# Modeling Sea Ice Thickness using Machine Learning and Remote Sensing Modalities

John Baker

Ira A. Fulton Schools of Engineering

School of Computing and Augmented Intelligence

Approved:

Dr. Douglas Cochran ______________________________, Director

Dr. Hua Wei ______________________________, Second Reader

Accepted:

______________________________

Dean, Barrett, the Honors College

Arizona State University

# Modeling Sea Ice Thickness using Machine Learning and Remote Sensing Modalities

Honors Thesis

Barrett, The Honors College

*Author:*

John Baker

*Committee:*

Dr. Douglas Cochran

Dr. Hua Wei

April 2024

# Acknowledgement

I hereby declare that I can carry out the present work independently without outside help. and used only the sources and aids indicated. I assure Furthermore, that I have not yet submitted this thesis to any other examination board.

# Abstract

Little is known about the state of Arctic sea ice at any given instance in time. The harshness of the Arctic naturally limits the amount of in-situ data that can be collected, resulting in gathered data being limited in both location and time. Remote sensing modalities such as Synthetic Aperture Radar (SAR) imaging and laser altimetry help compensate for the lack of data, but suffer from uncertainty because of its inherent indirectness. Furthermore, precise remote sensing modalities tend to be severely limited in spatial and temporal availability, while broad methods are more accessible at the expense of precision. This thesis focuses on the intersection of these two problems and explores the possibility of corroborating remote sensing methods to create a precise, accessible source of data that can be used to examine sea ice at a local scale.

//////////////////////////

English Abstract here. The abstract should provide a complete but concise description of your work. In brief, you should address the following: motivation, problem statement, approach, results, and conclusions.

# Contents

# Introduction

Why is sea ice important, and what are the applications of understanding it at a closer scale?

Background of how sea-ice currently measured: In-situ, LiDar, SAR. Discuss the advantages that SAR and LiDar have over In-situ measurements (e.g. availability), but how they are innately less confident in their accuracies.

Discuss the prevalent sources of data as it stands - NSIDC as a data repository, NASA's Sentinel(1/2), IceSat-2, CryoSat satellites as periodic sources of data. Transition into how despite how available these different mediums are, they are subject to inaccuracies from assumptions. In particular, without a direct auger core to measure ice-thickness it's common practice to infer thickness by the assumption of iso-static equilibrium (e.g. Assuming the ice body is floating, then deducing the mass and densities of the substances involved to achieve that state). This approach, while intuitive, is dependent on unknown mass, and variable density. Densities vary between the snow, sea water, and sea ice depending on factors like salinity, temperature, and age, all of which are precisely unknown at local scales both temporally and geographically.

Attempts have been made to improve this estimation [citing the IceSat-2 : CryoSat study, the In-Situ analysis], but the efforts have been consistently inconclusive at best. Ball-park estimations for sea-ice density currently range from foo to bar but is disputed to practically be even lower.

Transition into how SAR imaging can capture a return that indicates sea ice thickness, but it's low resolution from Sentinel-2 prevents it from being local in context. The intuition behind the remainder of the thesis is to coincide precise IceSat-2 LiDAR and data derivatives with widespread SAR imaging, to determine whether sea-ice thickness can be a learnable property from SAR image returns.

The problems with this include the availability of applicable data (such as high-resolution SAR, as the precision of IceSat-2 is immediately lost with low-resolution imagery) and also the coincidence of these two modalities, both spatially and temporally.

## 1.1 Preliminary Research

Preliminary research into the availability of ice thickness data yields three notable organizations which may act as sources of data. These three organizations are the National Snow and Ice Data Center (NSIDC), the National Aeronautics and Space Administration (NASA), and the European Space Agency (ESA). Although all organizations offer large amounts of data, they differ in approach. The NSIDC acts as a historic repository, accumulating previous studies into a single location for easy access. NASA offers a similar service, but additionally allows researchers the opportunity to query information from any of their active orbiting satellites. The ESA, like NASA, offers near real-time satellite data but also offers to sponsor data from external companies with the completion of a project proposal.

### NSIDC

Here is where the NSIDC links and data retrieval should be discussed. Limitations of the data or other notes about it's composition should be included according to the details in the abstract saved on file. Preliminary exploration will be where the figures will be uploaded (Or maybe we show one figure showing all of the points, and the exploration will be for the single instance?)

