Focus Group Proposal Whitepaper

ASR Quantification of Uncertainty in Cloud Retrievals (QUICR) Focus Group

Mission Statement

The mission of the ASR Quantification of Uncertainty in Cloud Retrievals (QUICR) Focus Group (FG) is to develop a methodology for characterizing and quantifying uncertainties in current and future ARM cloud retrievals (VAPs and PI products), separately for different cloud regimes, in support of both retrieval algorithm improvement and cloud modeling study.

Motivation

One primary goal of ARM is to obtain detailed observations of cloud microphysical properties using active and passive remote sensors (e.g., millimeter-wave cloud radar, dual frequency radar, scanning radars, Raman lidar, HSRL, micropulse lidars, microwave radiometers, etc.). With the ASR/ARM funding support, scientists have developed various algorithms to retrieve vertical profiles of cloud microphysical properties for a variety of cloud types. As a result, multiple retrieved cloud properties datasets exist in either the ARM archive or its individual PI's groups for a few selected ARM sites and periods. Large uncertainties in these retrievals, however, have been identified in many studies. The uncertainties may arise from instrument limitations, measurement errors, sampling errors, retrieval algorithm deficiencies in assumptions, as well as inconsistent input data and constraints used by different algorithms.

Given the uncertainty in current cloud retrievals, developing a best estimate of cloud microphysical properties with error bars from these existing cloud retrievals has long been desired by the ASR cloud modeling community for both cloud modeling studies and better constraining climate models. However, this has been proven very difficult to achieve due to the complexity of clouds and our insufficient knowledge on cloud microphysical processes and their retrievals. In addition, the need to balance retrieval complexity (often via multiple measurements) vs. ease and broadness of application makes even harder for producing "best estimate" retrieval products. The lack of an integrated effort to target this difficult problem by the ASR science community is also part of the reasons for the slow progress made in quantifying the uncertainty in ARM cloud retrieval products. Therefore, we propose to form a focus group that consists of ASR scientists (PIs) with expertise in cloud retrievals and ARM infrastructure teams so that we can make an integrated effort to address the uncertainty issue associated with ARM cloud retrievals. The proposed "Quantification of Uncertainty in Cloud Retrievals (QUICR)" Focus Group will be under the ASR Cloud Lifecycle Working Group structure.

Objectives

The science objectives of the Focus Group include the follows.

- Systematically analyze differences in current widely accepted cloud retrieval algorithms and their products.

- Understand potential uncertainties in each of the selected algorithms.
- Evaluate the accuracy of and test assumptions made in these cloud retrievals with in-situ aircraft data and observed surface and top of the atmosphere (TOA) radiative fluxes, as well as observation system simulation experiment (OSSE) datasets.
- Evaluate the value of implementing more complex retrievals (which can only be applied for a certain cases) compared to simple ones for easy and broad application.
- Apply advanced statistical methods to quantify uncertainties in these cloud retrievals for different cloud regimes.

Approaches

To achieve these science goals, we propose the following research activities:

1) Create an ensemble dataset for cloud retrievals

The first step is to create a cloud retrieval ensemble dataset (CRED) from current available ARM retrieved cloud microphysical properties. Nine major ARM ground-based cloud retrievals, which are available for multi-years at the five ARM permanent research sites from different research groups, are being assembled into a single dataset with a uniform temporal and vertical resolution by the ARM infrastructure team at Lawrence Livermore National Laboratory (LLNL) for accelerating the uncertainty analysis and easy use by data users. The CRED data can be used to systematically analyze differences and explore issues with current ARM retrievals for further analysis. In addition, CRED could provide a rough estimate of uncertainties in these cloud retrievals based on current instruments and retrieval techniques provided that the used algorithms were reasonably designed to retrieve cloud properties for a certain type of cloud.

One problem with the current CRED is that the uncertainty in each of the ensemble members is unknown. To address this issue, our next step is to generate the ensemble dataset for each of the algorithms, which can be achieved by perturbing key parameters and/or changing key assumptions used in these selected retrieval methods. This process could be helped with the use of the uncertainty quantification analysis software (PSUADE - Problem Solving environment for Uncertainty Analysis and Design Exploration) developed by LLNL, which allows to effectively perform in-depth uncertainty analyses for uncertainty assessment, sensitivity analysis, parameter estimation, and calibration and optimization. The goal is to improve our understanding of the uncertainty associated with each of these retrieval methods and provide necessary uncertainty information for further quantifying the uncertainty using advanced statistical methods such as the Bayesian approach.

