Nicholas Milikich
ACMS 60855 Spatio-Temporal Statistics for Environmental Applications
March 5, 2020

Midterm 1

Exercise for All

This report is concerned with the sea ice extent in Baffin Bay, the body of water separating Baffin island of Canada from Greenland [1]. The dataset details the average monthly sea ice extent in Baffin bay, measured in square kilometers. It is taken from the National Snow and Ice Data Center FTP Archives provided by the National Snow and Ice Data Center through the University of Colorado Boulder [2].

The data on sea ice extent is collected using satellite microwave imaging [2]. It is possible to calculate the concentration of sea ice in a region of the ocean from brightness temperature data taken by microwave imaging because water and sea ice have different passive microwave brightness temperature signatures. Sea ice extent is calculated by dividing the body of water of interest into a grid and calculating the concentration of sea ice in each cell; the cell is considered to be "ice-covered" if the concentration of sea ice in that cell is at least 15%, and not ice-covered if not [3]. The sea ice extent of the body of water is the total area of the cells that are considered to be ice-covered.

Data on sea ice extent in Baffin Bay began being collected in October of 1978, but the collection has not remained constant. The first instrument used, starting in October 1978, was a Nimbus-7 Scanning Multichannel Microwave Radiometer, but beginning in August 1987, Special Sensor Microwave Imager and Imager/Sounder satellite instruments became available [2]. Before this change, measurements were only recorded every other day, but daily after the switch [2]. Additionally, some measurements of sea ice area (which is calculated slightly differently than sea ice extent) of nearby bodies of water at times around this switch are considered unreliable (due to the instruments' different handling of the arctic pole hole) [2]. However, there are no indications from the providers of the data that these considerations should affect the reliability of the data.

However, a few concerns about the reliability of the data remain. Due to satellite problems, data were not recorded from December 3, 1987 to January 13, 1988 [2]. These days are simply excluded when calculating the monthly average sea ice extent. This leads to smaller sample sizes for those months which may not be representative of the entire month. For example, if the sea ice extent tends to increase throughout the month of December, then using the values for December 1-3 only would tend to underrepresent the true average sea ice extent for the month. Similarly, the data for October, 1978 and February, 2020 are incomplete: data collection began on October 28, 1978 [2], and data was truncated for collection on February 23, 2020, which might impart bias to the average sea ice extent for those months.

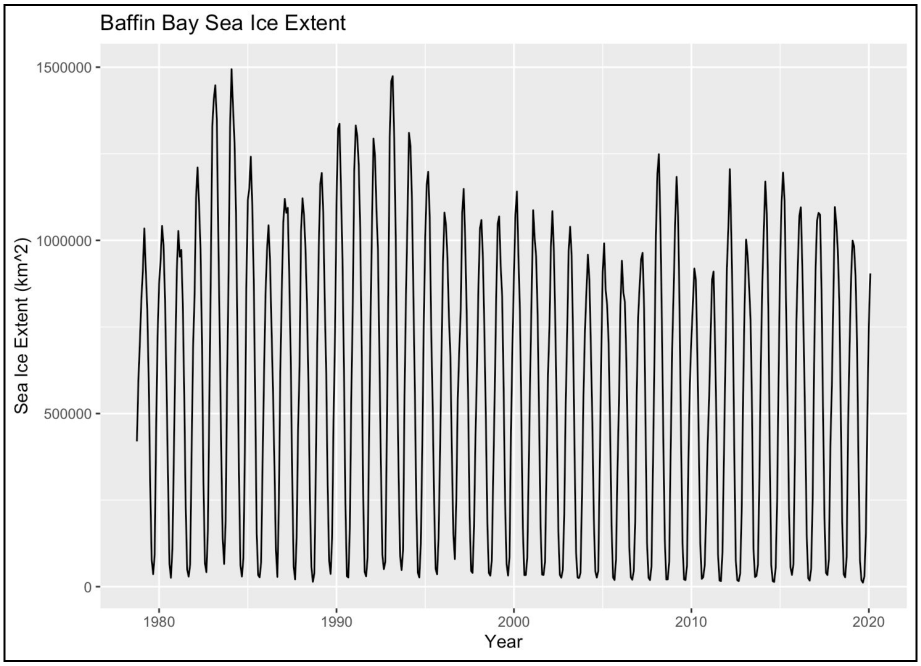


Figure 1. Average monthly sea ice extent in Baffin Bay, measured in km^2, from October, 1978 to February, 2020.

The Baffin Bay sea ice extent data are graphed in figure 1. The data include a monthly average for each month from October, 1978 to February, 2020. From this graph, none of the concerns raised above present obvious discrepancies in the data. The data for October, 1978 and February, 2020 do not noticeably differ from the trend; nor do the data for December, 1987 or January, 1988. These points are kept in the data for model fitting and forecasting.

