

# Autoregressive Integrated Moving Average Model for Polar Seas Ice Prediction

Safak Kayikci

Bolu Abant Izzet Baysal University,  
Department of Computer Engineering, 14030  
Bolu, Turkey

**Abstract**—Sea ice predictions are very important for the future of polar climates and play a significant role in ecosystems. Models are the simulated representations that have been set up to research systems. To advance model forecasts, researchers require improved parameterizations that are formed by the assembling and analysis of convenient observations. In this study, an Autoregressive Integrated Moving Average (ARIMA) model is proposed to predict the Arctic and Antarctic sea ice extent. The data is gathered from the National Snow and Ice Center (NSIDC) between 01. Jan.1979 and 30. Jun.2020. The fitted data between 2017 and 2020 matches the observed data very closely with the overlap is firmly within the 95% confidence band shows the success of the model.

**Keywords**— Antarctic sea, Arctic sea, Autoregressive Integrated Moving Average (ARIMA), Sea ice prediction

## I. INTRODUCTION

POLAR seas is the whole expression for the Arctic Ocean and the south portion of the Southern Ocean. They occupy about 15% of Earth's oceans together. The Arctic Ocean has the lowest acreage among other oceans. It is the most shoal, and the coldest one. It is partially covered by ice all over the year and pretty much fully in winter. Its surface temperature and salinity change at particular seasons of the year regarding ice melting and freezing [1]. The Southern Ocean, also known as the Antarctic Ocean is the second-smallest ocean. In winters, sea ice may surround 13% of Earth's total surface.

After the seawater freezes, they float on the ocean surface and form sea ice. They are formed in the Arctic and Antarctic seas in winter, their amount in the summer shrinks but they do not disappear completely. Sea ice has an intense effect on the polar medium, regulate ocean cycle, weather, and territorial climate.

Ice crystals extract salt from underneath water when they are formed. This concentrated, cold and salted water can go down the ocean floor as it moves back towards the equator. Sea ice also prevents erosion by reducing wind and waves from the shore. Another benefit is that it creates insulation that will reduce evaporation on the ocean surface and reduce heat loss to the atmosphere. Consequently, the air in the ice-filled

areas becomes cold and dry than the others. Besides, sea ice has an important place in the food chain by the melting and freezing that helps the development of phytoplankton and the formation of nutrients.

## II. SEA ICE DATASET

The Sea Ice index ensures sea ice variations at Arctic and Antarctic seas. Its spatial coverage is N: -39.23, S: -90, E: 180, W: -180 and N: 90, S: 30.98, E: 180, W: -180 with a spatial resolution of 25 km x 25 km. In this study, data between dates 02.01.1979 and 30.06.2020 are studied for both hemispheres. DMSP, DMSP 5D-3/F17, DMSP 5D-3/F18, Nimbus-7 are the platforms and SMMR, SSM/I, SSMIS are the source sensors for data [2].

## III. METHODS AND MATERIALS

From this data, we can see that there are independent maxima and minima for the northern and southern hemisphere data sets, as is to be expected from the change of season. In addition to this, there is on average more sea-ice in the southern hemisphere than the north.

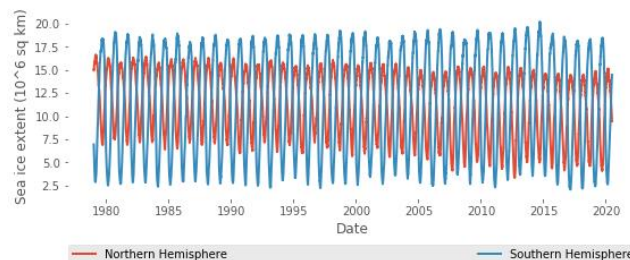


Fig. 1 Daily ice extent of both hemispheres

Here in Figure 1, we can observe the general trend that there has been a steady decrease in the extent of sea-ice in the northern hemisphere, the extent of sea-ice in the southern hemisphere was relatively constant until the early 2000s, and since then has been increasing, at first this seemed to be a very strange result, this has been known for some time and is due to a large number of complicated factors. One of the key trends observed in the rise of global warming is that seasonal phenomena seem to be shifting to earlier and earlier in the year, let's see if we can observe this in the melting and re-freezing of sea-ice in the polar regions. In this figure the older data are colored green, shifting towards blue in the current

data.

From this data we can see that there is a shift in the positions of the maxima and minima of the sea ice extent in both the northern and southern hemispheres, indicating that the seasons have been gradually shifting over time in Figure 2.

