Plynlimon Automatic Weather Station (AWS) and stream gauge data editing notes
J.W. Kirchner 08 April 2021, with minor update 07 January 2024

The files of primary interest are as follows (all are tab-delimited ASCII text):

"Plynlimon compiled AWS and hourly Q.txt" contains:

- a) hourly average discharge in mm per hour for the four Wye gauges (Iago, Gwy, Cyff, and Cefn Brwyn) and the six Severn gauges (Upper Hore, Hore, Upper Hafren, Hafren, Tanllwyth, and Severn). These columns are labeled "<gauge>_avgQ_mm.hr", where <gauge> is the stream gauge name.

- b) estimates of hourly catchment-averaged precipitation rates for each of those catchments. These columns are labeled "P_<gauge(s)>_mm.hr", where <gauge(s)> are the names of one or more stream gauges to which these precipitation rates apply. Infilled precipitation records (see item d below for more detail on this) from the following AWS sites were used for each stream gauge:
    - Upper Hafren gauge: Carreg Wen
    - Lower Hafren, Upper Hore, Lower Hore, and Severn gauges: Mean(Carreg Wen, Tanllwyth)
    - Iago and Gwy gauges: Gwy/Iago if available, otherwise Eisteddfa Gurig
    - Cyff gauge: Mean(Eisteddfa Gurig, Cefn Brwyn) if available, otherwise Gwy/Iago
    - Wye (Cefn Brwyn) gauge: Mean(Eisteddfa Gurig, Cefn Brwyn, Gwy/Iago)

- c) composites of hourly AWS data for the five main AWS locations (Carreg Wen, Tanllwyth, Eisteddfa Gurig, Gwy/Iago, and Cefn Brwyn), combining all measurements from co-located or nearby AWS's, and averaging when multiple sensors are simultaneously available (wind direction is vector-averaged; all other sensors are arithmetically averaged). These columns are labeled "<site>_combined_<sensor>", where <site> is one of the five AWS sites enumerated above:
  Carreg – Carreg Wen
  Tan – Tanllwyth
  Eistg – Eisteddfa Gurig
  Gwy_Iago – Gwy/Iago
  Cefn – Cefn Brwyn
  and <sensor> is one of the following:
  SN – solar radiation in W/m^2
  RN – net radiation in W/m^2
  TD – temperature depression (difference between wet and dry bulb temp's) in °C
  T – air temperature (dry bulb) in °C
  WS – wind speed in m/s
  WD – wind direction in degrees (vector average)
  RF – tipping bucket precipitation in mm/hr
  VPD – vapor pressure deficit in kPa (calculated from T and TD)

- d) in-filled composite hourly AWS data for the five main AWS locations. These columns are labeled "<site>_infilled_<sensor>", where <site> and <sensor> are as described above. These in-filled columns create the most complete possible record by in-filling missing data using other sites. Those alternate sites were chosen in order of the strength of the pairwise correlations between their rainfall rates and rainfall rates at the target site during overlapping time steps (the precedence table can be found in the notes below from 24 November 2020).

- e) an "all sites" composite, averaging together all of the five main AWS locations for which data are available for any given hour. These columns are labeled "AllSitesAWS_<sensor>, where <sensor> is as described above.

"All sites discharge edited 1973-2010.txt" contains 15-minute discharge data in mm per hour and cubic meters per second for four Wye gauges (Iago, Gwy, Cyff, and Cefn Brwyn) and six Severn gauges (Upper Hore, Hore, Upper Hafren, Hafren, Tanllwyth, and Severn). Original unedited data as supplied by CEH have the suffix "RDA" (Raw Data Archive). Columns without the suffix "RDA" are edited data as described in the notes from 22 January 2012 below. Data editing consisted almost entirely of removing suspect points, and corrections for obvious factor-of-10 conversion errors.

"All_AWS_stacked_RDA_master_archive.txt" contains the archive of raw, unedited AWS data received from CEH.

"All_AWS_stacked_edited_v06.txt" contains the archive of AWS data that was edited as described in the notes below dated 28 February 2014. Data editing consisted almost entirely of removing suspect points, correcting clock errors, and correcting registration errors (wrong time, date, or site). The edited records can be linked to the raw data archive using the "Original_sort_order" column, if users want to see the editing that was applied.

The notes below describe the data cleaning and compilation that was performed. There are three main sets of notes:

- **discharge data cleaning notes dates 22 January 2012,**
- **AWS data cleaning notes dated 28 February 2014, and**
- **a final set of AWS data editing notes dated 24 November 2020 that describes how the records for individual stations were combined to site averages and how missing data were in-filled.**

Plynlimon stream gauge discharge data cleaning notes
J.W. Kirchner 22 January 2012

I worked from files supplied by Mark Robinson in the late 1990's, in 2009, and again in 2012. The AWS data run from 1975 through 2010, depending on the site.

The original time series as supplied by CEH have been retained in the data set, in columns whose names end with "RDA" (for Raw Data Archive). Column headers for the edited time series do not carry the RDA designation. Because I started with working files that I compiled in the late 1990's, I may already have partially edited some periods before 2000. I have not verified whether the RDA columns are fully consistent with the CEH database before 2000; these may reflect my earlier edits of the CEH data (including corrections for timestamp errors and so forth).

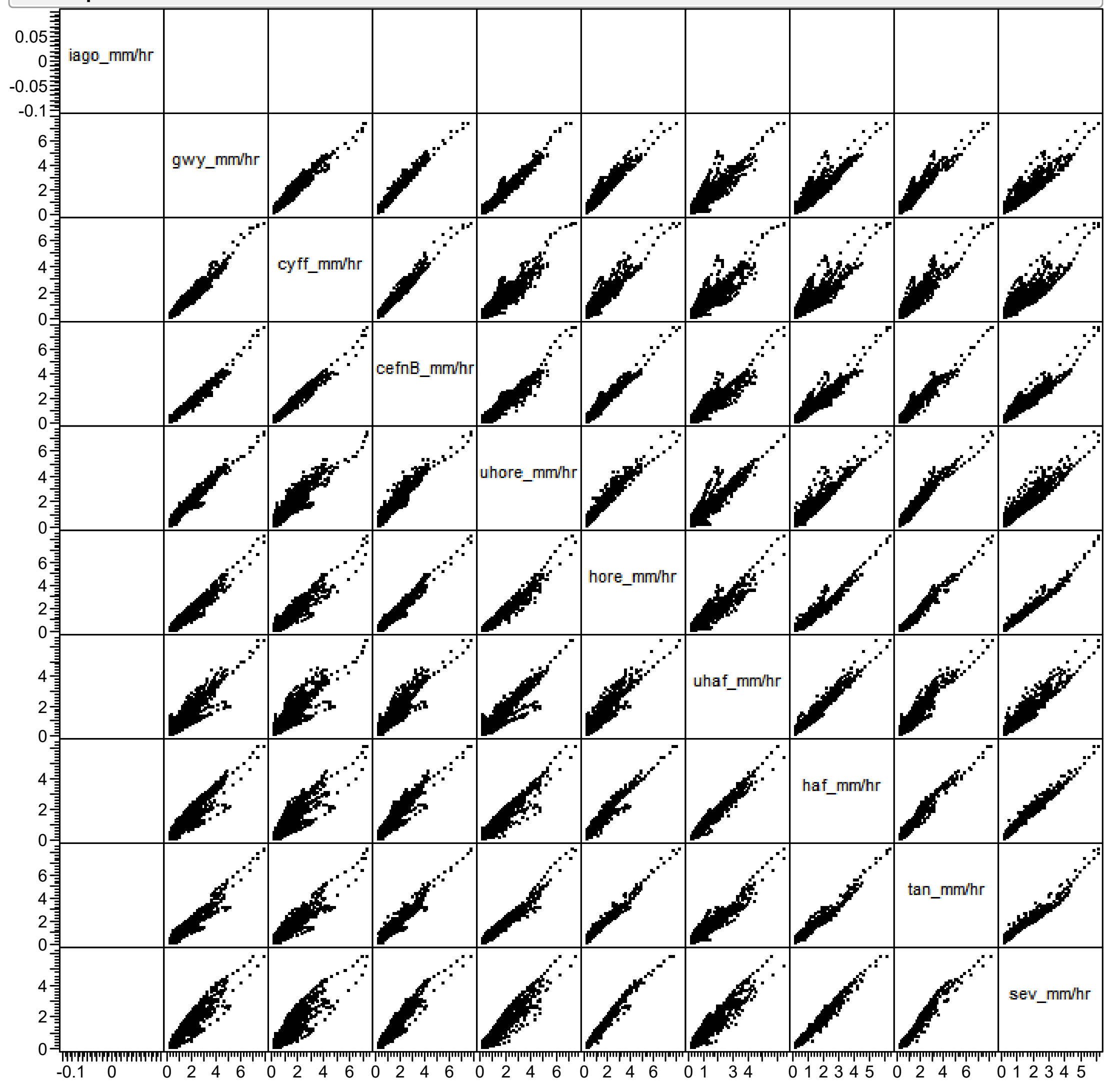
I visually inspected the entire 15-minute record for each site, plotted as quarterly (3-month) strip charts on log axes (these are about 1500 plots, each with up to 8800 points or so). I looked at the discharge time series in mm/hr units to simplify comparisons among sites. All sites were visually scanned together, permitting eyeball identification of gross anomalies. All values of exact zero were culled, because they did not plausibly correspond to true zero flows.

I also inspected annual scatterplot matrices. An example is shown below for 2007 (this one shows good correspondence among the sites).

Multivariate yr=2007

No multivariate correlations due to all rows missing

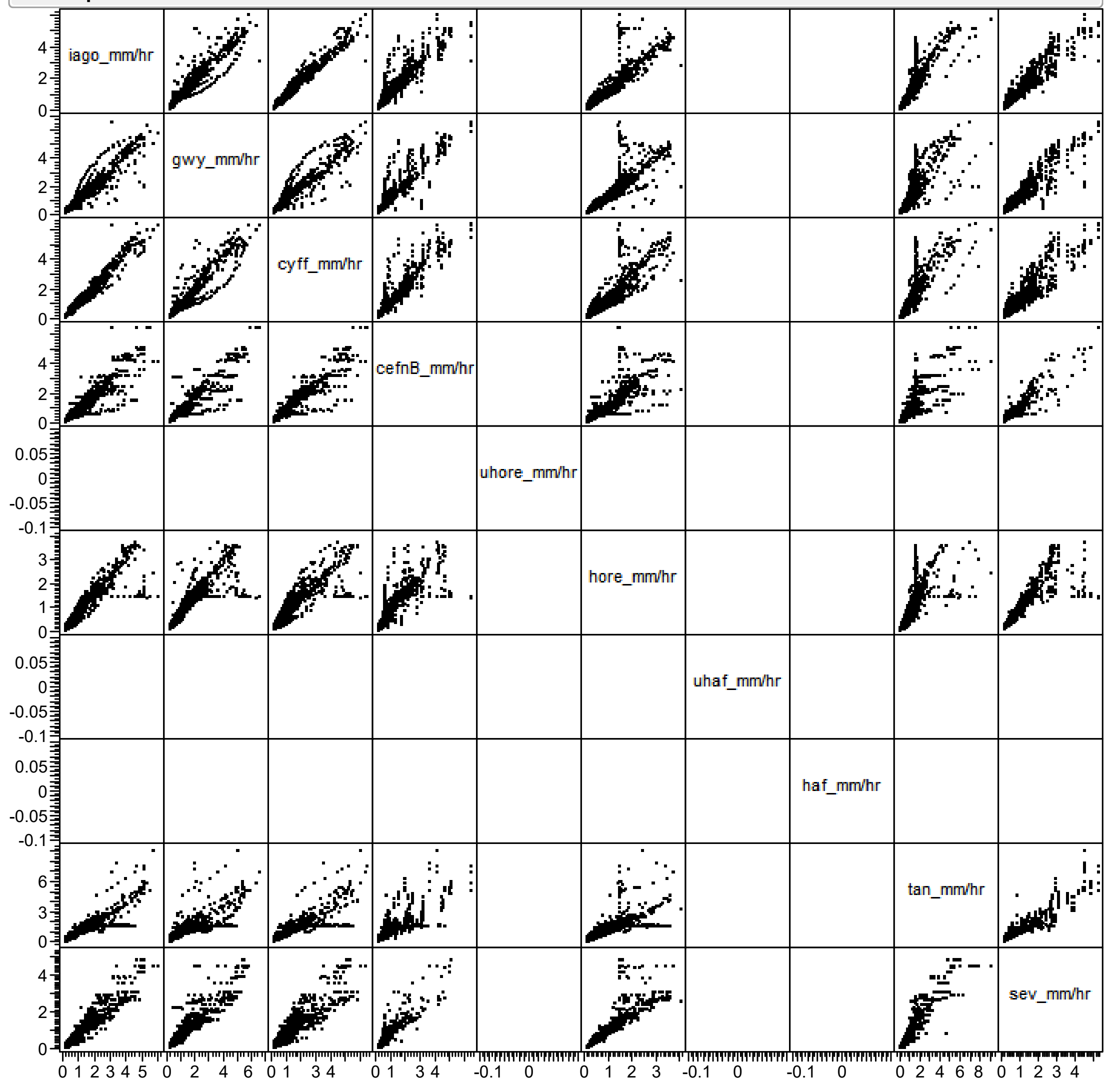
Scatterplot Matrix



Multivariate yr=1973

No multivariate correlations due to all rows missing

Scatterplot Matrix

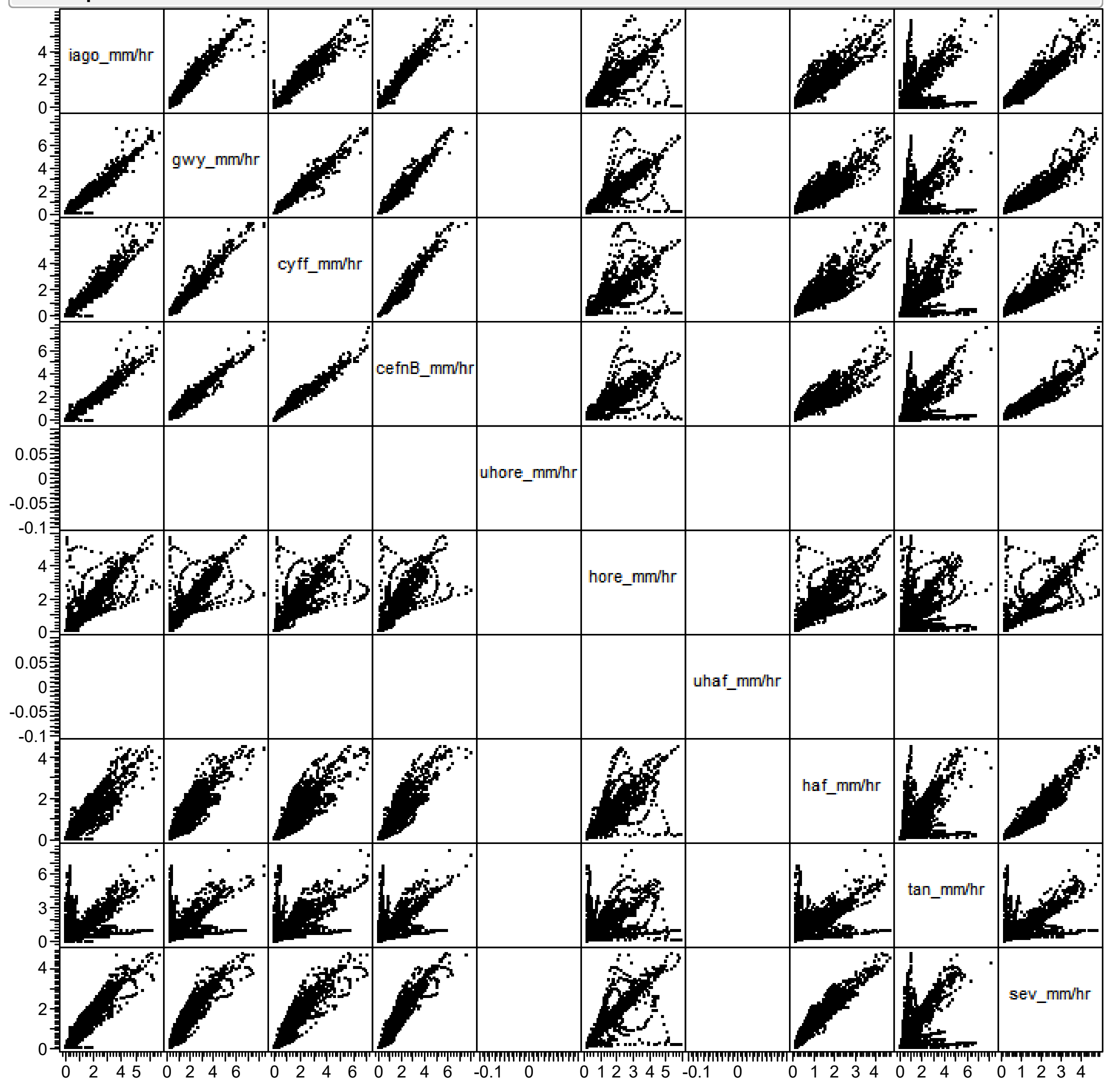


By contrast, the scatterplot above shows the noisy records from 1973. Note that some obvious problems are immediately evident, like the "stripes" of fixed values at Hore and Tanllwyth.

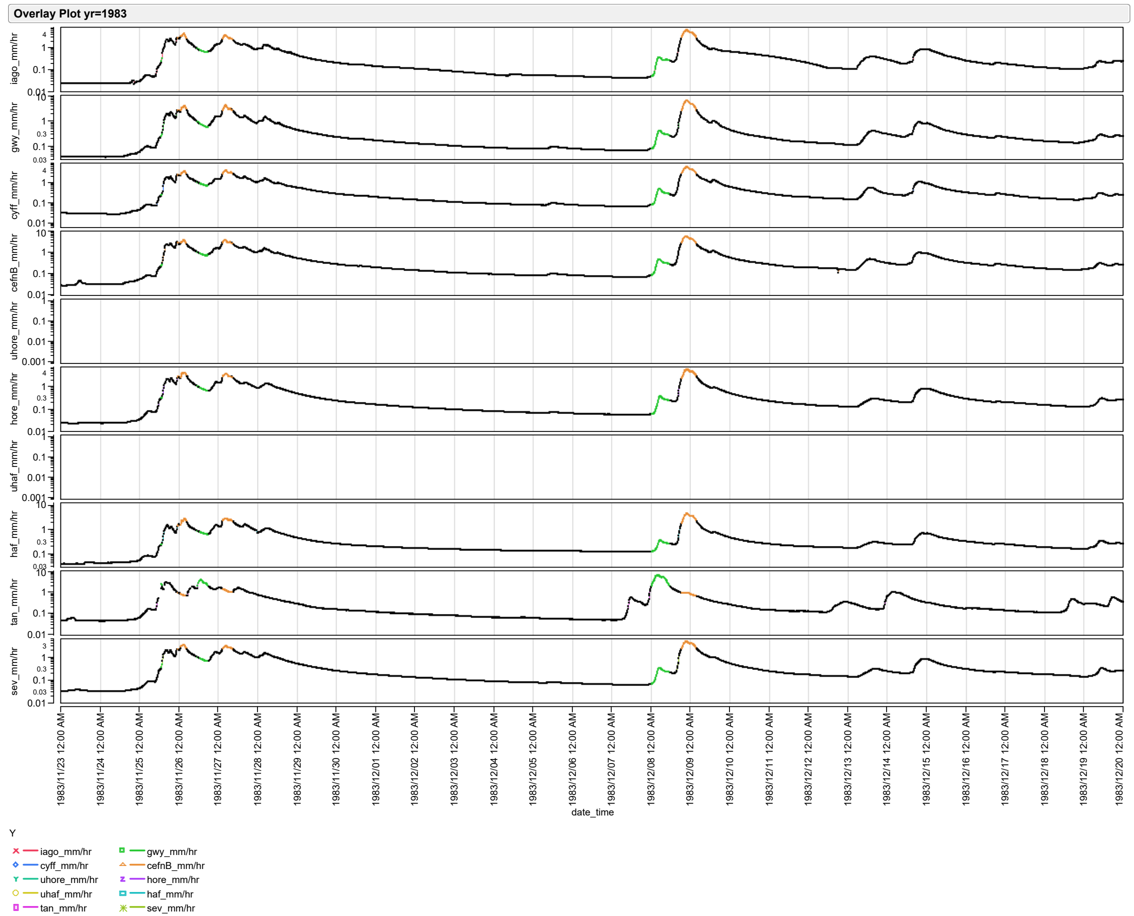
Multivariate yr=1983

No multivariate correlations due to all rows missing

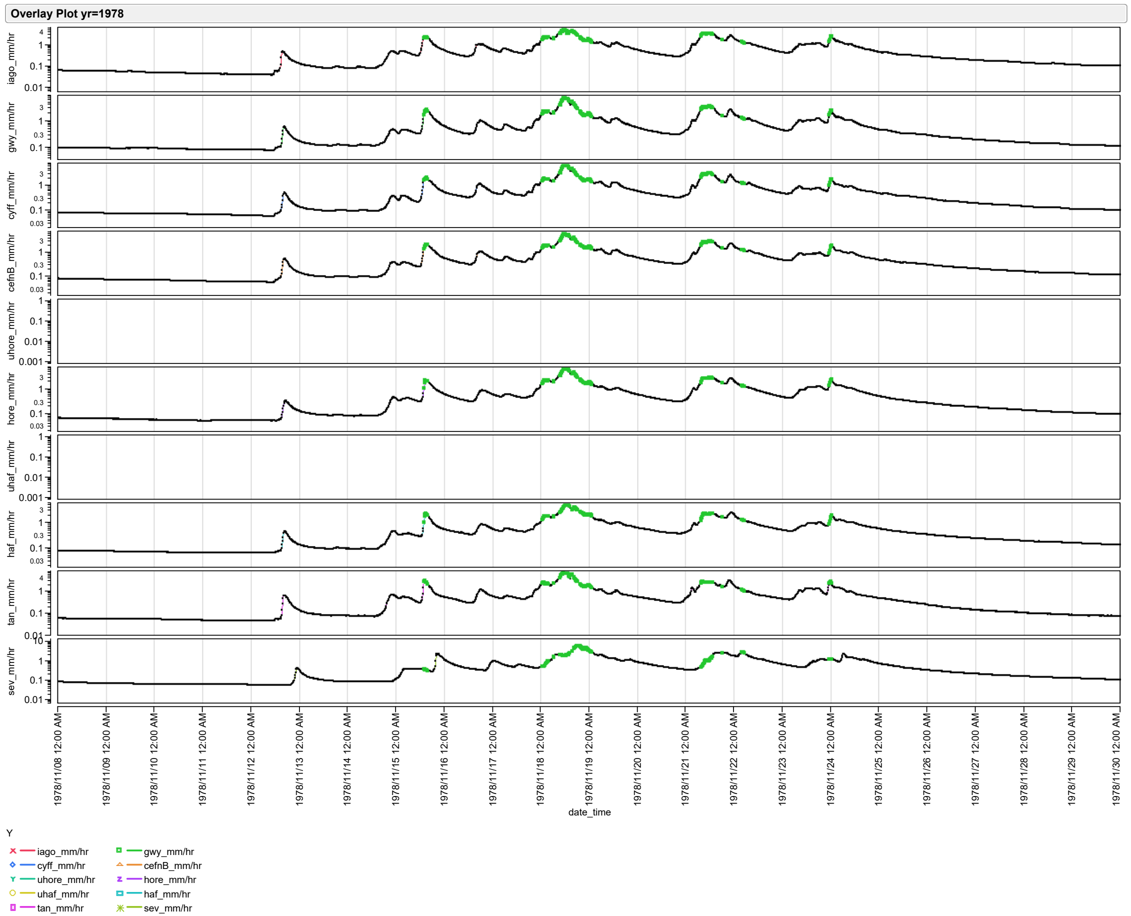
Scatterplot Matrix



Above is another example, from 1983, showing problems at Tanllwyth. Note the "wings" in the plots against all the other sites. These arise from a clock error, in which the datalogger clock has been initialized wrong sometime in early December 1983, in the middle of the time series shown below (Tanllwyth is the next-to-last line, and vertical lines are midnight). Note that elsewhere in the record, Tanllwyth is never so far out of synch with the other gauges.



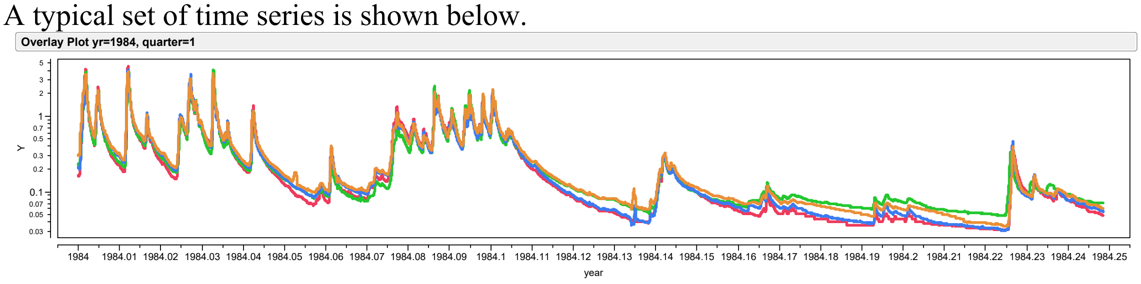
Likewise the timing has gone off at Severn in late 1978 (see example below – Severn is the last line). Note that elsewhere in the record, Severn is never so far out of synch with the other gauges.



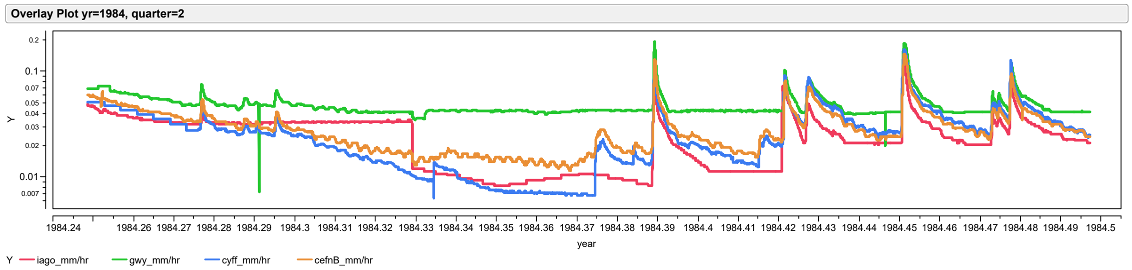
Still earlier in the record, timing errors like these are more common. I would advise great caution in using the streamflow data from the 1970's for cases where timing is important (for example, in rainfall-runoff relationships).

At low flows in the mid-winter, icing can apparently cause artifactual changes in stage in the flumes, and thus erroneous discharge readings. The most visually obvious examples of this problem were culled.

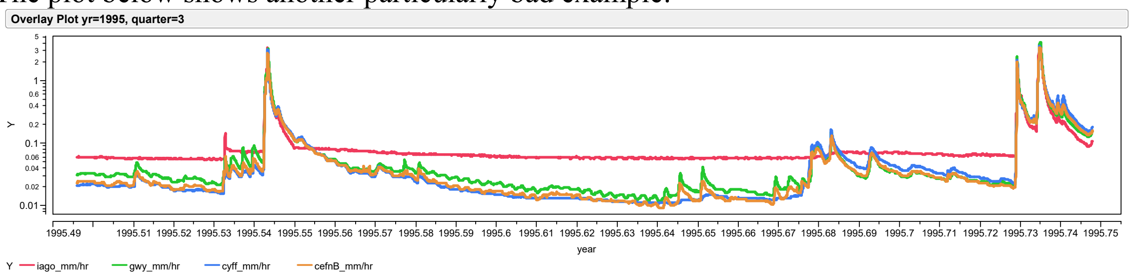
A typical set of time series is shown below.



The following quarter illustrates a problem that arises during dry periods in the Wye catchment. The gauged flow goes flat at a lower limit:



The plot below shows another particularly bad example:



The most severe such cases of flat-lining have been culled. These mostly involve the Iago, as well as the Gwy in the 1980's and earlier.

Similar flat-lining is also observed, to a much more limited extent, in the Severn catchment, mostly at the Severn and Hore gauges, particularly in 1984 and before.

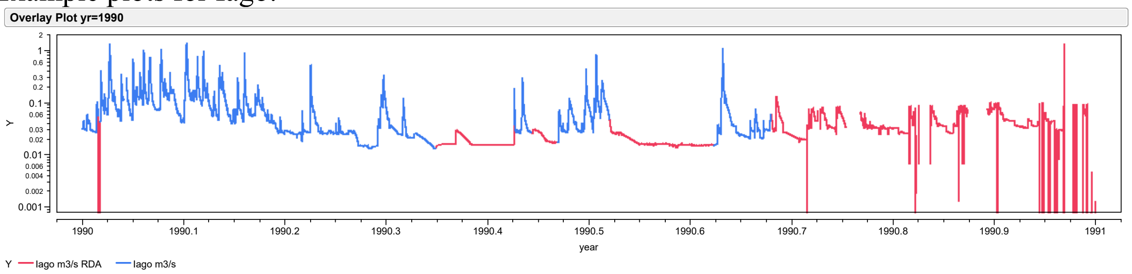
Another major glitch is an apparent conversion error at the Upper Hore. Starting on 2010/01/09 at 1:45 PM (SQNUM=55) and continuing until the end of the time series, reported flows at the Upper Hore are inflated by a factor of 10. Also, starting on 1996/06/17 at 9:15 AM and continuing until 1996/07/01 at 1:00 PM, reported flows at the Upper Hore are a factor of 10 too small. These obvious conversion errors have been corrected in the edited data.

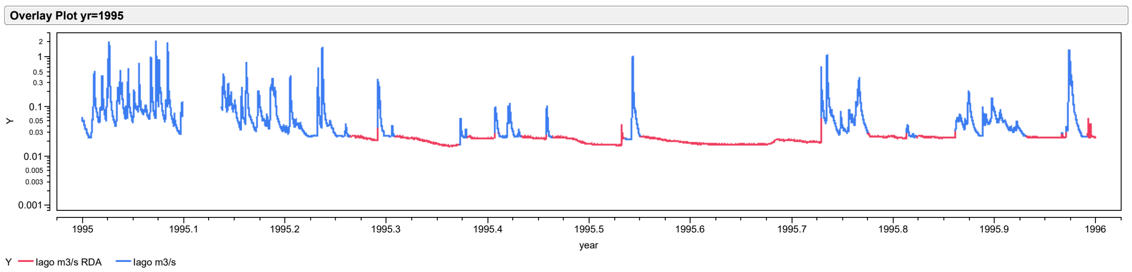
The fractions of the 15-minute flow records that have been culled at each of the gauges are as follows. These do not include the 515 culls of doubled time stamps (see farther below):

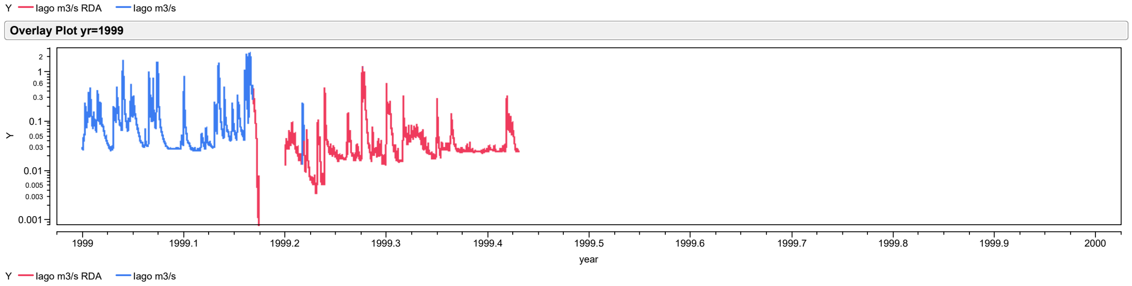
Iago	15.6 percent
Gwy	3.6 percent
Cyff	1.2 percent
Cefn Brwyn	1.3 percent
Upper Hore	1.4 percent (does not include the conversion errors that have been corrected)
Hore	2.5 percent
Upper Hafren	<0.01 percent
Hafren	0.13 percent
Tanllwyth	1.3 percent
Severn	1.5 percent

Examples of some of the most heavily culled years from each site are shown below; as always, culled records are shown in red and kept records are shown in blue. In some cases the problems that have led to data being culled will not be obvious unless the records are compared on mm/hr axes, whereupon it becomes clear that the behavior of one stream gauge is inconsistent with the others.

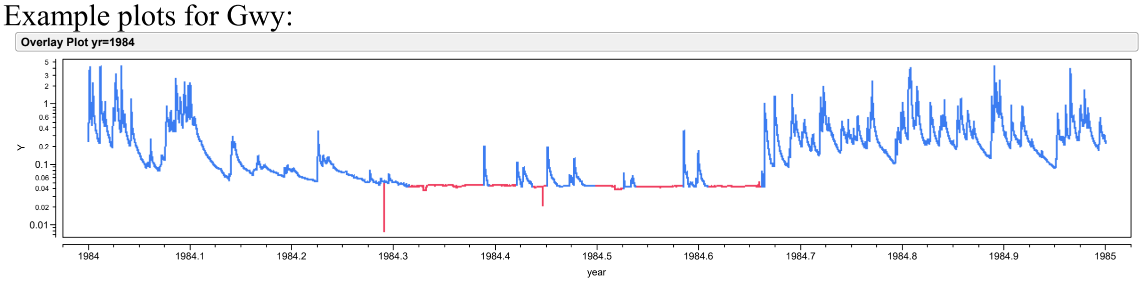
Example plots for Iago:

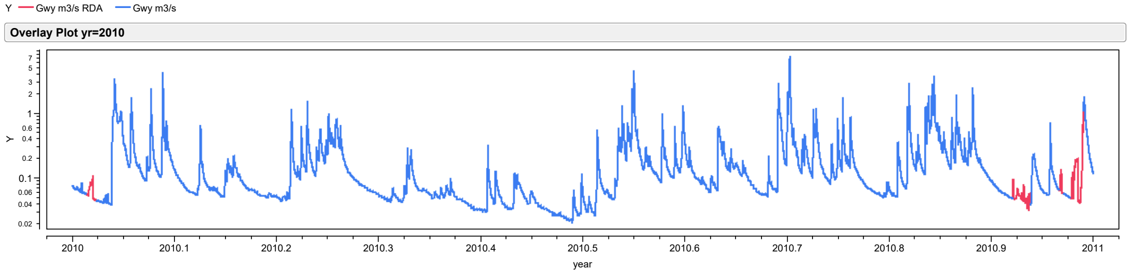


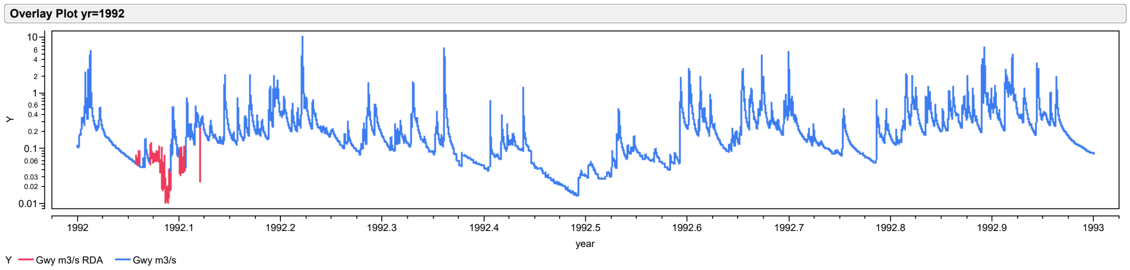




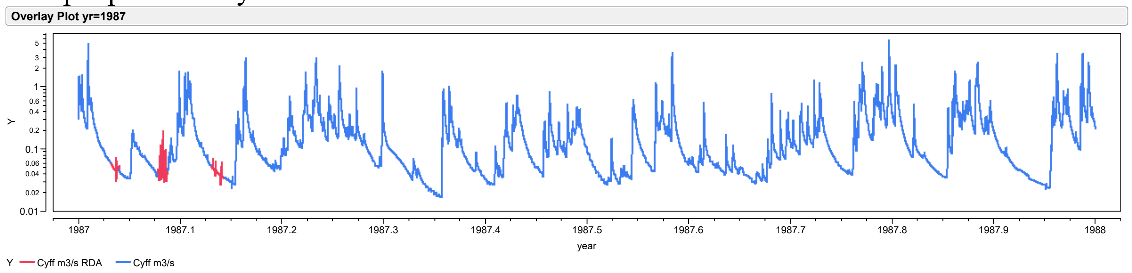
Example plots for Gwy:

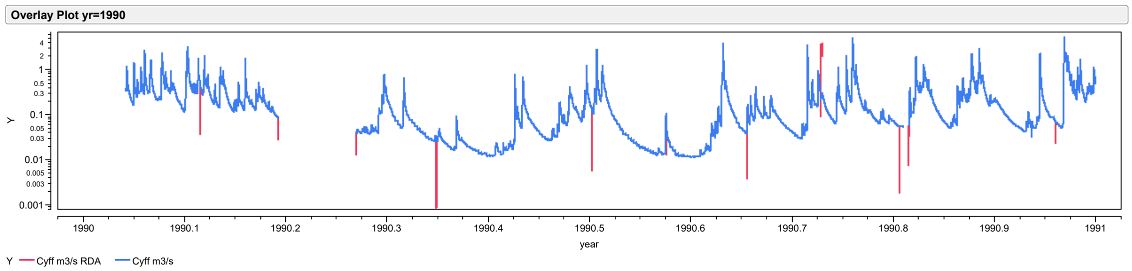


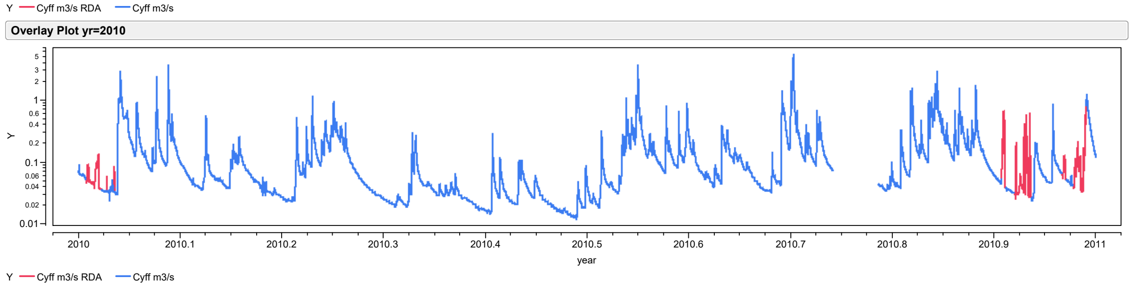



Overlay Plot yr=1992
Y
year
Y Gwy m3/s RDA Gwy m3/s

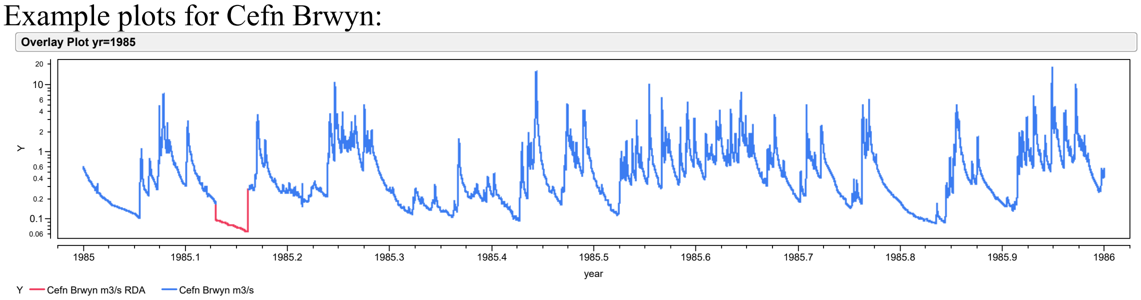
Example plots for Cyff:

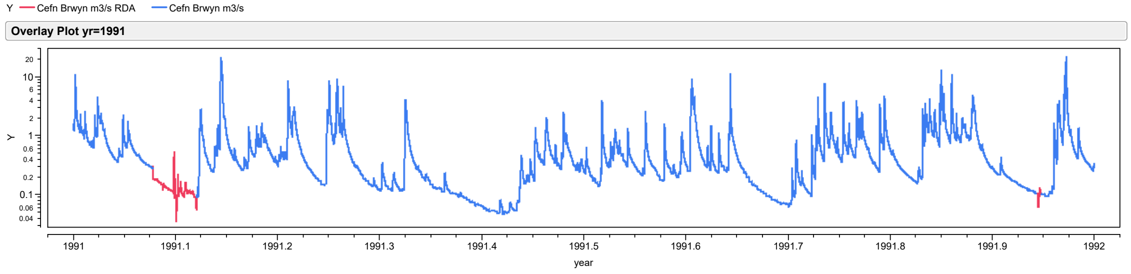


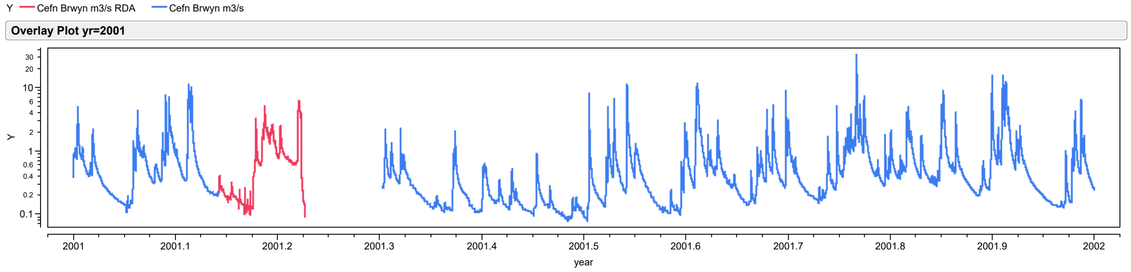




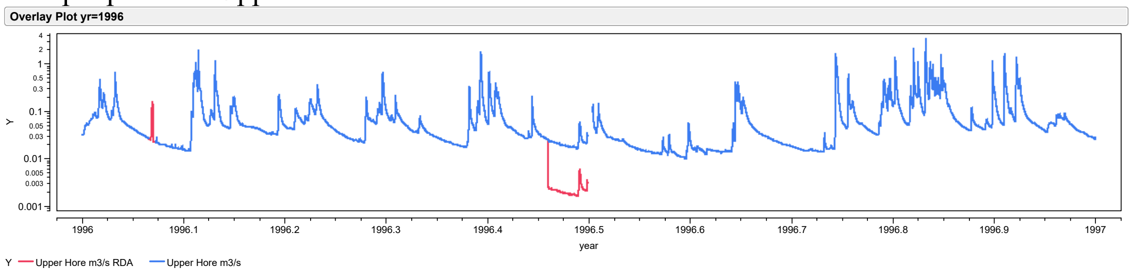
Example plots for Cefn Brwyn:

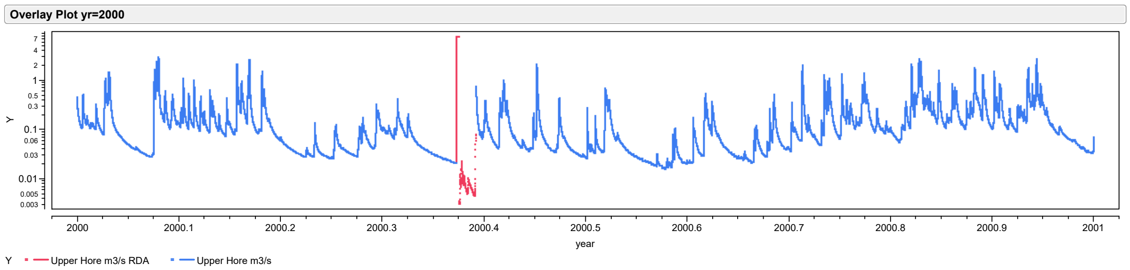


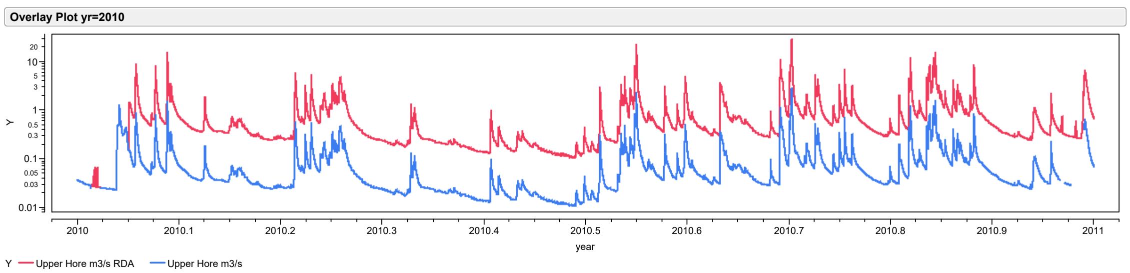




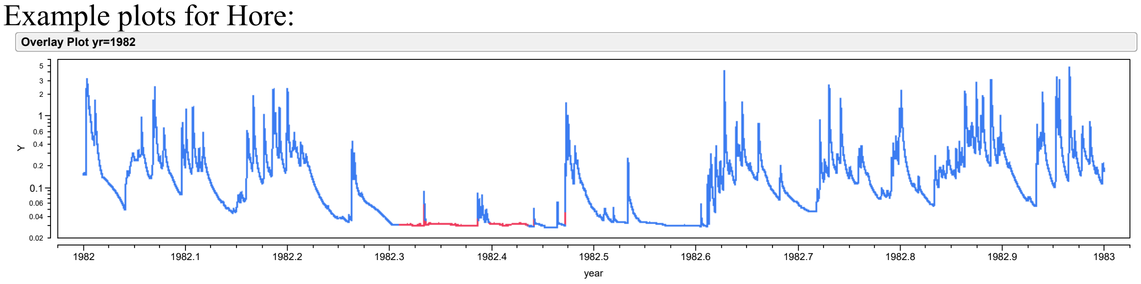
Example plots for Upper Hore:

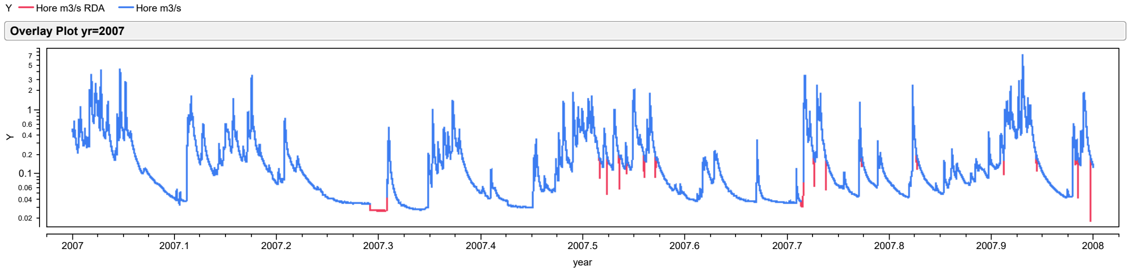


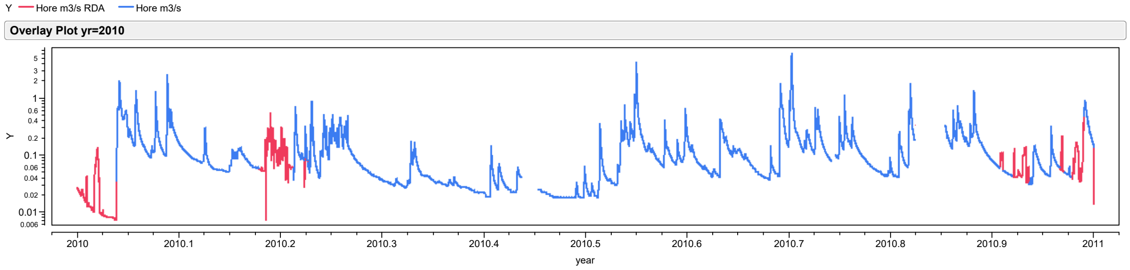




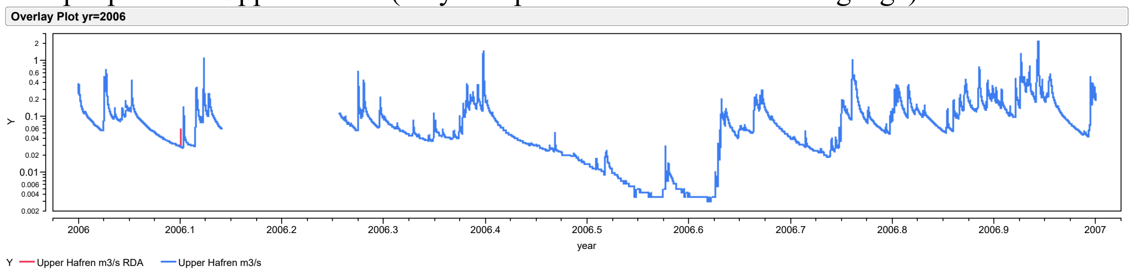
Example plots for Hore:

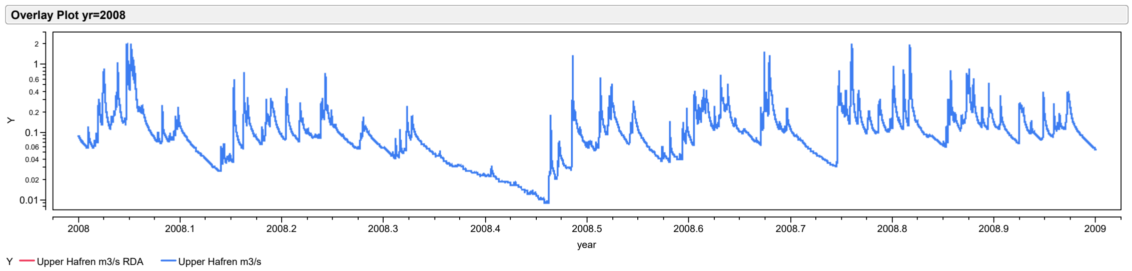


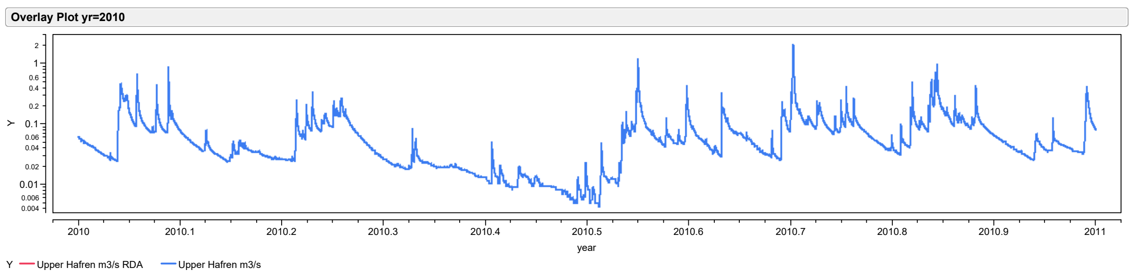




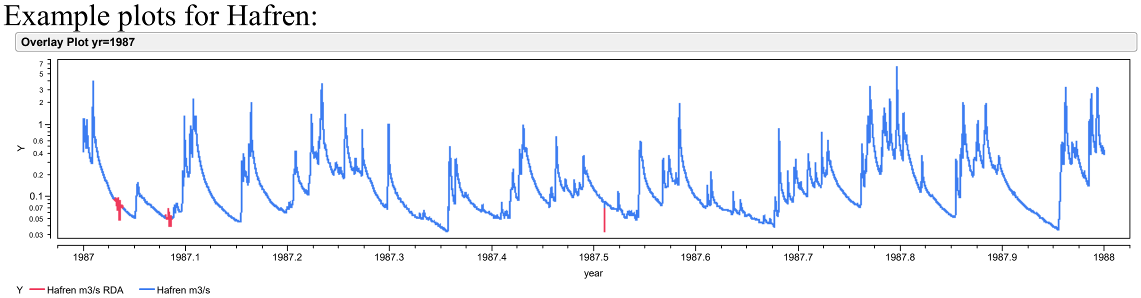
Example plots for Upper Hafren (very few points were culled for this gauge):

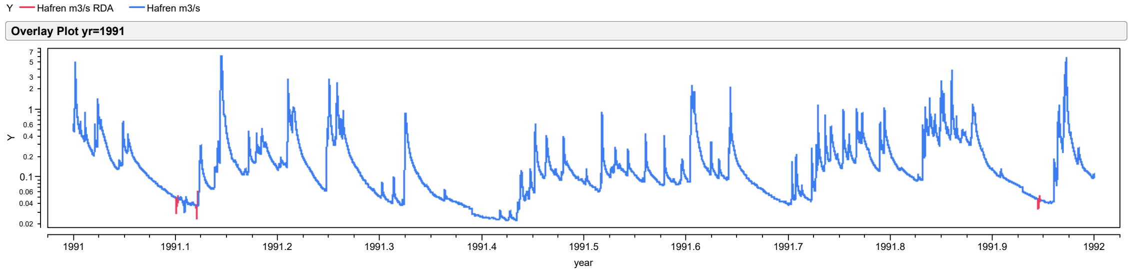


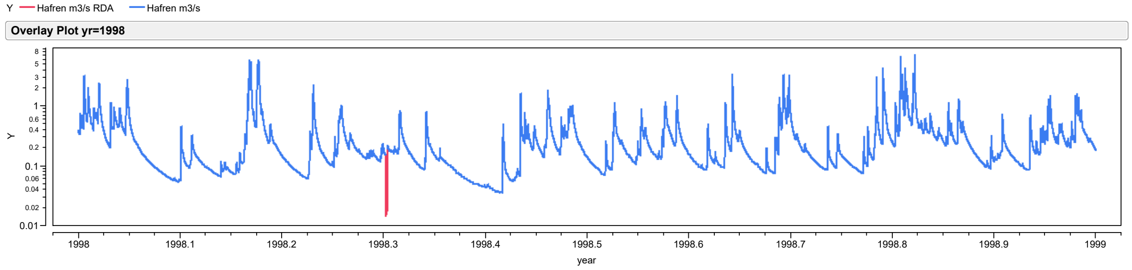




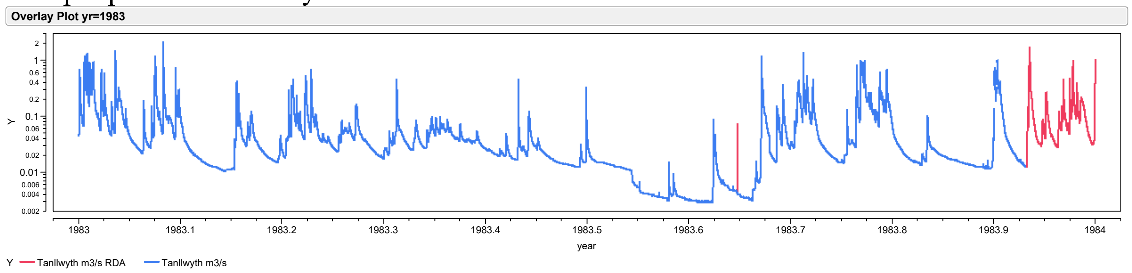
Example plots for Hafren:

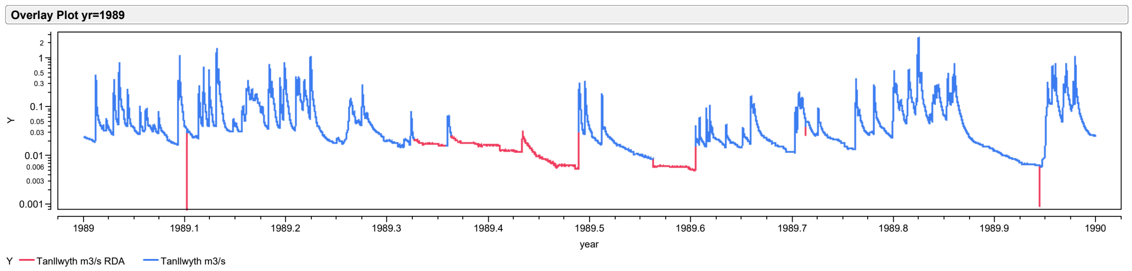


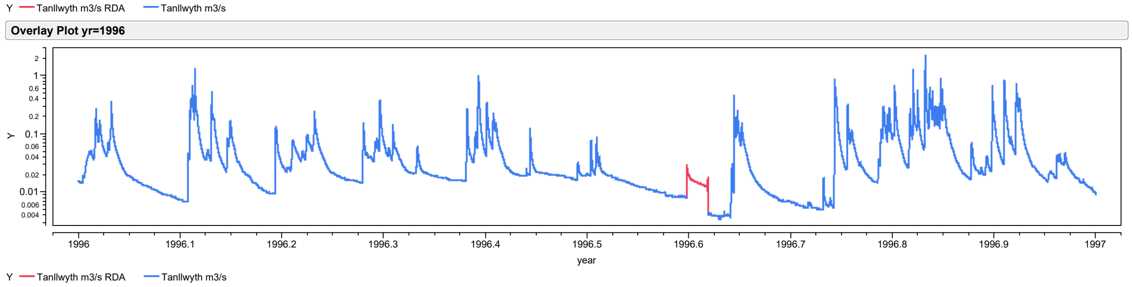




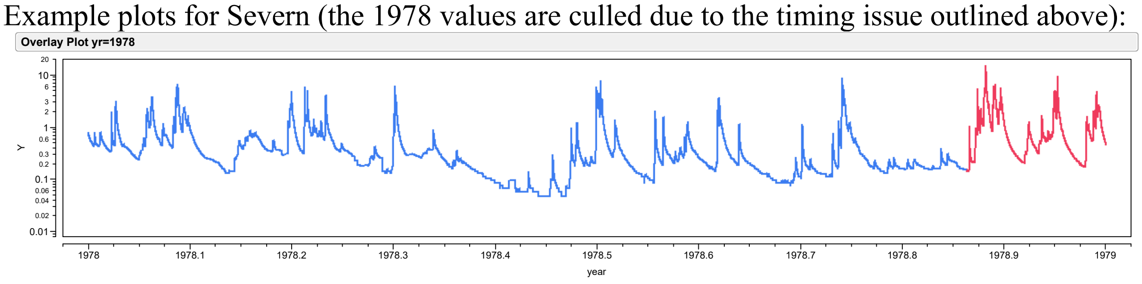
Example plots for Tanllwyth:

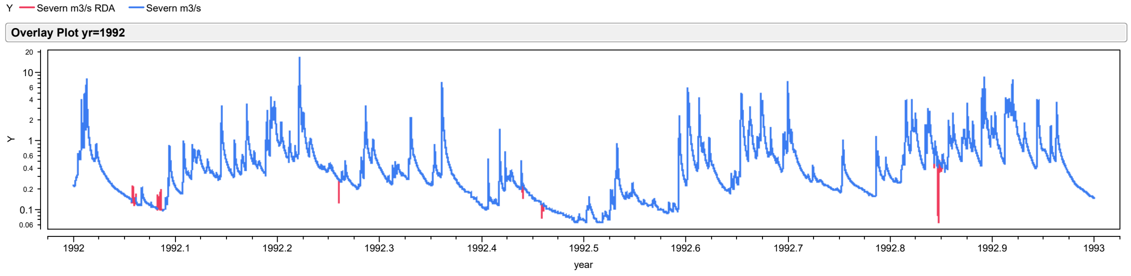


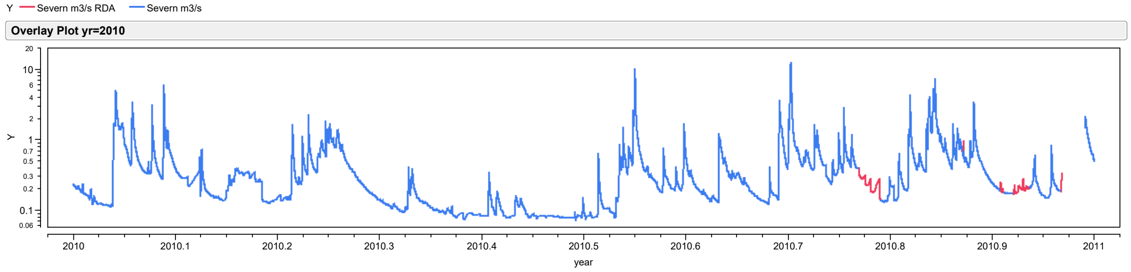




Example plots for Severn (the 1978 values are culled due to the timing issue outlined above):







There are 515 pairs of doubled date/time stamps in the discharge timestamp. One of these, 2006/12/31 11:45PM, appears to be a throw-forward of a year for measurements at Hore, Hafren, and Severn (but not the other gauges) on 2005/12/31 at 11:45PM. The thrown-forward values were returned to their correct date (and culled on the wrong date).

For all the other doubled time stamps with duplicate discharge values, one of each pair was culled.

Plynlimon AWS data cleaning notes
J.W. Kirchner 28 February 2014

I started with Oracle export files supplied by Matt Fry (Wallingford), and original AWS data logger files supplied by Dave Norris (Bangor). The Wallingford Oracle files cover the period of record at each AWS, up to 20 November 2012 at the latest. AWS logger files are available for three AWS's (Carreg Wen – ID3, Eisteddfa Gurig – ID7, and Gwy – ID49) after that date, running to January 29, 2014 (the most recent download) at Carreg Wen, and until June 5, 2013 and July 2, 2013 at Gwy and Eisteddfa Gurig, respectively (permanent logger failure at both of these sites).

I extended the Oracle export data with the logger files, where available, after converting the latter for compatibility with the former. This conversion entailed:
- modifying date and time formats (loggers are reported to be in GMT; times in Oracle database are one hour earlier than logger times)
- rounding solar flux, net radiation, and wind direction to integer values
- calculating temperature depression from wet and dry bulb temperatures

I also compiled the original logger files for two AWS's that do not appear at all in the Wallingford Oracle database. The first of these is logger code "TS", site numbers 61 and 192, which was installed at Tanllwyth near the site of the original station (and the Aberystwyth AWS, see below) from June 2003 until February 2004. This AWS was subsequently relocated to a new site at "mid-Tanllwyth", farther from tree encroachment and with a lower risk of vandalism.

The second group of original logger files came from the Aberystwyth University AWS that was located at Tanllwyth from 2000 to 2004 (logger code "TW, site #107). The original data columns are: site number, Julian day, hour, temperature (°C), relative humidity (%), windspeed (m/s), rainfall (mm/hr), wind direction (degrees), and battery voltage; there are no records of solar flux or net radiation. Because I could not find documentation for the relative humidity sensor, I used the Sprung psychrometric formula to infer the expected wet-bulb depression, and then cross-calibrated this against the observed wet-bulb depression at the Tanllwyth "TS" AWS (see above) during the periods where the two AWS records overlap. I then used the cross-calibration equation, in combination with the Sprung formula, to infer wet-bulb depression from relative humidity. The converted data contain only temperature, inferred wet-bulb depression, wind speed and direction, and rainfall. Logger times were converted assuming that they were synchronous with the CEH loggers.

Columns in the resulting files are:
SITENO – site or logger number (from logger file)
STIME – date and time
SN – solar radiation in W/m^2
RN – net radiation in W/m^2
TD – temperature depression (difference between wet and dry bulb temp's) in °C
T – air temperature (dry bulb) in °C
WS – wind speed in m/s
WD – wind direction in degrees
RF – tipping bucket precipitation in mm/hr

There have been multiple AWS's at some sites at various times over the years, and some AWS locations have changed, by amounts ranging from a few meters to hundreds of meters (at Tanllwyth). The naming conventions have not been fully consistent and this can be a source of confusion. The following notes outline my understanding of the histories of each of the AWS data sets, and describe the AWS codes that appear in the edited data sets. These AWS codes often correspond to those used in the logger data files, but not always.

CarregWen_AWS_CW_ID4_Oracle_export_20140217.tsv (3 Jan 1976 – 19 Jan 2005), site #4: this was the primary Carreg Wen AWS until 2005, when it was discontinued due to tree encroachment. I will denote this as "CW".

CarregWen_AWS_CW1_ID3_to1981_Oracle_export_20140217.tsv (8 Jan 1976 – 12 Oct 1981), site #3: this was a backup station co-located near CW. I will denote this as "CW2".

CarregWen_AWS_CW2_ID3_from2001_Oracle_export_20140217.tsv (27 Jul 2001 – 29 Jan 2014, still in operation), site #3: this is a replacement site, selected to avoid the tree encroachment problems at CW. Note that although this carries the same site number as CW2, I do not believe that it occupies the same physical location. I will denote this as "CW3".

CefnBrwyn_AWS_CF_ID6_Oracle_export_20140217.tsv (2 Apr 1975 – 10 Jun 2003), site #6: This was the primary Cefn Brwyn AWS. It was considered a poor site and was closed in 2003, and the station hardware was moved to Gwy. I will denote this as "CF".

CefnBrwyn_AWS_CF1_ID5_Oracle_export_20140217.tsv (8 Jan 1975 – 18 Aug 1982), site #5: This was a secondary Cefn Brwyn AWS. I will denote this as "CF2".

EisteddfaGurig_AWS_EG_ID7_Oracle_export_20140217.tsv (7 Jan 1976 – 2 July 2013), site #7: This was the primary Eisteddfa Gurig AWS. I will denote it as "EG".

EisteddfaGurig_AWS_EG2_ID8_Oracle_export_20140217.tsv (21 Jan 1976 – 22 Apr 1982), site #8: This was a backup/supplementary AWS near EG. I will denote this as "EG2".

Gwy_AWS_ID49_Oracle_export_20140217.tsv (6 Jul 1999 – 5 Jun 2013), site #49: This was the sole AWS at Gwy. I will denote this as GY (GW is already used for the stream gauge).

NantIago_AWS_ID55_Oracle_export_20140217.tsv (23 Feb 1978 – 2 Dec 1982), site #55: I will denote this as "NI".

Tanllwyth_AWS_Tan_ID50_to1999_Oracle_export_20140217.tsv (25 Apr 1975 – 5 May 1999), site #50: This was the original Tanllwyth AWS, located near the Moel Cynedd manual met site. I will denote this as "TN".

Tanllwyth_AWS_Tan2_ID51_Oracle_export_20140214.tsv (25 Aug 1976 – 21 Apr 1982), site #51: This was a secondary backup AWS, also located near the Moel Cynedd manual met site (I believe). The Oracle export file for this site includes five variables that do not appear in any of the other files: Specific humidity (SH), specific humidity deficit (SHD), and Penman-Monteith PET estimates for water, grass, and forest (PM_W, PM_G, and PM_F). Because these do not appear for the other files, I have not carried them forward into the merged file for all the AWS's. I will denote this AWS as "TN2".

The Tanllwyth AWS with logger files "TS", also located near the Moel Cynedd manual met site (11 June 2003 – 17 Feb 2004), site #61 and 192. This is the AWS described above, which was installed near the site of the old Tanllwyth AWS before being moved to the new mid-Tanllwyth site. Note that this has sometimes been called "Tan2", with potential for confusion with the AWS that was in operation from 1976 to 1982 (site #51). Note also that two site numbers appear in the logger files, but I believe that this just reflects the use of two loggers with two ID numbers. I will denote this as "TN3".

The Aberystwyth University AWS, also located near the Moel Cynedd manual met site (18 May 2000 – 13 April 2004), site #107: There are no Oracle records for this AWS, and the original logger files were converted as described above. I will denote this as "TW".

Tanllwyth_AWS_TanMid_ID50_from2005_Oracle_export_20140217.tsv (1 Jan 2005 – 15 Feb 2012), site #50: This was located at the "Tan Mid" site, located some 100's of meters from the original Tanllwyth met site. Note that in the Oracle database, this AWS apparently shares the same site number (50) with the original Tanllwyth site, but it is not the same physical location. I will denote this as "TM".

The list above comprises 14 AWS's in total. All of the available records from all of the AWS's have been concatenated in a single file, comprising 1,630,091 records, each with up to seven sensor values plus the site number and timestamp.

The original time series as supplied by CEH have been retained in the data set, in columns whose names begin with "RDA" (for Raw Data Archive). Column headers for the edited time series do not carry the RDA designation. The master raw data archive file is *All_AWS_stacked_RDA_master_archive.jmp*. Because the sample times (STIME), site codes (ID) and site numbers (SITENO) have in some cases been duplicated or mis-assigned (see below), I created a unique identifier, *Original_sort_order*, that can be used to connect edited records to the raw data archive. The site and sampling time are unreliable for this purpose, because they may need to be changed for some records. The original sample timestamps, site codes, and site numbers are recorded as RDA_STIME, RDA_ID, and RDA_SITENO.

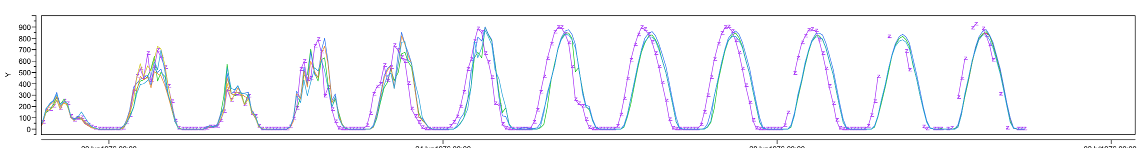
Correction of timestamp errors

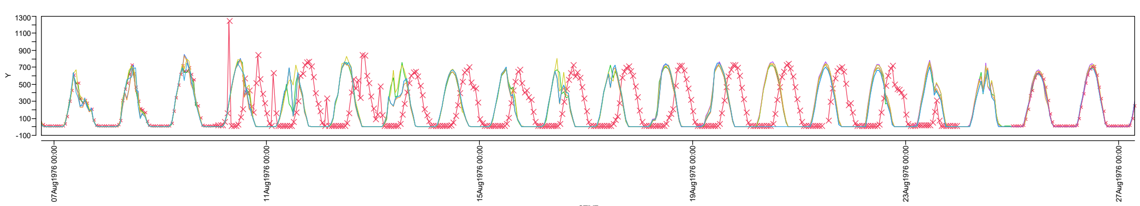
In total, there are 1594 **duplicate timestamps**. These include:

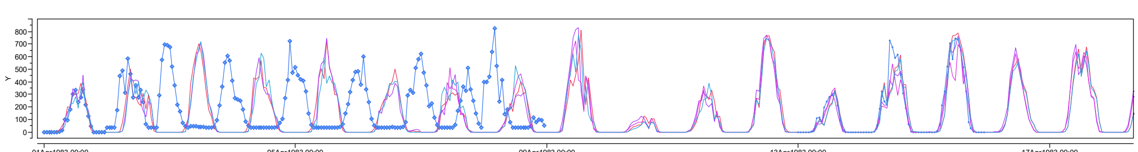
Carreg Wen (CW): 339 records
Carreg Wen (CW3): 1 record
Eisteddfa Gurig (EG): 1251 records
Gwy (GY): 1 record
Tanllwyth (TN): 1 record

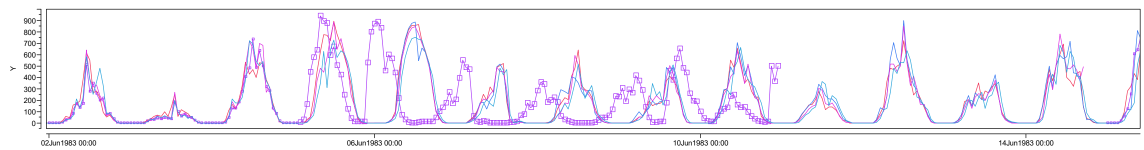
CW: hourly duplicates from 03Nov1998 9h00 to 17Nov1998 8h00 – deleted
CW: 29Jun2002 0h00, one copy should be 30Jun2002 – reassigned
CW: 03Jun2001 3h00, one copy doesn't fit – deleted
CW: 03Jun2003 3h00, one copy doesn't fit – deleted
CW3: 30Dec2004 23h00, one copy doesn't fit – deleted
EG: hourly duplicates from 25Oct1999 8h00 to 15Dec1999 14h00 – deleted
EG: 31Dec2006 23h00, one copy should be 31Dec2005 – reassigned
EG: hourly duplicates on 18Dec2008 (0h00-14h00) – one copy deleted
EG: hourly duplicates on 20Dec2008 (20h00-23h00) – one copy deleted
GY: 31Dec2007 23h00, one copy should be 31Dec2006 – reassigned
GY: 31Dec2006 23h00, one copy should be 31Dec2005 – reassigned (this duplicate does not appear in the original file, but is created when the duplicate on 31Dec2007 is reassigned to 2006).
TN: 01Dec1977 22h00, one copy doesn't fit – deleted
TN: 03Oct1978 22h00, one copy blank – deleted

There are intermittent **clock shift errors** in several records, most likely due to incorrect initialization after downloading. These are most clearly seen in the solar flux time series, because the solar flux measurements should be tightly correlated across the different AWS's. Where one stands out with a persistent time shift (as shown in the examples below), it clearly indicates a clock shift error.









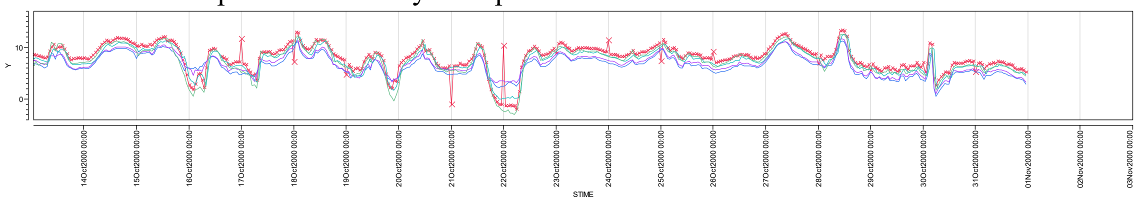
These **clock shift errors were corrected**, wherever this could be done unambiguously by shifting the sampling times by a fixed amount, so that a given site's solar flux pattern matched that of the other AWS's. If the solar flux patterns could not be unambiguously matched, the data with timing errors were deleted (for all sensors).

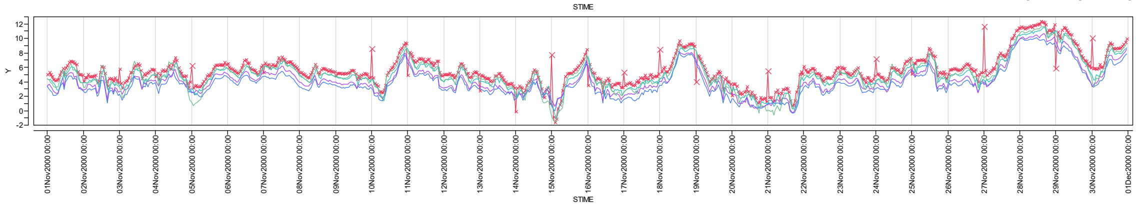
Two weather stations (Cefn Brwyn CF and Eisteddfa Gurig EG) exhibit **midnight throwback errors**, in which midnight values are incorrectly assigned to the previous day. These problems are most clearly seen in the temperatures, since solar fluxes are zero at midnight. The likely origin is incorrect handling of 24h00 and/or 0h00 conventions between the loggers and the subsequent database uploading.

These "throwback" errors are as follows:

- At Eisteddfa Gurig, midnight records have been "thrown back" one day starting on 1998/05/08 (actually 1998/05/09 which was thrown back to 1998/05/08), and continuing until 1998/12/31 (for which the midnight value is missing, having been thrown back to 1998/12/30).
- At Cefn Brwyn, midnight records have been "thrown back" one day starting on 2000/10/17 and continuing until 2000/12/31 (which again is missing, having been thrown back to 2000/12/30).
- Also at Cefn Brwyn, 11PM records have been "thrown <u>forward</u>" one day for the entire year of 2001. 2001/01/01 is missing, having been thrown forward to 2001/01/02.

Below is an example of Cefn Brwyn temperature records with the "throwback" error:





These throwback/throw forward errors were corrected by recalculating the date and time appropriately. The original date and time stamps are preserved in the data column "RDA STIME".

The working file with timing errors corrected is *All_AWS_stacked v03.jmp*.

<u>General (multi-sensor) data editing</u>

Data blocks with **zeroes in all sensors** (or all-but-one sensors) were presumed to reflect logger failure, and were deleted.

In many data blocks at Eisteddfa Gurig, the database shows a Swiss cheese of missing values, zero values, and trivially small, near zero values across all sensors. These blocks were presumed to reflect logger failure, and was deleted. This "**Swiss cheese effect**" is seen at Eisteddfa Gurig from 9Jan1984 to 23Jan1984, from 19Mar1984 to 02Apr1984, 07Apr1984 to 16Apr1984, 02Oct1984 to 29Oct1984, and 07Nov1984 to 12Nov1984. Curiously, this effect is always seen in blocks of whole days.

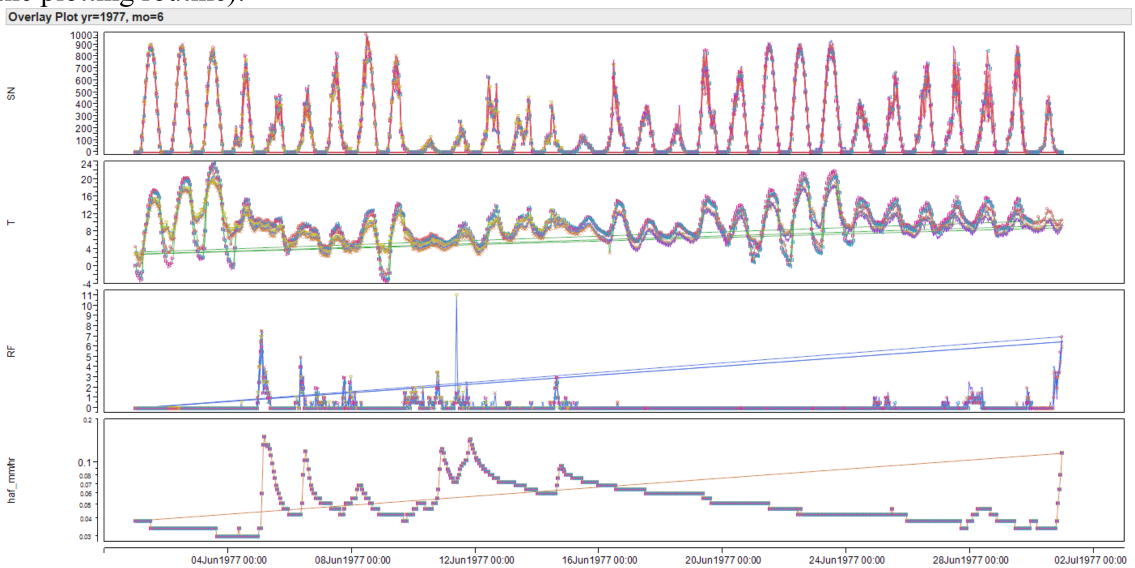
All values that were **grossly out of range** (e.g., negative rainfall totals, or large negative net radiation values) or that had **missing data codes** (-6999) were deleted without detailed inspection.

The working file after this general data editing is *All_AWS_stacked v04.jmp*.

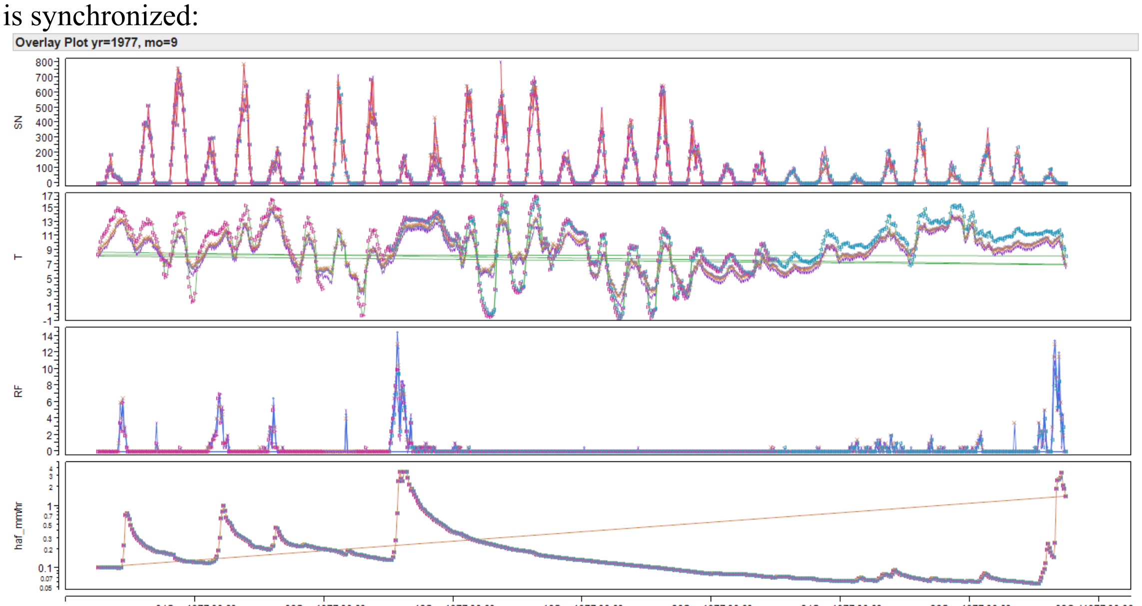
Detection and correction of registration errors

Registration errors, in which the data values are plausible but they are assigned to the wrong dates and times, sometimes can be found, particularly in the early years. Most of these errors have likely been created by joining individual logger files together in the wrong order, and then overwriting the original timestamps.

