



Supplement of

Characterization of liquid cloud profiles using global collocated active radar and passive polarimetric cloud measurements

Yutong Wang et al.

Correspondence to: Huazhe Shang (shanghz@radi.ac.cn) and Husi Letu (husiletuw@hotmail.com)

The copyright of individual parts of the supplement might differ from the article licence.

S1. Methodology Details for Cloud Base Height Retrieval

This method primarily retrieves cloud bottom height (CBH) by utilizing a deep neural network (DNN) based on multi-angle passive satellite observations—POLDER/Parasol data in the Oxygen A (O-A) band.

S1.1. Data and Matchups

The data used in this study consist of four months (March, June, September, and December) of Parasol L1 products from 2007 and the corresponding CloudSat-CALIPSO 2B-CLDCLASS-LIDAR joint product. Spatial collocation between active and passive sensors maintains an accuracy within 0.01° , and temporal differences are negligible owing to the near-simultaneous measurements provided by the A-Train satellite constellation. During data preprocessing, only pixels that are identified as cloudy in Parasol data and simultaneously classified as single-layer clouds in the CloudSat-CALIPSO product are selected.

S1.2. Predictors/Features

The training data used is shown in the table below.

Table S1. The training data and related parameters

Satellite	Product Name	Parameters	Note
Parasol	L1	longitude	
		latitude	
		elevation	
		cloud indicator	Cloud detection
		I763NP (14 view angles)	OA band
		I765NP (14 view angles)	OA band
CloudSat-CALIPSO	2B-CLDCLASS-LIDAR	Longitude	
		Latitude	
		CloudLayerBase	Cloud base height
		Cloudlayer	

S1.3. Model Architecture and Hyperparameters

This method primarily employs deep neural networks to train the combined active and passive data, with specific parameter settings detailed in the table below.

Table S2. Parameter settings for training the model

Parameter Category	Parameter Name	Value / Setting
Randomness control	random_state	42
	dense_units	[1024, 512, 256, 64]
Network Architecture Parameters	dense_activations	['relu', 'relu', 'relu', 'relu']
	output_activation	'relu'
	batch_size	21000
	epochs	44
Training Parameters / Hyperparameters	optimizer	'adam'
	loss_function	'mae'
	metrics	['mae']

S1.4. Training/Validation Design and Leakage Controls

For each month, the last seven days of data serve as the test set, while the remaining days form the training set.

S1.5. Metrics by Regime

The validation results of the model for predicting CBH across different cloud phases and various cloud types are as follows.

Table S3. The validation results for different cloud phases.

Cloud phase	Error metric	Value
Water cloud	N	188282
	MAE	0.3831
	Bias	-0.0532
	RMSE	0.9361
	R	0.6295
Ice cloud	N	76507
	MAE	0.6403
	Bias	-0.1291
	RMSE	1.1839
	R	0.6877
Mixed cloud	N	87373
	MAE	1.7632
	Bias	1.1279
	RMSE	3.1291
	R	0.7329

Table S4. The validation results for different cloud types.

Cloud type	Error metric	Value	Cloud type	Error metric	Value
Cirrus	N	36577	Stratus	N	37464
	MAE	2.6885		MAE	1.4915
	Bias	2.1800		Bias	0.5277
	RMSE	4.4079		RMSE	2.0529
	R	0.1909		R	0.6539
Altostratus	N	27637	Stratoculumus	N	9850
	MAE	1.1790		MAE	0.1520
	Bias	0.2194		Bias	0.0006
	RMSE	1.7428		RMSE	0.2630
	R	0.6458		R	0.8246
Altoculumus	N	153395	Culumus	N	47891
	MAE	0.2653		MAE	0.5238
	Bias	0.0051		Bias	-0.3217
	RMSE	0.4765		RMSE	1.4175
	R	0.7447		R	0.2962

S1.6. Inference Steps

The detailed steps for performing cloud base height prediction using the proposed method are as follows:

- Step 1. Read Parasol L1 data.
- Step 2. Apply quality control flags to filter cloudy pixels.
- Step 3. Extract all predictor variables specified in the "feature list" from the data.
- Step 4. Preprocess the extracted features exactly as done during the training stage.
- Step 5. Feed the preprocessed feature matrix into the loaded pre-trained DNN model.
- Step 6. Perform inference to obtain the estimated CBH for each pixel.