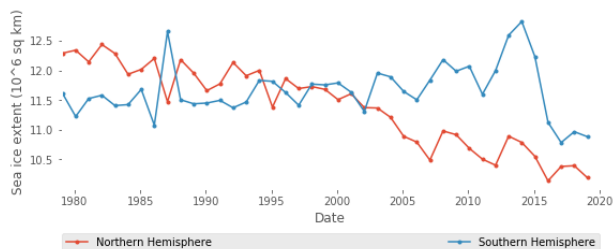


Fig. 2 Annual average sea ice extent

In the northern hemisphere data, there is a marked difference between the first and final year data, clearly highlighting the decrease in the extent of sea ice year on year [3]. The southern hemisphere has been much more uniform over the same period, indicating that the extent of sea-ice melting over this period has been much lower.

#### IV. ARIMA MODEL

Autoregressive Integrated Moving Average (ARIMA) is a prevalent technique for time series forecasting that ensures collateral solutions like exponential smoothing. Exponential smoothing deals with the characterization of tendency whereas ARIMA purposes to define the autocorrelations in data.

ARIMA(p,d,q) is the canonical representation where parameters are integer values. These parameters are defined as in Table 1.

p	order of AR term which is the number of lags observations in model
d	order of MA term which is the differentiation count in pure data
q	order of moving average

Table 1 Parameters of ARIMA model

In the basic Auto Regressive (AR only) model,  $Y_t$  relies only on self-delays. That is:

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_p Y_{t-p} + \varepsilon_t \quad (1)$$

$Y_{t-1}$  is the lag of the series,  $\beta$  is the coefficient of lag that the

model predicts and  $\alpha$  is the predicted intercept term.

In the same way, in a basic Moving Average (MA only) model,  $Y_t$  depends only on the delayed estimate faults.

$$Y_t = \alpha + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} \quad (2)$$

Error terms are the faults of the autoregressive models of the corresponding delays. ARIMA model is applied where time series varies at least once to perform stability. By combining AR and MA functions, the equation becomes:

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_p Y_{t-p} \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} \quad (3)$$

ARIMA model in explanation predicted  $Y_t$  is the sum of maximum p number linear combination of lags of Y and maximum q number of a linear combination of delayed forecast errors.

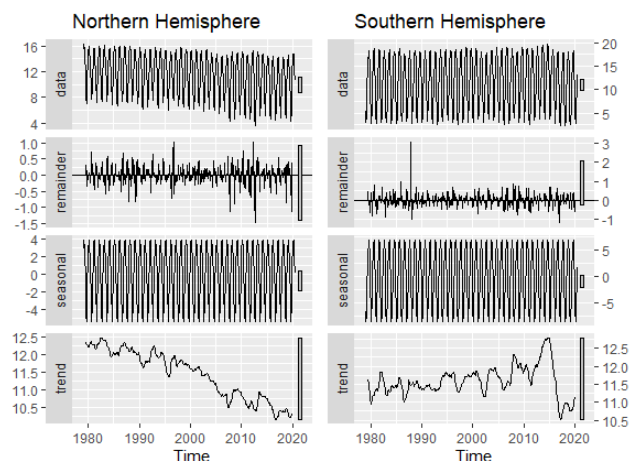


Fig. 3 Seasonal oscillations from the overall trend

Equations 1,2 and 3 are derived from machine learning web site [4]. R programming language is used for ARIMA modeling. The auto.arma() function applies a different version of Hyndman-Khandakar algorithm that assembles unit root tests, and minimize AICc and MLE to get an ARIMA instance [4]. The parameters of the function supply many contrasts on the model.

