

---

# IceCloudNet: Cirrus and mixed-phase cloud prediction from SEVIRI input learned from sparse supervision

---

Kai Jeggle<sup>1,2,\*</sup>, Mikolaj Czerkawski<sup>2</sup>, Federico Serva<sup>3,2</sup>,  
Bertrand Le Saux<sup>2</sup>, David Neubauer<sup>1</sup>, and Ulrike Lohmann<sup>1</sup>

<sup>1</sup>Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland

<sup>2</sup>Φ-lab, European Space Agency (ESA), Frascati, Italy

<sup>3</sup>Consiglio Nazionale delle Ricerche - Istituto di Scienze Marine (CNR-ISMAR), Rome, Italy

\*Corresponding author: Kai Jeggle, kai.jeggle@env.ethz.ch

## Abstract

Clouds containing ice particles play a crucial role in the climate system. Yet they remain a source of great uncertainty in climate models and future climate projections. In this work, we create a new observational constraint of regime-dependent ice microphysical properties at the spatio-temporal coverage of geostationary satellite instruments and the quality of active satellite retrievals. We achieve this by training a convolutional neural network on three years of SEVIRI and DARDAR data sets. This work will enable novel research to improve ice cloud process understanding and hence, reduce uncertainties in a changing climate and help assess geoengineering methods for cirrus clouds.

## 1 Introduction

**Clouds containing ice** particles cover 22 % of Earth’s surface at any moment [9]. They modulate incoming and outgoing radiation [14] and contribute to the majority of global precipitation [18]. Yet, we lack understanding about cloud formation and evolution which leads to large uncertainties in climate projections. Clouds containing ice particles can be distinguished into cirrus clouds and mixed-phase clouds. Cirrus clouds contain only ice crystals, are typically optically thin, occur at high altitudes at temperatures below  $-38^{\circ}\text{C}$  [20], and have on average a warming effect [9]. Mixed-phase clouds contain a mix of ice crystals and supercooled cloud droplets. They are thicker and usually exert a cooling effect [11]. Key to process understanding are properties such as ice water content (IWC) and ice water path (IWP). The former is a measure of the ice mass per unit volume and is vertically resolved while the latter is its integration along a vertical column. Ice properties are projected to change in a warming climate and may amplify or dampen global warming [16].

**Current observational constraints** are mainly two types of satellite retrievals. On the one hand, multiple studies e.g. [13, 21, 10] have used polar-orbiting active satellite instruments like CALIPSO’s lidar [23] and CloudSat’s radar [22] to analyze microphysical properties of ice clouds. These instruments retrieve cloud properties by emitting electromagnetic waves and measuring the backscattered signal. The main benefit is that active instruments are able to provide a vertical profile of cloud structures. Due to the narrow swath of the satellite overpass CALIPSO and CloudSat have a revisiting time of 16 days which is a much coarser temporal scale than necessary for studying clouds where processes act between seconds to hours [17]. In this work, we will use the DARDAR [3] data set which combines CALIPSO and CloudSat retrievals. On the other hand, passive geostationary satellite instruments such as SEVIRI onboard the Meteosat satellite [2] retrieve a top-down 2D view of Earth’s

surface every 15 minutes by measuring the reflected solar radiation and intensities of terrestrial infrared radiation. Previous studies have combined geostationary and actively sensed retrievals for IWP predictions [1, 12], but the context of cloud regime was not considered unlike in this work.

**The objective** here is to provide regime-dependent IWP with high spatio-temporal coverage. To this end, we train a convolutional neural network (CNN) that predicts IWP for cirrus and mixed-phase clouds from SEVIRI data which is supervised by co-located IWP retrievals from DARDAR. Unlike previous work, the model proposed here provides insight into both ice regimes independently.

