A High-Resolution Flood Inundation Archive (2016–the Present) from Sentinel-1 SAR Imagery over CONUS

Qing Yang, Xinyi Shen, Emmanouil N. Anagnostou, Chongxun Mo, Jack R. Eggleston, and Albert J. Kettner

ABSTRACT: Most existing inundation inventories are based on surveys, news, or passive remote sensing imagery. Affected by spatiotemporal resolution or weather conditions, these inventories are limited in spatial details or coverage. Satellite synthetic aperture radar (SAR) data have recently enabled flood mapping at unprecedented spatiotemporal resolution. However, the bottleneck in producing SAR-based flood maps is the requirement of expert manual processing to maintain acceptable accuracy by most SAR-driven mapping techniques. To fill the vacancy, we generate a high-resolution (10 m) flood inundation dataset over the contiguous United States (CONUS) from nearly the entire Sentinel-1 SAR archive (from January 2016 to the present), using a recently developed automated Radar Produced Inundation Diary (RAPID) system. RAPID uses U.S. Geological Survey (USGS) water watch system and accumulated precipitation to identify SAR images that potentially overlap a flood event. The dataset include inundation events with the temporal scale from several days to months. Concluded from all 559 overlapping images in the period from 2016 to the first half of 2019, the comparison of the proposed dataset against the USGS Dynamic Surface Water Extent (DSWE) product yields an overall, user, producer agreements, and critical success index of 99.06%, 87.63%, 91.76%, and 81.23%, respectively, demonstrating the high accuracy of the proposed dataset and the robustness of the automated system. We anticipate this archive to facilitate many applications, including large-scale flood loss and risk assessment, and inundation model calibration and validation.

KEYWORDS: Databases; Radars/Radar observations; Remote sensing; Flood events

<https://doi.org/10.1175/BAMS-D-19-0319.1>

Corresponding author: Xinyi Shen, xinyi.shen@uconn.edu

In final form 31 December 2020

©2021 American Meteorological Society

For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#).

AFFILIATIONS: Yang—College of Civil Engineering and Architecture, Guangxi University, Nanning, Guangxi, China, and Department of Civil and Environmental Engineering, University of Connecticut, Storrs, Connecticut; Shen and Anagnostou—Department of Civil and Environmental Engineering, University of Connecticut, Storrs, Connecticut; Mo—College of Civil Engineering and Architecture, Guangxi University, Nanning, Guangxi, China; Eggleston—WMA Hydrologic Remote Sensing Branch, U.S. Geological Survey, Leetown, West Virginia; Kettner—Dartmouth Flood Observatory, Community Surface Dynamics Modeling System, Institute of Arctic and Alpine Research, University of Colorado Boulder, Boulder, Colorado

A high-spatial-resolution (5–10 m) database of inundation extent at the national scale is vital for analyzing inundation patterns and risk yet is currently not available. Survey- and report-based inundation inventories are limited in spatial details and dynamics (Gourley et al. 2013; NCEI 2016; Sahoo and Bhaskaran 2018). While optical imagery might be able to characterize large-scale and long-term flood events (Pekel et al. 2016; Jones. 2015, 2019; Li et al. 2018; Almukhtar et al. 2019), they are likely to underrepresent flood peaks or miscapture events shaded by the cloud, a common weather condition during flood events (Aires et al. 2018). While penetrating cloudy conditions, microwave radiometer imagery also offers inundation extents at relatively high temporal resolution (daily to weekly). However, the coarse spatial resolution they provide (5 to 25 km) (Prigent et al. 2007; Schroeder et al. 2014; Du et al. 2017) and the large uncertainties of the downscaled products (Chouinard et al. 2015; Du et al. 2018) render them inadequate to capture local-scale flooding. The recently emergent freely available Sentinel-1 synthetic aperture radar (SAR) imagery providing a reasonable spatiotemporal resolution (10 m, ~6 days) and is not disturbed by cloud cover (Prigent et al. 2016; Aires et al. 2017). Consequently, SAR imagery has gain popularity in delineating flood events. However, due to the algorithm complexity and the requirement of expert manual editing, existing flood archives only respond to emergency [European Commission Joint Research Centre (EC JRC); EC JRC 2015; JPL 2017] or a few major cases (Zeng et al. 2020; Cerrai et al. 2020). No method has yet facilitated a national-scale inundation extent dataset. This primarily because fully automated retrieval algorithms with acceptable accuracy have only been recently developed (Shen et al. 2019a), which has limited the use of these data in flood events.

In this study, we present an unprecedented 10-m-resolution flood inundation archive over the contiguous United States (CONUS), generated from the entire Sentinel-1 SAR archive for the period from January 2016 to the present, based on the Radar Produced Inundation Diary (RAPID) algorithm (Shen et al. 2019b). By combining radar statistics and machine-learning methods, with the integration of multisource remote sensing data and product, RAPID achieves full automation and high-level accuracy with zero manual postprocessing or expert knowledge. The RAPID system is driven by Sentinel-1 SAR imagery provided by the European Space Agency (ESA), which are the only freely available satellite SAR data with global coverage. By applying an automatic processing chain, the method could be further applied to more sources of SAR data, such as the soon to be launched Surface Water and Ocean Topography (SWOT) and NASA–Indian Space Research Organisation (ISRO) SAR (NISAR), which is expected to deliver the next generation of global high-quality surface water data (Frasson et al. 2019; NASA 2019).

The production of the procedure is described in the second section. The deliverables and final formation of the data are detailed in the third section. We assess the accuracy of the dataset by visual and quantitatively comparison with the National Oceanic and Atmospheric Administration (NOAA) event reports, the Federal Emergency Management Agency (FEMA) derived floodplain maps, and the water extent from the U.S. Geological Survey (USGS)

Dynamic Surface Water Extent (DSWE) product in the fourth section. More application and limitation of the dataset are discussed in the fifth section.

Automated flood mapping system

We establish a fully automated flood mapping system that requires no human interference from the initial flood events discovery to the final flood map production. The system consists of a mapping trigger and a kernel algorithm.

