Program for Climate, Ecosystem and Fire Applications

RAWS Data Quality Check and Estimation

Phase I

Beth L. Hall
Timothy J. Brown

Desert Research Institute
Division of Atmospheric Sciences

CEFA Report 05-01 April 2005

Forward

In November 2000 an Assistance Agreement 1422RAA000002 was established between the Bureau of Land Management National Office of Fire and Aviation and the Desert Research Institute. This report describes the activities at the DRI Program for Climate, Ecosystem and Fire Applications (CEFA) under Task Order 17 – RAWS Data Quality Check and Estimation (Phase I) that covers September 2004 through November 2004. For further information regarding this report or the projects described, please contact either:

Dr. Timothy Brown
Program for Climate, Ecosystem and Fire Applications
Desert Research Institute
2215 Raggio Parkway
Reno, NV 89512-1095
Tel: 775-674-7090
Fax: 775-674-7016
Email: tim.brown@dri.edu

Beth Hall
Program for Climate, Ecosystem and Fire Applications
Desert Research Institute
2215 Raggio Parkway
Reno, NV 89512-1095
Tel: 775-674-7174
Fax: 775-674-7016
Email: beth.hall@dri.edu

Paul Schlobohm
USDI Bureau of Land Management
National Office of Fire and Aviation
3833 South Development Ave.
Boise, ID 83705
Tel: 208-387-5196
Fax: 208-387-5179
Email: Paul_Schlobohm@nifc.blm.gov

RAWS Data Quality Check and Estimation
Phase 1

by
Beth L. Hall and Timothy J. Brown

Program for Climate, Ecosystem and Fire Applications
Desert Research Institute

March 2005

1. INTRODUCTION

This report describes the results of the task elements and deliverables of Task Order 17: RAWS Data Quality Check and Estimation – Phase I. The project went from October 2004 through November 2004 under the Bureau of Land Management (BLM) national Office of Fire and Aviation and the Desert Research Institute (DRI) cooperative Assistance Agreement (AA) 1422RAA000002.

Historical weather information from Remote Automated Weather Stations (RAWS) has many uses for fire management, such as applications for fire planning analyses, assessments of fire danger, fire severity, prescribed burn planning and ecosystem health to name a few. In order to use this information effectively, data need to be “clean” with respect to erroneous values. In some cases, estimates for missing or erroneous values may be obtainable to achieve a more complete record; in other cases, original values, if unreasonable, may have to be changed to missing (Brown, et al 2002).

A new fire planning analysis process is being developed to assess the fire program needs of local fire agency units using an interagency approach. The first module of the Fire Program Analysis System (FPA) was implemented in October 2004. Critical to this effort was the availability of high quality weather data. The primary sources of these data are the archives in the USDA-managed National Interagency Fire Management Integrated Database (NIFMID; daily) and the Western Regional Climate Center (WRCC; hourly). Neither of these archives applies a rigorous data quality filter to the original data. To optimize the performance of FPA, a process of data quality checking and estimation was necessary.

Recently, Brown, et al (2002) performed this type of data analysis for the fire agencies in California. In fact, through this effort, the quality check (QC) portion of the proposal has already been done for all but 2002 and 2003 for all RAWS stations in California. The QC process described by Brown et al (2002) requires hourly data, even for an assessment of the daily entry archived in NIFMID. Brown and McCurdy (1998) conducted a data estimation analysis for the prototype national fire planning process at the time. Many of the techniques and code developed for these two efforts were applied to this project and established a foundation to enable this quality assurance processing to be performed routinely on an annual basis.

To prepare for the implementation of FPA in March 2005, a quality checked (QC'd) data set was needed by November 2004 to assess the effectiveness of the chosen methodology and product output. FPA had four prototype areas identified in Alaska, California, Mississippi, and Oregon. To meet the short- and long-term needs of the FPA, there were two objectives for this project.

Objective 1. Provide data analysis for the four prototype areas by 1 November 2004, as agreed to by the participating agencies. The priority will be assuring data quality for the 1300-hr observations for stations provided by the agencies.

