IMPACTS OF MESONET OBSERVATIONS ON METEOROLOGICAL SURFACE ANALYSES

Daniel P. Tyndall[1,2] and John D. Horel
Department of Atmospheric Sciences, University of Utah
Salt Lake City, UT

Submitted to: *Weather and Forecasting*
March 20, 2012

[1] Current affiliation: National Research Council Postdoctoral Fellow
[2] Corresponding author address:
Daniel P. Tyndall
U.S. Naval Research Laboratory
7 Grace Hopper Avenue
Monterey, CA 93943
Email: dan.tyndall.ctr@nrlmry.navy.mil

Abstract

Given the heterogeneous equipment, maintenance and reporting practices, and siting of surface observing stations, subjective decisions that depend on the application tend to be made to use some observations and avoid others. This research determines objectively high impact surface observations of 2-m temperature, 2-m dewpoint, and 10-m wind observations using the adjoint of a two-dimensional variational surface analysis over the contiguous United States. The analyses reflect a weighted blend of 1-h numerical forecasts used as background grids and available observations. High impact observations are defined as arising from: low observation quality, observation representativeness errors, or accurate observed weather conditions not evident in the background field.

The impact of nearly 20,000 surface observations is computed over a sample of 100 analysis hours during 25 major weather events. Observation impacts are determined for each station as well as within broad network categories. For individual analysis hours, high impact observations are located in regions of significant weather—typically, where the background field fails to define the local weather conditions. Low impact observations tend to be ones where there are many observations reporting similar departures from the background. When averaged over the entire 100 cases, observations with the highest impact are found within all network categories and depend strongly on their location relative to other observing sites and the amount of variability in the weather, e.g., temperature observations in urban areas such as Los Angeles, CA, where observations are plentiful and temperature departures from the background grids are small, have reduced impact.

1. Introduction

Mesoscale surface observations are vital data sources for applications in many different meteorological subfields, including operational forecasting, wind power management, transportation safety, wildfire management, dispersion modeling, and defense applications (Dabberdt et al. 2005; Horel and Colman 2005). Two recent reports (National Academy of Sciences 2009; 2010) recommend that existing and future mesoscale observations be integrated into a network of networks. The heterogeneous nature of the available mesoscale surface observing networks within the United States (e.g., varying sensor quality, maintenance and reporting practices, and siting) can limit their potential benefits. A critical recommendation in both reports is to improve the metadata that defines the sensor and station characteristics within the aggregated networks. Users of the national network would then be able to select the types of stations that meet their specific needs.

Obviously, stations with higher quality equipment that are properly sited and maintained are likely to be of greater value for all applications. However, does a low-cost station with lower-quality standards in a largely data void region have greater value than one of several expensive high quality stations located close to each other? Observation value is clearly a function of the observation's total cost, the availability of comparable data nearby, and its benefit to diverse potential applications, including use by human forecasters and integration into atmospheric analyses and numerical forecasts.

This study does not attempt to address the broader issues of the relative value of observations obtained from different networks. Rather, the scope is limited to assessing the utility of an objective metric to identify the characteristics of observations and networks that

76 strongly influence mesoscale surface analyses over the contiguous United States (CONUS).
77 We address the extent to which the impact of observations depends "dynamically" on the
78 synoptic, mesoscale, and local weather situation relative to the "static" underlying siting and
79 standards of the various observing networks. This work is motivated by the pressing need to
80 develop automated quality control procedures for mesonet observations from heterogeneous
81 sources.

82 Domestic and international research efforts have led to improved mesoscale analysis
83 and data assimilation systems. Current operational products include: MatchObsAll, which is
84 used at many National Weather Service offices; Vienna Enhanced Resolution Analysis
85 (VERA; Steinacker et al. 2006); Integrated Nowcasting through Comprehensive Analysis
86 (INCA; Haiden et al. 2010); Mesoscale Surface Analysis System (MSAS; Glowacki et al.
87 2011); Space and Time Multiscale Analysis System (STMAS; Xie et al. 2011); Real-Time
88 Mesoscale Analysis (RTMA; de Pondeca et al. 2011), and Rapid Refresh and High
89 Resolution Rapid Refresh Systems of the National Centers for Environmental Prediction.
90 These systems rely on analysis techniques ranging from Cressman and spline interpolation
91 methods to advanced variational techniques. Many of these high resolution (1-12 km)
92 analyses are not intended to initialize subsequent model forecasts, but instead used as end
93 products to help diagnose current conditions or verify prior forecasts. Examination of
94 sequences of such surface analyses allows users to grasp the spatial and temporal variability
95 of weather situations more readily than direct inspection of conditions at hundreds of specific
96 observing sites. We will introduce the University of Utah Variational Surface Analysis

97 (UU2DVar) in this paper as an efficient research tool comparable in characteristics and
98 performance to these operational systems.

99 Prior work has demonstrated that the near-surface boundary layer in the CONUS
100 remains under sampled (Myrick and Horel 2008; Horel and Dong 2010). Figure 1 illustrates
101 the need for additional observing capabilities, particularly in mountainous and coastal areas,
102 in terms of an integrated data influence analysis (IDI; Uboldi et al. 2008). The IDI analysis is
103 generated by assuming that the background grid values are 0 everywhere and setting the
104 observation values at the station locations available in this study to 1. As discussed by Horel
105 and Dong (2010) and later in Section 2, the IDI analysis is a measure of station density and
106 depends on the assumed observation and background error covariances. Areas with many
107 stations and enhanced data coverage appear in dark shades in Fig. 1, while areas with few
108 observations available appear white or light grey.

109 Figure 1 should be viewed as a static illustration of data density. An underlying
110 assumption of the recommendations related to improving station metadata made in the
111 National Academy of Sciences reports is that by simply knowing the locations of existing
112 mesonet stations and their basic operating characteristics as part of an improved national
113 database of station metadata, then it will be possible to infer where additional stations are
114 needed by identifying the apparent data voids. However, IDI analyses or examination of
115 network metadata do not take into consideration the spatial and temporal variability arising
116 from weather and may not identify where observations are most needed.

117 Observation impacts have often been evaluated through cross-validation experiments,
118 in which a control analysis using all observations is compared to an analysis in which

observations of interest are excluded from the assimilation (Seaman and Hutchinson 1985; Zapotocny et al. 2000; Myrick and Horel 2008; Benjamin et al. 2010; Tyndall et al. 2010; Horel and Dong 2010). Generally, groups of observations are withheld (e.g., observations from differing types of sensor systems or networks). Horel and Dong (2010) applied this technique extensively by sequentially withholding each of ~3,000 observations evaluated in their study from ~9,000 analyses, resulting in over 570,000 cross-validation experiments. While this methodology provides individual impacts for each observation, it is far too cumbersome and expensive to use operationally.

A more efficient approach to determine observation impacts utilizes the analysis adjoint that relates the analysis sensitivity to input location. Adjoints of forecast models are used routinely to assess where "targeted" observations might reduce forecast errors (Palmer et al. 1998, Buizza and Montaini 1999; Langland et al. 1999; Zhu and Gelaro 2008). Baker and Daley (2000) utilized the adjoint of a simple data assimilation system to explore analysis sensitivity on an analytical function. In this study, a similar approach is applied to surface mesonet observations throughout the entire CONUS domain to illustrate an approach to estimate which types of networks tend to have high observation impact.

The approach used in this study is described in Section 2. Application of the methodology to a single case and aggregate statistics over a sample of 100 analyses are presented in Section 3. Additional discussion and ongoing work related to automated quality control procedures for mesonet observations are described in Section 4.

2. Method

a. Analysis

This study utilizes the University of Utah Variational Surface Analysis (UU2DVar) to determine impacts of mesonet observations. The UU2DVar is a univariate two-dimensional variational (2DVar) analysis tool that generates meteorological analyses of 2-m air temperature, 2-m dewpoint temperature, 10-m u and v wind components, and surface pressure (Tyndall 2011). The analysis tool is designed to use the observation space form of the variational cost function equation (Lorenc 1986; Daley and Barker 2001):

$$\boldsymbol{y_o} - \mathbf{H}(\boldsymbol{x_b}) = \left(\mathbf{H}\mathbf{P_b}\mathbf{H}^{\mathbf{T}} + \mathbf{P_o}\right)\boldsymbol{\eta} \tag{1}$$

$$\boldsymbol{x_a} = \boldsymbol{x_b} + \mathbf{P_b}\mathbf{H}^{\mathbf{T}}\boldsymbol{\eta} \tag{2}$$

and computes the analysis, $\boldsymbol{x_a}$, by spreading information from observations ($\boldsymbol{y_o}$) across the background grid ($\boldsymbol{x_b}$) by the background ($\mathbf{P_b}$) and observation ($\mathbf{P_o}$) error covariances, and the forward operator ($\mathbf{H}$). The term $\boldsymbol{\eta}$ is computed by iteratively solving (1), and is used to yield the analysis in (2). The observation space form of the cost function equation is very efficient for undersampled problems such as mesoscale surface observation assimilation (the number of observations is much less than the number of gridpoints of the analysis). The resolution of the UU2DVar analyses is dependent upon the resolution of the background fields, which in this study are Rapid Update Cycle (RUC) 1-hr forecasts (Benjamin et al. 2004) downscaled from 13 km to 5 km resolution. These RUC backgrounds also serve as the background fields for the RTMA (de Pondeca et al. 2011).

