Supplement of Earth Syst. Sci. Data, 18, 1287–1305, 2026
https://doi.org/10.5194/essd-18-1287-2026-supplement

Open Access Earth System Science Data

Supplement of

FYAI: a Fengyun satellite-based dataset for atmospheric ice water path

Yifan Yang et al.

Correspondence to: Tingfeng Dou (doutf@ucas.ac.cn)

Table S1. Channel characteristic parameters of FY-3A/B MWHS-I

Channel Number	Center Frequency (GHz)	Polarization	Bandwidth (MHz)	NEΔT (K)
1	89	QH	1500	1
2	118.75±0.08	QV	20	3.6
3	118.75±0.2	QV	100	2
4	118.75±0.3	QV	165	1.6
5	118.75±0.8	QV	200	1.6

Table S2. Channel characteristic parameters of FY-3C/D MWHS-II

Channel Number	Center Frequency (GHz)	Polarization	Bandwidth (MHz)	NEΔT (K)	Calibration Accuracy (K)	3dB Beamwidth
1	89	QH	1500	1	1.3	2.0°
2	118.75±0.08	QV	20	3.6	2	2.0°
3	118.75±0.2	QV	100	2	2	2.0°
4	118.75±0.3	QV	165	1.6	2	2.0°
5	118.75±0.8	QV	200	1.6	2	2.0°
6	118.75±1.1	QV	200	1.6	2	2.0°
7	118.75±2.5	QV	200	1.6	2	2.0°
8	118.75±3.0	QV	1000	1	2	2.0°
9	118.75±5.0	QV	2000	1	2	2.0°
10	150	QH	1500	1	1.3	1.1°
11	183.31±1	QV	500	1	1.3	1.1°
12	183.31±1.8	QV	700	1	1.3	1.1°
13	183.31±3	QV	1000	1	1.3	1.1°
14	183.31±4.5	QV	2000	1	1.3	1.1°
15	183.31±7	QV	2000	1	1.3	1.1°

Table S3. Channel characteristic parameters of FY-3E/F MWHS-II

Channel Number	Center Frequency (GHz)	Polarization	Bandwidth (MHz)	NEΔT (K)	Calibration Accuracy (K)	3dB Beamwidth
1	89	QH	1500	0.4	1	2.0°
2	118.75±0.08	QV	20	2.2	2.4	2.0°
3	118.75±0.2	QV	100	1	1.2	2.0°
4	118.75±0.3	QV	165	0.8	1.2	2.0°
5	118.75±0.8	QV	200	0.8	1.2	2.0°
6	118.75±1.1	QV	200	0.8	1	2.0°
7	118.75±2.5	QV	200	0.8	1	2.0°
8	118.75±3.0	QV	1000	0.5	1	2.0°
9	118.75±5.0	QV	2000	0.5	1	2.0°
10	166	QH	1500	0.4	1	1.1°
11	183.31±1	QV	500	0.6	1	1.1°
12	183.31±1.8	QV	700	0.6	1	1.1°
13	183.31±3	QV	1000	0.5	1	1.1°
14	183.31±4.5	QV	2000	0.5	1	1.1°
15	183.31±7	QV	2000	0.5	1	1.1°

Table S4. Data time spans for each satellite

Satellite	Instrument	Time Span
FY-3A	MWHS-I	2010-01-01-2013-12-31
FY-3B	MWHS-I	2010-11-18-2019-12-24
FY-3C	MWHS-II	2014-01-01-2024-11-27
FY-3D	MWHS-II	2019-01-01- 2024-12-31
FY-3E	MWHS-II	2023-01-01- 2024-12-31
FY-3F	MWHS-II	2024-01-11- 2024-12-31

Table S5. Hyperparameters of MWHS-I/II model

Model Hyperparameters	Value
input_dim	13, 18*
dim_model	256
num_blocks	8
dim_ff	512
dim_attn_bottleneck	64
dim_scale_bottleneck	64
output_intermediate_dim	256
input_dropout	0.2
res_dropout	0.2, 0.24

*denotes the MWHS-II model.

