# Enhancing understanding of wetland plant community spatial arrangement, ecotones, and functioning through Machine Learning and Probabilistic Classifiers.

**Danielle, Afton Seymour**
Stellenbosch University
**Helen, Margaret De Klerk** ( hdeklerk@sun.ac.za )
Stellenbosch University

**Research Article**

**Keywords:** wetland mapping, ecotones, probabilistic classification, machine learning, random forest, Sentinel-2 MSI

**Posted Date:** June 24th, 2023

**DOI:** https://doi.org/10.21203/rs.3.rs-2999653/v1

## Abstract

Context.

Mapping wetlands presents challenges due to the fluctuating hydroperiod conditions and difficult underfoot conditions. Mapping wetland plant communities distributions provide insights into wetland structure and function.

Objectives.

We aim to use wetland plant spectral signatures to produce maps of spatial arrangements of wetland plant community distribution and transitions (ecotones) to help understand how the study wetland functions in terms of regulating water flow and sediment trapping.

Methods.

We used binary, Machine Learning Support Vector Machine (SVM) and Random Forest (RF) algorithms to map wetland plant communities, and the Naive Bayesian Probabilistic Classifier to map wetland ecotones. Field based plant community data is collected to train the algorithms to classify the remotely sensed optical Sentinel imagery of the Du Toits River wetland embedded within the terrestrial Fynbos Mediterranean ecosystem.

Results.

The RF algorithm accurately mapped wetland plant communities (overall accuracy (OA) of 76% and kappa 0.72). Results distinguished two peat wetland vegetation types, namely *Prionium serratum* and *Psoralea pinnata*. The Probabilistic Classifier identified abrupt ecotones between 1) peat wetland vegetation groups, 2) peatland, sclerophyllous, and fynbos communities, and 3) within the sclerophyllous wetland vegetation. These correspond to a fine spatial matrix of deep channels interspersed with areas of sediment deposition (peatland) and formation of sediment islands (sclerophyllous plants) as the plants slow water run-off and regulate nutrient cycling.

Conclusions.

Remote sensing algorithms capture the spatial distribution patterns of wetland plant communities linked to function. This improved understanding of wetland ecology provides useful insight for effective wetland management and conservation strategies.

## Highlights

- Machine Learning and probabilistic classifiers provide useful insight into wetland community spatial structure and function.
- The combination of methods identifies complex spatial patterns of plant community distribution which informs understanding of the ecological functioning of this wetland
- Palmiet (peat) wetlands consist of two distinct plant communities, *Prionium serratum* and *Psoralea pinnata*
- Probabilistic map transect graphs elucidate fine-scale ecotones that link plant community structure to wetland function.
- Wetland ecotones are identified as sharp, narrow, abrupt, spatially disjunct, and patchy.

## Introduction

Wetlands are crucial for providing various ecosystem services, including carbon storage, flood control, sediment retention, regulation of stream flow, removal of phosphates and nitrates, and serving as habitats for unique fauna and flora. Additionally, they offer aesthetic benefits, such as education and research opportunities, tourism, and cultural significance (Van Deventer et al. 2019; Fischer et al. 2019; Ramsar Convention on Wetlands 2018). However, wetlands are disappearing at an alarming rate globally. For instance, peatland wetlands, which cover 3% of the earth's surface, are at risk with 50 million hectares having been drained for agriculture, infrastructure, forestry, or peat extraction (Hooijer et al. 2010). Moreover, in the Western Cape of South Africa, approximately 87% of wetlands are under threat and have been modified or degraded due to various threats including agricultural expansion, urbanization, and invasive species (Helme & Rebelo 2016). It is vital to understand wetland plant community structure and function so that appropriate conservation and management measures can be implemented to safeguard these vital ecosystems (Gallant 2015; Job et al. 2018; Rebelo, Finlayson & Nagabhatla 2009; Richards 2001; Sieben, Mtshali & Janks 2014).

Wetlands consist of various plant communities. Boundaries between these plant communities, and between wetland and terrestrial plant communities, as with boundaries between ecosystems at any scale, will seldom be exact. Ecotones are “zones of transition between adjacent ecological systems, uniquely defined by space and time scales and the strength of interactions between these systems” (Holland 1988; Holland, Whigham & Gopal 1990). Ecotones occur at various scales, from biomes to smaller transition zones that separate different ecological systems or fine scale plant communities. Ecotones are essential for various landscape functions, including the movement of resources such as nutrients, seeds, and animals, high biological diversity, speciation, high rates of primary and secondary production, and refuge areas for species under changing conditions such as climate and human-induced impacts (Kark 2005; Kark 2007; Walker et al. 2003; Williams 1996; Hufkens). They often occur along ecological gradients created by spatial shifts in elevation, soil, climate, and other environmental parameters (Kark 2005). In the case of wetlands, ecotones may occur at various points of hydrogeomorphic units (Ollis et al. 2013). Holland, Whigham, and Gopal (1990) note that wetlands, like all other ecosystems, have internal and external boundaries that separate distinct vegetation patches, with various ecological processes and transfers occurring at these boundaries or ecotones. For this study, internal wetland ecotones among vegetation communities are of interest in an alluvial fan wetland dominated by dense *Prionium serratum* stands. The wetland is embedded within a Cape Floral Kingdom

Fynbos system with alluvial channels that deposit and retain varying amounts of sediment, water, and organic material (Fischer et al. 2019; Snaddon et al. 2018).

Remote Sensing (RS) of the Earth's surface has long been recognized in ecology as a timesaving, non-labour-intensive, and consistent long-term means of monitoring ecosystems and their surrounding environment at different scales, both locally and globally (Pettorelli et al. 2017; Wang et al. 2010). RS is particularly useful where field-based techniques are limited due to difficult terrain or dangerous underfoot conditions (Buchanan et al. 2017), such as deep channels often concealed, and/or with fast-flowing water, typical of channelled wetlands. This study will use binary supervised machine learning, Support Vector Machine and Random Forest on optical satellite imagery to map wetland plant community classes, and probabilistic classification to map and characterise fine-scale, internal ecotones in a peat (palmiet), alluvial fan fynbos embedded wetland in the mountains of the south-Western Cape, South Africa. Results should elucidate processes forming these spatial distribution patterns and so wetland functions, such as sediment deposition, which will aid in effective conservation and restoration practices being identified and implemented.

## Materials and methods

# Study Area

The Du Toits River Wetland is an alluvial fan typically formed by sediment deposition from the channelled nearby upstream (Fischer et al. 2019, Grenfell et al. 2019). Its characteristics are shaped by a distributary network from an upstream node with areas closest to the river's origin having more concentrated sediment deposition than areas farther away. The water flow in the distant areas of the wetland is more spread out and less concentrated, which causes small channels to branch off from the main river channel and diffuse over time (Grenfell et al. 2019). The wetland starts off upstream as a channelled valley-bottom wetland (Fischer et al. 2019) and becomes weakly channelled and sometimes unchanneled in the middle course. In the lower reaches (toe) it becomes a major alluvial fan with multiple channels and tributaries that feed the fan with very fine sediment that would have had channel reforming.

The Du Toits river wetland is situated in a high rainfall (total rainfall of 1241 mm/year) and high rainfall intensity catchment (Snaddon et al. 2018). It is located within the Cape Floristic Region (CFR) consequently experiencing a Mediterranean-type climate, with wet winters and moderately intense drought-prone summers (Midgley et al. 2003; Rebelo et al. 2017; van Wilgen 1984).

This wetland is located northwest of Theewaterskloof Dam (Fig. 1) and is one of three key wetlands in the area that largely contributes to enhancing water quality entering the dam. The dam is a crucial water supply for human use, primarily for domestic and industrial, water supply into Cape Town metro as well as for irrigation more locally (Fischer et al., 2019, p.23). Most of the wetland is managed and protected by CapeNature as part of the Hottentots-Holland Nature Reserve Complex (CapeNature, 2017) and is located almost entirely within the Theewaterskloof World Heritage site (Fischer et al. 2019; Snaddon et al. 2018).

The wetland is surrounded by Elgin Shale, Hawequas Sandstone, and Kogelberg Sandstone Fynbos vegetation units (Rebelo et al. 2006), but the wetland itself is a mosaic of dominant of palmiet (*Prionium serratum*, *Psoralea pinnata)*, which is analogous to peat, and sclerophyllous wetland vegetation embedded within upland Fynbos vegetation (Sieben et al. 2017; Sieben, Mtshali & Janks 2014). The wetland is said to be in a pristine condition as there are relatively low impacts affecting the system (Fischer et al. 2019, Snaddon et al. 2018).

