Small dams: determining the minimum waterbody surface area that can be successfully detected using Sentinel-1 SAR imagery

M von Fintel[1], J Kemp[1],

[1]Department of Geography and Environmental Studies, Stellenbosch University

DOI: http://dx.doi.org/10.4314/sajg.v11i2.9

Abstract

Water is a scarce resource in South Africa, and approximately 62% of the water used in South Africa is for irrigation. This water is stored in many small dams scattered across the country. If not managed correctly, they could have a negative effect on catchment areas and on the availability of water. As such, there is a need for a new monitoring and management system to be developed. This study determined the minimum surface area that would be required for a waterbody to be detected on Sentinel-1 Synthetic Aperture Radar imagery. A Random Forest classifier was used to detect waterbodies on a Sentinel-1 image calculated from a time series of imagery taken over a period of three months. Steep incidence angles outperformed shallow incidence angles, with the classification having an overall accuracy of 80%. Detection rates were almost 90% for waterbodies of one hectare and greater, with no false positives, and a 10% false negative rate. These findings provide the foundation for developing a detection and monitoring system, which would allow for the better management of water resources in South Africa.

1. Introduction

South Africa has a semi-arid climate and water is a scarce resource (Friedrich, Pillay & Buckley, 2009). Of all the water used in the country, 62% is used for irrigation (Claassen, 2010) and this figure is likely to rise with increasing pressure on the agricultural sector to produce more food (De Lange, 2010). In semi-arid climates, dams facilitate dry season water storage for irrigation (Stephens, 2010). In South Africa, most dams are typically smaller than 2.5 hectares (ha) (DWAF, 1986). Owing to the increasing gap between the need for and the availability of water in the Western Cape province, there is mounting pressure to build more dams (du Plessis 2017; van der Zaag & Gupta, 2008). The construction of these small dams limits the total amount of water available downstream (Tingey-Holyoak, 2010), reduces streamflow (Habets *et al.*, 2018), increases the salinity of the water (Bugan, Jovanovic & De Clercq, 2012) and causes siltation of the hydrological system (van der Zaag & Gupta, 2008).

There is, therefore, a need to implement a dam monitoring system to ensure that the agricultural sector can fulfil its role, without causing permanent damage to the hydrological system. This will allow the Department of Water and Sanitation (DWS) and the Western Cape Government (WCG) to

manage and maintain the system in a more efficient and effective manner. While DWS does in fact maintain a list of dams in South Africa, it only contains those registered as a safety risk, and the list is not complete. Hence, the total number of dams in South Africa is not accurately known (Boardman *et al.*, 2009), and a new method of monitoring these small dams is needed (WCG, 2018). There is also a need for more research on the effects of small dams on the hydrological system in the South African context (Boardman *et al.*, 2009).

Remote sensing has proven useful in many studies relating to the mapping and monitoring of water resources (Mahdavi *et al.* 2018; Smith, 1997). Specifically, there is often little data available on the smallest of dams, to the effect that the methods provided by remote sensing for gaining information on this topic are valuable. Synthetic Aperture Radar (SAR) has been shown to be a favourable system for waterbody mapping since open water produces a distinctive return that can be clearly identified (Brisco *et al.*, 2009). SAR imagery is often preferred above optical imagery for waterbody mapping as it is not limited by atmospheric conditions and is also independent of illumination from the sun (Lettenmaier, 2007). Unlike optical imagery, SAR imagery also provides surface structure information and, since water is structurally different from most other land cover (LC) classes, it can easily be differentiated (Muro *et al.*, 2016). SAR sensors transmit a microwave pulse, which interacts with the surface of the Earth and then records the intensity of the pulse that is backscattered (Moreira *et al.*, 2013). The backscatter intensity is dependent on the texture and structure of the observed surface (Curlander & McDonough, 1991), as well as the incidence angle of the incoming pulse and the dielectric properties of the surface (McCandless & Jackson, 2004). Surfaces, such as waterbodies, that are smooth relative to the wavelength of the incoming pulse, cause specular reflection and, as very little of the pulse is returned to the sensor, generally appear as dark patches on SAR images (Liebe *et al.*, 2009). Despite these benefits, SAR imagery can be expensive, has a high processing time and there is often a long delay before the data become available. The Sentinel-1 constellation overcomes these drawbacks, as the European Space Agency (ESA) freely provides imagery with a high temporal resolution and a short time period between capture and availability (Torres *et al.*, 2012). Co-polarised horizontal (HH) imagery is generally considered superior to other polarisations for waterbody mapping (Henry *et al.*, 2006; Twele *et al.*, 2016), but, unless a special request is made, Sentinel-1 does not provide this information. Nonetheless, Sentinel-1 has been shown to be a suitable sensor for many waterbody mapping applications (Amitrano, di Martino, *et al.*, 2014, Bangira *et al.*, 2019, Janse van Rensburg & Kemp, 2022).

The most common method of using SAR for waterbody mapping is to use a threshold to divide backscatter imagery into regions of similar intensities (Bioresita *et al.*, 2018; Heine *et al.*, 2014; Li & Wang, 2015). Another common method is to use machine learning classification algorithms to perform the threshold splitting (Bangira *et al.*, 2019; Long, Fatoyinbo & Policelli, 2014; Pôssa & Maillard, 2018). Studies that have shown the most promising results have also used various methods of filtering and smooth labelling to reduce speckle and to take advantage of spatial

autocorrelation (Amitrano, Ciervo, *et al.*, 2014; Bioresita *et al.*, 2018). While many studies have extracted waterbodies successfully from SAR imagery (Amitrano, Ciervo, *et al.*, 2014; Brisco *et al.*, 2009; Heine *et al.*, 2014), limited research has been undertaken to detect waterbodies of less than 1ha.

In this study, small dams are classified using Sentinel-1 imagery, and a determination is made of the minimum waterbody surface area that can be successfully detected from the classification. The impact of incidence angle on the detection rate is also assessed. Waterbodies are classified using a pixel-based Random Forest (RF) algorithm applied to a Sentinel-1 image which has undergone multitemporal filtering. The surface areas of the waterbodies in the study area range from 0.01ha to 12ha. Since the smallest dams are often covered by only one or two pixels, smooth labelling or filtering was not used.

2. Data and Methods

The data and methods used are outlined in the following sections.

2.1. Study area

The chosen study area is situated in the Western Cape province of South Africa, between Cape Town and Malmesbury, and is illustrated in Figure 1.

The study area is located in the overlap of two Sentinel-1 orbits. There are many farm dams of various sizes distributed across the study area that are used mainly to store freshwater for irrigation (Boardman *et al.*, 2009). The study area is covered predominantly by agricultural farmland, with wheat, canola and grapes being the main crop types (Halpern & Meadows, 2013). The remainder of the study area is mostly urban or covered in natural vegetation, the latter often found in areas of greater topographic relief. The area experiences a Mediterranean climate, with a winter rainfall of approximately 450mm per annum (CWDM, 2017).

