USING PRINCETON PRECIPITATION CLIMATOLOGY TO PREDICT FUTURE PRECIPITATION EVENTS

Tyrone Zhang

A SENIOR THESIS

PRESENTED TO THE FACULTY

OF PRINCETON UNIVERSITY

IN CANDIDACY FOR THE DEGREE

OF BACHELOR OF ARTS

RECOMMENDED FOR ACCEPTANCE

BY THE

DEPARTMENT OF GEOSCIENCES

Adviser: Frederik J. Simons

March 7, 2021

This paper represents my own work in accordance with University regulations,

Tyrone Zhang

Abstract

Princeton's climate has four seasons, with strong temperature variations, and precipitation occurring throughout the year. Statistically, precipitation event sequences can be characterized as drawn from exponential distributions in the three variables *duration*, *intensity* (the total precipitation divided by the duration), and interevent *separation*. The shortest and least intense precipitation events are the most frequent. Analyzing the precipitation measured from 2017 to the present day by a Vaisala WXT530 weather station located on the roof of Guyot Hall, I first summarize the data in terms of exponential distributions and their parameters, by season and by year. Subsequently, I evaluate the skill in predicting the arrival, duration and intensity of precipitation events solely based on this local "climatology", before including other variables logged by the weather station.

Acknowledgements

I would like to acknowledge my senior thesis advisor Frederik J. Simons for giving me constant feedback on my work as well as providing me with the data that he is collecting on top of Guyot Hall.

Contents

List of Figures

List of Tables

Introduction

The climatology of Princeton is one that belongs to the mid-latitudes, which is characterized by having four seasons that results in a large variation in temperature throughout a year. In terms of the average calculated between 1981 and 2010, Princeton gets an average of 1227 mm of precipitation annually, and the precipitation distribution throughout the year is fairly even, with less precipitation in the winter (PRISM, 2020). According to the Koppen-Geiger Climate Classification, Princeton, NJ lies in the classification Cfa, which denotes a temperate climate, with no dry season, and hot summers defined as reaching 22°C or higher (Peel et al., 2007). Princeton having no dry season means that precipitation is well spread out throughout the year.

Our weather station on the top of Guyot Hall is Vaisala weather transmitter WXT530 series. It measures six weather parameters of air pressure, temperature, humidity, rainfall, wind speed, and wind direction. The rainfall is measured using an acoustic Vaisala RAINCAP Sensor, which helps avoid the complications of flooding, wetting, and evaporation losses (Vaisala, 2017).

Air temperature is the temperature that the thermometer measures when exposed to the air, while sheltered from direct solar radiation. Atmospheric Pressure is measuring the pressure exerted by the atmosphere due to gravitation attraction on the air column above a point in question Glickman (2000). Relative Humidity is the ratio of vapor pressure to saturation vapor pressure with respect to water Glickman (2000). This means that two air parcels can the same amount of water vapor, but if one air parcel is a warmer temperature than the other parcel, the higher temperature air parcel has a lower relative humidity than the lower temperature air parcel, since warmer air has a higher saturation vapor pressure for water. Precipitation accumulation is the amount of water substance that has fallen at a given point over a specified period of time Glickman (2000).

Shown in Figure 1, data is being recorded every minute and air pressure, air temperature, relative humidity, and Precipitation accumulation is being plotted. This day happens to have precipitation, thus we can see some of the variables that could lead to precipitation actually occuring. It appears that relative humidity becomes higher before a precipitation event, and that

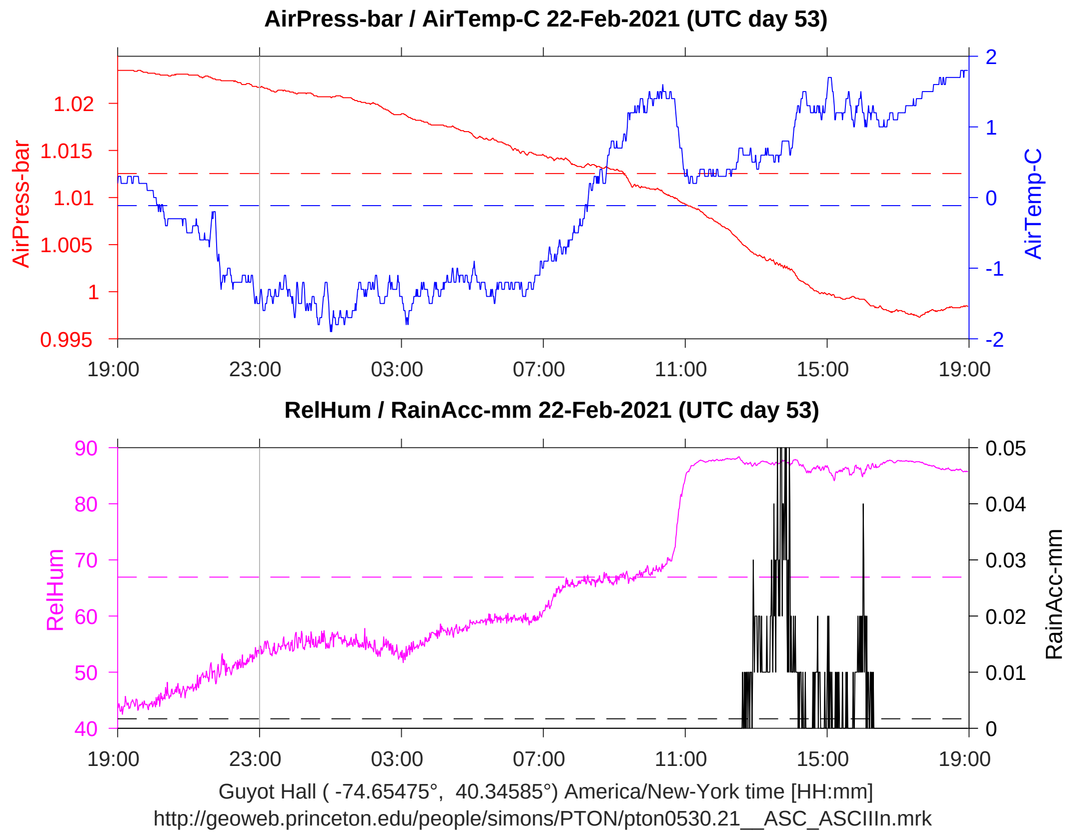

Figure 1: This is a daily plot of the meteorological variables that the Vaisala station measures on top of Guyot Hall.

the air pressure is low, while a precipitation event occurs. However, this is only one event and many precipitation events will be

From the data that I analyzed, the following is the climatology of Princeton from 2017 to 2021. Looking at Figure 2, we see that we have temperature and precipitation that varies throughout the year. Temperatures start increasing before peaking in July, before decreasing until January. This is consistent with the fact that Princeton has four seasons, which shows cold winters, warm summers, and mild springs and falls. At the same time, there is more precipitation in the spring and summer compared to the relative lack of precipitation in the fall and winter.

The month with the lowest average precipitation appears to be January at an average monthly precipitation of 44 mm. Whereas, the month with the highest average precipitation is July with 123 mm.

When looking at Figure 3, we see that the ranges of temperature are not equal in throughout the months. For example winter months of January and Feburary have a big range that span from about 20°C to -10°C. Meanwhile we have summer months such as July and August whose range of temperatures is smaller, from about 30°C to about 15°C.

We also see that relative humidity for Princeton is higher in the months of August to October, whereas the lowest relative humidity seems to be in the months of March and April. With Figure 4, we can see that all the months have similar upper extremes of about 90% for relative humidity, which can indicate precipitation. At the same time, the range for relative humidity is largest in the Spring time, with March, April, and May seem to have a range between 15% to 90% for relative humidity, with caution that this is just data from 3 years.

For atmospheric pressure, the lowest monthly air pressure are found in the months of April and June, while the highest monthly air pressure are found in the month of Feburary. Though the values for the mean atmospheric pressure are all pretty close to each other ranging from about 1.00 bar to 1.01 bar. The more interesting aspects of atmospheric pressure comes from the range of atmospheric pressure, where the months of May to September all seem to have a small range that vary between 0.995 bar to about 1.02 bar. Meanwhile we have months like October, which have a larger range of atmospheric pressure, which ranges from 0.97 bar to 1.03 bar. Clearly these low atmospheric pressure must come from an intense low pressure that passed near Princeton and it turns out that for the month of October from 2017-2020, we see that the values between 0.97 bar to 0.985 bar represent the 0 to 1st percentile of all atmospheric pressures for the month of October, so they are extremely low pressures. Other months such as November to February have relatively large ranges for atmospheric pressure as well.

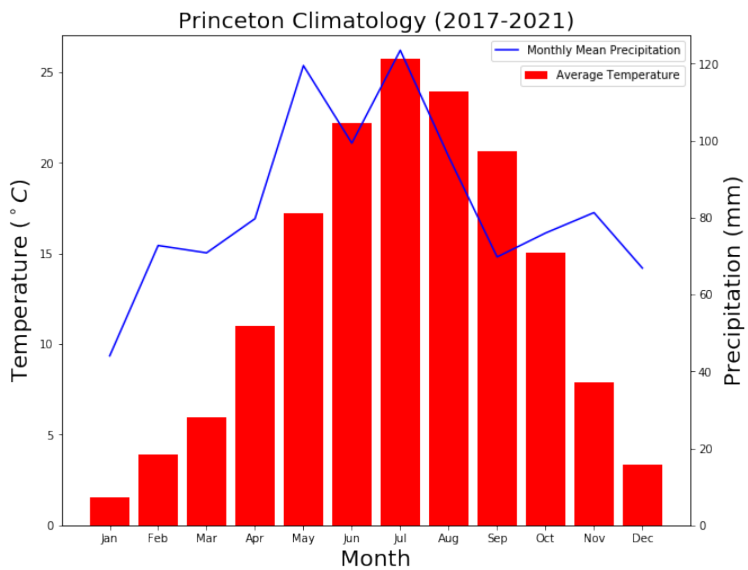

Figure 2: Climatology shown for Princeton from August 2017 to January 2021. The black interval lines show the 25-75th interpercentile range for temperature of the month.

