Graphical Analysis

Analyzing the accuracy of satellite precipitation measurements with Rain Gauges across different geographic and topographic regions in US.

Ankit Ghanghas and Pin-Ching Li

Lyles School of Civil Engineering, Purdue University

Datasets Description

**GPM (Global Precipitation Measurement)**

Our study involves precipitation datasets from on ground measurement and satellite project. GPM is known as NASA Global Precipitation Measurement. The satellite measurement GPM is calibrated with an algorithm called Integrated Multi-satellitE Retrievals for GPM (IMERG). IMERG interpolate all satellite precipitation with microwave-calibrated infrared (IR) satellite estimates. There are three steps in IMERG system. Final run of IMERG is chosen to generate the research-level products.

The major variable of GPM dataset applied in our study is CalPrecipitation. The long name of CalPrecipitation is Daily accumulated precipitation (combined microwave-IR) estimate. The unit of CalPrecipitation is mm.

GHCN (Global Historical Climatology Network)

In the United States, rainfall data is collected by National Oceanic and Atmospheric Administration's (NOAA's) and is available for download from the National Centers for Environmental Information (NCEI). Our daily precipitation dataset is from a integrated database, which contains summaries across the globe: GHCN (Global Historical Climatology Network). NCEI is the database storing the GHCN and can be accessed by people with registered token.

