4.5 Comparison of weather data from the Remote Automated Weather Station network and the North American Regional Reanalysis

Beth L. Hall and Timothy. J. Brown
DRI, Reno, NV

ABSTRACT. The North American Regional Reanalysis (NARR) is an assimilated dataset at a 32-km spatial and 3-hour temporal resolution. Due to its completeness, it offers an opportunity to estimate missing and erroneous data from other atmospheric datasets. Based upon multiple data sets including rawindondes, aircraft, and surface weather stations, NARR assimilates this data into modeled output. However, one of the datasets that NARR does not include is the Remote Automated Weather Station (RAWS) network (Zachariassen et al 2003). This is a network of over 2000 currently active weather stations throughout the US operated by federal and state agencies for the dominant use of wildfire applications. Therefore, many of these stations are placed in high-elevation and/or remote locations. The Fire Program Analysis (FPA) system is a multi-agency effort to plan, budget, and evaluate the effectiveness of alternative fire management strategies, and is highly dependent upon historical weather data from fire prone regions. Unfortunately, RAWS data has periods of missing and/or erroneous values for each station in the network, and FPA requires a complete weather dataset of best possible data. In the project described here, the NARR dataset has been integrated statistically with the RAWS data for FPA. Since RAWS was not used as one of the input datasets in NARR, the correlations between NARR and RAWS are not always strong. This analysis correlates temperature, humidity, precipitation, and wind speed data between NARR and RAWS at varying spatial and temporal scales. These are important variables for fire management. The results are intended to be informative on the discrepancies and similarities that occur when mid- and high-elevation remote location weather data are not integrated into, but compared to NARR.

1. INTRODUCTION

The National Center for Environmental Prediction's (NCEP's) North American Regional Reanalysis (NARR) is a long-term, dynamically consistent, high-resolution, high-frequency, atmospheric and land surface hydrology dataset for the North American domain (Mesinger et al 2006). At a 32-km spatial resolution and 3-hourly temporal resolution, the NARR provides a complete dataset that has the potential to be used as a source for replacing missing or erroneous data at observational locations. The NARR is a modeled dataset from 1979-present that incorporates data from rawinsondes, dropsondes, pibals, aircraft, selected surface stations, and geostationary satellites. It also incorporates high-resolution data from a variety of other sources such as the NCEP / Climate Prediction Center (CPC), Canadian, and Mexican precipitation network including modeled data from the Parameter-elevation Regression on Independent Slopes Model (PRISM) (Daly et al 1994).

The wildfire community in the US is heavily dependent upon surface weather information from the RAWS network. These stations take precedence over the more commonly used observational data networks that NARR has integrated due to their remote, high-elevation locations where wildfires are most likely – especially in the western US. The RAWS database, however, has periodic missing and erroneous data that affects climatological analysis of weather in these unique locations. To potentially use NARR data as estimates for RAWS, it is important to assess how well the NARR data correlates with the RAWS data. If the correlations are high, then perhaps NARR could be used as direct proxy for any missing or erroneous RAWS. Even with modest correlations, it may be possible (and perhaps even desirable), to generate regression equations with NARR variables as the predictors for a particular RAWS predictand.

RAWS data are predominantly recorded hourly at point locations. NARR is provided

* Corresponding author address: Beth L. Hall, Desert Research Institute, 2215 Raggio Pkwy, Reno, NV 89512; email: Beth.Hall@dri.edu

