Natural Hazards
and Earth System
Sciences

Open Access

EGU

Supplement of

Exposure of real estate properties to the 2018 Hurricane Florence flooding

Marco Tedesco et al.

Correspondence to: Marco Tedesco (mtedesco@ldeo.columbia.edu)

1   Supplementary material

In this Supplementary material, we report a synthesis and a discussion of the SAR-based approach. Limitations in the case of SAR can arise from features mimicking the behavior of flooded regions (such as some roads or airport runways) and from the presence of natural or manmade structures (such as trees or buildings) that tend to mask the signal associated with the presence of water on the surface (e.g., Shumann et al., 2018; Notti et al., 2018). SAR data has been shown to be able to detect the presence of flooded areas by means of approaches ranging from automated ones (e.g., Bazi et al., 2005; Moser and Serpico, 2006; Huang et al., 2018; Twele et al., 2016) to sophisticated techniques, such as active contour models (ACM) or change detection (e.g., Landuyt et al., 2019), to simpler ones such as threshold-based approaches (e.g., Shumann et al., 2018). In this study, we detect flooded areas from SAR assuming that the distribution of the recorded backscattering values can be approximated with a bimodal distribution in which dry (wet) pixels belong to the right (left) normal distribution (e.g., Otsu, 1979; Chini et al., 2017). The threshold value separating the two distributions can be computed through the minimization of a cost function that reflects the amount of overlap between the Gaussian density functions of the foreground and background classes (e.g., Kittler and Illingworth, 1989). As an example, Figures S1a and S1b shows maps of, respectively, backscattering coefficient values σ_{0VH} (Figure S1a, VH polarization) obtained by Sentinel-1A collected on 2 September, 2018 at 23:05:52 UTC and land use land cover attributes obtained from the National Geospatial Data Asset (NGDA) Land Use Land Cover (LULC) dataset (https://www.sciencebase.gov/catalog/item/581d050ce4b08da350d52363, Figure S1b) from the National Land Cover Database (NCLD, e.g. Jin et al., 2011; Xian et al., 2011) over a subset region where flooded due to Florence occurred. Specifically, we use the "Open Water - All areas of open water" class (Class # 11) to map the presence of permanent water bodies. The original 30 m dataset is projected on the same coordinate system as the radar images (WGS84) and it is resampled to 10 m through bilinear interpolation to match the SAR data spatial resolution. Figure S1c shows the histogram (and the fitted bimodal distribution) of the backscattering data in Figure S1a, together with the values of the mean and standard deviation of the two normal distributions, the computed threshold

coefficients as well as the Ashman coefficient (Ashman, 1994). This coefficient allows to automatically quantifying the degree of separation between the two Gaussian distributions: the higher the Ashman coefficient the more the two normal distributions are separated. Here, we adopt the minimum value of 2 on the Ashman coefficient to ensure separability (Chini et al., 2017). For those tiles where the Ashman coefficient is below 2, we use the mean values of the coefficients surrounding that tile. We assess also the use of two ways to compute the threshold value on the radar backscattering coefficients: the first one (σ_{0Thr1}) is obtained from the identification of the value where the two fitted Gaussian distributions intersect; the second threshold value is computed as $\sigma_{th2} = \mu2 - 2*\sigma2$, with $\mu2$ and $\sigma2$ being, respectively, the mean and standard deviations of the "dry" distribution. More details about the SAR-based approach and its assessment are reported in the Supplementary Material.

