Arctic Explorer Process Book

Ethan Ransom & Dylan Wootton

Table of Contents

Overview and Motivation

The concentration of sea ice plays an important role in creating global climate models due to its ability to influence solar radiation absorption. As such, climate researchers are interested in understanding patterns of change in ice concentration.

Our tool, Arctic Explorer, displays data relevant to the concentration of sea ice. This tool helps researchers understand both the spatial and temporal nature of sea ice change. Ice data has traditionally been viewed using geographic projections. While visually engaging, these maps often make it difficult to explore temporal trends in the data.

By providing both a spatial view and a temporal view of data, Arctic Explorer allows researchers to understand how sea ice changes over time. Our collaborators in the Golden Lab, have expressed difficulty understanding temporal patterns in the data. This tool was built to address this aim.

Using these patterns, our collaborators plan to adjust models to account for how the Arctic impacts the global climate. Such improvements would provide a better prediction of climate change and inform strategies to prepare for such changes.

Related Work

We were introduced to the dataset and problem space by the [Golden lab](). During a scoping meeting with a researcher in the Golden Lab, Delaney Mosier, we were introduced to the data in the form of a video recording of concentration values over the course of a year. Below is a screenshot from this video.
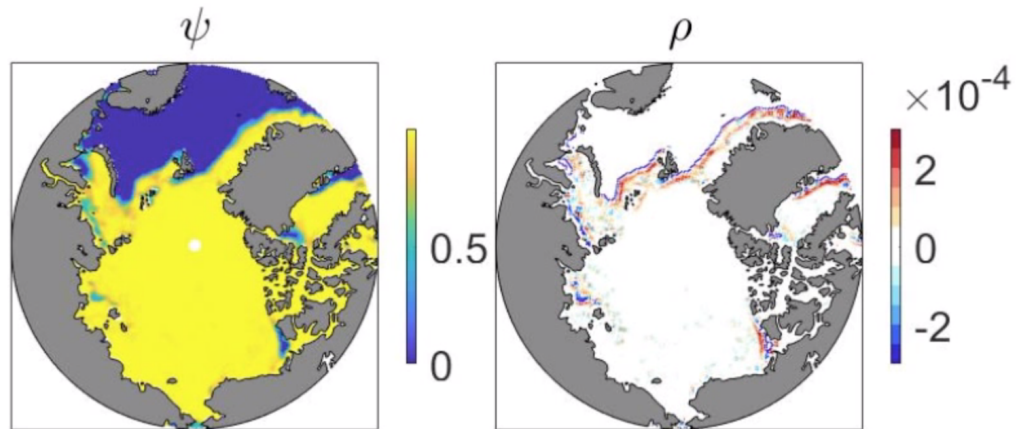
Questions

Our project currently aims to answer the following questions:
- How does global sea ice concentration change over months in a year? In years in a decade?
- Do temporal patterns exist in total concentration? Do temporal patterns exist at specific latitudes and longitudes?
- How do more recent melting trends compare to older trends? Is sea ice melting faster now? Is it growing slower? If so, by how much?

The following questions were our past question list:
- How does global sea ice concentration change over months in a year? Over years in a decade?
- On average how does sea ice concentration change throughout a year? Throughout a decade?
- Are there specific locations that are changing more rapidly than others? Do they exhibit any unique temporal patterns?
- At a given latitude and longitude, what months of a year is that location free of ice? Is that pattern consistent across years?

The original questions were created from an initial interview with Delaney. From this meeting, we realized that the lab currently lacked the functionality to analyze the temporal aspects of the sea ice data. However, during our exploratory data analysis, we realized that it was necessary to be more specific with the questions we were asking. This included specifying which aspect of the data we should analyze (ie total concentration, average concentration, total coverage, *etc*...) and how we would analyze it (by comparing it to older time points, by looking at the values isolated, *etc*...).

Data

Data was obtained from the [NOAA Sea Ice Extent Dataset](). The data was obtained from the FTP server via Batch Link Downloader. A MatLab script, SateliteData.m was obtained from the researchers and modified. This involved parsing the .nc file and transforming it into JSON encodings.

Two data structures were built using these maps. The first one is a three dimensional JSON Matrix. Each 'slice' of the matrix represents the 2D map of concentration values at given latitudes and longitudes. This representation is ideal for displaying the concentration on a map as it is easy to access all of the concentration values at a given time.

The second representation was to build an object that contained one field per "gridsquare", with each field representing one latitude and longitude pair. These fields store the time series values for that location. This representation is ideal for displaying how concentration changes over time. Additionally, this representation requires less space as we were able to compress many of the arrays due to repeated values. For example, the values that were far away from the arctic had 0% concentration year round.

Exploratory Data Analysis

We used Delaney's ice concentration analysis an an initial visualization. This animation confirmed the quality of the dataset. It was a continuous, gap-free sampling of ice concentrations with very high temporal and spatial resolution. As the animation plays through the course of a year, we could clearly see the retreat and advance of the ice through summer and winter, respectively.