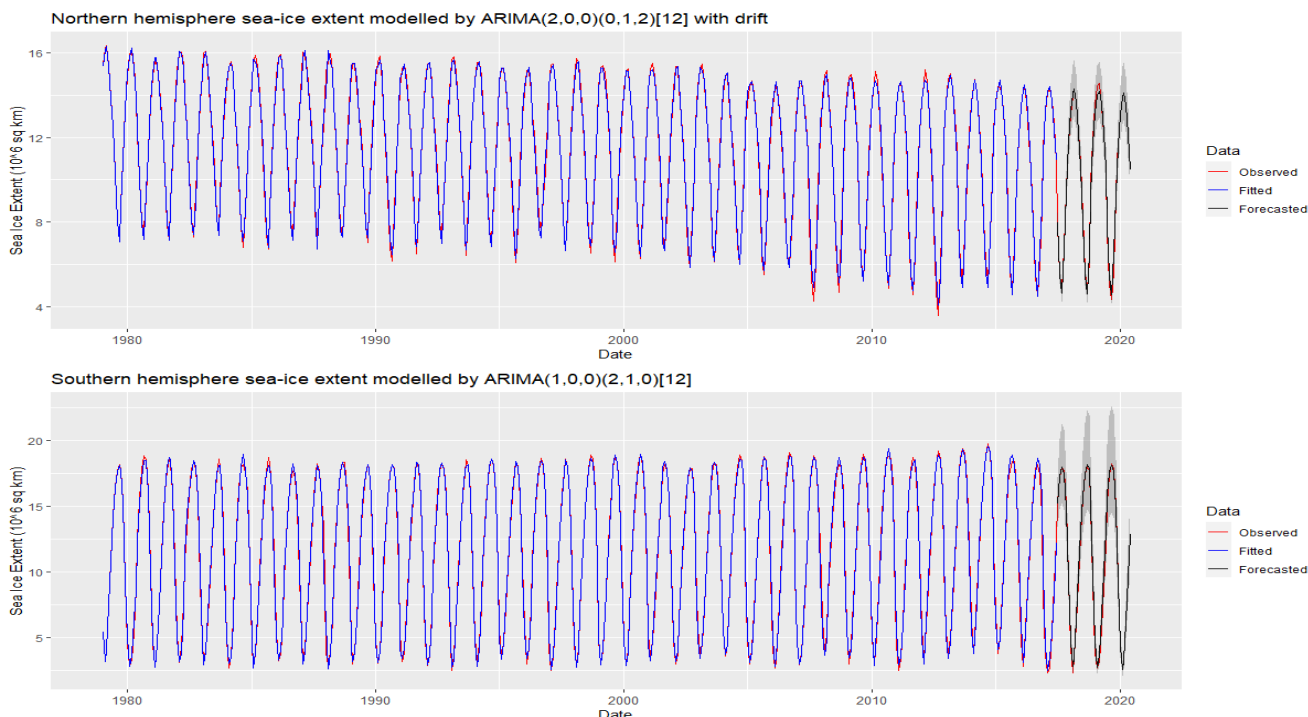


Fig. 4 Real and fitted data between 2017 and 2020

The Ljung-Box testing is applied to test residuals from a fitted model that has unwanted autocorrelation [5]. If autocorrelation exists in the residuals, then presumably a model with more parameters can be fitted to the original data and explain more of the structure it contains.

The test statistic is:

$$Q = n(n + 2) \sum_{k=1}^l \frac{\hat{r}_k^2}{n-k} \quad (4)$$

$$\hat{r}_k = \frac{\sum_{i=k}^{n-1} (v_i - \bar{v})(v_{i-k} - \bar{v})}{\sum_{i=0}^{n-1} (v_i - \bar{v})^2} \quad (5)$$

where  $n$  is the length of  $v$  and  $l$  is the number of lags. The variance of the statistic slightly exceeds the variance of the chi-squared distribution, but it still is a fairly good test with a reasonable computational cost.

