AMSI VACATIONRESEARCH SCHOLARSHIPS 2022–23

SET YOUR SIGHTS ON RESEARCH THIS SUMMER

Non-stationary dynamics of climate statistics and its visualization

Yihong Mei

Supervised by Sourav Das

James Cook University

AMSI AUSTRALIAN MATHEMATICAL SCIENCES INSTITUTE

Content

Abstract

The project was created using the Time Series model to simulate the pattern of glacier changes over time. The model was used to analyze the reasons for changes in glacier mass over time in Antarctica and Greenland and whether there is an interaction between the two glaciers. The model was also used to predict the change in glacier mass over time.

1 Introduce

The loss of mass of polar ice Sheet are some of the most important diagnostic tools of global warming, that are related to climate change. The Greenland and Antarctic ice sheets are the largest bodies of ice in the world and play an important role in the global climate system. Both ice sheets have been losing mass at an alarming rate, which has contributed one third of the global sea level rise over this period. It is important to find the main reason of mass losing by comparing those two places' data. We can use those data to analyse the same and different reasons of mass losing and whether there is an interaction between those two places.

After done a module in descriptive statistics, including Fisher's tests of significance, followed by a module in linear regression. The data on mass of polar ice –underlying its dynamics are monthly time series. To investigate this phenomena we would require statistical analysis of time series. And we can primarily investigate the following three things:

1. Statistical visualization of ice sheets changing in both locations.
2. Using regression methods to model the dynamics of ice sheets changing mass.
3. Compare the changing in the dynamics between the two locations.

2 Data

2.1 Data Collection

The original data was the real data of ice sheet data of Greenland and Antarctica which measured the ice sheet mass between the 2002 and 2022.The data is monthly time series information on the changing in the ice sheet mass measured in Gigatonnes. Data obtained from NASA's satellite GRACE measurements.

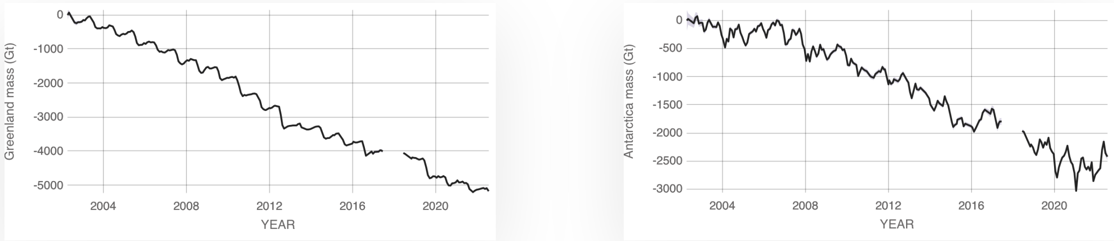

Figure 1: Trends in glacier mass over time in Greenland(Left) and Antarctica(Right)

2.2 Missing Value

There were missing values in data because NASA changed the satellites between 2017 and 2018 and there are also a few months that have not been measured.

Here we assume that observations are missing at random and are Gaussian. We can impute missing values using average of same months each year or use the observed data in the same month to model a linear regression to estimate the missing value. The following is the linear predictive equation:

$$Y_{m,t} = m_1 + m_2\, t + \; e_t$$

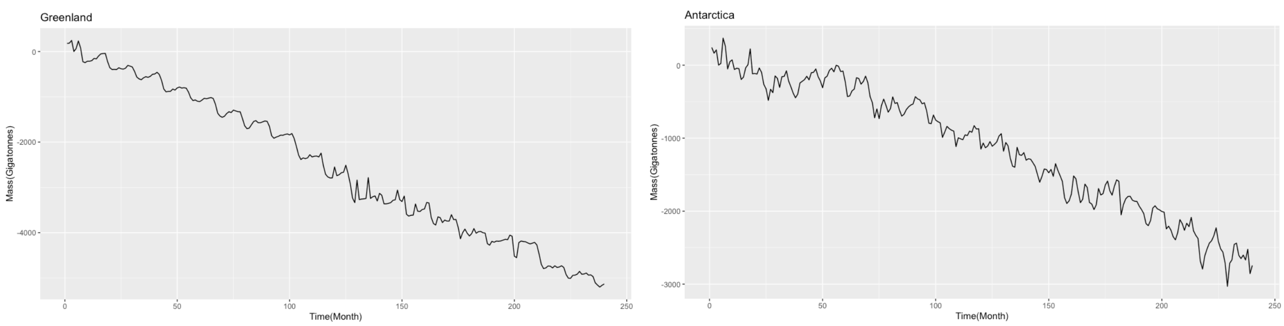

Figure 2: Trends in glacier mass over time in Greenland(Left) and Antarctica(Right) after filling the missing value

