3-2021

Comparison of Spatial Precipitation Forecasts with a Satellite Dataset

Andrew C. Siebels

COMPARISON OF SPATIAL PRECIPITATION FORECASTS WITH A SATELLITE DATASET

THESIS

Andrew C. Siebels, Captain, USAF

AFIT-ENP-MS-21-M-136

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT-ENP-MS-21-M-136

COMPARISON OF SPATIAL PRECIPITATION FORECASTS WITH A SATELLITE DATASET

THESIS

Presented to the Faculty

Department of Engineering Physics

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Atmospheric Science

Andrew C. Siebels, BS

Captain, USAF

March 2021

AFIT-ENP-MS-21-M-136

COMPARISON OF SPATIAL PRECIPITATION FORECASTS WITH A SATELLITE DATASET

Andrew C. Siebels, BS

Captain, USAF

Committee Membership:

Lt Col Robert C. Tournay, Ph.D.
Chair

Lt Col Hsienliang R. Tseng, Ph.D.
Member

Maj Peter A. Saunders, Ph.D.
Member

AFIT-ENP-MS-21-M-136

Abstract

The purpose of this research is to analyze and compare global precipitation data from the Climate Forecast System Version 2 (CFSv2) with the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN)-Climate Data Record (CDR) to improve long term precipitation forecasting. The CFSv2 has a 0.5-degree resolution which will provide model data for precipitation forecasts. The PERSIANN-CDR is a satellite derived daily 0.25-degree dataset with 37 years of global precipitation coverage 60 N to 60 S. The 0-to-10, 15-to-25, 55-to-65, and 80-to-90 day forecast time frames will then be analyzed for accuracy, and a quantile mapping (QM) technique will be applied to correct precipitation amounts for the CFSv2. The QM procedure requires both training and test datasets from the CFSv2 and PERSIANN-CDR. Finally, the forecast correction results for the CFSv2 may be used to improve medium range precipitation forecasts by the operational meteorological community.

Acknowledgments

I express my sincere gratitude to my advisor, Lt Col Robert Tournay, for guidance along the way with this research. I also would like to thank my fellow students for help with questions and ideas. My wife has also been of enormous support during this research and writing. She kept me motivated to focus on the end goal and remain calm. My mother has been a tremendous support, providing me constant insight as I progressed forward with my thesis. In honor of my grandparents, who provided the love and upbringing to get me to this point in my life. Thank you for everything.

Andrew C. Siebels

Table of Contents

Page

List of Figures

Page

List of Tables

Page

COMPARISON OF SPATIAL PRECIPITATION FORECASTS WITH A SATELLITE DATASET

I. Introduction

General Issue

The Climate Forecast System version 2 (CFSv2) attempts to forecast long range precipitation with reliable results. This research aims to improve the medium-range forecast of the CFSv2 using quantile mapping (QM) from the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Climate Data Record (PERSIANN-CDR) and short range of the CFSv2.

Problem Statement

The CFSv2 is only reliable in the 0-to-10 day range using traditional forecast and verification approaches. The CFSv2 is the current climate model that is run at the National Centers for Environmental Prediction (NCEP) for operational forecasting (Yuan et al., 2013). In order to support there is a need for longer range outlooks upon which risk-based assessments and decisions are made. Using statistical methods and the satellite-derived data, PERSIANN-CDR, we will attempt to apply corrections to sub-seasonal CFSv2 precipitation forecasts towards creating valuable long-range outlooks.

Precipitation measurements on climatological time scales are essential to understanding environmental processes at a sub-seasonal time. Precipitation patterns at these time scales are important to analyze on a global scale to accurately determine the location of events such as floods and droughts. Climatology of hydrologic events, including flooding, influences Department of Defense (DoD) planning and execution

(Gangrade et al., 2020). An example of this is the placement of operating locations. It is necessary for operational decisions to be based on accurate scientific information, which is a driving factor in this research. The runway flooding at Offutt AFB in 2019 brought to light how impactful these events can be (Gangrade et al., 2020). Many overseas locations deal with flooding as well, especially in the Korean Peninsula, which impacts the full range of operations when the monsoon arrives every summer.

The opposite climate extreme is droughts from lack of significant rainfall for a long duration, sometimes lasting years. All of these real-world implications show the importance of forecasting and analyzing precipitation across the globe. Another critical piece is understanding how quickly precipitation falls in a particular area, whether from a severe thunderstorm in the Great Plains of the U.S. or monsoonal rains in India (Trenberth et al., 2003). The Southwest monsoon in the U.S. delivers much more rain than a thunderstorm and on a much longer time scale. Although much research has been accomplished, the challenge of forecasting precipitation accurately remains.

II. Literature Review

Chapter Overview

This chapter gives a background understanding of the current state of satellite-derived datasets and climate forecast models. Various techniques for precipitation analysis were also reviewed in previous research.

PERSIANN-CDR

The PERSIANN-CDR is a relatively new satellite dataset (Ashouri et al., 2015). The PERSIANN-CDR dataset is of sufficient length for climatological studies and for enabling both statistical climate outlooks, dynamical climate modeling refinement and

postprocessing (Ashouri et al., 2015). PERSIANN-CDR allows for higher resolution comparisons when looking at model forecasts of differing lengths.

According to the World Meteorological Organization (WMO), 30 years of historical weather data is required for climatological studies to be conducted (Ashouri et al., 2015). Once the requirements meet with the various data sets, various climate agencies implement the climatological research's long-standing records. PERSIANN-CDR has been suitable for estimating spatial and temporal precipitation globally from 1983 to present (Ashouri et al., 2015). The resolution and global coverage of the PERSIANN-CDR were compared to other satellite derived precipitation datasets (Figure 1).

Product	Temporal resolution	Spatial resolution	Period	Coverage
GPCP	Monthly/Pentad	2.5°	1979–(delayed) present	90°S–90°N
GPCP-1DD	Daily	1°	1996–(delayed) present	90°S–90°N
CMAP	Monthly/Pentad	2.5°	1979–(delayed) present	90°S–90°N
TMPA v7	3 hourly	0.25°	1998–(delayed) present	50°S–50°N
CMORPH	0.5 h	~0.07°*	2002–present	60°S–60°N
PERSIANN	0.5 h	0.25°	2000–present	60°S–60°N
PERSIANN-CCS	0.5 h	0.04°	2003–present	60°S–60°N
PERSIANN-CDR	**Daily**	**0.25°**	**1983–(delayed) present**	**60°S–60°N**

* CMORPH resolution is ~8 km.

Figure 1. Temporal and spatial resolution of global and near-global precipitation datasets (Ashouri et al., 2015).

The PERSIANN algorithm was developed in 1997 by combining high spatial resolution low earth orbit satellite (LEO) data with high temporal-frequency geostationary satellite data (Nguyen et al., 2018). PERSIANN is based on a multilayer neural feedforward network consisting of two processes (Nguyen et al., 2018). In the first part, the infrared images are transformed into the hidden layer to form the self-organizing feature map (SOFM) through an automatic clustering process (Nguyen et al.,

2018). The second part is when the SOFM detects and classifies patterns in the input data, mapped to the continuous space of outputs such as rainfall (Nguyen et al., 2018). Parameter estimation is used in these processes, incorporating passive microwave rainfall from LEO satellites (Nguyen et al., 2018).

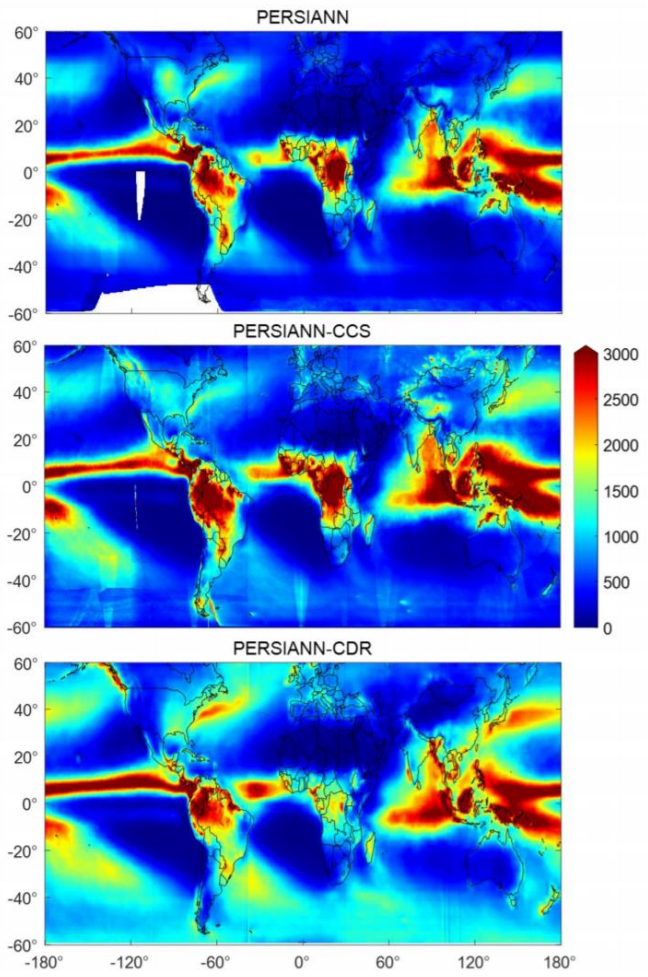

Figure 2. Mean annual precipitation (mm) for PERSIANN, PERSIANN-CCS, and PERSIANN-CDR (Nguyen et al., 2018).

PERSIANN-CDR uses Stage IV hourly precipitation to train the artificial neural network (ANN) model, which allows the algorithm to run with fixed parameters (Lin et al., 2005). Stage IV precipitation is a mosaic of regional multi-sensor analysis (Lin et

al.,2005). The differences in annual mean precipitation between PERSIANN, PERSIANN-Cloud Classification System (CCS), and PERSIANN-CDR is small (Figure 2). The PERSIANN-CCS is another dataset which obtains precipitation amounts from cloud coverage and the PERSIANN is a shorter data record at 17 years compared to 37 from the PERSIANN-CDR (Nguyen et al., 2018). The zonal and meridional rainfall comparisons for the three different PERSIANN satellite products depict the variation of precipitation in latitude and longitude (Figures 3 and 4). These different outputs from the various datasets support the claim that the PERSIANN-CDR can improve the CFSv2 (Figure 3).

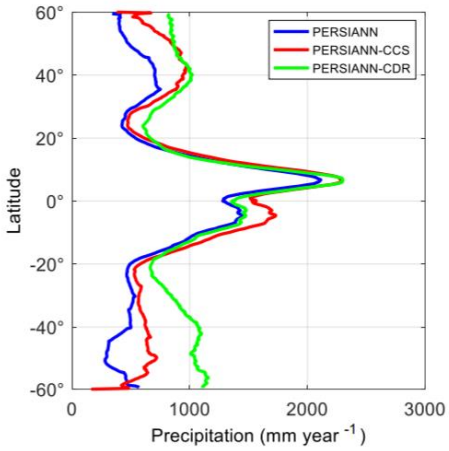

Figure 3. Mean annual zonal precipitation (mm/yr) for PERSIANN, PERSIANN-CCS, and PERSIANN-CDR (Nguyen et al., 2018).

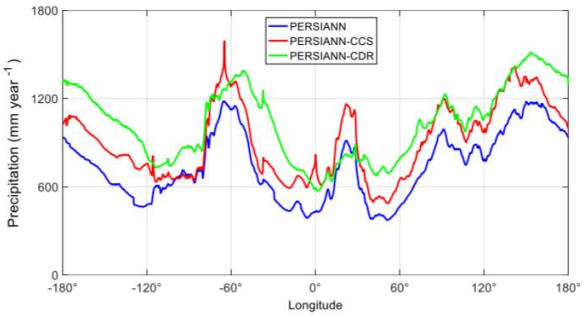

Figure 4. Mean annual meridional precipitation (mm/yr) for PERSIANN, PERSIANN-CCS, and PERSIANN-CDR (Nguyen et al., 2018).

The PERSIANN-CDR was developed to estimate rainfall from geosynchronous satellites every half hour (Sorooshian et al., 2000). Geosynchronous satellites provide extensive coverage of the Earth at high spatial and temporal resolution which aid in detailed retrievals for the PERSIANN-CDR (Sorooshian et al., 2000). Artificial neural network (ANN) models within PERSIANN provide useful information for hydrologic and meteorologic impacts, specifically precipitation (Sorooshian et al., 2000). ANNs are tools for processing precipitation data for specific results similar to the human brain's biological framework with an example of supervised learning (Sorooshian et al., 2000). The combination of high-resolution satellite retrievals and complex neural network processing provide the improved precipitation results for the PERSIANN-CDR

(Sorooshian et al., 2000). Neural network group techniques for the PERSIANN-CDR estimate convective precipitation reduced to less than 10% (Zhang et al., 1994).

The PERSIANN-CDR system operates in two modes: simulation and updates (Sorooshian et al., 2000). The simulation generates regular rainfall rate output, and updates improve the product's quality (Sorooshian et al., 2000). The region of interest Sorooshian et al. (2000) uses for their study includes the Tropical Pacific which analyzes the performance of precipitation rates from the PERSIANN-CDR. There is still a need for improved spatial and temporal accuracy of global precipitation (Sorooshian et al., 2000).

Extreme rain events have been studied using PERSIANN-CDR with indices such as percentile, absolute threshold, and maximum (Miao et al., 2014). Since these events do not occur regularly, it is essential to capture them in climatic or long-term studies. Miao et al. (2014) used five different absolute thresholds to account for correlations between the observations and the PERSIANN-CDR. The evaluation of the PERSIANN-CDR was obtained by using 11 extreme precipitation indices over the China region (Miao et al., 2014). Many different methods have been used in better forecasting these high-end rain events, with little improvement that was noted because of the complexity of parameterization within the models for the long term beyond 30 days.

CFSv2

The CFS was the first quasi-global, fully coupled atmosphere-ocean-land model used at NCEP for seasonal forecasting (Saha et al., 2014). The performance of the CFS has been broken down into seasonal and sub seasonal scales to determine forecast skill previously (Saha et al., 2014). The CFSv2 has limited skill beyond the 30-day mark, but

promising increased skill has been shown for precipitation with longer range forecasts at the seasonal scale (Yuan et al., 2011). The predictive skill of precipitation can provide useful information for flood and drought forecasting (Yuan et al., 2011). Physical processes and data assimilation improvement have compensated for the relatively low precipitation forecasting skill beyond one month (Yuan et al., 2011).

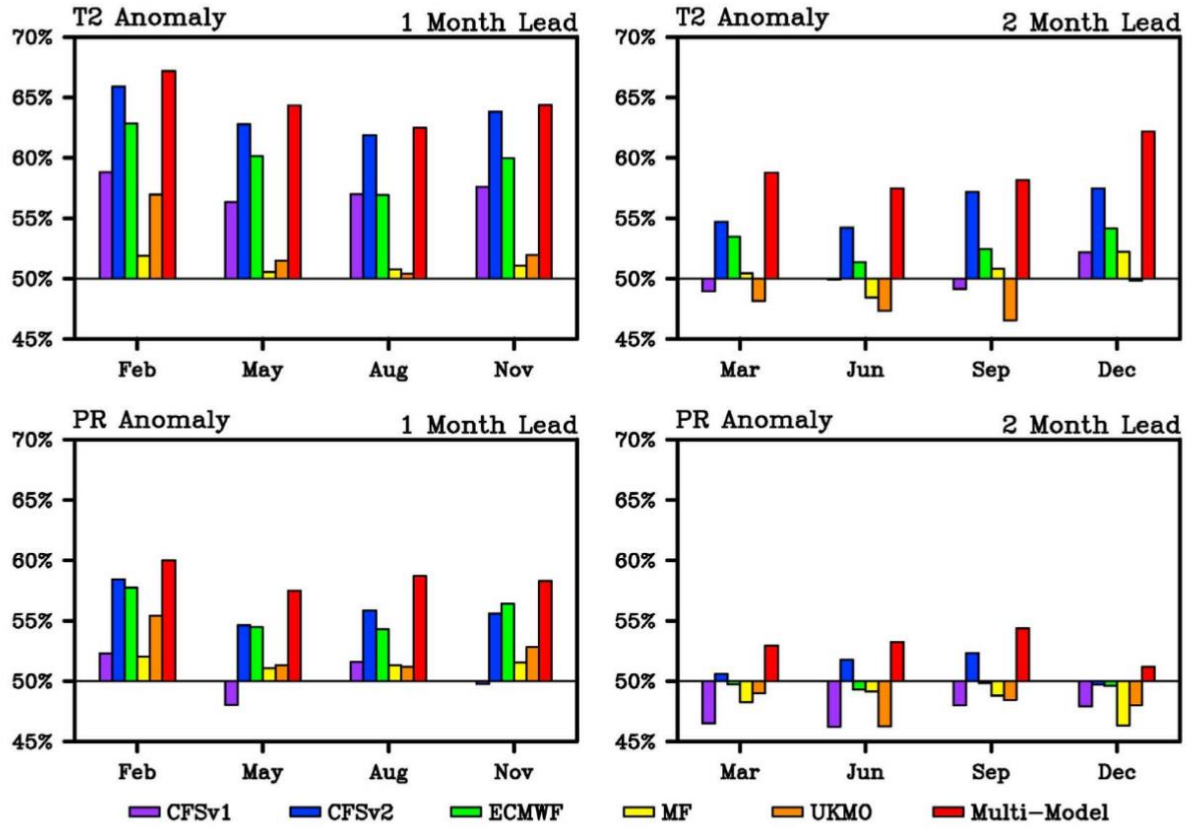

Figure 5. Percentage of positive Ranked Probability Skill Score (RPSS) comparing the CFSv1, CFSv2, ECMWF, MF, UKMO, and Multi-Model for monthly temperature and precipitation over global land areas (Yuan et al., 2011).

The highest predictive skill for precipitation of the CFSv2 appeared to be over the Amazon, Europe, and the Middle East (Yuan et al., 2011). The CFSv2 has outperformed the CFSv1 in temperature and precipitation anomalies at the 1-and-2 month periods which is the reason for the v1 being obsolete (Yuan et al., 2011). Ensembles are

implemented to accurately analyze forecasts with different start dates (Yuan et al., 2011). The Ranked Probability Skill Score (RPSS) produced the forecasts probabilistic quality (Yuan et al., 2011). If the RPSS is equal to one, that meant a near-perfect forecast while a value of zero would indicate that the forecast is inferior to climatology (Yuan et al., 2011). As far as precipitation anomalies go, the CFSv2 compared equally to the ECMWF, with skill considerably dropping beyond 30-days (Yuan et al., 2011). The need for accurate precipitation forecasting is to have real skills in the medium to long-range time scales (Yuan et al., 2011).

A generated dataset previously used was the Climate Forecast System Reanalysis (CFSR) which eliminated fictitious trends caused by model and data assimilation changes (Saha et al., 2010). This dataset was created mainly to simulate initial conditions for the coupled atmosphere (Saha et al., 2010). Artificial reanalysis model data sets were used, such as EMCWAF, JMA, and GFS/NCAR in comparison to the CFSR (Saha et al., 2010). The model-generated precipitation is replaced by observation precipitation for added realism (Saha et al., 2010). Historical and operational archives are required for reanalysis projects; the data came from various sources such as aircraft, satellite, and observations (Saha et al., 2010). The significant improvement was due to increased resolution in the horizontal and vertical directions (Saha et al., 2010). Further analysis of the coupling between the atmosphere and the ocean is essential for further understanding impacts (Saha et al., 2010).

Analysis Techniques

The PERSIANN-CDR dataset is used to assess economic impacts of China's drought conditions, comparing it with 32 years of China monthly Precipitation Analysis

Product (CPAP) data (Guo et al., 2016). The Standardized Precipitation Index (SPI) was used at various time scales between 1 and 12 months for detecting drought events (Guo et al., 2016). Since droughts can last months and years, climatological studies are the most useful for this phenomenon (Guo et al., 2016). The accuracy of the precipitation estimation directly impacted the SPI (Guo et al., 2016). Relative bias and Pearson linear correlation coefficient are error indices used to determine the accuracy of the PERSIANN-CDR in this study (Guo et al., 2016). A comparison between the China monthly Precipitation Analysis Product (CPAP) and the PERSIANN-CDR have been noted (Figure 6).

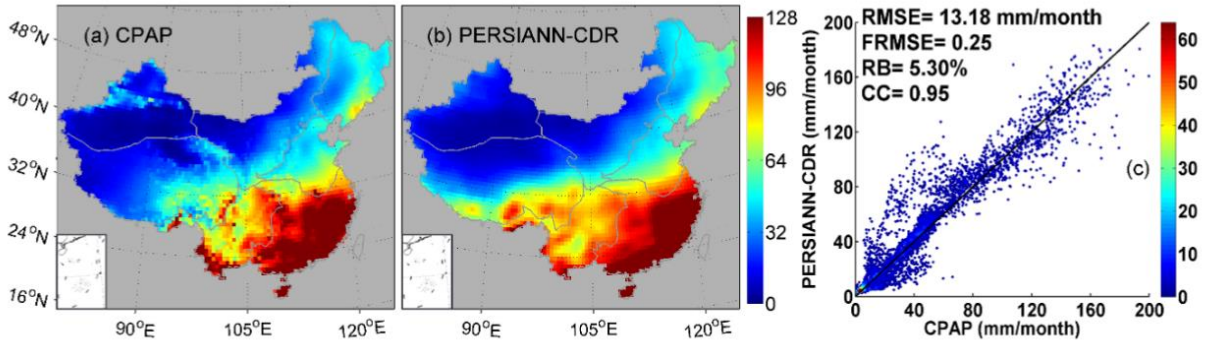

Figure 6. Spatial distribution of 32 years monthly mean precipitation data between the CPAP and PERSIANN-CDR along with a statistical analysis line of regression between the same datasets (Guo et al., 2016).

Various methods for analyzing the precipitation from models and using Cumulative Distribution Functions (CDFs) are standard (Wang et al., 2013). CDFs have long been used to bias-correct raw modeled precipitation (Wang et al., 2013). A CDF is the probability that a precipitation event is less than or equal to one (Wang et al., 2013). Much work has been done in investigating various post-processing techniques from simple additive and scaling corrections with CDFs (Wang et al., 2013). QM is an

effective way to adjust CDFs to agree with current observations in a given reference period (Wang et al., 2013).

Kadioglu et al. (2000) analyzed regional precipitation trends in Turkey, primarily using correlation. The downside of using correlation is the sensitivity to the non-normality of data, adjusted with the correlation matrix to assume a normal distribution (Kadioglu et al., 2000). Each observation station was analyzed using quantile to quantile plots, useful for visualizing transforms (Kadioglu et al., 2000). The most considerable drawback noted in this research was the strong seasonal signals of precipitation, which drive the values upwards or downwards, which affect the accuracy of the data (Kadioglu et al., 2000).

Quantile mapping (QM) is a commonly used technique to correct distributional biases in precipitation outputs from climate models compared to observations (Cannon et al., 2015). A drawback with quantile mapping is that corrupt model-projected trends are possible with precipitation, which requires a bias correction algorithm (Cannon et al., 2015). QM outperformed several empirical statistical downscaling (ESD) techniques (Rajczak et al., 2016). This technique well represented the tails of precipitation probability distribution functions in extreme rain events, which reduced biases in regional climate models (Trinh-Tuan et al., 2018).

The statistical analysis of extreme precipitation events is essential, and parametric probability distributions fit these climatological records' values in Wilks' et al. (1993). These fitted distributions smoothed the data, and extrapolation was obtained beyond the identified graphical region (Wilks et al., 1993). The depiction of the difference between partial duration and annual extreme precipitation events frequency is important to

differentiate (Wilks et al., 1993). The partial duration series only allowed events that exceeded a specific predetermined value beneficial for forecasting flash floods (Figure 7).

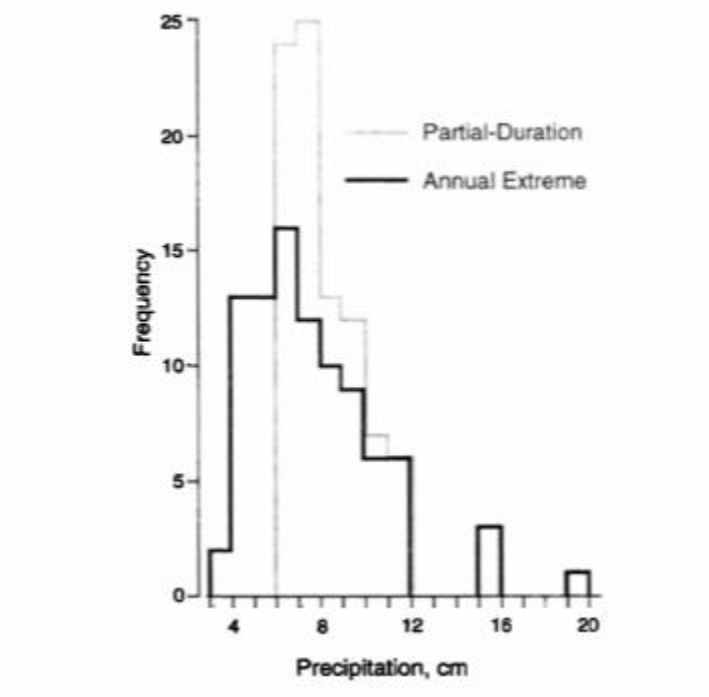

Figure 7. Histograms of 1-day annual extremes for Baltimore, MD (Wilks et al., 1993).

III. Methodology

Chapter Overview

This chapter describes the data in this research and the methods of analysis.

Summary

The data used in this research includes the CFSv2 and the PERSIANN-CDR. The CFSv2 and PERSIANN-CDR data are manipulated to fit equal resolution grids of 10-degree boxes such that direct comparisons could be made. The atmospheric model of the CFSv2 has a spectral triangular truncation of 126 waves in the horizontal and uses finite differencing vertically with 64 sigma pressure levels (Saha et al., 2014). The CFSv2 is

0.5-degree resolution and the PERSIANN-CDR is 0.25-degree resolution so coarser resolutions are required for this analysis.

The CFSv2 forecast data is broken up into forecast time scales of 0-to-10, 15-to-25, 55-to-65, and 80-to-90 days. General comparisons will be drawn from both datasets comparing the means, standard deviations, maximum value, and coverage percentages globally (Figure 8). Datasets are created with CFSv2 from 2016 through 2019 and the PERSIANN-CDR from 1983 through 2019. The PERSIANN-CDR required more data compared to the CFSv2 because the corrections applied were more reliable for climatological results.

Acronym	Description	Unit
Mean	Mean daily precipitation	mm
SD	Standard deviation of daily precipitation	mm
CS	Coefficient of skewness of daily precipitation	-
PRCPTOT	Annual total precipitation in wet days (daily precipitation ≥ 1 mm)	mm
SDII	Annual precipitation divided by the number of wet days	mm/day
Rx1day	Annual maximum 1-day precipitation	mm
Rx5day	Annual maximum 5-day precipitation	mm
R95pTOT	Annual total rainfall when daily precipitation > 95 percentile	mm
R99pTOT	Annual total rainfall when daily precipitation > 99 percentile	mm

Figure 8. Various statistical indices to represent characteristics of precipitation (Heo et al., 2019).

