Validating NASA's Remotely Sensed Rainfall Analysis with Insitu Florida Automated Weather Network observations

Cameron Mathis

Department of Earth, Ocean and Atmospheric Science
Florida State University
Prospectus for Honors in the Major (Meteorology)

Committee members:

Dr. V. Misra (Chair)

Dr. Alyssa Atwood

Dr. Harrison Prosper

Table of Contents

Abstract

This study evaluates the performance of the Integrated Multi-satellite Retrievals for GPM version 7 (IMERGv7) precipitation product against high-resolution ground-based observations from the Florida Automated Weather Network (FAWN). Over a 16-year period (2005–2020), precipitation data from eight strategically selected FAWN stations across northern and southern Florida were used to verify IMERG's gridded satellite estimates. Station coordinates were utilized to spatially align FAWN observations with IMERG grid cells, ensuring robust comparison at multiple temporal scales.

Ensemble Empirical Mode Decomposition (EEMD) was employed to decompose the precipitation datasets into distinct temporal components, including interannual, subseasonal, synoptic, and diurnal scales. Comparative analysis across these timescales reveals that IMERG systematically underestimates rainfall variance, particularly at synoptic (2–10 days) and subseasonal (20–90 days) scales, with these discrepancies becoming more pronounced during the convectively active spring, summer, and fall seasons. Interannual variability shows mixed performance with notable spatial inconsistencies. The greatest shortcomings emerge at the diurnal scale, where IMERG significantly underestimates the amplitude of daily rainfall fluctuations and consistently exhibits a temporal phase lag, particularly evident in the summertime, with satellite-derived precipitation peaks occurring later in the day compared to observed ground-based maxima.

These results suggest that while IMERG provides valuable insight into long-term climatic trends, it is limited in accurately capturing localized, short-duration convective events characteristic of subtropical climates like Florida's. Consequently, high-resolution ground-based monitoring networks such as FAWN remain indispensable for accurate validation and enhancement of satellite-derived precipitation products.

1 Introduction

Accurate precipitation estimation is essential for effective hydrologic forecasting, agricultural management, climate monitoring, and water resource planning. Satellite-based rainfall products, such as NASA's Integrated Multi-satellite Retrievals for Global Precipitation Measurement (IMERG), provide near-global coverage with high spatial (~0.1-degree) and temporal (30-minute) resolution, making them invaluable, particularly in data-sparse regions (Tan et al., 2019). IMERG integrates observations from multiple satellite microwave sensors and is calibrated against gauge observations to deliver continuous rainfall estimates (Huffman et al. 2023, 2024). Despite their utility, satellite products like IMERG rely on indirect measurements and are vulnerable to errors arising from temporal sampling gaps, misinterpretations of cloud microphysics, and algorithmic interpolation challenges (Tan et al., 2019).

Florida is a medium-sized U.S. state, ranking 22nd in area. However, despite its moderate size, Florida receives the greatest total volume of rainfall among states in the continental United States (Misra et al., 2017). This distinctive rainfall regime is largely attributed to Florida's peninsular geography and its tropical latitude location, which create pronounced seasonality in precipitation. The state's rainfall variability is particularly evident at diurnal scales, driven by strong day-night temperature contrasts typical of marine environments at low latitudes. These temperature differences induce density variations, producing sea breeze circulations and associated convective rainfall events that dominate Florida's hydrological cycle (Bastola and Misra, 2013).

Seasonal changes in Florida's hydroclimate are significantly influenced by the North Atlantic Subtropical High (NASH), also known as the Azores or Bermuda High. This high-pressure anticyclone migrates westward over the subtropical Atlantic basin during summer, then splits into

two separate high-pressure systems over eastern North America and northwestern Africa during winter. Movements of the NASH influence large-scale moisture advection, low-level divergence, and atmospheric stability, thereby impacting Florida's seasonal precipitation patterns (Li et al., 2013; Selman et al., 2013; Misra et al., 2011; Chan and Misra, 2010). Furthermore, Misra et al., 2018) show that the evolution of Florida's rainy season coincides closely with the warming of the Gulf of Mexico, the dynamics of surface ocean currents such as the Loop Current, and seasonal weakening of Atlantic trade winds. Thus, Florida's rainy season is inherently a coupled of ocean-atmosphere phenomenon, evolving gradually within both oceanic and atmospheric components.

Broader climatic phenomena further modulate Florida's hydroclimate variability. Variations in the El Niño-Southern Oscillation (ENSO) strongly affect Florida's rainfall, primarily during the boreal winter season (Ropelewski and Halpert, 1986, 1987). In the summer months, interannual variations in rainfall are linked, to some extent, to fluctuations in the Atlantic Warm Pool (Misra and DiNapoli, 2013). Even during winter, rainfall variability remains crucial for managing freshwater resources during Florida's subsequent dry spring season (Misra et al., 2023).

Given the state's complex rainfall dynamics across multiple temporal scales—interannual, subseasonal, synoptic, and particularly diurnal, Florida provides a natural and robust platform for validating satellite-derived rainfall datasets. This validation is crucial because ground-based, in situ rain gauge data at high temporal resolutions, especially at diurnal scales, are rarely available in most regions. The Florida Automated Weather Network (FAWN), operated by the University of Florida, fills this gap by providing high-frequency, sub-daily in situ meteorological observations across the state (Julie et al., 2023). Recent methodological enhancements to FAWN's data

processing have resulted in a reliable, gap-free observational dataset from 2005 to 2020, ideal for satellite validation.

In this context, this study aims to rigorously validate the IMERGv7 rainfall product against high-resolution FAWN observations from eight strategically selected stations across northern and southern Florida for the period 2005–2020. Using Ensemble Empirical Mode Decomposition (EEMD; Wu and Huang 2009), rainfall variability is separated into distinct temporal components, facilitating detailed comparisons at interannual, subseasonal, synoptic, and diurnal scales. This validation exercise not only identifies systematic biases and temporal limitations within IMERG's rainfall estimates—particularly during convectively active summer months—but also provides practical insights for researchers and practitioners relying on satellite precipitation products. Ultimately, the findings will enhance confidence in using NASA's IMERG rainfall dataset for hydroclimatic research, forecasting, and water management applications in tropical and subtropical environments.

2 Data and Methods

The NASA's Integrated Multi-Satellite Retrievals for Global Precipitation Mission version 7 (IMERG) dataset provides precipitation data with a 0.1° grid resolution at half hourly intervals (Huffman et al. 2023, 2024). This data is available from 2001 to the present. It includes early, late, and final run products with latencies of about 4 hours, 12 hours, and 3.5 months, respectively (Misra et al. 2022). The early run uses far less data processing than the late run, which requires additional microwave data from satellite overpasses. The final run also includes data from atmospheric reanalysis (including MERRA2 and ERA-5) and rain gauge analyses. Studies show that the IMERG dataset, particularly the 3.5-month final run is often superior to its predecessor, TRMM (Tang et al. 2016; Xu et al. 2019). This research focuses on assessing the fidelity of the final run of IMERG over Florida.

The Florida automated weather network (FAWN) is a local Florida dataset (Lusher et al. 2008). The FAWN dataset covers thirty stations across Florida (Lusher et al. 2008; Peeling et al. 2023). The FAWN dataset also has a higher temporal resolution at 15 minutes interval. In contrast the IMERG dataset is at 30-minute interval. Peeling et al. (2023) filled temporal gaps in the FAWN dataset to provide an unprecedented gap-free data set spanning from 2005 to 2020. The 30 FAWN stations cover diverse land types over Florida: 16 in forested areas in the north, 9 in savanna regions in the south, 4 in urban areas, and 1 in cropland. FAWN applies initial quality control and filtering to its data, with annual tests for sensor repair or replacement based on EPA guidelines. Gaps in data arise from equipment issues or power failures. Additional quality control measures were implemented such as flagging temperature differences of over 5°C within 15-minute intervals and verifying wind speeds above 30 mph, relative humidity (RH) values of 0% were also marked as data gaps and filled (Peeling et al. 2023).

For comparing with the IMERG rainfall analysis, we initially used two approaches: we pick the nearest grid point in the IMERG rainfall analysis to the geographical co-ordinates of the FAWN station. The other approach is to compute the average of the nearest 10 grid points in the IMERG analysis (IMERG10) to the geographical coordinates of the FAWN station. IMERG10 was generated to examine if it is a better representation of IMERG rainfall for validation of rainfall with the FAWN station in regions of strong precipitation gradient. In Figs. 2 and 3, we observe that relationship between IMERG and IMERG10 is extremely similar (with correlation coefficient > 0.9) between the two timeseries. Therefore, we abandon the approach of IMERG10 and use the IMERG timeseries that correspond to the closest grid point to the FAWN station for validation henceforth.

In contrast, the correlation of the timeseries of either IMERG or IMERG10 with the corresponding timeseries of FAWN for each of the 8 stations are extremely small (≤ 0.05). Therefore, a direct comparison of the raw timeseries of FAWN and IMERG would be futile. Hence, we breakdown the raw timeseries to discrete temporal scales and then make the comparison between FAWN and IMERG.

In this study we use the ensemble empirical mode decomposition (EEMD; Wu and Huang 2009) approach to analyze the data. EEMD is a data adaptive approach that does not use pre-determined kernel like Fourier transform. The EEMD can decompose the local features of the temporal variations into complete sets of near-orthogonal components called intrinsic mode functions (IMFs). The IMFs can be thought of as basic functions, which are determined by the time series itself rather than predetermined kernels. Thus, EEMD is a self-adaptive signal processing method, which is most suited for nonlinear and nonstationary time series. After decomposing the IMERG and FAWN timeseries separately to various frequencies, we then use

correlations and root mean square error (RMSE) to objectively assess the verification at the distinct timescales. This comparison will be conducted in each of the 8 FAWN stations. As an illustration of the discretization using EEMD we show a breakdown of the timeseries over Marianna into synoptic scale (Fig. 4), subseasonal scale (Fig. 5), interannual scale (Fig. 6) for a sample year of 2005.

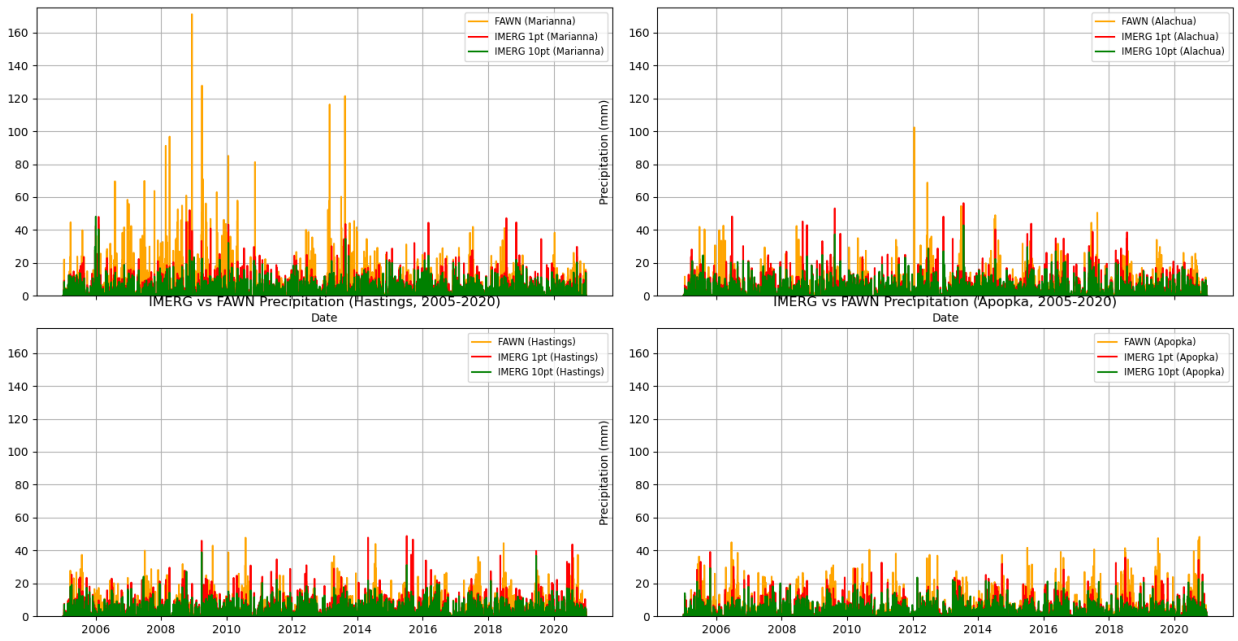

Figure 2: Timeseries of rainfall from IMERG (red) to the nearest grid point of FAWN station of a) Marianna, b) Alachua, c) Hastings, and d) Apopka overlaid with timeseries of rainfall averaged from nearest 10 grid points to the FAWN station (green) and timeseries of rainfall from FAWN (yellow) at intervals of 30 minutes and 15 minutes for IMERG and FAWN, respectively.

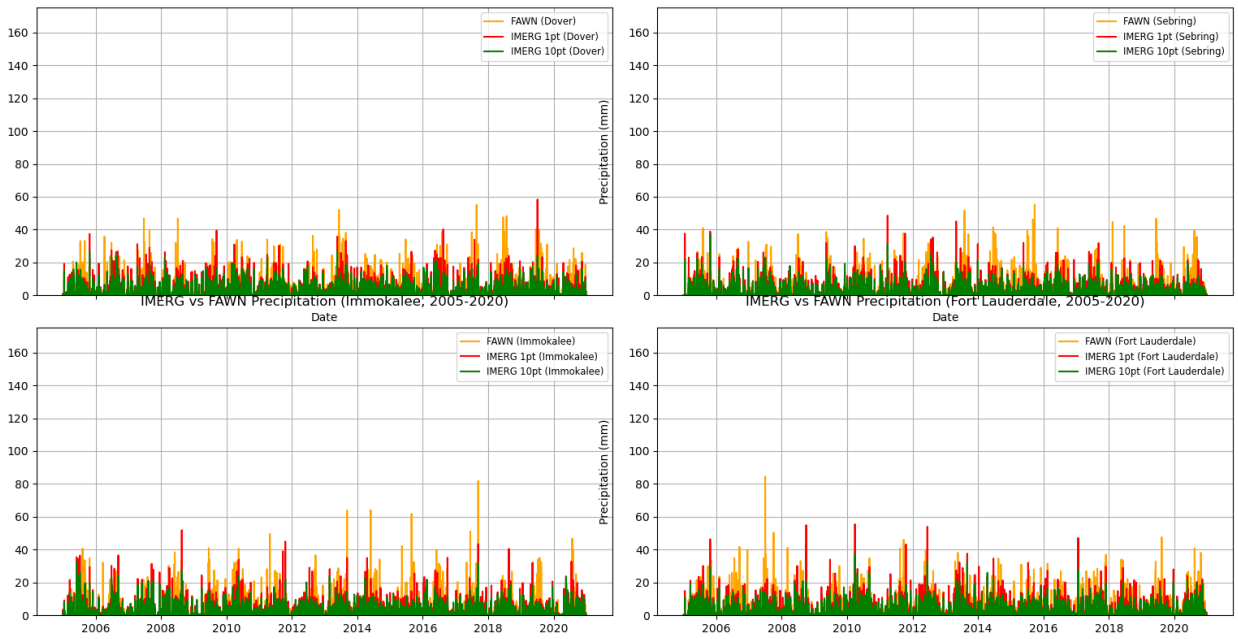

Figure 3: Timeseries of rainfall from IMERG (red) to the nearest grid point of FAWN station of a) Dover, b) Sebring, c) Immokalee, and d) Ft. Lauderdale overlaid with timeseries of rainfall averaged from nearest 10 grid points to the FAWN station (green) and timeseries of rainfall from FAWN (yellow) at intervals of 30 minutes and 15 minutes for IMERG and FAWN, respectively.

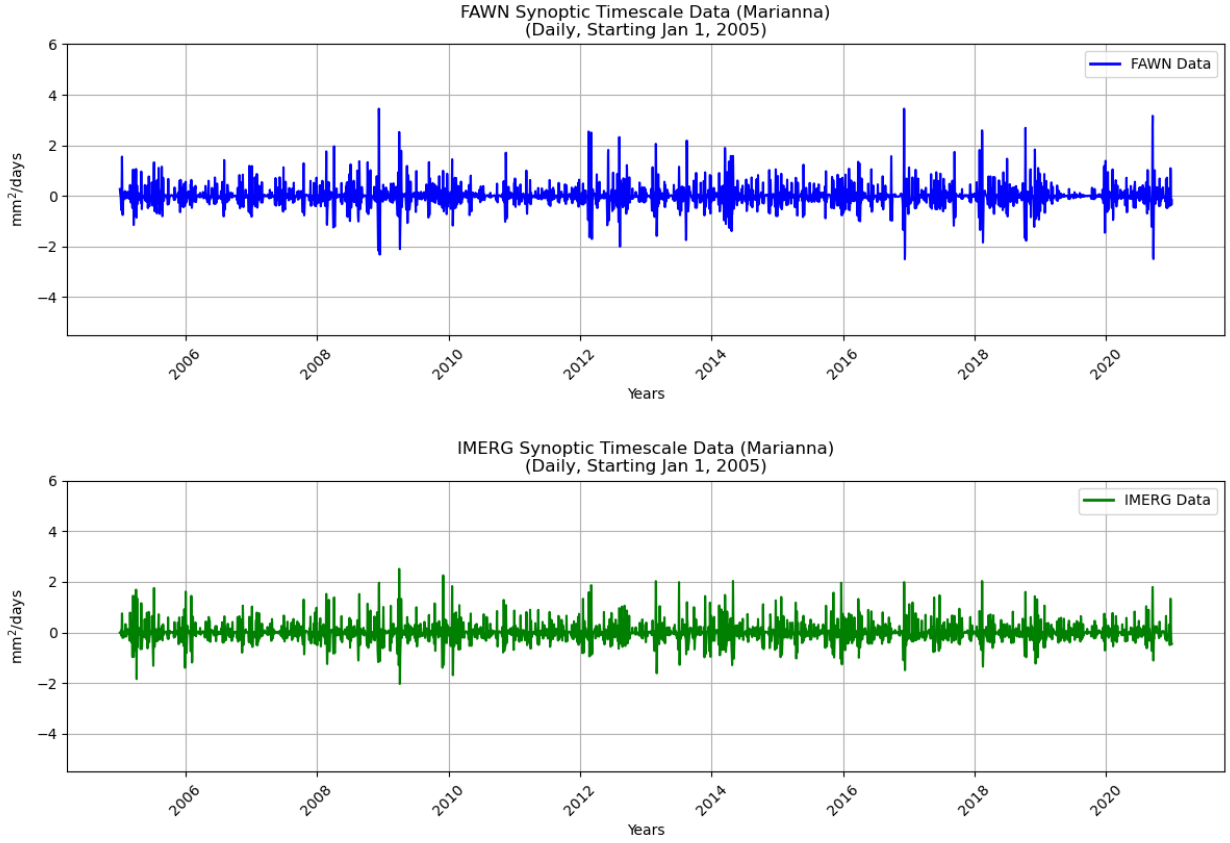

Figure 4: Filtered time series of rainfall (mm/day) isolating synoptic scales (2-10 days) over Marianna from a) FAWN and b) IMERG.

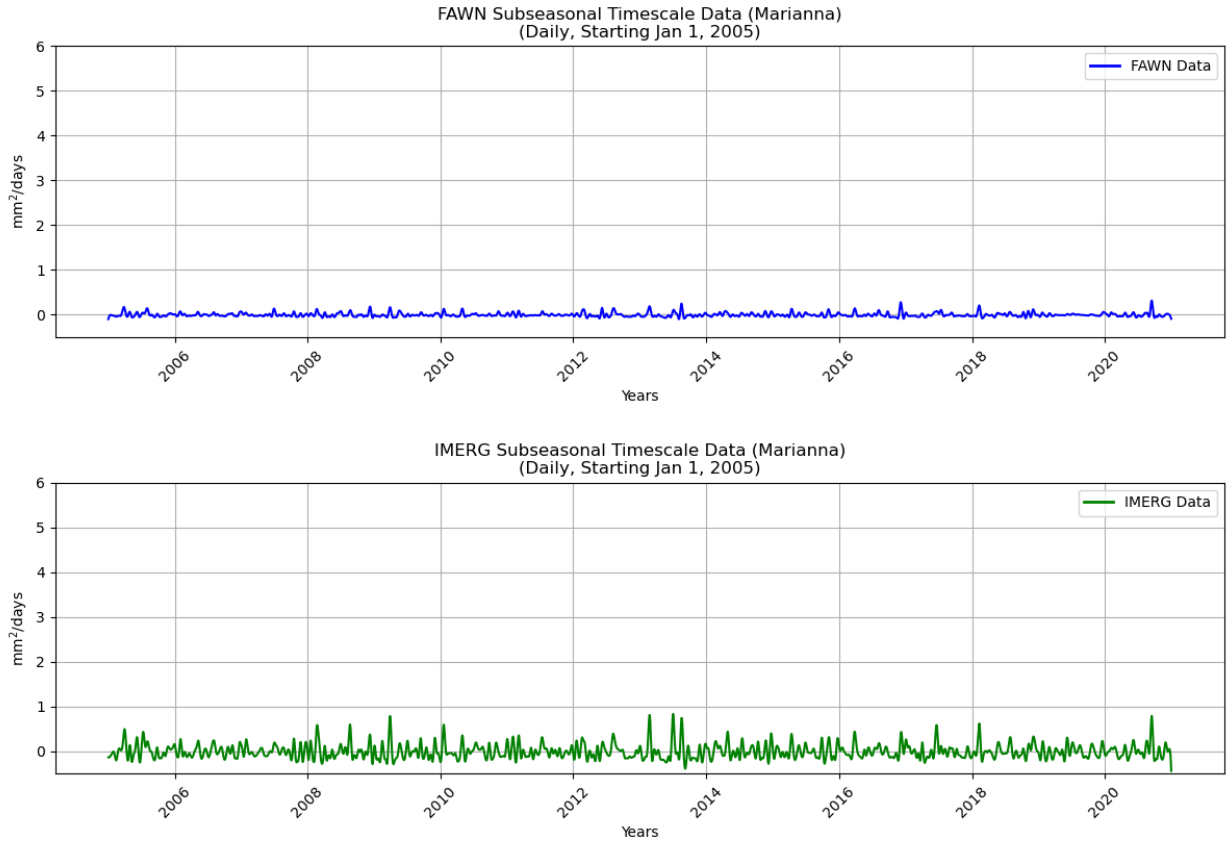

Figure 5: Filtered time series of rainfall (mm/day) isolating subseasonal scales (20-90 days) over Marianna from a) FAWN and b) IMERG.

