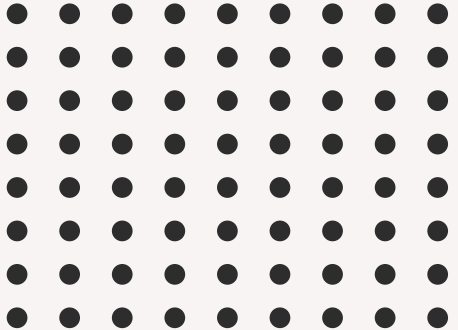
Exploring Causal Inference in Earth Science



Mentor: Lelys Bravo de Guenni
Mentees: Tanaiya Sunter, Samuel Park

What did we do?

Week 1~2

Read The Book of Why / Correlation vs. Causation / What is Causal Diagram

Week 3

Imported the Data into R and identify predictors and response

Week 4

Made graphs of Autocorrelation and Cross-Correlation / Filtering time series

Week 5

Empirical Dynamic Models / Convergent Cross Mapping

Causal Inference

Basically questioning. Why?

It is the strategies and pathways we take to draw causal conclusions based on data

Correlation and Causation

Correlation is when two or more variables relate to each other. But one variable doesn't cause or changes the other.

We often hear that "correlation does not imply causation"

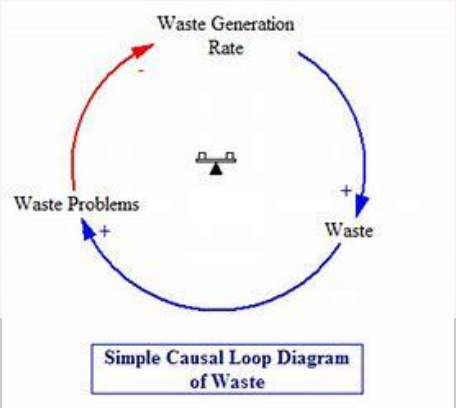
Example: Typically, on average the more educated a person is, the greater their income will be

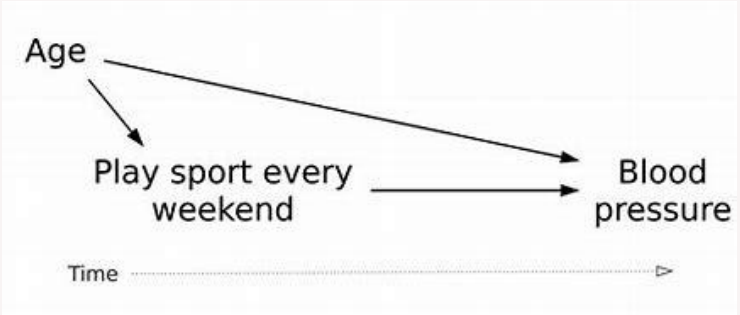
Causation is when one variable cause another variable to change.
There is a causal relationship between these variables

Example: if I leave my chocolate outside in the heat it will melt. The heat caused my chocolate to melt

Causal Diagram

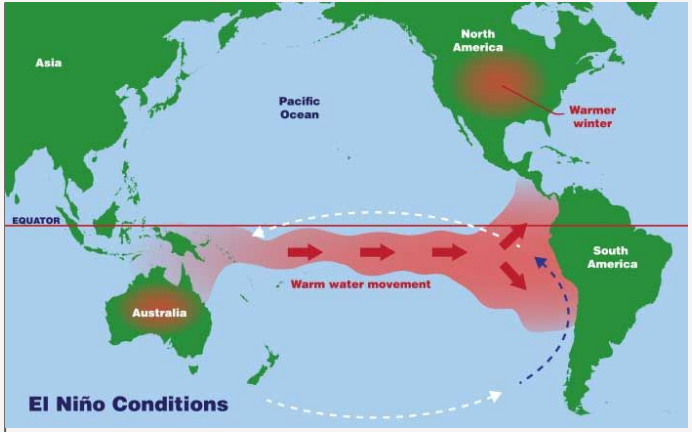
A simple bow and arrow graph
Putting facts into a diagram
Visually displays causal relationships/relationships between variables
Shows how different variables are related and the outcome
Can show the relationship of cause and effect





El Niño

- El Niño is a complex climate phenomenon that happens inconsistently in the equatorial Pacific Ocean
- The normal atmospheric circulation patterns are disturbed during an El Niño event, changing the patterns of precipitation
- El Niño may cause more rain to fall and fewer drought conditions in some areas (southwestern United States and western South America)



Data Explanation

Data Collection
(Data is given by our mentor)

SPEI was measured locally which is from Urbana, Illinois from 1950 to 2021

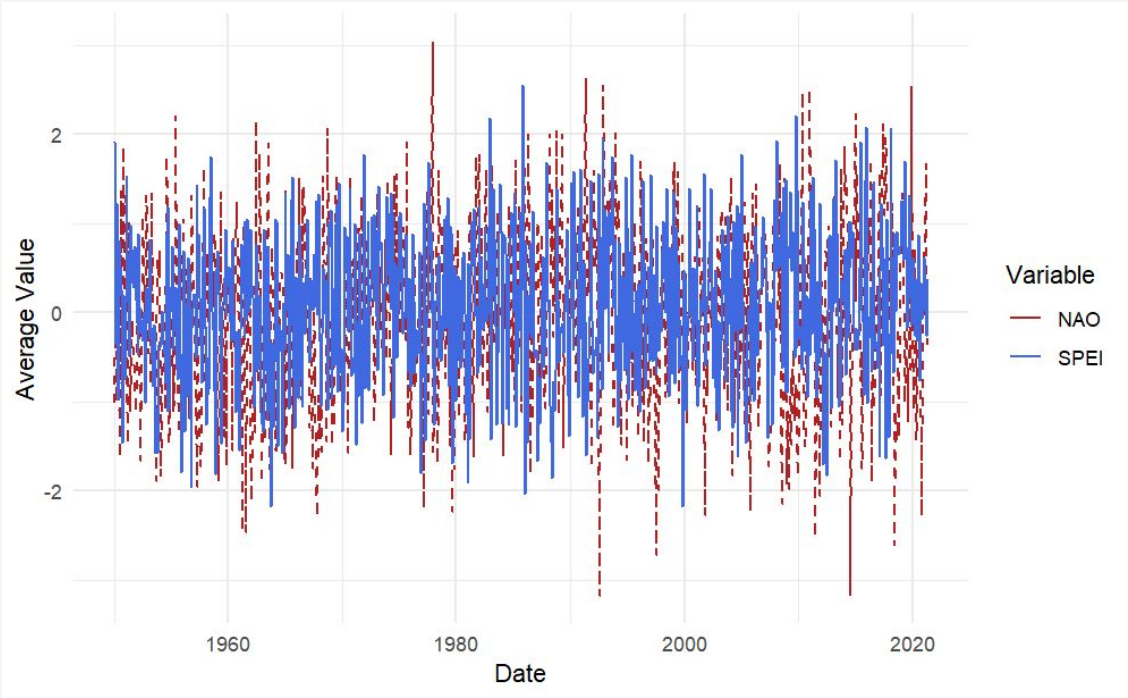
But other drivers were measured globally from 1950 to 2021

