Some considerations for computing large-scale forcing (LSF) dataset

Paul E. Ciesielski and Richard H. Johnson

Colorado State University

12 September 2011

An important goal of MC3E is to compute a well-constructed, high-quality forcing dataset for input into CRM simulations.

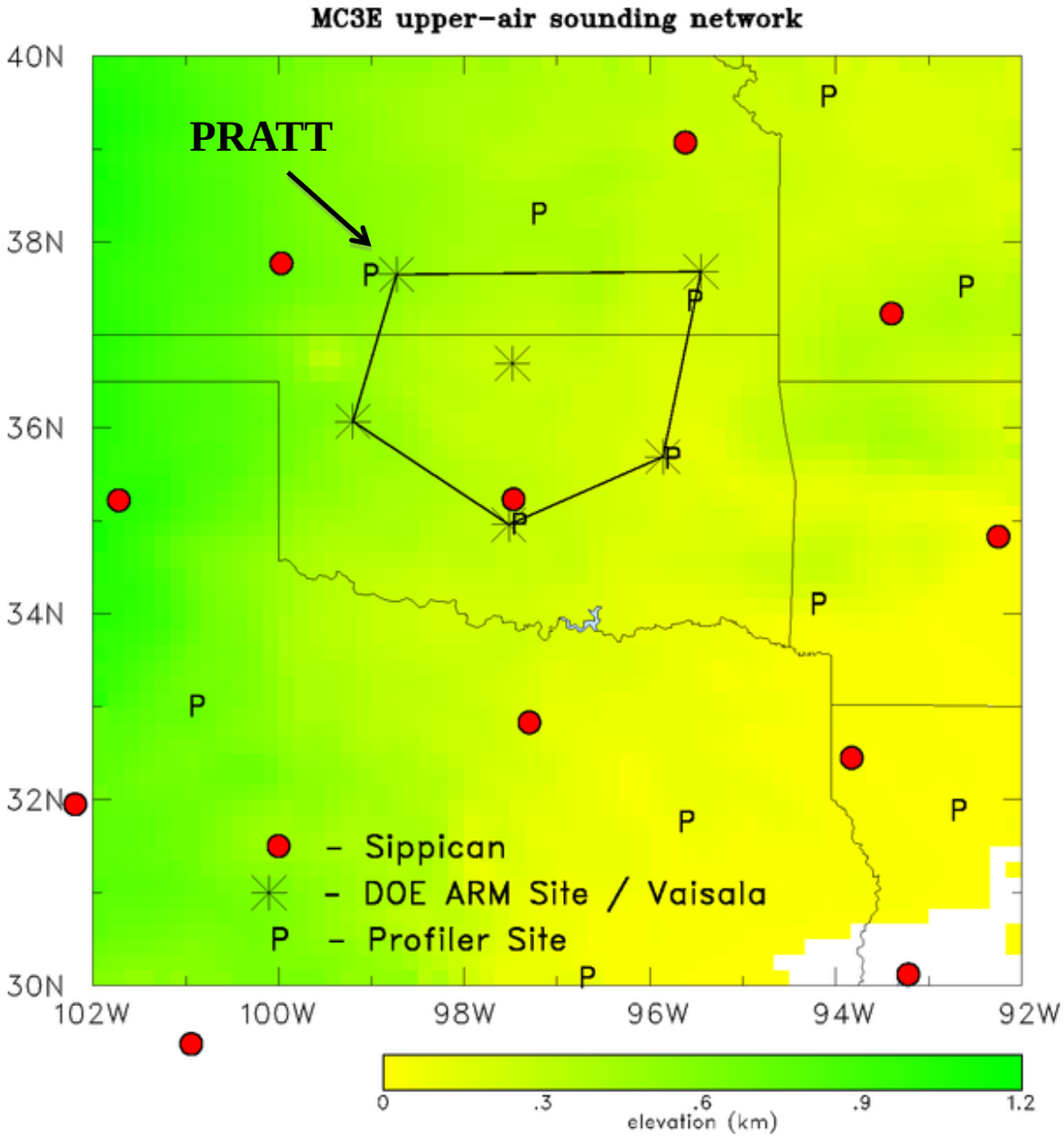
Achieving this goal begins with a high-quality sonde dataset. As eloquently stated by Vic Ooyama in his GATE sounding analysis paper, “To make gold, one must start with gold.” (Ooyama 1987)

How can these sonde observations taken during MC3E be refined to meet the scientific objectives of the experiment?

Two issues come to mind:

-humidity biases in the Vaisala sondes used at the 6 enhanced sites

-problems with the surrounding operational network



Request hi-res sonde data from NOAA from 8-10 operational sites surrounding MC3E domain.

MC3E sonde network embedded within the NWS operational sonde network and NOAA profiler network

Sites in the enhanced network used Vaisala RS92 sondes which have a well-documented daytime dry bias. (Collocated ground based GPS retrievals of PW can be used to identify these errors and validate their corrections.)

In 2005, the NWS began upgrading its sites with Sippican systems.

