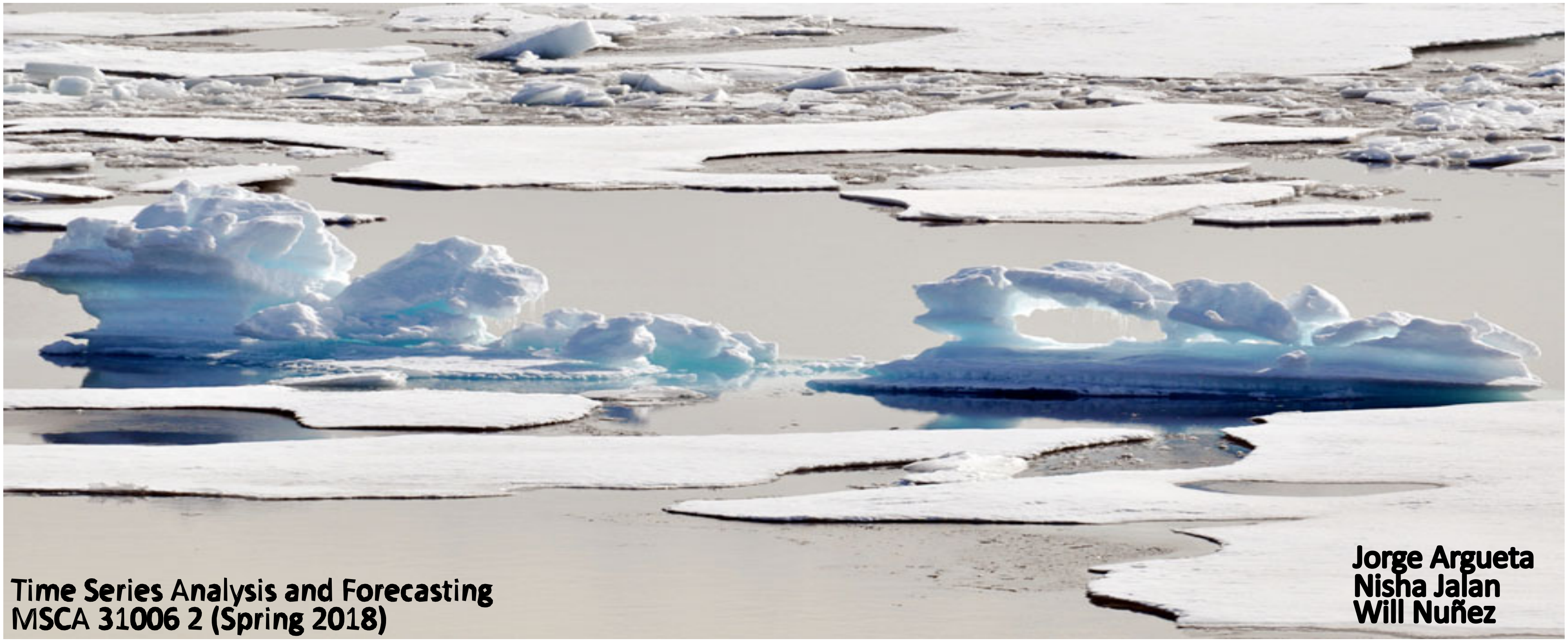
THE UNIVERSITY OF CHICAGO

Sea Ice Extent Forecast



Time Series Analysis and Forecasting
MSCA 31006 2 (Spring 2018)

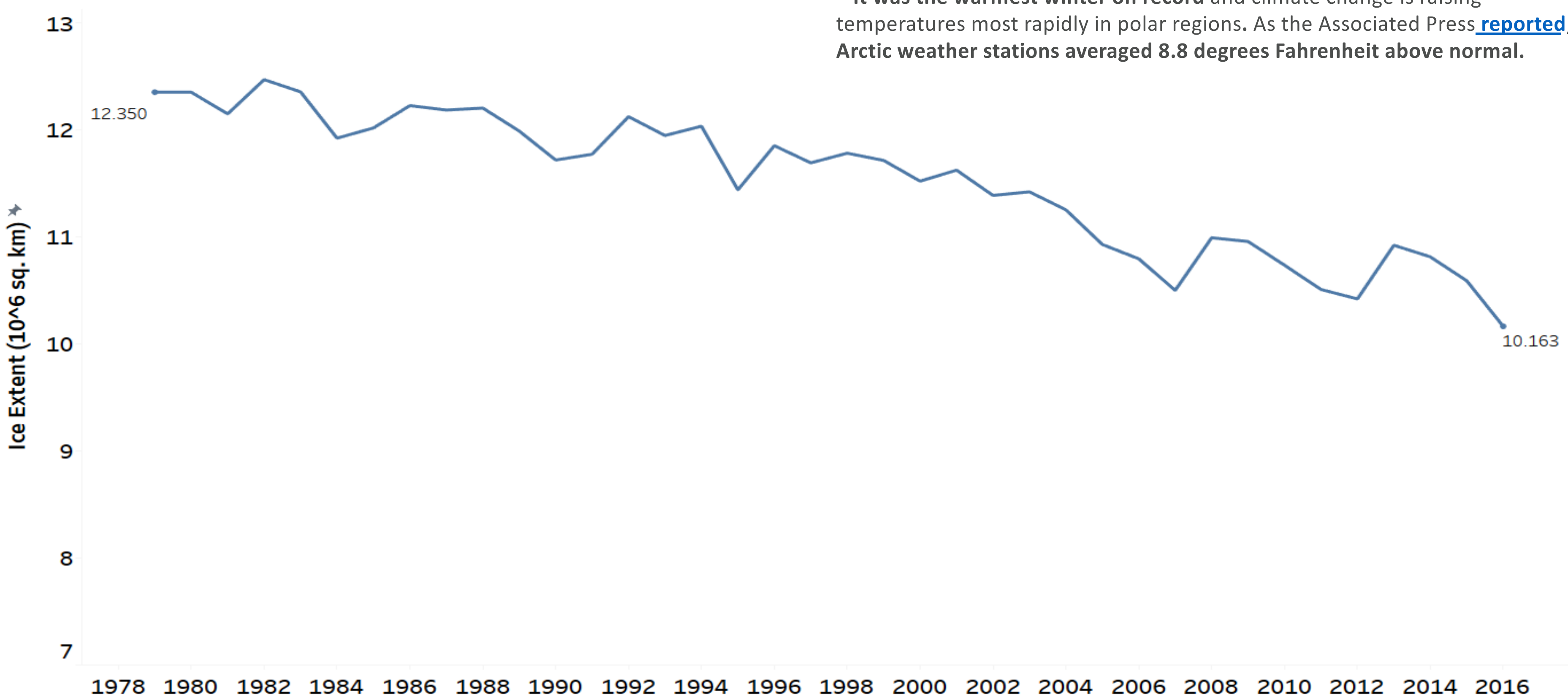
Jorge Argueta
Nisha Jalan
Will Nuñez

Agenda

- **Introduction**
  - Background
  - Problem Statement
  - Dataset Description
- **Exploratory Analysis**
- **Modeling**
  - Model Building
  - Model Selection
  - Forecasting
- **Results**
- **Future Work**

