Article

InSAR Multitemporal Data over Persistent Scatterers to Detect Floodwater in Urban Areas: A Case Study in Beletweyne, Somalia

Luca Pulvirenti 1,*, Marco Chini 2 and Nazzareno Pierdicca 3

1 CIMA Research Foundation, I-17100 Savona, Italy
2 Environmental Research and Innovation Department (ERIN), Luxembourg Institute of Science and Technology (LIST), L-4422 Belvaux, Luxembourg; marco.chini@list.lu
3 Department of Information Engineering, Electronics and Telecommunications (DIET), Sapienza University of Rome, I-00184 Rome, Italy; nazzareno.pierdicca@uniroma1.it
* Correspondence: luca.pulvirenti@cimafoundation.org; Tel.: +39-019-2302724

Abstract: A stack of Sentinel-1 InSAR data in an urban area where flood events recurrently occur, namely Beletweyne town in Somalia, has been analyzed. From this analysis, a novel method to deal with the problem of flood mapping in urban areas has been derived. The approach assumes the availability of a map of persistent scatterers (PSs) inside the urban settlement and is based on the analysis of the temporal trend of the InSAR coherence and the spatial average of the exponential of the InSAR phase in each PS. Both interferometric products are expected to have high and stable values in the PSs; therefore, anomalous decreases may indicate that floodwater is present in an urban area. The stack of Sentinel-1 data has been divided into two subsets. The first one has been used as a calibration set to identify the PSs and determine, for each PS, reference values of the coherence and the spatial average of the exponential of the interferometric phase under standard non-flooded conditions. The other subset has been used for validation purposes. Flood maps produced by UNOSAT, analyzing very-high-resolution optical images of the floods that occurred in Beletweyne in April–May 2018, October–November 2019, and April–May 2020, have been used as reference data. In particular, the map of the April–May 2018 flood has been used for training purposes together with the subset of Sentinel-1 calibration data, whilst the other two maps have been used to validate the products generated by applying the proposed method. The main product is a binary map of flooded PSs that complements the floodwater map of rural/suburban areas produced by applying a well-consolidated algorithm based on intensity data. In addition, a flood severity map that labels the different districts of Beletweyne, as not, partially, or totally flooded has been generated to consolidate the validation. The results have confirmed the effectiveness of the proposed method.

Keywords: flood; urban areas; Sentinel-1; persistent scatterers

1. Introduction

The frequency and intensity of urban flooding are likely to increase in the future because heavy rainfall caused by storms is increasing in frequency and severity due to climate change. Floods caused by intense rainfall events are responsible for numerous casualties and several million euros of damage every year [1]. Many cities around the world are affected by recurrent floods. The capability to detect floodwater in urban settlements may help in identifying areas of the cities more prone to flooding. Identification may be useful for urban development plans to prevent further building in flood-prone areas. In addition, insurance companies require detailed global and national disaster impact databases to identify areas vulnerable to flooding and to support parametric insurance, i.e., to trigger payouts in cases of emergency using predefined criteria [2]. Operational flood mapping can also be useful for relief management when a flood event occurs [3] or when intense rainfall is forecasted.
Synthetic Aperture Radar (SAR) represents the most suitable instrument for flood mapping because of its capability to provide high spatial resolution images in almost all weather conditions. Several algorithms from the literature are presently available to produce near real-time (NRT) inundation maps useful for operational flood relief management (e.g., [4–7]). However, although new insights were provided on mapping floods in urban environments ([3,8–11]), this task still represents a challenge and deserves further investigations. Mapping urban floods using SAR data is the subject of this paper, which presents a new method to perform it.

In urban settlements, specular reflection, as well as double and multiple bounces, may occur (e.g., walls and street forming dihedral reflectors) depending on the alignment of roads and adjacent buildings with respect to the direction of the SAR line of sight [12]. Moreover, urban structures may cause shadowing and layover. Hence, urban flooded areas (UFAs) do not have a unique and unambiguous backscattering signature that distinguishes them from other targets. This implies that UFAs are generally not detected by NRT algorithms used for rapid mapping based on SAR intensity data.

The high sensitivity of the interferometric SAR (InSAR) phase to small surface changes can help in dealing with the ambiguity of the radar backscatter signature of UFAs. Past studies ([12–15]) showed that the presence of floodwater during the interval between the observation times of two SAR images acquired in repeat pass configuration produces a decrease in magnitude of the complex cross correlation (i.e., the coherence) between the images. Three factors, namely spatial (related to the spatial baseline), thermal (related to the signal to noise ratio), and temporal (related to physical changes in the observed target) decorrelations, contribute to the total observed decorrelation [16]. Under standard non-flooded conditions, an urban area is a typically coherent target, because, for manmade structures, the spatial arrangement of the individual scatterers does not change in time (low spatial decorrelation) and the double-/multiple-bounce effect typically produces a high backscatter (low thermal decorrelation) [17]. Conversely, water surfaces show low coherency due to temporal decorrelation (signal scattered backward by Bragg waves generally decorrelates within tens of milliseconds) and low backscattering (unless wind roughens the water surface) [18].

Studies making use of the data provided by very high-resolution instruments such as COSMO-SkyMed (CSK) [12,14] and TerraSAR-X (TSX) [19] showed that SAR interferometry has potential to detect UFAs by searching for decreases in InSAR coherence (\(\gamma\)) due to temporal decorrelation. The detection of the decreases is performed by comparing the coherence of a co-event pair (one image acquired before the flood and one image of the flood) to that of a pre-event (or pre-flood) pair (two images acquired before the flood). More recently, the suitability of an approach based on \(\gamma\) data for mapping UFAs has been assessed for Sentinel-1 (S1) data too [2,20].

Even approaches based on analysis of the coherence may have limitations. Firstly, the analysis should preferably be limited to interferometric pairs having short spatial baselines (\(B_s\)) to minimize spatial decorrelation [21]. This requirement on \(B_s\) is fulfilled by S1 thanks to its narrow orbit tube [17]. Moreover, it must be considered that \(\gamma\) is computed over a moving window (generally a square window including several pixels) and that pixels where \(\gamma\) fluctuates because of causes like the presence of vegetation or unpaved roads inside the urban area may be included in the window. In these targets, changes in the spatial arrangement of the scatterers can occur even under non-flooded conditions. These changes may give rise to low values of coherence even for urban pixels of the pre-flood pair; therefore, low decreases in the coherence and thus omission errors can take place.