Time series will then be compared for PERSIANN-CDR at two different points globally, whereas CFSv2 forecasts for raw signal and raw mean are found using a series of filters, which include the Savitzky-Golay (Savgol) to smooth the data while the integrity remained. The Savgol filter will be used because data peaks are not flattened compared to other smoothing functions which was important for the seasonal variational analysis of precipitation (Schafer et al., 2011). Also, the Savgol filter applies the smoothing to a polynomial function which is effective in continuous sets of data such as

yearly precipitation amounts (Schafer et al., 2011). The points are in the Tropical Pacific near the international dateline and the equator and the U.S. near Colorado.

A technique used for verification is a Pearson correlation function from the numerical python (numpy) library. Pearson correlation is the covariance of two variables divided by the product of their standard deviations (Benesty et al., 2008). Correlation plots with the Savgol filter are shown to explain the accuracy of the CFSv2 dataset forecasts. The Savgol filter uses two different window lengths and polynomial order numbers for the correlation plots. The correlation technique applied will evaluate the CFSv2's linear relationships with the observations for the different forecast periods with mean and coverage statistics.

Verifications from observations of the CFSv2 ensure that accurate representation is attained with the values. Observations are applied to the CFSv2 forecasts to ensure the accuracy of the data. Contingency tables created determine the false alarm, hits, and misses for the different forecast periods to determine the highest skill of the CFSv2 compared to CFSv2 observations as quantitative percentages.

Geometrically spaced bins create plots that are used for histogram and CDF analyses for precipitation events from 0 to 30. A histogram method then places the precipitation values into individual bins. Seasons are broken up each year to gather the smallest temporal scale for this research. The seasons are defined as 91 Julian day periods with winter 1 to 92, spring was 93 to 184, summer was 184 to 275, and fall was 276 to 365. Then CDFs are plotted to identify an accurate representation of significant rainfall events compared to the percentage of occurrences annually. Precipitation time series are then created using the QM correction method. The mean, standard deviation, maximum,

and coverage of the PERSIANN-CDR and CFSv2 0-to-10 day forecast are applied to the CFSv2 forecast for 15-to-25 days to show improved precipitation forecast ability for the QM. This comparison allows for the raw precipitation, and quantile mapped to be observed side by side.

QM provids corrections to the CFSv2 15-to-25 day forecasts using both the PERSIANN-CDR and the CFSv2 0-to-10 day forecasts statistics. This QM method uses a test dataset, a year, and a training dataset that used three years of data from the CFSv2. The data is placed into the quantile correction function. The quantile correction function equates CDFs of the observed, modeled, and bias correction data within a specific time period (Cannon et. al., 2015). The training dataset is used as the modeled from the CFSv2 data. The QM data is manipulated based on the training set which used 2016, 2018, and 2019 from the CFSv2.

The test data is from the beginning and end of 2017, which is a full continuous year worth of CFSv2 data used as the bias corrected for the QM. That is the most effective way to break up the CFSv2 data to apply the QM function. The importance of using QM for this research is to apply any corrections to the CFSv2 biases in sub-seasonal precipitation trends. The general comparison provides insight for the accuracy of the high-resolution data with the QM implementing correction methods for training data from the observations for 2017.

IV. Analysis and Results

Chapter Overview

The analysis of the PERSIANN-CDR and the CFSv2 are conducted and described for specific findings in this section.

Summary

The 37-year period for all days with the leap days removed because of added complexity looking at the mean, standard deviation, maximum values, and total coverage for the PERSIANN-CDR. The mean aids in determining where the highest average precipitation values fall globally, while the standard deviation identifies regions of variability in the 10 by 10-degree boxes. The PERSIANN-CDR maximum values depict the heaviest rain events, especially in the tropical regions. The coverage demonstrates how much in each grid box precipitation has fallen globally, which is measured in percentage compared to millimeters for the other three statistical variables.

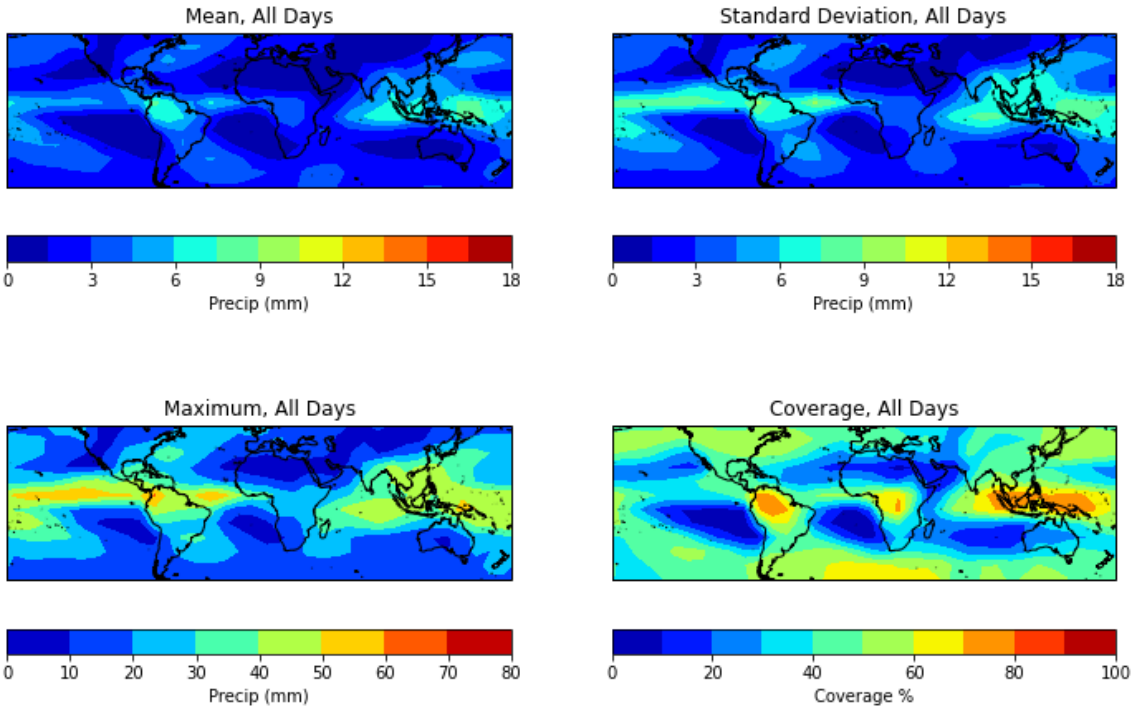

Figure 9. PERSIANN-CDR 37 year mean, standard deviation, maximum, and coverage for all days from 1983 to 2019.

The standard deviation and maximum values of precipitation from the PERSIANN-CDR annual analysis depict the Inter Tropical Convergence Zone (ITCZ)

region (Figure 9). The maximum precipitation fallen depicts the ITCZ in a broader scale compared to the mean and standard deviation. The coverage for the PERSIANN-CDR is most useful in identifying the desert regions of the globe. The coverage also highlights the tropical regions over the land along with the Indo-Pacific for all days.

The seasonal analysis is then examined for the PERSIANN-CDR. The importance of the data being broken down into the seasons is to identify signals, primarily large-scale rain events known to impact regions such as the Indian Monsoon. The identification of the signals in the satellite dataset can aid in the corrections used for the QM. The climatology of the PERSIANN-CDR seasonal signals will improve the CFSv2 forecasts.

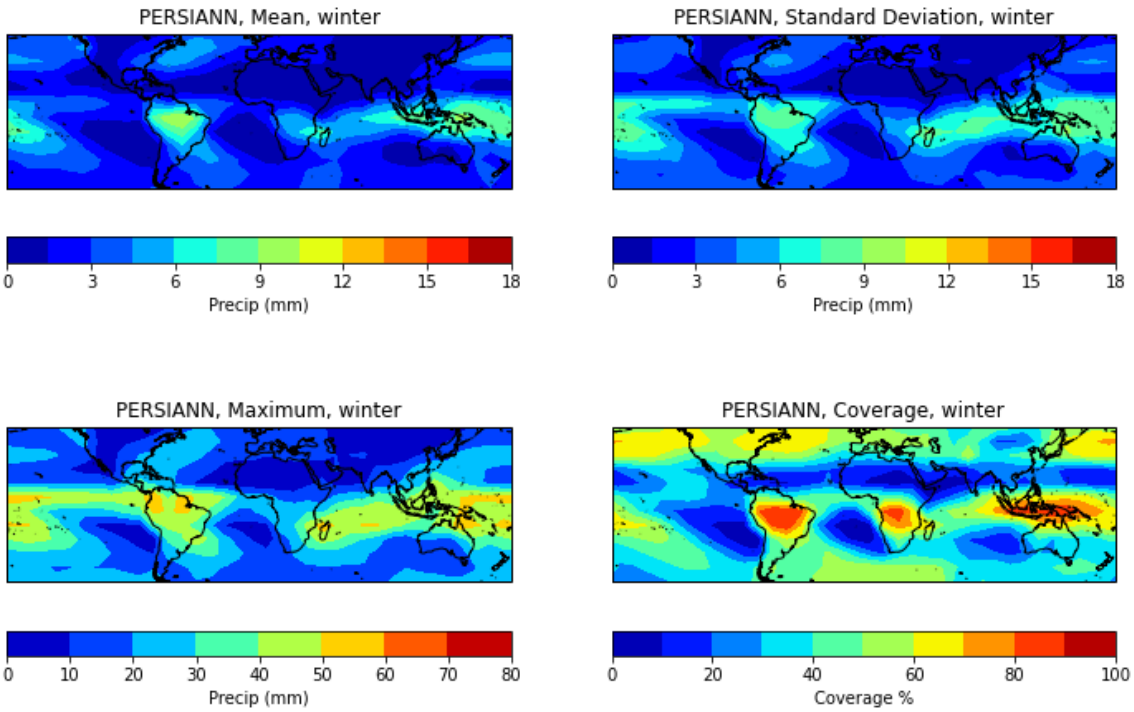

Figure 10. PERSIANN-CDR 37 year mean, standard deviation, maximum, and coverage in winter from 1983 to 2019.

The winter season for the PERSIANN-CDR depicts a larger area of mean precipitation over the northern part of South America along with the central Pacific compared to the mean (Figure 10). The standard deviation of winter does not show a continuous flow of higher precipitation variability in the Tropics, connected across the entire eastern Pacific compared to the all days analysis. The maximum values of rainfall have shifted further west in the Pacific than all days, with two peak regions located over northern South America. The coverage values are significantly higher in the annual statistics of the PERSIANN-CDR over South America, Africa, and the Indonesia region of the Pacific, which is associated with the ITCZ, which shifts south during the winter season.

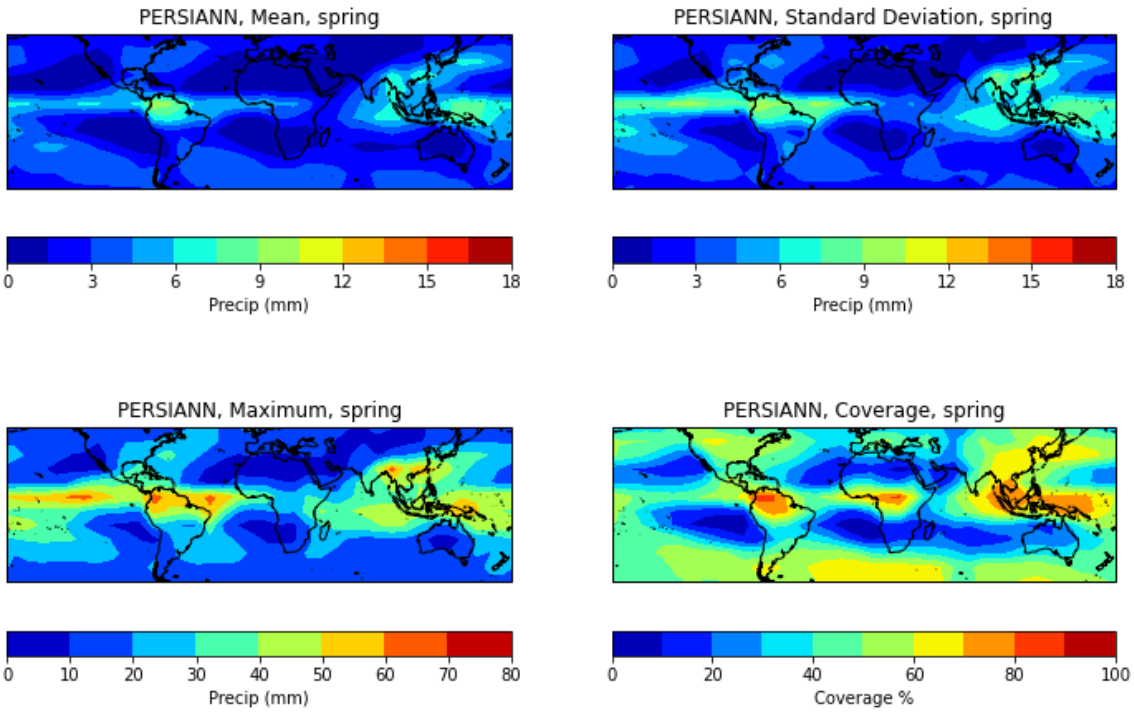

Figure 11. PERSIANN-CDR 37 year mean, standard deviation, maximum, and coverage in spring from 1983 to 2019.

The spring mean values of PERSIANN-CDR have highlighted the ITCZ more distinctly than winter, which had washed-out appearances from other seasonal signal impacts (Figure 11). The standard deviation remains about the same, but there are higher values of the maximum over the globe's eastern Pacific and Atlantic regions. The southeast Asia region has higher values over the land and just to the north of Papua New Guinea associated with the South Pacific Convergence Zone (SPCZ). The coverage values for spring precipitation are higher over northern South America and central Africa.

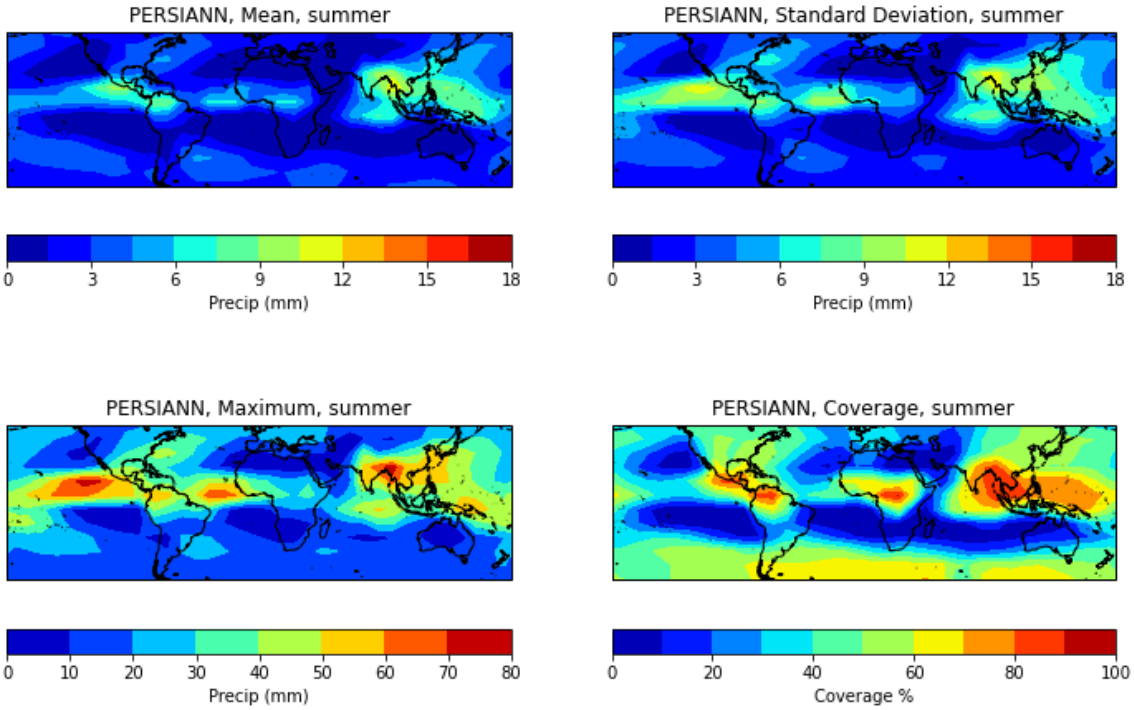

Figure 12. PERSIANN-CDR 37 year mean, standard deviation, maximum, and coverage in summer from 1983 to 2019.

The PERSIANN-CDR summer has a much different pattern from the previous seasons, especially the coverage and the maximum, which yields much higher values than the winter and spring (Figure 12). A prominent peak over southeast Asia is associated with the Indian Summer Monsoon, which typically lasts from June to September. In

addition, Central and South America have high coverage percentages related to increased ITCZ activity and the hurricane season in the eastern Pacific. The maximum summer values over the east Pacific and southeast Asia depict significant rainfall amounts, which line up with the synoptic meteorological activity.

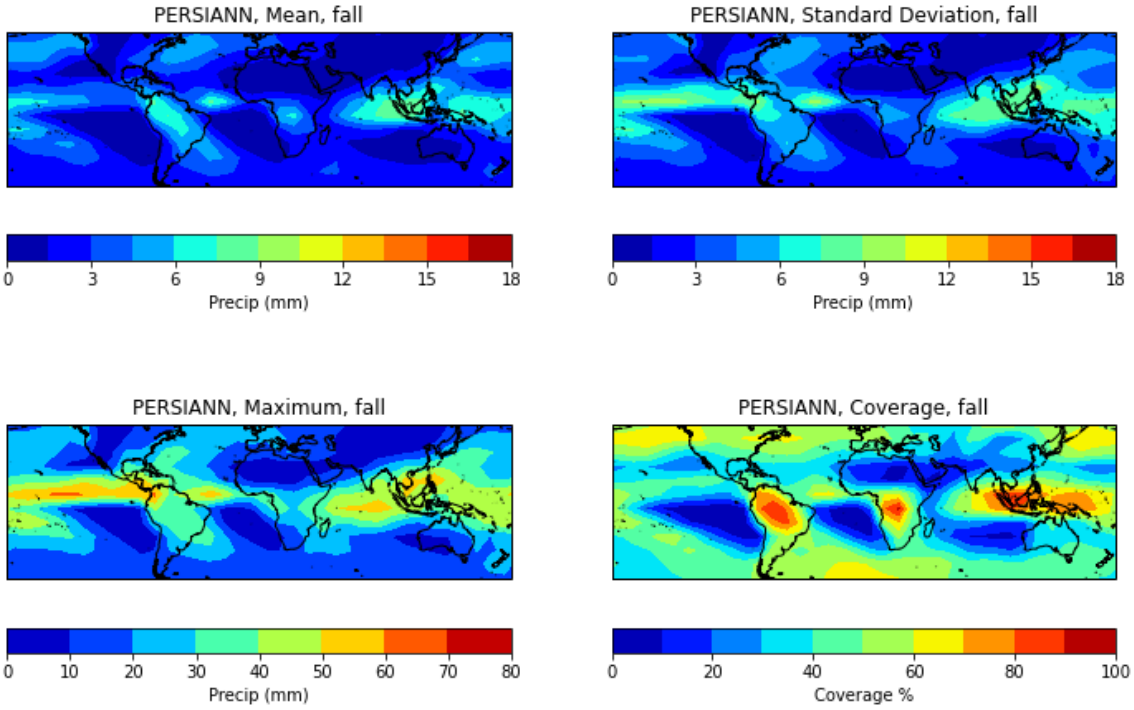

Figure 13. PERSIANN-CDR 37 year mean, standard deviation, maximum, and coverage in fall from 1983 to 2019.

The fall season of the PERSIANN-CDR depict higher mean values over the central and western Pacific compared to the summer (Figure 13). The standard deviation values are shifted further west, similar to the mean. It was also noted that the higher values of precipitation in the northern hemisphere are further south compared to the summer, which is plausible with fall being a transitional season. The maximum values in fall were much lower compared to summer which show the Indian Monsoon declining along with the hurricane season in the eastern Pacific. Finally, the coverage values are

shifted further south, similar to the standard deviation, which supports the large-scale meteorological events for this season.

The PERSIANN-CDR data differs from the CFSv2, which has plots at different forecast time scales. The same basic statistics were compiled on the CFSv2 to directly compare with the exact resolution and latitude values since PERSIANN-CDR only covers up to 60 north/south. The CFSv2 (00-24 hr) were created for an annual analysis to compare PERSIANN-CDR. The 00-24hr forecast did not change throughout the different forecast periods, which was expected with almost identical values. The 00-24hr forecast depicted the verification values for the CFSv2 which was essential to compare directly with the PERSIANN-CDR statistical parameters (Figure 14).

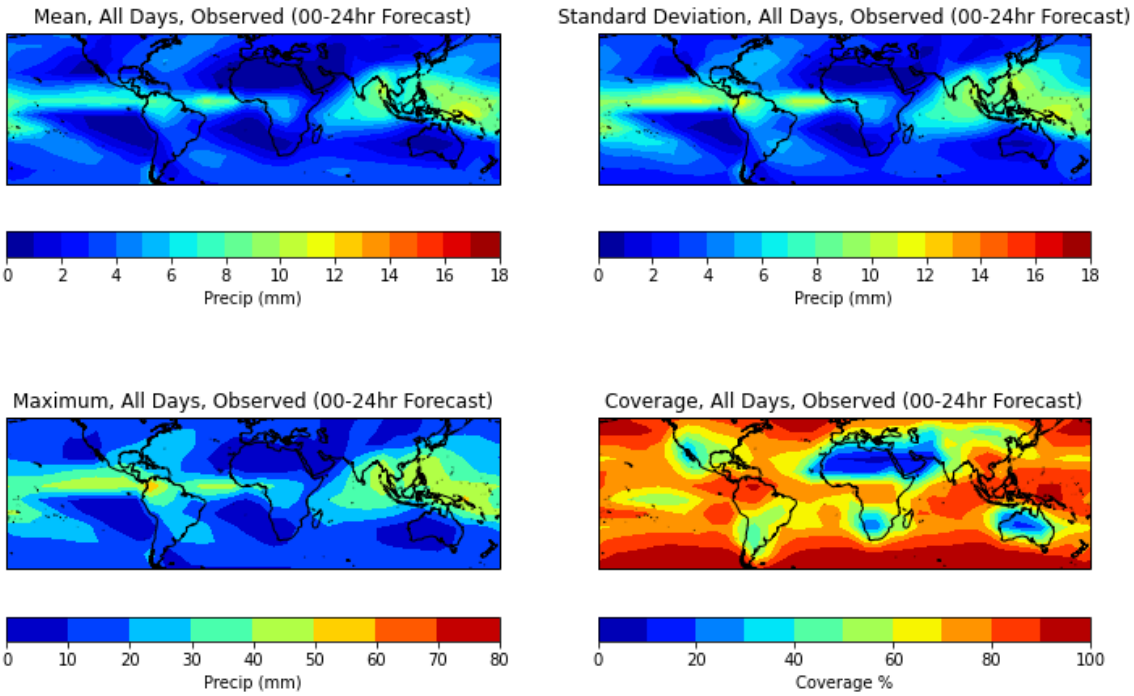

Figure 14. CFSv2 all days observations mean, standard deviation, maximum, and coverage for all days from 2016 to 2019.

The CFSv2 plots compared to the PERSIANN-CDR all days show a significant difference in the mean and the coverage values. The standard deviation and the

maximum of CFSv2 are similar to PERSIANN-CDR, with the main difference over central Africa and the central Pacific, the ITCZ is more intense across a broader area of the Tropics. The coverage percentage show the absence of precipitation over the globe's desert regions, including the southwestern United States, the Sahara, and the Outback in Australia. The CFSv2 overestimates the amount of precipitation coverage globally compared to the PERSIANN-CDR.

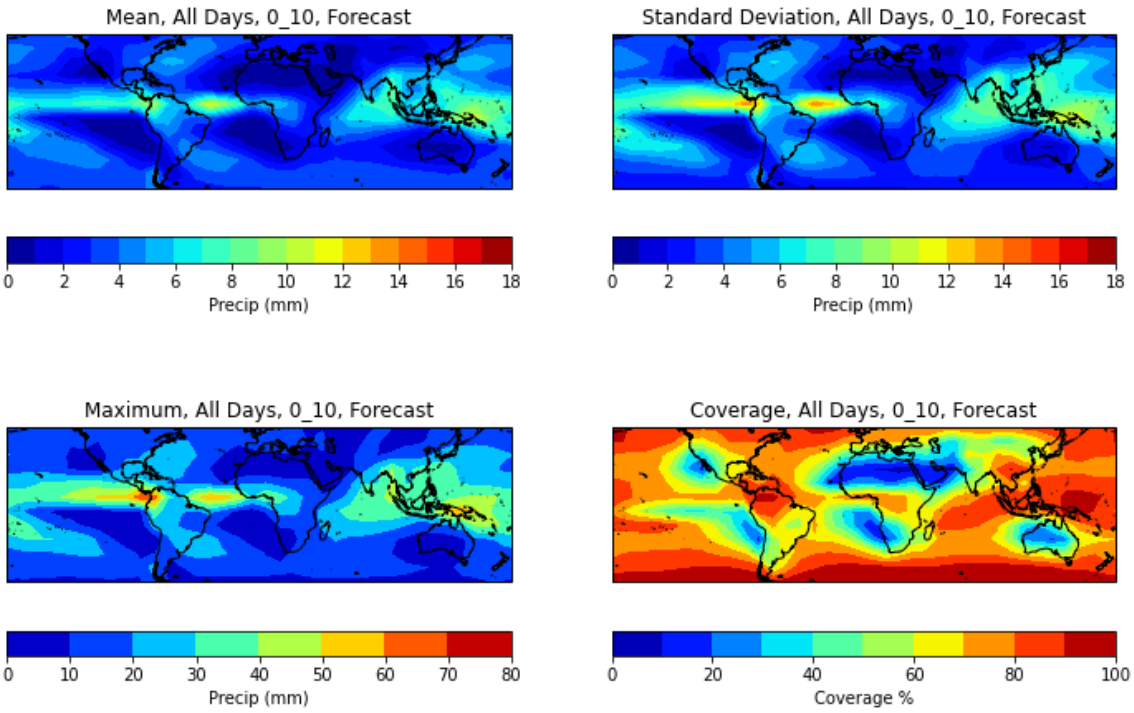

Figure 15. CFSv2 0-to-10 day forecasts mean, standard deviation, maximum, and coverage for all days from 2016 to 2019.

The most exciting features in the CFSv2 0-to-10 day forecasts are the higher values of the standard deviation, which show the model's variability with precipitation forecasts near the tropical regions (Figure 15). The Tropics tend to provide most of the global higher-end precipitation events. The coverage values line up closely with the observed values but over forecasted the percentage globally. This may have been caused

by insufficient raw data from the set and limitations compared to the PERSIANN-CDR. The mean showed a more distinctive ITCZ feature than the PERSIANN-CDR but lined up closely over the Indonesia and southeast Asia regions. As the forecasts move further out, a more distinctive analysis occurs when the CFSv2 has less reliable results.

