ARM Data Epochs

JENNIFER COMSTOCK[1], GIRI PRAKASH[2], KEN KEHOE[3]

1- Pacific Northwest National Laboratory
2- Oak Ridge National Laboratory
3- University of Oklahoma

ARM/ASR PI Meeting, June 2019

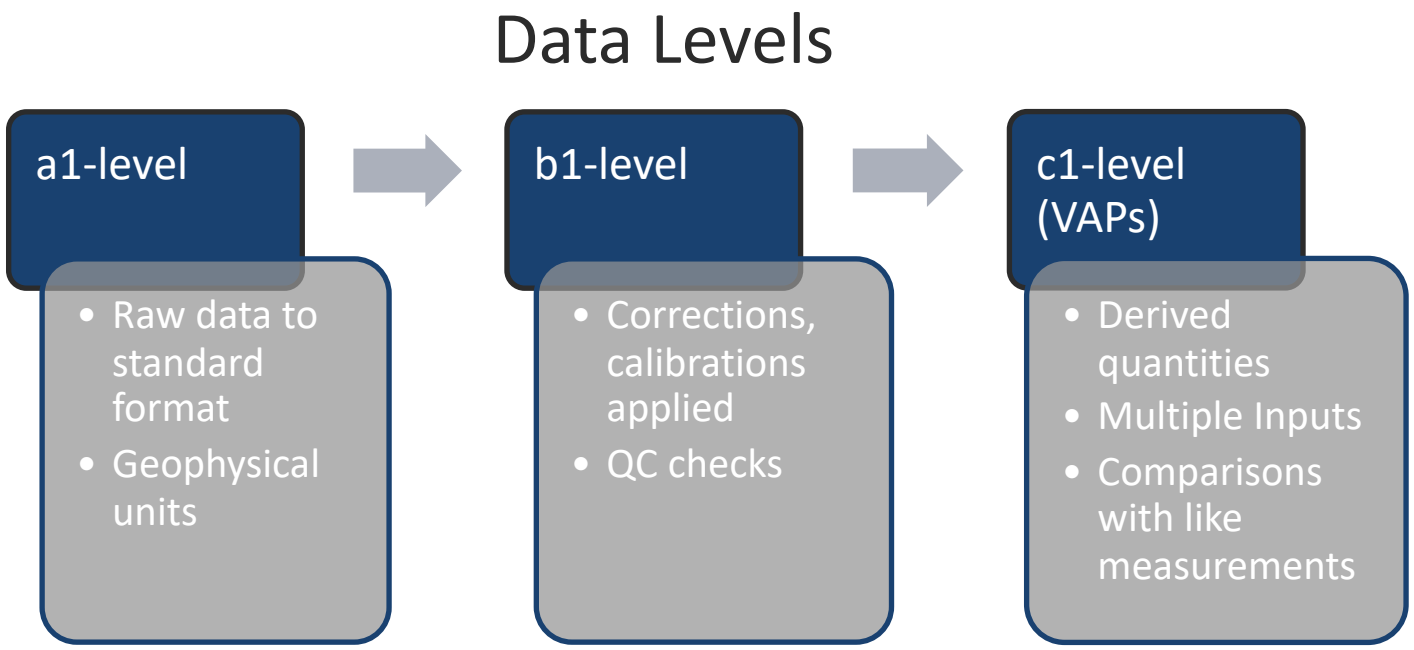
Motivation

- ARM processes over a thousand datastreams each year including raw, processed (ingests), and Value Added Products (VAPs), external
  - Example: Aerosol Optical Depth
- Difficult for users to identify 'best' data source for their science focus
- Several user groups and workshops have recommended that ARM identify 'good' datasets and time periods
  - Radar Science and Engineering
  - ARM-ACME-ASR Workshop
  - ARM User Executive Committee (UEC)

Goal: Make it easier for users to find high quality data of interest

Data Quality – What ARM is doing now

- Data Quality Office (DQO)
  - Data Quality Assessment on a1-level ingests – visual inspection
  - Submit Data Quality Problem Reports
    - DQPRs reviewed
    - Issues are described to users with data order
- b1-level data
- 'QC' VAPs
  - QCRAD, QCECOR, QCAOD*NEW*
  - *QME – Quality Measurement Experiments*
- Recommended datastreams