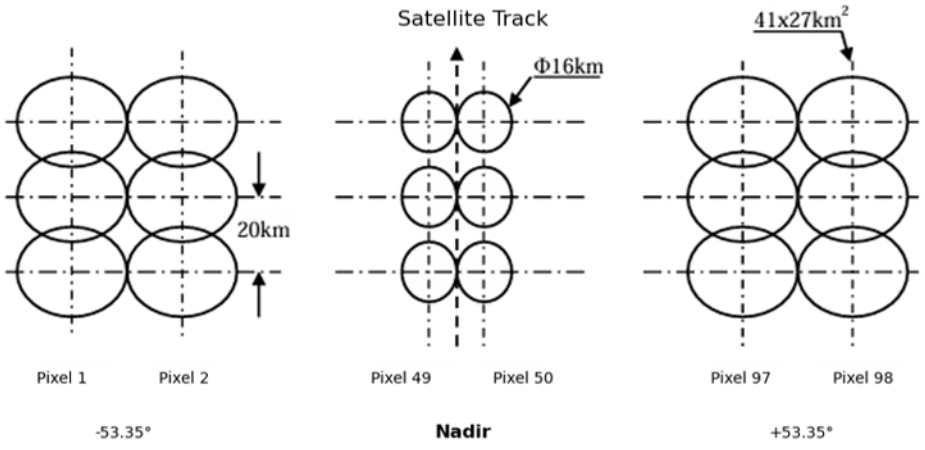

Figure S1. Schematic diagram illustrating the variation of MWHS-II spatial resolution with scan angle.

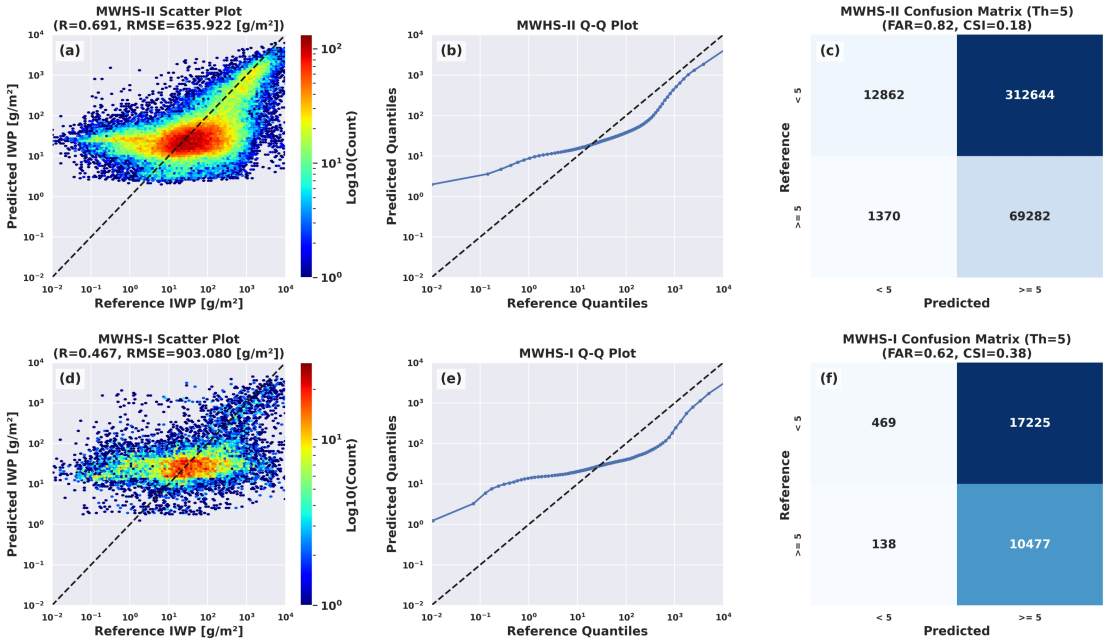

Figure S2. Performance metrics of the QRNN model on the SIWP test dataset. (a) scatter plot of mode-retrieved SIWP values versus reference values on MWHS-II; (b) Q-Q plot of predicted values versus reference values on MWHS-II; (c) confusion matrix for MWHS-II using an SIWP threshold of 5 g/m²; (d) analogous to (a) but for MWHS-I; (e) analogous to (b) but for MWHS-I; (f) analogous to (c) but for MWHS-I.

