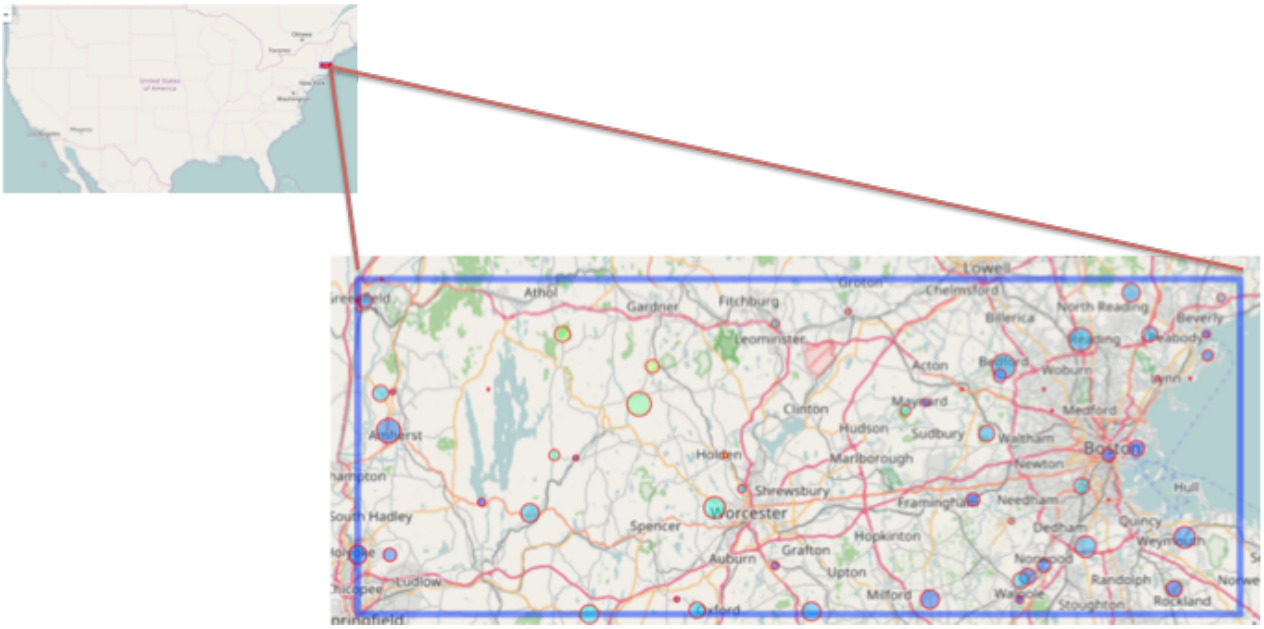
Precipitation and Snow Data Analysis in Noth East US

Backgroup

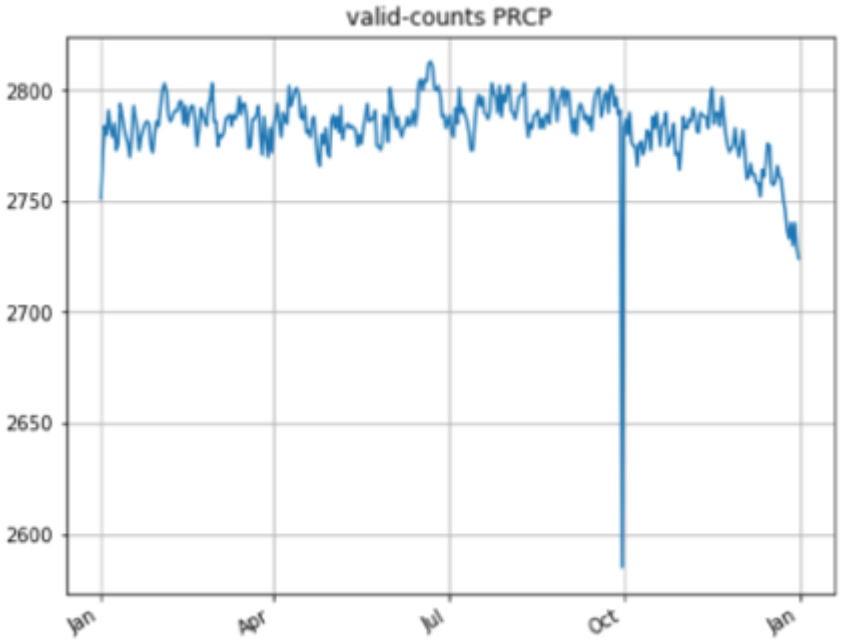
Three different climate variables are analyzed in this report, and these variables are including snow (SNOW), snow depth (SNWD), and precipitaiton (PRCP). Our research region located in north east US (please see the location map). In this report, these data analysis including yearly distribution of available data amount, seasonal variation of climate variables, principle components analysis are all based on the measurment of USS0005K26S station (location: 39.5667, -105.8), and some data analysis are based the measurement of all the station in this region. The yearly data of USS0005K26S station are the mean of 7 years data (1999, 2000, 2001, 2005, 2006, 2007, 2009). Our objective of this report is to analyze the major controlling factor of SNOW, SNWD and PRCP.

