Evaluation of Precipitation Products in Mid-Atlantic and Southern New England

Amalie Rebstock (amalier1@umbc.edu)[1], Ali Tokay [1,2]

[1]Geography & Environmental Systems, University of Maryland Baltimore County, Baltimore, MD

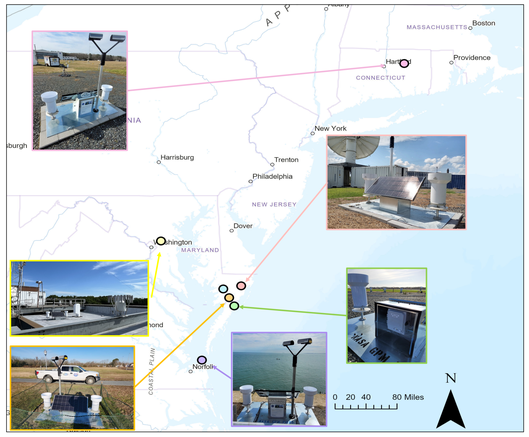
[2] Mesoscale Atmospheric Processes Laboratory, NASA, Goddard Space Flight Center, Greenbelt, MD

Motivation and Objective

Precipitation is a driving force for the water cycle and is one of the key sources of weather extremes. The changing global climate and consequent increase in extreme weather causes life to become more reliant on precipitation. The drought in Maryland is evident, as only 48 mm (1.9") of rainfall fell during the first 50 days of Summer 2024. This information is based on a gauge report at Baltimore airport, which hosts a complete weather station. Unfortunately, most of the global land is not equipped with precipitation measuring devices. This is due to topography, land use coverage, cost, and remoteness. With the addition of global ocean coverage (>70%), precipitation climatology relies on spaceborne precipitation retrievals and model outputs. Thus, there is a need to evaluate the accuracy of the global and regional precipitation products, a motivation of this study. NASA's Global Precipitation Measurement (GPM) ground validation program has been deploying Platforms for In-situ Estimation Rainfall Systems (PIERS) at granted institutes across the US. PIERS+ includes a laser optical PARSIVEL disdrometer and two tipping buckets. It is not part of any precipitation products, making it a resource to validate the precipitation products independently, a main objective of this study.

Introduction

Site Locations



Storrs, CT 41.809 -72.295	Newark, MD 38.262, -75.341
Greenbelt, MD 38.992, -76.840	Wallops, VA 37.9344, -75.470
Pocomoke, MD 38.067, -75.566	Chesapeake Bay Bridge-Tunnel, VA 37.036, -76.076

This study uses six (6) PIERS+ sites across Mid-Atlantic with and addition of a site in Storrs, CT. Three (3) of the sites are in Delmarva Peninsula and they are at distances of 17, 29 and 38 km to each other. The Greenbelt, MD and Chesapeake Bay Bridge-Tunnel sites are located to the northwest and south-southwest of triple Delmarva sites, respectively. Both sites are quite far distance, 151 and 113 km, from Delmarva sites but it is not uncommon a west-to-east progressing frontal system or northward moving off-shore low pressure could bring precipitation in all five sites during winter and early spring. Even, the same weather system may migrate northeastward and precipitate at the New England site. These Miller-A and Miller-B systems are replaced by relatively short-lived convective outbursts during late spring. This study covers the precipitation measurements for the first five (5) of 2024.

Ground Instrumentation



PARSIVEL: a laser-optical disdrometer, operates at near red edge of the visible spectrum (650 nm). The raw output is the 32 x 32 size vs. velocity matrix where the bin width ranges from 0.1 to 3.0 mm for particle size and 0.1 to 3.2 m s^{-1} for fall velocity. PARSIVEL cannot measure sub drizzle particles (< 0.2 mm in diameter) and its measurements are affected by secondary drops which are filtered out through independent fall velocity measurements. This filtering is not fully effective for small drops where the fall velocities are within terminal fall speed. The Chesapeake Bay Bridge-Tunnel site has high amounts of salt particles which interfered the PARSIVEL's ability to measure total rainfall.



