working to advance road weather information systems technology

The Pavement Precipitation Accumulation Estimation System—Further Development

http://aurora-program.org

Aurora Project 2009-05

Final Report
May 2013

About Aurora

Aurora is an international program of collaborative research, development and deployment in the field of road and weather information systems (RWIS), serving the interests and needs of public agencies. The Aurora vision is to deploy RWIS to integrate state-of-the-art road and weather forecasting technologies with coordinated, multi-agency weather monitoring infrastructures. It is hoped this will facilitate advanced road condition and weather monitoring and forecasting capabilities for efficient highway maintenance, and the provision of real-time information to travelers.

Disclaimer Notice

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. The opinions, findings and conclusions expressed in this publication are those of the authors and not necessarily those of the sponsors.

The sponsors assume no liability for the contents or use of the information contained in this document. This report does not constitute a standard, specification, or regulation.

The sponsors do not endorse products or manufacturers. Trademarks or manufacturers' names appear in this report only because they are considered essential to the objective of the document.

Non-Discrimination Statement

Iowa State University does not discriminate on the basis of race, color, age, religion, national origin, sexual orientation, gender identity, genetic information, sex, marital status, disability, or status as a U.S. veteran. Inquiries can be directed to the Director of Equal Opportunity and Compliance, 3280 Beardshear Hall, (515) 294-7612.

Iowa Department of Transportation Statements

Federal and state laws prohibit employment and/or public accommodation discrimination on the basis of age, color, creed, disability, gender identity, national origin, pregnancy, race, religion, sex, sexual orientation or veteran's status. If you believe you have been discriminated against, please contact the Iowa Civil Rights Commission at 800-457-4416 or the Iowa Department of Transportation affirmative action officer. If you need accommodations because of a disability to access the Iowa Department of Transportation's services, contact the agency's affirmative action officer at 800-262-0003.

The preparation of this report was financed in part through funds provided by the Iowa Department of Transportation through its "Second Revised Agreement for the Management of Research Conducted by Iowa State University for the Iowa Department of Transportation" and its amendments.

The opinions, findings, and conclusions expressed in this publication are those of the authors and not necessarily those of the Iowa Department of Transportation or the U.S. Department of Transportation Federal Highway Administration.

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
Aurora Project 2009-05		

4. Title and Subtitle	5. Report Date
The Pavement Precipitation Accumulation Estimation System—Further Development	May 2013
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
Mark Askelson, Jeff Tilley, Ed Townsend, and Christopher Theisen	Aurora Project 2009-05

9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
John D. Odegard School of Aerospace Sciences University of North Dakota 3890 Campus Road, Stop 9007 Grand Forks, ND 58202-9007	**11. Contract or Grant No.**

12. Sponsoring Organization Name and Address		13. Type of Report and Period Covered
Aurora Program Iowa Department of Transportation 800 Lincoln Way Ames, Iowa 50010	Federal Highway Administration U.S. Department of Transportation 1200 New Jersey Avenue SE Washington, DC 20590	Final Report **14. Sponsoring Agency Code** Pooled Fund SPR 72-00-0003-042

15. Supplementary Notes

Visit www.intrans.iastate.edu for color pdfs of this and other research reports.

16. Abstract

Wintertime precipitation has profound impacts on surface transportation. Surface transportation ramifications include decreased public safety, compromised traveler mobility, diminished productivity of roadway users, and adverse environmental effects due to the need to chemically treat icy roads. The winter road maintenance community can benefit from having more information about wintertime precipitation, including where precipitation is occurring and precipitation rate, leading to improvements of safety and greater cost effectiveness.

Determining wintertime precipitation occurrence and intensity is challenging. The Pavement Precipitation Accumulation Estimation System (PPAES) was designed such that precipitation field analyses could be produced using individual sources of information and multiple sources of information. When multiple sources are utilized, PPAES algorithms are designed to take advantage of data strengths and mitigate data weaknesses to maximize analysis quality.

This study, in addition to providing new techniques for fusing information from multiple sources to produce enhanced wintertime precipitation analyses, identifies future research areas that may lead to even more advanced techniques that will serve the road weather community.

17. Key Words	18. Distribution Statement
algorithm performance—precipitation analyses—wintertime precipitation	No restrictions.

19. Security Classification (of this report)	20. Security Classification (of this page)	21. No. of Pages	22. Price
Unclassified.	Unclassified.	169	NA

Form DOT F 1700.7 (8-72) **Reproduction of completed page authorized**

THE PAVEMENT PRECIPITATION ACCUMULATION ESTIMATION SYSTEM—FURTHER DEVELOPMENT

Final Report
May 2013

Authors
Mark Askelson, Associate Professor, Atmospheric Sciences
John D. Odegard School of Aerospace Sciences, University of North Dakota

Jeffrey Tilley, Research Faculty
Regional Weather Information Center, John D. Odegard School of Aerospace Sciences

Ed Townsend, Graduate Research Assistant Atmospheric Sciences
John D. Odegard School of Aerospace Sciences, University of North Dakota

Christopher Theisen, Radar Research Meteorologist,
Regional Weather Information Center, John D. Odegard School of Aerospace Sciences

Sponsored by
Aurora Program, Iowa Department of Transportation,
And Federal Highway Administration (FHWA)
FHWA Pooled Fund Study SPR-3(042): AK, IA, IL, IN, KS, MI, MN, ND, NV, NY, OH, Ontario MOT, PA, Quebec MOT, SD, Sweden NRA, TN, UT, VA, WI

Preparation of this report was financed in part
through funds provided by the Iowa Department of Transportation
through its Research Management Agreement with the
Institute for Transportation
(Aurora Project 2009-05)

A report from
Aurora Program
Institute for Transportation
Iowa State University
2711 South Loop Drive, Suite 4700
Ames, IA 50010-8664
Phone: 515-294-8103
Fax: 515-294-0467
www.intrans.iastate.edu

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

This research was conducted under the Federal Highway Administration (FHWA) Pooled Fund Study SPR-3(042). The authors would like to express their gratitude to the FHWA, the Aurora Program partners, and the Iowa Department of Transportation (lead state) for their financial support and technical assistance.

The authors thank, foremost, Aurora for having the vision to fund this research. We also thank Leon Osborne and John Nordlie of the Regional Weather Information Center (RWIC). Their hard work and gracious assistance assured that the vast volumes of data that were required to test the algorithms developed herein were both available and timely.

EXECUTIVE SUMMARY

Wintertime precipitation has profound impacts on surface transportation. Surface transportation ramifications include decreased public safety, compromised traveler mobility, diminished productivity of roadway users, and adverse environmental effects owing to the need to chemically treat icy roads. Furthermore, the winter road maintenance community can benefit from having more information about wintertime precipitation, including where precipitation is occurring and precipitation rate, leading to improvements of safety and greater cost effectiveness.

Individual sources of information regarding wintertime precipitation suffer from significant limitations. With radars, specifically the National Weather Service WSR-88D Doppler Radar network, the primary issues are overshooting and inherent uncertainties in estimated precipitation rate values. Overshooting arises because, under typical propagation conditions, radar beams rise to increasing altitudes with increasing range. This results in radars being unable to sense precipitation beyond a certain range. When this is coupled with the fact that wintertime precipitation systems tend to be shallower than summertime systems and with the spacing of radars within the U.S. (which was based partially upon the depths of summertime systems), the result is wintertime systems not being detected with the existing radar network over large areas of the U.S. The second principal limitation of radar data is inherent uncertainties in precipitation rates. These can arise, in extreme cases, from virga, in which precipitation is believed to be occurring based upon radar data but is not reaching the ground. They also arise owing to inherent uncertainties in the relation between what is measured with a radar, radar reflectivity factor, and the desired quantity, liquid-water-equivalent precipitation rate.

Surface data also suffer from significant limitations when estimating precipitation rates. These include poor spatial resolution of the precipitation field, which can be an exceptionally severe problem in regions where surface observation station density is low. They also include inherent uncertainty in liquid-water-equivalent precipitation rates, which again results from uncertainties in the relation between these rates and what is used to estimate them (visibility).

Although they utilize sophisticated data assimilation methods, model precipitation fields are known to suffer from both phase (location) and amplitude (intensity) errors. These errors arise both from inadequacies in observed fields and from weaknesses inherent with the modeling systems.

Consequently, use of one source of information regarding wintertime precipitation will result in significant inaccuracies. The combination, or fusion, of information from multiple sources, however, may enable enhanced analysis performance. The Pavement Precipitation Accumulation Estimation System (PPAES) was designed such that precipitation field analyses could be produced using individual sources of information and using multiple sources of information. When multiple sources are utilized, PPAES algorithms are designed to take advantage of data strengths and mitigate data weaknesses so as to maximize analysis quality.

The purpose herein was to extend previous PPAES development to explore 1) the fusion of radar and surface analyses—including the use of both ASOS/AWOS (Automated Surface Observing System/Automated Weather Observing System) data and Clarus data, 2) the fusion of radar and model data, and 3) the performance of an enhanced radar algorithm and a radar+model blending algorithm in complex terrain.

In order to quantify algorithm performance, data are verified using ASOS/AWOS data. While these data have limitations, some of which are highlighted in this study, they do provide the best readily-available verification data for this study. Verification for analyses that include the use of ASOS/AWOS data was accomplished through employment of a data denial scheme. In addition to performance metrics like Probability of Detection (POD), Equitable Threat Score (ETS), and Root Mean Square Error (RMSE), for surface analyses the statistical significance of performance differences were also analyzed.

The radar+surface blending algorithm was based upon the premise that radars would generally provide relatively accurate depictions of precipitation fields at locations that are near them, whereas surface data would be needed at distant ranges. Based upon this, algorithms for identifying the effective range of a radar, within which radar data are used as the primary information source, were developed. These algorithms initially produce effective ranges that are based upon radar data alone. These effective ranges are then updated using surface observations. Once this process is complete, precipitation fields based upon radar data and surface data are blended together to remove aesthetically-displeasing jumps in the analysis field. It is noted that in the surface analysis algorithm, less smoothing is used in areas where surface observation station density is high so that finer-scale information regarding the precipitation field in those areas is retained.

The purpose of the radar+model blending algorithm is to enhance analysis performance in complex terrain, where radar beam blockage and low surface observation station density can present significant challenges. The first generation algorithm developed herein utilizes a relatively simple approach in which radar-estimated precipitation fields are replaced with model-estimated precipitation fields if the altitude of the data upon which the model-precipitation estimates are based are significantly below (user controlled) the altitudes of the radar data.

In addition to the development of blending algorithms, the PPAES radar algorithm was updated such that the user can request terrain clearance. When requested, any radar data that are within a user requested altitude of the ground are not used in the analyses.

The PPAES radar+surface blending module provided an improved probability of detection compared to the other analyses and this difference in performance is statistically significant at the 5% level. However, other measures of accuracy, including False Alarm Ratio (FAR) and ETS indicated that surface analyses alone performed better than radar and radar+surface analyses. Despite this, the radar+surface analyses do seem to provide a more coherent picture of precipitation fields since they provide fine-scale information near radars and precipitation filling from surface data at locations where radar overshooting is a significant problem.

Inclusion of *Clarus* data degraded analysis quality. While the exact cause of this is not known, possible reasons for this include different performance characteristics relative to ASOS/AWOS precipitation sensors and equipment maintenance practices and standards. This is an issue that requires further examination.

With the radar+model blending approach and the limited number of test cases used herein, the addition of model data resulted in degradation of performance from the standpoint of the metrics used herein. However, the use of model data does have the advantage of presenting a more continuous and spatially-coherent representation of the precipitation field to the user. The primary issue with the approach used herein is that radar spacing is large enough in the Utah domain such that the model field is utilized over a significant portion of that domain. This results in overfilling. A useful approach in the future may be restriction of use of model data to areas that are poorly covered by both radar and surface data.

Because 1) WSR-88D radars were sited such that radar beam blockage at the lowest elevation is relatively minor, resulting in relatively small differences between radar-based analyses produced with and without terrain clearance, and 2) the surface observation network available for verification is relatively sparse, no significant enhancement in performance was observed when terrain clearance was used. This capability, however, could significantly enhance analyses in non-standard radar ray propagation conditions.

Determining wintertime precipitation occurrence and intensity is challenging. This study, in addition to providing new techniques for fusing information from multiple sources to produce enhanced wintertime precipitation analyses, identifies future research areas that may lead to even more advanced techniques that will serve the road weather community.

1 INTRODUCTION

Wintertime precipitation has profound impacts on surface transportation. Surface transportation ramifications include decreased public safety, compromised traveler mobility, diminished productivity of roadway users, and adverse environmental effects, owing to the need to chemically treat icy roads (OFCM 2002; OFCM 2005; OFCM 2006). Furthermore, the winter road maintenance community can benefit from having more information about wintertime precipitation (herein defined as precipitation in its solid phase), including both where precipitation is occurring and precipitation rate (Mahoney 2003), leading to improvements of safety and greater cost effectiveness.

An improved depiction of precipitation can be gained by fusing observations from multiple platforms together in a precipitation estimation system (Seo 1998; Breidenbach and Bradberry 2001). Remote sensing (radar) and in situ (surface observation) observation platforms serve as the primary data sources used in this study. However, the use of model reflectivity data is also examined. Radars provide information regarding the timing, structure, and spatial extent of wintertime precipitation (Rinehart 2004, 199-204). Surface observations can enhance the accuracy of precipitation estimations through the elimination of false precipitation indications (virga) and the identification of areas where precipitation is not sensed by other observation platforms (hole-filling). Model reflectivity may provide a means to estimate precipitation in areas where no observations are available (i.e., complex terrain and remote regions). Satellite data could also provide additional information in complex terrain and remote regions, albeit with lower temporal resolution, and needs to be examined further before inclusion into PPAES for this purpose.

This study describes efforts directed at blending model data and observations obtained with radars and in-situ surface instruments to advance the development of the University of North Dakota's Pavement Precipitation Accumulation Estimation System (PPAES). Individually, each observation platform can provide information regarding wintertime precipitation – but because of their own respective limitations, one's ability to estimate wintertime precipitation with just one type of platform is limited. When compared to a single platform – radar or surface observations – the blended product is expected to provide a more accurate depiction of precipitation occurrence and rate.

1.1 Motivation

This research is motivated by the need for accurate estimates of wintertime precipitation to mitigate serious impacts of wintertime precipitation on surface transportation. In an effort to provide a more accurate depiction of precipitation occurrence and rate, algorithms are designed to leverage each platform's strengths, while minimizing impacts of each platform's limitations.

The two primary shortcomings of surface observations are sensor limitations and data distribution. Surface observations can provide local measurements of precipitation occurrence and rate. The accuracy of these measurements, however, depends upon several factors that include, but are not limited to, precipitation type (frozen versus liquid precipitation) and

environmental conditions [e.g., winds (Brock and Richardson 2001, 178-180; Dingman 2002, 169-170)].

The accuracy of surface observations depends upon a sensor's efficacy in collecting and measuring precipitation, which can be quite low--especially for snowfall. Issues that affect sensor accuracy when measuring snowfall include: low density of snow resulting in lack of weight in the sensor collection location (e.g., not enough weight to tip the bucket in a tipping-bucket gauge), the creation of a thermal plume that can deflect falling snow in heated sensors, capping of the gauge orifice (occurring pre-event or over the course of an event), snow sticking inside the gauge, evaporation effects, and wind driven effects (Super and Holroyd 1998; Brock and Richardson 2001, 178-180; Dingman 2002, 109-114). Wind driven effects are one of the greatest sources of error and can result in significant underestimates of snowfall. Experiments have shown that with non-shielded gauges the reduction is on the order of 20% for a 5 to 10 m s-1 wind and over 80% for winds greater than 10 m s-1 (Brock and Richardson 2001, 178-180).

The data distribution issue involves the spacing of available surface stations. The limited number of available stations and the frequency with which they report, as well as the area over which observation stations are distributed, control which spatial scales can be resolved and thereby limit the degree of detail that can be resolved (Daley 1991, 432-435). Thus, analyses derived from a limited number of surface observations are likely to be overly smooth, limiting their usefulness for estimating precipitation in real-time. Because of the limited number of surface stations available and the concomitant station spacing, the information gained from these surface stations is generally associated with larger-scale precipitation structures and not with what is occurring at smaller scales. As the distance between observation stations increases, the ability to resolve information at smaller scales becomes more limited. For example, the area covered by a precipitation gauge is on the order of 10^{-6} km^2, but the distribution of gauges may mean that the area closest to a particular gauge might be on the order of 10^3 km^2 (Brock and Richardson 2001, 179-180). For the purposes of this study, the mean and median data spacings for continental United States surface stations, provided by the Meteorological Assimilation Data Ingest System (MADIS), are ~38 km and ~65 km, respectively. As such, the spatial scales that can be resolved are on the order of 102 km [i.e., the Nyquist wavelength obtained by multiplying the mean data spacing by two (Daley 1991, 432-435)].

In an ideal scenario, information would be obtained using a regularly-spaced gridded observation network, but this is not the case here (Baer and Tribbia 1976; Doswell and Lasher-Trapp 1997). The irregular distribution of surface observation stations makes it difficult to extract information regarding certain scales, as that information is available in some areas and is not in others (Baer and Tribbia 1976; Doswell and Lasher-Trapp 1997), with potentially strong impacts on the fidelity of an objective analysis scheme (Doswell and Lasher-Trapp 1997). Therefore, surface observations do not serve well as the primary data source in a real-time precipitation estimation system such as PPAES (Super and Holroyd 1998; Askelson 2008). The surface observation platform's role in PPAES lies in the enhancement of precipitation information accuracy.

Radars, specifically S-band (10 cm wavelength) Weather Surveillance Radar-1988 Doppler (WSR-88D) systems, are used to measure radar reflectivity factor (herein reflectivity), which can

be used to estimate snowfall rates. As the other primary observation platform considered here, radars allow one to monitor large areas (e.g., area of approximately 70,000 km^2 at 150 km range; Brock and Richardson 2001, 182-183) at fine spatial resolutions (e.g., a data spacing of ~ 0.44 km at 25 km range and ~2.62 km at 150 km range assuming a 1 degree beamwidth). In addition, radars allow one to monitor precipitation frequently given its fine temporal resolution (i.e., time scale of five to ten minutes to complete a volume scan; Crum et al. 1993). Note, however, that the data spacing in the azimuthal direction increases with increasing range away from the radar.

Although improved spatial coverage can be gained through the use of radars, they suffer from two principal limitations when measuring snowfall: overshooting of precipitation by the radar beam and inherent inaccuracies of estimated snowfall rates (Smith et al. 1996; Fulton et al. 1998; Askelson 2008). We discuss each of these in turn. An electromagnetic beam transmitted by a radar typically rises relative to the surface of the Earth with increasing distance from the radar (Rinehart 2004, 64–68). Consequently, precipitation is commonly "overshot" at increasing distances from a radar, due to either the beam being fully above the precipitation phenomena or through incomplete beam-filling. The result is either an underestimate of the precipitation or no return at all (Smith et al. 1996; Fulton et al. 1998; Askelson 2008). This problem is exacerbated with relatively shallow storms, such as wintertime systems which commonly have a vertical extent of a few kilometers (e.g., Rinehart 2004, 149)]. For example, a radar beam at an elevation angle of 0.5° would overshoot a winter snowstorm with a 2 km cloud top height at ranges beyond 120 km under standard atmospheric refractive conditions (Rinehart 2004, p. 149). In comparison, summertime thunderstorms frequently exceed 10 km in height and can be observed using radars at much greater ranges.

In addition to overshooting issues, estimated snowfall rates, which herein are computed as liquid water equivalent precipitation rates, inherently suffer from shortcomings related to either how precipitation rates are computed or other known limitations associated with radar.

Reflectivity is subsequently related to an approximate precipitation (rainfall or snowfall, depending on temperature and other atmospheric conditions) rate. This is accomplished through the use of a relationship that generally assumes the form of a power law : $Z_e = \alpha S^{\beta}$, where Ze is equivalent radar reflectivity factor (hereafter denoted as Z) and S is the liquid water equivalent precipitation rate for snowfall (Doviak and Zrnić 1993, 223-229; Super and Holroyd 1998). Error is inherent to S estimates obtained using radar data due to variability in hydrometeor properties (size distribution, density, and terminal velocity) in the real atmosphere and the fact that Z and S have different dependencies on hydrometeor size distribution, density, and terminal velocity (Doviak and Zrnić 1993, 218-224; Lee and Zawadzki 2005). Thus, it is theoretically possible to alter the number of hydrometeors that exist within each bin of a set of binned size ranges in such a way such that *Z* is unchanged but that the value of *S* would change. This problem will be discussed in more detail in Section 2.

Other shortcomings of radar include mischaracterization of precipitation owing to virga, beam blockage, attenuation, precipitation advection, bright band and wet snow problems, incomplete beam filling, evaporation, errors induced by vertical motion, anomalous propagation, radar calibration, determination of Z-R/S relationships (R is rainrate) in the presence of mixed-phase

precipitation (let alone a single phase), and ground clutter (Doviak and Zrnić 1993, p. 225; Super and Holroyd 1998; Fulton et al. 1998).

2 BACKGROUND

A brief review of the state of PPAES prior to the work described in this report (hereafter referred to as PPAESv1) is presented in this section. In addition, related research efforts, an introduction to objective analysis concepts, and snowfall rate and size-distribution variability are also discussed.

2.1 PPAESv1

PPAESv1 was developed by the University of North Dakota (UND) to provide highly accurate estimates of wintertime precipitation occurrence, rate, and accumulation for surface transportation applications (Askelson 2008). As previously stated, PPAESv1 was designed to emphasize the strengths of numerous data collection platforms while at the same time mitigating their shortcomings. Most of the PPAESv1 software was previously tested for real-time applications during the 2005-2006, 2006-2007, and 2007-2008 winters (Askelson 2008). PPAESv1 includes radar processing routines to compute instantaneous equivalent liquid water precipitation rate, satellite processing routines to fill in observation holes associated with the WSR-88D network, an algorithm for estimating precipitation type via the Bourgouin method, and real-time execution and processing routines (Bourgouin 2000; Askelson 2008).

The characteristics of PPAESv1 and its analysis grid are provided in Table 1. PPAES uses a latitude-longitude grid with its southwestern corner located at a latitude and longitude of 20.0° and -130.0°, respectively. The PPAES horizontal latitude and longitude grid spacing is approximately 2.13 km by 2.00 km, respectively. The PPAES domain is designed to cover the entire continental United States (CONUS) and parts of Canada, Central and South America.

2.2 Previous Studies

Previous work that utilized both in situ and remote measurements to estimate snowfall has been limited. More effort has been directed at estimating summer-time precipitation – especially when combining radar rainfall or accumulations with surface rain gauge data (Seo 1998; Seo and Breidenbach 2002; Kondragunta et al. 2005; Root et al. 2009). The previous work completed by Seo (1998), Seo and Breidenbach (2002), and Kondragunta et al. (2005) has since been utilized in Multisensor Precipitation Estimator (MPE) software for use by the National Weather Service (NWS). Most recently, Kondragunta et al. (2005) have worked to integrate satellite-based rainfall estimates into multi-sensor estimates in the MPE. Lastly, a novel approach has been taken by Root et al. (2009) with regards to utilizing surface observations, including the use of other variables (temperature, pressure, humidity) to estimate rainfall rates using an artificial neural network. However, some work has been completed in regard to wintertime precipitation.

A significant effort was undertaken by the Bureau of Reclamation during1995-1998 to use radar measurements to estimate snowfall. Through these efforts the snow accumulation algorithm (SAA) that utilizes NEXRAD (NEXt generation weather RADar) Level II radar data was developed (Super and Holroyd 1998). The basis of the SAA is a power law relationship between radar-measured equivalent radar reflectivity factor at the lowest radar sweep and snow liquid-water-equivalent precipitation rate. Notable shortcomings of the SAA are 1) the algorithm was

developed for dry snow (i.e., not partially melted and/or mixed with rain) and 2) the algorithm inherently suffers from the shortcomings of radar (e.g., overestimation owing to virga and underestimation owing to overshooting of precipitation) (Super and Holroyd 1998). Their work is closely related to this research as both utilize Level II radar data and, similarly, compute liquid-water-equivalent precipitation rates from a Ze-S relation.

Table 1. PPAES specifications

PPAES Analysis Grid Characteristics and Control Variables		
Variable	**Value**	**Comment**
Number of Rows	1838	Number of latitudes in the analysis grid
Number of Columns	3662	Number of longitudes in the analysis grid
Westernmost Longitude (°)	-130.0	
Southernmost Latitude (°)	20.0	
Longitude Resolution (°)	0.01912046	
Latitude Resolution (°)	0.017964	
Alpha Beta	150.0 2.0	Variables in the precipitation rate Estimation power law equation
Half-width of sfc data time window (minutes)	30	Size of time window within which data must reside
Analysis delta-t (minutes)	10	Time step for analysis output files
Mean spacing, Δn (km)	40.839626	Of nearest available observation stations[1]
Median spacing, Δn (km)	34.865108	Of nearest available observation stations[1]
Random spacing, Δn_r (km)	127.856483	Of nearest available observation stations[1]
Area of PPAES Domain (km^2)	22608288.0	
Transition buffer distance (km)	100.0	Distance when transition from radar to surface analyses in the PPAES Blending module

[1]Using only MADIS stations residing in the PPAES domain and the CONUS

In addition, the Multi-sensor Precipitation Estimator (hereafter referred to as MPE) developed by the Office of Hydrologic Development of the NWS was deployed at NWS Weather Forecast Offices (WFOs) and River Forecast Centers (RFCs) during 2001 (Breidenbach and Bradberry 2001; Kondragunta et al. 2005). The MPE provides a multi-sensor accumulation precipitation analysis at 1- and 6-hour intervals in real-time from radar precipitation estimates and hourly rain gauge data (Lin and Mitchell 2005). The MPE serves as a potential input to multi-sensor analyses created at RFCs that themselves are inputs to the NCEP Stage III and Stage IV analysis products (NCDC 2010).

Despite the fact that both PPAES and MPE utilize multiple observations platforms, a fair comparison between PPAES and MPE cannot be made because of different purposes and associated differences in design. With respect to purpose, the MPE provides analyses of precipitation accumulation estimates, which are used in NWS hydrological and meteorological applications (input into models, verification, and operational monitoring) after being converted into NCEP analysis products (Lin and Mitchell 2005). By contrast, PPAES' purpose is to provide accurate estimates of wintertime precipitation (in real-time) to limit the serious impacts wintertime precipitation has on the surface transportation community.

There are also differences in specific approaches, data streams, and overall methodologies between PPAES and MPE. Table 2 summarizes the details of the differences between the MPE and PPAES algorithms.

Table 2. Comparison of MPE and PPAES algorithm characteristics (Kondragunta et al. 2005 and Lin and Mitchell 2005)

Characteristic	MPE	PPAES
Input Fields	Accumulation reports, 1-hr total generated from individual radars, satellite estimates	Present weather, visibility, instantaneous radar reflectivity
Output Fields	Precipitation accumulation	Instantaneous precipitation rate
Quality Control	Human analyst	Automated procedure
Spatial Resolution	4 km	1-2 km
Temporal Resolution	1, 6, 24 hours	10 min
Data Fusion Methodology	Optimal estimation theory (Seo 1998)	Blending via Barnes scheme with assumptions about field structures.

2.3 Objective Analysis

Objective analysis (hereafter referred to as OBAN) is the estimation of fields on a fixed, regular-spaced grid using irregularly spaced observations, inherently providing a filtering of such irregularly-spaced observations (Daley 1991, Ch. 1, 13; Askelson et al. 2000; Kalnay 2003, Ch. 5). OBAN techniques range from the computationally simple (nearest neighbor and function fitting; e.g., Daley 1991, Ch. 1-2) to more sophisticated and computationally expensive techniques [statistical objective analysis; optimal interpolation; variational objective analysis; Daley 1991, Ch. 4,5, 8; Kalnay 2003, Chapt. 5 ; Askelson et al. 2000]. Between these extremes lies a group of techniques referred to as distance-dependent weighted averaging (DDWA) schemes (e.g., Cressman 1959; Barnes 1973; Koch 1981; Daley 1991, Ch 2, 3). In this study, two types of OBAN schemes are used to estimate and filter surface observation data to estimate precipitation occurrence and rate – a nearest neighbor algorithm and a Barnes-type scheme.

Numerous factors affect how well an OBAN scheme performs. These factors include observation distribution characteristics (e.g., Koch 1981; Smith et al. 1986; Barnes 1994; Doswell and Lasher-Trapp 1997; Askelson et al. 2005b), observation errors (e.g., Daley 1991), assumptions made within the scheme and tuning of scheme parameters (e.g., Barnes 1994), and analysis grid specifications (Koch 1981). When identifying which OBAN scheme to use for a particular task it is important to also consider the scales resolved in the observational data and to be retained in the analyses (Koch 1981; Daley 1991, 432-425; Barnes 1994; Spencer et al. 2007).

2.3.1 Nearest Neighbor Algorithm

A nearest neighbor algorithm is used to estimate precipitation occurrence. This scheme is considered to be one of the simplest objective analysis techniques and is commonly employed in the analysis of radar data (Trapp and Doswell 2000; Askelson et al. 2000). Algorithm details,

including how precipitation occurrence analyses are derived in PPAES, are provided in the Methodology section.

The nearest neighbor algorithm was selected for precipitation occurrence because this algorithm is consistent with the binary nature of precipitation occurrence (e.g., zero for no occurrence and one for occurrence). If an analysis gridpoint is closest to an indication of precipitation occurrence, it is assigned a "1" but if it is closest to a non-occurrence, it is assigned a "0". This binary nature cannot be easily retained using other schemes because they "generate" values that are not present in the observations--in this case values would be generated between zero and one.

Note that the method used in the PPAES radar module to estimate precipitation occurrence and rate on the analysis grid is a pseudo-nearest neighbor method. This method involves assigning, for a given analysis grid point, the value of the range resolution volume (RRV) closest to the observation location, under the constraint that the center of that RRV must itself lie within the grid box associated with that analysis location. As such, the scheme is not a "true" nearest neighbor scheme as the spatial influence of an observation is limited by the size of an analysis grid box rather than distances to surrounding observations.

2.3.2 Barnes Algorithm

The Barnes algorithm is a Distance Dependent Weighted Averaging (DDWA) OBAN technique that uses a specified Gaussian weight function [Barnes 1964; Koch et al. 1981]. The Barnes algorithm was chosen over other techniques for multiple reasons, including: 1) the scheme is computationally simple; 2) it can be used when a background field is absent (as one is not required); 3) it can be used without knowledge of error correlation fields, and 4) response function characteristics can be easily estimated once the smoothing parameter is known[1] (Barnes 1973; Koch et al. 1981; Daley 1991, 90-93; Spencer et al. 2007). The smoothing parameter is derived from the theoretical response function, which is consistent with the Barnes algorithm operating as a low-pass filter and takes the form of Gaussian function (Barnes 1964; Barnes 1973; Koch et al. 1981) given by:

$$D_o = e^{\left[-\kappa_o\left(\frac{\pi}{\lambda}\right)^2\right]} \tag{1}$$

where D_o denotes the Barnes (1964) response function, κ_o denotes the smoothing parameter, and λ denotes the horizontal wavelength of the input signal. Note that the theoretical response is defined as the response characteristics produced by the OBAN scheme given that the data domain is infinite and continuous; this is different from the actual response that is affected by the discrete data distribution. The smoothing parameter controls the filtering characteristics of the Barnes scheme and is used to remove noise from the input while retaining signal (e.g., Barnes 1973; Koch et al. 1981).

[1] A commonly used assumption when estimating the response function is that is that the data domain is infinite and continuous. This is of course never the case and has major impacts that have been previously explored by, for instance, Achtemeier (1986), Pauley and Wu (1990), Doswell and Lasher-Trapp (1997), Askelson et al (2005b), and Spencer et al. (2007) to name a few.

Other characteristics of the Barnes scheme include: (1) it is a univariate scheme and as such does not leverage physical relations between variables to produce improved analyses and (2) it assumes the observations are "perfect". Because of 2), analyses converge to the observations with successive analysis passes (Barnes 1964; Daley 1991, Ch. 3).

The Koch et al. (1981) implementation of the Barnes algorithm consisted of a single correction pass as they determined that additional passes were not needed as due to how rapidly their analyses converged (i.e., on the second pass). In their scheme the first pass acts to suppress small-scale noise while building in the larger scales (e.g., Koch et al. 1981; Pauley and Wu 1990). The second pass estimates a correction at the observation locations to build in details from smaller scales (e.g., Koch et al. 1981; Pauley and Wu 1990; Daley 1991). In the second pass the amount of convergence desired at smaller scales (shorter wavelength features) within the final analysis can be controlled via modification of the smoothing parameter (Barnes 1973; Koch et al. 1981).

Errors or distortions in the analyzed field sometimes occur due to a combination of assumptions made with the Barnes scheme and how the observations are distributed (data boundaries, irregularities in the data distribution, differences between the actual and theoretical response functions when applying the scheme to discrete data; e.g., Achtemeier 1986; Smith et al. 1986; Pauley and Wu 1990; Askelson et al. 2005b). One important effect is the aliasing that occurs with smaller-scales that are not effectively resolved with the observation network, resulting in energy from unresolved scales appearing in the larger scales (Daley 1991, 432-425).

In the Barnes scheme analysis values are computed as distance-weighted averages using the Gaussian weight function

$$w_m = \exp\left(-\frac{r_m^2}{\gamma \kappa_0}\right) \tag{2}$$

where w_m denotes the weight function, r_m denotes the distance between the observation location (x_m, y_m) and the analysis grid point (i,j), κ_o denotes the weight parameter or smoothing parameter, and γ denotes the numerical convergence parameter. The smoothing and convergence parameters are present in the denominator of (2), where the convergence parameter is used to change the value of the smoothing parameter on successive passes (if desired). The smoothing parameter controls the shape of the response function. The numerical convergence parameter controls the convergence of the scheme. A constraint on the convergence parameter is that it must range from zero to one and be equal to one in the first pass (Koch et al. 1983). First-pass analysis values are computed using

$$g_o(i, j) = \frac{\sum_{m+1}^{M} w_m f(x_m, y_m)}{\sum_{m+1}^{M} w_m} \tag{3}$$

where $g_o(i, j)$ denotes the analysis grid point value at the (i,j) location, $f(x_m, y_m)$ denotes the observation value at the (x_m , y_m) observation location, and M denotes the total number of observations.

As previously stated, with the Barnes scheme successive corrections can be used, with one additional correction pass commonly being applied. In the correction pass, a new weight function w_m', having the same form as (2) but with $\gamma < 1$, is used (Koch et al. 1981). The second pass involves the computation of differences between observation values and the first pass analysis values estimated at observation locations and adding weighted values of these to the result of the first pass. The final value at the end of the correction pass is computed as the sum of the weighted averages from the two passes:

$$g_1(i,j) = \frac{\sum_{m+1}^{M} w_m f(x_m, y_m)}{\sum_{m+1}^{M} w_m} + \frac{\sum_{m+1}^{M} w'_m [f(x_m, y_m) - g_o(x_m, y_m)]}{\sum_{m+1}^{M} w'_m} \tag{4}$$

where $g_1(i, j)$ denotes the new correction pass value at the analysis grid point and $g_o(x_m, y_m)$ are the first pass values estimated at the observation locations.
The Koch et al. (1983) implementation of the Barnes scheme employs a cutoff distance. All weights for observations beyond this distance are set equal to zero. The cutoff distance in Koch et al. is given by

$$R_c = (20\kappa_0)^{1/2} \tag{5}$$

Although a cutoff distance is not needed with the Barnes scheme, we employ one in this study to decrease the scheme's computational cost, which is desirable for real-time applications. Fortunately, the fidelity of the Barnes scheme is not sensitive to a finite cutoff distance provided that the cutoff distance is sufficiently large. Otherwise, noise distortion in the form of ringing is possible if the weight function does not decrease to negligible values at this cutoff distance (Trapp and Doswell 2000; Askelson et al. 2000). Without a cutoff distance all of the observations would be used to compute analysis values at each analysis location and observations having extremely small weights would be needlessly used at high computational cost (Koch et al. 1981).

The value of the smoothing parameter κ_o allows the analyst to control the scale retention and removal characteristics *a priori* (Spencer et al. 2007). Oftentimes the smoothing parameter is chosen such that a particular small percentage of the amplitude of the minimally-resolved Nyquist wavelength (the "2Δx wave") is retained in the final analysis (Barnes 1973; Daley 1991, 432-435; Spencer et al. 2007). With regard to the Nyquist wavelength, it can be understood in terms of the "2Δx wave" as described by Daley (1991, p. 433): "if Δx is the distance between equally spaced observation stations, the 2Δx wave is the shortest wavelength that can be resolved by the network."

The smoothing parameter is derived from the Barnes (1964; 1973) theoretical response function. The Barnes scheme filters short wavelengths by attenuating their signal, and, thus, acts as a low-pass filter (i.e., filters high frequencies/short wavelengths while retaining low frequencies/long wavelengths) (Koch et al. 1981). The smoothing parameter is given by

$$\kappa_0 = -\left(\frac{\lambda}{\pi}\right)^2 \ln D_0(\lambda) \tag{6}$$

where λ denotes the wavelength of the input wave and $D_0(\lambda)$ denotes the desired amplitude response for the first pass of the scheme (Barnes 1973). Koch et al. (1981) specified $D_0(\lambda) = 0.0064$ for the "2Δn wave".

Koch et al.'s variable "Δn" can be computed as the mean observation spacing (also known as Δn_c) for maximum detail or can be computed as a "random data spacing" Δn_r for severely irregular data distributions (Koch et al. 1981; Daley 1991, App. H). Δn_r is different from Δn_c because Δn_r only takes into account the number of observations and the area of the analysis domain. Δn_r can serve as a guide to the analyst in selecting the proper "Δn" on the basis that if the observational data spacing is severely irregular, Δn_r will greatly exceed Δn_c (Koch et al. 1983). The motivation for Koch et al. (1981) selecting the desired response (of 0.0064) was to produce a maximum response on the correction pass for the "2Δn wave".

In this study, we have chosen not to utilize a correction pass in the analysis--partly due to the computational cost but mostly due to the marginal benefit a correction pass is expected to provide given the known irregularity in the distribution of surface stations. In the Barnes scheme's theoretical response function, the observational domain is assumed to be continuous and infinite (Barnes,1964). The irregularity of the surface station distribution does not satisfy this constraint, leading to significant differences between actual response characteristics and theoretical ones (Pauley and Wu 1990; Askelson et al. 2005b; Spencer et al. 2007). Such differences can lead to adverse impacts in the resulting analyses (e.g., Doswell and Trapp 1997). Given this, the benefits of a correction pass are expected to be limited and we have chosen to focus efforts on dealing with data irregularity and the effects such irregularity has on the analyses.

Note that the Barnes Gaussian weight function was chosen over another common DDWA weight function (Cressman), as the Cressman weight function does not asymptotically approach zero. Instead, it becomes zero abruptly at the radius of influence (ROI) (Cressman 1959; Koch et al. 1983), which can produce discontinuities or noise distortion in the analysis [e.g., ringing (Koch et al. 1983; Trapp and Doswell 2000; Askelson et al. 2000)].

2.4 Snowfall Rate and Size Distribution Variability

Hydrometeor size-distribution variability is responsible for some of the error inherent to *S* estimates obtained using radar data. This occurs because reflectivity factor and snowfall rate are different moments of the hydrometeor size distribution (Doviak and Zrnić 1993, 219-222), which

can be exacerbated by potentially large hydrometeor density variations inherent to snowflakes (Rasmussen et al. 1999).

Radar reflectivity factor (Z) is the 6th moment of the hydrometeor size distribution and is computed by summing across all of the hydrometeors in a unit volume (Doviak and Zrnić 1993, 219-222). A general form for Z, denoted as Z_{ia}, can be formulated for a population of snowflakes (sometimes referred to as ice-air hydrometeors), by using the Debye (1929) model to estimate the dielectric constant (K) for an ice-air mixture (Bohren and Battan 1980; Askelson 2002). The dielectric constant of a hydrometeor is important as Z_{ia} depends upon K, and can vary significantly depending upon the substance being observed (K is ~ 0.93 and ~0.176 for liquid water and ice, respectively; Smith 1984). With a mixture, K is dependent upon the relative amounts and the dielectric constants of each material in that mixture (Bohren and Battan 1980; Askelson 2002). Using Debye's (1929) K relation with a spherical shape assumed, the reflectivity factor for snowflakes/ice-air hydrometeors (Z_{ia}) is equal to that of a spherical, solid-ice hydrometeor that occupies the same volume as the ice in a snowflake (Smith 1984; Askelson 2002):

$$Z_{ia} = \int_{D_{ia}=D_{ia\_\min}}^{D_{ia}=D_{ia\_\max}} N(D_{ia}) \left(\frac{\rho_{ia}}{\rho_i} \right)^2 D_{ia}^6 dD_{ia} \tag{7}$$

where D_{ia} denotes the diameter of a snowflake, ρ_{ia} and ρ_i denote the density of an snowflake and a solid-ice hydrometeor, respectively, and D_{ia_min} and D_{ia_max} are the minimum and maximum diameters of the snowflakes (Askelson 2002).

In contrast to Z_{ia}, the *liquid-water equivalent* snowfall rate (S) is approximately the 3.31th moment of the snowflake size distribution (assuming aggregated snowflakes and a constant density for snowflakes as a function of diameter; Doviak and Zrnić 1993, 219-222). In order to compute the *liquid-water equivalent* snowfall rate S, the diameter a snowflake would have if melted into a sphere with density, ρ_1 =1 g cm^{-3} is needed (Askelson 2002). This diameter, D_1, is given by

$$(D_1)^3 = \left(\frac{\rho_{ia}}{\rho_1} \right) D_{ia}^3 \tag{8}$$

which is derived from mass conservation principles. With this, S is given by

$$S = \frac{\pi}{6} \int_0^\infty N(D_{ia}) \left(\frac{\rho_{ia}}{\rho_1} \right) D_{ia}^3 w_t(D_{ia}) dD_{ia} \tag{9}$$

where w_t denotes the terminal velocity. (9) can be easily compared to (7) as they are formulated in a similar manner. Furthermore, (9) can be rewritten in terms of the density D_1 as

$$S = \frac{\pi}{6} \int_0^{\infty} N(D_{ia}) D_1^3 w_t(D_{ia}) dD_{ia} \tag{10}$$

Snowflake density can vary over two orders of magnitude (Rasmussen et al. 1999); such variations can affect (9) and (10). Because the terminal velocity,

$$w_t = \left(\frac{4g\rho_{ia} D_{ia}}{3C_D \rho_a} \right)^{1/2} \tag{11}$$

[where g is the acceleration due to gravity, ρ_a is air density of air, and C_D is a drag coefficient] also depends upon snowflake density, such density variations also impact the terminal velocity (Rasmussen et al. 1999]. In (9) and (10), terminal velocity's dependency upon ρ_{ia} is manifested by an implicit assumption that the snowflake is a spherical particle.

As a result of Z_{ia} and S being different moments of the hydrometeor size distribution, variability, and thus, error, is inherent in S estimates obtained from radar data, even if there are no density variations. This is evident from comparing (7) with (9). A result of such variability is that a single Z value may not provide a unique S estimate (Doviak and Zrnić 1993, p. 224). Consider two different size distributions that produce the same Z_{ia} value. In this situation, Z_{ia} can remain unchanged even when the number of hydrometeors that exists within each bin of a set of binned size ranges is altered. However, the two distributions would have different S values.

Unfortunately, an indefinite number of parameters is required to know all that is needed regarding a snowflake size distribution to correctly compute snowfall rate. Generally only one parameter is available (one Z_{ia} value which is used in a Z_{ia}-S power law relation) (Doviak and Zrnić 1993, 222-224).

In addition, the variability of size distributions is multifaceted. Snowflake size distribution depends upon the type of hydrometeors, their environment, and associated microphysical processes (Doviak and Zrnić 1993, 210-222; Lee and Zawadzki 2005). Size distribution can vary with geographic origin (i.e., maritime versus continental airmasses), time, space, precipitation type (e.g., rain, snowflakes, hail), and microphysical processes (e.g., warm-rain mechanism) (Rogers and Yau 1989, 170-172; Doviak and Zrnić 1993, 210-216; Lee and Zawadzki 2005). Lee and Zawadzki (2005) further indicate that (drop) size distribution variability has been found to explain 30-50% of errors in rainfall intensity estimations when a single Z/rainfall rate relationship is used. Due to the aforementioned issues related to snow, these errors are expected to be compounded, resulting in greater inaccuracies.

As with Z_{ia}, S is also a function of density, which, for snow, is itself dependent upon several factors, including: (1) size – snowflake density is inversely proportional to snowflake diameter (Rasmussen et al 1999); (2) environmental air temperature and supersaturation – which determine the habit and growth rate of ice crystals (Rogers and Yau 1989, 158-163; Pruppacher and Klett 1997, 40-43); (3) degree of wetness of the snowflake [where 'dry ' snowflakes are

assumed to be unmelted and unrimed and ‘wet’ snowflakes are assumed to be snowflakes that are partially melted and/or rimed (Rasmussen et al. 1999)]; and (4) degrees of aggregation, riming, and/or melting (Rasmussen et al. 1999). ‘Wet’ snowflakes have higher densities than ‘dry’ snowflakes, due to the existence of pore spaces in such snowflakes that are filled with other ice crystals or liquid water when wet (Rasmussen et al. 1999). Thus, ‘wet’ snowflakes have higher liquid water contents, and, therefore, higher densities, than dry snow, since the pore spaces of dry snow are filled with air rather than water (Rasmussen et al. 1999).

3 DATA

As previously stated, three types of observation platforms (*in situ* surface instruments, radar, and satellite) are used in this study, as are model data. For these data sources, five datasets are available and are used in PPAES – two surface datasets and a single dataset each for radar, satellite, and model. The five datasets are: (a) MADIS surface observation data, (b) *Clarus* surface observation data, (c) NEXRAD Level II radar data, (d) Geostationary Operational Environmental Satellite (GOES) satellite data, and Rapid Refresh model data. For brevity, the NEXRAD Level II and GOES Satellite data will be discussed briefly as the main focus of this study is the integration of surface data into PPAES and fusion with the radar data.

3.1 MADIS Surface Observation Data

MADIS surface observation data sets contain observations from many surface observation networks and providers, including aviation automated and non-automated weather reports (METARs), Canadian Surface Aviation Observation (SAO) reports, mesonet reports, and observations from climate and other networks (MADIS 2010). MADIS provides quality control processing of the surface data it distributes (MADIS 2010). Of the various MADIS datasets available, METAR stations comprise the only dataset used in this study. The METAR-reporting dataset includes both Automated Surface Observation System (ASOS) stations and Automated Weather Observing System (AWOS) stations. We limit our scope to METAR-reporting stations because of their prescribed reporting frequency and the specific weather elements required by the PPAES algorithm. Other surface data exist, but either do not provide the necessary information in real-time (e.g., the NWS Cooperative Observer Program) or do not have the fields needed by PPAES.

Specifically, METAR reports contain the atmospheric fields of interest to PPAES: temperature, wind speed, visibility, and present weather. METAR reports are provided in a uniform format and with routine reporting schedules (every hour and at times when triggered by certain meteorological conditions). Additionally, the large number of METAR-reporting stations provides good geographic coverage over the bulk of the United States as well as other areas of the globe, with well-established siting requirements and known instrument performance (NWS 1998).

MADIS quality checking procedures are completed before data are disseminated, which minimizes the need for application of post-acquisition quality checking. Challenges associated with MADIS observations include reference time selection (defined as the central time of the bin that contains the largest number of observations after they are sorted into time bins), time-window selection, the need to develop a scheme for use of non-automated stations,[2] and relationships between precipitation phenomena reported as present weather and precipitation rates. The MADIS quality checking results and how they are utilized in this study are discussed in detail in Section 4.

[2] Reports from automated stations are labeled AO1 and AO2 to indicate the absence or presence of a precipitation discriminator, respectively, while no information as to whether a station can provide precipitation information is provided if it is not automated.

To find the reference time for a data file, we sort the observation records into time bins. A total of 15 bins is used, with 14 bins separated by five-minute intervals and a "catch-all" bin for records that fall outside the PPAES spatial domain or outside the defined time bins. The bins are defined around a given METAR hour, with the first bin starting at 40 to 45 minutes before the given METAR hour. The reference time is then set to the central time of the bin that contains the greatest number of records. The reference time has been routinely seen to be seven minutes before the reporting hour (e.g., if the METAR hour is 16 UTC, the reference time would likely be 15:53 UTC).

For the current study, we prescribe a 30 minute time window within which observations are used. Thus, a station's record must fall within ±30 minutes of the reference time to be used. The basis for the use of a ±30 min time window is that METAR reports are provided every hour (up to 15 minutes before the hour and 44 minutes after the hour) and we wish to include as many reports as possible to create a comprehensive surface analysis of precipitation occurrence and rate. When multiple observation records for the same station are provided, the observation record that is nearest to the reference time of the analysis is used.

A scheme to use both manned stations and AWOS stations (that lack clarity in their capabilities), is needed since not all manned stations report present weather and information regarding which manned stations do and do not report this information is not readily available (metadata is lacking). Without such information, these stations cannot be separated into stations that can report precipitation information and those that cannot, leading to biased analyses since if we did not know which stations actually provided precipitation information, we would have to assume that a non-report means no precipitation. Underrepresentation of the spatial extent of precipitation would be the likely result of such an approach. Therefore, a present-weather station climatology routine was developed to catalog which manned stations report present weather so that they can be properly included into PPAES analyses. The challenge associated with developing a present-weather station climatology is discussed in more detail in Section 4.

Finally, the challenge of relating present-weather precipitation reports to individual precipitation rates is handled through the use of multiple logic trees to derive precipitation rates from present weather group information. Additional information regarding this is also provided in Section 4.

3.2 Clarus Surface Observation Data

The Clarus Initiative of the U.S. Department of Transportation (DOT) Federal Highway Administration (FHWA) Road Weather Management Program and the Intelligent Transportation System (ITS) Joint Program Office was intended to reduce the impact of adverse weather conditions on surface transportation users by enhancing the utility of Road Weather Information System (RWIS) data (Clarus Initiative 2010). Through Clarus, the utility of RWIS platforms is improved through the enhancement (via quality checking routines) and dissemination of data from the collective component of an RWIS – the Environmental Sensor Station (ESS) (Manfredi et al. 2005; Clarus Initiative 2010). The dissemination of these data, provided by state and provincial transportation agencies, through the organization's web-portal (http://www.clarus-system.com/). As of 2010, 37 states and three Canadian provinces participate in this voluntary system, with additional states' participation pending or being considered (Clarus Initiative 2010).

Given recent Clarus Initiative efforts, it is prudent to incorporate RWIS ESS data into PPAES, as the use of additional surface observation data could result in improved analyses of precipitation occurrence and rate. The inclusion of RWIS data is not redundant in this study because they are not present in the specific MADIS METAR dataset previously described (MADIS 2010). Compared to MADIS observations, Clarus surface observations are different in several ways, including limited spatial coverage over the U.S. (Figure 1), the fact that observations are reported at the discretion of the volunteering agency (with no predetermined reporting time interval used), and finally that ESS installations do not have mandatory siting requirements--different state agencies utilize different instruments in their ESS installations (Manfredi et al. 2005; Clarus Initiative 2010).

Several challenges have been encountered with Clarus surface observation data, including inconsistencies in information provision owing to the system being voluntary; selection of a reference time; time-window selection; inconsistencies/bugs encountered with Clarus data; and relating Clarus precipitation information to precipitation rates. These challenges are the same or similar to those associated with the use of MADIS data.

A caveat related to the Clarus system is how and how often the information is shared; irregular sharing limits the utility of some RWIS data. Therefore, if not enough information is available to derive a precipitation rate or, if data records are reported but are outside of the surface data time window, then such RWIS data are not utilized in the PPAES surface analysis algorithm.

The method for selecting a reference time for Clarus data is identical to that used with MADIS data except that nine bins are used, with eight 15-minute bins and one bin for records that fall outside of these eight bins. The bins are defined to be set around the hour of a Clarus report, with the first bin starting at the top of the previous hour. A lesser number of bins having larger widths is utilized for Clarus data because observation records are provided by the Clarus system regardless if they are late or on time – such data represents new information reported to the Clarus system during the time period within which it was delivered (Brenda Boyce 2010, personal communication). The result of late observation records is that they are included even though they are outside of the intended reporting hour. This requires a filtering step so that only Clarus data relevant to the larger MADIS dataset are used. In this way, we can fuse the two datasets together to provide a comprehensive precipitation analyses. Thus, a larger time interval and a smaller number of bins are used with Clarus data to account for the fact that Clarus records can be distributed over a greater period of time before and during the reporting hour.

The reference time for Clarus data is set to the central time of the bin that contains the greatest number of records. This reference time is then compared with the ‘an_dt’ value set in the ppaes surface control file to assign a date/time stamp to the analysis file name. The default an_dt value is 20 minutes. Since MADIS data only arrive every hour, the analysis file name generally will have a timestamp at the top of the hour as long as the MADIS reference time is between ±10 minutes before and after the top-of-the-hour, respectively. However, since at UND Clarus data are organized in files spaced 20 minutes apart, the surface module would then produce analysis files at 0, 20, and 40 minutes after the hour. Since no MADIS data files are available for those 20

and 40 minute analysis files, they will only contain analyses of Clarus data. Therefore, these two analysis files (20 and 40 minutes after the hour) should not be used.

One could change the an_dt value to generate only one analysis file every hour, but then one has the risk of having an analysis file that has hour-old data that would misrepresent the current precipitation field. It is suggested, therefore, that one use a 20 minute an_dt value and thus only process the Clarus data file that is generated 20 minutes after the hour (as this file would normally produce an analysis time at the top-of-the-hour). This way, the Clarus dataset (~750 stations that can report precipitation) supplements the larger MADIS dataset (~1500 possible precipitation reports).

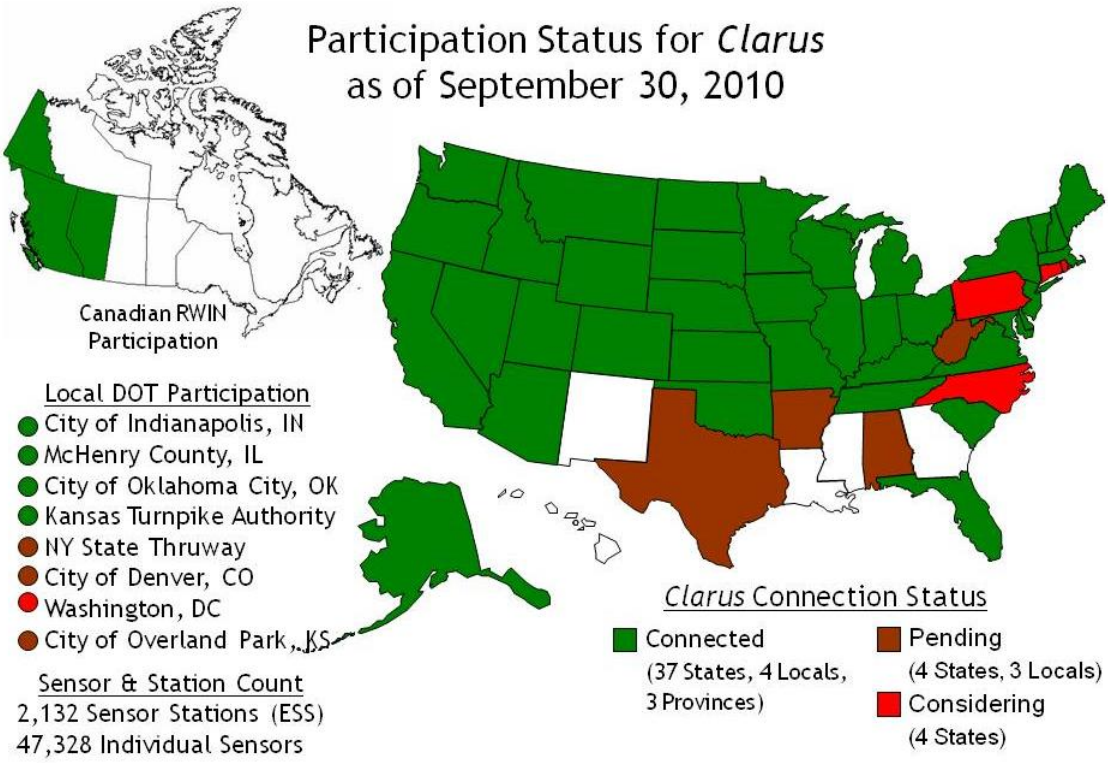

Figure 1. State, provincial, and local DOTs participating in Clarus initiative

Note that Considering means that a state or local DOT has shown significant interest (www.clarusinitiative.org/ accessed November 13, 2010).

3.3 NEXRAD Level II Radar Data

Software, processing routines, and quality checking measures associated with the PPAES radar module were initially reported by Askelson (2008). The radar data used in PPAES are from the S-band WSR-88D network, containing over 160 sites (Fulton et al. 1998). Specifications for WSR-88D radars can be found in Crum et al. (1993) and Rinehart (2004, p. 382). From these radars, Level II equivalent radar reflectivity factor, Z_e, data are used to compute instantaneous precipitation rates using the power law relation $Z_e = \alpha S^{\beta}$, where α and β are equal to 150 and 2.0

and the units of Z_e and S are $mm^6\ m^{-3}$ and $mm\ hr^{-1}$, respectively (Super and Holroyd 1998). The results presented by Super and Holroyd (1998) suggest that this Z_e-S relation is adequate for use for radar sites in metropolitan Minneapolis and Cleveland. Furthermore, the above Z_e-S relationship was tested over five radar sites (Aberdeen, SD; Grand Forks, ND; Minneapolis, MN; Bismarck, ND; Duluth, MN) and was shown to be a practical Z_e-S relation for use in the Great Plains (Holroyd 1999).

Twenty eight WSR-88D sites are used for testing and evaluation. These radar sites encompass the Great Plains and the Midwestern states, as shown in Figure 2. Station call letters are in white letters.

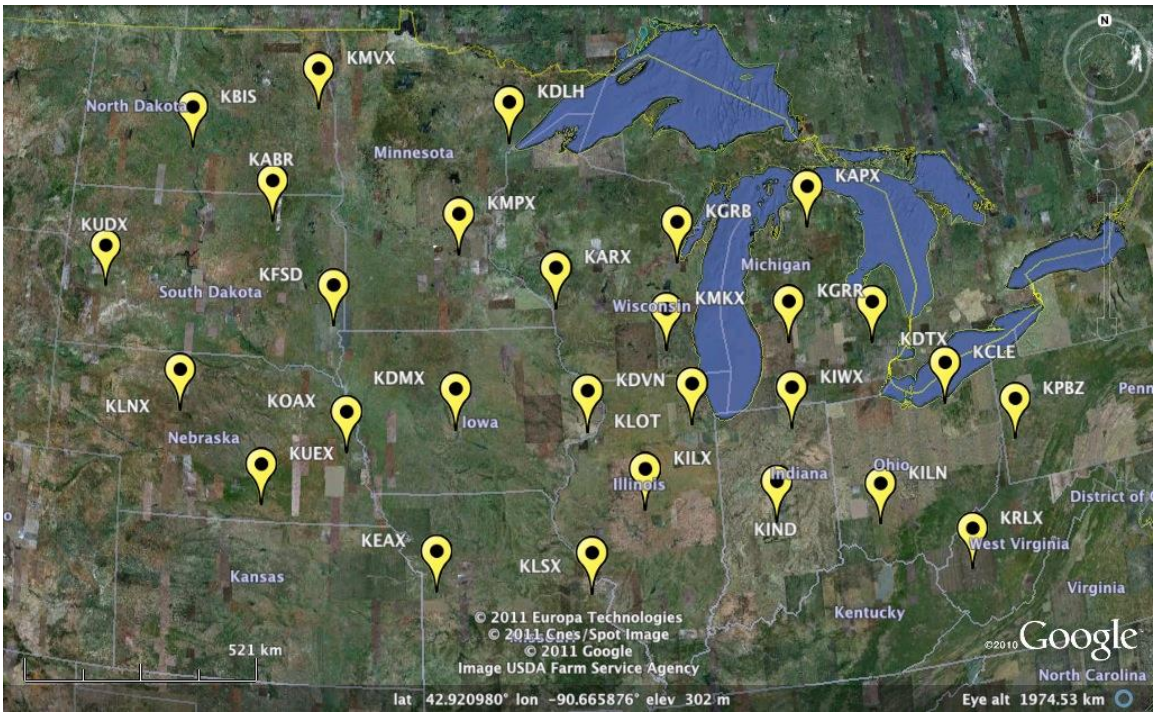

Figure 2. NEXRAD WSR-88D radar sites used in study

3.4 GOES Satellite Data

Software, processing routines, and quality checking measures associated with the PPAES satellite module were also initially described by Askelson (2008). The satellite data used in this research are from GOES sounder-derived Cloud Top Pressure (CTP) products over North America and the adjacent oceans. GOES CTP products are derived from GOES radiance measurements in the CO_2 absorption band using a CO_2-slicing or absorption technique (Hawkinson et al. 2005). Although GOES CTP data have not been extensively used to estimate precipitation rate, they are a viable source of information and have been used frequently in the meteorological community for initializing regional numerical models and for regional studies of cloud amount and type. A short summary of characteristics for the GOES sounder is provided by Hawkinson et al. (2005).