S1. Feature Gating and Bottleneck Attention Mechanism

To enhance the model's sensitivity to key physical information within the input features while suppressing background noise, we incorporated a lightweight channel-wise Attention Block at the output of the Gated MLP within each residual block. Unlike sequence-based attention, this module functions as a feature recalibration mechanism. It computes importance weights (ranging from 0 to 1) for each feature dimension and applies an element-wise gating operation to the feature map. Specifically, a Sigmoid activation function is employed to generate a mask that dynamically selects and emphasizes the feature components most relevant to IWP retrieval, thereby improving the model's robustness in handling complex non-linear mappings.

Structurally, to balance model expressivity with computational efficiency, the attention module adopts a Bottleneck Architecture. The high-dimensional features (d_{model}) are first compressed into a lower-dimensional bottleneck space ($d_{bottleneck}$) via a linear layer, followed by a ReLU activation, and then projected back to the original dimension. This "compress-and-expand" design significantly reduces the number of parameters while forcing the network to distill compact and representative feature representations. Furthermore, a learnable Adaptive Weight is introduced into the residual connection, allowing the model to automatically adjust the contribution of the attention branch during training, facilitating a smooth transition from identity mapping to refined feature extraction.

S2. Performance Evaluation of SIWP Retrieval

Based on the test results shown in Fig. S2, we provide an objective assessment of the QRNN model's performance in retrieving suspended ice water path (SIWP), with a focus on the inherent challenges and limitations associated with this physical variable. Compared to total ice water path (IWP), SIWP signals are generally weaker and more strongly influenced by complex cloud phase states, resulting in a notable reduction in overall regression accuracy. Quantitative analysis reveals that FY-3D (MWHS-II) achieves a correlation coefficient (R) of 0.70 and an RMSE of 635.92 g/m². Although these metrics are superior to those of FY-3B (R = 0.47, RMSE = 903.08 g/m²), the scatter plots for both sensors (Fig. S2a and Fig. S2d) exhibit considerable dispersion, indicating significantly increased uncertainty in passive microwave sensing of SIWP. This retrieval uncertainty is further elucidated by the Q-Q plots (Fig. S2b and Fig. S2e). Unlike the IWP results, the predicted quantiles for SIWP deviate noticeably from the reference line in the low-value regime (approximately $< 10^{-1}$ g/m²), manifesting as a distinct positive bias (overestimation). This suggests that, constrained by the sensitivity thresholds of microwave channels to liquid or mixed-phase water, the model tends to systematically overestimate trace amounts or thin layers of supercooled water, struggling to resolve values below the sensor's detection limit.

These physical limitations are particularly evident in the binary classification assessment using a 5 g/m² threshold. While FY-3D retains an exceptionally high False Alarm Ratio (FAR = 0.82), resulting in a low Critical Success Index (CSI) of only 0.18 (Fig. S2c). Conversely, FY-3B (Fig. S2f), despite a higher miss rate, maintains a relatively lower FAR (0.62) and a slightly higher CSI (0.38). This implies that for SIWP retrieval, simply increasing the number of channels (as in MWHS-II) improves regression correlation in strong-signal regions but does not necessarily resolve the challenge of background noise suppression in weak-signal regimes. Ultimately, while the model effectively identifies stronger SIWP events, its quantitative accuracy in the low-value range and detection capability for weak signals remain limited by the intrinsic physical characteristics of the sensors.