Triggering the event mapping. Considering the scarcity of the floods in both spatial and temporal domain and the big size of high-resolution SAR data, the system needs to know where or when the flood happens and which SAR image might overpass the inundation. For this purpose, we implement a triggering mechanism that relies on both in situ stream stage observations and satellite precipitation estimation to initially identify potential flooded zone (PFZ) (the maximal extent that may contain flood inundation) within which we acquire and process overpassing SAR images. The flood trigger detects two types of flooding, fluvial and pluvial, as depicted by Fig. 1. For the fluvial flooding, it applies the National Weather Service (NWS) flood stage threshold to USGS stream stage measurements to identify the daily flood status (flooded or unflooded) at around 4,455 stations. By subtracting the drainage areas draining to unflooded upstream stations from the flooded area pouring to a downstream station of flood status, the fluvial trigger confines the PFZ to the subwatershed level. The drainage area for a given outlet is delineated by running the watershed algorithm (Tarboton 1989, 1997) over the flow direction map at 30-arc-s resolution contained in the Hydrological Data and Maps based on Shuttle Elevation Derivatives at Multiple Scales (Lehner 2013). The trigger locates pluvial PFZ by applying a threshold to the maximal daily accumulated precipitation within a 3-day window, a possible delay between precipitation and runoff. We select 60 mm day^{-1} as the threshold according to the lower bound of moderate rainfall intensity (OFCM, 2019) to the IMERG (Huffman et al. 2015), a precipitation field of 0.1° horizontal resolution. Finally, the two types of PFZs are combined to form the daily PFZ.

We propose a clustering algorithm to delineate the maximum potential daily coverage of a flood event, based on the spatial proximity and temporal continuity of the daily PFZs. This algorithm traces daily PFZs to determine which PFZs can be merged into one flood event, is similar to the method used to delineate extreme hydrological events (e.g., Andreadis et al. 2005;

Fig. 1. The automated process of the RAPID system, from flood event discovery to production of an inundation map.

Sheffield et al. 2009; Zhan et al. 2016; He et al. 2020). To avoid to overly merge only marginally overlapped events, we apply the following rules for merging PFZs on consecutive days:

- 1) Merge two spatially disconnected PFZs into one if a pair of points exists in the two PFZs that their distance is equal to or less than 50 km.
- 2) For two PFZs on a day and the next, we associate them to the same event if the fraction of the intersected area is no less than 70% of the PFZ on either day.
- 3) Update the maximal flood extent by uniting all PFZs within the latest 5 days.
- 4) Terminate the event if the flood zone is less than 10% of the previous 5-day maximal flood extent.

Within a given PFZ, we acquire for retrieval processing the SAR images sensed on the day of flooding and, as persistent water body, multiple images obtained from the same Sentinel-1 ground track with a certain overlapping, sensed on previous dry days. The RAPID kernel algorithm requires approximately five dry-time images for each SAR image acquired on the flood day to reduce the error caused by noise-like speckle. Level-1 dual-polarized (VH + VV or HV + HH) Sentinel-1 SAR images in Interferometric Wide (IW) swath and strip map (SM) modes and Ground Range Detected (GRD) format are preprocessed via orbit correction, radiometric correction, and terrain correction using the Sentinel Application Platform (SNAP) and then normalized by the incidence angle using the cosine law (Mladenova et al. 2013).

The RAPID kernel algorithm. The preprocessed SAR image is regularized to $10\text{ m} \times 10\text{ m}$ grid resolution when inputting to the RAPID kernel algorithm for flood map delineation through four steps: 1) classification of water and land pixels based on statistics from polarimetric backscattering; 2) form water bodies by pixel connectivity, label them as connected or not connected to a known water body, and then apply different criterion set to reject less credible water bodies; 3) generation of a buffer region around the identified water bodies to reduce omission error using water mask derived from step 1 with less restrictive thresholds; and 4) correction of the pixel classification error based on their water probability predicted through a machine-learning approach. Besides the input SAR imagery, ancillary data include water surface occurrence, land-cover classification, topography, hydrography, and river bathymetry, as detailed in the RAPID kernel algorithm (Shen et al. 2019b).

The resulting inundation extent raster images are binary water masks, with pixels labeled as water or nonwater. Persistent water bodies are delineated as the maximal extent of the water masks on dry days. Therefore, a user can choose either to highlight the inundated area or to use the total water area.

Deliverables

To enable wide dissemination of the dataset, we archive it to a publicly accessible portal (<https://rapid-nrt-flood-maps.s3.amazonaws.com/index.html>) provided by the Amazon Web Service (AWS). The final product contains two subdatasets. The first subdataset is a flood event collection stored as multiple time series in an ESRI shapefile. Each series represents one event containing several days of multipolygon features with each represent the PFZ of a day. Each multipolygon feature contains a unique event ID, and the date as fields. The second subdataset contains binary flood extent raster files with each pixel labeled as 1 (flooded) or 0 (nonflooded). We also generate a separate list to associate the raster file name of each flood extent to the event ID to facilitate eventwise queries. We further link our archive to the Global Active Archive of Large Flood Events database produced by the Dartmouth Flood Observatory (DFO) (Brakenridge et al. 2010; Adhikari et al. 2010) to extend the flood death and displaced estimates caused by related events. The proposed dataset can, therefore, facilitate various applications, including flood

monitoring, inundation model calibration, and verification (Afshari et al. 2018; Zeng et al. 2020), flood damage and risk assessment (Wing et al. 2017), and mitigation management (Wing et al. 2020).

Results and discussion

Overview of flood events and maps.

The triggering system has detected 21,589 flood events from January 2016 to June 2019, with Fig. 2 showing the distribution of the duration and maximal extent. For these events, the system uses 1,897 SAR images acquired on flood days and 8,252 acquired on dry days out of 36,860 SAR images overpassing CONUS during the same period. Figure 3a and 3b provide the spatial distribution of fluvial and pluvial PFZs and Sentinel-1 acquisitions over CONUS. Significantly more frequent flooding occurred in Central and eastern CONUS, with the highest occurrence in the lower Mississippi River region (Fig. 3c). Although more Sentinel-1 images were acquired in the western than in the eastern CONUS, a smaller portion in the west overlapped with flooding (Figs. 3d,e). A total of only 635 triggered flood events were processed by the RAPID system for inundation map production, with a median duration and maximal triggering extent of 4 days and 8,900 km². The maximal PFZ of a flood event affects its probability of being captured by Sentinel-1, as plotted in Fig. 4. The archive

Fig. 2. Density-colored histogram plot of duration as a function of maximal extent area for all 21,589 events between January 2016 and June 2019.