Once water-covered flooded areas are identified, permanent water bodies and flooded areas are separated into two distinct classes are excluded using features from the United States Fish and Wildlife Service's National Wetlands Inventory and the United States Geological Survey's National Hydrography Dataset (https://fwsprimary.wim.usgs.gov/wetlands/apps/wetlands-mapper/). The financial analysis is, then, performed only over those areas that are covered by water but that do not belong to the permanent water body class. Whenever available, digitized water features from local sources are incorporated in order to create the most detailed water boundary delineations possible. To produce the final file used to mask permanent water, the classification fields representing deep existing water ("Estuarine and Marine Deepwater"), Ponds ("Freshwater Ponds"), Lakes, and Rivers ("Riverine") were merged into a single mask. The "Estuarine and Marine Wetland", "Freshwater Emergent Wetland", and "Freshwater Forested/Shrub Wetland" were not included in the original mask as these represent land classification types that are not permanently wet.
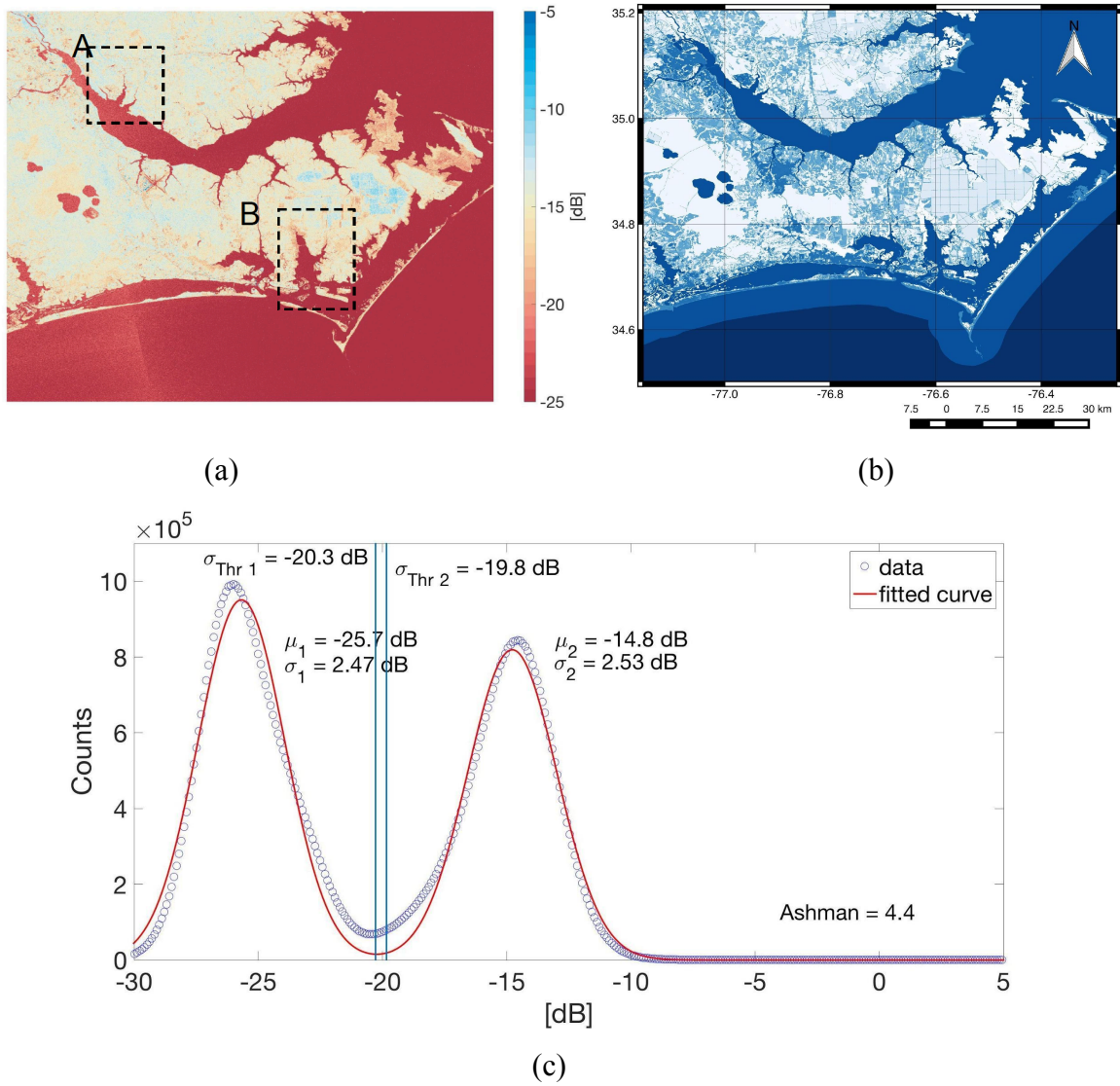
(a)

