AN EVALUATION OF THE DISTRIBUTION OF REMOTE AUTOMATED WEATHER STATIONS (RAWS)

John D. Horel [1] and Xia Dong
Department of Atmospheric Sciences, University of Utah
Salt Lake City, UT

Revision Submitted to: *Journal of Applied Meteorology and Climatology*
February 15, 2010

[1] Corresponding author address:
University of Utah
Department of Atmospheric Sciences
135 South 1460 East, Rm. 819
Salt Lake City, UT 84112-0110
Email: john.horel@utah.edu

Abstract

This study estimates whether surface observations of temperature, moisture, and wind at some stations in the continental United States are less critical than others for specifying weather conditions in the vicinity of those stations. Two-dimensional variational analyses of temperature, relative humidity, and wind were created for selected midday hours during summer 2008. This set of 8,925 control analyses was derived from 5 km resolution background fields and Remote Automated Weather Station (RAWS) and National Weather Service (NWS) observations within roughly 4° x 4° latitude-longitude domains. Over 570,000 cross validation experiments were completed to assess the impact of removing each RAWS and NWS station.

The presence of observational assets within relatively close proximity to one another is relatively common. The sensitivity to removing temperature, relative humidity, or wind observations varies regionally and depends on the complexity of the surrounding terrain and the representativeness of the observations. Cost savings for the national RAWS program by removing a few stations may be possible. However, nearly all regions of the country remain undersampled, especially mountainous regions of the western United States frequently affected by wildfires.

1. Introduction

The Remote Automated Weather Station (RAWS) network, which includes nearly 2200 stations in the United States, has been developed for a wide range of applications by federal and state land management government agencies (Forest Service, Bureau of Land Management, National Park Service, etc.). In addition to their core uses for monitoring and predicting weather conditions and fire behavior, RAWS data are also used extensively for smoke and related air quality issues, resource management and many research applications (Zachariassen et al. 2003, Hoadley et al. 2004, Huang et al. 2009). The National Wildfire Coordinating Group for the land management agencies requested during 2008 an evaluation of the weather information provided by the RAWS network in the context of agency needs and the expense of installing, maintaining, and operating those weather stations. The replacement cost for the entire RAWS network in the continental United State is estimated to be over $30 million while annual maintenance costs of the existing network are on the order of $3 million.

A recent report by the National Academy of Science (National Research Council 2009) summarizes the needs for improved monitoring of atmospheric conditions near the earth's surface across the United States as well as improved coordination among the government agencies, commercial firms, and educational institutions that provide access to those observations. The land management agencies through their support of the RAWS network will play a critical role in any future national "network of networks" because those agencies are particularly interested in the atmospheric state in typically undersampled remote locations, as well as within increasingly important urban-wildland interface regions.

Myrick and Horel (2008) evaluated the impact of RAWS observations on winter temperature and wind analyses in the western United States. That study demonstrated the considerable positive impact of the RAWS observations on winter analyses, which results from the sparse distribution of other stations in the mountainous regions of the West. Tyndall et al. (2010) developed a local surface analysis (LSA) system that can be used for cross validation studies to examine the analysis sensitivity to the distribution of observations.

This study is part of a broader effort underway in collaboration with researchers at the Desert Research Institute that is intended to provide guidance on the present distribution of RAWS and identify regions of the nation with redundant observing capabilities as well as those with large observing gaps. The goal of this phase of the broader study is not to make recommendations to remove or add specific RAWS, as that requires consideration of a wide range of factors. Rather, our objective is to develop procedures to estimate the impact of removing RAWS observations from analyses of temperature, moisture, and wind. We examine the impact of removing RAWS observations on such analyses relative to the impact of removing Automated Surface Observing System (ASOS) and Automated Weather Observing System (AWOS) observations from primarily airport locations that are disseminated by the National Weather Service (NWS). For convenience, we will refer to the ASOS and AWOS observations as NWS observations hereafter. The land agencies are interested in estimates of potential savings arising from eliminating RAWS located close to other RAWS as well as RAWS located near NWS sites. Beyond the scope of this project is an assessment of the utility of the many (>15,000) additional surface observing stations available around the country that are often used for fire weather applications, for example, through the web portal referred to as the Real-Time Observation Monitor and Analysis Network (http://raws.wrh.noaa.gov).

The data and methods used in this study are presented in the following section. The conditions observed at the time of a major fire outbreak in the foothills of the Sierra Mountains in northern California (10 July 2008) are used to illustrate our LSA system in Section 3. Results from data denial experiments for this region of northern California are also presented in Section 3 as well as summary statistics for the continental United States as a whole. A summary and discussion follow in Section 4.

2. Data and Method

a. RAWS and NWS Data

The RAWS network is the largest government-supported observing network in the western United States with widespread coverage in other federally and state managed lands around the country as well. We divided the continental United States into 51 roughly 4^o x 4^o latitude-longitude domains as shown in Fig. 1. The counts of RAWS in each subdomain available during the summer 2008 period are provided in Fig. 1, with as many as 124 in southern California and as few as 8 in sections of Texas. While many of the RAWS are permanent sites located to be representative of weather conditions within the surrounding region, others are transient and are moved to meet agency needs within local land management areas. In addition, some are repositioned as needed to locations anywhere within the nation in support of fire suppression operations. Temperature, relative humidity, wind speed and direction, peak wind speed, and precipitation are typically measured at all RAWS with additional solar radiation, pressure, soil moisture, and fuel temperature and moisture sensors at many sites. Maintenance and reporting standards are established on a national basis with once per hour reports transmitted by GOES satellite (National Wildfire Coordinating Group 2008).

NWS stations tend to be located at or near airports and the recommended routine reporting time is immediately prior to the top of each hour. As shown in Fig. 1, more RAWS are available in western subdomains (as well as an Appalachian region and the western Lake Superior region) compared to NWS stations while relatively few RAWS are located over the plains. Table 1 contrasts some of the key differences between NWS and RAWS observations.

While it is misleading to simply evaluate the spatial coverage of the RAWS and NWS networks on the basis of lateral separation between stations, Fig. 2 shows the distance between each RAWS and its nearest RAWS or NWS neighboring station. The median distance between a RAWS and its nearest neighboring RAWS is 29 km, but that distance drops to 23 km if NWS stations are considered. There are 146 pairs of RAWS located within 10 km of one another while an additional 127 RAWS have a NWS station within 10 km. There is often considerable justification for having stations within relatively close proximity to one another, especially in mountainous areas. However, there is also a need to examine whether operational costs might be reduced by eliminating some stations located close to other observing assets.

RAWS and NWS observations nearest to and within ±30 min of 1800 UTC (roughly midday across the country) on 35 days between 25 May and 17 September 2008 were accessed from the MesoWest data archive for this study (Horel et al. 2002). A total of 1688 (1589) RAWS (NWS stations) are used in this study and over 90% of them reported 32 of the 35 possible times. Obviously, a sample of 35 afternoons is not sufficient to describe all of the synoptic, mesoscale, and local weather conditions observed during summer across the entire country. However, the number of cross validation experiments that could be reasonably performed with the available computational resources limited the sample size. Results from one domain for an entire 24 h period will be used to corroborate the core results.

As a means to quality control the observations, some critical thresholds are set for observations of temperature, wind, relative humidity (Table 2). In Table 2, y_i is the i^{th}

observation, x_b^i and x_c^i are the corresponding background and control analysis values respectively described in the next section. If the absolute value of the difference between an observation and background (control analysis) is greater than the threshold, then the observation is discarded and not used in the data denial experiments. As summarized in Table 2, less than 1% of the observations are discarded for any of the variables with wind observations rejected the least.

b. Local Analyses

Following Tyndall et al. (2010), control analyses were computed using MATLAB software for temperature, relative humidity, wind speed, and zonal and meridional wind components for each of the 35 cases and for each of the 51 domains shown in Fig. 1. Hence, a total of 8925 control analyses were computed for this study. The univariate analyses of zonal and meridional wind are used to compute the analysis wind direction. Tyndall et al. (2010) computed the analyses for the limited domain using the General Minimum Residual method (Saad and Schultz 1986) to solve the basic 2DVar analysis equations:

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

$$\mathbf{x_a} = \mathbf{x_b} + \mathbf{P_b}\mathbf{v} \tag{2}$$

In these two equations, $\mathbf{y_o}$ is the observation dataset, $\mathbf{x_b}$ is the background field, $\mathbf{H}$ is the linear forward operator used to transform analysis gridpoints to the observation locations, and $\mathbf{P_b}$ and $\mathbf{P_o}$ are the background and observation error covariance respectively. The term $\mathbf{v}$ is solved iteratively requiring typically ~50 iterations to yield the analysis, $\mathbf{x_a}$. Because of the very large number of analyses to be completed for this study, reducing the computation time necessary for each analysis was a high priority. In order to do so, we took advantage of the sparse nature of $\mathbf{P_o}$ (diagonal elements only) and $\mathbf{H}$ (all 0 except for one value of 1 in each column). We define $\mathbf{T} = \mathbf{P_b}\mathbf{H}^\mathbf{T}$ and compute that matrix without performing matrix multiplication. Defining $\mathbf{u} = \mathbf{P_b}\mathbf{v}$, then (1) can be rewritten:

$$(\mathbf{P_o} + \mathbf{TH})\mathbf{u} = \mathbf{T}(\mathbf{Y_o} - \mathbf{H}(\mathbf{x_b})) \tag{3}$$

That linear system can be solved in the same manner as before in ~7 iterations. The cumulative effect of all the changes reduces the required computation time by a factor of roughly four compared to the method used by Tyndall et al. (2010). The resulting analyses are identical to those obtained by solving (1).

These local control analyses use the 1-h forecasts from the Rapid Update Cycle (RUC), which is the National Centers for Environmental Prediction short-range operational weather prediction system (Benjamin et al. 2004). The downscaling procedures using a 5-km resolution terrain field are described by Jascourt (2010). The background fields provide spatial and temporal physical consistency over the continental United States since they are derived from a three dimensional atmospheric weather prediction model that assimilates observations each hour.

The local analyses depend on the specification of the observation and the background error covariance. Tyndall et al. (2010) examined the sensitivity of local analyses to the parameters used to specify those error covariances: the ratio of observation to background error variance and the horizontal and vertical decorrelation length scales used to define the background error covariance fields. We chose to use the original values examined by Tyndall et al. (2010) (i.e., a 1:1 observation to background error variance ratio and 40 km horizontal and 100 m vertical decorrelation length scales). We will discuss in Section 3a the sensitivity of our results to these choices.

c. Cross validation experiments

In order to assess the impact of removing each RAWS, cross validation experiments were conducted in which all observations used in each domain's control analysis were used except for those at the one RAWS location. Then, this leave-one-out cross validation procedure (Wilks 2006) was repeated removing the next RAWS observation, etc. A similar set of cross validation experiments was completed removing sequentially each NWS station from the analyses. Thus, a total of over 570,000 local analyses were created for this study.

As a means to synthesize the results obtained from these large sets of data denial experiments, three metrics are calculated related to analysis accuracy, sensitivity, and degradation as defined as follows. Following many cross validation studies (Steinacker et al. 2006; Uboldi et al. 2008), the accuracy of objective analyses is estimated by calculating the cross-validation (CV) score, i.e. the root-mean-square difference between the values of the analyses at the i^{th} withheld observation location and the observations at that location:

$$CV^i = \sqrt{\frac{\sum_N \left(y_o^i - x_w^i\right)^2}{N}} \tag{4}$$

where x_w^i is the value at the i^{th} location of the withheld analysis, y_o^i is the observation at that location, and N is the total number of observations (a maximum of 35). If the withheld observations tend to be close to the analysis values that don't use those observations, then the CV score is small. However, the CV score values depend on variable and location, i.e., temperature CV scores along coasts are generally lower than those over mountains simply because there is less temperature variability in the former relative to the latter. In addition, all of the metrics, depend on the assumptions used to generate the analyses regarding the observational and background error covariances.

Following Zapotocny et al. (2000), analysis sensitivity is defined as the root mean square difference between control and withheld analyses computed over the entire grid for the i^{th} set of denial experiments:

$$s^i = \sqrt{\sum_N \sum_{j=1}^{L} \frac{\left(x_i^j - x_c^j\right)^2}{NL}} \tag{5}$$

where x_c^j is the j^{th} analysis value from the control analysis that uses all observations, x_i^j is the j^{th} analysis value for the i^{th} data withholding experiment, and L is the total number of analysis grid points in the subdomain. The analysis sensitivity indicates the magnitude of the change over the entire analysis domain resulting from withholding data; a small value of S implies that the analysis is largely unaffected by the removal of the observations. As with analysis accuracy, the sensitivity values depend on variable and location.

