Data Substitution Among Weather Stations in Yellowstone National Park: Defining The Scope of the Problem

Michael Tercek, Ph.D.
Walking Shadow Ecology
PO Box 1085
Gardiner, Montana 59030
Tercek@YellowstoneEcology.com
www.ClimateAnalyzer.org

October 2015

Introduction

Recently, it was noticed that a large number of the observations in Yellowstone's manually operated Cooperative Observer Network (COOP) weather stations were identical to observations recorded by nearby automated weather stations. Subsequent interviews with the staff that record observations in Yellowstone and with National Weather Service personnel have confirmed that copying data between weather station records has in some cases become routine.

Under ideal circumstances, nearby weather stations would record identical measurements, but in practice, there are always significant discrepancies due to the use of different instrumentation, different times of observation, different data handling methodology, and site specific conditions, such as shading and wind exposure (Holder et al. 2006, Leeper et al. 2015). Mixing data from different weather stations introduces significant, artificial biases that make it impossible to conduct meaningful analysis, particularly if the data substitutions are not documented. The National Weather Service confirms this on its web site:

> "A cooperative station may be collocated with other types of observing stations such as standard observations stations, Flight Service Stations, etc. In these cases, that portion of the station observing program supporting the cooperative program's mission is treated and documented independently of the other observational and service programs." Follow this link to see the web page.

COOP stations have some of the longest running climate records in the country. The value of these observations is, ostensibly, their consistency. Using mostly unchanged methods for the entire record of observation helps ensure that trends seen in the data are as free as possible from artificial biases. As the web site cited above says : "COOP data plays a critical role in efforts to recognize and evaluate the extent of human impacts on climate from local to global scales."

Occasionally, thermometers, methodology, and site conditions at COOP stations do change, and peer-reviewed statistical methods have been developed by scientists (including some NOAA employees) to detect and fix the discontinuities and bias introduced into the record by these changes. Crucially, these methods either use comparisons among nearby weather stations to detect the discontinuities and

/ or they rely on the fact that the data from the manual COOP stations has been maintained separately and that changes to the station have been properly documented (Schaal and Dale 1977, Quayle et al 1991, Peterson and Easterling 1994, Easterling and Peterson 1995, Durre et al 2010, Oyler et al 2015). Mixing data among weather stations will invalidate these correction methods.

Remote Automated Weather Stations (RAWS), which in the present case have provided the data that has been pasted into the COOP records, are established for a variety of purposes and by a variety of different agencies, but most commonly they are used for fire prediction and modeling. Depending on the purpose of a particular RAWS, it likely has less stringent or variable standards governing the accuracy of instrumentation used and the condition of the site in which the instruments are placed. More information on RAWS are available here. And here. A map of weather stations in Yellowstone is located here.

This report uses simple comparisons of COOP vs. RAWS daily temperature data to quantifiy the amount of data substitution that has taken place in Yellowstone since 2009. Defining the extent of the problem will help park staff determine the best solution. Appendix 2 provides a description of conditions at the south gate of Yellowstone, where it was found that two thermometers only 11m apart have recorded complexly different temperature regimes – an example that serves as a warning against assuming that two data sources are equivalent.

Methods

The weather stations under consideration are shown in Table 1.

Station Location	COOP ID	Automated ID	Years Considered
Mammoth Hot Springs	489905	YLAW	2009 - 2015
Tower Ranger Station	489025	TFAW4	2009 - 2015
Lamar Ranger Station	485355	LMAW4	2014 - 2015
Old Faithful	486845	OFAW4	2009 - 2015
Lake Yellowstone	485345	KP60 (ASOS)	2009 - 2015
Snake River Station (South Gate)	488315	SEYW4	2010 - 2015

Table 1: Weather stations considered in this report. In every case, the COOP station (2nd column) was the recipient of the automated data (3rd column). The length of time considered varied because of data availability. The Lamar Automated Station was established in October 2014. The Lake automated station is an ASOS/MET station rather than a RAWS. There is also a RAWS (LKAW4) at Lake, but the amount of similarity between the COOP and this station was quite low.