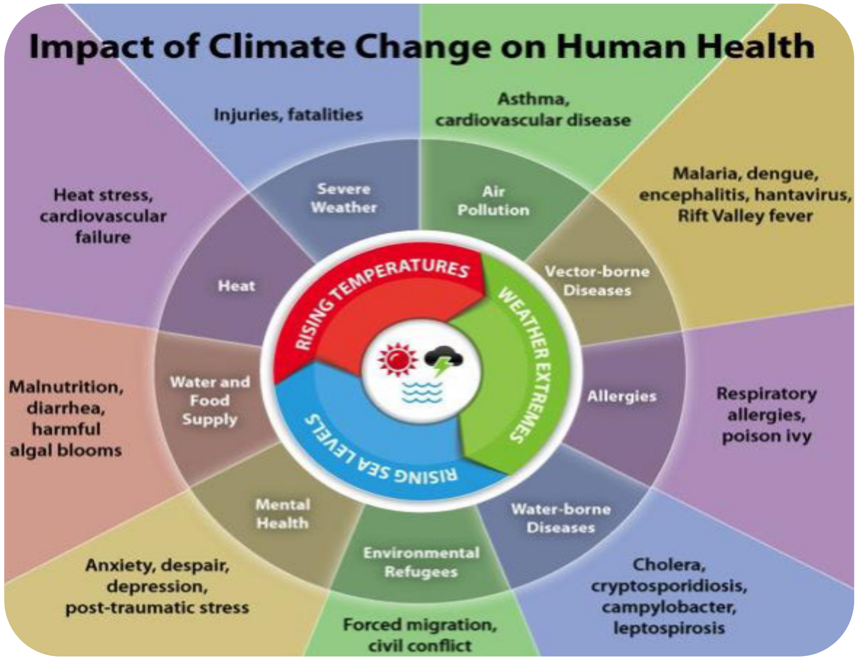
In 2017, the Arctic's sea ice extent was the smallest it's been in nearly 40 years of satellite records

* **It was the warmest winter on record** and climate change is raising temperatures most rapidly in polar regions**.** As the Associated Press [reported](), **Arctic weather stations averaged 8.8 degrees Fahrenheit above normal.**

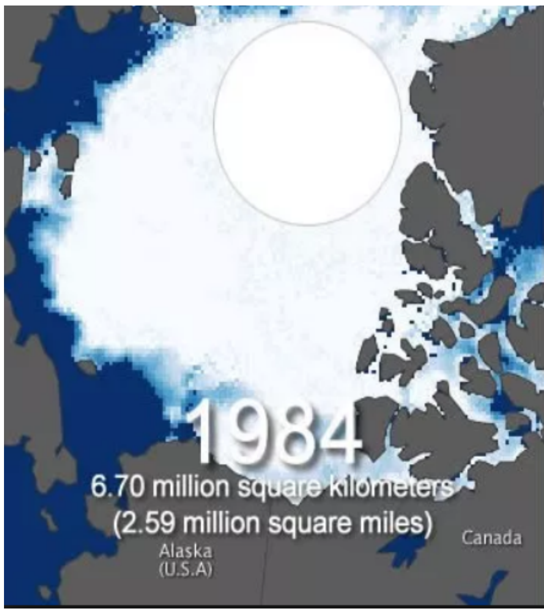


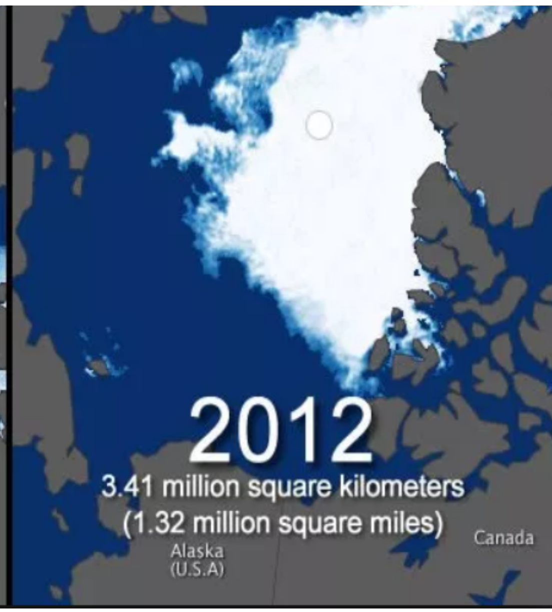
In a few decades, scientists predict that the Artic will become virtually ice-free with less than 1 million sq.km of ice (a bit larger than the size of Texas) during the peak of summer

- Gradually warming temperatures melt sea ice over time, fewer bright surfaces are available to reflect sunlight back into space, more solar energy is absorbed at the surface, and temperatures rise further:
  - This chain of events starts a cycle of warming and melting
  - This cycle is temporarily halted when the dark days of the polar winter returns, but it starts again in the following spring
  - Even a small increase in temperature can lead to greater warming over time, making the polar regions the most sensitive areas to climate change on Earth
- The fast shrinking Arctic ice cap is increasingly thought to have **major impacts on extreme weather patterns much further south**, due to its **influence on the jet stream**
  - Floods, heatwaves and severe winters in Europe, Asia and North America have all been linked to the Arctic meltdown



Our goal is to build forecasting models to predict the ice extent for the Northern Hemisphere in order to raise awareness of global warming





Future?

(July, 2017 to June, 2018)

The dataset consists of the Northern Hemisphere (Arctic) sea ice extent

- Sea ice is simply frozen ocean water that forms, grows, and melts in the ocean. In contrast, icebergs, glaciers, ice sheets, and ice shelves all originate on land
- The data set is the snow and ice product produced by the National Ice Center (NIC) and the input product is archived at the National Snow and Ice Data Center (NSIDC)
- The NIC constructs the data set using satellites and the NSIDC supports scientific research into our world's frozen realms by managing and distributing data that educates the public about the cryosphere
- The data are provided in the polar stereographic projection at a grid cell size of 25 x 25 km

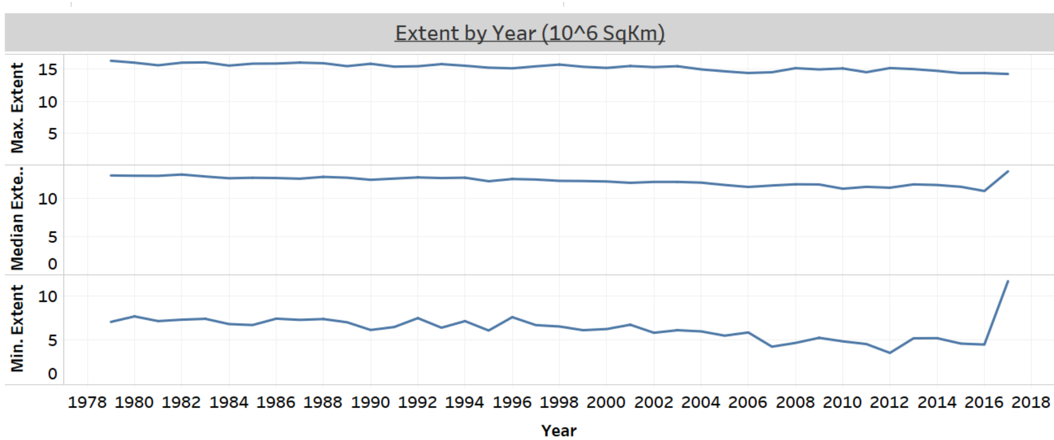
Initial overview

Variables

- Year *(Jan, 1979 to June, 2017)*
- Month
- Day *(mix of alternate day and daily)*
- Ice Extent *(unit is 10^6 sq. km)*

