Comparison of CHIRPS rainfall estimates and ground station measurements over Haiti

13/06/2021

Summary

Daily rainfall data from seven in Haiti were compared with satellite data estimates from CHIRPS (https://www.chc.ucsb.edu/data/chirps). The comparisons were done on total rainfall and number of rain days on a monthly and annual scale. On an annual scale the whole year as well as just the hurricane season (June to November) were considered. A variety of metrics were used, which are described here https://www.cawcr.gov.au/projects/verification/#Methods_for_foreasts_of_continuous_variables.

The total yearly rainfall is quite well estimated on average by CHIRPS across the stations, as is shown by the low bias values (average difference). However, there is low correlation between the yearly values, showing that CHIRPS does not appear to estimate the same years with above or below average rainfall. CHIRPS appears to better estimate total rainfall within the hurricane season. Observing the monthly graphs for total rainfall and number of rain days, as well as the fitted probabilities of rainfall, we see that CHIRPS seems to fail to detect many rainfall events outside the hurricane season.

Within the hurricane season the average total rainfall is similar from CHIRPS. At this time, CHIRPS overestimates the number of rain days, as is often observed from gridded data in comparison with point data. The bias in the number of rain days in the hurricane season appears to be highly seasonally and spatially dependent. To achieve similar distributions of number of rain days as in the station data, bias correction could be applied to CHIRPS. If so, bias correction which varies seasonally and spatially could be considered.

Within the hurricane season the results give some confidence in the use of CHIRPS rainfall totals. Number of rain days from CHIRS follows a similar seasonal pattern, but generally overestimates the number of rain days in the rainy season, hence a bias correct could be considered to achieve similar distributions to the station data.

Outside the hurricane season there is less confidence in the use of rainfall totals or number of rain days as CHIRPS appears to miss many rainfall events. This could be investigated further in discussion with the CHIRPS team and meteorologists within Haiti to understand potential causes of this.

Stations

Daily data from seven stations at the locations below were kindly provided by UHM for this analysis. The corresponding 0.05 x 0.05 (~5km square) pixels closest to these points from the CHIRPS daily rainfall data were extracted to compare with the daily data records. Some

quality control was carried out on the station data. In particular, extended periods of zero rainfall stretching over multiple months were removed. Also, very large daily rainfall values (> 200mm) were removed to prevent a small number of values having a large effect on the results. Comparison and analysis of extreme rainfall events from CHIRPS could be the topic of a further study.

station

latitude

longitude

cap_haitien

19.77

-72.20

damien

18.60

-72.28

quanaminthe

19.55

-71.71

petitionville

18.50

-72.28

jacmel

18.23

-72.53

jeremie

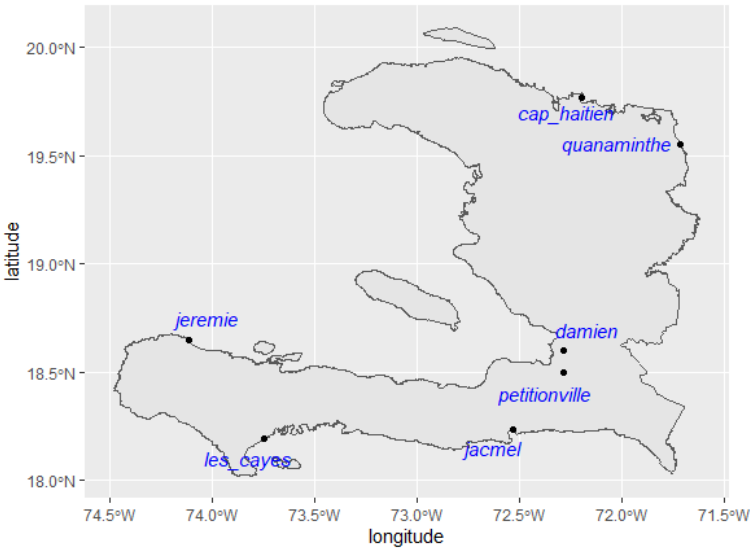
18.65

-74.12

les_cayes

18.19

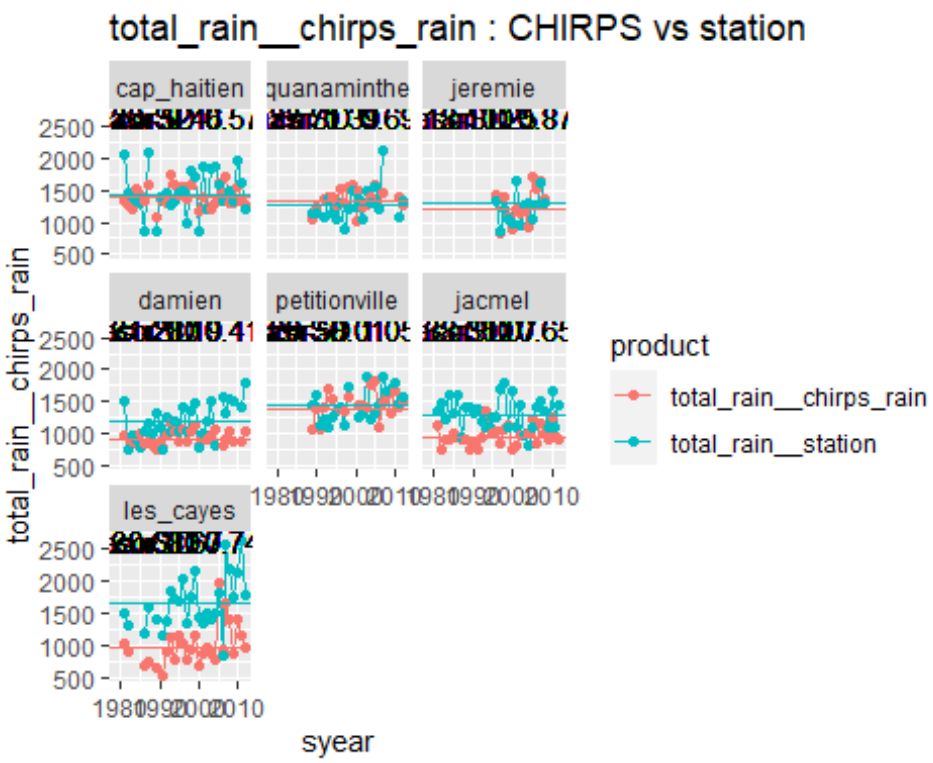
-73.75



Yearly Comparisons

Comparison Statistics for Total Yearly Rainfall

The correlation between the station and CHIRPS yearly total rainfall ranges from an (absolute) low of 0.01 for petitionville to a high of 0.67 for les cayes. We note that correlation is on the low side for most stations but slightly positive. This likely means that we would have difficulty using the CHIRPS as an estimator of total rainfall in a particular year compared to other years. The (absolute) bias is quite low for the stations: Cap Haitan, Quanaminthe, Jeremie and Petitionville, with all of them within a 100mm/year of the station value. It is highest for les cayes which has a bias of -680mm/year. This suggests in some locations CHIRPS estimates the average total rainfall reasonably well, but in some locations the bias is large as a proportion of the total rainfall. It is worth noting that Jacmel has 38 years of comparison and Jeremie has only 18 with the rest of the stations distributed between those two values, as shown in the table below.



Total yearly rainfall in mm from the station data (blue) and CHIRPS (red). The blue and red horizontal lines represent the average yearly total rainfall from the station data and CHIRPS respectively. The graphs are also annotated with comparison metrics.

n = Number of years compared

bias = average difference in mm/year, perfect value = 0

cor = Correlation coefficient, perfect value = 1

rSD = ratio of standard deviations: sd_CHIRPS/sd_station, perfect value = 1

Number of years compared

station

n

cap_haitien

36

quanaminthe

23

jeremie

18

damien

31

petitionville

28

jacmel

38

les_cayes

30

Number of years compared between the station and CHIRPS data for each station.

