A Database of Weekly Sea Ice Parcel Tracks Derived from Lagrangian Motion Data with Ancillary Data Products

Matthew Tooth ¹,*,†,‡ and Mark Tschudi ²‡

¹ Department of Aerospace Engineering Sciences, University of Colorado, Boulder, CO 80309, USA
² Colorado Center for Astrodynamics Research, University of Colorado, Boulder, CO 80309, USA;
  mark.tschudi@colorado.edu
* Correspondence: matthew.tooth@colorado.edu; Tel.: +1-303-492-3105
† Current address: ECNT 320, 431 UCB. University of Colorado, Boulder, CO 80309-0431, USA.
‡ These authors contributed equally to this work.

Received: 13 June 2017; Accepted: 8 August 2017; Published: 15 August 2017

Abstract: Arctic sea ice has been on the decline over the past several decades, and multi-year sea ice has decreased significantly in its areal share of the overall sea ice cover. Changes in several key variables such as radiative balances, albedo, ice surface temperature, and ice thickness have driven much of the decline, but the motion of sea ice makes studying the effects of these variables on individual parcels difficult. Previous studies have observed changes in the means of these variables and their impacts on sea ice concentration, but an accessible database of Lagrangian tracked data is not yet available for study. In order to address this, a database has been developed at the University of Colorado Boulder that performs Lagrangian tracking on individual sea ice parcels and saves coincident ancillary thermodynamic and dynamic variables for each parcel on a weekly timescale.

Data Set: https://doi.org/10.1594/PANGAEA.871925

Data Set License: CC-BY

Keywords: sea ice; data descriptor; climate; Lagrangian tracking; cryosphere; EASE-Grid

1. Summary

As the Arctic continues to transition toward an ice-free summer, stakeholders in the region will benefit from a greater understanding of the variables that drive the loss of sea ice in the region. Multi-year sea ice is of particular interest in studying changes in the Arctic, as it is typically thicker and stronger than younger sea ice floes, and is more likely to survive the summer melt season. The decline of overall ice extent in the Arctic has corresponded with a significant drop in the percent areal share of multi-year sea ice floes, which has weakened the Arctic ice against future climate forcing [1,2]. Further investigation of the variables that influence these changes is necessary in order to better forecast the melting of Arctic sea ice, and to quantify the conditions of the remaining ice.

The near-continual movement of most sea ice parcels in the Arctic exposes them to different regimes and climate conditions throughout a given year. This creates a need to better quantify how the ice responds to the variable conditions it encounters in order to improve forecasts and to better understand the changing ice conditions in specific regions. Lagrangian tracking of key variables that influence the survivability and conditions of the sea ice will help quantify the effects of these changing regimes on ice conditions. This is particularly of interest in regions such as the Beaufort Sea, which has become a sink for ice in the summer [3–5] and is of economic and political interest [6].
The database presented herein contains weekly sea ice parcel locations combined with ancillary dynamic and thermodynamic datasets spanning from the start of 2001 to the end of 2016. The data provide an avenue to study the changes occurring in the Arctic sea ice through a Lagrangian tracking approach that provides histories of individual 12.5 km sea ice parcels during the years of study. It has been created to provide a comprehensive set of data that will assist in studying sea ice changes in the Arctic by combining sea ice tracks, surface temperatures, concentration, thicknesses, and radiative variables into one searchable database for each year.

