Aberystwyth University

Automatic Detection of Open and Vegetated Water Bodies Using Sentinel 1 to Map African Malaria Vector Mosquito Breeding Habitats

Hardy, Andy; Ettritch, Georgina; Cross, Donall; Bunting, Pete; Liywalii , Francis ; Sakala , Jacob ; Silumesii , Andrew ; Singini , Douglas ; Smith, Mark W.; Willis, Thomas; Thomas, Chris

Article

Automatic Detection of Open and Vegetated Water Bodies Using Sentinel 1 to Map African Malaria Vector Mosquito Breeding Habitats

Andy Hardy [1],* , Georgina Ettritch [1], Dónall E. Cross [2], Pete Bunting [1], Francis Liywalii [3], Jacob Sakala [3], Andrew Silumesii [4], Douglas Singini [5], Mark Smith [6], Tom Willis [6] and Chris J. Thomas [2]

[1] Department of Geography and Earth Sciences, Aberystwyth University, Aberystwyth SY23 3DB, UK; gee25@aber.ac.uk (G.E.); pfb@aber.ac.uk (P.B.)

[2] Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Aberystwyth SY23 3DA, UK; dec8@aber.ac.uk (D.E.C.); cjt@aber.ac.uk (C.J.T.)

[3] Provincial Health Office, Western Province, P.O. Box 910022, Mongu, Zambia; franliywali@yahoo.com (F.L.); sakalajac57@gmail.com (J.S.)

[4] Ministry of Health, P.O. Box 30205, Lusaka, Zambia; asilumesii@gmail.com

[5] Limulunga District Health Office, P.O. Box 910022, Limulunga, Zambia; dsmlup@gmail.com

[6] School of Geography, University of Leeds, Leeds LS2 9JT, UK; m.w.smith@leeds.ac.uk (M.S.); t.d.willis@leeds.ac.uk (T.W.)

* Correspondence: ajh13@aber.ac.uk

Abstract: Providing timely and accurate maps of surface water is valuable for mapping malaria risk and targeting disease control interventions. Radar satellite remote sensing has the potential to provide this information but current approaches are not suitable for mapping African malarial mosquito aquatic habitats that tend to be highly dynamic, often with emergent vegetation. We present a novel approach for mapping both open and vegetated water bodies using serial Sentinel-1 imagery for Western Zambia. This region is dominated by the seasonally inundated Upper Zambezi floodplain that suffers from a number of public health challenges. The approach uses open source segmentation and machine learning (extra trees classifier), applied to training data that are automatically derived using freely available ancillary data. Refinement is implemented through a consensus approach and Otsu thresholding to eliminate false positives due to dry flat sandy areas. The results indicate a high degree of accuracy (mean overall accuracy 92% st dev 3.6) providing a tractable solution for operationally mapping water bodies in similar large river floodplain unforested environments. For the period studied, 70% of the total water extent mapped was attributed to vegetated water, highlighting the importance of mapping both open and vegetated water bodies for surface water mapping.

1. Introduction

Providing timely and accurate maps of surface water is valuable for a range of applications. Mapping flood water using satellite earth observation technologies is a routine practice, assisting emergency services as well as informing flood mitigation strategies, e.g., [1–5]. Additionally, satellite imagery has been successfully used to delineate and characterise wetland areas, providing spatial information for monitoring and managing wetland habitats, e.g., [6–10]. There is also considerable potential for using satellite-derived water body products for operational use in public health.

The burden of malaria remains a significant global challenge. In 2017 there was an estimated 219 million cases worldwide, resulting in approximately 435,000 deaths [11]. Malaria is an infectious disease that is intrinsically linked to water as mosquito vectors (a variety of species of the genus *Anopheles*) rely upon aquatic breeding sites. Yet operational tools to map relevant water bodies for public health have yet to be developed. In sub-Saharan Africa, which bears 80% of global mortality [11], the prevalent vectors breed in open or vegetated water bodies that are detectable either as an individual unit (e.g., bodies of water within floodplains and dry river channels, borrow pits, spring fed ponds, irrigated fields) or as a habitat containing smaller water bodies (e.g., water-filled depressions such as hoofprints).

Previous studies have demonstrated that remote sensing imagery can be used to map these features and predict spatial variations in transmission risk [12–15]. As such, these water bodies, or aquatic habitats, could be targeted with disease control interventions such as larviciding–essentially, restricting the number of mosquito larvae that develop through to the adult mosquito stage, thereby reducing the density of mosquitoes that might transmit malaria and other diseases [16]. This approach, often referred to as Larval Source Management (LSM), has been used successfully in the past (see for example Killeen [17]). However, in the seasonal climates prevalent across much of Africa, mosquito-producing water bodies are often highly variable and ephemeral, varying both within and between years [18]. As mosquitoes have short development times from eggs to adults (10–14 days) the efficacy and efficiency of interventions such as LSM is reliant on our ability to provide up to date, accurate information on the location of water bodies [19]. Synthetic aperture radar imagery has the potential to provide this information [20], but for use in many sub-Saharan settings significant image processing issues remain in discriminating both open and vegetated water-bodies.

A number of approaches for mapping open water are based on multi-temporal imagery, identifying inundated areas with reference to their radar backscatter during drier periods [2,5,21,22]. While providing an accurate delineation of flood water (reported accuracies ~95%), this approach is of limited use for mapping water bodies during the dry season. This represents an important deficiency for use in mapping mosquito aquatic habitats for LSM: larviciding is likely to be more efficient during the dry season when aquatic habitats are smaller and less abundant [23] and therefore less labour and larviciding product is needed. A dry season larviciding campaign therefore targets both the spatial 'pinch-point' where vector habitats are most restricted, and the temporal 'pinch-point' in vector population dynamics when the population is smallest and most vulnerable to intervention. As such, it is important to design an approach for mapping mosquito aquatic habitats using single input epoch images that is appropriate for mapping water bodies in the dry season. In addition, automated classification is an important step towards operational feasibility by African malaria elimination campaigns.

Radar imagery, unlike optical sources, is not impaired by cloud cover and from systems such as the European Space Agency's Sentinel-1, is freely available in a timely manner (6 days revisit period) for much of the globe. This effectively guarantees sufficient temporal resolution to capture the highly dynamic nature of many aquatic habitats. Due to their relatively smooth texture, open water bodies can often be distinguished from other land cover types due to their low backscatter response. Consequently, there is a large body of literature proposing methods for automatically detecting open water using satellite radar systems [1–5,24–30]. Many approaches with operational capability are based on automatic image histogram thresholding [1,2,5,19,22,23,25,26] and when used in combination with ancillary data (e.g., land cover maps and terrain-based derivatives) and other image-based refinements, can achieve a high degree of accuracy (often >95%). However, these approaches are only concerned with mapping open water. In public health, mosquito aquatic habitats can often be found in vegetated water bodies including irrigated cropland, such as rice paddies, e.g., [31–34] and more naturally occurring water bodies with riparian vegetation, such as springs, the fringes of river channels and areas of topographic convergence, e.g., [35–37].

When water bodies are vegetated, they can often be characterised by relatively high backscatter due to either volumetric scattering from the canopy or the effect of double-bounce backscattering [20,38,39], when accounting for other factors such as the degree of canopy penetration, related to polarisation and wavelength and density of the vegetation canopy [20,38,40–43]. Schlaffer et al. [7] generated thematic maps of wetlands in Zambia using multi-temporal radar imagery, exploiting seasonally variable backscatter characteristics of different land cover types including open water, vegetated water and dry land. Despite the success of this approach - similar to many flood mapping systems discussed previously - there is a reliance on dry season imagery to provide a reference for change. As such, this approach is not suitable for mapping aquatic mosquito habitats in the dry season.

A successful attempt to extract both open and vegetated water bodies has been demonstrated by Plank et al. [44]. Their approach combines C-band Sentinel-1 and L-band PALSAR-2 imagery. While Sentinel-1 is used to map open water, PALSAR-2 is used to map vegetated water bodies, exploiting the greater canopy penetration offered by the longer wavelength L-band and multi-polarity imagery. However, the reliance on the PALSAR-2 data limits the application of this approach to operational mosquito vector-borne disease control because the imagery is not currently freely available. Until such time as widely available L-band imagery is available, it would be beneficial to explore the use of Sentinel-1 imagery to map both open and vegetated water bodies, despite its deficiencies in terms of canopy penetration. Tsyganskaya et al. [45] demonstrated a successful approach for detecting both open and vegetated water bodies in the Caprivi Strip, Namibia using Sentinel-1 imagery. This was achieved by generating metrics characterising temporal change in backscatter, enabling the delineation of temporary vegetated water bodies. Despite its success (overall accuracy: 85%), this approach is tailored towards the detection of water during flood events–its reliance on reference dry season imagery means that this approach too is not readily applicable to the detection of water bodies during the dry season that is of key importance for application to public health.

For a given radar wavelength, the high double-bounce backscatter signal from the vegetation-water surface will be lost where the canopy density reaches a critical level [40,46]. As such, vegetated water bodies can expect to be detected in a radar image up to a particular level of canopy closure. In the case of relatively low penetrating C-band Sentinel-1 imagery, it is expected that detectable vegetated water bodies will consist of relatively low canopy density [20,40,47]. This represents a significant drawback for water body mapping in many floodplain environments, e.g., dense canopy mangrove forests in tropical coastal floodplains that cannot be penetrated by C-band radar [48]. Although water bodies with dense vegetation canopies are likely to be omitted by general water body mapping programmes, in terms of mosquito control, water bodies with relatively sparse vegetation are characteristic of important malaria vector (*Anopheles gambiae sensu lato*) aquatic habitats in sub-Saharan Africa [35,37,49,50] (Figure 1). Furthermore, other vector species (e.g., *Anopheles coustani, An. squamosus, An. ziemanni*) are found even in more densely vegetated water bodies [51–54] reinforcing the need to include these in the classification.

Classification approaches using machine learning have proven to be successful in mapping open and vegetated water bodies using radar imagery [8,25,44,55]. Machine learning tends to be a supervised approach, requiring sufficient training data that are traditionally labour intensive to collect. However, some have developed methods for automatic extraction of water body training data through refinement of existing datasets such as the MODIS-derived static water mask and Pekel et al.'s [56] Landsat-derived water occurrence layer [25,55,57]. This approach has strong potential for use in developing an operational monitoring technique but so far, the use of pre-existing datasets for training machine learning classifiers has not been applied to vegetated water bodies.
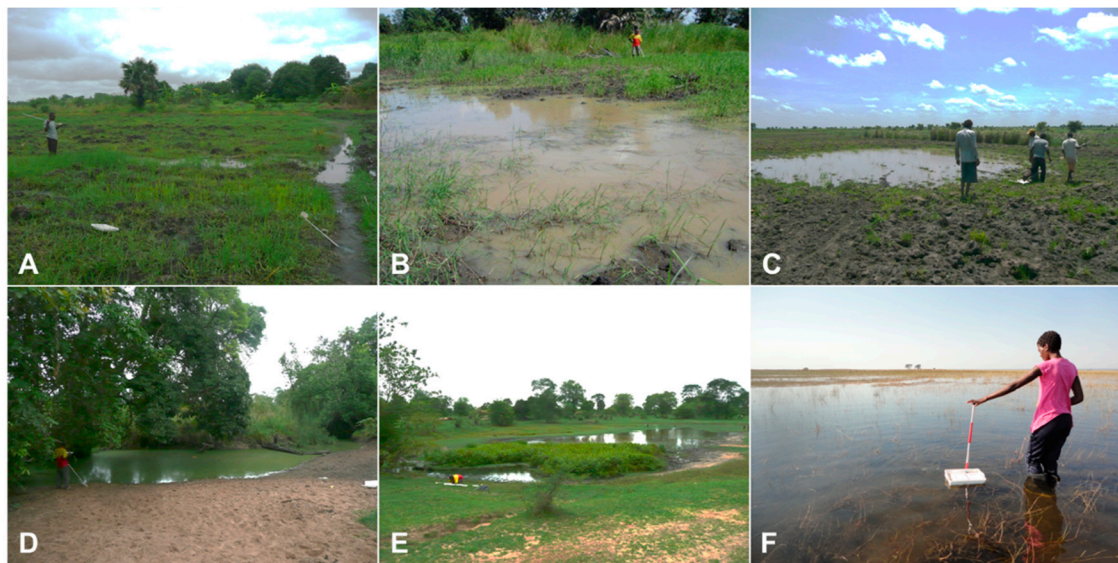
Figure 1. Example malarial mosquito aquatic habitats including (**A**) saturated areas driven by topographic convergence of subsurface moisture; (**B**) depressions within floodplains of active river channels with well-developed levees; (**C**) water bodies within relict palaeochannel systems; (**D**) pools located in perennial or seasonally active river channels; (**E**) spring-fed pools; and (**F**) river flood inundation zone. Images A-E taken within a malaria endemic area in Tanzania, from Hardy et al. [35] and F in Zambia. Each site contained malaria vector larvae.