Comparison statistics: Total Yearly Rainfall

station

r

ME

PBIAS %

MAE

NSE

rSD

cap_haitien

0.46

-51.68

-3.6

242.63

0.18

0.57

quanaminthe

0.39

70.73

5.6

203.97

-0.01

0.69

jeremie

0.25

-103.10

-7.9

273.12

-0.44

0.87

damien

0.19

-277.52

-23.1

317.97

-1.00

0.41

petitionville

-0.01

-58.15

-4.0

285.46

-1.17

1.05

jacmel

0.07

-348.03

-26.7

380.88

-2.40

0.65

les_cayes

0.67

-676.83

-41.0

692.05

-2.15

0.74

Comparison statistics for total yearly rainfall.

r = Correlation coefficient, perfect value = 1

ME = mean error (bias) in mm/year, perfect value = 0

PBIAS = percentage bias, perfect value = 0

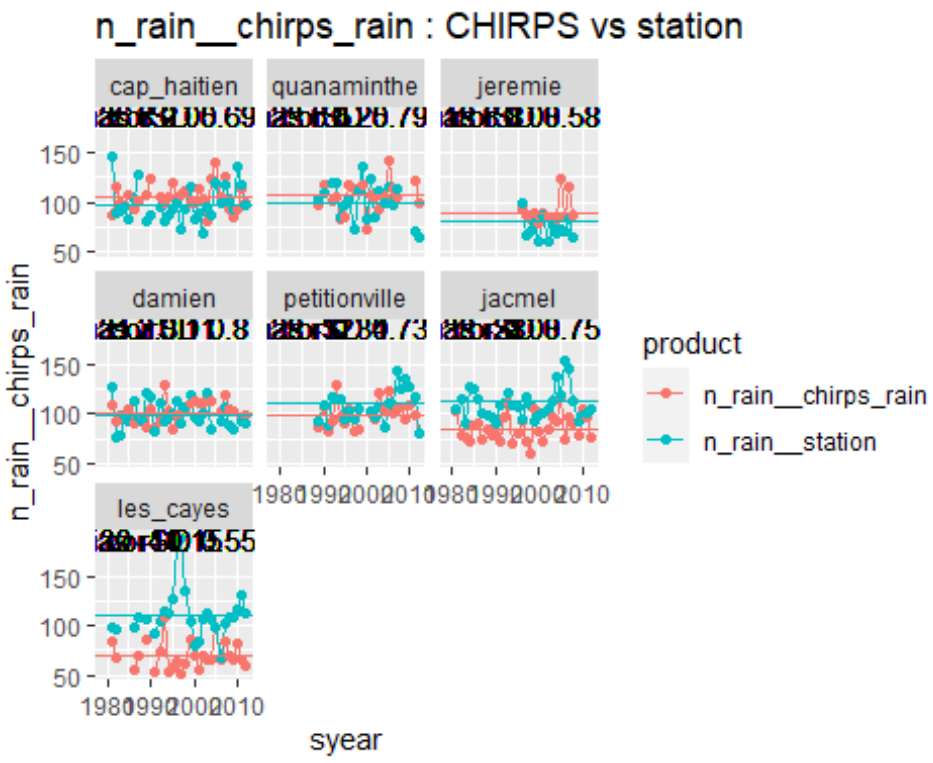
MAE = Mean absolute error in mm/year, perfect value = 0

NSE = Nash-Sutcliffe efficiency, perfect value = 1

rSD = ratio of standard deviations: sd_CHIRPS/sd_station, perfect value = 1

Comparison Statistics for Number of rainy days

Correlation on the whole is very low between CHIRPS and station values across all the stations. The bias is also quite low on average across the year. However, as will be seen on a monthly basis, this is hiding an overestimation and underestimation in different parts of the year. Hence, these results are not representative of the performance of CHIRPS.



Yearly number of rain days from the station data (blue) and CHIRPS (red). The blue and red horizontal lines represent the average yearly number of rain days from the station data and CHIRPS respectively. The graphs are also annotated with comparison metrics.

n = Number of years compared

bias = average difference in days/year, perfect value = 0

cor = Correlation coefficient, perfect value = 1

rSD = ratio of standard deviations: sd_CHIRPS/sd_station, perfect value = 1

Comparison statistics: number of rain days

station

r

ME

PBIAS %

MAE

NSE

rSD

cap_haitien

0.05

8.25

8.5

18.86

-0.62

0.69

quanaminthe

0.25

6.52

6.5

17.22

-0.35

0.79

jeremie

0.08

8.83

11.0

19.17

-0.45

0.58

damien

0.11

2.29

2.3

13.45

-0.50

0.80

petitionville

0.34

-11.57

-10.5

15.79

-0.53

0.73

jacmel

0.08

-27.92

-24.7

28.92

-3.07

0.75

les_cayes

-0.15

-40.93

-36.9

40.93

-3.32

0.55

Comparison statistics for total yearly number of rain days.

r = Correlation coefficient, perfect value = 1

ME = mean error (bias) in days/year, perfect value = 0

PBIAS = percentage bias, perfect value = 0

MAE = Mean absolute error in days/year, perfect value = 0

NSE = Nash-Sutcliffe efficiency, perfect value = 1

rSD = ratio of standard deviations: sd_CHIRPS/sd_station, perfect value = 1

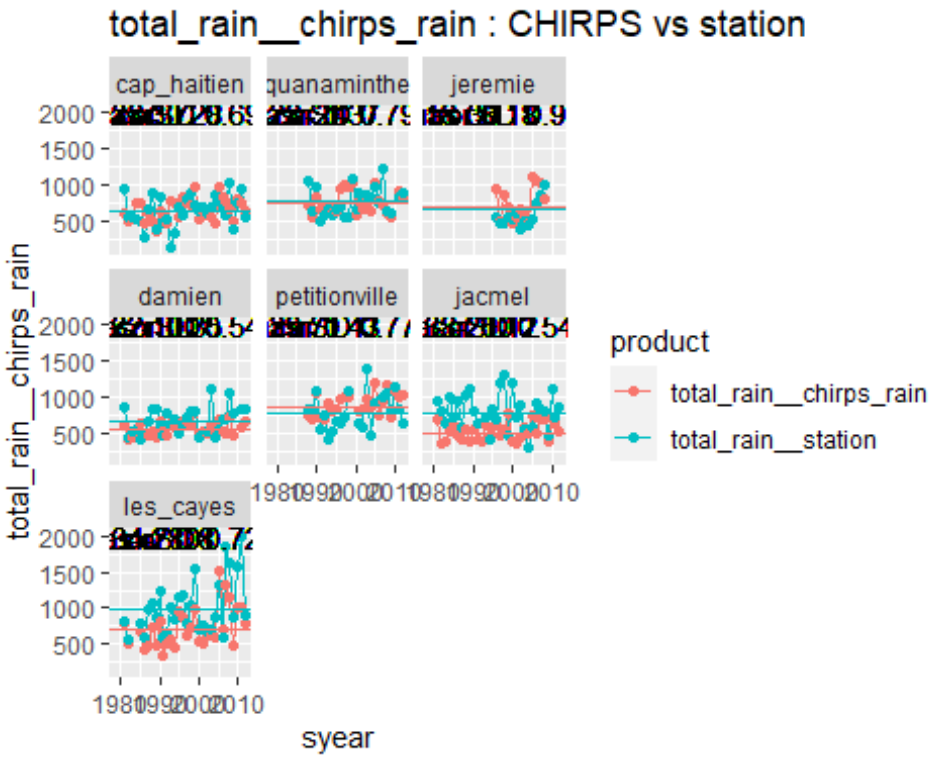
Hurricane Season Comparisons (June to November)

We also consider the data just within the hurricane season of June to November. The same graphs and statistics for total and rainfall and number of rain days were produced just for this 6 month period.

For total rainfall the (absolute) bias is lower for all stations in the hurricane season, except petitionville, compared to the full year. Hence, a large proportion of the bias in total rainfall comes outside the hurricane season.

For number of rainy days there is generally a consistent overestimation from CHIRPS at most stations, although there is location variability. Overestimation of number of rain days by a gridded data product (satellite data) compared to point based measurements (station data) is common and somewhat expected due to the areal nature of gridded data which represents the average rainfall over an area compared to at a point. There are established statistical methods to correct this overestimation if needed.

We also note that the results for Les Cayes and Jacmel are very similar to each other and have a similar latitude so perhaps looking further at the Southern coast climate could prove valuable.