(b)

(c)

Figure S1 Map of VH backscattering coefficients obtained from Sentinel-1 collected on 2 September, 2018 at 23:05:52 UTC over a region impacted by Florence b). Same as a) but showing land cover attributes from the National Geospatial Data Asset (NGDA) Land Use Land Cover (LULC) dataset data record (https://www.sciencebase.gov/catalog/item/581d050ce4b08da350d52363). Here, water is represented as medium blue. c) Histogram and fitted bimodal distribution of the backscattering data in a). In the panel, the mean and standard deviation of the two normal composing the bimodal distributions are reported, together with computed thresholds and the Ashman coefficient.

3

The list of the radar images used in this study to map the flood extent is reported in Fig. S2. The naming convention for the files obtained from https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-1-sar/products-algorithms/level-1-product-formatting is also reported in the figure.
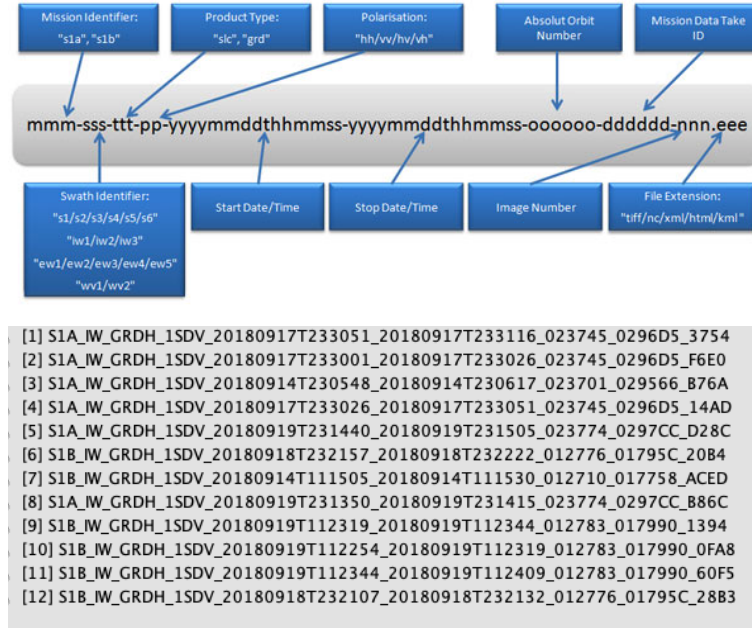


```
[1]  S1A_IW_GRDH_1SDV_20180917T233051_20180917T233116_023745_0296D5_3754
[2]  S1A_IW_GRDH_1SDV_20180917T233001_20180917T233026_023745_0296D5_F6E0
[3]  S1A_IW_GRDH_1SDV_20180914T230548_20180914T230617_023701_029566_B76A
[4]  S1A_IW_GRDH_1SDV_20180917T233026_20180917T233051_023745_0296D5_14AD
[5]  S1A_IW_GRDH_1SDV_20180919T231440_20180919T231505_023774_0297CC_D28C
[6]  S1B_IW_GRDH_1SDV_20180918T232157_20180918T232222_012776_01795C_20B4
[7]  S1B_IW_GRDH_1SDV_20180914T111505_20180914T111530_012710_017758_ACED
[8]  S1A_IW_GRDH_1SDV_20180919T231350_20180919T231415_023774_0297CC_B86C
[9]  S1B_IW_GRDH_1SDV_20180919T112319_20180919T112344_012783_017990_1394
[10] S1B_IW_GRDH_1SDV_20180919T112254_20180919T112319_012783_017990_0FA8
[11] S1B_IW_GRDH_1SDV_20180919T112344_20180919T112409_012783_017990_60F5
[12] S1B_IW_GRDH_1SDV_20180918T232107_20180918T232132_012776_01795C_28B3
```

Fig. S2 Naming convention and names of the Sentinel-1 files used in this study.