The regular acquisitions of S1 data in the frame of the Copernicus program guarantee a continuous operational service and allow users to analyze a long time series of data. Through this analysis, pixels that, under standard conditions, are expected to have little likelihood of decorrelation may be identified. To be more confident that \(\gamma\) decreases are produced by nothing but the presence of floodwater, attention can be restricted to this kind of pixels, known as persistent scatterers (PSs) [22]. PSs are targets characterized by
high backscatter combined with high temporal stability. They can be artificial or even natural elements present on the terrain whose electromagnetic features (e.g., complex radar reflectivity) do not vary significantly in time. Examples of PSs are buildings, metal structures, and exposed rocks; therefore, it is likely to find a large density of PSs in urban areas. PSs are basically insensitive to the effects of temporal (and thermal) decorrelation, thus generally preserving the phase information in time.

Even focusing on coherence in PSs can have limitations for flood mapping in built-up areas. Coherence accounts not only for the difference between the phases of the master and the slave images (i.e., the phase of the interferogram from which \( \gamma \) is derived), but also for the magnitude of the reflectivity of the pixels in the two images. For some PSs, this magnitude can be so high that it tends to be dominant in the calculation of \( \gamma \), while the phase of the other pixels included in the window over which \( \gamma \) is computed has limited influence [23]. In this case, phase variations due to the presence of floodwater in the window may not produce a significant decrease in \( \gamma \); this can lead to omission errors. In [23], the use of interferometric phase statistics instead of coherence has been proposed to reduce these errors.

The idea underpinning the present study is based on the consideration that the PS interferometry technique is well-established and commonly used to monitor surface displacements (e.g., [24]). It can therefore be expected that, in the near future, updated maps of PSs will be available for many urban areas in the world, thanks to the availability of Information and Communication Technology (ICT) for big data management and high-performance computing. These maps could be complemented by metadata containing, for each PS, reference data for \( \gamma \) and for the interferometric phase. Moreover, updated inundation maps in rural areas at a global scale will soon be available thanks to one of the automatic flood detection algorithms available in the literature (e.g., [4]). Hence, in the case of detection of a flooded rural/suburban area surrounding an urban one, it will likely be possible to quickly verify the occurrence of significant anomalies, with respect to reference values, of the coherence and/or the interferometric phase in some PSs. If many PSs can be identified in an urban area, they may represent a sample of where floodwater is present in the urban area. The flood map could therefore be complemented by a map of flooded PSs.

This paper investigates the feasibility of the approach described above and presents a new way to tackle the problem of the detection of floodwater in urban settlements using SAR data based on the identification of flooded PSs. The identification is carried out through a joint analysis of interferometric coherence and phase statistics, namely the spatial average of the exponential of the interferometric phase, over the PSs. For this purpose, a long time series (from June 2017 to July 2020) of IW5S1 data over the town of Beletweyne (Somalia) has been processed and analyzed. Beletweyne is an area that is recurrently affected by floods; during the period June 2017–July 2020, three big floods took place in April–May 2018, October–November 2019, and April–May 2020. Beletweyne represents a test site suitable to accomplish this study because, for the events listed before, maps showing flooding severity in the different districts of the town have been made available by the United Nations Institute for Training and Research (UNITAR) under the United Nations Operational Satellite Applications Programme (UNOSAT) (https://unitar.org/maps/all-maps). One of these maps has been used for calibration purposes, while the other maps have been used for validation purposes.

It is worth pointing out that, to the best of our knowledge, the use of the PS technique for a systematic monitoring of floods in an urban settlement was never investigated in the past. Moreover, in view of a possible operational use of InSAR data for the purpose of flood mapping in urban settlements, this study reliably evaluates commission errors by applying the proposed method to a stack of data acquired not only under flooded conditions, but also under non-flooded ones. Conversely, most of the previous studies making use of the coherence dealt only with the capability of InSAR data to detect the presence of floodwater.

Section 2 synthetically describes the available dataset and the processing carried out to identify the PSs. Then, it presents the method designed to detect flooded PSs and explains
its rationale. Section 3 presents the flood maps obtained by applying the proposed method. Section 4 analyzes the results and Section 5 draws the main conclusions.

2. Materials and Methods

2.1. Study Area and Test Cases

Beletweyne is a city in central Somalia placed in the Shabelle Valley near the Ethiopia border. It is the capital of the Hiiraan province, and its population exceeds 400,000 people. Figure 1 shows the geographic location of Beletweyne together with the main areas (districts) of the town. The map of the districts has been derived from one of the maps made available by UNOSAT.

Figure 1. Geographic location of Beletweyne (left panel) and main districts of the town (derived from the maps available at https://unitar.org/maps/map/3054). The lower panel is a zoom on the urban area, where the districts are shown with transparency.
Beletweyne is crossed by the Shabelle River that, in recent years, frequently broke its banks, causing severe floods. Most of them were caused by the Gu rains, which are monsoon-like rainfall events that usually take place in late March–April and last until June. A second flood season is known as the Dayr rains and occurs in October–November.

The flood that hit Beletweyne in April–May 2018 was caused by heavy Gu rains. Reportedly, more than 120,000 people were displaced from Beletweyne town and surrounding riverine villages. Floodwater inundated houses and crops and affected many infrastructures. The flood that occurred in October–November 2019 was caused by the increase in the level of the Shabelle River after the onset, in early October, of the Dayr rains. An estimated number of 72,000 people from Beletweyne were moved to higher ground in surrounding areas. The flood that took place in April–May 2020 was again due to the Gu rains that significantly increased their intensity in the last week of April. This increase caused a sharp rise in the level of the Shabelle river. Reportedly, more than 115,000 people were displaced from Beletweyne.