For simplicity, this report focuses only on daily maximum temperature (Tmax). Because the COOP stations report data at 8am and the automated stations calculate daily maxima over calendar days (midnight - midnight), it was necessary to shift the COOP Tmax record back in time one day for a proper comparison. Without this shift, no identical values were found, confirming that the shift matched the substitution procedure used by the operators.

Graphs showing the difference between corresponding daily measurements were used to detect patterns of data substitution. In every case, difference was calculated as Automated$_i$ - COOP$_i$, where COOP$_i$ and Automated$_i$ are corresponding daily observations. A positive difference indicated that the automated station was warmer. Because the automated data was rounded to the nearest degree before substitution (station operators, personal communication), any corresponding values that differed by less than +/- 0.5 degrees Fahrenheit were considered identical.

In order to estimate the amount of bias that is being introduced by the data substitution at each calculation, Mean Absolute Error (MAE) was calculated for the days that remained after all the substituted observations had been removed:

$$\text{MAE} = \frac{1}{n} \sum |COOP_i - Automated_i|$$

MAE measures the average difference among daily measurements, regardless of which station has the greater value.

The RAWS data source at Yellowstone's south gate is located about 400m from the COOP in a location that is known to be warmer. Because of this known bias the RAWS values were lowered by variable amounts before they were copied into the COOP record (NWS personnel, personal communication). Because of this adjustment, it was impossible to estimate the amount of substitution that occured at the south gate. Appendix 2 provides more detail on the siting differences found at the south gate.

Results

The substitutions for Mammoth Hot Springs occurred primarily in the summer (Figure 1). During June - September 2015, for example, 90 consecutive days were identical in the COOP vs. automated stations at Mammoth. A contrasting example is Old Faithful, where the substitution is more uniform (Figure 1). Graphs for the other locations are shown in Appendix 1.

