Evaluating Model Precipitation Products for Point-Based Locations: How Much Is It Going To Rain Tonight?

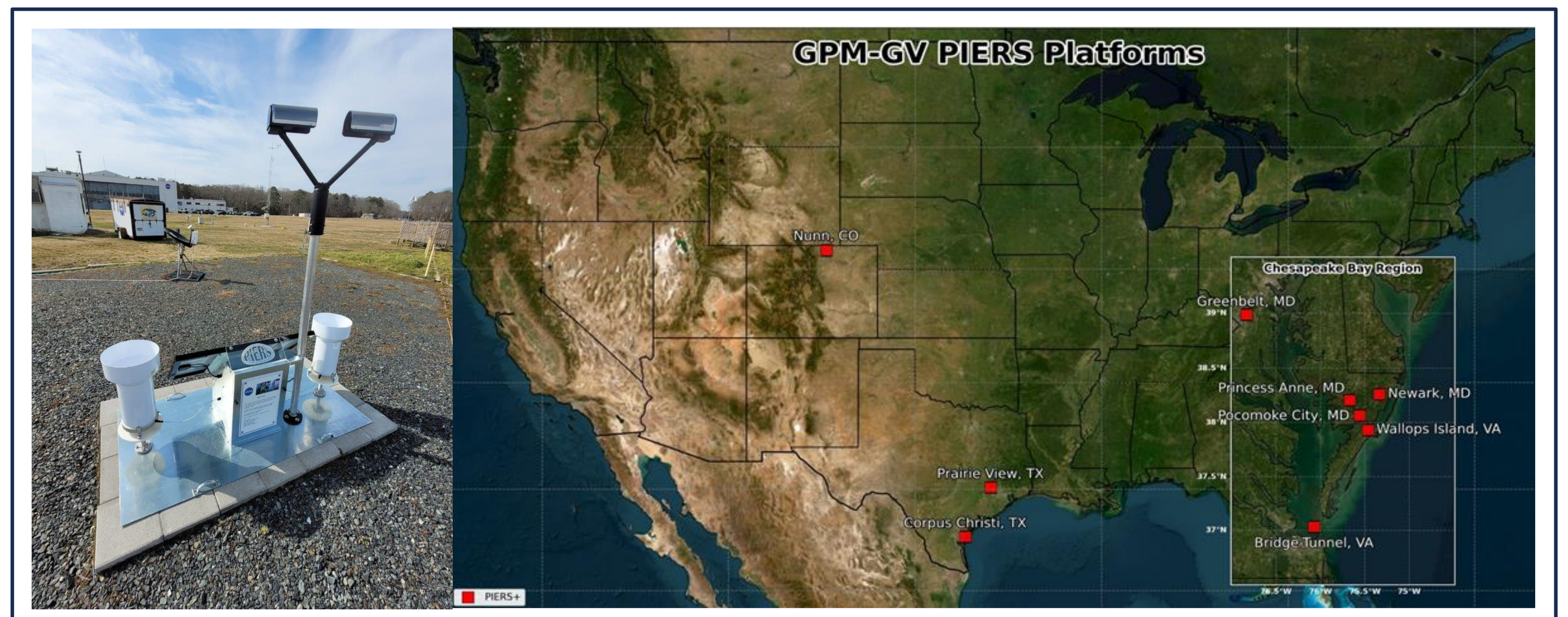
Alex Zielinski[1], Ali Tokay[2,3]

[1]University of Illinois at Urbana-Champaign, [2]UMBC, GESTARII-Geography and Environmental Systems (GES), [3]NASA Goddard Space Flight Center (GSFC)