3 Model

3.1 Additive decomposition of Time Series

There are various patterns in time series data, and the main ones among these are trend, seasonality, and periodicity. When decomposing a time series into components, it is common to combine trend and period into a trend-period component (often referred to simply as trend). The time series can then be thought of as being made up of three components: trend (trend-periodic), seasonal and residual (the rest of the component beyond the first two, also known as the stochastic component).

There are two methods of classical decomposition:

Additive decomposition: $Y_t = T_t + S_t + R_t$

Multiplicative decomposition: $Y_t = T_t \times S_t \times R_t$

Where the T_t is the trend cyclical component, S_t is the seasonal component, and R_t is the remaining random (residual) component after removing the first two.

If the magnitude of seasonal fluctuations and changes in trend cycles do not vary with the level at which the time series is taken, an additive decomposition is appropriate. The multiplicative decomposition is more appropriate when the change in the seasonal model or the change in the trend of the cycle is proportional to the level at which the time series is taken (Shumway & Stoffer, 2017). As the seasonal variation in glacier mass stabilizes over time, we use additive decomposition.

3.2 Trend and Seasonality

To estimate the trend and seasonality of the annual time series we use linear functions of time (t), respectively. That is,

$$T_t = a_1 + b_1 t$$

$$S_t = b_2 \, Months$$

3.3 Residual

The residual is the value that remains after fitting the model. For most time series models, the value of the residual is equal to the difference between the observed value and the corresponding fitted value. The residuals can be used to test whether the model adequately captures the information in the time series and are an important way we can use to test whether the reduction in glacier mass in Greenland and Antarctica is intrinsically related.

We estimate the stochastic component R_t by subtracting estimates of trend and seasonality from the observed time series.

$$\hat{R}_t = Y_t - \hat{T}_t - \hat{S}_t$$

One of ways measuring the stochastic component is the serial autocorrelation. The serial or autocorrelation of lag h for a time series R_t at location s with n samples is defined as:

$$\rho_s(|h|) = \frac{\sum_{t=1}^{n-h}(X_t - \bar{X})(X_{t+h} - \bar{X})}{n - |h|}, \; h = 0,1, \dots, n-1$$

And we compute Pearson's correlation between the autocorrelations up to lag 24 (2 years), at the two locations.

$$\rho = \frac{\sum_{h=1}^{24}(\rho_s(|h|) - \overline{\rho_s})(\rho_u(|h|) - \overline{\rho_u})}{24} / (s_s \, s_u)$$

where,

$$s_u = \sqrt{\frac{\sum_{u=1}^{24}(\rho_u(|h|) - \overline{\rho_u})^2}{24 - 1}}$$

4 Analysis

4.1 Exploratory analysis

For trend, the overall trend of ice sheet mass is showing a decreasing trend: This is true, both the Antarctic and Greenland ice sheets have been losing mass in recent decades due to a combination of factors such as warmer ocean and air temperatures, changes in winds, and ice dynamics. The latest estimate shows that between 2002 and 2022, Antarctica lost near 3000 billion tons of ice, while Greenland lost near 5000 billion tons of ice.

The mass of Antarctica has not declined as much as that of Greenland: While it is true that the Antarctic ice sheet has lost less ice than the Greenland ice sheet, it is still losing ice at an accelerating rate. In fact, recent research shows that Antarctica is losing ice six times faster than it was in the 1990s (NASA, 2020).

For Seasonality, the mass of the ice sheet fluctuates over each year: This is also true, the mass of the ice sheet goes through seasonal variations due to changes in temperature and precipitation. In general, there is more snow accumulation during the winter months and more melting during the summer months.

The mass annual fluctuations in Greenland are greater than in Antarctica: This is partially true. While both ice sheets experience seasonal fluctuations, the amplitude of these fluctuations is generally larger in Greenland than in Antarctica. This is because Greenland is a smaller ice sheet and is more sensitive to changes in temperature and precipitation.