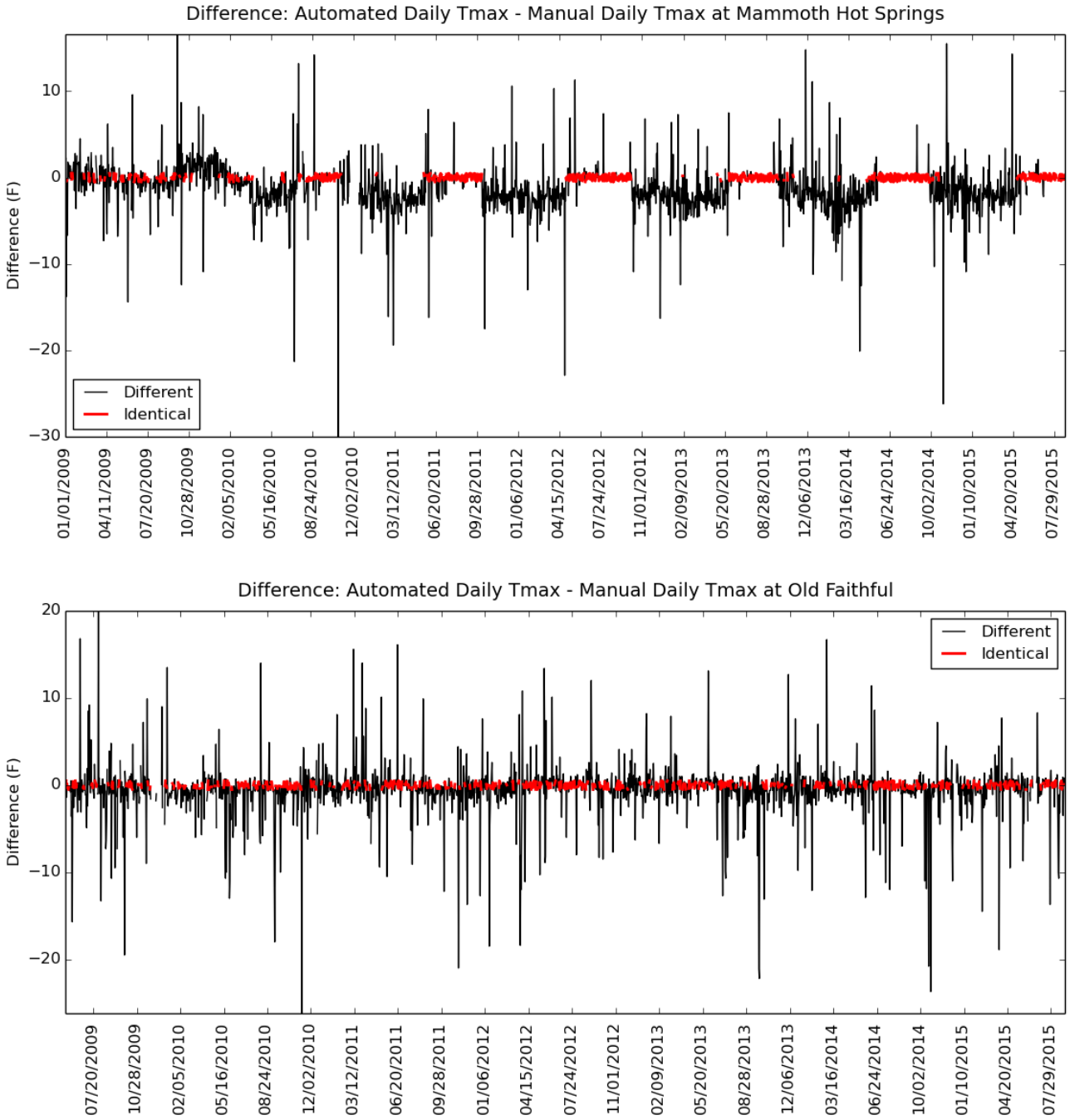


Figure 1: Difference in daily maximum temperature at Mammoth Hot Springs and Old Faithful. Positive values indicate that the automated station was warmer.

The amount of substitution and the estimated daily bias for each location are shown in Table 2.

Location	% values identical	MAE (Degrees F)
Mammoth Hot Springs	40%	2.5
Tower Ranger Station	23%	2.4
Lamar Ranger Station	30%	2.6
Old Faithful	48%	2.5
Lake Yellowstone	21%	3.1
Snake River Station (South Gate)	N/A	N/A

Table 2: The amount of data substitution and estimates of introduced bias at each location. Different time periods were considered at each location (Table 1). It was impossible to quantify the amount of substitution at the south gate because the RAWS data were modifed before copying.

Discussion

The desire to infill missing data is understandable, particularly in the remoter parts of Yellowstone where seasonal staffing changes make it impossible to collect data during the off-seasons (roughly early November until mid-December and mid-March until late April). Datasets that have large numbers of missing values are more difficult to use and make unconvincing graphs. Nevertheless, the desire to use observations that are not genuine as substitutes should be resisted because, as described in the introduction, the contamination will convert a difficult dataset into an unusable dataset. The solution is to use a data product that has been professionally infilled, e.g. TopoWx (Oyler et al. 2015) or PRISM (More information on PRISM is available here), and to keep in mind that those data products require properly documented source data in order to be accurate.

The Mean Absolute Errors (MAEs) reported here (Table 2) are absolute values. Consequently, an MAE of 2.5 degrees F indicates that two corresponding daily values differ by plus or minus 2.5 degrees, or that the range of differences might be as much as 5 degrees. These MAE values are also means, which by definition do not capture the extreme differences that sometimes occur among corresponding daily values. Figure 1 shows that the differences are frequently more than 10 degrees.