Additionally, exploratory analysis was conducted on the temporal aspect of the data. Using Matlab, a simple script was used to visualize one latitude and longitudes data over a given time period. A compilation of 3 location's time series data is displayed:
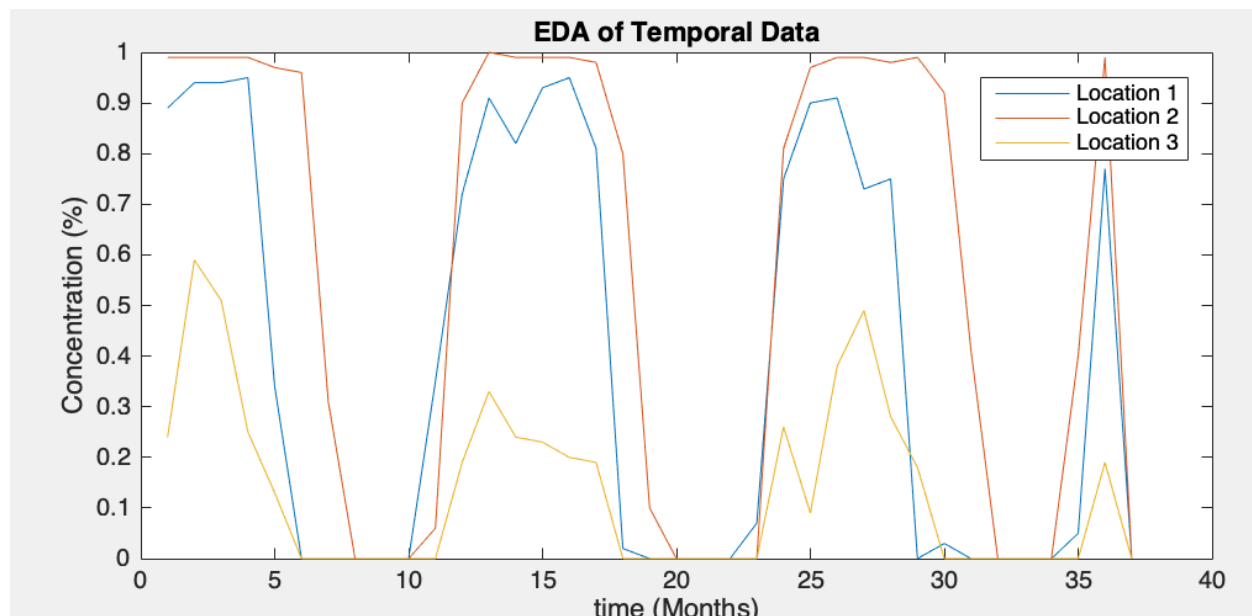


From both of these explorations, we could see that there were strong seasonal differences between the summer and winter months. Given the periodicity of this, we decided to implement a feature to let our user select which allows for them to select the datapoints they want to view (such as only view data from January, only view data from the winter months, or only view data from a given year).
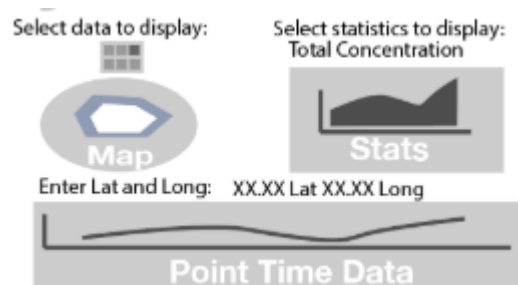
Design Evolution

Throughout the design process, we've iterated through a variety of designs.

Project Proposal Designs:
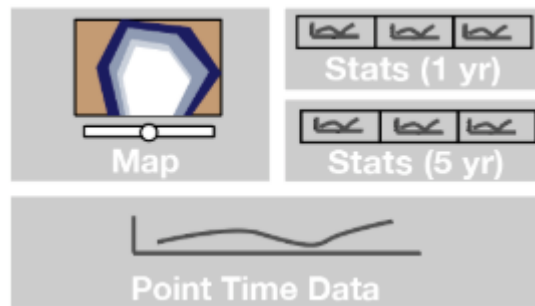
This design was our original design after meeting with Delaney. It captured the fundamental features that are necessary to convey relevant information to the user.

This design was another iteration that we believed would offer more control to the user. Rather than have the slider view (which could make it difficult to select a given data point), we provide the user with the ability to click a calendar and select the specific time point they're interested in. While we didn't utilize this approach in our final proposal design, it was the starting point for our 'heat map' data selection tool that is in our current design.

Contour map of sea ice concentration.

Point selector to select the lat and lon to view the data

Toggle button to fix the month and day. Allows researchers to see changes across years.

Arctic Sea Ice Concentration

☐ Display Sources and Sinks

Toggle to add sources and sinks overlay on map.

March 22nd, 1999

Time slider to allow user to select which dataset to view.

☐ Hold Month and Day

Summary Statistics for March 22nd and the Past 12 months

Change in total coverage	Change in total concentration	Average rate of change

Past year of aggregated statistics.

Summary on March 22nd for the past 5 years

Change in total coverage	Change in total concentration	Average rate of change

Past 5 years of aggregated statistics.