Tipping Bucket (TB): an automated *rain* gauge which records the time of the tip (0.254 mm). The raw output can be set to number of tips per time interval (e.g. one-minute). It underestimates the heavy rainfall due to missing tips and windy conditions. It does not resolve light rain and subject to underestimation due to evaporation.

Precipitation Products

- Multi-Radar Multi-Sensor (MRMS): The NOAA's Multi-Radar Multi-Sensor (MRMS) product has been widely employed to validate satellite and model precipitation estimates among many other applications. MRMS has a spatial resolution of 0.01 degree (~1 km) and temporal resolution of 2 min, although a 30-minute resolution was used in this project.
- High Resolution Rapid Refresh (HRRR): **:** HRRR is a model that covers the continental United States and Alaska. HRRR was created by the National Oceanic and Atmospheric Association (NOAA)/National Centers for Environmental Prediction (NCEP). HRRR has a spatial resolution of 3 km and temporal resolution of 1 hour.
- European Centre for Medium-Range Weather Forecasts Reanalysis v5 (ERA5): ERA5 is produced by the Copernicus Climate Change Service (C3S) at ECMWF. ERA5 has a spatial resolution of 25 km and temporal resolution of 1 hour.
- Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA2): MERRA2 is produced by NASA's Global Modeling and Assimilation Office (GMAO) at NASA Goddard Space Flight Center. MERRA2 has a spatial resolution of 0.5˚ x 0.625˚ (~50 km) and a temporal resolution of 1 hour.

Methodology

Rain Events

A rain event is defined based on three (3) or longer hours of rain-free period in PARSIVEL rain rate time series. In addition, the rain must have a minimum of 2 mm accumulation. The second instrument or product must also have 1 mm accumulation indicating that both instruments observed precipitation. PARSIVEL precipitation phase algorithm is used to disregard the events that included frozen or mixed precipitation. It is feasible that one of the instruments or products fails to record precipitation. This was noticed in short (< 2 hours) rain events in late Spring. For example, PARSIVEL at Greenbelt, MD recorded 44 mm rainfall in 47 minutes where maximum one-minute rain rate was 186.2 mm h^{-1}. The collocated gauges had 5-12% less totals showing the importance of dual or multiple gauges at a given site. The MRMS underestimated this event by 30%, while HRRR recorded only 25% of the rainfall. ERA5 and MERRA2 had no rainfall for this event. It is also feasible that the products could report false rainfall. The precipitation versus no-precipitation is the and subsequent precipitation phase are the two fundamental entity prior to precipitation amount. Both topics are beyond the scope of this study.

Rainfall Statistics

$$\text{Correlation Coefficient (CC)} = \frac{\Sigma(\text{RT ref1}-\text{avg(RT ref1)})(\text{RT ref2} - \text{avg(RT ref2)})}{\sqrt{(\Sigma(\text{RT ref 1}- \text{avg(RT ref1)})^2(\text{RT ref2}- \text{avg(RT } ref2))^2}}$$

$$\text{Correlation Coefficient (CC)} = \frac{\Sigma(\text{RT ref}-\text{avg(RT ref)})(\text{RT est} - \text{avg(RT est)})}{\sqrt{(\Sigma(\text{RT ref} - \text{avg(RT ref)})^2(\text{RT est}- \text{avg(RT est)})^2}}$$

$$\text{Bias} = \frac{\Sigma\ (\text{RT ref 1}- \text{RT } ref2)}{\frac{\Sigma\ \text{RT ref1} + \text{RT } ref2}{2}}$$

$$\text{Bias} = \frac{\Sigma(\text{RT est} - \text{RT } ref)}{\Sigma \text{RT } ref}$$

$$\text{Mean Absolute Error (MAE)} = \frac{\Sigma\ |(\text{RT ref1} - \text{RT ref2})|}{\frac{\Sigma\ \text{RT ref1} + \text{RT } ref2}{2}}$$

$$\text{Mean Absolute Error (MAE)} = \frac{\Sigma\ |(\text{RT est} - \text{RT ref})|}{\Sigma\ \text{RT } ref}$$

