National Aeronautics and
Space Administration

Commercial Satellite Data Acquisition Program

Tomorrow.io Radar Quality Assessment Report

Commercial Satellite Data Acquisition Program Tomorrow.io Radar Quality Assessment Report

Signature/Approval Page

Approval by:

03/09/2026

Melissa Yang Martin
Commercial Satellite Data Acquisition Program Manager
Earth Science Division
Headquarters/NASA

Date

Accepted by:

03/09/2026

Dana Ostrenga
Commercial Satellite Data Acquisition Project Manager
Earth Science Division
GSFC/NASA

Date

Preface

This document is under CSDA Project configuration control. Once this document is approved, CSDA approved changes are handled in accordance with Class I and Class II change control requirements described in the CSDA Configuration Management Procedures based on NASA standard configuration practices, and changes to this document shall be made by document change notice (DCN), documented in the Change History Log or by complete revision.

Abstract

The evaluation summarized in this report was conducted by subject matter experts (SMEs) funded by NASA's Commercial Satellite Data Acquisition (CSDA) Program. The SMEs evaluated the quality of Tomorrow.io precipitation radar data for the NASA Earth science research and applications community. The results of the evaluation help to inform NASA program management on the quality of the data for NASA science.

Authored and prepared by

Sarah Ringerud
Goddard Space Flight Center
National Aeronautics and Space Administration
United States of America

Daniel Watters
University of Oklahoma
United States of America

Pierre Kirstetter
University of Oklahoma
United States of America

Mircea Grecu
Morgan State University
National Aeronautics and Space Administration
United States of America

Jaime Nickeson
Science Systems and Applications Inc
National Aeronautics and Space Administration
United States of America

Frederick Policelli
Goddard Space Flight Center
National Aeronautics and Space Administration
United States of America

Change History Log

Revision	Effective Date	Description of Changes
1.0	03/10/2026	First reviewed version for public release

Table of Contents

List of Figures

Acronyms & Abbreviations

AI	Artificial Intelligence
ATBD	Algorithm Theoretical Basis Document
AWS	Amazon Web Service
CIWRO	Cooperative Institute for Severe and High-Impact Weather Research and Operations
CONUS	Contiguous United States
CSDA	Commercial Satellite Data Acquisition Program
CSI	Critical Success Index
DEM	Digital Elevation Model
EDAP	Earthnet Data Assessment Pilot
EIA	Earth incidence angle
ESA	European Space Agency
ESD	Earth Science Division
EULA	End-User License Agreement
FAR	False Alarm Ratio
GFS	Global Forecast System
GPM	Global Precipitation Mission (NASA)
GV	Ground Validation
HSS	Heidke Skill Score
JAXA	Japan Aerospace Exploration Agency
ML	Machine Learning
MRMS	Multi-Radar/Multi-Sensor
NASA	National Aeronautics and Space Administration
NRCS	Normalized Radar Cross Section
NetCDF	Network Common Data Form
NEXRAD	Next Generation Weather Radar
NN	Neural Network
NOAA	National Oceanic and Atmospheric Administration
NSSL	National Serve Storms Laboratory
NWP	Numerical Weather Prediction
PDF	Probability Density Function
PIA	Path Integrated Attenuation
POD	Probability of Detection
PRF	Pulse Repetition Frequency
QC	Quality Control
QRNN	Quantile Regression Neural Network
SME	Subject Matter Expert
SNR	Signal-to-Noise Ratio
TR1	Tomorrow.io Radar 1 (TR1) spacecraft
TR2	Tomorrow.io Radar 2 (TR1) spacecraft
UMM	Unified Metadata Model

Executive Summary

The Commercial Satellite Data Acquisition (CSDA) Program was established to identify, evaluate, and acquire data from commercial sources that support the National Aeronautics and Space Administration (NASA) Earth science research and application goals. NASA's Earth Science Division (ESD) recognizes the potential impact commercial satellite constellations may have in encouraging/enabling efficient approaches to advancing Earth System Science and applications development for societal benefit. Commercially acquired data may also provide a cost-effective means to augment and/or complement the suite of Earth observations acquired by NASA and other U.S. government agencies and those by international partners and agencies.

Products obtained from Tomorrow.io evaluated here include Level 1C reflectivities as well as the derived precipitation products. There were 1200 Granules (defined by Tomorrow.io as a discrete payload “task”) obtained from data collected in October and November of 2023 distributed in NetCDF files. Observations within each granule have identical radar sampling parameters, however the sampling is irregular in time.

This evaluation of Tomorrow.io precipitation radar performance was carried out by NASA subject matter experts (SMEs) that were enlisted to evaluate the fundamental quality of the Tomorrow.io data following the draft Joint NASA/European Space Agency (ESA) assessment guidelines developed in 2025.

Only documents provided to NASA by Tomorrow.io were considered as part of the assessment, which included a Calibration Report, an algorithm theoretical basis document (ATBD), and a File Specifications Document. Additional documentation with more detailed description of the calibration and validation procedures may be available online or in the literature, but any document not listed in this report was not considered for the evaluation. The product information provided in the vendor documentation guides and the product metadata together provided adequate information to work with the data.

The quality assessment was performed on the Level 1 reflectivity profile data products as well as the Level 2 Precipitation products. File formats are NetCDF with easily readable headers and available metadata. Documentation thoroughly describes pre-launch calibration activities for both radars as well as ongoing post-launch assessments. Equations for calculating and implementing the calibration factor are given in the calibration document and ATBD, as are expressions for reflectivity error, minimum detectable reflectivity, and geolocation error. The Level 2 algorithm is well described including surface reference technique calculations for Path Integrated Attenuation (PIA) and a neural network (NN) for precipitation retrievals. Flags are defined and included for quality control. Documentation in general is well-presented and thorough.

An independent quality assessment of the precipitation product and geolocation accuracy was performed by the SME team. The quality analysis results were generally in agreement with the analysis provided by Tomorrow.io in the ATBD. The geolocation assessment shows excellent correlation of 0.98 with a digital elevation model (DEM) reference. Comparisons to ground radar were in good agreement for both radars, with correlations to ground radar of 0.73 and 0.93. R2 showed slightly higher accuracy than R1, with biases of –22% (R1) and –6% (R2). Based on the positive quality and utility evaluation results, we recommend that Tomorrow.io precipitation radar data be considered for NASA scientific use, contingent upon alignment with specific science objectives and objectives and application needs.

1 Background

The Commercial Satellite Data Acquisition (CSDA) Program was established by NASA's Earth Science Division (ESD) in 2020 following the successful completion of the Private-Sector Small Constellation Satellite Data Product Pilot. The program's primary objective is to identify, evaluate, and acquire commercial remote sensing data that enhances NASA's Earth science research and applications. CSDA provides structured on-ramping opportunities for emerging commercial satellite data vendors, enabling NASA to continuously integrate innovative data sources as the private sector evolves. By leveraging these partnerships, NASA's ESD aims to accelerate scientific discovery and expand applications of Earth observation data for societal benefit.

Since the initial pilot, the Program has conducted three on-ramp activities, resulting in the addition of several vendors into sustainment. In 2024, CSDA streamlined its evaluation process by introducing high-quality, SME-led data assessments, accelerating reviews and strengthening NASA's engagement with the rapidly growing commercial data ecosystem. This evaluation framework not only ensures NASA gains timely access to high-quality, mission-relevant commercial data, but also provides valuable feedback to private-sector providers, fostering innovation, improved data products, and alignment of industry capabilities with NASA's evolving scientific needs.

1.1 Tomorrow.io Constellation Evaluation

In this report, CSDA provides an evaluation of the quality of data provided by the Tomorrow.io Ka-band Precipitation Radars R1 and R2 for advancing NASA's Earth system science research and applications. Tomorrow.io is a weather modeling company founded in 2016 providing real-time weather forecasts utilizing AI and proprietary weather satellites including Ka-band radar and passive microwave sounding radiometers.

The Tomorrow.io radar spacecraft are designated Tomorrow-R1 (TR1) and Tomorrow R2 (TR2) and were launched in April and June of 2023. The TR1 and TR2 spacecraft fly at an altitude of around 550 km in sun-synchronous orbits and utilize a spacecraft attitude determination and control system to vary the antenna pointing and sampling geometries. The instruments are all-solid-state Ka-band (35.5-36 GHz) radars with 1.2 m solid Cassegrain reflector antenna with peak transmit power of roughly 20 W. They are described in the documentation as having high detection sensitivity calculated to be about 8.2 dBZ. Nyquist sampling of the ground footprint is attained with 4.5 km resolution at an orbital height of about 500 km and a vertical resolution of 250 m. The instruments each employ a fixed, non-scanning reflector antenna with pencil beam antenna pattern and are described by Tomorrow.io as a pathfinder/technology demonstration for a future constellation of 8-12 scanning radars. Spacecraft maneuvers allow for variable pointing within the constraints of the pencil beam antenna pattern. In concert with a plan for 18 passive radiometers, the future Tomorrow.io constellation could offer excellent global coverage for modeling applications and precipitation nowcasting.

