THE NORTH ATLANTIC OSCILLATION HAS AN EFFECT ON WESTERN TURKEY'S PRECIPITATION PATTERN

Merih BOZBURA
bozbura@itu.edu.tr, No: 601161005

Abstract. The North Atlantic Oscillation (NAO) is a large-scale natural climate variability that has important impacts on the weather and climate of the North Atlantic region and surrounding continents, especially Europe. Strong positive phase is disposed to be connected with lower than normal temperatures over southern Europe and the Middle East and it is disposed to be associated with lower than normal precipitation over southern and central Europe. Unlike strong positive phase of NAO , negative phase is disposed to be connected with higher than normal temperature over southern Europe and the Middle East and it is disposed to be associated with higher than normal precipitation over southern and central Europe. In this study, relationship between NAO and annual precipitation of Turkey will be conducted. Exploratory data analysis, linear regression, and Pearson correlation are applied to precipitation and NAO indexes datasets to assess whether there is a relationship between NAO and precipitation or not. Only, Marmara region rejects the null hypothesis. Also, results from Pearson correlation matches with results from the regression analysis. In the light of these informations, precipitation is affected by North Atlantic Oscillation can be said.

1 Introduction

The North Atlantic Oscillation (NAO) is a large-scale natural climate variability that has important impacts on the weather and climate of the North Atlantic region and surrounding continents, especially Europe. NAO is a difference in sea-level pressure of Azores High and Subpolar Low. The positive phase of NAO indicates higher than normal heights and pressure over Europe and the negative phase indicates lower than normal heights and pressure over Europe. Strong positive phase is disposed to be connected with lower than normal temperatures over southern Europe and the Middle East and it is disposed to be associated with lower than normal precipitation over southern and central Europe as it is shown in Figure 1. Unlike strong positive phase of NAO , negative phase is disposed to be connected with higher than normal temperature over southern Europe and the Middle East and it is disposed to be associated with higher than normal precipitation over southern and central Europe as it is shown in Figure 2 (Climate Prediction Center Internet Team, 2012). Therefore, NAO is expected to have an impact on Turkey. Weather and climate conditions are controlled by NAO in Mediterranean basin along with Turkey. The geographical and temporal variations and anomalies in the precipitation for Turkey is strongly associated with strong NAO phases Türkeş and Erlat (2003). In this study, relationship between NAO and annual precipitation of Turkey will be conducted. Exploratory data analysis, linear regression, and Pearson correlation are applied to precipitation and NAO indexes datasets to assess whether there is a relationship between NAO and precipitation or not.

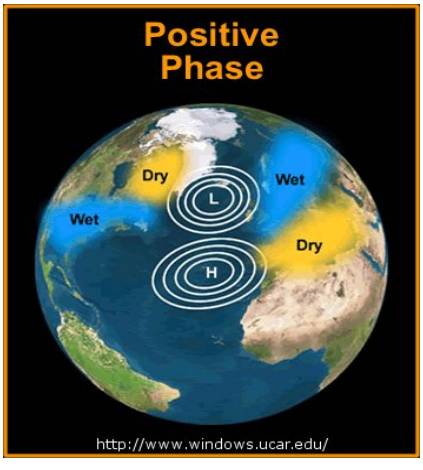

Figure 1. Positive phase of NAO (Global Patterns, n.d).

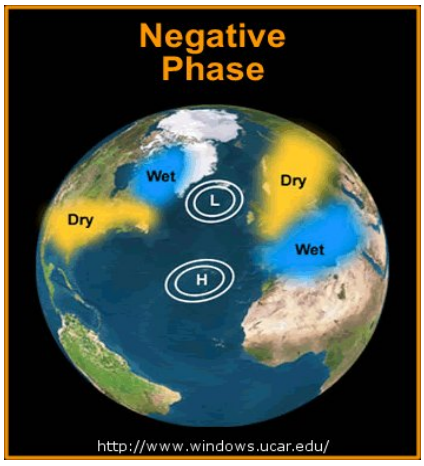

Figure 2. Negative phase of NAO (Global Patterns, n.d).