Total hurricane season rainfall in mm from the station data (blue) and CHIRPS (red). The blue and red horizontal lines represent the average hurricane season total rainfall from the station data and CHIRPS respectively. The graphs are also annotated with comparison metrics.

n = Number of years compared

bias = average difference in mm/season, perfect value = 0

cor = Correlation coefficient, perfect value = 1

rSD = ratio of standard deviations: sd_CHIRPS/sd_station, perfect value = 1

Comparison statistics: Total Yearly Rainfall

station

r

ME

PBIAS %

MAE

NSE

rSD

cap_haitien

0.28

3.67

0.6

174.89

-0.09

0.69

quanaminthe

0.37

-24.18

-3.2

146.10

-0.06

0.79

jeremie

0.18

36.37

5.5

237.36

-0.51

0.90

damien

0.35

-103.95

-15.9

150.37

-0.26

0.54

petitionville

0.43

70.78

9.2

192.08

-0.04

0.77

jacmel

0.12

-264.14

-34.5

303.87

-1.41

0.54

les_cayes

0.80

-283.61

-28.7

301.69

0.05

0.72

Comparison statistics for total hurricane season rainfall.

r = Correlation coefficient, perfect value = 1

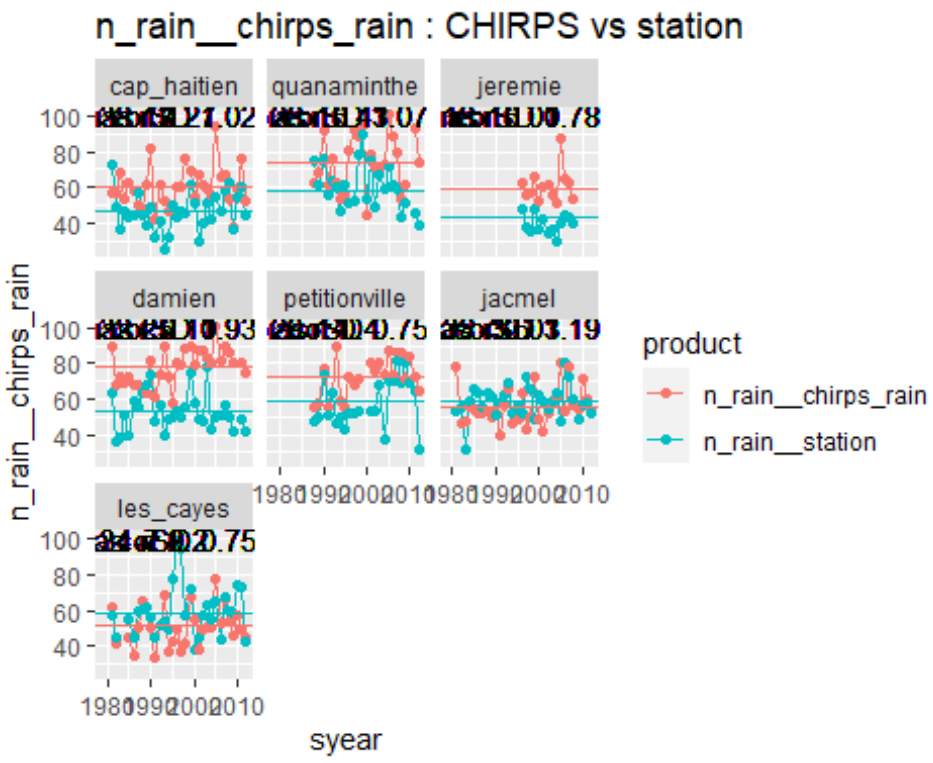
ME = mean error (bias) in mm/season, perfect value = 0

PBIAS = percentage bias, perfect value = 0

MAE = Mean absolute error in mm/season, perfect value = 0

NSE = Nash-Sutcliffe efficiency, perfect value = 1

rSD = ratio of standard deviations: sd_CHIRPS/sd_station, perfect value = 1



Hurricane season number of rain days from the station data (blue) and CHIRPS (red). The blue and red horizontal lines represent the average hurricane season number of rain days from the station data and CHIRPS respectively. The graphs are also annotated with comparison metrics.

n = Number of years compared

bias = average difference in days/year, perfect value = 0

cor = Correlation coefficient, perfect value = 1

rSD = ratio of standard deviations: sd_CHIRPS/sd_station, perfect value = 1

Comparison statistics: number of rain days

station

r

ME

PBIAS %

MAE

NSE

rSD

cap_haitien

0.27

13.42

28.9

15.26

-2.16

1.02

quanaminthe

0.43

15.11

25.9

16.96

-1.54

1.07

jeremie

0.01

15.11

34.6

18.67

-2.46

0.78

damien

0.11

25.16

47.5

25.47

-6.50

0.93

petitionville

0.40

13.72

23.4

15.59

-1.04

0.75

jacmel

0.03

-3.55

-6.0

10.82

-1.51

1.19

les_cayes

0.20

-7.18

-12.3

11.35

-0.54

0.75

Comparison statistics for total hurricane season number of rain days.

r = Correlation coefficient, perfect value = 1

ME = mean error (bias) in days/season, perfect value = 0

PBIAS = percentage bias, perfect value = 0

MAE = Mean absolute error in days/season, perfect value = 0

NSE = Nash-Sutcliffe efficiency, perfect value = 1

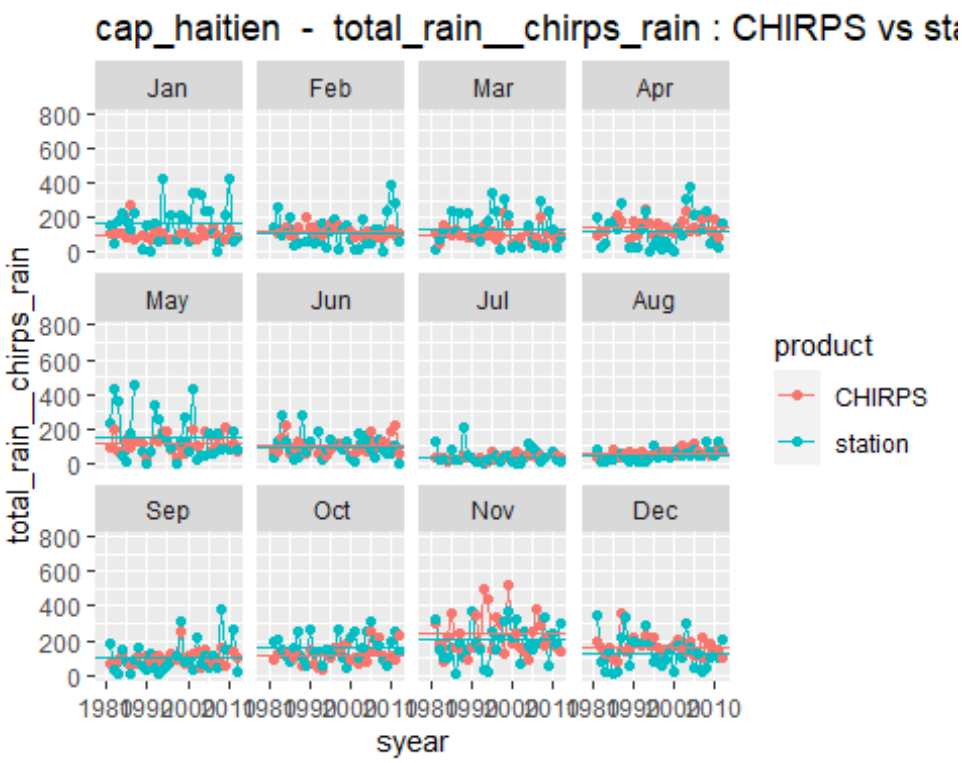
rSD = ratio of standard deviations: sd_CHIRPS/sd_station, perfect value = 1