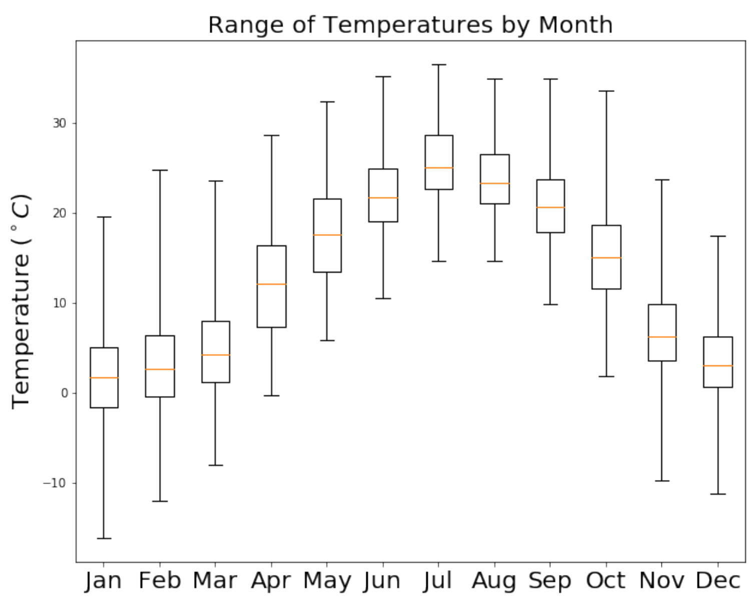

Figure 3: Shows the interpercentile ranges as a bar and the mean temperatures as a line graph. The whiskers goes out to the 0th percentile and 100th percentile respectively.

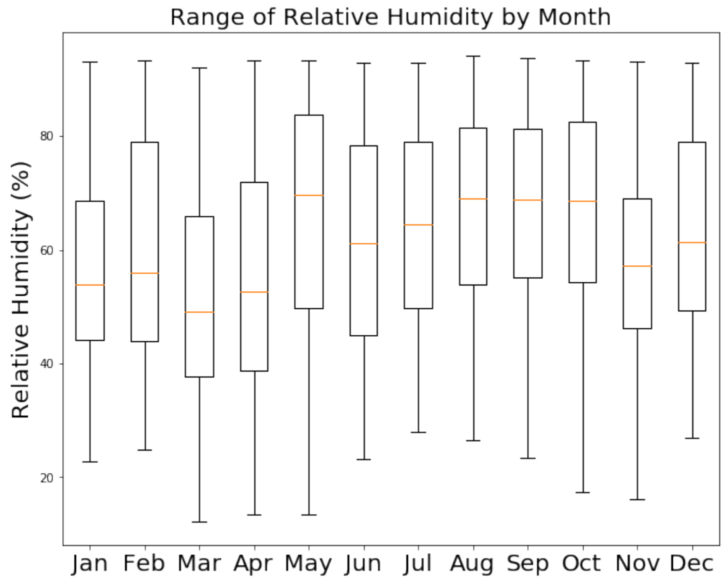

Figure 4: Range of relative humidity for Princeton with the mean relative humidity for the month indicated by the orange line.

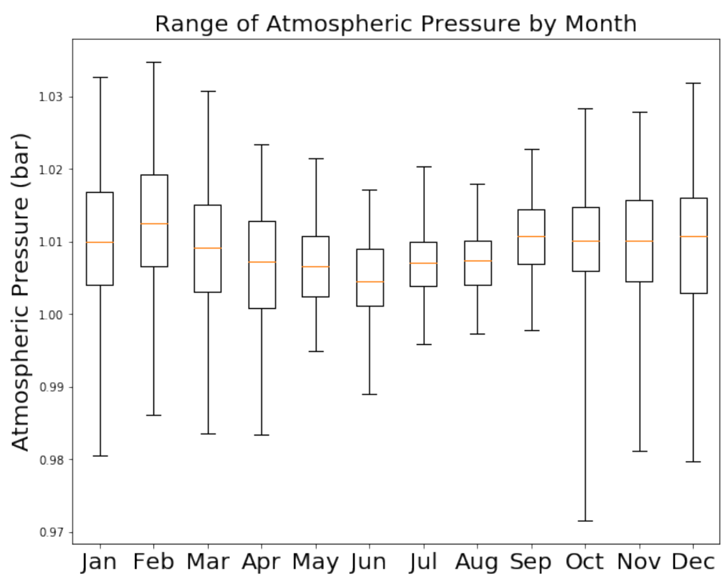

Figure 5: A similar graph with 4 looking at range of atmospheric pressure through out the months.

Descriptive Climatology of Precipitation

Despite the information given to us in looking the ranges of meteorological variables like temperature, relative humidity, and atmospheric pressure, precipitation events are no where to be seen. I shall define the following terms. The time series of **precipitation** as recorded by the instrument is denoted e_i, where i indexes the measurement intervals, each 60 s long. I define a precipitation **event** E_j^τ as a sequence of **duration** $d_j \geq \tau$ containing contiguous nonzero precipitation measurements $e_i > 0$, flanked left and right by zeros, $e_i = 0$, and where τ is in minutes Eagleson (1978).

Furthermore, I define a precipitation **non-event** N_j^τ, as having a contiguous set of zeros, $e_i = 0$, whose combined duration exceeds τ, flanked left and right by non-zero values, $e_i = 0$ Eagleson (1978). One more term to define is **precipitation intensity**, which for a precipitation event E_j^τ is the total amount of precipitation divided by its duration, i.e.,

$$I_j^\tau = \frac{\sum_i e_i}{d_j}, \quad \text{for } i \text{ belonging to the event } E_j^\tau. \tag{1}$$

For further analysis, I am breaking down the year into seasons, as different seasons may have different characteristics with regards to precipitation. I will define the seasons as follows: Winter will be December, January, and February. ‘Winter’ of a certain year contains December of the previous year. Spring will be March, April, and May. Summer is June, July, and August. Fall is September, October, November.

These following histograms show the distribution of precipitation event duration in terms of minutes and shows the different duration distributions per season. For precipitation and non-precipitation events, $\tau = 1$ min, which means we will be recording all precipitation events that are 1 minute or longer in duration. This covers all potential precipitation events.

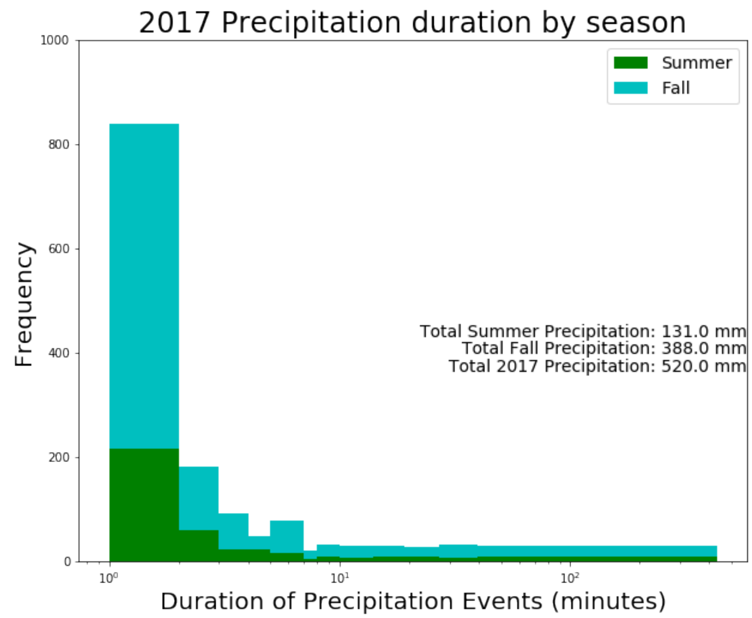

Figure 6: Distribution of precipitation duration in 2017, with a breakdown by seasons. The minimum duration is $\tau = 1$ min. Note that data collection began on 16 July 2017. Which means we only have part of the summer and all of the fall data.

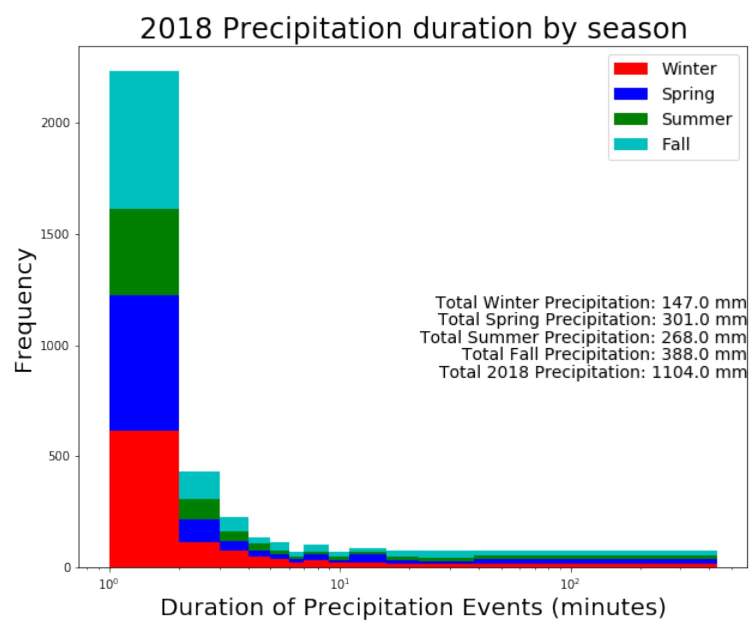

Figure 7: Distribution of precipitation duration in 2018, with a breakdown by seasons. 2018 is the first full year of data collections. The minimum duration is $\tau = 1$ min.