We pre-process Sentinel – 1 data using the Science Toolbox Exploitation Platform (SNAP) Toolkit developed by ESA (http://step.esa.int/main/toolboxes/snap/). Specifically, we first apply orbit and radiometric corrections, followed by filtering the dataset using a Lee-Sigma filter (Lee et al., 2009) with a 5 x 5 kernel to reduce speckle noise and terrain flattening and correction. Lastly, we convert backscatter intensities into a decibel-scale (dB) and export the image as a GeoTIFF with a 10 m spatial resolution.

As mentioned in the description of methods, a major issue can arise from the absence of a clear bimodality of the distribution of the recorded backscattering values, as the algorithm might not be able to select a realistic threshold for separating flooded and not-flooded regions. To mitigate this problem, following the approaches by Bovolo and Bruzzone (2007) and Martinis et al. (2015), we divide the SAR image in tiles. In order, to estimate an optimal number of tiles we compute the Ashman coefficient (Ashman et al., 1994).

The higher the Ashman coefficient the more the two normal distributions are separated. For this study, we adopt a minimum value of 2 on the Ashman coefficient to ensure separability (Chini et al., 2017). For those tiles when the Ashman coefficient is below 2, we use the mean values of the coefficients surrounding that tile.

We use the data collected before Florence on 2 September, 2018 to quantify the omission and commission errors of the radar-based technique by comparing the flooded areas identified by Sentinel-1 with those present in the LULC map from the National Land Cover Database (NCLD, e.g. Jin et al., 2011; Xian et al., 2011). In Table S1, the term "Match" refers to the percentage of the areas identified as water by Sentinel-1 matching the water