GOES CTP products are currently used in PPAES to estimate precipitation rates where observation 'holes' are present in the radar network (Askelson 2008). The technique involves merging CTP information with PPAES radar module instantaneous precipitation rates after the two datasets are spatially matched. Important requirements for the satellite-based routine are that a prescribed threshold of CTP and radar precipitation-rate pairs must exist so that a relationship between CTP and snowfall rate can be developed, that hole filling is performed only in locations where snow is expected, and that hole filling is performed only within a limited distance of a radar (Askelson 2008). In addition, because GOES CTP retrievals only occur hourly, one is limited in how often the PPAES satellite module can be used to produce analyses (Hawkinson et al. 2005).

3.5 Rapid Refresh Model Data

Use of rapid refresh model (RAP) radar reflectivity to refine analyses based upon radar data, especially in areas of complex terrain and remote regions, is examined herein. The purpose is somewhat similar to that for satellite data, except with model data one can also apply virga correction (when precipitation is indicated at high altitudes but model data indicate that it is not reaching the ground). The Rapid Refresh is supported by the Earth System Research Laboratory (ESRL) and replaced the RUC as the NOAA (National Oceanic and Atmospheric Administration) hourly-updated assimilation/modeling system (ESRL 2013).

The horizontal grid spacing of the RAP data used herein is 13.545 km. While higher resolution data are available with the High-Resolution Rapid Refresh (HRRR) model (3 km horizontal grid spacing), owing to time constraints these data were not evaluated herein. With RAP output, estimated radar reflectivity fields are provided at 1 km and 4 km AGL. While both of these fields are ingested into PPAES when radar and model data are blended together, only the 1 km data are used because they are the most relevant to surface precipitation. In the future, data from both 1 and 4 km could be used to estimate properties like the Vertical Profile of Radar Reflectivity (VPRR) to enhance diagnosis of precipitation.

In addition to the radar reflectivity data, surface geopotential height data provided in RAP data files are also used to estimate the MSL heights of of RAP 1 km AGL reflectivity fields so that the relative heights of radar-based precipitation estimates and model-based precipitation estimates can be determined. With this, one can determine which (radar or model) estimate is closer to the ground and, thus, provides a better estimate of what is happening at the surface.

It is noted that fields like surface precipitation are provided with the RAP model. Such fields could be used to estimate precipitation occurrence and intensity. However, models are known to suffer from significant phase (location) and amplitude (intensity) errors when representing precipitation. By combining model data with radar observations, the hope is that these errors will be somewhat mitigated while providing a more coherent precipitation picture that is not always available with relatively widely-spaced radars.

4 METHODOLOGY

The main focus of this study is advancing the development of PPAES through the blending of analyses derived from *in situ* surface and radar observations. The main developmental areas are (a) aggregating and blending multiple surface observation datasets (MADIS and *Clarus*) together, (b) deriving precipitation occurrence and rate from surface observations, and (c) blending the two analyses – radar and surface. Other areas that were examined were the use of model and satellite data to aid in providing precipitation information in areas of complex terrain. However, due to time constraints, only model data were included in the complex terrain algorithm. Further analysis is needed to include satellite data in this algorithm.

4.1 Surface Observation Data Quality Checking

The primary surface data quality measures employed in PPAES utilize the quality control (QC) checks or quality checking (QCh) services provided by the two data providers, MADIS and *Clarus*. Both datasets allow the user to decide whether or not to utilize a particular measurement based on the QC or QCh results. After surface data are acquired, PPAES uses the QC and QCh results from each dataset, in addition to specific checks employed by PPAES, to filter each dataset. A station's record is only used in PPAES after it passes the required checks described below. Because each dataset has different quality measures, each dataset is handled separately.

4.1.1 MADIS Quality Measures

MADIS performs two categories of QC checks: static (i.e., single-station and single-time QC checks employed without knowledge of the current meteorological situation) and dynamic (i.e., checks that take advantage of other available hydrometeorological information) (MADIS 2010). As such there is a maximum of three possible levels (the higher the level – the higher the complexity of the QC checks) for which an observation can pass a quality check (MADIS 2010).

The three levels of automated QC for MADIS include:

- Level I checks for validity and position consistency that compare the measurement against a specified set of tolerance limits (Haskins 2006; MADIS 2010).
- Level II checks for field consistency; these may include internal (i.e., meteorological relationships among specific observations), temporal (field changes over time), and/or statistical spatial consistency (i.e., weekly statistics such as an observation must pass a specific percentage of any QC check) checks (Haskins 2006; MADIS 2010).
- Level III checks for spatial consistency that utilize an optimal interpolation scheme to compute the difference between the scheme-estimated value and the measured value (Haskins 2006; MADIS 2010). If this difference is small enough, then the Level III check is passed. Otherwise, observations are then inspected to figure out which value (the scheme-estimated or measured) is suspect.

Table 3 displays the atmospheric fields of interest to this study along with each field's corresponding maximum MADIS QC level and required level for use in PPAES analyses. In PPAES, requirements for the minimal QC level are set using PPAES control variables. As seen

in Table 3, the automated QC level 1 must be passed for visibility and 1-hour precipitation accumulation observations. Temperature and wind speed were required to pass automated QC level 2. Note the absence of the present-weather MADIS field. Unlike most measurements, present weather measurements are not included in MADIS QC processing. For a more detailed description of all of the automated QC checks, refer to the surface QC section of the MADIS website (available online at http://madis.noaa.gov/madis_sfc_qc.html) or Haskins (2006).

Table 3. Required MADIS quality control levels for atmospheric fields of interest in PPAES

MADIS Dataset Measurement	Max Possible Level	Minimum Required Level
Visibility	1	1
Temperature	3	2
Wind Speed	3	2
1 hr Precipitation Accumulation	1	1

The levels (1, 2 ,3) correspond to the passing of automated QC level 1, passing of automated QC levels 1 and 2, and passing of automated QC levels 1, 2, and 3, respectively.

4.1.2 Clarus Quality Measures

The Clarus system provides quality checking for the Clarus dataset so that a user of such information can decide whether or not to utilize an observation. As with the MADIS dataset, not all measurements have the same checks applied. A total of ten quality checks are employed in the Clarus System and four results are possible: pass; fail; available but does not run; check not supposed to run.

The ten quality check (QCh) services are: Barnes spatial QCh, climate range QCh, dewpoint temperature QCh, like-instrument QCh, persistence QCh, sea level pressure QCh, sensor range QCh, step QCh, manual flag, and sequence complete check (Clarus 2010). Similar to the MADIS QC checks, Clarus QCh services range from simple threshold and temporal checks to more complex checks (eg., the Barnes Spatial QCh) (U.S. DOT 2005). For details regarding Clarus QCh services, refer to the Clarus System Detailed System Requirements – Appendix B (available at: http://www.clarusinitiative.org/documents/Final_Clarus_System_Detailed_Requirements.pdf, U.S. DOT FHWA 2005).

Table 4 displays the Clarus QCh services and requirements for use of an observation in PPAES. Akin to MADIS, the requirement for an observation record to be used in PPAES is that they must pass the QC criteria set forth in the PPAES control variables. In this version of PPAES, Clarus observation records are required to not fail any of the ten QCh services for any given observation record. Note, MADIS and Clarus have different requirements attributable to the two having different checks applied as well as the fact that MADIS does not apply all of the QC checks to every surface variable.

Table 4. Clarus quality checking services and associated PPAES requirements

Clarus Quality Check	PPAES Requirement
Barnes Spatial	P, -, or /
Climate Range	P, -, or /
Dewpoint Temperature	P, -, or /
Like-instrument	P, -, or /
Persistence	P, -, or /
Sea Level Pressure	P, -, or /
Sensor Range	P, -, or /
Step	P, -, or /
Manual	P, -, or /
Sequence Complete	P, -, or /

Notes:
P denotes QCh passing
N denotes QCh not passing
- denotes QCh is available but does not run
/ denotes QCh is not supposed to run

4.2 Surface Observation Density Analysis

A metric was developed to serve as a measure of local station spatial density, defined here as the as the average minimum spacing between surface observation stations. Values of this metric dictate the region of influence of an observation in the precipitation occurrence/non-occurrence algorithm. The region of influence, in turn, serves as an input value for the computation of the smoothing parameter in the Barnes scheme (Barnes 1964; Koch et al. 1981,1983) and controls whether or not the observation value can affect the value of the analysis at a given grid point. In PPAES, a surface observation modifies an analysis value if the distance between the analysis grid point and the observation location is less than the average minimum spacing between stations (previously computed) at the observation location multiplied by a factor.

A guard barrier is defined by Doswell and Lasher-Trapp (1997) as a finite area where only the information inside that region are considered, "such that the results near the edge of the guard barrier are indistinguishable from those deeper within the data lattice." As such, the local data density information is used to define a guard barrier that is used in the surface analysis of precipitation occurrence/non-occurrence. We selected a guard barrier of $4\Delta d$, where Δd is the mean data spacing. This value was chosen following Doswell and Lasher-Trapp (1997) who point out that a guard barrier value larger than $4\Delta d$ utilizes a larger part of the data domain in the objective analysis with only a negligible improvement in the analysis. They also note that a guard-barrier less than $4\Delta d$ can result in not enough of the domain being utilized, resulting in adverse impacts to the analysis, especially with irregular data (Doswell and Lasher-Trapp 1997).

Note that unlike in Doswell and Lasher-Trapp (1997), where the guard barrier was applied such that the objective analysis scheme was not impacted adversely by "feeling" internal data boundaries due to irregular data spacing), the guard barrier's function herein is different. In

PPAES, the guard barrier value serves as the region of influence value used in the precipitation occurrence/non-occurrence analysis. The value of 4Δd ensures that no gaps will exist between values in the precipitation occurrence/non-occurrence analysis.

Another key aspect of the precipitation occurrence algorithm involves determining the number of stations, N, that is representative of a given region. Unfortunately, there is no established way for determining this number and, as such, it was determined through experimentation guided by Petterson and Middleton's (1963) result that a horizontal wavelength must exceed 2Δn (Δn denoting average observation spacing) to be resolved. In addition, since "five data points are required to describe the wave and its derivatives" (Petterson and Middleton 1963), it can be argued that a minimum of five observations is needed to compute the average minimum data spacing. As a result, the average minimum data spacing plays an important role in what can and cannot be resolved.

Since a minimum of five stations is required to describe a wave, a way to maximize the wave's description is desired. However, the irregularity of the observation network will impede the sampling process and the analysis scheme (Doswell and Lasher-Trapp 1997; Spencer et al. 2007). As such, a number greater than five stations is likely more suitable for the problem. Walters (2000) and Pielke Sr. (2001) both share the opinion that "as many as 10 grid points," or 10 sampling points, may be necessary to reasonably represent a wave's amplitude. This is because errors in sampling – in amplitude, phase, or both- will limit the capability to resolve the wave or the numerical solution trying to be described (Walters 2000). If a wave is sampled out of phase with the observation network, then it will not be resolved accurately. From Walters' (2000) argument, 10 stations is a better number than 5 for approximating the average minimum data spacing.

The number of stations within a given area used to compute the average local data density is important. If too few stations are utilized in the computation (e.g., three), then an area may be labeled sparse or clustered when that may not be the case. Similarly, if too many stations are utilized (e.g., 25), the value obtained may not be indicative of that local area. Both scenarios lead to a misrepresentation of the local data density.

Because of these issues, tests were performed to determine which value of N is best suited for this application – N=5 or N=10. The test cases involved two precipitation scenarios, one where radar overshoot had been identified from past PPAES case logs and another where a large precipitation shield was present where overshot was not a concern. Four events total were used to evaluate the efficacy of using these two values of N.

The method for computing the local station spatial density is as follows:

- For each analysis grid point, the N nearest observations that are closest to that point are located. Only stations reporting present weather information are included. Thus, only stations eligible to be used in the surface analysis are utilized in this metric.
- For each individual station, the distances between that station and the other N-1 stations are

computed to find the closest station. That distance is then saved for that station. Subsequently, the nearest station is identified for each of the remaining stations.

- The mean of the resulting N distances is then computed and is used as the local data density for that analysis grid point. Thus, the average minimum distance between observation stations surrounding the analysis location is represented by this value. This is completed for each analysis grid point on the analysis grid.

The local station spatial density metric was computed only for MADIS surface data, owing to the far greater number of stations, greater coverage area, and greater consistency in reporting time (right before the hour) compared to the *Clarus* dataset. Once computed, the local station density field is saved for future use and is not recomputed unless substantial changes in the observing network occurs.

Local station spatial density values for N=5 and N = 10 are shown in Figures 3 and 4, respectively. A "stretching" of values occurs outside the CONUS over the oceans and Mexico. This is a result of analysis grid points over the oceans or Mexico being associated with identical sets of stations, resulting in an unrealistic local data density field in these regions. However, this does not impact this study because these issues only occur outside of the domain of interest at present.

As previously mentioned, the local station spatial density metric is used in the precipitation occurrence surface analysis for computing the guard-barrier. After testing was completed, N=10 stations was determined to be best suited for this study. In regard to the completed tests, the surface test analyses were similar in how they depicted precipitation; this can be surmised from how similar Figures 3 and 4 are, especially over ND, SD, NE, MN, IA, WI, and IL. However, minor differences are apparent between the N = 5 and 10 figures that give use reason to prefer one value of N for the local station density computation. Notably, the N = 5 stations plot (Figure 3) is less smooth and would result in greater spatial variability in the smoothing properties of the analysis than desired, specifically in this study's area of interest (the Great Plains and Midwestern States). The N = 10 stations plot (Figure 4) is smoother and increases the local station spatial density values in areas where they were less than 20 km for N = 5.

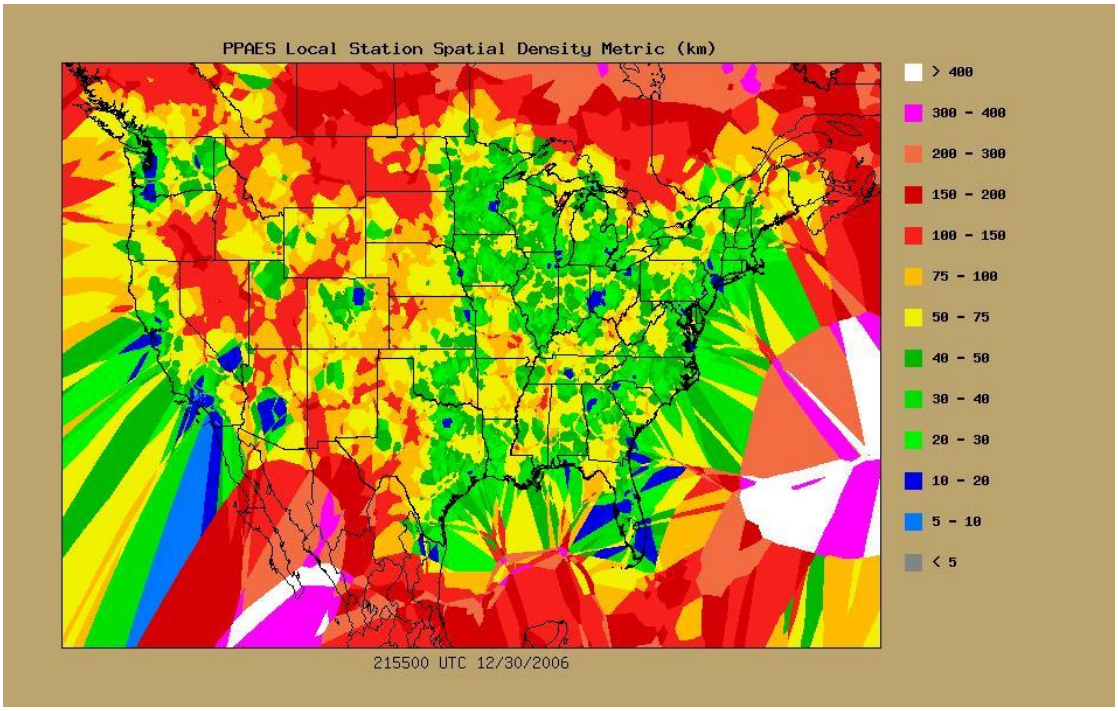

Figure 3. Local station spatial density (km) for MADIS data with N=5 from 2006

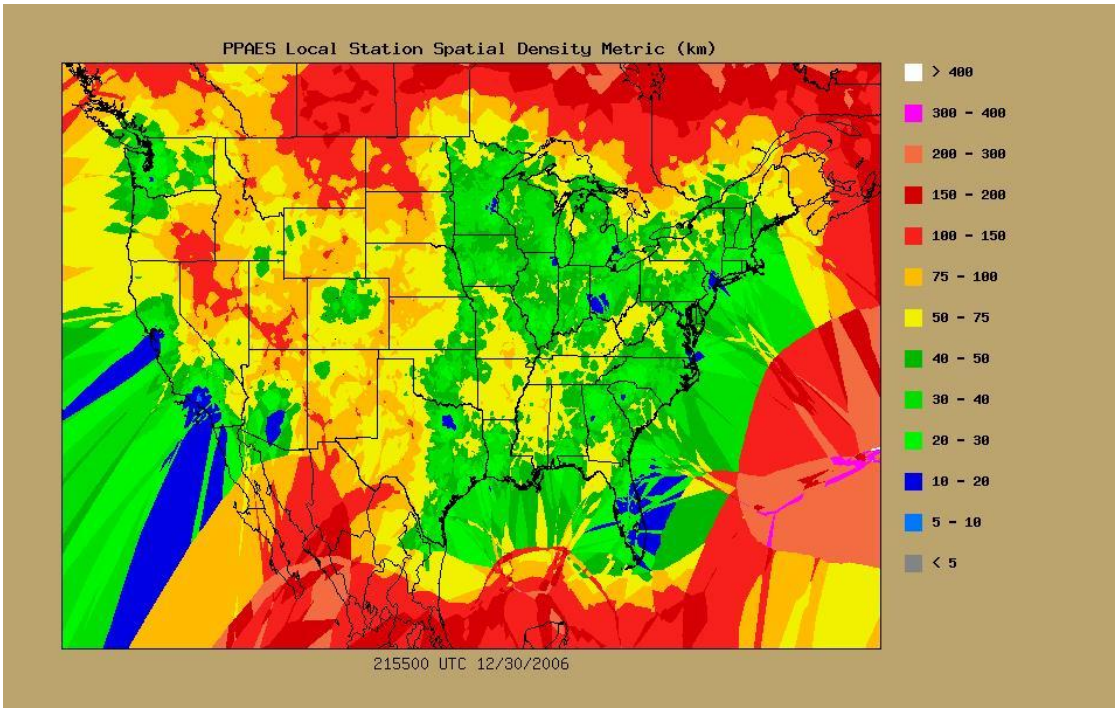

Figure 4. Local spatial density (km) for MADIS data with N=10 from 2006

Note that some areas (e.g., MT and SD) saw a decrease in the local station spatial density with N = 10. The local station spatial density values were also seen to increase or decrease in areas of where local spatial densities were either ample (< 20 km) or sparse (e.g., > 50 km), respectively,

in the N = 5 and 10 local station spatial density analyses. Because the N = 10 analysis is smoother and there are benefits of having a greater number of stations for sampling (Walters 2000), we chose N = 10 and feel that with N = 10 and a guard barrier equal to $4\Delta d$, we should avoid excessively extrapolating at the edges of the observational domain as well as creating empty areas in the analysis.

Towards the end of this project, the computer code used to generate station spatial density was reorganized and converted into a separate utility that can be run separate from the main surface analysis module. With this reorganization, the code was modified to increases efficiency. Some changes to the logic were also incorporated. Instead of picking the closest 10 stations to each grid point and calculating the spatial density solely using those 10 stations, the closest 30 stations were chosen for each grid point. Then, out of these 30 stations, the closest 10 stations were identified through a sorting procedure. For each of these 10 closest stations, distances to each of the other 29 stations were then computed. From that list of distances, the 10 stations closest to the analyzed station at the desired grid point were used to calculate the spatial reach and overall spatial density at the grid point. This process helps capture stations outside of the initially selected 10 stations, as the closest station to any one of the 10 initial stations may not be one of the 10 initial stations.

Due to the large amount of data points, stations, and processing needed to calculate the spatial density, the code takes approximately 12 hours to complete—depending upon computer power and number of closest stations being used. This is much more efficient than the previous code, as code required many days to complete. To test the new code, a new spatial density grid was calculated using MADIS data from 2012. The updated spatial density output is shown in Figure 5.

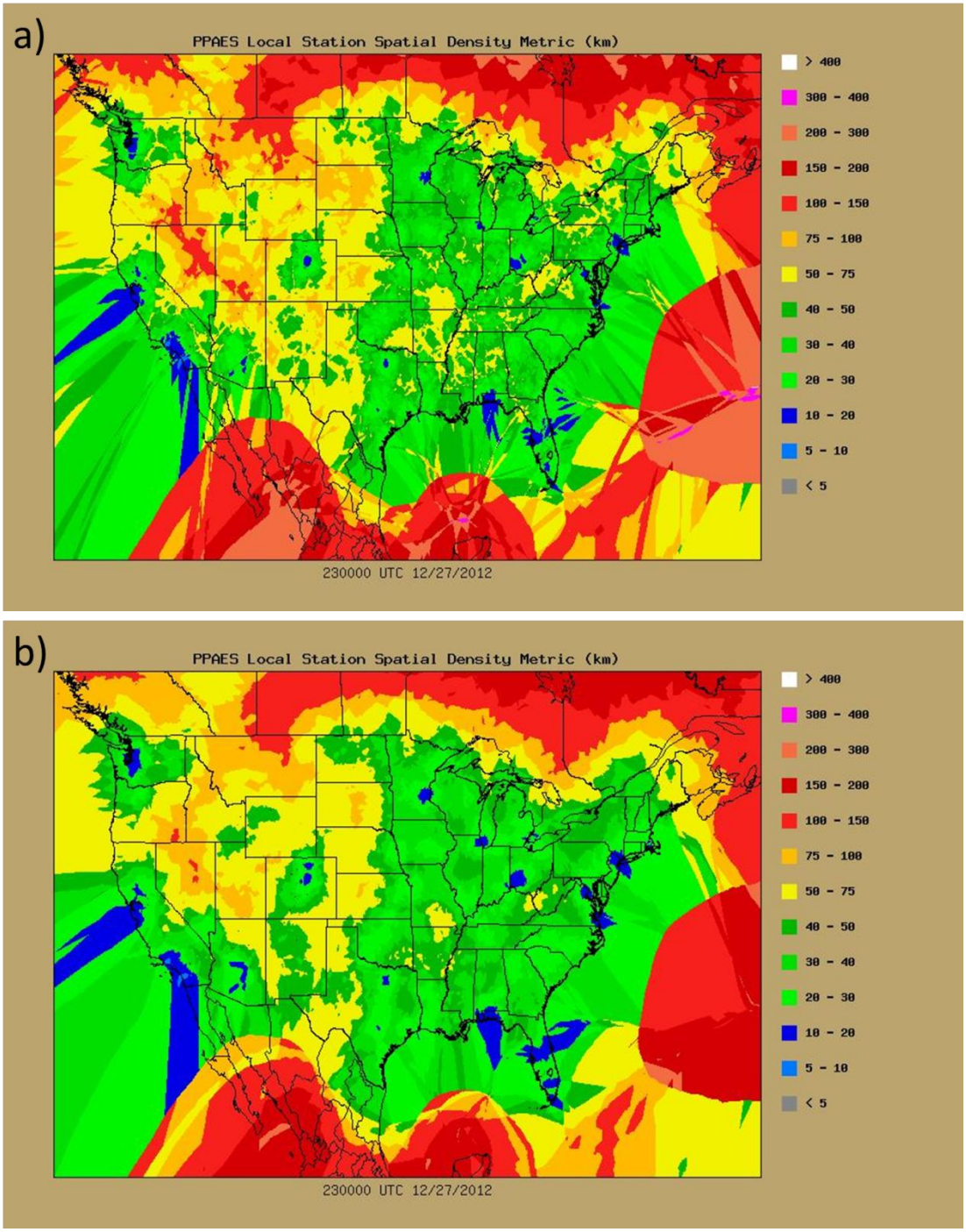

Figure 5. Local station spatial density (km) for MADIS data with N=10 from 2012 using (a) previous method for computing station density and (b) new method with 30 closest stations around each grid point used in computations

4.3 MADIS Surface Observation Climatology

Not all stations report present weather; determining which stations have this capability is not easy. As such, an automated/present weather reporting station climatology was created to

identify stations that have the capability to report present weather. It is known that AO2 stations report present weather but other stations (manned and AWOS stations) may or may not report this information and existing metadata is limited. A routine was developed to identify the stations that report present weather information and, thus, that can be used in the surface analyses. The MADIS surface observation climatology routine has three steps:

1. The directory where real-time METAR data are stored is checked every 10 minutes for new or updated METAR data files. METAR data files are updated if new reports for a METAR hour become available through the local data manager.
2. Each METAR data file is parsed to identify stations that report present weather but do not report whether it is automated (signified by AO1, AO2, or AO2A in the remarks section of a METAR). If a station is found to qualify (reports present weather but does not denote its type), information regarding this station information is stored in a climatology file along with the date of the observation.
3. Finally, whenever the PPAES surface module is executed, the surface module uses the station list maintained with the climatology routine to determine whether stations not designated as A02 can be utilized in the analysis.

The climatology module was run continuously and in real-time from 24 October 2010 through April 2012. As of 20 UTC 14 March 2011, 3,401 stations had been logged in the climatology file. This file maintains information for stations on a global scale and, as such, the climatology file contains many station IDs that are not currently used in PPAES surface analyses. Therefore, the value 3,401 does not truly represent the number of stations currently used.

4.4 Deriving Surface Precipitation Occurrence

A surface precipitation occurrence/non-occurrence analysis is the first of two analyses produced by the PPAES surface module. This is produced, in part, to mask where precipitation is occurring and not occurring for the instantaneous precipitation rate analysis.

The precipitation occurrence/non-occurrence analysis can be produced separately for each of the MADIS and Clarus surface observation datasets or jointly in one aggregated analysis. For MADIS data, this analysis is derived from each METAR's present weather measurement. For Clarus, this analysis is derived from specific precipitation-related fields. For each dataset, the precipitation occurrence/non-occurrence variable is derived using a logic tree unique to each dataset, (e.g., Figure 6). Although the framework for each logic tree is the same, differences exist between the trees relative to each dataset, mainly as a function of what variables are parsed in determining whether precipitation is occurring or not. Because of the binary nature of the field, precipitation occurrence is denoted as either as a zero for "no occurrence" or as a one for "occurrence".

4.4.1 Nearest Neighbor Algorithm

The nearest neighbor scheme is simplistic in that it assigns values to analysis grid points using the spatially closest observation, as long as the distance is less than the guard barrier at that location. The aforementioned binary nature of precipitation occurrence/non-occurrence

necessitates the use of a nearest neighbor scheme. It should be noted that the PPAES surface module was designed with the capability to aggregate and fuse additional surface datasets if they are available. To help facilitate this, additional analysis fields are saved (e.g., computed Barnes weights, distances from each analysis grid point to the nearest observation station, etc.) in case additional runs with other surface datasets are performed.

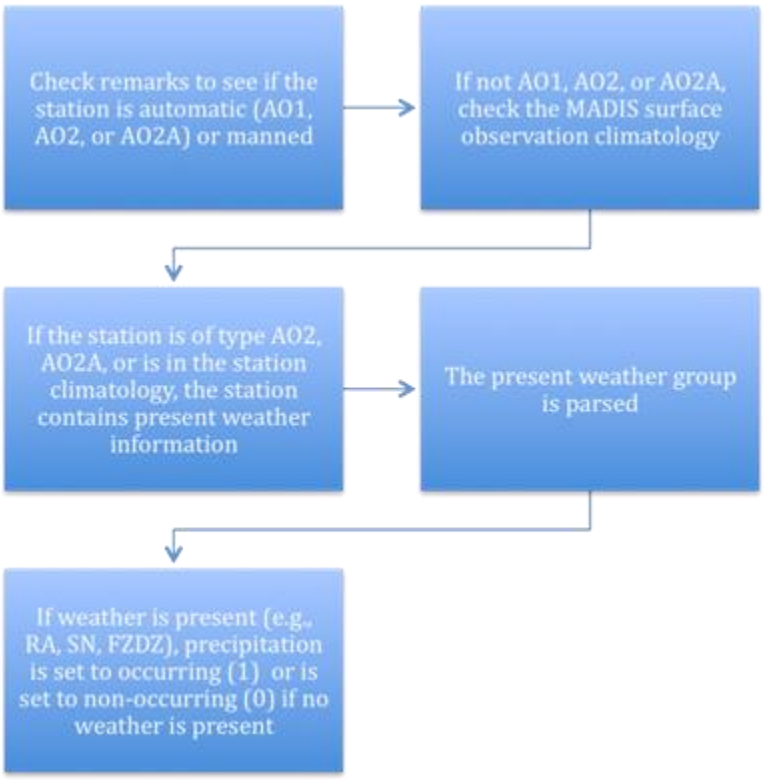

Figure 6. MADIS precipitation occurrence logic tree

The Clarus precipitation occurrence logic tree follows a similar structure, minus the need to use a station climatology, plus the fact that specific precipitation-related fields are checked in place of a present weather group.

Variables that are pertinent to the surface precipitation occurrence/non-occurrence algorithm were discussed in Section 2--analysis grid specifics, other control variables, and their values are also provided in that section. The nearest neighbor algorithm contains the following steps:

1. All surface observation data are handled station-by-station and are passed through the algorithm one at a time.
2. For MADIS, stations are compared to the present weather climatology list and are assigned a temporary variable to allow processing to proceed if necessary.
3. A subroutine is executed to determine if a station report contains relevant present weather information. If so, a station is assigned a “1” if precipitation is occurring or a “0” if precipitation is not occurring.

4. The station report is then compared with an updated and accumulating list of stations that have been previously processed to avoid using duplicate station reports.
5. Conditional statements are used to check if an observation is within the prescribed spatial domain and time window (as described in Section 2).
6. A subroutine is then executed to determine if there are any reports that are closer to an analysis' reference time. If so, that station report is used in lieu of the original report. Furthermore, the station identifier is then added to the processed station list.
7. A circular region of influence is defined to see which analysis grid points may be assigned this station's precipitation occurrence or non-occurrence value. In using a circular region of influence centered on an observation, the observation's influence is applied isotropically. The region of influence is centered on the observation location with the radius set equal to the guard barrier. The circular shape is appropriate given that not one direction (in the x- or y-axis) is desired to hold more weight.
8. Great circle distances are then computed from the observation location to each grid point to refine which of the candidate analysis points are actually assigned the observation value. If the distance is less than the value of the guard barrier, the analysis grid point is then eligible to be assigned the value from that station report.
9. Lastly, one of two requirements must be satisfied for the analysis grid point to be assigned that station's precipitation occurrence/non-occurrence value. These requirements are:
    a) No analysis value is present at that analysis grid point.
    b) The previous analysis value present at the analysis grid point is from a surface station that is farther away from this analysis point than the current station.

4.5 Estimating Instantaneous Surface Precipitation Rate

Instantaneous surface precipitation rate analyses are produced after precipitation occurrence/non-occurrence analyses. This field is produced for both MADIS and Clarus surface datasets. The objective analysis scheme used to produce these analyses is the Barnes scheme (e.g., Barnes 1964; Koch et al. 1983). Only stations utilized in the surface precipitation occurrence/non-occurrence analysis are utilized in this analysis. A rate of 0.0 mm hr-1 is used for stations that do not report an occurrence.

An analysis of instantaneous precipitation rate can be produced separately for each of the MADIS and Clarus surface observation datasets, or jointly in one aggregated analysis. Present weather fields and precipitation fields, from MADIS and Clarus, respectively, are used in order to gather the information to derive an instantaneous liquid water equivalent precipitation rate (Figure 7). Each possible present weather phenomenon reported coincides with an intensity category and a prescribed categorical precipitation rate. All rates within a category are treated as being equally likely, with the value that would produce the smallest average error being used – the middle of a category. Since all rates within a category are treated as being equally likely, a uniform distribution is assumed to describe the precipitation rate field (for any one category of present weather precipitation intensity) A prescribed rate is then used for each surface station report in the Barnes scheme.

Table 5 provides the precipitation rates used in the PPAES surface module and their associated respective precipitation types and intensities. Rates were derived from the American

Meteorological Society Glossary of Meteorology (2000) and Rasmussen et al. (1999). As a result of only using a small subset of possible values for precipitation rates versus a more realistic continuous distribution, quantization errors can occur, and will be discussed later.

The Clarus dataset precipitation rate variable is derived from specific precipitation-related field(s) from each station. The Clarus fields use identical prescribed categorical rates for each possible phenomenon that can be reported. As with the MADIS dataset, the algorithm provides a rate associated with multiple weather phenomena that are reported. In such cases, the predominant precipitation intensity is used for the other weather phenomena.

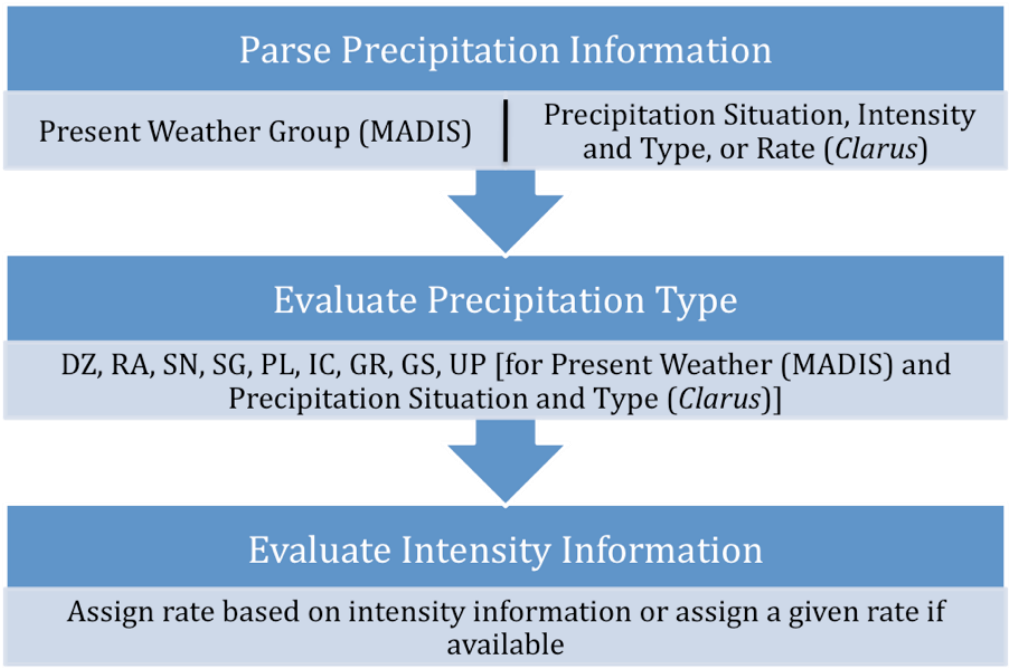

Figure 7. Logic tree for determining precipitation rate information from MADIS and Clarus station reports

When snow is occurring, a revised set of rates (Rasmussen et al 1999) is used that takes into account visibility and temperature (if available and passes the QC criteria). The basis for using a revised set of rates is that the intensity criteria (and their associated rates) for snow are misrepresentative, as they do not measure the actual amount of water in the precipitation (Rasmussen et al. 1999). Three possible sets of rates, depending on the current temperature and visibility, are used for the liquid equivalent snowfall rate (Table 5).

4.5.1 Precipitation Rate Algorithm

The PPAES surface module was designed to incorporate multiple surface datasets if they are available. To accomplish this, numerous variables (e.g., computed Barnes weights, distances from each analysis grid point to the nearest observation station, etc.) are saved in files when each analysis is performed. The Barnes algorithm includes the following steps:

1) Obtain needed parameters, including:
   - A unique Δn (local average observation spacing) for each observation. The Δn is extracted from the local station spatial density value that exists at the grid point nearest to the station location.
   - A smoothing parameter, using the previously extracted local station spatial density.
   - A cutoff distance is computed for each surface observation from the previously computed smoothing parameter. With discrete computation of the smoothing parameter at each observation location, this scheme is adaptive[3].
   - The required amplitude response for the minimally resolved wave is set to the value in Koch et al. (1981; 1983), $D_0(\lambda) = 0.0064$. As noted by Trapp and Doswell (2000), the smoothing parameter is based on a desired theoretical response related to the local station spatial density. The value of 0.0064 was selected by Koch et al. (1983) to give a baseline value for the theoretical minimum resolvable wavelength of 2Δn.
2) All surface observation data are handled station-by-station, one at a time.
3) MADIS stations are compared to the present weather climatology list and are assigned a temporary variable to allow processing to proceed if necessary. Clarus station reports omit this step.
4) The station report is then compared with an updated and accumulating list of stations that have been previously processed, to remove duplicate station reports.
5) Conditional statements are used to determine if the observation is within the prescribed domain and time window.
6) A subroutine is executed to determine if any other reports from the same station exist closer to the analysis' reference time. If so, that station report is used in lieu of the original report and the station identifier is added to the processed station list.
7) A circular radius of influence is identified to determine which analysis values may be influenced by the station's instantaneous precipitation rate. The circular area is centered on the observation location and its size determined by the previously computed cutoff distance. Outside this cutoff distance, the observation has zero weight.
8) An instantaneous precipitation rate is obtained from the list of categorical intensity-based precipitation rates from the observation's present weather report (MADIS) or observed phenomena report/rate (Clarus).
9) As with the surface precipitation occurrence/non-occurrence analysis, the distance between an analysis location and the observation is computed to determine which analysis locations reside in the radius of influence.
10) Analysis values are computed using a single iteration of the Barnes scheme.
11) For each computed analysis value the number of observations used to estimate that value is recorded.

[3] Previous efforts investigating adaptive objective analysis schemes, which adapt to the local station density, have noted the caveat that the detail resolved in richer data density areas will not be resolved in poor data density areas (Askelson et al. 2000). This is important to keep in mind when analyzing the PPAES produced analyses.

Table 5. Precipitation types, intensities, and corresponding rates used for MADIS and Clarus reports in PPAES surface module

Precipitation Type	Intensity	PPAES Rate (mm hr^{-1})	Precipitation Type	Intensity	PPAES Rate (mm hr^{-1})
Drizzle (DZ)	Light	0.15	Snow Grains (SG)	Light	0.15
	Moderate	0.40		Moderate	0.40
	Heavy	0.60		Heavy	0.60
Rain (RA)	Light	1.25	Ice Pellets (PL)	Light	1.25
	Moderate	5.10		Moderate	5.10
	Heavy	10.10		Heavy	10.10
Ice Crystals (IC)	Single Intensity	0.08	Hail (GR)	Single Intensity	7.74
Unknown Precipitation (UP)	Single Intensity	1.75	Small Hail and/or Snow Pellets (GS)	Single Intensity	1.26

Precipitation Type	Intensity	PPAES Rate (mm hr^{-1})	Visibiliy Criteria for Dry Snow[1] (in statute miles)	Visibiliy Criteria for Wet Snow[2] (in statute miles)
Snow (SN)	Light	0.50	Visibility > 0.875	Visibility > 1.125
	Moderate	1.75	Visibility ≤ 0.875 and > 0.375	Visibility ≤ 1.125 and > 0.625
	Heavy	3.25	Visibility ≤ 0.375	Visibility ≤ 0.625

[1] *Dry snow is defined as less than -1 °C from Rasmussen et al. (1999)*

[2] *Dry snow is defined as greater than or equal to -1 °C from Rasmussen et al. (1999)*

After the surface dataset has been completely ingested, a consistency subroutine is executed. The purpose of this routine is to use the previously completed surface precipitation occurrence/non-occurrence analysis and set instantaneous precipitation rates to zero where precipitation was not analyzed to be occurring.

4.6 Complex Terrain Algorithm for Radar Analysis

The quality of radar-based analyses can decrease significantly in complex terrain environments. This generally results from two causes: (1) increased clutter from ground returns and (2) blockage of radar beams that prevents data acquisition. While the original version of the PPAES radar module included a precipitation identification algorithm that attempts to delineate precipitation echoes from non-precipitation echoes (Kessinger et al. 2005), such algorithms are imperfect. Moreover, the original version of the PPAES radar module only utilized data from the lowest elevation angle. By not utilizing data from higher elevation angles, this version of the PPAES radar module is expected to perform poorly when significant beam blockage occurs at the lowest elevation angle. For these reasons, an improved version of the PPAES radar algorithm that accounts for beam altitude relative to the terrain was developed.

This algorithm utilizes a Digital Elevation Model (DEM), the location of the radar, radar-relative data locations, and coordinate transformation equations to determine AGL heights of radar data. The first step in this process was the transformation of DEM data from a set of points that corresponds to a commonly-used map projection to radar-relative locations. A program that produces MSL elevations as a function of radar azimuth and great circle distance from a radar was developed. An azimuth/great-circle-distance reference system is natural one for this type of problem that enables rapid retrieval of ground elevations that are relevant any particular ray of radar data. The use of great-circle-distance has the additional advantage of enabling easy use of more sophisticated ray-tracing programs that provide more accurate estimations of radar data locations under non-standard propagation conditions. While such ray-tracing approaches were not utilized herein, the code structure makes their incorporation relatively simple.

Once ground elevation data were available in a convenient format, the PPAES radar module was modified such that the user can require terrain clearance, if desired. It is noted that this is a common theme in PPAES in that users can easily alter PPAES module behavior by changing options within text files. If terrain clearance is requested, only radar data that clear the terrain by a user-specified amount, which herein is 0.1 km for the bottom of the radar beam, are used in the analysis. This has the advantage of cleaning-up some of the residual ground clutter that is commonly present in PPAES analyses very near each radar, which is present even when the precipitation identification option is turned on because of imperfections in such algorithms, and of enabling the production of analyses values over areas that for which radar beams are blocked at the lowest elevation angle.

It is noted that these modifications, while requiring significant changes to the code, are not expected to fundamentally alter radar-based analyses. Rather, they are expected to produce relatively small-scale changes in the analyses. This is because WSR-88D radars were sighted carefully so as to minimize beam blockage, even at the lowest elevations.

4.7 Fusion of Radar Measurements with Surface Observations

A key objective of this research was designing a way to blend PPAES radar analyses with the PPAES surface analyses. The method desired was one that would leverage the strengths of the both data sources while overcoming a fundamental weakness of radar data --- beam overshoot of shallow precipitation.

The process that was developed involves three steps prior to the actual blending of the two analyses: (1) deriving a radar effective range from the lowest available radar scan, (2) modifying the effective range such that it varies to account for overshooting, and (3) adjusting the radar and surface analysis values near the radar. A description of each of these steps is provided below; the algorithm utilizes a multi-tiered logic tree.

4.7.1 Define Radar Effective Range

First, the radar-derived effective range is determined by finding the farthest range where a contiguous echo has a predefined size (three range gates and five azimuths) that exceeds a reflectivity threshold (0 dBZ). The purpose of this is to determine the maximum range at which each radar can detect meteorological echoes. The reflectivity in the defined echo must be greater than 0 dBZ to delineate between atmospheric and non-atmospheric phenomena, though on rare occasions, light snow and freezing drizzle do produce reflectivities below this detection threshold due to the small hydrometeor sizes involved (Ikeda and Rasmussen 2003). The prescribed echo size was chosen because a continuous echo was desired to reject spurious reflectivity values.

4.7.2 Modify Radar Effective Range

After a defined effective range is determined for each radar, those effective ranges are then potentially modified based upon tests using surface observations. For this study, two modification methods were examined. The first method modified effective ranges according to radar quadrants while the second method modified effective ranges continuously in the azimuthal direction. These methods are described in more detail later in this section. The new modified ranges are utilized because even if an echo seems to be present at greater distances, it is still possible for overshooting to occur at lesser distances. These modified effective ranges are used in the blended algorithm as they indicate at what ranges one should transition from analysis values based on radar data to analysis values based on surface observations. The modified effective ranges are computed through the following steps:

1) For each surface observation that indicates the occurrence of precipitation, the distance and bearing (with respect to each radar) from that observation to each of the individual radars used in the analysis is computed. If this distance is less than the original effective range or the previously computed modified effective range, that surface observation is compared to the radar analysis value that is present at that location.
2) If the radar is indicating no precipitation, this distance becomes the effective range.
3) The new effective range is then stored for use in the analysis.

4.7.2.1 Quadrant Effective Range Method

The range and bearing to the radar are calculated for each station. That bearing was used to classify which quadrant (0°-90°, 90°-180°, 180°-270°, and 270°-360°) the station lies in for a select radar. The effective range is then modified according to the logic described earlier in this section and modified for all azimuths within the calculated quadrant.

4.7.2.2 Azimuthal Effective Range Method

It was determined that a better method for modifying effective ranges may be interpolating between azimuths based upon the number of stations within the scan range of the select radar such that the effective range varies continuously in the azimuthal direction. This method was examined in an effort to create a more realistic modified effective range field that can be used in the analysis to better capture the effects of different precipitation scenarios on radar electro-magnetic wave propagation characteristics.

The range and bearing are calculated between a radar and each station. If the logic (described above in this section) determines that the effective range needs to be modified, then the effective range is only modified at this azimuth and the azimuth is recorded as having a station that modified its effective range.

The algorithm then loops through all 360 degrees of the selected radar to find any other azimuths that have recorded stations that modified their effective ranges. If there is only one station identified for that radar, then the effective range for all 360 degrees is set to the effective range of that one station.

If multiple stations exist for a radar, then a linear interpolation method is used to modify the effective range at each azimuth between azimuths that were directly modified by the presence of a surface station.

4.7.3 Local Adjustment of Radar and Surface Analysis

The final step is a local (defined as inside the respective radar site's effective range) adjustment of the radar and surface analyses. This local adjustment:

1) Removes some error associated with quantization that is inherent to surface observation precipitation rate values determined using categorical present weather information.
2) Removes discrete jumps when transitioning from regions based upon radar data to those based upon surface observations.

The local adjustment proceeds as follows. For each individual radar envelope, all of the co-located surface-radar observation pairs with precipitation (e.g., Klazura et al. 1999) are identified inside the modified effective range. For each co-located pair, the difference between the surface observation-based analysis value and the spatial average of the radar analysis values centered on that location is computed. A spatial average of radar analysis values is used: (a) to avoid the situation where no precipitation is observed at the radar analysis location but is seen at the

surrounding locations (indicating that precipitation there is likely); (b) to account for possible errors due to horizontal displacement of radar-measured snowfall as radar measurements are collected aloft, above the surface observation; and (c) to account for uncertainty regarding which radar range gate contributed to the surface instrument's precipitation measurement (Fulton et al. 1998; Fulton 1999). A 3 x 3 bin box centered on the analysis location is used for this procedure, following Fulton et al. (1998).

Next, a quality control step is implemented in that a set number of co-located pairs must exist to compute and apply an adjustment value. PPAES uses a threshold of six co-located pairs (Fulton et al. 1998; Fulton et al. 1999). A set threshold is necessary to ensure a reasonable sample size when computing the adjustment value (Klazura et al. 1999). Too few stations would result in a potentially unrepresentative adjustment, whereas the larger the sample – the more representative the adjustment should be (Fulton et al. 1998,1999; Klazura et al. 1999). When the threshold condition is satisfied, an adjustment is applied uniformly to the radar-based analysis values [similar to the "mean field" bias adjustment and its application in Fulton et al. (1998)] and to the neighboring surface analysis values. For each individual radar, the adjustment value is computed by averaging the co-located mean differences and dividing the sum by two to remove the total mean amount of difference between the two sets of analyses (radar and surface). When for instance, the adjustment value for a radar is positive, that value is then added to radar analysis values and, similarly, subtracted from surface analysis values.

4.7.4 Blending Algorithm

The fusion of the two analyses involves a multi-tiered logic tree utilizing several analysis fields and variables, including: (a) modified effective ranges for each radar, (b) mean co-located differences for each individual radar, (c) the radar ID analysis field (containing IDs of radars used in the analysis), (d) the radar instantaneous precipitation rates, (e) the surface precipitation occurrence/non-occurrence field, (f) the surface instantaneous precipitation rates, and (g) specific PPAES execution control variables (cf. Section 2).

First, a transition distance, d_t, is used to blend the radar- and surface-based analyses linearly over a distance. Currently, d_t is set to 100 km. Since overshooting is not instantaneous when it occurs, overshooting is proceeded by range degradation or the reduction of reflectivity with increasing altitude (i.e., radar underestimation), which occurs up to where the radar beam completely overshoots the precipitating cloud (Fulton et al. 1998; Super and Holroyd 1998). A 100 km transition distance was decided upon because a large d_t will result in a loss of spatial resolution in the radar-based analysis regions. With an overly small d_t, non-physical gradients can develop owing to radar-based analyses transitioning too rapidly to regions where analysis values come from surface observations. Test cases were run to see the effects of different d_t*'s*, specifically 50 km and 100 km (not shown). Based upon these test cases, a final distance of 100 km was chosen for the work presented here.

The blending algorithm proceeds according to the following steps:

1) Before execution, the radar analysis is checked to determine if the analysis falls within the time window set for blending with the surface analysis. If more than one analysis falls within

the window selected, as can be the case when PPAES is run after the fact for historical cases, the radar analysis nearest to the date/time of the surface analysis is selected.

2) The previously mentioned analysis fields and variables are passed or computed.
3) Each analysis grid point is checked to determine if both a radar and a surface analysis value are present, only one analysis value is present, or neither is present. The logic tree splits depending upon the result of this check. Each analysis grid point is then assigned an analysis value depending upon which information is available for that analysis grid point.
4) If both radar and surface analysis information is available at the analysis grid point, the distance and bearing between the analysis grid point to the radar from which the present radar analysis value comes from is computed. The modified effective range along the azimuth in which the analysis grid point lies, relative to the corresponding radar, is then obtained. Depending upon the distance from the analysis grid point to the radar, numerous paths are possible. The optimal situation occurs when enough co-located positive radar-surface pairs are available to compute an adjustment for the radar and nearby surface analysis values (Figure 8). The following decision tree, which corresponds to the optimal situation, then applies:
    a) If the computed distance is less than or equal to the modified effective range minus d_t, the adjusted radar value (if applicable) is assigned to the blended analysis value at that analysis grid point.
    b) If the analysis value's distance is greater than the modified effective range minus d_t and is less than or equal to the modified effective range (i.e., it is within the transition zone), the blended analysis value is assigned a linearly-weighted value that takes into account the adjusted radar and surface analysis values at that analysis grid point. A linear weighting scheme is used such that at the effective range minus d_t, the analysis value is given by the adjusted radar-based value; at the effective modified range, the analysis value is given by the adjusted surface-based value.
    c) If the analysis value's distance is greater than the modified effective range and less than or equal to the modified effective range plus d_t, the blended analysis value is assigned a linearly-weighted value. This linearly-weighted value is derived using the adjusted surface analysis value and the un-adjusted surface analysis values at that analysis grid point. At the effective modified range and the effective modified range plus d_t, the analysis values are given the adjusted and non-adjusted surface analysis values, respectively.
    d) If the analysis value's distance is greater than the modified effective range plus d_t, the blended analysis value is assigned the non-adjusted surface value at that analysis grid point. The non-adjusted surface value is assigned because at distances far away from the radar, an adjustment based upon data collected within a radar envelope is not necessarily applicable.
    e) When no modified effective range can be computed, the blended analysis value is assigned a missing value. This step is a post-processing adjustment and removes precipitation rates that may have resulted from anomalous propagation, ground clutter, or non-atmospheric phenomena (Fulton et al. 1998).

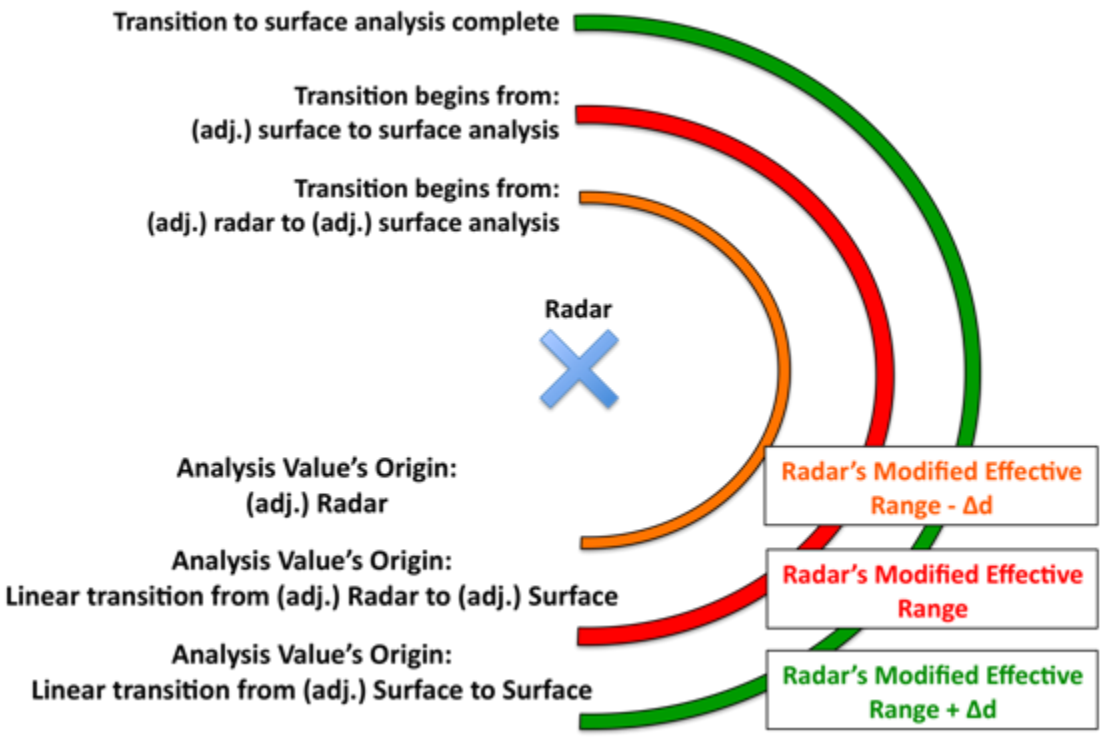

Figure 8. Optimal course of action when both surface and radar analysis values are present on analysis grid

5) When only radar-based analysis information is available at an analysis grid point, the distance and bearing between the analysis grid point and the radar from which the analysis value comes from are computed. The modified effective range along the azimuth in which the analysis grid point lies, relative to the corresponding radar, is then obtained. Depending upon the distance from the analysis grid to this radar, numerous paths are possible:
   a) If the computed distance is less than or equal to the modified effective range minus d_t, the blended analysis value is assigned the adjusted radar value at that analysis grid point. Similarly, if the analysis value's distance is greater than the modified effective range minus d_t and less than or equal to the modified effective range, the blended analysis value is assigned the adjusted radar value linearly blended with the surface precipitation rate (which is 0 mm hr^{-1} in this situation).
   b) When no modified effective range can be computed, the blended analysis value is assigned 0 mm hr^{-1}. This step serves as a post-processing adjustment for the blending algorithm and removes precipitation rates that resulted from anomalous propagation, ground clutter, or non-atmospheric phenomena (Fulton et al. 1998).
   c) If the analysis value's distance is greater than the modified effective range of the radar, 0 mm hr^{-1} is assigned.
6) If only a surface analysis precipitation value is available at the analysis grid point, the distance and bearing between the analysis grid point and the nearest radar is computed. Next, the nearest radar's modified effective range along the azimuth in which the analysis grid point lies, relative to the corresponding radar, is obtained. Depending upon this distance, several paths are possible:

a) If the computed distance of the analysis value at the analysis grid point is less than or equal to the modified effective range minus d_t, the blended analysis value is assigned 0 mm hr^{-1} at that analysis grid point.
b) If the computed distance of the analysis value at the analysis grid point is greater than the modified effective range minus d_t and less than the modified effective range, the blended analysis value is assigned the adjusted surface analysis-based value linearly blended with the radar-based precipitation rate (0 mm hr^{-1}).
c) If the distance of the analysis value at the analysis grid point is computed and it is greater than the modified effective range and less than the modified effective range plus d_t, the blended analysis value is assigned a linearly-weighted value. This value takes into account the adjusted surface-based analysis value and the un-adjusted surface-based analysis value at that analysis grid point. The linear-weighted value is specified such that at the effective modified range and the effective modified range plus d_t, the analysis values assigned are the adjusted and non-adjusted surface-based analysis values, respectively.
d) If the computed distance of the analysis value at the analysis grid point is greater than the modified effective range plus d_t or a modified effective range does not exist, the un-adjusted surface-based analysis value is assigned to the analysis grid point.

7) Lastly, when no precipitation information (radar or surface) is present at an analysis grid point, 0 mm hr^{-1} is assigned to that analysis grid point.

4.8 Fusion of Radar and Model Data

Relative to the fusion of radar and surface data, the approach to fusing radar and model data is pretty simple. As previously indicated, RAP data provide estimates of 1 and 4 km AGL radar reflectivity factor and geopotential heights at the ground. The first step in the fusion process is estimation of these fields on the PPAES output grid. This is accomplished through sampling—the location of the center of each PPAES analysis grid point is determined in the RAP grid. The value of the RAP field is then the model value in the PPAES analysis grid.

Because the model data are considered to be modifying the base radar analysis, the next step is the initialization of the analysis field using radar-based analysis values. After this, the more complicated step in the fusion process starts. This step is completely user controlled. First, if the user requested this behavior, at locations where no radar values are available but model values are, model values are used. Next, if both radar and model values are available at the same location:

- If the user requested this behavior (virga correction), if the MSL height of the model reflectivity value is lower than the MSL height of the radar data from which the radar-based value came by a user-specified amount, and if the radar indicates that precipitation is occurring while the model indicates that no precipitation is occurring, then the model value replaces the radar-based value.
- If the user requested this behavior (precipitation filling), if the MSL height of the model reflectivity value is lower than the MSL height of the radar data from which the radar-based value came by a user-specified amount, and if the radar indicates that no precipitation is occurring while the model indicates that precipitation is occurring, then the model value

replaces the radar-based value.

It is noted that at this time, if both radar and model data indicate that precipitation is occurring at a point and if the model data come from a lower altitude, nothing is done (the radar-based analysis value is not replaced with the model-based value). In the tests conducted herein, model data were used when no radar data were available; virga correction was applied when the radar and model data indicated it was occurring and radar data altitudes were greater than model data altitudes by at least 0.3 km, and precipitation filling was applied when radar and model data indicated it was appropriate to do so and radar data altitudes were greater than model data altitudes by at least 0.3 km.

The current algorithm uses simple replacement and does not attempt to blend between areas dominated by radar values and areas dominated by model values. This algorithm is at an early stage of development and because methods like those used to smoothly transition between radar- and surface-based analyses are not employed, abrupt transitions between radar- and model-based regions are expected. Furthermore, these are expected to be exacerbated by the significant difference in the resolutions of the two data sources.

4.9 Evaluation and Validation of PPAES-Produced Products

To fully test PPAES, three domains were chosen to capture regional characteristics that may affect PPAES results. These domains are classified as the Northern Great Plains (NGP), Midwest (MW), and the Utah region (UT) and are shown in Figure 9. The NGP and MW domains were chosen due to their relatively flat terrain and because of the Interstate 94 (I-94) and Interstate 80 (I-80) corridors that transit through each respective domain. The Utah region was chosen to specifically test PPAES in areas of complex terrain.

Several case studies were identified for each region and are shown in Table 6. For the initial evaluation and validation, 19 cases were examined where only MADIS METAR datasets were utilized (Clarus data were not available for these events). Nine additional cases for which both MADIS and RWIS data were available were examined using the MW domain and ten such cases were examined using the NGP domain. For the Utah region, five cases were examined to study the complex terrain problem as well as the inclusion of Clarus data within the surface analysis.

Separate approaches are used to verify PPAES precipitation occurrence/non-occurrence and instantaneous precipitation rate as dictated by the following types of analyses: (1) use of only the surface module with MADIS only data; (2) use of only the surface module with MADIS and Clarus data; 3) use of only the radar module (with or without terrain clearance), the model module, or the radar+model module; and (4) use of the radar+surface module. A primary goal was to determine how the PPAES platform-specific modules perform relative to the radar+surface blending module. This is needed to assess the value that is added by blending the different types of analyses together.

The observation element that is used to verify occurrence and instantaneous precipitation rate is present weather. For both single (surface, radar, and model) and multi-sensor (blended) module

validation, surface MADIS ASOS/AWOS observations are assumed to be "truth" for precipitation occurrence and instantaneous precipitation rate. The rationale for this choice is that surface observation errors are fairly well understood and at any one location, they are expected to be the most accurate representation of sensible weather. To ensure independence of the evaluation and validation of the PPAES products, a data denial scheme is utilized whenever MADIS ASOS/AWOS observations are used an input in an analysis. Thus, data denial was used for (1), (2), and (4), whereas no data denial was required for (3).