Parallelization of the UU2DVar originally using Matlab software for this study and now using open-source Python libraries has enabled the system to be run in real-time, and

hourly analyses are currently generated over the CONUS domain as part of quality control procedures for MesoWest (Horel et al. 2002). Usage of parallelization allows efficient computation of the background error covariance ($\mathbf{P_b}$), which is horizontally and vertically spatially dependent through the product of two inverse exponential functions that depend on horizontal (R) and vertical (Z) decorrelation length scales:

$$\mathbf{P_b}_{ij} = \sigma_b \exp\left(-\frac{r_{ij}^2}{R^2}\right) \exp\left(-\frac{z_{ij}^2}{Z^2}\right) \quad (3)$$

where σ_b is the background error variance and r_{ij} and z_{ij} are the horizontal and vertical distances between gridpoints i and j. The vertical term in the equation adds terrain anisotropy, which prevents an observation in a valley from influencing areas of the analysis over nearby mountain ridges. The vertical term also controls the influence of coastal observations, as water analysis gridpoints have their elevation reduced by 500 m in the covariance computation to prevent land observations from influencing them. The horizontal and vertical decorrelation length scales used in this study were 80 km and 200 m, respectively. These length scales were estimated by Tyndall et al. (2010) for the same RUC background fields. These assumed decorrelation length scales do influence some of the results of this study, as they define a priori the spatial scales over which an observation is most likely to influence the analysis. Since the functional form of the covariance in (3) only asymptotes to 0 at long distances, correlations between gridpoints separated by more than 300 km are set to 0 to improve computational efficiency.

b. Observations

Nearly all observations from the 125 mesonet networks used in this research are part of the MesoWest database of publicly accessible observations (Horel et al. 2002). A continually updated list of networks available in MesoWest is available online. Two proprietary networks (WeatherFlow's largely coastal network and the Oklahoma Mesonet, McPherson et al. 2007) were added with permission as they help to illustrate networks that would likely be included in a national network of networks. The networks used in this study are deployed by entities for specific purposes and can be broadly separated into the 10 categories listed in Table 1. The median IDI values computed from the locations of all stations within each category reflect differences in station density, e.g., primarily urban stations (IDI values > 0.95 for the PUBLIC and AQ categories) vs. rural or isolated stations (IDI values < 0.90 for the RAWS, EXT, and HYDRO categories).

Based on subjective experience working with data from these networks and prior research (Tyndall et al. 2010), we have attempted to classify loosely the relative magnitude of mesonet observation errors compared to the background errors. These assumptions are made separately for temperature, moisture, and wind observation errors for each of the 10 network categories. Given the subjectivity of these assumptions, our approach can be viewed as a sensitivity study in which the available stations are assigned varying error ratios to evaluate the dependence of our results on these assumptions. Higher ratios reduce the influence an observation will have on the resulting analysis in the vicinity of that observation. As a baseline, observations from the NWS/FAA reporting stations (NWS category) and other federal and state networks (FED+ category) that tend to adopt

203 standardized installation, siting, and maintenance procedures are assigned observation to
204 background error variance ratios of 1.0. This ratio implies that for an isolated observation,
205 the resulting analysis in the vicinity of the observation will approximate roughly the average
206 of the observation and background values.

207 Higher observation to background ratios of 1.5 and 2.0 are assigned for some
208 observation types and network categories to reflect the characteristics of many of these
209 networks or as an attempt to account for representativeness errors. For example, the higher
210 2.0 observation to background error variance ratio is used for wind observations from the
211 Remote Automated Weather Station (RAWS) category because: (1) the RAWS standard for
212 wind sensor height (6 m as opposed to the 10 m standard for NWS and FED+ categories)
213 leads to lower observed wind speeds relative to the 10 m background wind and (2) many of
214 the RAWS stations are sited for fire weather applications in rugged locations with highly
215 variable terrain and vegetation for which the observations may not represent the conditions
216 over nearby 5 km analysis grid boxes. Similarly, higher wind observation errors are also
217 assumed for the AG and PUBLIC categories since many agricultural networks rely on 3 m
218 towers and PUBLIC sensors are often mounted on or near residences with nearby
219 obstructions commonplace.

220 To prevent clearly erroneous observations from entering the analyses in this study, we
221 used a manual blacklist and observation innovation (observation minus background) checks.
222 The blacklist was prepared by subjectively rejecting observations from stations exhibiting
223 both large mean observation innovations combined with large mean impacts (as defined in

the next subsection) when the 100 analyses were computed using all available observations. The innovation control check is summarized by Equation 4:

$$|\boldsymbol{y_o} - \mathbf{H}(\boldsymbol{x_b})| \leq \max\left[\varepsilon_m \cdot \text{stdev}\left(\boldsymbol{x_b}^{r\leq 40\text{ km}}\right), \tau_{qc}\right] \quad (4)$$

where ε_m is a tunable coefficient, τ_{qc} is a tunable quality control threshold (in units of the observations), and $\boldsymbol{x_b}^{r\leq 40\text{ km}}$ are background field values (in units of the observations) no further than 40 km from the observations. The functions max and stdev are the maximum and standard deviations operating on the arguments, respectively. The tunable quality control threshold was added to prevent observations from being rejected in situations such as offshore areas where the variance in the background field is usually small. In this study, ε_m is 10 for all variables, and τ_{qc} is set to 3°C, 4°C, and 7.5 m/s for air temperature, dewpoint temperature, and wind values (both components and speed), respectively. Observation innovations that were likely due to representativeness errors (e.g., temperature observations at Mt. Rainier, WA where the analysis grid elevation is much lower than the sensor elevation) were not removed explicitly unless they were so egregious as to fail the innovation control check.

An additional quality control step was applied to wind observations to reject very light winds when the background values were much higher, i.e., wind observations were not used when the observed speed was less than 1 m/s and the background speeds exceeded 5 m/s. Such situations arise frequently due to a variety of siting and reporting reasons, and exacerbated at night during periods of weak synoptic flows. For example, the lowest reported wind speed for NWS observations is 1.25 m/s; hence, a “calm” report when the background suggests a relatively strong wind is ignored.

c. Observation Impact

As mentioned in the Introduction, the adjoint of a data assimilation tool can be used to find individual observation impacts in a much more efficient manner than traditional cross-validation experiments using data withholding. The measure of observation impact used in this study follows that used by Langland and Baker (2004) and Zhu and Gelaro (2008). This methodology measures observation impact with respect to a sensitivity cost function, $\boldsymbol{\mathcal{J}}$, which is often a forecast variable of interest over a particular subdomain. In this research, $\boldsymbol{\mathcal{J}}$ is specified to be the squared differences between the analysis and the background field [$\boldsymbol{\mathcal{J}} = 0.5\ (\boldsymbol{x_a} - \boldsymbol{x_b})^2$]. Following Baker and Daley (2000), observation sensitivity ($\partial\boldsymbol{\mathcal{J}}/\partial\boldsymbol{y_o}$) is found using the chain rule:

$$\frac{\partial \boldsymbol{\mathcal{J}}}{\partial \boldsymbol{y_o}} = \frac{\partial \boldsymbol{x_a}}{\partial \boldsymbol{y_o}} \frac{\partial \boldsymbol{\mathcal{J}}}{\partial \boldsymbol{x_a}} = \left[\left(\mathbf{H}\mathbf{P_b}\mathbf{H^T} + \mathbf{P_o}\right)\mathbf{H}\mathbf{P_b}\right][\boldsymbol{x_a} - \boldsymbol{x_b}] \tag{5}$$

The measure of observation impact, $\boldsymbol{\mathcal{G}}$, follows that previously defined by Zhu and Gelaro (2008) as the scalar product of observation sensitivity and the observation innovations:

$$\boldsymbol{\mathcal{G}} = \frac{1}{2}\left\langle \frac{\partial \boldsymbol{\mathcal{J}}}{\partial \boldsymbol{y_o}}, \boldsymbol{y_o} - \mathbf{H}(\boldsymbol{x_b}) \right\rangle \tag{6}$$

where $\langle \cdot \rangle$ denotes the scalar product operator.

Observation impacts were computed over 100 analysis hours for all available surface observations. The 100 analysis hour sample was composed of the 00, 06, 12, and 18 UTC analyses corresponding to multi-million dollar severe weather damage between October 2010 and April 2011. The dates of each event, type of severe weather, and areas affected are listed in Table 2.

263 The wide range of values of observation impact (as a function of network category,
264 variable, location, and date) led us to rank their values from smallest to largest for each
265 variable. Negative impacts occur where the deviation of an observation from the background
266 differs in sign relative to its neighbors, which may reflect either an erroneous observation or a
267 realistic weather phenomenon on a scale smaller than that assumed a priori for the background
268 errors. The percentile ranks of the observation impacts are aggregated for each station or for
269 all stations in a network category for each variable for either individual cases or over the
270 entire 100 analysis sample. Hence, a station's temperature observations are "high impact" if
271 its observation percentile rank impact values are large relative to the percentile impact values
272 for temperature at other stations.
273
274 **3. Results**
275 *a. 0000 UTC 04 April 2011*
276 The observation impact approach is illustrated using an analysis at the start of a two-day (04
277 and 05 April 2011) severe weather period. The severe weather on these two days produced
278 numerous tornado, hail, and high wind reports over much of the Midwest and southeastern
279 United States particularly later on 04 April, and caused 3 fatalities and over 20 injuries. The
280 severe weather was associated with prefrontal conditions ahead of the cold front stretching
281 from Iowa to Arizona evident in the 0000 UTC 04 April 2011 weather summary of the
282 Hydrologic Prediction Center (Fig. 2).
283 At 0000 UTC 04 April, clear contributions of mesonet observations to the
284 temperature, dewpoint, and wind speed analyses were evident near the front and dry line

285 locations as well as elsewhere within the CONUS (Fig. 3). As shown by the analysis
286 increments in the right panels of Fig. 3, the addition of the observations led to: higher
287 temperatures and lower dewpoint temperatures than the background south of the warm front
288 in the Midwest and lower wind speeds north of the warm front. Additionally, the
289 observations led to substantive adjustments to the background in mountainous areas of the
290 West, e.g., higher temperatures and dewpoints over the southern Sierra Nevada Mountains in
291 California and lower wind speeds over portions of Arizona and New Mexico.

292 The analysis increments in Fig. 3 show the collective influence of all observations on
293 the analysis. Figure 4 illustrates the influence of each observation for selected network
294 categories for 2-m temperature based on the analysis adjoint methodology. Red (blue) circles
295 at station locations in Fig. 4 denote percentile ranks of observation impact in the top 25%
296 (lowest 25%) for this temperature analysis. The 4 network categories shown in Fig. 4 reflect
297 the diversity of network density, observation siting, and sensor quality. The NWS category
298 consists of professional grade equipment sited primarily at airports around the country while
299 the FED+ category represents an aggregate of national and regional federal networks (e.g.,
300 Climate Reference Network and Modernized Cooperative Observer Program), local federal
301 networks (e.g., Field Research Division of the NOAA Air Resources Laboratory in eastern
302 Idaho), and state networks (e.g., Oklahoma, West Texas, New Jersey). The RAWS category
303 discussed previously is composed of stations that tend to be located in remote locations and
304 often exhibit representativeness errors while the PUBLIC category constitutes the largest
305 number of stations that are often densely distributed in urban areas and rely frequently on
306 lower-grade sensors sited on residences.