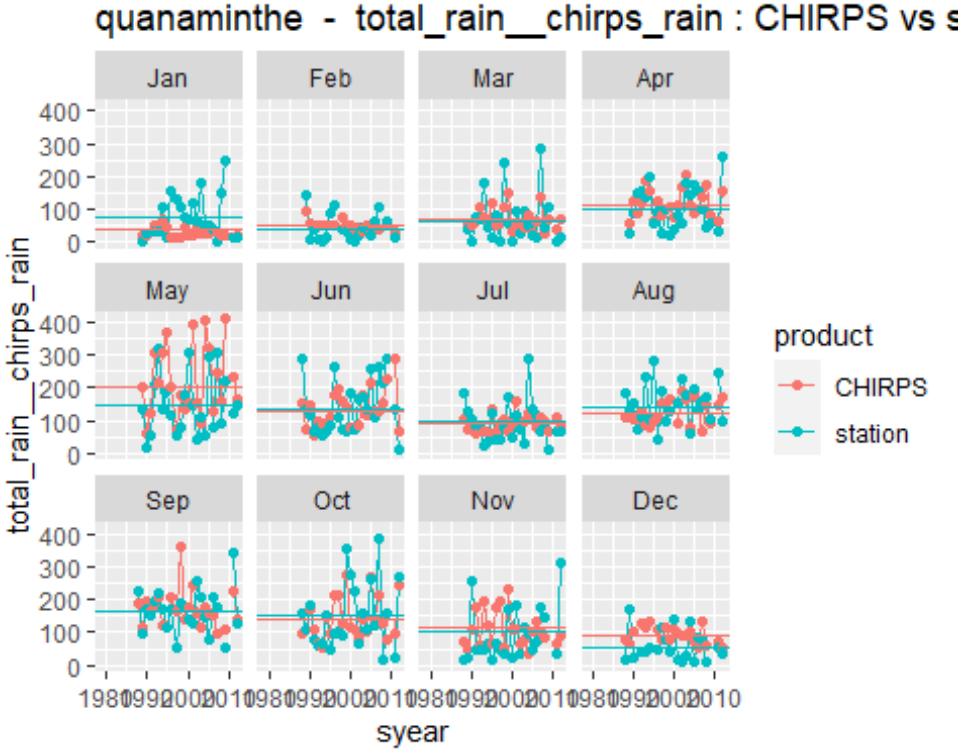
Monthly Comparisons

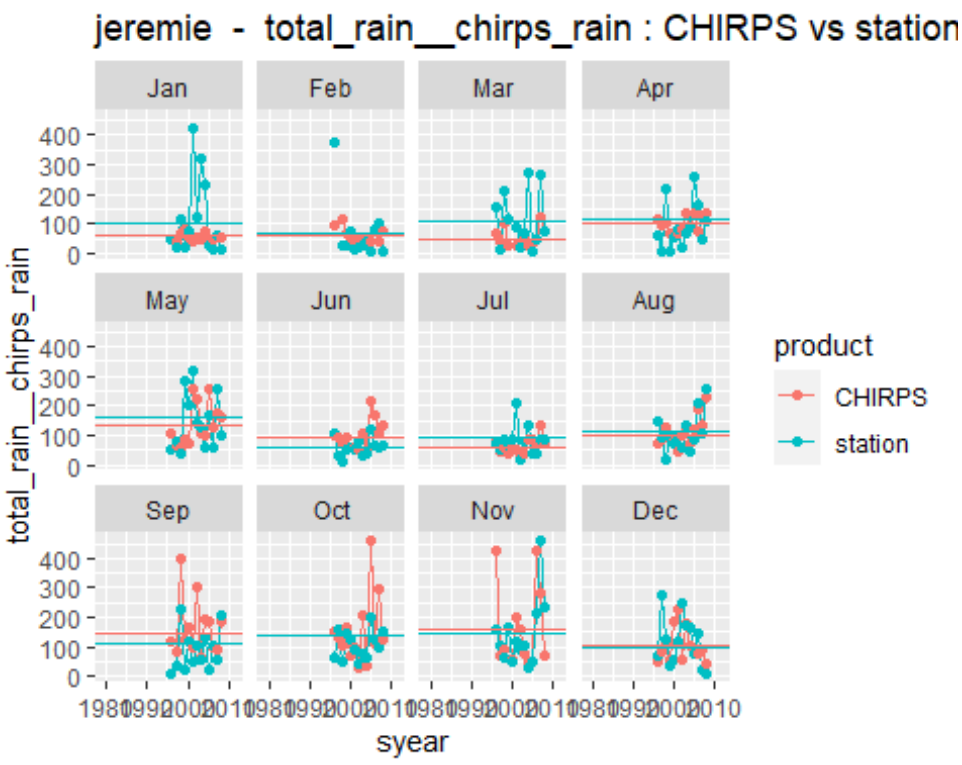
Comparison statistics for total monthly rainfall

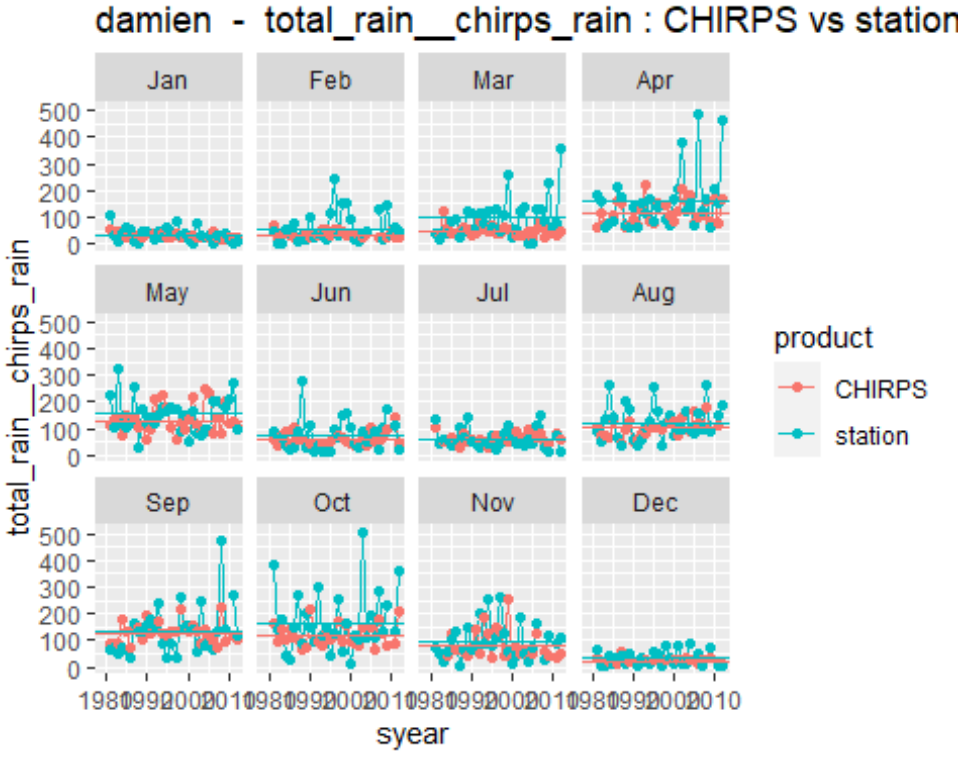
On a monthly basis, we observe some of the differences between inside and outside the hurricane season. This is particularly prominent at Les Cayes, where almost all rainfall events between January and March are missed, any underestimation in December and April.