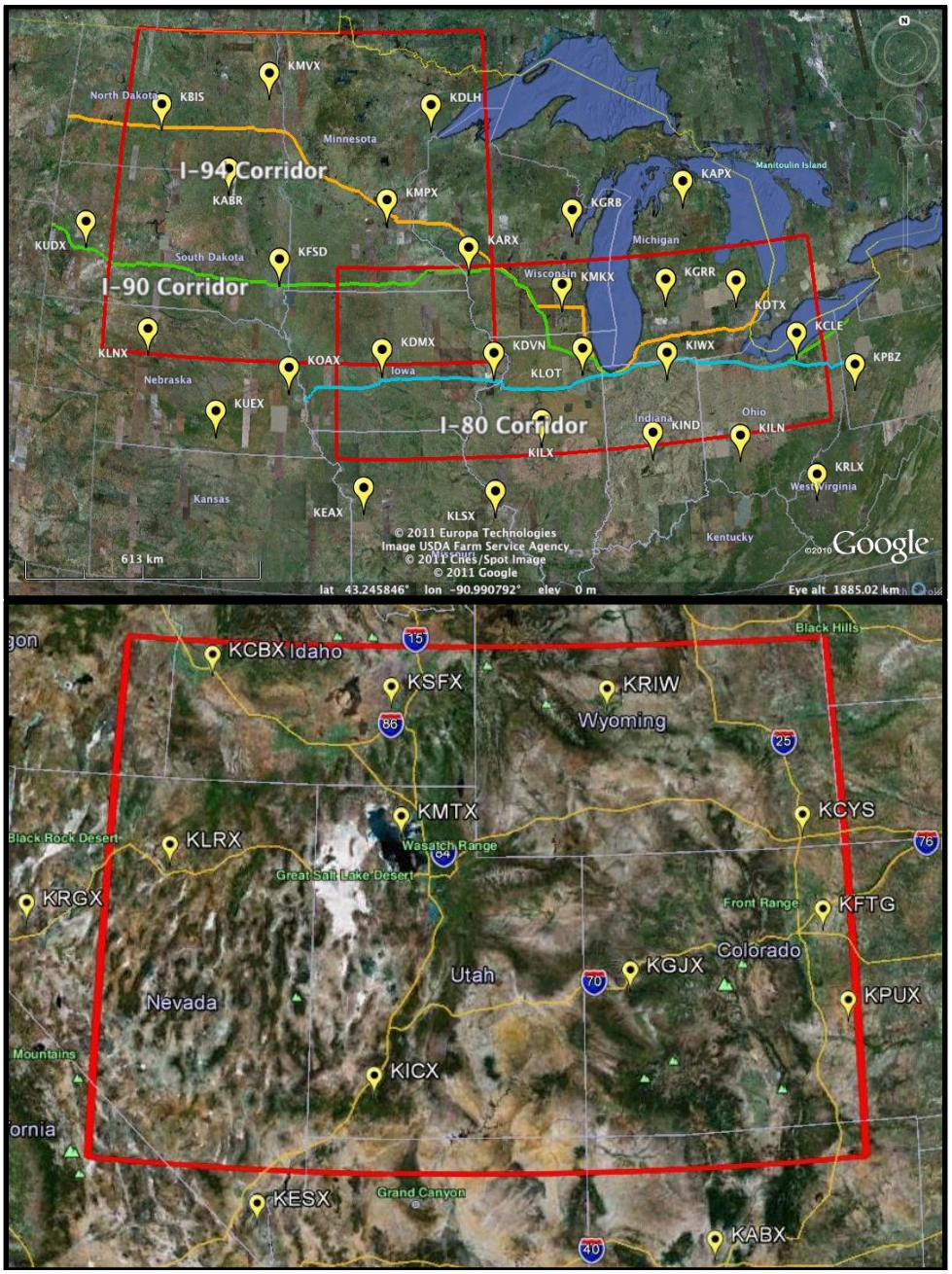

Figure 9. Two of the evaluation/verification domains (outlined in red) (top) and Utah evaluation domain (bottom)

In the top map, the three Interstate highways are denoted and the individual NEXRAD WSR-88D radar sites used in this study are the yellow markers. I-94 is orange, I-90 is green, and I-80 is blue.

Table 6. Historical events spanning from January 2006 to January 2013 used for PPAES testing

Initial Testing of MW & NGP Regions using MADIS (only) Data			
ID	Case Date	ID	Case Date
1	1/17/2006 - 1/18/2006	11	2/28/2007 - 3/2/2007
2	12/30/2006 - 12/31/2006	12	3/15/2007 - 3/16/2007
3	1/4/2007 - 1/5/2007	13	11/26/2007 - 11/27/2007
4	1/28/2007 - 1/29/2007	14	12/1/2007 - 12/2/2007
5	1/31/2007	15	12/16/2007
6	2/1/2007	16	2/11/2008 - 2/12/2008
7	2/2/2007	17	2/13/2008 - 2/16/2008
8	2/3/2007	18	3/4/2008 - 3/5/2008
9	2/6/2007	19	3/21/2008 - 3/23/2008
10	2/24/2007 - 2/27/2007		
Testing of MW Region using MADIS & Clarus Data			
ID	Case Date	ID	Case Date
1	12/8/2009 - 12/10/2009	6	1/14/2011 - 1/15/2011
2	2/1/2010 - 2/2/2010	7	12/8/2011 - 12/9/2011
3	2/7/2010 - 2/10/2010	8	2/28/2012 - 2/29/2012
4	12/10/2010 - 12/13/2010	9	12/8/2012 - 12/9/2012
5	12/23/2010 - 12/25/2010		
Testing of NGP Region using MADIS & Clarus Data			
ID	Case Date	ID	Case Date
1	12/8/2009- 12/9/2009	6	1/14/2011
2	2/1/2010 - 2/2/2010	7	3/22/2011 - 3/23/2011
3	2/7/2010 - 2/9/20107	8	11/19/2011 - 11/20/2011
4	12/10/2010 - 12/12/2010	9	2/28/2012 - 2/29/2012
5	12/23/2010 - 12/24/2010	10	12/8/2012 - 12/9/2012
Testing of Utah Region using MADIS, Clarus, Radar, and Model Data			
ID	Case Date	ID	Case Date
1	11/20/2010 - 11/23/2010	4	12/26/2012 - 12/27/2012
2	11/27/2010 - 11/29/2010	5	1/26/2013 - 1/28/2013
3	12/24/2012		

Most data denial experiments involve withholding an entire set of observations to: (a) determine if either a particular observation type or subset has an impact on an operational numerical model predication system or (b) evaluate the usefulness of different observation types (e.g., Arnold and Dey 1986; Atlas 1997; Benjamin et al. 2004; Benjamin et al. 2010). Thus, the motivation seen in

the literature for using a data denial scheme is different than within this study. For this study, data denial is used as a means of independent verification. As such, selecting how many stations to deny is a difficult issue in order for PPAES analyses to not be adversely impacted by denial of too many stations.

We have chosen to withhold ~10% of the stations used in the analysis in our data denial scheme. Since the Midwest and Northern Great Plains domains contain a similar number of METAR stations, the same fixed number of stations (15) is withheld for both. For the Utah region, we decided to withhold 10 stations for verification as its area was smaller than those of the other two domains. Potential sensitivity to this value will be addressed in the Discussion section.

In order to not skew the statistics obtained from the verification suites, a fixed set of stations is withheld over the duration of each case. A quasi-random approach is used to identify which stations are withheld. For the NGP and MW regions, seven stations are randomly chosen and eight stations are subjectively chosen. The eight subjectively chosen stations are split such that four are in precipitation areas and four are in non-precipitation areas. This is done to ensure that a portion of the stations will be placed in areas that will receive precipitation to produce meaningful statistics. For the Utah region, 6 stations were randomly picked and 4 stations were subjectively picked..

Contingency-table-based performance metrics are used to verify precipitation occurrence/non-occurrence. For instantaneous precipitation rate validation, statistics employed commonly in the atmospheric sciences (mean error, mean absolute error, mean squared error, root mean square error, multiplicative bias, and correlation) are used.

Because the individual cases have different event durations, to ensure that results for the cases listed in Table 6 can be compared with each other, results for each case are averaged. The average of each statistic is computed over the duration of each individual event.

4.9.1 Precipitation Occurrence/Non-occurrence Analysis Validation

A contingency table for precipitation occurrence is used to compute the performance metrics that are shown in Table 7.

Table 7. Contingency table used in this study (CAWCR 2011)

	Observed		
Forecast	Yes	No	Total
Yes	*hits*	*false alarms*	*forecast yes*
No	*misses*	*correct negatives*	*forecast no*
Total	*observed yes*	*observed no*	*total*

The performance metrics (Jolliffe and Stephenson 2004; CAWCR 2011) include:

Bias score: BIAS = (hits + false alarms)/(hits + misses). (12)

Probability of detection: POD = hits/(hits + misses). (13)

False alarm ratio: FAR = false alarms/(hits + false alarms). (14)

Odds ratio OR = (hits * correct negatives)/(misses * false alarms). (15)

Probability of false detection (POFD):

POFD = false alarms/(correct negatives + false alarms). (16)

Equitable Threat score (ETS):

$$ETS = (hits - hits_{random})/(hits + misses + false\ alarms - hits_{random}), \quad (17)$$

where

$$hits_{random} = (hits + misses)(hits + false\ alarms)/total. \quad (18)$$

Threat score TS = (hits)/(hits + misses + false alarms). (19)

Accuracy = (hits + correct negatives)/total. (20)

4.9.2 Instantaneous Precipitation Rate Analysis Validation

Validation of instantaneous precipitation rate involves the use of summary scores (Jolliffe and Stephenson 2004; CAWCR 2011). In addition to the computed performance metrics, scatter plots are created to visualize performance metrics' event-mean values. Summary scores computed, using the observation values (*O*), forecast values (*F*), and the total number of observation and forecast values (*N*), are:

$$\text{Mean error} = \frac{1}{N}\sum_{i=1}^{N}(F_i - O_i) \quad (21)$$

$$\text{Mean absolute error MAE} = \frac{1}{N}\sum_{i=1}^{N}|F_i - O_i| \quad (22)$$

$$\text{Mean squared error MSE} = \frac{1}{N}\sum_{i=1}^{N}(F_i - O_i)^2 \tag{23}$$

$$\text{Root mean square error RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(F_i - O_i)^2} \tag{24}$$

$$\text{Multiplicative bias BIAS} = \frac{\frac{1}{N}\sum_{i=1}^{N}F_i}{\frac{1}{N}\sum_{i=1}^{N}O_i} \tag{25}$$

$$\text{Correlation coefficient } r = \frac{\sum(F-\overline{F})(O-\overline{O})}{\sqrt{\sum(F-\overline{F})^2}\sqrt{\sum(O-\overline{O})^2}} \tag{26}$$

4.9.3 Significance Testing

Significance testing was used in the initial set of verification computations to compare the population (i.e., the PPAES surface, radar, and blending modules) means of performance metrics based on the set of historical cases. By doing so, differences between event-average performance metrics associated with different modules can be evaluated and one can determine if such differences are statistically significant.

Two significance test methods were utilized, the ANOVA (analysis of variances) and the two-sample *t*-test (with equal variances assumed). The performance metrics tested include event-averages of accuracy, POD, FAR, and ETS (case). The POD, FAR, and accuracy performance metrics were selected as they can be used to gain insight into general algorithm performance as well as into behavior of the ETS statistic. ETS (case) was selected because it is a popular composite score and indicates overall performance with regard to precipitation occurrences. The significance level (denoted as α) for the ANOVA and t-tests was set to 0.05 (the 5% significance level) so that the probability of falsely rejecting the null hypothesis (that the population means are the same) is 5% (Wilks 2006).

The one-way ANOVA test compares the means of the three populations—the PPAES radar, surface, and blending modules. The one-way ANOVA compares the variation in the sample data and both the variations among the sample population means and within the samples are tested to determine if the means of the populations are not equal (Weiss 2008). The alternative hypothesis for this method is that not all of the means are equal. The assumptions for one-way ANOVA are that the sample means are random, independent, normally distributed, and that the populations have equal population standard deviations, which are identical to the assumptions made in the two-sample *t*-test (Weiss 2008).

The test statistic for the one-way ANOVA is the *F*-statistic; before this statistic can be obtained, quantitative measures of the variation among the sample means and measures of the variation

within the samples must be computed. The variation among the sample means is denoted as the treatment mean square (*MSTR*) and is given by

$$MSTR = \frac{\sum_{j+1}^{j} n_j (\bar{x}_j - \bar{x})^2}{k-1} \tag{27}$$

where $\bar{x}$ denotes the overall mean from the total number of observations from all samples, *k* denotes the number of populations being sampled, $\bar{x}_j$ denotes the mean of a sample population, and n_j denotes the sample size for one of the invoking samples (Weiss 2008). Next, the error mean square (*MSE*) is computed as a measure of the variation within the samples and is given by

$$MSE = \frac{\sum_{j+1}^{j} (n_j - 1) s_j^2}{n-k} \tag{28}$$

where *n* denotes the total number of observations and s_j^2 denotes the variance of a sample (Weiss 2008). The *F*-statistic is then computed as

$$F = \frac{MSTR}{MSE} \tag{29}$$

With the *F*-statistic, the critical value can then be obtained from a table using the chosen significance level (α = 0.05) or numerically to see if the null hypothesis can be rejected or not.

The test statistic for the two-sample *t*-test with equal variances assumed is the *t*-statistic. The null and alternative hypothesis for this method are one of: (a) the means are equal (null) and not all of the means are equal (two-tailed test) or (b) that the first mean is greater (right tailed) or less than (left tailed) the other (Weiss 2008). Only two populations can be examined at a time when performing this test.

The *t*-test test statistic requires that the pooled sample standard deviation be computed. The pooled sample standard deviation is given by

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} \tag{30}$$

where *n* denotes the sample size of one of the samples being used and s^2 is the variance of one the sample being used. The test statistic is then given by

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{s_p \sqrt{1/n_1 + 1/n_2}} \tag{31}$$

where $\bar{x}$ denotes the mean of an independent sample and *n* denotes the total sample size of an independent sample. The test statistic can then be compared to the critical value at the alpha level ($\alpha = 0.05$) to see if the null hypothesis is rejected or not.

Both of these methods share the same assumptions regarding the populations of the samples. Of these assumptions, the most suspect are the normality and equal population standard deviation assumptions. The former can be affected by outliers and the latter is robust only if the sample sizes are nearly identical (as is the case here) (Weiss 2008). Both the one-way ANOVA and two-sample *t*-test are robust to moderate violations (Weiss 2008). To test the normality and equal population standard deviation assumptions, normal probability plots, box plots, and ratios of the larger to the smaller sample standard deviation ("rule of 2") were constructed (Weiss 2008). If these tests revealed an outlier, they were run with and without the outlier to see if the test statistic was adversely affected from the standpoint of obtaining statistical significance.

Normal probability plots depict the observed values of the variable sorted versus normal scores (the observations that would be expected for a variable that has a standard normal distribution) (Weiss 2008). Given the small sample size (19 cases in the initial set of tests), a normal probability plot is ideal for identifying outliers (Weiss 2008). Interpreting the normal probability plot is simple--if the plot is linear, the data are approximately normally distributed (Weiss 2008). A box plot's size, shape, and whiskers offer another assessment as they show the center of the dataset and the variation within the dataset. Lastly, the ratio of the larger to the smaller standard deviation ("rule of 2") was used as a way to determine whether the samples had approximately equal population standard deviations (Weiss 2008). If the ratio of the larger to the smaller standard deviation is less than two, the condition can be considered met (Weiss 2008).

5 RESULTS

5.1 Initial Validation

PPAES surface, radar, and blending modules were run for the 19 historical cases (Table 8). Event-average values of accuracy, bias, POD, FAR, POFD, TS, ETS, and odds ratio were computed for precipitation occurrence/non-occurrence. Event-average values of mean error, MAE, MSE, RMSE, multiplicative bias, and correlation were computed for precipitation rate. The averages include only analysis hours when all of the PPAES modules (surface, radar, and blended) were run.

The historical cases were categorized according to event type (Table 8), using the University Corporation for Atmospheric Research's (UCAR) Meteorological Case Study Selection Kit (http://www.mmm.ucar.edu/imagearchive/) and previously documented PPAES case logs (M. Askelson 2011, personal communication). The grouping of historical cases was accomplished subjectively using surface analyses, surface observation plots, radar precipitation rate analyses (and reflectivity), composite analyses (satellite, radar, and surface observations), and previous case documentation.

Synoptic features (e.g., cold front, low pressure system, etc.) were also used to classify each event type in the historical dataset. Several cases had different synoptic features present in the domain at the same time (e.g., a cold front to the west and a surface trough to the east) or an additional feature that occurred later in the event. For these cases, the main synoptic feature (defined as the primary synoptic feature that drove the precipitation in the verification domain) is used (always listed as the first characteristic) and other features are listed in order of precedence.

The three main synoptic features used to stratify the cases included (1) surface low pressure systems, (2) frontal boundaries, and (3) surface troughs. Surface troughs were defined as elongated areas of low pressure that pass over the verification domain, but do not exhibit frontal characteristics. Surface fronts were identified as the transition region between two air masses that exhibit a temperature gradient and strong wind shifts in a narrow elongated zone (Bluestein 1993; Glickman 2000). Low pressure systems were identified as centers of areas exhibiting cyclonic circulation and pressure minima on surface analyses (Bluestein 1993; Glickman 2000). Since fronts usually have some type of low pressure system associated with them, a distinction is made when a surface front passes through the domain but not necessarily the low pressure system itself – this type would be classified as frontal. A distinction is also made between fronts and low pressure systems because different processes often drive the rising motion that is associated with them. An in-depth classification and analysis using mesoscale characteristics is beyond the scope of the current study.

Precipitation intensity was also identified for each case (Table 8) and was defined as the intensity that occurred most frequently and was most widespread spatially. The predominant precipitation intensity was identified using PPAES case logs, surface maps, METARs, and radar reflectivity factor and precipitation rate analyses. Precipitation intensities were categorized as light (0.0 – 1.0 mm hr^{-1}), moderate (1.0 – 2.5 mm hr^{-1}), and heavy (2.5 mm hr^{-1} and greater) (Rasmussen et al.

1999). As previously mentioned, the Z_e-S relationship used in PPAES is $Z_e = \alpha S^{\beta}$, where α and β are constants and are set to 150.0 and 2.0, respectively. The above intensity interval endpoints of 0.0, 1.0, and 2.5 mm hr^{-1} correspond to approximately -18, 22, and 30 dBZ.

Included are the date(s) of the events, event type, predominant precipitation intensity, and any relevant comments. The event type denotes the type of synoptic feature that drove the precipitation in the verification domains.

Table 8. List of initial historical cases used in this study

ID	Case Dates	Verification Domain[1]	Event Type	Predominant Intensity	Comments
1	1/16-1/19/2006	NGP	Surface Trough	Light	Persistent snow across the Red River Valley
		MW	Low Pressure System	Moderate	Large frontal passage affecting Indiana and eastward
2	12/29/2006-1/1/2007	NGP	Frontal	Moderate	Large areal coverage mixed precipitation event
		MW	Frontal	Moderate	Large areal coverage rain event
3	1/4-1/6/2007	NGP	Frontal	Moderate	Mostly rain/freezing rain precipitation event
		MW	Frontal	Moderate	Rain event with widespread precipitation
4	1/27-1/30/2007	NGP	Frontal	Light	Mixed precipitation event
		MW	Surface Trough/Frontal	Light	Light snow, mostly localized to the lower Great Lakes and southward
5	1/31/2007	NGP	Low Pressure System	Light	Alberta clipper-type event
		MW	High Pressure System/Frontal	Light	Event begins with lake effect snow in eastern half of domain before a subsequent front
6	2/1/2007	NGP	Frontal/Low Pressure System	Light	Alberta clipper-type event
		MW	Frontal	Light	Cold front
7	2/2/2007	NGP	Frontal	Light	Very weak precipitation event involving a secondary cold front (with radar overshoot)
		MW	Frontal	Light	Cold front with limited widespread precipitation
8	2/3/2007	MW	Frontal	Moderate	Cold front with limited widespread precipitation
9	2/6-2/7/2007	NGP	Frontal/Surface Trough	Heavy	Warm front event with heavy snow at times
		MW	Surface Trough/Frontal	Moderate	Pre-frontal, ahead of a warm front
10	2/23-2/27/2007	NGP	Surface Trough	Heavy	Long lasting storm that brought significant snow accumulations in the Red River Valley
		MW	Frontal/Low Pressure System	Heavy	Mixed event with frontal passage, varying precipitation intensity
11	2/28-3/4/2007	NGP	Surface Trough	Heavy	Winter storm that dropped snow amounts of 1 ft total in the Red River Valley
		MW	Frontal	Heavy	Event starts as rain and then transitions to snow
12	3/13-3/16/2007	NGP	Low Pressure System/Frontal	Moderate	Small-scale Alberta clipper-type event
		MW	Frontal	Moderate	Cold front with widespread rain, some mixed precipitation at the end of the event
13	11/25-11/27/2007	NGP	Low Pressure System/Frontal	Moderate	Alberta clipper-type event that resulted in reduced visibility at times (near zero miles)
		MW	Low Pressure System/Frontal	Moderate	Widespread precipitation with some mixed/snow precipitation
14	12/1-12/3/2007	NGP	Low Pressure System/Surface Trough	Heavy	Colorado low pressure system that brought large snowfall amounts to Minnesota
		MW	Frontal	Heavy	Warm front event that started as snow and then transitioned to widespread rain
15	12/15-12/16/2007	NGP	Frontal	Light	Very light and shallow snow event, notably weak radar reflectivity
		MW	Low Pressure System	Moderate	Widespread snow event with surface low passing to the southeast
16	2/10-2/12/2008	NGP	Surface Trough/Frontal	Light	Light event involving a surface trough and followed by a warm front
		MW	Surface Trough	Light	Surface trough, predominantly light event that had moderate intensity at times
17	2/13-2/15/2008	NGP	Low Pressure System/Frontal	Moderate	Alberta-clipper type snow event
		MW	Frontal	Light	Remnants of previous event then a subsquent frontal passage
18	3/2-3/6/2008	NGP	Low Pressure System	Light	Alberta-clipper system
		MW	Low Pressure System/Frontal	Moderate	Surface low with widespread precipitation, mostly a snow event
19	3/20-3/23/2008	NGP	Surface Trough/Low Pressure System	Light	Slow moving Low pressure system moving south of the domain
		MW	Low Pressure System	Moderate	Predominately a snow event with some mixed precipitation

[1] NGP and MW denotes the Northern Great Plains and Midwest verification domains, respectively.

5.1.1 Northern Great Plains Domain

5.1.1.1 Historical Case Validation

Eighteen cases were used to evaluate PPAES precipitation occurrence/non-occurrence and rate in the Northern Great Plains domain (Table 8). These 18 cases include five surface trough events, seven frontal boundary events, and six low pressure system events (Table 8). In addition, nine of these cases were characterized predominantly by light precipitation intensities, five cases were characterized predominantly by moderate precipitation intensities, and four cases were characterized by heavy precipitation intensities. Event-averages for each performance metric were computed over the duration of each case (Table 9).

Values of the performance metrics previously discussed in the Methodology section were computed for each analysis. Scatter-plots of event-average values of mean error, RMSE, accuracy, FAR, POFD, POD, and mean/case-computed ETS are provided in Figures 10-16. Note that two different methods of computing ETSs are used. The first method uses ETS values computed for each analysis and then averaging the resulting values over the whole event. Method two uses the summed totals of the contingency table (hits, false alarms, misses, correct negatives) entries over the duration of each case and uses the totals to compute a case ETS value. The former is denoted as "mean" and the latter is denoted as "case" in Figure 16 to distinguish between these two different methods.

Several trends are apparent in Table 9 and Figures 10 through16:

1) For **mean error** (Figure 10), the surface module outperformed the other modules in 10 out of 18 cases. The blending module outperformed the other modules in 6 out of 18 cases. There were only five occurrences where a module had a mean error greater than zero (all from the blending module). Therefore, in a domain-averaged sense, this suggests that in the Northern Great Plains domain PPAES precipitation rates are underestimated.
2) **RMSE** (Figure 11) was smallest for the surface module analyses in 11 out of 18 cases. The blending module performed better than the radar module in every case, but only had an event-average RMSE smaller than the surface modules' average RMSE in 7 out of 18 cases. These results suggest that the typical average magnitudes of the estimate's errors are smaller in the surface module in most cases.
3) **Accuracy** (Figure 12): The surface module outperformed the radar and blending modules in 17 out of 18 cases. The blending module had a higher average accuracy score than the radar module and a lesser average accuracy score than the surface module in 14 out of 18 cases. As such, the surface module had higher average accuracy and a larger fraction of its estimates correct.
4) **FAR** (Figure 13) scores were smallest in the surface analyses in every case. In 16 out of 18 cases, the blending module outperformed the radar module and produced lower average FAR scores.
5) **POFD** (Figure 14): The surface module outperformed the radar and blending modules in 17 out of 18 cases. For Case 9, the radar module outperformed both the surface and blending modules. Of those 17 cases for which the surface module performed best, 15 cases had the radar module performing better than the blending module.

6) **POD** (Figure 15): The blending module outperformed the surface and radar modules in 17 out of 18 cases. Of those cases, the surface module outperformed the radar module in 10 out of 18 cases. A single case existed in which the radar module outperformed both the blending and surface modules (Case 7).
7) **Mean ETS** (Figure 16): The surface module outperformed the radar and blending modules in 16 out of 18 cases. For the other two cases, the blending module performed best and always outperformed the radar module. Thus, the surface module has a better correspondence of estimated to observed precipitation occurrence.
8) **Case ETS** (Figure 16): In 16 of 18 cases, the surface module outperformed the other modules. The case ETS values were similar to the mean ETS values, although the cases where the surface module did not outperform the other modules were not the same. The blending module outperformed the radar module in 17 out of 18 cases. Note that for both ways ETS values were computed (Mean and Case), there were less than 20% instances of the case ETS values being less than corresponding mean ETS values. Therefore, ETS values generally increase when computed using the second (Case) method.

In Table 9, event-average values of accuracy, bias, POD, FAR, POFD, TS, ETS, and odds ratio were computed for precipitation occurrence/non-occurrence analyses. Event-average values of mean error, MAE, MSE, RMSE, multiplicative bias, and correlation were computed for precipitation rate analyses. Case 8 is absent because it did not affect the Northern Great Plains verification domain.

In the figures, a positive value implies overestimation and a negative value implies underestimation.

Table 9. Event-average values of performance metrics for each event in Northern Great Plains domain

ID	PPAES Module	ACCURACY	BIAS	POD	FAR	POFD	TS	ETS (mean)	ODDS RATIO	MEAN ERROR	MAE	MSE	RMSE	MULTIPLICATIVE BIAS	CORRELATION	ETS (case)
1	Surface	0.851	1.287	0.651	0.555	0.118	0.355	0.275	7.522	-0.012	0.086	0.057	0.187	1.003	0.480	0.333
	Radar	0.771	1.426	0.621	0.658	0.188	0.263	0.181	4.689	0.002	0.141	0.159	0.329	0.791	0.369	0.175
	Blend	0.760	2.015	0.837	0.656	0.263	0.322	0.224	3.611	0.035	0.131	0.128	0.289	1.311	0.449	0.242
2	Surface	0.878	1.149	0.826	0.209	0.136	0.675	0.540	11.953	-0.036	0.338	0.569	0.610	0.981	0.629	0.578
	Radar	0.853	1.104	0.754	0.301	0.162	0.593	0.441	10.410	-0.178	0.352	0.728	0.652	0.580	0.600	0.512
	Blend	0.879	1.333	0.907	0.291	0.187	0.672	0.531	15.263	-0.065	0.321	0.620	0.593	0.830	0.662	0.594
3	Surface	0.972	0.735	0.500	0.563	0.016	0.271	0.263	-	-0.016	0.036	0.051	0.127	0.460	0.868	0.331
	Radar	0.909	1.471	0.559	0.852	0.082	0.120	0.105	0.000	-0.007	0.052	0.059	0.168	0.544	0.443	0.129
	Blend	0.913	1.735	0.647	0.836	0.081	0.151	0.137	0.000	0.003	0.047	0.048	0.144	0.833	0.599	0.156
4	Surface	0.804	0.890	0.520	0.429	0.150	0.367	0.230	4.860	-0.070	0.143	0.134	0.279	0.585	0.335	0.286
	Radar	0.757	0.917	0.427	0.578	0.163	0.261	0.131	2.906	-0.137	0.153	0.140	0.289	0.146	0.278	0.182
	Blend	0.771	1.256	0.661	0.502	0.237	0.377	0.218	4.583	-0.087	0.148	0.132	0.277	0.492	0.377	0.273
5	Surface	0.889	1.037	0.675	0.297	0.107	0.499	0.389	11.000	-0.022	0.078	0.058	0.174	0.810	0.556	0.454
	Radar	0.825	1.012	0.526	0.509	0.129	0.313	0.207	7.036	-0.077	0.105	0.085	0.211	0.242	0.417	0.231
	Blend	0.869	1.352	0.831	0.434	0.161	0.486	0.372	4.756	-0.039	0.087	0.068	0.183	0.560	0.506	0.439
6	Surface	0.834	1.232	0.682	0.297	0.104	0.521	0.436	8.031	-0.047	0.106	0.062	0.222	0.770	0.596	0.401
	Radar	0.748	0.685	0.344	0.449	0.105	0.289	0.196	2.870	-0.134	0.152	0.107	0.301	0.124	0.287	0.148
	Blend	0.802	1.485	0.734	0.359	0.169	0.509	0.393	8.000	-0.058	0.122	0.073	0.247	0.664	0.527	0.353
7	Surface	0.910	0.508	0.114	0.879	0.041	0.078	0.056	0.250	-0.043	0.058	0.083	0.168	0.181	0.183	0.043
	Radar	0.792	2.538	0.341	0.914	0.183	0.075	0.036	0.604	-0.046	0.057	0.093	0.173	0.185	0.156	0.054
	Blend	0.873	0.841	0.144	0.907	0.084	0.072	0.044	0.444	-0.045	0.057	0.088	0.169	0.152	0.112	0.036
9	Surface	0.803	0.947	0.591	0.420	0.196	0.398	0.253	13.093	-0.047	0.128	0.085	0.260	0.610	0.512	0.361
	Radar	0.754	1.185	0.514	0.511	0.166	0.340	0.208	5.901	-0.111	0.152	0.112	0.293	0.423	0.359	0.240
	Blend	0.773	1.606	0.779	0.489	0.282	0.466	0.298	6.815	-0.054	0.129	0.085	0.252	0.685	0.531	0.343
10	Surface	0.792	1.466	0.737	0.412	0.264	0.496	0.285	6.343	-0.031	0.218	0.225	0.365	1.229	0.463	0.384
	Radar	0.731	1.760	0.745	0.546	0.347	0.399	0.192	4.548	-0.067	0.257	0.748	0.472	1.246	0.479	0.288
	Blend	0.728	2.055	0.860	0.556	0.419	0.428	0.202	4.165	-0.023	0.236	0.332	0.404	1.308	0.521	0.309
11	Surface	0.850	1.036	0.730	0.253	0.189	0.596	0.356	12.096	-0.086	0.284	0.344	0.499	0.850	0.452	0.536
	Radar	0.793	1.273	0.768	0.365	0.239	0.534	0.281	10.345	-0.207	0.319	0.390	0.529	0.592	0.411	0.412
	Blend	0.845	1.498	0.913	0.339	0.296	0.634	0.385	8.779	-0.116	0.280	0.326	0.485	0.812	0.482	0.529
12	Surface	0.952	0.785	0.456	0.388	0.021	0.325	0.308	5.900	-0.021	0.041	0.042	0.133	0.654	0.646	0.302
	Radar	0.834	2.816	0.662	0.824	0.156	0.161	0.128	3.000	0.006	0.059	0.046	0.153	1.042	0.464	0.125
	Blend	0.843	3.039	0.833	0.743	0.156	0.238	0.202	6.167	0.009	0.052	0.035	0.133	1.113	0.654	0.171
13	Surface	0.949	1.222	0.444	0.667	0.030	0.222	0.184	5.500	-0.035	0.043	0.065	0.141	0.351	0.557	0.229
	Radar	0.897	2.278	0.611	0.786	0.091	0.182	0.156	0.000	-0.025	0.044	0.048	0.132	0.549	0.488	0.159
	Blend	0.894	3.000	0.778	0.741	0.103	0.217	0.176	1.500	-0.024	0.041	0.046	0.125	0.595	0.704	0.194
14	Surface	0.890	1.018	0.726	0.351	0.221	0.560	0.284	4.200	-0.081	0.252	0.249	0.374	0.780	0.552	0.627
	Radar	0.786	1.372	0.730	0.501	0.310	0.440	0.166	6.127	-0.201	0.366	0.507	0.522	0.482	0.348	0.395
	Blend	0.799	1.656	0.857	0.466	0.389	0.504	0.191	4.958	-0.110	0.332	0.418	0.468	0.784	0.445	0.429
15	Surface	0.878	0.793	0.418	0.494	0.055	0.286	0.243	3.182	-0.056	0.085	0.105	0.231	0.462	0.411	0.238
	Radar	0.813	1.015	0.225	0.792	0.098	0.131	0.083	0.820	-0.081	0.091	0.108	0.235	0.113	0.107	0.059
	Blend	0.834	1.348	0.451	0.690	0.108	0.224	0.167	3.044	-0.055	0.088	0.106	0.233	0.472	0.326	0.175
16	Surface	0.870	1.012	0.671	0.385	0.104	0.465	0.372	14.904	-0.046	0.107	0.099	0.243	0.706	0.555	0.450
	Radar	0.794	1.340	0.623	0.585	0.178	0.327	0.214	8.458	-0.094	0.129	0.111	0.265	0.427	0.429	0.284
	Blend	0.803	1.596	0.747	0.581	0.214	0.363	0.242	11.647	-0.065	0.118	0.101	0.247	0.632	0.516	0.340
17	Surface	0.848	1.430	0.696	0.464	0.123	0.440	0.367	8.398	-0.045	0.114	0.103	0.265	0.776	0.566	0.360
	Radar	0.822	1.345	0.643	0.564	0.156	0.365	0.273	6.543	-0.073	0.117	0.111	0.275	0.420	0.561	0.303
	Blend	0.810	1.939	0.860	0.534	0.210	0.432	0.327	6.160	-0.043	0.118	0.100	0.260	0.740	0.605	0.335
18	Surface	0.908	0.951	0.620	0.359	0.067	0.429	0.351	10.639	-0.026	0.073	0.049	0.160	0.690	0.601	0.478
	Radar	0.822	1.649	0.579	0.630	0.147	0.257	0.172	7.573	-0.053	0.096	0.064	0.185	0.543	0.415	0.258
	Blend	0.830	1.990	0.763	0.612	0.185	0.307	0.215	9.188	-0.025	0.085	0.051	0.170	0.777	0.522	0.334
19	Surface	0.860	0.998	0.672	0.276	0.116	0.554	0.438	13.215	-0.075	0.204	0.231	0.402	0.755	0.471	0.517
	Radar	0.772	1.440	0.775	0.450	0.226	0.461	0.300	9.180	-0.157	0.237	0.270	0.430	0.502	0.462	0.339
	Blend	0.789	1.732	0.918	0.455	0.275	0.519	0.361	8.152	-0.093	0.214	0.214	0.392	0.770	0.500	0.401

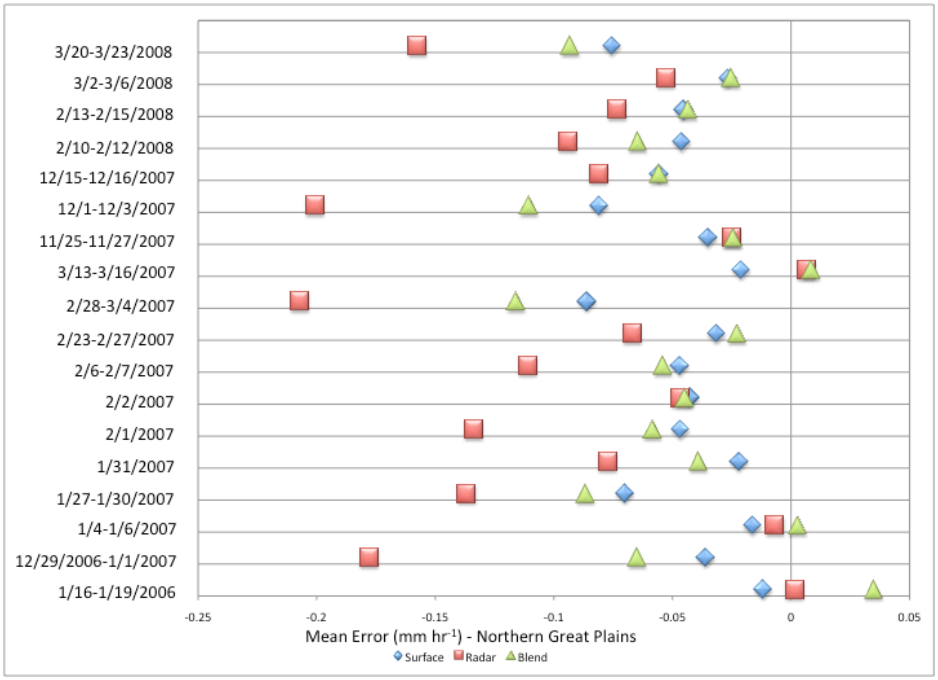

Figure 10. Northern Great Plains mean errors averaged over duration of each case (perfect score=0)

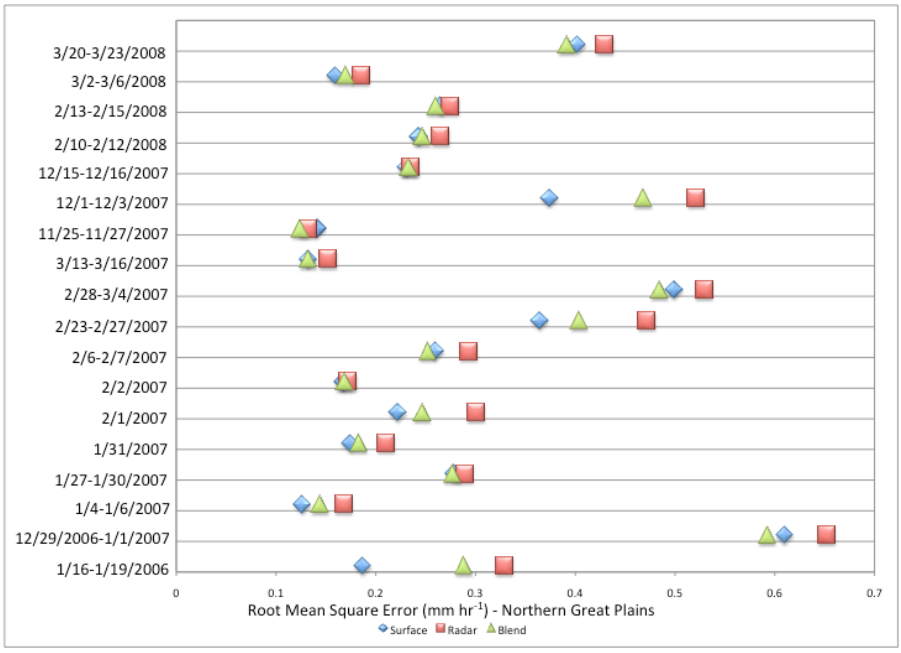

Figure 11. Northern Great Plains RMSE statistics averaged over duration of each case (perfect score=0)

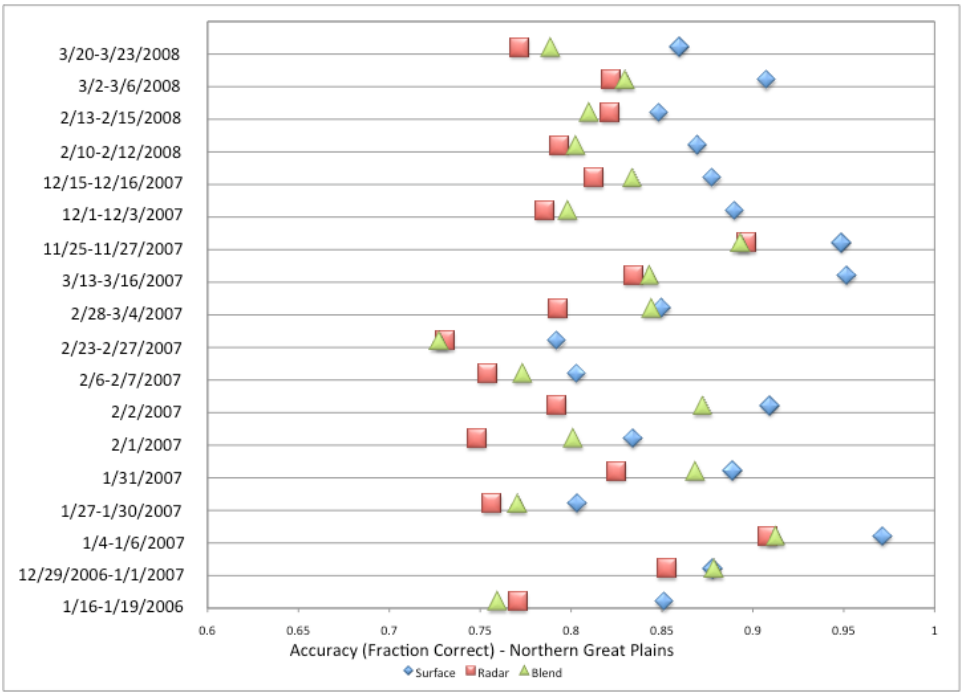

Figure 12. Accuracy averaged over duration of each case for Northern Great Plains domain for surface, radar, and blending modules (perfect score=1)

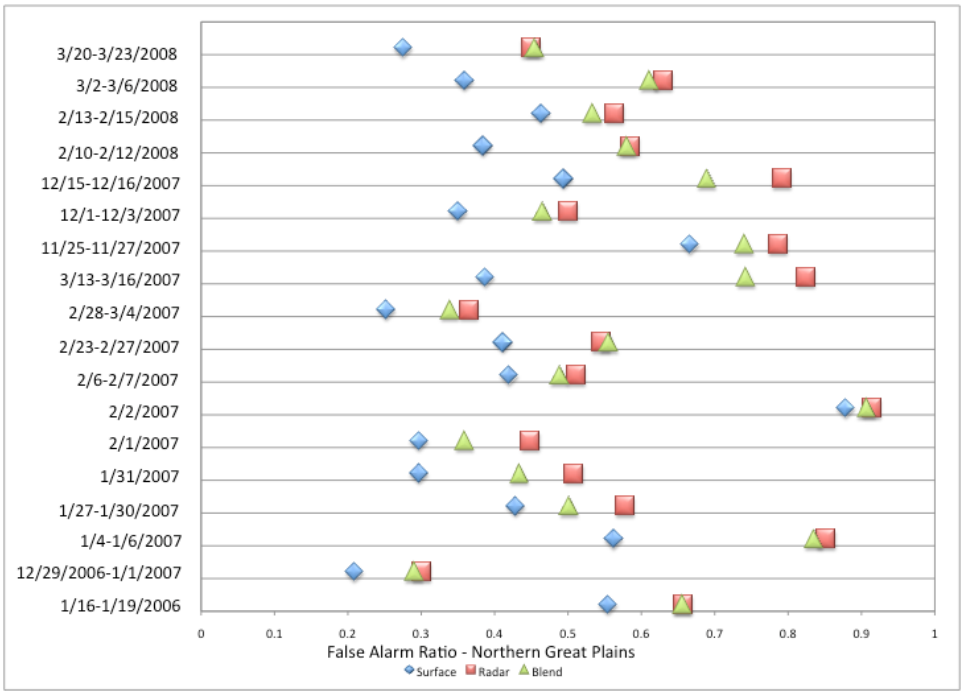

Figure 13. Northern Great Plains FAR statistics averaged over duration of each case for surface, radar, and blending modules (perfect score=0)

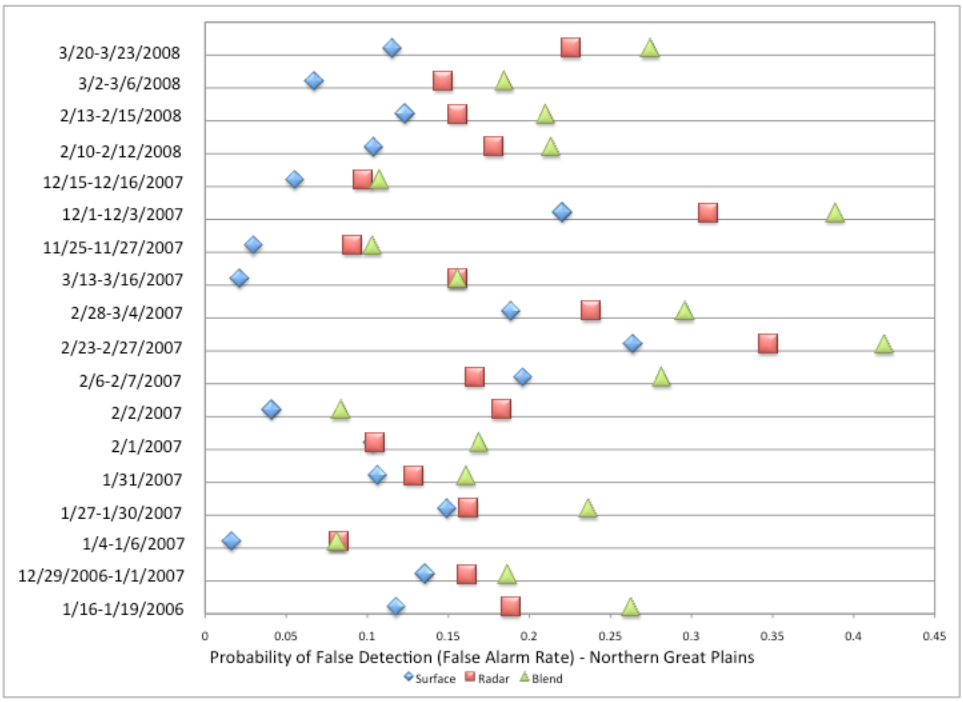

Figure 14. Northern Great Plains POFD scores averaged over duration of each case for surface, radar, and blending modules (perfect score=0)

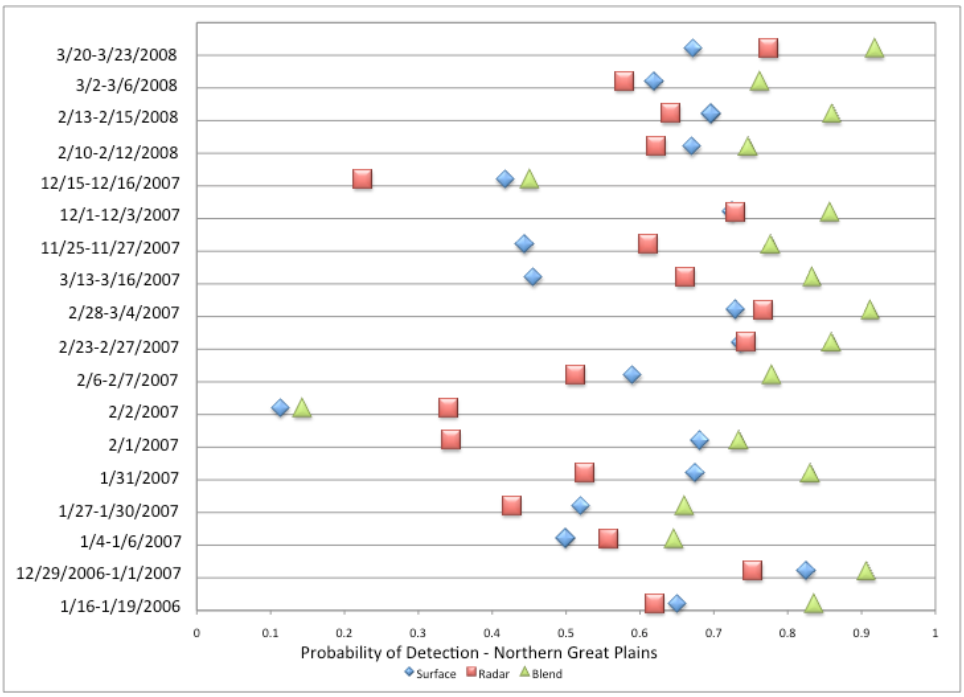

Figure 15 Northern Great Plains POD scores averaged over duration of each case for surface, radar, and blending modules (perfect score=1)

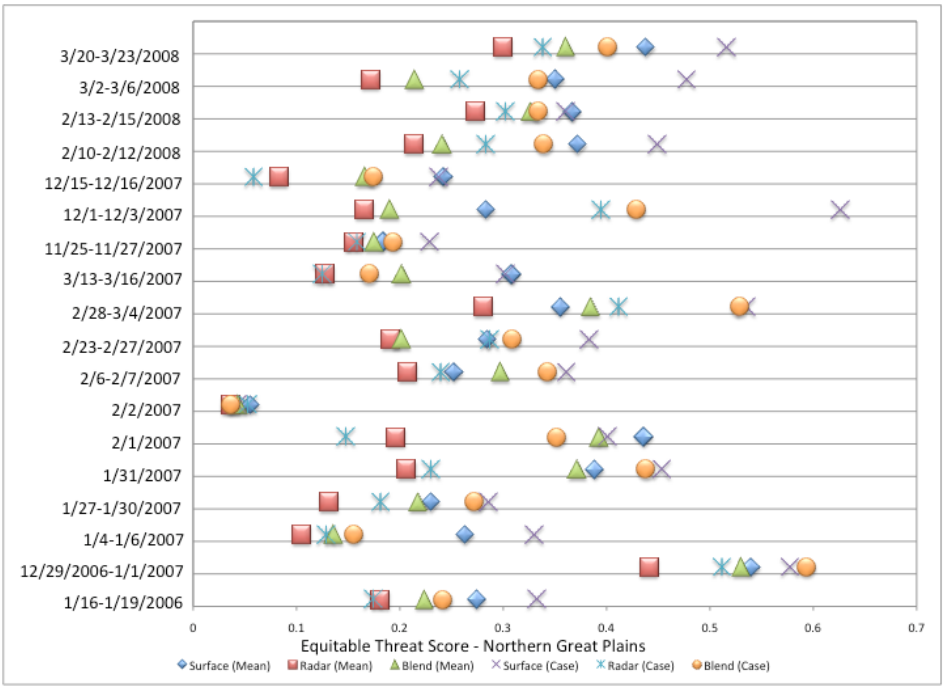

Figure 16. Northern Great Plains ETS averaged over duration of each case (perfect score=1)

Results in Figure 16 are for the computed average and case ETS (as defined in text).

5.1.1.2 Case Stratification

Cases were stratified by event type and predominant precipitation intensity to see if performance is related to these characteristics. Comparisons were made for: a) RMSE for precipitation rate and b) POD, FAR, and ETS for precipitation occurrence/non-occurrence (Figure 17). Note that the sample sizes are small for the various stratification groups (no group has a sample size greater than 10). As such, this limits how the stratification results can be interpreted.

With regard to event type, surface trough cases had a higher average event-average RMSE than frontal boundary or low pressure system cases for this set of historical cases (Figure 17a). Additionally, the surface module had a lower event-average RMSE, regardless of event type. This is similar to the results observed in Figure 11, where the surface module outperformed the other modules in 11 out of 18 cases. For each of the event types, the event-average RMSEs increased, in order by module, from surface to blending to radar. Event-average RMSEs were highest when the predominant intensity was heavy for each of the PPAES modules (Figure 17b). For all of the event-average RMSEs, the radar produced the highest RMSEs regardless of the event type of predominant precipitation intensity.

Results for POD are shown in Figures17c and d. The blending module tends to have a higher POD – regardless of the event type or predominant precipitation intensity. This is consistent with the full results discussed in the previous subsection (Figure 15).

Results for FAR are shown in Figures 17e and f. The surface module has a lesser number of false alarms than the other modules regardless of event type or predominant precipitation intensity. Note, however, that the blending module did outperform the radar module in both stratification groups.

Average values of case ETS were highest for the surface module for all event types and predominant precipitation intensities (Figures 17g and h). Furthermore, the blending module outperformed the radar module regardless of event type or predominant precipitation intensity. The surface module consistently had the best correspondence of estimated precipitation occurrence with observed precipitation occurrence.

Only the primary synoptic feature (event type) and precipitation characteristic (predominant precipitation intensity) for each case were used. The frequencies of event types for the Northern Great Plains cases are five, seven, and six for surface trough, frontal, and low pressure system events, respectively. The frequencies of predominant precipitation intensity types for the Northern Great Plains cases are nine, five, and four for light-, moderate-, and heavy-type events, respectively.

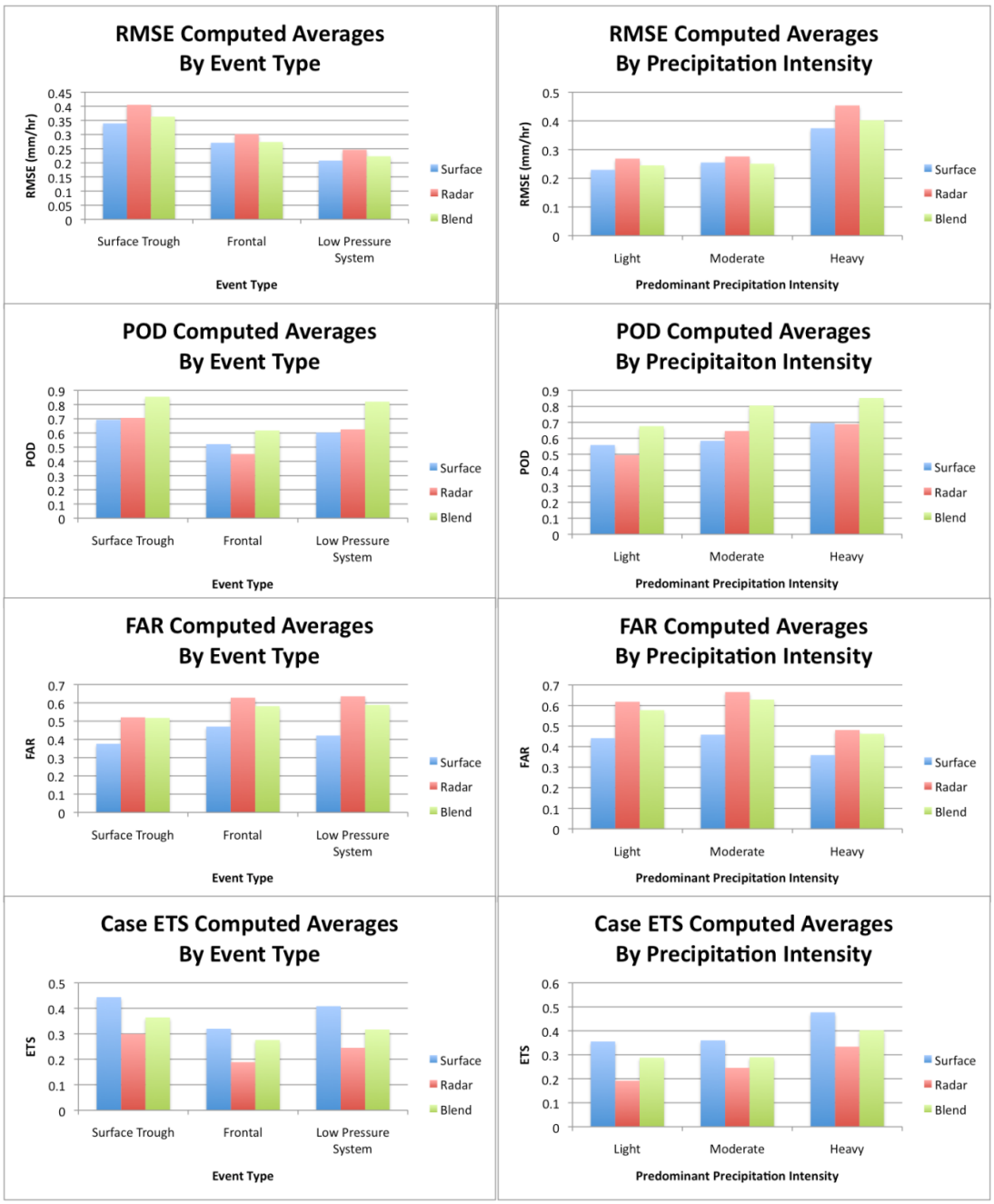

Figure 17. Eevent-average RMSE (a) and (b), POD (c) and (d), FAR (e) and (f), and case ETS (g) and (h) for all events per category for surface, radar, and blending modules

5.1.2 Midwest Domain

5.1.2.1 Historical Case Validation

Nineteen cases were used to evaluate PPAES products in the Midwest verification domain for both precipitation occurrence/non-occurrence and rate. The cases include one high pressure system, three surface troughs, ten frontal systems, and five low pressure system events (Table 10). Of these 19 events, six, ten, and three events were identified as having a predominant

precipitation intensity of light, moderate, and heavy. Event-average values were computed for each performance metric over the duration of each event (Table 10). Event-average values of mean error, RMSE, accuracy, FAR, POFD, POD, and ETS (mean- and case-computed) for all of the cases are shown in Figures 18 through 24.

Several trends are apparent from Table 10 and Figures 18 through 24:

1) **Mean Errors** (Figure 18) were smallest in either the surface or blending module analyses in 16 out of 18 cases. As with the Northern Plains domain, the surface module outperformed the other modules in the majority (12 out of 19) of cases. The modules all had a positive mean error five times, suggesting that PPAES precipitation rates are underestimated in a domain-averaged sense.
2) **RMSEs** (Figure 19) were smallest for the surface module in 14 out of 19 cases. The blending module performed better than the radar module in every case except one. As with the Northern Great Plains Domain, the radar module generally performed worse than the other modules (14 out of 19 cases). These results suggest that the typical average error magnitudes are smaller with the surface module.
3) **Accuracy** scores (Figure 20) were observed to be largest in the surface module in 18 out of 19 cases. The blending module outperformed the radar module in 15 out of 19 cases. Thus, the surface module has higher average accuracy and, as such, a larger fraction of its estimates correct.
4) **FAR** scores (Figure 21) were smallest with the surface module in every case. In 13 out of 19 cases, the blending module outperformed the radar module and had lower average FAR values.
5) **For POFD** (Figure 22), the surface module outperformed the radar and blending modules in every case. Also, the blending module outperformed the radar module in 10 out of 19 cases.
6) **POD** (Figure 23) is larger for the blending module in 17 out of 19 cases. The radar module outperformed the surface module in 15 out of 19 cases and performed the best in two cases (Cases 6 and 17). Thus, the blending module had the highest fraction of correct.
7) **Mean ETS** (Figure 24): The surface module outperformed the radar and blending modules in 16 out of 19 cases. The blending module outperformed the radar module in 15 out of 19 cases. As observed with the Northern Plains domain, the surface module more often corresponds estimated precipitation occurrence to observed precipitation occurrence.
8) **Case ETS** (Figure 25): The surface module outperformed the other modules in 17 out of 19 cases. Similar to the Northern t Plains domain, the blending module outperformed the radar module in 16 out of 19 cases. As such, the surface module performed better in regard to case ETSs and had a better correspondence of estimated to observed precipitation occurrences). In 14% of the possible realizations, the case ETS value was less than the corresponding mean ETS value.

In Table 10, event-average values of accuracy, bias, POD, FAR, POFD, TS, ETS, and odds ratio were computed for precipitation occurrence/non-occurrence analyses along with a case ETS. Event-average values of mean error, MAE, MSE, RMSE, multiplicative bias, and correlation were computed for precipitation rate analyses.