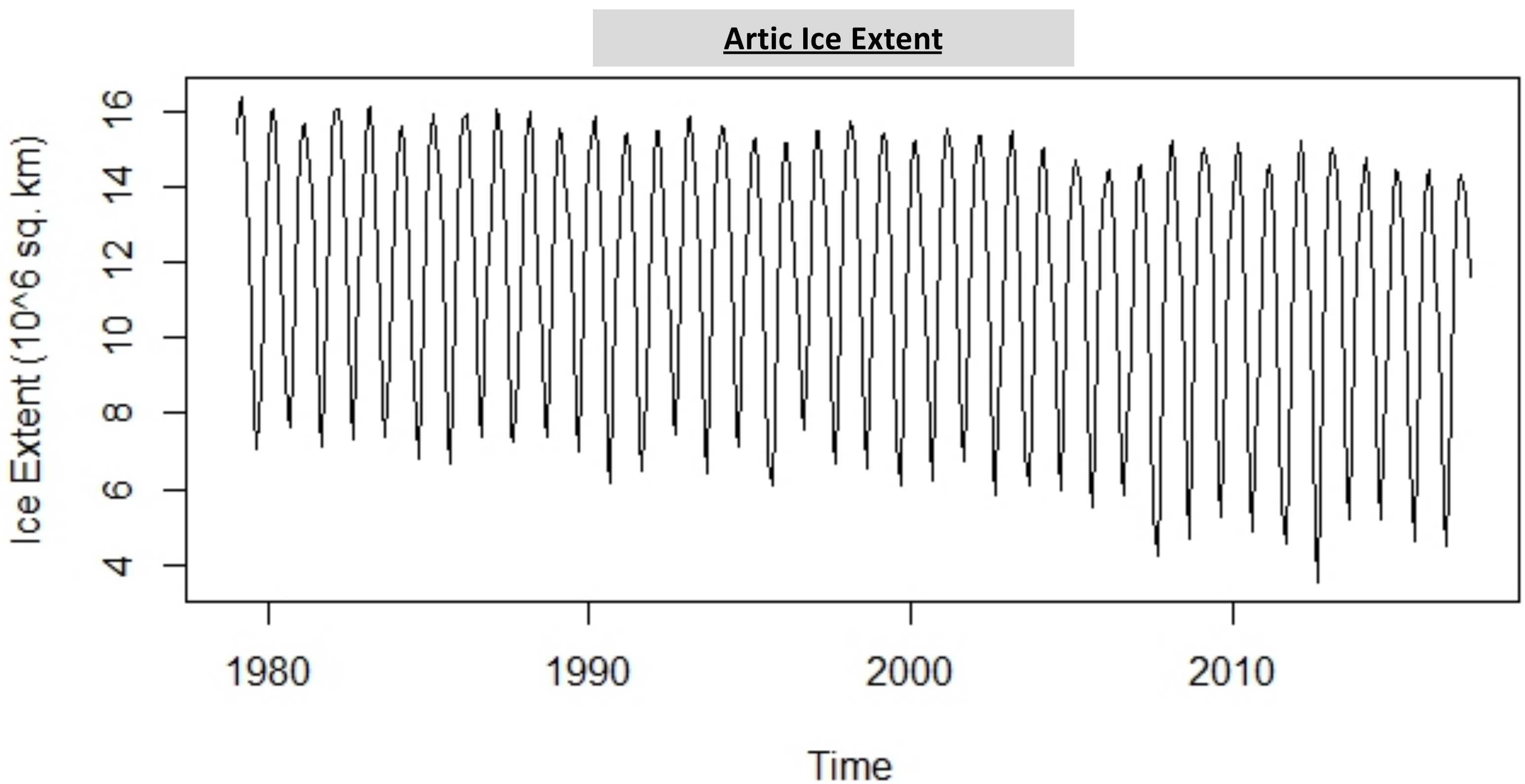
Assumptions

- The data was rolled up to a monthly level by taking the mean of inconsistent data *(mix of daily and every other day)*
- We assume that mean is the best representation of monthly ice-extent



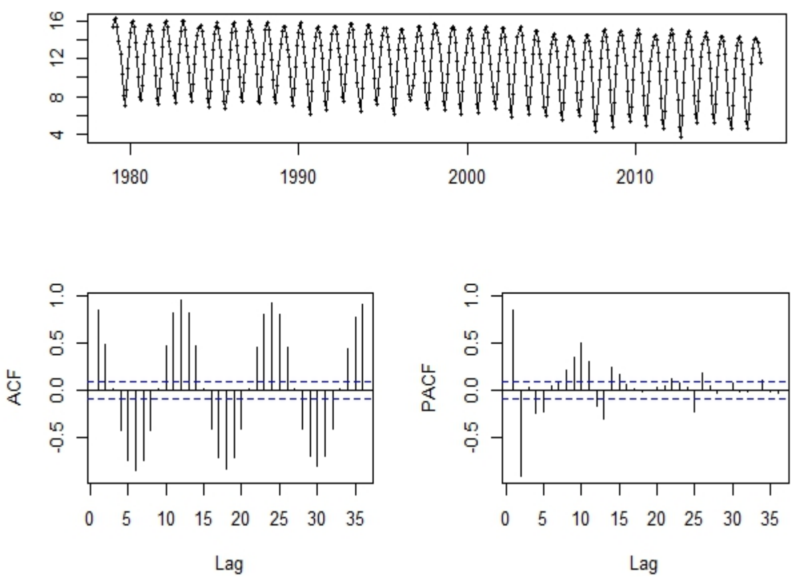
Exploratory Analysis

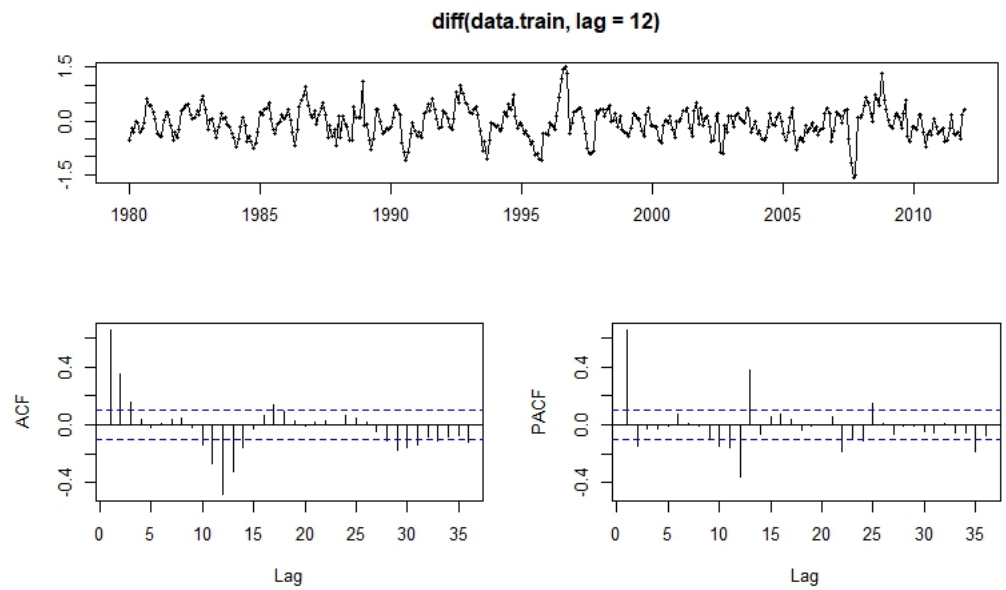
Arctic Ice Extent seems to be declining YoY; the time series shows clear seasonality



The time series in not stationary as seen in ACF plot; seasonal differencing leads to stationarity

