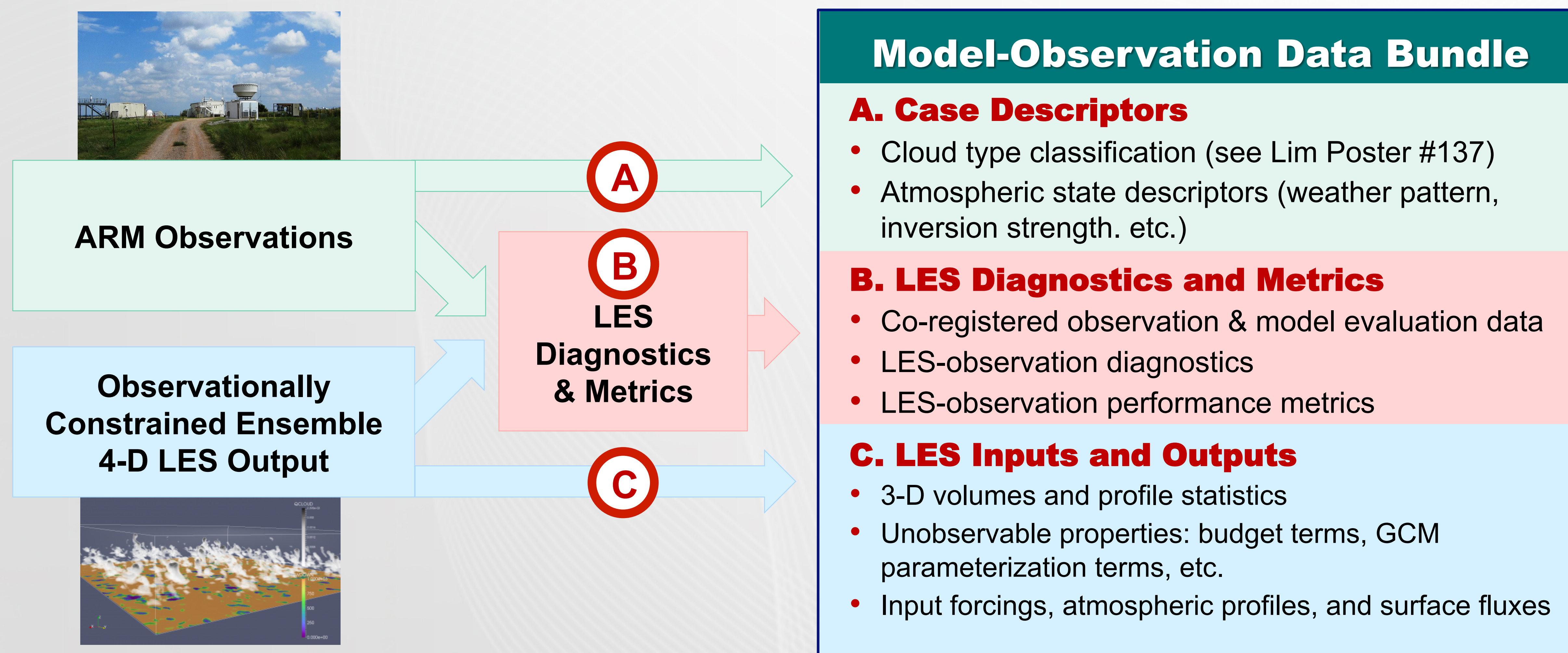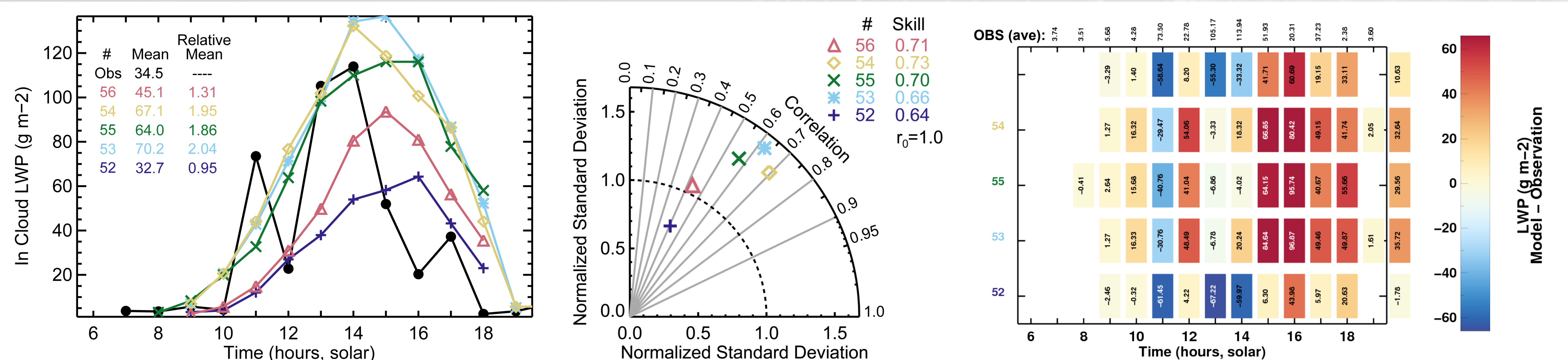# LES ARM Symbiotic Simulation and Observation (LASSO) Workflow: Model-Observation "Data Bundles"

