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

November 12, 2020

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

Tyrone Zhang

Abstract

Princeton's climate is one that has four seasons and a high temperature variation through the year. The precipitation in Princeton is spread out throughout the year. Precipitation events are often characterized by an exponential distribution of both the duration and the total precipitation per event. The shortest precipitation events and the smallest precipitation totals are the most frequent, while the longer the precipitation event, the less likely it is to occur at any given point. By analyzing the precipitation that is measured from Professor Simons' Vaisala weather station on the top of Guyot Hall from 2017 to present day, I can first summarize the data that is being characterized, then start using this climatology to start predicting precipitation events based on other variables that are observed in 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 clearly one that belongs to the mid-latitudes, which is characterized by having four seasons that results in having a large variation in temperature throughout a year. In terms of the average calculated between 1981 and 2010, Princeton gets an average of 48.27 inches (1227 mm) of precipitation annually, and the precipitation distribution throughout the year is fairly even, with less precipitation in the winter (PRISM Project). 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 $^{\circ}C$ or higher Peel et al. (2007). Princeton having no dry season means that precipitation should be well spread throughout the year.

The weather station that I am getting my data from is a Vaisala weather transmitter WXT530 series. It measures the 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 (accessed Septemeber 26, 2020). By analyzing the precipitation that is measured from Professor Simons' Vaisala weather station on the top of Guyot Hall from 2017 to present day, I can first summarize the data that is being characterized, then start using this climatology to start predicting precipitation events based on other variables that are observed in the weather station.

Methods

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.

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, $e_i = 0$.

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, breaking the year down into seasons since different seasons will 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.

Figure 2 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. 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, I have used similar procedure in which the distribution of the duration of precipitation events for events E_j^1 are graphed and an exponential fit is used for events E_j^2 and correspondingly graphed.

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				
2018	392	-0.59								
2019	308	-0.54								
2020	408	-0.65								

Table 1: This is the co-efficients found from the yearly distribution of precipitation as well as the seasonl distribution of precipitation. Italics refer to values obtained using incomplete information. NaN means there was no information. Blanks mean have not calculated it yet.

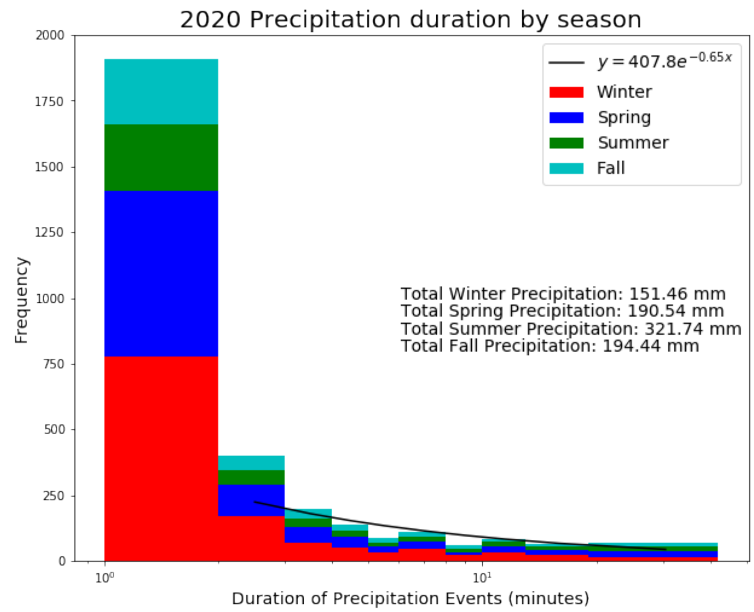

Figure 1: Distribution of precipitation duration in 2020, with a breakdown by seasons. 2020 is also incomplete as collection still ongoing. The layout is as in Figure 2.

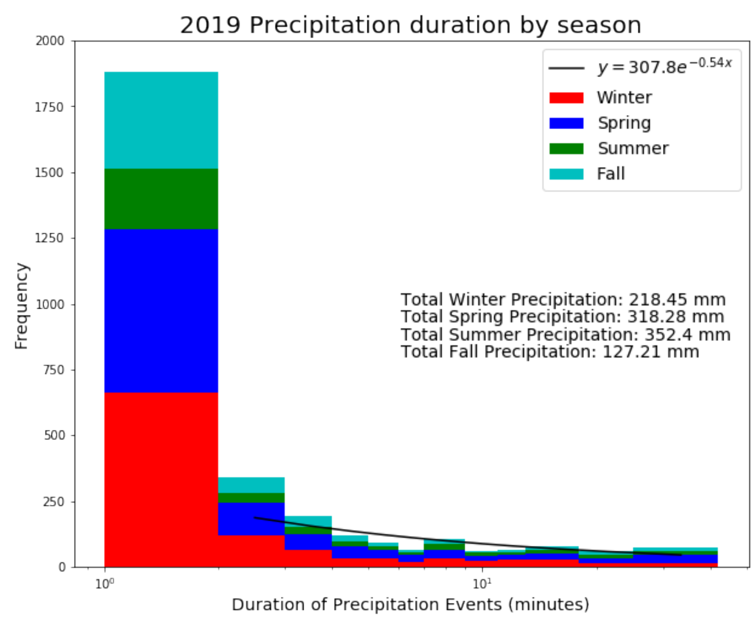

Figure 2: Distribution of precipitation duration in 2019, with a breakdown by seasons. The minimum duration, $\tau = 1$ min, and the maximum duration over the data set is 368 min, but the horizontal axis was limited to the 98th percentile of the durations, 42 min, for clarity. Superposed is a least-squares fit of a line in log-log space of duration versus frequency for the entire year, excluding the 1–2 min bin, quoted as the equivalent exponential in this space.