GHCN dataset contains daily maximum and minimum temperature, precipitation, snowfall, and snow depth. Daily precipitation dataset is the variable we chose for analysis of satellite precipitation. The unit of GHCN Daily Precipitation is tenth of millimeter. There are hundreds of gage stations within a state, while some of them are poor in data coverage. For example, though there are 631 stations within Indiana or nearby Indiana, there are many stations without enough data coverage. As the histogram shown below, more than 300 stations suffer from insufficient data coverage (less than 95%).
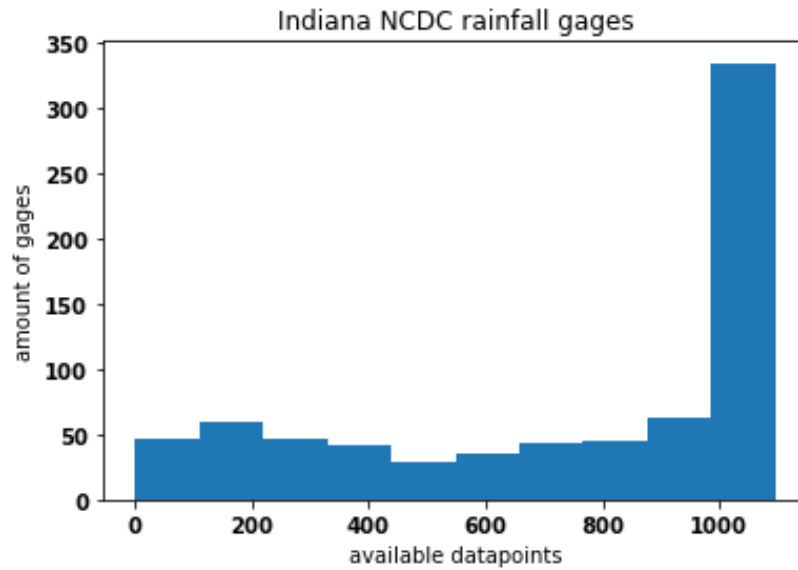
The scatter plot above shows the distribution of gages across Indiana state. All the gages are put on the basemap with shape of states in the USA. It is not necessary to keep gages with poor data coverage in our study. There are hundreds of gages with enough observations in this region. Therefore, the stations with data coverage less than 95% would be eliminated from this study. The histogram of available data points with better data coverage stations is shown below. Precipitation dataset from 196 sites in Indiana and 365 sites in Texas are applied to our study.
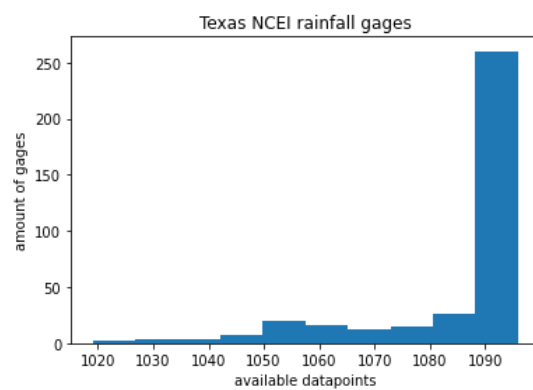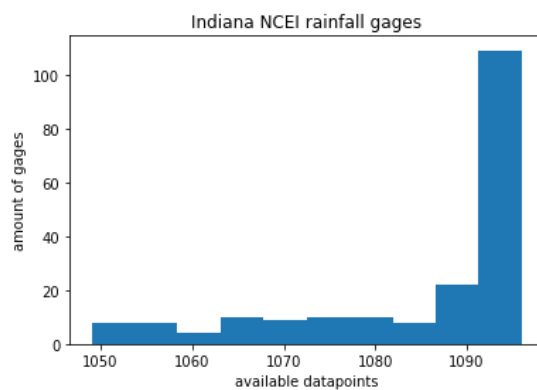
GHCN Daily Rainfall in Indiana
2016-2018
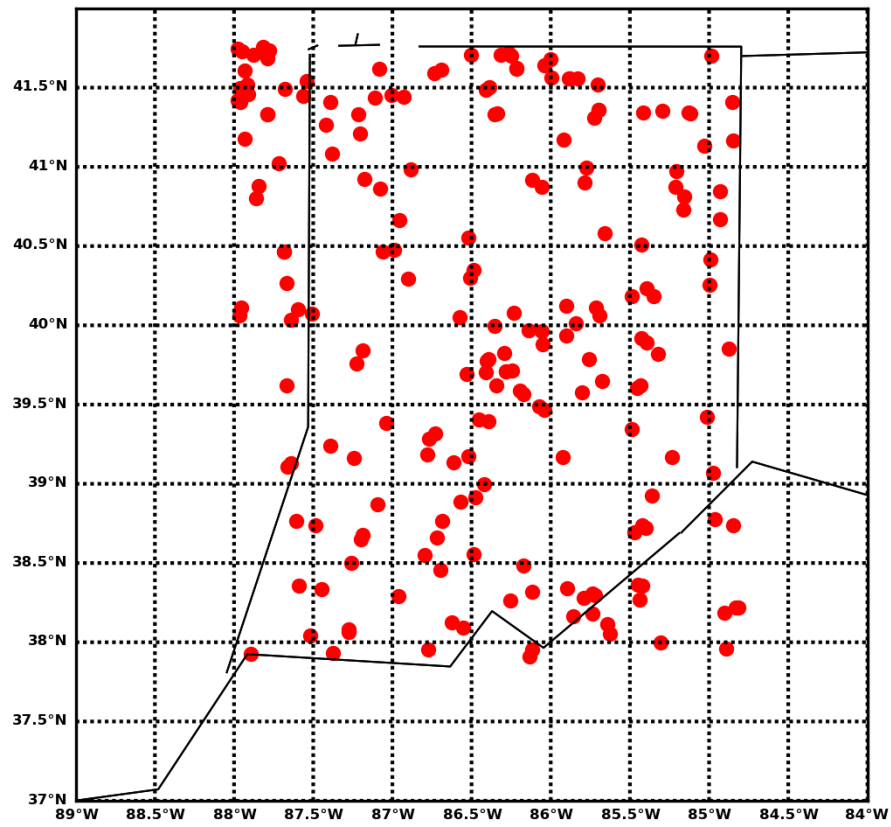
GHCN Daily Rainfall in Texas
2016-2018
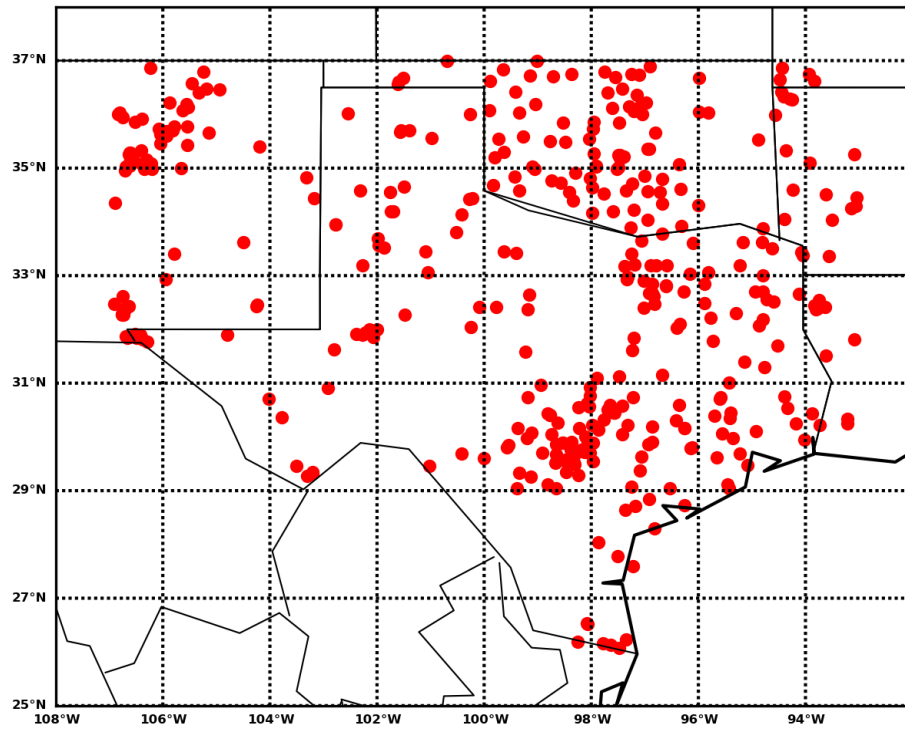
Graphical Analysis