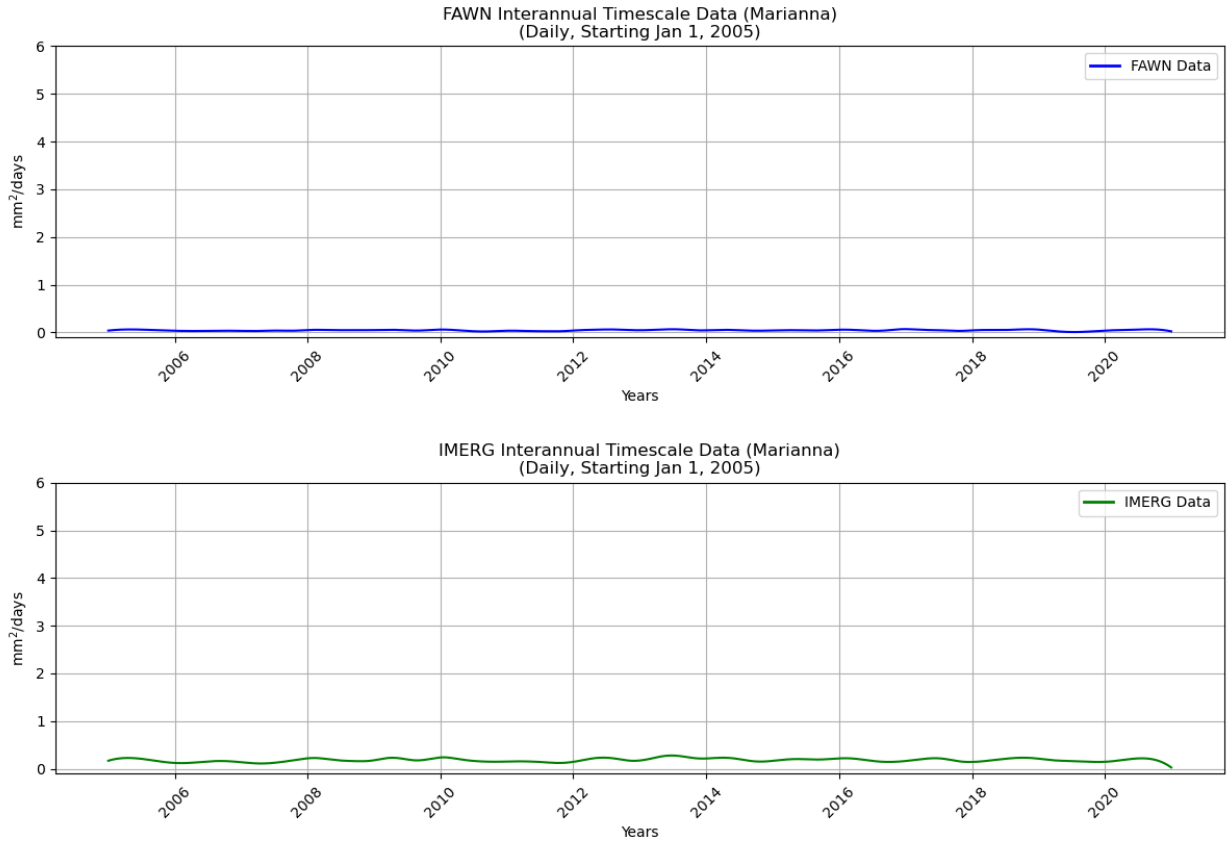

Figure 6: Filtered time series of rainfall isolating interannual scales (1-5 years) over Marianna from a) FAWN and b) IMERG.

3 Results

The analysis of the timeseries from IMERG and FAWN over the 8 stations (Fig. 1) are discussed in this section. We begin with synoptic (2 to 10 days) variability, followed by subseasonal variability (20-90 days), and interannual variability (1 to 5 years). Finally, we also verify the diurnal variations in the two timeseries.

a) Synoptic variability

(i) Marianna

The synoptic variability for Marianna is indicated in Figs. 4a and b for FAWN and IMERG, respectively. The variance is comparable in the two timeseries, although IMERG timeseries underestimates. The seasonal variability of the synoptic variance is similar in both, with winter exhibiting the most followed by the spring season with the least in the summer season. Again, systematically IMERG underestimates the variance in all seasons. The correlation of the timeseries between IMERG and FAWN is 0.73.

(ii) Alachua

The synoptic variability for Alachua is indicated in Figs. 7a and b for FAWN and IMERG, respectively. Unlike Marianna, the variance is grossly underestimated by IMERG. The seasonal variability of the synoptic variance is similar in both, with winter exhibiting the most followed by the spring season with the least in the summer season. But systematically IMERG underestimates the variance in all seasons by a significant amount. The correlation of the timeseries between IMERG and FAWN is insignificant.

(iii) Hastings

The synoptic variability for Alachua is indicated in Figs. 8a and b for FAWN and IMERG, respectively. The synoptic variance is underestimated by IMERG. Interestingly, in both timeseries, the seasonal variability of the synoptic variance although similar in both is most in the fall season followed by summer and least in winter season. But systematically IMERG underestimates the variance in all seasons. The correlation of the timeseries between IMERG and FAWN is 0.63.

(iv) Apopka

Like Hastings, the synoptic variability for Apopka is highest in fall season followed by summer (Figs. 9a and b). However, in FAWN the synoptic variance is least in winter while in IMERG it is most in the spring season. Like the other stations, IMERG underestimates the variance across all seasons. The correlation between the two timeseries is 0.62.

(v) Dover

The synoptic variance for Dover is comparable between FAWN and IMERG across all seasons except in the summer when it is almost less by half (Figs. 10a and b). Incidentally, summer is the

season with the highest synoptic variance over this region. The correlation between the two timeseries is 0.70.

(vi) Sebring

Over Sebring, we find that synoptic variance is slightly underestimated by IMERG (Figs. 11a and b). Again, in the summer, when FAWN displays the most variance in the year, IMERG underestimates it by nearly half as much. In the other seasons it is far more comparable with IMERG slightly overestimating the variance in the winter season. The correlation between FAWN and IMERG is 0.61.

(vii) Immokalee

Over Immokalee, the variance in IMERG is significantly underestimated (Figs. 12a and b). Although the correlation between the two timeseries is 0.61, there is significant underestimation of the variance in the summer and fall season by IMERG, when it is the highest in the year. However, in the winter, the variance is slightly overestimate.

(viii) Fort Lauderdale

The synoptic variance from the IMERG timeseries over Ft. Lauderdale is comparable to that in FAWN (Figs. 13a and b). The seasonal cycle of the seasonal variance is similar in the two timeseries and the correlations between them is 0.61. Nonetheless like most other stations, IMERG underestimates the synoptic variance compared to FAWN across all seasons except in the winter season.

In summary, IMERG generally shows a tendency to underestimate the synoptic variance across all stations and over most seasons. The one exception is a slight overestimation of the synoptic variance in the winter in a few of the stations, where the seasonal cycle of the synoptic variance is highest in the summer or the fall season, which are the stations in south Florida. The temporal correlations of the timeseries of FAWN with IMERG over all stations except over Alachua is over 0.6. In terms of spatial patterns, it is seen that central Florida (Apopka, Dover, and Sebring) show the least synoptic variance compared to stations located north or south of it.

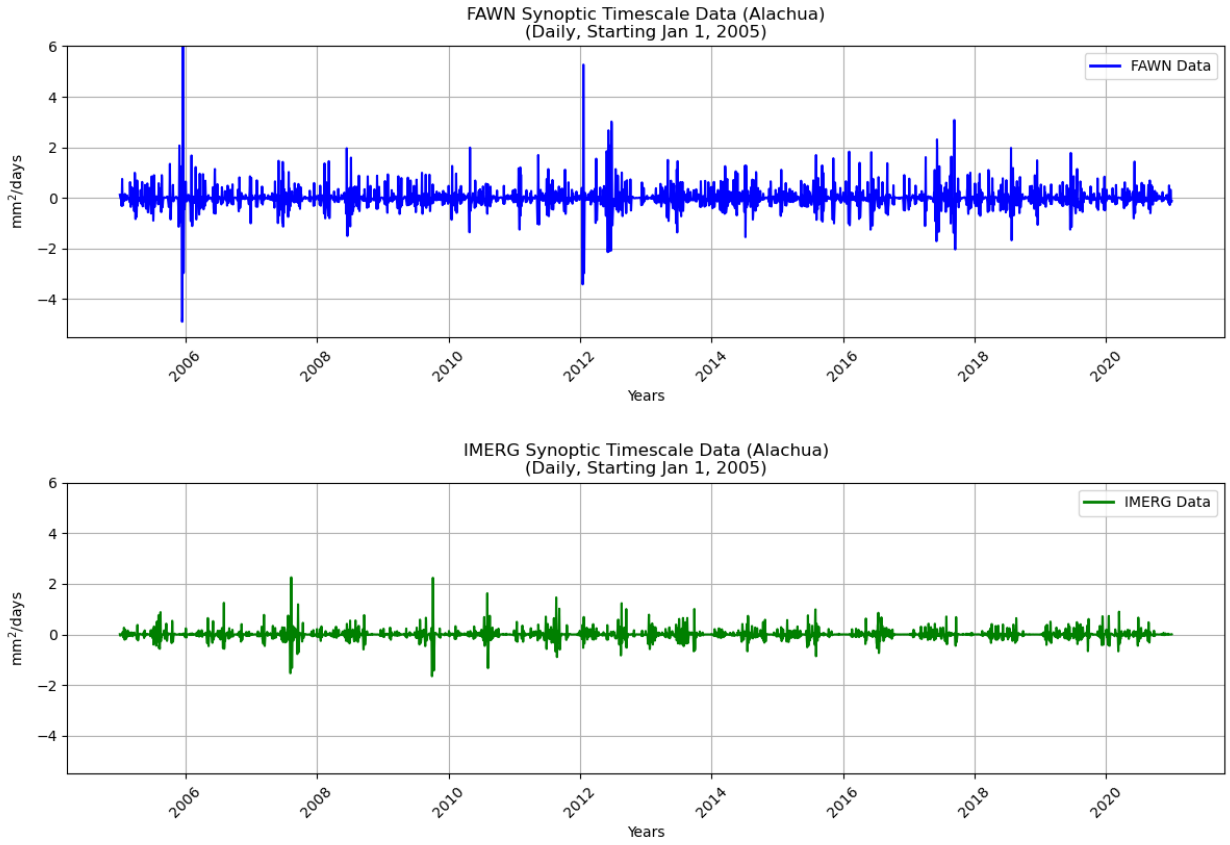

Figure 7: Filtered time series of rainfall isolating synoptic scales (2-10 days) over Alachua from a) FAWN and b) IMERG. The units of the rainfall anomalies along the ordinate is mm/day.

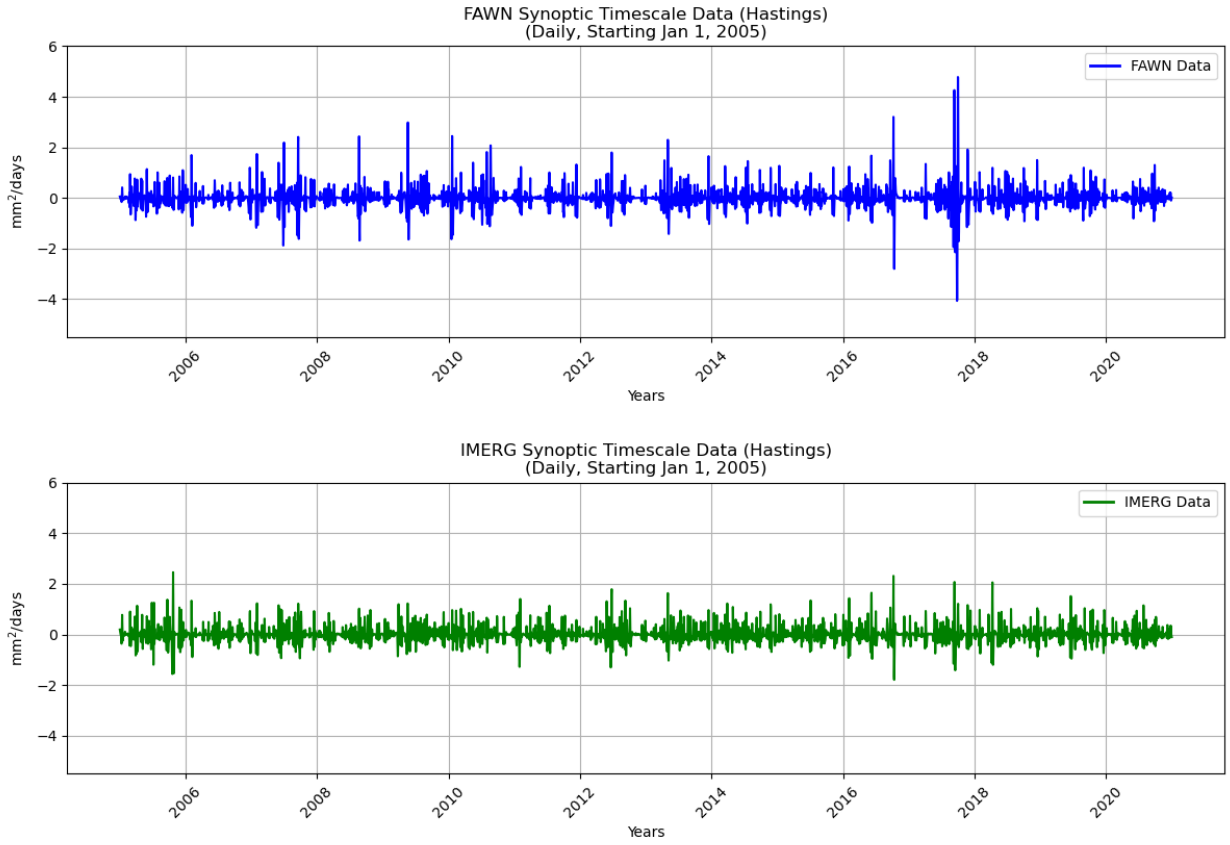

Figure 8: Filtered time series of rainfall isolating synoptic scales (2-10 days) over Hastings from a) FAWN and b) IMERG. The units of the rainfall anomalies along the ordinate is mm/day.

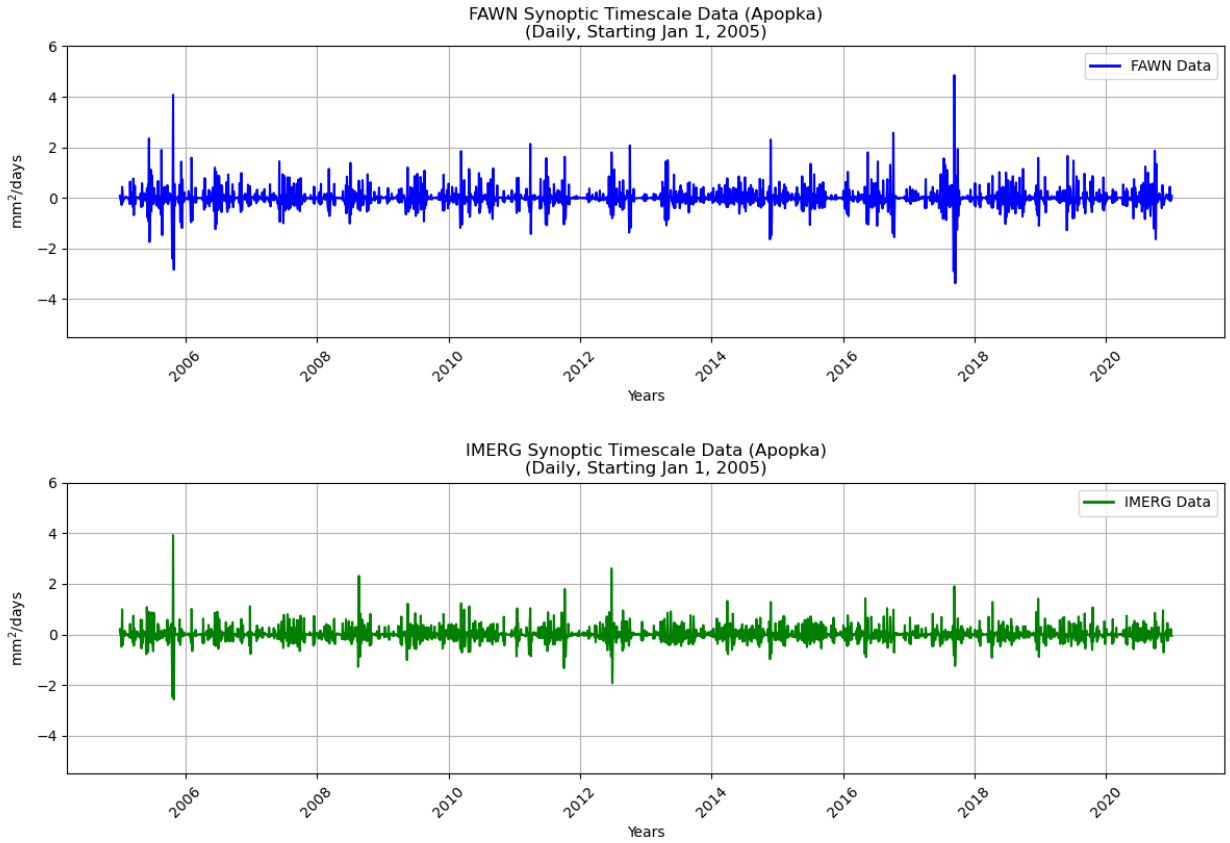

Figure 9: Filtered time series of rainfall isolating synoptic scales (2-10 days) over Apopka from a) FAWN and b) IMERG. The units of the rainfall anomalies along the ordinate is mm/day.

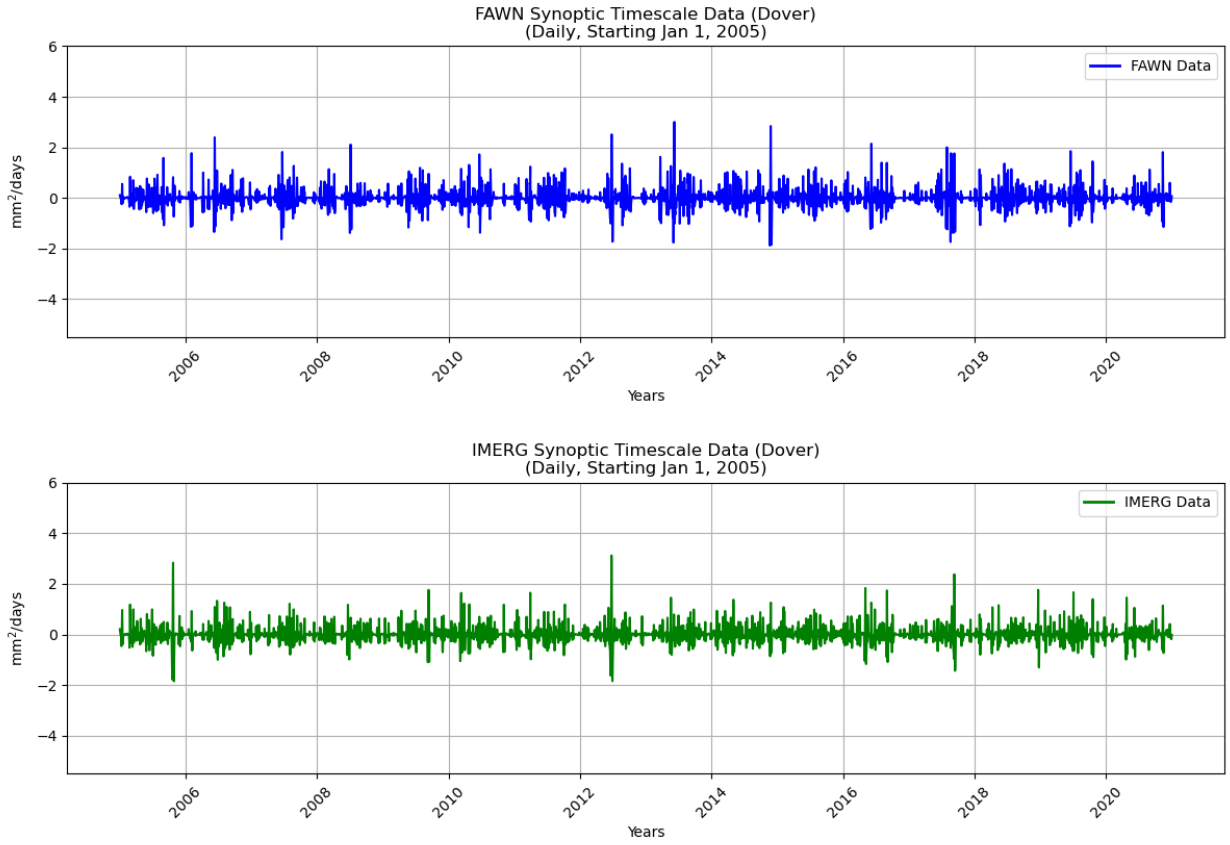

Figure 10: Filtered time series of rainfall isolating synoptic scales (2-10 days) over Dover from a) FAWN and b) IMERG. The units of the rainfall anomalies along the ordinate is mm/day.