2) Systematically evaluate cloud retrievals using BBHRP and in-situ data, as well as OSSE datasets

In-situ microphysical measurements and radiative closure checks using surface and TOA fluxes are often used to assess the quality of cloud retrievals. Both approaches have different advantages and shortcomings, which we intend to exploit in this focus group.

Recent in-situ datasets do include both particle size distribution and bulk cloud microphysics measurements over or near the ARM sites, which allows for extensive checking of cloud retrieval assumptions and performances, but on a limited number of meteorological situations. An important task that will be undertaken in this focus group will be to liaise with the PIs producing those in-situ datasets in order to produce a cloud retrieval test case library from each field experiment.

The radiative closure approach can be applied to long times series over the ARM sites to produce a statistical estimate of retrieval accuracy in different meteorological situations, but the approach itself does not allow for a vertically-resolved check of the cloud retrieval accuracy. The created cloud retrieval ensemble dataset will significantly accelerate the process to evaluate and inter-compare these retrieval products, especially with the help of the ARM BroadBand Heating Rate Profile (BBHRP) testbed. This could be done by integrating the CRED data into the Radiative Important Parameters Best Estimate (RIPBE) value-added product that ARM is developing for providing all the required parameters to run BBHRP (led by Sally McFarlane at PNNL).

In order to feed this 2-prong approach, we plan to build-up a case library that contains uniform input data files suitable for running these retrieval algorithms as well as in-situ data and surface and TOA radiative fluxes suitable for testing the performance of these algorithms in retrieving cloud properties over different cloud regimes. The cloud retrieval test case library will be very useful for future retrieval algorithm development and improvement. Knowledge learned from such validation and intercomparison studies will be used to improve the algorithms and guide the uncertainty quantification of these cloud retrievals.

Another potential approach would be to develop observation system simulation experiment (OSSE) datasets in addition to the in situ and flux closure approaches. These would include something like bin-resolved microphysics, dynamics, etc that could be forward modeled into measurements. Then the measurements could be run through algorithms and compared to reality. In the future, new algorithms that use new data systems could be tested with the OSSE data.

3) Implement advanced statistical methods to quantify uncertainties in these cloud retrievals for different cloud regimes

CRED/PSUADE analysis in Step 1) allows the creation of an ensemble of values of a cloud parameter at a given spatial-temporal location. Probabilistic models and non-parametric statistical analysis can be applied to the ensemble to estimate the critical statistical indices of a cloud property, such as probability density function (PDF), mean, median, variance, skewness, and kurtosis. The data evaluation using BBHRP and in-situ observations in Step 2) will lead to a case library which will provide cloud properties at given locations with specific temporal resolution. The case library allows two important analyses: (a) statistical inference to decide whether any two given retrieval results are significantly different, and to discern the differences in physical and mathematical assumptions used in the development of the retrieval algorithms; and (b) optimal blending of multiple retrieval results to form the “best estimate” of cloud parameters at a given location and specific temporal resolution. In (b), the in-situ case library dataset will be used to determine the prior distribution, and the others will be used to build

likelihood functions. The optimally blended results will be obtained via a Bayesian posterior estimate or other statistical methods and are expressed in terms of PDFs. The PDFs will be very useful for testing climate models' sensitivity to cloud regimes and properties.

Therefore, the powerful Bayesian posterior estimation approach or other statistical methods contained in PSUADE will explicitly provide uncertainty quantification of these retrievals based on the knowledge gained from the uncertainty analyses with respect to physical assumptions, instrument types, mathematical inverse problems in retrieval algorithms, statistical inference, and various statistical sampling methods in PSUADE. A suite of advanced statistical techniques will be applied separately for different cloud regimes, since the retrieval uncertainty is known to be very highly dependent on the type of cloud. As a starting point, this focus group will emphasis on high level ice clouds and boundary layer overcast clouds.

