . This maximum radius of influence was chosen somewhat arbitrarily, and reducing it further will decrease both wall clock time and memory requirements and may have little impact on the resulting analyses.

Figure 3.2 depicts the difference between an IDI temperature analysis computed without variational localization and one with variational localization, using a western

United States domain centered over Utah. As shown by Figure 3.2, differences between the two analyses are negligible, with a maximum difference on the order of 10^{-6} . The majority of the largest differences are located in data sparse and data void regions; however, these differences are extremely small. Furthermore, the usage of variational localization is supported by other analysis systems that also utilize the technique, e.g., the RTMA (M. de Pondeva, personal communication). Usage of variational localization is supported by the functional form of Equation 3.1, which assumes background error correlations asymptote to 0 at large distances.

In addition to reducing the memory requirements to store $\mathbf{P}_b\mathbf{H}^T$, there is also a need to significantly reduce its computation time. Computation of $\mathbf{P}_b\mathbf{H}^T$ on a single processor for a continental scale problem can take days; however, the wall clock time can be significantly reduced by only computing required elements of the background error covariance matrix \mathbf{P}_b that correspond to the nonzero rows of \mathbf{H}^T . As illustrated in Figure 3.3, only the first and fifth columns of \mathbf{P}_b actually need to be computed to yield the full $\mathbf{P}_b\mathbf{H}^T$ matrix for the simple example depicted in Figure 2.1. Wall clock time using the true covariance and forward operator is significantly reduced from days to less than 30 minutes, as the number of surface observations is typically 2-3 orders of magnitude smaller than the number of analysis gridpoints.

The wall clock time needed to compute the background error covariance matrix can further be reduced through parallel computing. The computation of \mathbf{P}_b is typically classified as an embarrassingly parallel computing problem, as the only interprocessor communication is at the start of the routine, to distribute pieces of information used to compute \mathbf{P}_b to individual workers, and at the end of the routine, to gather up the final

result from the individual workers to assemble a full matrix. Embarrassingly parallel computing problems typically have near perfect speedup; i.e., wall clock time is reduced by half when the number of available processors doubles. Because the UU2DVar utilizes sparse matrices in its computation of the background error covariance matrix, the problem is more complicated than the simple parallel for loop, as large amounts of full matrices must be computed and then converted to sparse matrices, instead of making all of these computations using sparse matrices (due to reallocation of memory). For completeness, the algorithm used to compute $\mathbf{P}_b \mathbf{H}^T$ (and $\mathbf{H} \mathbf{P}_b$) is depicted in Figure 3.4 and described below (variable names used by the code are italicized):

1. Nonzeros rows \mathbf{H}^T are identified to determine which columns of \mathbf{P}_b must be computed.
2. Resulting indices from (1) are divided into equal parts by the number of processors (*nprocs*) used by the analysis.
3. Each individual section of indices (owned by a particular processor) is divided into further subsections, based on the value of the user tunable variable *numpbrowscomptonce*. This particular variable controls how many columns of \mathbf{P}_b the computer as a whole (not individual processor) is allowed to operate on at once. Therefore, the length of each subsection each individual processor may operate on at one time is *numpbrowscomptonce/nprocs*.
4. Each processor computes each subsection of its assigned indices of \mathbf{P}_b column by column, through Equation 3.1, using a full matrix to store the resulting computations. When a processor reaches the end of the subsection, the entire

\mathbf{P}_b subsection is converted from a full matrix to a sparse matrix, and the processor moves on to the next subsection.

5. When all subsections have been completed, each individual section of \mathbf{P}_b owned by each processor is gathered into a single sparse matrix. Note that this matrix is not the full covariance matrix, as only elements that would not be reduced to 0 by multiplication of the transpose of the forward operator are computed.
6. The matrix product $\mathbf{P}_b\mathbf{H}^T$ is computed. The matrix storing the needed elements of \mathbf{P}_b is transposed, which is required for the computation of \mathbf{HP}_b .

Figure 3.5 shows the speedup (black thin line) of the algorithm used to compute $\mathbf{P}_b\mathbf{H}^T$ and \mathbf{HP}_b as a function of the number of processors. The speedup is a ratio of the computer time required for a code to run on a single processor versus the time required to run on multiple processors. Parallel algorithms with perfect speedup (depicted in Figure 3.5 as a thick grey line) have wall clock times that are halved when the number of processors used to compute the algorithm is doubled. Perfect speedup can be difficult to achieve due to communication overhead between processors. The speedup depicted in Figure 3.5 measures the average of 10 trials computing $\mathbf{P}_b\mathbf{H}^T$ and \mathbf{HP}_b for all CONUS temperature observations for the 1400 UTC 27 October 2010 analysis (approximately 14,000 observations over 740,000 gridpoints). Speedup of the UU2DVar's $\mathbf{P}_b\mathbf{H}^T$ and \mathbf{HP}_b computation is significantly less than ideal for larger numbers of processors because all processors used as part of this test share the same memory. This forces each processor to operate on a smaller piece of the background error covariance matrix at one time. Speedup can be improved by scaling the *numbrowscomptonce* with the number of

processors, but this requires increasing the memory of the system with the number of processors as well. Although the computer system used in this research has significantly more memory than other compute nodes, it was decided that the UU2DVar would be tested using *numpbrowscompatonce* corresponding to a moderately powered compute node. Although the speedup of the UU2DVar's computation of $\mathbf{P}_b\mathbf{H}^T$ and $\mathbf{H}\mathbf{P}_b$ is not perfect, the parallel implementation of this code still scales reasonably well and allows the entire tool to be run within real-time constraints. Computation of a single set of $\mathbf{P}_b\mathbf{H}^T$ and $\mathbf{H}\mathbf{P}_b$ arrays takes approximately 4 minutes on 8 processor cores for approximately ~15000 observations.

Background, Terrain, and Land/Water Mask

The UU2DVar is designed to use the background fields, topography, and land/water mask of the RTMA. This background field consists of the 12-km resolution Rapid Update Cycle (RUC; Benjamin et al. 2004) 1 h forecast downscaled to either 2.5-km or 5-km resolution, depending on the resolution of the analysis. The downscaling process of the RUC background field attempts to modify the meteorological fields based on differences between the 12-km and the 2.5-km or 5-km terrain; this process is fully described by Benjamin et al. (2007) and Jascourt (2007). The terrain field used in this research for computing the background error covariance is modified from its original format; the elevation of gridpoints that are classified as water points as specified by the land/water mask is lowered 500 m for the 2-m temperature and 2-m dewpoint background error covariance computation. This technique is similar to that used by the

RTMA. The terrain field and land/water mask of the entire domain used in this study is shown in Figure 3.6.

Input of the background fields is accomplished through a Network Common Data Form (NetCDF; Rew and Davis 1990) interface within MATLAB. The background fields can also be retrieved from the University of Utah server running Thematic Realtime Environmental Distributed Data Services (THREDDS; Caron et al. 2006) Server. Although CONUS domain background fields have been used in this research, the UU2DVar has also been configured to use background fields over an Austrian domain for comparison to the INCA system (Haiden et al. 2011).

Usage of Observations and Quality Control

Observations of 2-m temperature and 10-m u and v winds are used by the UU2DVar without any additional pre-processing and are assimilated in terms of their metric units ($^{\circ}\text{C}$ for temperature, m/s for winds). Pressure observations are assimilated in mb; however, the UU2DVar can either assimilate the raw observation or apply an elevation correction term, as is sometimes necessary when there are large differences between the observation elevation and the analysis gridpoint elevation. This pressure correction modifies an individual raw surface pressure observation to a corrected surface pressure (y_o^{pc}) by using the hypsometric equation:

$$y_o^{pc} = y_o^{ps} \exp\left(\frac{g}{R \cdot y_o^t}\right) (y_o^z - x_b^z) \quad (3.4)$$

where R is the ideal gas constant for dry air, g is the constant of gravitational acceleration, y_o^t and y_o^z correspond to the 2-m temperature and elevation of the pressure observation respectively, and x_b^z is the elevation of the nearest analysis gridpoint to the observation.

Since moisture is analyzed in terms of dewpoint temperature, mixing ratio values provided by some sources are converted to relative humidity to be consistent with the majority of mesonet observations available in terms of relative humidity. Because surface pressure is not available for all reports and to provide a consistent conversion from relative humidity to dewpoint temperature, relative humidity observations (y_o^{rh}) are converted to dewpoint temperature observations (y_o^{td}) using an empirical formula:

$$y_o^{td} = y_o^{rh \frac{1}{8}}(112^\circ\text{C} + 0.9y_o^t) + 0.1y_o^t - 112^\circ\text{C} \quad (3.5)$$

The UU2DVar can be configured to use one of three different observation sources: (1) the MesoWest database (Horel et al. 2002), (2) the observation data file used by the RTMA, or (3) a flat file generated by the user. Observations acquired using the UU2DVar's MesoWest interface must fall within a ± 30 min time window centered about the analysis hour. The UU2DVar uses the time window used by the RTMA for observations acquired from the RTMA's observations data files. The UU2DVar does not use a time window for the usage of an observation flat file; therefore, the time window used by the analysis is configurable by the user in this instance. For all three of these configuration options, only one observation is used per station. In the case of stations that record several observations within the time window, only the observation closest to the

analysis hour is used. If two observations from a station are separated by an equal amount of time about the analysis hour, the later observation is used.

All observations also undergo quality control during the assimilation cycle within the UU2DVar. Temperature, dewpoint, and pressure observations undergo a simple quality control check which rejects any observations that fail to meet the criteria

$$|y_o - x_b| \leq \varepsilon_m \text{stdev}(x_b) \quad (3.6)$$

where ε_m refers to a tunable error multiplier factor and the function $\text{stdev}(x_b)$ is the standard deviation of the background field over the entire domain (note that as in Equation 3.4, x_b is the value of a single background value nearest to the observation). Although this quality control may be rudimentary, it is effective in removing gross errors from the observation dataset. The error multiplier ε_m is set to 3 for temperature, dewpoint, and wind speed in this research.

An additional quality control step for wind observations is also available for UU2DVar. Wind observations still must meet the requirements as specified by Equation 3.6 (u and v wind components and wind speed must satisfy the criteria of Equation 3.6, or the entire observation is rejected), but additional light wind observations can be rejected if

$$y_o^{ws} < \tau_o^{ws} \cup x_b^{ws} > \tau_b^{ws} \quad (3.7)$$

where y_o^{ws} is the observation wind speed, x_b^{ws} is the value of the nearest background gridpoint to the observation, and τ_o^{ws} and τ_b^{ws} are wind speed observation and background quality thresholds, respectively. This additional quality control check helps to identify erroneously calm winds where the background field is specifying stronger winds. Although the asymmetric wind quality control was not used in the research presented here, it is mentioned here to present a complete description of the UU2DVar.

A minimum and maximum threshold quality control can also be applied to pressure observations within the UU2DVar. When this quality control is used, pressure observations that are used by the analysis must meet the criteria

$$\tau_{min}^{ps} \leq y_o^{ps} \leq \tau_{max}^{ps} \quad (3.8)$$

where τ_{min}^{ps} and τ_{max}^{ps} are minimum and maximum surface pressure thresholds (as specified by the user), respectively, and y_o^{ps} is an observation's surface pressure. This quality control was added after evaluation of several pressure analyses, as the quality control specified by Equation 3.6 fails to remove many unphysical surface pressure observations. Surface pressure analyses are not studied within this research, but this particular quality control is mentioned here as well to present a complete description of the UU2DVar.

A simple forward operator is used to interpret analysis values to observation locations, as well as a simple observation error covariance to assimilate the observations. The forward operator, \mathbf{H} , interprets analysis values using a nearest neighbor approach. As mentioned earlier in this chapter, the simplicity of this forward operator allows it to be

exploited as a filter to reduce the storage required for the background error covariance matrix. The observation error covariance matrix, \mathbf{P}_o , is simply a diagonal matrix, and in this research, the observation error variance is set to 1°C for temperature and dewpoint and to 1 m/s for u and v wind components and wind speed. Setting all off diagonal elements of \mathbf{P}_o to 0 assumes that all observation errors are uncorrelated with each other; however, this may not be true for all observations. Errors of observations within individual mesonets may actually be correlated with each other through siting or instrumentation biases of a particular mesonet. As will be discussed in Chapter 6, it is possible to use the methodology presented in Chapter 4 to help determine mesonet biases on the basis of large samples of analyses. Using the same value for all diagonal elements of observation error covariance matrix also implies that all observations have equal observation errors, which may also not be accurate. A particular mesonet may have significantly higher observations errors than others; the same may also apply for an individual observation when compared to other observations within the mesonet. While it is straightforward to implement varying observation errors dependent upon the mesonet or the individual observation in UU2DVar, the tuning of \mathbf{P}_o requires extensive research and additional testing that was beyond the scope of this study.

UU2DVar General Characteristics

and Analysis Cycle

As a result of the simple specification of the background error covariance by Equation 3.1, the UU2DVar produces univariate surface analyses. Because wind is a vector measurement, it was believed assimilating u and v wind components separately

would yield analyses that would be less accurate than analyses generated using a multivariate method. This approach involves generating u and v unit wind component analysis fields, along with analyses of wind speed. The vector wind field is produced by multiplying the unit u and v wind components analyses by the wind speed analysis. In the case of the wind speed analysis, the rare negative values within the analysis are set to 0 after the analysis has been computed.

The approximations and parallelization techniques listed in this chapter allow the UU2DVar to generate 2-m temperature, 2-m dewpoint, surface pressure, and 10-m u and v wind component analyses in approximately 25 minutes when run on a compute node of 8 processors. This is comparable to the computation cycle of the RTMA, which needs approximately 15 minutes when run on 16 processors on the NCEP development supercomputer (M. de Pondeca, personal communication). Computing the observation sensitivity and observation impact (which is defined in the next chapter) for all fields adds an additional 8 minutes to the UU2DVar's computation cycle. The complete computation cycle, along with the parallelization scheme across the entire cycle, is depicted in Figure 3.7.

While many data assimilation tools and systems, such as the RTMA, are written in FORTRAN, the UU2DVar is written in MATLAB. The MATLAB programming language offers several advantages over FORTRAN compilers commonly supported at research universities. MATLAB has built-in support for sparse matrix mathematics and optimizations through the Linear Algebra Package (LAPACK; Anderson et al. 1999) and the Basic Linear Algebra Subprograms (BLAS; Dongarra et al. 1988). Relative to the commonly used Message Passing Interface (MPI), MATLAB's parallel computing

implementation is also easier to use since MATLAB does not require a developer to explicitly code communication and work distribution between processors. The MATLAB programming language also offers intrinsic subroutines essential for solving the variational assimilation equation, including an efficient GMRES function used to solve Equation 2.18. The cross-platform compiling abilities of MATLAB allowed the UU2DVar to be developed and tested on two different operating systems (Windows and Linux) with very little additional development work. The ability to compile the MATLAB code using the freely available MATLAB Compiler Runtime also makes the UU2DVar available on systems without MATLAB licenses. MATLAB also offers built-in support of the standard NetCDF file format, which allows efficient input/output of the background and analysis. Although there were initially some reservations regarding computational overhead requirements of MATLAB, this research has demonstrated those reservations to be unfounded as the UU2DVar's wall clock time is comparable to the FORTRAN-based RTMA.

Figure 3.8 contrasts the analysis increments (analysis minus background) for 2-m air temperature at 1400 UTC 27 October 2010 for the UU2DVar (Figure 3.8a) and the RTMA (Figure 3.8b). For this example, the decorrelation length scales used by the UU2DVar have been set to match the RTMA's equivalent decorrelation length scales along with using the same observation dataset as the RTMA. Figure 3.8c shows the difference field between the RTMA and the UU2DVar analysis increments. Many of the differences between the two analyses are in regions of orography, i.e., along the Sierra-Nevada, Rocky, and Appalachian mountain ranges. While differences in the quality control procedures may cause some differences in the increments, the majority of the

differences are due to the coarse computation grid used by the recursive filters within the RTMA as opposed to the approach used by the UU2DVar. The coarse computation grid requires specifying a smoothed terrain for the background error covariance (M. de Pondeva, personal communication), which causes many of the minor terrain features to have limited impact on the computation of analysis increments within the RTMA.

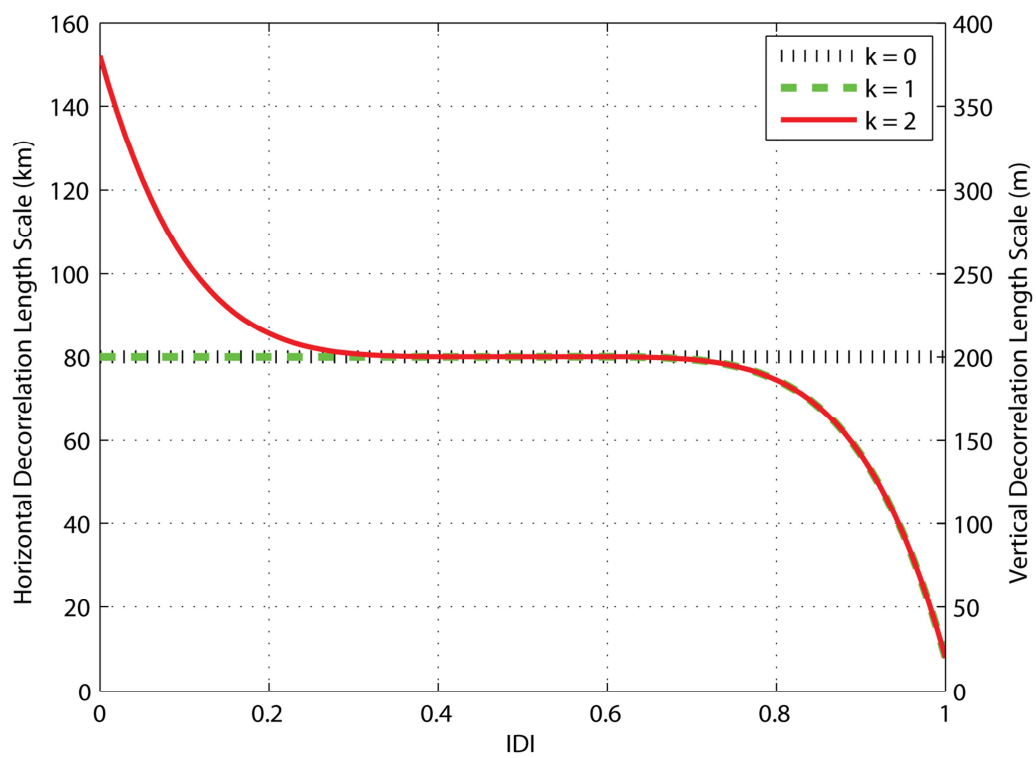


Figure 3.1. Horizontal (\mathcal{R}_k) and vertical (\mathcal{Z}_k) decorrelation length scales specified by Equations 3.2 and 3.3.

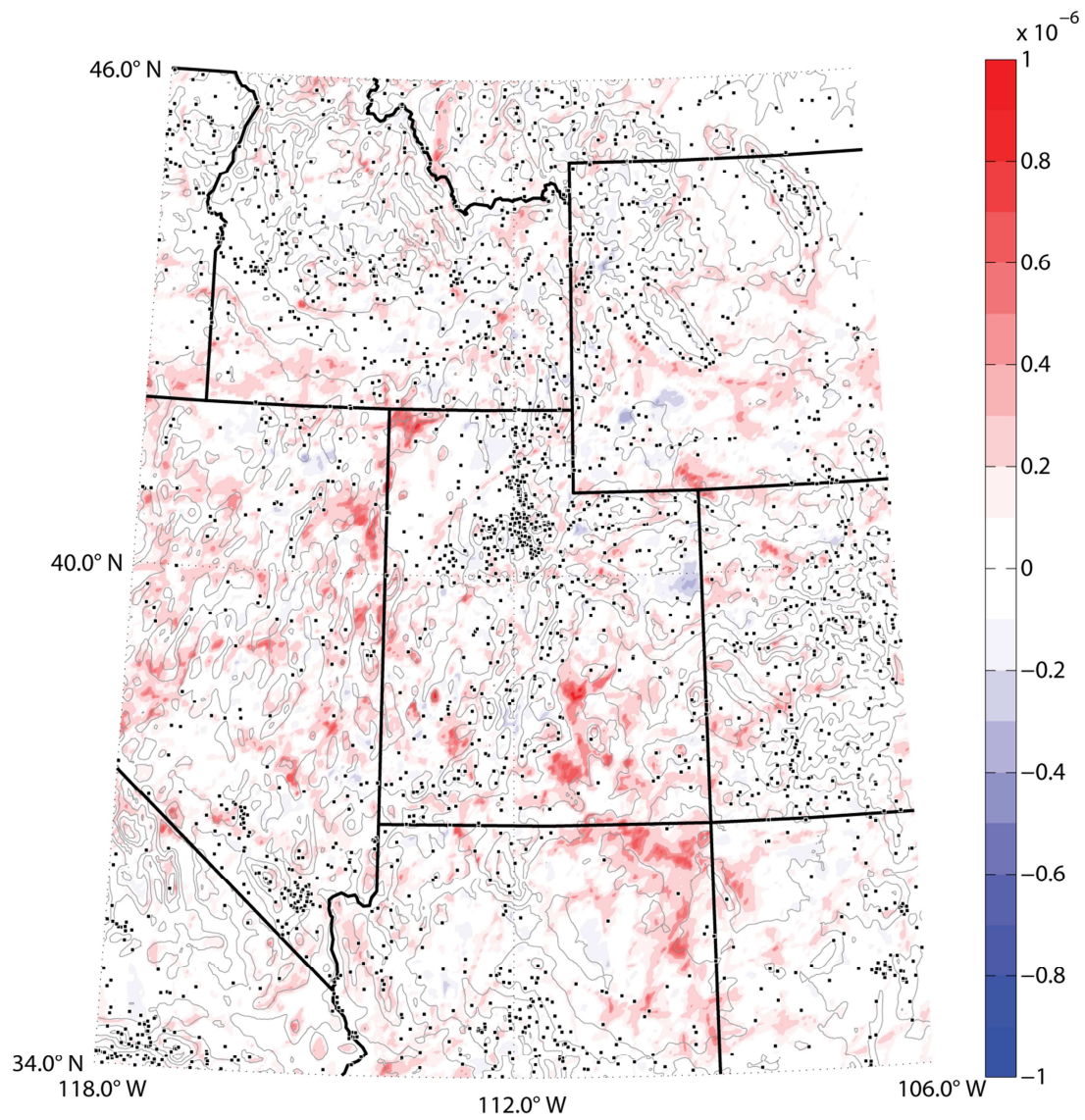


Figure 3.2. Difference (shaded) of IDI fields of temperature observations without and with localization. Grey contours denote 500 m terrain contours and black squares denote locations of temperature observations.

$$\begin{array}{c}
 \left[\begin{array}{cccccc}
 \sigma_b & A & F & J & M & O \\
 A & \sigma_b & B & G & K & N \\
 F & B & \sigma_b & C & H & L \\
 J & G & C & \sigma_b & D & I \\
 M & K & H & D & \sigma_b & E \\
 O & N & L & I & E & \sigma_b
 \end{array} \right] \times \begin{array}{c} \left[\begin{array}{cc}
 1 & 0 \\
 0 & 0 \\
 0 & 0 \\
 0 & 0 \\
 0 & 1 \\
 0 & 0
 \end{array} \right] \\
 \mathbf{H}^T
 \end{array} = \begin{array}{c} \left[\begin{array}{cc}
 \mathbf{P}_b \mathbf{H}^T_{1,1} & \mathbf{P}_b \mathbf{H}^T_{1,2} \\
 \mathbf{P}_b \mathbf{H}^T_{2,1} & \mathbf{P}_b \mathbf{H}^T_{2,2} \\
 \mathbf{P}_b \mathbf{H}^T_{3,1} & \mathbf{P}_b \mathbf{H}^T_{3,2} \\
 \mathbf{P}_b \mathbf{H}^T_{4,1} & \mathbf{P}_b \mathbf{H}^T_{4,2} \\
 \mathbf{P}_b \mathbf{H}^T_{5,1} & \mathbf{P}_b \mathbf{H}^T_{5,2} \\
 \mathbf{P}_b \mathbf{H}^T_{6,1} & \mathbf{P}_b \mathbf{H}^T_{6,2}
 \end{array} \right]
 \end{array}
 \end{array}$$

$$\begin{array}{ll}
 \mathbf{P}_b \mathbf{H}^T_{1,1} = \sigma_b & \mathbf{P}_b \mathbf{H}^T_{1,2} = M \\
 \mathbf{P}_b \mathbf{H}^T_{2,1} = A & \mathbf{P}_b \mathbf{H}^T_{2,2} = K \\
 \mathbf{P}_b \mathbf{H}^T_{3,1} = F & \mathbf{P}_b \mathbf{H}^T_{3,2} = H \\
 \mathbf{P}_b \mathbf{H}^T_{4,1} = J & \mathbf{P}_b \mathbf{H}^T_{4,2} = D \\
 \mathbf{P}_b \mathbf{H}^T_{5,1} = M & \mathbf{P}_b \mathbf{H}^T_{5,2} = \sigma_b \\
 \mathbf{P}_b \mathbf{H}^T_{6,1} = O & \mathbf{P}_b \mathbf{H}^T_{6,2} = E
 \end{array}$$

Figure 3.3. Pictorial demonstration of exploitation of the forward operator using the example presented in Figure 2.1. Only a fraction of the columns of \mathbf{P}_b (shaded) actually need to be calculated to compute $\mathbf{P}_b \mathbf{H}^T$; these columns correspond to nonzero rows of \mathbf{H}^T .

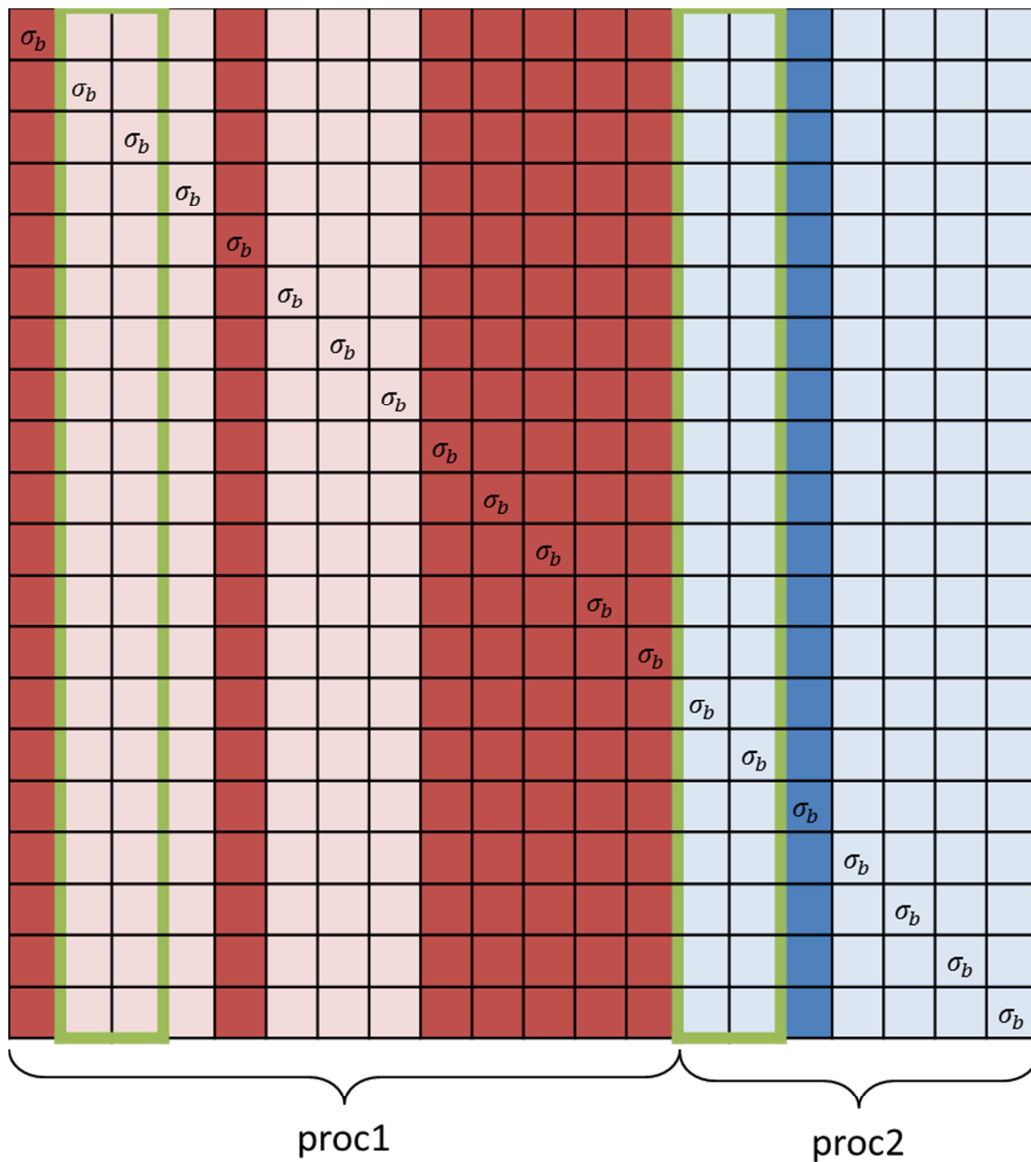


Figure 3.4. Slicing of background error covariance matrix for computations and storage, using a domain of 20 gridpoints on 2 processors with *numbrowscompatonce* set to 4. Distribution of the work to the processors is based on an equal division of the number of columns that must be computed (light red and light blue), and not on a division of the entire covariance matrix itself. Dark red and dark blue columns of the background error covariance are not computed. Each processor works on computing two columns at a time (outlined in green) using full matrices before converting to a sparse matrix type.

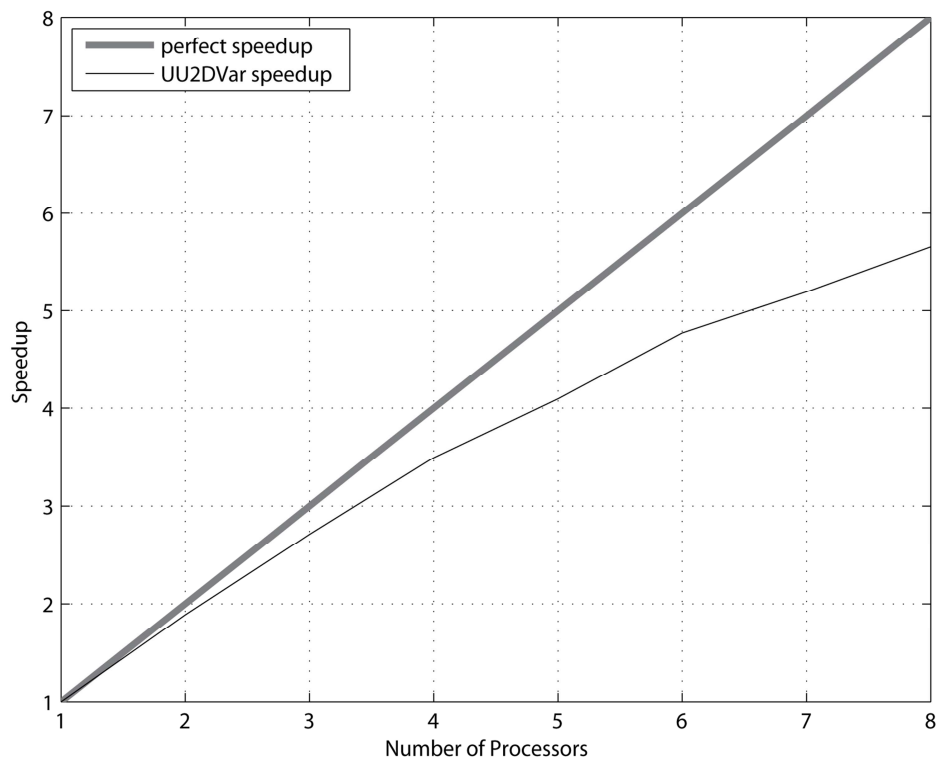


Figure 3.5. Speedup of background error covariance computation (thin black line), plotted against perfect speedup (thick grey line).

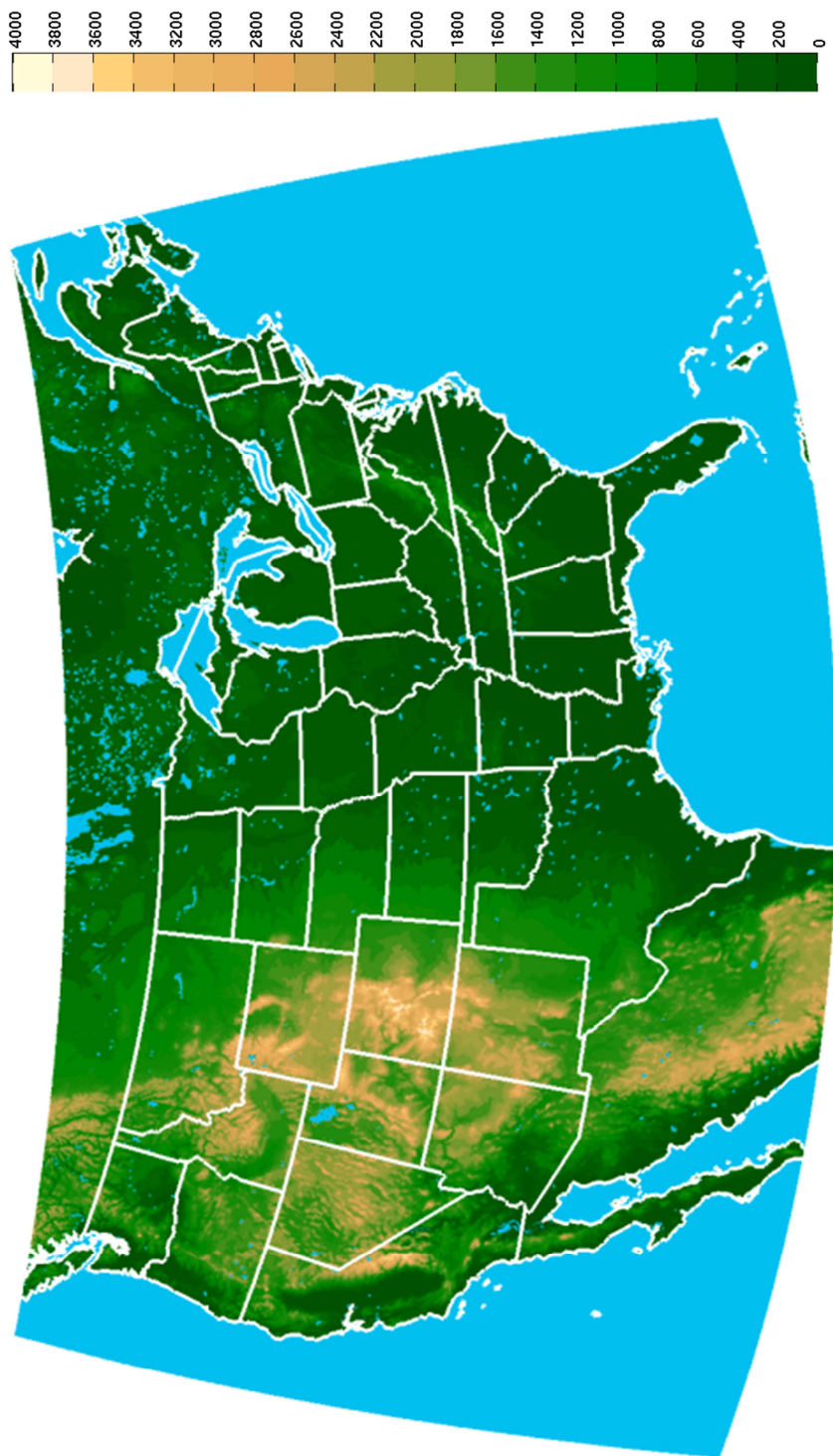


Figure 3.6. CONUS 5-km terrain and land/water mask used by UU2DVar in this study. Gridpoints in blue denote water grid points.

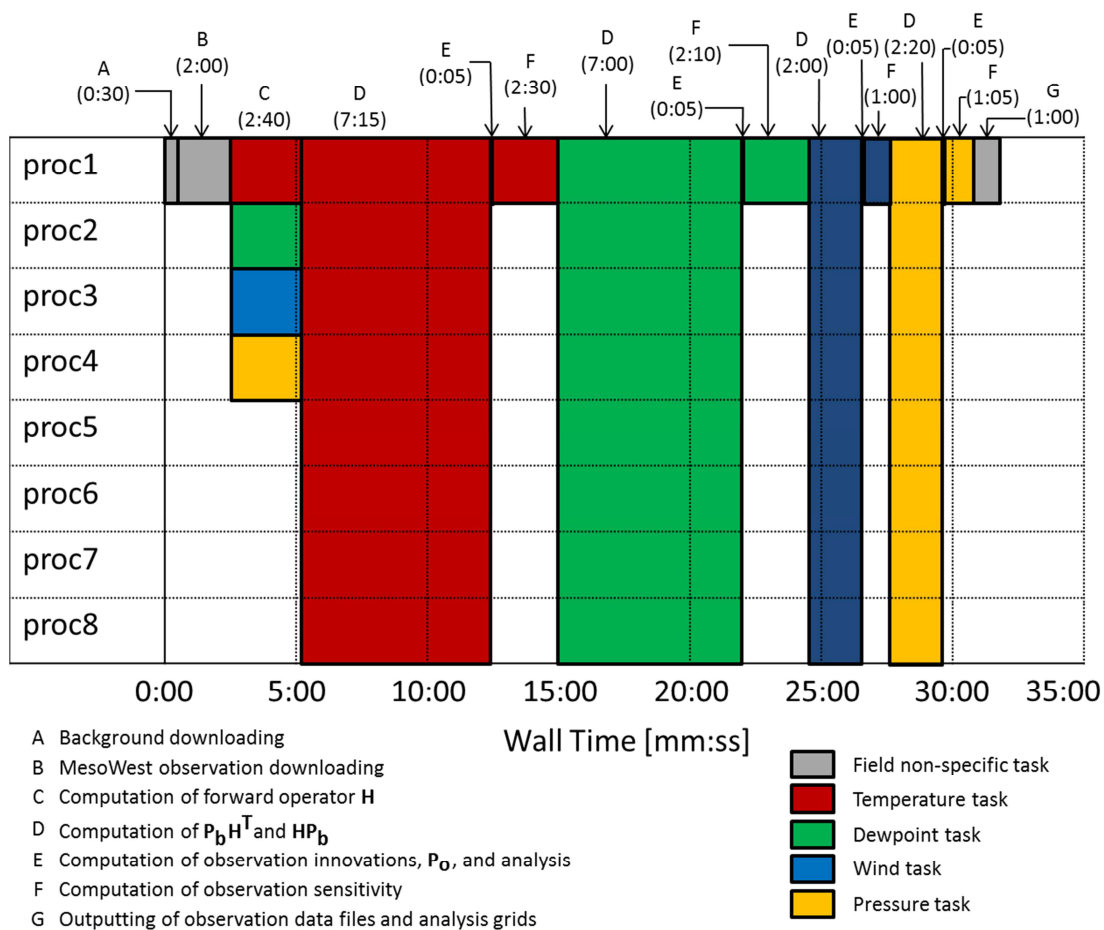


Figure 3.7. Computation cycle and parallelization within UU2DVar, computing temperature, dewpoint, wind, and surface pressure 5-km resolution analyses for the CONUS domain using 8 processors using the MesoWest observation dataset valid for 1400 UTC 27 October 2010. Times for each subtask were computed from averages of 10 trials, rounded to the nearest 5 s.