# Field data collection and plant community classes

The first set of field data collection took place between 21–23 October 2020 (spring) and the second took place between 7–9 June 2021 (winter) for training and verification that corresponds with the imagery date being used in classification (Wegman et al. 2016), and to cover a range of hydroperiods. Initially, transects were laid in an east-west direction across the wetland to sample the heterogeneity of the landscape. However, many of these were not accessible (water logging, deep channels, unsafe underfoot conditions, and private property). A handheld Garmin GPS (eTrex 10) was used to collect geographic coordinates at the northeast corner of each transect. Along each transect sampled, 1x1 metre quadrats were placed approximately 20m or more apart to ensure that plots represented different pixels in the 10-metre Sentinel-2 MSI imagery used in the classifications. Ideally plots should be further apart to reduce spatial autocorrelation of spectral signatures of pixels corresponding to field plots, but accessibility was very limited. Within each quadrat, each species was identified and the frequency of occurrence (%) recorded.

A preliminary plant community assessment of these data identified three wetland plant communities, *Prionium Serratum*, *Psoralea pinnata* (both are palmiet wetland types), and Sclerophyllous Wetland Vegetation (*Pteridium aquilinum, Restio paniculatus*, and *Merxmuellera cincta*), and a Fynbos class. Three additional land cover classes occurred in the study area, namely Bare soil/sandstone, Water, and Degraded. This yielded seven land cover classes with four distinct plant communities. Due to the low number of field points collected during the first set of fieldwork (Supplementary Material Appendix B, Table 1), a second set of points was gathered using simple random sampling (Wegman et al. 2016) across the full extent of the wetland in a second field visit (Supplementary Material Appendix B, Table 3). A random sampling map was created in ArcGIS 10.7.1. where field samples were collected as close as possible to these randomly generated points, given access constraints. Plant community or land cover type, as defined in round 1 of field data collection, was assigned to these points based on plant species composition and dominance. A sum of 40 samples was collected across both field sets. These points are used in the binary classifications.

# Multispectral Imagery Classification of plant communities

To investigate the distribution of plant communities across the wetland, we used binary supervised, machine learning classifiers namely Random Forest and Support Vector Machine, on multispectral Sentinel-2 MSI Level-2A (henceforth referred to as S2A) imagery for two distinct periods: winter rainfall (June 2020, July 2020, and August 2020) and summer dry period (December 2020, January 2021, and February 2021) (Van Wilgen 1984, CapeNature 2017). Imagery was sourced from Google Earth Engine (GEE), after being atmospherically corrected and surface reflectance normalized. The imagery was further processed by cloud masking, conversion to UTM zone 34S (EPSG: 32734) and clipping to the demarcated wetland area of interest. We selected bands 2 (Blue), 3 (Green), 4 (Red), and 8 (NIR) as they have the highest spectral response for vegetation (Wegman et al. 2016) (see code in Appendix A).

The Support Vector Machine (SVM) classification method (Cortes and Vapnik 1995), uses a hyperplane to maximize the separation between data points (pixels) belonging to different classes. SVM is relatively memory-efficient and has been shown to maintain high performance even with a limited number of training samples (Mountrakis et al. 2011). Previous studies have demonstrated the utility of SVM in a range of land cover analyses (Khatami et al., 2016). We implemented SVM using the R package, caret (version 6.0–90, Kuhn et al., 2021) and opted for a linear kernel (svmLinear method) as it is fast. Moreover, our preliminary analysis using the non-linear radial kernel function resulted in low accuracies. Before applying the linear kernel, we normalized the data by setting the TuneLength to select 10 random values for the cost parameter (C parameter).

Random Forest (RF) is a machine learning algorithm that fits decision trees to different subsets of training data. Once many trees are generated, the most popular tree or class is identified and classified, resulting in high accuracy (Breiman 2001; Wegman et al. 2016). RF is commonly used for satellite image classification, and several studies have reported its effectiveness (Bargiel and Herrmann 2011; Fu et al. 2017; Mellor et al. 2013; Poona et al. 2016; Xie, Sha and Yu 2008). We implemented RF in R using the randomForest package (Liaw and Wiener, 2002) where 200 decision trees were generated and four predictors per split were randomly selected. An 80/20 split of the field data was used for training and testing points for both algorithms. See Appendix B for the full description and characteristics of each land cover class, and the code for RF and SVM.

## Description of ecotones using probability classification

To investigate plant community transitions i.e. ecotones between three wetland and one fynbos plant community, we used ArcMap's Bayesian-based Class Probability (ArcMap 10.7.1. ESRI, Redlands) classifier. The Bare Soil, Open Water, and Degraded classes were excluded as they were not the focus of the ecotone analysis. We conducted analyses on the summer imagery, as it provided the most accurate results in the binary classification analyses.

For the Bayesian Class Probability classifier, training data polygons were captured based on the field collected data points used in the binary RF and SVM classifications. A total of 120 polygons were digitized and split using an 80/20 split, where 80% was used as training data and 20% as testing data, resulting in 96 training and 24 verification polygons distributed randomly for each of the four vegetation classes. Accuracy assessments i.e. overall accuracy, kappa, producer's accuracy, and user's accuracy, sensitivity, and specificity (sensu Fielding and Bell, 1997; Longley et al. 2015; Wegman et al. 2016) was performed using the caret package in R (version 4.1.1, see Appendix C for the full code) (Kuhn, 2008).

## Transects

To illustrate the nature of internal wetland ecotones, six transects were digitized and placed subjectively in such a way that each transect covered each of the four distinct vegetation types at gradients observed in field. Transects 1 & 2 were placed in the lower eastern parts of the wetland where upland Fynbos conditions were dominant. Here soils were dryer, coarser, and had less organic matter present as observed in the field. (See supplementary data Appendix D for soil profiles collected). Transects 3 & 4 focused on ecotones between peatland vegetation and sclerophyllous wetland vegetation, while transects 5 & 6 focused on the northern wetland areas dominated by peatland vegetation. Transects were 1.6 km in length and 200 m wide. Class probabilities were binned every 50 m, giving 32 readings per transect. Transects generally run from east to west, except for Transect 5 and 6 which run north-south due to the narrow size of the head of the wetland.

### Results And Discussion

## Classifying wetland plant communities:

Overall, the RF classification yielded the most accurate results i.e. 0.81 overall accuracy, and 0.78 kappa (see Table 1) for summer images, which is consistent with other vegetation mapping exercises comparing RF and SVM (Sehic and Latifi 2019, Boori, Vozenilek and Choudhary 2019, Chen, Li and Wang 2019, Lu and Weng 2007). With few exceptions, this trend also holds for user's and producer's accuracy, and sensitivity per target vegetation class. SVM on a summer image sometimes produced slightly higher results, such as a Producer's Accuracy of 90.6% for *Prionium serratum* compared to 86% for RF and Fynbos (see Table 1). Previous research has demonstrated the utility of summer images for wetland classification in winter rainfall areas (Rebelo 2017) which supports our findings. Specificity results are less conclusive. Similarly, the Class Probability map has an overall accuracy of 82.7% (Table 2) which is a moderately good accuracy. The high producer’s and user’s accuracy for *Prionium serratum* and *Psoralea pinnata* for the binary classifiers (Table 1) and Class Probability classification (Table 2) indicate that these two peatland vegetation classes are well separated which is confirmed by Rebelo et al. 2019.

Analysis of the occurrence of plant communities in this wetland using both binary classifiers (Fig. 4 and Fig. 5) on summer and winter images and the Class Probability map (Fig. 6) showed that the head and centre of the wetland where the main channel flows are dominated by peatland vegetation (*Prionium serratum* and *Psoralea pinnata*) on inundated soils that are rich in organic material or sediment (consistent with field observations). These two vegetation types also occur sporadically southward adjacent to the middle wetland channel, with smaller plants such as *Zantedeschia aethiopica* growing underneath their dense canopies. *Prionium serratum* was also classified as occurring in patches further southwest and southeast of the wetland, likely where narrow tributary channels occur.

