Detecting Floods from Cloudy Scenes: A Fusion Approach Using Sentinel-1 and Sentinel-2 Imagery

Qiuyang Chen
School of GeoSciences
University of Edinburgh
Q.Chen-43@sms.ed.ac.uk

Xenofon Karagiannis
Earth-i Ltd.
xenofon.karagiannis@earthi.co.uk

Simon M. Mudd
School of GeoSciences
University of Edinburgh
simon.m.mudd@ed.ac.uk

Abstract

As the result of climate change, extreme flood events are becoming more frequent. To better respond to such disasters, and to test and calibrate flood models, we need accurate real-world data on flooding extent. Detection of floods from remote sensed imagery suffers from a widespread problem: clouds block flood scenes in images, leading to degraded and fragmented flood datasets. To address this challenge, we propose a workflow based on U-Net, and a dataset that detects flood in cloud-prone areas by fusing information from the Sentinel-1 and Sentinel-2 satellites. The expected result will be a reliable and detailed catalogue of flood extents and how they change through time, allowing us to better understand flooding in different morphological settings and climates.

1 Overview

Floods are the most frequent and widespread natural disasters that can cause more than $40 billion in damage each year across the globe. Remote sensing images are commonly used in flood detection to map flood extent and indicate the probability of flood recurrence. In recent years, the Sentinel-1 (S1) and Sentinel-2 (S2) missions have received great interest for their potential to generate global flooding maps accurately from freely accessible data. The S1 Synthetic Aperture Radar constellation provides all-weather day-and-night imagery. S1's shortcoming is that it detects water by identifying a smooth water surface, meaning that flood mapping under conditions that increase water roughness, such as windy and vegetated areas, becomes a challenge. The S2 mission can detect open water using the near-infrared band. However, it can only observe floods during the daytime and in cloudless areas.

Extracting flood extents from S1 and S2 imagery traditionally uses thresholding to differentiate water from other pixels. This often introduces classification errors and requires manual adjustment. Recent advances show potential for the application of deep learning algorithms, especially convolutional neural networks (CNN), for enhancing flood mapping. This leads to our research question: Can deep neural networks be used in flood detection to leverage the complementary information from a fusion of data from the S1 and S2 sensors?

Our project builds on a preliminary exploration using U-Net models for delineating the extent of floods. The preliminary results show that S1 data can supplement S2 data by improving flood mapping results to almost perfect accuracy in cloudless scenes. We further identify the most challenging task in flood mapping and detection: how to use deep learning algorithms to learn from S1 data when S2

Tackling Climate Change with Machine Learning: workshop at NeurIPS 2022.