For this particular temperature analysis, many of the NWS temperature observations around the CONUS fall in the "typical" impact range found for most observations, i.e., percentile ranks between 25% and 75% (upper left panel of Fig. 4). The NWS stations at this time with the largest impacts (red colors) tend to be clustered near the warm front in the Upper Midwest where the background underestimated the strength of the warm front. The NWS temperature observations with the least impact are found in Virginia and Maryland and many coastal areas. The FED+ stations in Maine tend to have high impact while those in eastern Idaho exhibit individually modest impact at this time (upper right panel of Fig. 4).

Even though we a priori "trust" RAWS temperature observations less than NWS of FED+ observations, many RAWS temperature observations fall in the upper quartile of impact particularly in the Sierra Nevada and Cascade Mountain ranges of the Pacific states (lower left panel of Fig. 4). These observations lead to the upward (downward) adjustment of the background temperature in the Sierras (western Colorado) evident in Fig. 3. It is obvious that in most major urban areas, the impacts of observations in the PUBLIC category on the analyses are small, e.g., Los Angeles, CA, San Francisco, CA, Seattle, WA, Dallas, TX, coastal FL, and Washington D.C. (lower right panel of Fig. 4). This results from the large number of observations in those areas that reduces the observation sensitivity (5) as well as our assumption of higher observation to background error ratios for these stations. However, the limited impact of these observations is as much determined by the weather at this time in those locations, as NWS observations in those locations also tend to have small impact. Since the PUBLIC observations differ from the background substantively in Illinois, Indiana, and

328 Ohio, those observations influence the analysis as much as other stations that we "trust"
329 more.

330 Hence, observation impact as defined in this study depends on a complex blend of the
331 assumptions regarding observation and background errors, local observation density, and the
332 local, regional, and synoptic-scale weather taking place at any one time. To reduce the
333 dependency on specific weather features at specific times, we now examine observation
334 impact over a representative sample of 100 analyses.

335

336 *b. Observation Impacts Over 100 Analysis Hours*

337 Although the 100 analyses used in this section are associated with severe weather in specific
338 locales (Table 2), they reflect a reasonable mix of typical weather conditions when viewed on
339 the CONUS scale. Following the previous case study, the percentile ranks of the observation
340 impacts are computed for each variable and analysis. Then, the median percentile rank over
341 the sample of 100 cases is determined for the observations at a specific station. Those
342 stations that consistently influence the analyses more will have median values in the upper
343 quartile while most stations will have less effect or influence analyses on only a few
344 occasions within this sample.

345 Figure 5 summarizes the impact of temperature, dew point temperature, and wind
346 observations from the 4 network categories previously highlighted in Fig. 4. Not all NWS
347 stations exhibit high impact on temperature analyses (Fig. 5a). For example, temperature
348 observations from NWS stations near Los Angeles and San Francisco, CA exhibit some of
349 the lowest impacts. Hence, the cumulative number of observations combined with the

generally limited variability in temperature in those areas requires less adjustment of the background field and reduces the impact of any one station. On the other hand, NWS observations have large impact in North and South Dakota and many locations in Colorado and offshore areas in the Gulf of Mexico. In addition, many of the stations in the NOAA Special Operations and Research Division (SORD) network in southern Nevada included in the FED+ category have a large impact on the temperature analyses.

When evaluated over the entire 100 analysis sample, RAWS temperature observations frequently have a large impact in rugged areas (Fig. 5a) even though we have assumed those observations to have larger errors. The impact of temperature observations from the PUBLIC network category are clearly tied to the local density of observations with lower impacts in many metropolitan areas but substantive impacts in many other locales around the country.

NWS dew point and wind observations have larger impacts relative to temperature observations across the central swath of the country in Figs. 5b and 5c respectively. This suggests greater sensitivity where dew point temperatures and wind speeds vary greatly from day-to-day and large adjustments to the background fields are common. The impact of dew point temperature and wind speed observations in the FED+ category from the Oklahoma Mesonet and West Texas Mesonet stand out for similar reasons. The impact of RAWS dew point and wind observations tend to reflect representativeness issues with higher impacts in regions of complex terrain.

To examine the influence of observations from individual stations from all 10 network categories on the analyses, Figs. 6 and 7 show the impact of temperature

observations in northern Utah and wind observations in southern California respectively. As with the results shown in Fig. 5, these percentile impacts in Figs. 6 and 7 are relative to all CONUS stations. The dependence on local station density is quite apparent with no one station in any category having a large impact on the temperature analyses in the Salt Lake Valley (near SLC, labeled in the EXT panel). Not surprisingly, the stations over the Great Salt Lake from the air quality (AQ) and local network categories have a large impact as they help to take into consideration the unique weather conditions over that water body. Temperature observations from transportation networks exhibit a wide range of sensitivities in this region with low impact near SLC and high impact from some rail network stations in southeastern Idaho as well as from road weather stations across the southern portions of the domain.

The impact of surface wind observations in southern California suggest that the specification of higher observation error for some network categories compared to others is not a controlling factor. For example, most coastal stations have low impact, implying that the background fields are not substantively adjusted from either NWS or PUBLIC stations. However, further into the interior of the San Diego and Los Angeles basins, most stations have moderate impact, independent of whether they are NWS, RAWS, AQ, or PUBLIC stations. The analyses in this subdomain are most strongly affected by the relatively few stations in nearly all categories in the high desert regions north of the San Gabriel and San Bernardino mountains across the center of this subdomain.

The salient results of our study are summarized in Fig. 8. The fraction of the total number of stations in each of the 10 network categories that have observation impacts in the

upper quartile accumulated over all 100 cases are shown. (The percentage of the total station count in each category is used because the total number of stations in each network category varies widely; see Table 1.) If every station had a comparable impact, then all the bars would lay along the thick black line, i.e., 25% of the reports from any network would be in the upper quartile. Network categories with fractions greater (less) than 25% have more high (low) impact observations. For reference, the assumed observation to background error variance ratio is indicated in Fig. 8 for each network category and variable as well. If these assumptions controlled our results, then it would be expected that networks that are "trusted" less (i.e., error ratios of 1.5 or 2.0) would have fewer stations with large impacts than those that serve as the baseline (i.e., error ratios of 1.0). However, that is not the case.

Consider first the impact of the network categories shown in Figs. 4 and 5: NWS; FED+; RAWS; and PUBLIC. As expected, a larger fraction of observations from the NWS and FED+ network categories have substantive impacts on all three variables relative to many other network categories. Figure 6 also reinforces the spatial displays in Figs. 4 and 5 that a larger (smaller) fraction of RAWS (PUBLIC) stations have large impacts. These results are consistent with the interplay between the ability of the background to resolve the observed weather, the local station density, and representativeness errors as opposed to simply the specification of the observation error variance.

Now consider the cumulative statistics for the other 6 network categories, beginning with the networks intended primarily for agricultural applications (AG). The AG impacts hover around what might be expected for an "average" station, close to 25%. As mentioned before, AG category wind observations are assigned larger observation errors primarily

because common practice for agricultural applications is to use 3 m towers. Air quality networks (AQ) tend to have lower than expected impact, presumably due to their locations in urban areas where many other stations are available (i.e., similar behavior to PUBLIC stations). They have been assigned intermediate observation errors primarily due to the Environmental Protection Agency standard for reporting observations in terms of time averages as long as 60 min in contrast to shorter averaging periods used for stations in other network categories.

Canadian, Mexican, and offshore observations in the EXT category exhibit the highest fraction of large impact observations for all three variables. The EXT observations exhibit a tendency to substantively influence the analyses for all three variables, which is an indication that the observations tend to depart substantively from the background fields in those areas. The stations aggregated into the hydrological category (HYDRO) are often located in remote high elevation sites and report primarily precipitation and temperature (see Table 1). Hence, of most interest is the large impact of the HYDRO temperature observations even though they have been assigned a higher observation to background error variance ratio to compensate for representativeness errors.

The LOCAL network category is the most complex and consists of a diverse mix of networks available to MesoWest (e.g., NWS Weather Forecast Office local networks, commercial, and other local networks). We have assigned these to have intermediate error characteristics reflecting differences in siting and reporting standards. The LOCAL and PUBLIC network categories are loosely comparable in overall characteristics and, hence, it is not too surprising that the impact of LOCAL observations falls above only the PUBLIC

438 observations. Finally, the transportation network category (TRANS) consists of commercial
439 rail and state Road Weather Information System (RWIS) stations. Overall, the impact of
440 TRANS stations is high, but that is regionally dependent. RWIS and rail stations in the
441 eastern half of the country tend to exhibit behavior common to other stations in urban areas
442 while those in the west tend to be located in a mix of urban and rural locations and have
443 larger impact on the analyses (not shown).
444
445 **4. Discussion**
446 Two-dimensional analyses on a 5 km grid over the continental United States of surface
447 temperature, moisture, and wind were computed for a sample of 100 cases using the
448 observation space form of the variational cost function equation (1). This study used the 1-h
449 forecast grids from the RUC data assimilation system downscaled by NCEP for the RTMA 5
450 km analyses as background fields combined with observations from nearly 20,000 locations.
451 These UU2DVAR analyses are now being computed routinely every hourly on a 2.5 km grid
452 using 2.5 km downscaled grids from the Rapid Refresh data assimilation system that is
453 undergoing testing for use by the RTMA.
454 The fundamental result of our study can be summarized by the oft-repeated phrase:
455 location, location, location. Our metric of impact draws attention to the location of the
456 observations in terms of the interplay between the weather conditions observed there, the
457 ability of the background field to diagnose the variability of those conditions, and the
458 proximity of other nearby observations. Of lesser importance are the types of network from

459 which the observations are obtained and the assumptions made regarding their observational
460 errors.

461 Solely in terms of their impact on high resolution analyses, observations in major
462 metropolitan areas tend to have reduced impact simply because there are so many other
463 stations tending to suggest similar adjustments to the background fields. On the other hand,
464 observations in more remote locations tend to have a higher impact. However, the
465 observation impact metric by itself can't distinguish between observations that have high
466 impact due to gross errors, representativeness errors, or failure of the background fields to
467 diagnose local weather conditions.