## 2 Dataset

The dataset used for this study contains three years of multi-spectral images from SEVIRI and vertically-resolved DARDAR IWC swaths. We identify the matching DARDAR overpasses for every SEVIRI image and resample them to the native SEVIRI horizontal resolution of 3km x 3km. Due to the high temporal resolution of 15 minutes and the long revisiting times of DARDAR, any SEVIRI image contains at most one DARDAR overpass, resulting in a minor fraction of a given SEVIRI image that can be co-located with DARDAR data. Note, that many SEVIRI images do not contain any DARDAR overpass. Figure 1 visualizes an exemplary SEVIRI image with its matching DARDAR swath. Instead of using the whole vertical column of DARDAR IWC as target data, we integrate the IWC for cirrus and mixed-phase regimes resulting in an ice-regime dependent IWP. With this approach, we keep a key aspect of vertical cloud structure and ice distribution through the atmosphere while reducing the output from 419 vertical levels to 2. We chose a domain from 30°W to 30°E and 30°N to 30°N, resulting in 1984×1792 pixels per SEVIRI image in its native resolution. For training and validating the neural network, we create non-overlapping patches of size 64×64 and select only the patches that contain co-located DARDAR data, resulting in 160,137 patches for the years 2007-2009. Formally, our dataset is described as:

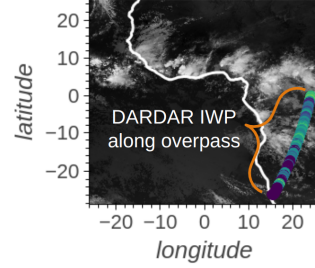


Figure 1: 10μm channel for single SEVIRI observation from 2007-10-24 00:12:43 with co-located DARDAR IWP. Note that the DARDAR swath is magnified for better visibility.

- $X \in \mathbb{R}^{160137 \times 64 \times 64 \times 9}$  for the nine infrared channels of SEVIRI. Three visible channels are omitted as they are only available at daytime.
- $Y_{\text{sparse}} \in \mathbb{R}^{160137 \times 64 \times 64 \times 2}$  for the IWP of cirrus and mixed-phase regime derived from DARDAR ice water content. Note, that while being on a 64×64 grid, only a small subset of the grid points contain data, i.e. along the DARDAR overpass.
- $M = \mathbb{Z}_2^{160137 \times 64 \times 64 \times 1}$  for the binary overpass mask indicating the location of the satellite overpass.

A considerable challenge from the machine learning perspective is the sparsity of the target data. Yet, our model, once trained and validated, is able to produce IWP predictions for cirrus and mixed-phase clouds at high spatio-temporal coverage for all available SEVIRI images, resulting in full image predictions for 9.1 million patches of size 64×64.

## 3 Methodology

### 3.1 Problem setting

The outlined task can be formally described as learning a mapping  $f : \mathcal{X} = \mathbb{R}^{64 \times 64 \times 9} \rightarrow \mathcal{Y} = \mathbb{R}^{64 \times 64 \times 2}$ , where  $x \in \mathcal{X}$  is representing SEVIRI channels and  $y \in \mathcal{Y}$  the IWP of cirrus and mixed-phase clouds. Note that at training time, a prediction  $\hat{y} \in \mathcal{Y}$  is masked to a narrow swath using the corresponding overpass mask  $m \in M$  and the loss  $\mathcal{L}(\hat{y}_{\text{sparse}}, y_{\text{sparse}})$  is calculated on the sparse data only.

### 3.2 IceCloudNet architecture

The backbone of IceCloudNet is a U-Net [19] architecture made up of ConvNeXt [15] blocks. ConvNeXt blocks are state-of-the-art convolutional modules based on the ResNet [7] architecture, improved with multiple macro and micro design choices inspired by transformer models, such as ViT [5]. Figure 2 visualizes the processing pipeline, where SEVIRI input data  $x \in X$  is fed into the network predicting the values of cirrus and mixed-phase IWP  $\hat{y} \in \mathcal{Y}$ . The orange lines in the predictions show the DARDAR overpass mask  $m \in M$ , i.e. where target data for supervision is available. Despite being trained only on sparse target data, the model learns to predict the full spatial image, expanding the spatio-temporal coverage of cirrus and mixed-phase IWP considerably.