Table 10. Event-average values of performance metrics for each event in Midwest domain

ID	PPAES Module	ACCURACY	BIAS	POD	FAR	POFD	TS	ETS (mean)	ODDS RATIO	MEAN ERROR	MAE	MSE	RMSE	MULTIPLICATIVE BIAS	CORRELATION	ETS (case)
1	Surface	0.913	0.865	0.701	0.158	0.039	0.635	0.534	14.667	-0.028	0.129	0.164	0.287	0.890	0.710	0.575
	Radar	0.839	1.772	0.845	0.516	0.192	0.447	0.325	10.417	-0.035	0.162	0.176	0.321	1.029	0.644	0.432
	Blend	0.854	1.862	0.928	0.503	0.186	0.477	0.362	19.717	-0.006	0.155	0.169	0.317	1.241	0.704	0.474
2	Surface	0.874	0.891	0.599	0.331	0.096	0.493	0.391	8.939	-0.040	0.385	0.980	0.767	0.898	0.561	0.481
	Radar	0.877	1.159	0.711	0.422	0.104	0.513	0.426	9.728	-0.232	0.323	0.859	0.681	0.465	0.649	0.505
	Blend	0.874	1.226	0.739	0.441	0.125	0.509	0.415	8.671	-0.162	0.319	0.808	0.667	0.586	0.644	0.509
3	Surface	0.888	1.042	0.804	0.219	0.095	0.644	0.539	12.977	0.013	0.458	1.475	0.931	1.162	0.667	0.566
	Radar	0.839	1.228	0.749	0.411	0.163	0.508	0.392	6.967	-0.182	0.427	1.448	0.898	0.771	0.634	0.448
	Blend	0.850	1.427	0.905	0.399	0.195	0.566	0.440	7.750	-0.051	0.414	1.241	0.843	1.081	0.696	0.495
4	Surface	0.866	0.993	0.718	0.215	0.088	0.606	0.492	13.857	-0.092	0.168	0.187	0.353	0.741	0.604	0.518
	Radar	0.832	1.116	0.723	0.311	0.134	0.572	0.446	11.301	-0.172	0.197	0.222	0.389	0.370	0.556	0.447
	Blend	0.847	1.332	0.851	0.310	0.171	0.626	0.493	14.361	-0.120	0.172	0.179	0.345	0.629	0.637	0.499
5	Surface	0.922	1.375	0.750	0.503	0.062	0.477	0.437	6.000	-0.011	0.049	0.027	0.132	0.929	0.618	0.425
	Radar	0.824	2.069	0.667	0.733	0.160	0.234	0.179	5.297	-0.035	0.058	0.045	0.161	0.438	0.528	0.189
	Blend	0.871	2.208	0.833	0.649	0.127	0.332	0.275	10.300	-0.019	0.053	0.035	0.140	0.817	0.665	0.310
6	Surface	0.789	1.086	0.595	0.350	0.140	0.471	0.347	9.313	-0.048	0.116	0.068	0.238	0.744	0.464	0.323
	Radar	0.713	1.895	0.860	0.498	0.352	0.464	0.273	8.955	-0.094	0.133	0.072	0.246	0.427	0.534	0.273
	Blend	0.699	1.903	0.830	0.515	0.364	0.448	0.256	5.277	-0.051	0.119	0.060	0.219	0.723	0.557	0.251
7	Surface	0.891	0.700	0.473	0.364	0.036	0.413	0.368	4.188	-0.051	0.077	0.046	0.190	0.456	0.524	0.387
	Radar	0.756	1.942	0.580	0.704	0.219	0.250	0.147	4.630	-0.062	0.091	0.060	0.203	0.409	0.367	0.181
	Blend	0.837	1.436	0.639	0.497	0.136	0.404	0.331	8.371	-0.054	0.079	0.047	0.192	0.432	0.554	0.325
8	Surface	0.902	0.895	0.592	0.483	0.069	0.380	0.331	15.417	-0.063	0.133	0.135	0.291	0.575	0.812	0.452
	Radar	0.838	1.751	0.656	0.614	0.151	0.332	0.258	10.833	-0.193	0.203	0.441	0.475	0.136	0.524	0.318
	Blend	0.862	1.695	0.726	0.636	0.133	0.336	0.278	6.333	-0.132	0.164	0.244	0.365	0.338	0.651	0.389
9	Surface	0.824	0.847	0.508	0.385	0.122	0.420	0.288	8.012	-0.051	0.151	0.117	0.282	0.596	0.485	0.419
	Radar	0.765	1.174	0.498	0.543	0.215	0.361	0.206	5.610	-0.150	0.208	0.222	0.367	0.374	0.393	0.320
	Blend	0.777	1.460	0.639	0.526	0.255	0.428	0.266	5.628	-0.089	0.173	0.142	0.308	0.594	0.461	0.364
10	Surface	0.820	0.975	0.683	0.312	0.153	0.517	0.357	7.830	-0.140	0.301	0.528	0.561	0.669	0.492	0.424
	Radar	0.820	1.258	0.821	0.388	0.205	0.536	0.384	13.794	-0.094	0.342	0.895	0.671	0.698	0.531	0.446
	Blend	0.817	1.281	0.833	0.397	0.227	0.535	0.381	12.884	-0.106	0.317	0.739	0.606	0.661	0.514	0.446
11	Surface	0.864	1.041	0.757	0.239	0.151	0.629	0.428	12.969	-0.117	0.302	0.569	0.505	0.804	0.509	0.551
	Radar	0.754	2.054	0.884	0.488	0.405	0.481	0.226	7.116	-0.201	0.355	0.753	0.595	0.652	0.439	0.352
	Blend	0.777	2.063	0.942	0.482	0.416	0.503	0.240	13.598	-0.144	0.315	0.627	0.534	0.784	0.526	0.399
12	Surface	0.939	0.850	0.717	0.280	0.044	0.592	0.501	6.000	-0.076	0.257	1.018	0.523	0.851	0.658	0.577
	Radar	0.714	2.521	0.971	0.779	0.341	0.216	0.129	11.133	-0.014	0.267	0.929	0.569	1.721	0.781	0.207
	Blend	0.718	2.580	0.975	0.780	0.343	0.217	0.125	2.444	0.005	0.274	0.989	0.576	1.749	0.769	0.213
13	Surface	0.930	1.054	0.736	0.329	0.047	0.547	0.493	10.667	-0.006	0.172	0.296	0.384	1.064	0.706	0.584
	Radar	0.836	1.805	0.747	0.653	0.174	0.333	0.254	4.006	0.025	0.234	0.695	0.507	0.896	0.564	0.353
	Blend	0.852	1.891	0.877	0.608	0.168	0.376	0.297	3.875	0.044	0.230	0.627	0.503	1.186	0.644	0.400
14	Surface	0.851	0.837	0.625	0.176	0.076	0.537	0.424	13.071	-0.197	0.440	1.126	0.815	0.740	0.617	0.482
	Radar	0.763	1.350	0.589	0.480	0.243	0.424	0.249	6.809	-0.238	0.435	1.004	0.802	0.600	0.416	0.324
	Blend	0.803	1.445	0.745	0.413	0.225	0.521	0.351	7.866	-0.179	0.410	0.929	0.750	0.848	0.585	0.409
15	Surface	0.912	1.009	0.831	0.169	0.083	0.728	0.606	23.893	-0.107	0.252	0.404	0.474	0.855	0.673	0.677
	Radar	0.797	1.522	0.861	0.412	0.214	0.517	0.365	17.492	-0.171	0.323	0.630	0.610	0.726	0.540	0.412
	Blend	0.807	1.512	0.867	0.404	0.212	0.536	0.395	12.377	-0.156	0.305	0.583	0.575	0.776	0.561	0.434
16	Surface	0.831	1.168	0.723	0.347	0.202	0.552	0.374	8.040	-0.045	0.158	0.120	0.284	0.806	0.569	0.460
	Radar	0.758	2.075	0.831	0.474	0.400	0.490	0.242	3.028	-0.097	0.187	0.158	0.312	0.628	0.526	0.353
	Blend	0.752	2.225	0.903	0.484	0.427	0.499	0.243	3.043	-0.065	0.182	0.141	0.295	0.760	0.571	0.349
17	Surface	0.920	1.300	0.500	0.549	0.056	0.365	0.340	0.550	-0.020	0.048	0.042	0.141	0.701	0.538	0.196
	Radar	0.829	3.267	0.850	0.810	0.173	0.189	0.153	0.450	-0.015	0.053	0.043	0.154	0.717	0.591	0.154
	Blend	0.835	3.117	0.833	0.806	0.164	0.191	0.155	1.139	-0.013	0.053	0.042	0.152	0.808	0.512	0.156
18	Surface	0.918	0.923	0.703	0.196	0.039	0.612	0.569	13.588	-0.064	0.203	0.395	0.472	0.876	0.752	0.570
	Radar	0.774	1.997	0.701	0.550	0.222	0.394	0.289	7.956	-0.058	0.272	0.701	0.613	0.851	0.453	0.269
	Blend	0.801	1.982	0.786	0.523	0.208	0.435	0.333	8.505	-0.068	0.250	0.616	0.575	0.883	0.550	0.332
19	Surface	0.926	1.098	0.703	0.293	0.048	0.560	0.512	13.808	-0.022	0.115	0.096	0.248	0.985	0.672	0.639
	Radar	0.748	2.607	0.775	0.636	0.288	0.334	0.220	8.471	-0.006	0.198	0.197	0.349	1.228	0.509	0.283
	Blend	0.751	2.690	0.818	0.616	0.285	0.356	0.249	10.194	0.012	0.178	0.158	0.324	1.379	0.535	0.291

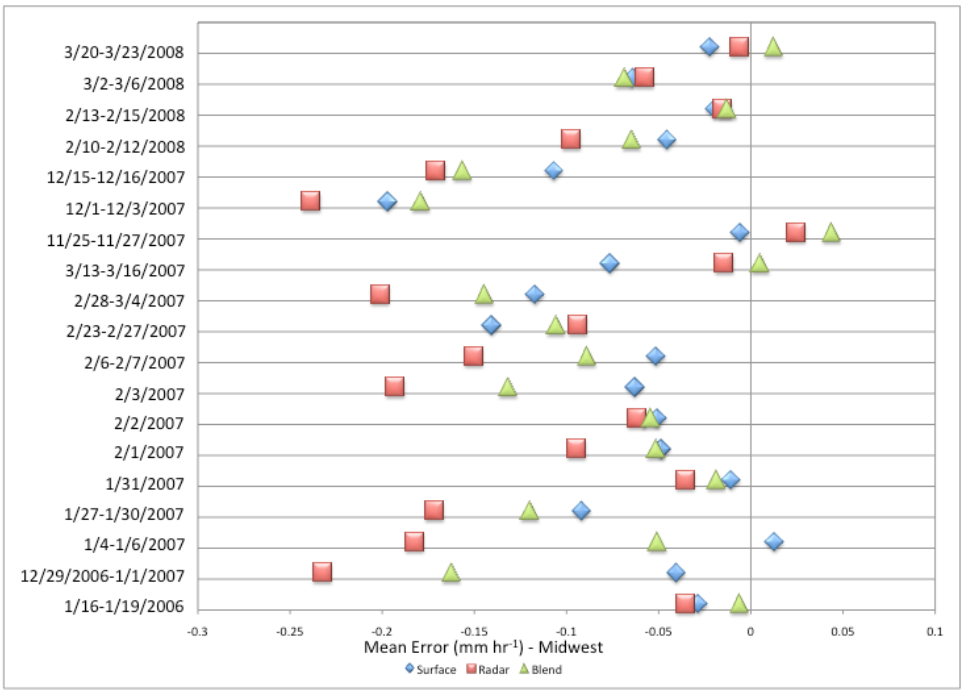

Figure 18. Midwest domain mean errors averaged over duration of each case for surface, radar, and blending modules (perfect score=0)

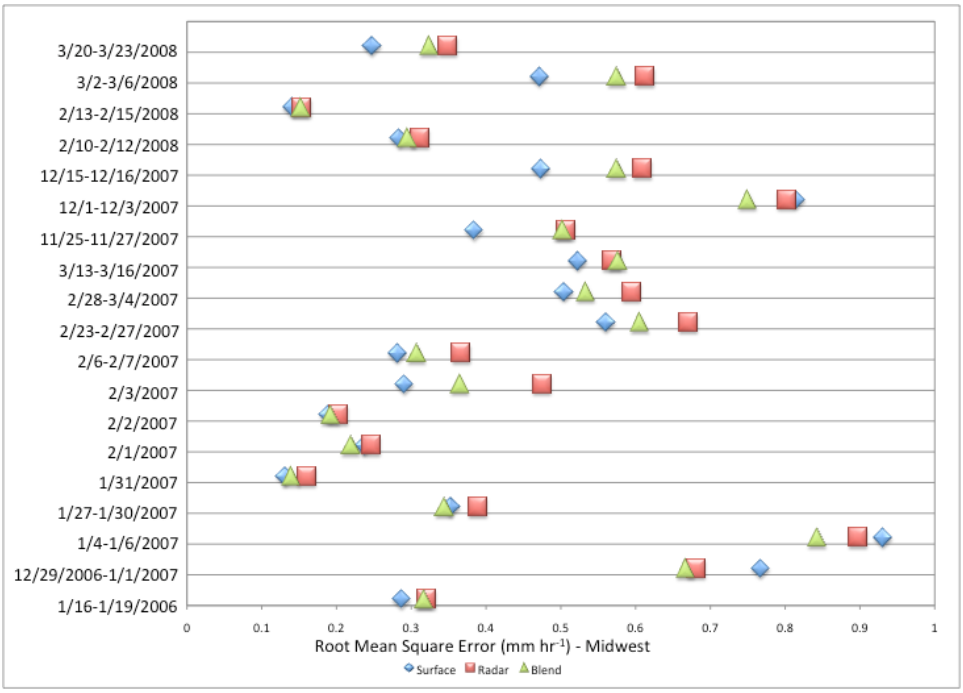

Figure 19. Event-average root mean square error for surface, radar, and blending modules in Midwest domain (perfect score=0)

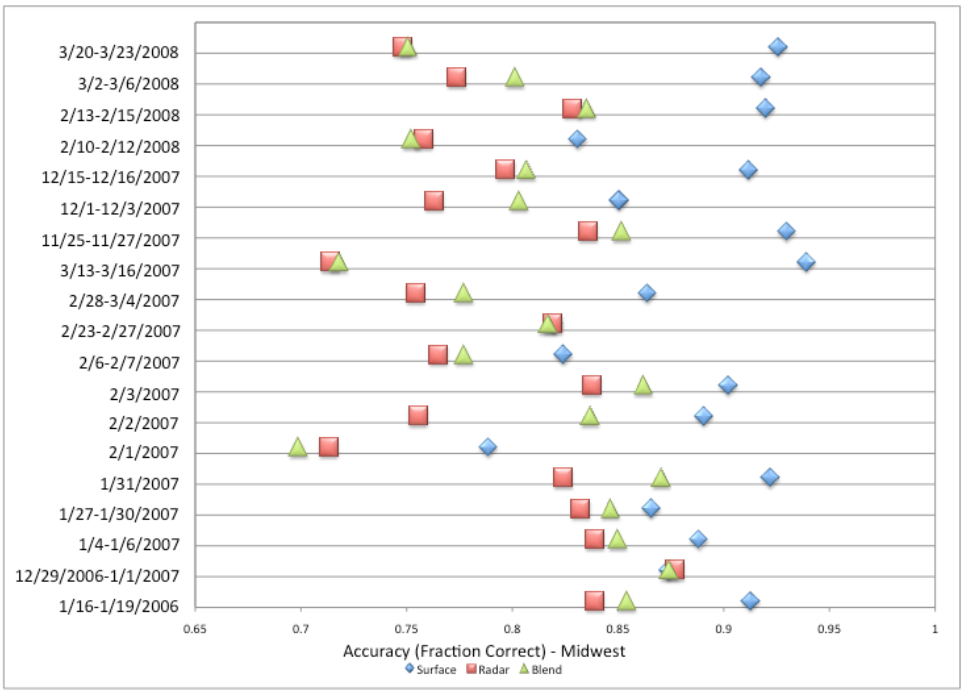

Figure 20. Event-average accuracy in Midwest domain for surface, radar, and blending modules (perfect score=1)

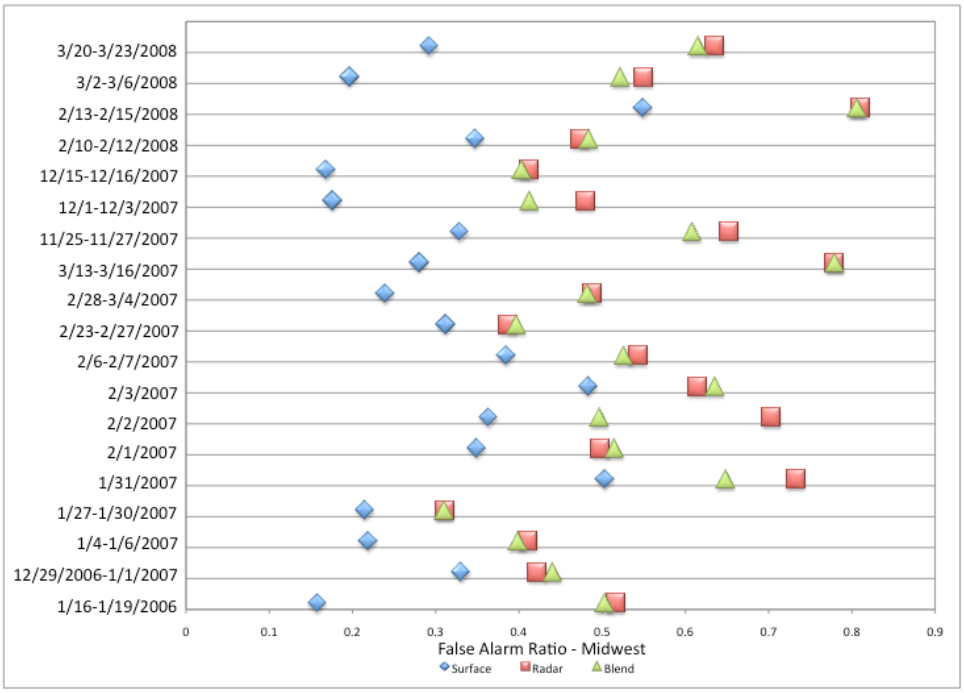

Figure 21. Event-average false alarm ratio in Midwest domain for surface, radar, and blending modules (perfect score=0)

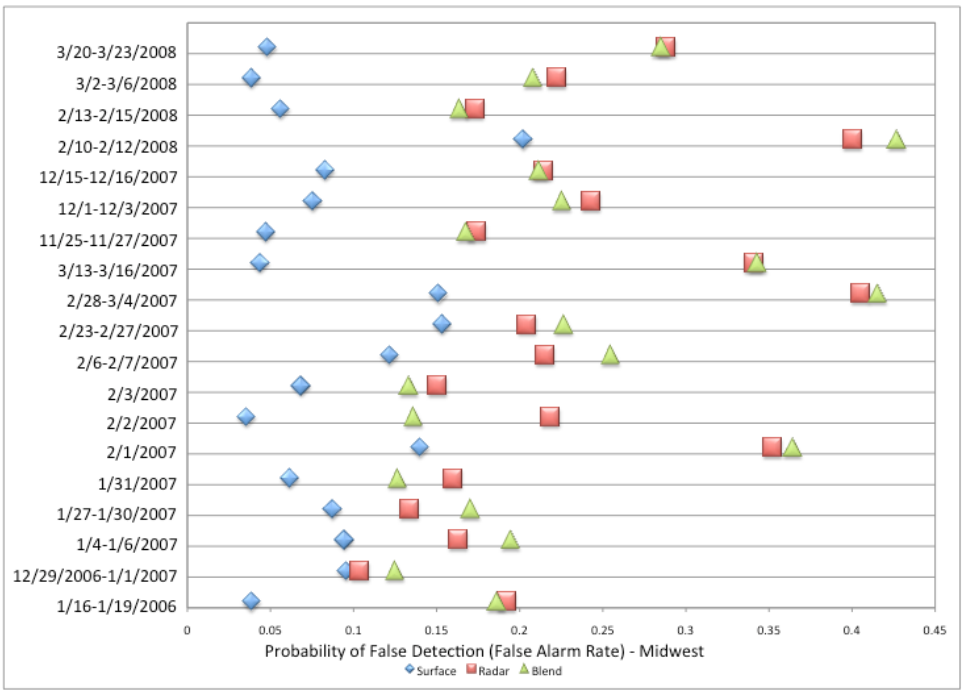

Figure 22. POFD in Midwest domain for radar, surface, and blending modules (perfect score=0)

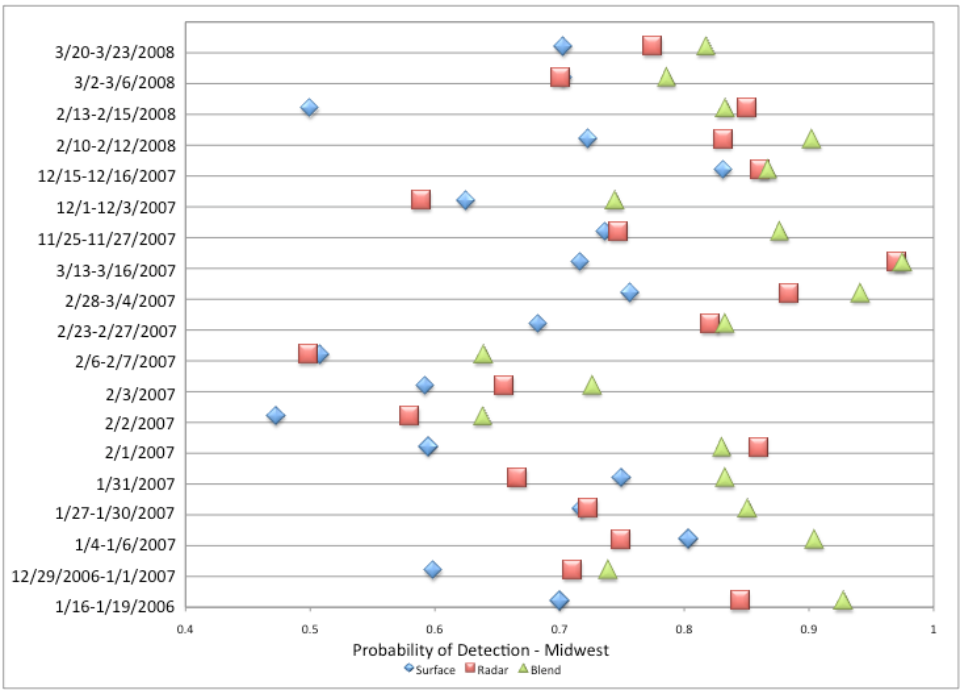

Figure 23. Event-average probability of detection in Midwest domain for radar, surface, and blending modules (perfect score=1)

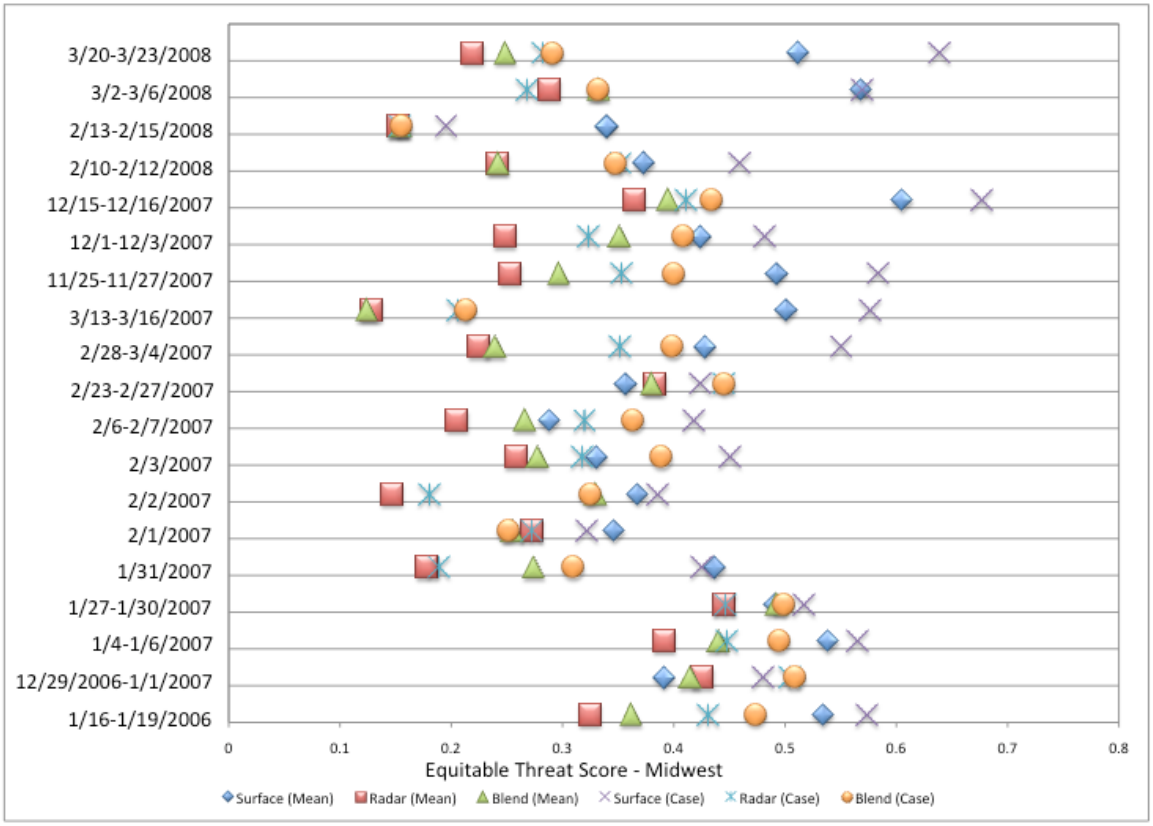

Figure 24. Event-average equitable threat scores (average and case) in Midwest domain for surface, radar, and blending modules (perfect score=1)

5.1.2.2 Event Stratification

The historical cases in the Midwest domain were also stratified to look for possible similarities among events and predominant precipitation intensities (Figure 25). In this domain, one event type (high pressure) existed that was not seen in the Northern Plains domain. As this was the only occurrence, the high pressure case was excluded from the event type comparisons. As with the Northern Great Plains domain, the sample sizes were small and this limits how the stratification results can be interpreted.

Cases for which the main synoptic feature was a surface trough had a lower event-average RMSE (as seen in Figure 25a). This result is different from what was observed in the Northern Plains domain. With regard to predominant precipitation intensity, RMSE values were highest for heavy intensities for each of the PPAES modules (Figure25b), similar to what was seen for the Northern Great Plains domain (Figure 17b). As such, cases where the predominant precipitation intensity was greater tended to have higher RMSEs, regardless of event type, for all modules.

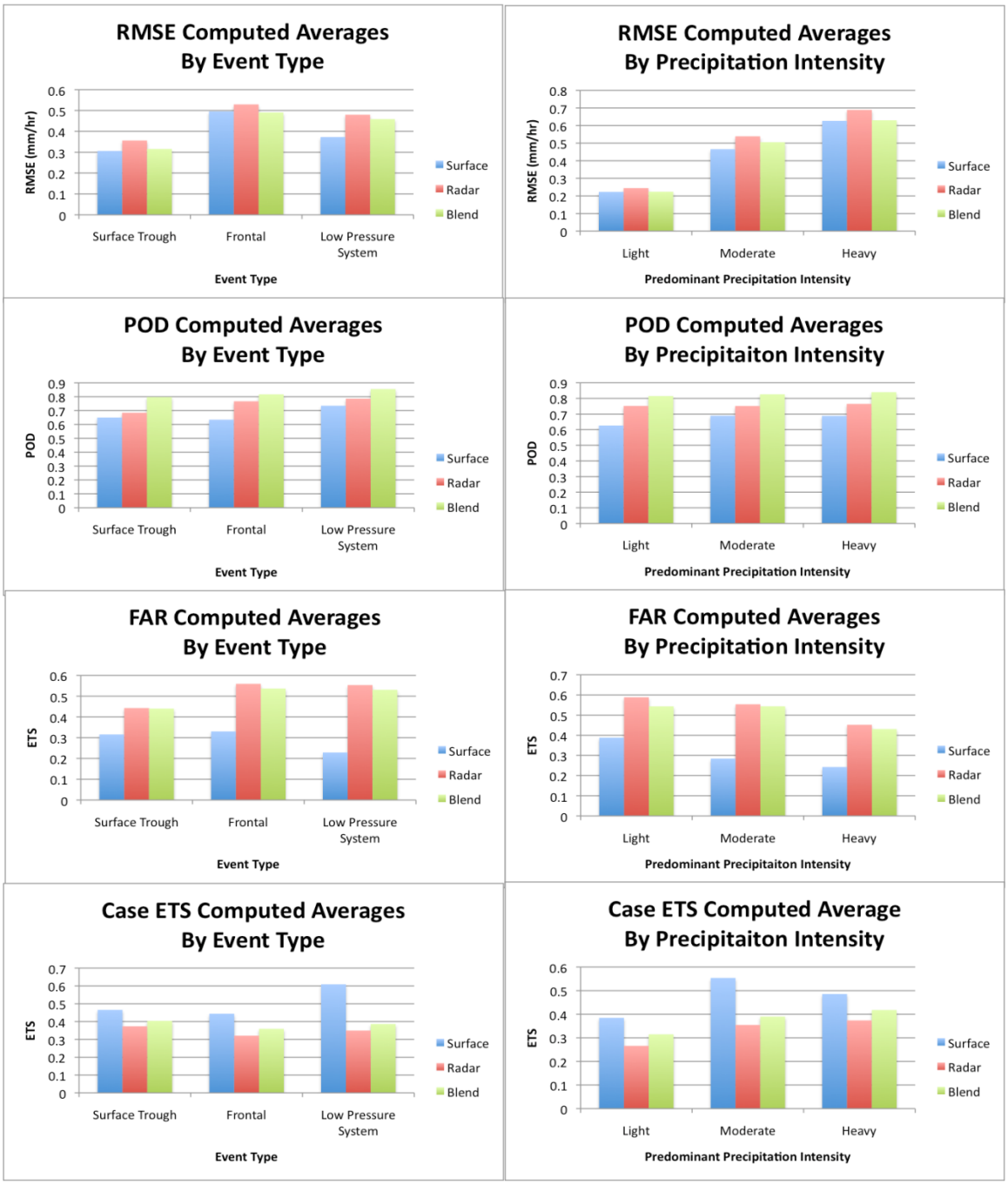

Figure 25. Average event-average RMSE (a) and (b), POD (c) and (d), FAR (e) and (f), and case ETS (g) and (h) for all events per category for surface, radar, and blending modules

Figures 25c and d show results for the POD statistic and indicate that the blending module tends to have a higher POD regardless of the stratification. A difference seen with this domain relative to the Northern Great Plains domain is that each module had an average POD greater than 0.6. Event-average POD was seen to increase from the surface module (lowest) to the blending module (highest) regardless of event type or predominant precipitation intensity.

Figures 25e and f show results for the FAR statistic and imply that the surface module performed better than the other two modules. The blending module outperformed the radar module, by a small amount, regardless of the stratification.

Figures 25g and h show results for mean ETS values. ETS values were again highest for the surface module for all event types and predominant precipitation intensities, and, as with the Northern Great Plains domain, the blending module outperformed the radar module regardless of event type or predominant precipitation intensity. Thus, the surface module has a better correspondence of estimated to observed precipitation events.

Only the primary synoptic feature (event type) and precipitation characteristic (predominant precipitation intensity) for each case were used. The frequencies of event types for the Midwest cases are three, ten, and five for surface trough, frontal, and low pressure system events, respectively. The frequencies of predominant precipitation intensity types for the Midwest cases are six, ten, and three for light-, moderate-, and heavy-type events, respectively.

5.1.3 Significance Testing

A one-way ANOVA was completed for POD, FAR, accuracy, and case ETS to see if differences in PPAES module performance were statistically significant. The computed *F*-statistic for event-averages of POD, FAR, accuracy, and ETS (case) were all greater than the critical value at the $\alpha = .05$ level (Table 11). As such, the three population means can be considered different for each of the performance metrics. However, a one-way ANOVA does not indicate which population mean is the smallest or largest.

The two-sample *t*-test (with equal variances assumed) was used to obtain information regarding the relationship between event-average performance metrics and the direction of the population mean differences (Weiss 2008). As with the one-way ANOVA test, the test was performed at the 5% significance level. The results of the two-sample *t*-test showed that nine out of the twelve two-sample *t*-test combinations were statistically significant. Therefore, the null hypothesis that the means are equal can be rejected for each of these module combinations (Table 12).

The main conclusions drawn from significance testing regarding the historical cases are:

1) The blending module performed best with respect to event-average POD.
2) The surface module performed best with respect to event-average accuracy.
3) The surface module performed best with respect to event-average FAR.
4) The surface module performed best with respect to event-average case ETS.

In Table 11, population means are denoted as μ_s, μ_r, and μ_b for the PPAES surface, radar, and blending modules, respectively. The degrees of freedom for the numerator and denominator for all hypothesis tests were 2 and 54, respectively. The sample size for each module was 19.

In Table 12, population means are denoted as μ_s, μ_r, and μ_b for the PPAES surface, radar, and blending modules, respectively. The degrees of freedom for all hypothesis tests were 36. The sample size for each module was 19.

Table 11. One-way analysis of variance (ANOVA) using a critical-value approach

One-way Analysis of Variance (ANOVA)				
Performance Metric	**Null Hypothesis (H_o)**	**Alternative Hypothesis (H_a)**	***F*-statistic (F)**	**Critical Value - 5% (F_a)**
Probability of Detection	$\mu_s = \mu_r = \mu_b$	Not all the means are equal.	10.9535	3.1682
Accuracy	$\mu_s = \mu_r = \mu_b$	Not all the means are equal.	19.8007	3.1682
False Alarm Ratio	$\mu_s = \mu_r = \mu_b$	Not all the means are equal.	19.2546	3.1682
Equitable Threat Score (Case)	$\mu_s = \mu_r = \mu_b$	Not all the means are equal.	14.4778	3.1682

Table 12. Two-sample *t*-test with equal variances assumed using a critical value approach

Two-sample *t*-test with Equal Variances Assumed						
	Null Hypothesis (H_o)	**Alternative Hypothesis (H_a)**		***t* Critical Value - 5%**		
Performance Metric	***Sample* μ_1 = *Sample* μ_2**	**Two Tailed**	**One Tailed**	**Test Statistic (t)**	**Two Tailed**	**One Tailed**
Probability of Detection	$\mu_r = \mu_s$	$\mu_r \neq \mu_s$	$\mu_r > \mu_s$	1.1141	2.0281	1.6883
	$\mu_b = \mu_s$	$\mu_b \neq \mu_s$	$\mu_b > \mu_s$	4.4202	2.0281	1.6883
	$\mu_b = \mu_r$	$\mu_b \neq \mu_r$	$\mu_b > \mu_r$	3.2951	2.0281	1.6883
Accuracy	$\mu_s = \mu_r$	$\mu_s \neq \mu_r$	$\mu_s > \mu_r$	6.0160	2.0281	1.6883
	$\mu_s = \mu_b$	$\mu_s \neq \mu_b$	$\mu_s > \mu_b$	4.5988	2.0281	1.6883
	$\mu_b = \mu_r$	$\mu_b \neq \mu_r$	$\mu_b > \mu_r$	1.3470	2.0281	1.6883
False Alarm Ratio	$\mu_s = \mu_r$	$\mu_s \neq \mu_r$	$\mu_s < \mu_r$	5.7040	2.0281	1.6883
	$\mu_s = \mu_b$	$\mu_s \neq \mu_b$	$\mu_s < \mu_b$	5.2440	2.0281	1.6883
	$\mu_b = \mu_r$	$\mu_b \neq \mu_r$	$\mu_b < \mu_r$	0.7248	2.0281	1.6883
Equitable Threat Score (Case)	$\mu_s = \mu_r$	$\mu_s \neq \mu_r$	$\mu_s > \mu_r$	5.2864	2.0281	1.6883
	$\mu_s = \mu_b$	$\mu_s \neq \mu_b$	$\mu_s > \mu_b$	3.2741	2.0281	1.6883
	$\mu_b = \mu_r$	$\mu_b \neq \mu_r$	$\mu_b > \mu_r$	2.0682	2.0281	1.6883

5.2 Surface Analyses with Clarus Data

PPAES was modified to analyze Clarus station data in the surface analysis module. The idea was that by including Clarus data in the surface analysis, the number of 'truth' observations would drastically increase, resulting in better analyses. As Clarus data were not available for the previous 19 cases examined in the initial analysis, new cases were chosen for each domain as indicated in Table 6. Similar statistics (accuracy, FAR, POD, ETS, and RMSE) were calculated for each of these cases using the same verification process. Because the focus of this analysis was determining the value of adding Clarus to the surface analysis, no runs that involved blending radar and surface data were performed. These tests compare the resuts obtained using only MADIS data and using MADIS+Clarus data. Tests were performed using each of the three domains.

5.2.1 Northern Great Plains Domain

Event-average values of bias, FAR, POD, and ETS for the Northern Great Plains are shown in Figure 26. Several trends are apparent in these plots. The bias drops in all cases when Clarus data are used. However, POD and ETS also decrease in all cases. FAR results are mixed, with inclusion of Clarus data both increasing and decreasing values. It is apparent that addition of Clarus data may have had some benefits, but overall it appears as if including Clarus data decreased performance. This may result from ESS instrumentation not being monitored/calibrated as often as MADIS instrumentation or it may result because the performance characteristics of ESS and MADIS instrumentation are different. These findings are reinforced by RMSE results (Figure 27), which show an overall increase in RMSE values for all NGP cases except one (for which this value was essentially unchanged).

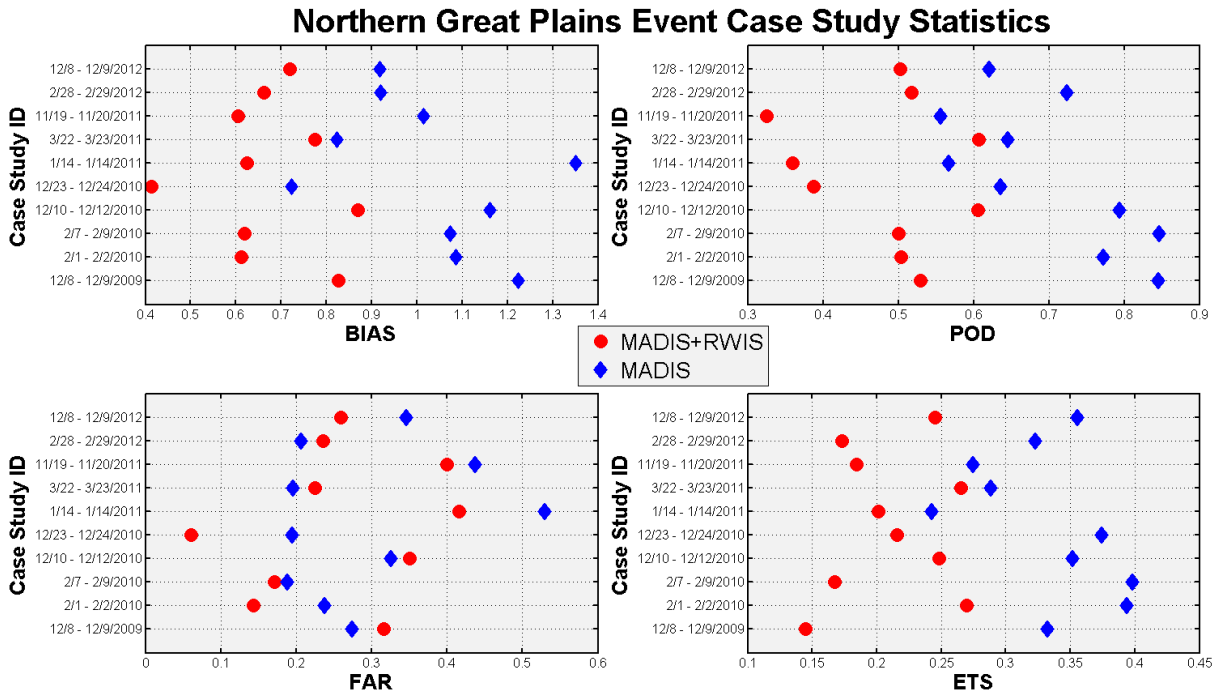

Figure 26. Event-average BIAS, POD, FAR, and ETS for Northern Great Plains domain using MADIS only and MADIS + Clarus data in surface analysis

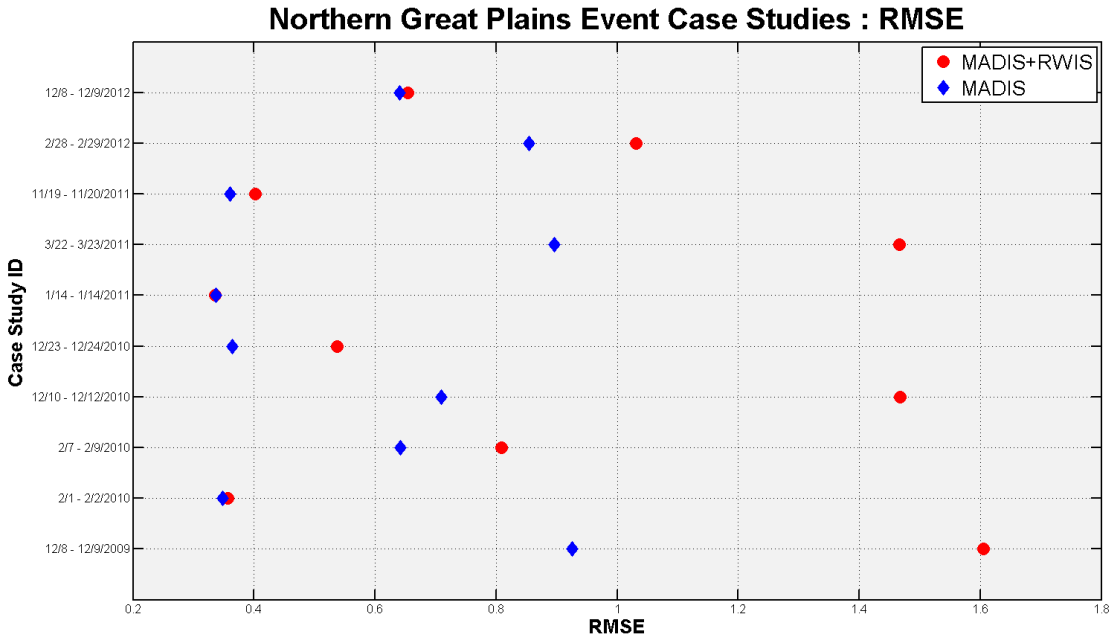

Figure 27. Event-average root mean squared error for Northern Great Plains domain using MADIS only and MADIS + Clarus data in surface analysis

5.2.2 Midwest Domain

Similar trends are seen in the event-average values of bias, FAR, POD, ETS, and RMSE for the Midwest domain (Figures 28 and 29). However, the spread between the two (MADIS only and MADIS+RWIS) is lower for all statistics. Bias results are mixed. The POD and ETS generally decreased across the board and FAR and RMSE increased.

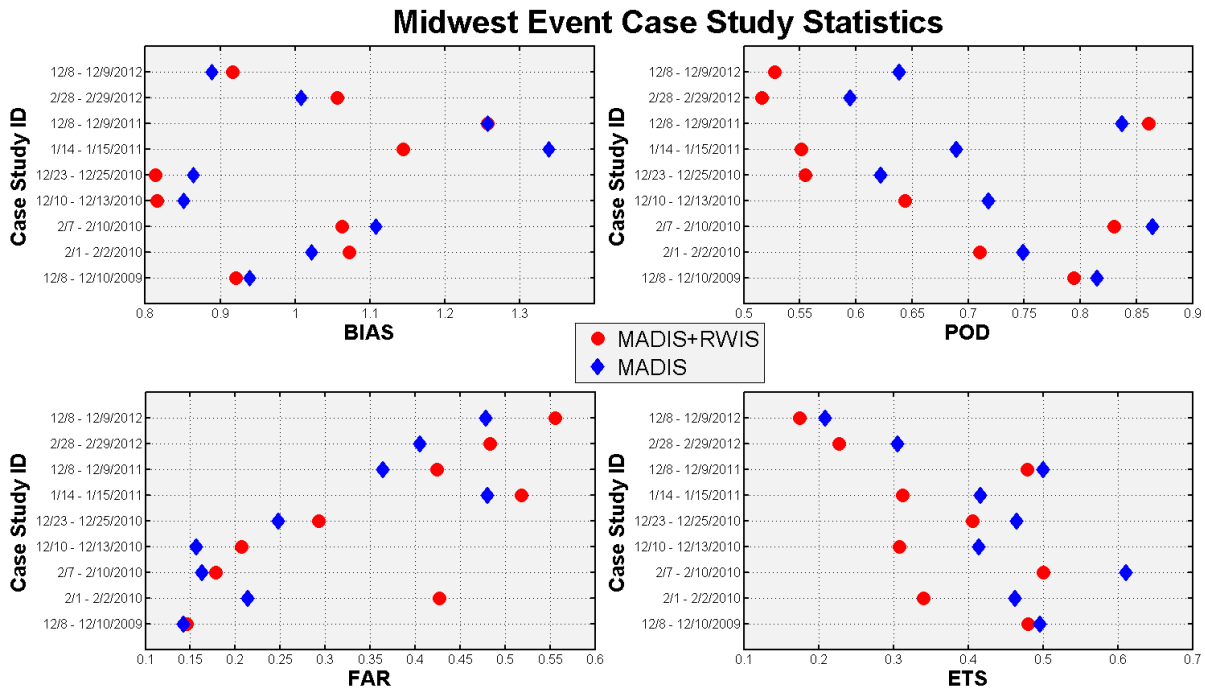

Figure 28. Event-average BIAS, POD, FAR, and ETS for Midwest domain using MADIS only and MADIS + Clarus data in surface analysis

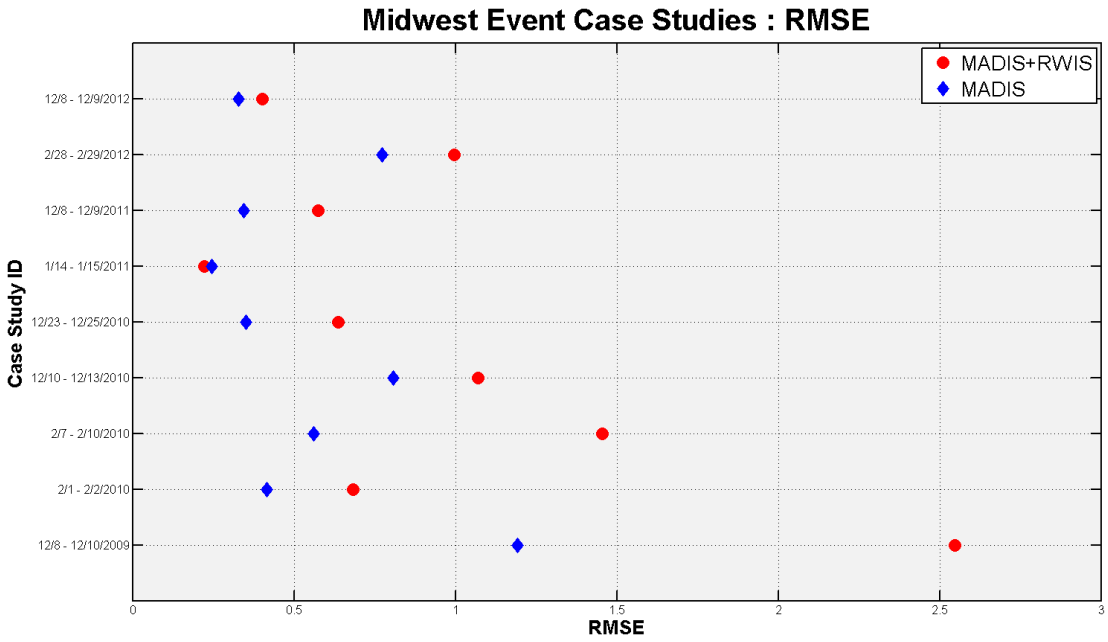

Figure 29. Event-average root mean squared error for Midwest domain using MADIS only and MADIS + Clarus data in surface analysis

5.2.3 Utah Domain

The same type of analysis was performed for the Utah cases and the statistics are shown in Figures 30-31. Fewer cases are considered with this domain and each case involved relatively light snow and had a smaller spatial extent relative to cases considered with the other two domains. Results shown in Figure 30 are consistent with those from the MW and NGP domains, as inclusion of Clarus data does not necessarily improve results. However, the results are more mixed, with ETS increasing for some events and decreasing for others. Moreover, the spread in the results is less relative to the results for the NGP and MW domains. These differences may be a result of the types of cases analyzed in this region, as the Utah region cases were light snow events, while the NGP and MW cases were more significant events. These differences could also result from differences in ESS maintenance practices. Either way, further examination of inclusion of Clarus data is needed, including the best way to use, improve, and discard data from the PPAES surface analysis.

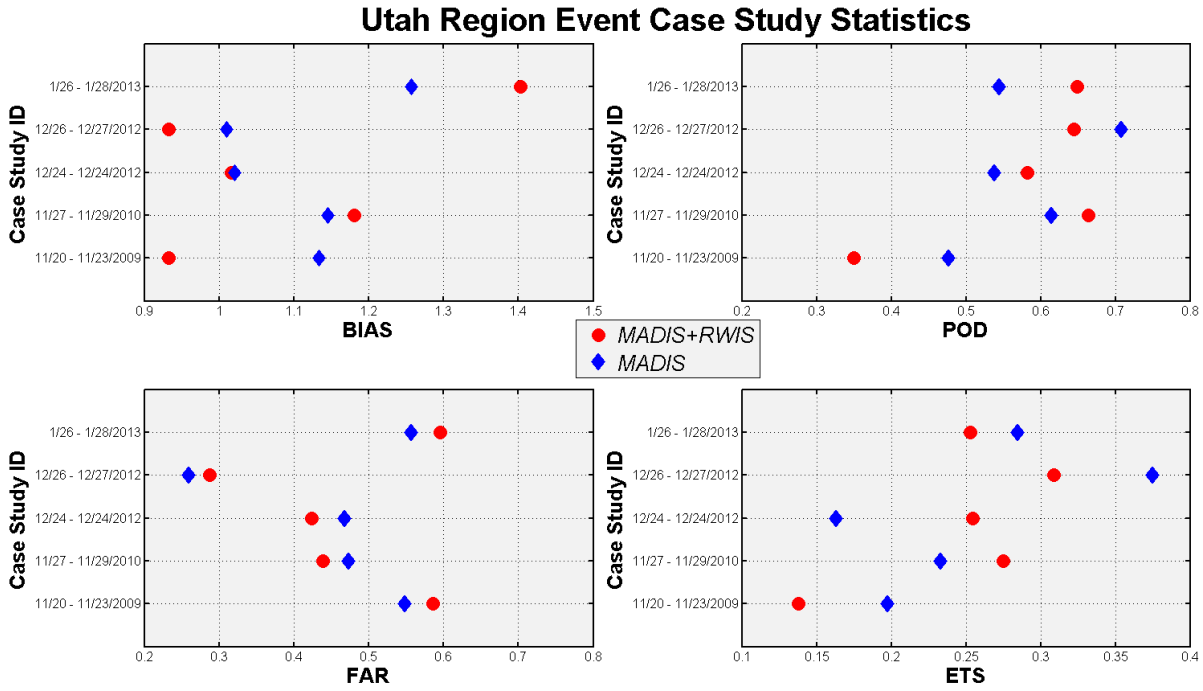

Figure 30. Event-average BIAS, POD, FAR, and ETS for Utah domain using MADIS only and MADIS + Clarus data in surface analysis

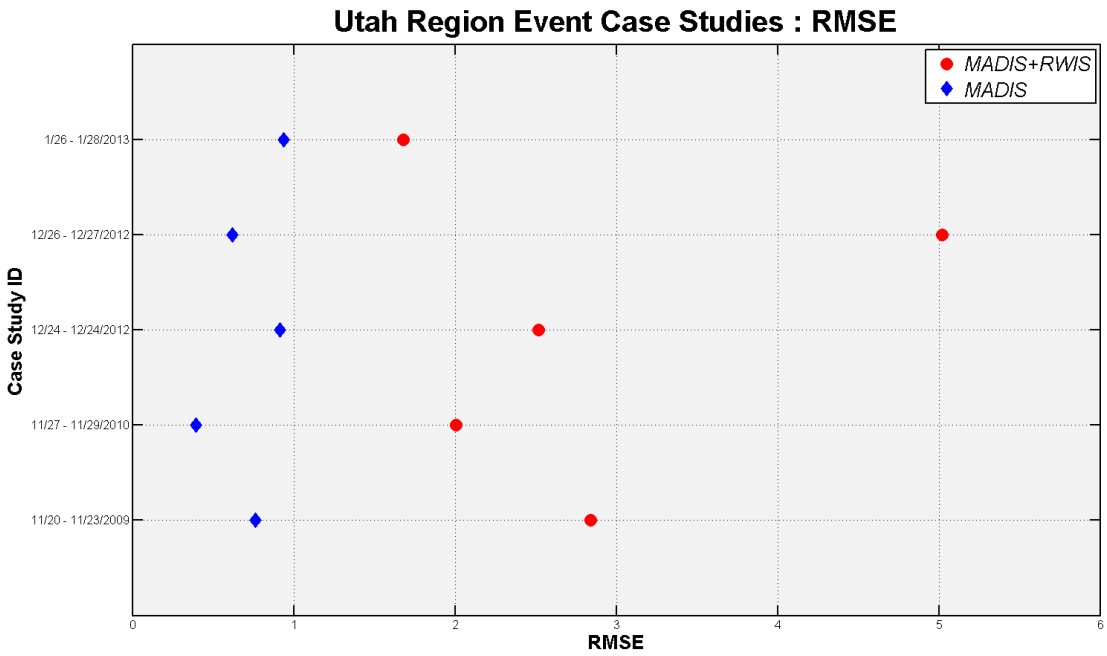

Figure 31. Event-average root mean squared error for Utah domain using MADIS only and MADIS + Clarus data in surface analysis

5.3 Complex Terrain Radar Data Analysis

In previous versions of PPAES, only radar data from the base scan elevation angle were used in the analysis. This causes problems in regions of terrain. One of the goals of this project was to add a complex terrain algorithm to the radar processing to aid in determining precipitation in such situations. The focus region for these tests was selected to be Utah region. This region has a good selection of radars, complex terrain, and an acceptable surface station network to provide truth data.

As is indicated in Figures 32 and 33, the results are mixed, with the addition of terrain clearance sometimes improving performance from a BIAS (two cases), POD (three cases), FAR (three cases), and RMSE (two cases) standpoint, but decreased performance for four of the five cases from an ETS standpoint. The latter is interesting considering that for the 12/24/2012 case the average POD increased and the average FAR decreased, which seemingly should produce, on average, a higher ETS. This suggests that outliers may be significantly affecting the average results.

To explore whether outliers significantly affected the results, event-median values are plotted in Figures 34 and 35. As indicated by these figures, when median values are considered, the performance with terrain clearance is generally as good as or better than without terrain clearance, with the exception being FAR. With terrain clearance, for example, ETS values are higher for four of the five events. Furthermore, terrain-clearance RMSE values are smaller than non-terrain-clearance values for three of the five events. This suggests that outlier values are having significant impacts.

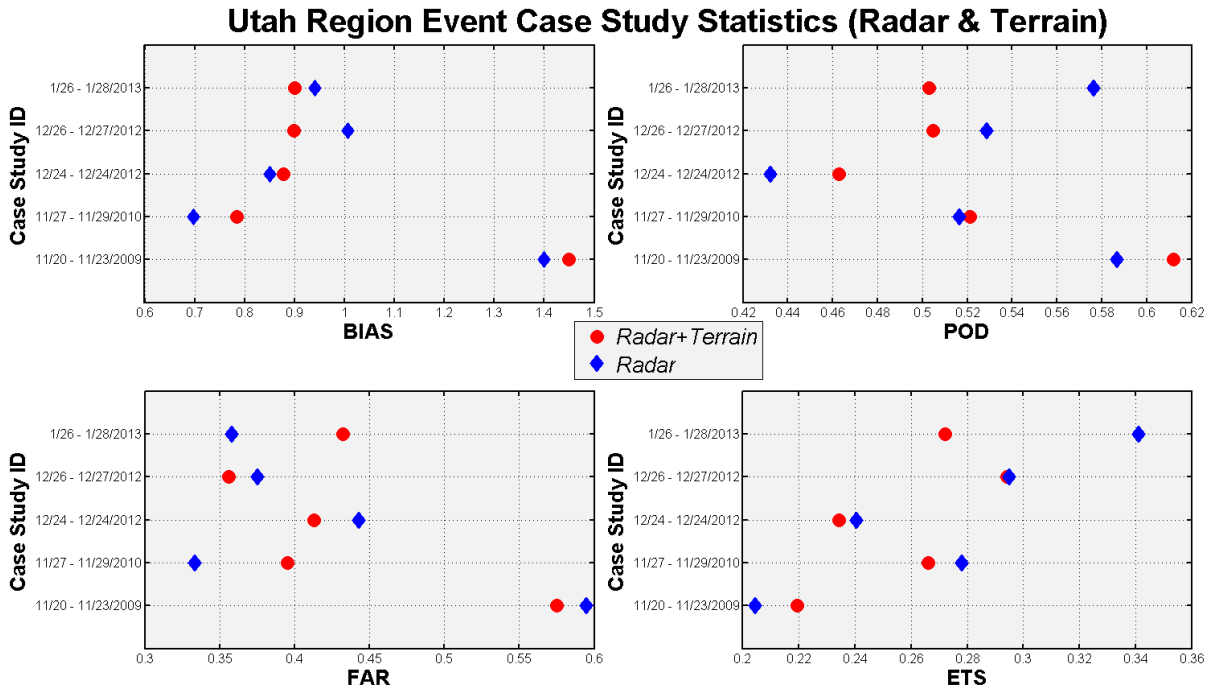

Figure 32. Event-average BIAS, POD, FAR, and ETS for Utah domain using Radar only and Radar + Terrain Clearance in radar analysis

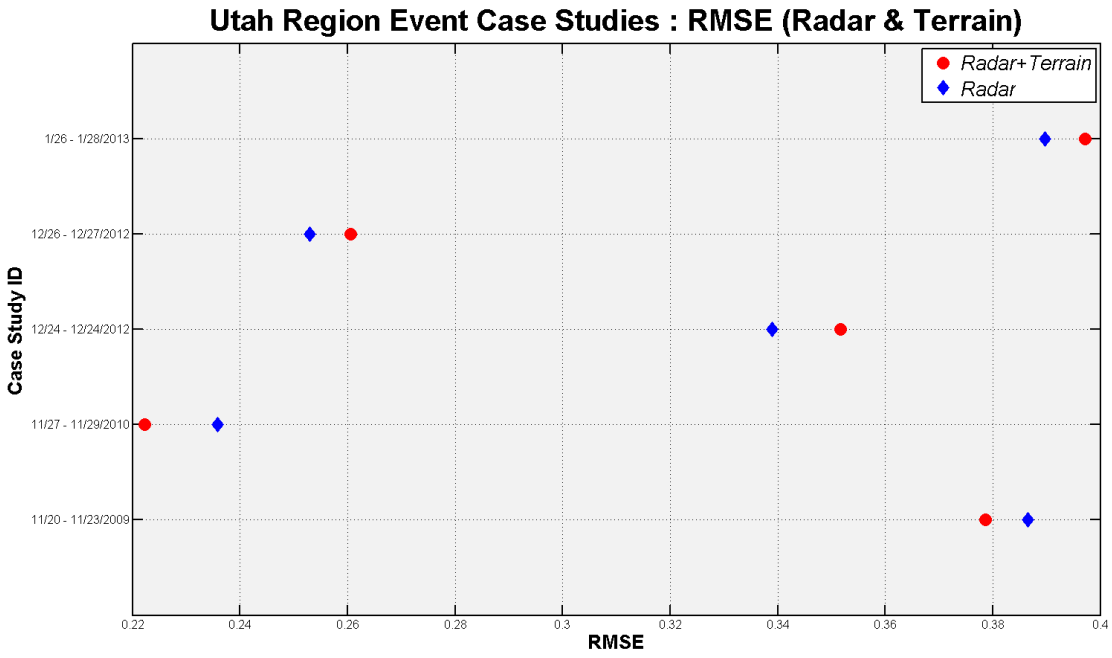

Figure 33. Event-average root mean squared error for Utah domain using Radar only and Radar + Terrain Clearance in radar analysis

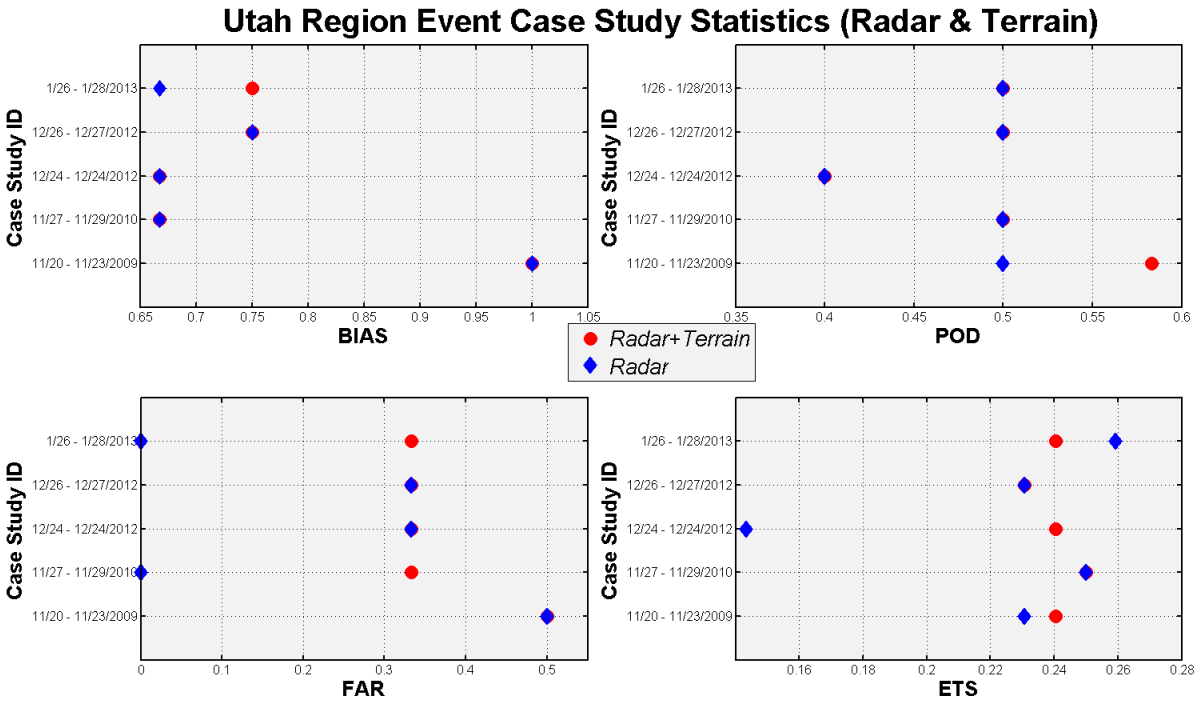

Figure 34. Event-median BIAS, POD, FAR, and ETS for Utah domain using Radar only and Radar + Terrain Clearance in radar analysis

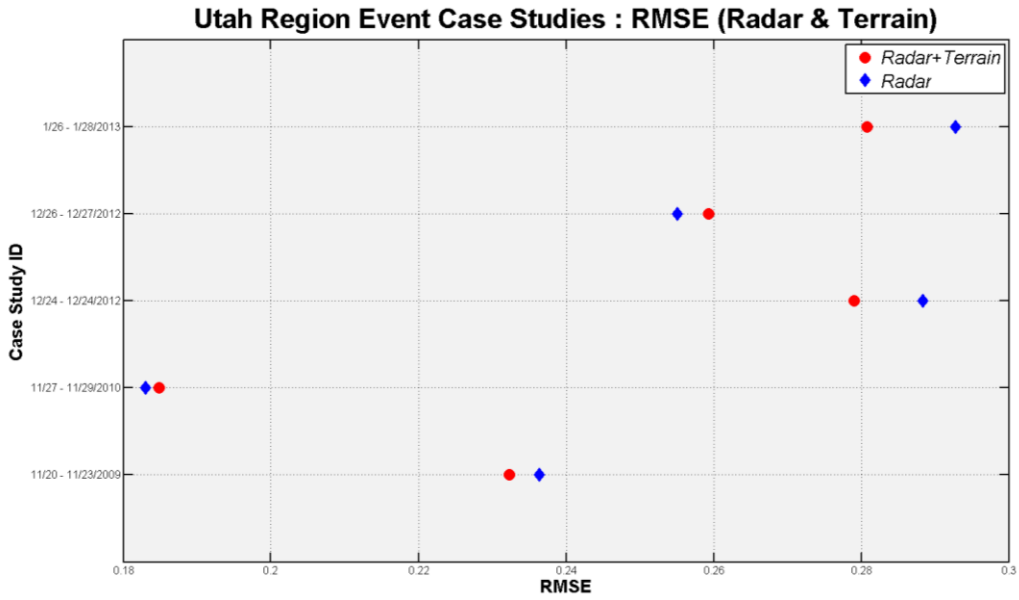

Figure 35. Event-median root mean squared error for Utah domain using Radar only and Radar + Terrain Clearance in radar analysis

In Figure 34, median values of zero arise owing to distributions that contain fewer non-zero values than zero values.