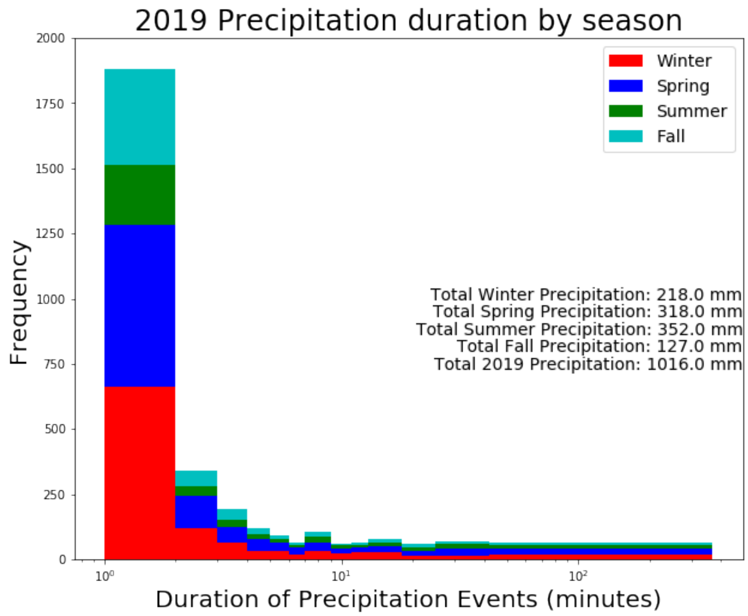

Figure 8: Distribution of precipitation duration in 2019, with a breakdown by seasons. The minimum duration is $\tau = 1$ min.

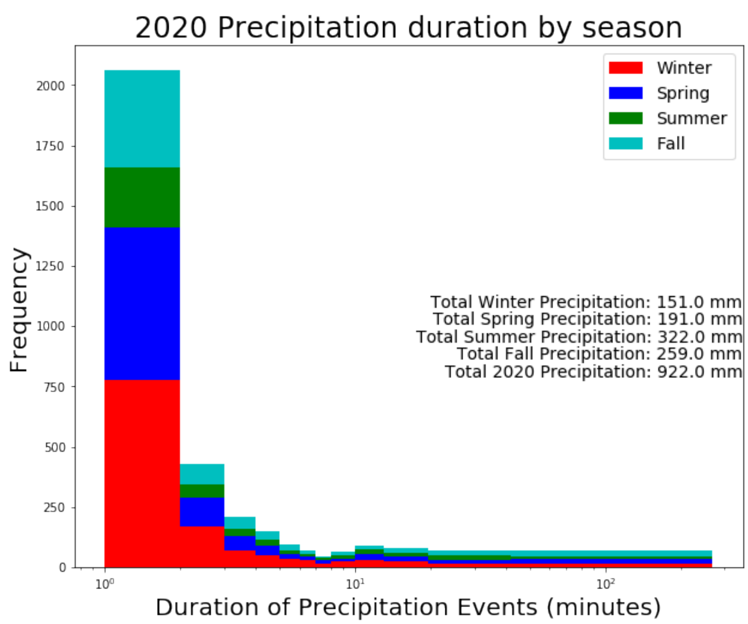

Figure 9: Distribution of precipitation duration in 2020, with a breakdown by seasons. The minimum duration is $\tau = 1$ min.

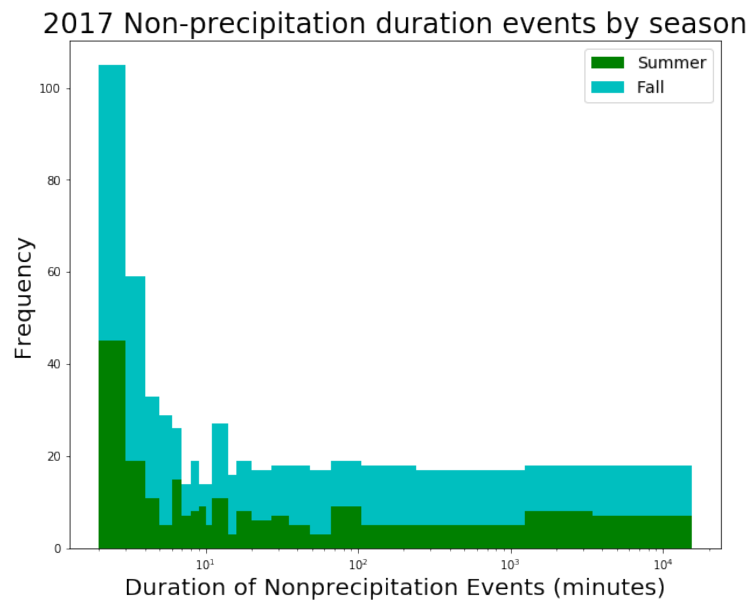

Figure 10: Distribution of non-precipitation event duration in 2017, with a breakdown by seasons. The minimum duration is $\tau = 1$ min. Note that data collection began on 16 July 2017, which means that Summer is only partially complete.

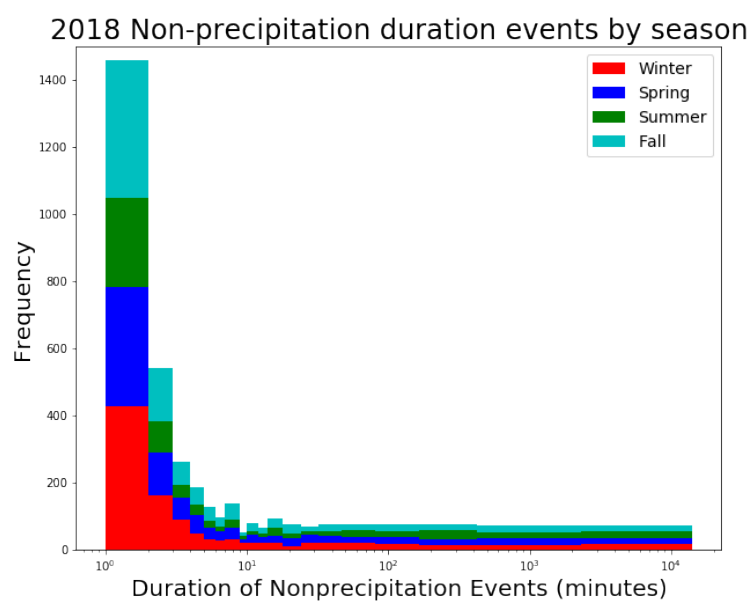

Figure 11: Distribution of precipitation duration in 2018, with a breakdown by seasons. The minimum duration is $\tau = 1$ min.

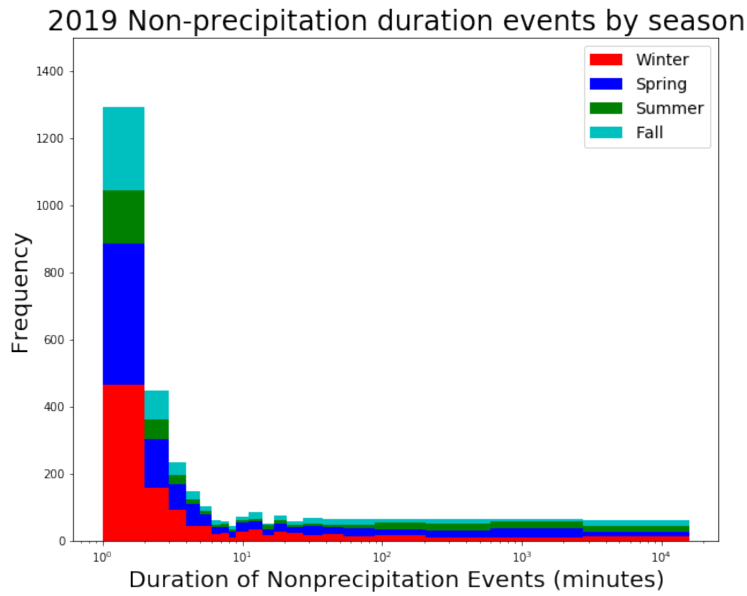

Figure 12: Distribution of non-precipitation event duration in 2019, with a breakdown by seasons. The minimum duration is $\tau = 1$ min.