Motivation and Objective

Precipitation is one of the two main components of the water cycle and global climate. The changing global climate has led to an increase in extreme weather, which includes high precipitation events. In July 2025, in just a few hours, the Texas Hill Country received over 51 cm of rainfall, surging the Guadalupe River to nearly 9 meters and causing 139 total deaths. Weather forecasting models performed poorly for predicting this extreme rainfall, giving residents very little time to prepare. It is imperative to validate the models against ground truth, which itself has its own shortcomings. Ground-based in-situ observations are known as the gold standard, but their spatial distribution and record length are heterogeneous. As an alternative, the precipitation products such as NOAA NSSL's Multi-Radar Muti-Sensor (MRMS) and NOAA RFC's Stage IV are heavily used as a ground truth. For point-based rainfall, known as the farmer's approach, these gauge-adjusted radar products inherit their own errors, missing short and light rain events and underestimating heavy rain events. Starting in late 2023, NASA's Global Precipitation Measurement (GPM) Ground Validation (GV) program has been deploying Platforms for In situ Estimation Rainfall Systems (PIERS) at granted institutes across the US. PIERS+ includes a PARSIVEL disdrometer and two tipping bucket gauges. This study evaluates four weather forecasting models using nine (9) PIERS+ sites, which are primarily in the Mid-Atlantic region, with two in Texas and one in Colorado at higher elevation. This poster will only focus on the HRRR model, but three other models were evaluated as well (NAM, MPAS-HN, WRF-NSSL). The study focuses on event rainfall totals for calendar year 2024. This aspect differs from the previous studies, which mainly focused on fixed time intervals (e.g. hourly, daily, monthly). It is well known that precipitation is highly variable in both spatial and temporal scale and does not necessarily end at the end of hour, day or even month. The fixed temporal scale past studies therefore inherit additional uncertainty.

Instrumentation



There are currently nine (9) PIERS+ sites that have continuous data records across the US. This study utilizes these sites to evaluate the forecasting models at six (6) sites (US sites) where weather systems and local conditions differ from each other. The site at Nunn, CO represents higher elevation, while the sites at Corpus Christi, TX and Chesapeake Bridge-Tunnel, VA are coastal. The site at Prairie View, TX is subject to mesoscale convective complex, while nor'easters primarily influence Bridge-Tunnel, VA and Newark, MD, It is not uncommon that the summertime convective storms die over Chesapeake Bay after passing Greenbelt, MD but not reaching Newark, MD. This study also focuses on *regional* evaluation of forecasting models by incorporating four (4) sites at Delmarva Peninsula (Delmarva sites). This contributes to the previous regional studies of evaluation of precipitation products (e.g. Major et al. 2026).

PARSIVEL (PARticle Size VELocity) is a laser-optical disdrometer that measures the size and fall velocity of individual hydrometeors and outputs in a 32 x 32 size versus fall velocity matrix. It underestimates the small raindrops (< 1 mm in diameter) and cannot distinguish the size of very large drops (> 5 mm in diameter). It also underestimates the fall velocity of mid-size drops (1-3 mm in diameter). This study uses OTT PARSIVEL as a reference instrument for event rainy minutes and event definition. A 12" orifice MetOne tipping bucket gauge which records time of each tip (0.254 mm) is the reference for event rain totals. In comparison to rain gauge, PARSIVEL provides accurate rain rate time series at one-minute resolution. PARSIVEL is a robust instrument requiring minimal maintenance. The tipping bucket gauge often fails for various reasons, and it is suggested that the dual tipping bucket gauge is minimum for long term (> 1 year) operation. One caveat of PIERS+ is the lack of power for heating PARSIVEL. This causes fogging on PARSIVEL lenses, an underestimation of rain intensity as observed in Corpus Christi, TX in winter. A second battery unit is now added for the future PIERS+ units.

Methodology

A rain event is defined for a) ≥ 3 hours no precipitation following PARSIVEL rain rate time series b) > 1.5mm (6 tips) of rain recorded by the tipping bucket. Rainfall statistics are run once for all events, and once for events where the model rainfall ≥ 1.2 mm. The second statistics filter out events where the model missed the event. The statistical package includes Normalized bias (Nbias), Normalized Mean Absolute Error (NMAE), and Pearson correlation coefficient (R). The Nbias measures the under or over-estimation of model rainfall, but its low value alone is not indicative of the performance of the model. A high performance of a model requires low Nbias and high correlation. As a single statistics, the NMAE or its alternative Normalized root mean square error is the measure of the performance of the model.

