Exploring Sentinel-1 and Sentinel-2 diversity for Flood inundation mapping using deep learning

Goutam Konapala*[1,2], Sujay V. Kumar[1], Shahryar Khalique Ahmad[1,3]
[1]Hydrological Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, USA
[2]Universities Space Research Association, Greenbelt, MD, USA
[3]Science Applications International Corporation, Greenbelt, MD, USA
*Corresponding author: goutam.konapala@nasa.gov

Abstract

Identification of flood water extent from satellite images has historically relied on either synthetic aperture radar (SAR) or multi-spectral (MS) imagery. MS sensors are limited to cloud free conditions, whereas SAR imagery is plagued by noise-like speckle. Prior studies that use combinations of MS and SAR data to overcome individual limitations of these sensors have not fully examined sensitivity of flood mapping performance to different combinations of SAR and MS derived spectral indices or band transformations in color space. This study explores the use of diverse bands of Sentinel 2 (S2) through well-established water indices and Sentinel 1 (S1) derived SAR imagery along with their combinations to assess their capability for generating accurate flood inundation maps. The robustness in performance of S-1 and S-2 band combinations was evaluated using 446 hand labeled flood inundation images spanning across 11 flood events from Sen1Floods11 dataset which are highly diverse in terms of land cover as well as location. A modified K-fold cross validation approach is used to evaluate the performance of 32 combinations of S1 and S2 bands using a fully connected deep convolutional neural network known as U-Net. Our results indicated that usage of elevation information has improved the capability of S1 imagery to produce more accurate flood inundation maps. Compared to a median *F1* score of 0.62 when using only S1 bands, the combined use of S1 and elevation information led to an improved median F1 score of 0.73. Water extraction indices based on S2 bands have a statistically significant superior performance in comparison to S1. Among all the band combinations, HSV (Hue, Saturation, Value) transformation of S2 bands provides a median *F1* score of 0.9, outperforming the commonly used water spectral indices owing to HSV's transformation's superior contrast distinguishing abilities. Additionally, U-Net algorithm was able to learn the relationship between raw S2 based water extraction indices and their corresponding raw S2 bands, but not of HSV owing to relatively complex computation involved in the latter. Results of the paper establishes important benchmarks for the extension of S1 and S2 data-based flood inundation mapping efforts over large spatial extents.

Introduction

Reliable, accurate and near-real time mapping of flood inundated areas is important for flood protection and preparedness (Mosavi et al., 2018). In this context, satellite remote sensing is invaluable in providing estimates of location, extent, and severity of flooding over political boundaries and complex terrain. Attempts to map flood extents using remote sensing data have focused on using either spectral data from optical sensors or backscatter data from synthetic aperture radar (SAR) (see Huang et al., 2018 and Shen et al., 2019a for review). Spectral data from optical sensors are highly correlated with open water surfaces (Irwin et al., 2017). Therefore, spectral data is the preferred source for flood inundation mapping under cloud-free weather conditions (Pekel et al., 2014, Gevaert et al., 2015). However, one of the serious limitations of these bands derived from optical imagery is their inability to penetrate cloud cover that is usually present during large scale floods caused by precipitation (McNairn et al., 2009, Clement et al., 2018). Synthetic Aperture Radar (SAR) which records data based on energy reflectance can detect water irrespective of cloud interference and is also effective both during day and nighttime (Shen et al., 2019a). However, limitations such as the inability to differentiate between water and water-like surfaces, noise-like speckle, and geometric correction may challenge the SAR's viability for global flood

inundation mapping applications (Shen et al., 2019b). Combinations of MS and SAR data are found to perform better in case of urban area extents (Iannelli and Gamba, 2018), invasive plant detection (Rajah et al., 2018), mapping soil moisture (Gao et al., 2017), wildfire assessment (Colson et al., 2018), crop monitoring (Betbeder et al., 2014, Dusseux et al., 2014), land cover classification (Ienco, 2019) and also in flood inundation mapping (Manakos et al., 2020; Bioresita et al., 2019, Slagter et al., 2020). Thus, combining both multi-spectral imagers and SARs may provide complementary information useful for flood mapping efforts.

Initial applications of flood inundation mapping involved the use of visible spectral bands and their combinations such as Normalized Difference Vegetation Index (NDVI) from coarse-scale VHRR (Very High Resolution Radiometer) and AVHRR (Advanced Very High-Resolution Radiometer) optical imagery at 1km resolution (Wiesnet, 1974; Barton and Bathols, 1989). With the advent of MODIS and Landsat, not only the spatial and temporal resolution have increased, but infrared spectral bands were also made available, significantly improving flood monitoring capabilities. Therefore, several water-based indices such as Normalized Difference Water Index (NDWI; McFeeters, 1996), Modified Normalized Difference Water Index (MNDWI; Xu et al., 2006), Automatic Water Extraction Indices (AWEI; Feyisha et al., 2014) were derived based on optical bands of Landsat and MODIS to delineate water extent. Consequently, SAR data from various satellites has also been widely used for flood mapping and management (Oberstadler et al., 1997; Matgen et al., 2007; Matgen et al., 2011; Martinis et al., 2013). Specifically, these studies indicate that dual-polarized SAR (VV, VH and HH polarization) data is suitable for flood mapping.

More recently, the European Space Agency's Sentinel-1 (S1) satellite mission which was launched in October 2014 started providing Synthetic Aperture Radar (SAR) data at 10 m resolution with 6-day revisit period. Subsequently, Sentinel-2 (S2) satellite mission which provides multi-spectral (MS) optical sensor with a 5-10 day revisit period with 10 m - 60 m spatial resolution was launched in June 2015. Since the spatial and temporal resolution of sentinel products is a significant improvement over previous satellite products, majority of recent studies have typically used Sentinel products for flood mapping (Amitrano et al., 2019; Twele et al., 2016, Li et al., 2018, Devries et al., 2020; Goffi et al., 2020; Jain et al., 2020). Since S1 and S2 have different sensors on board with S1 being an all-weather radio imaging mission and S2 a multispectral (MS) imaging satellite, the S1 and S2 measurements offer fundamentally different information sources. In addition, the high spatial and temporal resolution of S1 and S2 are of particular interest in the context of flood inundation mapping.

Most of the studies which investigate the combinations of S1 and S2 for flood inundation mapping have blended SAR sensor data using NDWI as the index. For instance, Manakos et al., (2020) used pixel based random forest classifier to identify the mean day difference between the target S1 image and available S2 high accurate inundation maps. Decision-level fusion rules was used by Bioresita et al., (2020) to integrate S1 and S2 derived NDWI bands. More recent studies have introduced additional indices for flood mapping. For example, Slagter et al., (2020) used multi-level random forest classification scheme to fuse S1 and S2 derived NDWI / MNDWI bands for flood inundation segmentation. Goffi et al., (2020) have used a rule-based framework to weight the different water indices such as NDWI, MNDWI and AWEI for water extent mapping. Even though use of NDWI with SAR has been successful to some extent in flood mapping, fusion approaches have not systematically investigated the sensitivity of inundation mapping performance metrics to different combinations of S1 and S2 derived spectral indices or band transformations in color space. Understanding the differences in performance resulting from such choices is important because the S2 based spectral indices have their own set of limitations such as their inability to capture shallow water surfaces, sensitivity to inundation of land cover classes, and misclassification due to shadows of clouds, mountains and buildings (Li et al., 2013; Boschetti et

al., 2014). Therefore, even though the combinations of S1 and S2 imagery have been applied to map flood inundations, a systematic investigation on the optimal combination of S1 and S2 is still needed. Also, prior knowledge about an optimal combination will save time and computational resources when generating flood inundation mapping in near real time settings.

Recently, there have been significant advancements in computer vision tasks such as classification and segmentation due to the growth of deep learning algorithms (Hinton and Salakhuditnov, 2006) and more particularly due to the introduction of Convolutional Neural Networks (CNNs; Jia et al., 2014). Unlike pixel-based learning approaches, CNNs can also leverage the spatial structure of target segment (such as flood inundation mask). Similarly, CNNs have been applied successfully in various flood mapping studies (Gebrehiwot et al., 2019; Nemni et al., 2020; Peng et al., 2019; Wang et al., 2020; Li et al., 2020; Potnis et al., 2019; Mateo-Garcia et al., 2021; Rambour et al., 2020; Bonafilia et al., 2020). Given the relative success of CNNs across a range of domains, we argue that a systematic investigation of the application of CNNs for blending the diverse bands of S1 and S2 will likely help us determine the optimal combination for improved flood inundation mapping. Combinations of optical and radar satellite images based on deep learning techniques have been proposed to address tasks such as land cover change detection (Liu et al., 2018), river discharge estimation (Tarpanelli et al., 2019) and land cover (Ienco et al., 2019). However, the same opportunity has not been fully exploited yet in the context of flood inundation mapping tasks.

Therefore, in this study we apply a special type of CNN known as U-Net [Ronneberger et al., 2015] successful in binary segmentation tasks to learn flood inundation mapping from S1 and S2 bands utilizing hand labeled flood inundation data covering 11 flood events across the globe [Bonafilia et al., 2020]. By training U-Net on different combinations of S1 and S2 on a dataset which spans across multiple land cover regions, we develop a classification approach more robust than prior studies which have trained on less extensive data and simpler classification algorithms. Overall, to the best of our knowledge, no other study has jointly evaluated 1) the sensitivity of combination of S1 and S2 bands on the performance of flood inundation mapping in the context of deep learning and 2) the robustness of performance of the band combinations for flood inundations across a dataset with diverse land cover spanning across 5 continents.

Through this study, we aim to address the following questions: 1) What are the optimal combination of S1 and S2 bands for flood inundation mapping through deep learning approaches? 2) Does the combination of S1 and S2 perform better than the individual performance of S1 and S2? To address these questions we structure the rest of the article as follows: the study sites and its associated data are introduced in Section 2; Section 3 describes the deep learning architecture for flood inundation mapping from radar/optical data and the experimental settings; Section 4 discusses the evaluation results. Finally, Section 5 concludes the work.