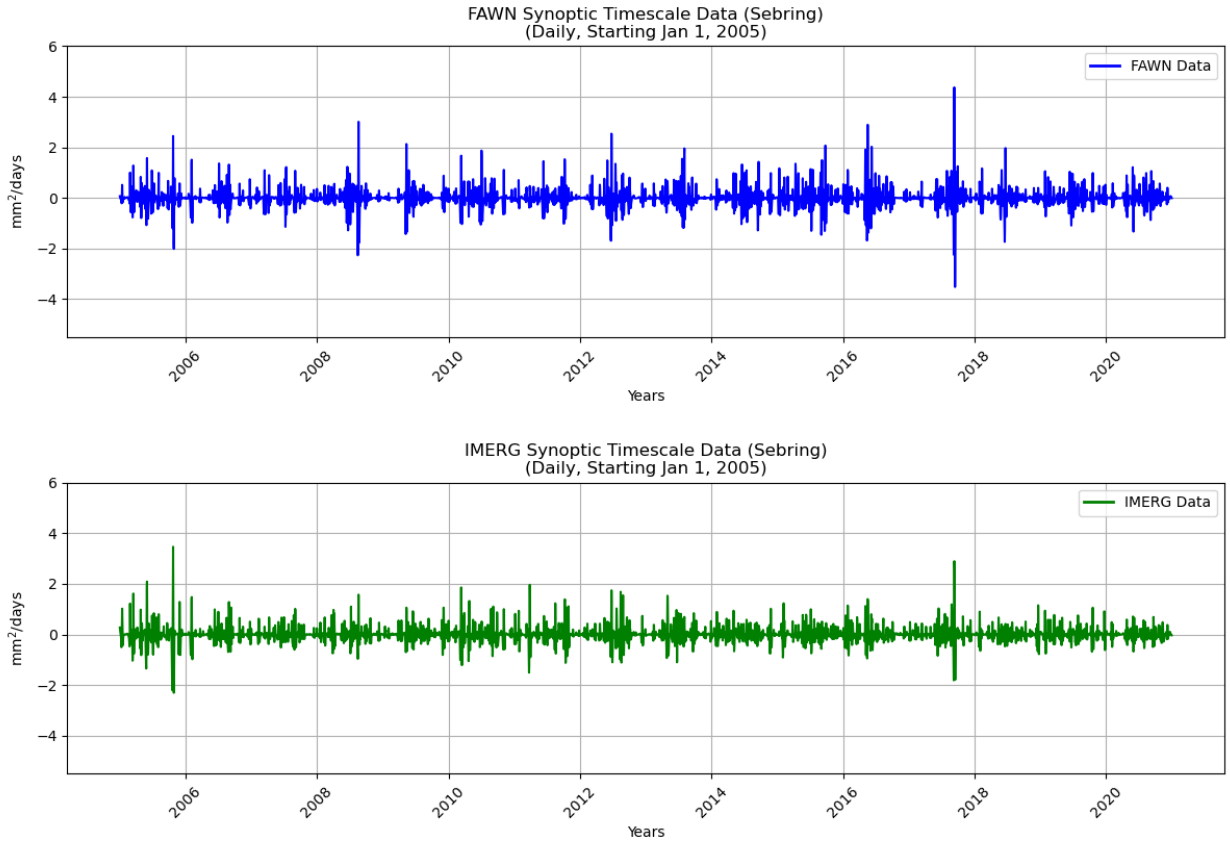

Figure 11: Filtered time series of rainfall isolating synoptic scales (2-10 days) over Sebring from a) FAWN and b) IMERG. The units of the rainfall anomalies along the ordinate is mm/day.

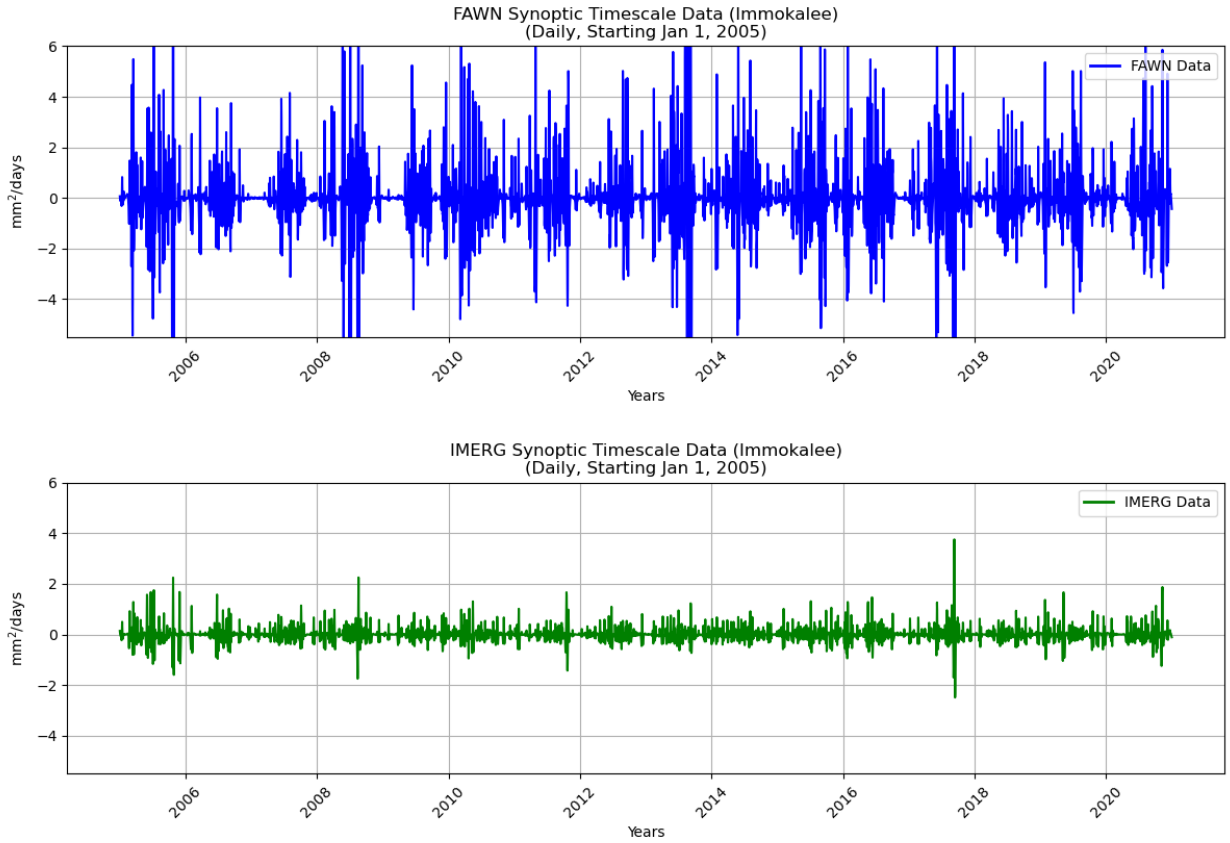

Figure 12: Filtered time series of rainfall isolating synoptic scales (2-10 days) over Immokalee from a) FAWN and b) IMERG. The units of the rainfall anomalies along the ordinate is mm/day.

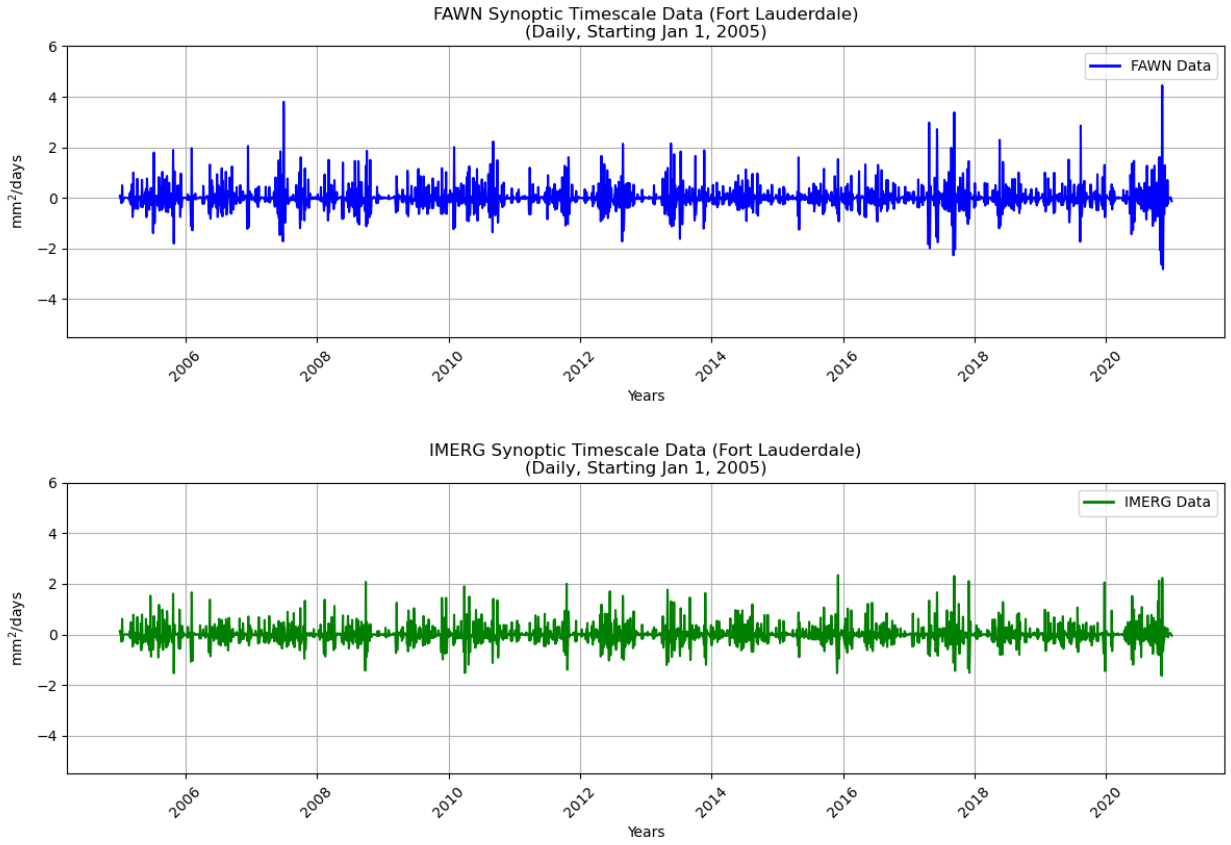

Figure 13: Filtered time series of rainfall isolating synoptic scales (2-10 days) over Fort Lauderdale from a) FAWN and b) IMERG. The units of the rainfall anomalies along the ordinate is mm/day.

b) Subseasonal variability

(i) Marianna

The timeseries of subseasonal variability between the two datasets are comparable in terms of their variance and the correlation between them is 0.82 (Figs. 5a and b). The seasonal variability of the subseasonal variance is also comparable with the highest variance in the winter, least in the fall season and are comparable in summer and spring seasons.

(ii) Alachua

Even at subseasonal timescales, Alachua continues to display insignificant correlation between the two timeseries (Figs. 14a and b). The variance explained by IMERG is significantly lower than in FAWN although both timeseries show the highest variance in the summer season.

(iii) Hastings

The subseasonal variance is comparable between FAWN and IMERG over Hastings (Figs. 15a and b). The temporal correlation between them is 0.61. Although the subseasonal variance between the two are comparable across all seasons, their seasonal cycle is slightly different. For instance, FAWN has the highest variance in the summer followed by spring. But IMERG shows highest variance in spring and summer seasons followed by fall season as in FAWN.

(iv) Apopka

The subseasonal variations over Apopka between FAWN and IMERG are comparable to each other in terms of their sub-seasonal variance and their seasonal variability (Figs. 16a and b). The temporal correlations between the two timeseries is 0.75.

(v) Dover

The subseasonal variations over Dover is slightly underestimated by IMERG, although the temporal correlations with FAWN is 0.70 (Figs. 17a and b). The seasonal variability of the subseasonal variance in both the timeseries are also comparable with the highest variance in the summer.

(vi) Sebring

Over Sebring, the subseasonal variance is comparable between the two timeseries (Figs. 18a and b). The temporal correlations between the two timeseries is 0.73. Furthermore the seasonal variability of the subseasonal variance is similar, with summer showing the highest variance. However, this followed by fall season in FAWN but in IMERG it is spring season that displays the next highest variability at sub-seasonal timescales.

(vii) Immokalee

Over Immokalee, the subseasonal timeseries from IMERG is comparable with FAWN (Figs. 19a and b). The total variance on the sub-seasonal timescales is quite comparable with the temporal correlation between them being 0.68. The seasonal variability of the subseasonal variance is also similar in the two timeseries.

(viii) Fort Lauderdale

The subseasonal variability over Ft. Lauderdale is similar between IMERG and FAWN (Figs. 20a and b). The correlation between the two timeseries is 0.66. The seasonal variability of the subseasonal variance is also similar between the two timeseries.

In summary, the validation of IMERG with FAWN at subseasonal timescales is slightly better than at synoptic scales as expected in terms of temporal correlations between them. Furthermore, the subseasonal variance is less than synoptic variance by about 1/3rd across all stations in both timeseries. However, Alachua still displays poor validation for IMERG at subseasonal timescales. In terms of spatial patterns, it is seen that central Florida (Apopka, Dover, and Sebring) show the least subseasonal variance compared to stations located north or south of it.

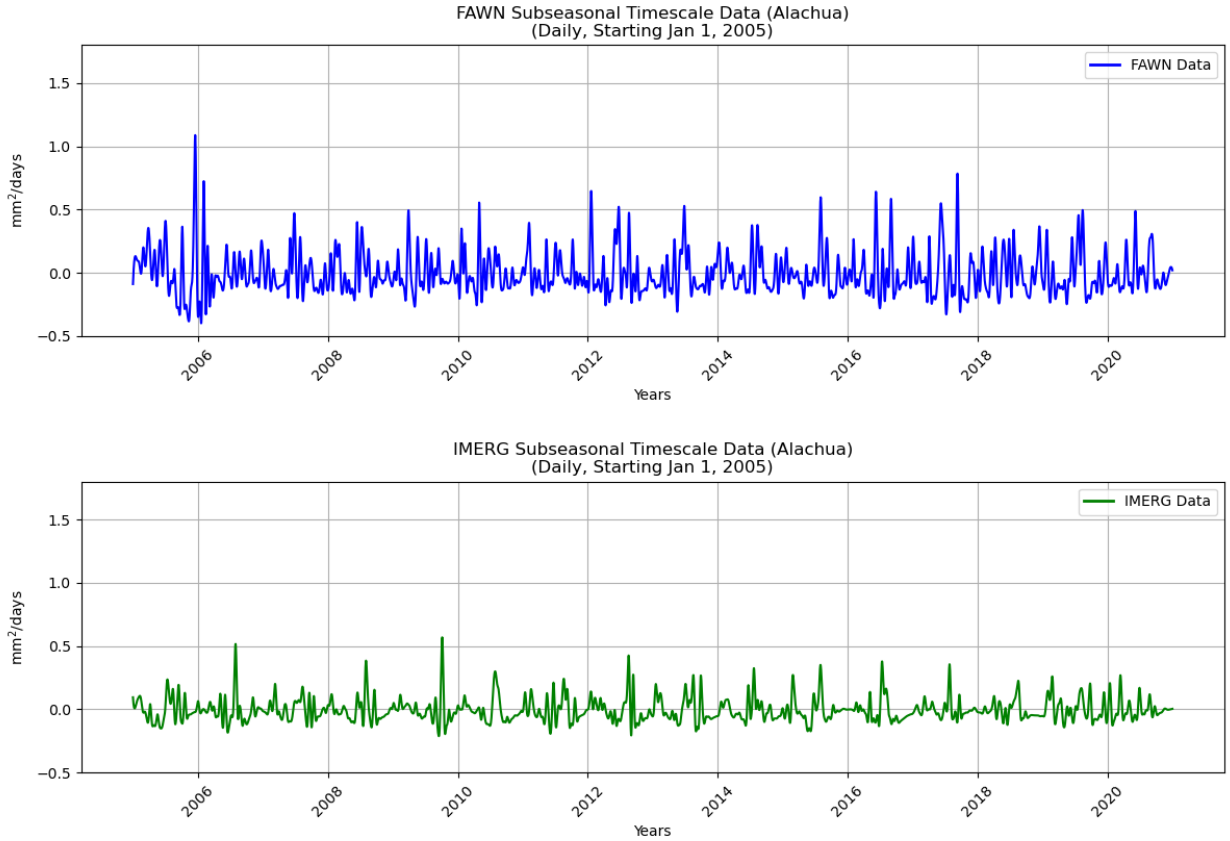

Figure 14: Filtered time series of rainfall isolating subseasonal scales (20-90 days) over Alachua from a) FAWN and b) IMERG. The units of the rainfall anomalies along the ordinate is mm/day.

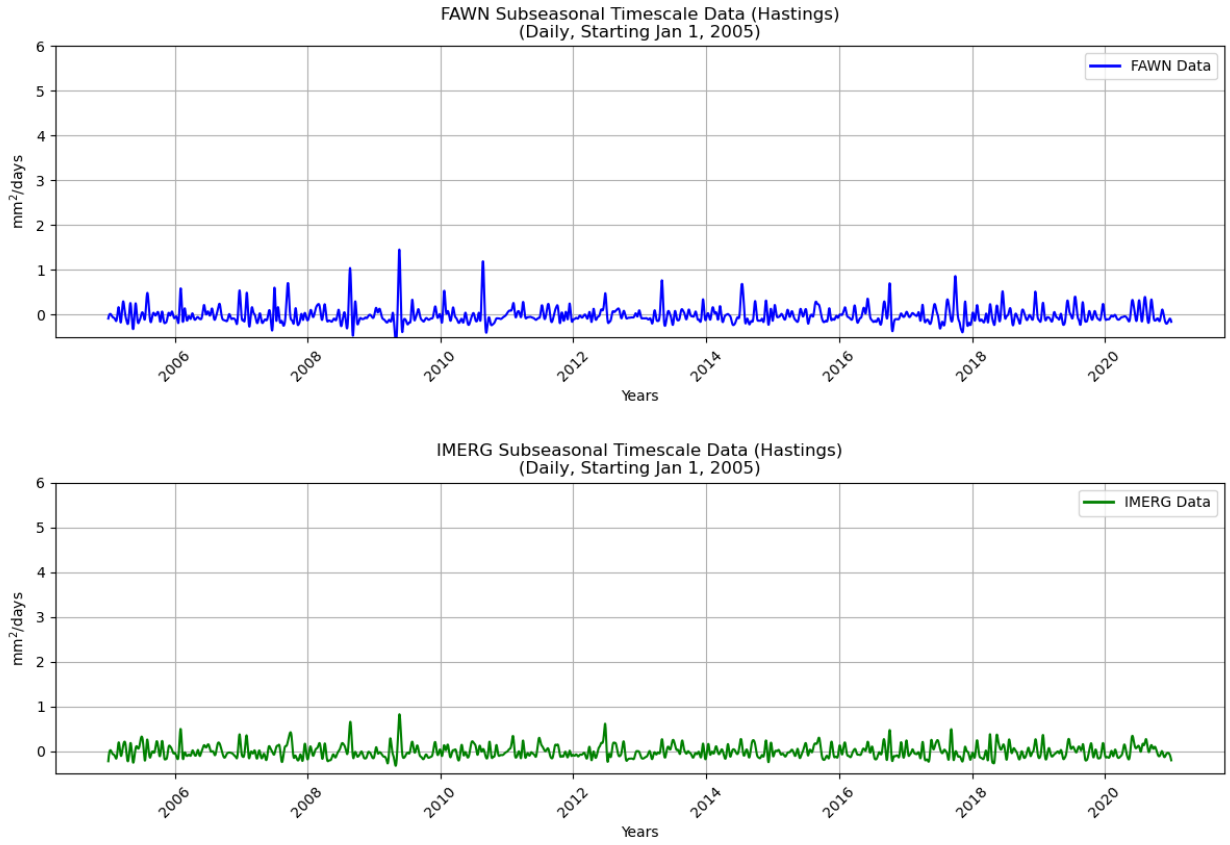

Figure 15: Filtered time series of rainfall isolating subseasonal scales (20-90 days) over Hastings from a) FAWN and b) IMERG. The units of the rainfall anomalies along the ordinate is mm/day.