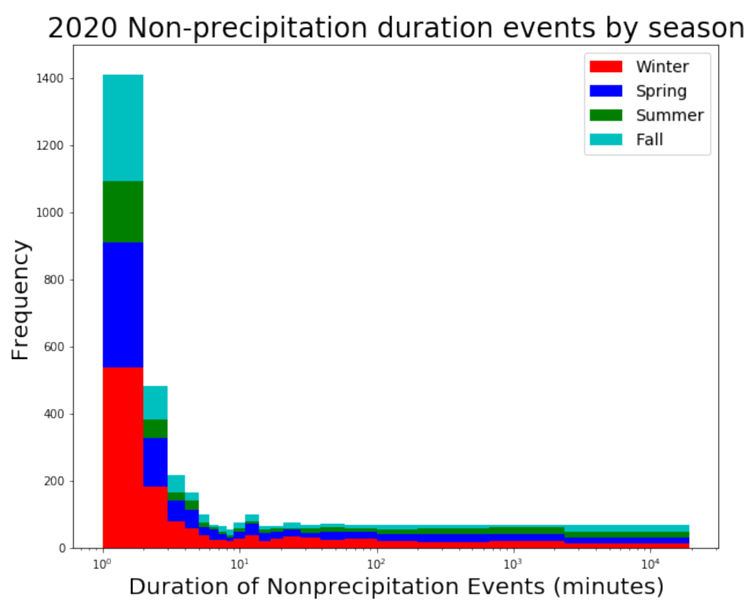

Figure 13: Distribution of non precipitation event duration in 2020, with a breakdown by seasons. The minimum duration is $\tau = 1$ min.

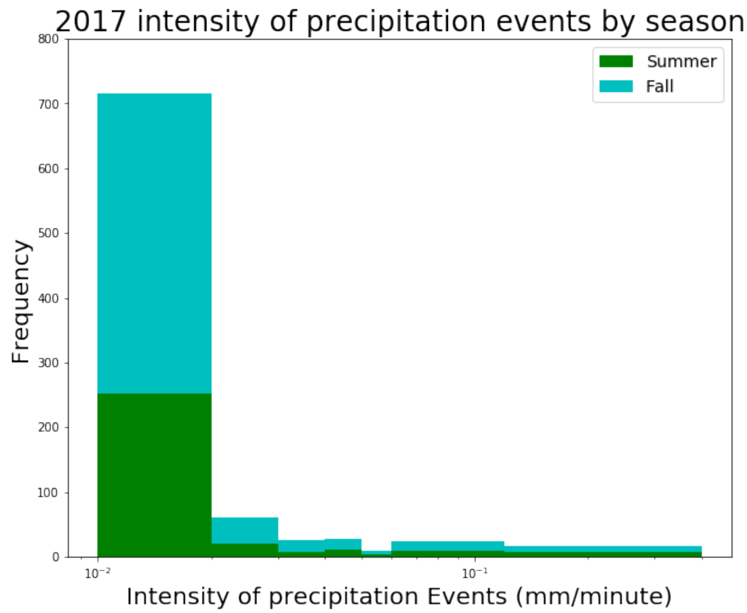

Figure 14: Distribution of precipitation intensity in 2017, with a breakdown by seasons. Note that data collection began on 16 July 2017. The minimum intensity is $I = 0.01$ mm/min

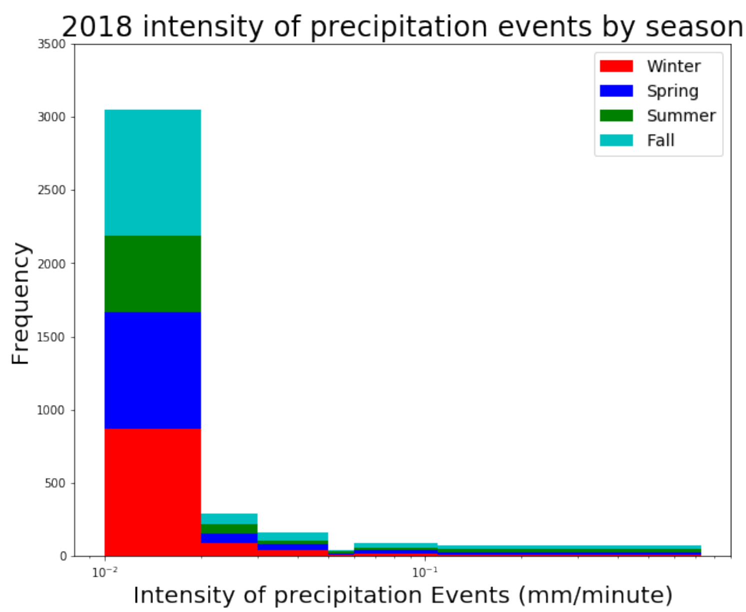

Figure 15: Distribution of precipitation intensity in 2018, with a breakdown by seasons. The minimum intensity is $I = 0.01$ mm/min

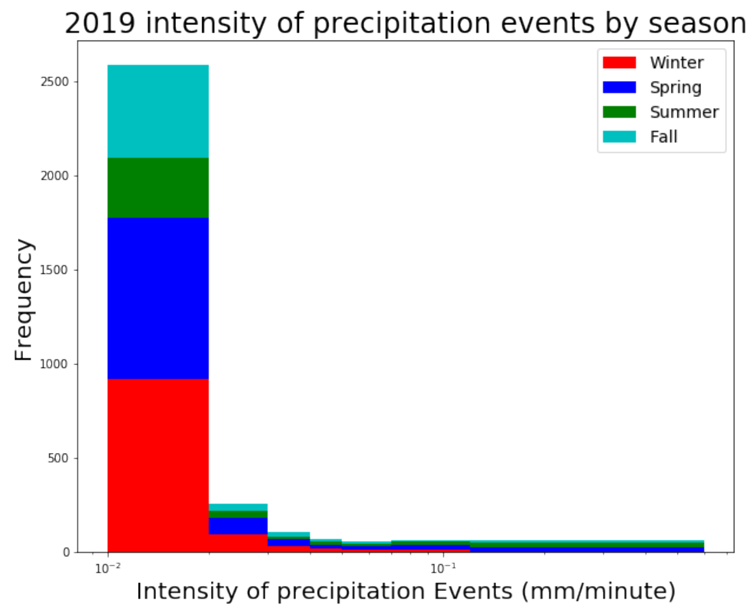

Figure 16: Distribution of precipitation intensity in 2019, with a breakdown by seasons. The minimum intensity is $I = 0.01$ mm/min.

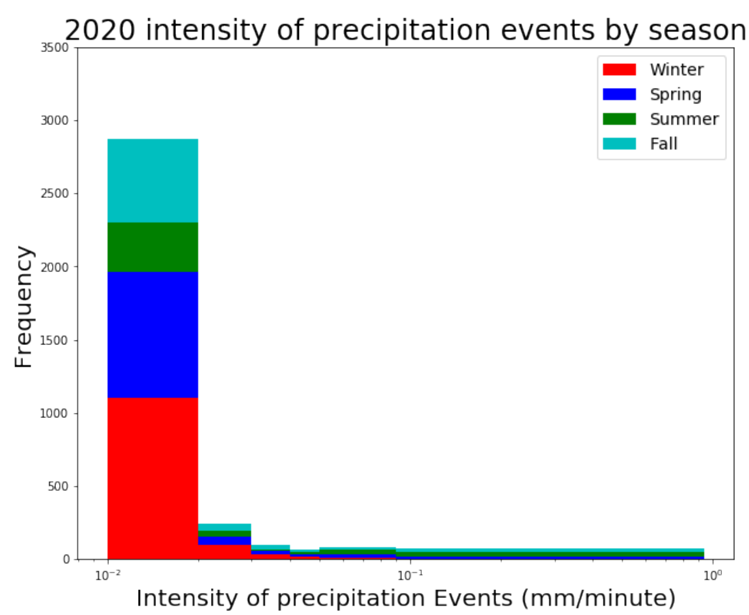

Figure 17: Distribution of precipitation intensity in 2020, with a breakdown by seasons. The minimum intensity is $I = 0.01$ mm/min.

Looking at these histograms, we can see that the smallest duration events for both nonprecipitation and precipitation events are the most numerous as seen with the largest bar in the bins that contain 1 minute to 2 minute durations. We can also see the frequency of duration decreases as we increase the duration bins. At the same time, there are differences between the precipitation and non-precipitation events, as we see that the upper limit for duration of precipitation events are about 300 to 400 minutes. However, the upper limit for duration of non-precipitation events is about an order of magnitude greater at about 2000 minutes. Basically non-precipitation events can be as long as one or two whole days, where as precipitation events can only last as long as 5-6 hours.

The intensity histogram shows an even more aggregation in the lowest intensity bin of 1 mm/hour to 2 mm/hour. It's interesting cause the small intensity bin can include events that are long but less intense and short, less intense precipitation events.

Analysis of Precipitation

Methods

Once we have characterized the climatology of precipitation, we will start moving towards buidling predictive models for precipitation. In order to see how accurate such models are in comparison to observed precipitation. Such accuracy we shall use is to see whether the model and the observed data match in terms of whether there exists precipitation at a given minute. We ignore the non-precipitation when looking at accuracy because by matching non-precipitation minutes in both the model and the observed data, we end up getting an accuracy of over 90%, which not useful. So, if the model and the observed data do not match in terms whether there exists precipitation, this contributes to the model being deemed less accurate. We can have a stricter definition of accuracy, in which we set up the precipitation condition, in addition to saying that the intensity must match too, otherwise we can not say the model and observed data match. This stricter definition of model accuracy might be used when thinking about

Results

Figure 8 shows the distribution of durations of 3198 precipitation events E_j^1, i.e. E_j^τ where $\tau = 1$ min for the year 2019, broken down by season. In order to make bins that contain non-zero values, I created bins using duration percentiles. I used unique values obtained from using a range of percentiles from 0 to 100, in 2 and 5 percent intervals. For one analysis, I stopped at 98 % believing this would be the best approach in terms of fitting. At the same time, I also made sure to analyze the data including the 100 percentile, to see how it differs when including the extremes. Such purpose is also to realize that excluding such extremes produced modelled precipitation that did not last long as well as well as the gaps between precipitation events being too small.