By observe the months for Greenland and Antarctica, we can know the following things:

For Antarctica, Jan, Feb, and Mar are summer so there is a significant decline. There are also declines in Jun and Jul that may be affected by Greenland: This is not entirely accurate. While it is true that Jan, Feb, and Mar are summer months in Antarctica, the ice sheet does not experience a significant decline during this time. In fact, the summer months are when the ice sheet gains the most mass due to increased snowfall. The declines in Jun and Jul may be related to other factors such as increased melting or changes in ocean circulation.

For Greenland, Jun and Jul are summer months so there is a significant decline: This is true. The summer months are when the Greenland ice sheet experiences the most melting, leading to a significant decline in mass.

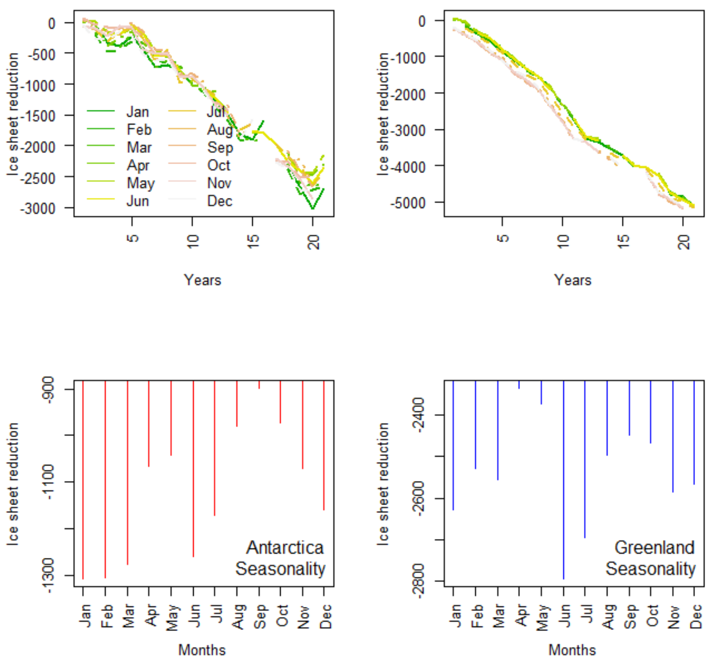

Figure 3: Seasonal variation in glacier mass in Greenland and Antarctica

4.2 Location specific model

We can using the estimates model to find the coefficient:

$$Model\ estimates:\ y_t = \widehat{a_1} + \widehat{b_1} t_1 + \widehat{\boldsymbol{b}_2}\ Months + X_t$$

Antarctica				
Variables	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	342.05	39.75	8.60	0.00
time	-12.21	0.15	-82.04	< 2e-16
monthsFeb	-171.24	50.44	-3.40	0.00
monthsJan	-132.11	50.44	-2.62	0.01
Adjusted R-squared:	0.97			
Multiple R-squared:	0.97			
Residual standard error: 159.5 on 227 degrees of freedom				

Greenland				
Variables	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	317.22	37.37	8.49	0.00
time	-22.93	0.14	-163.91	< 2e-16
monthsAug	-236.42	47.42	-4.99	0.00
monthsDec	-113.86	47.43	-2.40	0.02
monthsJan	-103.73	47.42	-2.19	0.03
monthsNov	-171.61	47.43	-3.62	0.00
monthsOct	-178.45	47.43	-3.76	0.00
monthsSep	-256.38	47.43	-5.41	0.00
Adjusted R-squared:	0.99			
Multiple R-squared:	0.99			
Residual standard error: 150 on 227 degrees of freedom				

According to the table we can get the model:

Antarctica : $y_t = 342.049 - 12.206t_1 \ - 132.107\ Jan - 171.235\ Feb + X_t$

Greenland: $y_t = 317.22 - 22.929t_1 - 236.423\,Aug - 256.383\,Sep - 178.45\,Oct - 171.607\,Nov - 113.856\,Dec - 103.727\,Jan + X_t$

Both locations have negative trend. But Greenland's trend is almost twice that of Antarctica. Also, on average, each year Greenland is witnessing significant loss in mass of ice across six months, compared to two months in Antarctica.

