Towards Operational Flood Monitoring in Flanders Using Sentinel-1

Lisa Landuyt ⓘ, Frieke M. B. Van Coillie ⓘ, Bram Vogels, Joost Dewelde, and Niko E. C. Verhoest ⓘ

Abstract—As floods pose an increasing threat to our society, insights into their occurrence and dynamics are of major importance for emergency relief, damage assessment, the optimization of predictive models, and spatial planning. Due to their capability of providing synoptic observations independent of cloud cover and daylight, synthetic-aperture radar (SAR) sensors are an invaluable tool for flood mapping and monitoring. In this study, the potential of SAR, and more specifically Sentinel-1, for automated flood monitoring in Flanders is assessed. Its capability to detect floods with varying characteristics is investigated, and an approach for automated monitoring is presented. This approach, combining thresholding and region growing, requires a SAR image pair and several ancillary data layers, including elevation, land cover, and flood risk, as input. The resulting map discriminates permanent water, open flooding, long-term flooding, possible flooding, flooded vegetation, and possibly flooded forests from dry land. Invisible forested areas are indicated as well. A quantitative and qualitative accuracy assessment, based on 17 and 138 flood maps, respectively, highlights the approach's robustness and improved accuracy compared to benchmark techniques. Furthermore, main sources of confusion are identified and suggestions for future improvements are listed.

Index Terms—Change detection, floods, monitoring, synthetic-aperture radar (SAR), Sentinel-1.

I. INTRODUCTION

CLIMATE change is affecting our environment and society in various ways. Amongst others, the intensification of the hydrological cycle is leading to more extreme drought and precipitation events, and eventually an increase in the frequency and intensity of flood events [1]–[3]. As these already are the most frequently occurring of all natural disasters [4], insights into their occurrence and dynamics are of paramount importance for an effective disaster management as well as for the calibration

and validation of flood prediction models, and the optimization of spatial planning [5]–[7]. Thanks to their ability to provide systematic, synoptic and timely observations, spaceborne remote sensing systems have become the main observation source for large scale flood events [8], [9].

Whereas optical sensors can be hampered by cloud cover, which is often persistent during flood events, synthetic aperture radar or synthetic-aperture radar (SAR) sensors provide observations independent of cloud cover and daylight. As such, they are an invaluable tool for flood mapping and monitoring. However, most operational SAR sensors do not acquire data systematically and/or do not provide freely available imagery. In this context, the Sentinel-1 constellation, launched in 2014 and 2016 by the European Space Agency (ESA) in the frame of the European Union's Copernicus Programme, was a true game changer. Consisting of two identical satellites (Sentinel-1A and Sentinel-1B), it provides freely available, dual-polarized imagery with a repeat time of 6 to 12 days across the globe [10]. As such, it has unlocked the potential of SAR for automated, near real-time flood monitoring even further.

Throughout the past years, a vast number of SAR-based flood mapping approaches has been developed [11]. These can be categorized based on the input data as well as on the classification techniques used. With respect to input, a single SAR scene [12]–[14], an image pair [15]–[17] or a time series [18]–[20] can be used. These SAR data are mostly complemented by ancillary data, such as topography, land cover or flood recurrence maps [12], [21]–[23]. Data fusion with optical data has been suggested too [24], [25], while interferometric and polarimetric SAR are mainly used to improve flood mapping in urban [26], [27] and vegetated areas [22], [28]. Classification techniques include but are not limited to automated thresholding [21], [29], probability distribution function (PDF) fitting [15], [30], active contour models [31], [32], region growing [15], [33], graph-based techniques [34], [35], fuzzy theory [16], [36], and supervised machine learning techniques [37], [38]. However, the majority of these approaches require user intervention and/or use parameters optimized for a specific case, thus hampering their applicability for automated flood monitoring.

A limited number of flood mapping approaches has been developed specifically for operational purposes. Westerhoff *et al.* [18] suggested a globally applicable approach providing backscatter-based flood probabilities, making use of locally fitted histograms based on training data generated using a 3-year time series of Envisat ASAR data and the NASA Shuttle Radar Topography Mission (SRTM) water body data [39]. Martinis *et*

Niko E. C. Verhoest is with the Hydro-Climate Extremes Lab (H-CEL), Ghent University, 9000 Ghent, Belgium (e-mail: niko.verhoest@ugent.be).

Lisa Landuyt is with the Hydro-Climate Extremes Lab (H-CEL), Ghent University, 9000 Ghent, Belgium, and also with the Remote Sensing | Spatial Analysis Lab (REMOSA), Ghent University, 9000 Ghent, Belgium (e-mail: lisa.landuyt@ugent.be).

Frieke M. B. Van Coillie is with the Remote Sensing | Spatial Analysis Lab (REMOSA), Ghent University, 9000 Ghent, Belgium (e-mail: frieke.vancoillie@ugent.be).

Bram Vogels and Joost Dewelde are with the Flanders Environment Agency, 9300 Aalst, Belgium (e-mail: b.vogels@vmm.be; j.dewelde@vmm.be).
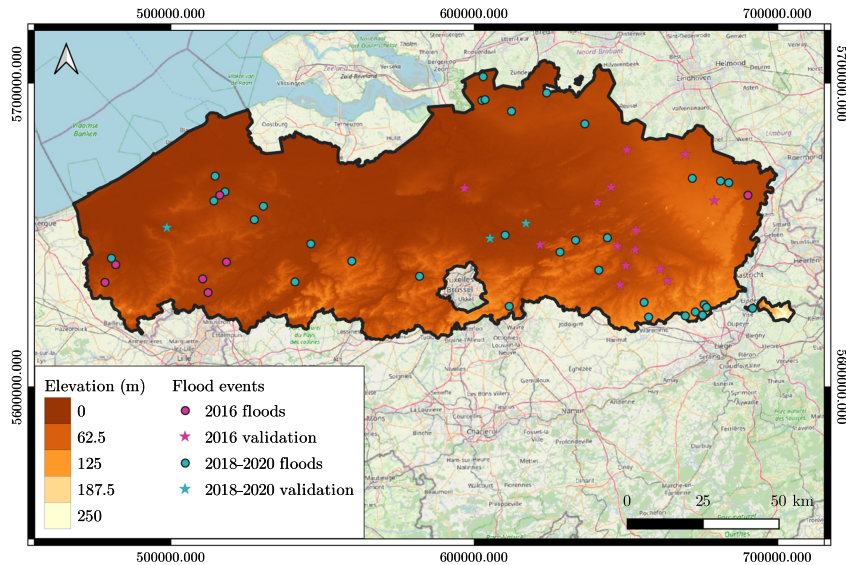
Fig. 1. Situation of the ROI and the flood events considered in this study. Floods events that occurred in 2016 are indicated in pink while those that occurred in 2018–2020 are indicated in blue. Flood events for which a quantitative validation is done are indicated by means of a star, while the others are indicated by means of a circle. Within the ROI, elevation is displayed. Used coordinate system: WGS84 UTM zone 29 N (EPSG:32631).

al. [40] and Twele *et al.* [29] developed approaches based on automated thresholding on single scenes and a fuzzy logic-based refinement for resp. TerraSAR-X and Sentinel-1 imagery. Chini *et al.* [41] automated an approach combining thresholding, region growing and change detection [15] by means of a hierarchical split-based approach. Finally, Shen *et al.* [20] suggested an approach based on PDF fitting, multilevel thresholding, morphological processing, and a machine learning-based refinement. These approaches consider full SAR scenes as input and are thus suited for automated flood mapping on large areas across the globe. Yet, these methods do not fully take advantage of locally available ancillary data and might miss small scale flooding. The region of interest (ROI) of this study, i.e., Flanders, is a rather small region with a heavily regulated river system. As a result, large-scale fluvial flooding occurs only very rarely. Moreover, rivers and retention basins are intensely monitored. Despite the high number of near real-time measurements, small-scale pluvial flooding remains hard to monitor. To inventory these, reports from local water managers and photographs made by residents are mostly used. Recently, the Flanders Environment Agency (Vlaamse Milieumaatschappij, VMM) also started to use drones to map critical flooding. Although providing high resolution observations, this is a very time-consuming task. Moreover, in case of widespread flooding, only a selection of locations can be observed. Therefore, this study aims to investigate to what extent Sentinel-1 imagery can contribute to flood monitoring in Flanders. The research objective is twofold as follows.

1) To investigate to what extent both small and large-scale flooding in Flanders is detectable using Sentinel-1, a medium resolution C-band SAR sensor, and identify the main factors influencing the detectability.
2) To investigate to what extent flooding in Flanders can be mapped in an automated way, and to suggest a flood mapping approach that fully takes advantage of locally

available data and can be deployed for operational flood monitoring.

II. STUDY AREA AND DATA

A. Study Area

The ROI of this study is Flanders, a region in the north of Belgium, displayed in Fig. 1. It is a rather small (ca. 13 600 km^2), densely populated area, and has a heavily fragmented and heterogeneous landscape compared to other regions in the world. Agricultural fields are the main land cover type, forests are small and dispersed. Topographical variations are limited, with heights varying between 0 and 290 m. As mentioned earlier, the hydraulic system is heavily monitored and regulated. With impervious surfaces accounting for about 16% of the total surface, a dense sewage and channel system and sluices in critical locations, precipitation is usually discharged fast toward the main rivers. Moreover, several retention basins have been installed to protect inhabited and agricultural zones from flooding. Both these retention basins and rivers are intensely monitored by means of water level measurements.

B. SAR and Ancillary Data

Since its launch in 2014 and 2016, the Sentinel-1 constellation provides C-band SAR imagery with repeat times of 6 to 12 days globally. In Europe, the repeat time is 6 days, while some sites can be revisited in less than a day. However, care should be taken when analyzing or comparing images from different orbits, as differences in incidence angle can severely affect backscatter intensities and geometric effects. Over land, interferometric wide swath is the main mode, producing dual-polarized (VV, VH) imagery distributed as single look complex (SLC) or ground range detected (GRD) product. The Terrascope

platform, developed and maintained by the Flemish Institute for Technological Research (Vlaams Instituut voor Technologisch Onderzoek, VITO) provides the Sentinel-1 GRD data covering Flanders in a ready-to-use format. A processing chain comprising orbit correction, border noise removal, thermal noise removal, radiometric calibration, and range-Doppler terrain correction is applied to all scenes, and the resulting imagery with a pixel spacing of 10 m is provided in GTIFF format. For all flood events, a Sentinel-1 image pair is selected in order to allow change detection. In this study, several isolated events that took place throughout 2018–2020, as well as a large flood event that took place in June 2016 are analyzed. For the 35 isolated events, the Sentinel-1 image acquired closest in time but after the preceding rainfall event was used. For the 2016 event, all Sentinel-1 images in the critical period (26th May—9th June) were analyzed for all 21 communities where flood damages were reported. To minimize computation time and complexity, the image acquired from the same orbit track and 12 days prior to the flood image is used as the reference image. This choice could be optimized in the future by means of an automated selection procedure [17], [42].