Analysis degradation is a nondimensional measure similar to Wilmott's index of agreement (Paimazumder and Molders 2009) and proportional to the ratio of squared errors of the control analyses to the squared error of the withheld analyses, where x_c^i is the control analysis value at the i^{th} location:

$$AD^i = 100 * \left(1 - \sqrt{\sum_N \frac{\left(y_o^i - x_c^i\right)^2}{\left(y_o^i - x_w^i\right)^2} \Big/ N}\right) \tag{6}$$

Small analysis degradation implies that the differences between the observations and withheld analysis values at the observation locations are of similar magnitude to the differences between the observations and the control values at the same locations. If the control analysis adequately describes the local conditions, then small analysis degradation indicates that withholding those observations does not have a large effect on the resulting analyses. However, a small analysis degradation may result as well from observations that are representative of local weather conditions (of interest for fire weather applications), which often are not captured adequately in the control analyses.

3. Results

a. Northern California domain

Lightning on 20 June 2008 led to a major outbreak of wildfires in northern California. High winds and temperatures combined with low relative humidity during most of the next six weeks contributed to over 12,000 km^2 burned in central and northern California, more than 500 structures destroyed, and 15 lives lost. Near the peak period of this fire outbreak, 10 July, thousands were evacuated and over 50 homes destroyed near Paradise, CA in the western foothills of the northern Sierra Mountains. Figure 3a shows the analysis domain (of the 51 such regions) that encompasses the Sacramento Valley, central and northern Sierra Mountains, and location of the Paradise fire.

We begin by showing in Fig. 3b the Integral Data Influence (IDI) field as defined by Uboldi et al. (2008). This field is computed by assuming background values are everywhere zero and all observation values are one. The IDI depends on the characteristics of the terrain in combination with the assumptions regarding observation and background error covariances. Large (small) IDI values define areas of the control analysis likely to be strongly (weakly) influenced by observations. In other words, white regions are data voids in our study where control analyses remain close to the background fields. Tests were conducted increasing the extent to which the background errors remain correlated as a function of spatial distance (not shown). The influence of discrepancies between observations and the background away from the location of interest can thereby be increased, at the risk of missing local weather conditions.

Figure 4 shows the control analyses of temperature, relative humidity, and wind valid at 1800 UTC 10 July 2008. RAWS and NWS observations used in the control analyses are also shown. As expected, the control analyses do not exactly match the observations but do generally conform to them. The highest temperature, lowest relative humidity, and relatively strong wind speeds are observed in the foothills of the central Sierra Mountains roughly at elevations between 500-1500 m. The root mean square error (RMSE) between RAWS or NWS observations and the control analyses at this particular time averaged over the domain are on the order of 1^{o}C for temperature, 5% for relative humidity, and 1 m s^{-1} for wind speed or zonal and meridional wind components (Table 3). All of those values are of comparable magnitude to those typically assumed for instrumentation errors.

The CV scores obtained from the roughly 4900 cross validation experiments for temperature observations in this domain are shown in Fig. 5a. Values range from ~1^{o}C in the Sacramento Valley to several ^{o}C in some of the more remote locations of the domain as well as in the Diablo Range to the east of San Francisco Bay. The anomalously high temperature CV score at Emigrant Gap (station code KBLU and labeled in Fig. 3a) is due to very low temperatures observed at this site relative to the background as well as other nearby observations (e.g., Fig. 4a). However, the relative humidity and wind differences at this station are comparable to that at other nearby stations (not shown). The sensitivity of the entire analyses to removing single temperature observations is small as expected (Fig. 5b). However, the analyses exhibit higher sensitivity in the Diablo Range.

Figure 5c shows the degradation of the temperature analyses when observations are withheld. Most of the degradation values are greater than 45%, which suggests that the control analyses near those locations are very dependent on the withheld observations. Since the analysis degradation is a nondimensional parameter, it is possible to aggregate values obtained from cross validation experiments for different variables as shown in Fig. 5d. The analysis degradation values obtained separately for temperature, relative humidity, and wind speed tend to be very close to one another at each station, and, thus, differences between the values in Figs. 5c and 5d are small.

Table 4 summarizes the cross validation metrics for the 10 stations with the lowest median analysis degradation. The locations of these stations are indicated in Fig. 5d. Five of the ten stations are RAWS located in the Diablo Range in the lower left corner of the domain; three others have NWS stations nearby with two of those having NWS stations in very close proximity (MSAC1, Mt. Shasta, and RRAC1, Redding). The weighted proximity metric in Table 4 scales the horizontal and vertical distances between each RAWS station and its nearest neighbor using the Gaussian weights defined to specify the background error covariance (see Tyndall et al. 2010). The weighted proximity is defined as $e^{-(\frac{h^2}{40^2})} e^{-(\frac{z^2}{100^2})}$ where h is the horizontal separation (km) and z is the vertical separation (m). A weighted proximity of 1 indicates the RAWS station has another station available in the same location while a value of 0 implies a large separation in either the horizontal or vertical directions or both. (An alternative to the weighted proximity metric that we did not use due to its computational expense is to compute the IDI at each station after withholding that station's "observation" as discussed by Uboldi et al. 2008.)

It should not be surprising that 5 of the 10 stations with the lowest analysis degradation have weighted proximities greater than 0.90, which suggests that the impact of removing the station observation is mitigated by the availability of another station nearby. Of particular interest are three of those five stations (MOWC1, RRAC1, and GRSC1) that also have smaller CV scores when the observations are withheld at those locations (where smaller CV score is defined for each variable as follows: temperature < 2°C, relative humidity < 8%, and wind speed < 2.5 ms^{-1}). The combination of low median degradation and low CV score indicates that the control analyses tend to be closer to the observed values near those stations as well.

On the other hand, the 5 RAWS in the Diablo Range are not in immediate proximity to one another and have low median degradation but often have high CV scores when the observations are withheld. Hence, the control analyses deviate substantively from the observations near these locations. These are examples of locations where the observations likely describe localized weather phenomena and it would not be appropriate to consider those observations to be adequately reproduced by using background analyses in combination with other observations nearby.

To assess the dependence of our results on time of day, 120 control analyses and 16,800 cross-validation experiments were completed within the 24 hour period from 0000 to 2300 UTC 10 July 2008. The results obtained from these cross validation experiments are summarized in Table 5 for the ten stations with the lowest median analysis degradation within this domain. Eight of the ten stations with the lowest median analysis degradation are the same as those listed in Table 4. However, the CV scores tend to be larger when observations are withheld at all hours of the day, which reflects greater sensitivity of the analyses to the available observations.

The availability of stations within relatively close horizontal proximity to one another does not guarantee that their observations are redundant. For example, consider the Mt. Diablo (MDAC1) RAWS, labeled by the number 8 in Fig. 5d. Although MDAC1 is located in an area with many data assets compared to much of the rest of the domain, this station is located near the crest of the Diablo Range to the east of the San Francisco Bay region and its elevation is much higher than any of the other stations nearby (hence, the weighted proximity is 0.0). Very large relative humidity CV scores are evident for both sets of analyses (Tables 4 and 5). In the case of 10 July (Table 5), the relative humidity observed at MDAC1 was low during most the day since the top of the marine layer tended to remain below the station. Since the downscaled background relative humidity grids do not resolve the inland penetration of the marine layer well and the nearby observations are located at lower elevations within the marine layer, removing the MDAC1 observations strongly affects the quality of the analyses.

Additional control and cross-validation analyses for all variables for the 35 days were computed using other observation to background error variance ratios and background error decorrelation length scales, including those recommended by Tyndall et al. (2010), i.e., an increase in the ratio to 2:1 and an increase in the horizontal and vertical decorrelation length scales to 80 km and 200 m, respectively. The observation to background error variance ratio limits how large the analysis degradation can become (Uboldi et al. 2008). The upper limit is 50% for our parameter choices, i.e., in data sparse regions removing isolated observations results in a maximum difference between the observations and the values at those locations in the withheld analyses of double that of the differences between the observations and the values at those locations in the control analyses. Increasing the observation to background error variance ratio to 2:1 lowers the upper limit value to 33%, since the reduced confidence given to the observations limits the impact that they can have upon the analyses. Lengthening the background error decorrelation length scales tends to reduce by a small amount the CV scores for the stations of interest to this study (e.g., those listed in Table 4), since the impact of removing those stations is controlled by the presence of other stations nearby. The CV scores for stations that are more isolated from others are affected more by lengthening the background error decorrelation length scales, but not enough to recommend that the local observations are unimportant.

b. National Statistics

We now present results obtained from the 570,000 cross validation experiments for the 51 domains within the continental United States. Additional results and graphics are available at http://mesowest.utah.edu/raws, including the background and control analyses for each case for all 51 domains. As an indicator of the distribution of observations combined with the assumptions regarding observation and background error covariance, Fig. 6 shows the IDI computed in each of the 51 domains. Regions where observations are more abundant appear in the darker shades while data voids appear white. It should not be surprising that the combination of the existing NWS and RAWS networks is not sufficient to provide adequate coverage of surface conditions in all regions of the continental United States (National Research Council 2009).

Table 6 summarizes a number of metrics computed from all stations, cases, and domains within the continental United States. First, the degree to which the control analyses agree with the observations is listed separately for RAWS and NWS stations. As shown for the single case and domain in Table 3, the control analyses are constrained on average to be close to the observations. The median CV scores for all three variables based on the 1688 RAWS are slightly higher than those for the 1589 NWS stations, which suggests that removing the RAWS observations is more detrimental to the analyses. The higher sensitivities to removing NWS observations of relative humidity and wind depends to some extent on the siting of NWS stations in relatively uniform terrain that allows for their wider spatial influence compared to the limited influence of many RAWS in mountainous regions. Overall, the lower median analysis degradation and CV scores for all three variables for all NWS stations suggests that removing the

NWS observations has a reduced impact on the analyses compared to the impact of removing the RAWS.

As discussed in the previous subsection, the weighted proximity values summarized in Fig. 7 help to define pairs of stations that are located relatively close to one another. Values greater than 0.367 (e^{-1}) reflect station pairs that are close to one another in both the horizontal and vertical directions. A larger fraction of pairs of stations are in close proximity to one in southern and southeastern United States and Great Lakes domains while larger numbers of close pairs of stations are also found in California and Oregon.

The impact of removing observations as evaluated using the CV score depends on the variable and differs regionally as evident in Figs. 8a-c. Smaller temperature differences between the observations and the analyses from which those observations are withheld are more common in coastal and plains regions than mountainous ones. Removing relative humidity observations tends to have less impact in arid regions of the west during summer compared to other regions of the country (Fig. 8b). Removing wind observations tends to have the greatest effect in regions of complex terrain (Fig. 8c). For example, the wind speed CV score is greater than the 2.5 m s^{-1} threshold for all stations in central Nevada.

The RAWS with relatively low values of the median analysis degradation for temperature, relative humidity, and wind speed are summarized in Fig. 8d. The largest total number of stations with low analysis degradation is found in southern California with high fractions of the stations in the eastern coastal domains also having low analysis degradation. All of the RAWS in Wyoming have values of median analysis degradation greater than the 40% threshold, which suggests the analyses are substantially degraded when those stations are removed.

Finally, Fig. 9 summarizes the relatively small number of RAWS (110) in the continental United States that exhibit low CV scores when the observations of temperature, relative humidity, and wind are removed combined with low median analysis degradation values. As discussed earlier, these stations are ones where the control analyses tend to be close to the observations and yet removal of those observations does not diminish the analysis quality as much compared to the removal of observations at other stations. Five stations in the northern California domain satisfy all of these criteria (including MOWC1, RRAC1, and GRSC1 as evident in Table 4). As shown in Fig. 9, the nearest neighboring stations to these RAWS are for the most part NWS stations. For example, 24 of the 26 RAWS in Minnesota, Wisconsin, and upper Michigan that meet these criteria have a NWS station nearby. Not surprisingly, domains that contain considerable mountainous terrain tend to have few stations that meet all of the criteria (e.g., Montana), even if those domains have a number of pairs of stations within relatively close proximity to one another (e.g., western Oregon).

4. Discussion

Identifying the optimal locations for stations for wildland fire applications has been of interest for many years (Fujioka 1986). More generally, there are many needs for improved selection of surface observing sites for a variety of weather and climate applications (Husain et al. 1986; Vose and Menne 2004; Paimazumder and Molders 2009; National Research Council 2009). Siting new stations as well as maintaining existing ones requires consideration of a broad range of factors including access, security, obstructions, and the representativeness of the site compared to the surrounding area.

This study is directed towards assessing the relative impact of removing existing RAWS on the quality of variational analyses of temperature, relative humidity, and wind. These analyses depend on the interplay between station spacing, terrain, the values of the observations and the background fields (and how those deviate from the unknown truth), and assumptions about the observation and background errors. The use of NWS observations to supplement those available from the RAWS network is considered, even though there are differences in the siting, observing, and reporting characteristics between the two networks.

