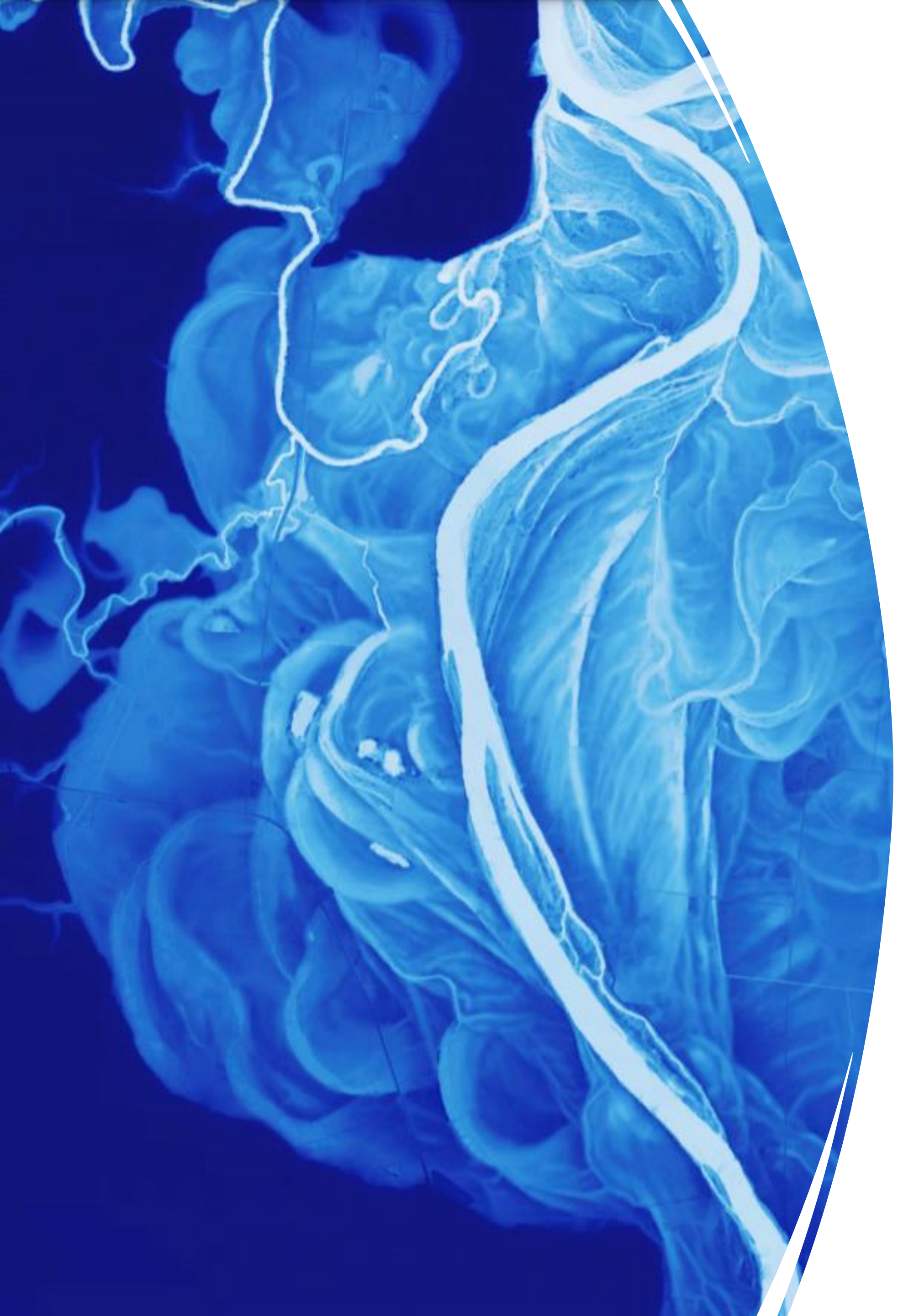


Detecting Floods from Cloudy Scenes: A Fusion Approach Using Sentinel-1 and Sentinel-2 Imagery