Sea Ice Concentration at Selected Point Over Time

Concentration

Time

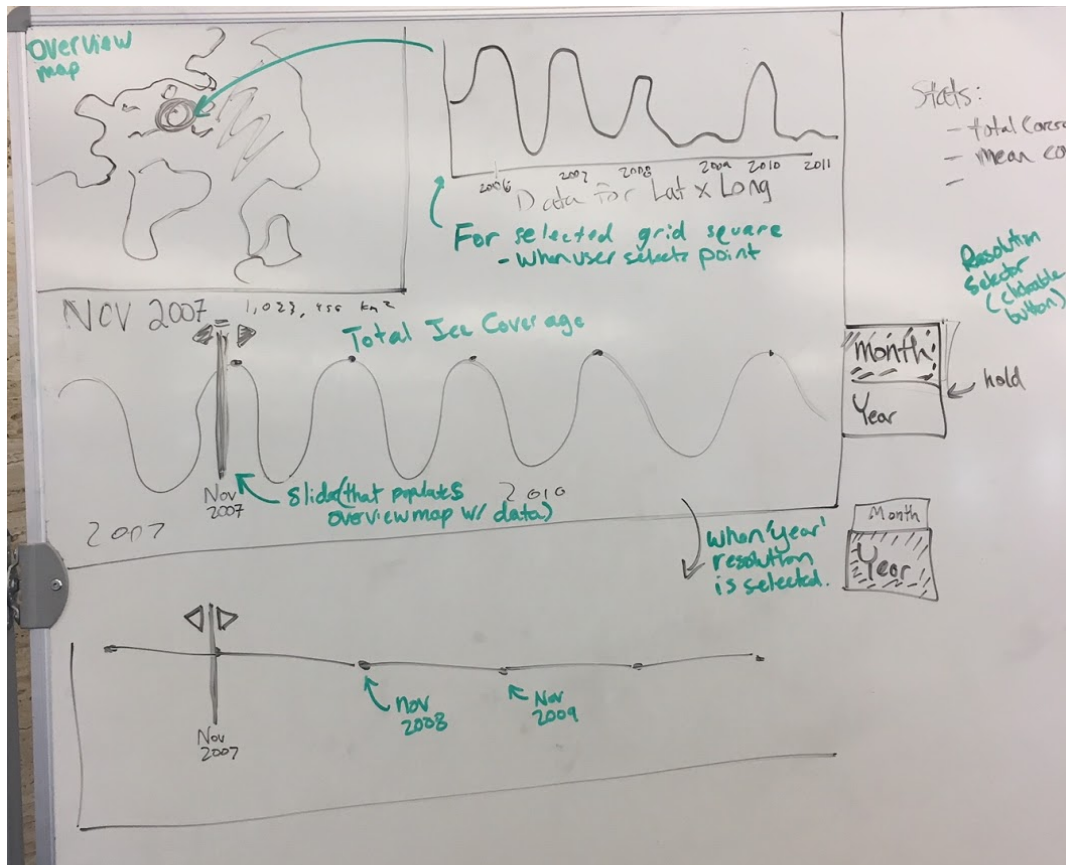Time series chart of concentration at a given point. Possibly display other info (sunshine or ocean temp)
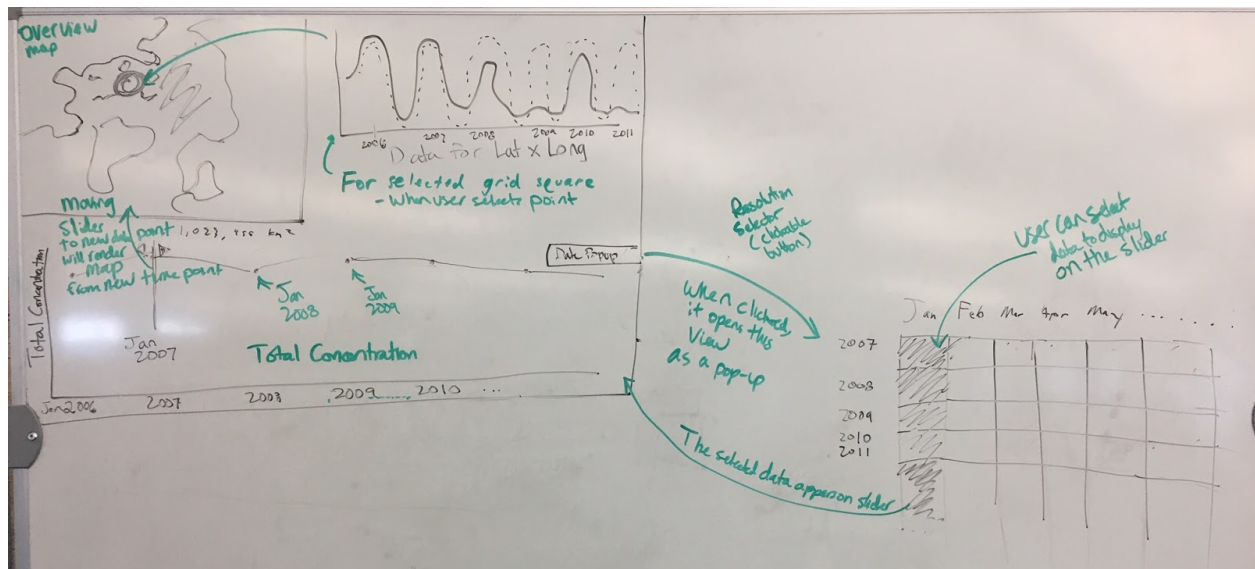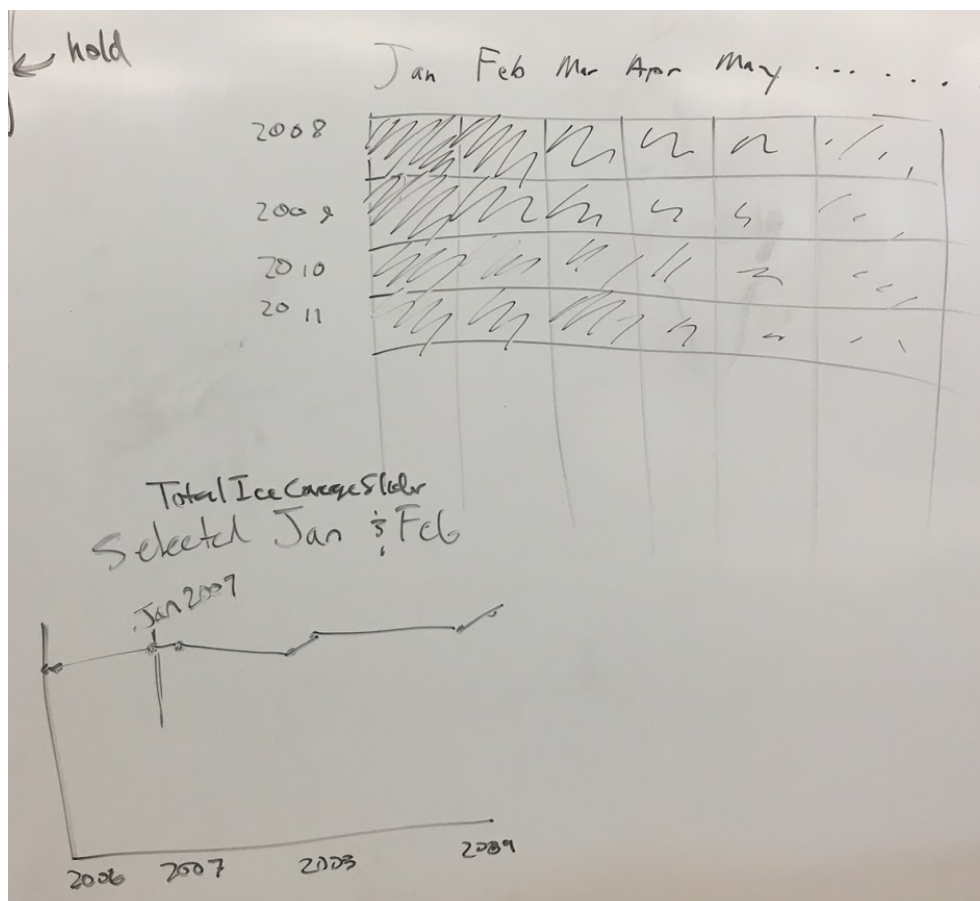
Additional Designs:

