



RESEARCH ARTICLE

10.1029/2024JH000521

Machine Learning Detection of Melting Layers From Radar Observations

Yan Xie¹ , Fraser King¹, Claire Pettersen¹ , and Mark Flanner¹ ¹Department of Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor, MI, USA**Key Points:**

- A binary semantic segmentation U-Net model has been developed for automatic detection of melting layers from radar observations
- The U-Net model detects more accurately and is applicable to complex weather conditions compared to a traditional method
- The U-Net model provides uncertainty estimates based on ensemble predictions, offering a quantitative measure of model confidence

Correspondence to:Y. Xie,
yanxieyx@umich.edu**Citation:**

Xie, Y., King, F., Pettersen, C., & Flanner, M. (2025). Machine learning detection of melting layers from radar observations. *Journal of Geophysical Research: Machine Learning and Computation*, 2, e2024JH000521. <https://doi.org/10.1029/2024JH000521>

Received 15 NOV 2024

Accepted 7 MAR 2025

Author Contributions:**Conceptualization:** Yan Xie, Fraser King, Claire Pettersen**Data curation:** Yan Xie**Formal analysis:** Yan Xie, Fraser King, Claire Pettersen, Mark Flanner**Funding acquisition:** Claire Pettersen**Investigation:** Yan Xie**Methodology:** Yan Xie, Fraser King**Resources:** Fraser King, Claire Pettersen**Supervision:** Fraser King,

Claire Pettersen, Mark Flanner

Validation: Yan Xie, Fraser King,

Claire Pettersen, Mark Flanner

Visualization: Yan Xie**Writing – original draft:** Yan Xie

Abstract Melting layers in the atmosphere signify where falling ice hydrometeors melt into raindrops, and can be identified by discernible radar signatures. Accurate detection of melting layers is crucial to improving quantitative precipitation estimation, weather forecasts, microwave communication, and aviation risk assessments in a changing climate. Traditional detection algorithms based on fixed thresholds or a priori assumptions lack general robustness across diverse weather conditions, which can be addressed by leveraging machine learning techniques. This study presents a binary semantic segmentation U-Net model for automatic detection of melting layers, using Ka-band vertical profiling ground radar observations collected at the North Slope of Alaska between March 2015 and February 2016. An interactive data extraction tool, ClickCollect, has been developed to generate a labeled data set of melting layer boundaries from radar observations during all seasons. Results show that the U-Net effectively detects 96% of the melting layer cases, and is applicable to complex weather conditions including heavy precipitation with velocity folding, multiple layer melting, and near-surface melt layers. Compared to a traditional detection method, the U-Net model increases the Probability of Detection by 57% and improves the mean Dice-Sørensen coefficient from 0.69 to 0.91. Furthermore, the U-Net model provides additional information of detection uncertainty based on ensemble predictions. The U-Net model and the data extraction tool can be applied to similar profiling radar instruments in different regions of the world, contributing to an enhanced understanding of the distribution and variations of melting layers.

Plain Language Summary Melting layers refer to levels in the atmosphere where snow melts into rain, which have clear signals in radar observations. Melting layer detection affects the environment and society in multiple ways via weather forecasts, precipitation estimates, risk assessments for aviation and communication systems. Traditional detection methods struggle to perform across different weather conditions. To address this issue, we implemented a machine learning model, called a U-Net, to automatically detect melting layers from radar observations collected in northern Alaska between March 2015 and February 2016. An interactive tool, ClickCollect, was developed to extract information of melting layer boundaries from radar observations, which were then used as input to the U-Net model. Results show that the U-Net model successfully detects most melting layers and performs well under extreme weather conditions. The U-Net model detects melting layers more accurately than a traditional method. Additionally, uncertainties are estimated from ensemble predictions to assess the reliability of U-Net model outputs. The presented detection framework can be applied to similar radar systems globally to help us better understand melting layers and their impacts on the environment and society in a changing climate.

1. Introduction

Melting layers refer to the phase transition zone in the atmosphere where falling ice particles melt into raindrops. Melting layers are crucial features to characterize the vertical structures in stratiform rainfall (e.g., Billault-Roux et al., 2023; Fabry & Zawadzki, 1995; Klaassen, 1988; Willis & Heymsfield, 1989), which is widespread in northern mid- and high-latitude regions (e.g., Houze, 1997; Mülmenstädt et al., 2015; Schumacher & Funk, 2023). Accurate detection of melting layers plays a key role in improving quantitative precipitation estimation, weather forecasts, microwave communication, and aviation risk assessment (e.g., Giangrande et al., 2008; Lackmann et al., 2002; Pujol et al., 2012; Zhang et al., 1994). In addition, latent heat release during phase transitions in melting layers can modify local temperature profiles substantially and affect the ambient atmospheric stability (e.g., Hu et al., 2024; Lackmann et al., 2002; Tao et al., 2010).

Studies have found that melting layers are usually located several hundred to a few thousand meters above the ground, with a typical thickness around a few hundred meters (e.g., Matrosov, 2008; Shin et al., 2000; Song

© 2025 The Author(s). *Journal of Geophysical Research: Machine Learning and Computation* published by Wiley Periodicals LLC on behalf of American Geophysical Union.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

Writing – review & editing: Yan Xie,
Fraser King, Claire Pettersen,
Mark Flanner

et al., 2021). The boundary heights and thickness of melting layers are determined by multiple environmental factors and provide valuable insights into the precipitation development (Fabry & Zawadzki, 1995). The melting layer heights are dependent on surface temperature, specific humidity, and atmospheric lapse rate, varying with geolocation and season (e.g., Heymsfield et al., 2021; Shates et al., 2023; Song et al., 2021). Meanwhile, the thicknesses of melting layers are associated with riming conditions, precipitation intensity, cloud microphysics, and orographic effects (e.g., Hu et al., 2024; Klaassen, 1988; Kumjian et al., 2016). As the climate warms, stratiform rainfall and melting layers will undergo considerable changes (Dougherty et al., 2023; Prein & Heymsfield, 2020). Prein and Heymsfield (2020) demonstrated that melting level heights over global land areas increased by approximately 100 m in altitude as temperatures rose between 1979 and 2010, leading to more frequent and intense rainfall.

Melting layers exhibit discernible features in radar observations and can be detected based on these characteristics. At the top of melting layers, partially melted snowflakes become more reflective due to the formation of liquid water shells over ice particles, which leads to enhanced reflectivity in vertical radar profiles (e.g., Austin & Bemis, 1950; Fabry & Zawadzki, 1995; Zhang et al., 2008). The snowflakes fully melt into raindrops at the bottom of melting layers, reduce in size, and fall faster toward the surface (e.g., Atlas et al., 1973; Klaassen, 1988). This acceleration is associated with Doppler velocity increases and is distinct in the profiles of Doppler velocity gradients with respect to height (e.g., Shates et al., 2023; White et al., 2002; Williams et al., 1995). Additional variables in dual polarization radar observations, including differential reflectivity, linear depolarization ratio, and copolar correlation coefficient, are also applied to identify melting layers (e.g., Giangrande et al., 2008; Li & Moiseev, 2020; Matrosov et al., 2007).

Melting layer detection algorithms have been also developed for cloud and precipitation radars onboard satellites of Tropical Rainfall Measurements Mission, CloudSat, and Global Precipitation Measurement (Le & Chandrasekar, 2013; Sassen et al., 2007; Shin et al., 2000). Satellite radars provide continuous measurements of melting layers across large spatial scales, but can be contaminated by ground clutter and become unreliable near the surface (Bennartz et al., 2019; Maahn et al., 2014). Therefore, ground radar observations provide complementary insights to satellite measurements, especially for shallow melting layers that are frequently observed in the mid-latitudes during winter and in high-latitudes generally (e.g., Cooper et al., 2022; Hu et al., 2024; Pettersen et al., 2020; Shates et al., 2023).

Previous studies have implemented various methods to detect melting layers based on observations of vertically pointing ground radars (e.g., Brast & Markmann, 2020; Fabry & Zawadzki, 1995; Li & Moiseev, 2020; Shates et al., 2023; Song et al., 2021). Fabry and Zawadzki (1995) used the gradient of radar reflectivity to detect melting layer boundaries, which was later improved in Li and Moiseev (2020) by incorporating copolar correlation coefficient thresholds. Song et al. (2021) determined the top and bottom heights of melting layers using the vertical gradient of Doppler velocity, along with multi-step criteria of dual polarization variables. Recently, Shates et al. (2023) demonstrated a framework to effectively determine rain-snow levels using linear thresholds of radar reflectivity, doppler velocity and its vertical gradient adopted from White et al. (2002). These melting layer detection algorithms are based on fixed thresholds or empirical assumptions, however, lack broad applicability and may not perform well in heavy precipitation due to attenuation (Brast & Markmann, 2020; Song et al., 2021). Brast and Markmann (2020) demonstrated the feasibility of using machine learning techniques, specifically a Multilayer Perceptron (MLP) model, to detect melting layers across various weather conditions. However, standard MLP structures are less effective in capturing spatial structures compared to other deep neural networks (e.g., Botalb et al., 2018; Driss et al., 2017), such as the Convolutional Neural Networks (CNN).

The U-Net models are a particular category of CNN known for their accuracy, efficiency, and flexibility in identifying objects with subtle details from digital images (e.g., Huang et al., 2020; Ronneberger et al., 2015; Zhou et al., 2018). U-Net models can capture nonlinear relationships between input features and the target objects, and have shown potential in a wide range of applications, such as remote sensing, precipitation forecasts, and gap-filling of missing data (e.g., Chase et al., 2024; Han et al., 2022; King et al., 2024; Ren et al., 2022). This work aims to develop a machine learning method, that is, a U-Net model, for the automatic detection of melting layers from vertically profiling ground-based radar observations. Radar reflectivity and Doppler velocity observations collected under various weather conditions between March 2015 and February 2016 at the Department of Energy's North Slope of Alaska instrument site are used to train, validate, and test the U-Net model. An interactive tool is also developed and implemented to facilitate the manual labeling of melting layers. The trained U-Net

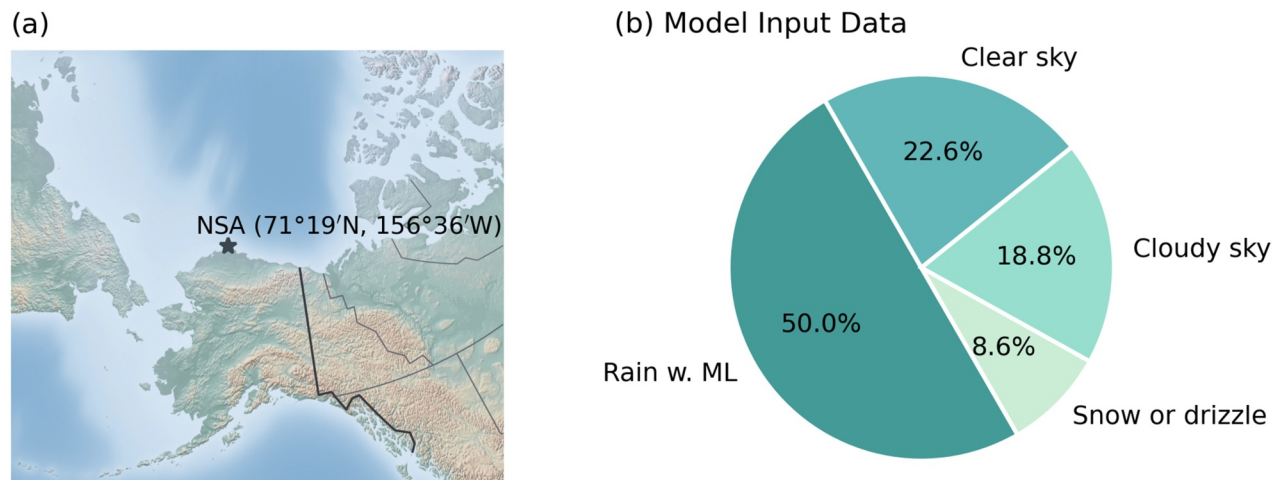


Figure 1. Details of the ground-based Ka-band Zenith Radar observations used to develop the model. (a) Geographic location of the North Slope of Alaska (NSA) site. (b) Composition of the 2,218 model inputs which contain labeled radar observations under various weather conditions between March 2015 and February 2016.

model is compared to a traditional detection method based on empirical thresholds to evaluate its performance in detecting the melting layers in terms of occurrence, boundary heights, and thickness. Section 2 introduces the radar observations and the manually labeled training data set. Section 3 describes the methods for melting layer detection and metrics for performance evaluation. Section 4 presents the melting layer detection results, comparisons between the U-Net model and the traditional method, and interpretations of the U-Net model. Section 5 summarizes and concludes the key results.

