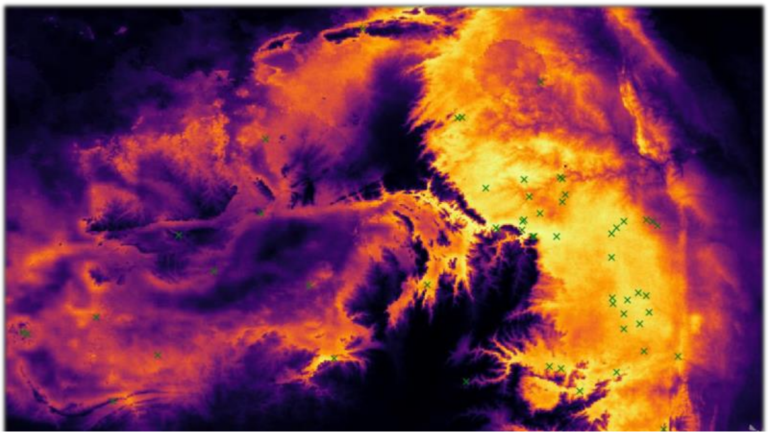

Marula distribution probability map, north-eastern South Africa

Task 2
A GIS-based approach to a resource assessment at a national (SA) and regional scale: findings, limitations and recommendations

Author(s): Ms Gillian M[c] Gregor and Mr David Kinsler
Contact details: Dept. of Geography, Rhodes University, Makhanda
Email: g.k.mcgregor@ru.ac.za
Date submitted: 9 December 2020
Project reference/contract: PN15.2014.7-018.11

This report was prepared by independent, external experts and reflects their opinions and evaluations.

For:

Deutsche Gesellschaft für
Internationale Zusammenarbeit (GIZ) GmbH
The ABS Capacity Development Initiative (ABS Compliant Biotrade in Southern Africa)

Contents

Table of Figures

Table of Tables

Acronyms and abbreviations

BIOCLIM	Bioclimatic variables
CD:NGI	Chief Directorate: National Geo-spatial Information (South Africa)
DEA	Department of Environmental Affairs (South Africa)
DEFF	Department of Environment, Forestry and Fisheries (South Africa)
GBIF	Global Biodiversity Information Facility
GSD	Ground survey distance
GIS	Geographic Information System
MCS	Multiple criteria selection
SDM	Species distribution model
SRTM…DEM	Shuttle Radar Topography Mission …. Digital Elevation Model
WWF	World Wildlife Fund

Executive summary

This report presents and demonstrates available desk-top methods, based on GIS and remote sensing for carrying out a resource assessment of a plant species at different scales, with marula (*Sclerocarya birrea* subsp*. caffra*) as the focus species. We briefly examine the literature to a) understand what GIS-based resource assessment methods have been used in the past, b) to collect information on marula to aid in understanding its autecology and, c) to provide inputs for yield estimations and development of a monitoring approach.

A national and regional species distribution model' (SDM) is produced for the probability of distribution of marula in South Africa and the region. This result is combined with GIS-based multiple criteria selection (MCS) to give a best estimate of distribution. The South African map is further refined with expert input. Although SDM has limitations, it provides a methodical approach based on readily available datasets and software which are easy to use. It further provides a standardised starting point for mapping for resource assessment for any species. We also demonstrate that rough resource yield estimations are possible, but not recommended at this scale and resolution of data. The methods for desk-top resource mapping at a local scale are promising. Availability of new software and high-resolution data open up possibilities for greatly improved mapping accuracy which could be expanded to a broader scale. This would require investment in software, data, and expertise. We also consider a basic approach to resource monitoring based on where the most species overlap and where there might be existing research centres at a national and regional scale.

In principle, the combination of desk-top methods provides a methodical approach to resource mapping at any scale, with some limitations. Desk-top analysis can only go so far in providing useful resource data as the results are simply rough estimations that require further verification, and do not provide essential information such as species density and condition. Any desk-top analysis must be verified and supported by field data.