Design 2a:

This design incorporates a slider that slides across a chart of total concentration. The user can move this slider back and fourth, and when they move it to a new month, it populates the overview map with the corresponding data.
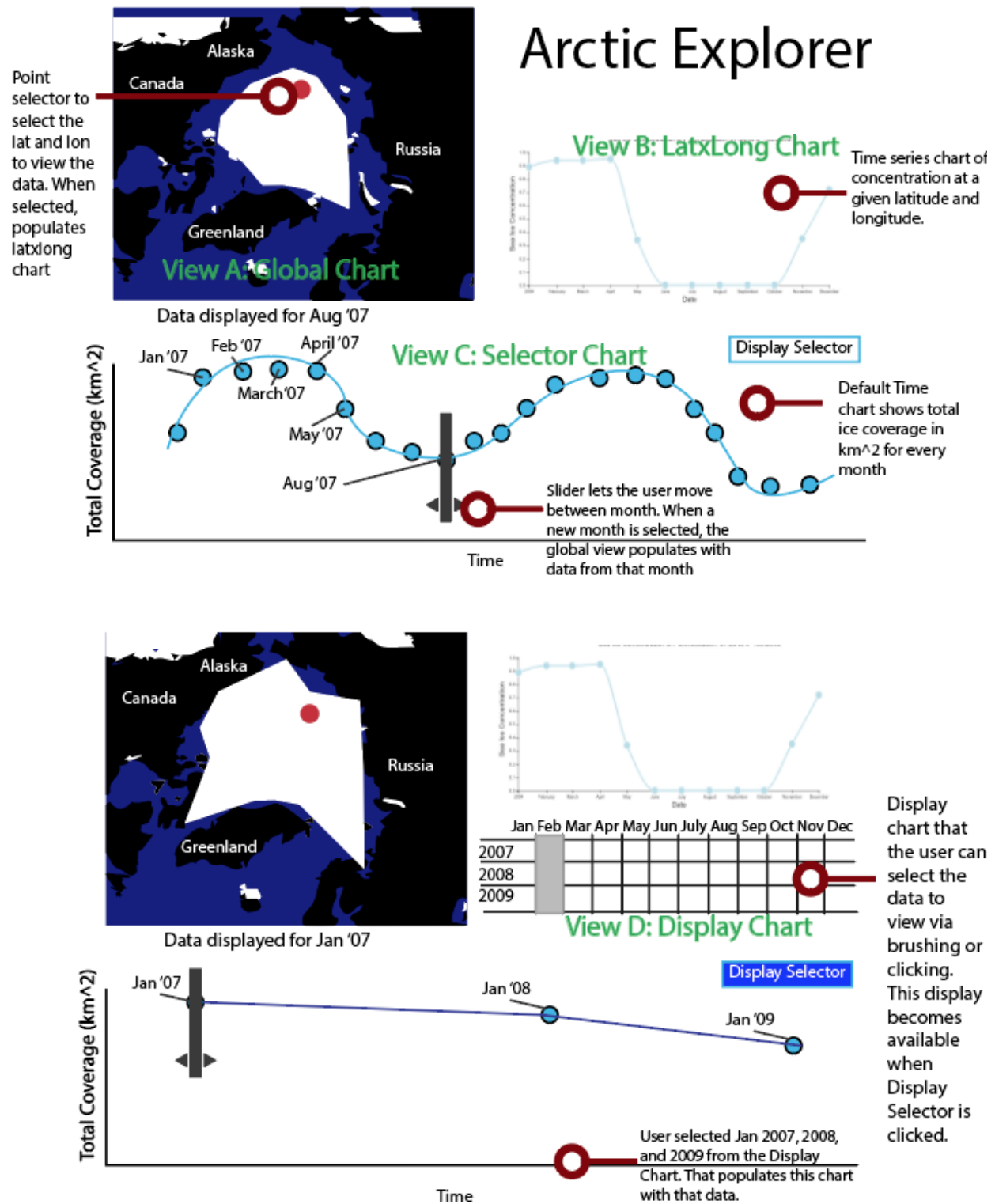
Design 2b: Heatmap Incorporation

This view allows the user to manually select or to brush and select the data to display. The user can click on the 'heatmap chart' (displayed to the right) to select the months/years they want to view.

Display chart in more detail.

Design 2b: Digitized

Arctic Explorer

Point selector to select the lat and lon to view the data. When selected, populates latxlong chart

Alaska
Canada
Russia
Greenland

View A: Global Chart

Data displayed for Aug '07

View B: LatxLong Chart

Time series chart of concentration at a given latitude and longitude.

Sea Ice Concentration
Date

View C: Selector Chart

Total Coverage (km^2)

Jan '07
Feb '07
March '07
April '07
May '07
Aug '07

Time

Display Selector

Default Time chart shows total ice coverage in km^2 for every month

Slider lets the user move between month. When a new month is selected, the global view populates with data from that month

Alaska
Canada
Russia
Greenland

Data displayed for Jan '07

Sea Ice Concentration
Date

	Jan	Feb	Mar	Apr	May	Jun	July	Aug	Sep	Oct	Nov	Dec
2007												
2008												
2009												

View D: Display Chart

Display chart that the user can select the data to view via brushing or clicking. This display becomes available when Display Selector is clicked.

Total Coverage (km^2)

Jan '07
Jan '08
Jan '09

Display Selector

Time

User selected Jan 2007, 2008, and 2009 from the Display Chart. That populates this chart with that data.

12

This final design ended up being nearly identical to the final visualization tool. This design allows for the user to have as much control over the data as possible, and it allows for them to do custom queries which were not possible in past designs.

This design also leverages position to encode data about the total coverage. This will allow for our users to quickly understand patterns inside of the data. The query selector helps to bolster this ability as it allows the user to 'ignore' other data that may make it difficult to see relationships over time (such as the decline in total coverage as is displayed in the past design.