The details within the control analyses and the analyses from which observations are withheld depend first on the quality of the observations. We removed a relatively small number of suspect observations as defined in Table 2. Some observations included in the analyses may differ substantially from the unknown truth on the scale of the analysis grid, due to instrumentation errors as well as the observations being representative of localized conditions compared to the surrounding regions.

The analyses also depend on our assumptions regarding the observation to background error variance ratio and the spatial dependence of the background error covariance on the underlying terrain. The national coverage due to the combined RAWS and NWS networks as estimated by the IDI field in Fig. 6 could be assumed to be improved by specifying that the background error covariance decreases less rapidly with distance both horizontally and vertically (Uboldi et al. 2008; Tyndall et al. 2010). In particular, using a larger vertical length scale would extend the influence of discrepancies between the observations and the background fields farther up- and down- slopes in mountainous regions. However, for the goals of this study, using broader decorrelation length scales would tend to lessen the impact of withholding a particular station. In addition, complex local temperature, moisture, and wind patterns in many regions such as the Diablo Range discussed in Section 3a are not necessarily handled better by assuming that the background errors at higher elevations are strongly related to the errors at lower elevations in the surrounding San Francisco Bay or Sacramento/San Joaquin Valley regions. Further, greater fidelity of the control analyses to the observations can be forced by assuming that the observational errors are small, but that can lead to overfitting problems in data void areas (Daley 1991; Tyndall et al. 2010).

A large number of cross validation experiments were evaluated using several objective metrics. The CV score computed using the observations withheld from the analyses is combined with a nondimensional estimate of the analysis degradation arising from withholding the observations. While the root mean square error of the control analyses could have been shown explicitly as a further diagnostic (beyond the values shown in Tables 3 and 6), it can be deduced readily from this combination of metrics. Thus, the removal of a station is considered to have reduced impact compared to others if the cross validation scores are small combined with small analysis degradations. That combination implies that the control analyses remain close to the observations as well. On the order of 100 RAWS within the continental United States satisfy relatively conservative thresholds for temperature, relative humidity, and wind RMSE and low analysis degradation. Meeting these criteria for nearly all of those stations depends to a large extent on the availability of observations from nearby NWS stations.

As mentioned in the Introduction, this project is part of a broader effort to evaluate the distribution of RAWS in the United States in order to identify both areas where present observations may be less critical than others as well as data gaps. Our objective metrics of station redundancy must be placed in the context of many other factors, including basic issues such as the length of record available from one station compared to another. Simply assuming that a nearby NWS observation can be substituted for that of a RAWS observation does not take into consideration the additional sensors for soil and fuel moisture and temperature usually available at RAWS sites that are not available at NWS ones. The ability to compute National Fire Danger Rating System indices in a systematic and consistent way must also be considered (National Wildfire Coordinating Group 2002; Hoadley et al. 2004). In addition, we have not addressed the needs for sufficient precipitation observations to monitor the often localized nature of summer season precipitation. Hence, it must be simply stated that nearly all fire prone regions of the country remain critically undersampled. Figures 3 and 6 provide a gross indication of some of the largest data voids.

Consistent with Myrick and Horel's (2008) results for winter temperature over the western United States only, the removal of RAWS on average has a larger detrimental impact on analyses than removal of NWS stations (Table 6). Hence, the "value" of RAWS as examined here from the standpoint of analysis impact is higher than the value of NWS stations. This results from the reduced ability of the background fields to be representative of the conditions in the often remote locations near RAWS where few other observational assets are likely to be located nearby. As part of the planning process for integrating the many disparate mesonet observations around the country into a national network of networks (National Resources Council 2009), it will be important to evaluate further the relative impacts of the many different data resources available.

Acknowledgments. We would like to thank Dan Tyndall for his many contributions to this project. This research was supported with funding provided by the National Park Service administered through contract #AG7604C080033 between the USDA Forest Service and the

University of Utah as well as by grant NA07NWS4680003 provided by the NOAA/NWS CSTAR program.

References

Benjamin, S. G., G. A. Grell, J. M. Brown, T. G. Smirnova, and R. Bleck, 2004: Mesoscale weather prediction with the RUC hybrid isentropic–terrain-following coordinate model . *Mon. Wea. Rev*., **132**, 473-494.

Daley, R., 1991: *Atmospheric Data Analysis*. Cambridge University Press. 456 pp.

Fujioka, F., 1986: A method for designing a fire weather network. *J. Atmos. Ocean. Tech*., 3, 564-570.

Hoadley, J., K. Westrick, S. Ferguson, S. Goodrick, L. Bradshaw, P. Werth, 2004: The effect of model resolution in predicting meteorological parameters used in fire danger rating *.J. Appl . Met*., 43, 1333-1347.

Horel, J., M. Splitt, L. Dunn, J. Pechmann, B. White, C. Ciliberti, S. Lazarus, J. Slemmer, D. Zaff, J. Burks, 2002: MesoWest: Cooperative mesonets in the western United States. *Bull. Amer. Meteor. Soc*., **83**, 211-226.

Huang, C., Y. Lin, M. Kaplan, J. Charney, 2009: Synoptic-scale and mesoscale environments conducive to forest fires during the 200 extreme fire event in southern California. *J. Appl. Clim. Meteor*., **48**, 553-579.

Husain, T., M. Ukayli, H. Khan, 1986: Meteorological network expansion using information decay concept. *J. Atmos. Ocean. Tech.,* 3, 27-35.

Jascourt, S. cited 2010: Real-time mesoscale analysis: What is the NCEP RTMA and how can it be used? MetEd/COMET. [available online at https://www.met.ucar.edu/loginForm.php?urlPath=nwp/RTMA.]

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

National Wildfire Coordinating Group, 2002: Gaining a basic understanding of the National Fire Danger Rating System—A self-study reading course. National Wildfire Coordinating Group, 73 pp. [Available online at *http://www.nationalfiretraining.net/ca/nctc/prework/nfdrs_pre_study.pdf.*]

National Wildfire Coordinating Group, 2008: NWCG Fire Weather Station Standards. PMS 426-3. 49 pp. [Available online at *http://www.fs.fed.us/raws/standards/FireWxStds_final_revMay08.pdf.]*

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

Paimazumder, D., and N. Molders, 2009: Theoretical assessment of uncertainty in regional averages due to network density and design. *J. Appl. Meteor. Clim*, **48**, 1643-1666.

Saad, Y., and M. H. Schultz, 1986: GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Stat. Comput.,* 7, 856-869.

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

Tyndall, D., J. Horel, and M. de Pondeca, 2010: Sensitivity of surface air temperature analyses to background and observation errors. *Wea. Forecasting*. In press.

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

Vose, R., and M. Menne, 2004: A method to determine station density requirements for climate observing networks. *J. Clim*., **17**, 2961- 2971.

Willks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences*. 3rd Edition. Elsevier, Inc. 627 pp.

Zachariassen, J., K. F. Zeller, N. Nikolov, and T. McClelland, 2003: A review of the Forest Service Remote Automated Weather Station (RAWS) network. Gen. Tech. Rep. RMRSGTR-119, Rocky Mountain Research Station, U.S. Forest Service, Fort Collins, CO, 153 pp. [Available online at http://www.fs.fed.us/rm/pubs/rmrs_gtr119.pdf.]

Zapotocny, T. H., and Coauthors, 2000: A case study of the sensitivity of the Eta data assimilation system. *Wea. Forecasting*, **15**, 603-621.

Figure Captions

1. The 51 analysis domains. The numbers to the left (right) represent the number of RAWS (NWS) stations located within each domain.
2. Median distance (km) between each RAWS station and its nearest neighboring RAWS (left) and RAWS or NWS (right) station.
3. a.) Terrain (shaded and contoured at intervals of 50 m, 500 m, 1500 m, and 2500 m) in the northern California domain. Locations of RAWS (circles) and NWS (squares) observations are shown as well as selected landmarks. b.) As in (a) except for the Integral Data Influence (IDI) field (shaded) computed by specifying the background everywhere as zero and RAWS (circles) and NWS (squares) observations set to one.
4. a.) The control temperature analysis (shaded; ^{o}C) valid at 1800 UTC 10 July 2008. Shaded circles (squares) indicate the temperature observations from RAWS (NWS) stations using the same scale. Approximate location of the Paradise, CA fire denoted by the star. b.) As in (a) except for the control relative humidity analysis (shaded; %). c.) As (a) except for the control wind speed (shaded; m s^{-1}) and vector wind analysis (red vectors). Green (blue) vectors indicate the wind observations from RAWS (NWS) stations.
5. a.) Cross validation (CV) scores of temperature (^{o}C) based on cross validation experiments removing RAWS (NWS) observations denoted by circles (squares). b.) As in (a) except for sensitivity of temperature (^{o}C). c.) As in (a) except for analysis degradation of temperature (%). d.) As in (a) except for the median of the analysis degradation (%)

for temperature, relative humidity, and wind speed. The ten RAWS with lowest median analysis degradation are numbered.

6. As in Fig. 3b except for all 51 domains.
7. Number of RAWS within each domain with weighted proximity to its nearest neighbor > 0.367 (left number) and total number of RAWS (right number).
8. a.) Number of RAWS within each domain with temperature CV scores < 2^{o}C (left number) and total number of RAWS (right number). b.) As in (a) except for relative humidity CV scores < 8%. c.) As in (a) except for wind speed CV scores < 2.5 m s^{-1}. d.) As in (a) except for median analysis degradation < 40%.
9. Number of RAWS that meet all four criteria shown in Fig. 8. The numbers to the left (right) indicate the number of stations with a RAWS (NWS station) nearby.

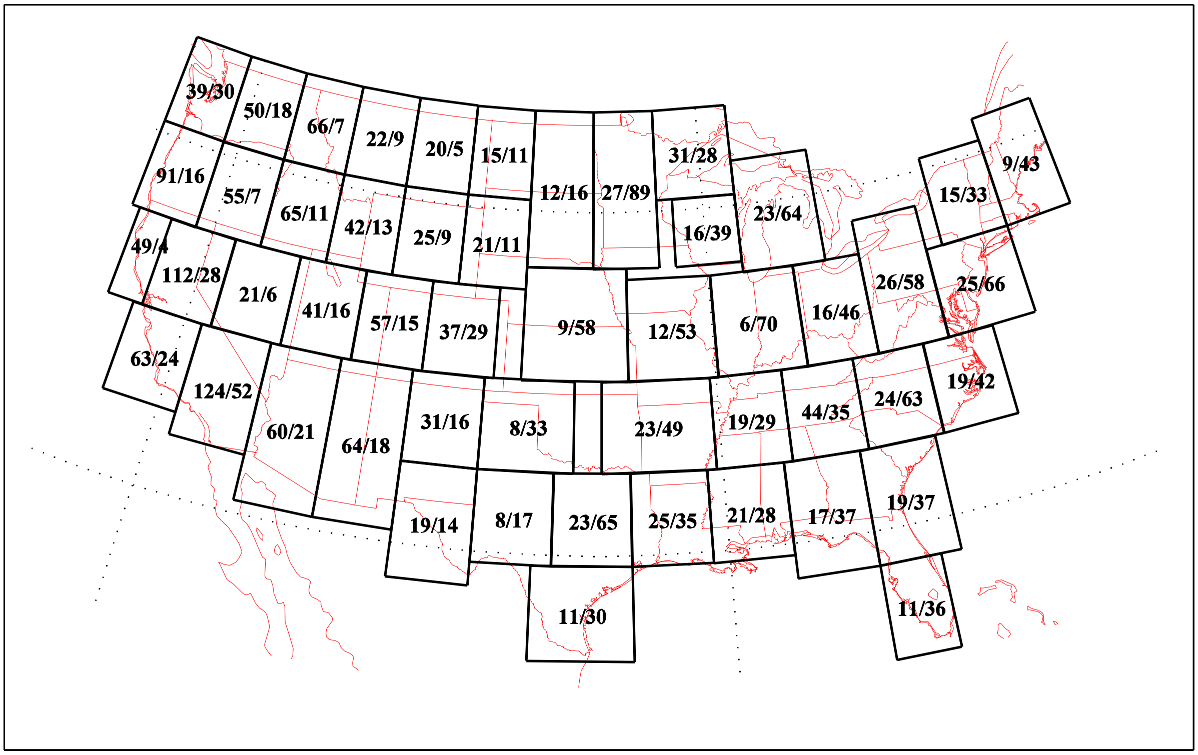

Figure 1. The 51 analysis domains. The numbers to the left (right) represent the number of RAWS (NWS) stations located within each domain.

