جامعة الملك عبدالله للعلوم والتقنية
King Abdullah University of Science and Technology

Biostatistics at KAUST

Arctic Sea Ice:
Exploring global warming

Hector G. Flores Rodriguez

Outline

- Motivation
- Background
- Approaches
- Preliminary Analysis

Motivation

- UCI Data Science Climate Hackathon
  - Investigate: California's drought, Arctic sea ice, or West Coast wildfires.
- Why Arctic sea ice?
  - Interested in exploring if climate change is predominantly caused by human influences.

Background

The data are provided in NetCDF-4 file

Variable	Description
latitude	Latitude in degrees
longitude	Longitude in degrees (0 to 360)
seaice_conc	Sea ice concentration in percent with values from 0 to 100, inclusive, Land is indicated by -1.
seaice_source	Describes the source of the data for each month of data beginning with January 1850
time	Time of the data observation in days since 1850-01-01 00:00:00

Contributing data sources

- From 1850 to 1978, analog sources were used:
  - Charts, yearbooks, ice extent grids, whaling ship logs, and etc.
- From 1979 to present:
  - Sea ice concentration from satellite passive microwave data.

Merging data sources

- Ranking
  - Each possible sources for a concentration value was given a rank (higher numbers outrank lower).
  - Satellite data is the exception to the ranking rule.
- Temporal consistency
  - Monthly data points represent ice at the 15th or 16th of each month.
  - Mid month value was selected instead of monthly average.

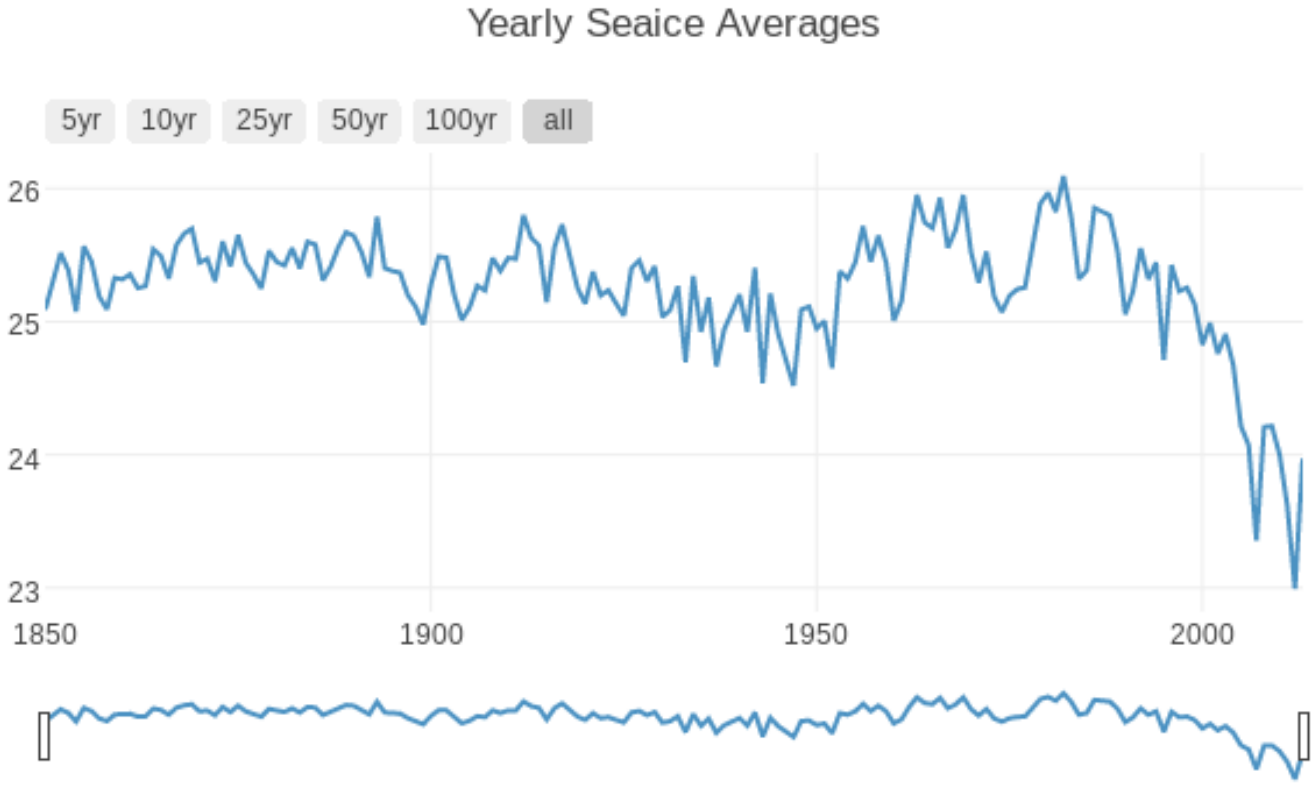
Approaches

One of the most visible of signs of warming is the retreat of Arctic sea ice.