One of the key challenges in mapping open water using radar imagery is the backscatter similarities with flat unvegetated or sparsely vegetated areas [29,58]. Such errors have been handled in flood mapping routines through the use of a frequently low backscattering mask, or Sand Exclusion Layer [29]. However, this approach relies on relatively coarse-scale products (e.g., Shuttle Radar Topography Mission (SRTM) Water Body Data (SWBD), the Water Indication Mask (WAM) [58], or the global water occurrence layer [56]) to delineate permanent water that would also be contained with the mask. These datasets are too coarse to account narrow canals or waterways in floodplain areas that often function as main transport routes for health practitioners. Equally, small water bodies can often act as important aquatic habitats for vector mosquitoes [49]. Additionally, some water mask products, e.g., [56] include large amounts of missing data due to persistent cloud cover. One approach that has not explored for reducing confusion between open water and smooth unvegetated surfaces is the use of a consensus approach to classification (e.g., see Dargie et al. [59]), applying a measure of certainty to refine the classification.

This study aims to develop a novel approach for mapping both open and vegetated water bodies using single epoch Sentinel-1 images. Automatically derived training data will be used to drive an object-oriented machine learning classification, which is run multiple times for the same image with the resulting classification certainty being used to refine the final classified map. The approach is developed for a malaria endemic region in sub-Saharan Africa with very high prevalence of the disease and with a number of other public health challenges relating to surface water extent–not least the challenge of providing health care to rural communities where access to health facilities is frequently compromised by significant flooding.

2. Materials and Methods

2.1. Study Site

The water body mapping approach was developed using the Barotse Floodplain of the Zambezi River in Western Province, Zambia as a test case (Figure 2). The floodplain is located on a deposit of Kalahari sand [60] covering an area of ~5500 km^2, though the inundated area is estimated to extend to

a maximum of 10,750 km^2 when the floodplains of tributaries such as the Luena River are included [61]. An outcrop of resistant basalt south of the floodplain constricts the Zambezi channel, promoting the upstream development of these extensive wetlands. The Upper Zambezi catchment extends into Angola and North-west Zambia; river discharge is distinctly seasonal, reflecting the increased rainfall to these areas between October and April from the shifting Inter-Tropical Convergence Zone (ITCZ). Following the rainy season peak in January-February, seasonal flood peaks arrive in April and recede through May to July, with flows immediately downstream at Victoria Falls ranging between 400 m^3 s^{-1} and 2300 m^3 s^{-1} [62]. Seasonal inundation to the edge of the escarpment occurs annually; however, a pronounced decadal periodicity in the magnitude of inundation is observed, superimposed on a gradually increasing trend [63].
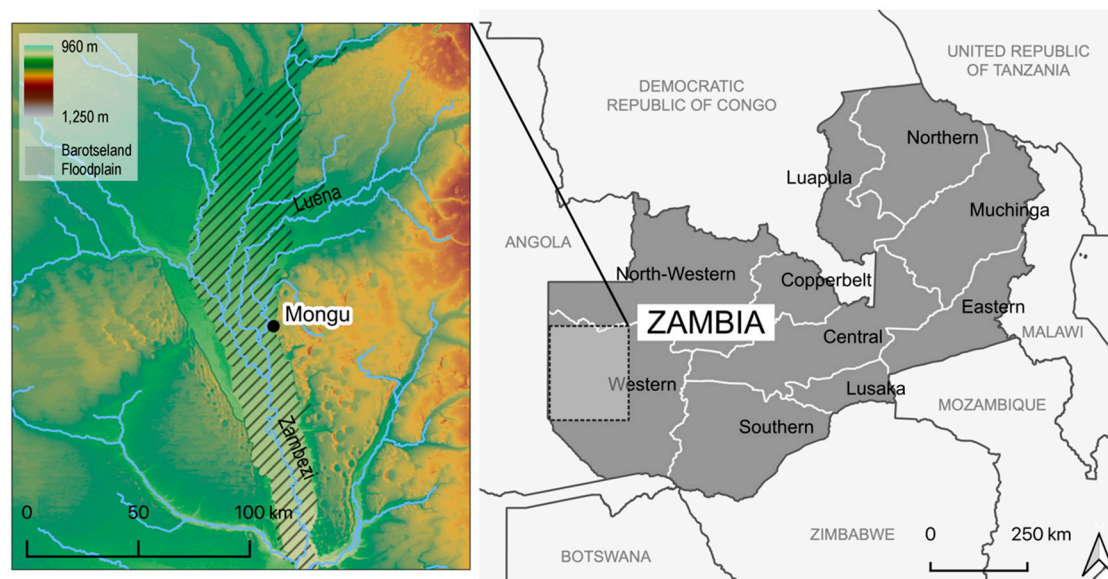
Figure 2. Location of the study site, Barotseland, Western Zambia. Digital elevation data from SRTM 3-second arc product. Location of the Zambezi and Luena Rivers are denoted, alongside the extent of the Barotse Floodplain that forms the Upper Zambezi catchment.

The floodplain itself extends up to 50 km in width following the course of the Zambezi for 230 km (Figure 2). As with all such large floodplains, complex hydrological regimes are observed as flows interact with topographically complex surfaces producing an extensive collection of geomorphological features and processes [64]. The Zambezi itself exhibits an anabranching planform, as seen with many such large rivers [65], with both active and abandoned accessory channels visible throughout the floodplain. The Kalahari sands are interspersed with silt and mud deposited by flood events [66] through a complex depositional web [67].

Yet water storage on the floodplain is spatially and temporally variable and is not supplied exclusively from overbank flooding [68]. Tributaries from the surrounding landscape enter the floodplain and run along the fringes of the escarpment on fine muddy deposits converted into rice paddies. Several of these tributaries are supplied by small shallow waterlogged depressions that form on peat deposits elevated above the Barotse floodplain, known as 'dambos' [69]. Connectivity of the floodplain to these surface flows and also subsurface flows through the floodplain plays an important role in surface water dynamics [70]. As such, a complex mosaic of dry season water bodies acting as potential mosquito refugia is observed. Many such water bodies, including the extensive Luena flats, where the anabranching tributary meets the left bank of the main channel floodplain, are covered with riparian vegetation.

The floodplain has high economic value: it is productive for rearing cattle and crop cultivation and supports a human population of at least 225,000 [66]. Malaria is endemic throughout Zambia,

with transmission in the Western Province occurring year-round with a distinct seasonal maximum after the flood peak [71]. Vegetation within the floodplain is typically comprised of seasonally flooded grasses locally termed 'mulapo' or seasonally cultivated grasses known as 'sitapa' [72]. During the wet season, matted floating grasses and other wetland vegetation types are abundant in inundated areas including papyrus, rushes, reeds, water hyacinth, water lettuce, kariba weed and the water fern [73] (examples shown in Figure 3). Forested areas are not extensive, with wooded areas (typically isolated stands of mango trees) being mainly limited to floodplain islands known as 'mazulu' [72].
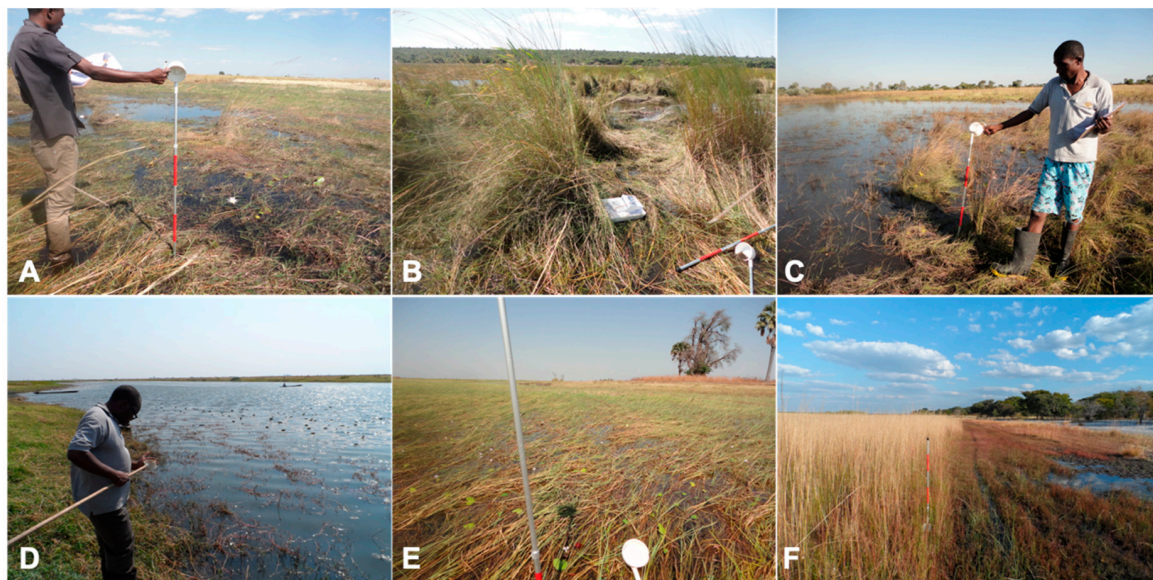

Figure 3. Example water bodies found within the study site. (**A**) mixture of matted grass, reeds, rushes and water hyacinth; (**B**) rushes, papyrus and reeds, (**C**) partially open water body; (**D**) open water; (**E**) matted grasses and water hyacinth; (**F**) edge of permanently inundated 'dambo' feature.

2.2. Datasets and Classification Procedure

Dual polarised (VV and VH) Sentinel-1 Ground Range Detected (GRD) scenes were acquired from Google Earth Engine (https://earthengine.google.com/). All scenes were pre-processed using the following steps: 1. Thermal noise removal, 2. Radiometric calibration and 3. Terrain correction. Within the Google Earth Engine Javascript API the VV/VH ratio was calculated and added to each image before exporting. A total of 59 scenes was downloaded covering the period 2016–2018. A summary of the datasets used for the classification and validation is given in Table 1. An overview of the classification approach is given in Figure 4.

All subsequent processing was scripted in Python using the open source libraries RSGISLib [75], GDAL and scikit-learn [76]. For each VV, VH, VV/VH stacked image, a 3 × 3 adaptive Lee filter was applied that reduces speckle noise using coefficient of variation applied to a moving window while retaining textural information by classifying pixels as homogenous (pixel is replaced by an average calculated within the window), heterogenous (a weighted window average is used) or as a point target (pixel value retained) [77,78]. The filtered images were then segmented using the Shepherd segmentation [79] and various statistics were added to the raster attribute table for each band, VV, VH and the VV/VH ratio (minimum, maximum, mean and standard deviation). Visual analysis showed that the segmentation was able to simplify the radar images into a manageable number of image objects, while retaining much of the backscatter information regarding different land cover types, including water bodies, although some very narrow waterways (typically < 2 m) were lost.
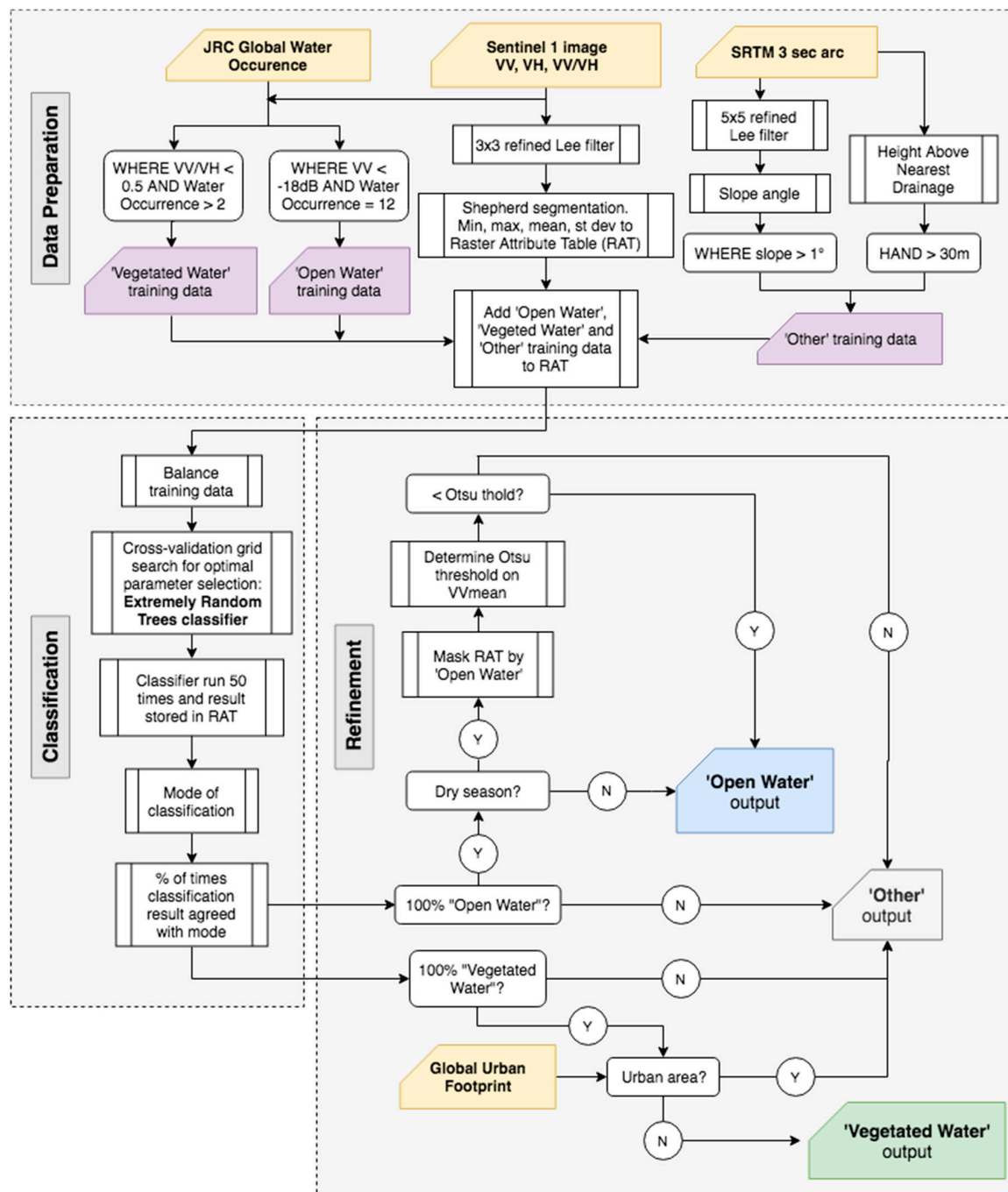
Figure 4. Flow diagram summarising the classification routine for open and surface water bodies using Sentinel 1 radar imagery.