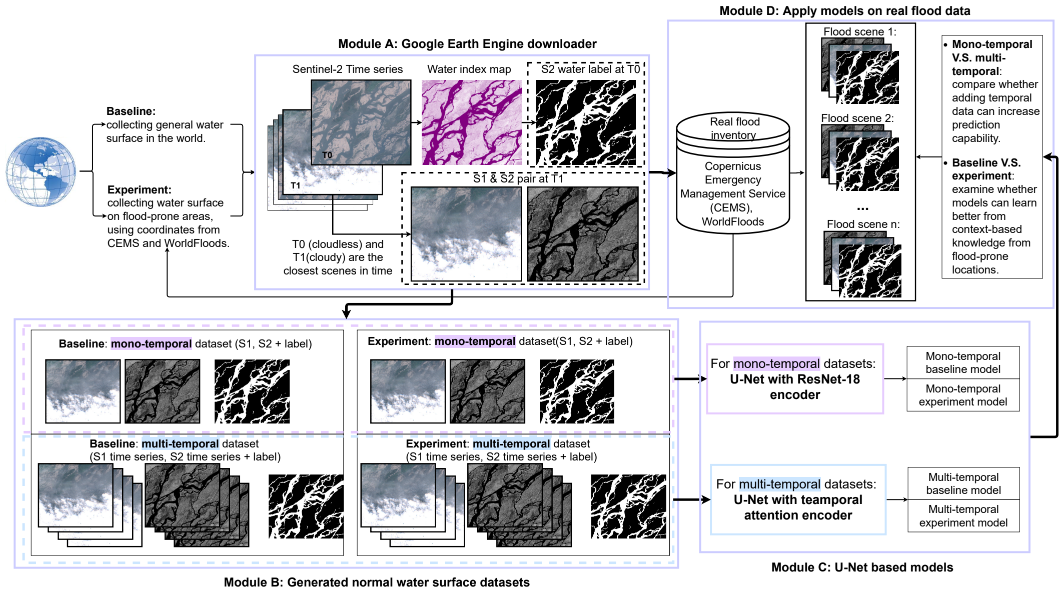

Figure 1: Overview of proposed workflow. Module A: generating normal water surface dataset with cloudless S2 water index label at $T0$ and cloudy S1 and S2 image pair at $T1$. Module B: generating normal water surface datasets. Module C: to train two U-Net based models on generated normal surface water datasets. Module D: Applying fine-tuned models of normal surface water to real flood scenes.

images are obscured by clouds. Our proposal intends to address this problem in two parts: *a.* We propose a dataset that can provide accurate S2 generated water surface labels for both S1 and S2 image pairs, including both cloudy and cloudless scenes. *b.* We intend to use this dataset to train U-Net based models and then test the applicability of the trained models to real flooding events.

2 Previous Work

Tasks. Cloudy images inhibit accurate flood detection and delineation (see summary in Figure 2; Appendix A.1). In optical imagery (S2), floods, as a type of open water surface, can be extracted either by calculating water indexes using water-sensitive spectral bands [1, 2, 3, 4] or by CNNs [5, 6]. While S2 imagery is a highly effective data source for flood detection, cloud cover in S2 imagery cannot be mitigated by a traditional cloud removal approach. Traditional cloud removal selects the closest (in time) images with cloudless conditions to restore information in the target cloudy scenes [7, 8]. However, this is not suitable for detecting a short-lived event such as a flood which is likely to have many pixels covered in cloud. Although not influenced by clouds, S1 imagery alone struggles to differentiate between water and non-water pixels on major areas of interest (especially urban areas) in flood mapping. Using deep learning to enhance the prediction of S1 imagery remains an open challenge (with competitions held by NASA [9] and Microsoft [10]). One possible reason of the low accuracy from S1-based methods lies in the training label quality. Labels, typically generated by thresholding S2 water indices, inherit the cloudy pixels in S2 images as "no data" [11, 10, 12]. Thus clouds in optical images can propagate to machine learning tasks that use only S1 images. Alternatively, a multi-sensor approach selects similar pixels from different sensors at the same time, which has the potential to reconstruct missing information covered by clouds. This approach is successfully adopted by [13, 14] using deep neural networks to fuse S1 and S2 imagery.

Datasets. For the purpose of training deep learning models to detect floods with fused S1 and S2 data, a small dataset **Sen1Floods11** can be potentially used for our proposed research. The dataset provides mono-temporal S1 and S2 image pairs sampled within two days from 11 flood events at a global scale, providing semi-manually annotated labels where each pixel is classified as either water, no water or no data (cloud-covered pixels in S2). In addition to a small number of manual labels, the dataset additionally includes a greater number of weakly labelled masks generated from S1 thresholding. Other large-scale datasets are not directly available for training a multi-model machine learning architecture, but can serve as flood event inventory data, which can be further developed for our purpose. **WorldFloods** [6] provides a global-scale flood dataset covering 119 floods events between 2015 and 2019 derived from S2 imagery. **Copernicus Emergency Management Service** (CEMS) provides accurate manual labels of delineated flood events by field experts and S1-based fully automated flood maps.

Models. In recent years, U-Net architecture [15] has been intensively used for water segmentation tasks in a variety of studies [6, 16, 17, 18]. For flood mapping specifically, modified U-Net models are used to delineate accurate boundary between permanent water and temporary flood extent [19] and to detect flood from multiple sensors [20]. ResNet [21] is a popular backbone in many flood segmentation tasks for scene understanding [22, 11, 12] by enhancing residual learning between shallow and deeper layers of neural networks. Comparisons of different CNN versions for flood detections suggest model selection has a small influence on accuracy, suggesting efforts should be concentrated on improving training data [12]. To mitigate the scarcity and poor quality of manual labels, a new trend of deep learning in flood detection is to develop semi-supervised [18] and self-supervised [23] models. They are still in the early stage of development, showing an attempt to derive values from automated labels.

