

CryoSat-2 waveform classification for melt event monitoring

Martijn Vermeer ^{*1}, David Völgyes¹, Malcolm McMillan², and Daniele Fantin¹

¹Science and Technology AS, Oslo, Norway

²Centre of Excellence in Environmental Data Science and UK Centre for Polar Observation & Modelling, Lancaster University, UK

Keywords: deep learning, surface mass balance, melt dynamics, Greenland, ice sheet, CryoSat-2, MODIS

Abstract

Measuring the mass balance of ice sheets is important with respect to understanding among others sea level rise, glacier dynamics, global ocean circulation and marine ecosystems. One important parameter of the mass balance is surface melt, which can be estimated from different satellite data sources. In this study we investigate the potential of utilizing machine learning techniques for CryoSat-2 (CS2) radar altimeter waveform classification in order to derive melt information. Training data is derived by spatio-temporally matching of CS2 measurements with MODIS land surface temperature measurements. We propose a time convolution network with a fully connected classifier tail for CS2 waveform classification. In addition a non-deep learning model is implemented, providing a baseline. One of the main challenges is the high class imbalance, as surface temperatures on the interior of Greenland rarely reach the freezing point. The model performance is measured by several metrics: F₁ score, average recall and Matthews correlation coefficient. The results of this proof of concept study indicate feasibility.

1 Introduction

1.1 Background

The Greenland and Antarctic ice sheets are major components within Earth's climate system, knowl-

edge about ice sheet dynamics is critical to understand climate change. The surface mass balance (SMB) is the net mass balance of accumulation and ablation processes operating on the surface of the ice sheet. Together with basal melt and glacial discharge SMB forms the total mass balance of the ice sheet. Studying the ice sheet mass balance is of key importance to predict sea level rise [1], furthermore it also directly affects glacier dynamics, global ocean circulation and marine ecosystems.

1.2 Problem statement

One important SMB parameter is surface melt. Previous studies on melt monitoring have utilized among others scatterometers [15] and spectroradiometers [2]. However, a continuous record of melt information is currently not available, among others because spectroradiometers such as the Moderate Resolution Imaging Spectroradiometer (MODIS) [4] are impacted by clouds. Future sensors could address the shortcomings of the current data acquisitions, but they will not be able to fill the gap in historical data. CryoSat-2 [10] is a radar altimeter launched in 2010, it has a 250m wide footprint fully covering the arctic regions at monthly intervals. Radar measurements are invariant to clouds, hence CryoSat-2 has the potential of providing a seamless melt record since 2010.

1.3 Scope of the study

It has been found that the presence of liquid water in the snowpack causes subtle changes in the CS2 altimeter waveform [12]. Deriving melt information from CS2 waveforms is complex as there are many snowpack properties (density, water content, grain size and shape, layering etc.) influencing the

*Corresponding Author: vermeer@stcorp.no

waveform. In this study we evaluate whether traditional machine learning (ML) and deep learning (DL) techniques can be exploited for the extraction of melt information from CS2 waveforms. We are not aware of previous studies utilizing ML/DL techniques for classifying CS2 waveforms to derive melt information.

1.4 Machine learning literature

DL [7] has dramatically improved the state of the art in a large variety of complex classification tasks. Convolutional neural networks (CNNs) [11][3] in particular have revolutionized the field of computer vision, including data extraction from remotely sensed imagery. Although less common, 1D CNNs have similarly successfully been applied on a large variety of different applications [5]. For signal classification 1D variants of common CNN architecture can be exploited such as the VGG-like architecture presented by [13]. To the best of our knowledge there are no studies exploiting 1D CNNs for the classification of CS2 waveforms.

ML approaches on the other hand have been exploited to classify CS2 waveforms for sea ice lead detection [8]. Non-deep learning based ML approaches are widely used in remote sensing. They demand fewer compute resources but require feature engineering and hyper-parameter search. Due to the lightweight nature of traditional machine learning, many automated machine learning approaches have been published, including among others TPOT [6] and Auto-Weka [14]. These frameworks do extensive architecture search next to the hyper-parameter search while they try to avoid overfit by using cross-validation.

2 Methodology

We propose a proof of concept deep learning approach, but given that we did not find comparable results, we decided to set up a baseline traditional machine learning approach too. In the following subsections the four main parts of the methodology will be described: 1) the data generation, preparing a dataset of CS2 waveforms labeled as melt or no melt, 2) deciding on data splits and evaluation metrics, 3) the ML baseline model, 4) the proposed DL model.