Table 1. Overall summary statistics of how each classifier (RF and SVM) performed using Sentinel-2: MSI, Level-2A imagery (S2A). Top scorers are in bold, and top scores and those within 2% thereof are shaded light grey.

| | | | Random Forest | | Support Vector Machine Linear | |
|---|---|---|---|---|---|---|
| | | | **S2A Winter** | **S2A Summer** | **S2A Winter** | **S2A Summer** |
| | **OA**: | | 0,76 | **0,81** | 0,69 | 0,73 |
| | **KS**: | | 0,72 | **0,78** | 0,64 | 0,68 |
| **Landcover class**: | *Prionium serratum* | User's Accuracy | 78 | **83,1** | 79,7 | 81,4 |
| | | Producer's Accuracy | 79,3 | 86 | 85,5 | **90,6** |
| | | Sensitivity | 77,9 | **83** | 79,6 | 81,3 |
| | | Specificity | 96 | 97 | 97,3 | **98,3** |
| | *Psoralea pinnata* | User's Accuracy | 81,7 | **86,7** | 78,3 | 78,3 |
| | | Producer's Accuracy | 83,1 | **85,2** | 83,9 | 81 |
| | | Sensitivity | 81,6 | **97** | 78,3 | 78,3 |
| | | Specificity | 96,7 | **97** | **97** | 96,4 |
| | Sclerophyllous Wetland Vegetation | User's Accuracy | 68,4 | **82,5** | 57,9 | 75,4 |
| | | Producer's Accuracy | 63,9 | **69,1** | 62,3 | 57,3 |
| | | Sensitivity | 68,4 | **82** | 57,9 | 75,4 |
| | | Specificity | 92,8 | **93** | **93,5** | 89,6 |
| | Fynbos | User's Accuracy | 72,3 | 70,2 | 61,7 | **74,5** |
| | | Producer's Accuracy | 29,2 | **91,7** | 50 | 77,8 |
| | | Sensitivity | 72,3 | 70 | 61,7 | **74,4** |
| | | Specificity | 95,6 | **99** | 90,1 | 96,8 |
| | Bare soil/sandstone | User's Accuracy | **80,4** | **80,4** | 57,1 | 65,2 |
| | | Producer's Accuracy | **80,43** | 75,5 | 65,3 | 63,8 |
| | | Sensitivity | **80** | **80** | 70 | 65 |
| | | Specificity | **97** | 96 | 95 | 95 |
| | Degraded | User's Accuracy | 75 | **85,4** | 69,4 | 63,3 |
| | | Producer's Accuracy | 72 | **83,7** | 61,5 | 70,5 |
| | | Sensitivity | 75 | **85** | 67 | 71 |
| | | Specificity | 96 | **97** | 94 | **97** |
| | Water | User's Accuracy | 77,6 | 75,5 | **84,2** | 70,8 |
| | | Producer's Accuracy | **86,36** | 80,4 | 79,1 | 77,3 |
| | | Sensitivity | **77,5** | **76** | 69,3 | 63,2 |
| | | Specificity | **98,1** | **97** | **97,1** | **95,9** |

OA = Overall Accuracy; KS = Kappa Statistic

High probabilities of Sclerophyllous Wetland Vegetation were observed toward the middle sections of the wetland (seen in Figs. 4,5 and 6), surrounding the *Prionium serratum* and *Psoralea pinnata* (east and west edges), and occurred toward the southern ('nearest') edge (toe) of the wetland approaching the open dam water. Additionally, Sclerophyllous Wetland Vegetation was found abutting the outer Fynbos edges as the soils dry out in the transition from peatland to upland Fynbos vegetation. At a local scale within the wetland, the Sclerophyllous Wetland group may be regarded as the transitional area from pure peat wetland conditions to the drier, sandier Fynbos conditions. These areas have a mixture of damp or sometimes dry, and sandy to sandy loam soil conditions, which are very different from those in the pure peatland areas where *Prionium serratum* and *Psoralea pinnata* are found on deeper, wetter, peat-accumulated soils. Fynbos vegetation was classified as having high probabilities toward the outer edges of the wetland boundary, with distinct soil conditions (drier, coarser, and sandier soils). However, there were instances where Fynbos was found next to the peatland communities. There was some spectral confusion between the sclerophyllous wetland vegetation and Fynbos classes, with around 50% of Fynbos reference data being misclassified as Sclerophyllous Wetland Vegetation by the Class Probability map (Table 2). Patches of Degraded and Bare soil classes (only in the binary classifications) occur almost throughout the wetland, but predominantly on the drier areas towards the eastern or western wetland edges.

Table 2
Confusion matrix with accuracy metrics for the four vegetation classes, Class Probability classifier using Sentinel-2 MSI: Level-2A, summer 2020/2021 imagery (Fig. 6).

| | | **Classified data** | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | *Prionium serratum* | *Psoralea pinnata* | Sclerophyllous Wetland Vegetation | Fynbos | **Row Totals** | **EO %** | **PA %** |
| **Reference data** | *Prionium serratum* | 98 | 2 | 4 | 0 | **104** | 5.8 | 94.2 |
| | *Psoralea pinnata* | 2 | 98 | 0 | 2 | **102** | 3.9 | 96.1 |
| | Sclerophyllous Wetland Vegetation | 0 | 0 | 44 | 7 | **51** | 13.7 | 86.3 |
| | Fynbos | 0 | 0 | 52 | 91 | **143** | 36.4 | 63.6 |
| | **Column Totals** | **100** | **100** | **100** | **100** | **400 (TP)** | | |
| | **EC %** | 2 | 2 | 56 | 9 | | | |
| | **CA %** | 98 | 98 | 44 | 91 | | | |
| | **Overall Accuracy %**: | 82.75 | | | | | | |
| | **Kappa**: | 0.77 | | | | | | |

EO = Errors of Omission; PA = Producer's Accuracy; EC = Errors of Commission; CA = User's Accuracy; TP = Total Pixel

# Ecotones between plant communities within the wetland on probabilities:

The ecotones between plant communities in this wetland were identified based on their locations and nature. Each plant community class was assigned a colour: *Prionium serratum* (red), *Psoralea pinnata* (green), Sclerophyllous Wetland Vegetation (blue), and terrestrial Fynbos (medium sand colour). Pixels presenting mixed coloration or very dark hues (sometimes black) were considered mixed pixels. The probability map (Fig. 6) shows the presence of mixed pixels with hues that vary between the designated class colours. In remote sensing literature, mixed pixels are generally considered a problem when using coarser (30m or larger pixel) resolution data as they are less sensitive to spatial complexity or heterogeneity (Rocchini 2007). However, in this study, mixed pixels in the probabilistic classification are suggested to be a mixture of classes with varying probability values of at least two vegetation types (or more), identifying an ecotone pixel where there is a transition from one vegetation type to another over the space of a pixel width (10 m). This is consistent with observations in the field. For example, in the multi-layered raster with four band layers (one band per plant community), a pixel may have a 55% probability value for the red band (*Prionium serratum*) and a 44% probability value for the green band (*Psoralea pinnata*), indicating that the pixel represents an ecotone that has high probabilities of both peatland species present, with transitions occurring sharply and abruptly from *Prionium serratum* to *Psoralea pinnata* within the 10 m resolution of Sentinel-2 pixel. The probabilistic approach supports the identification of co-occurrence of species within the frame of medium spatial resolution data, and their rapid turnover within the same geographical space over a very fine spatial scale, i.e., an ecotone pixel. Another example is a pixel that includes probability values for all four classified classes: Sclerophyllous Wetland Vegetation (blue) = 4%, *Psoralea pinnata* (green) = 43%, *Prionium serratum* (red) = 50%, and Fynbos (alpha band) = 2%. This example indicates an abrupt and rapid turnover from sclerophyllous wetland conditions into high-probability peatland conditions, and then into lower-probability Fynbos conditions.

Ecotones in this alluvial fan wetland are patchy, narrow, generally sharp, and abrupt, which leads to nonlinear behaviour, emphasizing that these transitions are ecotones in the strict sense di Castri, Hansen and Naiman 1988. The probability graphs (Fig. 7, see Appendix E for the code developed to generate the graphs) reveal how the dominance of vegetation types change along the length of the identified transects. Three types of ecotones were identified: 1) abrupt, sharp ecotones between peatland vegetation groups (Transect 5 and 6); 2) sharp, narrow ecotones under 10m between high probability peatland, sclerophyllous and fynbos communities (Transect 3 and 4); and 3) complex ecotones with both slow and rapid turnover between sclerophyllous wetland vegetation and fynbos vegetation (Transect 1 and 2).

In Transect 1 (Fig. 7a), which was located in the far south-eastern corner of the wetland and focused on sclerophyllous, and fynbos dominated areas, the graph shows high probability values for Fynbos vegetation as expected. At the western-most point of the transect, there are negligible probability values for *Prionium serratum* and *Psoralea pinnata*, followed by a few hundred meters of high probabilities of peatland occurrence at approximately 500–800 m of the transect. This abrupt transition may be due to hidden channels and tributaries, leading to spatial plant community shifts over different elevations, water levels, and peat soils. Field observations support this argument as narrow, hidden channels were observed within the wetland. The presence of *Prionium serratum* in random patches reinforces what the graphs suggest: species turnover occurs rapidly, and dominant vegetation zones transition abruptly, indicating spatial heterogeneity. Transect 2 (Fig. 7b), which occupies similar geographical space as Transect 1 shows the same complex transitions between sclerophyllous wetland vegetation and Fynbos. This may be due to the two vegetation types having similar structural traits such as growth form and architecture (Sieben, Mtshali & Janks 2014) which from remote sensing may be difficult to differentiate at medium to coarse scales of 10 m.