Monthly data

Variables

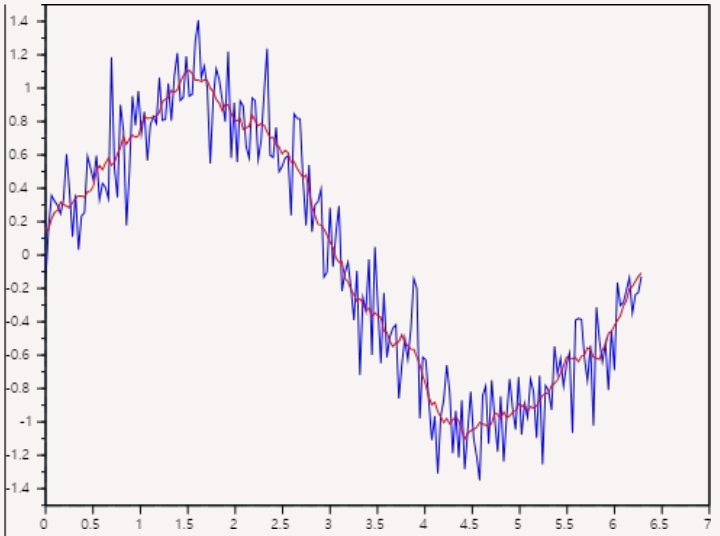
NAO: North Atlantic Oscillation Index
PDSI: Palmer Drought Severity Index
AMO: Atlantic multi-decadal oscillation Index
ONI: Oceanic Nino Index
PDO: Pacific Decadal Oscillation Index
SPEI: Drought Index

Graph of NAO and SPEI

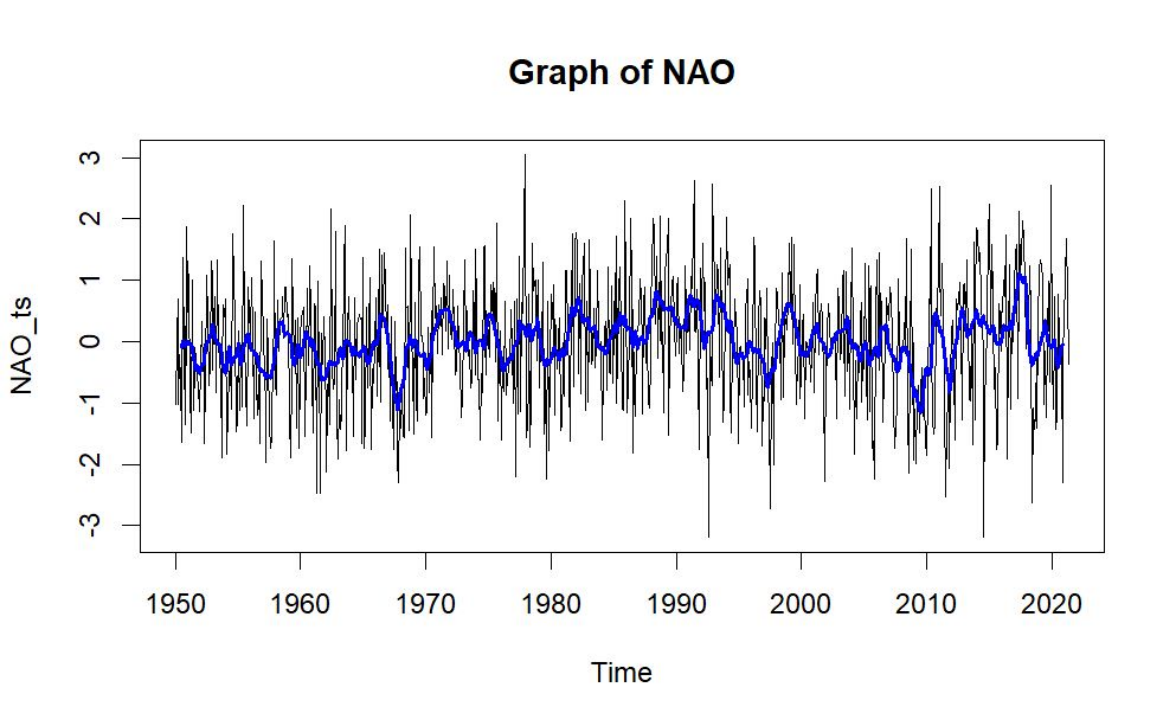


Moving Average Filtering

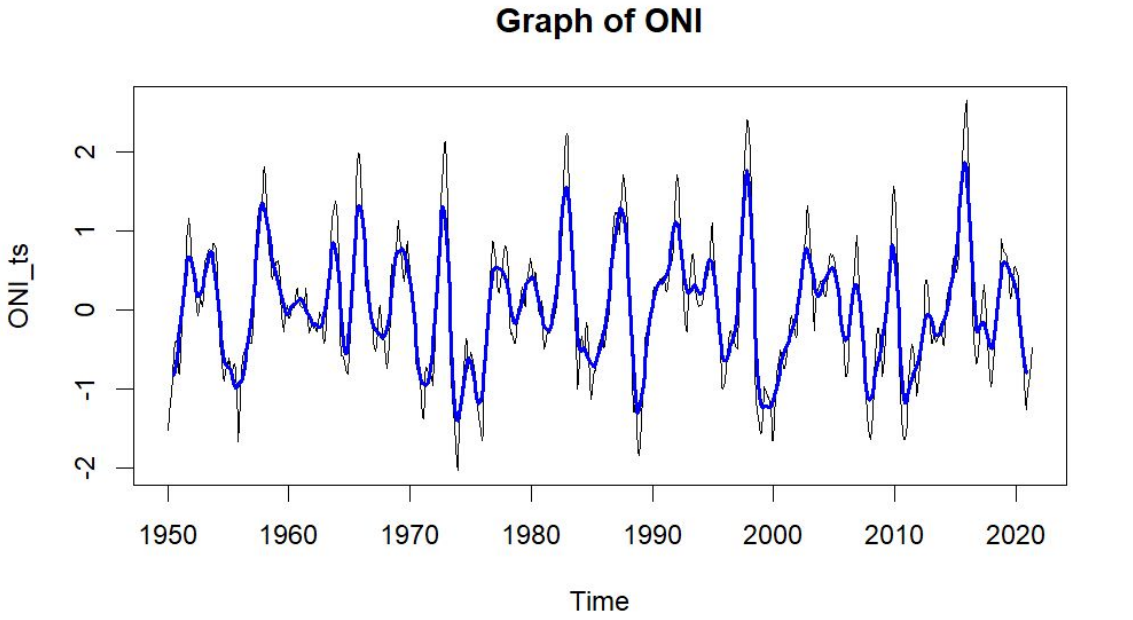
- The method of moving average is done by calculating the average by adding the value right before and after each value.