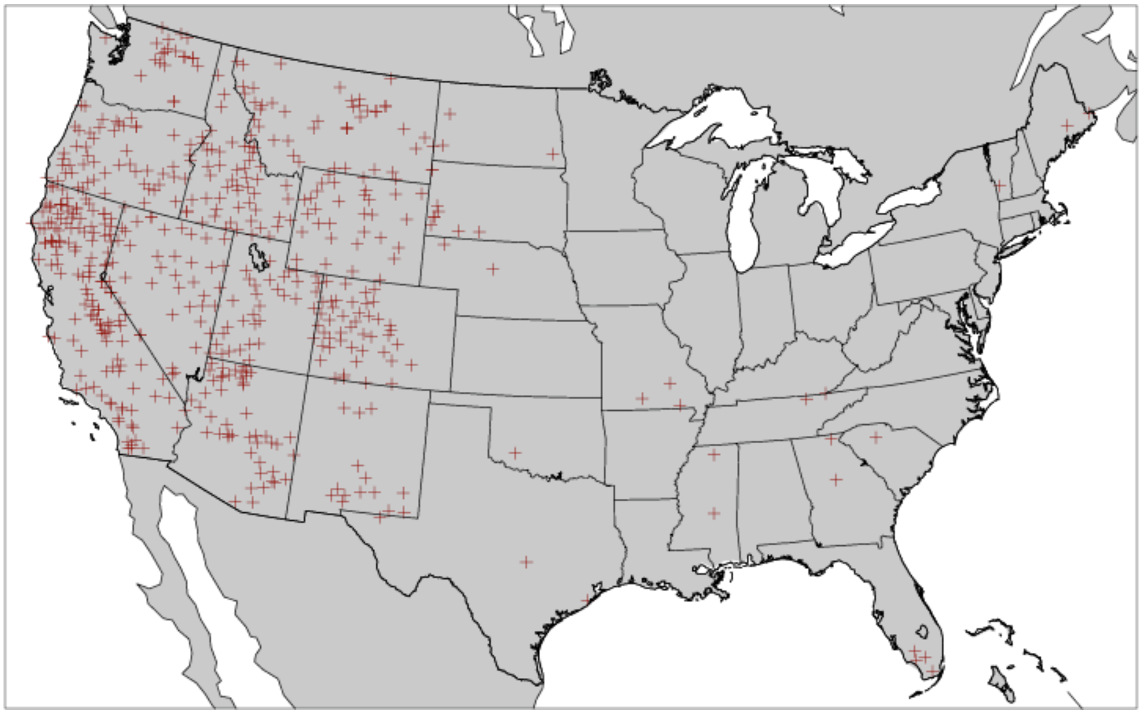

Figure 1. RAWS locations used in the correlation analysis.

at 3-hourly intervals on a 32-km grid. In order to replace actual observations with NARR values, some hourly interpolation will need to be made within the 3-hourly intervals. The correlations between NARR and near-surface observations (including rawinsondes) showed that monthly correlations were very high, even for highly variable precipitation amounts. This paper analyzes correlations between NARR and RAWS temperature, relative humidity, wind speed and precipitation amounts at an hourly, daily, 10-day, and 30-day time interval to determine if NARR and RAWS surface data are correlated sufficiently for NARR to be used in place of missing or erroneous RAWS.

2. DATA AND METHODS

RAWS surface temperature, humidity, wind speed, and 3-hourly accumulated precipitation amount are correlated to the same variables from NARR at the nearest grid point from 1990-2004. For unit comparison, NARR precipitation amount was converted to inches and NARR wind speed was converted to miles per hour. There were 591 stations that were selected based upon their length and quality of data record (Figure 1). If all 4 surface variables were available from RAWS, both the NARR and RAWS values were collected for analysis. Pearson correlations were computed for every third hour in RAWS that corresponded to NARR (in UTC time). Hourly correlations considered every observation (record) collected between the two datasets. Daily correlations were based on the average of 8 consecutive records, though not necessarily for a consecutive 24-hour period, in case there were missing records within that time. Correlations for 10-day and 30-day periods were based upon the averages of 80 and 240 consecutive records, respectively. A minimum of 20, 30-day records was required for including a station in the analysis (i.e., a minimum of 4800 hourly observations).