Data:

For this study, we used a recently introduced georeferenced flood label data i.e. Sen1Floods11 (Bonafiell et al., 2020) which provides flood inundation labels spanning over 11 flood events across the world (Figure1 and Table 1). This dataset contains human annotated flood labels generated for 446 images at 10-meter resolution at 512 × 512 dimensions. The dataset further provides corresponding Sentinel 1 and Sentinel 2 bands for the labeled flood events. For our study, we use all the 446-human supervised images for our deep learning model evaluation. These 446 images are selected by stratified sampling from a larger pool of 4385 images for the above mentioned 11 flood events. As Sentinel 2 has bands at varying resolutions, all bands are resampled linearly to 10 m resolution for common comparison. Further details on the Sen1Floods11 dataset characteristics are provided in Bonafiell et al., (2020). In addition to the data from Sen1Floods11, we also use elevation data derived from Shuttle Radar Topography Mission (SRTM) with a resolution of 30-meters. SRTM used two single-pass interferometers to produce a global digital elevation model (DEM) [Yang et al., 2011].

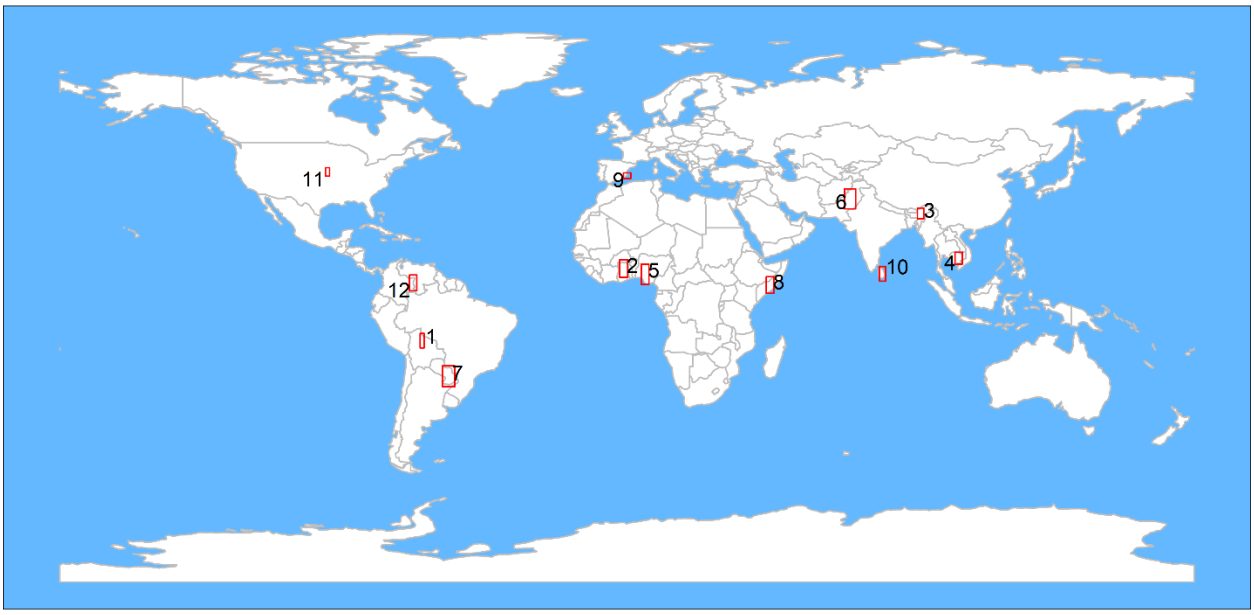

Figure 1: Locations of flood events sampled in Sen1Floods11 (adopted from Bonafilia et al., 2020)

Table 1: Flood events sampled in Sen1Floods11 dataset

ID	Country	S2 Date	S1 Date	Rel. orbit	Orbit	VH Threshold
1	BOLIVIA	2/15/2018	2/15/2018	156	Descending	<-20.44
2	GHANA	9/19/2018	9/18/2018	147	Ascending	<-22.81
3	INDIA	8/12/2016	8/12/2016	77	Descending	<-21.56
4	VIETNAM	8/4/2018	8/5/2018	26	Ascending	<23.06
5	NIGERIA	9/20/2018	9/21/2018	103	Ascending	<-21.94
6	PAKISTAN	6/28/2017	6/28/2017	5	Descending	<-19.56
7	PARAGUAY	10/31/2018	10/31/2018	68	Ascending	<-19.94
8	SOMALIA	5/5/2018	5/7/2018	116	Ascending	<-21.06
9	SPAIN	9/18/2019	9/17/2019	110	Descending	<-25.13
10	SRI LANKA	5/28/2017	5/30/2017	19	Descending	<-21.69
11	USA	5/22/2019	5/22/2019	136	Ascending	<-22.62

The 30-meter SRTM DEM is void-filled using elevation data from various other sources. To match the resolution of sentinel bands (both S1 and S2), SRTM DEM data was resampled linearly to 10-meter resolution. This resampled data is used as an ancillary input for our deep learning algorithm.

Methodology:

In this section, we describe the following aspects: (1) Preprocessing of S2 data, (2) The experimental setup, (3) Architecture of the CNN model, and (4) Evaluation of the deep learning model performance.

Preprocessing of S2 data:

As S1 SAR has only two available bands with VV and VH polarizations, we use them directly in the machine learning setup without making any

Table 2: Sentinel 2 Water Spectral indices formulation and their corresponding references

Index	Reference
MNDWI = (GREEN – SWIR1) / (GREEN + SWIR1)	**Xu, 2006**
NDWI = (GREEN – NIR) / (GREEN + NIR)	**McFeeters, 1996**
AWEI = 4* (GREEN − SWIR1) − (0.25 * NIR + 2.75 ∗ SWIR2)	**Feyisa et al., 2014**
AWEISH = BLUE + 2.5 * GREEN – 1.5 * (NIR + SWIR1) – 0.25 * SWIR2	**Feyisa et al., 2014**

further combinations/indices. However, the 12 bands in S2 allow the possibility to specify several combinations of bands. Boschetti et al., [2014] have indicated that spectral indices which contains the bands of visual spectra with wavelength (µm) between 0.5 to 0.7 and near infrared spectra with a wavelength(µm) 0.7 to 1.1 have been successful in delineating water on land. Therefore, we evaluated the spectral indices based on a set of S2 bands which can detect water features within the specified wavelength ranges (Table 2). In addition to the spectral indices, S2 Short Wave Infrared Radiation (SWIR2), Near Infrared Radiation (NIR), red bands were transformed to Hue, Saturation and Value (HSV) color space. Water can be effectively delineated by defining a relation between H, S and V components (Pekel et al., 2016, 2014). HSV transformation was initially developed for spectral bands of MODIS (Pekel et al., 2014). Subsequently, Landsat data is manually classified based on HSV values to map global water (Pekel et al., 2016). The SWIR2, NIR, red values are assigned to red, green, and blue colors (RGB) respectively and transformed into the HSV color space using a standardized colorimetric transformation (Smith, 1978), as shown in Eqs. (1), (2), (3).

$$V = \max(R, G, B) \quad (1)$$

$$S = V - \min(R, G, B) \quad (2)$$

$$H = \begin{cases} 0, \ if\ V = \min(R,G,B) \\ \left(60° \times \frac{G-B}{V-\min(R,G,B)} + 360°\right) mod\ 360°, if\ V = R \\ 60° \times \frac{B-R}{V-\min(R,G,B)} + 120°, if\ V = G \\ 60° \times \frac{R-G}{V-\min(R,G,B)} + 240°, if\ V = B \end{cases} \quad (3)$$

Hue (H), is expressed in degrees with range of 0°–360°, representing spectral composition of color. Value (V) can be defined as the brightness of color. Saturation (S) quantifies the distance of a color from a grey of equal brightness.

Experimental setup:

The 446 images used in this study are derived from different events spanning across 11 countries. For this study, we adopt a modified k-fold cross validation approach to train and test our experimental setup. This modified k-fold approach is implemented as follows

1. From the pool of eleven countries, randomly select nine countries
2. The training dataset for each of repetition in k-fold approach is augmented fourfold by flipping the S1, S2 and ground truth images up, down, right and left.
3. Train the deep learning algorithm on all the flood events for the selected nine countries
4. Evaluate the performance of trained deep learning algorithm on all the flood images from remaining two countries

5. Repeat the process for k = 10 by again randomly selecting nine countries for training and two for testing with a different combination.
6. Summarize the skill of the approach by using a median of model evaluation scores across all the test samples for all k=10 folds.

By using a testing subset from countries which are not used in training, we avoid spatial autocorrelation which is likely in the context of geo-spatial segmentation [Scratz et al., 2019]. In addition to that, for each of the 10 folds, the deep learning model is trained/tested on different subset decreasing the overall bias. For all the combinations, a separate set of experiments with DEM as an additional input to combination bands of S1 and S2 are also run. By using DEM data as an additional input, we can evaluate whether the S1 and S2 combinations through deep learning algorithms can distinguish the drainage areas without any elevation information.

Firstly, in case of Sentinel 1, we use both VV/VH bands to establish the benchmark performance of S1 for flood inundation mapping. In case of S2, the combined spectral indices of NDWI and MNDWI (referred as cNDWI), AWEIsh and AWEInsh (cAWEI) and their combination is