Implementation

This figure shows the final state of the project.
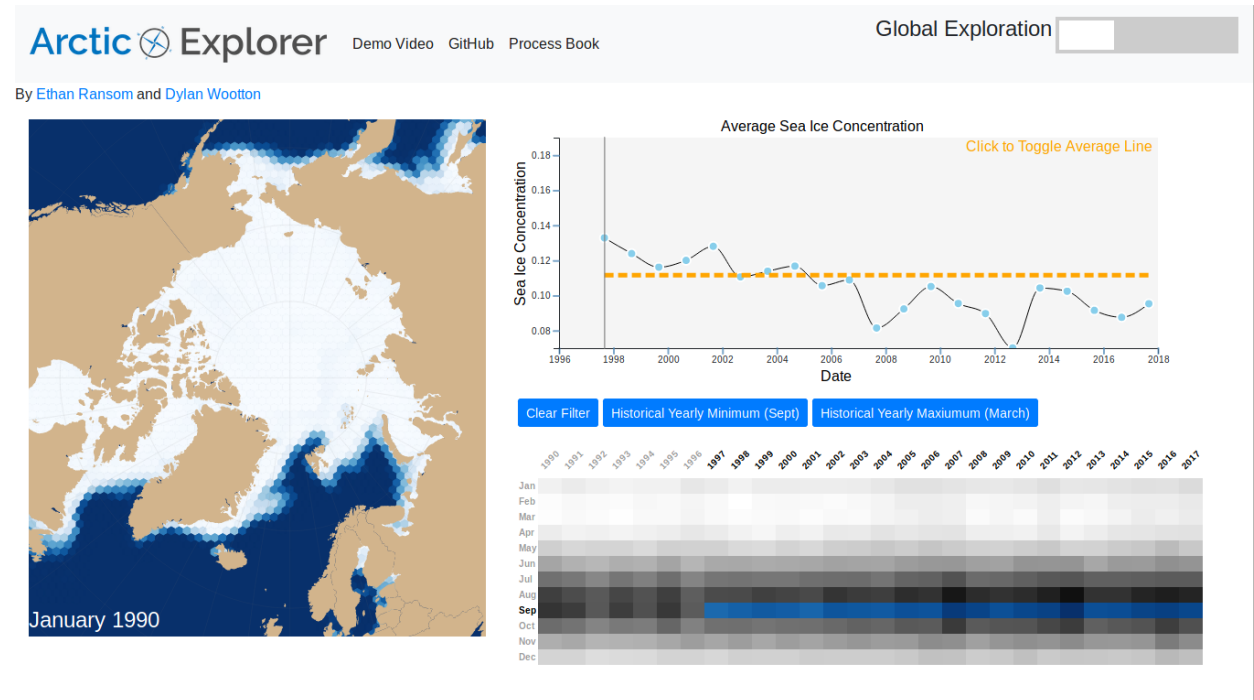
The main components are as follows:

- **Map View:** a map of the arctic region, found on the left side of the screen. This map shows the concentration of sea ice, as captured at a given instant by NOAA satellites. Blue denotes open ocean, while pure white denotes 100% sea ice coverage.
- **Average Concentration View:** a line chart found in the top right, labelled "Average Sea Ice Concentration". Clicking on a particular datapoint will update the Map View to display the state of the arctic at that time. By default, the line chart displays a dashed orange line. This line shows what the historical average is for a given month.
- **Heatmap View:** a heatmap, displayed on the bottom right. This view displays the average concentration for each month by shading a cell in a table. Clicking and dragging on the table cells as well as the year and month headers allows the user to filter the date ranges displayed on the Average Concentration line plot. For instance, to plot sea ice concentration each March for the last 10 years, the user could click on the cell corresponding to March of 2007, and then keep the mouse button depressed while sliding the mouse to the right until they reach March of 2017. Two preset filters are available, the "Historical Yearly Minimum (Sept)", and the "Historical Yearly Maximum (March)". Pressing "Clear Filter" will clear any filtering and display all available datapoints on the line plot.

Navigator Mode

We worked with sea ice researchers in creating this tool, and they expressed interest in being able to explore the status of important routes through the arctic. To investigate this, users can click the toggle slider in the top right corner of the application to switch the app into "Navigation" mode. In this mode, clicking on the Map View allows the user to place waypoints in a path. Ice concentration values between the waypoints are interpolated and their corresponding concentration values are displayed in the new concentration map chart. (Replaces the line chart in the top right portion of the page.) The y direction on the chart corresponds to the distance along the path and by hovering over a heatmap cell, you can see more information, such as the location and concentration value. Clicking on a cell renders that month's data.
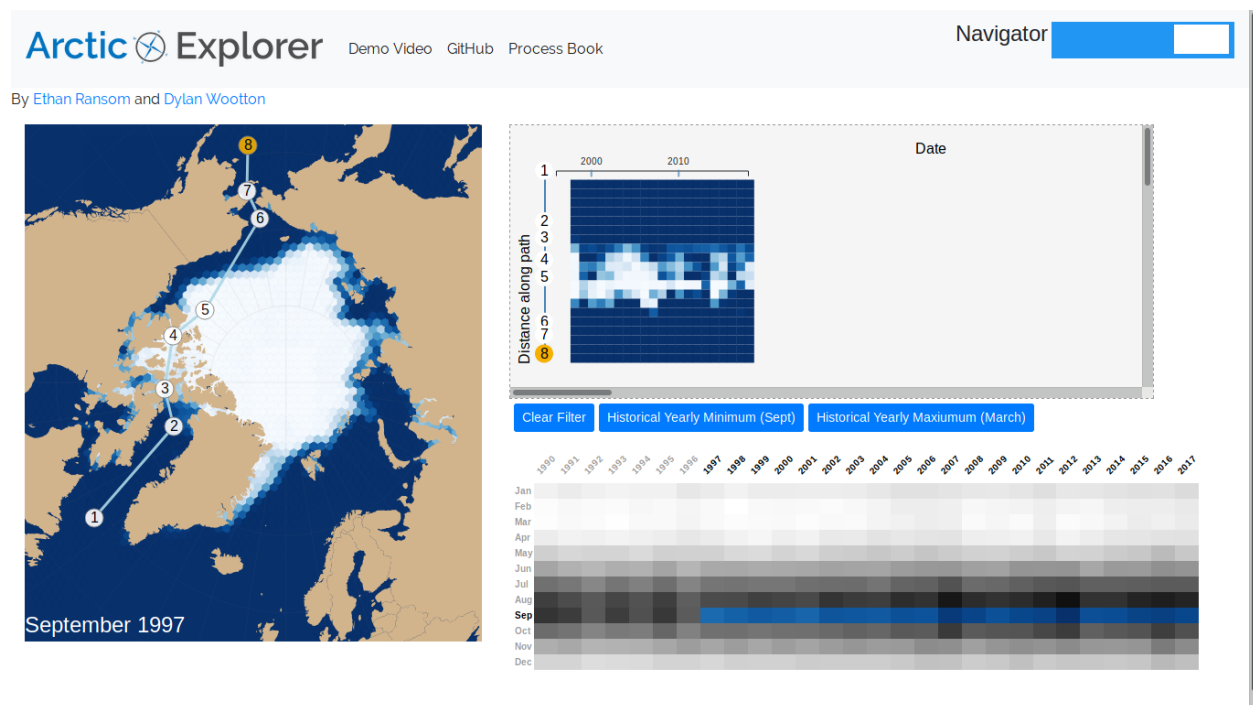