Recent analyses conducted by the current author and others in Yellowstone (in press) indicate that the observed rates of long-term temperature increase in Yellowstone are roughly 0.7 - 1.5 degrees (F) per decade, with considerable variation in this rate among seasons of the year and from location to location. In this context, the amount of potential error introduced into the datasets is large (Table 2). It is particularly unfortunate that the substitutions occurred in recent years, which were also some of the hottest on record.

Some of the identical values flagged in the graphs (Figure 1 and Appendix 1) *might* be indentical by coincidence, and it is therefore possible that the amount of cross-mixing has been over-estimated. This is more likely at Old Faithful (Figure 1), where the distribution of indentical values is more or less uniformly spread throughout the record instead of being aggregated into discrete time periods as at the other locations. Perhaps the two thermometers at that location are very close to each other and have very similar measuring characteristics.

An alternative method for detecting the amount of substitution in the records would be to count the number of days that contain observations in the COOP for Tmax, Tmin and Precipation but NOT snow depth. Since the RAWS do not record snow depth, the days infilled by NWS personnel contain missing markers for the snow depth parameter (NWS personnel, personal communication). Unfortunately, an examination of the Mammoth record, where the periods of infilling are in very clear seasonal chunks, shows that snow values were sometimes estimated or inferred during days that were infilled by NPS personnel. If the validity of this method could be better established, it might be employed at the South

Entrance, where it was not possible to directly compare the data.

Literature Cited

- Durre, I., M. Menne, B. Gleason, T. Houston, and R. Vose. 2010. Comprehensive Automated Quality Assurance of Daily Surface Observations. Journal of Applied Meteorology and Climatology 49: 1615 - 1633.
- Easterling, D.R, and T.C. Peterson, 1995: A new method for detecting and adjusting for undocumented discontinuities in climatological time series. Internation. Journal of Climatology, 15, 369-377.
- Holder, C., R. Boyles, A. Syed, D. Niyogi, and S. Raman. 2006. Comparison of collocated automated (NCECONet and Manual (COOP) Climate Observations in North Carolina. Journal of Atmospheric and Oceanic Technology 23: 671-682.
- Leeper, R.D., J. Rennie, and M. Palecki. 2015. Observational Perspectives from the US Climate Reference Network (USCRN) and Cooperative Observer Program (COOP) Network: Temperature and Precipitation Comparison. Journal of Atmospheric and Oceanic Technology 32: 703 - 721.
- Oyler, J. W., S. Z. Dobrowski, A. P. Ballantyne, A. E. Klene, and S. W. Running. 2015. Artificial amplification of warming trends across the mountains of the western United States, Geophys. Res. Lett., 42, 153–161, doi:10.1002/2014GL062803.
- Peterson, T.C., and D.R. Easterling, 1994: Creation of homogeneous composite climatological reference series. International Journal of Climatology, 14, 671-679.
- Quayle, R.G., D.R. Easterling, T.R. Karl, and P.Y. Hughes. 1991. Effects of recent thermometer changes in the cooperative station network. Bulletin of the American Meterological Society 72: 1718 - 1723.
- Schaal, L. and R. Dale. 1977. Time of observation temperature bias and climatic change. Journal of Applied Meterology 16: 215 - 222.

Appendix 1 - Patterns of substitution at Tower, Lamar and Lake.

See Figure 1 for the Mammoth and Old Faithful patterns.