Objective 2. Develop a prototype of a complete hourly archive for all federal fire agency fire weather stations containing QC'd original observations and, where needed, estimated data. Develop a prototype process that may be applied operationally in the future to meet agency needs (e.g. daily, annually) for maintaining dataset currency.

This project was a collaborative effort between the Desert Research Institute (DRI) Program for Climate, Ecosystem, and Fire Applications (CEFA), the Western Regional Climate Center (WRCC) and the five federal wildland fire agencies (BLM, BIA, FWS, NPS, USFS). Specific task elements, project accomplishments and deliverables are given in the next section.

2. Task elements

Six specific task elements were required to assess the quality of the existing data, produce a complete alternative dataset of clean and estimated values, and provide documentation of this effort.

1) Utilize lists of weather stations provided by participating agencies. This may include manual and non-satellite telemetered automatic fire weather stations as well as the more common RAWS. Create a working set of existing period of record data for these stations. This working set will be manipulated into the final clean dataset without compromising either source archive (NIFMID and WRCC). Station identification records will be matched with physical installation location to develop a complete site record.
2) Wrote and implemented software code to perform an assessment of the data for impossible and unlikely values in order to establish the overall quality of the dataset and the resulting workload required to provide a complete quality record. Much of this code was adapted from previous efforts mentioned earlier. Data were categorized as Acceptable, Questionable, or Impossible, according to criteria defined in Brown et al (2002). For example, it is unreasonable to expect a relative humidity value of less than 0% or greater than 100%.
3) Acceptable data could be used as is. Impossible data are outside the realm of possibility and were marked as missing. Further analysis was necessary to

determine if Questionable data was actually Acceptable or Impossible. An example would be 24 consecutive hours of unchanging temperature, which is possible, but unlikely. Depending on the amount of Questionable data and the complexity of its analysis, this task could be intensive and impose a significant impact on the timetable of this project. For this reason, the assessment in this task and its projected impact on the project schedule was reported prior to proceeding to task element four. If necessary, the agencies would determine criteria for limiting the scope of this analysis in order to produce as much station data as possible within the schedule.

4) To produce a complete data set, missing data were estimated. Data were estimated for the 1300-hour (daily) observation, where gaps exist. Hourly data were addressed during objective 1 activities, if time allows, and objective 2 activities. Estimates would be made for state of the weather, dry-bulb temperature, relative humidity, wind direction, wind speed, 24-hour maximum temperature, 24-hour minimum temperature, 24-hour maximum relative humidity, 24-hour minimum relative humidity, precipitation duration, and precipitation amount based upon Brown and McCurdy (1998).
5) Wrote and implemented software code to place the complete dataset in data file formats required for agency use, including weather observation data transfer format (fwx or fw9) and text (comma delimited) files. The fwx/fw9 formats allows for the observations to be input into PCHA and Fire Family Plus.
6) Wrote a final report documenting the processes used, criteria employed, and descriptions of the confidence the agencies can place in the estimated data.

2.1 Acquire a list of weather stations

Lists of weather stations in the four prototype regions were provided in September 2004 (see Appendix 1). There were 206 stations in the original list. Only 117 of those could be processed during this first phase. The primary reason why a station had to be excluded from the analysis was because the matching station data could not be identified or located within the WRCC RAWS database. The WRCC database uses a 4-character identifier code for each RAWS station. This WRCC code is used to download the historical weather data for each station. Without the code, correlations could not be computed between the hourly RAWS and the daily data available in NIFMID that were necessary for estimating missing data.

2.2 Coarse quality assessment of original hourly RAWS data

Coarse data quality assessment was performed on all available prototype RAWS using similar criteria as that found in CEFA Report #02-01 *Quality control of California RAWS historical data* (Brown 2002). Appendix 2 outlines the checks made on temperature, relative humidity, wind speed, wind direction, and precipitation. Changes from the California RAWS project allowed for potential climatological differences from state to state for maximum and minimum temperature value limits and precipitation amount.