3 Proposed Method

The missing piece of the puzzle: a data fusion of S1 and S2. Given a set of accurate labels of water extent, could a deep neural network learn from S1 imagery and then transfer that knowledge to cloud-covered pixels in S2 imagery? The study of [20], using complete (that is, with the cloud-covered area filled) flood labels, shows a supervised multi-modal CNN can effectively transfer the learnt feature from S1 to S2 pixels under clouds. As shown in Figure 2 in Appendix, for semantic segmentation tasks in mono-temporal flood scenes, flood detection consists of separating permanent water surfaces from flood water.

With time series data, flood detection is done through change detection of water surfaces: in an ideal scenario, we can extract accurate, cloud-free water surfaces from each time point and delineate flood extent by differencing the extracted water masks from pre- and post-event imagery. Are there any existing datasets to support such research? No. We propose to automate production of this data.

The proposed research aims at examining the transferability of U-Net based models trained on a large-scale water dataset collected during cloud-free periods (which will mostly during non-flooded periods) to detecting real flood scenes. The overall workflow is shown in Figure 3 in Appendix A.2. To compensate for the imbalance and scarcity of flood water labels, we first propose a multi-temporal bi-modal dataset of normal water surfaces consisting of S1 and S2 image pairs and complete labels (i.e., with no missing data). Second, we propose U-Net based models to train S1 and S2 image pairs with cloudy scenes on complete labels, and then fine tune the trained model on real flood events data from CEMS dataset. Our proposal builds on the assumption that the knowledge learnt from normal water surfaces can be transferred to flood surfaces, which is supported by [12] using small-scale data in China and Peru.

Preliminary Results. We trained a U-Net model with ResNet-18 as encoder on S1 and S2 imagery as model input, using labels from Sen1Floods11 dataset. We compared the results of the U-Net model trained on S1 and S2 with those of the model trained only on S2 data. We also used the same performance metrics to calculate the accuracy of the provided weak S1 labels in Sen1Floods11 dataset to provide a baseline. The comparison is shown in Table 1 and Figure 2 in Appendix A.1. We found multi-modal U-Net architecture can detect better flood extent in cloudless scenarios.

Table 1: U-Net metrics for mono-temporal imagery, compared with S1 Otsu thresholding labels

Metrics	U-Net on S1, S2	U-Net on S2	Otsu on S1
Precision	**0.971**	0.965	0.916
Mean IoU	**0.861**	0.836	0.696

Table 2: UTAE metrics for multi-temporal bi-modal imagery

Metrics	Macro
Precision	0.873
IoU	0.791

In order to explore whether multi-temporal information can improve model performance in cloudy areas, we enriched the temporal data of Sen1Floods11 dataset, extending the mono-temporal image pairs on the flooding day to multi-temporal data from one month before the event. The new dataset is generated by Google Earth Engine Python API and trained with a U-Net with Temporal Attention Encoder (UTAE, adapted from [24]) to extract spatial-temporal features. Without any tuning on the model, the overall performance of the UTAE is shown in Table 2. We found that by simply feeding multi-temporal data, the model cannot effectively learn information from cloudy pixels, although the position of these pixels are changing in the temporal stack. We believe the missing data from labels in Sen1Floods11 dataset contribute to this issue.

Proposed Dataset. To generate high-quality water surface labels, we apply Otsu thresholding on water index (e.g., MNDWI) of cloudless S2 images at reference date $T0$. Then we find cloudy scenes at $T1$ that are closest to $T0$, deriving S1 and S2 image pairs during cloudy scene. We assume that for normal water surfaces, the extent of water does not change significantly during a short period time, which can potentially be confirmed by more frequent rainfall data. The difference of the water surface on $T0$ and $T1$ will be manually checked using Planet Scope 3 m resolution imagery. By doing so, we can generate a global scale dataset with complete water labels and a pair of S1 and S2 raw images.