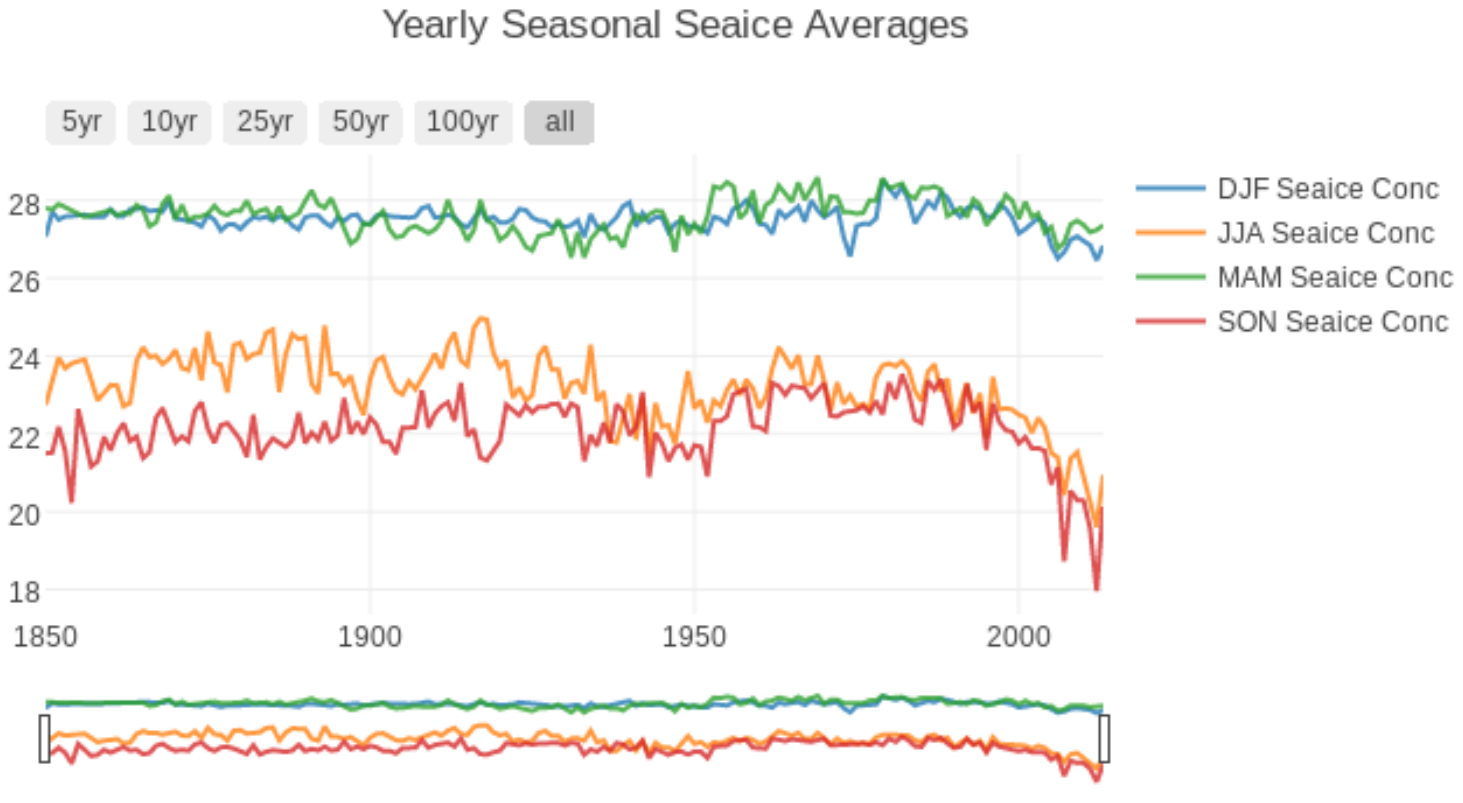
Develop visualizations to help analyze the spatial and temporal similarity:

- Yearly averages
- Seasonal averages
- Trend analysis
- Spatio-temporal clustering