<u>Artic Sea Ice Extent</u>

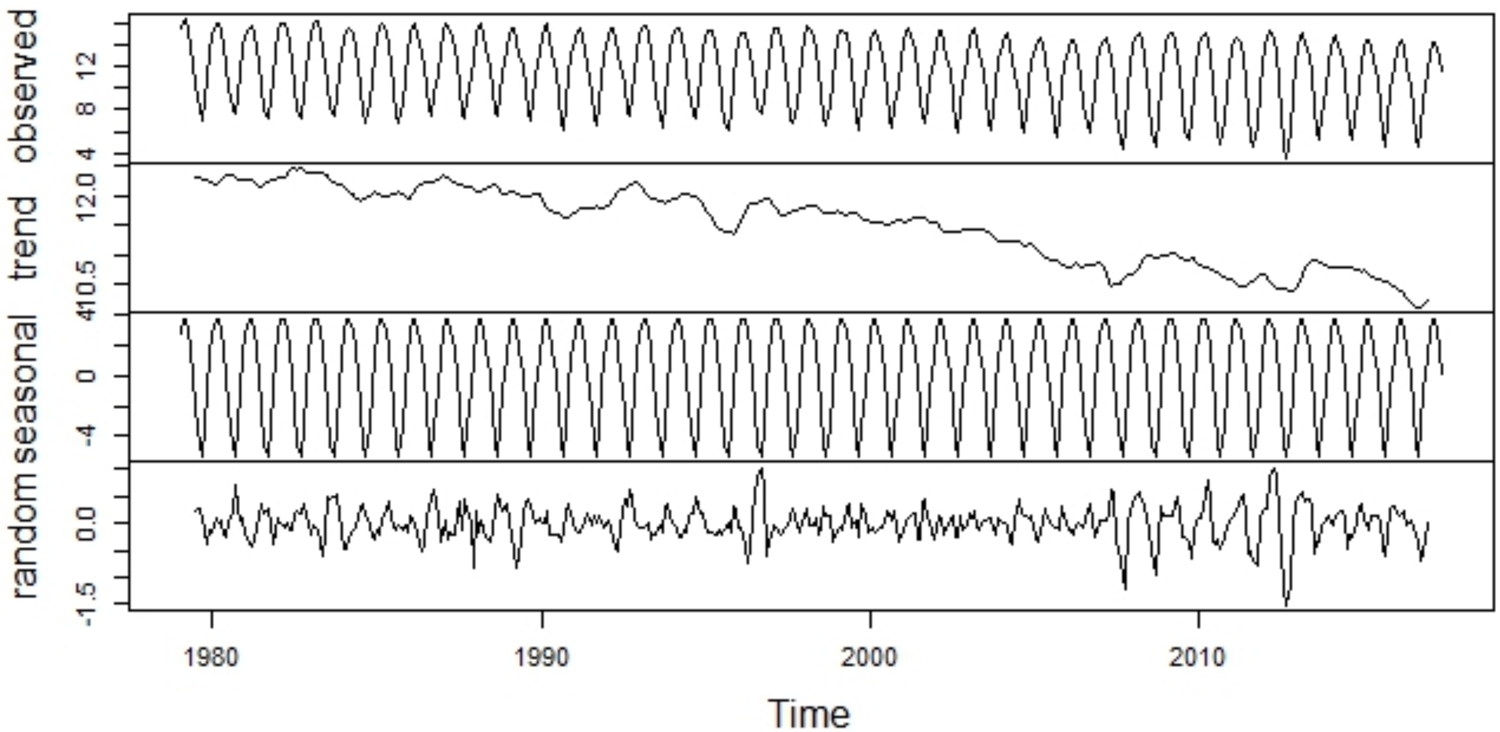




KPSS Test: p=0.1
ADF Test: p=.01

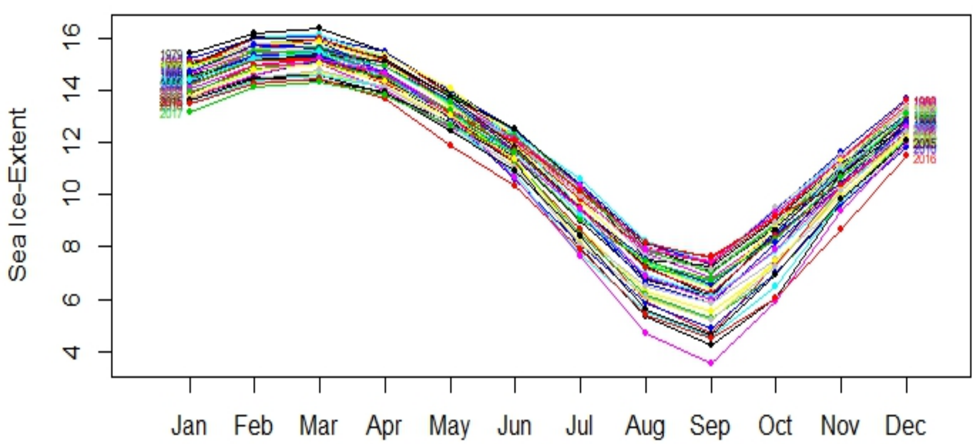
The seasonality exhibited by the time series is additive; the trend is on a decline

Decomposition of Additive Time Series

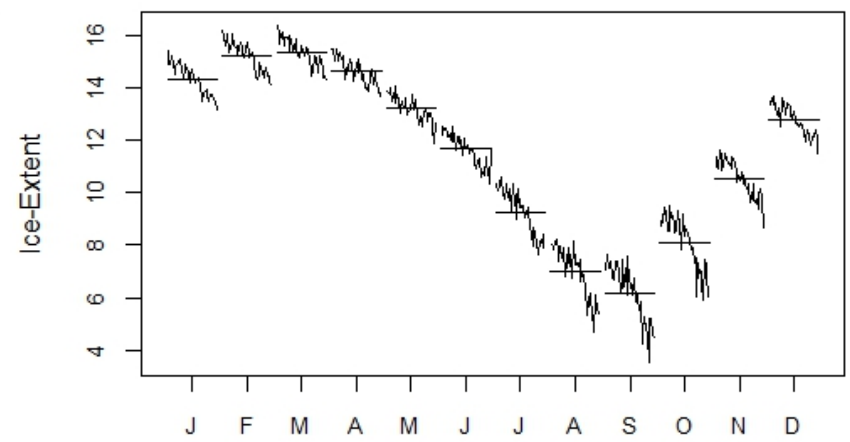


The seasonal plot shows that there is a comparatively steeper decrease in Ice extent in the months of August to October

