Use of Long Time-Series ACRF Measurements to Improve Data Quality Analysis

Sean Moore
Mission Research and Technical Services
Santa Barbara, CA

ARM Data Quality Office
University of Oklahoma
Norman, OK

ARM Data Quality Office

Full Time Staff	*Student Analysts*
•*Ken Kehoe*	•*Nathan Hiers (Sr)*
•*Randy Peppler*	•*Stephen Mullens (Sr)*
•*Karen Sonntag*	•*Kimberly Rabon (Jr)*
•*Justin Monroe*	•*Lacey Evans (Jr)*

ARM Data Quality Office, National Weather Center
The University of Oklahoma, Norman, OK

Sean Moore
Mission Research and Technical Services
Santa Barbara, CA

Daily Quality Checks

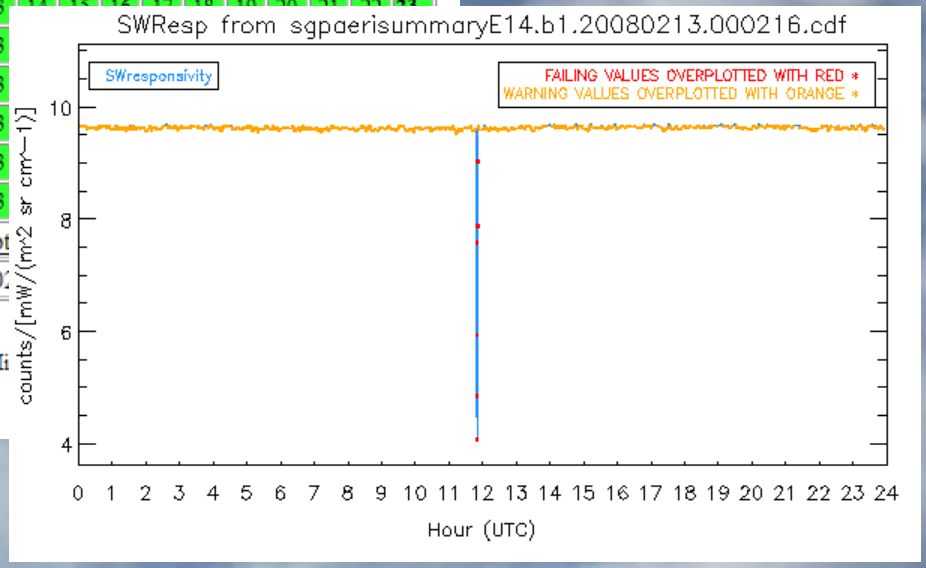
- Automated software checks every measurement for outliers against some pre-defined limits.
- Suspicious data are flagged
- Statistics are collected regarding how often and when data are flagged
- Diagnostic plots are generated
- Analysts review metrics and plots

DQ HandS Metrics and Plots

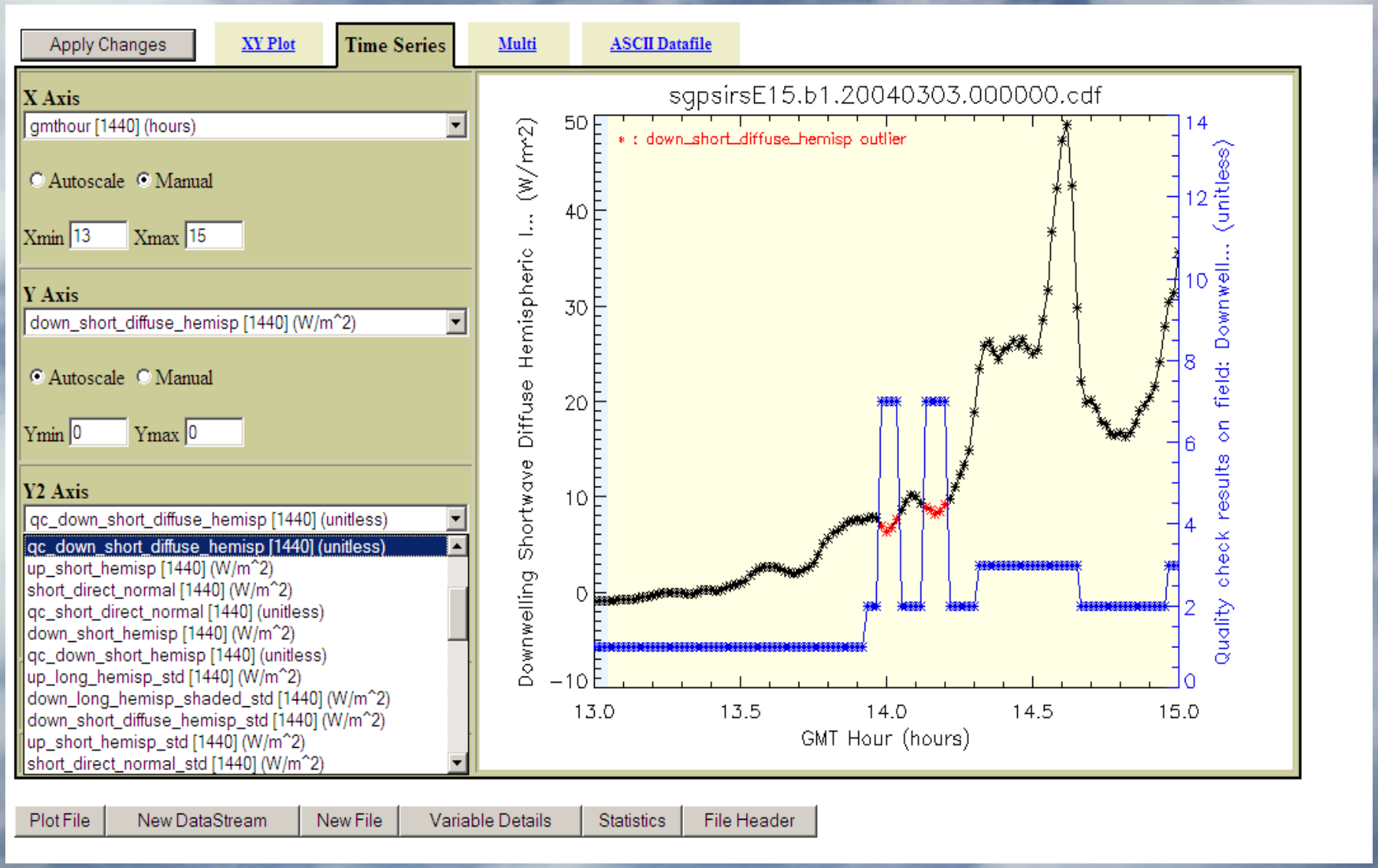
Move the cursor over any bold cell to identify failures