NASA acquired a subset of data for the two Tomorrow.io pathfinder radars (TR1 & TR2) for data quality evaluation. Tomorrow-R1 data were provided for October and November 2023 (61-day period), whilst Tomorrow-R2 data were provided for a shorter period of October 10-30, 2023 (21-day period).

Measurement samples from the Tomorrow.io radars over the contiguous United States (CONUS) domain are visualized in Figures 1 & 2, respectively. The 5 km resolution radar footprints are binned at a comparable scale of 0.05° x 0.05° in each figure. Note that samples sizes per grid box can exceed 1, which is explained by the oversampling pattern of the radar beam with a factor of 3.5, as showcased in the Tomorrow.io Pathfinder ATBD. Each Tomorrow.io radar samples across the CONUS, with a higher concentration of samples over what appears to be a ground calibration site in Massachusetts, and a sparsity of measurements elsewhere across New England (and the states of Minnesota, Iowa and Missouri for Tomorrow-R2 only).

The Tomorrow.io radars orbited over the CONUS several hundred times within the evaluation period (TR1: 655 orbits; TR2: 391 orbits) with approximately 20% of each radar's orbits including at least one rainfall measurement (TR1: 142 rainy orbits; TR2: 80 rainy orbits). Precipitation occurrence across all footprint samples for each Tomorrow.io radar is about 5% (TR1: 26,188 rainy footprints; TR2: 13,718 rainy footprints).

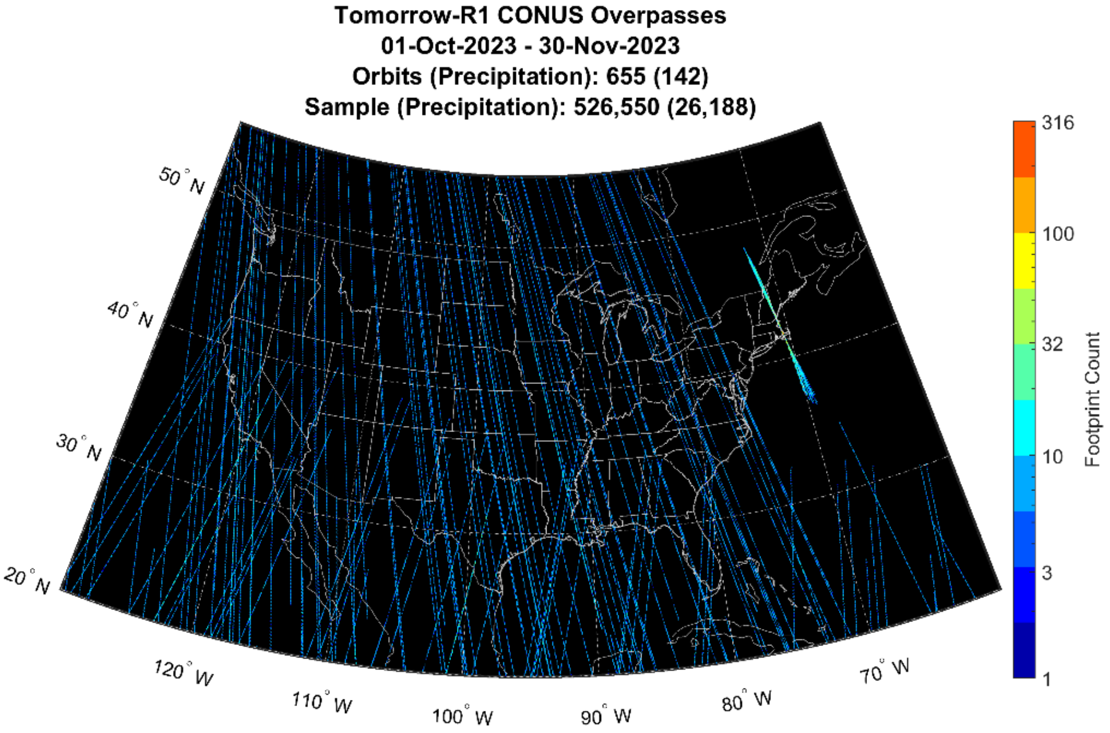


Figure 1. Measurement sample of Tomorrow-R1 over the CONUS domain. The 5 km resolution footprints are binned at a resolution of 0.05° x 0.05°.

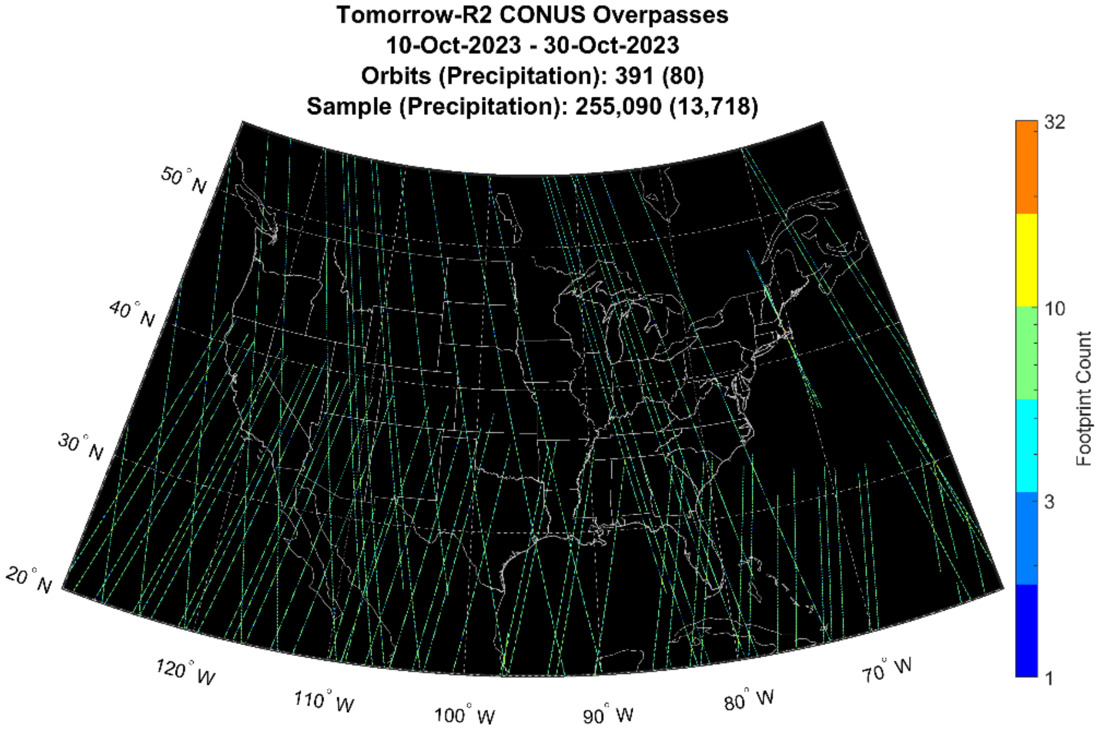
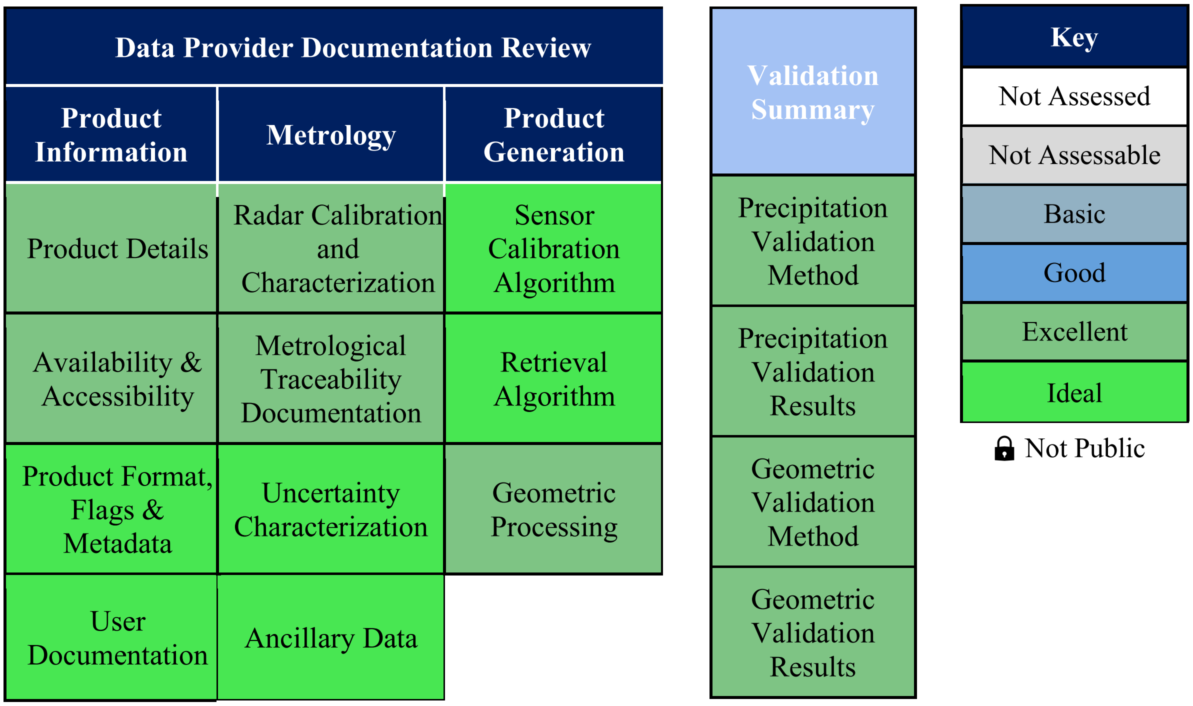


Figure 2. Analogous to Figure 1 except for Tomorrow-R2.