2.1 Data Generation

Over the interior and margins of the Greenland ice sheet CS2 operates in Low resolution mode (LRM) and SAR Interferometry Mode (SARIn) respectively. In this study we focus on the LRM measurements over the interior of Greenland (see Figure 1), with the aim of monitoring melt in the interior of the ice sheet. The distinct leading edge arises from the reflection from the icesheet surface, reflections from greater depth hold information on the properties of the snowpack. Note that in this study we are not interested in the timing of the leading edge to derive surface elevation, we are merely interested in the waveform shape to detect melt. In situ data on melt is not available, therefore ground truth reference data is derived by spatio-temporal matching of the CS2 data with daily land surface temperature (LST) MODIS measurements. Not all data can be matched as MODIS can only measure LST in cloud free conditions. A total of 3.7 million CS2 waveforms are successfully matched for the complete 2011-2020 archive, taking only summer months (Jun-Aug) into account. Outside summer months temperatures rarely reach the freezing point, hence there are almost no melt events. More data from winter months would just increase the already significant class imbalance between melt and no-melt.

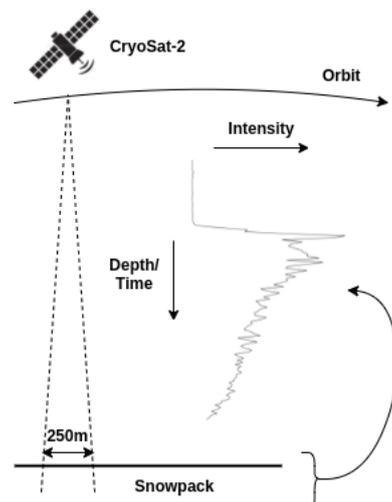


Figure 1: CryoSat-2 waveform collection over the Greenland ice sheet.

LST is thresholded at -2°C yielding roughly 6%

	CS2 (all)	CS2 (melt)	years (20xx)
train	2,608,999	148,734	11, 13-16, 18-19
val	722,046	48,682	12, 17
test	337,182	8,202	20

Table 1: Training, validation and test data splits.

of the CS2 waveforms, which are labeled as melt. Due to time lag and a measurement footprints of 250m and 1km for CS2 and MODIS respectively, occurrence of melt can be expected for temperatures slightly below the freezing point. In final form the dataset consists of radar altimeter waveforms from CS2 labeled as melt or no-melt using MODIS LST.

2.2 Data splits and Evaluation metrics

Most of the data is used for training but 3 seasons, 2012, 2017 and 2020, are left out for validation and testing purposes, see Table 1. 2012 was an extreme year with extensive melt happening on Greenland, whilst 2017 was a more average year. 2020 was left out in order to test how well our approach can extrapolate outside the training period. The main evaluation metric was Matthews Correlation Coefficient (MCC), which is a popular metric when dealing with a class-imbalanced classification problem. In addition we report on F_1 and the average recall of the melt and no-melt classes. The average recall is specifically added since we are mostly interested in the share of the actual melt and no-melt waveforms that was predicted correctly, rather than the share of predictions that was actually correct (precision). The latter is heavily influenced by class imbalance, which changes from season to season in our specific case.

2.3 Machine learning baseline

Due to the lack of reference results, a traditional machine learning based approach was implemented based on the TPOT AutoML framework [6]. TPOT provides robust results via extensive architecture and hyper-parameter search. Internally TPOT makes use of k-fold cross validation, for this both the training and validation sets described

above are used. As a baseline we extract 13 waveform features capturing properties of the entire waveform, leading edge, trailing edge and waveform noise, subsequently we use TPOT to find an optimized ML pipeline (see Listings 1) for classifying the CS2 waveforms based on the extracted features. The extracted features were the following: `std`, `mean`, `argmax`, `slope_min`, `slope_max`, `slope_argmin`, `slope_argmax`, `slope_std`, `slope_mean`, `noise`, `le_start`, `te_slope_mean`, `te_slope_std`.