Annual Trend



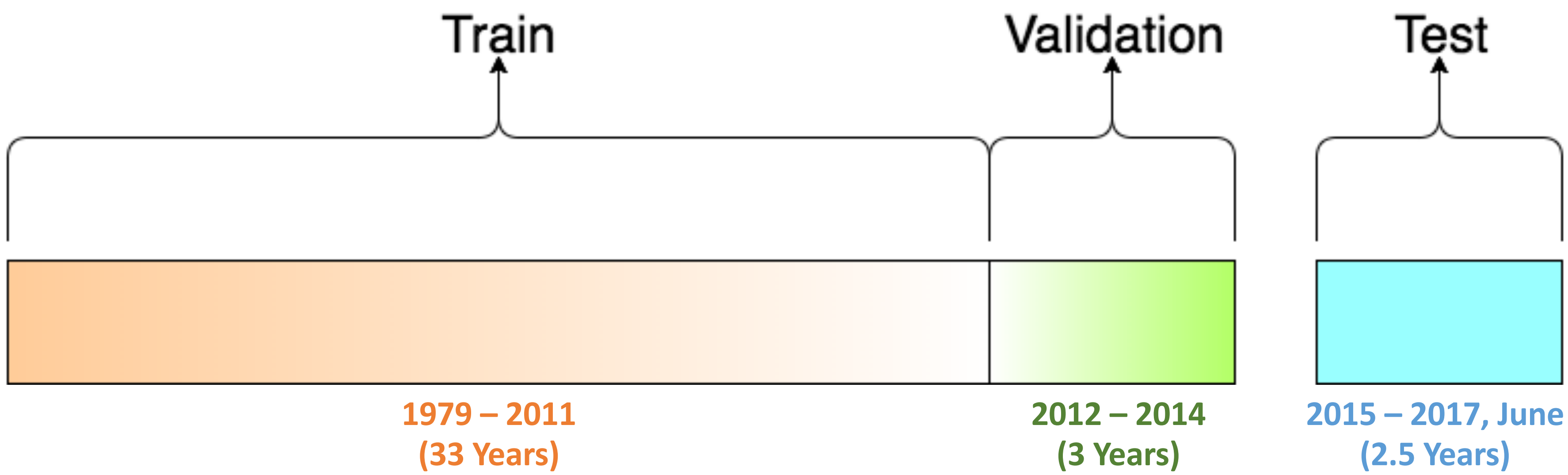
Monthly Trend



Modeling

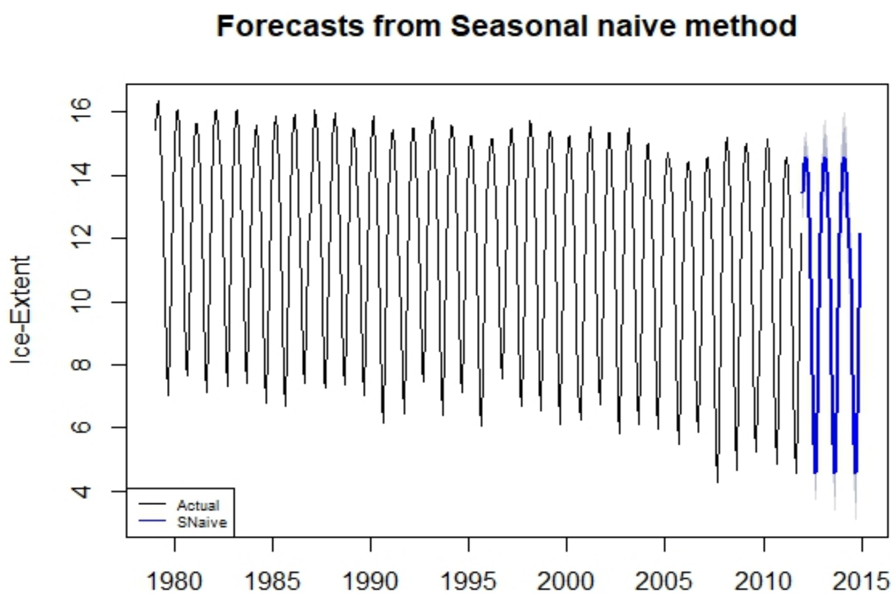
Pre-Model: Data Set-up

- The data was divided into 3 parts
- ~85% of the data was kept for training
- The remaining ~15% was split into validation and test



Benchmark: Seasonal Naïve

- The forecast is set to be equal to the last observed value from the same season of the year



Model	Train Error Measure
SNaïve	RMSE: 0.43

Model 1: Exponential Smoothing Models

1 Determination – Choosing the right model

SIMPLE EXPONENTIAL SMOOTHING
(for series with no trend & no seasonality)

$$s_i = \alpha x_i + (1-\alpha)s_{i-1}, \text{ with } 0 \le \alpha \le 1$$

- x_i is the actual value at time i.
- α is the mixing parameter. How much new vs. old information is used.
- s_i is the smoothed value at time i.

HOLT'S EXPONENTIAL SMOOTHING
(for series with trend & no seasonality)

$$s_i = \alpha x_i + (1-\alpha)(s_{i-1} + t_{i-1})$$
$$t_i = \beta(s_i - s_{i-1}) + (1-\beta)t_{i-1}$$

- t_i is the trend at time i. It is the difference between subsequent values in the series.
- β is a mixing parameter for the trend. How much new vs. old trend information to include.

HOLT WINTER'S EXPONENTIAL SMOOTHING
(for series with trend & seasonality)

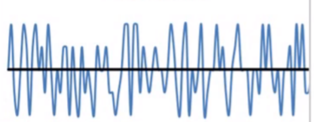
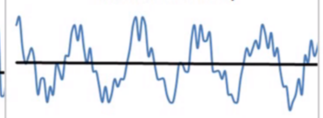
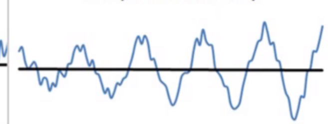
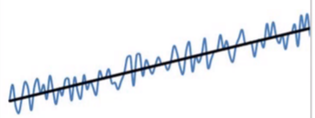
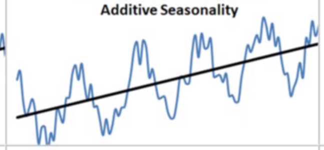
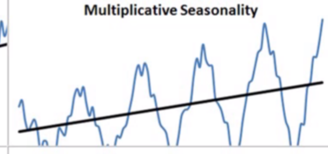
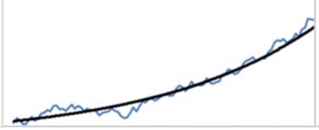
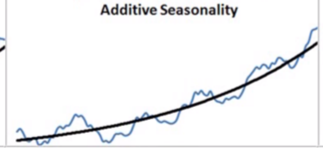
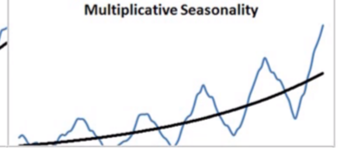
$$s_i = \alpha(x_i - p_{i-k}) + (1-\alpha)(s_{i-1} + t_{i-1})$$
$$t_i = \beta(s_i - s_{i-1}) + (1-\beta)t_{i-1}$$
$$p_i = \gamma(x_i - s_i) + (1-\gamma)p_{i-k}$$

- k is the length of the period.
- γ is a mixing parameter for the seasonality.
- s_i and t_i are to be thought of as 'doubly-smoothed' values, without taking seasonality into account.
- p_i is the seasonal part.

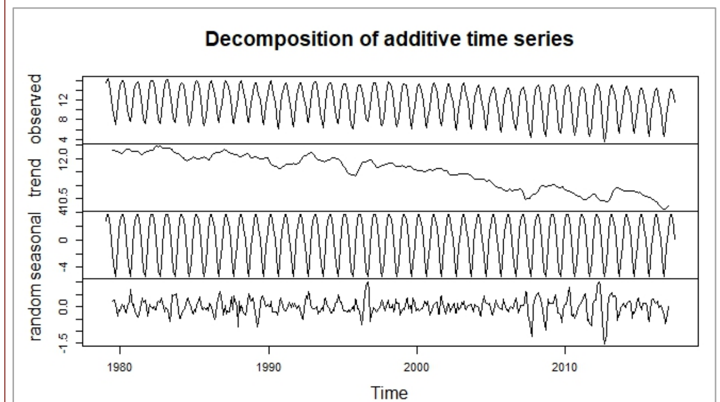
To forecast beyond the last data point: $x_{i+h} = s_i + ht_i + p_{i+h-k}$

Trend & Seasonality

2 Specification – Identifying Model Specifications

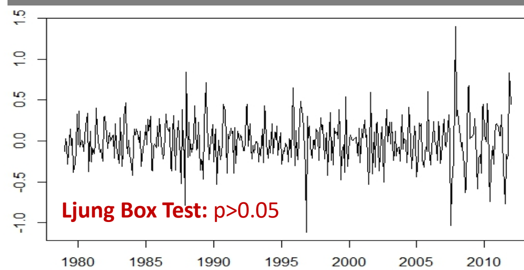
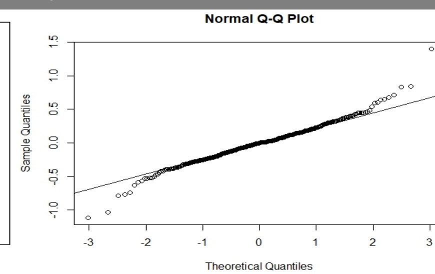
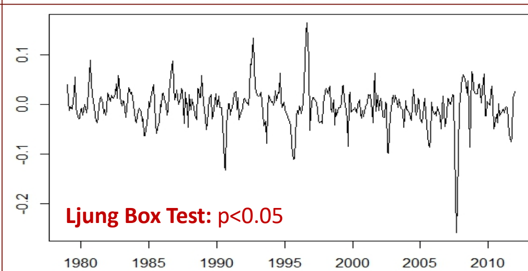
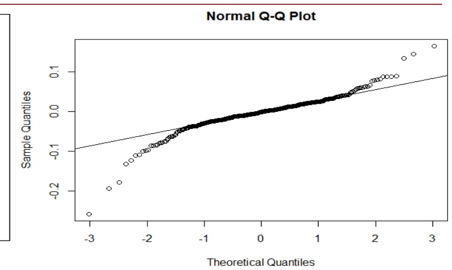
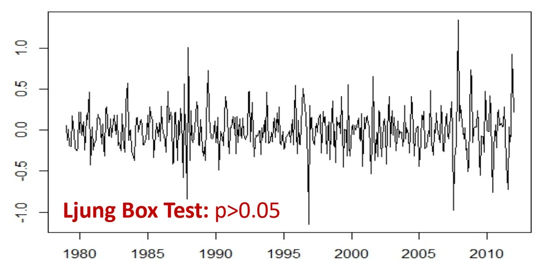
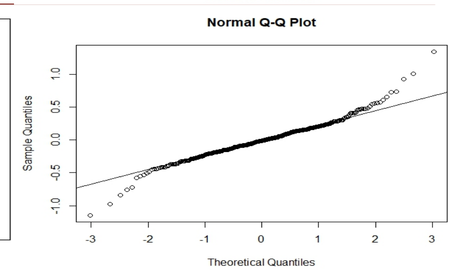
	No Seasonality	Additive Seasonality	Multiplicative Seasonality
No Trend	Constant Trend Non-seasonal 	Constant Trend with Additive Seasonality 	Constant Trend with Multiplicative Seasonality 
Additive Trend	Upward Linear Trend Non-seasonal 	Upward Linear Trend with Additive Seasonality 	Upward Linear Trend with Multiplicative Seasonality 
Multiplicative Trend	Upward Exponential Trend Non-seasonal 	Upward Exponential Trend with Additive Seasonality 	Upward Exponential Trend with Multiplicative Seasonality 