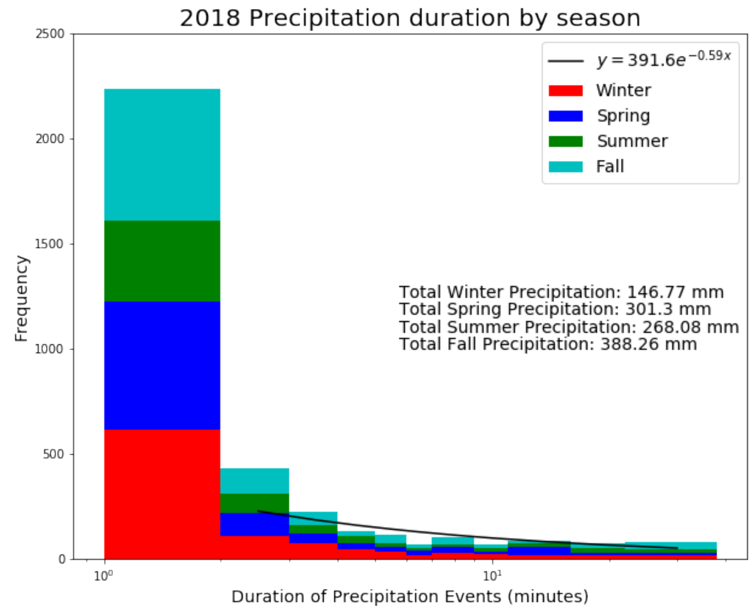

Figure 3: Distribution of precipitation duration in 2018, with a breakdown by seasons. The layout is as in Figure 2.

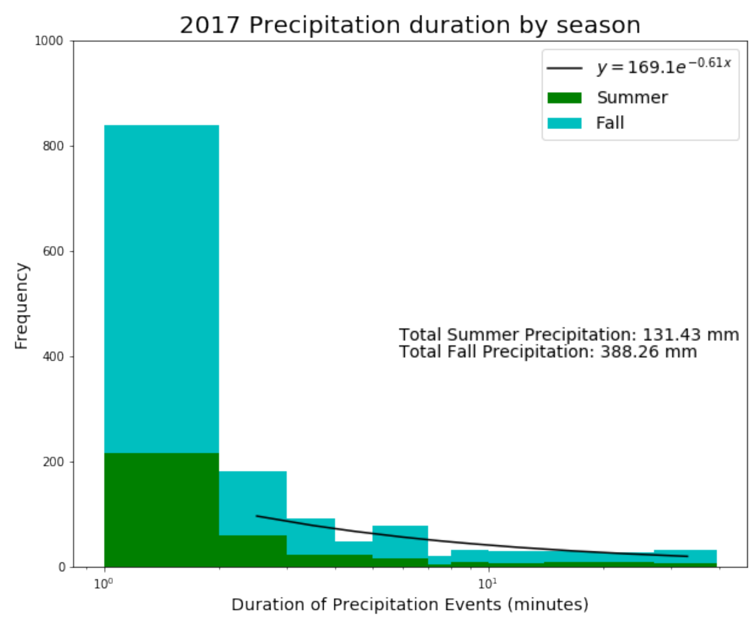

Figure 4: Distribution of precipitation duration in 2017, with a breakdown by seasons. The layout is as in Figure 2. Note that data collection began on July 16th of this year.

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 1/min). The Table 1 shows the coefficients and the exponential coefficients from the looking at the yearly frequency of precipitation duration. Based on this table, the β values are similar to each other with the exception of 2017, which only had partial data starting from 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.

Results

This section is blank for now. (Quite likely that the stuff in Methods may constitute results)

Discussion

This section is blank for now.

Conclusions

This section is blank for now.

References

Eagleson, P., 1978. Climate, soil, and vegatation: 2. the distribution of annual precipitation derived from observed storm sequences, *Water Resources Research*, **14**(5), 713–721.

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

Scher, S. & Messori, G., 2018. Predicting weather forecast uncertainty with machine learning, *Q J R Meteorol Soc.*, **144**, 2830–2841.

Vaisala, accessed Septemeber 26, 2020. Vaisala weather transmitter wxt530 series, `https://www.vaisala.com/sites/default/files/documents/WEA-MET-G-WXT530%20Datasheet__Low.pdf`.