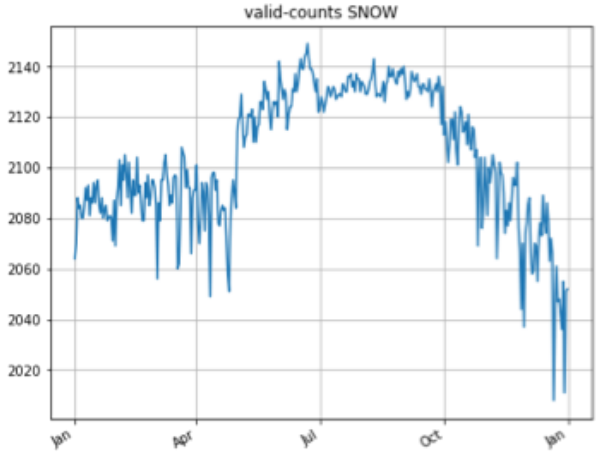


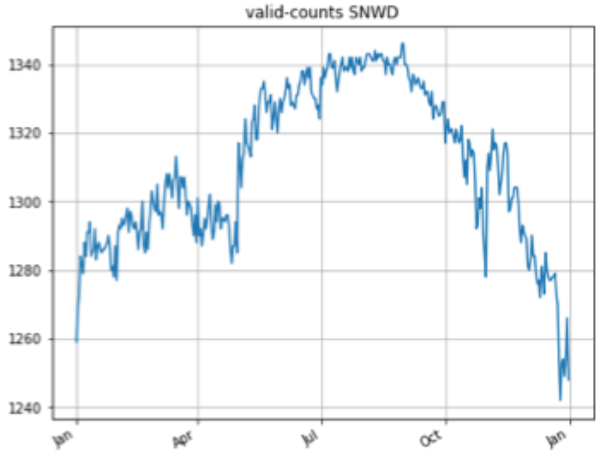
In the map, the dot is the location of observational station and dot size represent the amount of available data from this station.

Available data amount of each climate variables

Available data is always the most important issue for data analysis. Firstly, we made a yearly distribution about the amount of availabe data for each climate variables, then to understand the data quality. The statistical plots illustrated that in the whole year, the available data amount of PRCP (above) are aound of 2800, the available data amount of SNOW (below left) are around of 2100, and as to SNWD (below right) the available data are around 1300.