2. Data

2.1. Ground-Based Radar Observations

This study uses ground-based vertical profiling radar observations from the Department of Energy (DOE) Atmospheric Radiation Measurement (ARM) research facility at the North Slope of Alaska (NSA) site. The NSA site (71°19'N, 156°36'W, 8 m above sea level) is located along the northern coast of Alaska as shown in Figure 1a. The Ka-band Zenith Radar (KAZR; Feng et al., 2011; Kollias et al., 2020) at the NSA site is a ground-based zenith-pointing radar that operates at 35 GHz. The KAZR measures vertical profiles of clouds and precipitation up to 12 km above ground level (AGL), with high temporal resolution around 4 s and vertical resolution of 30 m. To mitigate the ground clutter contamination, KAZR measurements are collected starting from the lowest reliable range bin at 190 m AGL. KAZR observations of radar reflectivity and Doppler velocity between March 2015 and February 2016 are utilized to determine and detect melting layers in this study. Figure 1b displays the composition of KAZR measurements used to develop the U-Net model. The melting layer presence in each case can be identified when applying ClickCollect to KAZR observations. There are in total 1,109 melting layer cases, which have been all included in the data set (Figure 1b). An equal number of cases without melting layers are then randomly selected, resulting in 2,218 cases under various weather conditions. This balanced data set contains equal representations of cases with and without melting layers, which helps to train the neural network to focus on radar observation features most relevant to classification rather than become biased toward the dominant category (Mazurowski et al., 2008; Ustuner et al., 2016).

2.2. Manually Labeled Melting Layer Data Set by ClickCollect

An interactive tool, named ClickCollect, has been developed to generate a labeled data set of melting layer boundaries using KAZR observations. As shown in Figure 2, ClickCollect first produces a visual display of the time series of radar reflectivity (Z_e), doppler velocity (V_d) and its vertical gradient (the derivative of V_d with respect to height is referred to as dV_d), allowing users to manually identify the presence of a melting layer. If melting layers exist, ClickCollect prompts users to click on the upper and lower boundaries of one or multiple melting layers. The Doppler velocity gradients are displayed to facilitate the selection of melting layer boundaries which are signified by horizontal bands of negative Doppler velocity gradients. In addition, the ClickCollect

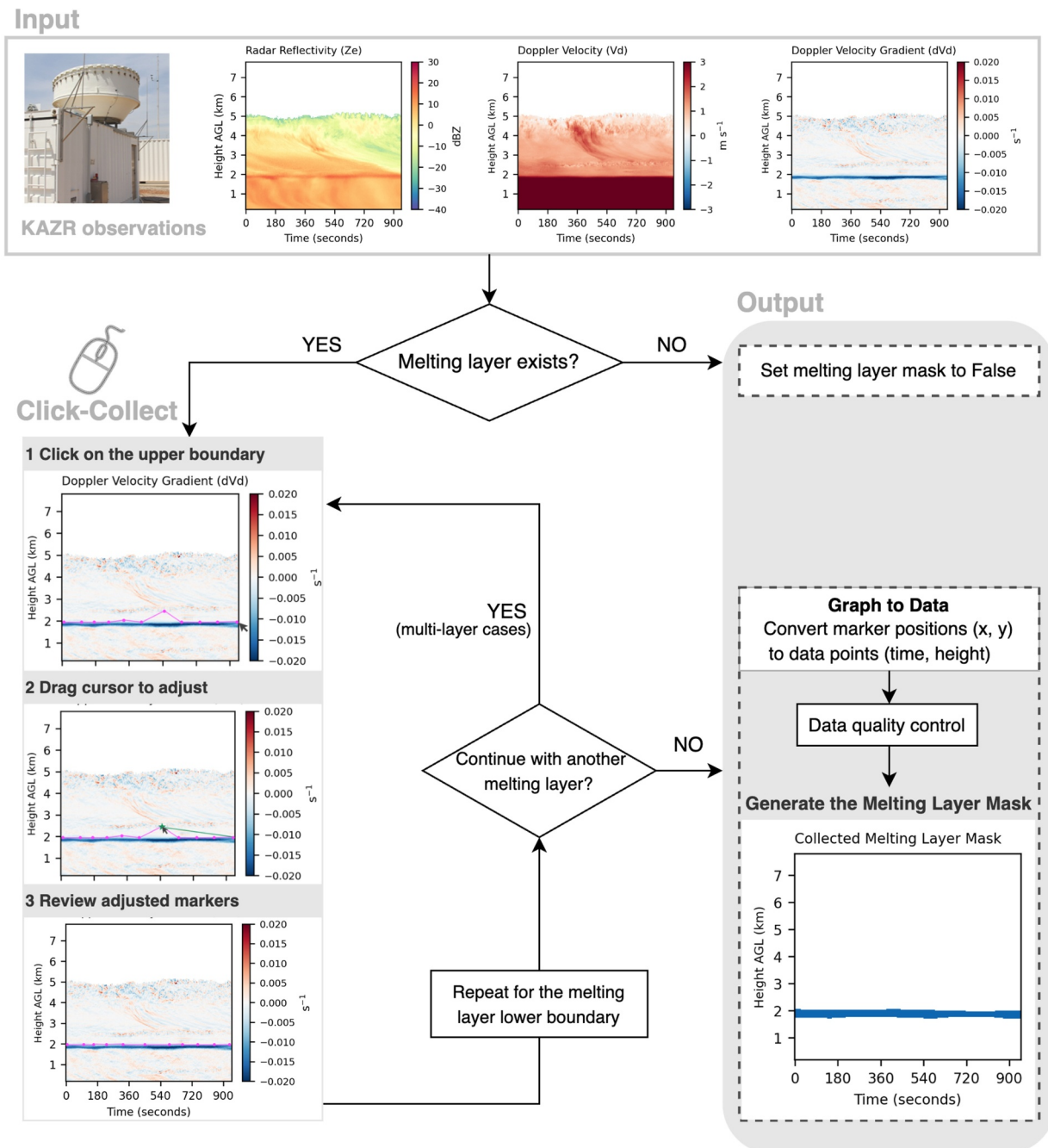


Figure 2. Flowchart of an interactive data extraction tool, named ClickCollect, which helps users to collect information of melting layer boundaries from KAZR observations and generate binary segmentation masks of melting layers. The labeled melting layers are used as ground-truths to train, validate and test the U-Net model.

allows users to precisely adjust positions of selected points by dragging as shown in the lower left panel of Figure 2, offering exceptional flexibility in refining melting layer boundaries. This tool accommodates errors in cursor placement by enabling easy corrections, while tracking cursor movements and converting graphical inputs into information of melting layer boundaries. Collected data points of selected melting layer boundary heights are then interpolated to the original KAZR temporal resolution and examined to ensure the upper boundaries are above the lower ones. Each output of the ClickCollect includes the time series of the upper and lower boundary heights and a binary melting layer mask with the same dimension as the input data set. The ClickCollect is not tied to a specific location or radar instrument, making it adaptable for use by other researchers with their own setups.

In total, over 700 hr of KAZR observations covering all seasons between March 2015 and February 2016 are processed by the ClickCollect. KAZR observations spanning 256 vertical range bins from 0.19 to 7.84 km are used for melting layer detection. The KAZR observations together with the melting layer masks are then segmented into intervals with 256 time steps to generate the 256×256 chunks used as model input, which ensures a balance between computation efficiency and sufficient spatial and temporal dimensions to capture the melting layers. Each input corresponds to a duration of approximately 16 min and is separated from other inputs by temporal gaps of at least 3 min. Among the 2,218 model input chunks, 50% of the cases were identified to have melting layers of various shapes and heights, 8.6% cases had snow or drizzle, and 41.4% had no precipitation, composed of 22.6% clear-sky and 18.8% cloudy-sky situations (Figure 1b).

2.3. Data Preprocessing

Prior to training the neural network, a data preprocessing pipeline illustrated in Figure 3a is applied to ensure the quality and consistency of the model input data. To enhance the input data quality, missing values and background noises in KAZR observations are masked and treated as clear-sky conditions. KAZR observations are then standardized between -1 and 1 based on the maximum and minimum values of the input data. This standardization process transforms the input observations to consistent scales, which reduces biases across input features and helps to train the model more efficiently. The standardized observations are then combined with corresponding melting layer masks detected by ClickCollect to generate image tensors with shape of $256 \times 256 \times 4$. The first two dimensions of the image sensor represent height and time, while the last dimension refers to three input variables (Z_e , V_d , dV_d) and the ground-truth melting layer mask.

The 2,218 image tensors are then split into training (85%), validation (5%), and testing (10%) data sets. The data sets are split in a non-shuffle approach that: (a) sorts tensors by time from the beginning to the end of each month, and (b) assigns the first 85% of the tensors in each month to the training data set, the following 5% to the validation data set, and the last 10% to the testing data set. The training data set includes cases in the first 25 days of each month, while the testing data set contains cases occurring at the end of each month. This non-shuffle data split strategy together with the temporal gaps between cases help to avoid data leakage due to autocorrelation, which would occur if the model was trained by cases temporally correlated with those in the validation or (and) testing data set(s) and leads to unreliable model performance (Honari et al., 2019; King et al., 2022). Flip augmentation has been widely used in training convolutional neural networks with a limited data set and performs well in reducing overfitting (Shorten & Khoshgoftaar, 2019). The training data set is horizontally flipped to generate a mirror version of the original input. The final training data set is composed of the original input and its flipped version to increase the size of the training data set and to address potential overfitting issues during the model training process. In addition, the horizontal flip augmentation introduces more variability in melting layer orientations while maintaining the physical interpretation, which helps to improve the model generalization (Wu et al., 2020).

3. Methods

3.1. Melting Layer Detection Techniques

3.1.1. Traditional Detection Method

A traditional detection method based on empirical thresholds is applied as a baseline method for comparison against the U-Net model. Shates et al. (2023) outlines a framework to detect rain-snow levels using a vertically pointing radar, which is then adapted for KAZR observations to detect stratiform rainfall at the North Slope of Alaska (Xie et al., 2024c). The traditional method is built on this framework and consists of three empirical criteria, which is essentially a simple decision tree with three categorical variables. First, radar reflectivity at the lowest reliable bin, 190 m AGL, larger than -10 dBZ is used to identify precipitation (Chandra et al., 2015). Second, a conservative threshold of Doppler velocity larger than 3 m s^{-1} at 190 m AGL is applied to distinguish rainfall from snowfall as raindrops fall faster than snowflakes (Atlas et al., 1973). Additionally, the process of snowflakes melting into raindrops is apparent in the profiles of Doppler velocity gradients with respect to height (dV_d), and can be effectively detected using empirical thresholds (e.g., Shates et al., 2023; White et al., 2002). A threshold of dV_d less than $-7.5 \times 10^{-3} \text{ s}^{-1}$ following Shates et al. (2023) is implemented to determine the heights of melting layers.

