Methodology for Developing SNOTEL Profiles

Last Revised: 27 May 2014

Christopher Daly, Keith Olson, Joseph Smith, Zlatko Dimcovic

PRISM Climate Group/NACSE

A major task in the development of the operational SNOTEL QC system was the creation of station profiles. Profiles are station-specific curves that define acceptable ranges of selected elements for each day of the year. Procedures for creating profiles are detailed in this document.

A suite of period-of-record profiles was developed for each station. The suite included ten profiles:

1. TMAX upper temperature
2. TMAX lower temperature
3. TMIN upper temperature
4. TMIN lower temperature
5. TRANGE upper temperature range (TMAX-TMIN)
6. Maximum IP (incremental precipitation) increase
7. Maximum ISWE (incremental snow water equivalent) increase
8. Maximum ISWE (incremental snow water equivalent) decrease
9. Maximum ISNWD (incremental snow depth) increase
10. Maximum ISNWD (incremental snow depth) decrease

Profiles for accumulated precipitation, snow water equivalent and snow depth were not developed, because the SNOTEL QC system focuses on daily incremental values, rather than accumulations. Whenever possible, TMAX, TMIN, and IP values that had already been subjected to the PRISM QC system (SSQC, SPQC) were used to construct the profiles. These QC steps removed a large amount of invalid data from the record.

Approach to Developing Profiles

Statement of the Problem

The problem presented in developing profiles was how to use extreme station observations of unknown quality to develop profiles for determining acceptable values. Since a profile is, by definition, an annual curve of daily extreme values, any of these values could be erroneous, and not appropriate for determining the validity of other values on the same day of the year (DOY). Initially, we attempted to follow the methods developed by the NRCS to develop profiles, but found that they apparently used subjective methods to determine *a priori* which values to believe and which to throw out, because the profiles appeared to accommodate the most extreme values available. It was not feasible for us to do this manually for ~900 stations, so a method that was not reliant on extreme values from any particular station was sought.

Approach Taken

After much experimentation with various statistics, we found that the deviation of an extreme value for a DOY from a smoothed mean extreme value curve, expressed in standard deviations (called the sd_distance), was a useful metric for determining validity. These thresholds were determined by: (1) calculating sd_distances of each daily value from each station's smoothed mean extreme value curve for that element; (2) identifying the most extreme sd_distances; (3) investigating these extreme cases and determining if they were valid or invalid; and (4) determine a universal sd_distance threshold for each profile that would distinguish valid from invalid data for all stations. These sd_distance thresholds are listed in Table 1.

In addition, a cap was placed on the 31-day average standard deviation (STDEV) for each temperature profile to minimize the possibility of several outliers acting together to artificially elevate the standard deviation (i.e., data scatter), thus allowing erroneous observations to slip under the threshold. This sometimes occurred when erroneous extreme values were measured on a string of consecutive days. These STDEV caps are also listed in Table 1.

Table 1. Profile sd_distance thresholds and STDEV caps.

Profile	Sd_distance Threshold	STDEV Cap	Comments
TMAX Lower	-4.7	5.0	STDEV cap reduced from 7.3 max legit value to minimize large winter windows
TMIN Lower	-4.13	5.5	STDEV cap reduced from 6.7 max legit value to minimize large winter windows
ISWE Decrease	-5.2	None	
ISNWD Decrease	-5.2	None	
TMAX Upper	3.75	3.85	STDEV cap reduced from 4.3 max legit value to minimize autumn "bumps"
TMIN Upper	4.4	2.5	STDEV cap reduced from 3.5 max legit value to minimize large winter windows
TRANGE Upper	5.1	4.1	STDEV cap reduced from 5.0 max legit value to minimize large winter windows
IP Increase	5.7	None	
ISWE Increase	6.0	None	
ISNWD Increase	7.0	None	

Period of Record Issues

The longer the period of record (POR), the more comprehensive (and hence extreme) a station profile becomes, because the station has had an opportunity to accumulate larger numbers of extreme events. This presents an issue when creating profiles using historical data from stations with varying periods of record. For example, the upper TMAX profile of a station with just one year of record will not be nearly as warm as one from a co-located station with thirty years of record. But observations from both stations should clearly be held to the same profile QC criteria, regardless of POR.

The largest changes in a station's profile were typically seen when the POR increased from less than 5 years to more than 10 years. We developed initial statistical relationships to estimate the changes needed to adjust a short-POR station's profile to approximate long-term expectations. These relationships were developed by regressing all SNOTEL PORs, regardless of location, against profile means and standard deviations. The results contained a lot of scatter, but useful relationships were found. Figure 1 shows the relationship between the annual average extreme maximum temperature and the station POR for all SNOTEL stations in the western conterminous US. The resulting power function was used to adjust the profile curves to approximate a 30-year POR. Other temperature adjustment functions had similar power formulations, with the extreme lowest TMAX and TMIN requiring the largest adjustments (Figure 2). The lower temperature adjustments are the result of rare but severe arctic outbreaks that cause more significant temperature deviations from the average extremes than major heat waves, for example.

POR adjustments for IP, ISWE and ISNWD profiles were found to be multiplicative, rather than additive, as for temperature. Figure 3 gives an example of a POR adjustment curve for extreme ISWE increase. In this case, the direct relationship between mean annual extreme ISWE and POR was not very strong or realistic, mainly because of the wide variation in ISWE increase during the year. However, we found the relationship between the percent standard deviation (PSTDEV) and POR to be a useful and realistic one. PSTDEV varies strongly with the mean; (high PSTDEVs are associated with small means, and low PSTDEVs are associated with large means). This allowed us to use the changes in PSTDEV with POR as surrogates for changes in the mean with POR. Mean profile adjustments (called A_{adj}) are given in Table 2.

We also found relationships between the STDEV and POR for temperature profiles (e.g., Figure 4). The STDEV typically decreases as POR increases, because as more data are accumulated, occasional extreme values occurring on various DOY's begin to "fill in" the profile to create a more consistent curve with less day-to-day variation. The effects of infilling on the STDEV become minimal beyond PORs of ten years, and our adjustments become zero at this point. Mean profile adjustments (called S_{adj}) are given in Table 3.

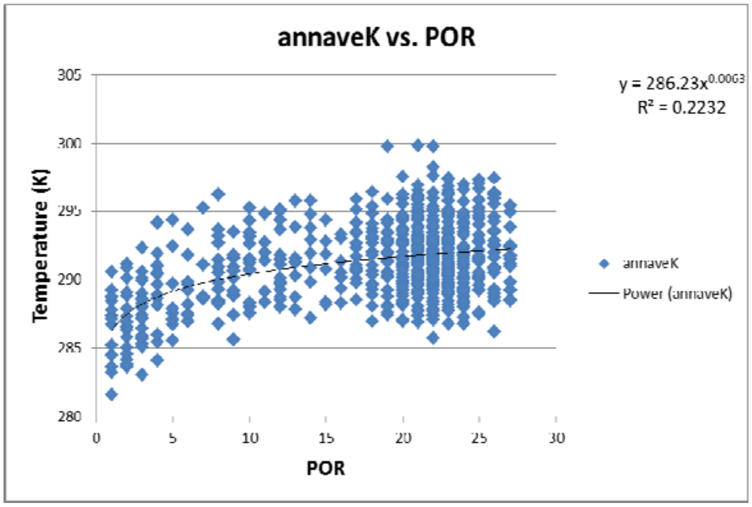

Figure 1. Relationship between annual average extreme maximum temperature and period of record. Data are for all SNOTEL stations in the western conterminous US. The resulting power function was used to adjust the upper tmax profile for stations with less than 30 years of record.

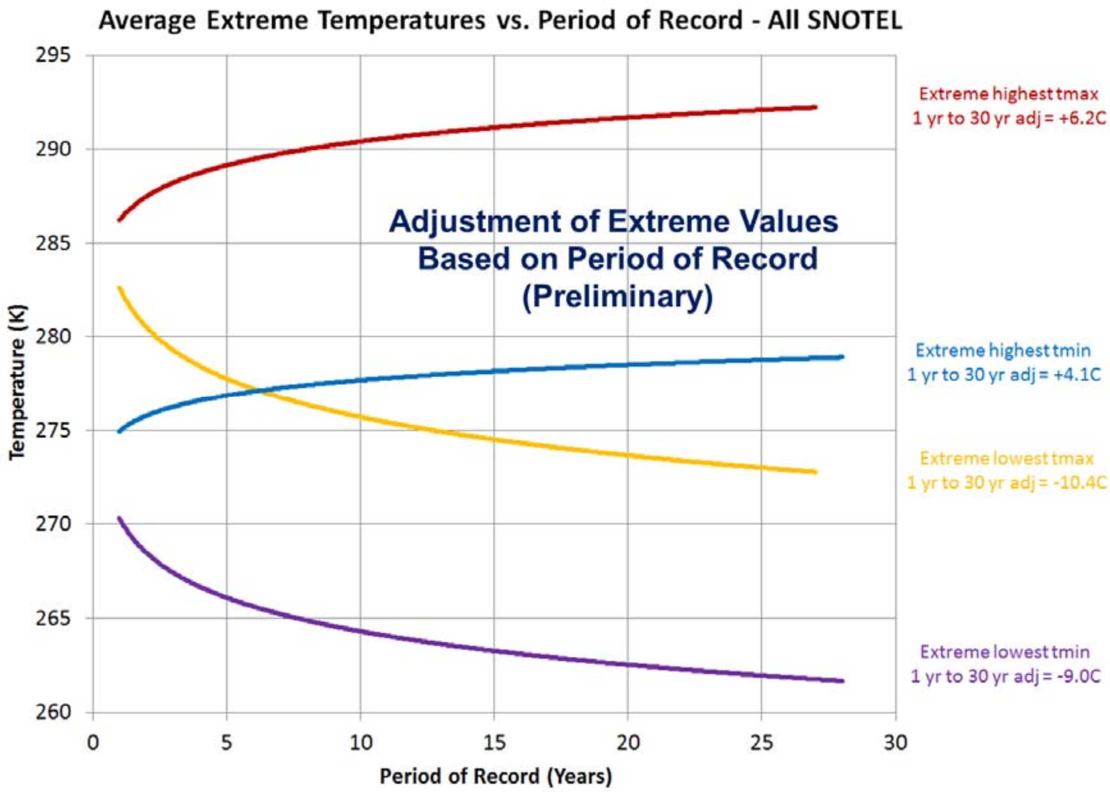

Figure 2. POR-adjustment curves for extreme temperature profiles.