A number of serious problems have been identified with the Sippican systems. Documented by Bob Maddox at *http:/www.madweather.com*, these problems include:

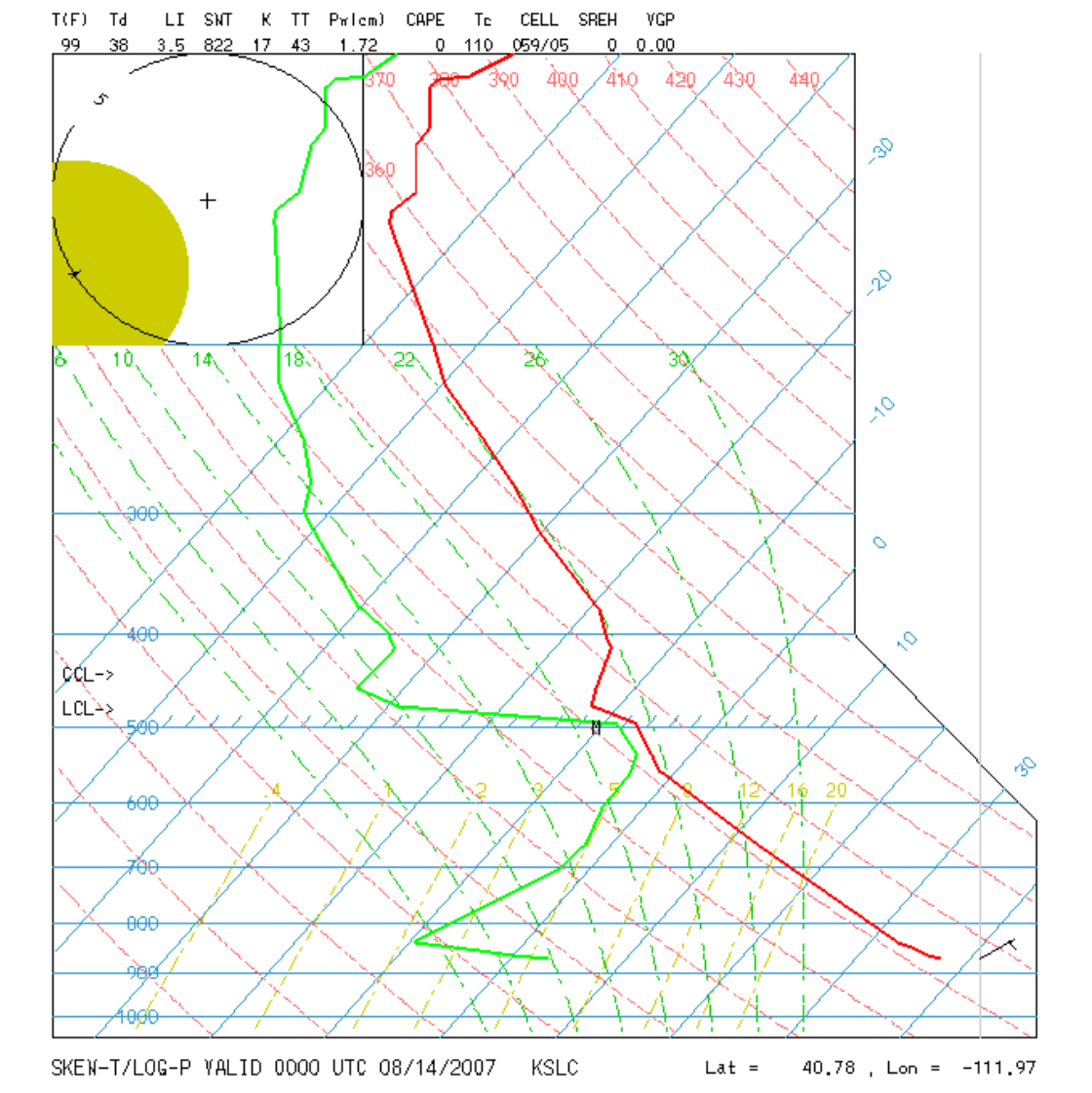
- extremely dry layers above the surface with a slow recovery leading to soundings that are too dry

- slow response to dry layers when a sonde is launched in a relatively humid environment, leading to occasional soundings are too wet

- both the thermistor and hygristor are poorly ventilated so that one or both are easily wetted, leading to a multitude of problems caused by the latent heat effects on the sensor. Depending on the site, Maddox found that 25%-49% sondes were impacted showing super-adiabatic lapse rates above cloud layers.

- soundings are sometimes missing winds through significant layers which may be related to sonde's internal electronic components.

SLC Sippican sounding on 14 August 2007



- missing winds except at surface
- too dry in low-levels
- superadiabatic layer above cloud layer at 500 hPa caused by "wetbulbing" i.e., rapid cooling that occurs when water/ice evaporate/sublimate from the thermistor as it passes out of the cloud.

Unless flagged these problems with Sippican moisture, temperature, and wind fields will adversely impact analyses in the MC3E network.