468 Attribution for the source of discrepancies between observations and background
469 grids is not clear cut and is all too often assumed to depend strongly on the type of station
470 rather than the dominant effect of location and the weather experienced there. For example,
471 the NWS observations of wind speeds at Cape Hatteras, NC (KHSE) are typically several
472 meters per second weaker than the background wind speeds (not shown explicitly but evident
473 in Fig. 5c as one of the high impact stations immediately offshore North Carolina). This
474 situation would be considered a representativeness error, since the reduced wind speeds only
475 reflect conditions observed in a narrow 1-3 km strip of land. It is unlikely that any data
476 assimilation system would reject the KHSE observations due to its bias because NWS
477 observations are generally considered to be well maintained and accurate. However, the only
478 other observing site on the Cape is located 8 km to the northeast and is a PUBLIC station
479 (Buxton, NC, D6557). Although that station would often be assumed a priori to provide
480 inferior observations, the bias of the observations relative to the background grid and their

impact are comparable to those at KHSE. The consistency in bias and impact between the two observing sites suggests that D6657 can provide valuable information for a variety of applications.

This study has been directed in part towards developing improved automated quality control algorithms for mesonet observations. The manual blacklist used to identify egregiously poor observations within the 100 h sample used in this study is not practical for routine use. The RTMA analyses generated by the National Centers for Environmental Prediction do rely in part on manual blacklists maintained and updated by National Weather Service forecasters around the country (de Pondeca et al. 2011). A problem with manual blacklists is the difficulty to determine when observations from a rejected station may no longer be in error.

As summarized by Fiebrich et al. (2010), many mesonets and mesonet aggregators (e.g., MesoWest and the Meteorological Assimilation Data Ingest System, MADIS) use automated quality control checks to identify erroneous observations. One common approach is to perform "buddy" checks that compare observations from one location with others nearby. While it is possible to do such checks for the 20,000 stations in the MesoWest database, the UU2DVAR analyses and associated bias and impact metrics provide a more efficient alternative approach to perform such spatial consistency checks. In areas of high data density (IDI values > 0.9 for example), biases and impacts at a station that disagree greatly from values at nearby stations indicate likely erroneous observations at that station.

The development of a national network of networks as recommended by the National Academy of Sciences panels (2009, 2010) should be a high priority. There is considerable

potential to develop a cost effective national network that takes advantage of the surface observations collected by hundreds of agencies, commercial firms, educational institutions, and the public. Understanding the strengths and weaknesses of the existing networks requires improved metadata combined with studies on the relative impact of those networks.

Acknowledgements

We wish to thank Xia Dong for her participation in the application of the UU2DVAR, Zachary Hansen for his assistance with the manual blacklisting of the observations, and Matthew Lammers for his investigations of observation bias and impact. This research was supported by the National Ocean and Atmospheric Administration under grant NA10NWS468005 as part of the CSTAR program.

References

Baker, N. L., and R. Daley, 2000: Observation and background adjoint sensitivity in the adaptive observation-targeting problem. *Q. J. R. Meteorol. Soc.*, **126,** 1431-1454.

Benjamin, S., and Coauthors, 2004: An hourly assimilation–forecast cycle: The RUC. *Mon. Wea. Rev.*, **132,** 495-518.

——, B. D. Jamison, W. R. Moninger, S. R. Sahm, B. E. Schwartz, T. 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.

Buizza, R., and A. Montani, 1999: Targeting observations using singular vectors. *J. Atmos. Sci.,* **56,** 2965-2985.

Dabberdt, W. F., and Coauthors, 2005: Multifunctional Mesoscale Observing Networks. *Bull. Amer. Meteor. Soc.*, **86**, 961–982.

Daley, R. and E. Barker, 2001: NAVDAS: Formulation and diagnostics. *Mon. Wea. Rev.*, **129,** 869-883.

de Pondeca, M. S. F. V., and Coauthors, 2011: The Real-Time Mesoscale Analysis at NOAA's National Centers for Environmental Prediction: Current status and development.

532 ——, and Coauthors, 2011: The Real-Time Mesoscale Analysis at NOAA's National Centers
533 for Environmental Prediction: Current status and development. *Wea. Forecasting,* **26**,
534 593-612.

535 Fiebrich, C. A., C. R. Morgan, A. G. McCombs, P. K. Hall, R. A. McPherson, 2010: Quality
536 Assurance Procedures for Mesoscale Meteorological Data. *J. Atmos. Oceanic*
537 *Technol.*, **27**, 1565–1582.

538 Glowacki, T. J., Y. Xiao, P. Steinle, 2012: Mesoscale Surface Analysis System for the
539 Australian Domain: Design Issues, Development Status, and System Validation. *Wea.*
540 *Forecasting*, **27**, 141–157.

541 Haiden, T., A. Kann, C. Wittmann, G. Pistotnik, B. Bica, and C. Gruber, 2011: The
542 Integrated Nowcasting through Comprehensive Analysis (INCA) system and its
543 validation over the eastern Alpine region. *Wea. Forecasting*, **26,** 166-183.

544 Horel, J., and Coauthors, 2002: Mesowest: Cooperative mesonets in the western United
545 States. *Bull. Amer. Meteor. Soc.*, **83,** 211-225.

546 ——, and B. Colman, 2005: Real-Time and retrospective mesoscale objective analyses. *Bull.*
547 *Amer. Meteor. Soc.*, **86**, 1477–1480.

548 ——, and X. Dong, 2010: An evaluation of the distribution of Remote Automated Weather
549 Stations (RAWS). *J. Appl. Meteor. Climatol.*, **49,** 1563-1578.

550 Langland, R. H., R. Gelaro, G. D. Rohaly, and M. A. Shapiro, 1999: Targeted observations in
551 FASTEX: Adjoint-based targeting procedures and data impact experiments in IOP17
552 and IOP18. *Q. J. R. Meteorol. Soc.*, **125,** 3241-3270.

553 ——, and Baker, N. L, 2004: Estimation of observation impact using the NRL atmospheric
554 variational data assimilation adjoint system. *Tellus A*, **56**, 189–201.

555 Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Q. J. R. Meteorol.*
556 *Soc.*, **112,** 1177-1194.

557 McPherson, R. A., and Coauthors, 2007: Statewide Monitoring of the Mesoscale
558 Environment: A Technical Update on the Oklahoma Mesonet. *J. Atmos. Oceanic*
559 *Technol.*, **24**, 301–321.

560 Myrick, D. T. and J. D. Horel, 2008: Sensitivity of surface analyses over the western United
561 States to RAWS observations. *Wea. Forecasting*, **23,** 145-158.

562 National Academy of Sciences, 2009: *Observing Weather and Climate from the Ground Up:*
563 *A Nationwide Network of Networks.* National Academy Press, 234 pp.

564 ——, 2010: *When Weather Matters: Science and Service to Meet Critical Societal Needs.*
565 National Academy Press, 199 pp.

566 Palmer, T. N., R. Gelaro, J. Barkmeijer, and R. Buizza, 1998: Singular vectors, metrics, and
567 adaptive observations. *J. Atmos. Sci.*, **55,** 633-653.

568 Seaman, R. S., and M. F. Hutchinson, 1985: Comparative real data test of some objective
569 analysis methods by withholding observations. *Aust. Meteorol. Mag.*, **33,** 37-46.

570 Steinacker, R., and Coauthors, 2006: A mesoscale data analysis and downscaling method
571 over complex terrain. *Mon. Wea. Rev.,* **134,** 2758-2771.

572 Tyndall, D. P., J. D. Horel, and M. S. F. V. de Pondeca, 2010: Sensitivity of surface air
573 temperature analyses to background and observation errors. *Wea. Forecasting*, **25,**
574 852-865.

575 ——, 2011: Sensitivity of surface meteorological analyses to observation networks. Dept. of
576 Atmospheric Sciences, University of Utah, 155 pp. [Available online at
577 http://content.lib.utah.edu/cdm4/document.php?CISOROOT=/us-
578 etd3&CISOPTR=61114.]

579 Uboldi, F., C. Lussana, and M. Salvati, 2008: Three-dimensional spatial interpolation of
580 surface meteorological observations from high-resolution local networks. *Meteor.*
581 *Appl.*, **15,** 331-345.

582 Xie, Y., S. Koch, J. McGinley, S. Albers, P. E. Bieringer, M. Wolfson, and M. Chan, 2011:
583 A space–time multiscale analysis system: A sequential variational analysis approach.
584 *Mon. Wea. Rev.,* **139,** 1224-1240.

585 Zapotocny, T. H., and Coauthors, 2000: A case study of the sensitivity of the Eta Data
586 Assimilation System. *Wea. Forecasting*, **15,** 603-621.

587 Zhu, Y., and R. Gelaro, 2008: Observation sensitivity calculations using the adjoint of the
588 Gridpoint Statistical Interpolation (GSI) analysis system. *Mon. Wea. Rev.*, **136,** 335-
589 351.

590 **Figure Captions**

591 FIG. 1. IDI analysis over the CONUS domain for surface observations that reported at least
592 50% of the 100 analysis hours used in this study.

593 FIG. 2. Hydrological Prediction Center 00 UTC 04 April 2011 surface analysis showing
594 position of frontal structures and dry line over CONUS.

595 FIG. 3. Meteorological surface analyses (left panels) and analysis increments (right panels)
596 for 00 UTC 04 April 2011 over CONUS.

597 FIG. 4. Observation impact percentile for 2-m temperature observations used in the 0000
598 UTC 04 April 2010 temperature analysis from 4 network categories.

599 FIG. 5 (a) Median impact percentiles for temperature observations computed over 100
600 analysis hours for 4 selected network categories. (b) As in (a) except for dew point
601 temperature. (c) As in (a) except for wind speed.

602 Fig. 6. As in Fig. 5a except for northern Utah for all 10 network categories. 5-km analysis
603 terrain shaded in m, with blue areas denoting water gridpoints. Location of the METAR
604 observation at Salt Lake International Airport is marked in the EXT panel.

605 FIG. 7. As in Fig. 6 except for wind speed observations over southern California. Locations
606 of the METAR observations at Los Angeles and San Diego International Airports marked in
607 the FED+ panel.