I used an exponential fit to the frequency-duration histograms for all 1253 events E_j^2, i.e. E_j^τ where $\tau = 2$ min. For all the other years, as shown in Figures 9, 7 and 6, I used a similar procedure.

Excluding the first interval shown, focusing on events of duration greater than or equal to 2 min, we propose an exponential model for the histogram, with the following equation:

$$F = \beta e^{\alpha d}, \tag{2}$$

where F is the frequency and d the duration, and with β the unitless frequency coefficient and α is the exponential coefficient (in units of min^{-1}). Table 1 shows the coefficients β and the exponential coefficients α from looking at the yearly frequency of precipitation duration.

We shall also propose the following equations which will also describe an exponential model for the histogram regarding non-precipitation events, which is described by the following equation:

$$F_{np} = \gamma e^{\delta D}, \tag{3}$$

where F_{np} is the frequency of non-precipitation events, D being duration, and γ being the unitless frequency coefficient and δ is the exponential coefficient. Table 3 shows the coefficients γ and exponential coefficients δ from looking at yearly frequency of non-precipitation event durations.

Another similar equation for precipitation intensity can be described by the following equation:

$$F_{inten} = \varepsilon e^{\zeta I}, \tag{4}$$

where F_{inten} is the frequency of intensity of precipitation events, I is the intensity of the precipitation events, ε is the unitless frequency coefficient, and ζ is the exponential coefficient. Table 5 shows the coefficients ε and the exponential coefficients ζ from looking at yearly frequency of intensity of precipitation events.

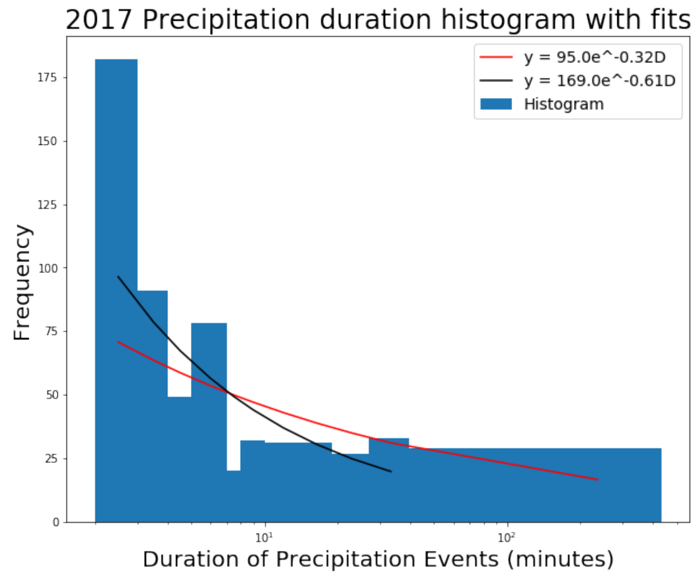

Figure 18: Curve fitting of the precipitation histogram excluding 1 minute duration events for the incomplete 2017 data. The red curve denotes the curve that fits to the 100th percentile, while the black curve fits the data to the 98th percentile. Has similar trends to the other years, though 2017 has incomplete data.

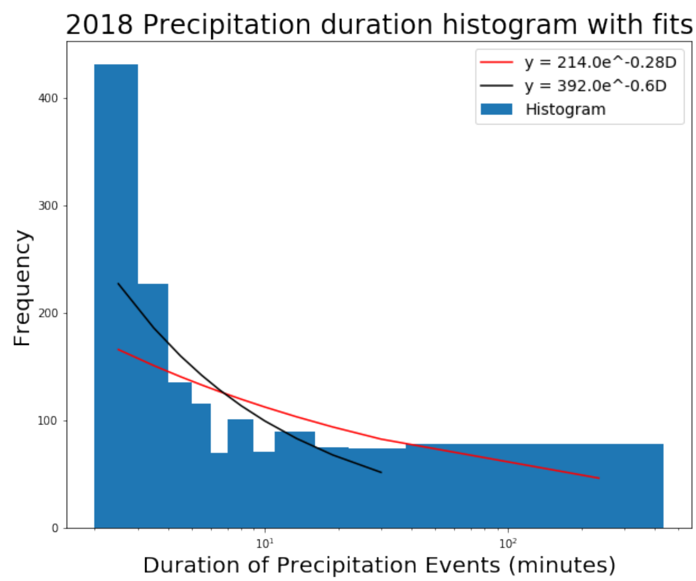

Figure 19: Curve fitting of the precipitation histogram excluding 1 minute duration events for the entire 2018 data. Format is the same as Figure 18, with the understanding that 2018 has complete data.

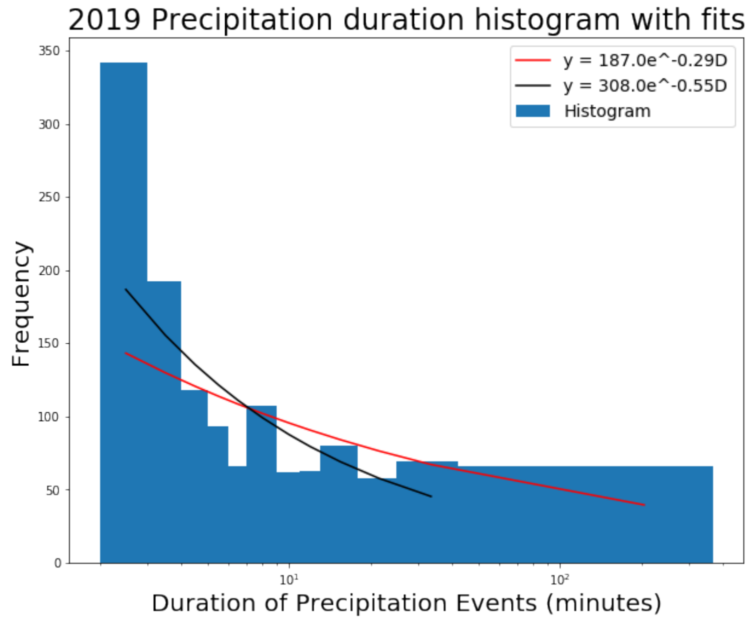

Figure 20: Curve fitting of the precipitation histogram excluding 1 minute duration events for the entire 2019 data. The layout is the same as seen in Figure 18, with 2019 having complete data like 2018. 2018 and 2019 have similar trends and similar curves for both best fit curves.

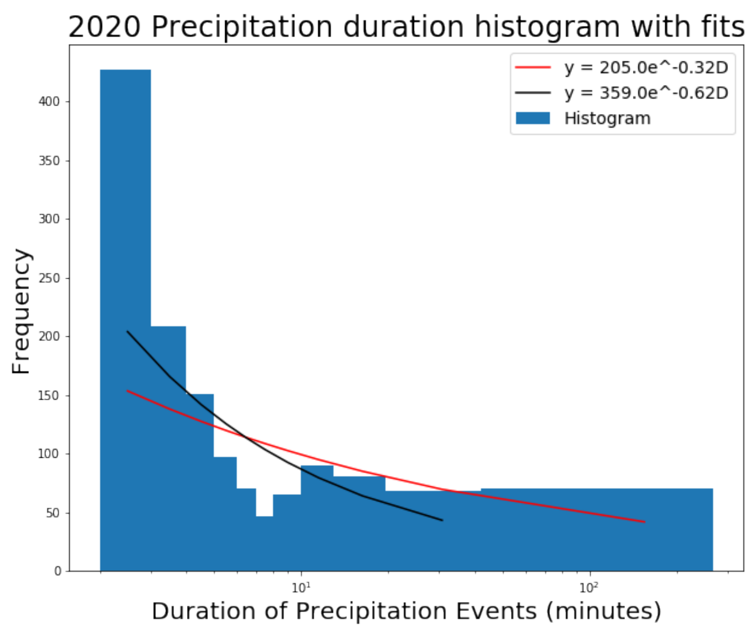

Figure 21: Curve fitting of the histogram excluding 1 minute duration events for the entire 2020 data. The layout is the same as seen in Fig18 with 2020 having completed data as well. 2018, 2019, and 2020 all look similar looking at the histogram and the best fit curves.

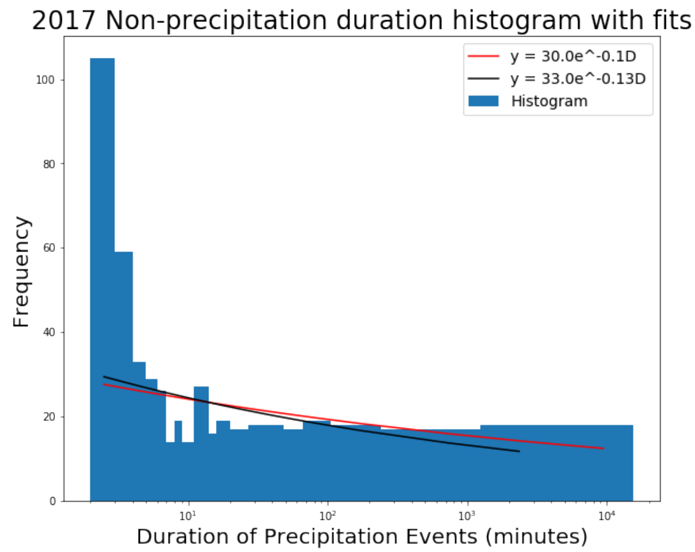

Figure 22:
Curve fitting of the histogram excluding 1 minute duration events for the incomplete 2017 data. The red curve denotes the curve that fits to the 100th percentile, while the black curve fits the data to the 98th percentile. Has similar trends to the other years, though 2017 has incomplete data.