Stages to developing a research-quality sonde dataset

Convert high-vertical resolution soundings in various formats into a single easily-read format. ➠ **Level 1**

Process high resolution sondes with automated software (e.g., ASPEN) which removes unreliable data as determined by quality control (QC) checks. ➠ **Level 2**

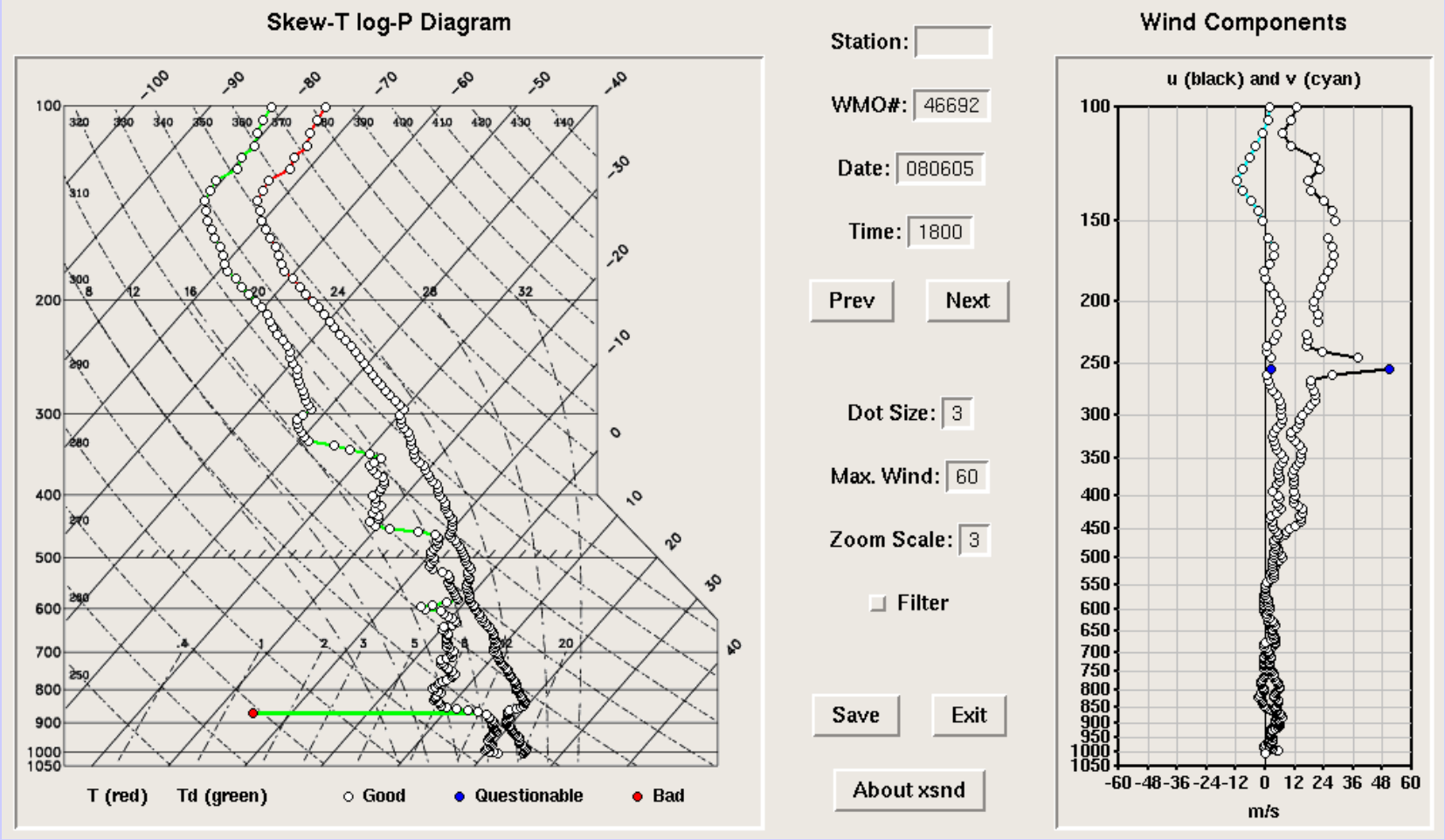
Identify biases in data and correct them if possible (e.g. humidity biases can be reduced using a variety of methods such as intercomparison studies, statistical techniques, lab-developed algorithms, etc.). ➠ **Level 3**