Listing 1: The pipeline found by TPOT

```
ExtraTreesClassifier (
  MinMaxScaler (
    BFSampler(input_matrix ,
              gamma=0.35)),
  bootstrap=True ,
  criterion=entropy ,
  max_features=0.65 ,
  min_samples_leaf=15 ,
  min_samples_split=17 ,
  n_estimators=100)
```

2.4 Proposed deep learning approach

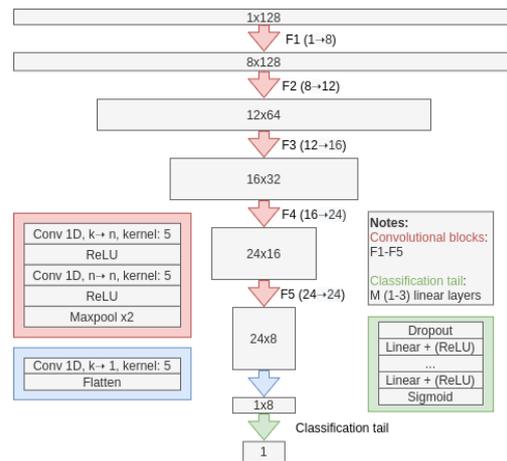


Figure 2: Architecture of the CS2-net for classification of CryoSat-2 waveforms.

As there are no previous studies on DL for classifying CS2 waveforms a basic design is selected. The

final VGG-16 inspired architecture, referred to as CS2-net, consists of 1D time convolution decoder and a classification tail (see Figure 2). The kernel size is fixed to 5 and the number of convolutional blocks to 5, this ensures that the field of view incorporates the entire waveform. Each convolutional block consists of two convolutional layers followed by ReLU activations. The convolutional blocks are connected through 2-wide maxpooling layers. The hyper-parameters of the network are the number of filters in the 5 convolutional blocks, the number of the linear layers in the classification tail, and the number of hidden units in the linear layers. The figure shows the final architecture, after hyper-parameter search.

As there are only a limited number of waveforms labeled as melt available, the model complexity is kept to a minimum in order to avoid overfit. In addition waveforms are augmented by shifting back and forth a maximum of 5 bins along the time dimension, since we are not interested in the timing of the first reflection. In order to deal with the class imbalance we use a class balanced focal loss [9] with a gamma of 1. The focal loss puts emphasis on the hard samples and reduces the risk of overfitting on easy samples. For learning rate we used $1e-4$, decaying to $1e-5$ over time. Each training ran for 400 epochs, and we selected the model with the smallest validation loss.

3 Results

Table 3 gives an overview of the performance of CS2-net and the TPOT derived ML pipeline.

The proposed DL method outperforms the baseline classic ML model for both test and validation sets looking at the selected metrics. Both F1 and MCC are much lower for the test set than for the validation set, this is due to the increased class imbalance in the test set. Since the validation set incorporates 2012, which was a year with extreme melt the imbalance is less pronounced. The average recall is robust against this imbalance and therefore similar for both validation and test sets, indicating minimal overfit. Table 2 shows the complete confusion matrix for the CS2-net on the test set.

Figure 4 shows the seasonally averaged prediction results for 2020 alongside MODIS LST and local F1. The spatial trend of LST and pre-

		Predicted	
		no-melt	melt
Actual	no-melt	0.80	0.20
	melt	0.25	0.75

Table 2: Confusion matrix CS2-net.

dicted melt is similar, with high temperatures/melt mostly towards the margins of the ice sheet. The 2012 summer season was characterized by extreme melting [12], in figure 3 monthly LST predictions are visualized capturing this event. Changing predictions for the same location over time show the model predicts melt events and not just topography.

4 Discussion

The results indicate that the mapping of melt events by classifying CS2 waveforms is feasible based on a comparison with independent MODIS LST data, both using traditional machine learning and deep learning.

4.1 Limitations of the data

One of the main limitations of the presented approach is that LST does not correspond one on one with melt, although they are closely related. Due to the lack of actual ground truth data it is impossible to precisely quantify the performance of actual melt detection. The main issue with thresholding of the LST to determine melt is that the actual melt in the snowpack is lagging behind LST. It should also be noted that there might be local variations within the MODIS pixels and footprint of the CS2 measurements.

We expect that several improvements can be gained from changing the data. First of all the LST data could be used in another way, for example by formulating the problem as a regression problem and predicting LST directly and by doing so avoiding the need for thresholding. However, this is also an imperfect method, since the radar waveform will not only be affected by the surface, while the MODIS measurements provide data about the surface only. Another option would be to exploit a timeseries of LST measurements at each location to get a more detailed picture of the

Metric	Validation		Test	
	TPOT (ML)	CS2-net	TPOT (ML)	CS2-net
MCC	0.24	0.35	0.17	0.21
avg. recall	0.71	0.78	0.73	0.77
F ₁	0.56	0.62	0.49	0.52

Table 3: Classification results, showing the proposed DL method outperforms the baseline classic ML model.