Figure: Using the Navigator mode to explore the historical coverage of ice along the famous Northwest Passage.

Please see Appendix A: Progress Journal at the end of this document if you are interested in a full description of the progress of the implementation up to this point.

Evaluation

We performed several informal user studies to evaluate the usability of the design. These studies informed a decision to remove scroll-to-zoom functionality from the average concentration line chart, as it was frequently triggered by accident by our users. Many users were impressed by the dataset and the features available to them, however the discoverability of these features could be improved. Future work on this tool may therefore include directed storytelling functionality, both to tell the story of widespread sea ice decline in the arctic, and to inform the user of the features available to them.

Appendix A: Progress Journal

After a cursory investigation, we were unable to find any information on the ".nc" format that the raw NOAA data is provided in. As a workaround, we modified Delaney's MATLAB script (which can read the .nc files) to output three .csv files for a single timepoint, one for lat data, one for lon data, one for psi data. The row and column indicies could be used to tie the files together and get a {lat, lon, psi} tuple for each gridpoint. We built a simple script that could "zip these csvs together into one csv, with each row representing a gridsquare.

By taking these csv files, as well as the world.json file from Homework 4 and using the Gringorten Quincuncial projection we were able to render the first iteration of our map view:
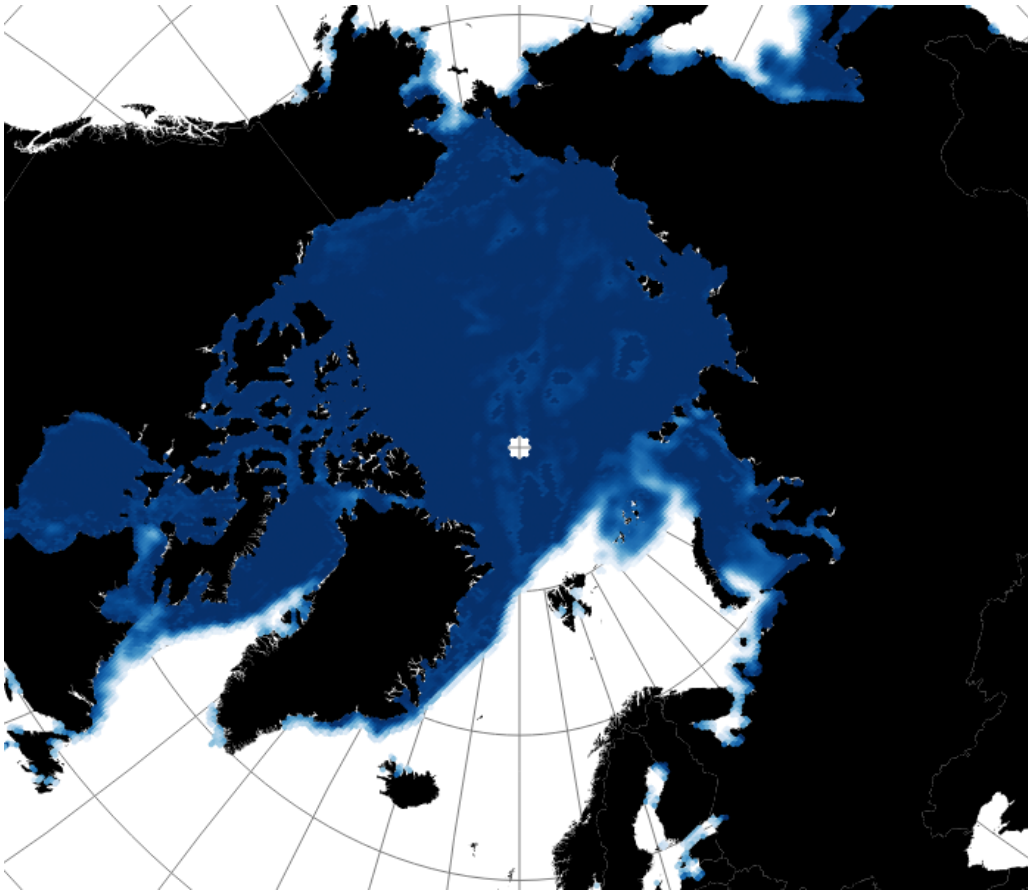

Figure: first iteration of the map view. Counterintuitively, blue is ice and white is ice-free ocean.

It was at this point that we began to run into issues with both the space and time runtime efficiency of our visualization. The map would take several seconds to render, and the ice grid squares would take upwards of 10 seconds. To solve the map issues we used the Mapshaper tool (https://mapshaper.org/) to clip the map features to north of 30 degrees latitude. This helped only marginally. We also used Mapshaper's simplify feature to downsample the detail of the

landmasses. This lead to a major improvement in map rendering times. To fix the rendering of the gridsquares, we took another look at the data. The data contained "NaN" values for gridsquares who fell onto a landmass. Additionally, for many datapoints, especially those in summer months, a large number of gridsquares contained zeroes, which for rendering purposes could be ignored. This cut down the number of gridsquares to render from ~130,000 to only ~20,000 for our initial timepoint.