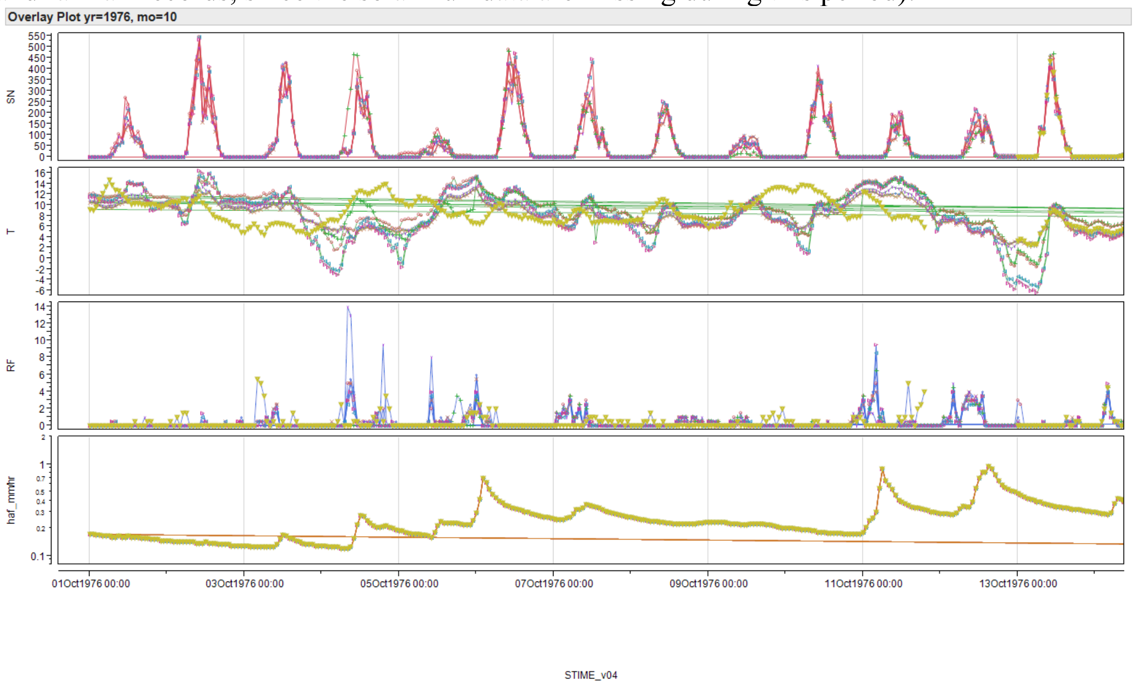
These registration errors can be diagnosed by comparing solar flux, temperature, and rainfall patterns among the different sites. When all the records are correctly registered, as in most of the later years, these variables rise and fall in synchrony, except in particular unusual situations (valley fog, snow on solar flux sensors, etc.). Here is an example of typical levels of synchrony during summer weather (ignore the long straight lines, which are an artifact of the plotting routine):



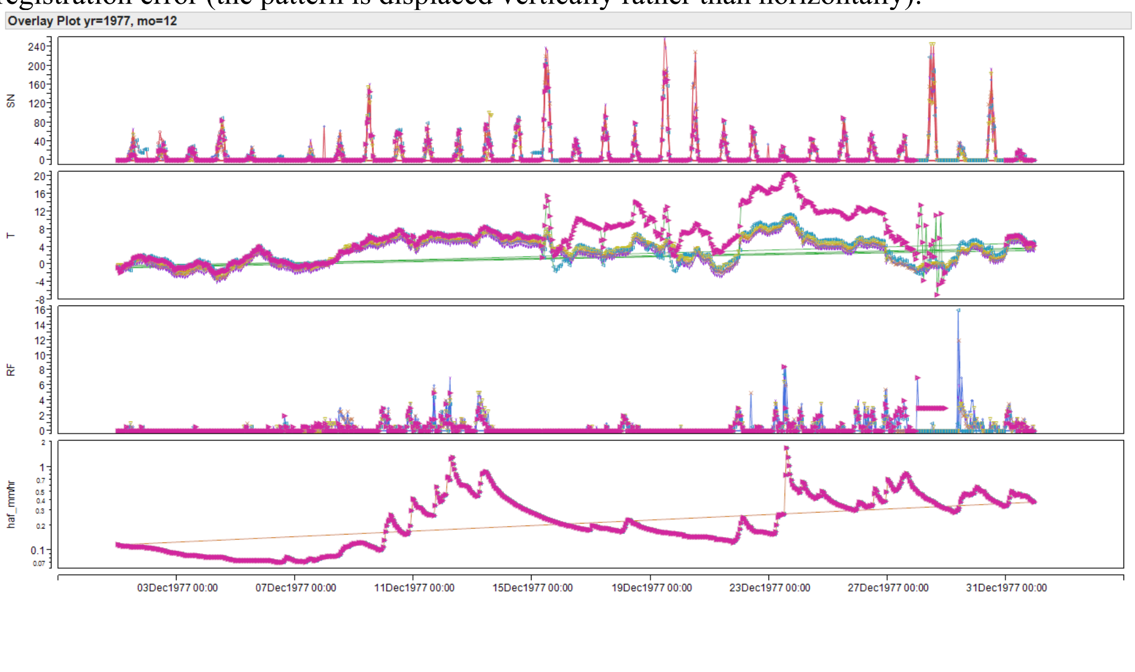
Below is another example. The temperature swings at TN and TN2 are bigger, but everything is synchronized:



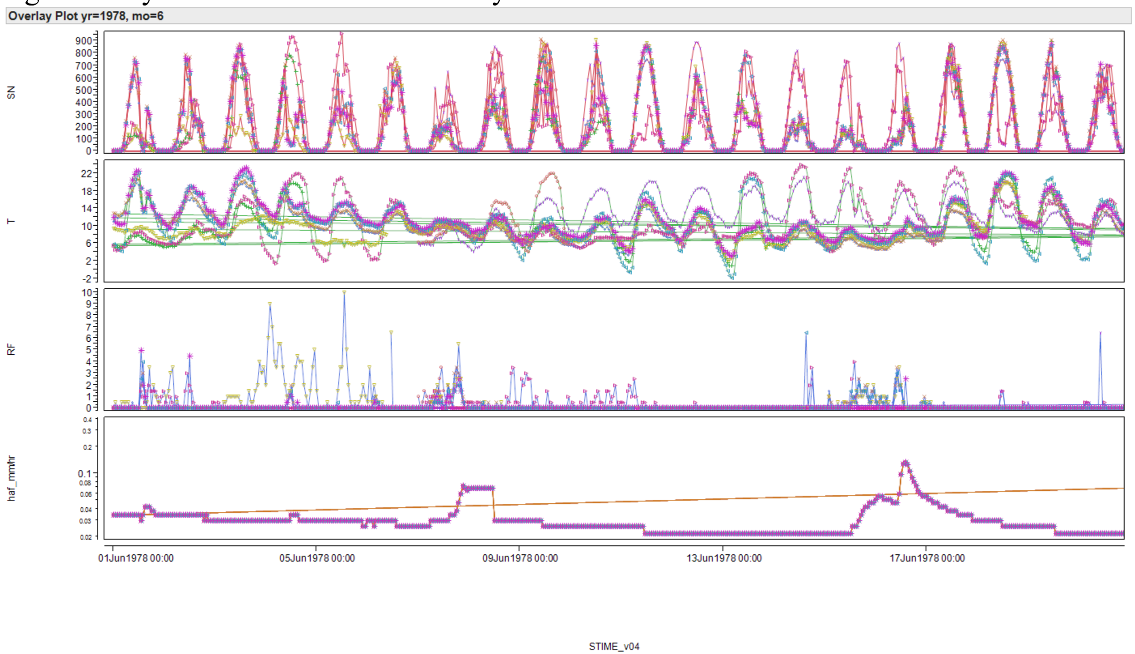
Here is a simple registration error that is easy to detect. EG2 (highlighted in yellow/green) has been shifted about 1 day early, relative to the other sites (as evident from the temperature and rainfall records, since the solar flux data are missing during this period):



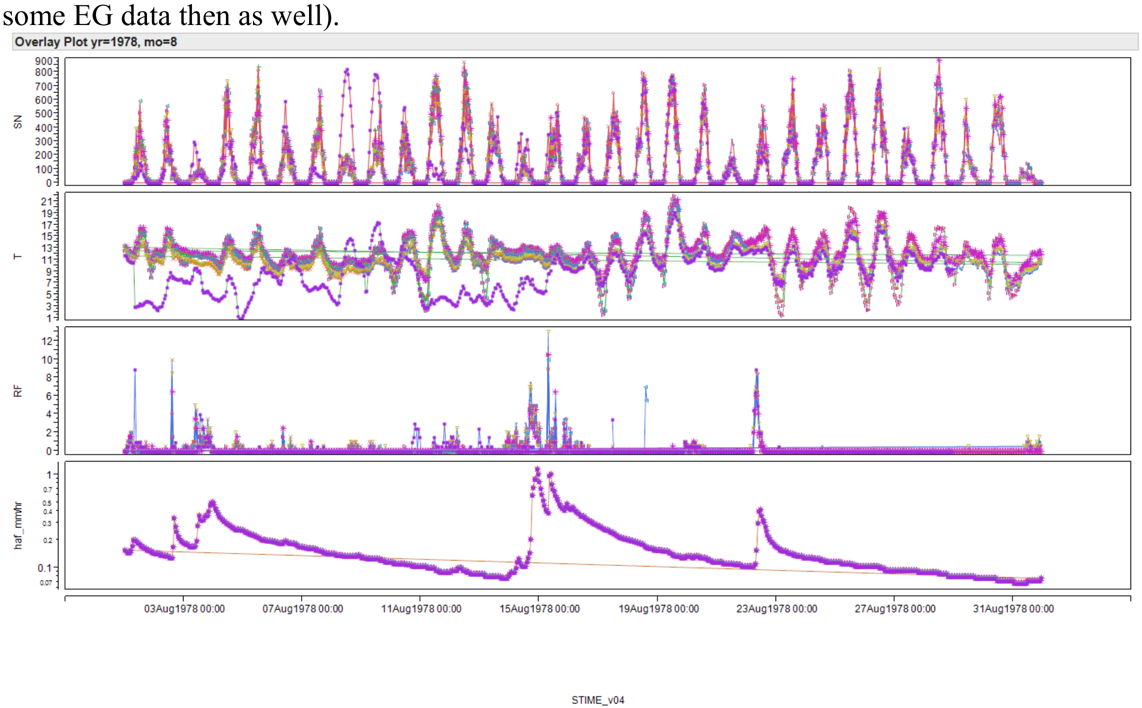
Here is a temperature calibration problem, at TN2 (highlighted in pink). This is not a registration error (the pattern is displaced vertically rather than horizontally):



Here, on the other hand, is an unmitigated registration error catastrophe. EG, EG2, TN2, CF2 and possibly others are all time-shifted to varying degrees and by varying amounts. Note, for example, how EG2 (yellow) shows a significant rainstorm that does not occur in the other AWS's, as well as a different temperature and solar flux pattern. Some of these problems begin in May 1978 and continue until July 1978.



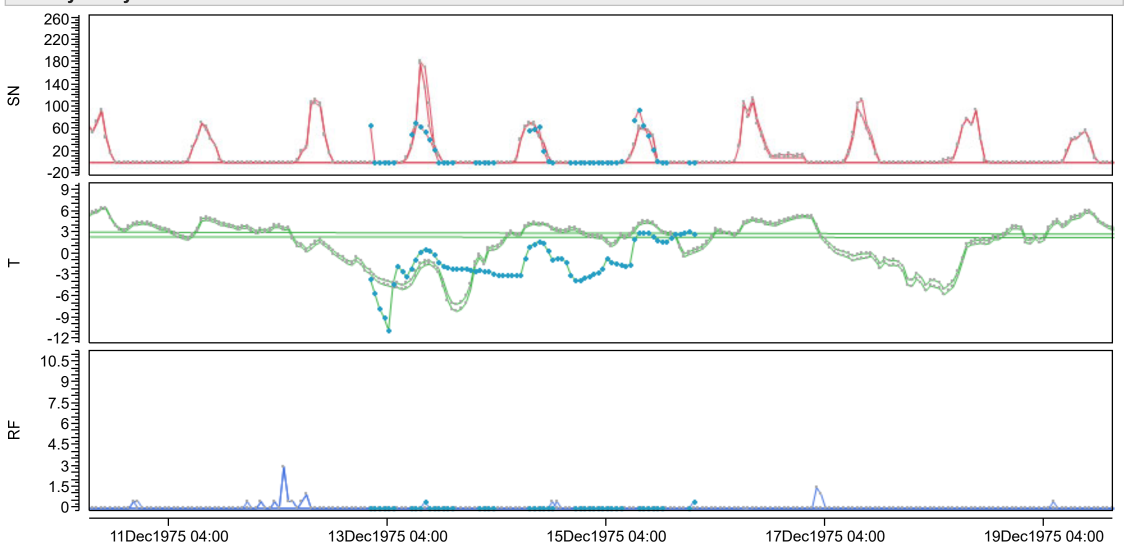
In the first two weeks of August 1978, EG shows a pattern (highlighted in purple below) that is completely different from the other sites. This arises from a huge time shift; these values belong to early MAY of 1978, and possibly to another site besides EG (because there are some EG data then as well).



Below I enumerate all the identified date shift errors, and the steps I took to address them. All dates and times refer to the original timestamps, with the (relatively rare) corrections or adjustments that have already been made, up to *All_AWS_stacked_v04.jmp*. The old timestamps are called *STIME_v04*; the new, corrected timestamps are called *STIME_new* or just *STIME* (depending on the version of the file).

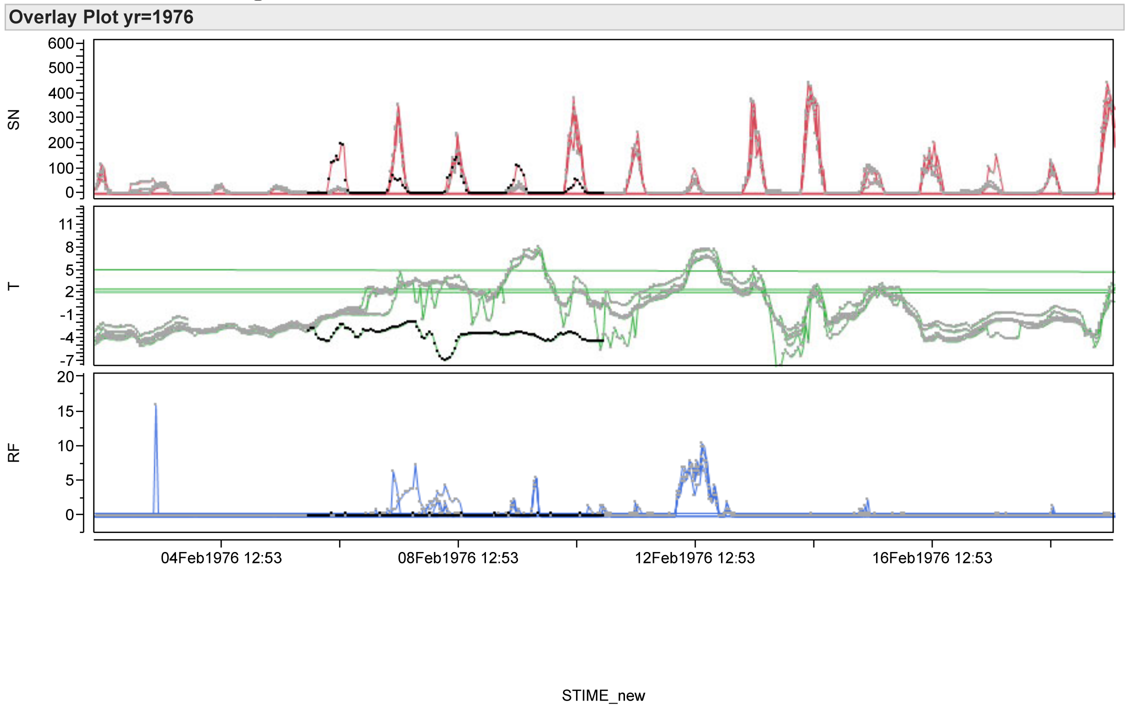
<u>December 1975:</u> From 13 to 15Dec1975, TN is clearly not registered to the correct date/time. This could not be resolved (too short of a record) and these points were deleted:

Overlay Plot yr=1975

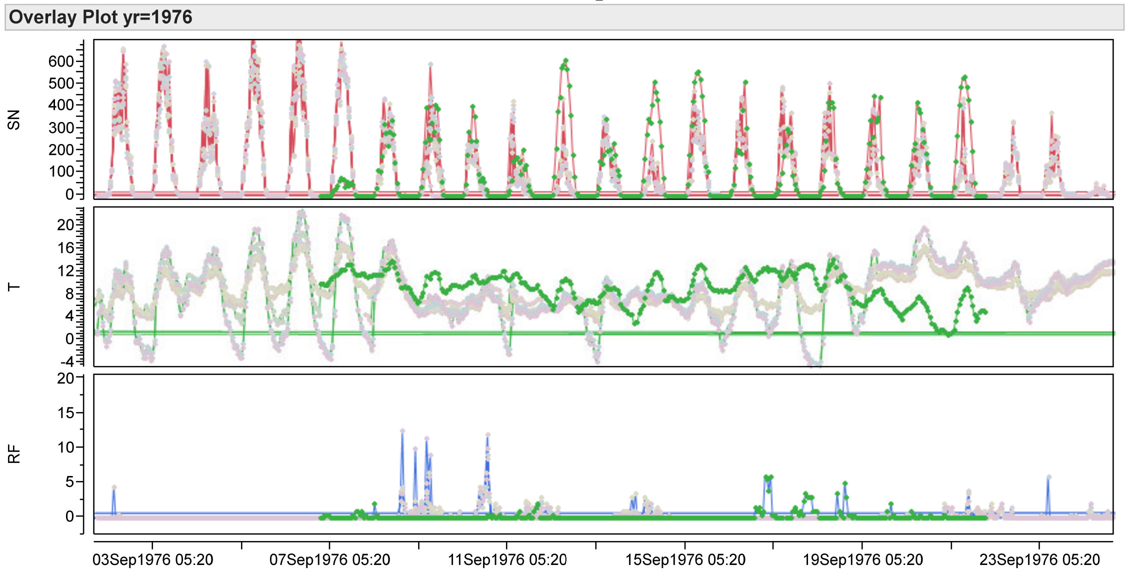


<u>February 1976:</u> From 6 to 10Feb1976, CF2 clearly has incorrect dates/times. This could not be resolved and these points were deleted:

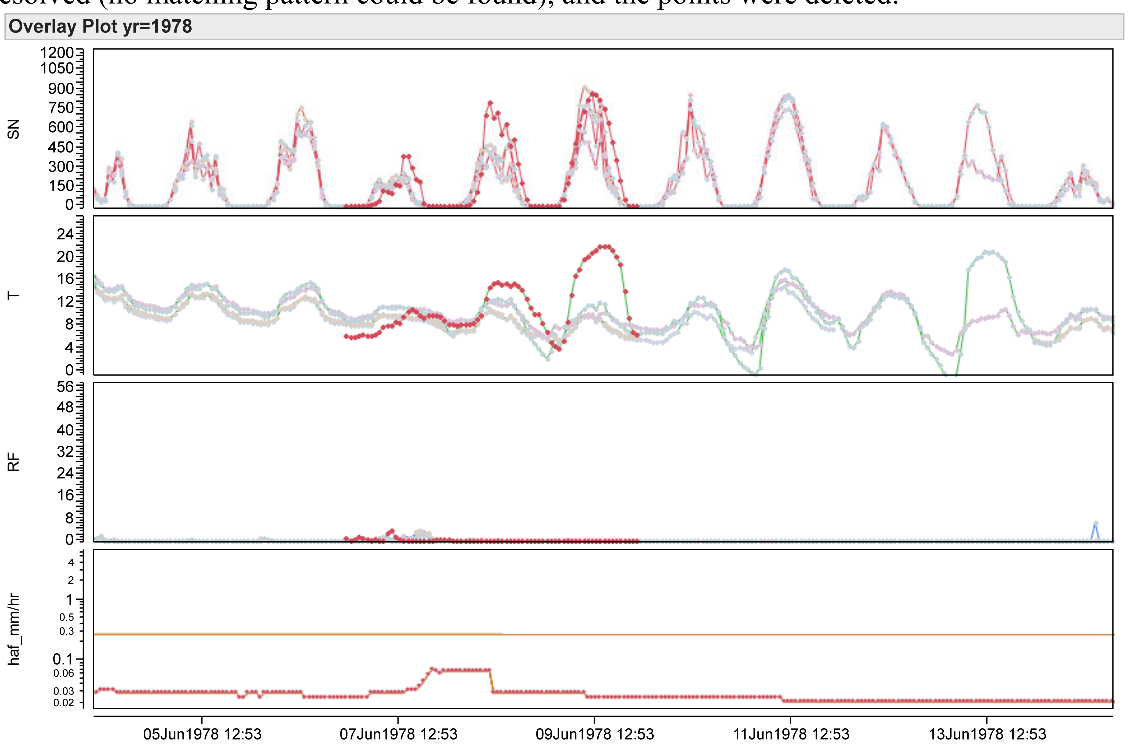
Overlay Plot yr=1976



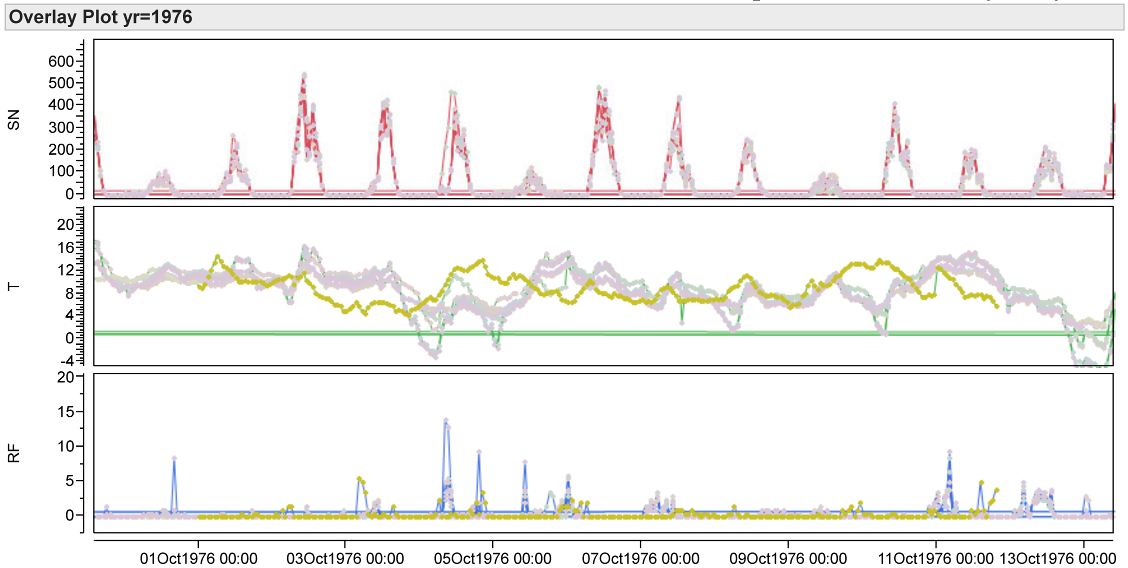
September 1976: From 07Sep1976 through 21Sep1976, CF2 is desynchronized from the other time series. This could not be resolved, and these points were deleted:



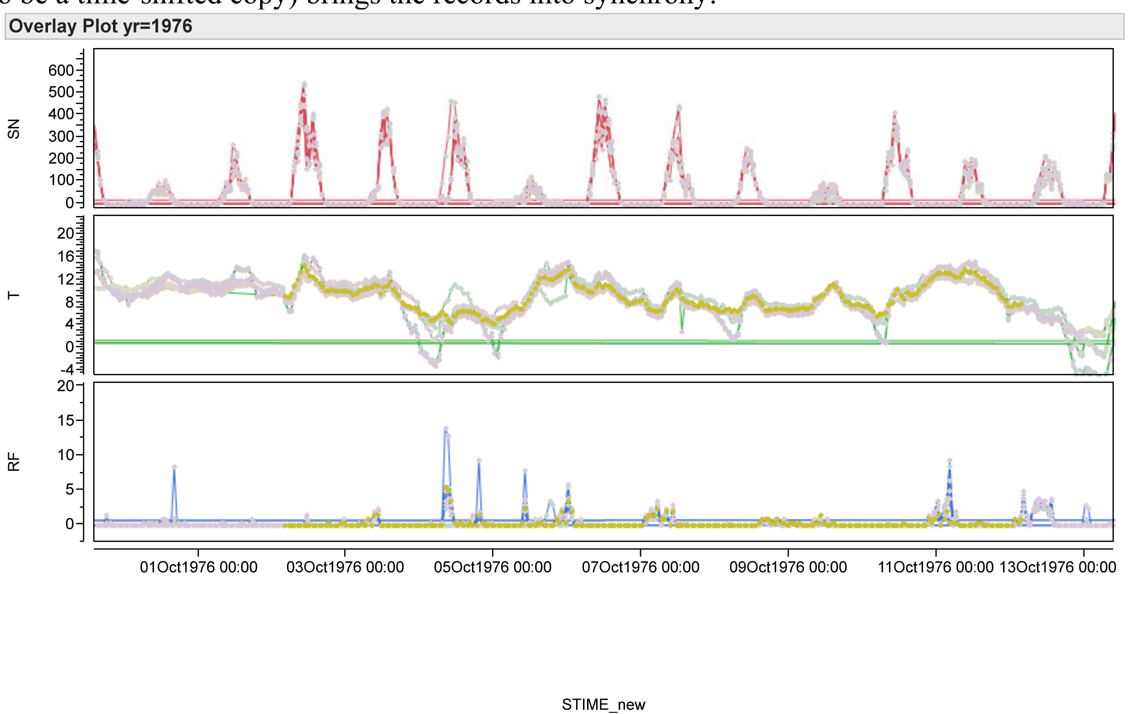
June 1978: CF is out of synch with the other sites from 7 to 9 Jun1978. This could not be resolved (no matching pattern could be found), and the points were deleted:



October 1976: From 01Oct1976 to 11Oct1976, EG2's timestamps are about one day early:

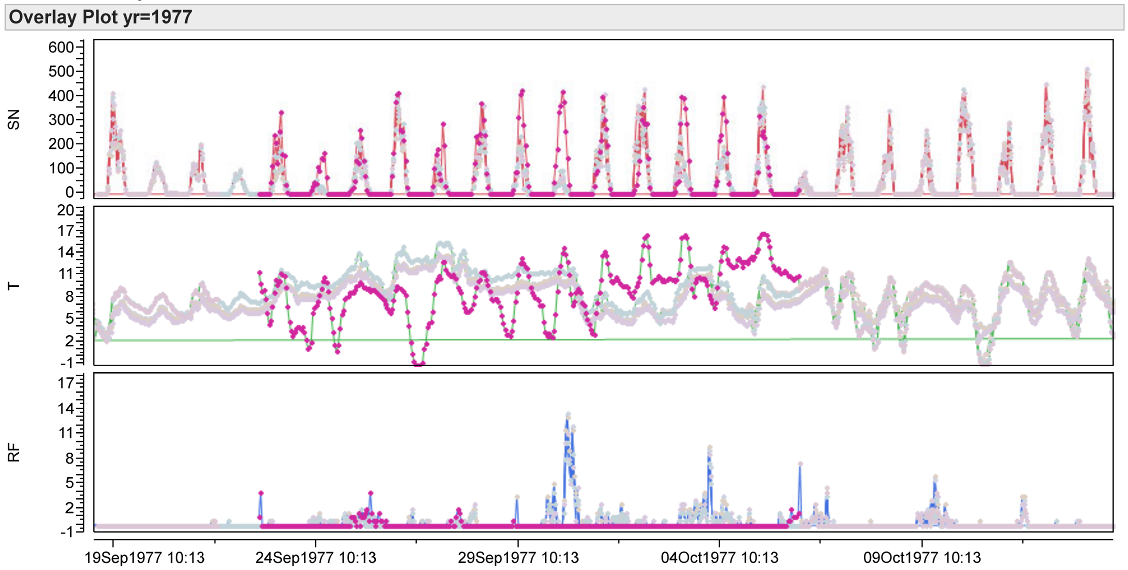


Applying a time shift of 28 hours, and deleting the last few points (which themselves appear to be a time-shifted copy) brings the records into synchrony:

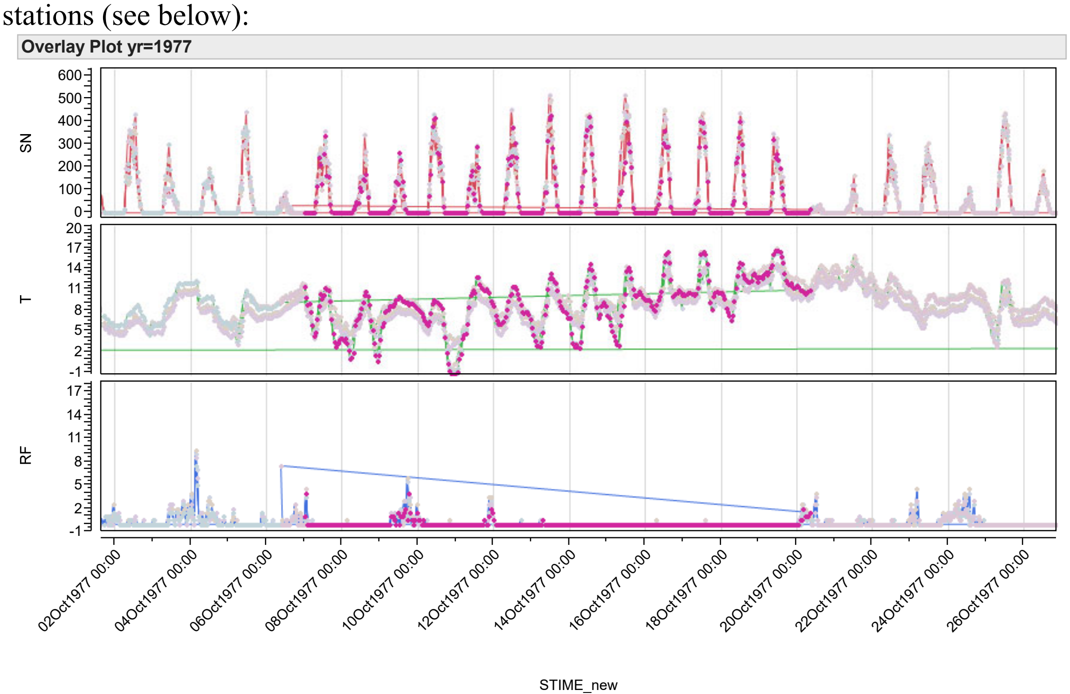


The timestamps have been corrected accordingly.

September 1977: Records from TN2 with timestamps from 23Sep1977 00h to 06Oct1977 08h are obviously inconsistent with the other AWS's:

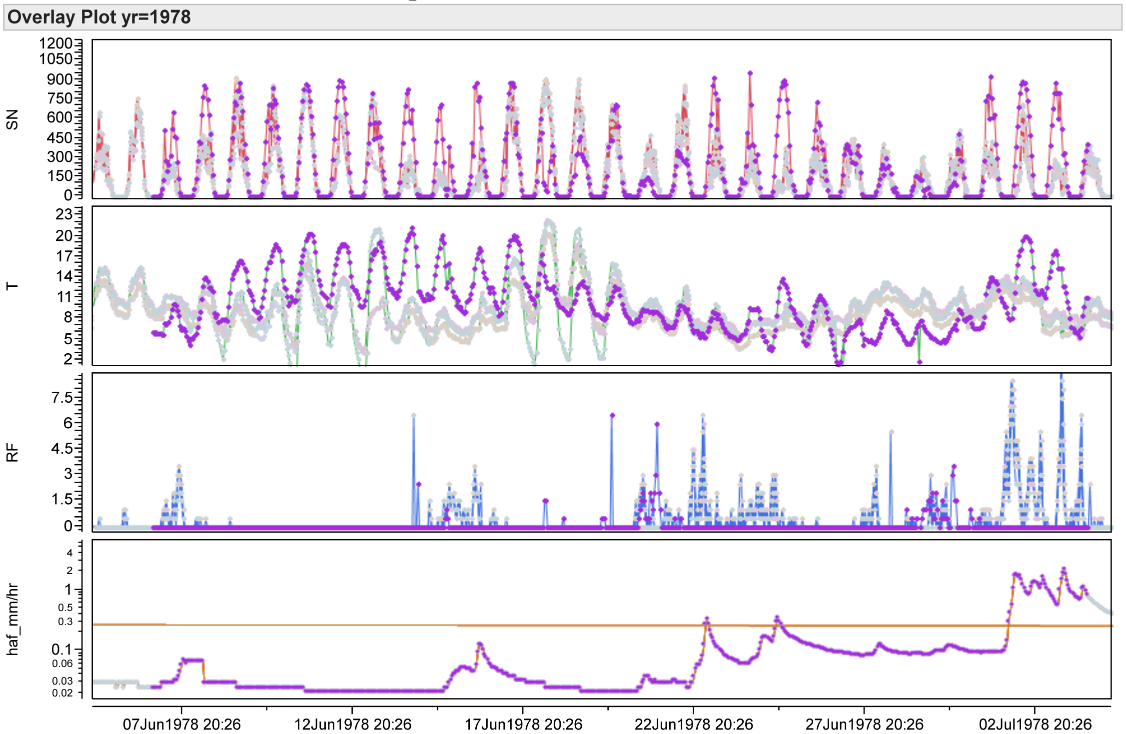


Shifting these values forward by 14 days brings them into good correspondence with the other stations (see below):

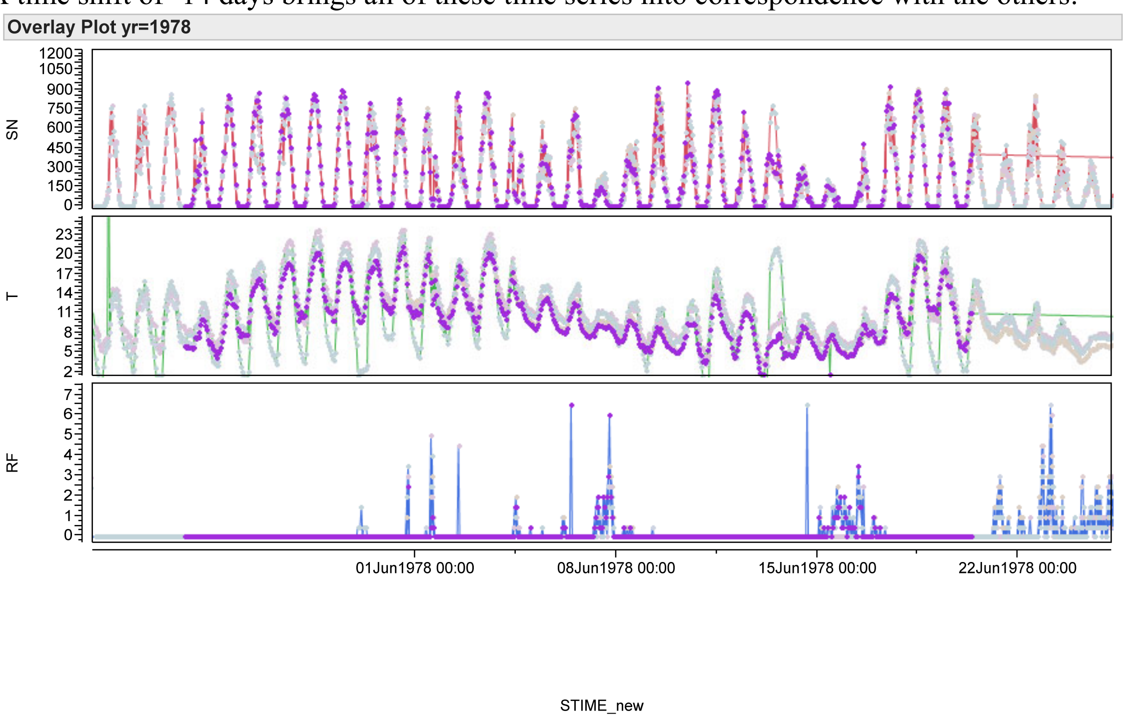


When time-shifted in this way, data for all sensors are numerically identical to TN2 records that already exist at the later times. From this I conclude that this block of records was somehow copied and given incorrect timestamps, two weeks earlier than the original timestamps. Because the misplaced data already exist with their correct timestamps, I have deleted the misplaced values.

June-July 1978: EG is badly out of agreement with the other AWS's from 07Jun1978 0h to 04Jul1978 9h. The SN, T, and RF patterns all are mismatched:

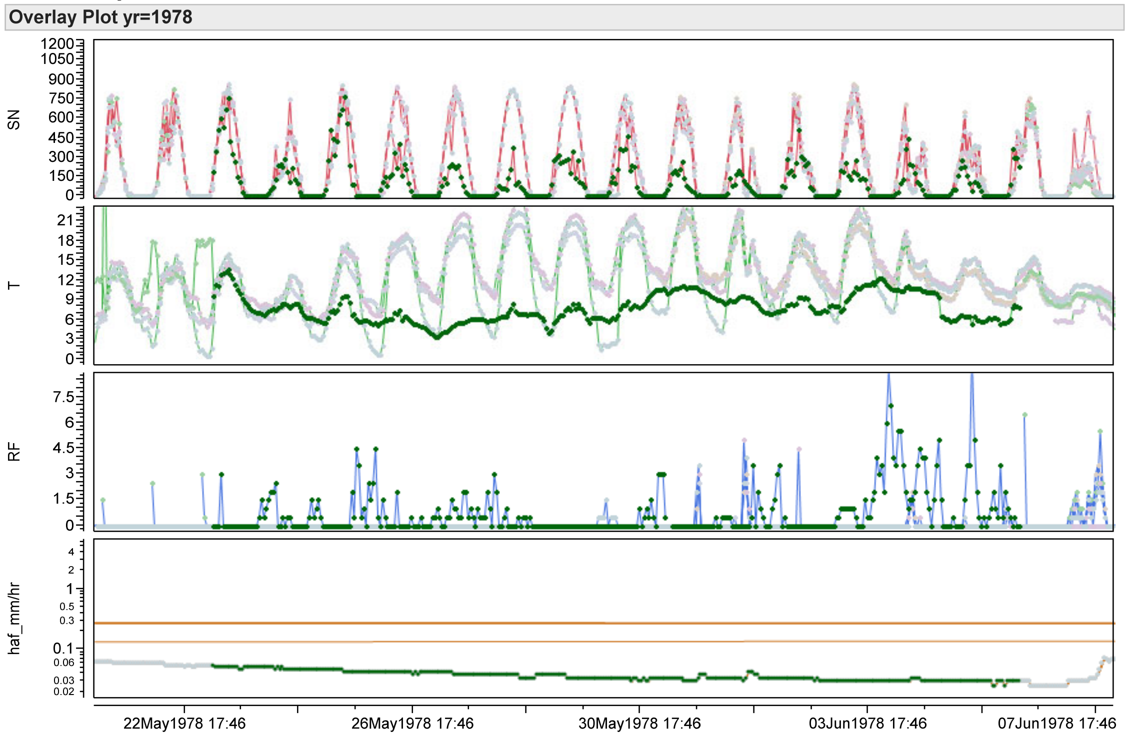


A time shift of -14 days brings all of these time series into correspondence with the others:

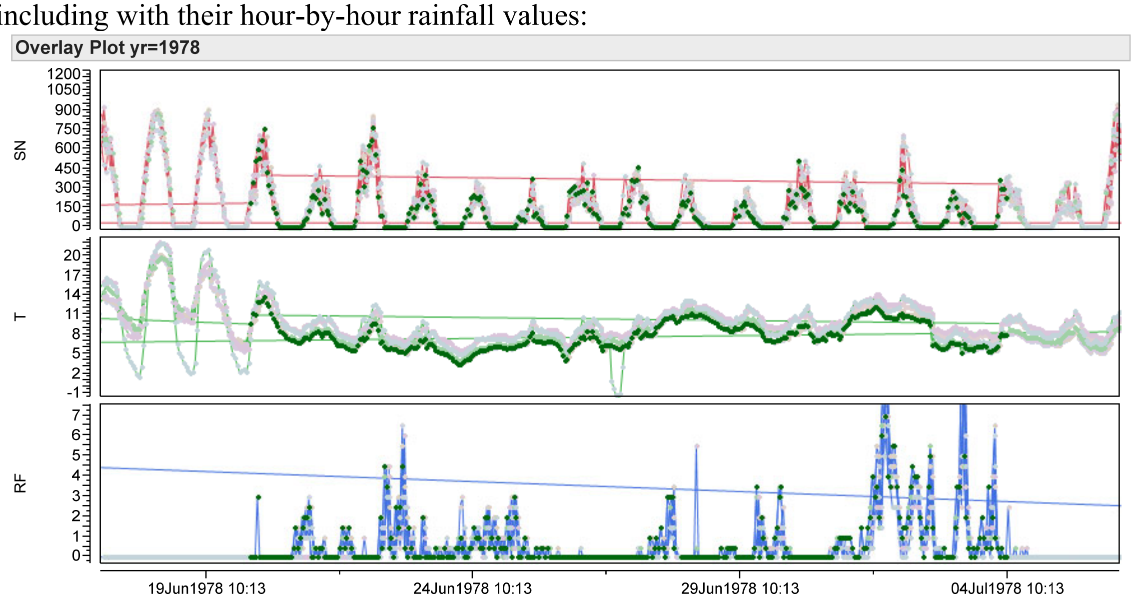


The timestamps have been adjusted acordingly.