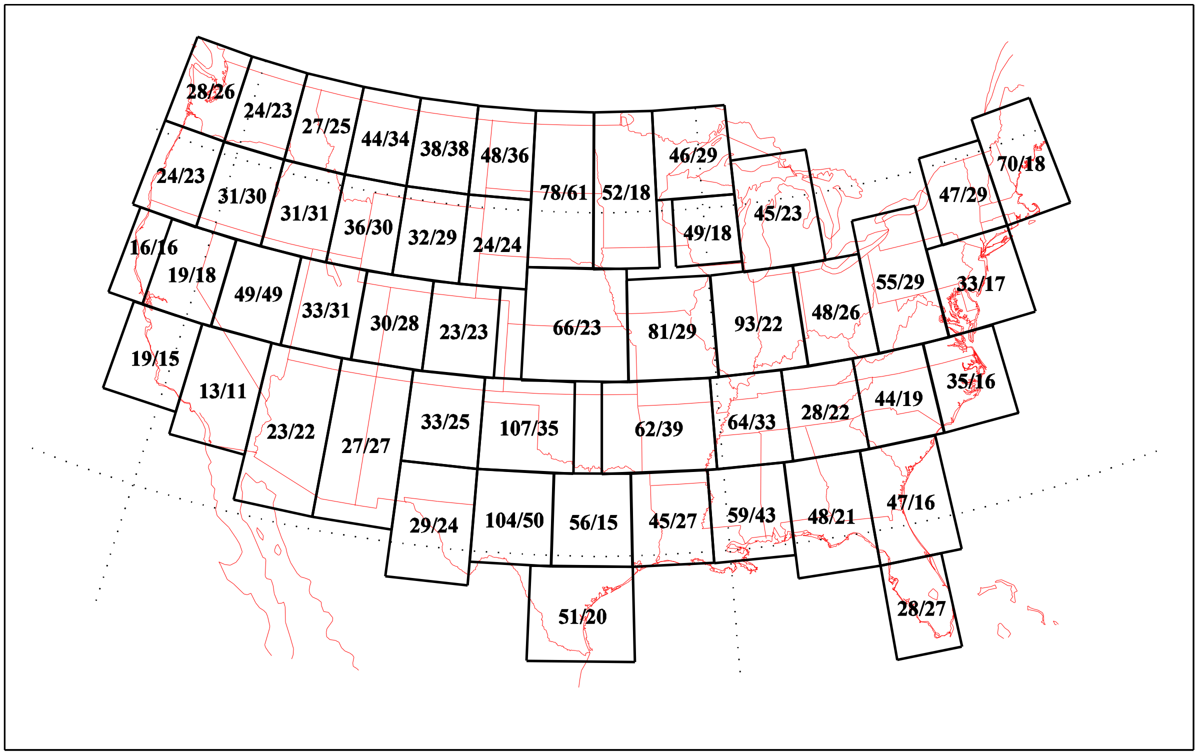

Figure 2. Median distance (km) between each RAWS station and its nearest neighboring RAWS (left) and RAWS or NWS (right) station.

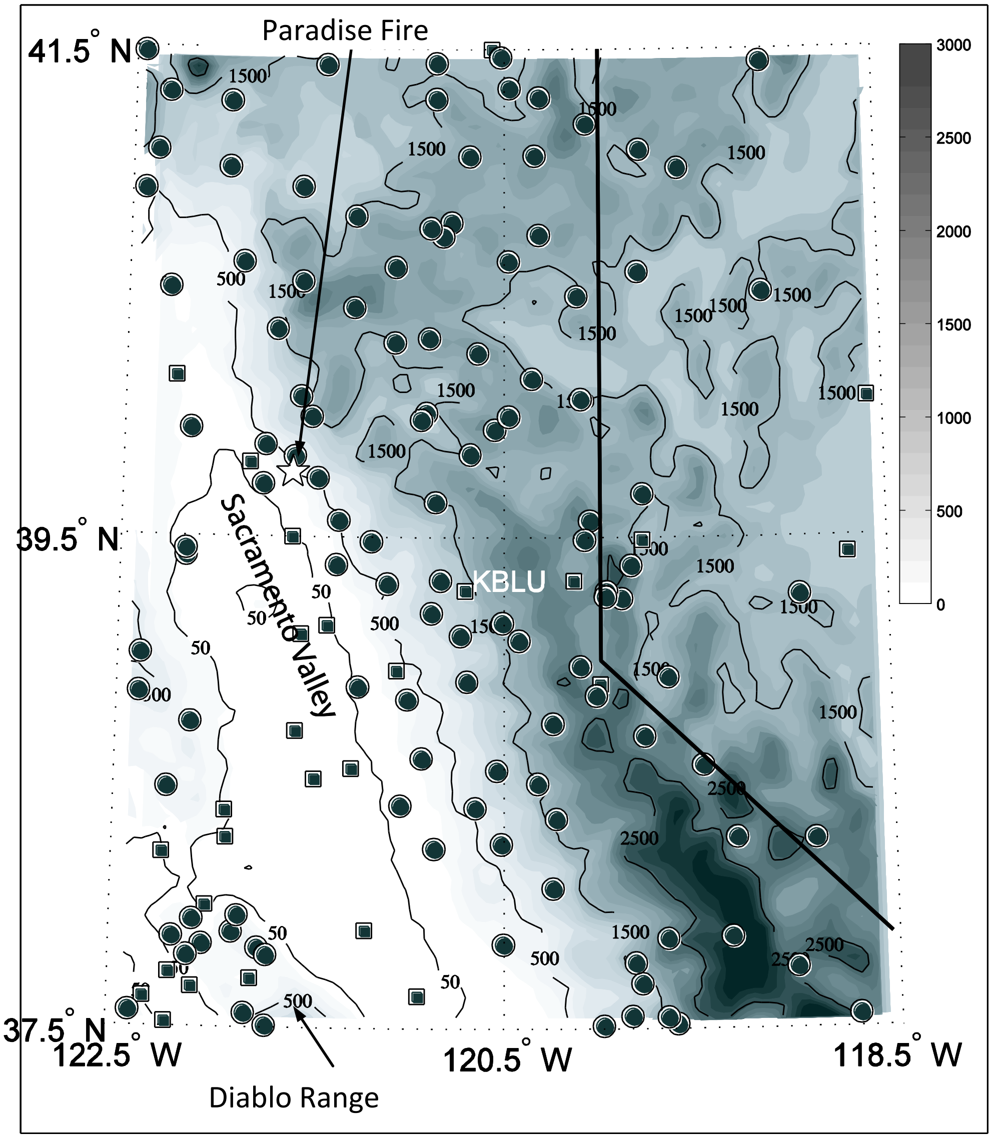

Figure 3a. Terrain (shaded and contoured at intervals of 50 m, 500 m, 1500 m, and 2500 m) in the northern California domain. Locations of RAWS (circles) and NWS (squares) observations are shown as well as selected landmarks.

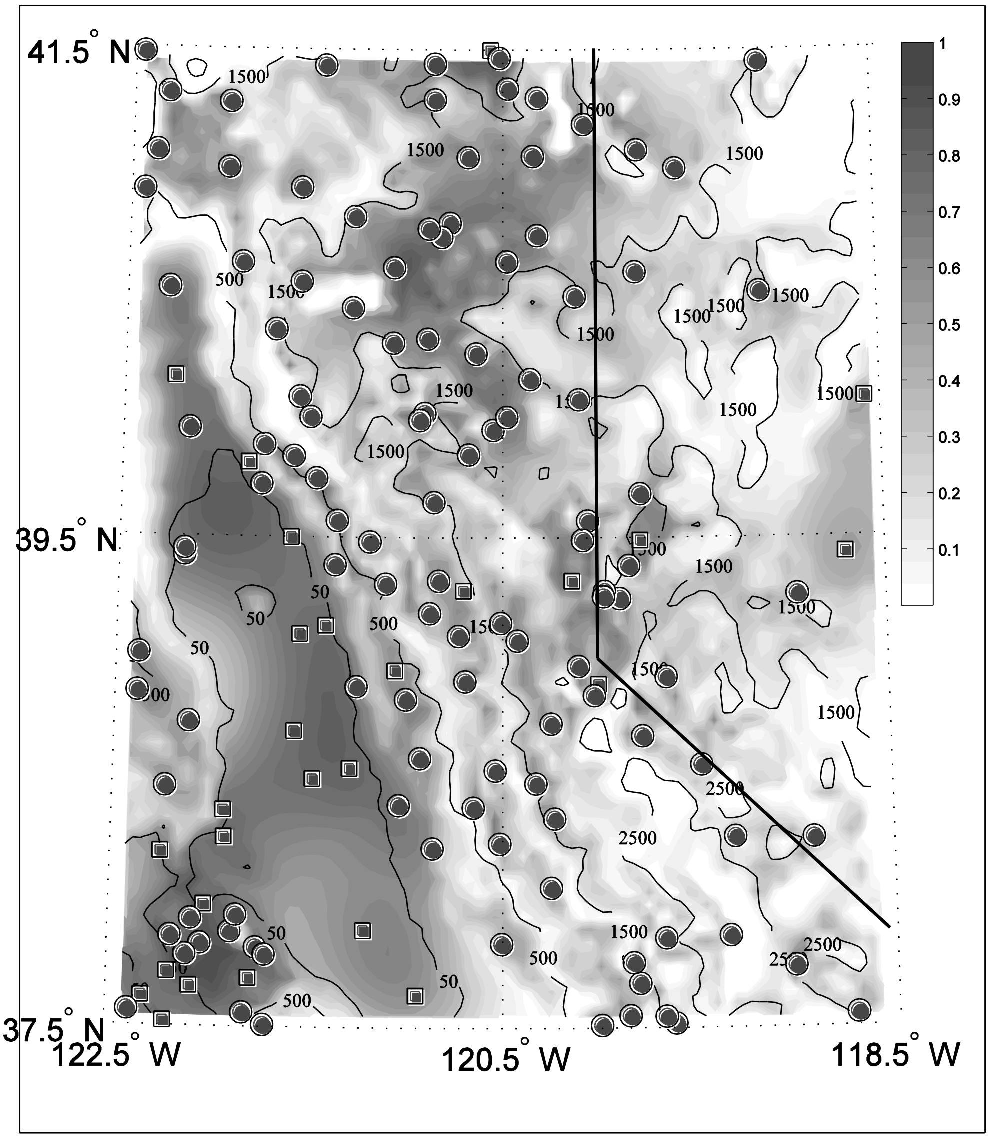

Figure 3b. As in Fig. 3a except for the Integral Data Influence (IDI) field (shaded) computed by specifying the background everywhere as zero and RAWS (circles) and NWS (squares) observations set to one.

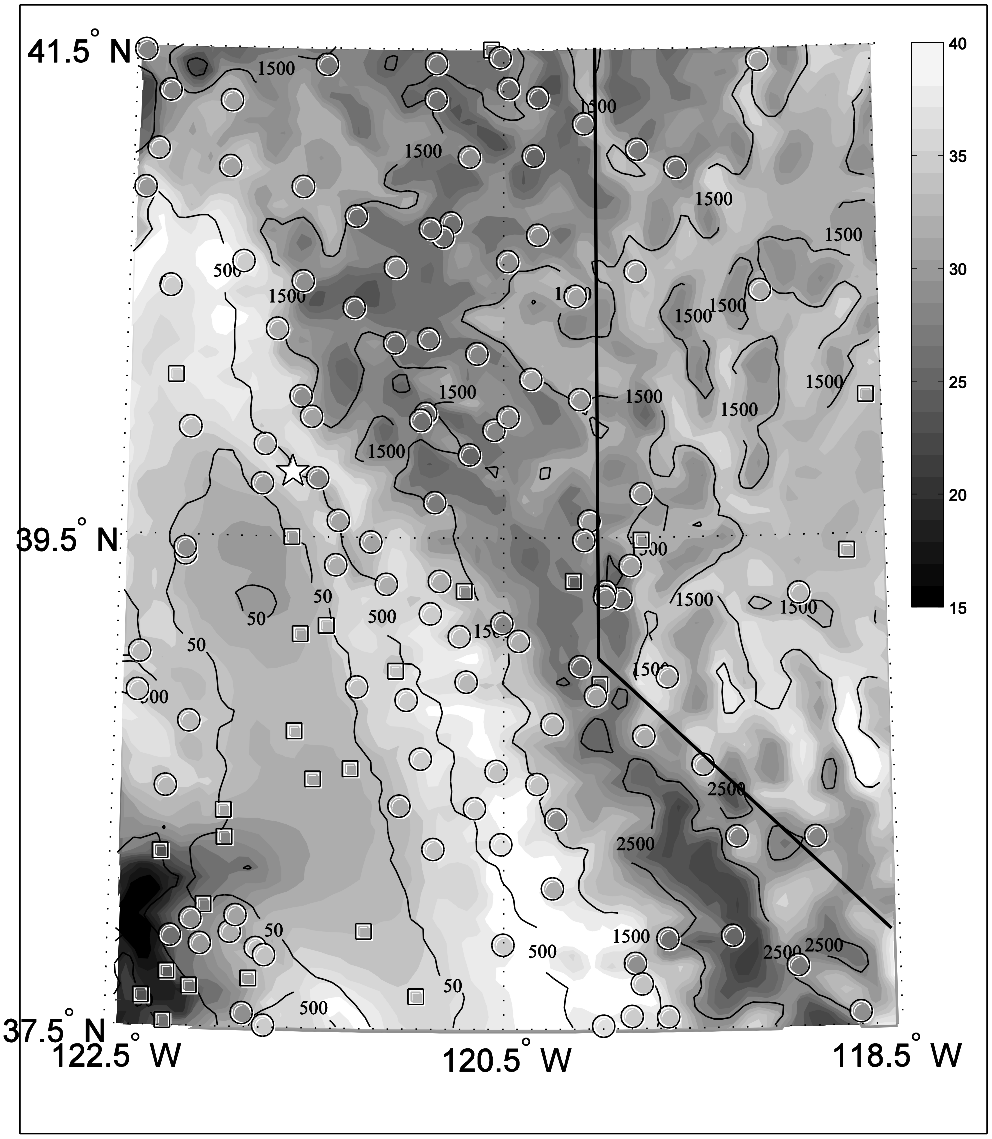

Figure 4a. The control temperature analysis (shaded; °C) valid at 1800 UTC 10 July 2008. Shaded circles (squares) indicate the temperature observations from RAWS (NWS) stations using the same scale. Approximate location of the Paradise, CA fire denoted by the star.