Fig. 3. Accumulative count of grids ($0.1^\circ \times 0.1^\circ$) during the periods from January 2016 to June 2019 for (a) fluvial trigger based on the USGS stage gauging, (b) pluvial trigger based on the IMERG precipitation, (c) flood trigger by combining (a) and (b), (d) total Sentinel-1 acquisitions, and (e) acquisitions in days of flooding.

Fig. 4. The maximal potential flooded zone of an event vs its probability of being covered by Sentinel-1 images (blue line with marks); average Sentinel-1 acquisition intervals for each $0.1^\circ \times 0.1^\circ$ pixel (green columns); accumulative number of flooding SAR images in every 3 months (orange line with marks) from January 2016 to June 2019 over CONUS. The abruptly increase of potential flooded images in 2019 is related to the Midwestern Great Flood.

contains flood events ranging from shorter than a week (e.g., the Massachusetts flood event in March 2018) to as long as several months (such as the so-called 2019 Midwestern Great Flood). Events of shorter duration and smaller extent are more likely to be included in the more recent part of the archive because Sentinel-1 increased the average acquisition intervals from 9.6 days (2016–18) to 5.4 days (January–June 2019), as shown in Fig. 4. Additionally, in the presence of major flood events (e.g., Hurricane Harvey in 2017), the maximal revisiting frequency (1–2 days) might be implemented from the joint use of the Sentinel-1 constellation (*S1A* and *S1B*) if activated by ESA.

Mapping the dynamics of the event. One distinguishing feature of the proposed flood mapping system is the capability of capturing flood extent dynamics. Figure 5 presents four flood maps associated with a flood event near Sacramento, California occurred in February 2017. In Fig. 5, the spatial distribution of both the PFZs and inundation shows a clearly temporal trend of flood rising and recessing. Compared to the existing flood event dataset (Figs. 5a–d), the PFZs generated from the triggering system provide more situational awareness information about where the flood generally impacted and when comes the peak of the event. The high spatial resolution, temporal continuity, and automatic detection/generation of the inundation maps (Figs. 5e–h) form the unique advantages of the proposed flood mapping chain in near-real-time applications.

To further demonstrate the event formation and detection of inundation on major events, we choose four events—the 2019 Midwestern flood, Hurricane Florence (2018), Hurricane Harvey (2017), and Hurricane Matthew (2016)—as examples (Fig. 6). These events, each caused over a billion dollars of disaster loss, according to NCEI (2020). The maximal PFZ (the solid bold boundaries in Fig. 6) and the event duration for each are consistent with the NOAA or USGS event reports (Stewart 2017; Stewart and Berg 2019; Blake and Zelinsky 2018; NOAA 2019). It should be noted that, after a storm dissipates, inundation may remain for a

Fig. 5. (a)–(d) Dynamics of the potential flood zone and (e)–(h) the detail inundating area associated to a flood event happened between 18 Feb and 2 Mar 2017 at California. The solid green boundary is the event affected area according to Global Active Archive of Large Flood Events database [DFO event ID: 4344, data source: Brakenridge et al. (2010)].

while, making the duration of the flood event longer than that of the triggering storm. For visualization, the maximal inundated area is aggregated into $0.01^\circ \times 0.01^\circ$ grids (Fig. 6). The inundation fraction of a grid, an indicator of inundation severity, is defined as the inundated area (excluding persistent water bodies) over the total grid area. Moreover, the majority of high fraction locations are distributed along the river (Fig. 6a) and the coast (Figs. 6b–d), which agrees with the high stream level (NOAA 2019) and records of storm surge during the events (Stewart 2017; Stewart and Berg 2019; Blake and Zelinsky 2018), respectively. Such detailed information with respect to the inundation severity, extent, and duration of major flood events are not reported by existing databases.

Visual comparison with optical maps and hydraulic simulations. The visual comparison of the RAPID open water extent with the DSWE product (water with high and moderate confidence) shows strong overall agreement, with some differences in the regions where vegetation is concentrated (Fig. 7). For the area covered by woody plants (i.e., forest), the inundation on the ground surface can greatly increase the double-bounce backscattering (Lang et al. 2008; Shen et al. 2019a), featured as strong signals in a SAR image. In this environment, a pixel would be detected as obstructed by RAPID (Shen et al. 2019b), while DSWE may still classify the pixel as water or wetland depending on the open water fraction (Feyisa et al. 2014; Jones 2015), as shown in Fig. 7a. However, for a small river or stream across a forest, DSWE

Fig. 6. Four events from the inundation archive at regional scale with locations and duration of the formed events: the flood inundation fraction and the available SAR acquisition during the events of (a) Midwest flood and Hurricanes (b) Matthew, (c) Florence, and (d) Harvey. Hurricane best track and surge inundation estimation from NOAA reports were added in (b)–(d) for comparison. Characters inside black circles represent 0000 UTC intensity of hurricanes (D: tropical depression; S: tropical storm; 1–5: hurricane classification based on the Saffir–Simpson scale).

tends to misclassify it as partial surface water or even cloud shadow (see Fig. 7a). This might be related to the light absorption effect by the forest canopy layer and the limitation of Landsat's resolution (Huang et al. 2014; Jin et al. 2017; Jones 2019). In the area dominated by herbaceous vegetation (as shown in Fig. 7b), on the other hand, our map reports a larger area

Fig. 7. Inundation comparison between (left) proposed archive (RAPID) and (center) Dynamic Surface Water Extent (DSWE), during the event of (a) Midwestern Great Flood on 18 Mar 2019 and (b) Hurricane Harvey caused flood on 31 Aug 2017, at the presence of (right) different type of vegetation [National Wetland Inventory (NWI), data source: U.S. Fish and Wildlife Service (2020)]. In each panel, the extent of comparison was marked as the zoom in location in Fig. 6.

of open water since the surface backscattering property in those areas has higher chance to be confused with water bodies due to the low surface roughness.