3. RESULTS

Figure 2 shows the distribution of correlations for temperature, relative humidity, wind speed and average 3-hourly accumulated precipitation amount based upon the values shown in Table 1. These box plots graphically show the extent of the values with the whiskers extending all the way to the maximum and minimum values within the distribution. These whiskers are not to the statistical steps with extreme outliers noted by circles. At the finest time scale (hourly), temperature has the strongest correlation between RAWS and NARR (median r=0.91). The median correlation increased to 0.93, 0.98, and 0.99 (out to 2 significant digits) for daily, 10-day, and 30-day time periods, respectively. Relative humidity, though lower at the hourly time scale compared to temperature (median r=0.75) increases to 0.86 and 0.89 at the 10- and 30-day time periods, respectively. Wind speed had the lowest hourly correlations overall (median r=0.22). The 10- and 30-day correlations improved marginally over the hourly correlations (median r=0.33 and 0.36, respectively). This is possibly due to RAWS being placed in complex terrain with potential exposures in windy or sheltered locations. Average precipitation amount had higher correlations than for wind speed, but not as strong as for temperature and relative humidity. The hourly, median correlation for precipitation amount was 0.29, however the 10-day and 30-day median correlations increased to 0.76 and 0.80, respectively. These results conflict with those of Mesinger et al (2006), who compared the monthly average precipitation amount from the input datasets to the gridded, monthly precipitation values from NARR. Their results indicate very high correlations at the monthly time scale for the continental US. The results using RAWS observations suggest that some stations do correlate well with NARR at the 30-day time scale, but this is not consistent throughout the country.