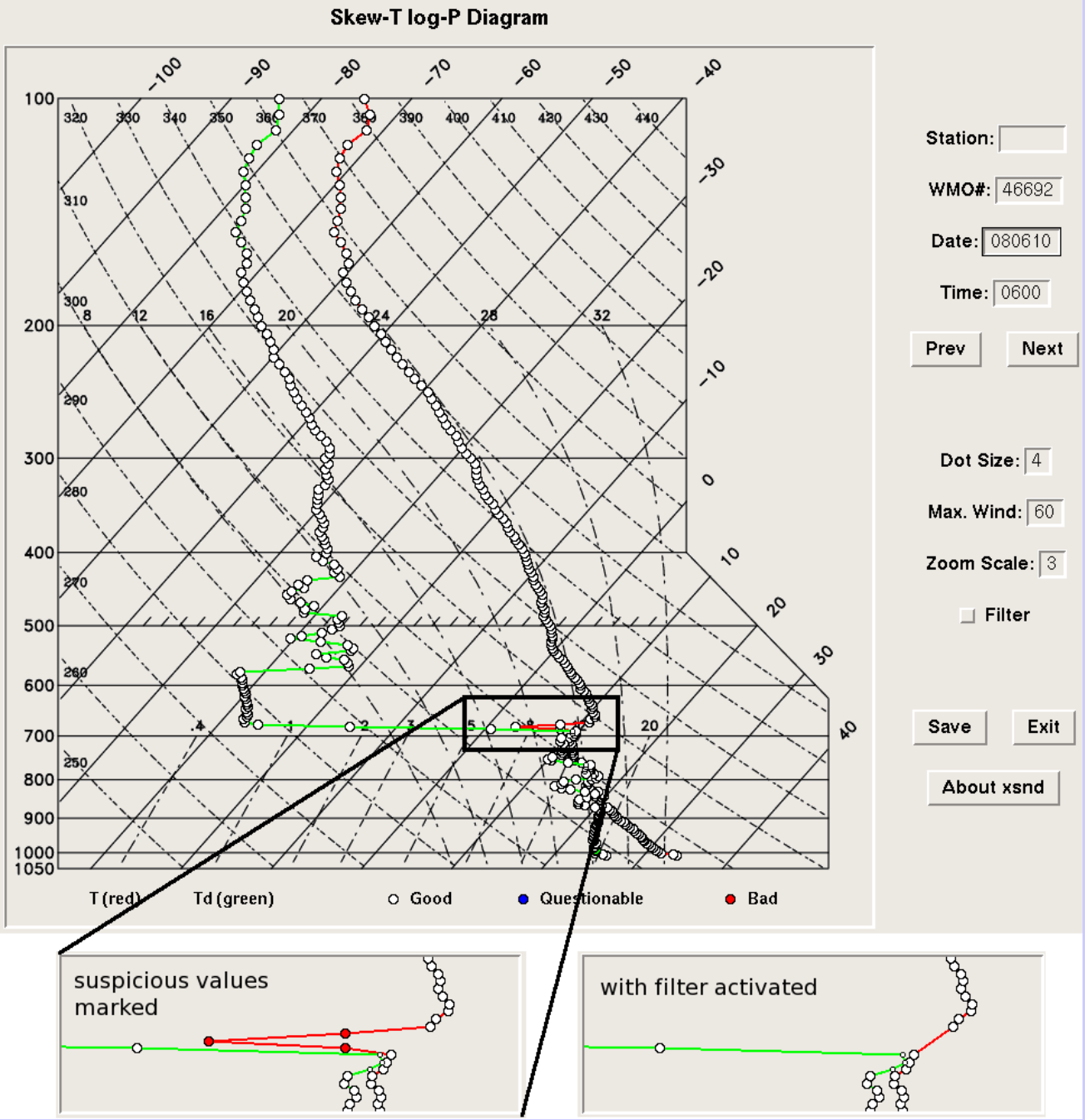
Create a user-friendly dataset (uniform vertical resolution with QC flags), where flags are generated through application of objective QC checks then adjusted by visual inspection of sondes (e.g., using *xsnd* tool). ➠ **Level 4**

Ciesielski et al. 2011 (BAMS)

Visual inspection of sondes, while tedious, is necessary to ensure a research quality dataset since subtle errors in sonde data are often difficult to detect with objective procedures.

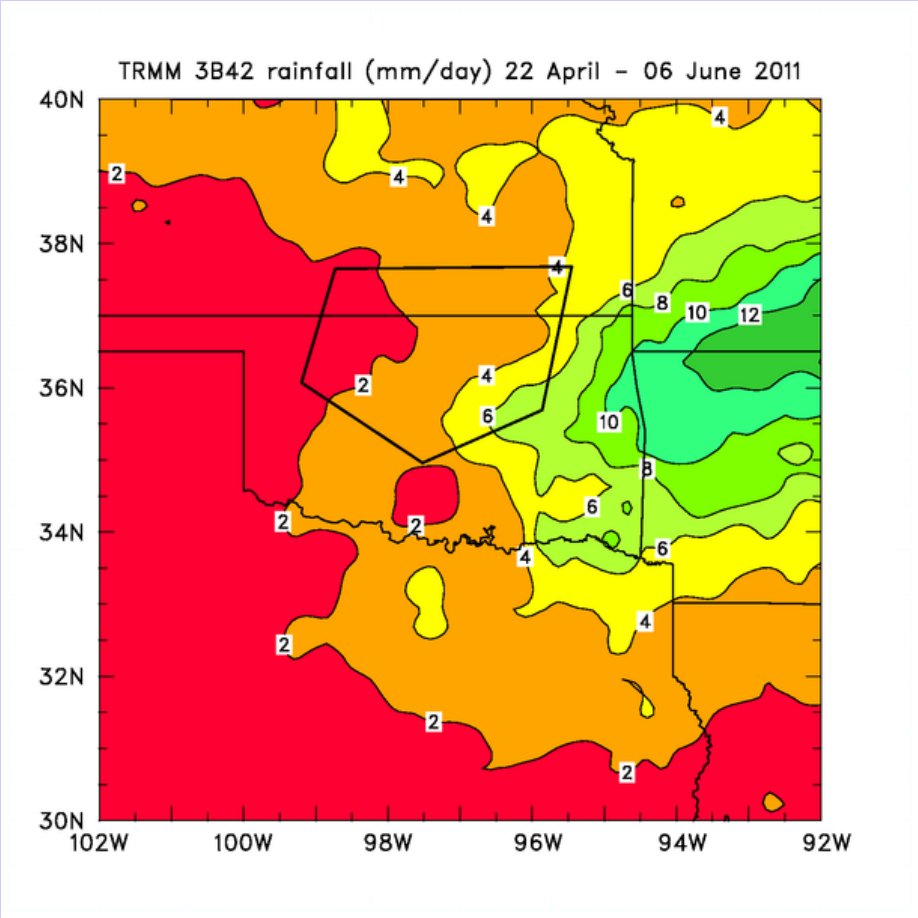
To facilitate this processing, we've developed a software tool (its display is shown below) which allows one to visually inspect a sonde and easily flag suspicious data.