Transect 3 (Fig. 7c) and Transect 4 (Fig. 7d) were positioned to cover all four vegetation types but were predominantly situated in large patches of Sclerophyllous Wetland Vegetation, as noted from field observations. In Transect 3, the graph corroborates this observation, with high probability values for Sclerophyllous Wetland Vegetation and Fynbos in the western section of the transect. Moreover, the transitions in this transect reveal different patterns along the transect length, where there is a gradual shift from high *Prionium serratum* values to low values for Sclerophyllous Wetland Vegetation, followed by an abrupt and sharp transition into *Psoralea pinnata*, before gradually moving back into high probabilities of Sclerophyllous Wetland Vegetation. The graph shows an almost rippling effect, with an intricate range of sharp changes between *Prionium serratum*, Sclerophyllous Wetland Vegetation, and Fynbos occurring between 0-300 m of the transect length. Moderate occurrences of Fynbos are present around 20–400 m, and there is no occurrence of Fynbos approaching the easternmost point.

Transect 4 shows considerably low probability values for Fynbos, with a sudden surge in occurrence at around the 1500 m mark. Within this transect, there are greater instances of peatland vegetation, which undergo an abrupt transition into Sclerophyllous Wetland Vegetation for several hundred meters, before sharply transitioning back into areas of high Fynbos, interspersed with patches of *Psoralea pinnata*. This occurrence is likely due to the presence of 'sediment islands' formed between meandering alluvial channels.

The classified map and associated graphs suggest that transitions between and within wetland vegetation, i.e. *Prionium serratum* and *Psoralea pinnata*, are abrupt with high probability values for either vegetation type. This phenomenon is most prominent in the main wetland channel (northern area of the wetland in Fig. 7), where ecotone pixels occur over several pixels (mixed green and red pixels in Transect 5 and Transect 6). Transect 5 (Fig. 7e) shows an abrupt transition line from high probabilities of *Prionium serratum* to higher probabilities of *Psoralea pinnata* across the length of the transect. At approximately 1500 m, this transitions rapidly from high occurrence of *Psoralea pinnata* into abrupt high occurrence of *Prionium serratum*. When another vegetation type such as the sclerophyllous class intersects (at approximately 100–200 m), a sudden island of sclerophyllous conditions is present, followed by very high probabilities of *Psoralea pinnata*. This may indicate that wetland species such as *Priounium serratum* and *Psoralea pinnata* are clustering wetland vegetation types, which often cause monodomination in a system (Gallant 2015) and spatially compete with smaller, finer wetland vegetation species.

Soils in the peatland areas are distinctly different, with a layer of damp, accumulated peat generally forming due to decaying animal and plant matter (Job 2014), which may account for the low probability value of Fynbos vegetation within this transect. Upland Fynbos and terrestrial species may not survive in these permanently wet and peat conditions. Transect 6 (Fig. 7f), occurring in the same geographical space as Transect 5 shows similar ecotonal conditions for the peatland vegetation group. Here, the graph shows relatively high probabilities for both peatland vegetation types interchanging across the transect, with low occurrences of sclerophyllous wetland vegetation. These interchanging inferences may be related to the main channel being deepest in Transect 5, possibly resulting in higher erosion control, sediment trapping, and increased accumulation of peat than in adjacent areas. Together these transitions help us understand spatial interactions of wetland plant communities, and we can gain insights into how they contribute to important ecosystem functions such as silt deposition and soil formation. There are also hidden channels where peatland communities 'show up' as patches of peatland plant communities between sclerophyllous areas. This further highlights the relationship between the hydrological patterns, specifically the flow of the river, and the formation and changes in these ecosystem functions. For some wetlands, it can be challenging to distinguish where one patch ends and another begins, while for others, it is more apparent (Holland, Whigham & Gopal 1990). In the case of the Du Toits River wetland, internal ecotones are intricate but distinguishable, as shown by the probability graphs and the mixed pixels. *Prionium serratum* and *Psoralea pinnata* have distinct spectral properties that differentiate them from other vegetation, enabling a clear distinction between peatland vegetation, sclerophyllous wetland vegetation, and fynbos communities. Holland, Whigham, and Gopal (1990) refer to these types of transitions as wetland-wetland ecotones, where surficial or diffuse flow transfers across vegetation zones with each zone dominated by a specific species. This study suggests that ecotonal areas in the wetland may have distinct hydrological and sedimentary properties due to their varying probabilities of comprising at least two vegetation types. This contrasts with areas with low species diversity that are dominated by one vegetation community. The different hydrological and sedimentary properties of ecotonal areas may affect the ecosystem services provided by this peatland system, such as water flow regulation (e.g., storage and flood attenuation), climate regulation (e.g., carbon storage, energy exchange), and water quality regulation (e.g., retention/removal of excess nutrients or pollutants, and biogeochemical transformations) (Rebelo et al. 2019). Flood attenuation and sediment trapping (and peat accumulation) properties may be entirely different in peatland vegetated areas, due to the extensive root systems of *Prionium serratum*, than in sclerophyllous and Fynbos areas, where smaller and finer plants belonging to these communities may not efficiently attenuate flows or trap sediment. Therefore, these results highlight the importance of understanding wetland-wetland (i.e. internal) ecotones to ensure the preservation of the valuable ecosystem services they provide.

As hydrology is recognized as the primary driving force in wetland ecosystems, changes in hydrologic conditions can have significant impacts on both biotic and abiotic characteristics such as salinity, nutrient availability, soil anaerobiosis, and vegetation composition (Holland, Whigham & Gopal 1990; Tiner 1999). In this alluvial fan peatland, several factors including flow velocity, flow direction, and the zones of vegetation and their associated ecotones through which it flows can affect wetland ecological processes. During periods of high flooding, water may move across the ecotones between peatland, sclerophyllous, and fynbos areas, potentially altering the regulating ecosystem services such as water flow regulation, erosion control, and sediment trapping. To better understand these distinct properties, in-situ measurements of water flow and quality, sediment, and nutrient levels are necessary to confirm this hypothesis. Wetland ecotones can also act as important buffers in a landscape by regulating and reducing water flows through the wetland. They can slow down overland runoff, soak, and store rainwater to replenish the groundwater table, bind soil together, reduce soil erosion, and intercept or trap sediment and silt from land runoff, thereby filtering and purifying water flowing through the wetland (Richards 2001).

## Conclusions

Both RF and SVM with a linear kernel performed well in delineating the distribution of different wetland plant communities in the Du Toits River wetland. Probability classification map and graphs of ecotones facilitate understanding of the interplay between plant communities, relating to the wetland function. Results highlight the presence of two distinct types of peatland vegetation, namely *Prionium serratum* and *Psoralea pinnata*. Other wetland vegetation, such as the sclerophyllous group of grasses, ferns, and restios, along with Fynbos vegetation, are mainly found on the outer edge of the wetland, although patches occur sporadically on 'sediment islands' formed by silt deposition between meandering channels.

This study demonstrates that this wetland is a complex spatial mosaic of plant communities with distinct properties of species composition, species function within the wetland and thus sediment trapping resulting in differing soil composition.

# Declarations

### Ethical Approval

The Stellenbosch University Research Ethics Committee: Social, Behavioural and Education Research (REC: SBE) reviewed the project at the proposal stage before any data collection began and approved the project with project approval number GEO-BA-2020-18511. The project adhered to the approved methodology at all times.

### Competing interests

This work is based on the research supported wholly by the National Research Foundation of South Africa (Grant Numbers: 122729) under the Masters & Doctoral Bursaries - Innovation, NRF Free Standing and Scarce Skills Scholarship 2020.

### Authors' contributions

Both authors conceptualised and wrote the paper. D. Seymour performed the majority of the analyses and field work as part of her masters thesis.

### Funding

This work is based on the research supported wholly by the National Research Foundation of South Africa (Grant Numbers: 122729) under the Masters & Doctoral Bursaries - Innovation, NRF Free Standing and Scarce Skills Scholarship 2020.

### Availability of data and materials

All field data and code can be accessed in the Supplementary Material, as well as github link for scripts (HelendKlerk/WetlandMapping)

**Acknowledgements:** Thanks to Dr Suzanne Grenfell and Nancy Job for useful discussions and insight on wetlands in drylands, and Blessing Khavu for all the technical assistance in debugging of codes.