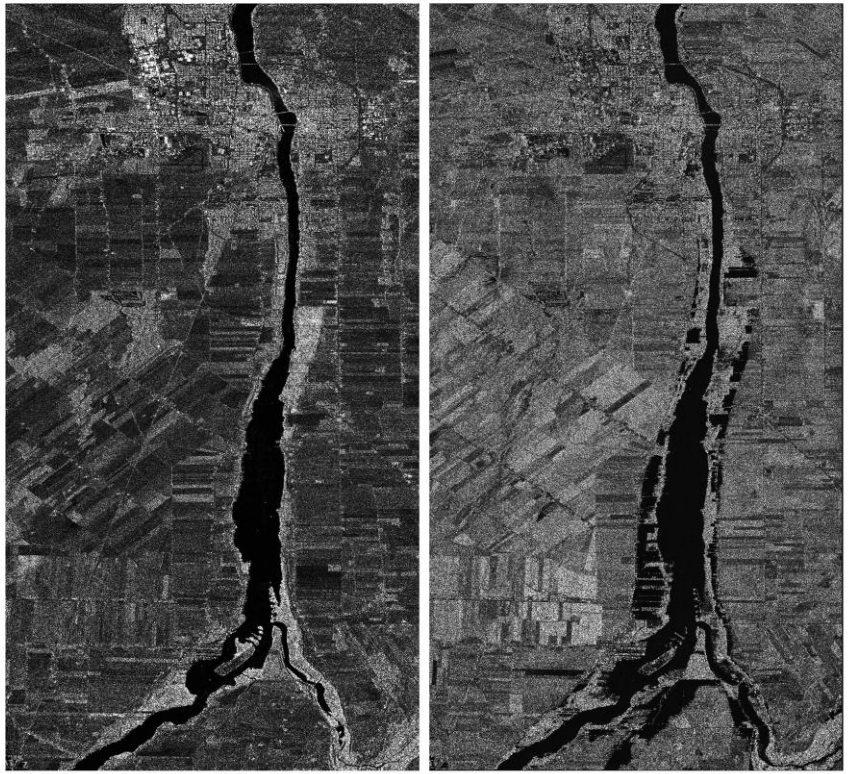
Qiuyang Chen – University of Edinburgh

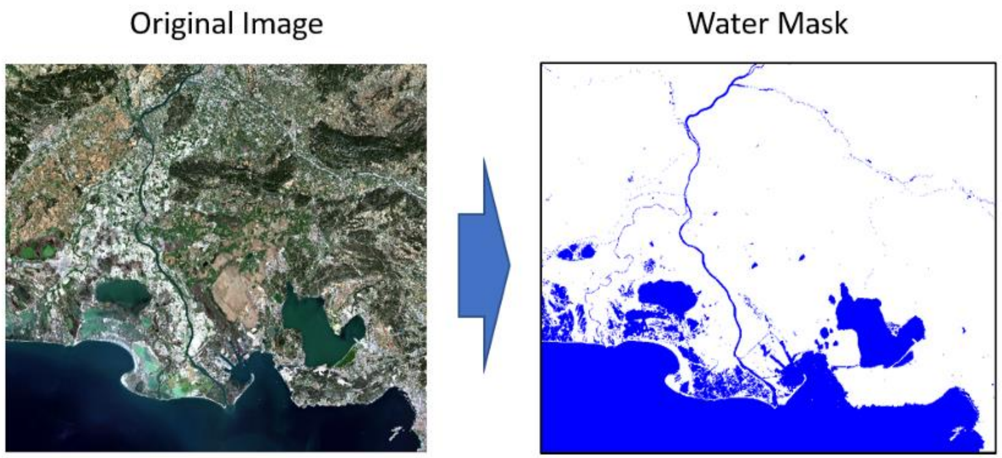
Xenofon Karagiannis – EARTH-i LTD

Simon M. Mudd – University of Edinburgh

Mapping accurate inundated areas matters

- Evaluate and quantify losses.
- Constrain flood models.
- Make better rescue plans to save life and properties.





Sensor Overview

Sentinel-1 (S1) mission:

- C band SAR constellation of two satellites
- All-weather day-and-night supply of imagery with 10 m of spatial resolution and 6-day revisit time.
- Smooth water surface shows low backscatter values in Sentinel-1 imagery.
- **Limit**: It struggles detecting flood under windy conditions or the presence of vegetation.

Sentinel-2 (S2) mission:

- Satellites with optical sensors
- 13 spectral bands at 10 - 60 m spatial resolution.
- It can detect flood surface by thresholding water indices (e.g., NDWI, MNDWI).
- **Limit**: during flooding events it is very like to be blocked by clouds.

Can deep neural networks leverage the complementary information from a fusion of data from the S1 and S2 sensors?