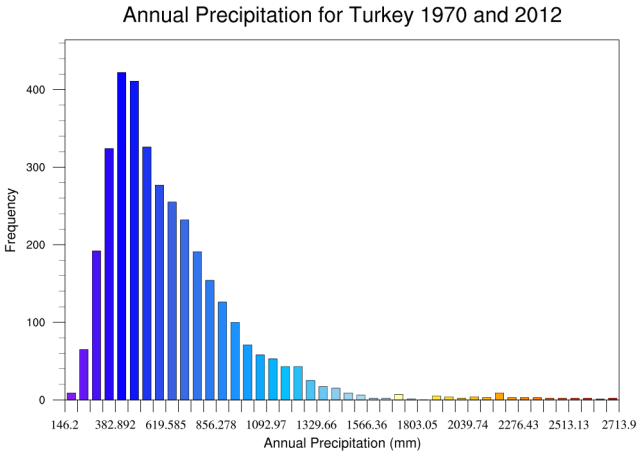

Figure 3. Histogram of Turkey.

2 Data and Method

Hypothesis is determined as there is a relationship between precipitation and NAO for this paper. The North Atlantic Oscillation Indexes and annual precipitation of Turkey datasets are used in this paper. Annual precipitation is an observation data and it includes all provinces of Turkey. These datasets are between 1970 and 2012. Firstly, exploratory data analysis are applied to understand fundamentals features of precipitation data. Secondly, linear regression is applied to see whether there is a relationship between NAO and precipitation or not. Finally, Pearson correlation are applied to the relationship. This study is conducted with NCL.

3 Exploratory Data Analysis

Exploratory data analysis is an approach to analyse a data. Exploratory data analysis supports to find out essential properties of the data. So, in this section histograms of Turkey and three regions and a barplot of average precipitation of Turkey will be examined.

Histogram of Turkey has right-skewed distribution as it is stated in the Figure 3. A few larger values bring the mean upwards. Marmara region is the closest region to normal distribution compared to the other regions. According to Figure 4, Figure 5, and Figure 6 other regions also have right-skewed distribution.

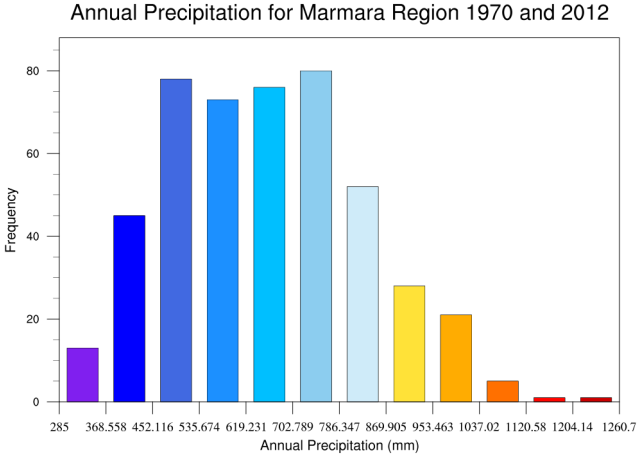

Figure 4. Histogram of Marmara region.

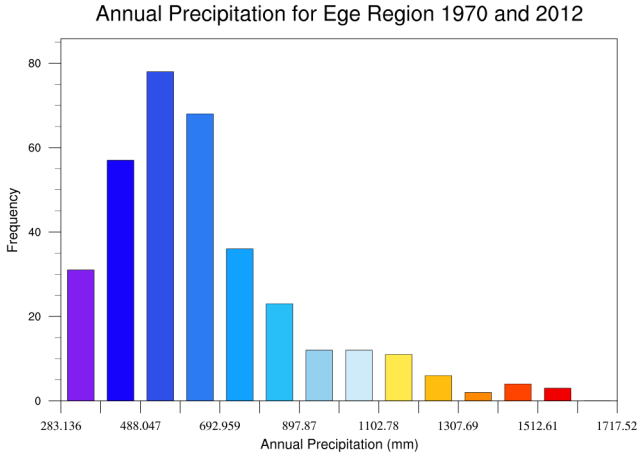

Figure 5. Histogram of Ege region.