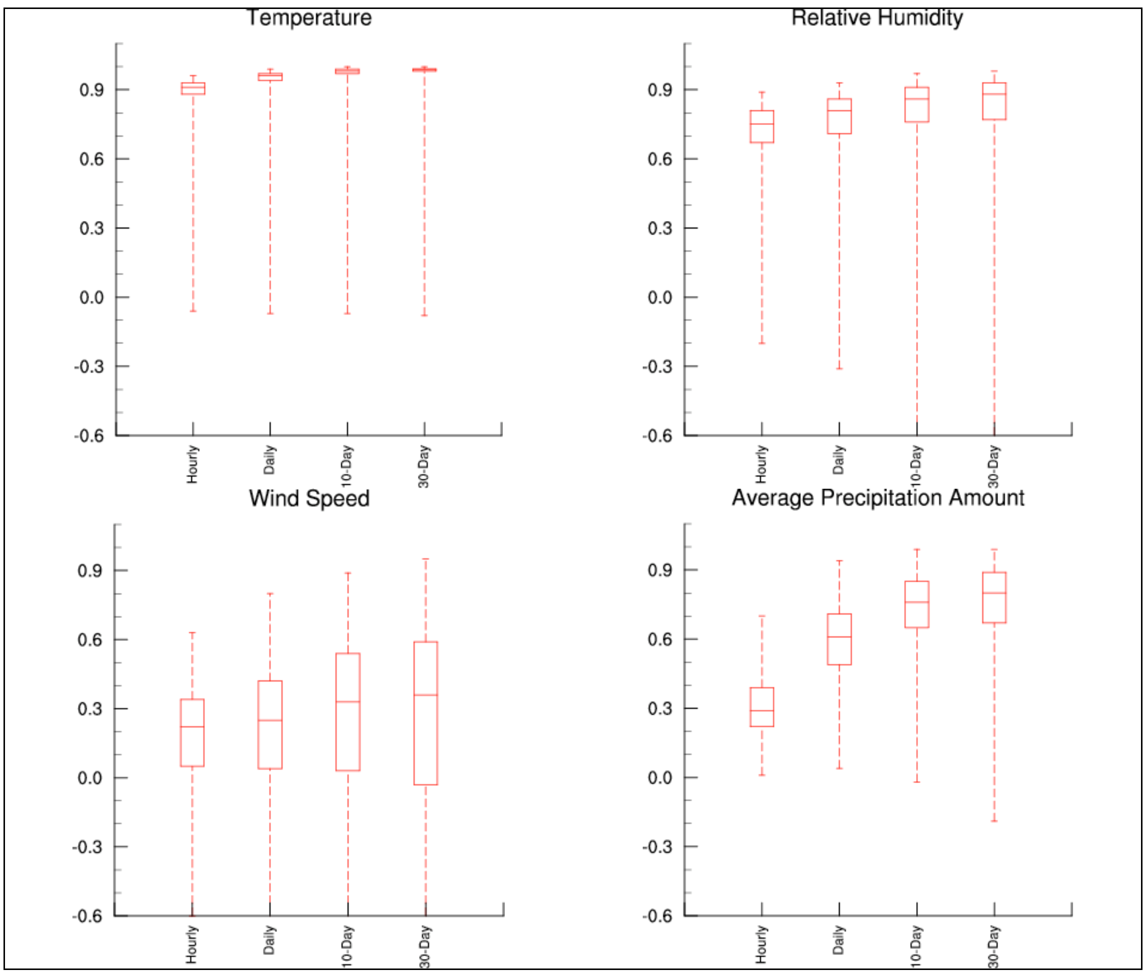

Figure 2. Box plots showing the distribution of correlations among the 591 RAWS stations

Along with the correlations, bias was also computed for each variable to understand whether or not NARR is typically greater or less than RAWS (Table 2). For all variables examined, the median bias is very close to zero indicating that the values between NARR and RAWS are often very close. Even though wind speed had the lowest correlations between NARR and RAWS, the median bias for all time scales was approximately 1.3 mph. This indicates that though NARR may not always be the same value as RAWS, the wind speed difference between the two datasets is still relatively close. Depending upon the needs of the users of RAWS data, a wind speed difference of less than 5 mph may be acceptable for most estimations.

Temperature						
Time period	**Min.**	**Q1**	**Q2 (Med.)**	**Q3**	**Max.**	**Avg.**
Hourly	-0.06	0.88	0.91	0.93	0.96	0.88
Daily	-0.07	0.94	0.96	0.97	0.99	0.94
10-day	-0.07	0.97	0.98	0.99	1.00	0.97
30-day	-0.08	0.98	0.99	0.99	1.00	0.97
Relative Humidity						
Time period	**Min.**	**Q1**	**Q2 (Med.)**	**Q3**	**Max.**	**Avg.**
Hourly	-0.20	0.67	0.75	0.81	0.89	0.71
Daily	-0.31	0.71	0.81	0.86	0.93	0.75
10-day	-0.64	0.76	0.86	0.91	0.97	0.79
30-day	-0.83	0.77	0.88	0.93	0.98	0.79
Wind Speed						
Time period	**Min.**	**Q1**	**Q2 (Med.)**	**Q3**	**Max.**	**Avg.**
Hourly	-0.60	0.05	0.22	0.34	0.63	0.19
Daily	-0.71	0.04	0.25	0.42	0.80	0.21
10-day	-0.82	0.03	0.33	0.54	0.89	0.26
30-day	-0.91	-0.03	0.36	0.59	0.95	0.26
Precipitation (Average Amount)						
Time period	**Min.**	**Q1**	**Q2 (Med.)**	**Q3**	**Max.**	**Avg.**
Hourly	0.01	0.22	0.29	0.39	0.70	0.31
Daily	0.04	0.49	0.61	0.71	0.94	0.60
10-day	-0.02	0.65	0.76	0.85	0.99	0.73
30-day	-0.19	0.67	0.80	0.89	0.99	0.75

Table 1. Quartile values of the correlations between NARR and RAWS.