2.2. Initial Processing of the Data

The initial processing of the data aimed at generating the following data: (A) preliminary flood maps based on a standard algorithm exploiting the S1 intensity data only [4]; (B) PSs in the Beletweyne area; (C) InSAR coherence and spatial average of the exponential of the interferometric phase; (D) reference maps derived from the shapefiles produced by UNOSAT. The production of the aforementioned data is described in the following subsections.

2.2.1. SAR Data

S1 Level-1 Single Look Complex (SLC) products acquired in Interferometric Wide Swath (IWS) mode have been used to carry out this study. In general, S1 observations of Beletweyne are available every 12 days from the Sentinel-1A platform (S1A); 81 images, acquired by S1A between the beginning of June 2017 and the beginning of July 2020, have been gathered (relative orbit 35) and the acquisition days are listed in Table 1. Only co-polarized data (VV polarization) have been considered in this study.

| Table 1. Acquisition days of the Sentinel-1 data. Italic font: images used for calibration; bold font: images used to produce flood maps in rural/suburban areas; underlined font: images excluded from the persistent scatterers (PSs) identification; blue text: master image for the PSs identification. |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| 06/06/2017     | 18/06/2017      | 30/06/2017      | 12/07/2017      | 24/07/2017      | 05/08/2017      | 17/08/2017      | 29/08/2017      |
| 22/09/2017     | 04/10/2017      | 16/10/2017      | 28/10/2017      | 09/11/2017      | 21/11/2017      | 03/12/2017      | 15/12/2017      |
| 08/01/2018     | 20/01/2018      | 01/02/2018      | 13/02/2018      | 25/02/2018      | 09/03/2018      | 21/03/2018      | 02/04/2018      |
| 26/04/2018     | 08/05/2018      | 20/05/2018      | 01/06/2018      | 13/06/2018      | 25/06/2018      | 07/07/2018      | 19/07/2018      |
| 28/11/2018     | 22/12/2018      | 03/01/2019      | 27/01/2019      | 08/02/2019      | 20/02/2019      | 04/03/2019      | 16/03/2019      |
| 09/04/2019     | 27/05/2019      | 08/06/2019      | 20/06/2019      | 02/07/2019      | 14/07/2019      | 26/07/2019      | 07/08/2019      |
| 22/03/2020     | 03/04/2020      | 27/04/2020      | 09/05/2020      | 21/05/2020      | 02/06/2020      | 14/06/2020      | 26/06/2020      |
| 22/03/2020     | 03/04/2020      | 27/04/2020      | 09/05/2020      | 21/05/2020      | 02/06/2020      | 14/06/2020      | 26/06/2020      |

2.2.2. Flood Maps in Rural/Suburban Areas

Flood maps in rural/suburban areas have been generated by applying, to the images acquired on April 26th, 2018, May 8th, 2018, October 30th, 2019, and May 21st, 2020 (bold font in Table 1), the algorithm proposed in [4], which works on SAR intensity data (backscattering). A smooth water surface reflects incident radiation mainly in the specular direction (low radar return), while rougher surfaces have a more diffuse scattering pattern [25]. Considering that in a scene observed by SAR, not only calm water but also other surfaces like asphalt are smooth, the algorithm developed in [4] searches for areas where the backscattering is low and decreases with respect to the backscattering
of a pre-flood image, thus performing a change detection approach. These areas are detected through a parametric thresholding of the flood image and of the difference image (difference between flood and pre-flood backscattering values in dB units). Readers are referred to [4] for more details about the algorithm.

To generate the difference images, even the S1 data acquired on 14 April, 2018, 6 October, 2019, and 9 May, 2020 (i.e., the closest in time to the flood images listed above) have been processed. The SARscape COTS software package has been used to calibrate, multilook, and geocode the SLC images. A multilooking factor of 4 (range) × 1 (azimuth) has been applied, obtaining a pixel size of 15 × 15 m². The freely available Shuttle Radar Topography Mission (SRTM) Digital Elevation Model (DEM) at 1 arcsec resolution (SRTM1), resampled at 15 × 15 m², has been used to geocode (Lat/Lon projection, datum WGS84) the SAR images. As an example, in Figure 2, the areas classified as flooded on 26 April, 2018 are indicated in blue color. As expected, an intensity-based flood detection algorithm searching for areas of low backscatter in SAR images hardly detects floodwater in the urban settlements. Hence, the idea is to complement the intensity-based flood maps with a map of flooded PSs.

Figure 2. Rural/suburban areas classified as flooded on 26 April, 2018 by the algorithm developed in [8] (blue) and persistent scatterers (red dots) identified in the Beletweyne area. The boundaries of the districts shown in Figure 1 are highlighted in black.

2.2.3. Identification of the PSs

A spatial subset has been initially created in order to reduce the image extent to the area of interest (AOI: Beletweyne town, in this case), thus speeding up the processing time. To extract the PSs, the S1 data acquired from June 2017 to December 2018 have been selected (italic font in Table 1), except those acquired in April–May 2018 and on 1 June, 2018 (6 images, underlined font in Table 1) in order to exclude from the processing stack the images acquired when, reportedly, floodwater was present in Beletweyne. The rationale is that phase fluctuations due to the presence of floodwater have an impact on the PSs identification, which has therefore been performed using 41 images (the total number of
images of Beletweyne acquired by S1A in orbit 35 during the period June 2017–December 2018 is 47).

The PS module available from SARscope [26], based on [22], has been used to identify the PSs. Firstly, all the images (slaves) have been co-registered onto a selected master image. Among the stack of 41 images used by the PS tool, the 21st one has been selected as the master. Then, interferograms have been generated for each slave image using the master one. The spatial and temporal baselines of the interferograms are shown in Figure 3. To perform interferogram flattening, the SRTM1 DEM previously introduced has been used. Successively, the amplitude dispersion index $D$, defined as the ratio between the temporal standard deviation of the backscattered radar signal amplitude for the single pixel and the temporal mean [22], has been computed. A first selection of the PSs has been performed based on the $D$ value, considering that a pixel that constantly has a similar, relatively large amplitude during all acquisitions is expected to have a small phase dispersion; therefore, $D$ is a proxy of the phase stability. Pixels with small $D$ (a threshold of 0.25 was proposed in [22]) are therefore expected to have a small phase standard deviation.