5.4 Rapid Refresh Model Data Analysis

Because the development of the model blending routine is at a much earlier stage, tests were conducted using a smaller number of events (two instead of five). Despite this, the results are consistent.

As indicated by Figures 36 and 37, generally the radar analyses performance was best, followed by radar+model, and finally model. The lone exception is POD, for which model and radar+model performance was significantly better than that of radar alone. However, the increased POD came at the price of increased FAR, thus reducing model and radar+model performance relative to that obtained using radar data alone.

This behavior is somewhat expected, as the grid spacing of the RAP (13.5 km horizontally), likely results in fine-scale precipitation systems being "smeared-out" and thus appearing as if they have cover larger areas than they actually do. One might expect that this could result from phase (positioning) errors. If this were the case, however, then POD should not be as incredibly high as it is since positioning errors should produce a significant number of misses.

It is noted that the RMSE results do not necessarily indicate that the RAP has a wet bias. The higher RMSE values that occur when model data are utilized may result from situations where no precipitation was observed but the model indicated precipitation occurrence. This may be associated with a smearing effect, with the integrated amount of RAP-produced precipitation being consistent with observed amounts.

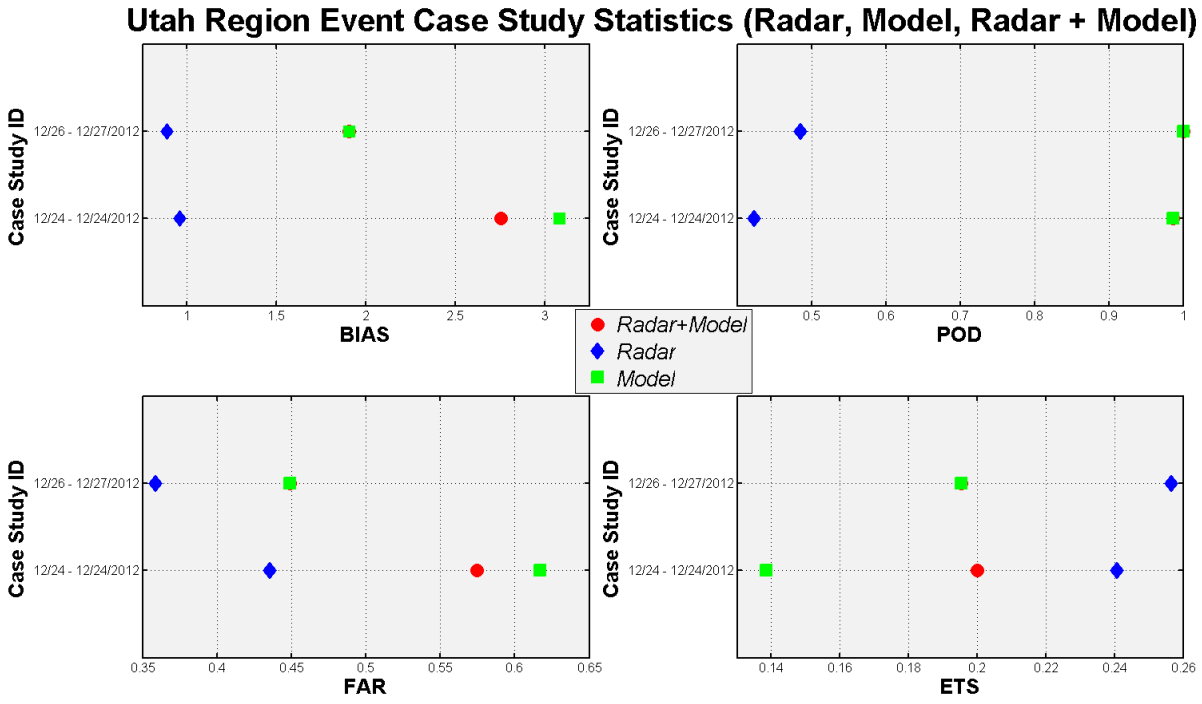

Figure 36. Event-average BIAS, POD, FAR, and ETS for Utah domain using Radar only, Model only, and Radar + Model data in analysis

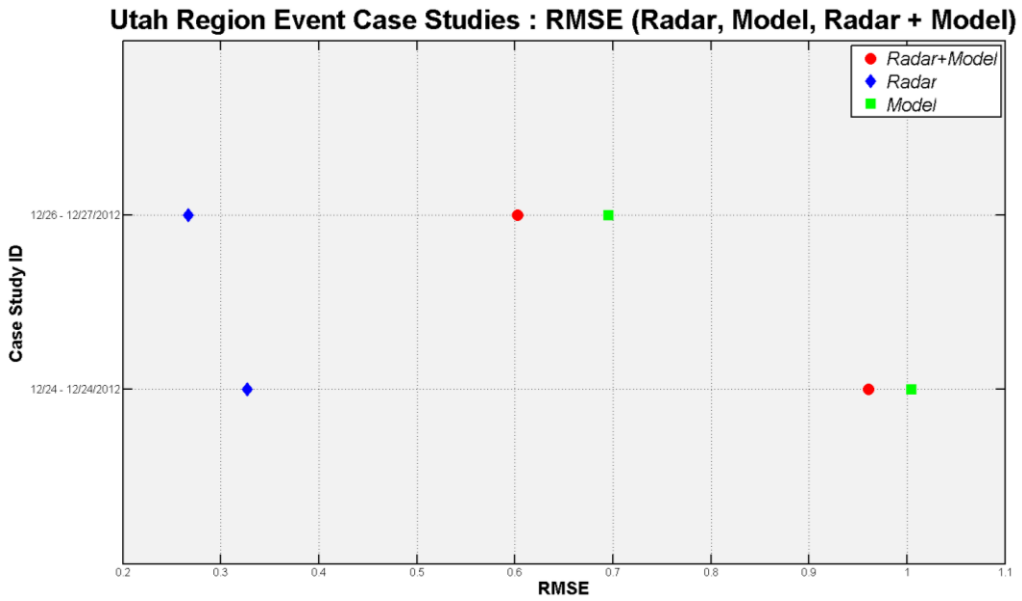

Figure 37. Event-average root mean squared error for Utah domain using Radar only, Model only, and Radar + Model data in analysis

6 DISCUSSION

6.1 Initial Analysis

6.1.1 Case Summary Statistics

The general event-average trends for both domains were:

1) RMSE and Mean Error values were lowest for the surface module.
2) Accuracy scores were highest for the surface module.
3) FAR and POFD scores were lowest for the surface module.
4) POD scores were highest for the blending module.
5) Case-computed ETS scores were highest for the surface module.
6) PPAES precipitation rates were generally underestimated.

In this section we discuss the likely reasons behind these statistical results.

Event-average RMSE values were lower in the surface module in 11 out of 18 cases in the Northern Great Plains domain (Figure 11) and in 14 out of 19 cases in the Midwest domain (Figure 19). This is not surprising as surface observations from MADIS were used to verify the precipitation rate analyses. Because the surface analyses use fixed precipitation rates derived from present weather, surface analysis estimates will inherently be closer to these observed values. Similarly, in Figures 10 and 18, the surface module was observed to perform better than the other modules in the majority of cases with respect to mean error. Insight into the module performance can be gained through examining how RMSE and mean error are computed.

Mean error does not consider the magnitude of individual errors while RMSE does, giving greater weight to larger errors through the squared operator. From examining the event-average RMSEs, we must have larger individual errors present in the blended and radar analyses than in the surface analyses. Furthermore, the difference in module performance with regard to event-average mean error and event-average RMSE can be attributed to the fact that the blending module precipitation rates are derived from both surface and radar analyses--where values may be adjusted or not adjusted. This suggests that the blending module inherits some of the radar and surface module analysis errors.

Figures 11 and 19 show that the largest average number of estimated errors occurred within the radar module. However, the radar module would be expected to contribute more errors based upon the radar module event-average RMSEs, with higher event-average RMSEs in 33 out of 37 events (18 and 19 events in the Northern Plains and Midwest domains, respectively) and an event-average RMSE of approximately 0.40 mm hr^{-1}, as compared to the surface module average of 0.34 mm hr^{-1}. Further insight as to where errors arise in the blended analyses may be gleaned from the other performance metrics.

Both ETS and TS indicate overall performance and indicate how well the estimated precipitation occurrences corresponded to observed precipitation occurrences. Both ETS and TS take into account hits, misses, and false alarms, with ETS also taking into account hits due to random

chance. Although these skill scores do not distinguish where the source of "error" comes from, as they both utilize false alarms and miss information, average values of POD and FAR can be examined to gain an understanding of ETS and TS. This is because a high POD and a low FAR result in high ETS or TS and vice-versa.

As noted previously, the surface module outperformed the radar module with respect to FAR for every case. The surface module outperformed the radar module with respect to POD for 10 out of 18 cases for the Northern Plains domain, while the radar module outperformed the surface module in the Midwest domain for 15 out of 19 cases. Because the blended analyses are derived from both the surface- and radar- modules, it appears as if the majority of the false alarms are originating from the radar module. Figures 13 and 21 support this contention. Because the surface and radar modules are similar with regard to event-average POD performance, false alarms appear to primarily influence the ETS or TS. Therefore, the false alarms from the radar analysis were influencing the blended analyses more often than false alarms from the surface analyses. With the radar module always having a higher event-average FAR, the subsequent blended analyses ETSs (case-computed) were usually intermediate between the radar and the surface module (as seen in 31 out of a total of 37 cases).

Similar arguments can be applied to the TS performance metric because TS is impacted similarly to ETS by high/low POD and FAR scores. As such, the primary effect contributing to the blended analysis TS values being less than those for surface analysis TS values is the higher rate of false alarms associated with the radar analyses.

6.1.2 Equitability and Hedging as Pertains to Performance Metrics

Many performance metrics are available to evaluate skill in a forecast, where skill is defined as the relative accuracy of the PPAES modules compared to some reference forecast such as persistence or random chance (e.g., Jolliffe and Stephenson 2004). Each metric has positive and negative attributes. Two important factors to consider when selecting performance metrics are equitability and the capability to limit hedging (Jolliffe and Stephenson 2004; Hogan et al. 2009).

Equitability is the characteristic where all random estimates and all constant estimates of the same category in a categorical statistic such as the ETS produce the same expected score (e.g., Jolliffe and Stephenson 2004). Otherwise, a random system could be incorrectly ranked and placed ahead of a system with some skill.

Hedging is a process by which a forecast that differs from ones "true belief" is given in order to improve a performance metric score (Hogan et al. 2009). In this study, hedging would be associated with estimates that are essentially tuned to get the best score for a specific metric. Rewarding hedging would be undesired, as the metric would then really not be expressing skill. When such tuning/hedging is present in a forecasting or estimating system, that system may perform worse when scored using other metrics (e.g., Hogan et al. 2009; Hogan et al. 2010).

In the context of PPAES, the ETS allows one to see how well the estimated precipitation occurrence events correspond to observed precipitation occurrence events. In the literature, ETS

has been noted to be "generally" equitable (e.g., Hogan et al. 2010), which allows one to assess PPAES module performance relative to a system that produces random estimates. However, Hogan et al. (2010) recently provided evidence that ETS is not equitable in the sense that all constant estimates of precipitation occurrence/non-occurrence and random estimates would not give the same score: zero (Hogan et al. 2010). As such, the applicability of using ETSs in this study is questionable.

Ultimately, Hogan et al. (2010) noted that ETS was asymptotically equitable, which means that ETS is only equitable provided a large sample of estimates (> 30) or forecasts exist. Recall that ETSs were computed two ways in our validation: (1) for each analysis hour and then averaged over the duration of a case; and (2) utilizing the summed totals of all of the combinations from the contingency table over the duration of each case. Method (1) will not ever pass the constraint set by Hogan et al. (2010) because only 15 surface observation stations are withheld for verification and, thus, does not satisfy the need for a "large sample of stations". Therefore, using method (1), ETS associates PPAES analyses with constant estimates of the same category (e.g., all precipitation or no precipitation occurring at all verification stations) with no skill and the usefulness of the Method (1) ETS is questionable.

A performance metric like TS would be more advantageous for PPAES validation *if* one uses Method (1) to compute TS. Although TS is also not equitable and does not account for hits due to chance, TS does not suffer from the issues associated with ETS. Alternatively, computing ETS with Method (2) is the preferred approach, as the requirement of $n > 30$ is satisfied. However, with Method (2) no information on performance changes over time can be gained. TS or other performance metrics can be used to gain insight into temporal changes in performance, as they do not require a minimum sample size.

6.1.3 Significance Testing

Significance testing with α set to the 0.05 level was completed to determine whether skill score differences between the modules were statistically significant. Each case was assumed independent of the others and, where necessary, the event-average metric values from the Northern Plains and the Midwest domains were averaged together for the same cases. Therefore, 19 events in total were used for significance testing. Significance testing was completed for the event-average POD, accuracy, FAR, and ETS (case) using the two-sample *t*-test and the one-way ANOVA.

For the one-way ANOVA tests, the hypotheses were that all of the means are equal (null) and not all of the means are equal (alternative). The ANOVA results allow us to reject, at a significance level of 0.05, the null hypothesis that the POD, accuracy, FAR, and case ETS statistics for all three modules are equal (Table 11). However, the one-way ANOVA test does not provide any information regarding the relationships between the populations associated with the three PPAES modules (Weiss 2008). Therefore, the two-sample *t*-test (with equal variances assumed) was used to learn more about the performance of the different modules.

One-tailed and two-tailed two-sample *t*-tests were executed for the same dataset. The hypotheses were that the population means are equal (null) and that the population means are different (two-tail) or greater than/less than (one-tail) the other population mean, respectively. One-tailed tests were run to determine the direction of population mean differences. Because only two samples can be used in a test at a time, three tests were needed to cover all PPAES module combinations (surface vs. radar, surface vs. blended, and blended vs. radar).

Three of the tests failed to provide evidence for rejection of the null hypothesis for POD, accuracy, and FAR (Table 12). The three tests that failed to provide evidence were: radar and surface for POD, radar and blending for accuracy, and radar and blending for FAR. For those three tests, the study cannot say that one module performed better or worse than the other module with confidence at the chosen significance level α.

Regarding the two-sample *t*-tests, most of the results were statistically significant at the α=.05 level. As such, this supports several assertions laid forth earlier in this section:

1) For POD, the blending module performed best.
2) For accuracy, the surface module performed best.
3) For FAR, the surface module performed best.
4) For ETS, the surface module performed best.

However, it is not possible to distinguish between those three aforementioned cases where the null hypothesis could not be rejected. Furthermore, as with the one-way ANOVA tests, these results depend upon none of the four assumptions being violated significantly (Weiss 2008).

When stratifying the cases by event type or precipitation intensity, the sample sizes are all very small ($\leq$10). This limits the robustness of any significance testing as the results could be a result of chance instead of clear differences in performance between the individual modules. Because of this, the event-average performance metrics are emphasized hereafter rather than the stratification groups (event type and predominant precipitation intensity).

6.1.4 Case Study: 1-3 December 2007 (Northern Great Plains)

In this case, several PPAES- and wintertime precipitation-related issues are seen, relating to the two platforms used in this study – radar and surface observations. These issues include but are not limited to: radar overshooting (e.g., Smith et al. 1996; Super and Holroyd 1998), falsely detecting virga (e.g., Fulton et al. 1998; Hunter et al. 2001), blowing snow conditions impacting the surface measurements (NWS 1998), light precipitation failing to be detected at the surface (NWS 1998; Wade 2003), and irregular observation station distribution (Baer and Tribbia 1976; Doswell and Lasher-Trapp 1997) impacting the blending algorithm.

This case also includes significant radar and surface analysis precipitation rate discrepancies that have also been observed in other historical cases. Specifically, an unusually large discrepancy exists between the surface and radar analysis, resulting in a large adjustment for radar site KMVX (+ 0.66 mm hr^{-1}). Such a large adjustment can impact the overall blended analysis, both

performance-wise and aesthetically. However, from an aesthetic standpoint, the greatest issue involved radar site KBIS, and was associated with observation station density. We will discuss this in more detail later in this section.

The 1-3 December 2007 is used for two reasons: (1) the case was a significant heavy snow event that also featured periods of light precipitation, and (2) the case includes many of the limitations associated with the various PPAES modules and the instrumentation used to observe wintertime precipitation. The case provides insight into the problems encountered with heavy and light precipitation that are important to the road maintenance community (e.g., Mahoney 2003). Specifically, light snowfall (rates around 0.1 mm hr^{-1}) affects how the road weather communities maintain and treat roads because of the difficulty in such amounts being predicted or detected (Mahoney 2003).

We focus the analysis at 21:55:00 UTC on 12/01/2007. At this hour, a large precipitation shield covered much of the Northern Plains domain. Note that this particular analysis hour did not share all of the typical performance metric characteristics previously discussed. Specifically, the radar module had lower FAR and POFD values than the surface and the blended modules, and the surface module had a false alarm where the radar module did not for this analysis hour. The performance of the blending module was seen to be either intermediate or worse relative to the other modules. The performance of the radar module, with regard to false alarms, was different from what was seen in the event-average results and discussed above. Nonetheless, this analysis is valuable as it demonstrates the capabilities of PPAES and highlights limitations.

A station blacklist was used for this particular analysis; such a blacklist contains a list of stations for which precipitation observations have been determined to be incorrect. The blacklist used here only contains one station –KFGN, but still illustrates the concept which we will include as a standard feature in the next version of PPAES. Station KFGN was blacklisted because it frequently reported rain during this case (15, 25, and 10 occurrences of +RA, RA, -RA, respectively, vs. only seven occurrences of -SN), which was deemed to be very unlikely given neighboring METAR present weather reports plus an observed maximum temperature over the course of the event of -8 °C.

6.1.4.1 Case Background

The 1-3 December 2007 case featured a cyclonic (low pressure) system (Figure 38) that produced precipitation within the Northern Plains domain from ~ 0500 UTC 12/01/2007 to 0600 UTC 12/03/2007. Pertinent synoptic features will now be briefly discussed using NOAA Hydrometeorological Prediction Center (Figure 38) and Rapid Update Cycle version 2 (hereafter denoted as RUC, Figures 39-41) analyses (e.g., Benjamin et al. 1998) for 2200 UTC 1 December 2007. The RUC provides high temporal frequency (hourly) analyses and incorporates an hourly data assimilation cycle utilizing many observation types (Benjamin et al. 1998).

The long wave pattern at the time of the analysis featured a trough/ridge to the west/east of the Northern Plains domain, respectively (Figure 38). The jet stream is positioned within the long wave pattern with two jet streaks evident over Nebraska/Kansas and Wisconsin. A short wave

trough at 500 hPa extended from South Dakota to Kansas; this trough included a vorticity lobe over South Dakota, with cyclonic vorticity advection occurring downstream into Minnesota (Figure 36). At the surface, the center of the cyclone was directly south of the Northern Plains domain at the analysis time, with a low pressure trough extending into northwestern Minnesota (Figures 38, 41).

The cyclonic system produced large snowfall amounts across central and southern MN with more moderate amounts elsewhere in the Northern Plains. The Red River Valley received 3" – 6" and other locations within the domain, including Fargo, Grand Forks (both ND), Aberdeen, Sisseton (both SD), and Chanhassen and St. Cloud (both MN) received 6-9" of snowfall (NWS 2011a; NWS 2011b; NWS 2011c). The heaviest snow moved out of the domain by approximately 05 UTC 12/02/2007. After this time, light to moderate snowfall persisted in localized bands until the end of the event. With respect to the event classifications developed earlier, this event is a low pressure system snow event with predominantly heavy precipitation.

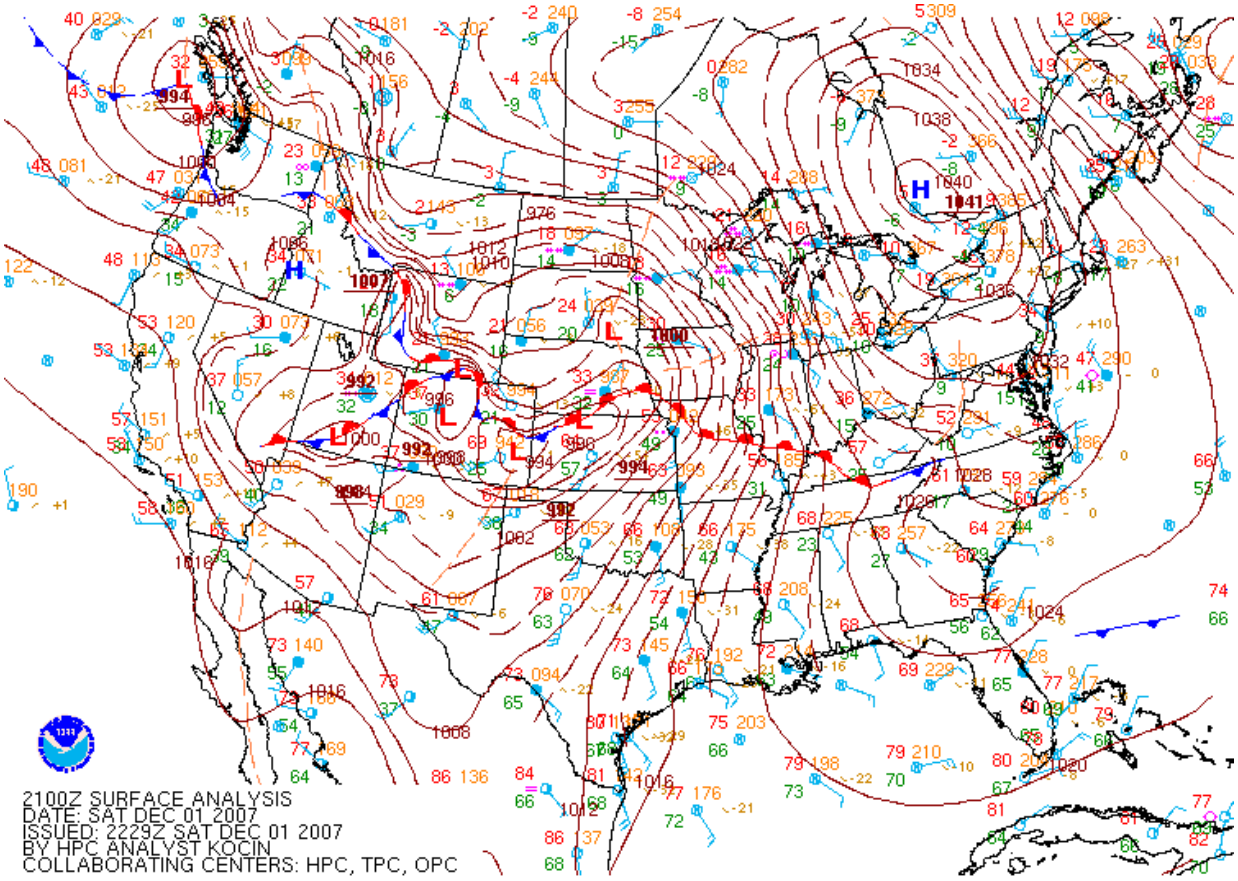

Figure 38. Surface analysis from Hydrometeorological Prediction Center at 2100 UTC December 1, 2007

For verification, the station denial scheme detailed in Section 4 was used, with seven stations randomly chosen and eight stations chosen subjectively. As discussed at the end of this section, performance measures may be sensitive to the way in that the withheld stations were selected

(random or manually, and this fact should be kept in mind when interpreting the statistical results).

With the exception of surface stations K9V9 and KD07 (the former did not report during the course of the event and the latter's present weather sensor was inoperable), all stations were used in this analysis. The nature of the surface observation station distribution varies considerably within the Northern Plains domain, as ND and SD stations are sparse while MN and IA stations are much more numerous (Figure 42).

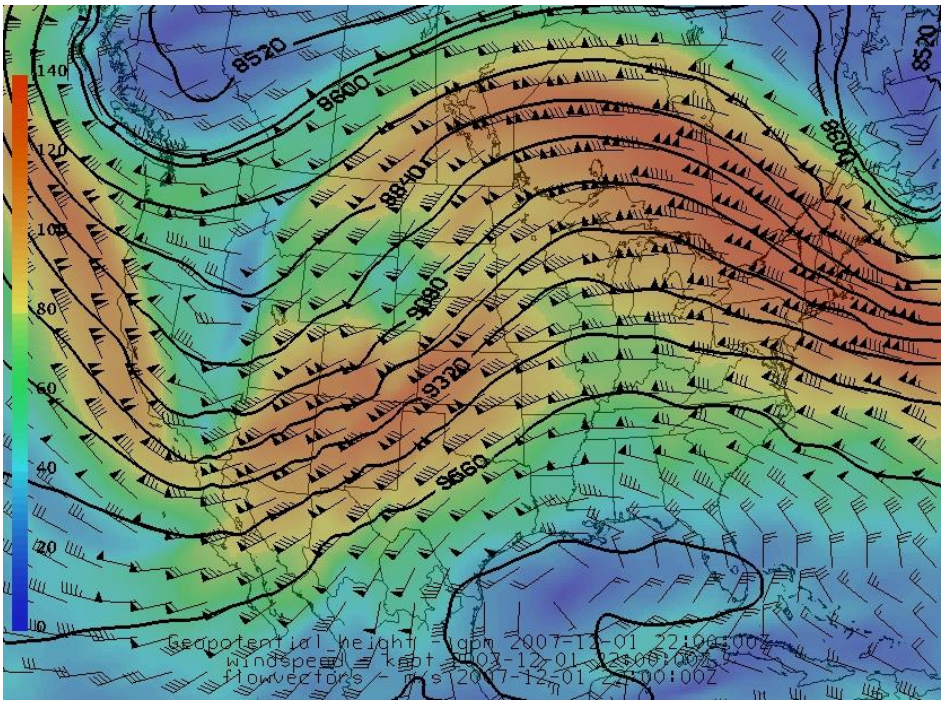

Figure 39. Geopotential heights (gpm), wind barbs (m s^{-1}), and wind speed (knots, color-filled) at 300 hPa at 2200 UTC December 1, 2007 from RUC-2 data assimilation system

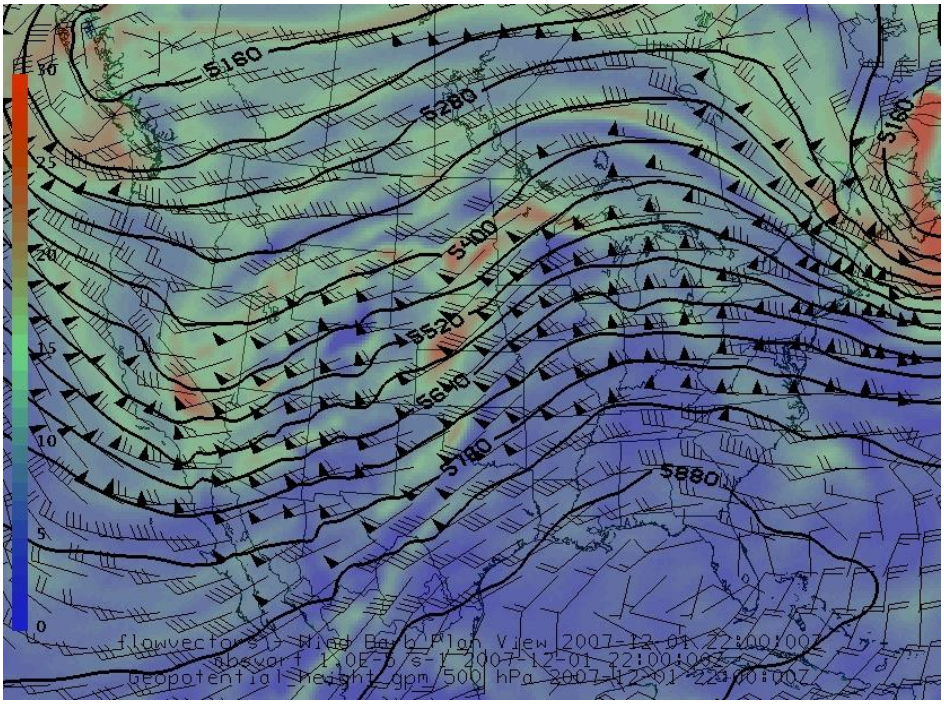

Figure 40. Geopotential height (gpm), absolute vorticity ($1.0x10^{-5}$), and wind barbs ($m\ s^{-1}$) at 500 hPa at 2200 UTC December 1, 2007 from RUC-2 data assimilation system

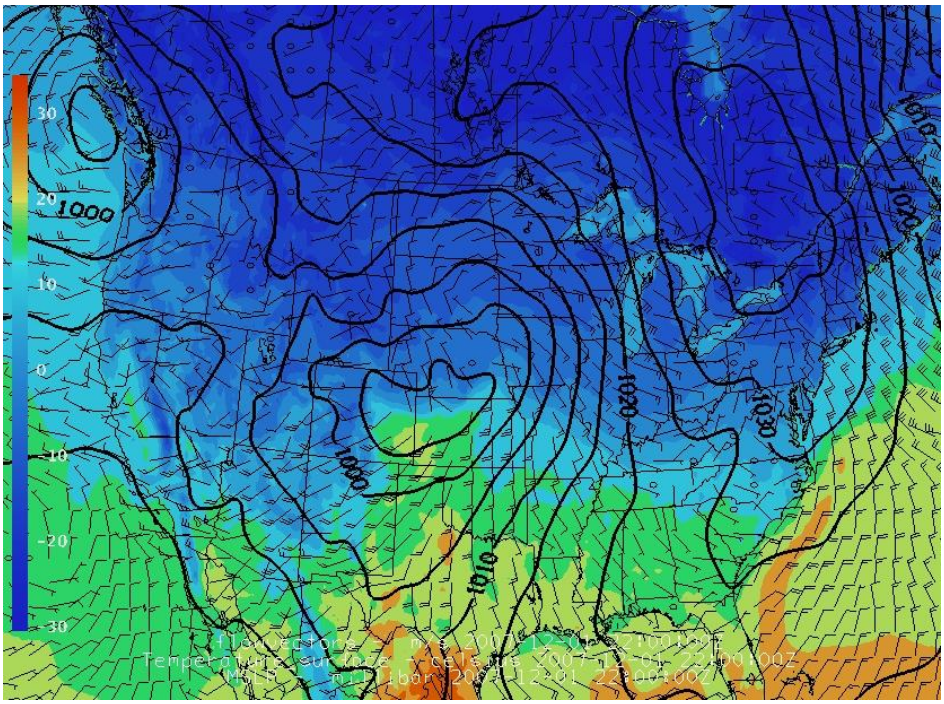

Figure 41. RUC mean sea level pressure (contours hPa), 10-m wind (barbs m/s), and 2-m temperature (colors °C) fields valid at 2200 UTC December 1, 2007

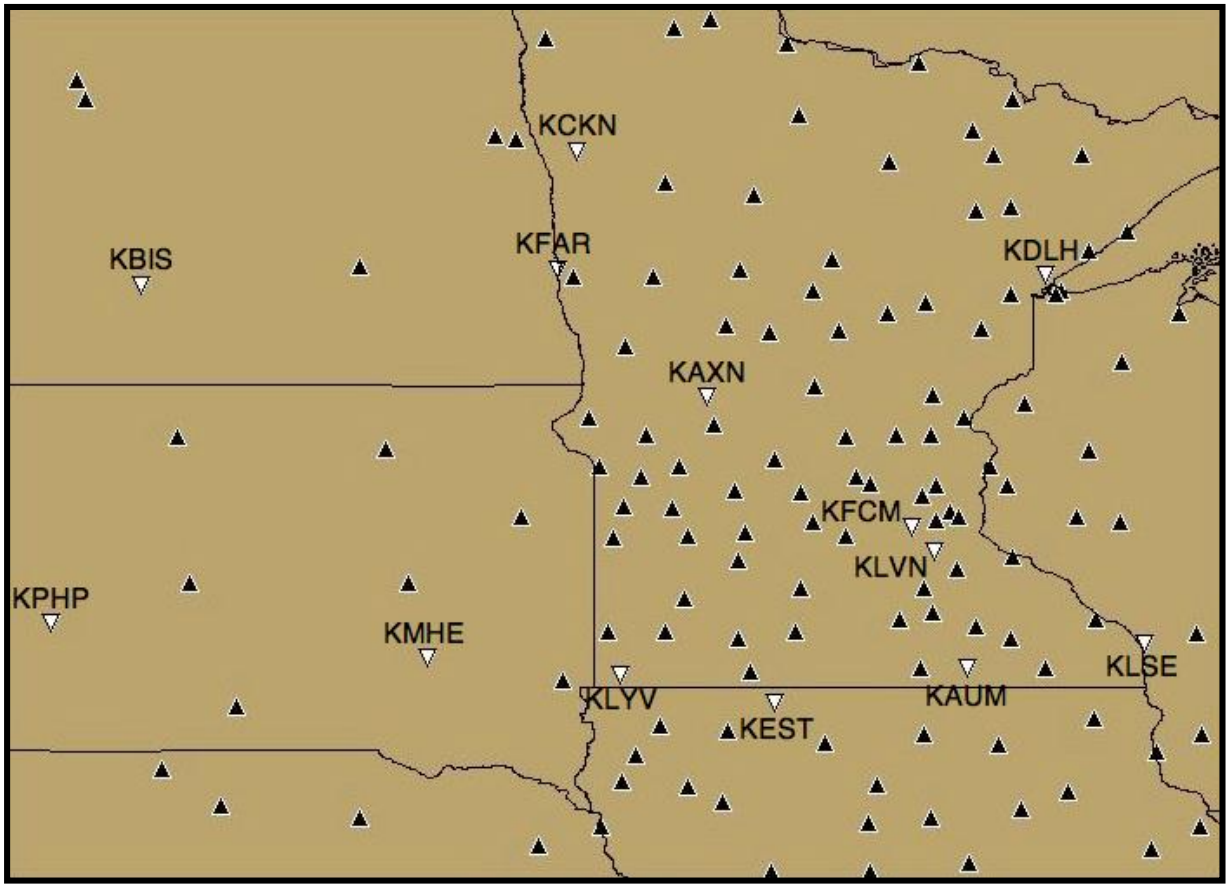

Figure 42. Northern Plains domain stations used and withheld from surface module verification stations for this case

In Figure 42, two stations, K9V9 and KD07, could not be used. The black triangles and 13 white inverted triangles in Figures 42, 43, and 44 denote surface stations used and withheld from the surface and blended analyses, respectively.

6.1.4.2 Analysis: 22 UTC December 1, 2007

The PPAES surface (Figure 43) and radar (Figure 44) module analyses for 21:55:00 UTC 1 December 2007 both depict a precipitation shield covering most of the domain; it is also apparent that the radar analysis has finer resolution than the surface analysis. This result is expected because the radar module uses Level II equivalent reflectivity from the S-band WSR-88D radars with horizontal data spacing of ~0.44 km at 25 km range and 2.62 km at 150 km range (Crum et al. 1993). Further, the radar module pseudo nearest neighbor objective analysis scheme preserves much of the fine detail in the radar dataset because the scheme assigns the value of the range resolution volume that is: a) closest to an analysis location and b) located within an analysis grid box associated with that location. In contrast, the surface module uses surface stations as its sole data source and the ability to resolve fine scale detail is limited given that the minimum surface station spacing is approximately 41 km in the domain. Also, the surface module uses a Barnes objective analysis scheme, which results in smoothing of fine detail (e.g., Barnes 1973; Koch et al. 1981).

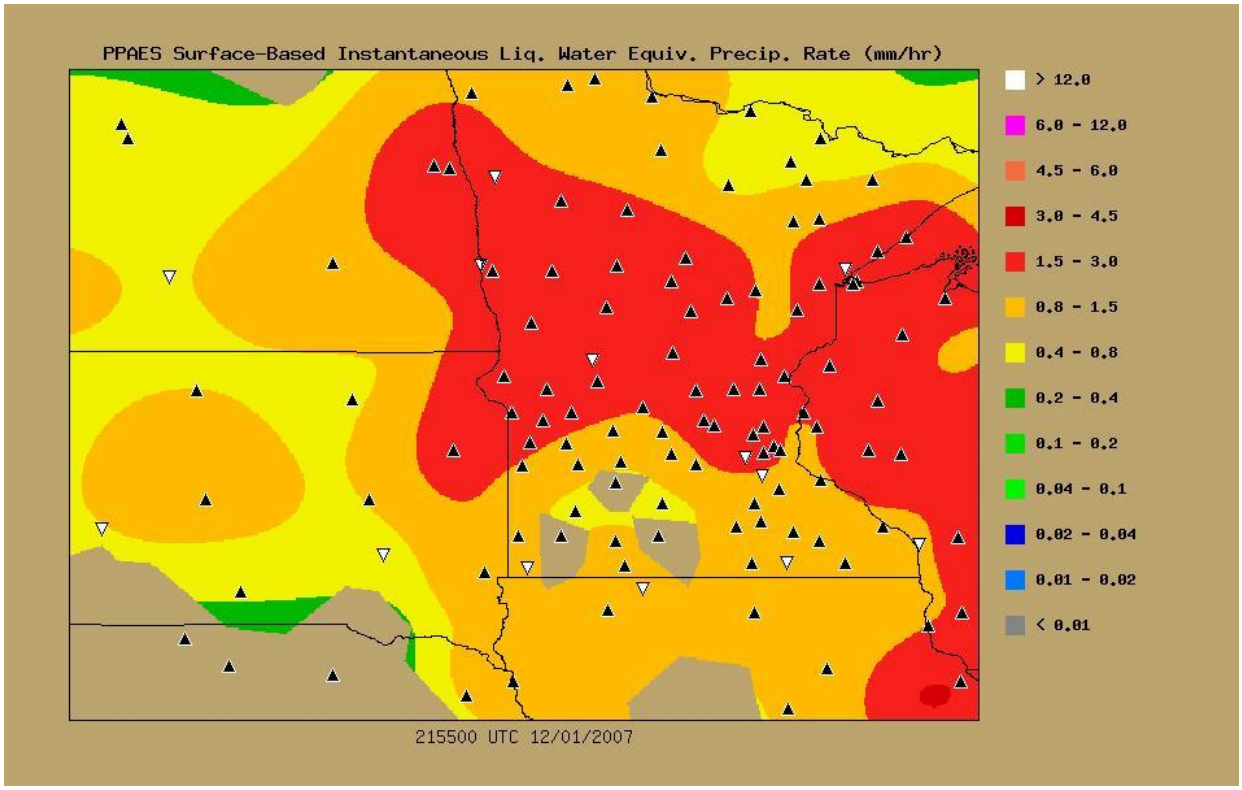

Figure 43. PPAES surface analysis at 21:55:00 UTC December 1, 2007 over Northern Plains domain

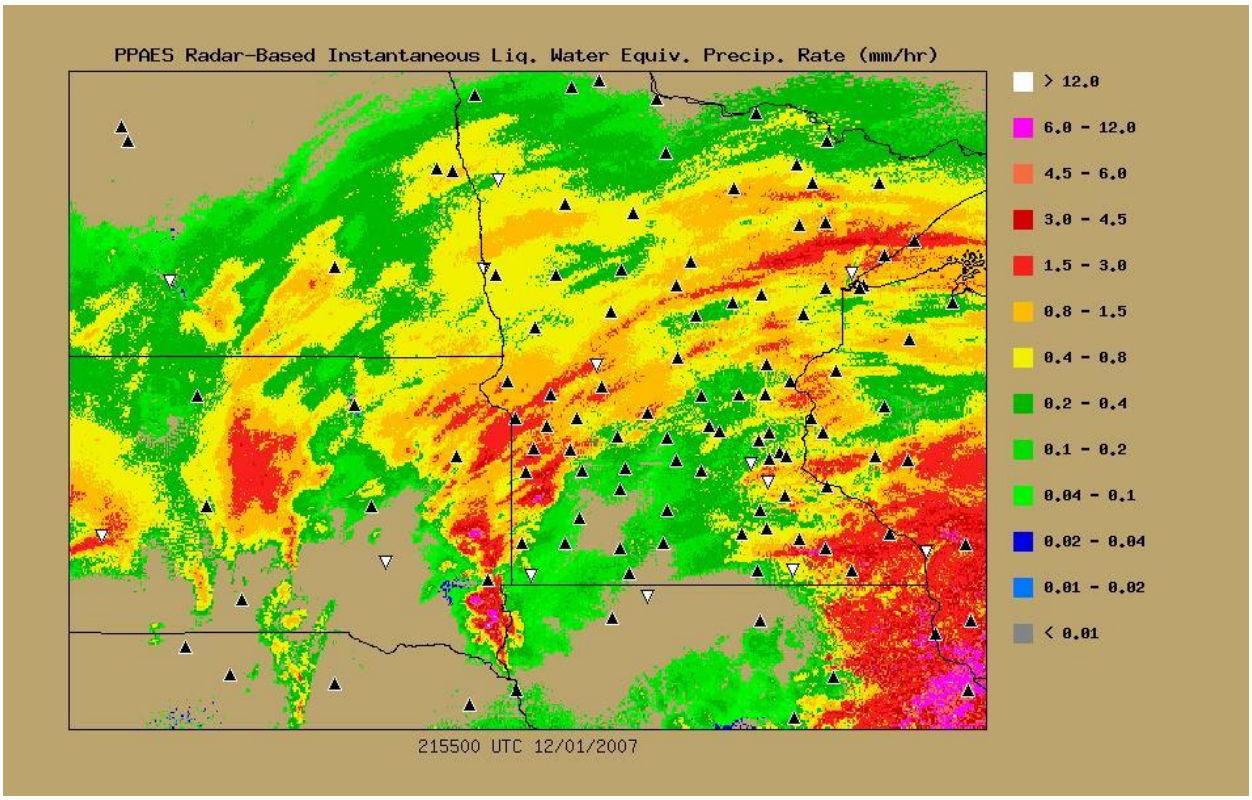

Figure 44. PPAES radar analysis at 21:55:00 UTC December 1, 2007 over Northern Great Plains domain

Precipitation occurrence/non-occurrence and precipitation rate performance metric scores are provided in Table 13. Bold values denote the module that performed the best for a particular metric. A perfect score for accuracy, bias, POD, TS, multiplicative bias, and correlation is one. A perfect score for FAR, POFD, mean error, MAE, MSE, and RMSE is zero.

Table 13. Precipitation occurrence/non-occurrence and precipitation rate performance metrics for the analyses at 21:55:00 UTC December 1, 2007

		Precipitation Occurrence/Non-occurrence Performance Metrics					
Module	Date and Time	Accuracy	Bias	POD	FAR	POFD	TS
Surface	22 UTC 12/01/2007	**0.92**	1.08	**1.00**	0.08	1.00	**0.92**
Radar	22 UTC 12/01/2007	**0.92**	0.92	0.92	**0.00**	**0.00**	**0.92**
Blending	22 UTC 12/01/2007	0.85	**1.00**	0.92	0.08	1.00	0.85

		Precipitation Rate Performance Metrics			
Module	Date and Time	Mean Error	RMSE	Multiplicative Bias	Correlation
Surface	22 UTC 12/01/2007	**-0.06**	**0.70**	**0.96**	**0.53**
Radar	22 UTC 12/01/2007	-0.79	1.12	0.45	0.40
Blending	22 UTC 12/01/2007	-0.43	0.95	0.70	0.27

From the analyses, several differences regarding precipitation occurrence and rate exist, including:

1) Precipitation occurrence differences owing to possible overshooting located near: Minot, ND; north central IA; the Lake of the Woods region in northwestern MN (Figure 44).
2) Discrepancies with regard to the location of precipitation rate maxima. Such maxima are located in the radar module analysis near Pierre, SD; east of radar KUDX; along the eastern border of SD; and along the eastern border of IA and WI (Figure 44).
3) Holes in the surface analysis where the radar analysis exhibits precipitation rates greater than 0.04 mm hr^{-1}. These occur near Slayton, MN (KDVP); Saint James, MN (KJYG); and Redwood Falls, MN (KRWF) (Figure 45).

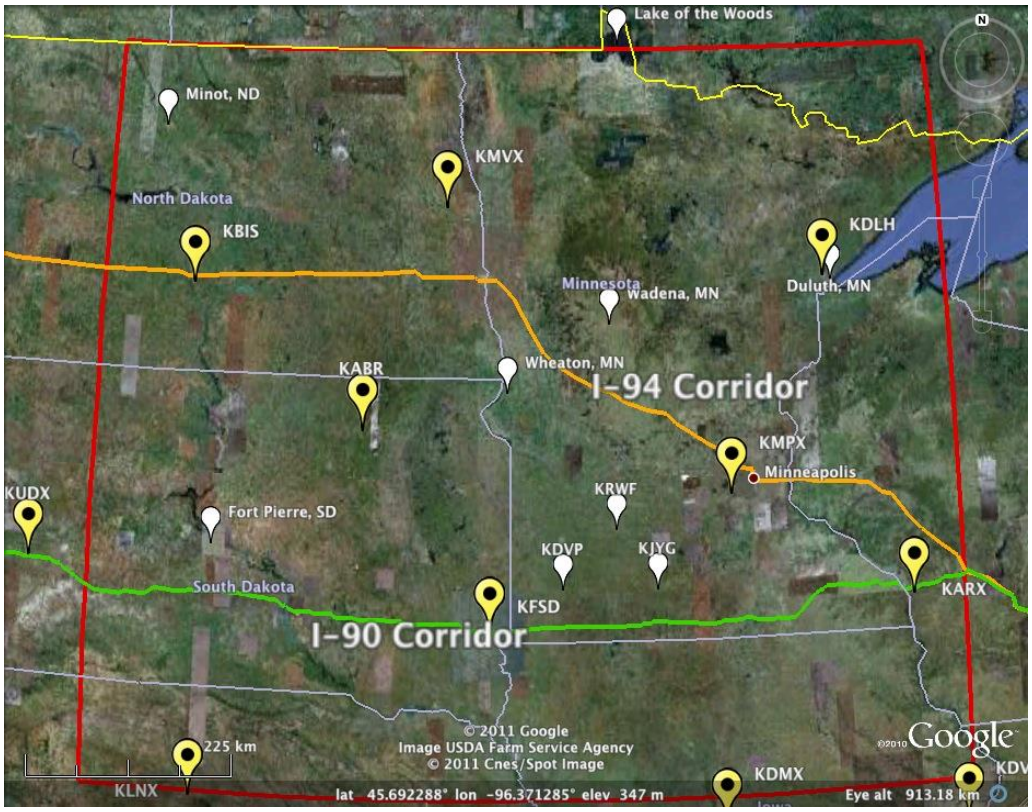

Figure 45. Locations in Northern Plains domain (red box) showing WSR-88D radar sites (yellow markers)

6.2 Overshooting Assessment

Differences between the radar and surface precipitation analyses can be caused by several factors. For instance, the effective radar coverage at 2 and 3 km above the surface is limited in several locations in the Northern Plains (Maddox et al. 2002, their Figures 5a and 5c). This problem is focused near the international border in north central ND and north central MN, where the effective coverage at 2 km above the surface is non-existent. An altitude of 2 km AGL, assuming standard atmospheric refractive conditions, corresponds to a range of ~120 km. Thus, shallow precipitation with a vertical extent of a few kilometers may be overshot at these

ranges. Since wintertime precipitation is often shallow, this can be a significant problem (e.g., Smith et al. 1996; Super and Holroyd 1998).

Figures 43 and 44 suggest that overshooting appears to be occurring in the Lake of the Woods region in northwestern MN, because surface stations in that area [Warroad, MN (KRRT); Roseau, MN (KROX); and Baudette (KBDE)] reported light snow at locations where the radar analysis had a precipitation rate of < 0.1 mm hr^{-1} (Figure 46). At stations KRRT and KBDE, the radar analysis indicates no precipitation at all. The radar beam may be partially or completely overshooting the phenomenon, resulting in inaccurate estimates.

The approximate distances from the nearest radar, KMVX, to KBDE, KROX, and KRRT are 240 km, 187 km, and 212 km, respectively. Assuming standard atmospheric refractive conditions, the heights of radar beams originating from KMVX at KBDE, KROX, and KRRT are ~ 5.5 km AGL, 3.7 km AGL, and 4.5 km AGL, respectively. Thus, overshooting appears possible given the approximate heights of the radar beam *if* standard atmospheric refraction conditions are assumed. However, this assumption is not always valid. To evaluate this possibility and to estimate the location of the radar beam, a RUC sounding was extracted to compute an index of refraction profile for a location in the Lake of the Woods region to examine the impact of the environmental conditions and refractivity (*N*) on the estimated ray path. The results of the ray tracing calculations are discussed in a later subsection.

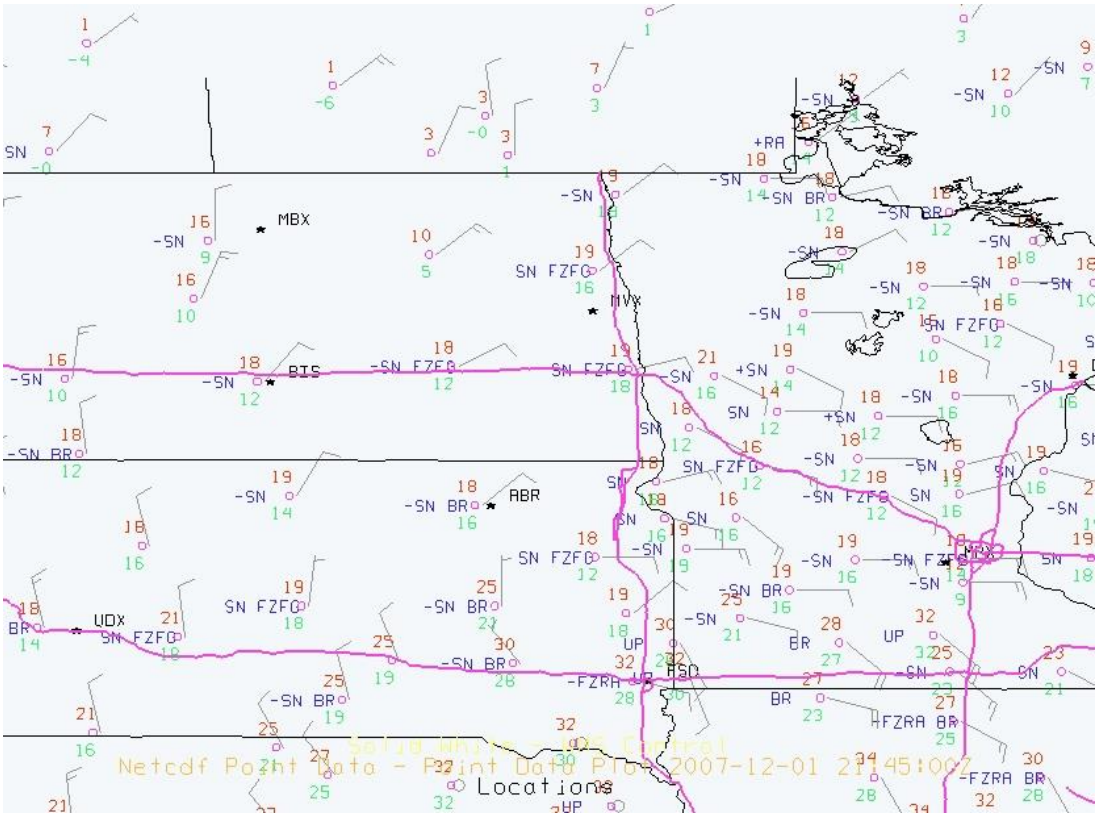

Figure 46. Station plot from 22 UTC December 1, 2007 over Northern Great Plains domain showing weather, wind barbs (knots), temperature (°F), and dewpoint (°F)

To help diagnose whether the if radar beam overshooting occurred near the Lake of the Woods region, METAR reports were examined for stations KBDE, KROX, and KRRT:

KBDE:

KBDE 012153Z AUTO 08010KT 3/4SM -SN BR VV014 M08/M11 A2997 RMK AO2 SLP175 P0000 T10831106

KROX:

KROX 012155Z AUTO 07012KT 1 1/4SM -SN OVC005 M06/M08 A2993 RMK AO2

KRRT:

KRRT 012155Z AUTO 09014KT 3/4SM -SN BKN001 OVC006 M08/M10 A2997 RMK AO2

These reports reveal that in the Lake of the Woods region, cloud ceilings were low – conditions ranged from broken coverage at 100 feet to an indefinite ceiling at 1,400 feet (KBDE). The report of a vertical visibility, VV, at KBDE indicates that there was an indefinite ceiling caused by either the reported mist or snow. The low ceilings suggest that the lower levels were very moist, as confirmed by a RUC sounding for 48.9° N, -95.0° W (Figure 47). Thus, overshooting may have occurred here or, at least, partial beam filling due to the moist conditions. Further examination of the refractive index profile and subsequent ray path computed from this sounding will further our evaluation of this hypothesis.

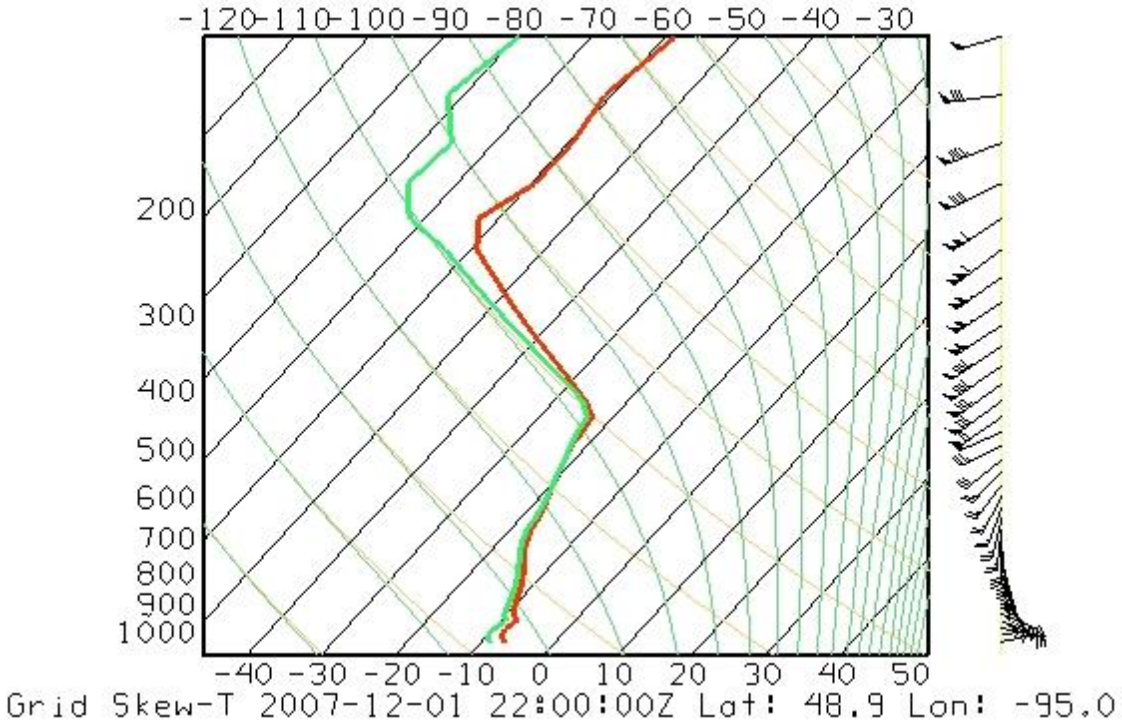

Figure 47. RUC skew-T profile for 48.9° N, -95.0° W showing environmental temperature (°C) (red line) and dewpoint (°C) (green line)

Another area that may have been affected adversely by overshooting is the area centered on Minot, ND. Both the Minot Air Force Base (KMIB) and Minot Airport ASOS (KMOT) reported

light snow. The distances from the KBIS radar (the nearest radar) to KMIB and KMOT are ~190 km and 170 km, respectively. With standard atmospheric refractive conditions assumed, the heights of radar beams from KBIS at KMIB and KMOT are ~3.8 km AGL and 3.2 km AGL, respectively. Unfortunately, surface stations are very sparse in this part of ND and absent from the radar analysis is the site located at the Minot Air Force Base in Minot, ND (KMBX). Radar data from KMBX cannot be used in the PPAES radar module because Level II data from this radar are not disseminated to the meteorological community. Only Level III radar data are available, which are not compatible with this version of the PPAES radar module. To this end, overshooting with KMVX and KBIS did seem to occur because the KMBX radar did indicate precipitation in the central ND area (Figure 48) and, at KMIB and KMOT, precipitation was reported.