4. CONCLUSIONS

Overall, NARR correlates well with RAWS at the 10-day and 30-day time scales, with the exception of wind speed. From a climatological perspective therefore, NARR could be used directly in place of RAWS as an estimate when trying to determine the average surface conditions of temperature, relative humidity, and average precipitation at a weekly to monthly time scale. At the hourly time scale, however, NARR should be used with caution for replacing hourly observations of RAWS, with the exception of temperature. Some of the low hourly correlations are likely due to the high-elevation and remote locations of RAWS that are far removed from more common observational data that was used as input into the NARR data model.

Future work will examine correlations within a diurnal and seasonal perspective, continuing to use the hourly, daily, 10-day, and 30-day control periods. In other words, how are the daily correlations when not considering the potentially full 24-hour periods within a day, but the 12-hour periods that represent daytime and nighttime separately? An example of the seasonal consideration would be to examine the correlations for 10-periods that fell within the months of December through February compared to June through August. Future work will also include the additional variable of incoming shortwave radiation. Radiation can be used to estimate state of the weather (e.g., extent of cloudiness) and green-up dates – two parameters that are essential for fire danger models in the US.

Temperature (K)						
Time period	**Min.**	**Q1**	**Q2 (Med.)**	**Q3**	**Max.**	**Avg.**
Hourly	-12.17	-0.58	1.27	3.14	14.65	1.30
Daily	-12.17	-0.58	1.27	3.14	14.65	1.30
10-day	-12.18	-0.59	1.27	3.15	14.67	1.31
30-day	-12.18	-0.60	1.29	3.16	14.74	1.32
Relative Humidity (%)						
Time period	**Min.**	**Q1**	**Q2 (Med.)**	**Q3**	**Max.**	**Avg.**
Hourly	-29.22	-2.27	1.32	5.97	52.90	2.32
Daily	-29.22	-2.27	1.33	5.97	52.90	2.32
10-day	-29.24	-2.30	1.32	6.01	52.93	2.32
30-day	-29.31	-2.37	1.30	5.96	52.93	2.30
Wind Speed (m.p.h.)						
Time period	**Min.**	**Q1**	**Q2 (Med.)**	**Q3**	**Max.**	**Avg.**
Hourly	-3.89	-0.17	0.41	1.23	5.50	0.61
Daily	-5.62	-0.88	0.00	1.02	7.41	0.14
10-day	-7.58	-1.42	-0.34	1.02	11.06	-0.13
30-day	-10.26	-1.69	-0.35	1.36	24.75	-0.08
Precipitation (Average Amount (in.))						
Time period	**Min.**	**Q1**	**Q2 (Med.)**	**Q3**	**Max.**	**Avg.**
Hourly	0.02	0.07	0.10	0.19	0.74	0.14
Daily	0.00	0.04	0.07	0.12	0.60	0.09
10-day	-0.06	0.02	0.04	0.08	0.46	0.06
30-day	-0.14	0.02	0.03	0.07	0.38	0.05

Table 2. Quartile values of the bias between NARR and RAWS (NARR average – RAWS average).

Where NARR shows strong correlations with RAWS, the opportunity exists to use NARR as proxy for missing or erroneous RAWS. Another application of NARR, if NARR correlates well with RAWS data, would be the development of national gridded fire danger indices based upon historical NARR data. Though forecast gridded fire danger indices are produced from model data (e.g., http://www.wfas.net), being able to develop historical fire danger data at a 32-km resolution since 1979 would be useful to assess fire danger seasonality and variability, and examine relationships with global climate variables such as pressure and sea surface temperatures.

5. REFERENCES

Daly, C., R. P. Neilson, D. L. Phillips, 1994: A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *J. of Appl. Meteor.,* **33**, 140-158.

Mesinger, F. and Co-authors, 2006: North American Regional Reanalysis, *Bull. Of Amer. Meteor. Soc.*, March, 343-360.

Zachariassen, J., K. Zeller, N. Nikolov, and T. McClelland, 2003: *A review of the Forest Service Remote Automated Weather Station (RAWS) network.* USDA Forest Service, Gen. Tech. Rep., RMRS-GTR-119, 153 pp.