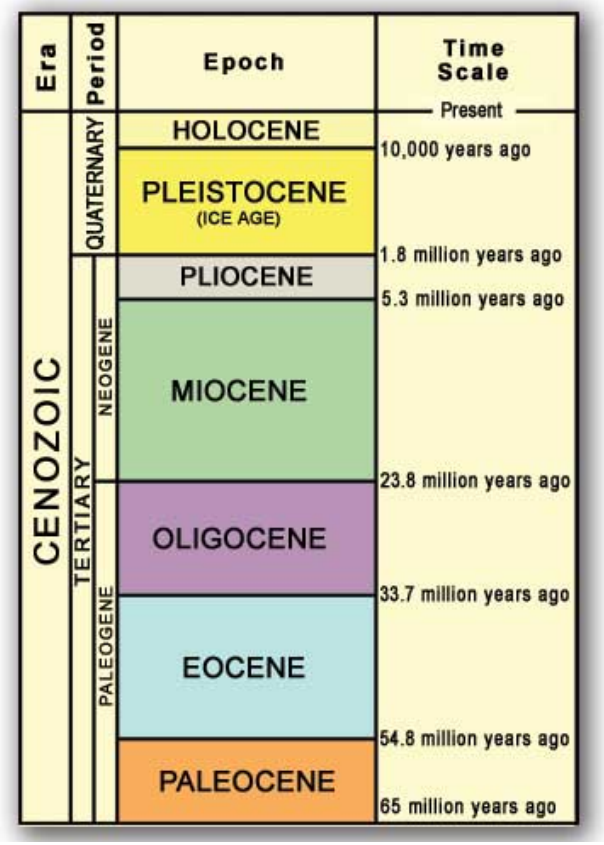
What is a Data Epoch?

- *Epoch:* A distinct period of time characterized by particular attribute (i.e. geologic time periods)

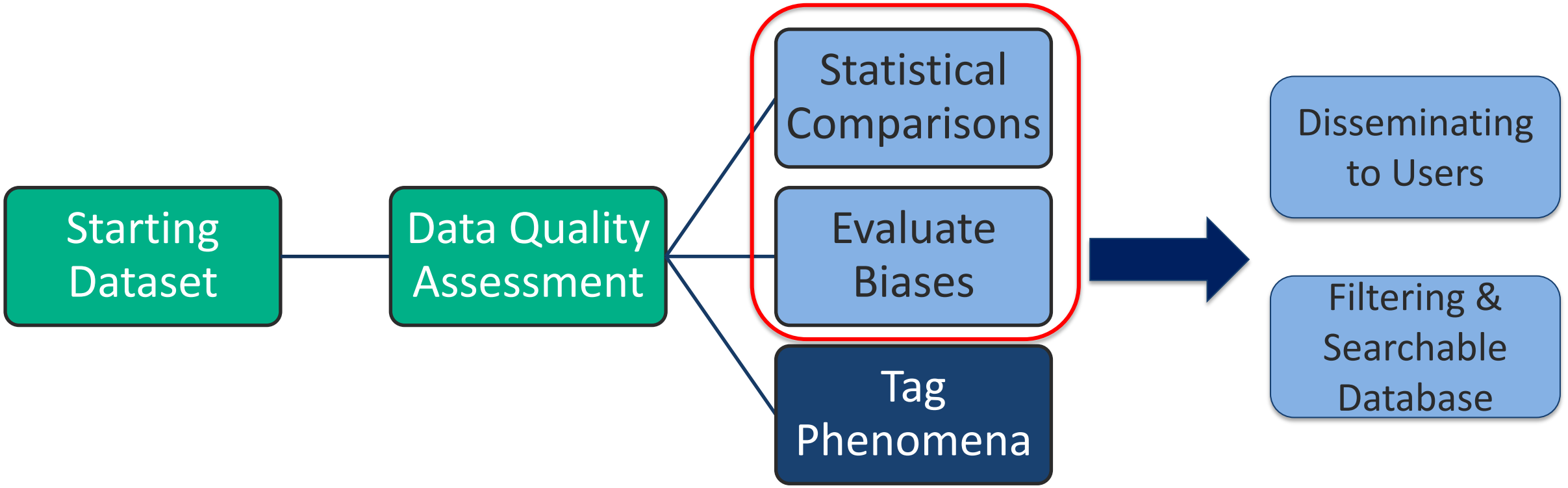
- *Data Epoch*: A specific time period for a particular measurement:
  - With well-characterized data quality
  - Scientifically interesting
  - Quantified uncertainties

- *Suggestions*:
  - Bounded by calibration period or instrument changes
  - Focus around recent field campaigns
  - Time periods with consistent measurements (such as consistent noise floor)

