Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Observational Ice-Free Periods within the Arctic Ocean:

a Study Using Linear Regression

Kaufhold, Christine
christine.kaufhold@studium.uni-hamburg.de
Institut für Meereskunde, Universität Hamburg

Sea Ice Physics Report
February 9, 2021

Abstract

Arctic ice-free periods are a good indicator for climate change. This is due to the feedback processes sea ice experiences with increasing temperatures. In this study, inner and outer ice-free periods are studied using linear regression using observational data over 39 years. Patterns within the data, such as regions of interest, are identified. The rate of change over the 39 years was calculated, and used to provide hindcasts and forecasts. A brief discussion into the effectiveness of linear regression was touched on, and it's relevance to Arctic sea ice periods.

Contents

List of Figures

List of Tables

1 Introduction

Arctic sea ice science is a complex, interdisciplinary field of study, which covers everything from phyiscs to biology. While perhaps obvious, sea ice refers to ice that forms on open water, in comparison to glaciers, icebergs, and ice sheets/shelves. Although we know that it is intimately tied with other processes, most our modern understanding of sea ice dynamics comes after October 1978 [10]. This was the time when satellite retrivals of Arctic sea ice data commenced. With this rise in technology, the sea ice concentration (SIC) and total sea-ice extent (SIE) in the Arctic was quantified in a way which it was never before.

Sea ice presents an opportunity to study the effects of climate change. While the polar regions are generally far from the equator and cool in temperature, they are especially sensitive to changes in the climate. A small increase in temperature can create large feedback effects. This is because sea ice has a brighter surface than the open ocean (and thus a higher albedo, α), and most of the sunlight is reflected. With less sea ice able to reflect the solar energy, it is able to absorb more and more, starting a warming-melting cycle. At the present time, sea ice grows the most during the winter months, and melts during the summer months. While some regions contain sea-ice that remains all year long, these places have been getting smaller due to the aforementioned feedback cycles.

1.1 Research goals

In this study, patterns within the ice-free periods will be analyzed within the Arctic ocean. Observational data will be used, as it is often closer to the 'truth' compared to simulations. Linear regression will be performed on the data to determine such patterns. As a first step, quantifying Arctic sea ice coverage will be done using visual analysis and using histograms. After this, hindcasts and forecasts will be generated for the data using the predictive capabilities of the linear regression model. In this way, the abilities of the linear regressor will be tested, and compared against observational data or other predictions. Finally, regions of interest within the Arctic ocean will be identified using the change of rate determined by the regression model. Changes which experience more change than others will be discussed, and again compared against observations. As a final result, the linear regression model will be used on some individual regions to determine if there is significant change in the change of rate or the forecasts.

2 Experimental Design & Procedure

For this particular problem, appropriate data was sought after. Data from NOAA (National Oceanic and Atmospheric Administration) and NSIDC (National Snow and Ice Data Center) was found for Arctic sea ice melt periods. The publishers of this data calculated the sea ice concentration from brightness temperature (T_b) observations. These were collected using the following passive microwave satellites: a Scanning Multichannel Microwave Radiometer (SMMR), Special Sensor Microwave/Imager (SSM/I), and Special Sensor Microwave Imager/Sounder (SSMIS). The authors point out that passive microwaves are sensitive to the state of water, and therefore, it makes ice melt detectable. Other input parameters for their calculations were accessed by the publishers through another NOAA/NSIDC databank.

2.1 Data structure

The aforementioned data was provided in a single netCDF (.nc) file. It spans from March 1st, 1979 up until February 28th, 2018 for a total of 39 years. Each year contained 5 different parameters of interest, as seen in Table 1. The normal gridded resolution was 25 [km] $\times$ 25 [km]; the number of rows was 448, and the number of columns was 304. This is of importance for the discussion of regression techniques later on.

Table 1: Different variables available in the NSIDC dataset. The following dataset spanned the years from 1979 and 2017 was obtained for this analysis. See [8] for more information.

Variable	Name	Description	Unit
OIFP	Outer ice-free period	Period under 80% concentration	Days
IIFP	Inner ice-free period	Period under 15% concentration	Days
SLIP	Seasonal loss of ice period	Period of decreasing concentration between 80-15 %	Days
SGIP	Seasonal gain of ice period	Period of increasing concentration between 15-80 %	Days
SIZ	Seasonal ice zone		Days

The Python library `netCDF4` was used to open the dataset. From `matplotlib`, the toolkit `basemap` was utilitzed for geophysical visualizations. The code to plot the above listed variables with basemap can be found on Christine Kaufhold's GitHub, in the 'Sea Ice Physics' repository. Please visit [4] for further information.

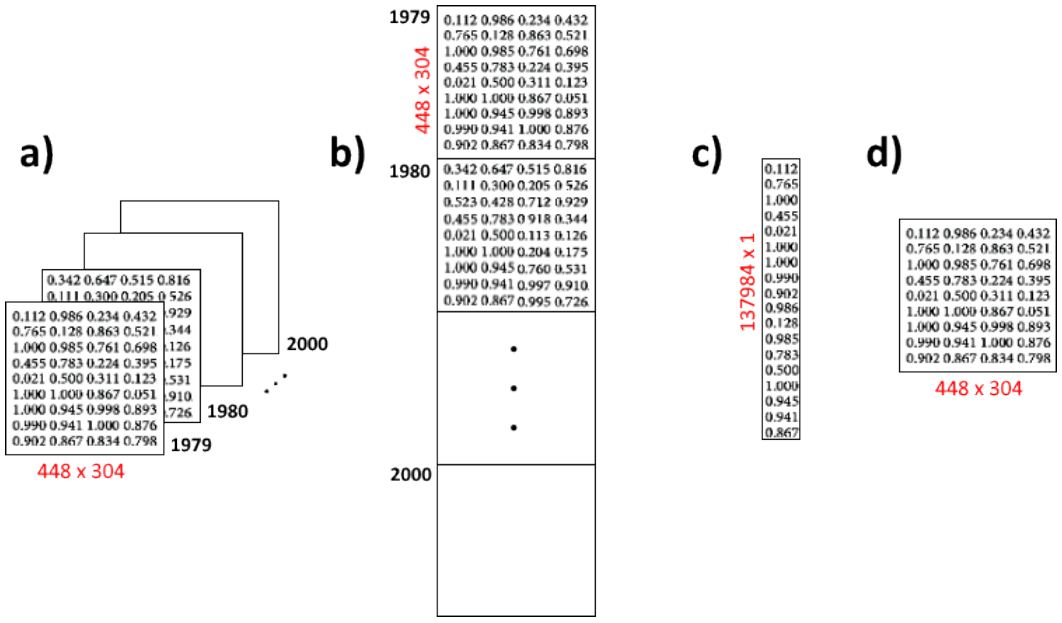
2.2 Methods

Linear regression was identified as first step in performing explanatory analysis. It can be used on any sized dataset, and can also make predictions – a topic which will be discussed later in this report. Linear fits for nonlinear problems, however, often come with many problems. This includes topics such as climate variability. But as a first step, it is good to further understand the problem at hand. Once linear regression was decided, the library `scikit-learn`, or `sklearn` was chosen for this task. Ultimately, it was chosen due to it's computational agility and ability to work with multi-dimensional (as in, longitudinal and latitudinal) data.

The variables in the NOAA/NSIDC file, whether it be IIFP, OIFP, or another one, was a matrix with a size of $39 \times 448 \times 304$. In Figure 1, this is best visualized as a). In this form, the variable in each individual year could be plotted using `basemap` and the longitudinal and latitudinal data. To use linear regression, however, the matrix was reshaped and flattened into a matrix of matrices (of length 39) such that it then became of size 39×137984. This is best represented as b) in Figure 1. Special care needed to be taken in order to flatten the matrix properly as it would have to be undone later.

After this, the linear regression model was built using the years and the flattened variable as inputs. With this done, information could be extracted from it. Firstly, the coefficients of the regression model (analogous to the 'slope', or rate of change of the problem when concerning one dimension) could be pulled. In Figure 1, the slope could be represented as something in the form of c). As multiple targets were passed into the regressor model, c) then represented the estimated coefficients for the linear regression problem over the entire grid for the span of the 39 years. This could then be unflattened using the opposite technique from a) to b), producing d). This could be plotted onto the `basemap` in the same way that a) was.

Figure 1: The data structure used in this experiment; where a) was the original variable of interest (IIFP, OIFP, etc.), b) was the flattened variable, c) was the calculated coefficients, or 'slope' using `sklearn`, and d) was the reshaped slope for plotting. For the visualization of d) please refer to Figures 4 and 5.



3 Results & Analysis

With the data located and the model theoretically set-up, all that was necessary was to collect the results. Variables of interest were particularly the IIFP and the OIFP, as it pertained to our question of exploring the ice-free periods in the Arctic. The raw data of the IIFP and OIFP have been plotted for the years of 1979, 2000 and 2017. These can be viewed in Appendices A and B. Using the information given by [1] and [8], numbers under a certain criteria were ignored. This was due to the authors of the dataset putting in placeholders to represent other information such as pole hole, missing data, or a land mask. When visualizing, a `numpy` mask array was applied to best see the changes in the actual data. It is for this reason why there is missing fields at the pole. Interestingly enough, when comparing the IIFP and OIFP between the years of 1979, 2000 and 2017, the pole hole gets smaller over time. This is likely due to more area closer to the pole experiencing non-zero ice-free periods.