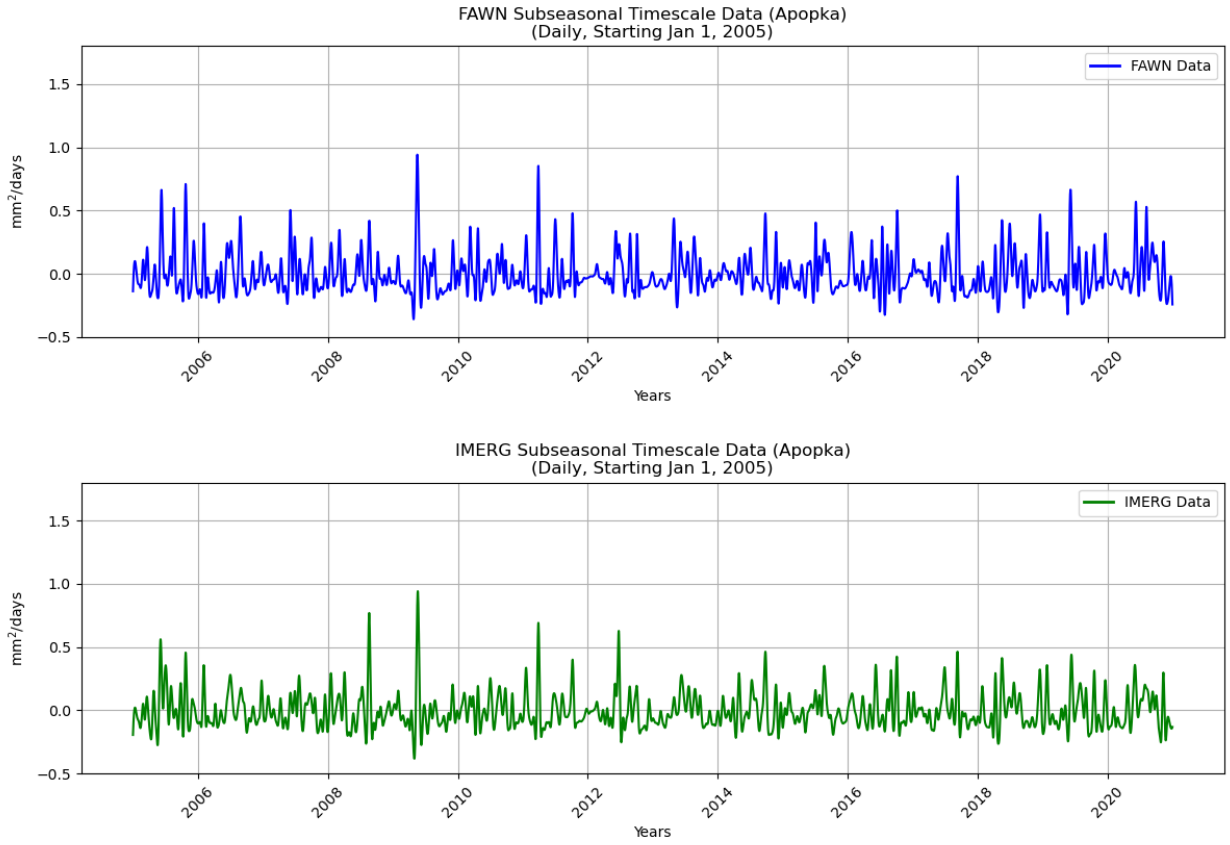

Figure 16: Filtered time series of rainfall isolating subseasonal scales (20-90 days) over Apopka from a) FAWN and b) IMERG. The units of the rainfall anomalies along the ordinate is mm/day.

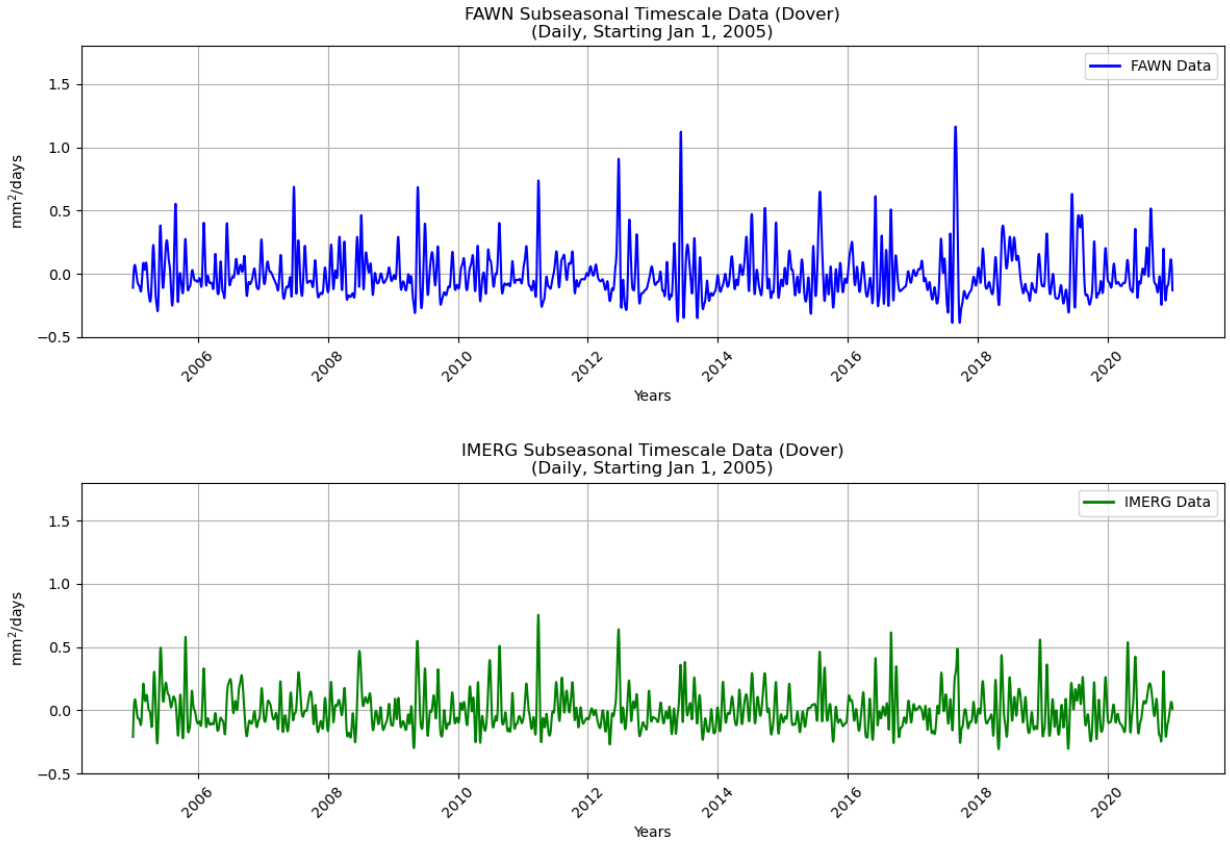

Figure 17: Filtered time series of rainfall isolating subseasonal scales (20-90 days) over Dover from a) FAWN and b) IMERG. The units of the rainfall anomalies along the ordinate is mm/day.

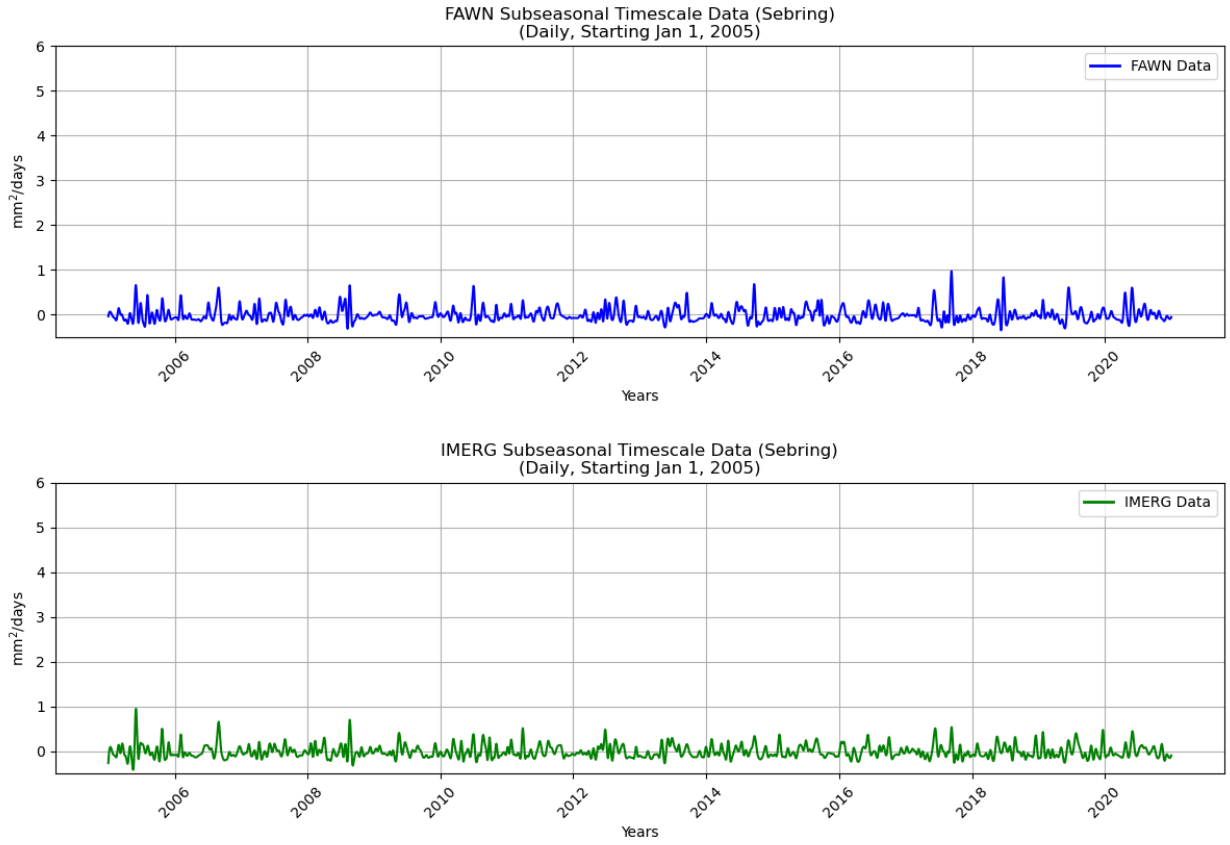

Figure 18: Filtered time series of rainfall isolating subseasonal scales (20-90 days) over Sebring from a) FAWN and b) IMERG. The units of the rainfall anomalies along the ordinate is mm/day.

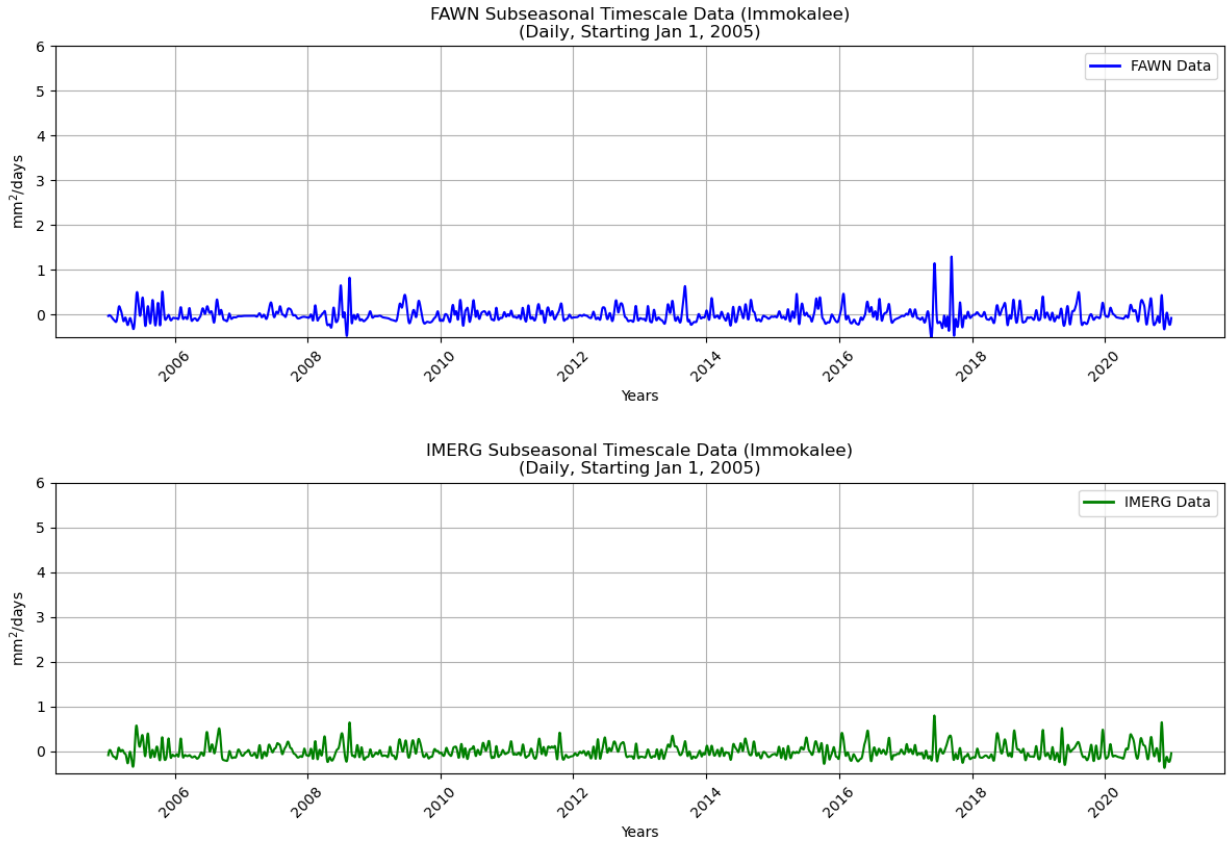

Figure 19: Filtered time series of rainfall isolating subseasonal scales (20-90 days) over Immokalee from a) FAWN and b) IMERG. The unit of the rainfall anomalies along the ordinate is mm/day.

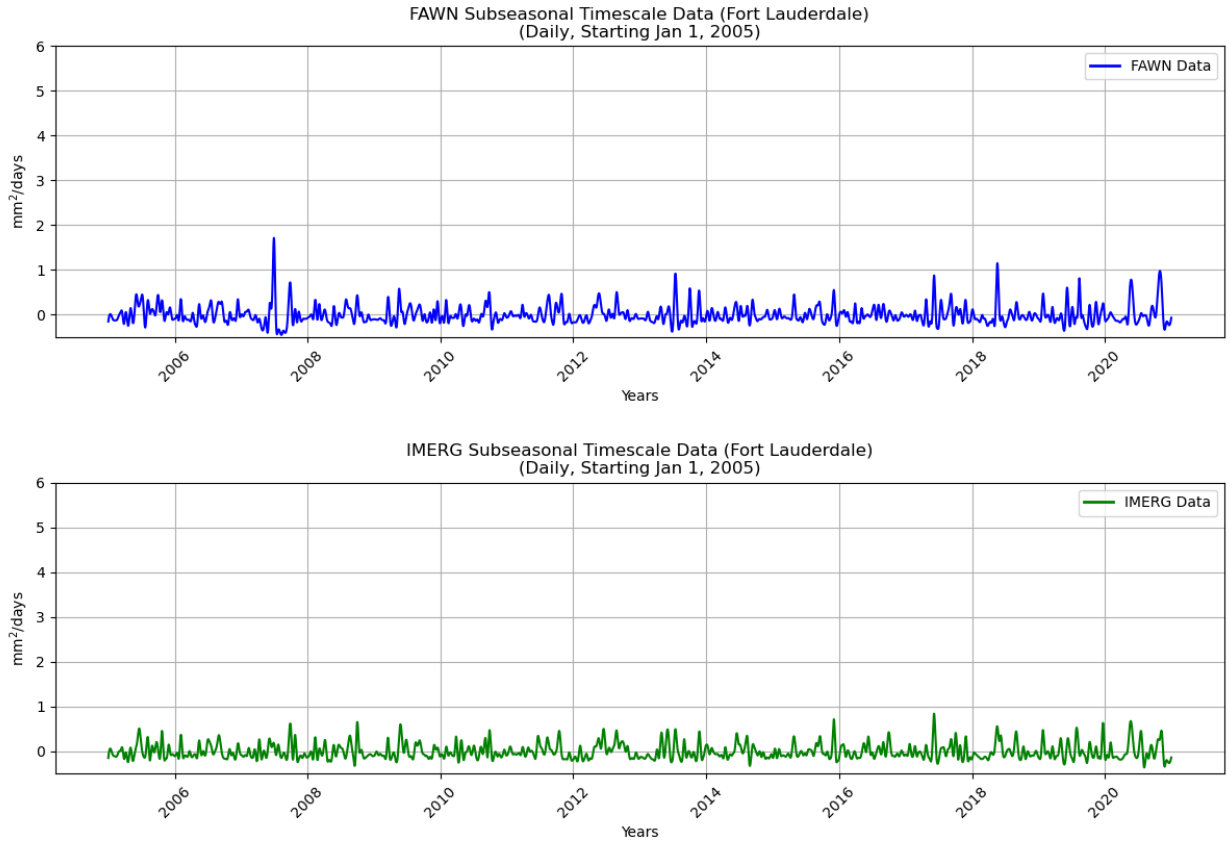

Figure 20: Filtered time series of rainfall isolating subseasonal scales (20-90 days) over Fort Lauderdale from a) FAWN and b) IMERG. The unit of the rainfall anomalies along the ordinate is mm/day.

c) Interannual variability

(i) Marianna

Over Marianna, the interannual timeseries of IMERG is quite comparable to corresponding FAWN timeseries (Figs. 6a and b). The temporal correlation between them is 0.71. The seasonal variability of the interannual variance however are not alike. For example, FAWN shows the highest interannual variance in spring followed by winter. Although, the variance is comparable across seasons in FAWN. The seasonal variability of interannual variance is stronger in IMERG with spring season showing the highest and fall season showing the least.

(ii) Alachua

The mismatch between the timeseries of FAWN and IMERG over this station continues to be an issue even at interannual scales (Figs. 21a and b). The temporal correlation between the two time series is 0.3. Although the total interannual variance is slightly higher in IMERG compared to FAWN. Furthermore, FAWN shows highest interannual variance in summer while IMERG shows highest in the winter.

(iii) Hastings

The interannual timeseries of FAWN and IMERG are comparable over Hastings (Figs. 22a and b). The temporal correlation between them is 0.77 and the total variance is quite comparable. The seasonal variability of the interannual variations is also comparable between FAWN and IMERG.

(iv) Apopka

The interannual timeseries of FAWN and IMERG are comparable over Apopka (Figs. 23a and b). The temporal correlation between them is 0.84 and the total variance is quite comparable. The seasonal variability of the interannual variations is also comparable between FAWN and IMERG.

(v) Dover

The temporal correlations at interannual scales between FAWN and IMERG for Dover is 0.92 and the total variance is also comparable (Figs. 24a and b). Furthermore, the seasonal variability of the interannual variance is also comparable between the two timeseries.

(vi) Sebring

The fidelity of the IMERG timeseries at interannual scale for Sebring is quite high (Figs. 25a and b). The temporal correlations with FAWN timeseries are 0.92 and the total variance is nearly the same in the two timeseries. Furthermore, the seasonal variability of the interannual variations are also similar in the two timeseries

(vii) Immokalee

Similarly, the IMERG timeseries shows a strong verification with FAWN over Immokalee (Figs. 26a and b). The temporal correlations with FAWN timeseries are 0.89 and the total variance is nearly the same in the two timeseries. Furthermore, the seasonal variability of the interannual variations are also similar in the two timeseries.

(viii) Fort Lauderdale

The temporal correlations at interannual scales between FAWN and IMERG for Ft. Lauderdale is 0.79 and the total variance is also comparable (Figs. 27a and b). Furthermore, the seasonal variability of the interannual variance is also comparable between the two timeseries.

In summary, we find that the fidelity of IMERG when compared with FAWN is the best at interannual timescales relative to synoptic and sub-seasonal timescales across all 8 stations. In terms of spatial variability, the interannual variance is comparable across all stations in both FAWN and IMERG. Alachua continues to show the least correlations and largest discrepancy in terms of variance between the two timeseries even at the interannual timescales.

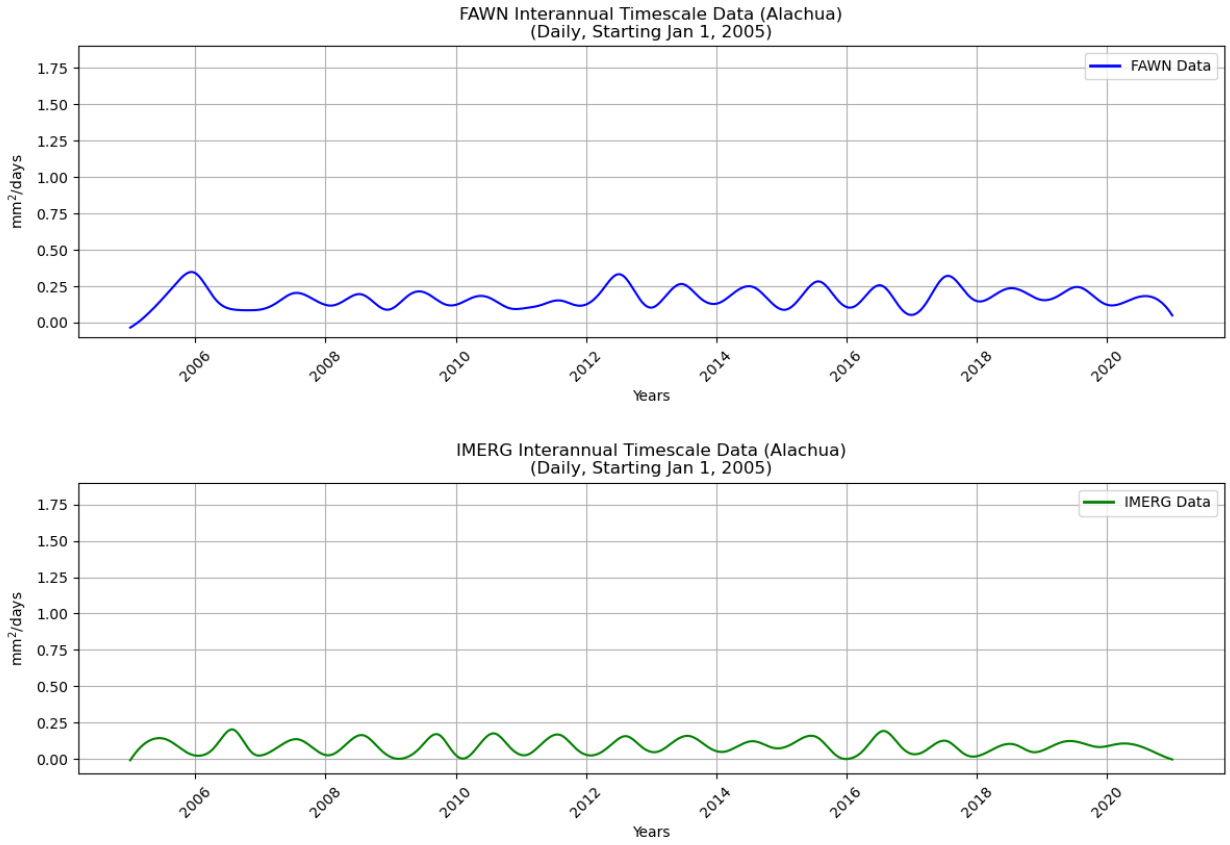

Figure 21: Filtered time series of rainfall isolating interannual scales (1-5 years) over Alachua from a) FAWN and b) IMERG. The unit of the rainfall anomalies along the ordinate is mm/day.