Data Visualizations

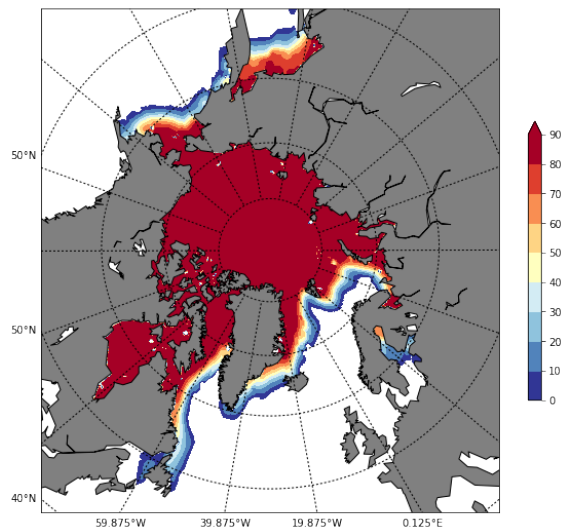


Data Visualizations

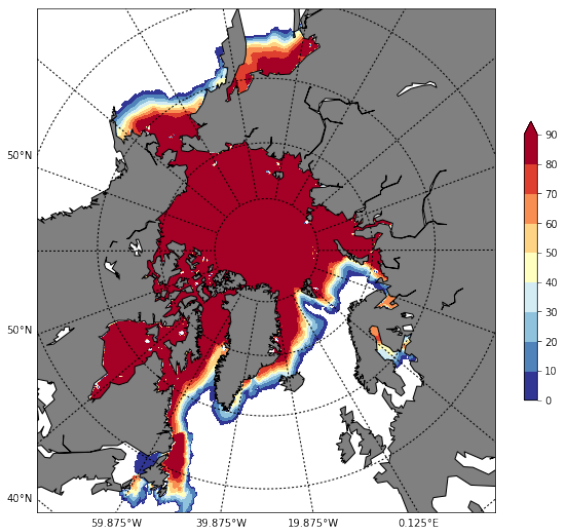


Data Visualizations

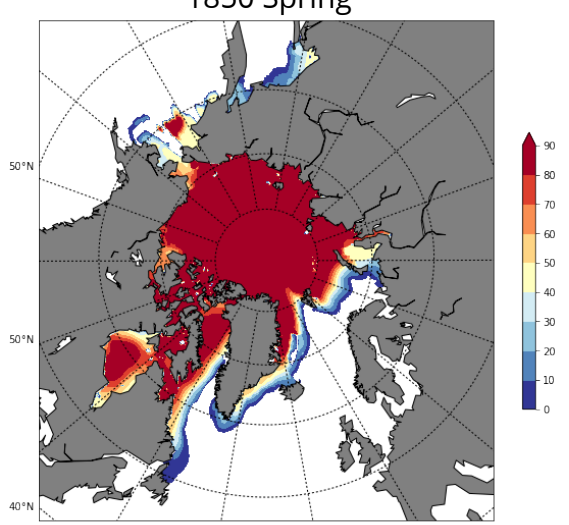
1850 Fall



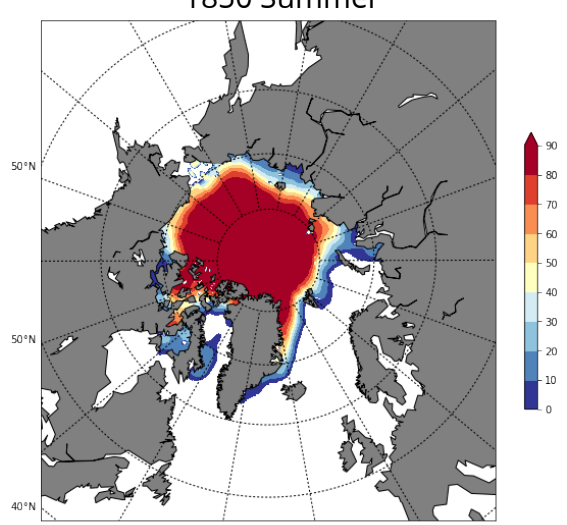
1850 Winter



1850 Spring

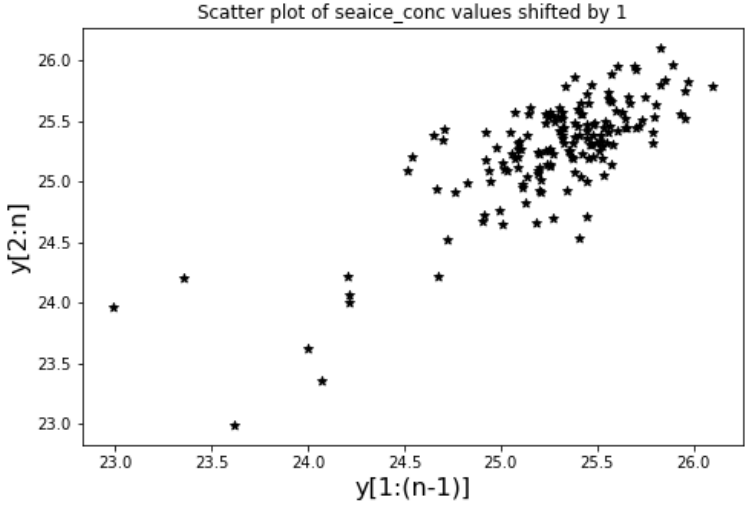


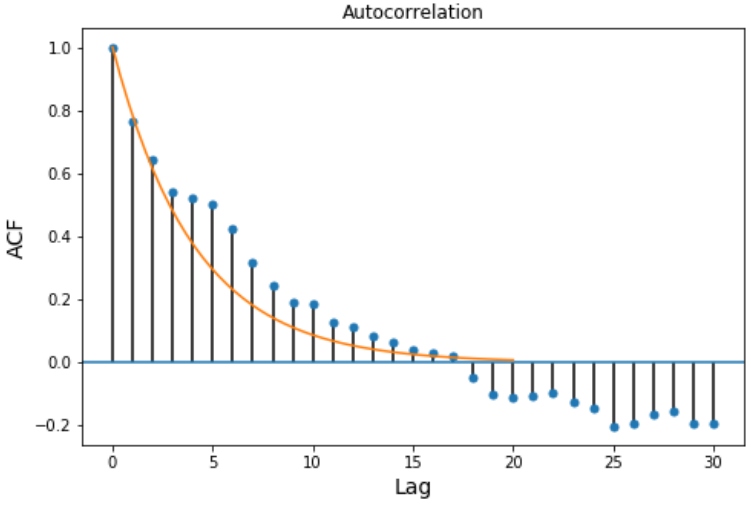
1850 Summer



Preliminary Analysis

Estimate autocorrelation function from time series data.





Preliminary Analysis

Simple trend analysis with linear regression

```
Residuals:
     Min       1Q   Median       3Q      Max