Challenges:

- **Cloud cover:** For all flood events in Europe from 2014: "*On average the 58 % of flood events are potentially observable by Sentinel-1 and only the 28 % by Sentinel-2 due to the cloud coverage.*" (Tarpanelli, et al., 2022)
- **Sensor fusion**: Overlapping of two missions at the same location is possible, but not very often! Spatial resolution of bands are not always the same.
- **Available datasets + labels**:
  - **Sen1Floods11**: 11 global flood events, semi-manually annotated labels where each pixel is classified as either water, no water or no data (cloud-covered pixels in S2).
  - **WorldFloods**: global-scale flood dataset covering 119 flood events from 2015 to 2019 from only 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.

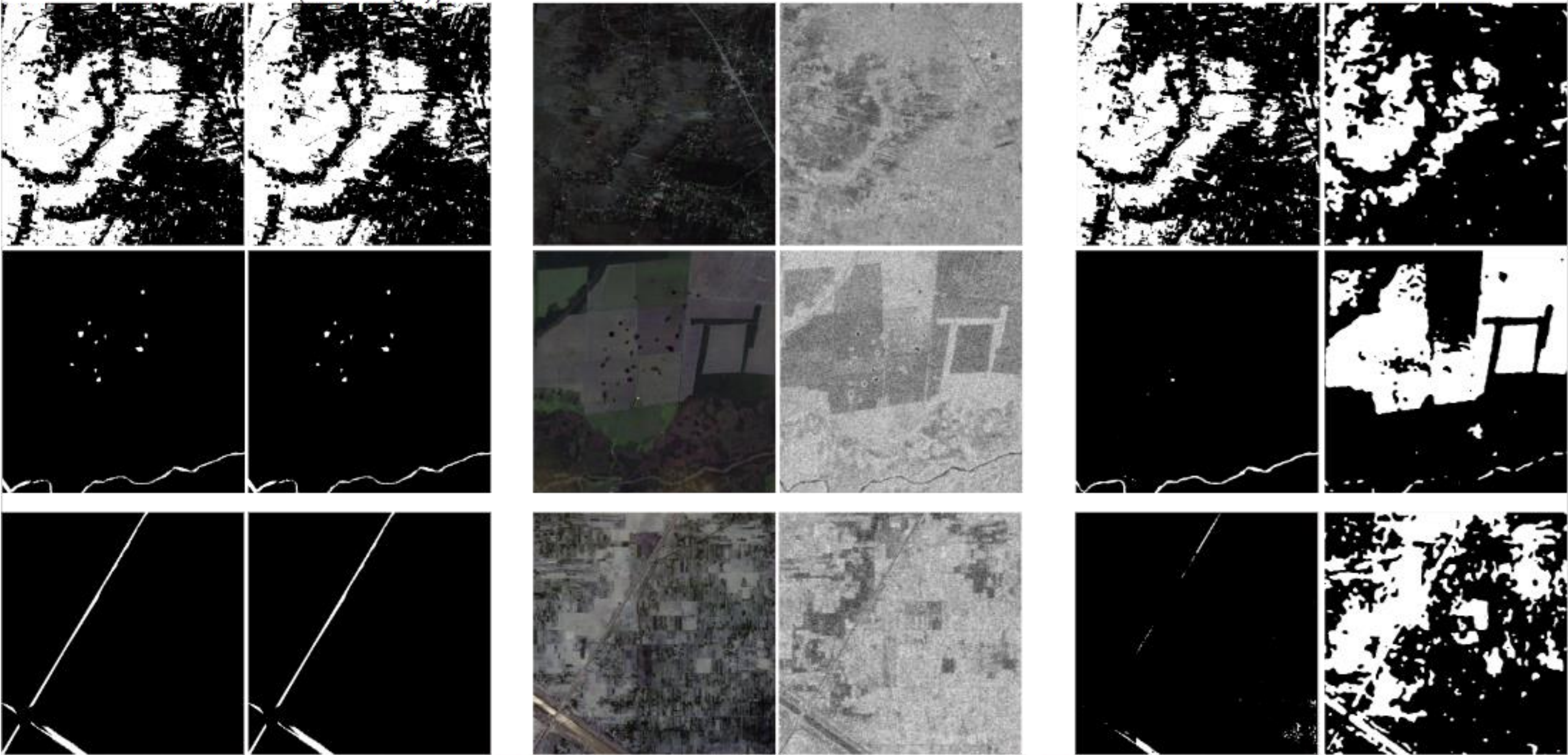
Model 1