Besides DSWE, the 100-yr floodplain delineated by FEMA using high-quality local hydraulic/hydrodynamic models (FEMA 2016), is selected to verify the proposed dataset in the frequency domain. As shown in Figs. 8a–c, most detected flooding areas are within the FEMA 100-yr floodplain because no floods were reported greater than 100 years (Stewart 2017; Stewart and Berg 2019; NOAA 2019). However, in Fig. 8d, a significant amount of floodwater is found outside of the 100-yr floodplain near the Huston region during Hurricane Harvey, which agrees with the experienced flow that exceeds 100 years return period (Blake and Zelinsky 2018).

Quantitatively assessment. To quantitatively evaluate the overall accuracy of the inundation archive, we compare the overlapping areas pixel by pixel using DSWE as the reference. Here, the “overlapping area” refers to the common pixels covered by both DSWE and the proposed dataset on the same day. We exclude any pixels identified as cloud, cloud shadow, shaded relief, missing pixels by the scanline corrector, and other types of error

Fig. 8. Inundation map (blue and red layer) overlap with FEMA-derived flood zone (yellow layer) at the local scale: (a) the Midwestern Great Flood on 14 Mar 2019, (b) Hurricane Matthew–caused flood on 11 Oct 2016, (c) Hurricane Florence–caused flood on 14 Sep 2018, (d) Hurricane Harvey–caused flood on 29 Aug 2017. In each panel, the extent of comparison or overlap was marked as the zoom in location in Fig. 6.

recorded by the DSWE mask band in the “overlapping area.” We resample the DSWE pixel to the resolution of Sentinel-1: 10 m × 10 m. Additionally, we also exclude pixels labeled in DSWE as potential wetland or water (wetland) with low confidence (Zanter 2019) for the comparison. Four error metrics are used in the assessment: overall agreement (OA), user agreement (UA), producer agreement (PA), and critical success index (CSI):

$$OA = \frac{TP + TN}{TP + FP + FN + TN}, \quad (1)$$

$$UA = \frac{TP}{TP + FP}, \quad (2)$$

$$PA = \frac{TP}{TP + FN}, \text{ and} \quad (3)$$

$$CSI = \frac{TP}{TP + FP + FN}, \quad (4)$$

where TP, TN, FP, and FN stand for the true-positive, true-negative, false-positive, and false-negative results from the confusion matrix (Wing et al. 2017; Shen et al. 2019b; Grimaldi et al. 2020), respectively, and positive (negative) represent the wet (dry) pixels. Analyzing over 73 billion pixels, the two datasets agree well across all 559 overlapping images, with the cumulative OA, UA, PA, and CSI, at 99.06%, 87.63%, 91.76%, and 81.23%, respectively (Fig. 9). The high value of PA indicates that over 90% of the DSWE open water area is

captured by our archive. Meanwhile, represented by UA, only less than 15% of water pixels are “false alarms.” An overall tendency toward more detection against DSWE is reported by the higher PA compared to the UA, while the OA and CSI still confirm the consistency. Note that the flood peaks are often coupled with cloudy weather, which are likely to be excluded in this evaluation. However, since the purpose is to validate the quality of RAPID instead of comparing the flood mapping capability of RAPID and DSWE, the result still sufficiently serves the purpose.

To further explore the performance of our flood maps over various types of land covers, we overlay the comparison pixels to the global land-cover classification dataset (Gong et al. 2019). As shown in Table 1, the pixels of error, i.e., FP and FN, are mainly in vegetated areas (cropland, grassland, and forest). More than three quarters of FP pixels are located in the croplands and grasslands while most FN pixels are reported in the forest area. Since RAPID only detects nonobstructed (submerged) inundation while DSWE can detect partial water surfaces, the forest area tends to have more FN. Since the roughness of grassland and cropland (before sprout or after harvest) can be low enough to be confused with water surfaces, the FP tend to occur there. Figure 10 provides the spatial distribution of the temporally averaged OA, UA,

Fig. 9. Boxplot of overall, user, producer, and critical success index of 559 images overlapping with Dynamic Surface Water Extent observed from January 2016 to June 2019.

Table 1. Percentage distribution of pixel classification results in regions with different types of land cover vs DSWE.

Land-cover type	TP	FP	FN	TN
Cropland	1.78%	51.44%	15.62%	22.07%
Forest	0.89%	6.69%	43.90%	43.42%
Grassland	1.12%	26.83%	13.69%	31.58%
Shrubland	0.06%	0.28%	0.45%	0.40%
Wetland	1.19%	2.08%	6.71%	0.09%
Water	94.57%	8.17%	14.48%	0.03%
Impervious surface	0.32%	3.15%	4.85%	2.22%
Bare land	0.06%	1.36%	0.30%	0.19%
Snow/ice	0.00%	0.01%	0.00%	0.00%

Fig. 10. Spatial distribution of measures of fit against Dynamic Surface Water Extent.

PA, and CSI (upscaled to $0.1^\circ \times 0.1^\circ$ grids for visualization). Moreover, relative low agreements could be found in the upper Missouri, lower Mississippi, southeast coastal, and New England (Figs. 10b,d) regions with high wetland density (Cohen et al. 2016).

Flood inundation pattern in the CONUS. Figures 11a and 11b show the maximal inundation fraction and occurrence (the number of inundated days) computed from the proposed archive at the national scale. For visualization, we upscaled the $10\text{ m} \times 10\text{ m}$ binary result to $0.1^\circ \times 0.1^\circ$. The final inundation occurrence and fraction delineated from SAR images show consistent spatial patterns to the PFZs, which is not limited by SAR overpasses (Fig. 3d), indicating that the dataset well captured the overall CONUS inundation pattern. Regions with high fraction are concentrated in the midwestern and the riverine area of the lower Mississippi. This finding agrees with a conclusion drawn from the long-term Landsat-based water occurrence dataset (Chouinard et al. 2015; Aires et al. 2018)—that is, transitory water areas are dominated by high water occurrence, resulting in higher flood vulnerability. Additionally, our archive shows that the areas along the Gulf Coast and the Atlantic Coast have higher inundation occurrences, most likely caused by the landfall of hurricanes. By comparing Fig. 11c to Figs. 11a and 11b, we observe that areas with more severe inundation caused by hurricanes generally report more flood insurance claims than regions flooded by the transitory water bodies, which indicates that fewer householders tend to build properties in vulnerable flood zones created by large inland rivers than in vulnerable flood zones along the coast. This finding also agrees with Kousky and Michel-Kerjan (2017). Weak correlation between the inundation severity and claim are found in coastal regions in North Carolina, Pennsylvania, and Southern California, where light to moderate inundation fraction can sometimes overlap with high claim

numbers. One possible explanation is the relatively low SAR acquisition in these areas, which only covers 45.72% of the total 287,439 claims reported from 2016 to 2019 (FEMA 2019). Direct evaluation of property damage in the building-level based on our archive is not recommended, even for the event with decent SAR coverage, since the RAPID algorithm only works for nonobstructed inundation (Shen et al. 2019b). Therefore, a user must utilize our dataset for inundation estimation in urban areas with a certain degree of caution.