Climatological data from the WRCC was used to identify reasonable value limits for these parameters. Appendix 3 lists the limit values used for each state.

2.3 Hand checks on questionable data

The original intention was to do hand-checks on data that was considered to be questionable, but not impossible. Because of the large amount of data, and limited time to deliver the products, this task element was removed during the project period. Therefore, all data that was *not* originally considered to be acceptable was flagged for future estimation. Ideally, if the originally questionable data was reasonable, then the estimated value would have been similar to the original value so removing these from the original dataset should not have posed a problem.

2.4 Estimation of missing or non-acceptable observations

Estimation of missing or non-acceptable observations was made for temperature, relative humidity, wind speed, wind direction, and precipitation for every hour since each station was first made active. All estimations depended upon the existence of good atmospheric data at a nearby location to each station for the hour of interest. Since all surface weather stations are susceptible to periods of inactivity or data collection errors, there was a desire to identify an atmospheric data set that was known to have no gaps. The National Center for Environmental Prediction (NCEP) / National Center for Atmospheric Research (NCAR) reanalysis data set meets this requirement (Kalnay et al 1996). Data was acquired from three different data sets within reanalysis: surface level, pressure level, and flux (see Appendix 4). The surface and pressure level data set is at a 2.5 degree spatial resolution and the flux data is on a T62 Gaussian grid (approximately 1.875 degrees in the east-west direction and 1.9 degrees in the north-south direction). All datasets were at a 6-hourly temporal resolution.

A stepwise multiple regression routine was used to determine which reanalysis variables had the strongest correlation to the RAWS variable being estimated. There were originally 554 predictor variables considered in each regression analysis. These included the 48 possible pressure level variables, the 8 surface variables and 4 flux variables from reanalysis for the nine grid cells surrounding and including the RAWS location. This also included RAWS persistence of all 7 RAWS variables (temperature, humidity, speed, direction, precipitation, u and v vector components of the wind) for both the previous hour and the current hour from yesterday. The statistical software S-Plus package was programmed and used to perform the statistical analysis.

Output from S-Plus included equations and regression coefficients for each variable. Unique equations were collected for each station, for each of the 7 RAWS variables, for two seasons in the year (May through October; November through April), and for 4 time periods throughout the day (00-05 UTC, 06-11 UTC, 12-17 UTC, and 18-23 UTC). Therefore, each RAWS station had 56 equations associated with it that were used to compute any missing RAWS observation since the station was made active through

2003. Since there were 117 stations processed, this means that there were 6,552 equations ultimately produced.

Unfortunately, since the majority of equations relied upon at least one value of persistence, consecutive hours of missing observations became so heavily dependent upon the previous hour's estimated value that it was not uncommon for the estimation algorithms to produce meaningless results. If the persistence value was an estimated value, then a continuing trend in values (either increasing or decreasing) was possible. For example, temperature values would be increasing with each consecutive hour. A check was written into the software code that counted how many hours the same trend in values occurred. After the 20th hour, all estimated values from that point on were re-assigned to missing until the next good observation was encountered. For example, if temperature values increased from hour to hour for more than 20 hours, then those first 20 hours were retained as either the original observed or estimated values, but every consecutive estimated value after the 20th hour was re-assigned to missing until an hour with a good observation was reached. Using 20 hours as a cutoff threshold was determined by looking at historical data and determining the longest number of consecutive hours a constant trend (either increasing or decreasing values) occurred. None of the variables examined from historical data either increased or decreased in value for as many as 20 consecutive hours.

In addition to missing RAWS observations being estimated through various equations, the hourly state of the weather was also computed based upon a combination of data from the reanalysis dataset and the RAWS values (either estimated or observed). Reanalysis total cloud cover along with RAWS relative humidity, temperature, and precipitation was used to assign one of the 10 categories of state of the weather (Appendix 5). The algorithms used for each category do not compare directly with the definitions from NIFMID. This was to adjust for the unique distributions of values from the reanalysis data. Several analyses were performed to adjust the estimation algorithms so that the distribution of each estimated state of the weather category was similar to the distribution of actual state of the weather values for that station in NIFMID. Several stations in California were used for this frequency distribution analysis of the state of the weather in developing the state of the weather algorithms.