### ESA

Here is where you discuss the ESA's copernicus portal, finding the ICEYE company, looking through the proposal process, seeing superior SAR resolution from this private company than NASA's sentinel-2

### NASA

Here is where you should include sources to the ATL10 Data Specification pdf, and the associated CryoSat data specs. It should also be included where these links came from. It also should be included where, if applicable, specific data downloads were from. It should include the process of querying for data, which portals are made accessible, and how flexible they are. Do not forget to mention that they made an IceSat-2 mobile application for iOS.

#### ICESat-2-ATL10-Product

Include how NASA's Sentinel-2 delivers SAR imaging, but at 50m resolution. This is a modality that can be used for a CNN but the resolution is not meaningfully compatible with any source of ground truth we have. Thus, searching for alternatives led to the European Space Agency and its sponsorship for ICEYE's 1m resolution imaging.

Somewhere you should discuss IceSat-2 and its laser altimetry. ESPECIALLY that it's 17m footprint offers precision at a given location,

## 1.2 Preliminary Exploration

After researching the available data sources across these 3 agencies, the NSIDC's In-Situ dataset offered the greatest possibility of learning direct trends of sea ice. However, NASA's IceSat-2's predictable orbit combined with the possibility of obtaining ICEYE's high resolution SAR imaging, suggested that a temporally and spatially coincident dataset could be obtained. This intersection of remote sensed imagery with precise remotely sensed freeboard measurements, is conducive for a convolutional neural network applying IceSat-2's accuracy onto the expansiveness of SAR imaging. This model would be more dynamic in nature than the static accumulation of In-situ data, and would give more insight into the changing state of sea-ice. [1]

Importantly, it's important to recognize that laser altimetry is purely restricted to measuring Earth's surface - meaning the returned measurements are innately incapable of measuring sea-ice thickness, especially in the presence of snow on the surface. (Add source and put it in bibliography)

NASA's ICESat-2 L4 Along-Track Sea Ice Thickness highlights this problem, and addresses it using the assumption of hydro-static equilibrium. This equation, expressed as follows, uses the densities of water, sea-ice and snow, alongside the height of the snow to calculate the correlated thickness of the sea ice.

$$\rho_i \rho_w \rho_s$$

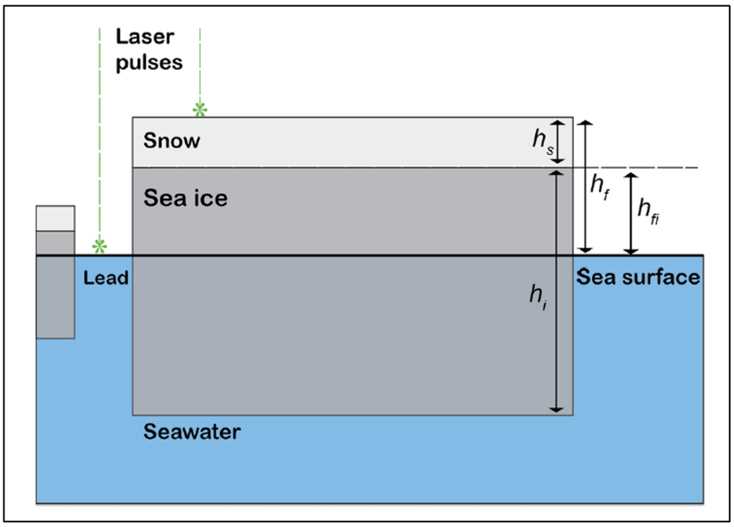


Figure 1.1: Laser Altimetry on a buoyant surface
[2]

## NSIDC

The "On-Ice Arctic Sea Ice Thickness Measurements by Auger, Core, and Electromagnetic Induction, from the Late 1800s Onward, Version 2" contains 69,750 rows of data, spanning 5 categorized regions; the Arctic Ocean, the Beaufort Sea, Greenland Coast, Prudhoe Bay, and Russian Coast. Filtering down to the Beaufort Sea region, a region of interest, yields 23 separate studies ranging between 1958 and 2016.