GPM

Maximum daily precipitation of GPM in three years showed us the pattern of spatial distribution of extreme rainfall. The maximum precipitation in Indiana distributed without a significant order, while the maximum precipitation in Texas shows a cluster along the shore. This evidence supports our knowing about the weather in Texas. The inland part of Texas is desert. Therefore, the maximum rainfall is close to zero at Western part of Texas. The variation of maximum daily precipitation is larger in Texas than in Indiana. Indiana has uneven distribution of maximum precipitation. It can be the influence of topography or other physical influences, such as wind speed.
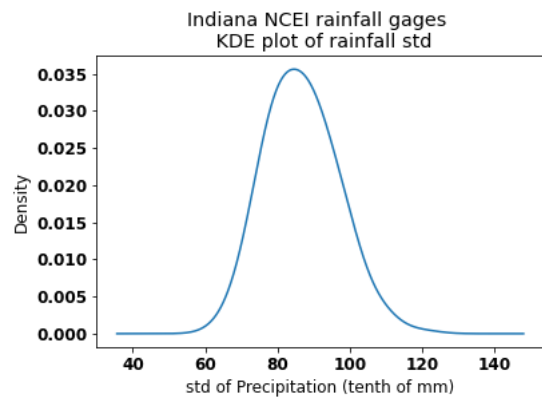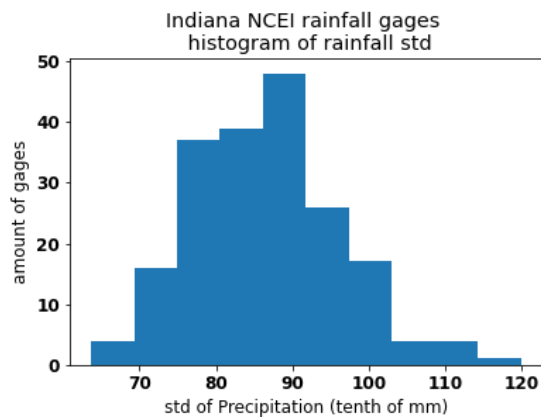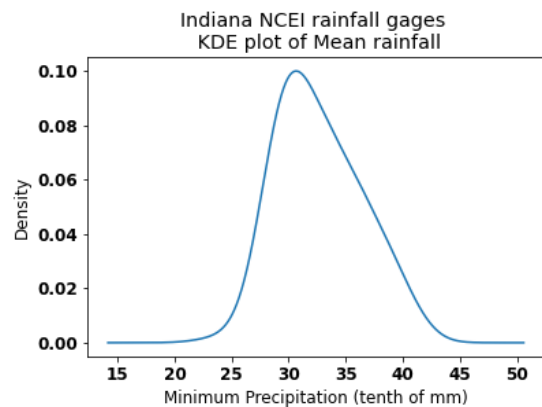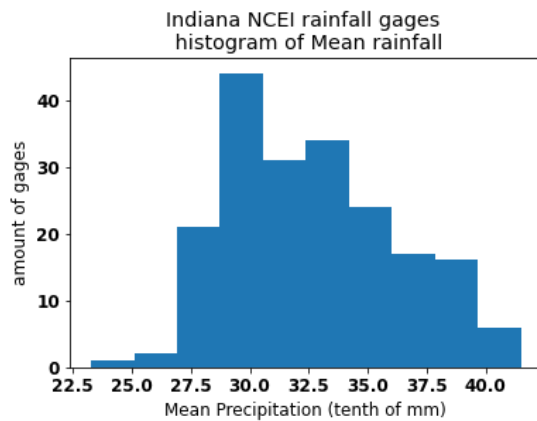