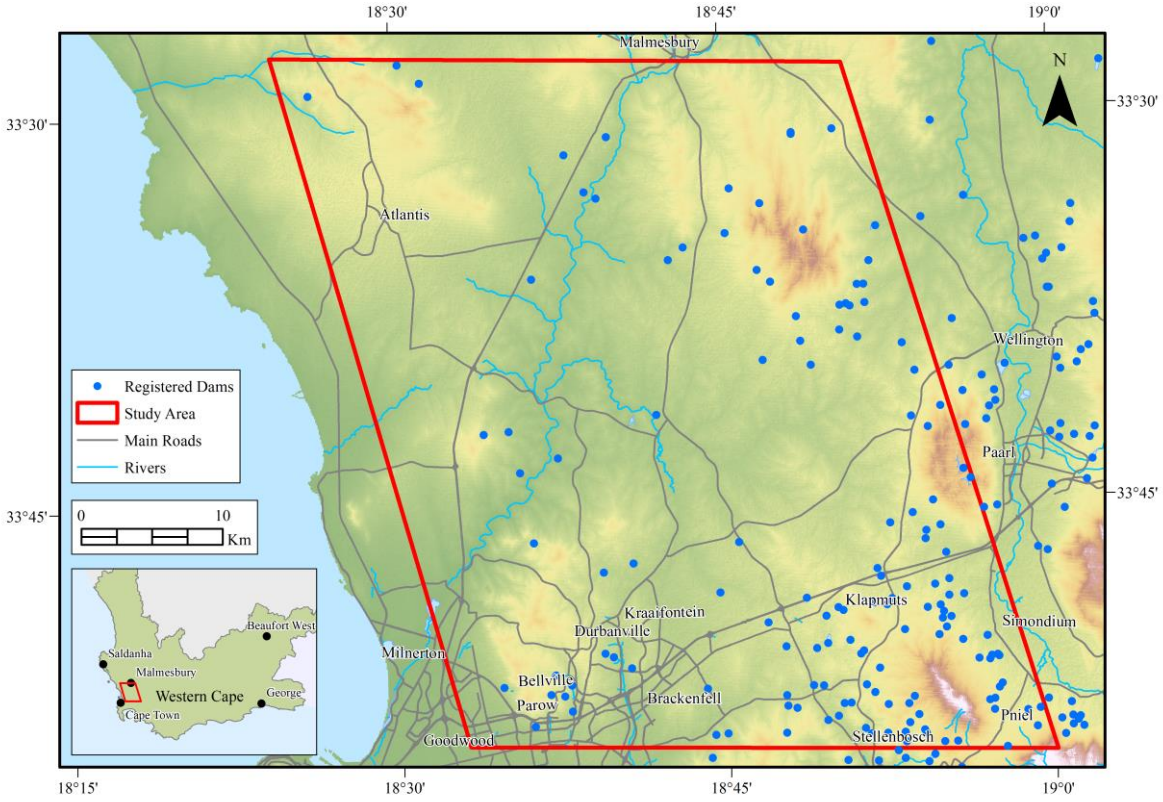

Figure 1. The study area is between the towns of Cape Town and Malmesbury. The blue dots represent dams that are present in the DWS list of registered dams.

2.2. Experimental design

The workflow of the study can be seen in Figure 2. Sentinel-1 images were processed and classified, and the accuracy of the classifications was assessed. Sections 2.3 and 2.4 describe these processing steps in greater detail.

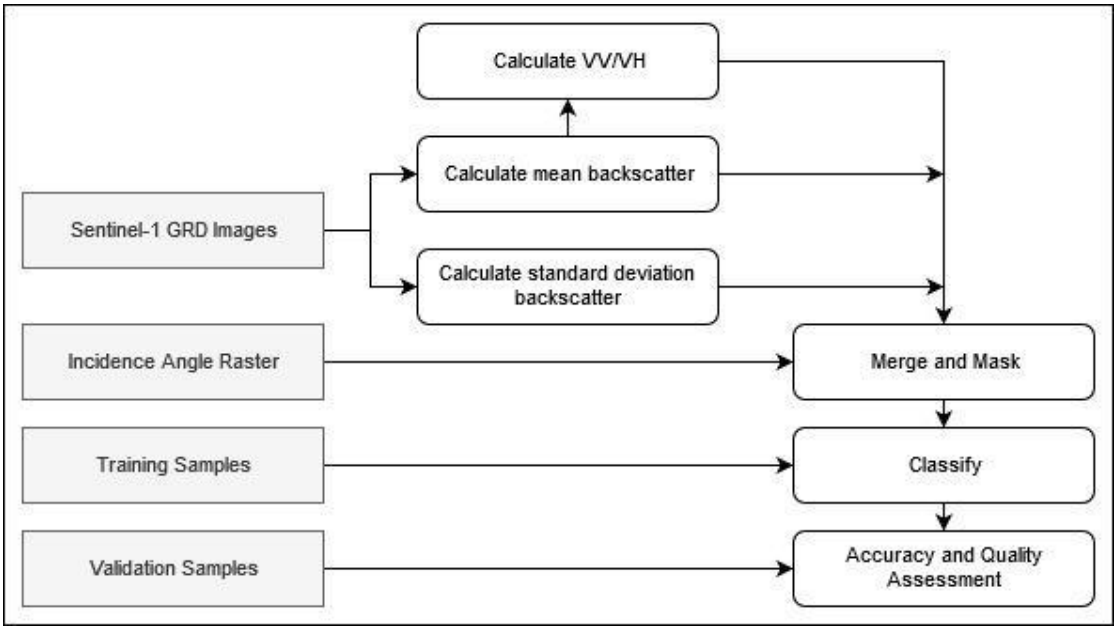

Figure 2. The processing workflow of the study.

2.3. Data collection and preparation

Sentinel-1 imagery was used for image classification, while ancillary Sentinel-2 and aerial imagery were used to collect training and validation datasets.

2.3.1. Sentinel-1 imagery

The study used C-band ($\lambda = 5.6$ cm) SAR imagery that was captured in 20m × 5m pixels and was resampled to a 10m resolution. The imagery had been captured in the Interferometric Wide (IW) swath mode in both co-polarised Vertical-Vertical (VV) and cross polarised Vertical-Horizontal (VH) polarisations. The backscattering coefficient, Sigma-naught (σ°), was converted to decibels (dB) via log scaling. As indicated in Table 1, images from six different dates and from two different orbits were used.

GEE's Sentinel-1 GRD ImageCollection was used (GEE 2019). Thermal noise removal, radiometric calibration and terrain correction had already been performed by GEE using the Sentinel-1 Toolbox. The images were all taken over the two months at the end of the wet season in 2018 when water levels in the dams are expected to be at their highest. A mean backscatter image was calculated for the time period, where each pixel value in this image represented the average backscatter intensity of that pixel over the six images. A standard deviation image was also calculated in the same manner. However, the standard deviation was used instead of the mean. This mitigated the effect of speckle, combined all the data and added greater definition to the shorelines.

Table 1. Product names of the Sentinel-1 and Sentinel-2 imagery that were used, along with acquisition dates and relative orbit numbers.