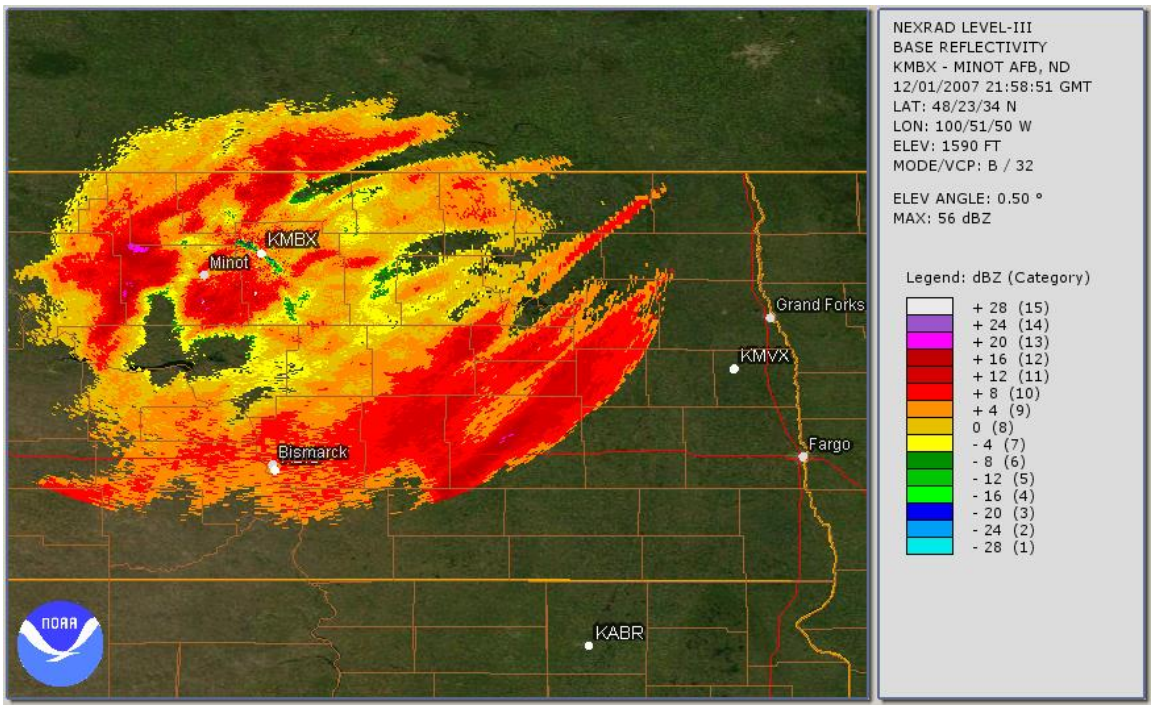

Figure 48. NEXRAD Level III base reflectivity at the Minot Air Force Base (KMBX) at 21:58 UTC December 1, 2007

As in the Lake of the Woods region, the ceilings at KMOT and KMIB were 1,100 feet and 800 feet, respectively, suggesting a moist lower troposphere. A RUC sounding for a point near the KMBX radar (48.3° N, -101.0° W, Figure 49). The Skew-T profile again suggests a moist lower troposphere. We will use this sounding to compute an index of refraction profile for this location to better describe radar ray propagation, which will allow us to more conclusively determine whether overshooting was occurring.

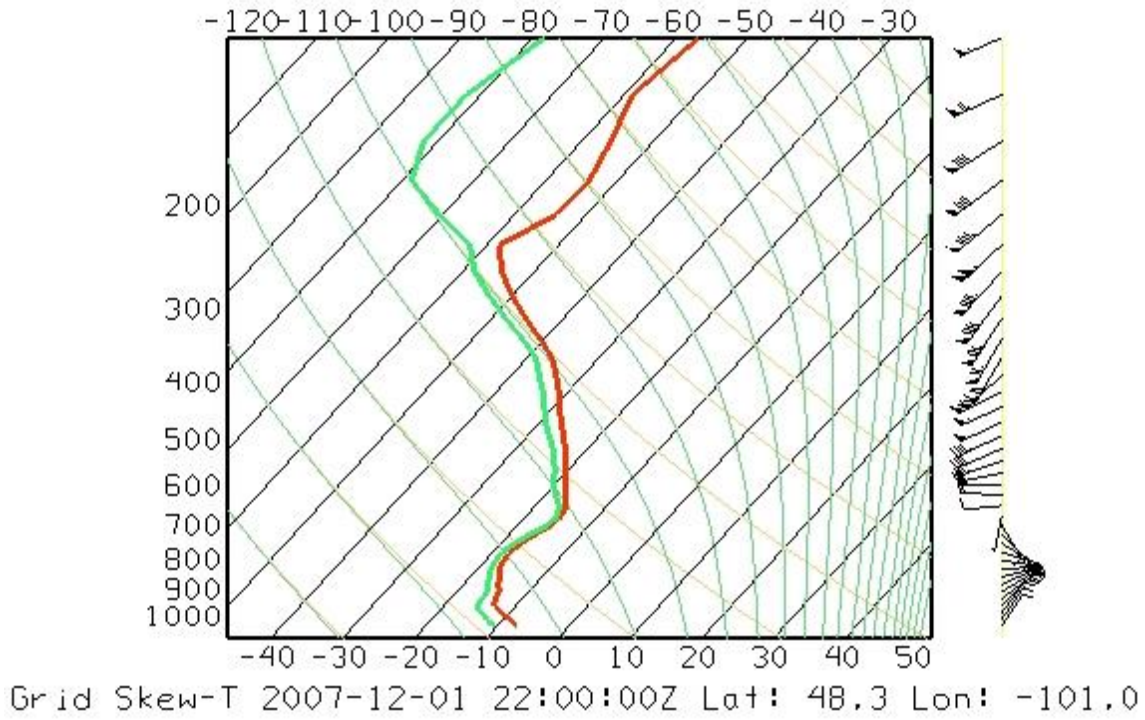

Figure 49. RUC skew-T profile for (48.3° N, -101.0° W) showing environmental temperature (°C) (red line) and dewpoint (°C) (green line)

6.3 Electromagnetic Propagation

Environmental conditions can influence radar ray paths such that assuming standard atmospheric refractive conditions is not viable. Non-standard propagation can occur in the boundary layer due to large vertical moisture gradients and temperature inversions (e.g., Doviak and Zrnić 1993, 16-23; Gao et al. 2008). These conditions result in the radar ray path being lower (superrefraction) or higher (subrefraction) than normal—where "normal" propagation is for standard atmospheric refractive conditions (e.g., Doviak and Zrnić 1993, 23-28). To examine this issue, vertical index of refraction profiles were computed for the soundings presented in Figures 47 and 49. These profiles were then used as input for an electromagnetic (EM) propagation model that can be used to compute ray paths in the atmosphere (M. Askelson 2011, pers. comm.). This EM propagation model assumes a spherically-stratified atmosphere such that only vertical refraction is handled.

The two resulting index of refraction profiles obtained were very similar and produced similar ray paths for 0.5° elevation (i.e., the lowest radar scan angle). These ray paths were compared to those obtained by using the theoretical effective Earth radius model, which assumes a constant vertical refractive index gradient and produces accurate ray paths for standard refractive conditions (Doviak and Zrnić 1993, 20-23). The estimated ray paths did not differ greatly from those obtained with the theoretical effective Earth radius model for both the Lake of the Woods (Figure 50) and Minot, ND, locations (not shown). The determination of the actual radar ray path allows ray heights to be compared with cloud ceilings at these METAR reporting sites.

The ray paths at the two locations – Minot, ND, and Lake of the Woods – suggest that overshooting occurred. This is because the cloud ceilings were considerably lower than the radar rays at these surface stations. At KMOT the ceiling and ray height for the radar KBIS were approximately 1,100 ft (0.34 km) and 3.38 km; at KMIB these were 800 ft (0.24 km) 4.01 km, respectively. At KROX, KRRT, and KBDE, the respective ceilings and radar ray beam heights for radar KMVX were 500 ft (0.15 km) and 3.9 km, 100 ft (0.03 km) and 4.75 km, and 1,400 ft (0.43 km) and 5.8 km, respectively. Thus, overshooting is suggested at both locations.

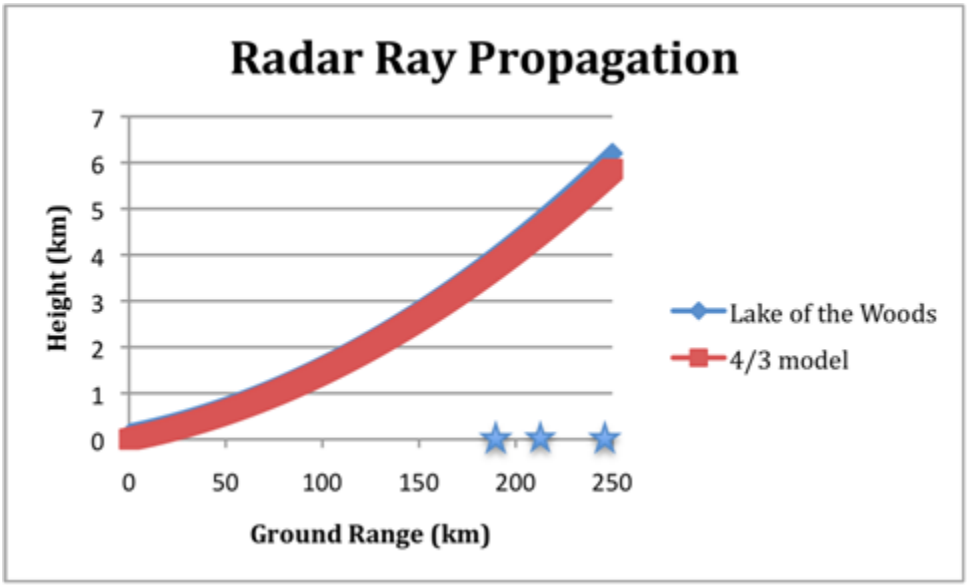

Figure 50. Electromagnetic ray path, at 0.5° elevation, from Lake of the Woods index of refraction sounding for 22:00 UTC December1, 2007

Red and blue lines represent the effective earth radius model (4/3 model) and the estimated ray path from the environmental sounding, respectively. The three stars represent the three stations of KROX, KRRT, and KBDE (from left to right) and their distances away from KMVX of 187 km, 212 km, and 240 km, respectively. KROX, KRRT, and KBDE had cloud ceilings of 500 ft (0.15 km), 100 ft (0.03 ft), and 1,400 ft (0.43 km), respectively.

6.4 Surface Analysis Discrepancy Assessment

For the surface and blending modules, limitations associated with surface stations can be a factor. One such limitation involves how present weather is determined using LEDWI sensors at automated stations. LEDWI specifications for monitoring precipitation include:

1) Must be capable of detecting precipitation 99% of the time when precipitation rates are greater than or equal to 0.25 mm hr^{-1},
2) Must be capable of identifying precipitation 97% and 90% of the time for solid and liquid precipitation respectively, and
3) False alarm rates are required to be less than 0.2% (Wade 2003).

Wade (2003) indicates that studies in the early 1990s revealed that LEDWIs met or exceeded these requirements. For the first requirement, the minimum threshold is 0.25 mm hr^{-1}, which can have significant consequences for measuring light precipitation since drizzle generally occurs and light snow can occur below this threshold. Thus, light snow oftentimes may not be detected. The situation is further compounded by the fact that there is no requirement for an ASOS to detect drizzle, although it is hoped that future work will make this possible (Wade 2003).

Stations KDVP, KJYG, and KRWF reported no precipitation during the event while the radar analysis had precipitation occurring (cf. Figures 44, 51). Possible theories as to why this occurred are: (a) LEDWI sensors failed to detect precipitation that occurred, (b) blowing snow occurred and caused a detection problem, (c) radars suffered from virga contamination, and (d) the time difference between the surface observations and the radar analysis resulted in differences regarding where precipitation was estimated to have occurred.

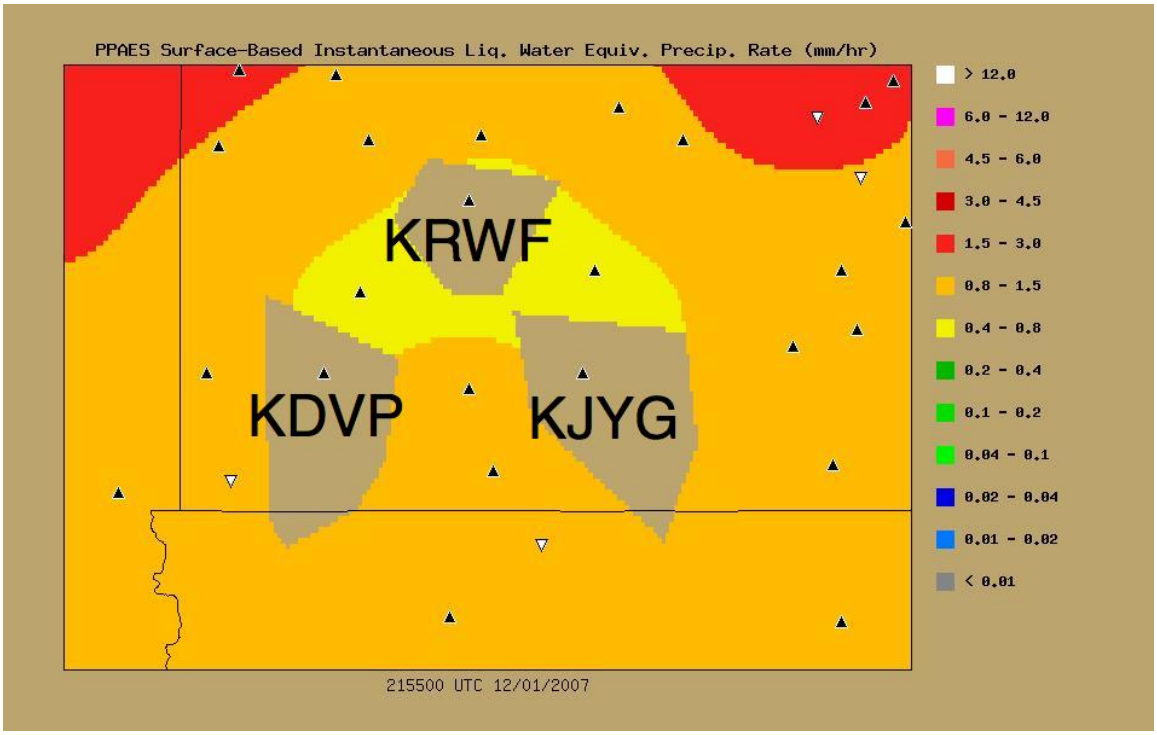

Figure 51. PPAES surface analysis at 21:55:00 UTC December 1, 2007 over Northern Plains domain

Stations KRWF, KDVP, and KJYG coincide with locations on the radar analysis at which precipitation was indicated despite no similar surface reports extant. Other conventions follow Figure 42.

To test if precipitation was too light for the LEDWI sensor to accurately detect snowfall at stations KRWF, KJYG, and KDVP, the METARs for these stations are examined below, with the METAR used in the surface analysis listed as (c).

KRWF:

a) KRWF 012123Z AUTO 09013G22KT 1SM -SN BR FEW011 OVC020 M07/M09 A2965 RMK AO2 P0000
b) KRWF 012151Z AUTO 09010G22KT 2 1/2SM -SN BR OVC016 M07/M09 A2963 RMK AO2 PRESFR P0000
c) KRWF 012153Z AUTO 09013G22KT 2 1/2SM BR OVC014 M07/M09 A2963 RMK AO2 SNE52 SLP049 P0000 T10671089

KJYG:

a) KJYG 012116Z AUTO 12022G28KT 2 1/2SM UP SCT007 OVC010 M01/M02 A2965 RMK AO2 P0001
b) KJYG 012136Z AUTO 12021G28KT 4SM BR OVC012 M01/M03 A2963 RMK AO2 P0002
c) KJYG 012156Z AUTO 12022G26KT 2 1/2SM BR OVC012 M02/M03 A2961 RMK AO2

KDVP:

a) KDVP 012116Z AUTO 13018G23KT 1 3/4SM BR BKN006 OVC011 00/M01 A2957 RMK AO2
b) KDVP 012136Z AUTO 13020KT 2SM BR OVC006 00/M01 A2955 RMK AO2
c) KDVP 012156Z AUTO 12020G24KT 3SM BR OVC006 00/M02 A2954 RMK AO2

Stations KDVP, KJYG, and KRWF all reported mist and no precipitation at 2155 UTC. The radar analysis indicates that light precipitation (0.1–0.2 mm hr^{-1}) was occurring at these sites (Figure 52). Because the LEDWI's minimum precipitation threshold is 0.25 mm hr^{-1}, it appears as if light precipitation could have been occurring and therefore was undetected or that precipitation was incorrectly identified as mist. In the literature, the misclassification of freezing drizzle as mist has been noted before (Ramsey 2002; Landolt et al. 2010). The difficulty lies in mist being defined as aggregates of water droplets being suspended in air, and mist cannot be routinely confirmed by examining droplet sizes but is assumed when specific criteria are met (Landolt et al. 2010). Specifically, if there is: a) an impediment to visibility; b) $(T\text{-}T_d) < 4°C$; and c) when the visibility is < 4 statute miles (Landolt et al. 2010). Thus, ASOSs and AWOSs can mistakenly classify light precipitation as mist.

The impact of mist can be significant in terms of false alarms resulting when stations report mist when precipitation is occurring. Conversely, a miss can result if mist is falsely identified through the above criteria when precipitation is occurring. The issue of mist in regard to precipitation is complex and is amplified in that human observers can often misreport mist and freezing drizzle (Ramsey 2002; Landolt et al. 2010).

Turning to blowing snow, it can cause the lenses of sensors to be blocked; alternatively, snowfall may be interpreted as blowing snow – a known problem (NWS 1998). An impact of blowing snow on the LEDWI receiver or projector is that it may contaminate the measurement and lead to inaccurate type/intensity reports (NWS 1998). Furthermore, when blowing snow conditions occur (wind speeds 10 kts or greater) LEDWIs can falsely indicate rain at temperatures less than 32 °F. The LEDWI blowing snow algorithm accounts for these inaccurate identifications and reports blowing snow or unknown precipitation. However, a consequence of this is that it is

possible for unknown precipitation to be reported when blowing snow is occurring. Although this did not occur at the analysis hour in question (but did occur earlier at KJYG), blowing snow may still be an issue here and, if so, could have impaired the measurement through contamination.

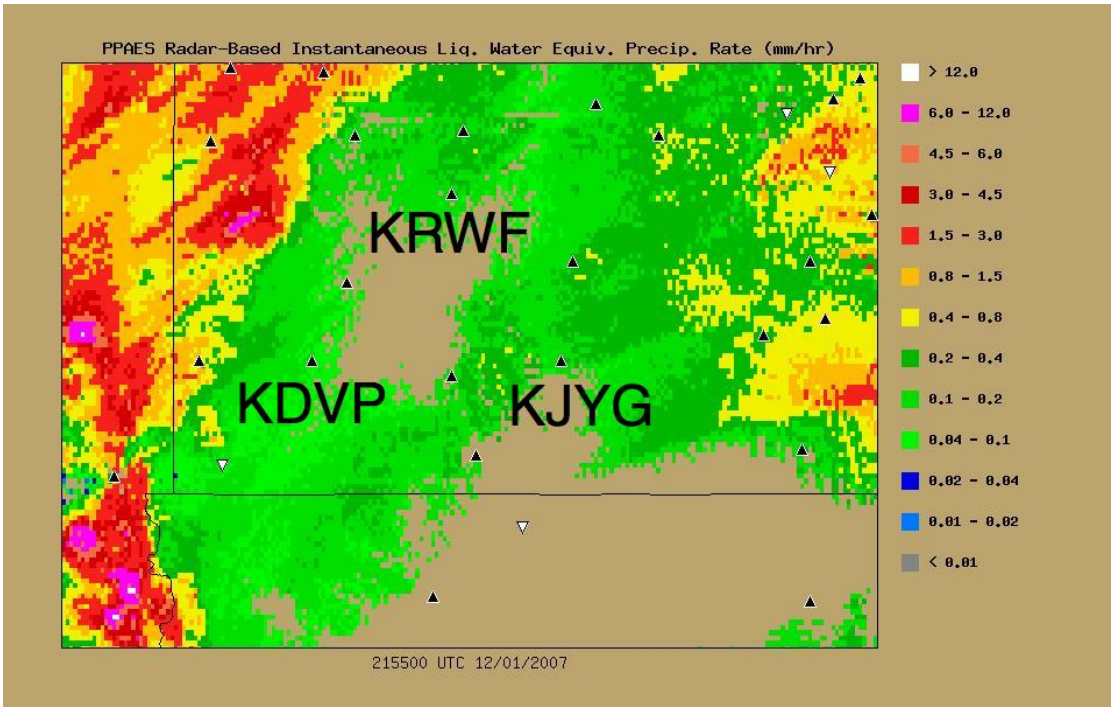

Figure 52. PPAES radar analysis at 21:55:00 UTC December 1, 2007 over Northern Great Plains domain

Stations KRWF, KDVP, and KJYG did not report precipitation. Other conventions follow Figure 42.

The conditions before and during the case are evaluated to examine if snow was on the ground before the event started, even though this is not a requirement for blowing snow to occur. This was accomplished by using National Operational Hydrologic Remote Sensing Center (NOHRSC) regional snow analyses following Carroll et al (2006) (Figures 53 and 54) for the period in question. The Carroll et al. snow analyses were derived from a combination of multiple data sources and numerical models.

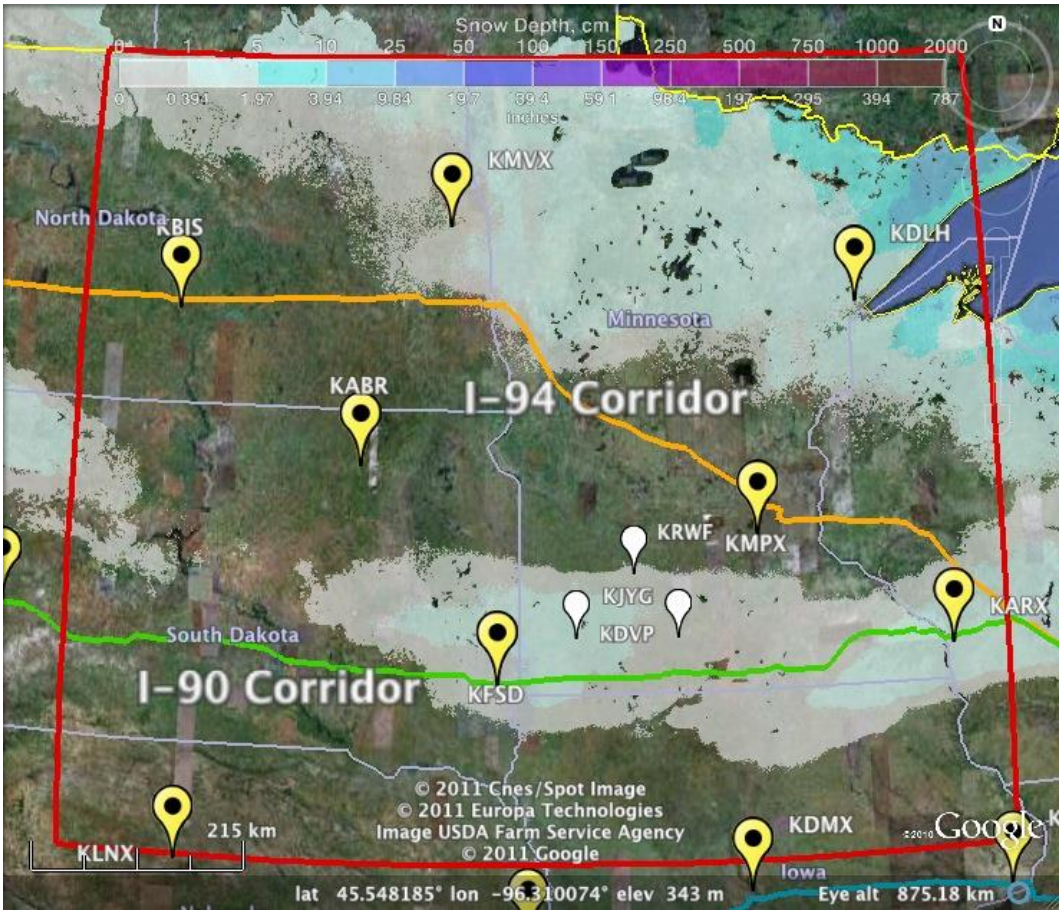

Figure 53. NOHRSC snow depth analysis (cm) for 0600 UTC December 1, 2007 showing NEXRAD WSR-88D radar sites (yellow markers) and Northern Plains domain (red line)

One limitation is that NOHRSC snow depth fields are limited to a single instantaneous snapshot at 0600 UTC daily and, as such, information between these hours must be inferred. The snow depth analyses showed that snow was on the ground at KRWF, KDVP, and KJYG at the start of the event (Figure 53). However, not every place in the domain had snow on the ground prior to the event, with snow absent along lines stretching from the radars KBIS-KABR-KMPX. Snow age can also be determined from examining the daily maps of snow depth, and the preceding snow depth analyses show that the snow on the ground near KRWF, KJYG, and KDVP was "new" (with the most recent snowfall occurring on 30 November), which favors blowing snow conditions.

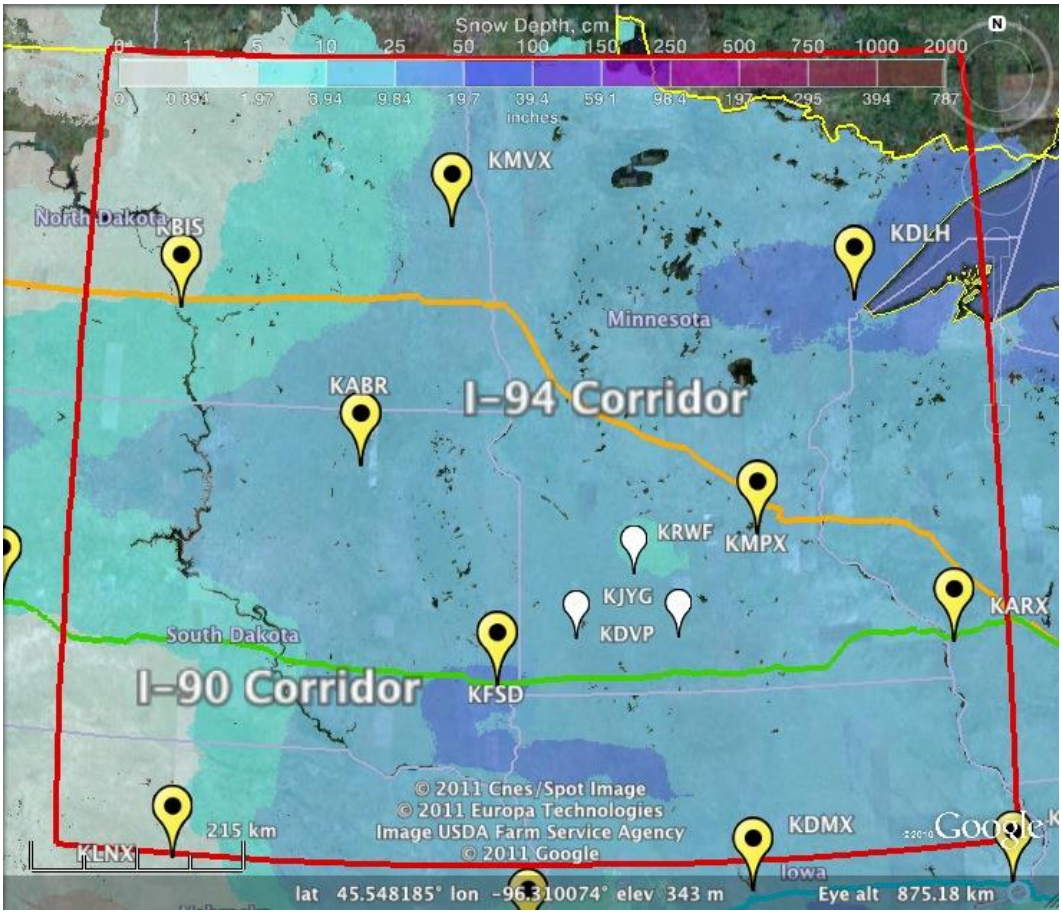

Figure 54. NOHRSC snow depth analysis (cm) for 0600 UTC December 2, 2007 showing. NEXRAD WSR-88D radar sites (yellow markers) and Northern Plains domain (red line)

Li and Pomeroy's (1997) blowing snow model is used to compute an appropriate threshold wind speed for the occurrence of blowing snow. For KDVP, KJYG, and KRWF, the threshold wind speeds for blowing snow were 18 kts, 17.5 kts, and 16 kts, respectively. The wind speeds observed at KDVP, KJYG, and KRWF at the time of observation were 20 kts (gusting to 24 kts), 22 kts (gusting to 26 kts), and 13 kts (gusting to 22 kts), respectively. Sustained wind speeds at two stations stations exceed the threshold criteria for blowing snow while the wind gusts are all higher than the computed thresholds. Note that this blowing snow threshold applies strictly only to transport of dry snow--defined as the case when the snow surface has not experienced temperatures above 0° C and no liquid precipitation has fallen (Li and Pomeroy 1997). Without a human observer at these sites (providing, for instance, a blowing snow remark in the METAR), blowing snow can only be identified as a possible factor using the Li and Pomeroy criteria.

Another possibility is that the radar analysis was contaminated by virga. The observed soundings at Aberdeen, SD (KABR; Figure 55), and Chanhassen, MN (KMPX; Figure 56), at 00 UTC 2 December 2007 can used to determine whether virga was likely.

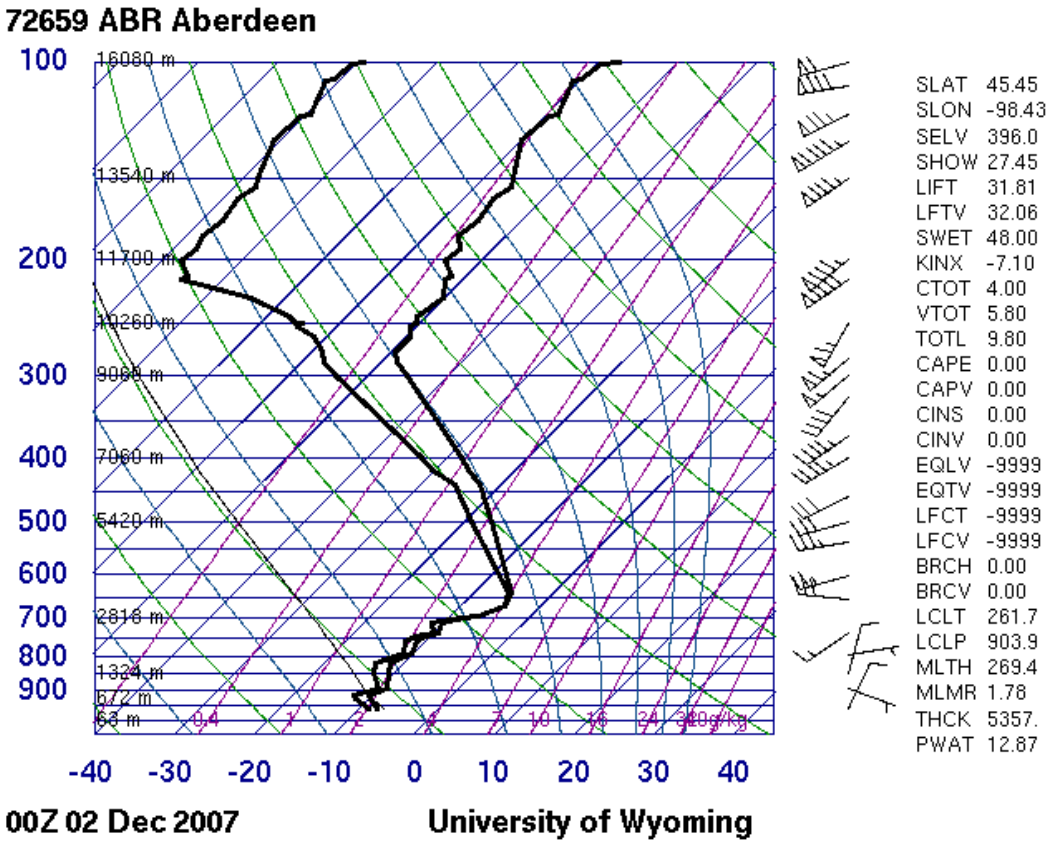

Figure 55. Skew-T Log-P sounding from Aberdeen, SD (KABR) at 00 UTC December 2, 2007

The sounding was obtained from the University of Wyoming at http://weather.uwyo.edu/upperair/sounding.html.

The soundings reveal that the layer near the surface was quite moist with mean relative humidities in the lowest 1.5 km of 88% and 93% for KABR and KMPX, respectively. The lowest 1.5 km is significant in that it is currently employed in the WSR-88D snow accumulation algorithm to diagnose virga (Hunter et al. 2001), with a threshold of 70% mean relative humidity in this layer resulting in initialization of a procedure to determine if virga is occurring.

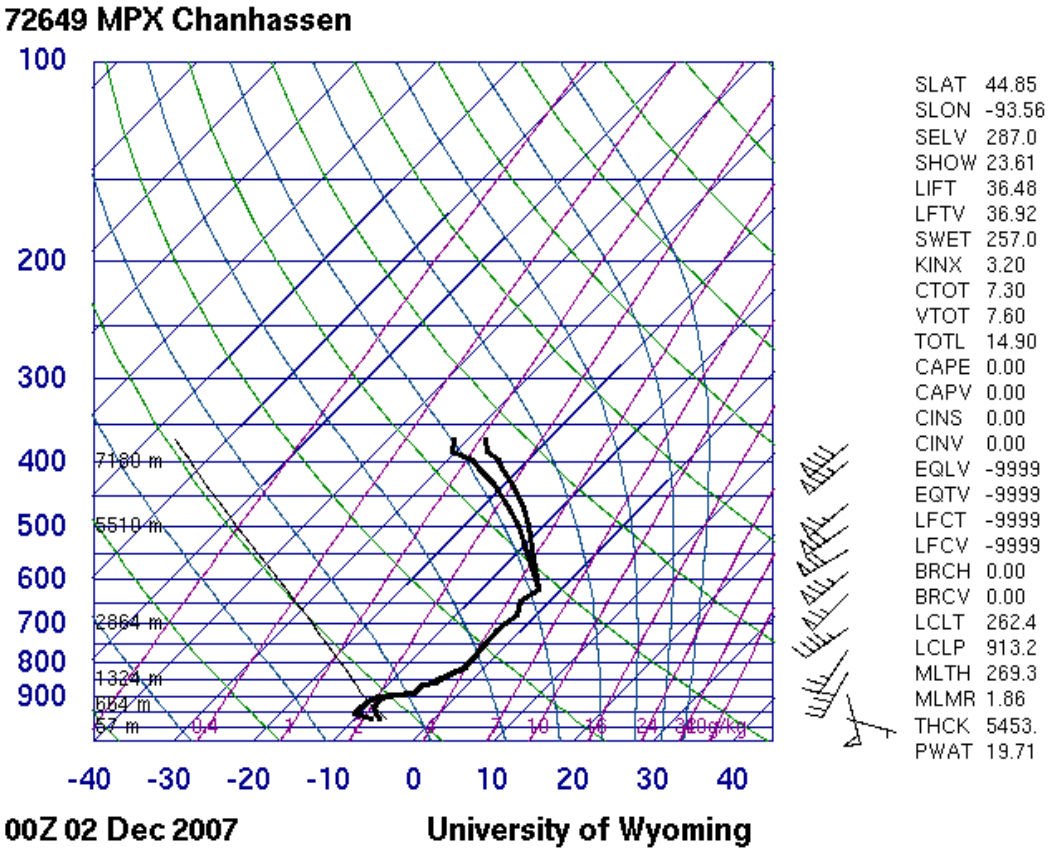

Figure 56. Skew-T Log-P sounding from Chanhassen, MN (KMPX) at 00 UTC December 2, 2007

The sounding was obtained from the University of Wyoming at http://weather.uwyo.edu/upperair/sounding.html.

Unfortunately, the KABR and KMPX upper air sites are relatively far away from the three surface stations in question (Figure 57). Thus, the relevance of these two relative humidity profiles is limited because of the differences in time (approximately two hours) and space (over 100 km separating the nearest surface station and the nearest upper air site) between the surface and upper air observations. Despite this, the two soundings do provide some insight into lower tropospheric conditions at the time of the event. To obtain a sounding that was much closer in both time and space, a RUC sounding for (44.3° N, -95.2° W, Figure 58) at 2200 UTC was examined. This location is at the approximate centroid of the three stations. This sounding also shows that the lower troposphere was quite moist and, thus, that virga was unlikely for these three stations.

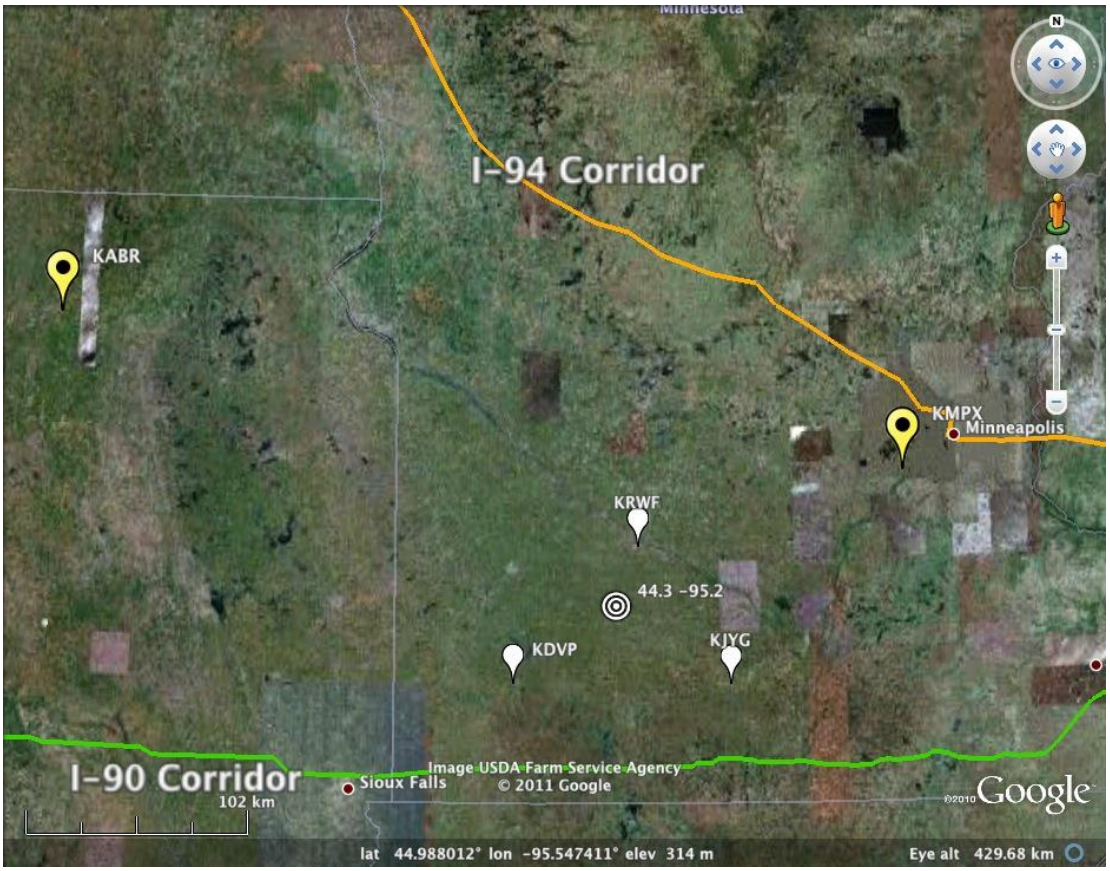

Figure 57. Relative spatial layout of the three METAR sites for the analysis hour of 22 UTC December 1, 2007 (KDVP, KJYG, and KRWF), the two nearest upper air sites (KABR and KMPX), and the location where a RUC-2 sounding was extracted

A surface station plot (Figure 46) was also examined for this period and showed that the dewpoint depression around the surrounding area was less than 5 °F. Because of this, we suspect that virga is not overly likely. Other independent sources were sought to confirm or deny the presence of virga, including various satellite sources. Unfortunately, the resolution of satellite soundings is generally too coarse and, in cloud, are not highly reliable regardless of the wavelengths utilized; imagery products are most useful for determining cloud top pressure and temperature, and even products derived from subtracting channels are not optimum for application to the virga problem.

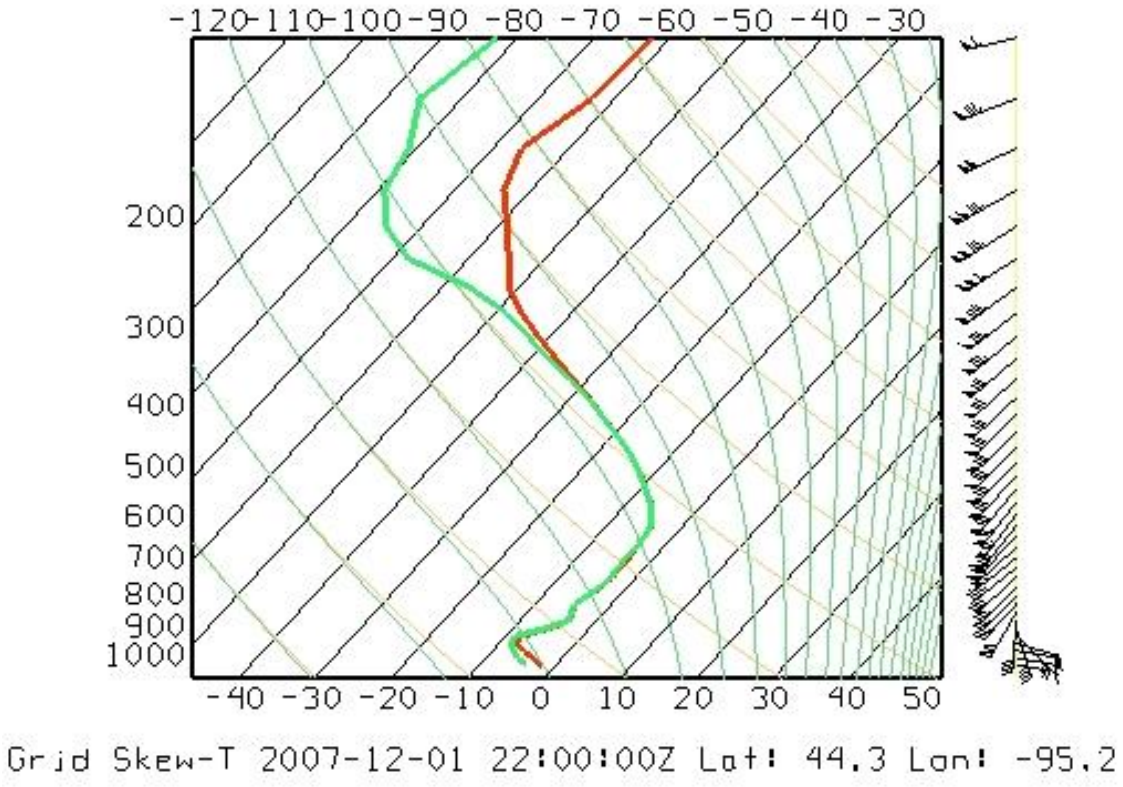

Figure 58. RUC-2 skew-T profile extracted from a grid point located at 44.3° N and -95.2° W showing environmental temperature (°C) (red line) and dewpoint (°C) (green line)

Another possible contributing factor to precipitation discrepancies is the difference in time between the surface observation reporting times and the radar analysis. Both the surface and radar analysis times were centered on 21:55 UTC 1 December 2007. The observations from KDVP, KJYG, and KRWF that are being considered were reported on 1 December 2007 at 21:56 UTC, 21:56 UTC, and 21:53 UTC, respectively. Per the LEDWI sensor's algorithm, precipitation intensity is determined from 1-minute samples from the past five minutes before the report (NWS 1998). The highest common intensity derived from three or more samples is the intensity that is reported.

The three nearest radars that may have provided information to the radar analysis for this area are KABR, KFSD, and KMPX. The radars KABR, KFSD, and KMPX had time stamps of 21:53 UTC, 21:54 UTC, and 21:56 UTC, respectively. These times correspond to the start of the first reflectivity sweep in a radar's volume scan. Therefore, a time difference between the radar analysis and the surface stations is present, albeit small [the information acquired spans from 21:48 UTC (surface station KRWF) to 21: 56 UTC]. Upon examining the radar loop in question, the time difference appears to be negligible in this situation.

MADIS and Clarus Surface Analysis

In Figure 31, the average RMSE value is abnormally high for the 12/26-12/27/2012 case of the Utah region study (5.02). This is because two values in the average calculation were extremely high (for reasons that need further examination). The median value, however, is 2.4, which is more consistent with values from Utah region cases (all RMSE values increased when Clarus data were added).

The inclusion of Clarus data into the PPAES surface analysis has been shown to decrease the quality of the PPAES surface analysis. There are many potential causes of this. Some possibilities are that the ESS and MADIS station instrumentation performance characteristics are significantly from each other or that the precipitation instrumentation on ESSs is not maintained in the same fashion as those on MADIS stations. Either of these could cause information from ESSs to degrade the surface analyses. Further examination into the use of RWIS data to improve analysis quality is still needed, especially regarding the use of quality control flags in combination with other data to help remove data from the analysis if they are considered to be of poor quality. In the current analysis, all checks must pass before using the data in the analysis except for the Barnes spatial analysis check, which is not used owing to uncertainty regarding its utility for precipitation. The data were processed when the check flags were set to pass, available but does not run, and not supposed to run. Further tests could include optimization of these flags, including testing the utility of the Barnes spatial check.

Radar Analyses in Complex Terrain

To better understand the impacts of using the terrain clearance option in the PPAES radar module and to illustrate differences that affect statistical results, situations in which performance is better without and with terrain clearance are shown in Figures 59 and 60, respectively. Triangles denote surface stations where precipitation was not occurring and squares denote surface stations at which precipitation was occurring. Ovals indicate locations where inclusion of terrain clearance altered the analysis field.

As is apparent in these figures, the differences that result from including terrain clearance are relatively small. This is because WSR-88D radars are sited such that beam blockage at the lowest elevation is minimized. While this may seem to make the incorporation of this capability relatively futile, it could have very large impacts in non-standard ray propagation conditions. While a ray-tracing model for such conditions has not as of yet been incorporated, the PPAES radar module has been designed to enable easy incorporation of such a capability.

Thus, the differences in performance seen in Figures 32 through 35 result from small changes in the analysis fields. It is possible that this is reinforced by the fact that the cases considered herein were relatively light snowfall events. If more vigorous events were considered, one might have more confidence in performance metric differences. However, relatively light events can oftentimes be the most challenging from a road maintenance standpoint.

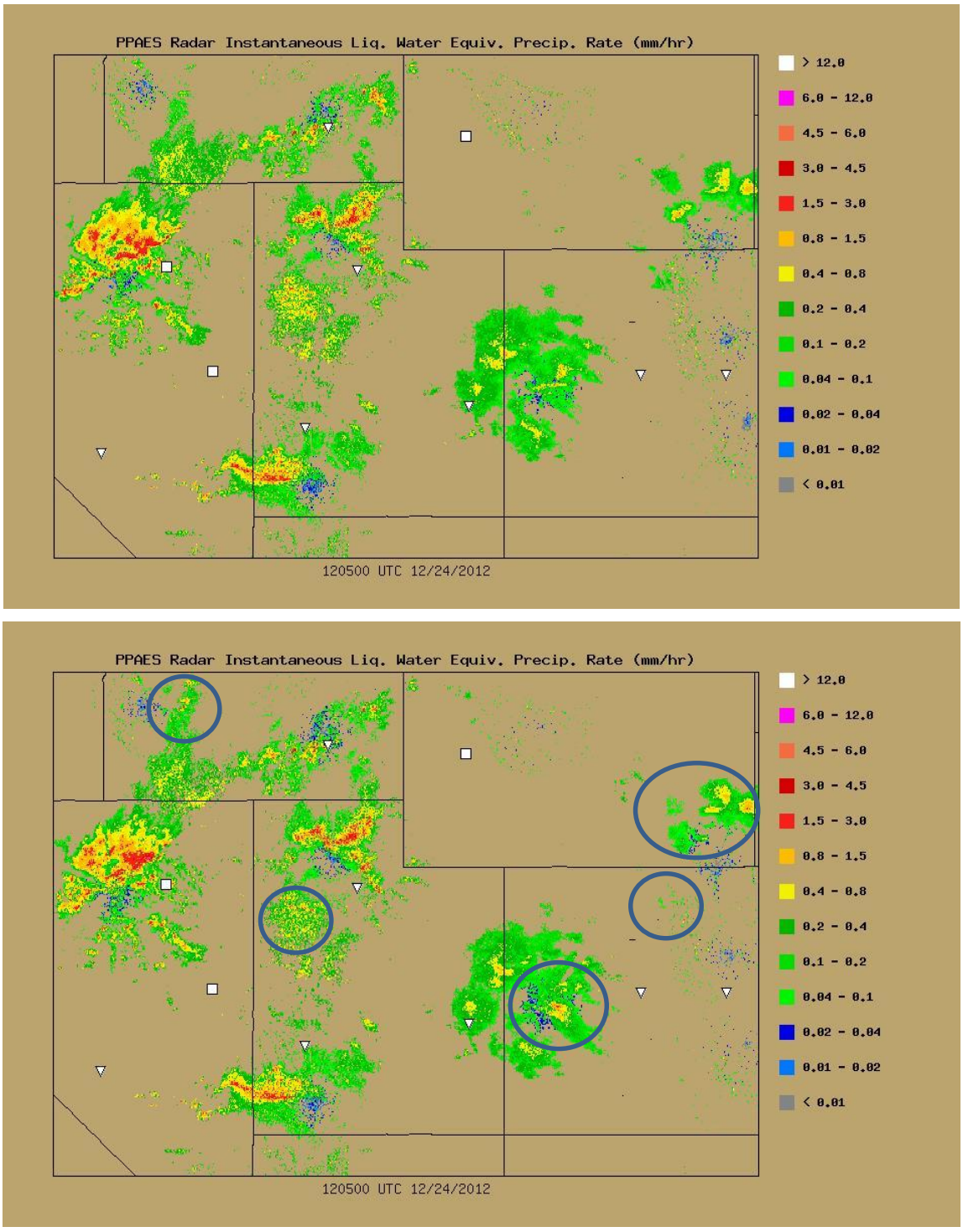

Figure 59. PPAES radar-based analyses at 120500 UTC December 24, 2012 without (top) and with (bottom) terrain clearance

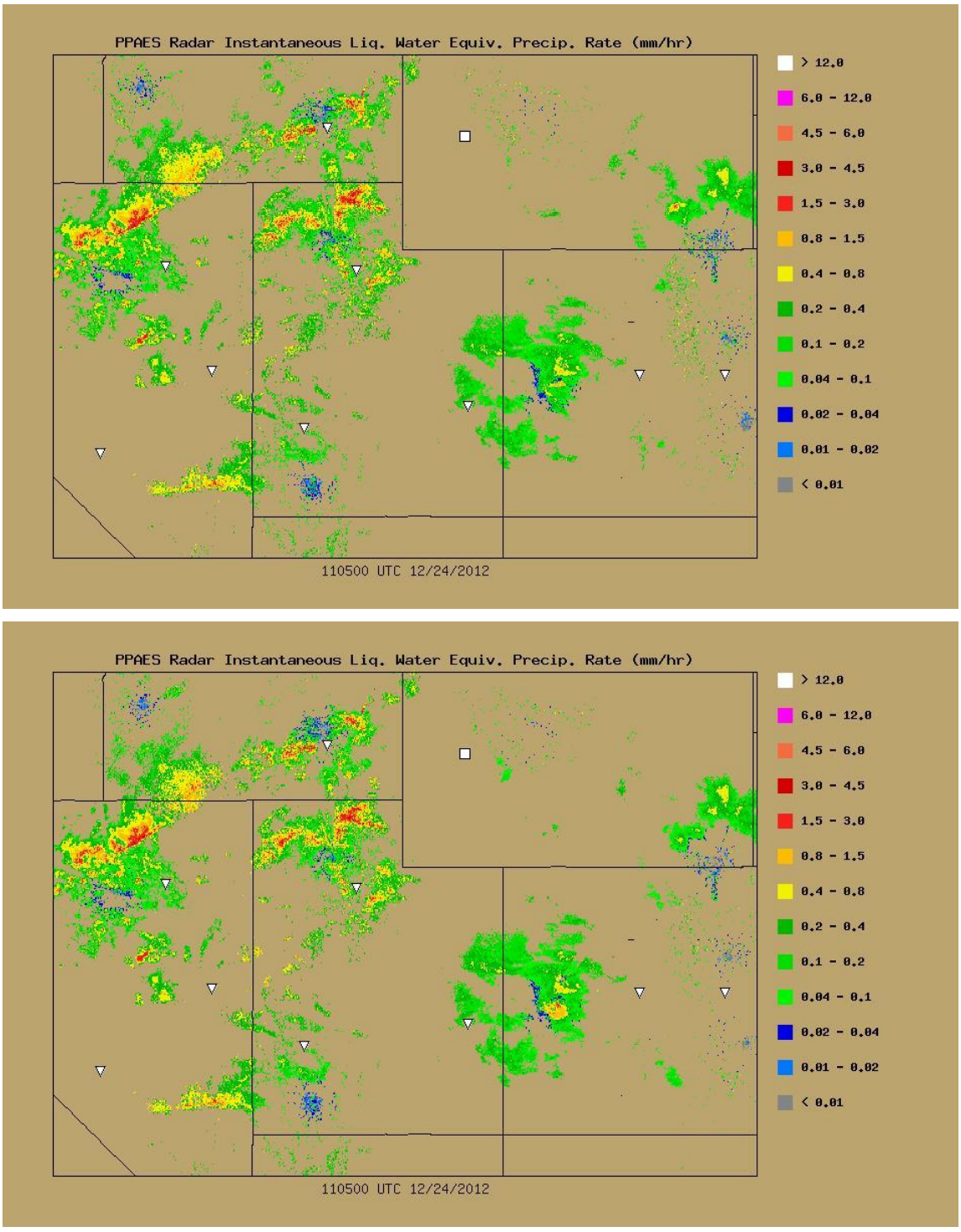

Figure 60. PPAES radar-based analyses at 110500 UTC December 24, 2012 without (top) and with (bottom) terrain clearance

The fields shown in Figure 60 underscore just how subtle differences can be. At this time, POD without terrain clearance is 0, but with terrain clearance is >0 (not shown). Because only one surface station indicated precipitation at this time, this difference has to arise because

precipitation is indicated at that station with terrain clearance but is not indicated at that station without terrain clearance. Comparing the images in Figure 60, however, one cannot discern a difference at that station. Thus, the precipitation at that location that is indicated with terrain clearance is covered by the square that is used to indicate that precipitation was observed at that location with a surface station.

The impact of including terrain clearance, therefore, is uncertain. Future efforts should focus on testing with a larger number of events, including stratification of events according to intensity and spatial extent. Moreover, tests of sensitivity to the height clearance requirement (i.e., how far above the terrain the bottom of rays must be) would be useful.

Radar and Model Analyses

An example of a radar+model blended analysis is provided in Figure 61. From that Figure, the difference between the resolution of the model output and that of radar data is immediately apparent. Because of this difference, regions in which analysis values primarily come from radar data exhibit much more fine scale variability than do those where analysis values primarily come from model data. This produces an aesthitically displeasing appearance, although it enables the user to rapidly discern the source of information in the analysis field.

It is apparent from Figure 61 that the model data produced a much larger precipitation shield that generally consisted of more intense precipitation. In addition, the continuity of the model field is better than that of the radar-based field. The lack of continuity in the radar analysis, with precipitation seemingly focused around radar locations, is a consequence of radars being too far apart and radar beams overshooting precipitation. The model-based analysis does not suffer from this issue.

The model-based analyses do contain errors, however, as is apparent in Figures 36 and37. In fact, as indicated in Figures 36 and 37, for these events the radar-based analyses, despite their lack of continuity and the coverage between radars, are superior from a POD, FAR, ETS, etc., standpoint. An example of an issue that can arise when using model data is provided in Figure 61. While the model indicates that precipitation is occurring in southeastern Idaho, the radar data clearly indicate that it is not. Thus, the RAP model is over-predicting precipitation occurrence in these events.

One interesting aspect here is the “re-use” of radar data. The RAP model, through its data assimilation system, incorporates information from both forecasts and observations, including radar data, to produce a physically-consistent representation of the atmosphere. Thus, blending radar data with RAP output results, in a way, involves re-using radar data. Despite this, the result is, from the standpoint of the performance metrics used herein, improvement relative to the use of model output alone, although from this standpoint the use of radar data alone is best.

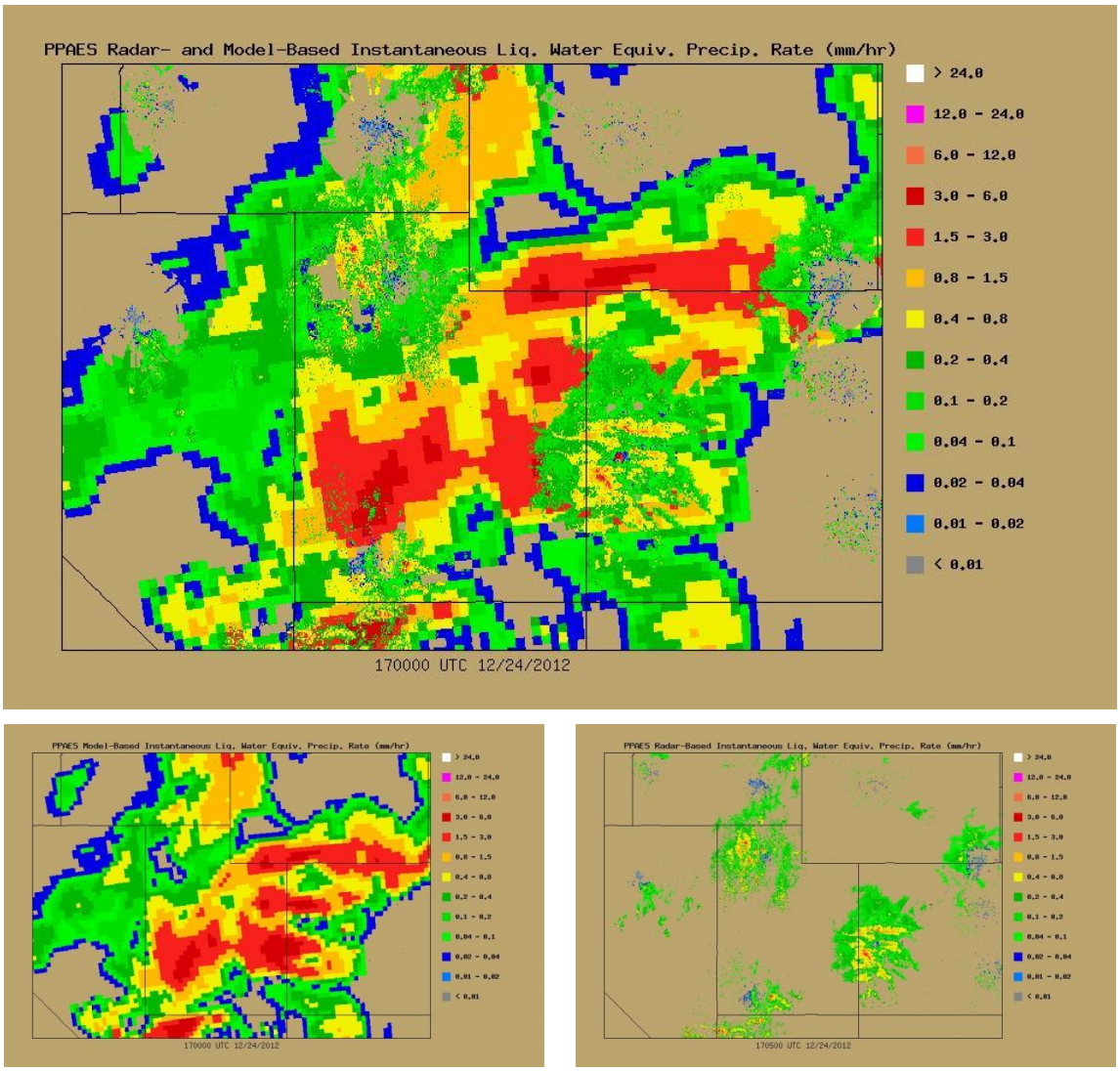

Figure 61. PPAES radar+model (top), model (lower left), and radar (lower right) analyses at 170000 UTC (170500 for radar) December 24, 2012

Blending of Surface, Radar, and Model Analyses

As shown in Table 13, the radar and surface analyses performed differently with regards to the performance metrics, with a higher FAR and POFD for the surface analysis. This difference occurred because KEST did not report precipitation and the surface analysis module estimated precipitation to be occurring (Figures 43 and 44) as a result of the precipitation occurrence/non-occurrence nearest neighbor scheme. The radar and surface analysis modules produced similar performance metric values for accuracy and TS (Table 13). The radar analysis had a miss at KMHE, which reported light snow (Figures 43, 44). For TS, both had identical scores of 0.92. Mean error values indicate that precipitation rates were generally underestimated in a domain-averaged sense for both analyses, though mean errors between the two modules differed by 0.73 mm hr^{-1}. The surface module performed better with regard to the precipitation rate metrics used in this study.

6.5 Radar/Surface Blending Analysis Assessment

The most noticeable differences occur between the blended and radar analyses for this case, particularly around the periphery of the radar-analysis precipitation shield, which in the blended analysis (Figure 62) is filled in with information gained from the surface analysis. This "filling-in" effect is most noticeable in northwest ND, northern MN, and IA, bordering or including areas where precipitation is likely being partially overshot or completely overshot.