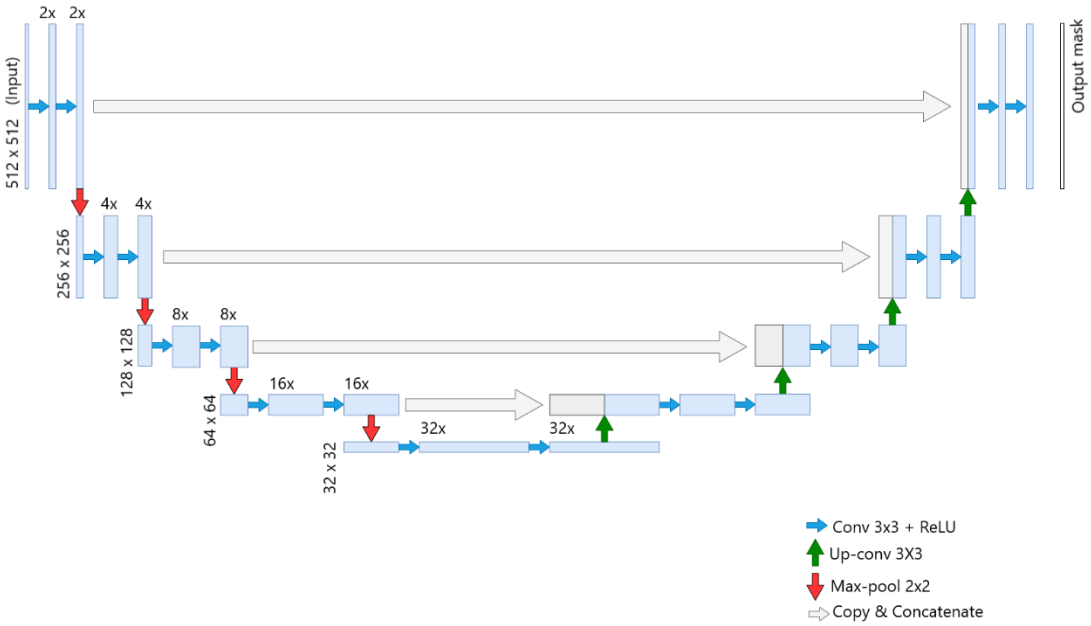

Figure 2: U-net architecture adopted for segmentation of water extents

provided as input to our deep learning algorithm (Table 2). Also, the HSV bands and their combination with spectral indices are used in our experiment. We also want to investigate if the raw bands used in cNDWI, cAWEI and HSV can also produce similar performance compared to their spectral index counterparts. Therefore, we also evaluate our deep learning algorithm using the raw bands named as rNDWI, rAWEI and rHSV and their combinations. Finally, all the S2 band configurations are combined with S1 to evaluate the effectiveness of S1 and S2 combinations. Overall, a total of 32 combinations are formulated in our experimental setup (Table 3)

Table 3: Input experiment design and their corresponding descriptions

Experiment ID	Input name	Input description
1	S1	VV and VH bands of S1
2	S1+DEM	VV and VH bands of S1 + 10 M resampled SRTM 30 M DEM
3	cAWEI	NDWI and MNDWI indices of S2
4	cAWEI+DEM	NDWI and MNDWI indices of S2 + 10 M resampled SRTM 30M DEM
5	cNDWI	AWEIsh and AWEInsh indices of S2
6	cNDWI+DEM	AWEIsh and AWEInsh indices of S2 + 10 M resampled SRTM 30M DEM
7	HSV	HSV transformation of S2 RGB Bands
8	HSV+DEM	HSV transformation of S2 RGB Bands + 10 M resampled SRTM 30M DEM
9	rNDWI	Bands used for computing NDWI and MNDWI from S2
10	rNDWI+DEM	Bands used for computing NDWI and MNDWI from S2 +10 M resampled SRTM 30M DEM
11	rAWEI	Bands used for computing AWEIsh and AWEInsh from S2
12	rAWEI+DEM	Bands used for computing AWEIsh and AWEInsh from S2 +10 M resampled SRTM 30M DEM
13	rHSV	Bands used for computing HSV from S2
14	rHSV+DEM	Bands used for computing HSV from S2 +10 M resampled SRTM 30M DEM
15	cAWEI+cNDWI	NDWI, MNDWI, AWEIsh and AWEInsh indices of S2
16	cAWEI+cNDWI+DEM	NDWI, MNDWI, AWEIsh and AWEInsh indices of S2 + 10 M resampled SRTM 30M DEM
17	HSV+cAWEI+cNDWI	HSV transformation of RGB Bands , NDWI, MNDWI, AWEIsh and AWEInsh of S2
18	HSV+cAWEI+cNDWI+DEM	HSV transformation of RGB Bands , NDWI, MNDWI, AWEIsh and AWEInsh of S2 + 10 M resampled SRTM 30M DEM
19	rAWEI+rNDWI	Bands used for computing NDWI, MNDWI, AWEIsh and AWEInsh indices of S2
20	rAWEI+rNDWI+DEM	Bands used for computing NDWI, MNDWI, AWEIsh and AWEInsh indices of S2 + 10 M resampled SRTM 30M DEM
21	rHSV+rAWEI+rNDWI	Bands used for computing HSV transformation, NDWI, MNDWI, AWEIsh and AWEInsh of S2
22	rHSV+rAWEI+rNDWI+DEM	Bands used for computing HSV transformation, NDWI, MNDWI, AWEIsh and AWEInsh of S2 + 10 M resampled SRTM 30M DEM
23	S1+cAWEI	VV , VH (S1), NDWI and MNDWI (S2)
24	S1+cAWEI+DEM	VV , VH (S1), NDWI and MNDWI (S2) + 10 M resampled SRTM 30M DEM
25	S1+cNDWI	VV , VH (S1), AWEIsh and AWEInsh (S2)
26	S1+cNDWI+DEM	VV , VH (S1), AWEIsh and AWEInsh (S2) + 10 M resampled SRTM 30M DEM
27	S1+cAWEI+cNDWI	VV , VH (S1), NDWI, MNDWI, AWEIsh and AWEInsh (S2)

28	S1+cAWEI+cNDWI+DEM	VV , VH (S1), NDWI, MNDWI, AWEIsh and AWEInsh (S2) + 10 M resampled SRTM 30M DEM
29	S1+HSV	VV , VH (S1), HSV, NDWI, MNDWI, AWEIsh and AWEInsh (S2)
30	S1+HSV+DEM	VV , VH (S1), HSV, NDWI, MNDWI, AWEIsh and AWEInsh (S2) + 10 M resampled SRTM 30M DEM
31	S1+cAWEI+cNDWI+HSV	VV , VH (S1) and HSV(S2)
32	S1+cAWEI+cNDWI+HSV+DEM	VV , VH (S1) and HSV(S2) + 10 M resampled SRTM 30M DEM

U-Net Architecture:

The CNN architecture applied for flood segmentation is U-Net as illustrated in Figure 1. U-Net is designed originally for the task of segmenting biomedical imagery [Ronneberg et al., 2016] targeted for applications which have relatively smaller number of training images and to yield segmentation at the same resolution as the input image. U-Net architecture consists of two parts (i.e. encoder and decoder) (Figure 1). Important features are extracted through the encoder part of the U-Net model through a downsampling process. The decoder uses these features as an input and builds back the spatial information of the input through an upsampling process. This upsampling and downsampling process of U-Net architecture can systematically combine low-level features and high-level features, while performing element-wise segmentation from multiple features. U-Net also incorporates skip connections to capture precise locations, at every step of the decoder as shown in the figure 1. These skip connections concatenate the output of the decoder layers with the feature maps from the encoder at the same level making the segmentation of pixels more precise. More details on the U-Net architecture and its recent applications can be found in Du et al., (2020).

We use Keras coupled with TensorFlow to train and test all our models. As our goal is to evaluate the S1 and S2 combination rather than train the best possible models, we do not perform an exhaustive hyperparameter search. We use the Adam optimizer with a base learning rate of 5e-4 and a weight decay coefficient of 1e-2 [KingMa and Ba, 2014]. Each of 32 combinations in Table 3 is trained separately using a single NVIDIA V100 GPU. All the models are trained for 500 epochs. Finally, models obtained at the end of epoch is selected for representing model performance.

Evaluation criteria:

In a binary segmentation study such as flood inundation, two outcomes which correspond to water and non-water regions are possible. The output can be classified as (1) True Positive (TP): where water pixels are correctly classified as water; (2) True Negative (TN), where non-water pixels are correctly classified as non-water regions; (3) False Positive (FP): non-water pixels incorrectly classified as water (4) False Negative (FN): water pixels incorrectly classified as non-water.

Based on these outputs, pixel *accuracy* which determines the percentage of pixels correctly classified can be computed. However, as accuracy computes this percentage irrespective of classes, it can be misleading when the class of interest (i.e. water) has relatively low number of pixels. To avoid this, *Precision*, *Recall*, and *F1* scores are commonly used. *Precision* and *Recall* are interdependent measures of over and under-segmentation Low values of *Precision* and *Recall* indicates over-segmentation and under-segmentation, respectively. *F1* score is the harmonic mean of *Precision* and *Recall* scores capturing both the aspects as a single metric.

Precision illustrates how many of the predicted water pixels matched the water pixels in the annotated labels. It can be calculated as

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

Whereas *Recall* denotes how many have been predicted as water pixels by our deep learning model. It can be defined as:

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

For an image to be classified accurately, both *Precision* and *Recall* should be high. For this purpose, *F1* score, is often used as a tradeoff metric to quantify both over- and under-segmentation into one measure.

$$F1\ Score = \frac{2\times Precision\times Recall}{Precision+Recall} \quad (6)$$

We evaluate the performance of 32 input combinations as follows

1) For each of the 32 input combinations, we generate flood inundation maps for test datasets belonging to all K =10 folds.
2) Compute *F1* score, *Precision* and *Recall* metrics separately for each individual image in test dataset across all the 10 folds.
3) Then for all 10 folds, median of all the three metrics is computed and reported.

Results:

Performance metrics of input combinations:

As highlighted before, ability of S1 and S2 data inputs to map flood inundation with and without DEM is measured by three metrics of *F1* score, *Precision* and *Recall* (Table 4). *F1* score, *Precision* and *Recall* metrics of each individual image in test dataset across all 10 folds was first computed. Then performance of each input combination is determined as median of individual metrics across all the folds. Further, non-parametric Kruskal–Wallis test is performed to determine if there is a statistically significant difference between medians of two band combinations. In our case, two band combinations are deemed to be significantly different, if the p-value is lesser than 0.05. Figure 3 presents results of Kruskal-Wallis test in form of matrix. Statistically significant differences between performance metrics on the rows and columns are indicated by the symbol ● in figure 3. Green (red) color indicates the performance of input combination on row is higher (lower) than that of input combination on the column. Performance metrics in Table 4, indicate that only S1 imagery as an input has the least F1 score of 0.62 among individual S1 and S2 inputs to

Table 4: Modified K-fold based median performance metrics for S1 and S2 inputs

F1	Precision	Recall	Input data	Type
0.62	0.59	0.88	S1	Original Bands
0.73	0.68	0.86	S1+DEM	
0.89	0.85	0.93	cAWEI	Feature engineered bands
0.88	0.86	0.93	cAWEI+DEM	
0.89	0.86	0.93	cNDWI	
0.88	0.86	0.93	cNDWI+DEM	
0.90	0.89	**0.94**	HSV	
0.90	**0.89**	0.93	HSV+DEM	
0.87	0.86	0.92	rNDWI	Original Bands
0.88	0.87	0.91	rNDWI+DEM	
0.88	0.87	0.92	rAWEI	
0.88	0.85	0.93	rAWEI+DEM	
0.86	0.86	0.89	rHSV	
0.87	0.85	0.91	rHSV+DEM	

U-Net. Low F1 score of S1 is a case of under-segmentation as expressed by a low precision score of 0.58. However, S1's performance improved with added use of DEM to a median F1 score of 0.73 from 0.62. This increase in performance of S1+DEM is due to improvement in under-segmentation as represented by a relatively better precision score of 0.68.