2 Quality Assessment Matrices

2.1 Summary Quality Assessment Matrix



Data Provider Documentation Review		
Product Information	**Metrology**	**Product Generation**
Product Details	Radar Calibration and Characterization	Sensor Calibration Algorithm
Availability & Accessibility	Metrological Traceability Documentation	Retrieval Algorithm
Product Format, Flags & Metadata	Uncertainty Characterization	Geometric Processing
User Documentation	Ancillary Data	

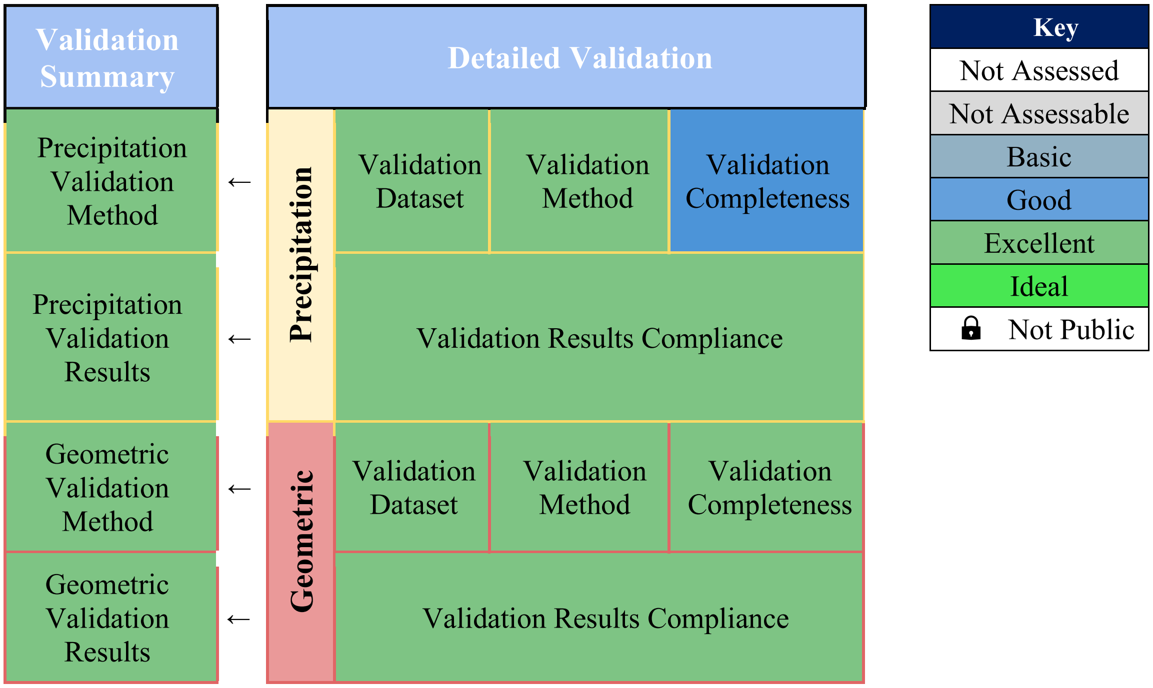
Validation Summary
Precipitation Validation Method
Precipitation Validation Results
Geometric Validation Method
Geometric Validation Results

Key
Not Assessed
Not Assessable
Basic
Good
Excellent
Ideal

🔒 Not Public

Figure 3. Summary Quality Assessment Matrix for Tomorrow.io radar data

2.2 Detailed Validation Maturity Matrix



Validation Summary

Precipitation Validation Method ←

Precipitation Validation Results ←

Geometric Validation Method ←

Geometric Validation Results ←

Detailed Validation

Precipitation

Validation Dataset | Validation Method | Validation Completeness

Validation Results Compliance

Geometric

Validation Dataset | Validation Method | Validation Completeness

Validation Results Compliance

Key

Not Assessed

Not Assessable

Basic

Good

Excellent

Ideal

🔒 Not Public

Figure 4. Detailed Validation Maturity Matrix for Tomorrow.io radar data, including validation activities for the reflectivity and precipitation products

3 Data Provider Documentation Review

3.1 Product Information

Product Details	
Grade: Excellent	
Justification	Product details are well defined in reference documentation and supporting metadata.
Product Name	L1C-GEOPROF, L2A-PRECIP, AUX
Sensor Name	TR1, TR2
Sensor Type	Ka-band Precipitation Radar
Mission Type	Constellation – 2 satellites
Mission Orbit	Sun Synchronous
Product Version Number	V00B
Product ID	Products descriptively named within NetCDF format
Processing level of product	Level 1 Reflectivity, Level 2 Precipitation
Measured Quantity Name	Reflectivity, Precipitation Rate
Measured Quantity Units	dBZ, mm/hr
Stated Measurement Quality	Maximum Reflectivity Error 2.2 dB
Spatial Resolution	4.5 km horizontal, 250m vertical
Spatial Coverage	Global/variable
Temporal Resolution	Orbit revisit information not included
Temporal Coverage	Two months of non-continuous data obtained for fall 2023
Point of Contact	Ben Kahler ben.kahler@tomorrow.io
Product locator (DOI/URL)	N/A
Conditions for access and use	Public release EULA
Limitations on public access	Public Release EULA
Product Abstract	N/A

Availability & Accessibility	
Grade: Excellent	
Justification	Data were staged on AWS and retrieved via script. File names include time and sensor description. Follows FAIR (Findable, Accessible, Interoperable, and Reusable) data principles
Compliant with FAIR principles	Yes
Availability Status	Not assessed for the general case

Product Format, Flags, and Metadata	
Grade: Ideal	
Justification	Files are NetCDF with easily readable headers and metadata. Products include QC flags described in the File Specification document. Specific flags denote each error type.
Product File Format	netCDF4
Metadata Conventions	UMM-Granule Profile (UMM-G) standard with .json file provided for each granule.
Analysis Ready Data?	Yes

User Documentation	
Grade: Ideal	
Justification	Vendor provided detailed ATBD, File Specification, and Calibration documents, similar in content and format to NASA documentation.
Document	**Reference**
Product User Guide	User Guide and File Specifications for Tomorrow-R1 and Tomorrow-R2 Version 00B Data Products (The Tomorrow Companies Inc., 2024)
ATBD	Algorithm Theoretical Basis Document for the Tomorrow-R1 and Tomorrow R-2 Version 00B Level 1C-GEOPROF and Level 2A-PRECIP Data Products (The Tomorrow Companies Inc., 2024)
Calibration Report	Tomorrow R1 and Tomorrow R2 Initial Calibration Report (The Tomorrow Companies Inc., 2024)

3.2 Metrology

Radar Calibration & Characterization	
Grade: Excellent	
Justification	Calibration methods pre-launch as well as ongoing post-launch activities are well described and documented. Equations for calibration factor and SNR are included along with in-orbit calibration methodology and validation plots.
References	• Tomorrow R1 and Tomorrow R2 Initial Calibration Report (The Tomorrow Companies Inc., 2024) • Algorithm Theoretical Basis Document for the Tomorrow-R1 and Tomorrow R-2 Version 00B Level 1C-GEOPROF and Level 2A-PRECIP Data Products (The Tomorrow Companies Inc., 2024)

Metrological Traceability Documentation	
Grade: Excellent	
Justification	In the Tomorrow.io ATBD, Figure 11 shows the L2 data processing algorithm flow including modules and dependencies. The nature of ML makes the actual retrieval mechanics slightly more opaque.
References	• Algorithm Theoretical Basis Document for the Tomorrow-R1 and Tomorrow R-2 Version 00B Level 1C-GEOPROF and Level 2A-PRECIP Data Products (The Tomorrow Companies Inc., 2024)