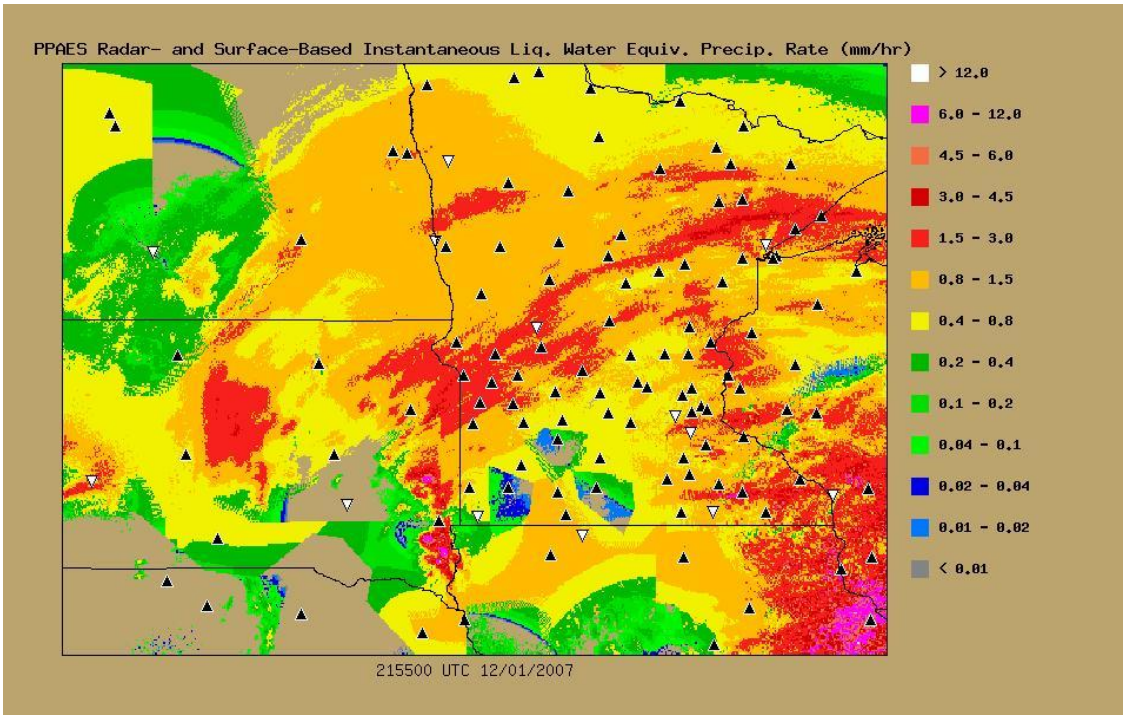

Figure 62. PPAES blended analysis at 21:55:00 UTC December 1, 2007 over Northern Great Plains domain

Black triangles and white inverted triangles denote surface stations used and withheld from the surface and blended analyses, respectively.

Another difference between the blended and radar analyses is that radar precipitation rates were adjusted upward for all of the radar sites in this domain except for one (KARX); the most noticeable change occurs with the KMVX radar data. Twelve surface observation-radar data pairs within the KMVX radar effective range resulted in a computed adjustment of approximately +0.66 mm hr^{-1}. Because of the large modified effective range for KMVX (~206 km) and because the precipitation was widespread, the radar-based analysis transitions cleanly to the surface analysis except for an area in central ND.

In addition, values derived from several other radars were also corrected: KDLH had an adjustment of +0.36 mm hr^{-1}, KMPX had an adjustment of +0.47 mm hr^{-1}, KABR had an adjustment of +0.37 mm hr^{-1}, and KARX had a negative adjustment of -0.27 mm hr^{-1}. An effect

of the adjustment at KARX is that several radar-estimated precipitation rates were corrected downward to near values of 0.01 mm hr^{-1}. Surface module precipitation rates around KARX were much smaller relative to the radar module precipitation rates, which resulted in a rather extreme adjustment (compare Figures 43 and 44).

The adjustment of KARX precipitation rates was a result of 11 co-located radar and surface station pairs being used to compute the resulting negative adjustment value. The radar-based values were higher because precipitation was not just snow but also freezing rain. The change in precipitation type can be traced to the increase in radar-estimated precipitation rate, as the *Z-S* relation used in this study was likely not appropriate given that a single hydrometeor species—snow—is assumed. Furthermore, a problem with mixed precipitation events is that bright band enhancement, associated with a melting layer in the cloud, may be taking place (Fulton et al. 1998; Pruppacher and Klett 1997, 61-64). The change of phase from solid to liquid causes the dielectric constant of hydrometeors to increase, such that the liquid water content enhances radar reflectivity values (Pruppacher and Klett 1997, 61-64).

The argument that the precipitation near KARX was not purely snow is supported by the radar reflectivity being > 45 dBZ in this region (Figure 63) and the station plots where ice pellets, unknown precipitation, freezing rain, and snow were reported. High reflectivity values were noticeable more than 120 km away from the radar, at which radar beam heights were greater than 2 km (assuming standard propagation conditions). A RUC sounding at (43.8° N, -89.6° W), 125 km east of KARX and near the radar reflectivity factor maximum, suggests that bright band enhancement in the area was possible but also does not confirm that it conclusively occurred since the temperature profile approaches but is never exceeds 0 °C (Figure 64).

The effective ranges of individual radars were modified due to the radars' failure to detect precipitation at large ranges. For example, the radar site KFSD had areas where it failed to sensed precipitation. The original unmodified effective range for KFSD obtained from the radar module was ~272 km. For this radar the effective ranges of the first (0° - 90°), second (90° - 180°), and third (180° - 270°) quadrants were modified the most, being truncated to 116 km, 132 km, and 150 km, respectively. The stations KTKC, KSPW, and KLCG were used to modify KFSD's effective ranges in the first, second, and third quadrants, respectively. The transition from radar-based to surface-based analysis values is noticeable in Figure 62 in the first, second, and third quadrants of KFSD.

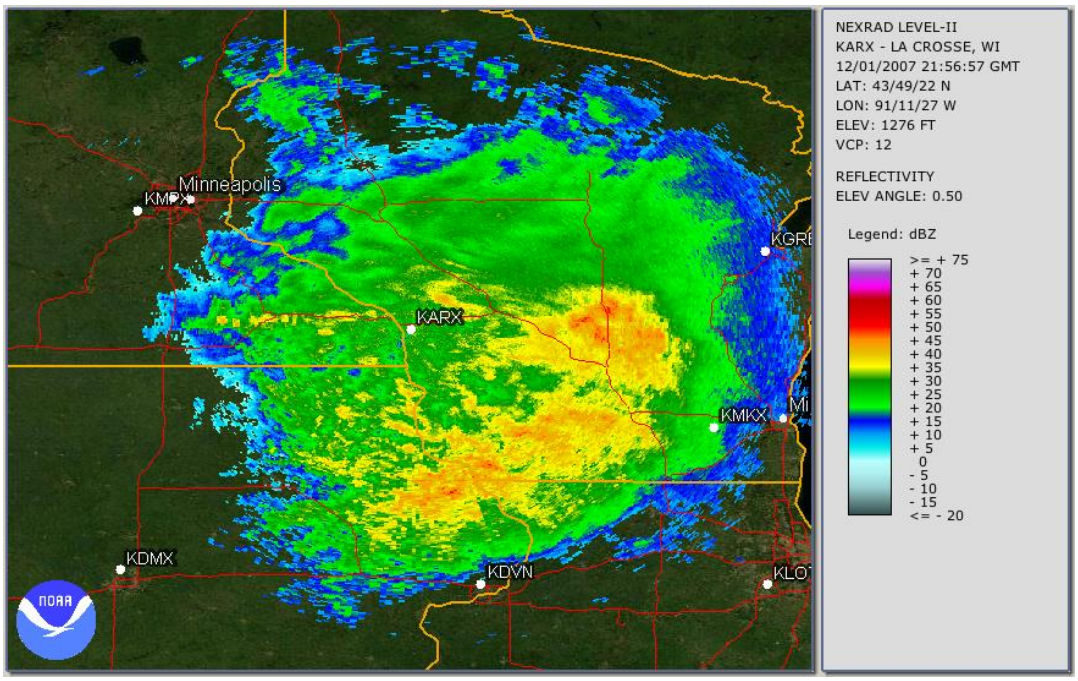

Figure 63. NEXRAD Level II base reflectivity at La Crosse, WI (KARX), at 21:56 UTC December 1, 2007

In KFSD's first quadrant, the surface-based analysis indicates no precipitation in the area where the radar-based analysis would be blended to the surface-based analysis. The blending algorithm began transitioning from the radar-based to the surface-based analysis at ~16 km away from the radar in this quadrant. Per the blending algorithm, the radar-based analysis values from KFSD (not adjusted since KFSD had too few pairs to compute an adjustment) were linearly blended to the surface-based analysis. Where the void was left by the surface-based analysis, the algorithm continued blending with a surface rate fixed at 0.0 mm hr^{-1}. As such, the blended analysis tapers off in this area where the surface analysis indicated no precipitation. Elsewhere, where surface-based and radar-based values are present (e.g., KFSD's second quadrant), the transition is much smoother and less noticeable. In the second quadrant of KFSD, the blended analysis transition can be seen as the values transition from the 0.4 – 0.8 mm hr^{-1} bin to the 0.8 – 1.5 mm hr^{-1} bin as part of the blending algorithm.

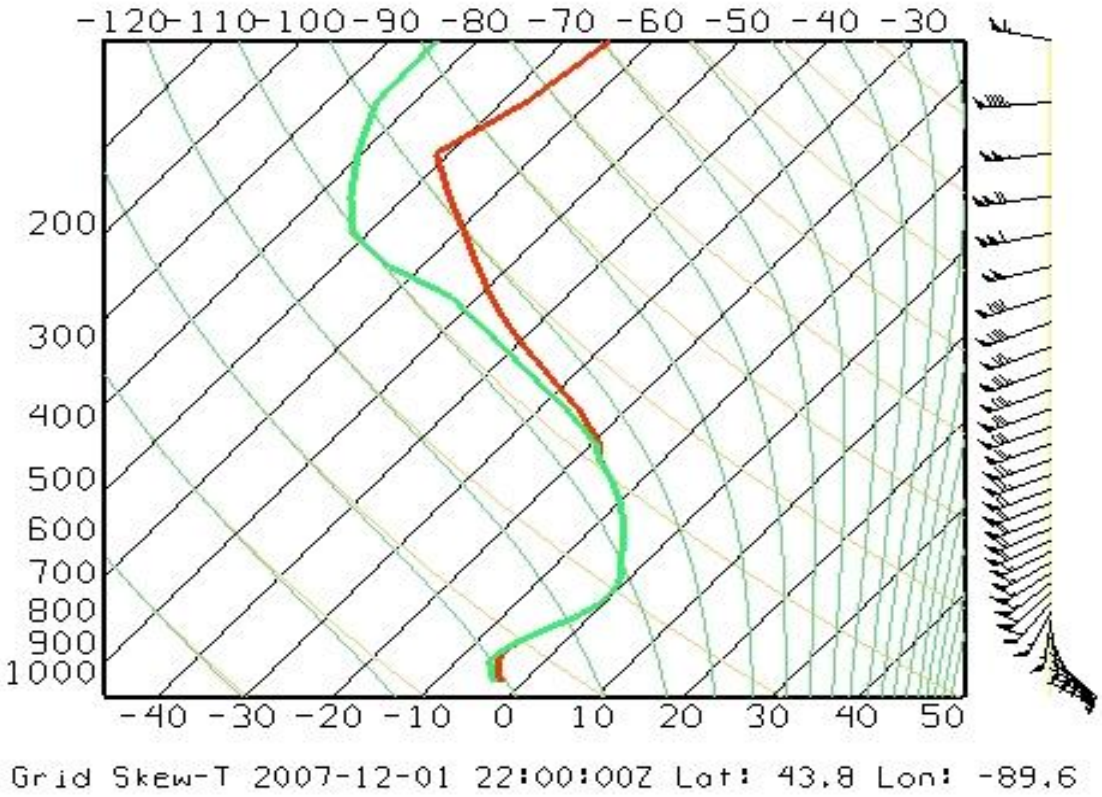

Figure 64. RUC-2 skew-T profile from a grid point extracted from 43.8° N and -89.6° W showing environmental temperature (°C) (red line) and dewpoint (°C) (green line)

Lastly, an aesthetically displeasing artifact is present near the radar site KBIS. The modified effective ranges for KBIS in the first and fourth quadrant were ~252 km and 155 km, respectively. While the fourth quadrant blended smoothly from the radar-based analysis to the surface-based analysis, the same cannot be said for the first quadrant. This effect occurs because where the radar blends to the surface, no precipitation is indicated by either analysis. Therefore, because the radar indicated no precipitation, a gradient appears as the surface-based analysis is given more weight as the distance away from the radar increases. Another strong-contrast difference exists centered on a line due north of KBIS. This results from the two modified effective ranges in these two neighboring quadrants being considerably different (by ~100 km).

As seen in Table 13, the blended analysis only performed better with regard to bias score. Otherwise, the blended analysis performance fell between the other two modules (mean error, RMSE, multiplicative bias), tied (POD, FAR, POFD), or below (accuracy, TS, correlation) that of the radar or surface analyses. The blended analysis did inherit the extra false alarm, KEST, from the surface analysis and inherited the miss (KMHE) present in the radar analysis.

The blended analysis failed to outperform the other analyses with regard to the precipitation rate metrics. In addition, the surface analysis outperformed the radar module for each of those four metrics. Because the surface module uses fixed precipitation rates derived from present weather, surface module values would inherently be closer to the observed values given the limited

number of categorical rates. Thus, the blended analysis appears to benefit from this and has precipitation rates that are closer to the withheld station's precipitation rates than the radar analysis estimates. As such, the question can be raised whether the verification routine favors one analysis over the others.

6.6 Radar/Surface Blending Method Comparison

Two methods, denoted as modes, exist for the blending algorithm and differentiate how the blending algorithm is run. The difference depends on whether surface-based values are used within the modified effective range before the algorithm starts the transition to the surface-based analysis. As described in the Methodology section, a radar's effective range is modified by identifying where precipitation overshooting is occurring. Values are then blended from radar-based values to surface-based values through a prescribed distance before the modified effective range is reached. In the original framework (denoted Mode 1), surface-based values are not used within the effective range of a radar because the radar-based analysis is trusted to be representative of what is occurring inside the radar envelope (the prescribed distance before the transition starts). The other mode (denoted Mode 2), fills in areas within the effective range of the radar where radar-based values are zero using surface-analysis values. Mode 2 favors precipitation occurrence. However, by design, the surface-based analysis fills in the analysis field when precipitation is denoted as occurring or not occurring using a guard-barrier. As such, when the station density is low, gaps between stations increase and the analysis becomes less reliable.

To test whether one mode is better than the other, the historical cases were run with both algorithms. Significance testing was completed to assess if the results were significantly different. The difference in performance was negligible (not statistically significant) for the set of historical cases. Event-average accuracy (Figure 65), FAR (Figure 66), POFD (Figure 67), POD (Figure 68), and ETS case (Figure 69) were similar for both domains. There are two possible reasons for these two approaches producing essentially the same results:

1) The change in the blending algorithm was not significant. More specifically, the area changed by the two modes is the prescribed distance before the transition starts, which can be relatively small depending on what the modified effective range is.
2) The validation scheme may not be capable of identifying the difference in performance because of the small sample (15 stations) used for verification.

Because the results showed no significant change in performance, Mode 1 was selected as the primary mode for comparisons and use in this study. Given how the algorithm was designed, it is logical to trust the radar inside the prescribed range before the transition starts. This is because overshooting is being identified in the algorithm and is used to demarcate the areas in which the radar data can be trusted.

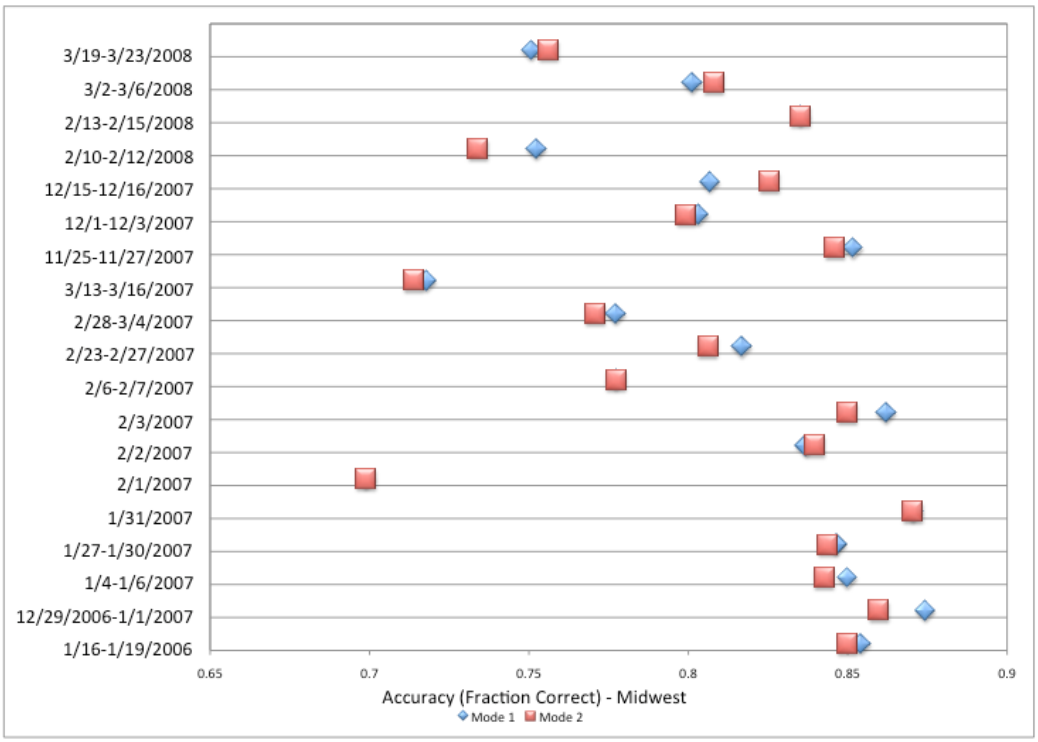

Figure 65. Event-average accuracy for Midwest domain depicting blended modes 1 and 2

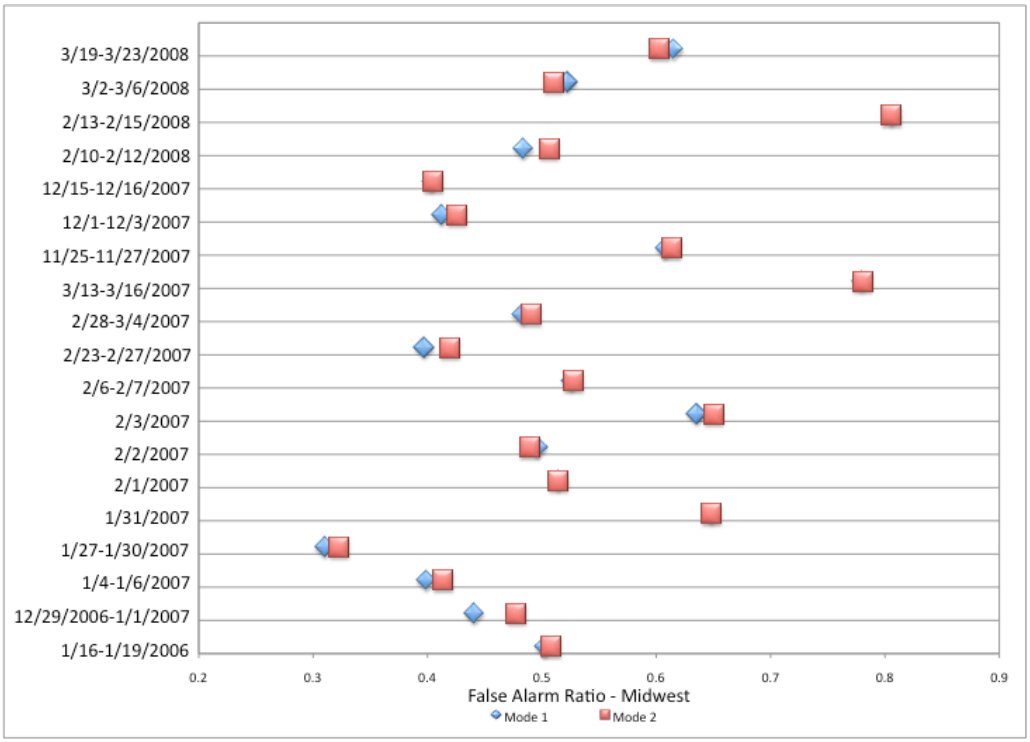

Figure 66. Event-average FAR for Midwest domain depicting blended modes 1 and 2

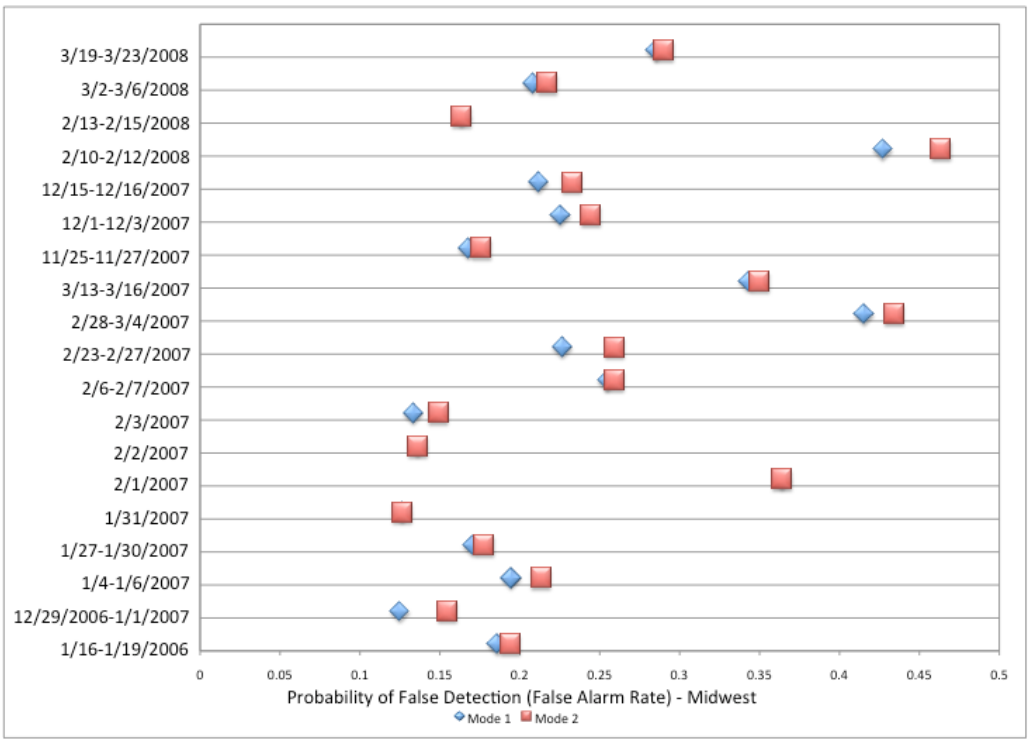

Figure 67. Event-average POFD for Midwest domain depicting blended modes 1 and 2

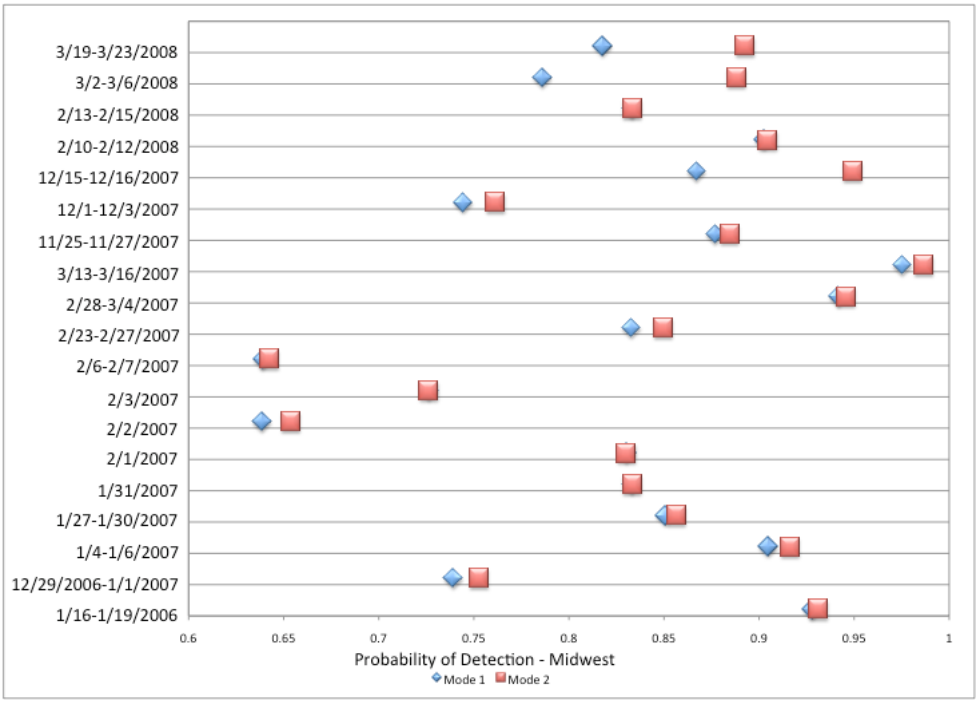

Figure 68. Event-average POD for Midwest domain depicting blended modes 1 and 2

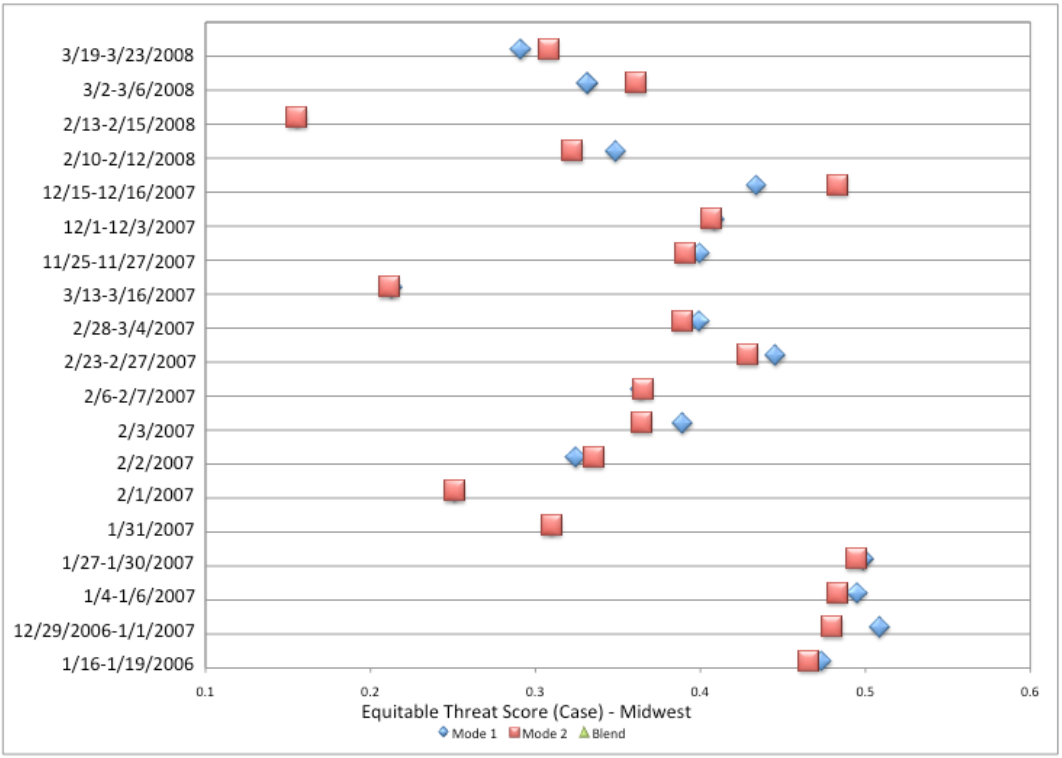

Figure 69. Event-average ETS (case) for Midwest domain, depicting blended modes 1 and 2

6.7 Radar/Surface Blending Module Performance Comparison

From the set of 19 historical cases, a wide variety of results was obtained. Both good and poor performance characteristics were seen from a performance metric and appearance standpoint. To provide more insight into the performance of the PPAES blending module, analyses at four additional times are considered. The PPAES radar and surface analyses that correspond to the blended analyses are also shown to provide insight as to how the PPAES blending module blended the two independent analyses. The four analyses are: 1455 UTC 18 January 2006 (Figure 70); 2155 UTC 31 January 2007 (Figure 71); 0955 UTC 1 February 2007 (Figure 72); and 1955 UTC 21 March 2008 (Figure 73).

In Figures 70 through 73, the blending module appears to work well, though locations of where the blended analysis used the radar and surface analyses can be identified due to the differences in resolution – with the surface analysis being coarser and the radar analysis being finer.

Figures 70 and 72 both show areas filled in by the surface analysis where the radar did not depict precipitation. This was desired as surface stations, by virtue of the produced surface analyses, were used to perform this function when precipitation was overshot. Also, when possible, the adjusted radar-based analysis values are used solely to retain the radar's improved resolution (Figure 73).

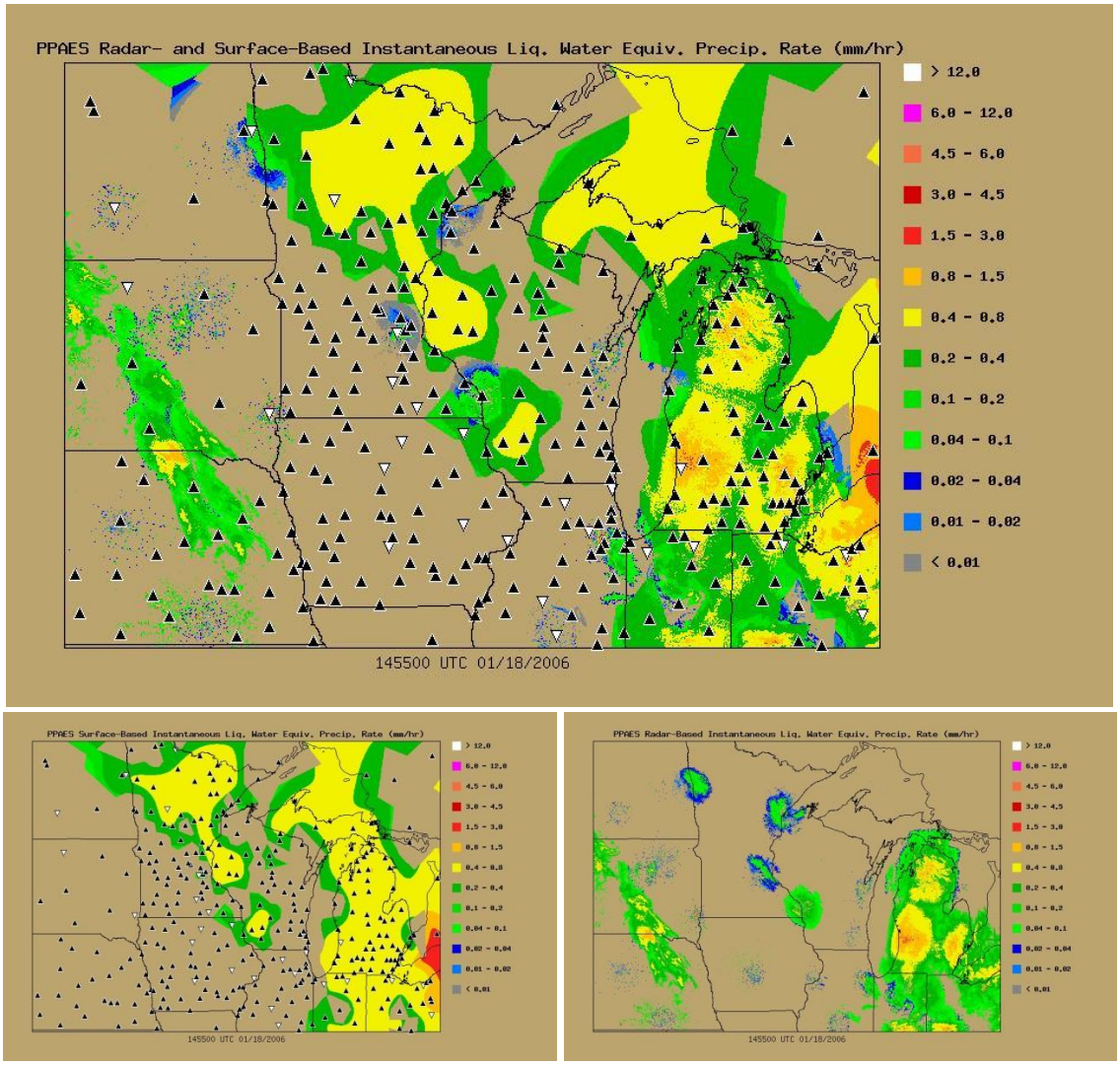

Figure 70. Analyses from 14:55:00 January 18, 2006 for PPAES blending module (top), surface module (lower-left), and radar module (lower-right)

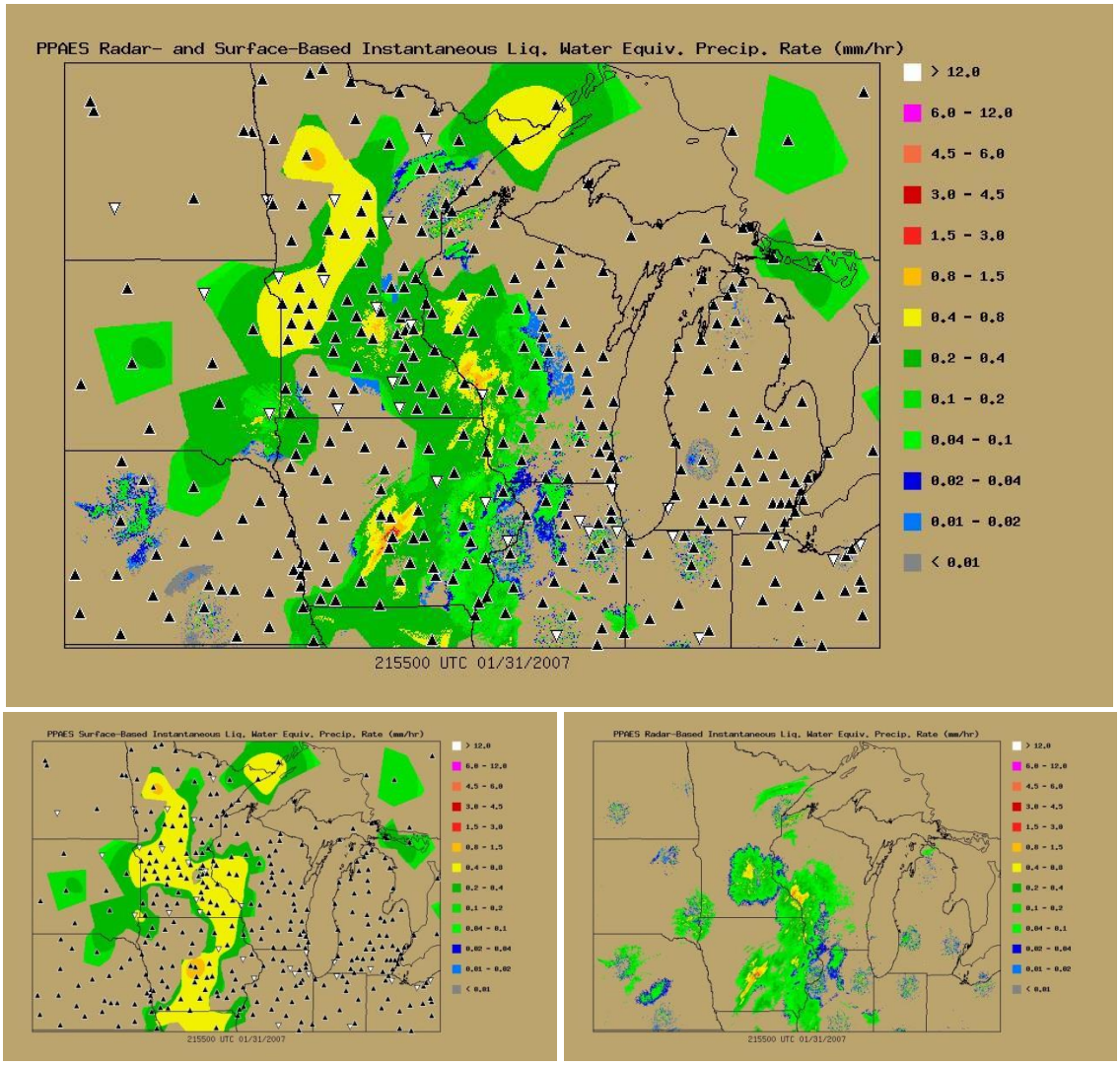

Figure 71. Analyses from 21:55:00 January 31, 2007 for PPAES blending module (top), surface module (lower-left), and radar module (lower-right)

In comparison to the analysis used for demonstrating the PPAES blending module (21:55:00 UTC 1 December 2007), these analyses lack the aesthetically displeasing artifact that was present in that analysis north of the Bismarck, ND, radar (KBIS). However, similar artifacts related to surface station density occur mainly in the Northern Great Plains domain, where station density is sparse.

Station density affects how well radar analyses can be blended with surface analysis by impacting the adjustment value and affecting effective range estimates. Station rich areas have more stations that could potentially modify the effective range and, from having more stations, have a corresponding surface analysis that allows smoother blending between the radar and the surface. Conversely, with sparse data, gaps between stations increase and the analysis becomes less reliable. In particular at 21:55:00 UTC on 1 December 2007, the lack of stations in central

ND limited the blending algorithm's effectiveness (Figure 62). Specifically, a station was not found to modify the effective range of KBIS in the first quadrant.

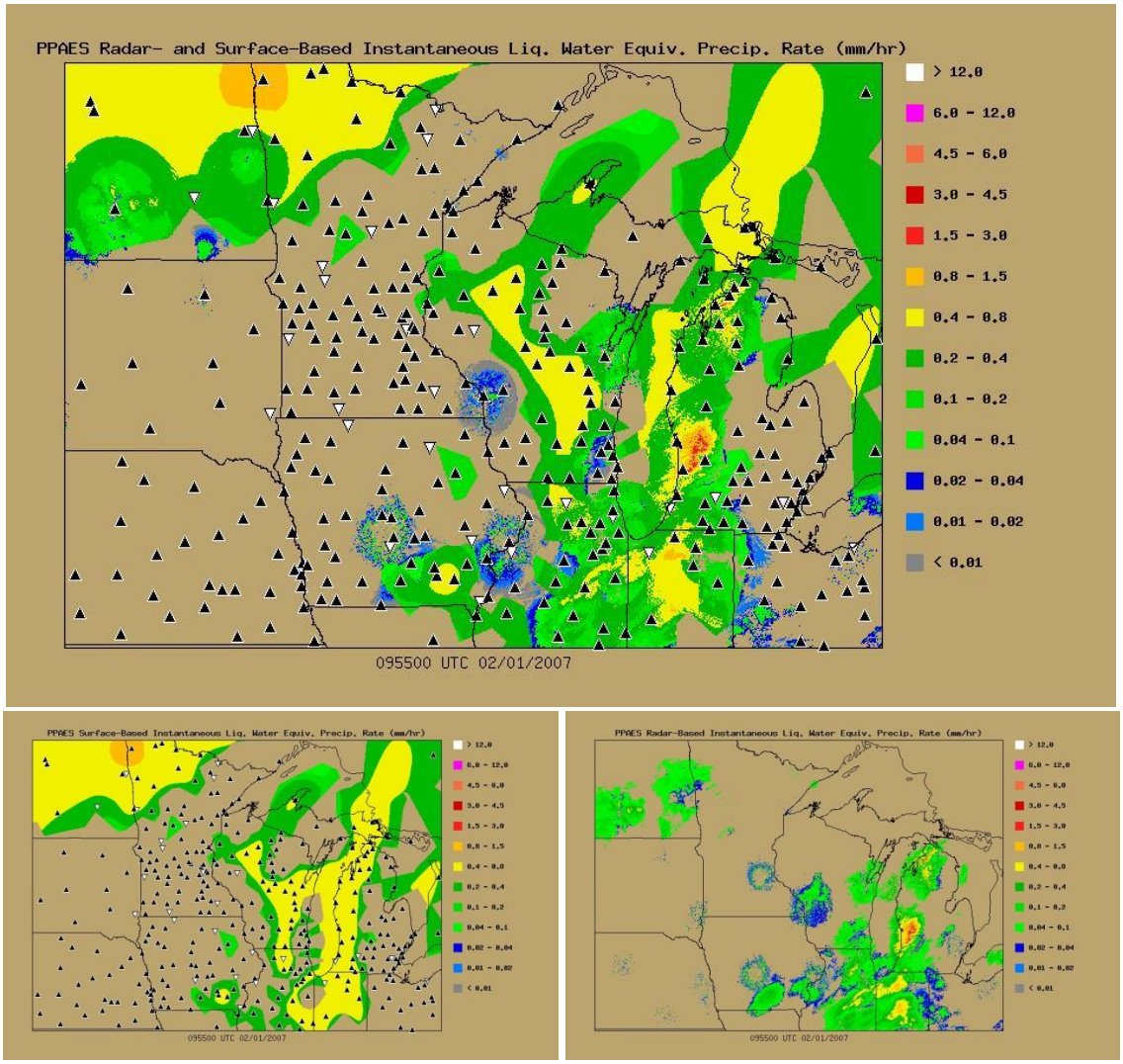

Figure 72. Analyses from 9:55:00 February 1, 2007 for PPAES blending module (top), surface module (lower-left), and radar module (lower-right)

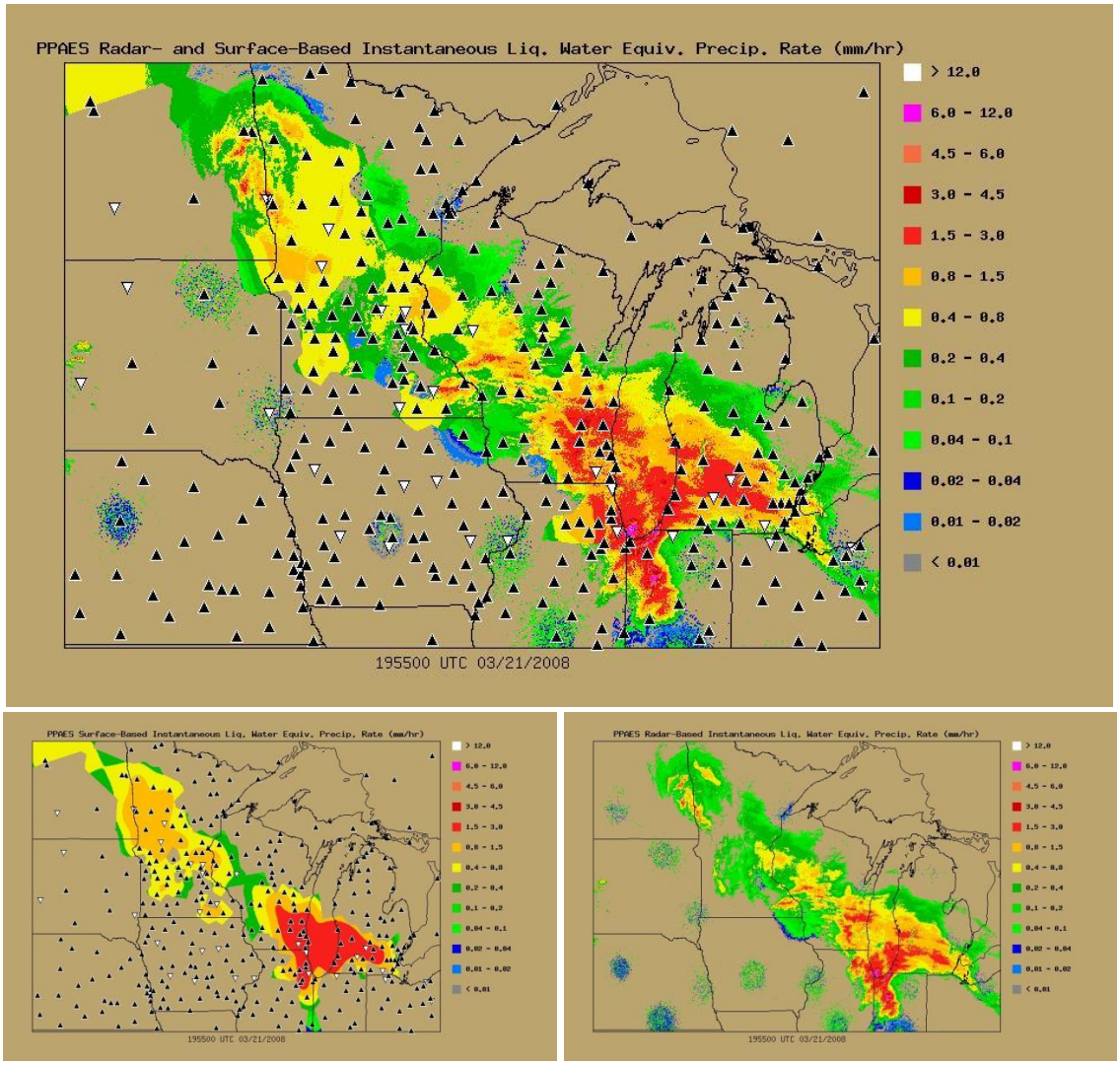

Figure 73. Analyses from 19:55:00 March 21, 2008 for PPAES blending module (top), surface module (lower-left), and radar module (lower-right)

Surface Transportation and PPAES

Information such as precipitation occurrence and rate are needed on the roadways since precipitation rate influences which treatment plan is used and when chemical roadway treatments will fail due to dilution (e.g., Mahoney 2003; Hallowel and Blaisdell 2003; Mahoney and Myers 2003). Chemical failure primarily depends on the road surface temperature and the liquid precipitation amount, with other secondary factors including runoff, chemical spray, and traffic (Hallowel and Blaisdell 2003; Mahoney and Myers 2003). The knowledge of where precipitation is occurring, among other variables, aids maintenance decision makers through its utility in Maintenance Decision Support Systems (MDSSs), which are designed to support roadway decision making through weather and road condition prediction combined with rules of practice for road anti-icing and deicing (Mahoney 2003).

PPAES currently provides liquid water equivalent precipitation rate information using areal plots. However, this information is critically needed along roadways. Such information can be obtained using PPAES routines to extract the liquid water equivalent precipitation rates from PPAES output files and map that information, using shape files, to Pooled Fund Study (PFS) MDSS test routes (Askelson 2008; Iteris 2011).

An example of the liquid water equivalent precipitation rate information derived for an ~34 mile length section of I-94 (Figure 74) is presented in Figure 75. The radar analysis values are lower than the blended analysis values because the values originating from KMVX had an adjustment for this hour of ~+0.66 mm hr^{-1}.

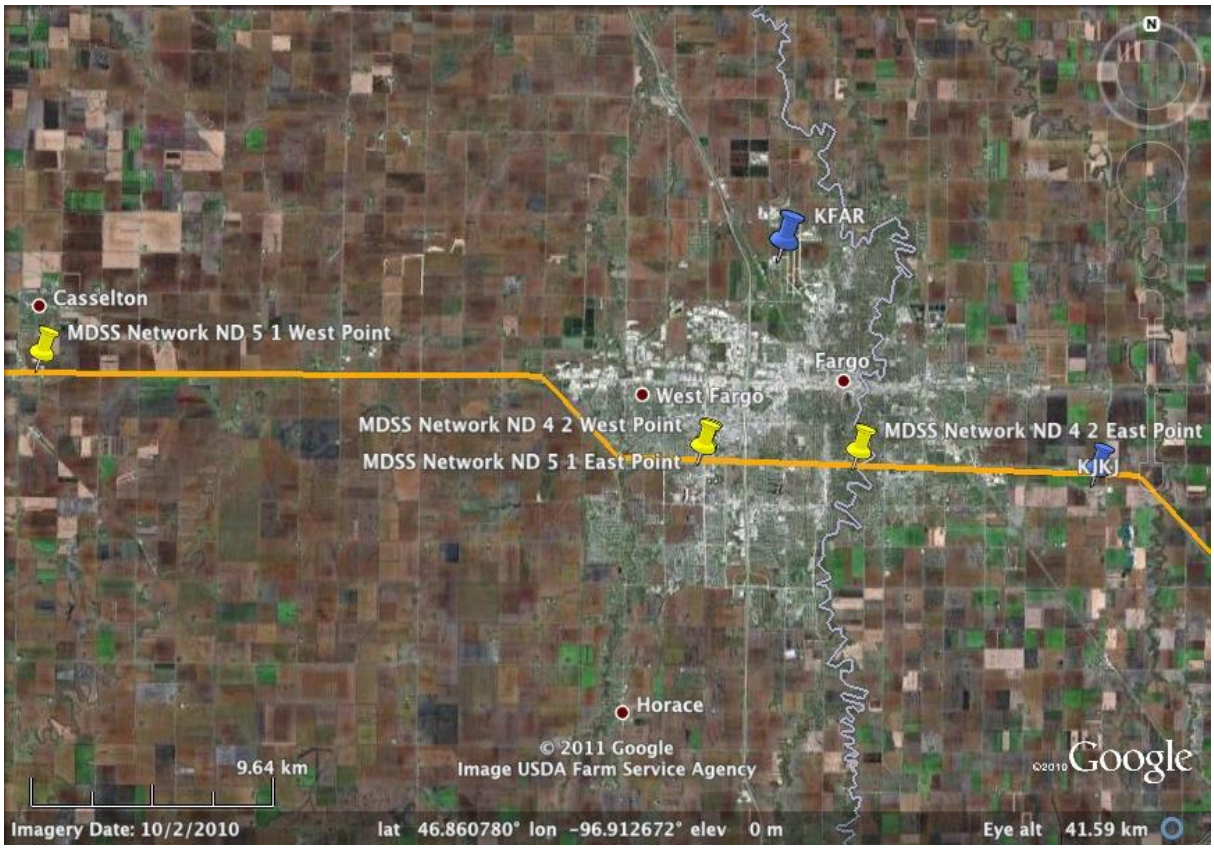

Figure 74. I-94 and associated I-94 PFS MDSS road segments

The two MDSS road segments of interest are plotted on I-94, their ends being denoted by their easternmost points.

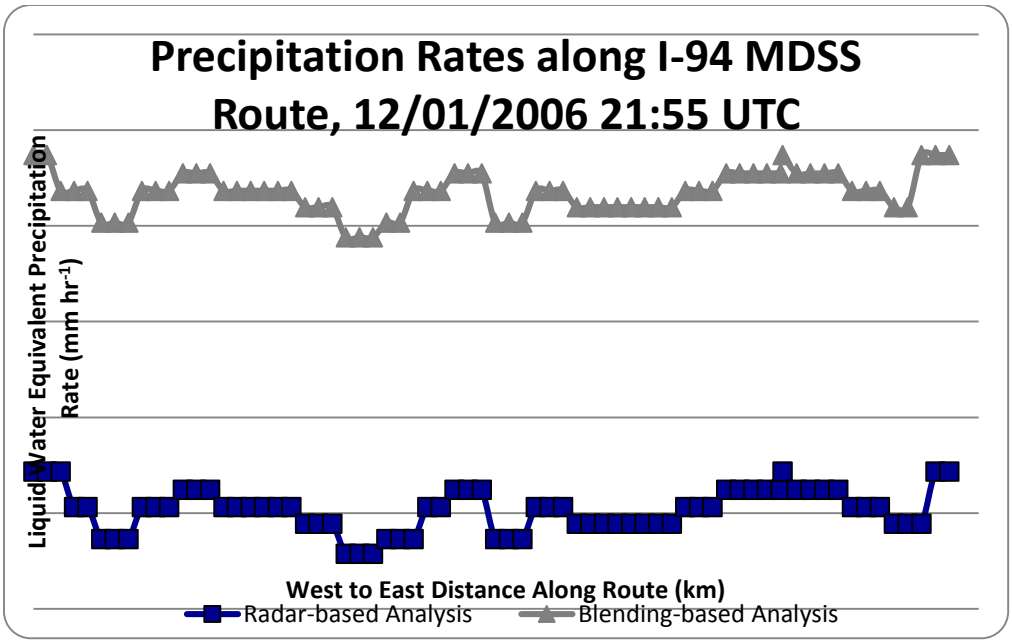

Figure 75. Liquid water equivalent precipitation rates extracted along an I-94 PFS MDSS route in ND from PPAES radar- and blending-based analyses

The liquid water equivalent precipitation rates are along ND-05-1 (from 0 km to 28 km along the route) and ND-04-2 (from 28 km to 34 km along the route), which are illustrated in Figure 72.

PPAES and Study Limitations

The PPAES modules developed in this study have several limitations. These limitations will be discussed, along with shortcomings of the validation approach.

6.8 Surface Analysis Precipitation Rates

Relative to radar analyses, surface analyses generally contain much less detail spatially (due to station density) as well as only having a maximum of three possible values to describe the precipitation rate. The station density in many areas appears to be generally insufficient to fully resolve the actual precipitation field, due to a lack of information. The problem is compounded further because present weather only has three corresponding intensities for most phenomena (RA, SN, DZ, and PE). As such, only three discrete rates can be used for phenomenon that in the real-world would have a continuous spectrum of values. In this study, all rates within an intensity category are treated as being equally likely, with the middle of the intensity category being used as the precipitation rate in an analysis. This assumes a uniform distribution, which results in the average precipitation rate producing the smallest average error. Accordingly, in the blending module this constraint and the spacing of the data result in discrete jumps, as can be seen when the surface analysis is blended with the radar analysis and vice-versa.

A possible way for providing more surface-observation-based values for snow is by using visibility and temperature information to compute rates (provided visibility and temperature pass the QC criteria). Rasmussen et al. (1999), for instance, provide a means for computing liquid equivalent snowfall rate using these two variables. This method involves assuming an average terminal velocity that depends upon the type of snow, which is estimated using temperature (Rasmussen et al. 1999). In addition, a constant based upon the assumed bulk snowflake density and diameter is used. Given this, Rasmussen et al. (1999) derived the relationship

$$S = \frac{1.3 C_3 \overline{V_t}}{Vis} \text{ cm s}^{-1} \tag{32}$$

where C_3 denotes a constant – the value of which is dependent on the type of snow (whether the snow is dry or wet), $\overline{V_t}$ denotes a constant average terminal velocity value in cm s^{-1} that is dependent on the type of snow (i.e., whether the snow is dry or wet), *Vis* denotes the visibility in cm, and *S* denotes the liquid equivalent snowfall rate in cm s^{-1}. For dry snow and wet snow, C_3 is equal to 0.017 g cm^{-2} and 0.072 g cm^{-2}, respectively. Average terminal velocity is equal to 100 cm s^{-1} and 200 cm s^{-1} for dry and wet snow, respectively. However, this equation assumes that the snowflakes are dry, spherical, and have a constant inverse relationship between their diameter and snowflake density (C_3) that depends on the type of snow (Rasmussen et al. 1999), which are rarely observed in the real world.

Alternatively, Rasmussen et al. (1999) provides theoretical relationships for various snow crystal types (over 20 types/variations). The results from two winter field sessions provided by Rasmussen et al. (1999) could be used to identify a best-fit line for a typical visibility-snowfall rate relationship. Such a relationship could be employed in place of a theoretical relationship for an individual crystal type. This could be advantageous considering the variability that can exist in an individual storm, specifically crystal types and whether aggregation is present or absent (Rasmussen et al. 1999).

Several drawbacks exist to these empirical equations as pointed out by Rasmussen et al. (1999). These drawbacks include: inner-storm variability during different time periods, variety of crystal types, degrees of riming, degrees of aggregation, degrees of crystal wetness, and natural variability of snow type (Rasmussen et al. 1999). A large degree of scatter was observed in the presented visibility-snowfall rate relationships, which can be attributable to the above complications. Similarly, Rasmussen et al. (1999) findings re-affirmed the large degree of scatter that exists in *Vis*-*S* relationships observed in past studies (e.g., Stallabrass 1985 and Fujioshi et al. 1983). As indicated by their work, the use of visibility to estimate liquid equivalent snowfall rate can be misleading due to the large degree of variability that can exist. Furthermore, *Vis*-*S* relationships are not one-to-one–different liquid equivalent snowfall rates can have the same visibility. This can be attributed to multiple factors also influencing visibility – wind speed and blowing snow, other phenomena (e.g., fog), and crystal size (Rasmussen et al. 1999). As such, the consequences of assuming a *Vis*-*S* relationship must be understood before using visibility as an indirect method of acquiring a liquid equivalent snowfall rate.

In future work it would be advantageous to explore the use of (32) or a different visibility-snowfall rate relation. Such an effort would provide extra detail in the surface analyses by ameliorating quantization-type errors.

6.9 Azimuthal Interpolation of Radar Effective Range

Two methods for modifying a radar's effective range were examined in this study: quadrant and azimuthal. Each method has advantages and disadvantages. The quadrant method is computationally easier, but produces some unrealistic artifacts that can be seen in Figure 62 in north central North Dakota (north of Bismarck, ND). Herein, it was determined that modifying the effective range azimuthally may be more realistic. The result for this example is shown in Figure 76. (Note that although the time indicated in Figure 76 is 220000 UTC as opposed to 215500 UTC in Fig. 62, the same surface data file was used to produce both. Prior to creating Figure 76, the method for calculating the reference time was modified, which produced this slight difference.) As is apparent in this figure, the analysis is smoother and these artifacts are not visible.

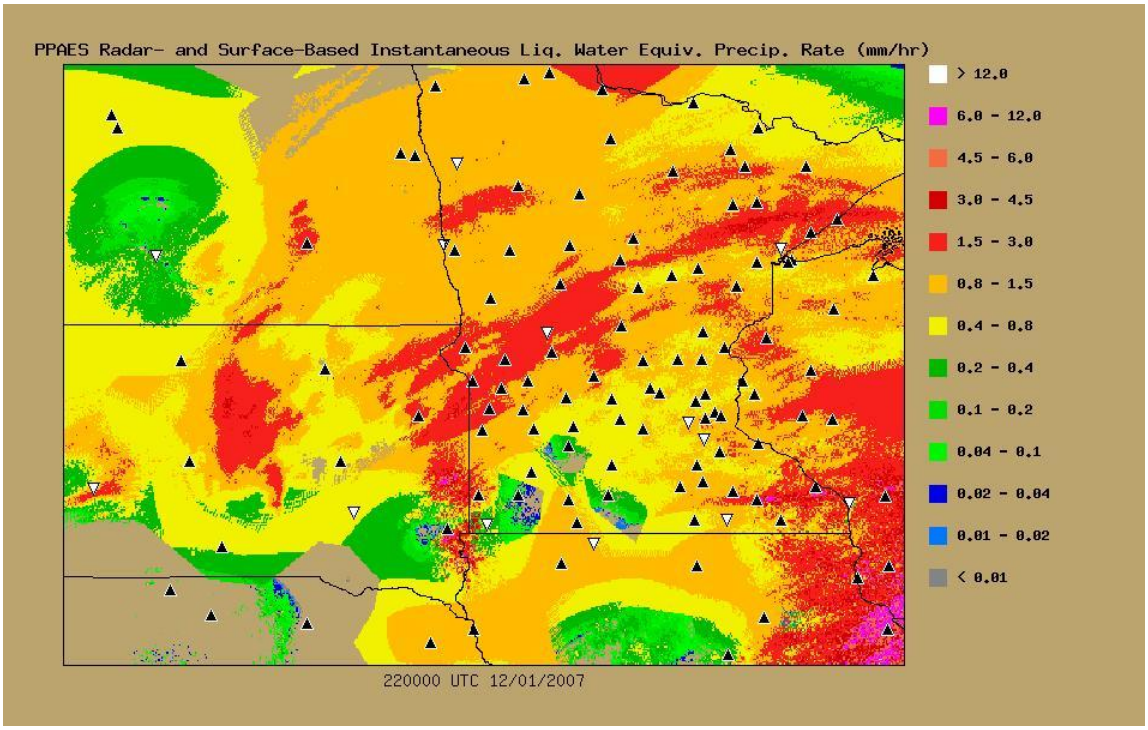

Figure 76. PPAES blended analysis at 22:00:00 UTC December 1, 2007 over Northern Great Plains domain using effective ranges modified azimuthally (annotation similar to Figure 62)

One limitation to the azimuthal modification approach for effective range is that if stations that are modifying the effective range are located relatively close azimuthally to each other and the difference in their modified effective ranges is large, a discontinuity can arise between those two azimuths. Further monitoring of these types of scenarios is needed. However, it has been found

that the azimuthal modification of effective range generally provides more acceptable results in the blending routine.