Next, we aimed to render multiple months' worth of data at a time. Our first iteration of a new MATLAB script outputted an array of 2-dimensional matrices, which the rows and columns of the matrix corresponding to the rows and columns of the gridsquares. (Because the gridsquares are stationary, the latitude and longitude of the gridsquares can be stored in a separate file.) This format proved to be somewhat space-inefficient. Our estimates showed that a file storing 30 years could be as large as 1 Gigabyte. We've moved to what we believe to be our final format, which means storing gridsquares first, with each gridsquare storing an array containing its value for each timepoint. In this format, gridsquares which fall onto land, and thus whose only value is "NaN", can be eliminated.

To complete our visualization up to the point of the milestone we implemented a slider to allow for year selection, and a line chart showing the sea ice coverage of a specific gridsquare. (In the final visualization the gridsquare will be chosen by the user.)
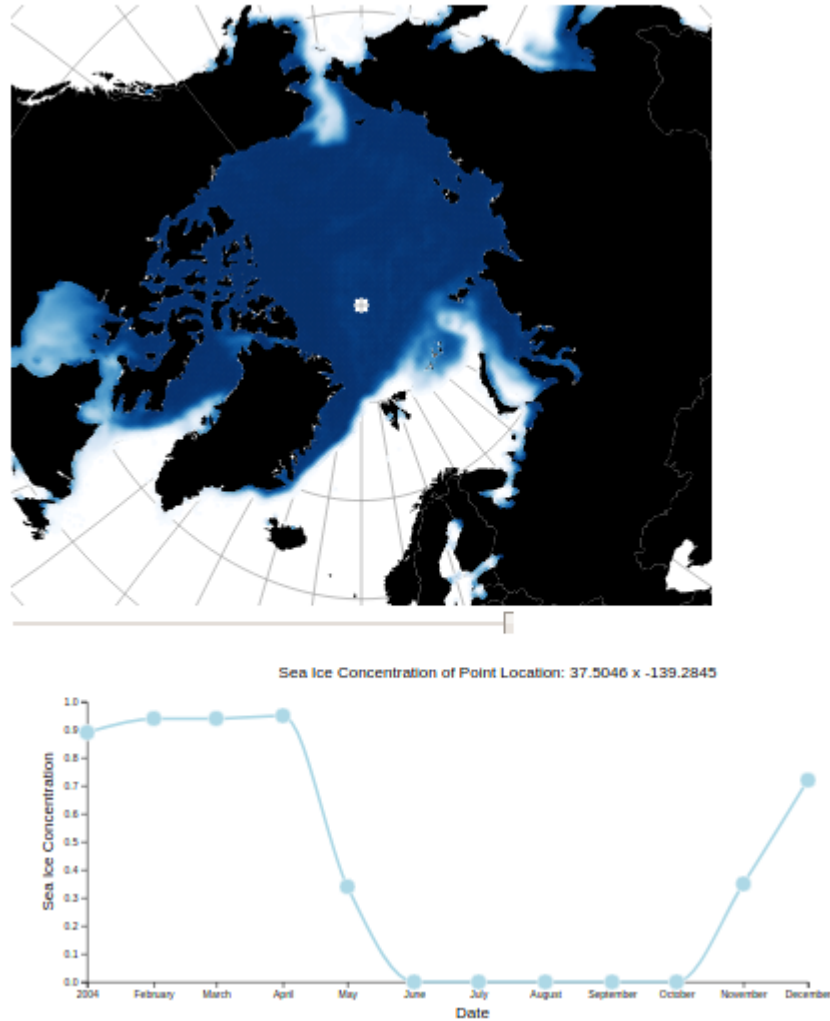
Figure: The state of our visualization at the milestone.

The top view is a map of a polar projection of the north pole. Counterintuitively, the deeper the blue, the higher the sea ice concentration. The dataset is the concentration at each month in the year of 2004. (12 datapoints.) The slider under the map can be used to select the month to be displayed on the map. The initial render of the landform outlines takes much less than a second, however the first render of the sea ice takes around 3 seconds. We are looking into optimizing the rendering, as well as techniques that could let us coalesce or downsample the data. Fortunately, consecutive renders take less than a second, we believe because d3 is reusing existing circle elements. The bottommost view is a line chart. It displays the concentration of ice of the grid square at the given latitude by month. In the final version of our visualization the user will be able to select which grid square is displayed here.

Updates from milestone:

After the milestone was completed, a variety of work was done to improve functionality and enhance the usability of this tool.

Changes to spatial map:

The original point based implementation of the spatial map was too slow to render- often taking multiple seconds to render one month's worth of data. To fix this, we implemented a contour map using d3 contours; however, due to the nature of our data in a non-rectangular form, we encountered a variety of problems seen below.
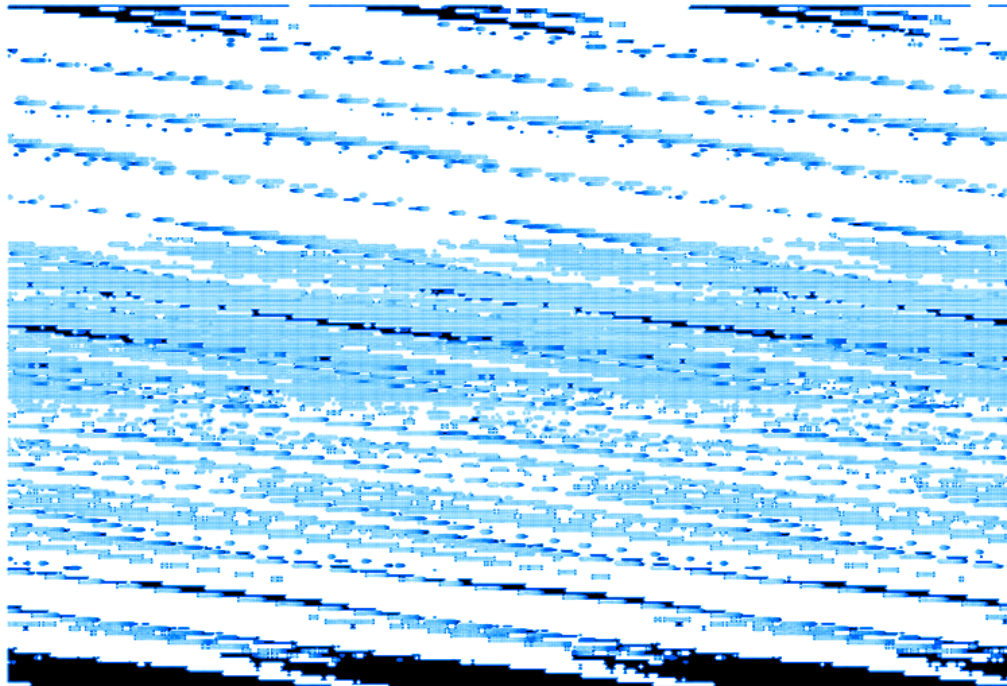