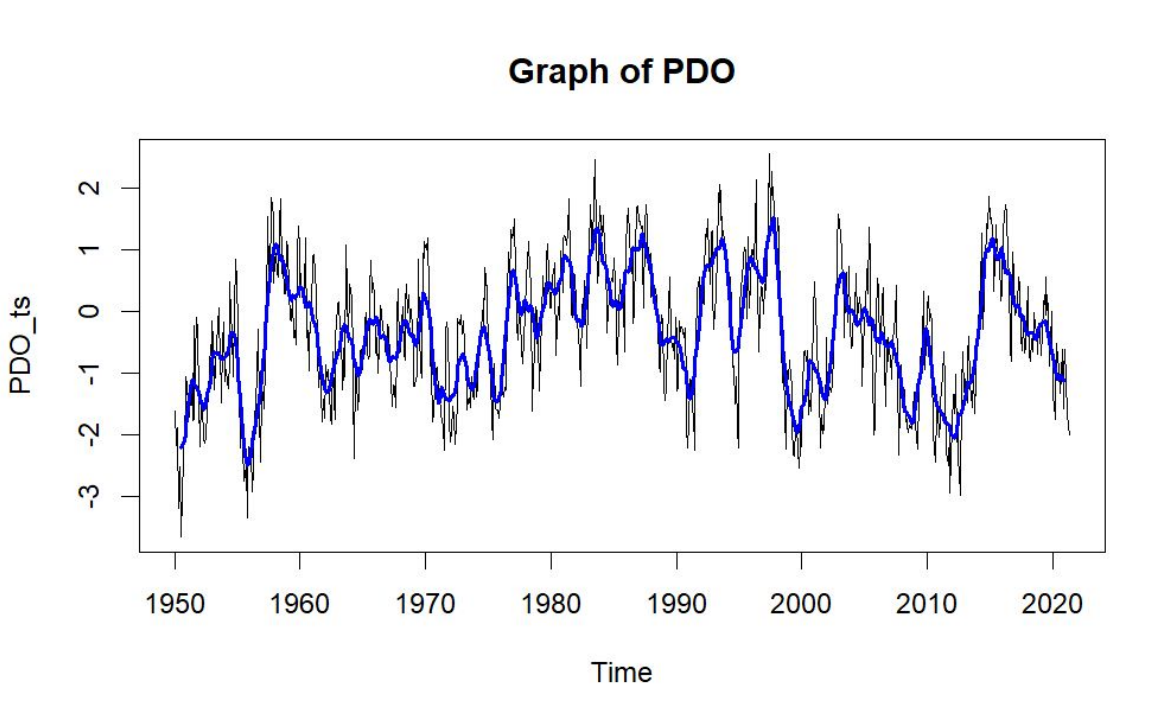
Smoothed Graphs



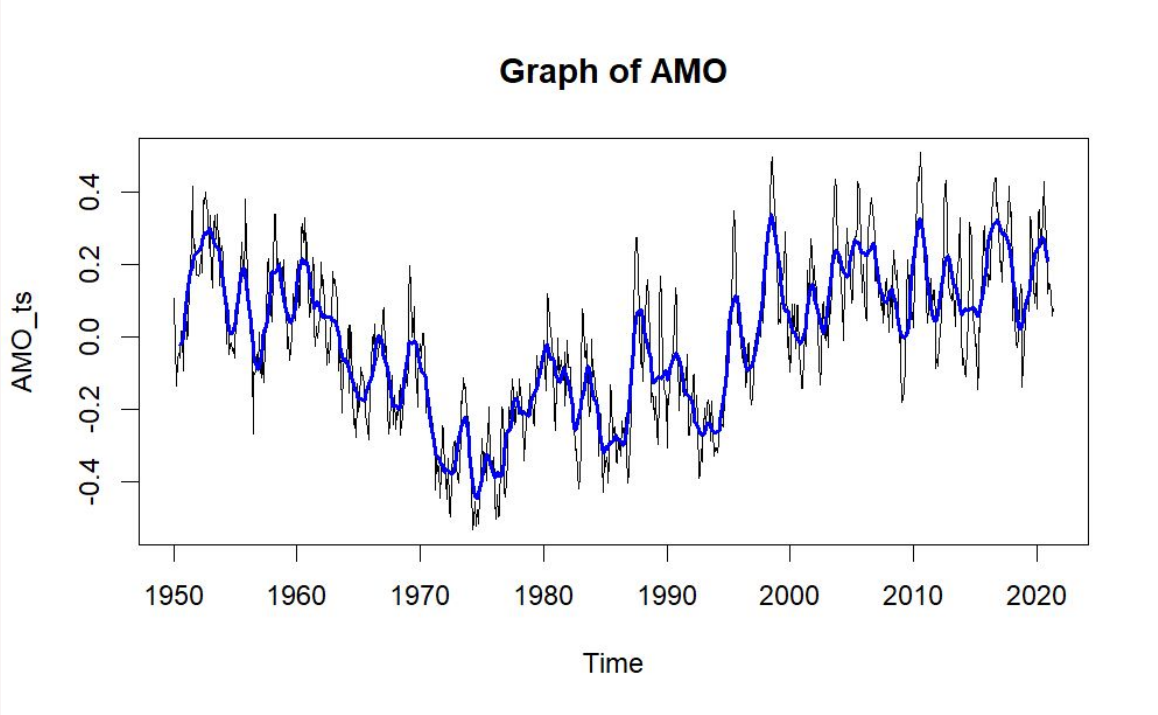
Smoothed Graphs



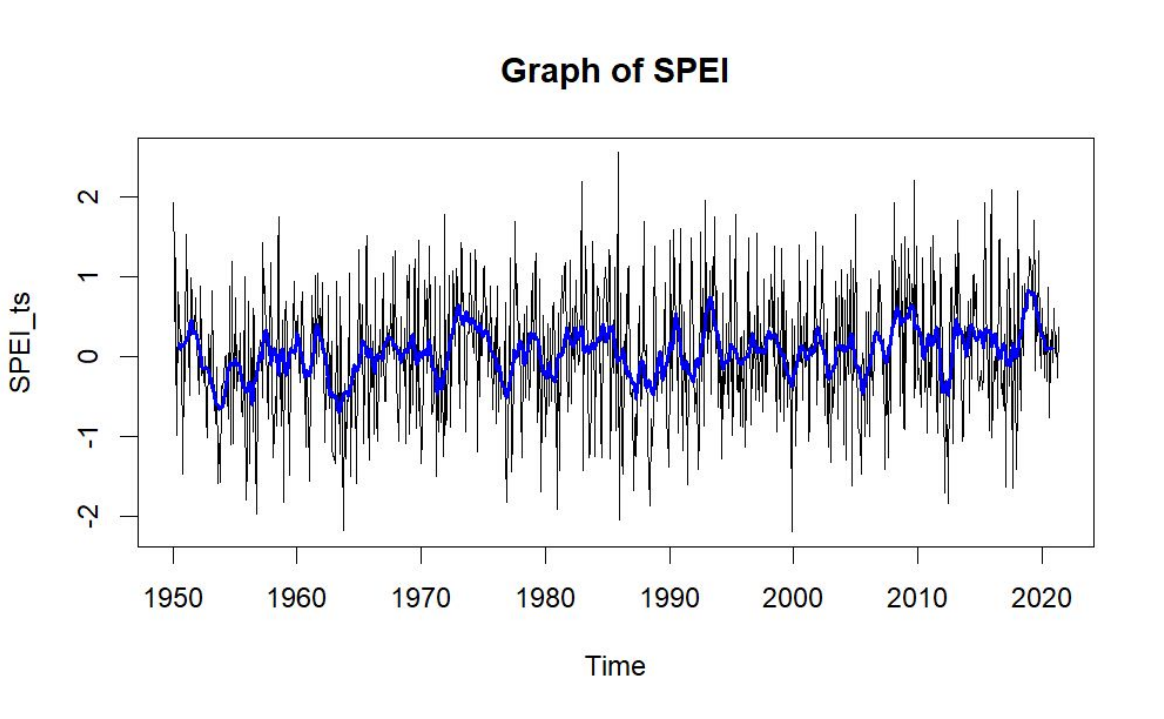
Smoothed Graphs



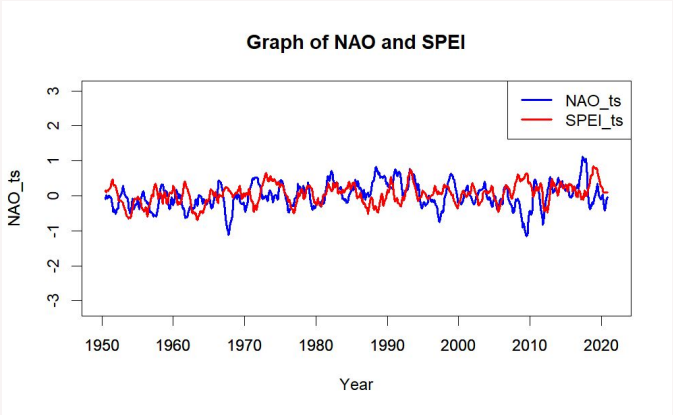
Smoothed Graphs

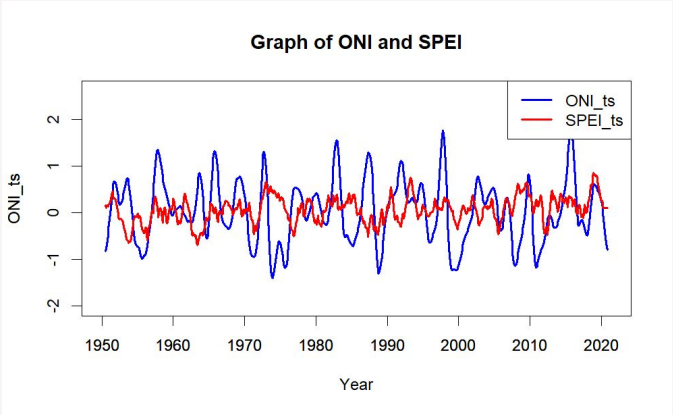


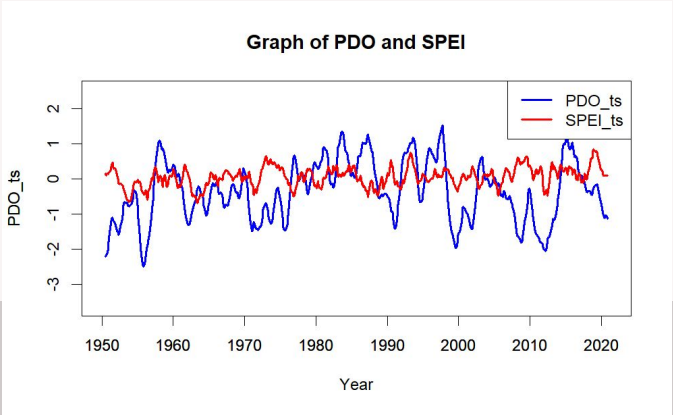
Smoothed Graphs

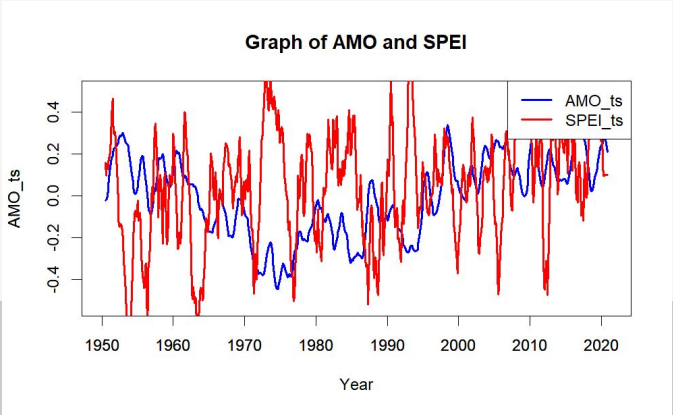


Graphs of Driver and Response









Autocorrelation and Cross-Correlation

Autocorrelation

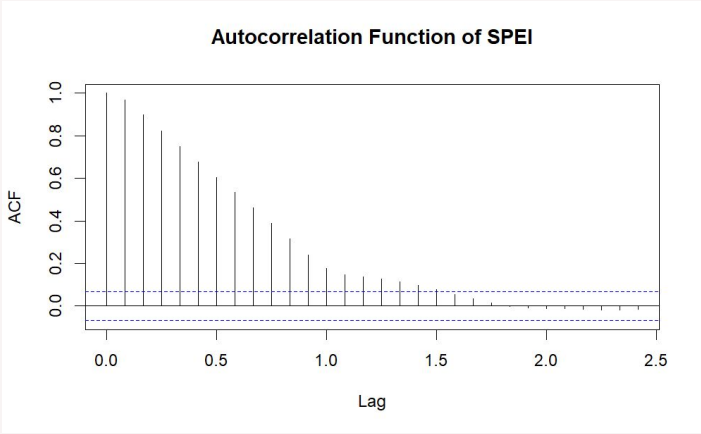