4.3 Combined model

After analysing the trend and seasonality of Greenland we can know that ice sheet mass decline overtime. To know the interaction between Antarctica and Greenland we can also fit a combined model:

Variables	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	385.31	33.61	11.46	< 2e-16
regionGreenland	-111.35	30.99	-3.59	0.00
time	-12.17	0.16	-77.12	< 2e-16
monthsAug	-90.13	37.84	-2.38	0.02
monthsDec	-70.42	37.85	-1.86	0.06
monthsFeb	-122.65	37.84	-3.24	0.00
monthsJan	-117.92	37.84	-3.12	0.00
monthsSep	-79.84	37.84	-2.11	0.04
regionGreenland:time	-10.80	0.22	-48.44	< 2e-16
Adjusted R-squared:	0.99			
Multiple R-squared:	0.99			
Residual standard error: 169.2 on 465 degrees of freedom				

The analysis of ice sheet mass trends reveals an overall significant negative trend, with a more significant decline in mass observed in Greenland compared to Antarctica. Specifically, four months Aug, Feb, Jan, and Sep showed a statistically significant association with the changing mass of ice, with all of them having negative slopes relative to Apr (baseline). Among these, Jan and Feb, which are winter months in the northern hemisphere, also showed statistically significant ice loss in Greenland. Interestingly, the average negative trend of Greenland is also statistically larger than that of Antarctica.

These findings suggest that the ongoing loss of ice in Greenland is a major concern, with a greater urgency to address the issue.

4.4 Autocorrelation function

The definition of a stationary process is that the mean and variance of the process are constant over time. Obvious our model is a non-stationary model.

The correlation of the sample autocorrelation between the Greenland and Antarctica is 0.665. We also see that there is periodic behavior in the sample ACF for Greenland. That is quite different from Antarctica which means other covariates need to be considered into model.

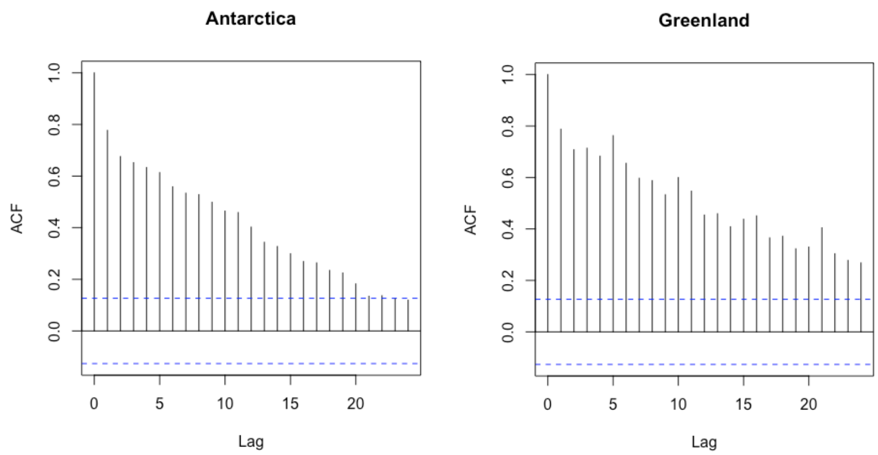

Figure 4: The autocorrelation function of Greenland and Antarctica

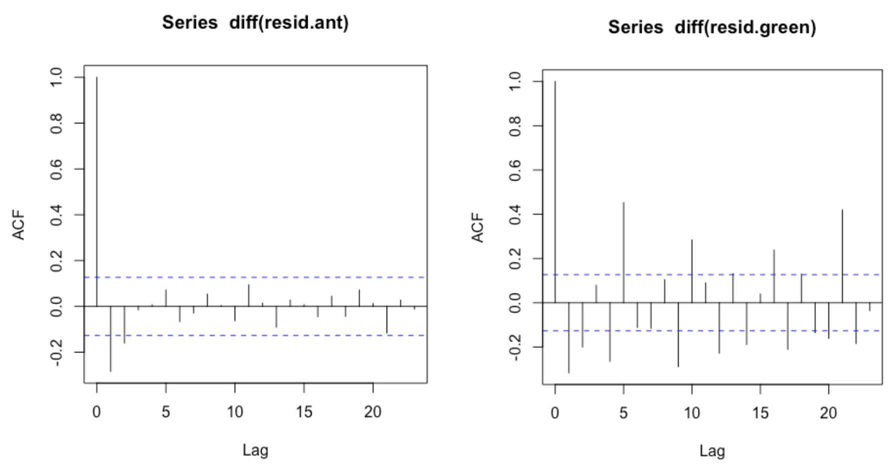

Figure 5: The autocorrelation function of Greenland(Right) and Antarctica(Left)

5 Conclusion

Despite being affected by seasonal variations, the mass of the Antarctic and Greenland ice sheets continues to decline each year, although with different characteristics. Greenland's negative trend is nearly twice that of Antarctica. While both regions experience substantial ice loss during the summer months, Greenland has been losing ice across multiple months for several years, which is concerning. The analysis also reveals statistical differences in trends and seasonality between the two regions. Furthermore, the residual auto-correlations are slowly decaying, indicating non-stationarity and the influence of other environmental factors. Taken together, these findings suggest the need for continued monitoring and study of ice sheet mass loss to better understand the impacts of climate change on the polar regions.