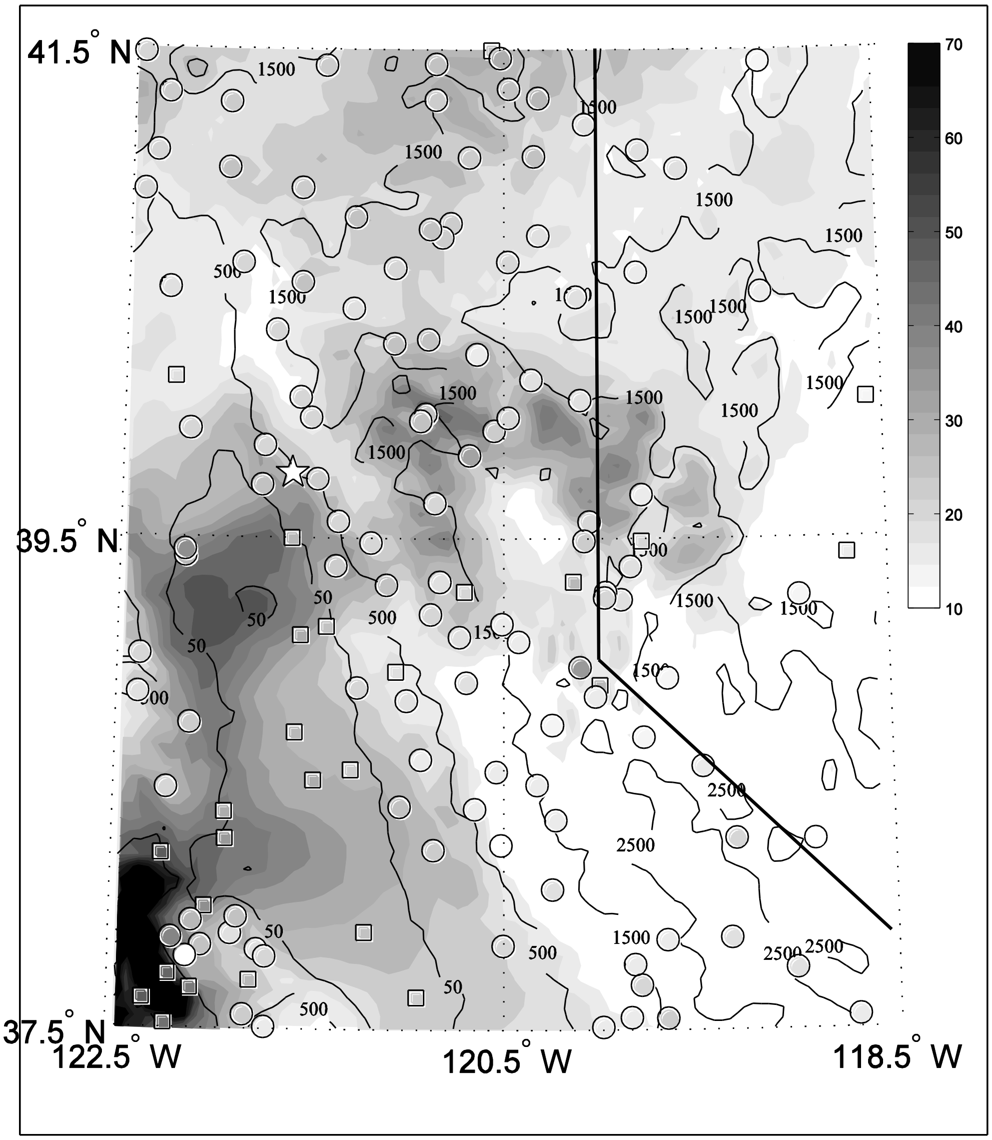

Figure 4b. As in Fig. 4a except for the control relative humidity analysis (shaded; %).

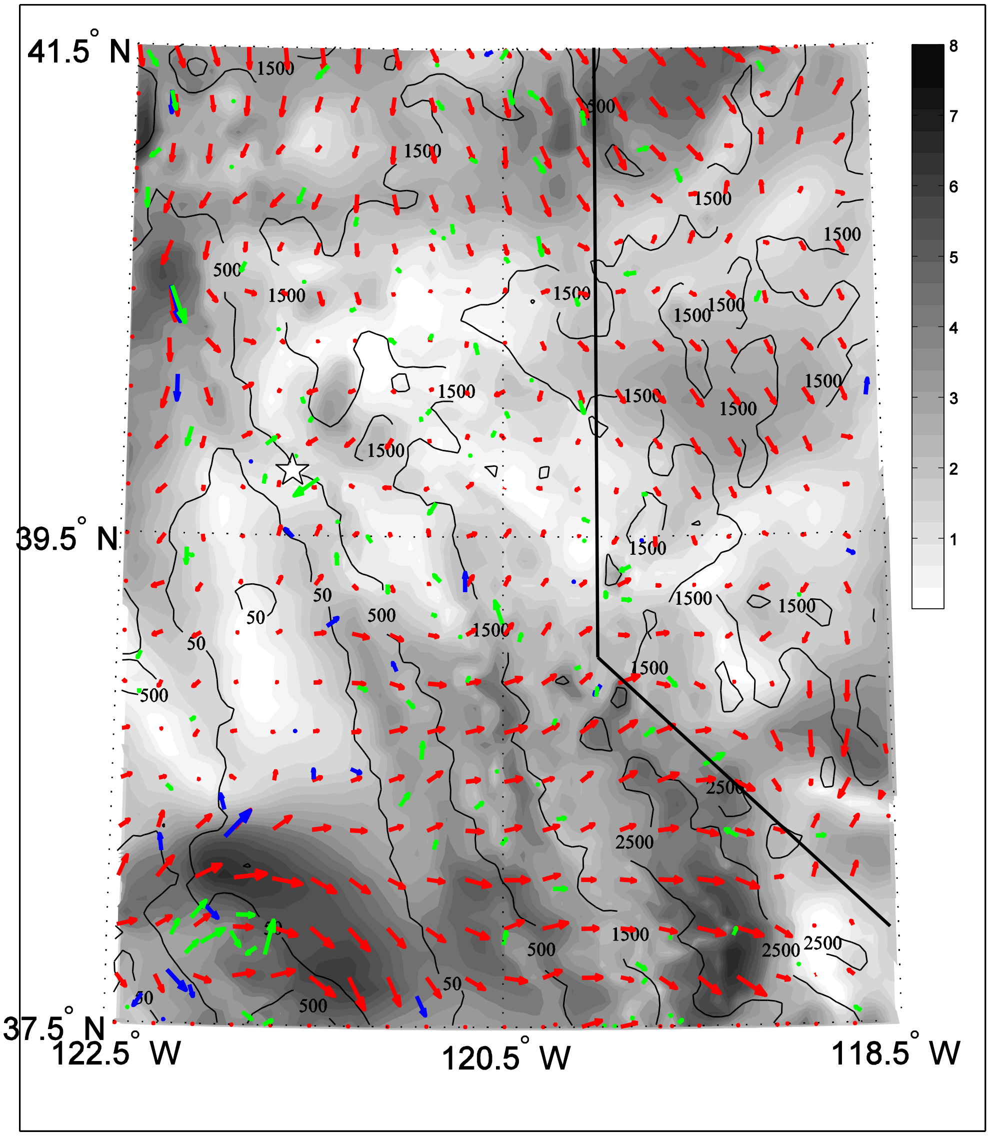

Figure 4c. As in Fig. 4a except for the control wind speed (shaded; m s^{-1}) and vector wind analysis (red vectors). Green (blue) vectors indicate the wind observations from RAWS (NWS) stations.

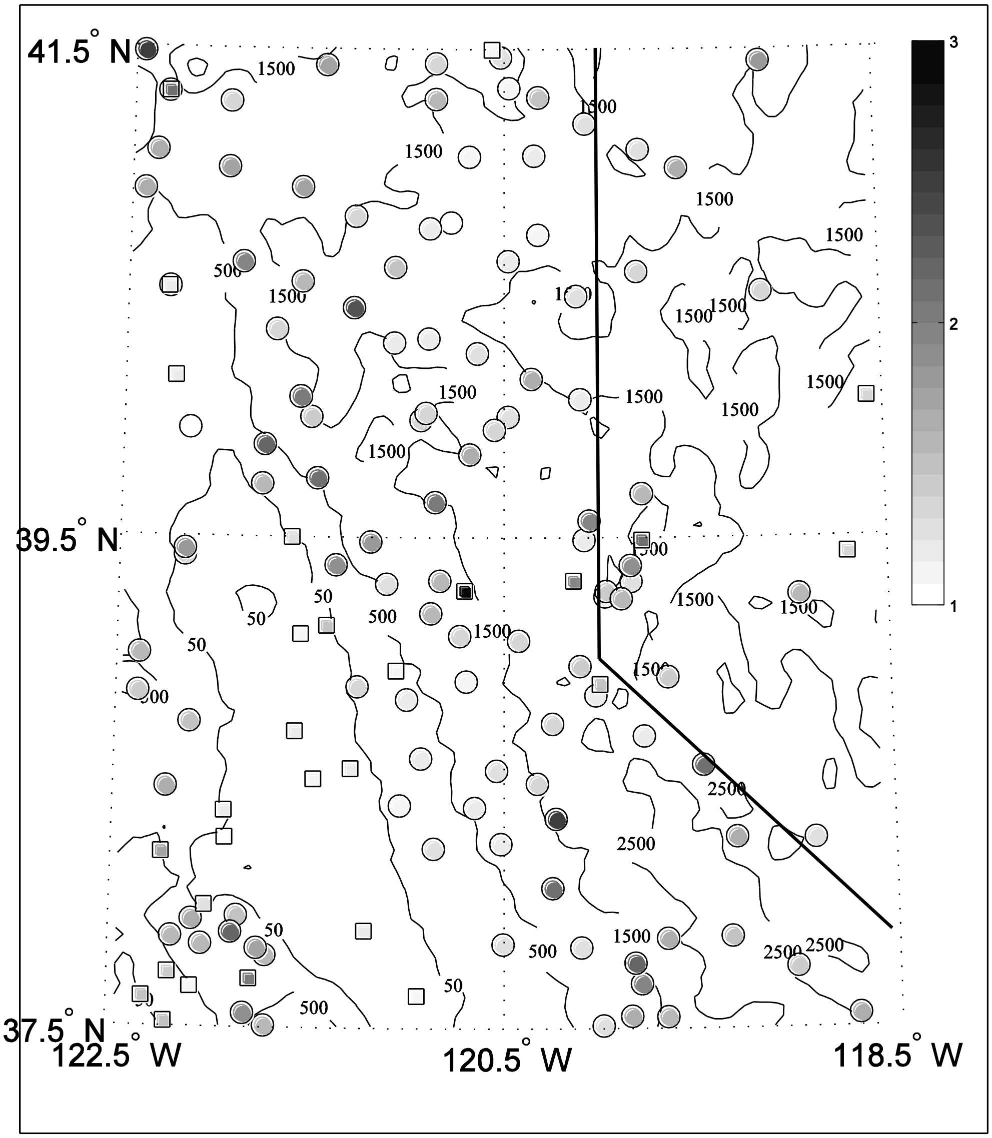

Figure 5a. Cross validation (CV) scores of temperature (^oC) based on cross validation experiments removing RAWS (NWS) observations denoted by circles (squares).