![Figure 3. Perpendicular baseline of the interferometric pairs used to identify the PSs versus the acquisition dates of the 41 SI images of the stack. The yellow diamond indicates the master image; the green diamonds indicate the slaves. The numbers represent an index identifying each image of the stack.](image)

Atmospheric artifacts (commonly known as atmospheric phase screen) have been removed using spatially low-pass and temporally high-pass filters, because the atmospheric component has high spatial correlation and low temporal correlation [27]. The final identification of the PSs has been performed based on a temporal analysis of their phase values, by searching for pixels having a stable temporal behavior of the phase.

Even the PSs product has been geocoded (Lat/Lon projection, datum WGS84, pixel size of $15 \times 15$ m$^2$) by using the SRTM1 DEM. In Figure 2, the red dots indicate location of the PS identified by the processing described above. The PSs placed outside the AOIs shown in Figure 1 will not be considered in the following analysis.

2.2.4. Interferometric Products

Besides the interferograms used to identify the PSs, other interferograms have been generated from the same stack of 47 S1 SLC data acquired throughout the period June 2017–December 2018. In particular, 46 interferometric pairs have been derived; by denoting as
$t_1, t_2, ..., t_{47}$ the acquisition times of the S1 images, pairs have been formed by images acquired at $t_i$ and $t_{i+1}$ ($i = 1:46$), i.e., at consecutive times. From the interferograms, the coherence has been firstly derived; it can be expressed as:

$$\gamma = \frac{|\langle s_1(x)s_2^*(x) \rangle|}{\sqrt{\langle |s_1(x)|^2 \rangle \cdot \langle |s_2(x)|^2 \rangle}} = \frac{\frac{1}{N} \sum_{x \in W} s_1(x)s_2^*(x)}{\sqrt{\frac{1}{N} \sum_{x \in W} |s_1(x)|^2 \cdot \frac{1}{N} \sum_{x \in W} |s_2(x)|^2}}$$

where $\langle \rangle$ indicates the expected value, $s_1$ and $s_2$ are two co-registered complex SAR images, $x$ represents a generic pixel, $W$ is the mobile window over which $\gamma$ is computed, and $N$ indicates the size of the window in number of pixels. In Equation (1), expected values are implemented by computing spatial averages.

A complex SAR image has module and phase components:

$$s(x) = \rho(x)e^{-j\beta R(x) - j\varphi(x)}$$

where $\rho(x)$ and $\varphi(x)$ represent the module and the phase of the reflectivity of pixel $x$, $\beta$ is the phase constant, and $R(x)$ is the sensor–pixel distance (i.e., the range). Then:

$$\gamma = \frac{\frac{1}{N} \sum_{x \in W} \rho_1(x)\rho_2(x)e^{-j\Delta \varphi(x)}}{\sqrt{\frac{1}{N} \sum_{x \in W} \rho_1(x)^2 \cdot \frac{1}{N} \sum_{x \in W} \rho_2(x)^2}}$$

where no variation of $R(x)$ between the InSAR acquisitions has been assumed (i.e., no surface displacements) and $\Delta \varphi = \varphi_1 - \varphi_2$. To take into account, beside $\gamma$, a quantity depending only on $\Delta \varphi$ and not on the modules, the magnitude of the average over $W$ of the interferometric phase, introduced in [23], has been computed too:

$$\zeta = \frac{1}{N} \sum_{x \in W} e^{-j\Delta \varphi(x)}$$

Hereafter, $\zeta$ will be denoted as the spatial average of the exponential of the interferometric phase (although, more rigorously, it is the magnitude of the spatial average).

A factor of 4 (range) $\times$ 1 (azimuth) looks has again been applied to generate multi-looked interferograms. A square window of $5 \times 5$ pixels has been used to derive the $\gamma$ and $\zeta$ maps from the flattened interferograms. All the maps have firstly been co-registered onto the geometry of the $(4 \times 1)$ multilooked master image selected to perform the PSs identification and successively geocoded using the SRTM1 DEM resampled at 15 m.

### 2.2.5. Reference Data

Flood maps made available by UNOSAT have been used as reference data for both calibration and validation of the proposed procedure. These maps, in shapefile format, provide information about two different classes of severity of flooding: totally flooded and partially flooded. The areas of partially flooded districts present affected roads and buildings but are not totally submerged by water like districts classified as totally flooded. The shapefiles have been rasterized taking as reference the geocoded maps generated as described in the previous sections. For the event that occurred in April–May 2018, the reference map has been derived by rasterizing the shapefile available from https://unitar.org/maps/map/2806, considering only the Beletweyne urban area. It has been produced by UNOSAT from the analysis of a GeoEye-1 very-high-resolution (VHR) optical image acquired on 30 April, 2018.

For the floods that took place in October–November 2019 and April–May 2020, the same procedure (i.e., rasterization of a shapefile) has been carried out to produce the reference maps. For the former event, the shapefile has been downloaded from https://unitar.org/maps/map/2972, while for the event that occurred in April–May 2020,
the shapefile has been downloaded from https://unitar.org/maps/map/3054. They have been generated by analyzing a VHR WorldView-1 image acquired on 1 November, 2019 and a VHR Worldview-2 image acquired on 13 May, 2020, respectively. Note that the shapefiles were produced using the first clear sky image available to map the flood. Hence, they do not correspond to the situation of maximum flooded area. The result of the rasterization of the shapefile produced for the event that occurred in 2018 (i.e., that used for calibration purposes) is shown in Figure 4.

![Figure 4](image)

**Figure 4.** Reference map derived from the shapefile made available by UNOSAT to provide information about the flood that occurred in Beletweyne in April–May 2018. The boundaries of the districts shown in Figure 1 are highlighted in black.