6.10 Virga

The primary effort in this study involved blending information from multiple platforms (radar and surface) to gain a more accurate depiction of wintertime precipitation occurrence/non-occurrence and rate. This was accomplished through alleviating a principle limitation of radar when measuring snowfall: overshooting. The blending algorithm employed in this study uses surface observations to determine where radars are overshooting, and radar analyses are blended to surface analyses starting nearby where overshooting is identified. However, this study did not account for virga – a significant limitation associated with radar. The utility of the current PPAES radar and blending algorithm can be impacted by this phenomenon. The effect of virga on radar estimates is well established. Virga produces false indications of precipitation and, thus, overestimated precipitation rates (Doviak and Zrnić 1993, p. 225; Super and Holroyd 1998; Hunter et al. 2001). Although addressing this issue is beyond the scope of this study, methods to identify and mitigate virga have been considered.

Virga stems from areas of dry air resulting in the total evaporation or sublimation of descending liquid- and ice-hydrometeors that would have otherwise fallen to the surface. This poses a problem for radar, as the radar samples hydrometeors above the ground and can indicate precipitation even though the sampled hydrometeors would not reach the ground. In such situations, radar analyses will then provide false indications of precipitation (Hunter et al. 2001). As such, several methods for mitigating this issue involving the use of radar data, as well as other sources of information, have been identified. These include the following:

1) Utilizing other surface observation datasets like Clarus as "ground truth" and eliminating these false indications of precipitation by modifying the effective radar range.
2) Utilizing model sounding data to diagnose dry air through an examination of the relative humidity profile to eliminate false indications of precipitation (Hunter et al. 2001).
3) Utilizing multiple radar elevation scans to eliminate false indications of precipitation (Hunter et al. 2001).
4) Utilizing derived satellite products like the GOES sounder temperature and moisture profiles (Schmit et al. 2002). Alternatively, moisture profiles from the COSMIC GPS-RO data provide another source of information (Anthes et al. 2000; Anthes et al. 2008; Lin et al. 2010).

Each of the above methods has limitations. Limitations with surface observations include determining if the surface observation can serve as "truth" given the limitations associated with precipitation gauges (e.g., wind driven effects; Super and Holroyd 1998) as well as with LEDWIs (e.g., difficulty with measuring light precipitation; Wade 2003). Model limitations are associated with shortcomings in modeling the planetary boundary layer (Atmospheric Policy Program American Meteorological Society 2003) and determining the humidity thresholds for virga production. Radar data limitations include incomplete sampling to discern whether the

phenomenon is virga or precipitation. Satellite limitations include a clear field of view being required for a GOES sounder profile to be obtained (Schmit et al. 2002).

6.11 Surface Station and Analysis Time Differences

A potential error has been identified in that some surface station reports may not be representative at the time of a PPAES analysis. This is a limitation in that a 30 minute (1,800 second) time window is currently used in the PPAES surface module. A 30 minute time window was used to ensure that the maximum amount of precipitation information provided by the surface dataset is used. Figures 77 and 78 display the absolute values of the analysis time minus the station time of report (in seconds) for each surface observation used in the PPAES analyses.

As seen in Figures 77 and 78, several stations provide reports outside of the typical reporting time window. Routinely, surface stations report 10 minutes before the intended reporting hour (FMH-1 2005). In this study, surface stations have been found to routinely report five and ten minutes before the hour as the majority of the time differences are <300 seconds. Thus, errors due to temporal non-representativeness are expected to be relatively small. Note that some stations do report late (or early). Such a case is evident in Figure 77, where a station in eastern ND reported ~14 minutes outside of the analysis time

Ways to mitigate this issue include using an objective analysis scheme that temporally weights the observations and reducing the size of time window. The former approach would increase the complexity of the objective analysis scheme used in the surface module. The latter could result in useful data being thrown out. A sensitivity study could be used to determine an optimal time window value. However, such an investigation is beyond the scope of this study.

Time differences (in seconds) are shown in Figures 77 and 78 using each observation station's computed region of influence.

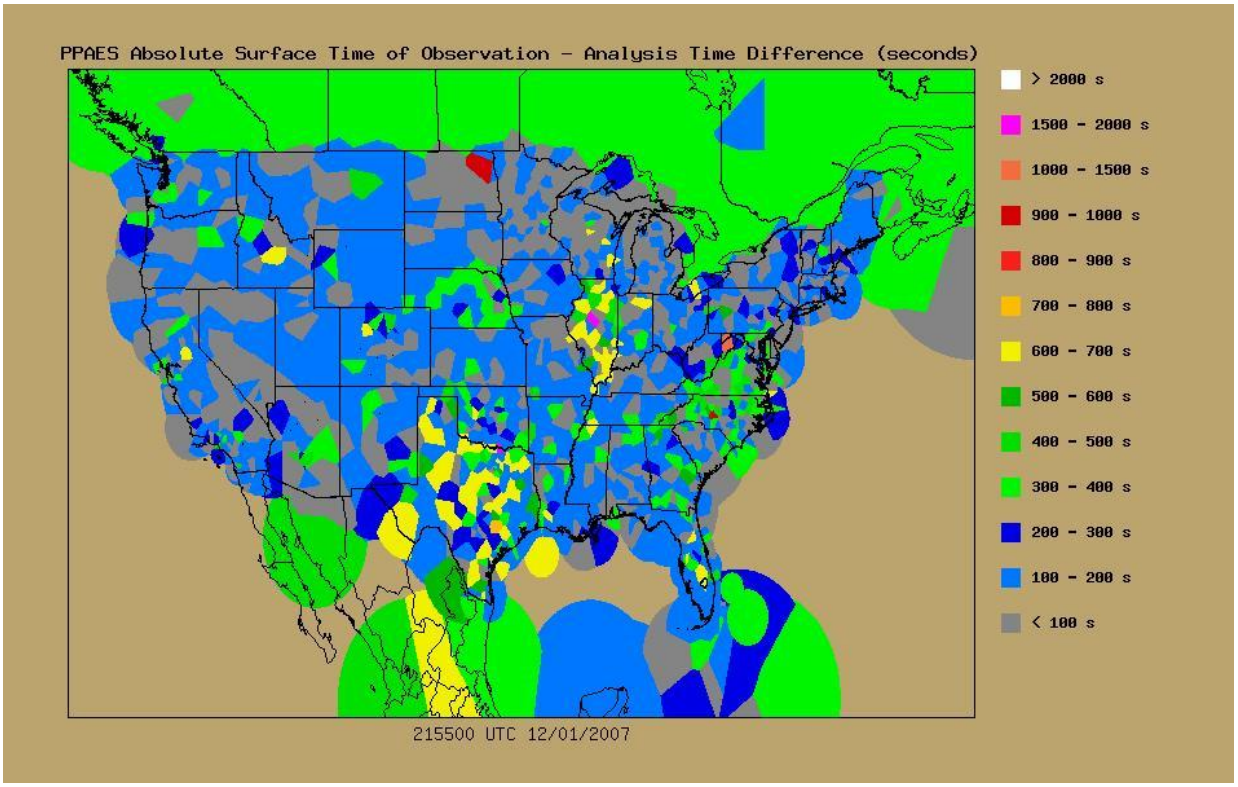

Figure 77. Absolute value of analysis time minus station time of report for PPAES surface module over the CONUS at 21:55 UTC December 1, 2007

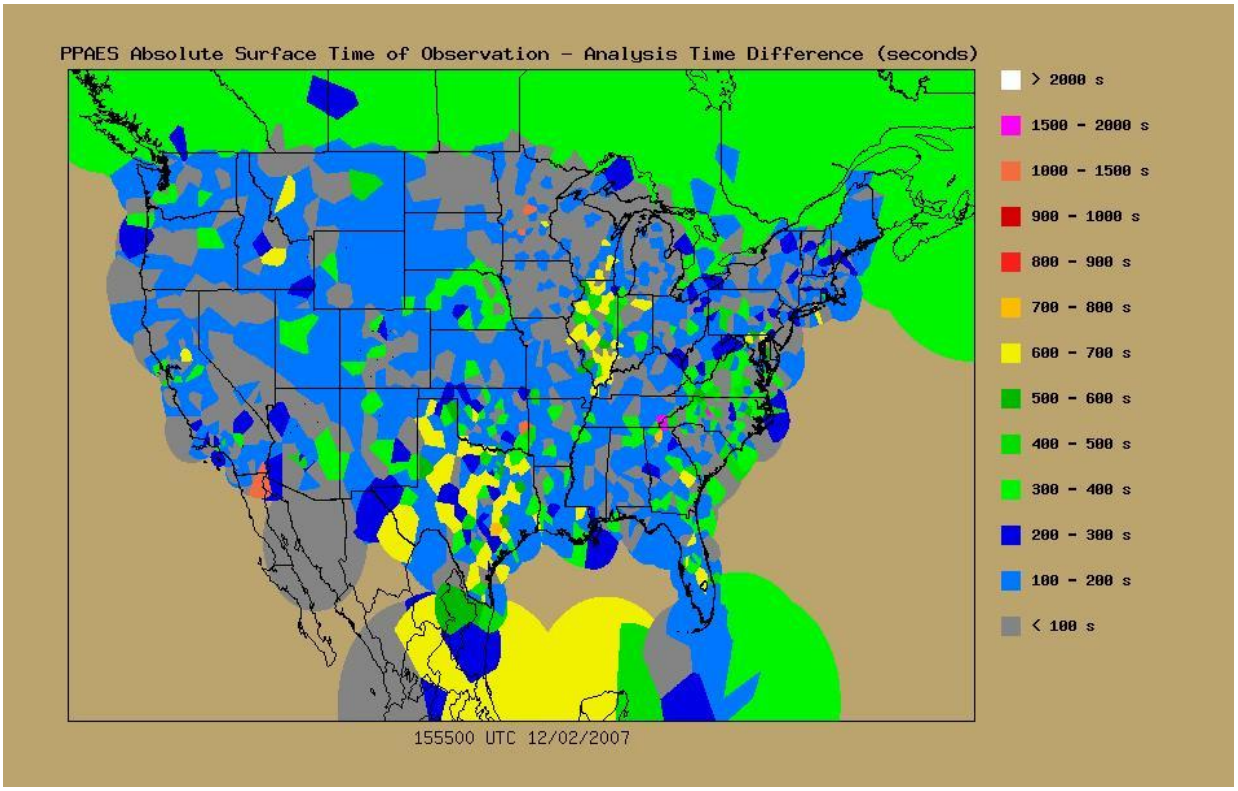

Figure 78. Absolute value of analysis time minus station time of report for PPAES surface module over the CONUS at 15:55 UTC December 2, 2007

6.12 Validation Limitations

The validation approach used has three limitations that must be acknowledged: significance testing, the subjective nature of the characterization of the cases, and an aspect of the verification methodology used. These areas are discussed to provide further insight and to suggest possible areas of future work to make the validation more robust.

Statistical tests were run for the set of historical cases. A limitation exists in that 19 cases were identified. This number provides a significant constraint on the statistical tests that can be used and what conclusions can be drawn from them. With a small sample, outliers can have a strong effect and assumptions may or may not be violated. If one of the underlying assumptions is significantly violated, a different statistical test should be used.

The two statistical tests utilized here are the two-sample *t*-test (with equal variances assumed) and one-way ANOVA. Both tests share similar assumptions: the samples are random, independent, come from normal populations, and have equal population standard deviations (or variances) (Weiss 2008). We examined the validity of these assumptions as they relate to these data. Specifically, the normality and equal population standard deviation conditions were checked through the "Rule of 2" and the standard deviation ratio test (e.g., Table 14), normal probability plots (e.g., Figure 79), and box plots (e.g., Figure 80) (Weiss 2008). These reveal that no gross violations existed. Furthermore, if an outlier was noticed that could affect a statistical

test, that test was run both with the case present and removed to see if the test statistic was adversely affected by the case. This occurred most notably with event-average POD, where the removal of the outlying case (Case 7 – 2/2/2007) actually increased the test statistic values in both significance tests.

Table 14. Ratio of standard deviations "Rule of 2" for event-average ETS (case) and FAR used to check equal standard deviation assumption in statistical significance tests

Ratio of Standard Deviations (Rule of Two)			
ETS (Case)	**Surface**	**Radar**	**Blend**
Surface	-	0.959	0.954
Radar	1.043	-	0.995
Blend	1.048	1.005	-
FAR	**Surface**	**Radar**	**Blend**
Surface	-	1.240	1.120
Radar	0.807	-	0.904
Blend	0.893	1.107	-

The condition is considered to be met if the ratio of the larger to the smaller sample standard deviation is less than two (Weiss 2008). For completeness, the ratio of the smaller to larger sample standard deviation value is also shown.

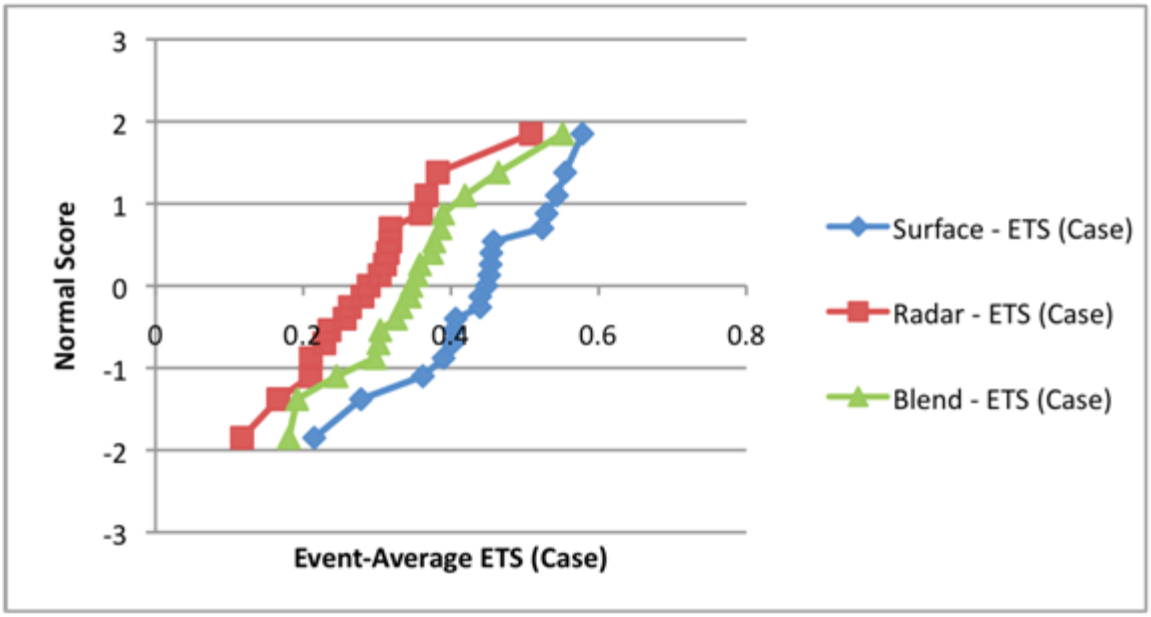

Figure 79. Normal probability plots for event-average ETS (case) used to check normality assumption in statistical significance tests

If the plot is linear, the normality condition can be considered met (Weiss 2008).

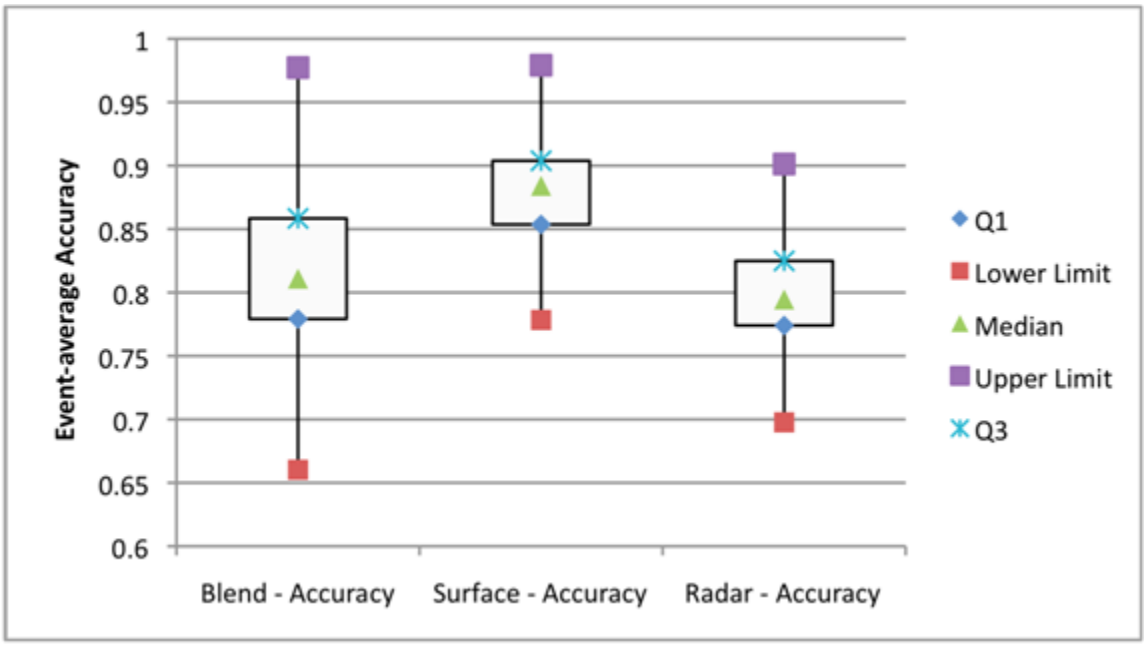

Figure 80. Box plots for event-average accuracy used to check normality assumption in the one-way ANOVA and the two-sample *t*-tests with equal variances assumed

For some two-sample *t*-tests, the test statistic landed in the non-rejection region of the *t*-curve. This likely resulted from not having a large enough sample of cases and we could not conclude, at the significance level α=.05, that the two modules' performance metric mean values came from two populations that were different.

With regard to the verification methodology, cases were subjectively characterized. As such, this is a limitation because different classifications could be obtained by using different classification methods. To mitigate this, several resources and criterion were used to characterize the events. However, difficulties still exist because of the classification's subjective nature. Furthermore, because an in-depth mesoscale analysis could not be completed we are limited in applying the results to events primarily driven by synoptic-scale processes, since different mesoscale processes that can drive precipitation can be present in an event.

Another important aspect of the verification methodology was the method for selection of stations that were withheld and used for verification. The method in this study entails withholding 15 stations, where seven and eight stations were selected randomly and manually. As such, the performance metrics may be sensitive to the method used to select the stations and the total number of stations withheld. For instance, a surface station used for verification may be located in an area prone to receive or not receive precipitation for a particular case via random selection. The impact of such a location is that in an area prone to not receive precipitation (e.g., a desert location), the accumulated statistics for a case could be inflated. Such a station could be expected to have higher scores through the accumulation of correct negatives. In contrast, choosing a station located in an area prone to receive more precipitation than other stations could cause lower performance metric scores because misses would be more likely.

Also, two scenarios that could adversely affect the performance of a station when precipitation is occurring include: (a) light precipitation falling below the LEDWI's threshold or (b) mist/blowing snow masking the precipitation. These scenarios could result in the surface station not reporting precipitation, which would adversely impact the statistics for the radar or blending analysis if they were to indicate precipitation.

Furthermore, recall that 15 stations were withheld from both domains to use for verification. The reason behind selecting 15 stations was that selecting too many stations could adversely impact the produced analyses. The selection of the number of stations used was somewhat arbitrary but, as was demonstrated, statistical significance was obtained, which suggests that this value is appropriate.

Overall, the surface module performed better than the other two modules. However, an issue may exist as to whether one analysis is favored over the others due to the verification methodology used. A shortcoming with the verification routine is that surface station present weather measurements are used solely. As such, only three distinct rates are possible for a given weather phenomena, in place of a possible continuous set of values that are naturally observed. With a *Vis-S* relationship, this problem can be mitigated. As such, with the data that are available, surface station's present weather provide valuable information for precipitation rate – albeit with caveats.

Other issues associated with LEDWI sensors such as blowing snow conditions, a light precipitation threshold, and the detection of non-atmospheric phenomena, can adversely affect a surface measurements used for verifying analyses. These surface station deficiencies bring in to question the validity of the performance metric scores--a false alarm or a miss can result from a verification station not detecting the present weather correctly.

6.13 Conclusions

This study used model output and observations from multiple observation platforms—radars and *in situ* surface stations—to develop PPAES algorithms for producing enhanced estimates of precipitation fields (specifically precipitation occurrence and liquid water equivalent precipitation rate). The methods that were developed blends independently produced precipitation rate analyses from model output, radar measurements, and surface observations into fused analyses. Given that model output and radar- and surface-based measurements suffer from their own shortcomings, the strategy was to leverage strengths from each of these data sources. Radar beam overshoot, a fundamental weakness associated with radar measurements in the wintertime (Fulton et al. 1998; Super and Holroyd 1998; Askelson 2008), was targeted as a primary limitation to be mitigated in this study. Through this effort, ways to blend PPAES radar, surface, and model analyses to provide improved estimates of precipitation rates and occurrences were developed.

To evaluate the radar+surface blending algorithm, a set of 43 historical cases covering two domains was used. Blended analyses were compared to corresponding surface and radar analyses to determine relative performance—skill scores and performance metrics were computed to

evaluate all three types of analyses. Significance testing was also performed to determine if differences for each module and for select performance metrics were statistically significant.

Five cases centered over the Utah region were used to evaluate algorithms that were developed to ameliorate negative impacts associated with complex terrain. These included an enhancement of the PPAES radar module that involved detection of the potential interaction of radar rays with the ground and utilization of higher-elevation data, and blending of radar analyses with model output. Skill scores and performance metrics were computed to evaluate the performance of these algorithms.

The main results obtained were as follows:

1) The PPAES radar+surface blending module provided an improved probability of detection compared to the other analyses and this difference in performance is statistically significant at the 5% level.

2) The PPAES surface module had a higher accuracy, lower probability of false detection, false alarm ratio, and higher ETS (case-computed) compared to the radar and radar+surface blending modules.

3) Regarding event-average performance overall, the PPAES radar module failed to outperform the PPAES radar+surface blending and surface modules.

4) Radar+surface blending module event-average performance overall was mixed: POD improved, but ETS (case-computed) event-averages were seen to be intermediate between the surface and radar analyses for this set of historical cases. Furthermore, several cases had event-average performance metrics where the radar+surface blending analysis performed worse relative to the other module-produced analyses.

5) Inclusion of Clarus data degraded analysis quality. While the exact cause of this is not known, possible reasons for this include different performance characteristics relative to ASOS/AWOS precipitation sensors and maintenance practices.

6) Use of a quadrant-based approach for defining the effective ranges of radars when blending radar and surface data can result in aesthetically displeasing artifacts. Consequently, an approach in which the effective range of a radar varies continuously with azimuth was developed. This approach removes the types of artifacts that arise with the quadrant method.

7) Because 1) WSR-88D radars were sited such that radar beam blockage at the lowest elevation is relatively minor, resulting in relatively small differences between radar-based analyses produced with and without terrain clearance, and 2) the surface observation network available for verification is relatively sparse, no significant enhancement in performance was observed when terrain clearance was used. This capability, however, could significantly enhance analyses in non-standard radar ray propagation conditions.

8) With the radar+model blending approach and the limited number of test cases used herein, the addition of model data resulted in degradation of performance from the standpoint of the metrics used herein. However, the use of model data does have the advantage of presenting a more continuous and spatially-coherent representation of the precipitation field to the user.

REFERENCES

Achtemeier, G. L., 1986: The impact of data boundaries upon a successive corrections objective analysis of limited-area datasets. *Mon. Wea. Rev.*, **114**, 40-49.

Anthes, R. A., and Coauthors, 2008: The COSMIC/FORMOSAT-3 mission: early results. *Bull. Amer. Meteor. Soc*., **89**, 313–333.

Anthes, R. A., C. Rocken, Y. –H. Kuo, 2000: Applications of COSMIC to meteorology and climate. *Terrestrial, Atmospheric and Oceanic Sciences*, **11**, 115-156.

Arnold, Charles P., Clifford H. Dey, 1986: Observing-Systems Simulation Experiments: Past, Present, and Future. *Bull. Amer. Meteor. Soc.*, **67**, 687–695.

Askelson, M. A., J. Aubagnac, and J. M. Straka, 2000: An adaptation of the barnes filter applied to the objective analysis of radar data. *Mon. Wea. Rev.*, **128**, 3050-3082.

Askelson, M. A., 2002: Kinematic, dynamic, and thermodynamic impacts of hook-echo hydrometeors, including explorations into the utilization of polarimetric radar data. Ph.D. Dissertation, University of Oklahoma, 258 pp.

Askelson, M. A. and J. M. Straka, 2005a: Response functions for arbitrary weight functions and data distributions. Part I: Framework for interpreting the response function. *Mon. Wea. Rev.*, **133**, 2117-2131.

Askelson, M. A., Pauley, P. A., J. M. Straka, 2005b: Response functions for arbitrary weight functions and data distributions. Part II: Response function derivation and verification. *Mon. Wea. Rev.*, **133**, 2132-2147.

Askelson, M. A., 2008: The Pavement Precipitation Accumulation Estimation System. *Extended Abstracts, Surface Transportation Weather and Snow Removal and Ice Control Technology*, Indianapolis, IN, Transportation Research Board, 544-557.

Baer, F., and J. J. Tribbia, 1976: Spectral fidelity of gappy data. *Tellus*, **28**, 215-227.

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

Barnes, 1973: Mesoscale objective analysis using weighted time-series observations. NOAA Tech. Memo. ERL NSSL-62, National Severe Storms Laboratory, Norman, OK 73069, 60 pp.

Barnes, 1994: Applications of the Barnes objective analysis scheme. Part I: Effects of undersampling, wave position, and station randomness. *J. Atmos. Oceanic Technol.*, **11**, 1433-1448.

Benjamin, S. G., J. M. Brown, K. J. Brundage, B. E. Schwartz, T. G. Smirnova, T. L. Smith, and L. L. Morone, 1998: RUC-2 – The Rapid Update Cycle Version 2. NWS Technical Procedures Bulletin 448, 18 pp. [Available online at http://www.weather.gov/om/tpb/448.pdf

Benjamin, Stanley G., Barry E. Schwartz, Steven E. Koch, Edward J. Szoke, 2004: The Value of Wind Profiler Data in U.S. Weather Forecasting. *Bull. Amer. Meteor. Soc.*, **85**, 1871–1886.

Benjamin, Stanley G., Brian D. Jamison, William R. Moninger, Susan R. Sahm, Barry E. Schwartz, Thomas W. Schlatter, 2010: Relative Short-Range Forecast Impact from Aircraft, Profiler, Radiosonde, VAD, GPS-PW, METAR, and Mesonet Observations via the RUC Hourly Assimilation Cycle. *Mon. Wea. Rev.*, **138**, 1319–1343.

Bluestein, H. B., 1993: *Synoptic-Dynamic Meteorology in Midlatitudes.* Vol. 2, *Observations and Theory of Weather Systems*. Oxford University Press, 594 pp.

Bohren, C. F., and L. J. Battan, 1980: Radar backscattering by inhomogeneous precipitation particles. *J. Atmos. Sci.*, **37**, 1821-1827.
Bourgouin, P., 2000: A method to determine precipitation type. *Wea. Forecasting*, **15**, 583-592.
Breidenbach, J. P. and J. S. Bradberry, 2001: Multisensor precipitation estimates produced by national weather service river forecast centers for hydrologic applications. *Extended Abstracts*. *Georgia Water Resources Conference,* Athens, GA, Georgia Water Resources Institute, 179-182. [Available online at http://www.gwri.gatech.edu/uploads/proceedings/2001/BreidenbachJ-01.pdf.]
Brock, F. V. and S. J. Richardson, 2001: *Meteorological Measurement Systems.* Oxford University Press, 290 pp.
Cardinali, Carla, Lars Isaksen, Erik Andersson, 2003: Use and Impact of Automated Aircraft Data in a Global 4DVAR Data Assimilation System. *Mon. Wea. Rev.*, **131**, 1865–1877.
Carroll, T., D. Cline, C. Olheiser, A. Rost, A. Nilsson, G. Fall, C. Bovitz, L. Li., 2006: NOAA's national snow analyses. *Proc.* 74th Annual Meeting of the Western Snow Conference, Las Cruces, New Mexico, Western snow Conference, 14 pp. [Available online at http://www.nohrsc.noaa.gov/technology/pdf/WSC_2006.pdf.]
CAWCR, cited 2011: NCAR advanced study program summer colloquium: forecast verification in the atmospheric sciences and beyond. [Available online at http://www.cawcr.gov.au/projects/verification/verif.html]
Chungu, L., G. L. Browning, 1998: The impact of observation errors on objective analysis. *J. Atmos. Sci.*, **55**, 1791-1807.
Clarus Initiative, cited 2010: Clarus initiative. [Available online at http://www.clarusinitiative.org/.]
Cressman, G. P., 1959: An operational objective analysis system. *Mon. Wea. Rev.*, **87**, 367-374.
Crum T. D., R. L. Alberty, and D. W. Burgess, 1993: Recording, archiving, and using WSR-88D data. *Bull. Amer. Meteor. Soc.*, **74**, 645-653.
Daley, R., 1991: *Atmospheric Data Analysis*. Cambridge Atmospheric and Space Science Series, Vol. 2, Cambridge University Press, 457 pp.
Dingman, S. L., 2002: *Physical Hydrology*. 2nd ed. Waveland Press, Inc., 646 pp.
Doswell, C. A. III and S. Lasher-Trapp, 1997: On measuring the degree of irregularity in an observing network. *J. Atmos. Oceanic Technol.*, **14**, 120-132.
Doviak R. J. and D. S. Zrnić, 1993: *Doppler Radar and Weather Observations*. 2d ed. Academic Press, 562 pp.
Durran, D. R., 2000: Comments on "The differentation between grid spacing and resolution and their application to numerical modeling." *Bull. Amer. Meteor. Soc.*, **81**, 2478.
ESRL, cited 2013: Earth System Research Laboratory—Rapid Refresh (RR). [Available online at http://rapidrefresh.noaa.gov/.]
Federal Aviation Administration, cited 2010: Surface weather observation stations – ASOS/AWOS. [Available online at http://www.faa.gov/air_traffic/weather/asos.]
Federal Meteorological Handbook No. 1, 2005: Surface weather observations and reports. U.S. Department of Commerce/National Oceanic and Atmospheric Administration, 104 pp.
Fulton, R. A., J. P. Breidenbach, D. Seo, D. A. Miller, T. O'Bannon, 1998: The WSR-88D rainfall alogithm. *Wea. Forecasting,* **13**, 377-395.
Gao, J. D., K. Brewster, and M. Xue, 2008: Variation of radio refractivity with respect to moisture and temperature and influence on radar ray path. *Adv. Atmos. Sci.*, **25**, 1098–1106.

Glickman, T., S., 2000: *AMS Glossary of Meteorology*. 2nd Ed. American Meteorological Society, 855 pp.

Grasso, L., D., 2000: The differentiation between grid spacing and resolution and their application to numerical modeling. *Bull. Amer. Meteor. Soc.*, **81**, 579-580.

Grotjahn, R., 1993: *Global Atmospheric Circulations*. Oxford University Press, 430 pp.

Hallowel, R.G. and G.L. Blaisdell, 2003: Automated forecasting of road conditions and recommended road treatments for winter storms. *19th Conf. on Interactive Information and Processing Systems,* Long Beach, CA, Amer. Meteor. Soc., 9 pp. [Available online at http://ams.confex.com/ams/annual2003/techprogram/paper_55538.htm]

Haskins, B., 2006: The design of a quality control algorithm for environmental sensor stations. *Extended Abstracts. 22nd International Conf. on Interactive Information Processing Systems for Meteorology, Oceanography, and Hydrology*, Atlanta, GA, Amer. Meteor. Soc., 4 pp. [Available online at http://ams.confex.com/ams/Annual2006/techprogram/paper_102365.htm.]

Hawkinson, W. Feltz, S. A. Ackerman, 2005: A comparison of GOES sounder- and cloud lidar- and radar-retrieved cloud-top heights. *J. Appl. Meteor.*, **44**, 1234-1242.

Hershey, B. W., 2008: The physical nature and prediction of blowing snow within the roadway environment. *Extended Abstracts, Surface Transportation Weather and Snow Removal and Ice Control Technology,* Indianapolis, IN, Transportation Research Board, 381-392.

Hogan, R. J., E. J. O'Connor, A. J. Illingworth, 2009: Verification of cloud-fraction forecasts. *Quart. J. Roy. Meteor. Soc.*, **135**, 1494-1511.

Hogan, R. J., C. A. T. Ferro, I. T. Jolliffe, D. B. Stephenson, 2010: Equitability revisited: why the "equitable threat score" is not equitable. *Wea. Forecasting,* **25**, 710-726.

Holroyd, E.W., 1999: Snow accumulation algorithm for the WSR-88D radar: supplemental report. Bureau of Reclamation Report R-99-11, Denver, Colorado, June, 30 pp.

Hunter, S., 1996: WSR-88D radar rainfall estimation: Capabilities, limitations and potential improvements. *NWA Digest,* **20** (4), 26-36.

Hunter, S. M., E. W. Holroyd III, C. L. Hartzell, 2001: Improvements to the WSR-88d snow accumulation algorithm. *Extended Abstracts. 30th International Conf. on Radar Meteorology*, Munich, Germany, Amer. Meteor. Soc., 3 pp. [Available online at http://ams.confex.com/ams/30radar/techprogram/paper_20852.htm.]

Ikeda, K., R. M. Rasmussen, 2003: Radar observations of a freezing drizzle case in Colorado. *Extended Abstracts. 30th International Conf. on Radar Meteorology.* Seattle, WA, Amer. Meteor. Soc., 4 pp. [Available online at http://ams.confex.com/ams/32BC31R5C/techprogram/paper_64788.htm.]

ISCCP, cited 2010: ISCCP cloud types. [Available online at http://isccp.giss.nasa.gov/cloudtypes.html.]

Jolliffe, I. T. and D. B. Stephenson, 2004: *Forecast Verification: A Practitioner's Guide in Atmospheric Sciences*. Wiley, 256 pp.

Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, 341 pp.

Kessinger, C., S. Ellis, J. V. Andel, J. Yee, and J. Hubbert, 2005: The AP ground clutter mitigation scheme for the WSR-88D. *Extended Abstracts, 21th Conf. Interactive Information Processing Systems for Meteorology, Oceanography, and Hydrology,* San Diego, CA, Amer. Meteor. Soc., 13 pp.

Klazura, G. E., J. M. Thomale, D. S. Kelly, P. Jendrowski, 1999: A comparison of NEXRAD WSR-88D radar estimates of rain accumulation with gauge measurements for high- and low-reflectivity horizontal gradient precipitation events. *J. Atmos. Oceanic Technol.*, **16**, 1842–1850.

Koch, S. E., M. desJardins and P. J. Kocin, 1981: The GEMPAK Barnes objective analysis scheme. NASA Tech. Memo. 83851, NASA/GLAS, Greenbelt, MD 20771, 56 pp.

Koch, S. E., M. desJardins and P. J. Kocin, 1983: An interactive Barnes objective map analysis scheme for use with satellite and conventional data. *J. Climate Appl. Meteor.*, **22**, 1487-1503.

Kondragunta, C., D. Kitzmiller, D.-J. Seo, K. Shrestha, 2005: Objective integration of satellite, rain gauge, and radar precipitation estimates in the multisensor precipitation estimator algorithm. *Extended Abstracts*. *19th Conf. on Hydrology*, San Diego, CA, Amer. Meteor. Soc., 6 pp. [Available online at http://ams.confex.com/ams/Annual2005/techprogram/paper_86219.htm.]

Landolt, S., Politovich M., R. Rasmussen, A. Gaydos, M. Dewey, 2010: A comparison of an automated freezing drizzle algorithm to human observations. *Extended Abstracts*. *14th Conf. on Aviation, Range, and Aerospace Meteorology*. Atlanta, GA, Amer. Meteor. Soc., 4 pp. [Available online at http://ams.confex.com/ams/90annual/techprogram/paper_164082.htm.]

Lee, G, and I. Zawadzki, 2005: Variability of drop size distributions: time-scale dependence of the variability and its effects on rain estimation. *J. Appl. Meteor.*, **44**, 241-255.

Li, L., and J. W. Pomeroy, 1997: Estimates of threshold wind speeds for snow transport using meteorological data. *J. Appl. Meteor.*, **36**, 205-213.

Lin, L., X. Zou, R. Anthes, Y-H. Kuo, 2010: COSMIC GPS radio occultation temperature profiles in clouds. *Mon. Wea. Rev.*, **138**, 1104–1118.

Lin, Y., and K. E. Mitchell, 2005: The NCEP stage II/IV hourly precipitation analyses: development and applications. *Extended Abstracts*. *19th Conf. on Hydrology,* San Diego, CA, Amer. Meteor. Soc., 4 pp. [Available online at http://ams.confex.com/ams/Annual2005/techprogram/paper_83847.htm.]

Maddox, R. A., J. Zhang, J. J. Gourley, K. W. Howard, 2002: Weather radar coverage over the contiguous United States. *Wea. Forecasting*, **17**, 927-934.

MADIS, cited 2010: Meteorological assimilation data ingest system. [Available online at http://madis.noaa.gov.]

Mahoney, W. P. III, 2003: Decision support systems for winter road maintenance: opportunities for radar data. Preprints, 31st Conference on Radar Meteorology. *Conference on Radar Meteorology*, Seattle, WA, Amer. Meteor. Soc., 13.7. [Available online at http://ams.confex.com/ams/pdfpapers/64468.pdf.]

Mahoney, W.P. III, and W. L. Myers, 2003: Predicting weather and road conditions: an integrated decision support tool for winter road maintenance operations. *Transportation Research Record: Journal of the Transportation Research Board*, 98-105. [Available online at http://pubsindex.trb.org/view.aspx?id=663402.]

Manfredi, J., T. Walters, G. Wilke, L. Osborne, R. Hart, T. Incrocci, T. Schmitt, 2005: Road weather information system environmental sensor station siting guidelines. U.S. DOT FHWA-HOP-05-026, 46 pp.

Menzel, W. Paul, Frances C. Holt, Timothy J. Schmit, Robert M. Aune, Anthony J. Schreiner, Gary S. Wade, Donald G. Gray, 1998: Application of GOES-8/9 soundings to weather forecasting and nowcasting. *Bull. Amer. Meteor. Soc.*, **79**, 2059–2077.

Myrick, David T., John D. Horel, 2008: Sensitivity of Surface Analyses over the Western United States to RAWS Observations. *Wea. Forecasting*, **23**, 145–158.

NCEP, cited 2010: National Stage IV QPE Mosaic at NCEP. [Available online at http://www.emc.ncep.noaa.gov/mmb/ylin/pcpanl/stage4/.]

NWS, 1998: Automated surface observing system user's guide. ASOS Program Office, National Weather Service, 72 pp.

NWS, cited 2011a: Preliminary Local Climatological Data. [Available online at http://www.weather.gov/climate/local_data.php?wfo=mpx.]

NWS, cited 2011b: Preliminary Local Climatological Data. [Available online at http://www.weather.gov/climate/local_data.php?wfo=fgf.]

NWS, cited 2011c: Preliminary Local Climatological Data. [Available online at http://www.weather.gov/climate/local_data.php?wfo=abr.]

OFCM, 2002: Weather Information for Surface Transportation: National Needs Assessment Report. FCM-R18-2002, 302 pp.

OFCM, 2005: Weather Information for Surface Transportation Initiative Document: First Steps to Improving the Nation's WIST Capabilities and Services. 43 pp.

OFCM, 2006: Weather Information for Surface Transportation: Update on Weather Impacts and WIST Results. FCM-R26-2006, 32 pp.

Pauley, Patricia M., Xiaihua Wu, 1990: The theoretical, discrete, and actual response of the Barnes objective analysis scheme for one- and two-dimensional fields. *Mon. Wea. Rev.*, **118**, 1145–1164.

Peterson, D. P., and D. Middleton, 1963: On representative observations. *Tellus*, **15**, 387-405.

Pisano, P., and L. C. Goodwin, 2002: Surface transportation weather applications. *Extended Abstracts. 2002 Institute of Transportation Engineers Annual Meeting*, Philadelphia, PA, Institute of Transportation Engineers, 11 pp. [Available online at http://ops.fhwa.dot.gov/weather/best_practices/ ITE2002_SurfTransWxAppl.pdf.]

Pielke Sr., R., A., 2001: Further comments on "The differentiation between grid spacing and resolution and their application to numerical modeling." *Bull. Amer. Meteor. Soc.*, **82**, 699-700.

Pruppacher, H. R., and J. D., Klett, 1997: *Microphysics of Clouds and Precipitation*. Kluwer Academic, 954 pp.

Ramsey, A. C., 2002: Freezing drizzle (FZDZ) identification from the Automated Surface Observing System (ASOS): Status of the ASOS multi-sensor FZDZ algorithm. *Extended Abstracts. 6th Symp. on Integrated Observing Systems*, Orlando, FL, Amer. Meteor. Soc., 7 pp. [Available online at http://ams.confex.com/ams/annual2002/techprogram/paper_27226.htm.]

Rasmussen, R. M., J. Vivekanandan, J. Cole, B. Myers, C. Masters, 1999: The estimation of snowfall using visibility. *J. Appl. Meteor.* **38**, 1542-1563.

Rinehart, R. E., 2004: *Radar for Meteorologists*. 4th ed. Rinehart, 428 pp.

Rogers, R. R. and M. K. Yau, 1989: *A Short Course in Cloud Physics*. 3rd ed. Butterworth-Heinemann, 293 pp.

Root, B., T-Y Yu, M Yeary, 2009: The added value of surface data to radar-derived rainfall rate estimation using an artificial neural network. *Extended Abstracts*. *23rd Conf. on Hydrology*, Phoenix, AZ, Amer. Meteor. Soc., 7 pp. [Available online at http://ams.confex.com/ams/89annual/techprogram/paper_147108.htm.]

Schaefer, J. T., 1990: The critical success index as an indicator of forecasting skill. *Wea. Forecasting*, **5**, 570-575.

Schmit, Timothy J., Wayne F. Feltz, W. Paul Menzel, James Jung, Andrew P. Noel, James N. Heil, James P. Nelson, Gary S. Wade, 2002: Validation and use of GOES sounder moisture information. *Wea. Forecasting*, **17**, 139–154.

Seo, D. -J., 1998: Real-time estimation of rainfall fields using radar rainfall and rain gage data. *J. Hydrol*., **208,** 37-52.

Seo, D. –J. and J. P. Breidenbach, 2002: Real-time correction of spatially nonuniform bias in radar rainfall data using rain gauge measurements. *J. Hydrol.*, **3**, 93-111.

Smith, D. R., 1984: Equivalent radar reflectivity factors for snow and ice particles. *J. Climate Appl. Meteor*., **23**, 1258-1260.

Smith, D. R., F. W. Leslie, 1984: Error determination of a successive correction type objective analysis scheme. *J. Atmos. Oceanic Technol.*, **1**, 120-130.

Smith, D. R., M. E. Pumphry, J. T. Snow, 1986: A comparison of errors in objective analyzed fields for uniform and nonuniform station distributions. *J. Atmos. Oceanic Technol.*, **3**, 84-97.

Smith, J. A., D. J. Seo, M. L. Baeck, M. D. Hudlow, 1996: An intercomparison study of NEXRAD precipitation estimates. *Water Resour. Res.*, **32**, 2035-2045.

Spencer, P. L., M. A. Askelson, C. A. Doswell III, 2007: Choosing the smoothing parameters within a multiple-pass Barnes objective analysis scheme: a cautionary note. *J. Atmos. Oceanic Technol.*, **24**, 713-726.

Stephenson, D. B., 2000: Use of the "odds ratio" for diagnosing forecast skill. *Wea. Forecasting*, **15**, 221-232.

Super, A. B. and E. W. Holroyd III, 1998: Snow Accumulation Algorithm for the WSR-88D Radar: Final Report. U.S. Department of the Interior, R-98-06, 75 pp.

U.S. DOT FHWA, 2005: *Clarus* Weather System Design – Detailed System Requirements Specification. 99 pp.

Wade, C. G., 2003: A multisensory approach to detecting drizzle on ASOS. *J. Atmos. Oceanic Technol.*, **20**, 820-832.

Walters, M. K., 2000: Comments on "The differentiation between grid spacing and resolution and their application to numerical modeling." *Bull. Amer. Meteor. Soc.,* **81**, 2475-2477.

Weiss, N. A., 2008: *Elementary Statistics.* 7th ed. Pearson Addison Wesley, 744 pp.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences*. 2nd ed. Academic Press, 627 pp.

Zapotocny, T. H., and Coauthors, 2000: A Case Study of the Sensitivity of the Eta Data Assimilation System. *Wea. Forecasting*, **15**, 603–621.

APPENDIX

This appendix is designed to give the user of PPAES a guidance regarding the operations and utility of the software program and its functionalities. This software has been developed for a Linux environment. Therefore, there are some apparent assumptions when working with the system (i.e., basic unix/linux commands to move, copy, and execute files, etc.).

Overarching Design

This system was designed as a real-time system. This means that scripts are written to identify when new data files exist and to execute the program to update the analysis.

makefile.inc

The makefile.inc file contains macro definitions and a list of path specific information that is used by the program (specifically the Makefiles when compiling the program). These paths are set based on specific computer location of the code. The items that are set in the makefile.inc include the following:

- PPAES home directory path
- Utility directory path
- PPAES radar module directory path
- PPAES surface module directory path
- PPAES precipitation type module directory path
- PPAES satillite module directory path
- PROJ directory path
- GD directory path
- SHAPELIB directory path
- NETCDF directory path

Environmental Variables and Path

Depending on the operational environment of the system, the environment variables PPAES_DIR and MADIS_DATA need to be set. The following shows an example of two lines that need to be included in the .cshrc file in the users home directory if that such user is using tcsh for a shell.

```
setenv PPAES_DIR ${HOME}/PPAES

setenv MADIS_DATA /data/MADIS
```

In this example, the main PPAES source code is found in the user's home directory under PPAES and the raw MADIS data files are found in a directory located at /data/MADIS/.

Main Analysis Programs

There are three modules that are considered the core of the PPAES software. These modules are the radar, surface, and blend modules. Within these three modules there are inclusions of different components described above in the report including the use of model data, satellite data, and different surface observations.

Radar Module

The radar module is located in

${PPAES_HOME}/src/lvl2/

and provides analysis of WSR-88D level 2 data for the PPAES domain. It is run using the command line:

./ppaes_lvl2 byte_swap_flag radar_identifier full_path_to_input_file full_path_to_output_directory

where byte_swap flag is either 1=swap or 0=don't swap, radar_identifier = four letter identifier of radar (e.g., KMVX), and full_path_to_input_file and full_path_to_output_directory are self-explanatory.

This module produces eight analysis files:

1) ppaes_lvl2_YYYYMMDD_HRMNSC_dcr.bin
2) ppaes_lvl2_YYYYMMDD_HRMNSC_dtr.bin
3) ppaes_lvl2_YYYYMMDD_HRMNSC_mgr.bin
4) ppaes_lvl2_YYYYMMDD_HRMNSC_mht.bin
5) ppaes_lvl2_YYYYMMDD_HRMNSC_processed_radars.txt
6) ppaes_lvl2_YYYYMMDD_HRMNSC_prt.bin
7) ppaes_lvl2_YYYYMMDD_HRMNSC_rnm.bin
8) ppaes_lvl2_YYYYMMDD_HRMNSC_rrf.bin,

where "YYYYMMDD_HRMNSC" is the date and time stamp of the analysis. The following are descriptions of each file:

*_dcr.bin – Distance to the closest range resolution volume from which the analysis value came.
*_dtr.bin – Time difference (seconds) relative to the reference time for the analysis.
*_mgr.bin – Maximum ground range at which a contiguous precipitation echo is present.
*_mht.bin – MSL height of data from which precipitation rate came.
*_processed_radars.txt – If ppaes_lvl2 was executed using one of the provided Perl scripts, then this file is present and provides a list of radars used to produce the analysis at that time.
*_prt.bin – Liquid water equivalent instantaneous precipitation rate.
*_rnm.bin – The number (identifying index) of the radar from which an analysis value came.

*_rrf.bin – Reflectivity values that correspond to instantaneous precipitation rates.

Terrain Evaluation

To combat the problem that complex terrain produces, an option was included in the processing to invoke terrain clearance. This algorithm utilizes a DEM, the location of the radar, radar-relative data locations, and coordinate transformation equations to determine AGL heights of radar data. The first step in this process was the transformation of DEM data from a set of points that corresponds to a commonly-used map projection to radar-relative locations. A program that produces MSL elevations as a function of radar azimuth and great circle distance from a radar was developed. Three versions of this program are present in the ${PPAES_HOME}/src/util_exe directory, with the version in ${PPAES_HOME}/src/util_exe/Radar_DEM_files_v3 being the fastest and, therefore, recommended version. Use of this program is straightforward and can be understood by executing the program with the "-h" flag, as it typical for PPAES executables.

Terrain clearance is controlled by modifying the appropriate flag in the file ${PPAES_HOME}/src/lvl2/parm/ppaes_lvl2_ctrl.txt. For terrain clearance to work, a data file that contains azimuths, great-circle distances, and elevations for the radar in question must first be created using ${PPAES_HOME}/src/util_exe/Radar_DEM_files_v3/radar_DEM_files_v3. Output from ${PPAES_HOME}/src/util_exe/Radar_DEM_files_v3/radar_DEM_files_v3 must be placed within ${PPAES_HOME}/geog/Radar_DEM_data.

Model and Radar + Model Blend

This code is located in

> ${PPAES_HOME}/src/radmdl/

and produces both model and radar+model analysis products. It is executed using

> blend_radmdl input_type(1=RAP; 2=LAPS) full_path_to_intermediate_text_input_file full_path_to_ppaes_lvl2_prt_file full_path_to_output_dir

where input type is the type of model input (currently only the RAP option is operational), full_path_to_intermediate_text_input_file is a file produced by the script that is normally invoked in order to execute blend_radmdl, full_path_to_ppaes_lvl2_prt_file is the full path to the ppaes lvl2 prt output file that is to be blended with the model data contained in full_path_to_intermediate_text_input_file, and full_path_to_output_dir is the full path to the output directory that will hold analysis files.

blend_radmdl produces the following output files:

*mdl*mht.bin: The MSL height from which the model-based data used to estimate liquid-water-equivalent precipitation rate came.
*mdl*prt.bin: The model-based liquid-water-equivalent precipitation rate.
*radmdl*mht.bin: The MSL height from with the radar+model based data used to estimate liquid-water equivalent precipitation rate came.
*radmdl*pds.bin: The source of liquid-water-equivalent precipitation rate data (1=radar; 2=model).
*radmdl*prt.bin: The radar+model-based liquid-water-equivalent precipitation rate.

The script that is used to execute blend_radmdl is located in ${PPAES_HOME}/scripts/radmdl. This script is invoked in the following manner:

> blend_radmdl.pl input_type(1=RAP; 2=LAPS) full_path_to_mdl_input_file_or_dir full_path_to_ppaes_radar_analysis_dir full_path_to_output_dir

where intput_type is the type of model input (currently only RAP is supported, full_path_to_mdl_input_file_or_dir is the full path to the RAP input file or to the LAPS input directory, full_path_to_ppaes_radar_analysis_dir is the full path to the directory that contains the PPAES analysis files from which the analysis file that will be blended with the model data will be chosen (if none matches the time of the model data within the prescribed time window the script will end with an error), and full_path_to_output_dir is the full path to the output directory. This script produces an intermediate text file that is fed to blend_radmdl, executes blend_radmdl, and then removes the intermediate text files such that the remaining output is the analyses files produces by blend_radmdl.

Surface Module

The surface module is located in

> ${PPAES_HOME}/src/sfc/

and is used as the main analysis for MADIS and RWIS data. Two important data files that are output from this module end with *_spo.txt and *_spr.txt. These files are the precipitation occurrence and rate files, respectively, and are used in the blending routine.

The surface module can be run using the following command line structure.

> ./ppaes_sfc full_path_to_input_file full_path_to_output_dir data_stream MADIS_stats_flag

where “full_path_to_input_file” is the path and file name of the data file desired for processing, “full_path_to_output_dir” is the path to the output directory without file name(s), “data” is

literally the text “data” (which defines the lat/lon of the analysis grid), “data_stream” indicates data type (MADIS = 0, RWIS = 1), and “MADIS_stats_flag” indicates if separate stations on MADIS data should be calculated and printed to the screen (NO = 0, YES = 1).

In order for this module to run properly, the user must set specific parameters in the ppaes_sfc_ctrl.txt file located in PPAES/src/sfc. This file contains specific control parameters for operation of the code and includes declaration of the verification station list, stations density file, data time windows, and other parameters.

If no verification is desired, the verification file name in ppaes_sfc_ctrl.txt must be changed to ‘no_verification.txt’ and the verification flag in that file must be set to ‘0’. This file also allows the user to set their own verification directory (default is PPAES/src/sfc/verification). Likewise, the user can also set their own spatial density location and file (default directory is located in PPAES/src/sfc/density_metric)

The following is a list of the output files and their descriptions:

*_bnu.txt - Numerator values from Barnes scheme computation for each grid point
*_b_n.txt - Number of observations used in the Barnes computation for each grid point
*_cla.txt - Latitude of the closest observation station at each grid point
*_clo.txt - Longitude of the closest observation station at each grid point
*_o2g.txt - Distance from a gridpoint to the nearest observation
*_sdt.txt – Time differential relative to the reference time
*_spo.txt - Surface precipitation occurrence
*_spr.txt - Surface precipitation rate
*_stations.txt - File containing all stations reporting precipitation
*_stn.txt - Output station location on analysis grid
*_stns_used.txt - File containing all stations used in the analysis
*_stats.txt – File containing the statistics from the verification process (optional)
*_typ.txt – File containing indicators at which types of stations were used in this analysis (0 = MADIS, 1=RWIS).
*_v_stns.txt – Verification stations withheld from analysis (optional)

Radar+Surface Blending Module

The radar+surface blending module is located in:

PPAES/src/radsfc

and is code that preforms the blending of the surface and radar analysis. It is run using a similar structure to the following command line entry.

./ppaes_blend full_path_to_surface_spo_file full_path_to_radar_analysis_dir full_path_to_output_dir virga_model_flag(1=yes;0=no) verification_flag(1=run verification;other=no verification)

where “full_path_to_surface_spo_file” is the full path and file name of the surface module *_spo.txt file that was produced from the surface analysis, “full_path_to_radar_analysis_dir” is the full path to the directory where the radar analysis files are stored, “full_path_to_output_dir” is the path to the location the user wants to blended analysis output files to be located, “virga_model_flag” is an indicator to turn on or off the virga processing (OFF = 0, ON = 1), and “verification_flag” turns on (1) or off (0) verification processing.

The blend module produces output files that end with:

*_prt.txt – Liquid water equivalent instantaneous precipitation rate.
*_stats_join.txt – This file is created only with the verification flag set to 1. It contains the verification statistics of the blended analysis.
*_stats_lvl2.txt – This file is created only with the verification flag set to 1. It contains the verification statistics of the input radar analysis.

The _prt.txt file can be plotted using code located in:

PPAES/src/gd_plotting/ipr/blend/.

Utility Programs

The utility programs are needed to calculate needed information for the use of PPAES. These programs are not needed to run the software as default files currently exist for its operations. These programs are used to update files such as the station density, present weather climatology for MADIS stations, and other files that are needed by PPAES.

Station Density

The station density code it located in the PPAES/src/util_exe/Stn_Density directory. It is used to calculate the spatial density at each grid point of the PPAES domain by examining the user specified number of closest stations to each grid point. The program can be executed by five command line entries which include:

1) Executable name
2) Full path to input file
3) Full path to present weather climatology file
4) Full path to output directory (generally this is located in PPAES/sfc/density_metric/)
5) Number of closest station to use
    a. This number is the initial number of stations gathered for each grid point and the code is defaulted to use only 10 of those initial stations. This helps to ensure that the closest stations for each of those 10 closest stations are also properly determined as the stations that are closer to a grid point may not be the 10 closest stations to any one of these stations.

The code will automatically calculate a file name using the following convention and place it in the user defined directory:

ppaes_stn_density_df_20121227_230000_cf_20120429_070000_s10_30.txt

"ppaes_stn_density_" – indicates this is a station density file.
"df_20121227_230000_" – indicates the data file date/time stamp (year month day_hour min sec).
"cf_20120429_070000_" – indicates the date/time stamp of the present weather climatology file where the date/time stamp is determined by the last MADIS station that was added/updated to the file (year month day_hour min sec).
"s10_30" – indicates that 10 stations were used to calculate the spatial density using the initial 30 stations gathered for each grid point.

This output can be plotted using code that resides in PPAES/src/gd_plotting/station_density/.

Present Weather Climatology

The present weather climatology code is used to gather a list of MADIS stations that can (and do) report weather conditions. As defined in the report, some MADIS stations other than AO2 or AO2a may report weather conditions. Therefore, they should be included in the surface analysis. This code is used in conjunction with a perl script that constantly examines a data directory for new MADIS files. When a new file is present, it processes it to determine if stations should be added to the climatology list. This perl script can be set up in a cron job to run continuously so that the climatology file is always up-to-date. It is best to let this code run for some time before using the climatology file that is produced so that the code has time to build an informative list of precipitation reporting stations.

The perl script to run the code is found in:

PPAES/scripts/sfc/metar_climo/preswx_climo/,

and the code can be found in:

PPAES/src/util_exe/metar/preswx_climo.

Radar/Model/Satellite Verification

The radar, model, and satellite verification code it located in the PPAES/src/util_exe/Rdr_Mdl_Sat_Verif directory. It is used to calculate the same statistics that are calculated for the surface analysis (which are output into the files ending in "_stats.txt"). This verification method uses a predetermined list of MADIS stations that report precipitation (verification stations list that is set in the control file and identical to the verification method used for the sfc analysis) and are used to compare with analyses from the radar, model, satellite, or any combination of these. The precipitation rates and occurrences are examined and used in the

statistic calculations. Each stations observational information is output into the file ending in "_v_stns.txt" and the average statistics for all stations are output in the file ending in "_stats.txt" (similar to the naming convention from the surface analysis). The program can be executed using four command line entries which include the following:

1) Executable name (./ppaes_verif)
2) Full path to analysis input file
3) Full path to raw MADIS data file that covers analysis reference time
4) Full path to output directory

These stats can then be aggregated to determine how well the radar/model/satellite analyses are performing based on truth (MADIS) surface data.

Util

The util directory, ${PPAES_HOME}/src/util, contains source code (*.c files) and header files (*.h) files that are used within multiple algorithms (surface, radar, model, etc.). Because of their commonality, they are all organized within this one folder, which enhances code maintainability.

Plotting Programs

All plotting software follows similar command line structure for execution and includes two methods for this execution. The following is an example and description of the general command line requirements.

> ./plot_dt_sfc full_path_to_plotting_control_file full_path_to_ppaes_sfc_data_file data full_path_to_output_file

-or-

> ./plot_dt_sfc full_path_to_plotting_control_file full_path_to_ppaes_sfc_data_file lon_left lon_right lat_bott lat_top full_path_to_output_file

1) The first entry is the executable for the program. This usually follows the convention of "plot_XXXX" where "XXXX" relates to the type of plotting being performed.
2) The second entry is the full path to the plotting control file. In each directory, there is a file that ends with "_ctrl.txt" and contains configurable settings of the plotting software.
3) The third entry is the full path to the analysis file being plotted.
4) The fourth entry is where the user can select to plot a specific domain, or the entire PPAES domain
   a. If the user wants the entire domain, then the text "data" should be placed in this position.
   b. If the user wants a specific domain, then the longitude and latitude of the domain must be specified in the command line in this order: left longitude, right longitude, bottom latitude, top latitude.

5) The last entry is the full path to the output file. This includes a file extension that is associated with an image (i.e., PNG, JPG, etc.).

dt_sfc

This directory contains code to plot the time difference in the surface analysis. The analysis files that should be used with this code end with "_sdt.txt".

ipr

Several plotting programs exist under the ipr directory. These programs are meant to plot the instantaneous precipitation rate for surface blended (blend), satellite (gctp), radar (lvl2), model (mdl), radar and model blended (radmdl), and surface (sfc) analyses.

prec_type

This code is used to plot data produced by the prec_type analysis (located in ${PPAES_HOME}/src/prec_type/).

rad_ground_range

This code is used to plot data produced by the radar module (located in ${PPAES_HOME}/src/lvl2/) and used with files ending in "_mgr.bin" or "_mgr.txt".

rad_num

This code is used to plot data produced by the radar module (located in ${PPAES_HOME}/src/lvl2/) and used with files ending in "_rnm.bin" or "_rnm.txt".

sfc

This code is used to plot data produced by the surface module (located in ${PPAES_HOME}/src/sfc/) and used with files ending in "_spo.txt".

Shapefiles

This directory contains GIS data files to be used during plotting.

station_density

This plotting software is used to plot the data produced by the station density code (located in ${PPAES_HOME}/src/util_exe/Stn_Density/) and used with files ending in "_s*_*.txt", where "*" is an integer (usually "_s10_30.txt").