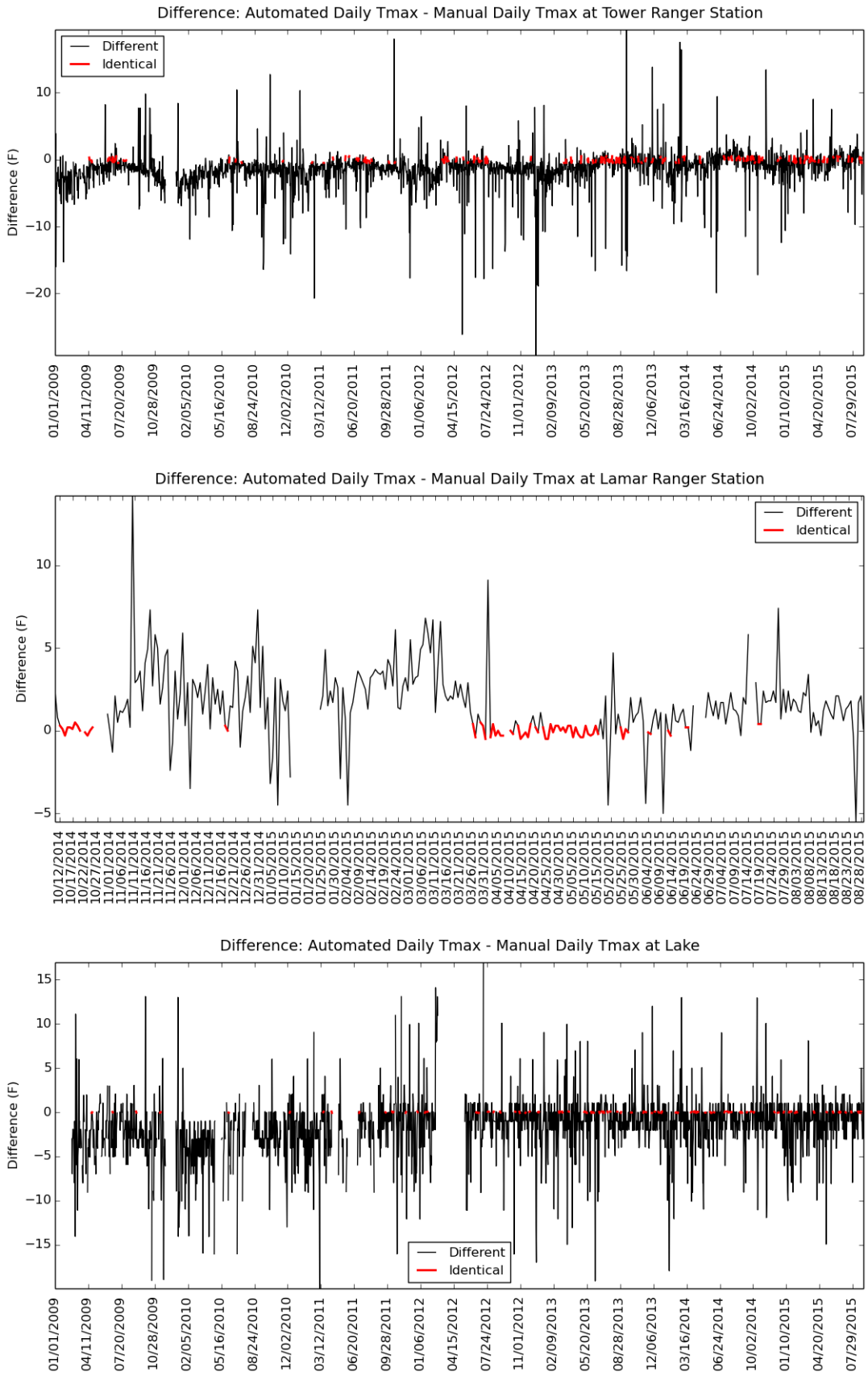


Figure 2: Patterns of data substitution at Tower, Lamar, and Lake.

Appendix 2 - Snake River Station (Yellowstone's South Gate) - An example illustrating the dangers of assuming that two weather stations are equivalent.