Geologic Time Periods



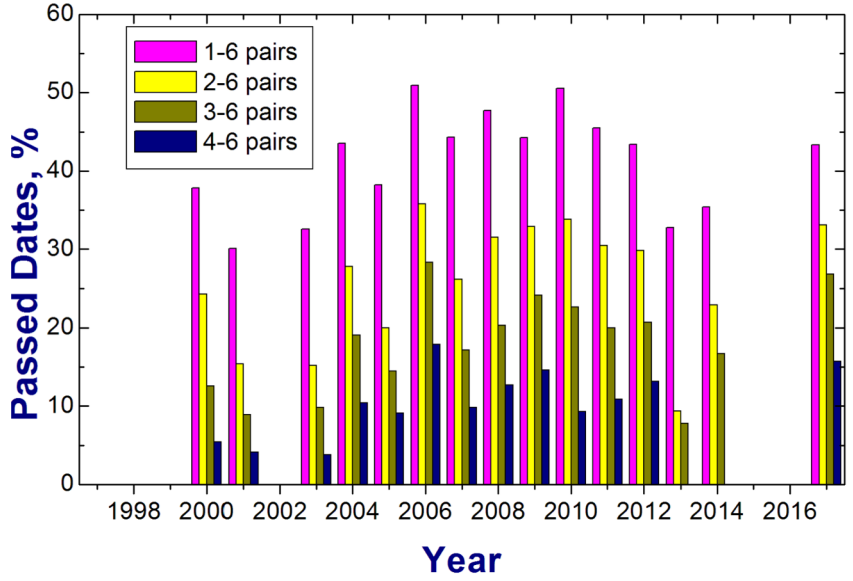
Process for Developing Data Epochs



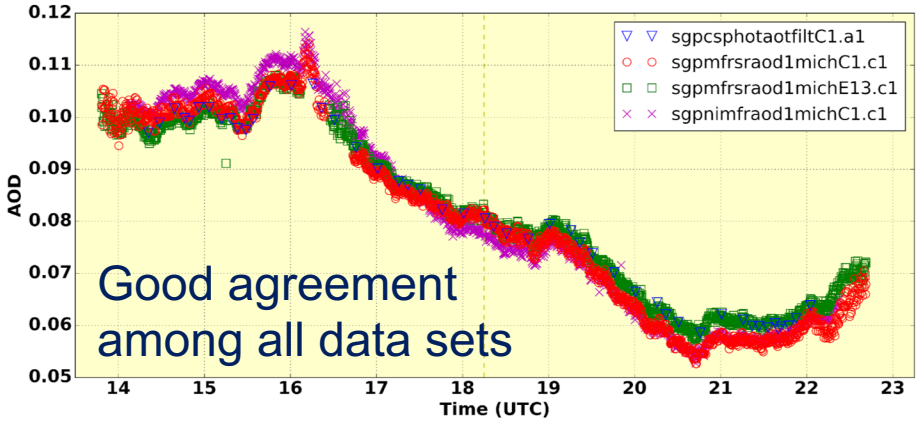
Example: Aerosol Optical Depth
QC-AOD and AOD Best Estimate at SGP

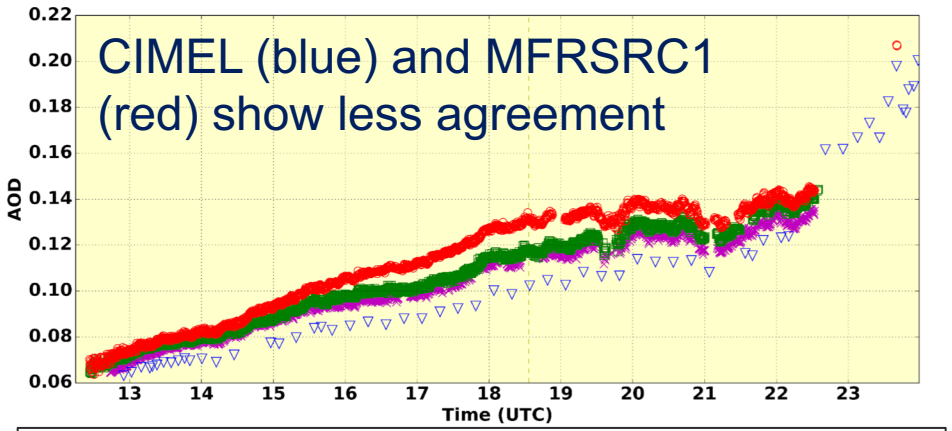
- 20 yr data set (1997-2017) at SGP
- Characterizing quality and estimating uncertainty
- Comparing 4 instruments: MFRSR-C1, MFRSR-E13, NIMFR, CIMEL
- Statistical comparison tests: slope, bias, R2 criteria