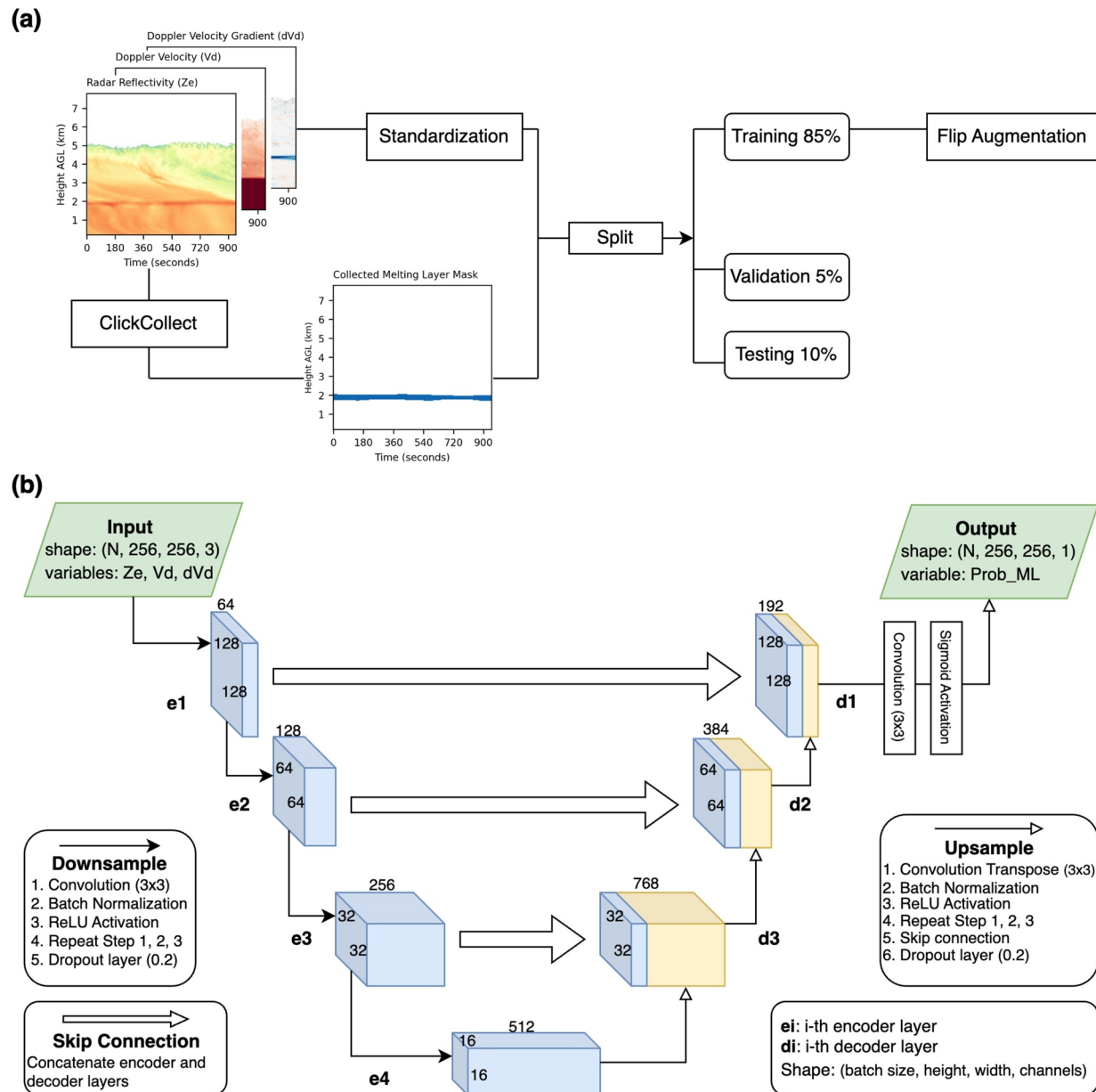


Figure 3. U-Net model implementation details. (a) Data preprocessing steps transforming original KAZR observations into U-Net input data. (b) Diagram of the U-Net model architecture.

3.1.2. U-Net Model

The U-Net model is a type of convolutional neural network specifically designed for image segmentation tasks (Ronneberger et al., 2015). U-Net models can accurately and efficiently label multiple distinct regions in a digital image (e.g., King et al., 2024; Siddique et al., 2021). Furthermore, U-Net models perform well even with a relatively small training data set, and are applicable to situations where labeled data is limited (Ronneberger et al., 2015; Zhao et al., 2020). In this study, a U-Net model has been developed to detect melting layers from radar observations (Figure 3b). The U-shaped architecture shown in Figure 3b consists of a downsampling and an upsampling path with multiple layers. The input radar observations are passed through encoder layers which trade spatial information with feature space as the layer resolution decreases and depth increases moving toward the bottleneck layer of the U-Net architecture. The decoder layers progressively reconstruct the spatial dimension of feature maps. Meanwhile, skip connections merge feature maps in decoder layers with spatial details in

Table 1
Summary of Hyperparameter Sweep Details and Final Tuned Values

| Hyperparameter | Parameter space | Tuned value |
|----------------|--|--------------------|
| Depth | [3, 4, 5, 6] | 4 |
| Batch size | [2, 4, 8, 16] | 16 |
| Learning rate | $[5 \times 10^{-2}, 5 \times 10^{-3}, 5 \times 10^{-4}, 5 \times 10^{-5}, 5 \times 10^{-6}, 5 \times 10^{-7}]$ | 5×10^{-3} |
| Filter count | [16, 32, 64] | 64 |
| Dropout rate | [0.01, 0.1, 0.2, 0.5] | 0.2 |
| Optimizer | [Adam, RMSprop, SGD] | RMSprop |

corresponding encoder layers, which helps to restore the original spatial details lost during downsampling and leads to more accurate and robust detection of melting layers. The output of U-Net is a probability map of melting layer existence for each pixel in the input image. A probability threshold of 50% is applied to generate a binary segmentation mask.

Three loss functions are compared during the U-Net training process: Binary Cross Entropy (BCE), Dice Loss, and their combination defined by Equation 1, where n represents the number of pixels, p and y refer to the predicted and true melting layer probability, $\sigma = 10^{-9}$ is a smooth parameter. The combo loss function combines the strengths of BCE in pixel-wise classification and of Dice Loss in characterizing the overall shapes (Jadon, 2020; Taghanaki et al., 2019):

$$\text{Loss}_{\text{combo}} = \text{Loss}_{\text{BCE}} + \text{Loss}_{\text{Dice}} = \left[-\frac{1}{n} \sum_{i=1}^n (y_i \cdot \log p_i + (1 - y_i) \cdot \log(1 - p_i)) \right] + \left[1 - \frac{\left(2 \cdot \sum_{i=1}^n y_i \cdot p_i \right) + \sigma}{\sum_{i=1}^n y_i + \sum_{i=1}^n p_i + \sigma} \right] \quad (1)$$

Evaluations show that BCE loss, even weighted, leads to a relatively low detection accuracy because melting layers account for a small fraction of the input images, which may result in a poor model performance on the minority category. The Dice Loss and the combo loss improve the accuracy by emphasizing the overlap between predicted and target melting layers. The combo loss shows a more efficient and robust performance than the Dice Loss, with comparable magnitudes of BCE and Dice losses, and is eventually applied to compile the U-Net model.

Hyperparameters of the U-Net model listed in Table 1 are tuned to optimize the model performance. Each hyperparameter sweep trains the U-Net model with a combination of hyperparameter values. The search for the best combination of hyperparameter values is efficiently conducted using a Bayesian optimization approach (Turner et al., 2021; Wu et al., 2019). In total, 50 hyperparameter sweeps are performed on a high-performance computer cluster (Table 2), with early stopping implemented to prevent overfitting (Prechelt, 1998; Ying, 2019). U-Net training is then performed with the tuned hyperparameters and stops after 67 epochs when the validation loss ceases to improve, effectively preventing overfitting.

Table 2
Summary of Computation Resources Used for Hyperparameter Tuning and Model Training With Tensorflow v2.17.0 on a Linux 4.18.0 Operating System Using the University of Michigan's High Performance Computing (HPC) Clusters <https://its.umich.edu/advanced-research-computing/high-performance-computing/great-lakes>

| Application | Hardware | GPU memory | Requested system memory | Computation time hh:mm:ss |
|-----------------------|---|------------|-------------------------|---------------------------|
| Hyperparameter tuning | 1 × NVIDIA A40 2 × 3.0 GHz Intel Xeon Gold 6154 | 48 GB | 40 GB | 18:04:09 |
| U-Net training | 1 × NVIDIA Tesla V100 2 × 3.0 GHz Intel Xeon Gold 6154 | 16 GB | 20 GB | 0:33:23 |

Gal and Ghahramani (2016) showed that dropout layers of the neural networks can be utilized as an approximate Bayesian inference to estimate the model uncertainty, which is referred as Monte Carlo dropout (MC dropout). Although subject to statistical assumptions, the MC dropout is a computationally efficient tool for identifying regions where neural network outputs are relatively less reliable (e.g., Camarasa et al., 2020; King et al., 2024). MC dropout is applied to the U-Net model with tuned hyperparameters. During each run, dropout layers in the model randomly deactivate 20% of neurons, as specified by the tuned dropout rate in Table 1. Ensemble predictions are collected by running the U-Net model on the testing data set 50 times. The mean of ensemble predictions is used to generate melting layer masks. U-Net detection uncertainty is quantified as the standard error of differences between the model ensemble predictions and the ground truths.

3.2. Evaluation Metrics

Multiple metrics are used to evaluate the performance of the traditional method and U-Net model on the testing data set. Dice-Sørensen coefficient (DSC) measures the overlap between predicted melting layers and the truth as defined in Equation 2, where A and B represent the set of pixels in the predicted and true segmentation mask:

$$DSC(A, B) = \frac{2 \cdot |A \cap B|}{|A| + |B|} \quad (2)$$

Intersection over Union (IoU) calculates the ratio between the intersection of the predicted and true sets to their union following Equation 3:

$$IoU(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad (3)$$

DSC and IoU were computed for each case and averaged over the testing data set. The two metrics both evaluate the similarity in shape and location of melting layers between predictions and target, with DSC more sensitive to small overlaps and IoU penalizing overall mismatch.