Andrew M Vogelmann[1], William I Gustafson Jr[2], Tami Toto[1], Satoshi Endo[1], Heng Xiao[2], Zhijin Li[3,4], and Xiaoping Cheng[3]

[1]Brookhaven National Laboratory, [2]Pacific Northwest National Laboratory, [3]University of California, Los Angeles, [4]NASA Jet Propulsion Laboratory

ARM
CLIMATE RESEARCH FACILITY

ARM Observations

Observationally Constrained Ensemble 4-D LES Output

(A)
(B)
(C)

LES Diagnostics & Metrics

Model-Observation Data Bundle

A. Case Descriptors
- Cloud type classification (see Lim Poster #137)
- Atmospheric state descriptors (weather pattern, inversion strength. etc.)

B. LES Diagnostics and Metrics
- Co-registered observation & model evaluation data
- LES-observation diagnostics
- LES-observation performance metrics

C. LES Inputs and Outputs
- 3-D volumes and profile statistics
- Unobservable properties: budget terms, GCM parameterization terms, etc.
- Input forcings, atmospheric profiles, and surface fluxes

Summary
- A "data bundle" is a unified package consisting of LASSO model output, observations, and model metrics aimed at providing the best description of the atmosphere for use by the research community.
- This poster highlights key features of the data bundles.
- **Suggestions are welcome!**

+ Ease of Use

Level 3 — Searchable Metrics, Quick-Look Plots

Level 2 — Evaluation Data, Statistical Summaries

Level 1 — 4-D Simulated Output, Forcing Dataset Inputs

1. Data bundles are searchable and include quick-look plots and filtering methods for users to find and order cases of interest. Where applicable, instrument simulators are used to provide model-observation comparable fields for 1:1 comparisons in the diagnostics and metrics (e.g., radar/lidar, cloud cover, etc.).

2. Three levels of data are provided in each bundle. Level 3 is the most processed and easiest to use, and lower levels provide more detailed information for the sophisticated user.