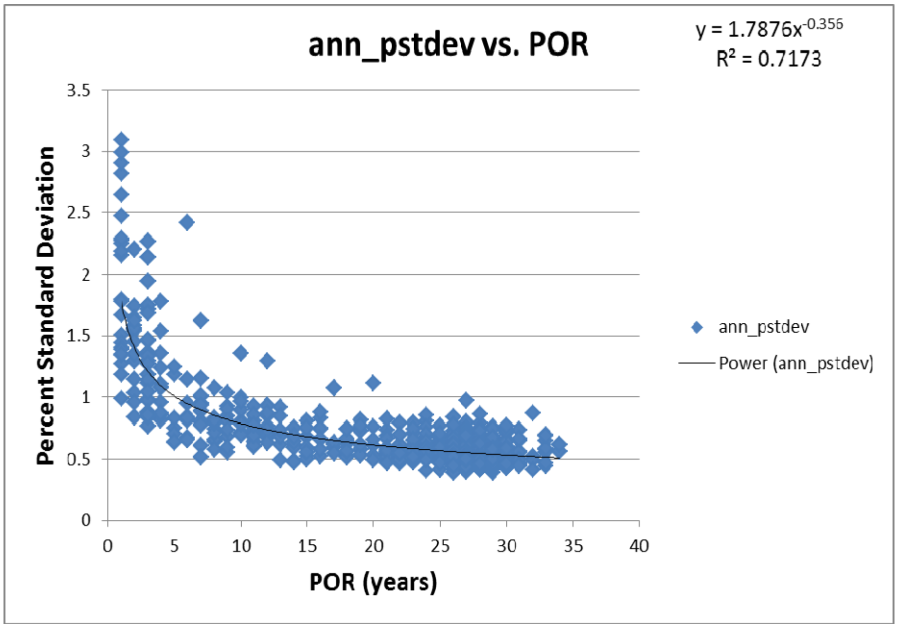

Figure 3. Relationship between the percent standard deviation (PSTDEV) of annual average extreme SWE increase and period of record. Data are for all SNOTEL stations in the western conterminous US. The resulting power function was used to adjust the mean of the ISWE maximum increase profile for stations with less than 30 years of record.

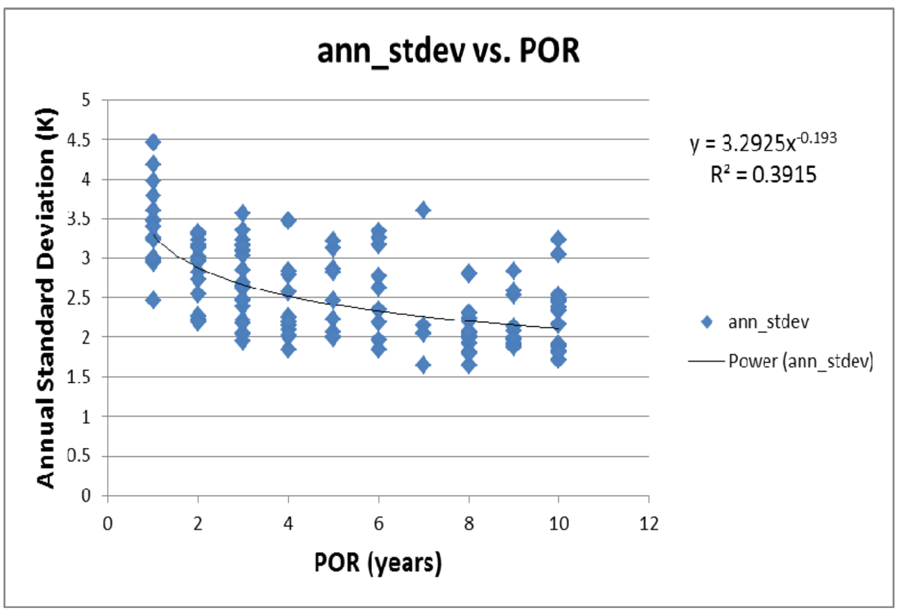

Figure 4. Relationship between the standard deviation (STDEV) of annual average extreme maximum temperature and period of record. Data are for all SNOTEL stations in the western conterminous US. The resulting power function was used to adjust the standard deviation on the upper tmax profile for stations with less than 10 years of record.

Table 2. Average POR adjustment (A_{adj}) equations and coefficients. A_{adj} is the adjustment factor, a is the coefficient, y is the exponent, base is the base POR above which no adjustment is needed, and POR is the period of record of the station in question. A_{adj} can be either multiplicative or additive, depending on the profile (see comments column).

Profile	A_{adj}	***a***	***y***	***base* (yrs)**	**Comments**
TMAX Lower	$a(base)^y - a(POR)^y$	282.61	-0.011	30	A_{adj} is additive; if $POR \geq base$, $A_{adj} = 0$
TMIN Lower	$a(base)^y - a(POR)^y$	270.3	-0.01	30	A_{adj} is additive; if $POR \geq base$, $A_{adj} = 0$
ISWE Decrease	$1 / (a(base)^y / a(POR)^y)$	1.2361	-0.239	35	A_{adj} is multiplicative; if $POR \geq base$, $A_{adj} = 1$
ISNWD Decrease	$1 / (a(base)^y / a(POR)^y)$	1.1169	-0.312	20	A_{adj} is multiplicative; if $POR \geq base$, $A_{adj} = 1$
TMAX Upper	$a(base)^y - a(POR)^y$	286.23	0.0063	30	A_{adj} is additive; if $POR \geq base$, $A_{adj} = 0$
TMIN Upper	$a(base)^y - a(POR)^y$	274.96	0.0043	30	A_{adj} is additive; if $POR \geq base$, $A_{adj} = 0$
TRANGE Upper	$a(base)^y - a(POR)^y$	13.052	0.1193	30	A_{adj} is additive; if $POR \geq base$, $A_{adj} = 0$
IP Increase	$1 / (a(base)^y / a(POR)^y)$	1.7334	-0.364	35	A_{adj} is multiplicative; if $POR \geq base$, $A_{adj} = 1$
ISWE Increase	$1 / (a(base)^y / a(POR)^y)$	1.7876	-0.356	35	A_{adj} is multiplicative; if $POR \geq base$, $A_{adj} = 1$
ISNWD Increase	$1 / (a(base)^y / a(POR)^y)$	2.3249	-0.461	20	A_{adj} is multiplicative; if $POR \geq base$, $A_{adj} = 1$

Table 3. STDEV POR adjustment power functions and coefficients. S_{adj} is the adjustment factor, a is the coefficient, y is the exponent, base is the base POR above which no adjustment is needed, and POR is the period of record of the station in question. S_{adj} is always multiplicative (see comments column).

Profile	S_{adj}	*a*	y	*base* (yrs)	Comments
TMAX Lower	$a(base)^y$ / $a(POR)^y$	4.1089	-0.17	10	S_{adj} is multiplicative; if $POR \geq base$, $S_{adj} = 1$
TMIN Lower	$a(base)^y$ / $a(POR)^y$	3.546	-0.117	10	S_{adj} is multiplicative; if $POR \geq base$, $S_{adj} = 1$
ISWE Decrease	NA	NA	NA	NA	NA
ISNWD Decrease	NA	NA	NA	NA	NA
TMAX Upper	$a(base)^y$ / $a(POR)^y$	3.2925	-0.193	10	S_{adj} is multiplicative; if $POR \geq base$, $S_{adj} = 1$
TMIN Upper	$a(base)^y$ / $a(POR)^y$	2.8753	-0.219	10	S_{adj} is multiplicative; if $POR \geq base$, $S_{adj} = 1$
TRANGE Upper	$a(base)^y$ / $a(POR)^y$	2.2864	-0.132	10	S_{adj} is multiplicative; if $POR \geq base$, $S_{adj} = 1$
IP Increase	NA	NA	NA	NA	NA
ISWE Increase	NA	NA	NA	NA	NA
ISNWD Increase	NA	NA	NA	NA	NA

Computational Steps for Developing Profiles

Below are the steps for developing profiles. Minimum values are used for TMAX and TMIN lower profiles, and ISWE and ISNWD decrease profiles. Maximum values are used for TMAX, TMIN, and TRANGE upper profiles, and IP, ISWE, and ISNWD increase profiles. Steps below are for lower/decrease profiles, with parenthetical alternatives for upper/increase profiles.

1. Find period-of-record (POR) minimum (maximum) values for each day of the year (DOY).

2. Apply a 31-day moving average to the time series of values in (1) to create a smooth curve.

3. Apply 5 passes of a 15-day moving average to the smoothed curve in (2) to create the smoothed_avg curve.

    a. Adjust the smoothed_avg curve for POR by calculating A_{adj} using the formulas in Table 2. Either add A_{adj} to the smoothed_avg curve or multiply A_{adj} by the smoothed_avg curve, depending on the profile, as instructed in Table 2 (comments column). POR is defined as the length of the actual data record, in years, of the DOY with the shortest data record.

4. Calculate a 31-day moving standard deviation (STDEV) for the values in (1).

    a. Adjust the 31-day moving STDEV for POR by calculating S_{adj} using the formulas and coefficients in Table 3. Multiply the STDEV for each DOY by S_{adj}.

    b. If the POR-adjusted 31-day moving STDEV for a DOY exceeds the standard deviation cap given above for that variable in Table 1, set that DOY's 31-day moving STDEV to the value of the cap for that profile.

5. Apply 5 passes of a 15-day moving average to the time series of 31-day moving STDEV values in (4) to create a smooth curve, called sd_smoothed.

6. To create the profile curve, start with the smoothed_avg curve in (3) and subtract (add) sd_smoothed in (5) multiplied by the sd_distance threshold given above. For example, if the sd_distance threshold is 4.0, the profile curve would be smoothed_avg – (+) (sd_smoothed * 4.0).