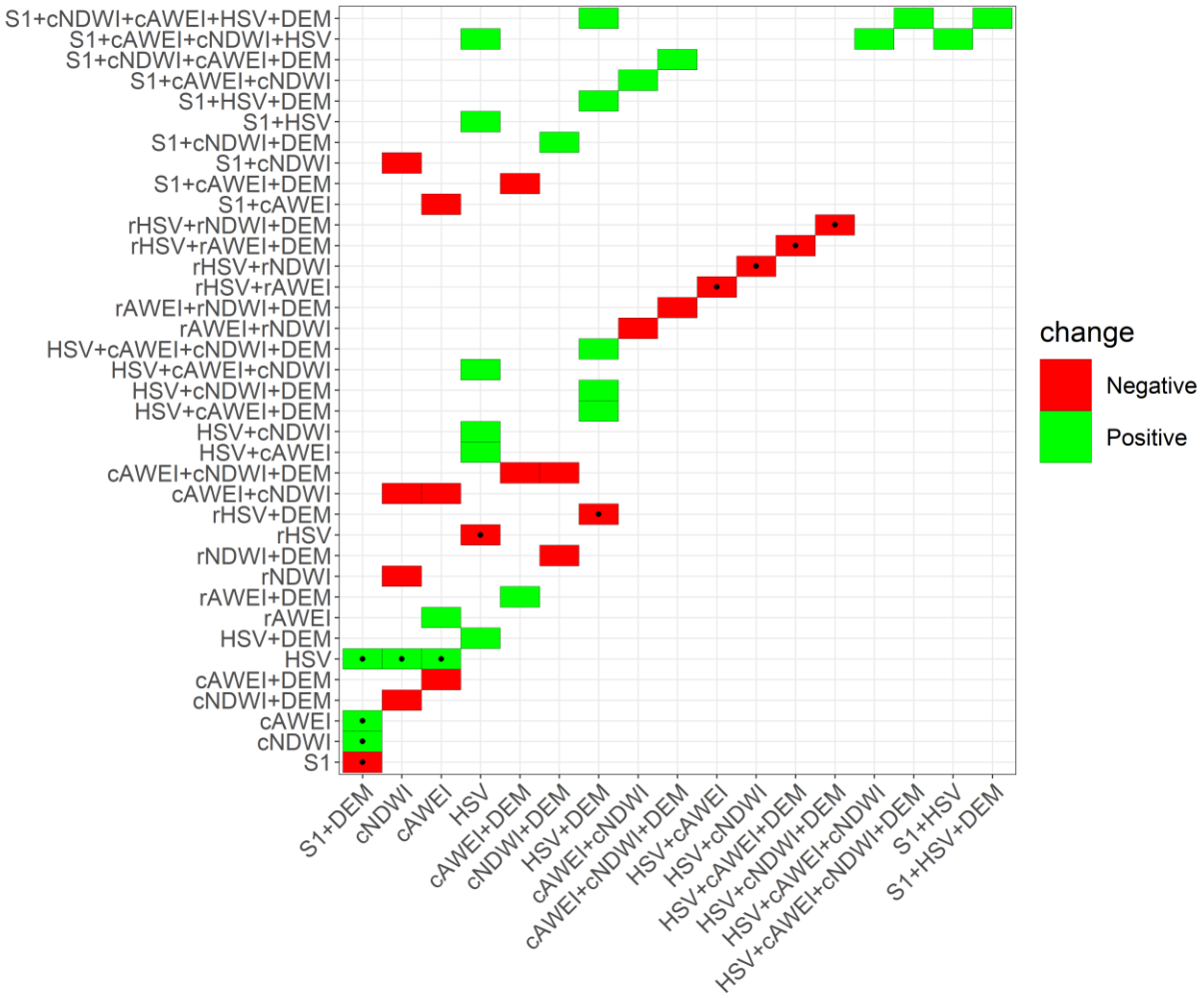

Figure 3: Outcome of Kruskal-Wallis test between the input combinations in rows and columns. The statistically significant differences between the performance metrics on the rows and columns are indicated by the symbol ●. Green (red) color indicates the performance of the input combination on the row is higher (lower) than that of the input combination on the column.

Our results indicating an improved S1 performance with added use of DEM is not surprising as elevation data distinguishes the potential flood plains from other regions and has been used previously for numerous flood inundation studies (Saksena & Merwade, 2015; Zheng et al., 2018; Musa et al., 2015). Note that the use of S1 SAR data without DEM only provides moderate skill in mapping the flood inundation. For instance, Figure 4 shows an instance of flood over a location in southeastern Paraguay in S1 images acquired on October 31, 2018 (Table 1 and Figure 1). The U-Net trained with S1 SAR data without DEM was unable to detect the flood extent in Paraguay despite

presence of a clear contrast in VV and VH backscatter (Figure 4 (B), (C), (D), (F)). However, using DEM (Fig 4 (A)) as an ancillary input dramatically improves the U-net ability to detect water extent indicating no deficiency in U-Net architecture (Figure 4 (E)). Further, speckle noise in S1 SAR images and its imperfect filtering can also lead to false positives in classified flood maps (Schmitt, 2020; Gulácsi, A., & Kovács, 2020).

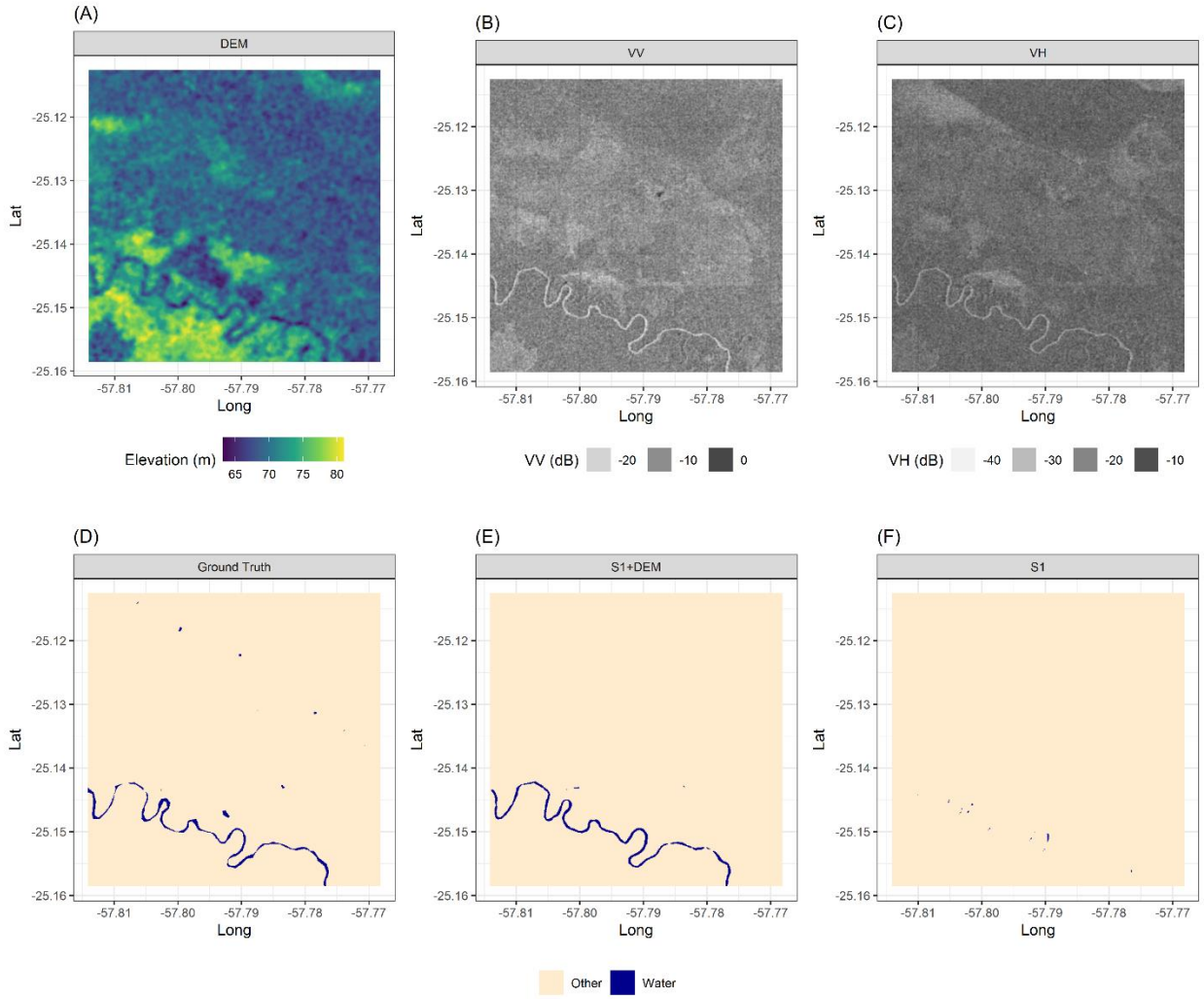

Figure 4: Flood inundation mapping inputs of (A) Elevation, (B) VV and (C) VH backscatters for a location in Paraguay. The lower panel represents the flood inundation extent by (D) ground truth, produced by U-Net when using (E) S1+DEM as input and (F) S1 as input.

In case of all S2 spectral indices with and without DEM, high *F1* scores with median F1 scores ranging between 0.88 to 0.9 values can be observed. These higher *F1* scores are due to both high *Precision* and *Recall* values indicating a balance between over and under-segmentation in case of S2 spectral indices. There does not appear to be any significant advantage in combining DEM with any of spectral indices unlike the case of S1 imagery. However, in comparison to S1, the performance of all the S2 based spectral indices have a statistically significant increase in terms of F1, precision as well as recall scores (Figure 3). However, we must note that *Recall* scores have increased by a lesser extent indicating an improvement in over-segmentation in few cases. Also, in comparison to water indices (cNDWI and cAWEI), the performance of flood inundation mapping using HSV transformation of S2

bands has a statistically significant better performance with median F1 score value of 0.9.