The Artic Ice Extent series shows additive trend and seasonality



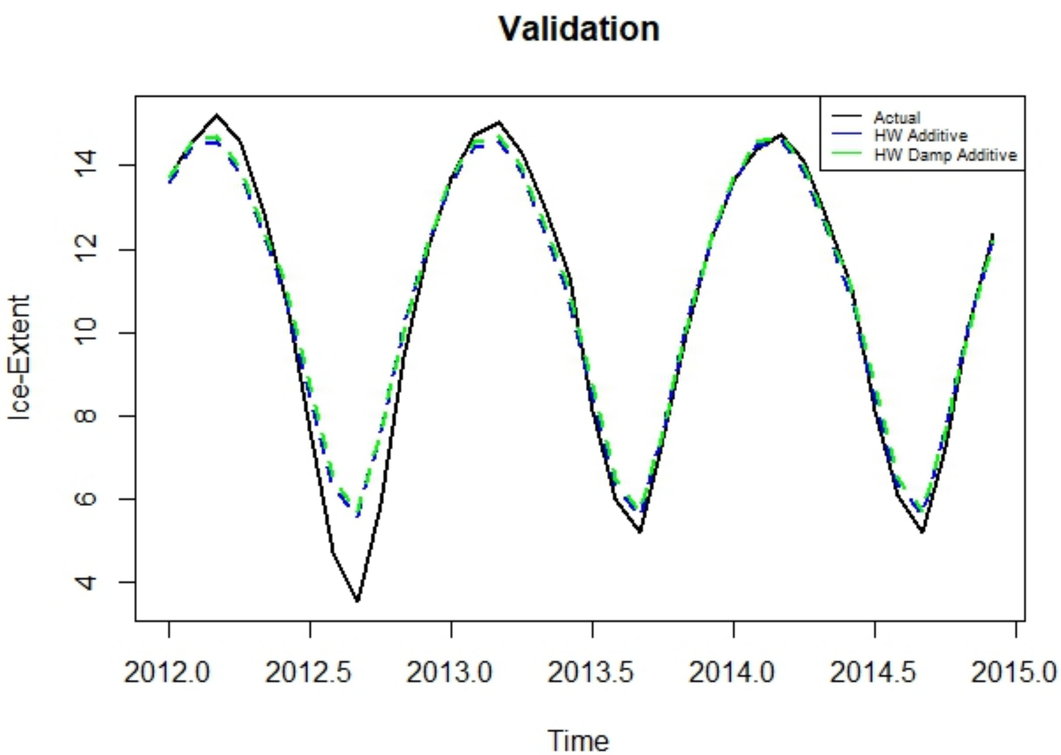
Model Fit & Residual Assessment

3 Estimation & Evaluation – Building Models & Model Assessment

	Smoothing Parameters	Training Error Measures	Residual – Homoskedasticity, Normality & Autocorrelation checks
HW Additive	• Alpha: 0.88 • Beta: 1e-04 • Gamma: 0.06	• RMSE: 0.27 **Model Fit Measures** • AIC: 1366.8 • BIC: 1434.5	 
HW Multiplicative	• Alpha: 0.02 • Beta: 1e-04 • Gamma: 0.39	• RMSE: 0.35 **Model Fit Measures** • AIC: 1770.9 • BIC: 1838.6	 
HW Damped Additive	• Alpha: 0.99 • Beta: .01 • Gamma: 1e-04 • Phi: 0.97	• RMSE: 0.26 **Model Fit Measures** • AIC: 1349.4 • BIC: 1421.1	 

Validation – Holt Winters

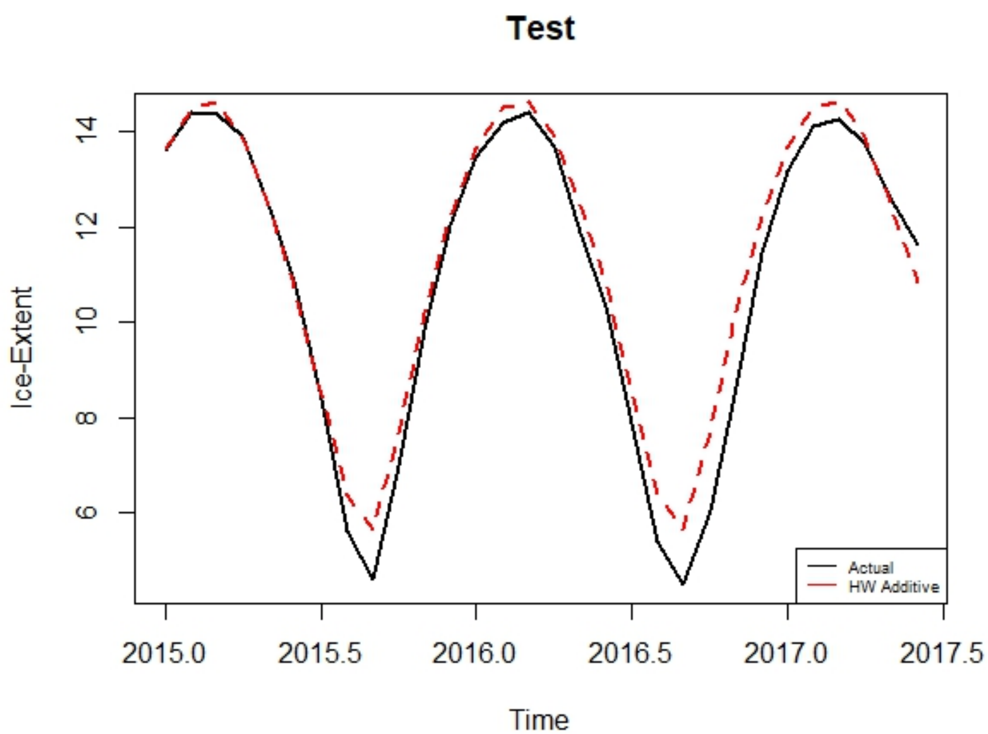
4 Validation



Model	Validation Error Measures
Holt Winters Additive	RMSE: 0.64
Holt Winters Damped Additive	RMSE: 0.66