A vast number of high resolution thematic layers is available for the region of Flanders. A selection of these is used to refine the flood maps. A digital terrain model (DTM) constructed based on LiDAR measurements acquired between 2013 and 2015 is available at multiple resolutions (1, 5, 25, 100 m). Additionally, a land cover map based on orthophoto imagery acquired in 2015 is available at 1 and 5-m resolution. In order to match the resolution of the Sentinel-1 data, both 5-m products are resampled to 10 m. Moreover, several flood risk layers are available at high resolution. These include a flood susceptibility map based on flood return periods, discriminating between zones not prone to flooding and zones with a small (1000-y return), medium (100-y return) and large (10-y return) chance of flooding. For each return period, a flood depth map is provided as well, discriminating between 1–25, 25–50, 50–100, 100–200, and >200-cm water depth. Separate layers are available for pluvial, fluvial, and coastal flooding. The fluvial maps are made based on a combination of the probability distributed model [43] or Nedbør-Afstrømnings model (NAM-model) [44], depending on the type of stream, and a hydraulic model constructed using Infoworks-RS or MIKE11, while for the pluvial maps direct rainfall modeling [45] was used. All maps, with an original resolution of 2 m, were resampled to 10 m. The flood mapping approach described in Section III-B requires a single flood risk and a single flood depth layer. The former was calculated as the maximum of the fluvial and pluvial flood risk, while the latter was calculated as the maximum of the fluvial and pluvial flood depth linked to a small chance of flooding.

C. Validation Data

Since 2018, VMM has deployed a drone team to map floods, in most cases caused by heavy precipitation. The RGB drone imagery of 35 flood events, that occurred in the period 2018–2020, is used in this study to investigate the flood detection capabilities of Sentinel-1 and assess the accuracy of the automated flood

mapping approaches. From the 35 flood events captured by drone footage, only three events were also captured by a Sentinel-2 image acquired less than 24 h after the drone image. For all other events, no Sentinel-2 image was acquired less than 24 h after the drone image, or clouds hampered the image acquisition. For the three events covered by a Sentinel-2 image, ground truth contours were manually digitized based on the drone and Sentinel-2 imagery. Additionally, airborne orthophoto imagery, acquired throughout March–April 2020, was considered as an indication of current land cover. The ground truth contours discriminate between dry land, flooding (both open flooding and flooded vegetation), and permanent water.

In June 2016, exceptional rainfall led to widespread flooding across Flanders. A list of 21 communities, in which water damage was reported, was extracted from the flood report [46]. For these communities, all Sentinel-1 images acquired in the critical period (26th May–9th June) are analyzed. As flooding was not present in all locations on all dates, these short time series also allow to assess the false positive rate in the absence of flooding, an important aspect to consider within a monitoring framework. Throughout the critical period, the 2016 flooding was captured by one cloud-free Sentinel-2 image. More specifically, an image acquired on the 7th of June—the same day as a Sentinel-1 acquisition—covered 14 of the 21 communities. Similarly to the 2018–2020 events, ground truth polygons were manually delineated for these 14 image subsets. The acquisition times of the drone, Sentinel-1, and Sentinel-2 imagery used for quantitative validation are summarized in Table I.

III. METHODS

First, the Sentinel-1-based detectability of the floods captured by drone imagery is assessed by means of visual interpretation. Next, a selection of automated flood mapping approaches as well as a newly developed method are applied on the Sentinel-1 image pairs. The resulting flood maps are assessed quantitatively or qualitatively depending on whether ground truth data are available. In the following sections, the used classification approaches as well as the assessment strategy are described.

A. Benchmark Methods

In the first place, four benchmark methods are applied. The first two benchmark methods encompass automated thresholding using the Kittler and Illingworth (KI) algorithm [47] in a split-based approach, as suggested by Martinis et al. [12], applied on both the pixel and object level. These methods will be further referred to as Thresh-px and Thresh-ob, respectively. The split-based approach first selects tiles with a low average and high standard deviation in backscatter, and only uses the selected tiles to automatically determine a threshold value using the KI algorithm. However, due to the urbanized character of the landscape, the thresholding algorithm had to be slightly modified. Instead of selecting the threshold based on the minimization of a cost function, the value corresponding to the global minimum is only withheld if it falls within the range of plausible threshold values [−30, −5 dB] [37], [41], [48]. If this is not the case, e.g., due to a high abundance of double bounce scatterers, the

Location	Drone	Sentinel-1	Sentinel-2
Barebeek	2020-03-12 09:15	2020-03-12 17:32	2020-03-13 10:37
Handzamevaart	2020-03-06 11:50	2020-03-06 17:33	2020-03-06 10:48
Putte-Bonheiden	2020-02-29 09:30	2020-02-29 17:32	2020-03-01 10:50
2016 communities	-	2016-06-07 17:32	2016-06-07 10:40

Times are in UTC. The 2016 communities include Aartselaar, Alken, Geetbets, Halen, Herk-De-Stad, Laakdal, Lummen, Maaseik, Meerhout, Mol, Oudsbergen, Pelt, Wellen, and Zoutleeuw.

value corresponding to the lowest local minimum in this range is selected as the threshold. Tiles for which a threshold value satisfying these conditions cannot be found, are not considered for the tile selection. Using the automatically determined thresholds, a three-class classification is obtained. Open flooding (OF) and permanent water (PW) are classified based on definitions considering both the VV and VH polarization, while pixels not satisfying either of these definitions are classified as dry land (DL). The following class definitions are used:

$$
\begin{cases}
\text{PW} \Leftrightarrow \text{VV}_{\text{ref}} < T_{VV} \text{ and } \text{VH}_{\text{ref}} < T_{VH} \text{ and} \\
\text{VV}_{\text{flood}} < T_{VV} \text{ and } \text{VH}_{\text{flood}} < T_{VH} \\
\text{OF} \Leftrightarrow \text{VV}_{\text{flood}} < T_{VV} \text{ and } \text{VH}_{\text{flood}} < T_{VH}
\end{cases}
\tag{1}
$$

where VV_{ref}, VV_{flood}, VH_{ref}, and VH_{flood} refer to the SAR sigma naught backscatter intensity values of the reference and flood image in the VV and VH polarization (dB scale). This benchmark method is applied both on the pixel and the object level.

Also a previously developed method is applied. This method, which will be further referred to as OBIAflood, applies unsupervised clustering on an object-based feature space comprising the SAR bands of the flood and reference image, optionally complemented by deduced SAR parameters, optical or land cover features, followed by an automated cluster classification and a region growing (RG)-based refinement. It discriminates open flooding, permanent water, and flooded vegetation (FV) from dry land, and makes use of globally available data only. More details concerning this approach can be found in [25]. Given the urbanized character of the landscape and the possible confusion between flooded vegetation and urban areas, the cluster classification for the FV class was modified. More specifically, objects with a main land cover equal to the "buildings" class are masked for the clustering and cluster classification processes. The results shown in this study are produced by using 10 clusters and a feature space comprising the SAR features (VH_{ref}, VV_{ref}, VH_{flood}, VV_{flood}) only.

The outcome of these methods will be used as a benchmark to assess a newly developed method, which will be further referred to as TerraFlood and is described in Section III-B. Compared to the previously described methods, this method uses enhanced classification techniques and defines additional flood classes. In order to illustrate the effect of the additional classes and the enhanced techniques separately, the outcome of the TerraFlood method is also compared with the outcome of a basic version of this method which will be referred to as Thresh-ob-AD, and encompasses object-based thresholding complemented by

a refinement of the PW class and the inclusion of two vegetation classes (flooded vegetation and possibly flooded forest). The latter approach is thus used as a fourth benchmark method. The refinement of the PW class is performed as described in Section III-B3, while the two vegetation classes are included using the approach described in Section III-B4.

B. TerraFlood: A Flood Mapping Approach for Flanders

In order to take full advantage of the available ancillary data, a newly developed method is presented based on similar concepts as the OBIAflood method. More specifically, thresholding is used to delineate a conservative core flood area, while region growing is used to refine the obtained extent and indicate additional flood classes. Objects and pixels have complementary advantages. For example, objects inherently deal with speckle and consider contextual information, while pixels allow for faster processing. In order to make use of the advantages of both, an object- and a pixel-based approach are applied in parallel and combined in the end. The final flood map discriminates between dry land (DL), permanent water (PW), open flooding (OF), long-term flooding (LTF), possible flooding (PF), flooded vegetation (FV), possibly flooded forests (PFF), and invisible forested areas (IF). The LTF class refers to flooding present in both the reference and the flood image, while the PF class refers to areas that are characterized by a higher uncertainty compared to the OF class. An overview of the approach is displayed in Fig. 2, while the different steps are described in the following sections. For all details, the reader is referred to the source code which is available through GitHub.[1]

1) Segmentation and Feature Extraction: In order to allow an object-based approach, objects are created through a segmentation procedure based on the quickshift algorithm [49]. The latter is similar to the mean shift algorithm but uses a kernel-based approximation instead of gradient-descent for mode seeking. More details about this approach are provided in [25]. The quickshift algorithm is applied on a feature space comprising four bands, i.e., the VV and VH bands of the SAR image pair. The two algorithm parameters ks and dist_{max} are set to a 7×7 window and a value of 3, respectively.