The Pearson correlation coefficient is widely used to determine the accuracy of an estimated (est) variable with respect to reference (ref) variable. In this study, the variable is the event rain total. The correlation between the two collocated precipitation measuring instruments is labeled as both references, while the correlation between the instrument and product is labeled as the reference and estimate. The formulation of bias and mean absolute error (MAE) therefore differ depending on the source of variables. If two instruments are used, then the normalization is the average value of the two variables, while the normalization is based on the reference instrument if the second variable is based on a product.

The correlation and bias themselves cannot resolve the outcome of the agreement between the two variables. The combination of high correlation and low bias is the signature of a good agreement. The MAE, on the other hand, could resolve the outcome of the agreement itself. When two collocated rain measuring instruments are compared, the following MAE criterion in % could be used: 0-5: Excellent; 6-10: Very Good; 11-15: Good; 16-20: Reasonable; >20: Poor (Tokay et al. 2010, JHM). When an instrument and product are compared, the MAE criterion in % may follow: 0-10: Excellent; 11-20: Very Good; 21-30: Good; 31-40: Reasonable; >40: Poor. The boundaries of these limits are subjective and open to discussion.

Acknowledgements

Funding for this project was received from the Maryland Space Grant Consortium. Special thanks to Jason Pippitt of SSAI/NASA-GSFC for providing the HRRR database and Mario F. L. de Quardo of Federal University of Santa Catarina, Brazil. Workplace provided by NASA Mesoscale Processing Laboratory, George Hoffman, Chief. Ground-based database was collected and provided by the NASA GPM ground validation office and the University of Connecticut led by Alexey Chibisov of McCallie Associates. Discussions with Charles N. Helms of ESSIC/NASA-GSFC and bi-weekly meetings with David Wolff of NASA-GSFC-WFF and Diego Cerrai of University of Connecticut were helpful. Acknowledgements extend to my co-interns, Connor Mahone and Sloane Poppei for supporting and providing feedback through my research.

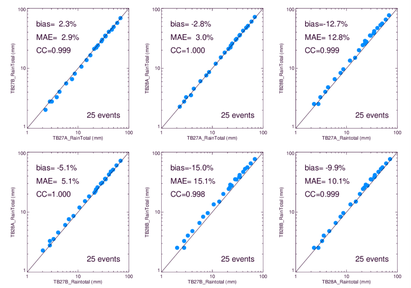
Summary & Future Research

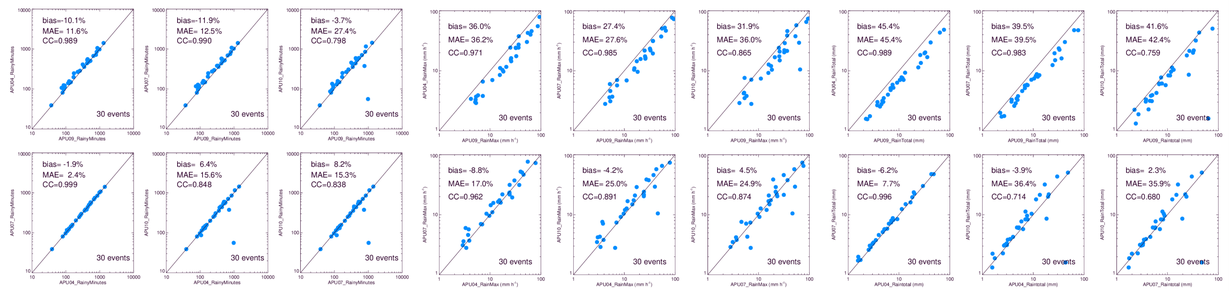
This study first questioned the reliability and accuracy of the ground-based reference instruments. The precipitation products, MRMS, HRRR, ERA5, and MERRA2, were then evaluated following PARSIVEL event definition and tipping bucket event rain total. The study is confined to the events that have been recorded by the products. The events that were missed by the products and falsely recorded were not included. The number of events (> 15) was sufficient to derive the meaningful statists but did not allow to subdivide the events based on rain intensity, synoptic settings. Future research consists of expanding the observation period to a year or longer and including IMERG and model outputs to the comparative study.