Sensor	Product Name	Acquisition Date	Relative Orbit Number
S1	S1A_IW_GRDH_1SDV_20180701T172601_20180701T172626_022603_0272EE_81AD	1 Jul 2018	131
S1	S1A_IW_GRDH_1SDV_20180713T172602_20180713T172627_022778_027822_AF30	13 Jul 2018	131
S1	S1A_IW_GRDH_1SDV_20180725T172602_20180725T172627_022953_027DA9_FF74	25 Jul 2018	131
S1	S1A_IW_GRDH_1SDV_20180806T172603_20180806T172628_023128_02831F_CC19	6 Aug 2018	131
S1	S1A_IW_GRDH_1SDV_20180818T172604_20180818T172629_023303_0288CA_3C32	18 Aug 2018	131
S1	S1A_IW_GRDH_1SDV_20180830T172605_20180830T172630_023478_028E58_0B25	30 Aug 2018	131
S1	S1A_IW_GRDH_1SDV_20180624T173408_20180624T173433_022501_026FE5_43AC	24 Jun 2018	29
S1	S1A_IW_GRDH_1SDV_20180706T173409_20180706T173434_022676_0274FE_0B82	6 Jul 2018	29
S1	S1A_IW_GRDH_1SDV_20180718T173409_20180718T173434_022851_027A5A_11BA	18 Jul 2018	29
S1	S1A_IW_GRDH_1SDV_20180730T173410_20180730T173435_023026_027FDF_226C	30 Jul 2018	29
S1	S1A_IW_GRDH_1SDV_20180811T173411_20180811T173436_023201_02856B_9DF6	11 Aug 2018	29
S1	S1A_IW_GRDH_1SDV_20180823T173411_20180823T173436_023376_028B14_1DA1	23 Aug 2018	29
S2	20180730T081559_20180730T083741_T34HBH	30 Jul 2018	

These images were then masked, using a projected local incidence angle raster, to exclude as many areas of radar shadow as possible (Prasad *et al.* 2019). The projected local incidence angle raster was generated in SNAP, using the 30m resolution SRTM DEM. The incidence angles for relative orbit 29 ranged from 42.9° to 44.9°, and for relative orbit 131, from 31.7° to 34.3°. Relative orbit 29 was therefore used for shallow incidence angles and relative orbit 131 for steep incidence angles.

2.3.2. Training and validation data

Three different datasets were created for training and validation use. One was used for training and the other two for validation. The first dataset, used to train the classifier, was digitised from the Sentinel-2 image. Table 2 shows the seven different LC classes identified in the study area, with 40 polygons digitised per class, as in Myburgh and Van Niekerk (2014). This dataset was used to train the classifier.

Table 2. Land cover classes represented in the study area.

Number	Class Name	Description
0	Bare Ground	Artificial and natural areas where the mineral layer is not covered by any vegetation or buildings
1	Urban	Man-made/built features, such as houses, farm buildings and large warehouses
2	Natural Vegetation	Brushy vegetation, consisting of Fynbos, dune plants and grasslands
3	Water	Areas of standing water, natural or man-made
4	Vineyards and Orchards	Woody crops planted in a defined pattern in large fields
5	Wheat and Canola Fields	Smaller crops planted in large fields; high temporal variability
6	Hard Man-made Surfaces	Surfaces such as parking lots, roads, covered reservoirs and airstrips

The second dataset, used for validation, consisted of polygons representing the respective water extents of the waterbodies in the study area. These polygons were digitised from high-resolution (0.5m) colour aerial photographs. As seen in Table 3, 120 waterbodies were digitised for each size class. Since these aerial images were captured in 2014, each polygon was checked against the Sentinel-2 image to ensure that the water extent was current.

Table 3. Waterbody extent size classes.

Number	Class Name	Abbreviation	Surface Area (ha)
0	Extra-small	XS	0.00 – < 0.20
1	Small	S	0.20 – < 0.50
2	Medium	M	0.50 – < 1.00
3	Large	L	1.00 – < 2.10
4	Extra-large	XL	≥ 2.10

The third dataset, which was the second dataset used for validation, consisted of polygons representing the non-water LC classes. Six hundred (600) points were randomly generated across the study area. A random selection was made and then buffered by a random distance. This was repeated multiple times until there were 120 polygons in each size range, as with each size class in the water extents dataset. The LC class for each polygon was then allocated from the Sentinel-2 image.

2.4. Classification and accuracy assessment

Using the mean backscatter image, a ratio index between VV and VH was calculated. This index, along with the mean and standard deviation VV and VH backscatter images, was used as input data to an RF classification algorithm in GEE. The number of classification trees per class was tested, and it was found that 30 trees per class performed best. As recommended by Breiman (2001), the other RF hyperparameters were left unchanged. This produced a binary water/not-water classification. The RF algorithm was run separately on imagery from each incidence angle, producing two different classifications.

Three different accuracy assessment and validation steps were performed. First, the accuracy of the classifications across all size classes was assessed using an error matrix. The error matrix was constructed using both the water extents dataset and the non-water dataset. A random sample of 80 water extent polygons and 80 non-water polygons was taken from each size class. The percentage of pixels correctly classified in each polygon was calculated and a threshold was set to determine whether the polygon was detected as a waterbody or not. This percentage was termed the detection rate of the dam. If the detection rate was above 30%, that polygon was considered a successfully detected waterbody. Polygons in the water extents dataset that were not detected were regarded as false negatives and polygons in the non-water dataset that were detected were regarded as false positives.

Second, using the 30% detection rate threshold, each polygon in the water extents dataset was marked as either detected or undetected. The smallest size class where more than 70% of the waterbodies in that class were detected, was considered the smallest class that could be successfully detected.

Third, error matrices were constructed for each size class to give more insight into false positive and negative rates within each size class. Since the number of samples per class was imbalanced, a random sample of 80 water extent polygons and 80 non-water polygons was taken from each class, and a confusion matrix was generated. This was repeated three times and the results averaged, so that a larger number of samples could be used without having an error matrix affected by the imbalanced class sizes (Sokolova & Lapalme, 2009). This yielded the final per-class accuracy metrics.

3. Results

The error matrices for steep incidence angles and shallow incidence angles can be seen in Table 4 and Table 5 respectively. The Overall Accuracy (OA), False Positive Rate (FPR) and False Negative Rate (FNR), from the error matrices can be seen in Table 6.

Table 4. Error matrix for classification produced from steep incidence angles (relative orbit 131).

		Classification		
		Not Dam	**Dam**	**Totals**
Reference	**Not Dam**	390	10	400
	Dam	146	254	400
	Totals	536	264	800

Table 5. Error matrix for classification produced from shallow incidence angles (relative orbit 29).

		Classification		
		Not Dam	**Dam**	**Totals**
Reference	**Not Dam**	386	14	400
	Dam	164	236	400
	Totals	550	250	800

Table 6. Accuracies of the classifications produced from shallow and steep incidence angles.