2) Delineation of Core Flooding: As a first step, the core flooding is delineated based on thresholding. For the object-based approach, PW and OF class maps on the object and pixel level are created by applying the class definitions in (1) on the original pixel-level SAR images and on the segmented images,

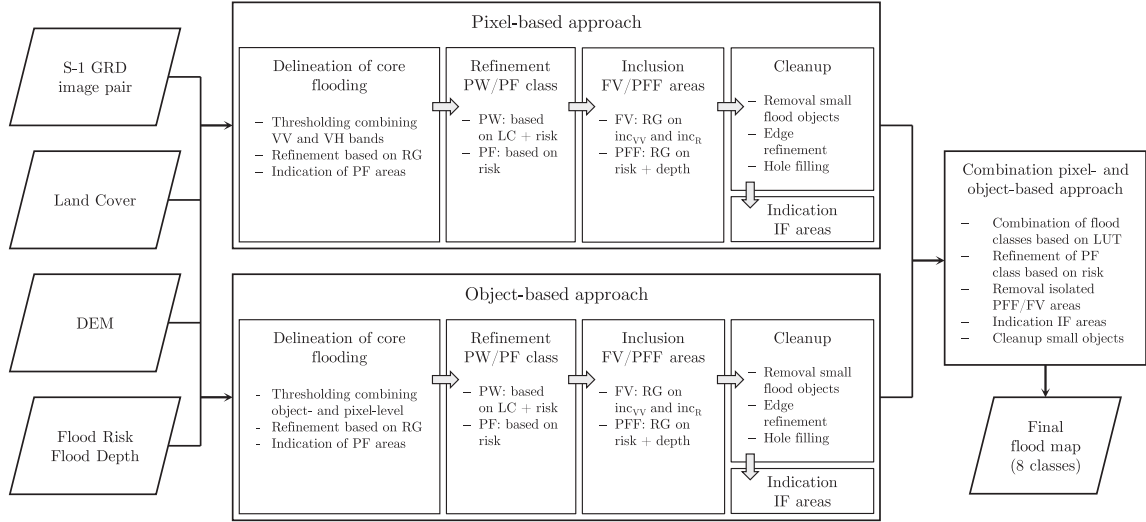[1][Online]. Available: https://github.com/h-cel/TerraFlood

Fig. 2.   Overview of the TerraFlood methodology. The algorithm requires a SAR image pair, land cover, digital elevation, and a flood risk and depth layer as input, and results in an 8-class flood map, discriminating between DL, PW, OF, LTF, PF, FV, PFF, and IF.

respectively. These maps are then combined by applying region growing using the intersection of both maps, on which a minimal mapping unit (MMU) of 10 pixels is applied, as seeds, and the union of both maps as stopping criterion. As such, areas only picked up on the pixel or object level are excluded while the maximum footprint of overlapping patches is maintained. This allows to correct for object edge uncertainties while eliminating noise. Areas that are only picked up on the pixel level and are larger than the MMU, are classified as PF. An algorithmic description of this step is provided in Algorithm 1.

For the pixel-based approach, the PW class is delineated using the class definition in (1). For the OF class, region growing is applied using the class definition in (1) as seed and a so-called "kind" definition, i.e., including pixels below the threshold in one of the polarizations and exceeding the threshold with less than 1 dB in the other polarization, as stopping criterion. Areas that only satisfy the "kind" definition, are not connected to the core flooding and are larger than the MMU of 10 pixels, are classified as PF. An algorithmic description of this step is provided in Algorithm 2.

3) Refinement PW and PF Classes: The PW class can lead to significant confusion due to other smooth surfaces and radar shadow. Smooth surfaces typically include artificial surfaces such as roads and flat roofs, but can include dry agricultural fields too. The latter can be an issue especially during the summer months. In order to reduce confusion, this class is filtered based on land cover and flood risk. More specifically, only PW objects of which more than half of the pixels belong to the PW class according to the land cover layer are retained. If not retained but more than half of the object is prone to flooding according to the flood risk layer, the object is reclassified as LTF. Otherwise, it is reclassified as DL.

Also the PF class is characterized by a significant uncertainty. Therefore, only PF patches of which more than half of the area is prone to flooding according to the flood risk layer are retained.

4) Inclusion FV and PFF Areas: Following the core flood delineation, two additional classes related to flooded vegetation

---

Algorithm 1: Object-Based Approach: Delineation of Core Flooding.

$PW_{px} = VV_{px,ref} < T_{VV}$ and $VH_{px,ref} < T_{VH}$ and $VV_{px,flood} < T_{VV}$ and $VH_{px,flood} < T_{VH}$;

$OF_{px} = VV_{px,flood} < T_{VV}$ and $VH_{px,flood} < T_{VH}$;

$PW_{ob} = VV_{ob,ref} < T_{VV}$ and $VH_{ob,ref} < T_{VH}$ and $VV_{ob,flood} < T_{VV}$ and $VH_{ob,flood} < T_{VH}$;

$OF_{ob} = VV_{ob,flood} < T_{VV}$ and $VH_{ob,flood} < T_{VH}$;

$core_{PW} = PW_{px} \cap PW_{ob}$ and $area_{patch} \geq MMU$;

while *number of changed pixel states > 0* **do**
  $core_{PW} = core_{PW}$ or [neighbouring $core_{PW}$ and $(PW_{px} \cup PW_{ob})$] ;
end

$core_{OF} = OF_{px} \cap OF_{ob}$ and $area_{patch} \geq MMU$;

while *number of changed pixel states > 0* **do**
  $core_{OF} = core_{OF}$ or [neighbouring $core_{OF}$ and $(OF_{px} \cup OF_{ob})$] ;
end

$PW \leftarrow core_{PW}$;

$OF \leftarrow core_{OF}$ and not $core_{PW}$;

$PF \leftarrow OF_{px}$ and $(area_{patch} \geq MMU)$ and not $(core_{PW}$ or $core_{OF})$;

---

are indicated. For both classes, RG is used to fully exploit contextual information and limit overestimation. If sufficient penetration of the SAR signal occurs, flooded vegetation is characterized by enhanced double bounce scattering. As such, an increase in the VV backscatter and an increase in the difference between the VV and VH backscatter (defined as the VV/VH log-ratio) is expected. Similarly to [25] and based on values reported in literature, a threshold value of 3 dB is used for both parameters [50]–[52]. In order to limit confusion with backscatter changes due to other land dynamics such as soil moisture variations, only pixels neighboring OF areas are considered. However, it should be noted that by doing so, only flooded vegetation adjacent to open flooding can be detected.

Algorithm 2: Pixel-Based Approach: Delineation of Core Flooding.

$\text{core}_{\text{PW}} = \text{VV}_{\text{px,ref}} < T_{VV}$ and $\text{VH}_{\text{px,ref}} < T_{VH}$ and $\text{VV}_{\text{px,flood}} < T_{VV}$ and $\text{VH}_{\text{px,flood}} < T_{VH}$ and $\text{area}_{\text{patch}} \geq \text{MMU}$;

$\text{core}_{\text{OF}} = \text{VV}_{\text{px,flood}} < T_{VV}$ and $\text{VH}_{\text{px,flood}} < T_{VH}$ and $\text{area}_{\text{patch}} \geq \text{MMU}$;

$\text{core}_{\text{OF,kind}} = (\text{VV}_{\text{px,flood}} < T_{VV} + 1$ and $\text{VH}_{\text{px,flood}} < T_{VH})$ or $(\text{VV}_{\text{px,flood}} < T_{VV}$ and $\text{VH}_{\text{px,flood}} < T_{VH} + 1)$;

while *number of changed pixel states > 0* **do**
   | $\text{core}_{\text{OF}} = \text{core}_{\text{OF}}$ or [neighbouring core_{OF} and $\text{core}_{\text{OF,kind}}$] ;

end

$\text{PW} \leftarrow \text{core}_{\text{PW}}$;

$\text{OF} \leftarrow \text{core}_{\text{OF}}$ and not core_{PW};

$\text{PF} \leftarrow \text{core}_{\text{OF,kind}}$ and $(\text{area}_{\text{patch}} \geq \text{MMU})$ and not $(\text{core}_{\text{PW}}$ or $\text{core}_{\text{OF}})$;

Algorithm 3: Hole Filling.

foreach *patch* $\in$ *class* **do**
  | **if** *[($\text{VV}_{\text{patch,ref}} < T_{VV} + 1$ and $\text{VH}_{\text{patch,ref}} < T_{VH}$) or ($\text{VV}_{\text{patch,ref}} < T_{VV}$ and $\text{VH}_{\text{patch,ref}} < T_{VH} + 1$)] and [($\text{VV}_{\text{patch,flood}} < T_{VV} + 1$ and $\text{VH}_{\text{patch,flood}} < T_{VH}$) or ($\text{VV}_{\text{patch,flood}} < T_{VV}$ and $\text{VH}_{\text{patch,flood}} < T_{VH} + 1$)]* **then**
    | **if** *PW $\in$ neighbouring classes* **then**
      | patch $\leftarrow$ PW
    | **else**
      | patch $\leftarrow$ PF
    | **end**
  | **else if** *($\text{VV}_{\text{patch,flood}} < T_{VV} + 1$ and $\text{VH}_{\text{patch,flood}} < T_{VH}$) or ($\text{VV}_{\text{patch,flood}} < T_{VV}$ and $\text{VH}_{\text{patch,flood}} < T_{VH} + 1$)* **then**
    | **if** *OF $\in$ neighbouring classes* **then**
      | patch $\leftarrow$ OF
    | **else**
      | patch $\leftarrow$ PF
    | **end**
  | **else if** $\text{area}_{\text{patch}} \leq \text{MMU}$ **then**
    | patch $\leftarrow$ most occuring neighbour class
  | **end**
end

As discussed in previous studies, C-band SAR does mostly not penetrate forest canopies sufficiently to lead to significant backscatter changes [25], [53], [54]. As such, ancillary data and contextual information are used to indicate possibly flooded forest areas in this method. More specifically, a RG approach is used, considering the OF class as seeds. For each seed object, the maximal overall flood risk, considering both pluvial and fluvial flooding, is calculated. Next, neighboring forest pixels are classified as PFF if their absolute water level (calculated as the sum of the elevation and the water depth) exceeds that of the neighboring seed pixel and if their flood risk exceeds or equals that of the initial seed object. In order to account for uncertainty in the elevation and flood depth layers, a buffer of 10 cm is used on the flood depth.

5) Cleanup: In order to improve the interpretability of the flood map, flood class edges are refined by reclassifying small class objects. For this refinement, a MMU of 5 pixels is used. Small objects of the flood classes (PW, OF, LTF, PF, FV) are reclassified based on the most occurring class across the neighboring pixels. Objects of the non-flood classes (DL, IF), which can be considered as holes, are reclassified depending on their size and backscatter properties. If their mean backscatter satisfies the "kind" PW definition, they are reclassified as PW if at least one of the neighboring pixels belongs to this class, and otherwise to PF. If their mean backscatter satisfies the "kind" OF definition, they are reclassified as OF if at least one of the neighboring pixels belongs to this class, and otherwise to PF. Holes that do not satisfy the aforementioned conditions and are smaller than or equal in size to the MMU of 5 pixels, are reclassified as the most occurring class across the neighboring pixels. An algorithmic description of this step is provided in Algorithm 3. Finally, all forested pixels not classified as PFF are labeled as IF.

6) Combination of Pixel- and Object-Based Approach: Finally, the intermediate results of the pixel- and object-based approach, further referred to as TerraFloodPx and TerraFloodOb,

TABLE II
LOOKUP TABLE SPECIFYING IN WHICH CLASS A COMBINATION OF CLASSES RESULTS

Flood map 1	Flood map 2	Combination
PW	PW	PW
OF	OF	OF
LTF	LTF	LTF
PF	PF	PF
FV	FV	FV
OF	FV	PF
LTF	FV	PF
PF	PW, OF, LTF, FV, DL, PFF, IF	PF
OF	PW, LTF, DL, PFF, IF	OF
PW	FV, DL, PFF, IF	PW
LTF	PW, DL, PFF, IF	LTF
FV	DL, PFF, IF	FV

The flood map 1 and 2 column headers refer to the pixel-based and object-based flood map, whose order can be switched.

are combined based on the lookup table provided in Table II. Both the pixel-based and object-based flood map can refer to the first or second column, depending on the combination of values. For example, a combination of PF in the object-based map and OF in the pixel-based map, or vice versa, will result in PF in the combined map. As can be seen in this table, the increased uncertainty of mismatching areas is indicated by means of the PF class. After combining the two maps, the PF class is filtered based on flood risk (cfr. Section III-B3) and FV and PFF objects that have become isolated (i.e., not neighboring an OF object) are reclassified as DL or IF depending on the presence of trees. Finally, class edges are refined by reclassifying small class objects (cfr. Section III-B5).