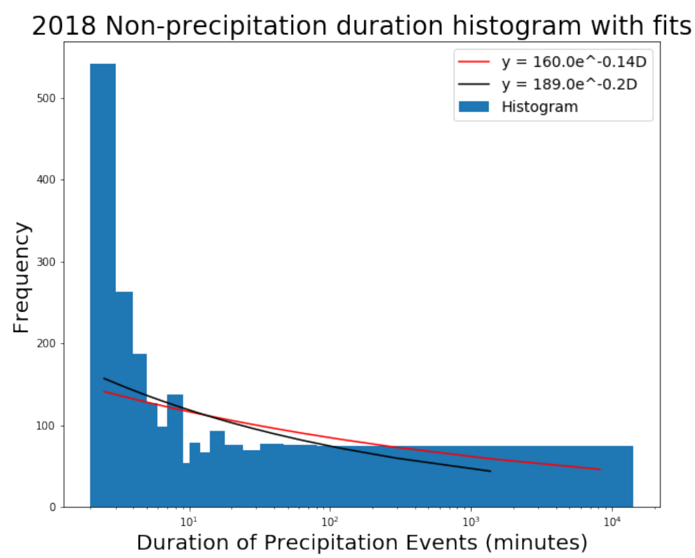

Figure 23: Shows curve fitting of the histogram excluding 1 minute duration events for the entire 2018 data. The red curve denotes the curve that fits to the 100th percentile, while the black curve fits the data to the 98th percentile. The 98 percentile curve fitting the smaller durations better, while the 100 percentile curve fitting the larger durations better. Perhaps the only real difference between 2017 and 2018 is the frequency, with 2018 having complete data.

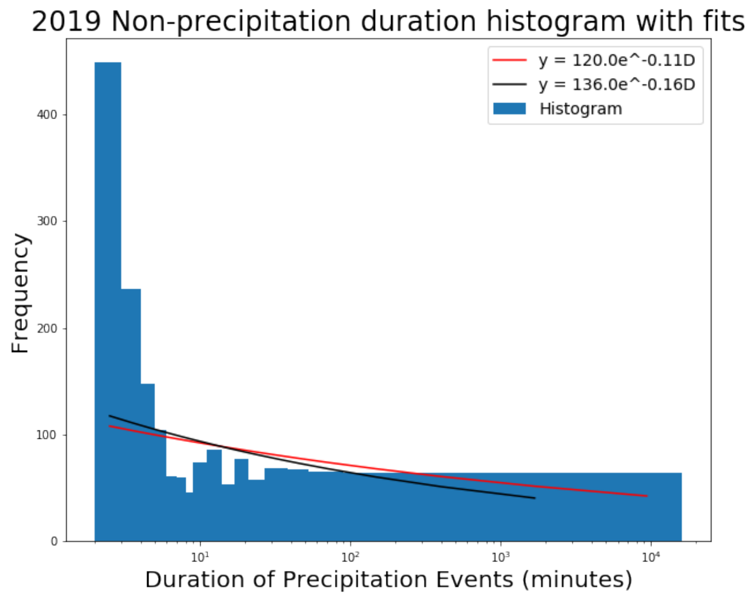

Figure 24: Shows curve fitting of the histogram excluding 1 minute duration events for the entire 2019 data. The layout is the same as seen in Figure 19. 2018 and 2019 have similar trends and similar curves for both best fit curves.

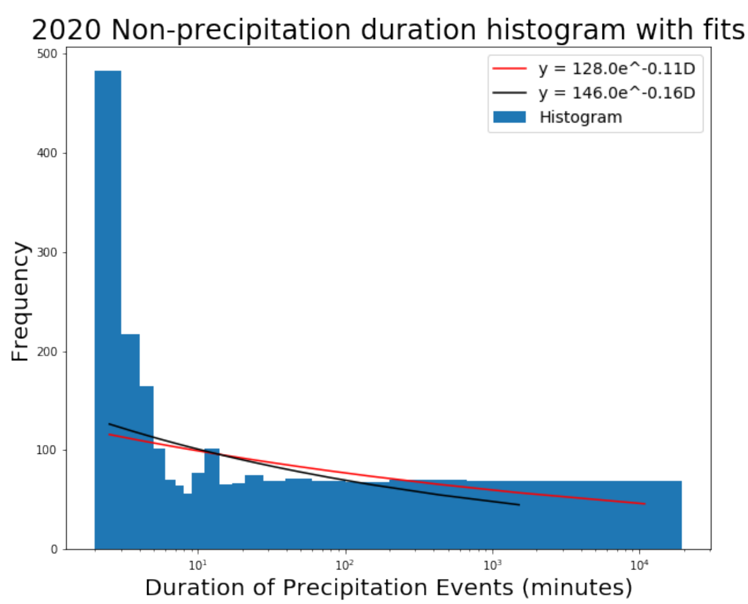

Figure 25: Curve fitting of the histogram excluding 1 minute duration events for the entire 2020 data. The layout is the same as seen in Figure 23. 2018, 2019, and 2020 all look similar looking at the histogram and the best fit curves.

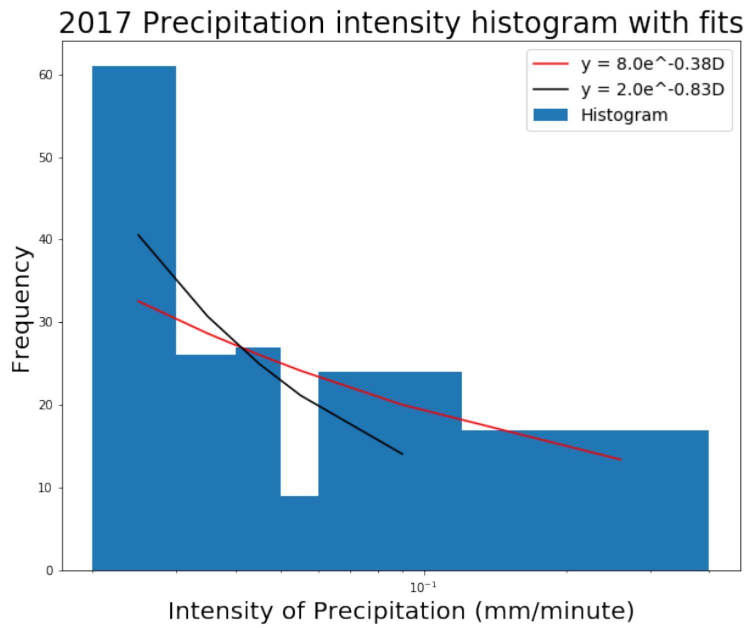

Figure 26: Curve fitting of the histogram excluding the 1 mm/minute intensity bin for the 2017 data that was collected. The black curve fits up to the 98th percentile bin. The red curve fits up to the 100th percentile bin. The black curve fits the lower bins better, but the red curve fits to include the extremes better.

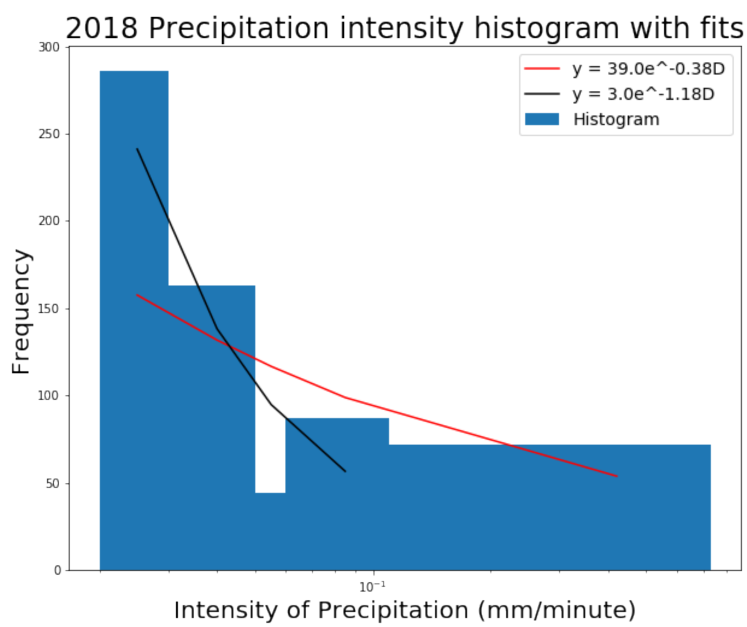

Figure 27: Curve fitting of the histogram excluding the 1 mm/minute intensity bin for the entire 2018 data. The layout is the same as Figure 26, with the red curve fitting the extremes better, while the black curve fits the lower intensity bins better.