For a simple analysis, the largest single dataset was chosen to conduct a simple test on the distribution of sea ice thickness. This distribution if normal will reasonably allow for future conducting of statistical tests, and if not normal will give insight as to what a reasonable assumption of ice thickness distribution should be. The largest dataset consists of 21,547 distinct points sourced from direct auger measurements and indirect electromagnetically sensed methods. Given that the majority of the data comes from the remote

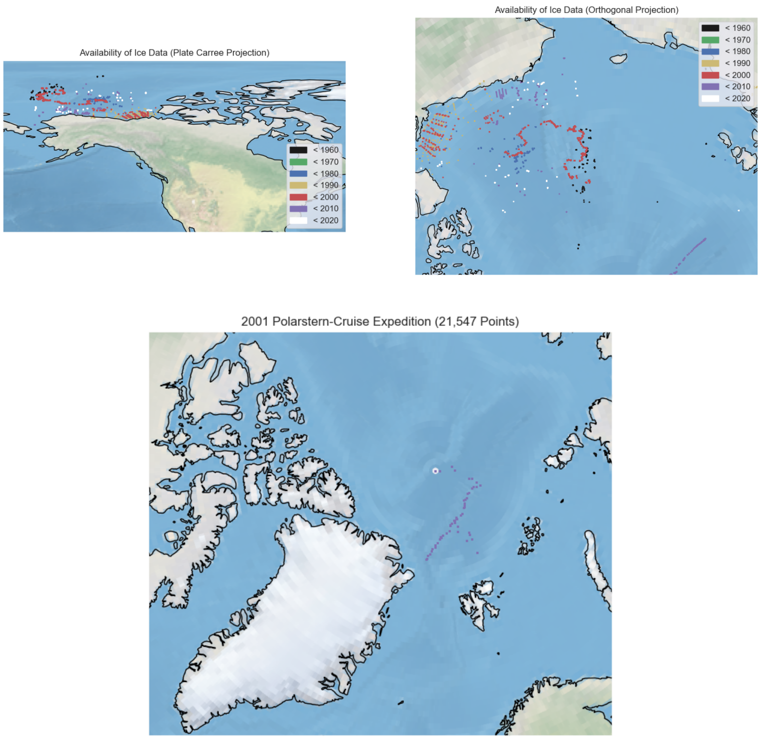


Figure 1.2: In-Situ Data Availability

sensed method, it was validated by plotting the sensed values against the auger values where applicable and seeing how correlated they were. The results of a linear regression test yielded a relationship of 0.92x + 0.18, with an adjusted $R\hat{2}$ value of 0.866 and a P Value of 2.31 * $10\hat{-}202$. The adjusted $R\hat{2}$ value and low P value demonstrate there is a high correlation between these variables, and our model captures a significant portion of the variance. The 0 line is plotted as a way to visualize how the residuals are symmetrically clustered around 0.

After accepting the validity of the electromagnetically sensed data, the points were plotted to visualize the strip of ice that was measured. Given that these data points were taken in a single linear track, the line graph shows a cross-sectional profile of ice in that axis. The distribution reveals a

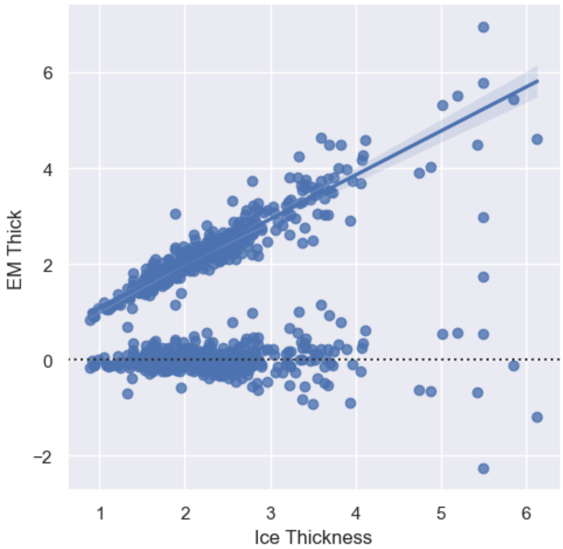


Figure 1.3: Electromagnetic vs Auger Sourced Relationship

slight tail on the left side of the curve, and a hypothesis test confirms that the distribution is not normal. Moving forward, it is reasonable to believe that ice distribution is not normal, and over a large enough sample there is expected to be a slight tail on either end.