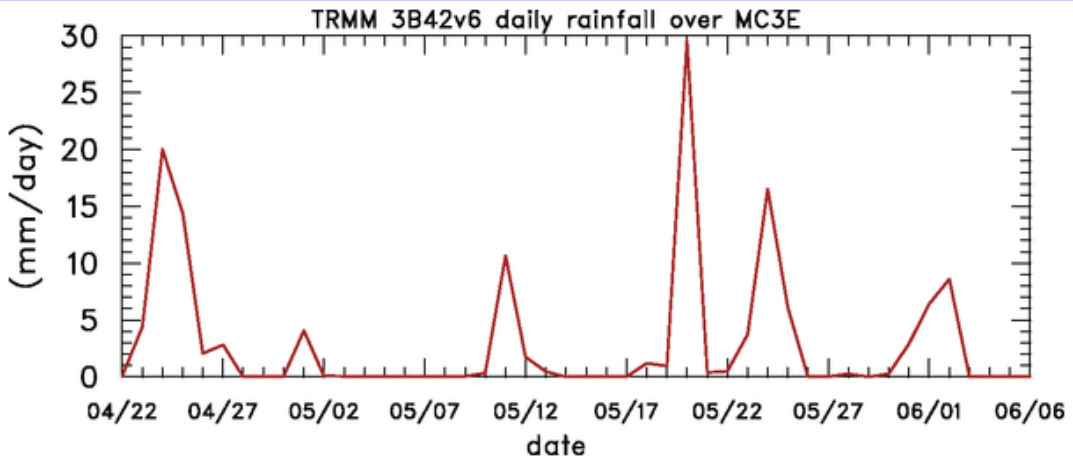
- using this tool to identify and mark suspicious data (a superadiabatic layer above a cloud layer.)
- marking points changes QC flags, not the data values.
- by flagging suspicious data, good data is easily retrievable with the user deciding what level of quality is acceptable for their analysis.

Tool is freely available at: *http://www.eol.ucar.edu/projects/sondeqc*



- compute Q_1 and Q_2 budgets and large-scale forcing fields associated with the convective systems that passed through MC3E network
- rainfall map for 46-day enhanced observing period (mean rate rain over domain was ~3.0 mm/day with heavy rain just east of domain)

- 5-6 convective events occurred during this 46-day period



- Different approaches for computing budgets (reviewed by Zhang et al. 2001)
  - line integral
  - least squares fitting
- Gridded methods
  - successive correction (Cressman 1959)
  - weighted-averaging interpolation (Barnes 1964)
  - statistical interpolation (Ooyama 1987)
  - multiquadric interpolation (Nuss and Titley 1994)
- More advanced approaches – constrained variational analysis (Zhang and Lin 1997)
- Techniques yield different results for non-regular, data-sparse fields

In traditional methods, terms on RHS are evaluated (e.g., finite differences using gridded data) and P_o and $<Q_R>$ are computed as residuals.

Heat and Moisture Budgets

Apparent heat source/moisture sink:

(Yanai et al. 1973)

$$Q_1 \equiv c_p[(\partial\bar{T}/\partial t + \bar{\mathbf{v}}\cdot\nabla\bar{T} + (p/p_0)^\kappa\bar{\omega}\partial\bar{\theta}/\partial p)]$$

$$Q_2 \equiv -L(\partial\bar{q}/\partial t + \bar{\mathbf{v}}\cdot\nabla\bar{q} + \bar{\omega}\partial\bar{q}/\partial p)$$

Integrated budgets:

[Definition: $< \quad > \equiv 1/g \int_{p_T}^{p_s} (\ )\, dp$]

$$\text{Heat}: \quad < Q_1 > \; = \; < Q_R > + LP + S \qquad (1)$$

$$\text{Moisture}: \quad < Q_2 > \; = \; L(P - E) \qquad (2)$$

(1) − (2) yields

$$< Q_1 > - < Q_2 > \; = \; < Q_R > + S + LE\ , \qquad (3)$$

where P = Precipitation rate, E = Evaporation rate, S = Sensible heat flux

These are powerful technique but computational details are crucial

In the variational approach, P_o and $<Q_R>$ are used, along with other analyses, to help constrain the budgets.