2. Data Description

The variables included in this dataset are summarized in Table 1. Ancillary data products include EASE-Grid Sea Ice Motion Vectors [7], MODIS Terra ice surface temperature (IST) [8], SSM/I & SSMIS ice concentration (IC) [9], PIOMAS ice thickness [10], EASE-Grid convergence values in the u & v directions, APP-X albedo, APP-X shortwave up/downwelling radiation, and APP-X longwave up/downwelling radiation [11]. Weekly values of each variable are stored for individually tracked parcels on a 12.5 km EASE-Grid. Data that are available on daily timescales are averaged for each week, and data that are not provided on EASE-Grid are converted to a 12.5 km EASE-Grid. Further background information on the ancillary data products in this database is provided in the Appendix A.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Type</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parcel ID</td>
<td>Int</td>
<td>N/A</td>
</tr>
<tr>
<td>Parcel Age</td>
<td>Int</td>
<td>Weeks</td>
</tr>
<tr>
<td>u Position</td>
<td>Int</td>
<td>12.5 km EASE-Grid Cell</td>
</tr>
<tr>
<td>v Position</td>
<td>Int</td>
<td>12.5 km EASE-Grid Cell</td>
</tr>
<tr>
<td>U Vector</td>
<td>Float</td>
<td>cm/s</td>
</tr>
<tr>
<td>V Vector</td>
<td>Float</td>
<td>cm/s</td>
</tr>
<tr>
<td>IST</td>
<td>Float</td>
<td>Kelvin (K)</td>
</tr>
<tr>
<td>Ice Concentration</td>
<td>Float</td>
<td>Percentage</td>
</tr>
<tr>
<td>Model Ice Thickness</td>
<td>Float</td>
<td>Meters (m)</td>
</tr>
<tr>
<td>Convergence in V Direction</td>
<td>Float</td>
<td>cm/s</td>
</tr>
<tr>
<td>Convergence in U Direction</td>
<td>Float</td>
<td>cm/s</td>
</tr>
<tr>
<td>Albedo, Morning</td>
<td>Float</td>
<td>N/A</td>
</tr>
<tr>
<td>Albedo, Afternoon</td>
<td>Float</td>
<td>N/A</td>
</tr>
<tr>
<td>Shortwave Downwelling Radiation, Morning</td>
<td>Float</td>
<td>W/m²</td>
</tr>
<tr>
<td>Shortwave Upwelling Radiation, Morning</td>
<td>Float</td>
<td>W/m²</td>
</tr>
<tr>
<td>Shortwave Downwelling Radiation, Afternoon</td>
<td>Float</td>
<td>W/m²</td>
</tr>
<tr>
<td>Shortwave Upwelling Radiation, Afternoon</td>
<td>Float</td>
<td>W/m²</td>
</tr>
<tr>
<td>Longwave Downwelling Radiation, Morning</td>
<td>Float</td>
<td>W/m²</td>
</tr>
<tr>
<td>Longwave Downwelling Radiation, Afternoon</td>
<td>Float</td>
<td>W/m²</td>
</tr>
<tr>
<td>Longwave Upwelling Radiation, Morning</td>
<td>Float</td>
<td>W/m²</td>
</tr>
<tr>
<td>Longwave Upwelling Radiation, Afternoon</td>
<td>Float</td>
<td>W/m²</td>
</tr>
</tbody>
</table>

The data are archived through the Pangaea Data Publisher, and are separated by year from the start of 2001 to the end of 2016 [12]. Each year’s data have been separated into individual files containing parcel IDs, grid positions, and variable data to keep file sizes below 100 MB. Further information on how to read the data files and example Python 2.7 codes are available in the product’s Pangaea archive [12]. Fill values of 9999 and 9999.0 have been used for missing data, and a fill value of 9998.0 has been used for the pole-hole present in some products. The files are stored in a tab delimited format, and are organized with a single parcel per line with the following format:

ParcelID, u1, v1, Var1, u2, v2, Var2, ..., u52, v52, Var52 (1)

3. Methods

The parcel database is generated using weekly mean motion vectors available from the National Snow and Ice Data Center (NSIDC) [7]. An initial field is sampled for the first week in January, and
then weekly changes in position of parcels are stored along with available ancillary data for the parcel’s location for the rest of the year through 31 December. Newly frozen parcels are also accounted for, along with parcels that have melted during a weekly time step. The main flow of the program is shown in Figure 1.

![Figure 1](image)

**Figure 1.** Procedural flow for generating a year of parcel tracks with ancillary data.

3.1. Initialization of Field

In order to initialize the ice parcel field, the program begins by searching for sea ice parcels that are of greater than 15% sea ice concentration in week one of the year. Parcels that meet these criteria are assigned a parcel ID and are given space in a Python dictionary object. The initial fields for the ancillary variables are also stored in the initialization phase for week one.

3.2. Weekly Data Acquisition

After the initialization steps are completed, the program loops through every week of the year to save the weekly ancillary data for each parcel. It begins by verifying that data are available, and loading in maps for each ancillary data product for that week. Once these maps are loaded, and any averaging of daily values to weekly values is performed, the program enters a loop that goes through each parcel to save their ancillary data.

3.3. Storing Parcel Data

After the ancillary data are loaded, the program loops through each parcel in the database in order to update each parcel’s position and ancillary data values. The previous week’s sea ice motion vectors are applied to the parcel’s EASE-Grid location, which provides a new location for the current week. Once this new location is saved, the grid coordinates are used to find and store ancillary data for that particular parcel. If data are missing, a fill value is assigned. After the ancillary parcel data are stored, the program also increments the weekly age value of the parcel by one.