Training data were drawn from the segmented image and stored in the raster attribute table using the image statistics as well as information from ancillary datasets including the global water occurrence layer [56] and SRTM terrain derivatives. This process was carefully guided by field observations made during fieldwork campaigns in September 2018, May 2018 and March 2017, including geo-tagged photographs and broad landscape characterization (presence of water, nature of vegetation etc.) and in some instances, point measures of water depth and vegetation height.

Table 1. Summary of the datasets used for developing the water body classification and for validating the resulting water body products.

Dataset	Dates	Area (UL, LR Coordinate WGS84)	Description
		Classification datasets	
Sentinel-1	October 2016 to September 2018 (n = 59)	22, −14; 24, −17	Level-1 Ground Range Detected (GRD), VV and VH polarisation (10 m)
JRC Global Water Occurrence	Static	22, −14; 24, −17	Water body frequency derived from optical Landsat imagery (30 m) [56]
SRTM 3-second arc	Static	22, −14; 24, −17	Digital surface model (90 m)
Global Urban Footprint	Static	22, −14; 24, −17	Urban mask (12 m) derived from X-band TerraSAR-X imagery [74]
		Validation datasets	
Pleiades	July and March 2017 (n = 2)	22.7, −15.6; 23.3, −15.1	Visible-Near Infrared optical imagery (2 m)
Sentinel-2	October 2016 to June 2018 (n = 11)	22.8, −15.5; 23.6, −14.5	Visible-Near Infrared optical imagery (10 m)

Image objects were assigned as training data for areas of open water where: a) the VV backscatter was less than −18 dB and b) the global water occurrence layer [56] identified the object as being permanently wet. Training data for vegetated water bodies were selected where the VV/VH ratio was less than 0.5 and global water occurrence layer identified the object as being wet for at least two months of the year. The optically-derived global water occurrence layer is designed to map open water only, but we found that some permanently wetted vegetated areas were also identified within this dataset, where water occurrence was recorded for at least two months of the year. As such we were able to use this dataset, alongside the VV/VH ratio to delineate sufficient vegetated water training objects.

Training data for non-wetted objects were extracted where the slope angle was greater than 1 degree. Slope angle was generated using the SRTM 3 arc second product with a 5×5 smoothing filter applied to supress fine scale variations. This mask was refined by selecting areas where the terrain derivative Height Above Nearest Drainage (HAND) [80] was greater than 30 m to ensure that areas at the fringes of floodplains (often marked by a defined change in elevation or escarpment) were not included. HAND has been used by other authors [21,44,81] to help eliminate false positive water body detections located above the drainage line.

The open water, vegetated water and dry training data were used to train an Extra Trees classifier (otherwise known as 'Extremely Random Forests'). This algorithm splits decision trees at random over the dataset range [76] and has been shown to outperform Random Forests as well as being computational more efficient [82]. Machine learning classifications can be affected by imbalanced training data, with under-predictions reported for classes with relatively few samples, typically a less abundant class [83]. To overcome this, before the classifier was initiated the training samples were balanced by selecting objects at random, matching the class with the least number of training samples. Similarly, parameter selection for input into the machine learning algorithm can affect the resulting classification [83]. In this study, a cross-validation grid search is used to determine optimal parameters for use in the Extra Trees classifier, testing the relative performance of variables such as number of trees used and the maximum number of features to consider when looking for the best split.

For each image stack, the balanced Extra Trees classifier was run 50 times. Tests demonstrated that running the classifier 50 times provided a balance between computational demand and a sufficient number of instances to determine classification certainty. The classification result for run was stored in the raster attribute table and the mode was calculated providing the most commonly assigned class per object. The percentage of occasions that the mode was predicted was also calculated, generating a metric of class certainty. For an object to be labelled as open water or vegetated water, there needed to

be 100% agreement with the mode. Essentially, this means that for objects where there was uncertainty over whether it was a water body (open or vegetated) or not, they would be relabelled as dry. In doing so, we found that the confusion between open water and other low backscattering objects was largely removed. Similarly, we found that the classification certainty refinement helped to remove much of the confusion between vegetated water bodies and other high backscattering objects such as rough water and urban settlements.

Outside the flood inundation period (July-January) a post-classification correction is applied for falsely identified open water areas (that were not sufficiently removed by the classification certainty refinement), a symptom of smooth low-backscatter fluvial deposits. For each image within this period, a mask is applied based on the predicted area of open water (that includes false positives). It was found that backscatter values between open water and low-backscattering areas were relatively distinct in VV polarisation (as illustrated in Figure 5). This distinction was exploited through Otsu thresholding, relabelling areas above the threshold as dry. Despite providing an improved classification, this refinement did result in some misclassification of open water as dry, particularly at the fringes of permanent water bodies such as the trunk Zambezi river. To resolve these issues, a permanent water mask was derived where open water was predicted in 100% of the scenes from the original, pre-refinement, classification results. Although the use of classification certainty was shown to eliminate much of the confusion between vegetated water bodies and urban settlements, following the example of Plank et al. [44] the open source dataset Global Urban Footprint (GUF), derived from TanDEM-x imagery [74], was used to mask out urban areas.
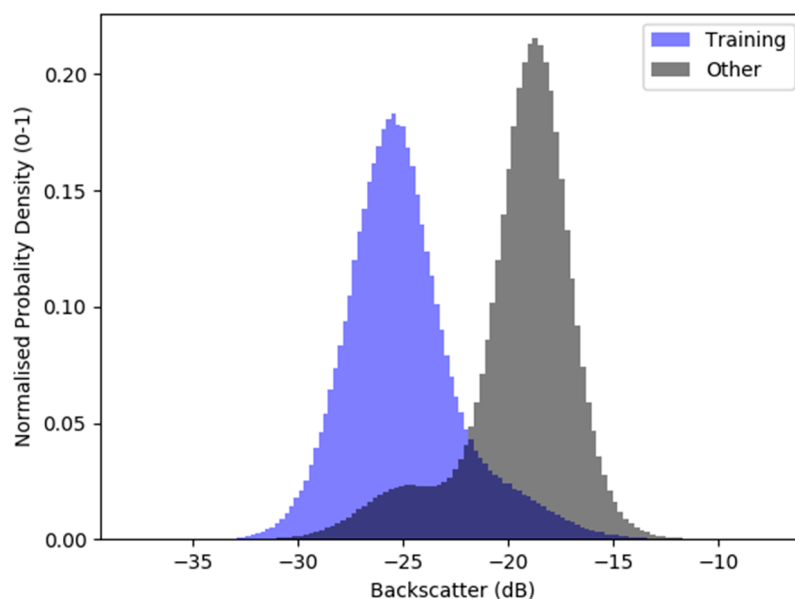
Figure 5. Example histogram for a relatively dry period showing the difference between open water pixels ('training') and lumped open water and low-backscatter areas ('other').

Resulting thematic classifications were validated using stratified random points, between 800–1000 per scene (total = 9880, open water = 3215; vegetated water = 3593; other = 3072), using scenes (n = 11) across the temporal range covering different hydrological conditions. Validations were made using 2 m Pleiades optical imagery acquired during the wet season (March 2017) and the beginning of the dry season (July 2017). Additional validations were made using 10 m Sentinel-2 imagery (February, April, May, June, August, September, December). Despite being used for mapping wetlands in the past [84–86], optical imagery is not an ideal tool for providing observations of vegetated water bodies particularly where vegetation canopies are completely closed, obscuring the water surface below. In this study, interpretation of optical validation imagery was supported by field observations made during fieldwork campaigns (September 2018, May 2018 and March 2017) that included water body

characterization and point measures of water depth and vegetation height (although not sufficient in number for use as an independent ground-truth dataset). In the case of fully vegetated water bodies, these features tend to be spectrally distinct with very high reflectance in the near infrared due to high leaf cell structure as well as planophile leaf orientation associated with matted grasses and floating vegetation. Combined with the context in which the feature occurs—for instance, positioned within a known marsh area—this enables the use of this imagery for identifying vegetated water bodies even where vegetation canopies are closed.

Standard accuracy assessment metrics were calculated, including user's and producer's % agreement scores for each class and corresponding overall accuracy (%) and Kappa scores. It is becoming increasingly accepted that Kappa has significant limitations rendering them potentially misleading [87]. Therefore, we have also calculated quantity and allocation disagreement, a proposed replacement for Kappa that describes how well the mapped products represent the area covered by each class (quantity disagreement) as well as the spatial agreement between mapped areas versus the validation data (allocation disagreement). Full description and discussion of these metrics can be found in Pontius and Millones [87] as well descriptions and applications in other reported studies [88–93].

3. Results

The approach for mapping both open and vegetated water bodies was applied to 59 Sentinel-1 radar images for Barotseland over the period 2016-18. The accuracy assessment revealed strong agreement with optical validation data, with a mean overall accuracy of 92% (kappa = 0.9) (Table 2). The accuracy ranged from 84% in November 2016 to 97% in April 2018. This study represents the first attempt to map vegetated water bodies using Sentinel-1 alone.

Table 2. Summary of the accuracy assessment results.

			User's Accuracy %			Producer's Accuracy %				
		Num	**Water**	**Other**	**VegWater**	**Water**	**Other**	**VegWater**	**Acc %**	**Kappa**
Jan	12/01/2018	903	98	82	99	99	96	85	92.7	0.89
Feb	05/02/2018	902	97	73	100	99	96	79	89.8	0.85
Mar	30/03/2017	903	100	74	100	100	100	80	91.4	0.87
Apr	18/04/2018	903	99	92	99	99	99	93	96.7	0.95
May	12/05/2018	903	95	69	99	99	92	77	87.8	0.82
Jun	17/06/2018	903	95	91	97	98	93	93	94.7	0.92
Jul	16/07/2017	903	99	89	97	99	96	91	95.2	0.93
Aug	09/08/2017	903	99	89	93	99	92	90	93.9	0.91
Sep	14/09/2017	903	98	88	93	99	91	90	93.1	0.9
Oct	13/10/2016	903	93	93	87	100	83	93	91.0	0.87
Nov	06/11/2016	851	95	89	69	99	71	88	84.4	0.77
Dec	Cloudy	-	-	-	-	-	-	-	-	-
	Mean		97.1	84.6	94.1	99.1	91.6	87.2	91.9	0.9
	St Dev		2.1	8.5	9.1	0.5	8.2	6.0	3.6	0.1
	Total points	9880								

Both the user's and producer's results for vegetated water bodies indicated strong agreement with the validation data with mean scores of 94% (user's) and 87% (producer's) (Table 2). Nine of the eleven assessed epoch demonstrated a broad agreement between Kappa/overall accuracy (%) and the overall agreement score (C: Table 3) that accounts for class quantity and allocation agreement, but for two epochs (6/11/2016 and 13/10/2016), Kappa and overall accuracy (%) underestimated the apparent accuracy of the classification. This was a direct results of Kappa/overall accuracy (%) not accounting for the relative area covered by each class. This is particularly important for the present study whereby dry season images are likely to be dominated (> 90% of total area) by dry land, and therefore the errors associated with this class should carry greater influence on the overall assessment of accuracy. The quantity and allocation disagreement metrics demonstrate a good overall agreement with the validation data (mean C = 0.86, s.d. = 0.07). The allocation disagreement scores demonstrate a strong ability for the developed system to accurately predict the spatial location of water (both open

and vegetated) and dry land. The quantity disagreement scores demonstrate that the main source of error lies with the ability to accurately predict the area of each class, particularly during wetter periods (i.e., Feb 2018, May 2018 and Mar 2017). This is due to false positives, associated with inundated vegetation being false classified as dry land, and should be a target for future development of the classification routine.