E14	sgpaerisummary status by hour for 20080213																							
ABBmaxTempDiff	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	**23**
HBBmaxTempDiff	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	**23**
HBBtempDrift	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	**23**
LW_HBB_NEN	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	**23**
LWresponsivity	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	**23**
SW_HBB_NEN	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	**23**
SWresponsivity	00	01	02	03	04	05	06	07	08	09	10	**11**	12	13	14	15	16	17	18	19	20	21	22	**23**
atmosphericPressure	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	**23**
atmosphericRelativeHumidity	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	**23**
calibrationCBBtemp	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	**23**
calibrationHBBtemp	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	**23**
frontEndFanAirTempDiff	00	01	02	03	04	05	06	07	08	09	10	11	12	13										
hatchOpen	00	01	02	03	04	05	06	07	08	09	10	11	12	13										
maxSampleStdDev	00	01	02	03	04	05	06	07	08	09	10	11	12	13										
outsideAirTemp	00	01	02	03	04	05	06	07	08	09	10	11	12	13										
rainSensorIntensity	00	01	02	03	04	05	06	07	08	09	10	11	12	13										
sceneMirPosEncoderDrift	00	01	02	03	04	05	06	07	08	09	10	11	12	13										
Diagnostic Plots													NCVweb Interactive Plot											
Data Source: sgpaerisummaryE14.b1 for 20080																								

= Metric Passing 100% = Metric Passing < 75% = Mi
= Metric Passing 75%-100% = Data Not Available



Interactive Plotting



Multi-Day, Multi-Facility Plot
EBBR Soil Temperature at SGP E2, E4, and E7

Apply Changes | XY Plot | Time Series | Multi | ASCII Datafile

Facilities:

☑ E2 ☑ E4 ☑ E7 ☐ E8
☐ E9 ☐ E12 ☐ E13 ☐ E15
☐ E18 ☐ E19 ☐ E20 ☐ E22
☐ E26 ☐ E27

Select All

X Axis

DOY [144] (days)

◉ Autoscale ○ Manual

Xmin 0 Xmax 0

Y Axis

ts1 [144] (C)