Visually, the data appear to follow a harmonic seasonal pattern, and appear to be trending downward over time. These trends are investigated by aggregating the data monthly and annually (figure 2). For the visualization of the inter-annual trend, 1978 and 2020 are not included since there is not data available for all months of those years and the annual average is therefore biased. (The annual average for 2020 is the highest of all available years, contrary to the downward trend, since it only includes the months of January and February, winter months when ice cover tends to be highest.) Data from these years is however included for model fitting and forecasting.

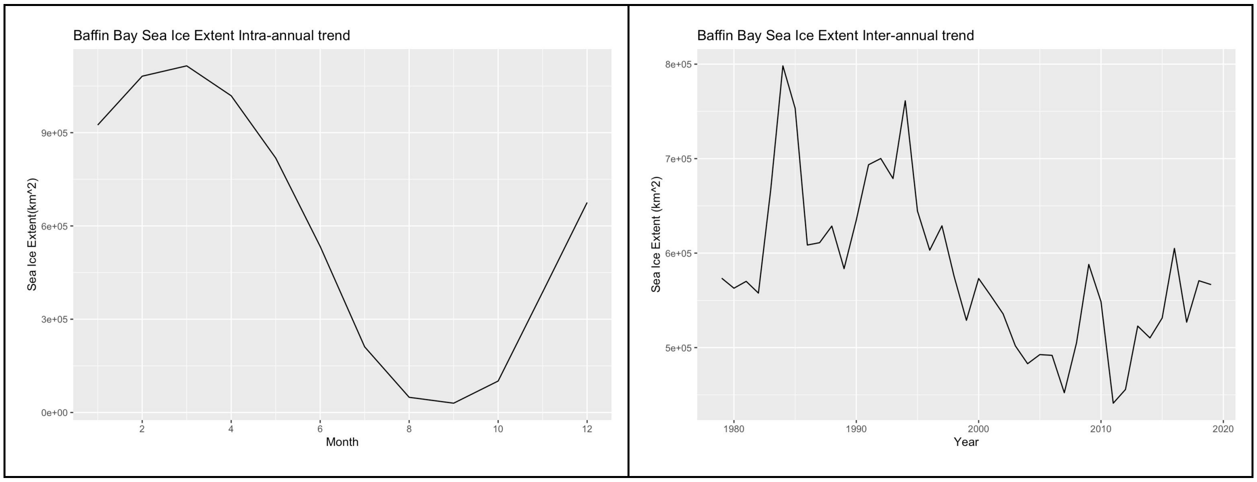


Figure 2. Intra-annual (left) and inter-annual (right) trends for average monthly sea ice extent in Baffin Bay. The intra-annual trend examines the average value for each month across all available years, and the inter-annual trend examines the average value for all months within each year. For the graph of the inter-annual trend, only complete years are included and graphed (i.e. 1978 and 2020 are not included when visualizing the trend since there is not complete data for those years).

The results of aggregating the data monthly (left) and annually (right) are shown in figure 2. These graphs confirm the presence of a harmonic seasonal trend and a general downward annual trend. The seasonal trend is smooth and sinusoidal, while the annual trend is more noisy and does not follow an obvious parameterizable shape.

The shape of the trends make intuitive sense in the context of the data. Intra-annually, sea ice extent follows a sinusoidal pattern peaking in March and reaching its lowest point in September. Baffin Bay is located in the northern hemisphere, where winter occurs between the months of December and March and summer occurs between the months of June and September. The extent of sea ice increases during the coldest months of the year until March, when the weather begins to warm again and the sea ice begins to melt. Inter-annually, sea ice extent in Baffin Bay has tended to decrease since 1978. During this time, human greenhouse gas emissions have caused the earth's average temperature to increase, reducing the average sea ice extent.

After confirming the presence of seasonal and yearly trends, these trends can be captured so that the remaining information in the time series is stationary and can be captured with a time series model. There being no other obvious choice for the shape of the trend, a linear trend was fit and found to be significant (R output included in appendix). The seasonality was also found to be significant. For the first harmonic, both terms are significant, and for the second harmonic, one term is significant (see appendix for R output); the second harmonic is included since one term is significant and since the time series is quite long so adding parameters is not a concern. The third harmonic is not included as only one term is marginally significant. The form of the equation chosen to fit the model is shown in equation 1. This formulation includes an intercept, a term that is linear with time, and harmonic terms up to order two.

$$y_t = a_0 + b_0 t + a_1 cos(\omega_0 t) + b_1 sin(\omega_0 t) + a_2 cos(2\omega_0 t) + b_2 sin(2\omega_0 t) \quad (1)$$

The fitted values for capturing the trend and seasonality are plotted in figure 3 (left). Visually, the fitted values appear to capture the overall trend of the data quite well. However, some problems with the model are apparent. Most significant is the fact that many fitted values are less than 0, a physical impossibility in the context of the data - it is impossible for the sea ice extent to be a negative area. One solution to this problem would be to set the floor for any forecasts to 0. This solution makes some intuitive sense in the context of the problem: sea ice extent is likely related to temperature, and the temperature can increase without sea ice extent following a proportional decrease (it can keep getting warmer, but the concentration of sea ice cannot get any lower once it all melts). If the temperature follows a sinusoidal pattern, then sea ice might reasonably follow a sinusoidal pattern with a floor at 0, as in this approximation. The mathematics of this solution will be discussed further later.