	Shallow (29)	**Steep (131)**
OA (%)	77.80	80.50
FPR (%)	3.50	2.50
FNR (%)	41.00	36.50

The classification of imagery taken at a steep incidence angle (relative orbit 131) performed better overall, and it didn't under-classify as much as the shallow incidence angle (relative orbit 29) did. Highlighted by the lower FPR and FNR, it produced fewer misclassifications. The classification of imagery from a shallow incidence angle did in fact produce marginally more false positives, however, the main contributor to the difference in OA was the much higher FNR.

Figure 3 provides more insight regarding the detection rate of each dam in the reference dataset for steep incidence angles.

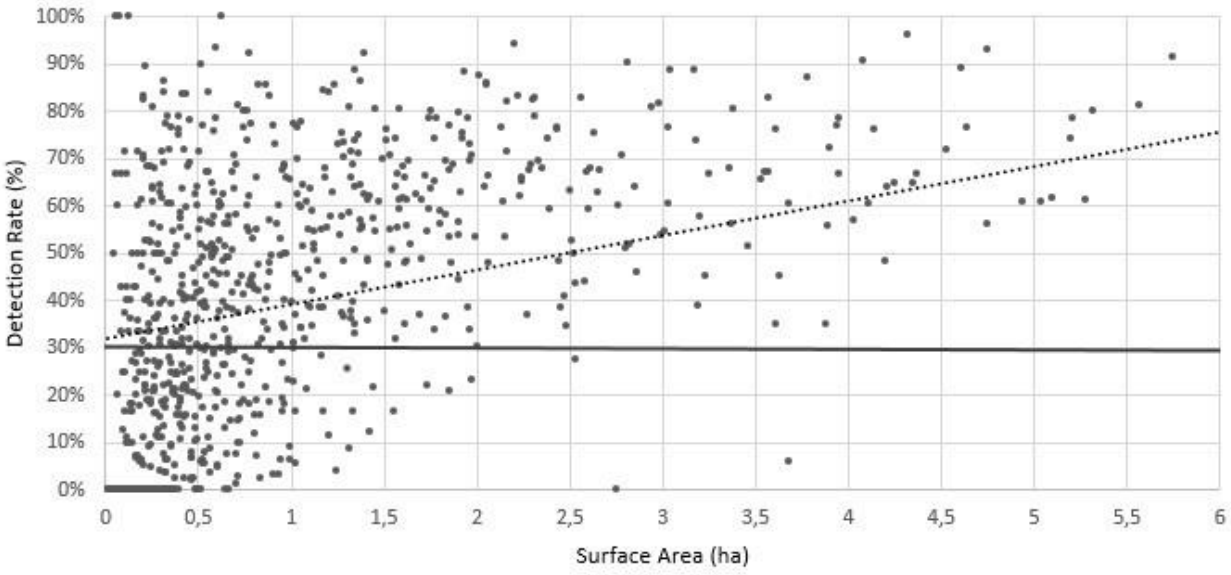

Figure 3. Correlation between per dam detection rate and surface area for the steep incidence angle. Only dams with a surface area of less than six (6) ha are shown here. The 30% detection rate threshold line is shown in bold. The dotted line shows the ordinary least squares regression line.

Despite a very weak positive correlation between detection rate and surface area (R^2 = 0.1976), most of the dams larger than one ha were found to exhibit a detection rate of above 30%. Dams lying below the 30% detection rate threshold were considered false negatives. Thirty percent (30%) was shown to be a suitable detection threshold since this is the point where the low correlation would cause detection rates to drop unnecessarily.

Dams that have no pixels classified as water lie on the 0% line and can be seen mainly between 0 and 0.5ha. Under-classification of smaller dams is an issue, as shown by the high FNRs in Table 6 and Table 7. However, as the surface area increases, fewer dams remain undetected. This can be seen by comparing the results for the different size classes. The difference in percentage dams detected per size class between the classifications from the two incidence angles is shown in Figure 4.

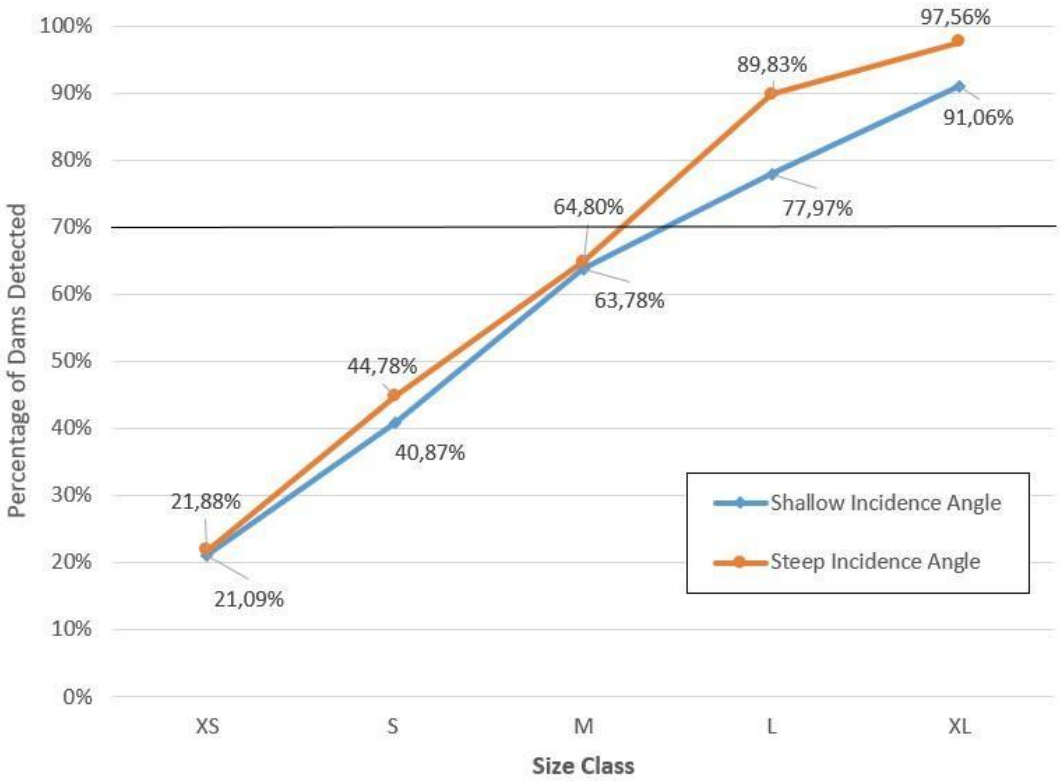

Figure 4. Percentage of dams detected in each size class for the two incidence angles. The 70% detection threshold is shown in bold.

The percentage of dams successfully detected per size class is comparable for the two incidence angles, up until the L class, where the steep incidence angles start to perform better. The only classes above the 70% threshold on both incidence angles are the L and XL classes. However, the M class is approaching that threshold. In the XL class, almost all the dams were successfully detected. A further breakdown of the results for steep incidence angles can be seen in Table 7 and Figure 5.

Table 7. Accuracies, FPR's and FNR's per size class for steep incidence angles.