Measurement Agreement



Two dates that passed all daily statistical comparison tests:

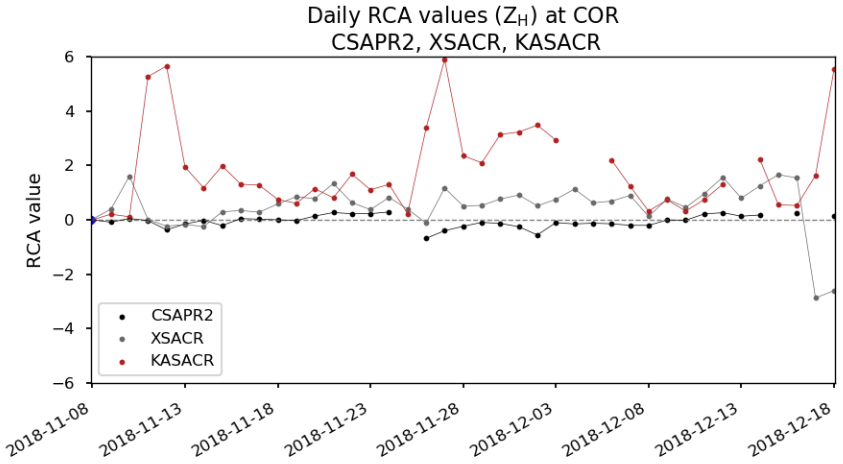


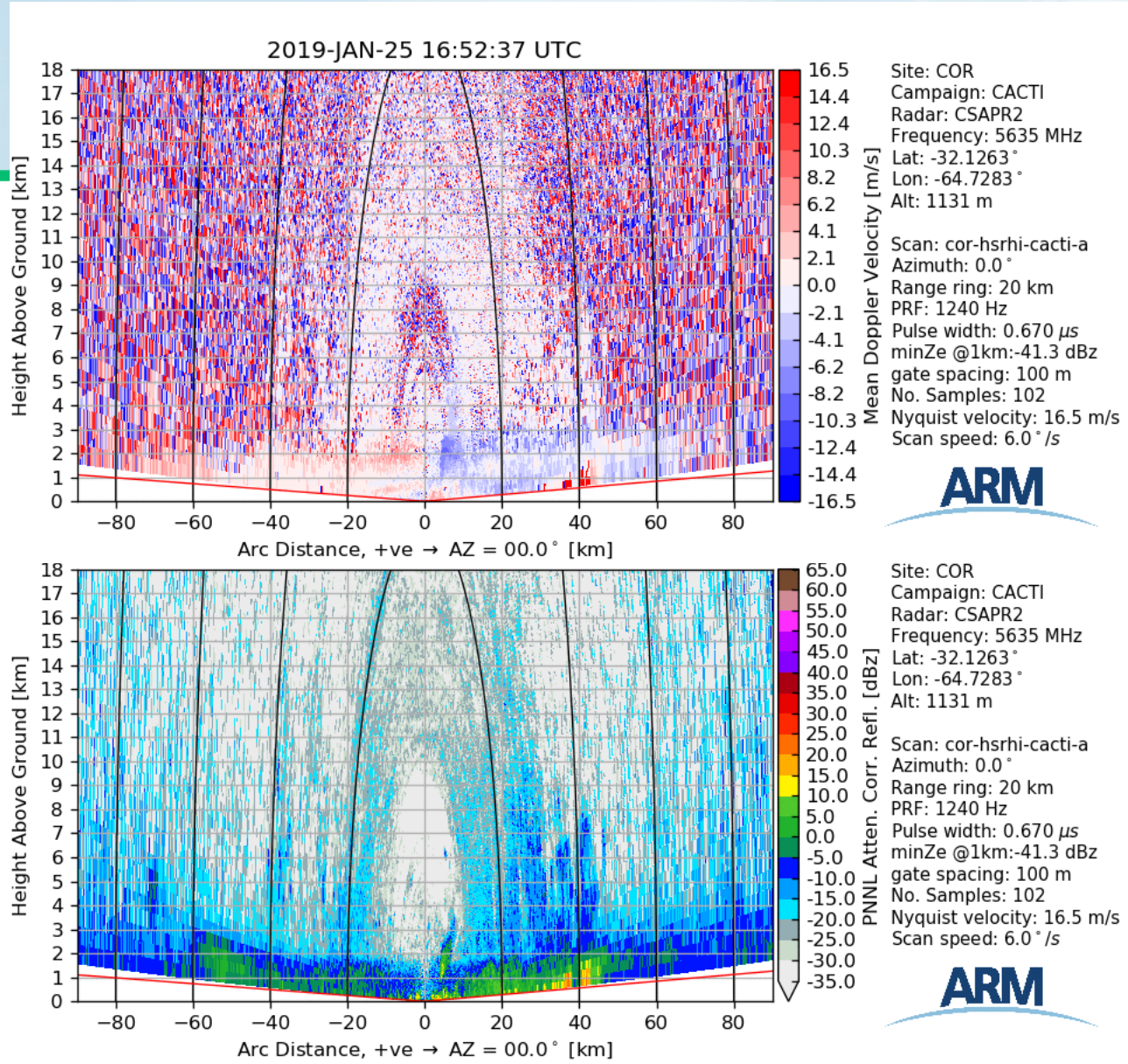


Work by Laura Riihimaki, Evgueni Kassianov, Justin Monroe, and Connor Flynn