<u>May-July 1978:</u> From 23May1978 06h to 06Jun1978 09h, EG2 is badly out of correspondence with the other AWS's. Among other things, it features a rainstorm that is not recorded by the other AWS's:

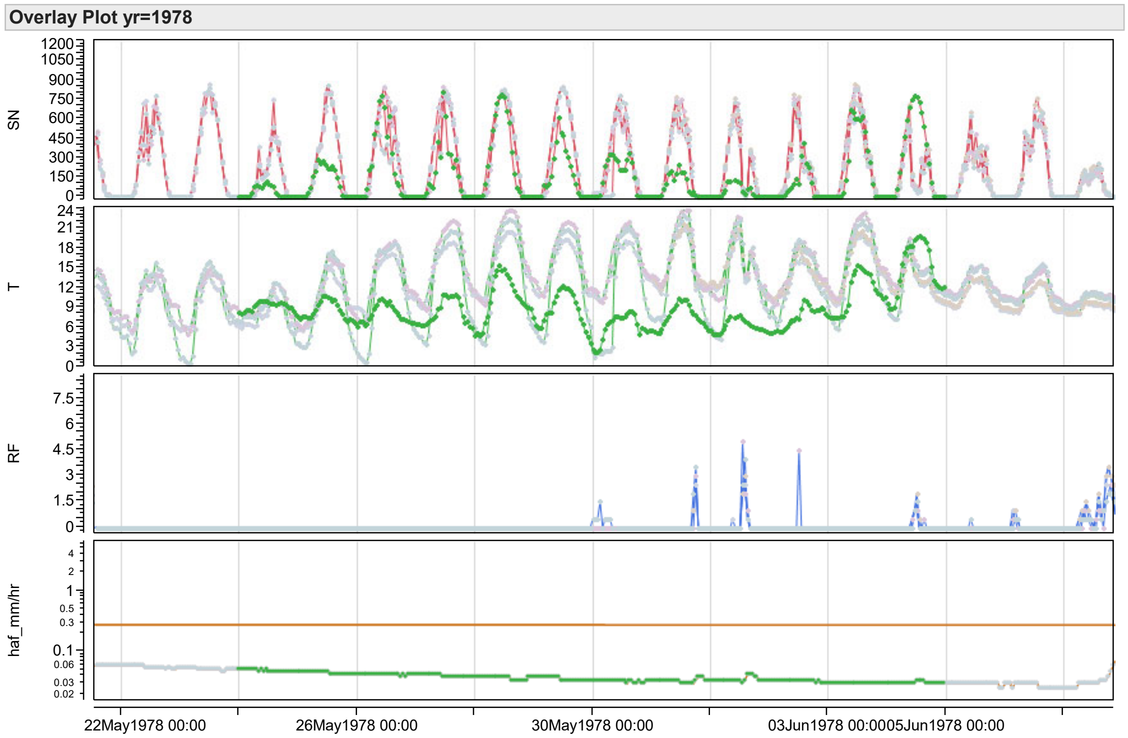


If this is shifted forward by 28 days it comes into exact correspondence with the other AWS's, including with their hour-by-hour rainfall values:

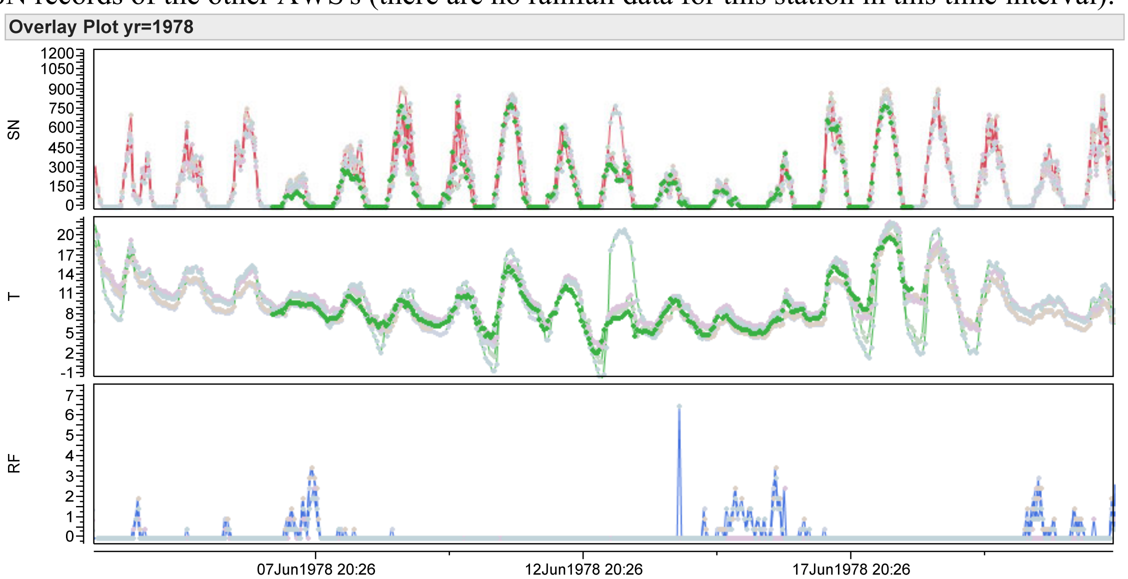


However, this cannot be EG2, because EG2 already has data over this interval, and their values are not exactly the same as the time-shifted "EG2" values. Instead, it appears that these are <u>*EG*</u> values, which were somehow mis-assigned to EG2, and to the wrong time interval. If these time-shifted "EG2" values are re-assigned to EG, they end exactly at the same time as the original block of EG values described on the previous page (which themselves have been time-shifted by 14 days in the opposite direction).

<u>May-June 1978:</u> From 24May1978 0h to 4Jun1978 23h, CF2 is badly out of correspondence with the other AWS's:

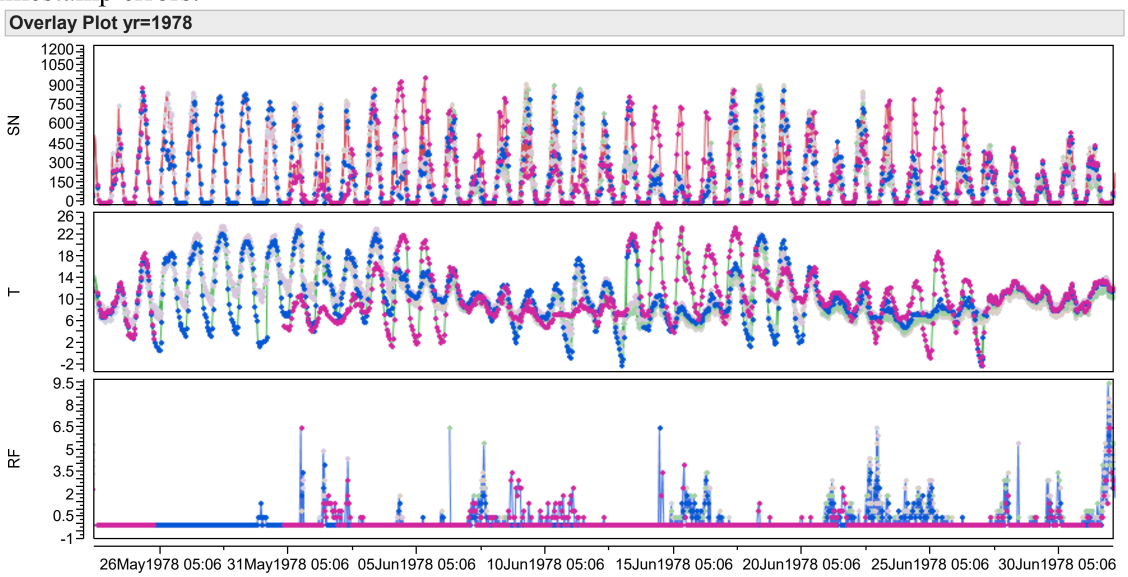


A time shift of 14 days plus 1 hour brings it into exact correspondence with both the T and SN records of the other AWS's (there are no rainfall data for this station in this time interval):

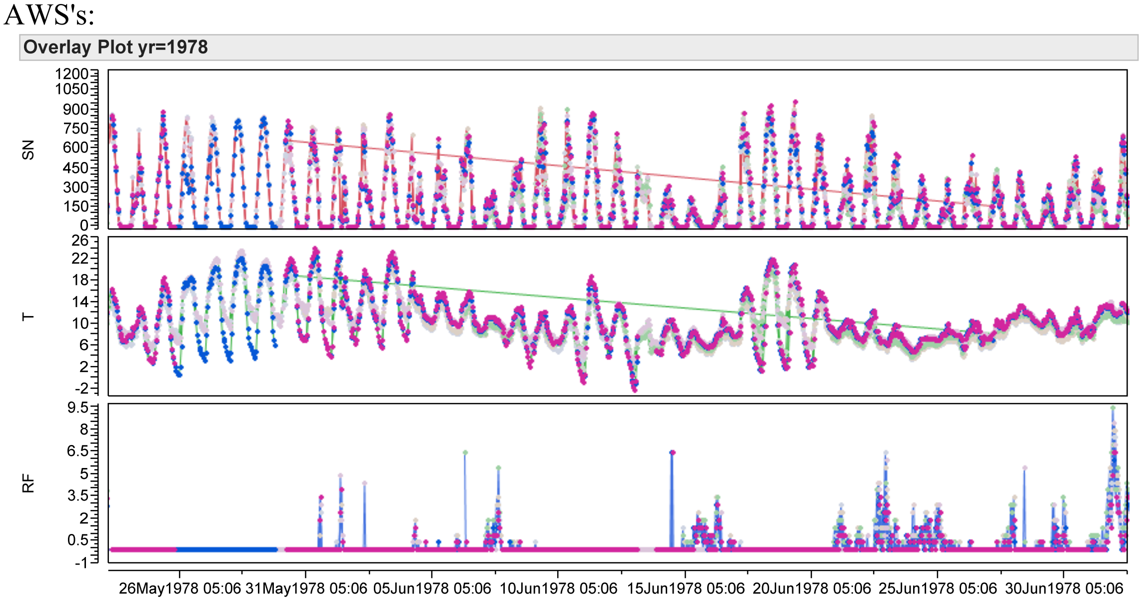


However, once again, this cannot be CF2 because there are already CF2 data during this interval. Instead, the numerical values coincide exactly with those of EG2 over this interval. Thus, apparently EG2 was copied with a time shift of -14 days plus -1 hour, and labeled as CF2 instead of EG2. Because these copied data still exist in EG2, they have been deleted from the CF2 time series.

June 1978: TN and TN2 (blue and red, respectively, in the figure below) show a complex and confusing pattern of variability in June 1978. Because TN and TN2 are sited in a tree-sheltered valley, they can exhibit wider temperature swings than the other AWS's can. However, there are also obvious mismatches in the solar flux data, strongly suggesting timestamp errors:

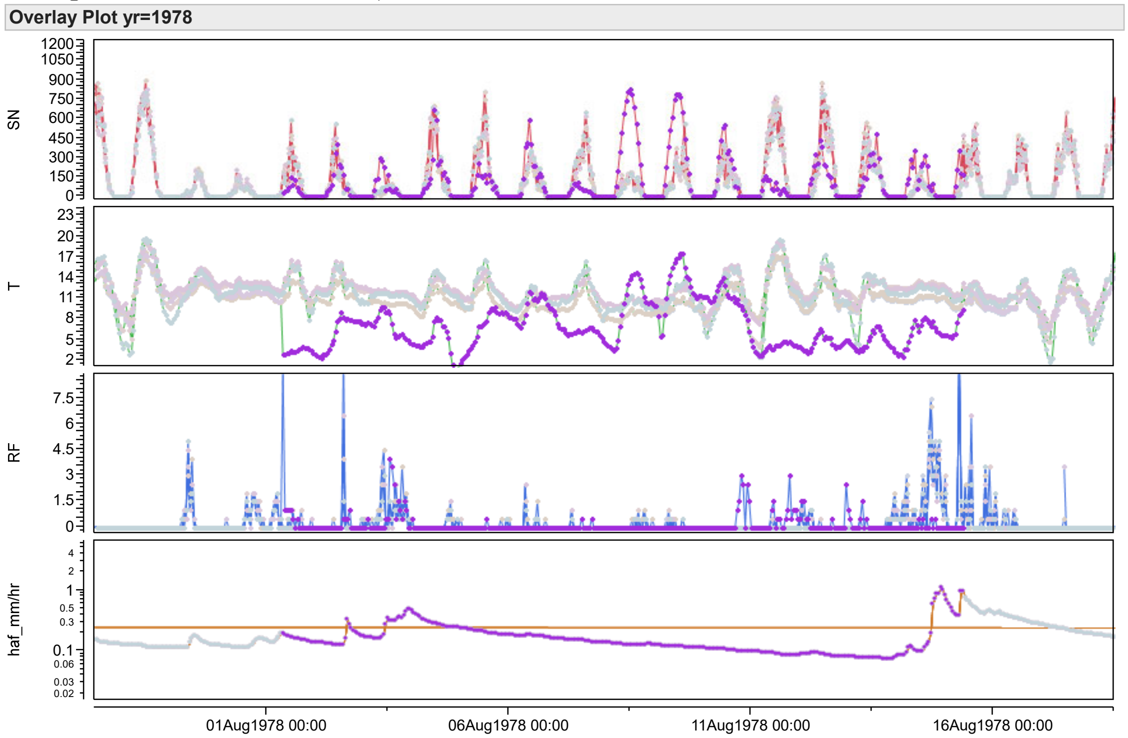


There are four different things going on here. First, the TN2 points that appear from 13Jun78 09h to 27Jun78 08h actually originate 14 days earlier, and need to be shifted back. Second, the TN2 points that appear from 31May78 0h to 13Jun78 8h actually originate 14 days later, and need to be shifted forward. Third, the TN values from 8h to 23h on 13Jun78 are a copy of identical points that originate from 09h to 24h on 30May78; because the originals are still there, these copies need to be deleted. Lastly, the TN values from 00h to 08h on 30May78 and from 00h to 07h on 27Jun78 are clearly out of place, and have been deleted. When all of these changes have been made, TN and TN2 are consistent with each other, and with the other AWS's:

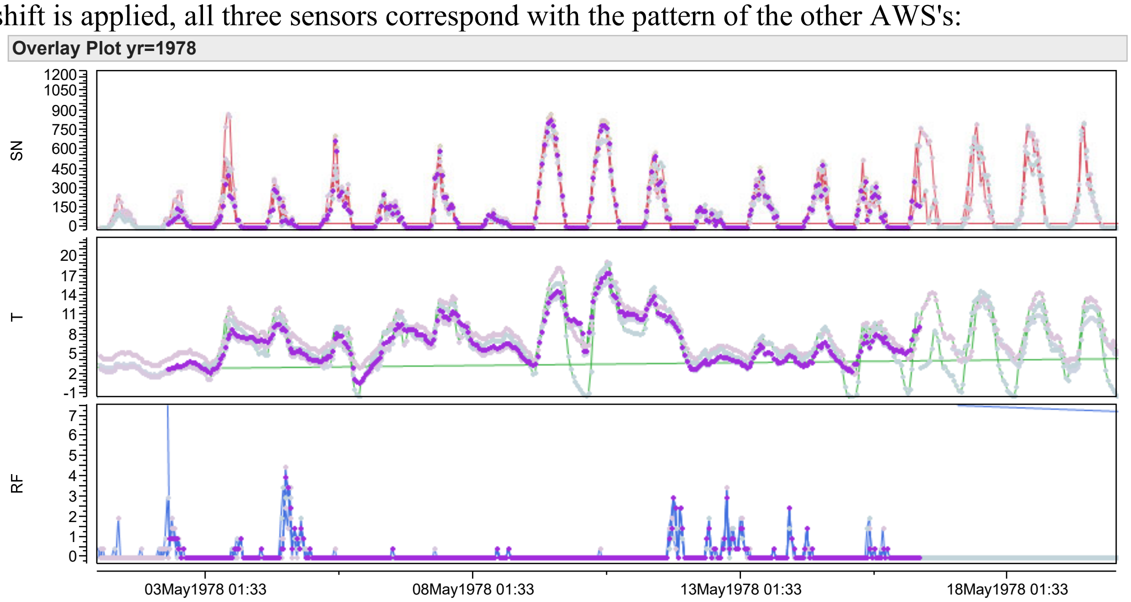


The timestamps have been corrected accordingly.

August 1978: From 01Aug1978 08h to 15Aug1978 09h, EG is badly out of correspondence with the other AWS's, in RN, T, and RF (note that the rainstorm on 14-15Aug is entirely missing from the EG time series):

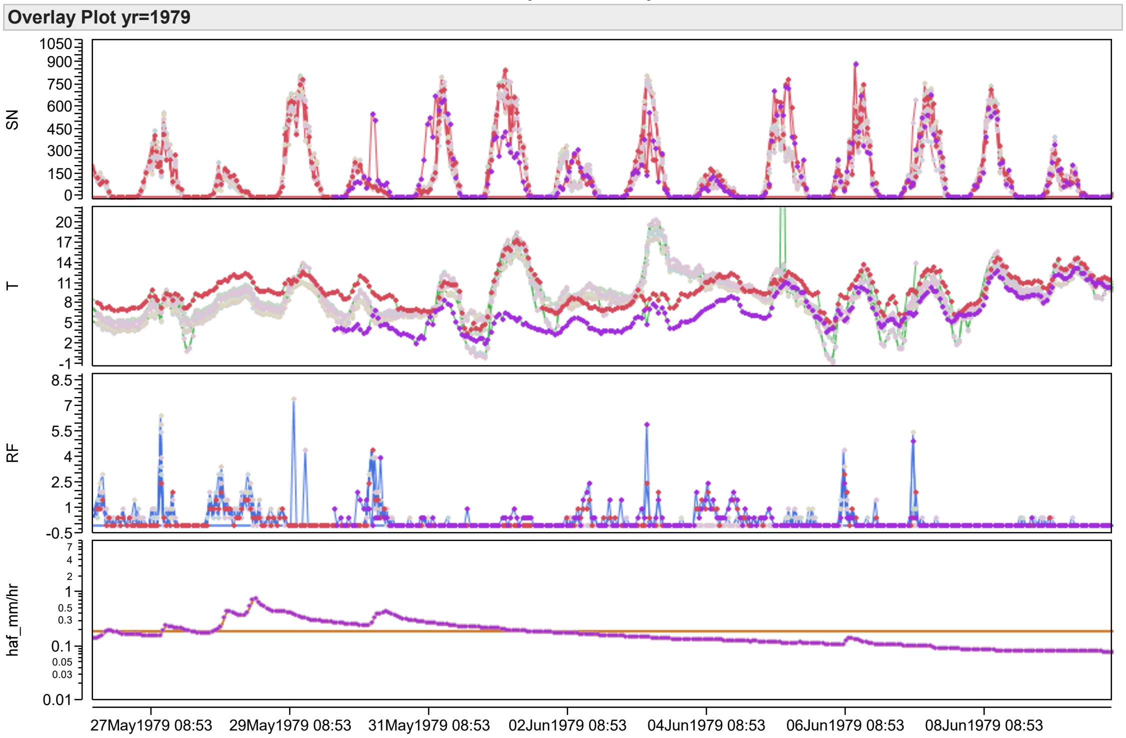


In this case, the necessary time shift is -91 days (from August 1 back to May 2). When this shift is applied, all three sensors correspond with the pattern of the other AWS's:

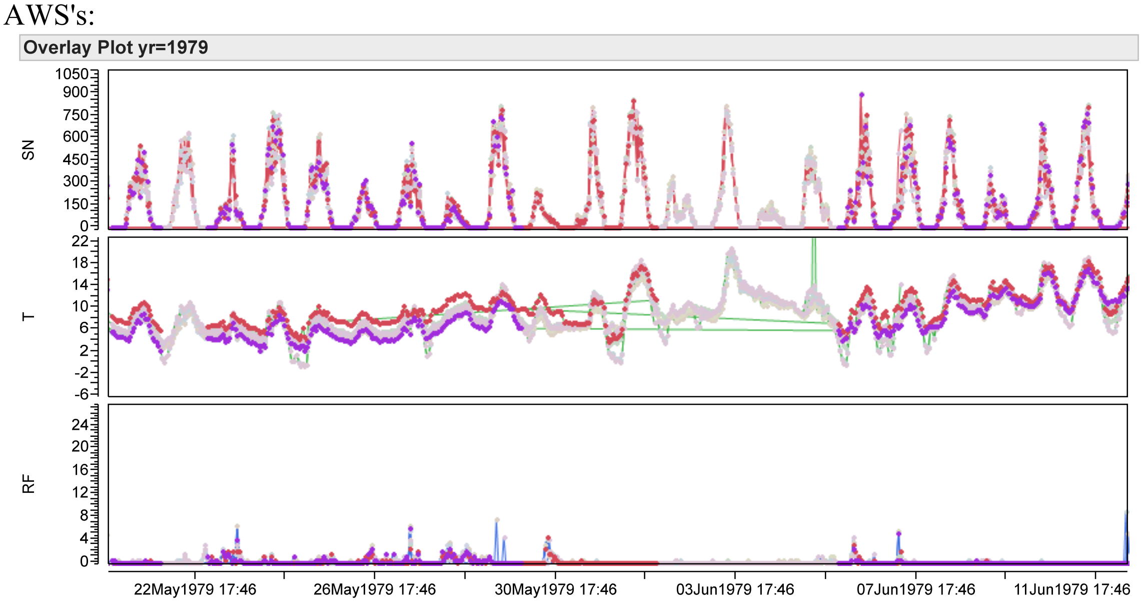


However, the resulting records cannot belong to EG, because there are already records for EG during in early May, and their numerical values are slightly different. Instead, the time-shifted "EG" records correspond exactly to the numerical values of *CW2*, at the timestamps where they overlap. Thus a block of CW2 data has been shifted by three months (minus one day) and mis-assigned to EG instead. Where the "original" CW2 data still exist, they have been retained; where they do not, the corresponding EG values have been time-shifted back to their original positions and the AWS ID has been reassigned to CW2.

May-June 1979: Both CF and EG (red and purple, respectively, in the figure below) are inconsistent with the other AWS's in late May and early June of 1979:

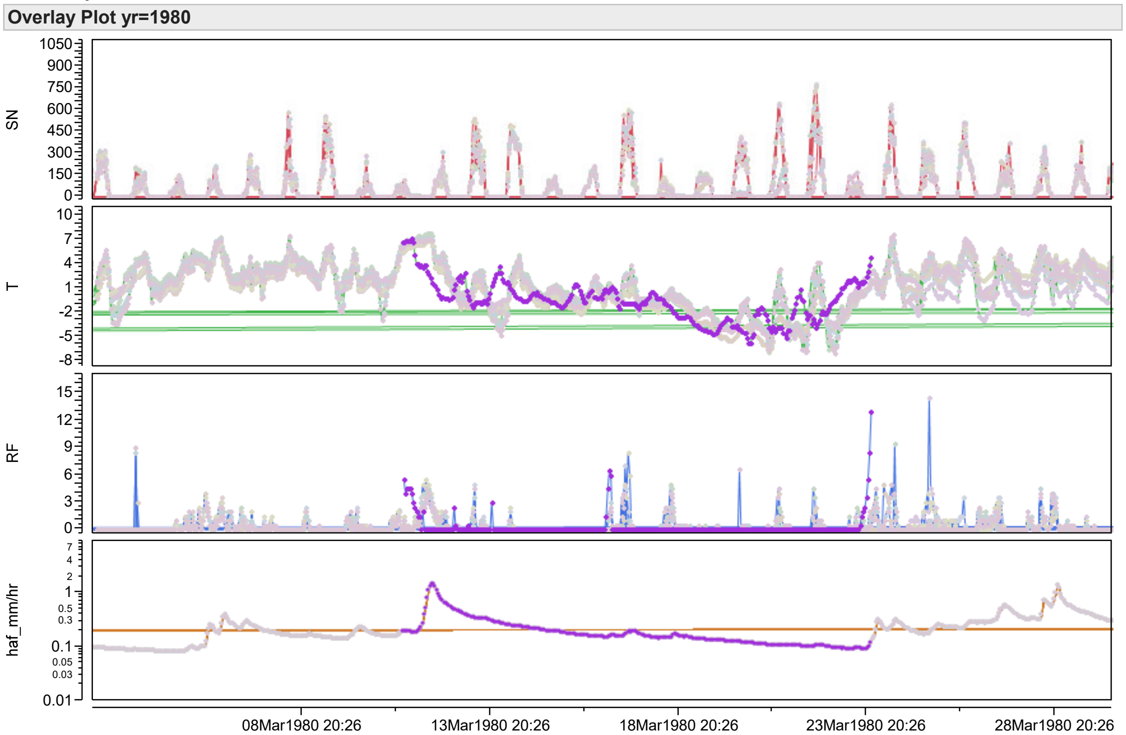


By pattern matching, one can determine that EG values with timestamps from 30May1979 00h to 05Jun1979 23h have been shifted 7 days forward in time, and that CF values with timestamps from 02Jun1979 to 05Jun1979 23h have also been shifted 7 days forward. When these shifts are reversed, one sees a very good match between CF and EG and the rest of the AWS's:

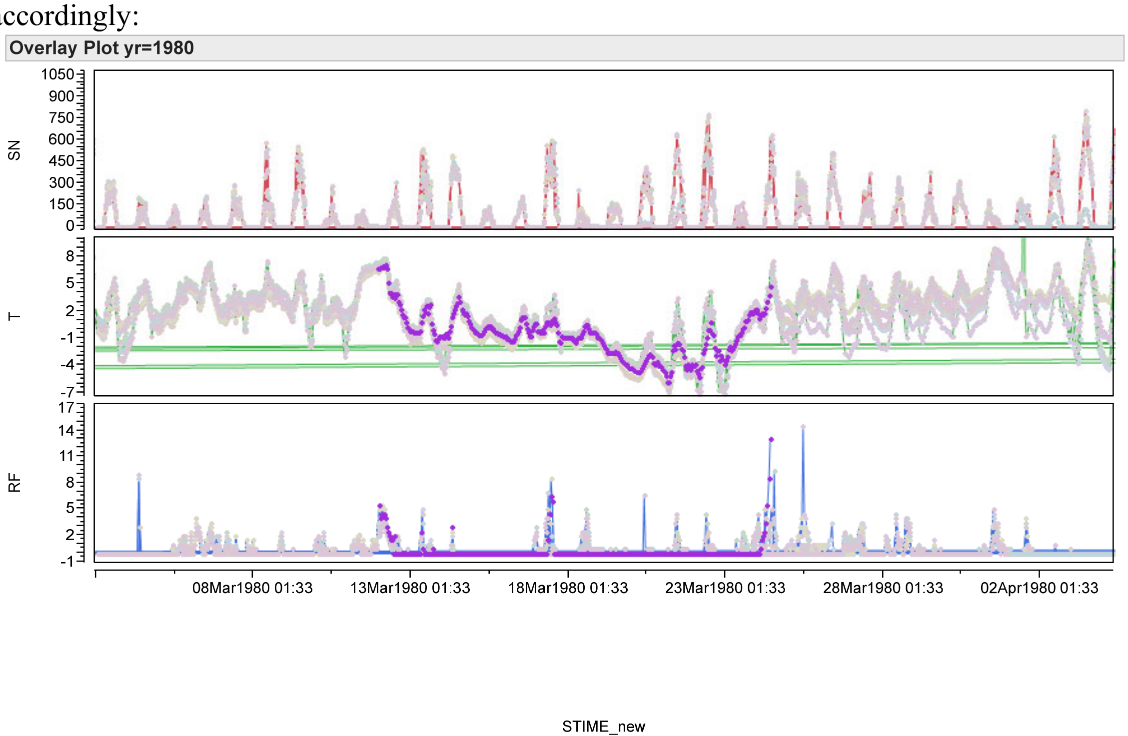


The time shift in the EG data has been reversed. The CF data already exist with their original time stamps, so the time-shifted copies have been deleted.

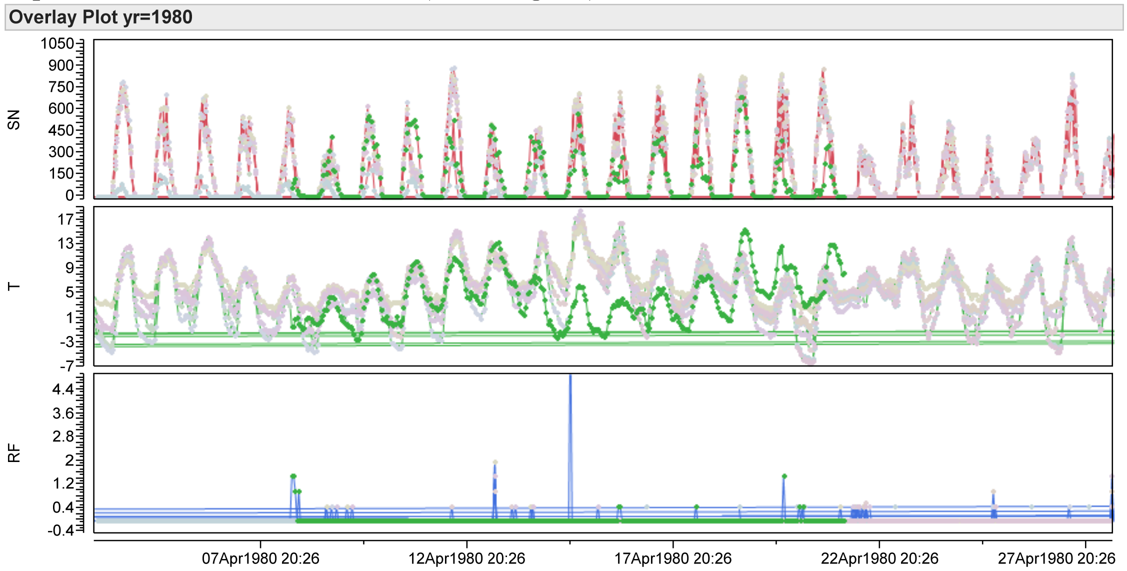
March 1980: EG values with timestamps from 11Mar1980 12h to 23Mar1980 23h have evidently been shifted 12 hours back in time:



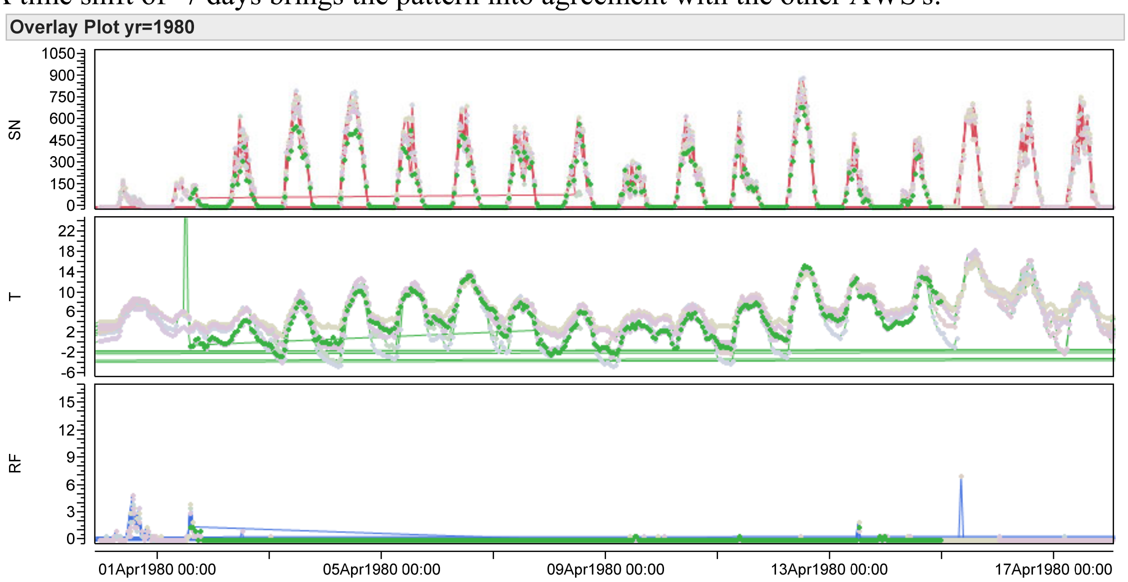
When this time shift is reversed, the temperature variations and rainfall pulses are synchronized exactly with the other AWS's, so the timestamps have been corrected accordingly:



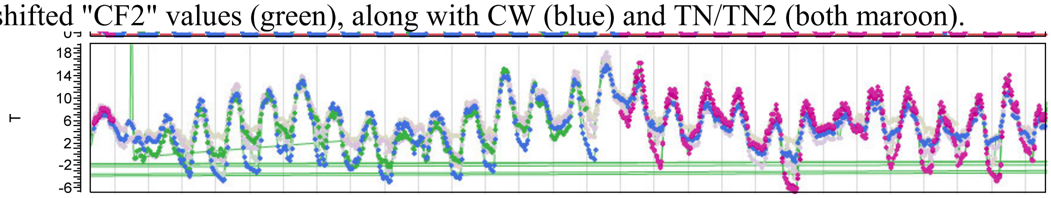
April 1980: CF2 values with timestamps from 14h on 08Apr1980 to 23h 21Apr1980 do not fit the pattern of the rest of the AWS's (including CF):



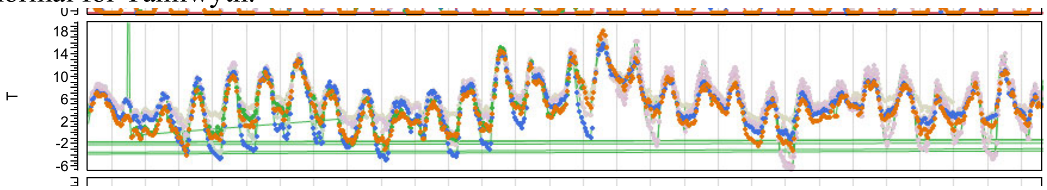
A time shift of -7 days brings the pattern into agreement with the other AWS's:



However, there are already records for CF2 during part of interval, and they are not identical to the misplaced records. The natural conclusion is that, not only have the timestamps been shifted, but the station has been mis-assigned. There are already also records for CF during part of this interval. Here is what I think has happened. The figure below shows the time-shifted "CF2" values (green), along with CW (blue) and TN/TN2 (both maroon).



In the second half of the graph, CW (blue) lies mostly well inside the large diurnal swings of TN and TN2 (these are distinctive of the Tanllwyth site). However, in the first half of the graph, it is CW that has the very large diurnal temperature swings, which are not at all distinctive of the Carreg Wen site. During this period, the "CF2" values behave much like the CW values in the second half of the graph. In the graph below, we see CW in blue, "CF2" in green, and EG in orange. In the second half of the graph, CW closely follows EG, as is usual. In the first half of the graph, however, "CF2" closely follows EG and CW swings widely, as is normal for Tanllwyth.

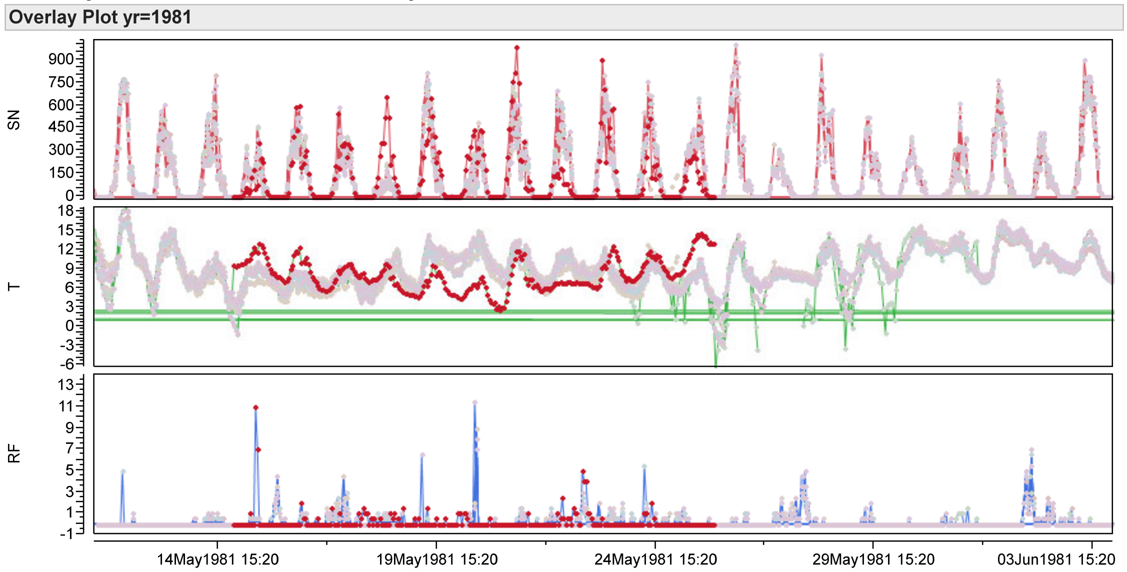


Finally, in the first half of the month, night-time wind speeds in the CW records regularly dip to near zero, which is highly uncharacteristic of Carreg Wen but characteristic of Tanllwyth. Starting roughly April 16th, night-time wind speeds in the CW records are rarely near zero (normal for Carreg Wen) but Tanllwyth wind speeds are frequently near zero.

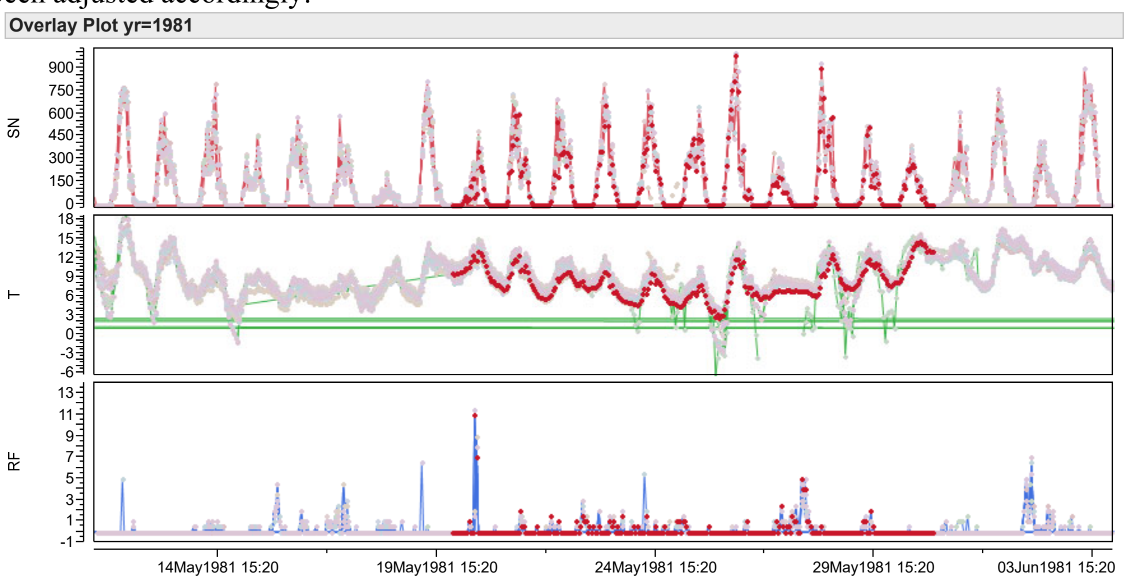
From all of these lines of evidence, I conclude that the records labeled as "CW" are really TN2 (not TN, because TN's wind speeds are much more noisy), from 09h on 01Apr1980 through 07h on 15Apr1980. I also conclude that the time-shifted records labeled as "CF2" are really CW2 from 14h on 08Apr1980 to 23h 21Apr1980 in the original, incorrect timestamps, which equals the interval from 14h on 01Apr1980 to 23h 14Apr1980 in the corrected timestamps. I conclude that "CF2" is really CW2 rather than CW, because (a) there is a gap in CW2 that ends at precisely 00h 15Apr1980, (b) the temperature, wet-bulb depression (TD), and windspeed are all fully consistent with the end of the "CF2" interval, and (c) the windspeed sensor at CW was apparently not functioning in early April, but windspeed data are found in "CF2". I have changed the site designations accordingly (however, I have not changed the SITENO field in these records).