Uncertainty Characterization	
Grade: Ideal	
Justification	Reflectivity: the ATBD describes random error sources due to configuration and thermal noise and the calculations used to compute relative error from the signal to noise ratio (SNR) which is included in the L1 output files. Follows the guide to the expression of uncertainty in measurement (GUM) approach. Precipitation: the Quantile Regression Neural Network (QRNN) retrieves a probability density of precipitation and reports quantiles to represent the probability density function (PDF), which provides a measure of uncertainty.
References	• Algorithm Theoretical Basis Document for the Tomorrow-R1 and Tomorrow R-2 Version 00B Level 1C-GEOPROF and Level 2A-PRECIP Data Products (The Tomorrow Companies Inc., 2024)

Ancillary Data	
Grade: Ideal	
Justification	The associated ancillary (AUX) files for each granule include geophysical variables from the National Oceanic and Atmospheric Administration (NOAA) Global Forecast System (GFS). These are used in generating the products but can also be helpful to users for analysis.
References	• User Guide and File Specifications for Tomorrow-R1 and Tomorrow-R2 Version 00B Data Products (The Tomorrow Companies Inc., 2024)

3.3 Product Generation

Sensor Calibration Algorithm	
Grade: Ideal	
Justification	A calibration report is included, documenting pre-flight far-field testing at Astro Digital for R1 and R2 and calibration parameter computations. In-orbit calibration is done via internal calibration using observed ocean normalized radar cross sections compared to a value calculated using Global Precipitation Mission (GPM) KaPR and Numerical Weather Prediction (NWP) model winds. Mean biases are found to be -0.4 dB for R1 and <0.1 dB for R2
References	• Tomorrow R1 and Tomorrow R2 Initial Calibration Report (The Tomorrow Companies Inc., 2024) • Algorithm Theoretical Basis Document for the Tomorrow-R1 and Tomorrow R-2 Version 00B Level 1C-GEOPROF and Level 2A-PRECIP Data Products (The Tomorrow Companies Inc., 2024)

Retrieval Algorithm	
Grade: Ideal	
Justification	L1: Reflectivity calculations from radar echo power are well described, including calibration. Relevant equations are included in an ATBD L2: Precipitation retrieval utilizes the surface reference technique to calculate path integrated attenuation following GPM algorithms, then precipitation is calculated using a neural network approach (RadarNNSolver). The model is trained using observed GPM Ka-band data and internally generated synthetic data. Inputs are reflectivity profiles, freezing level height, path integrated attenuation (PIA), and Earth incidence angle (EIA)
References	• Algorithm Theoretical Basis Document for the Tomorrow-R1 and Tomorrow R-2 Version 00B Level 1C-GEOPROF and Level 2A-PRECIP Data Products (The Tomorrow Companies Inc., 2024)

Geometric Processing	
Grade: Excellent	
Justification	Geolocation is done at L1 and passed through to L2. Equations for using spacecraft Attitude Determination and Control System for determining measurement position and geolocation are outlined in the ATBD
References	• Algorithm Theoretical Basis Document for the Tomorrow-R1 and Tomorrow R-2 Version 00B Level 1C-GEOPROF and Level 2A-PRECIP Data Products (The Tomorrow Companies Inc., 2024)

Mission-Specific Processing	
Grade: Not Assessed	

4 Detailed Validation – Precipitation

4.1 Validation Dataset, Method, and Completeness

The evaluation of the surface precipitation retrievals for the Tomorrow.io pathfinder radars were compared relative to ground validation of the Multi-Radar/Multi-Sensor (MRMS) surface radar network over the CONUS. The MRMS surface radar network is a NOAA-National Severe Storms Laboratory (NSSL) and the University of Oklahoma-Cooperative Institute for Severe and High-Impact Weather Research and Operations (CIWRO) product utilizing the Next Generation Weather Radar (NEXRAD) network and additional processing. This dataset constituted the reference data for the assessment and provides coverage only over the CONUS.

The methods used in this evaluation have been successfully utilized in past studies to evaluate the suite of GPM and TRMM precipitation products (e.g., Kirstetter et al., 2012, 2013, 2015, 2018; Watters et al., 2018, 2024, 2025).

Space-Ground Retrievals Matchup

The ground validated (GV)-MRMS framework ingests the MRMS ground-based surface precipitation estimates (0.01°, 2-minute resolution), principally from NEXRAD ground-based radars, rain gauges and models, and subjects them to a range of quantity and quality control procedures that are described in Kirstetter et al. (2012, 2014, 2018). This framework tailors the MRMS suite of hydrometeorological datasets for serving as a validation source for spaceborne precipitation products. The GV-MRMS processing also matches the ground-based estimates to the spatiotemporal resolution of the Tomorrow.io radars for direct comparison: the ground-based retrievals are upscaled from ~1 km to 5 km using a Gaussian weighting pattern (analogous to the spaceborne radar's antenna gain; see Kirstetter et al., 2012 for details), and the 2-minute period that encapsulates the satellite's overpass is selected. Consequently, the Tomorrow.io radar precipitation retrievals are evaluated at their native spatiotemporal resolution (5 km instantaneous).

Please note that the evaluated comparison sample for each Tomorrow.io radar is reduced from the number of available Tomorrow.io radar measurements (see Figures 1 & 2). This sample reduction is the result of the quality and quantity controls for GV-MRMS such that the Tomorrow.io radars are evaluated only in regions where the surface precipitation reference data are deemed the most accurate and trustworthy.

Statistical Comparison

A range of statistical measures are used to evaluate the detection and quantification ability of the Tomorrow.io radars relative to the GV-MRMS ground reference. A precipitation-no precipitation threshold of 0.1 mm/h is used to define the occurrence of precipitation in the subsequent analysis. This value roughly equates to the Tomorrow.io Ka-band radar sensitivity of 7 dBZ as listed in Table I of the Tomorrow.io ATBD.

To assess the detection skill, a contingency table is used. The four contingency table measures are: hit (*h*, satellite and reference both detect precipitation); miss (*m*, satellite detects no precipitation whilst reference detects precipitation); correct rejection (*c*, satellite and reference both detect no precipitation); and false detection f, satellite detects precipitation whilst reference detects no precipitation). With these contingency table measures, a range of key statistics can be deduced.

$$POD = \frac{h}{h+m} \tag{1}$$

The probability of detection (POD) measures the proportion, expressed as a ratio, of precipitation events reported by the reference that the satellite captures. POD has an optimal score of 1 (full detection of reference precipitation) and a worst score of 0 (no detection).

$$FAR = \frac{f}{f+h} \tag{2}$$

The false alarm ratio (FAR) determines the proportion of satellite-detected precipitation events for which the reference detects no precipitation. FAR has an optimal value of 0 (no false detections) and worst score of 1 (all false detections).

$$CSI = \frac{f}{f+h+m} \tag{3}$$

The Critical Success Index (CSI) is a measure of the satellite's ability to correctly detect precipitation. The CSI quantifies the proportion of all spaceborne-detected and reference-detected precipitation events for which the reference detects no precipitation, whilst the FAR provides a similar metric, though excludes the sample of reference-detected precipitation events from its proportion. It has an optimal score of 1 (completely accurate detection) and a worst score of 0 (completely inaccurate detection).

$$HSS = \frac{2(hc - fm)}{f^2 + m^2 + 2hc + (f+m)(h+c)} \tag{4}$$

The Heidke Skill Score (HSS) quantifies the satellite's ability to distinguish between precipitation and no precipitation, with an optimal score of 1 signifying that the satellite and reference completely align in their detection of precipitation, while a score of 0 signifies that the satellite can distinguish precipitation occurrence no better than random chance when compared to the reference. A HSS above 0.3 represents significant skill in satellite detection of precipitation (Derin et al. 2021).

To assess the quantification skill, the evaluation sample is reduced to precipitation events that are both detected by the spaceborne radar, i.e., both report precipitation rates above the detection threshold of 0.1 mm h^{-1} which corresponds to hits in the contingency table. In the following equations, $R_{sat,i}$ and $R_{ref,i}$ refer to the Tomorrow.io radar and GV-MRMS precipitation rate estimates for the i^{th} footprint, respectively. Systematic errors are measured using the bias and mean absolute bias statistics:

$$\text{Bias [\%]} = \frac{\sum_i^N \Delta R_i}{\sum_i^N R_{ref,i}} \times 100\% \tag{5}$$

$$\text{Mean Absolute Bias [\%]} \quad = \frac{\sum_i^N |\Delta R_i|}{\sum_i^N R_{ref,i}} \times 100\% \tag{6}$$

with $\Delta R_i = R_{sat,i} - R_{ref,i}$ and N is the total number of footprints evaluated for the respective Tomorrow.io radars.