Mass balancing the divergence:

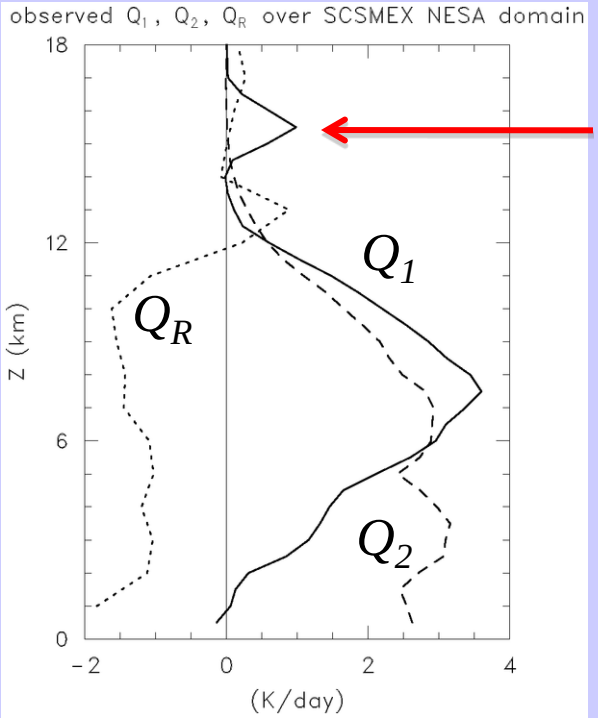
Upper boundary condition (locate tropopause)

$\omega = 0$ at tropopause

$d\theta/dt = 0$ at tropopause

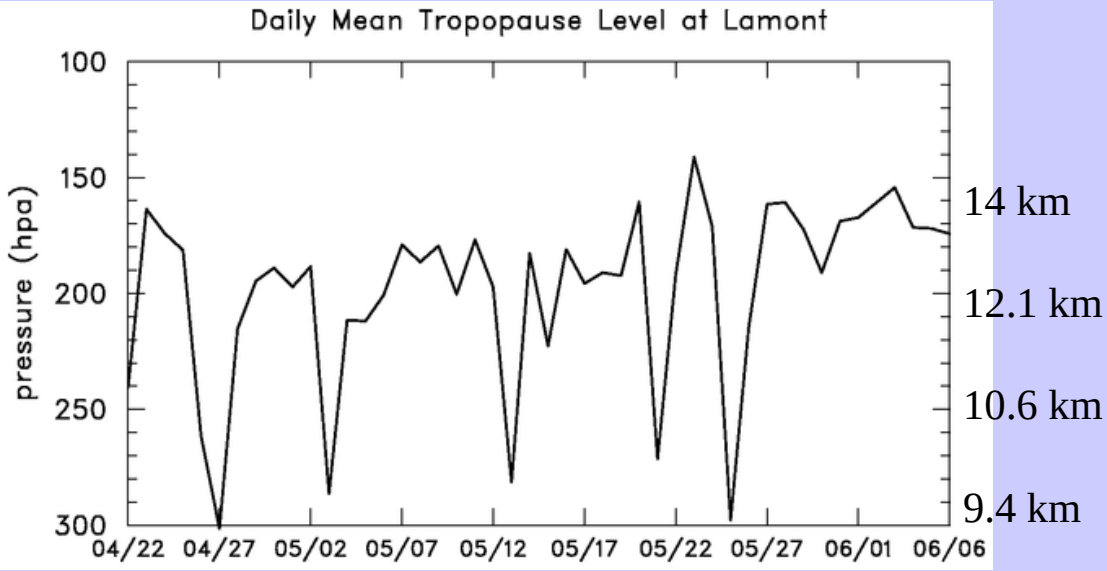
$d\theta/dt = Q_R$ at tropopause (best, Can ARM provide vertical profiles of Q_R ?)

High vertical resolution needed for $\partial\theta/\partial p$



In SCSMEX analysis, 25-hPa resolution data was inadequate to resolve tropopause level. We suspect that small vertical motions coupled with high $\partial\theta/\partial p$ near the trop resulting in a spurious upper-level heating peak.

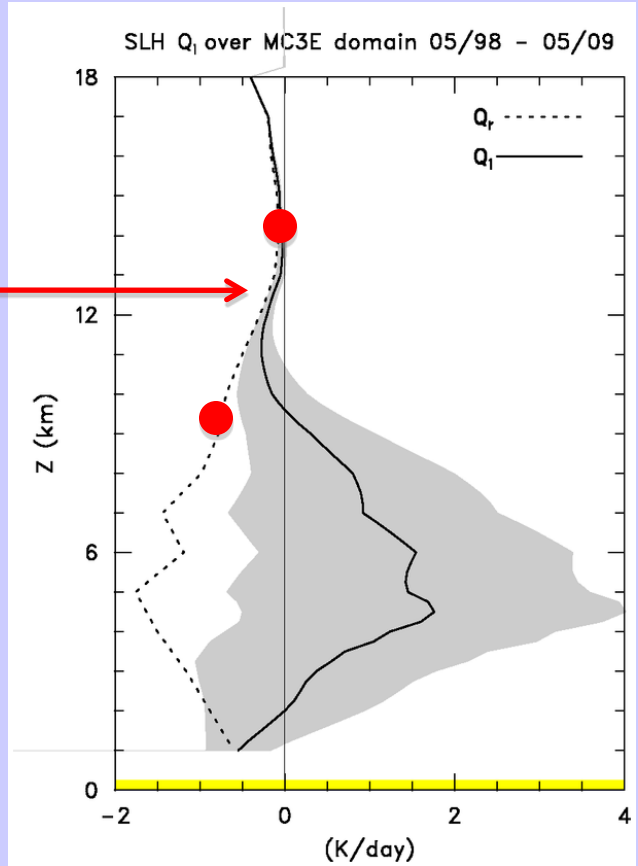
Note: in this case mean Q_R near trop (17 km) is small.