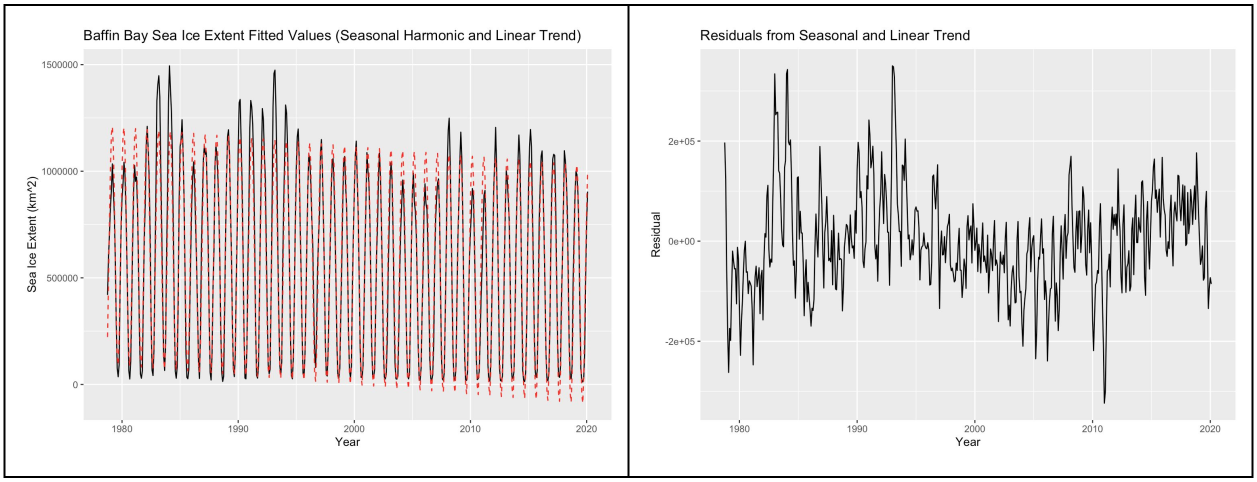


Figure 3. Left: fitted values for capturing trend and seasonality; the original data is shown in black and the fitted values are shown in dashed red. Right: residuals after removing the trend and seasonality shown in the left panel, i.e. information in the time series not captured by the trend and seasonality.

The residuals after removing the trend and seasonality from the data are shown in figure 3 (right). Detrending is intended to make the series stationary for further analysis, and the residual time series, based on visual inspection, could reasonably be assumed to be stationary. A non-seasonal ARMA model was fit to these residuals, and an ARMA(3,2) model was found to have the lowest AIC. The autocorrelation function of the residuals of this ARMA model is plotted in figure 4. The autocorrelation at lags 12 and 24 are significant, which is indicative of the fact that there is additional seasonality that could not be accounted for. (These autocorrelation values are still significant when additional harmonics are included in the seasonality.) Aside from these, no autocorrelation values are significant (lag 25 is marginally significant), indicating that the downward linear trend has been removed.

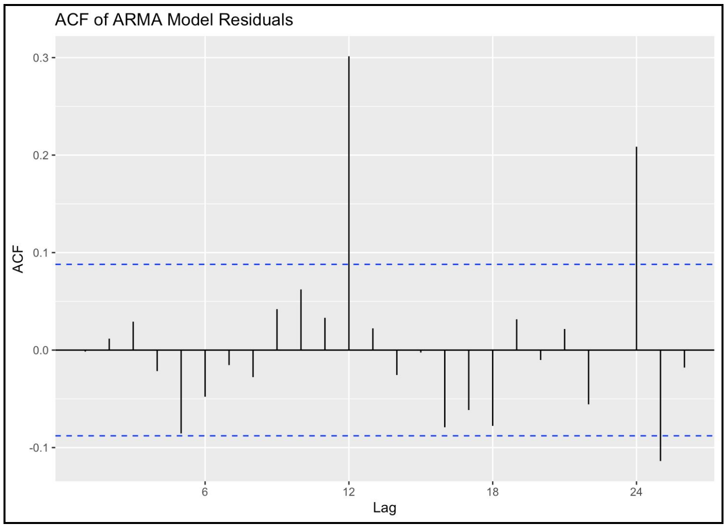


Figure 4. Autocorrelation function of the residuals of the ARMA(3,2) model which was fit to the residuals of the sea ice extent data after detrending.