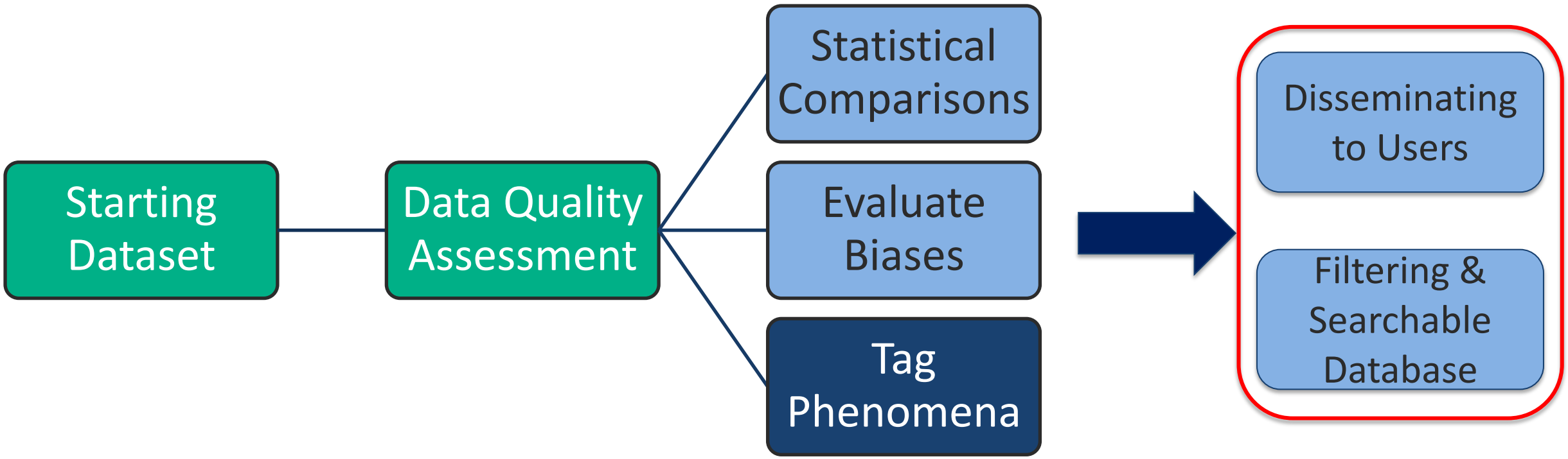
Example: Radar Products from CACTI

- Developing data epoch for CACTI
- b-level products
  - Quality control masks
  - Calibrated moments, KDP
- Relative Calibration Adjustment
  - Run real-time during CACTI
  - Daily calibration log