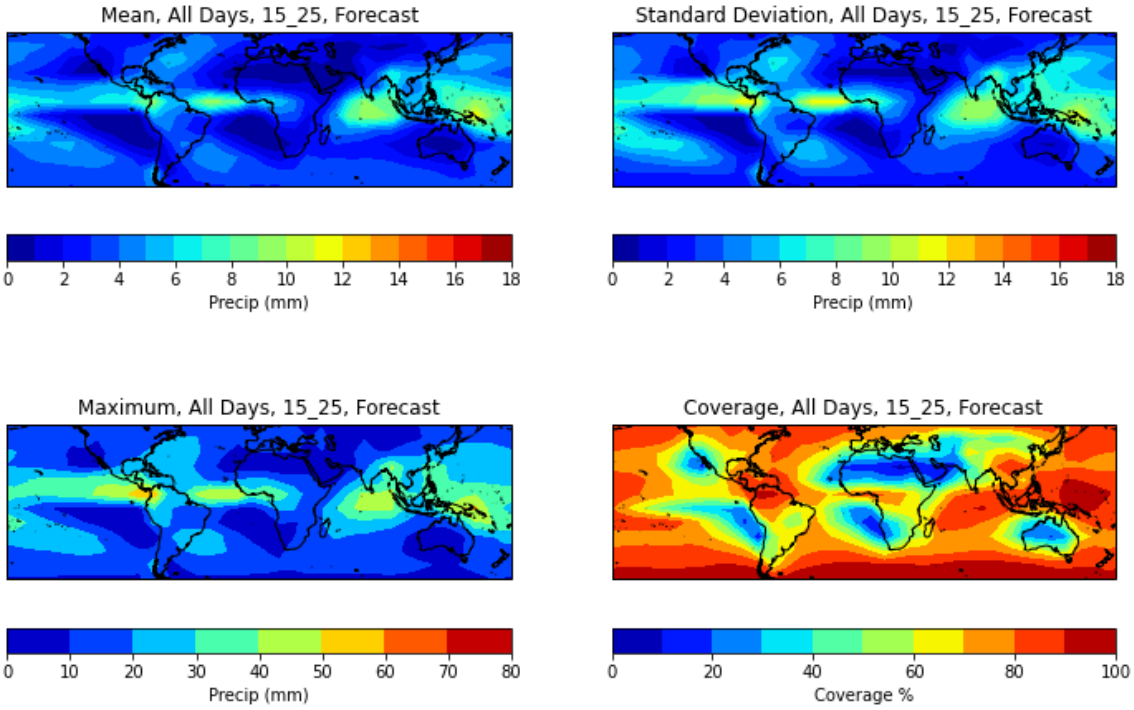

Figure 16. CFSv2 15-to-25 day forecasts mean, standard deviation, maximum, and coverage for all days from 2016 to 2019.

The statistics for CFSv2 for 15-to-25 day forecasts from 2016 to 2019 appears different than the 0-to-10 day (Figure 16). The statistical variables note the standard deviation, which shows slightly less variability near the tropical regions than the 0-to-10 day forecasts. The mean values are almost identical compared to the maximum and coverage values, with little to no changes between 0-to-10 and 15-to-25 day forecasts for all days. The ITCZ region is not as clearly depicted across the Tropical Pacific in the 15-to-25 day forecasts compared to the 0-to-10 day forecasts. The maximum values are

lower in the 15-to-25 day compared to the 0-to-10 day forecasts, which is associated with lower accuracy of hurricane season in the Atlantic and eastern Pacific regions.

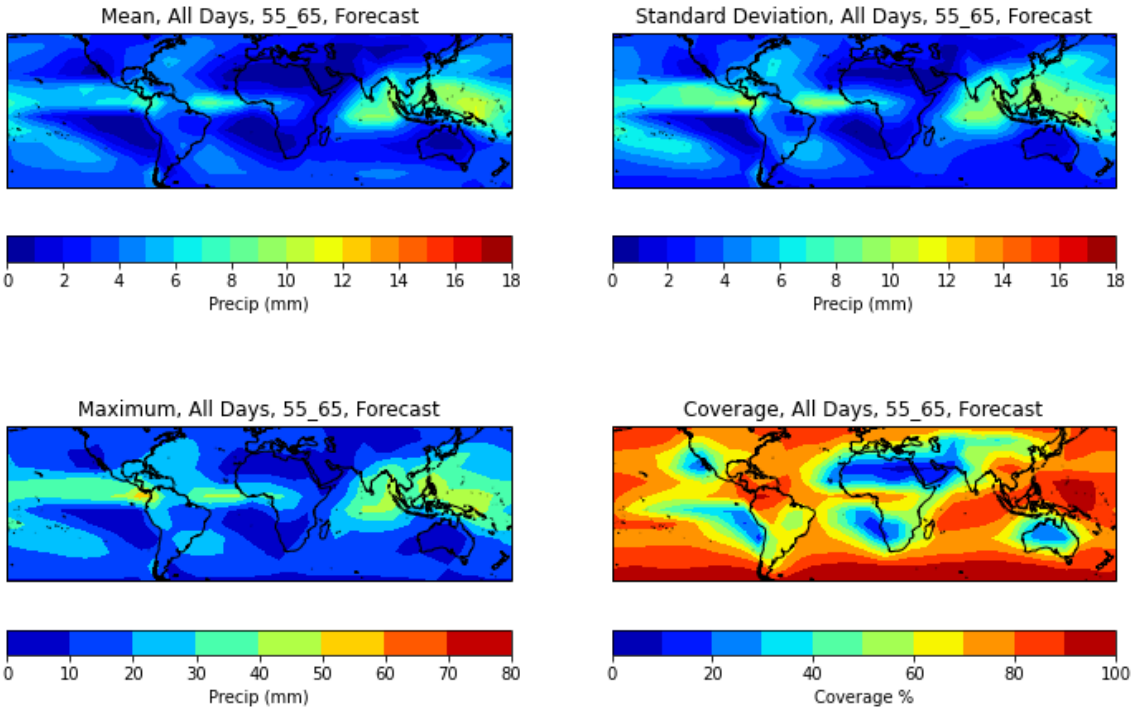

Figure 17. CFSv2 55-to-65 day forecasts mean, standard deviation, maximum, and coverage for all days from 2016 to 2019.

The 55-to-65 day forecasts of the CFSv2 depict consistency in precipitation amounts compared to the shorter-range forecasts (Figure 17). There is little change on these statistical parameters as the forecasts extend beyond 15-to-25 days. The main difference is with the standard deviation, which has lower variability compared to the 0-to-10 and 15-to-25 day forecasts. The mean in the 55-to-65 day forecasts has a weaker ITCZ signal compared to the 15-to-25 and the 0-to-10 day forecasts. The maximum amounts of precipitation are similar to the 15-to-25 day forecasts. Maximum values are more accurately represented the further out in the forecasts for the CFSv2 compared to the 0-to-10 day because those events are over predicted in the short-range. The coverage

remains consistent through all the different forecast time frames, suggesting that the model handles this variable the best.

The 80-to-90 day CFSv2 forecasts is the furthest out time period analyzed for all days with minimal differences from the shorter range forecasts mostly associated with the mean values over the eastern Pacific (Figure 18).

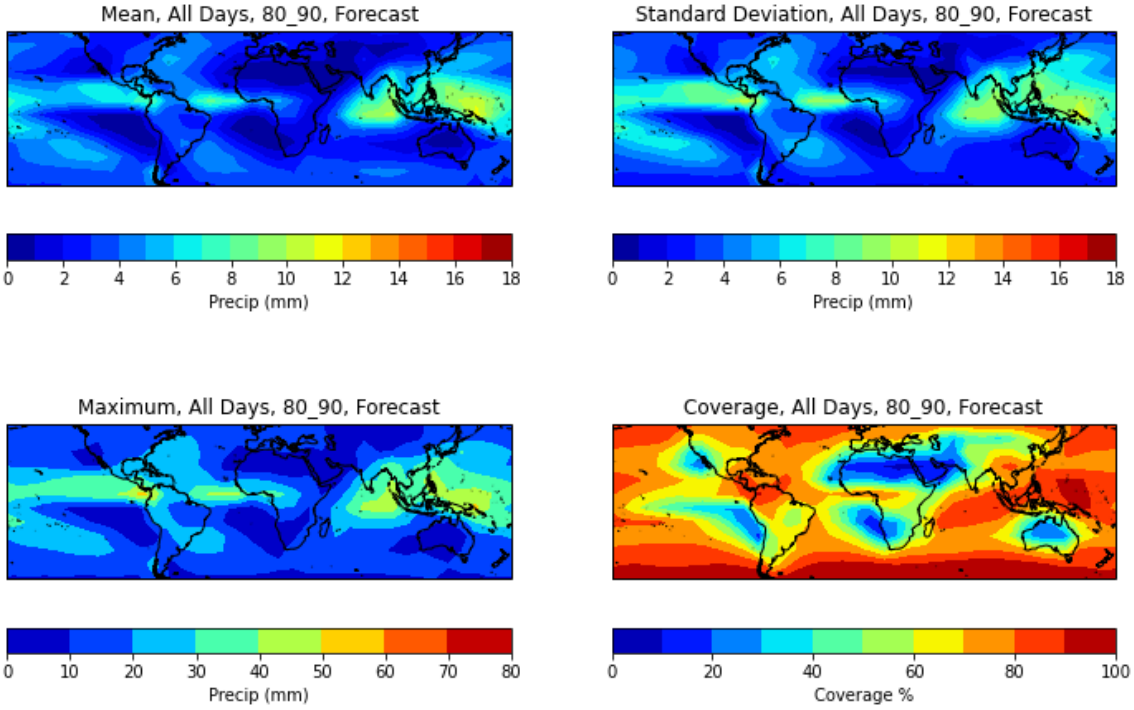

Figure 18. CFSv2 80-to-90 day forecasts mean, standard deviation, maximum, and coverage for all days from 2016 to 2019.

Many of the features are different, especially near the tropical regions, as the forecasts go out in time compared to the 0-to-10 day forecasts, which supports the idea that models are not reliable beyond 30 days. The mean for the 80-to-90 day forecasts is similar to 55-to-65 and 15-to-25 day forecasts, which attempt to represent the ITCZ, which is visible in the 0-to-10 day forecasts. The standard deviation in the 80-to-90 day forecasts are equivalent to the 15-to-25 and 55-to-65 day forecasts, which do not depict the high

variability of tropical precipitation in the globe's equatorial regions. The maximum is forecasted the best at the more extended range forecasts than the 0-to-10 days.

The coverage is the only statistical parameter that has remained consistent during all forecast time frames with an over-representation of global precipitation in the CFSv2 compared to the PERSIANN-CDR. It is essential to break all days into seasons, which was also done with the PERSIANN-CDR for direct comparisons. The seasonal signals within the CFSv2 for spring and summer are directly compared to the PERSIANN-CDR. The determination of seasonal biases forecasted in the spring and summer are identified. The lack of CFSv2 data do not allow for accurate analyses to be conducted with the winter and fall seasons.

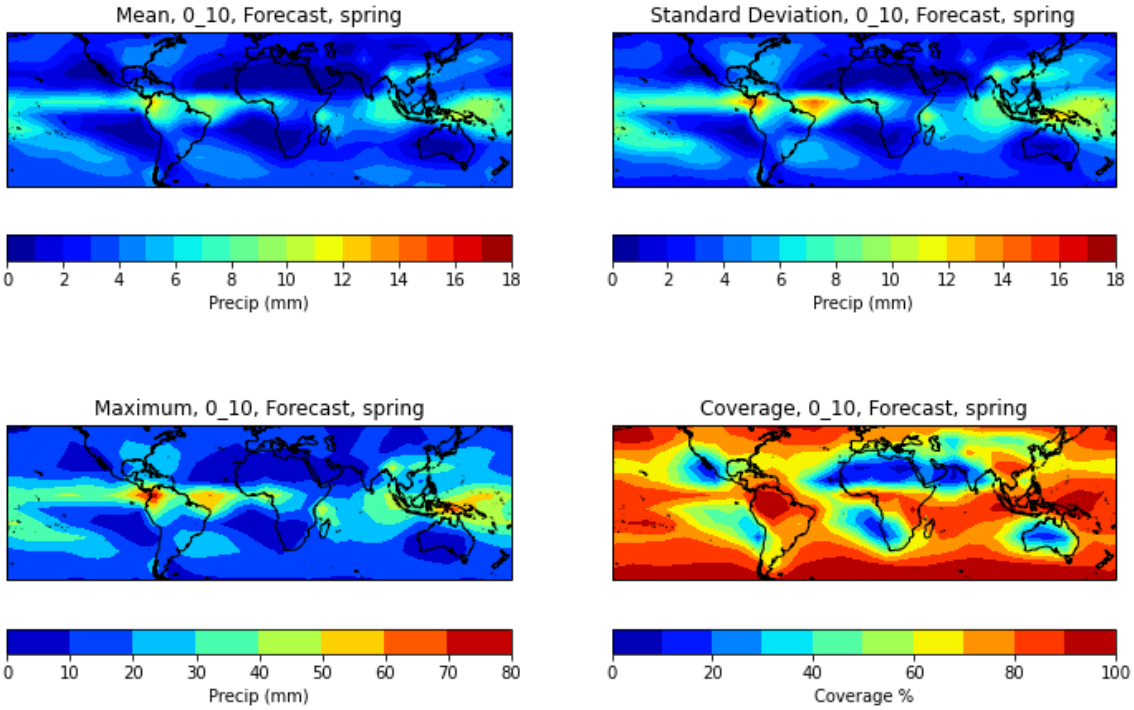

Figure 19. CFSv2 0-to-10 day forecasts mean, standard deviation, maximum, and coverage for spring from 2016 to 2019.

The main difference noted in the spring CFSv2 0-to-10 day forecasts compared to PERSIANN-CDR is the higher variability in precipitation over the tropical region (Figure 19). This higher variability is associated with the models' failure to accurately depict precipitation in this region of the globe compared to the PERSIANN-CDR. The ITCZ is not well depicted in the spring for the PERSIANN-CDR, which supports the transitional season precipitation complexity, while the CFSv2 for 0-to-10 day forecasts has a defined tropical precipitation band. The maximum values of the spring line up with all days for the CFSv2, which differ from the PERSIANN-CDR, showing peaks mostly over the Tropical Pacific. The coverage in the CFSv2 showed an over prediction of precipitation across the globe compared to the PERSIANN-CDR, which has peaks only over land in the Tropics.

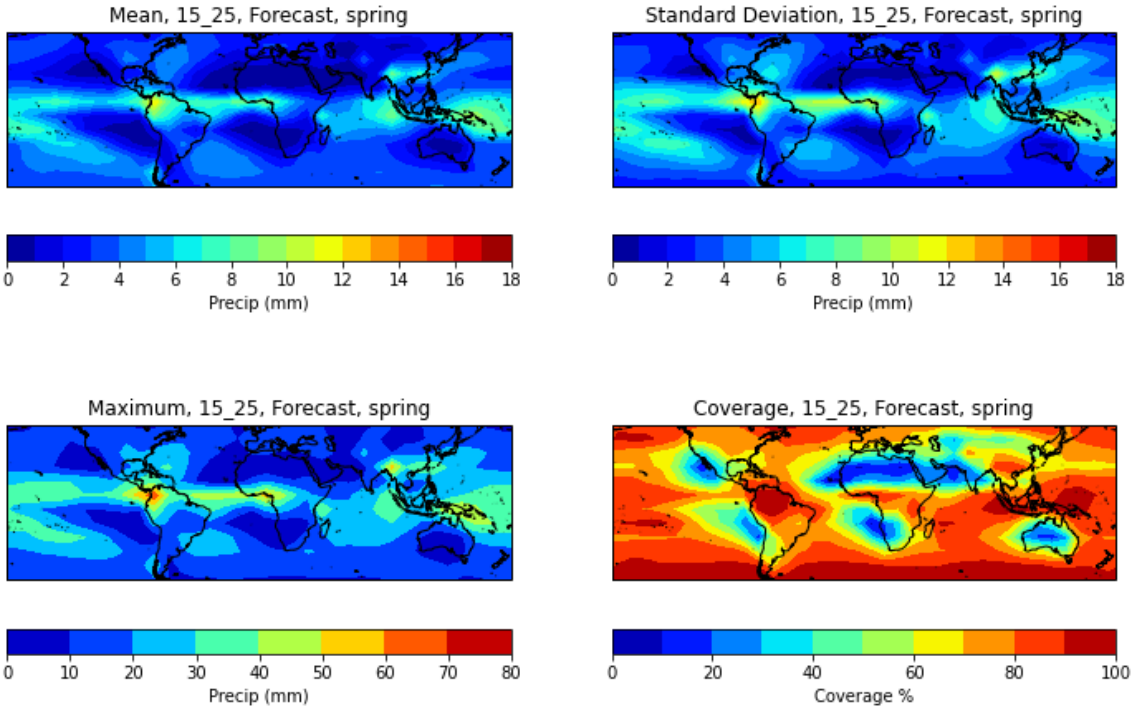

Figure 20. CFSv2 15-to-25 day forecasts mean, standard deviation, maximum, and coverage for spring from 2016 to 2019.

The 15-to-25 day spring mean values are weaker over the Tropical Pacific compared to the 0-to-10 day with a closer resemblance to the PERSIANN-CDR (Figure 20). The standard deviation of the 15-to-25 day forecasts has lower values across the ITCZ region than the 0-to-10 day forecasts for spring, reflecting the PERSIANN-CDR more closely. The maximum values peak over South America, which is the same location from the 0-to-10 day forecasts but significantly different from the PERSIANN-CDR. The coverage percentage for the 15-to-25 day forecasts are consistent with the 0-to-10 day CFSv2 forecasts but are much different from the PERSIANN-CDR.

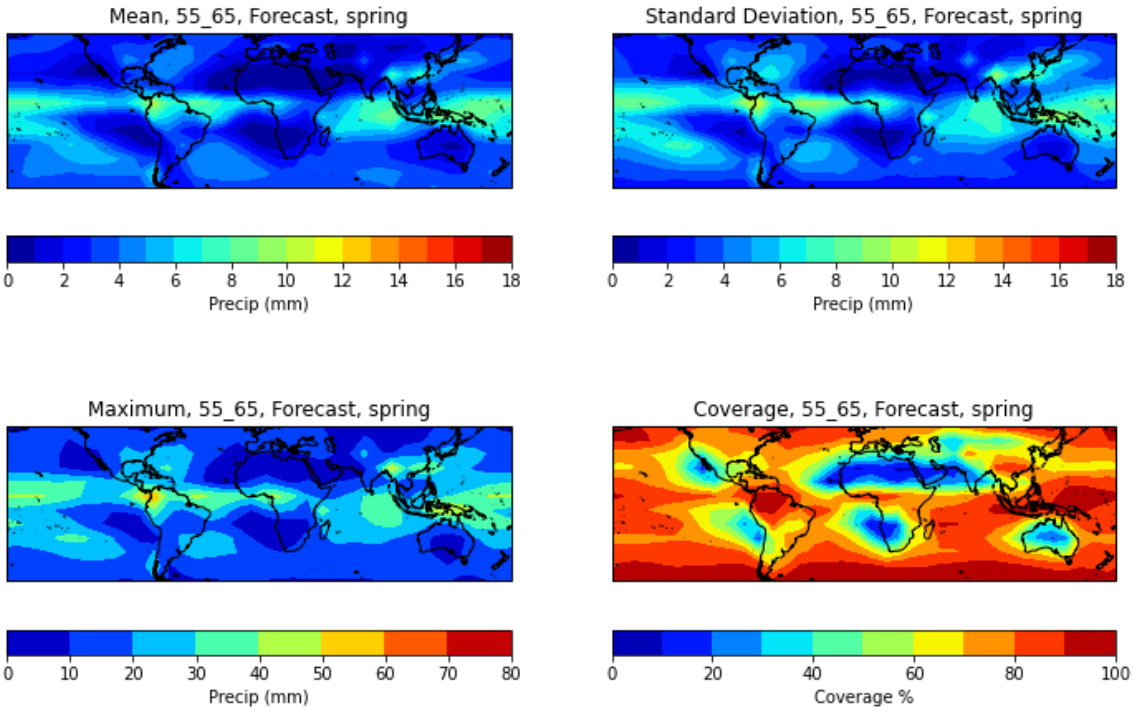

Figure 21. CFSv2 55-to-65 day forecasts mean, standard deviation, maximum, and coverage for spring from 2016 to 2019.

The 55-to-65 day forecasts for spring are almost equivalent in variability with the 15-to-25 day but lower than the 0-to-10 day forecasts (Figure 21). The mean shows an even closer representation compared to the PERSIANN-CDR with a weaker ITCZ signal.

The maximum values are lower than the previous CFSv2 forecast time frames for spring, widely different from the PERSIANN-CDR. The coverage percentage for the 55-to-65 day forecasts remain consistent with the previous CVSv2 spring forecast time frames, which do not line up closely with the PERSIANN-CDR.

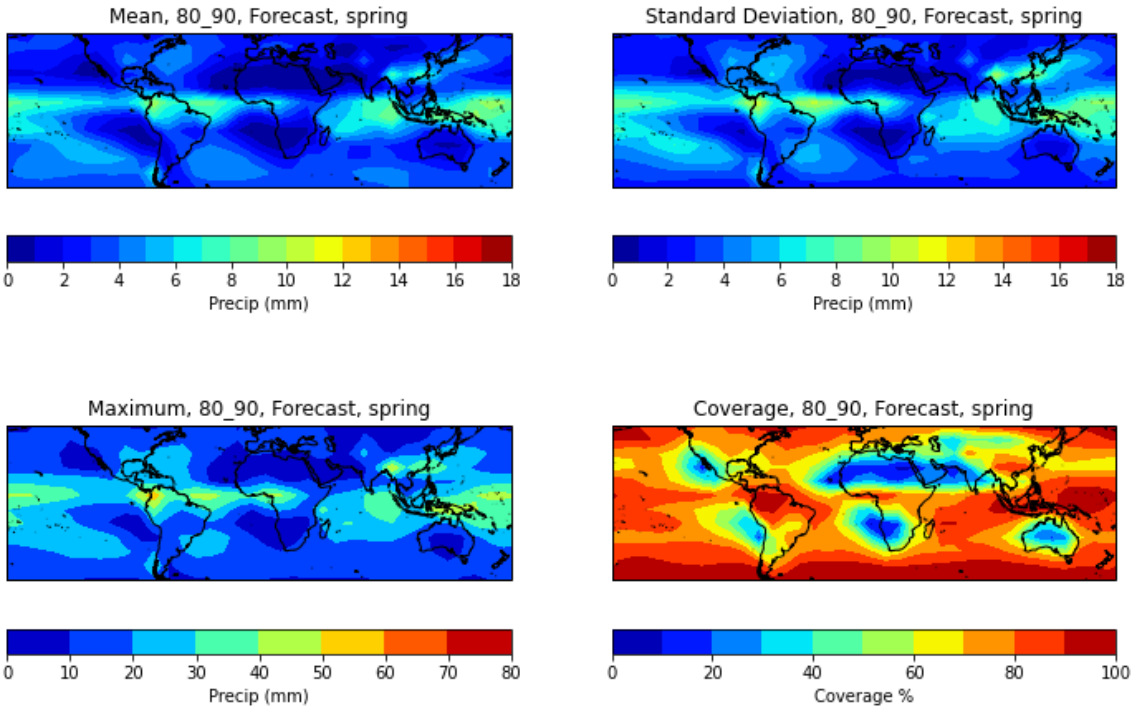

Figure 22. CFSv2 80-to-90 day forecasts mean, standard deviation, maximum, and coverage for spring from 2016 to 2019.

All of the 80-to-90 day forecasts statistical parameters are almost equivalent to the 55-to-65 day forecasts, further supporting the lack of variability in data beyond 15-to-25 day forecasts (Figure 22). The mean and standard deviation of the 80-to-90 day forecasts represent the PERSIANN-CDR more accurately. Forecasting heavier precipitation is difficult the further out from the current observational time, even more so in the Tropics. The spring season forecasts determine the difficulty that the CFSv2 has with the transitional precipitation features specifically in the Tropical Pacific.

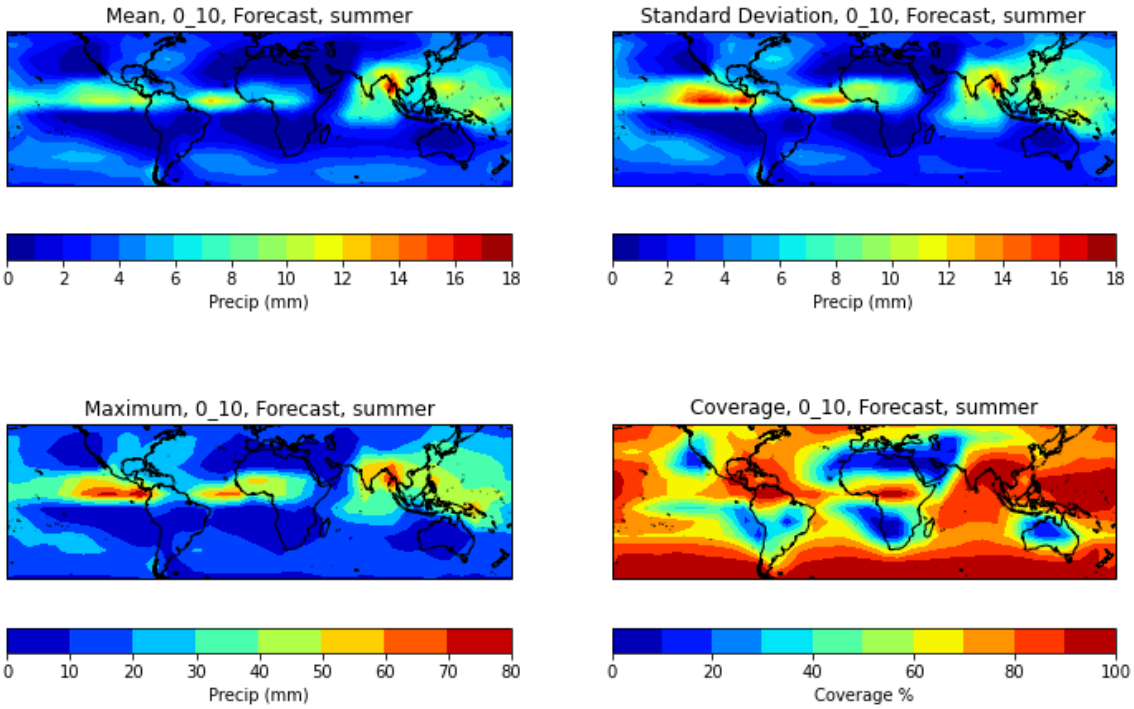

Figure 23. CFSv2 0-to-10 day forecasts mean, standard deviation, maximum, and coverage for summer from 2016 to 2019.

The summer CFSv2 forecasts has a stronger grasp on the ITCZ compared to spring (Figure 23). The 0-to-10 day forecasts of summer overpredict the amount of precipitation over the ITCZ region similar to the spring for the CFSv2 compared to the PERSIANN-CDR mean values. The standard deviation of the CFSv2 show very high variability compared to the PERSIANN-CDR over the eastern Pacific and Atlantic, which line up with hurricane season along with the Indian Monsoon. The maximum values are associated with the hurricane season and the monsoon near the Indian ocean, significantly larger than the PERSIANN-CDR. The coverage percentage remain high across the entire globe besides desert regions with the CFSv2 and a notable increase over the southeastern U.S. compared to the PERSIANN-CDR for the summer.

