PhD Seminar - Arctic Sea Ice Feature Detection

Alison Kleffner

11/19/21

Seminar Agenda

- Introduction to Problem and Data Set
- Previous Methods
- Our proposed Method: Bounding Box Approach
- Interpolation of Missing Data
- Future Work

Project 1: Arctic Sea Ice Feature Detection

Alison Kleffner

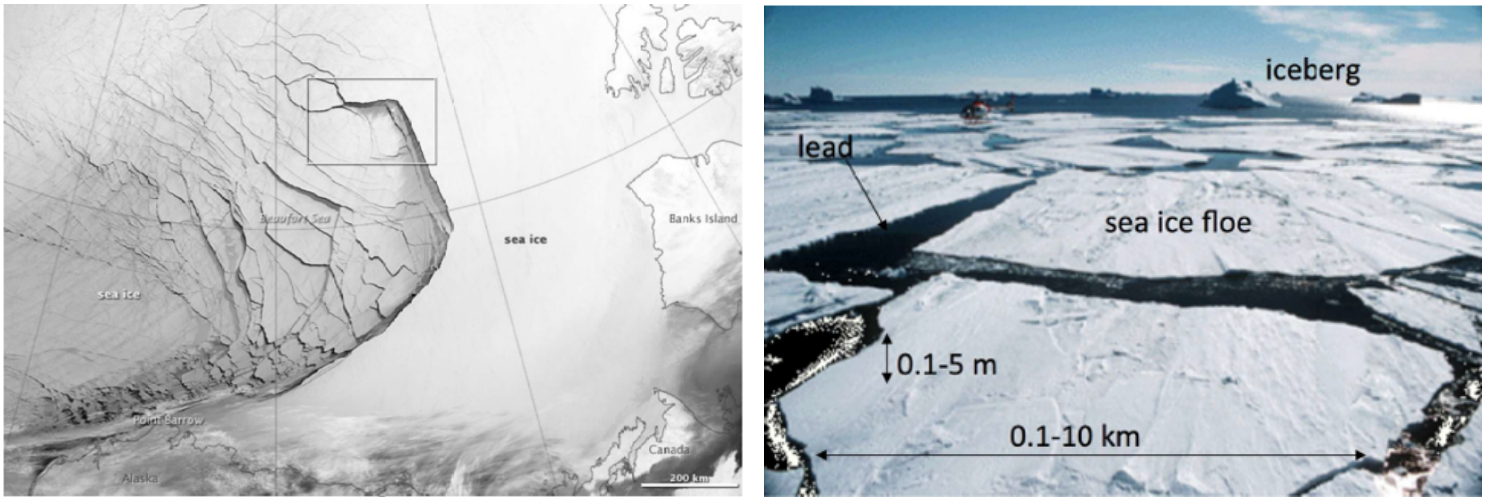


Figure 1: Figures of Ice Cracks

Motivation

- What are We trying to Do?
  - Develop a method to determine where possible Ice Cracks may form given only movement data
- Data Given
  - Gpid: Identify of part of ice chunk
  - Location of gpids (x/y)
  - Observation Time: Have 22 days worth of data
  - k: image index (sometimes will have multiple observations for a gpid on a day)