Results

Rain Gauge: A reference instrument for event rain total. Comparison of rain totals from four collocated tipping bucket gauges at Storrs, CT revealed that three units had an excellent agreement while the fourth unit (TB28B) overestimated rainfall for almost all rain events. The overestimation is attributed to the gauge calibration assuming there was no leveling issue during installation. These findings show the uncertainty in gauge measurements, confirming previous recommendations (Tokay et al. 2010, JHM): Two gauge is the minimum and three gauge is ideal for determining event rain total accurately.

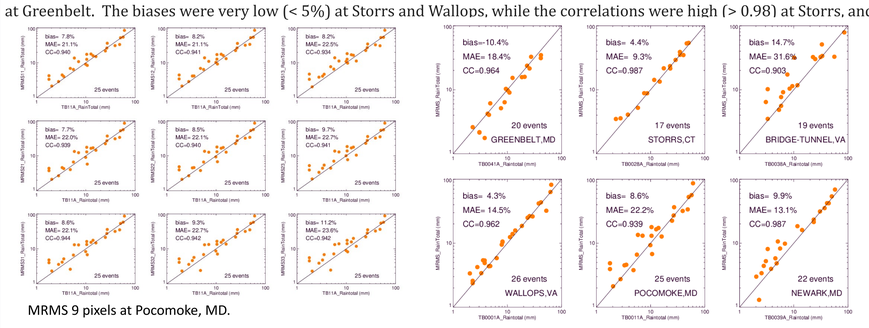
PARSIVEL: A reference instrument for defining rain events, rainy minutes, event maximum rain rate. It is a complementary instrument for event rain total. Comparison of a) rainy minutes, b) maximum rain rate, and c) rain total between four collocated Autonomous PARSIVEL Units (APU) at Wallops, VA revealed that the two APU (04 and 07) had the best agreement. The APU09 underestimated rainy minutes but overestimated maximum rain rate and rain total. The APU10 had two outlier events which prevented it to join the best agreement with the other two units. The outlier events resulted in lower correlations and higher biases.





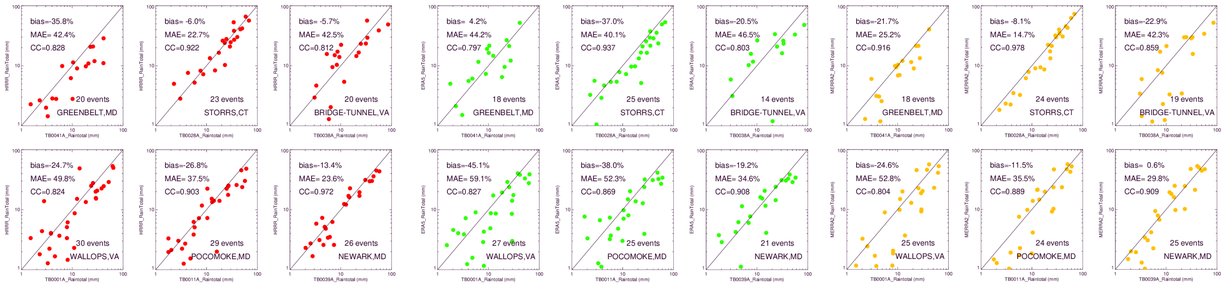
MRMS: The high spatial resolution of MRMS prompted us to examine its performance not only for the pixel where the gauge is located but also for the eight adjacent pixels. The difference in MAE was less than 3% at all sites with exception of the upper left and upper middle pixels in Greenbelt, MD and Wallops, VA. Here, the nine-pixel layout of the performance of MRMS is presented for Pocomoke, MD. The middle-pixel represents the MRMS performance well.