### NASA

A single delivery of IceSat-2's ATL10 data product yields a '.h5' file, incompatible with traditional spreadsheets. To access IceSat-2 data in a meaningful way involved developing a script to extract relevant information according to the types enumerated in the data product specification. Columns of interest include the latitude, longitude, time, and calculated freeboard height for each of the three beam pairs. // Consider adding perl and dynamically adding sample rows from the extracted .h5 file. This will give the reader context as to what's being gathered /

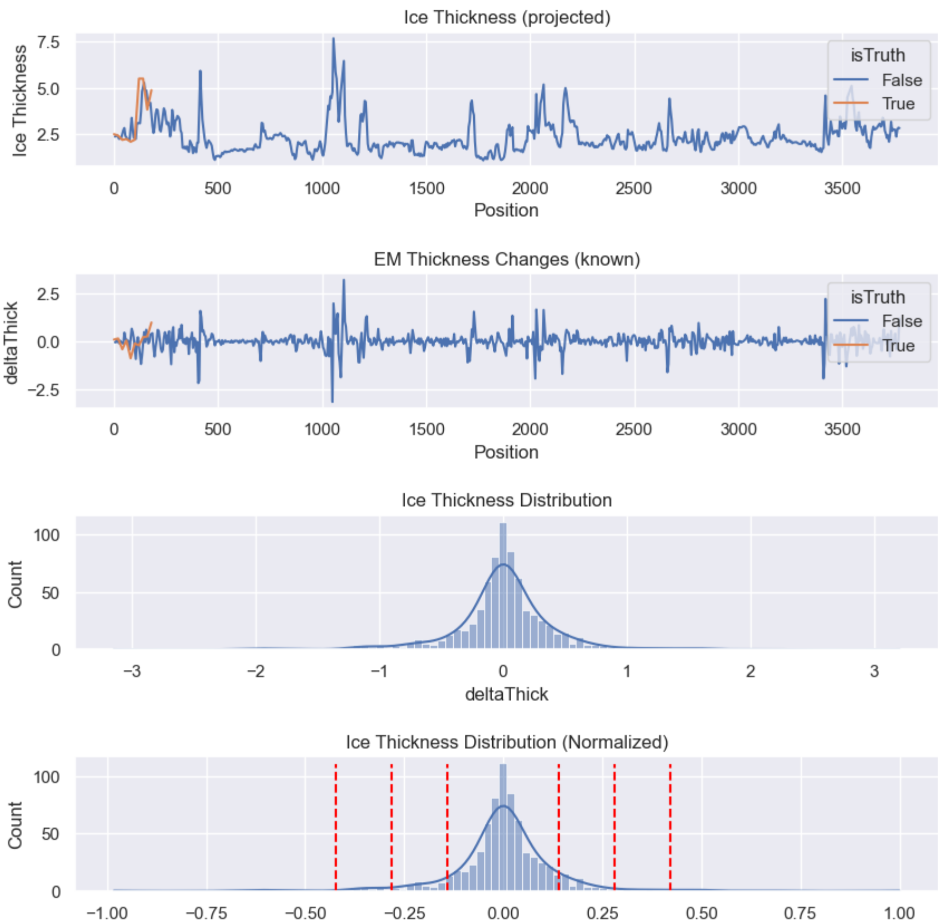


Figure 1.4: In-Situ Ice Thickness Distribution

## ESA

Filling out the proposal for ICEYE sponsored data. (eg: See Appendix A or something)

# Data Collection

In this section, include research relevant to the gathering of data. Include the different metrics that may be captured and the different options presented from different avenues. Much research was done earlier based on the opportunities and limitations of NASA satellites. Include documentation you captured about CryoSat if pertinent to this section.

It may be necessary to include information on SAR imaging in the introduction, so even the non-expert reader can get a functional understanding of the topic.

It may be possible to include figures and other data from the NSIDC In-Situ data repository and discuss the viability of that data as a 'ground-truth' and demonstrate that according to those studies, sea ice was non-normal in distribution. Acknowledge that the data from those studies spanned many miles, and the distribution may sheerly be by variance of ice thickness across the collection area.