Random errors are measured using the normalized random error and standard deviation metrics. The random error metric quantifies the first order deviation from the systematic bias and is less affected by large discrepancies between the satellite and ground retrievals, and the standard deviation quantifies the second order deviation from the systematic bias. Note that $\overline{\Delta R_i} = \frac{\sum_i^N \Delta R_i}{N}$.

$$\text{Random Error [\%]} \quad = \frac{\sum_i^N |\Delta R_i - \overline{\Delta R_i}|}{\sum_i^N R_{ref,i}} \times 100\% \tag{7}$$

$$\text{Standard Deviation [\%]} \quad = \frac{\sqrt{\frac{1}{N-1}\sum_i^N (\Delta R_i - \overline{\Delta R_i})^2}}{\frac{1}{N}\sum_i^N R_{ref,i}} \quad \times 100\% \tag{8}$$

Finally, the correlation coefficient is utilized as a measure of the linear relationship between the satellite and ground retrievals, with an optimal value of 1. Note that $\overline{R_i} = \frac{\sum_i^N R_i}{N}$.

$$\text{Correlation Coefficient} \quad = \frac{1}{N-1}\sum_i^N \left(\frac{R_{ref,i} - \overline{R_{ref,i}}}{\sqrt{\frac{1}{N-1}\sum_i^N (R_{ref,i} - \overline{R_{ref,i}})^2}}\right)\left(\frac{R_{sat,i} - \overline{R_{sat,i}}}{\sqrt{\frac{1}{N-1}\sum_i^N (R_{sat,i} - \overline{R_{sat,i}})^2}}\right) \tag{9}$$

While the validation was performed only over the Continental United States, and covers a very large area and many meteorological conditions, it is not ideal for global applications.Thus the grade of ‘Good’ in the detailed validation matrix for Validation Completeness.

4.2 Results Compliance

Detection of Precipitation

Figure 5 showcases the contingency table statistics for the Tomorrow.io radars. Both spaceborne radars detect most precipitation events that occur, with similar skill in distinguishing between precipitation and no-precipitation (hits + correct rejections; R1: 79.1%; R2: 77.5%). Tomorrow-R2 slightly exceeds Tomorrow-R1 in the proportion of correct precipitation detections (hits; R1: 68.4%; R2: 71.7%), though Tomorrow-R1 has a higher proportion of correctly identified precipitation detection.

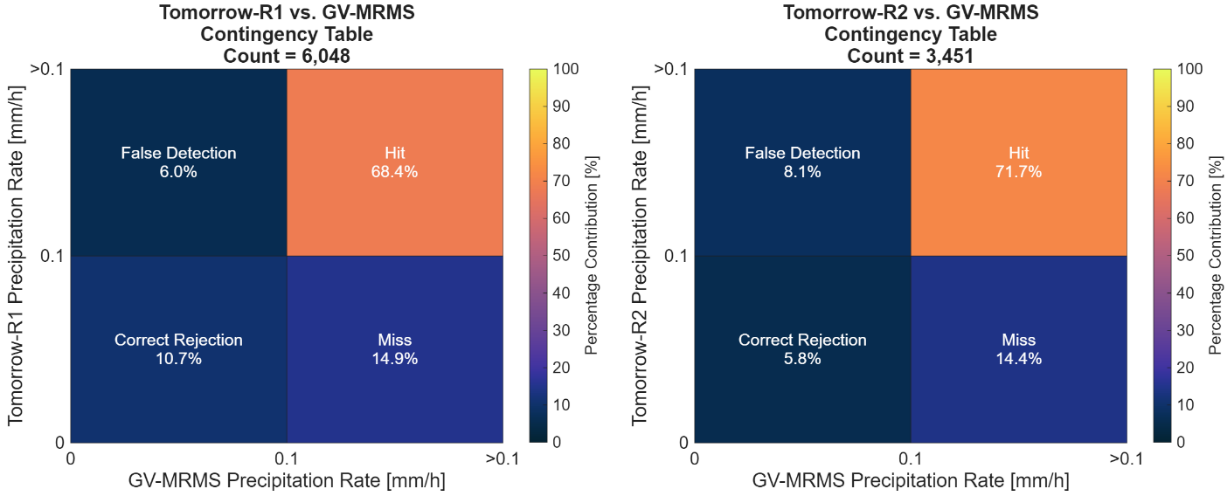


Figure 5. Contingency tables for Tomorrow-R1 (left) and Tomorrow-R2 (right) with respect to GV-MRMS.

Figure 6 further highlights the detection performance of the Tomorrow.io radars relative to GV-MRMS with POD, FAR and CSI. Both Tomorrow.io radars perform well in detecting precipitation and no precipitation, as evident from high POD, low FAR and high CSI scores. When judged by the HSS metric, the radars have good scores (in the [0.2 – 0.4] range), however, Tomorrow-R1 clearly outperforms Tomorrow-R2 due to its higher proficiency for correctly determining instances of no rain (see correct rejections in Fig. 5). The two Tomorrow.io radars have more similar scores for CSI than HSS because CSI does not account for correct rejections. Overall, Tomorrow-R1 slightly outperforms Tomorrow-R2 in terms of distinguishing rain and no rain events (superior FAR, HSS, and CSI), though Tomorrow-R2 is marginally more likely to detect precipitation when it occurs than Tomorrow-R1.

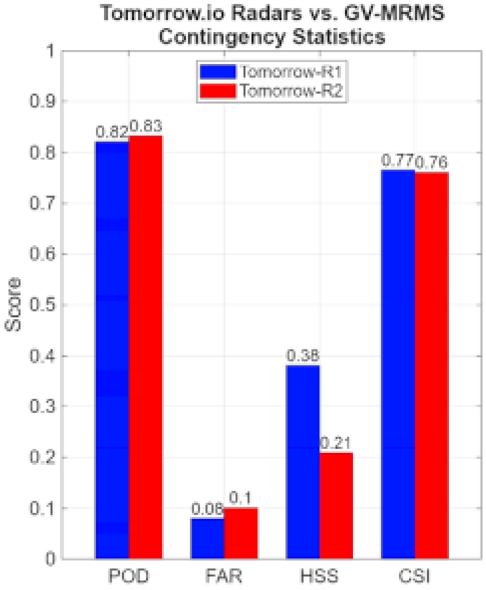


Figure 6. Bar chart of contingency statistics for the Tomorrow.io radars with respect to GV-MRMS.

Quantification of Precipitation

To provide insight into and intercompare the relative proportion of precipitation rates from GV-MRMS and the Tomorrow.io radars, Figure 7 showcases the probability density functions (PDF) for the spaceborne and ground reference retrievals. The precipitation range (R) in this analysis is defined by three categories: light ($0.1 \geq R > 1$ mm h^{-1}); intermediate ($1 \geq R > 10$ mm h^{-1}); and heavy ($R \geq 10$ mm h^{-1}). The Tomorrow.io radars both generally capture the PDFs of precipitation occurrence and volume according to the GV-MRMS reference. Tomorrow-R1 better captures the PDF of light and intermediate precipitation occurrence, whilst Tomorrow-R2 better captures the PDF of heavy precipitation occurrence; the results are reversed when considering precipitation volume.

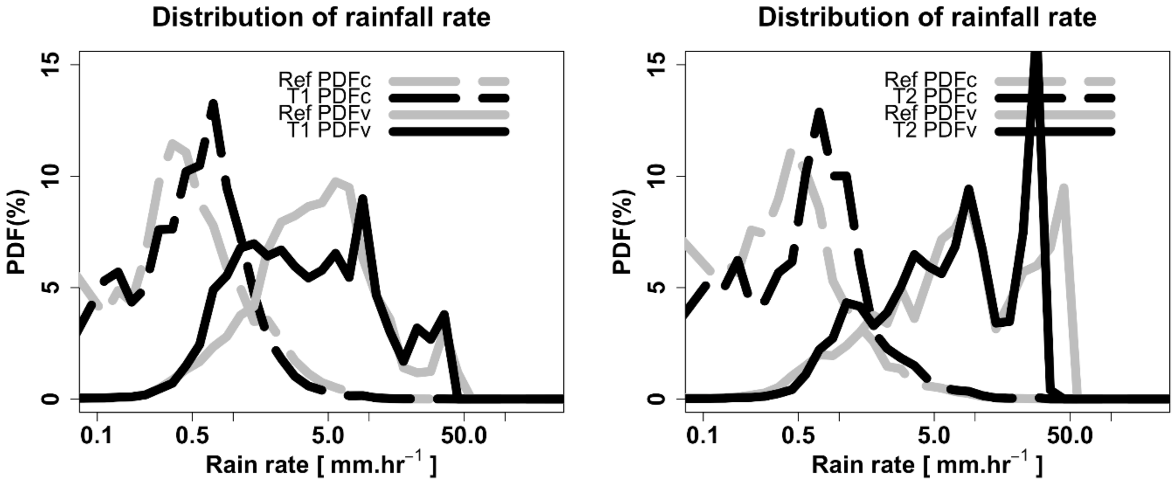


Figure 7. Probability density functions of precipitation occurrence (PDFc) and volume (PDFv) according to Tomorrow-R1 (T1; left) and Tomorrow-R2 (T2; right) relative to reference GV-MRMS (ref).