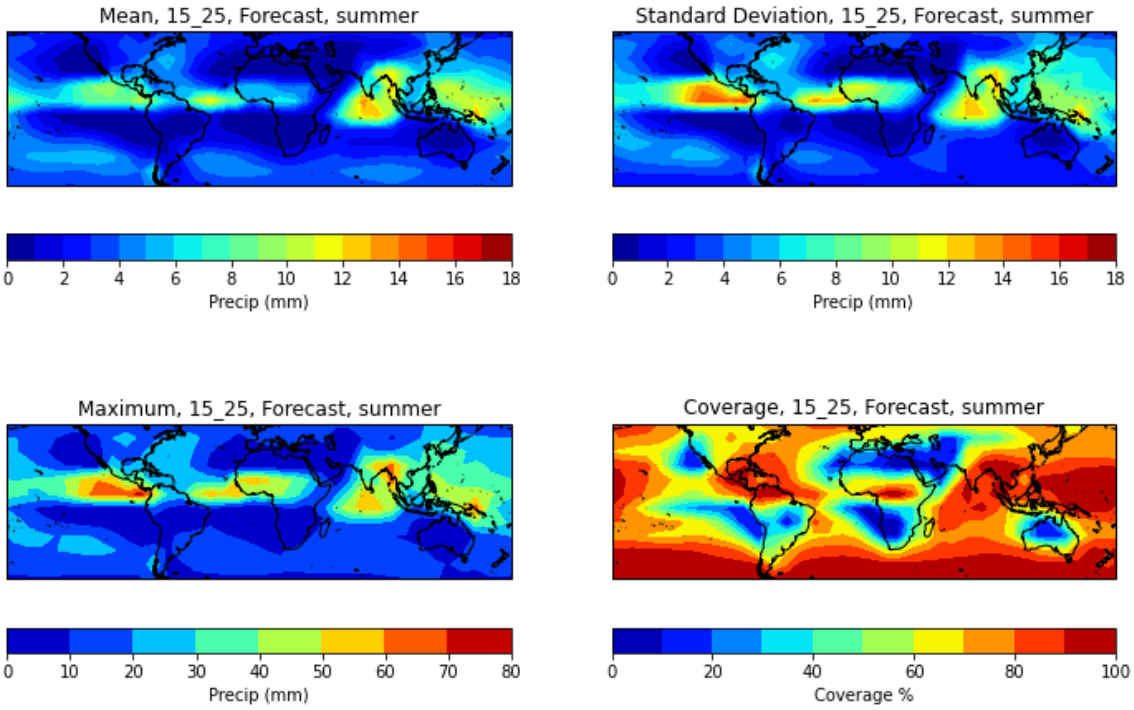

Figure 24. CFSv2 15-to-25 day forecasts mean, standard deviation, maximum, and coverage for summer from 2016 to 2019.

The 15-to-25 day summer forecasts of precipitation mean line up with the 0-to-10 day forecasts but much different compared to the PERSIANN-CDR (Figure 24). The standard deviation values are smaller in the 15-to-25 day forecasts compared to 0-to-10 day but remain higher than PERSIANN-CDR summer variability in precipitation. The maximum values are slightly lower in the 15-to-25 day compared to the 0-to-10 day but still more extensive than the PERSIANN-CDR, all associated with the mesoscale and synoptic meteorological patterns. The coverage percentage is about the same in the 15-to-25 day forecasts, with 0-to-10 day keeping the precipitation higher in the southeastern U.S. than the PERSIANN-CDR.

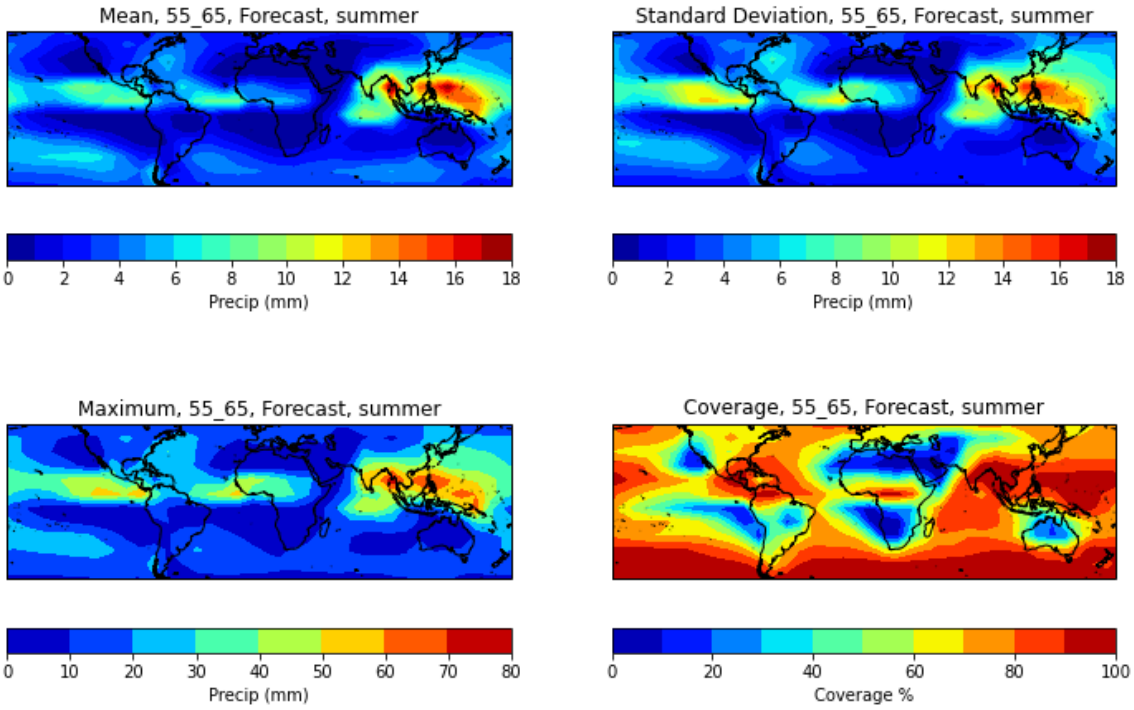

Figure 25. CFSv2 55-to-65 day forecasts mean, standard deviation, maximum, and coverage for summer from 2016 to 2019.

The mean in the 55-to-65 day forecasts is more closely represented with the PERSIANN-CDR, although the Indo-Pacific region has much higher values than the previous CFSv2 forecast time frames for summer (Figure 25). The increased precipitation values in the Indo-Pacific may be associated with typhoons and the monsoon season, which create an over prediction of rain for this part of the world. The maximum values for the 55-to-65 day forecasts represent the higher values that the mean has depicted with a decent quantitative spread from southeast Asia to Papua New Guinea. The coverage percentage remains mostly consistent from the previous CFSv2 forecast time frames of summer with an over-prediction globally compared to the PERSIANN-CDR.

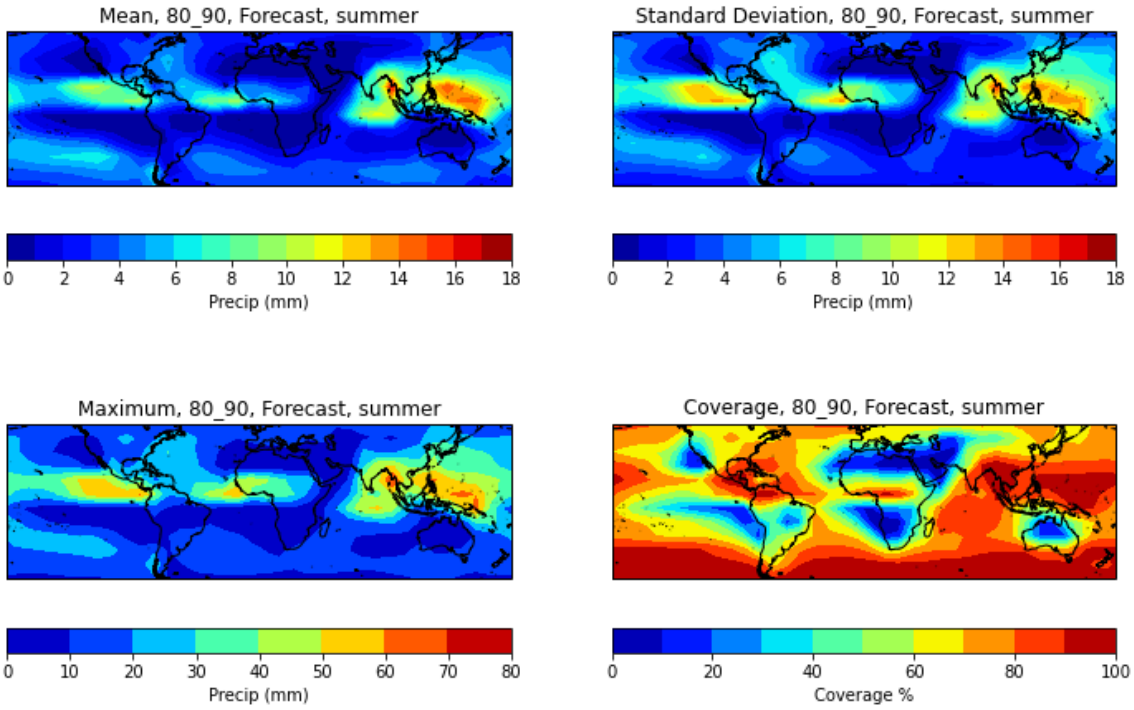

Figure 26. CFSv2 80-to-90 day forecasts mean, standard deviation, maximum, and coverage for summer from 2016 to 2019.

The 80-to-90 day forecasts shows an equivalent representation of the mean and standard deviation from the 55-to-65 day forecasts of summer with only slightly lower amounts over the Indo-Pacific (Figure 26). The maximum values are slightly lower as well, which support the decreased mean amount near southeast Asia. All of the statistics in the 80-to-90 day forecasts remain higher than the PERSIANN-CDR, which support the idea that the CFSv2 has stronger seasonal signals than a satellite dataset. The next section of analysis focuses on specific time series of individual point locations for the Tropics and the Mid-Latitudes, specifically Colorado, for different precipitation data manipulation types. The analysis of two different climate locations on the globe give stronger conclusions of seasonal precipitation patterns.

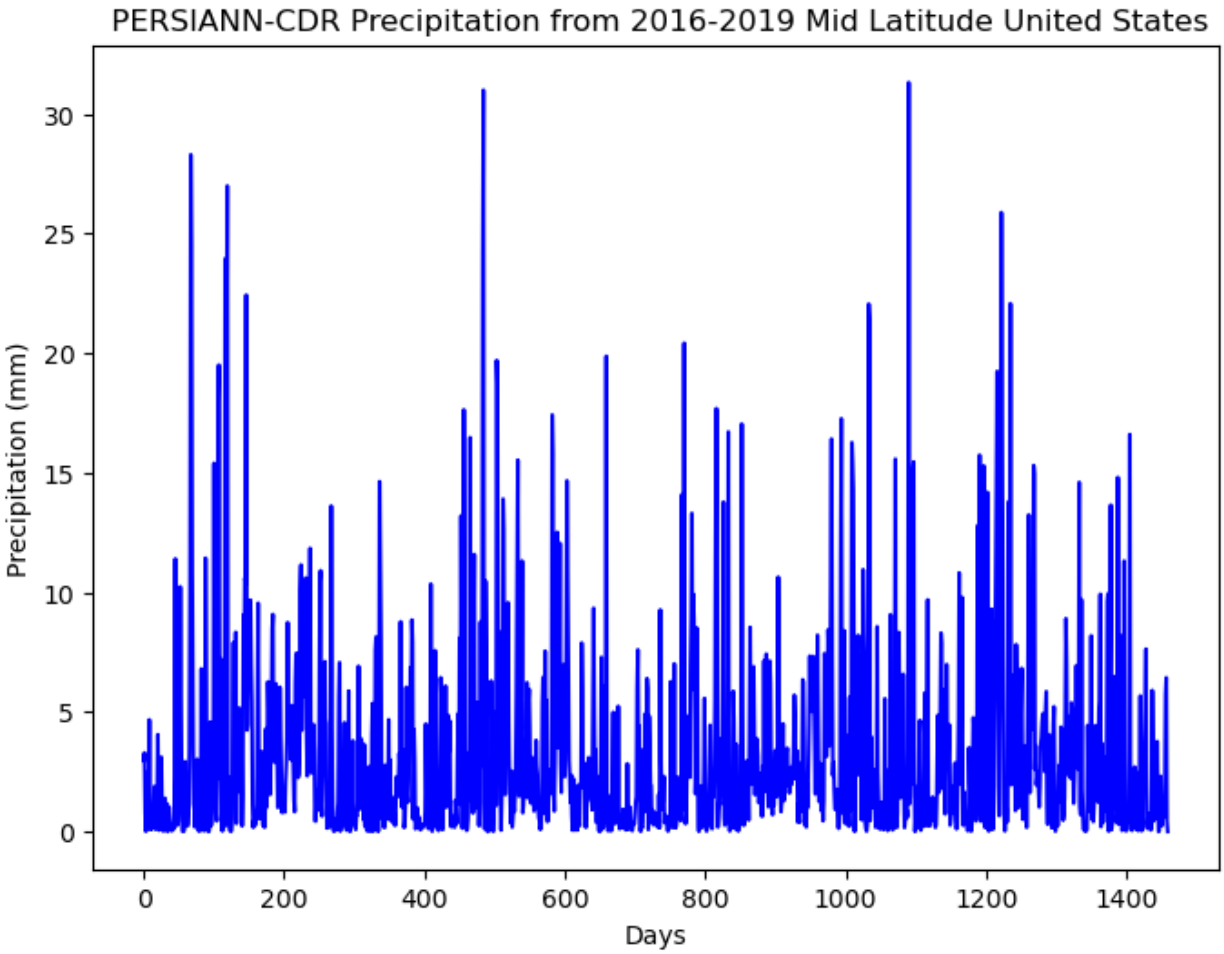

Figure 27. PERSIANN-CDR time series of raw mean precipitation for 2016-2019 in the U.S.

The last four years of the PERSIANN-CDR dataset raw mean precipitation depict an interesting trend in the United States (Figure 27). It is explicitly around the plains U.S. region with a datapoint of latitude 45 N and longitude 105 W, Colorado's state, and it can be seen that the highest precipitation values are in the late winter and early spring. This is due to more active weather patterns in the Mid-Latitudes during this time frame, with lower values tapering off towards the end of each year. The year 2018 does not peak compared to the other three years, indicating a less active heavy rainfall season.

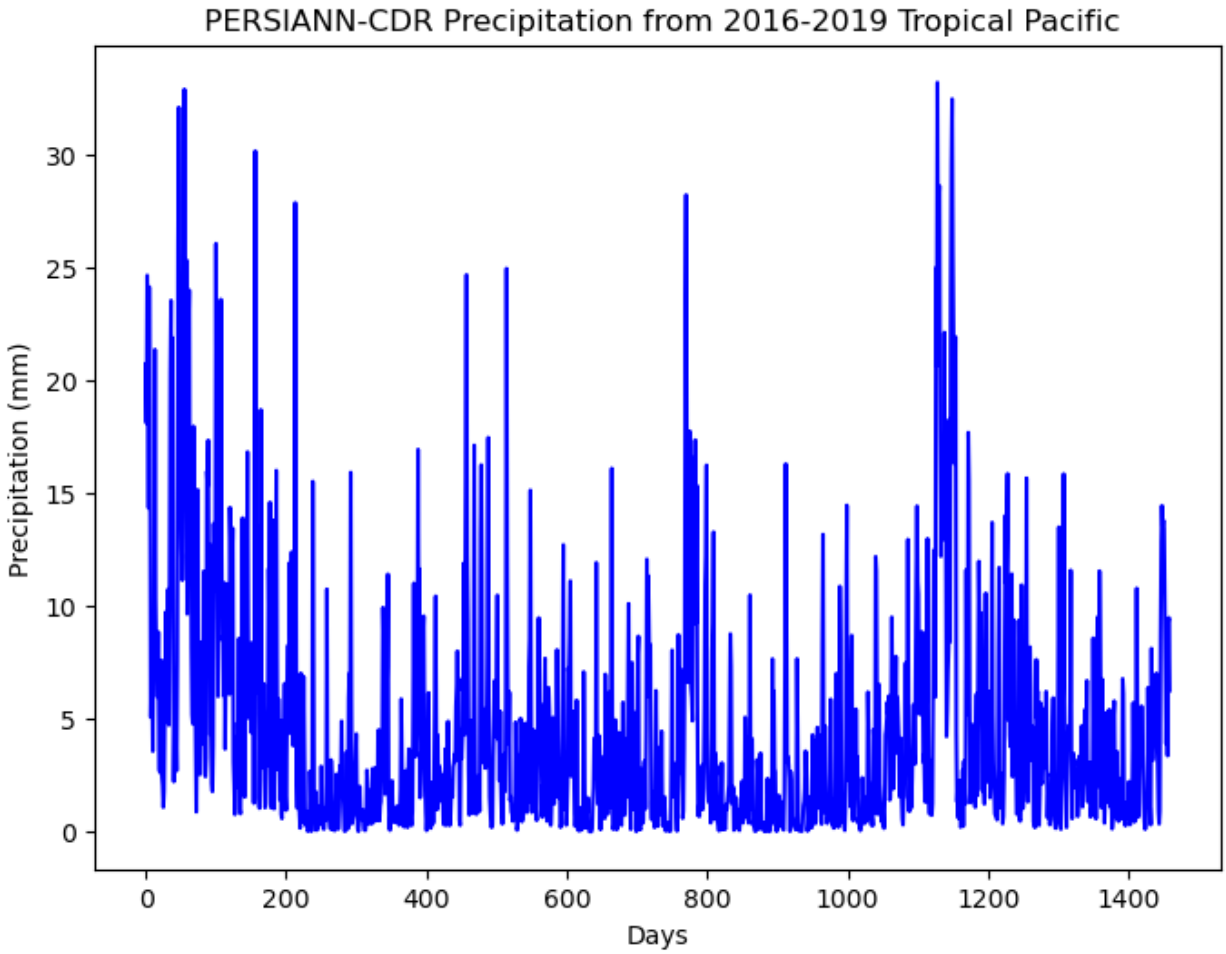

Figure 28. PERSIANN-CDR time series of raw mean precipitation for 2016-2019 in the Tropics.

The Tropics region near the international dateline with a latitude of 5 N and longitude 175 E have much higher precipitation peak values compared to the U.S. above 30 mm at the beginning of 2016 and 2019. There is much less variability in precipitation than in the U.S., which line up with the consistency of rainfall in this region of the globe (Figure 28). On average most of the precipitation remains at or below 15 mm daily, but the values are almost always above 0, which is different from the Mid-Latitudes, which depicted more days with 0 precipitation. The raw mean plots are useful in identifying trends during each season and throughout each year.

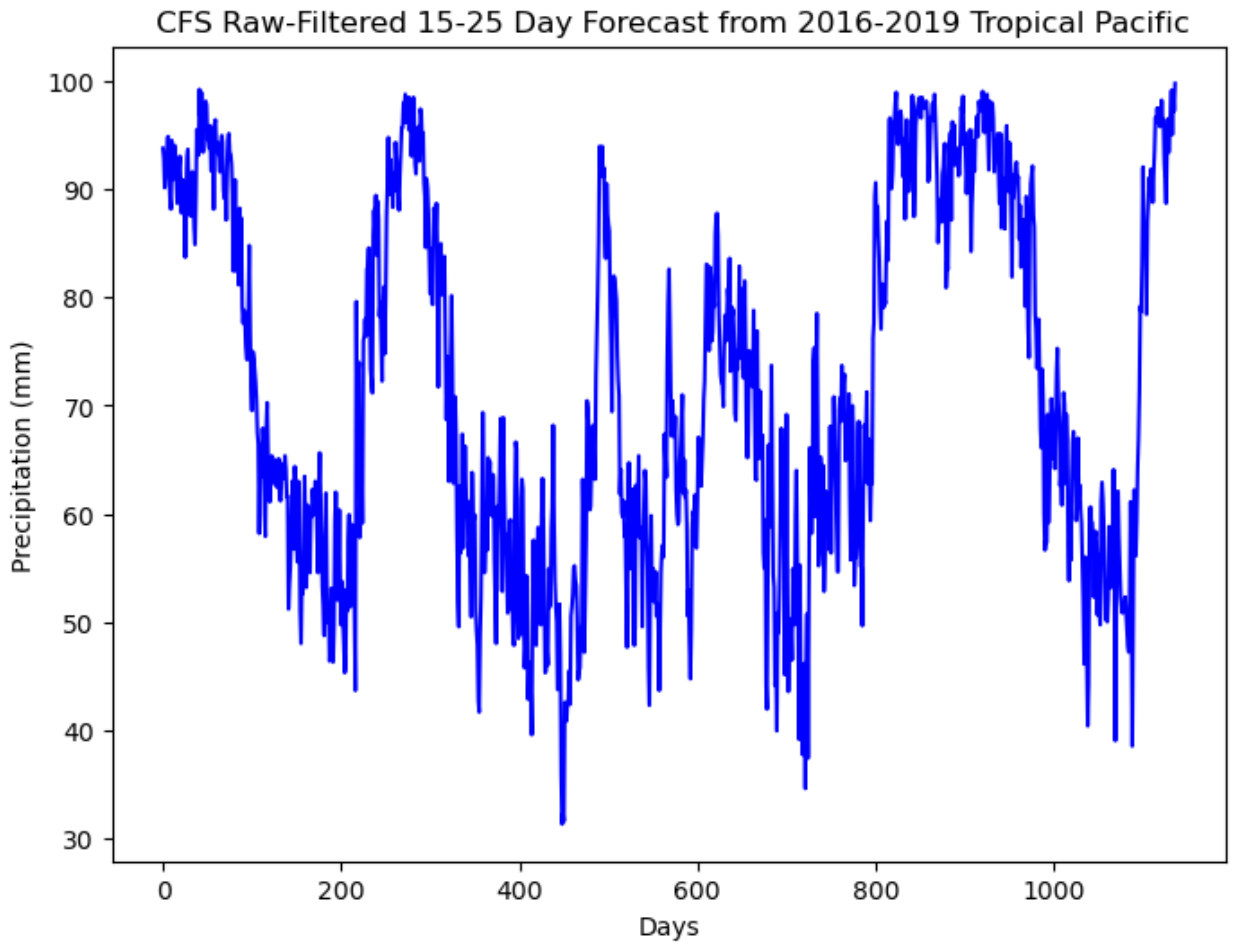

Figure 29. CFSv2 15-to-25 day forecasts of raw filtered time series from 2016 to 2019 in the Tropics.

The CFSv2 15-to-25 day forecasts is the only time period plotted for the raw and Savgol filters because of the accuracy of statistical values compared to the PERSIANN-CDR (Figure 29). The main difference in comparison to the PERSIANN-CDR is the number of days on the x-axis, which is because there is a decent amount of missing precipitation data from the CFSv2 for certain days. The missing days are masked while the raw filtered technique allow for specific seasonal trends more identifiable than the mean. In the Tropics, the peaks are around 100 mm, and the lowest value is above 30 mm, which supported the overestimation of model precipitation in this region. Each peak

is at the beginning of 2016, 2017, and 2019 with 2018 being the exception, which agreed with the PERSIANN-CDR. There is still a decent amount of variability but much less noise compared to the raw mean.

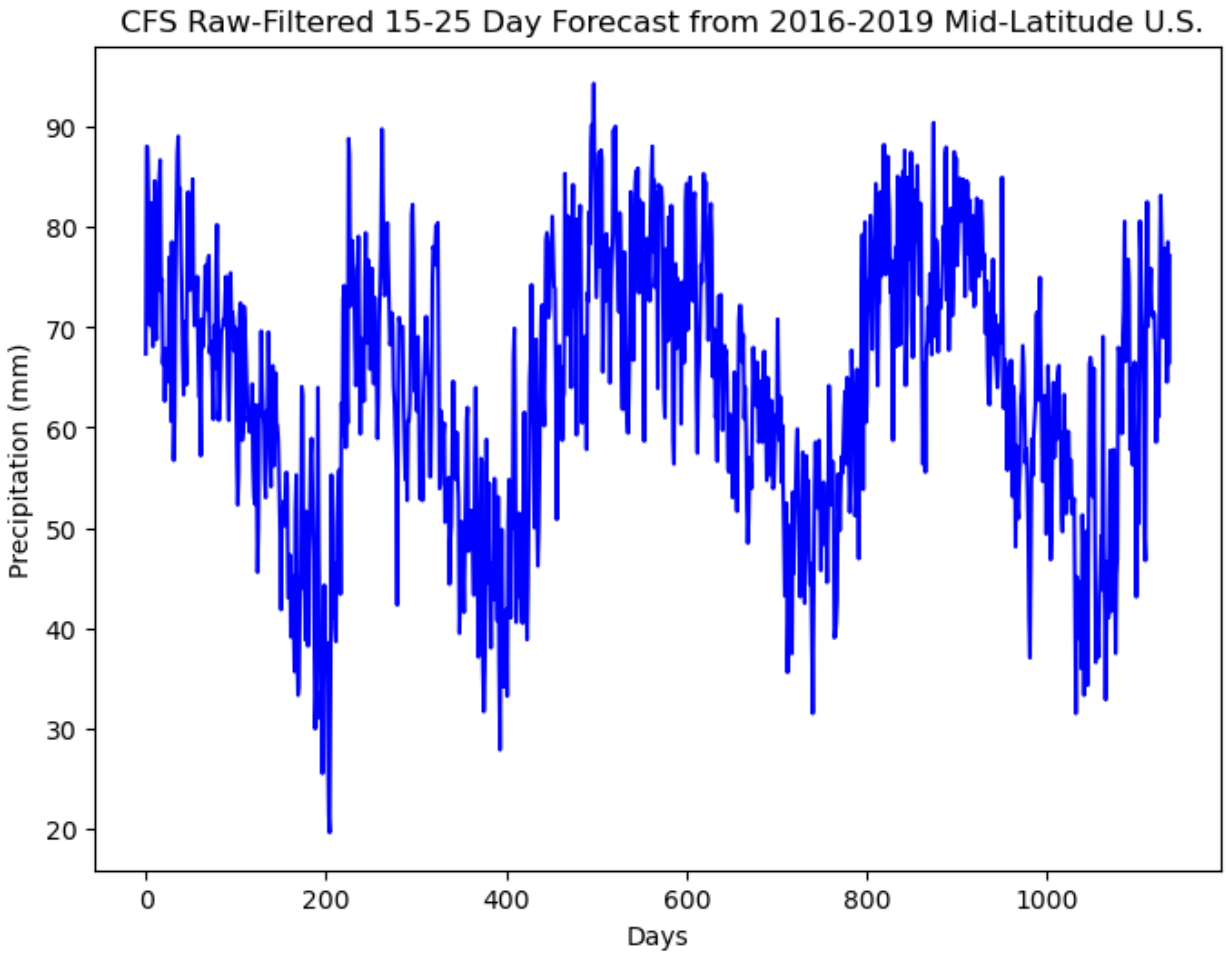

Figure 30. CFSv2 15-to-25 day forecasts of raw filtered time series from 2016 to 2019 in the U.S.