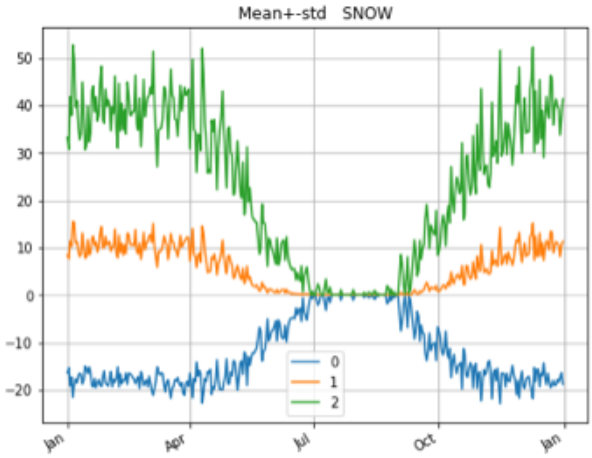


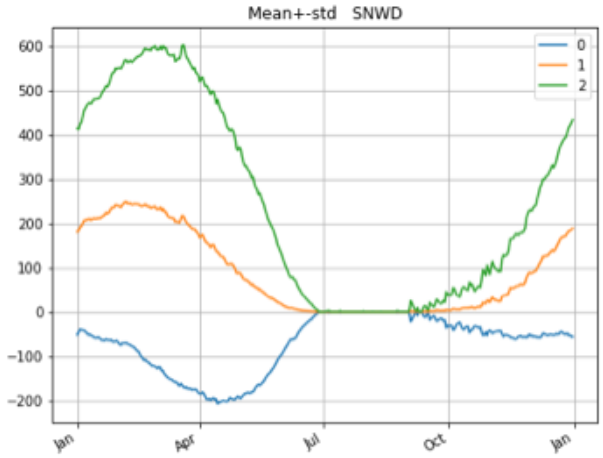


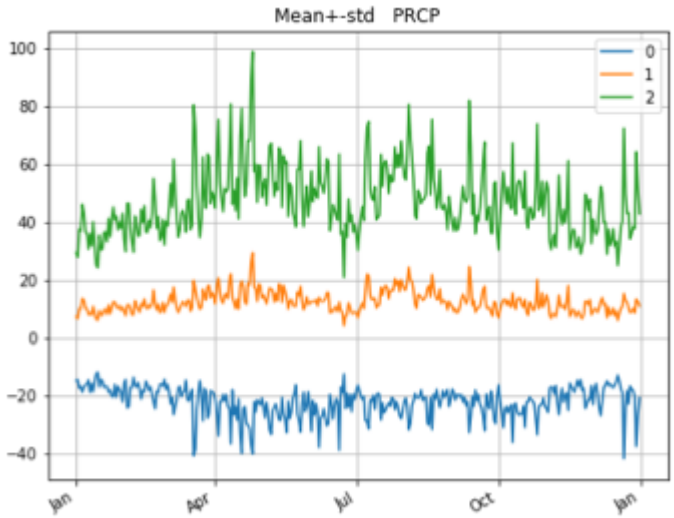


Seasonal variation of climate variables

The following Figs showed us that both SNOW (above left) and SNWD (above right) have obvious seasonal variation. For the figures, the orange line is the mean of variable, the green line is the mean plus standard deviation, and the blue line is the mean minus standard deviation. In summer, SNOW and SNWD are close to zero, and in winter, both of them attain maximum. and there is no much change around the year. and SNWD is highly correlated with SNOW. However, precipitation (below) has no obvious seaonal pattern, and the pattern of precipitation is pretty steable and there is no much change around the year.