Table 3. Summary of quantity and allocation disagreement results, alongside overall accuracy and Kappa scores for comparison. Q: Quantity disagreement; A: Allocation disagreement; C: Overall agreement; D: Overall disagreement.

		Acc %	Kappa	Q	A	C	D
Jan	12/01/2018	92.7	0.89	0.162	0.003	0.83	0.165
Feb	05/02/2018	89.8	0.85	0.241	0.001	0.75	0.245
Mar	30/03/2017	91.4	0.87	0.201	0.000	0.80	0.201
Apr	18/04/2018	96.7	0.95	0.058	0.003	0.94	0.062
May	12/05/2018	87.8	0.82	0.253	0.008	0.74	0.261
Jun	17/06/2018	94.7	0.92	0.084	0.002	0.91	0.085
Jul	16/07/2017	95.2	0.93	0.108	0.001	0.89	0.108
Aug	09/08/2017	93.9	0.91	0.104	0.001	0.89	0.105
Sep	14/09/2017	93.1	0.9	0.115	0.001	0.88	0.116
Oct	13/10/2016	91.0	0.87	0.044	0.004	0.95	0.048
Nov	06/11/2016	84.4	0.77	0.097	0.012	0.89	0.108
Dec	Cloudy	-	-	-	-	-	-
Mean		91.9	0.88	0.133	0.003	0.86	0.137
St Dev		3.6	0.05	0.071	0.004	0.07	0.072

Figure 6 demonstrates the relative importance of variables following the extra trees classification for each image. The standard deviation metrics tend not to be important, consistently contributing less than 10% to the classifier. This is due to the relatively small size of the objects—larger objects would make standard deviation more meaningful, but this would be a sacrifice of the detail in the subsequent classification product, with smaller water bodies being missed entirely. Average VV backscatter was the most important with an average contribution of more than 18%. This represents the ability of VV imagery to differentiate between high backscatter vegetated water bodies and low backscatter open water within an extra trees classifier. Equally, average VV backscatter offers the ability to differentiate between smooth open water and relatively rough non-water objects with higher backscatter. Variations over time existed with all variables (between 16-39% contribution for important variables), demonstrating the need to have adaptive strategies when delineating water, where hard thresholds set independently from time would likely lead to temporally dependent errors. Average VV backscatter becomes less important during March-April. These are the wettest months and include the greatest coverage of wetted vegetation. In these months, accurate mapping of both open and vegetated water bodies require contribution from other variables (bar standard deviation metrics), a result of the increasing complexity of the landscape.

The consensus approach to refining the classification output–classification certainty based on 50 iterations of the classifier—had variable influence over the image archive. On average, 7.2% of classified objects were changed due to classification uncertainty. The influence increased during the dry season with up to 20% of classified objects being changed due to uncertainty associated with dry flat sediments that would otherwise be classified incorrectly as open water. The post-classification refinement (classification certainty and dry season Otsu thresholding of water masks) was found to improve classification accuracies up to 7% and an improvement on errors of commission associated with open water areas of up to 33% (illustrated in Figure 7). Due to the lack of extensive urbanisation in the region, the addition of the Global Urban footprint had a negligible positive influence on the results.
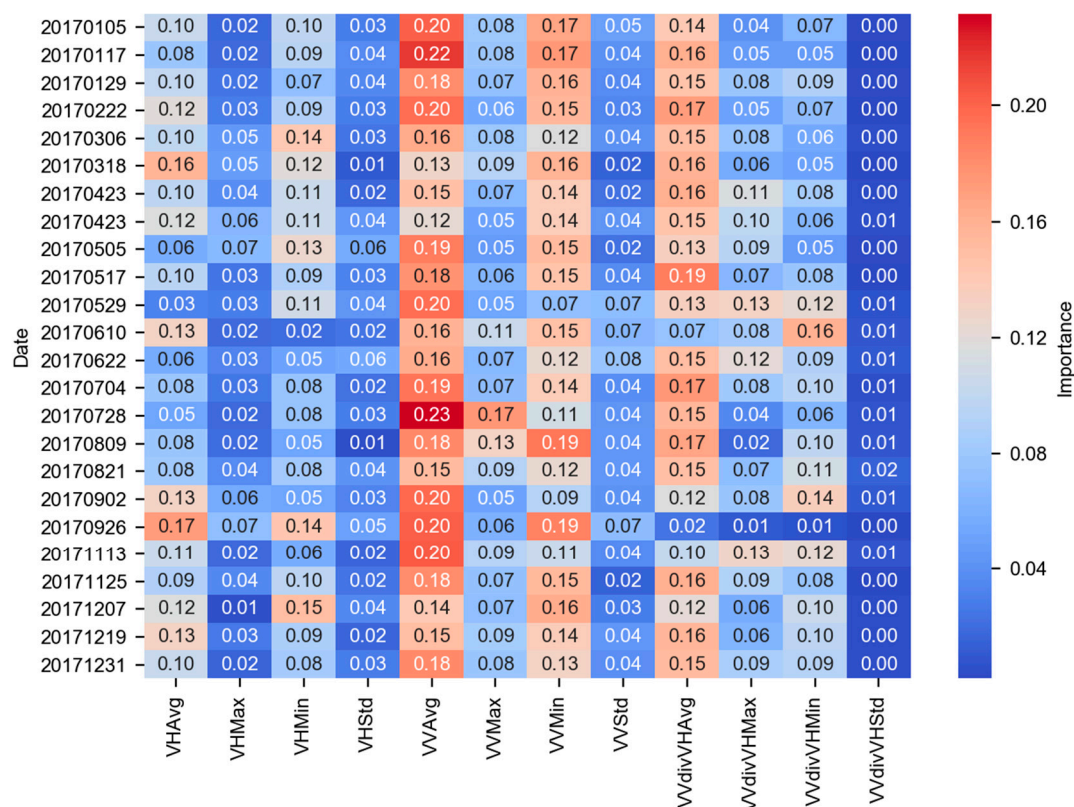
Figure 6. Variable importance for the Extra Tree classifier over 2017. VV = VV polarisation (dB); VH = VH polarisation (dB); VVdivVH = VV/VH ratio. Suffixes represent mean, maximum, minimum and standard deviation statistics respectively.
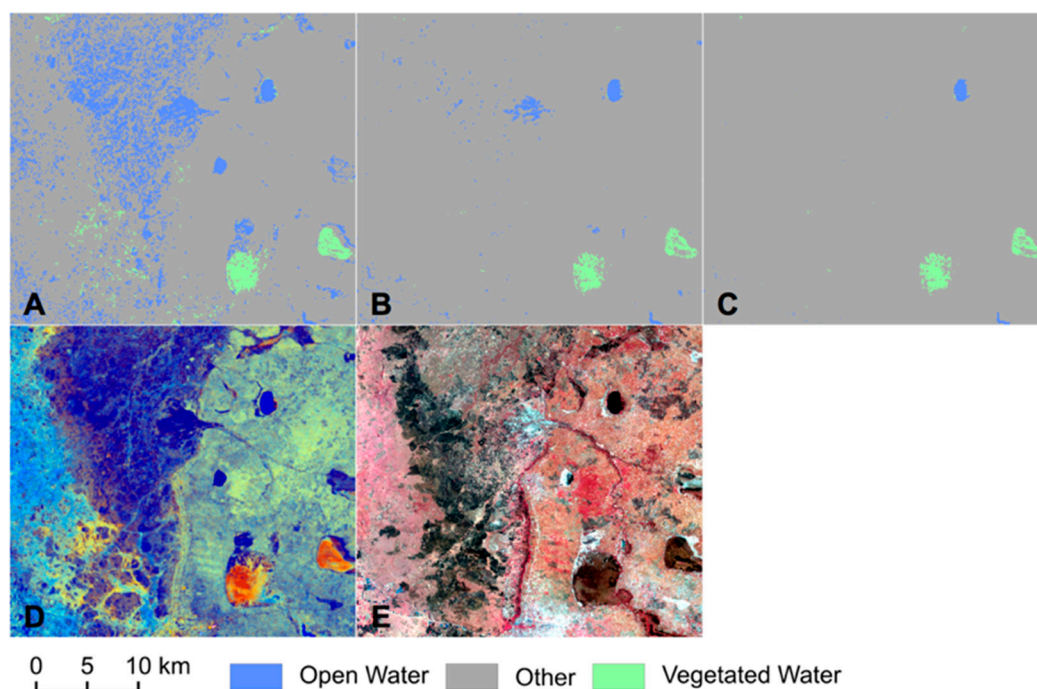
Figure 7. Progressive classification improvement following implementation of the classification certainty and dry season refinement. Example image from 2 Sept 2017 (dry season). (**A**) Pre-refinement classified image, (**B**) post-classification certainty refinement, (**C**) post-Otsu thresholding of open water mask, (**D**) input VV, VH, VV/VH Sentinel-1 image and (**E**) reference NIR, R, G Sentinel-2 optical image.

As outlined in the introduction, the C-band radar signal from vegetated water bodies is expected to saturate at a particular level of vegetation canopy density [40]. As a consequence, the classification scheme outlined in this paper is only applicable for vegetated water bodies with vegetation cover below a particular level. Figure 8 shows the relationship between VV and VH dB backscatter for 10,000 points over vegetated water bodies for a peak flood scene (11 May 2018). In this instance, almost 70% of points lying on vegetated water bodies were not identified by the classification routine, due to the saturation of the backscatter signal over relatively dense canopies.
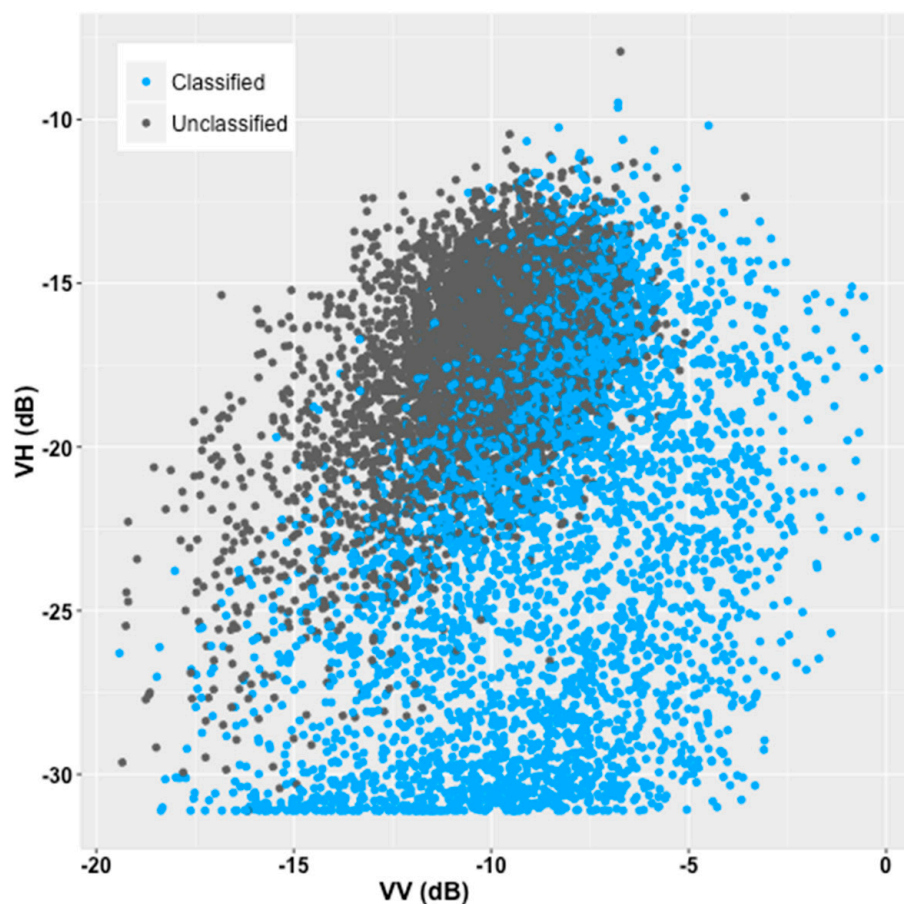
Figure 8. Scatterplot of VV and VH polarised dB backscatter values over vegetated water bodies for 11 May 2018.