Maximum daily rainfall_Indiana
2016-2018

Maximum daily rainfall_Texas
2016-2018
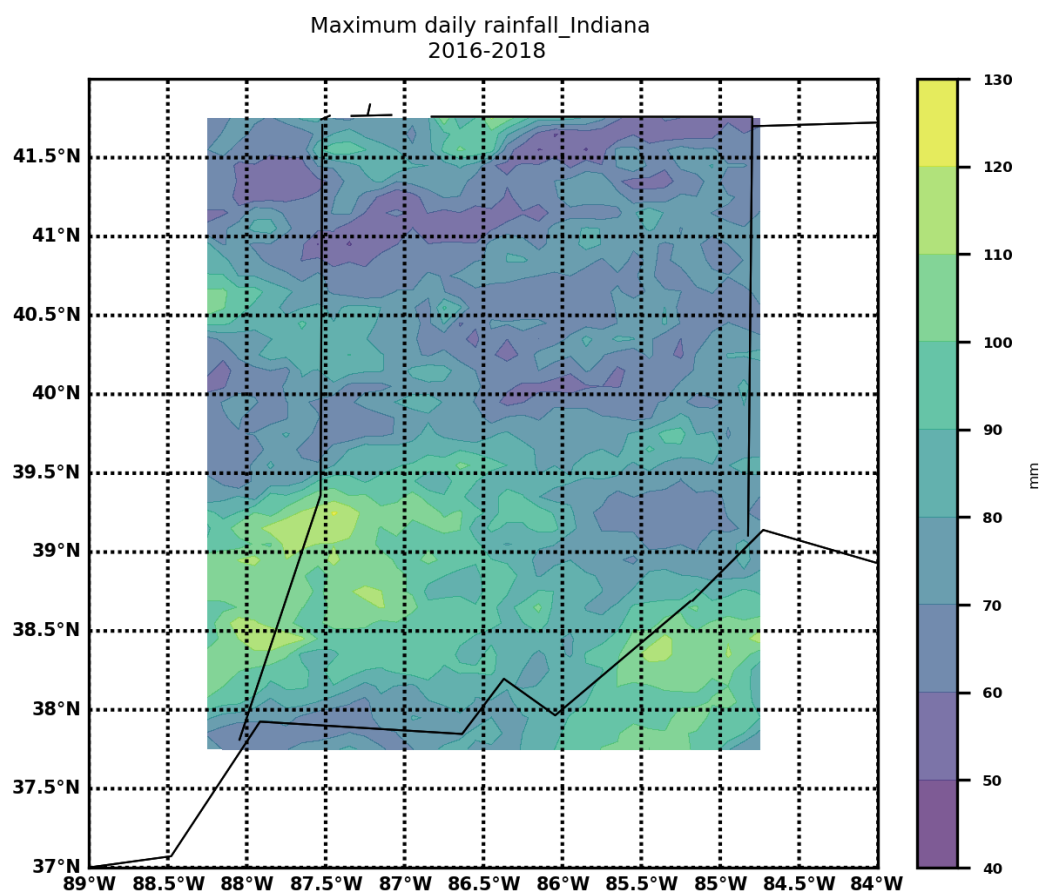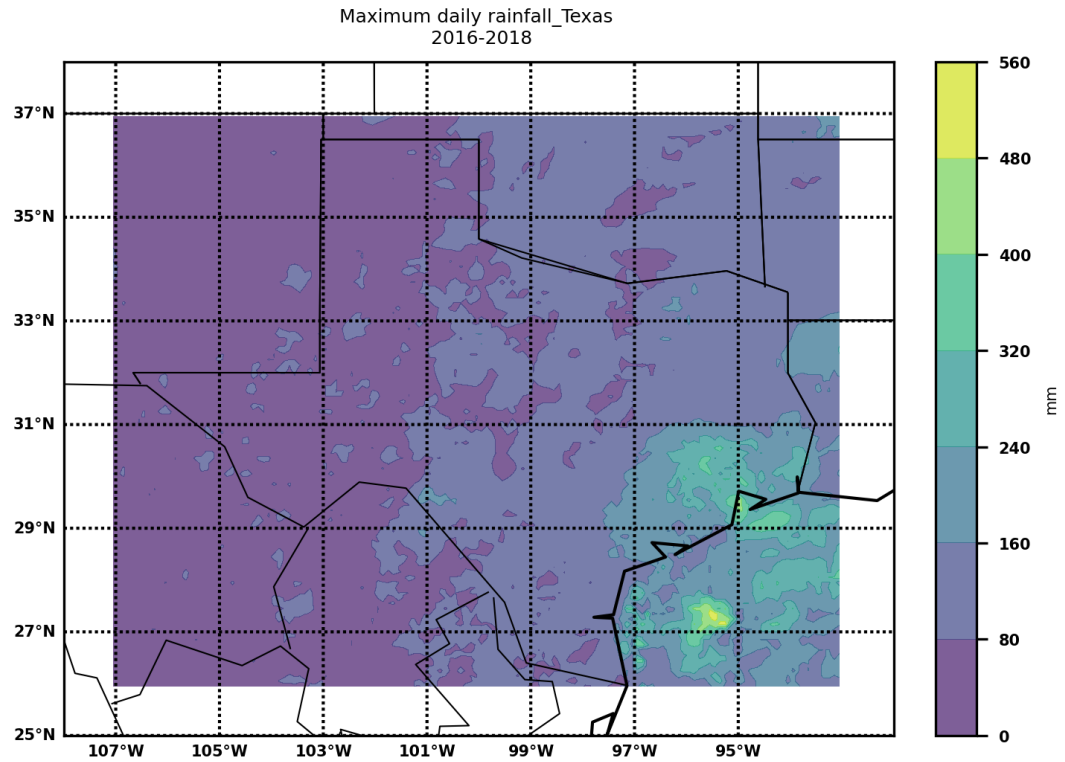
NCDC