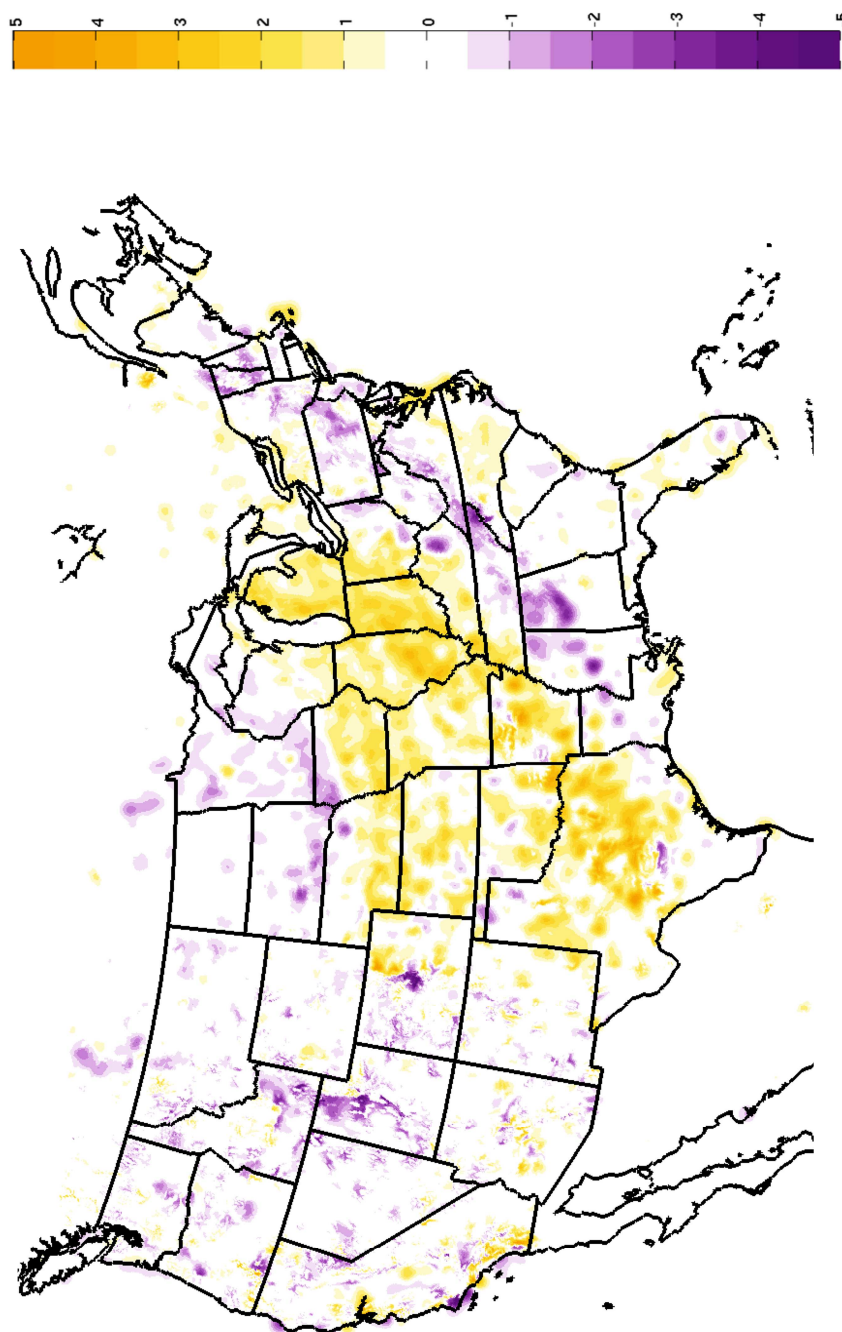


Figure 3.8. UU2DVar and RTMA 2-m air temperature analysis increments, and the difference between the two analysis increments for 1400 UTC 27 October 2010. a. UU2DVar analysis increments, using decorrelation length scales approximately equal to those used by the RTMA. Analysis increments shaded in $^{\circ}\text{C}$.

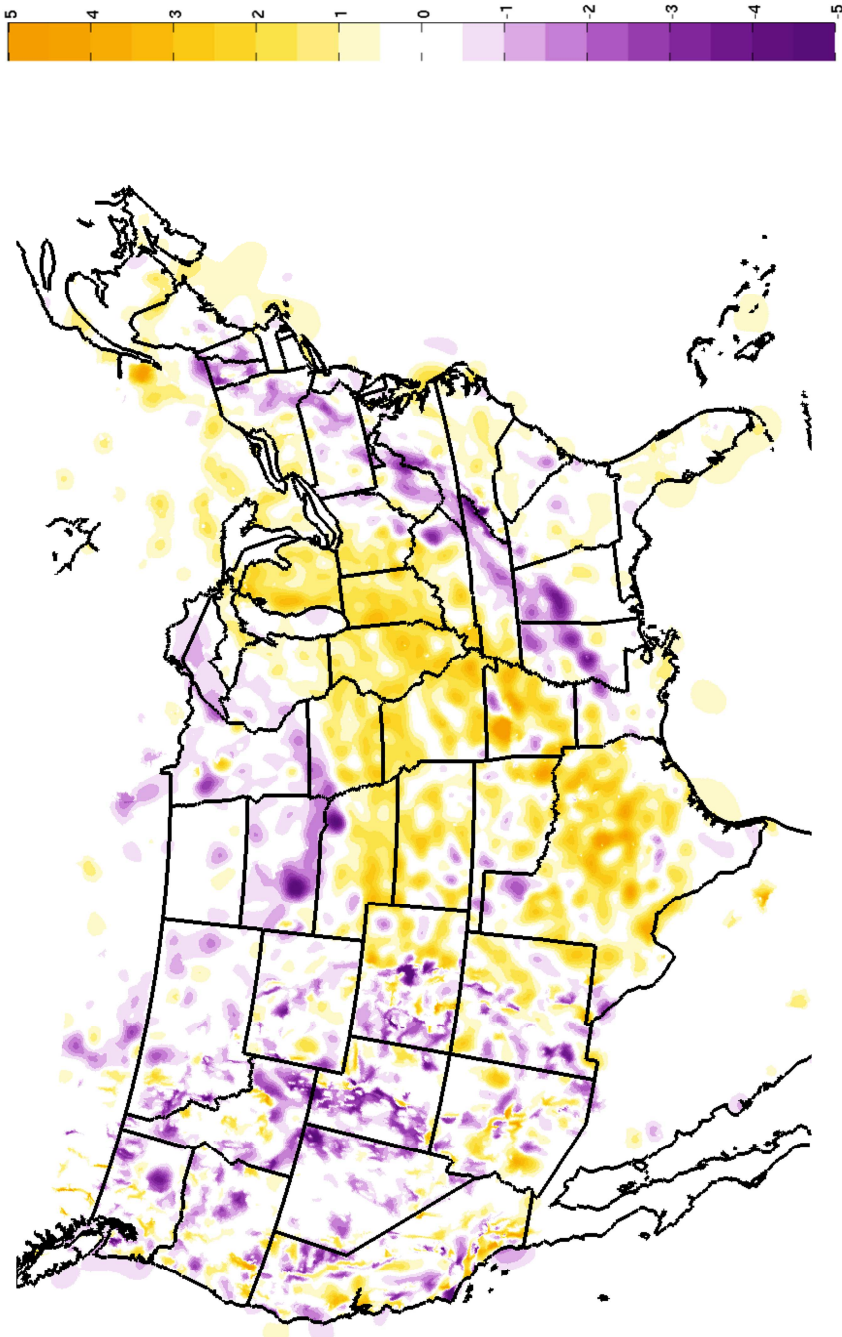


Figure 3.8. continued. b. RTMA analysis increments in °C.

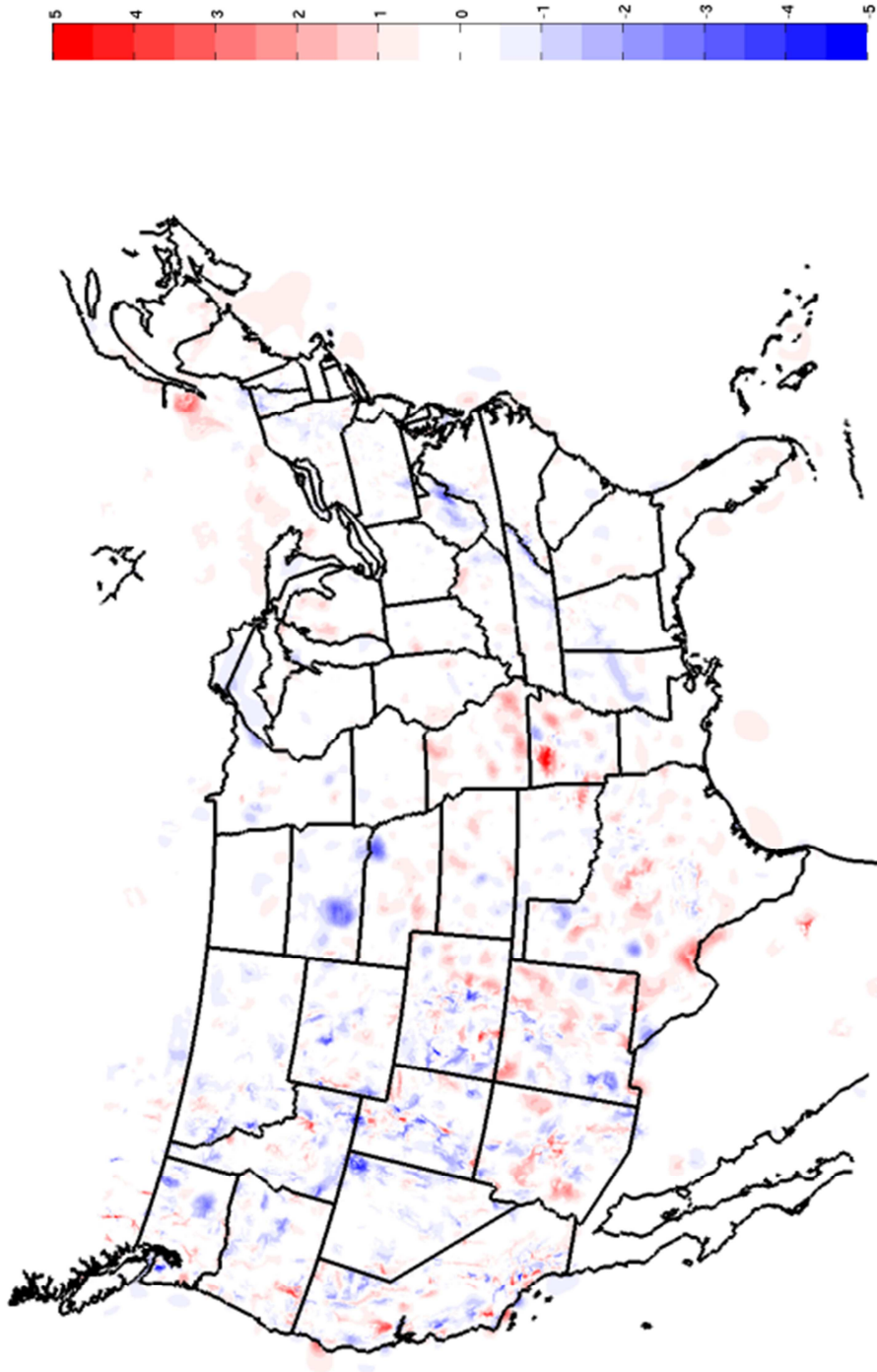


Figure 3.8. continued. c. Difference between RTMA and UU2DVar analysis increments in $^{\circ}\text{C}$.

CHAPTER 4

OBSERVATION IMPACTS AND SENSITIVITY

Introduction

The National Research Council (2009) discussed the need for assessing the extent to which observations collected from disparate sources can be used to meet a variety of needs. Building on prior work (Myrick and Horel 2008; Horel and Dong 2010; Tyndall et al. 2010), this study is aimed at providing a better foundation for addressing the sensitivity and impacts of observations as a function of observation source. The objective of this chapter is to describe an appropriate approach and the data sources that will be evaluated in the next chapter.

Computation

Withholding a subset of observations from analyses and comparing the differences between the withheld observations and the resulting analyses is commonly used to assess analysis accuracy and uncertainty as well as the impacts of different types of observations on the analysis (Seaman and Hutchinson 1985; Zapotocny et al. 2000; Hiemstra et al. 2006; Myrick and Horel 2008; Tyndall 2008; Horel and Dong 2010; Tyndall et al. 2010). The choice of which observations to withhold depends on the application. For example, Zapotocny et al. (2000) differentiated by instrumentation type,

Hiemstra et al. (2006) used observations from networks generally not used operationally in their data assimilation system, while Myrick and Horel (2008) randomly withheld 30% of all surface observations. Horel and Dong (2010) applied the most extensive approach by sequentially withholding each of ~3000 observations from each of ~9000 analyses resulting in over 500,000 cross-validation experiments. This leave-one-out cross validation approach (Wilks 2006) is far too computationally expensive to be used for real-time applications.

This research utilizes the analysis adjoint to efficiently compute analysis sensitivity to observations without the need to perform cross validation experiments. Adjoint of forecast models are now used routinely to assess where “targeted” observations might reduce model forecast errors (Palmer et al. 1998; Buizza and Montani 1999; Langland et al. 1999). Following the derivation by Baker and Daley (2000), Equations 2.18 and 2.19, which are used to compute the analyses within the UU2DVar, can be combined into a single equation (4.1):

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{P}_b \mathbf{H}^T (\mathbf{H} \mathbf{P}_b \mathbf{H}^T + \mathbf{P}_o)^{-1} [\mathbf{y}_o - \mathbf{H}(\mathbf{x}_b)] \quad (4.1)$$

The adjoint of an analysis system can be viewed as the adjustment to the background field by the observations necessary to yield the resulting analysis, i.e., solving for the right-most term in Equation 4.1 (Kalnay 2003). The sensitivity of the analysis to the observations is calculated from the derivative of Equation 4.1 with respect to the observation vector and through use of the chain rule:

$$\frac{\partial \mathbf{x}_a}{\partial \mathbf{y}_o} = \left[\mathbf{P}_b \mathbf{H}^T (\mathbf{H} \mathbf{P}_b \mathbf{H}^T + \mathbf{P}_o)^{-1} \right]^T = \mathbf{K}^T \quad (4.2)$$

\mathbf{K}^T is the transpose of the weight matrix \mathbf{K} , which can be simplified using the distributive property of the transpose operator and the symmetry of the background and observation error covariance matrices to yield:

$$\mathbf{K}^T = (\mathbf{H} \mathbf{P}_b \mathbf{H}^T + \mathbf{P}_o) \mathbf{H} \mathbf{P}_b \quad (4.3)$$

A cost function \mathcal{J} (different from the cost function J in Equation 2.1 used as the foundation for variational analysis) is specified that is a measure of a quantity of interest within the analysis domain. Forecast model adjoint sensitivity studies may choose a parameter such as air temperature or sea level pressure over a limited domain of interest in order to highlight what additional targeted observations might help reduce the forecast error of that parameter in that region (Langland et al. 1999; Baker and Daley 2000; Zhu and Gelaro 2008). Since the objective of this research is to assess analysis sensitivity to differences between observation networks, \mathcal{J} is defined with respect to the entire analysis domain as the squared differences between the analysis and the background field:

$$\mathcal{J} = \frac{1}{2} (\mathbf{x}_a - \mathbf{x}_b)^2 \quad (4.4)$$

The observation sensitivity vector, $\partial \mathcal{J} / \partial \mathbf{y}_o$, is computed using Equation 4.2 and the chain rule:

$$\frac{\partial \mathcal{J}}{\partial \mathbf{y}_o} = \frac{\partial \mathbf{x}_a}{\partial \mathbf{y}_o} \frac{\partial \mathcal{J}}{\partial \mathbf{x}_a} = \mathbf{K}^T (\mathbf{x}_a - \mathbf{x}_b) \quad (4.5)$$

The observation sensitivity is defined at each observation location and does not depend on the specific values of observations at those locations (Baker and Daley 2000), i.e., it is a measure of the sensitivity of the analysis to having observations at that location, not the sensitivity to the specific observation reported at that location.

The observation impact, \mathcal{G} , as defined by Zhu and Gelaro (2008), considers the value of observations as well as their locations:

$$\mathcal{G} = \frac{1}{2} \left\langle \frac{\partial \mathcal{J}}{\partial \mathbf{y}_o}, \mathbf{y}_o - \mathbf{H}(\mathbf{x}_b) \right\rangle \quad (4.6)$$

Since \mathcal{G} is the scalar product of the observation sensitivity and the observation innovations, the contribution of specific observations to the analysis can be assessed.

Observation Networks

As a test of the methodology described in the previous subsection, observations are examined for 1400 UTC 27 October 2010. The analysis is restricted to observations accessible via MesoWest that are publicly available via the Meteorological Assimilation and Data Ingest System (MADIS; Miller et al. 2005). With permission of the data providers, observations available to MADIS that are subject to usage restrictions from the Oklahoma Mesonet and WeatherFlow Inc. were added for this analysis as they reflect networks with very good maintenance, equipment, and siting standards. A total of 13,763

temperature, 11,201 humidity, and 11,728 wind observations were available for this analysis

As discussed by the National Research Council (2009), metadata on the equipment, siting, and reporting standards used by many data providers are incomplete. The MesoWest developers identify mesonets by their source, which in the cases of large networks is often in turn an aggregate of many different sources. For the purposes of this study, the source networks are grouped into 10 general categories based subjectively on the characteristics of the networks known to the MesoWest developers. Figure 4.1 depicts the locations of each observation available at this specific time and grouped by mesonet and network category. Providing such information routinely is one of the recommendations of the National Research Council (2009). Horel et al. (2002) provide a description of observing networks provided by MesoWest, but because those descriptions may no longer be up to date and because many new networks have been added to MesoWest since the publication of that article, the networks are described here for completeness.

Figure 4.1a shows station locations that are classified as primarily agricultural (AG). The AG networks monitor standard meteorological parameters and often report soil temperature and moisture as well. Wind sensors are typically mounted at 3 m to facilitate surface evaporation estimates. The majority of this network is made up of observations from the Soil Climate Analysis Network (Schaefer and Paetzold 2001), the California Irrigation Management Information System (Snyder 1984), and the U.S. Bureau of Reclamation AgriMet (U.S. Bureau of Reclamation 2011). Most of the

observations within this category are located within well irrigated areas and are collected in real-time.

Air quality monitoring networks are aggregated into the AQ network category, as shown in Figure 4.1b. The AIRNow network combines air quality stations from many state and local agencies. MesoWest and MADIS continue to access selected air quality networks directly that do not provide their complete suite of weather data to the AirNow program.

Figure 4.1c depicts stations that are primarily external (EXT network category) to the contiguous United States. This category includes observations from Environment Canada, which is the Canadian equivalent of the United States' NWS and Federal Aviation Administration (FAA) observing network. As with many networks, there are increasing concerns about the quality of these observations due to siting and quality control issues (Environment Canada 2008). Observations from Mexico are provided by the Servicio Meteorológico Nacional (SMN) de México. The SMN network is a synoptic-scale observation network. The limited available documentation suggests there are quality control and reliability issues also associated with this network (Servicio Meteorológico Nacional 2011). Observations from network providers that are primarily located along the coast or offshore (with a few interior exceptions from the commercial WeatherFlow network) are also included in this category with the majority provided by the National Oceanic and Atmospheric Administration.

The FED network category consists of land based observations from federal agencies, excluding observations maintained by the NWS/FAA and those used primarily for agricultural, hydrological, fire weather, or air quality purposes. As shown in Figure

4.1d, most of these observations come from local and regional networks, except for the nationwide MADIS Non-commissioned Automated Weather Observing System and the Climate Reference Network. Generally all of these networks have well-defined siting and maintenance procedures, but it should be recognized that even the highest quality networks can have standards different from what many users might expect. For example, wind direction is not available from the Climate Reference Network since the wind speed sensor at 3 m is intended to estimate under catch of precipitation (National Environmental Satellite, Data and Information Service 2003).

Hydrometeorological networks are grouped into the HYDRO network category (Figure 4.1e). The majority of observations are supplied by the Hydrometeorological Automated Data System (HADS), which is itself an aggregate of stations owned and maintained by many different agencies (Office of Hydrologic Development 2011). Many more observations in the HADS network report precipitation only and do not appear in this figure; however, the ones shown here report at least air temperature. The Snowpack Telemetry network observations (SNOTEL; Schaefer and Paetzold 2001) of the National Resources Conservation Service are a very important resource due to their locations generally at high elevation within the western United States. Due to the meteor burst data communication required for these remote locations, observations often are not available until a few hours after the valid time, and may not be available for the RTMA. In addition to precipitation measurements, all SNOTEL observations collect 2-m air temperature.

Stations supplied to MesoWest and MADIS by a number of local, state, and regional sources are aggregated into the LOCAL network category (Figure 4.1f). The largest network within this category is the Oklahoma Mesonet (Brock et al. 1995), which

has been described as the “most prominent state mesonet” in the country (National Research Council 2009) for its high quality instrumentation, siting practices, and documentation of observation metadata (Brock et al. 1995). The West Texas Mesonet (Schroeder et al. 2005) was modeled after the Oklahoma Mesonet and follows similar instrumentation and siting protocols. Some of the other LOCAL networks are aggregates of stations often including a mix of stations directly maintained by the network provider as well as stations managed by other local data providers. For example, there is considerable advantage to having WFOs work with local data providers to locally access their observations and then disseminate those observations to MesoWest and MADIS for other users.

Figure 4.1g depicts the NWS network category, which is composed of Automated Surface Observing Stations (National Oceanic and Atmospheric Administration 1998) and Automated Weather Observing Stations (Federal Aviation Administration 2011). The majority of these observations are located in the urban regions of the eastern United States. All stations within this network category are commissioned by the NWS and FAA (although AWOS observations are maintained by state or local agencies), and must meet ASOS or AWOS equipment and siting standards.

Observations from the Automatic Position Reporting System Weather Network/Citizen Weather Observing Program (APRSWXNET/CWOP; Gladstone 2000) make up the majority of the PUBLIC network category as shown in Figure 4.1h. APRSWXNET/CWOP stations are owned, installed and maintained typically by private citizens with varied siting and reporting practices (Chadwick 2005). Concerns over the quality of observations from the APRSWXNET/CWOP network have been raised in the

past, especially due to representativeness errors associated with siting issues (Tyndall et al. 2010). While Tyndall et al. showed that temperature observations appeared to be of similar quality to observations from the NWS network, concerns remain regarding wind measurements due to the frequent occurrence of nearby obstructions. The asymmetric wind quality control described in Chapter 3 was developed to mitigate some of these issues.

The Remote Automated Weather Stations (RAWS) network (Figure 4.1i) is designed for wildfire management applications and supported by the U.S. Forest Service and many other federal, state, and local land and wildfire management agencies (Zachairassen et al. 2003; Horel and Dong 2010). RAWS are often located in remote locations preferably on slight south-facing slopes with limited nearby vegetation. Wind sensors are located at 6 m instead of 10 m anemometer heights used by their NWS observation counterparts. The lower anemometer heights and 10 minute averaging interval have contributed to the perception that their wind speeds are less than what might be expected leading to many RAWS being excluded from the RTMA.

Finally, Figure 4.1j depicts the collective availability of stations located adjacent to roads and railways (TRANS network category). Many Union Pacific Railroad stations report 2-m air temperature only, as their primary interest for deploying the equipment is related to monitoring the expansion and contraction of the rails. Road Weather Information Systems (RWIS) generally report all standard meteorological variables as well as other measurements from additional road sensors.

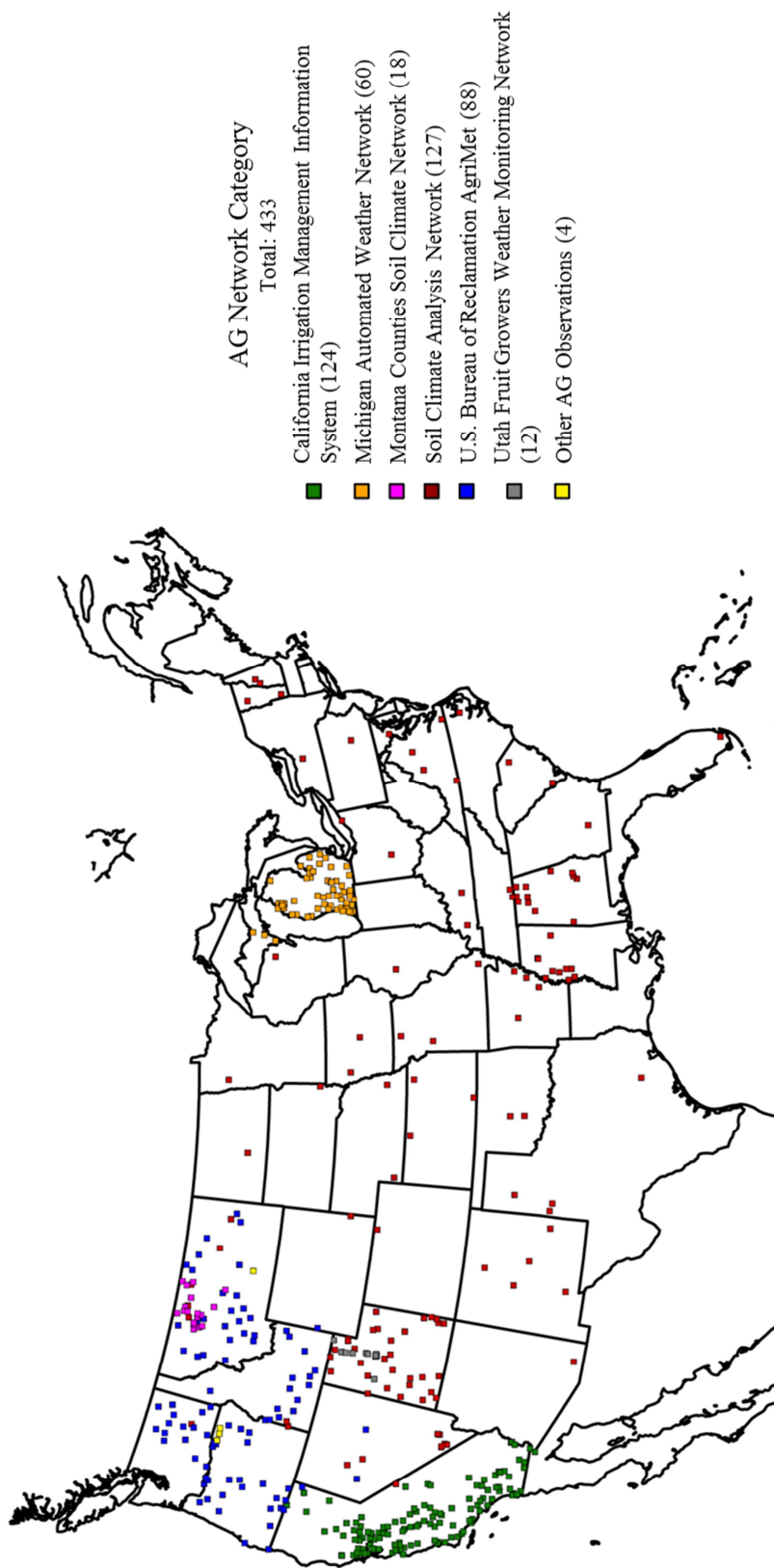


Figure 4.1. Observations available for 1400 UTC 27 October 2010 analyses for each of the 10 network categories used in this research. Quantities in parentheses denote the number of observation each individual network contributed. a. AG network category.

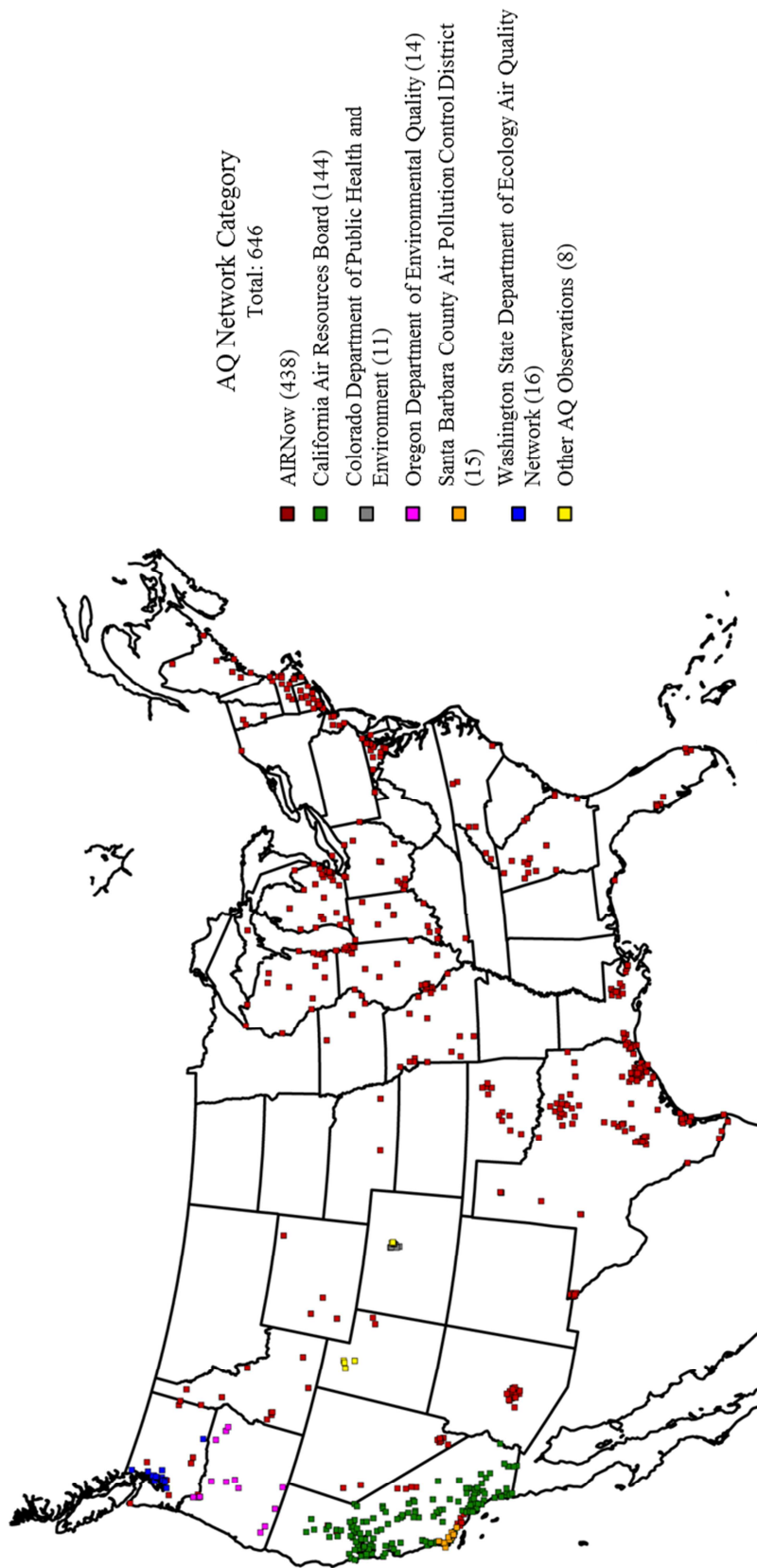


Figure 4.1. continued. b. AQ network category.

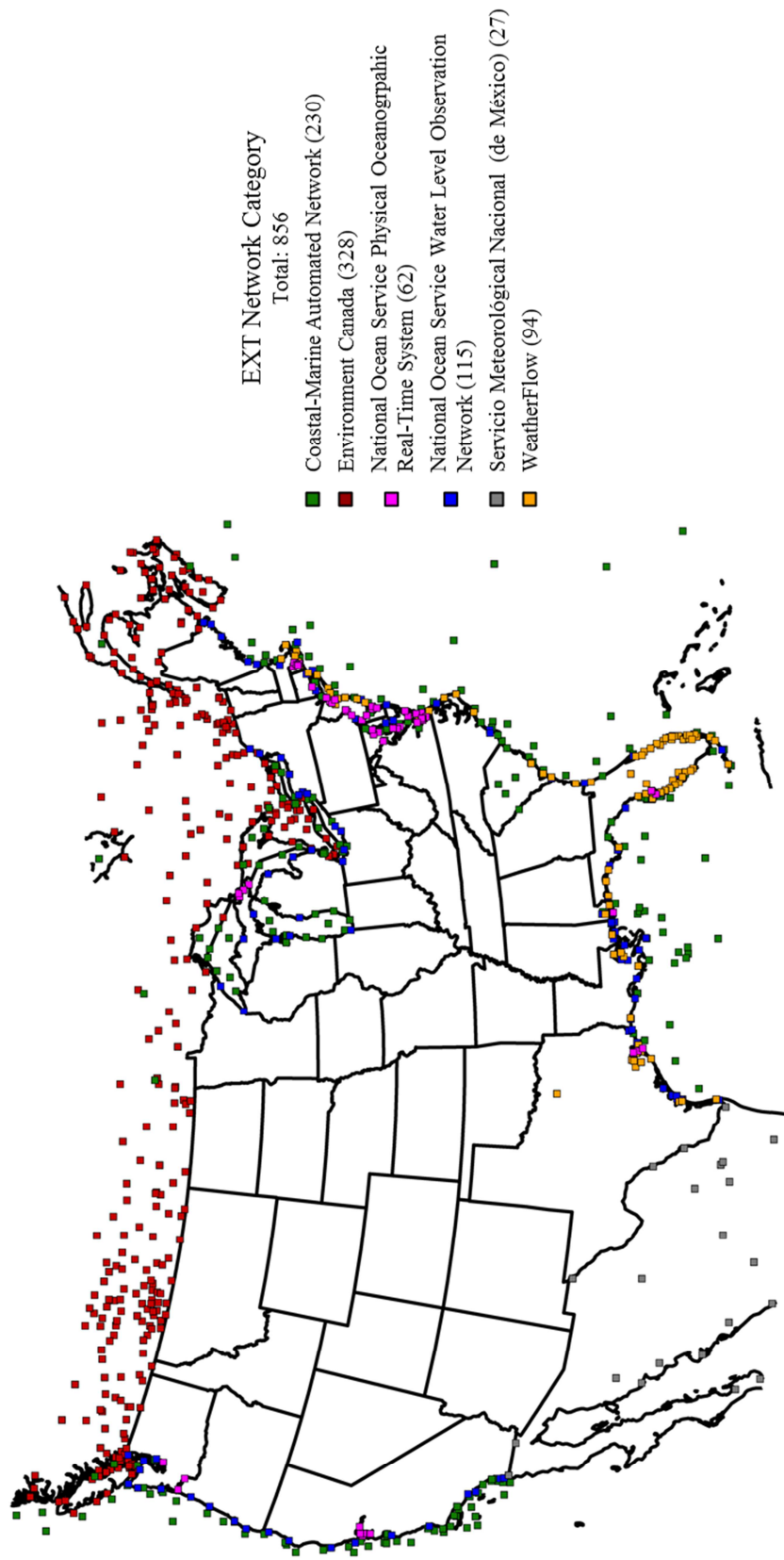


Figure 4.1. continued. c. EXT network category.

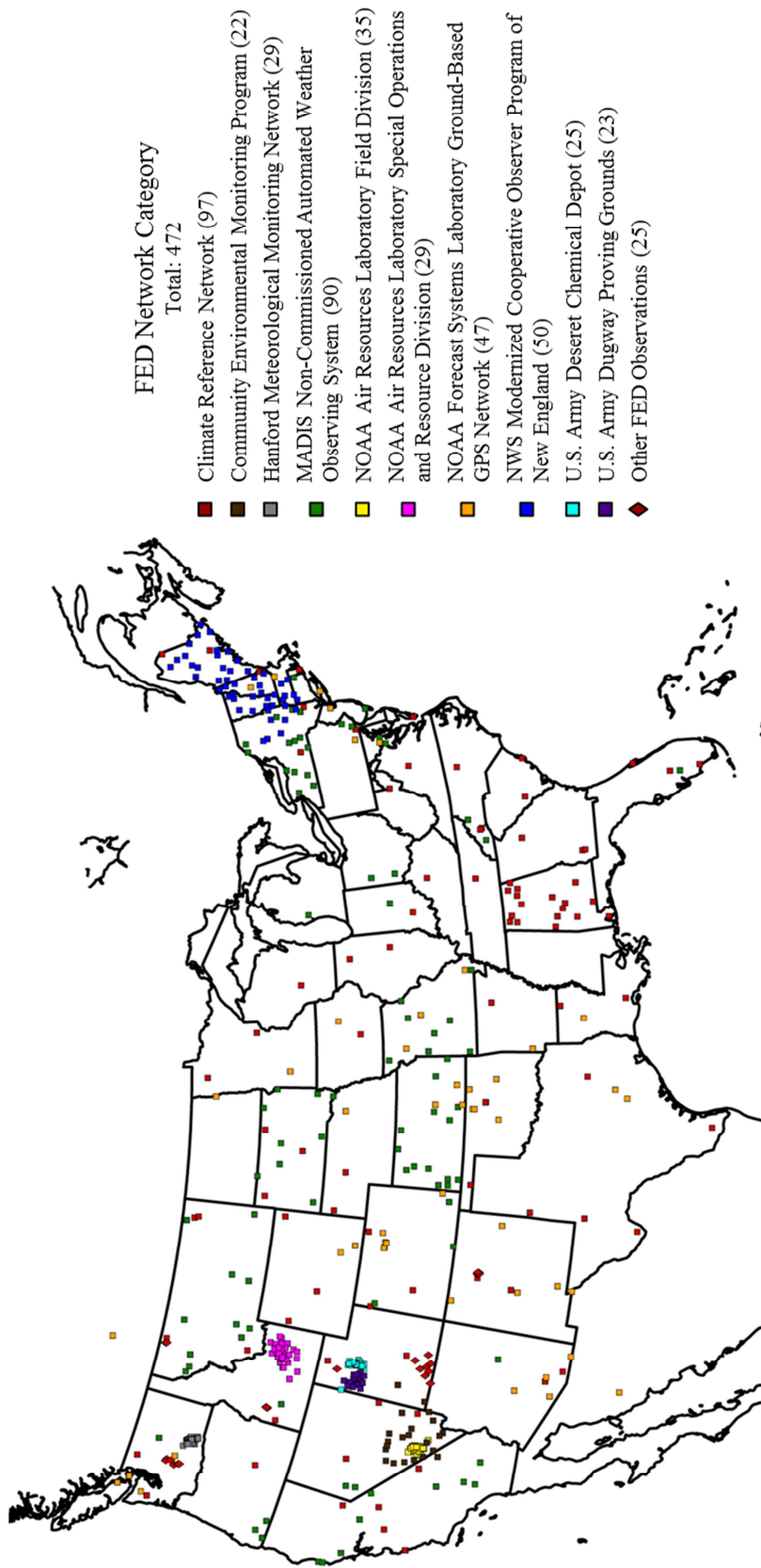


Figure 4.1. continued. d. FED network category.