Segmentation on mono-temporal S1 & S2 images

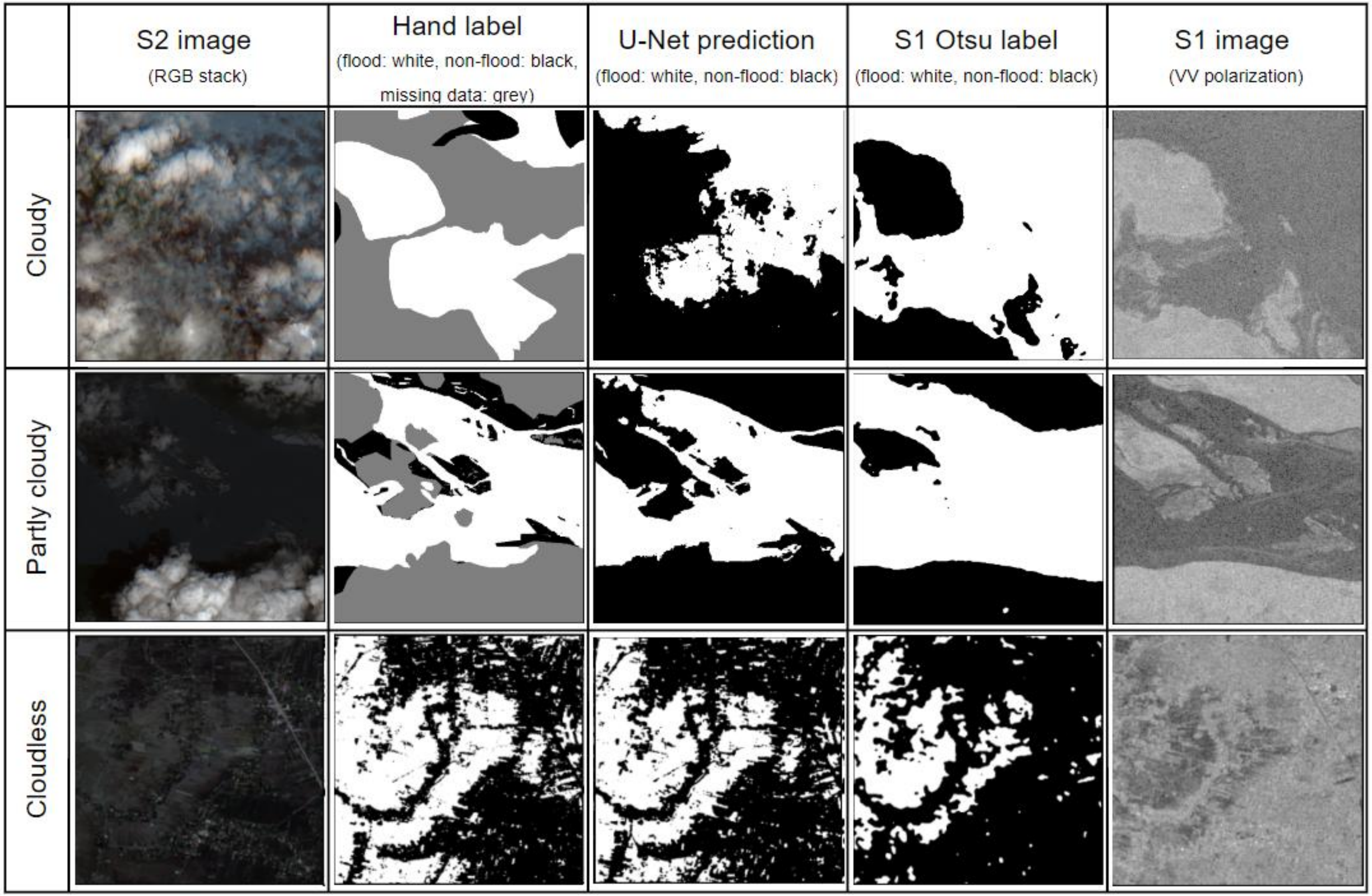
- Train input: data from Sen1Floods11. Sentinel-1 2 (VV, VH) polarizations, and Sentinel-2 13 bands with hand labels. Size of each image: 512*512.
- Architecture: U-Net backbone with ResNet-18 as encoder.
- Image transformation: applying random crop (224*224) to training input images, mask “no data” (-1) in the hand labels to “no water” (0).
- Data split: train (252 sets), validation (89 sets), test (90 sets).
- Pre-processing: remove images contain Nan values (missing swath, tile connections, mostly for Sentinel-1 images). After cleaning data: 229 train sets, 80 validation sets, 84 test sets.

Without clouds:

Urban inundation is tricky in flood detection! (one conclusion from the Sen1Floods11 paper).
But it’s more tricky for Otsu thresholding method!