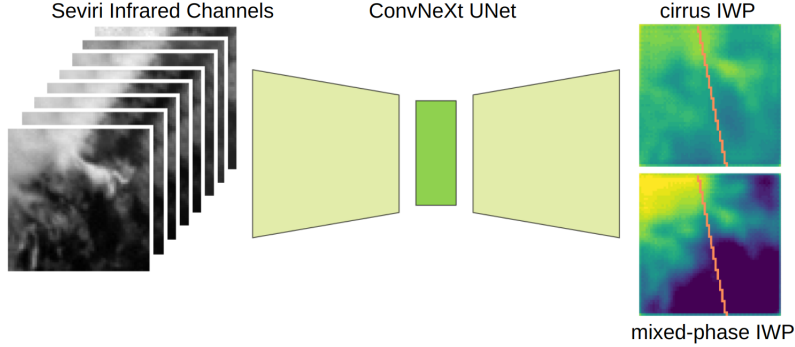


Figure 2: IceCloudNet architecture with sample inputs and predictions. At training, ground truth target data are narrow DARDAR IWP swaths shown as orange lines which are the only supervision for the network. At inference, dense images of the size of the SEVIRI inputs are predicted.

### 3.3 Experimental setup

We split the dataset into train, validation, and test splits in a 70%, 20%, 10% proportion. In order to prevent spatio-temporal auto-correlation in the splits, data from the same day will be assigned to only one split. We transform  $Y_{\text{sparse}}$  logarithmically with base 10 and normalize  $X$  with the training set statistics. We randomly apply rotations of  $90^\circ$  multiples for data augmentation and train IceCloudNet for 100 epochs and batch size 32 using the Adam optimizer and learning rate of  $10^{-6}$ .

## 4 Results

We compare the predictive performance of IceCloudNet with two baseline models: linear regression and gradient boosted regression trees (XGBoost) [4]. Unlike CNNs, these models are not capitalizing on the spatial structure of the input data, but are predicting the cloud properties pixel by pixel.

Table 1: IceCloudNet performance on independent test set. Note that MAE and correlation are computed only on available DARDAR ground-truth and are calculated on log transformed target variable.

	Cloud regime	MAE ↓	Pearson Correlation ↑	Accuracy (%) ↑
<b>Linear regression</b>	cirrus	0.83	0.75	77
	mixed-phase	0.66	0.75	84
<b>XGBoost</b>	cirrus	0.74	0.75	77
	mixed-phase	0.60	0.78	84
<b>IceCloudNet</b>	cirrus	0.49	0.82	86
	mixed-phase	0.47	0.83	88

Regression and classification metrics can only be calculated where DARDAR ground truth data is available. The performance on the test set is reported in Table 1. The accuracy is calculated based on a post-processed cloud mask where a pixel contains a cloud if the IWP  $> 10^{-5} \text{ kg m}^{-2}$ . IceCloudNet outperforms the baseline models as it is able to learn from the spatial structure in the input data. Figure 3 illustrates the predicted and target variables along the satellite overpass for a single patch

(a,b) as well as the ground truth IWC (c) from which the ice regime dependent IWP is derived. All ice above the height of the  $-38^{\circ}\text{C}$  isotherm (horizontal line in panel (c)) is in the cirrus regime, everything below in the mixed-phase regime. Despite the complex multi-layer cloud scene, IceCloudNet is able to separate the cirrus and mixed-phase regimes and successfully quantify the IWP for both regimes for the majority of the scene. More samples are shown in Appendix A.