in the LULC dataset. "False positive" refers to the areas where Sentinel-1 suggests water but this is not the case for the LULC dataset. In the Table we report the results in the case of no erosion/dilation of the Sentinel flooded maps and when erosion and dilation (in this order) are applied with a square filter of 5 x 5, 10 x 10 and 20 x 20 pixels. The matching between Sentinel-derived and LULC water masks exceeds 99 % in many of the cases using the VH polarization with false positive values being of the order of 5 %. Overall, best results are obtained in the case of VH polarization with a erosion/dilation window of 5 pixels, showing a matching percentage greater than 99 % and the lowest false positive value. False positive values are driven not only by the overestimation of flood extent by the radar due to noise or spatial features that might mimic the electromagnetic behavior of flooded regions but also to errors in the LULC map. Figures S3 and S4 show the LULC classes (coded according to the original LANDSAT LULC product) and the Sentinel-1 backscattering coefficients (VH) over two areas within the area highlighted in the squares in Figure 1a together the water masks obtained from the LANDSAT and from the Sentinel-1 using the σ_{th2} threshold. The visual inspection of the water masks from LANDSAT indicates that LANDSAT tends to misclassify water in those areas where water bodies such as relatively narrow streams or mixed pixels occur, as highlighted, for example, by the squares in the figures. This increases our confidence in the inundated areas maps obtained from Sentinel-1 using the σ_{th2} threshold value. Therefore, we used this option to extract the pas used to detect flooded areas and merge them with those obtained from FEMA.
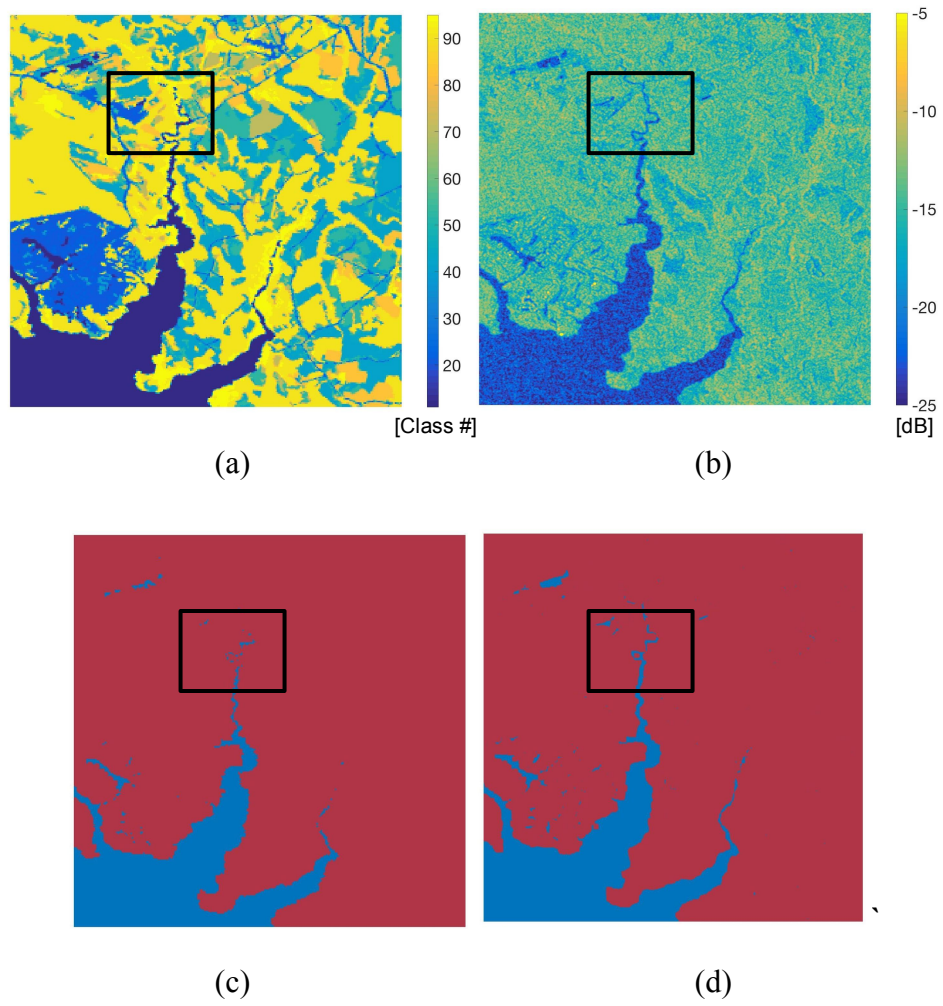
5