Figure 8 depicts density scatterplots for the Tomorrow.io radars relative to GV-MRMS, with the key systematic and random error metrics for the precipitation sample determined. Tomorrow-R2 outperforms Tomorrow-R1 in every metric of systematic and random errors, suggesting a greater skill in quantifying precipitation. The reason for the superior quantification performance of Tomorrow-R2 is unclear, given that the Tomorrow radars utilize the same neural network retrieval schemes trained on GPM combined radar-radiometer retrievals, and that both satellites have the same hardware design. An investigation based on reducing the Tomorrow-R1 61-day period to the same Tomorrow-R2 21-day period showed no significant difference in results.

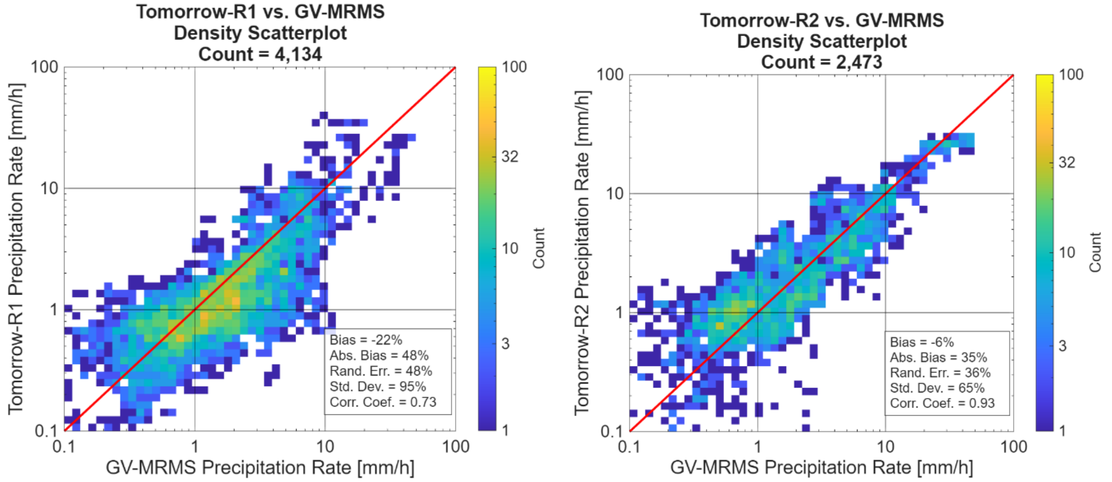


Figure 8. Density scatterplots of Tomorrow-R1 (left) and Tomorrow-R2 (right) relative to GV-MRMS.

Both of the Tomorrow.io radars tend to underestimate precipitation rates on average (R1: -22%; R2: -6% bias). Their random errors imply good consistency between the spaceborne and ground-based precipitation rates (R1: 48%; R2: 36% random error), though Tomorrow-R1 exhibits a small number of instances where radar precipitation rates can disagree with GV-MRMS by up to one order of magnitude. Tomorrow-R2 exhibits a very strong correlation (0.93) with GV-MRMS, far exceeding that of Tomorrow-R1 (0.73).

Figure 9 qualitatively showcases the systematic and random errors with the diagnostic percentile distributions for the Tomorrow.io radars with respect to GV-MRMS. The diagnostic percentile distribution considers a range of percentiles in the Tomorrow.io radar precipitation rate distribution as a function of GV-MRMS precipitation rates. It is apparent that the Tomorrow.io radars tend to overestimate light precipitation rates, whilst underestimating intermediate and heavy precipitation rates, with underestimates from Tomorrow-R2 being smaller than those from Tomorrow-R1.

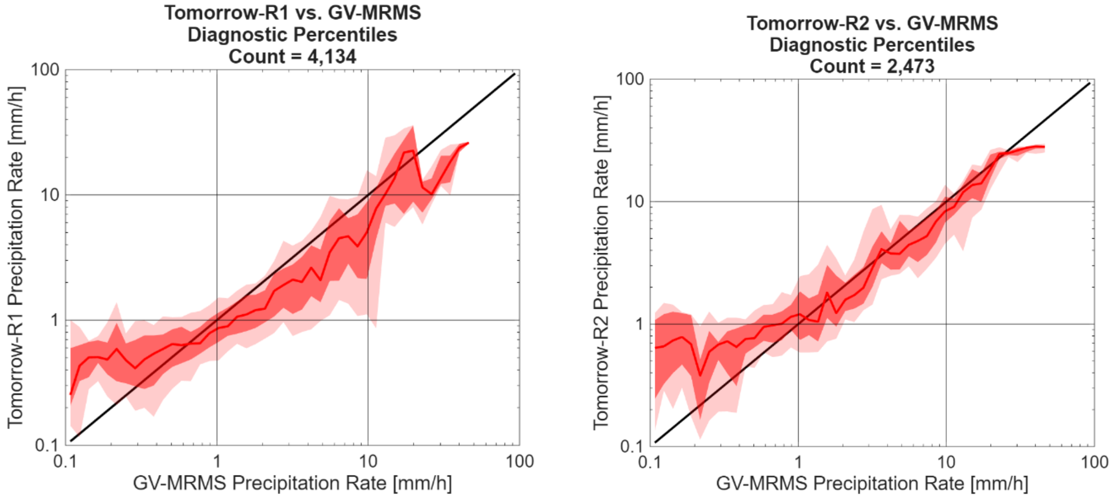


Figure 9. Diagnostic percentile distributions of Tomorrow-R1(left) and Tomorrow-R2 (right) relative to GV-MRMS. The light red shading represents the 10th-90th percentile range, the dark red shading represents the 25th-75th percentile range, and the red line represents the 50th percentile.

The Tomorrow.io radar errors are further quantified as a function of the precipitation spectrum in Figure 10, which showcases the mean systematic (Eq. 5) and random errors (Eq. 7) as functions of GV-MRMS precipitation rate. This shows that the random errors for the Tomorrow.io radars tend to decrease with increasing precipitation intensity; Tomorrow-R2 has greater random errors below 5 mm h^{-1}, whilst Tomorrow-R1 random errors are greater beyond 5 mm h^{-1}.

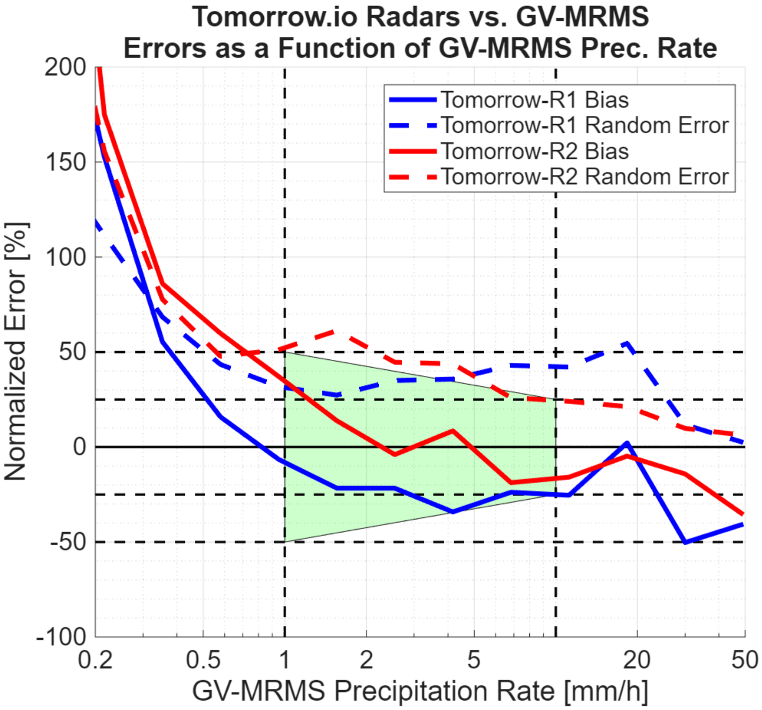


Figure 10. Systematic and random errors for the Tomorrow radars as a function of GV-MRMS precipitation rate at 5 km × 5 km resolution. The green shaded region highlights the range of errors and precipitation rates for which the GPM mission science requirements are satisfied at 50 km × 50 km resolution.

The GPM mission science requirements for precipitation rate retrievals at 50 km × 50 km scale (green region in Figure 10) require a bias and random error of <50% for 1 mm h^{-1} and <25% for 10 mm h^{-1} (Skofronick-Jackson et al., 2016). Figure 10 highlights the skill of the Tomorrow.io radar retrievals, such that their precipitation bias satisfies the GPM science requirements at a much finer resolution of 5 km × 5 km, whilst the random errors partially satisfy the requirements.

Figure 11 showcases the systematic and random errors for the Tomorrow.io radars as a function of the percentage of precipitation within the radar footprint. The Tomorrow.io radar bias and random errors tend to reduce as the radar footprint contains a high percentage of precipitation, with errors generally at their smallest when precipitation fills the footprint (100%). Errors across percentages are typically smaller for Tomorrow-R1 until the footprint is filled with precipitation; the better performance of Tomorrow-R1 suggests that variable precipitation fields within the radar footprint generally correspond to light precipitation rates of < 2 mm h^{-1}.