Sea Ice Motion Animation

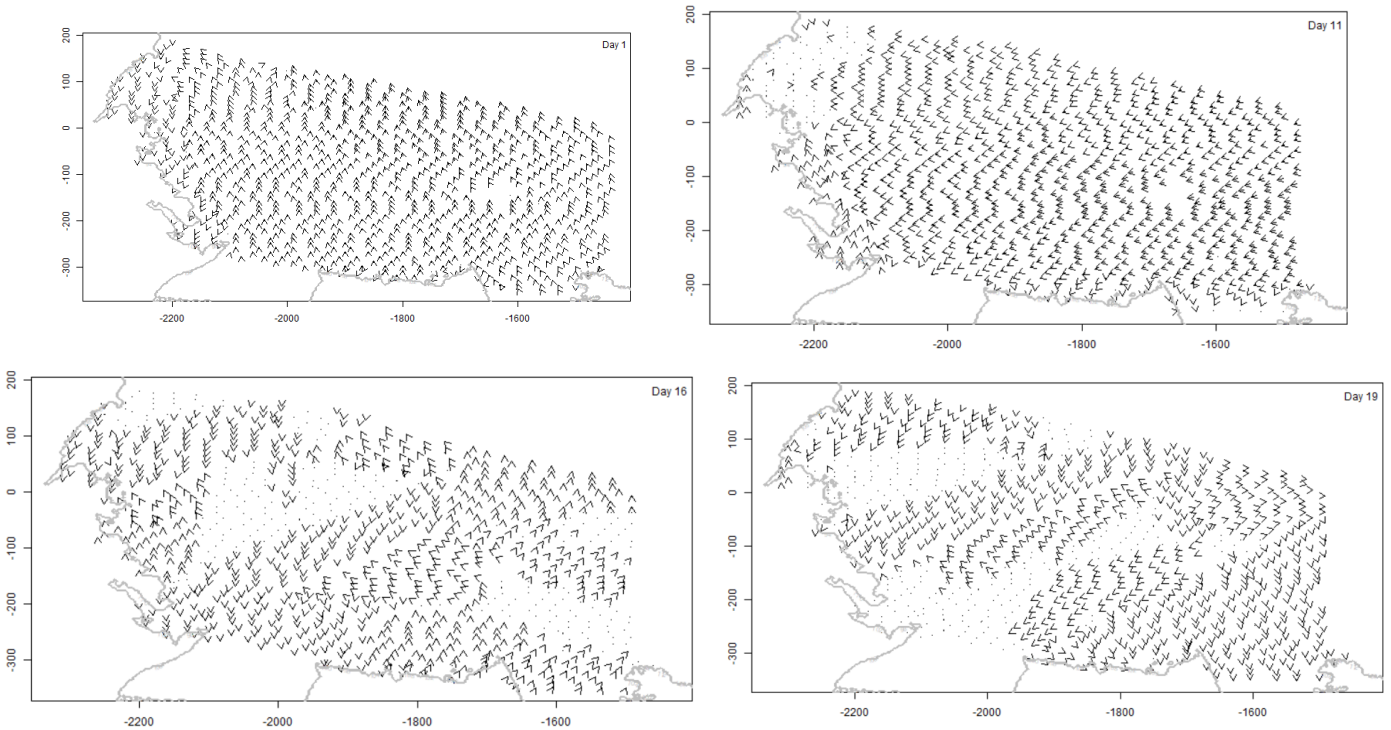
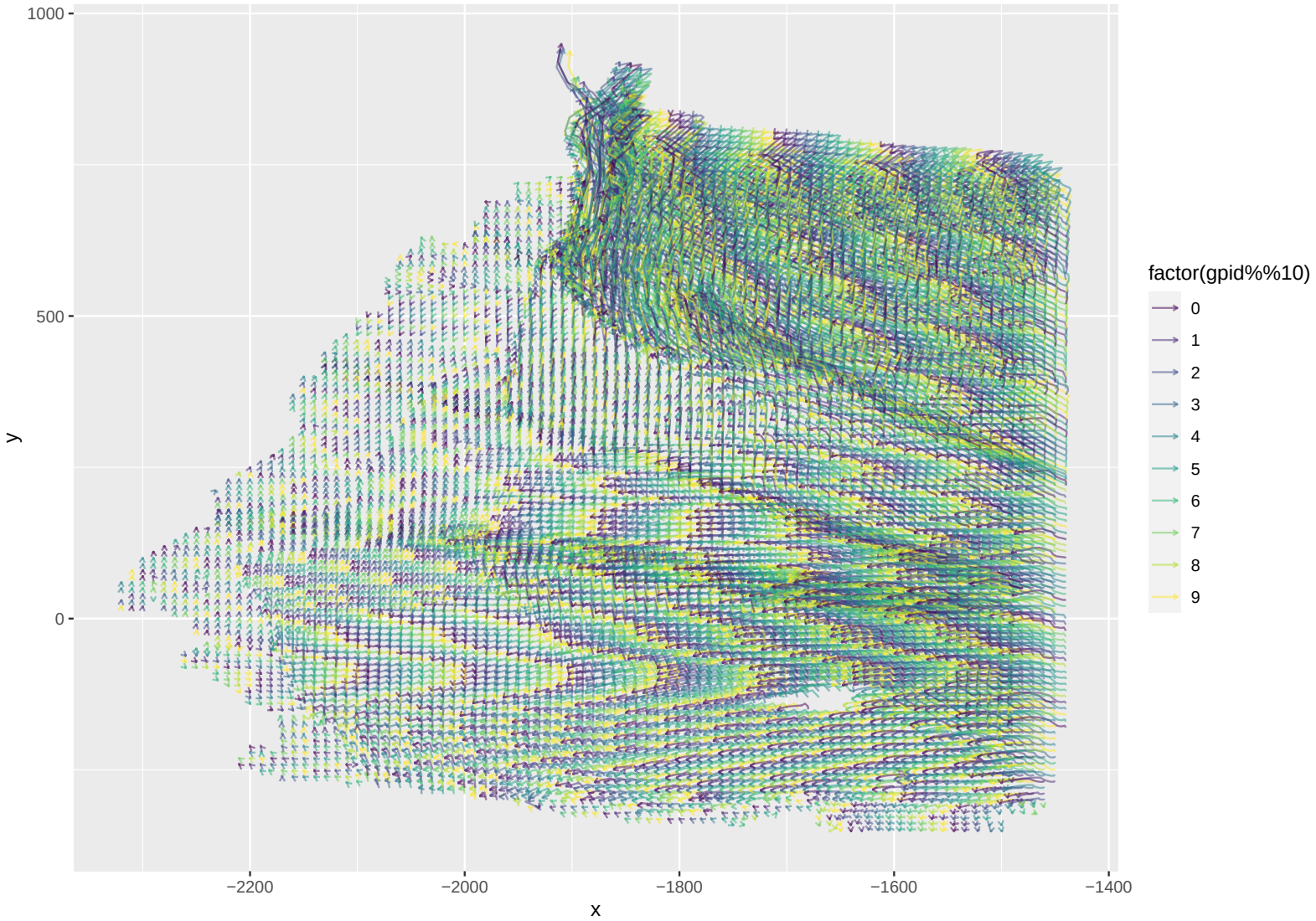