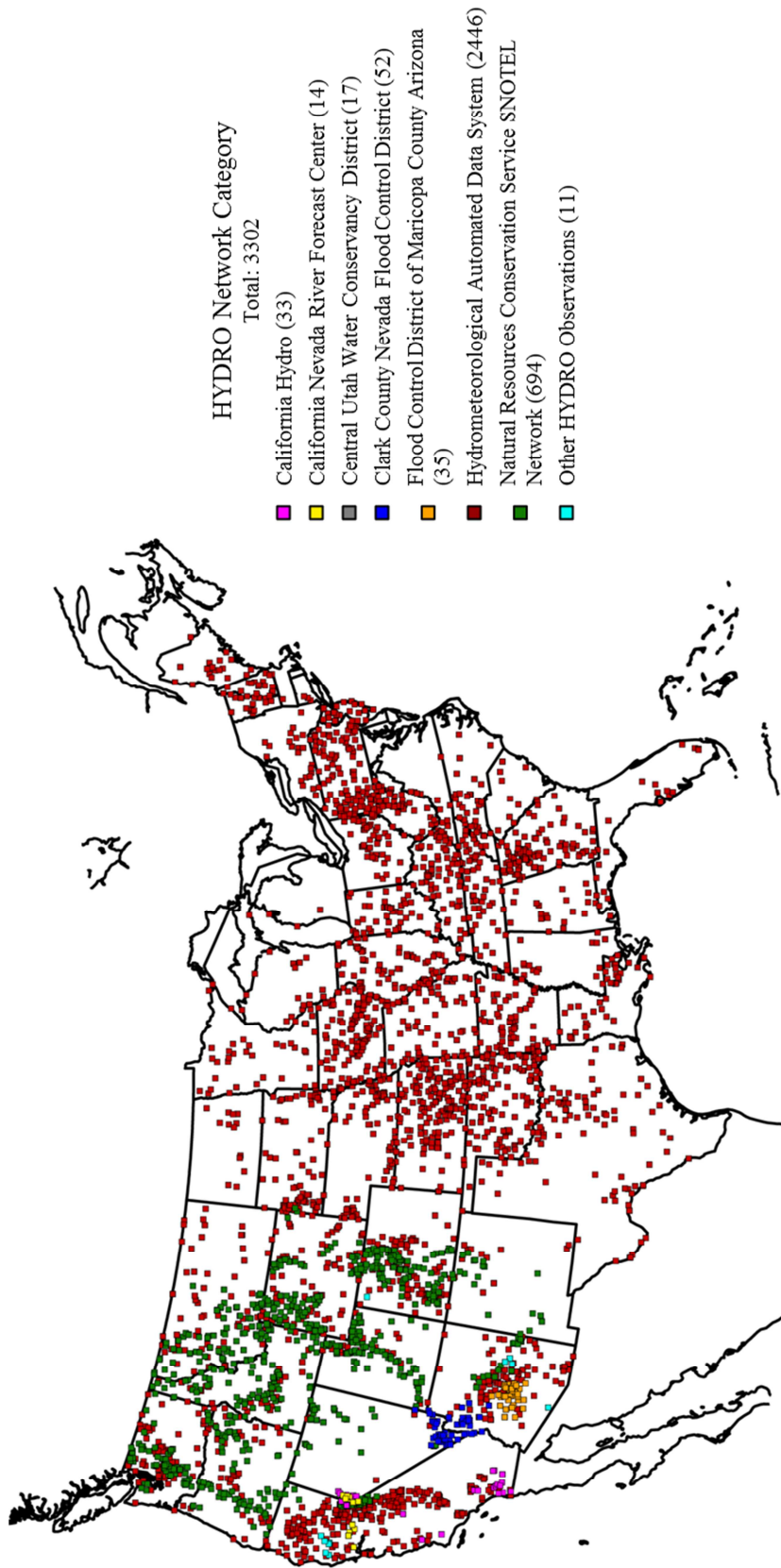


Figure 4.1. continued. e. HYDRO network category.

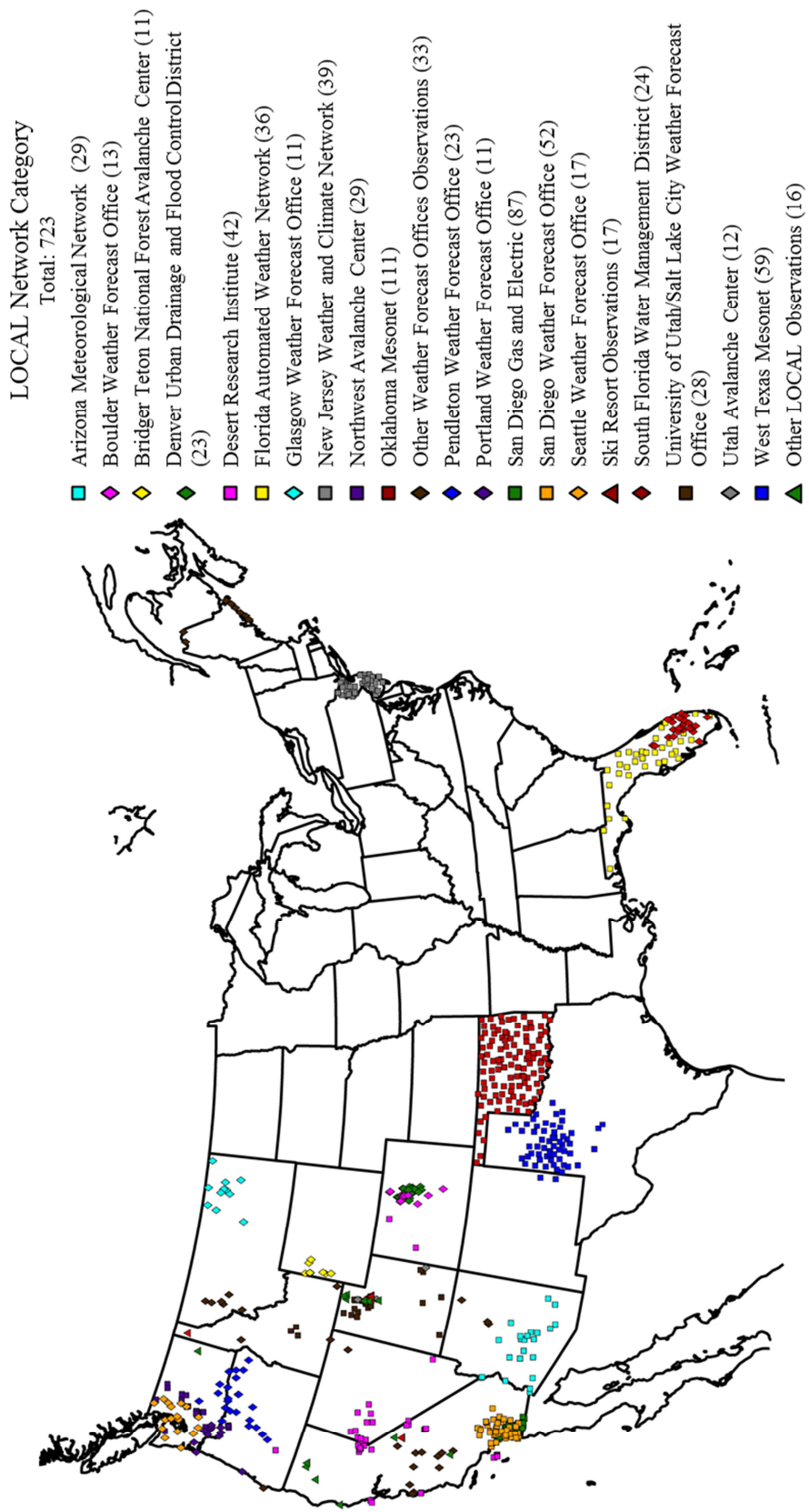


Figure 4.1. continued. f. LOCAL network category.

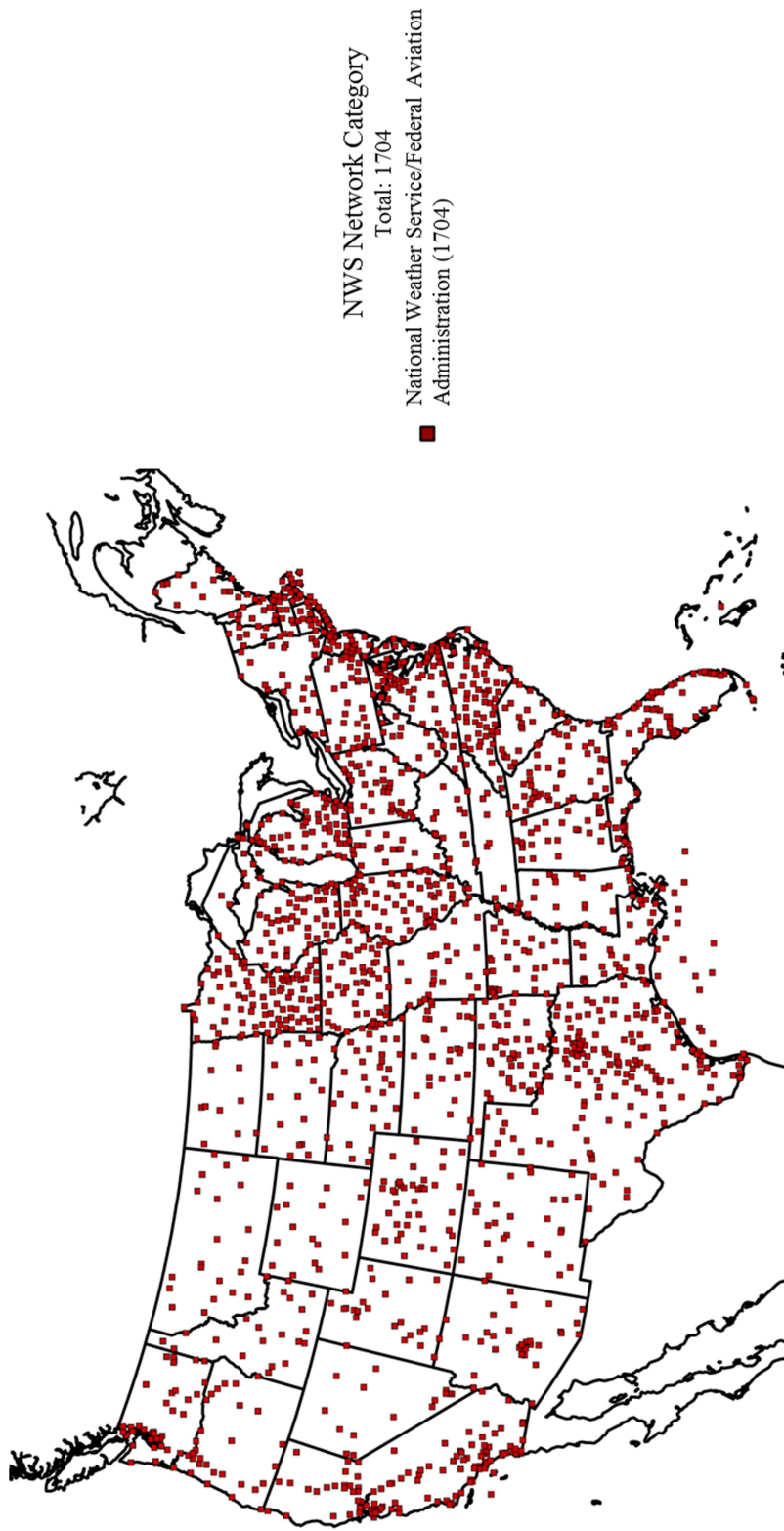


Figure 4.1. continued. g. NWS network category.

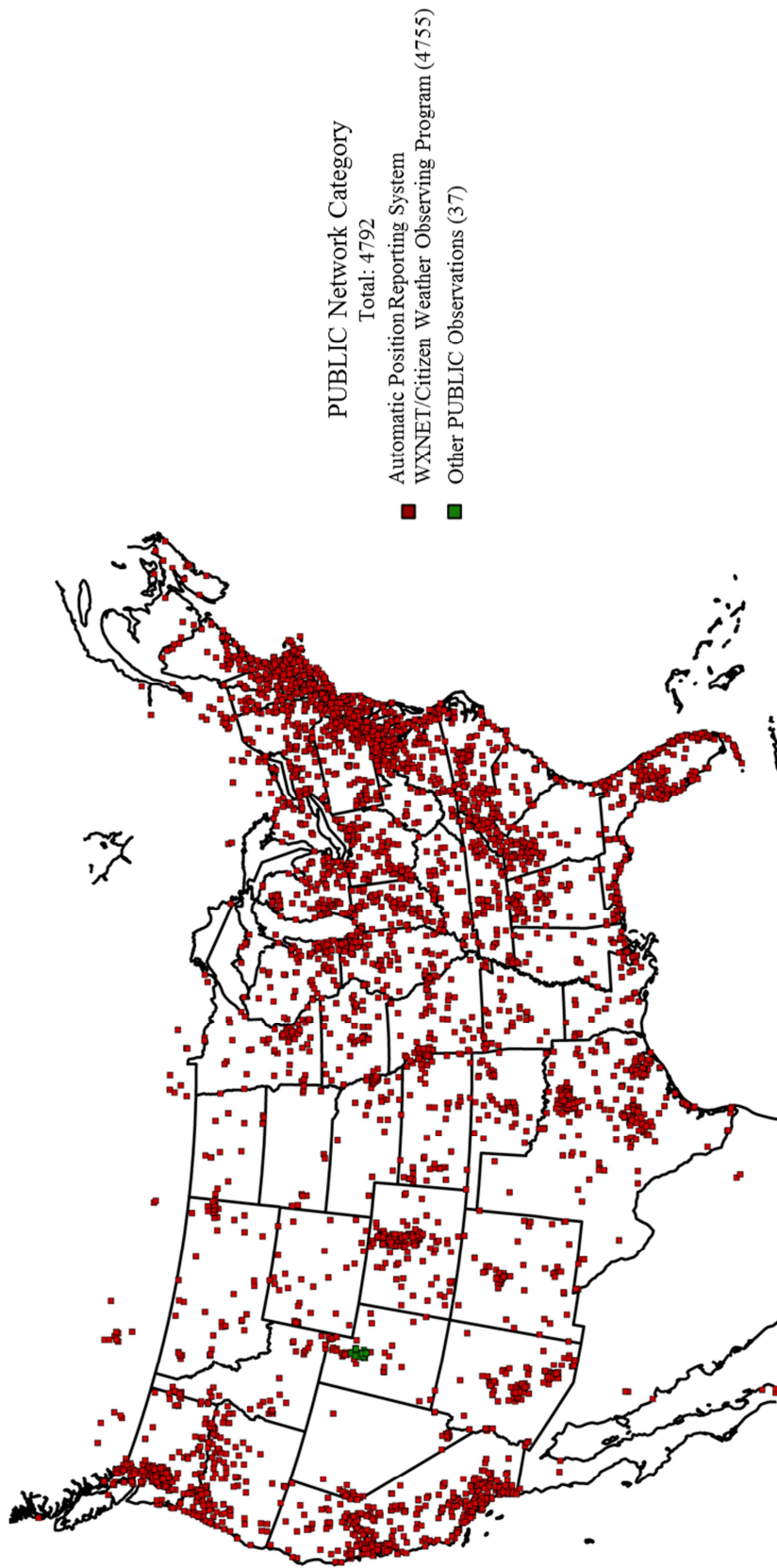


Figure 4.1. continued. h. PUBLIC network category.

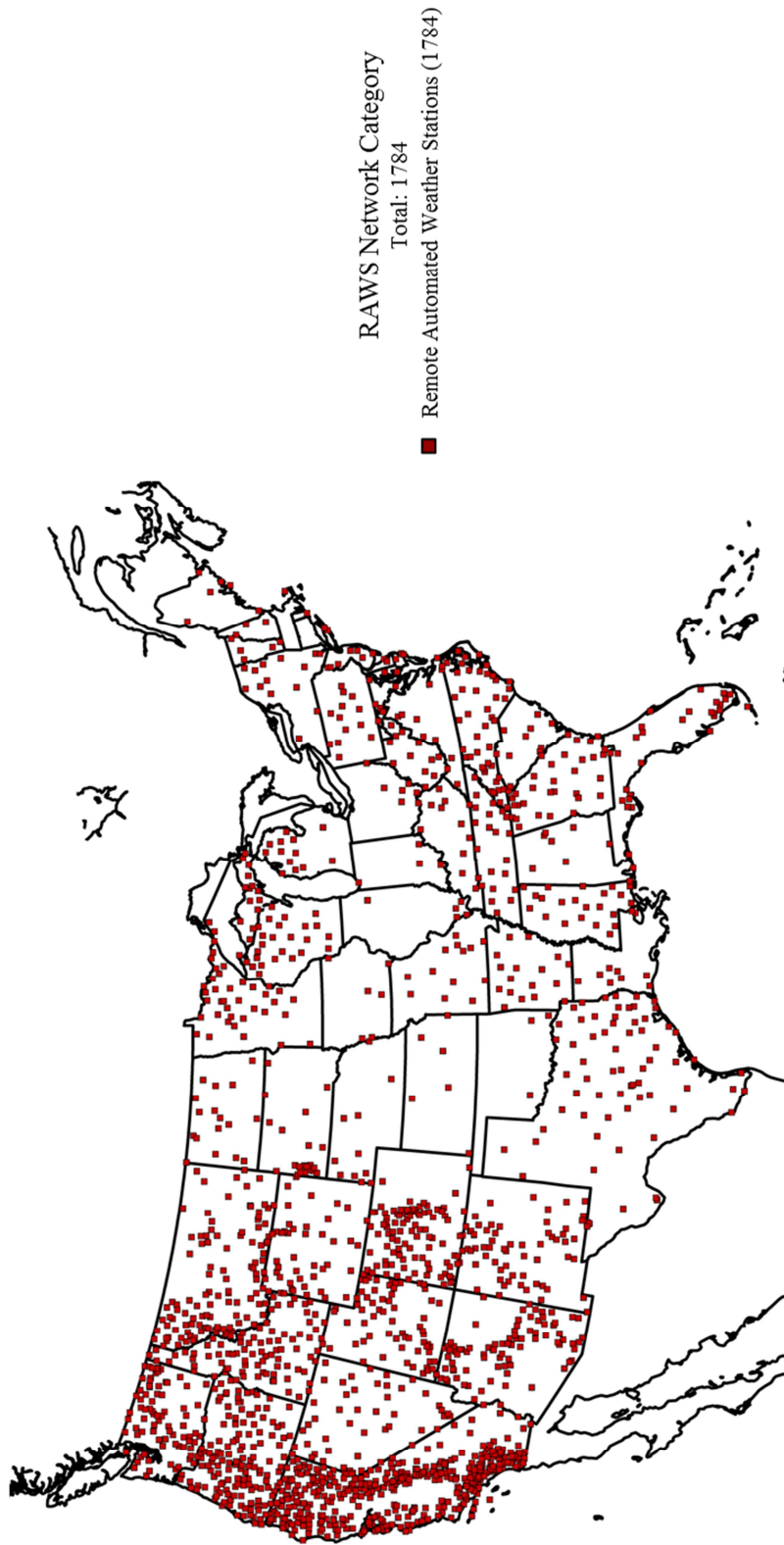


Figure 4.1. continued. i. RAWS network category.

- TRANS Network Category
 Total: 1477
- ▲ Arizona Department of Transportation (17)
 - California Department of Transportation (16)
 - Colorado Department of Transportation (69)
 - Idaho Transportation Department (72)
 - ▲ Indiana Department of Transportation (90)
 - ◆ Iowa Department of Transportation (51)
 - ◆ Kentucky Transportation Cabinet (18)
 - ◆ Maine Department of Transportation (6)
 - ◆ Maryland Department of Transportation (49)
 - ◆ Minnesota Department of Transportation (76)
 - Montana Department of Transportation (44)
 - ◆ Nebraska Department of Roads (40)
 - Nevada Department of Transportation (62)
 - ▲ New Hampshire Department of Transportation (13)
 - ◆ North Dakota Department of Transportation (19)
 - ◆ Ohio Department of Transportation (158)
 - Oregon Department of Transportation (65)
 - Union Pacific Railroad (333)
 - Utah Department of Transportation (65)
 - ▲ Vermont Agency of Transportation (13)
 - ▲ Virginia Department of Transportation (38)
 - Washington Department of Transportation (90)
 - ◆ Wisconsin Department of Transportation (43)
 - Wyoming Department of Transportation (30)

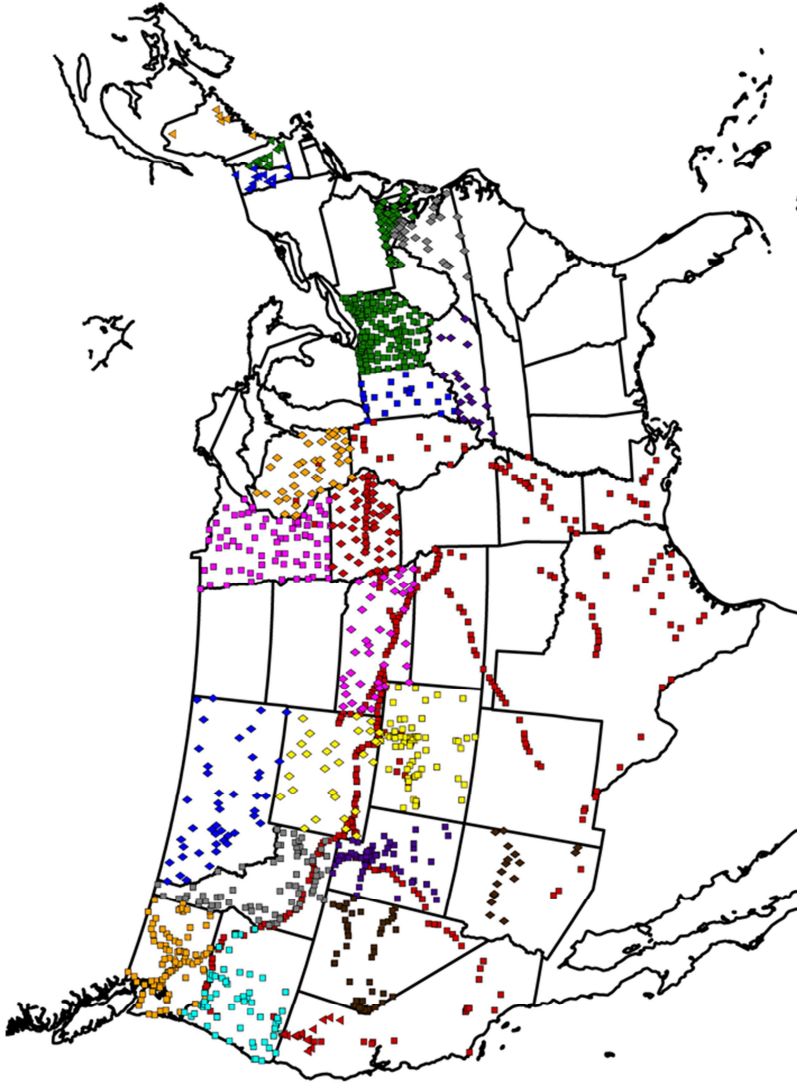


Figure 4.1. continued. j. TRANS network category.

CHAPTER 5

RESULTS AND DISCUSSION

Description of Case Study

This research relies on the 1400 UTC 27 October 2010 CONUS analysis to demonstrate the methodology presented in Chapter 4. This particular analysis occurred during a period of extremely active weather for the eastern United States caused by an extratropical cyclone that progressed across the Great Lakes region prior to the case study time period (Figure 5.1). This cyclone was one of the strongest noncoastal low pressure systems observed within the United States with the lowest sea level pressure (955.2 mb) recorded at Big Fork, MN. The storm was accompanied with sustained winds in excess of 20+ m/s over the northwestern Great Lakes and Dakotas regions and nearly 5 inches of rain in Minnesota. The cold front, which extended south from the low center, brought severe thunderstorms and tornados from the southern Great Lakes region down through the southeast United States.

Figure 5.2 shows the downscaled RUC background fields used in the case study for 2-m air temperature, 2-m dewpoint temperature, and 10-m wind speeds. The stationary and cold fronts are evident in the background fields in terms of gradients in temperature, moisture, and wind speed from central Mississippi to Virginia and from

Pennsylvania into Canada. The high winds associated with the cyclonic circulation across the Dakotas and Great Lake region are quite evident as well.

As mentioned in Chapter 4, there are more than 10,000 observations available to adjust the background fields. Figure 5.3 depicts those corrections to each of the background fields, as well as the resulting analyses using the control case ($k = 0$ for both \mathcal{R}_k and \mathcal{Z}_k from Figure 3.1). The adjustments to the background field appear relatively modest on the scale of the entire continental United States when comparing the final analyses (Figure 5.3b, 5.3d, and 5.3f) to the comparable background grids in Figure 5.2. However, the analysis increments shown in Figure 5.3a, 5.3c, and 5.3e are substantive throughout much of the analysis domain, reflecting the impact of including the observations. The background error decorrelation length scales used in this control case are designed to allow observation innovations to influence the analyses over relatively broad areas (~ 100 km) in areas without significant topographic relief. Positive (orange) analysis increments denote where temperatures or wind speeds in the background are too low, while negative (purple) increments indicate where the background fields are too high.

The largest analysis increments are concentrated near the areas of high impact weather (the strong extratropical cyclone over the Great Lakes and the stationary front in the southeast United States) as well as over the complex terrain of the western United States. Figure 5.3a shows that the background field underestimates the intensity of the stationary front across the southern states while overestimating the intensity of the cold front across New York. The background tended to be too cold throughout much of the Midwest and southern Plains with complex adjustments in temperature across the western

United States. The RUC background tended to be too moist in the upper Midwest and too dry over the Dakotas and most of Texas (Figure 5.3c). The wind speed analysis increments (Figure 5.3e) exhibit a general tendency to analyze lower wind speeds nearly everywhere with the most significant adjustments to the background field in the areas of high winds associated with the cyclone as well as on the warm side of the stationary front in the southeast.

Considerable uncertainty exists whether the general tendency for negative wind increments when surface wind observations are used here or in the RUC or RTMA results from the complex mix of issues related to wind sensor siting as well as the representativeness of those observations in forested and areas of complex terrain. The solution adopted for the RUC and RTMA has been to restrict severely the mesonet wind observations used in those analysis systems (de Pondeca et al. 2011). All possible wind observations were used in this study specifically to help investigate this issue. Evaluation of this case, as well as many others not shown here, suggests that wind observations assigned here to the PUBLIC category have a negative speed bias likely due to siting. However, the large wind speed innovations leading to the large analysis increments in Figure 5.3e in the Great Lakes regions are found not only in the PUBLIC category, but in nearly all network categories. Hence, this tendency for the analysis wind speeds to be less than the background wind speeds may reflect insufficient downscaling of the RUC winds to conditions appropriate for the 10-m level.

Evaluation of Data Density Constraints

Figure 5.4 shows the IDI fields for 1400 UTC 27 October 2010 for 2-m air temperature, 2-m dewpoint, and 10-m wind speed using the horizontal and vertical background error decorrelation length scales of 80 km and 200 m, respectively, that are used for the control analyses shown in Figure 5.3. (The IDI field for temperature in Figure 1.1 used horizontal and vertical background error decorrelation length scales of 40 km and 100 m that are used by the RTMA.) The IDI serves as a metric of observational data density but depends on the specifications of the observational and background error covariances (Equation 3.1). Not surprisingly, observational coverage as defined by the IDI in the eastern United States is greater than that in the western United States due to the smaller number of stations in the west as well as the assumption that background errors at two gridpoints become less related to one other when the two gridpoints are separated in elevation. Although not particularly evident in Figure 5.4, IDI values in the west for wind tend to be a bit smaller than for temperature or dewpoint temperature since many hydrologically-oriented networks (e.g., SNOTEL, Figure 4.1e) in the western United States are not equipped with anemometers. Many of the very small apparent data voids in the eastern United States in Figure 5.4a and 5.4b result from the assumption that background errors on- and off-shore of water bodies (whether oceans or small lakes) are unrelated to one another for temperature and moisture (Chapter 3). This land/water contrast assumption is not used for the wind background error covariance matrix due to the presumption that the background errors remaining well correlated across coastlines (M. de Ponca, personal communication).

As discussed in Chapter 3, the IDI field can be used as a constraint on the analysis such that the background errors in data rich regions are assumed to be less correlated with one another, hence allowing smaller scale deviations between the observations and the background to be reflected in the final analysis. The background errors in data voids can be assumed to be more correlated with one another, allowing deviations between the sparse observations in such regions and the background to be felt over broader distances. In the context of Figure 5.4, the decorrelation length scales are narrowed for gridpoints that are dark blue for the $k = 1$ and $k = 2$ polynomials described in Chapter 3, and broadened for gridpoints that are red for the $k = 2$ polynomial.

Figure 5.5 shows the increments and analyses for temperature, dewpoint, and wind speed using the $k = 2$ polynomial IDI constraint. Differences between Figures 5.3 and 5.5 can be seen in the increments of all fields from Texas to the northern Midwest, as well as from southern Appalachia to the northeast United States. Smaller scale increments, collocated with areas of high observation density, are evident in Figure 5.5, such as in Wisconsin, Iowa, Texas, and western California, which lead to small scale, but large differences between the analyses.

These differences are not necessarily associated with a better analysis since narrowing the background error decorrelation length scales tends to lead to overfitting. As described by Daley (1991) for polynomials and Tyndall et al. (2010) for variational analyses, overfitting artificially creates false maxima and minima that are not representative of the data, or in the case of variational analyses, the background field and the observations. In variational methods, the likelihood of overfitting errors appearing in data sparse areas increases as the analysis is constrained more tightly to the observations

(Tyndall et al. 2010). Overfitting errors are especially evident in Figure 5.5e as an offshore band of much larger negative increments from Virginia to Maine relative to that seen in Figure 5.3e. This band and the smaller secondary negative increment band farther offshore is a result of the extreme gradient of observation density from the coast to the offshore zone. While not shown here, these overfitting errors also appear, not surprisingly, in the same region of the wind analysis if the $k = 1$ polynomial is used to compute the horizontal and vertical decorrelation length scales since that applies the same constraint in data rich areas.

Besides the overfitting errors presented above, bull's-eye features throughout the Great Lakes region in the wind analysis of Figure 5.5f result from weak wind speed observations where the background wind speeds are quite large. This apparent noise is lessened in the control wind analysis in Figure 5.3f, which suggests that the data density constraint, if applied for wind, must be accompanied with effective removal of unrepresentative observations as part of a quality control procedure.

As applied here, the broadening of the background error decorrelation length scales in data sparse regions does not appear to have a large impact on the analysis increment fields (compare Figure 5.3 and 5.5). Most of the analysis increments that appear to be affected are located along the Mexico border and the United States coastline from buoy observations. Although data density is low in many mountainous areas of the western United States, the apparent limited impact of broader decorrelation length scales in many of those areas is likely due to the continued dominance of the elevation constraint on the decorrelation length scale.

However, broadening the decorrelation length scales (using the $k = 2$ polynomial instead of the $k = 1$ polynomial) comes at significant computational expense as well. Increasing the variation localization threshold distance (Chapter 3) increases the computational time to compute the arrays $\mathbf{P}_b\mathbf{H}^T$ and $\mathbf{H}\mathbf{P}_b$ and significantly increases the memory requirements of the analysis. For example, computing $\mathbf{P}_b\mathbf{H}^T$ and $\mathbf{H}\mathbf{P}_b$ with the broadened decorrelation length scales for 2-m air temperature for this particular case on 8 processors requires an additional 4 minutes and triples the memory requirements. Hence, the potential benefits of increasing decorrelation length scales in data sparse areas must be weighed against these increased computational requirements.

Sensitivity and Impact to Observation Networks

As described in Chapter 4, observation sensitivity ($\partial\mathcal{J}/\partial\mathbf{y}_o$) and observation impact (\mathcal{G}) were computed for every observation in the case study. Because of the large number of figures that would be required to describe the sensitivity and observation impacts for all variables and all networks, the approach will be demonstrated here in terms of observation impact for 2-m air temperature for the control case (Figure 5.3) and the $k = 2$ polynomial case (Figure 5.5). The impact of dewpoint and wind speed observations will be briefly summarized near the end of this section followed by an even more cursory summary of observation sensitivity results.

As specified by Equation 4.6, the observation impact, \mathcal{G} , is the product of an observation's innovation and its sensitivity, $\partial\mathcal{J}/\partial\mathbf{y}_o$. As a result of the cost function used in this study (Equation 4.4), the sensitivity depends on the analysis increment at the observation location. Hence, the impact will tend to be positive since the innovation and

sensitivity will usually have the same sign, i.e., in the case of an isolated observation, a positive (negative) innovation will tend to lead to locally a positive (negative) increment. Negative impacts are likely to occur only where the deviation of an observation from the background differs in sign and has a large magnitude relative to its neighbors, which may reflect either an observation in error or a realistic weather phenomenon on a scale smaller than that assumed a priori for the background errors. Since negative observation impacts are found to only comprise ~20% of the total and their magnitudes tend to be generally smaller than the corresponding positive ones, the impacts are ranked in terms of their absolute value from smallest to largest based on the entire sample of all observations.

The methodology to evaluate observation impact is illustrated in Figure 5.6, which plots the observation impact percentile by observing network category for the control case (Figure 5.3). Figure 5.6 focuses on observations in the upper and lower quartiles (i.e., observations that had the most and least impact on the analysis, respectively). Because the percentiles of observation impact depicted in Figure 5.6 may overlap each other in data rich areas, the impacts are plotted in order from least impact to the most impact in order to identify those regions where those particular observations tend to have the greatest affect in adjusting the background field.

Many of the first panels in Figure 5.6 tend to demonstrate the strong dependence of observation impact on station density as can be seen by comparing Figure 5.4 and Figure 5.6. The preponderance of high percentile (red) vs. low percentile (blue) impacts of agricultural (AG) temperature observations relative to the entire sample of observations is evident in Figure 5.6a. In contrast, air monitoring stations (AQ; Figure 5.6b) have more stations in the bottom quartile, i.e., AG stations tend to be in more

remote locations than AQ stations located in urban areas where many other data assets are generally available. In addition, it is possible to infer from Figure 5.3a that the analysis increments in this case tend to be small in the vicinity of the AQ observations, e.g., along the northeast United States coast, in the Central Valley of California, or along the coast of southeast Texas. Similarly, Figure 5.6c shows the frequently high impact of temperature observations from outside the continental United States (EXT category networks) where data density tends to be low. Offshore stations in the Gulf of Mexico tend to have relatively low impact, even though their sensitivity is high (not shown), since the observations and background do not differ substantively. Networks grouped into the FED category (Figure 5.6d) also have a higher percentage of high observation impact temperature observations than low impact observations. Many of these high impact observations are found in data sparse areas, specifically, in Nevada, Utah, and Idaho. The importance of observations in generally data sparse regions is particularly evident in Figure 5.6e for the HYDRO network category. SNOTEL observations at high elevations in the Sierra, Cascade, and Rocky Mountains tend to exhibit very high impact.

The broad range of networks aggregated into the LOCAL category exhibit regional dependencies due to the type of weather event underway at this time as well as station density (Figure 5.6f). The West Texas Mesonet contains the majority of the high impact temperature observations in this category due to their large positive observation innovations leading to large temperature analysis increments over much of this region (Figure 5.3a). In contrast, observations from the northern half of the Oklahoma Mesonet and Florida Automated Weather Network tend to have low impact because of the small departures in temperature from the background in those regions. The contribution of

NWS network observations to the analysis is seen by the high impact observations across the Midwest and southward into Texas (Figure 5.6g). However, the NWS category also has a large number of low impact observations concentrated along the coast of the northeast United States, which are the result of small observation innovations in the area and high data densities.

Although there is a prevailing tendency to dismiss observations provided by the general public through the CWOP program (PUBLIC network category), Tyndall et al. (2010) showed that the error characteristics for temperature observations from those stations are similar to those from other network sources. Similarly, Figure 5.6h shows that the impact of temperature observations from PUBLIC stations can be high and consistent in terms of their locales with those provided from other networks (e.g., compare to the NWS observations in Figure 5.6g). In other words, if the background field differs significantly from the actual weather, then PUBLIC observations can be quite useful, especially if there are relatively few other observations nearby. However, as will be shown later, the vast majority of PUBLIC temperature observations have low impact and those stations are simply obscured in Figure 5.6h.

The impact of temperature observations from the RAWS network is depicted in Figure 5.6i. Since RAWS temperature observations contribute significantly to the negative temperature analysis increments in the mountainous regions in the western United States (Figure 5.3a), many of those stations exhibit high impact. In addition, RAWS stations extending northeastward from eastern Texas have a large impact in this case consistent with those of NWS and PUBLIC stations along this swath. Finally, Figure 5.6j shows the impact of temperature observations from the transportation network

category. Due to their coverage in otherwise data void regions, TRANS temperature observations contribute frequently to the temperature analysis in the western United States as well as to the broad region of the positive temperature analysis increment found across the Midwest.

To evaluate the influence of the background error decorrelation length scales on observation impact, results from the $k = 2$ polynomial case (Figure 5.5) are now presented in Figure 5.7. Not surprisingly, this modification of the decorrelation length scales increases the observation impacts for networks that are primarily located in data rich regions. For example, the percentage of high impact observations within the PUBLIC network (Figure 5.7h) increases compared to the control case (Figure 5.6h), with observations near many urban areas (Dallas, TX; Detroit, MI; Chicago, IL, San Diego CA; San Francisco, CA; Denver, CO; Portland, OR) increasing their impact relative to all the other observations in this case. The improvement in the number of high impact observations in the PUBLIC network comes at the expense of the observation impacts of the RAWS, EXT, and HYDRO network categories. For example, the impact of RAWS temperature observations is significantly reduced along the Appalachian and Sierra Nevada Mountains (compare Figure 5.7i to Figure 5.6i) while the impact of HYDRO observations is reduced in the mountainous regions of the Intermountain West.

Figure 5.8 summarizes the percentile rank of observation impact for all available stations aggregated into the 10 network categories that are computed from analyses of the 3 variables (temperature, dewpoint temperature, and wind speed) using 5 distinct background error decorrelation length scales. The upper left panel of Figure 5.8a summarizes the results previously shown in Figure 5.6 while the lower right panel of that

figure summarizes the results shown in Figure 5.7. The count of stations (y axis) with observation impacts that fall into each decile category (x axis) is color coded for each of the 10 network categories.

First, the preponderance of observations available from the PUBLIC category tends to dominate all panels in Figure 5.8. For temperature (Figure 5.8a), there are larger numbers of PUBLIC observations that have low impact than high impact regardless of the assumptions related to the background error decorrelation length scale. Narrowing the horizontal decorrelation length scale in data rich areas (middle left and bottom left panels) slightly increases the number of high impact PUBLIC observations, at the expense of the number of high impact observations from other network categories, such as RAWS and NWS. Narrowing both horizontal and vertical decorrelation length scales (middle right and bottom right panels) further increases the number of high impact PUBLIC temperature observations.

In contrast to PUBLIC observations, the count of observations in each decile category is relatively flat for NWS observations (yellow bars) and to a large extent independent of the assumptions related to background error decorrelation length scale (Figure 5.8a). This is consistent with the results shown in Figure 5.6g where there was considerable regional dependency in observation impact for NWS observations. Many of the other networks exhibit similar tendencies. However, RAWS (magenta bars) tend to have more stations with high impact than lower impact.

Figure 5.8b summarizes the impact of dewpoint observations and the results are generally similar to those shown for temperature Figure 5.8a). (The total number of humidity sensors is lower for the AG, HYDRO, and TRANS network categories.)

Application of the data density constraints only slightly increases the impact of observations in the PUBLIC network category. Overall, the influence of the data density constraints on the impact of dewpoint observations appears to be much less than the influence on the impact on temperature observations.

The statistics obtained from wind speed observations shown in Figure 5.8c exhibit very different characteristics compared to the statistics from temperature or dewpoint observations (Figures 5.8a and 5.8b). For the control analysis (upper left panel), there are as many stations in the PUBLIC network category with high impact as low impact while the number of high impact observations increases as the background error decorrelation length scales shrinks. The increased impact of PUBLIC wind observations from the application of the constraints comes at the expense of the impact of the NWS and RAWS observations. The high observation impacts from the stations in the PUBLIC category are related to the aforementioned siting and representativeness issues of PUBLIC observations. In addition, occasional, possibly incorrect or misreported, calm winds obtained from PUBLIC stations produce strong negative observation innovations as well as strong negative analysis increments, and contribute to the high impacts of this category. Application of the asymmetric wind observation quality control discussed in Chapter 3 is one step towards mitigating for these issues, rather than the common operational practice of simply omitting all wind observations from the PUBLIC networks.