<u>April 1981:</u> From 10h on 12Apr1981 to 23h on 17Apr1981, TN2 values are inconsistent with the rest of the AWS's. These values appear to have been copied from early March. Since they are copies, they have been deleted:

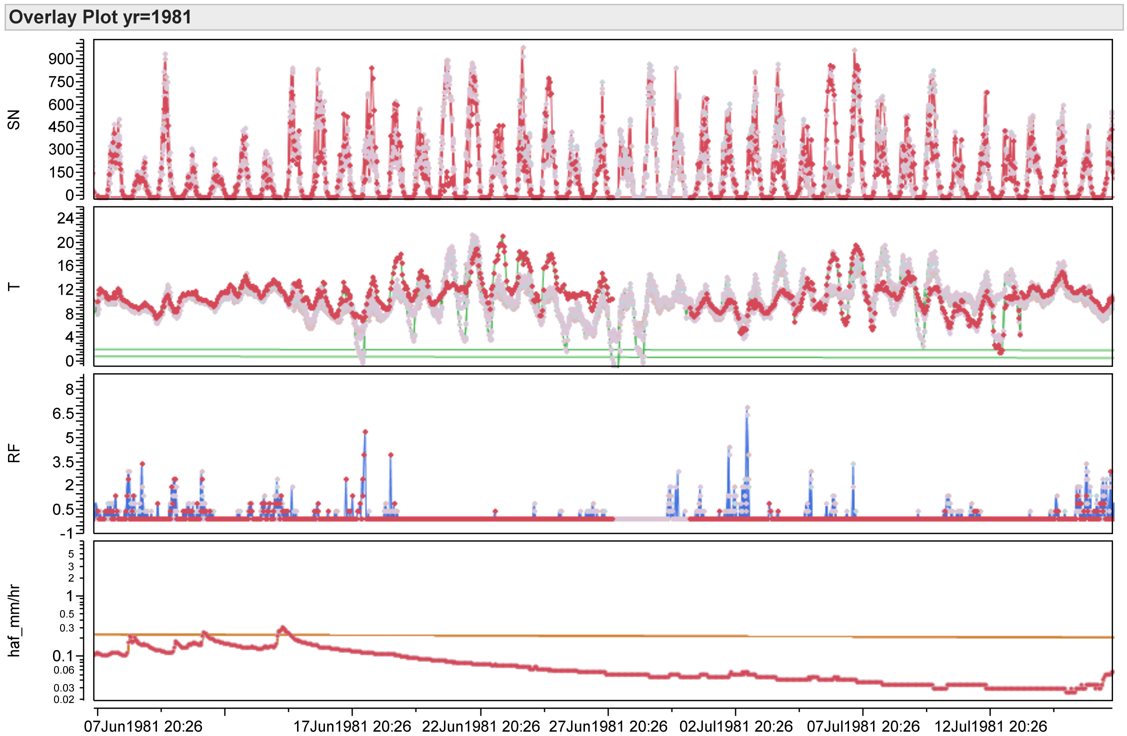
May 1981: EG2 values with timestamps from 00h on 15May1981 to 23h on 25May1981 have evidently been time-shifted five days earlier than their correct dates and times:



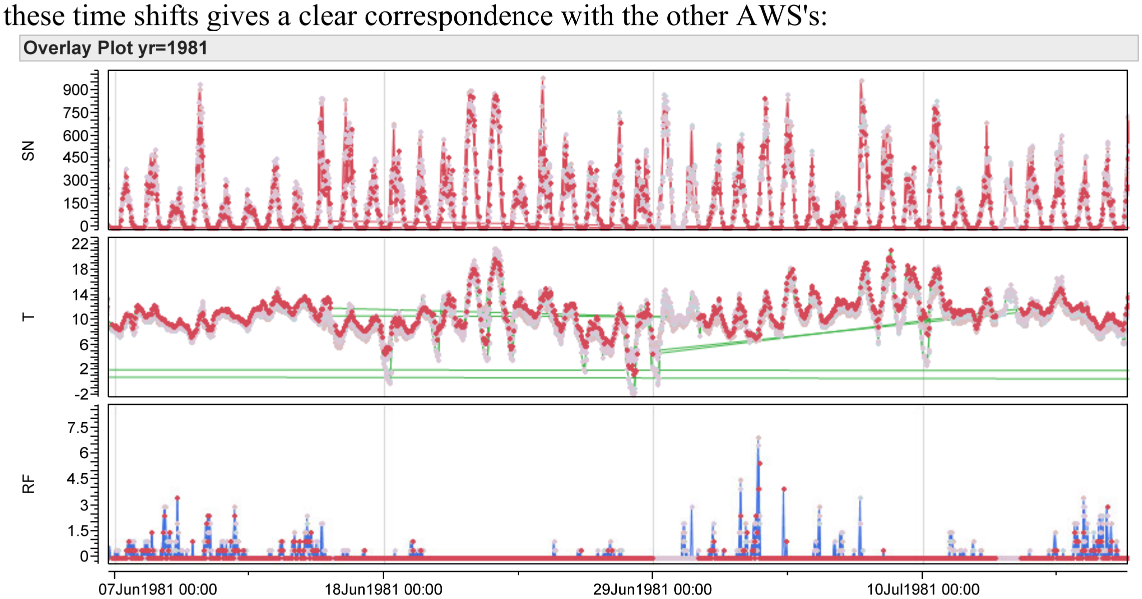
Reversing this time shift gives a very good match to the other AWS's, so the timestamps have been adjusted accordingly:



June-July 1981: Here, as in June 1978, several messy things are going on simultaneously. First, CF is inconsistent with the other AWS's in two blocks:

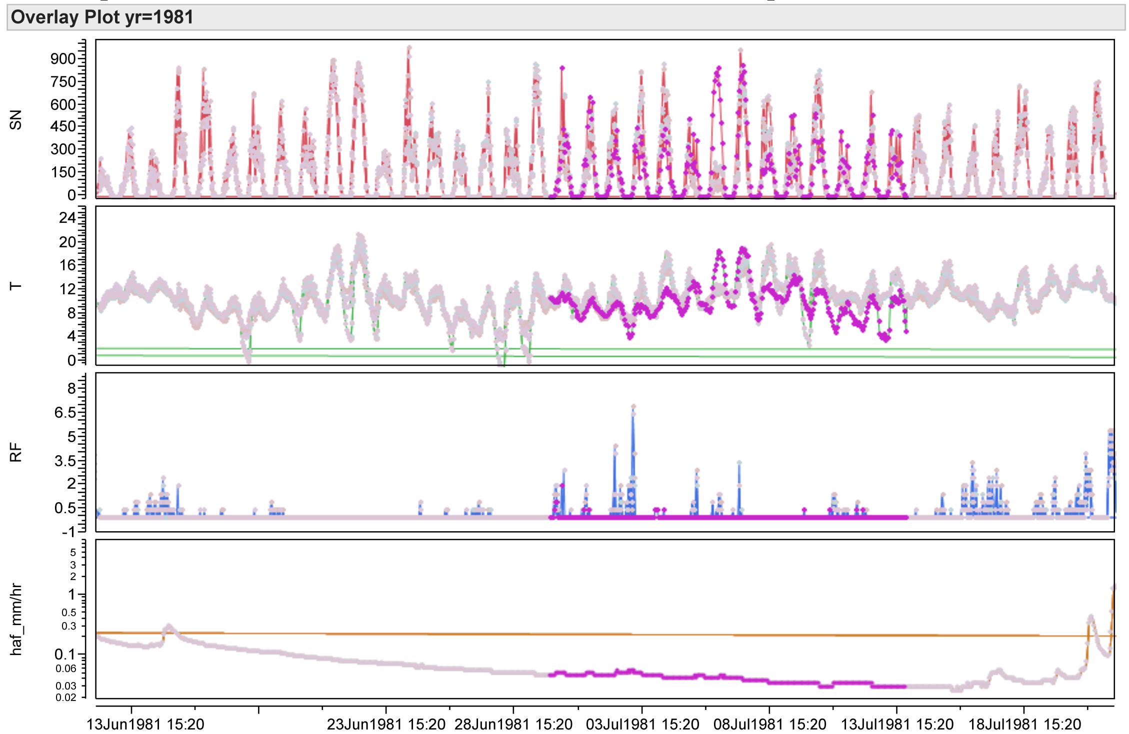


What has happened is that CF records with timestamps from 15Jun1981 20h to 27Jun1981 23h have been shifted back in time by 15 days, and CF records with timestamps from 01Jul1981 00h to 13Jul1981 23h have been shifted forward in time by 15 days. Reversing these time shifts gives a clear correspondence with the other AWS's:

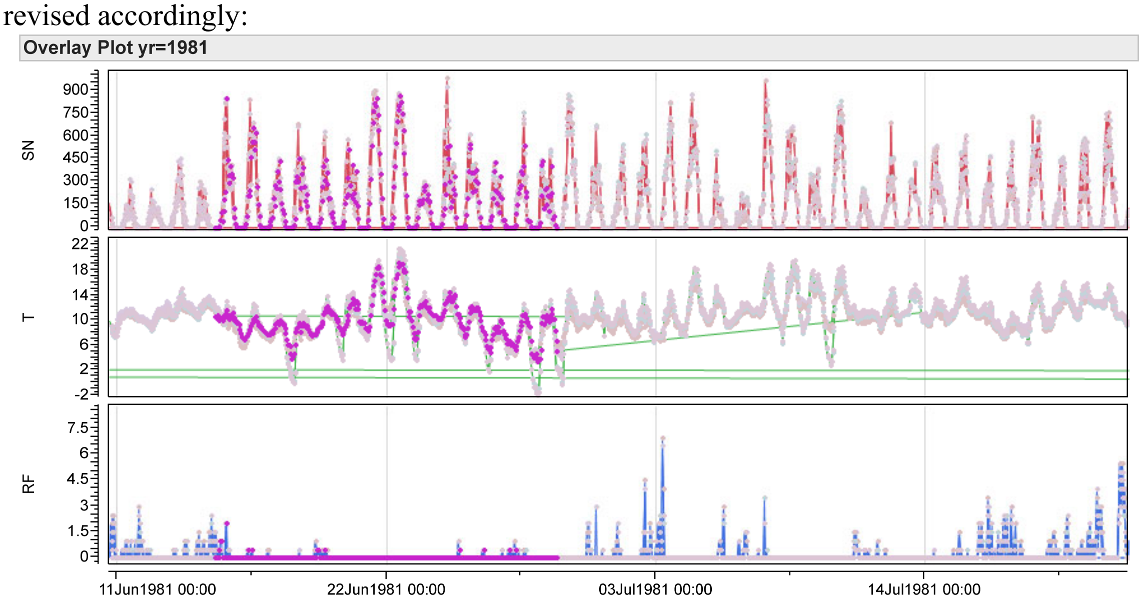


The timestamps have been corrected accordingly.

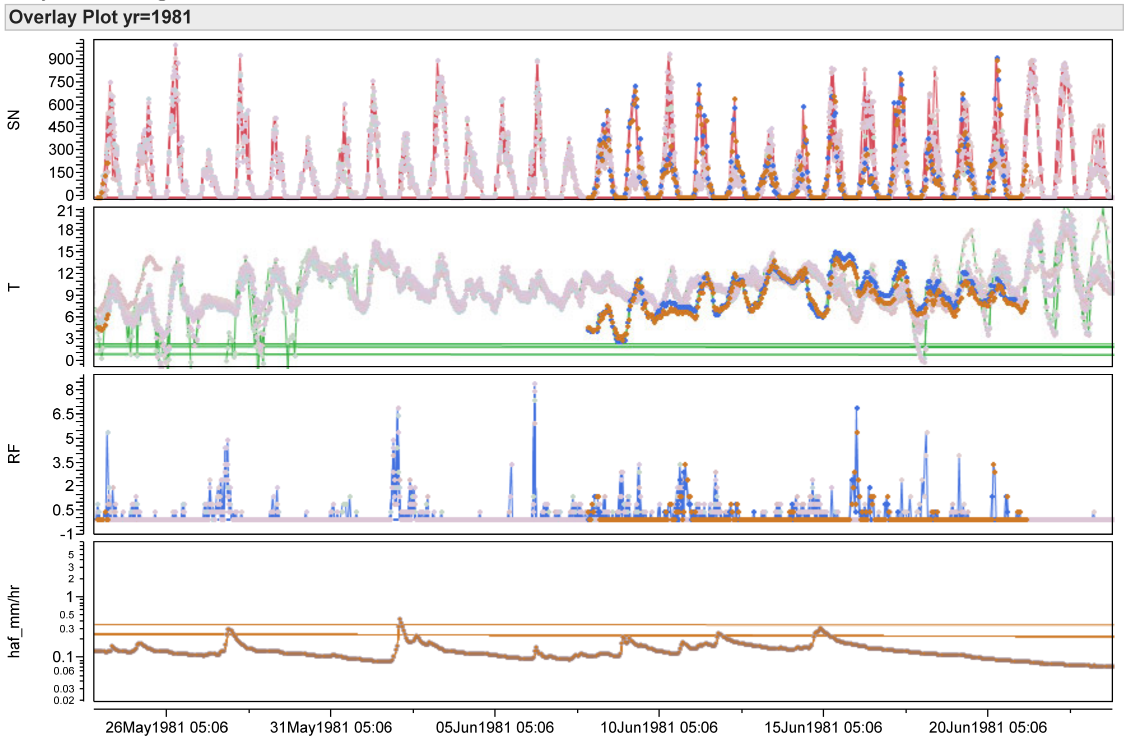
June-July 1981: In addition to the block shifts in CF as noted above, also the NI records with timestamps from 30Jun1981 00h to 13Jul1981 23h do not correspond to the other AWS's.



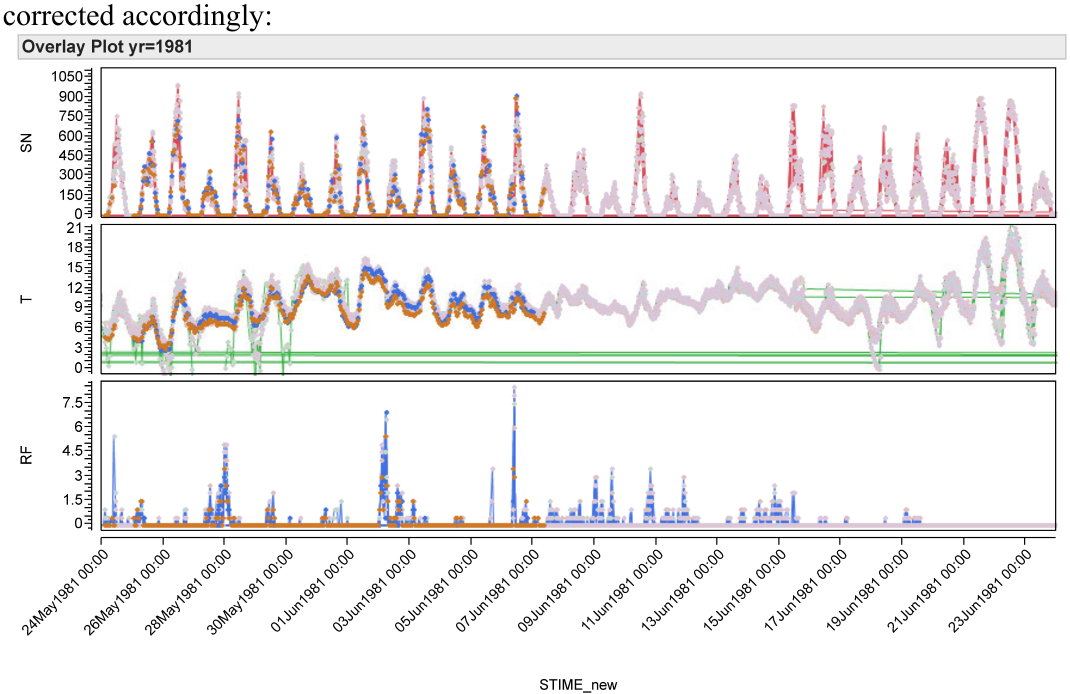
These records have apparently been shifted forward in time by 15 days. When this shift is reversed, they correspond almost exactly to the other AWS's, so the timestamps have been revised accordingly:



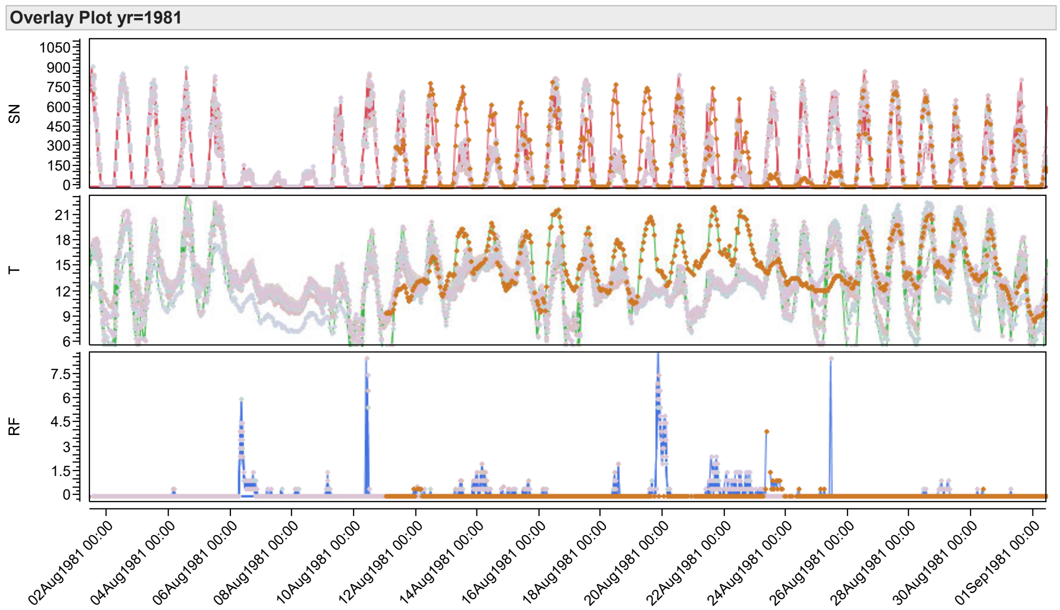
June 1981: Both CW and CW2 with timestamps from 08Jun1981 00h to 21Jun1981 08h are badly out of agreement with the other AWS's:



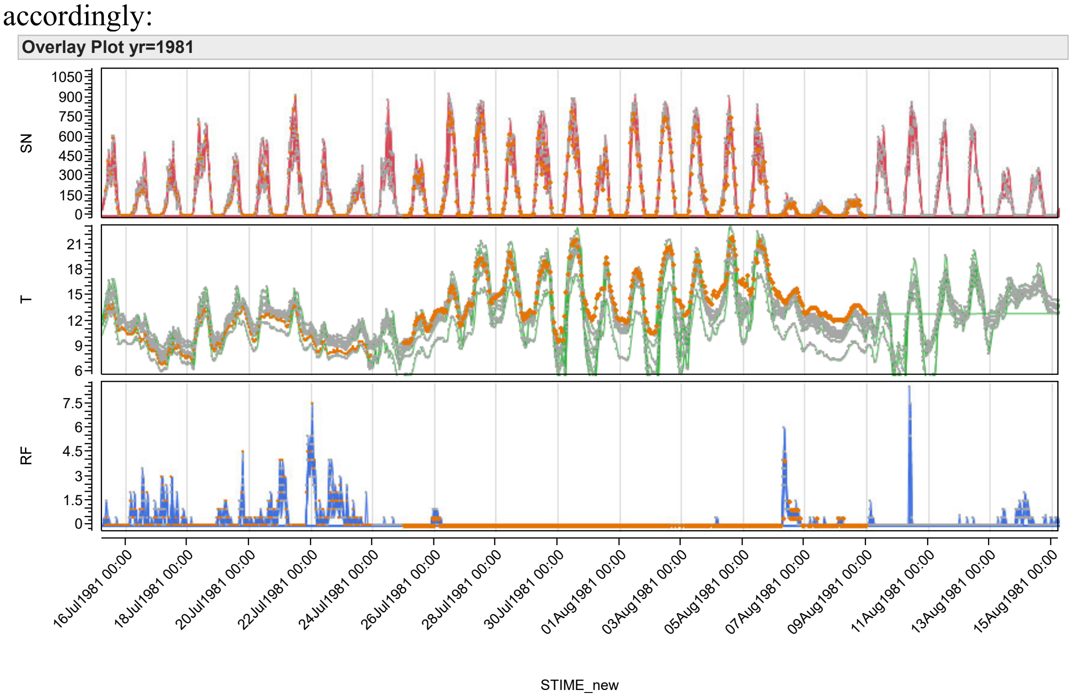
Applying a time shift of -13 days plus -23 hours (for CW) and -14 days (for CW2) brings these AWS's into exact synchrony with the other stations, and the timestamps have been corrected accordingly:



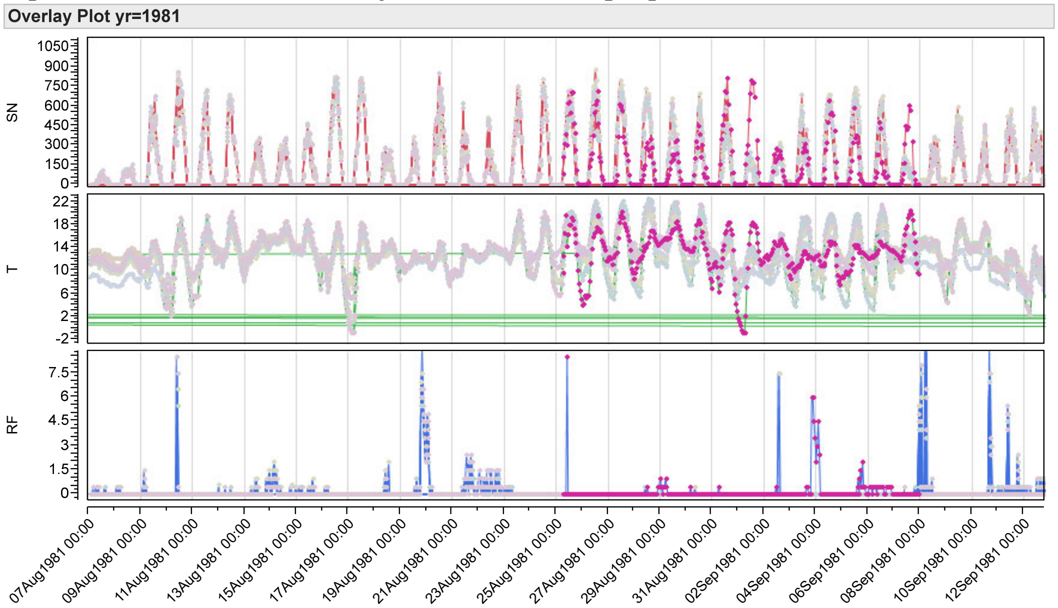
August 1981: For timestamps from 11Aug1981 00h to 25Aug1981 23h, CW2 is inconsistent with the other AWS's:



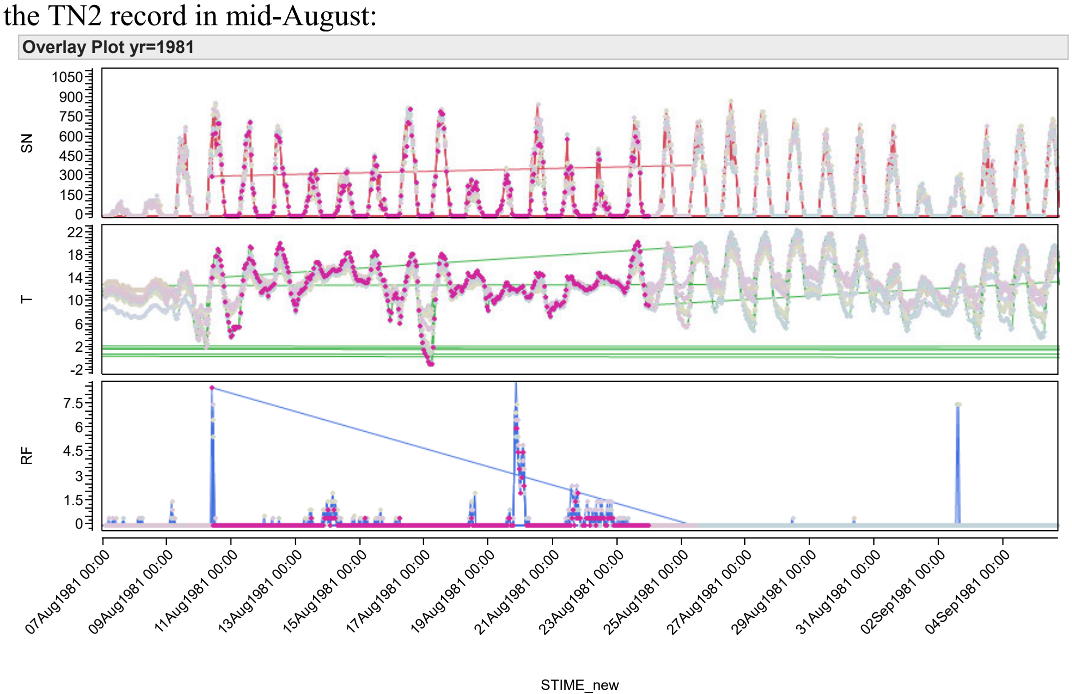
Applying a time shift of -17 days brings CW2 into synchrony with the other records, including with the rainfall event on 06 August, so the timestamps have been corrected accordingly:



<u>September 1981:</u> The pattern in TN2 with timestamps from 25Aug1981 10h to 07Sep1981 23h is inconsistent with the other AWS's (including TN). Although the temperature and solar flux data do not look too inconsistent during 25-26 August, note the 8 mm spike of rain on 25 August that does not occur in any of the other rain gauges:

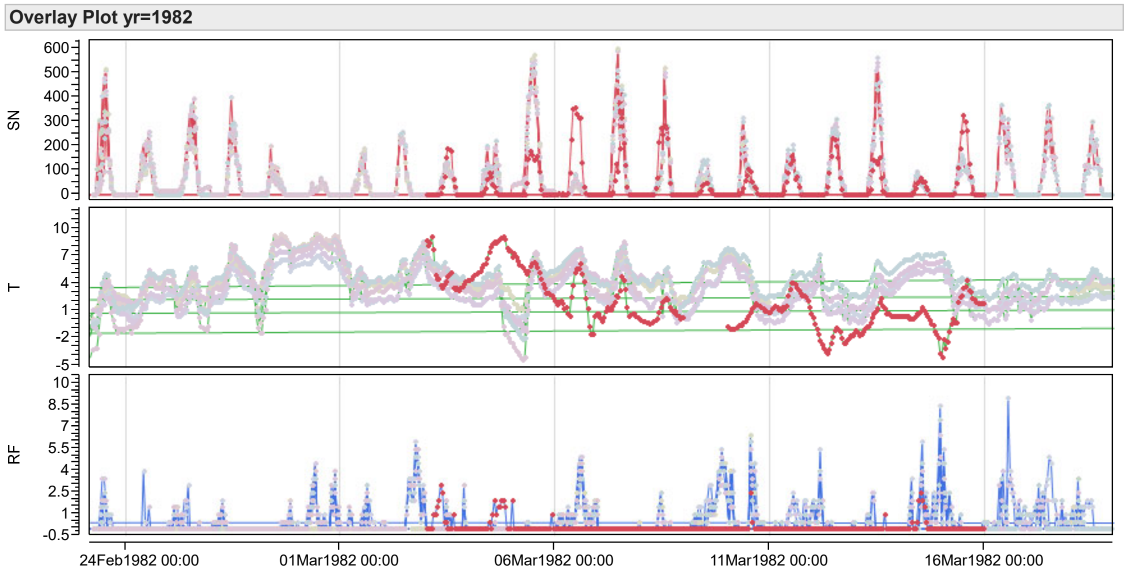


A time shift of -15 days plus -1 hour brings TN2 into exact correspondence with a segment of the TN2 record in mid-August:

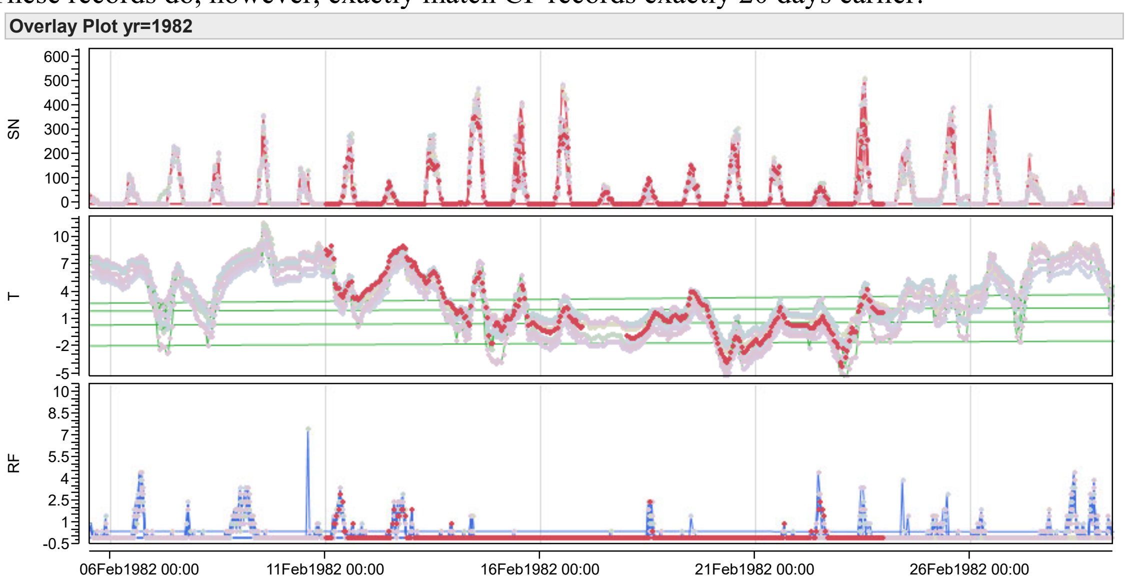


The match with the earlier record is numerically exact (and is also very strongly correlated with TN). Because this earlier record has apparently been copied and time-shifted, the later, misplaced copy has been deleted.

March 1982: CF does not match the other AWS's from 00h on 03Mar1982 to 23h on 15Mar1982:



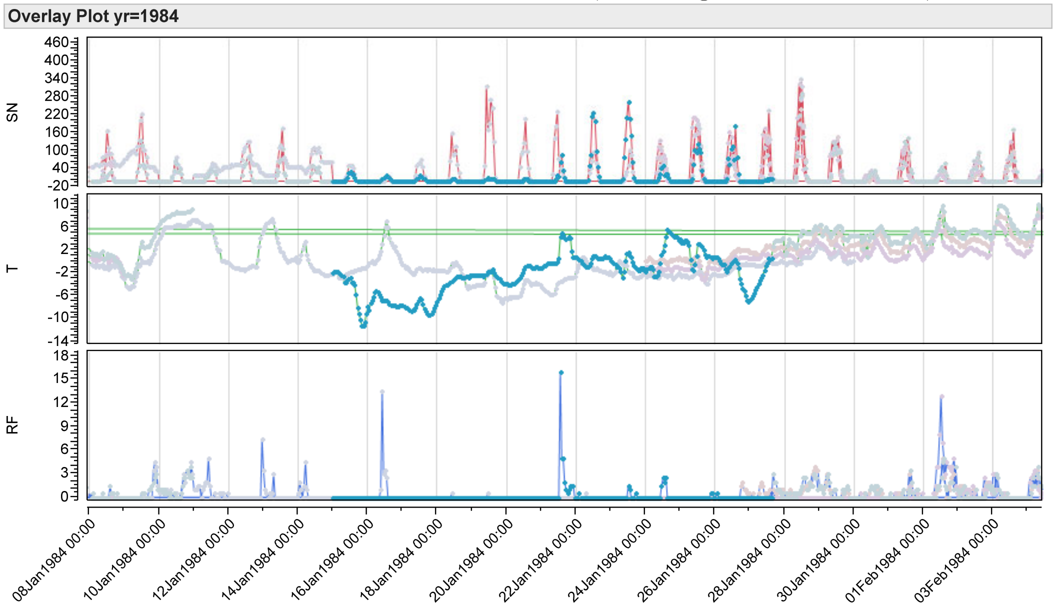
These records do, however, exactly match CF records exactly 20 days earlier:



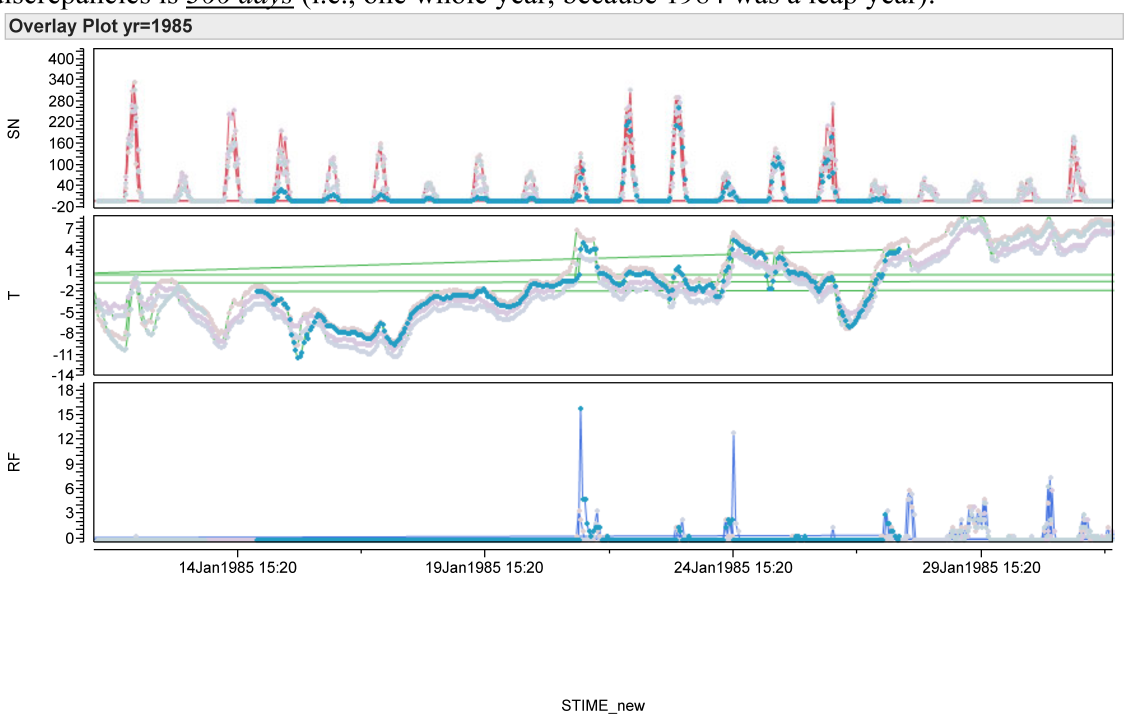
STIME_new

Because the misplaced records are exact copies of the original records, I have deleted the misplaced copies.

January 1984: Records from TN with timestamps from 00h on 15Jan1984 to 15h on 27Jan1984 are inconsistent with the other AWS's (including CW, EG, and CF):

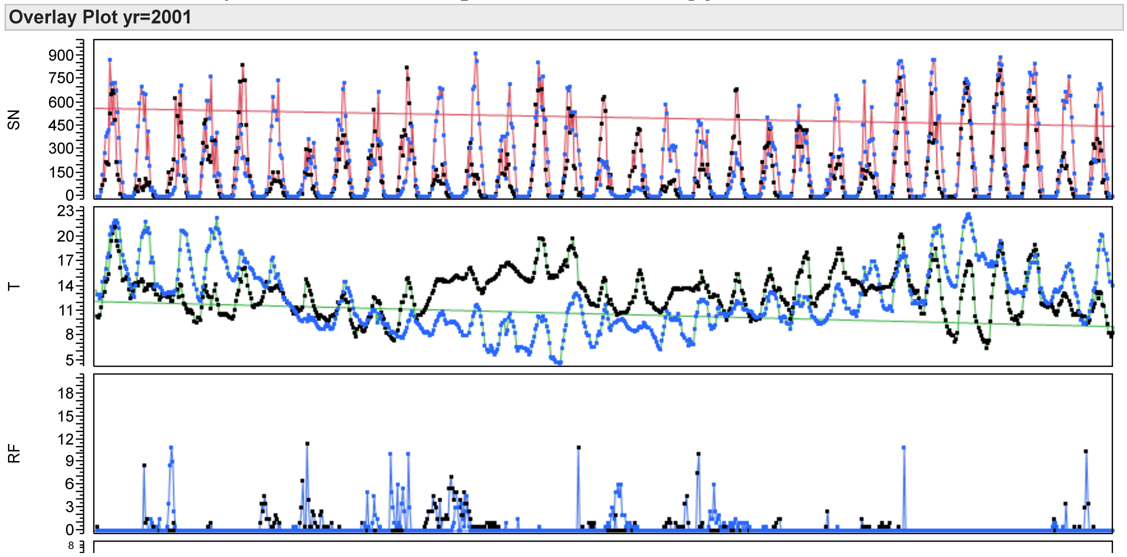


Note in particular the sharp precipitation event on 21Jan, which was not recorded by the other AWS's, and the mismatch with the solar flux data as well. The time shift that resolves these discrepancies is *366 days* (i.e., one whole year, because 1984 was a leap year):

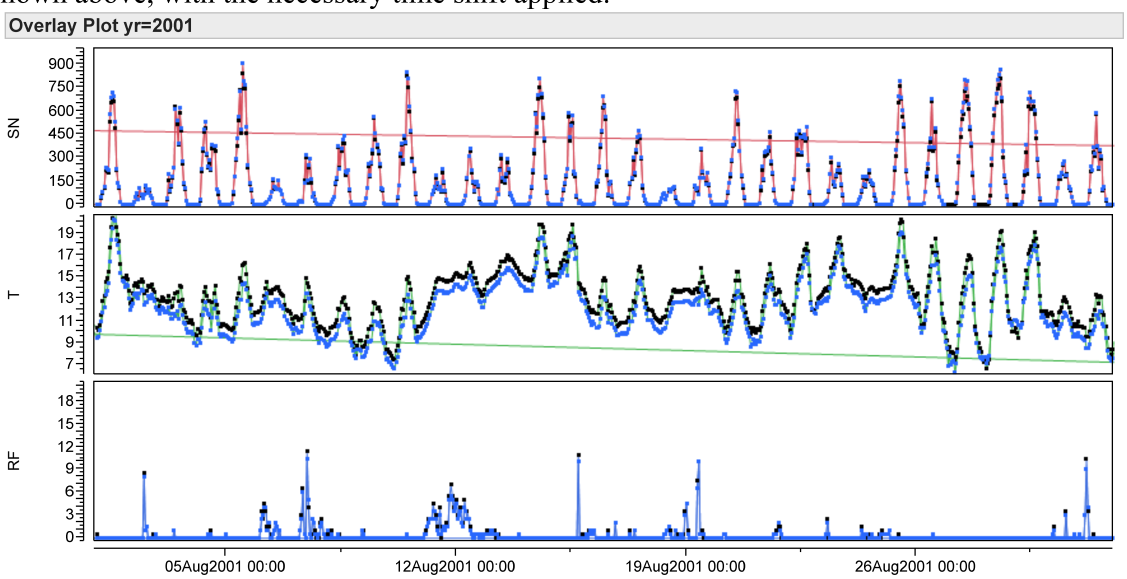


The timestamps have been corrected accordingly.