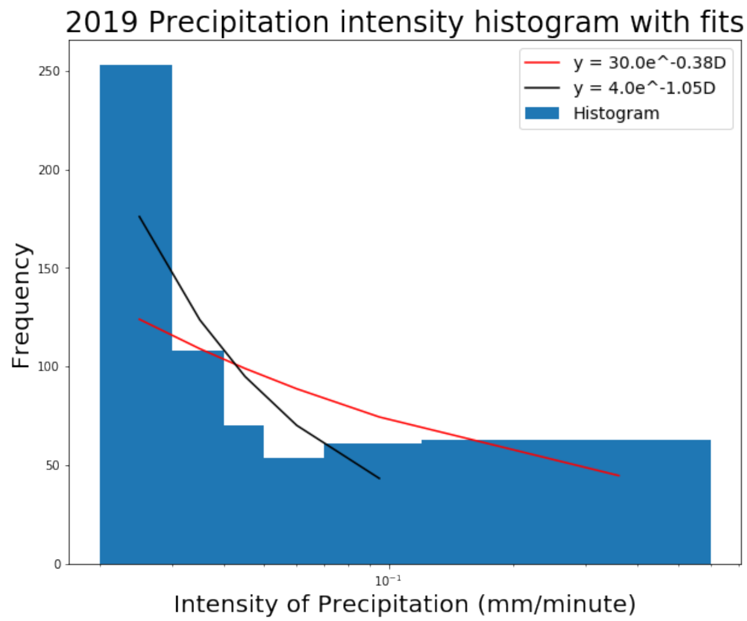

Figure 28: Curve fitting of the histogram excluding the 1 mm/minute intensity bin for the entire 2019 data. The layout is the same as Figure 26, with the red curve fitting the extremes better, while the black curve fits the lower intensity bins better.

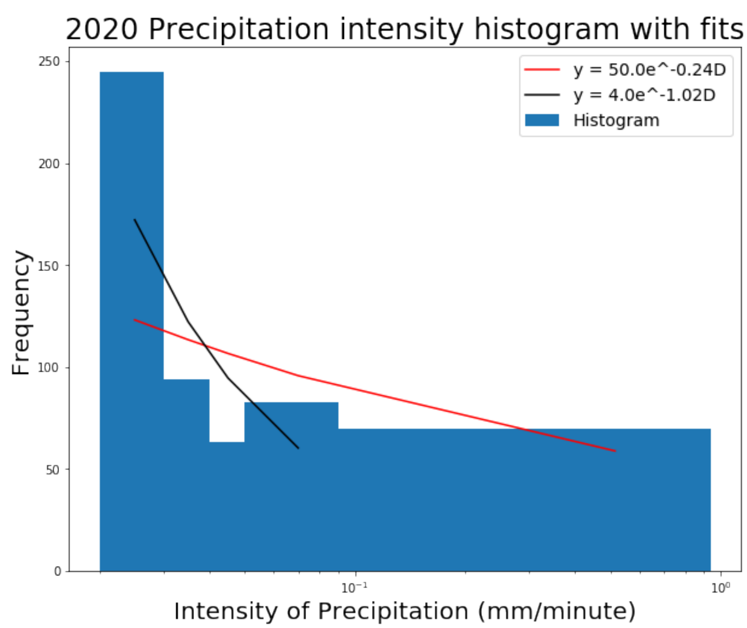

Figure 29: Curve fitting of the histogram excluding the 1 mm/minute intensity bin for the entire 2020 data. The layout is the same as Figure 26, with the red curve fitting the extremes better, while the black curve fits the lower intensity bins better.

Discussion

Season	Annual		Winter		Spring		Summer		Fall	
Year	β	α	β	α	β	α	β	α	β	α
2017	*169*	*-0.61*	NaN	NaN	NaN	NaN	*57*	*-0.70*	108	-0.57
2018	392	-0.59	148	-0.74	125	-0.76	65	-0.52	91	-0.50
2019	308	-0.54	118	-0.62	107	-0.58	32	-0.31	57	-0.60
2020	408	-0.65	201	-0.80	112	-0.66	48	-0.42	*64*	*-0.66*

Table 1: Coefficients found from the yearly distribution of precipitation duration (as in equation 2) as well as the seasonal distribution of precipitation duration. Italics refer to values obtained using incomplete information. NaN means there was no information. These coefficients were computed using the 0 to 98th percentile of precipitation duration.

Season	Annual		Winter		Spring		Summer		Fall	
Year	β	α	β	α	β	α	β	α	β	α
2017	*94*	*-0.32*	NaN	NaN	NaN	NaN	*32*	*-0.14*	68	-0.09
2018	214	-0.28	73	-0.37	46	-0.27	38	-0.24	55	-0.23
2019	224	-0.28	68	-0.34	60	-0.28	25	-0.19	33	-0.32
2020	234	-0.35	119	-0.51	60	-0.32	36	-0.27	*29*	*-0.23*

Table 2: Coefficients found from the yearly distribution of precipitation duration (as in equation 2) as well as the seasonal distribution of precipitation duration. Italics refer to values obtained using incomplete information. NaN means there was no information. These coefficients were computed using the all the data, from 0 to 100th percentile.

Based on Table 1, the β values are similar to each other with the exception of 2017, which only had partial data starting in the Summer. Even, with the partial data we have from 2017 and 2020, it is clear that the α values are similar to each other. Focusing on the years with complete data, 2018 and 2019, it is clear that yearly variations exists between them. All α values are negative.

In the seasonal variations, we see that Summer has α values that are less negative compared to the other seasons as well as having a lower β compared to the other seasons. The Spring appears to mimic the Winter in that there are a lot of preciptiation events, but also shares the quality of Summer in having fairly high precipitation totals. The Fall shares the Summer quality of having relatively few preciptation events, but tends to have smaller precipitation totals, so its intensity is less than the Summer, but greater than the winter precipitation intensity.

Compared to Table 1, Table 2 seems to show a β value that are consistently lower when considering the 100th percentile bins for precipitation event durations. The α values are also less negative, which better reflects the fact there are these extreme durations for precipitation events.

Season	Annual		Winter		Spring		Summer		Fall	
Year	γ	δ	γ	δ	γ	δ	γ	δ	γ	δ
2017	*33*	*-0.13*	NaN	NaN	NaN	NaN	*12*	*-0.15*	19	-0.12
2018	189	-0.20	59	-0.28	50	-0.29	26	-0.10	55	-0.22
2019	136	-0.16	63	-0.30	46	-0.16	11	0.03	24	-0.16
2020	140	-0.17	64	-0.23	46	-0.16	13	-0.02	20	-0.21

Table 3: Coefficients found from the yearly distribution of non-precipitation event duration (as in equation 3) as well as the seasonal distribution of non-precipitation event duration. Italics refer to values obtained using incomplete information. NaN means there was no information. Using non-precipitation event durations from the 0 to the 98th percentile.

Season	Annual		Winter		Spring		Summer		Fall	
Year	γ	δ	γ	δ	γ	δ	γ	δ	γ	δ
2017	*30*	*-0.10*	NaN	NaN	NaN	NaN	*11*	*-0.10*	18	-0.09
2018	160	-0.14	47	-0.19	43	-0.14	23	-0.05	46	-0.15
2019	120	-0.11	49	-0.21	44	-0.15	11	0.04	19	-0.08
2020	121	-0.11	57	-0.19	41	-0.12	12	0.01	15	-0.11

Table 4: Coefficients found from the yearly distribution of non-precipitation event duration (as in equation 3) as well as the seasonal distribution of non-precipitation event duration. Italics refer to values obtained using incomplete information. NaN means there was no information. Using non-precipitation event durations from the 0 to 100th percentile.

Looking at trying to extend the exponential fit towards non-precipitation event durations, the exponentials for such duration, δ is less negative overall compared to the exponentials from the precipitation event duration, α. Some of this is explained from the fact that the duration of non-precipitation events range from 1 minute to over 2000 minutes, whereas the duration of precipitation events range from 1 minute to just over 300 to 400 minutes. At the same time how we bin the durations for both precipitation and non-precipitation events will affect how we get the fits.

Once again, taking into the account the most extreme durations for non-precipitation events lowers the frequency value of γ and the exponential constant of δ is less negative, when comparing Table 3 and Table 4

Season	Annual		Winter		Spring		Summer		Fall	
Year	ε	ζ	ε	ζ	ε	ζ	ε	ζ	ε	ζ
2017	*0.06*	*-1.98*	NaN	NaN	NaN	NaN	*0.03*	*-1.94*	0.04	-2.00
2018	0.19	-2.14	0.02	-2.41	0.03	-2.24	0.10	-1.87	0.08	-2.05
2019	0.23	-2.01	0.03	-2.29	0.07	-2.02	0.25	-1.47	0.03	-2.16
2020	0.06	-2.38	0.002	-3.00	0.02	-2.38	0.10	-1.74	0.02	-2.18

Table 5: Coefficients found from the yearly distribution of precipitation intensity (as in equation 4) as well as the seasonal distribution of precipitation intensity. Italics refer to values obtained using incomplete information. NaN means there was no information. Using intensity from 0 to 98th percentile

Season	Annual		Winter		Spring		Summer		Fall	
Year	ε	ζ	ε	ζ	ε	ζ	ε	ζ	ε	ζ
2017	*1.6*	*-1.05*	NaN	NaN	NaN	NaN	*0.8*	*-0.94*	0.8	-1.12
2018	12	-0.919	0.7	-1.36	3	-0.91	6	-0.67	4	-0.91
2019	9	-0.94	0.10	-1.9	3	-0.94	7	-0.48	1	-1.0
2020	16	-0.76	0.5	-1.4	3	-0.86	13	-0.32	3	-0.69

Table 6: Coefficients found from the yearly distribution of precipitation intensity (as in equation 4) as well as the seasonal distribution of precipitation intensity. Italics refer to values obtained using incomplete information. NaN means there was no information. Using intensity from 0 to 100th percentile.