3.1 Difference in variables over time

The difference between 1979 and 2017 for both IIFP and OIFP can be seen below. This was simply done by subtracting the dataset in 1979 from the dataset in 2017. Subsequently, a visual analysis was performed. A significant change over the 39 years is clear within the Kara/Laptev sea, as well as the Siberian/Beaufort seas. The definitions of these regions can be viewed in Appendix E. The maximum value in Figure 2 (IIFP) was determined to be 186 days, and the mean difference was 33.6 days. The maximum value in Figure 3 (OIFP) was 196 days, and the mean was a 39.1 day difference. As more ice-free periods were observed in 2017 (than 1979), the difference between the two resulted in a positive change in days. Some negative change of days (corresponding to regions which had more ice-free days in 1979, than in 2017) existed, but were difficult to pinpoint. This may be due to the colourbar chosen, or because the regions which experienced less ice-free days in

2017 (more ice in 1979) are small and rare. As this was difficult to analyze visually, histograms were produced, and placed within Appendix C.

Figure 2: Difference of inner ice-free period between 2017 and 1979. This was done by taking the two sets of data and subtracting them. By doing this, the total change of the ice-free period between the two years can be visualized.

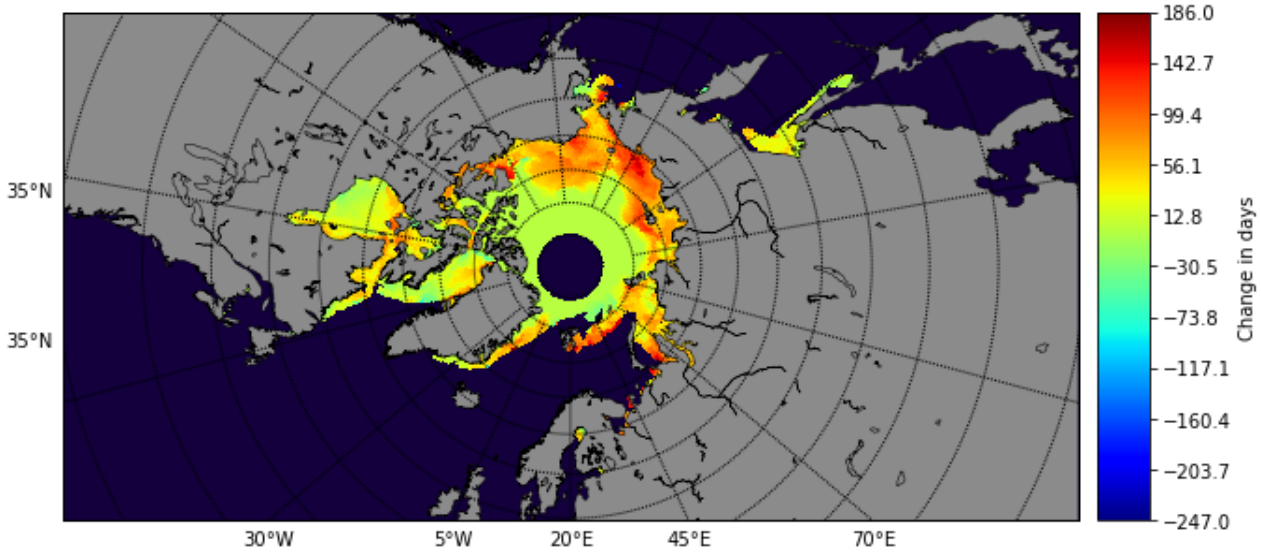
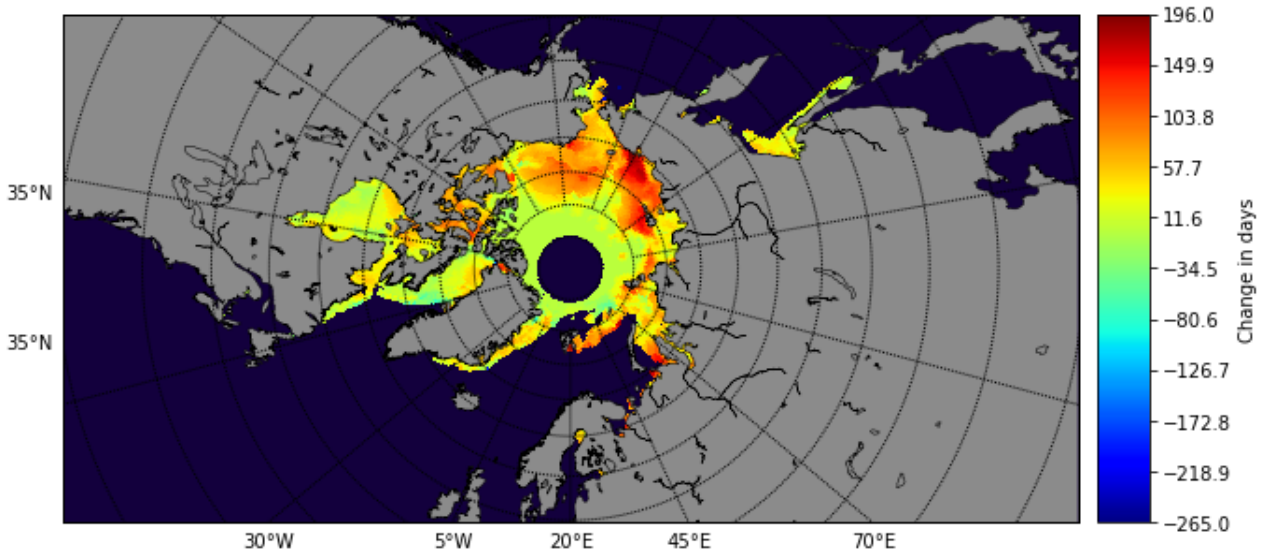


Figure 3: Difference of outer ice-free period between 2017 and 1979. A similar process for the inner ice-free period was performed. Compared to the inner ice-free period, more change in the outer ice-free period over the 39 years is evident in the Laptev/Siberian.



In the histograms, a significant shift (in both the IIFP and OIFP) is clear between the two years. Over time, a higher percentage of the area gained more ice-free days. A significant part of the area which experienced near 0 ice-free days did not change significantly, however, those that already experienced some ice-free days gained more. One must also keep in mind that areas in 2017 which have already completely melted, and have more than a 365 day ice-free period, are not accounted for in this data set. It is possible that the extent of the increase of ice-free areas is even

greater than one can see. The outer ice-free period has experienced accelerated increase compared to the inner-ice free period.

3.2 Areas of importance

Using the process discussed before, linear regression was used on IIFP and OIFP to determine the rate of change per year over the 39 years. This was done for a few reasons. Firstly, it was done to determine the average loss of sea-ice periods. Secondly, for finding regions of particular interest. And lastly, in order to make predictions on the dataset. See Figure 4 and 5 below. While linear regression is far from a perfect method, it provided a good basis for determining patterns within the data.

Figure 4: Rate of change of the inner ice-free period (IIFP) per year from 1979 to 2017. The values between $\pm 1E-3$ were masked over as the linear regressor did not handle the pole hole well.

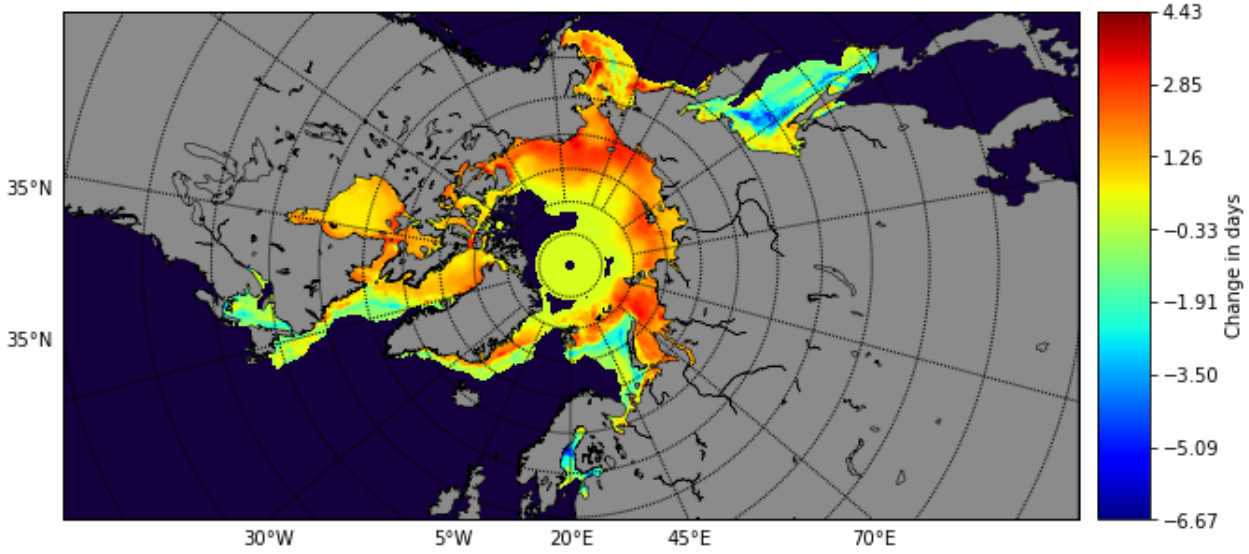
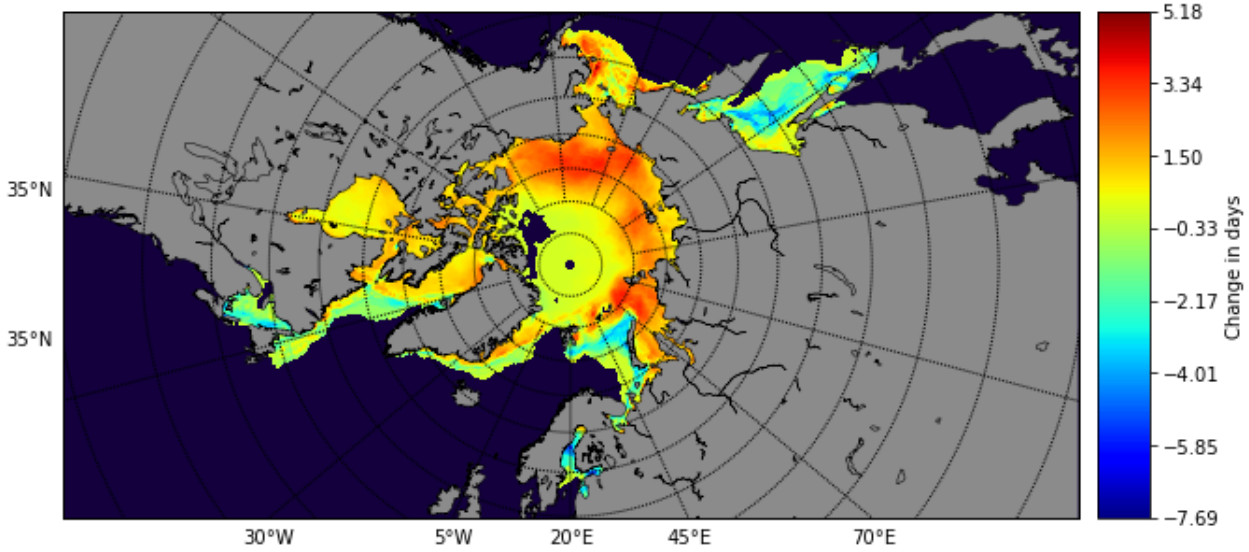