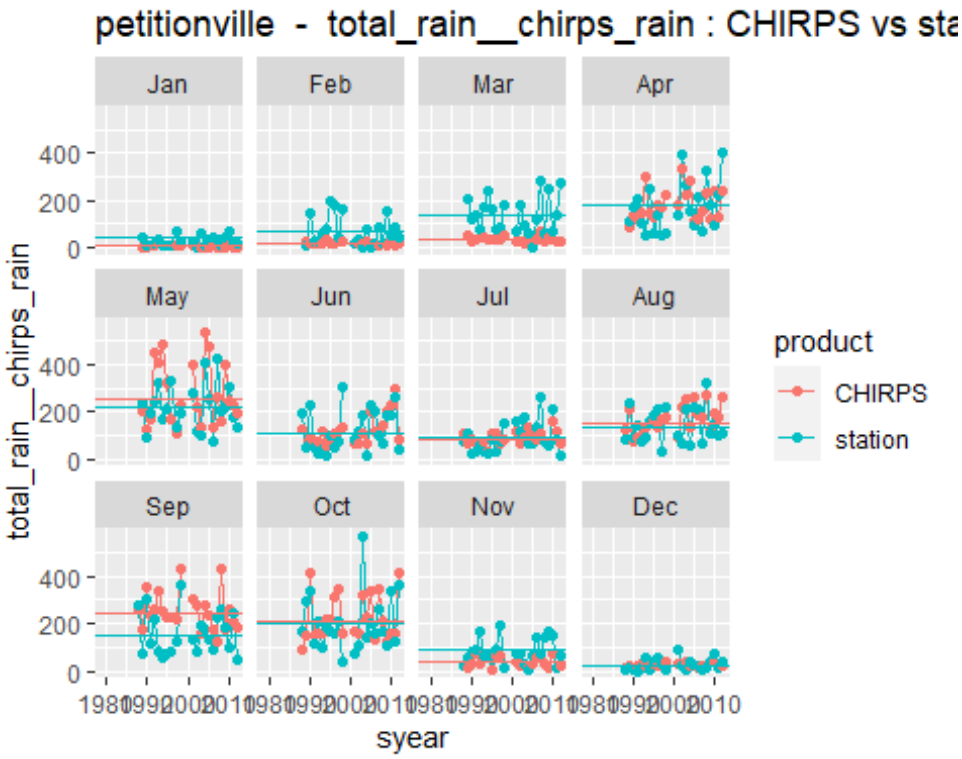
For Cap Haitian, in general the bias is quite low with peaks in January and May. Damien also has low biases except for March, April and Oct. The biases in Jacmel tend to be higher between the months of April and November whereas biases in Jeremie are on the low side, except for the month March. Les Cayes has very little rainfall between December and April so our chirps data is always overestimating, biases on the whole are quite low and there are no outliers that we can spot easily. For petitionville we have very little station rain between December and March, October 2003/2004 seems to have an outlying value of total rain. For quanaminthe, the biases are in general quite low with slightly larger ones for May. We have some outlying station values, most notably Jan 2008/2009 and Nov 2012/2013 where the station value seems very out of line with the CHIRPS estimate. A closer look at these outliers is likely needed.

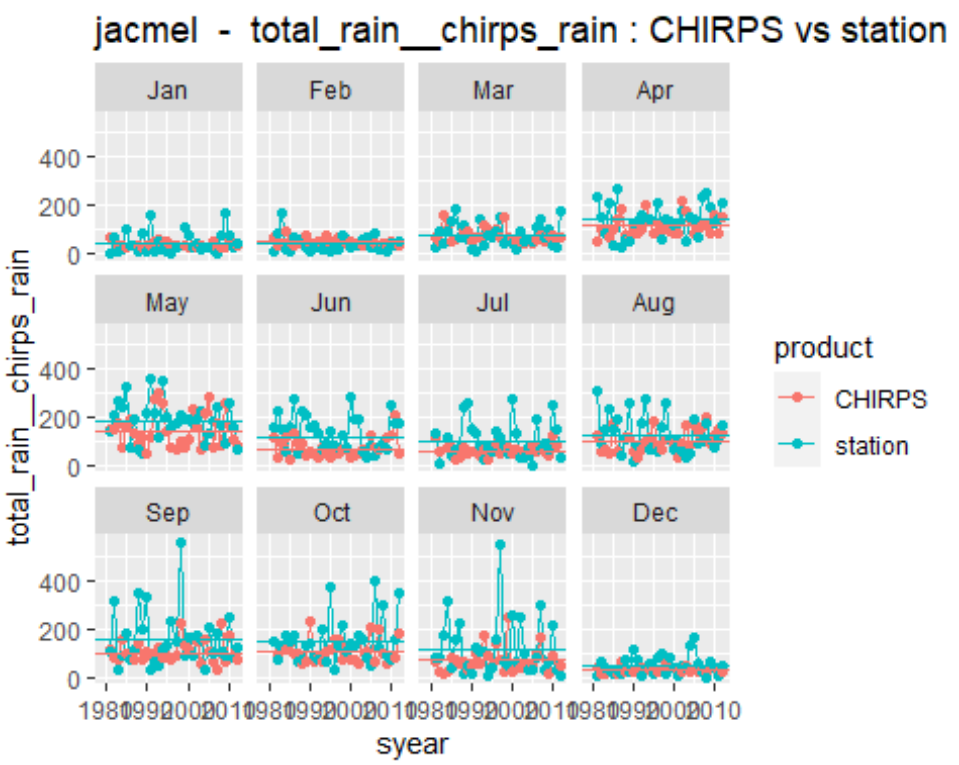

cap_haitien - total_rain__chirps_rain : CHIRPS vs sta
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
total_rain__chirps_rain
syear
product
CHIRPS
station

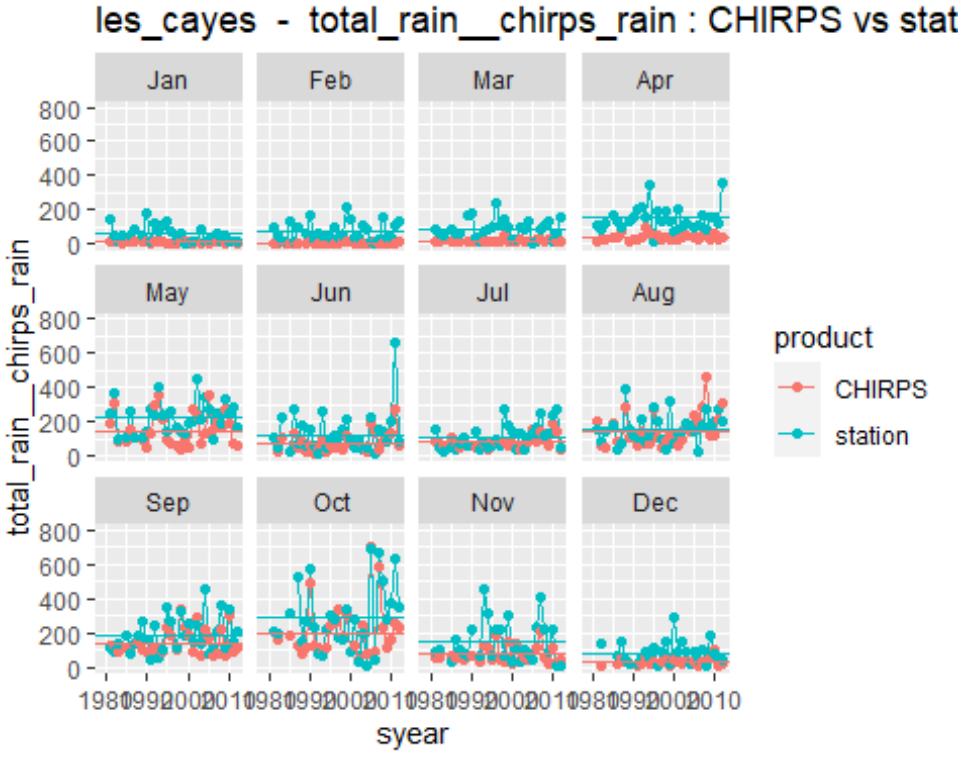

quanaminthe - total_rain__chirps_rain : CHIRPS vs s
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
total_rain__chirps_rain
syear
product
CHIRPS
station


jeremie - total_rain__chirps_rain : CHIRPS vs station
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
total_rain__chirps_rain
syear
product
CHIRPS
station


damien - total_rain__chirps_rain : CHIRPS vs station
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
total_rain__chirps_rain
syear
product
CHIRPS
station


petitionville - total_rain__chirps_rain : CHIRPS vs sta
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
400
200
0
total_rain__chirps_rain
1980
1990
2000
2010
syear
product
CHIRPS
station


jacmel - total_rain__chirps_rain : CHIRPS vs station
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
total_rain__chirps_rain
syear
product
CHIRPS
station