C. Flood Detectability Assessment

The drone dataset (Section II-C) is used to assess the detectability of the considered flood events using the Sentinel-1 constellation. This assessment is thus done in a qualitative way, by means of a visual comparison between the drone imagery and the Sentinel-1 imagery. More specifically, a label-based score ("bad", "acceptable," or "good") is given to each image. This score is given considering the visibility of open flood patches and contrast to their surroundings, the visibility of flooded vegetation (if present) and contrast to surrounding non-flooded vegetation, and the presence of non-flood features that could cause confusion (e.g., smooth surfaces or backscatter intensity decreases caused by agricultural activities).

D. TerraFlood Accuracy Assessment

Considering both validation datasets, ground truth data providing consistent spatial coverage are available for 17 images, i.e., images of three floods that occurred in 2018–2020 and an image of 14 communities hit by the 2016 flooding. For these images, a quantitative validation by means of binary measures is carried out.

In order to assess the general accuracy, a class-averaged F1 score is calculated. More precisely, the F1 score is first calculated for the DL, PW, and OF classes separately, and the arithmetic average of these class-specific values is then calculated. As such, a bias due to class imbalance is avoided. The F1 score is calculated according to the following equation:

$$F1 = \frac{2TP}{2TP + FP + FN} \qquad (2)$$

where TP, FP, and FN stand for true positives, false positives, and false negatives, respectively. As only DL, PW, and OF are considered here, LTF and FV are reclassified as OF. PF represents an inherently uncertain class and is thus excluded rather than reclassified as OF or DL. For the considered flood cases, this class represents 0.04 to 2.14 % of the total area. In order to indicate the potential over- or underestimation of the accuracy due to this exclusion, error bars indicate the minimum and maximum accuracy obtained across two scenarios, i.e., reclassifying PF as DL and reclassifying PF as OF. As information on the true flood state of forests is not available for all cases, this class is reclassified to DL in both the modeled and truth dataset, with the exception of PFF in the predicted map overlapping with OF in the truth map.

As will be described in more detail in Section IV, the sensitivity of the F1 score is strongly influenced by the (relative) size of the flooding. In order to overcome this issue and to provide more insights into omission and commission errors, the false positive rate (FPR) and false negative rate (FNR) are calculated:

$$FPR = \frac{FP}{FP + TN} \qquad (3)$$

$$FNR = \frac{FN}{FN + TP}. \qquad (4)$$

In order to provide a means of comparison, the aforementioned measures are also calculated for maps produced by pixel-based and object-based thresholding, the OBIAflood method and object-based thresholding complemented by a refinement of the PW class and the inclusion of FV and PFF areas (cfr. Section III-A).

Using the quantitative assessment strategy described above, 17 flood maps can be assessed. In order to include all the 138 flood maps (35 of the flood events in 2018–2020 and 4 or 5 maps per community that reported flood damage in 2016), these maps are assessed in a qualitative way, i.e., by assigning a label-based score ("bad", "acceptable," or "good") based on visual inspection. For the 2018–2020 cases, a comparison is made between the flood maps and the drone imagery as well as between the flood maps and the SAR imagery. While the former comparison gives an indication of how well the flood map represents the ground conditions, the latter comparison allows to assess how good the algorithm extracts the information present in the SAR imagery, irrespective of the flood detectability. For most of the 2016 maps, no ground truth data are available. As such, a label is assigned based on a visual comparison between the maps and the SAR imagery. The main criteria considered for this visual inspection are, in descending order of importance, the presence of false positives and negatives related to the flood classes, flood class confusion, the indication of the PFF class and edge mismatches.

IV. RESULTS AND DISCUSSION

A. Detectability of Flooding on SAR

In order to assess the suitability of medium resolution C-band SAR for flood monitoring in Flanders, the detectability of flooding is assessed as described in Section III-C. The label-based scores are provided in Table III. For 19 events, the flood detectability was assessed as "bad." Three main causes were identified for this lack of detectability. First, the time difference between the drone acquisition and closest Sentinel-1 acquisition (Table III) was often too large to capture the—often very time dynamic—flooding. Second, the observed flooding was often very shallow, in some cases difficult to discriminate from soil saturation based on the drone imagery, and/or small, e.g., along narrow streams. Third, local urban flooding was never detectable. In most cases, a time difference of several hours was enough for the flood water to recede. In one case, the drone imagery captured several flooded houses, and the Sentinel-1 image was acquired only 2.25 h later. However, also here no flood-induced backscatter intensity changes could be observed, indicating that not only the temporal but also the spatial resolution of Sentinel-1 as well as geometric effects are a limiting factor in urban areas. As SAR coherence is increasingly suggested for urban flood mapping [26], [27], a coherence time series was visually inspected (results not shown). Although a decrease in coherence was detected for the flooded area on the co-event image pairs compared to the pre- and post-event image pairs, a similar decrease was observed for other urban zones that were not flooded according to the drone imagery. As such, the decrease might not be induced by the flooding, and more
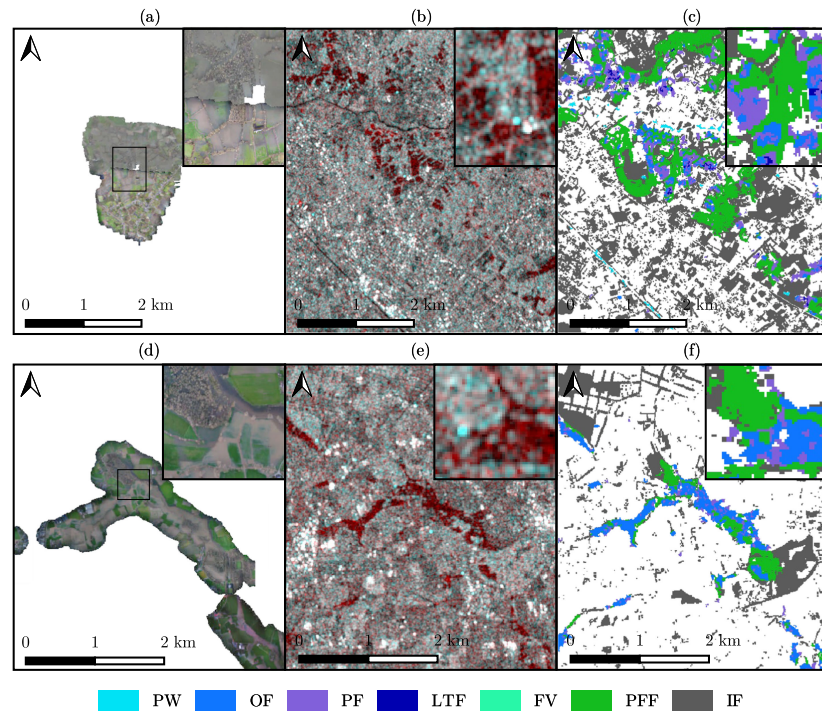
Fig. 3. Drone imagery, Sentinel-1 RGB composite ((VV_{ref}, VV_{flood}, VV_{flood}) with visualization limits (-20, 0) dB), and TerraFlood flood map for floods in Boortmeerbeek on 2020-02-29 (a, b, c) and along the Poekebeek on 2020-03-06 (d, e, f). Imagery and maps are displayed in WGS84 UTM Zone 31 N (EPSG:32631). In the flood map, PW stands for permanent water, OF for open flooding, PF for possible flooding, LTF for long-term flooding, FV for flooded vegetation, PFF for possibly flooded forests and IF for invisible forested areas.

research is needed to investigate the potential of SAR coherence for urban flood mapping in the Flemish context.

For 12 events the visibility was assessed as "acceptable," mostly because the flood extent evolved during the time between the drone and Sentinel-1 acquisition or because part of the flooding was invisible due to tree rows, forest patches, or protruding vegetation. Also disagreements due to shallow flooding were identified. For the remaining 4 events, the detectability was assessed as "good."

In Fig. 3, the drone image and Sentinel-1 composite for floods along the Poekebeek (panel a and b) and in Boortmeerbeek (panel d and e) are shown. For both events, the detectability was assessed as "acceptable." As can be seen in the displays, the open flood zones are characterized by a clear decrease in backscatter (red zones on S-1 composite). However, in both cases, tree rows and forest patches hamper the flood detection. Moreover, no significant difference between the flooded and nonflooded forest patches can be observed. Based on the investigated imagery, it can thus be confirmed that C-band SAR does not allow for sufficient penetration in broad-leaved forests, even under leaf-off conditions, to lead to flood-induced backscatter changes in forested areas.

B. Algorithm Accuracy and Comparison

A quantitative accuracy assessment and method comparison was carried out based on 17 flood maps. An overview of the accuracies obtained by benchmark pixel- and object-based

thresholding (Thresh-px and Thresh-ob), object-based thresholding followed by a refinement of the PW class, and inclusion of FV and PFF based on ancillary data (Thresh-ob-AD), the OBIAflood and the TerraFlood method (the pixel- and object-based intermediates TerraFloodPx and TerraFloodOb, as well as the final map TerraFlood), is provided in Fig. 4. Fig. 4(a)–(c) shows the F1 score, false positive rate, and false negative rate respectively. The F1 scores indicate that pixel- and object-based thresholding mostly perform similar, with low F1 scores for both methods. Several causes of confusion can be identified. First, the medium spatial resolution and speckle effect present in Sentinel-1 imagery hamper a correct delineation of small scale features. The ground truth dataset contains a considerable amount of small (permanent) water features, i.e., between 100 and 2000 m^2, which are often not visible in the SAR imagery and mostly not picked up by the automated algorithms. Second, backscatter changes due to agricultural activities lead to confusion related to the OF class for several cases, especially for images acquired in summer. Third, low backscatter areas such as dry agricultural fields, radar shadow areas, streets, and flat roofs lead to confusion related to the PW class. Fourth, long-term flooding causes confusion between the PW and OF class. When complemented by ancillary data, the accuracy of object-based thresholding increases significantly for several cases. This difference in accuracy highlights the added value of these data, which mainly eliminate confusion related to the PW class and due to tree rows.