Figure S3 a) LULC map of the area contained in the rectangle A in **Figure 1a. b)** map of Sentinel-1 backscattering coefficient over the same region as in **Figure 1. c)** and **d)** water masks (blue indicates the presence of water) obtained from c) the LULC Landsat and d) Sentinel-1 using the σ_{Thr1} as a threshold coefficient and dilating and eroding by 10 pixels using a square filter.
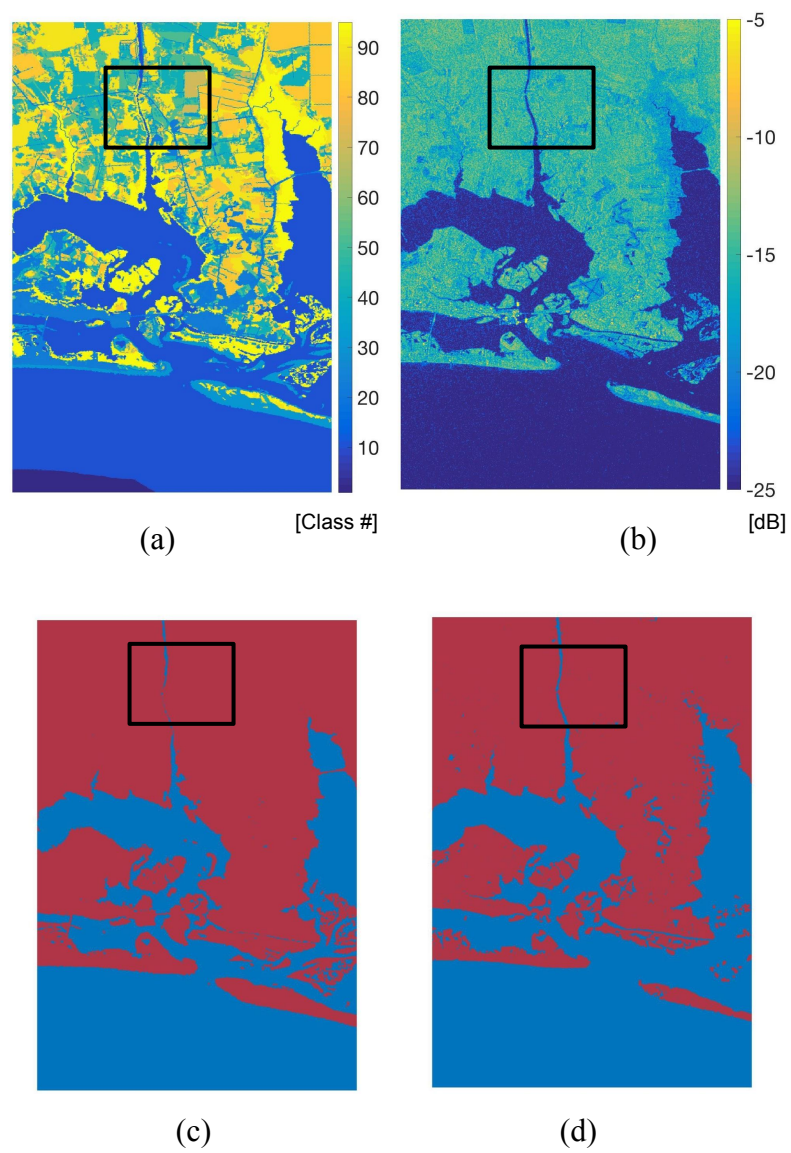
(a)   [Class #]   (b)   [dB]

(c)   (d)

Figure S4 Same as Figure 14 but for the area B indicated in Figure S1.

			No eros./dil.	5 x 5	10 x 10	20 x 20
VV	σ_{Thr1}	Match	97.23 %	98.34 %	98.37 %	98.24%
		False positive	5.43 %	5.22 %	6.43 %	9.60 %
	σ_{Thr2}	Match	98.11 %	99.14 %	99.30 %	99.48 %
		False positive	5.03 %	5.53 %	5.85 %	8.78 %
VH	σ_{Thr1}	Match	99.18 %	99.69 %	99.77 %	99.8 4%
		False positive	4.58 %	4.88 %	5.83 %	8.56 %
	σ_{Thr2}	Match	98.21 %	99.52 %	99.70 %	99.78 %
		False positive	4.08 %	4.33 %	4.96 %	6.86 %

Table S1 Percentage of matching and false positive between the LANDSAT-based water classification and the Sentinel-1 estimated water mask. Here, "match" refers to the case when both LANDSAT and Sentinel-1 indicate water is present. In this case, the percentage is computed using the total number of pixels classified as water by LANDSAT. "False positive" refers to the case when LANDSAT does not indicate water but Sentinel-1 does. In this case, the percentage is relative to the total number of pixels in the image minus the number of pixels classified by water in the LANDSAT dataset. No eros.dil. refers to the values obtained from the radar water masks without any dilation and/or erosion applied. The following columns indicate the values obtained when applying, in this order, dilation and erosion with a square of 5, 10 and 20 pixels, respectively.