Interestingly, U-Net based flood mapping performance using raw bands of S2 water indices (rNDWI and rAWEI) are similar as compared to cNDWI and cAWEI (Figure 3 and Table 4). This indicates that the U-Net algorithm was able to learn features that are at least as meaningful as spectral indices for flood segmentation. However, in case of HSV, raw band's performance (rHSV) is comparatively lower, but like the performance of spectral indices. Unlike the case of spectral indices, U-net algorithm was unable to capture the relation between HSV transformations and their corresponding raw bands leading to relatively lower F1 score than its feature engineered counterpart. This differential performance of a deep learning model might be related to degree of computation complexity involved in the considered spectral indices and HSV transformation. Table 2 and equation 1-3 indicate that the spectral indices computation is relatively simpler than HSV. Therefore, based on our training data size and hyper-parameters, U-net may have captured the representation of spectral indices but not that of HSV transformation.

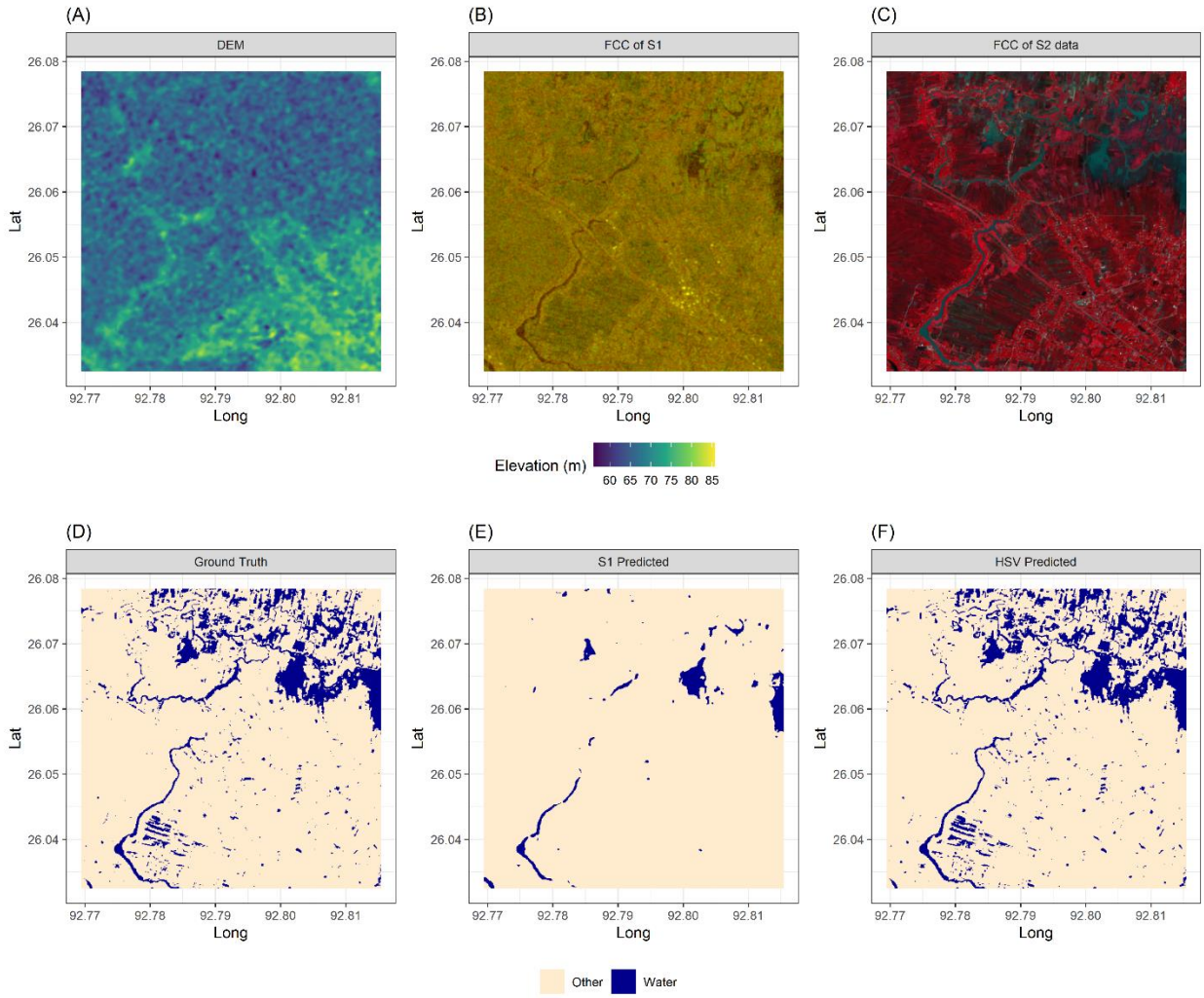

Figure 5: Flood inundation mapping inputs of (A) Elevation, (B) False color composite of S1 data and (C) False color composite of S2 for a location in India. The lower panel represents the flood inundation extent by (D) ground truth, produced by U-Net when using (E) S1+DEM as input and (F) HSV + DEM as input.

Distinct visual contrasts between water and non-water pixels in S2 based indices over S1 may be a plausible reason for the better performance of S2 based indices. For instance, Figure 5 shows an instance of flood in northeastern part of India from S1 and S2 images acquired on August 12, 2016 (Figure1 and Table 1). The false color composite (FCC) of S1 bands (Figure 5 (A),(B)) can clearly distinguish the meandering river in the south west region and the wetland in north east region, but was unable to capture the flooded agricultural lands which are in the northern regions. As a result, the U-net was able to produce flood extent only across those previously identified regions by S1's FCC image (Figure 5 (E)). However, the FCC of S2 bands was able to distinguish the flooded agricultural land, wetlands as well as the meandering rivers (Figure 5 (F)) and have produced a nearly identical map as shown in the hand labeled ground truth images (Figure 5 (D)). Hence, the resulting prediction of HSV based transformation of S2 data has captured the flood area extents across all the land types. This superior performance of HSV may be due to the better spectral correlation of band pixels with water characteristics (Huang et al., 2018).

Table 5: Modified K-fold based median performance metrics for combinations within S2 inputs

F1	Precision	Recall	Input data	Type
0.86	0.85	0.91	cAWEI+cNDWI	Feature Engineered bands
0.88	0.86	0.91	cAWEI+cNDWI+DEM	
0.90	**0.89**	0.92	HSV+cAWEI+cNDWI	
0.90	0.88	**0.93**	HSV+cAWEI+cNDWI+DEM	
0.87	0.85	0.91	rAWEI+rNDWI	Original bands
0.87	0.86	0.90	rAWEI+rNDWI+DEM	
0.88	0.85	0.91	rHSV+rAWEI+rNDWI	
0.86	0.84	0.92	rHSV+rAWEI+rNDWI+DEM	

Table 6: Modified K-fold based median performance metrics for combinations of S1 and S2 inputs

F1	Precision	Recall	Input data
0.88	0.86	0.90	S1+cAWEI
0.87	0.82	0.93	S1+cAWEI+DEM
0.88	0.88	0.91	S1+cNDWI
0.89	0.86	0.92	S1+cNDWI+DEM
0.88	0.85	**0.94**	S1+cAWEI+cNDWI
0.89	0.85	0.94	S1+cAWEI+cNDWI+DEM
0.90	0.90	0.92	S1+HSV
0.90	**0.90**	0.93	S1+HSV+DEM
0.90	0.90	0.92	S1+cAWEI+cNDWI+HSV
0.90	0.90	0.93	S1+cAWEI+cNDWI+HSV+DEM

Combinations within S2 imagery:

The previous performance results of individual spectral indices indicate that HSV transformation has superior performance in mapping floods across the globe. But, in these three indices, there are some non-overlapping S2 spectral bands and their combinations might result in further improvement in performance. Therefore, it makes sense to investigate how the combinations within S2 imagery such as combining cAWEI , cNDWI and HSV transformation would perform compared to the individual indices (Table 5). The combinations within S2 bands have indicated that, there does not appear to be any significant advantage by the combination approach. Even though there is a slight decrease in median F1 score for the combined spectral indices (cAWEI + cNDWI) when compared to the individual indices, it is still statistically insignificant (Figure 3). By combining all three S2 index approaches, we obtain a performance similar to the performance HSV alone. Therefore, our results indicate that overall, even if there are non-overlapping bands in our experimental setup, combining spectral indices and HSV has no added advantage in terms of performance.