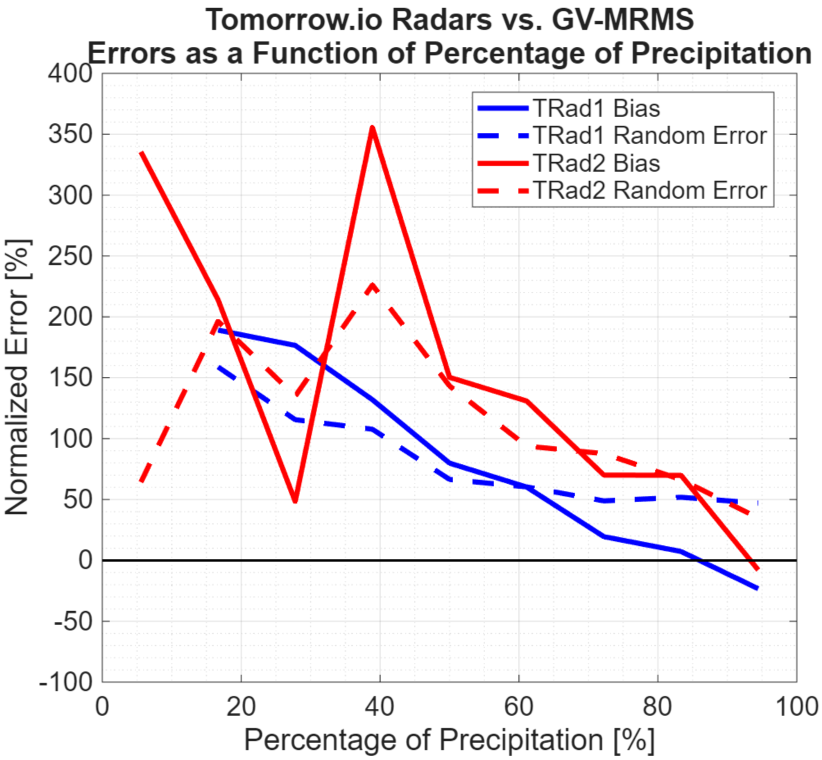


Figure 11. Systematic and random errors for the Tomorrow radars as a function of the percentage of precipitation with the 5 km Tomorrow footprint according to GV-MRMS.

5 Detailed Validation – Geometric

5.1 Validation Dataset, Method, and Completeness

The evaluation of the Tomorrow.io R1 and R2 pointing and positional accuracy quality was assessed globally using the 0.065-degree resolution GMTED2010 digital elevation model (https://www.temis.nl/data/gmted2010.html).

The evaluation utilized the TR1 and TR2 reported geolocation information in conjunction with view angle, satellite altitude, and estimated surface bin data. By using the known radar range bin, surface elevation was calculated and compared against surface elevation data from the DEM dataset.

5.2 Results Compliance

The radar-derived and DEM surface elevation estimates demonstrated very good agreement, with a correlation coefficient of around 0.98 for both TR1 and TR2 (Fig. 12). Occasional discrepancies exceeding several hundred meters (or even thousands of meters) were observed. These discrepancies are believed to be the result instrument rolls potentially associated with synchronization errors. Shown in Figure 13 is an example of observations with associated incidence angles for a case characterized by significant disagreement between the radar and DEM estimates, which are the result of a roll possibly associated with time-difference errors. While precise quantification of pointing and colocation errors remains challenging, the overall agreement suggests that the geolocation information is unbiased and accurate to within a few hundred meters, which meets the accuracy requirements for the radar system's intended applications.

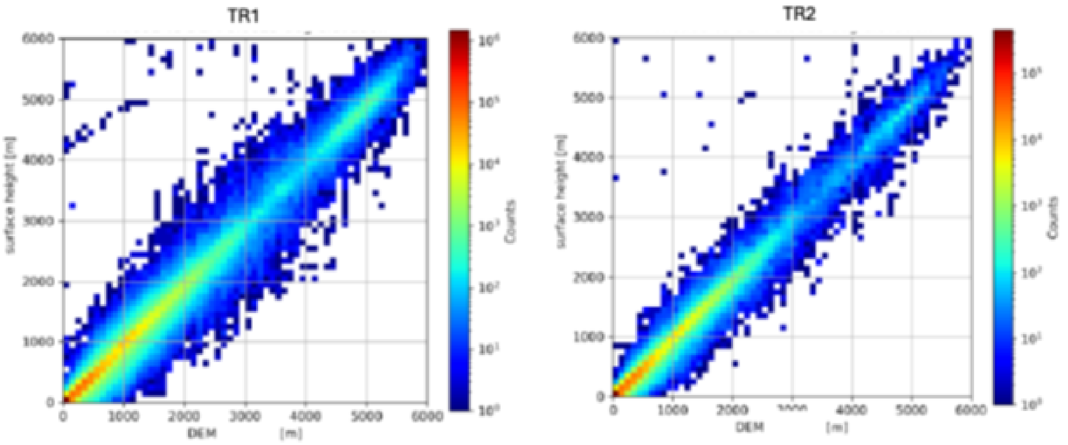


Figure 12. Joint distributions of surface height estimates from the Tomorrow.io radars and a digital elevation model (DEM).

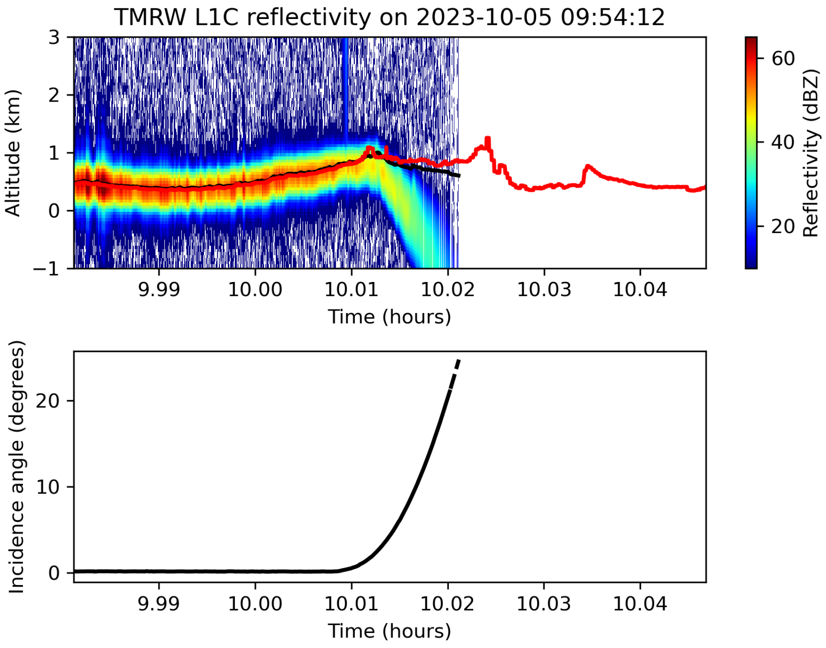
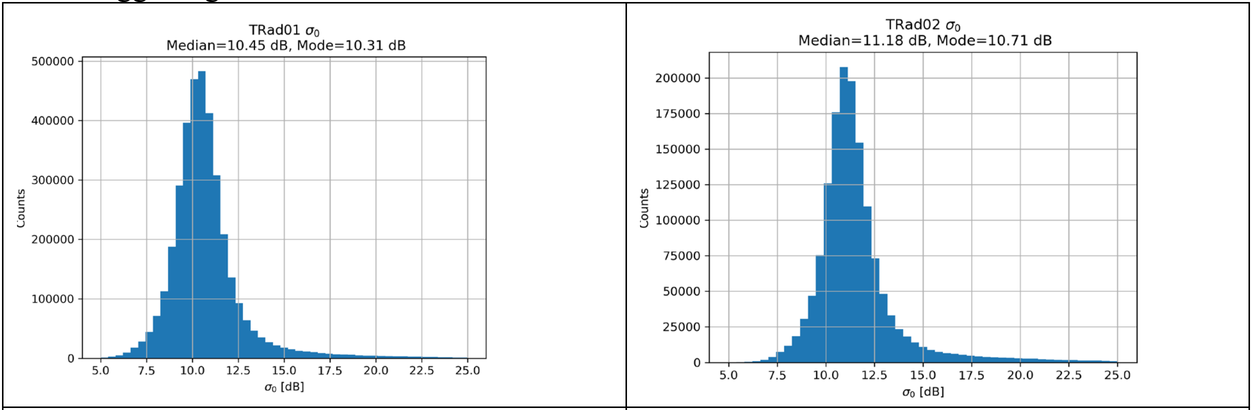


Figure 13. Example of TR1 observed reflectivity profiles (top). The black line indicates the radar estimated surface, while the red line is the surface from the DEM model. The incidence angle associated with the measurements shown in the bottom panel.

