National Aeronautics and Space Administration

FY24 Strategic Initiatives Research and Technology Development (SRTD)

# Foundations of Data Fusion for Future PBL Missions

**Principal Investigator:** Amy Braverman (398)**; Co-Investigators:** Margaret Johnson (398), Marcin Kurowski (329), Matthew Lebsock (329)

**Strategic Focus Area:** Next Earth Science Decadal Survey: Technology & Architecture for Planetary Boundary Layer (PBL)/ Surface Topography & Vegetation **| Strategic Initiative Leader:** Rashmi Shah

## Objectives

1. Develop, implement, and test prototype algorithms for fusing data from three notional future JPL PBL missions identified in the PBL Report (a sounder, a microwave, and a radar/lidar modeled on the DIAL and DAR instrument concepts).

2. Demonstrate that we can evaluate accuracies, uncertainties, and computational costs of fused data products that will result from different design choices.

## Background

1. One the most important conclusions of a recent NASA study on future PBL observation is that no single instrument can observe the PBL adequately, and that *"...the production suite of a spaceborne PBL mission...should combine each component's information content in and optimal manner…"*

2. Statistical methods for combining such heterogeneous data are based on spatial statistical models.

3. These methods are purely data driven. A physical model can be used incorporated to drive temporal dynamics (this is data assimilation), but that is not our objective here.

4. In this project, we built a spatial statistical model to quantify the relationship between the (three-dimensional) PBL spatial field and notional observations of it by three instruments: a sounder-like instrument, a microwave instrument, and a radar-like instrument flying down the center of the sounder/microwave swath.

5. The PBL variable studied here is relative humidity.

## Significance

This is crucial data processing technology for future PBL missions. Instead of performing trade studies for individual instruments in formulation, data fusion allows us to do so for combinations of instruments that provide complementary data. These results provide proof-of-concept for methods to simulate basic observing characteristics, perform data fusion, and evaluate results quantitatively.

## Approach and Results

1. Basic idea: We postulated a probabilistic model that describes the behavior of a full, 3D water vapor field. Then we added a probabilistic specification of how observed data relate to the full field, and finally used basic rules of probability to obtain the conditional probability distribution of the true-but-not-directly-observed field, given the observed data.

2. Data: We used data from the Regional Atmospheric Modeling System (RAMS; Cotton, W. R., and co-authors, 2003) for December 30, 2016 over the Democratic Republic of Congo in Africa: latitude -5.142 to -5.056, longitude 24.06 to 24.15, and up to 20 km altitude. Clouds (opaque to infrared) inherited.

3. Pseudo-observations from three instruments were constructed as overlapping spatial aggregates with Gaussian weighting functions, and independent multiplicative measurement errors of 15%. Resolutions represent two "edge cases": best-case and current case.

| Case | Radar | Sounder | Microwave |
|---|---|---|---|
| Best | 20km x NA x 300m | 500m x 500m x 1km | 5km x 5km x 1km |
| Current | 20km x NA x 300m | 10km x 10km x 1km | 20km x 20km x 2km |

4. Statistical model:

$$\mathbf{Z} = \mathbf{F}\mathbf{Y} + \boldsymbol{\epsilon},\ \boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{V}_{\epsilon}), \quad \mathbf{Y} = \mathbf{H}\boldsymbol{\eta} + \boldsymbol{\delta},\ \boldsymbol{\delta} \sim \mathcal{N}(\mathbf{0}, \mathbf{V}_{\delta}), \quad \boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \mathbf{K}).$$

Observation vector / Aggregation matrix / Vector of all voxels' "true" values / Vertical basis matrix / Coefficient vector / errors: $\boldsymbol{\epsilon}, \boldsymbol{\delta}$

5. Estimate mean vector and covariance matrix of ``True" vector given "Observation" vector => fused estimate and uncertainty at "target" resolution. Target resolution is finest instrument resolution in each direction. Estimate vertical basis matrix $\mathrm{H}$ from radar-like pseudo-observations, and horizontal spatial dependence $\eta$ from all pseudo-observations.

6. Fused estimates and their uncertainties obtained from usual Gaussian Process conditional mean vector and covariance matrix formulas.

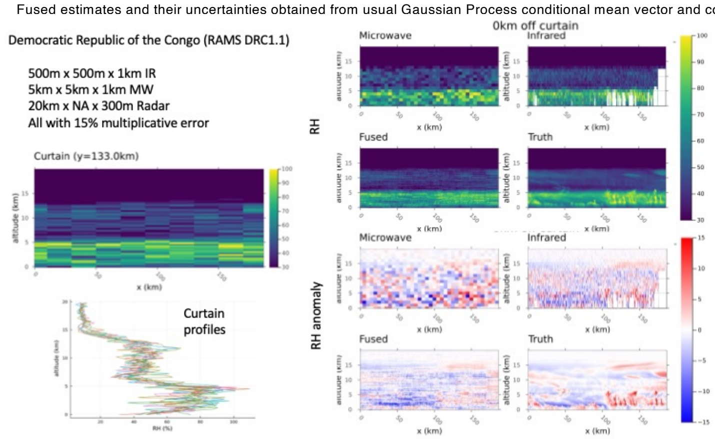
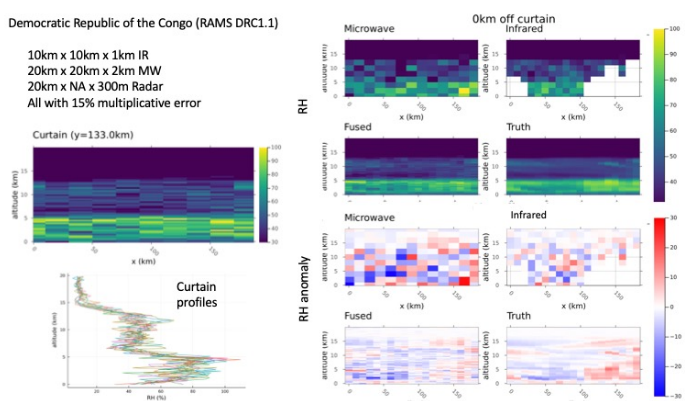



**Figure 1.** Best-case scenario, truth, pseudo-observations, and fused estimates. Left: edge-on view of the radar pseudo-observations ("Curtain") and a representative set of vertical profiles along curtain. Right, top two panels: edge-on views of microwave and infrared pseudo-observations down the center of the region of interest. Second row: fused estimates (left), truth at target resolution (right). Third and fourth rows: same as first and second rows but with mean removed.

**Figure 2.** Current-case scenario. Panels are the same as in Figure 1.




National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

www.nasa.gov

**PI/Task Mgr. Contact Information:**

*818-793-4606, Amy.Braverman@jpl.nasa.gov*

RPD-098 Clearance Number: CL#00-0000