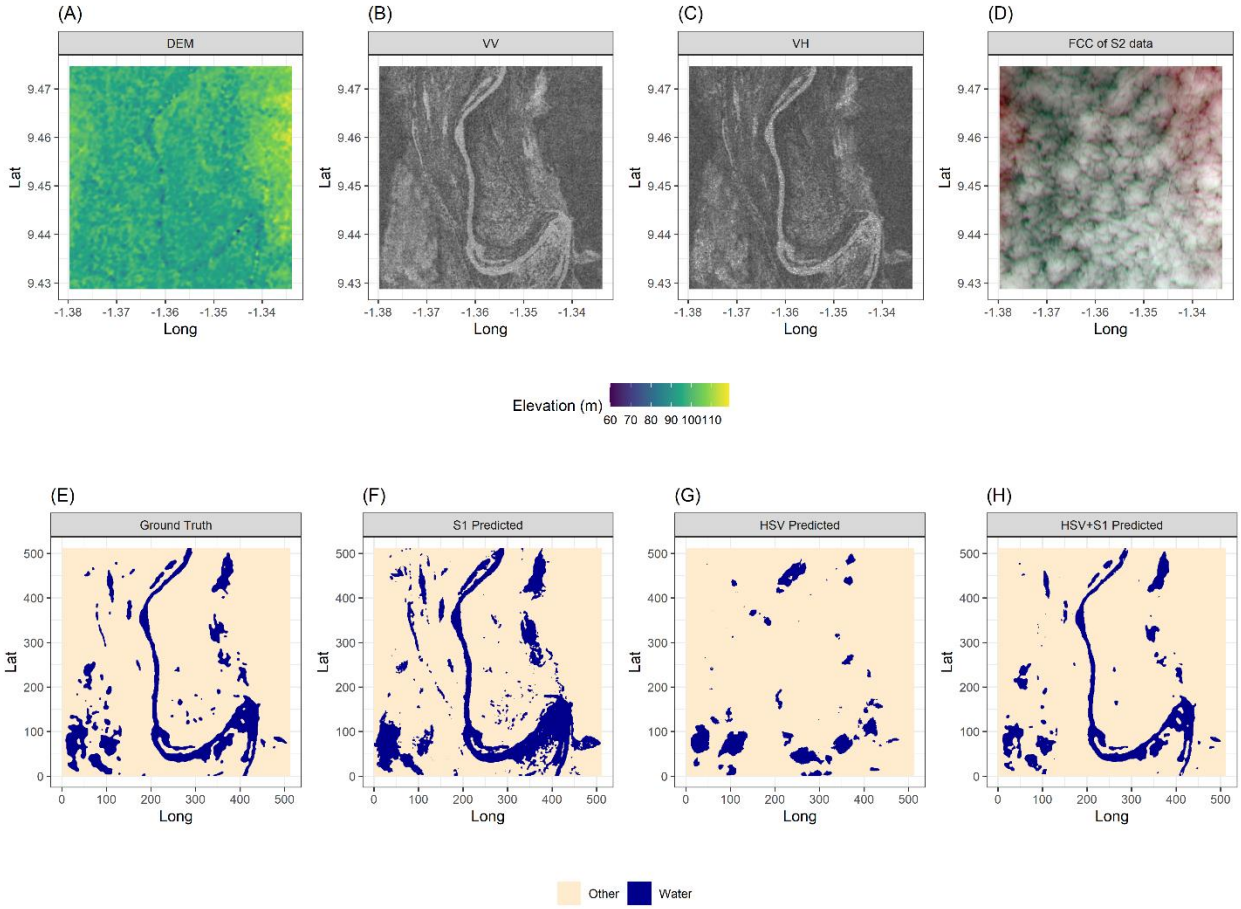

Figure 6: Flood inundation mapping inputs of (A) Elevation, (B) VV and (C)VH backscatter intensities of S1 data and (D) False color composite of S2 for a location in Ghana. The lower panel represents the flood inundation extent by (E) ground truth, U-Net produced flood masks when using (F) S1+DEM as input and (G) HSV + DEM as input and (H) fusion of HSV and S1+DEM as input.

Combinations of S1-SAR and S2-spectral imagery:

We combined S1 SAR and S2 based spectral indices to test whether there is a significant advantage in performance due to their combination and examine if there is an optimal combination of S1 and S2. Therefore, all the combinations of S1 and S2 spectral indices are tested (Table 6). Combining S1 and S2 indices indicated that their performance did not change significantly when compared to the performance of individual S2 indices. For instance, the performance of flood mapping when using cAWEI, cNDWI and HSV is not significantly different when compared to the performance of S1+ cAWEI , S1+ cNDWI and S1+HSV (Figure 3). All the best performing input combinations have HSV as a part of their combination. However, the performance of these input combinations is not significantly different from each other and are similar to the performance of HSV alone. Even though the combination of S1 and S2 data has relatively no performance improvement in comparison to individual S2 indices in detecting the flood water extents, it is important to clarify that the U-net trained with cloud obscured S2 images still cannot efficiently detect flood extents. An example instance is shown in Figure 12 for a flooding case over central Ghana using S1 and S2 images acquired on September 18-19, 2016 (Table 1 and Figure 1). The DEM shows a narrow strip of low elevation indicating a potentially narrow river (Figure 6 (A)). VV (Figure 6 (B)) and VH (Figure 6 (C)) backscatters show a complete picture of flood plain surrounding the river and other flood inundated regions across the plains. FCC of S2 indicated a total cloud coverage over a flooded region in Ghana (Figure 6 (D)). Consequently, the S1 predicted flooded regions (Figure 6 (F)) have captured the flooded extent similar to ground truth (Figure 6 (E)) with some flood plains areas being overestimated. Despite, the huge cloud coverage, the HSV transformation of S2 data could capture the flood plains in the Southwest region and some other patches across the map (Figure 6 (G)). But it could not capture the river segment due to the dense cloud coverage. However, the fusion of both S1 backscatter and S2's HSV transformation has resulted in a more accurate representation of flood extent without any overestimation across the flood plains (Figure 6 (H)). Since, Sens1Floods11 dataset is curated to exclude majority of the satellite imagery with clouds, we could not find the significant advantage in our experiments. However, in presence of clouds, the fusion of S1 and S2 will have an added advantage.

Discussion:

Our results indicate that even though the SAR data is not affected by cloud cover, poor contrast between VV and VH backscatter affects S1 data's flood inundation mapping performance. F1 score for S1 based flood water mapping varied between 0.65 to 0.91 in previous studies depending on the location and approach [Bioresita et al., 2018, Liang and Liu, 2020]. Therefore, the F1 score of 0.62 obtained in our case compares reasonably well to these studies. Our current configuration performs semantic segmentation with a single-trained model for the entire dataset. Previous studies have suggested that this may lead to poor accuracy due to unclear backscatter conditions in case of S1 of the inundated areas in different land cover features [Manakos et al., 2020]. Therefore, a land cover specific ensemble of deep learning models or pixel centric approaches can be used to improve the performance [Huang et al., 2018; Pham-Duc et al., 2017]. Additionally, the 10-meter resampled SRTM 30M - DEM improves S1 data's flood mapping abilities. This can be attributed to the identification of low lying flood plains (Manfreda et al., 2015; Samela et al., 2016) in the elevation data. Also, in urban areas, elevation data can help us to distinguish the areas of urban ground surface which may not be visible due to radar shadowing and building layovers preventing misclassification of flooded water extent [Soergel et al., 2003]. Further, DEMs can helps us distinguish roads and tarmac areas have a low backscatter similar to water creating a misclassification [Mason et al., 2014]. But given the dependence of flood inundation mapping on elevation data's spatial resolution (Haile and Rientjes, 2005; Fereshtehpour, and Karamouz, 2018), further studies should consider the sensitivity of flood inundation map's performance to the DEM spatial resolution as well as the data source.

S2 based spectral indices are generally better than S1 based SAR data for mapping flood extents due to stronger correlations between spectral features of S2 bands with respect to water surfaces in the cloud-free S2 images (Boschetti et al., 2014; Klein et al., 2017; Bonafilia et al., 2020). Also, our results which state that the S2 has better performance than S1 is similar to conclusion obtained by Bonafilia et al., 2020. In particular, the HSV transformation of RGB bands in S2 satellite data has a statistically significant superior performance in comparison to conventional water spectral indices probably due to improved contrast between water and non-water surface in HSV components (Pekel et al., 2014, 2016). However, the difference in performance between HSV and spectral indices is relatively low. The performance of S1 and S2 combinations for flood water mapping was not significantly different from the individual S2 indices' performance. This is not surprising in our case, as the Sen1Floods11 dataset has relatively a smaller number of satellite imagery with clouds. But, in satellite imagery with clouds, the fusion of S1 and S2 imagery has performed significantly better than individual S2 indices as demonstrated in figure 6. In addition to that, previous studies have also indicated in the fusion approach, S2 bands' inability in penetrating clouds is complemented by the SAR's ability to map water extent during cloud cover.

Interestingly, U-Net algorithm's performance using cNDWI/cAWEI in segmenting water inundated areas is similar to that of using raw bands used for computing NDWI/AWEI indicating the ability of U-Net to learn the features with similar characteristics as spectral indices for flood segmentation as a part of its encoder operations. However, the U-Net algorithm's performance using S2 bands transformed to HSV is superior compared to that of using raw S2 bands. This may be due to the relatively shorter data record for training the U-Net algorithm, or inability of the U-net to capture the relatively complex formulation of HSV. Therefore, relatively complex feature engineering such as HSV transformation may be performed before training the deep learning algorithm for flood inundation mapping.

U-net architecture (Ronneberger et al., 2015) has been previously proven to efficient in segmentation of binary classes. Even though our focus was to explore the diverse bands of S2 and S1 for flood mapping, our adopted configuration of U-Net has performed well in delineating the flooded regions. In this work, we modified the U-Net configuration to accept S1 and S2 bands instead of the traditional RGB bands to delineate flood floods using Multi-spectral and SAR imagery. However, in case of S1, our modified U-Net's inability to find a robust threshold which is sensitive enough to identify the contrast between water and non-water pixels based on SAR VV and VH backscatter frequencies has resulted in decreased performance in some flooded regions. This may be achieved through modification of convolution configurations, activation functions and loss functions in U-Nets paving way for future research. In addition, benchmarking different deep learning architectures for flood inundation mapping using the Sen1Floods11 dataset can be investigated.

Conclusions

In this paper we explored the diverse bands of S2 and S1 satellites along with combinations for flood inundation mapping through a deep CNN model known as U-Net to train, validate and test against manually annotated pixel level flood inundated images. Our results indicate that using DEM as an ancillary data can improve the performance of U-net when using S1 imagery as input. However, U-Net algorithm has shown a better performance when using S2 bands when compared to S1 bands, likely due to better spectral correlation between optical sensor output and water features. In addition, there is minimal influence of DEM ancillary on the median performance of S2 bands. Among the S2 metrics, the U-Net with the HSV transformation of RGB bands outperforms the established spectral indices such as AWEIsh, AWEInsh, NDWI and MNDWI owing to its superior visual contrast segmentation. The U-Net algorithm was able to learn the relationship between raw S2 bands and cNDWI and cAWEI, but not of HSV owing to relatively complex computation involved in the latter. Therefore, based on our training data size

and hyper-parameters, U-net may have captured the representation of spectral indices but not that of HSV transformation. These results also show that automatic flood detection is possible when an appropriate water index technique is being used. The extension of our approach to benchmark the performance of different deep learning architectures for flood water segmentation is left for a future work.