As illustrated in this section, observation impact, \mathcal{G} , appears to be a useful metric for assessing the relative role of observations in the development of analyses. An alternative metric is simply the sensitivity, $\partial\mathcal{J}/\partial\mathbf{y}_o$, which depends only on the locations

of the observations and the analysis increments for this case study. The concept of analysis sensitivity is of particular relevance for targeting observations in a complete data assimilation system where additional observations may be of particular importance for a future forecast, yet the value of the observation that would be obtained by that targeted observation is unknown at that time (Langland et al. 1999; Baker and Daley 2000).

The interpretation of observation sensitivity in this study has been judged to be of less relevance for evaluating the relative importance of observations from specific network categories. A large sensitivity could result at a station where an observation matches the background but is surrounded by observations with large deviations from the background, e.g., a station with a strong wind report surrounded by erroneously calm wind reports would be evaluated as having a large negative sensitivity. Summary statistics of the magnitude of sensitivity are presented in Figure 5.9 in a manner similar to that presented in Figure 5.8 for impact. Hence, large positive and negative sensitivities both appear in the highest percentile categories since there is no sign preference for the sensitivity metric.

The percentile distributions of temperature observation sensitivity in the control case (upper left panel of Figure 5.9a) have many similar features to those of observation impact (Figure 5.8a). A notable difference is the relative number of stations in the upper 20th percentile in the HYDRO category such that the HYDRO observations tend to have more stations with high impact (presumably due to the locally larger innovations in many remote mountainous areas) than with high sensitivity. Application of the constraints specified by the $k = 1$ and $k = 2$ polynomials increases the influence of the PUBLIC category more when measured by observation sensitivity instead of by observation

impact. This effect is especially enhanced when the horizontal and vertical decorrelation length scales are both narrowed.

The percentile distribution of dewpoint observation sensitivity is depicted in Figure 5.9b, and is similar to the distribution for temperature observations. As the decorrelation length scales narrow in high observation density regions, the counts of PUBLIC stations with high sensitivities tend to increase. The sensitivity metric tends to accentuate the importance of the RAWS networks compared to that of the impact metric.

The count of wind speed observation sensitivities in the upper 20th percentile from the control analysis (upper left panel of Figure 5.9a) is substantively less than that of the wind speed observation impacts (upper left panel of Figure 5.8a). Application of the data density constraints tends to homogenize the sensitivities from the PUBLIC network stations, in contrast to the increasing impact of the PUBLIC stations as the decorrelation length scales are narrowed. While a large percentage of NWS and RAWS stations have high wind observation sensitivities in the control analysis, the percentages again tend to remain relatively constant when the decorrelation length scales are decreased.

Another notable feature in the summary statistics for wind speed sensitivity is the very high percentage of stations in the top decile from the EXT networks (solid blue bars in Figure 5.9c). Nearly all of these highly sensitive locations are located offshore in relatively data void regions adjacent to onshore highly data rich areas. This high sensitivity may be due to the super-sensitivity artifact, first described by Baker and Daley (2000). Super-sensitivity typically occurs where sharp changes in observation density are found. The overfitting errors in the wind analysis increments seen in Figure 5.3e may result from the combination of super-sensitivity, the strong observation innovations of the

coastal observations, and constraining the analysis too tightly by narrow background decorrelation length scales. Clearly, there is a tradeoff between analysis quality and observation impact and sensitivity. It would be possible to force the observation impact and sensitivity to be high by drastically reducing the background error decorrelation length scales, which would in turn force the analysis to have many bull's-eyes in data rich areas and overfitting issues in data voids.

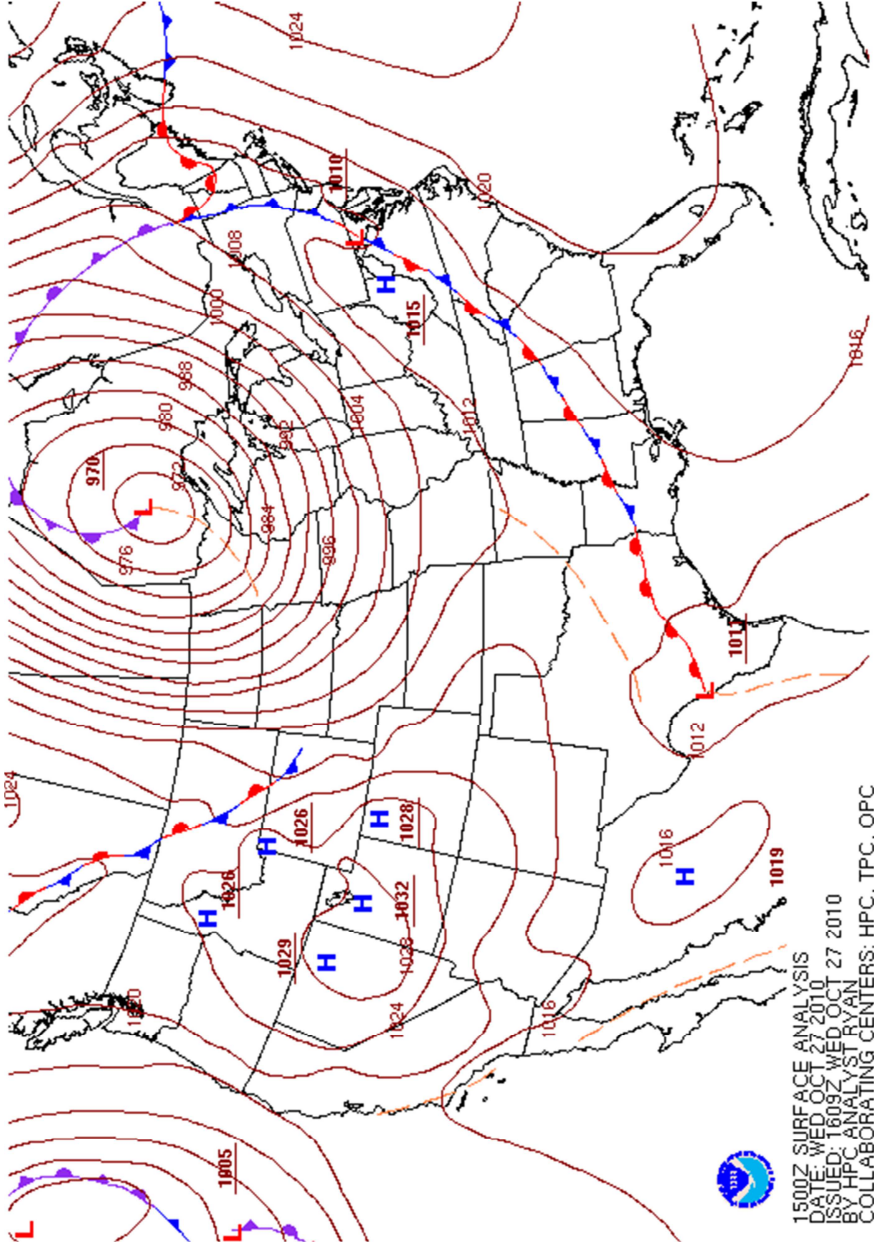


Figure 5.1. Hydrometeorological Prediction Center surface analysis valid for 1500 UTC 27 October 2010 (one hour after the case study presented in this research). Surface pressure isobars are plotted every 4 mb, along with surface fronts and maxima and minima in the surface pressure field.

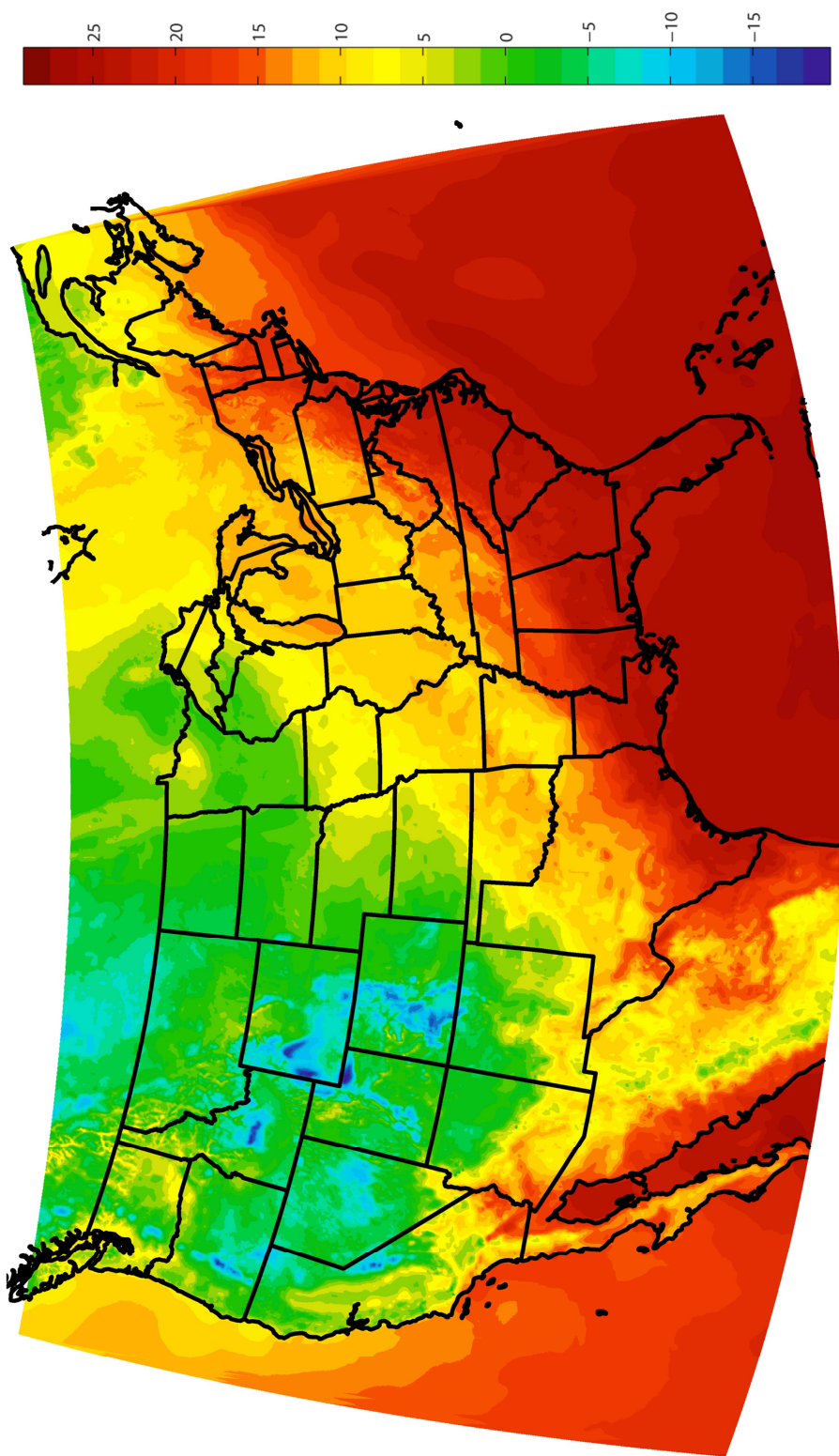


Figure 5.2. RUC 1-hr forecast background fields for 1400 UTC 27 October 2010 over the CONUS domain. a. 2-m air temperature in °C.

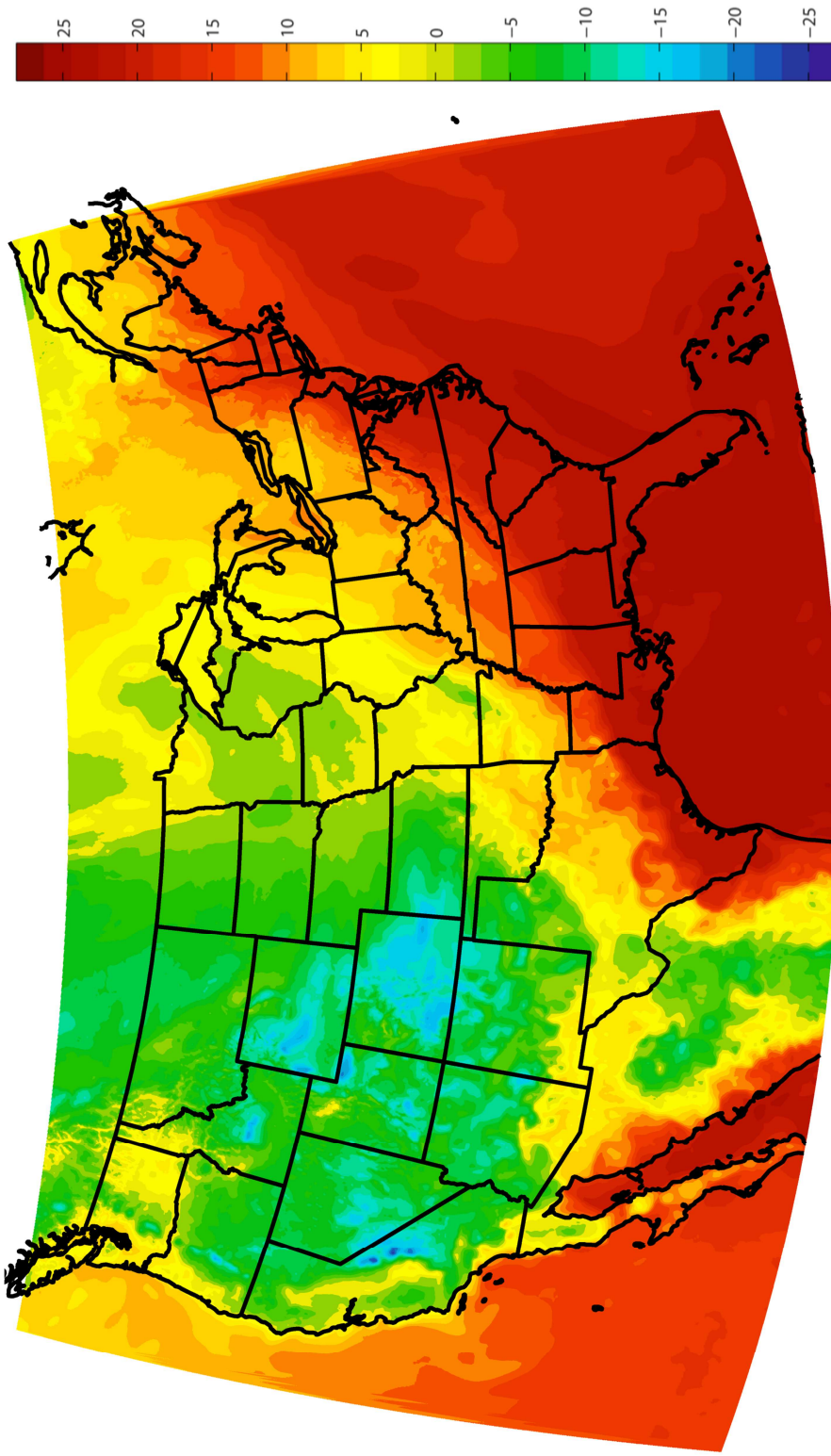


Figure 5.2. continued. b. 2-m dewpoint temperature in °C.

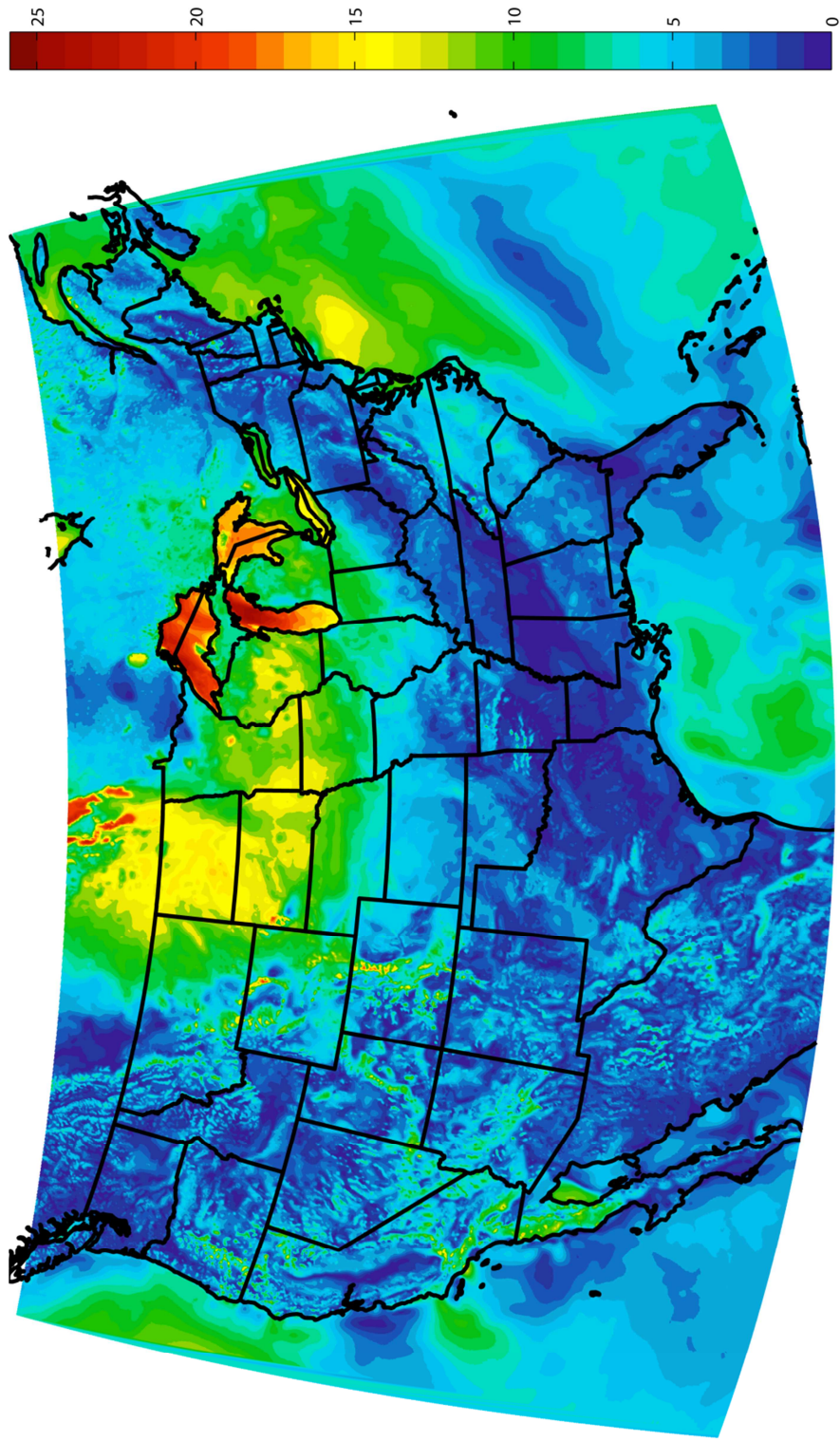


Figure 5.3. continued. c. 10-m wind speeds in m/s.

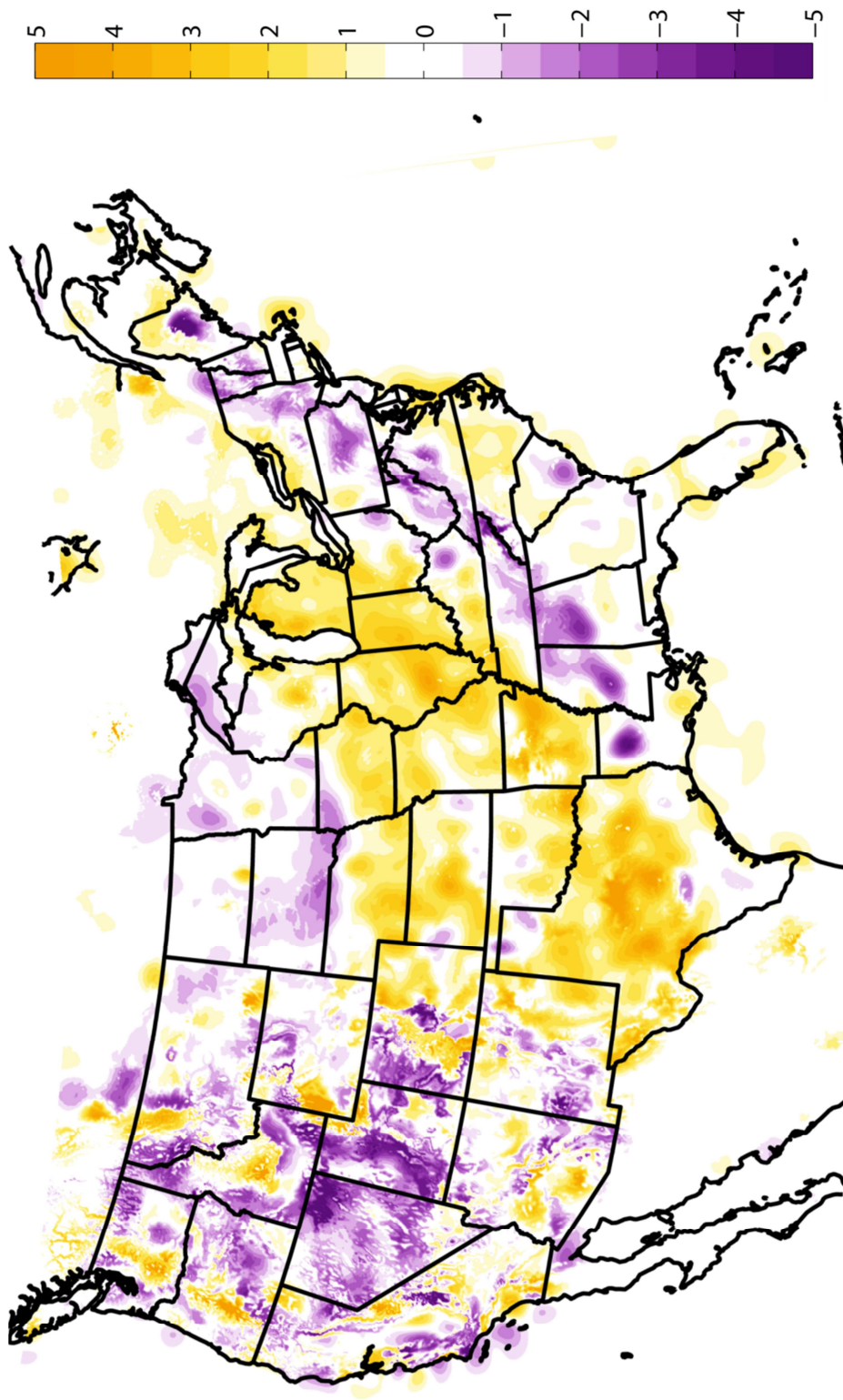


Figure 5.3. UU2DVar increments and analyses for 1400 UTC 27 October 2010 using constant horizontal and vertical decorrelation length scales. a. 2-m air temperature analysis increments in $^{\circ}\text{C}$.

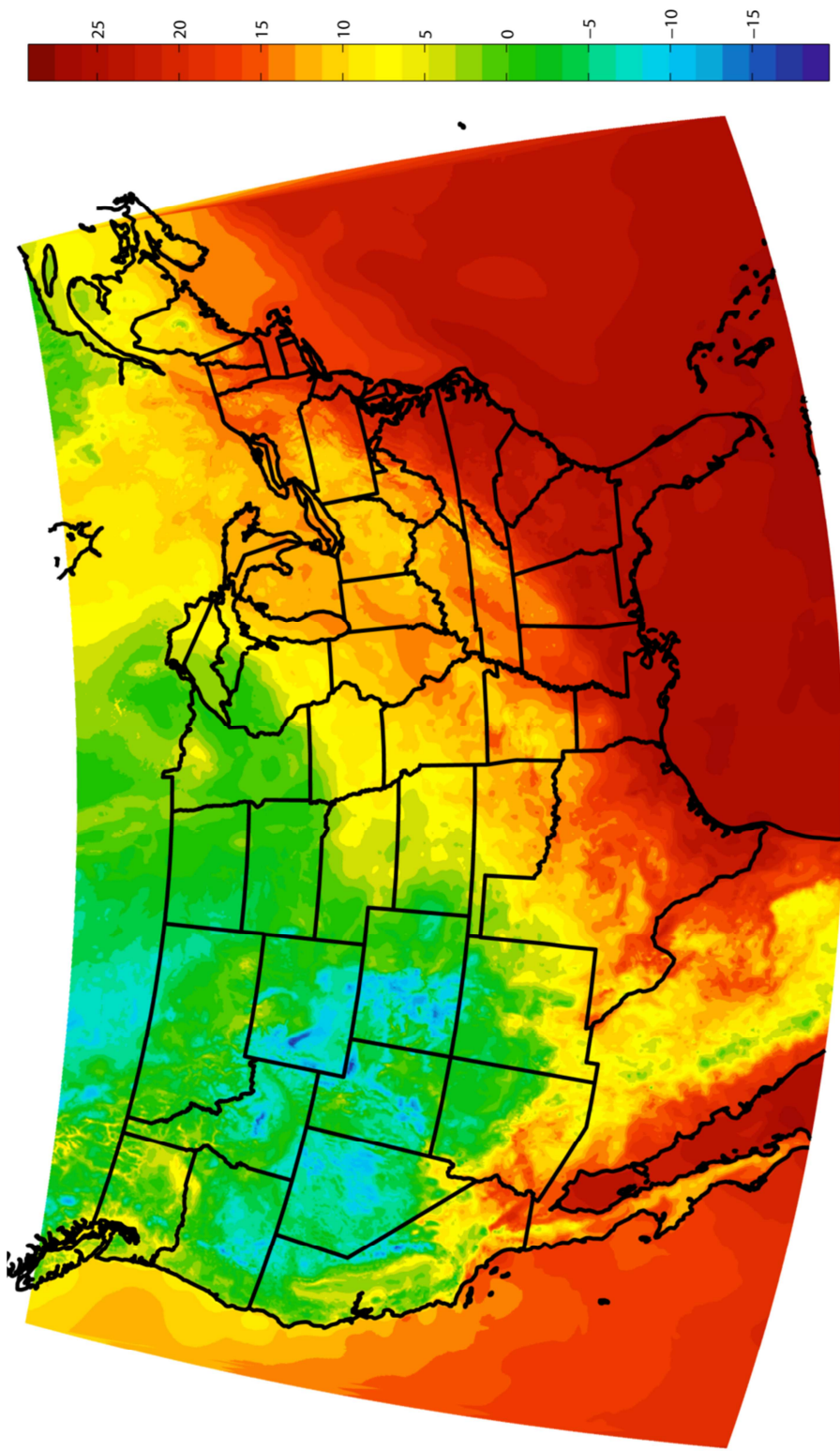


Figure 5.3. continued. b. 2-m air temperature analysis in °C.

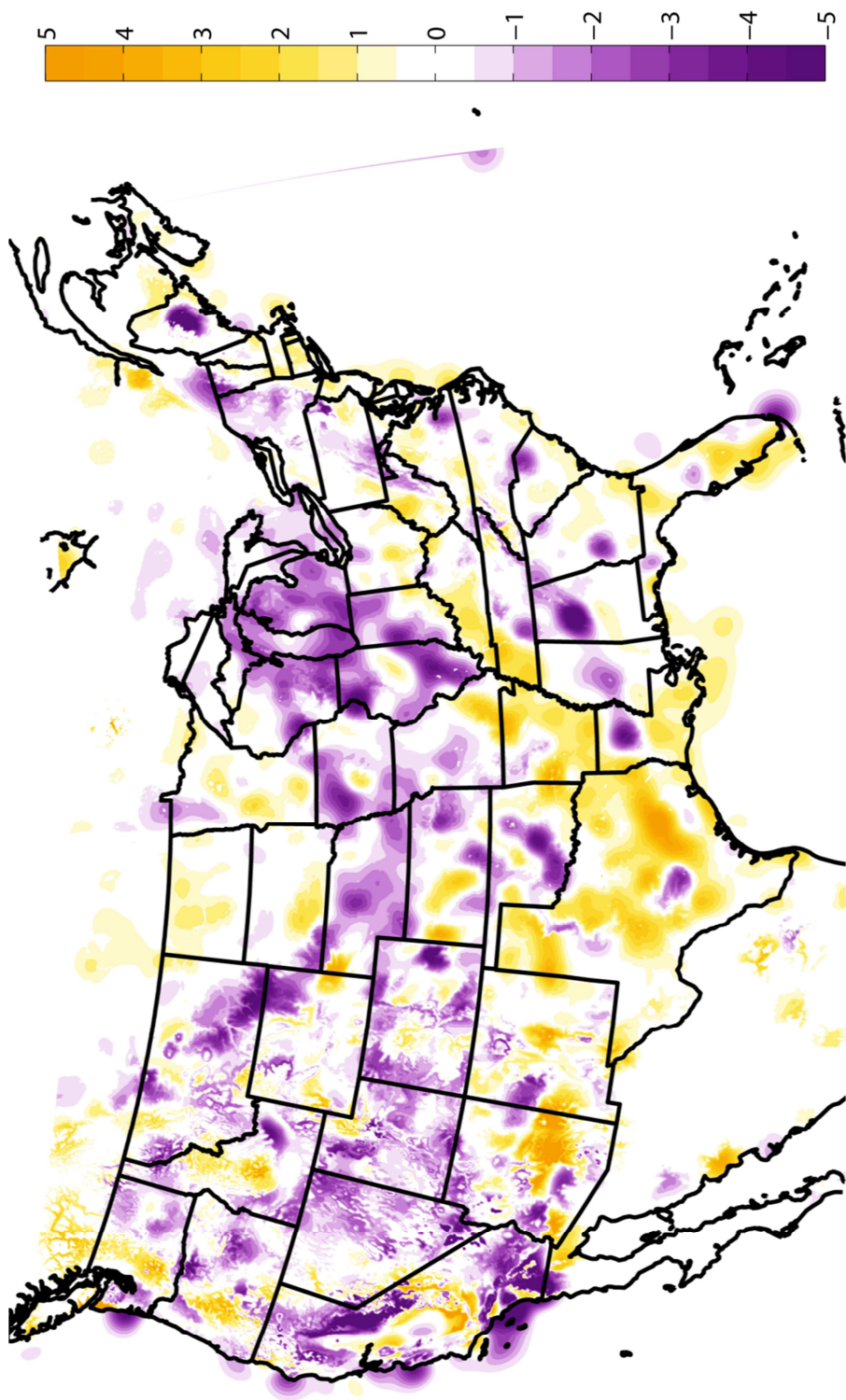


Figure 5.3. continued. c. 2-m dewpoint temperature analysis increments in °C.

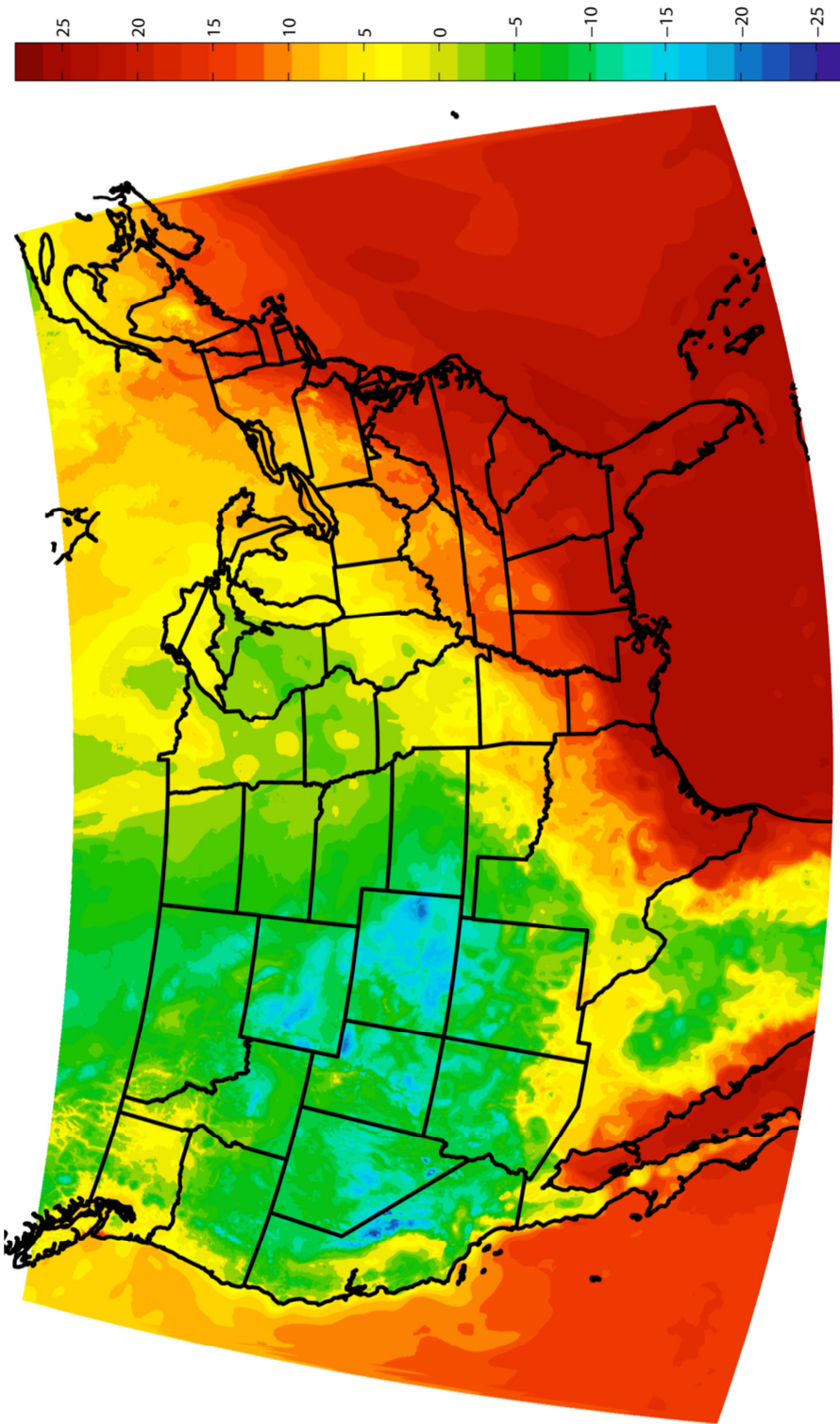


Figure 5.3. continued. d. 2-m dewpoint temperature analysis in °C.

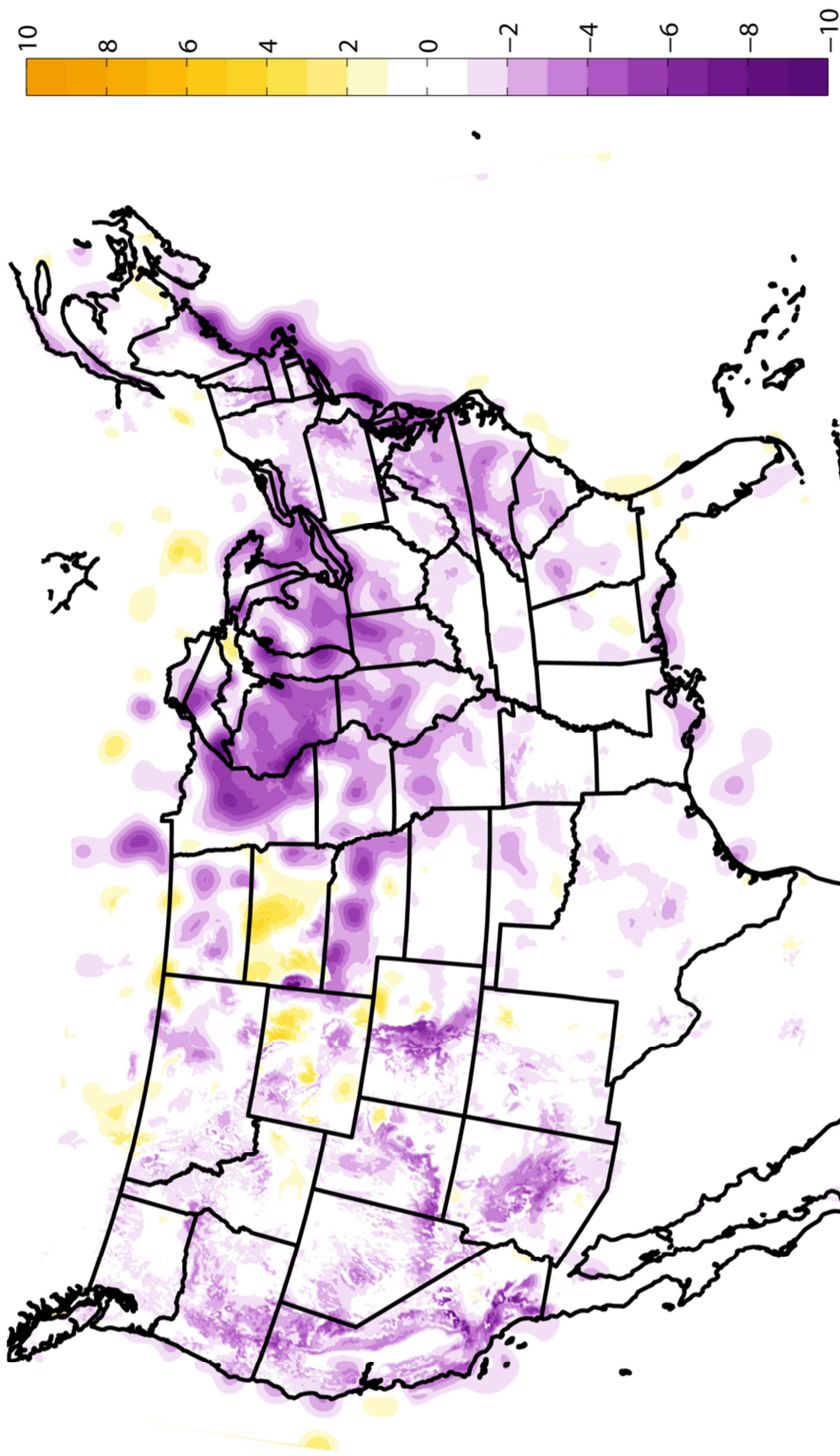


Figure 5.3. continued. e. 10-m wind speed analysis increments in m/s.