The variability of data for the U.S. is significantly higher than the Tropics in the filtered plots (Figure 30). The seasonal trends of precipitation are very evident with the peaks located in the winter and spring. The lowest values are in the summer months, especially in 2016, with 20 mm noted. The highest value is in 2017, above 90 mm in the

spring months. Another trend noted in the filtered plots is precipitation increased in the fall, which indicates a transition time from summer to winter.

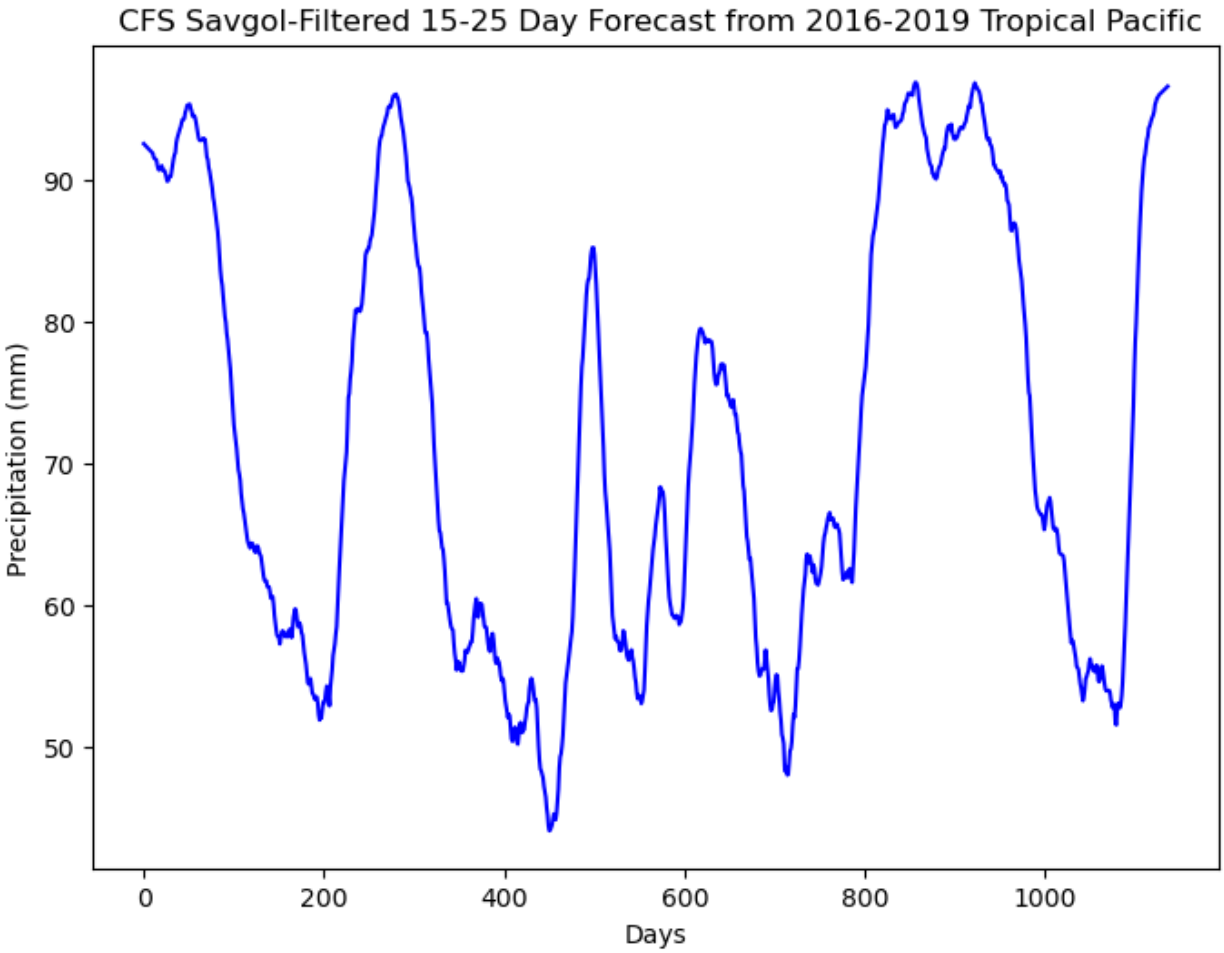

Figure 31. Savgol filter applied to the CFSv2 15-to-25 day forecasts for 2016-2019 in the Tropics.

The Savgol filter is finally applied to the CFSv2 data, which allows the time series to depict reduced noise but keep the integrity (Figure 31). The Savgol filter for the Tropics notes four distinct peaks above 90 mm of precipitation, which line up with the four years of data. The lowest values hover around 50 mm, with only two valleys dropping below this value, which indicates high precipitation events being reasonably

common. The 2018 data shows much lower rainfall values than the other years, which line up with the raw filtered CFSv2 data.

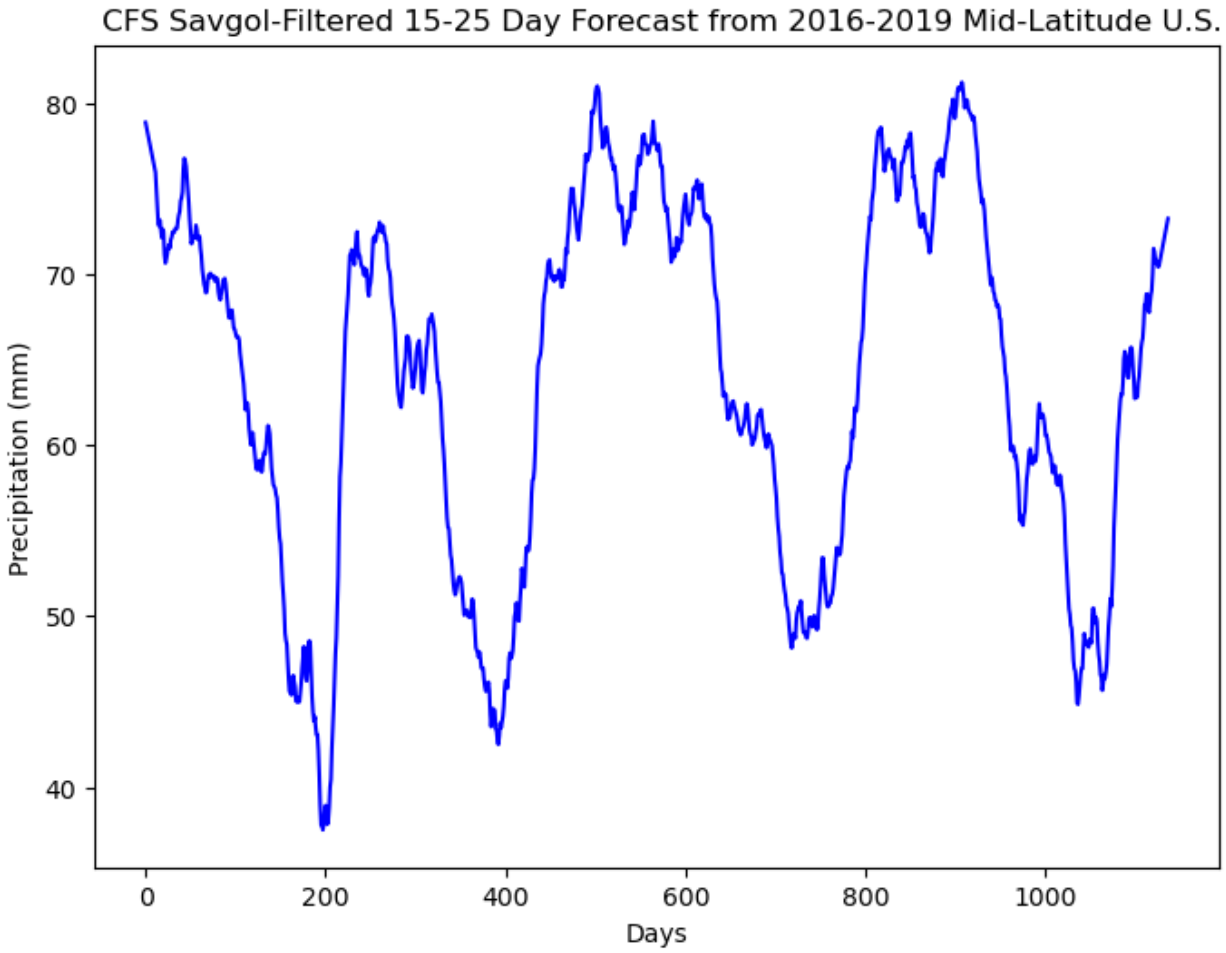

Figure 32. Savgol filter applied to the CFSv2 15-to-25 day forecasts for 2016-2019 in the U.S.

The Savgol filter applied to the Mid-Latitude data only has peak values around 80 mm and minimum values below 40 mm (Figure 32). All of the peaks are associated with each year of the dataset, which remains consistent from 2016 to 2019. The variability in the peaks and valleys is still evident with the filtered technique applied, which shows that it is near impossible to completely smooth Mid-Latitude daily precipitation forecast data. These plots have established the relationships between basic statistical parameters

between the observed and forecast data. Also, the raw mean, filtered, and Savgol filtered show how much more efficient in accuracy the quantitative results are in comparison to global analysis of precipitation patterns.

The verification statistics are necessary to prove the CFSv2's accuracy at the different forecast time frames 0-to-10, 15-to-25, 55-to-65, and 80-to-90 days. This is done by creating correlation plots to determine the statistical variables' relevancy compared to the CFSv2 observational data. It was necessary to determine the accuracy of forecast durations and lengths done with analytical, statistical plots rather than eyeballing on a global visual scale with significant errors that could be made. It is expected that the 0-to-10 day forecasts will yield higher significant values compared to the other three forecast time frames. These first sets of plots use the Savgol filter with window length 21 and order 1, which suggest that the smoothing of the data is more significant.

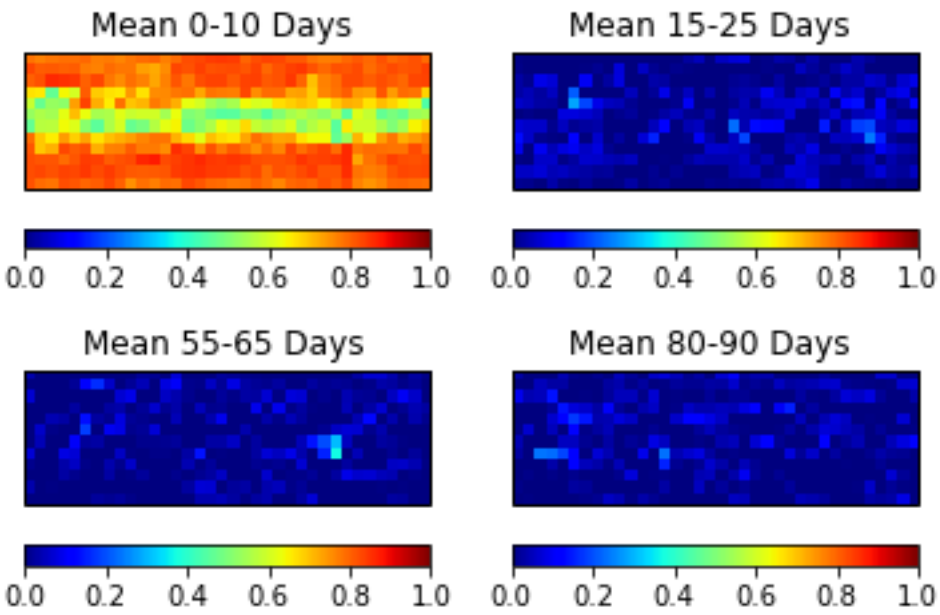

Figure 33. Pearson Correlation plot for CFSv2 all forecast day ranges of mean. (create a map and change color bar down to 0.5)

The highest correlation is located above and below the equatorial region for the 0-to-10 days, with the lowest values located in the Tropics region (Figure 33). While in contrast, the other forecast timeframes show little correlation, so this alone is not the most effective way to determine accurate verification results. These lower values in the 0-to-10 day may be associated with oceans and the higher values located over the land regions since more accurate precipitation data is located on the continental locations.

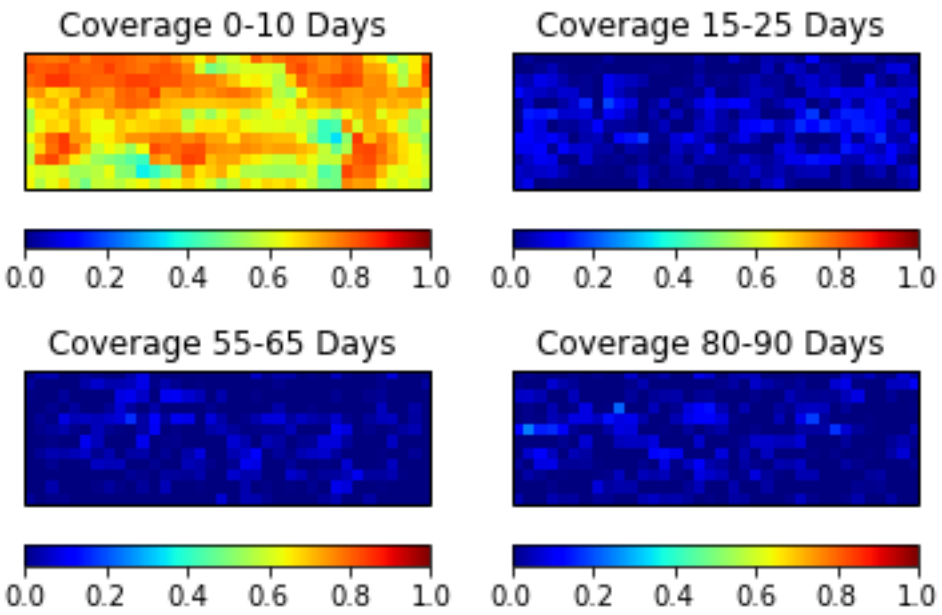

Figure 34. Pearson correlation plot for CFSv2 all forecast day ranges of coverage.

The coverage variable notes similar results from the mean with high correlation only evident in the 0-to-10 day period associated with the landmasses (Figure 34). It is challenging, but a considerable difference between the 0-to-10 day and 15-to-25 day plots is where the p-value is above the 95% confidence interval, indicating the statistical significance of the observed difference. The null hypothesis rejected is the CFSv2 forecast and observed correlation occurred by chance. The 0-to-10 day p value is 432, which indicates a lower chance to reject the null hypothesis compared to the 15-to-25 day value of 122, which is a much higher chance to reject the null hypothesis.

Only the p-value being slightly higher up to 148 for the coverage parameter for days 55 to 65 is calculated, and parallels from 15-to-25 days. The p-value is significantly lower, though, for this forecast period with a value of 95, which showed the best time frame to reject the null hypothesis. The fact that the correlation is zero beyond 30 days

supports the claim that models are unreliable with precipitation forecasts beyond one month from previous literature. The 80-to-90 day forecast plots do not show any different information with p values equal to or less than the 55-to-65 day forecast. The next set of plots have shown the Savgol filter with a window length of 15 and order 9, which indicates that the correlation values should be lower for the 0-to-10 day forecast compared to observed.

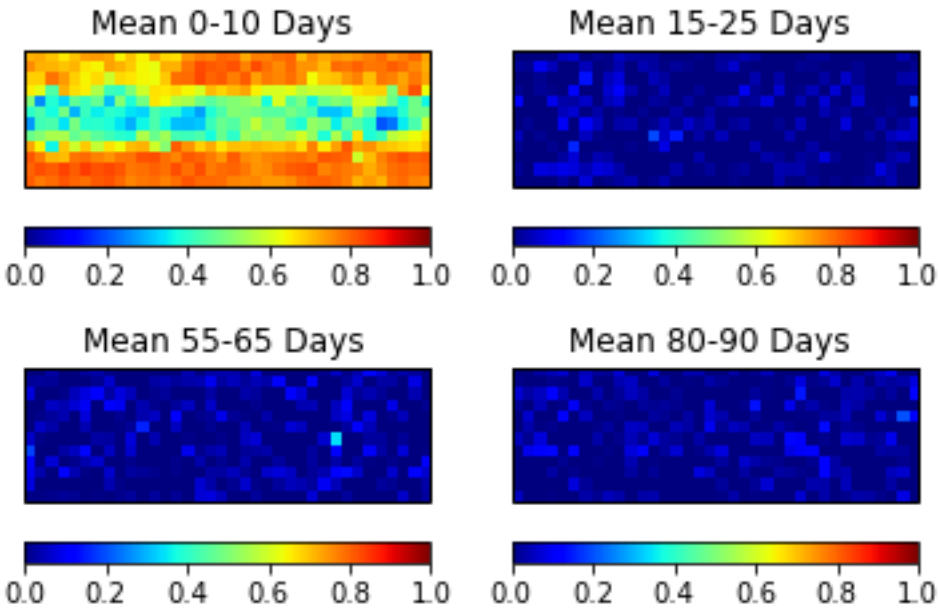

Figure 35. Pearson correlation plot for CFSv2 all forecast day ranges of mean.

The mean with the new window length and order applied depict a much larger area of low correlation in the Tropics, which spread to the lower parts of the Mid-Latitudes in both hemispheres (Figure 35).

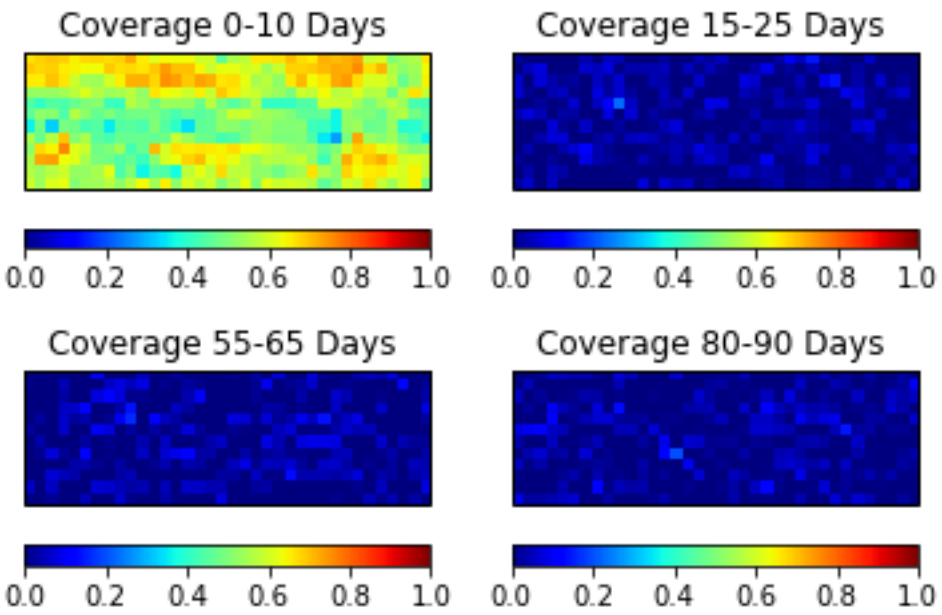

Figure 36. Pearson correlation plot for CFSv2 all forecast day ranges of coverage.

The coverages illustrate that most of the globe has low correlation values even in the Tropical regions for the 0-to-10 days (Figure 36). This is all driven by the window length and order of the Savgol filter, which show the importance of smoothing the data as efficiently as possible. The following forecast periods have little to no correlation with the p values dropping off significantly starting at the 15-to-25 day range. Around the 80-to-90 day range, there was a p-value of 80, which is the lowest value.

Contingency tables are then created to show the accuracy of forecasts in a confusion matrix, which depicts the specific metrics. This part of the analysis provides the most useful quantitative values for the performance of the CFSv2 at different forecast times. The different metrics are shown for every forecast period along with each of the statistical parameters such as mean, standard deviation, maximum, and coverage. The metrics include true positive and negative, which is the equivalent to a hit in operational

meteorology when the model and human forecasters are compared to the actual observations. The other metrics are false positive, which is a false alarm, and false-negative, which indicates that precipitation fell, but it is not forecasted.

Contingency Table for Forecast Days 0-10		
	Occurred	Did Not Occur
Forecasted	145731	64272
Not Forecasted	47920	269117

Table 1. False alarm, hits, and misses of precipitation for 0-to-10 day forecasts. (add map to each contingency table)

The metrics for the first forecast time of 0-to-10 days depicts a hit rate of 78% for rain or no rain within each grid box (Table 1). The false alarm rate is 13% when rain was forecasted to occur but did not from the observations. The miss rate of precipitation globally for the CFSv2 is 9%, which showed that short term forecasts are handled relatively accurately.

Contingency Table for Forecast Days 15-25		
	Occurred	Did Not Occur
Forecasted	87415	143509
Not Forecasted	101943	182509

Table 2. False alarm, hits, and misses of precipitation for 15-to-25 day forecasts.

The 15-to-25 day forecasts hit rate for precipitation is 53%, significantly lower than 0 to 10 days (Table 2). The false alarm rate is 27%, which is twice as high as the short-term forecast rate. The miss rate of precipitation is 20%, which is also twice as high as the previous forecast. The further out in time, the models' accuracy drops significantly, especially between the short and medium-range forecasts.

Contingency Table for Forecast Days 55-65		
	Occurred	Did Not Occur
Forecasted	80875	136513
Not Forecasted	97715	170465

Table 3. False alarm, hits, and misses of precipitation for 55-to-65 day forecasts.

The 55-to-65 day forecasts hit rate for precipitation is 52%, which is slightly worse than the 15-to-25 days but not significantly (Table 3). The false alarm rate for overpredicting rain globally is 28%, which is also not much lower than the previous medium length. The miss rate is 20%, equivalent to the 15-to-25 day range miss of precipitation events. It is surprising that the further out beyond 30 days of forecasting, the CFSv2 does not show significantly worse metrics, which indicates that the graphical depiction is exponential with a plateau about halfway across.

Contingency Table for Forecast Days 80-90		
	Occurred	Did Not Occur
Forecasted	80229	143509
Not Forecasted	95599	168424

Table 4. False alarm, hits, and misses of precipitation for 80-to-90 day forecasts.

The 80-to-90 day forecasts hit rate for precipitation is 51%, which is slightly worse than the 55-to-65 days (Table 4). The false alarm rate for the precipitation forecasts is 29%, which is slightly higher than the previous forecast. Finally, the miss rate of precipitation events globally is 20%, equivalent to the 15-to-25 and 55-to-65 day miss rates. The miss rates would depict a much more drastic exponential curve beginning with it rapidly flattening about halfway across. The main piece to note is that the hit rate dropped as we went further out in time while the miss rate increased, which is another piece of quantitative evidence that the model is not reliable beyond 30 forecast days. The

15-to-25 day forecast range verification results provide the best reason for why QM was only applied to this time frame.

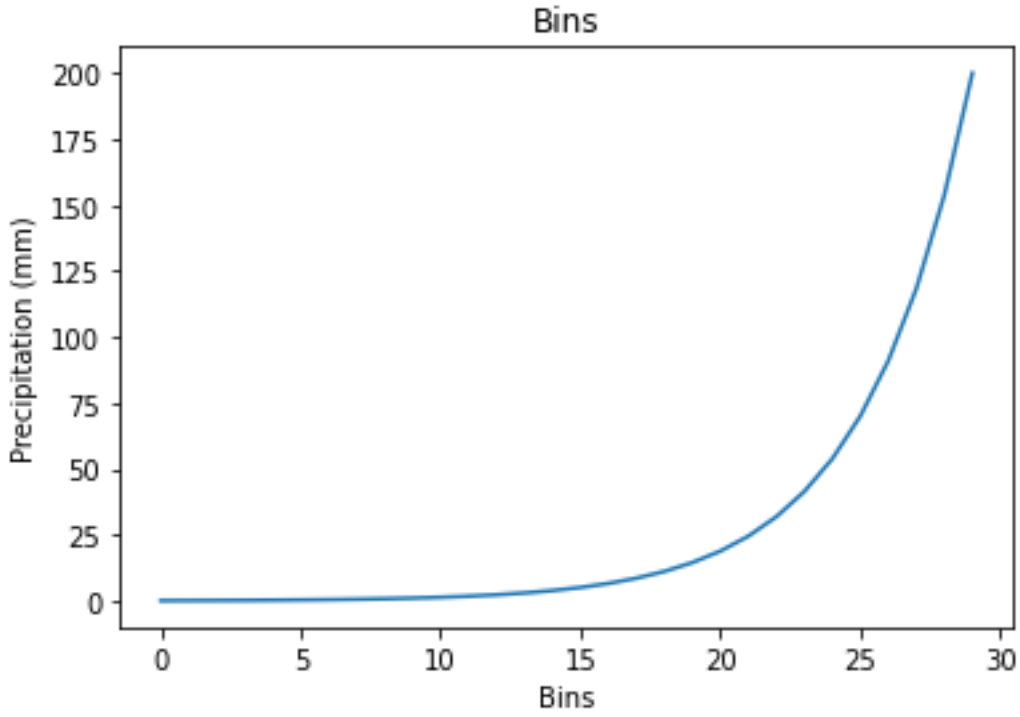

Figure 37. Line graph depiction of the 30 bins created for the CDF, histogram, and QM.

Bins are then created to produce histogram and CDF plots which is ultimately how the QM was applied to the CFSv2 data (Figure 37). There are 30 bins created, with each of them containing precipitation values as low as 0.1 mm up to 200 mm, which cover the entire spread of data for the PERSIANN-CDR and the CFSv2. The maximum value statistical parameter is used for this part of the analysis to ensure that all of the data is covered on the upper-end extreme.

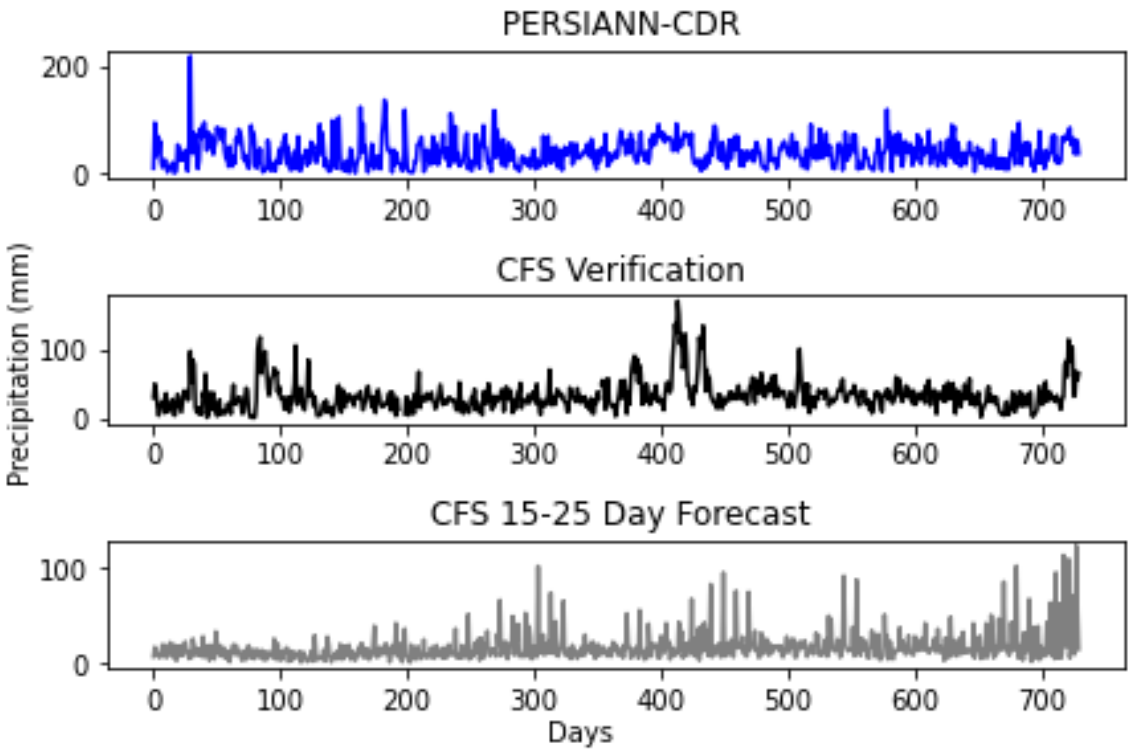

Figure 38. Three panel time series plots of raw data of PERSIANN-CDR, CFS verification, and CFS 15-to-25 day forecasts in the Tropical Pacific for 2018-2019.