The OBIAflood method outperforms the thresholding benchmarks for 4 events, but leads to lower F1 scores for the other
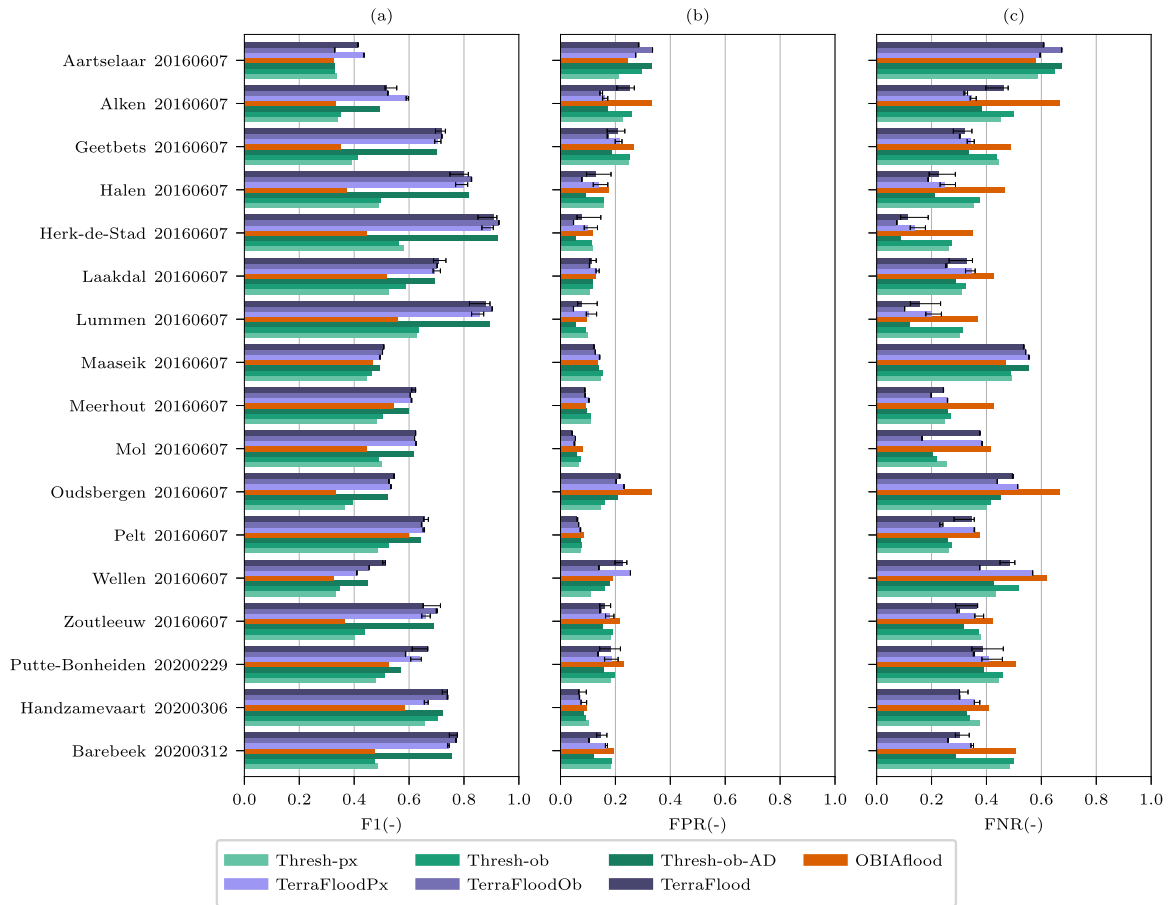
Fig. 4.   (a) Class average F1 score. (b) False positive rate. (c) False negative rate. For benchmark pixel- and object-based thresholding (Thresh-px and Thresh-ob), object-based thresholding complemented by ancillary data (Thresh-ob-AD) the OBIAflood method and the TerraFlood method (pixel- and object-based intermediates TerraFloodPx and TerraFloodOb as well as the final TerraFlood map). Error bars indicate the potential over- or underestimation of the accuracy due to the exclusion of the PF class.

13 events. Due to the abundance of small-scale features and the heterogeneity of the landscape, considering only the object level leads to severe edge mismatching. Moreover, the clustering approach appears to be sensitive to the size of the flooding, leading to large over- or underestimations (cfr. FPR and FNR subplots) especially in case of little to no flooding. Increasing the number of clusters improves the outcome for some cases, but not consistently. For the larger flood events the delineation is good, with the exception of PW-OF confusion. As such, it can be concluded that this approach performs well for larger scale flood mapping, but is less suited for heterogeneous landscapes on the one hand and for continuous flood monitoring on the other hand. With respect to the second point, further research could focus on improving the robustness of the clustering approach, e.g., by optimizing the number of clusters depending on the scene properties [55], [56]. Moreover, it should be noted that the class-averaged F1 score has shortcomings especially for the assessment of maps comprising little to no flooding, in this situation disadvantaging the OBIAflood method compared to the others, as will be further discussed in Section IV-C.

Fig. 4 also shows the TerraFlood method outperforms the others for all cases, with improvements in F1 of up to 0.30.

The main reason for this improvement is a reduction of the main sources of confusion, aided by ancillary data, as also illustrated by the improvement of Thresh-ob-AD compared to Thresh-ob. More specifically, the refinement of the PW and PF classes based on land cover and/or flood risk substantially improves the resulting flood maps, e.g., for Alken, Laakdal, Oudsbergen, and Zoutleeuw. The inclusion of the LTF class strongly reduces OF-PW confusion, leading to improved accuracies for the flood maps of Lummen, Herk-de-Stad, Halen, and the Barebeek, where long-term flooding was present. Moreover, the combination of conservative thresholding and region growing significantly reduces classification errors scattered across the ROI, as illustrated by the improvement of TerraFlood compared to Thresh-ob-AD. However, some classification errors persist or are introduced. For example, low backscatter areas in flood prone zones are classified as LTF, and thus, not correctly eliminated. Also classification errors caused by backscatter decreases due to soil moisture changes or agricultural activities persist. This source of confusion could be eliminated by filtering the OF class based on flood risk. However, this would also create a bias and would decrease the independence of the flood risk information and SAR-based flood maps. As VMM, the end-user of the
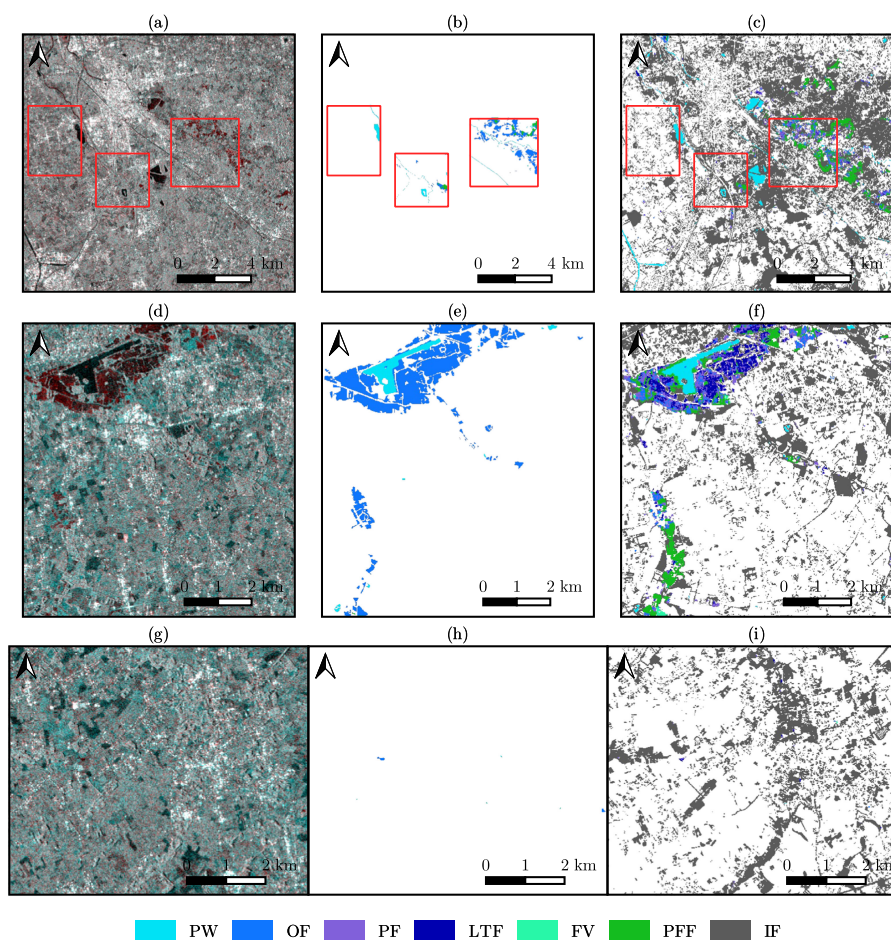
Fig. 5. (a)–(c) Sentinel-1 RGB composite ((VV_{ref}, VV_{flood}, VV_{flood}) with visualization limits (-20, 0) dB), ground truth map and TerraFlood flood map for floods along the Barebeek on 2020-03-12, (d)–(f) in the community of Herk-de-Stad on 2016-06-07, (g)–(i) in the community of Wellen on 2016-06-07. For the Barebeek flood, red rectangles indicate the subsets for which the quantitative assessment was carried out. Imagery and maps are displayed in WGS84 UTM Zone 31 N (EPSG:32631). In the flood map, PW stands for permanent water, OF for open flooding, PF for possible flooding, LTF for long-term flooding, FV for flooded vegetation, PFF for possibly flooded forests, and IF for invisible forested areas.

flood maps, intends to use these among others to investigate whether and where their flood models should be adjusted, this is undesirable. A dependence to the ancillary data is already created, as long-term flooding in areas not prone to flooding is erroneously eliminated from the flood maps. However, this occurred for a few cases only.

Also the accuracies of the pixel- and object-based intermediates are displayed in Fig. 4, as they provide more insight into the TerraFlood algorithm, and the differences between the pixel- and object-based subapproaches. While in general the differences between the two approaches are limited, the best performing approach depends on the event, which validates the choice to combine these intermediates. While the final map provides the highest accuracy for most cases, one of both intermediates is most accurate for several others. However, when considering all events, the final map is the most accurate overall. Moreover, it provides a better indication of uncertainty, as indicated by the error bars.