- Analysis is done for one time series
- Analyze the relationship or pattern within each time series data
- Pattern analysis, seasonality detection, forecasting, checkingg dependencies, etc can be done.
- If we can find some dependencies, it will be meaningful of conducting cross-correlation analysis

Cross-Correlation

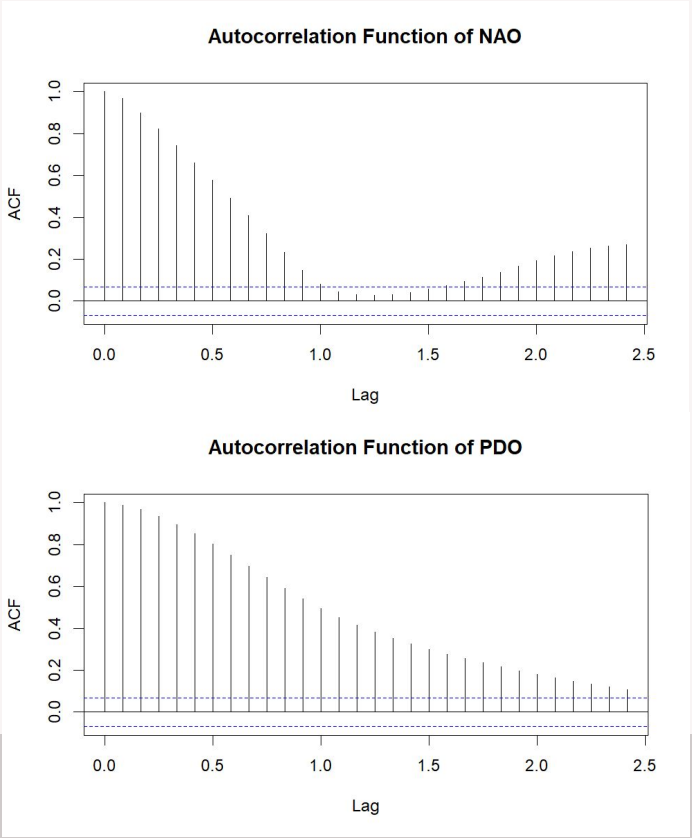
- Analysis is done for two different time series
- Measures the degree and direction of linear link when one series is moved or lagged in relation to the other