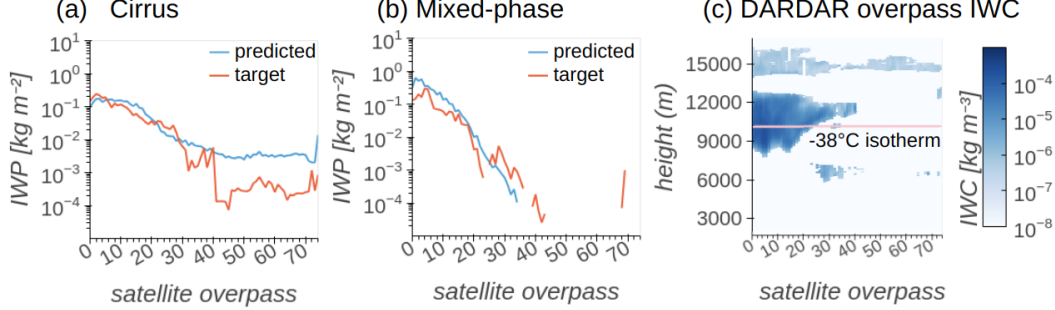


Figure 3: Sample predictions and ground truth along overpass for single patch shown in Figure 2. Panel (c) visualizes vertically resolved IWC ground truth and the  $-38^{\circ}\text{C}$  isotherm retrieved from ERA5 [8]. The x-axis is an incremental counter for pixels along the overpass.

In Figure 4, we show the prediction of IWP (a-b) of a full SEVIRI observation. Additionally, the post-processed cloud mask is shown in (c), which allows studying cirrus cloud origin [6]. Figure 1 displays SEVIRI input and the available DARDAR ground truth for this observation. With IceCloudNet we are able to expand the information of ice-regime dependent IWP from a narrow swath to the full image. We note that adding geographical information via a topographic map did not improve the prediction performance.

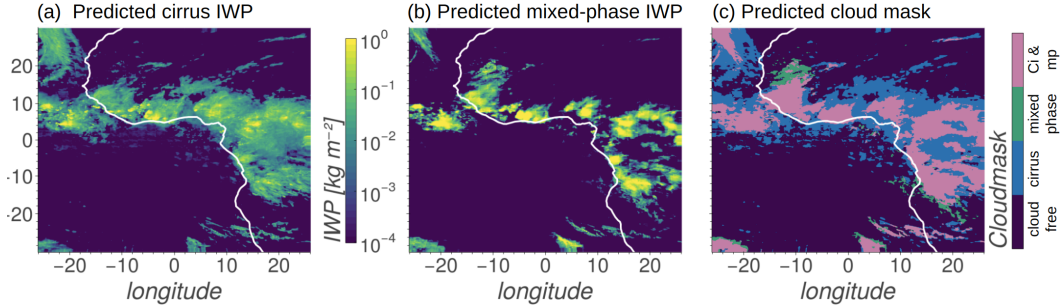


Figure 4: Sample prediction of IceCloudNet for cirrus (a) and mixed phase (b) IWP and cloud mask (c). One channel of the SEVIRI input for this prediction is shown in 1

## 5 Conclusions

We introduce a new way to obtain high-quality predictions for microphysical properties of cirrus and mixed-phase clouds with high spatio-temporal coverage. Trained on geostationary SEVIRI data and retrievals of actively measured DARDAR data, our machine learning based approach allows to supply the community with a new observational constraint that will enable novel research on ice cloud formation and improve understanding of the microphysical process by tracking and studying cloud properties through time and space, even beyond the lifetime of recently-ended satellite missions underlying DARDAR. We show that IceCloudNet can learn from sparse data and significantly outperforms baseline models. New findings enabled by our work will help to improve climate models reduce climate projection uncertainty and help assess the risk of geoengineering methods.

## **Acknowledgments and Disclosure of Funding**

This research was supported by grants from the European Union’s Horizon 2020 research and innovation program iMIRACLI under Marie Skłodowska-Curie grant agreement No 860100 and Swiss National Supercomputing Centre (Centro Svizzero di Calcolo Scientifico, CSCS; project ID s1144). KJ is grateful for the opportunity to being hosted as a visiting researcher at ESA  $\Phi$ -lab during spring 2023 which led to the initiation of this research project. The authors thank Peter Naylor for his ideas on improving the model architecture and three anonymous reviewers for their suggestions and constructive feedback.