3.4. Melting and Formation of Parcels

During the parcel sub-loop in the program, a check is performed for parcels that have dropped below 15% sea ice concentration. At that point, the parcel is considered melted, and is assigned fill
values in lieu of ancillary data values. Another data field in the database also stores the week in which that parcel melted for verification after the product run. In addition to searching for melted parcels, the program checks parcel positions against a landmask and terminates parcels that make landfall.

In order to account for parcel formation later in the year, the program checks for new parcels in areas that are not occupied during the year’s freeze-up period. After the minimum extent week is reached for a particular year, the program begins to search all non-occupied locations for sea ice parcels that contain greater than 15% sea ice concentration. These newly frozen parcels are assigned a space in the parcel database, and are tracked in subsequent weeks. In order to maintain the proper database size and spacing, weeks prior to the formation of the particular parcel are filled with fill values for the ancillary variables and locations.

3.5. Production of Final Database

Once all of the weeks have been sampled, and the database has been composed, the product saves the data in a combined CSV file for the year. Further steps break this file up into smaller files for individual ancillary variables in order to maintain the 100 MB size limit set by Pangaea.

3.6. Example Outputs from Database

An example of tracks and ancillary data from the 2014 data is shown in Figure 2. The major motion patterns of the Arctic are visible from the drifts of the example parcels in the left panel of the figure, while the right panel contains the IST and concentration history for an individual parcel that advected from the Canadian Archipelago region toward the Beaufort Sea. As the melt season began, the IST of this parcel rose to the melting point, which led to some of the parcel melting away and a corresponding decrease in concentration. This database enables quick searches for relationships like these for individual tracks and different variables from the beginning of 2001 to the end of 2016.

![Figure 2](image.png)

*Figure 2.* Selected example parcel drift tracks from 2014 with IST and concentration for the parcel indicated by an asterisk in the left panel.

Another example use of this database to gather bulk regional statistics for 2014 is shown in Figure 3. The left panel shows a user-defined Beaufort Sea region highlighted in red, while the right panel is a scatter plot of the ice surface temperatures of parcels in that region for the duration of 2014. Parcels that melted (red) experienced higher temperatures than their surviving counterparts (blue) prior to the start of the melt season, which may have played a role in their survivability during the summer. Searching regions like the one shown in Figure 3 enables users to examine relationships between different cohorts of parcels in areas of their own interest.
4. Usage Notes

These data are particularly useful for tracking changes in individual parcels or small groups of parcels through time. This Lagrangian tracking technique has the advantage of being able to both determine the effects of initial conditions on specific parcels and the outcome of forcing from the variables tracked in the database. Current studies in our research group use these tracks in airborne and satellite comparison studies, and are also aimed at determining the influence of ancillary variables on the survivability of sea ice parcels during the summer melt season. Additional work has been done to compare the parcel tracks in this database to available buoy data. Users can utilize the parcel tracks and ancillary data for comparison against ship track and buoy measurements that they have available to them.

Example programs that demonstrate how to read the data using Python 2.7 code are provided with the data on Pangaea [12]. These programs are well commented in order to allow the user to translate them to their preferred programming language, and demonstrate how to read the data.

The locations and vectors used in this database are on the 12.5 km EASE-Grid for ease of comparison against existing Arctic datasets. In order to compare against latitude/longitude measurements, users must convert the grid coordinates to latitude/longitude values. The authors utilize the Basemap package for Python 2.7 to do this, but there are numerous packages and tools available that users can implement in most commonly used programming languages. Further information about the EASE-Grid can be found in a related NSIDC publication [13].

It is important to note that parcel ID numbers are persistent between years. This allows the user to track an individual parcel across multiple yearly files by searching for the parcel ID of interest in subsequent years. This enables users to study a parcel for any date range within the database range by using the parcel ID to search multiple yearly files. In addition, parcels that have melted during a particular year do not appear in future files to reduce the size of each file.

The motion of the parcels is tracked by the program that produces it using fractional EASE-Grid cell locations, but the database described herein uses integer positions. Integer positions were chosen in order to allow users to easily compare against other EASE-Grid products that are widely available in the Arctic. As a result, some parcels may share the same integer position while occupying part of the same EASE-Grid cell.