Due to the huge study area we have, there could be considerable spatial difference between gages in the same state. Histograms and KDE plots with gaussian kernel (width=0.5) are drawn to find out how the precipitation influenced by spatial distribution within a state. If there is some especially wet or dry area, the skewness of precipitation distribution in KDE plot would be significantly presented. If the consistency of precipitation exists within this study area, the histogram and KDE tend to show a distribution of low skewness.

In Indiana, maximum rainfall seems to be consistent and the distribution shows like a bell shape. Mean rainfall has a distribution with positive skew. Standard deviation of rainfall seems to distribute in a bell shape without skewness. The distribution of standard deviation shows that precipitation in Indiana has no significant difference among each station.

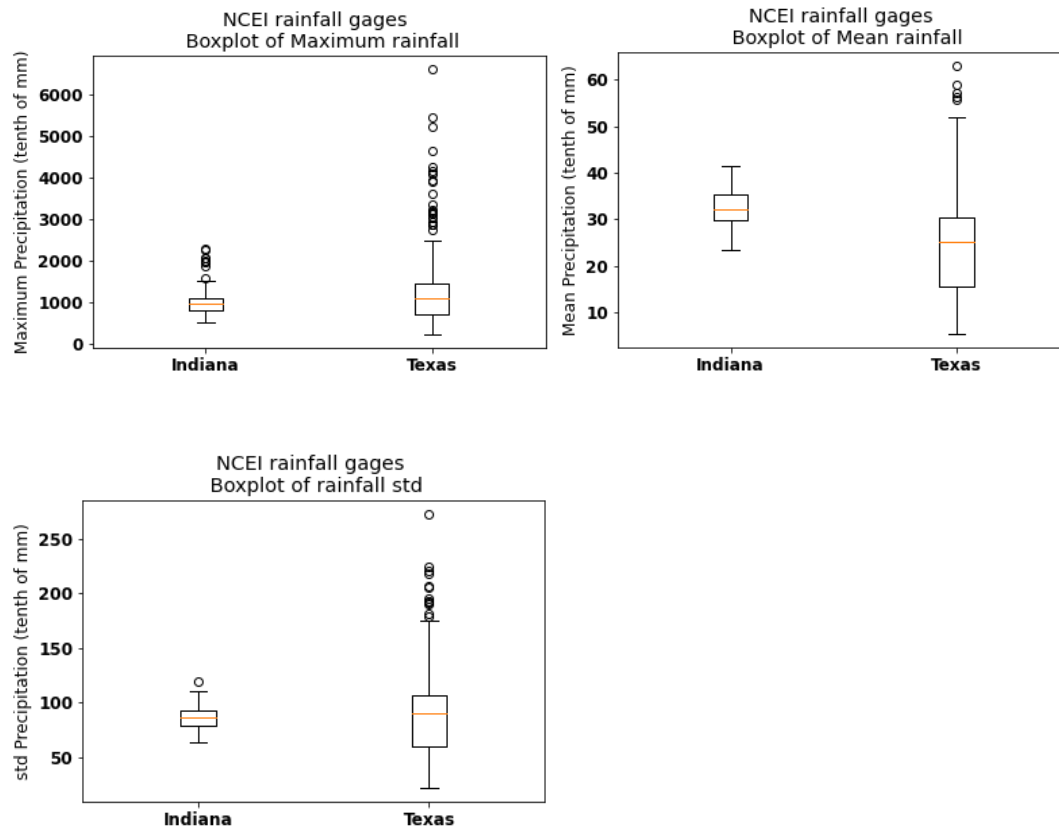Indiana NCEI rainfall gages — histogram of Maximum rainfall; Indiana NCEI rainfall gages KDE plot of Maximum rainfall; Indiana NCEI rainfall gages histogram of Mean rainfall; Indiana NCEI rainfall gages KDE plot of Mean rainfall; Indiana NCEI rainfall gages histogram of rainfall std; Indiana NCEI rainfall gages KDE plot of rainfall std.
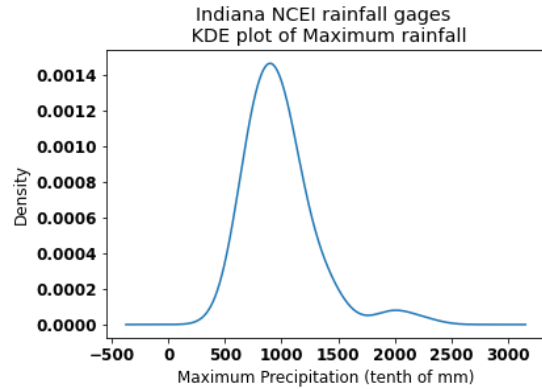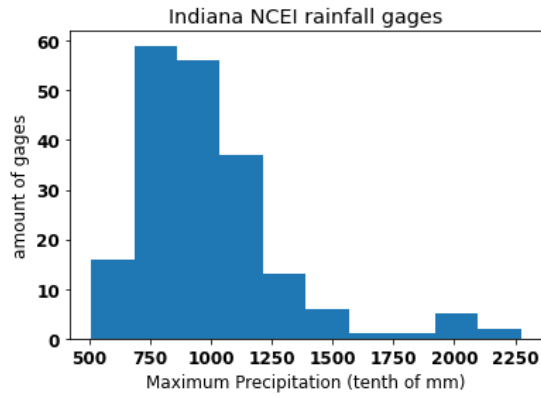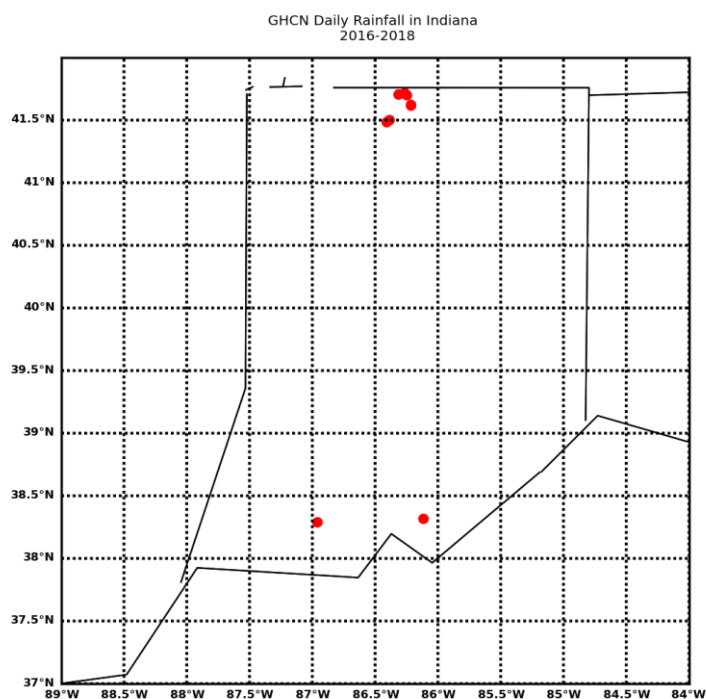
In Texas, maximum rainfall of stations seems to be consistent and the distribution shows like a bell shape. Mean rainfall has a distribution with slightly negative skewness. Standard deviation of rainfall seems to distribute in a bell shape with slightly negative skewness as well.