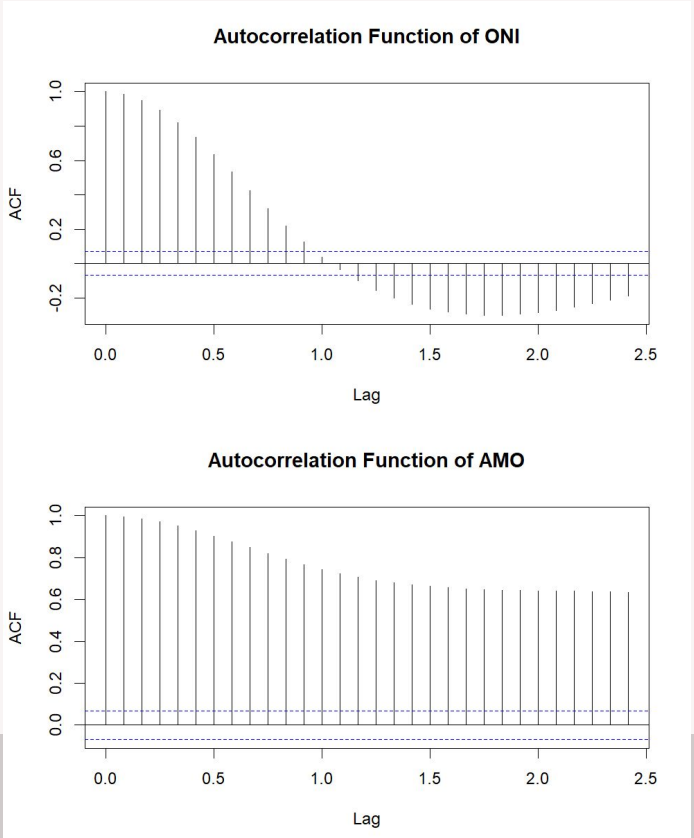
Autocorrelation Graph

- Autocorrelation graphs shows the dependence to its past data
- The x-axis implies the value of lag(month) and the y-axis represents the value of correlation coefficient
- When the y-value resides within the confidence interval, it means the relationship between the certain lag and another right before that lag is not significantly dependent
- On the other hand, if the y-value is located outside of the confidence interval, that means the relationship is significantly dependent
- We can see some significant dependencies from the graph



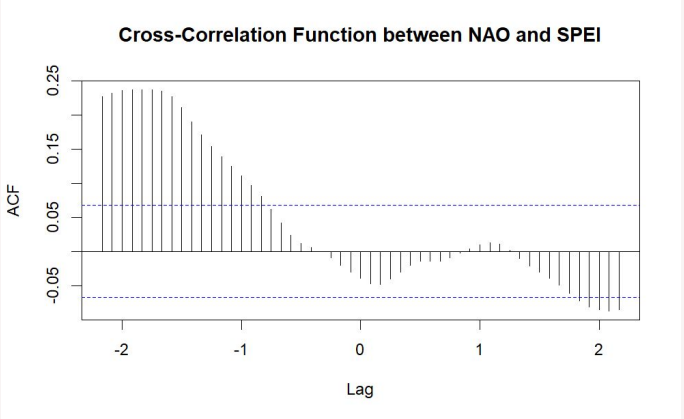
Autocorrelation Graphs



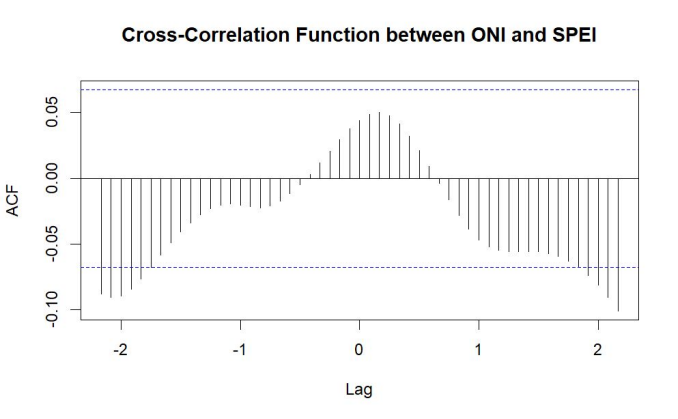


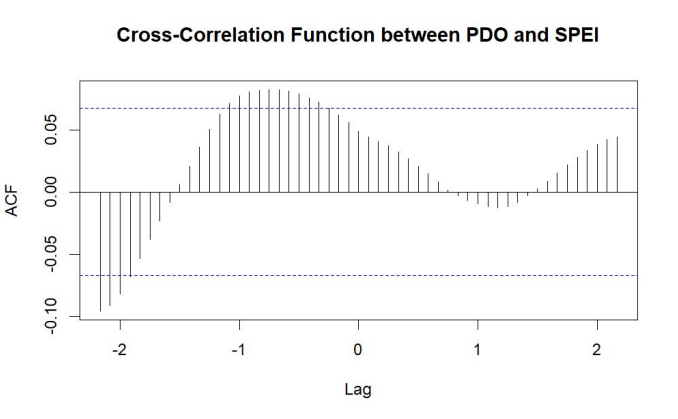
Cross-Correlation Graph

- Negative lag represents the relationship when second time series is shifted to later times while positive lag represents the relationship when the second time series is shifted to earlier times.
- When a value of correlation coefficient falls outside the confidence intervals, it indicates a statistically significant relationship between the two time series at that certain lag
- Negative lags can be more meaningful in cases where we expect the second time series to respond or be influenced by the first time series after a certain time delay
- We are more interested in the relationship between the drivers and drought after a certain time delay, so we are focusing on negative lags
- We can see some significant relationships from the graph.