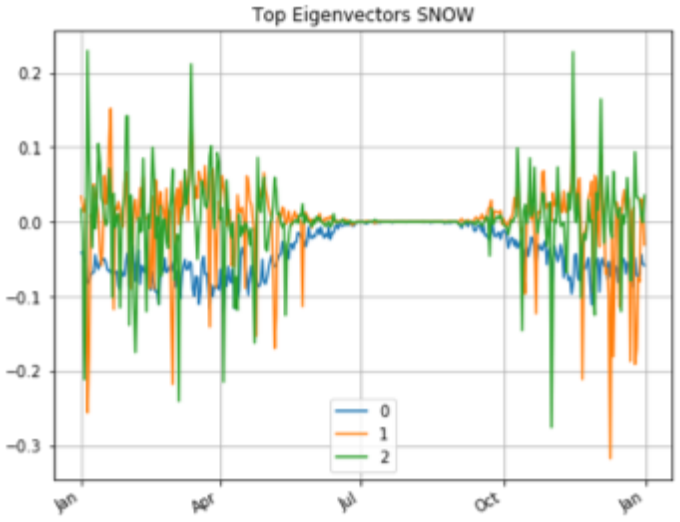


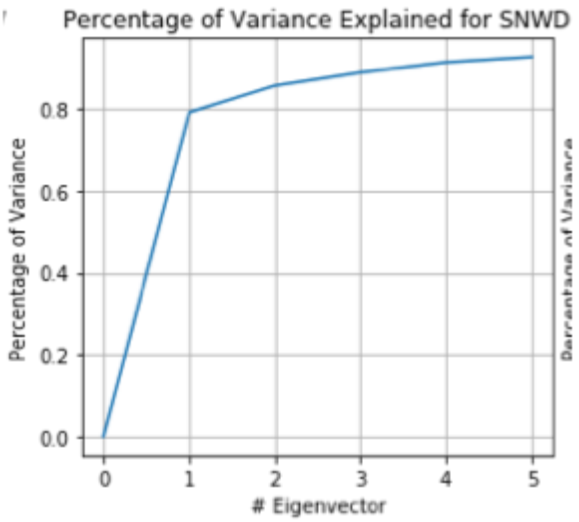
Priciple components analysis of climate variables

Principal component analysis (PCA) is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components (or sometimes, principal modes of variation). The number of principal components is less than or equal to the smaller of the number of original variables or the number of observations. This transformation is defined in such a way that the first principal component has the largest possible variance (that is, accounts for as much of the variability in the data as possible), and each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to the preceding components. The resulting vectors are an uncorrelated orthogonal basis set. PCA is sensitive to the relative scaling of the original variables (Wikipedia).

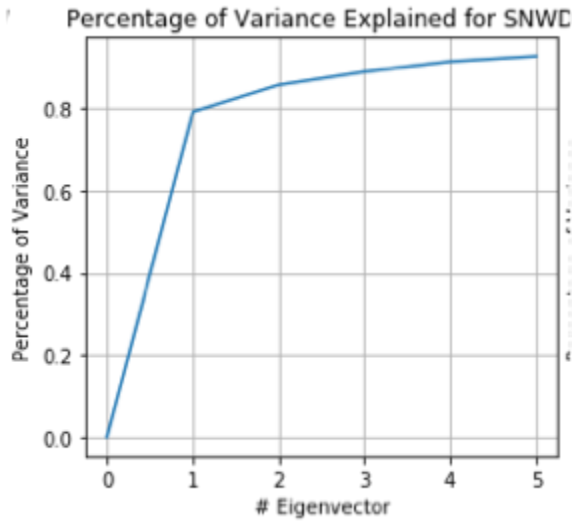
We all know that the snow is controlled by so many different factors, like air temperature, air humidity, and air movements ect. Here we presented the first three major components, the blue line is the first principal component, the orange line is the second principal component (left), and the green line is the third principal component. And according to the pattern of first principal component, we can infer that the temperature is major factor in controlling the amount of snow, and

the temprature contribute around 80% of the variance of SNOW. However, the temperature and snow have the negative correlation, it means that the amount of snow decrease with the increase of temperature.

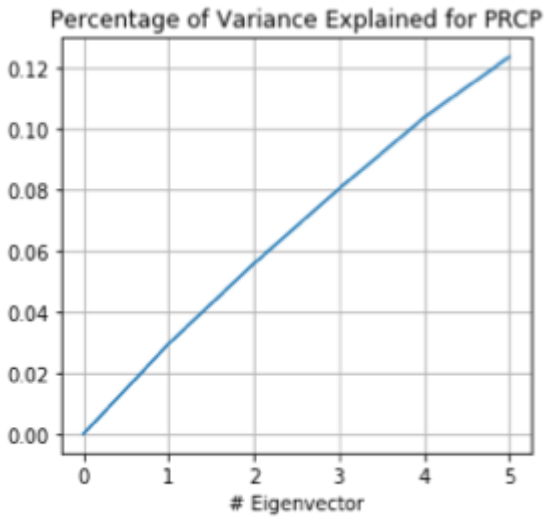




We analyzed the principal components of SNWD. The first three principal components are presented in the figure (left), and the first principal component can explain 80% variance of SNWD. It is obvious that the SNWD is mainly controlled by the amount of SNOW, and our results reflect this phenomena exactly. The blue line is the first principle component and the seasonal pattern is similar with SNOW.



For the precipitation, it is an very complicated phenomina and it is controlled by so many different factors. When we did the principle component analysis, we can not find any dominant factor controlling precipitation.



Conclusion

In summary, principal component analysis can work well in analyze snow and snow depth, however, because the precipitation is a very complicated phenomina, we can not use principle component analysis to discover the major factor in controlling precipitation.