7. For each DOY, calculate the difference of each DOY's value in (1) from the smoothed curve in (3) in terms of standard deviations (termed sd_distance). This is found by dividing each day's difference by sd_smoothed for that DOY from (5).

    a. If the calculated sd_distance is more negative (positive) than the sd_distance threshold given above, the DOY's value is assumed to be invalid. Flag the value and choose the next most extreme value for that DOY from the database. Repeat the process, if necessary, until a value that does not violate the threshold is found for that DOY.

8. If the sd_distance for any DOY in (6) violated the threshold, go back and repeat steps 1-6 to create a new profile curve. Repeat this process until no DOY's value violates the threshold. For each profile, keep a running count of the number of times the profile generation process was invoked.

Plotting Profiles

Profiles were plotted for each station and element with DOY on the x axis and value on the y axis.

Values plotted:

- Observations from (1)
- Smoothed average of the observations from (3)
- Adjusted smoothed average from (3a)
- "Raw" unadjusted 31-day moving average STDEV from (4)
- Smoothed STDEV curve from (5)
- STDEV cap from Table 1, and used in (4b)
- Final profile curve from (6), on the final iteration (if more than one was needed)

Header information included:

- Line 1. Name of profile and station, e.g., "Lower Temperature Profile (tmax) @ 03E05S"
- Line 2. POR information, including start and end dates and POR, e.g., "Start: 1998-08-13 End: 2013-08-01 Years: 17"
    - Note that "Years" should correspond with the POR used in the profile calculation steps. The POR is defined as the length of the actual data record, in years, of the DOY with the shortest data record. For example, if all DOYs have a POR of 5 years but only one DOY has a POR of 3 years, the resulting POR is three years.
- Line 3. Base POR, and average and STDEV adjustment factors, e.g., "Average Base Years: 25, Adj: 5.64 STDEV Base Years 10, Adj: 0.87"

- Note that ISWE, IP, and ISNWD have no STDEV adjustment, so those values are blank. If the POR of the station is at or above the base POR for either the average or the STDEV, also leave the respective adjustment factors blank.

Example Profiles

First, we present TMAX upper profiles for two nearby stations in the western Oregon Cascades:

- Jump Off Joe (22E07S), 1067 m, 23-year temperature record
- Smith Ridge (22E11S), 1015 m, one-year temperature record

Both should have about the same profile, given their close proximity, but one year of data from Smith Ridge, which is essentially just the daily weather at the station, cannot be used as-is to determine acceptable extreme values. Using the factors discussed above, adjustments were made to approximate a profile with a POR of 30 years. Shown in Figure 5, Jump Off Joe has 23 years or record, so little change is seen between the average curve (pink) and the POR-adjusted average curve (blue). The running standard deviation curve (yellow) does not change beyond 10 years of record, so is coincident with the adjusted (green) curve. Based on an examination of extremely warm tmax events at SNOTELs across the West, the acceptable profile limit was set at 3.75 standard deviations from the adjusted average curve (Table 1). In addition, a hard cap on the standard deviation curve (red line) was established at 3.85C to ensure that several bad values on successive days did not artificially elevate the running standard deviation curve, allowing other extreme values to get under the profile limit (Table 1). The final profile limit is 20-25°C in winter and 40-45°C in summer. These are relatively liberal limits, but the profiles are designed to be accepting of new extreme values.

Shown in Figure 6, adjustments were made to the Smith Ridge profile to approximate a profile with a POR of 30 years. Smith Ridge had only 1 full year of record for all DOYs at the time of this analysis, so there was a large upward shift (+6.2C) between the average curve (pink) and the POR-adjusted average curve (blue). The running standard deviation curve (yellow), which often exceeded the 3.75 cap, was lowered substantially, as shown by the (green) curve. The profile limit is about 25C in winter and 40-45C in summer, similar to those for Jump Off Joe, which we would expect to be a reasonable surrogate for Smith Ridge. But as more data are collected, Smith Ridge will develop its own "personality," and likely differ somewhat from Jump Off Joe.

For comparison, Figures 7 and 8 show TMIN lower profiles and adjustments for Jump Off Joe and Smith Ridge, respectively.

A complete suite of profiles for Jump Off Joe SNOTEL is provided in Appendix A.

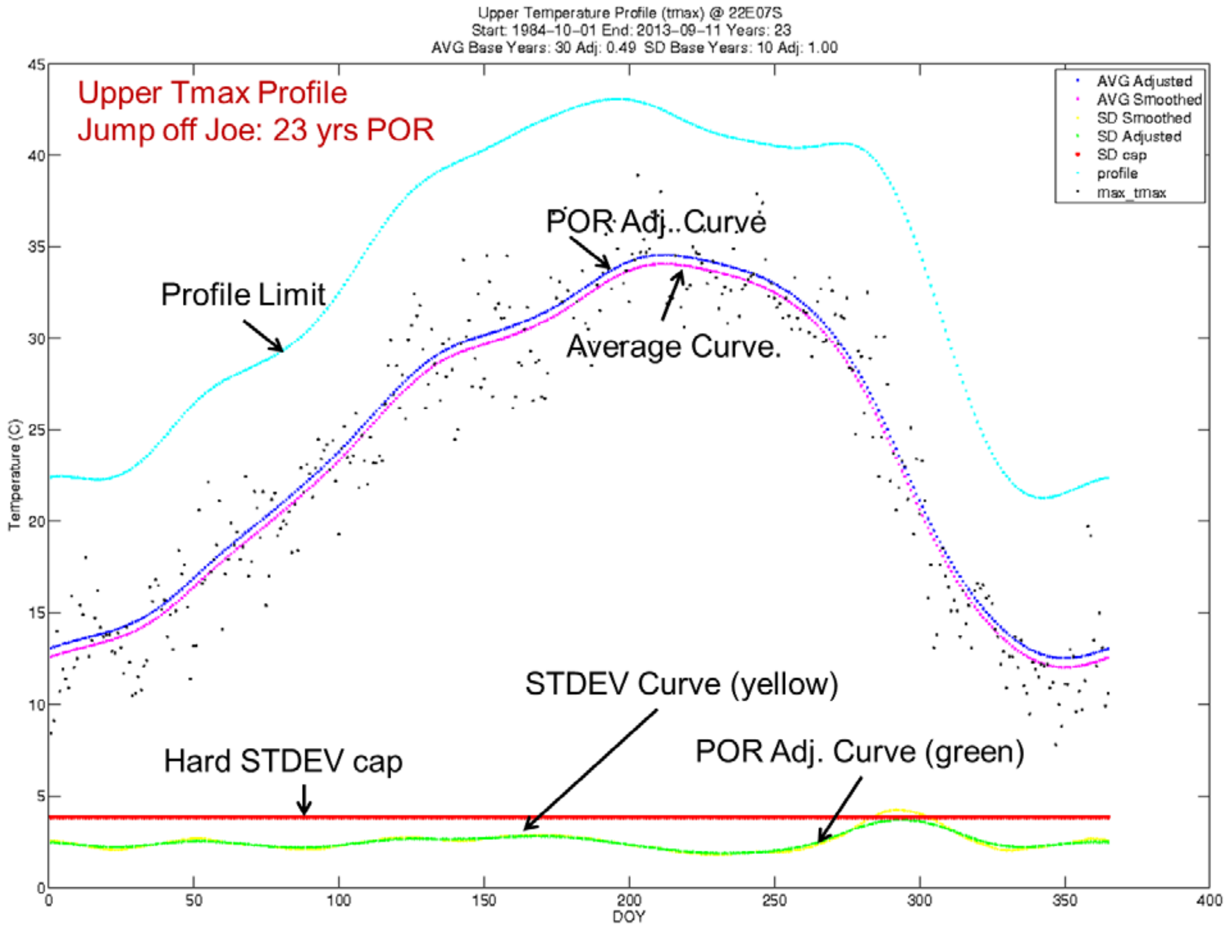

Figure 5. TMAX upper profile for Jump Off Joe SNOTEL, adjusted to approximate a 30-year POR. Given that the station has 23 years of TMAX record, the adjustment process has only minor effects on profile characteristics. (POR is based on the DOY with the shortest record.)

