

Breakout Session:  
Integration, Data Products, Accessibility, and Training

Topic: Software, Algorithms,  
and Data Processing

Session moderator: Mariko Oue, Jake Liu, Chris Burghart, Ting-Yu Cha

# Purposes of this session:

- Discuss the community needs and expectations for data integration, data processing, and helping users understand how to work with changes in radar analysis capabilities.
- Discuss in more detail the algorithm development and data processing aspects needed within the community.

# What are the community needs and lacks for advanced software and algorithms?

- Standardized formats and processing and detailed documentation are vitally important
  - There are multiple retrieval tools (3D wind, QPE, microphysics, hydrometeor ID). It's hard to know which tool to use, how to parameterize, uncertainty.
  - Many algorithms can depend on scan strategy (e.g. conventional radar scans), case, radar frequency, environments.
  - Software should be more explicit about tools, modularize, smart hooks
  - a shift away from a download centric data model due to institutional storage limitations
  - Versatile software. LROSE and XRADAR are good tools.
  - Keep good tracking on status of data processing - need reports or metadata to account for changes in processing; this is especially important as we move to new technology
  - There are underutilized datasets (e.g. dual pol data for data assimilation and microphysical retrievals)
  - Documentation of scans,

# What are the community needs and lacks for advanced software and algorithms?

- Developing a plan to work with future PAR
  - Data format (adaptive data format), how to visualize, how to document how data are collected and developed
  - The adaptive scanning capabilities could require new developments in data processing
  - PAR cost - heavily invested on hardware side. It would be nice to support more on software side

# What resources might be missing from the landscape to ensure true community progress?

- A Development Testbed Center (DTC) entity or a centralized repository to support the community for various radar-related software and tools and benchmark for testing and evaluating software products
  - Automated QC, VAP software using ML: dealiasing, recovering good data, 3D wind, microphysical retrievals
  - a shift away from a download centric data model due to institutional storage limitations
  - Virtual observatory (model, LES +radar simulator, IQ simulator) can be useful to evaluate changes in software
  - Intercomparisons of different algorithms
  - Test utilization of datasets in algorithms, retrievals, software, DA
  - Can be used to know what happens when working in a pre-field campaign, which algorithms and processing to use
  - LROSE, [openradarscience.org](http://openradarscience.org) type tools can help

# How can we achieve more seamless integration of algorithms into workflows and understand how they can adapt to changing or updated radar platforms?

- Utilize AI
  - to query data that go into algorithms
- Require diversity of datasets
  - More diverse cases (including edge cases) the more robust we can design your algorithm to different radar platforms.
- Label data that can be found by LLM search queries
  - Collaboration with search engine companies

# Products and applications that require accurate radar data

- Conventional data (e.g. NEXRAD) are still useful to estimate hail products with transition to PAR, which can provide different data where whether current algorithms and their assumptions work is unclear.
- AI/ML can help to develop new algorithms and evaluate conventional algorithms and codes

# Summary and Recommendations

- Standardized formats and processing and documentation are vitally important
- Having a centralized repository (e.g. A Development Testbed Center) for benchmark testing and evaluating software products and accessibility.
- We need to develop a plan to work with future PAR –Data format, data accessibility, software