The product is currently available from the start of 2001 to the end of 2016, but further updates are planned pending future support for the project. These new data will be added as a supplementary dataset to the Pangaea archive that hosts the current data. In addition, the code that generates this database can accommodate additional data products, which leaves the possibility of expanding this
product or changing ancillary data sources in the future. Further updates may include additional or new data sources as sources of information or points of comparison.

Acknowledgments: The authors would like to acknowledge support from NASA Grant #NNX16AQ41G for the work in this publication. The authors would also like to acknowledge Chuck Fowler for research that served as a foundation for the work described in this paper. The authors also acknowledge contributions from two anonymous reviewers of this publication, who’s comments greatly improved the quality of this manuscript.

Author Contributions: Mark Tschudi developed previous Lagrangian tracking programs that served as a predecessor to this work, and proposed that this new database be created. Matthew Tooth developed code to re-grid the ancillary data, and developed the program that produces the parcel database with advice and assistance from Mark Tschudi throughout its development. Matthew Tooth provided the additional support codes for the database, and published them to the Pangaea repository. Matthew Tooth and Mark Tschudi both contributed to this publication.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations
The following abbreviations are used in this manuscript:

- APP-X: Extended AVHRR Polar Pathfinder
- EASE: Equal-Area Scalable Earth
- IC: Ice Concentration
- IST: Ice surface temperature
- MODIS: Moderate Resolution Imaging Spectroradiometer
- NSIDC: National Snow and Ice Data Center
- PIOMAS: Pan-Arctic Ice Ocean Modeling and Assimilation System
- SSM/I: Special Sensor Microwave Imager
- SSMIS: Special Sensor Microwave Imager/Sounder

Appendix A. Ancillary Data Products

Appendix A.1. Polar Pathfinder EASE-Grid Sea Ice Motion Vectors

Polar Pathfinder EASE-Grid sea ice motion vectors produced at the University of Colorado Boulder are used in this product, and they serve as the foundation of determining ice parcel positions for each week. The data are hosted at NSIDC, and are further described on their documentation page [7]. The data are provided on a 25 km EASE-Grid, and are incorporated into this database’s 12.5 km grid through the use of a two-dimensional linear interpolation scheme. The product is based on ice motion products derived from satellite measurements, buoy drift tracks, and NCEP/NCAR wind data. The product is generated by merging motion fields from the various sources by using the expected accuracy of each source field as the weight during the merging process. The database described in this paper utilizes the weekly mean motions computed using those merged fields, as the weekly mean data reduce the contribution from noise present in each individual source. The individual motion field sources have RMS errors that range from 1–6 cm/s [14–18], which yield annual displacement errors around 50–100 km in some cases [19]. Further discussion of the error in the merged motion vector product can be found in [19] and the product’s documentation page [7].

At the time of publication, the database utilizes sea ice motion vectors that have been found to contain buoy-affected domain artifacts in the daily motion fields [20]. The impact of these artifacts on this database is mitigated through the use of weekly mean vectors for this product, but a new update to the motion data that utilizes a weighting scheme that eliminates the buoy artifacts is slated to be released in late 2017. The database will be updated to include the new version of the motion data once they are available, and those updated data will be uploaded to the product’s Pangaea archive [12].

Appendix A.2. MODIS Terra IST

ISTs from the MODIS instrument aboard NASA’s Terra satellite are incorporated in this parcel database. The temperatures indicate when individual ice parcels are at melting temperature,
and are provided to further the analysis of changes in other variables that are tracked. The data are hosted at NSIDC, and are further described on their documentation page [8]. The data are provided on a 4 km EASE-Grid, and are re-gridded to fit the 12.5 km EASE-Grid used in this product through the use of a two-dimensional linear interpolation scheme. The product has been compared to other sources and in-situ measurements, which have yielded an estimated accuracy of 1–3 Kelvin under ideal conditions [21]. The accuracy of the product potentially degrades in the presence of clouds and water vapor, which lends toward lower accuracy during the summer months due to more cloud cover. Further discussion of the product and its accuracy can be found in [22,23]. Valid ranges for this data product are 210 K to 313.2 K, pole-hole points are masked to 9999.8, and all other values are masked to 9999.0. It is important to note that, while other products may be available outside of the 2001 to 2016 date range, these data form a complete year starting in 2001. This limits the current version of the parcel database to years including and after 2001 due to its use of these data.