Class	**XS**	**S**	**M**	**L**	**XL**
OA (%)	57.27	68.37	81.27	94.80	98.77
FPR (%)	6.67	5.83	3.33	0.00	0.00
FNR (%)	78.75	57.50	34.17	10.42	2.50

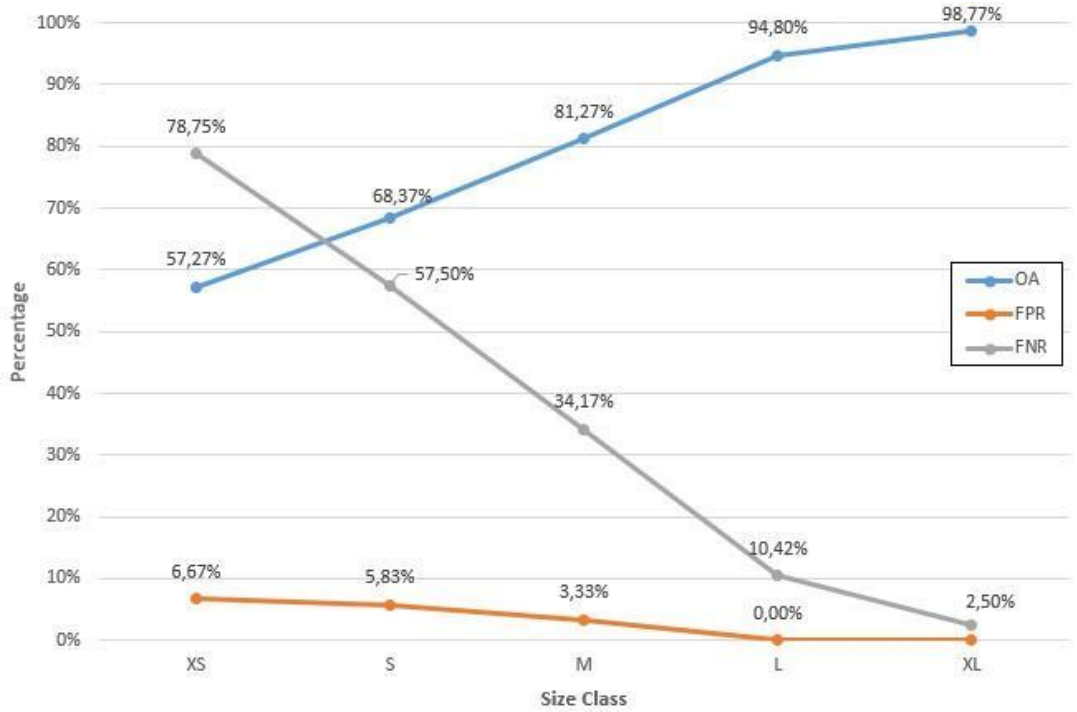

Figure 5. Accuracies, FPR's and FNR's per size class for steep incidence angles.

The effect of the under-classification is evident for the smaller classes, but it decreases as the surface area increases, as shown by the rapidly decreasing FNR. For the smaller classes, the FNR is very high (78.75% for the XS class), thereby contributing to lower OAs for those classes. The FPRs also decrease slightly, further leading to the OA increasing for each class. The M, L and XL classes all have overall accuracies above 80%. From the L class and above, almost all the dams are detected and there are no false positives.

Nine different areas of interest from the study area are shown in Figure 6. The red areas are pixels that were classified as water only on imagery taken at a steep incidence angle, the yellow areas are pixels that were classified as water only on imagery taken at a shallow incidence angle and the blue areas are pixels that were classified as water on imagery from both incidence angles (steep ∩ shallow).

Figure 6 B, E and G show dams of the L class that were classified sufficiently by each, although B and G were classified more successfully using steep incidence angles, as highlighted by the red pixels on the borders. Figure 6 A, C and F show dams that were classified more effectively by using steep incidence angles. Figure 6 H shows some false positives in fields that would have been covered in dense crop during the time period of the study and Figure 6 I shows a runway that was misclassified as water. Figure 6 F and D are examples of dams, from the XS and S classes respectively, that were not entirely classified and could be examples of dams where the water level is lower than that in the reference aerial imagery.

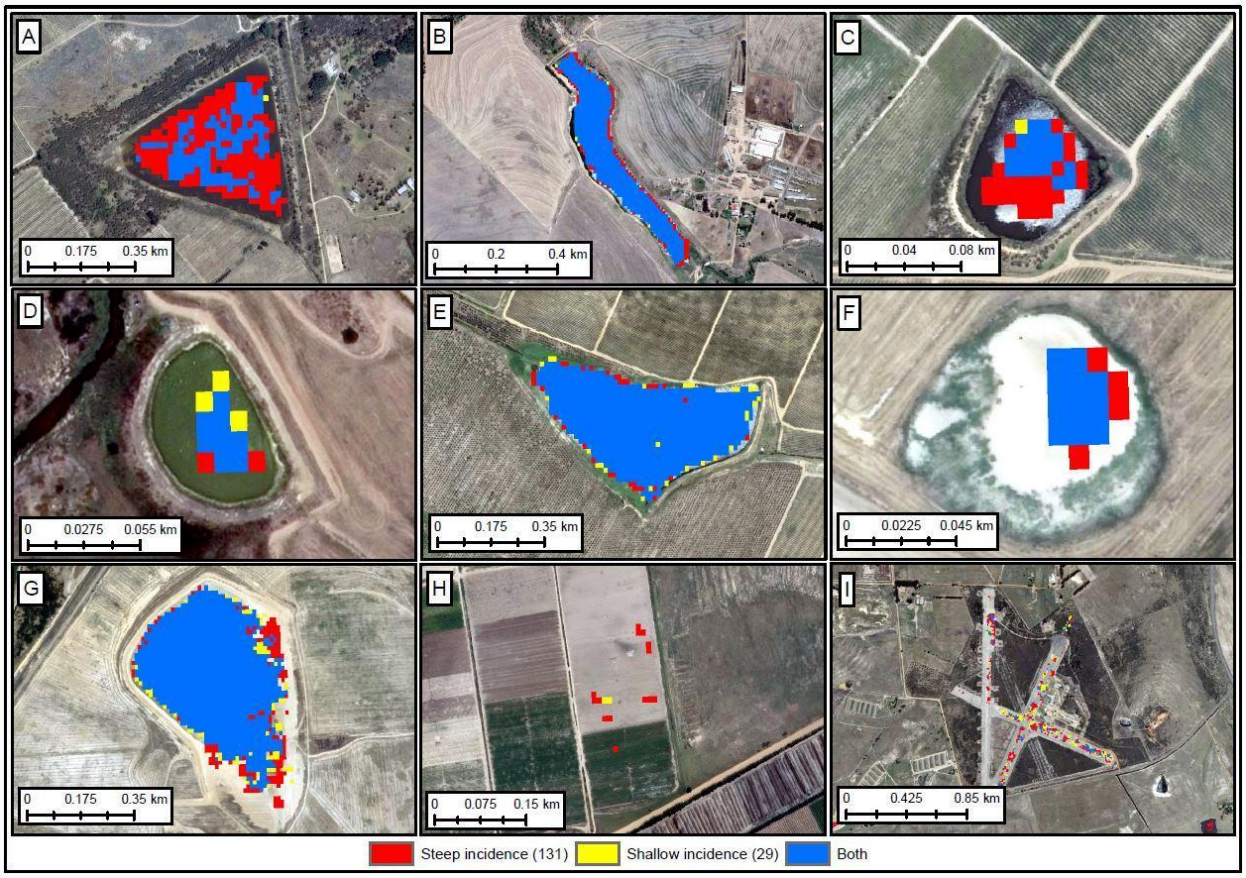

Figure 6. Examples of dam classifications in nine different areas in the study area. Note that the scale differs for each block.