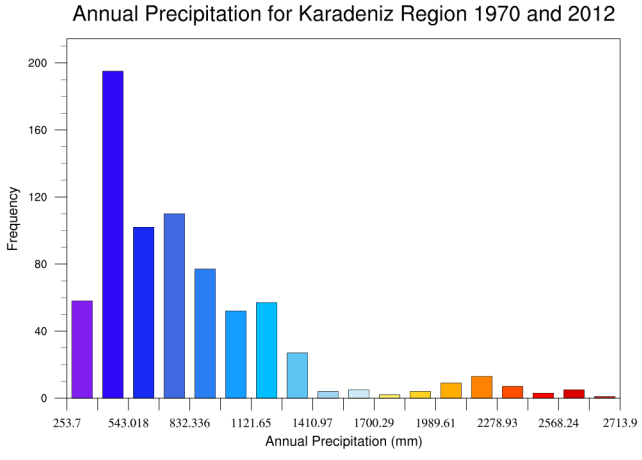

Figure 6. Histogram of Karadeniz region.

There are eighty one provinces in Turkey and as it is expected, 53^{rd} province which is Rize has the highest precipitation in Turkey between 1970 and 2012 as it is stated in Figure 7. Also, Iğdır has the lowest precipitation. Average precipitation of Turkey is around 630 mm.

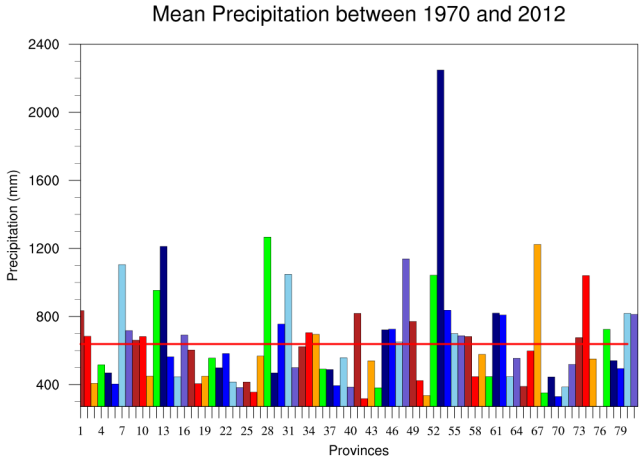

Figure 7. Average barplot for all provinces.

4 Linear Regression

Turkey has seven regions and linear regression is applied between each region and NAO index. P-values of Marmara, Ege, and Karadeniz regions are respectively 0.05056401, 0.1887776, 0.05120632. All regions do not reject the null hypothesis which is there is no relationship between precipitation and NAO. However, the region closest to rejecting the null hypothesis is Marmara. So, the summary of the linear model between Marmara region and NAO index is shown in Figure 8. Marmara region is the independent variable and NAO index is the dependent variable in this linear model. The equation of the model is $NAOindex = 7.98642 - 0.001753206 * Canakkale - 0.00451112 * Edirne - 0.007670015 * Istanbul + 0.00214023 * Tekirdag + 0.008831669 * Yalova + 0.002830141 * Kirklareli - 0.001584583 * Balikesir - 0.004196721 * Bilecik - 0.003053695 * Bursa - 0.00273433 * Kocaeli - 0.0005422276 * Sakarya.$ In addition to the Marmara region, Karadeniz and Ege regions are chosen to examine and linear regression is applied between each provinces of these three regions and NAO indexes. Provinces with significant p-values are obtained since they will be compared principle component analysis results to understand that pattern on precipitation belongs to NAO or not. Provinces with significant p-values are Balikesir, Bilecik, Bursa, Bolu, Canakkale, Edirne, Istanbul, Kocaeli, Tekirdag, Kirklareli, Amasya, Kastamonu, Rize, Izmir, Kutahya, and Manisa.

```
Variable: reg_marmara
Type: float
Total Size: 48 bytes
            12 values
Number of Dimensions: 1
Dimensions and sizes:   [12]
Coordinates:
Number Of Attributes: 31
  _FillValue :  9.96921e+36
  long_name :   multiple regression coefficients
  model :       Yest = b(0) + b(1)*X1 + b(2)*X2 + ...+b(M)*XM
  N :   43
  NP :  11
  M :   12
  bstd :        <ARRAY of 12 elements>
  SST : 162.7658
  SSE : 93.1009
  SSR : 69.66499
  MST : 3.875377
  MSE : 3.003255
  MSE_dof :     31
  MSR : 6.333181
  RSE : 1.73299
  RSE_dof :     30
  F :   2.108773
  F_dof :       ( 11, 31 )
  F_pval :      0.05056401
  r2 :  0.4280075
  r :   0.6542228
  r2a : 0.2250424
  fuv : 0.5719925
  Yest :        <ARRAY of 43 elements>
  Yavg :        0.3441861
  Ystd :        1.968598
  Xavg :        <ARRAY of 11 elements>
  Xstd :        <ARRAY of 11 elements>
  stderr :      <ARRAY of 12 elements>
  tval :        <ARRAY of 12 elements>
  pval :        <ARRAY of 12 elements>
(0)     7.98642
(1)     -0.001584583
(2)     -0.004196721
(3)     -0.003053695
(4)     -0.001753206
(5)     -0.00451112
(6)     -0.007670015
(7)     -0.00273433
(8)     -0.0005422276
(9)     0.00214023
(10)    0.008831669
(11)    0.002830141
```