Figure 2: Ice Motion

Explanation of Problem

PhD Seminar - Arctic Sea Ice Feature Detection

Alison Kleffner



Comparison: Previous Work by Guan et al (2019)

- There is another dataset that has more information (like derivations of estimates of ice deformation)
- Then they ran a kinematic analysis of the deformations.
  - Fit a jump in displacement that would account for the observed deformation in a cell.
  - So gives some indication of where cracks may form, and the level of opening of the crack.

Overview: Spatio-Temporal Clustering

- Ansari et. al (2019)
  - Event Clustering
  - Geo-Referenced data item clustering
  - Geo-Referenced time series clustering
  - Trajectory Clustering (Focus)
  - Moving Clusters
  - Semantic Based Trajectory Mining
- Clustering of sub-trajectories (Lee et al (2007))

Challenges

- How gpids are laid out (can't use density-based clustering)
- Missing chunks of data (issues with calculations of distances)
- Only motion data is observed
- Typical interpolation methods aren't suitable
  - Non-smooth spatial process
  - Nonstationarity due to ice moving as patches.

Our Proposed Method

- Cluster similar trajectories to identify patches of ice using information from a Bounding Box
  - A way to work around the missing data problem
- Space-time interpolation within each ice pack where ice movements are similar.

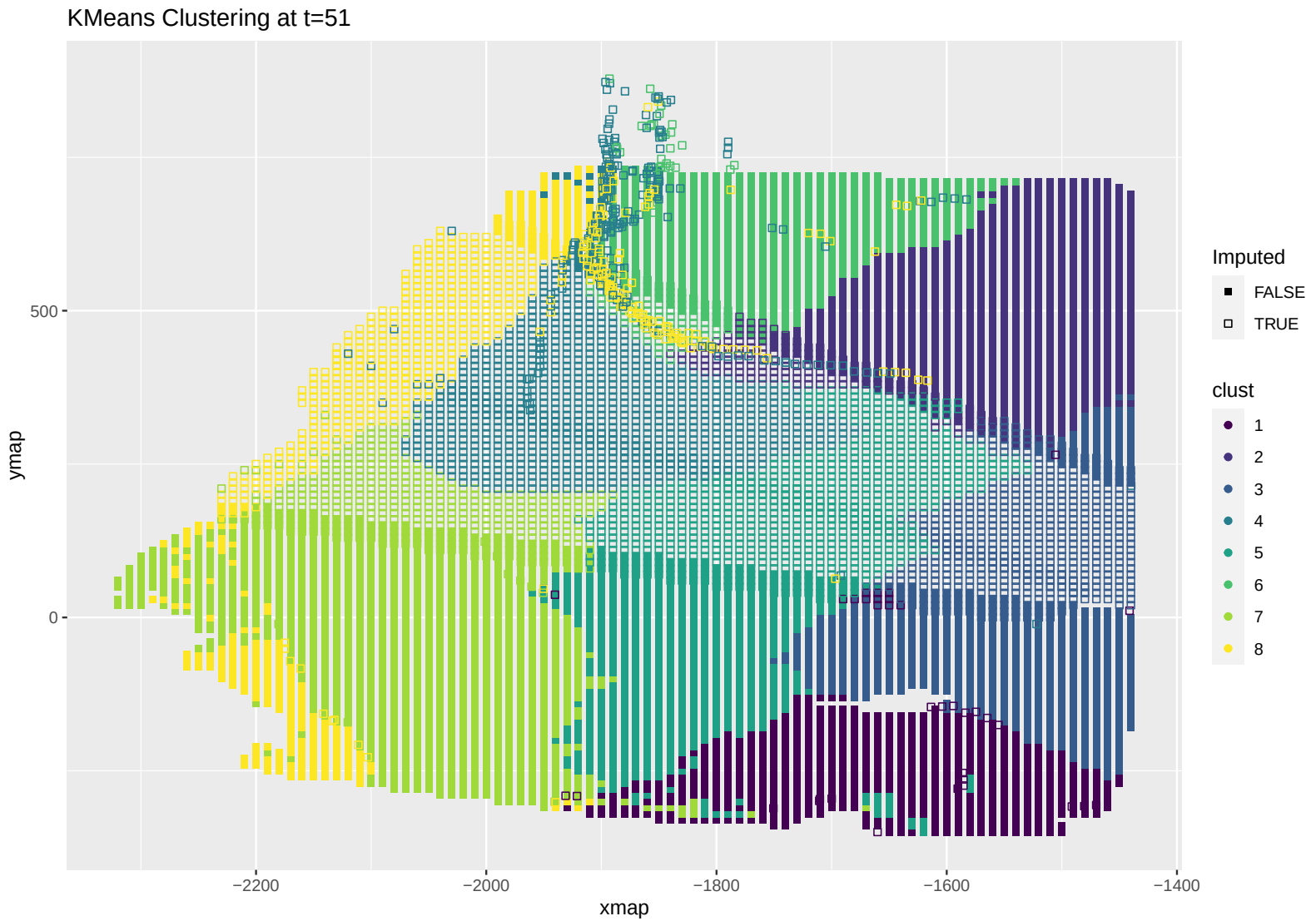
Clustering with Bounding Box

- Included in Bounding Box
  - Min/Max Latitude
  - Min/Max Longitude
  - Average Lat/Long
  - Length of Latitude
  - Length of Longitude
  - Angle/Direction Moved
- Use the features of the bounding box as inputs into KMeans Clustering
  - The boundaries of each cluster would be where the ice crack forms
  - The number of clusters was determined using the silhouette statistic

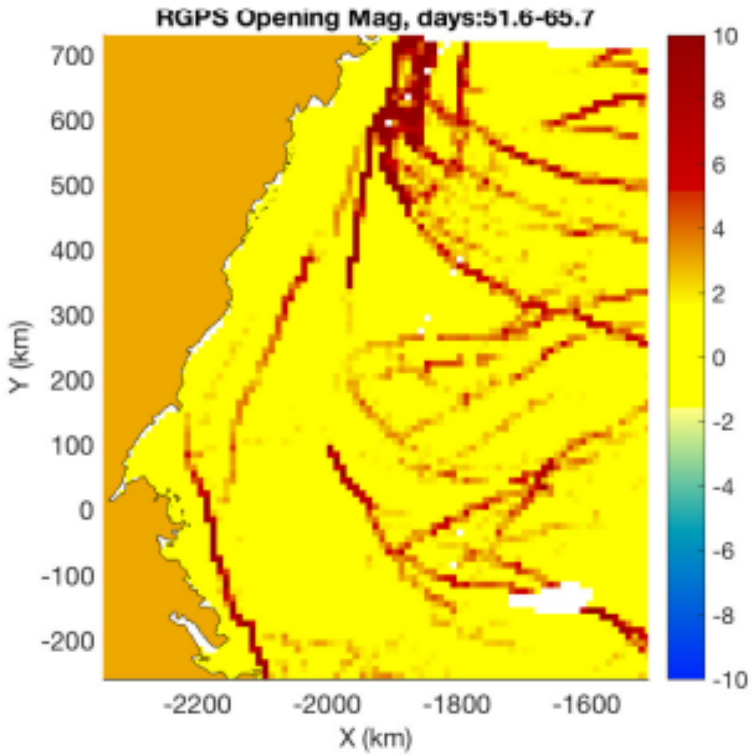
Results: Bounding Box of All Days

PhD Seminar - Arctic Sea Ice Feature Detection

Alison Kleffner



Comparison to Yawen's Previous Work



(b)