### 2.3. Temporal Trends of the Interferometric Data

Figures 5 and 6 show, for the calibration data acquired in 2017–2018, the temporal trends of the mean values of \( \gamma \) and \( \zeta \), respectively. They have been produced by computing, for each map of \( \gamma \) and \( \zeta \), the mean values of these quantities (denoted as \( \bar{\gamma} \) and \( \bar{\zeta} \)) over the AOIs shown in Figure 1 (upper panels) and only over the PSs included in the AOIs (lower panels). The AOIs with a small density of PSs have been discarded.
Figure 5. Temporal trend of the mean value of the coherence in the various districts of Beletweyne (upper panel) and in the PSs located in the districts (lower panel). The period June 2017–December 2018 is considered.

Figure 6. Same as Figure 5, but for the spatial average of the exponential of the interferometric phase.
The decrease in \( \gamma \) and \( \xi \) in April–May 2018 can be noted for several districts of Beletweyne town. This decrease is better detected in the lower panels (i.e., considering only the PSs), for which the trends of \( \gamma \) and \( \xi \) are quite stable for the other InSAR pairs. Conversely, larger fluctuations occur considering all the pixels included in the urban AOIs. As discussed in the introduction, these fluctuations can produce errors if, to detect floodwater, \( \gamma \) decreases are searched for, as commonly performed in the literature. This could represent a criticality in view of an operational use of the coherence for urban flood mapping.

As expected, since mean values are considered in Figures 5 and 6, the trends of \( \gamma \) and \( \xi \) are quite similar. More differences between these two quantities will emerge in the following analysis.

2.4. Detection of Flooded PSs Based on Coherence and Spatial Average of the Interferometric Phase

A method for identifying flooded PSs has been developed based on the outcomes of the previous section. As mentioned in the introduction, the goal is to detect anomalous values of \( \gamma \) and \( \xi \). For this purpose, reference values (\( \gamma_{\text{ref}} \) and \( \xi_{\text{ref}} \)) have been determined for each PS by computing the average over time of the values of \( \gamma \) and \( \xi \). The interferometric pairs used to generate Figures 5 and 6 (except those derived from the images acquired in April–May 2018 and on 1 June, 2018) “have been considered to perform this calculation. Then, we have defined the anomalies of \( \gamma \) and \( \xi \) as:

\[
\gamma_{\text{anom}}(x,i) = \gamma_{\text{ref}}(x) - \gamma(x,i)
\]

\[
\xi_{\text{anom}}(x,i) = \xi_{\text{ref}}(x) - \xi(x,i)
\]

where \( x \) represents a PS in this case and \( i \) denotes an interferometric pair.

To calibrate the new method for detecting flooded PSs, the quantities defined by Equations (5) and (6) have been derived for the following interferometric pairs acquired under non-flooded and flooded conditions: 9 March–21 March, 2018, (non-flooded) and 26 April–8 May, 2018 (flooded). The last pair of images has been chosen concerning that, for most of the districts, \( \gamma \) and \( \xi \) have a minimum just in correspondence to this pair (see Figures 5 and 6). For the Kutimbo district, the minimum of \( \gamma \) and \( \xi \) is reached for the pair 14 April–26 April, 2018. Therefore, for the PSs included in this AOI, the values of \( \gamma_{\text{anom}} \) and \( \xi_{\text{anom}} \) for flooded conditions have been extracted from the latter pair. Among the pairs acquired under non-flooded conditions belonging to the calibration stack, the pair 9 March–21 March is the closest in time to the pairs acquired under flooded conditions. Figure 7 shows the statistical distribution of the values of \( \gamma_{\text{anom}} \) and \( \xi_{\text{anom}} \) in the PSs, as well as the scatterplot of \( \gamma_{\text{anom}} \) versus \( \xi_{\text{anom}} \). Only the AOIs classified as totally flooded by UNOSAT (see Figure 4) have been considered to generate this figure.

Looking at the upper and central panels of Figure 7, good discrimination between the blue (non-flooded) and red (flooded) histograms can be noted, especially concerning \( \xi_{\text{anom}} \). To quantify the discrimination capability, the parametric separability index (5), used for instance in [28], has been used here too:

\[
S = \frac{\mu_{\text{non-flooded}} - \mu_{\text{flooded}}}{\sigma_{\text{non-flooded}} + \sigma_{\text{flooded}}}
\]

where \( \mu_{\text{non-flooded}} \) and \( \mu_{\text{flooded}} \) are the mean values of the samples belonging to the classes of non-flooded and flooded PSs, respectively, and \( \sigma_{\text{non-flooded}} \) and \( \sigma_{\text{flooded}} \) are the corresponding standard deviations. The separability between two classes is generally considered good if \( S > 1 \) and the higher the \( S \), the higher the separability [29]. The following \( S \) values have been obtained: 1.35 for \( \gamma_{\text{anom}} \) and 1.47 for \( \xi_{\text{anom}} \). Moreover, only 33 out of 5741 PSs classified as totally flooded in the reference data have \( \xi_{\text{anom}} < 0 \), whereas positive values are expected; this number increases up to 118 for \( \gamma_{\text{anom}} \). Hence, an improvement of the discrimination of the flooded PSs can be achieved also considering \( \xi \).
Figure 7. Statistical distribution of the values of $\gamma_{anom}$ (upper panel) and $\zeta_{anom}$ (central panel) in the PSs and scatterplot of $\gamma_{anom}$ versus $\zeta_{anom}$ (lower panel). Blue lines and dots: data derived from the pair of images acquired on 9 March and 21 March under non-flooded conditions; red lines and dots: data derived from the pair of images acquired on 26 April–8 May, 2018 under flooded conditions (except for the PSs included in the Kutimbo district whose data were derived from the pair 14 April–26 April, 2018).