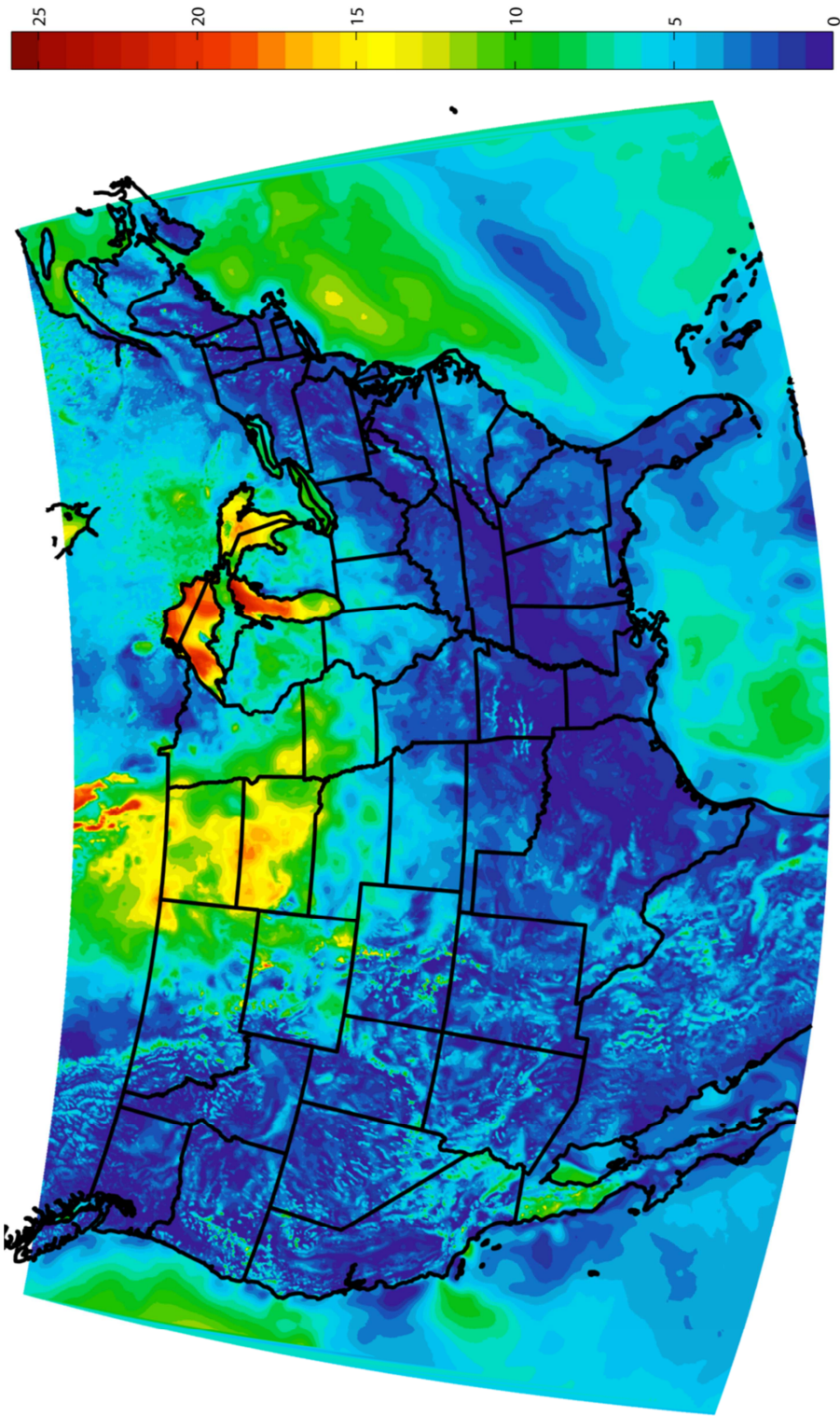


Figure 5.3. continued. f. 10-m wind speed analysis in m/s.

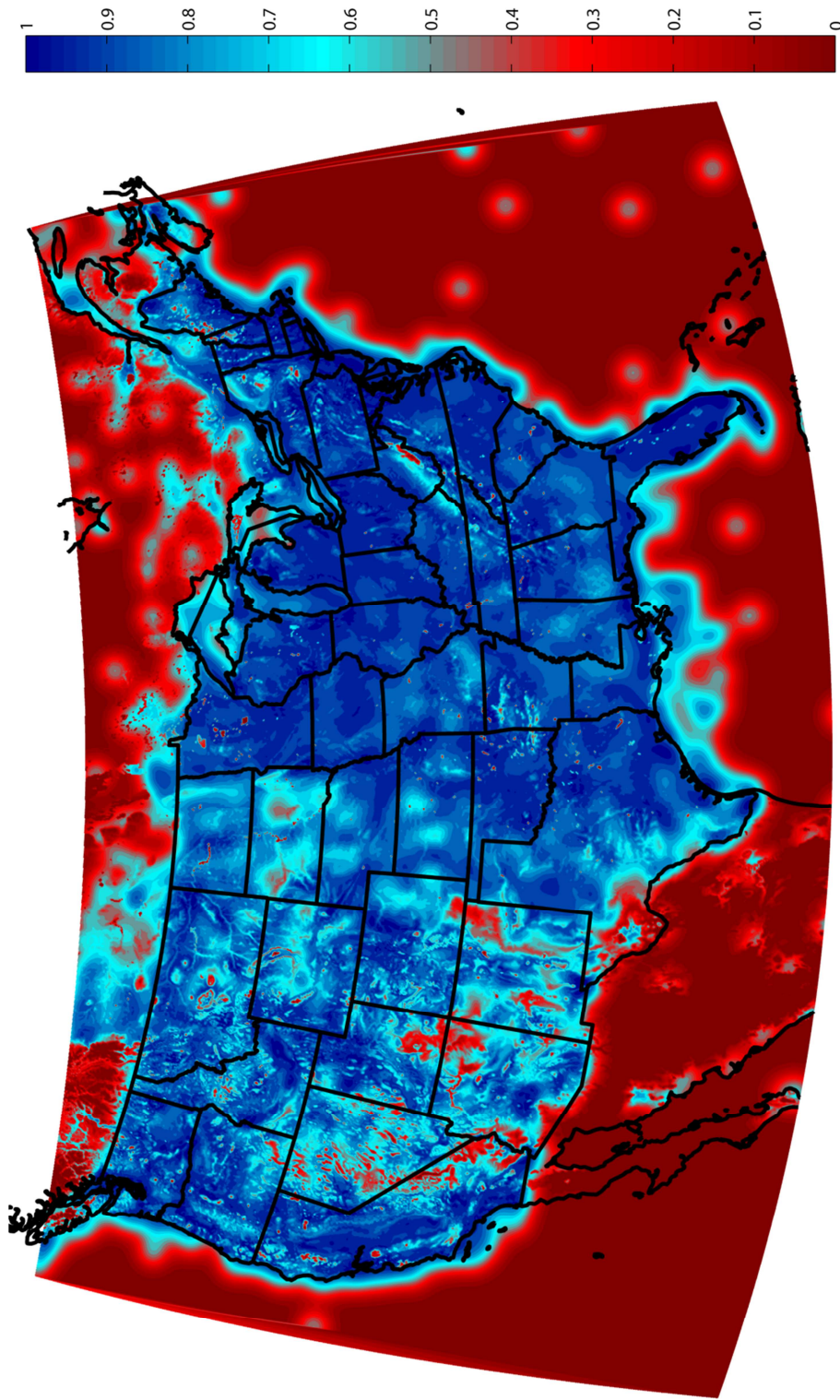


Figure 5.4. IDI analysis for all temperature, dewpoint, and wind speed observations used by the 1400 UTC 27 October 2010 analysis using horizontal and vertical decorrelation length scales of 80 km and 200 m, respectively. a. 2-m air temperature IDI analysis.

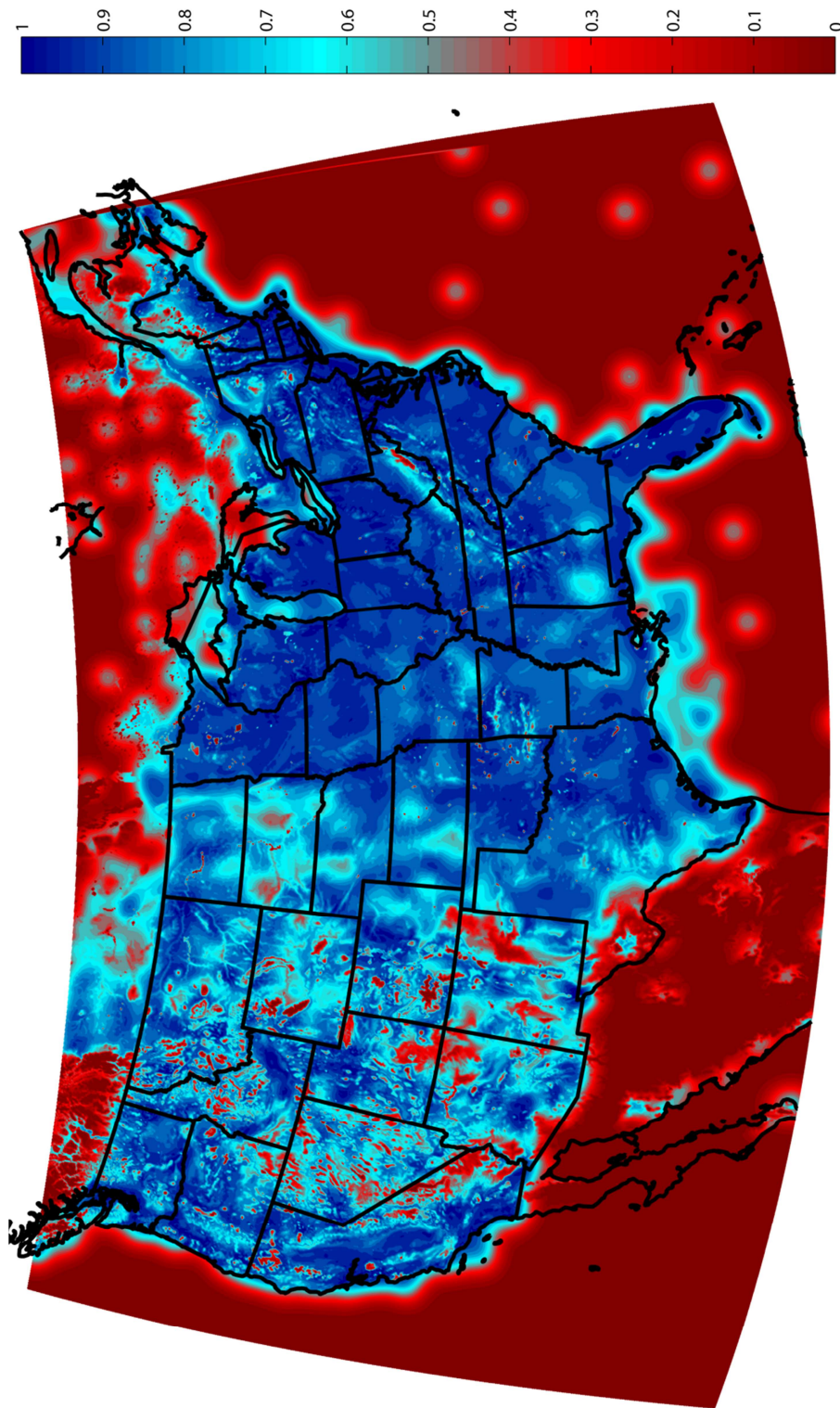


Figure 5.4. continued. b. 2-m dewpoint temperature IDI analysis.

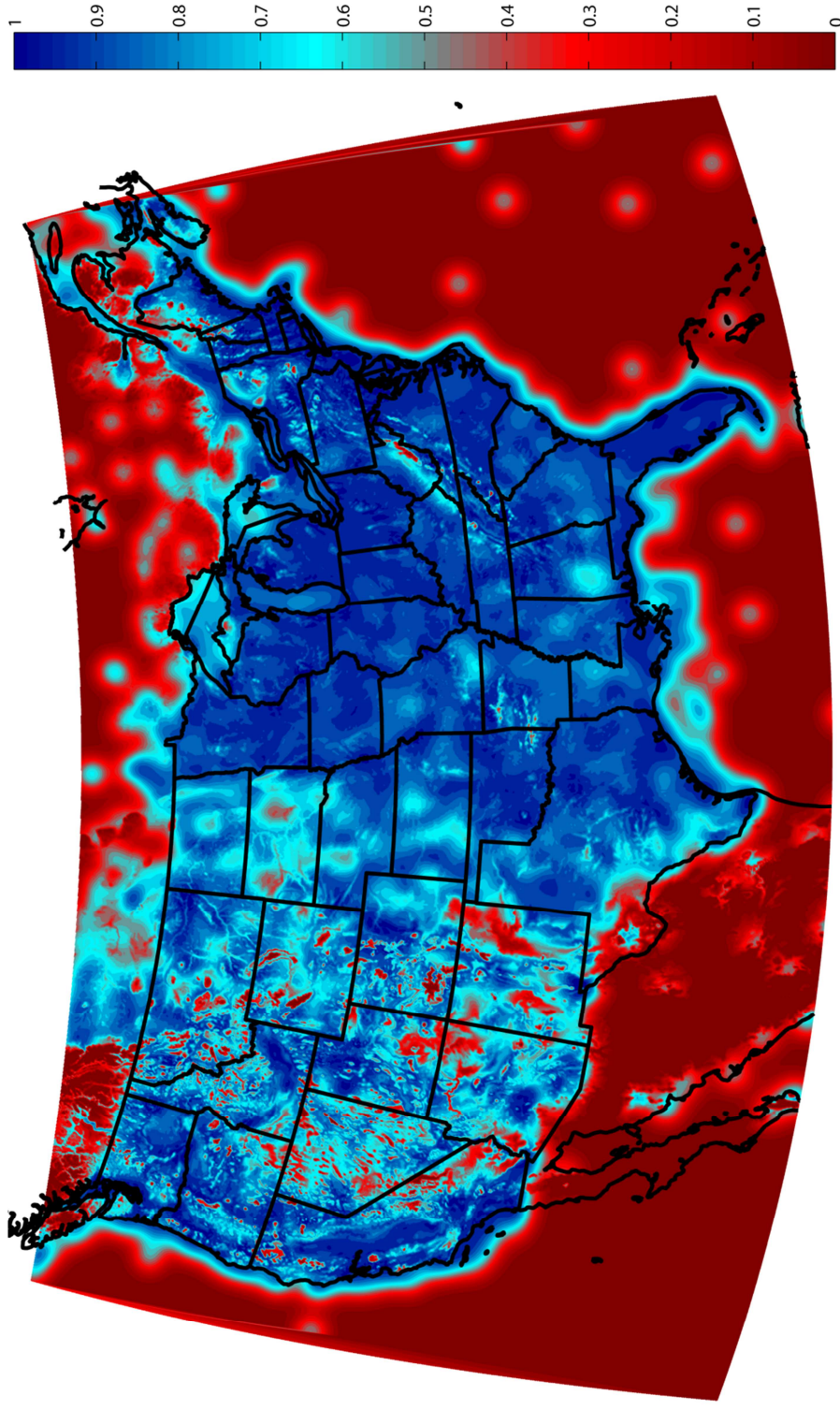


Figure 5.5. continued. c. 10-m wind speed IDI analysis.

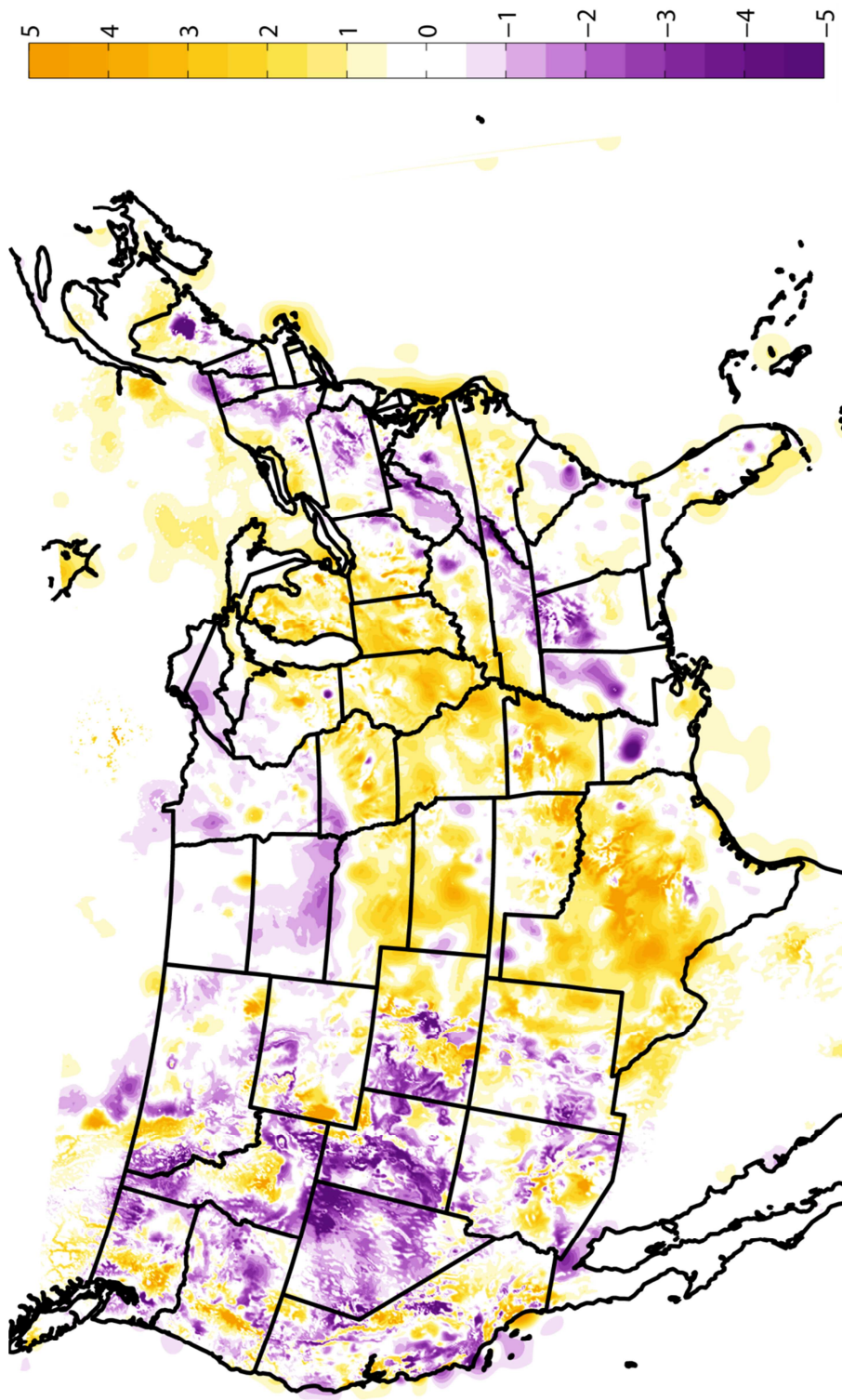


Figure 5.5. UU2DVar increments and analyses for 1400 UTC 27 October 2010 using the $k = 2$ polynomial to compute the vertical and horizontal decorrelation length scales. a. 2-m air temperature analysis increments in $^{\circ}\text{C}$.

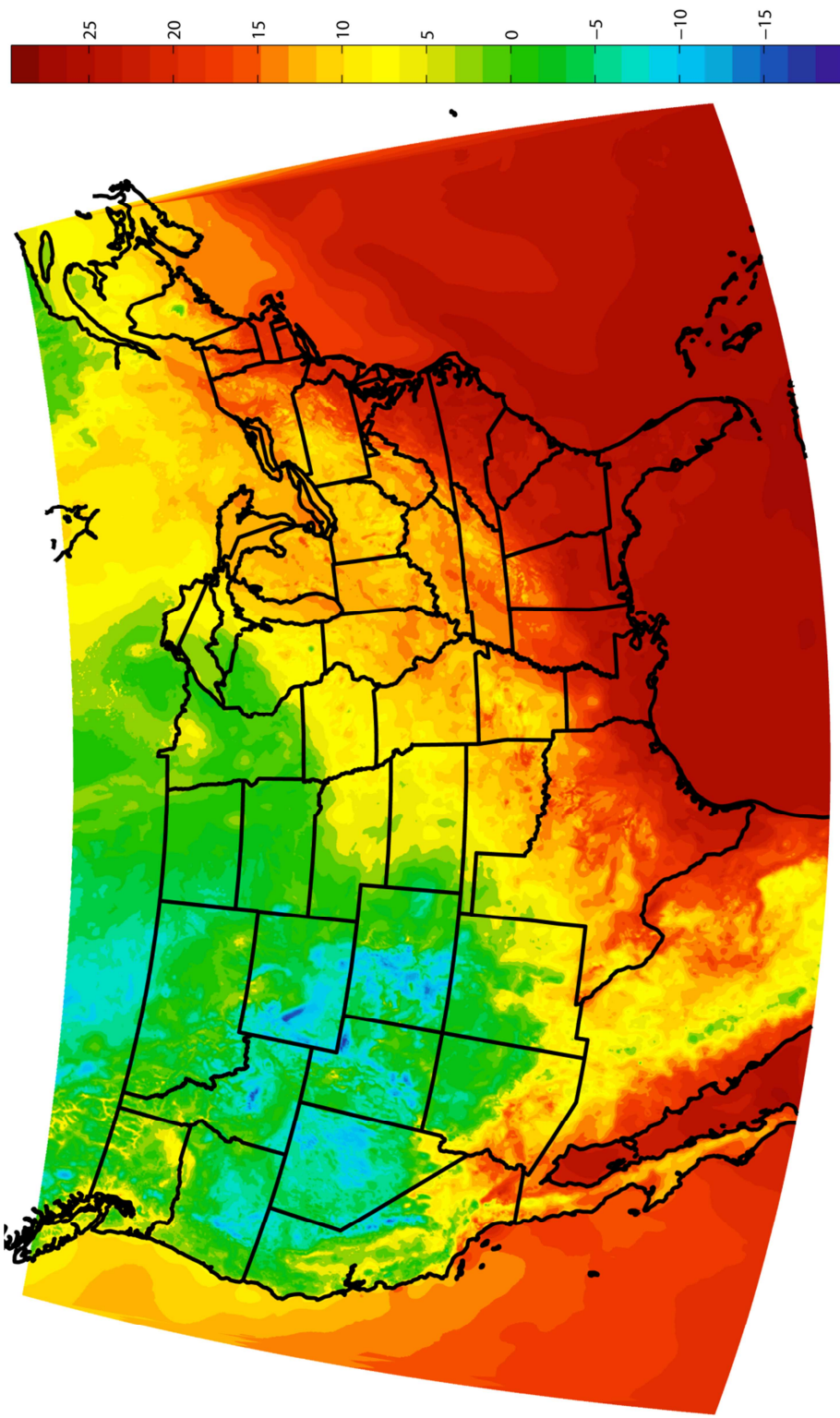


Figure 5.5. continued. b. 2-m air temperature analysis in °C.

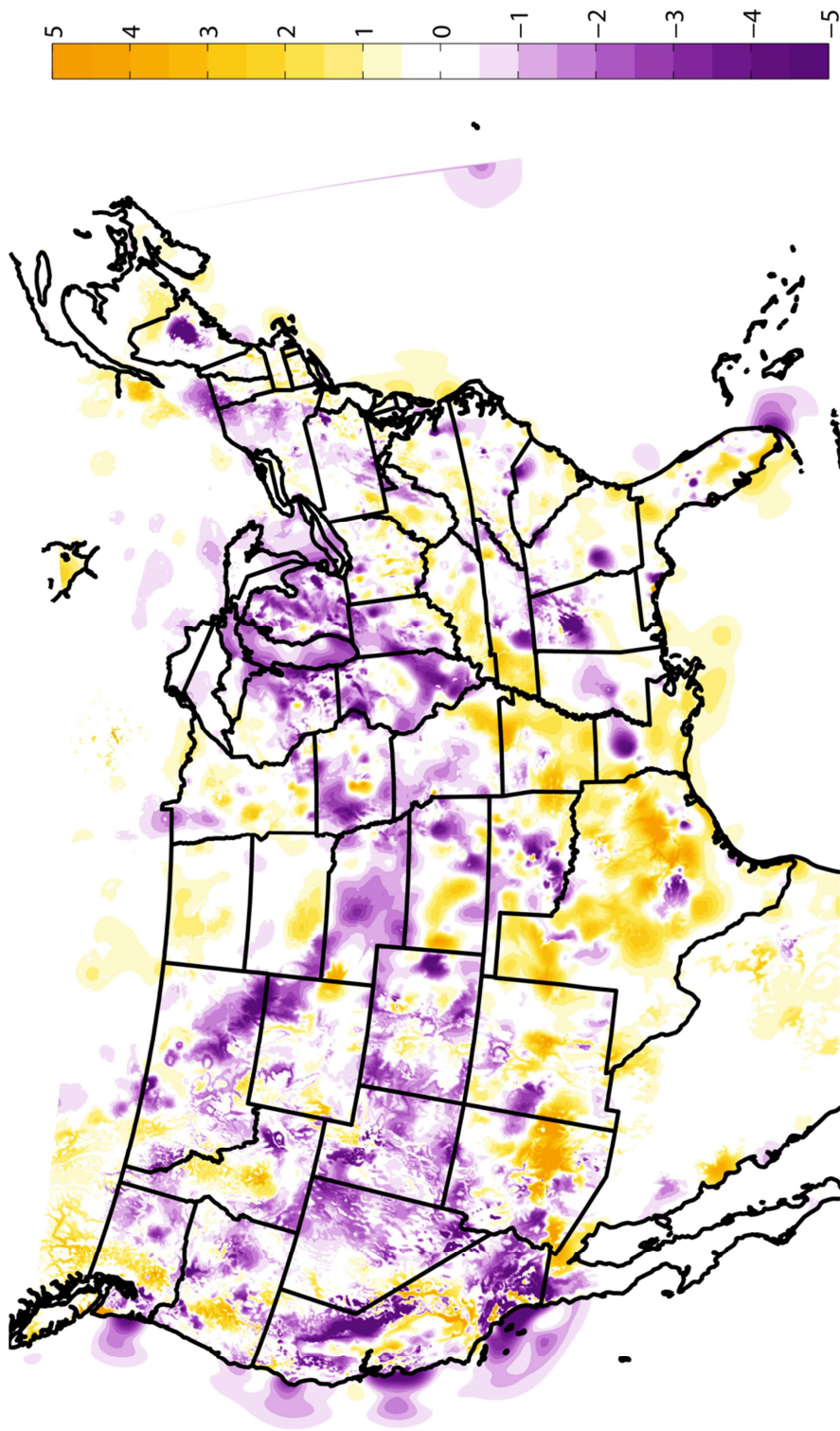


Figure 5.5. continued. c. 2-m dewpoint temperature analysis increments in °C.

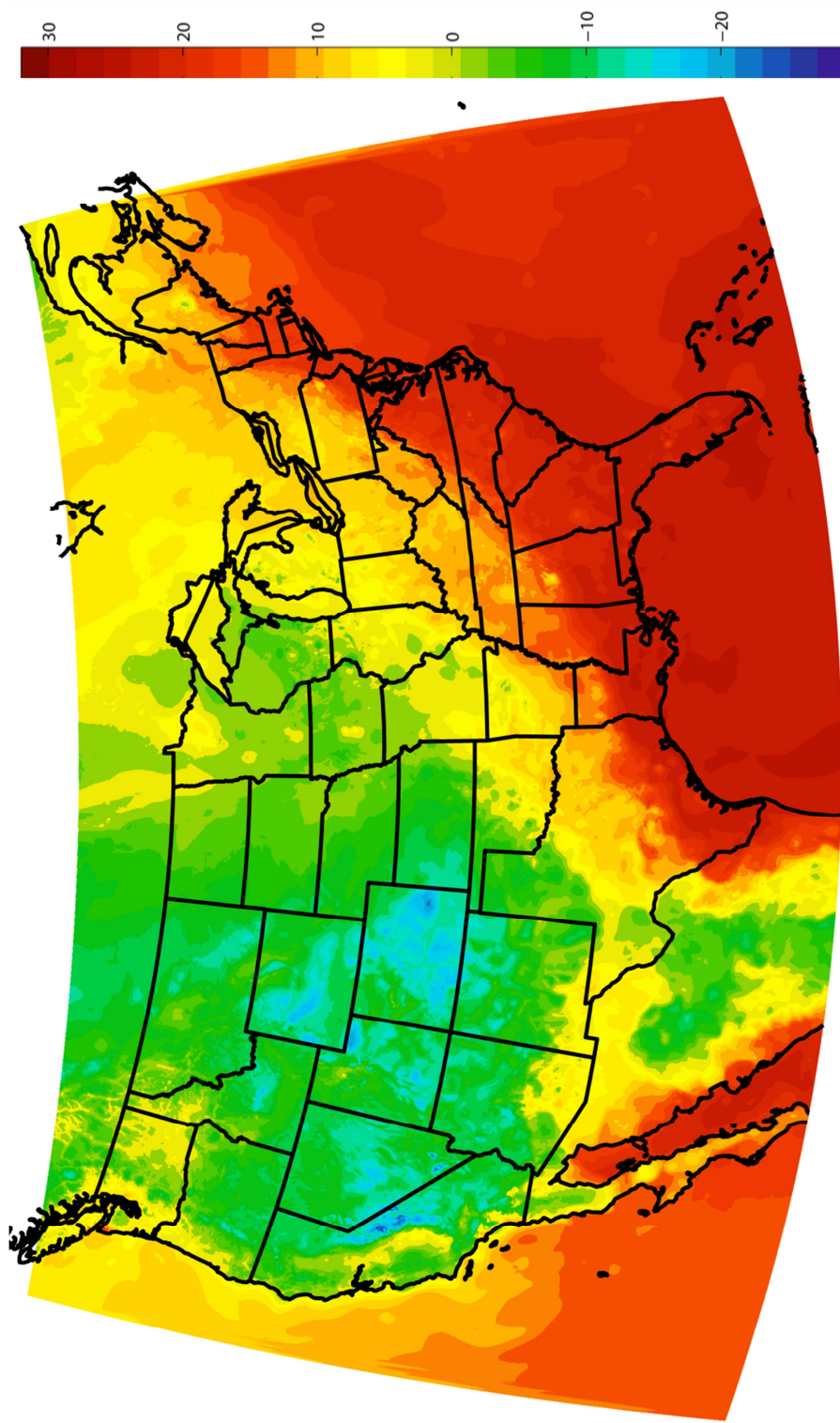


Figure 5.5. continued. d. 2-m dewpoint temperature analysis in °C.

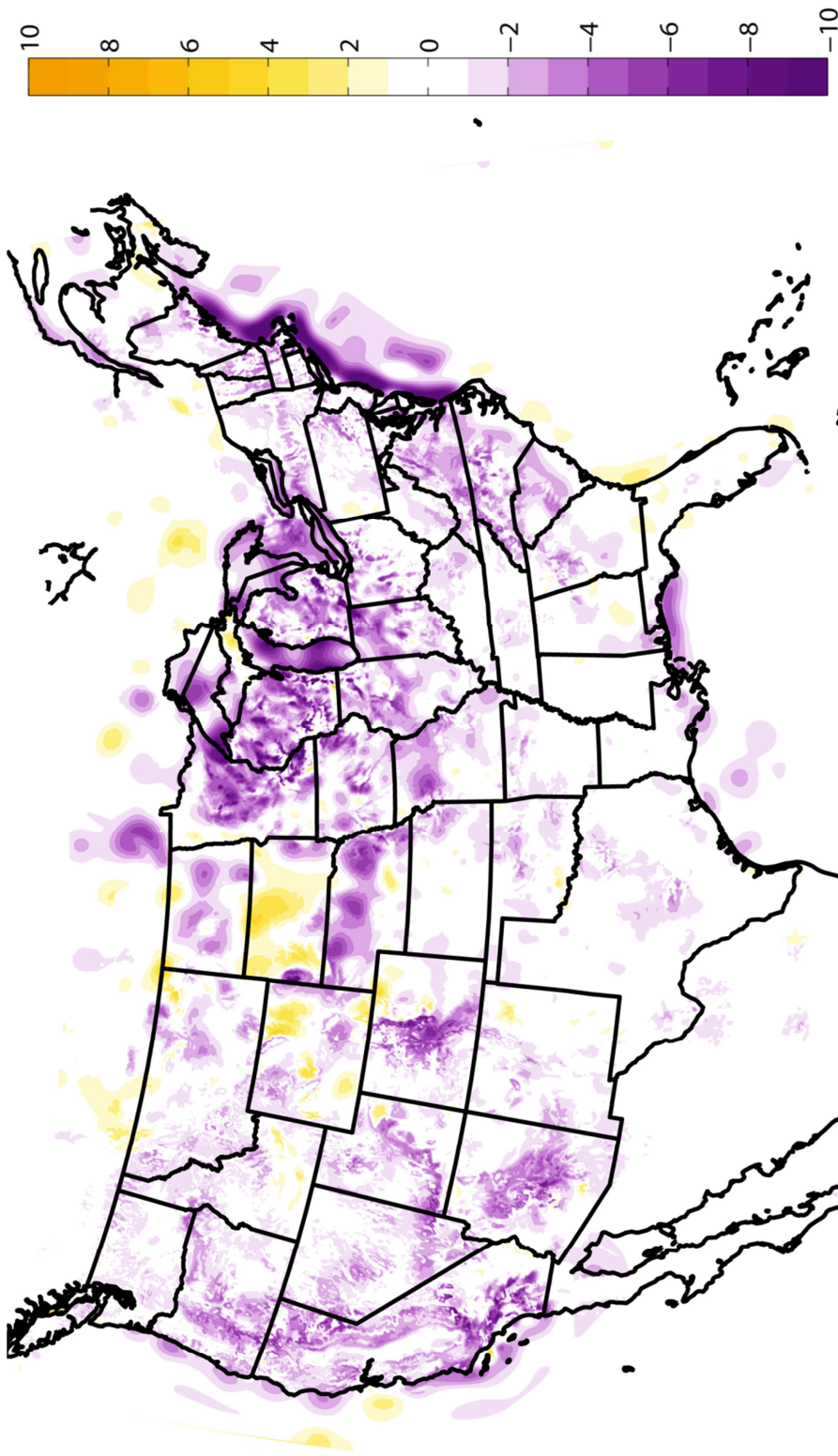


Figure 5.5. continued. e. 10-m wind speed analysis increments in m/s.

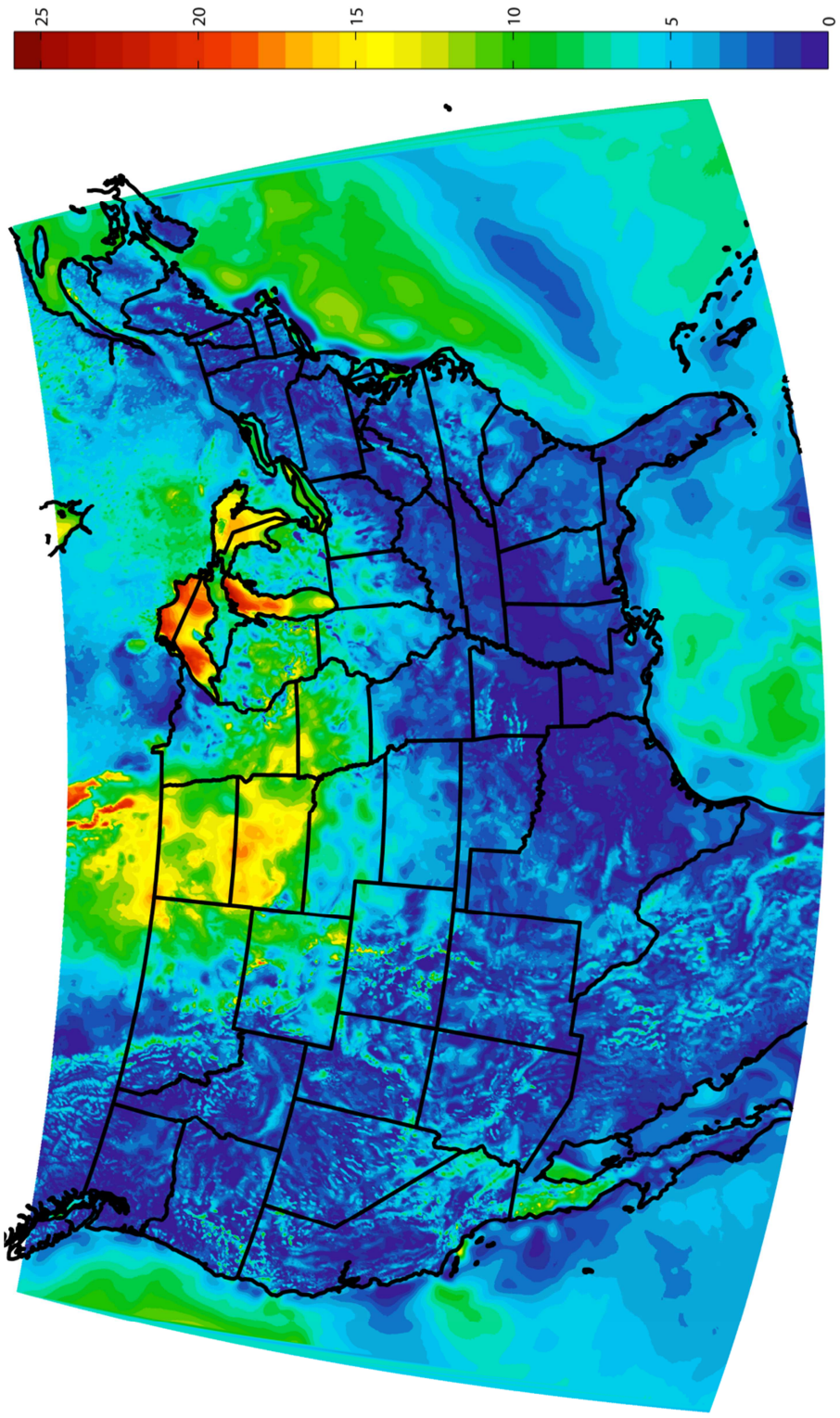


Figure 5.5. continued. f. 10-m wind speed analysis in m/s.

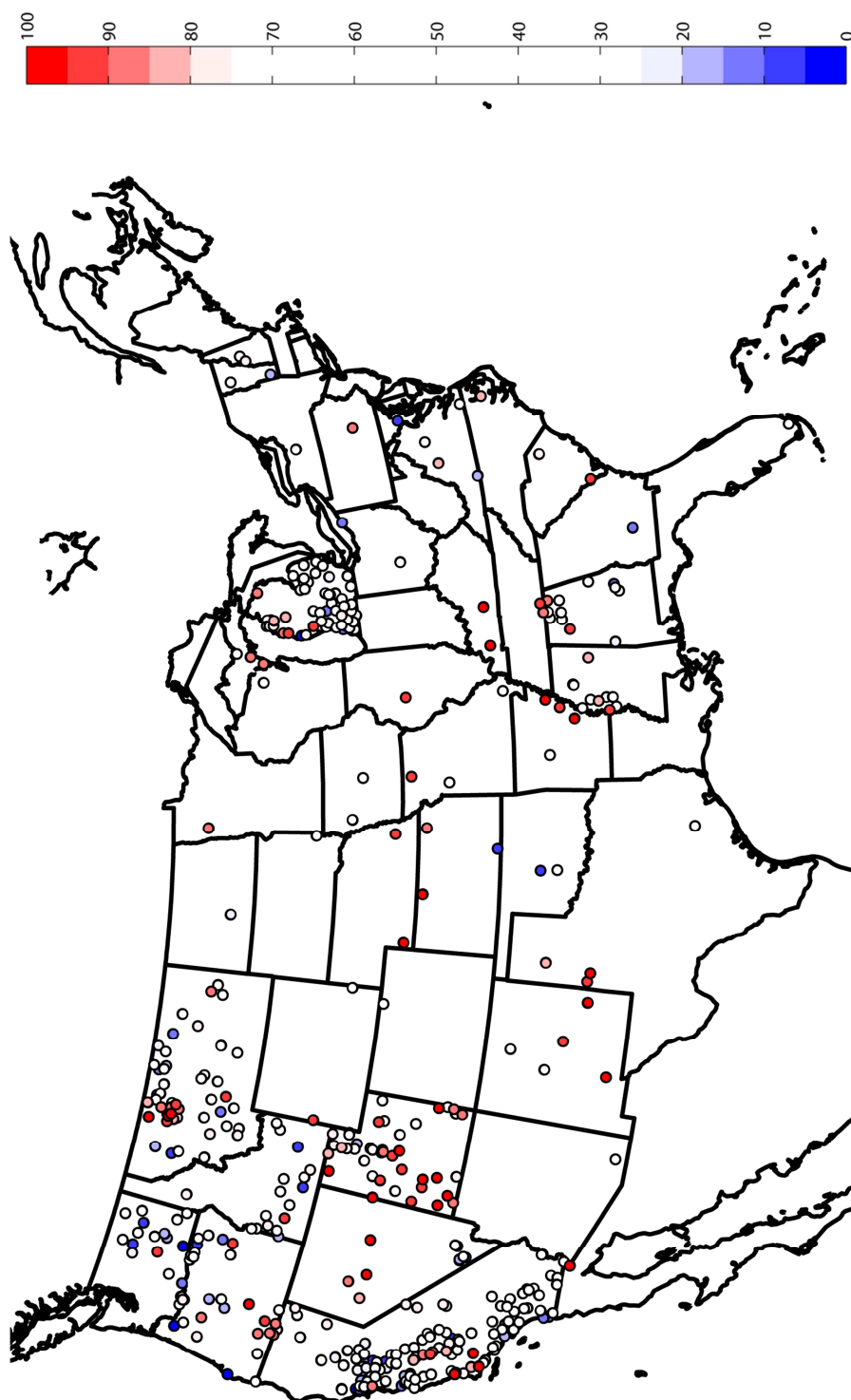


Figure 5.6. Observation impact percentile for 2-m air temperature for 1400 UTC October 2010 using constant decorrelation length scales, grouped by network category. a. AG network category.