Field observations show that the types of areas where vegetated water bodies were successfully identified are characterised by relatively low (generally < 20 cm tall) grasses with a degree of spacing between grass stands (~20 cm apart) (Figure 9). The identification of vegetated water bodies with taller grass and vegetation stands (generally > 50 cm) was not possible due to saturation of the radar backscatter signal. In many instances, the vegetation canopy completely obscured the water's surface below to the extent that water was not observable even in field photographs (Figures 10 and 11). Conversely, during drier periods false positives associated with vegetated water bodies occur where a mixture of stubble vegetation and flat underlying soil creates a high backscatter double-bounce effect.
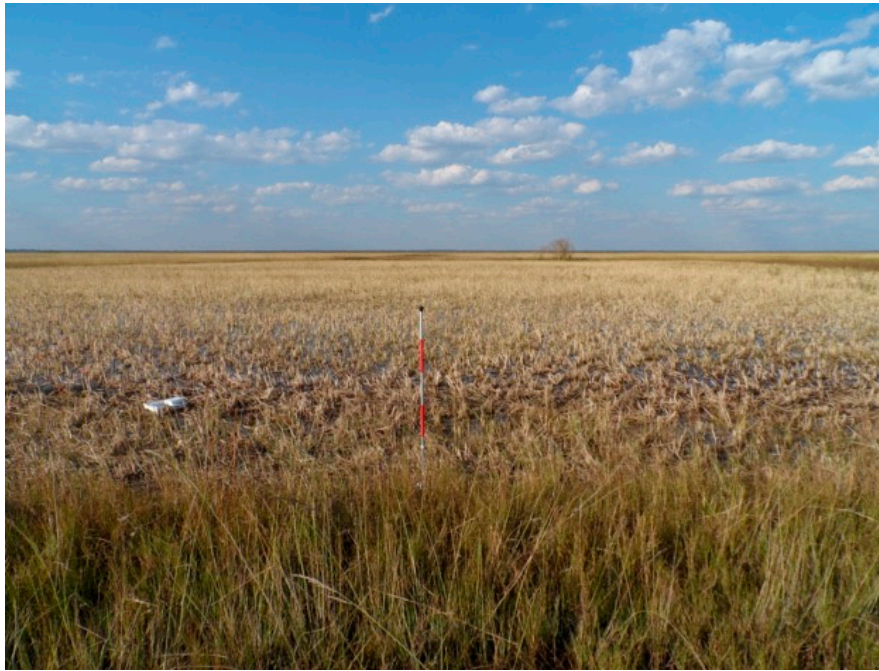
Figure 9. Example vegetated water body (raised peat bog or 'dambo') successfully classified, characterised by relatively short grasses. Image from June 2018 long: 23.32, lat: −15.14. Ranging pole with 5 × 20 cm sections from base. Water depth = 5 cm.


Figure 10. Example vegetated water body that was missed by the classification, characterised by dense grasses. Image from June 2018 long: 23.4, lat: −14.8. Ranging pole with 5 × 20 cm sections from base. Water depth = 40 cm.

Figure 11. Example vegetated water body that was missed by the classification, characterised by matted or floating grasses. Image from June 2018 long: 23.16, lat: −15.36. Ranging pole with 5×20 cm sections from base. Water depth = 10 cm.

Mapped water body extent per Sentinel-1 epoch demonstrated a wide seasonal range (Figure 12) with a total extent typically less than 1000 km^2 during the dry season (July-January) and a maximum extent between 7600 km^2 and 8500 km^2 at the peak of the wet season. The peak of the wet season was mid-March during 2017, shifting to early-April for 2018. The 2018 maximum water extent was 11% greater than the previous year. For the period studied, 70% of the total water extent mapped was attributed to vegetated water with the peak in the maximum extent of inundated vegetation being followed by a peak in open water extent approximately two weeks later.
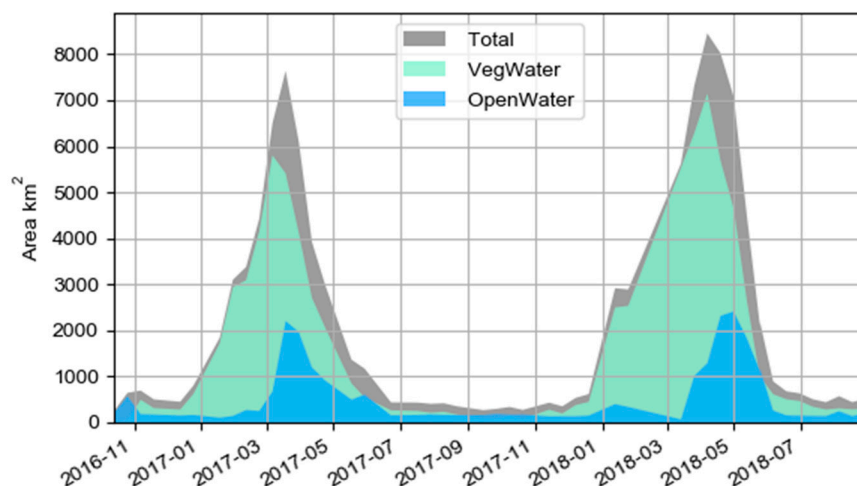

Figure 12. Water body extent derived from classified Sentinel-1 images for Barotseland over the study period.

Mapped water bodies over the study's time period indicate that the hydrological landscape of Barotseland is highly dynamic with maximum water extent expanding across much of the area and permanent water largely restricted to the trunk Zambezi River and raised peat bogs on the escarpment

known as 'dambos' (Figure 13; example dambo shown in Figure 9). In the northeast of the study area, water bodies associated with the Luena tributary sub-catchment tend to be vegetated. This is in contrast to the floodplain of the Zambezi that during the wet season typically consists a mixture of vegetated and open water bodies.
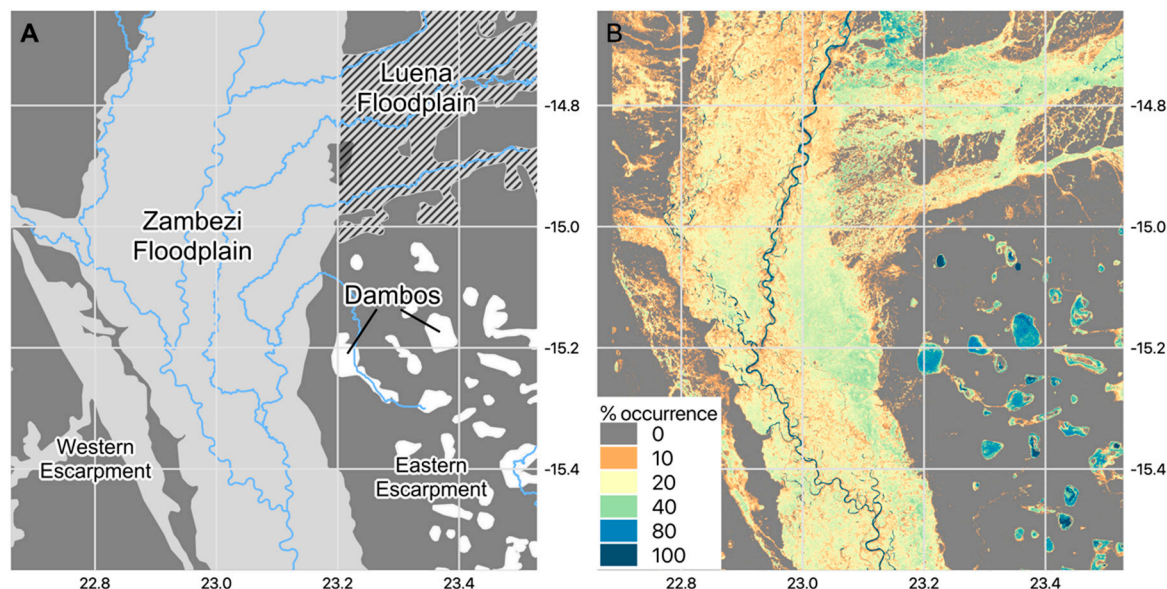
Figure 13. (**A**) Reference maps showing location of key hydrological units. (**B**) Percentage occurrence of mapped water over the studied period, 2016-2018.

4. Discussion

4.1. Water Body Detection

The results indicate that both open and water bodies vegetated with low (generally < 20 cm) grasses can be mapped with a high degree of accuracy in Barotseland, Western Zambia using Sentinel-1 radar imagery. For the period studied, the majority of total water extent mapped was vegetated water (Figure 12) highlighting the importance of mapping both open and vegetated water bodies for surface water mapping, particularly in the wet season, whereas many existing approaches focus only on mapping non-vegetated water bodies [1,2,5,22,24,27,28,30].

In most water body mapping exercises, one would expect the dominant land cover type, in terms of area, to be represented by dry land. In these instances, it is especially important that the relative area covered by each class is considered when assessing the accuracy of the classification, using measures such as the quantity and allocation disagreement [87]. This study demonstrated that % overall accuracy and Kappa statistic underestimated the apparent accuracy of the classification due to errors associated with the dominant dry and ('other') class.

Major sources of error in the classification system were identified as false positives: mapping dry land over areas of inundated vegetation. In part, this was due to the saturation of the radar backscatter signal over dense canopies. This follows findings from other studies [40,46] and highlights the importance of defining the types of vegetation over which water bodies are presumed to be mappable. In this instance, using C-band Sentinel-1, vegetation types were characterised as grasses growing in tussocks or clumps, enabling a double bounce signal from water's surface and vegetation stems to be observed. In the Barotseland study area this vegetation type was estimated to account for approximately three-quarters of areas prone to flooding, with extensive floating vegetation mats and thick, tall grasses (>50 cm) tending to obscure the water's surface below during the wet season. In this regard, it is likely that more vegetated water bodies will be mappable during the early stages of the wet season, during which water is extensive and vegetation canopies are less dense.

The inclusion of information from longer wavelength (e.g., L or P-band) systems would enable greater signal over water bodies with denser vegetation canopies, e.g., [44]. Similarly, there is evidence to suggest that the signal over vegetated water bodies may be enhanced through the decomposition of fully polarimetric radar data [9,44,94]. However, a key focus of this study was to exploit the freely available and high frequency Sentinel-1 imagery. The free availability of high temporal frequency (sub-weekly) L or P-band systems would represent a significant development for water body mapping. This type of data, as well as decomposition metrics can be easily incorporated into the object-based machine learning classification structure proposed in this paper.

Being dominated by a large seasonally inundated river floodplain, the Barotseland landscape is complex and varies significantly throughout the year. In this study, the machine learning classifier included a number of independent variables largely based on the radar image statistics. The individual importance of these variables varied overtime, reflecting the dynamic nature of the landscape but also highlighting the importance of using adaptive strategies for classifying flood inundation. In addition, the consensus approach used in this classification scheme (50 iterations) had a significant effect on the final classification (affecting the final class for up to 20% of image objects). This was valuable for limiting the degree of confusion between open water and other low backscattering objects such as flat dry sand, that are a source of considerable error in radar-based flood water mapping [29]. Further errors in this regard were handled efficiently by using Otsu thresholding to refine areas classified as open water. The combination of the consensus approach and refinement of individual classes (in this instance using Otsu-derived thresholding) helped to maximise the degree of separability between similar land cover types, providing a potential solution for other land cover mapping exercises beyond water body mapping.

This study did not carry out a formal evaluation of the segmentation approach. The accuracy assessment identified that a number of narrow channels (< 30 m) were lost due to a combination of the Lee image filter and the segmentation process. The absence of these channels in the resulting water body product has consequences for use in defining waterway transportation routes and subsequent health access. In effect, these narrow channels may represent critical waterways for accessing wider navigable channels such as the trunk Zambezi River. Future work should focus on the development of open-source segmentation techniques that preserve linear features.