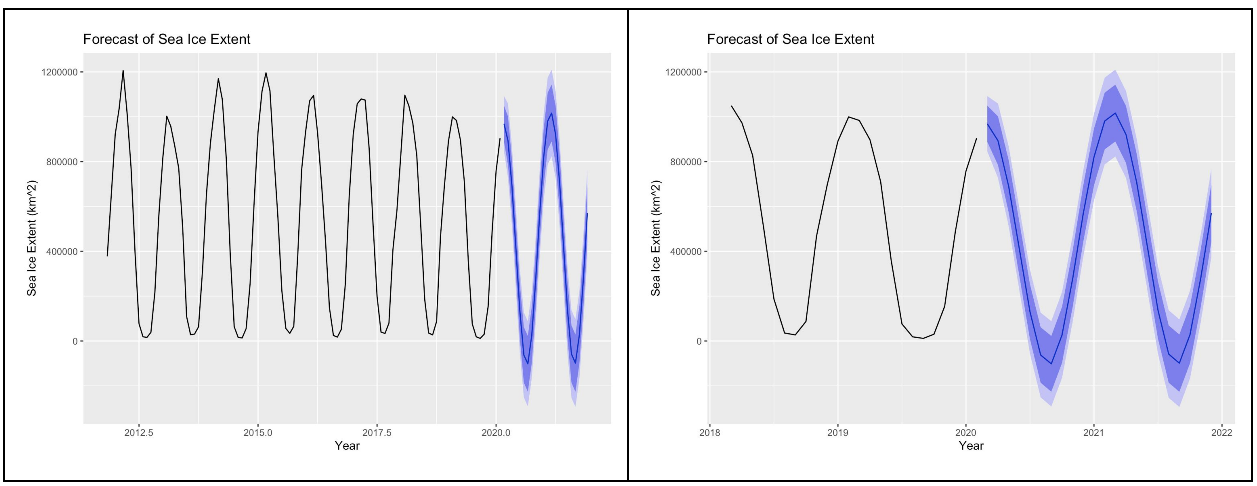


Figure 5. Forecasts for all months of 2020 and 2021 using the final model that is the summation of the trend and seasonality and the ARMA(3,2) model fit to the residuals of the data after detrending. Not all available data points are shown in these graphs; a complete graph is included in the appendix. Black lines represent point forecasts, dark blue bands represent 80% confidence intervals, and light blue bands represent 95% confidence intervals.

The final model is the summation of the trend and seasonality and the ARMA(3,2) model fit to the residuals after detrending, and this model can be used to forecast future values of sea ice extent. These forecasts are shown in figure 5. The dark blue band represents an 80% confidence interval for the forecasts, and the light blue band represents a 95% confidence interval for the forecasts. It is noteworthy that several predictions and confidence intervals include values less than 0, a physical impossibility. As mentioned, one solution would be to set the floor of any forecasts to 0. However, this solution becomes undesirable when considering predictions as probability distributions rather than point forecasts. It might be possible to truncate each probability distribution at 0 and renormalize, but then the shapes of forecast distributions would be inconsistent between neighboring points, which is undesirable. Solving this problem would require reformulating the model in a way that would make negative outputs impossible, or that would make asymmetric and constrained forecast distributions possible. If one is to accept the model as is, one must be willing to accept some unreliability in forecasts of small values.

There is also concern about the reliability of the predictions due to the behavior of the function. The left panel of figure 5 highlights the fact that each summer, sea ice extent decreases to approximately the same minimum value, while each winter it increases to a maximum that varies but has tended to decrease over time. This trend can also be seen in figure 1. This has the overall effect of heteroscedasticity, of decreasing the variance in measurements over time. The right panel of figure 3 is also suggestive of heteroscedasticity, as the magnitude of the residuals after detrending appears to decrease somewhat over time. The ACF of the residuals of the ARMA model (figure 4) corroborates this, suggesting that there is additional information in the residuals available to be captured. All of this together suggests that a more complex form of model may be required to capture the shape of the data. Taking the model as is, the presence of heteroscedasticity casts some doubt on the reliability of the forecast distributions, especially those at the minimums (during summer months). Visually, the forecast distributions for the peaks (winter months) appear to be reasonably in line with the trend and past years (figure 5), while the forecast distributions for troughs (summer months) are noticeably out of line from past values.