In order to provide more insight into the spatial distribution of the classification errors as well as how the PF, LTF, FV, and PFF classes perform, Sentinel-1 composites, ground truth

data, and TerraFlood maps are shown in Fig. 5 for the flooding along the Barebeek, in Herk-de-Stad, and in Wellen. Most of the flooding along the Barebeek (Fig. 5(a)–(c)) is correctly classified as LTF. The PF class covers both flooded and non-flooded zones, and thus, seems to have correctly received this uncertainty label. The PFF class is correctly indicated in some locations, while underestimated or overestimated in others. For this case, underestimations are mainly caused by the absence of OF zones neighboring the forest zones. However, PF and LTF were not included as seeds for PFF region growing, as this caused significant overestimations for other cases. PFF overestimations are generally caused by the combined uncertainty of the flood depth classes, especially for the larger depth classes, and resampling of the DTM. Similar conclusions can be drawn for the flooding in Herk-de-Stad (Fig. 5(d)–(f)). Open and long-term flooding are correctly discriminated from the lake, although a significant amount of the LTF areas belongs to the OF class. This confusion is caused by low backscatter caused by low soil moisture conditions in the reference image. For this case, most of the PF areas represent true flooding. The PFF class is indicated in expected places, although some overestimation occurs too.

An area of FV in the south is correctly classified, while some smaller patches in the north are erroneously assigned to this class.

All TerraFlood maps were also assessed in a qualitative way, as summarized in Tables III and IV. The maps of the 2018–2020 events were assessed by comparing them with the drone imagery on the one hand and the SAR imagery on the other hand. In comparison to the drone imagery 21 maps were assessed as "bad," 7 as "acceptable" and 7 as "good," while in comparison to the SAR imagery 11 maps were assessed as "bad," 10 as "acceptable" and 14 as "good." The difference between both assessments underlines the detectability of the flooding is a major hampering factor for the considered cases. The maps of the 2016 flooding were only compared with the SAR imagery. Among these maps, 28 maps were assessed as "bad," 31 as "acceptable" and 44 as "good." The maps assessed as "bad" suffer mainly from overestimations caused by low backscatter areas or backscatter decreases not induced by flooding. Similar classification errors, but to a lesser extent, occurred in the maps assessed as "acceptable." Also overestimations and underestimations of small-scale flood features or at flood edges, as well as over- and underestimations of flooded vegetation and forests were noted. In the flood maps considered as "good" classification errors were limited. Two of the qualitatively assessed maps are visualized in Fig. 3(c) and (f). The FV class occurs in neither of these maps. The flood map of the Poekebeek was assessed as "good," as visible flooding was correctly delineated and PFF indications corresponded very well with the drone imagery, with the exception of some slight underestimations in the southeast of the map. The flood map of Boortmeerbeek was assessed as "acceptable," as OF underestimations along the flood edge and possible PFF overestimations were noted. As such, it can be concluded that the PFF class provides a good indication of possibly flooded forests, but should be interpreted with care.

C. TerraFlood Monitoring Capabilities

The accuracy of flood mapping methods is typically assessed based on images comprising flooding. However, when the method is intended to be used for automated monitoring, its accuracy, when little to no flooding is present, is an important aspect to consider too. In this study, the monitoring capabilities of the algorithm were assessed based on the short time series of the 2016 flood event. A qualitative assessment of the resulting flood maps (Table IV) shows that low backscatter due to dry soil conditions or radar shadow, and backscatter decreases due to soil moisture variations or agricultural activities are the main sources of confusion when no flooding is present.

As the images used for the quantitative accuracy assessment of the 2016 event were acquired on the 7th of June, toward the end of the flood event, little to no flooding was left in several communities (Aartselaar, Meerhout, Mol, Maaseik, Wellen). The resulting F1 scores, shown in Fig. 4, illustrate the shortcomings of this measure when little to no flooding is present. It is known to be sensitive to the size of the flooding, as all binary measures are, and additionally have a bias toward overprediction [57]. This bias becomes more pronounced if little to no flooding is

present. In that case, both the numerator and denominator [cfr. (2)] are low and 1-pixel errors will reduce the resulting score comparatively more than for larger true class sizes. For example, the obtained F1 scores for Wellen are low for all methods, while differences between the true and predicted flood map [Fig. 5(h) and (i)] are very limited. Moreover, when little or no flooding is present, a slight underprediction of the flood extent can lead to a class F1 score of 0, if no true positives are detected. As a result, also the class average F1 score is strongly reduced. This is the case for the PW and OF class in the OBIAflood maps of Alken and Oudsbergen, among others.

D. Limitations and Future Perspectives

Both the quantitative and qualitative accuracy assessment illustrate the potential of the presented approach for automated flood monitoring in Flanders, and the improvements obtained compared to benchmark methods. Nevertheless, some remaining limitations can be identified. First of all, the applicability of the approach remains limited by the detectability of flooding by Sentinel-1, driven by its spatial and temporal resolution. As such, small-scale or time dynamic flooding may not be observed. Second, several classification errors persist. These include OF-DL confusion due to backscatter decreases not induced by flooding, and LTF-DL confusion due to low backscatter in flood prone zones. Third, a bias to the ancillary land cover and flood risk information is introduced, leading to additional classification errors compared to the benchmarks when long-term flooding occurs in areas not indicated as prone to flooding or when permanent water areas are not included in the land cover map. However, this was only observed for a limited number of cases. Fourth, the method depends on ancillary data, which hampers its transferability to other regions. Excluding these data or reducing the resolution of these data is expected to decrease the accuracy of the outcome. Yet, the goal of this study was to optimally make use of the data available for Flanders in the development of a flood monitoring approach for this region.

Several pathways could be explored to overcome the aforementioned limitations in the future. For example, optimizing the reference image selection procedure could reduce confusion due to backscatter changes not induced by flooding. Moreover, the inclusion of crop type information [58] could support the identification of backscatter changes induced by agricultural activities and reduce confusion caused by these. Furthermore, the inclusion of time series could improve the discrimination of gradual backscatter changes caused by soil moisture or vegetation dynamics from abrupt changes caused by flooding. Finally, when imagery from other SAR sensors becomes available, it can be used to improve the classification of flooded vegetation if acquired using another wavelength and close in time to the Sentinel-1 image, and/or to increase the observation frequency of the monitoring system.

To conclude, the automation and demonstrated monitoring accuracy of the presented approach unlock several future research opportunities, when combined with the growing archive of consistent Sentinel-1 SAR acquisitions. These include past flood dynamics analyses, such as the retrieval of flooded area through

time charts and water/flood occurrence maps, as demonstrated in [37] and [38].

V. CONCLUSION

In this study, the potential of Sentinel-1 for automated flood monitoring in Flanders was assessed. First, the capability of Sentinel-1 to detect floods in Flanders was investigated based on several case studies. A qualitative assessment based on visual comparison indicated that floods are generally well detectable by Sentinel-1, although the temporal and spatial resolution of the constellation limit the detectability of small-scale and time-dynamic floods. Moreover, floods in urban and forested zones are difficult to impossible to detect. Second, an automated flood mapping approach was suggested for the Flanders region, as it makes use of locally available ancillary data. It combines thresholding and region growing, and discriminates six flood-related classes. Whereas benchmark methods severely suffer from classification errors due to low backscatter caused by low soil moisture, radar shadow, or impervious surfaces, and due to backscatter decreases caused by agricultural activities or soil moisture/vegetation dynamics, these are reduced significantly in

the suggested approach. Also flooded forests and tree rows are indicated rather accurately by combining elevation, land cover, and flood risk information. However, classification errors due to the aforementioned backscatter dynamics persist, albeit to a lesser extent, while classification errors caused by the uncertainty of the ancillary data are introduced in a limited number of cases. Further improvements could be obtained by extending the ancillary dataset or the SAR input, i.e., by considering a longer time series instead of an image pair or by including images acquired by other SAR sensors. The approach is currently tested by the local water manager VMM, and intended to be used in an operational context.

ACKNOWLEDGMENT

The authors would like to sincerely thank VITO for providing access to analysis-ready Sentinel-1 data and a virtual machine to do the necessary calculations, also for future operational use, as well as providing technical assistance when needed.

APPENDIX A

See Table III–IV.

TABLE III
QUALITATIVE ASSESSMENT OF THE FLOOD DETECTABILITY AND OF THE FLOOD MAPS PRODUCED BY THE TERRAFLOOD ALGORITHM FOR THE 35 EVENTS OBSERVED BY BOTH DRONE AND SAR IMAGERY

Location	Date	Time diff. (h)	Detectability	TF map - drone	TF map - SAR
Aalter	2018-05-25	-2.19	Acceptable	Acceptable	Bad
Barebeek	2020-03-12	8.29	Acceptable	Acceptable	Acceptable
Bellebeek	2020-03-12	29.54	Acceptable	Acceptable	Good
Boekhout	2019-06-10	2.28	Bad	Bad	Good
Boortmeerbeek	2020-02-27	13.59	Acceptable	Acceptable	Good
Boortmeerbeek	2020-02-29	30.54	Acceptable	Acceptable	Acceptable
Brecht	2018-05-31	23.96	Bad	Bad	Good
Bree	2019-05-21	34.83	Bad	Bad	Bad
Bree	2019-05-24	3.65	Bad	Bad	Acceptable
DePinte	2019-04-08	-2.10	Bad	Bad	Acceptable
Diest	2018-06-03	33.04	Bad	Bad	Bad
Diets-Heur	2018-06-03	52.78	Bad	Bad	Bad
Essen	2109-06-22	-5.91	Bad	Bad	Acceptable
Handzamevaart	2020-03-06	5.71	Good	Good	Good
Hertsbergebeek	2020-03-06	2.05	Good	Good	Good
Hoeleden	2020-02-29	28.54	Acceptable	Good	Good
Ieper	2019-06-04	-1.33	Bad	Bad	Acceptable
Kalmthout	2019-06-20	-5.91	Good	Good	Good
Kerkebeek	2020-03-06	9.30	Good	Acceptable	Good
Lauw	2018-06-03	28.78	Bad	Bad	Bad
Mark	2020-03-12	29.54	Acceptable	Acceptable	Acceptable
Overijsse	2109-06-10	3.66	Bad	Bad	Good
Oudenaarde	2019-06-09	-2.50	Bad	Bad	Bad
Peer	2019-05-21	36.83	Bad	Bad	Bad
Poekebeek	2020-03-06	0.30	Acceptable	Good	Good
Poperingevaart	2020-03-06	8.80	Acceptable	Bad	Good
Putte-Bonheiden	2020-02-29	8.04	Acceptable	Bad	Acceptable
Rivierbeek	2020-03-06	2.05	Acceptable	Good	Good
Rotselaar	2020-02-29	32.54	Acceptable	Good	Acceptable
Rutten	2018-06-03	49.28	Bad	Bad	Bad
Sint-Truiden	2019-06-11	3.66	Bad	Bad	Bad
Tielt-Winge	2018-06-03	47.04	Bad	Bad	Acceptable
Turnhout	2019-06-13	15.97	Bad	Bad	Good
Voeren-Moelingen	2018-06-03	48.53	Bad	Bad	Bad
Vreren	2018-06-03	50.78	Bad	Bad	Bad

The TerraFlood maps are compared with both the drone imagery (TF map–drone) and the SAR imagery (TF map–SAR), as described in Section III-D. The acquisition date of the SAR image as well as the time difference (in hours) between the SAR and drone imagery are indicated too. For the latter, positive values indicate the SAR imagery was acquired after the drone imagery.