July-December 2001: Dates/times are incorrect for CW3 from the beginning of the record until the end of the year. Here is a sample month, showing just CW and CW3:



Notice that nothing matches! The CW3 timestamps need to be shifted earlier by 30 days plus one hour in order to coincide with CW (and the other AWS's). Here is the same time interval shown above, with the necessary time shift applied:



The timestamps have been corrected accordingly.

The results at this stage of the analysis are contained in the file *All_AWS_stacked v05.jmp*.

Level 1 editing of time series from individual sensors

The editing performed so far has been focused on correction or deletion of entire records that contain obvious junk data, or that have been linked to the wrong site and/or timestamp. Next, I undertook "level 1" editing of the time series of individual sensors at individual sites. By "level 1" editing, I mean the identification and removal of clearly incorrect data, which does not require a significant degree of subjective expert judgment.

- spikes in T

- flat zeroes in T

- flat zeroes in SN

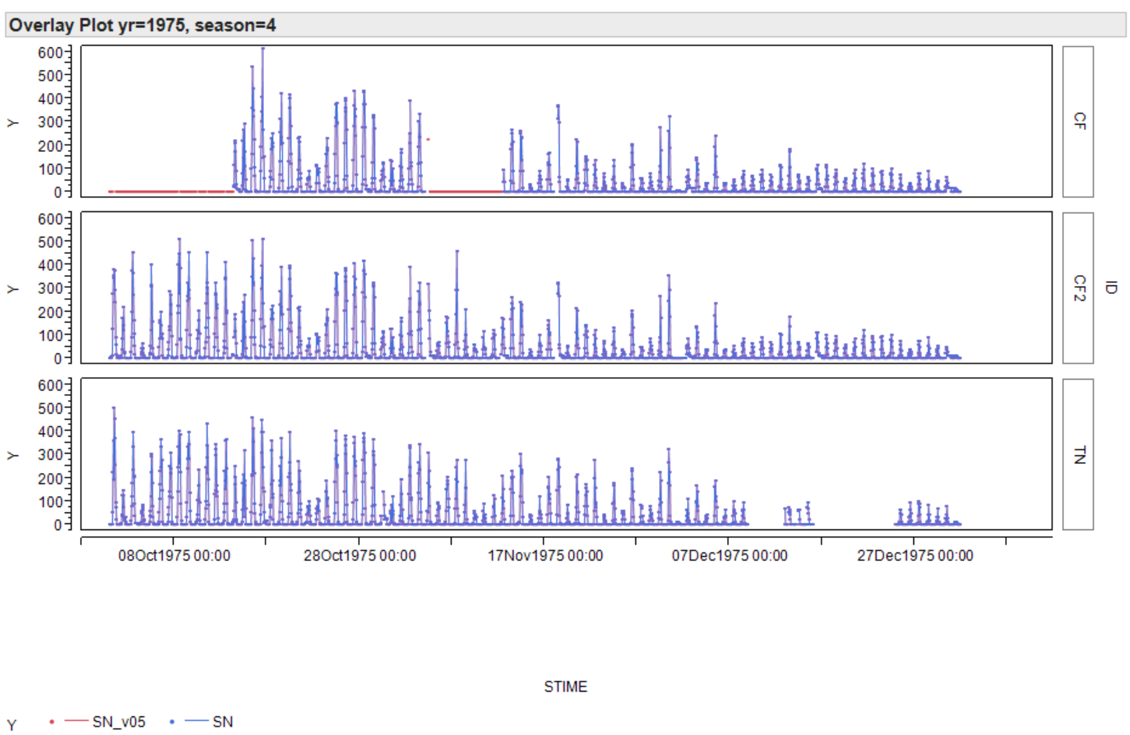
- night-time elevated levels in SN

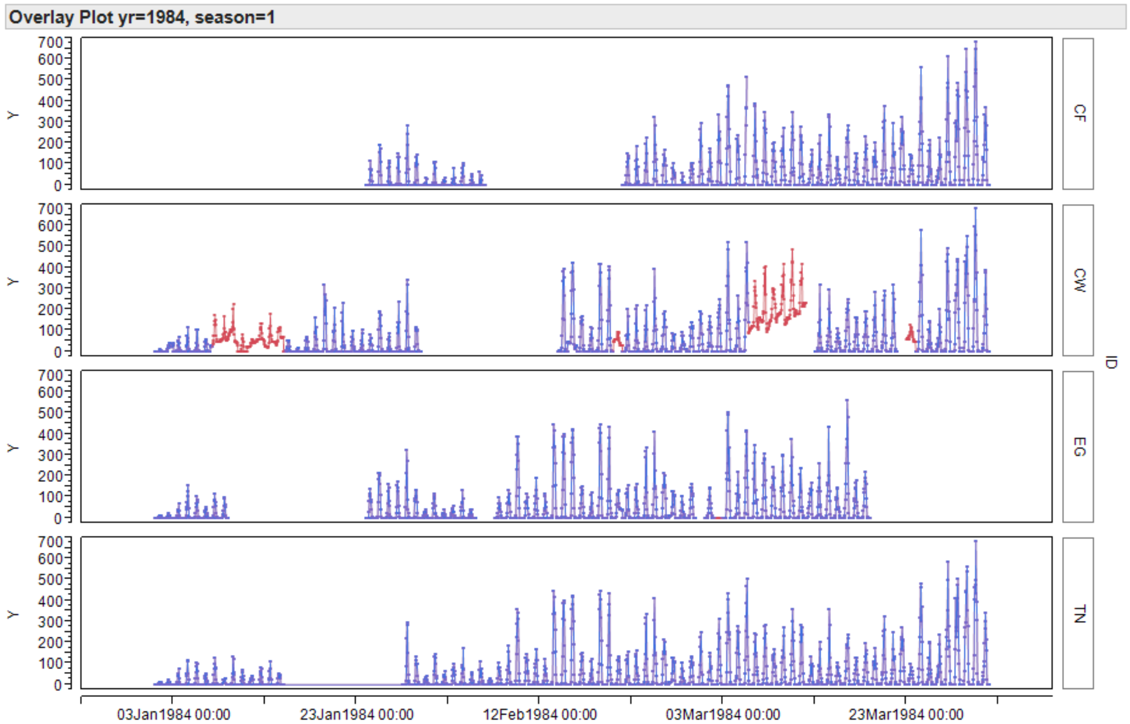
- whacko period in SN at EG2 November 1981

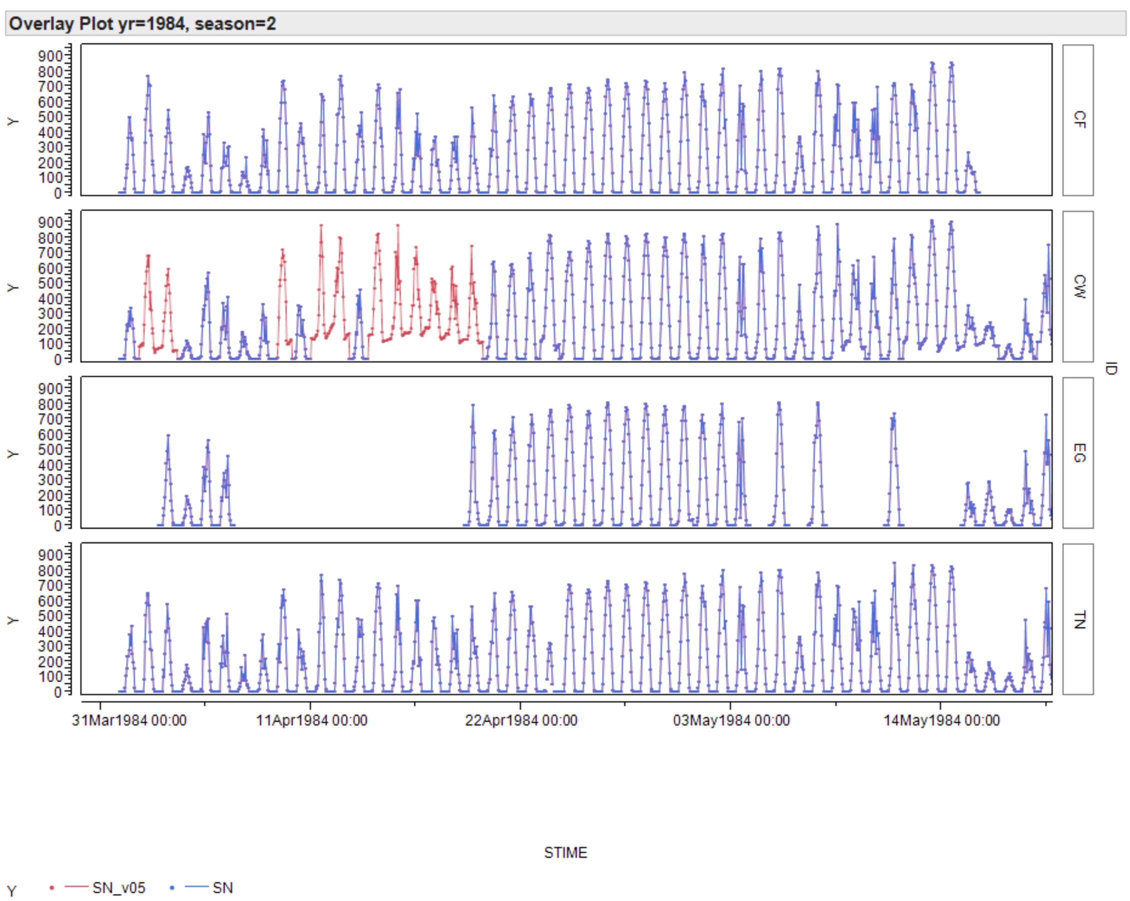
- flat zeroes in SN at EG2 Jan 1982

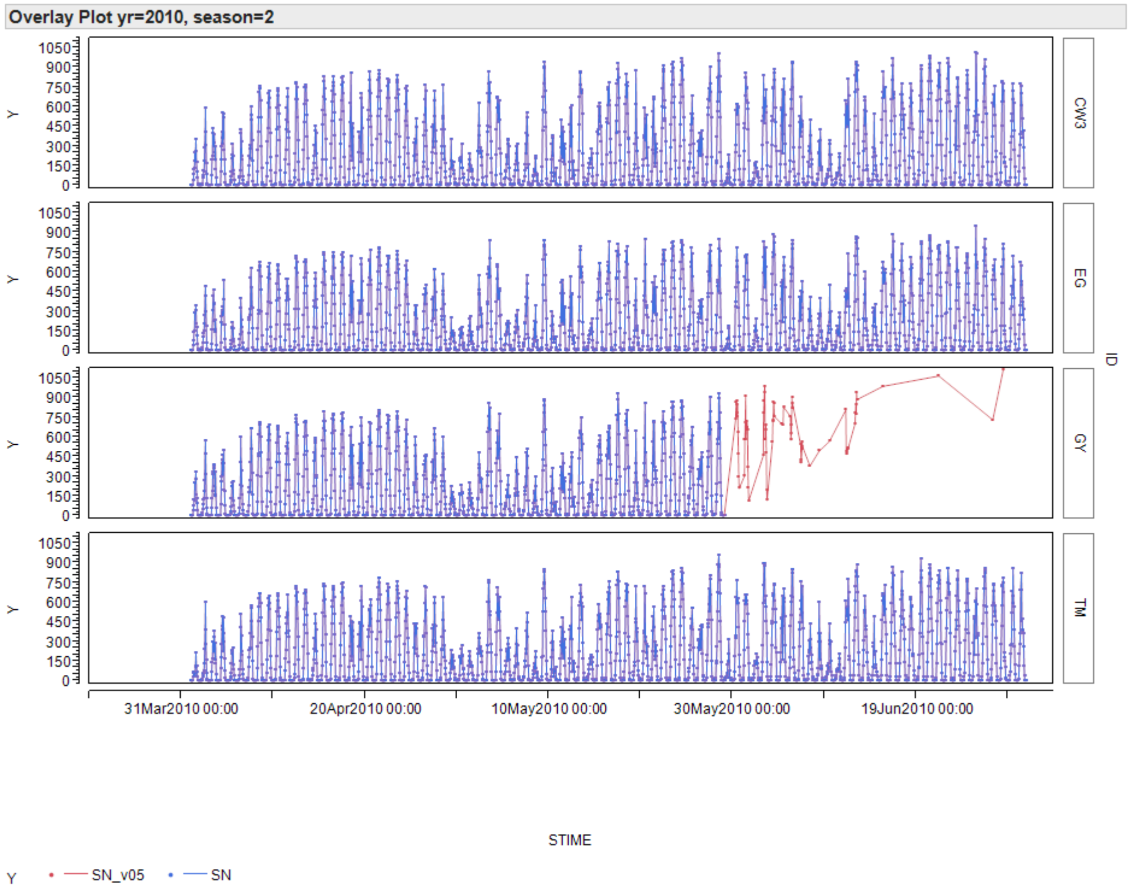
SN CW March 1984

TD: culled (many) spikes extending below 0 degrees (physically impossible) or more than about 0.5C below "baseline" of rest of TD (to allow for calibration drift, which is not corrected at Level 1). Note that these spikes do not occur before 1988, suggesting they are an artifact of an instrumentation change.

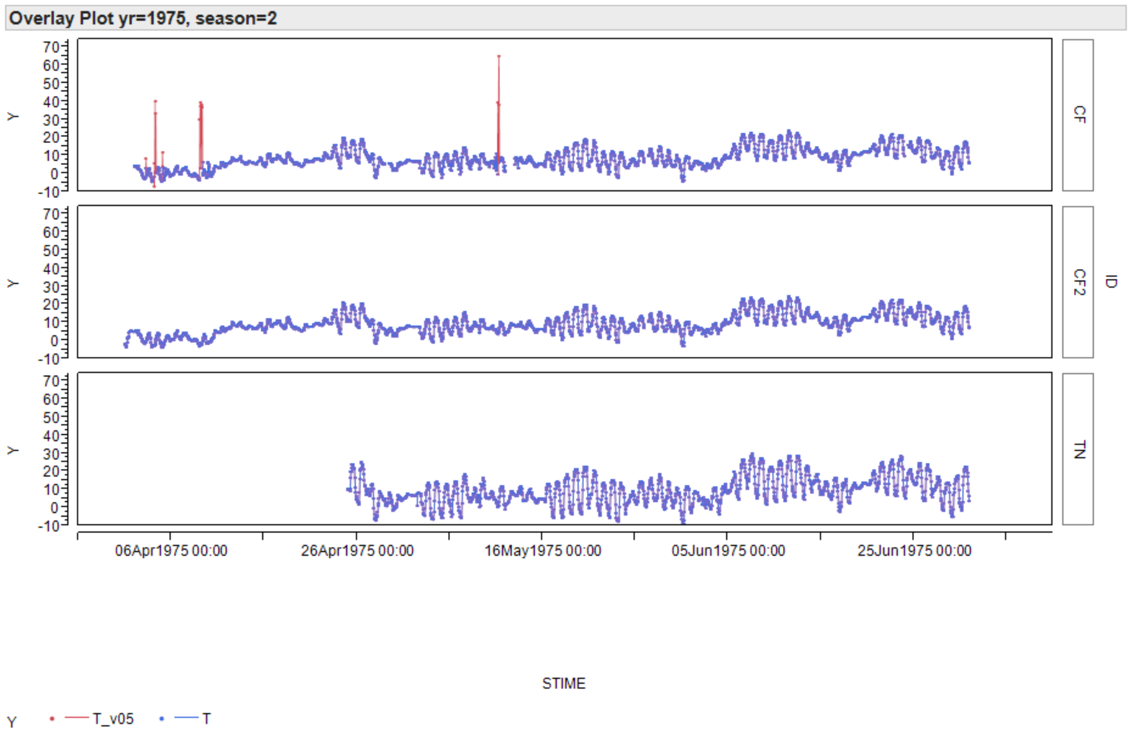


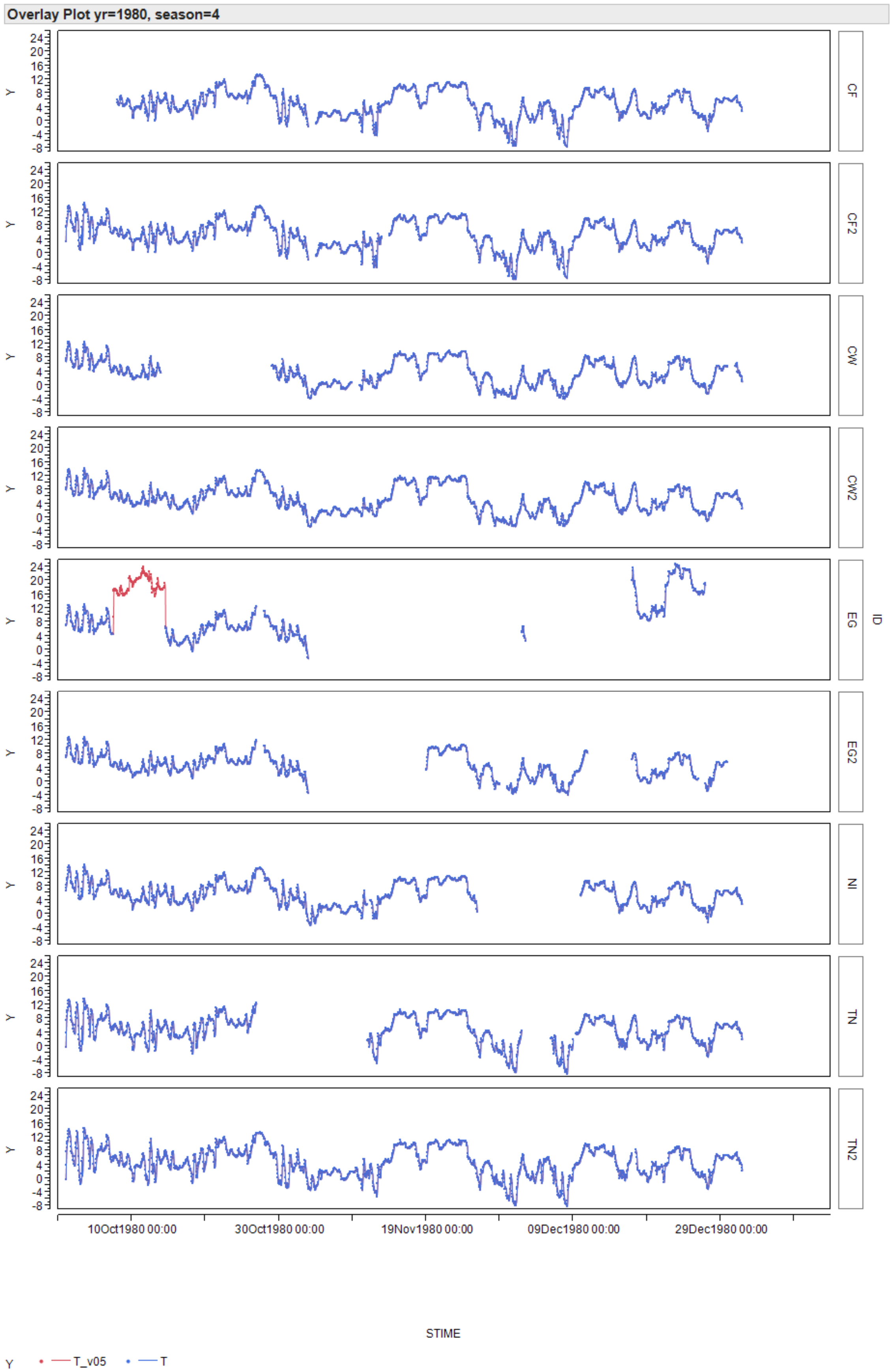

Overlay Plot yr=1984, season=1
Y
CF
CW
EG
TN
ID
03Jan1984 00:00
23Jan1984 00:00
12Feb1984 00:00
03Mar1984 00:00
23Mar1984 00:00


Overlay Plot yr=1984, season=2
Y
CF
CW
EG
TN
ID
31Mar1984 00:00
11Apr1984 00:00
22Apr1984 00:00
03May1984 00:00
14May1984 00:00
STIME
Y SN_v05 SN

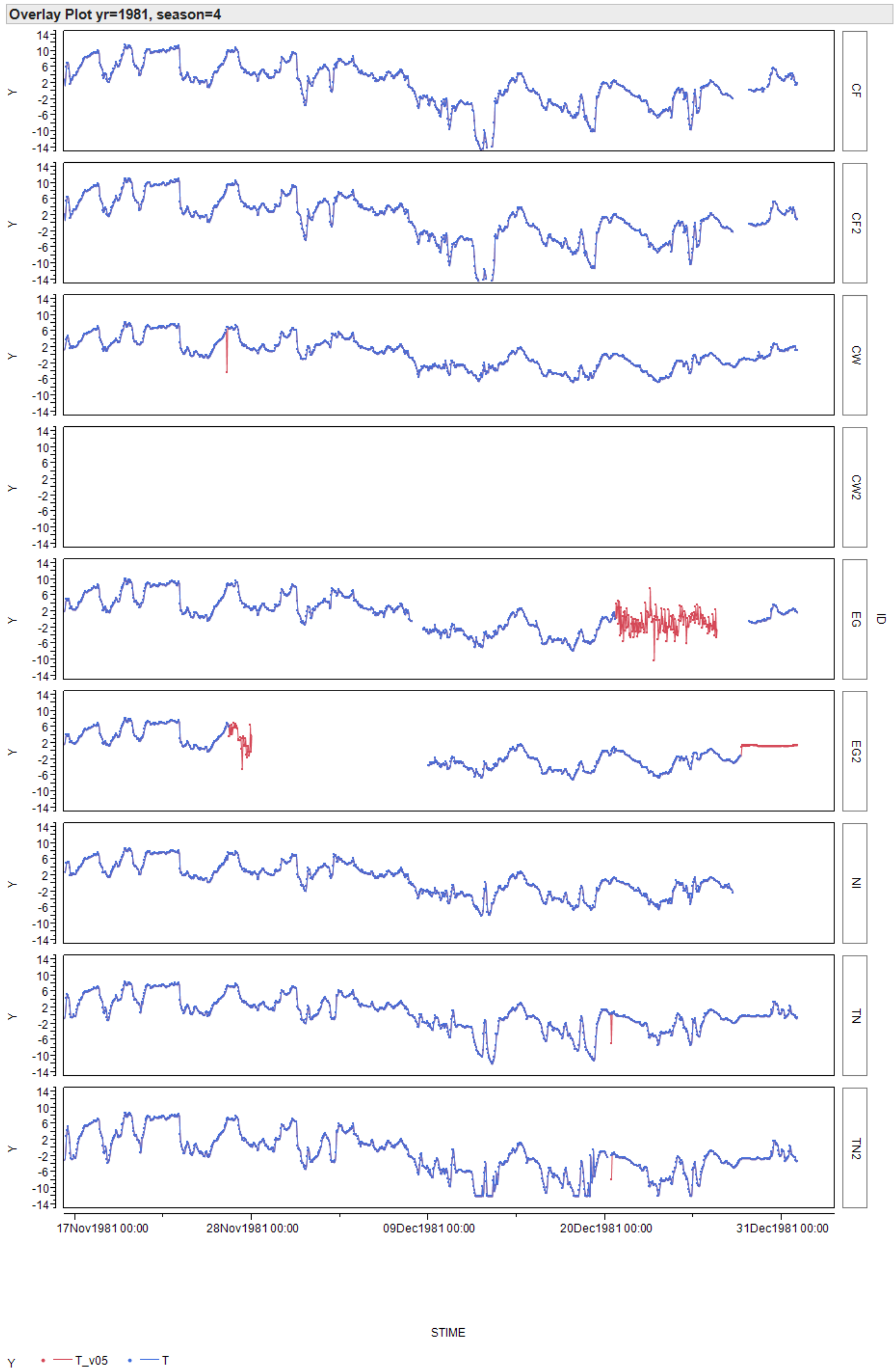


Overlay Plot yr=2010, season=2
Y
1050
900
750
600
450
300
150
0
CW3
EG
GY
TM
ID
31Mar2010 00:00
20Apr2010 00:00
10May2010 00:00
30May2010 00:00
19Jun2010 00:00
STIME
Y
SN_v05
SN

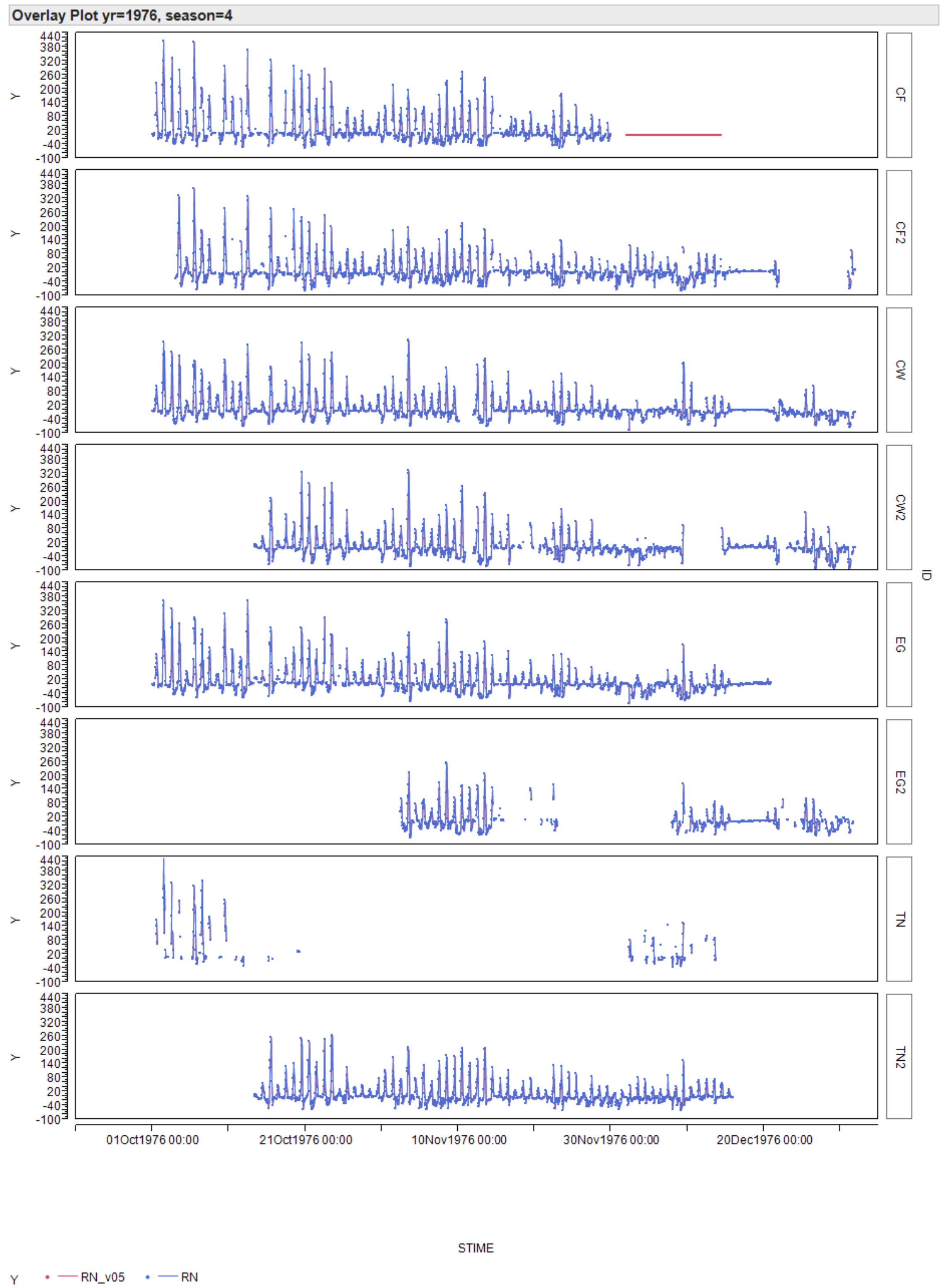

Overlay Plot yr=1975, season=2
Y
CF
CF2
TN
ID
06Apr1975 00:00
26Apr1975 00:00
16May1975 00:00
05Jun1975 00:00
25Jun1975 00:00
STIME
Y T_v05 T



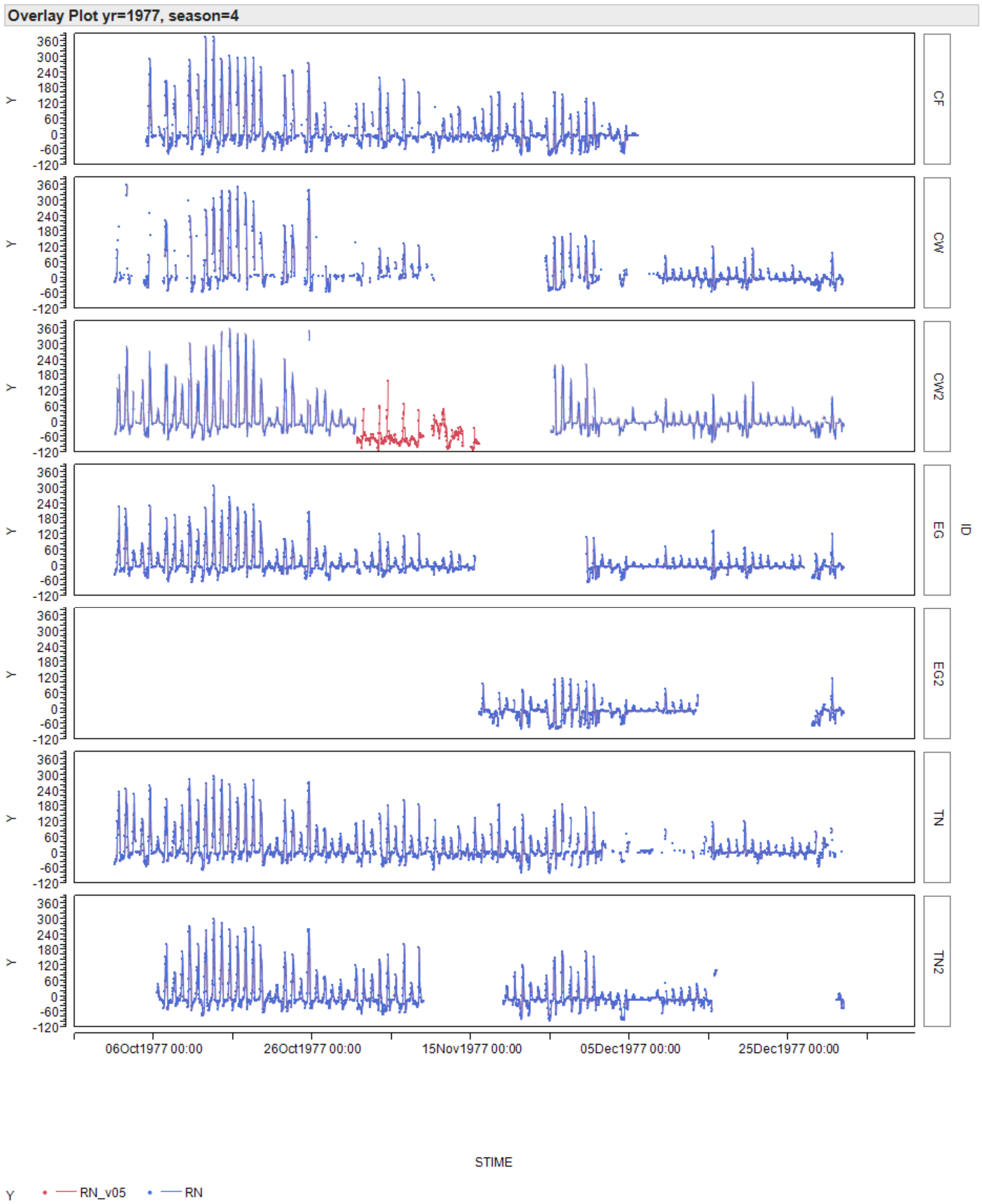
Overlay Plot yr=1980, season=4
Y
CF
CF2
CW
CW2
EG
EG2
NI
TN
TN2
ID
10Oct1980 00:00
30Oct1980 00:00
19Nov1980 00:00
09Dec1980 00:00
29Dec1980 00:00
STIME
Y
T_v05
T


Overlay Plot yr=1981, season=4
Y
14
10
6
2
-2
-6
-10
-14
CF
CF2
CW
CW2
EG
EG2
NI
TN
TN2
ID
17Nov1981 00:00
28Nov1981 00:00
09Dec1981 00:00
20Dec1981 00:00
31Dec1981 00:00
STIME
Y
T_v05
T

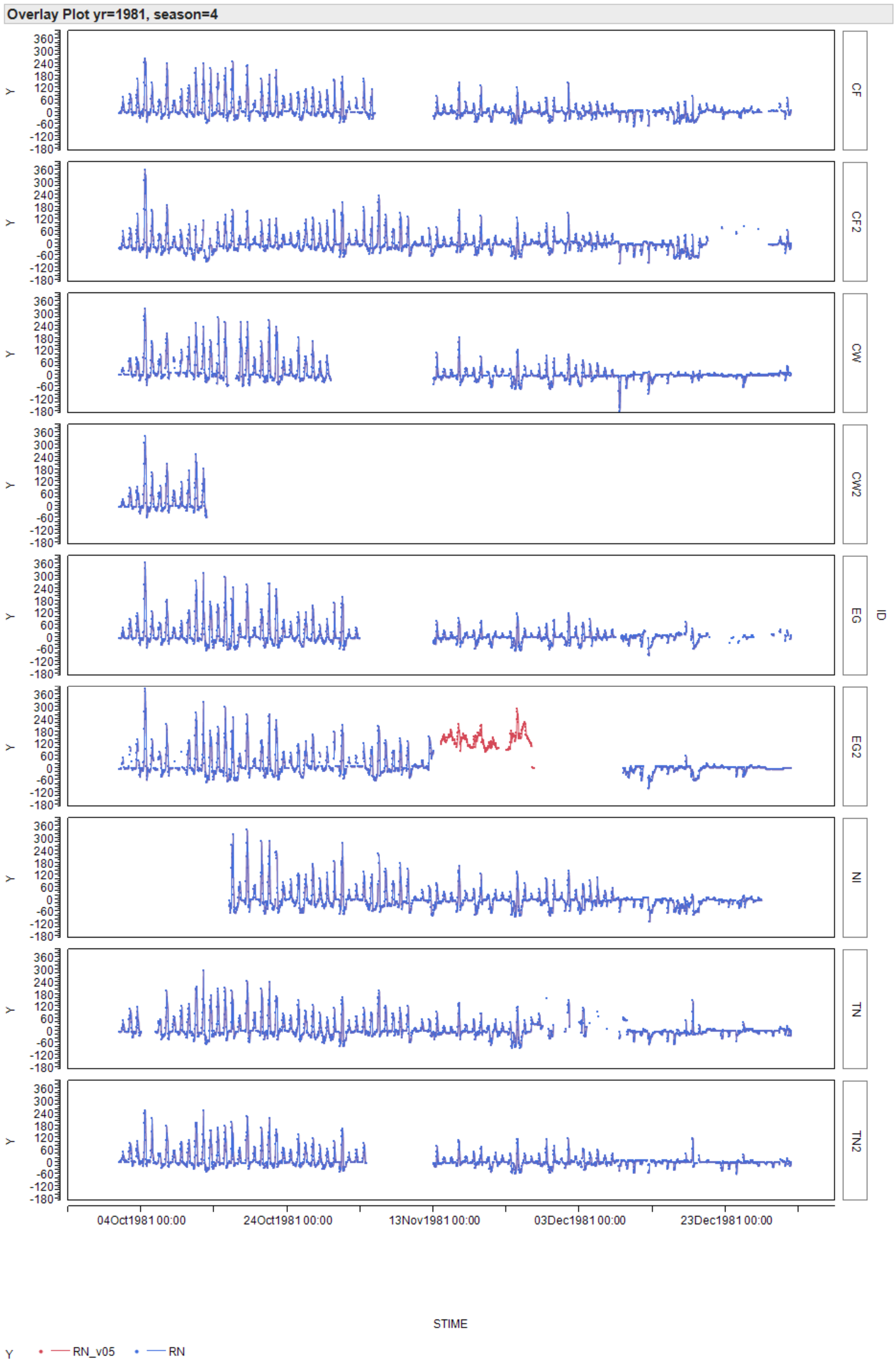
RN:



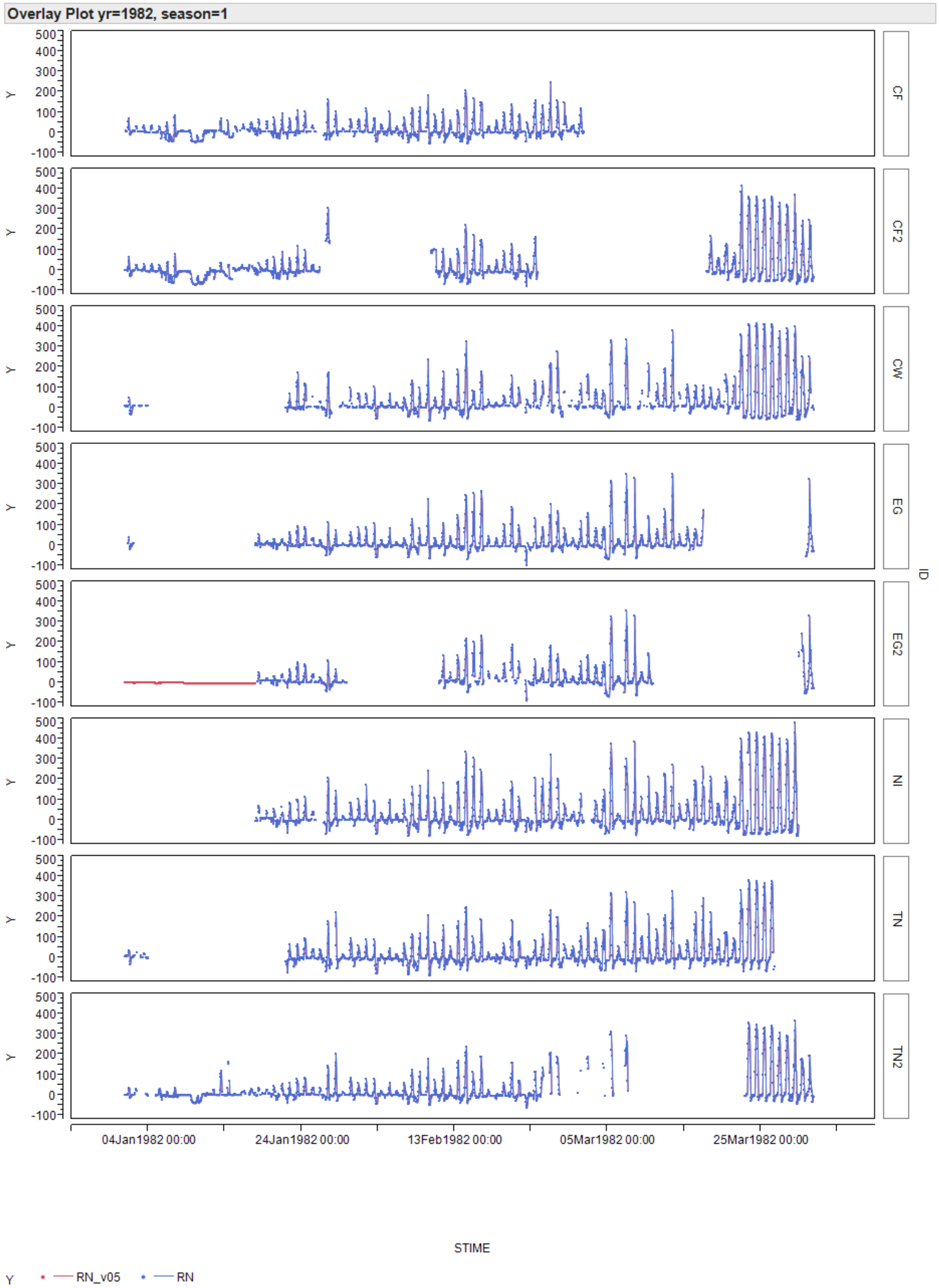
Overlay Plot yr=1976, season=4
Y
440
380
320
260
200
140
80
20
-40
-100
CF
CF2
CW
CW2
EG
EG2
TN
TN2
ID
01Oct1976 00:00
21Oct1976 00:00
10Nov1976 00:00
30Nov1976 00:00
20Dec1976 00:00
STIME
RN_v05
RN