## References

- [1] Adrià Amell, Patrick Eriksson, and Simon Pfreundschuh. Ice water path retrievals from Meteosat-9 using quantile regression neural networks. *Atmospheric Measurement Techniques*, 15:5701–5717, October 2022.
- [2] D M A Aminou. MSG’s SEVIRI instrument. *ESA Bulletin*(0376-4265), (111):15–17, 2002.
- [3] Quitterie Cazenave, Marie Ceccaldi, Julien Delanoë, Jacques Pelon, Silke Groß, and Andrew Heymsfield. Evolution of DARDAR-CLOUD ice cloud retrievals: new parameters and impacts on the retrieved microphysical properties. *Atmospheric Measurement Techniques*, 12(5):2819–2835, May 2019. Publisher: Copernicus GmbH.
- [4] Tianqi Chen and Carlos Guestrin. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, August 2016.
- [5] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, and Neil Houlsby. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale, June 2021. arXiv:2010.11929 [cs] version: 2.
- [6] B. Gasparini, A. Meyer, D. Neubauer, S. Münch, and U. Lohmann. Cirrus Cloud Properties as Seen by the CALIPSO Satellite and ECHAM-HAM Global Climate Model. *Journal of Climate*, 31(5):1983–2003, March 2018.
- [7] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep Residual Learning for Image Recognition, December 2015. arXiv:1512.03385 [cs].
- [8] Hans Hersbach, Bill Bell, Paul Berrisford, András Horányi, Joaquín Muñoz-Sabater, Julien Nicolas, Carole Peubey, Raluca Radu, Dinand Schepers, Adrian Simmons, Cornel Soci, Gionata Biavati, Dick Dee, and Jean-Noël Thépaut. ERA5 hourly data on pressure levels from 1959 to present. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*, 2018.
- [9] Andrew J. Heymsfield, Martina Krämer, Anna Luebke, Phil Brown, Daniel J. Cziczo, Charmaine Franklin, Paul Lawson, Ulrike Lohmann, Greg McFarquhar, Zbigniew Ulanowski, and Kristof Van Tricht. Cirrus Clouds. *Meteorological Monographs*, 58:2.1–2.26, January 2017.
- [10] Kai Jeggle, David Neubauer, Gustau Camps-Valls, and Ulrike Lohmann. Understanding cirrus clouds using explainable machine learning. *Environmental Data Science*, 2:e19, January 2023.
- [11] A. Korolev, G. McFarquhar, P. R. Field, C. Franklin, P. Lawson, Z. Wang, E. Williams, S. J. Abel, D. Axisa, S. Borrmann, J. Crosier, J. Fugal, M. Krämer, U. Lohmann, O. Schlenczek, M. Schnaiter, and M. Wendisch. Mixed-Phase Clouds: Progress and Challenges. *Meteorological Monographs*, 58(1):5.1–5.50, January 2017. Publisher: American Meteorological Society Section: Meteorological Monographs.
- [12] S. Kox, L. Bugliaro, and A. Ostler. Retrieval of cirrus cloud optical thickness and top altitude from geostationary remote sensing. *Atmospheric Measurement Techniques*, 7(10):3233–3246, October 2014. Publisher: Copernicus GmbH.
- [13] Martina Krämer, Christian Rolf, Anna Luebke, Armin Afchine, Nicole Spelten, Anja Costa, Jessica Meyer, Martin Zöger, Jessica Smith, Robert L. Herman, Bernhard Buchholz, Volker Ebert, Darrel Baumgardner, Stephan Borrmann, Marcus Klingebiel, and Linnea Avallone. A microphysics guide to cirrus clouds – Part 1: Cirrus types. *Atmospheric Chemistry and Physics*, 16(5):3463–3483, March 2016.
- [14] Kuo-Nan Liou. Influence of Cirrus Clouds on Weather and Climate Processes: A Global Perspective. *Monthly Weather Review*, 114(6):1167–1199, June 1986.
- [15] Zhuang Liu, Hanzi Mao, Chao-Yuan Wu, Christoph Feichtenhofer, Trevor Darrell, and Saining Xie. A ConvNet for the 2020s, March 2022. arXiv:2201.03545 [cs].