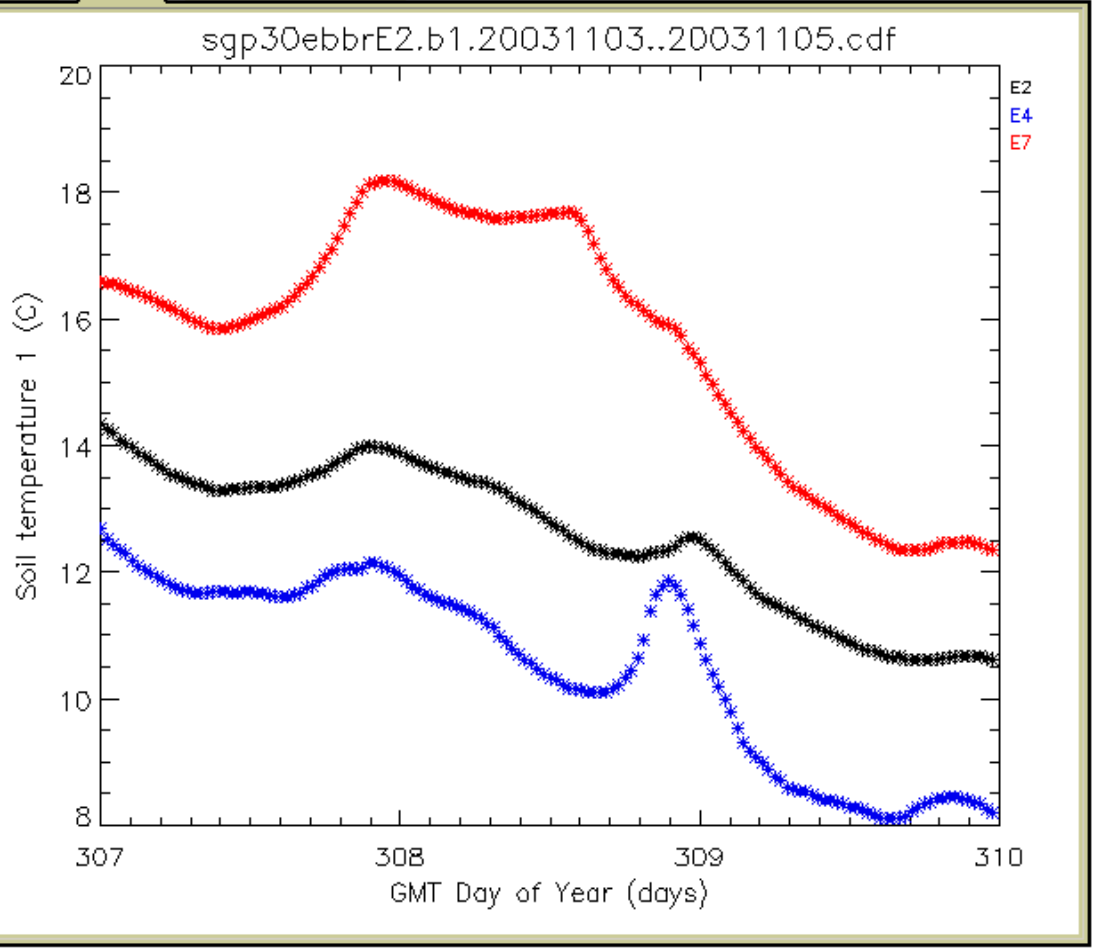
◉ Autoscale ○ Manual

Ymin 0 Ymax 0

Symbol: ○ None ○ Plus ◉ Star ○ Dot

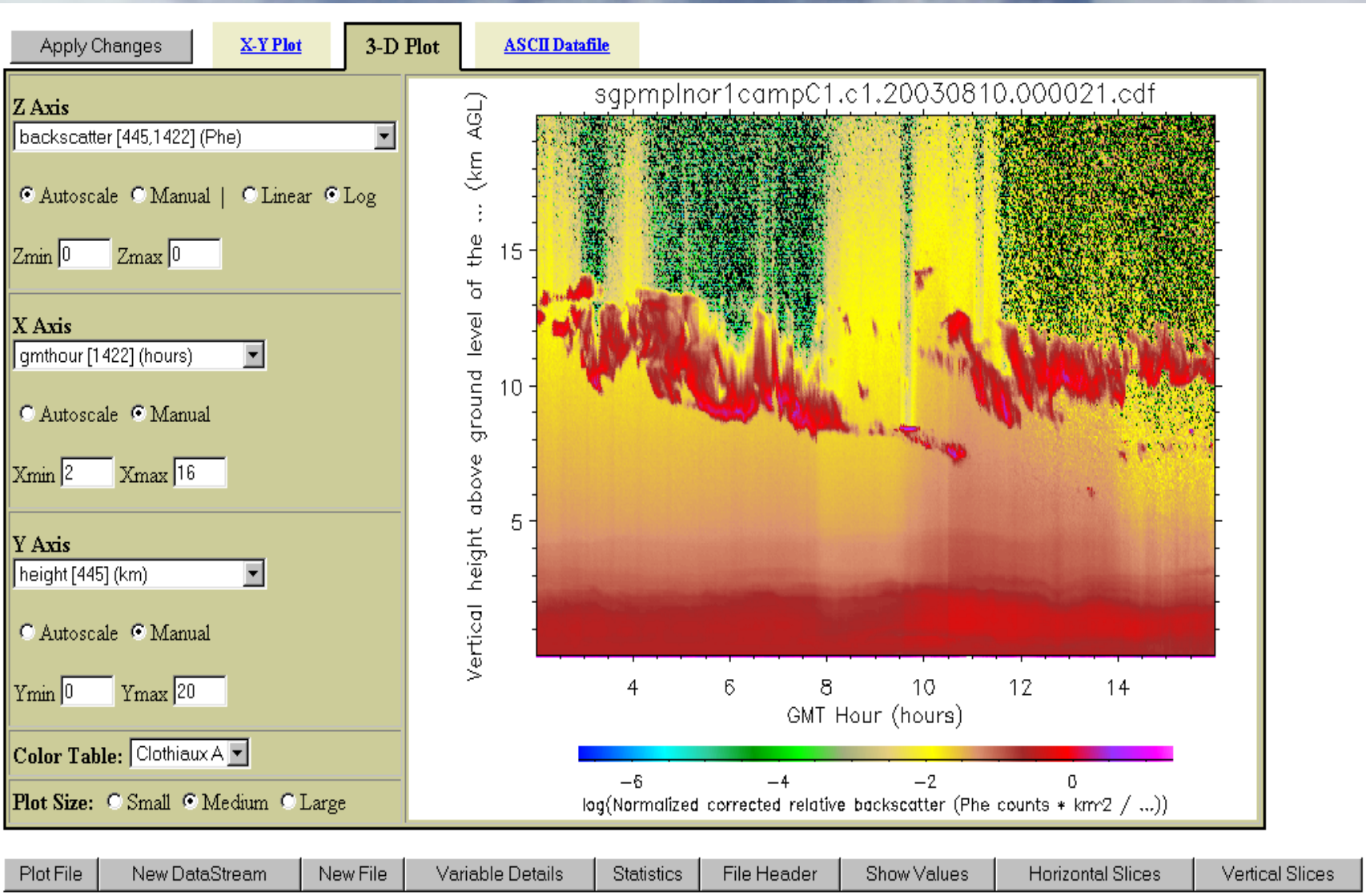
Line: ☐

Plot Size: ○ Small ◉ Medium ○ Large

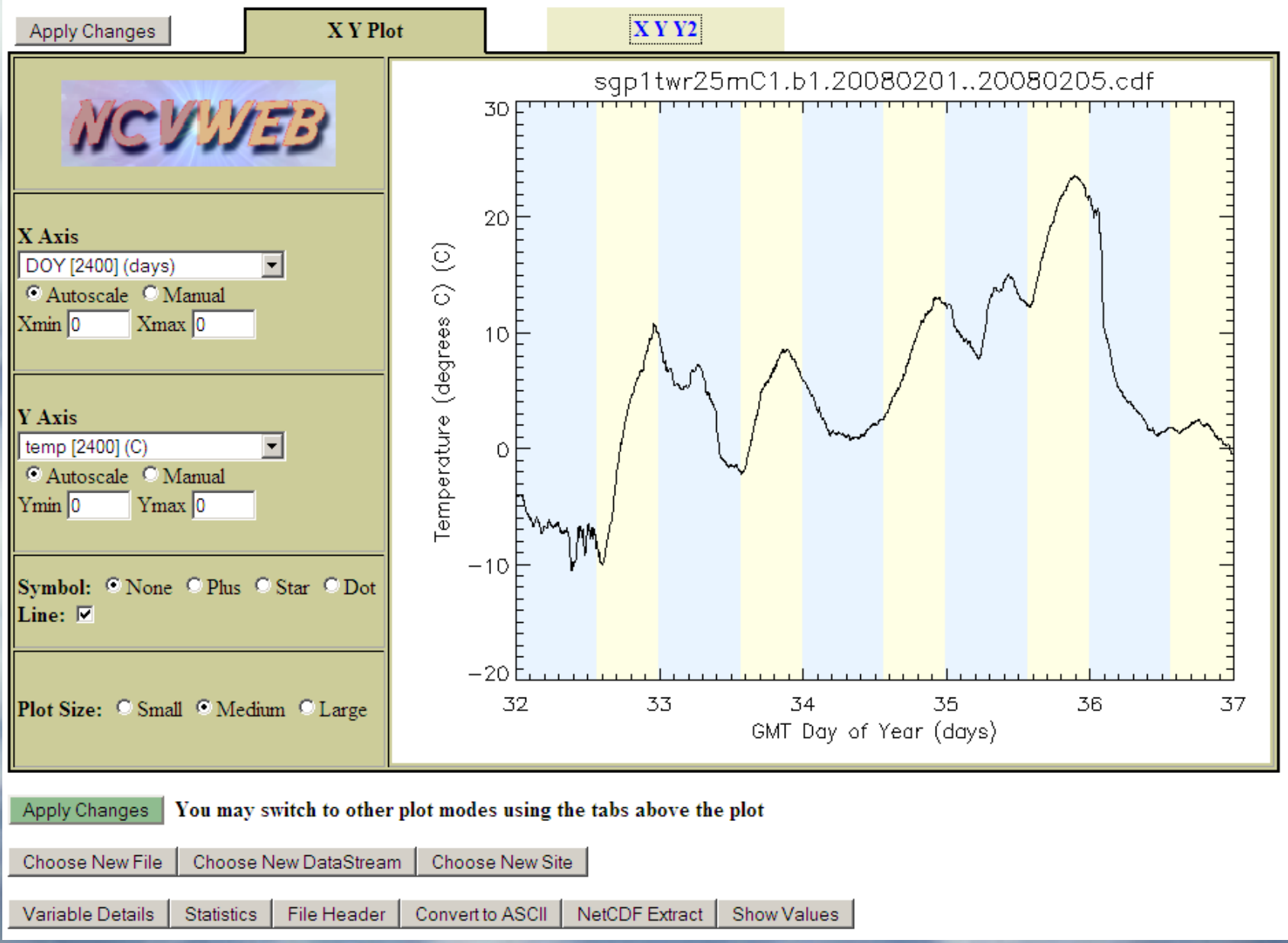


Plot File | New DataStream | New File | Variable Details | Statistics | File Header

Multi-dimensional Data



Custom Data Extraction



DQ HandS Plot Browser

ARM

Setup Search

Search Site

FKB
NIM
NSA
PYE
SGP
TWP

Datastream

fkbaerich1
fkbaerich2
fkbaeriengineer
fkbaerisummary
fkbgndrad60s
fkbmet
fkbmfrsr
fkbmplpol
fkbmwrhf
fkbmwrlos
fkbmwrp
fkbnfov2ch
fkbskyrad60s
fkbsondewnpn
fkbtsiskycover