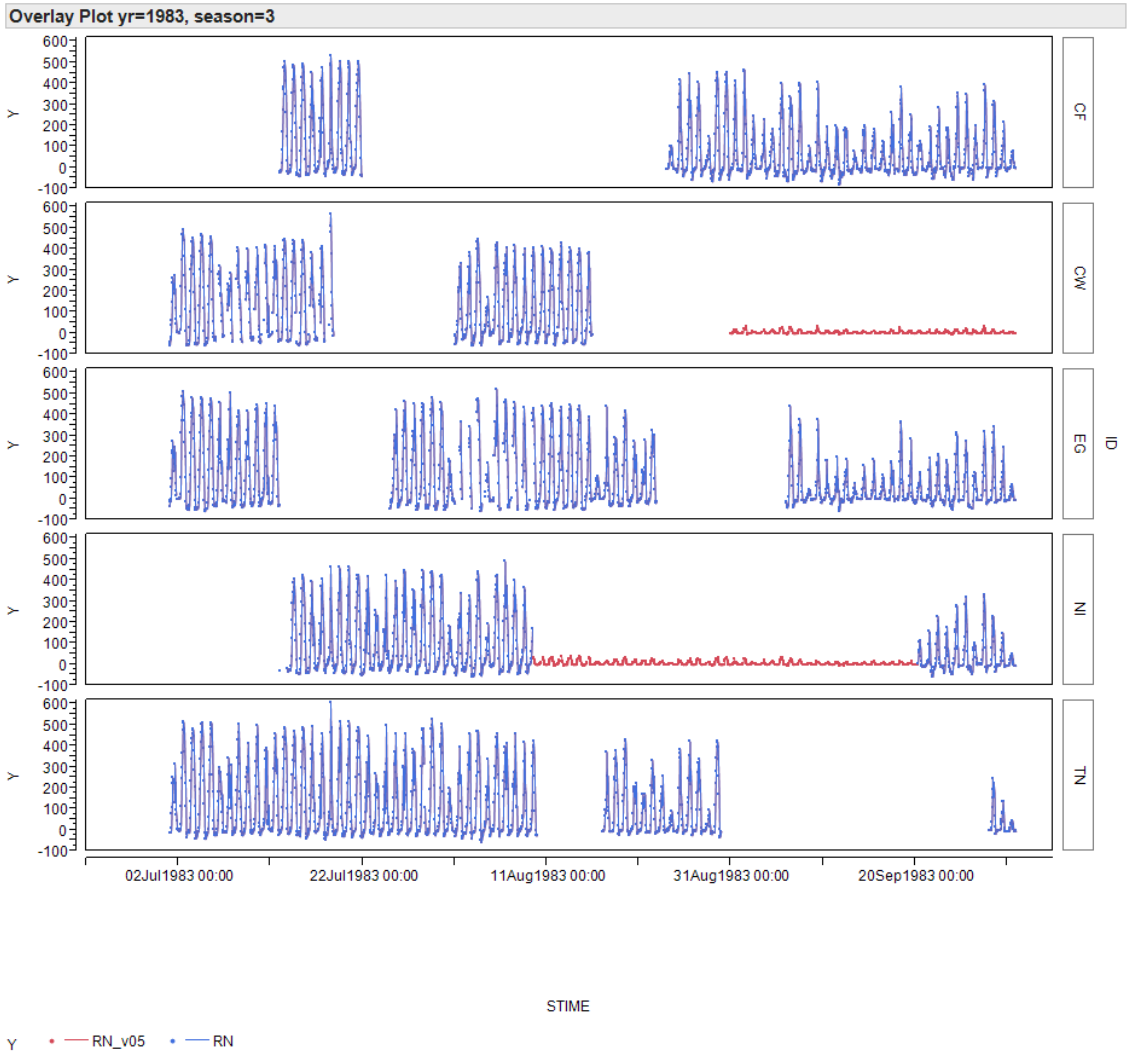


Overlay Plot yr=1977, season=4
Y
360
300
240
180
120
60
0
-60
-120
CF
CW
CW2
EG
EG2
TN
TN2
ID
06Oct1977 00:00
26Oct1977 00:00
15Nov1977 00:00
05Dec1977 00:00
25Dec1977 00:00
STIME
RN_v05
RN

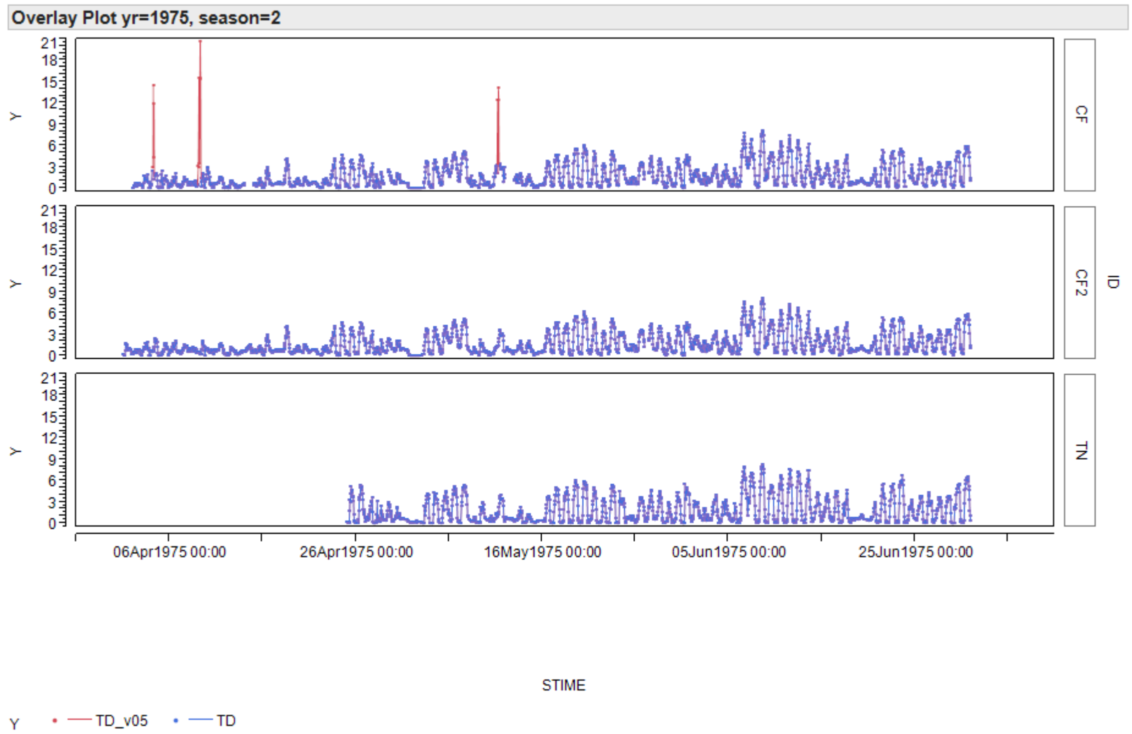


Overlay Plot yr=1981, season=4
CF
CF2
CW
CW2
EG
EG2
NI
TN
TN2
ID
Y
04Oct1981 00:00
24Oct1981 00:00
13Nov1981 00:00
03Dec1981 00:00
23Dec1981 00:00
STIME
RN_v05
RN

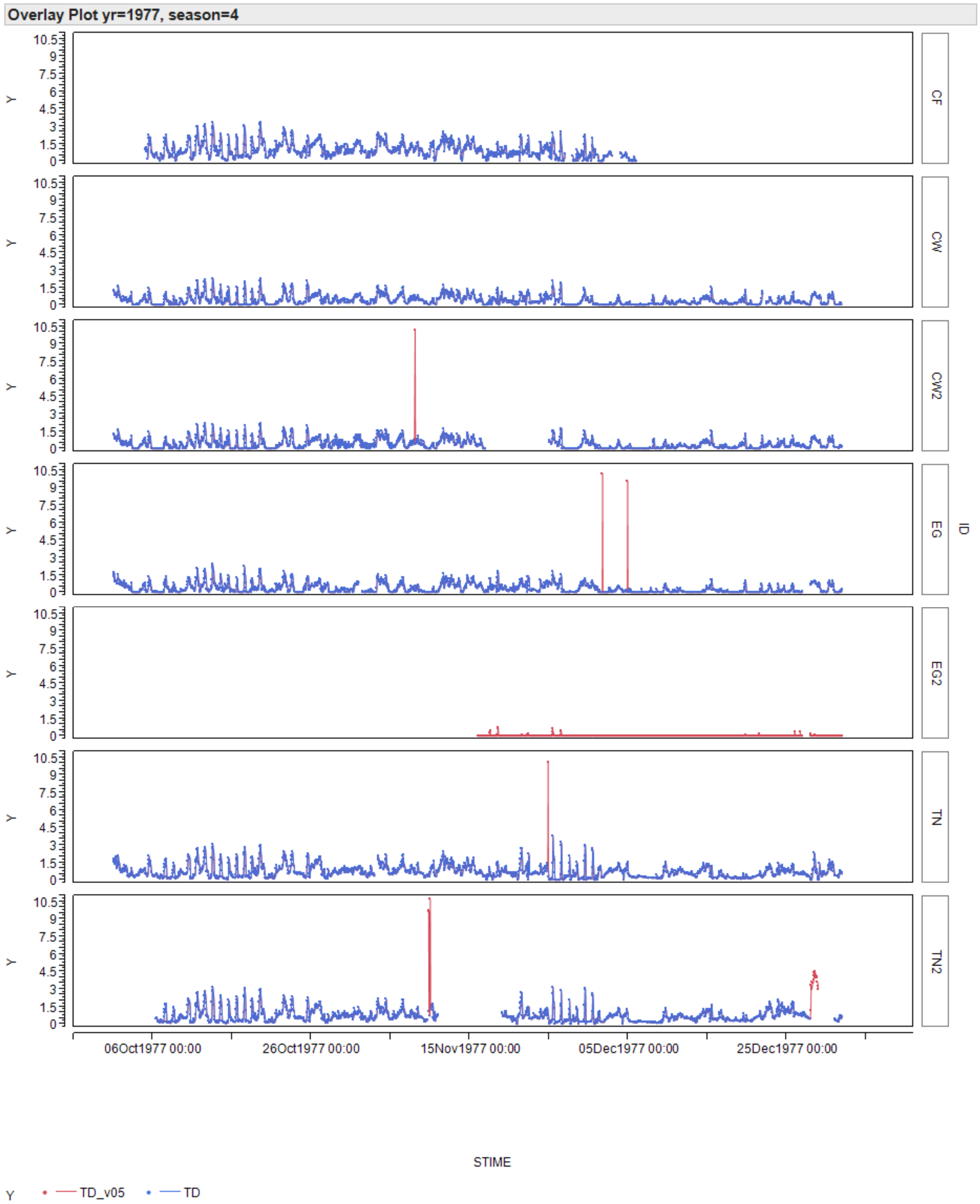

Overlay Plot yr=1982, season=1
Y
500
400
300
200
100
0
-100
CF
CF2
CW
EG
EG2
NI
TN
TN2
ID
04Jan1982 00:00
24Jan1982 00:00
13Feb1982 00:00
05Mar1982 00:00
25Mar1982 00:00
STIME
Y
RN_v05
RN


Overlay Plot yr=1983, season=3
Y
600
500
400
300
200
100
0
-100
CF
CW
EG
NI
TN
ID
02Jul1983 00:00
22Jul1983 00:00
11Aug1983 00:00
31Aug1983 00:00
20Sep1983 00:00
STIME
Y
RN_v05
RN

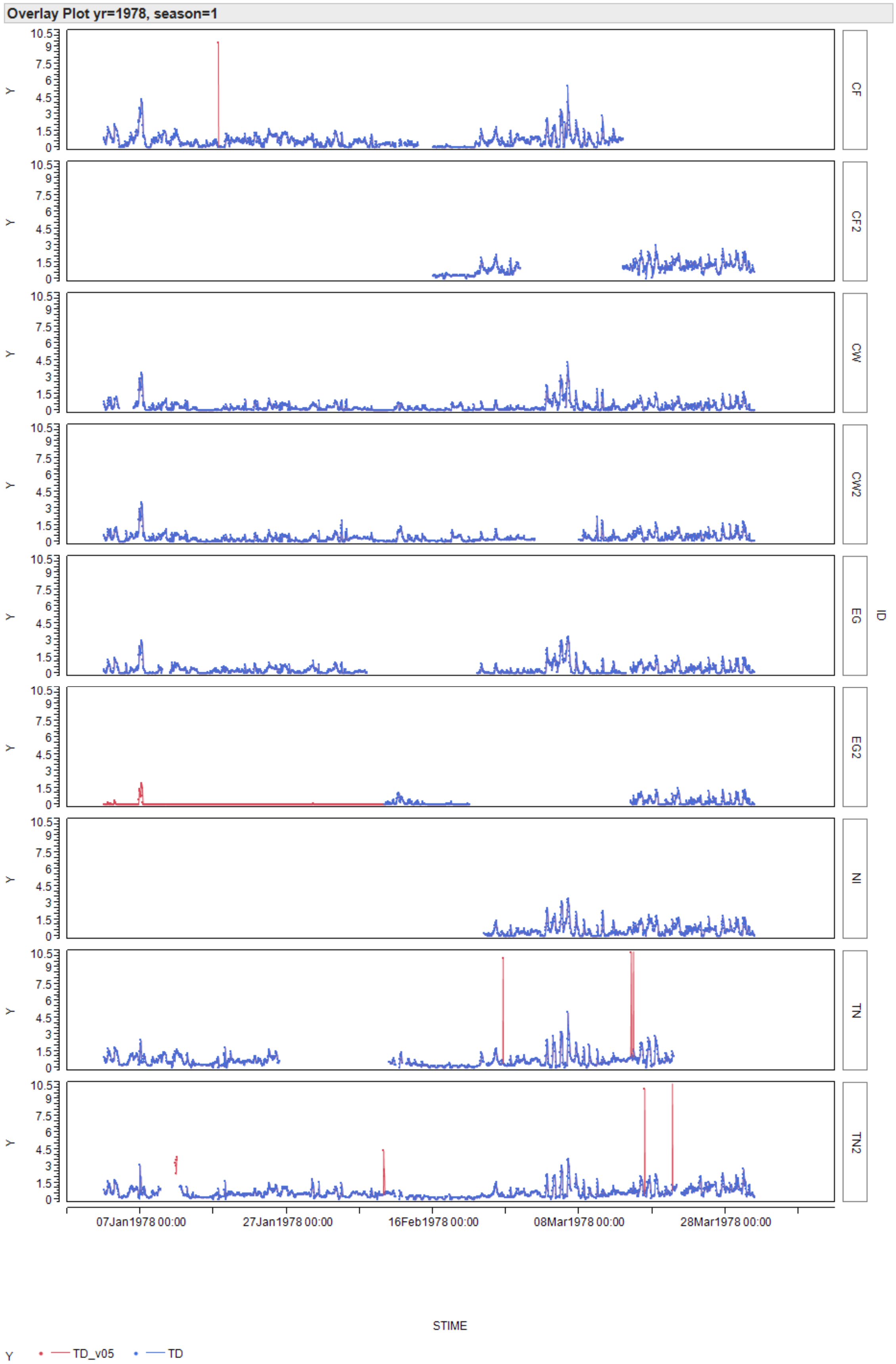
TD:

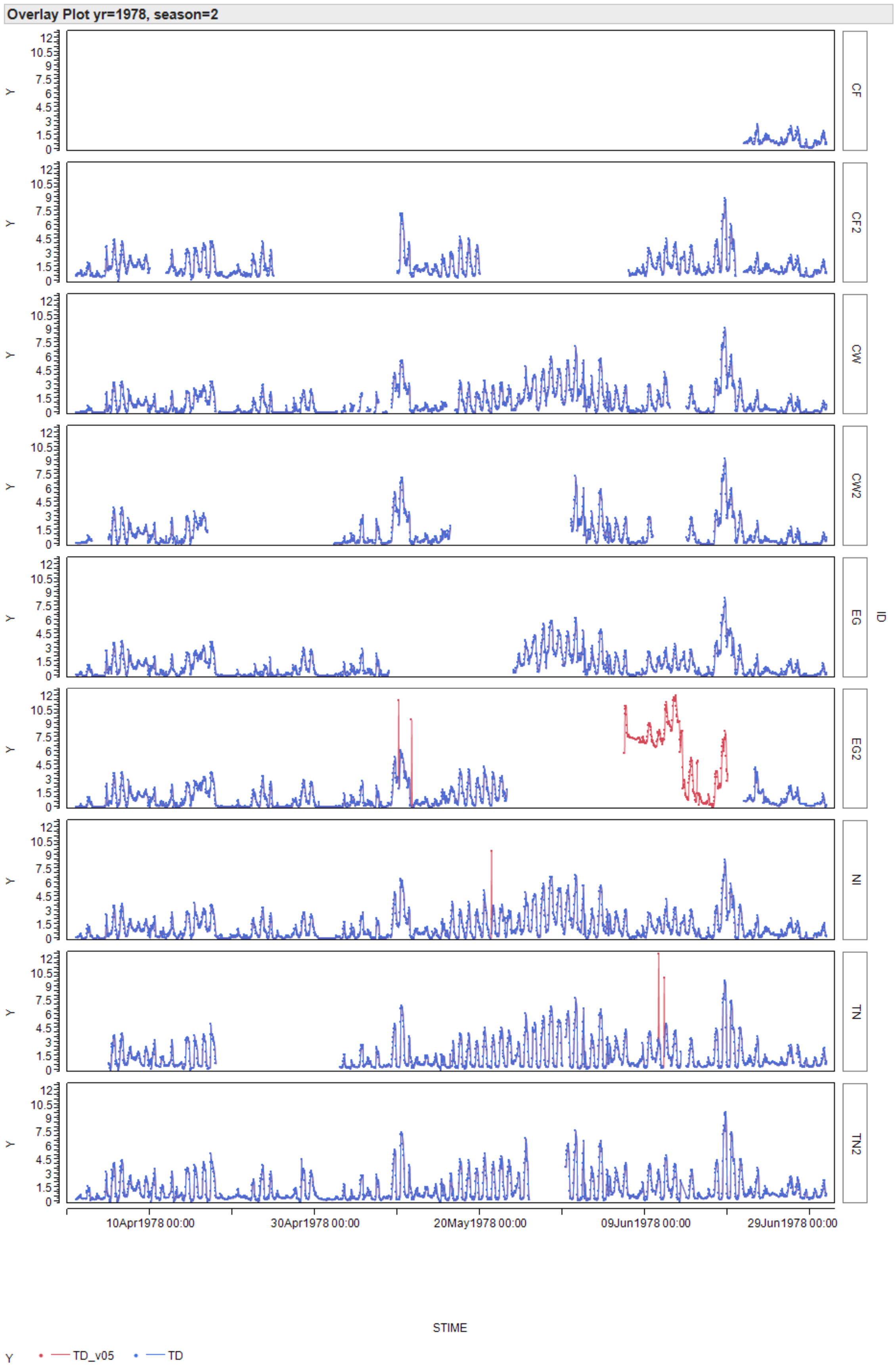


Overlay Plot yr=1975, season=2
Y
CF
CF2
TN
ID
06Apr1975 00:00
26Apr1975 00:00
16May1975 00:00
05Jun1975 00:00
25Jun1975 00:00
STIME
TD_v05
TD

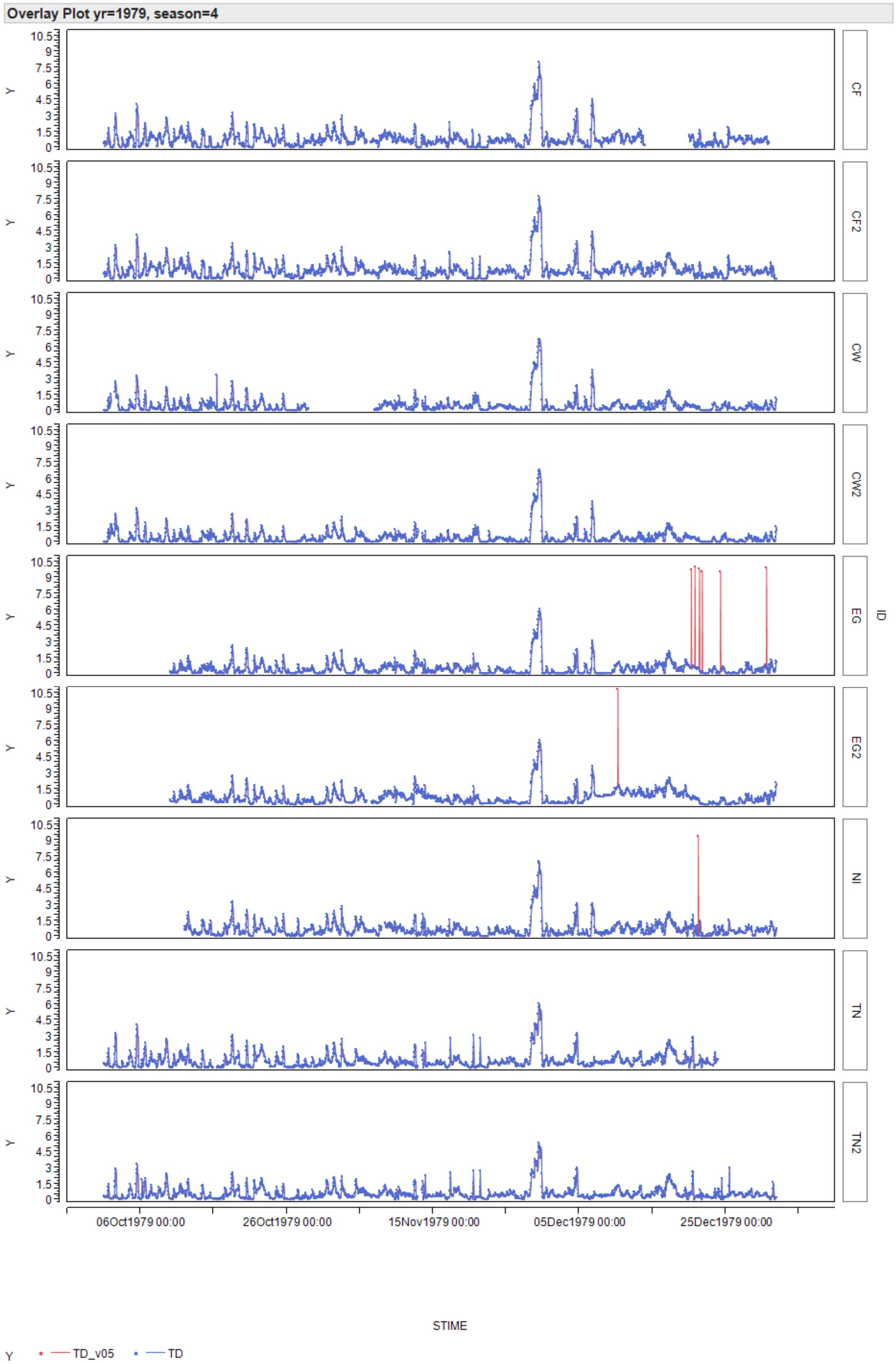


Overlay Plot yr=1977, season=4
Y
ID
CF
CW
CW2
EG
EG2
TN
TN2
06Oct1977 00:00
26Oct1977 00:00
15Nov1977 00:00
05Dec1977 00:00
25Dec1977 00:00
STIME
TD_v05
TD

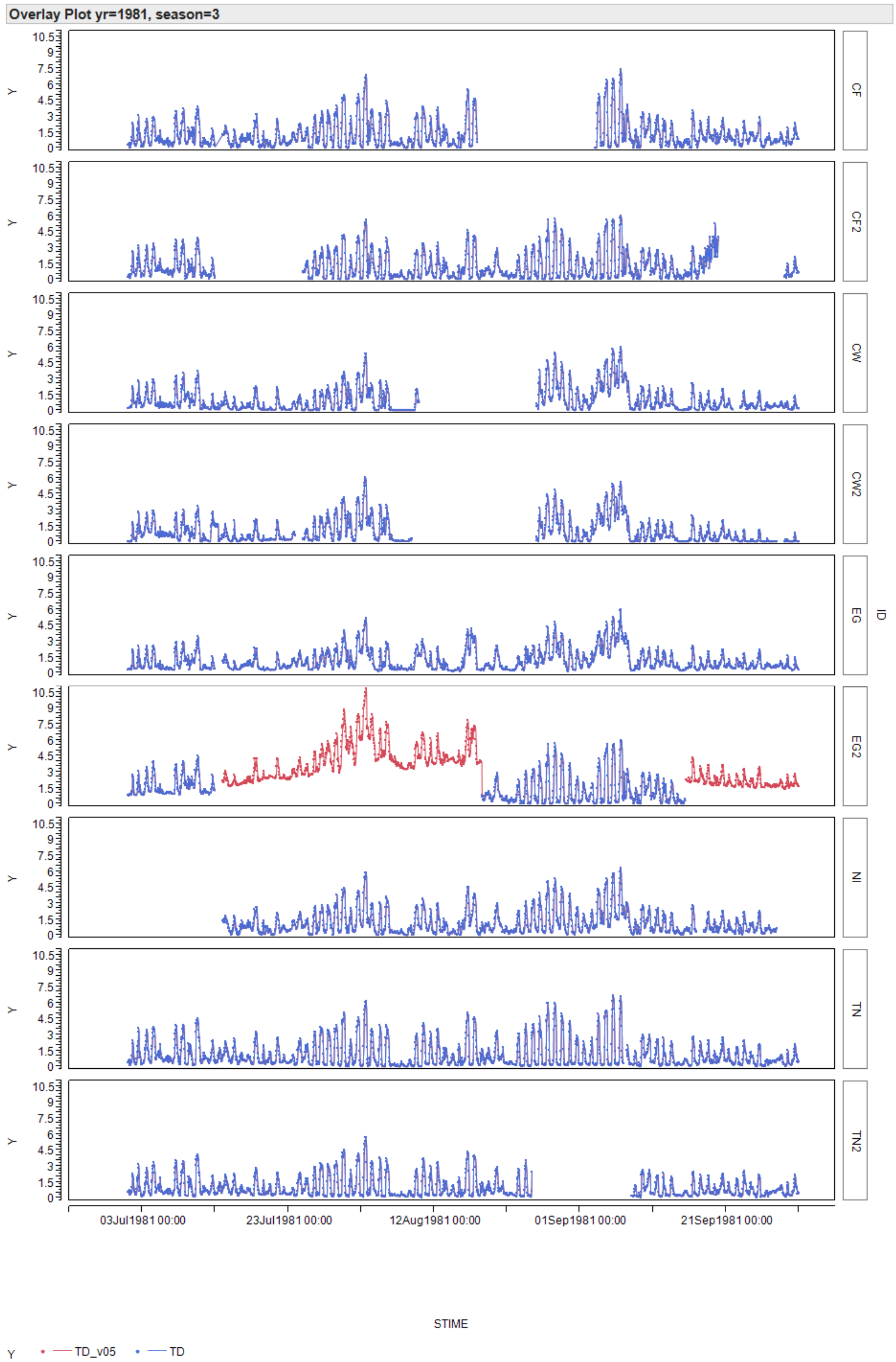

Overlay Plot yr=1978, season=1
Y
10.5
9
7.5
6
4.5
3
1.5
0
CF
CF2
CW
CW2
EG
EG2
NI
TN
TN2
ID
07Jan1978 00:00
27Jan1978 00:00
16Feb1978 00:00
08Mar1978 00:00
28Mar1978 00:00
STIME
Y
TD_v05
TD

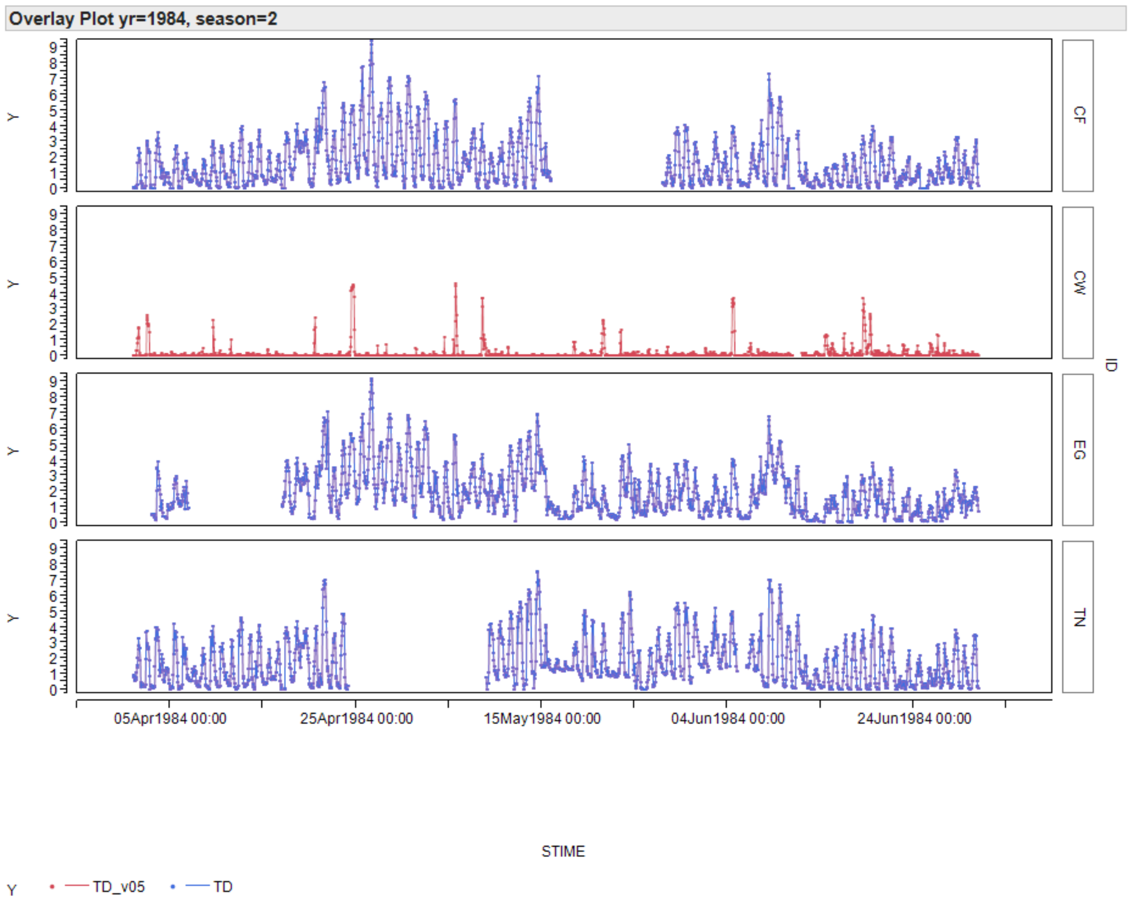


Overlay Plot yr=1978, season=2
Y
CF
CF2
CW
CW2
EG
EG2
NI
TN
TN2
ID
12
10.5
9
7.5
6
4.5
3
1.5
0
10Apr1978 00:00
30Apr1978 00:00
20May1978 00:00
09Jun1978 00:00
29Jun1978 00:00
STIME
TD_v05
TD

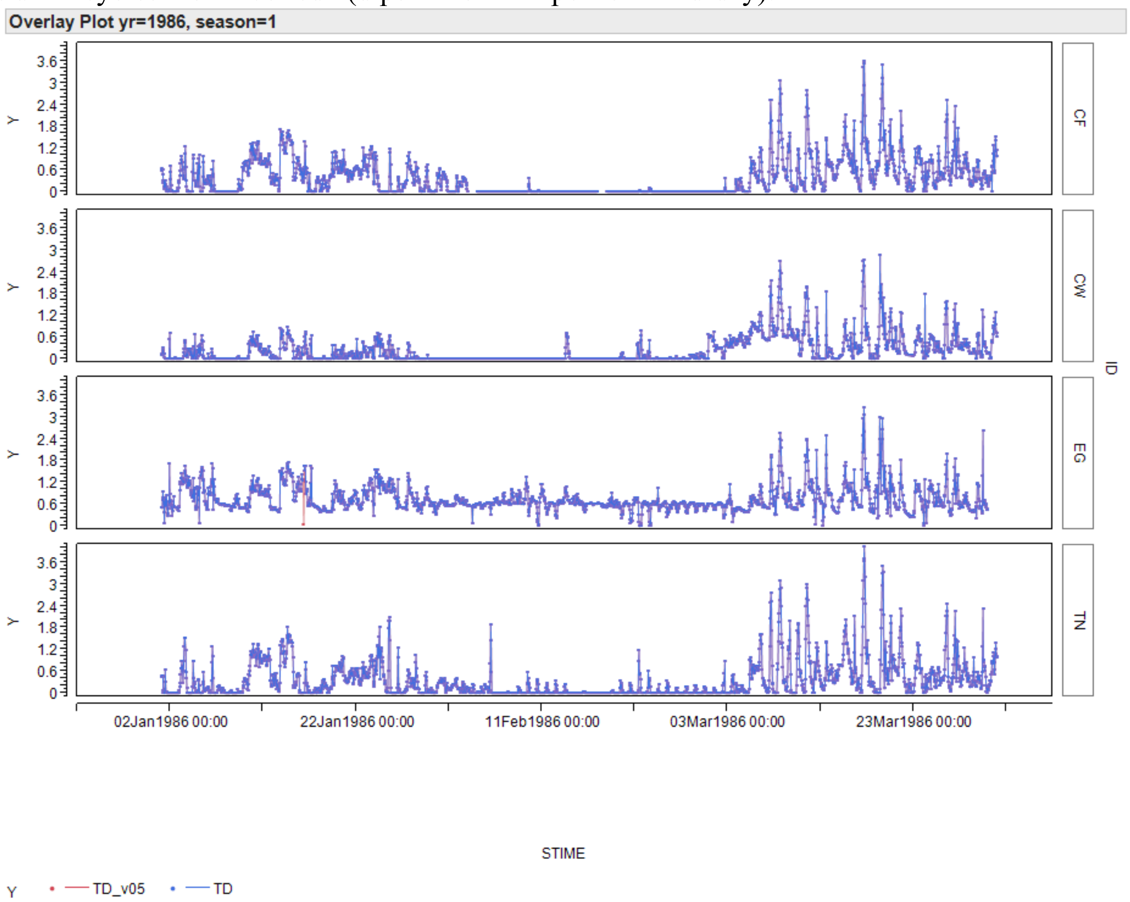


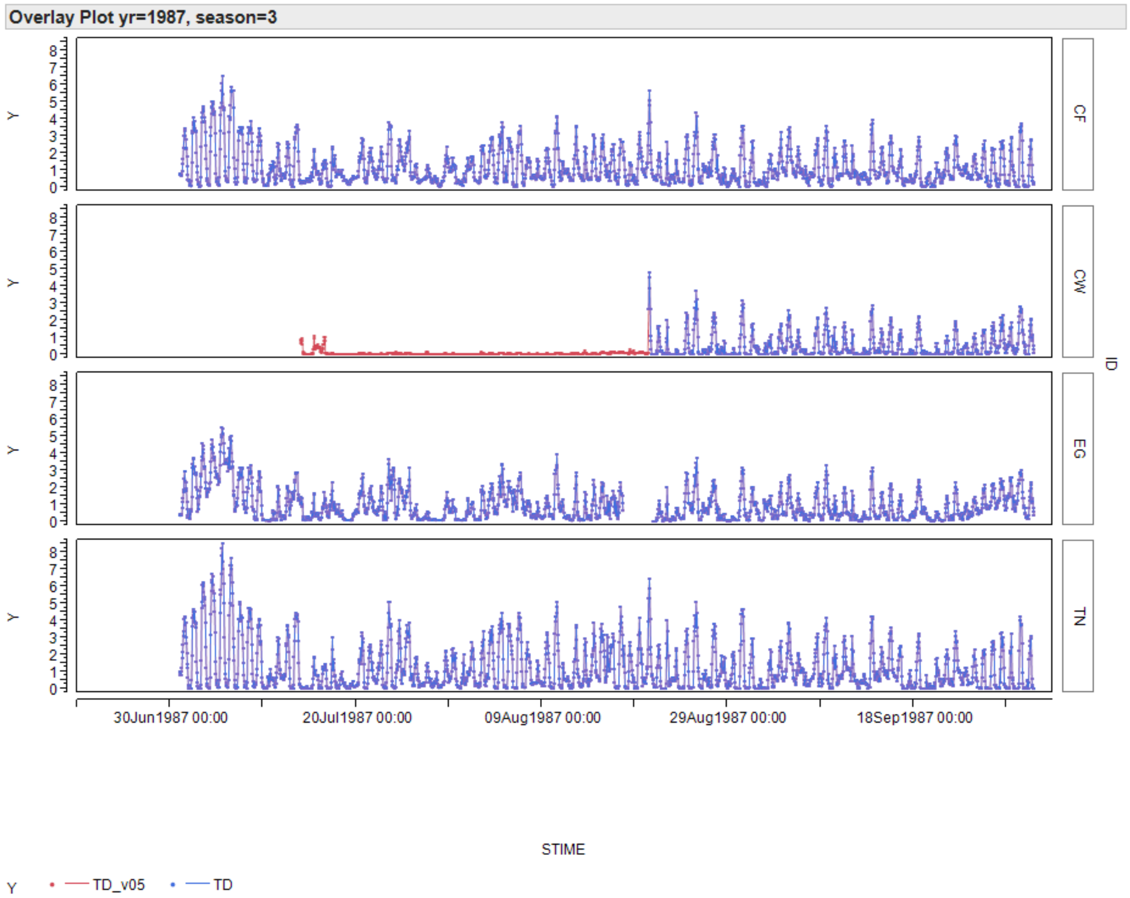
Overlay Plot yr=1979, season=4
Y
10.5
9
7.5
6
4.5
3
1.5
0
CF
CF2
CW
CW2
EG
EG2
NI
TN
TN2
ID
06Oct1979 00:00
26Oct1979 00:00
15Nov1979 00:00
05Dec1979 00:00
25Dec1979 00:00
STIME
Y
TD_v05
TD


Overlay Plot yr=1981, season=3
Y
10.5
9
7.5
6
4.5
3
1.5
0
CF
CF2
CW
CW2
EG
EG2
NI
TN
TN2
ID
03Jul1981 00:00
23Jul1981 00:00
12Aug1981 00:00
01Sep1981 00:00
21Sep1981 00:00
STIME
TD_v05
TD

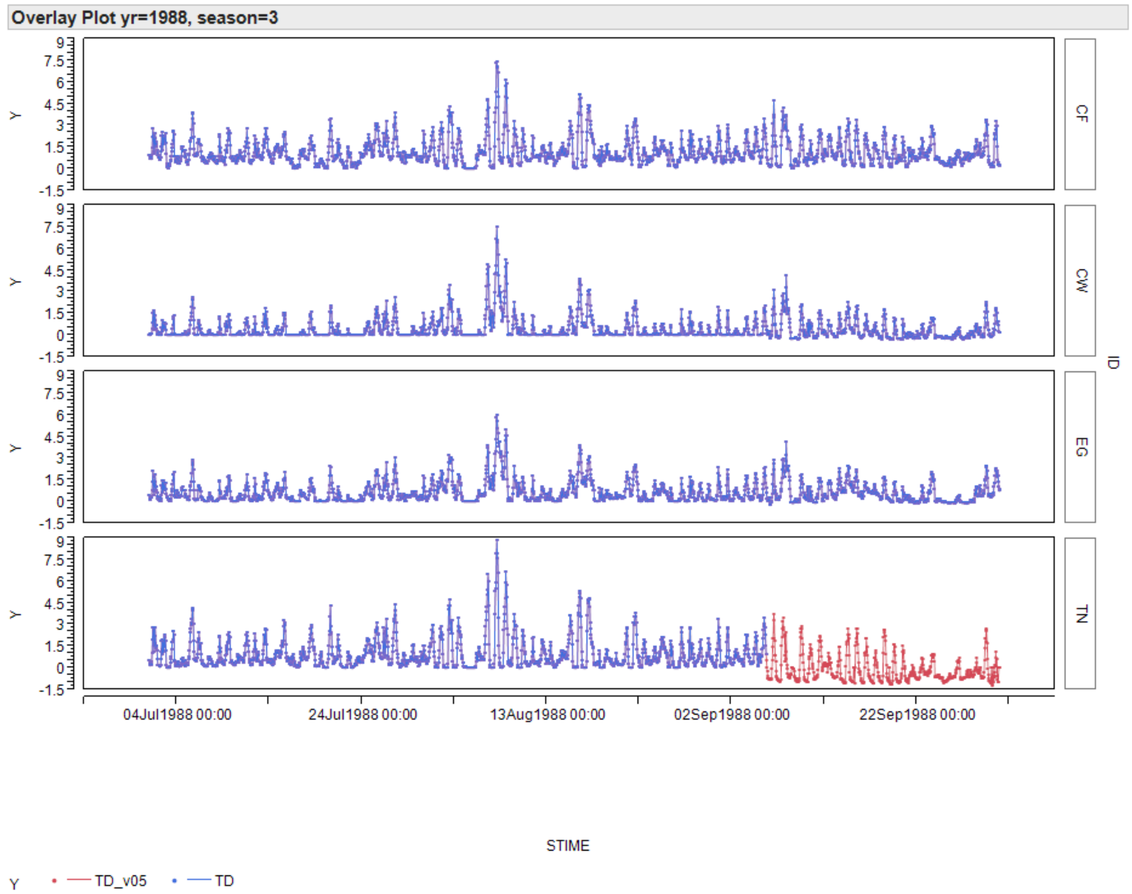

Overlay Plot yr=1984, season=2
Y
CF
CW
EG
TN
ID
05Apr1984 00:00
25Apr1984 00:00
15May1984 00:00
04Jun1984 00:00
24Jun1984 00:00
STIME
Y
TD_v05
TD