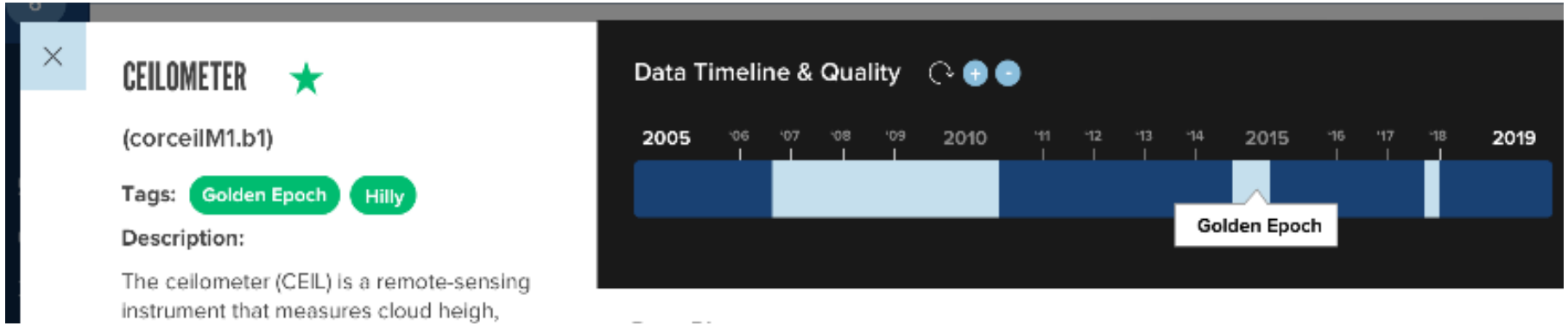




Process for Developing Data Epochs



Discovering Data Epochs Through Data Ordering



The new Data Discovery highlights epoch data using the timeline, data highlights, and tags.

Tags: Make it easier to search for specific epochs or phenomena

Timeline: Highlighting epochs visually

Data Highlights: Recent data epochs will be listed on the home page

Documentation: Describes methodology for developing individual epochs

Giri Prakash, Ranjeet Devarakonda

Today: Working Lunch – Data Discovery Updates

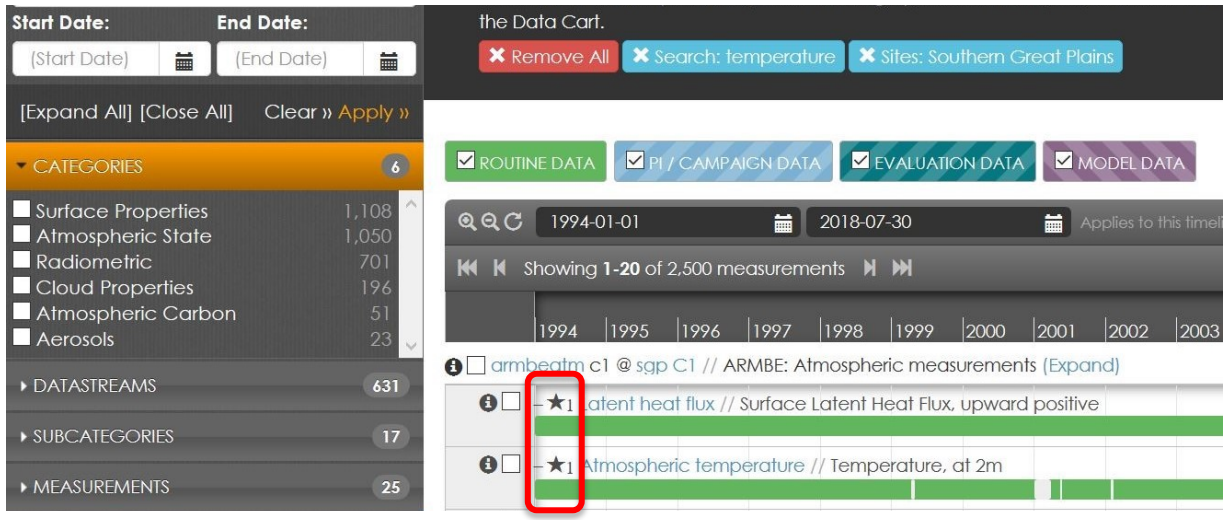
Updating Recommended Datastreams

- 2017 ARM Triennial Review identified need for assisting users to find 'best' source of ARM measurements
- Original list identified in 2012
- Updating as part of the Data Discovery upgrades
- Working with Translators to make new assignments

Future Plans

- Enhance display of recommendations in Data Discovery
  - Use recommendations for choosing what datastreams are initially displayed in Data Discovery
  - Display recommendation rationale and datastream characteristics
- Data Epochs will appear as a source of recommended datastreams