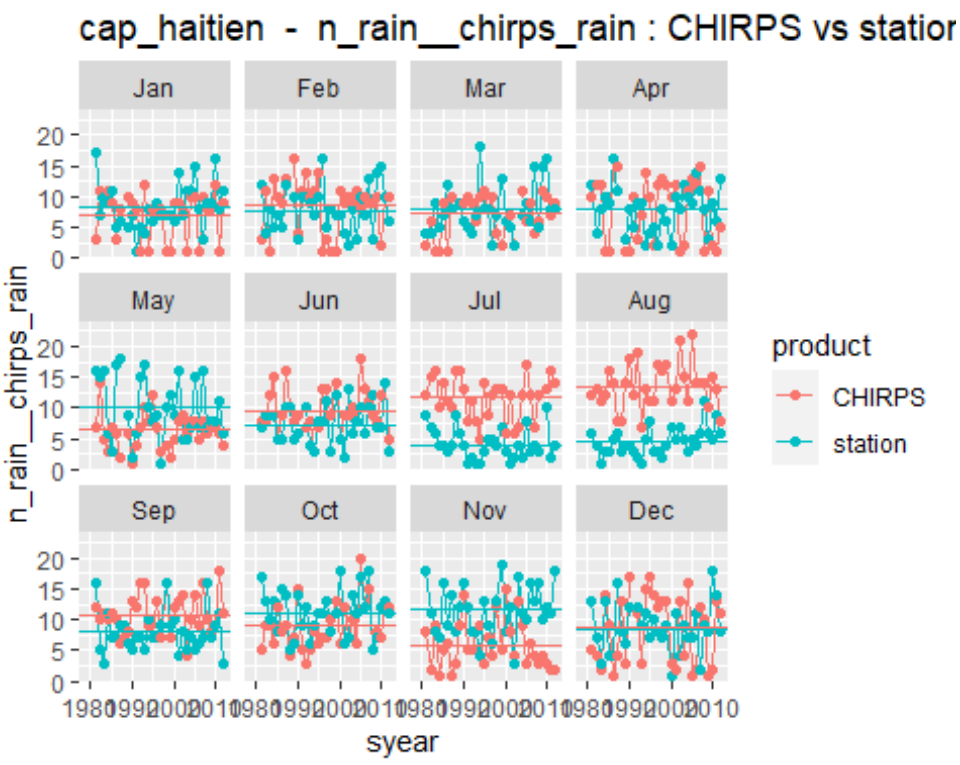

les_cayes - total_rain__chirps_rain : CHIRPS vs stat
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
total_rain__chirps_rain
syear
product
CHIRPS
station

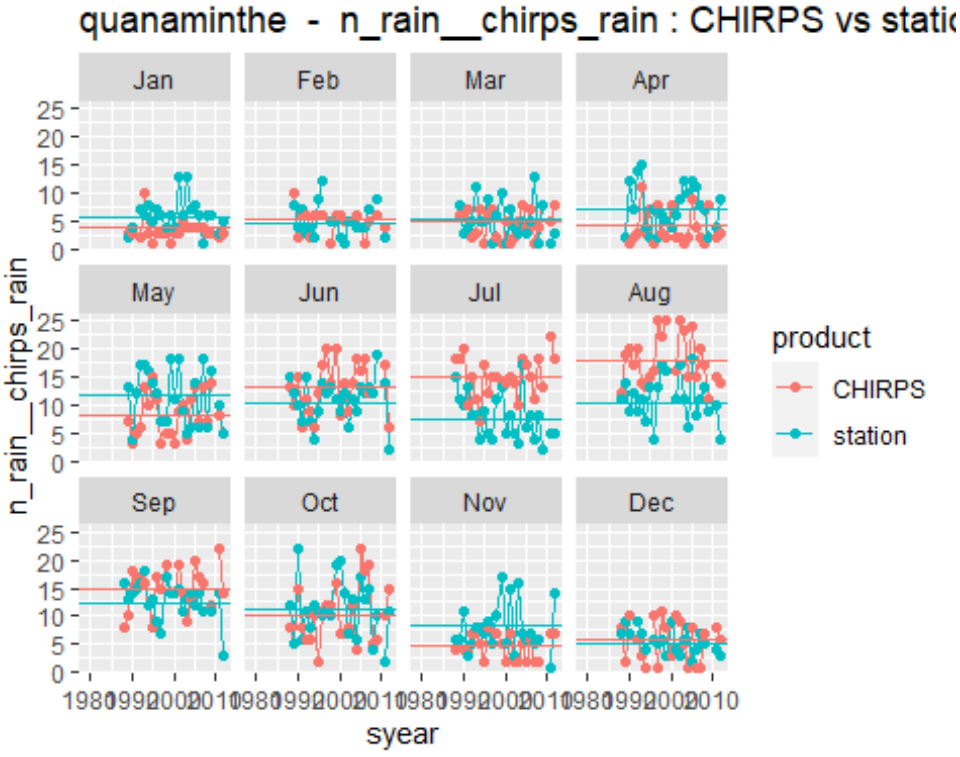
Total monthly rainfall in mm from the station data (blue) and CHIRPS (red). The blue and red horizontal lines represent the long term average monthly total rainfall from the station data and CHIRPS respectively.

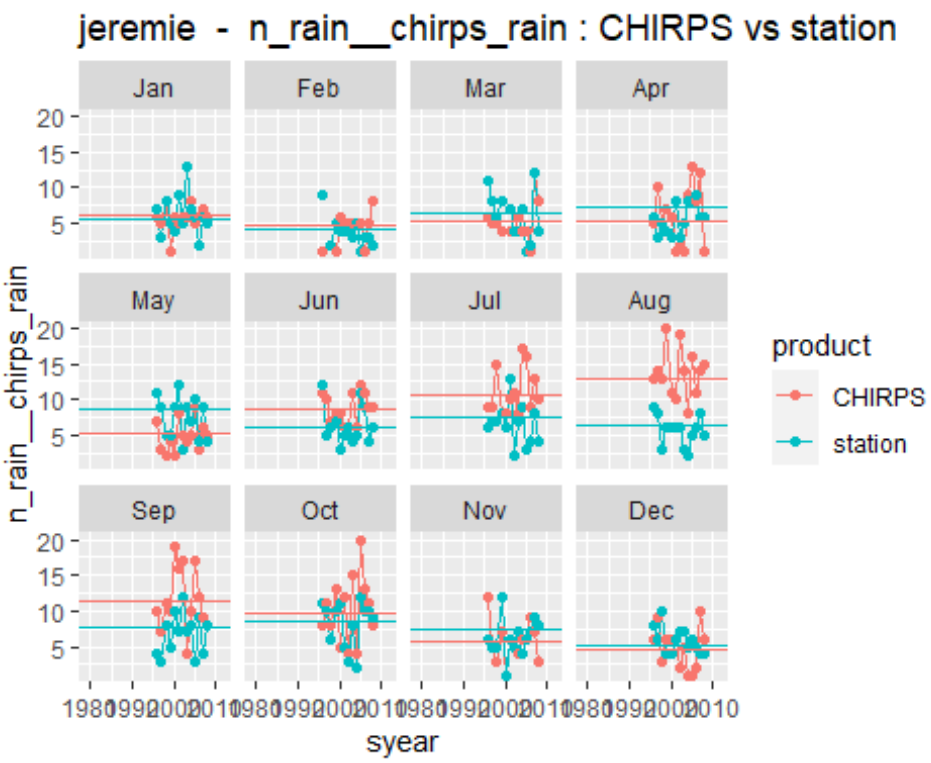
Comparison statistics for number of monthly raindays

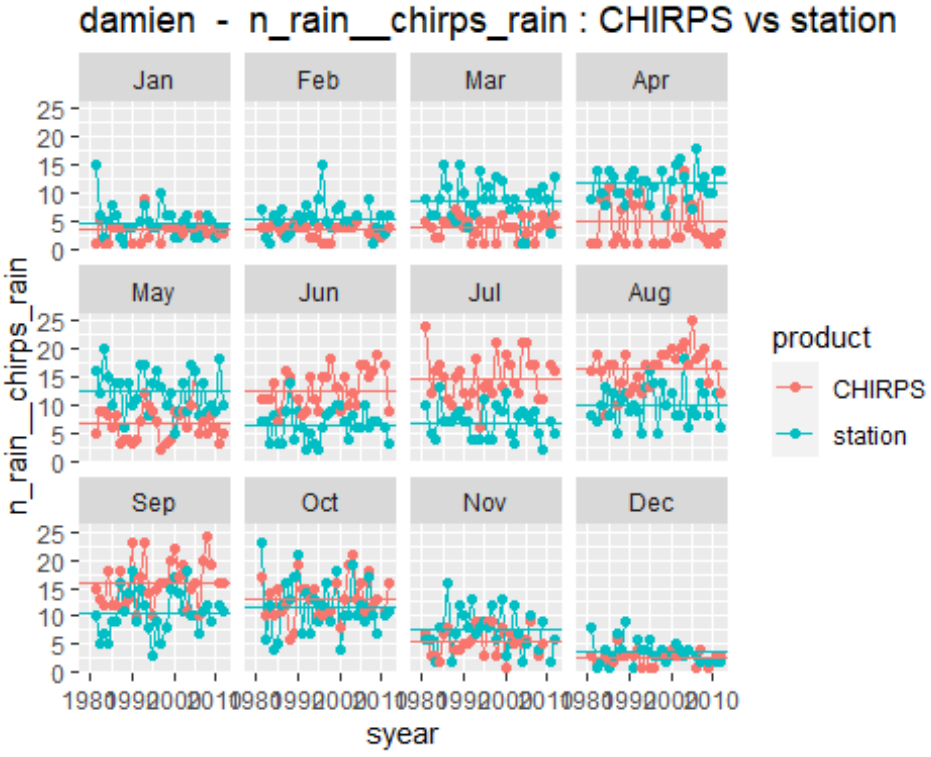
Here we again see substantial seasonal variation. A general pattern in some stations is overestimation of number of rain days in the hurricane season, and understimation outside. This gives the appearance of good performance on a yearly basis, but this is because positive and negative biases have cancelled out.