$$\text{Nbias} = \frac{\Sigma(M-O)}{\Sigma(O)} \qquad \text{NMAE} = \frac{\Sigma|M-O|}{\Sigma(O)} \qquad \text{R} = \frac{\Sigma(M-avg(M))(O_i-avg(O))}{\sqrt{\Sigma(M-avg(M))^2} * \sqrt{\Sigma(O-avg(O))^2}}$$

M: Model value O: Observed value

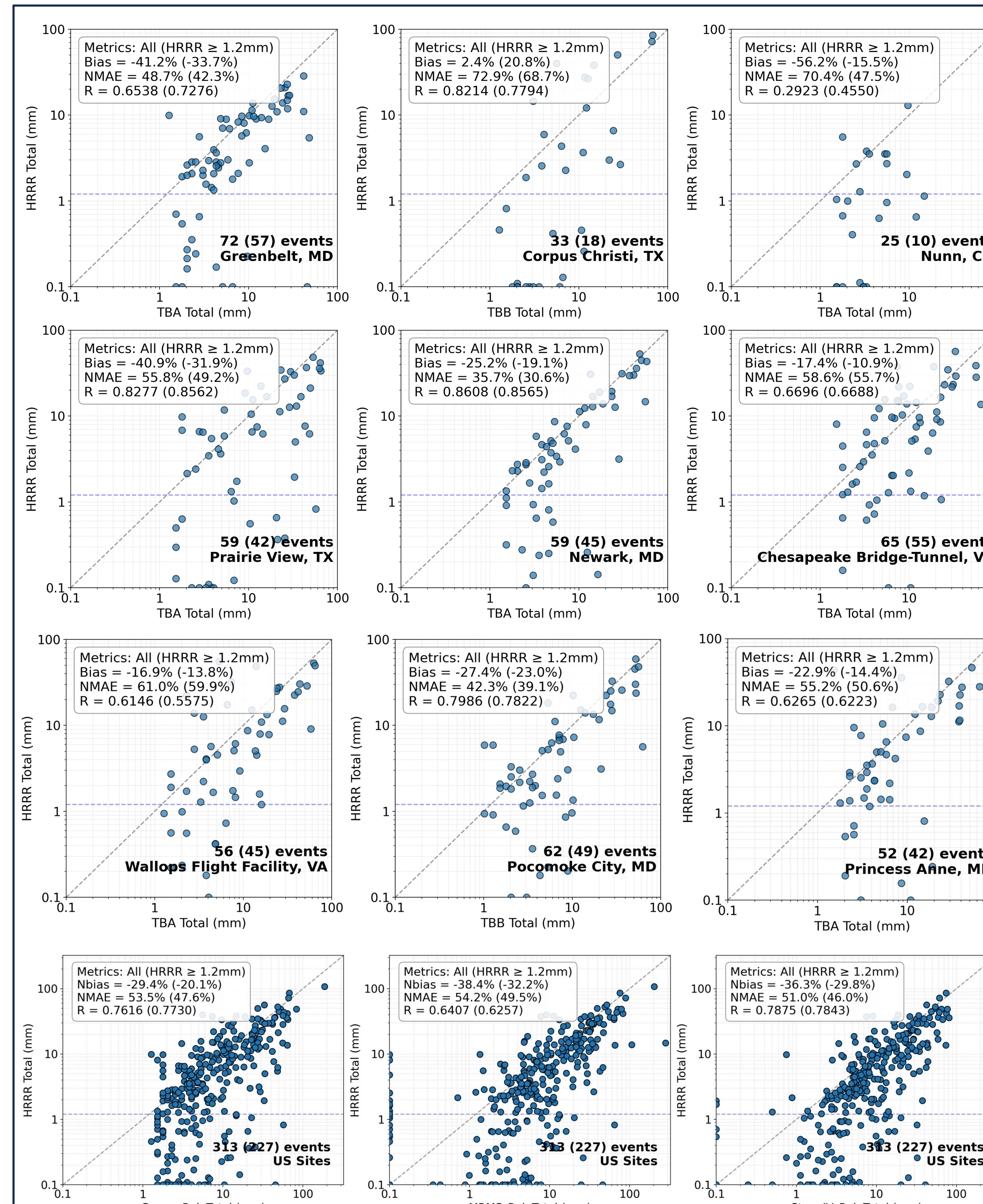
Conclusions & Future Research

Overall, HRRR, which is initialized the most frequently (every hour), outperformed NAM, MPAS, and WRF. HRRR missed 15% less events and had 19-24% less rainfall error than the other models. The events with longer duration (>5 hours), higher rainfall (> 10 mm) and moderate intensity (2-5 mm h^{-1}) had less rainfall error regardless of the models. The findings of this study are limited to the observational sites and the length of the dataset. Future work will include evaluating the newly operational MPAS-based Rapid Refresh Forecast System (RRFS) when the data becomes available and blends of models.