Figure 5: Rate of change of the outer ice-free period (OFIP) per year from 1979 to 2017. The same procedure applied to inner ice-free period (IIFP) was applied to this plot. Small variances between the two rates of change exist.

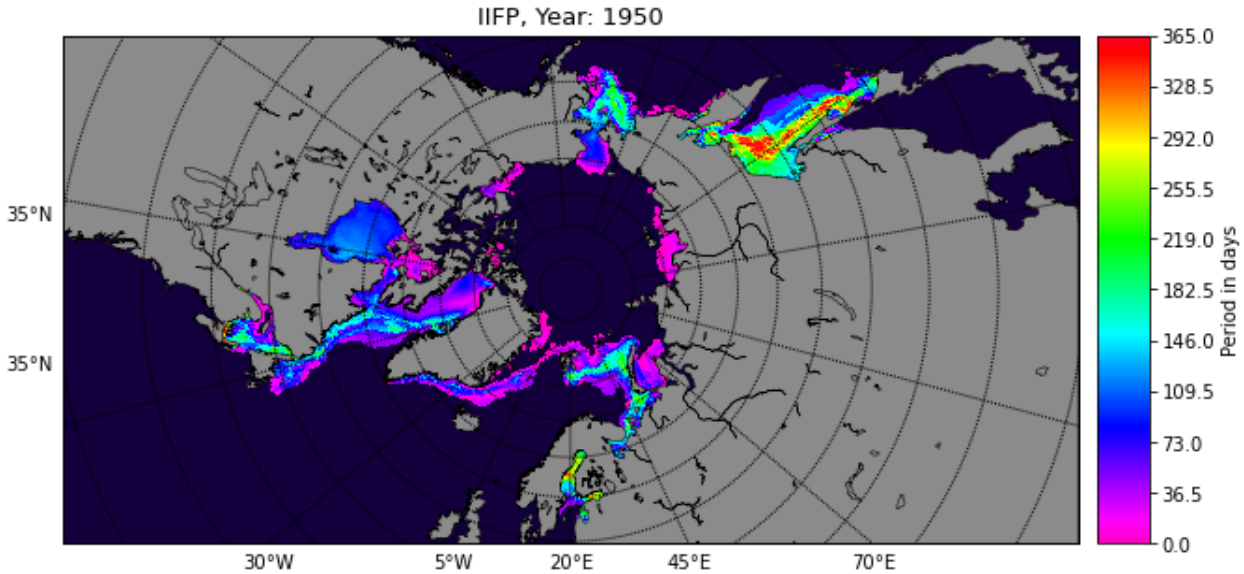


The colourbar represents the average amount of days either lost or gained over the 39 year period. The inner and outer ice-free periods experienced similar results from the linear regression. As seen in Figure 4 and 5, the Arctic ocean closest to Alaska and Russia experienced the largest gain in ice-free periods. This is especially interesting as it gained more ice-free days compared to the area which the Arctic ocean connects with the Atlantic ocean (the Greenland and Barents sea). Sub-polar gyre areas, such as the northern-most part of the Gulf Stream and the Kuroshio Current, appeared as if they accquired less ice-free periods (meaning they gained more sea ice days). This was another unexpected result in this data.

3.3 Hindcasting and forecasting

Using the linear regression coefficients calculated (Figures 4 and 5) forecasts could be built for on-coming years. Instead of calling `coef`, `predict()` could instead be used on the model for an array of years. While this technique is best used for points within our training data set (from 1979 to 2017) it can also be tested for points outside of it. First, the linear regression model could be tested by hindcasting, and checking if it matched with expectations. In addition to this, hindcasts may perhaps give insight into some of the regions of importance. The hindcasts for the inner ice-free period can be seen in Figure 6. The outer ice-free period hindcasts are located in Appendix D.

One obvious problem of the hindcasts, on first glance, is the treatment of the pole holes. As no data was recorded at these locations, the regression method generated noise between the parts with data, and the parts without data. For the best visualization, this noise was trimmed from the data before plotting. Because of this, the pole holes represented in the hindcasts are misshapen, however, they are best understood as areas at the pole which experience *no* change at all (i.e., their value is 0). In general, the hindcasts match up to expectations (of decreasing ice-free periods further back in time). In the figure from 1950, the Japan/Okhotsk sea was predicted to have inner/outer ice-free periods of 365 days. The result was interesting, as larger ice-free periods was surrounded by smaller ice-free periods in that area. Intuitively, the opposite should be true due to the nature of sea ice melt starting from the edge of ice forms. This was almost certainly an artifact due to the decreasing rate of change in Figure 4, and signals a heavy bias towards it. However, with this result, it has become a region of interest.



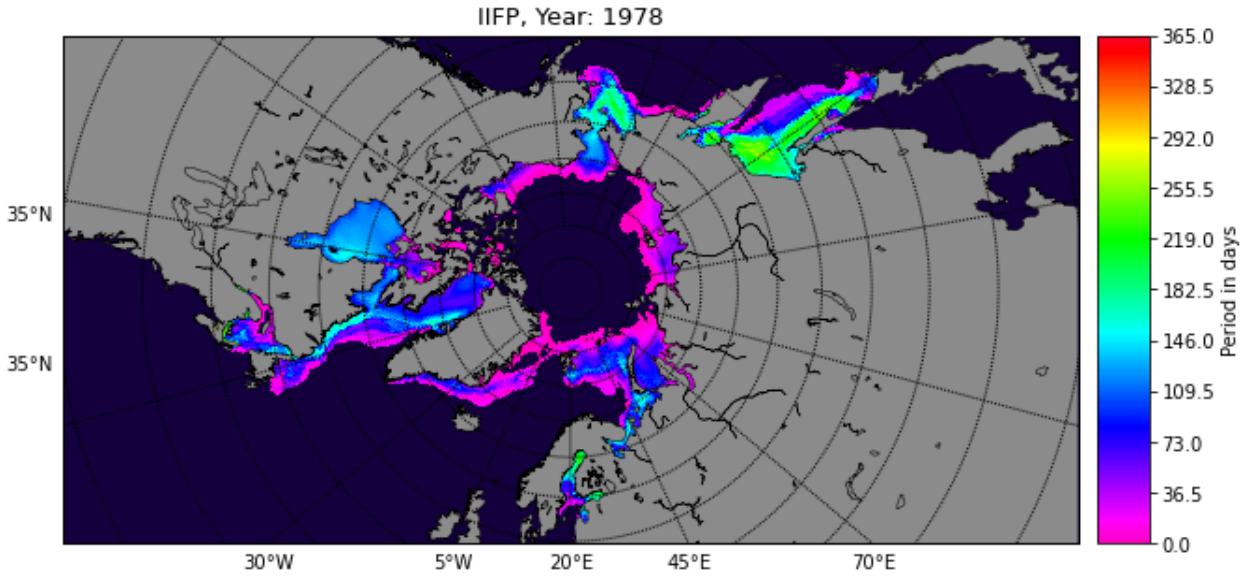
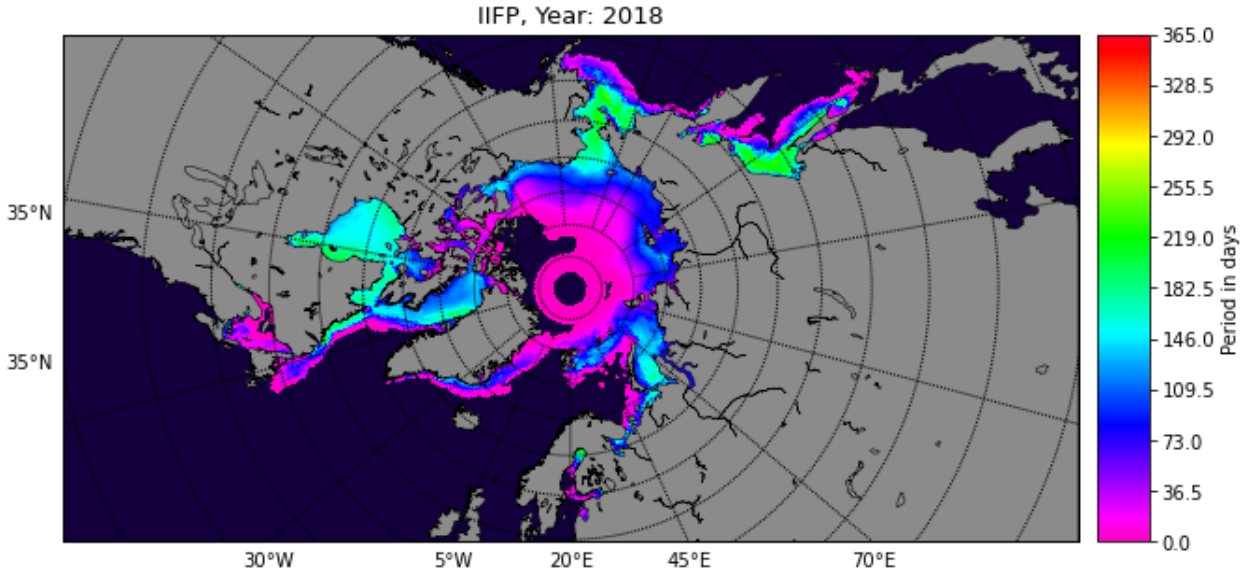


Figure 6: Generated hindcast for the inner ice-free period for the years of 1950 and 1978. This was done using the generated linear regression model for the IIFP and using it's predictive capabilities to test outside of it's range on observations. More ice-free periods are present in 1978 compared to 1950, as expected.