Proposed Model:

1. Baseline: no preference on location of dataset collection. Dataset will be collected on general water surface in the world. With the generated dataset, we will apply U-Net with ResNet-18 encoder to mono-temporal S1 and S2 pairs, and apply UTAE model to multi-temporal image pairs. Two differently encoded U-Net models will be able to compare the usefulness of multi-temporal data in segmenting water surface during cloudy scenes.
2. Experiment: only collect dataset on flood-prone areas. We will use the polygons in existing flood datasets (using long-term flood risk map, or past flooded locations from WorldFloods and CEMS datasets) to locate these areas, and then generate water surface image pairs and labels from non-flooding days. This experiment is designed to examine whether the models can learn context-base knowledge from flood-prone locations. The following treatments of collected data using basic U-Net and UTAE models are the same as the proposed baseline architecture.

4 Pathway to Tackle Climate Change

With a wealth of satellite imagery acquired in the past few decades, only a subset is usable in flood detection due to the limitations of clouds. According to a study collecting information on all flood events in Europe from 2014-2021 [25], on average 58% of flood events are potentially observable by S1 and only 28% by S2 due to cloud coverage. With the proposed method, we can unlock the potential to re-use cloud-polluted images and reliably, accurately and automatically recover flood extent from them. This method can considerably expand the size, quality and temporal resolution of the catalogue of flood extents for a range of flood events, including extreme flood events that are becoming more frequent and are the biggest threat to people and property as a result of climate change.

References

[1] B. Gao, “Ndwi—a normalized difference water index for remote sensing of vegetation liquid water from space,” Remote Sensing of Environment, vol. 58, no. 3, pp. 257–266, 1996.

[2] H. Xu, “Modification of normalised difference water index (ndwi) to enhance open water features in remotely sensed imagery,” International Journal of Remote Sensing, vol. 27, no. 14, pp. 3025–3033, 2006.

[3] J.-F. Pekel, C. Vancutsem, L. Bastin, M. Clerici, E. Vanbogaert, E. Bartholomé, and P. Defourny, “A near real-time water surface detection method based on hsv transformation of modis multi-spectral time series data,” Remote Sensing of Environment, vol. 140, pp. 704–716, 2014.

[4] A. Goffi, D. Stroppiana, P. A. Brivio, G. Bordogna, and M. Boschetti, “Towards an automated approach to map flooded areas from Sentinel-2 MSI data and soft integration of water spectral

features," International Journal of Applied Earth Observation and Geoinformation, vol. 84, p. 101951, Feb. 2020.

[5] T. James, C. Schillaci, and A. Lipani, "Convolutional neural networks for water segmentation using sentinel-2 red, green, blue (rgb) composites and derived spectral indices," International Journal of Remote Sensing, vol. 42, no. 14, pp. 5338–5365, 2021.

[6] G. Mateo-Garcia, J. Veitch-Michaelis, L. Smith, S. V. Oprea, G. Schumann, Y. Gal, A. G. Baydin, and D. Backes, "Towards global flood mapping onboard low cost satellites with machine learning," Scientific Reports, vol. 11, p. 7249, Mar. 2021.

[7] F. Ramoino, F. Tutunaru, F. Pera, and O. Arino, "Ten-meter sentinel-2a cloud-free composite—southern africa 2016," Remote Sensing, vol. 9, p. 652, 07 2017.

[8] D.-C. Tseng, H.-T. Tseng, and C.-L. Chien, "Automatic cloud removal from multi-temporal spot images," Applied Mathematics and Computation, vol. 205, no. 2, pp. 584–600, 2008. Special Issue on Advanced Intelligent Computing Theory and Methodology in Applied Mathematics and Computation.

[9] NASA-IMPACT, "ETCI 2021 Competition on Flood Detection." `https://nasa-impact.github.io/etci2021/`[Accessed: 2022-09-14].

[10] DrivenData, "STAC Overflow: Map Floodwater from Radar Imagery." `https://www.drivendata.org/competitions/81/detect-flood-water/`[Accessed: 2022-09-14].

[11] D. Bonafilia, B. Tellman, T. Anderson, and E. Issenberg, "Sen1floods11: a georeferenced dataset to train and test deep learning flood algorithms for sentinel-1," in 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 835–845, 2020.

[12] M. Helleis, M. Wieland, C. Krullikowski, S. Martinis, and S. Plank, "Sentinel-1-based water and flood mapping: Benchmarking convolutional neural networks against an operational rule-based processing chain," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 15, pp. 2023–2036, 2022.

[13] A. Meraner, P. Ebel, X. X. Zhu, and M. Schmitt, "Cloud removal in sentinel-2 imagery using a deep residual neural network and sar-optical data fusion," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 166, pp. 333–346, 2020.