3. Evaluation diagnostics based on ARM observations of environmental and cloud properties are used to assess the LES simulations including: Time series, Regressions, Taylor diagrams, Heat maps, and Phase space relationships.
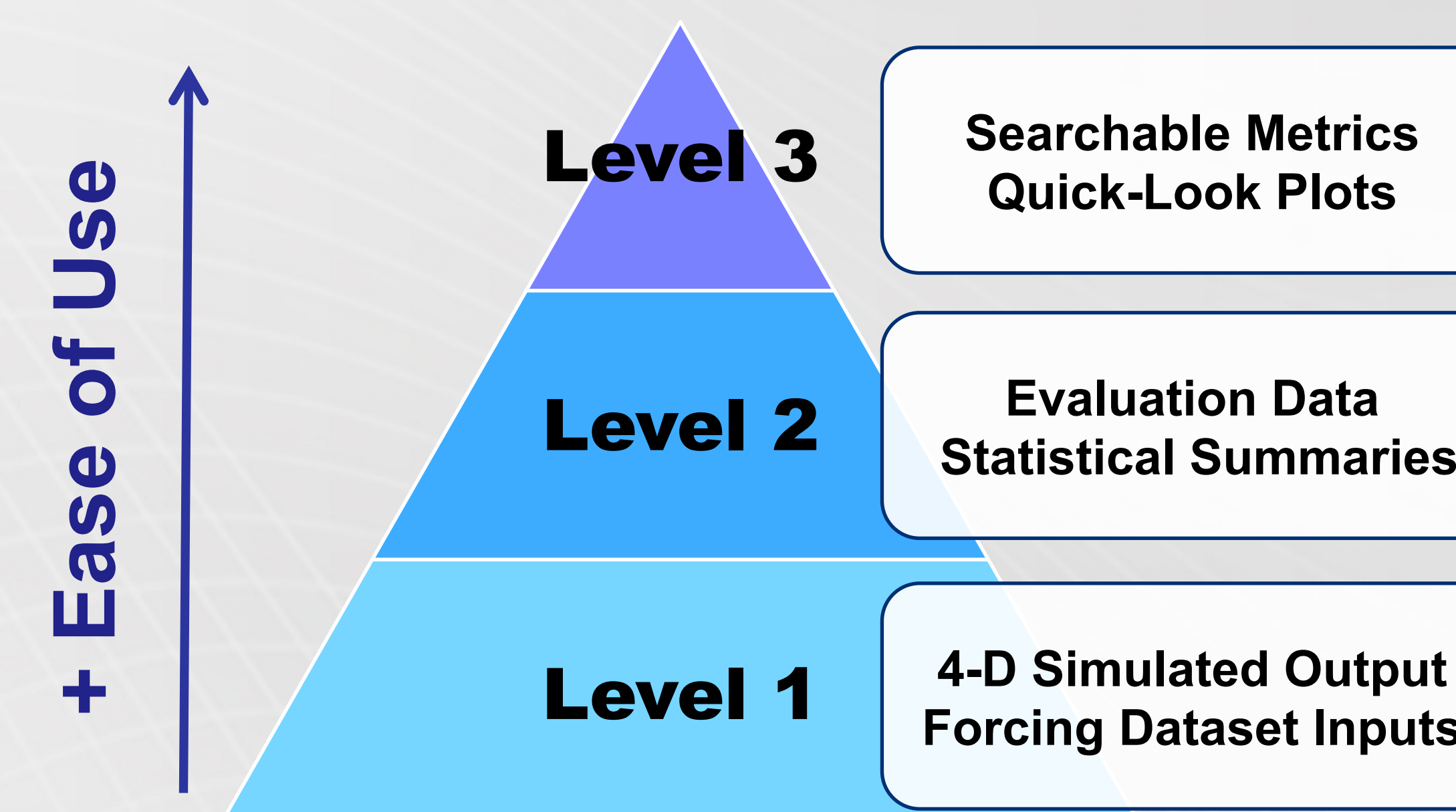
4. Metrics quantify simulation quality using the 2-D cloud mask, LWP, and TSI cloud fraction (CF). (Above) The skill for the simulated 2-D mask is quantified using the Equitable Threat Score (ETS) and bias. (Below) The skill for LWP and cloud fraction is based on the Taylor skill and skill of the relative mean. A single, combined skill score is computed for each of the 3 variables (2-D cloud cover, LWP, and CF). The LWP and CF skill scores are combined into a single "bulk" score that is used with the combined 2-D skill score to find the best simulations.