Summary

This article describes a newly generated inundation archive of CONUS at 10 m resolution from January 2016 to June 2019. This archive, based on the long-term SAR images, has more comprehensive temporal and spatial coverage than currently available CONUS flood inundation datasets. Another feature is the association of the inundation extent with the event information. The accuracy of the dataset has been thoroughly validated against other products at the national scale. The database will be updated every 6 months to incorporate the flood events and inundation maps produced during the last half year.

Limitations and future developments.

Notwithstanding many advantages, the archive is affected by limitations that mostly due to the Sentinel-1 acquisition program and the backscattering nature of SAR sensors, which are listed as follows:

- 1) The current revisiting frequency of Sentinel-1 might not capture flood events (such as flash floods) lasting less than a few days.
- 2) Obstructed inundation, such as wetlands and flooding surrounding tall buildings, is also excluded from flood detection in the current version owing to remaining challenges in the retrieval algorithm.

Fig. 11. (a) Area fraction of maximal inundation (exclude persistent water bodies) over $0.1^\circ \times 0.1^\circ$ cells from January 2016 to June 2019; (b) total number of observed flooded days for each cell; (c) number of flood insurance claims from National Flood Insurance Program (data source: FEMA 2019) during January 2016 to June 2019. In (b), a day of a given cell is flooded if its inundation fraction exceeds 1%.

- 3) Over desert areas, bare surfaces may be mistakenly classified as water using SAR images, so we excluded those images from detection.
- 4) The accuracy of the archive has not been verified completely since the benchmark data (DSWE) have their own error and uncertainty; this will be further addressed and validated when other high-quality datasets become available to serve as the reference.

To capture small flood events and more flood dynamics, other sources of SAR satellites than Sentinel-1 can be utilized, for instance, Advanced Land Observation Satellite (ALOS-2), and NISAR (NASA 2019) upon their accessibility. All available data sources such as SWOT and RADARSAT-2, even with lower revisiting frequency or spatial continuity than the satellites mentioned above, could still strengthen the depiction of extreme and long-duration events (Frasson et al. 2019; Bolanos et al. 2016). We are also working on integrating high-resolution (30 m) river centerline and width from other products such as the global river width and discharge database (Frasson et al. 2019; Lin et al. 2020), and the Global Reach-Level A Priori Discharge Estimates for SWOT (Lin et al. 2019), to further increase the flood detectability. Moreover, more efforts are needed in the change detection-based approaches, to better distinguish the backscattering of vegetated area in dry and flooded conditions (Tsyganskaya et al. 2018; Shen et al. 2019a; Grimaldi et al. 2020). To enable reliable urban flood detection over CONUS, future studies might also integrate the big data of building information (e.g., geometry, orientation, and even material of buildings) into building flood signatures such as backscattering enhancement and the loss of coherence (Chini et al. 2019; Li et al. 2019). A new algorithm to enable flood mapping over desert areas utilizing SAR data is under development since flooding also happened there, though infrequently. For example, the Somalian desert, flooded twice between 2018 and 2020 (UNITAR 2018, 2020), and New South Wales experienced a flood event after the 2019 wildfire. Finally, we intend to extend the proposed archive to cover the globe, by incorporating the water fraction products generated from passive microwave satellites data (Schroeder et al. 2014; Du et al. 2018) or directly utilizing the real-time global flood events identified using the social media data (Bruijn et al. 2019).

Applications and opportunities. We anticipate that the product will directly facilitate the calibration and validation of inundation modeling and monitoring and promote flood event characterization (Shen et al. 2017a), flood risk analysis (Shen et al. 2016, 2017b; Wing et al. 2017), and damage assessment (Wing et al. 2020). We believe this dataset could further accelerate the process of fusion research with optical sensor-based flood information to obtain a global long-term high-resolution dataset of water extent, as suggested by Westerhoff et al. (2013) and Aires et al. (2018). Furthermore, A potential application of the database can be the improvement of characterizing floodplain topography using the methods by Mason et al. (2016) or by Shastry and Durand (2019). The database could also be a game changer for data-poor regions without DEM of eligible resolution for flood inundation simulations.

Acknowledgments. The authors of this publication had research facility support from Eversource Energy Center at the University of Connecticut. This research was partially funded by Natural Science Foundation of China (NSFC), Regional Science Program, 51969004, Innovation Project of Guangxi Graduate Education (YCBZ2018023 and YCBZ2019022). The authors thank the Amazon Web Service for supporting us with the archiving of the flood map dataset.