608 FIG. 8. Fraction of reports from each network category with observation impacts in the upper
609 quartile over 100 analysis hours for temperature, dewpoint, and wind speed. The assumed
610 observation to background error variance ratio (see Table 1) is labeled at the top of each bar.

TABLE 1. Mesonet categories based on purpose and type of network, total number of stations and the median IDI value for that category, and the number of observations and assumed observation to background error variance ratio for each variable.

					Temperature		Dewpoint		Wind	
Category	**Group Purpose/Type**	**No. of Networks**	**Total No. of Stations**	**Median IDI**	**No.**	σ_o^2/σ_b^2	**No.**	σ_o^2/σ_b^2	**No.**	σ_o^2/σ
NWS	NWS/FAA	1	1814	0.94	1751	1.0	1747	1.0	1733	1.0
FED+	Federal and state networks	21	849	0.95	696	1.0	513	1.0	470	1.0
RAWS	Fire weather	1	1986	0.86	1736	2.0	1729	2.0	1674	2.0
PUBLIC	Primarily Citizen Weather Observing Program (CWOP)	3	6808	0.96	5263	1.5	5168	1.5	4842	2.0
AG	Agricultural	9	472	0.94	440	1.5	434	1.5	413	2.0
AQ	Air quality	8	796	0.96	522	1.5	227	1.5	650	1.5
EXT	Offshore, Canadian, Mexican	6	940	0.71	755	1.5	392	1.5	628	1.5
HYDRO	Hydrological	11	3580	0.85	1411	2.0	151	2.0	207	2.0
LOCAL	Commercial, state, and local	41	799	0.94	610	1.5	542	1.5	492	1.5
TRANS	Road and rail weather	24	1653	0.93	1538	1.5	1116	1.5	1076	1.5
	TOTAL	125	19697	-	14722	-	12019	-	12185	-

TABLE 2. Sample of 25 high impact weather days used to evaluate observation impacts.

Date	Weather	Region	Date	Weather	Region
1 Oct. 2010	Floods	NY, PA	15 Jan. 2011	Flood; Wildfire	WA, ID; SC
2 Oct. 2010	Hail	NM	16 Jan. 2011	Flood	ID, OR
5 Oct. 2010	Hail, Flash Flood	AZ, NV	1 Feb. 2011	Cold	TX
24 Oct. 2010	Tornado	TN, TX	20 Feb. 2011	Ice Storm	MI
26 Oct. 2010	Tornado	WI	27 Feb. 2011	Wildfire	TX
27 Oct. 2010	Tornado	MN; NC	28 Feb. 2011	Wildfire; Flood	TX; IN, OH, TN
22 Nov. 2010	Wind; Tornado	IL, WI	7 Mar. 2011	Flood	NJ
10 Dec. 2010	Winter Storm	MN	24 Mar. 2011	Flood	CA
12 Dec. 2010	Flood	WA	4 Apr. 2011	Tornado	AR, KY, LA, MS, OH, TN
19 Dec. 2010	Flood	CA, UT	5 Apr. 2011	Tornado	GA, NC, KY
21 Dec. 2010	Flood	AZ, CA, NV	25 Apr. 2011	Tornado	AL, AR, GA, LA, MS, NC, TN, TX, VA
31 Dec. 2010	Tornado	AR, IL, MS	26 Apr. 2011	Tornado	“”
			27 Apr. 2011	Tornado	“”

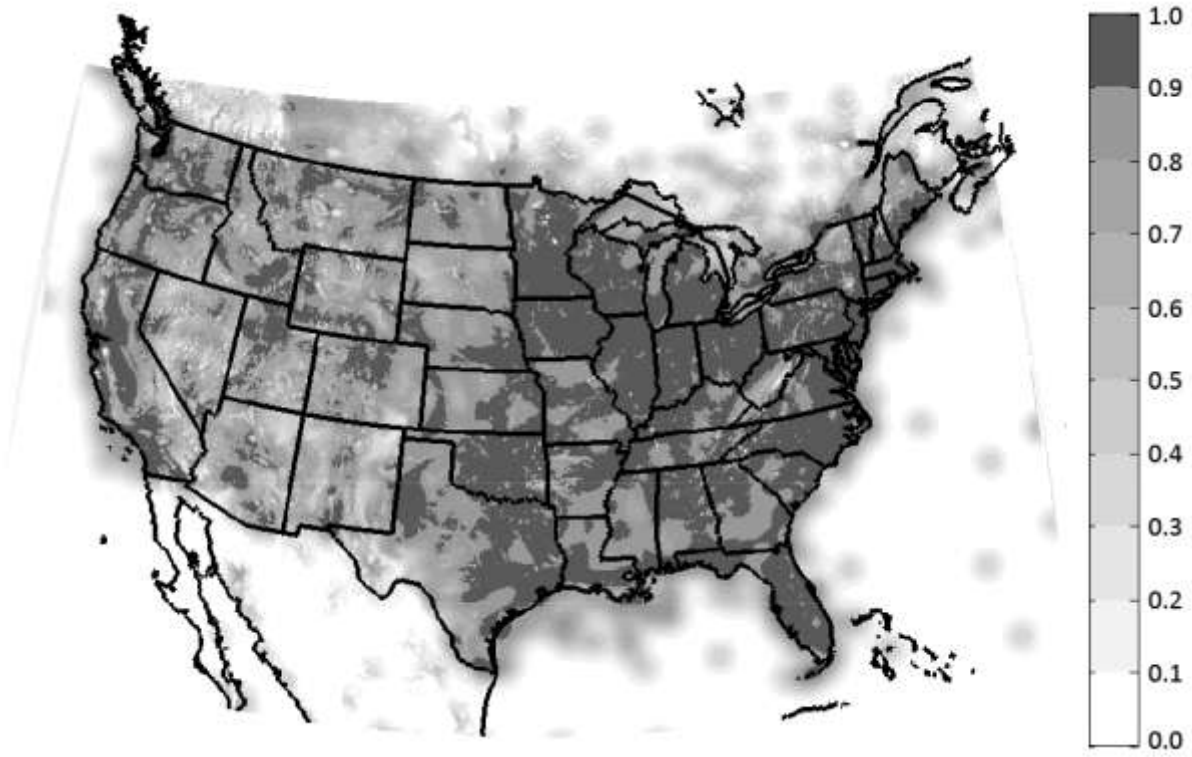

Fig. 1. IDI analysis over the CONUS domain for surface observations that reported at least 50% of the 100 analysis hours used in this study.

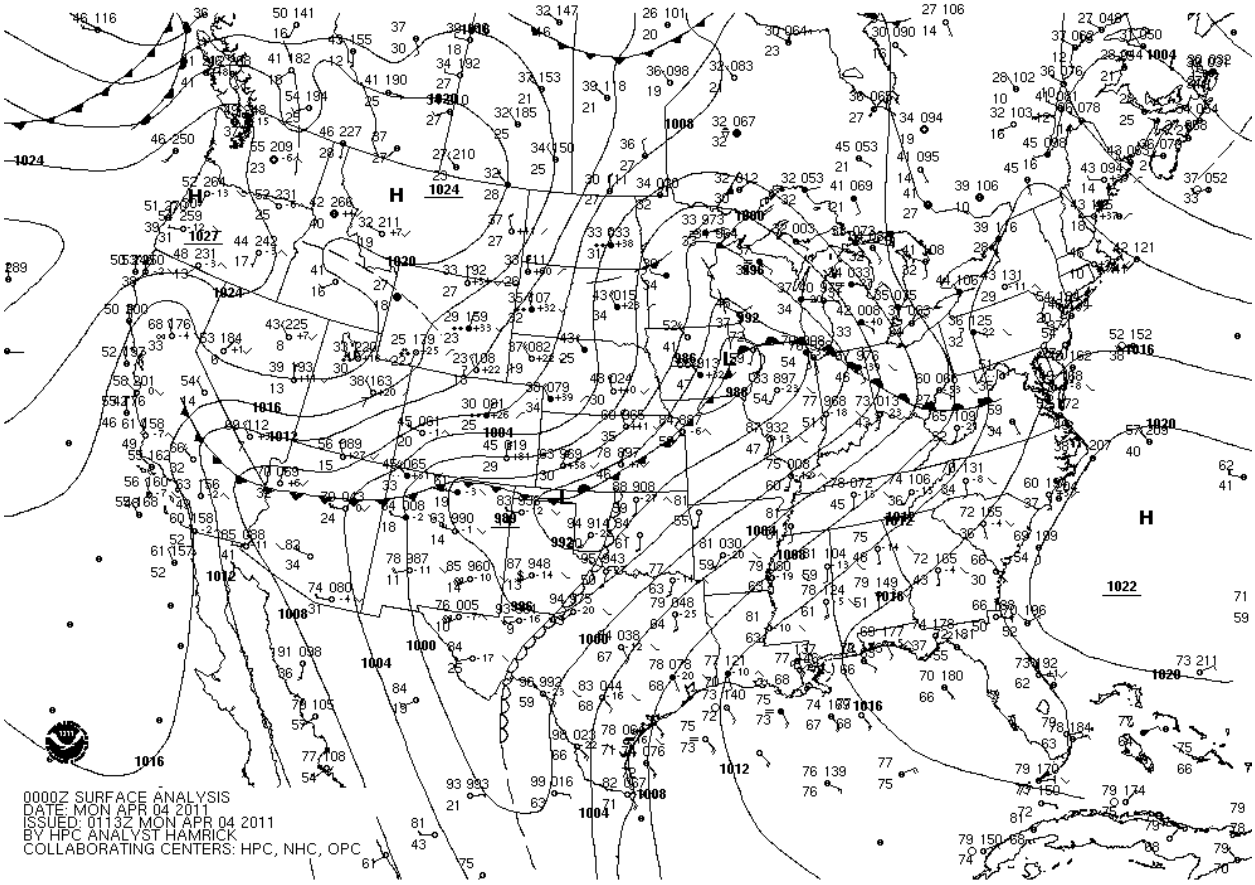

Fig. 2. Hydrological Prediction Center 00 UTC 04 April 2011 surface analysis showing position of frontal structures and dry line over CONUS.