There are three weather stations at Yellowstone's South Gate: (1) A COOP, (2) A RAWS, and (3) a Snow-Telemetry (SNOTEL) station. More information on the SNOTEL network is here. The COOP weather station is located behind the ranger station. The RAWS and SNOTEL are co-located behind the corrals in the residental area [Figure 3]. The SNOTEL and RAWS thermometers are only 11m apart. The SNOTEL thermometer is mounted at 5.5m height above ground, while the RAWS and COOP are 4.4m and 2.4m above the ground, respectively.

Figure 3: Location of weather stations at Yellowstone's south gate. S = SNOTEL. R = RAWS, C = COOP.

The amount of difference among these stations varies greatly by season. For average daily maximum temperature (Tmax), there is good agreement among all three stations for most of the year, but the SNOTEL has monthly Tmax averages 5 - 7 degrees warmer during the winter (Figure 4). For monthly average daily minimum temperature (Tmin), the SNOTEL and the RAWS agree well all year. The COOP Tmin is cooler than the other two for much of the year, except during the seasons when most of the infilling from the RAWS is likely occurring (October - November and March - April) (Figure 4).

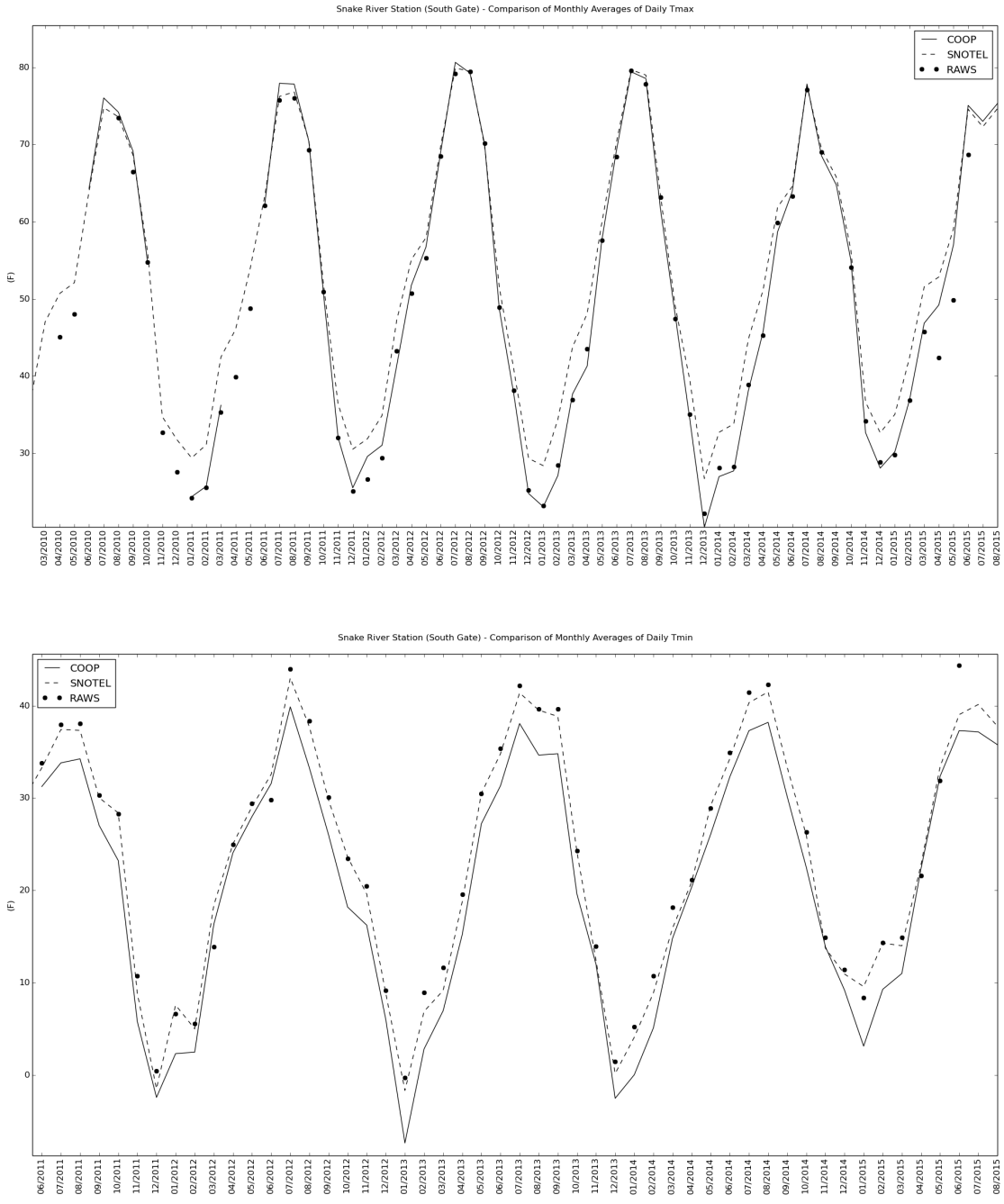