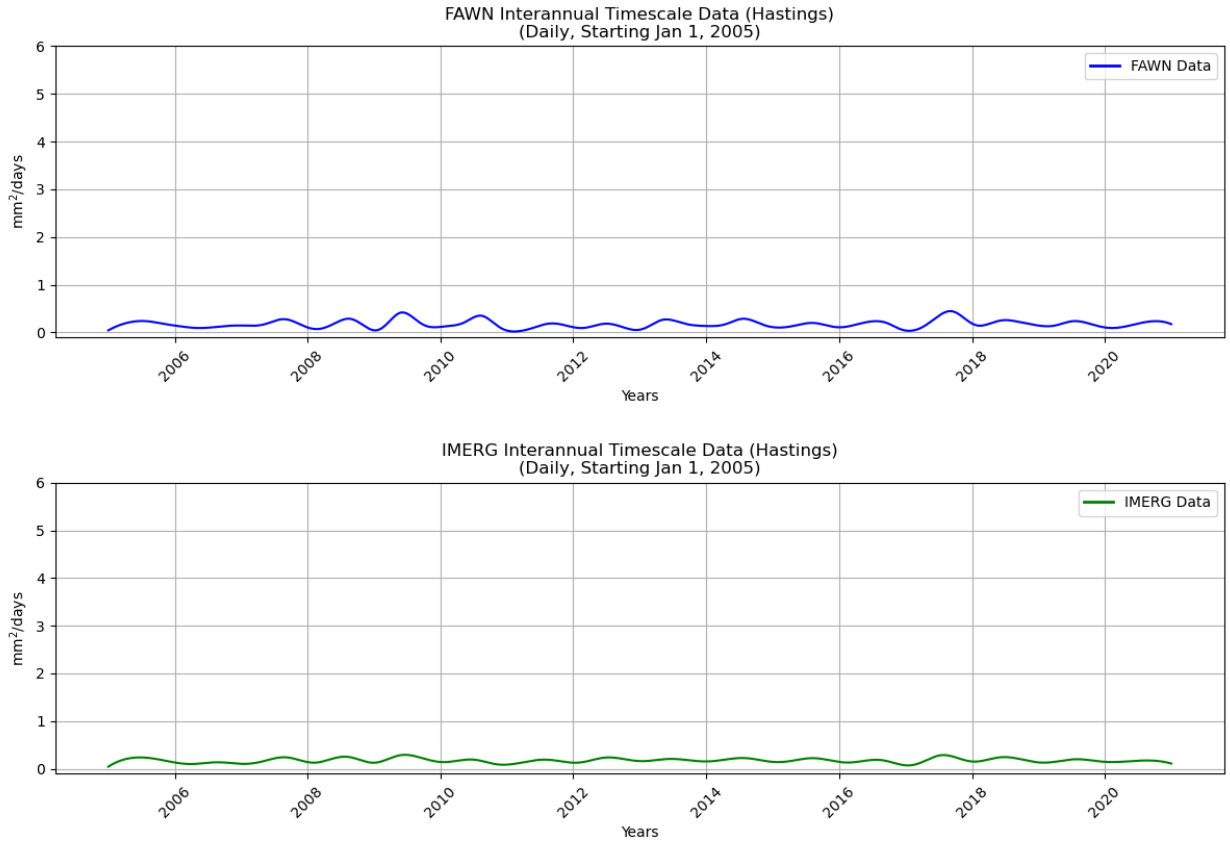

Figure 22: Filtered time series of rainfall isolating interannual scales (1-5 years) over Hastings from a) FAWN and b) IMERG. The unit of the rainfall anomalies along the ordinate is mm/day.

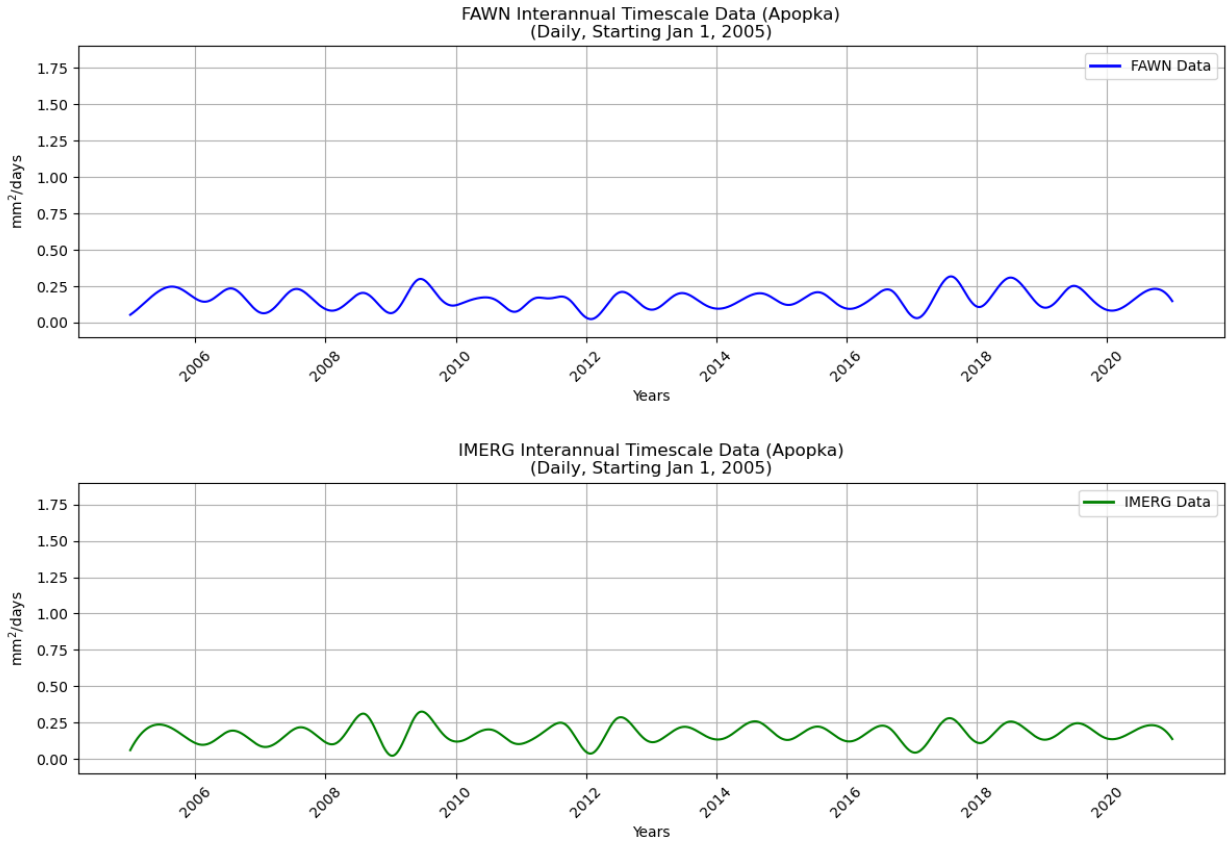

Figure 23: Filtered time series of rainfall isolating interannual scales (1-5 years) over Apopka from a) FAWN and b) IMERG. The unit of the rainfall anomalies along the ordinate is mm/day.

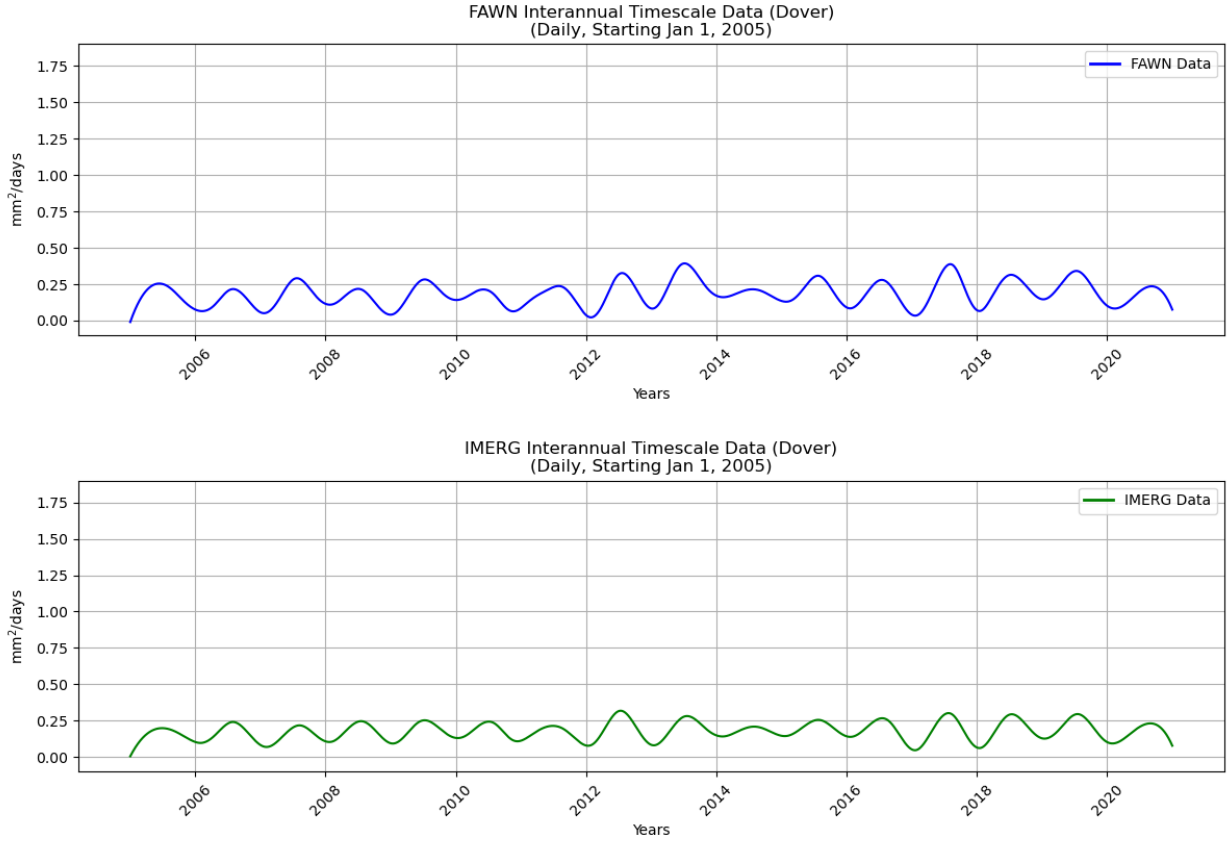

Figure 24: Filtered time series of rainfall isolating interannual scales (1-5 years) over Dover from a) FAWN and b) IMERG. The unit of the rainfall anomalies along the ordinate is mm/day.

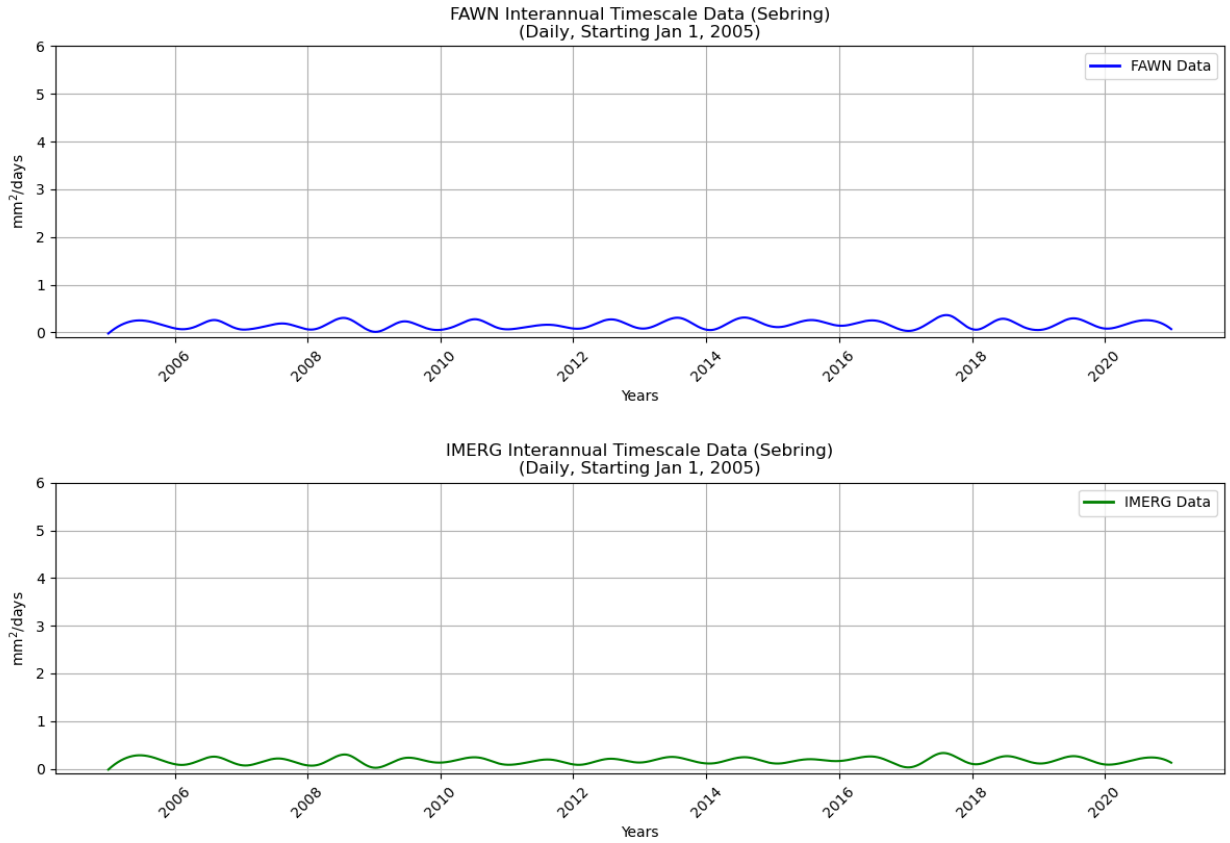

Figure 25: Filtered time series of rainfall isolating interannual scales (1-5 years) over Sebring from a) FAWN and b) IMERG. The unit of the rainfall anomalies along the ordinate is mm/day.

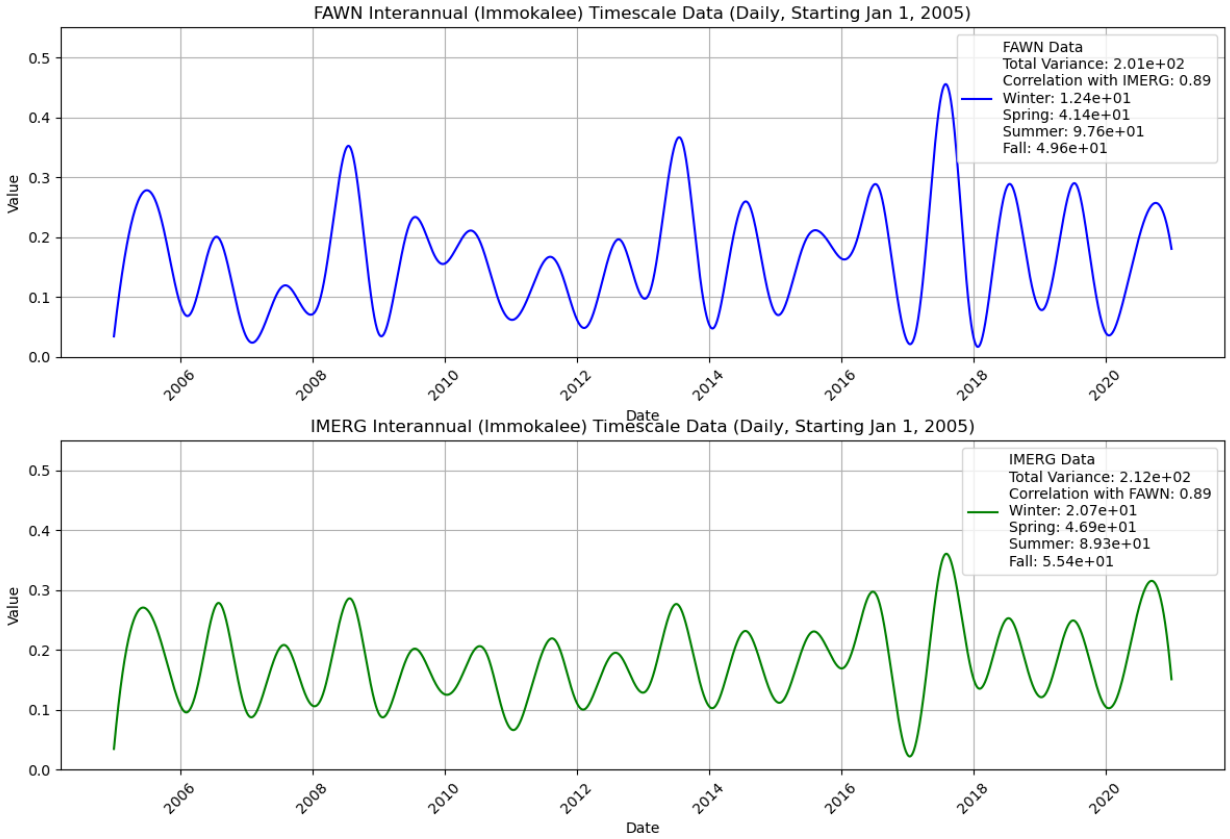

Figure 26: Filtered time series of rainfall isolating interannual scales (1-5 years) over Immokalee from a) FAWN and b) IMERG. The unit of the rainfall anomalies along the ordinate is mm/day.

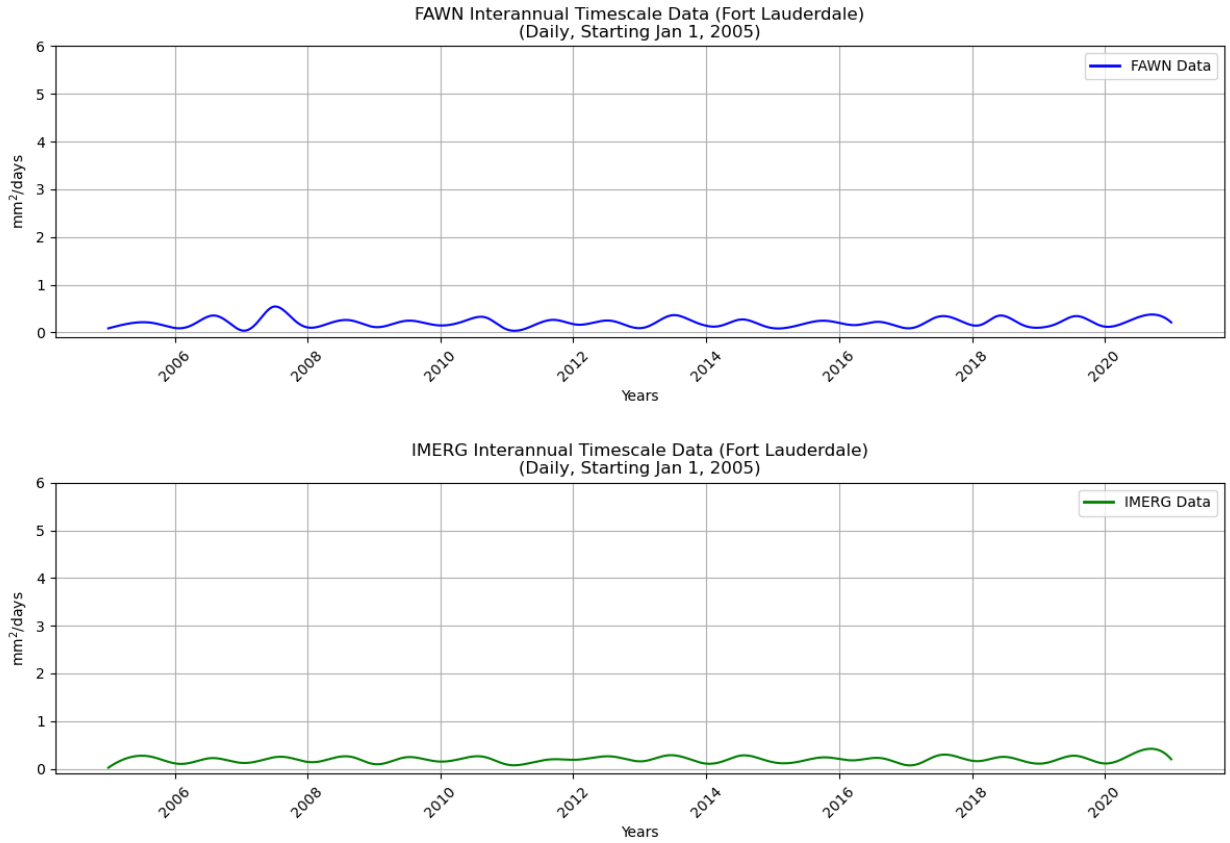

Figure 27: Filtered time series of rainfall isolating interannual scales (1-5 years) over Fort Lauderdale from a) FAWN and b) IMERG. The unit of the rainfall anomalies along the ordinate is mm/day.

d) Diurnal variability

EEMD also similarly isolates the diurnal variations like the other timescale. Since the temporal resolution of FAWN is 15 minutes and IMERG Is 30 minutes, they are both able to resolve the diurnal variations of precipitation. We compute the climatological diurnal cycle from the 15 years of data from both FAWN and EEMD for each day. The phase and amplitude for each day of the season is then computed from this diurnal cycle for comparison, which is discussed below.