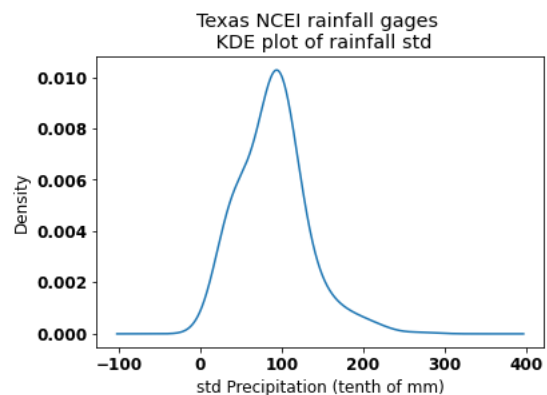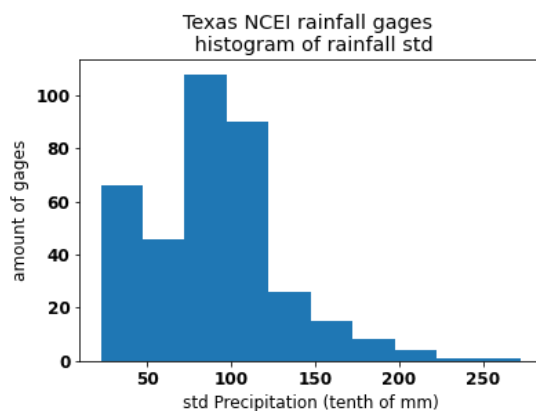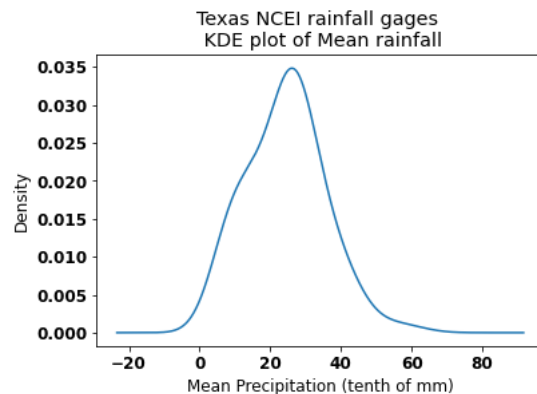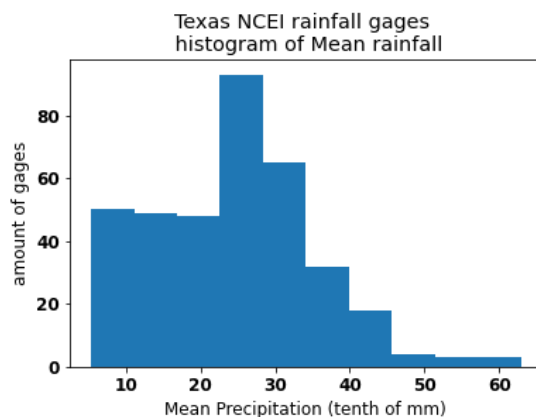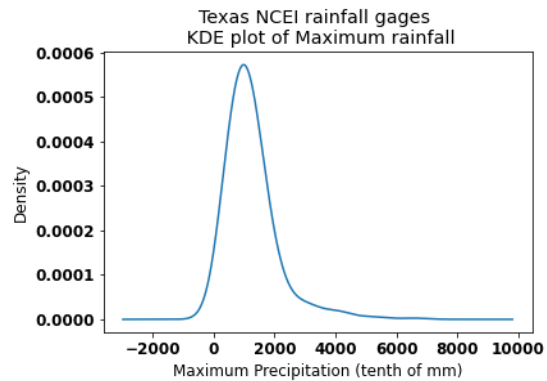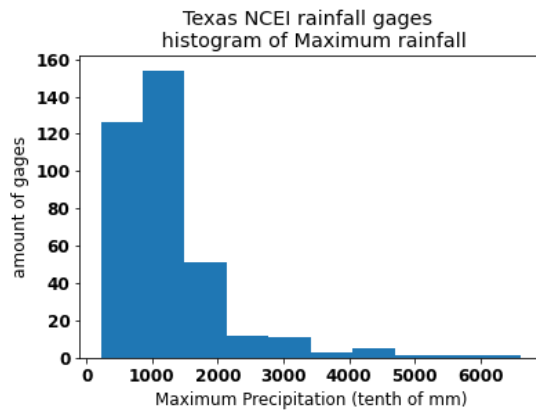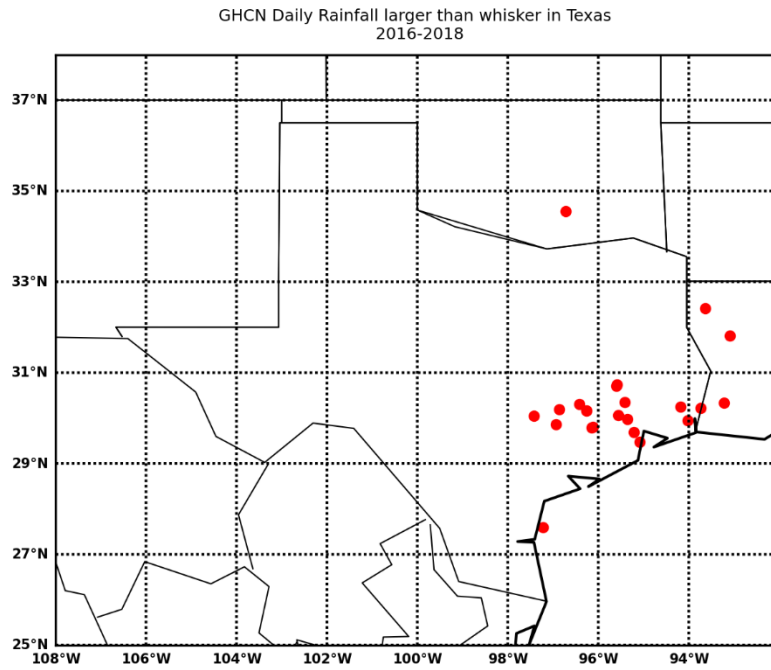
Box plots are used for checking the difference of max, mean, std of precipitation in different states. Indiana state tends to haven smaller box than Texas due to the less variation between values of precipitation variables. Texas has wider range of max, mean, and std precipitation. However, the median value of max, and std of precipitation is almost the same in these two states. Texas and Indiana have different size of shape with each other but share the similar central value of maximum precipitation distribution across the state. With the higher variance (or std), the shape of maximum precipitation of stations in Texas is wider, which indicates the IQR value (box shape) is larger.