Cross-Correlation Graphs





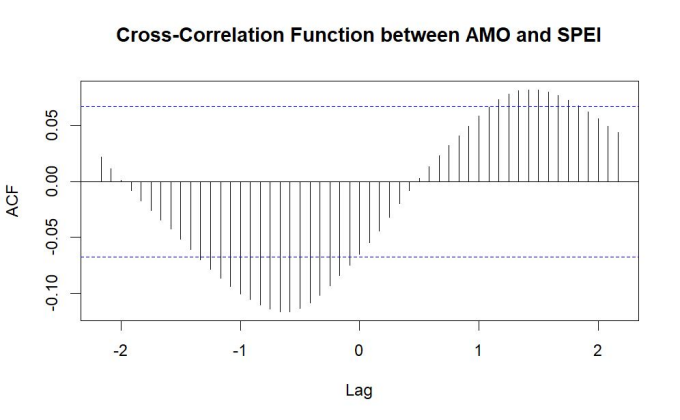
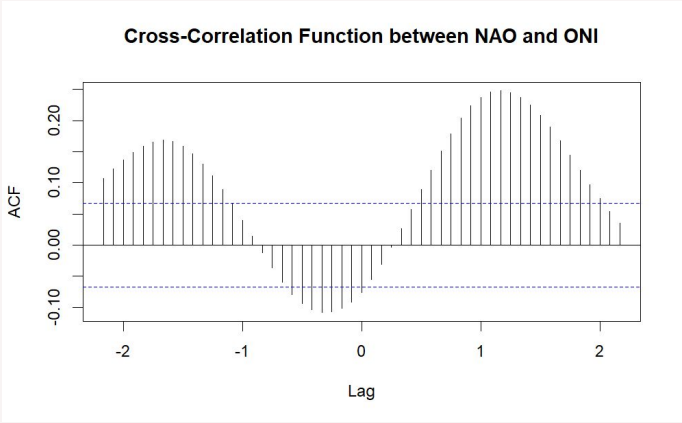
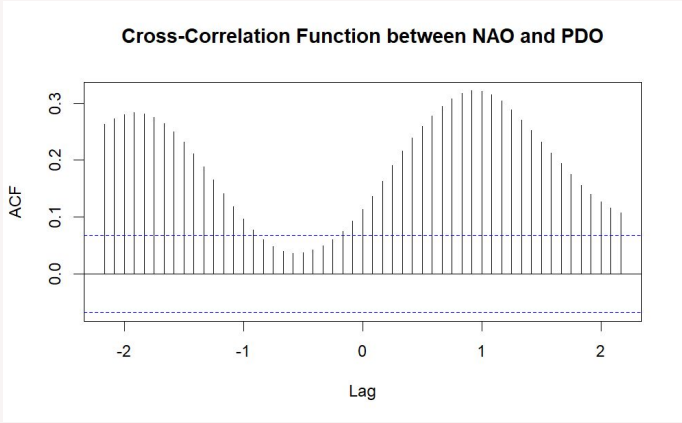


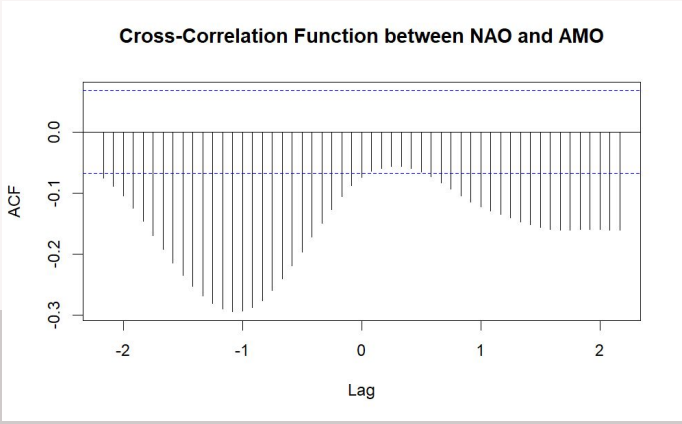
Table for max-Correlation Coefficient

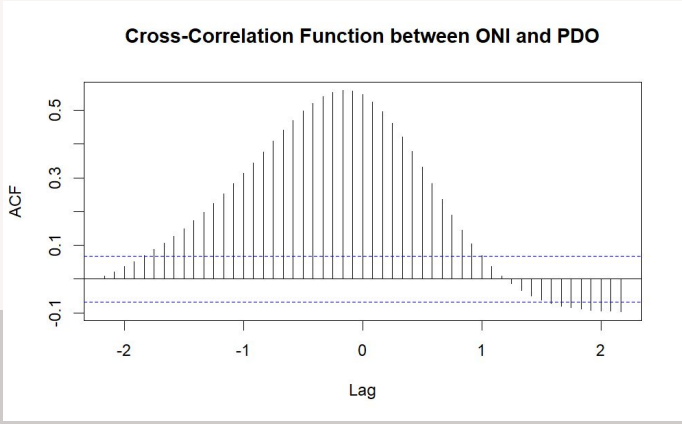
	SPEI	Correlation Coefficient
NAO	lag: −22	CC: 0.237
ONI	lag: −25	CC: −0.091
PDO	lag: −26	CC: −0.096
AMO	lag: −8	CC: −0.116

Cross-Correlation Graphs

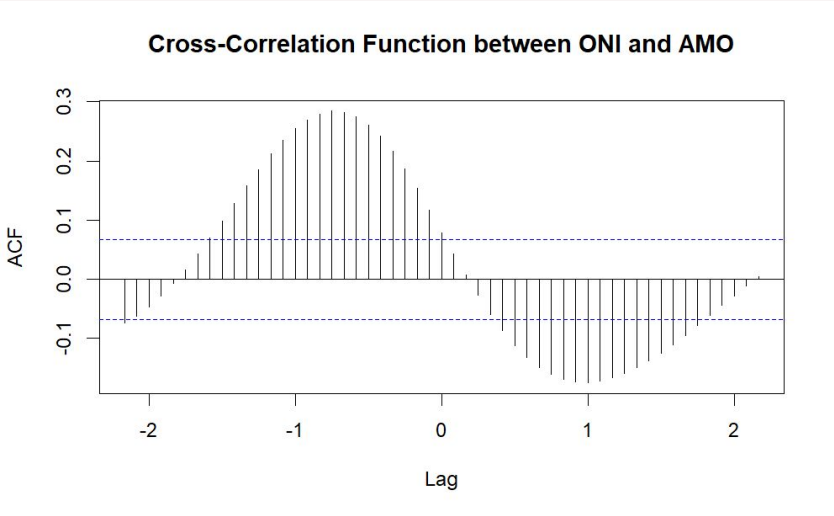


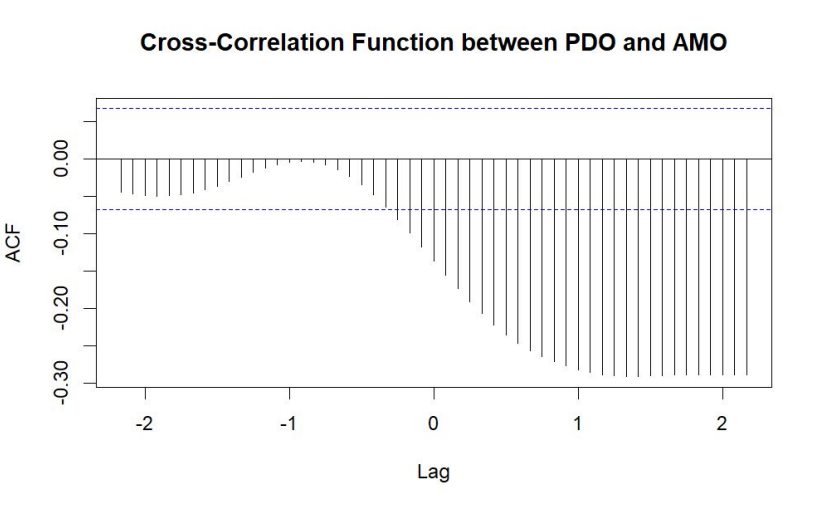






Cross-Correlation Graphs





Empirical Dynamic Models

- Developed by George Sugihara and colleagues
- They are a group of modeling techniques used to comprehend and analyze complex systems
- They aim to reconstruct the behavior of dynamic systems from time series and uncover the hidden variables and relationships

Convergent Cross Mapping (Sugihara Causality)

If X causes Y, we use the information of Y to predict X.