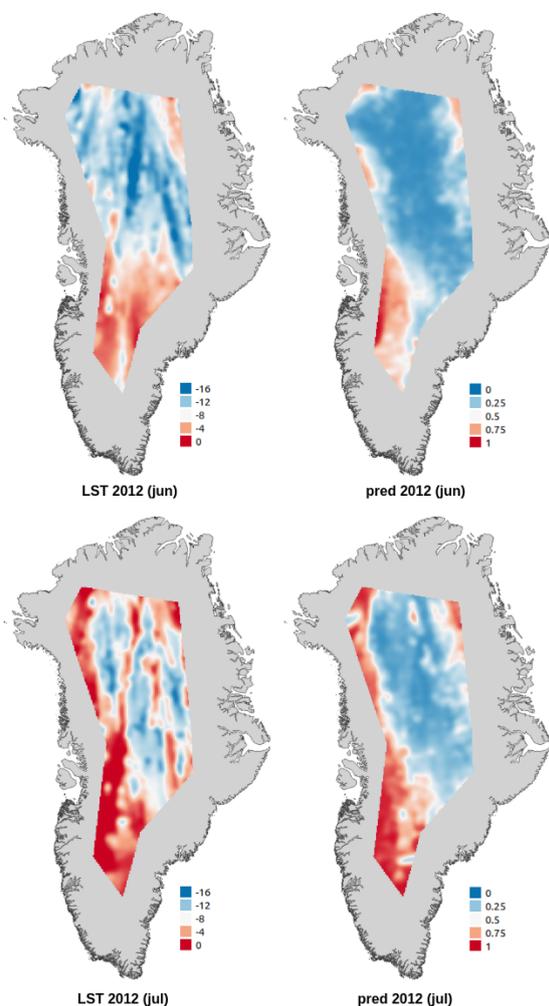


Figure 3: 2012 summer melt event CS2 prediction and MODIS LST.

expected conditions of the snowpack. In addition to MODIS data as training data source, information from other satellite sources and regional climate models could potentially be exploited.

4.2 Alternative deep learning methods

While the VGG-16 family proved to be robust in many applications, it might not be ideal base for this specific problem. Even inside this family, only a limited subgroup of 1D CNN architectures were evaluated in this study. Likely there is room for improvement by investigating different architectures. Recurrent neural networks (RNN) are frequently used in time series analysis and could be explored for this use case. In this study we chose to focus on more simple architectures first to give a proof of concept. For RNNs signal attenuation poses a significant challenge, the signal to noise ratio decreases with time/depth. Transformer based sequence modeling would face similar issues. These promising architectures can be explored in future work, but for clarity and brevity reasons the 1D CNN architecture was chosen.

4.3 Traditional machine learning

The feasibility of using traditional machine learning for melt event prediction has also been shown. One could argue that the performance difference between the deep learning and traditional method is not large, and more detailed feature engineering could bridge the gap. This is certainly possible, but manual feature engineering requires significant domain expertise and it is hard to generalize to other data sources, while the deep learning approach eliminates feature engineering step.

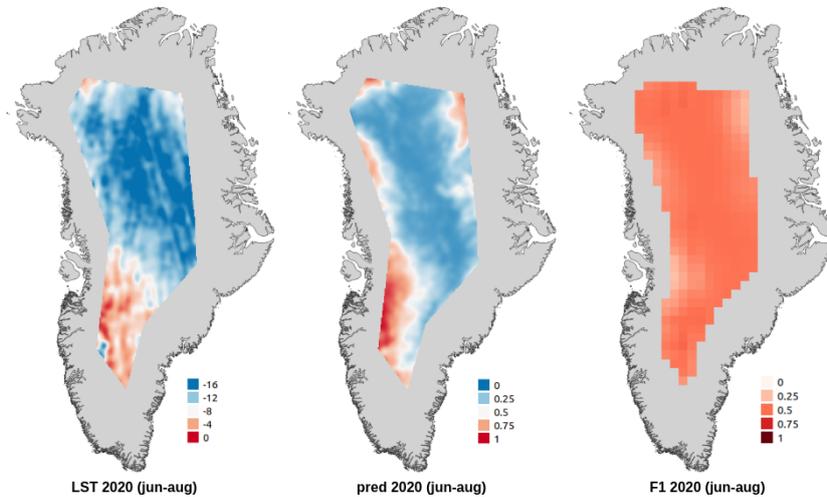


Figure 4: 2020 averaged LST, CS2-net prediction and F1 score.