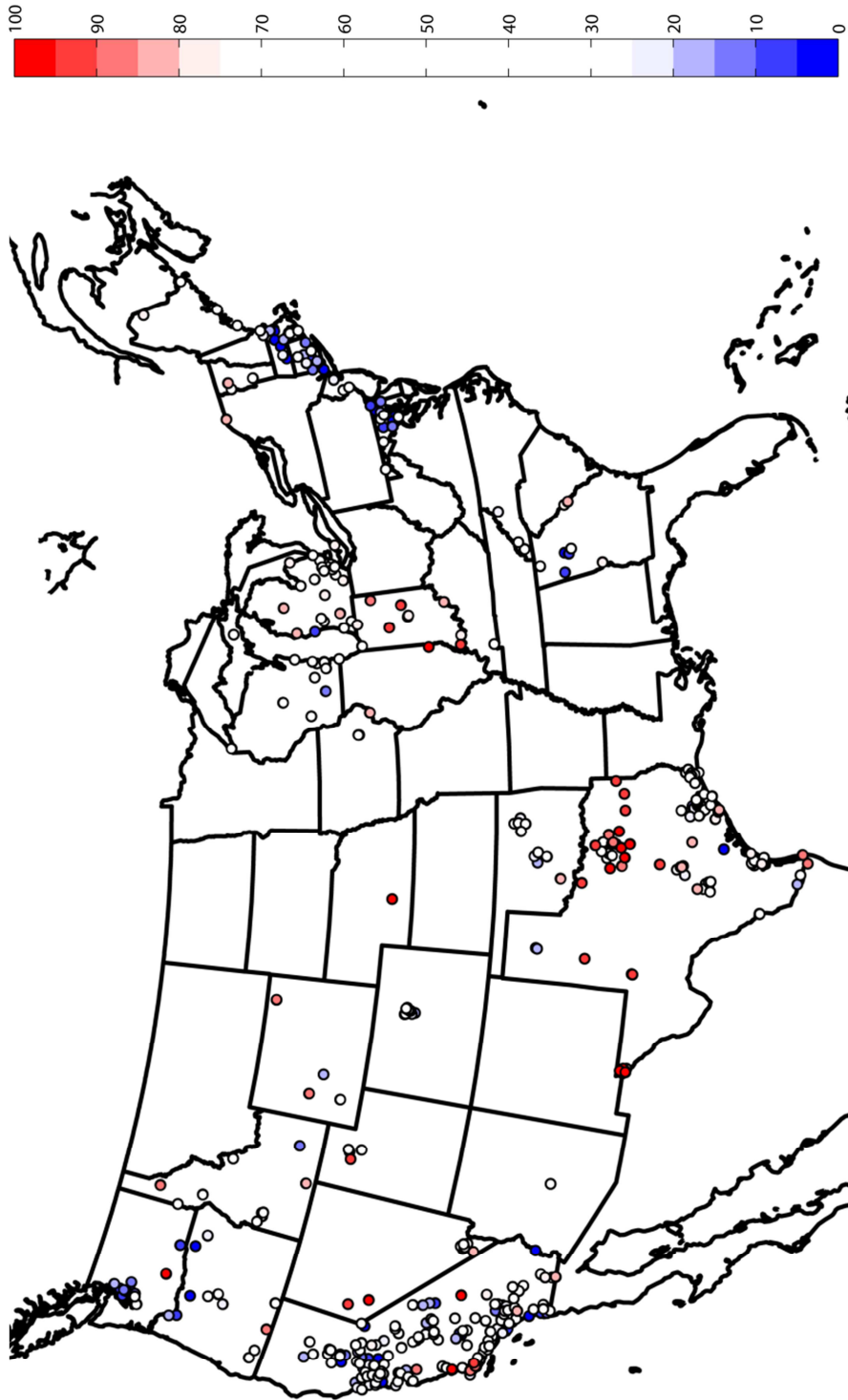


Figure 5.6. continued. b. AQ network category.

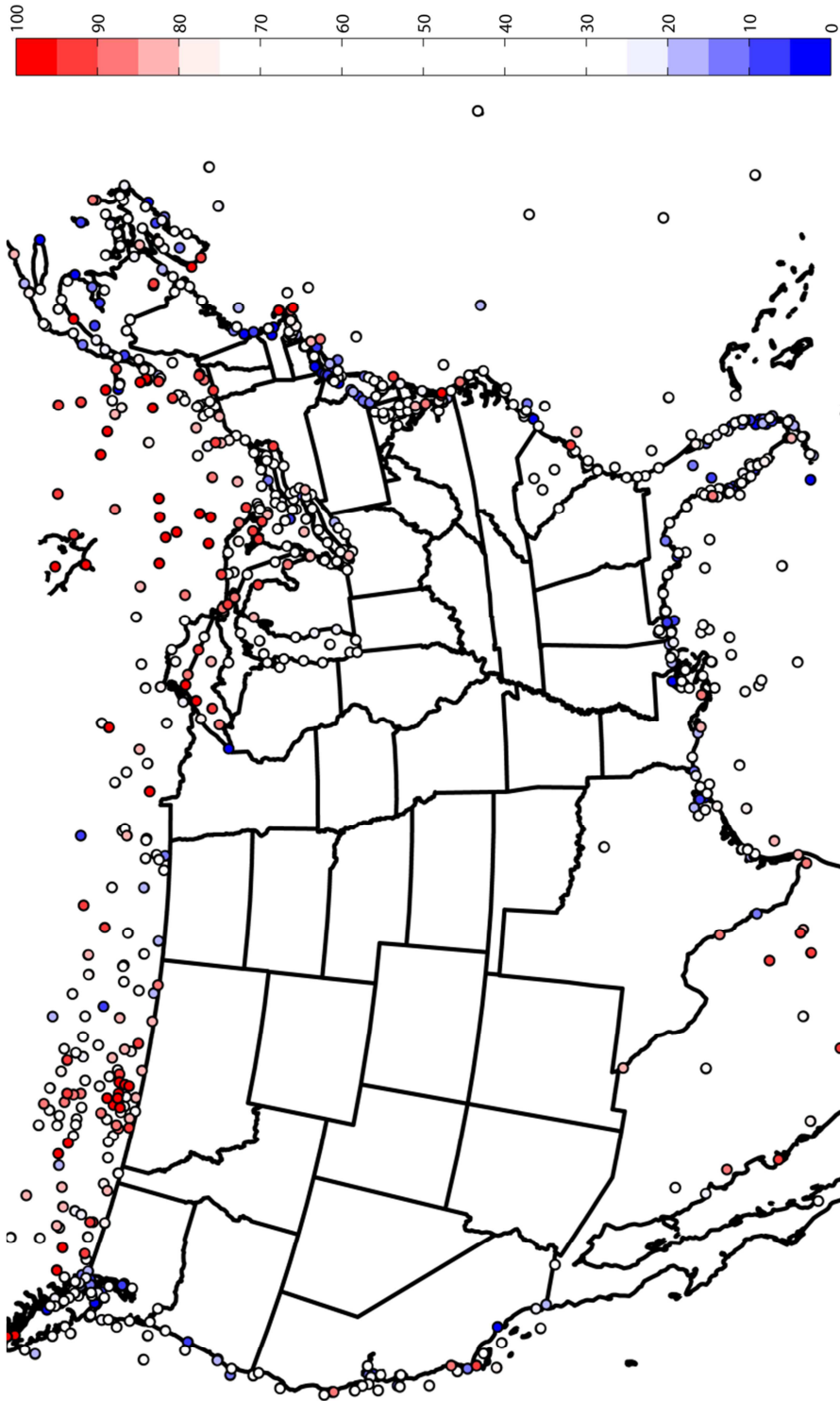


Figure 5.6. continued. c. EXT network category.

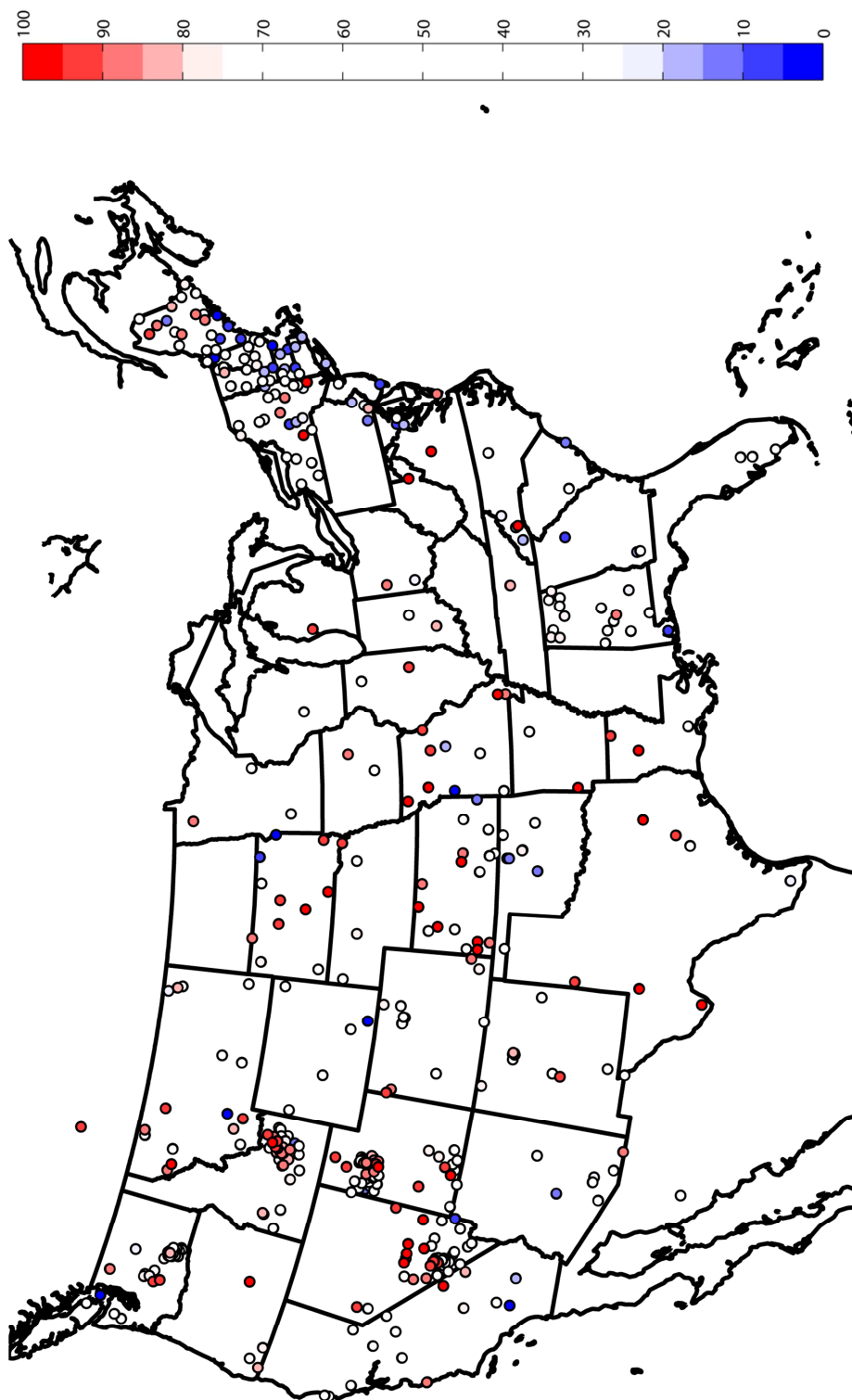


Figure 5.6. continued. d. FED network category.

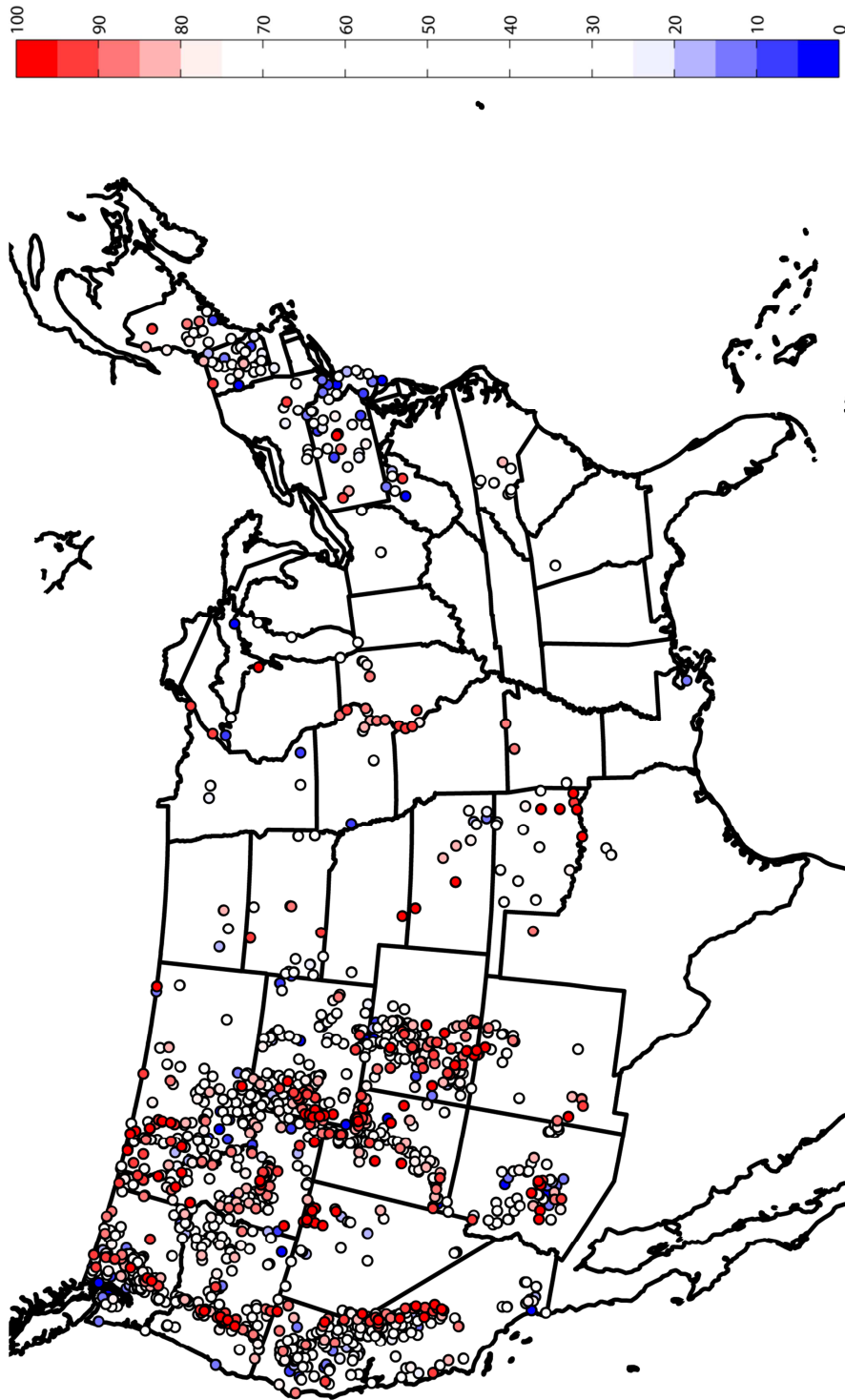


Figure 5.6. continued. e. HYDRO network category.

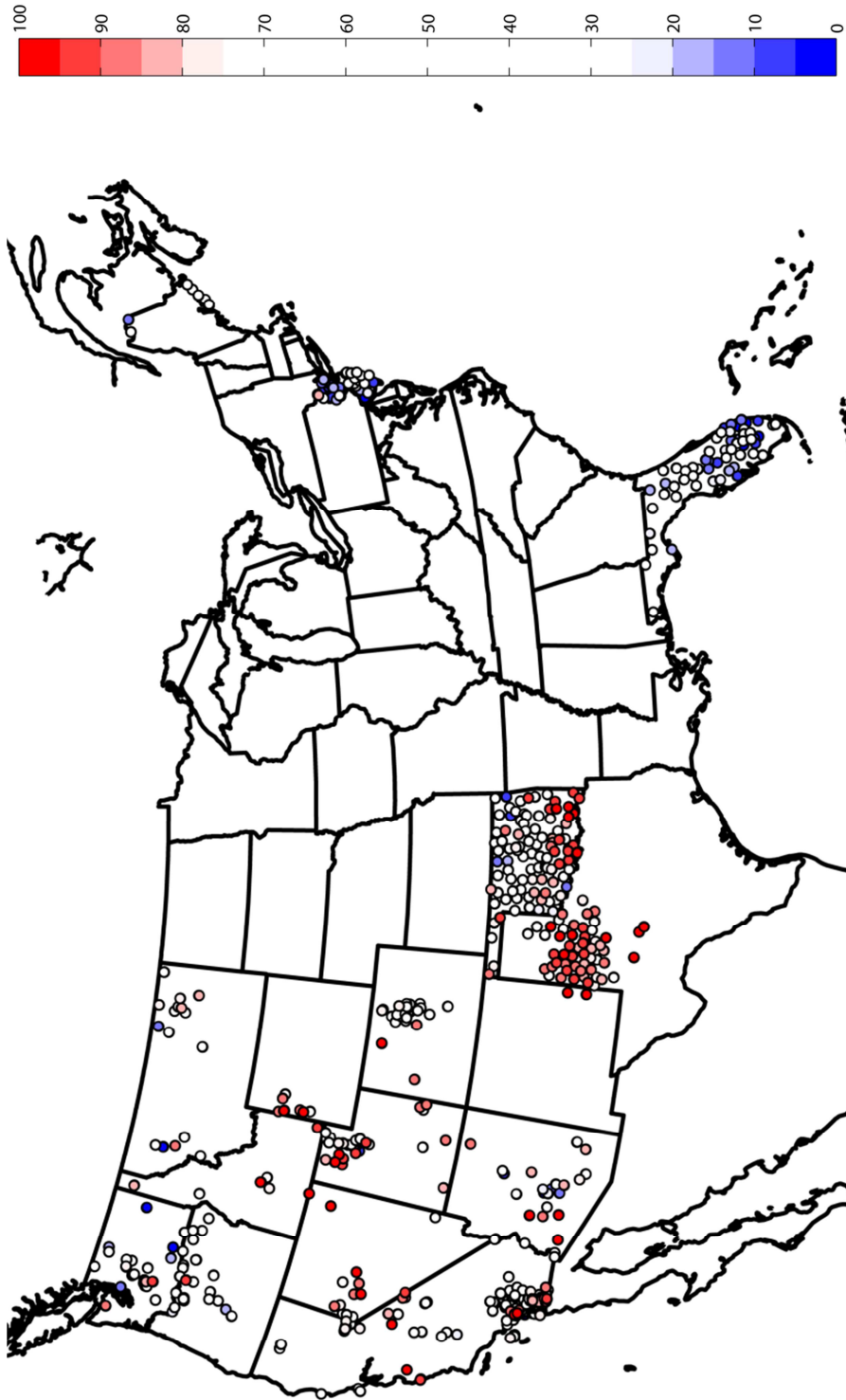


Figure 5.6. continued. f. LOCAL network category.

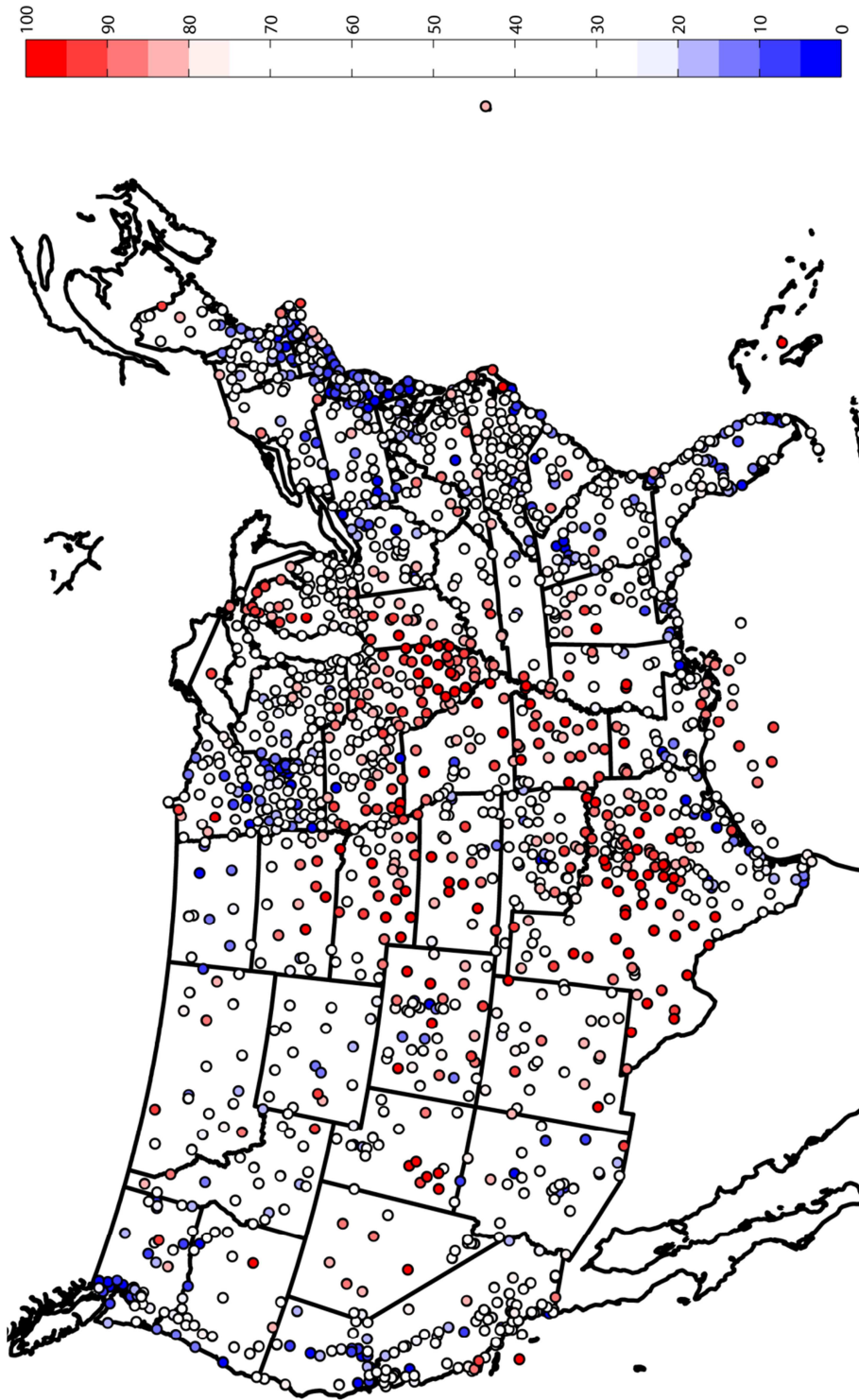


Figure 5.6. continued. g. NWS network category.

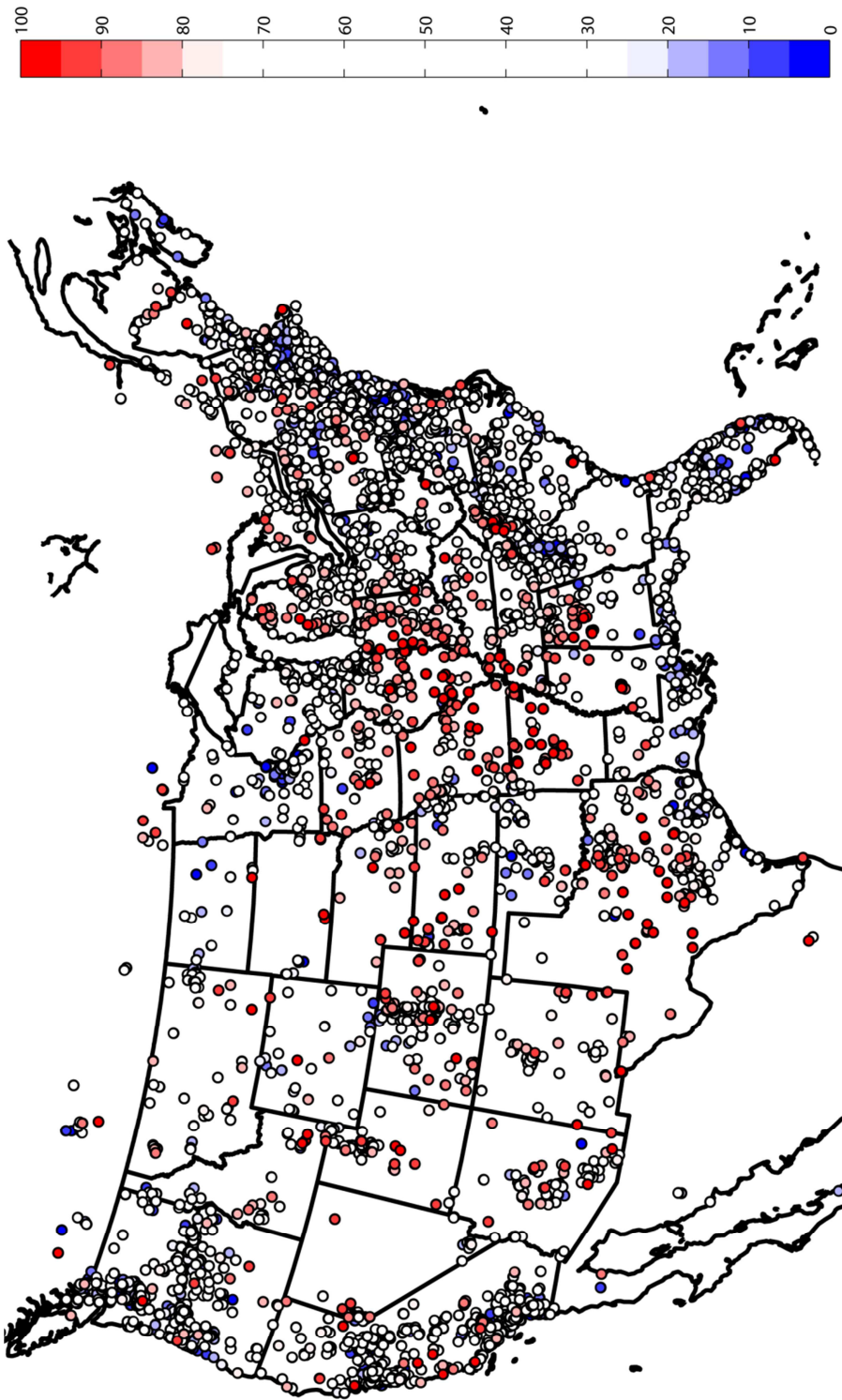


Figure 5.6. continued. h. PUBLIC network category.

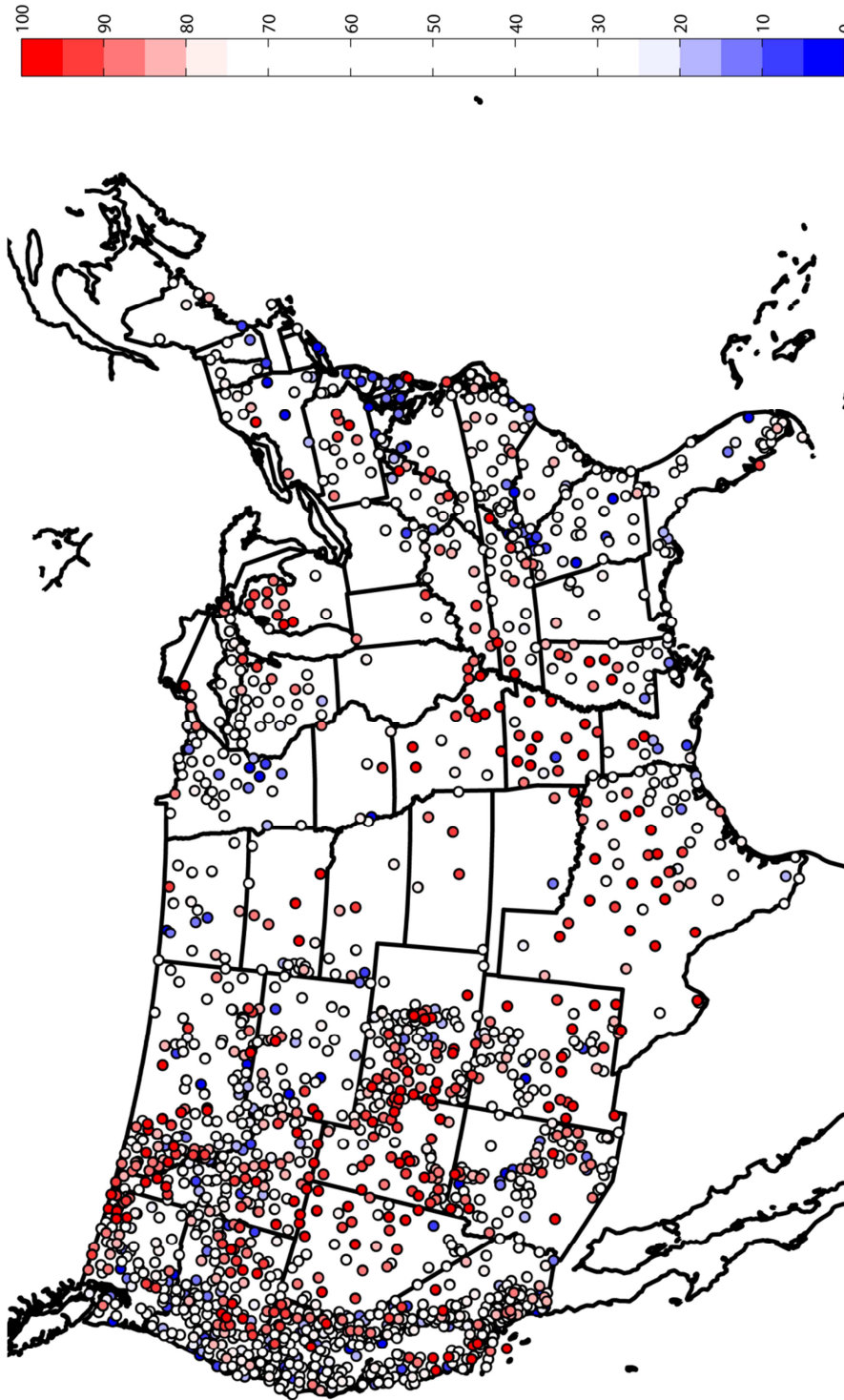


Figure 5.6. continued. i. RAWS network category.

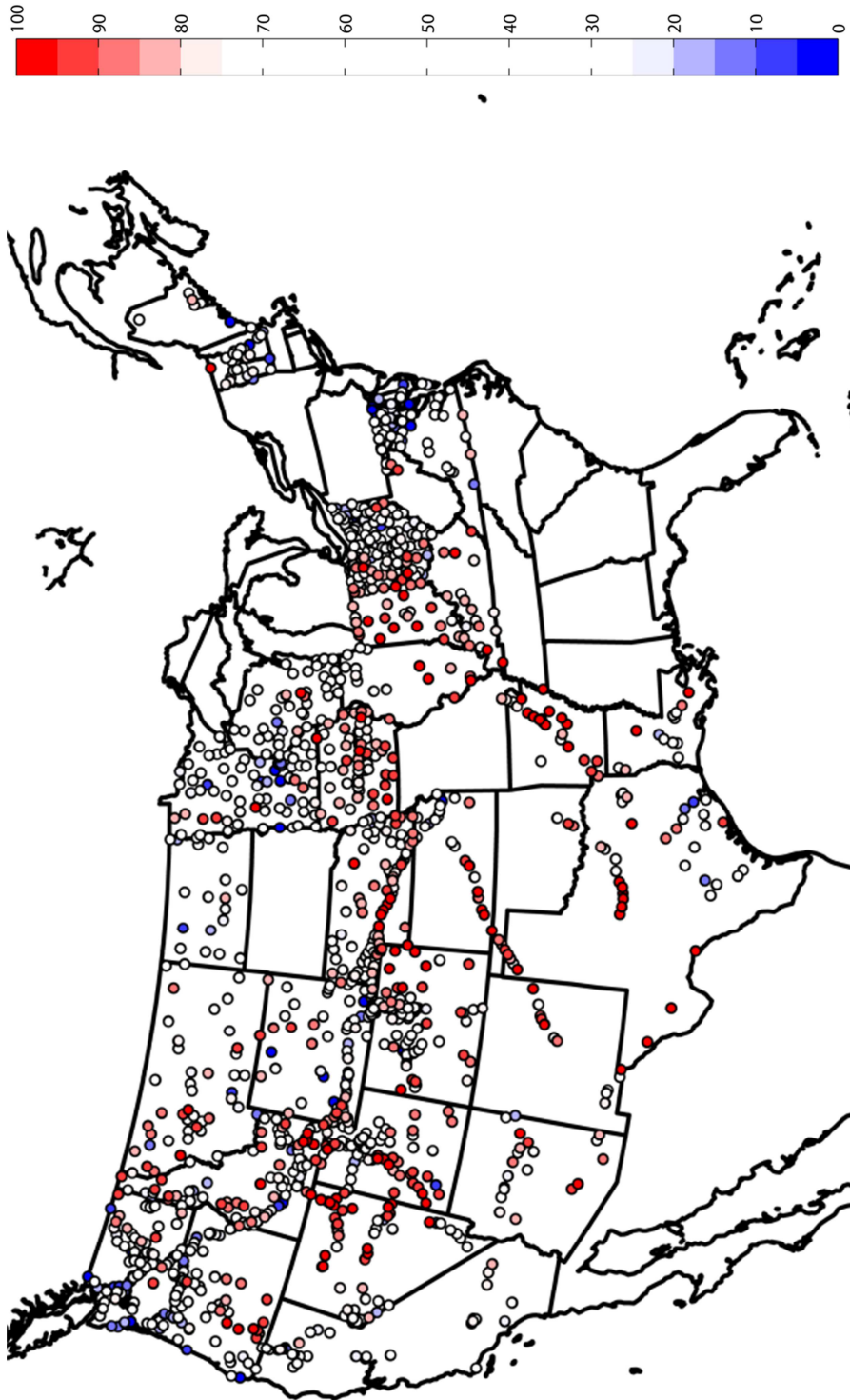


Figure 5.6. continued. j. TRANS network category.

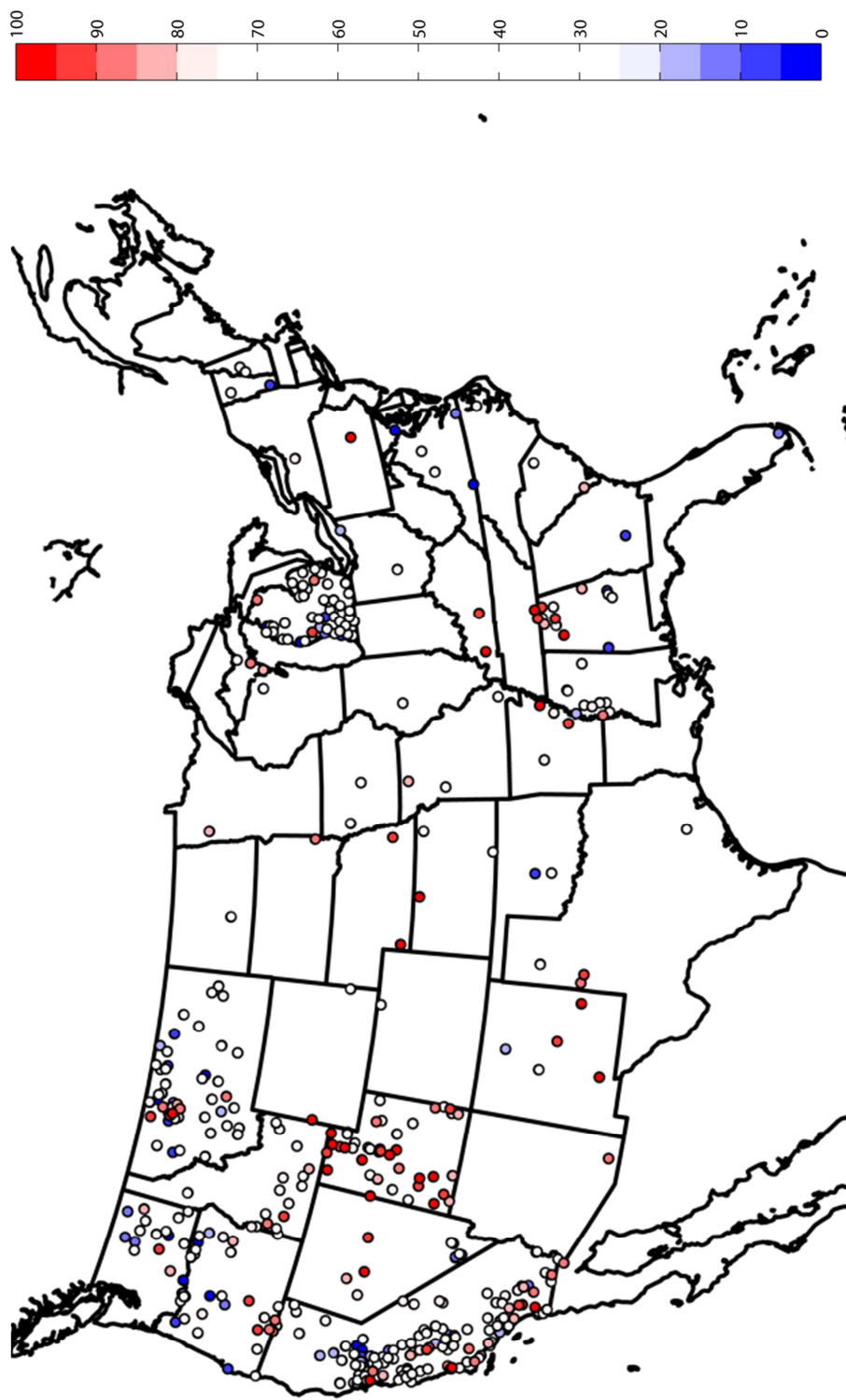


Figure 5.7. Observation impact percentile for 2-m air temperature for 1400 UTC October 2010 using the $k = 2$ polynomial to compute the vertical and horizontal decorrelation length scales, grouped by network category. a. AG network category.