# References

1. Bargiel, D., & Herrmann, S. (2011). Multi-temporal land-cover classification of agricultural areas in two European regions with high-resolution spotlight TerraSAR-X data. Remote Sensing 3(5), 859–877.
2. Boori, M. S., Vozenilek, V., & Choudhary, K. (2019). A comparative study of support vector machine and random forest classifier for land use/land cover mapping using sentinel-2A satellite imagery. Applied Sciences 9(14), 2873.
3. Breiman, L. (2001). Random forests. Machine Learning 45, 5–32. https://link-springer-com.ez.sun.ac.za/article/10.1023/A:1010933404324
4. Buchanan, G.M., Donaldson, L.J., Pollard, J. et al. (2017). The use of remote sensing to map habitats and locate endangered species in conservation projects. In R. S. Ambasht (Ed.), Remote Sensing of Biomass: Principles and Applications (pp. 87–108). Springer International Publishing. https://doi.org/10.1007/978-3-319-38974-0_5
5. CapeNature. (2017). Hottentots-Holland Nature Reserve Complex: Protected Area Management Plan 2017–2021.
6. Chen, C., Li, X., & Wang, C. (2019). A comparison of support vector machine and random forest classifiers for land-cover classification based on airborne LiDAR and hyperspectral data. Remote Sensing 11(15), 1845.
7. Cortes, C. and Vapnik, V. (1995). Support vector machine. Machine learning 20(3), 273–297.
8. di Castri, F., Hansen, A.J., Naiman, R.J. (1988). A new look at Ecotones: Emerging International Projects on Landscape Boundaries. Biology International, 17, 1–16.
9. ESRI 2011. ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems Research
10. Fielding, A.H. & Bell, J.F. (1997). A Review of Methods for the Assessment of Prediction Errors in Conservation Presence/Absence Models. Environmental Conservation, 24, 38–49. http://dx.doi.org/10.1017/S0376892997000088
11. Fischer T., Haasbroek B., Govender M., Kotze D., Marais D., & Horn A. (2020). Economic valuation of selected wetlands in the Breede Catchment Final Report to the Department of Environmental Affairs and Development Planning.10.13140/RG.2.2.10030.66884.
12. Fu, B., Wang, Y., Campbell, A., Li, Y., Zhang, B., Yin, S., Xing, Z., & Jin, X. (2017). Comparison of object-based and pixel-based Random Forest algorithm for wetland vegetation mapping using high spatial resolution GF-1 and SAR data. Ecological Indicators, 73, 105–117. DOI: 10.1016/j.ecolind.2016.09.029
13. Gallant A.L. (2015). The challenges of remote monitoring of wetlands. Remote Sensing 7 (8), 10938–10950.
14. Grenfell, S., Grenfell, M., Ellery, W., Job, N., & Walters, D. (2019). A Genetic Geomorphic Classification System for Southern African Palustrine Wetlands: Global Implications for the Management of Wetlands in Drylands. Frontiers in Environmental Science 7 (174), 1–23. doi.org/10.3389/fenvs.2019.00174
15. Helme, N. & Rebelo, T. (2016). Ecosystem Guidelines, Renosterveld Ecosystems: Incorporating Coast and Inland Renosterveld. [online]. Available from:
16. Holland, M.M, Whigham, D.F., & Gopal, B. (1990). The Characteristics of Wetland Ecotones. In The Ecology and Management of Aquatic-Terrestrial Ecotones, edited by Naiman, R. J. and Decamps, H., 171–198. Man and the Biosphere Series. Canforoth, UK: The Parthenon Publishing Group.

17. Hooijer, A., Page, S., Canadell, J. G., Silvius, M., Kwadijk J., Wösten, H., & Jauhiainen, J. (2010). Current and future CO2 emissions from drained peatlands in Southeast Asia. Biogeosciences 7(5), 1505–1514. doi:10.5194/bg-7-1505-2010.
18. Hufkens, K., Scheunders, P., & Ceulemans, R. (2009). Ecotones in vegetation ecology: Methodologies and definitions revisited. Ecological Research 24, (5), 977–986.
19. Job, N. (2014). Geomorphic origin and dynamics of deep, peat-filled, valley bottom wetlands dominated by palmiet (Prionium serratum) – a case study based on the Goukou Wetland, Western Cape. Master's thesis (Rhodes University).
20. Job, N., Mbona, N., Dayaram, A., & Kotze, D. (2018). Guidelines for mapping wetlands in South Africa. SANBI Biodiversity Series *28*.
21. Kark, S. (2005). Ecotones and Ecological Gradients. In Meyers RA (ed) Systems, Ecological, 272–273.
22. Kark, S. (2007). Effects of Ecotones on Biodiversity. Encyclopaedia of Biodiversity, September: 1–10.
23. Khatami, R., Mountrakis, G., & Stehman, S. (2016). A meta-analysis of remote sensing research on supervised pixel-based land-cover image classification processes: General guidelines for practitioners and future research. Remote Sensing of Environment, 177, 89–100. 10.1016/j.rse.2016.02.028.
24. Kuhn, M. (2008). Building predictive models in R using the caret package. J. Stat. Softw. https://doi.org/10.18637/jss.v028.i05
25. Kuhn, M., J. Wing, S. Weston, A. Williams, C. Keefer, A. Engelhardt, T. Cooper, et al. (2021). caret: Classification and Regression Training. https://cran.r-project.org/web/packages/caret/index.html.
26. Liaw, A., & Wiener, M. (2002). "Classification and Regression by randomForest." R News, 2(3), 18–22. https://CRAN.R-project.org/doc/Rnews/.
27. Longley, P.A., Goodchild, M.F., Maguire, D.J, & Rhind, D.W. (2015). Geographic information science & systems. Fourth edition. Hoboken, NJ: Wiley.
28. Lu, D. & Weng, Q. (2007). A survey of image classification methods and techniques for improving classification performance. International Journal of Remote Sensing, 28(5), 823–870. https://doi.org/10.1080/01431160600746456.
29. Mellor, A., Haywood, A., Stone, C. & Jones, S. (2013). The performance of random forests in an operational setting for large area sclerophyll forest classification. Remote Sensing, 5(6), 2838–2856; https://doi.org/10.3390/rs5062838
30. Midgley, G.F., Hannah, L., Millar, D., Thuiller, W. & Booth, A. (2003). Developing regional and species-level assessments of climate change impacts on biodiversity in the Cape Floristic Region. Biological Conservation 112, (1–2), 87–97. Available from: [Accessed 22 May 2021].
31. Mountrakis, G., Im, J., & Ogole, C. (2011). Support vector machines in remote sensing: A review. ISPRS Journal of Photogrammetry and Remote Sensing 66, 247–259. 10.1016/j.isprsjprs.2010.11.001.
32. Ollis, D., Snaddon, K., Job, N. & Mbona, N. (2013). Classification System for Wetlands and Other Aquatic Ecosystems in South Africa. *User Manual: Inland Systems*.
33. Pettorelli, N., Nagendra, H., Rocchini, D., Rowcliffe, M., Williams, R., Ahumada, J., De Angelo, C., Atzberger, C., Boyd, D., Buchanan, G., Chauvenet, A., Disney, M., Duncan, C., Fatoyinbo, T., Fernandez, N., Haklay, M., He, K., Horning, N., Kelly, N., de Klerk, H., Liu, X., Merchant, N., Paruelo, J., Roy, H., Roy, S., Ryan, S., Sollmann, R., Swenson, J. & Wegmann, M. (2017). Remote Sensing in Ecology and Conservation: three years on. Remote Sensing in Ecology and Conservation 3(2), 53–56.
34. Poona, N.K., van Niekerk, A., Nadel, R.L., & Ismail, R. (2016). Random Forest (RF) Wrappers for Waveband Selection and Classification of Hyperspectral Data. *Applied Spectroscopy* 70(2), 322–333. Available from: http://journals.sagepub.com/doi/10.1177/0003702815620545
35. Ramsar Convention on Wetlands. (2018). Global Wetland Outlook: State of the World's Wetlands and their Services to People. Ramsar Convention on Wetlands. (2018). 88. Available from: https://static1.squarespace.com/static/5b256c78e17ba335ea89fe1f/t/5b9ffd2e0e2e7277f629eb8f/1537211739585/RAMSAR+GWO_ENGLISH_WEB.pd
36. Rebelo A.G., Boucher C., Helme N., Mucina L., and Rutherford M.C. (2006). Biomes and bioregions of southern Africa: The vegetation of South Africa, Lesotho and Swaziland. In Strelitzia, 53–219.
37. Rebelo, A.J., Morris, C., Meire, P., & Esler, K.J. (2019). Ecosystem services provided by South African palmiet wetlands: A case for investment in strategic water source areas. Ecological Indicators 101, 71–80. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1470160X18309762
38. Rebelo, A.J., Scheunders, P., Esler, K.J., & Meire, P. (2017). Detecting, mapping, and classifying wetland fragments at a landscape scale. Remote Sensing Applications: Society and Environment. https://doi.org/10.1016/j.dib.2018.08.113
39. Rebelo, A.J., Somers, B., Esler, K.J., & Meire, P. (2018). Plant functional trait data and reflectance spectra for 22 palmiet wetland species. Data in Brief, 20, 1209–1219.
40. Rebelo, L.M., Finlayson, C.M., & Nagabhatla, N. (2009). Remote sensing and GIS for wetland inventory, mapping and change analysis. Journal of Environmental Management, 90(7), 2144–2153.
41. Richards L.T. (2001). A guide to wetland identification, delineation and wetland functions., January. [online]. Available from: https://ujdigispace.uj.ac.za/handle/10210/1999
42. Rocchini, D., (2007). Effects of spatial and spectral resolution in estimating ecosystem α-diversity by satellite imagery. Remote Sensing of the Environment, 111, 423–434.
43. Sehic, J., & Latifi, H. (2019). Comparison of machine learning algorithms for classification of land cover types using Sentinel-2 imagery. Journal of Applied Remote Sensing 13(3), 034513.
44. Sieben, E.J.J., Kotze, D.C., Job, N.M., & Muasya, A.M. (2017). The sclerophyllous wetlands on quartzite substrates in South Africa: Floristic description, classification, and explanatory environmental factors. South African Journal of Botany, 113, 54–61. https://doi.org/10.1016/j.sajb.2017.07.008
45. Sieben, E.J.J, Mtshali, H., & Janks, M. (2014). *National Wetland Vegetation Database: Classification and Analysis of the Wetland Vegetation Types for Conservation Planning and Monitoring*. WRC Report No. 1980/1/14.