2.5 Deliverable data files

Data for each station was available via DRI's anonymous FTP site. Notification of the data availability was made to Howard Roose and Susan Weber. The anonymous FTP site is cleaned out on a regular basis, so current access to the data is by request at no cost. Data was provided in 3 separate text/ASCII formats. Data files ending with an .fw9 extension consisted of all hourly data in the 1998 WIMS data format. Files ending with an .fwx extension consisted of the once-daily values for 1300 LT in the 1972 WIMS data format. The third data file format has a .dat extension and was comma delimited including all values and flags indicating whether (1) the value was the original, good observation, (2) the value was estimated, or (3) the value was made missing because the algorithm failed after 20 hours OR the estimation was physically unreasonable (e.g., relative humidity well above 100%).

3. FUTURE WORK

This project has been extended into a second phase that will include nearly all RAWS stations throughout the United States. Initial data will be provided with the phrase 'unvalidated' in the filename to indicate that statistical analysis of the confidence in all estimated values has yet to be statistically assessed. It is intended during this second phase that validation will take place and data files will be modified and re-distributed accordingly. The initial quality control and estimation on all US stations will be performed by 1 March 2005 with the validation process to begin thereafter.

Acknowledgements

Special thanks are extended to Howard Roose, FPA Business Coordinator, for his help in providing the list of RAWS stations and continued support for this project, to Greg McCurdy at the Western Regional Climate Center for providing the RAWS data and offering insight and suggestions into how to handle problematic RAWS data, and to Paul Schlobohm for agency project coordination.

References

Brown, T. J, B. L. Hall, and G. D. McCurdy. 2002. Quality Control of California RAWS Historical Data. CEFA Report #02-01.

Brown, T. J and G. D. McCurdy. 1998. National Fire Weather Data Retrieval Final Report. WRCC Report #98-01.

Kalnay, E. and Co-Authors, 1996: The NCEP/NCAR 40-year reanalysis project. Bulletin of the American Meteorological Society*,* **77**(3), 437-471.

Appendix 1

List of weather stations provided by the agencies represent the prototype regions of Alaska, Oregon, California, and Mississippi.

NIFMID ID	Latitude	Longitude	Elevation (ft)	NAME	WRCC code
500726	62.83	-141.474	1800	ALCAN HWY MI-1244	aalc
500742	65.017	-146.22	1100	ANGEL CREEK	aang
500423	65.593	-144.356	850	BIRCH CREEK	abck
500414	65.288	-143.068	1850	BEN CREEK	aben
500939	61.525	-149.915	100	BIG LAKE	abil
500966	61.949	-151	150	BENTALIT	abnt
500811	57.271	-154.559	171	BOOTH LAKE	aboo
500902	60.491	-149.779	625	BROADVIEW	abro
500740	65.186	-147.508	2517	CARIBOU PEAK	accr
500721	65.017	-148.593	1450	CHATANIKA	acha
500933	62.136	-142.085	3318	CHISANA	achi
500421	66.593	-144.339	450	CHALKYITSIK	achl
500949	62.559	-144.661	2300	CHISTOCHINA	achs
500945	61.525	-144.441	581	CHITNA	acht
500747	64.051	-141.932	2860	CHICKEN	ackn
500733	65.339	-155.949	1310	COTTONWOOD	acot
500725	64.779	-141.153	880	EAGLE	aeag
500741	64.847	-147.61	454	FAIRBANKS	afai
500624	62.729	-154.068	775	FAREWELL	afar
500625	62.83	-156.61	1480	FLAT	aflt
500748	63.847	-144.356	1525	GEORGE CREEK RAWS	ageo
500743	64.237	-145.271	1520	GOODPASTURE	agop
500416	67.034	-143.288	850	GRAPHITE LAKE	agra
501044	55.356	-132.695	1637	HAIDA	ahai
500731	67.746	-144.119	2800	HELMUT MTN.	ahel
500731	67.746	-144.119	2800	HELMUT MTN.	ahel
501042	55.525	-131.356	492	SHELTER COVE	ahlm
500417	66.746	-148.678	1075	HODZANA	ahod
500309	66.22	-155.678	685	HOGATZA RIVER	ahog
500965	59.746	-151.203	715	HOMER	ahom
501013	57.813	-135.136	450	HOONAH	ahon
500730	65.593	-163.407	1550	HOODOO HILL	ahoo
500615	63.39	-158.83	930	INNOKO FLATS	ainn
500936	62.61	-142.085	2300	JATAHMUND LAKE	ajat
501029	58.356	-134.576	25	JUNEAU RD	ajun
500322	64.424	-158.102	110	KAIYUH	akai
501026	56.983	-133.661	400	KAKE	akak