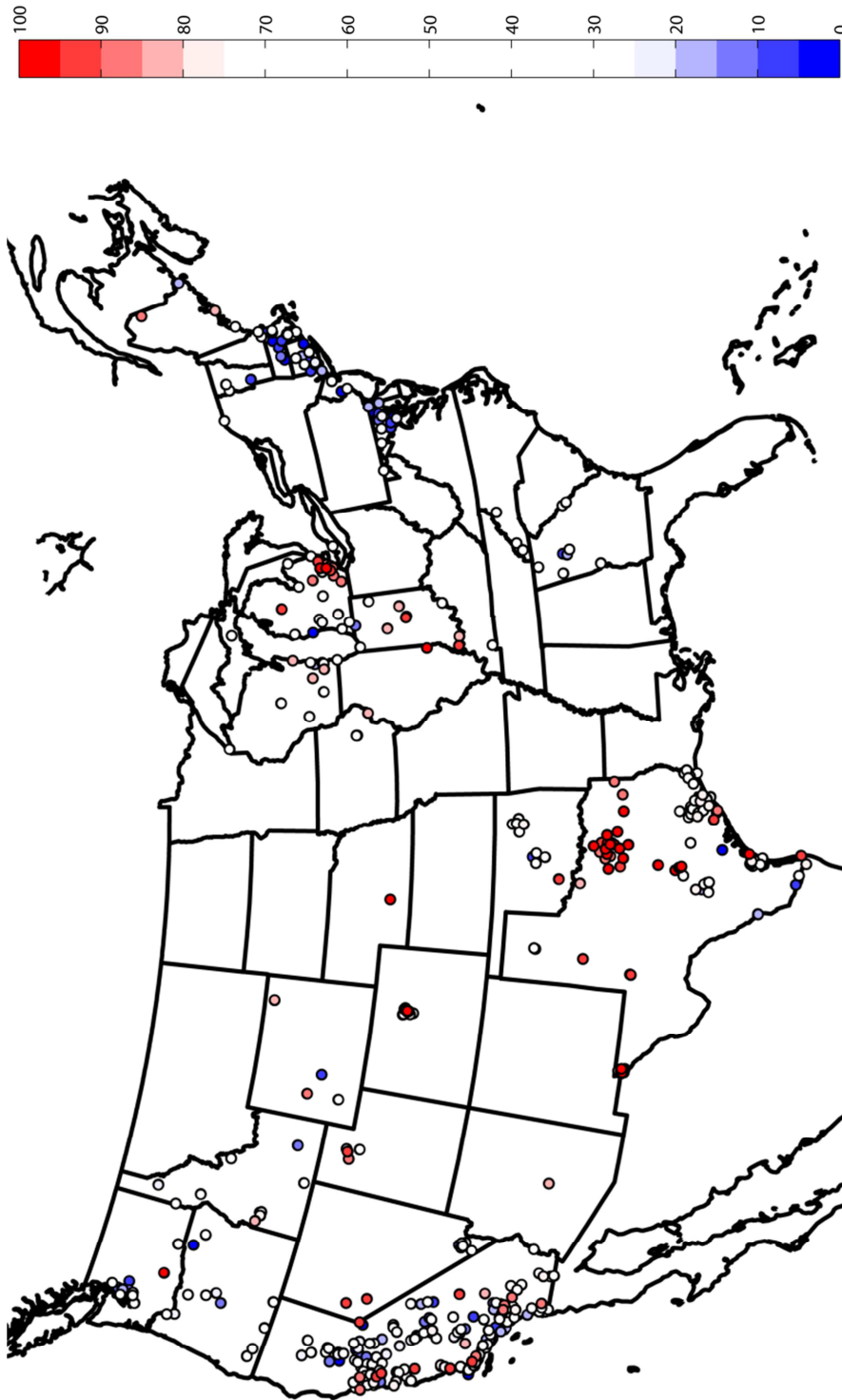


Figure 5.7. continued. b. AQ network category.

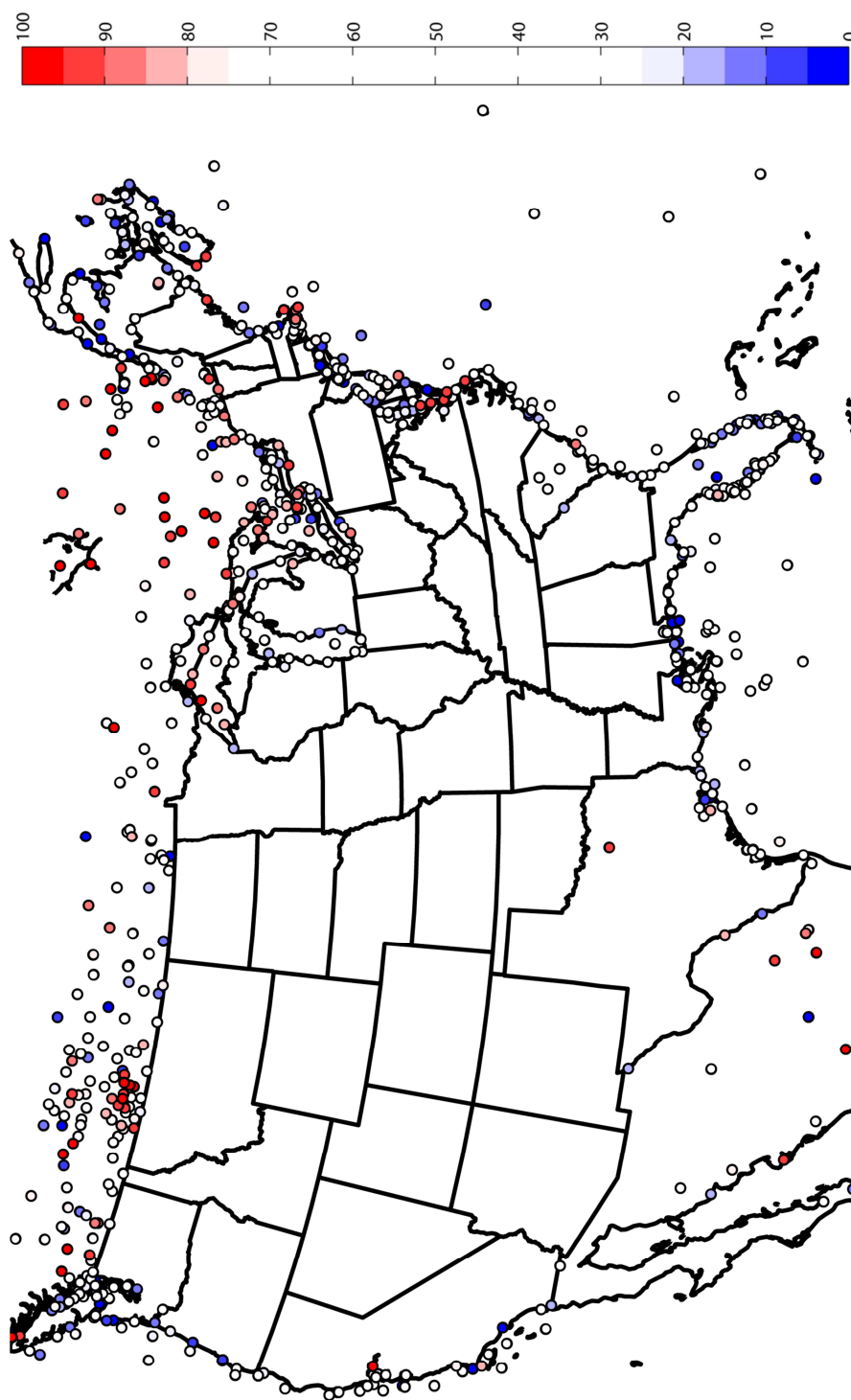


Figure 5.7. continued. c. EXT network category.

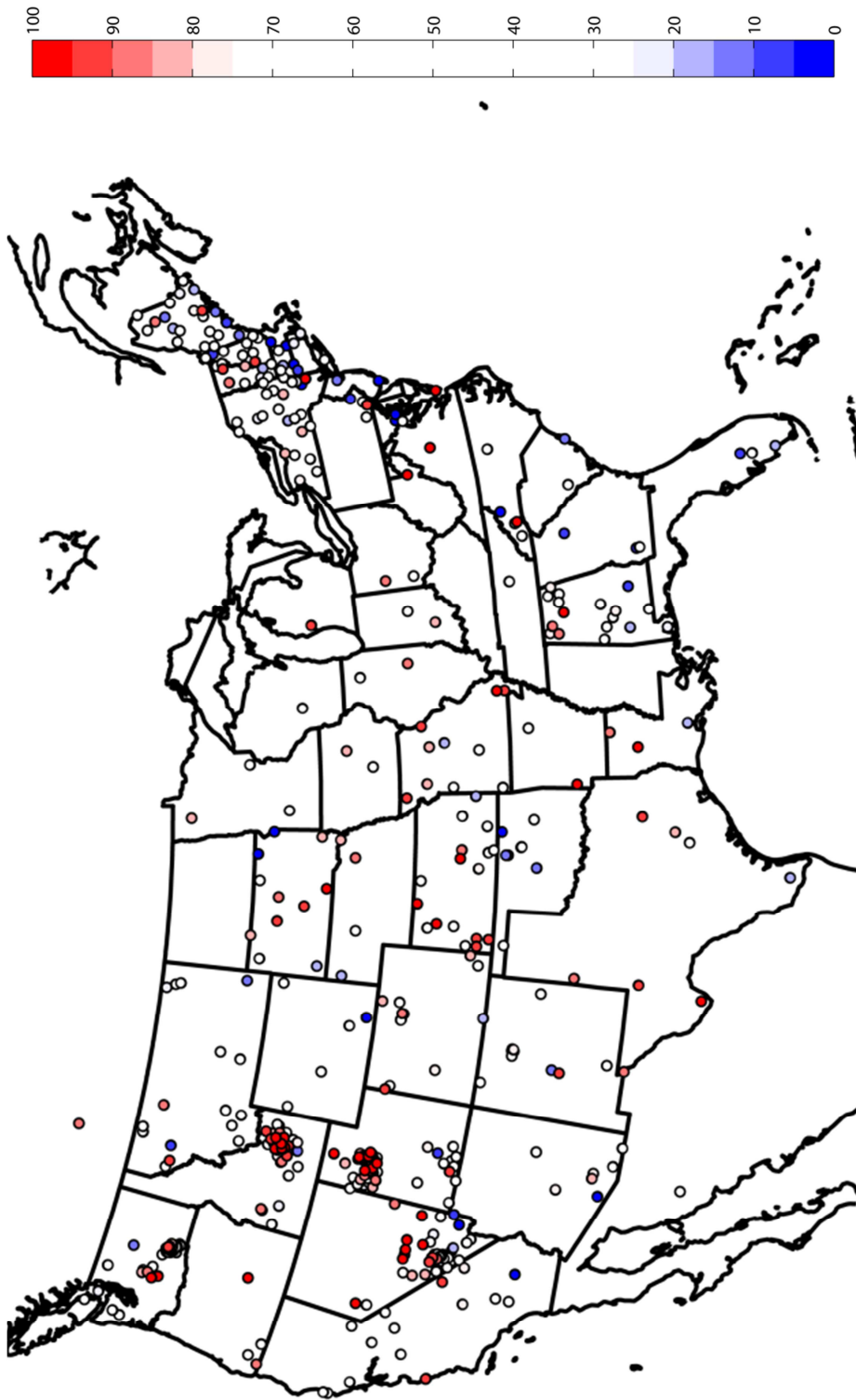


Figure 5.7. continued. d. FED network category.

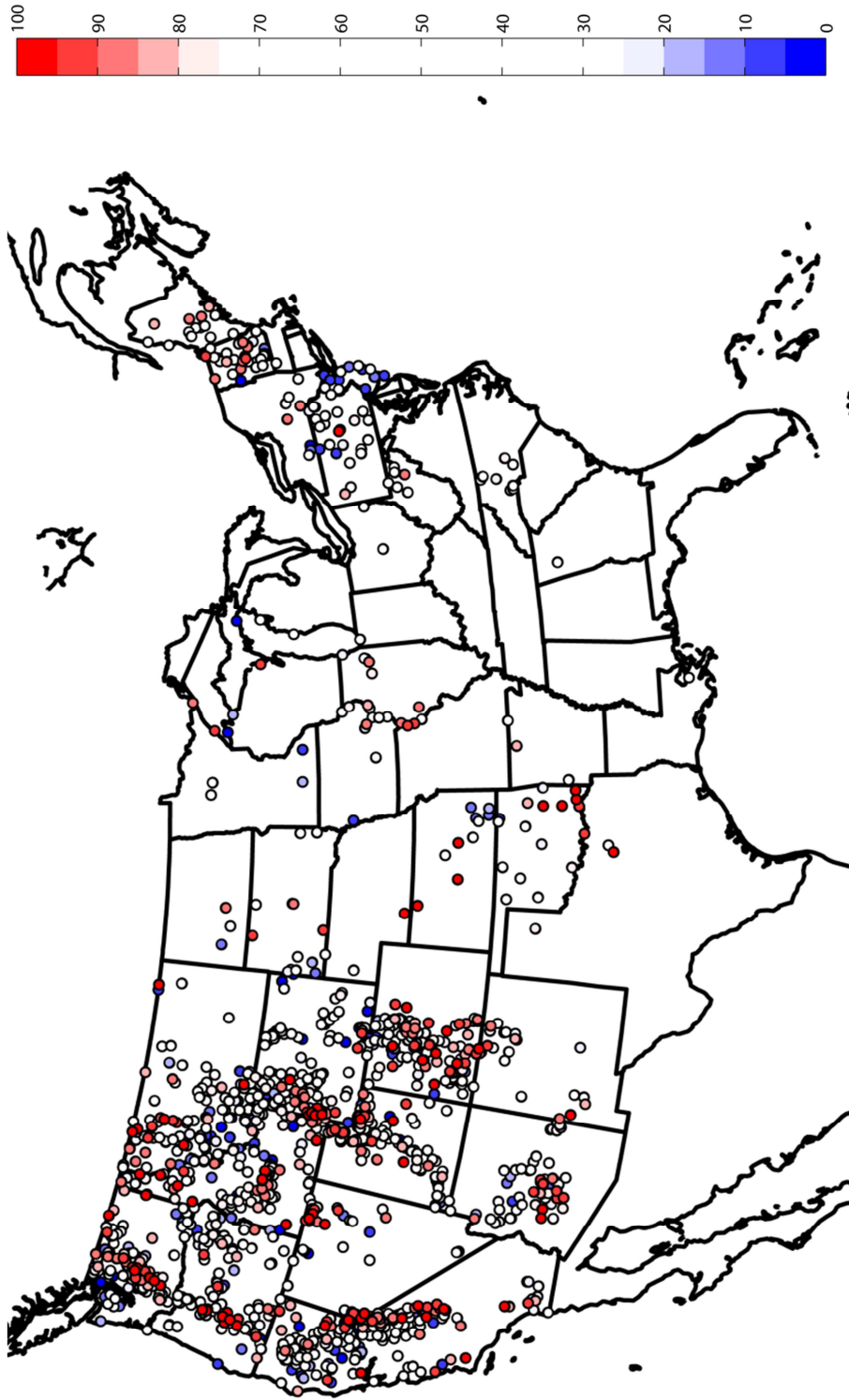


Figure 5.7. continued. e. HYDRO network category.

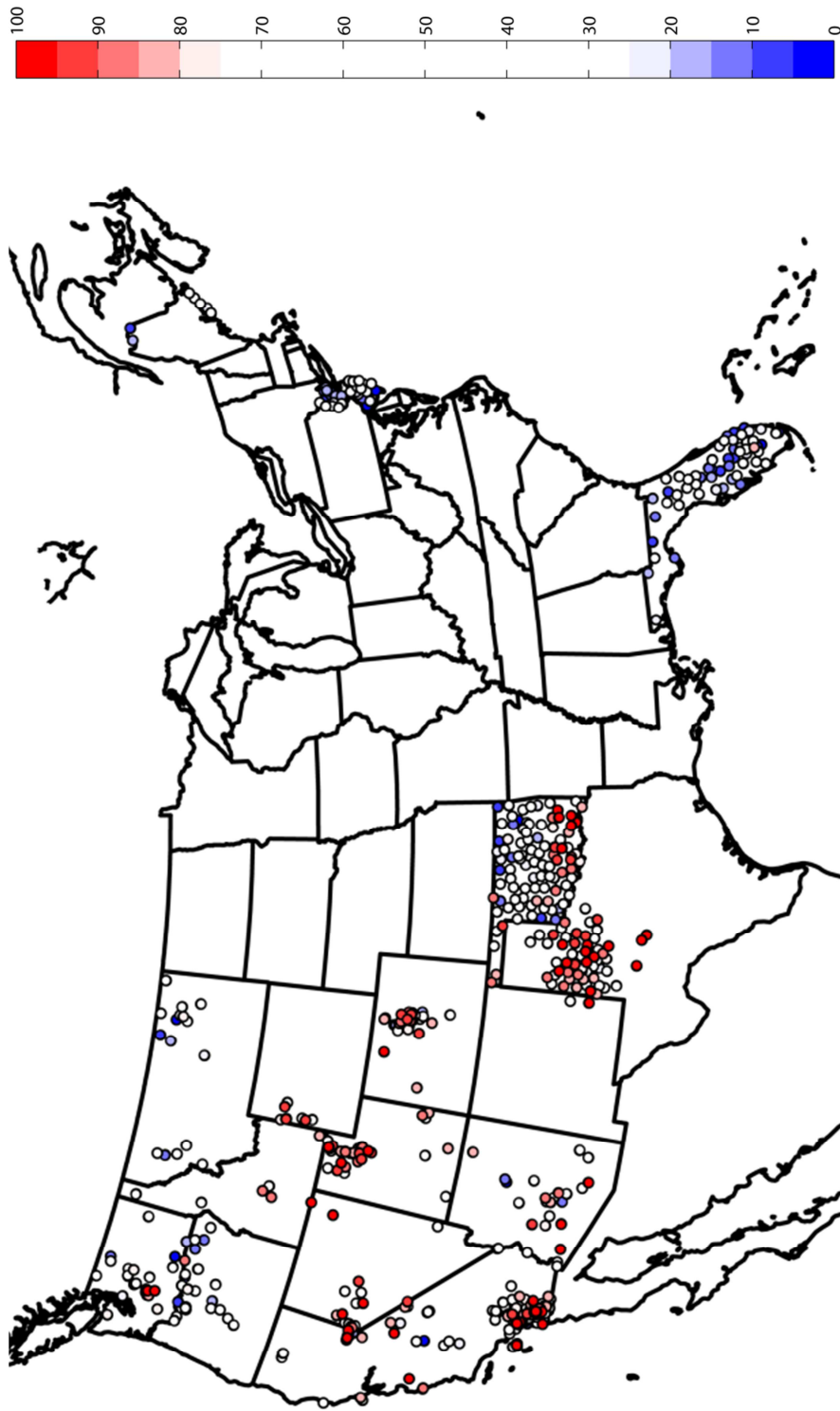


Figure 5.7. continued. f. LOCAL network category.

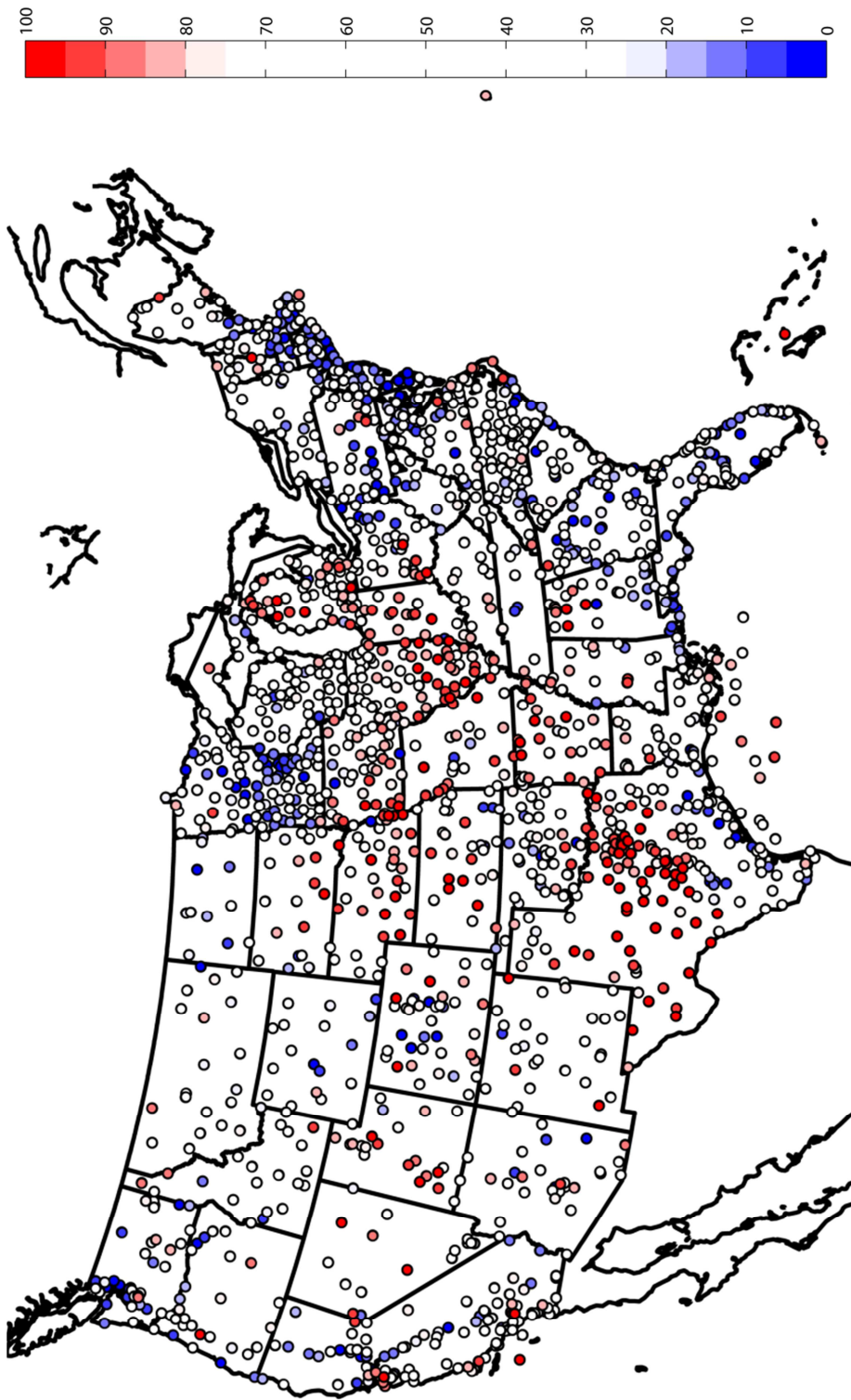


Figure 5.7. continued. g. NWS network category.

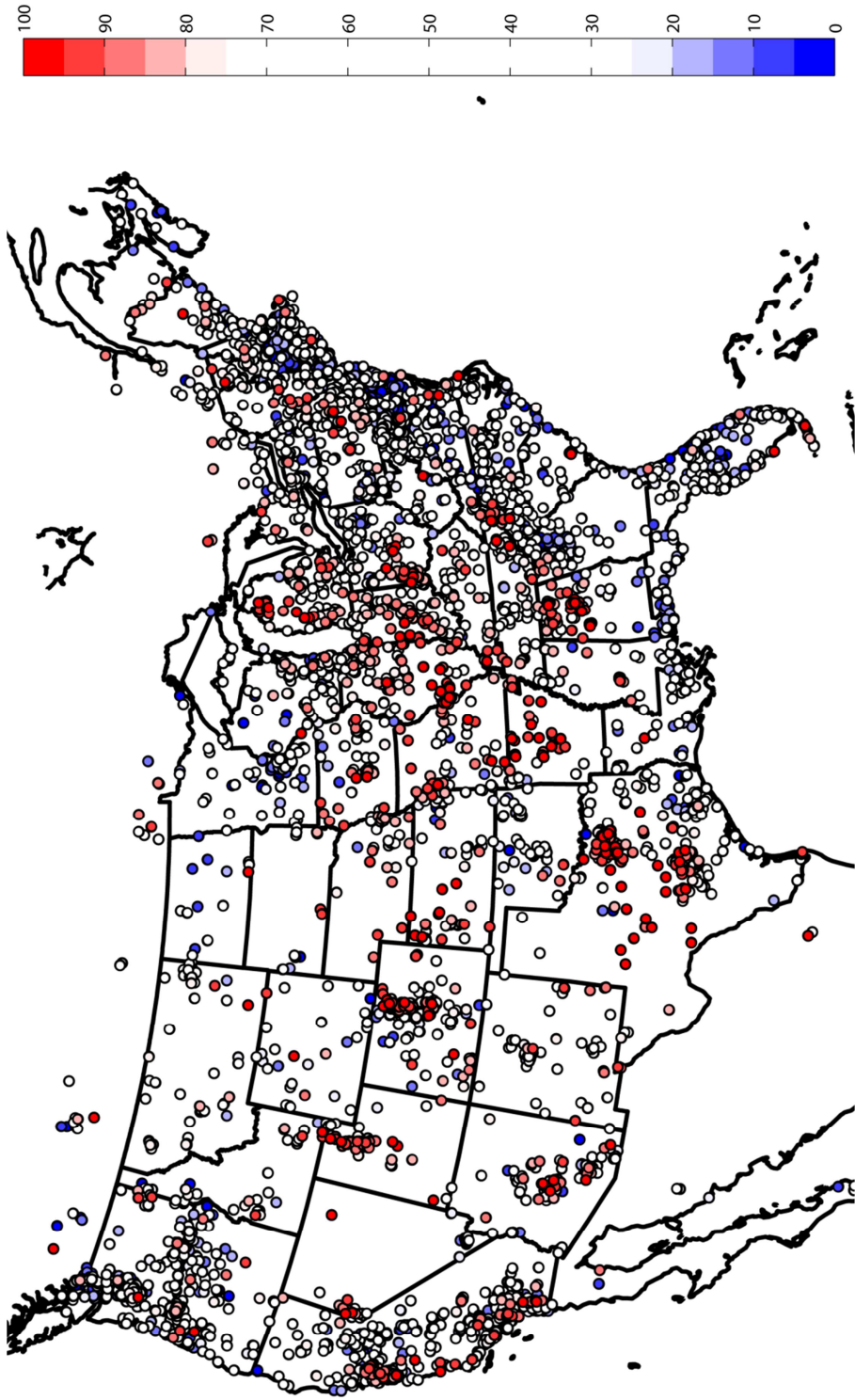


Figure 5.7. continued. h. PUBLIC network category.

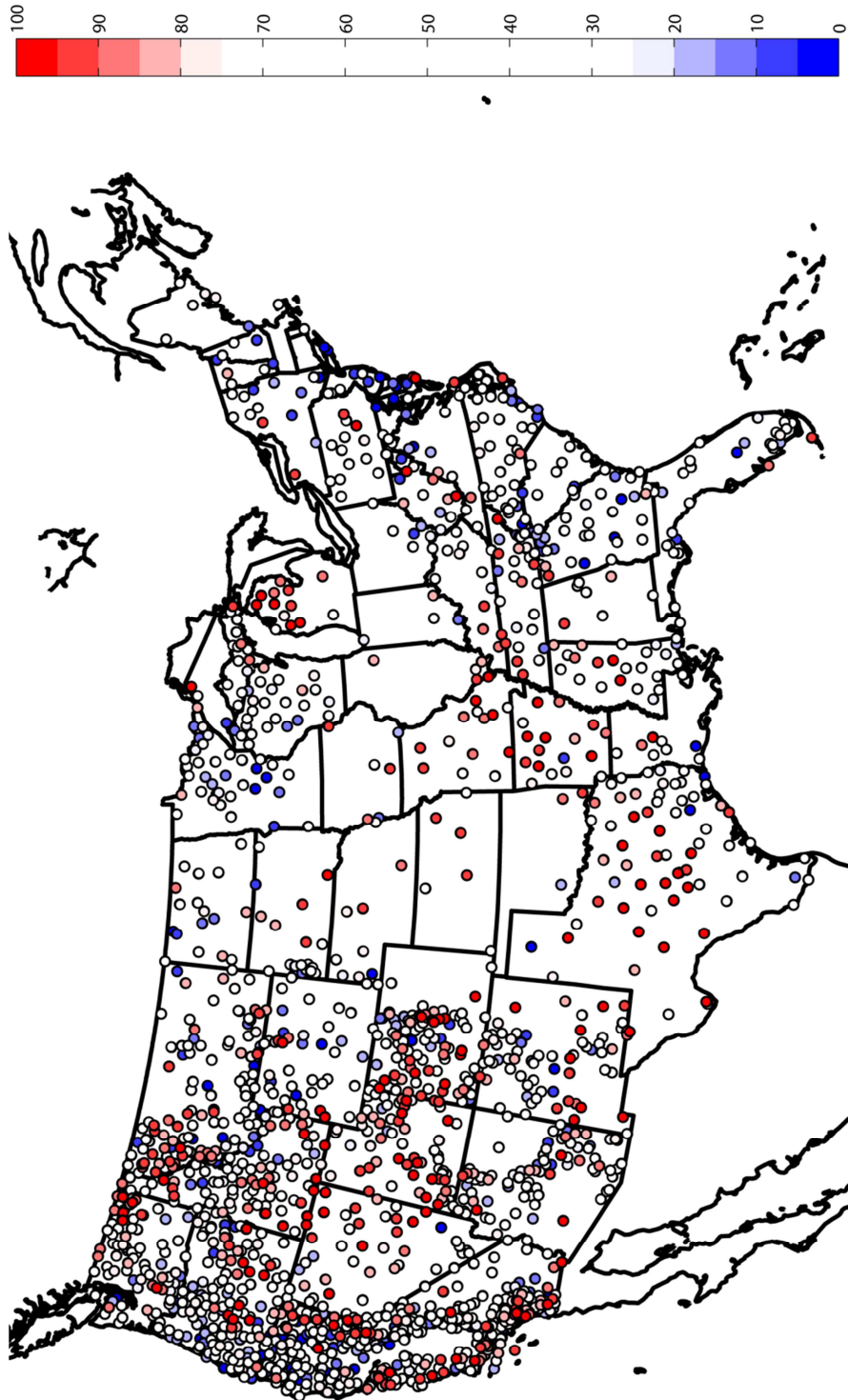


Figure 5.7. continued. i. RAWS network category.

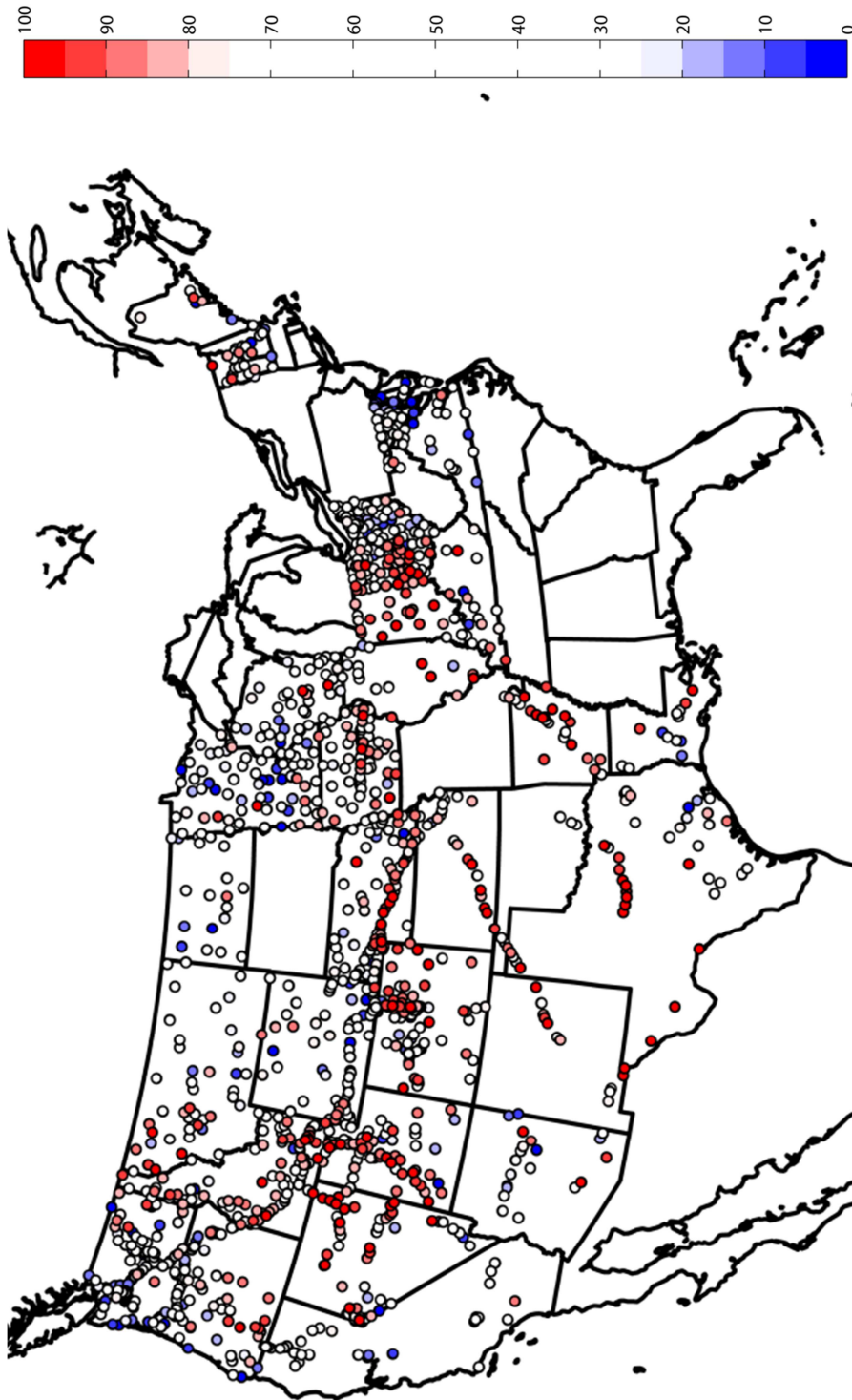


Figure 5.7. continued. j. TRANS network category.

2-m Air Temperature Observation Observation Impact

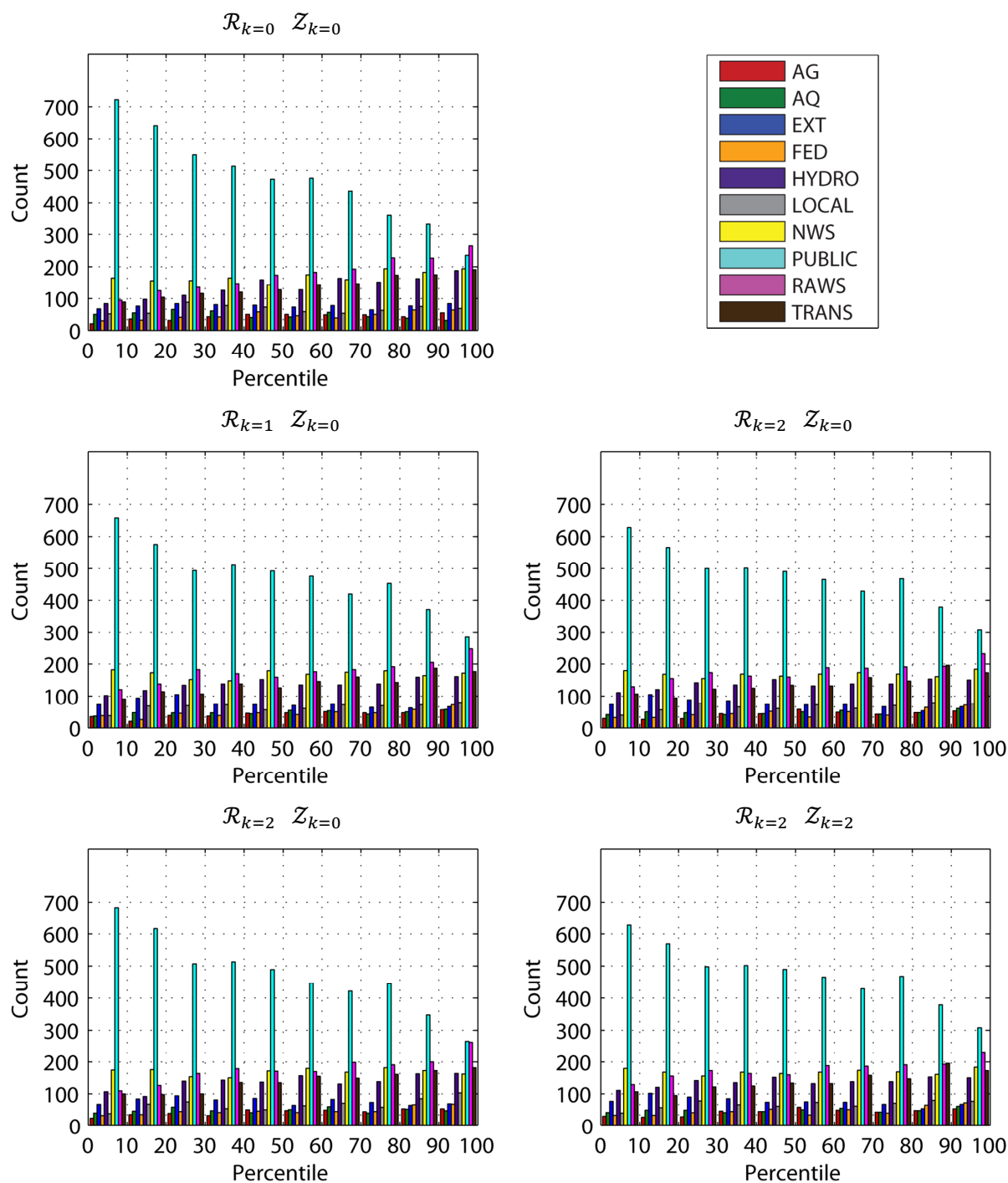


Figure 5.8. Observation impact percentile distribution by network category for 1400 UTC 27 October 2010 for temperature, dewpoint, and wind speeds for all 5 specifications of the background error covariance studied in this research. a. 2-m air temperature observation impact percentile distribution.

2-m Dewpoint Temperature Observation Impact

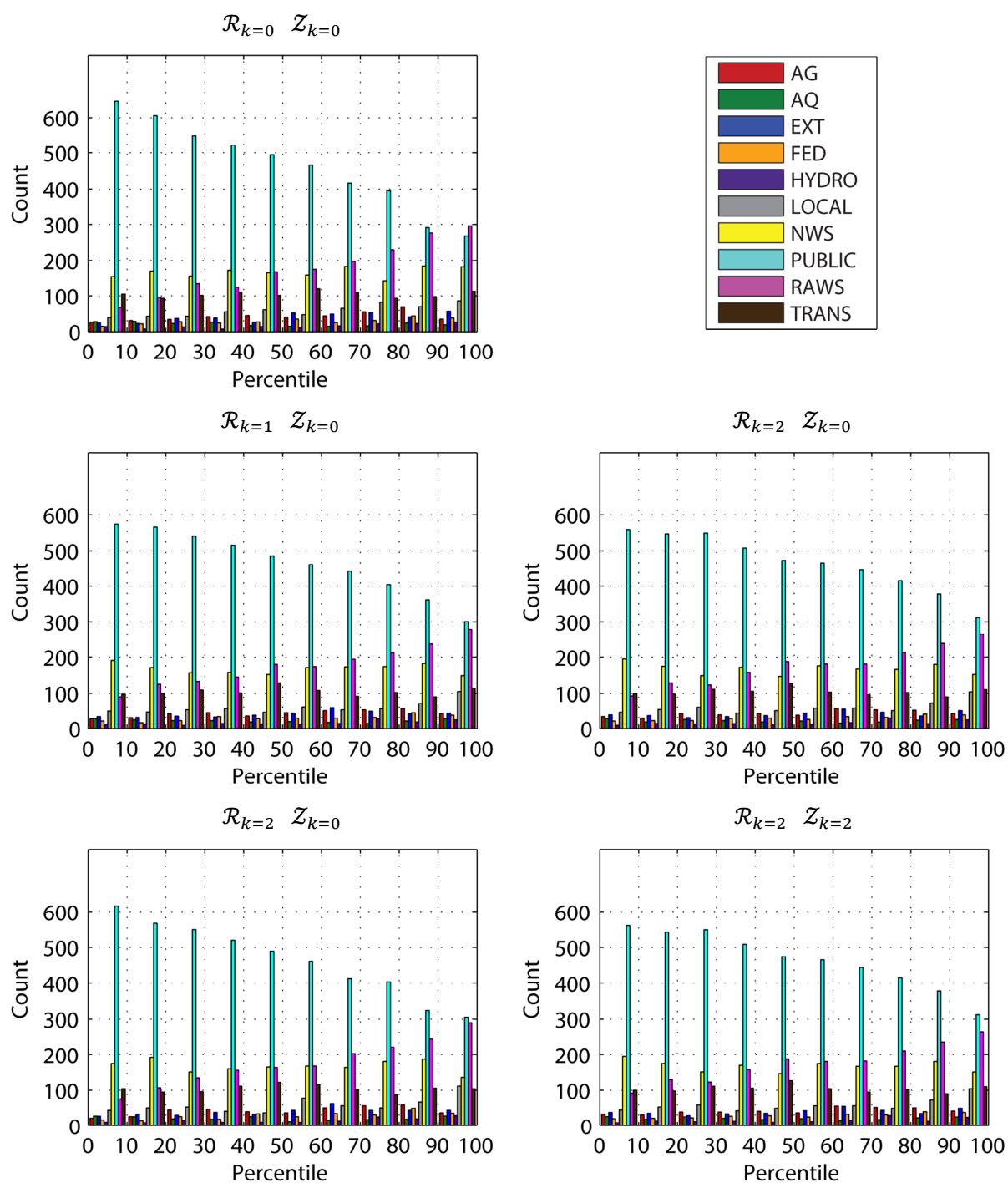


Figure 5.8. continued. b. 2-m dewpoint temperature observation impact percentile distribution.