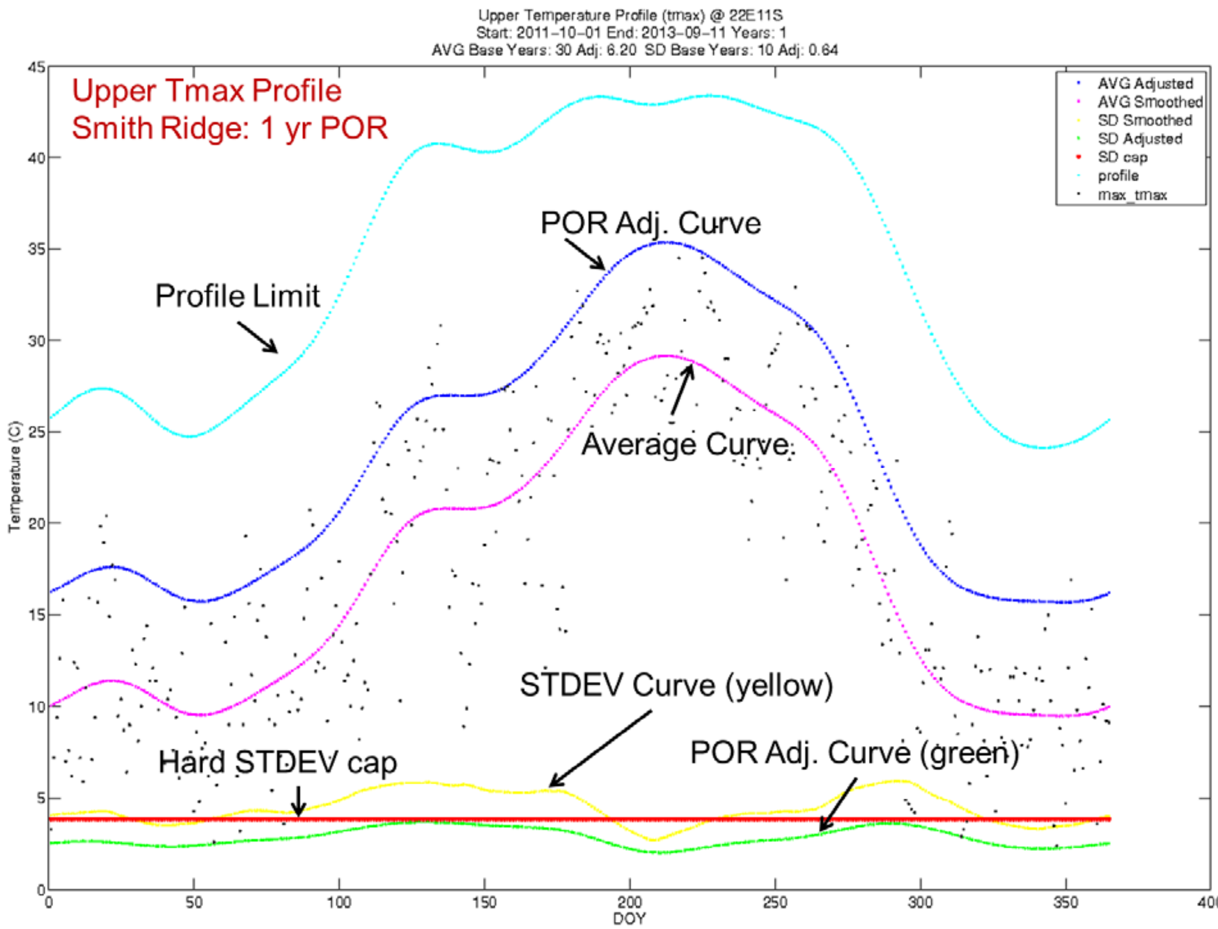

Figure 6. TMAX upper profile for Smith Ridge SNOTEL, adjusted to approximate a 30-year POR. Given that the station has only one year of temperature record, the adjustment process has major effects on profile characteristics, shifting the profile thresholds to much warmer values, and lowering the STDEV curve. (POR is based on the DOY with the shortest record.)

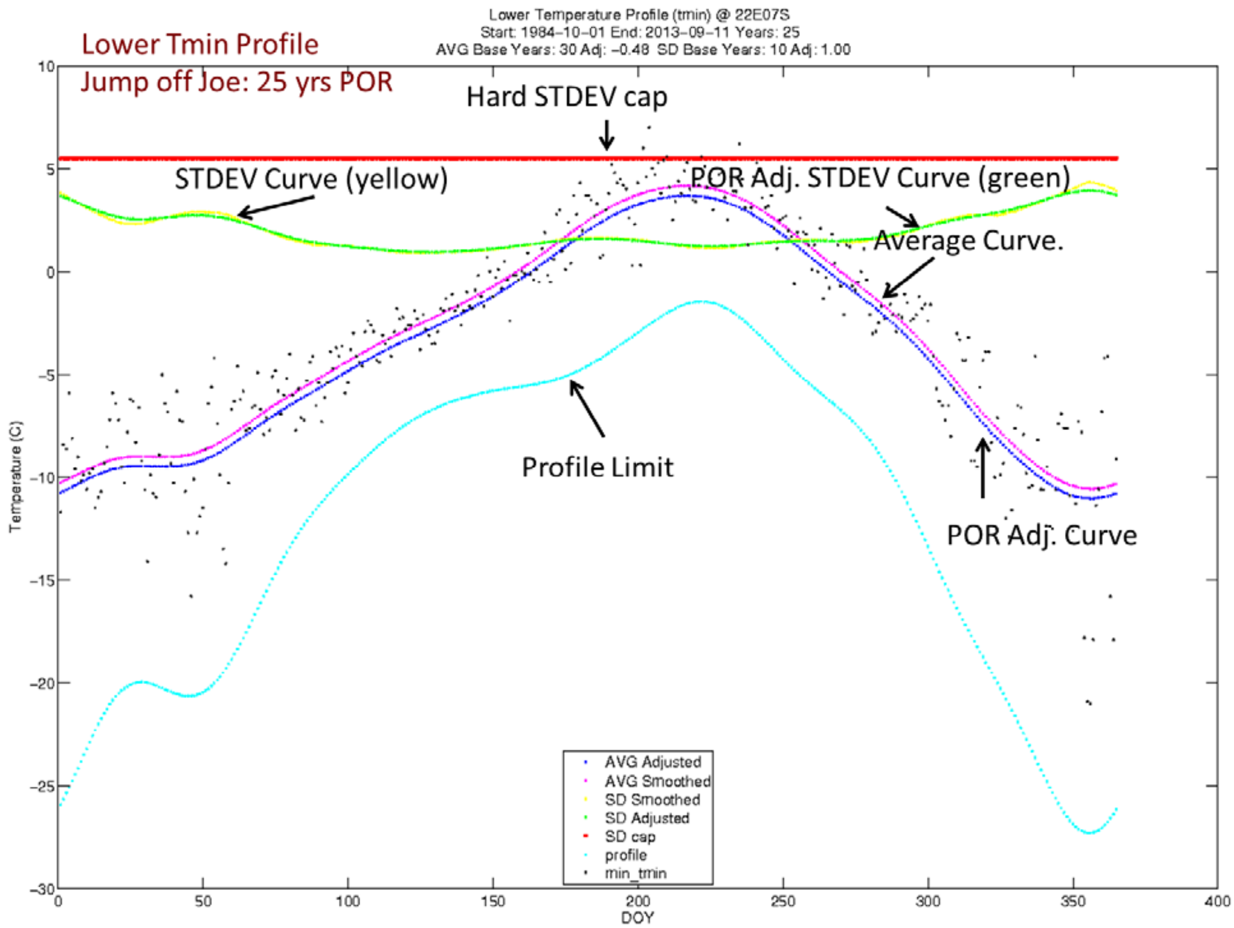

Figure 7. TMIN lower profile for Jump Off Joe SNOTEL, adjusted to approximate a 30-year POR. Given that the station has 25 years of temperature record, the adjustment process has only minor effects on profile characteristics. (POR is based on the DOY with the shortest record.)

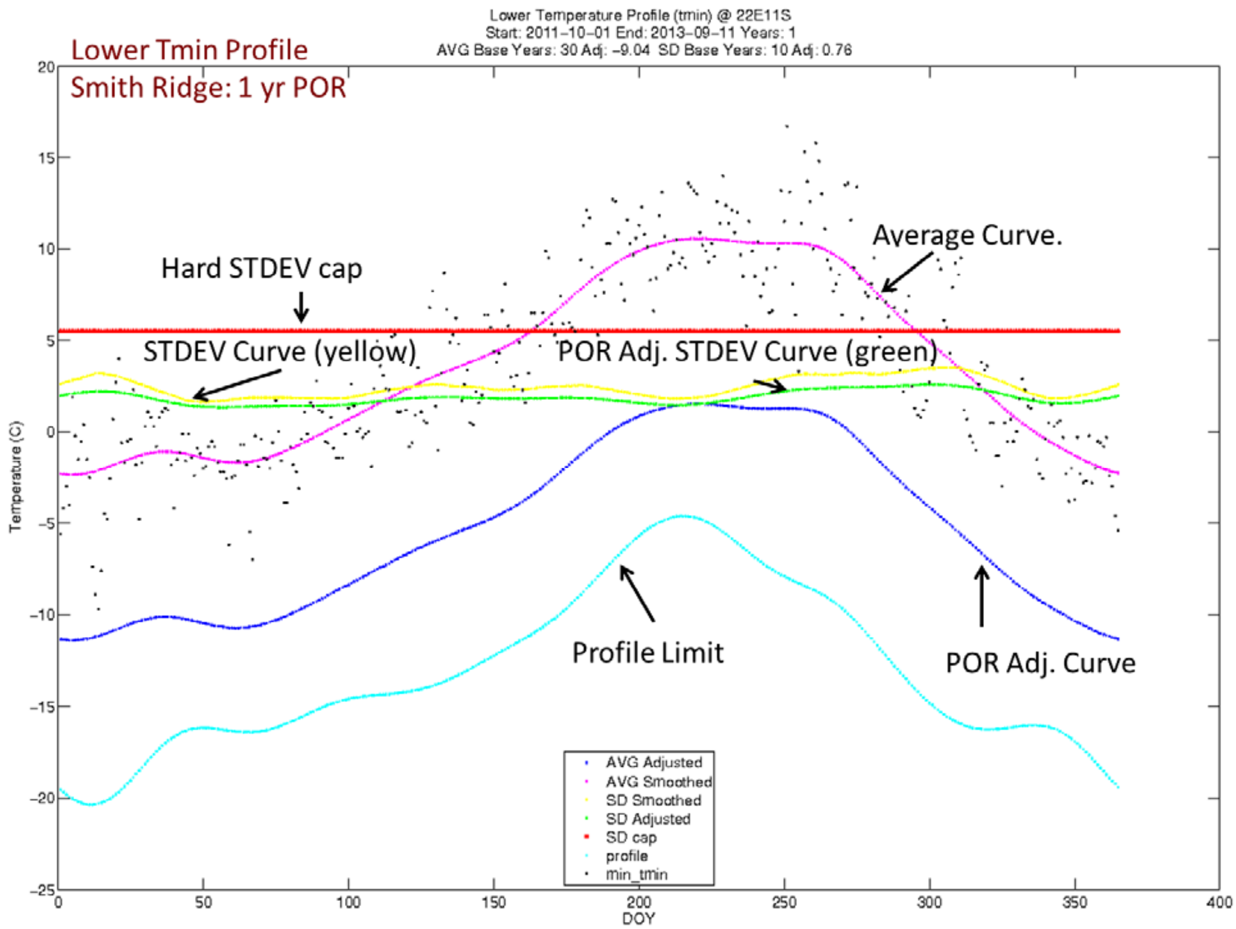

Figure 8. TMIN lower profile for Smith Ridge SNOTEL, adjusted to approximate a 30-year POR. Given that the station has only one year of temperature record, the adjustment process has major effects on profile characteristics, shifting the profile thresholds to much colder values, and lowering the STDEV curve. (POR is based on the DOY with the shortest record.)