Furthermore, the Success Ratio (SR), Probability of Detection (POD), Critical Success Index (CSI), and Heidke Skill Score (HSS) shown in Equations 4–7 are calculated from a case-level contingency table to assess the model skills in detecting melting layer occurrence. Detailed descriptions of these metrics can be found in previous research (e.g., Chase et al., 2022; Doswell et al., 1990; Roebber, 2009). TP denotes the amount of true positive cases, FN refers to false negative cases, and so on. All the evaluation metrics used in this study range from 0 to 1, with 1 representing perfect detection.

$$SR = \frac{TP}{TP + FP} \quad (4)$$

$$POD = \frac{TP}{TP + FN} \quad (5)$$

$$CSI = \frac{TP}{TP + FP + FN} \quad (6)$$

$$HSS = \frac{2 (TP \cdot TN - FP \cdot FN)}{[(TP + FN) \cdot (FN + TN) + (TP + FP) \cdot (FP + TN)]} \quad (7)$$

4. Results and Discussion

4.1. Melting Layer Detection Case Studies

Figure 4 shows the applications of ClickCollect in generating melting layer masks, which are used as ground-truth to assess the detection methods. ClickCollect enables manual labeling of melting layer positions and variations in three distinct cases featuring a discontinuous melting layer (Figure 4a), a melting layer with varying thickness (Figure 4b), and a near-surface melting layer (Figure 4c). The U-Net model shows superior performance in detecting melting layers compared to the traditional method. Figure 5 demonstrates six cases of KAZR

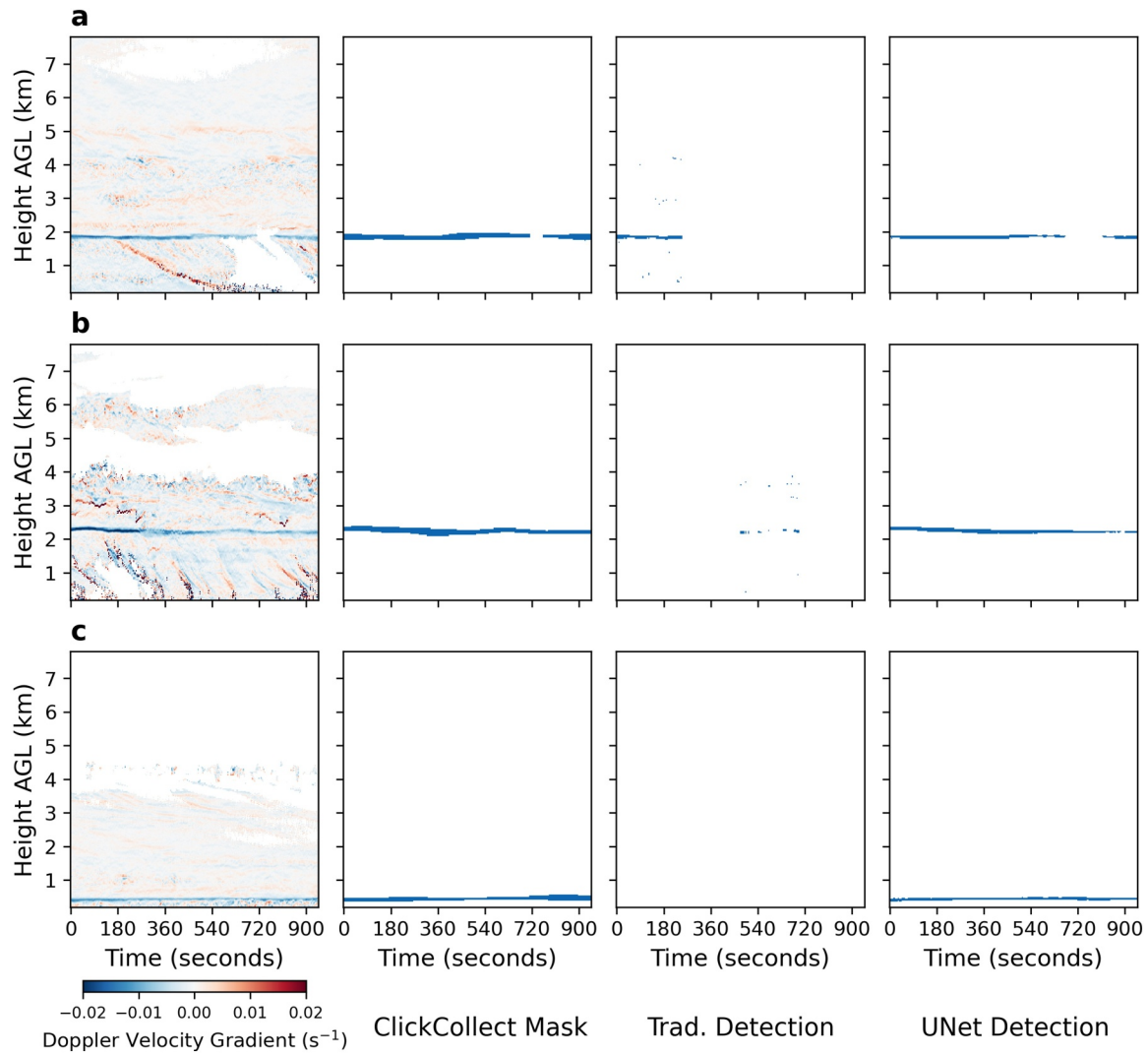


Figure 4. Application of ClickCollect to generate melting layer masks in three cases, (a) 2015-06-29 12:20:00 to 12:35:46, (b) 2015-06-29 15:00:43 to 15:16:28, and (c) 2015-08-30 13:49:16 to 14:05:02. ClickCollect would extract data points of melting layers based on KAZR observations of Doppler velocity vertical gradient (first column from left) and generate binary masks (second column). These melting layer masks were used as ground-truth to evaluate the detection performance of the traditional method (third column) and U-Net model (fourth column).

observations and the corresponding melting layer detection results under various weather conditions. Case a shows a typical stratiform rainfall with a discernible melting layer located around 1.5 km AGL. The traditional method and the U-Net model both successfully detect the melting layer. However, the traditional method misidentifies the Doppler velocity variations associated with microphysical processes at the cloud top edge as a part of the melting layer. This unrealistic result arises from the fact that the fixed threshold of the Doppler velocity gradient in the traditional method does not distinguish between the melting layer and the cloud echos. Furthermore, the traditional method will fail to detect melting layers when raindrops do not reach the ground, as shown in Case b. Meanwhile, the U-Net model is not limited by a priori assumptions or empirically derived thresholds of radar observations at the surface and provides a complete depiction of the melting layer during the precipitation development process.

The U-Net model can capture non-linear relationships between KAZR observations and the melting layers, and performs well in complex scenarios, such as with multiple melting layers (Case c), heavy precipitation (Case d), and near-surface melting layers (Case e). Case c shows a rare situation with two concurrent melting layers at different heights, which could arise from multiple freezing levels (Brast & Markmann, 2020) or different ice

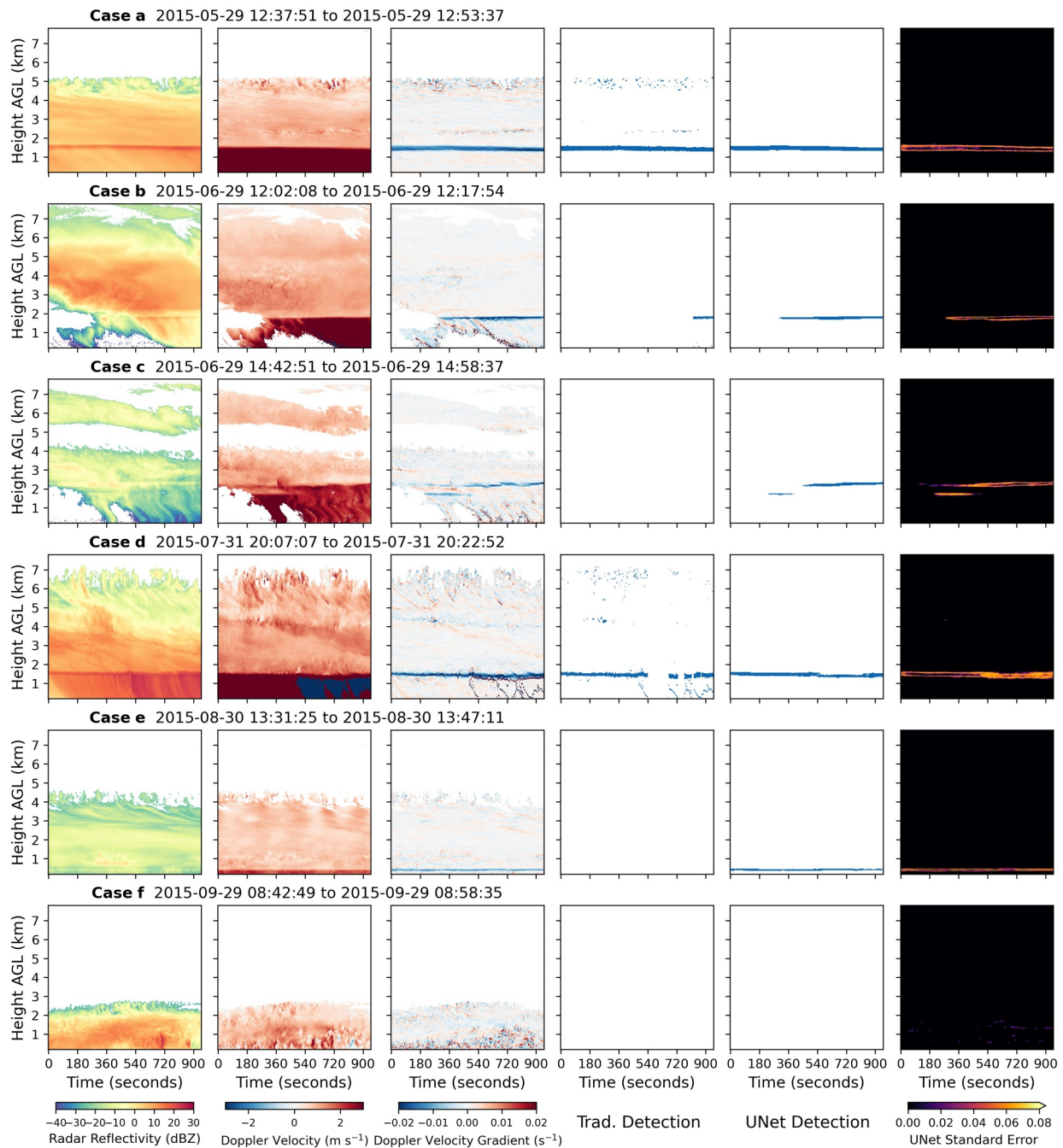


Figure 5. Examples of melting layer detection by the traditional method and the U-Net model. Each row represents a case observed at the NSA site. Columns from left to right refer to KAZR radar reflectivity, Doppler velocity, gradient of Doppler velocity with respect to height, traditional method results, U-Net model results and their uncertainty estimates. The U-Net detection results and uncertainties are based on ensemble predictions of pixel-wise model output melting layer probabilities and the ClickCollect ground truths.

particles (Li & Moisseev, 2020). Effective detection of concurrent melting layers by the U-Net model can help to better understand the vertical structures and microphysical processes within mixed-phase clouds. Case d presents a heavy stratiform rainfall event with a continuous melting layer at 1.6 km AGL. Note that the Doppler velocity values during the second half of the rainfall are negative due to velocity folding (Jay Miller et al., 1986), which occurs when the raindrops fall at a speed faster than the Nyquist velocity of the radar. Velocity folding is

Table 3

Contingency Table of Melting Layer Detections by the Traditional Method and the U-Net Model on the Testing Data Set

| Traditional method | | | | U-Net model | | | |
|--------------------|------------|-------------|-------|-----------------|------------|-------------|-------|
| Number of cases | Prediction | Observation | | Number of cases | Prediction | Observation | |
| | | ML | No-ML | | | ML | No-ML |
| Prediction | ML | 70 | 2 | Prediction | ML | 110 | 3 |
| | No-ML | 45 | 130 | | No-ML | 5 | 129 |

commonly observed when using cloud radars, such as KAZR, to detect moderate to heavy rainfall. The U-Net model proficiently detects melting layers from folded radar signals. Compared to the traditional method and a multi-step approach developed by Song et al. (2021), the U-Net model demonstrates greater applicability during heavy precipitation. Case e displays a near-surface melting layer located below 500 m AGL. The traditional method based on empirical criteria is applicable to persistent moderate rainfall (Case a) but fails to detect melting layers with relatively weak radar reflectivity at the surface (Case c & e). Given appropriate training, the U-Net model can efficiently detect shallow melting layers and provide insights into the distribution of melting layer heights at high latitudes and their seasonal

variations. Case f shows a precipitation process without a melting layer, which both the traditional method and the U-Net model successfully distinguish from cases with melting layers.