As a next step, forecasting was attempted. The years of 2018 to 2050 were used as the input years for the model. As 2030 and 2050 is a benchmark year for climate scientists and policy makers, it seemed natural to choose those dates. The same procedure for hindcasting was used. The results for the inner ice-free period are shown below in Figure 7, and the outer ice-free period in Appendix D. In these forecasts, a heavy bias towards the rate of change was evident. As an example, when the linear regression model calculated the rate of change close to the poles, the areas in the Arctic ocean which experienced zero inner ice-free periods was not effected. This would be the mathematical equivalent of having a slope of zero (no change).



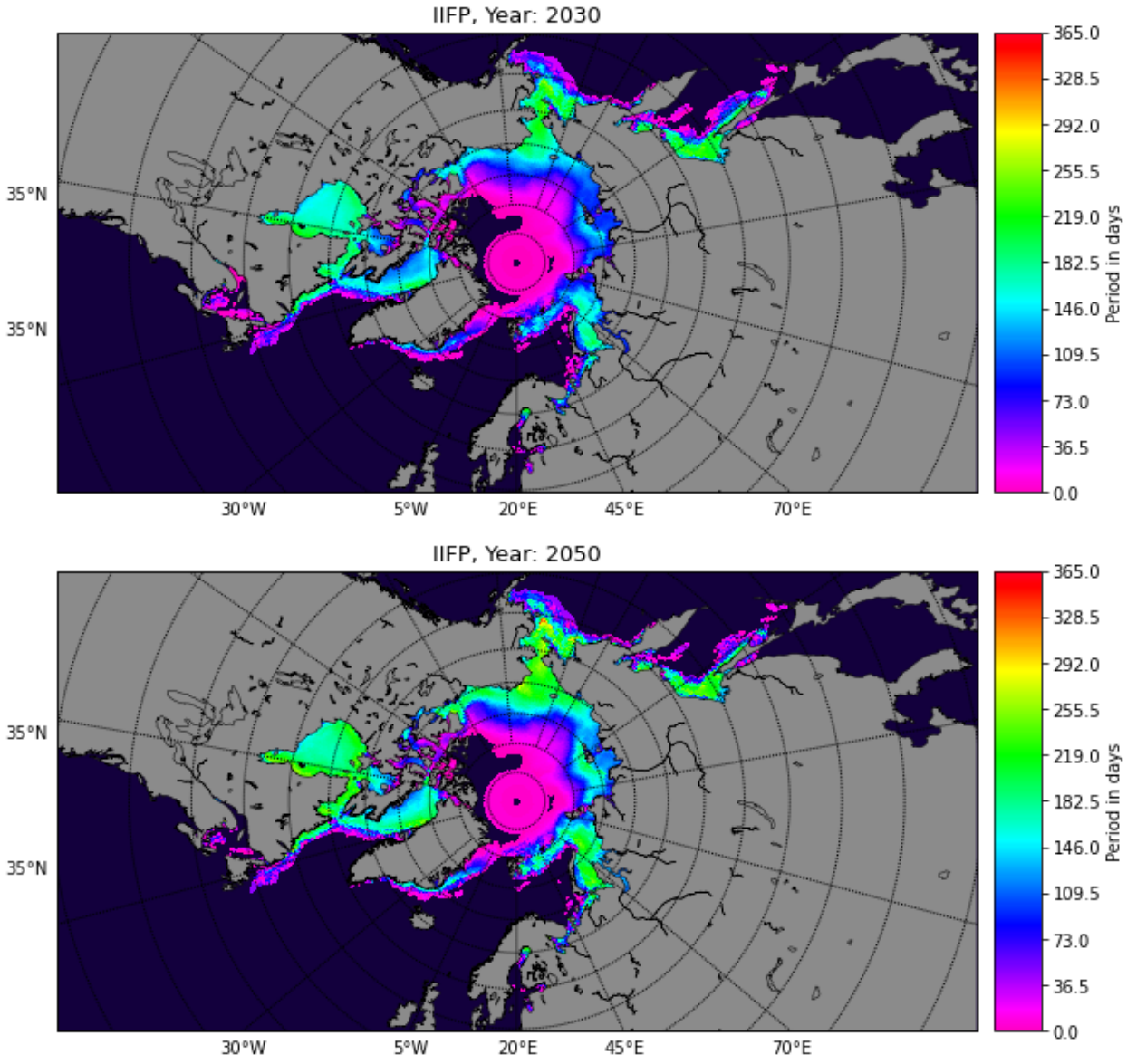


Figure 7: Forecasts for the inner ice-free period for the years of 2018, 2030 and 2050. Instead of using the linear regression model to hindcast, future input dates were selected. The range of the forecast years (32 years) was shorter than the training data (39 years), but longer than the one used for the hindcasts (28 years).

For this reason, there is very little change at the North pole in these forecasts. Although recent study by Nature Climate Change reports that everywhere in the Arctic will experience some ice-free days by 2035 [3], the linear regression model used in this report was not able to capture this. This result from the model could not be changed without serious modification to the methodology.

In addition to this: the ice-free periods near the sea of Japan/Okhotsk raise more questions than answers. Based on the rate of change (Figure 4), the area should have been losing ice-free days in the forecasts. Due to the recession away from the open ocean in Figures 7 and 12, the data implies the opposite. This could also be a problem with the data trimming, and the linear regressor producing noise in the interface between data and no data.

3.4 Further analysis

As a heavy bias on the change within the forecasts was determined, linear regression on each region was performed to try and improve predictions. Using Figures 4 and 5, the regions of most visual interest were split into 9 categories. For sake of brevity, the visuals of these regions have been pasted within Appendix E. The number of regions were chosen based on a previous paper working with the same dataset [2]. The regions were generated as a way to break down and attempt to quantify the areas with different changes. It was possible that the linear regressor was ineffective as regions experienced varying different rates of change. Three areas chosen for this further study was the Chukchi/Bering seas, the Japan/Okhotsk seas, and the Barents sea. These regions were chosen for their dynamic variability, and lack of clarity in the data. See Appendix E for the regional rate of change and forecast. The outer ice-free period rate of change was omitted for brevity, as it was similar to the inner ice-free period results.

With this being said, however, little change was seen between the original rate of change and the rate of change of each region. The regional forecasts of all three regions gave similar results as well. As the IIFP and OIFP were not different enough in extent and magnitude, compared to the original regression, no further analysis was performed.

4 Discussion & Conclusion

As a result of the analysis, it can be clearly stated that linear regression is not fit for modelling ice-free periods in the Arctic. While the application of the experiment was successful, mixed results were obtained. Firstly, hindcasts and forecasts generally predict the expected trend (according to [3]). There are less ice-free periods in the past, and more in the future. Since the dimension of the problem remains unknown, however, linear forecasts alone do not produce completely comprehensive results. Hindcasts, on the other hand, do produce something reasonable to expectations. Performing linear regression on smaller areas also did not improve prediction quality, but that may be an obvious result due to the nature of linear regression.

From this, however, regions of particular interest were successfully identified. Refer back to Figure 4 and 5. Areas on the western side of sub-polar gyres experienced less sea-ice period loss, and rather gained periods of sea ice. The eastern side of these gyres did not see any sea ice. This could be due to the affect of thermohaline circulation dynamics, and cold water being brought down from the Arctic on the west. In particular, the Japan/Okhotsk sea had a very interesting trend of ice-free days, and it is still unclear (based on these results) how they will continue to change. Additionally, the Arctic regions closer to the Pacific ocean have experienced more loss in periods of sea ice compared to the Atlantic. This is an interesting result as the Northern Atlantic is warmer and saltier [6]. Further analysis on the regions is heavily suggested.

From this, only a small part of the problem of hand has been understood. In the future, a different regression technique (such as polynomial/multivariate) is suggested. This could enhance prediction quality, and remove the strong linear bias. It could also, perhaps, aid with the regional forecasts as well. This has yet to be explored in the context of this data.

5 References

[1] M. Anderson and A. Bliss, *Snow Melt Onset Over Arctic Sea Ice from SMMR and SSM/I-SSMIS Brightness Temperatures*. https://nsidc.org/data/NSIDC-0105/versions/4, 2019.

[2] A. C. Bliss, M. Steele, G. Peng, W. N. Meier, and S. Dickinson, *Regional Variability of Arctic Sea Ice Seasonal Change Climate Indicators from a Passive Microwave Climate Data Record*, Environmental Research Letters, 14 (2019), p. 045003.