All of the time series plots for the raw precipitation data identified the CFSv2 15-to-25 day forecast accuracy and limitations compared to the PERSIANN-CDR (Figure 38). The Tropics location is analyzed for determination of forecast results. Toward the beginning of the PERSIANN-CDR dataset, there is a significant spike of over 200 mm of precipitation, which is not visible in either of the CFSv2 sets of data. The 15 to 25-day forecasts of the CFSv2 seem to underpredict the higher amounts of precipitation while accurately depicting little to no rain for a specific day. The end of the time series for the CFSv2 15-to-25 day forecasts depict accurate precipitation results.

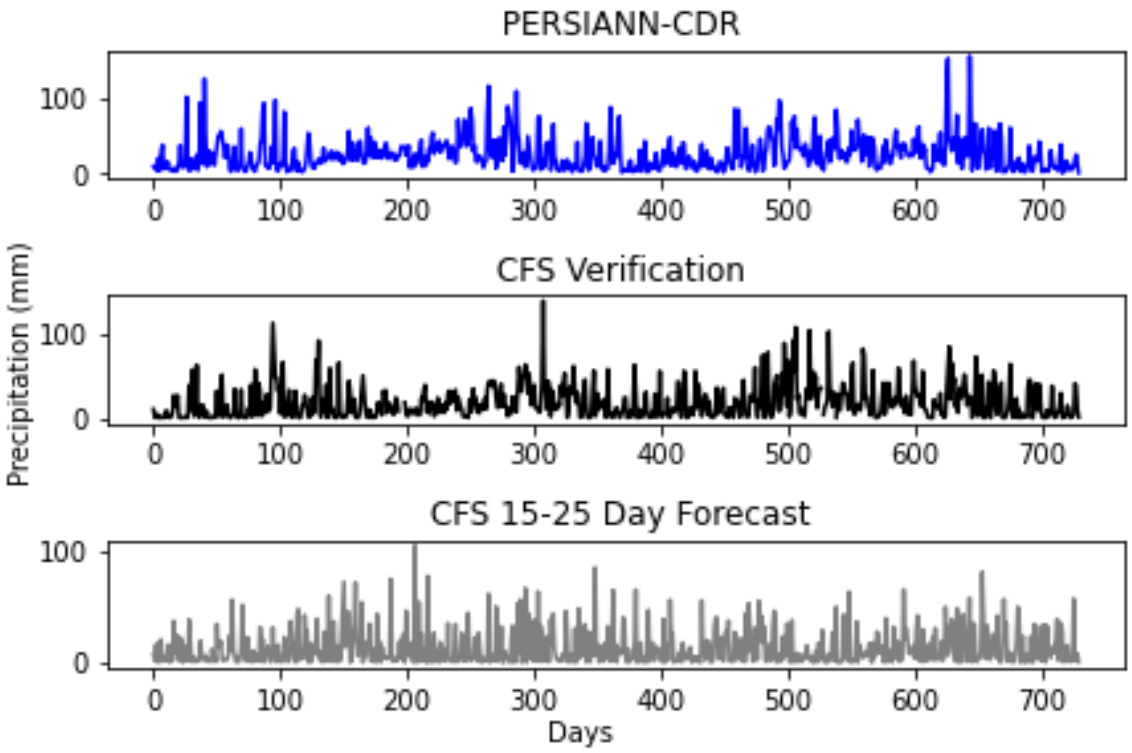

Figure 39. Three panel time series plot of raw data of PERSIANN-CDR, CFS verification, and CFS 15-to-25 day forecasts in the U.S. for 2018-2019.

The U.S. is similar to the Tropics, with the PERSIANN-CDR precipitation values higher compared to the CFSv2 (Figure 39). The 15-to-25 day forecasts attempts to represent the higher precipitation events, but these are not as common as a Tropical location, so the accuracy improves in the Mid-Latitudes. There is higher variability in the maximum and minimum values throughout each of the datasets, especially the PERSIANN-CDR and the CFSv2 verification. Since the raw data has been compared, it is necessary to then analyze the three datasets histograms and look to see the peak locations.

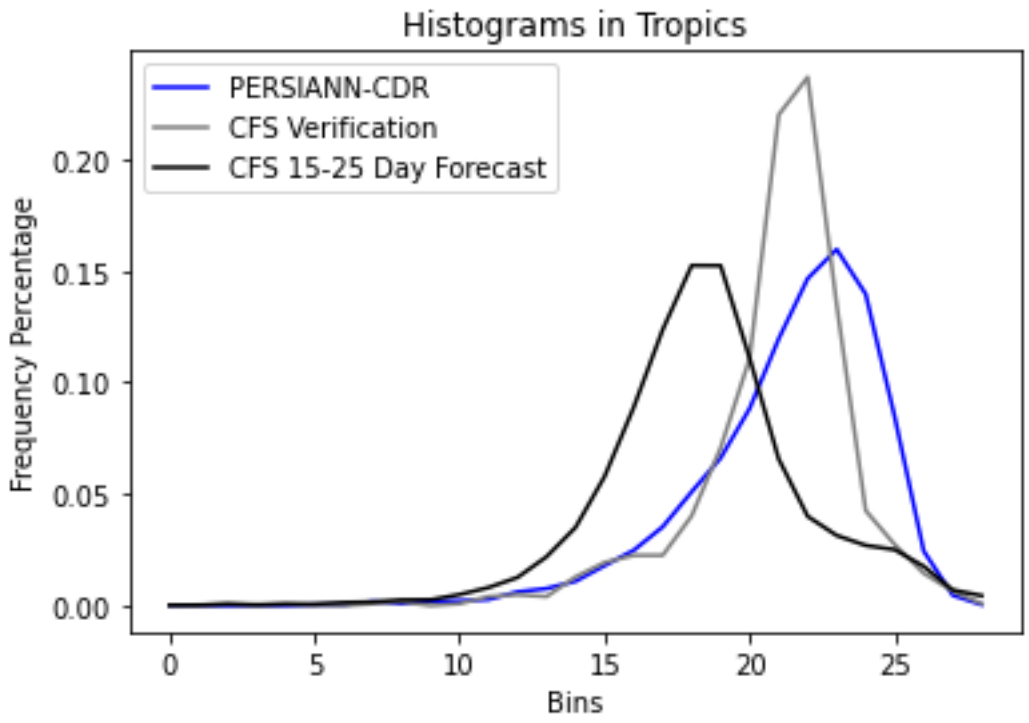

Figure 40. Histogram plots of PERSIANN-CDR, CFSv2 verification, and CFSv2 15-to-25 day forecasts in the Tropical Pacific.

Histogram plots are created for all three datasets, which identify peaks in the different bin locations (Figure 40). The higher the bin then, the larger the precipitation event that occurred, which is represented by that specific dataset. The 15-to-25 day CFSv2 forecast has a peak around the 18th bin while the verification is around the 22nd at a higher percentage. The PERSIANN-CDR peaks around the same height as the forecast but, around bin 24.

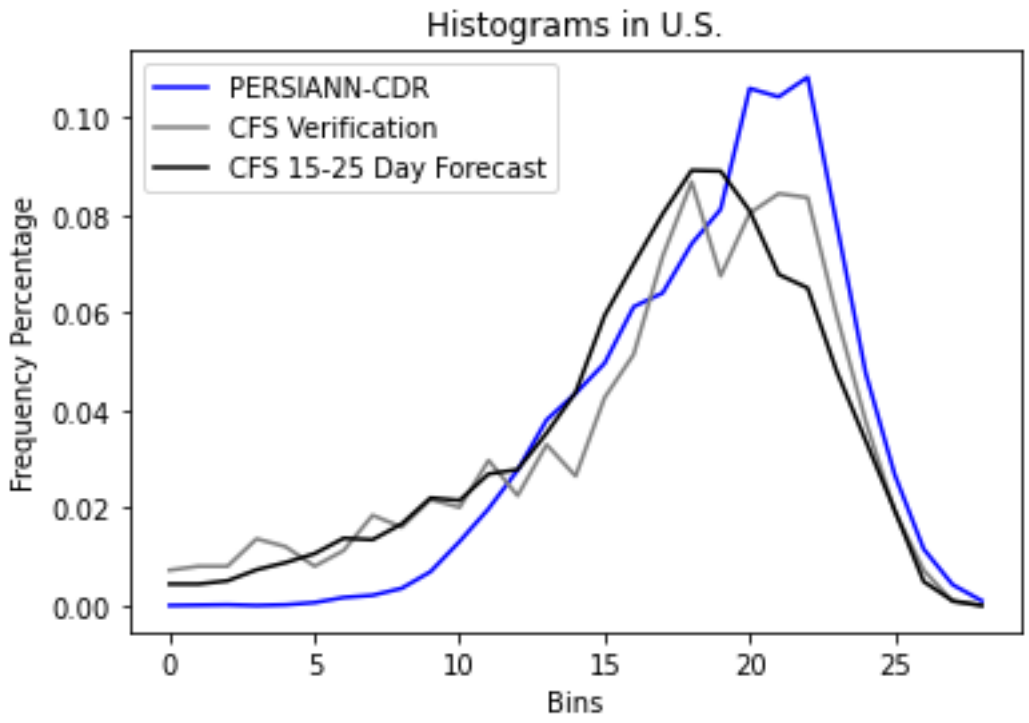

Figure 41. Histogram plots of PERSIANN-CDR, CFSv2 verification, and CFSv2 15-to-25 day forecasts in the U.S.

The main difference from the U.S. was that the histograms are noisy, which means there is much more variability in precipitation events (Figure 41). The CFSv2 verification plot depicted with the jagged line toward the bottom of the bins that there is no smoothed data, and the forecasted CFSv2 has some variability toward the lower end of bins created. The peaks are also sporadic for the PERSIANN-CDR and the CFSv2 verification, but the CFSv2 15 to 25 days are flat at the peak bin. The peak bins for all of the U.S. histogram plots are located at the 20th bin. The last set of 3-panel plots with these datasets consisted of the CDF, which is created by taking the cumulative sum of the histograms and dividing by the length of each dataset.

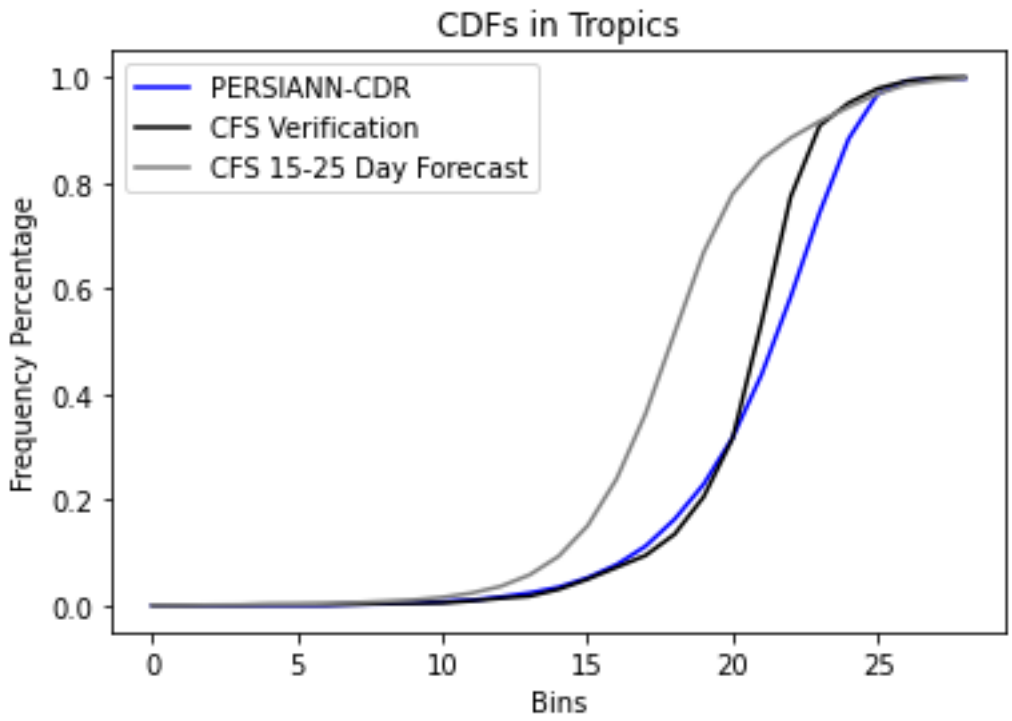

Figure 42. CDF plots of PERSIANN-CDR, CFSv2 verification, and CFSv2 15-to-25 day forecasts in the Tropical Pacific.

Comparison of the three different CDFs in the Tropics produced from the individual datasets suggests the only difference is the forecast CFSv2 did not capture heavy precipitation events as often, which is also noticed in the histograms (Figure 42). Based on the PERSIANN-CDR CDF, this dataset seems to capture precipitation events more accurately than the forecast CFSv2 in the Tropics, which agreed well with the statistical analysis performed on the global scale.

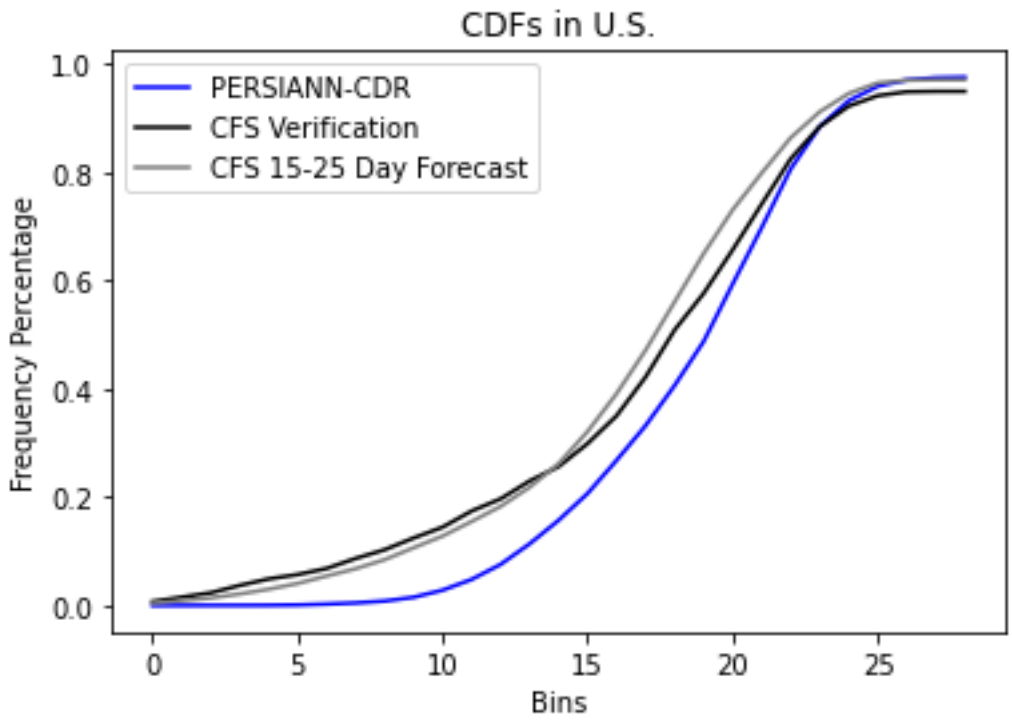

Figure 43. CDF plots of PERSIANN-CDR, CFSv2 verification, and CFSv2 15-to-25 day forecasts in the U.S.

The CDFs plotted for the U.S. are similar to the Tropics, with the exception of the CFSv2 forecast that has a more flattened appearance toward the top (Figure 43). This suggests that the forecast dataset handled the heavier precipitation events in the Mid-Latitudes better than the Tropical Pacific. Also, the CFSv2 verification has a lower amount of high-end precipitation events that are recorded. So, this would make the forecasted precipitation amounts over predicted compared to the Tropical Pacific, which makes sense since precipitation variability is much higher in the Mid-Latitudes. The final part of the analysis is focused on applying QM to the 15 to 25 day forecasts of the CFS using the PERSIANN-CDR and the 0 to 10-day forecast statistics. There is corrected, and the

difference between corrected and raw data plots is created looking at the same two geographic points.

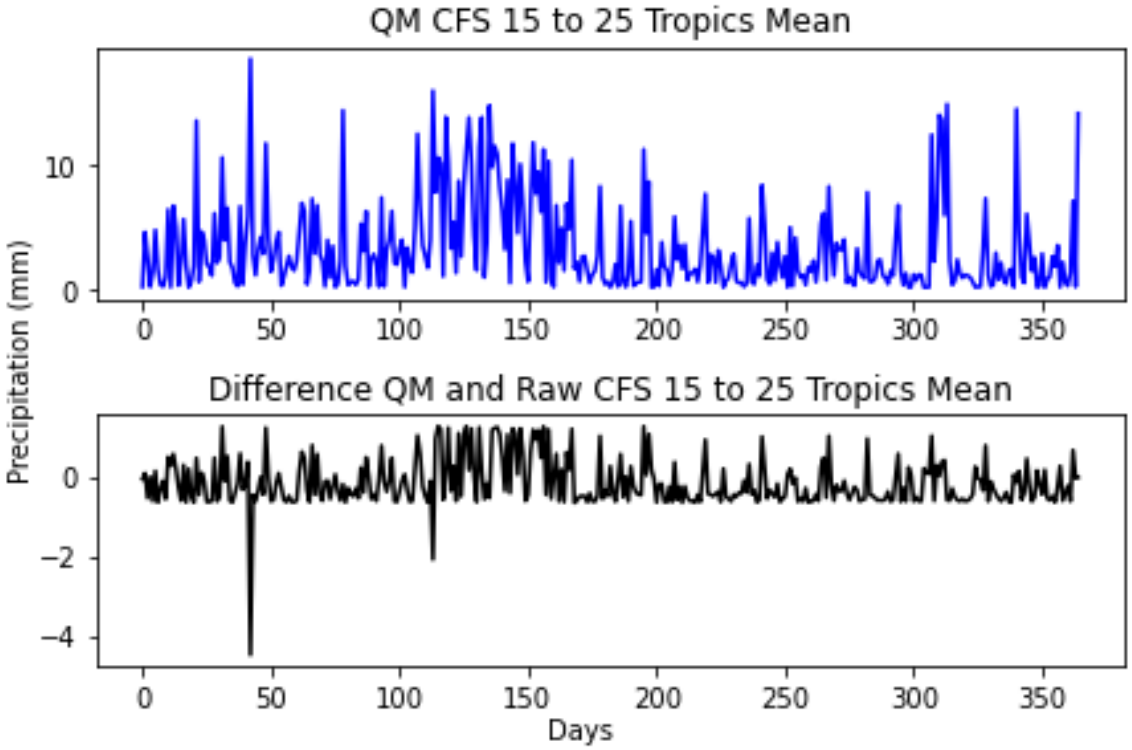

Figure 44. QM correction of 15-to-25 day forecasts CFSv2 mean for the Central Tropical Pacific utilizing PERSIANN-CDR as the observation corrector (top) and the difference between QM corrected minus raw mean (bottom).

The raw CFSv2 data under forecasted mostly throughout the time frame, with the QM having to raise the values slightly (Figure 44). The difference between the two datasets should remain close to 0 for mean values, proving that the data is not biased. There are only two areas that show the raw data is larger than the QM from PERSIANN-CDR corrections. This is supported from the CDFs, which depict a higher frequency of lower precipitation events for CFSv2 15-to-25 day forecasts than the PERSIANN-CDR with more significant rainfalls captured. The four-panel analysis charts for all days with PERSIANN-CDR and CFSv2 15-to-25 day forecasts shows similar results.

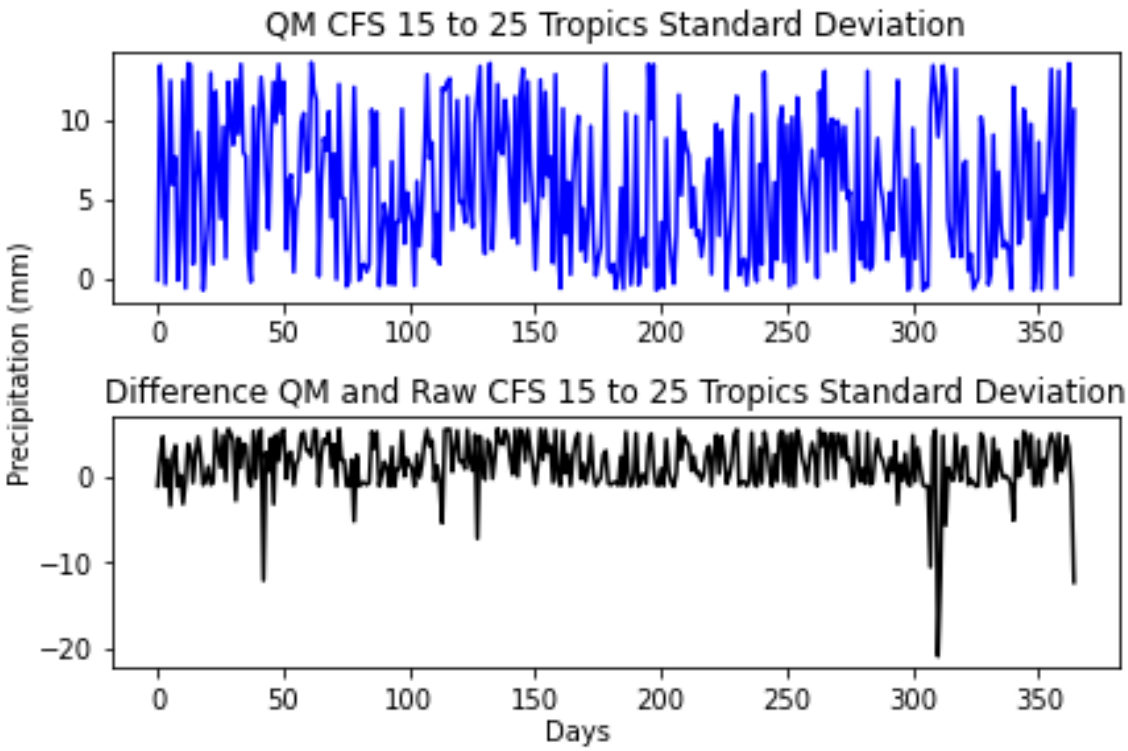

Figure 45. QM correction of 15-to-25 day forecasts CFSv2 standard deviation for the Central Tropical Pacific utilizing PERSIANN-CDR as the observation corrector (top) and the difference between QM corrected minus raw mean (bottom).

The standard deviation values for the Tropics showed large variability of precipitation QM corrections (Figure 45). There are eight noticeable peaks on the negative side of the difference, which means that the raw CFSv2 data has a higher variability. The four plot analyses analyzed for the CFSv2 and the PERSIANN-CDR supports these results over the Tropical Pacific. The QM plot depicts four different peaks

which resemble slight seasonal variation in precipitation.

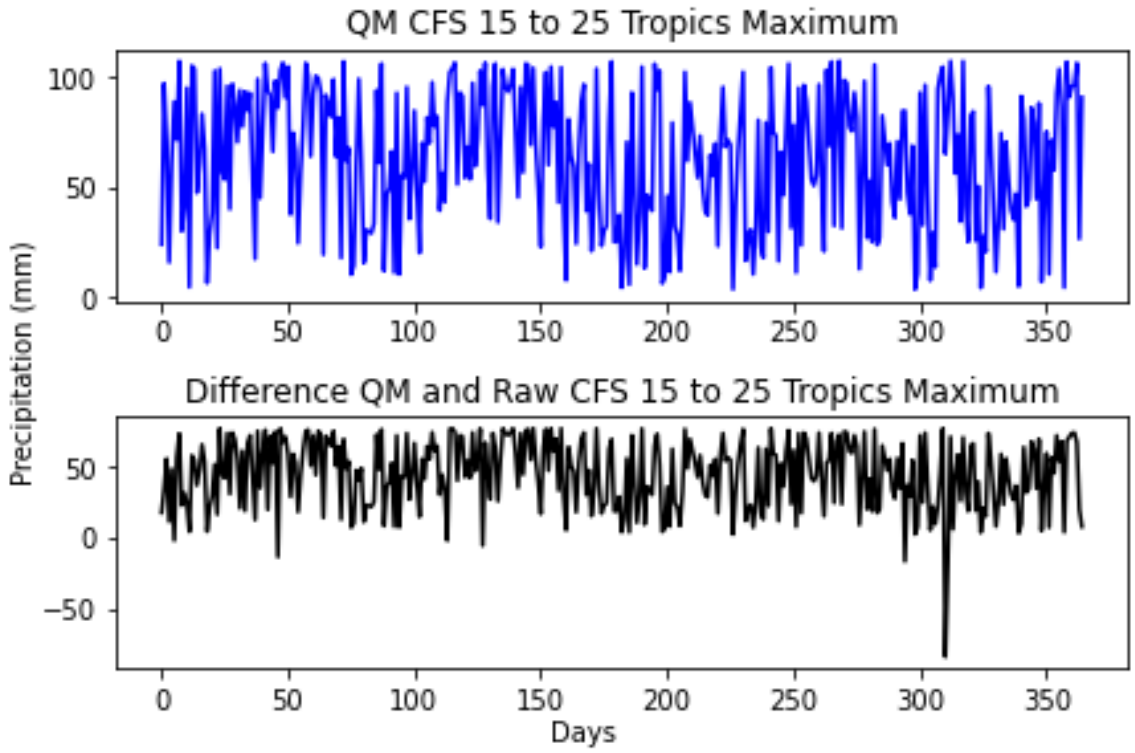

Figure 46. QM correction of 15-to-25 day forecasts CFSv2 maximum for the Central Tropical Pacific utilizing PERSIANN-CDR as the observation corrector (top) and the difference between QM corrected minus raw mean (bottom).

The maximum values corrections shows seasonal variability for the Tropics which suggests that the corrections are most noticeable with this statistical parameter (Figure 46). The QM of the CFSv2 depicts the same seasonal variations that the standard deviation picked up throughout the year. The difference shows that the QM has higher precipitation values than the raw CFSv2 except for one value below negative 50, which indicates a potential anomaly. The anomalous data associated with the corrections should be disregarded since there are very few and it does not impact the integrity of the CFSv2 dataset.