Future Work

The profiles presented here are the result of a first attempt to develop an automated process for creating station profiles. Below we discuss future improvements and updates.

Improved profile curves: Profile curves for IP, ISWE and ISNWD tend to be too responsive to historical extremes at the high of the distribution. It may be possible to create a temporal smoothing function that has a high effective wavelength at high values and a low effective wavelength at low values. Calculation of wavelength as a proportion of the mean value may be a viable method.

Relationships among profiles for different elements: At present, profiles for each element are developed independently of each other. Therefore, for a given DOY, it is theoretically possible, for example, for a TMAX lower profile value to be lower than a TMIN lower profile. Consistency checks should be made to address this possibility.

Accumulation profiles: At this point, profiles have been developed for daily incremental changes in observed elements. As discussed as future work in the MSQC documentation, there are situations where it would be advantageous to have profiles for accumulated SWE, SNWD, and possibly precipitation. This need arises when an observation changes to an incorrect value, and maintains that value for two or more days in sequence. Incremental QC would likely catch the initial large change to the erroneous value, but may not catch the subsequent days in the sequence, because the incremental changes are small or zero. See the future work section of the MSQC documentation for a detailed example.

POR adjustments: The initial POR adjustment functions developed here can be refined in the future. Potential improvements include:

- Developing the adjustment functions by progressively sampling fewer and fewer years of data from long-POR stations, and re-calculating profile statistics at those stations. This could be done by randomly omitting years of record, as in a Monte Carlo experiment, or by progressively shortening the POR by omitting early years first. Comparing the functions developed from the two methods could provide information on trends in the magnitude of extreme values over time. The initial methods developed here implicitly incorporate temporal trends, because nearly all long-period stations were established relatively early, while short-period stations were established recently.
- Developing adjustment functions for individual DOYs, and creating adjustment "curves" that vary across the year, so that adjustments can vary seasonally. A related issue is that

the current adjustment procedures for ISWE and ISNWD are multiplicative, and thus cannot change the span of consecutive days in summer when zero snow occurs. (Multiplying by zero still gives zero). It is likely that the span of zero days should contract as the POR increases, because of rare snowfalls that occur during the normally snow-free period.

Periodic Updates: Profiles should be updated as new data are collected. Given that the process is largely automated, updates can probably be performed on an annual basis. The thresholds used to divide valid from invalid data should also be re-visited, but on a less frequent basis.

Appendix A

Complete profile suite for Jump Off Joe SNOTEL, Oregon Cascades (22E07S, site #552)

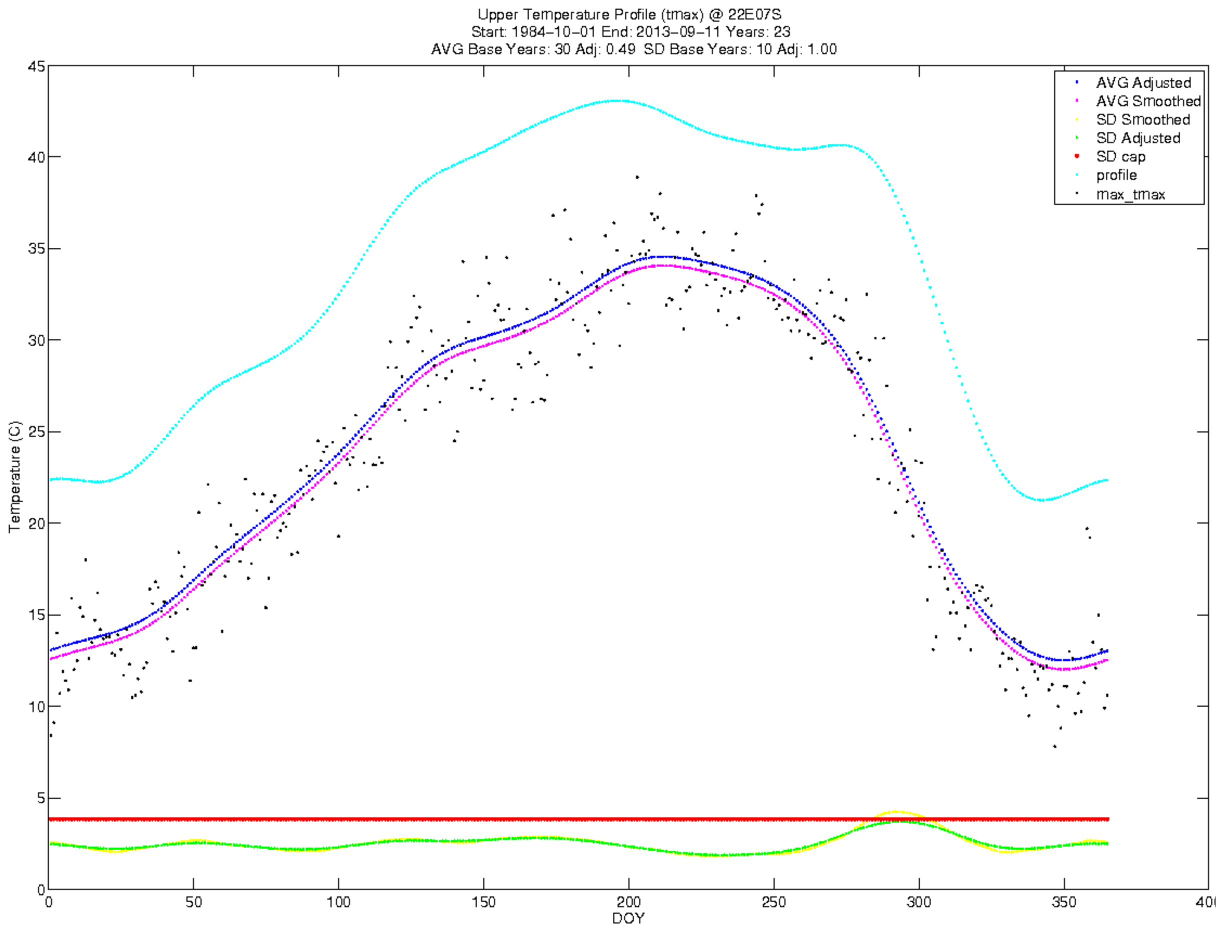

Figure A1. Jump Off Joe TMAX upper temperature profile.

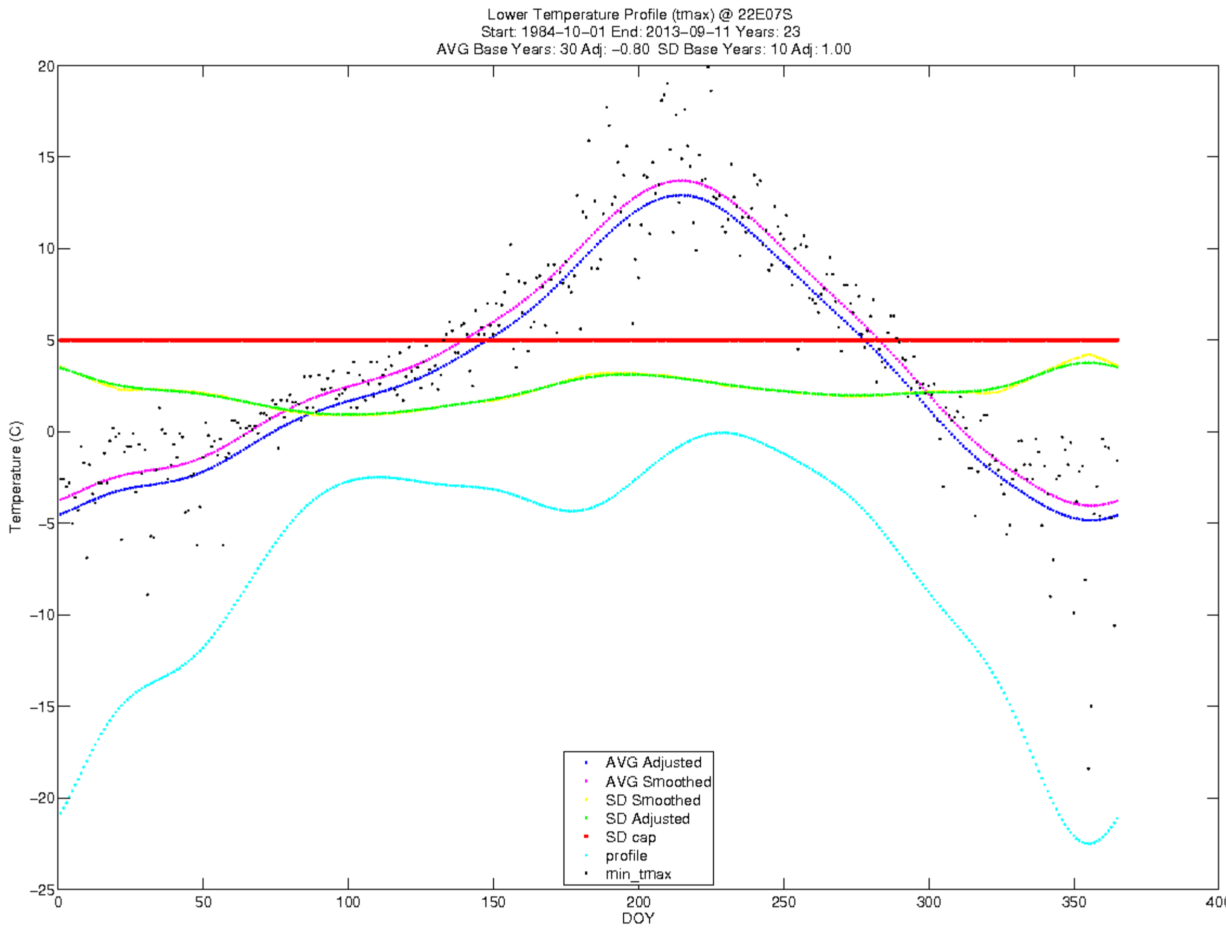

Figure A2. Jump Off Joe TMAX lower temperature profile.