Figure 4: Monthly averages of average daily maximum temperature (Tmax) and average daily minimum temperature (Tmin) from three stations at Yellowstone' south gate.

It is impossible to offer a definite explanation for the difference between the COOP and the other stations because the amount of infilling and data adjustment to the COOP record is currently unknown (see above). In contrast, the difference between the SNOTEL and RAWS thermometer is due to the fact that the winter sun's lower angle on the horizon allows nearby trees to shade the RAWS while the SNOTEL thermometer, which is higher and further away from the trees, receives sunlight from 11 - 12 am. During the summer, when the sun is higher above the horizon (and the trees) at mid-day, both thermometers receive direct sunlight. Figure 5 shows that the treeline to the south of the RAWS (indicated by yellow arrow), is only 10m distant, while the SNOTEL thermometer is about 30m from

the southern tree line and more than 1m higher.

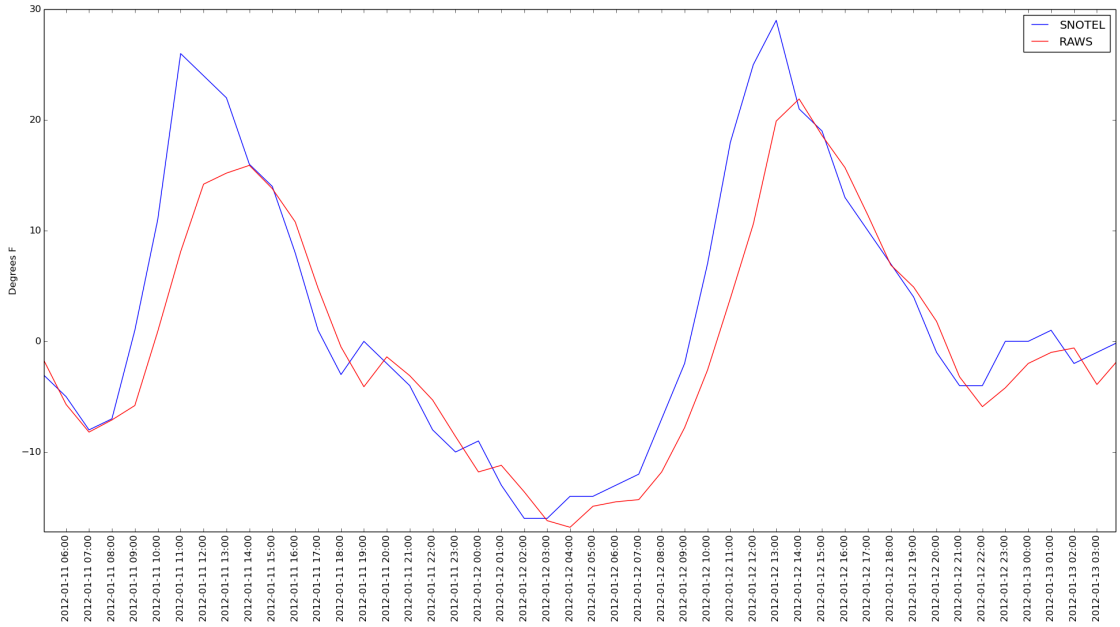
Figure 5: Photograph of the south entrance SNOTEL / RAWS installation. S = SNOTEL thermometer. R = RAWS thermometer. Yellow arrow = location of tree line to the south of the RAWS. The photographer is standing in front of the southerly tree that is closest to the SNOTEL thermometer.