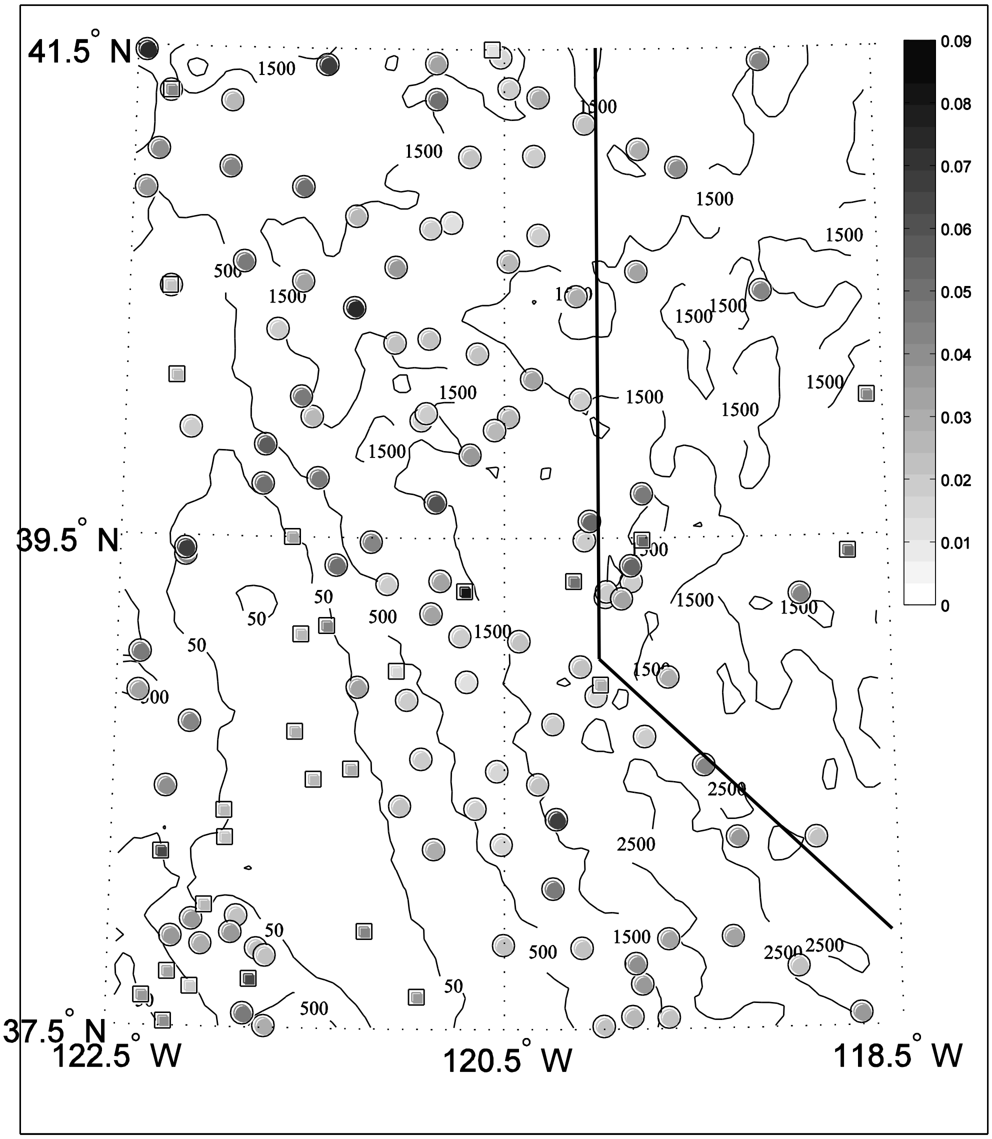

Figure 5b. As in Fig. 5a except for sensitivity of temperature (^{o}C).

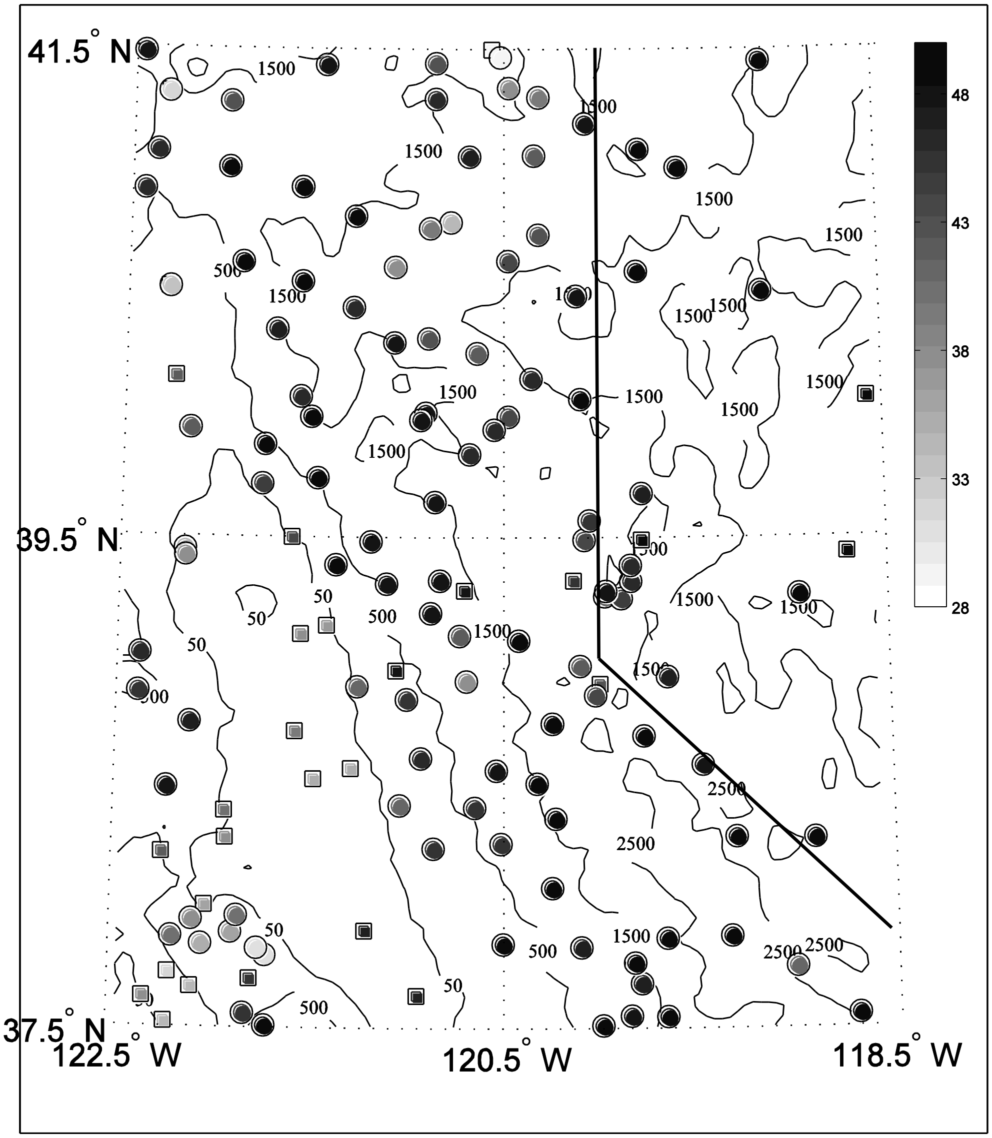

Figure 5c. As in Fig. 5a except for analysis degradation of temperature (%).

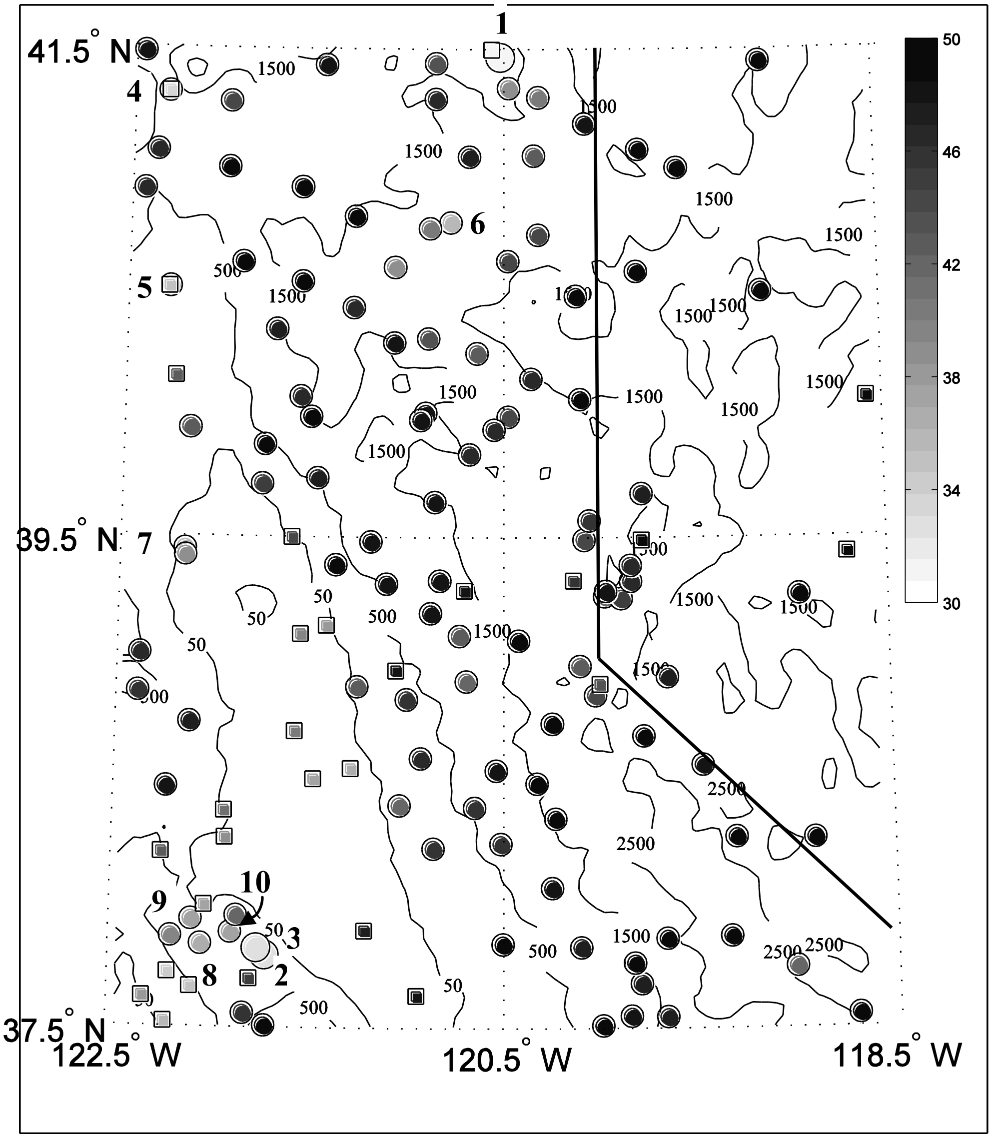

Figure 5d. As in Fig. 5a except for the median of the analysis degradation (%) for temperature, relative humidity, and wind speed. The ten RAWS with lowest median analysis degradation are numbered.