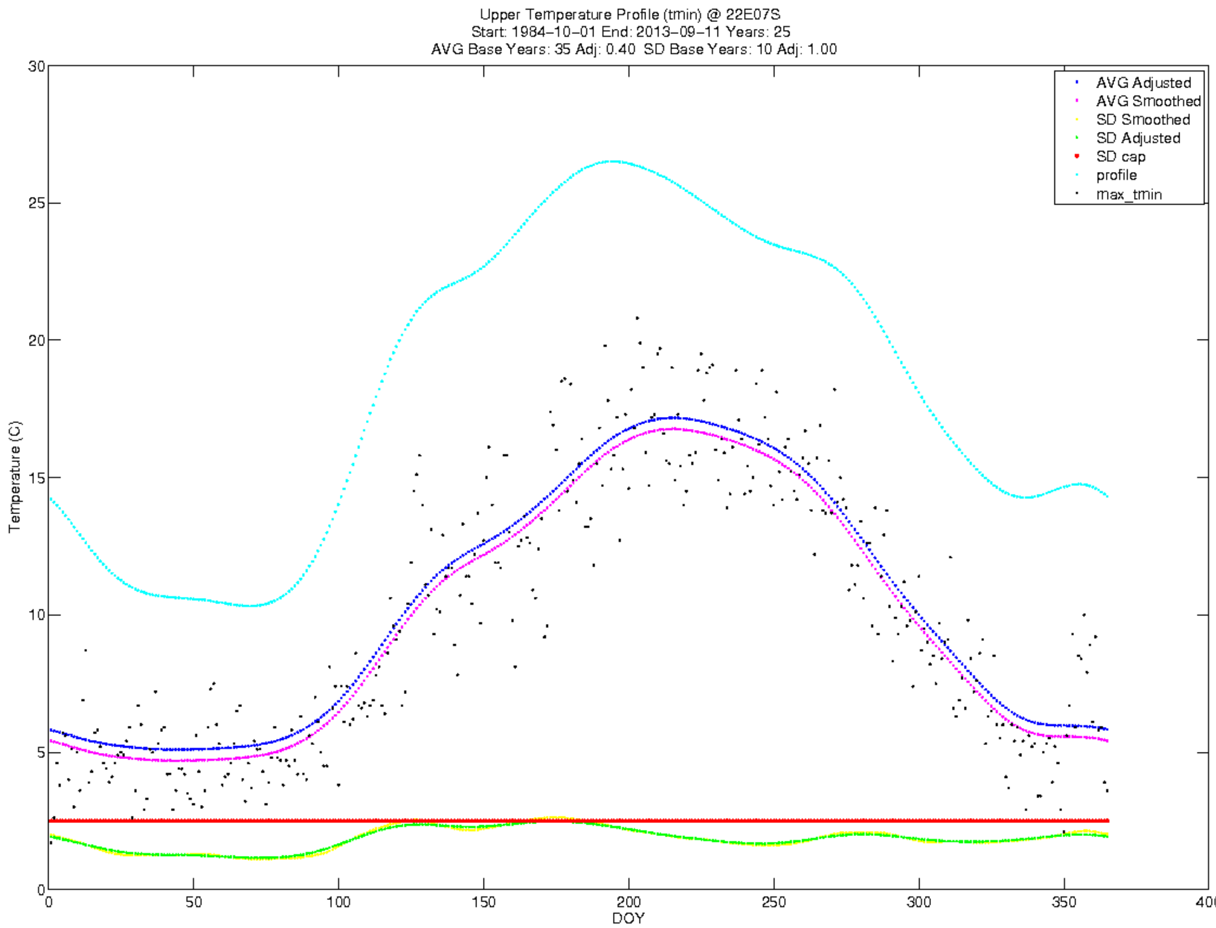

Figure A3. Jump Off Joe TMIN upper temperature profile.

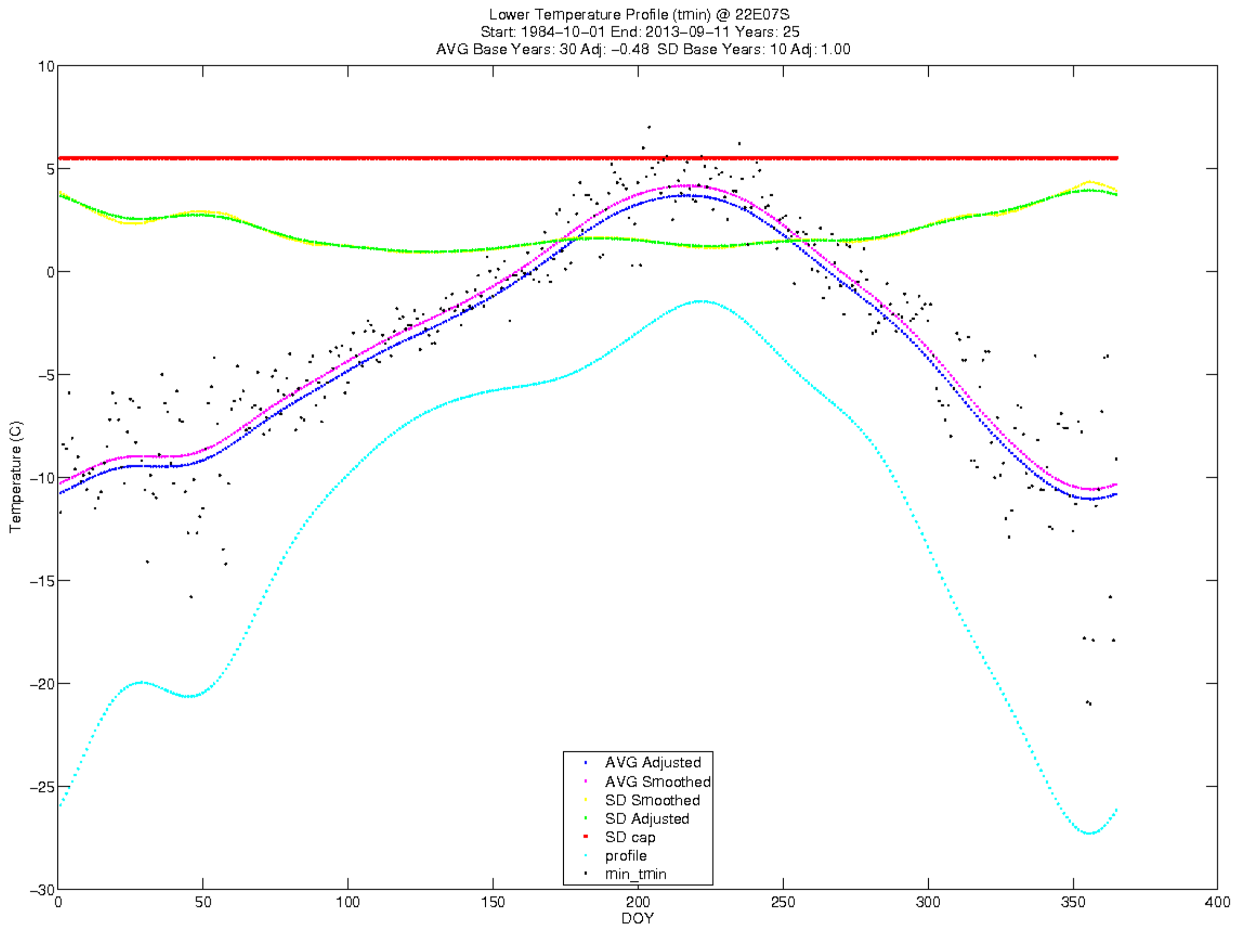

Figure A4. Jump Off Joe TMIN lower temperature profile.

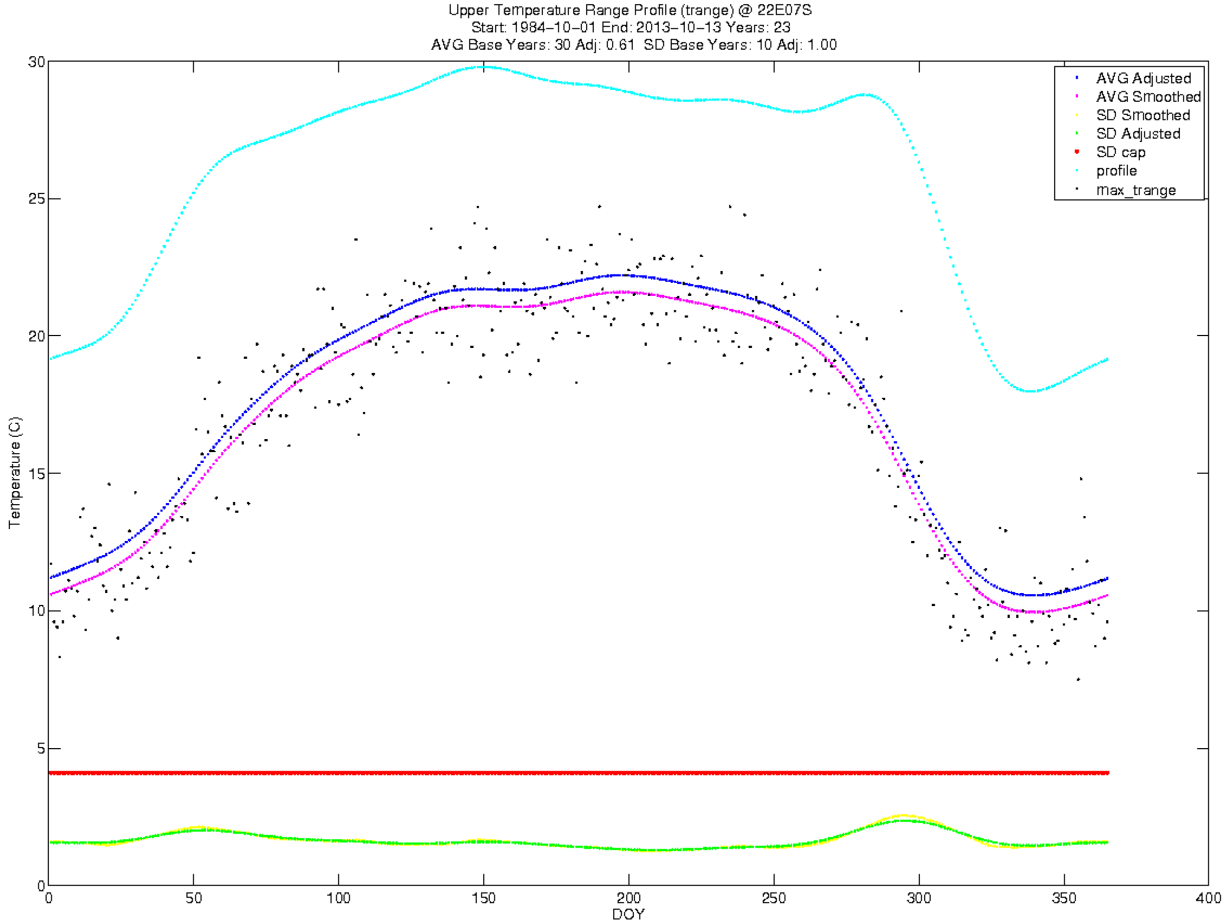

Figure A5. Jump Off Joe TRANGE (TMAX-TMIN) upper profile.