References

- Adhikari, P., Y. Hong, K. R. Douglas, D. B. Kirschbaum, J. Gourley, R. Adler, and G. R. Brakenridge, 2010: A digitized global flood inventory (1998–2008): Compilation and preliminary results. *Nat. Hazards*, **55**, 405–422, <https://doi.org/10.1007/s11069-010-9537-2>.
- Afshari, S., A. A. Tavakoly, M. A. Rajib, X. Zheng, M. L. Follum, E. Omranian, and B. M. Fekete, 2018: Comparison of new generation low-complexity flood inundation mapping tools with a hydrodynamic model. *J. Hydrol.*, **556**, 539–556, <https://doi.org/10.1016/j.jhydrol.2017.11.036>.
- Aires, F., L. Miolane, C. Prigent, B. Pham, E. F. Chouinard, B. Lehner, and F. Papa, 2017: A global dynamic long-term inundation extent dataset at high spatial resolution derived through downscaling of satellite observations. *J. Hydrometeorol.*, **18**, 1305–1325, <https://doi.org/10.1175/JHM-D-16-0155.1>.
- , C. Prigent, E. F. Chouinard, D. Yamazaki, F. Papa, and B. Lehner, 2018: Comparison of visible and multi-satellite global inundation datasets at high-spatial resolution. *Remote Sens. Environ.*, **216**, 427–441, <https://doi.org/10.1016/j.rse.2018.06.015>.
- Almukhtar, S., B. Migliozzi, J. Schwartz, and J. Williams, 2019: The great flood of 2019: A complete picture of a slow-motion disaster. *New York Times*, 11 September, www.nytimes.com/interactive/2019/09/11/us/midwest-flooding.html.
- Andreadis, K. M., E. A. Clark, A. W. Wood, A. F. Hamlet, and D. P. Lettenmaier, 2005: Twentieth-century drought in the conterminous United States. *J. Hydrometeorol.*, **6**, 985–1001, <https://doi.org/10.1175/JHM450.1>.
- Blake, E. S., and D. A. Zelinsky, 2018: Tropical cyclone report: Hurricane Harvey (AL092017). NHC Tech. Rep., 77 pp., www.nhc.noaa.gov/data/tcr/AL092017_Harvey.pdf.
- Bolanos, S., D. Stiff, B. Brisco, and A. Pietroniro, 2016: Operational surface water detection and monitoring using RADARSAT 2. *Remote Sens.*, **8**, 285, <https://doi.org/10.3390/rs8040285>.
- Brakenridge, G. R., and Coauthors, 2010: Global Active Archive of Large Flood Events. University of Colorado Dartmouth Flood Observatory, accessed 20 September 2020, <http://floodobservatory.colorado.edu/Archives/index.html>.
- Bruijn, J. A., H. Moel, B. Jongman, M. C. Ruiter, J. Wagemaker, and J. C. J. H. Aerts, 2019: A global database of historic and real-time flood events based on social media. *Sci. Data*, **6**, 311, <https://doi.org/10.1038/s41597-019-0326-9>.
- Cerrai, D., Q. Yang, X. Shen, M. Koukoulou, and E. N. Anagnostou, 2020: Brief communication: Hurricane Dorian: Automated near-real-time mapping of the “unprecedented” flooding in the Bahamas using synthetic aperture radar. *Nat. Hazards Earth Syst. Sci.*, **20**, 1463–1468, <https://doi.org/10.5194/nhess-20-1463-2020>.
- Chini, M., R. Pelich, L. Pulvirenti, N. Pierdicca, R. Hostache, and P. Matgen, 2019: Sentinel-1 InSAR coherence to detect floodwater in urban areas: Houston and Hurricane Harvey as a test case. *Remote Sens.*, **11**, 107, <https://doi.org/10.3390/rs11020107>.
- Chouinard, E. F., B. Lehner, L. M. Rebelo, F. Papa, and S. K. Hamilton, 2015: Development of a global inundation map at high spatial resolution from topographic downscaling of coarse-scale remote sensing data. *Remote Sens. Environ.*, **158**, 348–361, <https://doi.org/10.1016/j.rse.2014.10.015>.
- Cohen, M. J., and Coauthors, 2016: Do geographically isolated wetlands influence landscape functions? *Proc. Natl. Acad. Sci. USA*, **113**, 1978–1986, <https://doi.org/10.1073/pnas.1512650113>.
- Du, J., J. S. Kimball, L. A. Jones, Y. Kim, J. Glassy, and J. D. Watts, 2017: A global satellite environmental data record derived from AMSR-E and AMSR2 microwave Earth observations. *Earth Syst. Sci. Data*, **9**, 791–808, <https://doi.org/10.5194/essd-9-791-2017>.
- , J. Galantowicz, S. B. Kim, S. K. Chan, R. Reichle, L. A. Jones, and J. D. Watts, 2018: Assessing global surface water inundation dynamics using combined satellite information from SMAP, AMSR2 and Landsat. *Remote Sens. Environ.*, **213**, 1–17, <https://doi.org/10.1016/j.rse.2018.04.054>.
- EC JRC, 2015: Copernicus Emergency Management Service. Accessed 21 August 2020, <https://emergency.copernicus.eu/>.