- [16] Ulrike Lohmann and David Neubauer. The importance of mixed-phase and ice clouds for climate sensitivity in the global aerosol–climate model ECHAM6-HAM2. *Atmospheric Chemistry and Physics*, 18(12):8807–8828, June 2018. Publisher: Copernicus GmbH.
- [17] Chunsong Lu, Yangang Liu, Bin Zhu, Seong Soo Yum, Steven K. Krueger, Yujun Qiu, Shengjie Niu, and Shi Luo. On Which Microphysical Time Scales to Use in Studies of Entrainment-Mixing Mechanisms in Clouds. *Journal of Geophysical Research: Atmospheres*, 123(7):3740–3756, 2018. \_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/2017JD027985>.
- [18] Johannes Mülmenstädt, O. Sourdeval, J. Delanoë, and J. Quaas. 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, 2015. \_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/2015GL064604>.
- [19] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation, May 2015. arXiv:1505.04597 [cs].
- [20] Kenneth Sassen, Zhien Wang, and Dong Liu. Global distribution of cirrus clouds from CloudSat/Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) measurements. *J. Geophys. Res. Atmos.*, 113(D8), 2008.
- [21] Odran Sourdeval, Edward Gryspeerdt, Johannes Mülmenstädt, Martina Krämer, and Johannes Quaas. Satellite-based estimate of the climate forcing due to aerosol - ice cloud interactions. Technical Report EGU2020-17971, Copernicus Meetings, March 2020. Conference Name: EGU2020.
- [22] Graeme L. Stephens, Deborah G. Vane, Ronald J. Boain, Gerald G. Mace, Kenneth Sassen, Zhien Wang, Anthony J. Illingworth, Ewan J. O’connor, William B. Rossow, Stephen L. Durden, Steven D. Miller, Richard T. Austin, Angela Benedetti, and Cristian Mitrescu. THE CLOUDSAT MISSION AND THE A-TRAIN: A New Dimension of Space-Based Observations of Clouds and Precipitation. *Bulletin of the American Meteorological Society*, 83(12):1771–1790, December 2002. Publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society.
- [23] David M Winker, Chris A Hostetler, Mark A Vaughan, and Ali H Omar. [CALIOP Mission Overview] Part 1 : CALIOP Instrument, and Algorithms Overview. page 29, 2009.