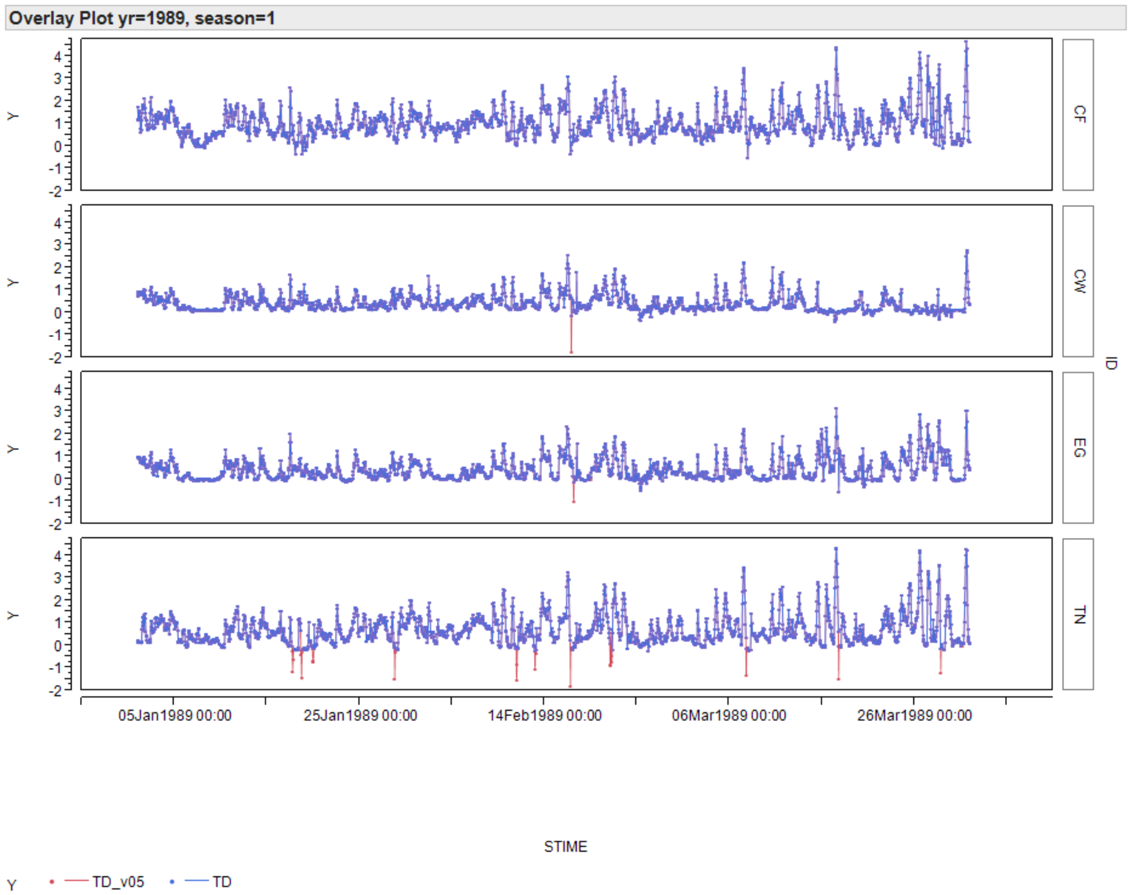
Note we've kept (for the moment) TD records like the one below for EG, even though the diurnal cycles are *inverted* (dips rather than peaks at mid-day):




Overlay Plot yr=1987, season=3
Y
CF
CW
EG
TN
ID
0
1
2
3
4
5
6
7
8
30Jun1987 00:00
20Jul1987 00:00
09Aug1987 00:00
29Aug1987 00:00
18Sep1987 00:00
STIME
Y
TD_v05
TD

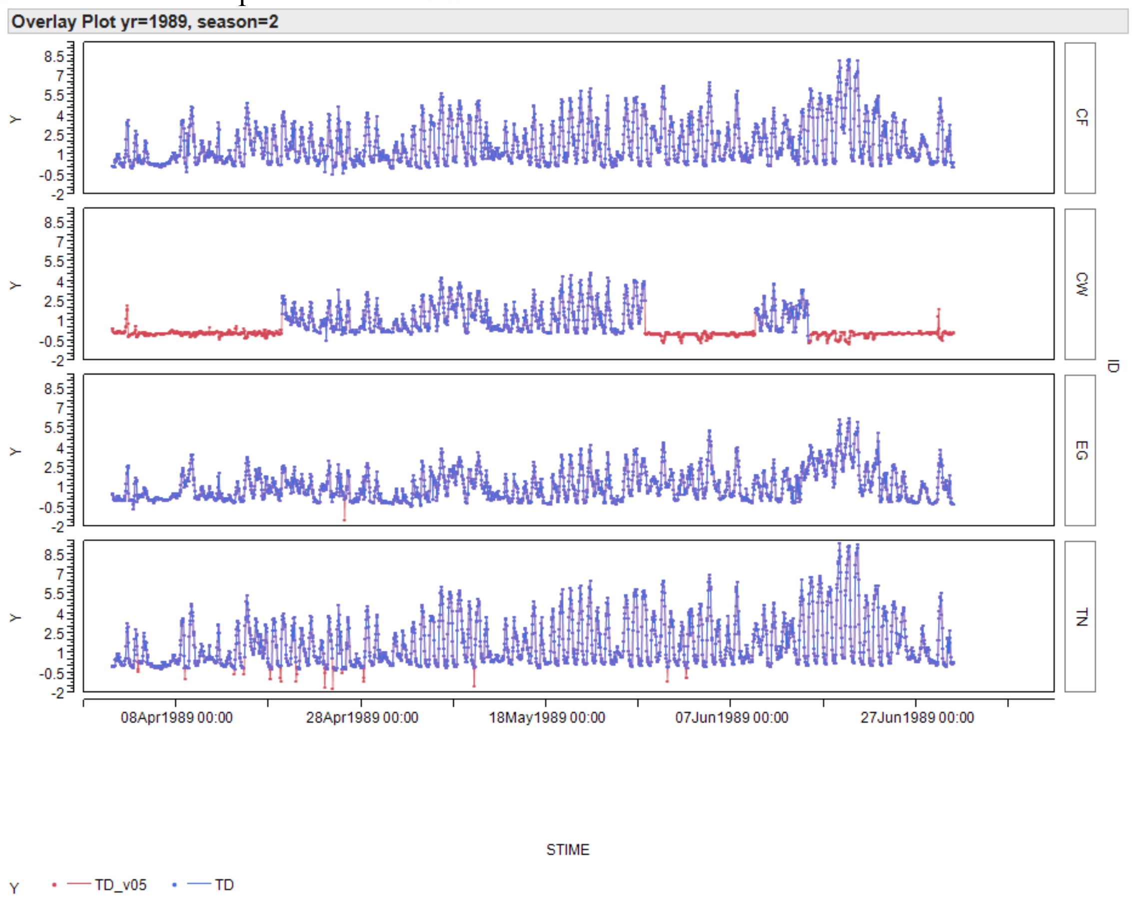
Obvious calibration shift in TN:

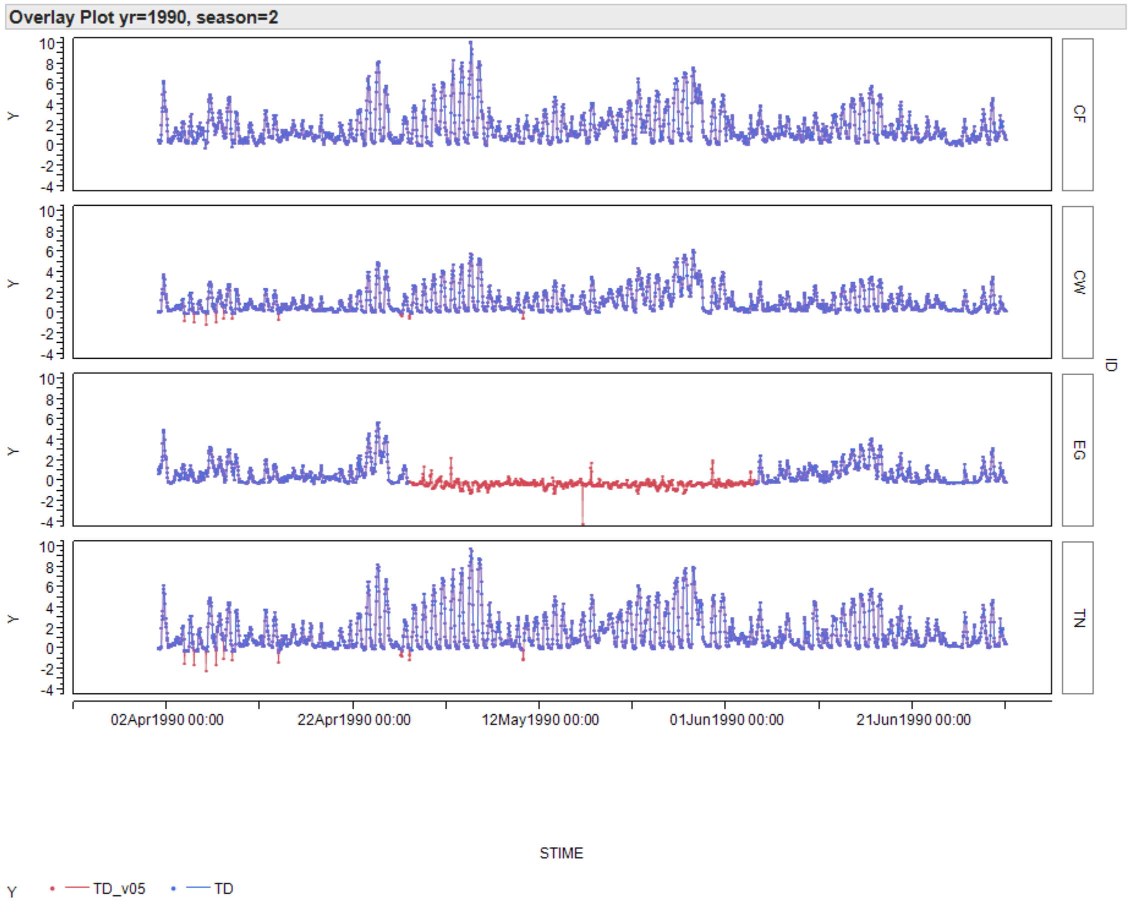




Overlay Plot yr=1989, season=1
Y
CF
CW
EG
TN
ID
05Jan1989 00:00
25Jan1989 00:00
14Feb1989 00:00
06Mar1989 00:00
26Mar1989 00:00
STIME
Y
TD_v05
TD

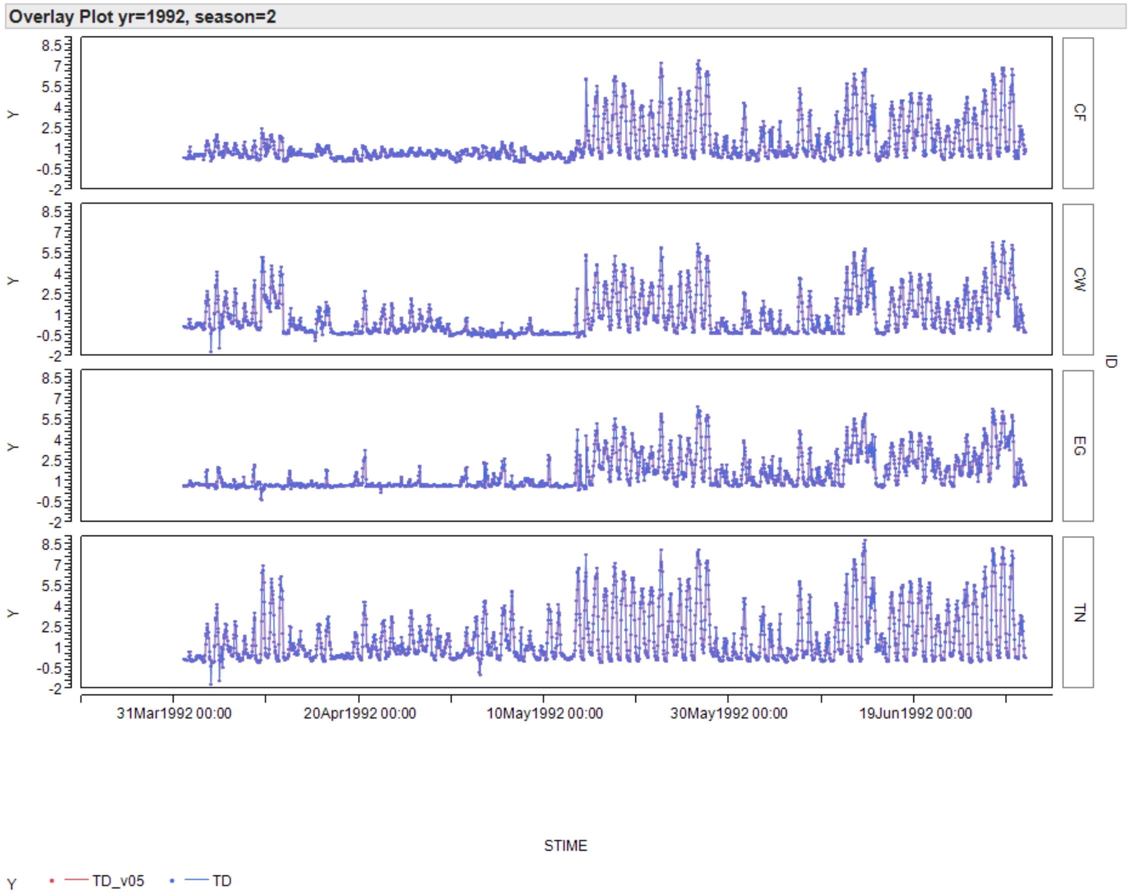
Obvious calibration problems with CW:



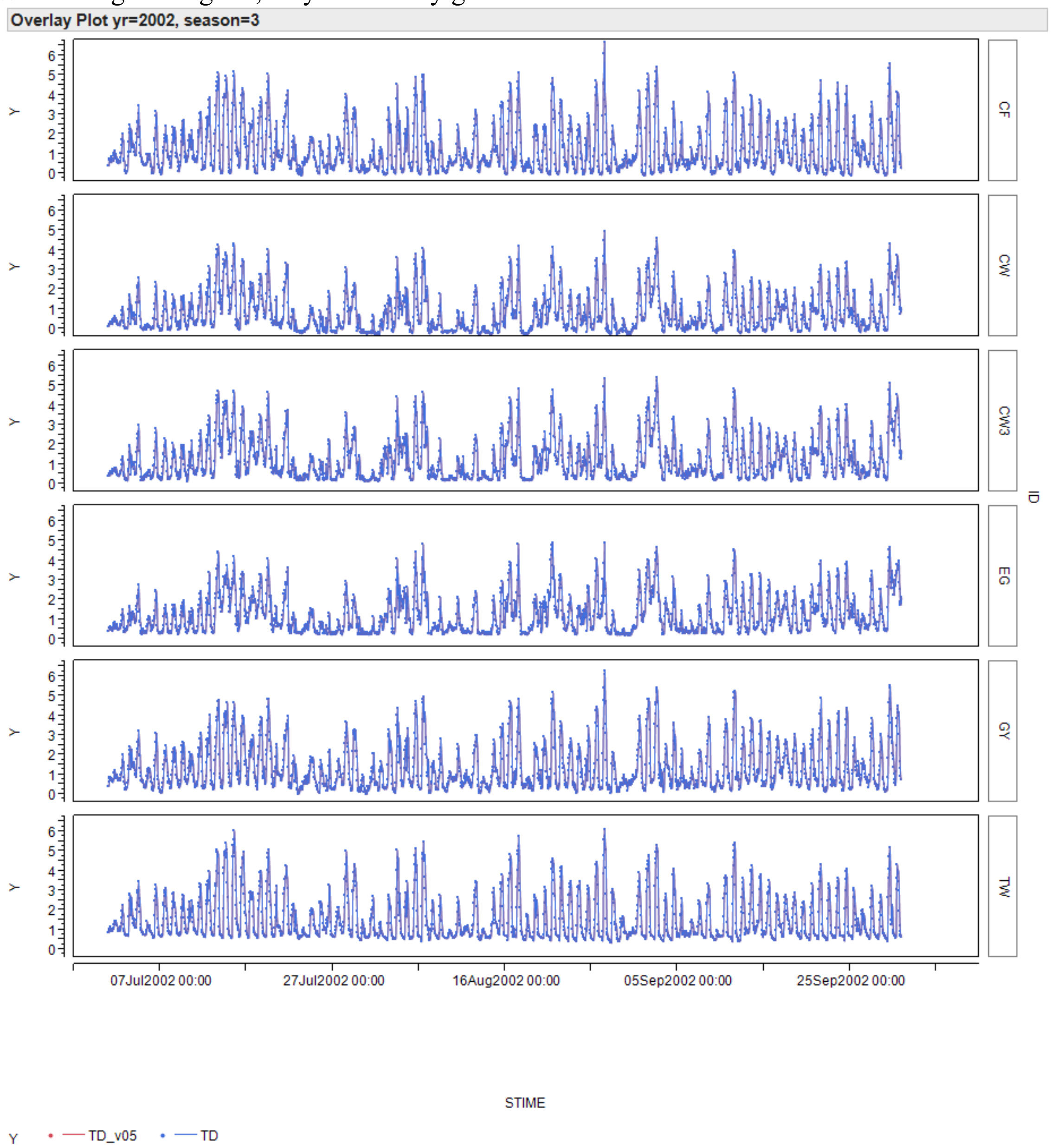


Overlay Plot yr=1990, season=2
Y
CF
CW
EG
TN
ID
02Apr1990 00:00
22Apr1990 00:00
12May1990 00:00
01Jun1990 00:00
21Jun1990 00:00
STIME
TD_v05
TD

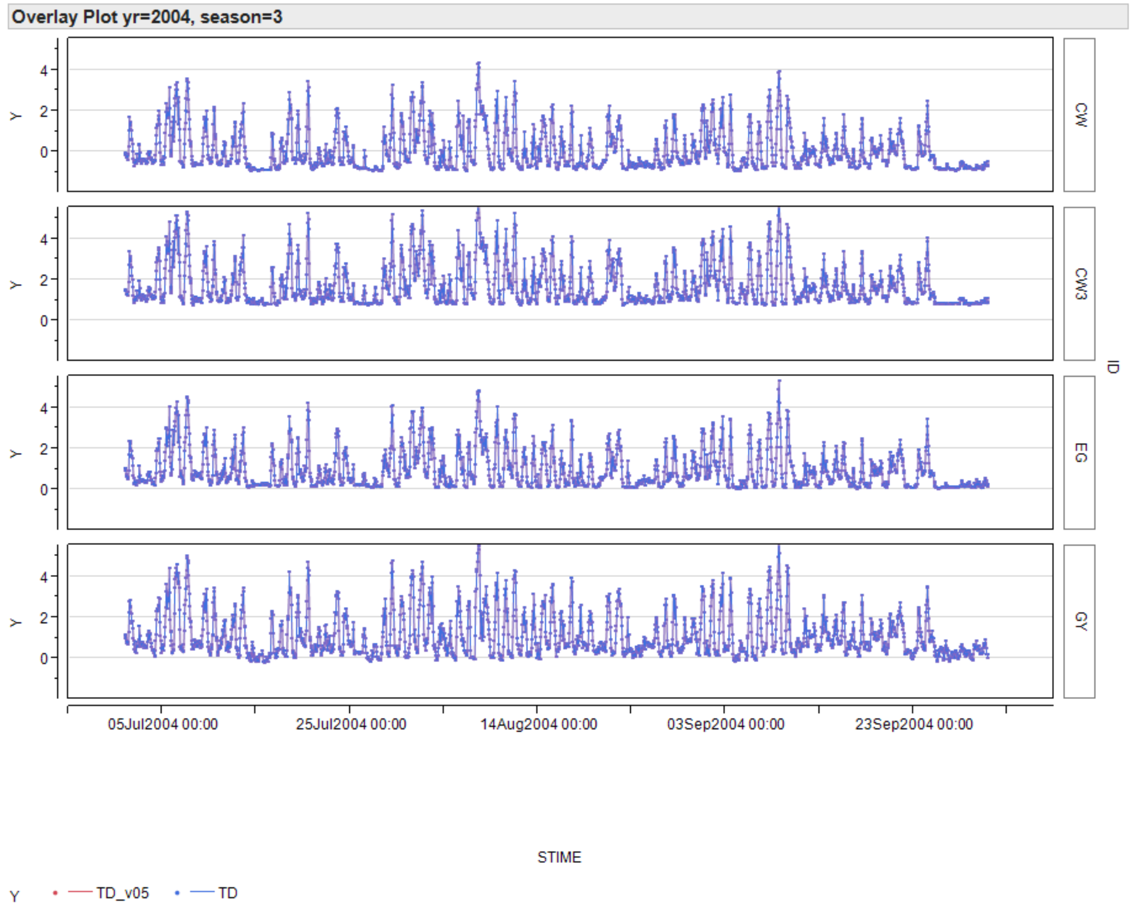
Unclear whether these are calibration problems, or just site-to-site differences... thus I left them alone:



When things look good, they look really good!



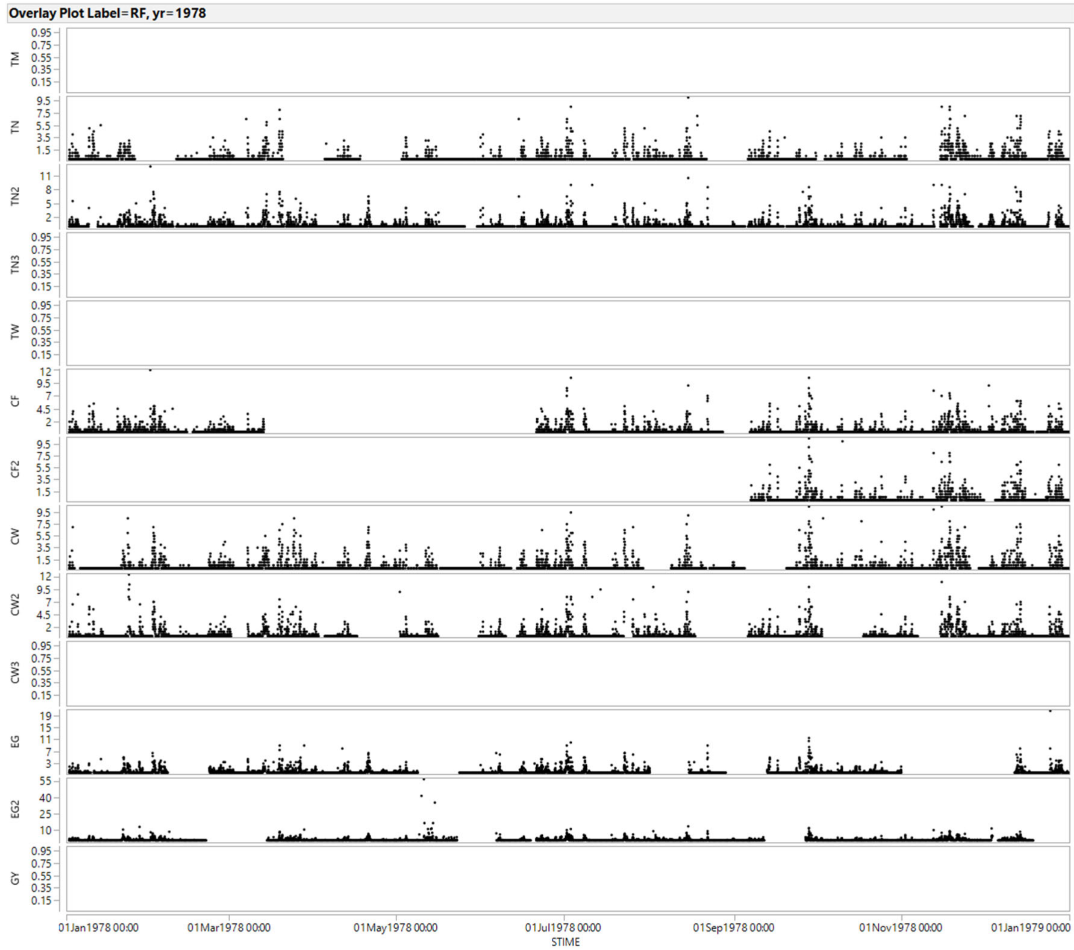
Here everything looks good, but note the different baselines (calibration differences), even at CW and CW3, which are co-located!



Plynlimon AWS data editing notes
J.W. Kirchner 24 November 2020

Combining co-located AWS's into longer-term records

I inspected stacks of strip charts of AWS time series for each year (see below for an example) and removed visibly anomalous data (usually flyers, or strings of zeroes, or obvious corruption by calibration drift).

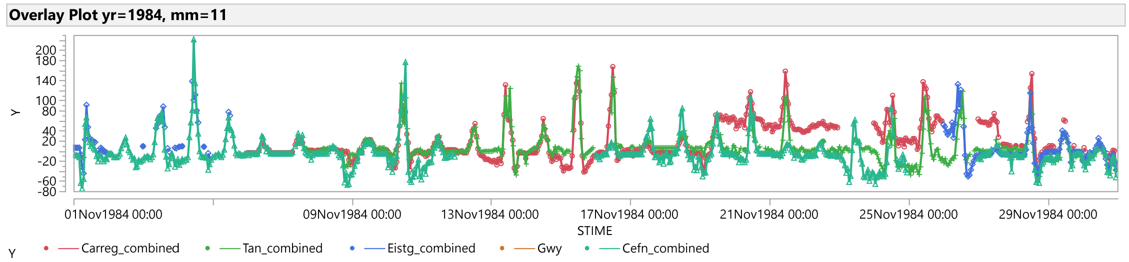


I then averaged the available values for each hour, using all AWS's for each site, as follows:

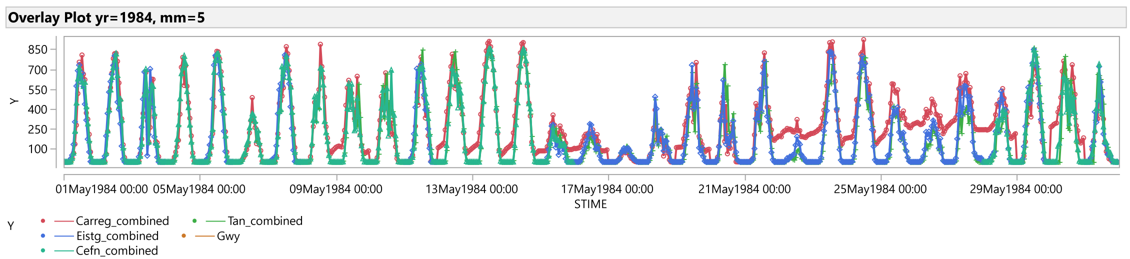
Carreg_combined – CW, CW2, CW3
Cefn_combined – CF, CF2
Eistg_combined – EG, EG2
Tan_combined – TN, TN2, TN3, TM, TW
Gwy_Iago_combined – GY, NI

An alternative approach, following the metadata notes of Mark Robinson, would be to create the following four "meta-stations":
Upper Severn – CW, CW2, CW3
Lower Severn – TN, TN2, TN3, TM, TW
Upper Wye – EG, EG2, NI
Lower Wye – CF, CF2, GY
That approach was not followed here.

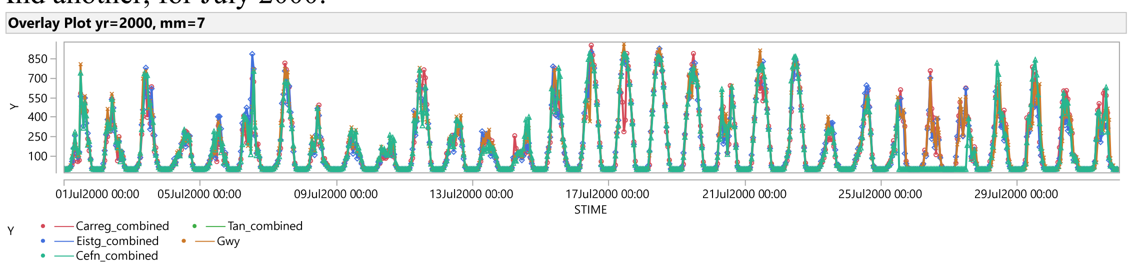
I then superimposed these composites, to check again for outliers. Here is one example, for net radiation in November 1984:

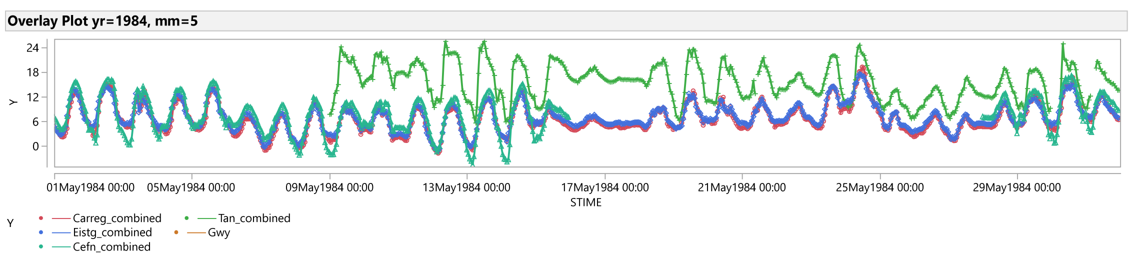


Here is another example, for solar radiation in May 1984:



And another, for July 2000:





Obvious outliers like these were removed.

I replaced about 50 missing precipitation values in 1983 and 1984, plus a few individual missing values in the 1970's, with zeroes (for time steps with no precip values at any AWS's), so that the precipitation record is complete. Although there was no AWS record to verify this choice, the streams showed no response, so if there were any precipitation it must have been minimal.

I then in-filled missing values, using the following orders of precedence (these are ordered by pairwise correlations between rainfall rates on overlapping time steps):

Carreg_infilled:

- Carreg Wen
- Tanllwyth

- Cefn Brwyn
- Gwy_Iago
- Eisteddfa Gurig

Tan_infilled:
- Tanllwyth
- Carreg Wen
- Cefn Brywn
- Gwy_Iago
- Eisteddfa Gurig

Eistg_infilled:
- Eisteddfa Gurig
- Cefn Brwyn
- Carreg Wen
- Tanllwyth
- Gwy_Iago

Cefn_infilled:
- Cefn_Brwyn
- Tanllwyth
- Eisteddfa Gurig
- Carreg Wen
- Gwy_Iago

Gwy_Iago_infilled:
- Gwy_Iago
- Tanllwyth
- Carreg Wen
- Cefn Brwyn
- Eisteddfa Gurig

Fractions of hourly rainfall values that were infilled using records from another AWS are:
Carreg Wen: 6.3%
Tanllwyth: 10.7%
Eisteddfa Gurig: 11.1%
Gwy_Iago: 10.0%
Cefn Brwyn: 6.4%

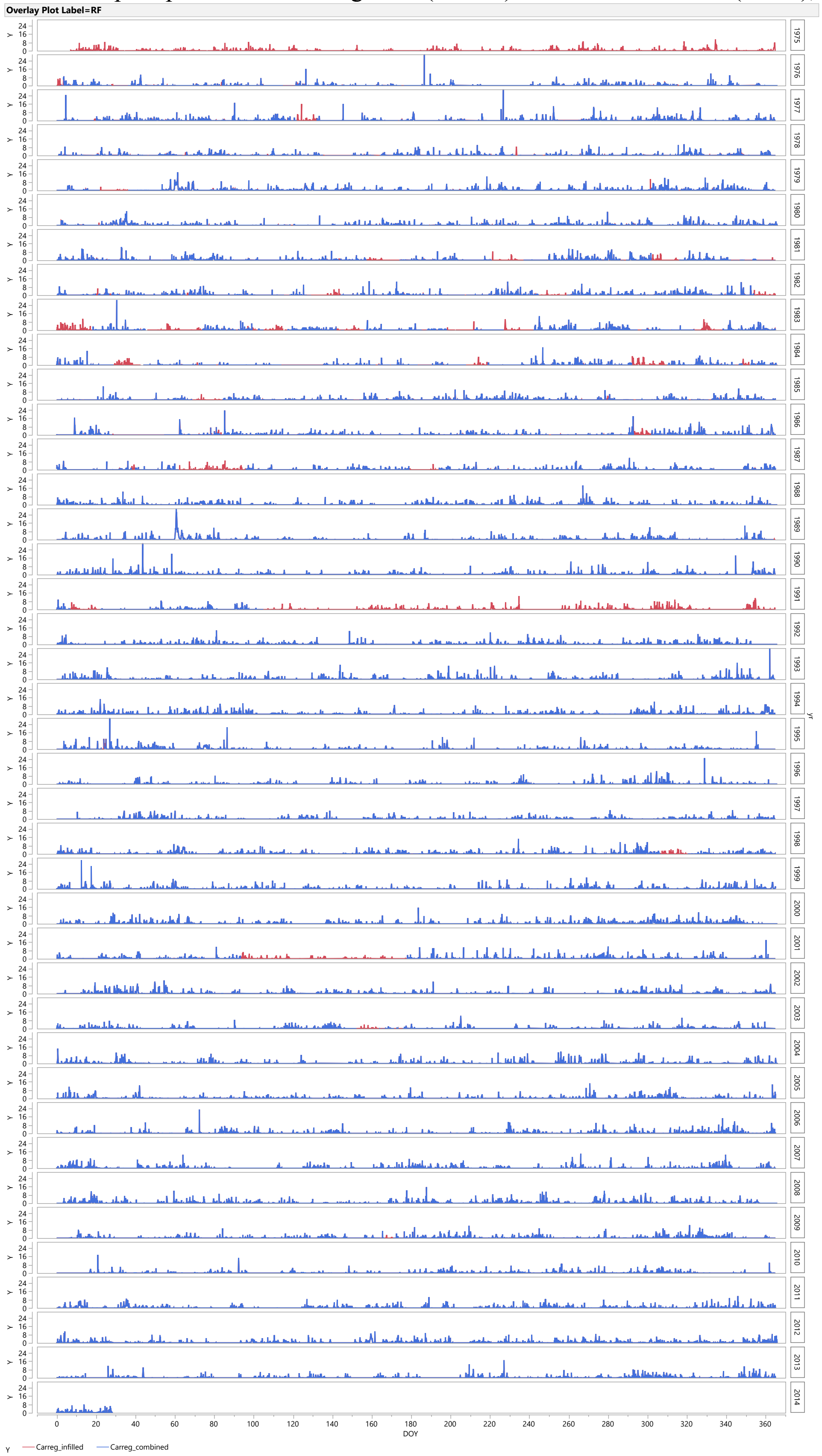
Lastly I trimmed each site's combined, in-filled data to correspond to the max/min date ranges of the "native" AWS records (but these are still included in the big blocks plotted below).

I also created a virtual record, "AllSitesAWS", which is the average of all five sites (whichever ones are available for any time step).

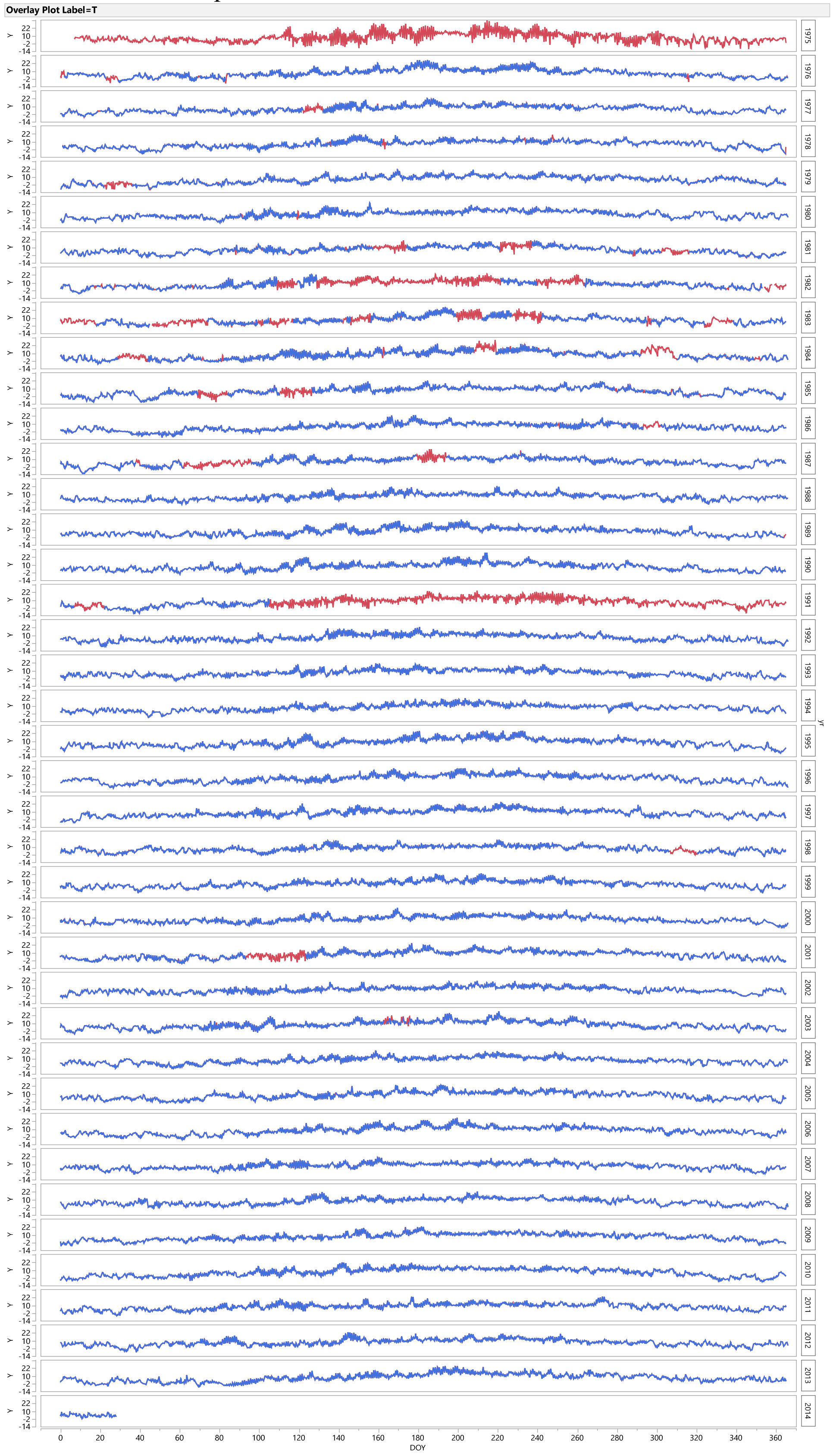
Note that I did not use any calibration equations to correct for any site-to-site differences (which I assume to be small).

Wind directions in the combined and in-filled data are vector averages (including windspeed), to prevent (for example) 5° and 345° arithmetically averaging to 175° instead of 355°.

Here is the precip record for Carreg Wen (in blue) with in-filled values (in red), for example:



And here are the temperature records:



Now, which precip records should be used for which discharge sites? Here are the priority orders:

Upper Hafren:
- Carreg Wen

Lower Hafren, Upper Hore, Lower Hore, Severn:
- Mean (Carreg Wen, Tanllwyth)

Iago, Gwy:
- Iago_Gwy
- Eisteddfa Gurig

Cyff:
- Mean(Eisteddfa Gurig, Cefn Brwyn)
- Iago_Gwy

Wye (Cefn Brwyn):
- Mean(Eisteddfa Gurig, Cefn Brwyn, Iago_Gwy)

Final file is "Plynlimon compiled AWS and hourly Q.jmp" or "Plynlimon compiled AWS and hourly Q.txt".

Note that 1970s-era AWS data have relatively more glitches than later years. The most reliable data appear to be from the 1990s and early 2000s.

Upper Hafren records are only available for 2005-2010 (6 years) in these data sets.