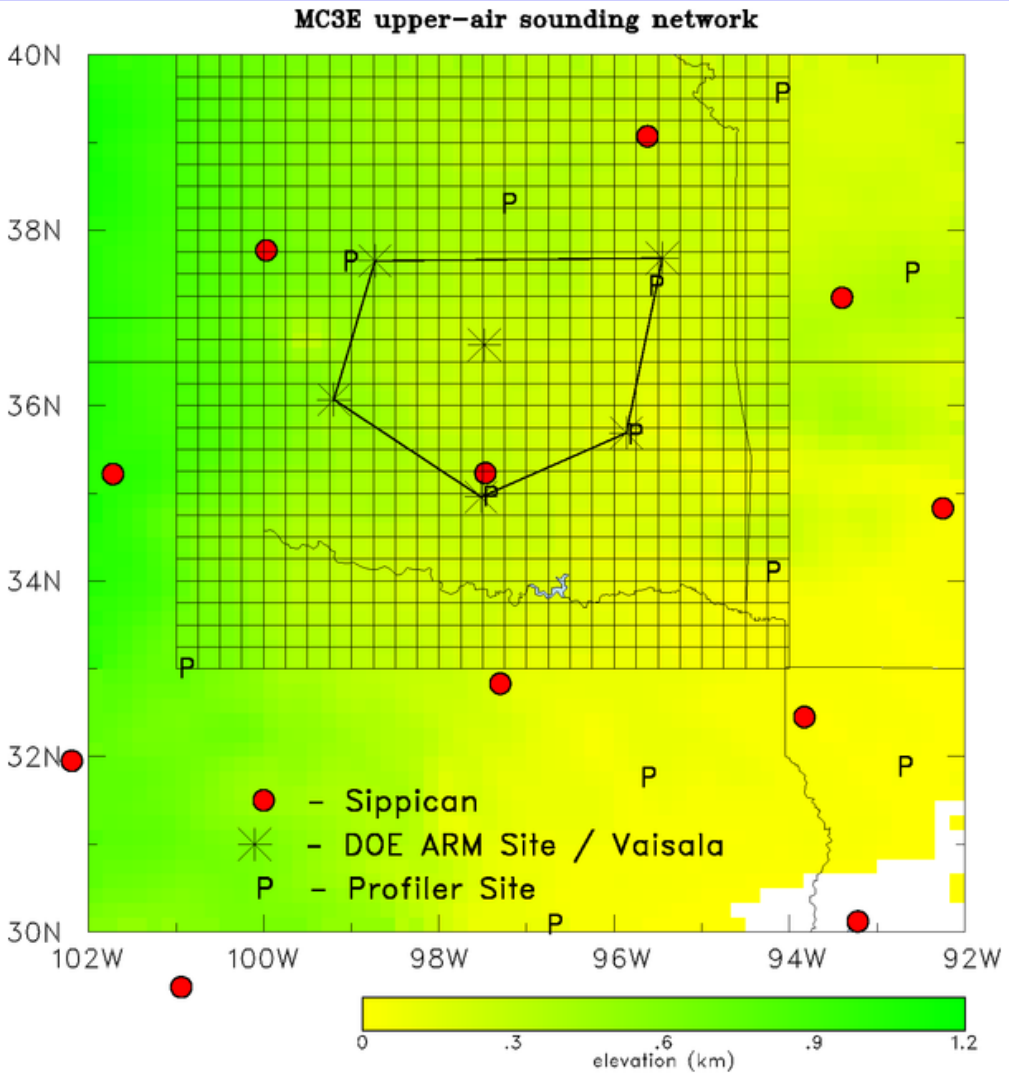


Trop height shows large variability during MC3E falling rapidly after passage of convective events (range from 9.5 to 14 km).

Mean Q_R over this layer varies from near 0 K/day at 14 km to near -0.8 K/day at 9.5 km.

In other words, correct mass balancing requires accurate identification of trop level and knowledge of Q_R at this level. Can ARM provide this Q_R information?





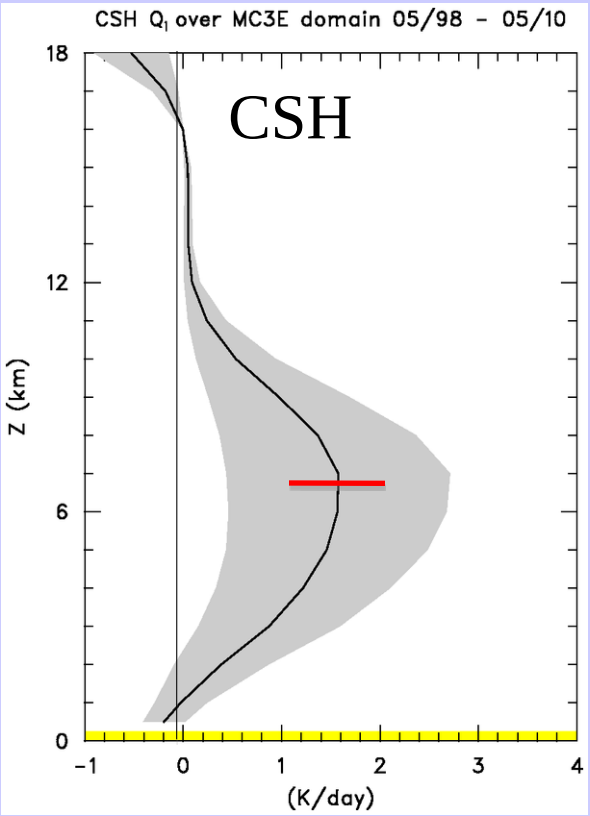
Our approach: use multiquadric interpolation with sonde and profiler data to produce a gridded dataset of basic fields with 0.25° horizontal resolution, 5 hPa in vertical and every 3 hours.