Search Date

2007 | 5 | 8

Search Style

List Thumb

4 days at a time

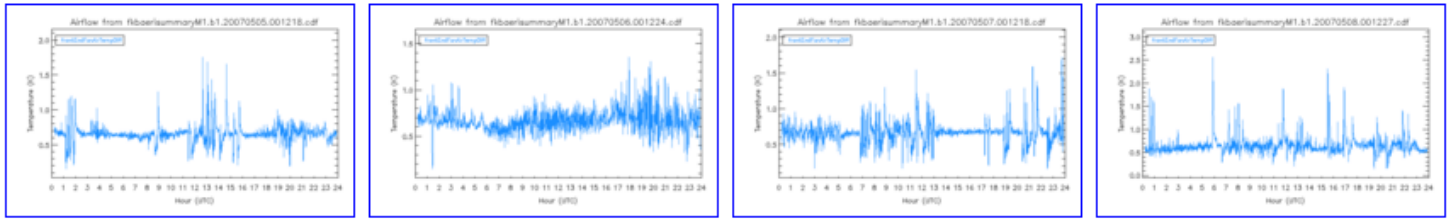
Get Plots Reset

Previous 4 days Next 4 days

Searching Dates: 20070505-20070508

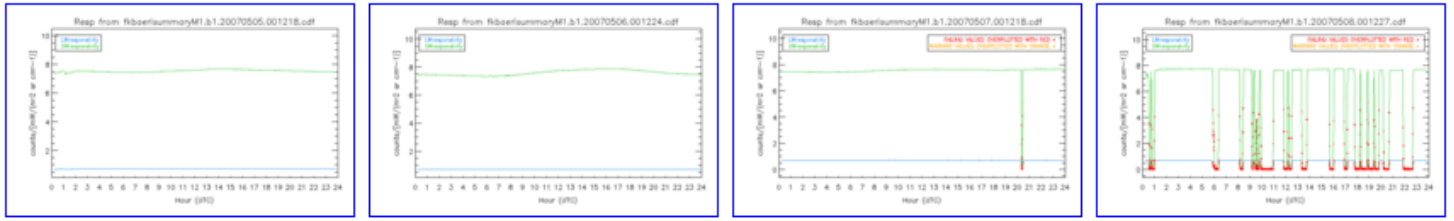
fkbaerisummary

M1.Airflow



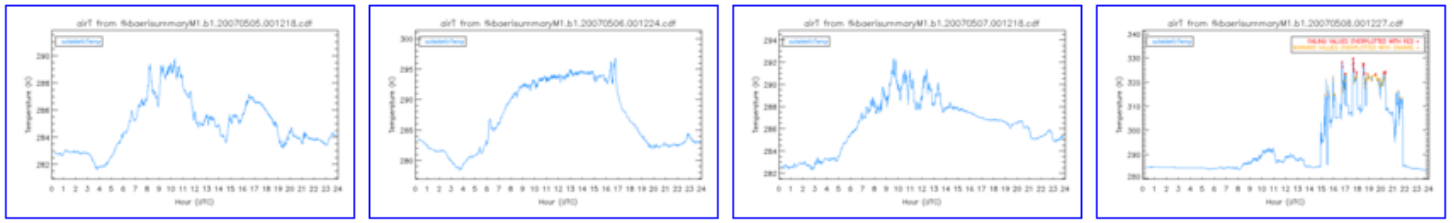
20070505 20070506 20070507 20070508

M1.Resp



20070505 20070506 20070507 20070508

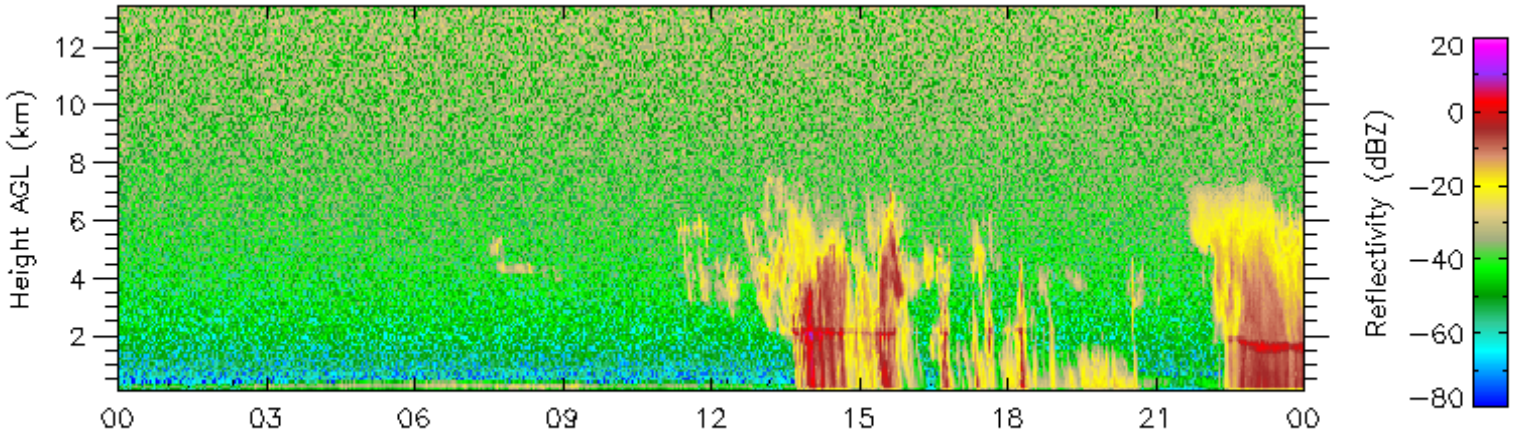
M1.airT



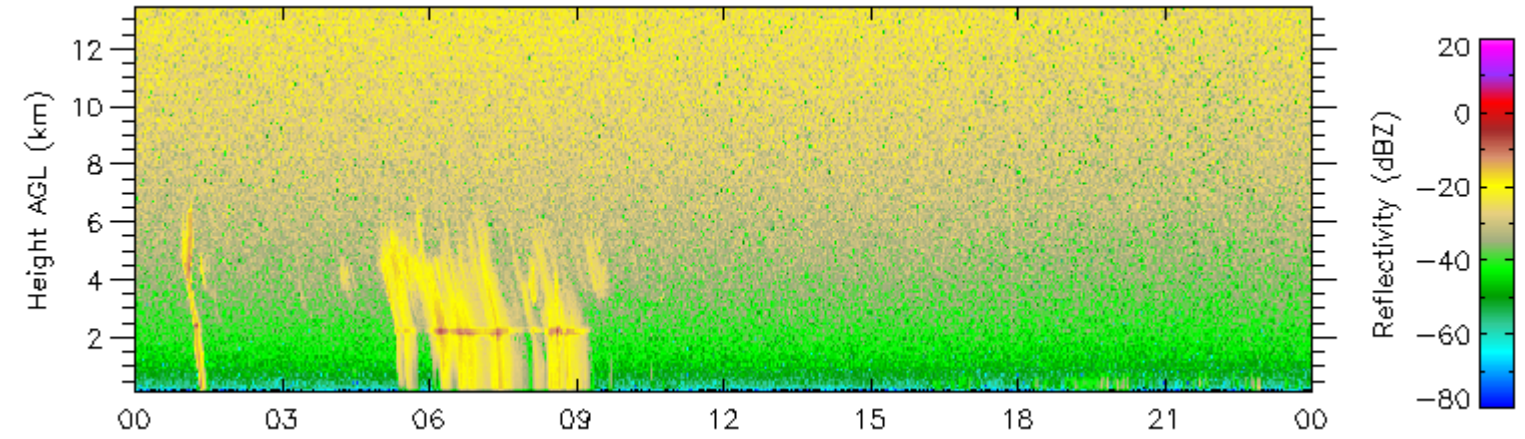