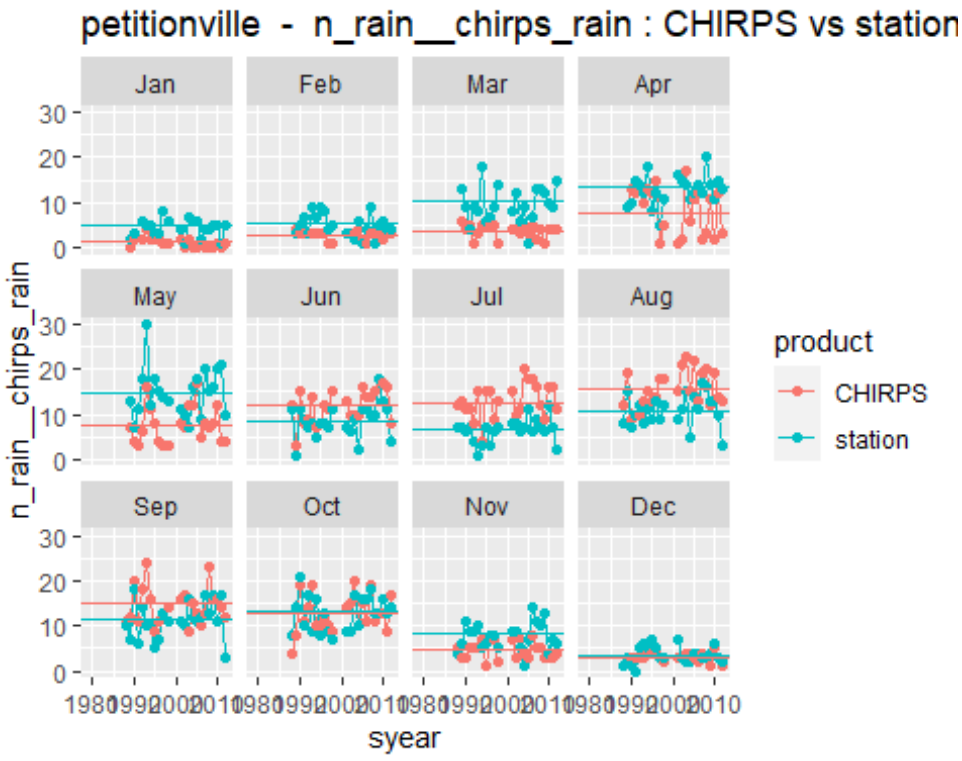
For Cap Haitian the bias is quite low between December and April, it is quite a bit larger during the remaining months and especially during July , August and November.For Damien, the bias is fairly low between October and February where it starts to increase. The months July and April have particularly high biases. For Jacmel we get that the bias is particularly high for the months March, April and May. For jeremie the bias is generally low with perhaps the exception of August, where it is a bit higher than other places.Les cayes has a small amount of (CHIRPS) rainy days between December and April and this leads to high biases in those months. The highest biases are in the months April and May. For petitionville, the biases vary considerably and none of them are clear outliers. Around May 1993 we see that it rained every day of the month for this station and that stands out a bit so perhaps further inspection of that would provide valuable insight. The biases for quanaminthe tend to be considerably higher in the months July and August.

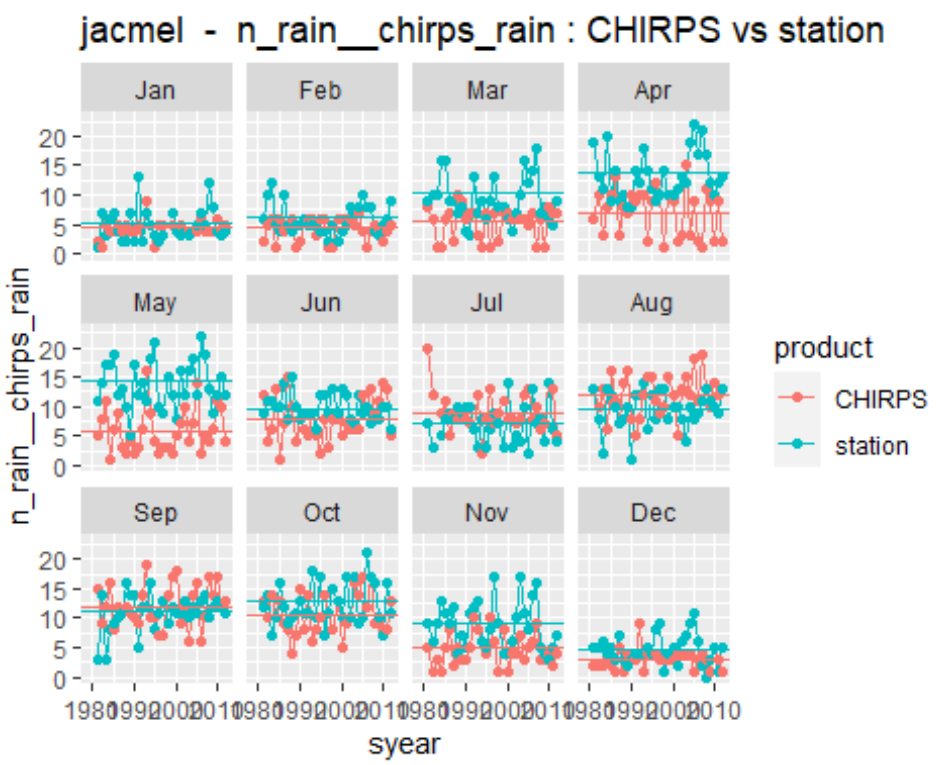

cap_haitien - n_rain__chirps_rain : CHIRPS vs statior
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
n_rain__chirps_rain
syear
product
CHIRPS
station

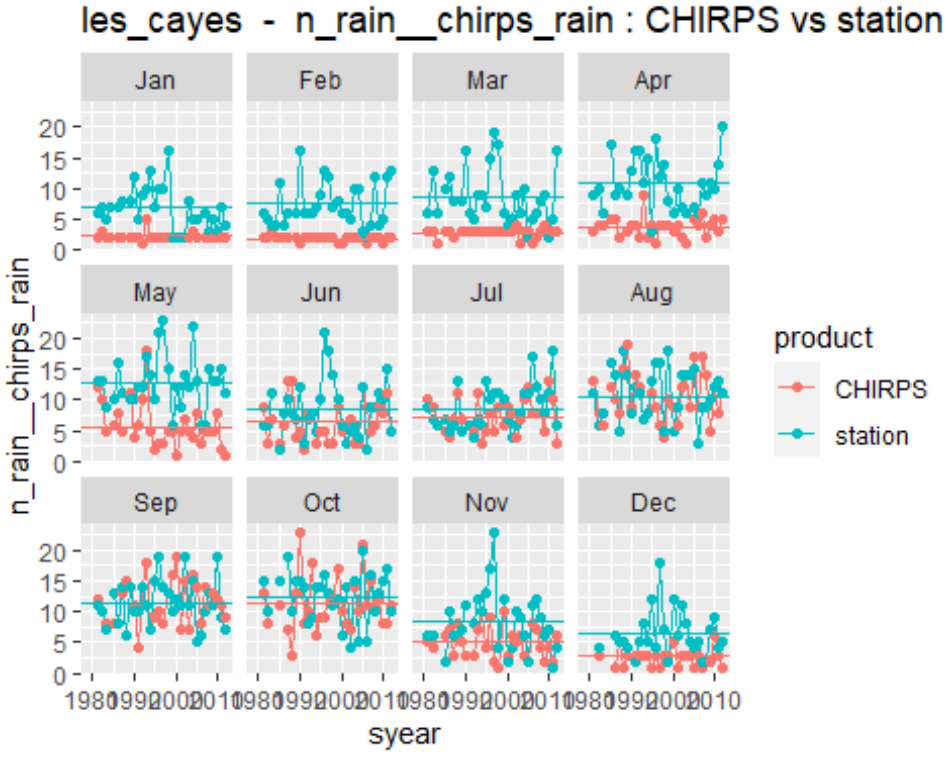

quanaminthe - n_rain__chirps_rain : CHIRPS vs static
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
n_rain__chirps_rain
syear
product
CHIRPS
station