From the lower panel of Figure 7, it can be deduced that the populations of flooded and non-flooded PSs form two different clusters. In particular, by applying a test-and-trial method, it has been found that the line that separates the clusters of red dots and blue dots (dashed line in the lower panel of Figure 7), i.e., that contemporary maximizes the
number of blue dots below the line and the number of the red dots above the line, has the following equation:

\[
\zeta_{anom}(x, i) = -0.84 \cdot \gamma_{anom}(x, i) + 0.27
\]  

(8)

Some blue dots above the line defined by Equation (8) are visible in the lower panel of Figure 7; they correspond to non-flooded PSs misclassified as flooded ones. To decrease the risk of making commission errors, flooded PSs identified through Equation (8) have been intersected with those identified by applying thresholds on \( \gamma_{anom} \) and \( \zeta_{anom} \). For this purpose, the 95th percentile of their distribution for the pair 9 March–21 March, 2018 (blue lines in the upper and central panels of Figure 7) has been computed, thus deriving the following combination of inequalities:

\[
\zeta_{anom}(x, i) > 0.25 \quad \text{OR} \quad \gamma_{anom}(x, i) > 0.20
\]  

(9)

Flooded PSs have been identified through a logical AND of Equations (8) and (9):

\[
[\zeta_{anom}(x, i) > -0.84 \cdot \gamma_{anom}(x) + 0.27] \quad \text{AND} \quad [\zeta_{anom}(x, i) > 0.25 \quad \text{OR} \quad \gamma_{anom}(x, i) > 0.20]
\]  

(10)

2.5 Validation

To validate the methodology described in the previous section, the reference maps derived from the shapefiles produced by UNOSAT for the floods that affected Beletweyne in October–November 2019 and April–May 2020 have mainly been used. They have been compared to the maps of flooded PSs obtained for the pairs 6 October–30 October, 2019 and 9 May–21 May, 2020. For the AOIs classified as partially flooded by UNOSAT, it was not possible to distinguish flooded pixels from non-flooded ones. Moreover, the class of non-flooded pixels is not included in the UNOSAT data. Consequently, we have been able to compute only the omission error (\( \epsilon_{om} \)) for the flooded PSs, corresponding to the PSs included in the AOIs labelled as totally flooded by UNOSAT that have been classified as non-flooded by our method.

As discussed in the introduction, in view of an operational use of the InSAR data for detecting UFAs, it is important to evaluate not only \( \epsilon_{om} \) but also the commission error (\( \epsilon_{comm} \)). For this purpose, a map of flooded PSs has been produced not only for the two InSAR pairs mentioned above, but also for the pairs of images acquired at consecutive times in the validation period January 2019–7 July, 2020 (normal font in Table 1), that is at \( t_i \) and \( t_{i+1} \) (\( i = 48, 80, 33 \) interferometric pairs). The percentage of PSs classified as flooded with respect to the total number of PSs has been calculated to determine \( \epsilon_{comm} \) considering only the period when no floods were reported (the pairs acquired in October–November 2019 and April–May 2020 have therefore been excluded). Moreover, to avoid working with long temporal baselines, even the pairs acquired on 9 April–27 May, 2019, 23 November, 2019–3 February, 2020, and 15 February, 2020–22 March, 2020 have been discarded.

To further consolidate the validation by comparing the maps produced by applying Equation (10) to the UNOSAT-derived ones, a simple way to label the AOIs has been designed. An AOI included in the validation maps has been classified as non-flooded, partially flooded, or totally flooded based on the percentage of the PSs classified as flooded with respect to the total number of PSs identified in each AOI (\( \text{PSpert} \)). The following thresholds have been used:

1. \( \text{PSpert} < 25\% \): non-flooded;
2. \( \text{PSpert} \geq 25\% \) AND \( \text{PSpert} \leq 75\% \): partially flooded;
3. \( \text{PSpert} > 75\% \): totally flooded.

For the two events considered for validation, an AOI-based confusion matrix has been calculated, by comparing, for each AOI in the reference maps, the UNOSAT-derived classification with that produced by applying the thresholds on \( \text{PSpert} \) listed above. The AOIs including few PSs (less than 5% of the total number of pixels) have not been classified except for the “Unknown” district (dark gray in Figure 1), which is basically a rural area.
The latter has been classified based on the percentage of open water pixels with respect to the total number of bare soil pixels.

3. Results

Figure 8 presents the maps of flooded areas derived by combining the maps of flooded PSs with the maps of flooded rural/suburban areas. The former has been produced by applying Equation (10) to the pairs 6 October–30 October 2019 and 9 May–21 May 2020. The latter has been generated by applying, to the images acquired on 30 October 2019 and 21 May 2020 (see Section 2.2.2), the algorithm designed in [4]. These combined maps represent an example of what can be derived from the procedure proposed in this study. The following values of $\epsilon_{\text{con}}$ have been obtained: 18% for 6 October–30 October, 2019 and 17% for 9 May–21 May 2020.

Figure 8. Maps of flooded areas for 30 October, 2019 (top panel) and 21 May, 2020 (bottom panel). Dark blue: rural/suburban areas; light blue: flooded PSs. The boundaries of the districts shown in Figure 1 are highlighted in black.
Figure 9 shows a visual comparison between reference maps and PSpert-derived ones, while the AOI-based confusion matrices are reported in Tables 2 and 3, which also report the values of the overall accuracy (OA) and the kappa coefficient (κ). Looking at the right panels of Figure 9, it can be noted that none of the AOIs have been classified as non-flooded by our method, while some small AOIs have not been classified in the PSpert-derived maps because of the low percentage of PSs (less than 5%).

Concerning the commission error, its value is shown in Figure 10 (blue line) for the pairs included in the validation period (except those acquired in October–November 2019 and April–May 2020). For most of the pairs, $\epsilon_{\text{comm}}$ is less than 10% and never exceeds 15% (maximum value 13.1%).

Figure 9. Comparison between the UNOSAT-derived reference maps (upper left: October–November 2019, lower left: April–May 2020) and PSpert-derived maps (upper right: pair 6–30 October 2019, lower right: 9–21 May 2020). The boundaries of the districts shown in Figure 1 are highlighted in black.
Table 2. Results of the comparison between UNOSAT-derived maps and PS-derived maps for the flood that occurred in October November 2019. OA: overall accuracy. Columns represent classes in the reference maps.

<table>
<thead>
<tr>
<th>October–November 2019</th>
<th>Totally Flooded</th>
<th>Partially Flooded</th>
<th>Total</th>
<th>OA</th>
<th>κ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Totally flooded</td>
<td>11</td>
<td>3</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partially flooded</td>
<td>0</td>
<td>5</td>
<td>5</td>
<td>84.2</td>
<td>0.66</td>
</tr>
<tr>
<td>Total</td>
<td>11</td>
<td>8</td>
<td>19</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Same as Table 2, but for the flood that occurred in April–May 2020.