## A Prediction samples

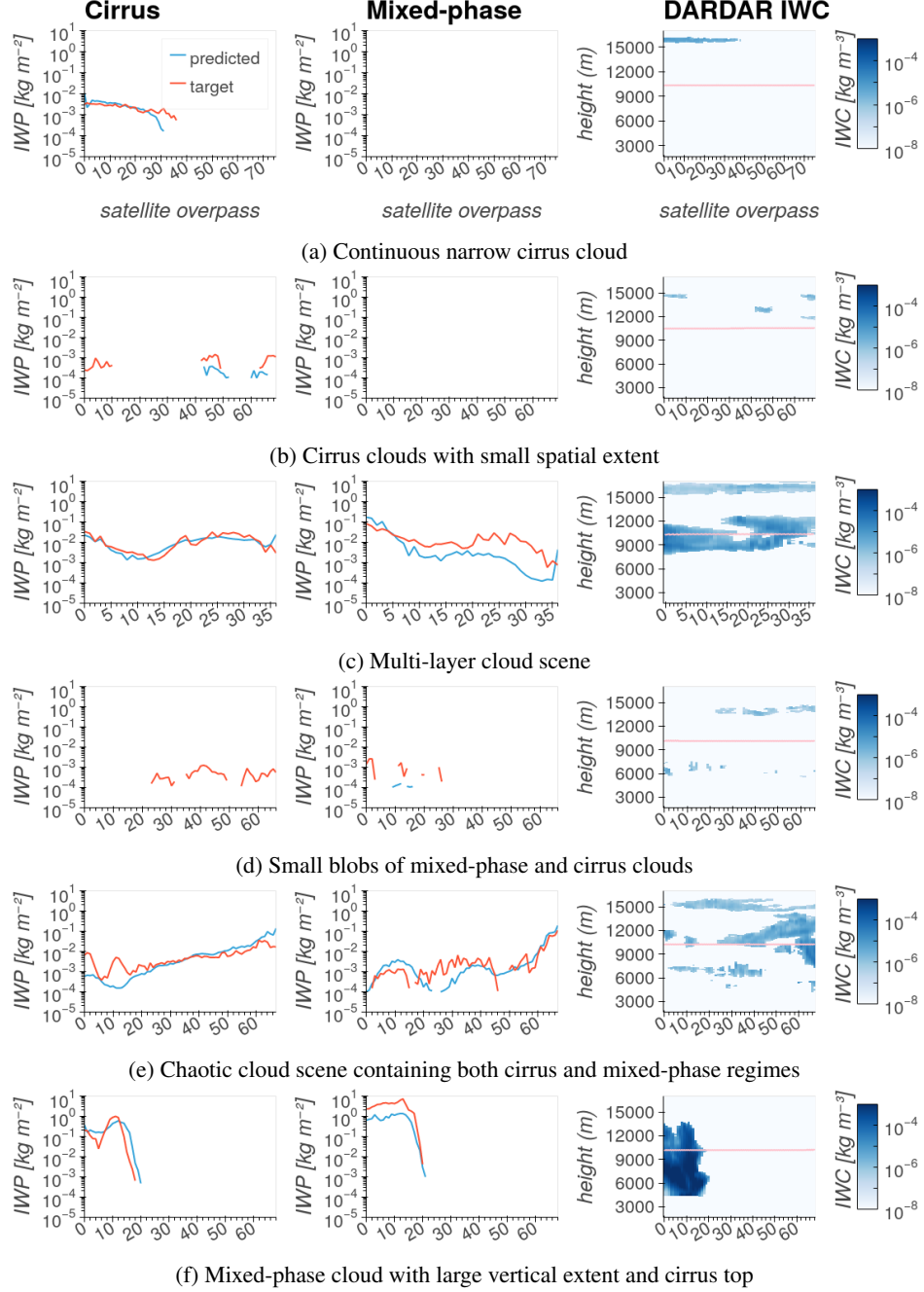


Figure A1: Sample predictions of IceCloudNet (blue lines) and DARDAR ground truth (red lines) for cirrus (left) and mixed-phase (center) IWP. Each row represents a random sample from the test set. The right plot in each row shows the DARDAR IWC ground truth along the overpass, i.e. the vertically resolved variable and is displayed to provide insight into the structure of the cloud scene. The horizontal orange line represents the ERA5 derived -38°C isotherm which acts as a border between cirrus and mixed-phase regimes. The x-axis an incremental counter for pixels along the overpass which can vary depending on the overpass angle.

Figure A1 shows IceCloudNet predictions for randomly sampled patches of the test set. From qualitative inspection, we note that IceCloudNet is able to detect and quantify regime-dependent IWP for larger cloud structures (e.g. a,c,f), even if the scene is very chaotic (e.g. d), but may struggle on



clouds with small horizontal extent (e.g. b,d). From a climate impact perspective, small clouds have minor relevance compared to large cloud systems. Nonetheless, we aim to improve the performance of IceCloudNet on smaller clouds in future versions.