Conclusion

This report is concerned with the sea ice extent in Baffin Bay, the arctic body of water separating Baffin island of Canada from Greenland. Sea ice extent in Baffin Bay is measured from orbiting satellites, which use microwave imaging to determine which regions of the ocean are ice-covered. Sea ice extent is a measure of the total area of the bay that is ice-covered. The satellite data provides the average sea ice extent in Baffin Bay for each month from October, 1978 to February, 2020. Sea ice extent fluctuates seasonally, peaking at the end of each winter and bottoming out at the end of each summer. However, even after removing the effect of seasonal fluctuations, there was found to be a statistically significant downward trend in the average monthly sea ice extent over time, and this trend was forecasted to continue in the next two years. This particular study did not investigate the causes of this trend, only the presence and nature of the trend itself, but the most likely cause is climate change due to greenhouse gas emissions during this time period.

Exercises for Graduate Students

When performing forecasting on time series data, it is necessary to quantify the uncertainty in the predictions. The prediction for each time point in the future is not a point forecast alone, but a probability distribution of the possible values the process might obtain given the process parameters estimated from the available information. However, from this description, one can see a potential problem: the forecast probability distributions are computed from the *estimated* parameters of the process, while the true values the process will adopt at those time points follow the *true* parameters of the underlying process. It is worth examining whether this assumption affects the reliability of the forecast distributions.

To examine the effect that this assumption has on forecast distributions, a simulation was performed. First, a time series of 1,000 time points $X_1, \ldots, X_{1000}$ was generated from a known ARMA process with fixed parameters. This time series served as the "true" data-generating process. The parameters for the model were estimated by fitting an ARMA model to time points $X_1, \ldots, X_{999}$ (it was assumed that the order of the data-generating process was known, but not the values of the parameters). From this estimated model, a 95% confidence interval for $\widehat{X}_{1000}$ was calculated. This process was repeated 1,000 times, and the proportion of trials in which the 95% confidence interval encompassed the true X_{1000} value was calculated. This entire process was repeated four times, with four different sets of known ARMA process parameters (listed in "Parameters" column of figure 6).

Order of Data-Generating ARMA Process	**True Parameters of Data-Generating ARMA Process**	**Proportion of 95% Confidence Intervals Encompassing True Value**
(1,0)	$\varphi_1 = 0.8$	0.951
(0,1)	$\theta_1 = 0.8$	0.947
(1,1)	$\varphi_1 = 0.6$ $\theta_1 = -0.2$	0.943
(2,2)	$\varphi_1 = -0.4$ $\varphi_2 = -0.6$ $\theta_1 = 0.8$ $\theta_2 = -0.2$	0.960

Figure 6. Results of the simulation study. The first two columns give the order and parameters for the true ARMA process that generates $X_1, \ldots, X_{1000}$. The third column gives the proportion, out of 1,000 simulations, of 95% confidence intervals for the forecast $\widehat{X}_{1000}$ (calculated from an ARMA model fit to $X_1, \ldots, X_{999}$) that encompass the true value X_{1000}.

The results of the described simulation study are summarized in figure 6. One would expect the proportion of 95% confidence intervals encompassing the true value, by definition, to be 0.95. This expectation is confirmed in the right column of figure 6: in all processes, the proportion of 95% confidence intervals that included the true value was reasonably close to 0.95. Therefore, from this simulation study, one can conclude that the effect of assuming the estimated parameters to be the true parameters is negligible. This assumption is, in a way, similar to the assumption one makes when performing hypothesis testing or statistical inference using the Wald statistic, that the standard error of the true process is equal to the standard error of the sample statistic. This assumption is generally reasonable and the Wald statistic normally gives consistent results. The effect of assuming the true parameters are equal to the estimated parameters is similarly often negligible and in practice does not often affect results or reliability of forecasts.

It is important to keep in mind that this conclusion is based off of a constrained example and may not be generalizable to all cases. In this simulation, the true data-generating process was exactly an ARMA process, the order of which was known. In practice, natural processes are not likely to follow models so exactly. An ARMA model is meant to be a simplification or approximation of the true processes and may not accurately capture the entire process, in which case assuming the estimated parameters to accurately represent the true process may have a larger effect.

Sources

1. Walter N. Meier, Julienne Stroeve, and Florence Fetterer. Whither arctic sea ice? A clear signal of decline regionally, seasonally and extending beyond the satellite record. Annals of Glaciology, 46:428-434, 2007.
2. Fetterer, F., K. Knowles, W. N. Meier, M. Savoie, and A. K. Windnagel. 2017, updated daily. Sea Ice Index, Version 3.0. Boulder, Colorado USA. NSIDC: National Snow and Ice Data Center. doi: https://doi.org/10.7265/N5K072F8. March 5, 2020.
3. Frequently Asked Questions on Arctic sea ice (2008) National Snow & Ice Data Center. URL https://nsidc.org/arcticseaicenews/faq/#area_extent. March 5, 2020.