(i) Marianna

In Fig. 28 we show the scatter of the diurnal amplitude for each of the four seasons for Marianna. The correlation between the two datasets is highest in the fall season and least in the summer. Generally, across all seasons, IMERG tends to overestimate the diurnal amplitude. The diurnal amplitude is the least in the winter and largest in the summer followed by that in fall season. The timing of the diurnal peak of rainfall shows very poor fidelity in IMERG across all four seasons (Fig. 29).

(ii) Alachua

In comparison to Marianna, Alachua shows poor verification for IMERG both for diurnal amplitude (Fig. 30) and phase (Fig. 31).

(iii) Hastings

In Fig. 32 we show the scatter of the diurnal amplitude for each of the four seasons for Hastings. The correlation between the two datasets is highest in the fall season and least in the summer. Generally, across all seasons, IMERG tends to overestimate the diurnal amplitude. The diurnal amplitude is the least in the winter and largest in the summer followed by that in fall season. The timing of the diurnal peak of rainfall shows very poor fidelity in IMERG across all four seasons (Fig. 33).

(iv) Apopka

In Fig. 34 we show the scatter of the diurnal amplitude for each of the four seasons for Apopka. The correlation between the two datasets is highest in the fall season and least in the summer. In the spring and winter seasons the correlations are 0.55 and 0.57, respectively. Generally, across all seasons, IMERG tends to overestimate the diurnal amplitude. The diurnal amplitude is the least in the winter and largest in the summer followed by that in fall season. The timing of the diurnal peak of rainfall shows very poor fidelity in IMERG across all four seasons (Fig. 35).

(v) Dover

In Fig. 36 we show the scatter of the diurnal amplitude for each of the four seasons for Dover. The correlation between the two datasets is highest in the winter season and least in the summer. In the spring and fall seasons the correlations are 0.43 and 0.62, respectively. Generally, across all seasons, IMERG tends to overestimate the diurnal amplitude. The diurnal amplitude is the least in the winter and largest in the summer followed by that in fall season. The timing of the diurnal peak of rainfall shows very poor fidelity in IMERG across all four seasons (Fig. 37).

(vi) Sebring

In comparison to the previous 5 stations, Sebring displays a poor fidelity of IMERG diurnal amplitude of precipitation (Fig. 38). The correlations across all seasons are much lower with highest correlation in the fall season at 0.34. IMERG continues to overestimate the diurnal amplitude relative to FAWN. The timing of the diurnal peak of rainfall shows very poor fidelity in IMERG across all four seasons (Fig. 39).

(vii) Immokalee

In Fig. 40 we show the scatter of the diurnal amplitude for each of the four seasons for Immokalee. The correlation between the two datasets is highest in the fall season (0.80) and least in the summer (0.21). In the spring and fall seasons the correlations are 0.46 and 0.65, respectively. Generally, across all seasons, IMERG tends to overestimate the diurnal amplitude. The diurnal amplitude is

the least in the winter and largest in the fall followed by that in summer season. The timing of the diurnal peak of rainfall shows very poor fidelity in IMERG across all four seasons (Fig. 41).

(viii) Fort Lauderdale

In Fig. 42 we show the scatter of the diurnal amplitude for each of the four seasons for Ft. Lauderdale. The correlation between the two datasets is highest in the spring season (0.54) and least in the fall (0.37). In the winter and summer seasons the correlations are 0.44 and 0.44, respectively. Generally, across all seasons, IMERG tends to overestimate the diurnal amplitude. The diurnal amplitude is the least in the winter and largest in the summer followed by that in the fall season. The timing of the diurnal peak of rainfall shows very poor fidelity in IMERG across all four seasons (Fig. 43).

In summary, IMERG displays reasonable fidelity of the diurnal amplitude of rainfall but comparatively much poor skills in the timing of the diurnal peak of rainfall across all stations. Over Alachua, the diurnal amplitude shows the poorest fidelity for IMERG compared to the remaining 7 stations.

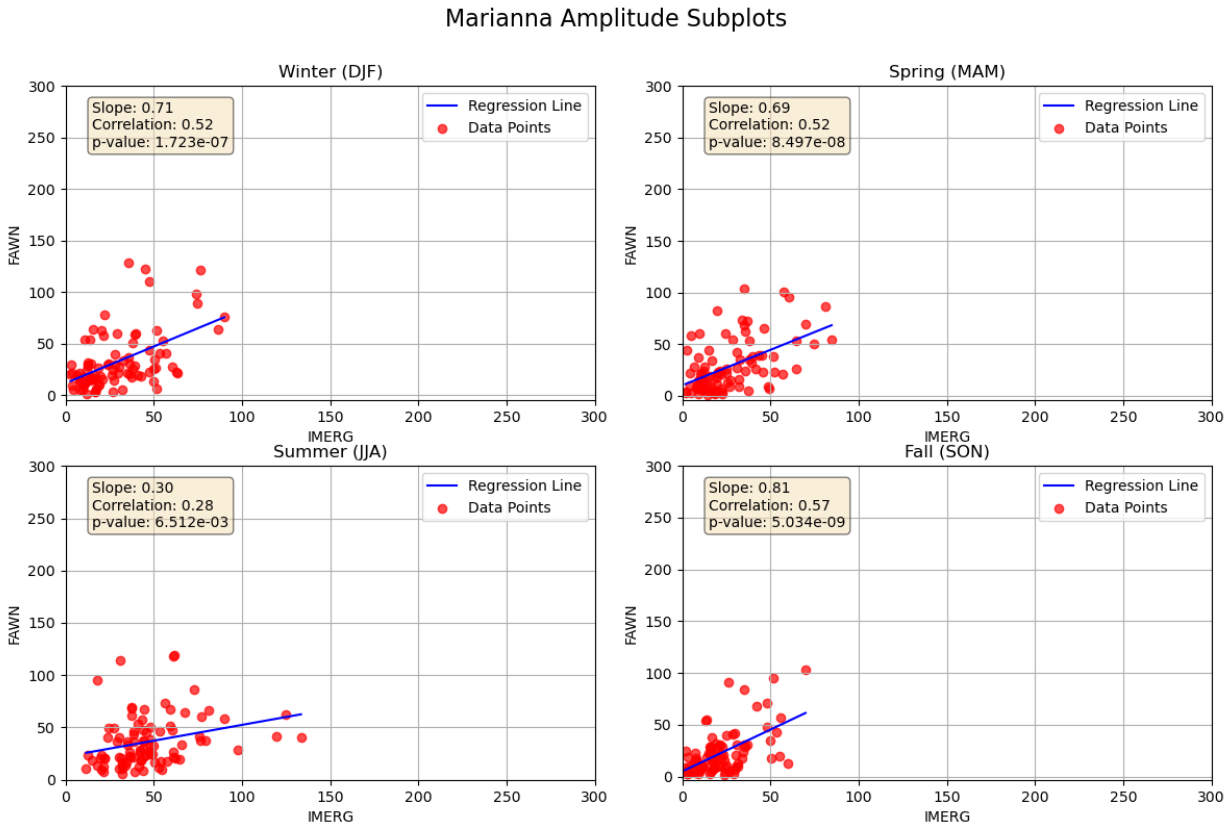

Figure 28: The scatter of the diurnal amplitude (mm/day) between FAWN and IMERG for Marianna for the four seasons of a) DJF, b) MAM, c) JJA, and d) SON.

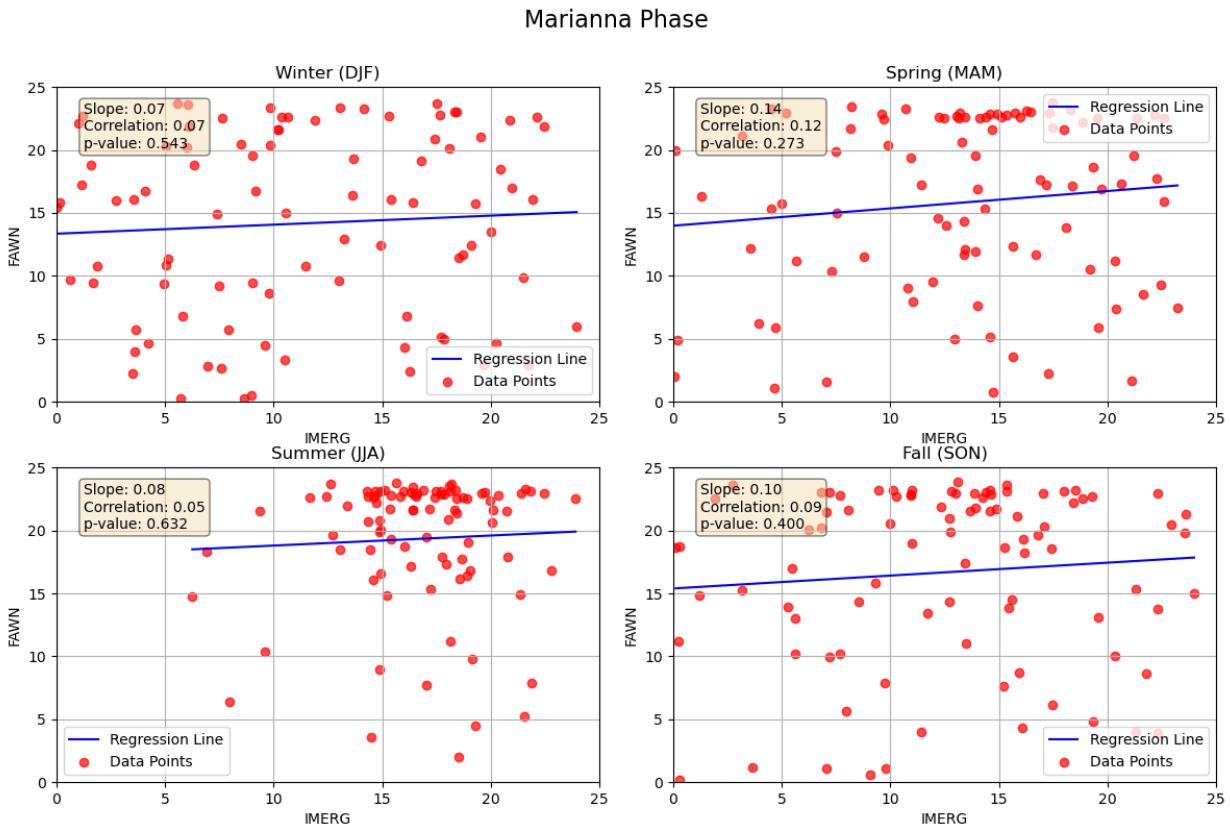

Figure 29: The scatter of the phase of the diurnal peak (local solar time) between FAWN and IMERG for Marianna for the four seasons of a) DJF, b) MAM, c) JJA, and d) SON.

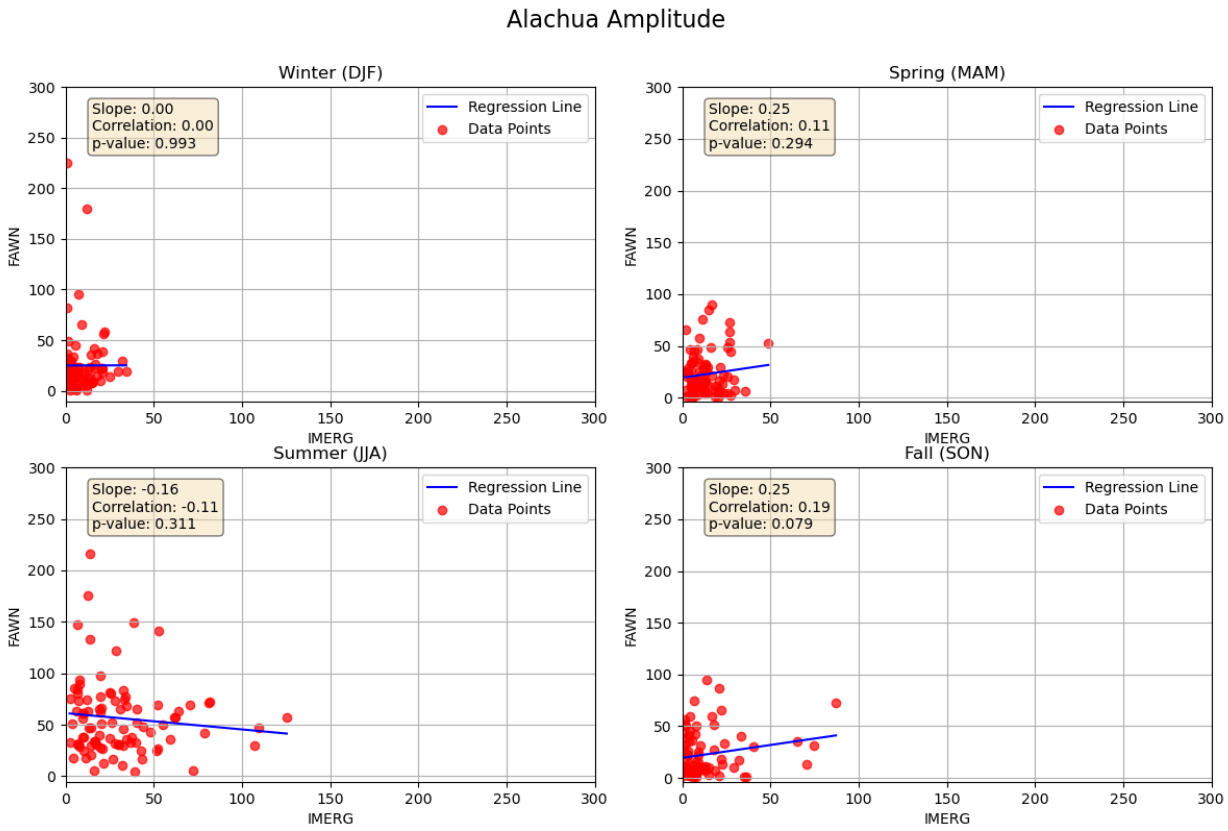

Figure 30: The scatter of the diurnal amplitude (mm/day) between FAWN and IMERG for Alachua for the four seasons of a) DJF, b) MAM, c) JJA, and d) SON.

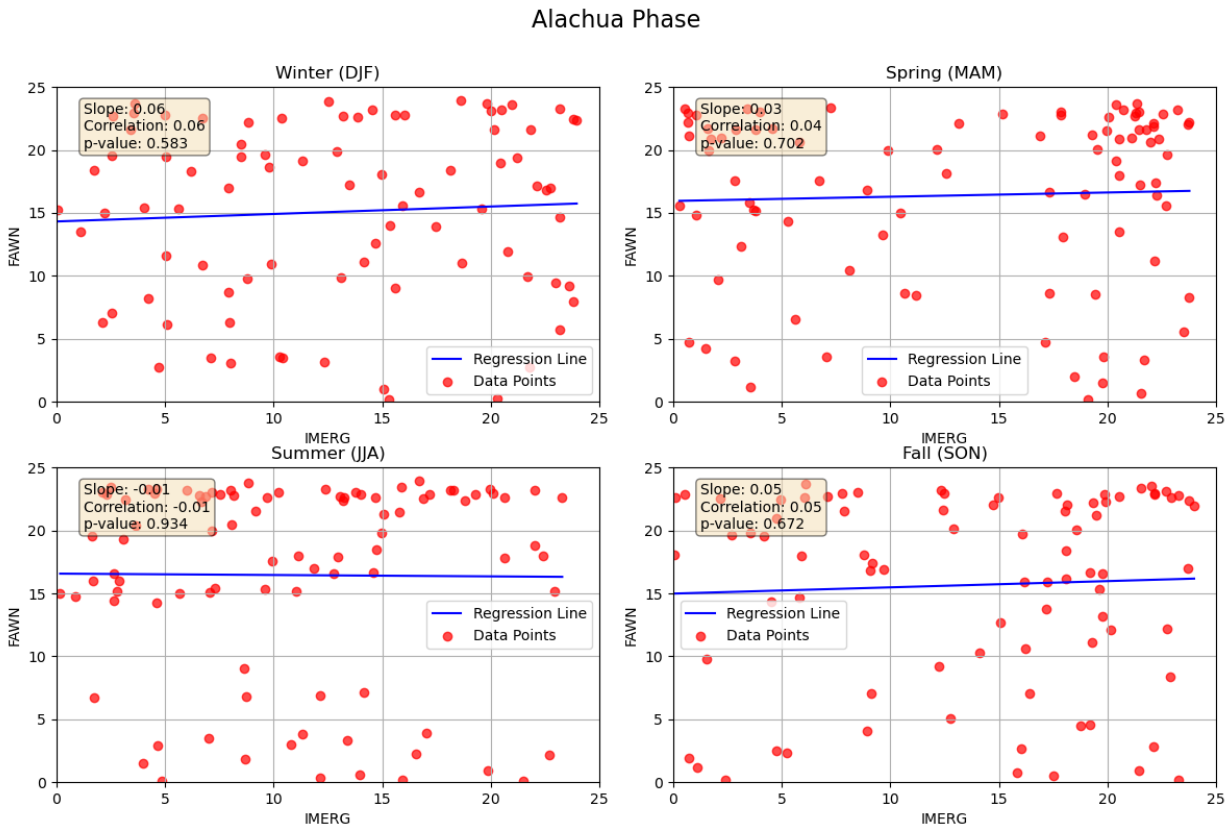

Figure 31: The scatter of the phase of the diurnal peak (local solar time) between FAWN and IMERG for Alachua for the four seasons of a) DJF, b) MAM, c) JJA, and d) SON.

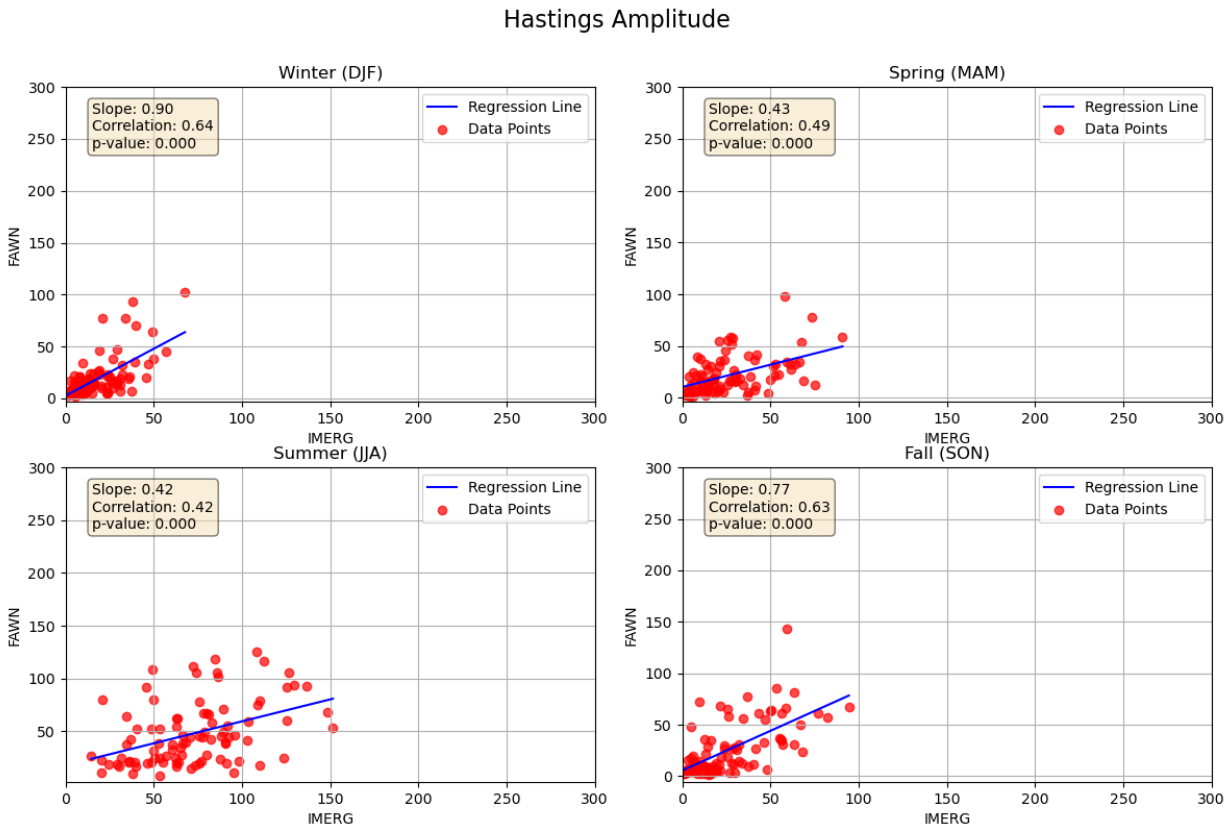

Figure 32: The scatter of the diurnal amplitude (mm/day) between FAWN and IMERG for Hastings for the four seasons of a) DJF, b) MAM, c) JJA, and d) SON.