Our "contour" map using d3-contours.

After this failed attempt, we utilized the d3.hexbin library to bin our data into representative hexagon and visualize the data. This is the current implementation of our map, and it improved run time by an order of magnitude.

Changes to Locational data:

In the milestone, we had accomplished the creation of a time series chart that displayed the sea ice concentration for one of our data points; however, we decided such information wasn't particularly interesting after discussions with our collaborators. Instead we worked to expand this utilizing the current heatmap implementation of path data. This involved creating a variety of prototype drawings as seen on the next page:
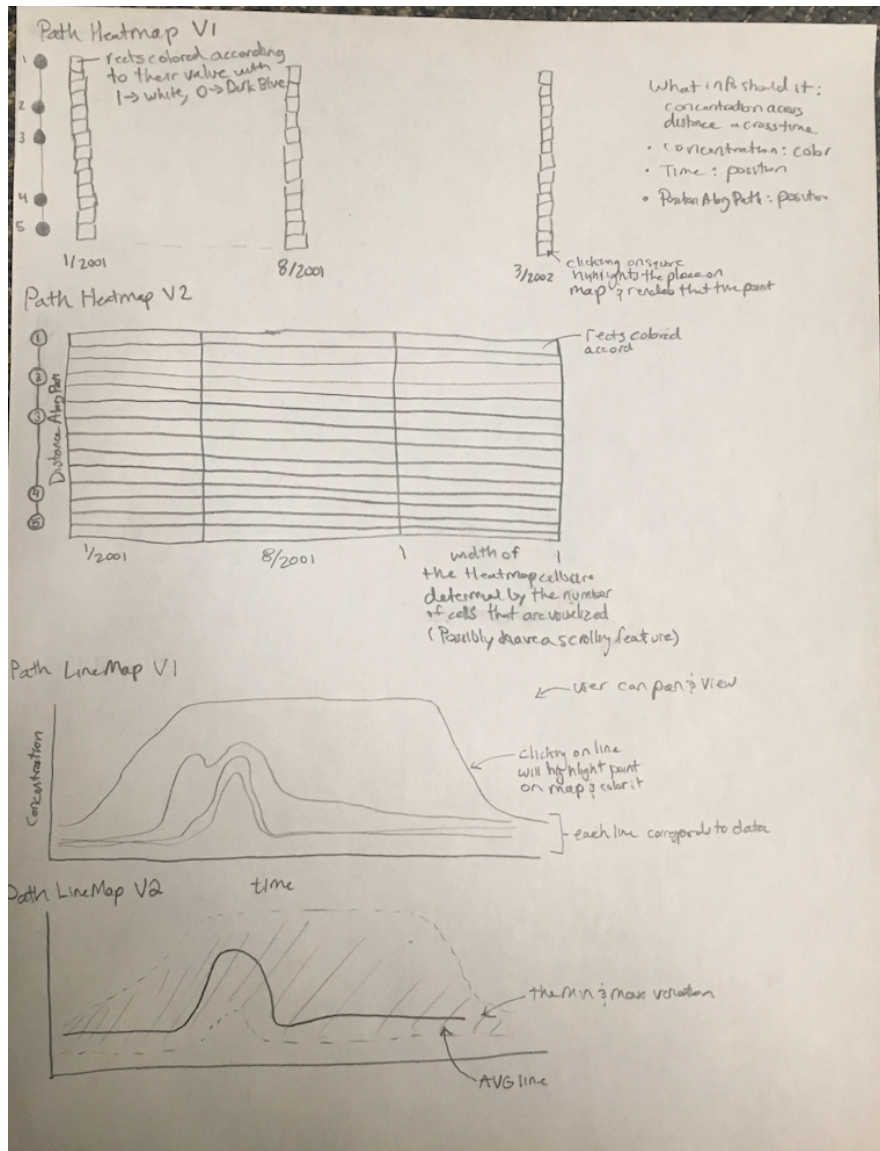
After creating these illustrations, we realized that the path heatmap V1 would best encode the relevant information and help our users better understand how ice concentration changes along a path.

Changes to time selector:

At the time of milestone our selector was a basic html slider; however, we realized that this slider would hamper our users ability to select relevant data. As such, we can up with a heatmap 'query selector' that could be used to look at ice concentration and drag/brush this map.

Changes to aggregated statistics:

While we were originally interested in a variety of statistics, throughout the development of this tool, we realized that the important information is truly captured by average sea ice concentration. This is a important measurement that many researchers use as its a good proxy for

overall arctic condition. As such, we implemented a time chart that would display how this average concentration changes over time.

Work after the milestone:
- Displayed a heatmap of average concentrations on the entire dataset
- Enabled brushing of this heatmap to filter the dates displayed on the line chart
- Used hexagonal binning to improve the rendering performance of the map view (approx 10x faster)
- Navigator Mode: Implemented the ability to place a route on the map view, and display the ice concentrations along this route on a view.