46. Sieben, E.J.J. (2012). Plant functional composition and ecosystem properties: the case of peatlands in South Africa. Plant Ecology 213, 5: 809–820. http://link.springer.com/10.1007/s11258-012-0043-3
47. Tiner, R. (1999). Wetland indicators: a guide to wetland identification, delineation, classification, and mapping, First edit. ed. Taylor & Francis. https://doi.org/10.5860/choice.37-1546
48. Van Deventer, H., Collins, N.B., Genthe, B., Grundling, P-L., Grundling, A., Grenfell, M., Hill, L., Impson, D., Lötter, M., Petersen, C., Smith-Adao, L.B., Snaddon, K., Tererai,, F., Van der Colff, D., & Van Rensburg, S. (2019). Chapter 4: Pressures on the Inland Aquatic Environment. In: Van Deventer et al. South African National Biodiversity Assessment 2018: Technical Report. Volume 2: Inland Aquatic (Freshwater) Component. Council for Scientific and Industrial Research (CSIR) and South African National Biodiversity Institute (SANBI): Pretoria, South Africa. CSIR report number CSIR/NRE/ECOS/IR/2019/0004/A: http://hdl.handle.net/20.500.12143/6230
49. Van Wilgen, B.W. (1984). Fire climates in the Southern and Western Cape Province and their potential use in fire control and management. South African Journal of Science, 80, 358–362.
50. Walker, S., Wilson, J.B., Steel, J.B., Rapson, G.L., Smith, B., King, W., & Cottam, Y. (2003). Properties of ecotones: Evidence from five ecotones objectively determined from a coastal vegetation gradient. Journal of Vegetation Science 14 (4), 579–590.
51. Wang, K., Franklin, S.E., Guo, X., & Cattet, M. (2010). Remote sensing of ecology, biodiversity, and conservation: A review from the perspective of remote sensing specialists. Sensors 10 (11), 9647–9667.
52. Wegman, B., Leutner, S., Dech, M., & Wegmann, M. (2016). Remote sensing and GIS for ecologists: using open-source software. Wegman M Leutner B Dech S & Wegmann M (eds). Exeter, England: Pelagic Publishing.
53. Williams, P.H. (1996). Mapping variations in the strength and breadth of biogeographic transition zones using species turnover. Proceedings of the Royal Society B: Biological Sciences 263 (370), 579–588.
54. Xie, Y., Sha, Z., & Yu, M. (2008). Remote sensing imagery in vegetation mapping: a review. Journal of Plant Ecology 1(1), 9–23.

## Figures

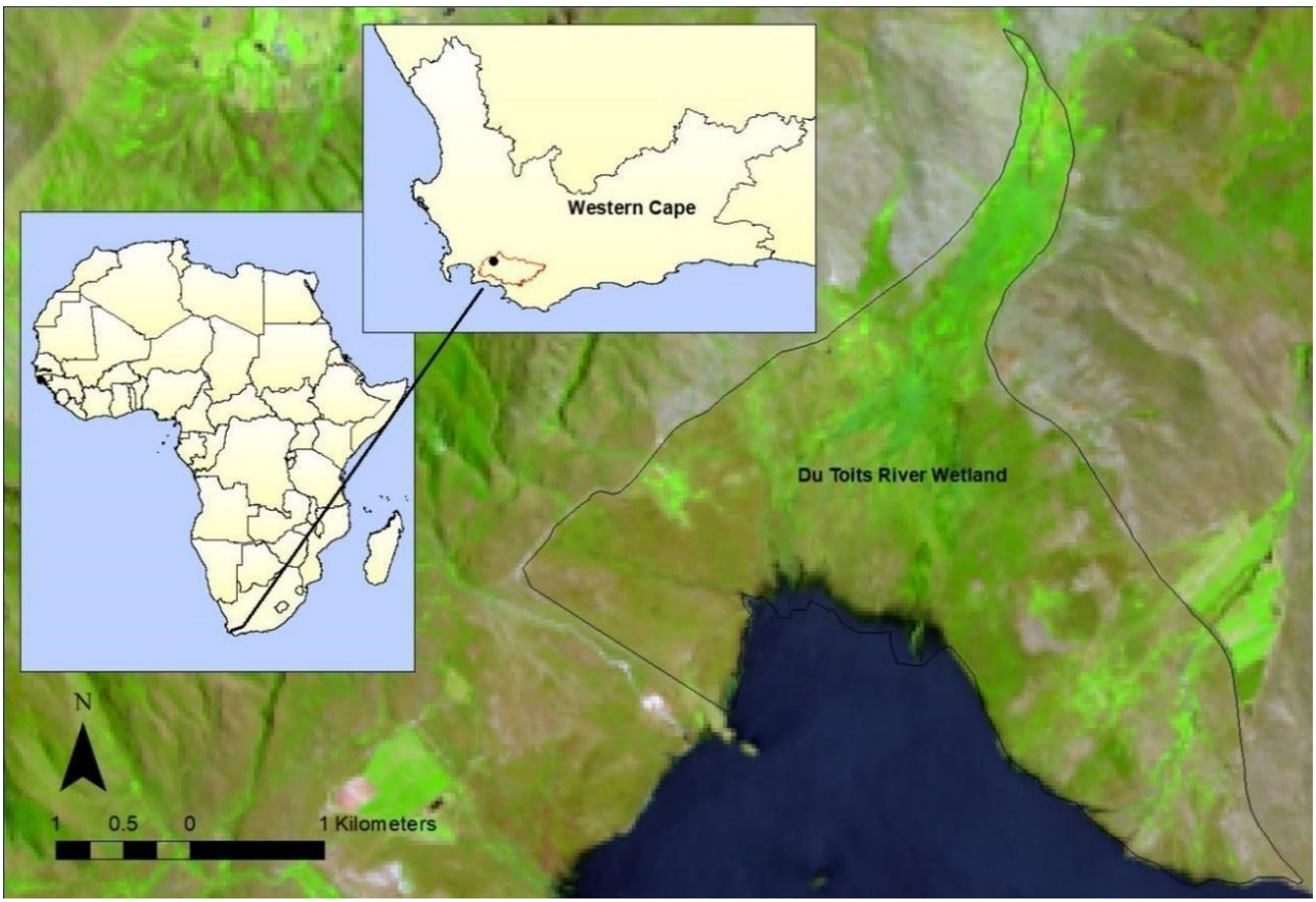



**Figure 1**

Study area of the Du Toit River Wetland (solid black line surrounding the alluvial fan wetland), with the Theewaterskloof Dam situated further south. The black dot in the map insert shows the wetland location within the Theewaterskloof catchment (displayed as a red outlined polygon) in the Western Cape province of South Africa.