Surface Return Analysis

Surface return characteristics were analyzed under clear-sky conditions and statistically compared to those from the GPM Ka-band radar, which operates at the same frequency. Although the GPM Ka-band radar operates with different parameters (PRF, pulse width, etc.) and identical surface returns were not expected, the comparison showed consistency between systems (Figure 14). Both evaluated radars exhibited normalized radar cross section (NRCS) distributions with modes and medians within 0.5 dB of the GPM Ka-band precipitation radar (KaPR) surface return values, which suggests good calibration relative to the KaPR.



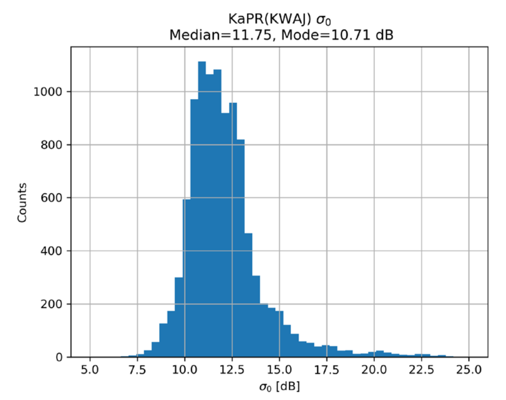


Figure 14. Distributions of sigma0 from the Tomorrow.io TR1 and TR2 systems (top row) and the GPM KaPR (bottom row). While there are sampling and signal processing differences between the Tomorrow.io radars and the GPM radar, the results are consistent with the results of (Yang et al. 2019), suggesting a mode of about 7 m/s in the surface winds and agreement in surface associated normalized radar cross sections (NRCS) measured by the two types of radars.

6 Conclusions

The conclusions of the assessment of the Tomorrow.io radar data are given below for precipitation and geolocation analyses.

Precipitation

- When assessing the precipitation detection capabilities of the Tomorrow.io radars with respect to GV-MRMS, the two radars show good performance in distinguishing between precipitation and no precipitation events over CONUS (POD ~ 0.82, CSI ~ 0.76).

- Tomorrow-R1 slightly outperforms Tomorrow-R2 with respect to precipitation detection ability, due to its enhanced capability to detect instances of no rain (HSS; R1 = 0.38; R2 = 0.21).
- When assessing the precipitation quantification capabilities of the Tomorrow.io radars with respect to GV-MRMS, the two radars show good performance in representing CONUS precipitation regimes, although they have a tendency to overestimate light precipitation rates and underestimate heavier precipitation rates (> 1 mm h^{-1}).
- Tomorrow-R2 outperforms Tomorrow-R1 when quantifying CONUS precipitation rates overall, with a smaller tendency to underestimate precipitation rates (bias; R1: -6%; R2: -22%) and a greater consistency in precipitation variability with the ground reference (random error and correlation; R1: 48% & 0.73; R2: 36% & 0.93). However, it should be noted that Tomorrow-R1 better quantifies light precipitation rates (< 1 mm h^{-1}).
- The reasons for the superior performance of Tomorrow-R2 are unclear, given that the radars use the same hardware and retrieval algorithm, and the difference when comparing the same sampling period shows no significant difference.

Geometric

- The evaluation demonstrates acceptable pointing and position accuracy for both radar systems, with the geolocation performance found to be suitable for the intended precipitation measurement applications.

7 References

Hou, A.Y., Kakar, R.K., Neeck, S., Azarbarzin, A.A., Kummerow, C.D., Kojima, M., Oki, R., Nakamura, K., Iguchi, T., 2014. The Global Precipitation Measurement Mission. Bulletin of the American Meteorological Society 95, 701–722. https://doi.org/10.1175/BAMS-D-13-00164.1

Kidd, C., Tan, J., Kirstetter, P.-E., Petersen, W.A., 2018. Validation of the Version 05 Level 2 precipitation products from the GPM Core Observatory and constellation satellite sensors. Quarterly Journal of the Royal Meteorological Society 144, 313–328. https://doi.org/10.1002/qj.3175

Kirstetter, P.-E., Hong, Y., Gourley, J.J., Cao, Q., Schwaller, M., Petersen, W., 2014. Research Framework to Bridge from the Global Precipitation Measurement Mission Core Satellite to the Constellation Sensors Using Ground-Radar-Based National Mosaic QPE, in: Remote Sensing of the Terrestrial Water Cycle. American Geophysical Union (AGU), pp. 61–79. https://doi.org/10.1002/9781118872086.ch4

Kirstetter, P.-E., Hong, Y., Gourley, J.J., Chen, S., Flamig, Z., Zhang, J., Schwaller, M., Petersen, W., Amitai, E., 2012. Toward a Framework for Systematic Error Modeling of Spaceborne Precipitation Radar with NOAA/NSSL Ground Radar–Based National Mosaic QPE. Journal of Hydrometeorology 13, 1285–1300. https://doi.org/10.1175/JHM-D-11-0139.1

Kirstetter, P.-E., Hong, Y., Gourley, J.J., Schwaller, M., Petersen, W., Cao, Q., 2015. Impact of sub-pixel rainfall variability on spaceborne precipitation estimation: evaluating the TRMM 2A25 product. Quarterly Journal of the Royal Meteorological Society 141, 953–966. https://doi.org/10.1002/qj.2416

Kirstetter, P.-E., Hong, Y., Gourley, J.J., Schwaller, M., Petersen, W., Zhang, J., 2013. Comparison of TRMM 2A25 Products, Version 6 and Version 7, with NOAA/NSSL Ground Radar–Based National Mosaic QPE. https://doi.org/10.1175/JHM-D-12-030.1

Kirstetter, P.-E., Petersen, W., Gourley, J.J., 2018. GPM Ground Validation Multi-Radar/Multi-Sensor (MRMS) Precipitation Reanalysis for Satellite Validation Product.

Mroz, K., Treserras, B.P., Battaglia, A., Kollias, P., Tatarevic, A., Tridon, F., 2023. Cloud and precipitation microphysical retrievals from the EarthCARE Cloud Profiling Radar: the C-CLD product. Atmospheric Measurement Techniques 16, 2865–2888. https://doi.org/10.5194/amt-16-2865-2023

Petersen, W.A., Kirstetter, P.-E., Wang, J., Wolff, D.B., Tokay, A., 2020. The GPM Ground Validation Program, in: Levizzani, V., Kidd, C., Kirschbaum, D.B., Kummerow, C.D., Nakamura, K., Turk, F.J. (Eds.), Satellite Precipitation Measurement: Volume 2. Springer International Publishing, Cham, pp. 471–502. https://doi.org/10.1007/978-3-030-35798-6_2

Simpson, J., Kummerow, C., Tao, W.-K., Adler, R.F., 1996. On the Tropical Rainfall Measuring Mission (TRMM). Meteorl. Atmos. Phys. 60, 19–36. https://doi.org/10.1007/BF01029783

Skofronick-Jackson, G., Huffman, G., Stocker, E., Petersen, W., 2016. Successes with the Global Precipitation Measurement (GPM) mission, in: 2016 IEEE International Geoscience and

Remote Sensing Symposium (IGARSS). Presented at the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 3910–3912. https://doi.org/10.1109/IGARSS.2016.7730015

Watters, D., Battaglia, A., Allan, R.P., 2021. The Diurnal Cycle of Precipitation according to Multiple Decades of Global Satellite Observations, Three CMIP6 Models, and the ECMWF Reanalysis. Journal of Climate 34, 5063–5080. https://doi.org/10.1175/JCLI-D-20-0966.1

Watters, D., Battaglia, A., Mroz, K., Tridon, F., 2018. Validation of the GPM Version-5 Surface Rainfall Products over Great Britain and Ireland. Journal of Hydrometeorology 19, 1617–1636. https://doi.org/10.1175/JHM-D-18-0051.1

Watters, D., Huffman, G., Gatlin, P., Kirstetter, P.-E., Bolvin, D., Joyce, R., Nelkin, E., Tan, J., Wolff, D., 2025. A Diagnosis of Oceanic Precipitation in IMERG-GMI. Journal of Hydrometeorology 1–21.

Watters, D.C., Gatlin, P.N., Bolvin, D.T., Huffman, G.J., Joyce, R., Kirstetter, P., Nelkin, E.J., Ringerud, S., Tan, J., Wang, J., Wolff, D., 2024. Oceanic Validation of IMERG-GMI Version 6 Precipitation Using the GPM Validation Network. Journal of Hydrometeorology 25, 125–142. https://doi.org/10.1175/JHM-D-23-0134.1

Yan, Q., Zhang, J., Fan, C. and Meng, J., 2019. Analysis of Ku-and Ka-band sea surface backscattering characteristics at low-incidence angles based on the GPM dual-frequency precipitation radar measurements. *Remote Sensing*, *11*(7), p.754.