10-m Wind Speed Observation Impact

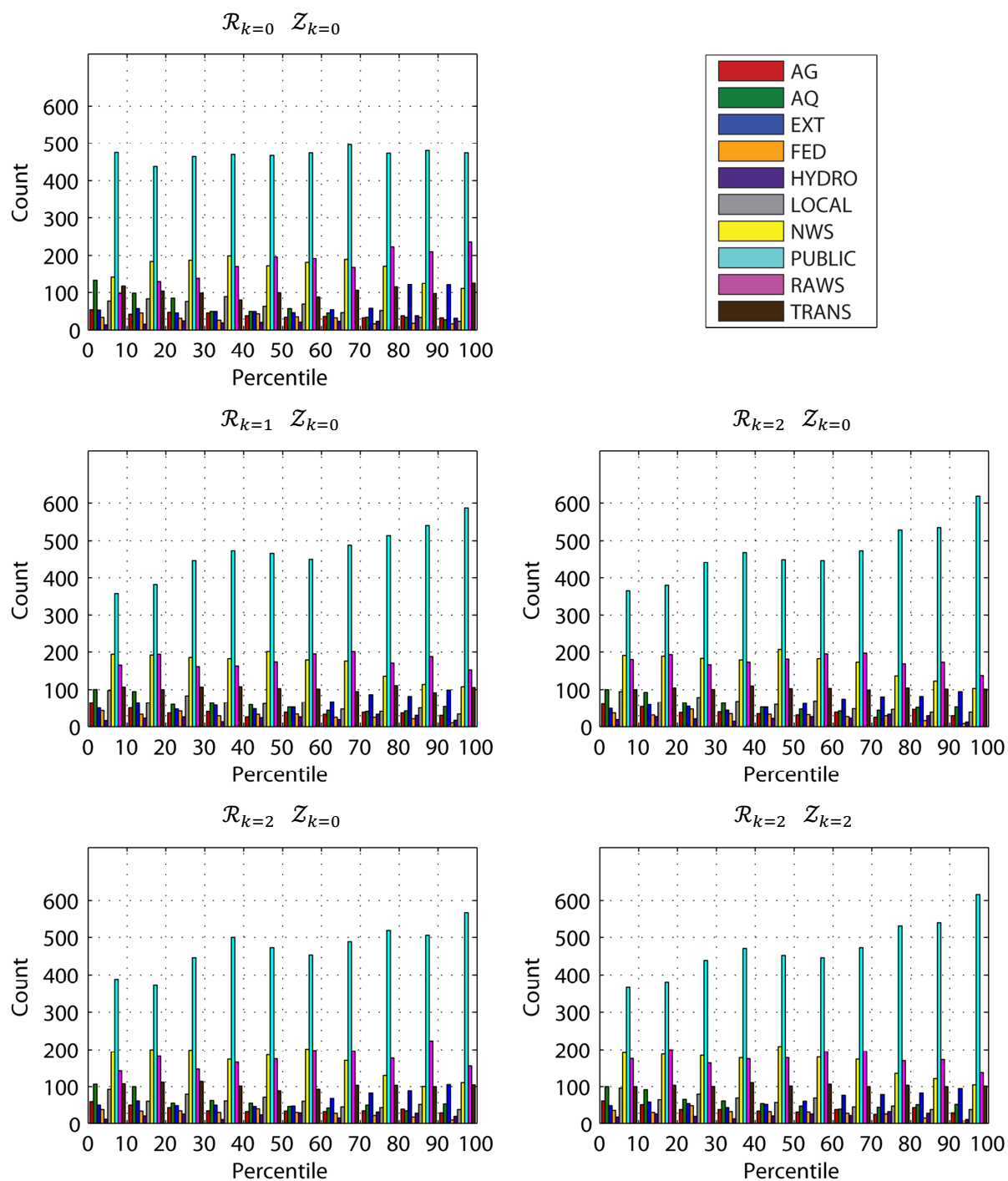


Figure 5.8. continued. c. 10-m wind speed observation impact percentile distribution.

2-m Air Temperature Observation Sensitivity

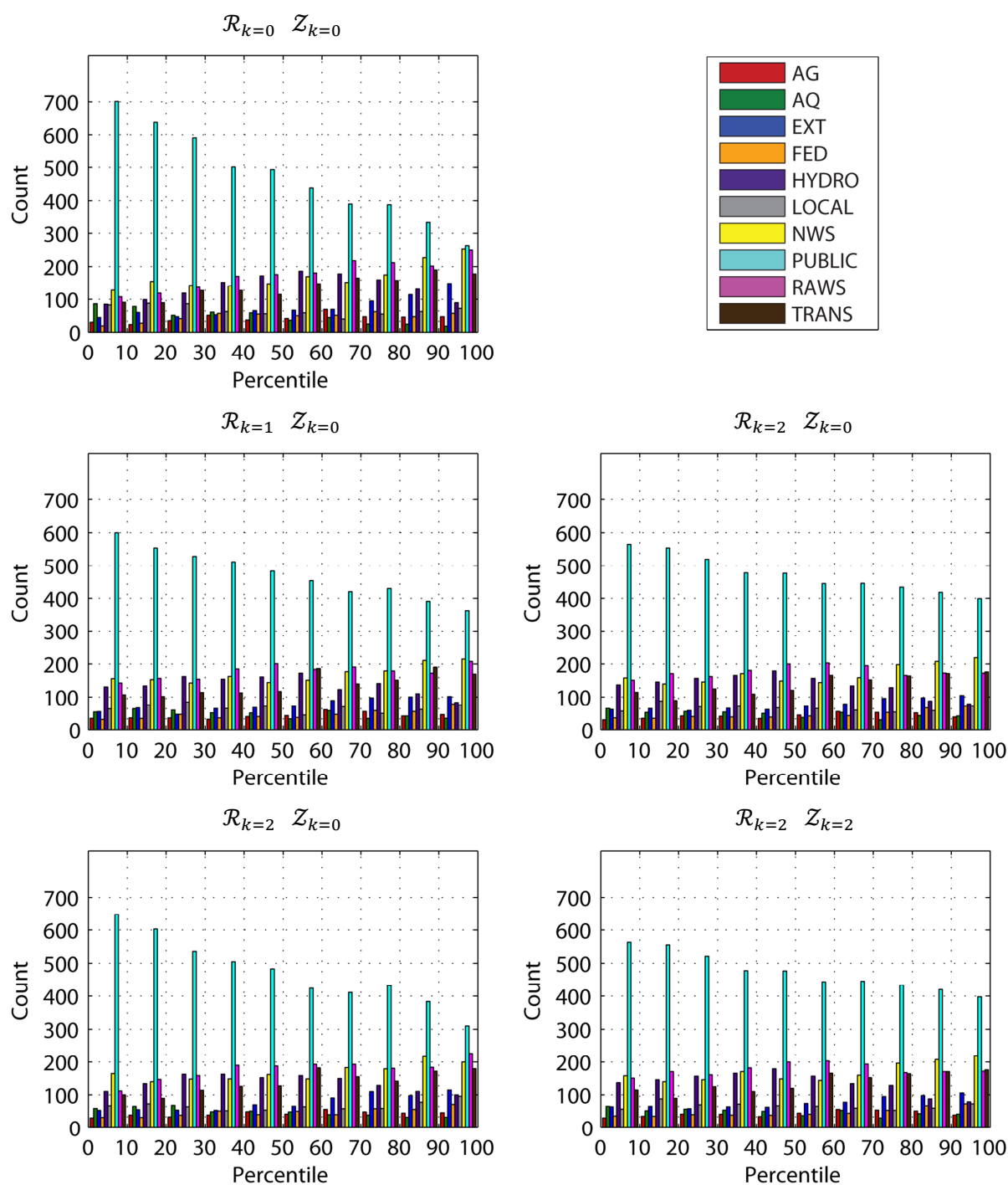


Figure 5.9. Observation sensitivity percentile distribution by network category for 1400 UTC 27 October 2010 for temperature, dewpoint, and wind speeds for all 5 specifications of the background error covariance studied in this research. a. 2-m air temperature observation sensitivity percentile distribution.

2-m Dewpoint Temperature Observation Sensitivity

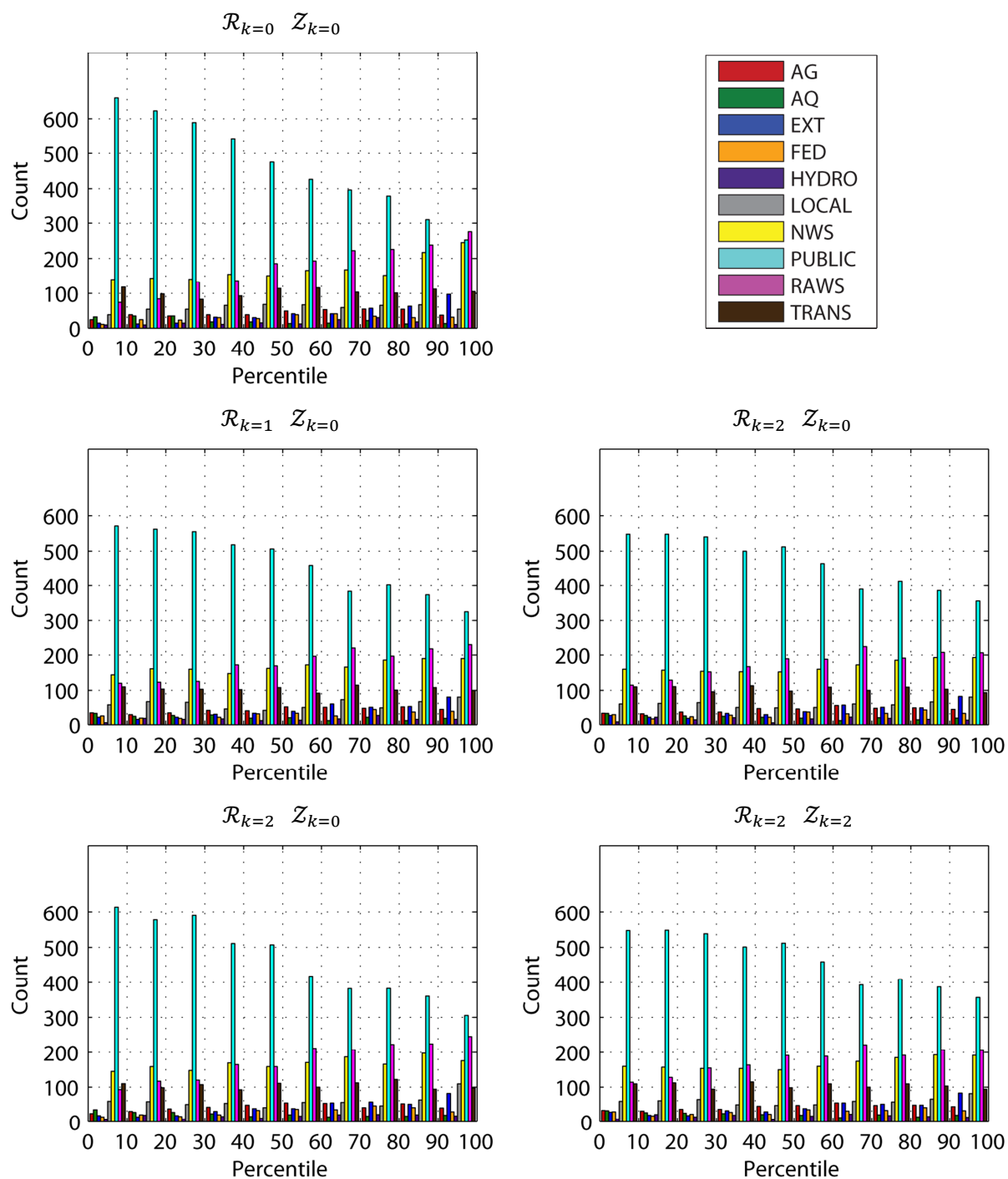


Figure 5.9. continued. b. 2-m dewpoint temperature observation sensitivity percentile distribution.

10-m Wind Speed Observation Sensitivity

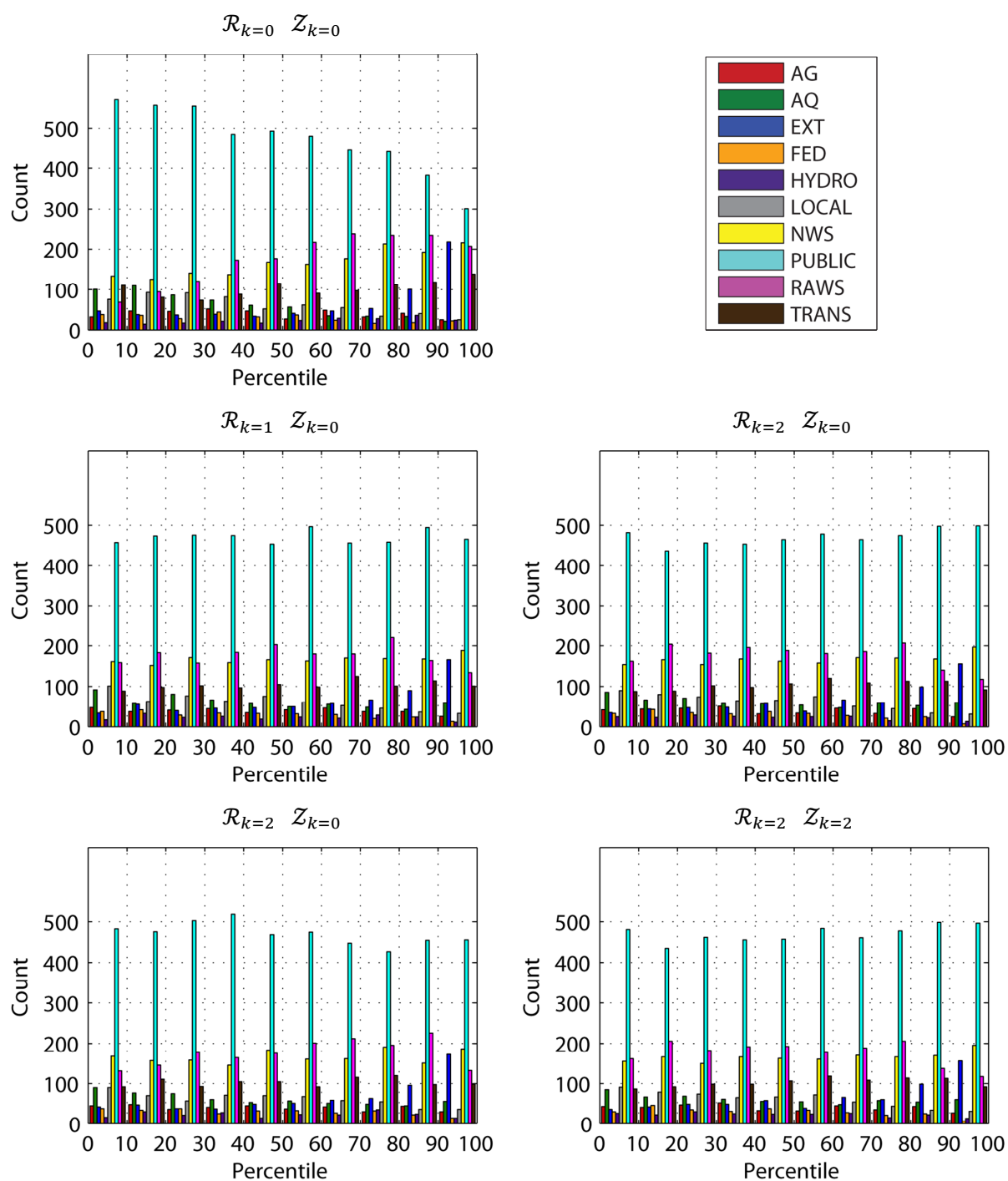


Figure 5.9. continued. c. 10-m wind speed observation sensitivity percentile distribution.

CHAPTER 6

CONCLUSION

Summary

High resolution spatial and temporal objective surface analyses are needed for many different mesoscale nowcasting and short-term forecasting needs. Unfortunately, model output from many operational numerical models is unable to fill this need due to their coarser resolution as well as their inability to appropriately model or parameterize many boundary layer processes. Accurate surface analyses can be created by using this model output as a first guess and using surface mesonet observations to correct this first guess through data assimilation.

This study presented the UU2DVar, a 2DVar analysis tool that can assimilate thousands of surface observations to produce surface analyses of 2-m air temperature, 2-m dewpoint temperature, 10-m u - and v - wind components, 10-m wind speed, and surface pressure. Unlike its predecessor (the LSA), the UU2DVar can be run over continental scale domains because it solves the variational cost function in observation space instead of analysis space, greatly reducing the necessary memory to store the background error covariances. The majority of the UU2DVar's routines are written to take advantage of parallel processing, greatly decreasing the computational time required to compute its background error covariances and allowing it to be run over a continental

scale domain within real-time constraints. The parallel speedup of the UU2DVar's functions to compute the background error covariances is a function of the amount of memory of the computer system used to run it; systems with more memory can take advantage of the processor computing larger blocks of the covariance array at once, increasing the actual speedup towards the idealized perfect speedup.

The UU2DVar is written using the MATLAB programming language, allowing it to be run with any operating system that supports the MATLAB software (Windows, Mac OS X, and Linux). Earlier versions of the UU2DVar have also been compiled using the MATLAB Compiler, which has allowed the code to be run on systems without MATLAB licenses using the freely available MATLAB Compiler Runtime as an executable binary. Users of the UU2DVar do not have to supply their own observations and background fields, as the tool is written to interface with the University of Utah THREDDS server and MesoWest database; however, users have the option to incorporate their own observation datasets.

The UU2DVar provides a flexible platform from which observations from heterogeneous surface mesonets can be examined objectively. The National Research Council (2009) recommended improved metadata, data quality control procedures, and understanding of the relative merits of differing data sources as ways to increase the utilization of such observations throughout the weather enterprise. The development of the UU2DVar and its adjoint was instigated with those goals in mind.

While the UU2DVar shares its background field as well as a similar 2DVar assimilation technique with the RTMA, its analyses differ from those of the NCEP system due the different assumptions used to generate the analyses. Because the

UU2DVar computes its analyses in observation space, the UU2DVar utilizes a terrain field closer to that observed to compute its background error covariance, which helps to explain why the largest differences between the UU2DVar and the RTMA analyses tend to be located in areas of complex terrain.

To illustrate the applicability of this system and expand on prior research (Horel and Dong 2010; Tyndall et al. 2010), a single case was examined in depth with particular attention placed on the dependence of the analysis system to variations in the background error covariance specified as a function of data density. Data density is computed in terms of the IDI field over the entire CONUS. Usage of the UU2DVar adjoint instead of leave-one-out data withholding experiments (as done by Horel and Dong [2010]) allows for an efficient methodology to determine the sensitivity and impacts to all observations. This study demonstrates that it is possible to use varying decorrelation length scales in specifying the background error covariance as well as the efficiency of the adjoint methodology to determine the impact of varying data assets. Further study is required to assess whether using a data density criterion to constrain the analysis is beneficial. However, it is clear from this single case that the extreme variations in data density over the continental United States are a challenge, since overfitting can result if the analysis is too tightly constrained. Additional research may show that a “flatter” polynomial, in which decorrelation length scales do not decrease as significantly in data dense regions, may prove more beneficial to the analysis. Furthermore, forecasters utilizing such analyses must help assess whether it is more beneficial to have a smoother analysis, or one that is able to resolve small scale features where the observing network is dense.

Observation impact appears to be a more robust metric for contrasting the influence of observations than observation sensitivity. Observations with high impact draw attention to locations with both high sensitivity as well as large innovations. However, it is difficult to distinguish through simple objective criteria between high impact observations resulting from meaningful deviations in local weather from the background, erroneous or unrepresentative observations, or erroneous features in the background itself. Alternative cost functions to that examined here (mean squared difference between the analysis and background over the entire grid) could be specified by Equation 4.4 in order to focus on other questions of interest, e.g., particular flow characteristics within limited domains.

For the analysis hour examined here, stations in data sparse regions where deviations from the background were large tended to have high impact, while stations in data rich urban areas tended to have lower impact. For example, the HYDRO and RAWS network categories, with many stations in remote locations, had larger numbers of high impact temperature observations than low impact observations. RAWS observations also had many stations with high dewpoint temperature impact as well, due to their strong observation innovations in many regions of the western United States. Stations in the PUBLIC category tended to have very high observation impacts on the wind speed analysis, which may be the result of sensor siting issues as well as unrepresentative and erroneously calm wind observations collocated with high background field wind speeds. Applying the four data density constraints to the analyses increased the number of high impact PUBLIC observations in all fields.

This research, which includes the UU2DVar analysis tool, as well as an investigation of observation impacts for an individual case study, helps to lay the foundation for additional research focusing on addressing issues associated with the development of the NNoN. A parallel study applying the methodology presented in Chapter 4 to 100 cases of significant weather events is already underway and some preliminary results have already been collected. Those results confirm the higher observation impacts of networks located in data sparse regions (such as the RAWS network) as seen with the single case study presented in Chapter 5. The implementation of observation sensitivity and impacts as a measure of quality control for observations part of the NNoN is also being discussed with MesoWest researchers.

Recommendations and Future Work

The development of the UU2DVar and usage of the variational adjoint to determine observation impact has led to a number of additional questions, as well as goals for future work. These recommendations and goals for future work are expanded upon here:

1. **Collection and regular updating of observation metadata is necessary for the production of high quality surface analyses.** As shown by the wind analyses computed as part of this research, the assimilation of poor quality observations can greatly reduce the quality of the analyses. Unfortunately, without complete siting information, as well observation maintenance and station instrument information, it is difficult to differentiate good quality observations from poor quality observations, even with the use of more

advanced quality control procedures. Providers of observation data should also make every effort to make network description publications available with the data, and large observing networks (especially federally funded networks) should be required to maintain documents describing the standards used within the network.

2. **Advanced quality control procedures should be implemented within the UU2DVar.** Implementation of more rigorous quality control procedures on the observations retrieved from the MesoWest database used by the UU2DVar will substantively improve the utility of this system. The quality control steps implemented within this version of the UU2DVar are limited to removing gross errors based on error characteristics assumed for the entire domain as a whole. A number of additional quality control procedures are under development by the MesoWest team and will help to remove many commonly occurring problems. For example, incorrect station elevation or incorrect base line pressure values are common in mountainous areas. A quality control step that uses the standard deviation of the background field immediately around the observation instead of across the entire domain as presently implemented would be a substantive improvement. The asymmetric quality control for wind observations described in Chapter 3 has already been investigated over a large sample of cases and appears to be effective at removing some of the erroneous observations found in the PUBLIC network wind observations. An additional quality control measure, in which observations are checked against nearby gridpoints from a preliminary analysis before they are actually assimilated by

the final analysis would also improve analysis quality, and would be straightforward to implement using legacy code from the earlier versions of the UU2DVar.

3. **Adjust observation error based on network characteristics.** Even without complete observation metadata, assumptions can be made about certain networks and the quality of their observations. Preliminary work that has already been completed for a sample of 100 cases demonstrates this approach is promising. Increasing observation errors for networks or network categories with representativeness or other recognizable errors appear to improve the analyses as an alternative to simply eliminating usage of those observations through commonly used network blacklists. In addition, variable observation errors increase the complexity of the observation sensitivity metric, as it becomes a function of observation locations, analysis increments, and the background and observation errors.
4. **Observation sensitivity and impact should be studied as a function of different cost functions over many cases to identify important characteristics of individual observations and mesonets.** Preliminary research using the aforementioned sample of 100 different analysis hours on days of high impact weather in the United States suggests that observation impact combined with data density are useful metrics for such studies. However, the domain-scale root-mean squared difference cost function, although commonly exploited in similar studies, is not the only possible choice. Examining whether regional analyses combined with cost functions of

interest to particular aspects of the weather enterprise may be warranted. For example, wind energy companies may be interested in capabilities of a local boundary layer analysis system to estimate wind increases or decreases on time scales of 5 minutes to an hour as a function of different data resources.

5. **Use the UU2DVar as an additional quality control tool for mesonet observations.** Simplified versions of the UU2DVar could be implemented for real-time quality control of many variables archived in the MesoWest database. Statistics of observation sensitivity and observation impact collected routinely from hourly analyses would help to determine poor quality observations within a network. Observations with consistently high impacts over a large sample of cases are most likely biased, either through instrumentation errors or representativeness errors. In addition, quality control procedures based on negative observation impacts may be warranted, since they reflect observation innovations that are located in a region of analysis increments of opposite sign.

REFERENCES

- Ancell, B. C., C. F. Mass, and G. J. Hakim, 2011: Evaluation of surface analyses and forecasts with a multiscale ensemble Kalman filter in regions of complex terrain. *Mon. Wea. Rev.*, **139**, 2008-2024.
- Anderson, E., and Coauthors, 1999: LAPACK Users' *Guide*. SIAM, 407 pp.
- Andersson, E., and Coauthors, 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). III: Experimental results. *Q. J. R. Meteorol. Soc.*, **124**, 1831-1860.
- Baker, N. L., and R. Daley, 2000: Observation and background adjoint sensitivity in the adaptive observation-targeting problem. *Q. J. R. Meteorol. Soc.*, **126**, 1431-1454.
- Barker, T., L. Colin, C. Gibson, and K. Pomeroy, cited 2011: ObjAnal: Version 2.0 User Document. [Available online at <http://www.mdl.nws.noaa.gov/~applications/STR/generalappinfoout.php3?appnum=1163>.]
- Bannister, R. N., 2008a: A review of forecast error covariance statistics in atmospheric variational data assimilation. I: Characteristics and measurements of forecast error covariances. *Q. J. R. Meteorol. Soc.*, **134**, 1951-1970.
- Bannister, R. N., 2008b: A review of forecast error covariance statistics in atmospheric variational data assimilation. II: Modelling the forecast error covariance statistics. *Q. J. R. Meteorol. Soc.*, **134**, 1971-1996.
- Benjamin, S., and Coauthors, 2004: An hourly assimilation–forecast cycle: The RUC. *Mon. Wea. Rev.*, **132**, 495-518.
- , J. M. Brown, G. Manikin, and G. Mann, 2007: The RTMA background -- Hourly downscaling of RUC data to 5-km detail. Preprints, *18th Conference on Numerical Weather Prediction*, Park City, UT.
- Bergthórsson, P., and B. R. Döös, 1955: Numerical weather map analysis. *Tellus*, **7**, 329-340.
- Brock, F. V., K. C. Crawford, R. L. Elliott, G. W. Cuperus, S. J. Stadler, H. L. Johnson, and M. D. Eilts, 1995: The Oklahoma Mesonet: A technical overview. *J. Atmos. Oceanic Technol.*, **12**, 5-19.

- Buehner, M., and M. Charron, 2007: Spectral and spatial localization of background-error correlations for data assimilation. *Q. J. R. Meteorol. Soc.*, **133**, 615-630.
- Buizza, R., and A. Montani, 1999: Targeting observations using singular vectors. *J. Atmos. Sci.*, **56**, 2965-2985.
- Caron, J., E. Davis, Y. Ho, and R. Kambic, 2006: Unidata's THREDDS data server. Preprints, *22nd International Conference on Interactive Information Processing Systems for Meteorology, Oceanography, and Hydrology*, Atlanta, GA.
- Chadwick, R., cited 2011: APRSWXNET: The Automatic Position Reporting System as a WX NETWORK. [Available online at <http://www.findu.com/citizenweather/>.]
- Courtier, P., and Coauthors, 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation. *Q. J. R. Meteorol. Soc.*, **124**, 1783-1807.
- Cressman, G. P., 1959: An operational objective analysis system. *Mon. Wea. Rev.*, **87**, 367-374.
- Daley, R., 1991: *Atmospheric Data Analysis*. Cambridge University Press, 457 pp.
- , and E. Barker, 2001: NAVDAS: Formulation and diagnostics. *Mon. Wea. Rev.*, **129**, 869-883.
- de Ponte, M. S. F. V., and Coauthors, 2011: The Real-Time Mesoscale Analysis at NOAA's National Centers for Environmental Prediction: Current status and development. Accepted to *Wea. Forecasting*.
- Dongarra, J. J., J. D. Croz, S. Hammarling, and R. J. Hanson, 1988: An extended set of FORTRAN basic linear algebra subprograms. *ACM Trans. Math. Softw.*, **14**, 1-17.
- Environment Canada, 2008: Degradation in Environment Canada's climate network, quality control, and data storage practices: A call to repair the damage. 72 pp.
- Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10143-10162.
- Federal Aviation Administration, 2011: Advisory circular on Automated Weather Observing Systems (AWOS) for non-federal applications, AC 150/5220-16D. 78 pp.
- Fisher, M., 2003: Background error covariance modelling. *ECMWF Seminar on Recent Developments in Data Assimilation for Atmosphere and Ocean*, Reading, UK, ECMWF, 45-63.
- Foisy, T., 2003: The MatchObsAll Analysis System, NOAA Tech. Att. 03-02. 4 pp.

- Gandin, L. S., 1963: *Objective Analysis of Meteorological Fields*. Israel Programme for Scientific Translations.
- Gilchrist, B., and G. P. Cressman, 1954: An experiment in objective analysis. *Tellus*, **6**, 309-318.
- Gladstone, P., cited 2011: Weather Quality Reporter. [Available online at [http://weather.gladstonefamily.net/.](http://weather.gladstonefamily.net/)]
- Glahn, H. R., and D. P. Ruth, 2003: The new digital forecast database of the National Weather Service. *Bull. Amer. Meteor. Soc.*, **84**, 195-201.
- Glowacki, T. J., Y. Xiao, and P. Steinle, 2011: Mesoscale surface analysis system for Australian domain. Design issues, development status, system validation. Accepted to *Wea. Forecasting*.
- Haiden, T., A. Kann, C. Wittmann, G. Pistotnik, B. Bica, and C. Gruber, 2011: The Integrated Nowcasting through Comprehensive Analysis (INCA) system and its validation over the eastern Alpine region. *Wea. Forecasting*, **26**, 166-183.
- Hamill, T. M., J. S. Whitaker, and C. Snyder, 2001: Distance-Dependent Filtering of Background Error Covariance Estimates in an Ensemble Kalman Filter. *Mon. Wea. Rev.*, **129**, 2776-2790.
- Hestenes, M. R., and E. Stiefel, 1952: Methods of conjugate gradients for solving linear systems. *J. Res. Nat. Bur. Stand.*, **49**, 409-436.
- Hiemstra, C. A., G. E. Liston, R. A. Pielke, D. L. Birkenheuer, and S. C. Albers, 2006: Comparing Local Analysis and Prediction System (LAPS) assimilations with independent observations. *Wea. Forecasting*, **21**, 1024-1040.
- Horel, J. D., and Coauthors, 2002: Mesowest: Cooperative mesonets in the western United States. *Bull. Amer. Meteor. Soc.*, **83**, 211-225.
- and B. Colman, 2005: Real-time and retrospective mesoscale objective analyses. *Bull. Amer. Meteor. Soc.*, **86**, 1477-1480.
- and X. Dong, 2010: An evaluation of the distribution of Remote Automated Weather Stations (RAWS). *J. Appl. Meteor. Climatol.*, **49**, 1563-1578.
- 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.>]
- Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation, and Predictability*. Cambridge University Press, 341 pp.

- Langland, R. H., R. Gelaro, G. D. Rohaly, and M. A. Shapiro, 1999: Targeted observations in FASTEX: Adjoint-based targeting procedures and data impact experiments in IOP17 and IOP18. *Q. J. R. Meteorol. Soc.*, **125**, 3241-3270.
- Lazarus, S. M., C. M. Ciliberti, J. D. Horel, and K. A. Brewster, 2002: Near-real-time applications of a mesoscale analysis system to complex terrain. *Wea. Forecasting*, **17**, 971-1000.
- Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Q. J. R. Meteorol. Soc.*, **112**, 1177-1194.
- , R. S. Bell, and B. Macpherson, 1991: The Meteorological Office analysis correction data assimilation scheme. *Q. J. R. Meteorol. Soc.*, **117**, 59-89.
- McPherson, R. D., 1975: Progress, problems, and prospects in meteorological data assimilation. *Bull. Amer. Meteor. Soc.*, **56**, 1154-1166.
- Miller, P. A., M. F. Barth, L. A. Benjamin, R. S. Artz, and W. R. Pendergrass, 2005: The Meteorological Assimilation and Data Ingest System (MADIS): Providing value-added observations to the meteorological community. Preprints, *21st Conference on Weather Analysis and Forecasting*, Washington, D.C., American Meteorological Society, P1.95.
- Myrick, D. T., J. D. Horel, and S. M. Lazarus, 2005: Local adjustment of the background error correlation for surface analyses over complex terrain. *Wea. Forecasting*, **20**, 149-160.
- and ———, 2008: Sensitivity of surface analyses over the western United States to RAWS observations. *Wea. Forecasting*, **23**, 145-158.
- National Environmental Satellite, Data, and Information Service, 2003: United States Climate Reference Network (USCRN) FY 2003 Annual Report. 24 pp.
- National Oceanic and Atmospheric Administration, 1998: Automated Surface Observing System (ASOS) user's guide. 74 pp.
- National Research Council, 2009: *Observing Weather and Climate from the Ground Up: A Nationwide Network of Networks*. National Academy Press, 234 pp.
- Office of Hydrologic Development, cited 2011: Hydrometeorological Automated Data System. [Available online at <http://www.weather.gov/oh/hads/>.]
- Palmer, T. N., R. Gelaro, J. Barkmeijer, and R. Buizza, 1998: Singular vectors, metrics, and adaptive observations. *J. Atmos. Sci.*, **55**, 633-653.
- Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747-1763.

- Protat, A., and I. Zawadzki, 1999: A variational method for real-time retrieval of three-dimensional wind Field from multiple-Doppler bistatic radar network data. *J. Atmos. Oceanic Technol.*, **16**, 432-449.
- Purser, R. J., W.-S. Wu, D. F. Parrish, and N. M. Roberts, 2003a: Numerical aspects of the application of recursive filters to variational statistical analysis. Part I: Spatially homogeneous and isotropic Gaussian covariances. *Mon. Wea. Rev.*, **131**, 1524-1535.
- , ———, ———, and ———, 2003b: Numerical aspects of the application of recursive filters to variational statistical analysis. Part II: Spatially inhomogeneous and anisotropic general covariances. *Mon. Wea. Rev.*, **131**, 1536-1548.
- Rew, R., and G. Davis, 1990: NetCDF: An interface for scientific data access. *IEEE Comput. Graphics Appl.*, **10**, 76-82.
- Saad, Y., and M. H. Schultz, 1986: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Stat. Comput.*, **7**, 856-869.
- Sasaki, Y., 1958: An objective analysis based on the variational method. *J. Meteorol. Soc. Jpn.*, **36**, 77-88.
- , 1970: Some basic formalisms in numerical variational analysis. *Mon. Wea. Rev.*, **98**, 875-883.
- Schaefer, G. L., and R. F. Paetzold, 2001: SNOTEL (SNOWpack TELEmetry) and SCAN (Soil Climate Analysis Network), *Automated Weather Stations for Application in Agriculture and Water Resources Management*, K. G. Hubbard, and M. V. K. Sivakumar, Eds., WMO/TD No. 1074. 187-194.
- Schroeder, J. L., W. S. Burgett, K. B. Haynie, I. Sonmez, G. D. Skwira, A. L. Doggett, and J. W. Lipe, 2005: The West Texas Mesonet: A technical overview. *J. Atmos. Oceanic Technol.*, **22**, 211-222.
- Seaman, R. S., and M. F. Hutchinson, 1985: Comparative real data test of some objective analysis methods by withholding observations. *Aust. Meteorol. Mag.*, **33**, 37-46.
- Servicio Meteorológico Nacional, cited 2011: Información Sinóptica de los Observatorios (Information on Synoptic Observations). [Available online at http://smn.cna.gob.mx/index.php?option=com_content&view=article&id=117&Itemid=90.]
- Steinacker, R., and Coauthors, 2006: A mesoscale data analysis and downscaling method over complex terrain. *Mon. Wea. Rev.*, **134**, 2758-2771.
- Snyder, R., 1984: California Irrigation Management Information System. *Am. J. Potato Res.*, **61**, 229-234.

- Talagrand, O., 1997: Assimilation of observations, an introduction. *J. Meteorol. Soc. Jpn.*, **75**, 191-209.
- Tyndall, D. P., 2008: Sensitivity of surface temperature analyses to specification of background and observation error covariances. M.S. thesis, Dept. of Meteorology, The University of Utah, 105 pp.
- , J. D. Horel, and M. S. F. V. de Pondeca, 2010: Sensitivity of surface air temperature analyses to background and observation errors. *Wea. Forecasting*, **25**, 852-865.
- U.S. Bureau of Reclamation, cited 2011: AgriMet: The Pacific Northwest Cooperative Agricultural Weather Network. [Available online at <http://www.usbr.gov/pn/agrimet/index.html>.]
- Uboldi, F., C. Lussana, and M. Salvati, 2008: Three-dimensional spatial interpolation of surface meteorological observations from high-resolution local networks. *Meteor. Appl.*, **15**, 331-345.
- Weaver, A., and P. Courtier, 2001: Correlation modelling on the sphere using a generalized diffusion equation. *Q. J. R. Meteorol. Soc.*, **127**, 1815-1846.
- Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences*. 3rd ed. Elsevier, 627 pp.
- Xie, Y., C. Lu, and G. L. Browning, 2002: Impact of formulation of cost function and constraints on three-dimensional variational data assimilation. *Mon. Wea. Rev.*, **130**, 2433-2447.
- , S. Koch, J. McGinley, S. Albers, P. E. Bieringer, M. Wolfson, and M. Chan, 2011: A space-time multiscale analysis system: A sequential variational analysis approach. *Mon. Wea. Rev.*, **139**, 1224-1240.
- Zachairassen, J., K. Zeller, N. Nikolov, and T. McClelland, 2003: A review of the Forest Service Remote Automated Weather Station (RAWS) network. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, 153 pp.
- Zapotocny, T. H., and Coauthors, 2000: A case study of the sensitivity of the Eta Data Assimilation System. *Wea. Forecasting*, **15**, 603-621.
- Zhu, J., and C. Yan, 2006: Nonlinear balance constraints in 3DVAR data assimilation. *Sci. China, Ser. D Earth Sci.*, **49**, 331-336.
- Zhu, Y., and R. Gelaro, 2008: Observation sensitivity calculations using the adjoint of the Gridpoint Statistical Interpolation (GSI) analysis system. *Mon. Wea. Rev.*, **136**, 335-351.