[3] M.-V. Guarino, L. C. Sime, D. Schröeder, I. Malmierca-Vallet, E. Rosenblum, M. Ringer, J. Ridley, D. Feltham, C. Bitz, E. J. Steig, E. Wolff, J. Stroeve, and A. Sellar, *Sea-Ice-Free Arctic during the Last Interglacial supports Fast Future Loss*, Nature Climate Change, 10 (2020), pp. 928–932.

[4] C. Kaufhold, *Github: Sea Ice Physics 1 Repository*. https://github.com/cekaufho.

[5] T. Markus, J. C. Stroeve, and J. Miller, *Recent Changes in Arctic Sea Ice Melt Onset, Freezeup, and Melt Season Length*, Journal of Geophysical Research: Oceans, 114 (2009).

[6] J. L. Reid, *On the temperature, salinity, and density differences between the atlantic and pacific oceans in the upper kilometre*, Deep Sea Research (1953), 7 (1961), pp. 265–275.

[7] D. M. Smith, *Recent Increase in the Length of the Melt Season of Perennial Arctic Sea Ice*, Geophysical Research Letters, 25 (1998), pp. 655–658.

[8] M. Steele, A. Bliss, P. Ge, W. Meier, and S. Dickinson, *Arctic Sea Ice Seasonal Change and Melt/Freeze Climate Indicators from Satellite Data, Version 1*. https://nsidc.org/data/NSIDC-0747/versions/1, 2018.

[9] J. Stroeve, T. Markus, W. N. Meier, and J. Miller, *Recent Changes in the Arctic Melt Season*, Annals of Glaciology, 44 (2006), p. 367–374.

[10] J. Stroeve and D. Notz, *Changing state of arctic sea ice across all seasons*, Environmental Research Letters, 13 (2018), p. 103001.

[11] J. C. Stroeve, T. Markus, L. Boisvert, J. Miller, and A. Barrett, *Changes in Arctic Melt Season and Implications for Sea Ice Loss*, Geophysical Research Letters, 41 (2014), pp. 1216–1225.

Appendix A - Inner Ice-Free Period Raw Data

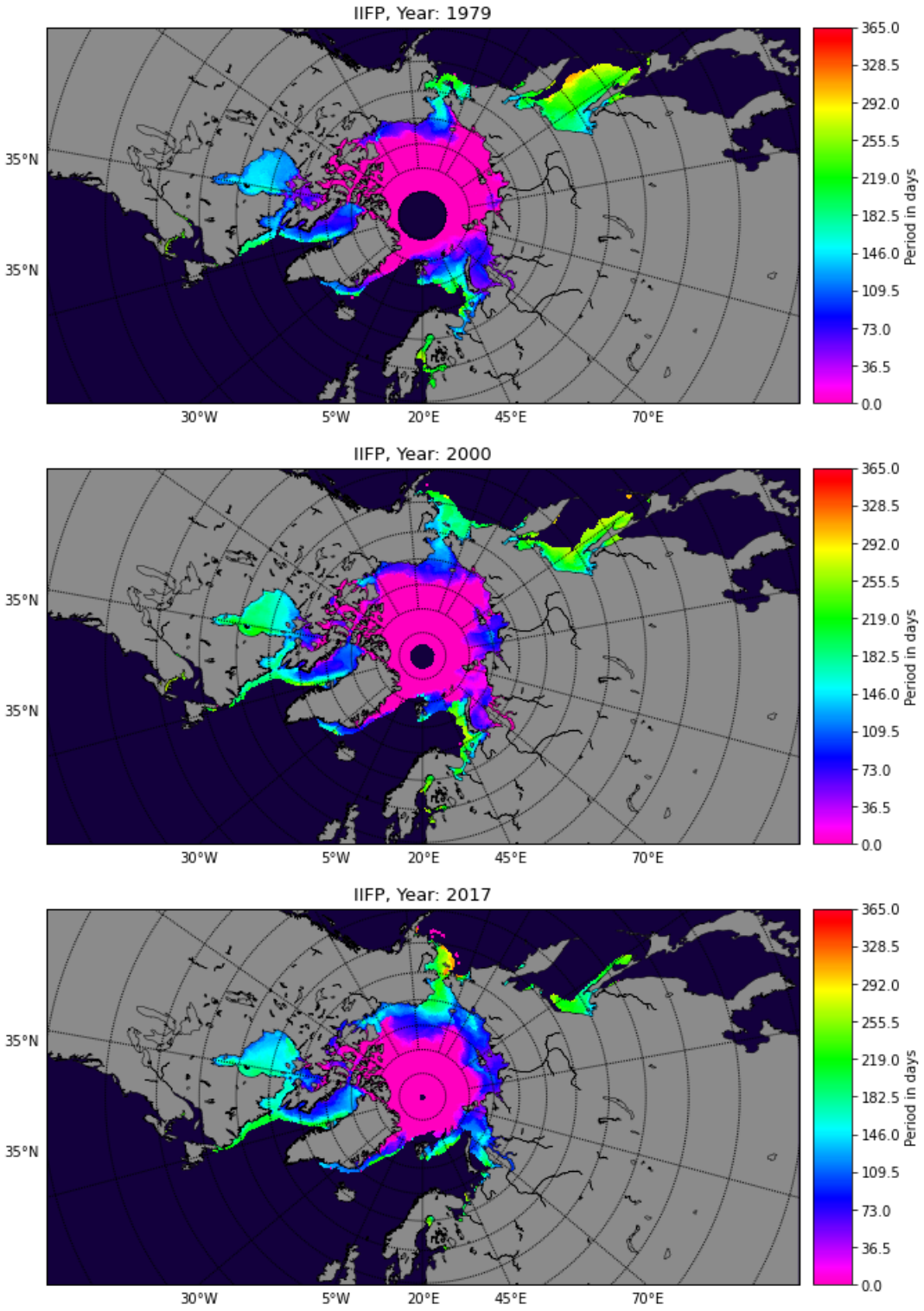


Figure 8: Inner ice-free period, the period in which there is less than 15% concentration of sea ice was plotted for 1979, 2000, and 2017. The pole hole represented an area with zero measurement as there was no change. These years, while random, were chosen to best reflect the changing ice-free period in the Arctic ocean. For more information on the data, one can visit [8] for more information.

Appendix B - Outer Ice-Free Period Raw Data

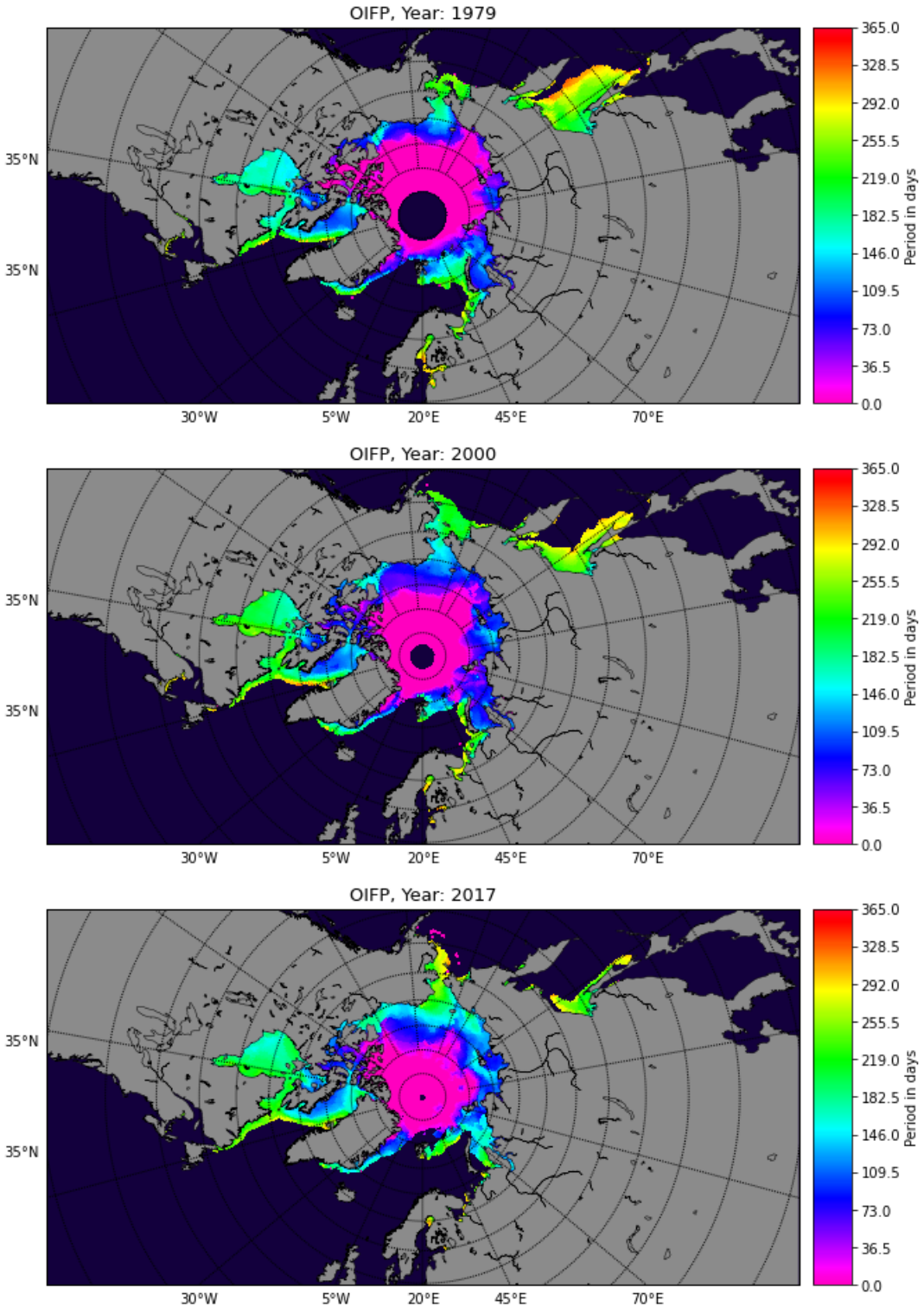


Figure 9: Outer ice-free period, the period in which there is less than 80% concentration of sea ice was plotted for 1979, 2000, and 2017. The pole hole represented an area with zero measurement as there was no change. These years, while random, were chosen to best reflect the changing ice-free period in the Arctic ocean. For more information on the data, one can visit [8] for more information.