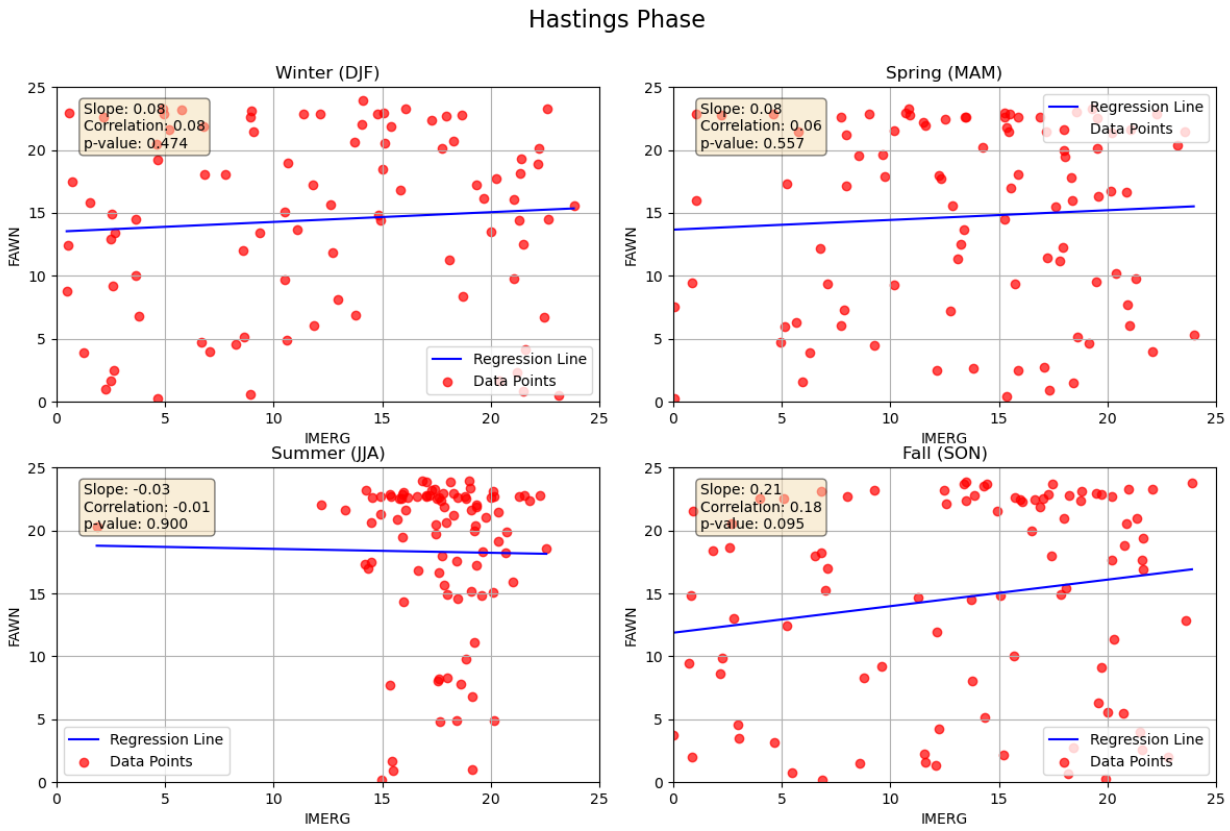

Figure 33: The scatter of the phase of the diurnal peak (local solar time) between FAWN and IMERG for Hastings for the four seasons of a) DJF, b) MAM, c) JJA, and d) SON.

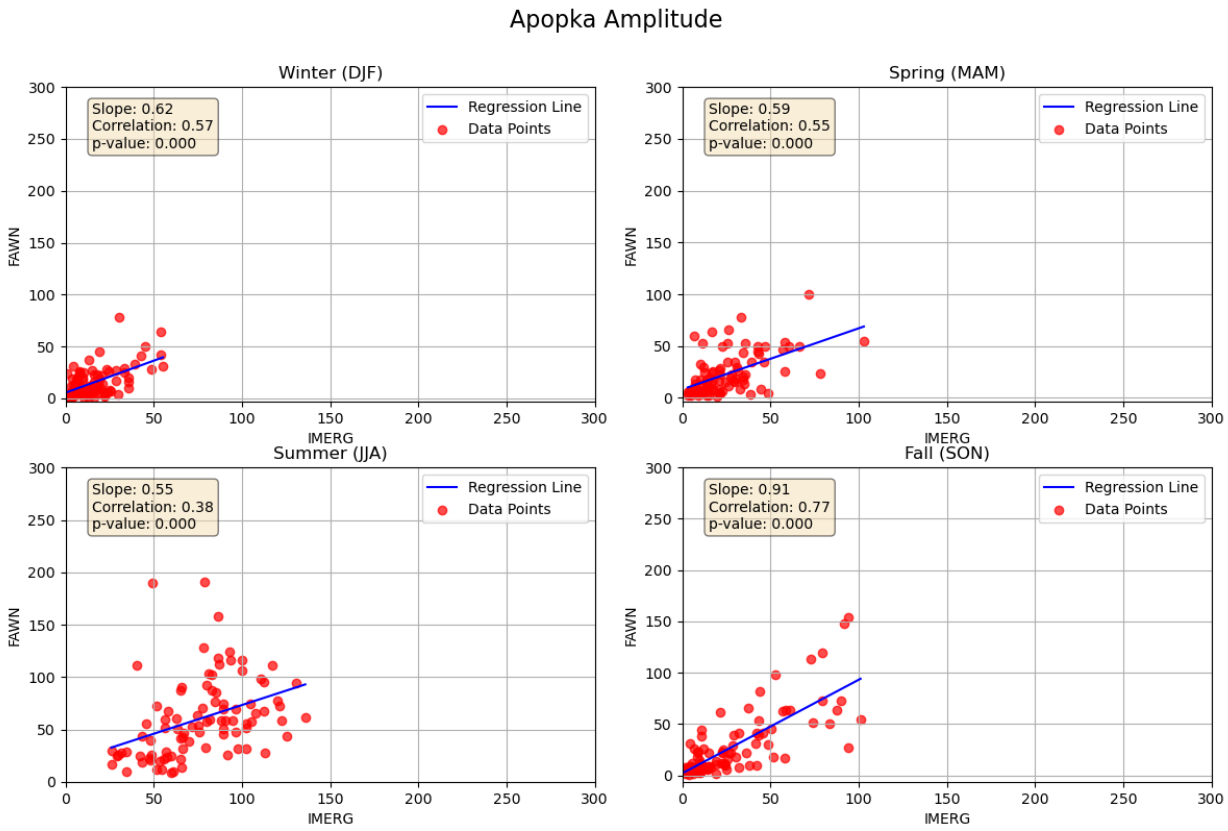

Figure 34: The scatter of the diurnal amplitude (mm/day) between FAWN and IMERG for Apopka for the four seasons of a) DJF, b) MAM, c) JJA, and d) SON.

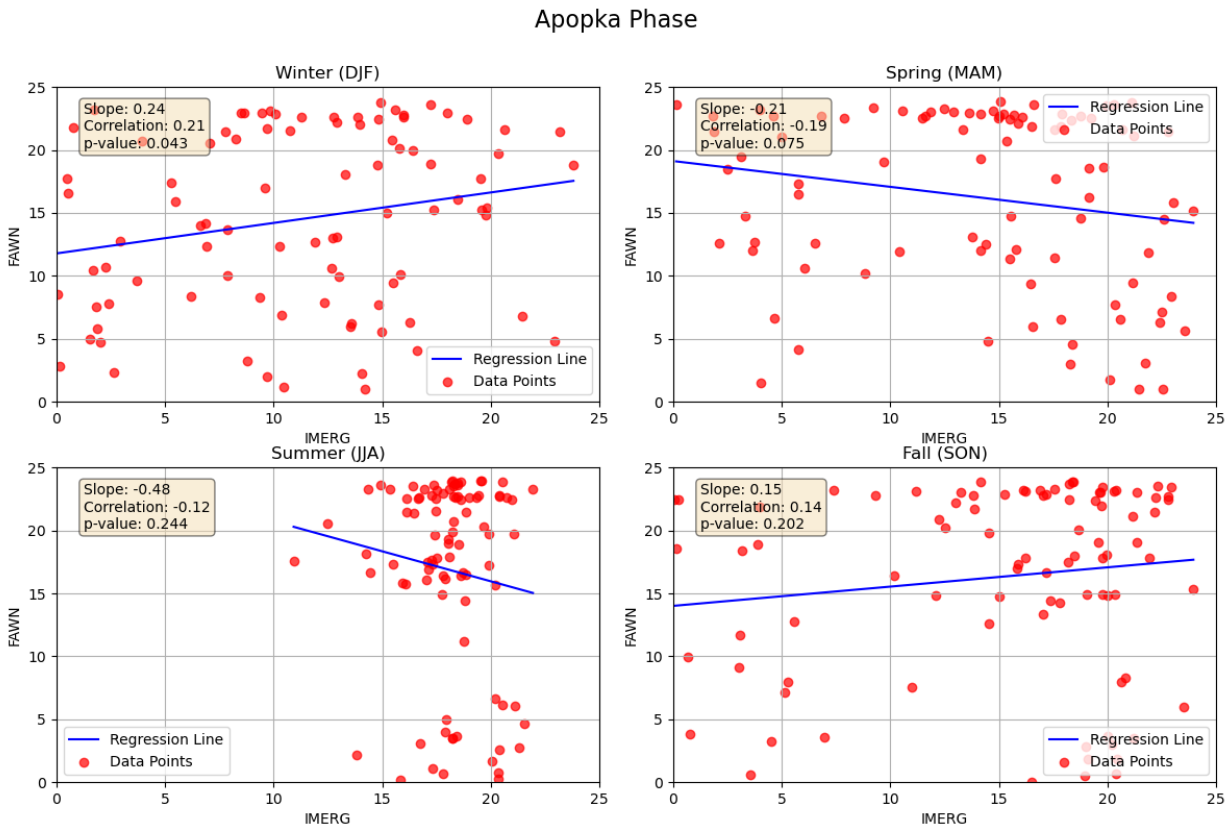

Figure 35: The scatter of the phase of the diurnal peak (local solar time) between FAWN and IMERG for Apopka for the four seasons of a) DJF, b) MAM, c) JJA, and d) SON.

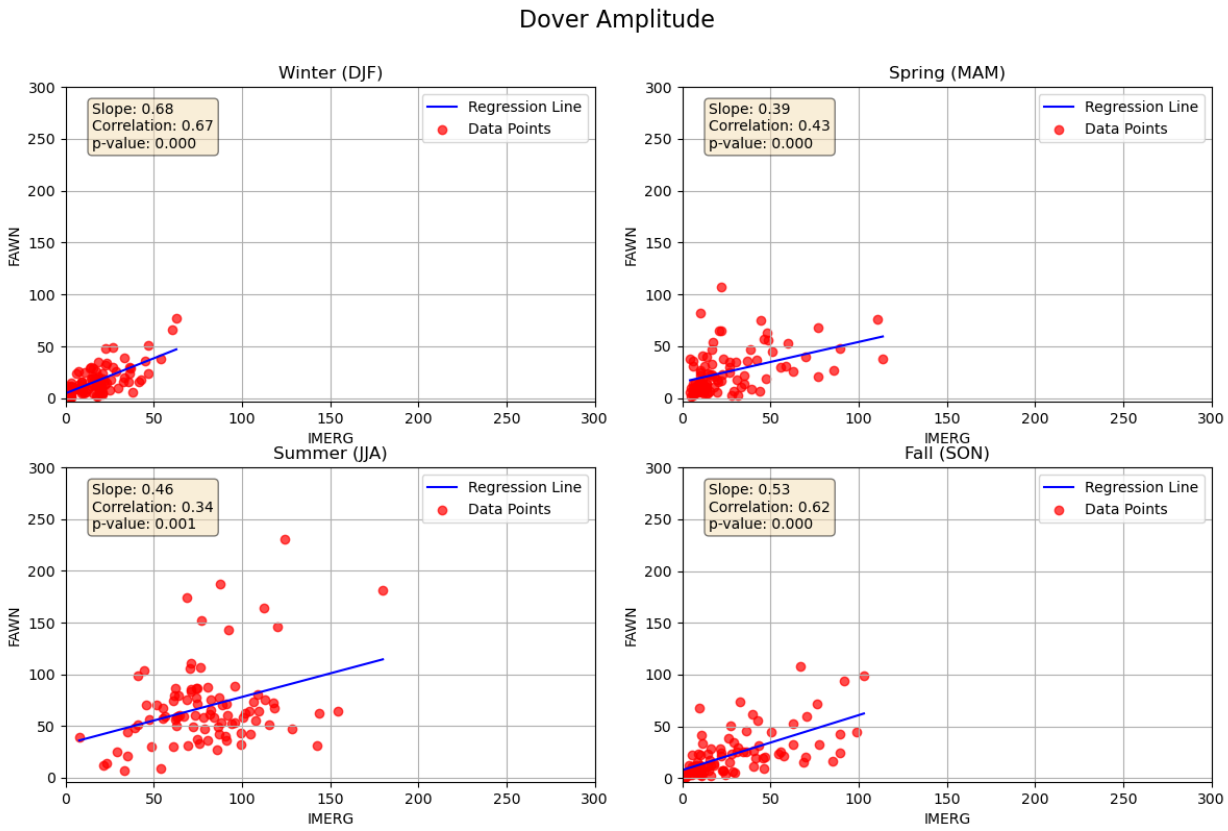

Figure 36: The scatter of the diurnal amplitude (mm/day) between FAWN and IMERG for Dover for the four seasons of a) DJF, b) MAM, c) JJA, and d) SON.

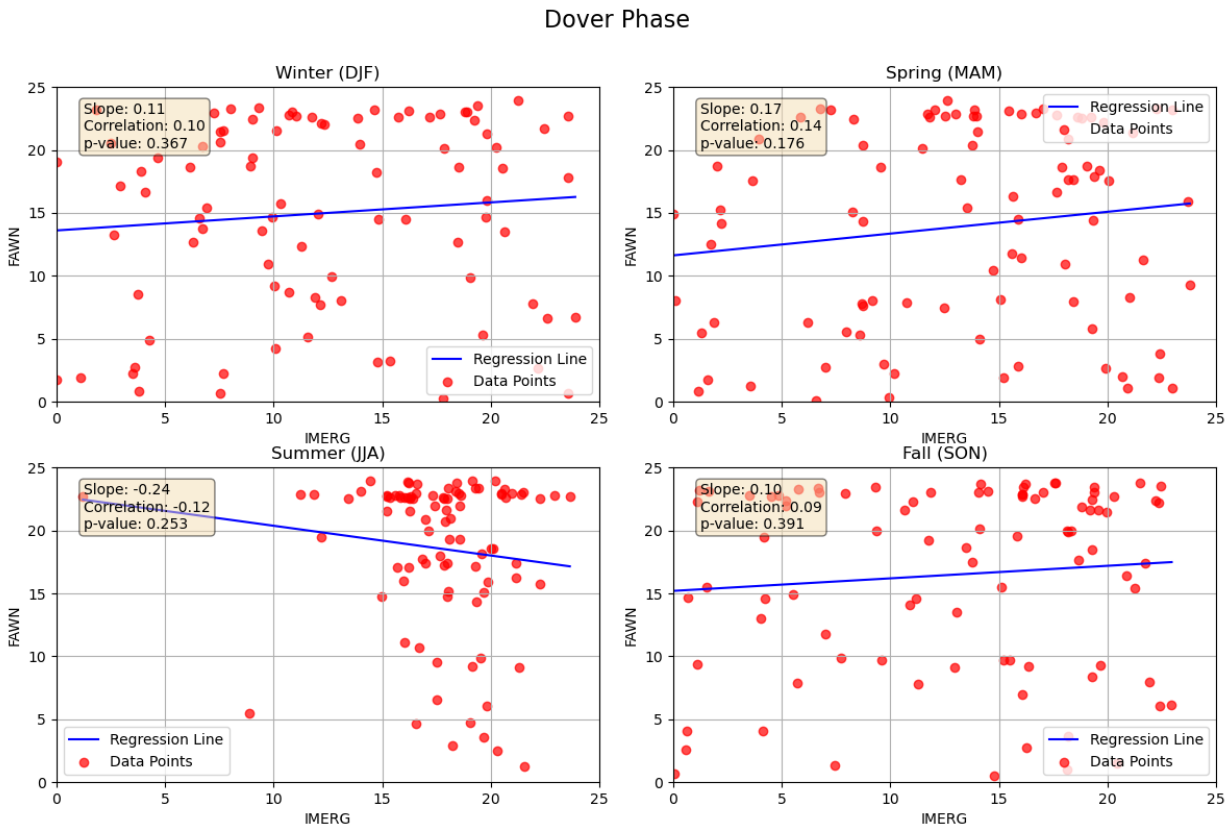

Figure 37: The scatter of the phase of the diurnal peak (local solar time) between FAWN and IMERG for Dover for the four seasons of a) DJF, b) MAM, c) JJA, and d) SON.

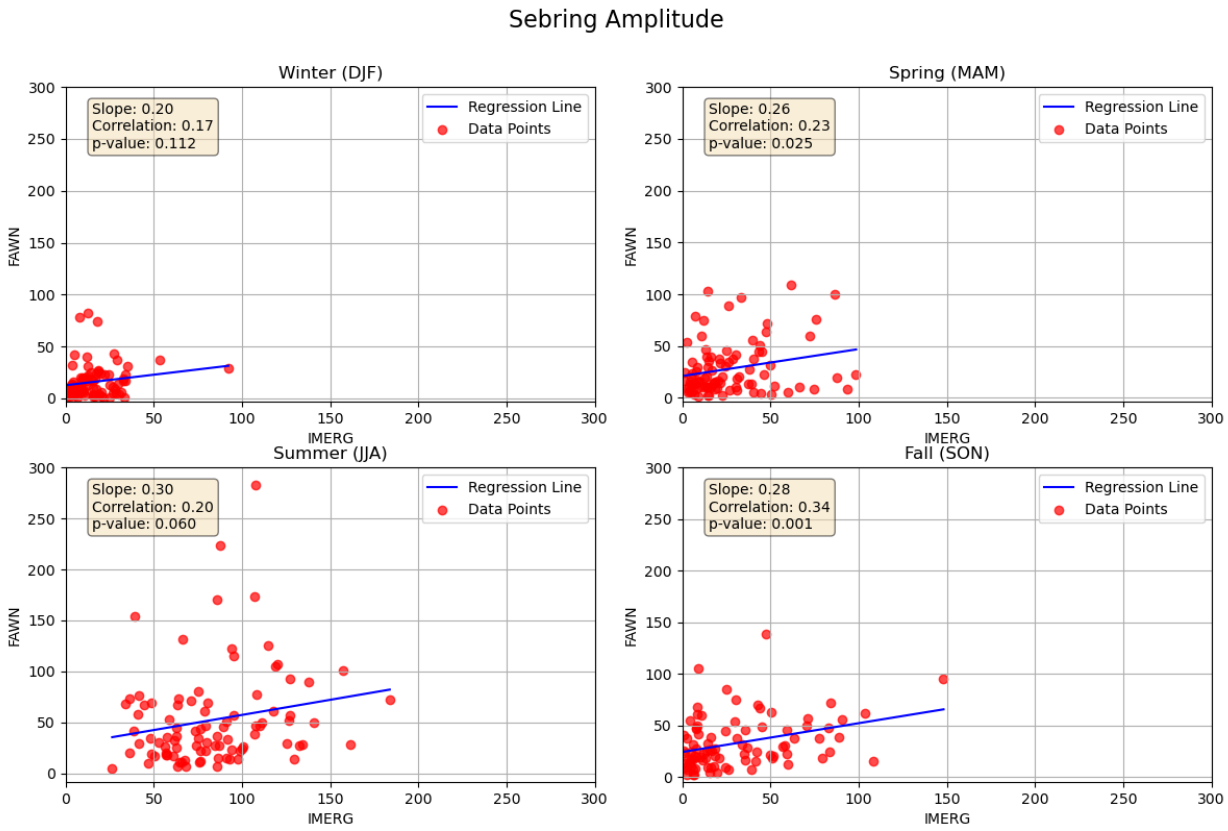

Figure 38: The scatter of the diurnal amplitude (mm/day) between FAWN and IMERG for Sebring for the four seasons of a) DJF, b) MAM, c) JJA, and d) SON.

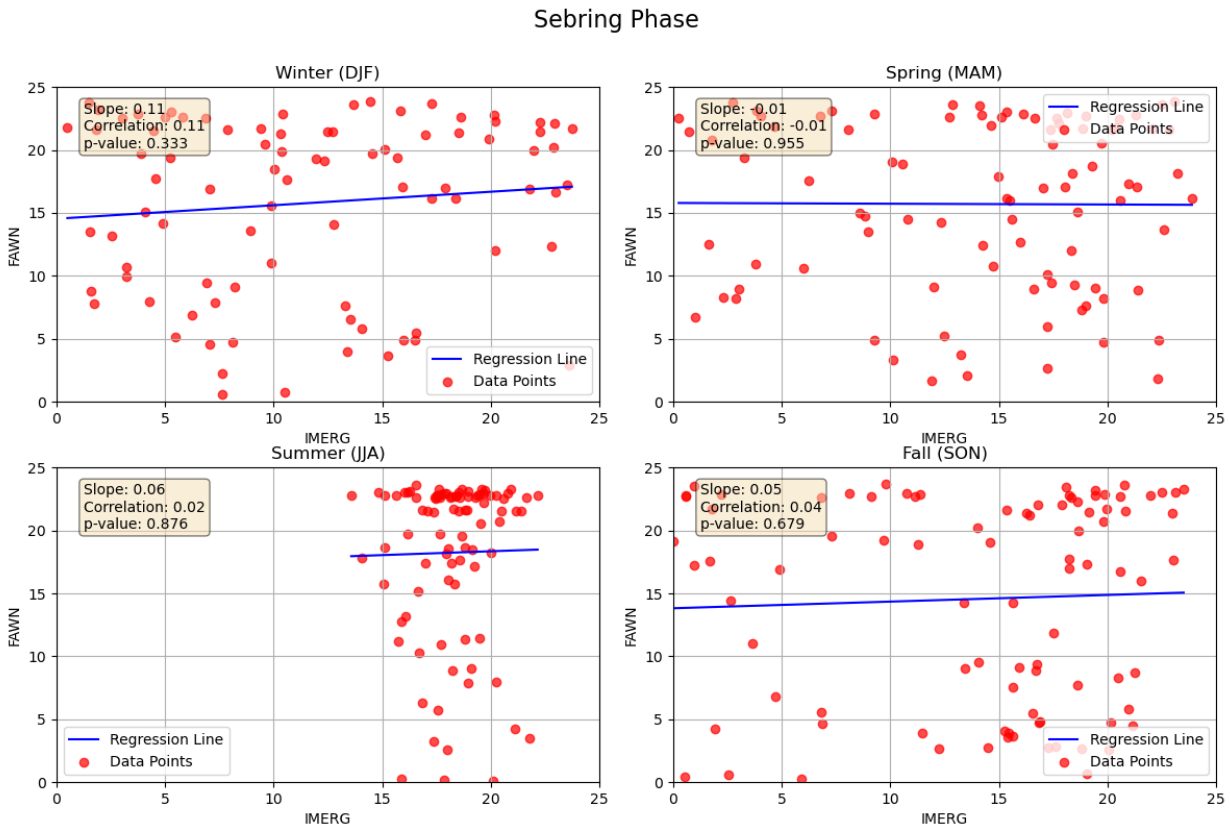

Figure 39: The scatter of the phase of the diurnal peak (local solar time) between FAWN and IMERG for Sebring for the four seasons of a) DJF, b) MAM, c) JJA, and d) SON.