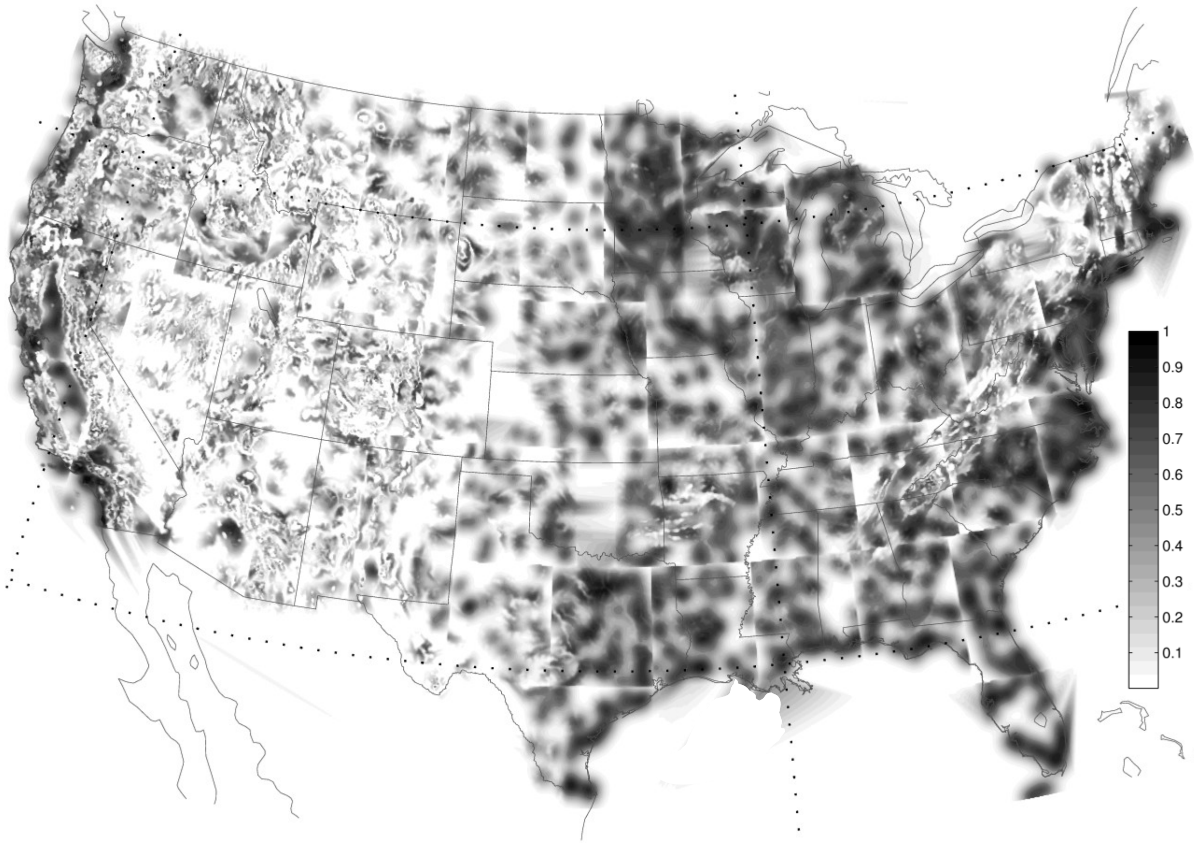

Figure 6. As in Fig. 3b except for all 51 domains.

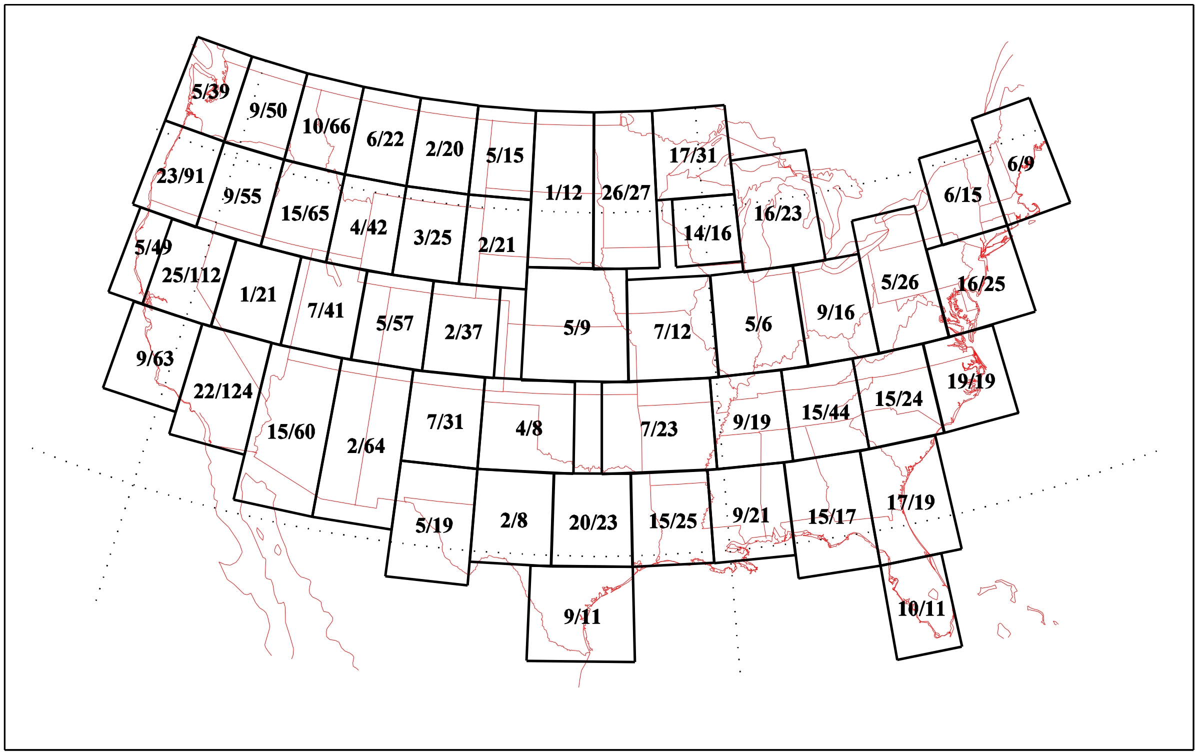

Figure 7. Number of RAWS within each domain with weighted proximity to its nearest neighbor > 0.367 (left number) and total number of RAWS (right number).

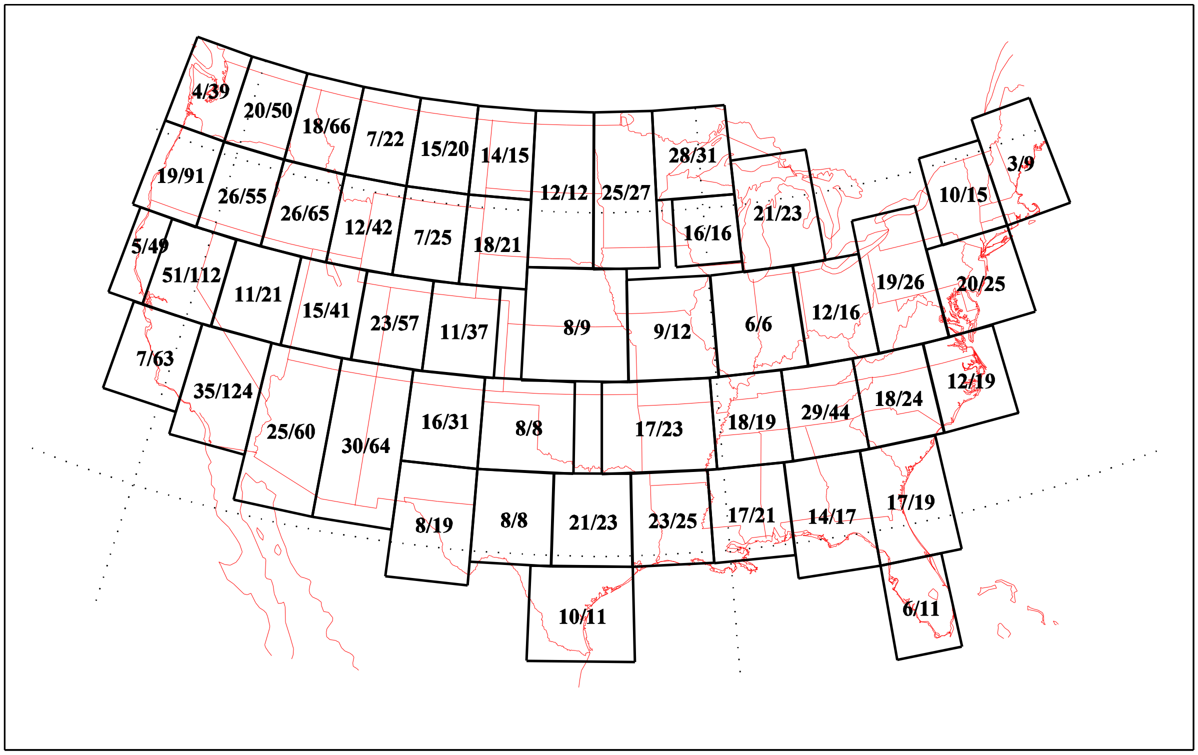

Figure 8a. Number of RAWS within each domain with temperature CV scores < 2°C (left number) and total number of RAWS (right number).

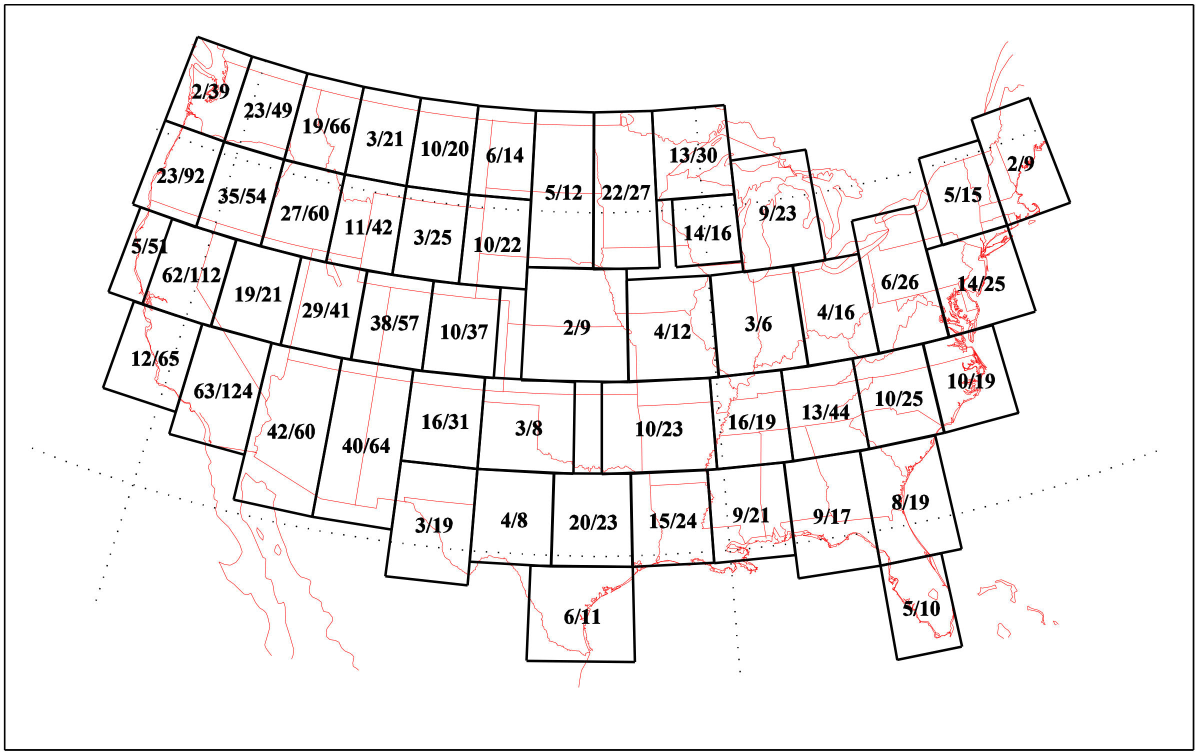

Figure 8b. As in Fig. 8a except for relative humidity CV scores < 8%.

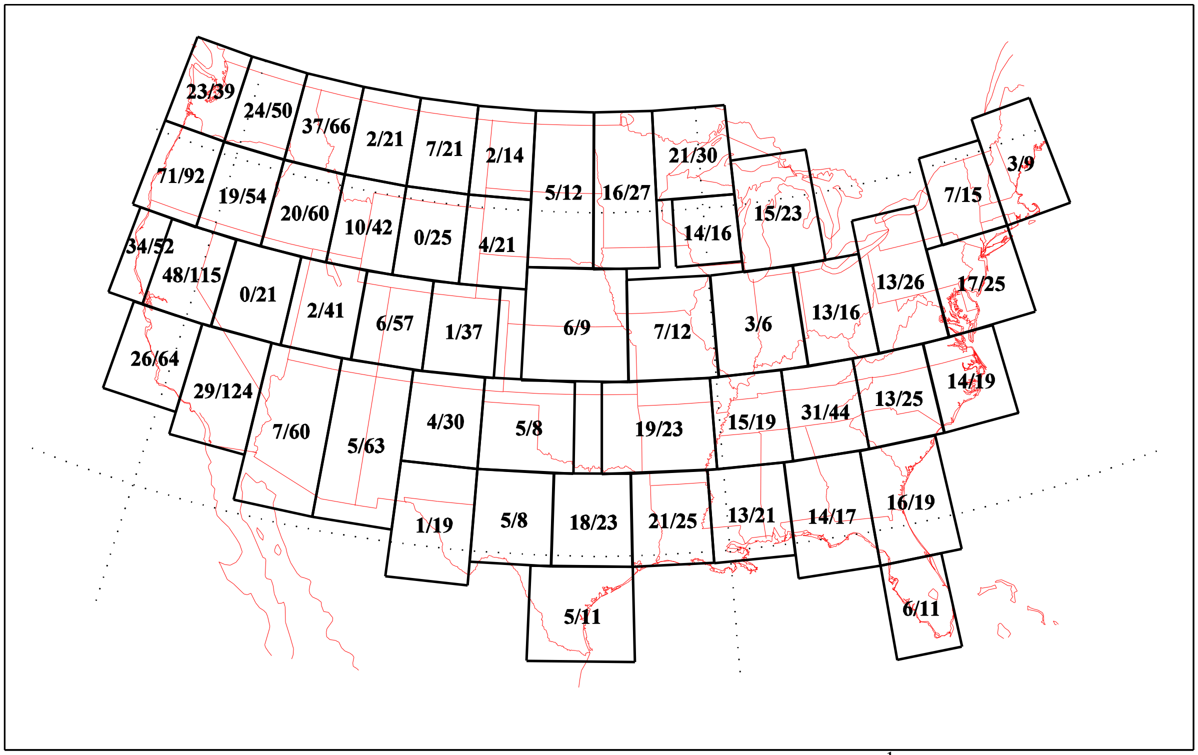

Figure 8c. As in Fig. 8a except for wind speed CV scores < 2.5 m s^{-1}.