-1.99912 -0.17707  0.00519  0.20418  1.00251

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 31.7106483  1.3827229  22.933  < 2e-16 ***
t1          -0.0033389  0.0007157  -4.665 6.41e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.4339 on 162 degrees of freedom
Multiple R-squared:  0.1184,    Adjusted R-squared:  0.113
F-statistic: 21.77 on 1 and 162 DF,  p-value: 6.412e-06
```

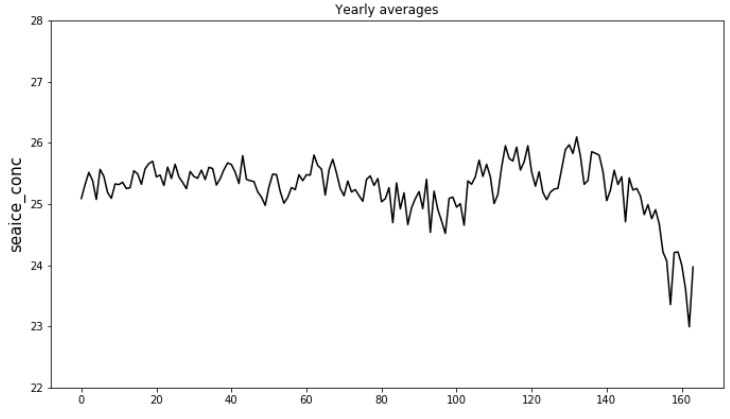
Preliminary Analysis

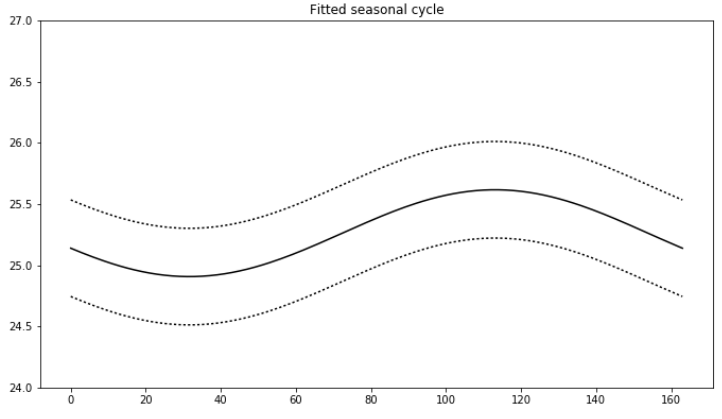
Trend analysis II: Linear regression with AR1 covariance