Forecast – Holt Winters

5 Forecasting



Model	Test Error Measures
Holt Winters Additive	RMSE: 0.66

Model 2: SARIMA

1 Model – Determination, Specification & Evaluation

Model Parameters

ARIMA(1,0,1)(0,1,1)[12] with drift

- ar1: 0.5966
- ma1: 0.2157
- sma1: -0.7956

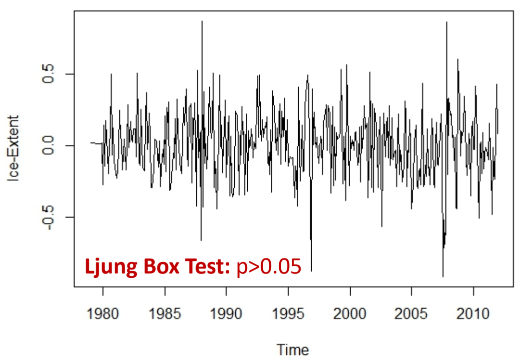
Training Error Measures

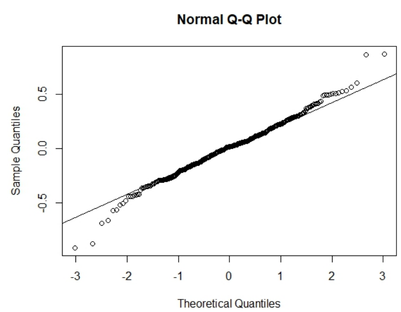
- RMSE: 0.24

Model Fit Measures

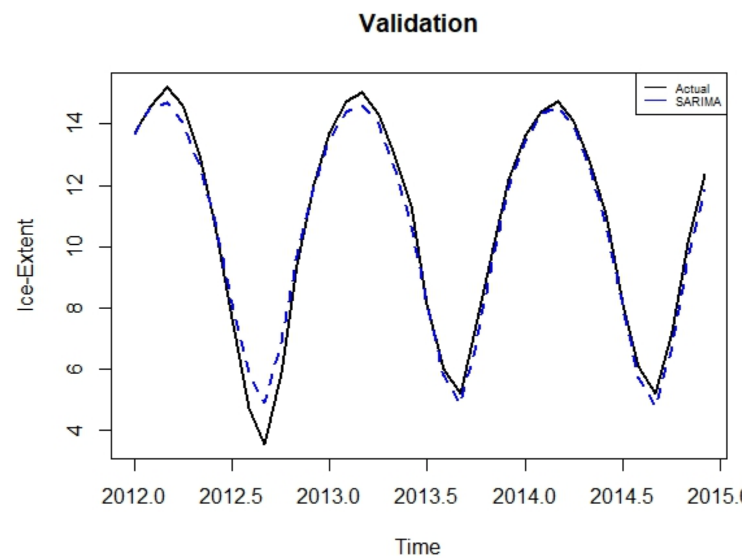
- AIC: 24.88
- BIC: 44.63

2 Residual – Homoskedasticity, Normality & Autocorrelation checks





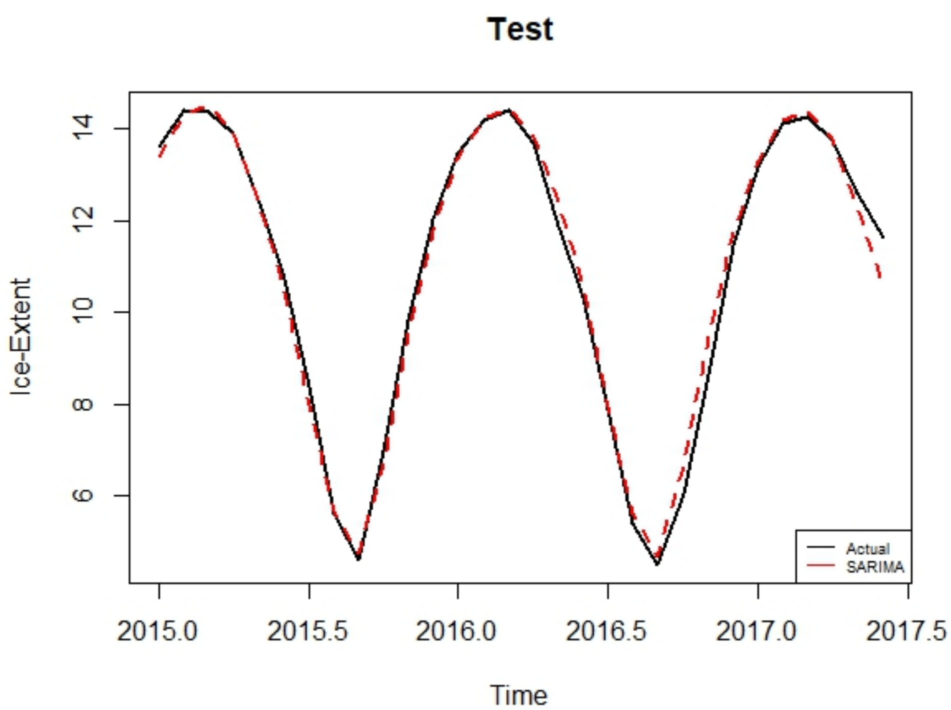
3 Model – Validation



Error Measure
RMSE: 0.47

Forecast – SARIMA

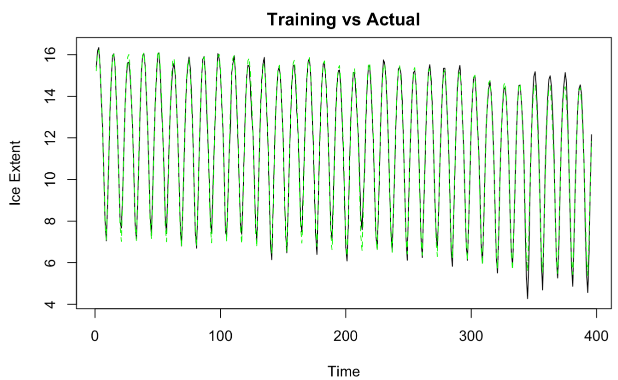
4 Forecasting



Model	Test Error Measures
SARIMA	RMSE: 0.34

Model 3: PROPHET

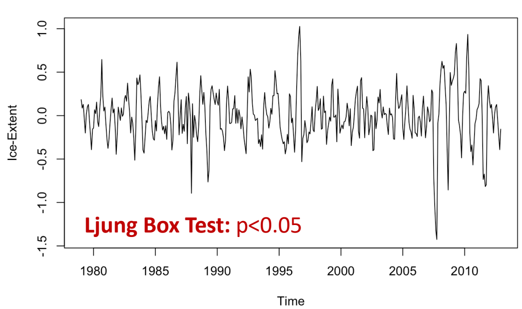
1 Model – Determination, Specification & Evaluation