Appendix C - Histograms

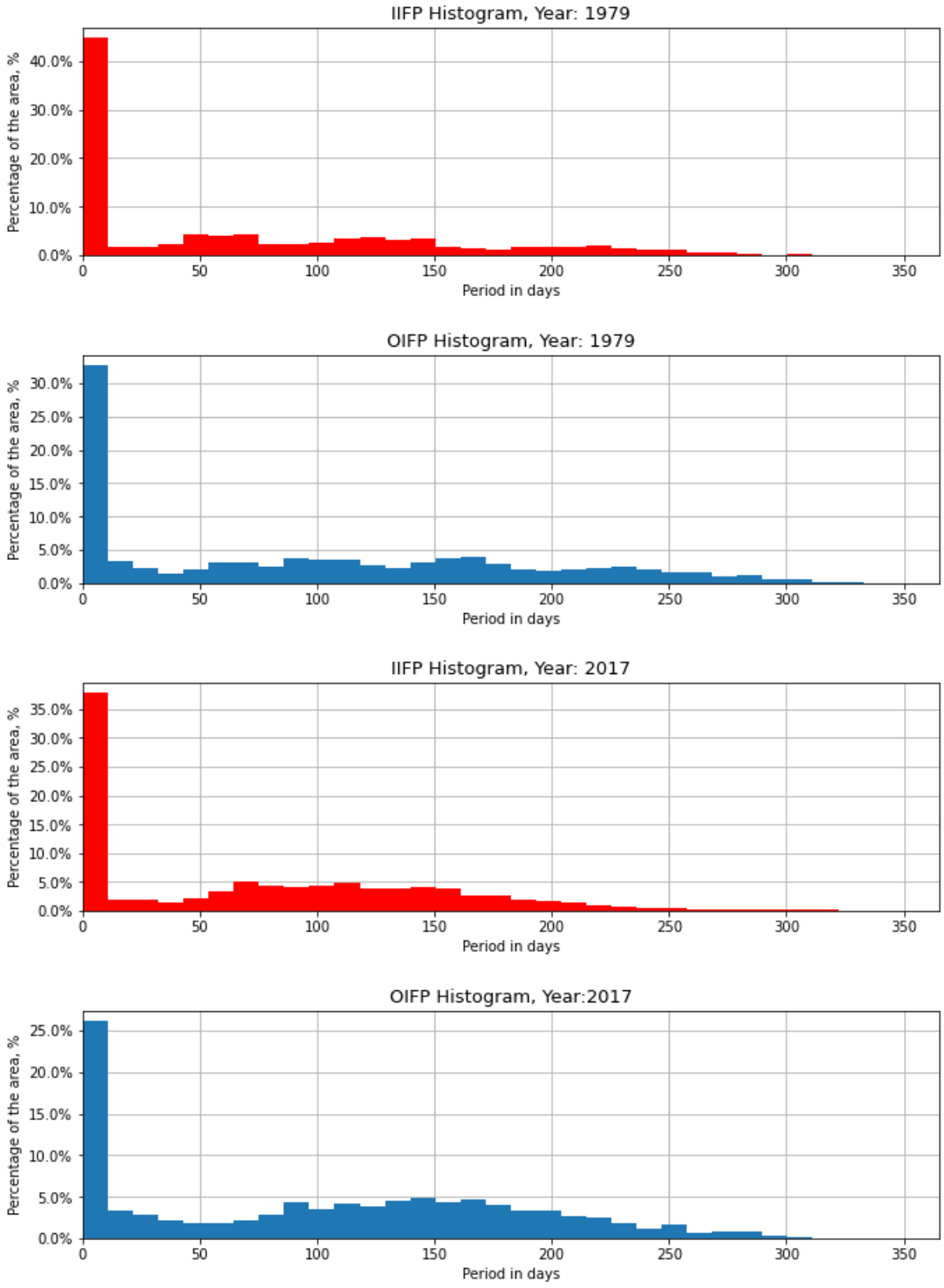


Figure 10: Histograms of the inner and outer ice-free periods for 1979 and 2017. This was done using `matplotlib`'s `hist` plotting function. The histogram shows the length of the ice-free periods as a percentage of the total area. Shifts between 1979 and 2017 are quite evident, signaling there is a trend towards more ice-free days.

Appendix D - Hindcasts and Forecasts

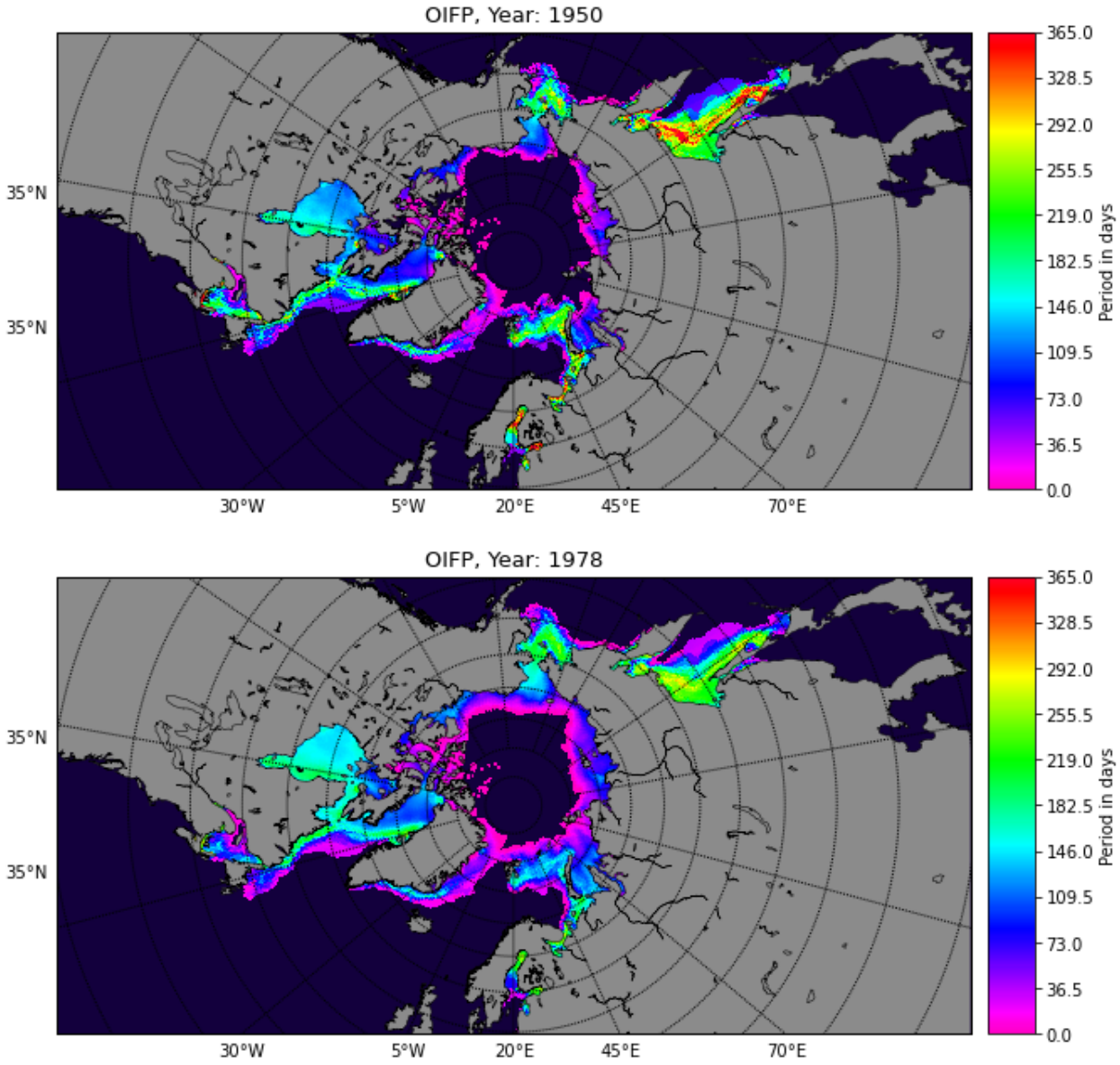


Figure 11: Generated hindcast for the outer ice-free period for the years of 1950 and 1978. A similar technique to **??** and 11 was performed. Again, similar structures between the two come up: the polar experiences distinctive shrinking between 1950 and 1978, and the sea of Japan/Okhotsk has interesting properties.