Appendix A.3. SSM/I and SSMIS Ice Concentration

SSM/I & SSMIS sea ice concentration values [9] are utilized by the product to determine when parcels are considered melted, and to search for new parcels later in the year. Changes in concentration are additionally useful for tracking the degradation of ice health during the melt season. The data are provided on a 25 km Polar-Stereographic grid, and are re-gridded to fit a 25 km EASE-Grid before use in this product. Each 25 km EASE-Grid cell is then quartered into 12.5 km EASE-Grid cells with equal values in order to fit the 12.5 km EASE-Grid. The concentration data are less accurate in the presence of thin sea ice, melt ponds, and near the ice edge where some ocean and atmospheric effects can be mistaken for sea ice. The seasonality of these error producing conditions lends to a ±5% accuracy during the winter and a ±15% accuracy during the summer [9]. Further discussion of the error and performance of the product can be found in [24,25]. Concentrations in this combined database range from 15% to 100%, with other values masked to 9999.0 and the pole-hole masked to 9998.0.

Appendix A.4. PIOMAS Ice Thickness

The PIOMAS ice thickness model data are incorporated into the product to provide volume loss estimates, and to enable tracking of changes in ice health. The data are provided in a curvilinear coordinate system that is re-gridded to the 12.5 km EASE-Grid prior to use in this database. The product is validated using submarine, mooring, and satellite observations in order to compare its model output to available data sources. An estimated RMS difference between the model output and independent submarine tracks is stated as 0.78 m. Further information about the product and error measurements can be found in [10,26].

Appendix A.5. EASE-Grid Convergence Values

Convergence values for the u and v EASE-Grid axes are also included in this database, and are produced using the weekly motion vectors that form the basis of this product. In order to compute a convergence value in each direction, an individual cell’s motion vector components are compared to the motion vector components of directly adjacent cells. The boundary convergence calculations are performed using the Equations (A1) and (A2) for the v direction, and Equations (A3) and (A4) for the u direction:

\[ C_{\text{top}} = V - V_{\text{above}} \]  

\[ C_{\text{bottom}} = V_{\text{below}} - V \]  

\[ C_{\text{right}} = U - U_{\text{right}} \]  

\[ C_{\text{left}} = U_{\text{left}} - U \]  

(A1)  

(A2)  

(A3)  

(A4)
where $C_{\text{top}} \& C_{\text{bottom}}$ represent the convergence between the parcel of interest and the parcels above and below it, and $C_{\text{left}} \& C_{\text{right}}$ represent the convergence between the parcel of interest and the parcels to the left and right of it. The $U$ and $V$ terms represent the vertical and horizontal velocity components of each parcel respectively.

The location of each of these equation elements, along with a representation of the parcels used in the calculation, are shown in Figure A1. In this grid system $V$ vectors represent the vertical component and are positive-upward, while $U$ vectors represent the horizontal component and are positive-right. It is important to note that the $u$ and $v$ coordinates are 0,0 at the top left corner of the grid, which lends to the $V$ vectors being inverted with respect to the grid coordinates. Total convergence values for the cell in the $u$ and $v$ directions are then derived by summing the obtained boundary values on each axis:

\[ C_v = C_{\text{top}} + C_{\text{bottom}} = V_{\text{below}} - V_{\text{above}} \]  \hspace{1cm} (A5)

\[ C_u = C_{\text{left}} + C_{\text{bottom}} = U_{\text{left}} - U_{\text{right}} \]  \hspace{1cm} (A6)

The resulting numbers represent the speed at which a parcel is approaching its neighboring parcels, where larger positive values of $C$ represent more convergent motion. This number serves as an indicator of greater potential for ice collision and deformation due to this convergence. Negative values of $C$ indicate that the neighboring parcels are receding, and that other processes such as lead formation may be more likely for that parcel. These data are provided to aid researchers in determining where relationships like these may occur.

**Figure A1.** EASE-Grid cell with vector components used to compute convergence values shown.

**Appendix A.6. APP-X Atmospheric and Surface Data**

Albedo and longwave/shortwave radiation data are incorporated into the data product in order to provide context of energy input into the sea ice parcels. The data are obtained from the Extended AVHRR Polar Pathfinder dataset, which is mapped to a 25 km EASE-Grid at local solar times of 0400 h and 1400 h in the Arctic. In order to use these data in the database, each 25 km cell is split into four equal 12.5 km EASE-Grid cells and incorporated into the product. The product’s reliance on AVHRR lends to decreased accuracy during the summer months due to the presence of more atmospheric
water vapor and clouds. Further discussion of these products can be found in their NOAA CDR documentation [11] and related publications [27–29].

References


© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).