Figure 8. The summary of the linear model between Marmara region and NAO index.

The R^2 of the model is 0.4280075 and it is not too small. So, the model explains about 43% of the variability of the dependent variable which is NAO index.

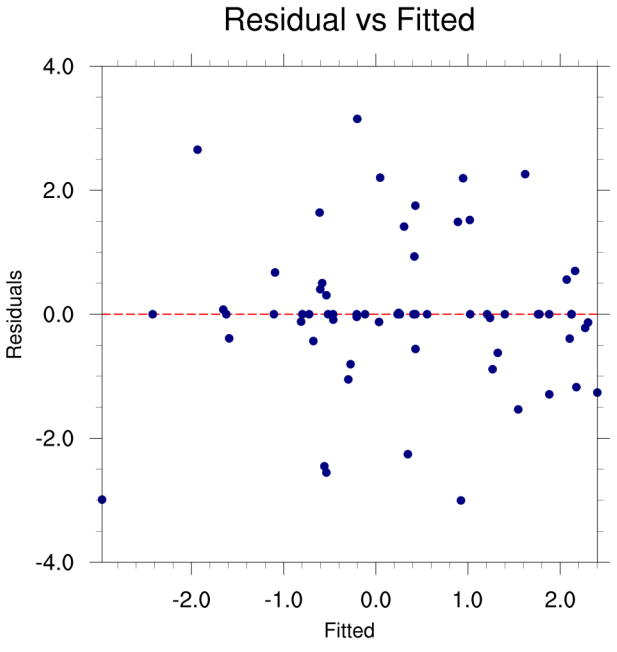

Figure 9. Residual vs Fitted plot.

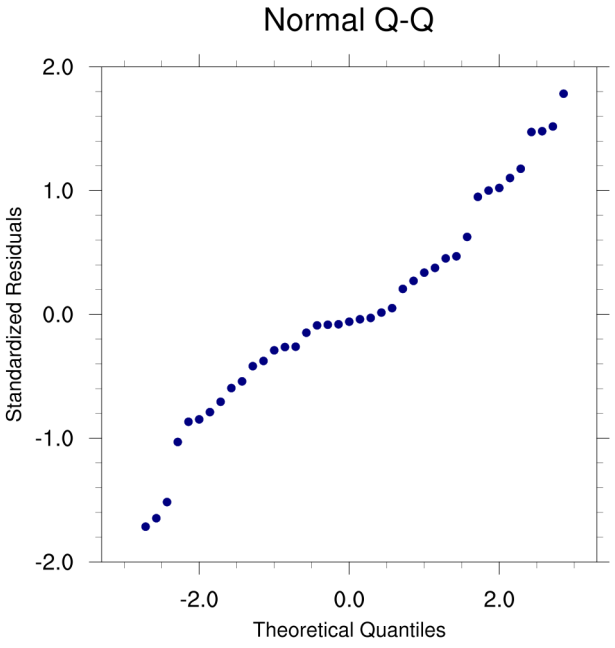

Figure 10. Normal Q-Q plot.