Here is where we'll discuss the selection of ICEYE's Data for it's high resolution imaging, and IceSat-2 for it's high-accuracy laser altimetry. It may be possible to include figures to demonstrate how these satellites work, but it may not be pertinent to the thesis (although it would help with understanding).

## 2.1 Laser Altimetry (IceSat-2)

## 2.2 SAR Imaging (ICEYE)

——Default Text ———-

Provide in brief the background information for your work/field keeping in mind that maybe your readers do not have experience with topics your reference or address in your thesis.

In the second part, provide a review of the state of the art relevant to your thesis. Here you present relevant research that relates to your work.

# Experimentation

Here you should present how you aimed to answering your question or solving the problem you identified in the introduction. You should present the methods and the instrument you use, the structure of your study or workplan and how you achieved the desirable outcome. The order is indicative, please feel free to rearrange as you see fit.

## 3.1 Methods

[3] [4]

## 3.2 Study setup

## 3.3 Data collection

## 3.4 Data analysis

# Results

Here you present the results of your study (if you carried out one), and any data analysis you may have performed to answer your question. You should consider splitting the results per research question, as these are presented in the introduction

# Discussion

In the Discussion section you should elaborate on the following points:

## 5.1 Research Questions

Here you will answer your research questions, as they appear in the introduction. Answer each question in a different section. Relate your answer to your results. Discuss if your findings support and align with related work or not. Explain why do you think this happens, especially if your findings contradict existing work. Discuss alternative interpretations of your findings.

## 5.2 Theoretical and Practical Implications

Here you should explain what is your contribution and how it promotes knowledge in the field both in terms of theory and practice. How do you envision your work to change or promote research in the area. How could it be used? How do you envision it to be used?

# Conclusion

The result of this thesis is a rough pipeline that links elusive data to a intuitive method of better modeling sea ice. IceSat-2's semi-frequent orbit provides data that can, with planning, be corroborated with other data sources to bridge the gap between different remote sensing methods. The results of the experimentation do not suggest the model is capable of deducing sea-ice thickness from mere SAR imaging.

## 6.1 Limitations

While intuitive, the fundamental approach to this problem lies on a set of assumptions that may not always be true. At the root of the deduction of sea-ice thickness using IceSat-2 freeboard measurement is the assumption of hydro-static equilibrium at the measured footprint. The equation relies on the densities of snow, ice, and sea water, which are constants that may vary seasonally and geographically. Furthermore, each footprint in this equation is assuming the absence of any other forces acting on the body. This thesis neither considered nor explored the dynamics of sea-ice floes, meaning that the calculated thickness at any given location may differ not only from error, but from a faulty equation derived from an incomplete physical understanding of the observed body.

## 6.2 Future Work

In the future, this topic can be further explored by examining different machine learning models and architectures that may be better suited on 2-channel, low-resolution imaging. To aid this, more data should be collected from both IceSat-2 and ICEYE.

# List of Figures

# Bibliography

[1] B. Holt, *On-ice arctic sea ice thickness measurements by auger, core, and electromagnetic induction, from the late 1800s onward, version 2*, 2019. DOI: `10.7265/wz0k-4p60`. [Online]. Available: `https://nsidc.org/data/G10011/versions/2`.

[2] N. K. R. K. T. M. Petty A. A. and T. A. Neumann., *Icesat-2 l4 along-track sea ice thickness, version 1*, 2022. DOI: `10.5067/JTI5YG3S6VAJ`. [Online]. Available: `https://nsidc.org/data/IS2SITDAT4/versions/1`.

[3] S. S. Basha, S. R. Dubey, V. Pulabaigari, and S. Mukherjee, "Impact of fully connected layers on performance of convolutional neural networks for image classification", *Neurocomputing*, vol. 378, pp. 112–119, 2020, ISSN: 0925-2312. DOI: `https://doi.org/10.1016/j.neucom.2019.10.008`. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S0925231219313803`.

[4] I. L. Jernelv, D. R. Hjelme, Y. Matsuura, and A. Aksnes, "Convolutional neural networks for classification and regression analysis of one-dimensional spectral data", *arXiv preprint arXiv:2005.07530*, 2020.

# Insert Appendix Name

## A.1 Insert subtitle