Contact: az22@illinois.edu

This work was supported by the **National Science Foundation**, Grant Number 2349543, Project Title: ***REU Site: EXPeriments in Earth and Atmospheric Science: Learning Opportunities and Research Experience (EXPLORE).***

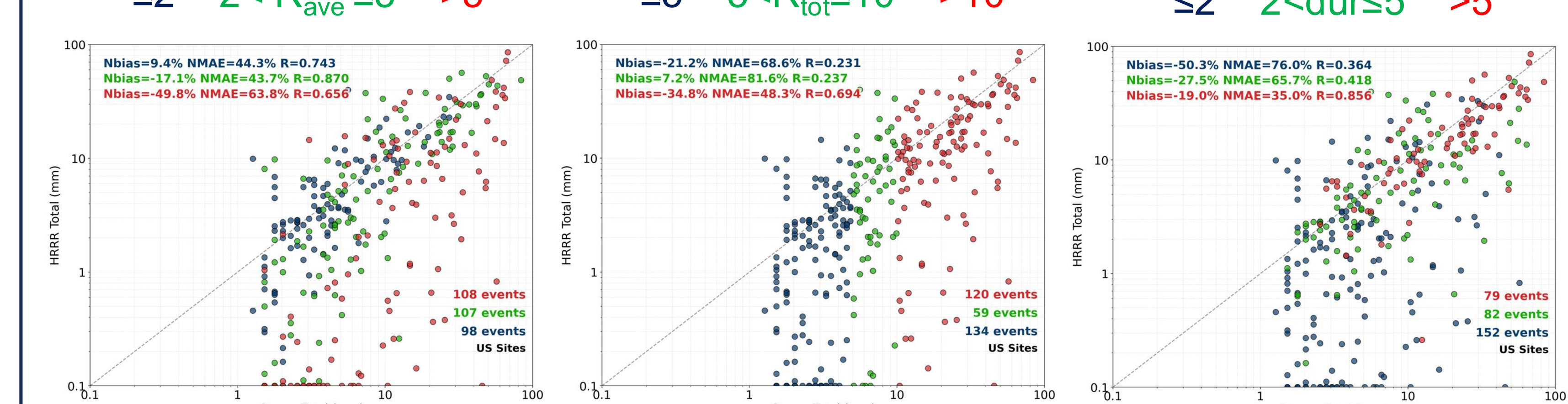
Results



Mean Rain Rate (mm/hr) ≤2 2< R_{ave} ≤5 >5

Total Rainfall (mm) ≤5 5<R_{tot}≤10 >10

Event Duration (hours) ≤2 2<dur≤5 >5



High-Resolution Rapid Refresh (HRRR): The NOAA/NCEP HRRR outputs hourly precipitation at 3km resolution over conterminous US. HRRR is initialized every hour assimilating operational radar data every 15-minutes (Benjamin et al. 2016 DOI: 10.1175/MWR-D-15-0242.1). This study showed that HRRR has a dry bias even after excluding missing rain events. HRRR event rainfall has 53% rainfall error with respect to rain gauges, only 1% lower and 2% higher than when MRMS and Stage IV are taken as reference, respectively.