[14] J. Li, C. Li, W. Xu, H. Feng, F. Zhao, H. Long, Y. Meng, W. Chen, H. Yang, and G. Yang, "Fusion of optical and sar images based on deep learning to reconstruct vegetation ndvi time series in cloud-prone regions," International Journal of Applied Earth Observation and Geoinformation, vol. 112, p. 102818, 2022.

[15] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in International Conference on Medical image computing and computer-assisted intervention, pp. 234–241, Springer, 2015.

[16] V. Katiyar, N. Tamkuan, and M. Nagai, "Near-real-time flood mapping using off-the-shelf models with sar imagery and deep learning," Remote Sensing, vol. 13, no. 12, p. 2334, 2021.

[17] E. Nemni, J. Bullock, S. Belabbes, and L. Bromley, "Fully convolutional neural network for rapid flood segmentation in synthetic aperture radar imagery," Remote Sensing, vol. 12, no. 16, p. 2532, 2020.

[18] S. Paul and S. Ganju, "Flood Segmentation on Sentinel-1 SAR Imagery with Semi-Supervised Learning," arXiv e-prints, p. arXiv:2107.08369, July 2021.

[19] Y. Bai, W. Wu, Z. Yang, J. Yu, B. Zhao, X. Liu, H. Yang, E. Mas, and S. Koshimura, "Enhancement of detecting permanent water and temporary water in flood disasters by fusing sentinel-1 and sentinel-2 imagery using deep learning algorithms: Demonstration of sen1floods11 benchmark datasets," Remote Sensing, vol. 13, no. 11, 2021.

[20] G. I. Drakonakis, G. Tsagkatakis, K. Fotiadou, and P. Tsakalides, “Ombrianet—supervised flood mapping via convolutional neural networks using multitemporal sentinel-1 and sentinel-2 data fusion,” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 15, pp. 2341–2356, 2022.

[21] J. Long, E. Shelhamer, and T. Darrell, “Fully convolutional networks for semantic segmentation,” 2014.

[22] C. Rambour, N. Audebert, E. Koeniguer, B. Le Saux, M. Crucianu, and M. Datcu, “Flood detection in time series of optical and sar images,” The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XLIII-B2-2020, pp. 1343–1346, 2020.

[23] K. A. Islam, M. S. Uddin, C. Kwan, and J. Li, “Flood detection using multi-modal and multi-temporal images: A comparative study,” Remote Sensing, vol. 12, no. 15, 2020.

[24] V. S. F. Garnot and L. Landrieu, “Panoptic segmentation of satellite image time series with convolutional temporal attention networks,” 2021.

[25] A. Tarpanelli, A. C. Mondini, and S. Camici, “Effectiveness of sentinel-1 and sentinel-2 for flood detection assessment in europe,” Natural Hazards and Earth System Sciences, vol. 22, no. 8, pp. 2473–2489, 2022.