Recommended Datastreams Displayed on Data Discovery Tool

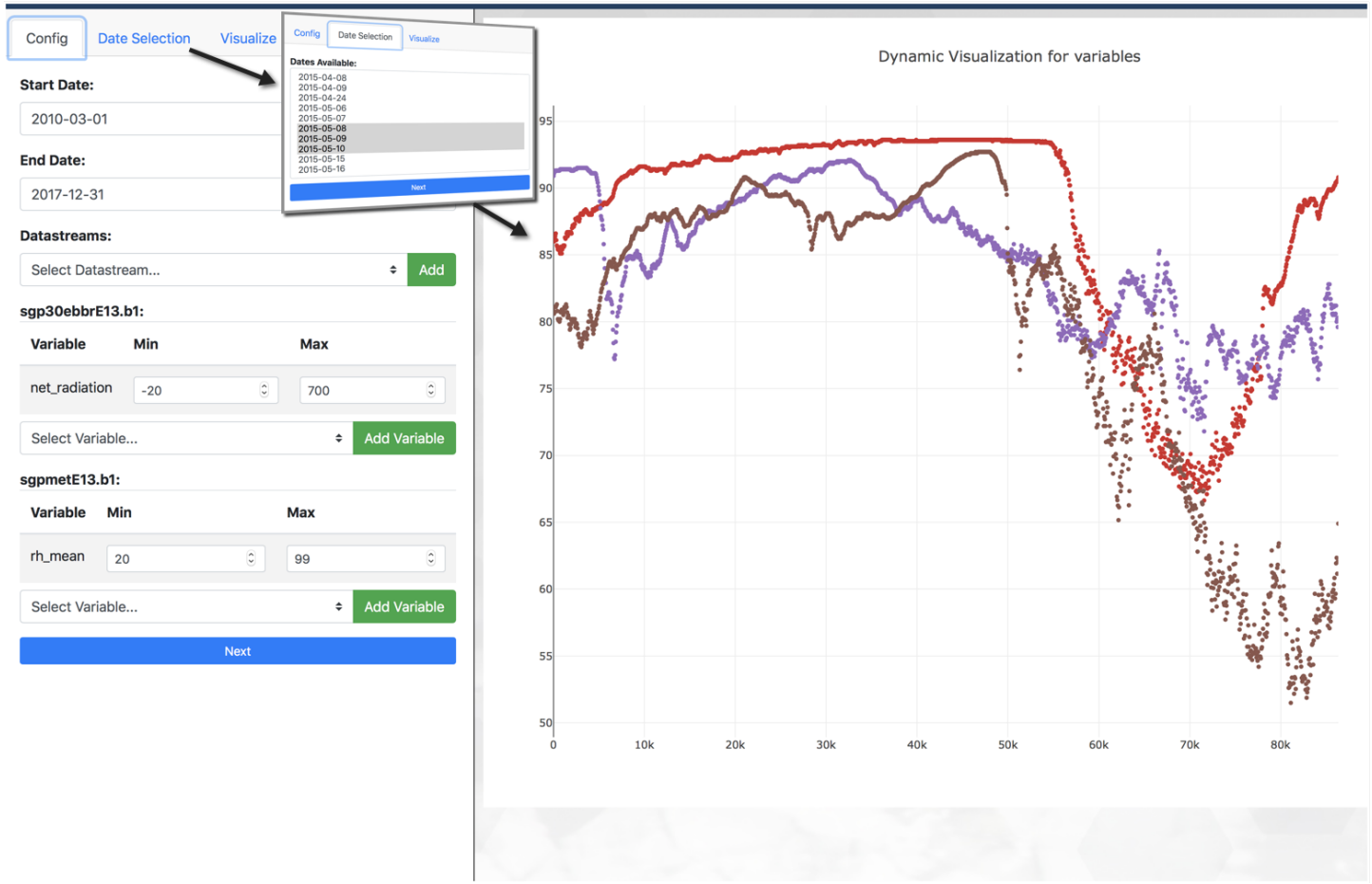


Ric Cederwall
Poster session: A2 Tues 5:00 - 6:30 pm #32

Identifying Data Epochs Through Data Search and Filtering

NoSQL Data Analytics Platform

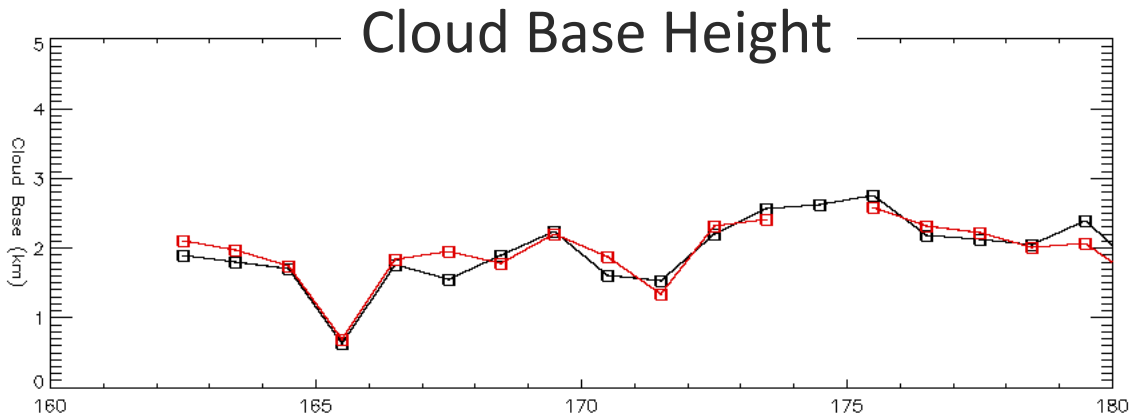
- Allows users to search and filter data (i.e. bundle browser) based on conditional queries
- Tag specific phenomena
- Capture metadata
- Couple with Data Consolidator and Data Discovery Tool for merging and ordering data

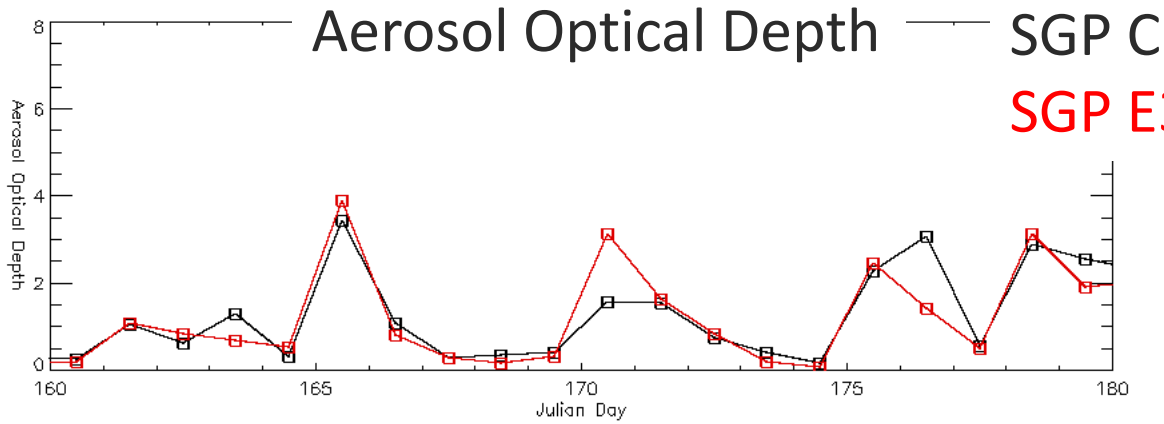


Tool developed by Bhargavi Krishna

Beyond the Data Epoch: Virtual Field Campaign

- Virtual Field Campaign
  - Observation driven modeling study centered around focused scientific objectives
  - Multiple measurement sources
    - Ground, satellite, in situ measurements
    - Data assimilation, large-scale forcing
- Example: HI-SCALE
  - Land-atmosphere and aerosol cloud interactions of shallow cumulus clouds
  - Compile measurements from SGP central facility and 4 boundary layer profiling sites
  - ARM Data Consolidator merges multiple datastreams into single ARM standard netcdf
  - In situ observations – G-1 aircraft





Krista Gaustad and Chitra Sivaraman

Data Consolidator Breakout Session – This afternoon

Currently Available Data Products and In Progress Epochs

Available b1 and c1 level products

b-level

- AOS Harmonization Products
  - Optical properties, CCN, trace gas
- Doppler Lidar Products

c-level

- AOP (aerosol optical properties)
- AERIoe (LASSO cases)
- QCRAD, QCAOD, QCECOR
- RLPROF (FEX, Mixing Ratio, Temperature)
- PARSQUANTS (disdrometer)
- ARMBE – Modeling Best Estimate
- ARSCL

Under Development or in Release

- b-level radar products – CACTI
- XSAPR data from SGP
- KAZR/MMCR ARSCL corrections using CloudSat
- MWR-RET v2 – auto bias correction for LWP
- MPLCMASK – improved corrections (β, δ)