jeremie - n_rain__chirps_rain : CHIRPS vs station
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
n_rain__chirps_rain
syear
product
CHIRPS
station


damien - n_rain__chirps_rain : CHIRPS vs station
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
n_rain__chirps_rain
syear
product
CHIRPS
station


petitionville - n_rain__chirps_rain : CHIRPS vs station
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
n_rain__chirps_rain
syear
product
CHIRPS
station


jacmel - n_rain__chirps_rain : CHIRPS vs station
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
n_rain__chirps_rain
syear
product
CHIRPS
station


les_cayes - n_rain__chirps_rain : CHIRPS vs station
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
n_rain__chirps_rain
syear
product
CHIRPS
station

Total monthly number of rain days from the station data (blue) and CHIRPS (red). The blue and red horizontal lines represent the long term average monthly number of rain days from the station data and CHIRPS respectively.

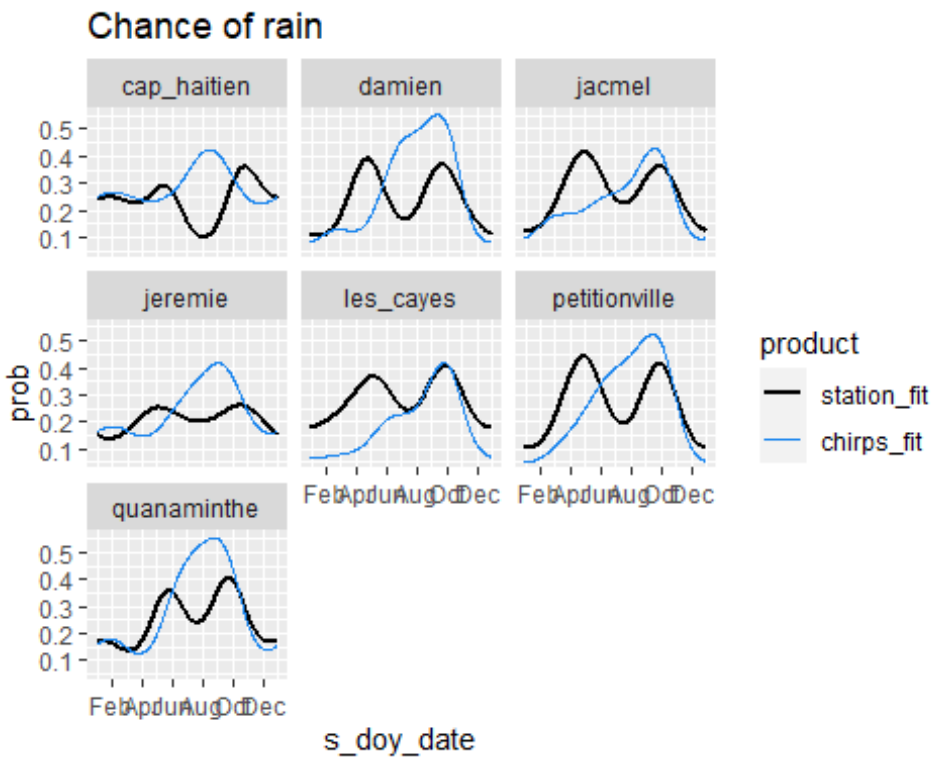
Modeled probabilities of rainfall

Zero-order Markov Chain models with three harmonics were fitted, using logistic regression, to the occurrence of rainfall (wet/dry, using a 0.85mm threshold for a rain day) to obtain fitted probabilities of rainfall through the 365/366 days of the year from the station and CHIRPS data

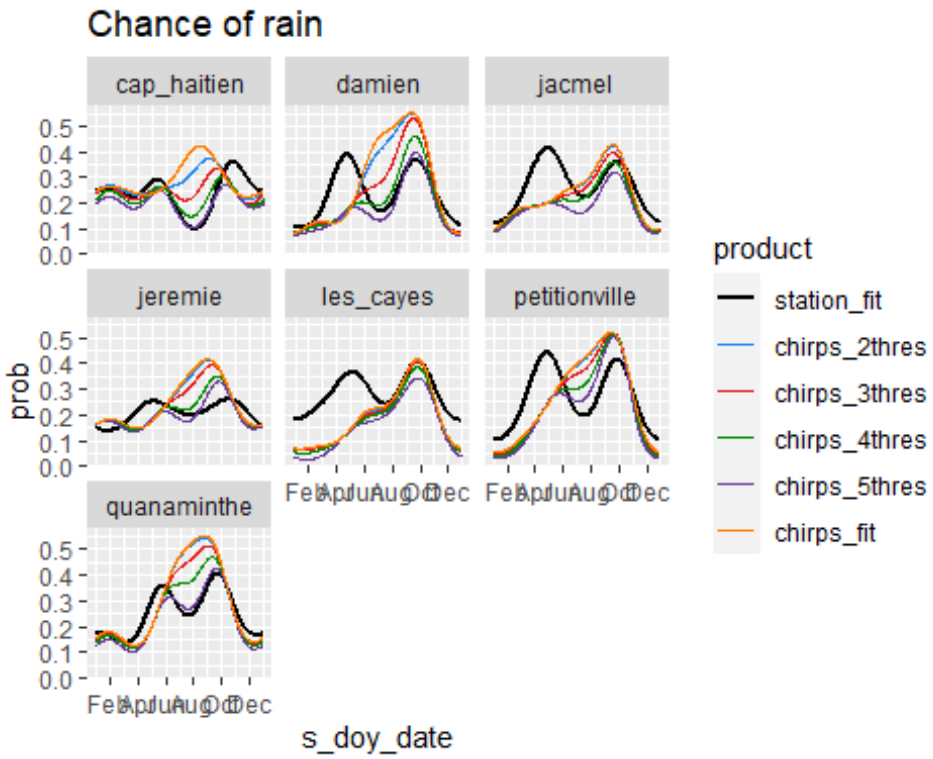
At all the stations, the CHIRPS curves have a similar seasonal pattern to the station data, however, at stations with a bimodal pattern, CHIRPS often misses or underestimates the first peak. For most stations the probabilities tend to be closer for the months Oct-Dec.

In the second set of curves, where multiple threshold on the CHIRPS data were applied, a threshold of 5mm seems to improve the fit to the station data in parts of the year for quanaminthe and cap haitan, however it does not make much of a difference to the remaining stations.

It is promising to see closer fits in the hurricane season at some stations when slightly higher thresholds are applied. Increasing the threshold for a rainy day can be used to correct biases in the number of rain days. However, this can only correct for overestimations, as are observed in the hurricane season, but not underestimation, which is seen outside the hurricane season. It appears that any correction to the rain day threshold should be seasonally dependent and location dependent, at least to some extent.



Fitted probabilities of rainfall throughout the year for the station data (black) and CHIRPS (blue). Values are fitted for each 365/366 days of the year on the x-axis with probability values on the y-axis.



Fitted probabilities of rainfall throughout the year for the station data (black) and CHIRPS (coloured line). Each coloured line represents the probabilities using a different threshold for a rain days. The threshold values used are: 0.85mm (orange), 2mm (blue), 3mm (red), 4mm (green), 5mm (purple). As the threshold increases, the probabilities decrease, hence the curves with larger thresholds are beneath those with lower thresholds. Values are fitted for each 365/366 days of the year on the x-axis with probability values on the y-axis.