5 Conclusion and future work

We have shown the feasibility of using CS2 waveforms to estimate surface melt events using machine learning, both by a traditional machine learning and deep learning approach. There are several areas where this proof of concept study could be improved, including but not limited to: increasing robustness with data augmentation, exploring alternative network structures, increasing training data quantity, using other data sources and reformulating the problem as regression task.

6 Disclosures and acknowledgments

The paper was published as part of the Earth Observation for Surface Mass Balance (EO4SMB) project, which is funded by the European Space Agency (ESA).

References

- [1] X. Fettweis, B. Franco, M. Tedesco, J. Van Angelen, J. T. Lenaerts, M. R. van den Broeke, and H. Gallée. Estimating the Greenland ice sheet surface mass balance contribution to future sea level rise using the regional atmospheric climate model mar. *The Cryosphere*, 7(2):469–489, 2013. DOI: 10.5194/TC-7-469-2013.
- [2] D. K. Hall, J. E. Box, K. A. Casey, S. J. Hook, C. A. Shuman, and K. Steffen. Comparison of satellite-derived and in-situ observations of ice and snow surface temperatures over Greenland. *Remote Sensing of Environment*, 112(10):3739–3749, 2008. DOI: 10.1016/j.rse.2008.05.007.
- [3] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, 2016. DOI: 10.1109/CVPR.2016.90.
- [4] C. Justice, J. Townshend, E. Vermote, E. Matsuoka, R. Wolfe, N. Saleous, D. Roy, and J. Morisette. An overview of MODIS Land data processing and product status. *Remote Sensing of Environment*, 83(1):3–15, 2002. DOI: 10.1016/S0034-4257(02)00084-6. The Moderate Resolution Imaging Spectroradiometer (MODIS): a new generation of Land Surface Monitoring.
- [5] S. Kiranyaz, O. Avci, O. Abdeljaber, T. Ince, M. Gabbouj, and D. J. Inman. 1D convolutional neural networks and applications: A survey. *Mechanical systems and sig-*

- nal processing*, 151:107398, 2021. DOI: 10.1016/j.ymsp.2020.107398.
- [6] T. T. Le, W. Fu, and J. H. Moore. Scaling tree-based automated machine learning to biomedical big data with a feature set selector. *Bioinformatics*, 36(1):250–256, 2020. DOI: 10.1093/bioinformatics/btz470.
- [7] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015. DOI: 10.1038/nature14539.
- [8] S. Lee, J. Im, J. Kim, M. Kim, M. Shin, H.-c. Kim, and L. J. Quackenbush. Arctic sea ice thickness estimation from CryoSat-2 satellite data using machine learning-based lead detection. *Remote Sensing*, 8(9):698, 2016. DOI: 10.3390/rs8090698.
- [9] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár. Focal Loss for Dense Object Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(2):318–327, 2020. DOI: 10.1109/TPAMI.2018.2858826.
- [10] T. Parrinello, A. Shepherd, J. Bouffard, S. Badessi, T. Casal, M. Davidson, M. Fornari, E. Maestroni, and M. Scagliola. CryoSat: ESA’s ice mission—Eight years in space. *Advances in Space Research*, 62(6):1178–1190, 2018. DOI: 10.1016/j.asr.2018.04.014.
- [11] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
- [12] T. Slater, A. Shepherd, M. McMillan, T. W. Armitage, I. Otosaka, and R. J. Arthern. Compensating changes in the penetration depth of pulse-limited radar altimetry over the Greenland ice sheet. *IEEE Transactions on Geoscience and Remote Sensing*, 2019. DOI: 10.1109/TGRS.2019.2928232.
- [13] R. A. Solovyev, M. Vakhrushev, A. Radionov, I. I. Romanova, A. A. Amerikanov, V. Aliev, and A. A. Shvets. Deep learning approaches for understanding simple speech commands. In *2020 IEEE 40th International Conference on Electronics and Nanotechnology (ELNANO)*, pages 688–693. IEEE, 2020. DOI: 10.1109/ELNANO50318.2020.9088863.
- [14] C. Thornton, F. Hutter, H. H. Hoos, and K. Leyton-Brown. Auto-WEKA: Combined selection and hyperparameter optimization of classification algorithms. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 847–855, 2013. DOI: 10.1145/2487575.2487629.
- [15] L. D. Trusel, K. E. Frey, S. B. Das, P. K. Munneke, and M. R. Van Den Broeke. Satellite-based estimates of Antarctic surface meltwater fluxes. *Geophysical Research Letters*, 40(23):6148–6153, 2013. DOI: 10.1002/2013GL058138.