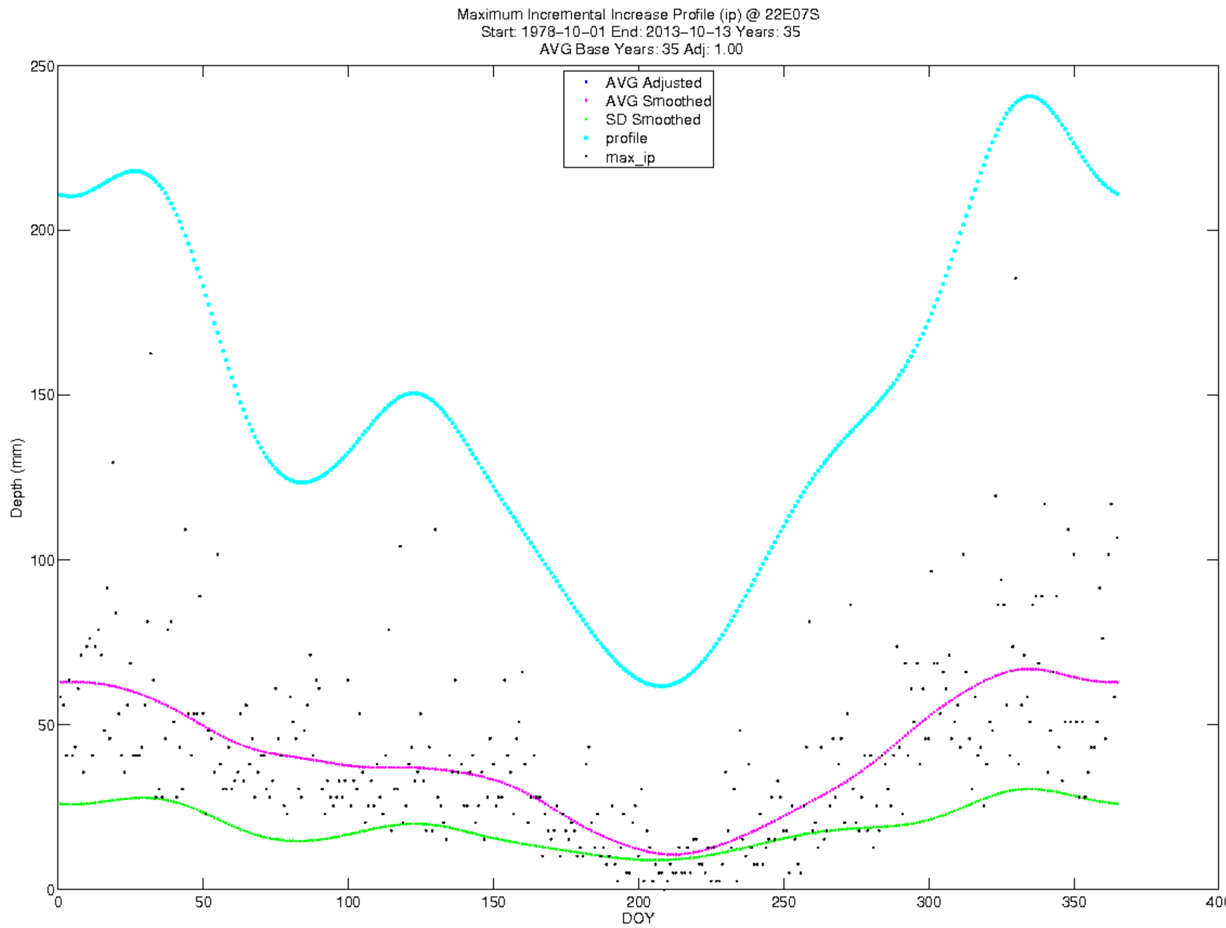

Figure A6. Jump Off Joe IP (incremental precipitation) increase profile.

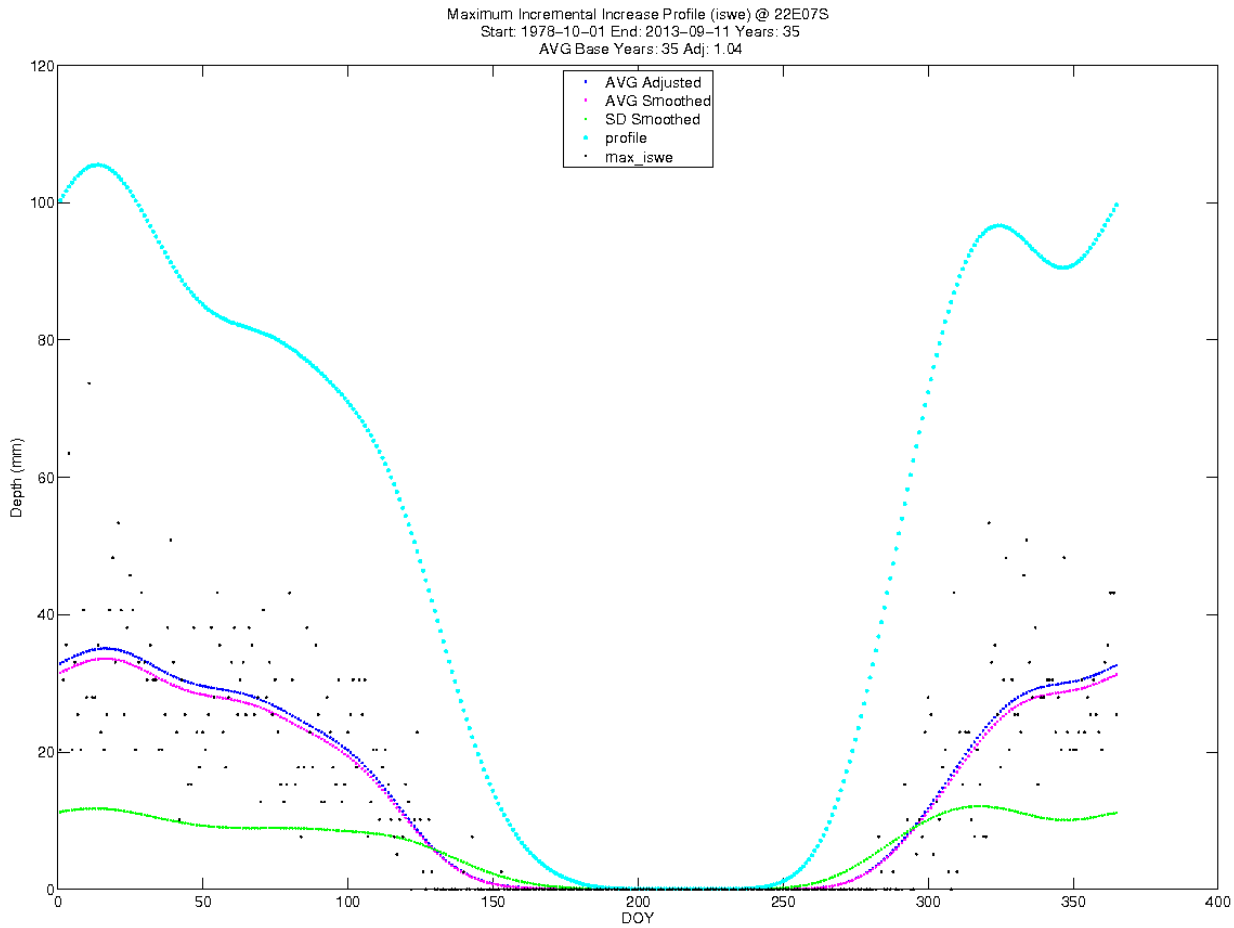

Figure A7. Jump Off Joe ISWE (incremental snow water equivalent) increase profile.