Furthermore, the U-Net detection for all cases shown in Figure 5 are average results of the ensemble predictions described in Section 3.1.2. Standard errors of differences between model ensemble predictions and the ground truths quantitatively estimate the U-Net detection uncertainty. As shown in the last column of Figure 5, the standard errors are most prominent at the melting layer boundaries where the U-Net model is trained to focus on. Also, when velocity folding is present (Case d), the U-Net model shows a lower confidence in regions within the melting layer.

4.2. Detection Performance Comparisons

The U-Net model is assessed and compared to the traditional method in terms of their performance on detecting the occurrence, boundary heights, and thickness of the melting layers. Table 3 displays the melting layer occurrence in 247 testing cases detected by the traditional method and the U-Net model. The U-Net model effectively identifies 96% of the melting layer cases, compared to 61% detected by the traditional method on the testing data set (Table 3). The U-Net model is more applicable under various weather conditions and greatly

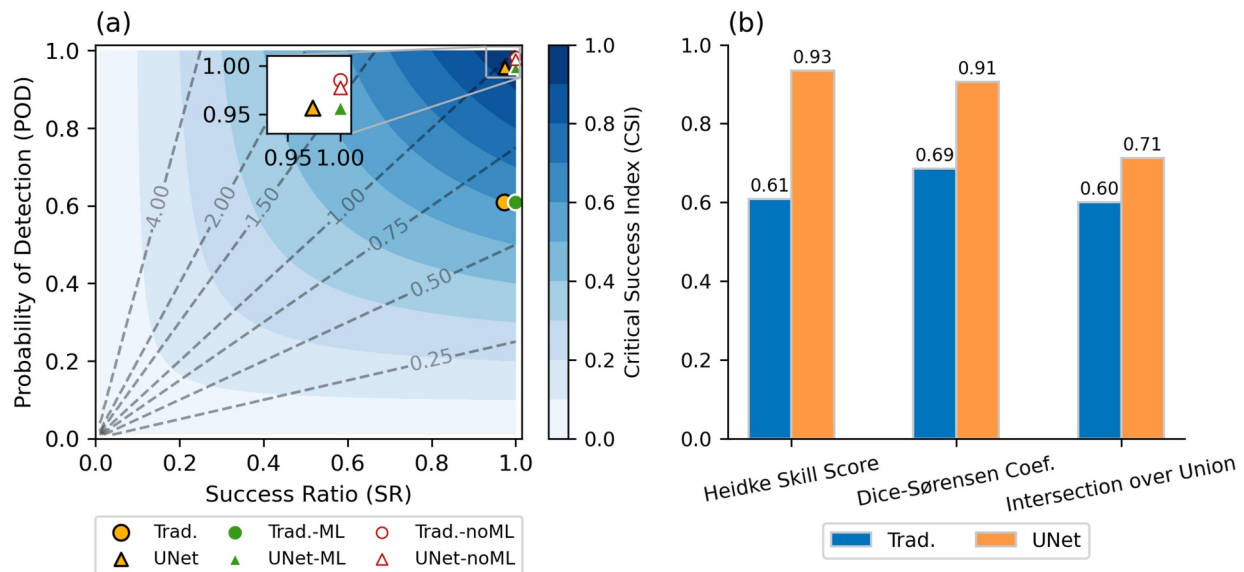


Figure 6. Comparisons of evaluation metrics between the traditional method and the U-Net model. (a) Detection performance diagram of the melting layer occurrence. Circle markers represent traditional method detections. Triangle markers refer to U-Net model detections. There are three different testing categories: all testing data indicated by filled markers with black edges, cases with melting layers (ML) indicated by filled markers with no edges, and cases without melting layers (no-ML) indicated by unfilled markers with red edges. Dashed lines represent model frequency biases. Bias smaller than 1 represents underprediction and larger than 1 indicates overprediction. (b) Comparisons of the forecast skill metric HSS, and two similarity metrics DSC and IoU which assess the techniques' proficiency in characterizing the heights and shapes of melting layers.

increases the amount of successfully detected melting layers. The traditional method shows a comparably good performance as the U-Net model in distinguishing cases without melting layers.

Figure 6a demonstrates the SR, POD, and CSI of the two detection techniques on different testing categories: all cases, melting layer cases, and cases without melting layers. The SR of the U-Net model and the traditional method are both higher than 0.97. The POD and CSI of the U-Net model is 0.96 and 0.93, showing an 57% increase in melting layer detection relative to the traditional method. The main difference between the U-Net model and the traditional method lies in the detection of melting layer cases, which is consistent with results shown in Table 3. The traditional method underpredicts the melting layer occurrence as indicated by the corresponding markers below the one-to-one bias line in Figure 6a. Meanwhile, the U-Net model detects the occurrence of melting layers accurately and effectively with its marker closely aligned to the one-to-one bias line (Figure 6a), indicating an almost unbiased performance.

The U-Net model improves the forecast skill score HSS by 0.32 compared to the linear method (Figure 6b). In addition, Figure 6b displays the mean DSC and IoU of the testing data set, both of which evaluate the similarity between detected melting layers and the target. Compared to the traditional method, the U-Net model improves the average DSC from 0.69 to 0.91 and the average IoU from 0.60 to 0.71. This demonstrates that the U-Net model is remarkably proficient at detecting pixel-wise details and characterizing the location and shape of melting layers.

Figure 7 further illustrates the distribution of melting layer boundary heights and thickness. The upper boundary heights of melting layers in the testing data set have a bimodal distribution, with peaks occurring near 1.5 and 2 km (Figure 7a). Similarly, the lower boundary heights also display a bimodal distribution peaking near 1.3 and 1.9 km (Figure 7b). The bimodal distributions result from the data splitting strategy (Section 2.3) and seasonal variations in melting layer heights associated with surface temperature changes (Hu et al., 2024; Shates et al., 2023). In the testing data set, melting layers occur at the high-latitude NSA site during May, June, July, and August. June testing cases features higher melting layers, with upper and lower boundary peaks near 2 and 1.9 km, respectively. Melting layers in May and July contribute to the mode with upper and lower peaks around 1.5 and 1.3 km. August testing cases exhibit the shallowest melting layers, with boundary heights distributed near 500 m (Figures 7a and 7b).

The U-Net model characterizes the lower and upper boundary heights well and effectively detects the bimodal distributions. Nevertheless, the median of upper boundary heights detected by the U-Net model is lower than the target value by approximately 150 m (Figure 7c). Also, the U-Net model tends to overestimate the height of lower boundaries by around 100 m on average. The underestimation of upper boundary heights and overestimation of lower boundary heights by the U-Net model therefore lead to a thinner-than-actual estimate of melting layer thicknesses. The melting layer thicknesses in the testing data set span from 30 to 510 m, with a mean depth of 180 m and a median depth of 150 m. The U-Net detections of melting layer thickness range between 30 and 270 m, with a mean value of 120 m and a median value of 135 m. It is worth noting that the sawtooth patterns in detected melting layer thickness shown in Figure 7c arise from KAZR's 30-m vertical resolution (Section 2.1), and the U-Net detection sample size is comparable to that of target values (Figure 7c).

This low bias in melting layer thickness detections can be caused by multiple factors. First, the melting layer is relatively narrow and accounts for a small region in the input image, which can make it difficult for the U-Net model to detect fine details of melting layers. In addition, the probability thresholds used to binarize the model output can affect the trade-off between increasing true positives and reducing false positives. Testing probability thresholds ranging from 30% to 70% showed minimal impacts on the U-Net detection performance. As a result, the 50% threshold is applied as a default trade-off baseline, representing equal likelihood for the two possible outcomes. Moreover, future research into improving the machine learning model architecture could further mitigate the loss of spatial information. On the other hand, the traditional method struggles to provide reliable estimates of upper boundary heights and thicknesses, although showing fairly good performance in detecting the lower boundary heights. It is worth noting that the outliers of upper boundary heights detected by the traditional method range from 3 km up to 6 km (Figure 7a), which are likely associated with cloud microphysical processes rather than melting layers as revealed in Figure 5 Case a. Melting layer thicknesses detected by the traditional method are therefore largely impacted by the inaccurate upper boundary heights and fall beyond the normal range (Figure 7c).

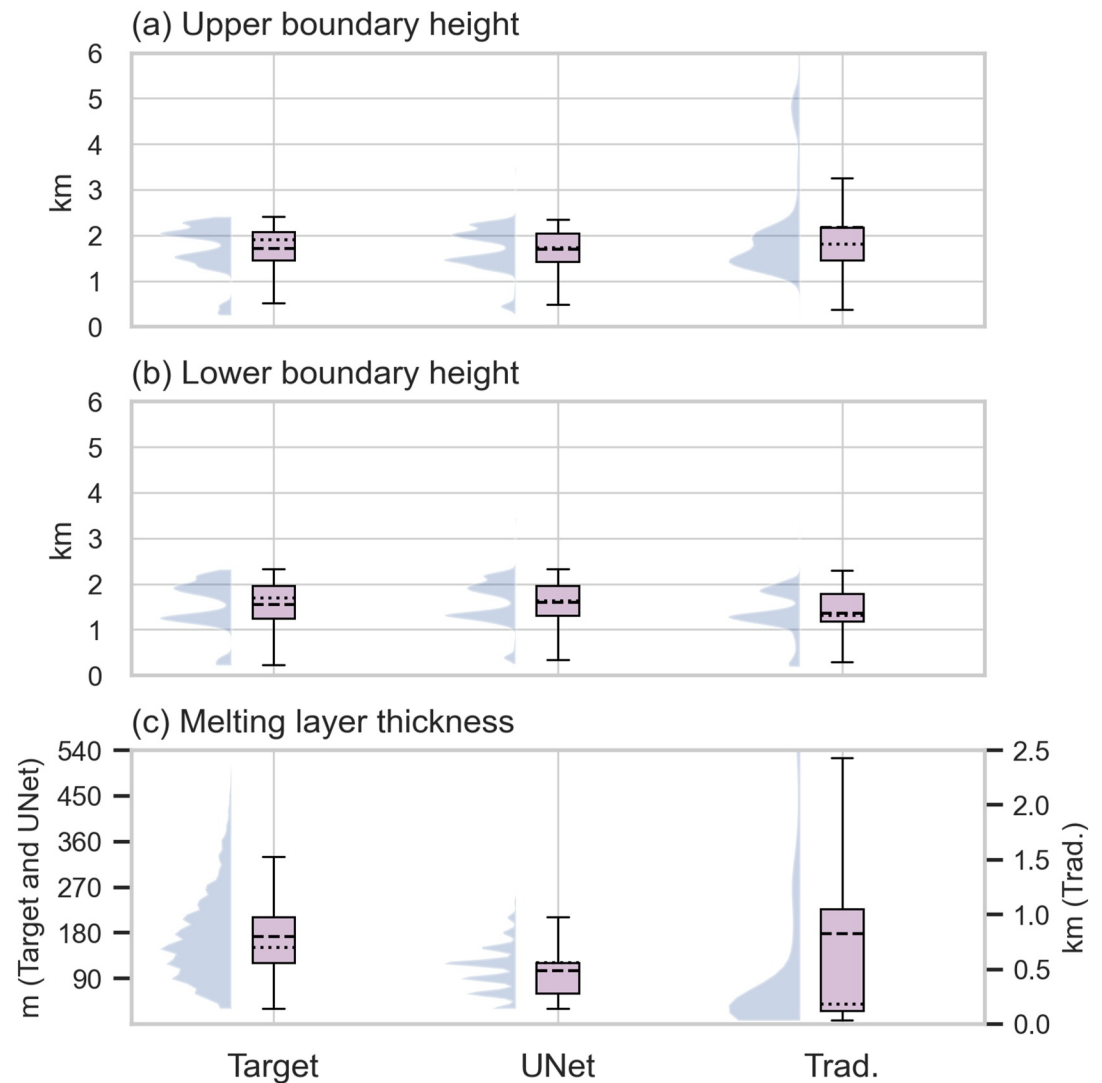


Figure 7. Distribution of melting layer boundary heights and thickness in the testing data set shown by density curves and boxplots. Rectangles of the boxplots denote the first and third quartiles. The dashed line and dotted line inside the boxes refer to the mean and median of the distribution. The boxplots exclude outlier data points which fall beyond the 1.5x interquartile range. (a) Heights of melting layer upper boundaries (b) Heights of melting layer lower boundaries (c) Thickness of melting layers. Note that since the traditional method detections are much thicker, thickness distributions are displayed on different scales. The target values and U-Net model detections are plotted against the left y-axis (m), while the traditional method results are plotted against the right y-axis (km).