Acknowledgements

This research was supported by the NASA Earth Science Technology Office sponsored New Observing System (NOS) project. Computing was supported by the resources at the NASA Center for Climate Simulation. The authors also acknowledge Cloud to Street for developing open-source flood label data available for access through Google Cloud Storage bucket at: gs://senfloods11/

Declaration of Competing Interest

None

References

Amitrano, D., Di Martino, G., Iodice, A., Riccio, D. and Ruello, G., 2018. Unsupervised rapid flood mapping using Sentinel-1 GRD SAR images. IEEE Transactions on Geoscience and Remote Sensing, 56(6), pp.3290-3299, doi: 10.1109/TGRS.2018.2797536

Auynirundronkool, K., Chen, N., Peng, C., Yang, C., Gong, J., & Silapathong, C. (2012). Flood detection and mapping of the Thailand Central plain using RADARSAT and MODIS under a sensor web environment. International Journal of Applied Earth Observation and Geoinformation, 14(1), 245-255, doi: 10.1016/j.jag.2011.09.017

Barton, I.J. and Bathols, J.M., 1989. Monitoring floods with AVHRR. Remote sensing of Environment, 30(1), pp.89-94, doi: 10.1016/0034-4257(89)90050-3

Betbeder, J., Rapinel, S., Corpetti, T., Pottier, E., Corgne, S., & Hubert-Moy, L. (2014). Multitemporal classification of TerraSAR-X data for wetland vegetation mapping. Journal of applied remote sensing, 8(1), 083648, doi: 10.1117/1.JRS.8.083648

Bioresita, F., Puissant, A., Stumpf, A., & Malet, J. P. (2019). Fusion of Sentinel-1 and Sentinel-2 image time series for permanent and temporary surface water mapping. International Journal of Remote Sensing, 40(23), 9026-9049, doi: 10.1080/01431161.2019.1624869

Bonafilia, D., Tellman, B., Anderson, T., & Issenberg, E. (2020). Sen1Floods11: a georeferenced dataset to train and test deep learning flood algorithms for Sentinel-1. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (pp. 210-211), doi: 10.1109/CVPRW50498.2020.00113

Boschetti, M., Nutini, F., Manfron, G., Brivio, P. A., & Nelson, A. (2014). Comparative analysis of normalised difference spectral indices derived from MODIS for detecting surface water in flooded rice cropping systems. PloS one, 9(2), e88741, doi: 10.1371/journal.pone.0088741

Clement, M. A., Kilsby, C. G., & Moore, P. (2018). Multi-temporal synthetic aperture radar flood mapping using change detection. Journal of Flood Risk Management, 11(2), 152-168, doi: 10.1111/jfr3.12303

Colson, D., Petropoulos, G. P., & Ferentinos, K. P. (2018). Exploring the potential of Sentinels-1 & 2 of the Copernicus Mission in support of rapid and cost-effective wildfire assessment. International journal of applied earth observation and geoinformation, 73, 262-276, doi: 10.1016/j.jag.2018.06.011

DeVries, B., Huang, C., Armston, J., Huang, W., Jones, J. W., & Lang, M. W. (2020). Rapid and robust monitoring of flood events using Sentinel-1 and Landsat data on the Google Earth Engine. Remote Sensing of Environment, 240, 111664, doi: 10.1016/j.rse.2020.111664

Du, G., Cao, X., Liang, J., Chen, X., & Zhan, Y. (2020). Medical image segmentation based on u-net: A review. Journal of Imaging Science and Technology, 64(2), 20508-1, doi: 10.2352/J.ImagingSci.Technol.2020.64.2.020508

Dusseux, P., Corpetti, T., Hubert-Moy, L., & Corgne, S. (2014). Combined use of multi-temporal optical and radar satellite images for grassland monitoring. Remote Sensing, 6(7), 6163-6182, doi:10.3390/rs6076163

Fereshtehpour, M., & Karamouz, M. (2018). DEM resolution effects on coastal flood vulnerability assessment: Deterministic and probabilistic

approach. Water Resources Research, 54(7), 4965-4982, doi: 10.1029/2017WR022318
Feyisa, G. L., Meilby, H., Fensholt, R., & Proud, S. R. (2014). Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery. Remote Sensing of Environment, 140, 23-35, doi: 10.1016/j.rse.2013.08.029
Gao, Q., Zribi, M., Escorihuela, M. J., & Baghdadi, N. (2017). Synergetic use of Sentinel-1 and Sentinel-2 data for soil moisture mapping at 100 m resolution. Sensors, 17(9), 1966, doi: 10.3390/s17091966
Gebrehiwot, A., Hashemi-Beni, L., Thompson, G., Kordjamshidi, P., & Langan, T. E. (2019). Deep convolutional neural network for flood extent mapping using unmanned aerial vehicles data. Sensors, 19(7), 1486.
Gevaert, C. M., Suomalainen, J., Tang, J., & Kooistra, L. (2015). Generation of spectral–temporal response surfaces by combining multispectral satellite and hyperspectral UAV imagery for precision agriculture applications. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8(6), 3140-3146, doi: 10.1109/JSTARS.2015.2406339
Goffi, A., Stroppiana, D., Brivio, P. A., Bordogna, G., & Boschetti, M. (2020). 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, 84, 101951, doi: 10.1016/j.jag.2019.101951
Gulácsi, A., & Kovács, F. (2020). Sentinel-1-imagery-based high-resolution water cover detection on wetlands, Aided by Google Earth Engine. Remote Sensing, 12(10), 1614, doi: 10.3390/rs12101614
Haile, A. T., & Rientjes, T. H. M. (2005). Effects of LiDAR DEM resolution in flood modelling: a model sensitivity study for the city of Tegucigalpa, Honduras. Isprs wg iii/3, iii/4, 3, 12-14
He, W., & Yokoya, N. (2018). Multi-temporal sentinel-1 and-2 data fusion for optical image simulation. ISPRS International Journal of Geo-Information, 7(10), 389, doi: 10.3390/ijgi7100389
Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. science, 313(5786), 504-507, doi: 10.1126/science.1127647
Hornik, K., Stinchcombe, M., & White, H. (1990). Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. Neural networks, 3(5), 551-560.
Huang, C., Chen, Y., Zhang, S., & Wu, J. (2018). Detecting, extracting, and monitoring surface water from space using optical sensors: A review. Reviews of Geophysics, 56(2), 333-360, doi: 10.1029/2018RG000598
Huang, G. B., Chen, L., & Siew, C. K. (2006). Universal approximation using incremental constructive feedforward networks with random hidden nodes. IEEE Trans. Neural Networks, 17(4), 879-892. Manakos, I., Kordelas, G. A., & Marini, K. (2020). Fusion of Sentinel-1 data with Sentinel-2 products to overcome non-favourable atmospheric conditions for the delineation of inundation maps. European Journal of Remote Sensing, 53(sup2), 53-66.
Huang, W., DeVries, B., Huang, C., Lang, M.W., Jo nes, J.W., Creed, I.F.,
& Carroll, M.L. (2018). Automated extraction of surface water extent from Sentinel-1 data. Remote Sensing, 10(5), 797. doi:10.3390/rs10050797
Iannelli, G. C., & Gamba, P. (2018, July). Jointly exploiting Sentinel-1 and Sentinel-2 for urban mapping. In IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium (pp. 8209-8212), doi: 10.1109/IGARSS.2018.8518172
Ienco, D., Interdonato, R., Gaetano, R., & Minh, D. H. T. (2019). Combining Sentinel-1 and Sentinel-2 Satellite Image Time Series for land cover mapping via a multi-source deep learning architecture. ISPRS Journal of Photogrammetry and Remote Sensing, 158, 11-22, doi : 10.1016/j.isprsjprs.2019.09.016
Irwin, K., Beaulne, D., Braun, A., & Fotopoulos, G. (2017). Fusion of SAR, optical imagery and airborne LiDAR for surface water detection. Remote Sensing, 9(9), 890, doi: 10.3390/rs9090890
Jain, P., Schoen-Phelan, B., & Ross, R. (2020, March). Automatic flood detection in SentineI-2 images using deep convolutional neural networks. In Proceedings of the 35th Annual ACM Symposium

on Applied Computing (pp. 617-623), doi: 10.1145/3341105.3374023

Jamali, A., Mahdianpari, M., Brisco, B., Granger, J., Mohammadimanesh, F., & Salehi, B. (2021). Comparing Solo Versus Ensemble Convolutional Neural Networks for Wetland Classification Using Multi-Spectral Satellite Imagery. Remote Sensing, 13(11), 2046.

Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., ... & Darrell, T. (2014, November). Caffe: Convolutional architecture for fast feature embedding. In Proceedings of the 22nd ACM international conference on Multimedia (pp. 675-678), doi: 10.1145/2647868.2654889

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

Klein, I., Gessner, U., Dietz, A. J., & Kuenzer, C. (2017). Global WaterPack–A 250 m resolution dataset revealing the daily dynamics of global inland water bodies. Remote sensing of environment, 198, 345-362, doi: 10.1016/j.rse.2017.06.045

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. Journal of the American statistical Association, 47(260), 583-621.

Li, S., Sun, D., & Yu, Y. (2013). Automatic cloud-shadow removal from flood/standing water maps using MSG/SEVIRI imagery. International journal of remote sensing, 34(15), 5487-5502. Doi: 10.1080/01431161.2013.792969

Li, Y., Martinis, S., & Wieland, M. (2019). Urban flood mapping with an active self-learning convolutional neural network based on TerraSAR-X intensity and interferometric coherence. ISPRS Journal of Photogrammetry and Remote Sensing, 152, 178-191.

Li, Y., Martinis, S., Plank, S., & Ludwig, R. (2018). An automatic change detection approach for rapid flood mapping in Sentinel-1 SAR data. International journal of applied earth observation and geoinformation, 73, 123-135, doi: 10.1016/j.jag.2018.05.023

Liang, J., & Liu, D. (2020). A local thresholding approach to flood water delineation using Sentinel-1 SAR imagery. ISPRS Journal of Photogrammetry and Remote Sensing, 159, 53-62.