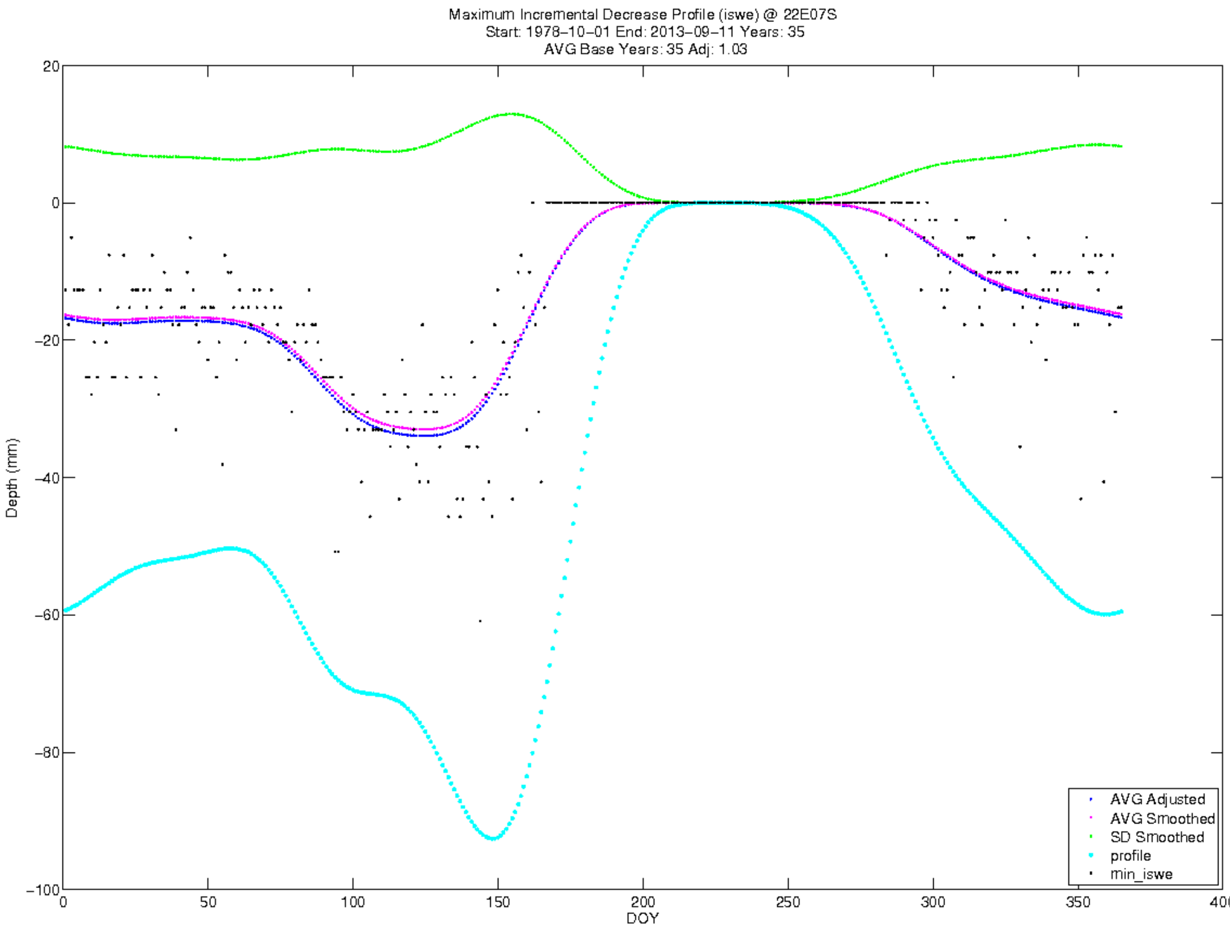

Figure A8. Jump Off Joe ISWE (incremental snow water equivalent) decrease profile.

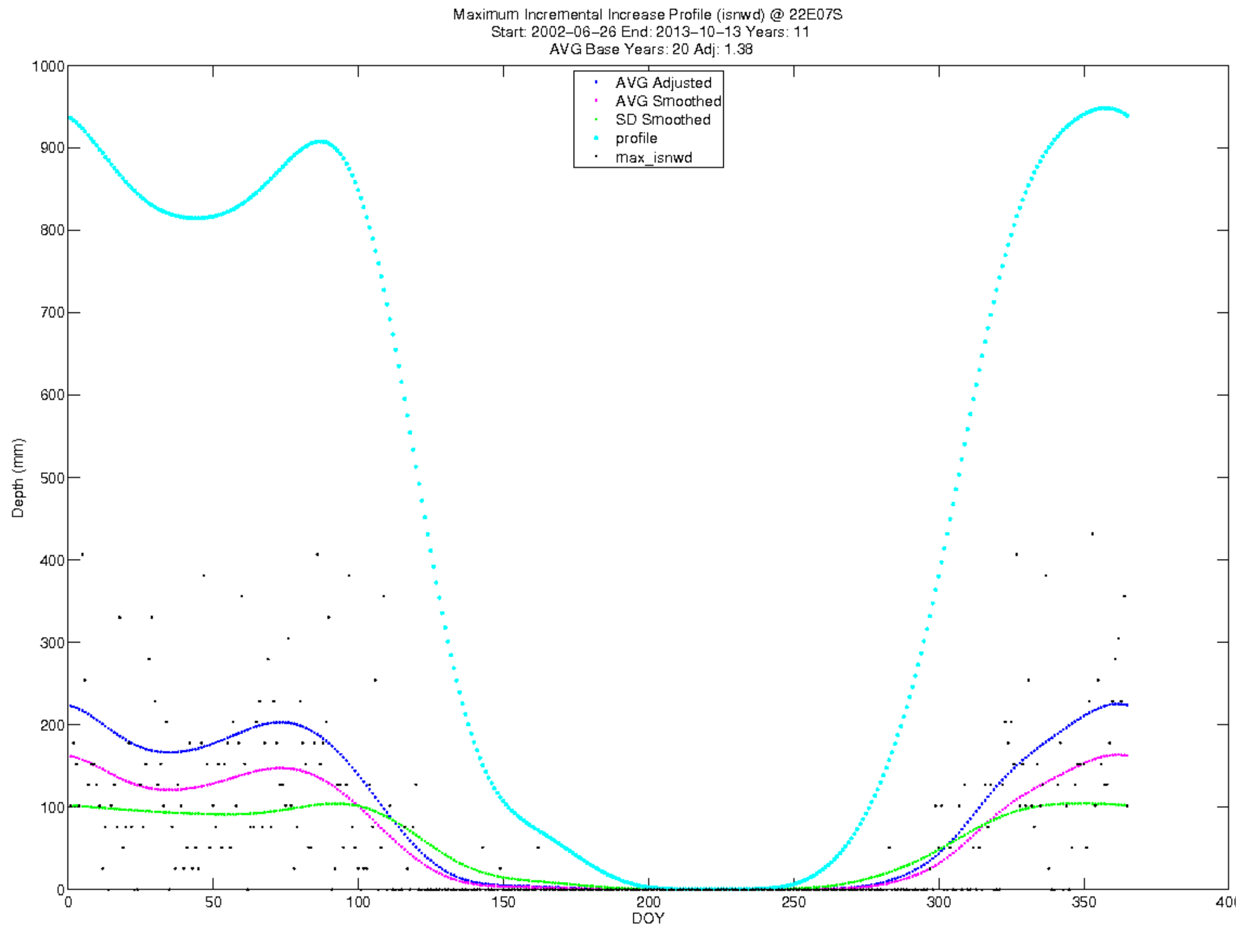

Figure A9. Jump Off Joe ISWND (incremental snow depth) increase profile.

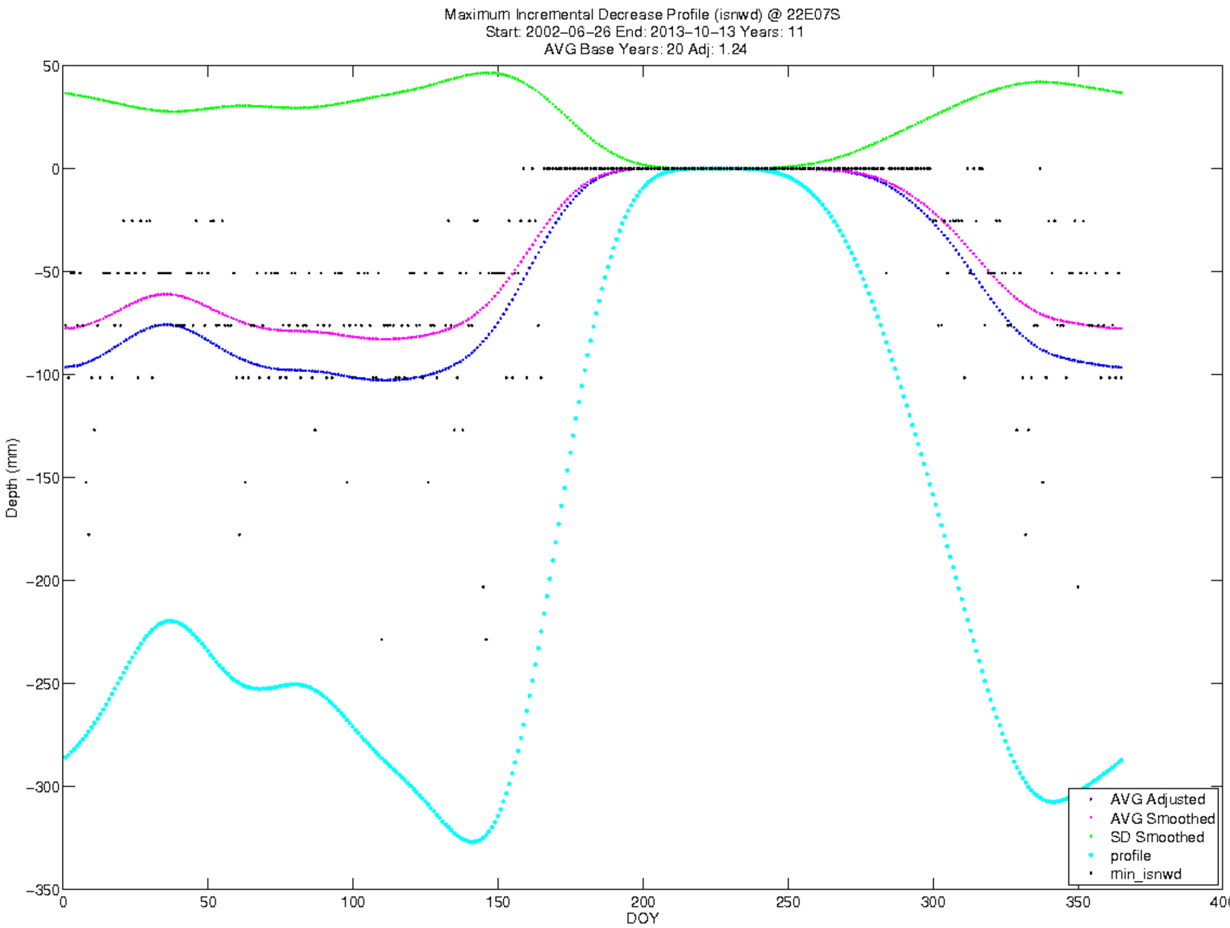

Figure A10. Jump Off Joe ISWND (incremental snow depth) decrease profile.