Appendix - R Codes

```
rm(list=ls())

library(ggplot2)
library(forecast)

this.dir = dirname(rstudioapi::getActiveDocumentContext()$path)
setwd(this.dir)

load("SeaIce.Rdata")

##################
# Exercise for all
##################

gp = autoplot(y.ts) + ggtitle("Baffin Bay Sea Ice Extent") +
xlab("Year") + ylab("Sea Ice Extent (km^2)")
gp

# Exploratory analysis

seaice.monthly = aggregate(c(y.ts), list(month = cycle(y.ts)), mean)
seaice.annual = aggregate(c(y.ts), list(year = floor(time(y.ts))),
mean)

seaice.monthly = ts(seaice.monthly[,2], start = 1)
autoplot(seaice.monthly) + ggtitle("Baffin Bay Sea Ice Extent
Intra-annual trend") + xlab("Month") + ylab("Sea Ice Extent(km^2)")

seaice.annual = ts(seaice.annual[,2], start = seaice.annual[2,1], end
= seaice.annual[42,1])
autoplot(seaice.annual) + ggtitle("Baffin Bay Sea Ice Extent
Inter-annual trend") + ylab("Sea Ice Extent (km^2)") + xlab("Year")

# Detrending seasonality and trend

year.ts = floor(time(y.ts))

X1 = fourier(y.ts, K = 1)
X2 = fourier(y.ts, K = 2)
X3 = fourier(y.ts, K = 3)
```

```
mod.h1 = tslm(y.ts ~ X1 + year.ts)
summary(mod.h1)
mod.h2 = tslm(y.ts ~ X2 + year.ts)
summary(mod.h2)
mod.h3 = tslm(y.ts ~ X3 + year.ts)
summary(mod.h3)

# Plot the fitted values

df = data.frame(X = time(y.ts), Y = mod.h2$fitted.values)
gp + geom_line(data = df, aes(X,Y), color = "red", linetype =
"dashed") + ggtitle("Baffin Bay Sea Ice Extent Fitted Values
(Seasonal Harmonic and Linear Trend)")

# Store and plot the residuals

res.ts = y.ts - mod.h2$fitted.values
autoplot(res.ts) + ggtitle("Residuals from Seasonal and Linear
Trend") + xlab("Year") + ylab("Residual")

# Fit a non-seasonal ARMA model to the residuals
# (the information left over after detrending)

seaice.res.fit = auto.arima(res.ts, d = 0, seasonal = FALSE, max.p =
4, max.q = 4, stepwise = FALSE)
summary(seaice.res.fit)
aic = NULL
for (p in 0:3)
{
  for (q in 0:3)
  {
    dummy = arima(res.ts, order = c(p, 0, q), include.mean = FALSE)
    aic = rbind(aic,cbind(p, q, dummy$aic))
  }
}
aic

ggAcf(seaice.res.fit$residuals) + ggtitle("ACF of ARMA Model
Residuals")

# Forecast for 2020-2021

horizon = 22
```

```
forc = forecast(seaice.res.fit, h = horizon, level = c(80, 95))

# Add back the trend

year.ts = c(kronecker(2019,rep(1,10)), kronecker(2020,rep(1,12)))
trend.predict = forecast(mod.h2, data.frame(fourier(y.ts, K = 2, h =
horizon), year.ts))

forc$x = forc$x + mod.h2$fitted.values
forc$mean = forc$mean + trend.predict$mean
forc$lower = forc$lower + trend.predict$mean
forc$upper = forc$upper + trend.predict$mean

# Plot final results
autoplot(forc) + xlab("Year") + ylab("Sea Ice Extent (km^2)") +
ggtitle("Forecast of Sea Ice Extent")
autoplot(forc, include = 100) + xlab("Year") + ylab("Sea Ice Extent
(km^2)") + ggtitle("Forecast of Sea Ice Extent")
autoplot(forc, include = 24) + xlab("Year") + ylab("Sea Ice Extent
(km^2)") + ggtitle("Forecast of Sea Ice Extent")

################################
# Exercise for graduate students
################################

set.seed(1)

len = 1000
nsim = 1000

results = NULL

# Test 1 - ARMA(1,0) where phi = 0.8

successes = 0
for (n in 1:nsim)
{
  test = arima.sim(model = list(ar = 0.8), n = len)
  fit = arima(test[1:(nsim-1)], order = c(1,0,0), include.mean =
FALSE)
  check = forecast(fit, h = 1, level = 95)
  check.low = check$lower[1]
  check.high = check$upper[1]
```

```
  if (check.low <= test[len] & test[len] <= check.high)
    successes = successes + 1
}
results = rbind(results, cbind(1, 0, successes/nsim))

# Test 2 - ARMA(0,1) where theta = 0.8

successes = 0
for (n in 1:nsim)
{
  test = arima.sim(model = list(ma = 0.8), n = len)
  fit = arima(test[1:(nsim-1)], order = c(0,0,1), include.mean =
FALSE)
  check = forecast(fit, h = 1, level = 95)
  check.low = check$lower[1]
  check.high = check$upper[1]
  if (check.low <= test[len] & test[len] <= check.high)
    successes = successes + 1
}
results = rbind(results, cbind(0, 1, successes/nsim))

# Test 3 - ARMA(1,1) where phi = 0.6 and theta = -0.2

successes = 0
for (n in 1:nsim)
{
  test = arima.sim(model = list(ar = 0.6, ma = -0.2), n = len)
  fit = arima(test[1:(nsim-1)], order = c(1,0,1), include.mean =
FALSE)
  check = forecast(fit, h = 1, level = 95)
  check.low = check$lower[1]
  check.high = check$upper[1]
  if (check.low <= test[len] & test[len] <= check.high)
    successes = successes + 1
}
results = rbind(results, cbind(1, 1, successes/nsim))

# Test 4 - ARMA(2,2) where phi1 = -0.4, phi2 = -0.6, theta1 = 0.8,
and theta2 = -0.2

successes = 0
for (n in 1:nsim)
{
```

```
  test = arima.sim(model = list(ar = c(-0.4, -0.6), ma = c(0.8,
-0.2)), n = len)
  fit = arima(test[1:(nsim-1)], order = c(2,0,2), include.mean =
FALSE)
  check = forecast(fit, h = 1, level = 95)
  check.low = check$lower[1]
  check.high = check$upper[1]
  if (check.low <= test[len] & test[len] <= check.high)
    successes = successes + 1
}
results = rbind(results, cbind(2, 2, successes/nsim))

results
```