The effect of the sun can be seen through an examination of hourly data from both thermometers. During the winter, the SNOTEL thermometer measures a sharp increase in temperature around 11am - 12pm, but the RAWS does not. During other times of the year, the two thermometers are in good agreement during an entire 24 hour cycle (Figure 6).

Fortunately, the data used to infill the COOP station has been taken from the RAWS thermometer (NWS personnel, personal communication) instead of the SNOTEL. If the SNOTEL thermometer had been the source, then the seasonally changing bias would have made the correction process more complex.

This example, in which two thermometers only 11m apart have recorded large and complexly different temperature regimes, should serve as a warning against the dangers of infilling data from different sources without documentation. Without detailed knowledge of the site, it might seem reasonable to assume that there could be very little difference between two such closely located instruments, but such an assumption is clearly not valid.

Winter Hourly Data - 48 hours



Summer Hourly Data - 48 Hours

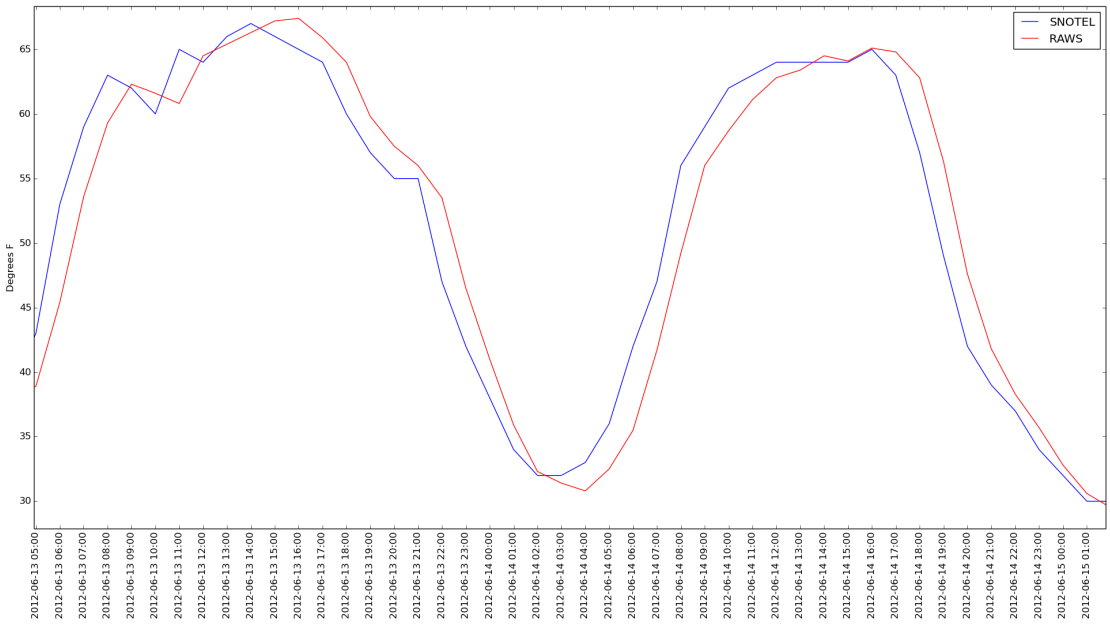


Figure 6: Comparison of diurnal fluctuations in temperature data from the SNOTEL vs. RAWS thermometers at Yellowstone's south gate.