Results 1


S2 image
(RGB stack)
Hand label
(flood: white, non-flood: black, missing data: grey)
U-Net prediction
(flood: white, non-flood: black)
S1 Otsu label
(flood: white, non-flood: black)
S1 image
(VV polarization)
Cloudy
Partly cloudy
Cloudless

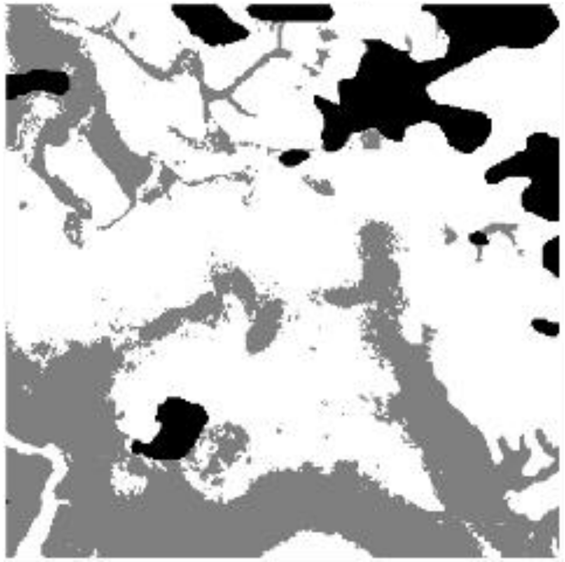
Dataset 2

Extended dataset on Sen1Floods11: multi-temporal dataset

Sen1floods11 Dataset

ID: Bolivia_129334
Flooding_date: 2018-02-15
Bound box: from label

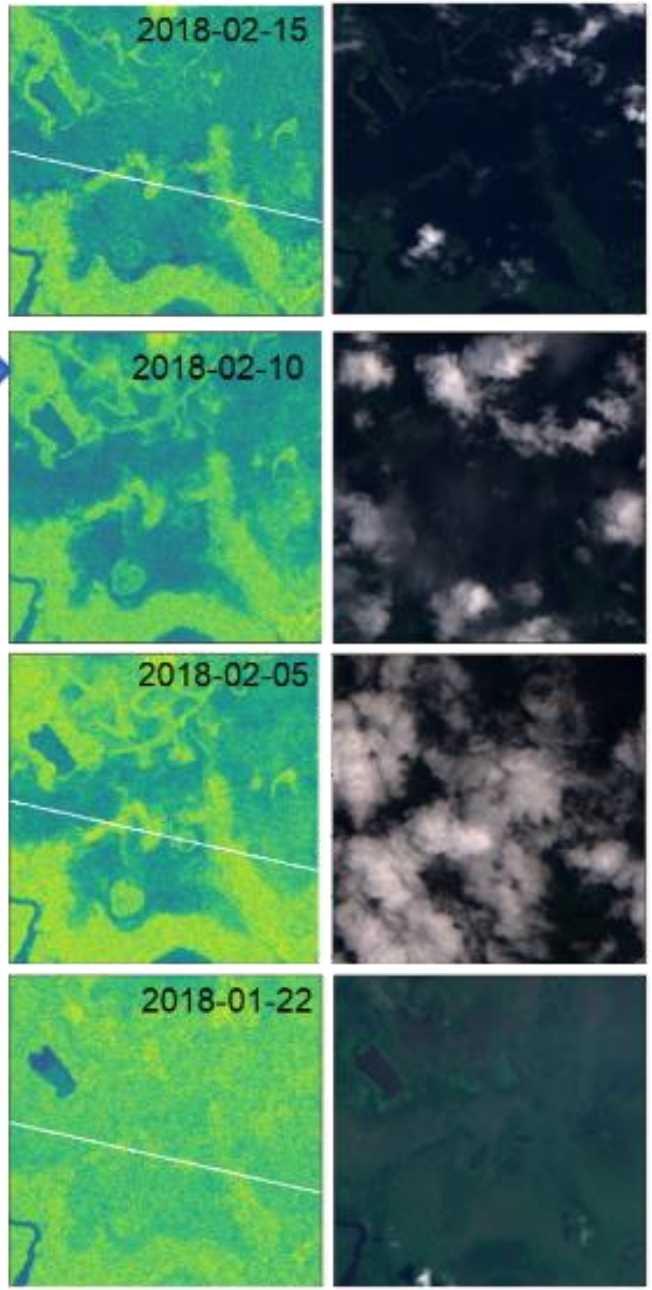
Hand_label:
(no data: grey, flood: white, non-flood: black)



My work: enriched flood event time series from Sentinel-1 and Sentinel-2

Using existing labels from Sen1floods11 dataset to prepare time series of flood events. The images in time are **filtered**, **preprocessed** (pairing S1&2 images by dates, resampling all bands to 10m resolution, interpolating for missing data, cropping) and **downloaded** by Earth Engine Python API.