- Federal Emergency Management Agency, 2016: FEMA policy standards for flood risk analysis and mapping. Federal Emergency Management Agency Doc., 79 pp., [www.fema.gov/media-library-data/1581015527397-57f85306e2a67abc6a852298a9bcae77/Standards_for_Flood_Risk_Projects_\(Nov2016\).pdf](http://www.fema.gov/media-library-data/1581015527397-57f85306e2a67abc6a852298a9bcae77/Standards_for_Flood_Risk_Projects_(Nov2016).pdf).
- , 2019: NFIP redacted claims data set. Accessed 20 April 2020, www.fema.gov/media-library/assets/documents/180374.
- Feyisa, G. L., H. Meilby, R. Fensholt, and S. R. Proud, 2014: Automated water extraction index: A new technique for surface water mapping using Landsat imagery. *Remote Sens. Environ.*, **140**, 23–35, <https://doi.org/10.1016/j.rse.2013.08.029>.
- Frasson, R. P. d. M., G. J. P. Schumann, A. J. Kettner, G. R. Brakenridge, and W. F. Krajewski, 2019a: Will the Surface Water and Ocean Topography (SWOT) satellite mission observe floods? *Geophys. Res. Lett.*, **46**, 10435–10445, <https://doi.org/10.1029/2019GL084686>.
- , and Coauthors, 2019b: Global relationships between river width, slope, catchment area, meander wavelength, sinuosity, and discharge. *Geophys. Res. Lett.*, **46**, 3252–3262, <https://doi.org/10.1029/2019GL082027>.
- Gong, P., and Coauthors, 2019: Stable classification with limited sample: Transferring a 30-m resolution sample set collected in 2015 to mapping 10-m resolution global land cover in 2017. *Sci. Bull.*, **64**, 370–373, <https://doi.org/10.1016/j.scib.2019.03.002>.
- Gourley, J. J., and Coauthors, 2013: A unified flash flood database across the United States. *Bull. Amer. Meteor. Soc.*, **94**, 799–805, <https://doi.org/10.1175/BAMS-D-12-00198.1>.
- Grimaldi, S., J. Xu, Y. Li, V. R. N. Pauwels, and J. P. Walker, 2020: Flood mapping under vegetation using single SAR acquisitions. *Remote Sens. Environ.*, **237**, 111582, <https://doi.org/10.1016/j.rse.2019.111582>.
- He, X., M. Pan, Z. Wei, E. F. Wood, and J. Sheffield, 2020: A global drought and flood catalogue from 1950 to 2016. *Bull. Amer. Meteor. Soc.*, **101**, E508–E535, <https://doi.org/10.1175/BAMS-D-18-0269.1>.
- Huang, C., Y. Peng, M. Lang, I. Y. Yeo, and G. McCarty, 2014: Wetland inundation mapping and change monitoring using Landsat and airborne lidar data. *Remote Sens. Environ.*, **141**, 231–242, <https://doi.org/10.1016/j.rse.2013.10.020>.
- Huffman, G. J., and Coauthors, 2015: NASA Global Precipitation Measurement (GPM) Integrated Multi-Satellite Retrievals for GPM (IMERG), version 4. NASA Algorithm Theoretical Basis Doc., 30 pp., https://gpm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V4.5.pdf.
- Jin, H., C. Huang, M. W. Lang, I. Y. Yeo, and S. V. Stehman, 2017: Monitoring of wetland inundation dynamics in the Delmarva Peninsula using Landsat time-series imagery from 1985 to 2011. *Remote Sens. Environ.*, **190**, 26–41, <https://doi.org/10.1016/j.rse.2016.12.001>.
- Jones, J. W., 2015: Efficient wetland surface water detection and monitoring via Landsat: Comparison with in situ data from the everglades depth estimation network. *Remote Sens.*, **7**, 12 503–12 538, <https://doi.org/10.3390/rs70912503>.
- , 2019: Improved automated detection of subpixel-scale inundation—Revised Dynamic Surface Water Extent (DSWE) partial surface water tests. *Remote Sens.*, **11**, 374, <https://doi.org/10.3390/rs11040374>.
- JPL, 2017: Advanced Rapid Imaging and Analysis Project. NASA, accessed 21 August 2020, <https://aria.jpl.nasa.gov/>.
- Kousky, C., and E. Michel-Kerjan, 2017: Examining flood insurance claims in the United States: Six key findings. *J. Risk Insur.*, **84**, 819–850, <https://doi.org/10.1111/jori.12106>.
- Lang, M. W., E. S. Kasischke, S. D. Prince, and K. W. Pittman, 2008: Assessment of C-band synthetic aperture radar data for mapping and monitoring coastal plain forested wetlands in the mid-Atlantic region, U.S.A. *Remote Sens. Environ.*, **112**, 4120–4130, <https://doi.org/10.1016/j.rse.2007.08.026>.
- Lehner, B., 2013: HydroSHEDS technical documentation version 1.2. World Wildlife Fund Tech. Doc. 26 pp., https://hydrosheds.org/images/inpages/HydroSHEDS_TechDoc_v1_2.pdf.
- Li, S., and Coauthors, 2018: Automatic near real-time flood detection using Suomi-NPP/VIIRS data. *Remote Sens. Environ.*, **204**, 672–689, <https://doi.org/10.1016/j.rse.2017.09.032>.