Model	Q*	df	p-value	df	lags
ARIMA (2,0,0)(0,1,2)[12]	28.863	19	0.06817	5	24

Table 2 Parameter values of for the Arctic Sea

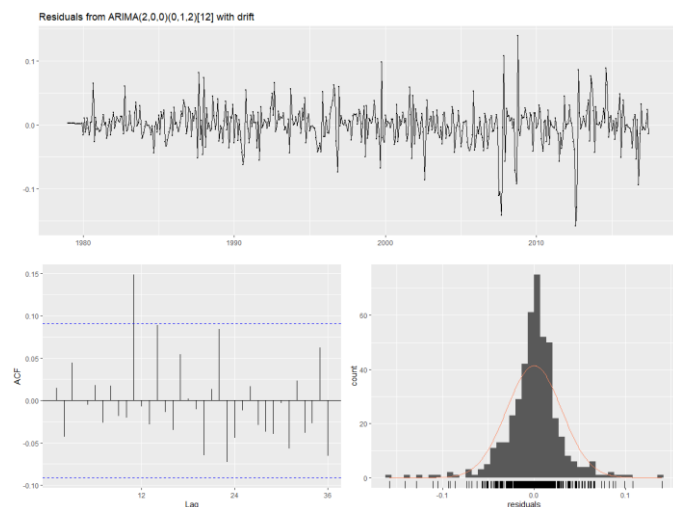


Fig. 5 Residuals from ARIMA for the Arctic Sea

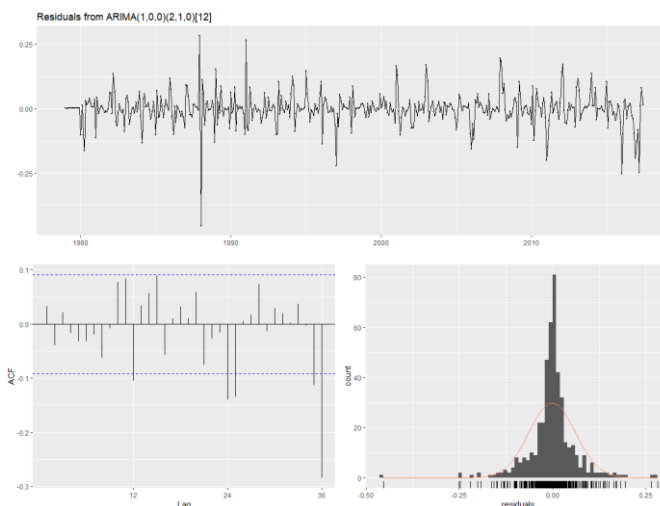


Fig. 6 Residuals from ARIMA for Antarctic Sea

Model	Q*	df	p-value	df	lags
ARIMA (1,0,0)(2,1,0)[12]	38.79	21	0.0104	3	24

Table 3 Parameter values for the Antarctic Sea

## V. RESULTS AND CONCLUSION

From this fit in Figure 4, we can see that all of the historic data is reproduced very faithfully, and the fitted data between 2017 and 2020 matches the observed data very closely too, giving us a good level of faith in the model, as the overlap is firmly within the 95% confidence band.

Arctic sea covers an ocean basin largely enclosed by land. Since there is not much ground at the North Pole, sea ice spreads all around the pole. Its ice is exposed to high oscillations between winter darkness and summer sunlight. In the same way, as the ocean basin is sided by land, ice is not free to move into below latitudes and dissolve. The main reason for atmospheric instability is the Arctic Oscillation. Most of the climate predictions indicate that the Arctic will lose most of the ice floe at the end of the 21st century. Related to the Arctic sea melting rate, sea ice could be at risk of natural variability like Arctic Oscillation.

The Antarctic is the opposite of the Arctic as it is surrounded by ocean. As a result of this geography, Antarctic sea ice distribution is wider than the Arctic in winter but lesser in summer. Also, its sea ice is more variable among years and climate oscillations don't modify ice as much as the other. Global warming and climate circumstances may influence each polar sea in different formats and measures. Sea ice can vary essentially from time to time. Comparison of only two locations in a specific time and analyzing tendencies over inadequate terms is not enough to figure out the long-term variations on sea ice. In this study, a prediction was made using the ARIMA model and the results of the last three years and the model created were compared. The predicted model has been shown to yield successful results on the sea ice forecast.

## References

- [1] Guemas, Virginie, et al. A review on Arctic sea-ice predictability and prediction on seasonal to decadal time-scales. *Quarterly Journal of the Royal Meteorological Society*, 2016, 142.695: 546-561.
- [2] Fetterer, F., K. Knowles, W. N. Meier, M. Savoie, and A. K. Windnagel. updated daily. *Sea Ice Index, Version 3*. Boulder, Colorado USA. NSIDC: National Snow and Ice Data Center, 2017
- [3] Wuebbles, D. J., Jain, A., Edmonds, J., Harvey, D., & Hayhoe, K. (2000). Global change: state of science. *Environmental Pollution* 100 (1999), 57-86. *Environmental Pollution*, 107, 473.
- [4] Selva Prabhakaran | ARIMA Model – Complete Guide to Time Series Forecasting in Python [online] Available: <https://www.machinelearningplus.com/time-series/arima-model-time-series-forecasting-python/>
- [5] Khandakar, Y., & Hyndman, R. J. *Automatic time series forecasting: the forecast package for R*, 2008
- [6] Burns, P. (2002). Robustness of the Ljung-Box test and its rank equivalent. Available at SSRN 443560.
- [7] Bushuk, M., Msadek, R., Winton, M., Vecchi, G., Yang, X., Rosati, A., & Gudgel, R., Regional Arctic sea-ice prediction: potential versus operational seasonal forecast skill. *Climate Dynamics*, 2019, 52(5-6), 2721-2743.
- [8] Contreras, J., Espinola, R., Nogales, F. J., & Conejo, A. J. ARIMA models to predict next-day electricity prices. *IEEE transactions on power systems*, 2003, 18(3), 1014-1020.
- [9] Stroeve, J., Hamilton, L. C., Bitz, C. M., & Blanchard-Wrigglesworth, E. Predicting September sea ice: Ensemble skill of the Search sea ice outlook 2008–2013. *Geophysical Research Letters*, 2014, 41(7), 2411-2418.
- [10] Wang, W., Chen, M., & Kumar, A. Seasonal prediction of Arctic sea ice extent from a coupled dynamical forecast system. *Monthly Weather Review*, 2013, 141(4), 1375-1394.

## **Creative Commons Attribution License 4.0 (Attribution 4.0 International, CC BY 4.0)**

This article is published under the terms of the Creative Commons Attribution License 4.0

[https://creativecommons.org/licenses/by/4.0/deed.en\\_US](https://creativecommons.org/licenses/by/4.0/deed.en_US)