NIFMID ID	Latitude	Longitude	Elevation (ft)	NAME	WRCC code
500321	66.085	-152.169	825	KANUTI NWR	akan
500217	67.136	-159.034	235	KAVET CREEK	akav
500934	67.949	-162.305	412	KELLY	akel
500963	60.593	-150.322	400	KENAI NWR	aken
500206	66.983	-160.441	150	KIANA	akia
500735	60.322	-160.203	1910	KILBUCK	akil
500958	62.136	-144.932	3100	KLAWASI	akla
500908	60.373	-149.407	475	KENAI LAKE	aklk
500319	66	-157.576	100	KOYUKUK NWR	akoy
500732	66.085	-143.373	1300	LITTLE BLACK	alib
500745	65.424	-148.729	450	LIVENGOOD	aliv
500623	63.898	-152.305	740	LAKE MINCHUMINA	almi
500406	66.034	-147.983	700	LOST CREEK	alos
500957	61.322	-142.593	1600	MAY CREEK	amay
500724	63.644	-151.644	840	MCKINLEY RIVER	amck
500962	60.034	-151.661	130	NINILCHIK	anin
500102	68.068	-158.712	985	NOATAK	anoa
500317	66.847	-154.339	800	NORUTAK LAKE	anor
500810	60.186	-154.322	260	PORT ALSWORTH	apal
500931	62.949	-145.508	2670	PAXSON	apax
500618	64.102	-155.559	935	POORMAN	apoo
500738	65.932	-145.017	1038	PREACHER CREEK	apre
500215	65.407	-164.661	427	QUARTZ CREEK	aqtz
500942	61.085	-149.729	1480	RABBIT CREEK	arab
500505	61.712	-162.661	140	REINDEER RIVER	arei
500736	64.695	-153.949	570	ROUND LAKE	arou
500412	66.813	-141.627	2210	SALMON TROUT	asal
500734	66.61	-159.102	105	SELAWIK	asel
500405	65.949	-149.864	823	SEVEN MILE	asev
500744	64.593	-146.136	1000	SALCHA	aslc
500956	61	-153.898	1250	STONEY	asto
500621	61.644	-156.441	265	STONEY RIVER	astr
500924	60.729	-150.881	280	SWANSON RIVER	aswa
500620	63.441	-153.356	650	TELIDA	atel
501040	55.746	-132.762	600	THORNE RIVER	atho
500723	63.763	-143.83	2073	T LAKE	atla
500749	62.966	-143.339	2300	TOK RIVER VALLEY	atok
500746	64.407	-148.458	556	GOLD KING	atta
500420	66.796	-146.712	525	VUNZIK LAKE	avun
500715	64.305	-151.085	1050	WEIN LAKE	awei
500710	63.491	-150.881	2120	WONDER LAKE	awon
501028	56.305	-132.847	900	ZAREMBO	azar