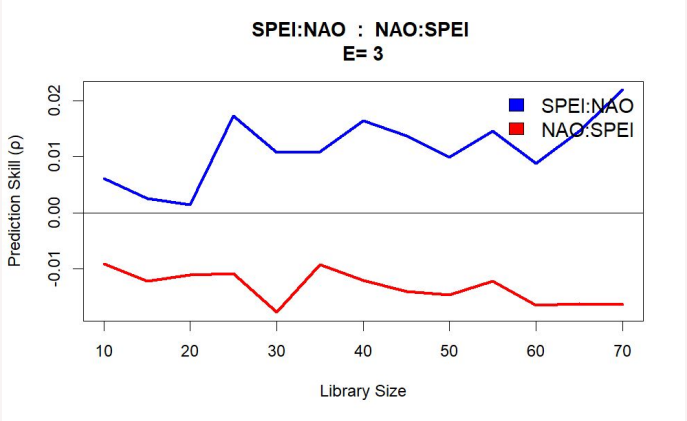
After predicting X, we compare it with the original data and measure the accuracy

We can find causality between variables even if the correlation is weak

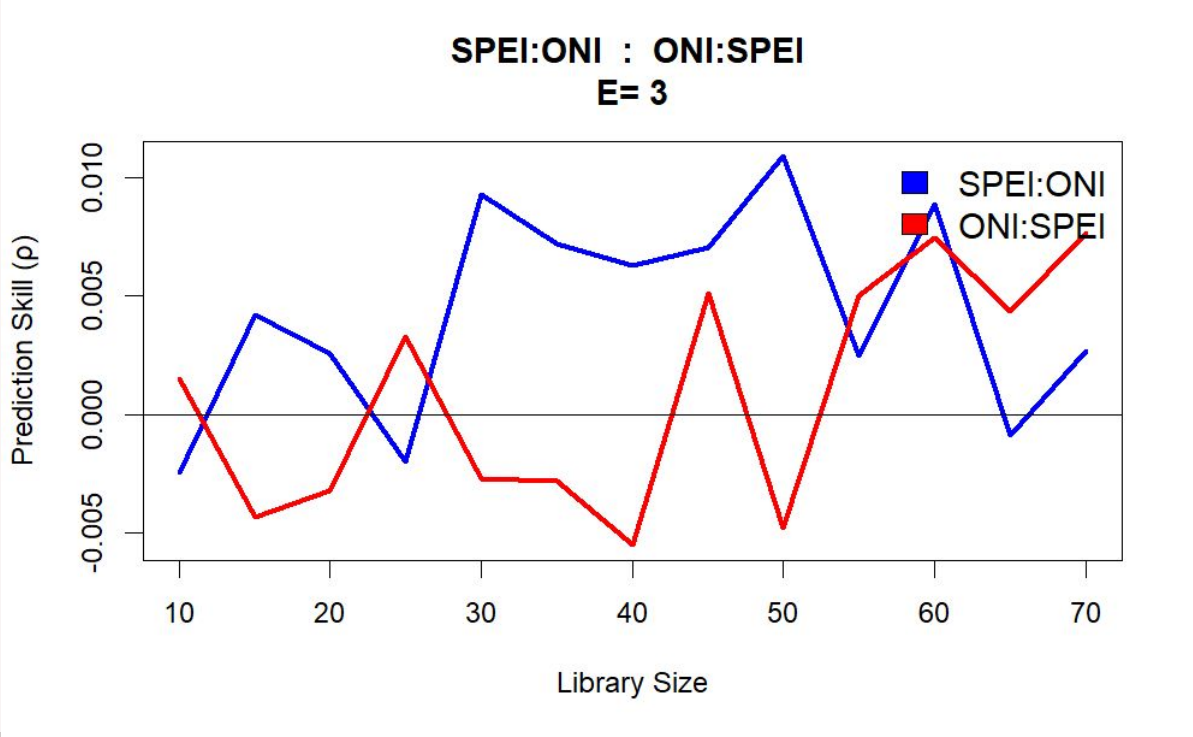
We can find if the causality happens in both way or one way.

Convergent Cross Mapping

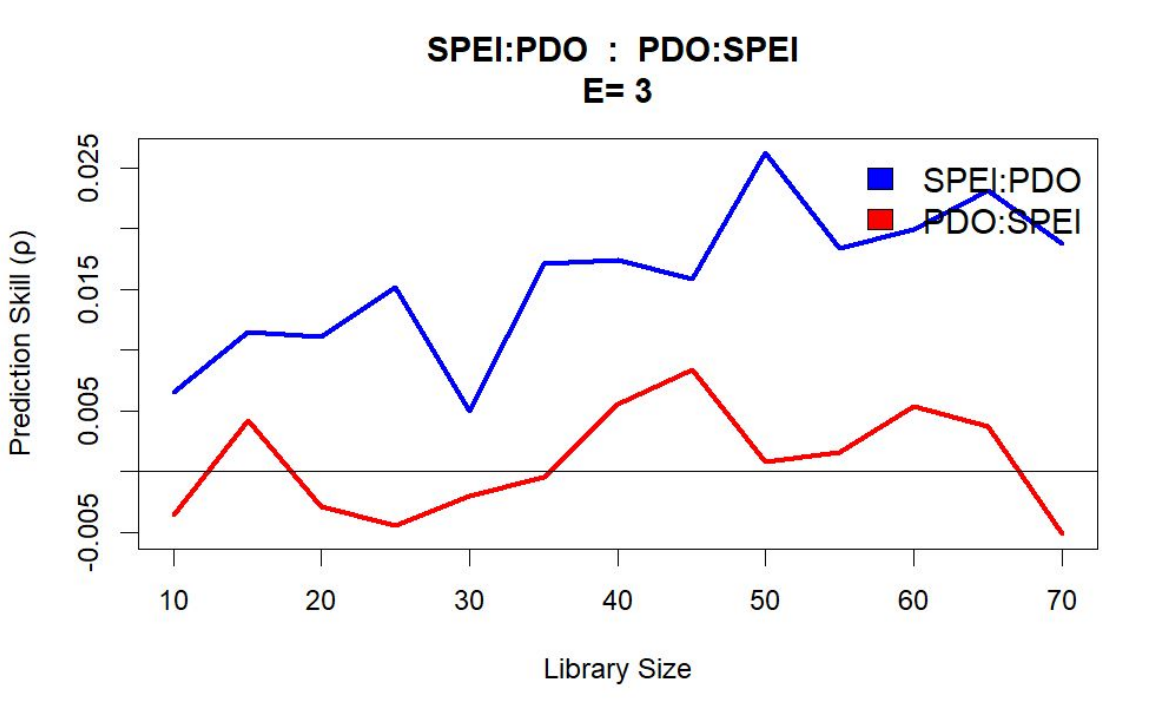
- If cross-variable state predictability converges to a positive value as more state-space information is provided, this indicates a causal link
- CCM indicates causal influence in the reverse direction (Y to X)
- From this graph, we see that cross mapping SPEI:NAO converges, indicating that NAO influences SPEI.
- Because average cross map skill less than 0 means there is no prediction skill, we infer from the graph of NAO:SPEI that SPEI do not affect NAO.