4.3. U-Net Model Feature Explainability

Understanding the U-Net model features and their potential connections to the physical properties of melting layers is crucial for enhancing model performance and explainability. Figure 8 demonstrates a case with input radar observations (Figure 8a) and the corresponding feature maps generated by 64 filters in the first encoder layer of the U-Net model (Figure 3b).

Certain filters outline the overall structure of precipitation systems from radar observations (Figure 8b panels 9, 10, 22, 26, 42, 50), while others highlight gradients in radar reflectivity and Doppler velocity (Figure 8b panels 2, 19, 41, 52, 60) to distinguish between the cloud and rainfall regions. Several feature maps capture detailed characteristics of clouds (Figure 8b panels 3, 4, 15, 32, 38) and rainfall (Figure 8b panels 8, 11, 20, 48, 57). The model also identifies the shape and location of the melting layer (Figure 8b panels 24, 35, 40, 55, 62). Interpretations of the feature maps may vary depending on the specific case. However, similar patterns occur

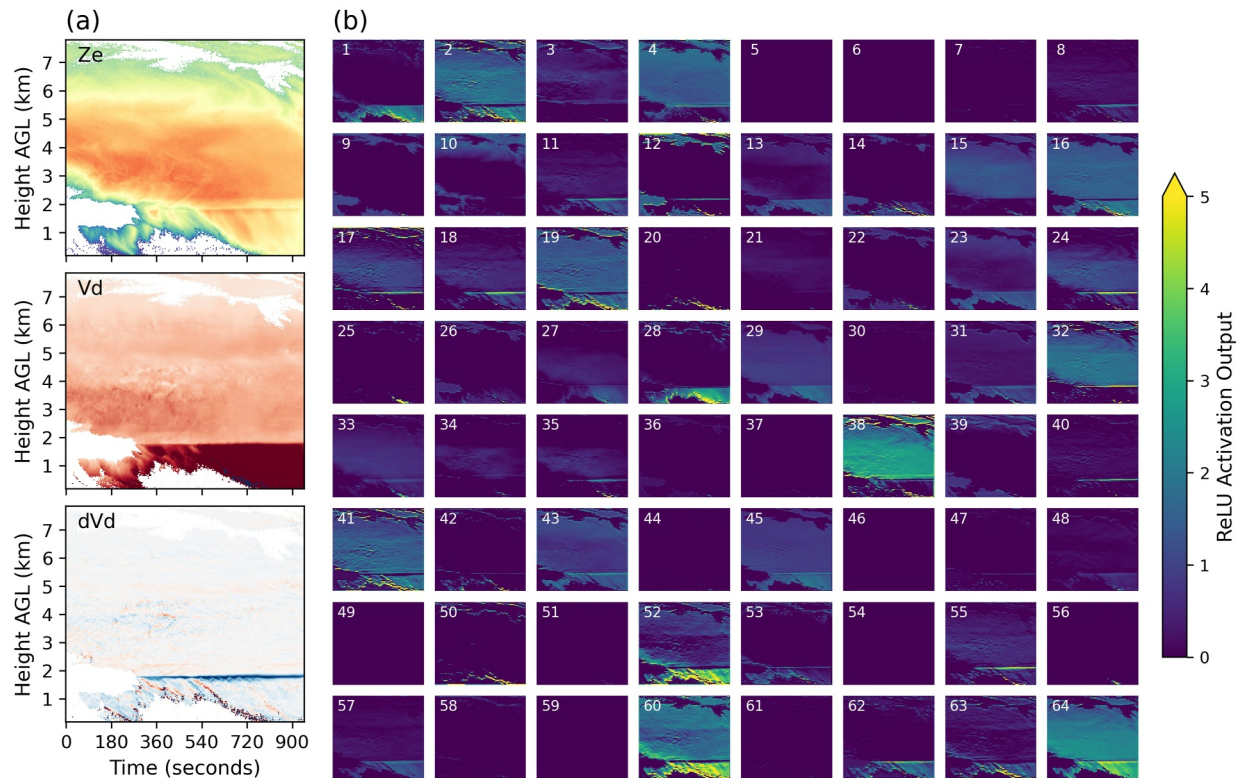


Figure 8. Feature map visualizations. (a) Example standardized input of radar reflectivity (Z_e), Doppler velocity (V_d), and Doppler velocity gradient (dV_d). (b) Corresponding feature map outputs of 64 filters from the ReLU activation function of the first encoder layer in the U-Net model.

frequently across the testing data set. Song et al. (2021) highlighted that the key step to detect melting layers using cloud radar observations is to distinguish melting layers from cloud and rain echoes. Figure 8 indicates that the U-Net model is capable of performing this task by effectively learning the spatial information of clouds, rainfall, and melting layers from the input images of radar observations.

5. Summary and Conclusion

This study presents a U-Net-style machine learning model designed for automatic detection of melting layers from vertically pointing ground radar observations. The U-Net model is trained by KAZR observations under various weather conditions collected from the DOE ARM facility at the NSA site between March 2015 and February 2016. An interactive tool, named ClickCollect, has been developed to generate a binary labeled data set of melting layer boundaries based on radar observations of Doppler velocity. The manually drawn melting layers are applied as the target to train, validate and test the U-Net model. A comparative study has been conducted to evaluate the performance of the U-Net model relative to a traditional method based on empirical thresholds. Results show that the U-Net model outperforms the traditional method and detects the occurrence, boundary heights, and thickness of melting layers accurately and effectively. The main contributions and conclusions of this paper are outlined below.

1. We provided an interactive tool, ClickCollect, which facilitates the process of manually identifying and digitizing melting layer boundaries in radar observations (Section 2.2). This tool is adaptable and can be applied to a wide range of applications that require extraction of data points from certain image features. ClickCollect is an open-source software available for download at our public repository (details in Data Availability Statement).
2. We developed a novel U-Net-style model with carefully tuned hyperparameters for melting layer detection using KAZR observations (Section 3.1.2). The U-Net model addresses the limitations of traditional methods (Brast & Markmann, 2020; Song et al., 2021) and performs better under various weather conditions, including heavy precipitation with velocity folding, concurrent melting layers, and shallow melting layers (Section 4.1).

3. The U-Net model effectively detects 96% of the cases with melting layers. Compared to a traditional method based on empirical thresholds, the U-Net model greatly increases the amount of successfully detected melting layers by 57% and is less biased in the binary segmentation task. Furthermore, the U-Net model shows remarkable proficiency in accurately identifying the location and shape of melting layers, displaying an improvement of more than 18% over the traditional method. The U-Net model performs well in detecting melting layer boundary heights, yet tends to underestimate the melting layer thickness (Section 4.2).
4. The U-Net model provides additional information of pixel-wise detection uncertainty based on ensemble predictions. The uncertainties in predicted probability are higher at the melting layer boundaries, with overall values below 8% (Section 4.1). In addition, explainability analysis indicates that the spatial distribution of melting layers, clouds, and rainfall in the input radar observations plays an essential role in the U-Net model outputs (Section 4.3). The uncertainty estimation and explainability analysis help to enhance the transparency and interpretability of the U-Net model.

Despite its excellent performance in detecting melting layers in most cases, the U-Net model may incorrectly classify stratified cloud structures, likely associated with turbulence or wind shear, as melting layers. It is also challenging for the U-Net model to accurately identify extremely shallow melting layers just above the surface. This issue could be addressed by incorporating more cases with shallow melting layers into the training data set. Additionally, further investigation into comprehensive uncertainty quantification of the U-Net model (Haynes et al., 2023) is necessary to establish an informative uncertainty threshold below which the U-Net detections are considered reliable. Evaluations of the model confidence thresholds could provide valuable insights into the underestimation problem of melting layer thickness, which would improve the reliability of U-Net model output and enhance our understanding of melting layers. The U-Net model developed in this study serves well as a novel U-Net-style machine learning method for melting layer detections. Meanwhile, implementing more advanced U-Net-style neural networks, such as UNet++ (Zhou et al., 2018), UNet3+ (Huang et al., 2020), and Vision Transformers (Thisanke et al., 2023) to explore the influence of model architectures on the accuracy and effectiveness of results represents a promising direction for future improvement.

The melting layer detection framework presented in this study, including the data extraction tool and the U-Net model, are highly adaptable and can be applied to similar profiling radar instruments at sites across the globe. The U-Net is expected to generalize effectively within the Rayleigh scattering regime, that is, at radar frequencies lower than Ka-band. However, at W-band, melting layer signals may be less distinct due to non-Rayleigh scattering effects and rain attenuation (e.g., Heymsfield et al., 2008; Li & Moiseev, 2020). When applied to radar observations at lower latitudes, there is a concern that the U-Net model may underestimate the occurrence frequency and boundary heights of melting layers, considering the impact of latitudinal dependence and local climatology (Heymsfield et al., 2021; Hu et al., 2024; Shates et al., 2023; Song et al., 2021). Therefore, it is necessary to validate the robustness of the U-Net model by conducting additional analyses using observations from different regions, which will be the focus of our future research. As the climate warms, variations of the melting layers remain uncertain, which can lead to extreme rainfall events with enhanced frequency and intensity (Prein & Heymsfield, 2020). The proposed U-Net model demonstrates great potential to enhance our understanding of the global distribution and seasonal variations of melting layers, contributing to improved predictions of their future changes.

Acknowledgments

We thank the Department of Energy ARM program for providing access to the radar data used to develop the model. We gratefully acknowledge the support of the University of Michigan's Great Lakes High Performance Computing (HPC) cluster (<https://its.umich.edu/advanced-research-computing/high-performance-computing/great-lakes>) for providing the computational resources used in this study. We also thank Dr. Julia Shates for suggestions to improve the figures. We thank the three reviewers for their thorough comments and suggestions which help to improve the manuscript. Yan Xie's efforts on this work were supported by the Rackham Predoctoral Fellowship provided by the University of Michigan Rackham Graduate School. Dr. King's contributions were supported by the Natural Sciences and Engineering Research Council of Canada (577912). Dr. Pettersen's contributions were supported by NASA Grant, award 80NSSC22K0789 and NSF Grant, award 2137152. This research was carried out at the University of Michigan, Ann Arbor.