6 Future work

To improve the accuracy of our analysis, we recommend adopting advanced time series regression models and exploring alternative methods for seasonality. This will allow us to better explain the data and capture any underlying patterns that may not be apparent using traditional methods. Additionally, we suggest experimenting with probabilistic methods for treating missing values in the data. This will help to reduce any bias or error introduced by missing data points. Furthermore, conducting separate time series analyses for summer versus winter seasons can provide us with a better understanding of the seasonal patterns and trends in the data. This approach will enable us to capture any unique seasonal effects that may not be evident in the overall trend analysis. Overall, these strategies can improve the accuracy and robustness of our analysis, enabling us to better understand the behavior of the ice sheets and their response to environmental changes.

Reference

National Aeronautics and Space Administration. (2022, September 22). Vital Signs of the Planet: Global Climate Change. Retrieved from https://climate.nasa.gov/vital-signs/ice-sheets/

NASA. (2020, March 10). Greenland, Antarctica Melting Six Times Faster Than in the 1990s. Retrieved from https://climate.nasa.gov/news/2958/greenland-antarctica-melting-six-times-faster-than-in-the-1990s/

Grolemund, G., & Wickham, H. (2017). R for Data Science. O'Reilly Media.

Artemieva, I. M. (2022). Antarctica ice sheet basal melting enhanced by high mantle heat. Earth-Science Reviews, 226, 103954.

Sproson, A. D., Yokoyama, Y., Miyairi, Y., Aze, T., & Totten, R. L. (2022). Holocene melting of the West Antarctic Ice Sheet driven by tropical Pacific warming. Nature Communications, 13(1), 1-9.

Wickham, H., & Grolemund, G. (2016). R for Data Science: Import, Tidy, Transform, Visualize, and Model Data. O'Reilly Media.

R Core Team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Shumway, R. H., & Stoffer, D. S. (2017). Time Series Analysis and Its Applications: With R Examples. Springer.

Acknowledgement

I would like to express my gratitude to the teachers at XUT and JCU who have played an instrumental role in shaping my research skills and have been a constant source of guidance and support throughout the entire research process. Their patience and dedication have enabled me to face and overcome scientific problems with confidence. Additionally, I would like to extend my thanks to NASA for providing the data that formed the foundation of my research. Without their support, this study would not have been possible. Lastly, I am thankful to AMSI for giving me the opportunity to conduct this research and present my findings, which has been a valuable experience in my academic and professional growth.

Appendix

Example of R Code:

```
write.csv(ant2.long, file = "ant.dat.csv")
write.csv(grn2.long, file = "green.dat.csv")
ant.dat = read.csv("ant.dat.csv", header = TRUE)
green.dat = read.csv("green.dat.csv", header = TRUE)
f1 = 1
T = 12
mod.ant = lm(Mass ~ time + months, data = ant2.long[1:240,])
summary(mod.ant)
mod.green = lm(Mass ~ time + months, data = grn2.long[1:240,])
summary(mod.green)
resid.ant = ant2.long[1:240,]$Mass - mod.ant$fitted.values
ant.acf = acf(resid.ant, lag = 24, main = "Antarctica")
ant.diff.acf = acf(diff(resid.ant))
pacf(resid.ant, lag = 12, main = "Antarctica")
resid.green = grn2.long[1:240,]$Mass - mod.green$fitted.values
green.acf = acf(resid.green, lag = 24, main = "Greenland")
green.diff.acf = acf(diff(resid.green))
pacf(resid.green, lag = 24, main = "Greenland")
library(ggplot2)
ggplot(data = ant.dat, aes(time, Mass)) +
  geom_line() +
  xlab("Time (Month)") +
  ylab("Mass (Gigatonnes)") +
  ggtitle("Antarctica")
ggplot(data = green.dat, aes(time, Mass)) +
  geom_line() +
  xlab("Time (Month)") +
  ylab("Mass (Gigatonnes)") +
  ggtitle("Greenland")
```