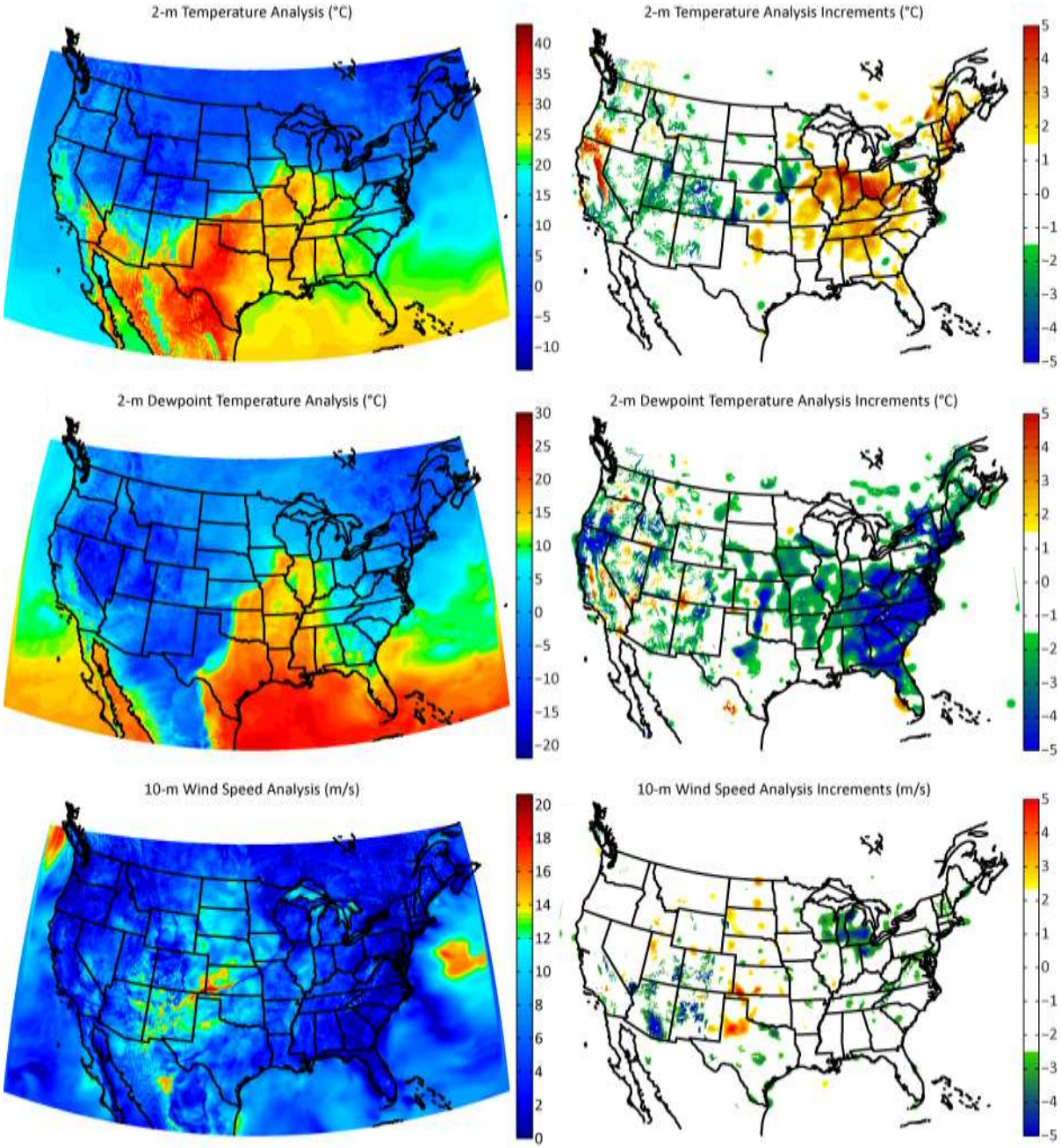

Fig. 3. Meteorological surface analyses (left panels) and analysis increments (right panels) for 00 UTC 04 April 2011 over CONUS.

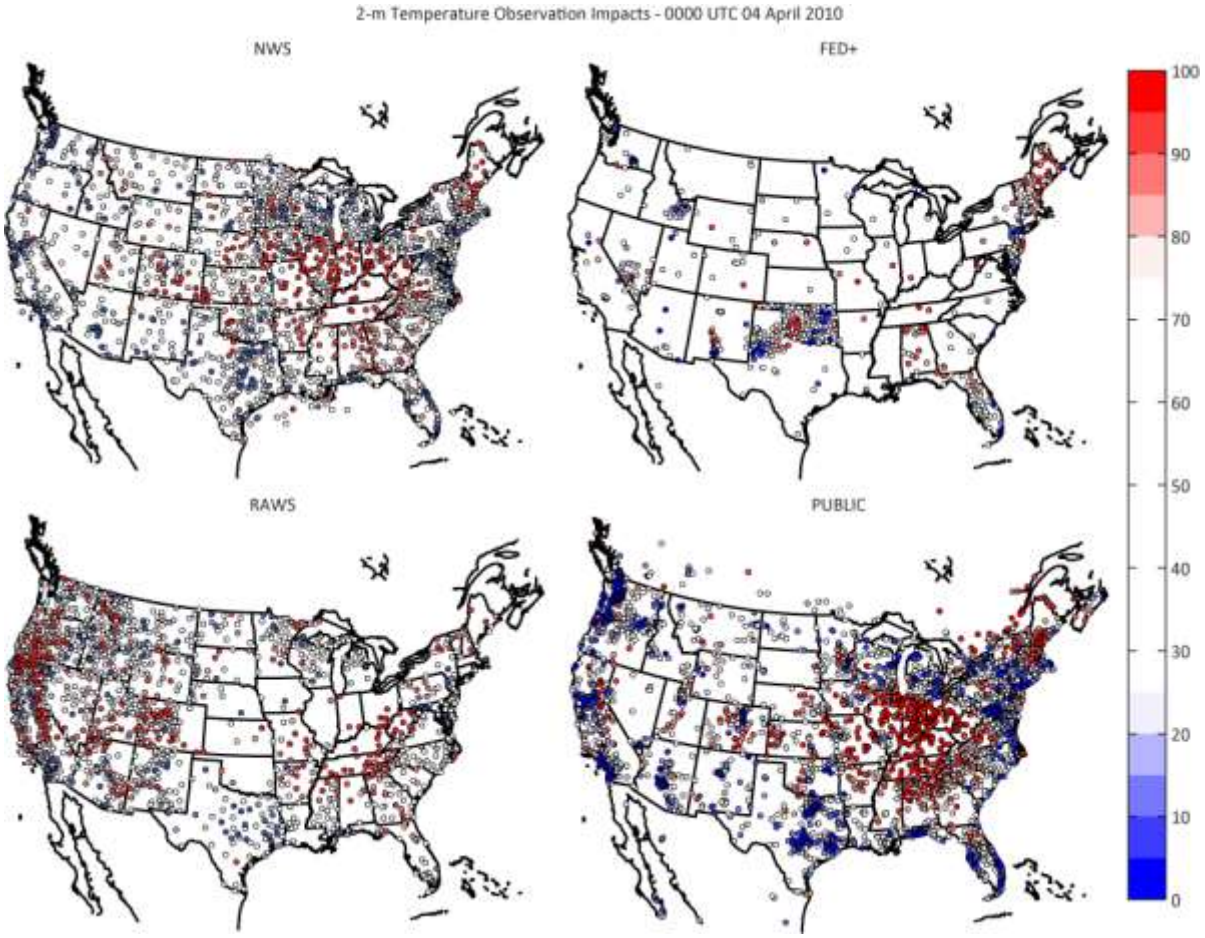

Fig. 4. Observation impact percentile for 2-m temperature observations used in the 0000 UTC 04 April 2010 temperature analysis from 4 network categories.

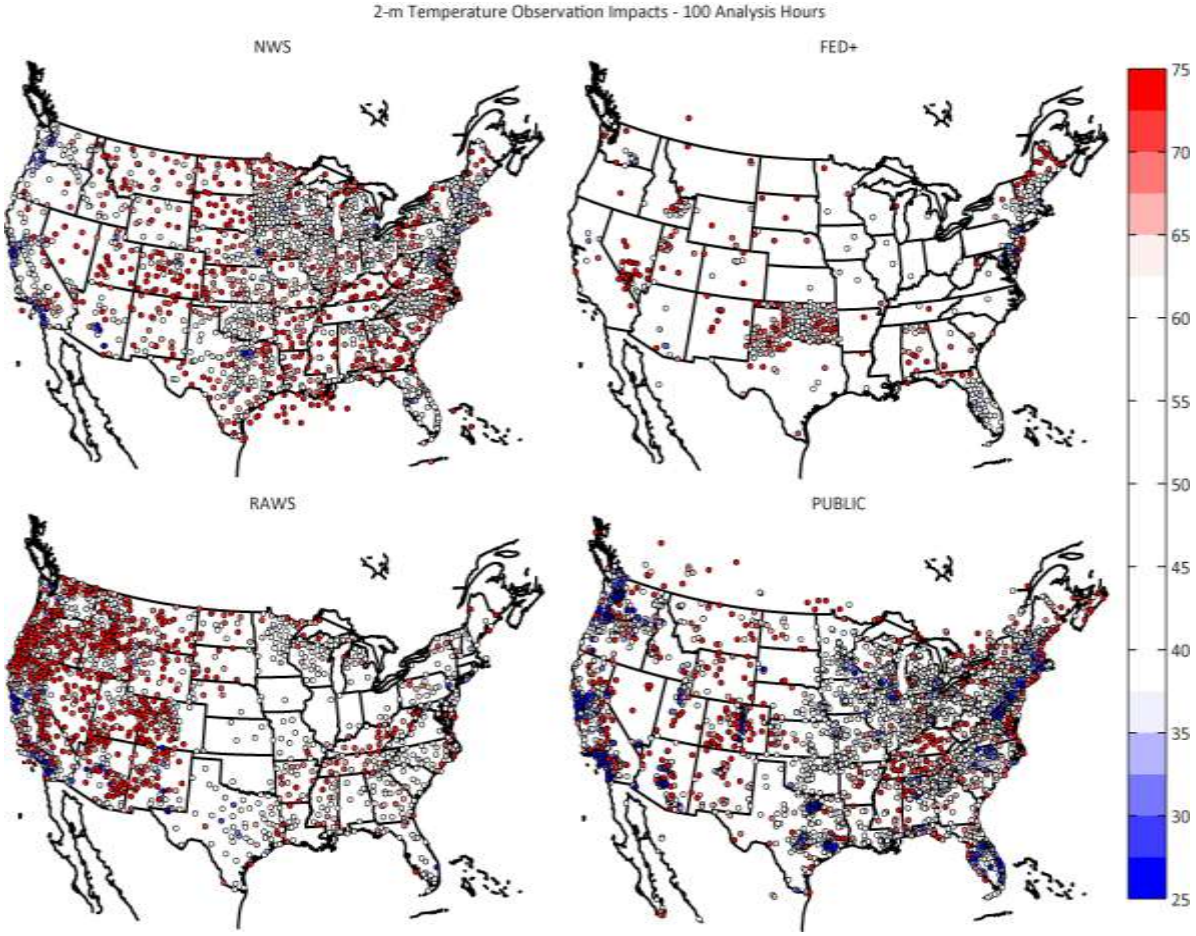

Fig. 5 (a) Median impact percentiles for temperature observations computed over 100 analysis hours for 4 selected network categories.