4.2. Hydrological Findings and Implications for Public Health

The Barotseland landscape is dominated by herbaceous wetlands, as opposed to open water, with emergent or floating vegetation. This is significant in terms of the ability of a water body to support juvenile mosquito development: the presence of vegetation within a water body can provide *Anopheles* mosquito larvae with a refuge from predators [50,95]. The dynamic nature of the hydrological landscape has important implications for vector mosquito ecology and subsequent disease transmission.

Specifically, water bodies in the north-eastern part of the study area (Luena floodplain), in some locations are shown to persist for much of the hydrological year, particularly near Sitoya (Figure 15). These persistent water bodies potentially provide a refuge for mosquito species throughout the dry season, where aquatic habitat availability is limited [23]. Furthermore, the Luena River is characterised as a highly diffuse anabranching river system due to deposition and subsequent channel avulsion. Together with persistent presence of vegetation, this is indicative of shallow, slow-moving or still water bodies, characteristics that are consistently associated with productive mosquito habitats [49,96–98]. Malaria case data indicate that this region sustained relatively high transmission rates in both the wet and dry seasons of 2017, particularly at Sitoya, Ushaa and Nangili that lie within, or next to this drainage system (Figure 15).

The dambo features located on the escarpment in the east of the study area remain inundated for much of the year (Figures 13–15). These water bodies are also characterised as shallow, sunlit and typically vegetated (Figures 3F and 9). As such, the dambos, coupled with the shallow, slow-moving

channels that connect them, act as important sources of malaria mosquito larvae causing widespread transmission in the 2017 wet season and sustaining relatively high numbers of cases in areas like Ikabako, Miulwe, Kalundwana and Luandui. Given their importance in driving transmission throughout the year, these hydrological features, as well as permanently flooded regions in the Luena system, could be an important target for larval source management interventions.
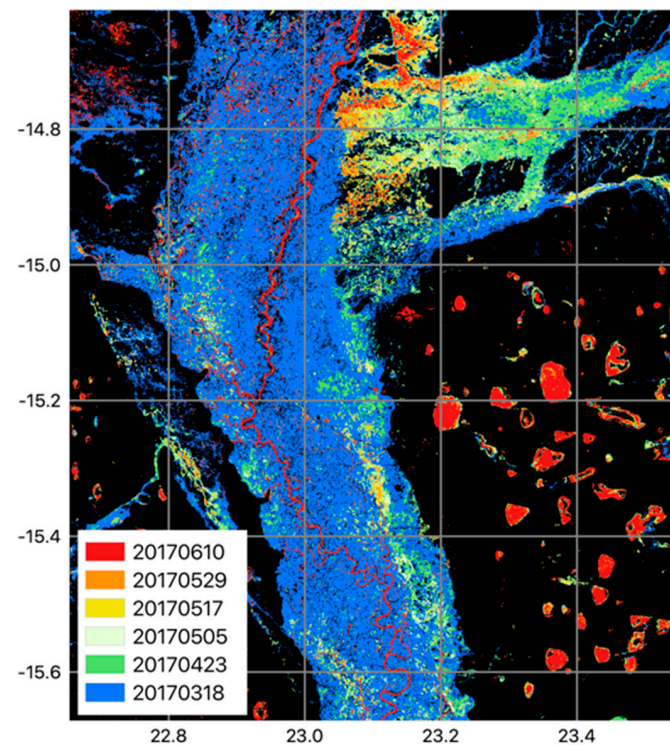
Figure 14. Sequential images representing the 2017 draw-down of water following the peak of the wet season (mid-March) to the start of the dry season (June). Colour represents the latest date on which water was present.
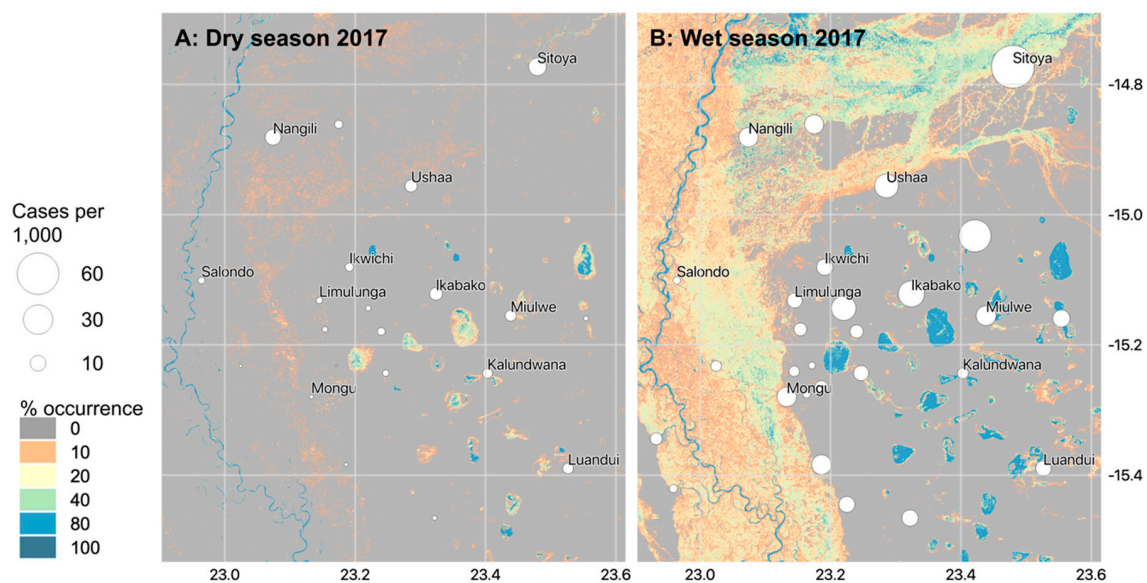
Figure 15. Malaria case data (total number of cases expressed per 1000 population: https://www.path. org/malaria/) reported by health facilities for the 2017 (**A**) Dry season (Sep-Nov) and (**B**) Wet season (Mar-May), alongside corresponding water occurrence for that period.

Towards the south-eastern flank of the Zambezi Floodplain, half a degree south of Mongu, contributions of water from the escarpment allow water to persist for longer than in areas close to the main Zambezi channel. This is illustrated by sequentially mapped inundated areas during the draw-down period in Figure 14 showing that almost three months after the peak of the 2017 wet season, water bodies persist in these areas. Again, water bodies in this area share characteristics with productive mosquito aquatic habitats (shallow, sunlit, still or slow moving) but despite their persistence, malaria transmission in 2017 is generally limited to the wet season (Figure 15). Field observations showed that these water bodies were relatively clear and low in temperature during the dry season, which has been shown in other studies to be less conducive to supporting oviposition and larval development [49,99,100] and may therefore not represent important vector aquatic habitats in this instance.

When developing tools for use in public health it is important to consider the degree to which the approach can be adopted by the people that need it most–in this instance, public health managers in malaria endemic regions. The classification system presented in this study makes use of a freely available, fully pre-processed radar imagery and globally available ancillary datasets. Furthermore, all processing was conducted within an open-source software environment, increasing the likelihood of this automated classification system to be used operationally. However, to fully ensure the use of this classification system and subsequent mapped products, future work should centre on incorporating these methods within a cloud-based service, such as Google Earth Engine, to avoid dependency on physical data storage and processing resources. It is also important not to over-interpret water body maps alone for delineating malaria risk and defining subsequent disease control strategies. Rather, careful consideration needs to made of the complex relationship between habitat location and persistence, alongside adult mosquito dispersal mechanisms, human population densities and socio-economic factors.

5. Conclusions

This paper presents an automated, tractable solution for mapping both open and vegetated water bodies in a large floodplain environment with a high degree of accuracy using Sentinel-1 radar imagery and openly available ancillary datasets. The use of open source data, combined with open source software provides an approach with considerable operation potential for use within public health programmes where resources may not permit the acquisition of commercial imagery.

Subsequent water body products demonstrate that the Barotseland area is highly dynamic, including large extents of inundated vegetation throughout the year, but particularly during the wet season. This dynamism highlights the need for developing reliable routines for mapping both open and vegetated water bodies to inform public health as well as other sectors such water resource management, riverine transportation and flood risk management.

Despite providing freely available, timely information regarding water body locations, Sentinel-1 C-band radar is limited in its ability to provide a signal from vegetated water bodies unless grasses are < 20 cm in height with sufficient gaps between grass tussocks or clumps. Future provision of freely available L or P-band imagery will help greatly improve classification of vegetated water bodies, particularly for regions with dense forest canopies [20]. The classification structure outlined in this paper is designed so that it can easily take this kind of data, with automatic generation of training data for both open water and flooded vegetation, and robust machine learning algorithms, including automatic selection of optimal parameters.

The developed water body products for Barotseland indicate a number of areas where surface water persists beyond the wet season, and in some cases throughout the year. The identification of these areas is significant in terms of public health risk mapping as these areas are likely to maintain mosquito densities throughout the year, facilitating longer malaria transmission seasons in these parts of Barotseland, and potentially forming 'source' populations for other areas when suitable habitats reappear. As well as representing an increased risk in terms of malaria transmission, these areas

represent a potential target for disease management strategies either directly against the vectors through LSM or by indicating high-transmission zones where patient-based interventions can also be focussed.

We believe that in many ecological settings, targeting interventions based on operational, dynamic mapping of the breeding habitat for mosquito vectors using radar remote sensing is both feasible and potentially an important element of campaigns for disease elimination. But for this approach to be effective and efficient, we need reliable maps of both open and vegetated water. The methods developed in this paper therefore could provide an important tool for public health in malaria-ecological settings found in the Upper Zambezi and other major flooding river systems in Africa. Our work also highlights the need for widely available, high space-time resolution L-band SAR.

Author Contributions: Conceptualization, A.H., C.J.T.; Data curation, A.H., D.E.C., J.S. and T.W.; Formal analysis, A.H.; Funding acquisition, A.H., F.L., A.S., M.S., D.E.C. and C.J.T.; Investigation, A.H., D.E.C., F.L., C.J.T. and M.S.; Methodology, A.H., G.E. and P.B.; Software, A.H.; Validation, A.H. and G.E.; Writing—original draft, A.H., D.E.C., C.J.T. and M.S.; Writing—review & editing, G.E., P.B., F.L., J.S., A.S. and D.S.

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

References

1. Bioresita, F.; Puissant, A.; Stumpf, A.; Malet, J.-P. A Method for Automatic and Rapid Mapping of Water Surfaces from Sentinel-1 Imagery. *Remote Sens.* **2018**, *10*, 217. [CrossRef]
2. Westerhoff, R.S.; Kleuskens, M.P.H.; Winsemius, H.C.; Huizinga, H.J.; Brakenridge, G.R.; Bishop, C. Automated global water mapping based on wide-swath orbital synthetic-aperture radar. *Hydrol. Earth Syst. Sci.* **2013**, *17*, 651–663. [CrossRef]
3. Martinis, S.; Kersten, J.; Twele, A. A fully automated TerraSAR-X based flood service. *ISPRS J. Photogramm. Remote Sens.* **2015**, *104*, 203–212. [CrossRef]
4. Pulvirenti, L.; Pierdicca, N.; Chini, M.; Guerriero, L. An algorithm for operational flood mapping from Synthetic Aperture Radar (SAR) data using fuzzy logic. *Nat. Hazards Earth Syst. Sci.* **2011**, *11*, 529–540. [CrossRef]
5. Giustarini, L.; Hostache, R.; Matgen, P.; Schumann, G.J.-P.; Bates, P.D.; Mason, D.C. A Change Detection Approach to Flood Mapping in Urban Areas Using TerraSAR-X. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 2417–2430. [CrossRef]
6. Behnamian, A.; Banks, S.; White, L.; Brisco, B.; Millard, K.; Pasher, J.; Chen, Z.; Duffe, J.; Bourgeau-Chavez, L.; Battaglia, M. Semi-Automated Surface Water Detection with Synthetic Aperture Radar Data: A Wetland Case Study. *Remote Sens.* **2017**, *9*, 1209. [CrossRef]
7. Schlaffer, S.; Chini, M.; Dettmering, D.; Wagner, W. Mapping Wetlands in Zambia Using Seasonal Backscatter Signatures Derived from ENVISAT ASAR Time Series. *Remote Sens.* **2016**, *8*, 402. [CrossRef]
8. Clewley, D.; Whitcomb, J.; Moghaddam, M.; McDonald, K.; Chapman, B.; Bunting, P. Evaluation of ALOS PALSAR Data for High-Resolution Mapping of Vegetated Wetlands in Alaska. *Remote Sens.* **2015**, *7*, 7272–7297. [CrossRef]
9. Brisco, B.; Li, K.; Tedford, B.; Charbonneau, F.; Yun, S.; Murnaghan, K. Compact polarimetry assessment for rice and wetland mapping. *Int. J. Remote Sens.* **2013**, *34*, 1949–1964. [CrossRef]
10. Bwangoy, J.-R.B.; Hansen, M.C.; Roy, D.P.; De Grandi, G.; Justice, C.O. Wetland mapping in the Congo Basin using optical and radar remotely sensed data and derived topographical indices. *Remote Sens. Environ.* **2010**, *114*, 73–86. [CrossRef]
11. *WHO World Malaria Report 2018*; WHO: Geneva, Switzerland, 2018.