NIFMID ID	Latitude	Longitude	Elevation (ft)	NAME	WRCC code
44730	35.884	-118.052	8228	BEAR PEAK	cbea
45009	35.452	-118.582	7548	BRECKENRIDGE	cbre
44722	36.093	-118.26	8200	BLACKROCK	cbrk
44719	36.791	-118.66	4720	CEDAR GROVE	ccdg
45002	35.532	-118.63	2356	DEMOCRAT	cdmc
44707	35.972	-118.545	4700	JOHNSONDALE	cjoh
45005	35.755	-118.417	2720	KERNVILLE	cker
44717	36.175	-118.702	3240	OAK OPENING	coko
44508	36.685	-119	4060	PINEHURST	cpih
44713	36.724	-118.943	7540	PARK RIDGE	cpkr
44728	36.412	-118.425	8600	RATTLESNAKE	crtl
45012	35.983	-118.583	3000	SOUTH FORK	csqs
44729	36.727	-118.675	8120	SUGARLOAF	csug
44712	35.889	-118.633	3720	UHL/HOT SPRINGS	cuhl
45014	35.581	-118.057	5572	WALKER PASS	cwal
44732	36.44	-118.702	5240	WOLVERTON	cwol
227802	30.848	-89.034	275	BLACK CREEK	mblc
352711	44.03	-120.4	-99	BADGER (CREEK)	obad
353342	43.528	-121.816	-99	BLACK ROCK	obla
352109	44.593	-119.278	-99	BOARD HOLLOW	oboh
352208	44.323	-119.767	-99	BRER RABBIT (BRIAR)	obri
353428	43.561	-120.249	-99	BROWNS WELL	obro
353402	43.5	-121.05	-99	CABIN LAKE	ocab
352619	43.78	-121.05	-99	CAMP 2	ocam
352701	44.35	-120.13	-99	COLD SPRINGS	ocod
352620	44.316	-121.606	-99	COLGATE	ocol
352107	44.45	-121.13	-99	HAYSTACK	ohay
350918	44.966	-121.491	-99	HEHE BUTTE	oheh
352618	43.93	-121.33	-99	LAVA BUTTE	olav
352110	44.627	-121.615	-99	METOLIUS ARM	omet
350916	45.03	-121.628	-99	MT WILSON	omtw
350917	44.926	-121.194	-99	MUTTON MTN.	omut
350915	45.028	-120.539	-99	NORTH POLE (RIDGE)	onpr
351001	45.322	-120.925	-99	PATJENS	opat
352605	43.764	-121.717	-99	ROUND MOUNTAIN	orou
352712	44.044	-120.666	-99	SALT CREEK	osal
352207	44.463	-120.294	-99	SLIDE	osli
350913	45.241	-121.453	-99	WAMIC MILL	owam
5009	62.949	-145.508	2670	PAXSON	
5007	67.746	-144.119	2800	HELMUT MTN.	
0	62.83	-141.39	2125	MILE POST 1243	
500964	60.729	-149.288	512	GRANITE	