4. Discussion

In this section, the results of the previous section are discussed.

4.1. The effect of vegetation on edge pixels

The main issue experienced with classifying small waterbodies in the current study was that the edge pixels of the dams were often misclassified. This negative effect becomes more pronounced as the surface area of the waterbody gets smaller. The edge pixels are either mixed pixels or are influenced by the vegetation on the edges of the waterbody. Mixed pixels occur when the boundaries of mapping units lie within a single pixel (Fisher, 1997). Pixels that cover both land and water would contain an aggregated backscatter value for both the land and the water, leading to a pixel that is darker than the surrounding LC, but lighter than water.

The second factor influencing edge pixels would be that, since SAR can penetrate vegetation cover, backscatter values would have a higher variability along the edges or margins of the waterbodies. This is due to the different scattering mechanisms caused by reeds, trees or submerged vegetation (Cazals *et al.*, 2016). Vegetation would also cause volume scattering and would, therefore, appear brighter, causing the edge pixels that overlap the vegetation surrounding the

waterbody and the water's edge to be misclassified (Amitrano, Ciervo, *et al.*, 2014; Heine *et al.*, 2014).

Since the proportion of edge pixels to pure pixels for a waterbody increases as the surface area of the waterbody decreases, classifications of smaller waterbodies would tend to exhibit a lower detection rate than larger dams, as can be seen in Figure 6 D and Figure 6 F. This would also explain the lack of correlation in the smaller sized waterbodies in Figure 3. The FNR decreases rapidly as the size classes get larger (Figure 5) since the larger classes would have a larger proportion of pure pixels. As the FNR is much higher than the FPR, it can be assumed that the decrease in the FNR is responsible for the increase in the OA and therefore that the OA is more dependent on the FNR than on the FPR. This would suggest that the ratio of edge pixels to pure pixels directly influences the OA.

4.2. The effect of incidence angle on waterbody classification

It can be observed from Figure 4 that steeper incidence angles performed more effectively across all size classes. Incidence angle plays a role in vegetation penetration, with steeper incidence angles penetrating vegetation more effectively (Lang, Townsend & Kasischke, 2008). Backscattered intensity from vegetation decreases as incidence angle increases since energy is lost from double bounce and volume scattering (Brisco *et al.*, 1992) which would also cause a greater contrast between water and vegetation for steeper incidence angles (Xu *et al.*, 2019). This would enable the classifier to better separate waterbodies from the surrounding LC classes. Shallow incidence angles also produced more false positives, often in lushly cropped wheat or canola fields. This is due to the lower backscatter experienced in vegetation when compared to that produced under steeper incidence angles.

The difference between the steep and shallow incidence angles becomes even more apparent for the L and XL classes. This could be explained by the presence of waves on larger waterbodies (Pôssa & Maillard, 2018). Large waterbodies are often influenced by wind-induced waves and these waves cause the backscatter to be higher than expected (Hahmann & Wessel, 2010). Since pulses are more likely to be reflected away from the sensor than back to it, steeper incidence angles are not as affected by waves as shallower incidence angles are (Martinis & Rieke, 2015). This would explain why the classification performed on imagery taken at a steeper incidence angle detected more dams in the two largest classes (L and XL). The dam shown in Figure 6 A is situated in an opening between two mountains causing wind to funnel through this gap, and it is noted how much better the steeper incidence angle performed. These observations illustrate the effect that wind and incidence angle have on waterbody classification using SAR imagery.

Unlike other studies, the study area is large, and therefore varying incidence angles over the area would also affect the classification of waterbodies. This is because changing incidence angles would cause a variety of backscatter intensities within the LC classes (Bioresita *et al.*, 2018).

4.3. Changes in the false positive rate

The steady decrease in the FPR as the size classes increased can be attributed to the use of a pixel-based classifier. Pixel-based classifiers often produce salt-and-pepper classifications, since individual pixels are used as mapping units (Manaf *et al.*, 2015).

In addition to water, other LC classes that exhibited low backscatter values were hard, man-made surfaces, such as wide expanses of tar or cement, which cause specular reflection (Li & Wang, 2015), and densely cropped maize or wheat fields which, owing to the intense attenuation of energy from volume scattering, produce low backscatter (Brisco *et al.*, 1992).

Apart from the small, often single-pixel false positives experienced in these classes, radar shadow was also misclassified. Radar shadow is intrinsic in SAR imagery captured in areas of topographic relief. The misclassification of radar shadow as water is a common occurrence when classifying SAR imagery (López-Caloca, Tapia-Silva & Rivera, 2017; Santoro *et al.*, 2015). The amount of radar shadow in the images was limited by the masking process, yet there were still areas of radar shadow in the unmasked areas. This could have been caused by using a DEM of a lower resolution than that of the Sentinel-1 imagery for the generation of the projected local incidence angle raster.

4.4. Determining the minimum surface area for a successful detection

The L class was the smallest class where more than 70% of the dams were detected. Note in Figure 3 that most of the waterbodies with a surface area of one ha and above were detected, but that from 1.5ha, almost all the waterbodies, apart from a few anomalies, were detected. This accuracy for such small waterbodies is on par with, if not slightly higher than what has been reported in the literature. Brisco *et al.* (2009) used 12m resolution Radarsat-1 imagery to map waterbodies, but to avoid misclassifications, ignored all waterbodies with a surface area below one ha. While the present study did use slightly higher resolution imagery, a large percentage of waterbodies in the M class were detected as well, emphasising that lower resolution SAR sensors can be used to detect and map small waterbodies. High-resolution TerraSAR-X imagery has been used to classify waterbodies of up to 0.7ha to very high accuracies (Heine *et al.* 2014), yet there were no waterbodies with a surface area smaller than that in the study area of that particular study. Limited research has been done to assess the accuracy of classifications of small waterbodies using Sentinel-1 imagery. In general, the dams in the study area of this study are very small, with even the larger classes being small compared to those of other studies (Li & Wang, 2015; López-Caloca, Tapia-Silva & Rivera, 2017).