Acknowledgements

We would like to acknowledge the valuable inputs given by Prof. Tony Palmer, Dr Tony Cunningham and Prof. Neil Crouch. Discussions were also held with Prof. Richard Cowling, Mr. Jan Vlok and Dr Nicky Allsopp, who as always were generous with their ideas. Prof. Charlie Shackleton gave expert input on the maps.

1. Introduction

The first in a list of 'Steps and standard elements of a management plan for Medicinal and Aromatic Plant utilisation' is a map-based resource inventory of population abundance and distribution (Schippmann, 1997 cited in Schippmann, 2002). Mapping of vegetation communities is a crucial part of biodiversity management and planning at local and global scales (Tierney et al., 2019) particularly with a view to the development and application of regulations for biodiversity management (Busby, 2002). As noted by the DEA (2019), one of the key actions in any biodiversity management plan is the creation of a distribution map of the species of concern, plus the location of sites of harvest, over-harvest and problem areas. The value of mapping, particularly at a local scale, is promoted by Cunningham (2002) who advocates the use of mapping approaches such as aerial photograph analysis and participatory mapping with local communities who use the resource to identify harvest sites, markets and patterns of use.

1.1 Understanding of brief

This task considers a range of suitable desk-top methods based on GIS and remote sensing which may be used for resource assessment mapping. By way of illustration it uses the case of marula, as a regional resource. It presents examples of methods used and results generated in the form of maps of marula distribution at local, national (SA) and regional scales. The analysis is conducted using readily and freely available datasets and software, promoting an approach which is accessible to any organisation involved in this kind of resource assessment.

1.2 Challenges

The task as described was substantial, and constrained by the time allocated. Examples of results from various approaches were generated, but not verified. The spatial setting is extensive: from local to cross-border scales and the nature and use of the resource varies tremendously across these scales. The resource is used and managed by a wide spectrum of stakeholders. It ranges from the harvesters on the ground through to internationally based industries, and to national and inter-national policy makers all of whom have different requirements, expectations and contributions to make in terms a resource assessment. It is difficult to meet the needs of such a wide group of stakeholders in a meaningful way. While there is much interest and research around natural resource use, there is an uneven distribution of interest and focus across the region where marula is distributed.

1.3 Outline of approach

The aim of this task is to develop and demonstrate a GIS-based approach for carrying out a resource assessment of a species (e.g.: marula) using the best available research, spatial data, and any other relevant existing materials at different scales. This task is about the principles and approaches to resource assessment rather than the actual results from applying these approaches.

To achieve the brief for this task the following steps were taken:

1. Development of desk-top methods through
    a) Collation of existing research including:
        - Other relevant resource assessments
        - Marula research
        - Spatial datasets
    b) Development of desk-top mapping and analysis approaches at a local, national and regional scale using:
        - Thematic mapping: GIS-based multiple criteria analysis (MCS)
        - Species distribution modelling (SDM) with MaxEnt
        - MCS and SDM combined with ‘expert’ input
2. A combination of objectives a) and b) for South Africa to illustrate an approach to deriving potential annual marula fruit yield.
3. Development of an approach to identifying sites for long term monitoring.
4. Concluding comments:
    - Principles for a desk-top approach to resource assessment and monitoring.
    - Limitations of the demonstrated approaches.
    - Resource assessment methodology workshop: Participant input.

2. Development of desk-top methods

2.1 Collation of existing resource assessments

A review of existing resource assessment type literature was necessary to identify and compare existing approaches with their specifics, strengths and limitations. The focus was on three completed resource assessments including: *Aloe ferox* (Palmer and Weiderman, 2020); *Pelargonium sidoides* (De Castro et al., 2010); and *Cyclopia intermedia* (McGregor, 2021).

Table 1: Resource assessments of other species from which useful approaches were identified.

Species	Methods	Comments
Pelargonium sidoides	Used quarter-degree square mapping, divided range into smaller, more manageable sites for field transects.	No desktop-based data collection. **Small plant suited to field-based survey**.
Aloe ferox	Desktop analysis with GIS, supported by **substantial field survey**. '**Super sites' for monitoring.**	Smaller range, therefore methods such as extensive field data collection possible. No climate modelling.
Cyclopia intermedia (Bergtee)	**MAXENT modelling, MCS with GIS, expert mapping,** mapping of permits. **Field surveys of populations** to get typical plant densities and population structure. **Harvest surveys** to get typical yield per plant and yield per hectare.	Species with quite specific environmental preferences and restricted distribution, formal industry with organised stakeholder community. 50-70% overlap of desktop data with field mapping.