Appendix - R Outputs

Exercise for All

```
> summary(mod.h1)

Call:
tslm(formula = y.ts ~ X1 + year.ts)

Residuals:
    Min      1Q  Median      3Q     Max 
-320891  -62854   -4633   54467  352924 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) 9544411.7   761303.6   12.54   <2e-16 ***
X1S1-12       73546.4     6434.5   11.43   <2e-16 ***
X1C1-12     -558875.2     6447.7  -86.68   <2e-16 ***
year.ts       -4485.1      380.8  -11.78   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 101500 on 493 degrees of freedom
Multiple R-squared:  0.9402,    Adjusted R-squared:  0.9399 
F-statistic:  2586 on 3 and 493 DF,  p-value: < 2.2e-16

> summary(mod.h2)

Call:
tslm(formula = y.ts ~ X2 + year.ts)

Residuals:
    Min      1Q  Median      3Q     Max 
-323766  -61143   -7569   56502  350277 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) 9519150.2   755881.9  12.593   <2e-16 ***
X2S1-12       73448.8     6387.8  11.498   <2e-16 ***
X2C1-12     -558940.9     6401.0 -87.321   <2e-16 ***
X2S2-12        7199.6     6388.5   1.127   0.2603    
X2C2-12      -18089.0     6400.3  -2.826   0.0049 ** 
year.ts       -4472.5      378.1 -11.828   <2e-16 ***
```