This study focused on detecting dams over a very large area, and it was therefore challenging to determine parameters and training data that suited all the different regions within the study area. This is novel, as often in the literature a single large dam or a few large dams are chosen as study areas, and then the classification boundary delineation is tested. Most significantly, this study detected more dams than what were known to be in the study area. As of November 2019, DWS's database contains only 142 dams (DWS, 2019), whereas this study detected more than 900 waterbodies in the study area.

4.5. Limitations of this study

One drawback of SAR imagery is its lack of spectral resolution, making it difficult to separate some LC classes and waterbodies. Cement often produces the same dark return that water does and therefore often results in false positives. This is the nature of SAR imagery, and is, therefore, a limitation if only SAR is used to detect waterbodies. Topographic relief causes radar shadow, layover and foreshortening which are distorted during terrain correction. Man-made dams that fall into these geometrically distorted areas would not then be classified successfully and shadow could also be misclassified as a waterbody.

SAR is also sensitive to water level fluctuations and this study was therefore limited by the fact that high-resolution optical reference imagery, used to create water extent polygons for detection, could not be attained for the same time period as the Sentinel-1 imagery. It was found that owing to the effects of edge pixels, the accuracy assessment techniques of this study also disadvantaged the smaller classes.

5. Conclusion

Water is a scarce resource in South Africa. Thus, it is necessary that improved monitoring and management systems should be developed to ensure that the limited water supply is used most optimally. More dams are being built as a result of pressure on the agricultural sector, but, if not managed correctly, could have a negative effect on the catchment system.

This study used a pixel-based RF classifier to detect waterbodies in the Western Cape. The classifier was run on SAR imagery from six different dates and two different incidence angles. The results indicate that the steep incidence angle outperformed the shallow, especially as the surface area of the dams increased. It was determined that waterbodies with an area greater than one ha could be successfully detected, with almost 90% of dams between the sizes of one ha and 2.1ha being detected.

The findings from this study could allow for a system to be developed to easily determine, to a high level of certainty, the number of dams in the Western Cape. Knowing how many dams lie in a

given catchment allows for better management and the more effective use of the water in that catchment. Further studies could create systems to extract dam properties and characteristics from these classifications and these details could be used to create databases, and to monitor the construction of new dams. The effects of these small dams could then be further studied, specifically in the South African context.

Future research could include an initial segmentation using SAR and optical imagery, followed by classification on fused SAR and optical image features. While the aim of this study was to test the limits of using only one SAR sensor, an operational waterbody mapping system would likely benefit from a multi-modal approach. This would mitigate the salt-and-pepper classification problem, create more homogenous waterbody classifications and potentially lower the FPR. Alternately, performing a logical AND of the classifications from both steep and shallow incidence angles (steep ∩ shallow) and lowering the 30% detection rate threshold would also provide a robust classification as it would eliminate most false positives. If only one relative orbit could be used, the orbit with steeper incidence angles should be chosen. This result is the first step to creating an effective monitoring and management system using remote sensing.

6. References

Amitrano D, Ciervo F, Di Martino G, Papa MN, Iodice A, Koussoube Y, Mitidieri F, Riccio D & Ruello G 2014, 'Modeling watershed response in semiarid regions with high-resolution synthetic aperture radars', *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* vol. 7, no. 7, pp. 2732–2745.

Amitrano D, di Martino G, Iodice A, Mitidieri F, Papa MN, Riccio D & Ruello G 2014, 'Sentinel-1 for monitoring reservoirs: A performance analysis', *Remote Sensing* vol. 6, no. 11, pp. 10676–10693.

Bangira T, Alfieri SM, Menenti M & van Niekerk A 2019, 'Comparing thresholding with machine learning classifiers for mapping complex water', *Remote Sensing* vol. 11, no. 11.

Bioresita F, Puissant A, Stumpf A & Malet JP 2018, 'A method for automatic and rapid mapping of water surfaces from Sentinel-1 imagery', *Remote Sensing* vol. 10, no. 2.

Boardman J, Foster I, Rowntree K, Mighall T & Parsons T 2009, 'Small farm dams: a ticking time bomb?', *Water Wheel* vol. 8, no. 4, pp. 30–35.

Breiman L 2001, 'Random forests', *Machine Learning* vol. 45, no. 1, pp. 5–32.

Brisco B, Brown RJ, Gairns JG & Snider B 1992, 'Temporal ground-based scatterometer observations of crops in Western Canada', *Canadian Journal of Remote Sensing* vol. 18, no. 1, pp. 14–21.

Brisco B, Short N, van der Sanden JJ, Landry R & Raymond D 2009, 'A semi-automated tool for surface water mapping with RADARSAT-1', *Canadian Journal of Remote Sensing* vol. 35, no. 4, pp. 336–344.

Bugan R, Jovanovic NZ & De Clercq WP 2012, 'The water balance of a seasonal stream in the semi-arid Western Cape (South Africa)', *Water SA* vol. 38, no. 2, pp. 201–212.

Cazals C, Rapinel S, Frison PL, Bonis A, Mercier G, Mallet C, Corgne S & Rudant JP 2016, 'Mapping and characterization of hydrological dynamics in a coastal marsh using high temporal resolution Sentinel-1A images', *Remote Sensing* vol. 8, no. 7.

Claassen M 2010, 'How much water do we have', in Oelofse S & Strydom W (eds.), *A CSIR Perspective on Water in South Africa - 2010*, CSIR, Pretoria. <https://researchspace.csir.co.za/dspace/bitstream/handle/10204/5760/Wall11_2010.pdf>.

Curlander J & McDonough R 1991, *Synthetic aperture radar: systems and signal processing*, John Wiley & Sons, Inc, New York.

CWDM (Cape Winelands District Municipality) 2017, *Climate Change Adaption Summary Report*.

DWAF (Department of Water Affairs and Forestry) 1986, Management of the Water Resources of the Republic of South Africa.

DWS (Department of Water & Sanitation) 2019, List of Registered Dams November 2019, viewed 8 December 2021, <https://www.dws.gov.za/DSO/Documents/Copy of List of Registered Dams November 2019.xls>.

Fisher P 1997, 'The pixel: A snare and a delusion', *International Journal of Remote Sensing* vol. 18, no. 3.

Friedrich E, Pillay S & Buckley CA 2009, 'Carbon footprint analysis for increasing water supply and sanitation in South Africa: a case study', *Journal of Cleaner Production* vol. 17, no. 1, pp. 1–12.

GEE (Google Earth Engine) 2019, 'Sentinel-1 SAR GRD: C-band synthetic aperture radar ground range detected, log scaling', viewed 13 August 2019, <https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S1_GRD>.

Habets F, Molénat J, Carluer N, Douez O & Leenhardt D 2018, 'The cumulative impacts of small reservoirs on hydrology: A review', *Science of the Total Environment* vol. 643, pp. 850–867.

Hahmann T & Wessel B 2010, *Surface water body detection in high-resolution TerraSAR-X data using Active Contour Models,* Proceedings of 8th European Conference on Synthetic Aperture Radar. Aachen. VDE., pp. 1–4.

Halpern ABW & Meadows ME 2013, 'Fifty years of land use change in the Swartland, Western Cape, South Africa: characteristics, causes and consequences', *South African Geographical Journal* vol. 95, no. 1, pp. 38–49.