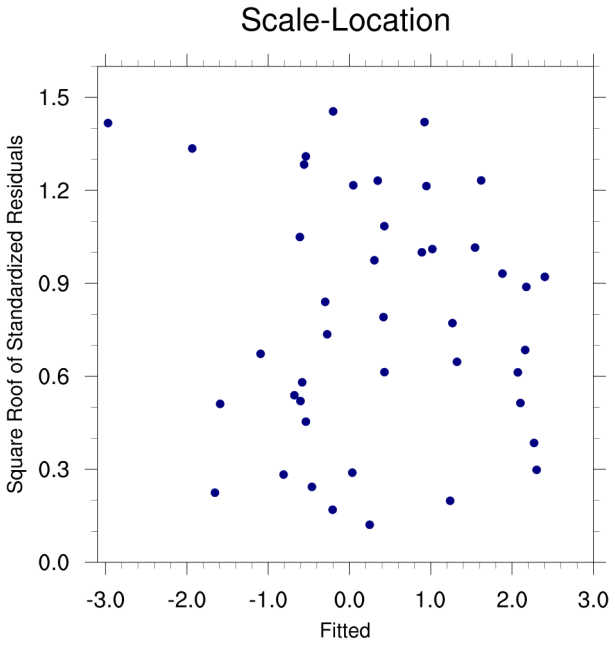

Figure 11. Scale Location plot.

Residual is the difference between fitted and actual dependent point. Fitted point is a predicted point by the model.There should be no discernible pattern around zero residual versus fitted plot. As it is stated in Figure 9, residuals and fitted values are almost randomly distributed around the zero line. Therefore, the model is not very good, but it fits the data. According to Figure 10, residuals nearly follow the nor-

mal distribution. Also, scale-location residuals are randomly distributed and there is no discernible pattern as it is shown in Figure 11.Cook's distance is used to find dominant points in independent variables. These points are far from the other points. In this case, 29^{th} and 40^{th} points are little away from the other points and 41^{st} point are far from the other points as it is shown in Figure 12. If the Cook's distance is greater than 0.5, that point can be influencial, so it is worthy to examine (Identifying Influential Data Points, n.d). Cook's distance of 41^{st} point is greater than 0.5. If exceeding point equals about two times of average of data, it should be examined (Jacoby, n.d). 41^{st} point is not greater or equal to the two times of average of data. So, it is not worthy for investigating.

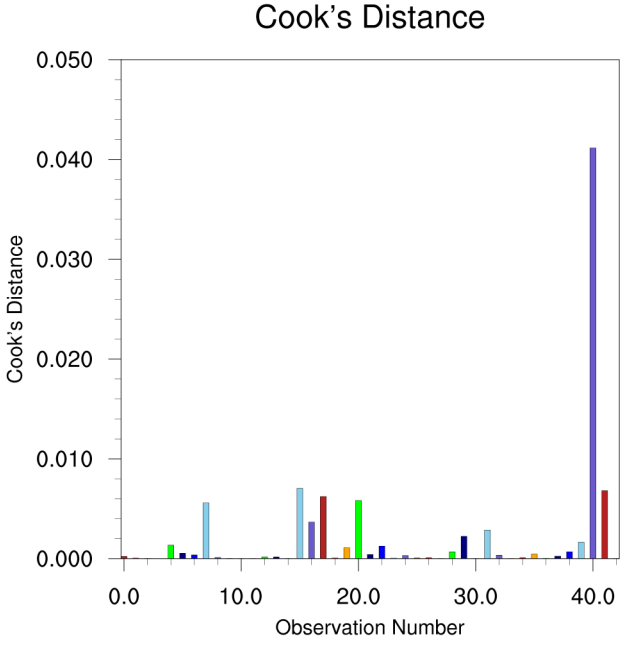

Figure 12. Cook's distance plot.

Figure 13 represents the provinces which have significant p-value with red circle and the provinces which have not significant p-value with dark green circles. As it is seen, the provinces which have significant p-values come together around northwest of Turkey.