- Li, Y., S. Martinis, M. Wieland, S. Schlaffer, and R. Natsuaki, 2019: Urban flood mapping using SAR intensity and interferometric coherence via Bayesian network fusion. *Remote Sens.*, **11**, 2231, <https://doi.org/10.3390/rs11192231>.
- Lin, P., and Coauthors, 2019: Global reconstruction of naturalized river flows at 2.94 million reaches. *Water Resour. Res.*, **55**, 6499–6516, <https://doi.org/10.1029/2019WR025287>.
- , M. Pan, G. H. Allen, R. P. de Frasson, Z. Zeng, D. Yamazaki, and E. F. Wood, 2020: Global estimates of reach-level bankfull river width leveraging big data geospatial analysis. *Geophys. Res. Lett.*, **47**, e2019GL086405, <https://doi.org/10.1029/2019GL086405>.
- Mason, D. C., M. Trigg, J. Garcia-Pintado, H. L. Cloke, J. C. Neal, and P. D. Bates, 2016: Improving the TanDEM-X Digital Elevation Model for flood modelling using flood extents from synthetic aperture radar images. *Remote Sens. Environ.*, **173**, 15–28, <https://doi.org/10.1016/j.rse.2015.11.018>.
- Mladenova, I. E., T. J. Jackson, R. Bindlish, and S. Hensley, 2013: Incidence angle normalization of radar backscatter data. *IEEE Trans. Geosci. Remote Sens.*, **51**, 1791–1804, <https://doi.org/10.1109/TGRS.2012.2205264>.
- NASA, 2019: NASA-ISRO SAR (NISAR) mission science users' handbook. NASA Doc., 139 pp., https://nisar.jpl.nasa.gov/system/documents/files/26_NISAR_FINAL_9-6-19.pdf.
- NCEI, 2016: Storm events database. NOAA, accessed 15 April 2020, www.ncdc.noaa.gov/stormevents/ftp.jsp.
- , 2020: U.S. billion-dollar weather and climate disasters. NOAA, <https://doi.org/10.25921/stkw-7w73>.
- NOAA, 2019: Spring flooding summary 2019. NOAA, www.weather.gov/dvn/summary_SpringFlooding_2019.
- OFCM, 2019: Surface weather observations and reports. OFCM Federal Meteorological Handbook 1, 101 pp., www.ofcm.gov/publications/fmh/FMH1/fmh1_2019.pdf.
- Pekel, J. F., A. Cottam, N. Gorelick, and A. S. Belward, 2016: High-resolution mapping of global surface water and its long-term changes. *Nature*, **540**, 418–422, <https://doi.org/10.1038/nature20584>.
- Prigent, C., F. Papa, F. Aires, W. B. Rossow, and E. Matthews, 2007: Global inundation dynamics inferred from multiple satellite observations, 1993–2000. *J. Geophys. Res.*, **112**, D12107, <https://doi.org/10.1029/2006JD007847>.
- , D. P. Lettenmaier, F. Aires, and F. Papa, 2016: Toward a high resolution monitoring of continental surface water extent and dynamics, at global scale: From GIEMS (Global Inundation Extent from Multi-Satellites) to SWOT (Surface Water Ocean Topography). *Surv. Geophys.*, **37**, 339–355, <https://doi.org/10.1007/s10712-015-9339-x>.
- Sahoo, B., and P. K. Bhaskaran, 2018: A comprehensive data set for tropical cyclone storm surge-induced inundation for the east coast of India. *Int. J. Climatol.*, **38**, 403–419, <https://doi.org/10.1002/joc.5184>.
- Schroeder, R., K. C. McDonald, B. Chapman, K. Jensen, E. Podest, Z. Tessler, T. J. Bohn, and R. Zimmerman, 2014: Development and evaluation of a multi-year inundated land surface data set derived from active/passive microwave remote sensing data. *Remote Sens.*, **7**, 16668–16732, <https://doi.org/10.3390/rs71215843>.
- Shastri, A., and M. Durand, 2019: Utilizing flood inundation observations to obtain floodplain topography in data-scarce regions. *Front. Earth Sci.*, **6**, 243, <https://doi.org/10.3389/feart.2018.00243>.
- Sheffield, J., K. M. Andreadis, E. F. Wood, and D. P. Lettenmaier, 2009: Global and continental drought in the second half of the twentieth century: Severity–area–duration analysis and temporal variability of large-scale events. *J. Climate*, **22**, 1962–1981, <https://doi.org/10.1175/2008JCLI2722.1>.
- Shen, X., H. J. Vergara, E. I. Nikolopoulos, E. N. Anagnostou, Y. Hong, Z. Hao, K. Zhang, and K. Mao, 2016: GDBC: A tool for generating global-scale distributed basin morphometry. *Environ. Modell. Software*, **83**, 212–223, <https://doi.org/10.1016/j.envsoft.2016.05.012>.
- , Y. Mei, and E. N. Anagnostou, 2017a: A comprehensive database of flood events in the contiguous United States from 2002 to 2013. *Bull. Amer. Meteor. Soc.*, **98**, 1493–1502, <https://doi.org/10.1175/BAMS-D-16-0125.1>.
- , E. N. Anagnostou, Y. Mei, and Y. Hong, 2017b: A global distributed basin morphometric dataset. *Sci. Data*, **4**, 160124, <https://doi.org/10.1038/sdata.2016.124>.
- , D. Wang, K. Mao, E. N. Anagnostou, and Y. Hong, 2019a: Inundation extent mapping by synthetic aperture radar: A review. *Remote Sens.*, **11**, 879–895, <https://doi.org/10.3390/rs11070879>.
- , E. N. Anagnostou, G. H. Allen, G. R. Brakenridge, and A. J. Kettner, 2019b: Near real-time non-obstructed flood inundation mapping using synthetic aperture radar. *Remote Sens. Environ.*, **221**, 302–315, <https://doi.org/10.1016/j.rse.2018.11.008>.
- Stewart, S. R., 2017: Tropical cyclone report: Hurricane Matthew (AL142016). NHC Tech. Rep., 96 pp., www.nhc.noaa.gov/data/tcr/AL142016_Matthew.pdf.
- , and R. Berg, 2019: Tropical cyclone report: Hurricane Florence (AL062018). NHC Tech. Rep., 98 pp., www.nhc.noaa.gov/data/tcr/AL062018_Florence.pdf.
- Tarboton, D. G., 1989: The analysis of river basins and channel networks using digital terrain data. Ph.D. thesis, Massachusetts Institute of Technology, 252 pp.
- , 1997: A new method for the determination of flow directions and upslope areas in grid digital elevation models. *Water Resour. Res.*, **33**, 309–319, <https://doi.org/10.1029/96WR03137>.
- Tsyganskaya, V., S. Martinis, P. Marzahn, and R. Ludwig, 2018: SAR-based detection of flooded vegetation—a review of characteristics and approaches. *Int. J. Remote Sens.*, **39**, 2255–2293, <https://doi.org/10.1080/01431161.2017.1420938>.
- UNITAR, 2018: International charter space and major disasters activation for floods in Somalia. Accessed 7 May 2018, <https://disasterscharter.org/web/guest/activations/-/article/flood-in-somalia-activation-570->.
- , 2020: International charter space and major disasters activation for floods in Somalia. Accessed 28 April 2020, <https://disasterscharter.org/web/guest/activations/-/article/flood-large-in-somalia-activation-650->.
- U.S. Fish and Wildlife Service, 2020: National wetlands inventory. Accessed 10 June 2020, www.fws.gov/wetlands/Data/Data-Download.html.
- Westerhoff, R. S., M. P. H. Kleuskens, H. C. Winsemius, H. J. Huizinga, G. R. Brakenridge, and C. Bishop, 2013: Automated global water mapping based on wide-swath orbital synthetic-aperture radar. *Hydrol. Earth Syst. Sci.*, **17**, 651–663, <https://doi.org/10.5194/hess-17-651-2013>.
- Wing, O. E. J., P. D. Bates, C. C. Sampson, A. M. Smith, K. A. Johnson, and T. A. Erickson, 2017: Validation of a 30 m resolution flood hazard model of the conterminous United States. *Water Resour. Res.*, **53**, 7968–7986, <https://doi.org/10.1002/2017WR020917>.
- , N. Pinter, P. D. Bates, and C. Kousky, 2020: New insights into US flood vulnerability revealed from flood insurance big data. *Nat. Commun.*, **11**, 1444, <https://doi.org/10.1038/s41467-020-15264-2>.
- Zanter, K., 2019: Landsat Dynamic Surface Water Extent (DSWE) product guide, version 3.0. USGS Rep., 21 pp., https://prd-wret.s3.us-west-2.amazonaws.com/assets/palladium/production/s3fs-public/atoms/files/LSDS-1331-Landsat-DynamicSurfaceWaterExtent-DSWE-ProductGuide-v3.0_%202019_03_19.pdf.
- Zeng, Z., Y. Gan, A. J. Kettner, Q. Yang, C. Zeng, G. R. Brakenridge, and Y. Hong, 2020: Towards high resolution flood monitoring: An integrated methodology using passive microwave brightness temperatures and Sentinel synthetic. *J. Hydrol.*, **582**, 124377, <https://doi.org/10.1016/j.jhydrol.2019.124377>.
- Zhan, W., K. Guan, J. Sheffield, and E. F. Wood, 2016: Depiction of drought over sub-Saharan Africa using reanalyses precipitation data sets. *J. Geophys. Res. Atmos.*, **121**, 10555–10574, <https://doi.org/10.1002/2016JD024858>.