TABLE IV
QUALITATIVE ASSESSMENT OF THE FLOOD MAPS PRODUCED BY THE
TERRAFLOOD ALGORITHM FOR ALL 103 SAR IMAGES CAPTURING THE
COMMUNITIES HIT BY THE 2016 FLOOD EVENT BETWEEN MAY 26 AND JUNE 9

Location	Date	TF map - SAR
Aartselaar	2016-05-26	Acceptable
Aartselaar	2016-05-30	Good
Aartselaar	2016-06-02	Good
Aartselaar	2016-06-04	Good
Aartselaar	2016-06-07	Bad
Alken	2016-05-26	Bad
Alken	2016-05-30	God
Alken	2016-06-02	Good
Alken	2016-06-04	Acceptable
Alken	2016-06-07	Acceptable
Geetbets	2016-05-26	Bad
Geetbets	2016-05-30	Good
Geetbets	2016-06-02	Good
Geetbets	2016-06-04	Good
Geetbets	2016-06-07	Acceptable
Halen	2016-05-26	Bad
Halen	2016-05-30	Good
Halen	2016-06-02	Good
Halen	2016-06-04	Good
Halen	2016-06-07	Good
Herk-de-Stad	2016-05-26	Bad
Herk-de-Stad	2016-05-30	Good
Herk-de-Stad	2016-06-02	Good
Herk-de-Stad	2016-06-04	Good
Herk-de-Stad	2016-06-07	Good
Ingelmunster	2016-05-26	Acceptable
Ingelmunster	2016-05-28	Acceptable
Ingelmunster	2016-05-31	Good
Ingelmunster	2016-06-04	Good
Ingelmunster	2016-06-07	Acceptable
Laakdal	2016-05-26	Bad
Laakdal	2016-05-30	Good
Laakdal	2016-06-02	Good
Laakdal	2016-06-04	Acceptable
Laakdal	2016-06-07	Acceptable
Ledegem	2016-05-26	Acceptable
Ledegem	2016-05-28	Acceptable
Ledegem	2016-05-31	Good
Ledegem	2016-06-04	Good
Ledegem	2016-06-07	Bad
Lummen	2016-05-26	Bad
Lummen	2016-05-30	Good
Lummen	2016-06-02	Good
Lummen	2016-06-04	Good
Lummen	2016-06-07	Acceptable
Maaseik	2016-05-26	Bad
Maaseik	2016-05-30	Good
Maaseik	2016-06-02	Acceptable
Maaseik	2016-06-07	Acceptable
Meerhout	2016-05-26	Bad
Meerhout	2016-05-30	Good
Meerhout	2016-06-02	Good
Meerhout	2016-06-04	Good
Meerhout	2016-06-07	Acceptable
Mol	2016-05-26	Bad
Mol	2016-05-30	Good
Mol	2016-06-02	Acceptable
Mol	2016-06-04	Bad
Mol	2016-06-07	Acceptable
Oostkamp	2016-05-26	Bad
Oostkamp	2016-05-28	Bad
Oostkamp	2016-05-31	Good
Oostkamp	2016-06-04	Good
Oostkamp	2016-06-07	Bad
Oudsbergen	2016-05-26	Bad
Oudsbergen	2016-05-30	Good
Oudsbergen	2016-06-02	Good
Oudsbergen	2016-06-07	Acceptable

TABLE IV
(Continued.)

Location	Date	TF map - SAR
Pelt	2016-05-26	Bad
Pelt	2016-05-30	Acceptable
Pelt	2016-06-02	Good
Pelt	2016-06-04	Acceptable
Pelt	2016-06-07	Good
Poperinge	2016-05-26	Bad
Poperinge	2016-05-28	Bad
Poperinge	2016-05-31	Good
Poperinge	2016-06-04	Acceptable
Poperinge	2016-06-07	Bad
Rotselaar	2016-05-26	Bad
Rotselaar	2016-05-30	Good
Rotselaar	2016-06-02	Acceptable
Rotselaar	2016-06-04	Acceptable
Rotselaar	2016-06-07	Bad
Vleteren	2016-05-26	Bad
Vleteren	2016-05-28	Bad
Vleteren	2016-05-31	Good
Vleteren	2016-06-04	Acceptable
Vleteren	2016-06-07	Bad
Wellen	2016-05-26	Bad
Wellen	2016-05-30	Good
Wellen	2016-06-02	Good
Wellen	2016-06-04	Good
Wellen	2016-06-07	Acceptable
Wevelgem	2016-05-26	Acceptable
Wevelgem	2016-05-28	Good
Wevelgem	2016-05-31	Good
Wevelgem	2016-06-04	Acceptable
Wevelgem	2016-06-07	Acceptable
Zoutleeuw	2016-05-26	Bad
Zoutleeuw	2016-05-28	Acceptable
Zoutleeuw	2016-05-31	Acceptable
Zoutleeuw	2016-06-04	Acceptable
Zoutleeuw	2016-06-07	Bad

REFERENCES

[1] K. E. Trenberth, A. Dai, R. M. Rasmussen, and D. B. Parsons, "The changing character of precipitation," *Bull. Amer. Meteorological Soc.*, vol. 84, no. 9, pp. 1205–1218, 2003.

[2] W. Ingram, "Extreme precipitation: Increases all round," *Nature Climate Change*, vol. 6, no. 5, pp. 443–444, 2016.

[3] P. C. D. Milly, R. T. Wetherald, K. Dunne, and T. L. Delworth, "Increasing risk of great sitafloods in a changing climate," *Nature*, vol. 415, no. 6871, pp. 514–517, 2002.

[4] "Natural disasters 2019," Centre for Research on the Epidemiology of Disasters, Brussels, Belgium, 2020.

[5] P. C. Oddo and J. D. Bolten, "The value of near real-time earth observations for improved flood disaster response," *Frontiers Environmental Sci.*, vol. 7, 2019, Art. no. 127.

[6] G. Di Baldassarre, G. Schumann, and P. D. Bates, "A technique for the calibration of hydraulic models using uncertain satellite observations of flood extent," *J. Hydrol.*, vol. 367, no. 3, pp. 276–282, 2009.

[7] J. Ran and Z. Nedovic-Budic, "Integrating spatial planning and flood risk management: A new conceptual framework for the spatially integrated policy infrastructure," *Comput., Environ. Urban Syst.*, vol. 57, pp. 68–l79, 2016.

[8] S. Voigt *et al.*, "Global trends in satellite-based emergency mapping," *Science*, vol. 353, no. 6296, pp. 247–252, 2016.

[9] G. J.-P. Schumann, G. R. Brakenridge, A. J. Kettner, R. Kashif, and E. Niebuhr, "Assisting flood disaster response with earth observation data and products: A critical assessment," *Remote Sens.*, vol. 10, no. 8, 2018, Art. no. 1230.

[10] ESA, Paris, France, "Sentinel-1 mission guide," 2020, Accessed: Dec. 1, 2020. [Online]. Available: https://sentinel.esa.int/web/sentinel/missions/sentinel-1

[11] X. Shen, D. Wang, K. Mao, E. Anagnostou, and Y. Hong, "Inundation extent mapping by synthetic aperture radar: A review," *Remote Sens.*, vol. 11, no. 7, 2019, Art. no. 879.

[12] S. Martinis, A. Twele, and S. Voigt, "Towards operational near real-time flood detection using a split-based automatic thresholding procedure on high resolution TerraSAR-X data," *Natural Hazards Earth Syst. Sci.*, vol. 9, no. 2, pp. 303–314, 2009.

[13] M. Tanguy, K. Chokmani, M. Bernier, J. Poulin, and S. Raymond, "River flood mapping in urban areas combining Radarsat-2 data and flood return period data," *Remote Sens. Environ.*, vol. 198, pp. 442–459, 2017.

[14] A. Dasgupta, S. Grimaldi, R. Ramsankaran, V. R. Pauwels, and J. P. Walker, "Towards operational SAR-based flood mapping using neuro-fuzzy texture-based approaches," *Remote Sens. Environ.*, vol. 215, pp. 313–329, 2018.

[15] P. Matgen, R. Hostache, G. Schumann, L. Pfister, L. Hoffmann, and H. H. G. Savenije, "Towards an automated SAR-based flood monitoring system: Lessons learned from two case studies," *Phys. Chem. Earth*, vol. 36, no. 7/8, pp. 241–252, 2011.

[16] L. Pulvirenti, N. Pierdicca, M. Chini, and L. Guerriero, "An algorithm for operational flood mapping from synthetic aperture radar (SAR) data using fuzzy logic," *Natural Hazards Earth Syst. Sci.*, vol. 11, no. 2, pp. 529–540, 2011.

[17] Y. Li, S. Martinis, S. Plank, and R. Ludwig, "An automatic change detection approach for rapid flood mapping in Sentinel-1 SAR data," *Int. J. Appl. Earth Observ. Geoinformation*, vol. 73, pp. 123–135, 2018.

[18] R. S. Westerhoff, M. P. H. Kleuskens, H. C. Winsemius, H. J. Huizinga, G. R. Brakenridge, and C. Bishop, "Automated global water mapping based on wide-swath orbital synthetic-aperture radar," *Hydrol. Earth Syst. Sci.*, vol. 17, no. 2, pp. 651–663, 2013.

[19] S. Schlaffer, P. Matgen, M. Hollaus, and W. Wagner, "Flood detection from multi-temporal SAR data using harmonic analysis and change detection," *Int. J. Appl. Earth Observ. Geoinformation.*, vol. 38, pp. 15–24, 2015.

[20] X. Shen, E. N. Anagnostou, G. H. Allen, G. Robert Brakenridge, and A. J. Kettner, "Near-real-time non-obstructed flood inundation mapping using synthetic aperture radar," *Remote Sens. Environ.*, vol. 221, pp. 302–315, 2019.

[21] M. A. Clement, C. G. Kilsby, and P. Moore, "Multi-temporal synthetic aperture radar flood mapping using change detection," *J. Flood Risk Manage.*, vol. 11, no. 2, pp. 152–168, 2018.

[22] I. Olthof and T. Rainville, "Evaluating simulated RADARSAT Constellation Mission (RCM) compact polarimetry for open-water and flooded-vegetation wetland mapping," *Remote Sens.*, vol. 12, no. 9, May 2020, Art. no. 1476.

[23] A. Hardy*et al.*, "Automatic detection of open and vegetated water bodies using Sentinel 1 to map African malaria vector mosquito breeding habitats," *Remote Sens.*, vol. 11, no. 5, Mar. 2019, Art. no. 593.

[24] B. DeVries, C. Huang, J. Armston, W. Huang, J. W. Jones, and M. W. Lang, "Rapid and robust monitoring of flood events using Sentinel-1 and Landsat data on the Google Earth Engine," *Remote Sens. Environ.*, vol. 240, 2020, Art. no. 111664.