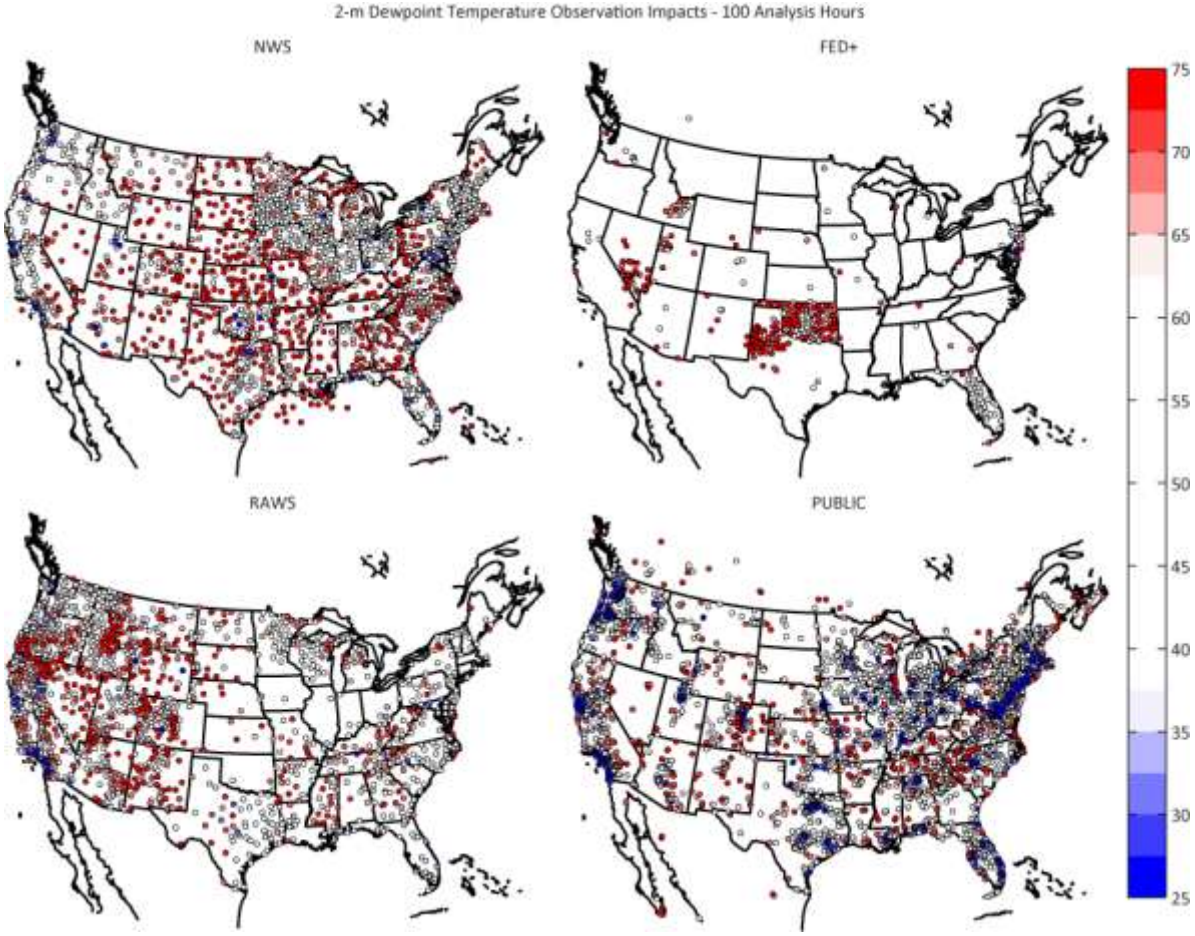

Fig. 5b. As in Fig. 5a except for dew point temperature.

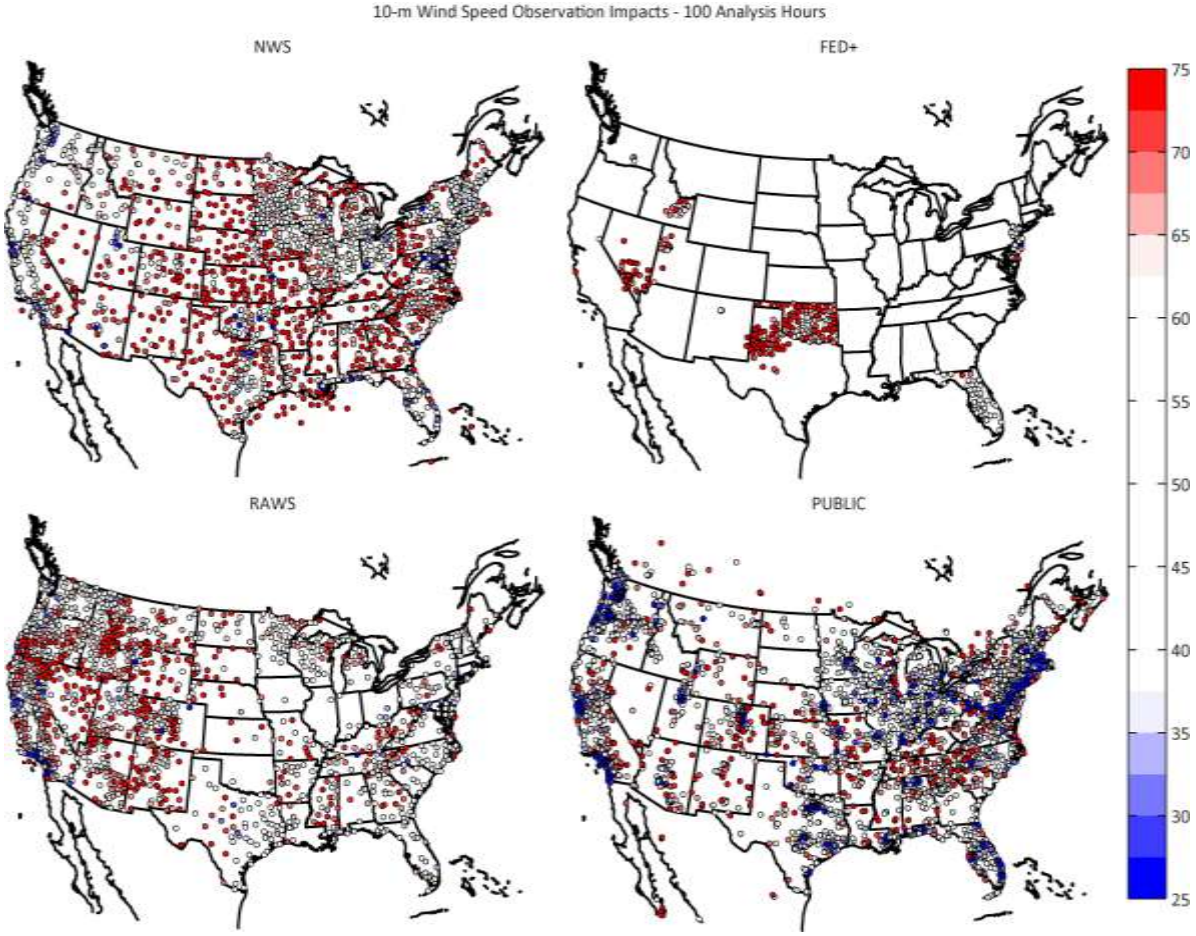

Fig. 5c. As in Fig. 5a except for wind speed.