20070505 20070506 20070507 20070508

Radar – Two day compare

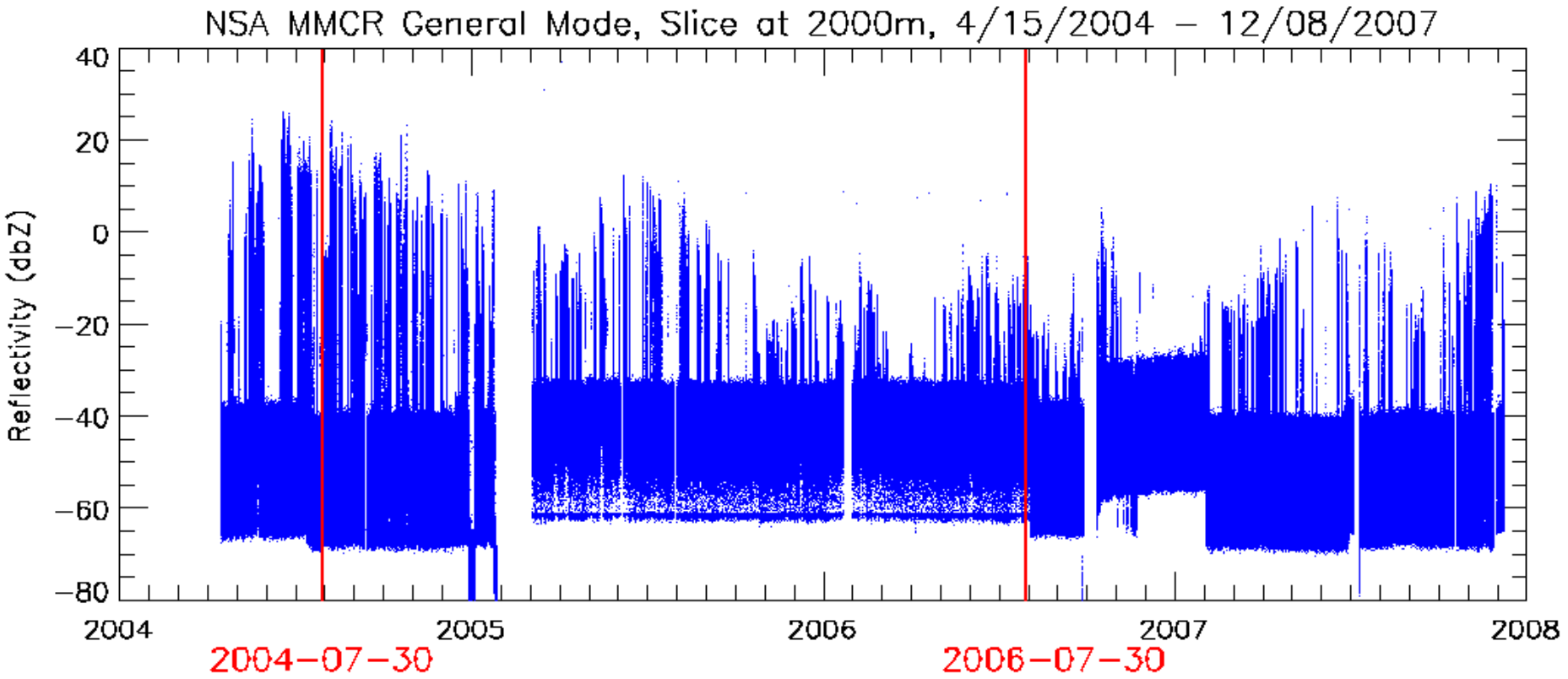
NSA MMCR 2004-07-30



NSA MMCR 2006-07-30



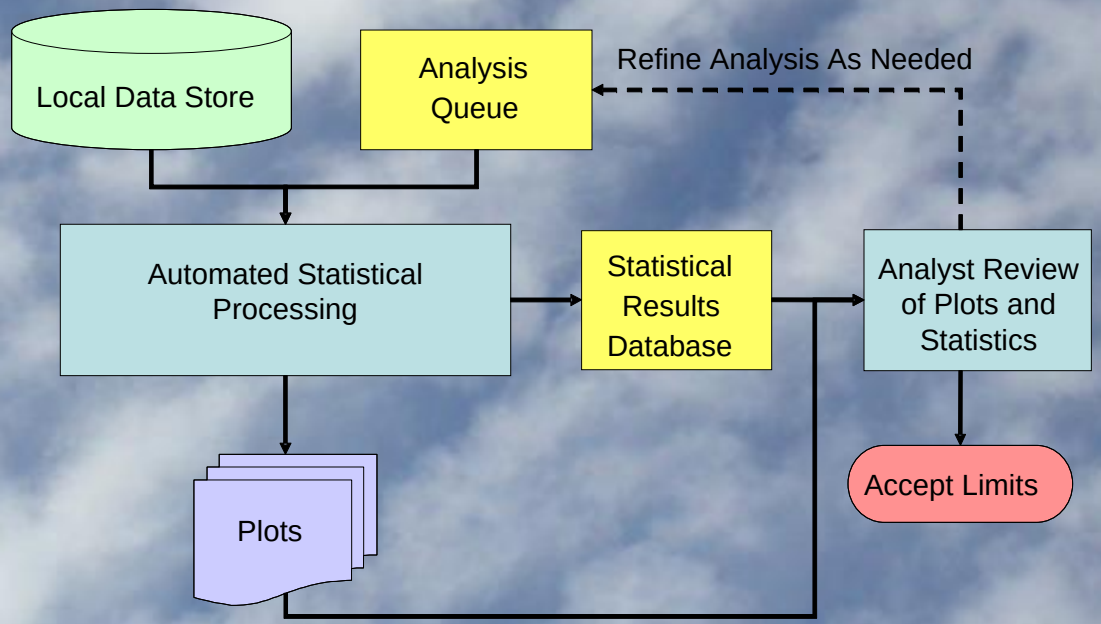
Four Year Slice of Reflectivity



ARM*STAR

ARM STatistical Analysis and Reporting System

- Combines data from many months, years or decades
- Computes statistics over custom time ranges
- Produces long-term plots
- Frequency Distributions
- Feedback loop with analysts to refine results
- New limits in a database