Heine I, Francke T, Rogass C, Medeiros PHA, Bronstert A & Foerster S 2014, 'Monitoring seasonal changes in the water surface areas of reservoirs using TerraSAR-X time series data in semiarid north eastern Brazil', *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* vol. 7, no. 8, pp. 3190–3199.

Henry J -B., Chastanet P, Fellah K & Desnos Y -L. 2006, 'Envisat multi-polarized ASAR data for flood mapping', *International Journal of Remote Sensing* vol. 27, no. 10, pp. 1921–1929.

Lang MW, Townsend PA & Kasischke ES 2008, 'Influence of incidence angle on detecting flooded forests using C-HH synthetic aperture radar data,' *Remote Sensing of Environment* vol. 112, no. 10, pp. 3898–3907.

De Lange W 2010, 'The water situation in South Africa: Some inconvenient truths', In Oelofse S & Strydom W (eds.), *A CSIR Perspective on Water in South Africa - 2010*, CSIR, Pretoria. <https://researchspace.csir.co.za/dspace/bitstream/handle/10204/5760/Wall11_2010.pdf>.

Janse van Rensburg G & Kemp J 2022, 'The Use of C-Band and X-Band SAR with Machine Learning for Detecting Small-Scale Mining', *Remote Sensing* vol. 14, no 4.

Lettenmaier DP 2007, 'Measuring surface water from space', *Reviews of Geophysics* vol. 45, no. 2, pp. 1–24.

Li J & Wang S 2015, 'An automatic method for mapping inland surface waterbodies with Radarsat-2 imagery', *International Journal of Remote Sensing* vol. 36, no. 5, pp. 1367–1384.

Liebe J, van de Giesen N, Andreini M, Steenhuis TS & Walter MT 2009, 'Suitability and limitations of ENVISAT ASAR for monitoring small reservoirs in a semiarid area', *IEEE Transactions on Geoscience and Remote Sensing* vol. 47, no. 5, pp. 1536–1547.

Long S, Fatoyinbo TE & Policelli F 2014, 'Flood extent mapping for Namibia using change detection and thresholding with SAR', *Environmental Research Letters* vol. 9, no. 3. [online]. Available from: https://iopscience.iop.org/article/10.1088/1748-9326/9/3/035002/pdf

López-Caloca A, Tapia-Silva F & Rivera G 2017, 'Sentinel-1 satellite data as a tool for monitoring inundation areas near urban areas in the Mexican Tropical Wet', *Water Challenges of an Urbanizing World*.

Mahdavi S, Salehi B, Granger J, Amani M, Brisco B & Huang W 2018, 'Remote sensing for wetland classification: a comprehensive review', *GIScience and Remote Sensing* vol. 55, no. 5, pp. 623–658.

Manaf SA, Mustapha N, Zulhaidi H, Shafri M, Sulaiman N & Husin NA 2015, *Comparison between pixel-based and object-based classifications using radar satellite image in extracting massive flood extent at northern region of peninsular Malaysia*, Proceedings of 5th International Conference on Computing and Informatics, Istanbul.

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

McCandless S & Jackson C 2004, 'Principles of synthetic aperture radar', In *SAR Marine User's Manual*, pp. 1–23. National Oceanic and Atmospheric Administration, Washington, viewed 8 December 2021, <http://www.sarusersmanual.com/ManualPDF/NOAASARManual_CH01_pg001-024.pdf>.

Moreira A, Prats-Iraola P, Younis M, Krieger G, Hajnsek I & Papathanassiou KP 2013, 'A tutorial on synthetic aperture radar', *IEEE Geoscience and Remote Sensing Magazine* vol. 1, no. 1, pp. 6–43.

Muro J, Canty M, Conradsen K, Hüttich C, Nielsen AA, Skriver H, Remy F, Strauch A, Thonfeld F & Menz G 2016, 'Short-term change detection in wetlands using Sentinel-1 time series', *Remote Sensing* vol. 8, no. 10, pp. 1–14.

Myburgh G & Van Niekerk A 2014, 'Impact of training set size on object-based land cover classification: A comparison of three classifiers', *International Journal of Applied Geospatial Research* vol. 5, no. 3, pp. 49–67.

du Plessis A 2017, 'Water scarcity and qther significant challenges for South Africa', In *Freshwater Challenges of South Africa and its Upper Vaal River: Current State and Outlook*, pp. 119–125. Springer International Publishing, Cham.

Pôssa ÉM & Maillard P 2018, 'Precise delineation of small water bodies from Sentinel-1 data using Support Vector Machine classification', *Canadian Journal of Remote Sensing* vol. 44, no. 3, pp. 179–190.

Prasad KA, Ottinger M, Wei C & Leinenkugel P 2019, 'Assessment of coastal aquaculture for India from Sentinel-1 SAR time series', *Remote Sensing* vol. 11, no. 3.

Santoro M, Wegmüller U, Lamarche C, Bontemps S, Defourny P & Arino O 2015, 'Strengths and weaknesses of multi-year Envisat ASAR backscatter measurements to map permanent open water bodies at global scale', *Remote Sensing of Environment* vol. 171, pp. 185–201.

Smith L 1997, 'Satellite remote sensing of river inundation area, stage, and discharge: A review', *Hydrological Processes* vol. 11, no. 10, pp. 1427–1439.

Sokolova M & Lapalme G 2009, 'A systematic analysis of performance measures for classification tasks', *Information Processing and Management* vol. 45, no. 4, pp. 427–437.

Stephens T 2010, 'Manual on small earth dams: A guide to siting, design and construction', *Food and Agriculture Organization of the United Nations*

Tingey-Holyoak JL 2010, 'Water sharing risk in agriculture: Perceptions of farm dam management accountability in Australia', *Agricultural Water Management* vol. 145, pp. 123–133.

Torres R, Snoeij P, Geudtner D, Bibby D, Davidson M, Attema E, Potin P, Rommen B, Floury N, Brown M, Navas I, Deghaye P, Duesmann B, Rosich B, Miranda N, Bruno C, Abbate ML, Croci R, Pietropaolo A,

Huchler M & Rostan F 2012, 'GMES Sentinel-1 mission', *Remote Sensing of Environment* vol. 120, pp. 9–24.

Twele A, Cao W, Plank S & Martinis S 2016, 'Sentinel-1-based flood mapping: a fully automated processing chain', *International Journal of Remote Sensing* vol. 37, no. 13, pp. 2990–3004.

WCG (Western Cape Government) 2018. *Western Cape Sustainable Water Management Plan 2017 — 2022: Towards a new norm for water resilience*.

Xu S, Qi Z, Li X & Yeh A 2019, 'Investigation of the effect of the incidence angle on land cover classification using fully polarimetric SAR images', *International Journal of Remote Sensing* vol. 40, no. 4, pp. 1576–1593.

van der Zaag P & Gupta J 2008, 'Scale issues in the governance of water storage projects', *Water Resources Research* vol. 44, no. 10, pp. 1–14.