Figure 3: RGPS opening magnitude from the kinematic algorithm

Results: Bounding Box By Week

PhD Seminar - Arctic Sea Ice Feature Detection

Alison Kleffner

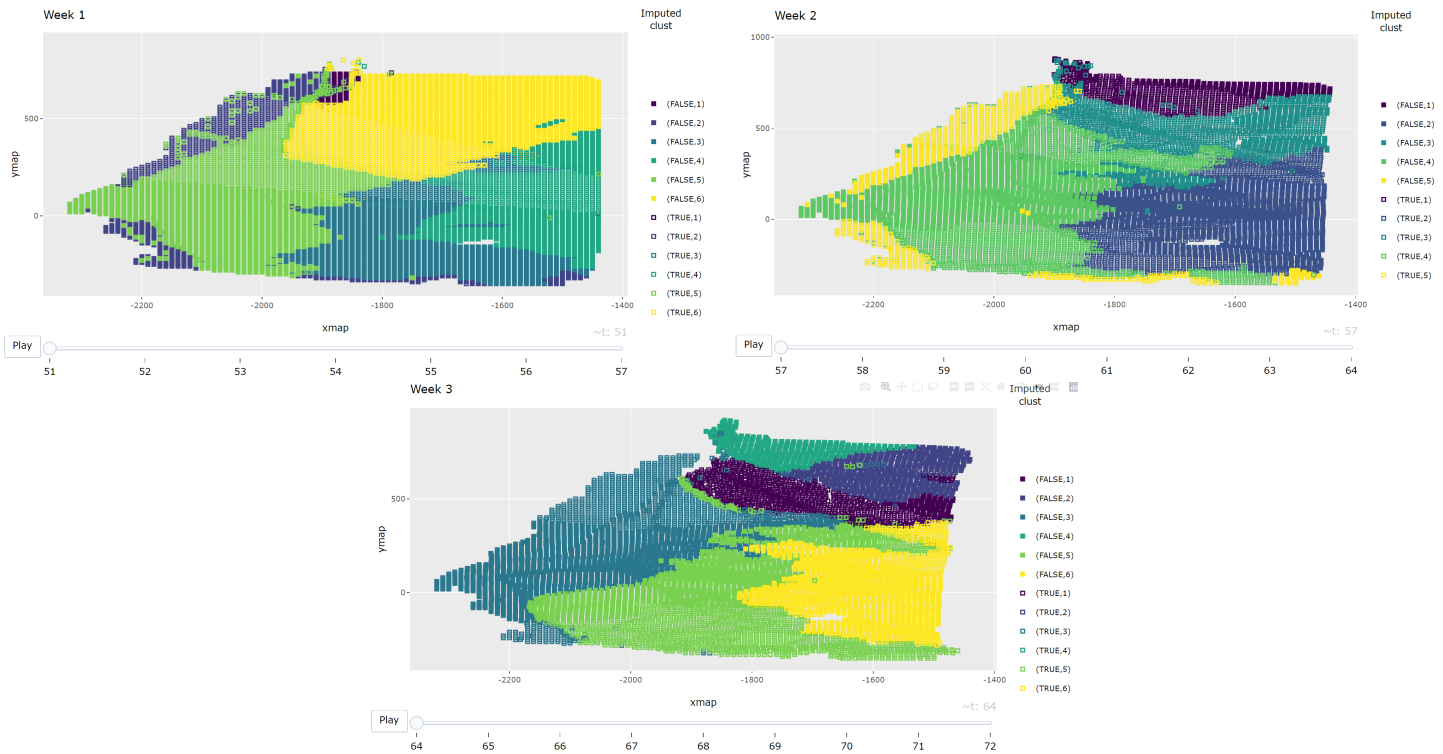


Figure 4: Clustering Bounding Boxes by Week

Next: Interpolation of Missing Information

- Want to be able to interpolate the missing x/y gpid information
  - Challenges:
    - When missing gpid information, missing it in chunks
    - For spatial- temporal interpolation, in order to calculate the distance matrix, need latitude and longitude.
- Our Method: Use of Polygon Intersections
  - Find Spatial and temporal neighbors and use these to interpolate onto a grid