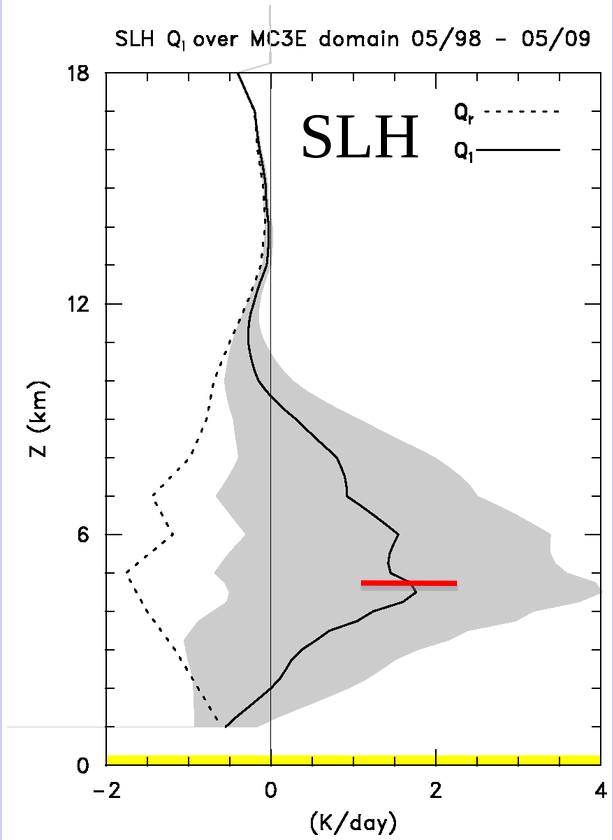
Balloon drift should be taken into account since drift distances are a non-negligible fraction of station spacing.

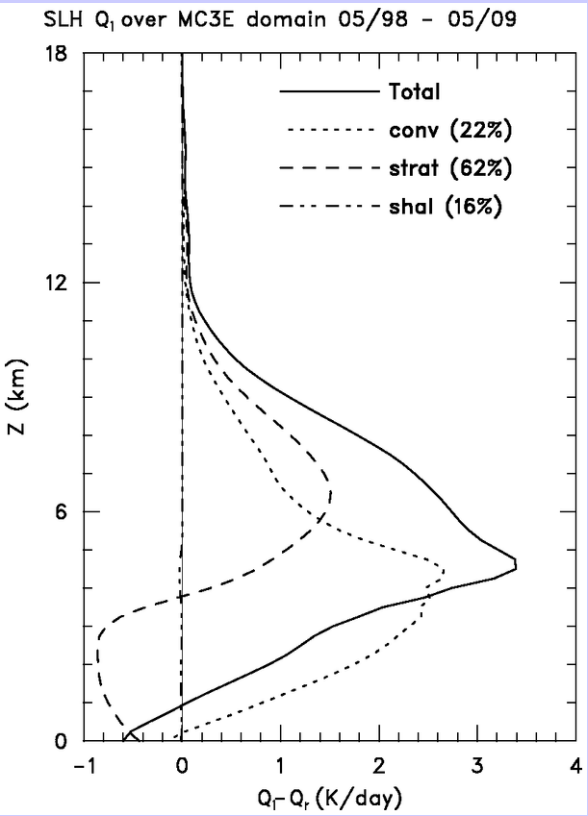
Compute divergence with gridded dataset, mass balance divergence (hopefully with diabatic upper-boundary condition) to compute vertical motion, then compute budget terms and average results over MC3E domain.

Use surface latent and sensible heating fluxes estimates to close budgets and compute P_o and $\langle Q_R \rangle$ as residuals.

Multi-year May-mean Q_1 profiles for different V6 TRMM estimates







- Magnitude of peaks are similar (~1.5K/day) but level of peaks are quite different.
- shading show standard deviation of estimates.
- MC3E results will help us evaluate this TRMM heating products.

Summary

creating a high-quality LSF dataset

- Produce a quality-controlled sounding dataset by reducing humidity biases in sondes and visually inspecting all sondes, particularly surrounding sites that use Sippican systems.
- Numerous tools and suggestions for QCing sonde datasets can be found at: http://www.eol.ucar.edu/projects/sondeqc
- Use balloon drift information in computing basic fields.
- Use accurate an accurate upper-level boundary condition for vertical motion which involves precisely locating tropopause and using Q_R data at trop level.

Computing budgets and LSF fields for MC3E with both the traditional and variational methods, using the same input data, will provide us an opportunity to compare their analyses and better understand their differences and limitations.