[25] L. Landuyt, N. E. C. Verhoest, and F. M. B. V. Coillie, "Flood mapping in vegetated areas using an unsupervised clustering approach on Sentinel-1 and -2 imagery," *Remote Sens.*, vol. 12, no. 21, 2020, Art. no. 3611.

[26] Y. Li, S. Martinis, M. Wieland, S. Schlaffer, and R. Natsuaki, "Urban flood mapping using SAR intensity and interferometric coherence via Bayesian network fusion," *Remote Sens.*, vol. 11, no. 19, Sep. 2019, Art. no. 2231.

[27] M. Chini, R. Pelich, L. Pulvirenti, N. Pierdicca, R. Hostache, and P. Matgen, "Sentinel-1 InSAR coherence to detect floodwater in urban areas: Houston and Hurricane Harvey as a test case," *Remote Sens.*, vol. 11, no. 2, Jan. 2019, Art. no. 107.

[28] B. Brisco, A. Schmitt, K. Murnaghan, S. Kaya, and A. Roth, "SAR polarimetric change detection for flooded vegetation," *Int. J. Digit. Earth*, vol. 6, no. 2, pp. 103–114, 2013.

[29] A. Twele, W. X. Cao, S. Plank, and S. Martinis, "Sentinel-1-based flood mapping: A fully automated processing chain," *Int. J. Remote Sens.*, vol. 37, no. 13, pp. 2990–3004, 2016.

[30] L. Giustarini *et al.*, "Probabilistic flood mapping using synthetic aperture radar data," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 6958–6969, Dec. 2016.

[31] M. Horritt and P. Bates, "Effects of spatial resolution on a raster based model of flood flow," *J. Hydrol.*, vol. 253, no. 1–4, pp. 239–249, 2001.

[32] R. Heremans *et al.*, "Automatic detection of flooded areas on EN-VISAT/ASAR images using an object-oriented classification technique and an active contour algorithm," in *Proc. Int. Conf. Recent Adv. Space Technol.*, 2003, pp. 311–316.

[33] N. Pierdicca, L. Pulvirenti, G. Boni, G. Squicciarino, and M. Chini, "Mapping flooded vegetation using COSMO-SkyMed: Comparison with polarimetric and optical data over rice fields," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 6, pp. 2650–2662, Jun. 2017.

[34] S. Martinis, A. Twele, and S. Voigt, "Unsupervised extraction of flood-induced backscatter changes in SAR data using Markov image modeling on irregular graphs," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 1, pp. 251–263, Jan. 2011.

[35] W. Cao, A. Twele, S. Plank, and S. Martinis, "A three-class change detection methodology for SAR-data based on hypothesis testing and Markov random field modelling," *Int. J. Remote Sens.*, vol. 39, no. 2, pp. 488–504, 2018.

[36] S. Grimaldi, J. Xu, Y. Li, V. R. Pauwels, and J. P. Walker, "Flood mapping under vegetation using single SAR acquisitions," *Remote Sens. Environ.*, vol. 237, 2020, Art. no. 111582.

[37] B. Pham-Duc, C. Prigent, and F. Aires, "Surface water monitoring within Cambodia and the Vietnamese Mekong delta over a year, with Sentinel-1 SAR observations," *Water*, vol. 9, no. 6, May 2017, Art. no. 366.

[38] I. Olthof and S. Tolszczuk-Leclerc, "Comparing landsat and RADARSAT for current and historical dynamic flood mapping," *Remote Sens.*, vol. 10, no. 5, May 2018, Art. no. 780.

[39] USGS, "Shuttle radar topography mission (SRTM) water body dataset," [Online]. Available: https://doi.org/10.5066/F79G5K8R

[40] S. Martinis, J. Kersten, and A. Twele, "A fully automated TerraSAR-X based flood service," *ISPRS J. Photogrammetry Remote Sens.*, vol. 104, pp. 203–212, 2015.

[41] M. Chini, R. Hostache, L. Giustarini, and P. Matgen, "A hierarchical split-based approach for parametric thresholding of SAR images: Flood inundation as a test case," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 12, pp. 6975–6988, Dec. 2017.

[42] R. Hostache, P. Matgen, and W. Wagner, "Change detection approaches for flood extent mapping: How to select the most adequate reference image from online archives?," *Int. J. Appl. Earth Observ. Geoinformat.*, vol. 19, pp. 205–213, 2012.

[43] R. J. Moore, "The PDM rainfall-runoff model," *Hydrol. Earth Syst. Sci.*, vol. 11, no. 1, pp. 483–499, 2007.

[44] S. A. Nielsen and E. Hansen, "Numerical simulation of the rainfall-runoff process on a daily basis," *Hydrol. Res.*, vol. 4, no. 3, pp. 171–190, 1973.

[45] C. J. Digman, T. Bamfor, D. J. Balmforth, N. M. Hunter, and S. G. Waller, *A Comparison of Modelling Methods for Urban Flood Assessment*. London, U.K.: Taylor & Francis Group, 2008, p. 67.

[46] V. Milieumaatschappij, "Rapport wateroverlast 27 mei - 8 juni 2016," Aalst, Belgium, 2016.

[47] J. Kittler and J. Illingworth, "Minimum error thresholding," *Pattern Recognit.*, vol. 19, no. 1, pp. 41–47, 1986.

[48] J. Zhao, R. Pelich, R. Hostache, P. Matgen, W. Wagner, and M. Chini, "A large-scale 2005–2012 flood map record derived from ENVISAT-ASAR data: United Kingdom as a test case," *Remote Sens. Environ.*, vol. 256, 2021, Art. no. 112338.

[49] A. Vedaldi and S. Soatto, "Quick shift and kernel methods for mode seeking," in *Proc. Eur. Conf. Comput. Vision*, 2008, pp. 705–718.

[50] L. L. Hess and J. M. Melack, "Remote sensing of vegetation and flooding on Magela Creek Floodplain (Northern Territory, Australia) with the SIR-C synthetic aperture radar," *Hydrobiologia*, vol. 500, pp. 65–82, 2003.

[51] S. Martinis and C. Rieke, "Backscatter analysis using multi-temporal and multi-frequency SAR data in the context of flood mapping at River Saale, Germany," *Remote Sens.*, vol. 7, no. 6, pp. 7732–7752, Jun. 2015.

[52] V. Tsyganskaya, S. Martinis, P. Marzahn, and R. Ludwig, "Detection of temporary flooded vegetation using Sentinel-1 time series data," *Remote Sens.*, vol. 10, no. 8, 2018, Art. no. 1286.

[53] M. W. Lang, P. A. Townsend, and E. S. Kasischke, "Influence of incidence angle on detecting flooded forests using C-HH synthetic aperture radar data," *Remote Sens. Environ.*, vol. 112, no. 10, pp. 3898–3907, 2008.

[54] N. Pierdicca, L. Pulvirenti, and M. Chini, "Flood mapping in vegetated and urban areas and other challenges: Models and methods," in *Flood Monitoring Through Remote Sensing*. Cham, Switzerland: Springer, 2018, pp. 135–179.

[55] M. Chini, A. Chiancone, and S. Stramondo, "Scale Object Selection (SOS) through a hierarchical segmentation by a multi-spectral per-pixel classification," *Pattern Recognit. Lett.*, vol. 49, pp. 214–223, 2014.

[56] A. Refice, M. Zingaro, A. D'Addabbo, and M. Chini, "Integrating C- and L-band SAR imagery for detailed flood monitoring of remote vegetated areas," *Water*, vol. 12, no. 10, 2020, Art. no. 2745.

[57] E. Stephens, G. Schumann, and P. Bates, "Problems with binary pattern measures for flood model evaluation," *Hydrological Processes*, vol. 28, no. 18, pp. 4928–4937, 2014.

[58] K. V. Tricht, A. Gobin, S. Gilliams, and I. Piccard, "Synergistic use of radar Sentinel-1 and optical Sentinel-2 imagery for crop mapping: A case study for Belgium," *Remote Sens.*, vol. 10, no. 10, 2018, Art. no. 1642.

Lisa Landuyt received the B.Sc. and M.Sc. degrees in bioscience engineering, and the Ph.D. degree in applied biological sciences from Ghent University, Ghent, Belgium, in 2014, 2016, and 2021, respectively.

In 2016, she joined the Hydro-Climate Extremes lab (H-CEL) and Remote Sensing | Spatial Analysis lab (REMOSA) with Ghent University. Her research interests include the retrieval of environmentally relevant information from remote sensing products, with a special interest for hydrology and SAR imagery.

Frieke M. B. Van Coillie received the M.Sc. and Ph.D. degrees in bioscience engineering from the Ghent University (UGent), Ghent, Belgium, in 1996 and 2003, respectively.

She was a Postdoc Researcher with the Department of Environment, UGent, from 2003 to 2015. She was the resp. Assistant Professor and Associate Professor of Remote Sensing and GIS with the Faculty of Bioscience Engineering, Ghent University, in 2015 amd 2020, respectively. Since 2017, she is heading the Remote Sensing/Spatial Analysis lab. Her research interests include remote sensing, forestry, and geoinformatics.

Bram Vogels received the M.Sc. degree in bioscience engineering from the Catholic University of Leuven, Leuven, Belgium, in 2006.

He first started working around environmental assesment for 1.5 a and since then he works with the Flanders Environment Agency (VMM), Aalst, Belgium, mostly around water assesment. Evaluating the impact of projects and plans on the watersystem is a core business and a very important part of this is the impact of floods on infrastructure. Good and accurate data on flooded zones is important and satellite imaginary is one of the sources fort this besides drone images, field observations, measurements, etc.

Joost Dewelde received the M.Sc. degree in civil engineering from the Catholic University of Leuven, Leuven, Belgium, in 2007.

He has been working with the Flanders Environment Agency (VMM), Aalst, Belgium, as a Hydrologist applying hydrologic and hydrodynamic models to generate flood hazard and risk maps, evaluate the impact of measures and forecast floods for the larger unnavigable rivers in Flanders. His research interests include automated flood mapping based on Sentinel-1 imagery in addition to the smaller scale and higher resolution flood mapping based on drone images as already available at VMM.

Niko E. C. Verhoest received the B.E. and Ph.D. degrees in applied biological sciences from Ghent University, Ghent, Belgium, in 1994 and 2000, respectively.

He was a Teaching Assistant and an Assistant Professor with the Laboratory of Hydrology and Water Management (currently Hydro-Climate Extremes Lab), Ghent University, from 1998 to 2000 and 2000 to 2002, respectively. In 2002, he became an Associate and Full Professor of Hydrology and Water Management with the Faculty of Bioscience Engineering, Ghent University. His research interests include hydrological applications of radar remote sensing.