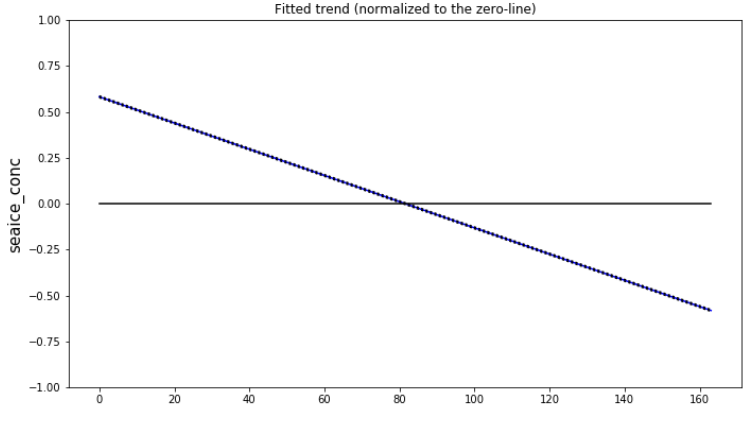
```
Generalized least squares fit by maximum likelihood
  Model: y ~ t1
  Data: NULL
       AIC      BIC    logLik
  59.81062 72.21009 -25.90531

Correlation Structure: AR(1)
 Formula: ~1
 Parameter estimate(s):
      Phi
0.7662343

Coefficients:
               Value Std.Error   t-value p-value
(Intercept) 32.43808  3.655355  8.874128  0.0000
t1          -0.00373  0.001892 -1.971625  0.0504

 Correlation:
   (Intr)
t1 -1

Standardized residuals:
        Min          Q1         Med          Q3         Max
-4.40980463 -0.35191840  0.05324723  0.53181699  2.38815117

Residual standard error: 0.4398229
Degrees of freedom: 164 total; 162 residual
```

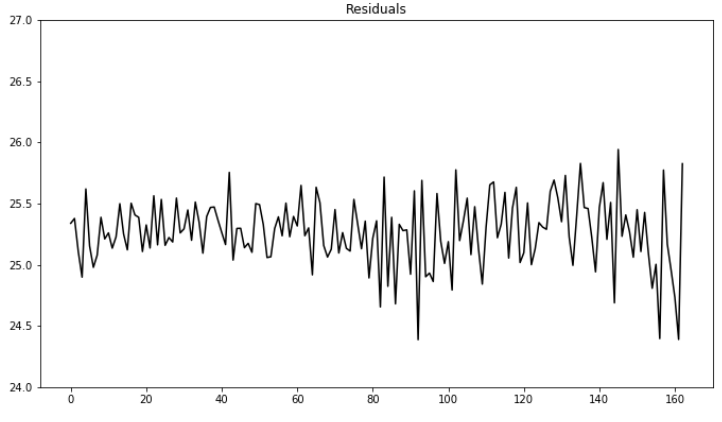
Preliminary Analysis

Time series decomposition









Preliminary Analysis

Trend analysis II: time series decomposition

```
Generalized least squares fit by maximum likelihood
  Model: y ~ sinf + cosf + t1
  Data: NULL
       AIC      BIC    logLik
  59.08643 77.68563 -23.54322

Correlation Structure: AR(1)
 Formula: ~1
 Parameter estimate(s):
     Phi
0.723617

Coefficients:
               Value Std.Error   t-value p-value
(Intercept) 39.03511  4.553353  8.572829  0.0000
sinf        -0.33283  0.164774 -2.019891  0.0451
cosf        -0.12260  0.108425 -1.130770  0.2598
t1          -0.00714  0.002357 -3.029358  0.0029

 Correlation:
     (Intr) sinf   cosf
sinf -0.732
cosf -0.001  0.000
t1   -1.000  0.732  0.000

Standardized residuals:
       Min         Q1        Med         Q3        Max
-3.8765989 -0.4946197  0.1250875  0.6931627  2.3477637

Residual standard error: 0.4037854
Degrees of freedom: 164 total; 160 residual
```

Future Work

- Investigate climate analysis and models
- Spatial Analysis
- Interactive map of sea ice concentration
- Spatio-temporal clustering
- Include "data source" into analysis