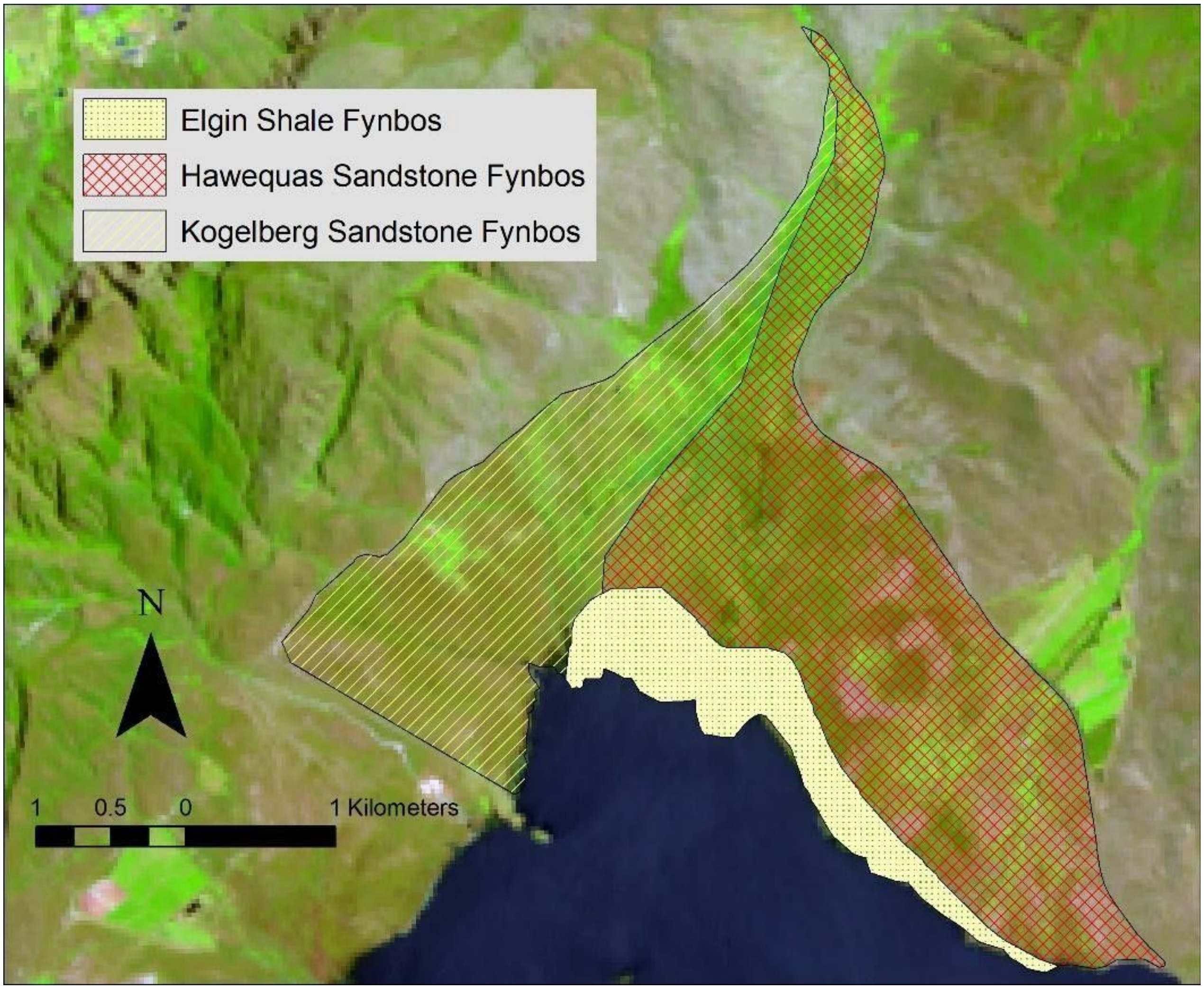
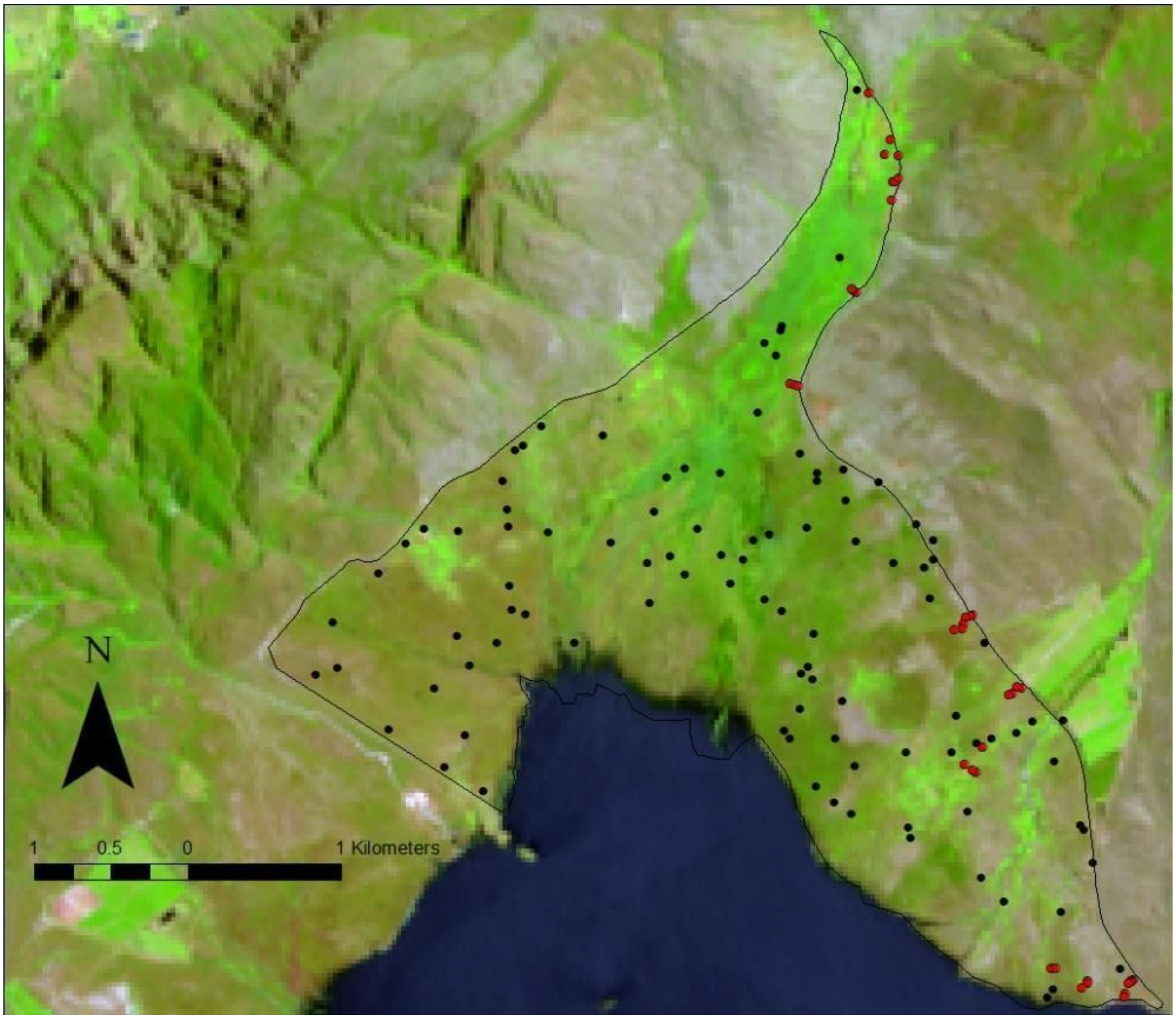


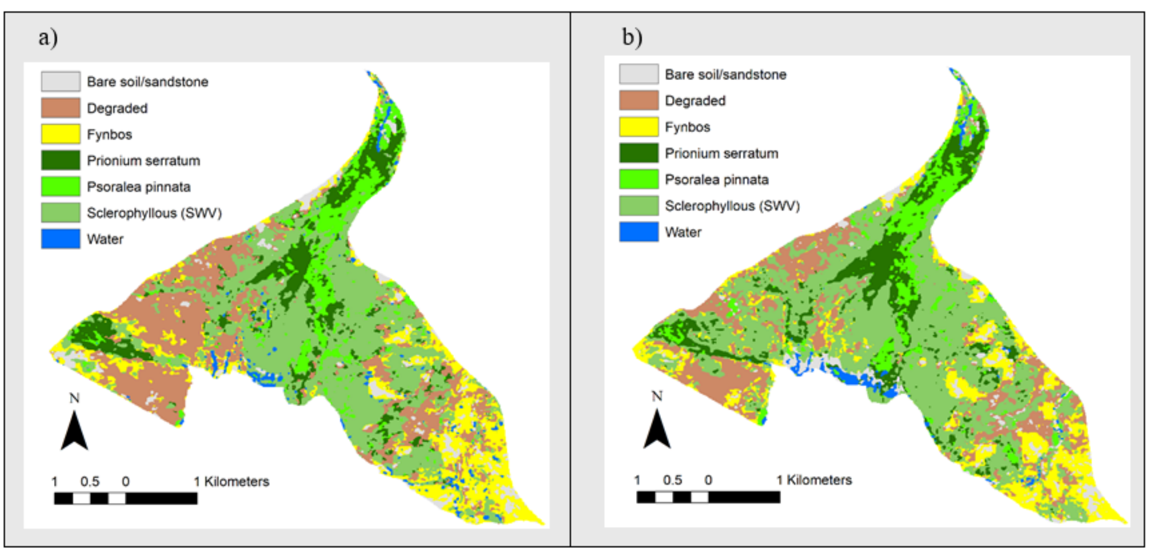

**Figure 2**

Three Fynbos Biome vegetation types that occur within the study area namely Hawequas Sandstone Fynbos, Kogelberg Sandstone Fynbos and Elgin Shale Fynbos. This Fynbos Vegetation Unit Classification is based on the Rebelo et al. 2006 National Vegetation Map for South Africa, Lesotho, and Swaziland. The data is provided as a shapefile available on the SANBI BGIS website and is the 2018 final version of the National Vegetation Map.