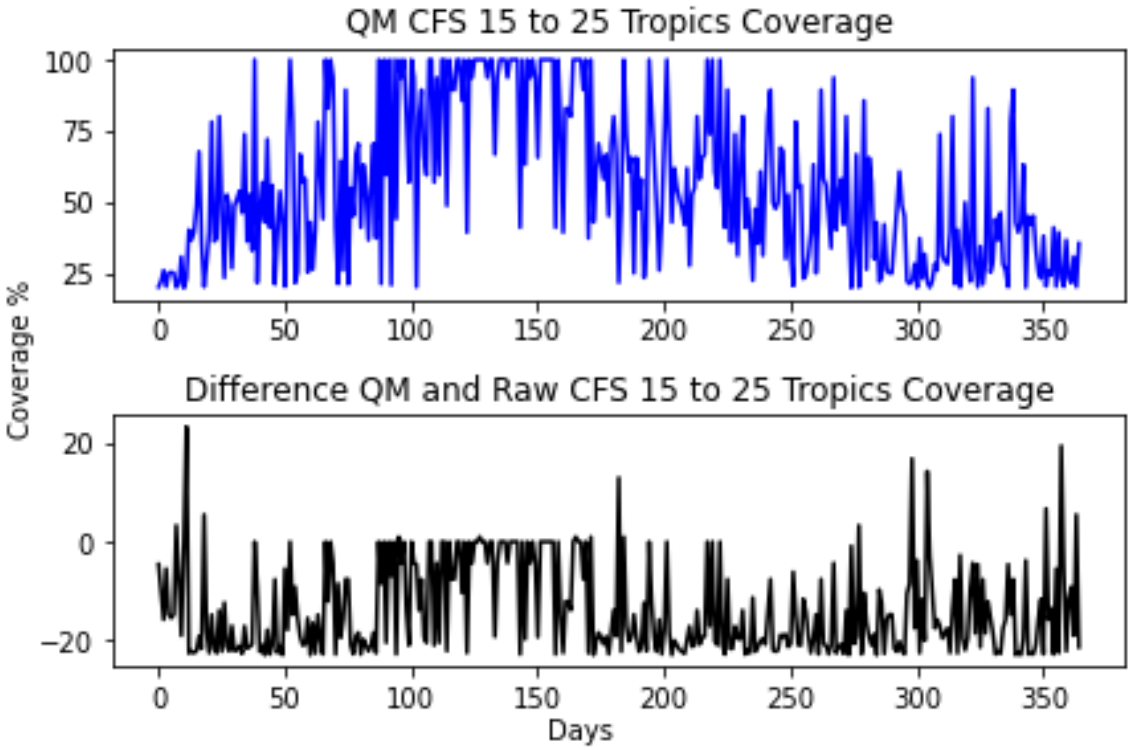

Figure 47. QM correction of 15-to-25 day forecasts CFSv2 coverage for the Central Tropical Pacific utilizing PERSIANN-CDR as the observation corrector (top) and the difference between QM corrected minus raw mean (bottom).

The coverage amounts of precipitation are depicted as a percentage which is the equivalent to the global analysis of the PERSIANN-CDR and CFSv2 (Figure 47). The QM peaks' coverage toward the middle is associated with no difference in the raw CFSv2. Throughout the rest of the year, the CFSv2 has over forecasted precipitation coverage in the Tropical Pacific by amounts of up to 20%. The four panel plots of PERSIANN-CDR coverage compared to CFSv2 15-to-25 day forecasts shows a difference of 30% above the observation amounts on a global scale. There are about five days that the QM values are higher than the CFSv2. The coverage percentage correction supports the over forecasted precipitation in the Tropics region.

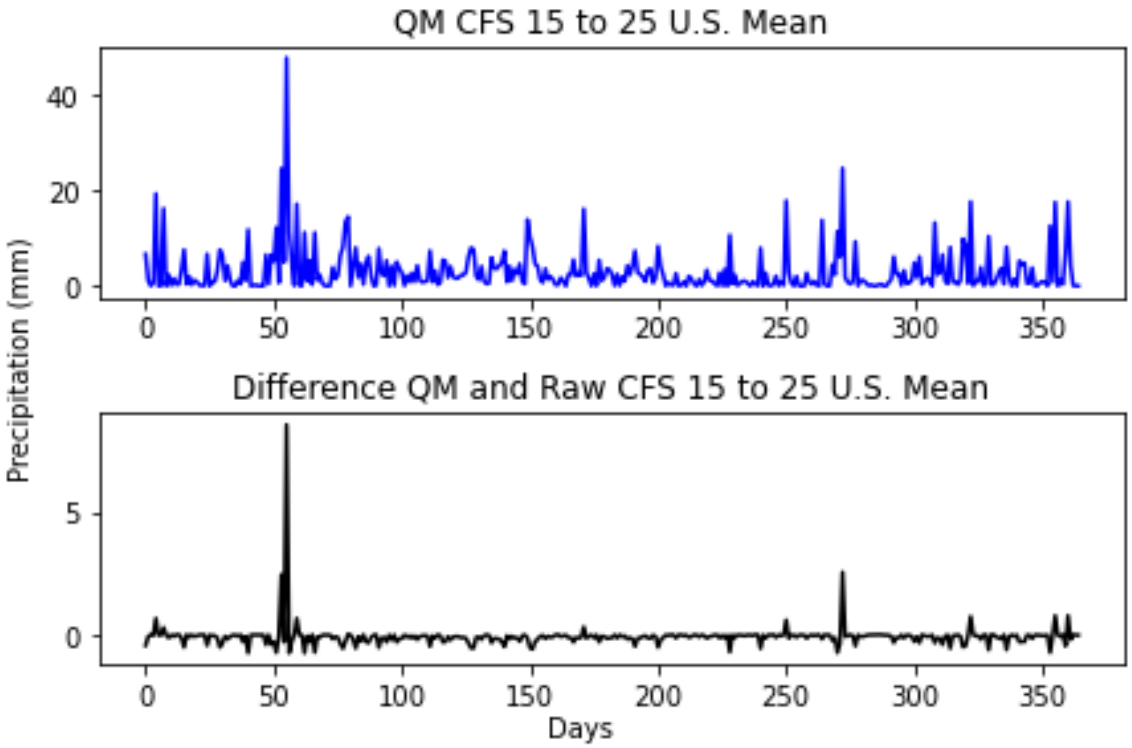

Figure 48. QM correction of 15-to-25 day forecasts CFSv2 mean for the U.S. utilizing PERSIANN-CDR as the observation corrector (top) and the difference between QM corrected minus raw mean (bottom).

The U.S. location QM depicts a lower mean amount of precipitation compared to the Tropics (Figure 48). There are more noticeable peaks throughout the year, with the largest toward the beginning of the dataset. The difference shows that the QM is higher than the CFSv2 raw data by more than 5 mm of precipitation. There is another peak toward the end of the dataset with less than 5 mm difference. The mean difference in the U.S. is closer to 0 compared to the Tropics, which supports the lower variability in high rainfall events.

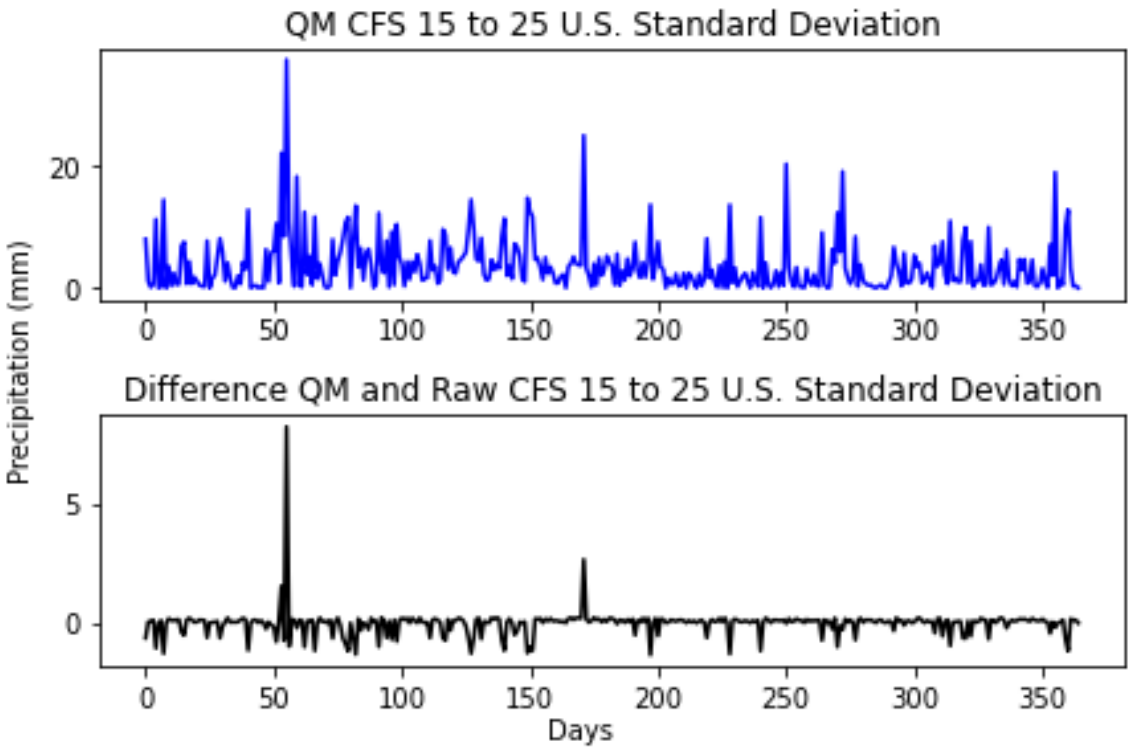

Figure 49. QM correction of 15-to-25 day forecasts CFSv2 standard deviation for the U.S. utilizing PERSIANN-CDR as the observation corrector (top) and the difference between QM corrected minus raw standard deviation (bottom).

The standard deviation values of the QM in the U.S. is much lower compared to the Tropics meaning less variability in rain throughout the year (Figure 49). There is a peak near day 52 along with day 175 for anomalously high precipitation from the QM correction. The four-panel analysis plots in the 15-to-25 day and PERSIANN-CDR supports rainfall variability over the central U.S. for all days. The heavy rainfall events that occurred in the U.S. location for this year impact the mean more than the standard deviation which is also why different statistical parameters are analyzed.

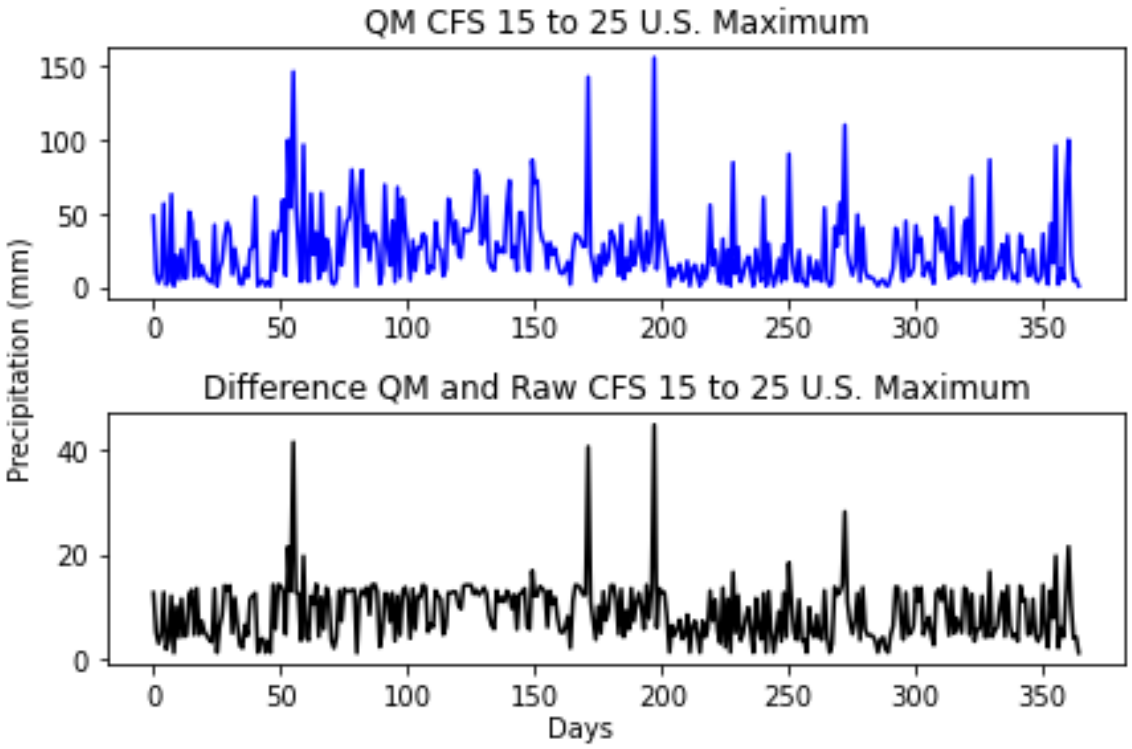

Figure 50. QM correction of 15-to-25 day forecasts CFSv2 maximum for the U.S. utilizing PERSIANN-CDR as the observation corrector (top) and the difference between QM corrected minus raw maximum (bottom).

The QM maximum values show that the U.S.'s significant rainfall events are more significant compared to the Tropics with a lower frequency (Figure 50). The U.S.'s maximum values are higher than the Tropics, with three peaks around 150 mm throughout the entire dataset. The differences depicted the raw CFSv2 under forecasted maximum precipitation amounts in the U.S. throughout most of the year. Maximum value corrections aided in identifying the accuracy of the significant rainfall events and the CFSv2 forecast limitations.

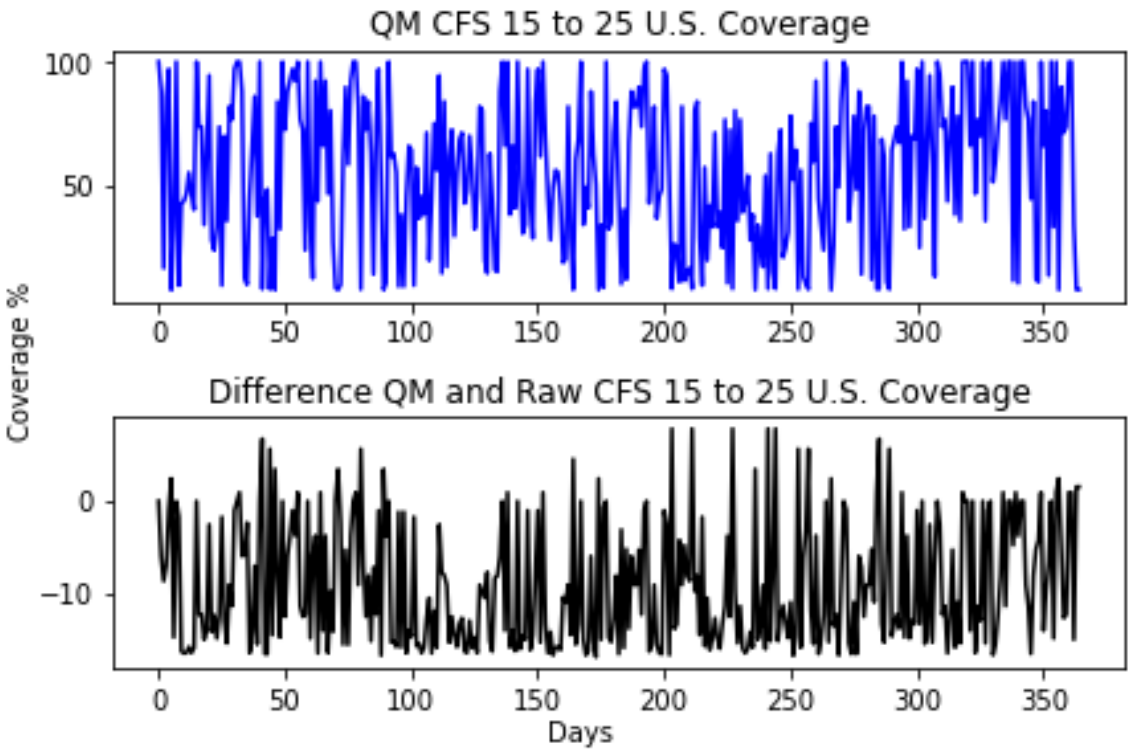

Figure 51. QM correction of 15-to-25 day forecasts CFSv2 coverage for the U.S. utilizing PERSIANN-CDR as the observation corrector (top) and the difference between QM corrected minus raw coverage (bottom).

The QM coverage percentages are higher for the U.S. than the Tropics, supported by the PERSIANN-CDR 4 panel analysis (Figure 51). The difference shows how much the raw CFSv2 over forecasted the coverage of precipitation in the U.S. There are some points in the difference between the QM and CFSv2; the rainfall coverage amounts are more extensive than the PERSIANN-CDR but not significantly. The next sets of plots use the CFSv2 0-to-10 day forecasts for the QM corrections on the CFSv2 15-to-25 day forecasts.

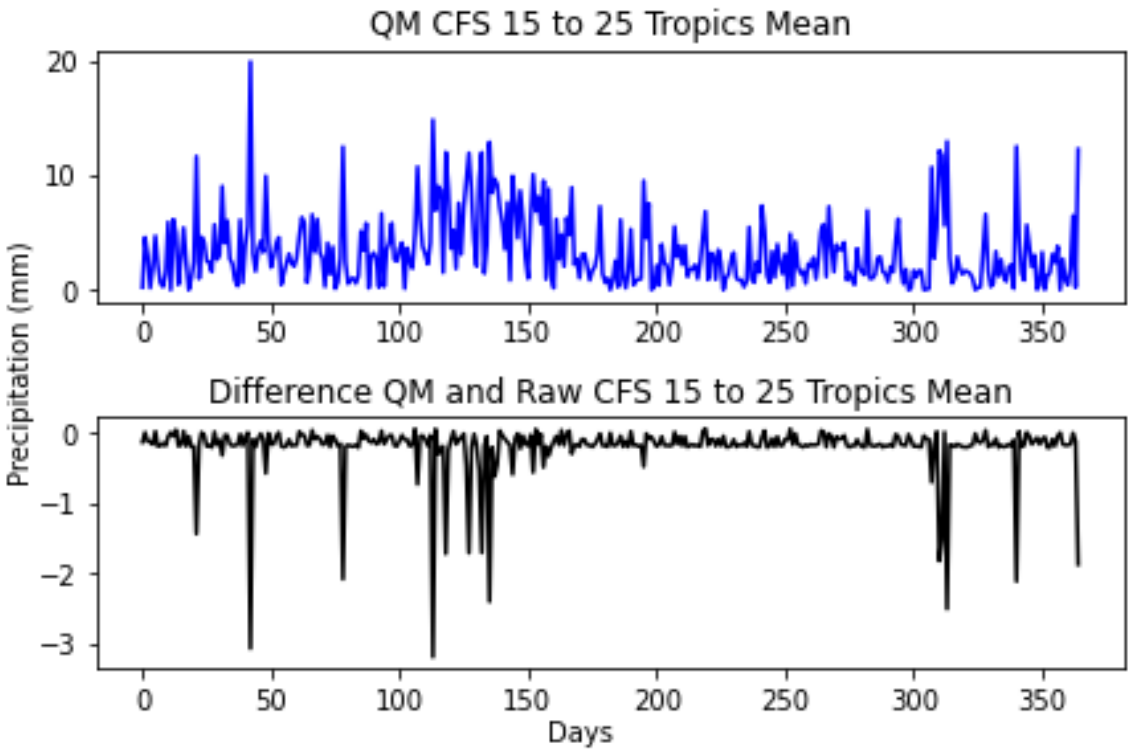

Figure 52. QM correction of 15-to-25 day forecasts CFSv2 mean for the Central Tropical Pacific utilizing CFSv2 0-to-10 day forecasts as the observation corrector (top) and the difference between QM corrected minus raw mean (bottom).

The QM which applied the CFSv2 0-to-10 day forecasts shows similar results of mean values in the Tropics from the PERSIANN-CDR with a peak of 20 mm, which is a significant amount of precipitation (Figure 52). The differences between the 0-to-10 day and the PERSIANN-CDR corrections depict the higher accuracy of the PERSIANN-CDR due to the negative peaks. The 0-to-10 day forecasts are limited with representing the significant precipitation events. The QM and raw are relatively equal throughout the year overall, with the mean precipitation values. The 0-to-10 day forecasts statistics appear to under forecast precipitation compared to the PERSIANN-CDR mean values.

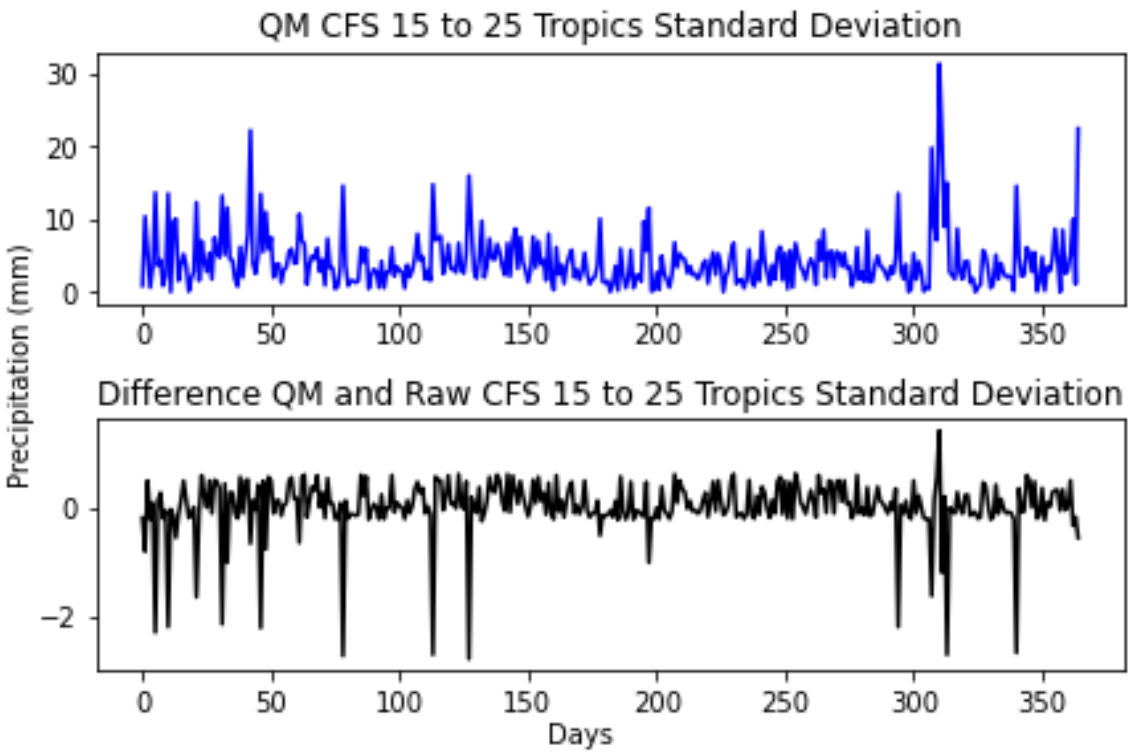

Figure 53. QM correction of 15-to-25 day forecasts CFSv2 standard deviation for the Central Tropical Pacific utilizing CFSv2 0-to-10 day forecasts as the observation corrector (top) and the difference between QM corrected minus raw standard deviation (bottom).

The standard deviation in the 0-to-10 day forecasts line up with the mean differences (Figure 53). The variability in the 0-to-10 day forecasts is higher compared to the PERSIANN-CDR QM correction, which suggests that the model has a difficult time capturing the amount of precipitation in the Tropics region. Another possibility is that the PERSIANN-CDR smoothed out the precipitation over the Tropical Pacific compared to the CFSv2. The raw data has higher variability than the QM, which indicates some improvement in corrections for the 15-to-25 day forecasts with the 0-to-10 day.

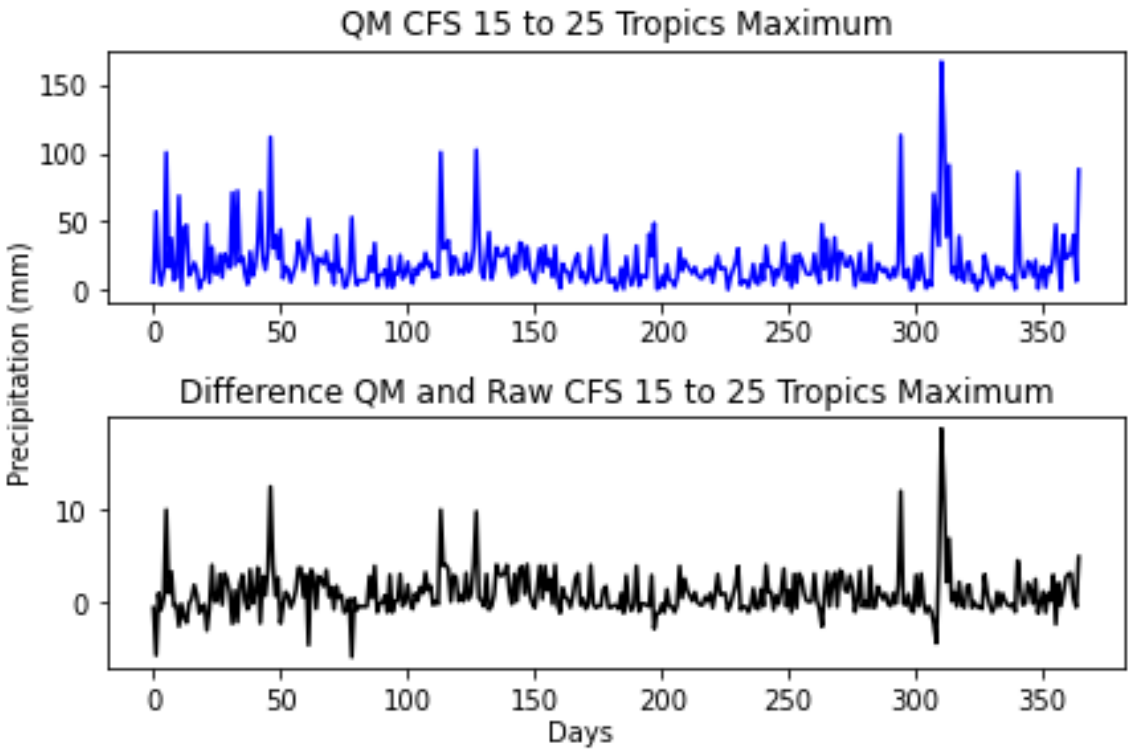

Figure 54. QM correction of 15-to-25 day forecasts CFSv2 maximum for the Central Tropical Pacific utilizing CFSv2 0-to-10 day forecasts as the observation corrector (top) and the difference between QM corrected minus raw maximum (bottom).

The 0-to-10 day forecasts maximum values are under forecasted with the peaks larger compared to the PERSIANN-CDR (Figure 54). The corrections are much less significant and support the need for the PERSIANN-CDR maximum values for the Tropics. The average maximum value appeared to remain around 25 mm for the QM with a difference of 5 mm above the raw CFSv2 data in the Tropical Pacific.