NIFMID ID	Latitude	Longitude	Elevation (ft)	NAME	WRCC code
0	64.763	-156.881	120	FWSAKRX3	
501040	55.746	-132.762	600	THORNE RIVER	
0	62.712	-146.627	3050	RENEE	
500746	64.407	-148.458	556	GOLD KING	
501031	56.508	-132.796	163	WOODPECKER	
501030	55.356	-132.474	800	POLK PEAK	
0	62.712	-150.542	3300	RUTH GLACIER	
0	66.271	-146.525	483	BEAVER (WBQ)	
0	64.847	-147.712	460	AKNPSEAST	
0	63.729	-148.932	1800	AKNPSWEST	
0	59.244	-135.517	0	HAINES	
0	59.46	-135.3	0	SKAGWAY	
0	66.57	145.246	0	FORT YUKON MBST	
0	59.345	151.842	0	PORT GRAHAM	
0	57.169	-157.271	175	BLACK LAKE	
0	60.136	-149.796	4200	HARDING ICEFIELD	
0	0	0	0	FISH CREEK	
0	63.729	-148.915	1800	DENALI VISITOR CENTER	
0	62.864	-145.627	2500	GULL RIVER	
0	63.339	-145.83	2700	DELTA	
0	69.779	-154.661	0	IKPIKPUK RIVER	
0	60.119	-143.288	75	BERING GLACIER	
500205	67.106	-157.854	0	AMBLER	
500952	61.594	-149.091	0	PALMER	
500959	57.75	-152.494	0	KODIAK	
500961	61.174	-149.996	0	ANCHORAGE	
500601	61.582	-159.543	0	ANIAK	
500103	71.286	-156.766	0	BARROW	
500301	66.915	-151.528	0	BETTLES	
500809	59.046	-158.503	0	DILLINGHAM	
500737	63.667	-144.533	0	DRY CREEK	
500713	64.664	-147.1	0	EILSON	
500702	64.816	-147.858	0	FAIRBANKS AIRPORT	
500701	63.995	-145.72	0	FORT GREELY	
500404	66.571	-145.25	0	FORT YUKON	
500905	62.155	-145.457	0	GULKANA	
500727	63.883	-149.017	0	HEALY	
500951	59.645	-151.478	0	HOMER MAN	
500805	59.753	-154.917	0	ILIAMNA	
500941	60.571	-151.248	0	KENAI	
500925	61.733	-145	0	KENNY LAKE	
501005	58.355	-134.576	0	JUNEAU	

NIFMID ID	Latitude	Longitude	Elevation (ft)	NAME	WRCC code
500212	66.884	-162.599	0	KOTZEBUE	
500606	62.953	-155.606	0	MCGRATH	
500211	64.512	-165.445	0	NOME	
500704	62.961	-141.929	0	NORTHWAY	
500954	62.717	-143.967	0	SLANA	
500929	60.475	-151.375	0	SOLDOTNA	
500305	65.174	-152.109	0	TANANA	
500955	62.05	-146.533	0	TAZLINA LODGE	
500947	62.042	-145.433	0	TAZLINA VILLAGE	
500720	63.328	-142.994	0	TOK	
500203	63.888	-160.799	0	UNALAKLEET	
0	60.491	-150.458	590	SKILAK GUARD STATION	
0	57.05	-135.367	0	SITKA	
0	57.169	-157.271	175	MOTHER GOOSE	
500703	64.547	-148.926	0	NENANA ASOS	
0	61.254	-149.525	3924	SITE SUMMIT	
5002	65.203	-161.153	177	HAYCOCK	
500501	60.78	-161.838	0	BETHEL	
500302	64.736	-156.937	0	GALENA AWOS	
501007	55.356	-131.714	0	KETCHIKAN	
500915	62.322	-150.094	0	TALKEETNA	
0	58.407	-152.898	75	BLACK CAPE	
0	57.729	-153.932	150	CHIEF COVE	
0	57	-153.542	100	CAPE KIAVAK	
352102	44.749	-121.614	-99	SHITIKE BT.	
350902	45.15	-121.583	-99	CLEAR LAKE	
350909	44.925	-121.535	-99	SIDWALTER	
352106	44.842	-121.233	-99	EAGLE BUTTE	
352108	44.775	-121.254	-99	WARM SPRINGS	
350920	44.956	-121.498	-99	HEHE 1	
352208	44.323	-119.767	-99	BRER RABBIT (BRIAR)	
352109	44.593	-119.278	-99	BOARD HOLLOW	
228102	30.38	-89.04	209	AIERY	
44512	36.8	-119.103	5156	DELILAH	
44701	36.492	-118.824	1700	ASH MOUNTAIN	
44721	35.9	-118	6240	CHIMNEY PEAK	
44726	36.073	-118.535	7167	PEPPERMINT	
45002	35.532	-118.63	2356	DEMOCRAT	
45009	35.452	-118.582	7548	BRECKENRIDGE	
45014	35.581	-118.057	5572	WALKER PASS	
500964	60.729	-149.288	512	GRANITE	
501031	56.508	-132.796	163	WOODPECKER	

NIFMID ID	Latitude	Longitude	Elevation (ft)	NAME	WRCC code
501030	55.356	-132.474	800	POLK PEAK	

Appendix 2 – Coarse quality assessment criteria

RAWS values were flagged for estimation if:

Air Temperature

a. The hourly value was less than historical period of record minimum temperature for the state or greater than the historical period of record maximum temperature for the state.
b. A single hourly value was missing.
c. 24 or more consecutive hourly values were identical.
d. 3 or more consecutive hours were less than the climatological minimum threshold.