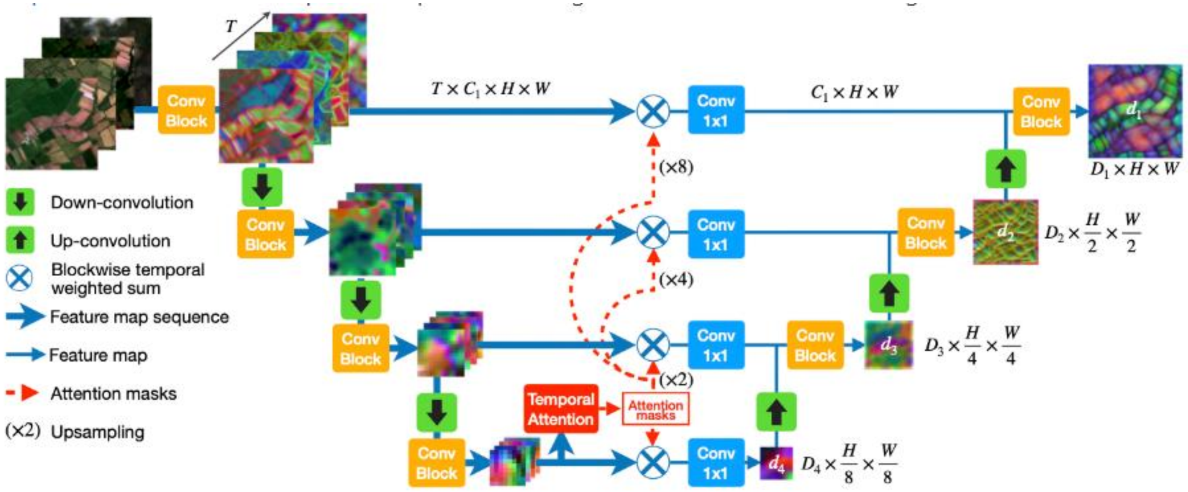
Pairing Sentinel-1 and Sentinel-2 images from one month before the flooding event



Model 2

Segmentation on multi-temporal S1 & S2 images

- Train input: Sentinel-1 2 (VV, VH) polarizations, and Sentinel-2 13 bands with hand labels. Size of each image: 512*512. For each location, the time sequence data vary from 2 to 10 pairs of S1/2 images in time.
- Architecture: UTAE, which is a U-Net backbone with Temporal Attention Encoder for satellite image time series (Garnot and Landrieu, 2021).
- Image transformation: mask “no data” (-1) in the hand labels to “no water” (0).
- Data contains 365 time sequences. Data split into 5 folds: train (3 folds, 226 sequences), validation (1 fold, 69 sequences), test (1 fold, 70 sequences).
- Pre-processing: remove images contain NaN values (missing swath, tile connections, mostly for Sentinel-1 images).



Model adapted based on UTAE architecture (Garnot and Landrieu, 2021)

Results 2

By simply adding time series data, the model cannot effectively learn information from cloudy pixels, although the positions of cloudy pixels are changing in the temporal stack. We believe the missing data from labels in Sen1Floods11 contribute to this issue.