2.2 Collation of useful material from existing marula research

To develop a GIS analysis approach, it is necessary to know the biogeographical characteristics for the species of interest (Table 2). The range and limits of these variables form the criteria which will define the distribution range of the species.

Table 2: Biogeographical variables for marula.

Biogeographical variables	Criteria/range	Source
General	Semi-arid, deciduous savannah Semi-deciduous forest, KZN seaboard Dominant keystone sp. Community dominant, 20% of woody biomass. Important in creating cooler, moister sub-canopy environment	Peters, 1988 in Shackleton et al., 2002 Johnson and Johnson, 1993 in Shackleton et al., 2002
Rainfall	200 – 1500 mm, typically, 400-1000 mm	In Shackleton et al., 2002: Shone 1979, Peters 1988, Bandeira et Al., 1999
Elevation	0 – 1600 m	Hall et al., 2002
Temperature	Frost sensitive. Preferred temperature is 19-26˚C.	Hall et al., 2002
Topography or terrain unit	Slopes and well-defined ridges preferred over valley bottoms. .	Hall et al., 2002
Soil	Sandy soils. Heavy soils are undesirable.	Hall et al., 2002
Lat/long	Min lat: 31˚S near Port Shepstone. Max lat: 10˚S in southern Tanzania.	Shackleton et al., 2002 Hall et al., 2002

Existing literature, although not necessarily explicitly about resource assessment, is a source of information about plant densities, abundance and harvest yields. The data can be tied to a location. The information is extrapolated to predict density, abundance and yields across the distribution range. It is refined to zones of low, medium and high production as it relates to, for example, rainfall (as recorded from the literature in Table 3 and extrapolated in Table 6). The resulting information can be used to map potential yield zones and the total potential yield for that resource per harvest season.

Table 3: Density and fruit yield for marula from published literature.

Locality	**MAR (mm)**	**Density (stems/ha)**	**Source (all cited in Shackleton et al., 2002. Unless otherwise noted)**	
Arid	500	16.8	Shackleton et al., 2002	
Semi-arid	670	107.5	Shackleton et al., 2002	
Mesic	>850	37.7	Shackleton et al., 2002	
Timbavati communal lands		7.5 (adult)	Shackleton, 1996	
Gottenberg communal lands, Bushbuckridge		8	Lombard et al., 2000	
			Fruit Yield	**Source**
Lebombo Mountains, Mozambique		37.5		Bandeira, et al., 1999
Cultivated yields - Israel			12yr old, 500 kg	Van Wyk and Gericke, 2000
Cultivated yields - Israel			4yr old, 27 kg	Nerd and Mizrahi, 1993
Zebdelia Estate (Tzaneen area?)			550 kg	Quin, 1959
NE Transvaal			270 kg	Shone, 1979
Botswana			550 kg	Peters, 1998
Not specified			570 kg	Roodt, 1998
Several localities			17,4 kg	Todd, 2001
North-central Namibia		4-5 females	596 kg	Botelle et al., 2002
Groblersdal, Mpumalanga			311 kg	Petje, 2008

*'No. of stems' is widely used in the literature – although this does not denote mature trees, it can be used as a measure of potential production, for the purposes of this exercise.

Other information on yield comes from Shackleton, 2002 (Table 4). He notes that: "Fruit production data for wild trees are scanty and often anecdotal," (Shackleton et al., 2002, p.30). He further notes that there is wide and inexplicable inter annual variation of fruit yields. Based on his figures and an estimate of there being approximately four to six female trees per hectare (assuming a sex ratio of 1:1) (Shackleton, 1996; Lombard et al., 2000), this gives an approximate fruit yield of 150 to 300 kg per hectare.

Table 4: From Shackleton (2002): Rainfall zones and marula yield (South Africa).