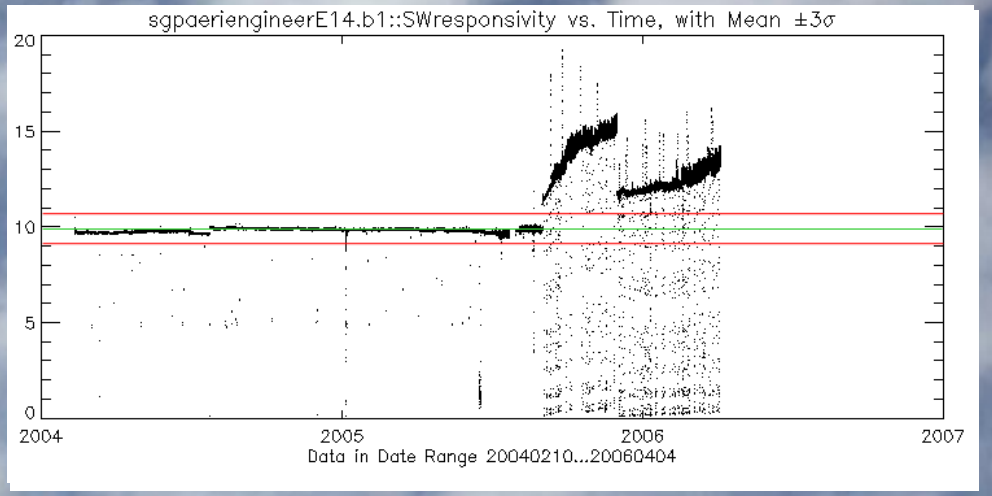


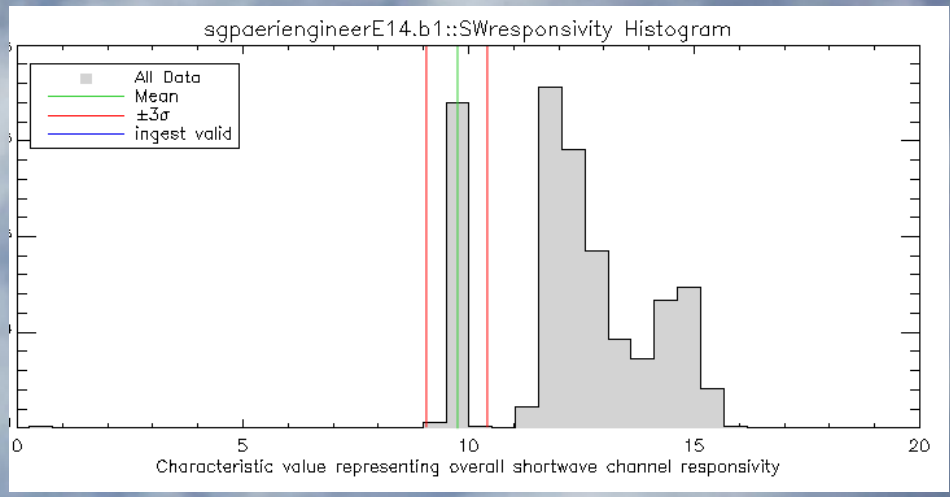
Determining New Limits

- Review historical data to decide on appropriate limits

- Detect instrument problems before problems affect primary measurements.

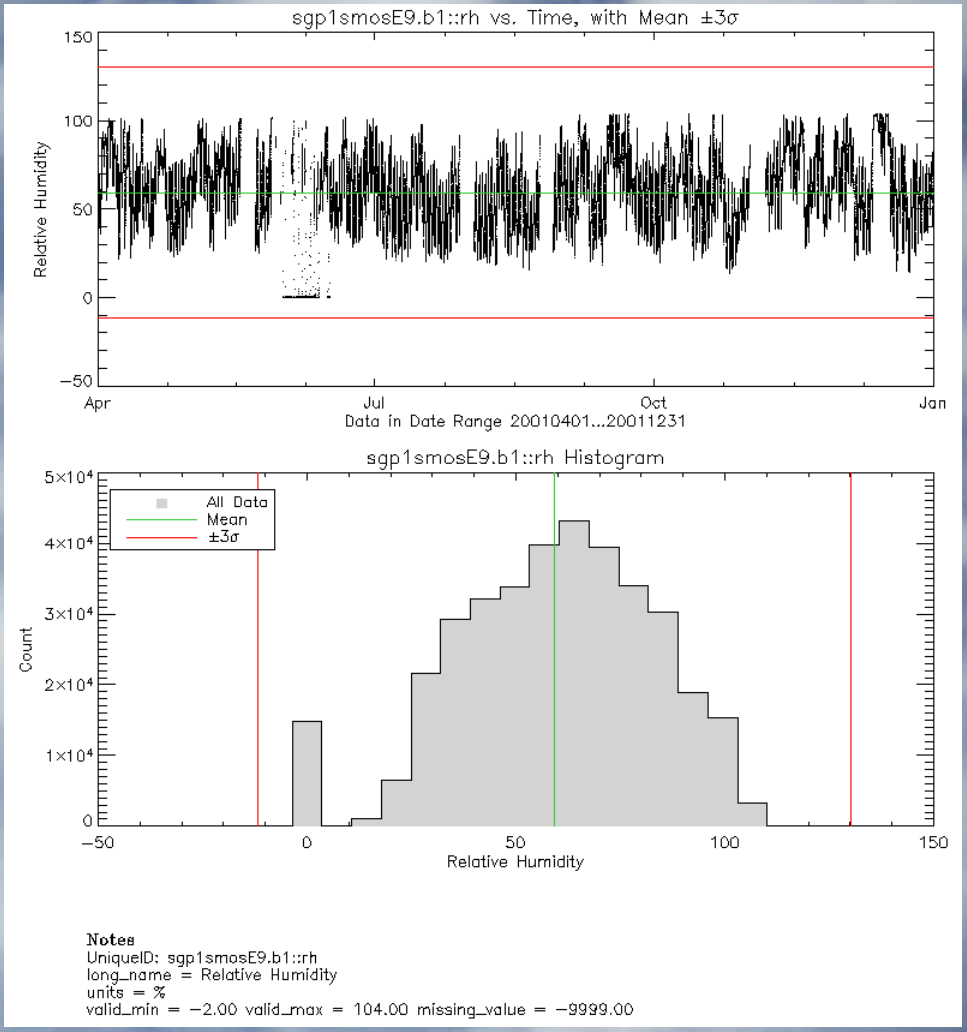
- Multimodal frequency distribution shows abrupt changes that warrant investigation