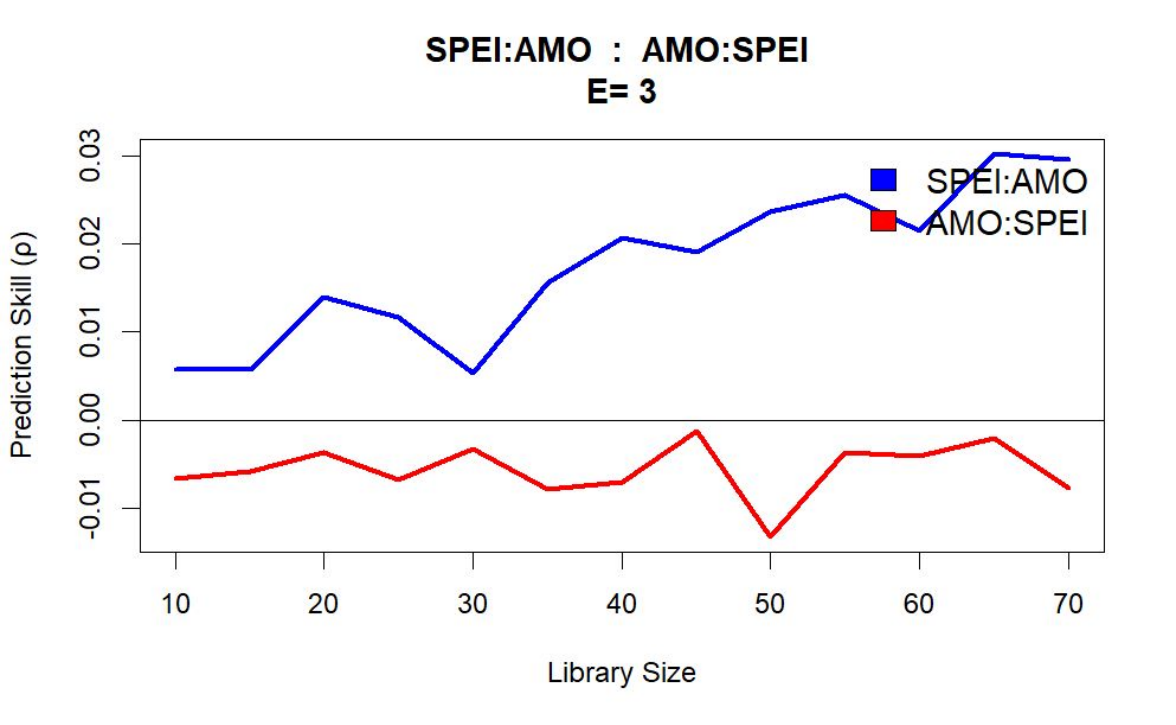
Convergent Cross Mapping



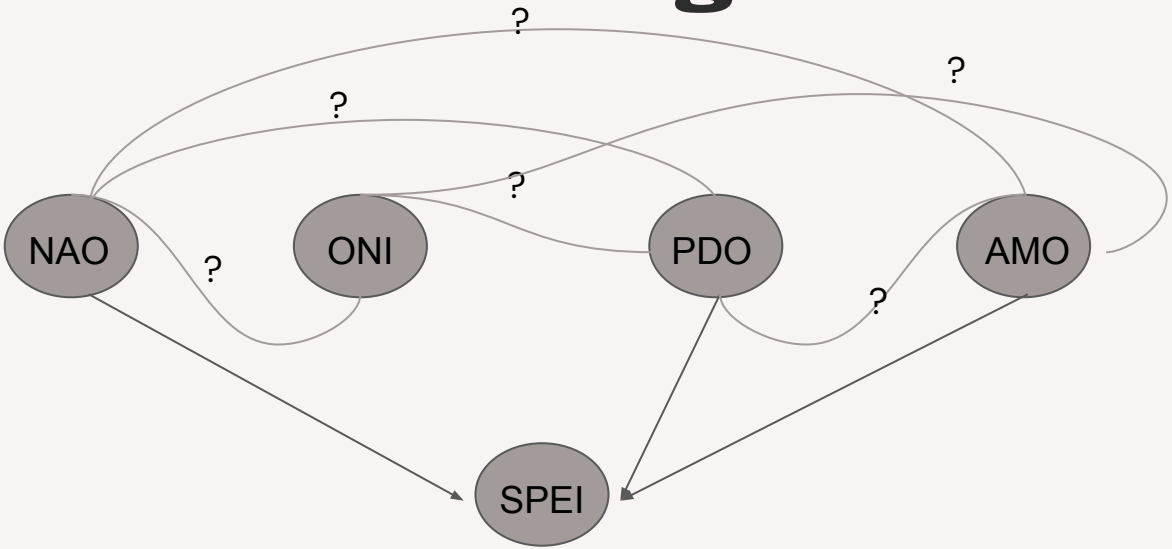
Convergent Cross Mapping



Convergent Cross Mapping



Causal Diagram



Conclusion and Application

- NAO, AMO, and PDO causes SPEI
- The Cross-Correlation graph can't be used to find out the causality even though we can see the significance, but the CCM can show the causality between the drivers and the response and even the direction of the causality
- We can come up with predictive model(machine learning) so that we can predict when the drought will happen. This can be helpful for farmers to prepare for the drought.
- We can apply the similar methods to find the relationship in biological sciences such as researching if temperature, rainfall, and El Niño causes abundance of mosquitoes (related to Malaria).

Work Cited Pages

Dahlman, R. L. A. L. (n.d.). *Climate variability: North atlantic oscillation*. NOAA Climate.gov. https://www.climate.gov/news-features/understanding-climate/climate-variability-north-atlantic-oscillation

Climate Data Guide. Palmer Drought Severity Index (PDSI) | Climate Data Guide. (n.d.). https://climatedataguide.ucar.edu/climate-data/palmer-drought-severity-index-pdsi

Lindsey, R. (n.d.). *Climate variability: Oceanic niño index*. NOAA Climate.gov. https://www.climate.gov/news-features/understanding-climate/climate-variability-oceanic-ni%C3%B1o-index

Climate Data Guide. Atlantic Multi-decadal Oscillation (AMO) | Climate Data Guide. (1870, January 1). https://climatedataguide.ucar.edu/climate-data/atlantic-multi-decadal-oscillation-amo

Climate Data Guide. Pacific Decadal Oscillation (PDO): Definition and Indices | Climate Data Guide. (1900, January 1). https://climatedataguide.ucar.edu/climate-data/pacific-decadal-oscillation-pdo-definition-and-indices

Sugihara Lab. Deep Eco. (n.d.). https://deepeco.ucsd.edu/

sp. (n.d.). *A platform to benchmark causal discovery methods*. CauseMe. https://causeme.uv.es/

Work Cited Pages

YouTube. (2019, December 19). *Casual interference with convergent cross mapping - cátia fortunato - barcamp LX*. YouTube. https://www.youtube.com/watch?v=JwzBK-KBhTY

Pearl, J., & Mackenzie, D. (2018).The Book Of Why: The New Science Of Cause and Effect. Basic Books.

Joseph Park [aut, cre] (). (2023, January 9). *REDM: Vignettes/redm-tutorial.rmd*. R Package Documentation. https://rdrr.io/cran/rEDM/f/vignettes/rEDM-tutorial.Rmd