Relative Humidity

a. The hourly value was less than 0% or greater than 100%.
b. A single hourly value was missing.
c. The hourly value remained constant for 24 or more consecutive hours.

Wind Speed

a. The hourly value was less than 0 mph.
b. The hourly value was greater than 200 mph.
c. Only a single hourly value was missing.
d. Up to 12 hourly values were missing and the two surrounding values were less than 3 mph.
e. The hourly value remained constant at 0 mph for 24 consecutive hours or more.
f. The hourly value was less than 2 mph and remained unchanged for 18 consecutive hours or more.
g. The hourly value was greater than or equal to 2 mph and remained unchanged for 12 consecutive hours or more.

Wind direction

a. The hourly value was less than 0 or greater than 360 degrees.
b. The hourly value remained constant for 8 consecutive hours or more.

Precipitation (running total, not hourly increment)

a. The hourly values were missing for up to and including 96 consecutive hours, and both hourly-accumulated values surrounding the missing period were identical.
b. An hourly value decreased from the previous hourly value and yet did not go lower than 0.2 (as a result of a counter reset); the hourly value was reset to the lower value.
c. There was an increase in an hourly value that was greater than 2 inches.

Appendix 3

Coarse data quality assessment limits used for maximum and minimum temperature (F) and precipitation amounts based upon climatological values acquired from the Western Regional Climate Center (WRCC).

State	Maximum Temperature (F)	Minimum Temperature (F)	Maximum Hourly Precipitation Amount (in)
Mississippi	115	-19	15
Oregon	119	-54	11
California	134	-45	26
Alaska	100	-80	15

Appendix 4 – Reanalysis variables considered for estimating RAWS values

Surface data

1. Air temperature
2. Relative humidity
3. Sea level pressure
4. Precipitable water
5. U-wind
6. V-wind
7. Wind speed (derived from u,v-winds)
8. Wind direction (derived from u,v-winds)

Surface Flux data

1. Downward long-wave radiation flux, in energy per area that reaches the ground
2. Downward short-wave radiation flux, in energy per area that reaches the ground
3. Precipitation rate

Other Flux data

1. Total cloud cover

Pressure Level data (1000 hPa, 925 hPa, 850 hPa, 700 hPa, 600 hPa, 500 hPa, 400 hPa, 300 hPa)

1. Air temperature
2. Relative humidity
3. U-wind
4. V-wind
5. Wind speed (derived from u,v-winds)
6. Wind direction (derived from u,v-winds)

Appendix 5 – State of the Weather Codes and Algorithms

Code	Description	Algorithm
0	Clear	Total cloud cover < 30%
1	Scattered clouds	Total cloud cover ≥ 30% and < 45%
2	Broken	Total cloud cover ≥ 45% and < 60%
3	Overcast	Total cloud cover ≥ 60%
4	Foggy	Total cloud cover ≥ 65% AND Relative humidity between 95% and 98%
5	Drizzling	Total cloud cover ≥ 65% AND Relative humidity ≥ 98%
6	Raining	Hourly precipitation > 0.0 inches AND Temperature > 32F AND Total cloud cover between 45% and 90%
7	Snow / sleet	Hourly precipitation > 0.0 inches AND Temperature ≤ 32F
8	Showering	Hourly precipitation > 0.0 inches AND Temperature > 32F AND Total cloud cover < 45%
9	Thunderstorms in progress	Hourly precipitation > 0.0 inches AND Temperature > 32F AND Total cloud cover > 90%