Liu, J., Gong, M., Qin, K., & Zhang, P. (2016). A deep convolutional coupling network for change detection based on heterogeneous optical and radar images. IEEE transactions on neural networks and learning systems, 29(3), 545-559, doi: 10.1109/TNNLS.2016.2636227

Manakos, I., Kordelas, G. A., & Marini, K. (2020). Fusion of Sentinel-1 data with Sentinel-2 products to overcome non-favourable atmospheric conditions for the delineation of inundation maps. European Journal of Remote Sensing, 53(sup2), 53-66, doi: 10.1080/22797254.2019.1596757

Manfreda, S., Nardi, F., Samela, C., Grimaldi, S., Taramasso, A. C., Roth, G., and Sole, A., (2014). Investigation on the Use of Geomorphic Approaches for the Delineation of Flood Prone Areas, Journal of Hydrology, 517, 863-876, doi: 10.1016/j.jhydrol.2014.06.009

Manjusree, P., Kumar, L. P., Bhatt, C. M., Rao, G. S., & Bhanumurthy, V. (2012). Optimization of threshold ranges for rapid flood inundation mapping by evaluating backscatter profiles of high incidence angle SAR images. International Journal of Disaster Risk Science, 3(2), 113-122,doi: 10.1007/s13753-012-0011-5

Martinis, S., Kersten, J., & Twele, A. (2015). A fully automated TerraSAR-X based flood service. ISPRS Journal of Photogrammetry and Remote Sensing, 104, 203-212, doi: 10.1016/j.isprsjprs.2014.07.014

Martinis, S., Twele, A., Strobl, C., Kersten, J. and Stein, E., 2013. A multi-scale flood monitoring system based on fully automatic MODIS and TerraSAR-X processing chains. Remote Sensing, 5(11), pp.5598-5619, doi: 10.3390/rs5115598

Mason, D. C., Speck, R., Devereux, B., Schumann, G. J. P., Neal, J. C., & Bates, P. D. (2009). Flood detection in urban areas using TerraSAR-X. IEEE Transactions on Geoscience and Remote Sensing, 48(2), 882-894, doi: r 10.1109/TGRS.2009.2029236

Mateo-Garcia, G., Veitch-Michaelis, J., Smith, L., Oprea, S. V., Schumann, G., Gal, Y., ... & Backes, D. (2021). Towards global flood mapping onboard low cost satellites with machine learning. Scientific

reports, 11(1), 1-12. doi: 10.1038/s41598-021-86650-z

Matgen, P., Hostache, R., Schumann, G., Pfister, L., Hoffmann, L. and Savenije, H.H.G., 2011. Towards an automated SAR-based flood monitoring system: Lessons learned from two case studies. Physics and Chemistry of the Earth, Parts A/B/C, 36(7-8), pp.241-252, doi: 10.1016/j.pce.2010.12.009

Matgen, P., Schumann, G., Henry, J.B., Hoffmann, L. and Pfister, L., 2007. Integration of SAR-derived river inundation areas, high-precision topographic data and a river flow model toward near real-time flood management. International Journal of Applied Earth Observation and Geoinformation, 9(3), pp.247-263, doi: 10.1016/j.jag.2006.03.003

McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. International journal of remote sensing, 17(7), 1425-1432, doi: 10.1080/01431169608948714

McNairn, H., Champagne, C., Shang, J., Holmstrom, D., & Reichert, G. (2009). Integration of optical and Synthetic Aperture Radar (SAR) imagery for delivering operational annual crop inventories. ISPRS Journal of Photogrammetry and Remote Sensing, 64(5), 434-449. doi: 10.1016/j.isprsjprs.2008.07.006

Mosavi, A., Ozturk, P. and Chau, K.W., 2018. Flood prediction using machine learning models: Literature review. Water, 10(11), p.1536, doi: 10.3390/w10111536

Musa, Z. N., Popescu, I., & Mynett, A. (2015). A review of applications of satellite SAR, optical, altimetry and DEM data for surface water modelling, mapping and parameter estimation. Hydrology and Earth System Sciences, 19(9), 3755, doi: 10.5194/hess-19-3755-2015

Nemni, E., Bullock, J., Belabbes, S., & Bromley, L. (2020). Fully convolutional neural network for rapid flood segmentation in synthetic aperture radar imagery. Remote Sensing, 12(16), 2532.

Oberstadler, R., Hönsch, H. and Huth, D., 1997. Assessment of the mapping capabilities of ERS-1 SAR data for flood mapping: a case study in Germany. Hydrological processes, 11(10), pp.1415-1425.

Ohki, M., Tadono, T., Itoh, T., Ishii, K., Yamanokuchi, T., Watanabe, M., & Shimada, M. (2019). Flood area detection using PALSAR-2 amplitude and coherence data: The case of the 2015 heavy rainfall in Japan. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 12(7), 2288-2298, doi: 10.1109/JSTARS.2019.2911596

Pekel, J. F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. Nature, 540(7633), 418-422, doi:10.1038/nature20584

Pekel, J. F., Vancutsem, C., Bastin, L., Clerici, M., Vanbogaert, E., Bartholomé, E., & Defourny, P. (2014). A near real-time water surface detection method based on HSV transformation of MODIS multi-spectral time series data. Remote sensing of environment, 140, 704-716, doi: 10.1016/j.rse.2013.10.008

Peng, B., Meng, Z., Huang, Q., & Wang, C. (2019). Patch Similarity Convolutional Neural Network for Urban Flood Extent Mapping Using Bi-Temporal Satellite Multispectral Imagery. Remote Sensing, 11(21), 2492.

Pham-Duc, B., Prigent, C.,
& Aires, F. (2017). Surface water monitoring within Cambodia and the Vietnamese Mekong Delta over a year, with Sentinel-1 SAR observations. Water, 9(6), 366. doi:10.3390/w9060366

Plank, S., Jüssi, M., Martinis, S., & Twele, A. (2017). Mapping of flooded vegetation by means of polarimetric Sentinel-1 and ALOS-2/PALSAR-2 imagery. International Journal of Remote Sensing, 38(13), 3831-3850, doi: 10.1080/01431161.2017.1306143

Potnis, A. V., Shinde, R. C., Durbha, S. S., & Kurte, K. R. (2019, July). Multi-class segmentation of urban floods from multispectral imagery using deep learning. In IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium (pp. 9741-9744). IEEE.

Rajah, P., Odindi, J., & Mutanga, O. (2018). Feature level image fusion of optical imagery and Synthetic Aperture Radar (SAR) for invasive alien plant species detection and mapping. Remote Sensing

Applications: Society and Environment, 10, 198-208, doi: 10.1016/j.rsase.2018.04.007
Rambour, C., Audebert, N., Koeniguer, E., Le Saux, B., Crucianu, M., & Datcu, M. (2020). Flood Detection in Time Series of Optical and SAR Images. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 43(B2), 1343-1346.
Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241). Springer, Cham. doi: 10.1007/978-3-319-24574-4_28
Saksena, Siddharth, and Venkatesh Merwade. "Incorporating the effect of DEM resolution and accuracy for improved flood inundation mapping." Journal of Hydrology 530 (2015): 180-194, doi: 10.1016/j.jhydrol.2015.09.069
Schmitt, M. (2020). Potential of Large-Scale Inland Water Body Mapping from Sentinel-1/2 Data on the Example of Bavaria's Lakes and Rivers. PFG–Journal of Photogrammetry, Remote Sensing and Geoinformation Science, 88, 271-289, doi: 10.1007/s41064-020-00111-2
Schratz, Patrick, Jannes Muenchow, Eugenia Iturritxa, Jakob Richter, and Alexander Brenning. 2019. "Hyperparameter Tuning and Performance Assessment of Statistical and Machine-Learning Algorithms Using Spatial Data." Ecological Modelling 406 (August): 109–20, doi: 10.1016/j.ecolmodel.2019.06.002.
Shen, X., Anagnostou, E. N., Allen, G. H., Brakenridge, G. R., & Kettner, A. J. (2019b). Near-real-time non-obstructed flood inundation mapping using synthetic aperture radar. Remote Sensing of Environment, 221, 302-315, doi: 10.1016/j.rse.2018.11.008
Shen, X., Wang, D., Mao, K., Anagnostou, E., & Hong, Y. (2019a). Inundation extent mapping by synthetic aperture radar: a review. Remote Sensing, 11(7), 879, doi: 10.3390/rs11070879
Slagter, B., Tsendbazar, N. E., Vollrath, A., & Reiche, J. (2020). Mapping wetland characteristics using temporally dense Sentinel-1 and Sentinel-2 data: A case study in the St. Lucia wetlands, South Africa. International Journal of Applied Earth Observation and Geoinformation, 86, 102009, doi: 10.1016/j.jag.2019.102009
Smith, A. R. (1978). Color gamut transform pairs. ACM Siggraph Computer Graphics, 12(3), 12-19.
Soergel, U., Thoennessen, U., & Stilla, U. (2003, May). Visibility analysis of man-made objects in SAR images. In 2003 2nd GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas (pp. 120-124). IEEE.
Tanguy, M., Chokmani, K., Bernier, M., Poulin, J., & Raymond, S. (2017). River flood mapping in urban areas combining Radarsat-2 data and flood return period data. Remote Sensing of Environment, 198, 442-459, doi: 10.1016/j.rse.2017.06.042
Tarpanelli, A., Santi, E., Tourian, M. J., Filippucci, P., Amarnath, G., & Brocca, L. (2018). Daily river discharge estimates by merging satellite optical sensors and radar altimetry through artificial neural network. IEEE Transactions on Geoscience and Remote Sensing, 57(1), 329-341, doi: 10.1109/TGRS.2018.2854625
Twele, A., Cao, W., Plank, S. and Martinis, S., 2016. Sentinel-1-based flood mapping: a fully automated processing chain. International Journal of Remote Sensing, 37(13), pp.2990-3004, doi: 10.1080/01431161.2016.1192304
Wiesnet, D.R., McGinnis, D.F., and Pritchard, J.A., 1974. Mapping of the 1973 Mississippi River Floods by the NOAA-2 Satellite. JAWRA Journal of the American Water Resources Association, 10(5), pp.1040-1049, doi: 10.1111/j.1752-1688.1974.tb00623.x
Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. International journal of remote sensing, 27(14), 3025-3033, doi: 10.1080/01431160600589179
Yang, L., Meng, X., & Zhang, X. (2011). SRTM DEM and its application advances. International Journal of Remote

Sensing, 32(14), 3875-3896, doi: 10.1080/01431161003786016

Yulianto, F., Sofan, P., Zubaidah, A., Sukowati, K. A. D., Pasaribu, J. M., & Khomarudin, M. R. (2015). Detecting areas affected by flood using multi-temporal ALOS PALSAR remotely sensed data in Karawang, West Java, Indonesia. Natural Hazards, 77(2), 959-985, doi: 10.1007/s11069-015-1633-x

Zheng, X., Maidment, D. R., Tarboton, D. G., Liu, Y. Y., & Passalacqua, P. (2018). GeoFlood: Large-Scale Flood Inundation Mapping Based on High-Resolution Terrain Analysis. Water Resources Research, 54(12), 10-013, doi: 10.1029/2018WR023457