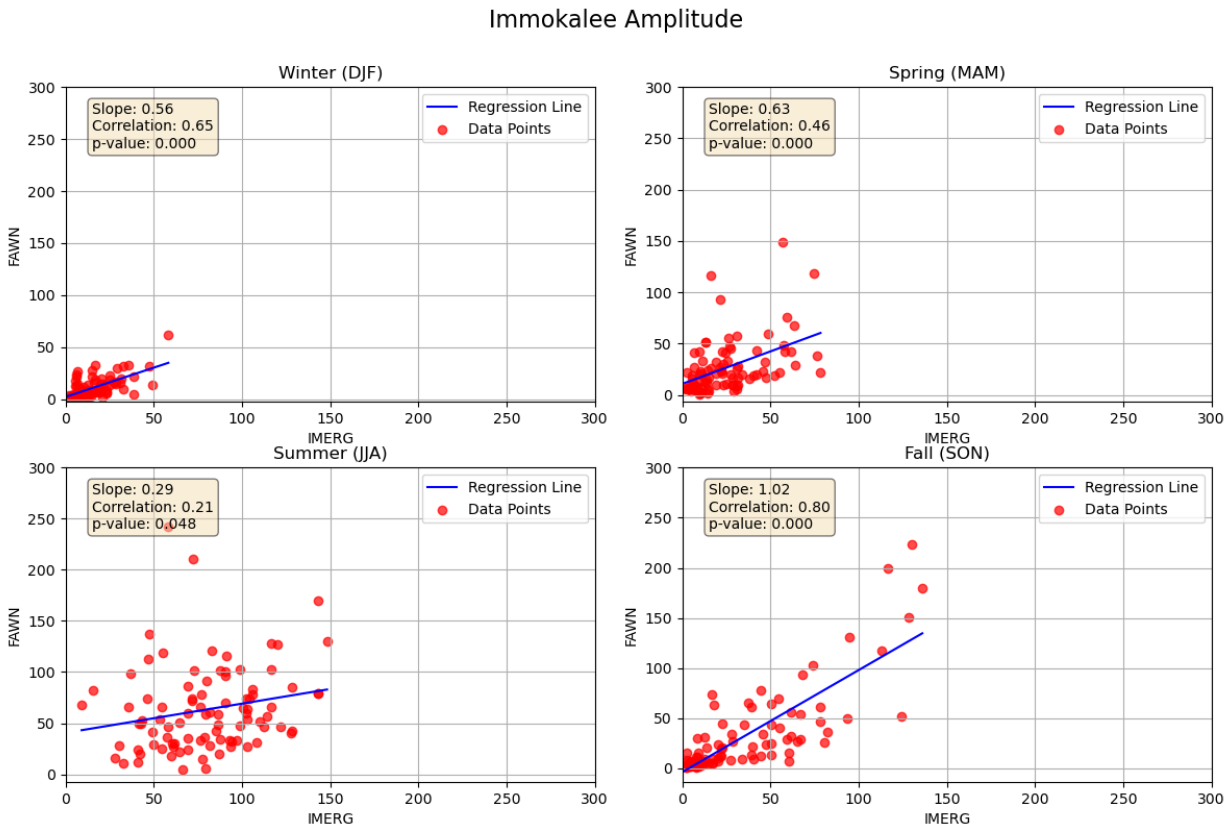

Figure 40: The scatter of the diurnal amplitude (mm/day) between FAWN and IMERG for Immokalee for the four seasons of a) DJF, b) MAM, c) JJA, and d) SON.

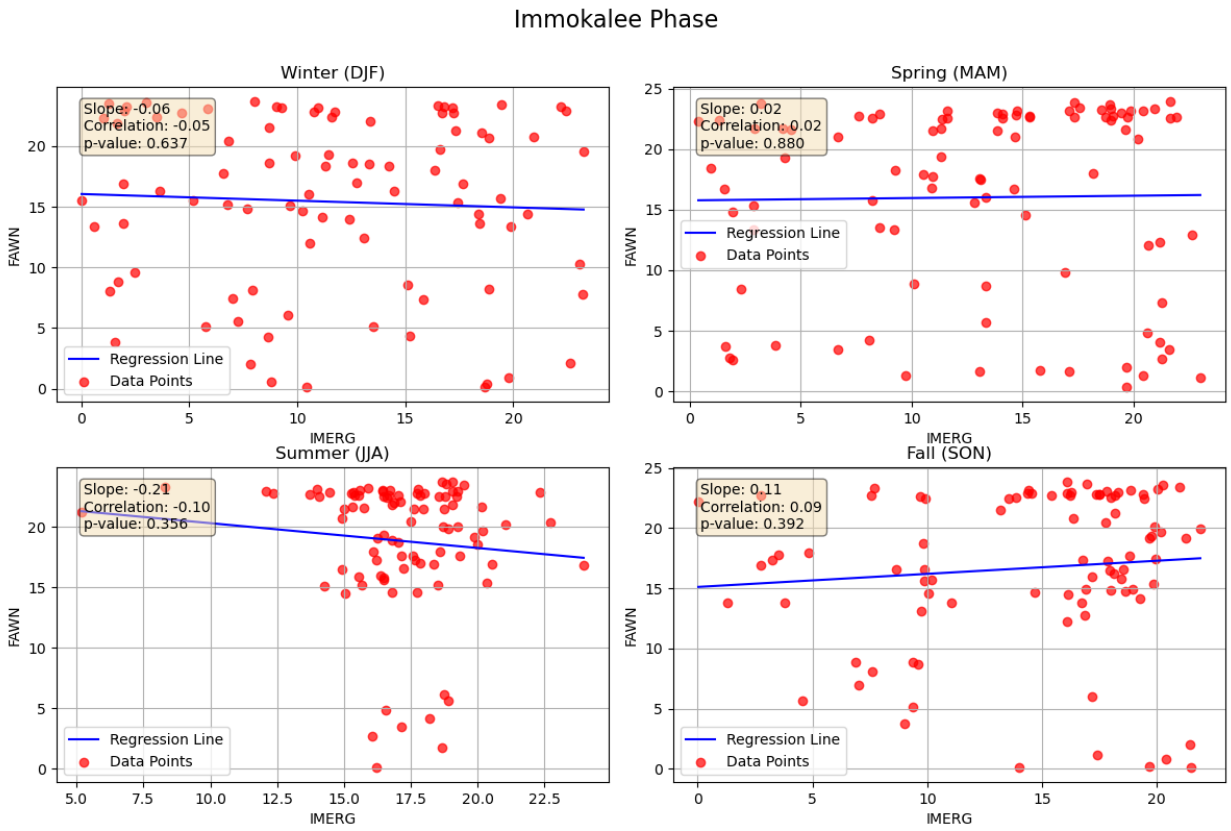

Figure 41: The scatter of the phase of the diurnal peak (local solar time) between FAWN and IMERG for Immokalee for the four seasons of a) DJF, b) MAM, c) JJA, and d) SON.

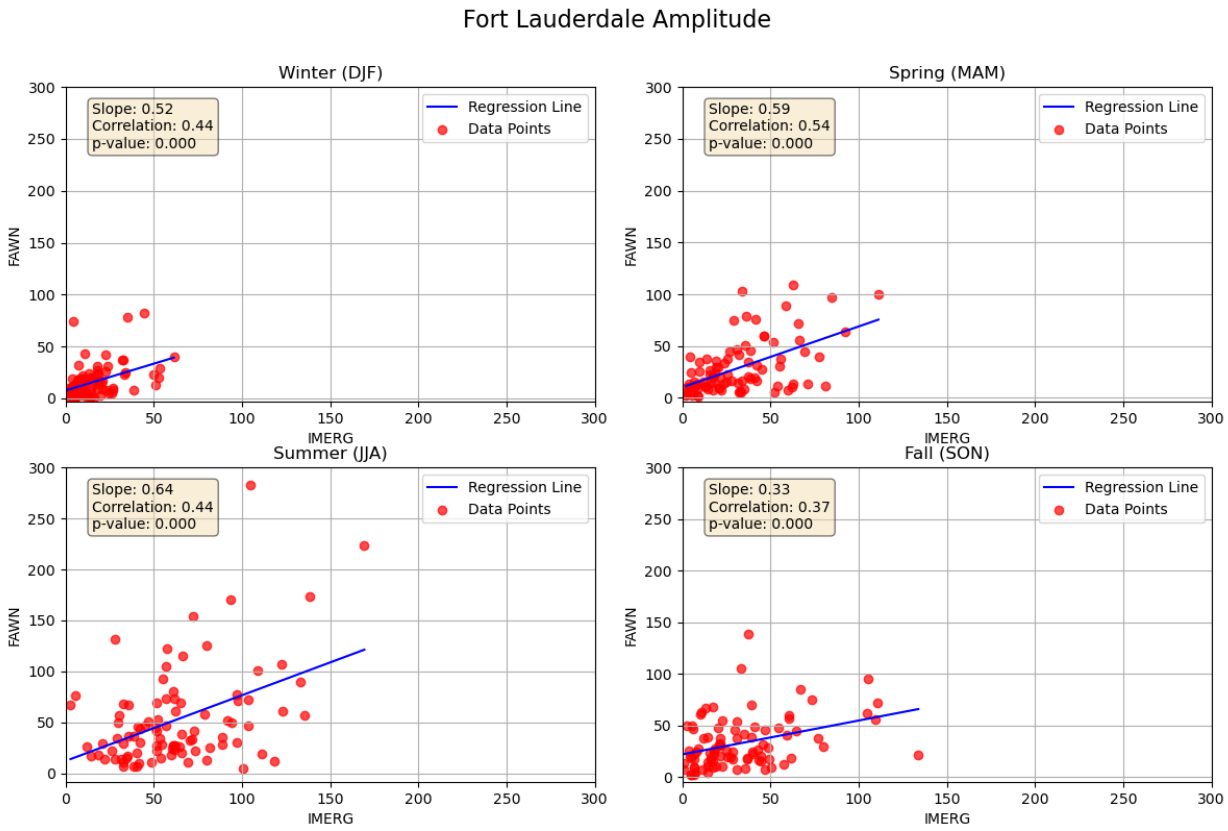

Figure 42: The scatter of the diurnal amplitude (mm/day) between FAWN and IMERG for Ft. Lauderdale for the four seasons of a) DJF, b) MAM, c) JJA, and d) SON.

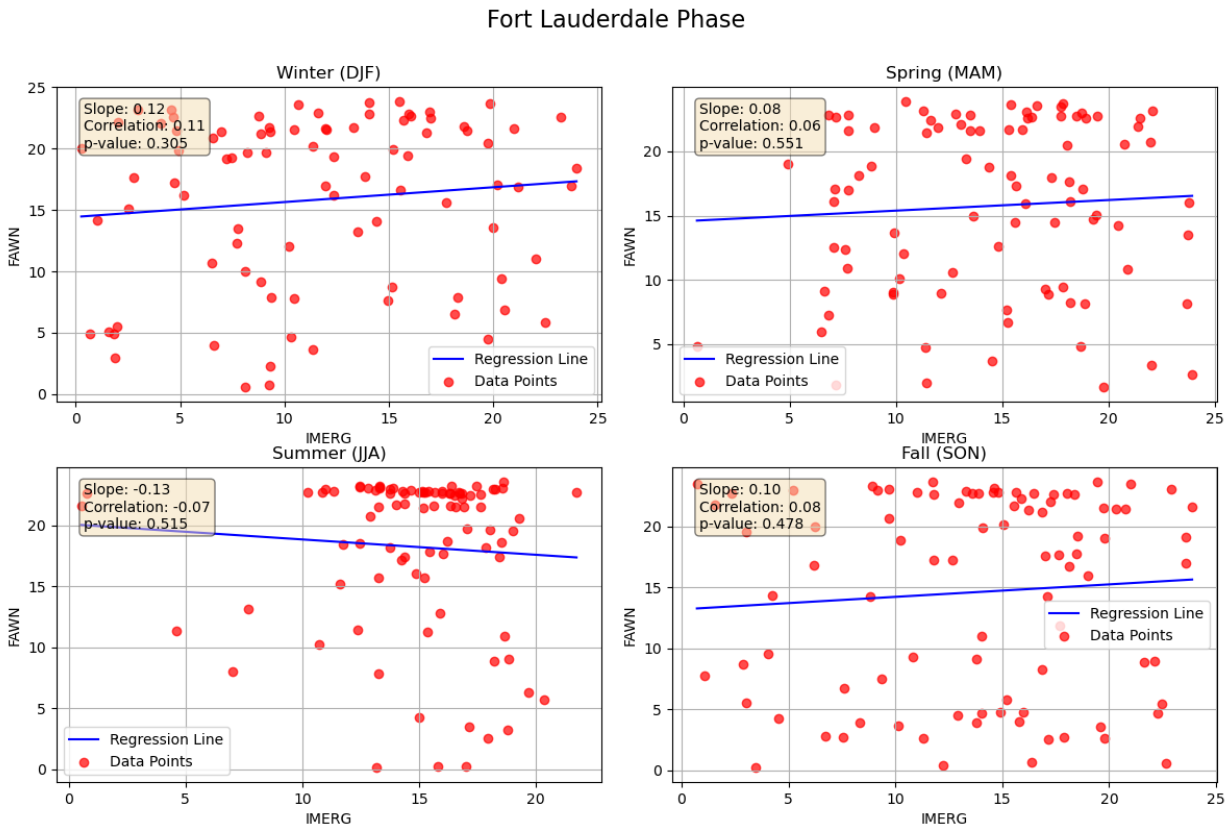

Figure 43: The scatter of the phase of the diurnal peak (local solar time) between FAWN and IMERG for Ft. Lauderdale for the four seasons of a) DJF, b) MAM, c) JJA, and d) SON.

4 Conclusions

This study has conducted an extensive verification of remotely sensed rainfall analysis of NASA called the Integrated Multi-satellite Retrievals for Global Precipitation Mission version 7 (IMERGv7) with in situ observations of rainfall collected under the Florida Automated Weather Network (FAWN). The FAWN rainfall observations are at 15 min temporal resolution while that of IMERGv7 rainfall analysis is at 30 minutes interval. The verification was conducted over an overlapping period of 16 years (2005-2020), when all the data gaps were filled and quality controlled for the FAWN observations. In this study, we selected 8 stations for verification out of the thirty available FAWN stations. These 8 stations covered north Florida (Marianna, Hastings, Alachua, and Apopka) and south Florida (Dover, Sebring, Immokalee, and Ft. Lauderdale). Furthermore, amongst these 8 FAWN stations, based on their proximity to the coasts were considered Gulf Coast for Marianna, Alachua, Dover, and Immokalee while Hastings, Apopka, Sebring, and Ft. Lauderdale were considered on the Atlantic coast of Florida. This comparison allowed for a comprehensive analysis of the accuracy and reliability of NASA's rainfall data which is part of its Global Precipitation Mission, providing valuable insights into the effectiveness of remotely sensed rainfall analysis in conjunction with ground-based observations.

The verification was conducted separately for interannual (year-to-year), sub-seasonal (10-90 days), synoptic (3-10 days), and diurnal (< 1 day) across these 8 stations. These

scales were separated in the time series using the data-adaptive technique called the Ensemble Empirical Mode Decomposition (EEMD). Our findings on IMERGv7 verification with FAWN are as follows:

1. The synoptic variance is highest across all stations in both IMERG and FAWN followed by that in interannual scales. The sub-seasonal variance is the least. In majority of the stations the synoptic variance is highest in the winter but some stations in south Florida display the highest synoptic variance in the summer in both IMERG and FAWN. The seasonal variability of the synoptic variance is comparable in IMERG with FAWN. The central Florida region displays the least synoptic variance in both datasets. The temporal correlations between FAWN and IMERG is in the range of 0.6 to 0.73.

2. The subseasonal variance is the least across most stations, about 1/3rd the variance of the synoptic scales. There is no discernible spatial patterns. The temporal correlations between FAWN and IMERG is in the range of 0.6 to 0.82. They are higher than at synoptic scales.

3. The verification at the interannual scales yielded the highest temporal correlations between IMERG and FAWN, in the range of 0.71 to 0.92. The interannual variance was also ~ 1/3rd that of synoptic variance but slightly higher than subseasonal variance. There was no discernible spatial pattern of the interannual variance.

4. The climatological diurnal amplitude shows reasonable and comparable fidelity across all stations. However, the verification of the diurnal phase is sobering. Furthermore, IMERG has a tendency to overestimate the diurnal amplitude. But the seasonal variations of the diurnal variability largely match with FAWN.

5. Alachua station shows the poorest verification at all time scales. The source of this error needs to be further investigated.

References

Bastola, S., & Misra, V. (2013). Seasonal hydrological forecasts for watersheds over Florida. *Hydrology and Earth System Sciences, 17*(4), 1401–1410. https://doi.org/10.5194/hess-17-1401-2013

Chan, S. C., & Misra, V. (2010). A diagnosis of the 1979–2005 extreme rainfall events in the southeastern United States using the NARR dataset. *Climate Dynamics, 36*(3–4), 911–923. https://doi.org/10.1007/s00382-010-0792-y

Huffman, G.J., D.T. Bolvin, R. Joyce, E.J. Nelkin, J. Tan, D. Braithwaite, K. Hsu, O.A. Kelley, P. Nguyen, S. Sorooshian, D.C. Watters, B.J. West, P. Xie, 2023: Algorithm Theoretical Basis Document (ATBD) NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG) Version 07. PPS, last updated 12 July 2023. 52 pp. https://gpm.nasa.gov/sites/default/files/2023-07/IMERG_V07_ATBD_final_230712.pdf

Huffman, G.J., D.T. Bolvin, R. Joyce, O.A. Kelley, E.J. Nelkin, A. Portier, E.F. Stocker, J. Tan, D.C. Watters, B.J. West, 2024: IMERG V07 Release Notes. PPS, last updated 26 November 2024, 17 pp. https://gpm.nasa.gov/resources/documents/imerg-v07-release-notes

Julie, P., Misra, V., Peeling, R., & Bhardwaj, A. (2023). Gap-free 16-year (2005–2020) sub-diurnal surface meteorological observations across Florida. *Florida Climate Institute*, 1–10.

Li, Y., Misra, V., Carlson, T. N., & Kumar, A. (2013). Predictability of seasonal precipitation over the southeast United States. *Journal of Climate, 26*(1), 108–122. https://doi.org/10.1175/JCLI-D-12-00077.1

Lusher, J., Clark, J., & Zierden, D. (2008). Florida Automated Weather Network (FAWN). *University of Florida IFAS Extension*. https://fawn.ifas.ufl.edu/

Misra, V., Bastola, S., & Chan, S. C. (2011). Seasonal predictability of rainfall over the southeast United States. *Climate Dynamics, 36*(9), 1991–2007. https://doi.org/10.1007/s00382-010-0887-5

Misra, V., Jayasankar, C. B., Beasley, P., & Bhardwaj, A. (2022). Operational monitoring of the evolution of rainy season over Florida. *Frontiers in Climate, 4*, 1–25. https://www.coaps.fsu.edu/~vmisra/operational.pdf

Misra, V., & DiNapoli, S. (2013). The sensitivity of the southeast US summer precipitation to Atlantic warm pool variability. *Climate Dynamics, 40*(11–12), 2941–2959. https://doi.org/10.1007/s00382-012-1474-4

Misra, V., Jayasankar, C. B., Bhardwaj, A., & Bastola, S. (2018). Rainy season characteristics over the state of Florida. *Climate Dynamics, 51*(5), 1863–1877. https://doi.org/10.1007/s00382-017-3971-1

Peeling, R., Julie, P., Misra, V., & Bhardwaj, A. (2023). Gap-free sub-daily meteorological dataset for Florida from 2005 to 2020. *Florida Climate Institute*. (Manuscript in preparation / submitted)

Ropelewski, C. F., & Halpert, M. S. (1986). North American precipitation and temperature patterns associated with the El Niño/Southern Oscillation (ENSO). *Monthly Weather Review, 114*(12), 2352–2362.

Ropelewski, C. F., & Halpert, M. S. (1987). Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation. *Monthly Weather Review, 115*(8), 1606–1626.

Selman, C., Misra, V., & Bastola, S. (2013). Seasonal predictability of precipitation over the southeastern United States using the NCEP Climate Forecast System. *Theoretical and Applied Climatology, 113*(1–2), 55–71. https://doi.org/10.1007/s00704-012-0766-1

Tan, J., Huffman, G. J., Bolvin, D. T., & Nelkin, E. J. (2019). Diurnal cycle of IMERG V06 precipitation. *Geophysical Research Letters, 46*(22), 13584–13592. https://doi.org/10.1029/2019GL085395

Tang, G., Ma, Y., Long, D., Zhong, L., Hong, Y., & Srinivasan, R. (2016). Evaluation of GPM Day-1 IMERG and TMPA Version-7 legacy products over mainland China at multiple spatiotemporal scales. *Journal of Hydrology, 533*, 152–167. https://doi.org/10.1016/j.jhydrol.2015.12.008

Wu, Z., & Huang, N. E. (2009). Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Advances in Adaptive Data Analysis, 1*(1), 1–41. https://doi.org/10.1142/S1793536909000047

Xu, W., Ma, Y., Wei, Z., & Wang, H. (2019). Evaluation of satellite-based precipitation products in the southeastern United States. *Remote Sensing, 11*(1), 47. https://doi.org/10.3390/rs11010047

Let me know if you need this formatted into your Word document or exported to BibTeX, EndNote, or another citation manager!