In contrast, Table 5 for intensity of precipitation events show a more negative exponential compared to either precipitation or non-precipitation event durations. This is partially due to the fact that there are less bins in the intensity of preciptiation events as seen that the vast majority of precipitation events are not particularly intense on average.

Season	Annual		Winter		Spring		Summer		Fall	
Year	T	I	T	I	T	I	T	I	T	I
2017	*520*	*0.07*	NaN	NaN	NaN	NaN	*131*	*0.07*	388	0.07
2018	1104	0.06	147	0.03	301	0.07	268	0.08	388	0.07
2019	1016	0.06	218	0.04	318	0.06	352	0.11	127	0.05
2020	*858*	*0.06*	151	0.03	191	0.04	322	0.12	*194*	*0.06*

Table 7: Total precipitation (in mm) and average intensity (in mm/min) of precipitation for each year and season. Italics refer to values obtained using incomplete information. NaN means there was no information.

Looking at the different between Table 5 and Table 6, intensity clearly has the biggest changes from using 98th percentile to the 100th percentile. First the ε value actually increases when you use the 100th percentile rather than the 98th percentile. Furthermore, we see that the ζ value also becomes less negative , but very dramatically compared to non-precipitation and precipitation event durations.

However looking at Table 7, we see that the summers also have a lot of precipitation. The fewer amounts of precipitation events in the Summer, but with a lot of precipitation gives a higher intensity for the Summer. For the Winter, the α are the furthest from 0 and the β are large. However, Winter tends to have the lowest values for total precipitation and the combination of lots of preciptation events and low precipitation totals results in the lowest intensities among the four seasons.

Simulation for Prediction

Method

To make a base line model for us to look at, we need to first use the exponentials that we calculated for the duration of the precipitation event, intensity of the precipitation event, and duration of the non-precipitation events. However, we also need to manually add back in the lowest durations and lowest intensities, since we excluded those values when calculating the exponential distributions. Since the frequencies are in the bins of values that are distributed by percentiles, we make use of the fact that there is an equal chance of pulling a value in the bin ranges. Thus once we have all the probabilities in place for precipitation event duration, non-precipitation event duration, and precipitation event intensity, we can start running a simulation based on the values that come from the data that we analyzed.

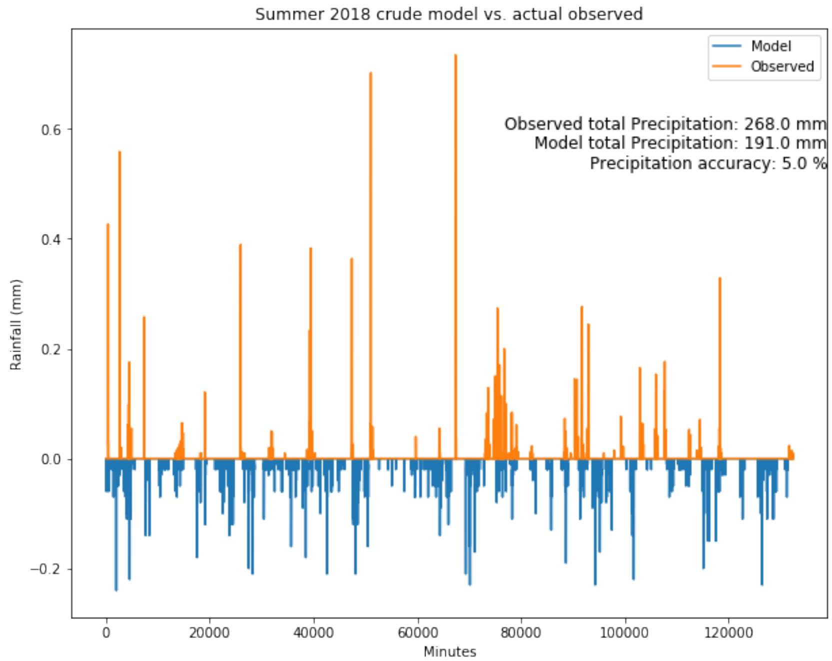

Figure 30:
One model run using the exponentials calculated from Summer 2018.

Results

Running the model one hundred times let us look at what the average precipitation total the model yields, it yields 195 mm compared to the Summer 2018 total of 268 mm that we used. Furthermore, the accuracy of model precipitation matching model precipitation is only 6%. Looking at 30, we see that the model precipitation total is lower than the actual precipitation observed in the Summer of 2018. Furthermore, the precipitation seem to be more frequent compared to the observed 2018 Summer data.

It is clear that despite the better fit using the curve fit of 98th percentile, we need to use the 100th percentile, since the extremes such as the big 2000 minute non-precipitation even is very important in the nature of the precipitation event. However, it is clear that the accuracy does not improve significantly. The fact that the model total precipitation is still less than from the initial observed precipitation shows how this model needs more improvement.

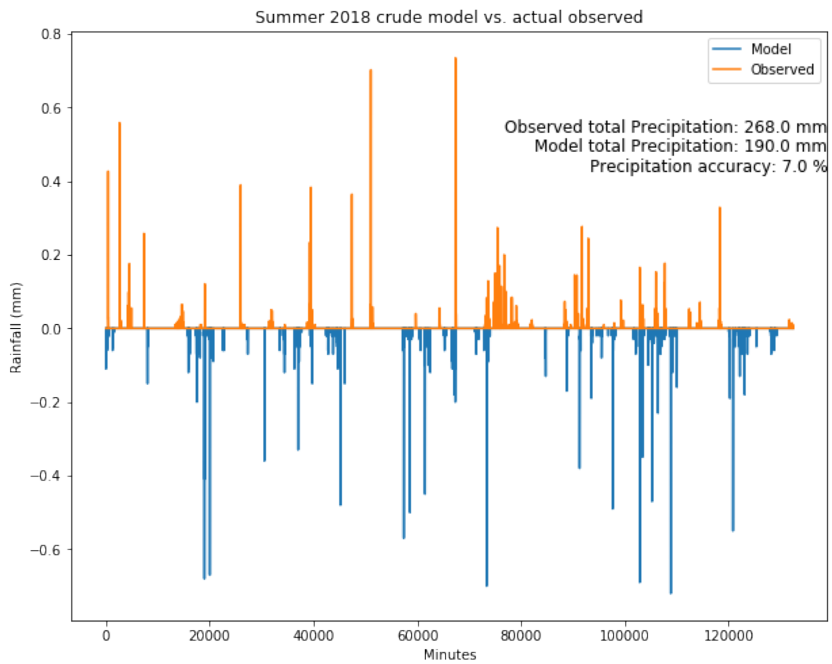

Figure 31: One model run using the exponentials calculated from Summer 2018. Made some adjustments to get a better match for the actual summer 2018 run.

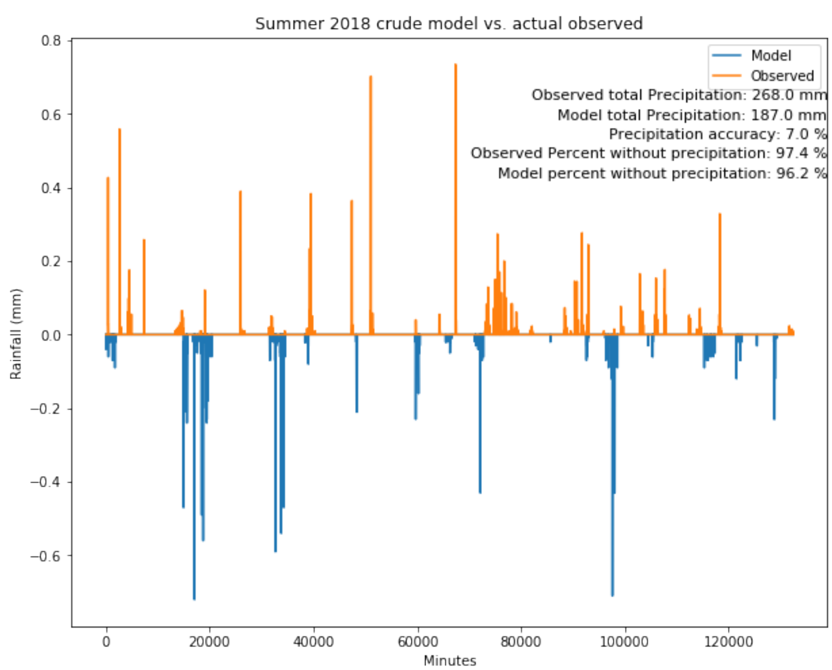

Figure 32: One model run using the exponentials calculated from Summer 2018. Put up some more information such as how much of the time was non-dominated by no precipitation.

Discussion

Machine Learning for Prediction

Method

Results

Discussion

Discussion and Conclusions

References

Eagleson, P., 1978. climate, soil, and vegetation: 2. The distribution of annual precipitation derived from observed storm sequences, *Water. Resource Res.*, **14**(5), 713–721.

Glickman, T. S., 2000. *Glossary of Meteorology*, Am. Meteorol. Soc., Boston, Mass, 2nd edn.

Peel, M., Finlayson, B., & McMahon, T., 2007. Updated world map of the Koppen-Geiger climate classification, *Hydrol. Earth Syst. Sci.*, **11**, 1633–1644.

PRISM, 2020. PRISM climate data, http://prism.oregonstate.edu.

Vaisala, 2017. Vaisala weather transmitter WXT530 series, `https://www.vaisala.com/sites/default/files/documents/WEA-MET-G-WXT530\%20Datasheet__Low.pdf`.