12. Bøgh, C.; Lindsay, S.W.; Clarke, S.E.; Dean, A.; Jawara, M.; Pinder, M.; Thomas, C.J. High spatial resolution mapping of malaria transmission risk in the Gambia, West Africa, using Landsat TM satellite imagery. *Am. J. Trop. Med. Hyg.* **2007**, *76*, 875–881. [CrossRef] [PubMed]

13. Mutuku, F.M.; Bayoh, M.N.; Hightower, A.W.; Vulule, J.M.; Gimnig, J.E.; Mueke, J.M.; Amimo, F.A.; Walker, E.D. A supervised land cover classification of a western Kenya lowland endemic for human malaria: Associations of land cover with larval Anopheles habitats. *Int. J. Health Geogr.* **2009**, *8*, 19. [CrossRef] [PubMed]

14. Mushinzimana, E.; Munga, S.; Minakawa, N.; Li, L.; Feng, C.; Bian, L.; Kitron, U.; Schmidt, C.; Beck, L.; Zhou, G.; et al. Landscape determinants and remote sensing of anopheline mosquito larval habitats in the western Kenya highlands. *Malar. J.* **2006**, *5*, 13. [CrossRef] [PubMed]

15. Pope, K.; Masuoka, P.; Rejmankova, E.; Grieco, J.; Johnson, S.; Roberts, D. Mosquito habitats, land use, and malaria risk in Belize from satellite imagery. *Ecol. Appl.* **2005**, *15*, 1223–1232. [CrossRef]

16. Fillinger, U.; Lindsay, S.W. Suppression of exposure to malaria vectors by an order of magnitude using microbial larvicides in rural Kenya. *Trop. Med. Int. Heal.* **2006**, *11*, 1629–1642. [CrossRef] [PubMed]

17. Killeen, G.F.; Fillinger, U.; Kiche, I.; Gouagna, L.C.; Knols, B.G.J. Eradication of Anopheles gambiae from Brazil: Lessons for malaria control in Africa? *Lancet Infect. Dis.* **2002**, *2*, 618–627. [CrossRef]

18. Smith, M.W.; Macklin, M.G.; Thomas, C.J. Hydrological and geomorphological controls of malaria transmission. *Earth Sci. Rev.* **2013**, *116*, 109–127. [CrossRef]

19. Smith, D.L.; Perkins, T.A.; Tusting, L.S.; Scott, T.W.; Lindsay, S.W. Mosquito Population Regulation and Larval Source Management in Heterogeneous Environments. *PLoS ONE* **2013**, *8*, e71247. [CrossRef] [PubMed]

20. Catry, T.; Li, Z.; Roux, E.; Herbreteau, V.; Gurgel, H.; Mangeas, M.; Seyler, F.; Dessay, N. Wetlands and Malaria in the Amazon: Guidelines for the Use of Synthetic Aperture Radar Remote-Sensing. *Int. J. Environ. Res. Public Health* **2018**, *15*, 468. [CrossRef]

21. Clement, M.A.; Kilsby, C.G.; Moore, P. Multi-temporal synthetic aperture radar flood mapping using change detection. *J. Flood Risk Manag.* **2018**, *11*, 152–168. [CrossRef]

22. Santoro, M.; Wegmuller, U. Multi-temporal Synthetic Aperture Radar Metrics Applied to Map Open Water Bodies. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 3225–3238. [CrossRef]

23. Charlwood, J.D.; Vij, R.; Billingsley, P.F. Dry season refugia of malaria-transmitting mosquitoes in a dry savannah zone of East Africa. *Am. J. Trop. Med. Hyg.* **2000**, *62*, 726–732. [CrossRef] [PubMed]

24. Mason, D.C.; Davenport, I.J.; Neal, J.C.; Schumann, G.J.-P.; Bates, P.D. Near Real-Time Flood Detection in Urban and Rural Areas Using High-Resolution Synthetic Aperture Radar Images. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 3041–3052. [CrossRef]

25. Chini, M.; Hostache, R.; Giustarini, L.; Matgen, P. A Hierarchical Split-Based Approach for Parametric Thresholding of SAR Images: Flood Inundation as a Test Case. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 6975–6988. [CrossRef]

26. Matgen, P.; Hostache, R.; Schumann, G.; Pfister, L.; Hoffmann, L.; Savenije, H.H.G. Towards an automated SAR-based flood monitoring system: Lessons learned from two case studies. *Phys. Chem. Earth Parts A/B/C* **2011**, *36*, 241–252. [CrossRef]

27. Martinis, S.; Plank, S.; Ćwik, K. The Use of Sentinel-1 Time-Series Data to Improve Flood Monitoring in Arid Areas. *Remote Sens.* **2018**, *10*, 583. [CrossRef]

28. Martinis, S.; Twele, A.; Voigt, S. Towards operational near real-time flood detection using a split-based automatic thresholding procedure on high resolution TerraSAR-X data. *Nat. Hazards Earth Syst. Sci.* **2009**, *9*, 303–314. [CrossRef]

29. Bovolo, F.; Bruzzone, L. A Split-Based Approach to Unsupervised Change Detection in Large-Size Multitemporal Images: Application to Tsunami-Damage Assessment. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 1658–1670. [CrossRef]

30. Huang, W.; DeVries, B.; Huang, C.; Lang, M.; Jones, J.; Creed, I.; Carroll, M. Automated Extraction of Surface Water Extent from Sentinel-1 Data. *Remote Sens.* **2018**, *10*, 797. [CrossRef]

31. Hardy, A.; Makame, M.; Cross, D.; Majambere, S.; Msellem, M. Using low-cost drones to map malaria vector habitats. *Parasites Vectors* **2017**, *10*. [CrossRef]

32. Hardy, A.J.; Mageni, Z.; Dongus, S.; Killeen, G.; Macklin, M.; Majambere, S.; Ali, A.S.; Smith, M.; Al-Mafazy, A.; Msellem, M.; et al. Mapping hotspots of malaria transmission from pre-existing hydrology, geology and geomorphology data in the pre-elimination context of Zanzibar, United Republic of Tanzania. *Parasites Vectors* **2015**, *8*, 41. [CrossRef] [PubMed]

33. Dongus, S.; Nyika, D.; Kannady, K.; Mtasiwa, D.; Mshinda, H.; Gosoniu, L.; Drescher, A.W.; Fillinger, U.; Tanner, M.; Killeen, G.F. Urban agriculture and Anopheles habitats in Dar es Salaam, Tanzania. *Geospat. Health* **2009**, *3*, 189–210. [CrossRef] [PubMed]

34. Mwangangi, J.M.; Shililu, J.; Muturi, E.J.; Muriu, S.; Jacob, B.; Kabiru, E.W.; Mbogo, C.M.; Githure, J.; Novak, R.J. Anopheles larval abundance and diversity in three rice agro-village complexes Mwea irrigation scheme, central Kenya. *Malar. J.* **2010**, *9*, 228. [CrossRef]

35. Hardy, A.J.; Gamarra, J.G.P.; Cross, D.E.; Macklin, M.G.; Smith, M.W.; Kihonda, J.; Killeen, G.F.; Ling'ala, G.N.; Thomas, C.J. Habitat Hydrology and Geomorphology Control the Distribution of Malaria Vector Larvae in Rural Africa. *PLoS ONE* **2013**, *8*, e81931. [CrossRef] [PubMed]

36. Ndenga, B.A.; Simbauni, J.A.; Mbugi, J.P.; Githeko, A.K.; Fillinger, U. Productivity of Malaria Vectors from Different Habitat Types in the Western Kenya Highlands. *PLoS ONE* **2011**, *6*, e19473. [CrossRef] [PubMed]

37. Minakawa, N.; Mutero, C.; Githure, J.; Beier, J.; Yan, G. Spatial distribution and habitat characterization of anopheline mosquito larvae in Western Kenya. *Am. J. Trop. Med. Hyg.* **1999**, *61*, 1010–1016. [CrossRef]

38. Kwoun, O.; Lu, Z. Multi-temporal RADARSAT-1 and ERS Backscattering Signatures of Coastal Wetlands in Southeastern Louisiana. *Photogramm. Eng. Remote Sens.* **2009**, *75*, 607–617. [CrossRef]

39. Moser, L.; Schmitt, A.; Wendleder, A.; Roth, A.; Moser, L.; Schmitt, A.; Wendleder, A.; Roth, A. Monitoring of the Lac Bam Wetland Extent Using Dual-Polarized X-Band SAR Data. *Remote Sens.* **2016**, *8*, 302. [CrossRef]

40. Dobson, M.C.; Ulaby, F.T.; Pierce, L.E.; Sharik, T.L.; Bergen, K.M.; Kellndorfer, J.; Kendra, J.R.; Li, E.; Lin, Y.C.; Nashashibi, A.; et al. Estimation of forest biophysical characteristics in Northern Michigan with SIR-C/X-SAR. *IEEE Trans. Geosci. Remote Sens.* **1995**, *33*, 877–895. [CrossRef]

41. Kasischke, E.S.; Smith, K.B.; Bourgeau-Chavez, L.L.; Romanowicz, E.A.; Brunzell, S.; Richardson, C.J. Effects of seasonal hydrologic patterns in south Florida wetlands on radar backscatter measured from ERS-2 SAR imagery. *Remote Sens. Environ.* **2003**, *88*, 423–441. [CrossRef]

42. Pope, K.O.; Rejmankova, E.; Paris, J.F.; Woodruff, R. Detecting seasonal flooding cycles in marshes of the Yucatan Peninsula with SIR-C polarimetric radar imagery. *Remote Sens. Environ.* **1997**, *58*, 157–166. [CrossRef]

43. Wang, Y.; Davis, F.W.; Melack, J.M.; Kasischke, E.S.; Christensen, N.L. The effects of changes in forest biomass on radar backscatter from tree canopies. *Int. J. Remote Sens.* **1995**, *16*, 503–513. [CrossRef]

44. Plank, S.; Jüssi, M.; Martinis, S.; Twele, A. Mapping of flooded vegetation by means of polarimetric Sentinel-1 and ALOS-2/PALSAR-2 imagery. *Int. J. Remote Sens.* **2017**, *38*, 3831–3850. [CrossRef]

45. Tsyganskaya, V.; Martinis, S.; Marzahn, P.; Ludwig, R.; Tsyganskaya, V.; Martinis, S.; Marzahn, P.; Ludwig, R. Detection of Temporary Flooded Vegetation Using Sentinel-1 Time Series Data. *Remote Sens.* **2018**, *10*, 1286. [CrossRef]

46. Tsyganskaya, V.; Martinis, S.; Marzahn, P.; Ludwig, R. SAR-based detection of flooded vegetation—A review of characteristics and approaches. *Int. J. Remote Sens.* **2018**, *39*, 2255–2293. [CrossRef]

47. Clarke, A.D.; Noone, K.J. Soot in the Arctic snowpack: A cause for perturbations in radiative transfer. *Atmos. Environ.* **1985**, *19*, 2045–2053. [CrossRef]

48. Ettritch, G.; Hardy, A.; Bojang, L.; Cross, D.; Bunting, P.; Brewer, P. Enhancing digital surface models for hydraulic modelling using flood frequency detection. *Remote Sens. Environ.* **2018**, *217*, 506–522. [CrossRef]

49. Gimnig, J.E.; Ombok, M.; Kamau, L.; Hawley, W.A. Characteristics of larval anopheline (Diptera: Culicidae) habitats in Western Kenya. *J. Med. Entomol.* **2001**, *38*, 282–288. [CrossRef]

50. Fillinger, U.; Sonye, G.; Killeen, G.F.; Knols, B.G.J.; Becker, N. The practical importance of permanent and semipermanent habitats for controlling aquatic stages of Anopheles gambiae sensu lato mosquitoes: Operational observations from a rural town in western Kenya. *Trop. Med. Int. Heal.* **2004**, *9*, 1274–1289. [CrossRef]

51. Gillies, M.T.; De Meillon, B. The Anophelinae of Africa south of the Sahara (Ethiopian Zoogeographical Region). *Publ. South. Afr. Inst. Med. Res.* **1968**, *54*, 1–343.