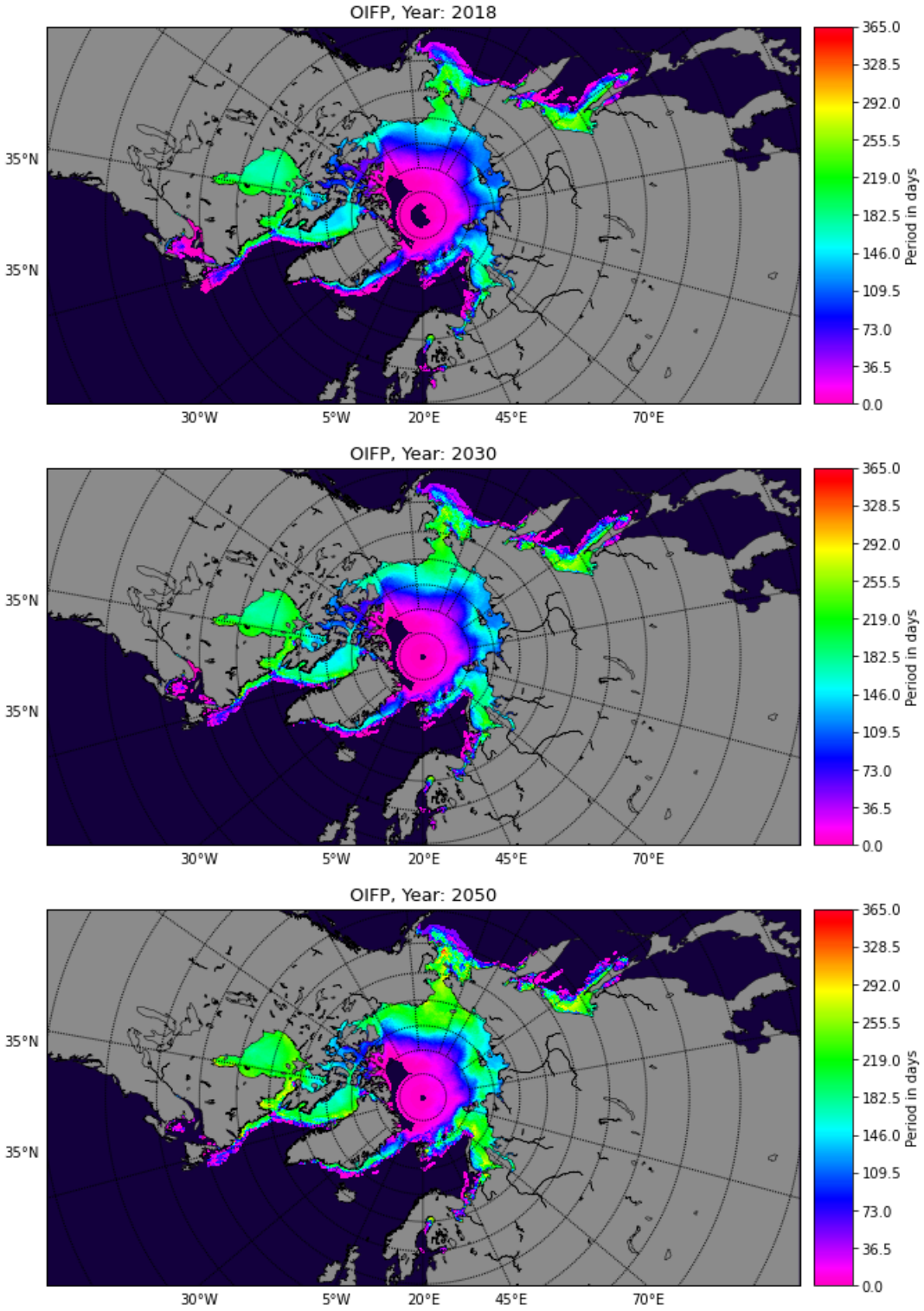


Figure 12: Forecasts for the outer ice-free period for the years of 2018, 2030 and 2050. A similar procedure for the inner ice-free period, as seen in Figure 7 was applied here.0

Appendix E - Identified Regions and Forecasts

Figure 13: This figure shows regions of importance as according to Figure 4. These areas have been identified as having interesting change of rates in their sea ice melt periods. The visuals below are to aid in comprehension, however, actual values in the physical realm may differ, as it depends heavily on the map projection chosen. The projection `gnom` in `basemap` was chosen to easily view the Arctic.

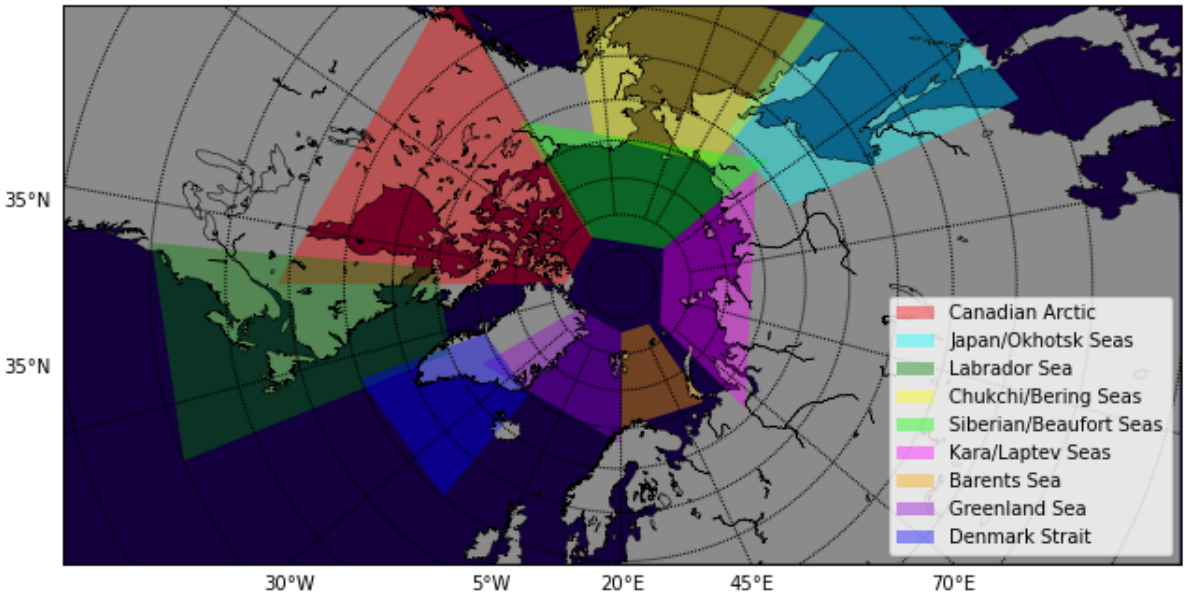


Table 2: This table shows the regions of importance, as plotted in Figure 2. The general values of longitude and latitude were taken from Google Maps. The total longitudinal range spanned from 180°W to 180°E. The total latitudinal range spanned from 42°N to 90°N. These values are pasted here for additional reference.

Region:	**Longitudinal Range:**	**Latitudinal Range:**
Canadian Arctic	70°W to 130°W	49°N to 82°N
Japan/Okhotsk Seas	135°E to 165°E	42°N to 65°N
Labrador Sea	48°W to 75°W	40°N to 65°N
Chukchi/Bering Seas	150°W to 162°E	50°N to 69°N
Siberian/Beaufort Seas	128°W to 150°E	64°N to 82°N
Kara/Laptev Seas	65°E to 150°E	66°N to 82°N
Barents Sea	20°E to 55°E	70°N to 83°N
Greenland Sea	40°W to 20°E	68°N to 83°N
Denmark Strait	50°W to 20°W	55°N to 70°N