<table>
<thead>
<tr>
<th>April–May 2020</th>
<th>Totally Flooded</th>
<th>Partially Flooded</th>
<th>Total</th>
<th>OA</th>
<th>κ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Totally flooded</td>
<td>7</td>
<td>1</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partially flooded</td>
<td>2</td>
<td>7</td>
<td>9</td>
<td>82.4</td>
<td>0.65</td>
</tr>
<tr>
<td>Total</td>
<td>9</td>
<td>8</td>
<td>17</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 10. Temporal trend of the commission error for the validation period (except October–November 2019 and April–May 2020). Blue: flooded PSs identified according to Equation (10); red: flooded PSs identified based on Equation (9); yellow: flooded PSs identified according to Equation (8).

4. Discussion

This paper proposes a method for tackling the problem of UFAs mapping based on the assumption that maps of PSs complemented by metadata derived from InSAR products are available.

S1 allows for a continuous streaming of SAR-derived added value products; since SAR intensity-based algorithms are presently mature for an operational application, one of the first products that is expected to be available at a global scale is represented by frequently updated maps of rural/suburban flooded areas. Considering also that PSs will be used in the future to provide an operational monitoring of surface displacement phenomena (an example is the new European Ground Motion Service recently presented in the framework of the Copernicus Land Monitoring Service: https://land.copernicus.eu/user-corner/technical-library/european-ground-motion-service), another product whose availability might be expected in the future is a global map of S1-derived PSs. If these two products are available, once a rural/suburban area surrounding an urban settlement is labelled as flooded, it can be soon verified whether an anomalous behavior of $\gamma$ and $\zeta$ occurs in the settlement.

With respect to previous papers dealing with the same topic [2,12,14,20], here a joint use of phase statistics and coherence data is carried out. The use of phase statistics has been previously suggested in [23], but instead of the coherence, not jointly. As expected, $\gamma$ and $\zeta$
are quite correlated (see the right panel of Figure 7); nonetheless, pixels whose backscattering is so high and stable even under flooded conditions that \( \gamma \) does not show significant changes can be detected by accounting for \( \zeta \). It has been found that, for the October–November 2019 event, more than 12% of flooded PSs has \( \zeta_{\text{anom}} > 0.25 \) and \( \gamma_{\text{anom}} < 0.20 \); they would have likely been classified as non-flooded by considering only coherence. Almost half of them have a high intensity (>2 dB) in the post event image (30 October 2019). Similar percentages (slightly smaller) have been obtained by considering the April–May 2010 event.

The advantage of focusing on the PSs identified by processing a stack of SAR data is that the analysis is based only on pixels that are not affected by variations of the coherence due to, for instance, the presence of vegetation inside the urban settlement. Considering that in the upper panels of Figures 5 and 6, average values of \( \gamma \) and \( \zeta \) are shown, even larger fluctuations of these quantities can be expected in some urban pixels not identified as PSs; this may affect the reliability of the detection of the possible presence of floodwater. By simply applying a threshold (0.2) on the difference between \( \gamma_{\text{pre-event}} \) and \( \gamma_{\text{co-event}} \) considering all the pixels included in the urban AOIs (not only the PSs), as commonly done when searching for UFAs using SAR data, omission errors larger than 30% have been obtained for both the test cases considered for validation. It has been found that most of these errors have been caused by pixels having low coherence in the pre-flood pairs. Note that even for this exercise, only the AOIs labelled by UNOSAT as totally flooded have been considered.

By comparing the maps shown in the left and right columns of Figure 9, fairly good agreement emerges between the reference maps and PSperc-derived ones. This agreement is confirmed by the values of OA (larger than 82%) and \( \kappa (\sim 0.65) \) reported in Tables 2 and 3. It must be considered that the different acquisitions times of the images used by UNOSAT and the S1 ones might have an impact on the results. In this case, the temporal scales of flood progression and recession were quite slow because the floods lasted several weeks. Hence, the impact on the results is limited. Note that similar results in terms of OA and \( \kappa \), as well as \( \epsilon_{\text{om}} \) have been obtained for the October–November 2019 flood and the April–May 2020 flood, even though for the former case study, the difference between acquisition times of reference and S1 data was only one day, while for the other test case, it was 8 days.

Equation (10) is a combination, through a logical AND, of Equations (8) and (9). The omission error obtained by applying Equation (10) is around 17–18%; \( \epsilon_{\text{om}} \) is due to point targets whose reflectivity is not very influenced by the presence of floodwater. Therefore, the anomalies of \( \gamma \) and \( \zeta \) are not high enough to be above the thresholds used in Equation (10), or to produce a point in the scatterplot of \( \zeta_{\text{anom}} \) versus \( \gamma_{\text{anom}} \) that is above the line defined by Equation (8). Note that by using only the thresholds, i.e., Equation (9), the reduction in \( \epsilon_{\text{om}} \) is not very significant (from 18% to 17% for the event that occurred in 2019 and from 17% to 16% for the event that occurred in 2020). A slightly larger reduction has been achieved by using only the classification based on Equation (8); in this case, values of \( \epsilon_{\text{om}} \) equal to 14% (October 2019) and 13% (May 2020) have been obtained. However, the reduction in \( \epsilon_{\text{om}} \) is compensated by an increase in \( \epsilon_{\text{comm}} \). Looking at Figure 10, it can be noted that quite large increases in this error have been obtained by using only Equation (8) or only Equation (9). In the literature, even better results have been reported (e.g., in [20]), but this was by either analyzing small AOIs within an urban area and/or also including rural/suburban areas in the evaluation of the performances.