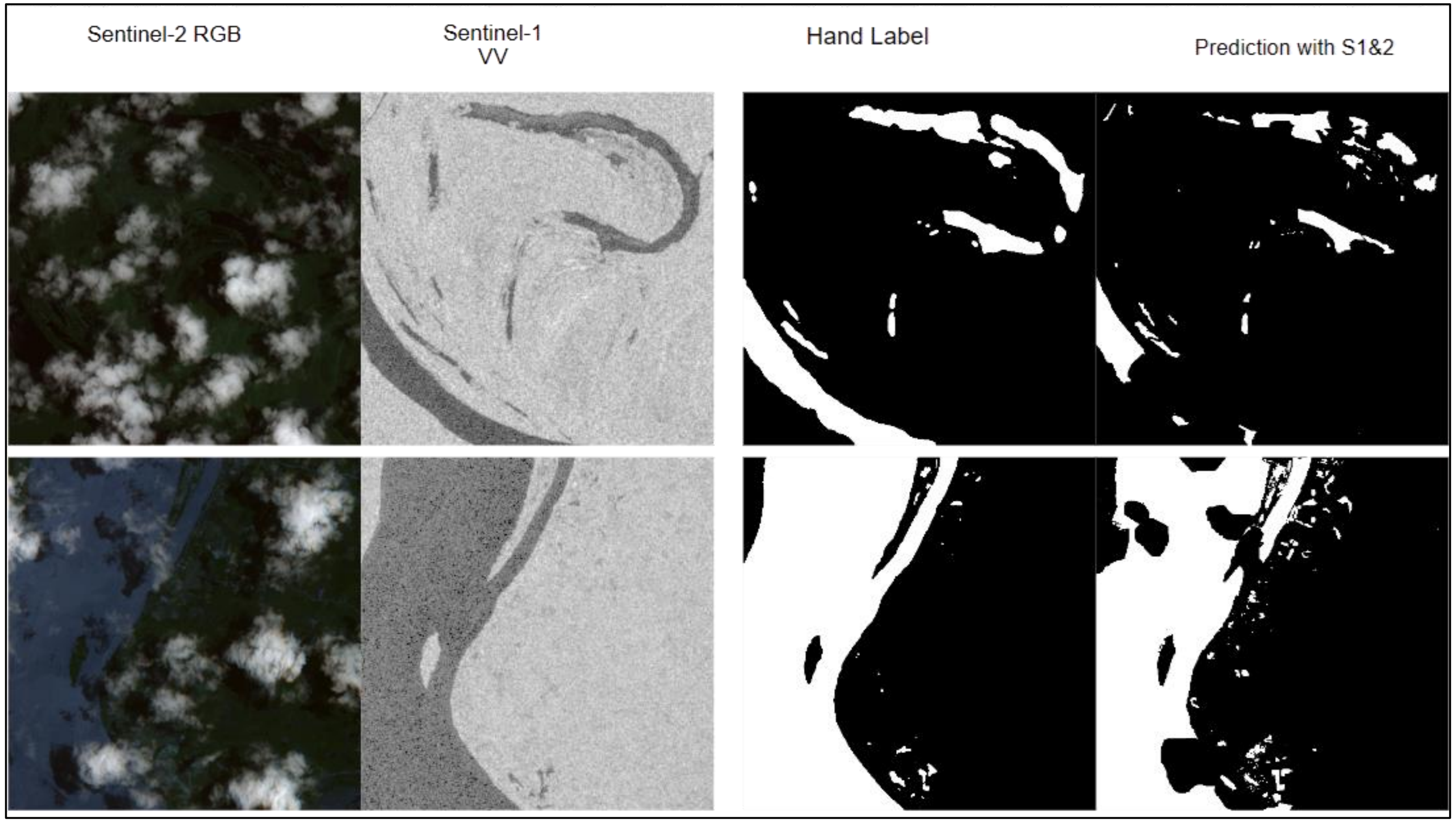


Table 1: U-Net metrics for mono-temporal imagery, compared with S1 Otsu thresholding labels