5. LES Outputs and Inputs are included in the data bundle with observations co-registered on the model grid. **Outputs** include 3-D fields of cloud and environmental variables (~20 min) and more frequent profile statistics, budget terms, and terms used in common GCM parameterizations. **Inputs** include the forcings, atmospheric profiles, and surface fluxes for modelers to perform sensitivity studies (for "good" forcings) using their own model settings.
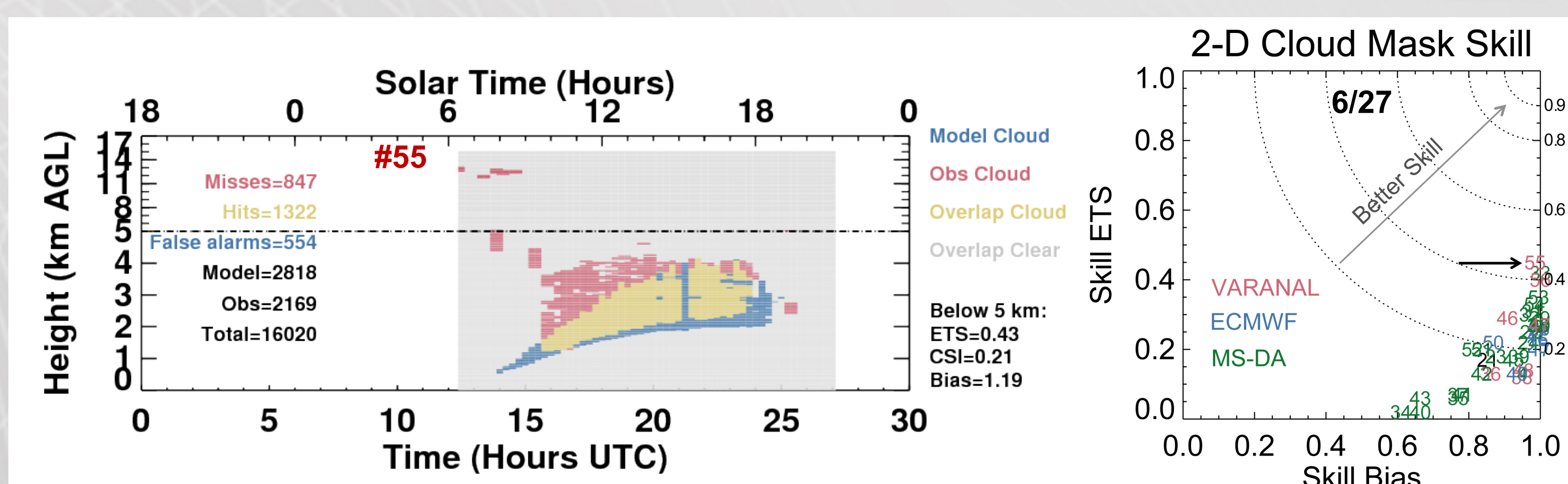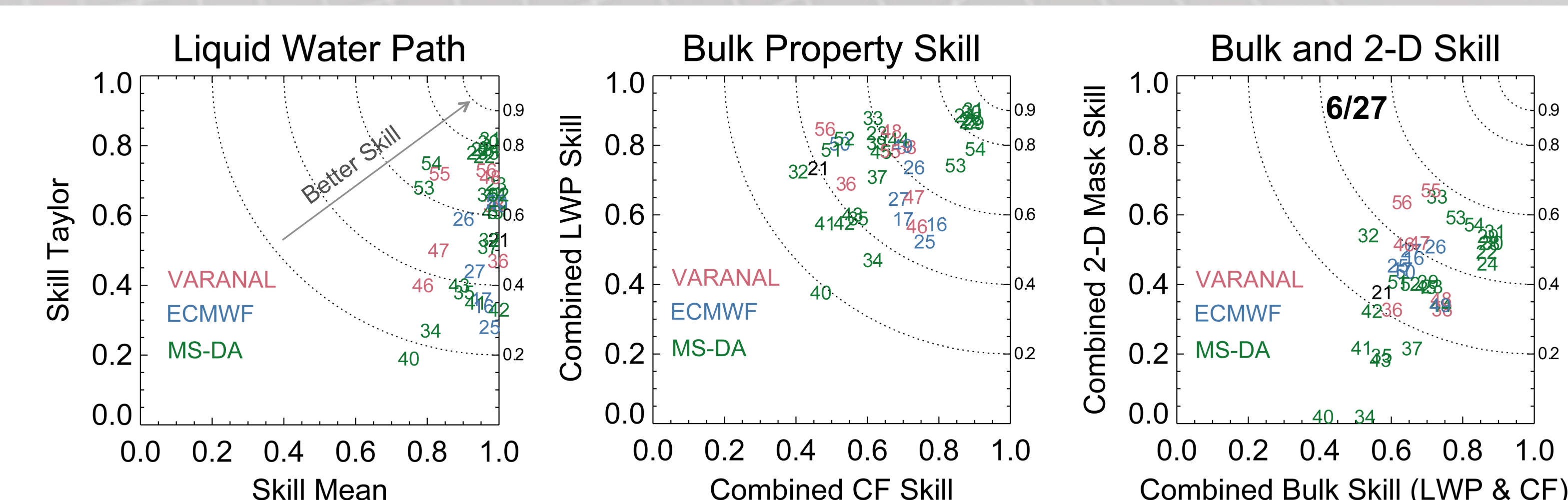
6. Bundle search and order is planned using 3 methods:
1) Easy-to-use web search and order using pre-computed diagnostics, metrics & quick-look plots,
2) Interactive processing and visualization for flexibility and cross-case comparisons (see Krishna poster #138 for a noSQL study)
3) Remote processing from a user's locale on an ARM system before moving the result to their system.