#### Figure 3

Samples collected in the field during October 2020 and June 2021 are presented as red dots. Random Sampling points created in ArcMap are represented as black dots.




#### Figure 4

Map outputs of the Random Forest classifier for a) winter 2020 composite and b) summer 2020/2021 composite with seven distinct landcover classes.

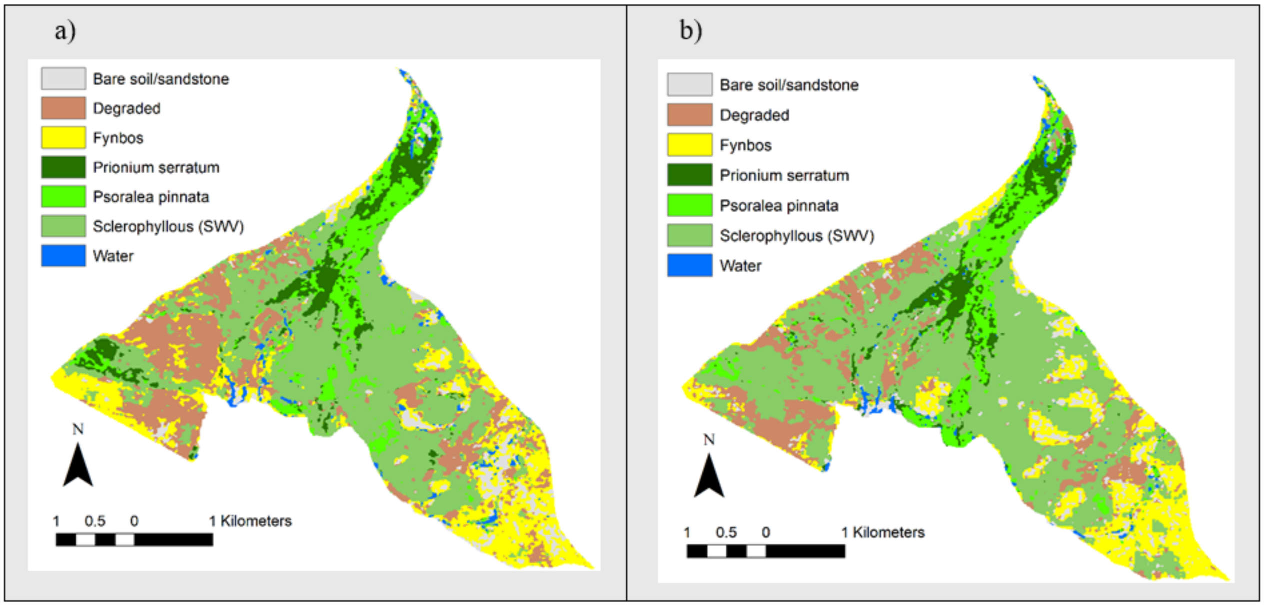



**Figure 5**

Map outputs of the SVM linear classifier for a) winter 2020 composite and b) summer 2020/2021 composite with seven distinct landcover classes.

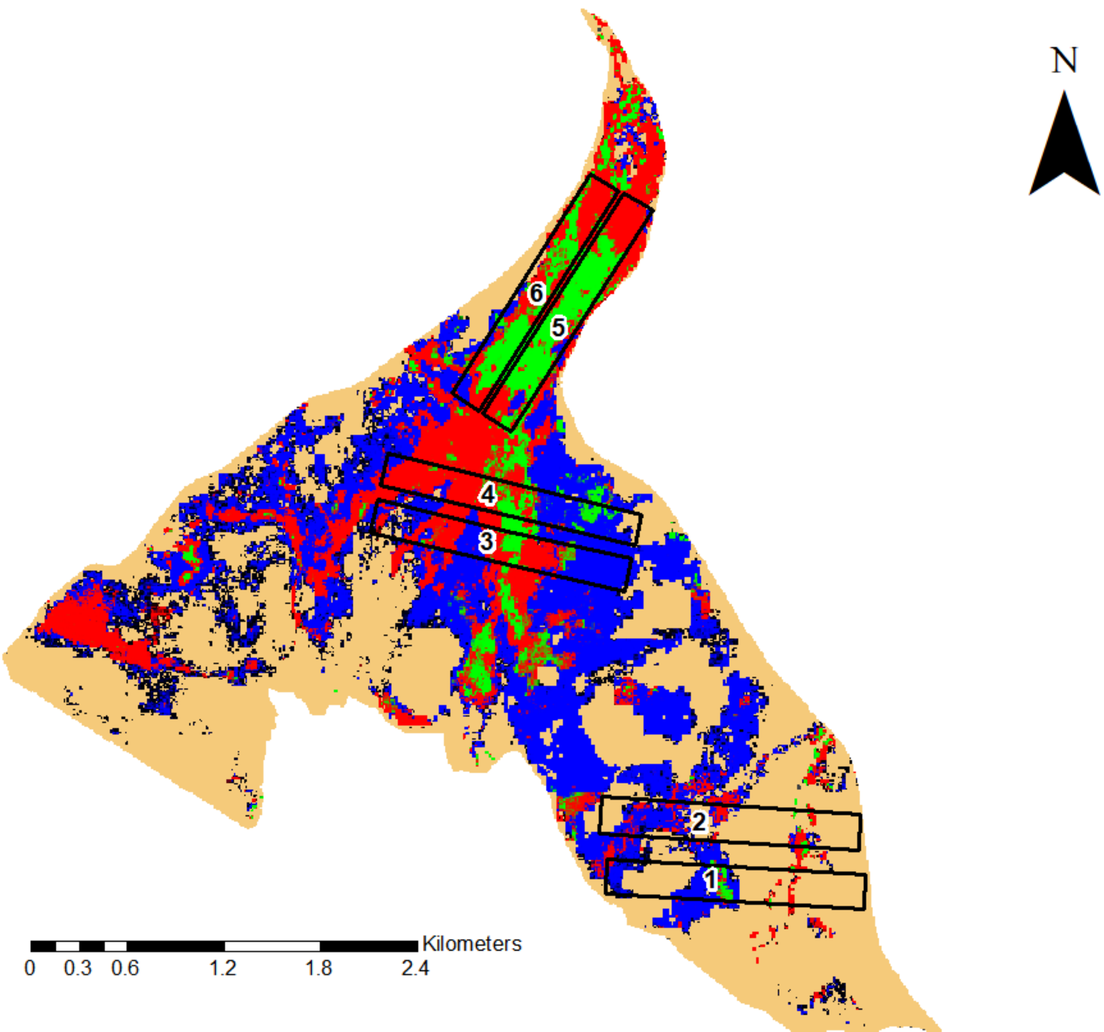



**Figure 6**

Class Probability value map (RGB) for the four distinct vegetation types in the Du Toits River Wetland i.e., *Prionium serratum* (red), *Psoralea pinna*ta (green), Sclerophyllous Wetland Vegetation (blue) and Fynbos (medium sand). Transects are outlined by black solid lines.

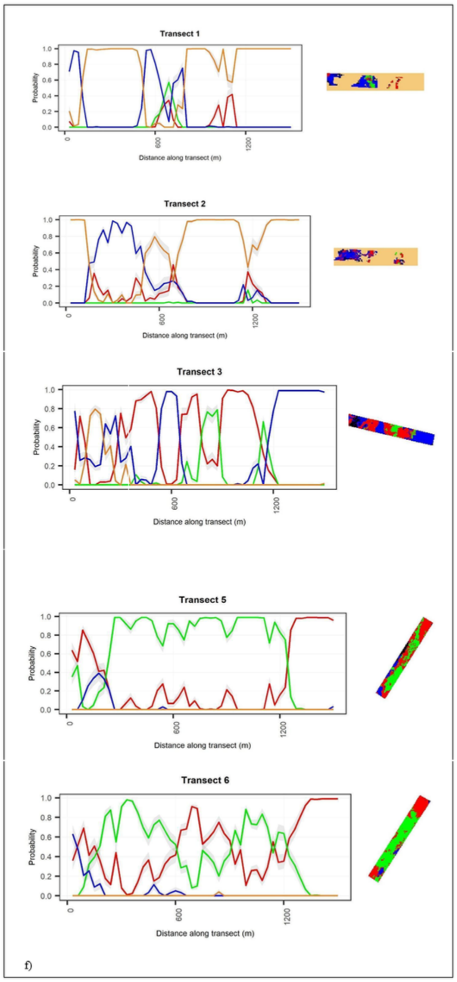



**Figure 7**

Probabilistic graphs (right) and associated map transects (left) of probability values for *Prionium serratum* (red), *Psoralea pinnata* (green), Sclerophyllous Wetland Vegetation (blue) and Fynbos (medium sand). For the graphs, class colours coincide with colours of the probability map in Figure 4.2. Values are binned at 50 m intervals over the 1600 m transect length. Ecotone pixels or 'mixed pixels' are displayed as varying saturated pixels within transects.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarymaterial20230608.docx