NCEI rainfall gages
Boxplot of Maximum rainfall

NCEI rainfall gages
Boxplot of Mean rainfall

NCEI rainfall gages
Boxplot of rainfall std

Stations containing maximum precipitation above the whisker of box plot are investigated. The location of them are shown below. It is not surprising the stations with large maximum precipitation located at the near shore or lake place in TX and IN. These points fit the maximum precipitation pattern in GPM contour.

GHCN Daily Rainfall in Indiana
2016-2018

GHCN Daily Rainfall larger than whisker in Texas
2016-2018

For standard deviation part, stations with higher standard deviation of precipitation are nearby the shore in Texas. The distance between station and shoreline seems to significantly influences the rainfall pattern in Texas.

GHCN Daily Rainfall with large std in Texas
2016-2018
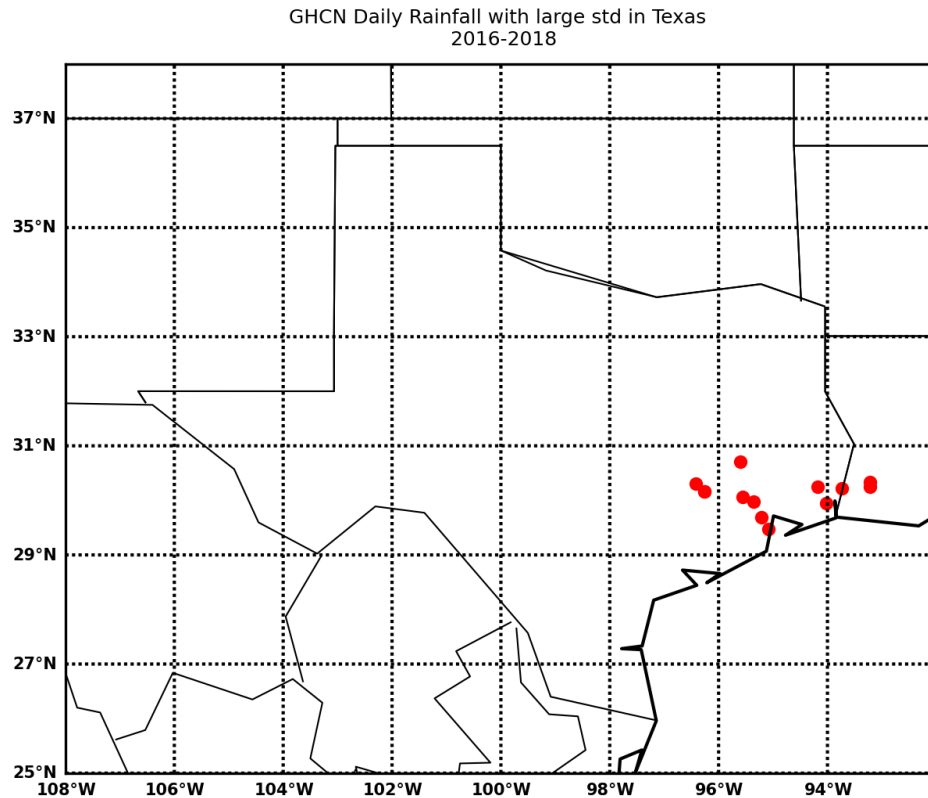
Maximum Precipitation of stations in IN and TX does not fit normal distribution well. The existing skewness and outliers make maximum precipitation across Texas having worse fitting with the 45-degree line (the line of normal distribution). PP plot of maximum precipitation across TX shows that the distribution of maximum precipitation across TX is not significantly close to normal distribution. PP plot shows a closer result between observation points and normal distribution.