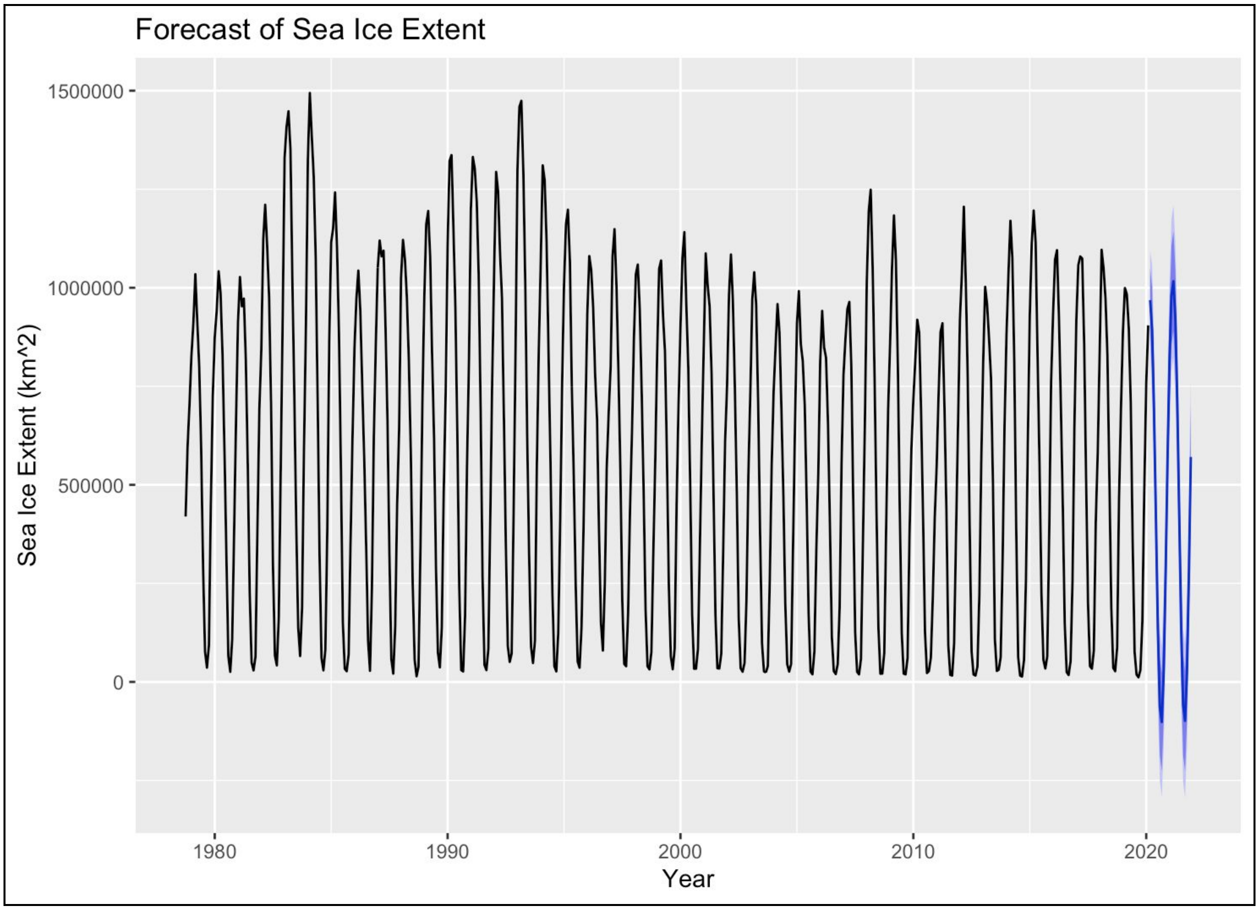
```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 100800 on 491 degrees of freedom
Multiple R-squared:  0.9414,    Adjusted R-squared:  0.9408
F-statistic:  1576 on 5 and 491 DF,  p-value: < 2.2e-16

> summary(mod.h3)

Call:
tslm(formula = y.ts ~ X3 + year.ts)

Residuals:
    Min      1Q  Median      3Q     Max
-334256  -59551   -6042   57373  352473

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 9529148.8   755028.1  12.621  < 2e-16 ***
X3S1-12       73448.4     6380.2  11.512  < 2e-16 ***
X3C1-12     -558899.9     6393.4 -87.418  < 2e-16 ***
X3S2-12        7271.8     6381.0   1.140  0.25501
X3C2-12      -18056.1     6392.9  -2.824  0.00493 **
X3S3-12       -4031.2     6380.0  -0.632  0.52778
X3C3-12       10642.8     6393.0   1.665  0.09660 .
year.ts       -4477.5      377.7 -11.855  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 100700 on 489 degrees of freedom
Multiple R-squared:  0.9417,    Adjusted R-squared:  0.9409
F-statistic:  1129 on 7 and 489 DF,  p-value: < 2.2e-16

> summary(seaice.res.fit)

Series: res.ts
ARIMA(3,0,2) with zero mean

Coefficients:
         ar1      ar2      ar3      ma1      ma2
      1.2551  -0.1289  -0.1534  -0.3555  -0.4812
s.e.  0.1563   0.2596   0.1226   0.1442   0.1340
```

```
sigma^2 estimated as 3.96e+09:  log likelihood=-6194.97
AIC=12401.93   AICc=12402.11   BIC=12427.19

Training set error measures:
                     ME     RMSE      MAE      MPE     MAPE      MASE
ACF1
Training set -332.8837 62612.62 48379.79 160.8815 314.4088 0.6357638
-0.001536567

> aic

      p q
 [1,] 0 0 12858.05
 [2,] 0 1 12529.76
 [3,] 0 2 12458.52
 [4,] 0 3 12430.68
 [5,] 1 0 12419.13
 [6,] 1 1 12405.34
 [7,] 1 2 12403.02
 [8,] 1 3 12403.95
 [9,] 2 0 12409.69
[10,] 2 1 12402.77
[11,] 2 2 12404.48
[12,] 2 3 12405.76
[13,] 3 0 12406.06
[14,] 3 1 12404.36
[15,] 3 2 12401.93
[16,] 3 3 12403.86
```

Exercises for Graduate Students

```
> results

     [,1] [,2]  [,3]
[1,]    1    0 0.951
[2,]    0    1 0.947
[3,]    1    1 0.943
[4,]    2    2 0.960
```

Appendix - Additional Graphs



Forecasts for all months of 2020 and 2021 using the final model that is the summation of the trend and seasonality and the ARMA(3,2) model fit to the residuals of the data after detrending. Forecasts are the same as in figure 5, but all available data is plotted for context.