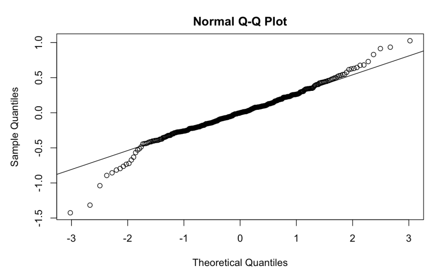


Training Error Measures

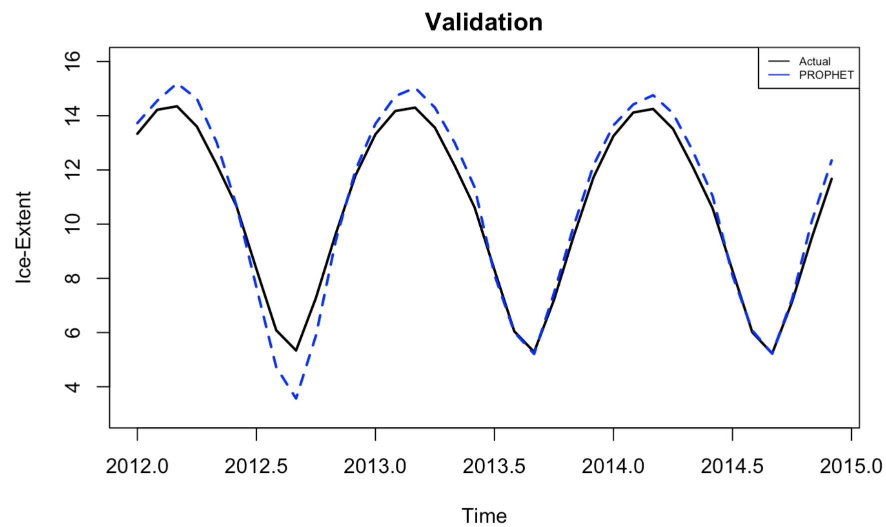
- RMSE: 0.31

2 Residual – Homoskedasticity, Normality & Autocorrelation checks





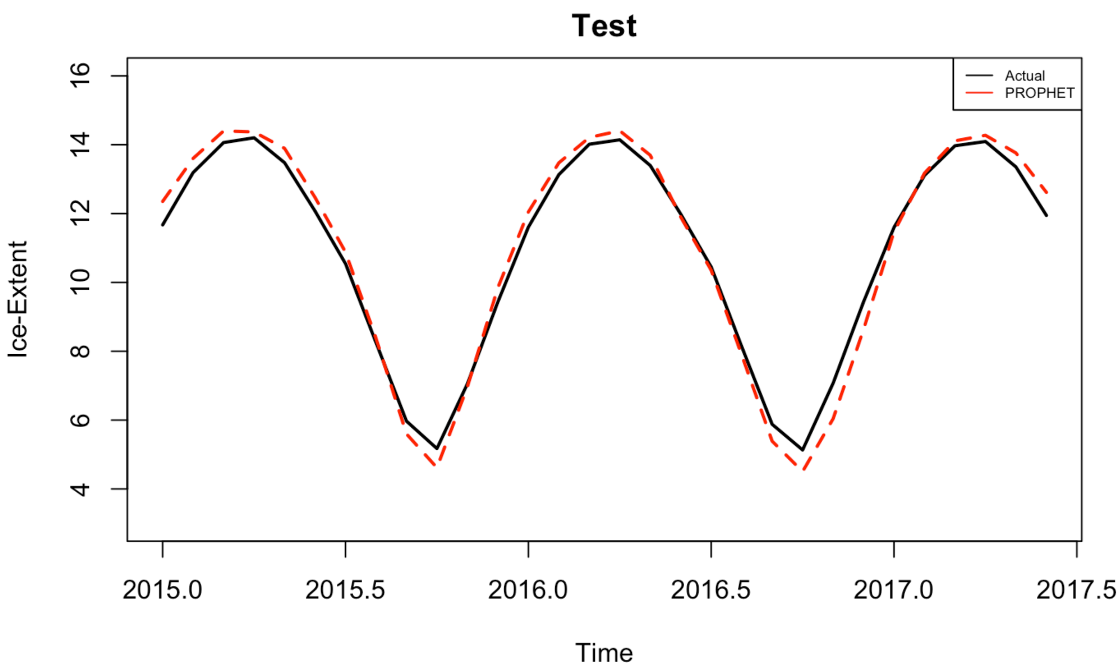
3 Model – Validation



Error Measure
RMSE: 0.66

Forecast – PROPHET

4 Forecasting

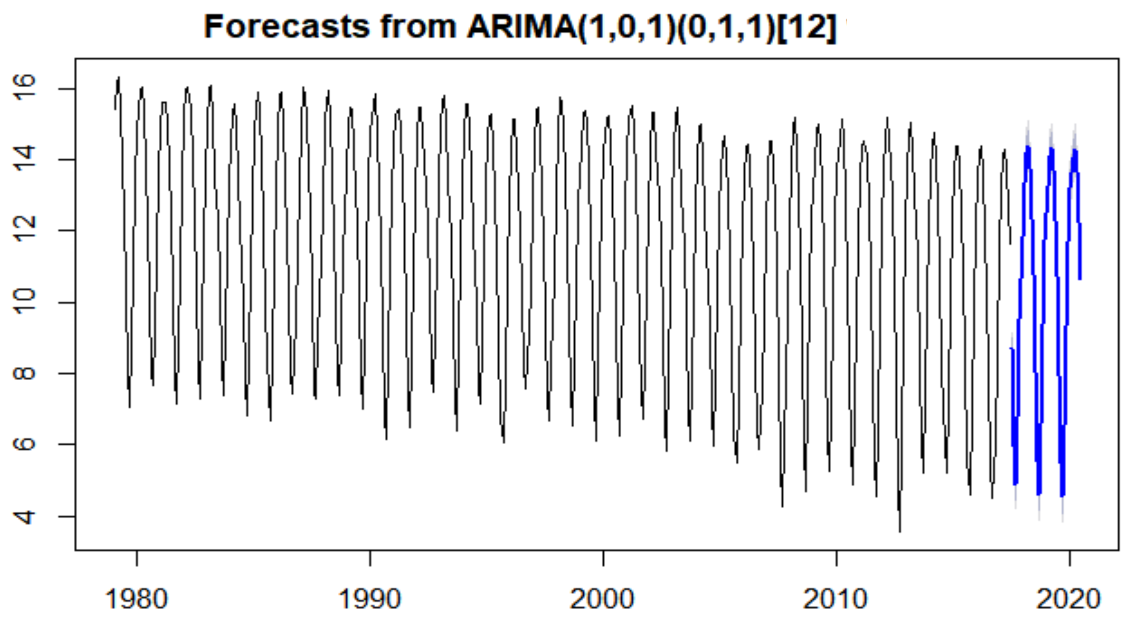


Model	Test Error Measures
PROPHET	RMSE: 0.42

Model Selection: SARIMA performs the best

Model	Train RMSE	Validation RMSE	Test RMSE	Ljung-Box (p-value)
Holt Winters Additive	0.27	0.64	0.65	>0.05
Seasonal ARIMA	0.24	0.47	0.34	>0.05
Prophet	0.31	0.66	0.42	<0.05

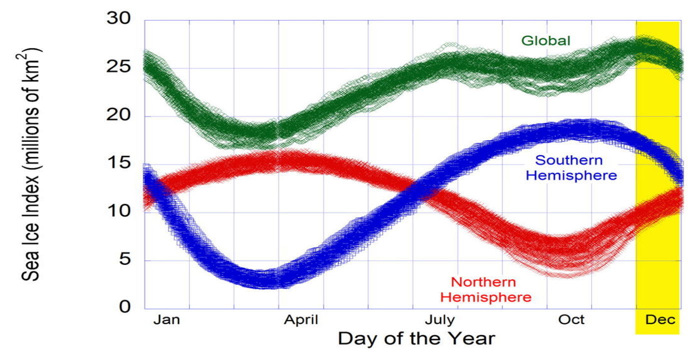
Forecast for 3 Years ahead – July 2017 to June 2020



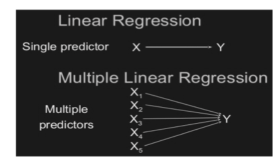
Conclusion and Future Consideration

The consequences of global warming are already here and we must continue to confront the realities of climate change. Continued research and analysis on the causes of global warming is essential to help lobby in support of "green" policy initiatives.

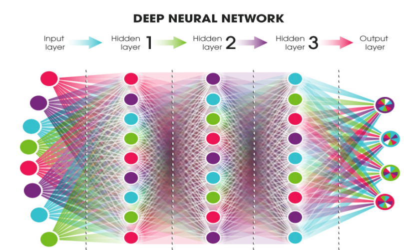
1. For comparative purposes, we could benefit from incorporating the Southern Hemisphere (Antarctic Region)



2. Adding additional variables, such as weather data (air/sea temperature, rain, winds, etc.) could also improve the model



3. Try Deep Learning (RNN/LSTM) to see if we get better predictions



Thank You!