Previous papers dealing with the detection of UFAs using coherence analyzed few interferometric pairs, typically 1–2 pre-event pairs (except [20], where 9 pre-event pairs were used), a co-event pair, and possibly a post event pair. In this study, considering only the validation set, 30 pairs have been processed. In this way, a continuous monitoring and mapping of the flood situation in Beletweyne throughout 2019–2020 has been provided as an example of the potentiality of the proposed method in view of an operational application. Moreover, commission errors have been reliably evaluated. It is understood that the analysis of data related to the flood phenomenon should be based on multi-annual data.
From this point of view, the research period is quite short. However, S1 observations of Beletwewyne are available only since 2015. In the period 2015–2020, reference data are available only for the three floods considered in this work. Hence, even enlarging the stack of S1 images, it would not have been possible to take into account other flood events to calibrate/validate the method.

As for the limitations of the proposed method, the first one is the necessity to have a dense grid of PSs to achieve a reliable evaluation of where UFAs are present. This problem could be overcome by complementing the PSs map with a map of the buildings, as proposed in [2], although a tool for building detection such as that developed in [2] would be needed in this case. In addition, since PSs maps are currently unavailable at a global scale, an amount of work to gather a stack of S1 data and extract the PSs in the built-up area has been performed. Such an effort is worthwhile only for urban areas prone to recurrent flooding, such as Beletwewyne.

The use of fixed thresholds for $\gamma_{\text{anom}}$ and $\zeta_{\text{anom}}$ might represent another critical point, although the identification of flooded PSs is not based only on these thresholds. The values chosen in this study have been derived from the analysis of the distributions of the values of $\gamma_{\text{anom}}$ and $\zeta_{\text{anom}}$. The use of the 95th percentile of the distributions obtained considering InSAR pairs acquired under non-flooded conditions allowed us to achieve a good compromise between omission and commission errors. Note that by increasing the thresholds used in Equation (10) to 0.30 and 0.25 for $\gamma_{\text{anom}}$ and $\zeta_{\text{anom}}$, respectively, the AOI-based OA reduces to 0.76 (2019) and 0.73 (2020), while obviously the omission error increases, reaching values of about 25% for both events. Conversely, by decreasing the thresholds to 0.20 and 0.15 for $\gamma_{\text{anom}}$ and $\zeta_{\text{anom}}$, respectively, the AOI-based OA reduces to 0.70 (2019) and 0.73 (2020), even though $\zeta_{\text{anom}}$ decreases as expected.

This paper proposes a general procedure to detect flooded PSs, which consists of the identification of the PSs and the analysis of two maps of $\gamma_{\text{anom}}$ and two maps of $\zeta_{\text{anom}}$ : a group of two maps derived from a pre-event pair and the other group with $\gamma_{\text{anom}}$ and $\zeta_{\text{anom}}$ derived from a co-event pair. The availability of reliable information about flood extent is fundamental for an accurate determination of the thresholds and of the line separating the two clusters shown in the lower panel of Figure 7. Only in urban areas prone to recurrent flooding a similar training set is expected to be available at present. On the other hand, the fact that by performing the calibration using a stack of data acquired in 2017–2018, good results have been obtained even using SAR observations performed three years later (2020), this might be considered an indication that Equation (10) can be applied even for other urban areas obtaining fairly good results.

5. Conclusions

The analysis of a time series of Sentinel-1 InSAR data has led to the proposal of a new method to monitor the presence of floodwater in flood-prone urban areas. The method takes advantage of the pre-programmed operation mode of Sentinel-1 which continuously provides consistent data acquired in Interferometric Wide Swath mode and allows users to generate a long-term archive of interferograms. The method assumes the availability of a map of PSs and requires a calibration step in which, for each PS, a reference value of the coherence and the spatial average of the exponential of the phase must be computed. Once this training phase has been accomplished, as soon as new S1 data are available, it is possible to quickly verify whether one or more clusters of PSs show an anomalous decrease in the coherence and/or the spatial average of the exponential of the phase. The comparison between the maps of flooded PSs generated by applying the proposed method and reference maps produced by UNOSAT by analyzing VHR optical images provided promising results in terms of omission error and commission error. The latter has been evaluated by considering a stack of Sentinel-1 data acquired under non-flooded conditions.

With respect to the use of an approach based on coherence change detection considering a whole urban area, by focusing on the PSs, the errors due to variations of the coherence caused, for instance, by the presence of unpaved roads among the buildings or vegetation
inside the urban settlement can be limited. Since, in some PSs, the coherence may remain high even in the presence of floodwater in the spatial window considered to compute it, adding the spatial average of the exponential of the interferometric phase has allowed us to reduce the omission error, without an excessive increase in the commission error.

Although focusing on the PSs, a smaller number of urban pixels are monitored, they can be considered a sparse grid of measurements giving an indication of the zones of an urban settlement where floodwater is present. From this point of view, the larger the PS density, the higher the reliability of the flood mapping.

This study has been intended to open the way for a new urban flood monitoring system consisting of the systematic processing of Sentinel-1 data based on the PS multipass approach, while, at present, only the routine elaboration of intensity data is carried out for operational flood mapping service in rural/suburban areas. Considering that, at present, an amount of work is required to gather a stack of Sentinel-1 data to identify the PSs (while the availability of PSs maps can be expected in the future) and determine the reference values for each PS, i.e., to perform the calibration, it is currently worthwhile to apply the proposed method to flood-prone urban areas like Beletwye

To consolidate the proposed methodology, a more comprehensive investigation considering a longer research period will be carried out in the future in order to overcome the time limitations of the data used in this study. In addition, the method will be tested on other flood-prone urban areas to further verify its performances.

Author Contributions: Conceptualization, L.P., M.C., and N.P.; methodology, L.P.; validation, L.P.; formal analysis, L.P.; investigation, L.P., M.C., and N.P.; resources, M.C.; data curation, L.P.; writing—original draft preparation, L.P.; writing—review and editing, L.P., M.C., and N.P.; supervision, N.P.; project administration, M.C.; funding acquisition, M.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Research Fund of Luxembourg (FNR) through the MOSQUITO (grant number C15/SR/10380137) and the CASCADE (grant number C17/SR/11682050) projects.

Data Availability Statement: Sentinel-1 data are available at https://scihub.copernicus.eu/dhus/#!/home. UNOSAT data are available at https://unitar.org/maps.

Conflicts of Interest: The authors declare no conflict of interest.

References