Locality (31° 0 E – 31° 35 E; 24° 30 – 25° 0 S)	**Rainfall**	**Yield**	**Notes**
Hoedspruit Nature Reserve (arid)	484mm	23.7kg	Adult trees, mean fresh mass of fruit per tree 36.8kg. Following season, almost no fruit. 20% below average rainfall in both seasons.
Wits Rural Facility (semi-arid)	651mm	55.9kg	
Bushbuckridge Nature Reserve (mesic)	870mm	34.3kg	

2.3 Identification of spatial datasets

The capability of spatial technologies and the availability of good quality spatial data has enabled the supply of information of high temporal and spatial resolution. It has enabled advances in biodiversity knowledge and management that have not been possible before (Foley, 2011; Diaz, 2019). In the context of this work this situation enables the production of accurate distribution maps at a regional to national scale with relatively low input costs in terms of software and data. However, production of local scale maps and resource assessments will inevitably be limited by the availability of spatial data and information on the nature and state of the resource.

For the national and regional GIS analysis QGIS (open source) and ArcGIS (commercial) software were used. The MaxEnt (open source) software was used for the species distribution modelling. For local scale mapping, QGIS was used for manual digitising and ArcGIS was used for image classification. For the object detection exercise, two deep learning software packages were tested: YOLOv4 on Google Colab (open source) and Picterra (commercial).

A list of commonly available data sets were identified (Table 5), with suitable resolution for national to regional scale mapping. The data layers cover themes relating mainly to the biogeographical variables.

Table 5: Available global to local scale spatial datasets suitable for mapping at regional, national (SA) and local scales.

Variable	Comment	Source	Data scale	Suitability of use
National vegetation map (SA)	450 vegetation communities that share similar biotic and abiotic features. Includes threat levels.	SANBI, 2018	1: 50 000	National and local
Protected areas (SA)	Formally protected areas of SA.	DEFF, 2020	1: 50 000	National and local
Protected areas (SADC)	Best available record of formally protected areas in the SADC.	Peace Parks Foundation, 2020	1: 50 000	International
National land cover (SA)	Based on Sentinel 2 satellite imagery, 20m resolution, classified into 73 classes according to the new gazetted land cover classification standard (SANS 19144-2)	NLC 2018, DEA	1: 50 000	National and local
Topographic features (SA)	Includes base map features: roads, rivers, contours, all man-made and natural features on the 1:50 000 map series.	CD:NGI, 2016. Provincial geodatabases.	1: 50 000	National and local
GSD 0,5m Colour aerial photography 2015 to 2019 (SA)	High resolution colour imagery (RGB) with 50cm pixel size. Also available on request as multi-spectral imagery.	CD:NGI	0,5m	National and local
Elevation		SRTM 1sec DEM	30m res.	International, national and local
Aspect		SRTM 1sec DEM	30m res.	International, national and local
Worldview2 Imagery	High resolution multi spectral satellite imagery with <50cm pixel size available for any date from 2009 on request.	DigitalGlobe (to purchase)	0,5m	International, national and local

Climatic variables	19 derived variables relating to temperature and rainfall.	BIOCLIM	250m and 500m	International and national
Species locality	A comprehensive collection of species locality records from national collections.	GBIF	Various	International, national and local
World ecosystems (incorporates landforms, vegetation and land use)	431 ecosystem units. 278 units classified as natural/semi-natural vegetation (forestlands, shrub-lands, grasslands, bare areas etc.) and 153 classes of settlements and croplands.	Sayre, R. et al. (2020)	250m res.	International and national
WWF ecoregions	867 terrestrial ecoregions, in 14 biomes. Represents original distribution of assemblages of species and communities, no anthropogenic classes.	Olson, D.M., et al. (2001)	Moderate	International and national
Climatic variables	19 derived variables relating to temperature and rainfall.	BIOCLIM	250m and 500m	International and national
Species locality	A comprehensive collection of species locality records from national collections.	GBIF	Various	International, national and local

2.4 Development of desk-top mapping and analysis approaches at a regional, national and local scale

2.4.1 Thematic mapping: GIS-based 'multiple criteria selection' (MCS)

The approach to this part of the analysis is illustrated in a cartographic model (Figure 1) which details the inputs, processes and outputs. Inputs are typically relevant spatial (raster and/or vector) and non-spatial data sets. The processes are standard GIS functions, in this case spatial join, selection by attribute', and 'intersection. The outputs take the form of maps (primarily) as well as data in the form of tables, spreadsheets and graphs.