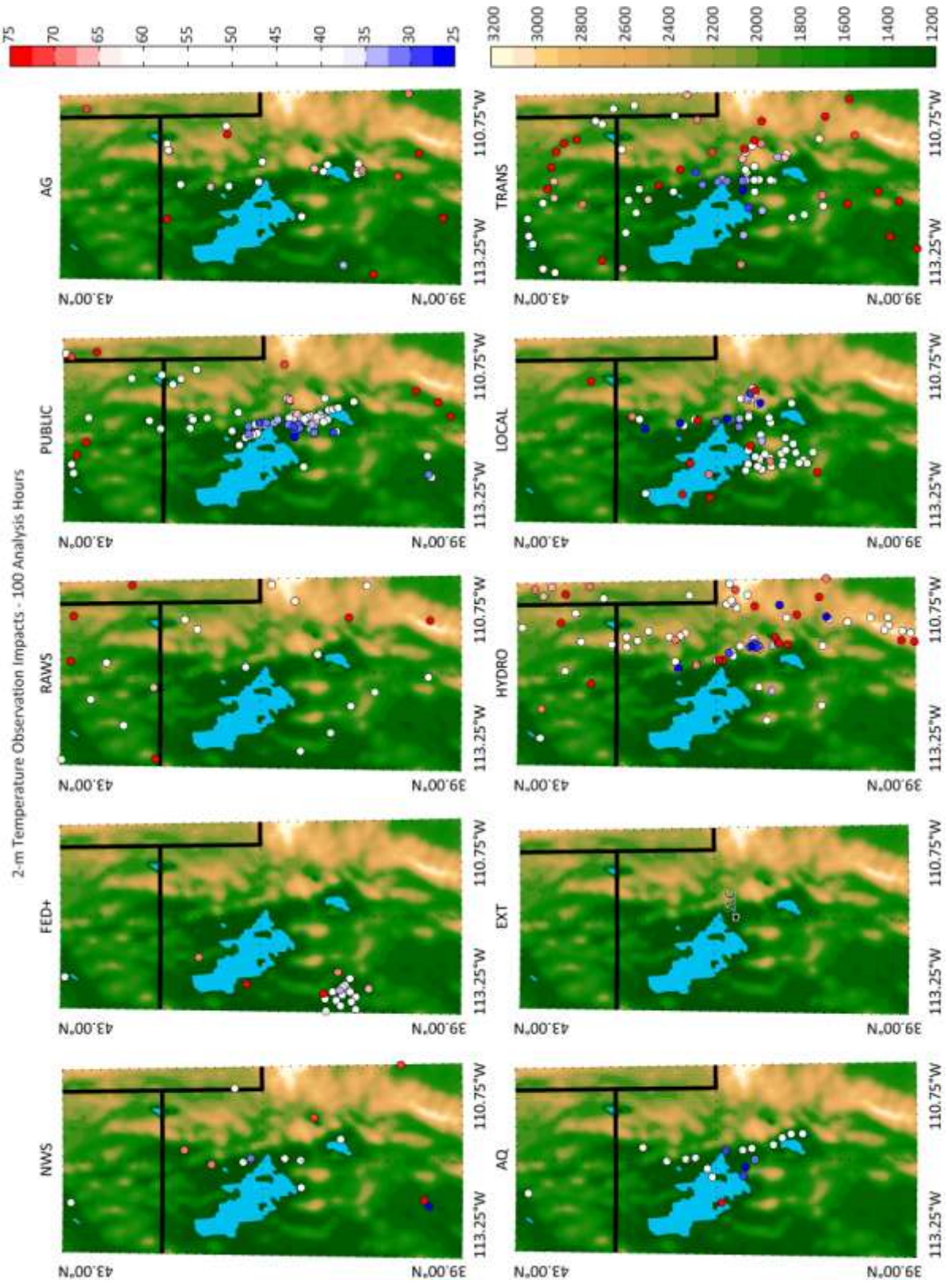
2-m Temperature Observation Impacts - 100 Analysis Hours
75
70
65
60
55
50
45
40
35
30
25
3200
3000
2800
2600
2400
2200
2000
1800
1600
1400
1200
AG
TRANS
PUBLIC
LOCAL
RAWS
HYDRO
FED+
EXT
NWS
AQ
43.00°N
39.00°N
113.25°W
110.75°W

Fig. 6. As in Fig. 5a except for northern Utah for all 10 network categories. 5-km analysis terrain shaded in m, with blue areas denoting water gridpoints. Location of the METAR observation at Salt Lake International Airport is marked in the EXT panel.

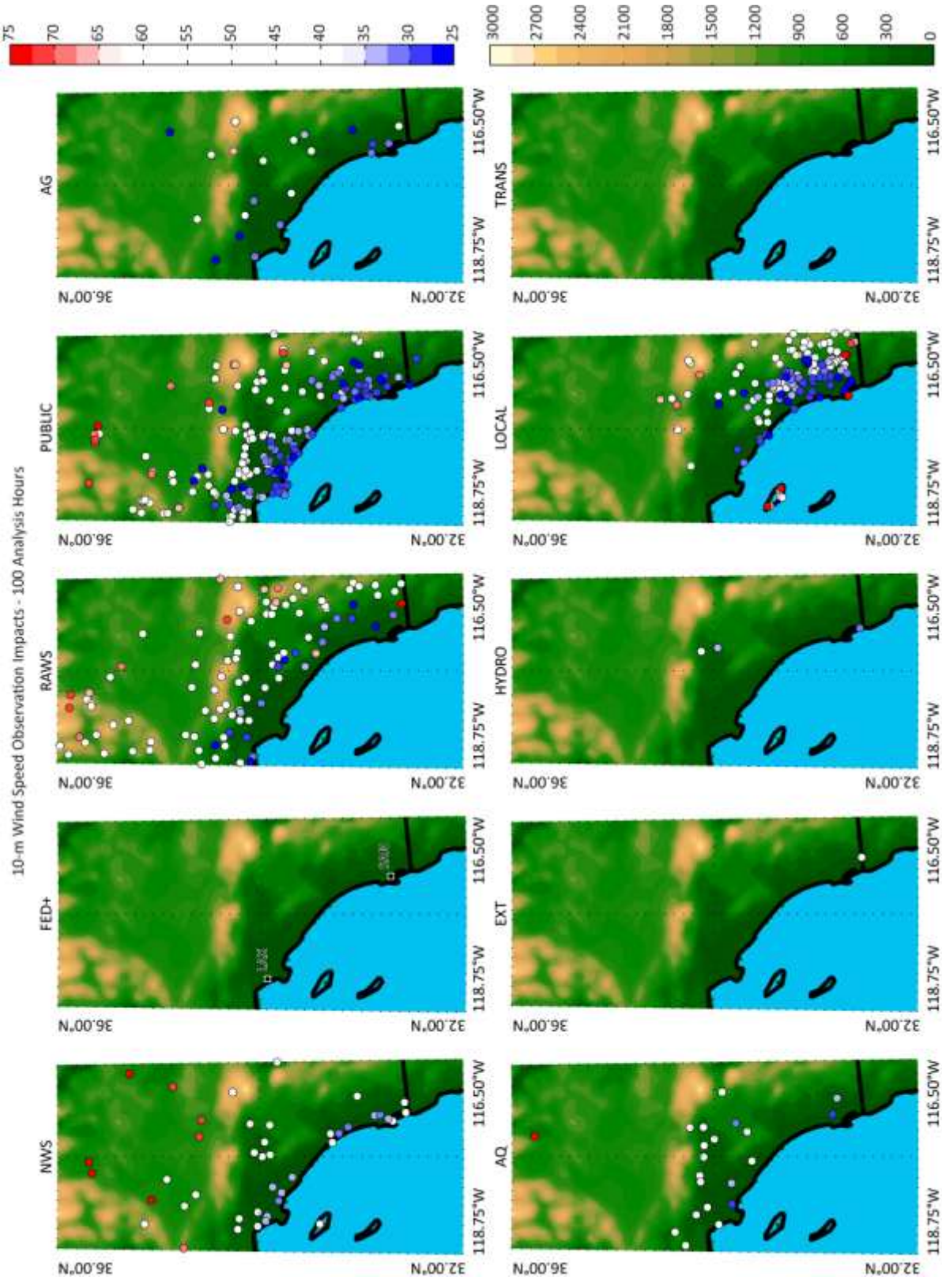
10-m Wind Speed Observation Impacts - 100 Analysis Hours
75
70
65
60
55
50
45
40
35
30
25
3000
2700
2400
2100
1800
1500
1200
900
600
300
0
NWS
FED+
RAWS
PUBLIC
AG
AQ
EXT
HYDRO
LOCAL
TRANS
36.00°N
32.00°N
118.75°W
116.50°W

Fig. 7. As in Fig. 6 except for wind speed observations over southern California. Locations of the METAR observations at Los Angeles and San Diego International Airports marked in the FED+ panel.

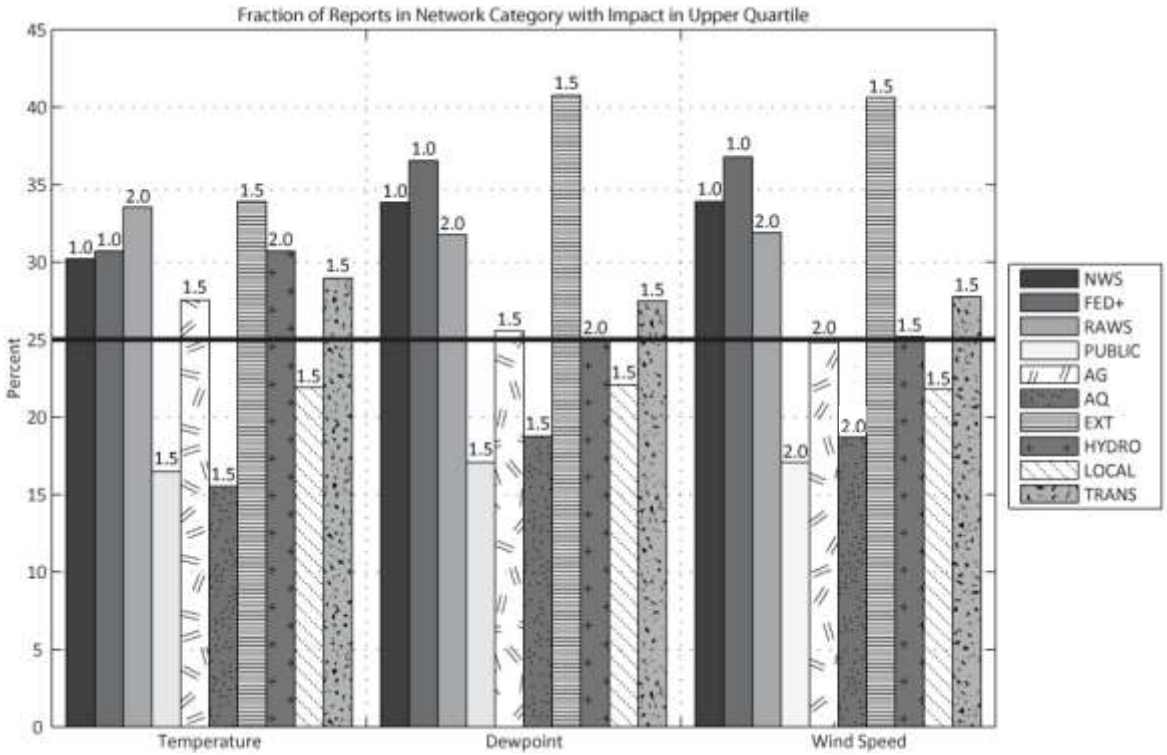

Fig. 8. Fraction of reports from each network category with observation impacts in the upper quartile over 100 analysis hours for temperature, dewpoint, and wind speed. The assumed observation to background error variance ratio (see Table 1) is labeled at the top of each bar.