Data Availability Statement

As an open-source project, codes used for the training and testing of the U-Net model are made publicly available via a Zenodo repository at (Xie et al., 2024b; ML2 version 1.0.1). In addition, the interactive data extraction tool ClickCollect is also publicly accessible on Zenodo for others to adapt at (Xie et al., 2024a; ClickCollect version 1.0.1). The DOE ARM ground profiling radar observations by KAZR can be accessed and downloaded from the ARM Data Discovery online archive via (Feng et al., 2011; data product code: kazrge.a1).

References

- Atlas, D., Srivastava, R. C., & Sekhon, R. S. (1973). Doppler radar characteristics of precipitation at vertical incidence. *Reviews of Geophysics*, 11(1), 1–35. <https://doi.org/10.1029/RG011i001p00001>
- Austin, P. M., & Bemis, A. C. (1950). A quantitative study of the "bright band" in radar precipitation echoes. *Journal of the Atmospheric Sciences*, 7(2), 145–151. [https://doi.org/10.1175/1520-0469\(1950\)007<0145:AQSOTB>2.0.CO;2](https://doi.org/10.1175/1520-0469(1950)007<0145:AQSOTB>2.0.CO;2)
- Bennartz, R., Fell, F., Pettersen, C., Shupe, M. D., & Schuettmeyer, D. (2019). Spatial and temporal variability of snowfall over Greenland from CloudSat observations. *Atmospheric Chemistry and Physics*, 19(12), 8101–8121. <https://doi.org/10.5194/acp-19-8101-2019>

- Billault-Roux, A., Grazioli, J., Delanoë, J., Jorquera, S., Pauwels, N., Viltard, N., et al. (2023). ICE GENESIS: Synergetic aircraft and ground-based remote sensing and in situ measurements of snowfall microphysical properties. *Bulletin of the American Meteorological Society*, 104(2), E367–E388. <https://doi.org/10.1175/BAMS-D-21-0184.1>
- Botalb, A., Moinuddin, M., Al-Saggaf, U. M., & Ali, S. S. (2018). Contrasting convolutional neural network (CNN) with multi-layer perceptron (MLP) for big data analysis. In *2018 International conference on intelligent and advanced system (ICIAS)* (pp. 1–5). <https://doi.org/10.1109/ICIAS.2018.8540626>
- Brast, M., & Markmann, P. (2020). Detecting the melting layer with a micro rain radar using a neural network approach. *Atmospheric Measurement Techniques*, 13(12), 6645–6656. <https://doi.org/10.5194/amt-13-6645-2020>
- Camarasa, R., Bos, D., Hendrikse, J., Nederkoorn, P., Kooi, E., Van Der Lugt, A., & De Bruijne, M. (2020). Quantitative comparison of monte-carlo dropout uncertainty measures for multi-class segmentation. *Uncertainty for Safe Utilization of Machine Learning in Medical Imaging, and Graphs in Biomedical Image Analysis*, 12443, 32–41. https://doi.org/10.1007/978-3-030-60365-6_4
- Chandra, A., Zhang, C., Kollias, P., Matrosov, S., & Szyrmer, W. (2015). Automated rain rate estimates using the Ka-band ARM zenith radar (KAZR). *Atmospheric Measurement Techniques*, 8(9), 3685–3699. <https://doi.org/10.5194/amt-8-3685-2015>
- Chase, R. J., Harrison, D. R., Burke, A., Lackmann, G. M., & McGovern, A. (2022). A machine learning tutorial for operational meteorology. Part I: Traditional machine learning. *Weather and Forecasting*, 37(8), 1509–1529. <https://doi.org/10.1175/WAF-D-22-0070.1>
- Chase, R. J., McGovern, A., Homeyer, C. R., Marinescu, P. J., & Potvin, C. K. (2024). Machine learning estimation of maximum vertical velocity from radar. *Artificial Intelligence for the Earth Systems*, 3(2), 230095. <https://doi.org/10.1175/AIES-D-23-0095.1>
- Cooper, S. J., L'Ecuyer, T. S., Wolff, M. A., Kuhn, T., Pettersen, C., Wood, N. B., et al. (2022). Exploring snowfall variability through the high-latitude measurement of snowfall (HiLaMS) field campaign. *Bulletin of the American Meteorological Society*, 103(8), E1762–E1780. <https://doi.org/10.1175/BAMS-D-21-0007.1>
- Doswell, C. A., III, Davies-Jones, R., & Keller, D. L. (1990). On summary measures of skill in rare event forecasting based on contingency tables. *Weather and Forecasting*, 5(4), 576–585. [https://doi.org/10.1175/1520-0434\(1990\)005<0576:OSMOSI>2.0.CO;2](https://doi.org/10.1175/1520-0434(1990)005<0576:OSMOSI>2.0.CO;2)
- Dougherty, E. M., Prein, A. F., Gutmann, E. D., & Newman, A. J. (2023). Future simulated changes in Central U.S. mesoscale convective system rainfall caused by changes in convective and stratiform structure. *Journal of Geophysical Research: Atmospheres*, 128(4), e2022JD037537. <https://doi.org/10.1029/2022JD037537>
- Driss, S. B., Soua, M., Kachouri, R., & Akil, M. (2017). A comparison study between MLP and convolutional neural network models for character recognition. *Real-Time Image and Video Processing 2017*, 10223, 32–42. <https://doi.org/10.1117/12.2262589>
- Fabry, F., & Zawadzki, I. (1995). Long-term radar observations of the melting layer of precipitation and their interpretation. *Journal of the Atmospheric Sciences*, 52(7), 838–851. [https://doi.org/10.1175/1520-0469\(1995\)052<0838:LROOT>2.0.CO;2](https://doi.org/10.1175/1520-0469(1995)052<0838:LROOT>2.0.CO;2)
- Feng, Y.-C., Bharadwaj, N., Lindenmaier, I., Johnson, K., Nelson, D., Isom, B., et al. (2011). Ka ARM Zenith Radar (KAZRGE) [Dataset]. *Atmospheric Radiation Measurement (ARM) User Facility*. <https://doi.org/10.5439/1984772>
- Gal, Y., & Ghahramani, Z. (2016). Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In *International conference on machine learning* (pp. 1050–1059). PMLR. <https://doi.org/10.48550/arXiv.1506.02142>
- Giangrande, S. E., Krause, J. M., & Ryzhkov, A. V. (2008). Automatic designation of the melting layer with a polarimetric prototype of the WSR-88D radar. *Journal of Applied Meteorology and Climatology*, 47(5), 1354–1364. <https://doi.org/10.1175/2007JAMC1634.1>
- Han, L., Liang, H., Chen, H., Zhang, W., & Ge, Y. (2022). Convective precipitation nowcasting using U-net model. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1–8. <https://doi.org/10.1109/TGRS.2021.3100847>
- Haynes, K., Lagerquist, R., McGraw, M., Musgrave, K., & Ebert-Uphoff, I. (2023). Creating and evaluating uncertainty estimates with neural networks for environmental-science applications. *Artificial Intelligence for the Earth Systems*, 2(2), 220061. <https://doi.org/10.1175/AIES-D-22-0061.1>
- Heysfield, A. J., Bansemer, A., Matrosov, S., & Tian, L. (2008). The 94-GHz radar dim band: Relevance to ice cloud properties and CloudSat. *Geophysical Research Letters*, 35(3), L03802. <https://doi.org/10.1029/2007GL031361>
- Heysfield, A. J., Bansemer, A., Theis, A., & Schmitt, C. (2021). Survival of snow in the melting layer: Relative humidity influence. *Journal of the Atmospheric Sciences*, 78(6), 1823–1845. <https://doi.org/10.1175/JAS-D-20-0353.1>
- Honari, H., Choe, A. S., Pekar, J. J., & Lindquist, M. A. (2019). Investigating the impact of autocorrelation on time-varying connectivity. *NeuroImage*, 197, 37–48. <https://doi.org/10.1016/j.neuroimage.2019.04.042>
- Houze, R. A., Jr. (1997). Stratiform precipitation in regions of convection: A meteorological paradox? *Bulletin of the American Meteorological Society*, 78(10), 2179–2196. [https://doi.org/10.1175/1520-0477\(1997\)078<2179:SPIROC>2.0.CO;2](https://doi.org/10.1175/1520-0477(1997)078<2179:SPIROC>2.0.CO;2)
- Hu, X., Ai, W., Qiao, J., & Yan, W. (2024). Insight into global climatology of melting layer: Latitudinal dependence and orographic relief. *Theoretical and Applied Climatology*, 155(6), 4863–4873. <https://doi.org/10.1007/s00704-024-04926-6>
- Huang, H., Lin, L., Tong, R., Hu, H., Zhang, Q., Iwamoto, Y., et al. (2020). Unet 3+: A full-scale connected unet for medical image segmentation. *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1055–1059. <https://doi.org/10.1109/ICASSP40776.2020.9053405>
- Jadon, S. (2020). A survey of loss functions for semantic segmentation. *2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, 1–7. <https://doi.org/10.1109/CIBCB48159.2020.9277638>
- Jay Miller, L., Mohr, C. G., & Weinheimer, A. J. (1986). The simple rectification to cartesian space of folded radial velocities from Doppler radar sampling. *Journal of Atmospheric and Oceanic Technology*, 3(1), 162–174. [https://doi.org/10.1175/1520-0426\(1986\)003<0162:TSRTCS>2.0.CO;2](https://doi.org/10.1175/1520-0426(1986)003<0162:TSRTCS>2.0.CO;2)
- King, F., Duffy, G., Milani, L., Fletcher, C. G., Pettersen, C., & Ebell, K. (2022). DeepPrecip: A deep neural network for precipitation retrievals. *Atmospheric Measurement Techniques*, 15(20), 6035–6050. <https://doi.org/10.5194/amt-15-6035-2022>
- King, F., Pettersen, C., Fletcher, C. G., & Geiss, A. (2024). Development of a full-scale connected U-net for reflectivity inpainting in spaceborne radar blind zones. *Artificial Intelligence for the Earth Systems*, 3(2), e230063. <https://doi.org/10.1175/AIES-D-23-0063.1>
- Klaassen, W. (1988). Radar observations and simulation of the melting layer of precipitation. *Journal of the Atmospheric Sciences*, 45(24), 3741–3753. [https://doi.org/10.1175/1520-0469\(1988\)045<3741:ROASOT>2.0.CO;2](https://doi.org/10.1175/1520-0469(1988)045<3741:ROASOT>2.0.CO;2)
- Kollias, P., Bharadwaj, N., Clothiaux, E. E., Lamer, K., Oue, M., Hardin, J., et al. (2020). The ARM radar network: At the leading edge of cloud and precipitation observations. *Bulletin of the American Meteorological Society*, 101(5), E588–E607. <https://doi.org/10.1175/BAMS-D-18-0288.1>
- Kumjian, M. R., Mishra, S., Giangrande, S. E., Toto, T., Ryzhkov, A. V., & Bansemer, A. (2016). Polarimetric radar and aircraft observations of saggy bright bands during MC3E. *Journal of Geophysical Research: Atmospheres*, 121(7), 3584–3607. <https://doi.org/10.1002/2015JD024446>
- Lackmann, G. M., Keeter, K., Lee, L. G., & Ek, M. B. (2002). Model representation of freezing and melting precipitation: Implications for winter weather forecasting. *Weather and Forecasting*, 17(5), 1016–1033. [https://doi.org/10.1175/1520-0434\(2003\)017<1016:MROFAM>2.0.CO;2](https://doi.org/10.1175/1520-0434(2003)017<1016:MROFAM>2.0.CO;2)