A Appendix

A.1 Overview of the problem

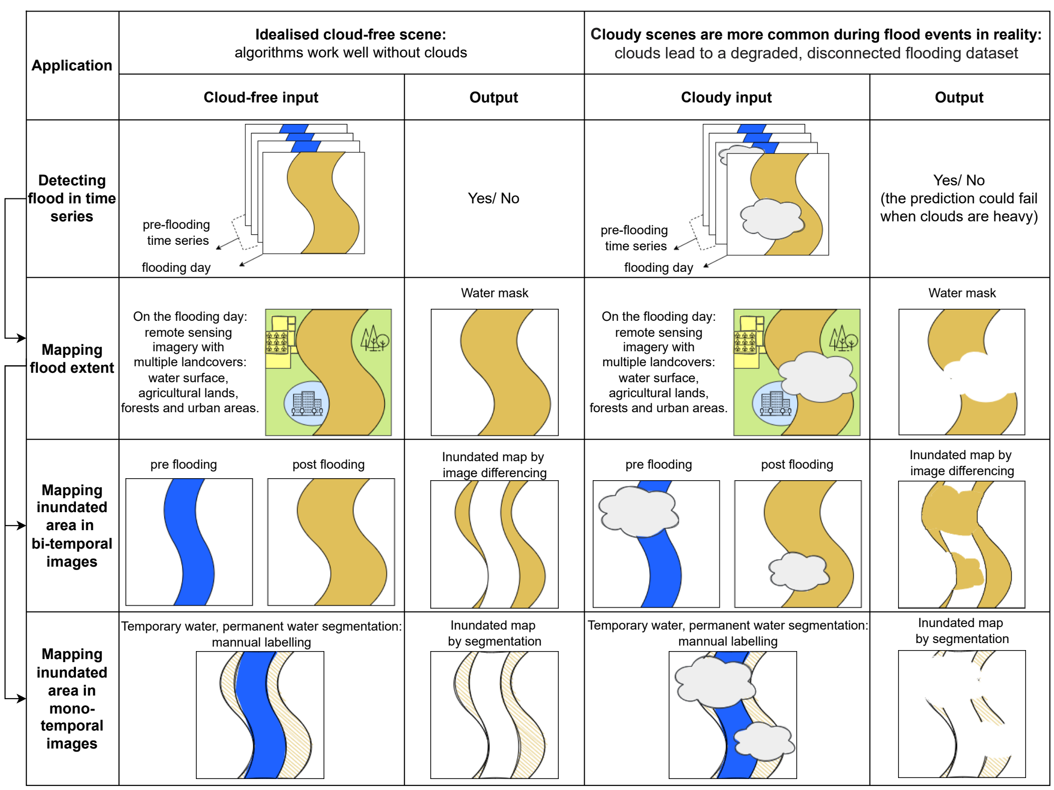

Figure 2: A diagram of pipelines for common flood detection applications and a comparison of how the existence of clouds challenges in those applications.

A.2 Examples showing problematic labels and their influence on U-Net predictions

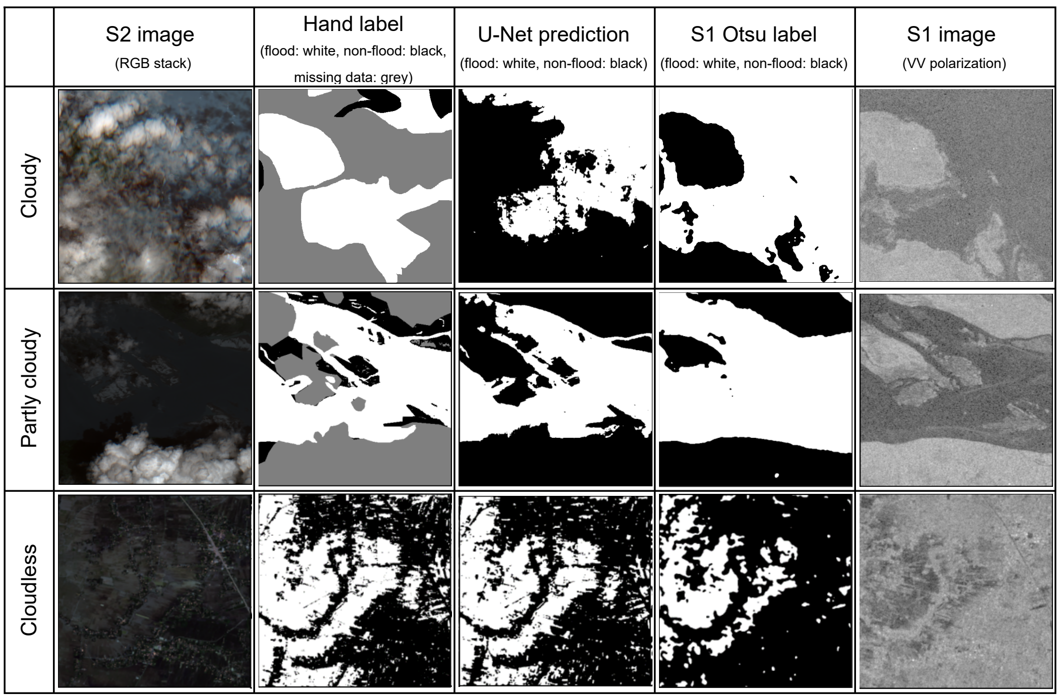

Figure 3: Examples of U-Net predictions on S1 and S2 image pairs, comparing with hand labels and S1 Ostu thresholding labels from Sen1Floods11 dataset.