MRMS performance ranged from excellent in Storrs, CT to reasonable at Chesapeake Bay Bridge-Tunnel, VA. The high performance at Storrs is attributed to the gauge adjustment since the site is far (97 km) from the closest operational radar in Boston, MA (KBOX). The relatively low performance at Bridge-Tunnel is attributed to the remoteness of the site and far (96 km) from the closest radar in Wakefield, VA (KAKQ). The performance of MRMS was similar between Wallops, VA and Newark, MD. The reasoning, however, was attributed to the nearly adjacent (0.5 km) gauge and close (26 km) to Dover, DE (KDOX) radar, respectively. MRMS overestimated rainfall at all sites except Greenbelt, MD. The overestimation was significant for several events where the reference gauge had rainfall less than 10 mm. This is also true for the underestimation at Greenbelt. The biases were very low (< 5%) at Storrs and Wallops, while the correlations were high (> 0.98) at Storrs, and Newark.



MRMS 9 pixels at Pocomoke, MD.

Site	Gauge & Radar Distance
Greenbelt, MD	KDEA (23 km) KLWX (59 km)
Storrs, CT	KIJD (12 km) KBOX (97 km)
Chesapeake Bay Bridge-Tunnel, VA	CRN - CC (26 km) KAKQ (96 km)
Wallops, VA	KWAL (0.5 km) KDOX (65 km)
Pocomoke, MD	KWAL (17 km) KDOX (54 km)
Newark, MD	KOCX (20 km) KDOX (26 km)



HRRR: HRRR underestimated total rainfall in comparison with reference gauge at each site. There is a wide variety of HRRR performance from one site to another. The combination of low bias and high correlation was evident in Storrs, CT. The MAE was also the lowest meeting *good* agreement conditions at the same site. The good agreement has also been noticed at Newark, MD where correlation and bias were higher than New England site. The site at Pocomoke, MD satisfied reasonable agreement with higher bias and lower correlation than the Newark and Storrs sites. The remaining three sites had correlations under 0.9 and MAE above 40%. HRRR in these sites were marked as poor performance even though low bias was evident in Chesapeake Bay Bridge-Tunnel, VA. The scatter diagrams also showed several low rain total events where HRRR's underestimation is severe. These events could be related to the small-scale convective rainfall. This study is subject to extend the longer period of observations where performance of HRRR can be evaluated for a season, synoptic setting.

ERA5: ERA5 had a poor performance (MAE > 40%) for all sites except a reasonable (MAE of 35%) performance at Newark, MD. It underestimated rainfall (bias < 0%) for all sites except at Greenbelt, MD. The negative bias was evident for the majority of the events across the rainfall spectrum at three sites (Pocomoke, Storrs, and Wallops). Nearly half of the events had a positive bias at Newark and Chesapeake Bridge-Tunnel even though overall bias was approximately -20% at these two sites. This is due to the high negative bias of a few events, which have also been noticed at three other sites. The correlations were mostly low (CC < 0.9) due to the large scatter around 1:1 line.

MERRA2: MERRA2 had a wide variety of performance ranging from very good (MAE of 15%) at Storrs, CT to poor at Chesapeake Bridge-Tunnel, VA and Wallops, VA (MAE > 40%). It underestimated rainfall (bias < 0%) for all sites except at Newark, MD where bias was near 0%. Correlations were also wide ranging from high (CC of 0.98) at Storrs to very low (CC of 0.8) at Wallops, VA. Despite coarser spatial resolution, MERRA2 had better performance than ERA5 and HRRR at several sites.