- Le, M., & Chandrasekar, V. (2013). Precipitation type classification method for dual-frequency precipitation radar (DPR) onboard the GPM. *IEEE Transactions on Geoscience and Remote Sensing*, 51(3), 1784–1790. <https://doi.org/10.1109/TGRS.2012.2205698>
- Li, H., & Moiseev, D. (2020). Two layers of melting ice particles within a single radar bright band: Interpretation and implications. *Geophysical Research Letters*, 47(13), e2020GL087499. <https://doi.org/10.1029/2020GL087499>
- Maahn, M., Burgard, C., Crewell, S., Gorodetskaya, I. V., Kneifel, S., Lhermitte, S., et al. (2014). How does the spaceborne radar blind zone affect derived surface snowfall statistics in polar regions? *Journal of Geophysical Research: Atmospheres*, 119(13), 13604–13620. <https://doi.org/10.1002/2014JD022079>
- Matrosov, S. Y. (2008). Assessment of radar signal attenuation caused by the melting hydrometeor layer. *IEEE Transactions on Geoscience and Remote Sensing*, 46(4), 1039–1047. <https://doi.org/10.1109/TGRS.2008.915757>
- Matrosov, S. Y., Clark, K. A., & Kingsmill, D. E. (2007). A polarimetric radar approach to identify rain, melting-layer, and snow regions for applying corrections to vertical profiles of reflectivity. *Journal of Applied Meteorology and Climatology*, 46(2), 154–166. <https://doi.org/10.1175/JAM2508.1>
- Mazurowski, M. A., Habas, P. A., Zurada, J. M., Lo, J. Y., Baker, J. A., & Tourassi, G. D. (2008). Training neural network classifiers for medical decision making: The effects of imbalanced datasets on classification performance. *Neural Networks*, 21(2–3), 427–436. <https://doi.org/10.1016/j.neunet.2007.12.031>
- Milnenstädt, J., Sourdeval, O., Delanoë, J., & Quaas, J. (2015). Frequency of occurrence of rain from liquid-mixed-and ice-phase clouds derived from A-Train satellite retrievals. *Geophysical Research Letters*, 42(15), 6502–6509. <https://doi.org/10.1002/2015GL064604>
- Pettersen, C., Kulie, M. S., Bliven, L. F., Merrelli, A. J., Petersen, W. A., Wagner, T. J., et al. (2020). A composite analysis of snowfall modes from four winter seasons in Marquette, Michigan. *Journal of Applied Meteorology and Climatology*, 59(1), 103–124. <https://doi.org/10.1175/JAMC-D-19-0099.1>
- Prechelt, L. (1998). Automatic early stopping using cross validation: Quantifying the criteria. *Neural Networks*, 11(4), 761–767. [https://doi.org/10.1016/S0893-6080\(98\)00010-0](https://doi.org/10.1016/S0893-6080(98)00010-0)
- Prein, A. F., & Heymsfield, A. J. (2020). Increased melting level height impacts surface precipitation phase and intensity. *Nature Climate Change*, 10(8), 771–776. <https://doi.org/10.1038/s41558-020-0825-x>
- Pujol, O., Mesnard, F., & Sauvageot, H. (2012). Effects of melting layer in airborne meteorological X-band radar observations. *IEEE Transactions on Geoscience and Remote Sensing*, 50(6), 2318–2324. <https://doi.org/10.1109/TGRS.2011.2172111>
- Ren, Y., Li, X., Yang, X., & Xu, H. (2022). Development of a dual-attention U-net model for sea ice and open water classification on SAR images. *IEEE Geoscience and Remote Sensing Letters*, 19, 1–5. <https://doi.org/10.1109/LGRS.2021.3058049>
- Roebber, P. J. (2009). Visualizing multiple measures of forecast quality. *Weather and Forecasting*, 24(2), 601–608. <https://doi.org/10.1175/2008WAF2222159.1>
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. *Medical image computing and computer-assisted intervention—MICCAI 2015: 18th International Conference, Munich, Germany, Proceedings, part III*, 18, 234–241. <https://doi.org/10.48550/arXiv.1505.04597>
- Sassen, K., Matrosov, S., & Campbell, J. (2007). CloudSat spaceborne 94 GHz radar bright bands in the melting layer: An attenuation-driven upside-down lidar analog. *Geophysical Research Letters*, 34(16), L16818. <https://doi.org/10.1029/2007GL030291>
- Schumacher, C., & Funk, A. (2023). Assessing convective-stratiform precipitation regimes in the tropics and extratropics with the GPM satellite radar. *Geophysical Research Letters*, 50(14), e2023GL102786. <https://doi.org/10.1029/2023GL102786>
- Shates, J. A., Pettersen, C., L'Ecuier, T. S., & Kulie, M. S. (2023). Multi-year analysis of rain-snow levels at Marquette, Michigan. *Journal of Geophysical Research: Atmospheres*, 128(2), e2022JD037132. <https://doi.org/10.1029/2022JD037132>
- Shin, D., North, G. R., & Bowman, K. P. (2000). A summary of reflectivity profiles from the first year of TRMM radar data. *Journal of Climate*, 13(23), 4072–4086. [https://doi.org/10.1175/1520-0442\(2000\)013<4072:ASORPF>2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013<4072:ASORPF>2.0.CO;2)
- Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, 6(60), 60. <https://doi.org/10.1186/s40537-019-0197-0>
- Siddique, N., Paheding, S., Elkin, C. P., & Devabhaktuni, V. (2021). U-Net and its variants for medical image segmentation: A review of theory and applications. *IEEE Access*, 9, 82031–82057. <https://doi.org/10.1109/ACCESS.2021.3086020>
- Song, J. I., Yum, S. S., Park, S.-H., Kim, K.-H., Park, K.-J., & Joo, S.-W. (2021). Climatology of melting layer heights estimated from cloud radar observations at various locations. *Journal of Geophysical Research: Atmospheres*, 126(17), e2021JD034816. <https://doi.org/10.1029/2021JD034816>
- Taghanaki, S. A., Zheng, Y., Zhou, S. K., Georgescu, B., Sharma, P., Xu, D., et al. (2019). Combo loss: Handling input and output imbalance in multi-organ segmentation. *Computerized Medical Imaging and Graphics*, 75, 24–33. <https://doi.org/10.1016/j.compmedimag.2019.04.005>
- Tao, W., Lang, S., Zeng, X., Shige, S., & Takayabu, Y. (2010). Relating convective and stratiform rain to latent heating. *Journal of Climate*, 23(7), 1874–1893. <https://doi.org/10.1175/2009JCLI3278.1>
- Thisanake, H., Deshan, C., Chamith, K., Seneviratne, S., Vidanaarachchi, R., & Herath, D. (2023). Semantic segmentation using vision transformers: A survey. *Engineering Applications of Artificial Intelligence*, 126, 106669. <https://doi.org/10.1016/j.engappai.2023.106669>
- Turner, R., Eriksson, D., McCourt, M., Kiili, J., Laaksonen, E., Xu, Z., & Guyon, I. (2021). Bayesian optimization is superior to random search for machine learning hyperparameter tuning: Analysis of the black-box optimization challenge 2020. *Proceedings of the NeurIPS 2020 Competition and Demonstration Track, PMLR*, 133, 3–26. <https://doi.org/10.48550/arXiv.2104.10201>
- Ustuner, M., Sanli, F. B., & Abdikan, S. (2016). Balanced VS imbalanced training data: Classifying rapidEye data with support vector machines. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 379–384. <https://doi.org/10.5194/isprs-archives-XLI-B7-379-2016>
- White, A. B., Gottas, D. J., Strem, E. T., Ralph, F. M., & Neiman, P. J. (2002). An automated brightband height detection algorithm for use with Doppler radar spectral moments. *Journal of Atmospheric and Oceanic Technology*, 19(5), 687–697. [https://doi.org/10.1175/1520-0426\(2002\)019<0687:AABHDA>2.0.CO;2](https://doi.org/10.1175/1520-0426(2002)019<0687:AABHDA>2.0.CO;2)
- Williams, C. R., Ecklund, W. L., & Gage, K. S. (1995). Classification of precipitating clouds in the tropics using 915-MHz wind profilers. *Journal of Atmospheric and Oceanic Technology*, 12(5), 996–1012. [https://doi.org/10.1175/1520-0426\(1995\)012<0996:COPCIT>2.0.CO;2](https://doi.org/10.1175/1520-0426(1995)012<0996:COPCIT>2.0.CO;2)
- Willis, P. T., & Heymsfield, A. J. (1989). Structure of the melting layer in mesoscale convective system stratiform precipitation. *Journal of the Atmospheric Sciences*, 46(13), 2008–2025. [https://doi.org/10.1175/1520-0469\(1989\)046<2008:SOTMLI>2.0.CO;2](https://doi.org/10.1175/1520-0469(1989)046<2008:SOTMLI>2.0.CO;2)
- Wu, J., Chen, X., Zhang, H., Xiong, L., Lei, H., & Deng, S. (2019). Hyperparameter optimization for machine learning models based on bayesian optimization b. *Journal of Electronic Science and Technology*, 17(1), 26–40. <https://doi.org/10.11989/JEST.1674-862X.80904120>
- Wu, S., Zhang, H., Valiant, G., & Re, C. (2020). On the generalization effects of linear transformations in data augmentation. *Proceedings of the 37th International Conference on Machine Learning, in Proceedings of Machine Learning Research*. <https://doi.org/10.48550/arXiv.2005.00695>

- Xie, Y., King, F., Pettersen, C., & Flanner, M. (2024a). Machine learning detection of melting layers from radar observations. Jan-16-2025, ClickCollect v1.0.1 [Software]. Zenodo. <https://doi.org/10.5281/zenodo.14675535>
- Xie, Y., King, F., Pettersen, C., & Flanner, M. (2024b). Machine learning detection of melting layers from radar observations. Jan-20-2025, ML2 v1.0.1 [Software]. Zenodo. <https://doi.org/10.5281/zenodo.14691918>
- Xie, Y., Pettersen, C., Flanner, M., & Shates, J. (2024c). Ground-observed snow albedo changes during rain-on-snow events in northern Alaska. *Journal of Geophysical Research: Atmospheres*, 129(16), e2024JD040975. <https://doi.org/10.1029/2024JD040975>
- Ying, X. (2019). An overview of overfitting and its solutions. *Journal of Physics: Conference Series*, 1168(2), 022022. <https://doi.org/10.1088/1742-6596/1168/2/022022>
- Zhang, J., Langston, C., & Howard, K. (2008). Brightband identification based on vertical profiles of reflectivity from the WSR-88D. *Journal of Atmospheric and Oceanic Technology*, 25(10), 1859–1872. <https://doi.org/10.1175/2008JTECHA1039.1>
- Zhang, W., Karhu, S. I., & Salonen, E. T. (1994). Predictions of radiowave attenuations due to a melting layer of precipitation. *IEEE Transactions on Antennas and Propagation*, 42(4), 492–500. <https://doi.org/10.1109/8.286217>
- Zhao, W., Jiang, D., Queralta, J. P., & Westerlund, T. (2020). MSS U-Net: 3D segmentation of kidneys and tumors from CT images with a multi-scale supervised U-Net. *Informatics in Medicine Unlocked*, 19, 100357. <https://doi.org/10.1016/j.imu.2020.100357>
- Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., & Liang, J. (2018). UNet++: A nested U-Net architecture for medical image segmentation. *4th Deep Learning in Medical Image Analysis (DLMIA) Workshop*, 3–11. <https://doi.org/10.48550/arXiv.1807.10165>