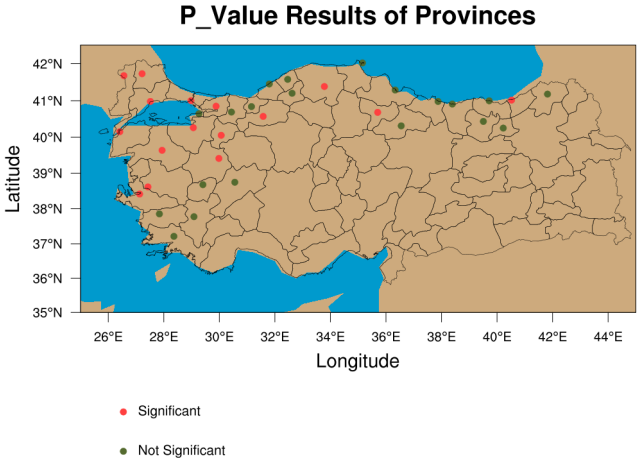

Figure 13. P-value of examined provinces.

5 Pearson Correlation

Correlation is a technique for investigating the relationship between two quantitative, continuous variables. Pearson's correlation coefficient (r) is a measure of the strength of the association between the two variables. Thus, Pearson correlation coefficients is calculated to examine the relationship between Nao and the precipitation.

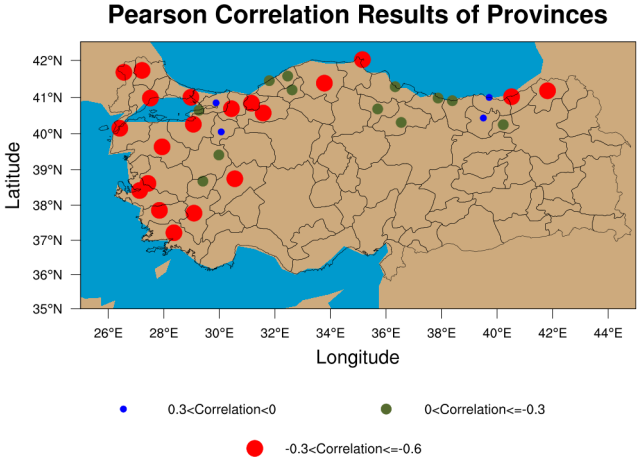

Figure 14. Pearson Correlation Results of Provinces

As it is seen in Figure 14, several provinces have higher correlation than the other which are Balikesir, Bilecik, Bursa, Bolu, Canakkale, Edirne, Istanbul, Kocaeli, Tekirdag, Kirklareli, Amasya, Kastamonu, Rize, Izmir, Kutahya, and Manisa. They are represented with red circles. At the same time, these provinces have significant p-values in the regression analysis.

6 Conclusions

Exploratory data analysis is applied to understand the datasets and linear regression analysis is applied to find out whether there is a relationship or not between precipitation and NAO index datasets. 16 out of 36 provinces reject the null hypothesis is found out and to investigate NAO's effect, Ege and Karadeniz regions are chosen in addition to Marmara region. Finally, Pearson correlation are applied to three region which contains thirty six provinces. As it is seen in section 4 , the provinces with significant p-values and correlations results matches. Also, majority of the provinces with significant p-value is from Marmara region. In the light of these informations, precipitation is affected by North Atlantic Oscillation can be said.

References

Abdi, H. and Williams, L. J.: Principal component analysis, Wiley Interdisciplinary Reviews: Computational Statistics, 2, 433–459, doi:10.1002/wics.101, 2010.

Climate Prediction Center Internet Team, .: North Atlantic Oscillation (NAO), http://www.cpc.ncep.noaa.gov/data/teledoc/nao.shtml, 2012.

Global Patterns, .: Global Patterns Arctic and North Atlantic Oscillations (AO and NAO), http://climate.ncsu.edu/climate/patterns/NAO.html, n.d.

Habing: Exploratory Factor Analysis, http://people.stat.sc.edu/habing/courses/530EFA.pdf, 2003.

Identifying Influential Data Points, .: Identifying Influential Data Points, University Lecture, https://onlinecourses.science.psu.edu/stat501/node/340, n.d.

Jacoby, .: Regression III:Advanced Methods, University Lecture, http://polisci.msu.edu/jacoby/icpsr/regress3/lectures/week3/11.Outliers.pdf, n.d.

Türkeş, M. and Erlat, E.: Precipitation changes and variability in Turkey linked to the North Atlantic oscillation during the period 1930-2000, International Journal of Climatology, 23, 1771–1796, doi:10.1002/joc.962, 2003.