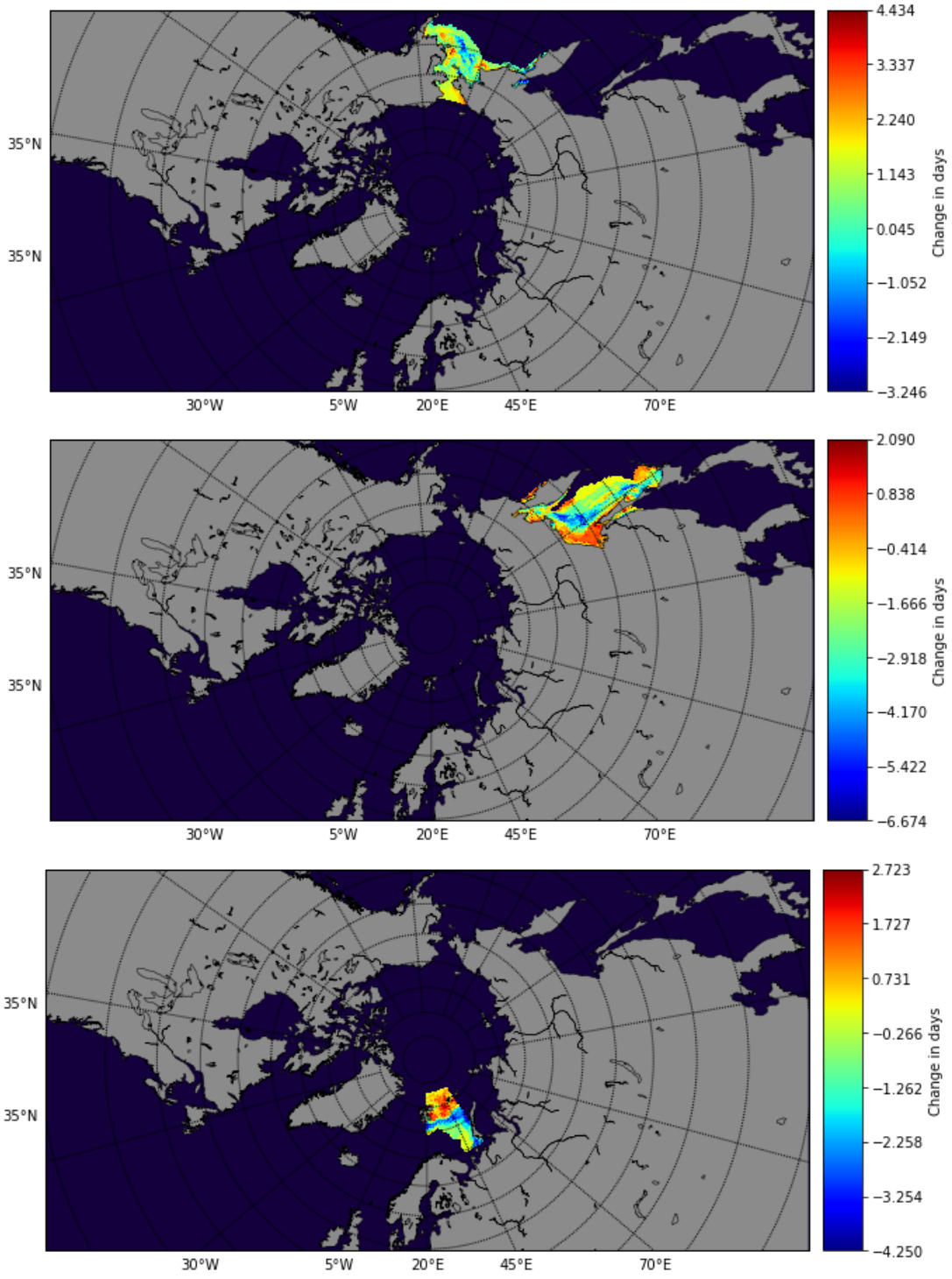


Figure 14: The inner ice-free period rate of change per year in different regions. The different regions chosen for another linear regression were the Chukchi/Bering seas, the Japan/Okhotsk seas, and the Barents sea. This was because all three regions exhibited interesting dynamics in Figure 4. The outer ice-free period rate of changes are not visualized for the individual regions for the sake of brevity.

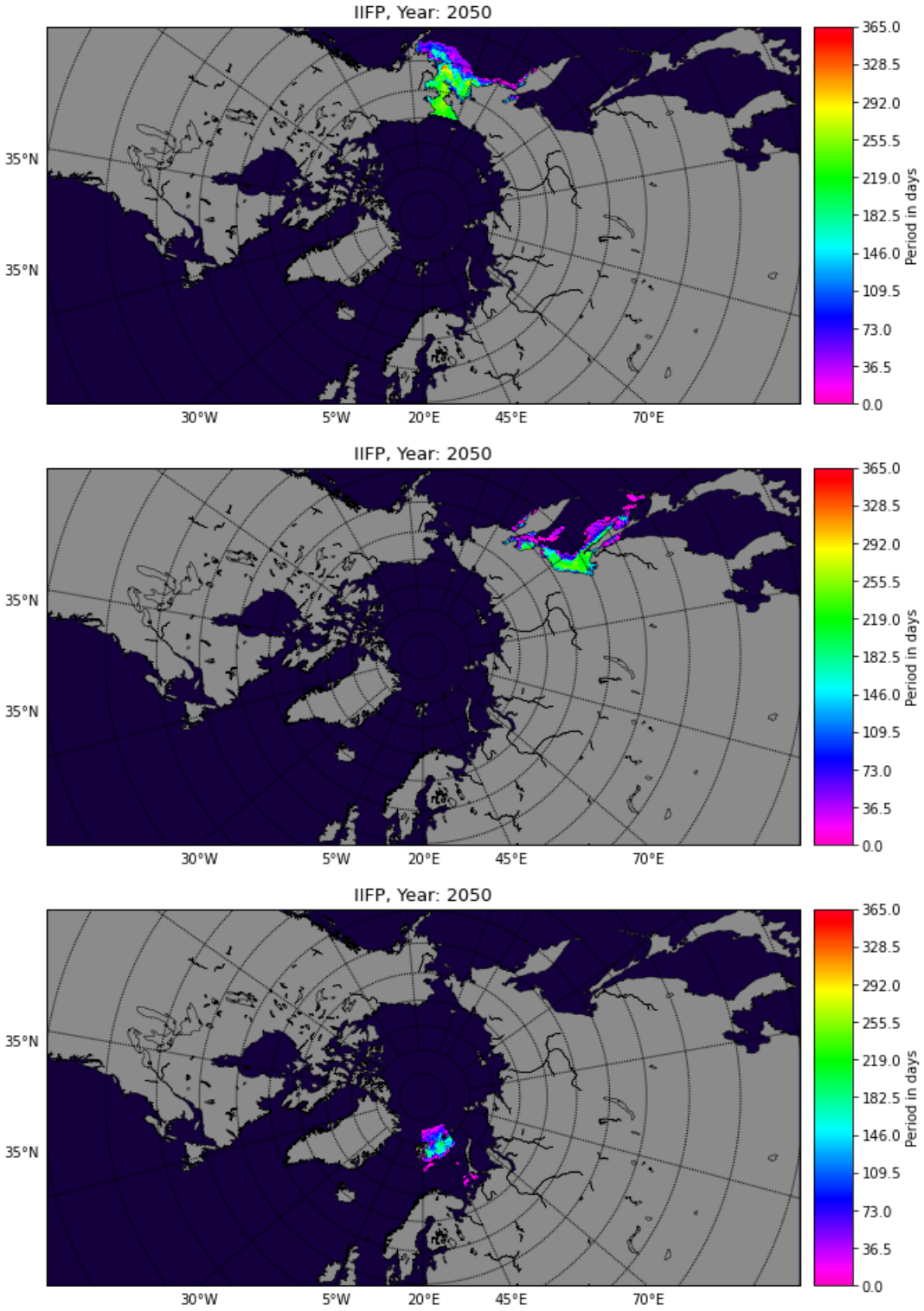


Figure 15: Forecasts of the inner ice-free period for different regions. This was done using the rate of change of each individual region that was calculated in 14.

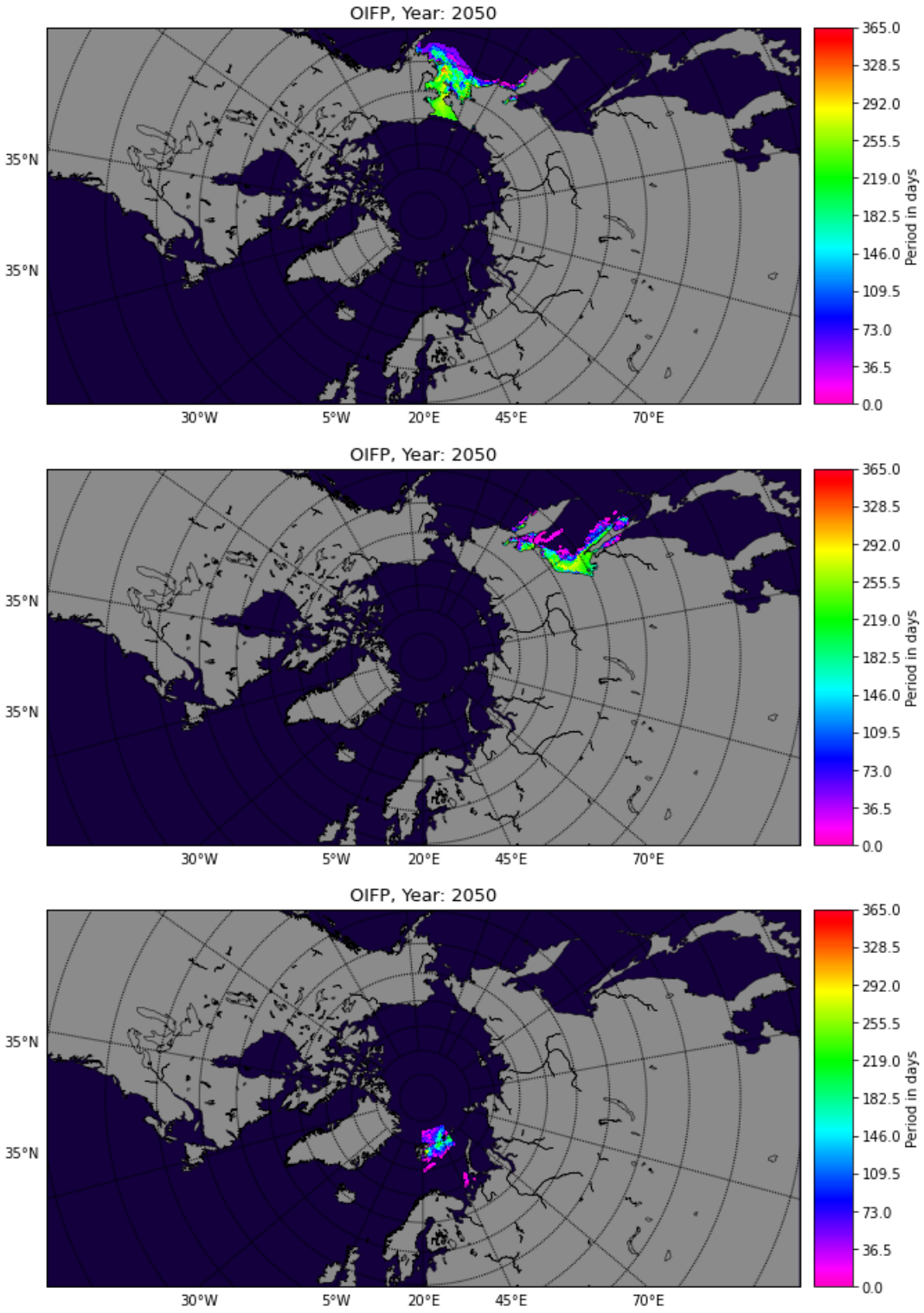


Figure 16: Forecasts of the outer ice-free period for different regions. Again, this was done similarly to the inner ice-free periods, using the rate of change calculated. The rate of changes for the outer ice-free periods are not located in this report for the sake of brevity, as it would be another repetition of 14.