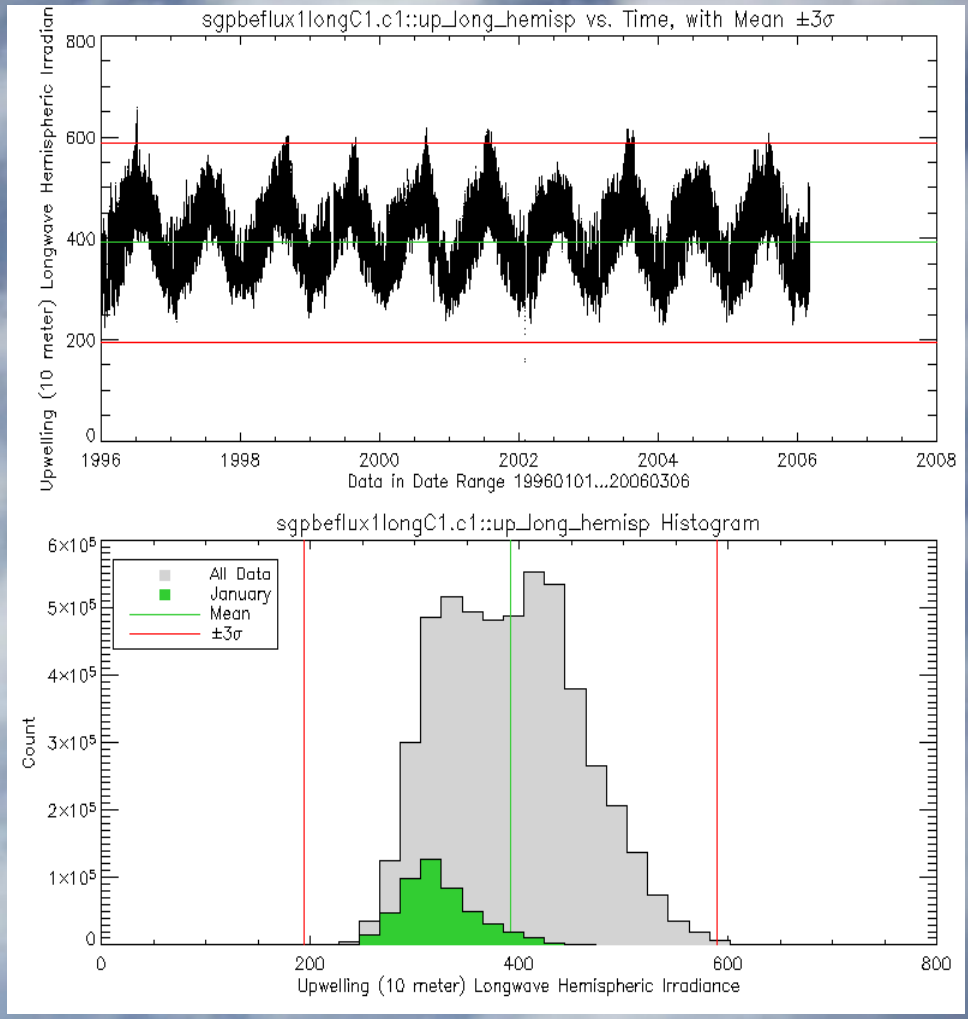
Example: Improve QC Limits

- SMOS b-level NetCDF files define valid rh range as -2 to 104%
- Frequency distribution plot helps to identify problems (spike on left)
- Analysis suggests existing range may be too broad.



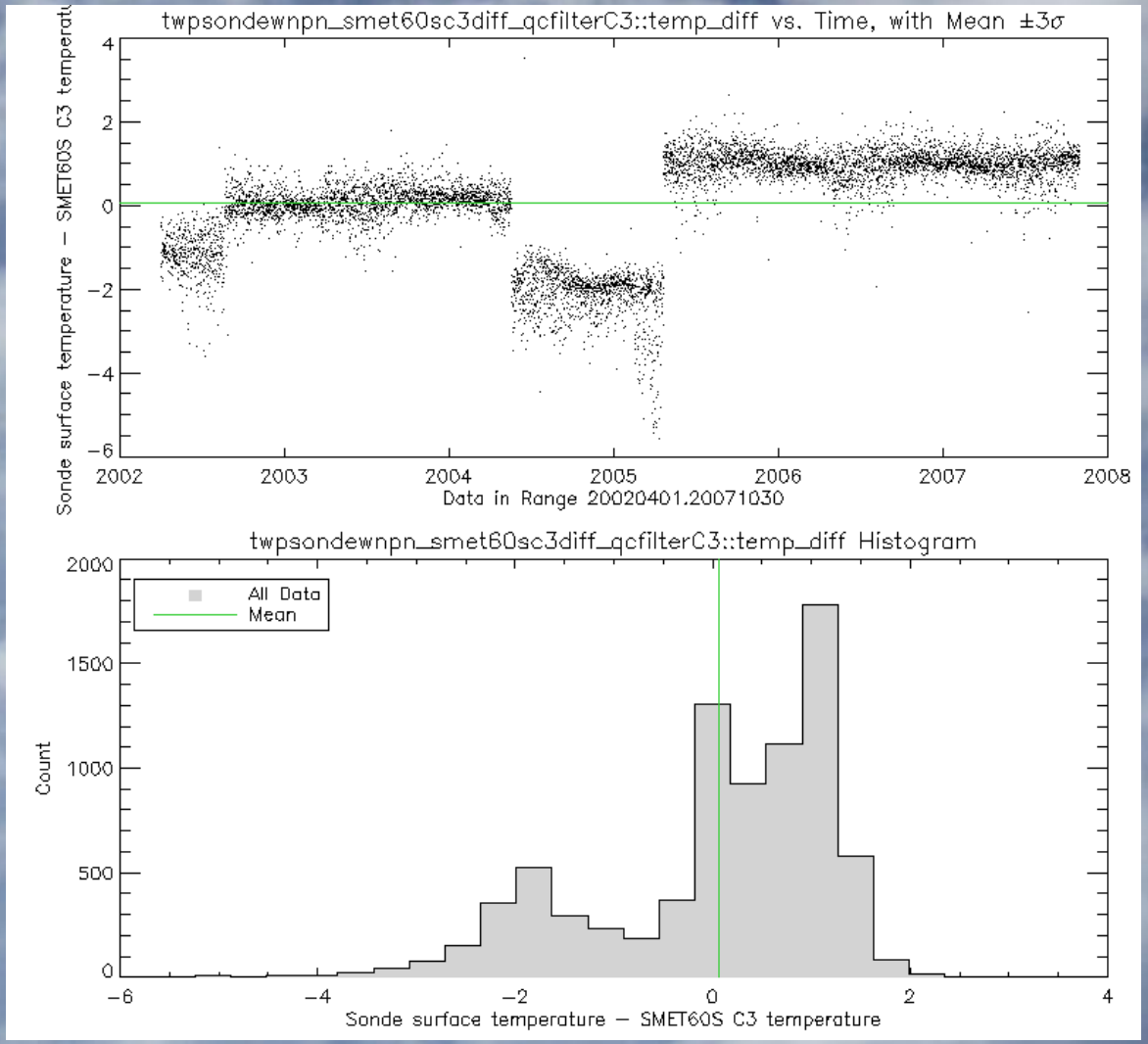
Example: Monthly Limits

- No limits defined in NetCDF files for BEFLUX VAP
- Quickly assess typical behavior with plot
- Frequency distribution shows monthly data range distinct from overall range (January colored green)
- Analysis suggests monthly varying limits may be appropriate for data



Instrument Comparisons

- Compare instruments to detect subtle problems
- Difference of Sonde surface temperature and SMET temperature
- 6 years of data
- Multiple issues discovered and resolved after this analysis



Summary

- Statistics from historical data are useful to set or improve data quality range limits used by our automated metrics software.
- Long time-series plots useful to visually detect trends and instrument degradation problems.
- Historical distributions are useful visual aids to compare current data to typical data.
- These techniques are being implemented to improve ACRF data quality analysis!