52. Fornadel, C.M.; Norris, L.C.; Franco, V.; Norris, D.E. Unexpected Anthropophily in the Potential Secondary Malaria Vectors Anopheles coustani s.l. and Anopheles squamosus in Macha, Zambia. *Vector-Borne Zoonotic Dis.* **2011**, *11*, 1173–1179. [CrossRef] [PubMed]

53. Lobo, N.F.; Laurent, B. St.; Sikaala, C.H.; Hamainza, B.; Chanda, J.; Chinula, D.; Krishnankutty, S.M.; Mueller, J.D.; Deason, N.A.; Hoang, Q.T.; et al. Unexpected diversity of Anopheles species in Eastern Zambia: Implications for evaluating vector behavior and interventions using molecular tools. *Sci. Rep.* **2016**, *5*, 17952. [CrossRef]

54. Amvongo-Adjia, N.; Wirsiy, E.L.; Riveron, J.M.; Chounna Ndongmo, W.P.; Enyong, P.A.; Njiokou, F.; Wondji, C.S.; Wanji, S. Bionomics and vectorial role of anophelines in wetlands along the volcanic chain of Cameroon. *Parasites Vectors* **2018**, *11*, 471. [CrossRef] [PubMed]

55. Feng, M.; Sexton, J.O.; Channan, S.; Townshend, J.R. A global, high-resolution (30-m) inland water body dataset for 2000: First results of a topographic–spectral classification algorithm. *Int. J. Digit. Earth* **2016**, *9*, 113–133. [CrossRef]

56. Pekel, J.-F.; Cottam, A.; Gorelick, N.; Belward, A.S. High-resolution mapping of global surface water and its long-term changes. *Nature* **2016**, *540*, 418–422. [CrossRef]

57. Klein, I.; Dietz, A.; Gessner, U.; Dech, S.; Kuenzer, C. Results of the Global WaterPack: A novel product to assess inland water body dynamics on a daily basis. *Remote Sens. Lett.* **2015**, *6*, 78–87. [CrossRef]

58. Wendleder, A.; Wessel, B.; Roth, A.; Breunig, M.; Martin, K.; Wagenbrenner, S. TanDEM-X Water Indication Mask: Generation and First Evaluation Results. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 171–179. [CrossRef]

59. Dargie, G.C.; Lewis, S.L.; Lawson, I.T.; Mitchard, E.T.A.; Page, S.E.; Bocko, Y.E.; Ifo, S.A. Age, extent and carbon storage of the central Congo Basin peatland complex. *Nature* **2017**, *542*, 86–90. [CrossRef]

60. Moore, A.E.; Cotterill, F.P.D.; Main, M.P.L.; Williams, H. The Zambezi River. In *Large Rivers: Geomorphology and Management*; John Wiley & Sons: Chichester, UK, 2007.

61. Zimba, H.; Kawawa, B.; Chabala, A.; Phiri, W.; Selsam, P.; Meinhardt, M.; Nyambe, I. Assessment of trends in inundation extent in the Barotse Floodplain, upper Zambezi River Basin: A remote sensing-based approach. *J. Hydrol. Reg. Stud.* **2018**, *15*, 149–170. [CrossRef]

62. Kling, H.; Stanzel, P.; Preishuber, M. Impact modelling of water resources development and climate scenarios on Zambezi River discharge. *J. Hydrol. Reg. Stud.* **2014**, *1*, 17–43. [CrossRef]

63. Cai, X.; Haile, A.T.; Magidi, J.; Mapedza, E.; Nhamo, L. Living with floods—Household perception and satellite observations in the Barotse floodplain, Zambia. *Phys. Chem. Earth, Parts A/B/C* **2017**, *100*, 278–286. [CrossRef]

64. Dunne, T.; Aalto, R. Large River Floodplains. In *Treatise on Geomorphology*; Shroder, J., Ed.; Academic Press: San Diego, CA, USA, 2013; pp. 645–678.

65. Latrubesse, E.M. Patterns of anabranching channels: The ultimate end-member adjustment of mega rivers. *Geomorphology* **2008**, *101*, 130–145. [CrossRef]

66. IUCN. *Barotse Floodplain, Zambia: Local Economic Dependence on Wetland Resources*; IUCN: Harare, Zimbabwe, 2003.

67. Day, G.; Dietrich, W.E.; Rowland, J.C.; Marshall, A. The depositional web on the floodplain of the Fly River, Papua New Guinea. *J. Geophys. Res.* **2008**, *113*, F01S02. [CrossRef]

68. Mertes, L.A.K. Documentation and significance of the perirheic zone on inundated floodplains. *Water Resour. Res.* **1997**, *33*, 1749–1762. [CrossRef]

69. Burrough, S.L.; Thomas, D.S.G.; Oriojemie, E.A.; Willis, K.J. Landscape sensitivity and ecological change in western Zambia: The long-term perspective from dambo cut-and-fill sediments. *J. Quat. Sci.* **2015**, *30*, 44–58. [CrossRef]

70. Lewin, J.; Ashworth, P.J. Defining large river channel patterns: Alluvial exchange and plurality. *Geomorphology* **2014**, *215*, 83–98. [CrossRef]

71. Zambia Ministry of Health. *Zambia National Malaria Indicator Survey*; Ministry of Health: Lusaka, Zambia, 2012.

72. Del Rio, T.; Groot, J.C.J.; DeClerck, F.; Estrada-Carmona, N. Integrating local knowledge and remote sensing for eco-type classification map in the Barotse Floodplain, Zambia. *Data Br.* **2018**, *19*, 2297–2304. [CrossRef]

73. Timberlake, J. *Biodiversity of the Zambezi Basin*; Biodiversity Foundation for Africa: Bulawayo, Zimbabwe, 2000.

74. Esch, T.; Taubenböck, H.; Roth, A.; Heldens, W.; Felbier, A.; Thiel, M.; Schmidt, M.; Müller, A.; Dech, S. TanDEM-X mission—New perspectives for the inventory and monitoring of global settlement patterns. *J. Appl. Remote Sens.* **2012**, *6*, 061702. [CrossRef]

75. Bunting, P.; Clewley, D.; Lucas, R.M.; Gillingham, S. The Remote Sensing and GIS Software Library (RSGISLib). *Comput. Geosci.* **2014**, *62*, 216–226. [CrossRef]

76. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.

77. Kwak, Y.; Yun, S.; Iwami, Y. A new approach for rapid urban flood mapping using ALOS-2/PALSAR-2 in 2015 Kinu River Flood, Japan. In Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 23–28 July 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 1880–1883.

78. Lopes, A.; Touzi, R.; Nezry, E. Adaptive speckle filters and scene heterogeneity. *IEEE Trans. Geosci. Remote Sens.* **1990**, *28*, 992–1000. [CrossRef]

79. Clewley, D.; Bunting, P.; Shepherd, J.; Gillingham, S.; Flood, N.; Dymond, J.; Lucas, R.; Armston, J.; Moghaddam, M. A Python-Based Open Source System for Geographic Object-Based Image Analysis (GEOBIA) Utilizing Raster Attribute Tables. *Remote Sens.* **2014**, *6*, 6111–6135. [CrossRef]

80. Rennó, C.D.; Nobre, A.D.; Cuartas, L.A.; Soares, J.V.; Hodnett, M.G.; Tomasella, J.; Waterloo, M.J. HAND, a new terrain descriptor using SRTM-DEM: Mapping terra-firme rainforest environments in Amazonia. *Remote Sens. Environ.* **2008**, *112*, 3469–3481. [CrossRef]

81. Twele, A.; Cao, W.; Plank, S.; Martinis, S. Sentinel-1-based flood mapping: A fully automated processing chain. *Int. J. Remote Sens.* **2016**, *37*, 2990–3004. [CrossRef]

82. Geurts, P.; Ernst, D.; Wehenkel, L. Extremely randomized trees. *Mach. Learn.* **2006**, *63*, 3–42. [CrossRef]

83. Maxwell, A.E.; Warner, T.A.; Fang, F. Implementation of machine-learning classification in remote sensing: An applied review. *Int. J. Remote Sens.* **2018**, *39*, 2784–2817. [CrossRef]

84. Grenier, M.; Demers, A.-M.; Labrecque, S.; Benoit, M.; Fournier, R.A.; Drolet, B. An object-based method to map wetland using RADARSAT-1 and Landsat ETM images: Test case on two sites in Quebec, Canada. *Can. J. Remote Sens.* **2007**, *33*, S28–S45. [CrossRef]

85. Harvey, K.R.; Hill, G.J.E. Vegetation mapping of a tropical freshwater swamp in the Northern Territory, Australia: A comparison of aerial photography, Landsat TM and SPOT satellite imagery. *Int. J. Remote Sens.* **2001**, *22*, 2911–2925. [CrossRef]

86. Michishita, R.; Gong, P.; Xu, B. Spectral mixture analysis for bi-sensor wetland mapping using Landsat TM and Terra MODIS data. *Int. J. Remote Sens.* **2012**, *33*, 3373–3401. [CrossRef]

87. Pontius, R.G.; Millones, M. Death to Kappa: Birth of quantity disagreement and allocation disagreement for accuracy assessment. *Int. J. Remote Sens.* **2011**, *32*, 4407–4429. [CrossRef]

88. Warrens, M.J. Relative quantity and allocation disagreement measures for category-level accuracy assessment. *Int. J. Remote Sens.* **2015**, *36*, 5959–5969. [CrossRef]

89. Comber, A.; Fisher, P.; Brunsdon, C.; Khmag, A. Spatial analysis of remote sensing image classification accuracy. *Remote Sens. Environ.* **2012**, *127*, 237–246. [CrossRef]

90. Haas, J.; Ban, Y. Urban growth and environmental impacts in Jing-Jin-Ji, the Yangtze, River Delta and the Pearl River Delta. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *30*, 42–55. [CrossRef]

91. Bennett, N.D.; Croke, B.F.W.; Guariso, G.; Guillaume, J.H.A.; Hamilton, S.H.; Jakeman, A.J.; Marsili-Libelli, S.; Newham, L.T.H.; Norton, J.P.; Perrin, C.; et al. Characterising performance of environmental models. *Environ. Model. Softw.* **2013**, *40*, 1–20. [CrossRef]

92. Olofsson, P.; Foody, G.M.; Stehman, S.V.; Woodcock, C.E. Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sens. Environ.* **2013**, *129*, 122–131. [CrossRef]

93. Olofsson, P.; Foody, G.M.; Herold, M.; Stehman, S.V.; Woodcock, C.E.; Wulder, M.A. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* **2014**, *148*, 42–57. [CrossRef]

94. Patel, P.; Srivastava, H.S.; Navalgund, R.R. Use of synthetic aperture radar polarimetry to characterize wetland targets of Keoladeo National Park, Bharatpur, India. *Curr. Sci.* **2009**, *97*, 529–537.

95. Thullen, J.S.; Sartoris, J.J.; Walton, W.E. Effects of vegetation management in constructed wetland treatment cells on water quality and mosquito production. *Ecol. Eng.* **2002**, *18*, 441–457. [CrossRef]

96. Oesterholt, M.; Bousema, J.T.; Mwerinde, O.K.; Harris, C.; Lushino, P.; Masokoto, A.; Mwerinde, H.; Mosha, F.W.; Drakeley, C.J. Spatial and temporal variation in malaria transmission in a low endemicity area in northern Tanzania. *Malar. J.* **2006**, *5*, 98. [CrossRef]

97. Fillinger, U.; Sombroek, H.; Majambere, S.; Van Loon, E.; Takken, W.; Lindsay, S.W. Identifying the most productive breeding sites for malaria mosquitoes in The Gambia. *Malar. J.* **2009**, *8*, 62. [CrossRef]

98. Sinka, M.E.; Bangs, M.J.; Manguin, S.; Coetzee, M.; Mbogo, C.M.; Hemingway, J.; Patil, A.P.; Temperley, W.H.; Gething, P.W.; Kabaria, C.W. The dominant Anopheles vectors of human malaria in Africa, Europe and the Middle East: Occurrence data, distribution maps and bionomic précis. *Parasites Vectors* **2010**, *3*, 117. [CrossRef] [PubMed]

99. McCrae, A. Oviposition by African malaria vector mosquitoes. II. Effects of site tone, water type and conspecific immatures on target selection by freshwater Anopheles gambiae. *Ann. Trop. Med. Parasitol.* **1984**, *78*, 307–318. [CrossRef] [PubMed]

100. Paaijmans, K.P.; Takken, W.; Githeko, A.K.; Jacobs, A.F.G. The effect of water turbidity on the near-surface water temperature of larval habitats of the malaria mosquito Anopheles gambiae. *Int. J. Biometeorol.* **2008**, *52*, 747–753. [CrossRef]