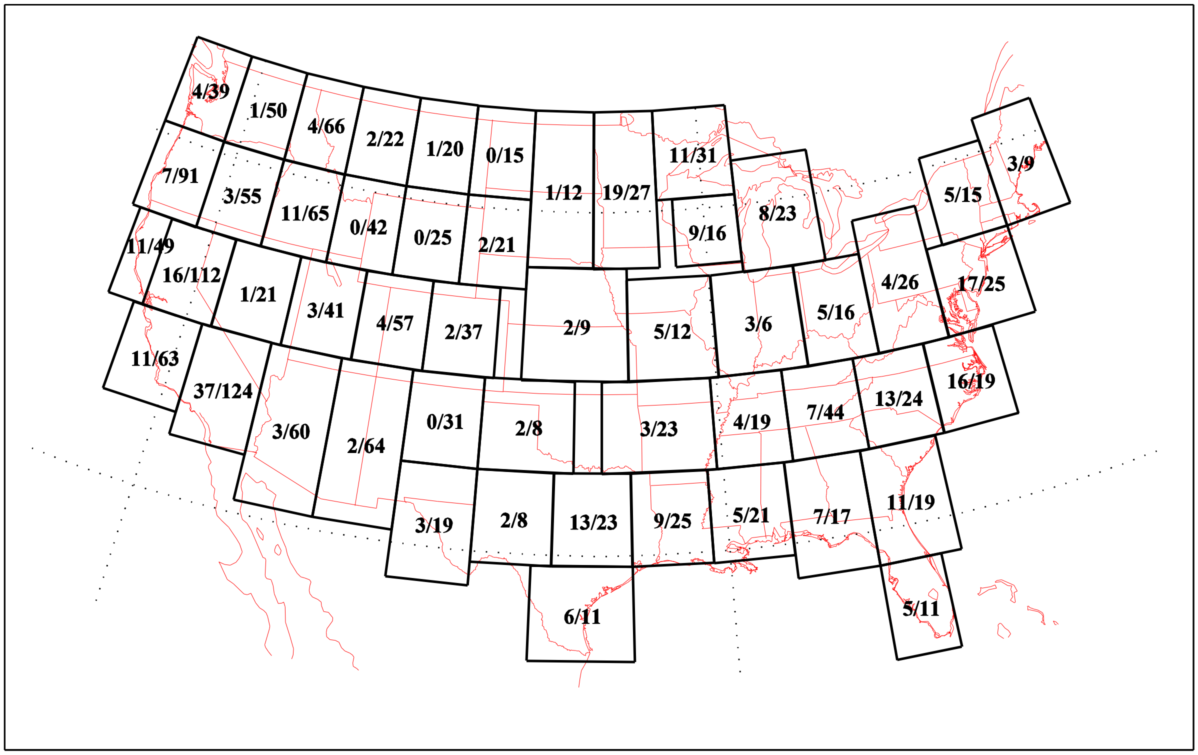

Figure 8d. As in Fig. 8a except for median analysis degradation < 40%.

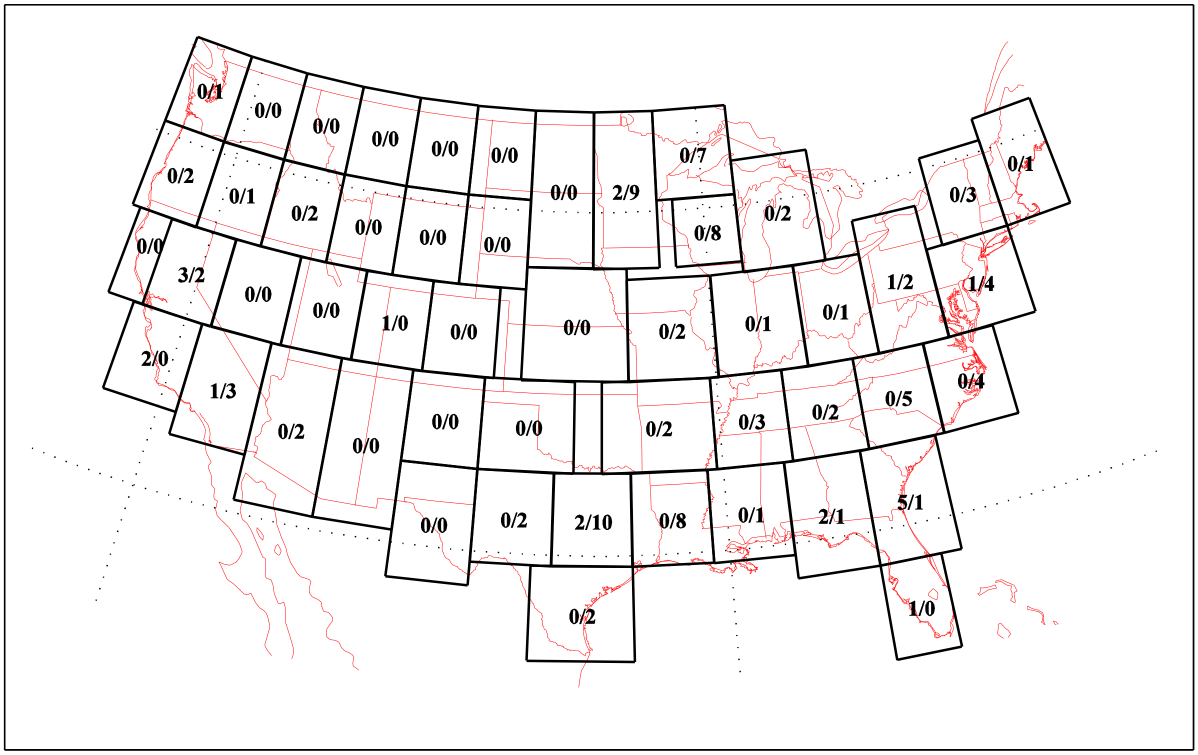

Figure 9. Number of RAWS that meet all four criteria shown in Fig. 8. The numbers to the left (right) indicate the number of stations with a RAWS (NWS station) nearby.

Table 1

Differences between NWS and RAWS observations

Characteristic	**NWS**	**RAWS**
Location	Often adjacent to airport runways	Preferred to be located in open areas on slight south facing aspects in mountainous/forested areas
Wind sensor height	10 m	6 m
Wind speed averaging interval	2 minute	10 minute
Temperature Aspiration	Yes	No
Routine reporting time	5-10 minutes before the hour	Hourly but scheduled to balance satellite transmissions throughout the hour

Table 2

Quality control thresholds set for observations

Threshold	**Temperature**	**Relative Humidity**	**Wind**
Relative to Background	$\left\|y_i - x_b^i\right\| > 10\ ^\circ C$	$\left\|y_i - x_b^i\right\| > 50\%$	$\left\|y_i - x_b^i\right\| > 15\ ms^{-1}$
Relative to Analysis	$\left\|y_i - x_c^i\right\| > 6\ ^\circ C$	$\left\|y_i - x_c^i\right\| > 40\%$	$\left\|y_i - x_c^i\right\| > 10\ ms^{-1}$
Percent Rejected	0.66%	0.31%	0.02%

Table 3

RMSE of control analyses valid at 1800 UTC 10 July 2008 for the northern California domain relative to RAWS and NWS observations

Network	**Temperature (^{o}C)**	**Relative Humidity (%)**	**Wind Speed (m s^{-1})**	**Zonal Wind (m s^{-1})**	**Meridional Wind (m s^{-1})**
RAWS	1.4	5.6	1.0	1.5	1.1
NWS	1.7	5.3	0.9	1.6	1.0

Table 4

Summary of cross validation measures for the stations within the northern California domain with the lowest median degradation

No.	RAWS Station	Nearest Neighboring Station (NWS in italics)	Horizontal Separation (km)	Vertical Separation (m)	Weighted proximity	Temper-ature CV score (°C)	Temper-ature Sensitivity (°C)	Temperature Degradation (%)	Relative Humidity CV score (%)	Wind CV score (ms^{-1})	Median Degradation (%)
1	MOWC1	*KAAT*	5	2	0.98	1.5	0.01	28	7.8	2.1	29
2	VAQC1	LVMC1	5	-287	0.00	2.6	0.02	30	7.4	5.1	30
3	LVMC1	VAQC1	5	287	0.00	3.1	0.02	31	12.4	3.8	30
4	MSAC1	*KMHS*	0	18	0.97	2.5	0.02	31	8.4	2.2	31
5	RRAC1	*KRDD*	1	-1	1.00	1.5	0.02	33	4.0	2.5	33
6	GRSC1	TR721	10	-13	0.93	0.8	0.01	34	3.5	1.6	34
7	NWQC1	NWRC1	3	-8	0.99	3.4	0.07	34	12.4	4.1	34
8	LTRC1	PLEC1	12	94	0.37	2.7	0.03	35	10.1	3.0	35
9	PLEC1	*KCCR*	9	435	0.00	2.9	0.04	37	14.9	2.1	36
10	MDAC1	PIBC1	8	686	0.00	4.7	0.04	36	22.6	5.7	36

Table 5

As in Table 4 except for the 24 h period from 0000- 2300 UTC 10 July 2008.

No.	RAWS Station	Nearest Neighboring Station (NWS in italics)	Horizontal Separation (km)	Vertical Separation (m)	Weighted proximity	Temper-ature CV score (°C)	Temper-ature Sensitivity (°C)	Temper-ature Degradation (%)	Relative Humidity CV score (%)	Wind CV score (ms^{-1})	Median Degra-dation (%)
1	MOWC1	*KAAT*	5	2	0.98	1.5	0.01	29	14.1	2.9	29
2	LVMC1	VAQC1	5	287	0.00	4.4	0.03	30	12.7	5.4	30
3	VAQC1	LVMC1	5	-287	0.00	3.3	0.03	30	7.3	8.1	30
4	MSAC1	*KMHS*	0	18	0.97	4.9	0.05	33	10.9	1.4	33
5	RRAC1	*KRDD*	1	-1	1.00	1.4	0.02	33	3.6	2.7	33
6	GRSC1	TR721	10	-13	0.93	2.6	0.03	34	5.0	2.4	34
7	TS678	RRRC1	9	188	0.03	2.3	0.02	34	9.2	2.7	34
8	MDAC1	PIBC1	8	685	0.00	3.9	0.03	36	19.8	7.8	36
9	NWRC1	NWQC1	3	-8	0.98	2.9	0.06	37	20.2	2.4	37
10	SETC1	TS678	15	73	0.51	2.3	0.02	38	7.6	2.5	38

Table 6

Median RMSE of control analyses and CV scores from withheld analyses, sensitivity, and degradation computed from the cross validation experiments withholding RAWS and NWS stations throughout the continental United States

	RAWS	NWS
Median Control RMSE		
Temperature (^oC)	1.4	1.0
Relative Humidity (%)	5.6	4.5
Vector Wind (m s^{-1})	1.2	1.0
Median Withheld CV scores		
Temperature (^oC)	2.0	1.3
Relative Humidity (%)	8.4	7.0
Vector Wind (m s^{-1})	2.7	2.4
Median Sensitivity		
Temperature (^oC)	0.03	0.03
Relative Humidity (%)	0.13	0.16
Vector Wind (m s^{-1})	0.04	0.06
Median Degradation	48%	44%