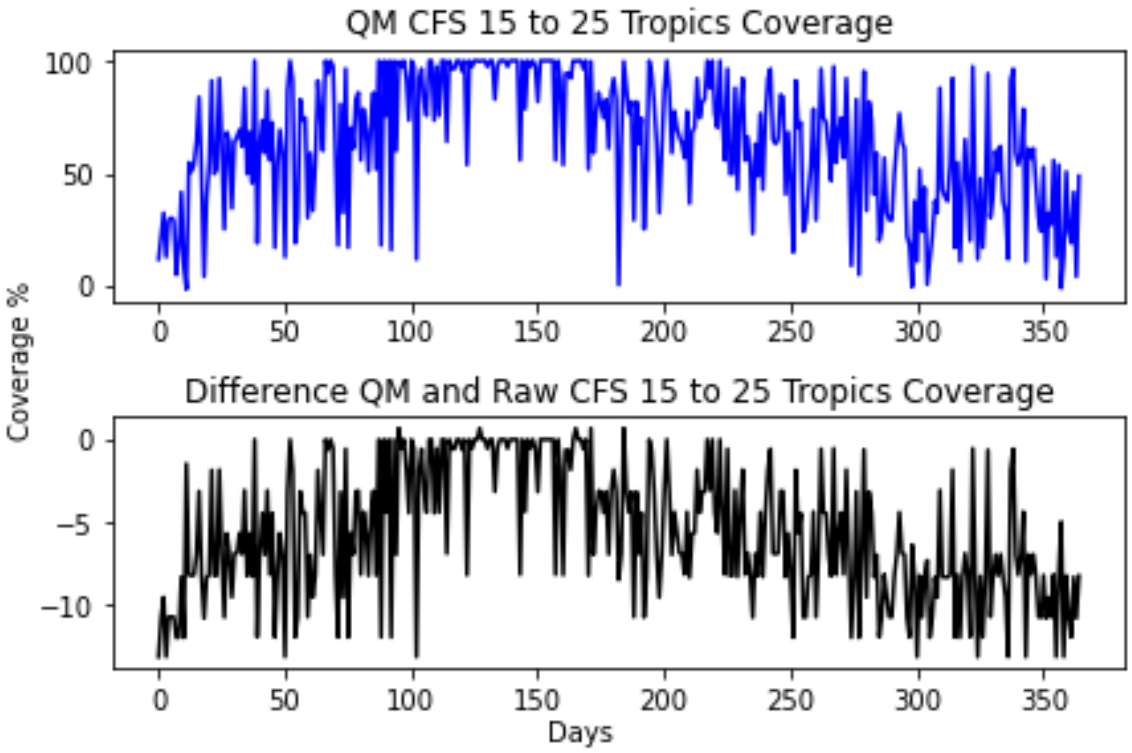

Figure 55. QM correction of 15-to-25 day forecasts CFSv2 coverage for the Central Tropical Pacific utilizing CFSv2 0-to-10 day forecasts as the observation corrector (top) and the difference between QM corrected minus raw coverage (bottom).

The coverage percentages for the 0-to-10 day forecasts corrections are similar to the PERSIANN-CDR (Figure 55). Compared to the PERSIANN-CDR, there are no days that the CFSv2 0-to-10 day forecasts has coverage values larger than the raw CFSv2 15-to-25 day forecasts. These observations are also noticeable in the four-panel analysis plots between the PERSIANN-CDR and the CFSv2 0-to-10 day forecast days. The Tropics have more improvement from the PERSIANN-CDR corrections for all of the statistical precipitation parameters.

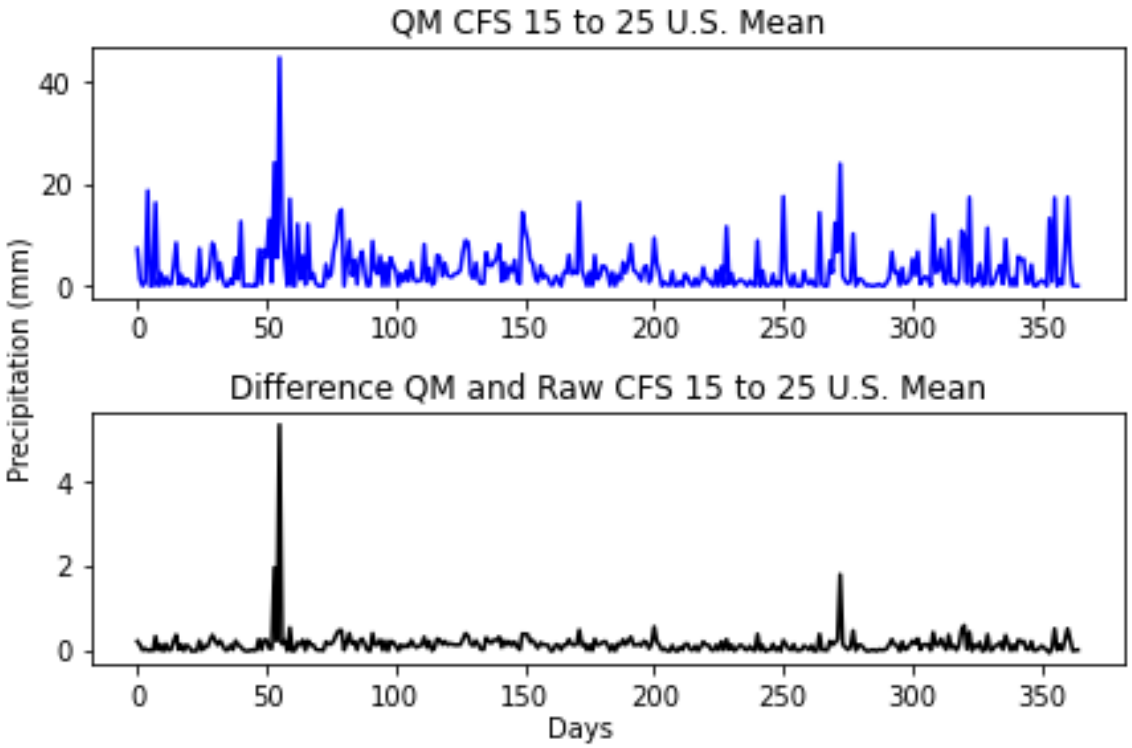

Figure 56. QM correction of 15-to-25 day forecasts CFSv2 mean for the U.S. utilizing CFSv2 0-to-10 day forecasts as the observation corrector (top) and the difference between QM corrected minus raw mean (bottom).

In comparison to the tropic's location the U.S. CFSv2 0-to-10 day forecasts corrections for the mean are almost equivalent (Figure 56). The only difference noted is that the CFSv2 0-to-10 day forecasts slightly under predicted the average amount of precipitation in the U.S. There is little change with the mean statistical parameter between the CFSv2 and the PERSIANN-CDR which indicates the need to identify any other possible discrepancies with the standard deviation, maximum, and coverage.

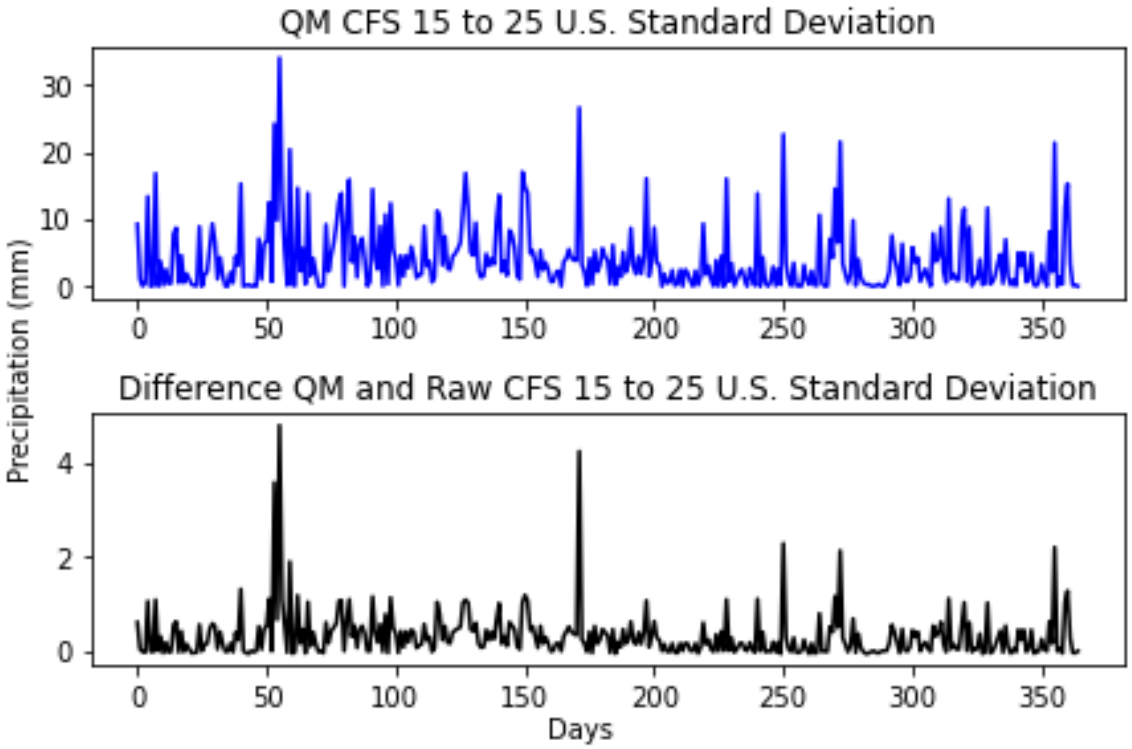

Figure 57. QM correction of 15-to-25 day forecasts CFSv2 standard deviation for the U.S. utilizing CFSv2 0-to-10 day forecasts as the observation corrector (top) and the difference between QM corrected minus raw standard deviation (bottom).

The variability of precipitation in the U.S. is slightly higher than the Tropics with the CFSv2 0-to-10 day forecasts corrections applied with a peak value of 34 mm which is larger than the PERSIANN-CDR (Figure 57). The CFSv2 0-to-10 day forecasts has higher variability in precipitation amounts for the U.S. The differences show that the QM has higher variation in precipitation amounts than the raw CFSv2 data, which mean that the local precipitation patterns are impacted in the longer-range forecasts of the CFSv2. The lower standard deviation values of the PERSIANN-CDR in the U.S. determine higher accuracy compared to the CFSv2 0-to-10 day forecasts.

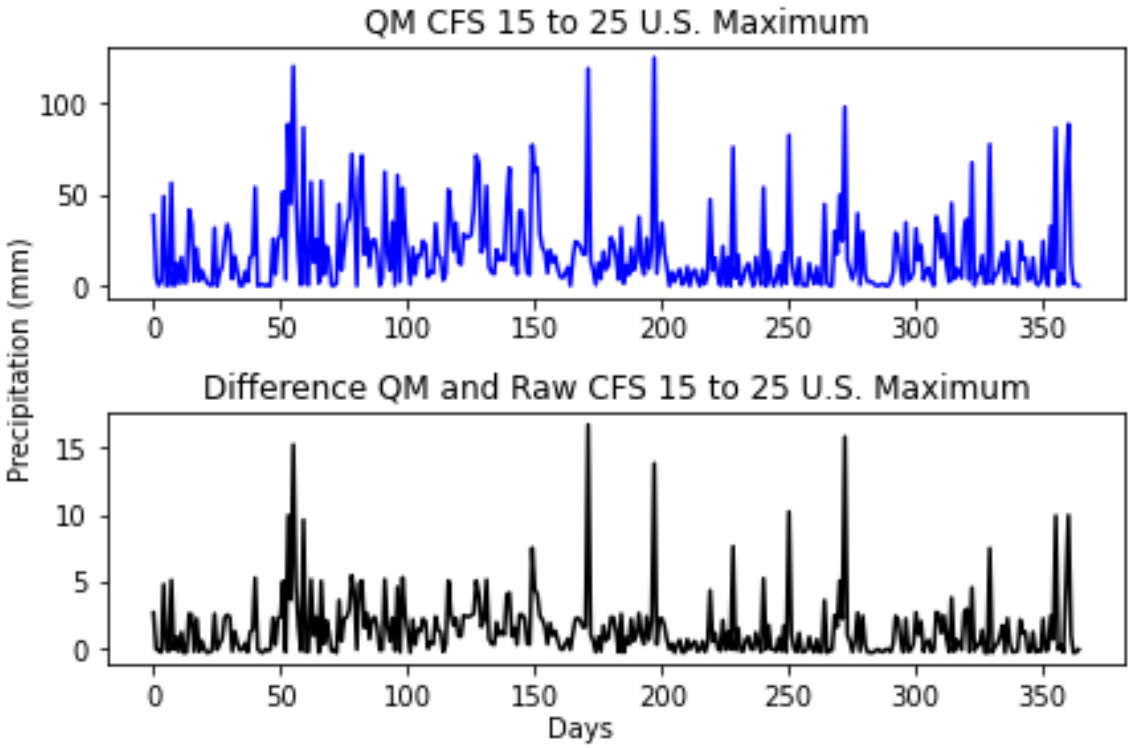

Figure 58. QM correction of 15-to-25 day forecasts CFSv2 maximum for the U.S. utilizing CFSv2 0-to-10 day forecasts as the observation corrector (top) and the difference between QM corrected minus raw maximum (bottom).

The maximum values has been under forecasted in the U.S. with the CFSv2 raw 15-to-25 day forecasts (Figure 58). The CFSv2 0-to-10 day forecasts are supported by the four-plot analysis of maximum values shown globally, which has the 0-to-10 day maximum values higher than the 15-to-25 day forecasts. The PERSIANN-CDR QM compared to the CFSv2 0-to-10 day QM are similar in the results of maximum precipitation values over the U.S. with the only difference of under forecasted peak amounts.

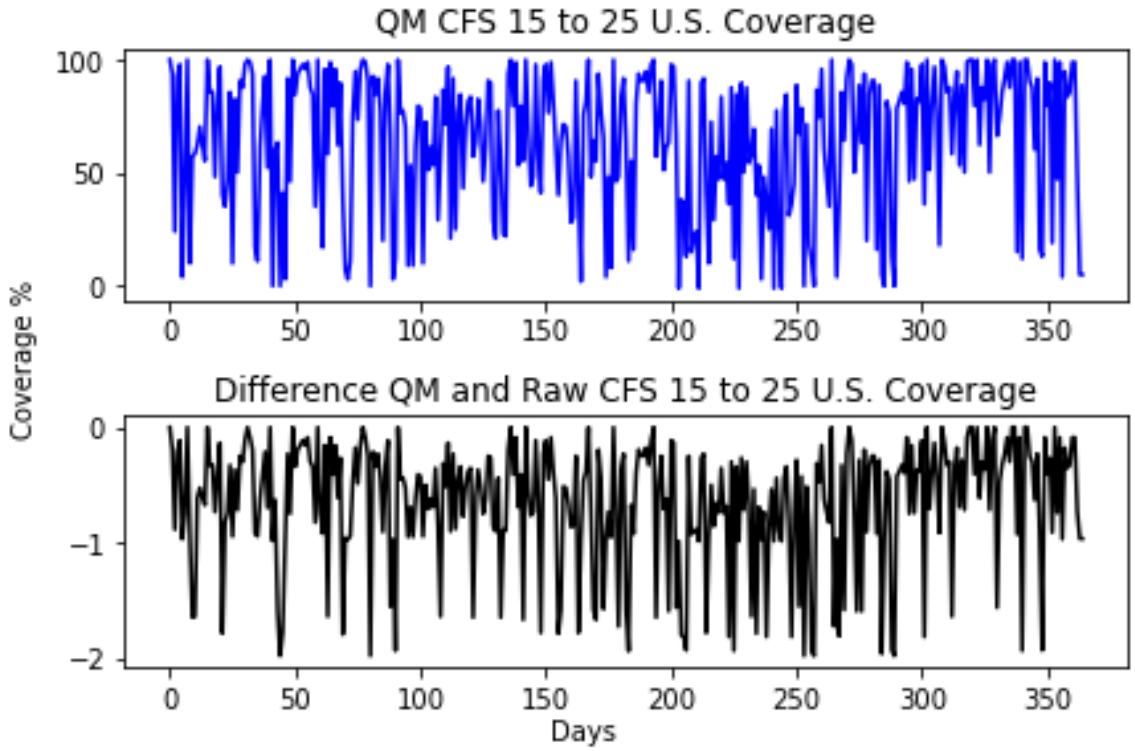

Figure 59. QM correction of 15-to-25 day forecasts CFSv2 coverage for the U.S. utilizing CFSv2 0-to-10 day forecasts as the observation corrector (top) and the difference between QM corrected minus raw coverage (bottom).

The final statistical parameter analyzed for the QM is the coverage percentage in the U.S. with the CFSv2 0-to-10 day forecasts (Figure 59). The coverage percentage from CFSv2 0-to-10 day forecasts shows higher fluctuations in value compared to the Tropics but, similar to the PERSIANN-CDR for the U.S. The difference plot shows that the raw CFSv2 data over forecasted precipitation coverage values for the entire year in comparison to the QM for the 0-to-10 day forecasts. The CFSv2 0-to-10 day forecasts has more similar values compared to the 15-to-25 day forecasts which supports that the PERSIANN-CDR provides more accurate corrections for the coverage of precipitation. The two specific geographic locations gave a more precise and accurate analysis of the CFSv2 data's QM correction limitations compared to the PERSIANN-CDR.

V. Conclusions and Future Research

Chapter Overview

The 15-to-25 day forecasts from the CFSv2 show the most promising results for future improvement in precipitation forecasts. Forecasting precipitation beyond 30 days show little skill meaning QM will have no appreciable effect.

Summary

The main conclusions about this analysis and research are that the PERSIANN-CDR statistics highlight the accuracy of the CFSv2 in the 0-to-10 day range. On the other hand, once the forecast days grow beyond 30 days, the correlation and accuracy values are much too low for promising results to apply the QM technique. The QM technique is most useful in the 15-to-25 day range due to the model having some skill at this range and the fact that the model solution shows drift and bias from the observed statistics. The 0-to-10 day forecasts does not need any correction applied since the accuracy is relatively high for most of the globe besides the Tropics regions. The QM from the PERSIANN-CDR statistics provides a more accurate representation of global precipitation values compared to the CFSv2 0-to-10 day forecasts. The accuracy percentages of hits and misses with the CFSv2 does not worsen significantly beyond the 15-to-25 day forecast range. The contingency tables and correlation plots for the different forecast time frames helped support the use of the 15-to-25 day range for the QM corrections.

Future research can implement the findings here in conjunction with precipitation impacts from the El Nino Southern Oscillation (ENSO) and the Madden Julian Oscillation (MJO). The Oceanic Nino Index (ONI) can then be applied to analyze years

above and below normal Sea Surface Temperatures (SSTs) present in the Pacific and compare annual scales of global precipitation patterns. The MJO can then be used to have all 8 phases broken down into four, so there will be enough useable days for comparison. Another useful direction to apply this data is by clustering dependent on precipitation distribution. K-Means clustering would allow the distributions to be separated by statistical parameters such as the mean into individual bins.

There are many different directions this research can head with the foundational understanding from the statistical values and improvements from the QM. The K-Means clustering will further break down the data to specifically highlight areas of the higher reliability of forecasts than lower from significant precipitation events for different parts of the globe. Another way to analyze the CFSv2 data is to break the QM down into the individual seasons with enough data available combined with the teleconnection patterns to identify peaks throughout the year and cut down on the seasonal signals from large scale meteorological events. Applying the QM to a larger time series dataset will also help compare different years of El Nino or La Nina events and the model's accuracy to handle precipitation pattern changes across the Tropical Pacific. Implementing another satellite data set, such as the Global Precipitation Climatology Project (GPCP), may help gather more statistical analysis combined with the PERSIANN-CDR. The PERSIANN-CCS can provide detailed information on assigning precipitation values to various cloud types at greater spatial and temporal resolution.

Bibliography

Ashouri, H., Hsu, K. L., Sorooshian, S., Braithwaite, D. K., Knapp, K. R., Cecil, L. D., Nelson, B. R., & Prat, O. P. (2015). PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies. *Bulletin of the American Meteorological Society*, *96*(1), 69–83.

Benesty, J., Chen, J., & Huang, Y. (2008). On the importance of the pearson correlation coefficient in noise reduction. *IEEE Transactions on Audio, Speech and Language Processing*, *16*(4), 757–765.

Cannon, A. J., Sobie, S. R., & Murdock, T. Q. (2015). *Bias Correction of GCM Precipitation by Quantile Mapping: How Well Do Methods Preserve Changes in Quantiles and Extremes?*

Gangrade, S., Morales-Hernandez, M., Tavakoly, A. A., Arsenault, K. R., Wegiel, J., McCormack, K., Wahl, M., Kumar, S. V., Peters-Lidard, C. D., Kao, S.-C., Evans, K. J., Gangrade, S., Morales-Hernandez, M., Tavakoly, A. A., Arsenault, K. R., Wegiel, J., McCormack, K., Wahl, M., Kumar, S. V., Evans, K. J. (2020). Towards the Development of a High-resolution, Global Streamflow and Flood Forecasting System - An U.S. Interagency Collaboration Effort.

Guo, H., Bao, A., Liu, T., Chen, S., & Ndayisaba, F. (2016). Evaluation of PERSIANN-CDR for Meteorological Drought Monitoring over China. *Remote Sensing*, *8*(5), 379.

Heo, J.-H., Ahn, H., Shin, J.-Y., Kjeldsen, T. R., & Jeong, C. (2019). *Probability Distributions for a Quantile Mapping Technique for a Bias Correction of Precipitation Data: A Case Study to Precipitation Data Under Climate Change*.

Kadioglu, M. (2000). Regional variability of seasonal precipitation over Turkey. *International Journal of Climatology*, *20*(14), 1743–1760.

Lin, Y., & Mitchell, K. E. (2005). *The NCEP Stage II/IV Hourly Precipitation Analyses: Development and Applications*. In Proceedings of the 19th AMS Conference on Hydrology, San Diego, CA (USA), 5–14.

Miao, C., Ashouri, H., Hsu, K. L., Sorooshian, S., & Duan, Q. (2015). Evaluation of the PERSIANN-CDR daily rainfall estimates in capturing the behavior of extreme precipitation events over China. *Journal of Hydrometeorology*, *16*(3), 1387–1396.

Nguyen, P., Ombadi, M., Sorooshian, S., Hsu, K., AghaKouchak, A., Braithwaite, D., Ashouri, H., & Rose Thorstensen, A. (2018). The PERSIANN family of global satellite precipitation data: A review and evaluation of products. *Hydrology and Earth System Sciences*, *22*(11), 5801–5816.

Rajczak, J., Kotlarski, S., & Schär, C. (2016). Does quantile mapping of simulated precipitation correct for biases in transition probabilities and spell lengths? *Journal of Climate*, *29*(5), 1605–1615.

Saha, S., Moorthi, S., Pan, H. L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Gayno, G., Wang, J., Hou, Y. T., Chuang, H. Y., Juang, H. M. H., Sela, J., & Goldberg, M. (2010). The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society*, *91*(8), 1015–1057.

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y.-T., Chuang, H., Iredell, M., Ek, M., Meng, J., Yang, R., Mendez, M. P., van den

Dool, H., Zhang, Q., Wang, W., Chen, M., & Becker, E. (2014). The NCEP Climate Forecast System Version 2. *Journal of Climate*, *27*(6), 2185–2208.

Schafer, R. W. (2011). What is a savitzky-golay filter? *IEEE Signal Processing Magazine*, *28*(4), 111–117. https://doi.org/10.1109/MSP.2011.941097

Sorooshian, S., Hsu, K. L., Gao, X., Gupta, H. V., Imam, B., & Braithwaite, D. (2000). Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. *Bulletin of the American Meteorological Society*, *81*(9), 2035–2046.

Trenberth, K. E., Dai, A., Rasmussen, R. M., & Parsons, D. B. (2003). The changing character of precipitation. In *Bulletin of the American Meteorological Society* (Vol. 84, Issue 9, pp. 1205-1217). American Meteorological Society.

Trinh-Tuan, L., Matsumoto, J., Tangang, F.T., Juneng, L., Cruz, F., Narisma, G., Jerasorn, S., Phan-Van, T., Gunawan, D., Aldrian, E., & Ngo-Duc, T. (2018). Application of Quantile Mapping Bias Correction for Mid-future Precipitation Projections over Vietnam. *The Meteorological Society of Japan.*

Wang, L., & Chen, W. (2013). Equiratio cumulative distribution function matching as an improvement to the equidistant approach in bias correction of precipitation. *Atmospheric Science Letters*, *15*(1), 1–6.

Wilks, D. S. (1993). Comparison of three-parameter probability distributions for representing annual extreme and partial duration precipitation series. *Water Resources Research*, *29*(10), 3543–3549.

Yuan, X., Wood, E. F., Luo, L., & Pan, M. (2011). A first look at Climate Forecast System version 2 (CFSv2) for hydrological seasonal prediction. *Geophysical Research Letters*, *38*(13).

Yuan, X., Wood, E. F., Roundy, J. K., & Pan, M. (2013). CFSv2-Based seasonal hydroclimatic forecasts over the conterminous United States. *Journal of Climate*, *26*(13), 4828–4847.

Zhang, M., & Scofield, R. A. (1994). Artificial neural network techniques for estimating heavy convective rainfall and recognizing cloud mergers from satellite data. *International Journal of Remote Sensing*, *15*(16), 3241–3261.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From – To)*
25-03-2021	Master's Thesis	August 2020 – March 2021

TITLE AND SUBTITLE

Comparison of Spatial Precipitation Forecasts with a Satellite Dataset

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Siebels, Andrew, C., Captain, USAF

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S)
Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 Hobson Way
WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT-ENP-MS-21-M-136

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
14th Weather Squadron
151 Patton Ave. Room 120, Asheville NC 28801
828-271-4291 and 14WS_SAR@us.af.mil

10. SPONSOR/MONITOR'S ACRONYM(S)

AFRL/RHIQ (example)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
DISTRUBTION STATEMENT A. APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES
This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

14. ABSTRACT

The purpose of this research paper is to analyze and compare global precipitation data from the Climate Forecast System Version 2 (CFSv2) with the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN)-Climate Data Record (CDR) to improve forecast capabilities. The comparison of statistical parameters of the satellite dataset with the CFSv2 will provide a useful foundational analysis on global precipitation. The various forecast time frames will then be analyzed for accuracy, and a quantile mapping (QM) technique will be applied to correct precipitation amounts from the CFSv2. QM requires a training and test dataset of the CFSv2 with the statistics of the PERSIANN-CDR used for corrections. Finally, the forecast corrections result for the CFSv2 may be used by a broad community of users specifically for the management of flood and drought-prone areas along with the scientific modeling community.

15. SUBJECT TERMS
Atmospheric Science, quantile mapping, forecast skill, climatology

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Tournay, Robert C., Lt Col, USAF, AFIT/ENP
U	**U**	**U**	**UU**	88	**19b. TELEPHONE NUMBER** *(Include area code)* (937) 255-6565 x 4743 (Robert.Tournay@afit.edu)

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18