Milestones

1st year:

- Create a cloud retrieval ensemble dataset (CRED) from current available ARM cloud retrievals
- Perform in-depth analysis to understand what cause the differences between these retrievals
- Explore issues associated with these cloud retrievals for further studies
- Release the CRED data to the ARM archive with detailed description of the data and algorithms.
- Write research paper(s) to document what we have learned from these research activities

2nd year:

- Select a few retrieval algorithms that could mostly represent current ASR/ARM retrieval state-of-the-art for further uncertainty analysis
- Implement the code of these selected methods on a LLNL computing platform
- Create uniform input data from ARM "best estimate" products suitable for running these retrieval algorithms
- Create ensemble datasets for each of the selected algorithms by using PSUADE, which will be also used to help uncertainty analysis
- Explore issues associate with these cloud retrievals for further studies
- Write research paper(s) to document what we have learned from these research activities

3rd year:

- Build-up a case library that contains (i) uniform input data files suitable for running these retrieval algorithms, (ii) in-situ data, (iii) surface and TOA radiative fluxes, and (iv) OSSE datasets suitable for testing the performance of these algorithms. The test will demonstrate the effectiveness of retrieving cloud properties over different cloud regimes

- Integrate CRED into RIPBE (Radiative Important Parameters Best Estimate) for BBHRP tests
- Define cases for cloud retrieval intercomparison studies
- Explore issues associate with these cloud retrievals for further studies
- Write research paper(s) to document what we have learned from these research activities

4th year:

- Improve retrieval algorithms based on what we have learned from these studies
- Produce cloud retrieval ensemble datasets using improved algorithms
- Evaluate the new data using BBHRP and aircraft data
- Write research paper(s) to document what we have learned from these research activities

5th year:

- Apply advanced statistical methods to perform in-depth uncertainty analysis and quantify the uncertainty in these cloud retrievals over different cloud regimes based on our current knowledge on these cloud retrieval products
- A cloud retrieval testbed is built-up for testing existing and new algorithms in the future

Participants

Shaocheng Xie (Co-Chair), Lawrence Livermore National Laboratory
Alain Protat (Co-Chair), The Centre for Australian Weather and Climate Research, Melbourne, Victoria, Australia, and Laboratoire ATmosphère, Milieux, Observations Spatiales, Guyancourt, France.
Chuanfeng Zhao, Lawrence Livermore National Laboratory
Scott Collis, Argonne National Laboratory
Jennifer Comstock, Pacific Northwest National Lab, Richland, WA, USA
Min Deng, University of Wyoming, Laramie, WY, USA
Maureen Dunn, Brookhaven National Lab, NY, USA
Dong Huang, Brookhaven National Lab, NY, USA
Robin Hogan, University of Reading, Reading, UK
Michael Jensen, Brookhaven National Lab, NY, USA
Stephen A. Klein, Lawrence Livermore National Laboratory, Livermore, CA, USA
Jay Mace, University of Utah, UT, USA
Renata McCoy, Lawrence Livermore National Laboratory, Livermore, CA, USA
Sally McFarlane, Pacific Northwest National Lab, Richland, WA, USA
Ewan Oconnor, University of Reading, Reading, UK
Samuel Shen, Department of Mathematics and Statistics, San Diego State University
Matthew Shupe, University of Colorado, CO, USA
David Turner, NOAA National Severe Storms Lab, OK, USA
Zhien Wang, University of Wyoming, Laramie, WY, USA

Performance Metrics

- A cloud retrieval ensemble dataset based on current ARM available cloud retrieval products will be delivered to the ARM archive in a short term (1st year).
- Cloud retrieval ensemble datasets produced from each of the selected retrieval algorithms will be generated by perturbing key parameters and/or changing assumptions used in these algorithms (2nd – 3rd year).
- A case library will be developed that contains (i) uniform input data files suitable for running these retrieval algorithms, (ii) in-situ data, (iii) surface and TOA radiative fluxes, and (iv) OSSE datasets suitable for testing cloud retrievals (3rd – 4th year).
- The uncertainty in these cloud retrievals will be quantified for different cloud regimes in the 4th-5th years. The best and probabilistic estimate will be made for needed cloud parameters.
- Improved retrieval algorithms will be developed and made available to the cloud research community. Research papers will be produced from the Focus Group research activities.