Interpolation Process

- Find Spatial-Temporal Neighbor groupings for each week.
  - Created Polygons for each week of the Clusters given previous (spatial neighbors)
  - Find intersection of polygons for the different weeks (temporal neighbors)
- Develop a grid for starting values if missing.
- At a time point, find the known data, and use this to develop a model using fit_model in the GpGp package
  - Exponential Space-Time Covariance Function
- Then predict the gpids x or y location using the developed model with the initial value being the grid cell.
  - Current Issues

Interpolation Pics

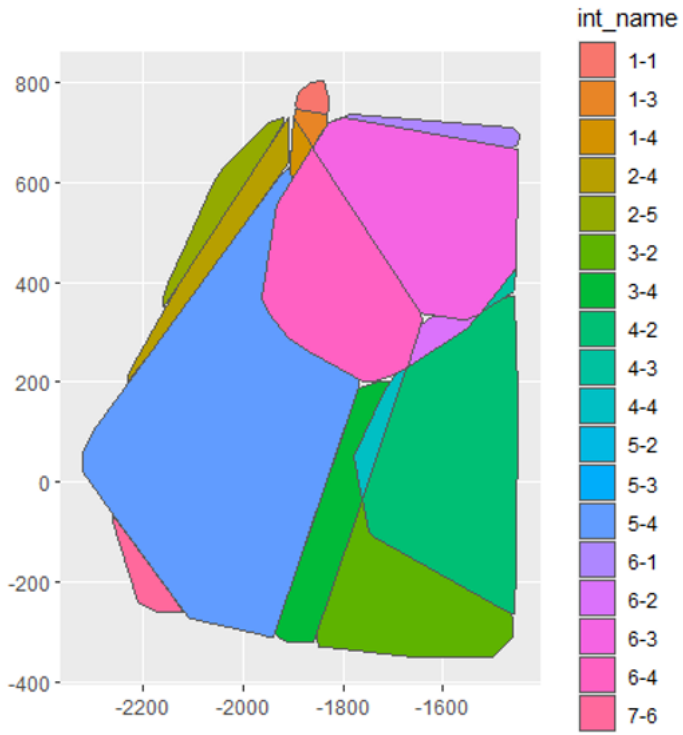


Figure 5: Spatial-Temporal Neighbors of Week 1

Current Work: Analyzing Nonstationary Spatial Data Using Gaussian Processes

- Find a method that can determine groupings, and also model our data in one step.
  - Methods on analyzing nonstationary spatial data using piecewise Gaussian processes
    - Voronoi Tesselation (Kim et al. (2005))
    - Bayesian Tree (Konomi et al. (2014))
- Problems so far:
  - Methods don't have a time component
  - Developing the Code.

Future Work

- Figure out errors in interpolation model
- Validation of my interpolation method
  - See how it does holding out known data
  - Comparison to Linear Interpolation
- Keep exploring the modeling of nonstationary data using Gaussian Processes.
- Create a pipeline so can become more automated (for example, if have more days)

Selected References

- Ansari, M.Y., Ahmad, A., Khan, S.S. et al. Spatiotemporal clustering: a review. Artif Intell Rev 53, 2381–2423 (2020). https://doi.org/10.1007/s10462-019-09736-1
- Bledar A. Konomi, Huiyan Sang & Bani K. Mallick (2014) Adaptive Bayesian Nonstationary Modeling for Large Spatial Datasets Using Covariance Approximations, Journal of Computational and Graphical Statistics, 23:3, 802-829, DOI: 10.1080/10618600.2013.812872
- Guan, Y., Sampson, C., Tucker, J.D. et al. Computer Model Calibration Based on Image Warping Metrics: An Application for Sea Ice Deformation. JABES 24, 444–463 (2019). https://doi.org/10.1007/s13253-019-00353-7
- Kim, H., B. Mallick, and C. Holmes (2005). Analyzing Nonstationary Spatial Data Using Piecewise Gaussian Processes. Journal of the American Statistical Association, 100(470), 653–668. http://www.jstor.org/stable/27590585

Selected References

- Lee, J. G., Han, J., & Whang, K. Y. (2007). Trajectory clustering: A partition-and-group framework.In SIGMOD 2007: Proceedings of the ACM SIGMOD International Conference on Management of Data (pp. 593-604). https://doi.org/10.1145/1247480.124754