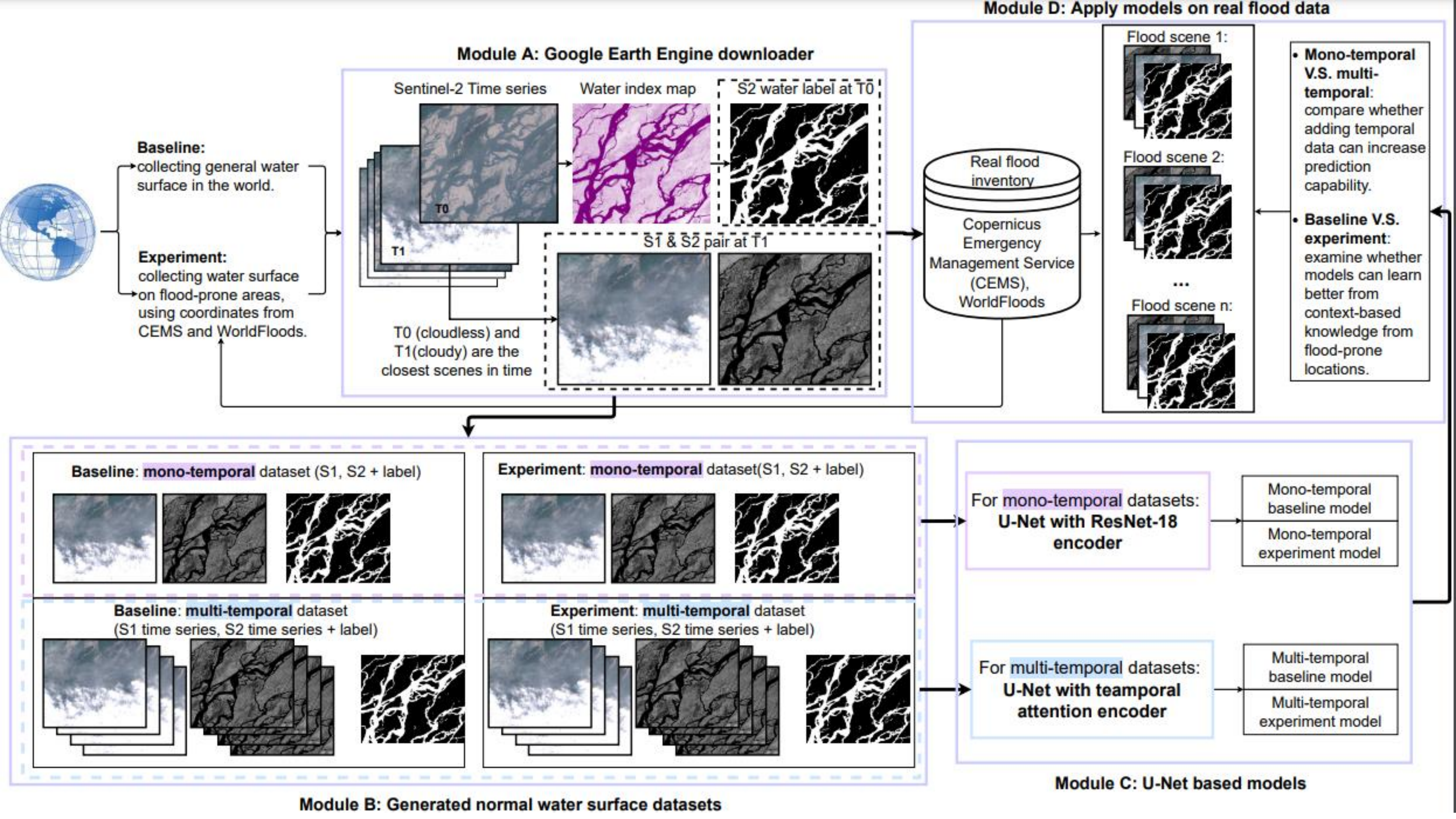
Metrics	U-Net on S1, S2	U-Net on S2	Otsu on S1
Precision	**0.971**	0.965	0.916
Mean IoU	**0.861**	0.836	0.696

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

Metrics	Macro
Precision	0.873
IoU	0.791

Proposed workflow: a dataset with complete water surface labels to force models to learn from S1 when S2 has cloudy scenes.

- **Proposed dataset** (Module A): Generate complete labels on cloudless S2 images at T0, and find cloudy scenes with S1/2 pairs at T1 that are closest to T0. Labels will be manually checked on Planet Scope high-resolution imagery.
- **Proposed model**:
  - **Baseline**: collect dataset on general water surface globally. Apply U-Net based models on the dataset.
  - **Experiment**: only collect dataset from non-flood scenes on flood-prone areas, using polygons in the existing flood dataset such as WorldFloods and CEMS datasets). Examine whether the U-Net based models can learn context-base knowledge from flood-prone locations.



Module D: Apply models on real flood data
Module A: Google Earth Engine downloader
Sentinel-2 Time series
Water index map
S2 water label at T0
T0
T1
S1 & S2 pair at T1
T0 (cloudless) and T1(cloudy) are the closest scenes in time
Baseline: collecting general water surface in the world.
Experiment: collecting water surface on flood-prone areas, using coordinates from CEMS and WorldFloods.
Real flood inventory
Copernicus Emergency Management Service (CEMS), WorldFloods
Flood scene 1:
Flood scene 2:
...
Flood scene n:
Mono-temporal V.S. multi-temporal: compare whether adding temporal data can increase prediction capability.
Baseline V.S. experiment: examine whether models can learn better from context-based knowledge from flood-prone locations.
Baseline: mono-temporal dataset (S1, S2 + label)
Experiment: mono-temporal dataset(S1, S2 + label)
Baseline: multi-temporal dataset (S1 time series, S2 time series + label)
Experiment: multi-temporal dataset (S1 time series, S2 time series + label)
For mono-temporal datasets: U-Net with ResNet-18 encoder
Mono-temporal baseline model
Mono-temporal experiment model
For multi-temporal datasets: U-Net with teamporal attention encoder
Multi-temporal baseline model
Multi-temporal experiment model
Module C: U-Net based models
Module B: Generated normal water surface datasets