In this example, each locality point (from the GBIF locality records) is combined with all the relevant underlying thematic data layers through a spatial join such that the most commonly occurring value for each theme is identifiable. These commonly occurring values are then used as criteria to identify other areas where marula is likely to occur based on elevation, landcover and bioregion. The overlap of areas that meet the set criteria for each of the thematic layers are then combined through an intersection to identify areas where marula is most likely to occur. The approach is applied at a national (green box) and regional (blue box) scale, using the best available data for each scenario to produce a broad scale likely distribution map of marula in South Africa (265 220 km^2) and the region (1.47 million km^2).

The map output for each of the thematic variables used for delineating the likely distribution area of marula in South Africa is shown in Figure 2 (A, B and C). Map D represents the final combined output of the processes illustrated in the cartographic model.

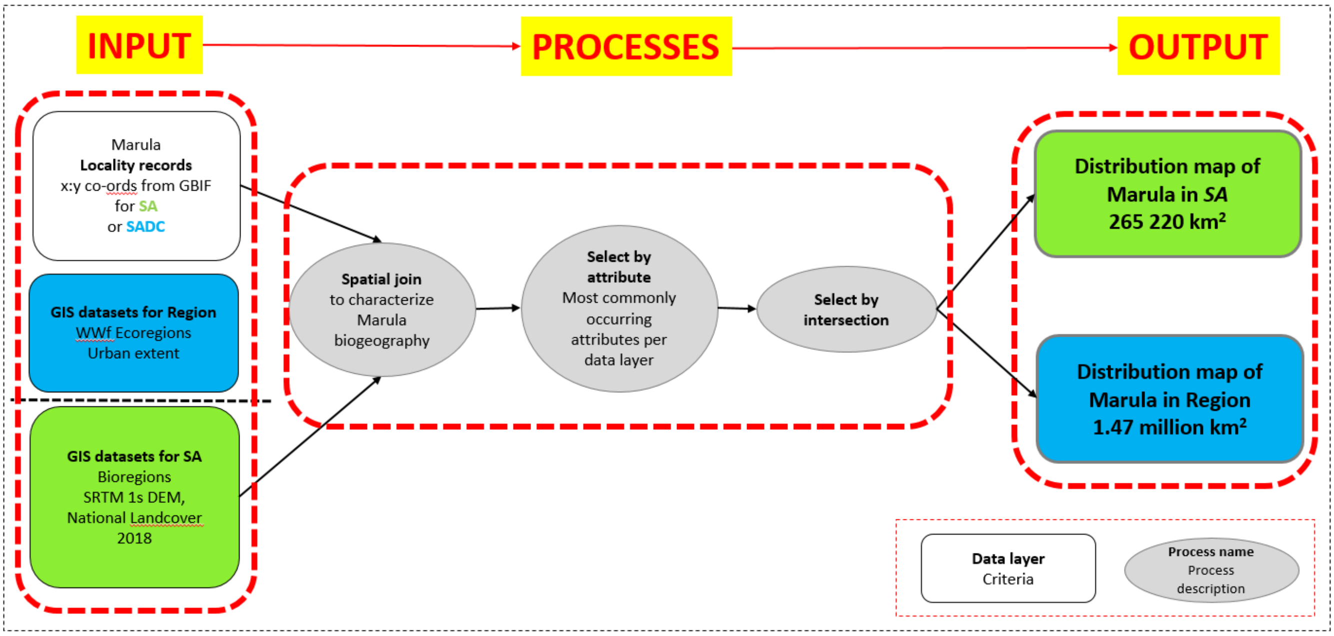

Figure 1: Cartographic model illustrating a GIS-based approach to delineating the potential distribution of marula.

.

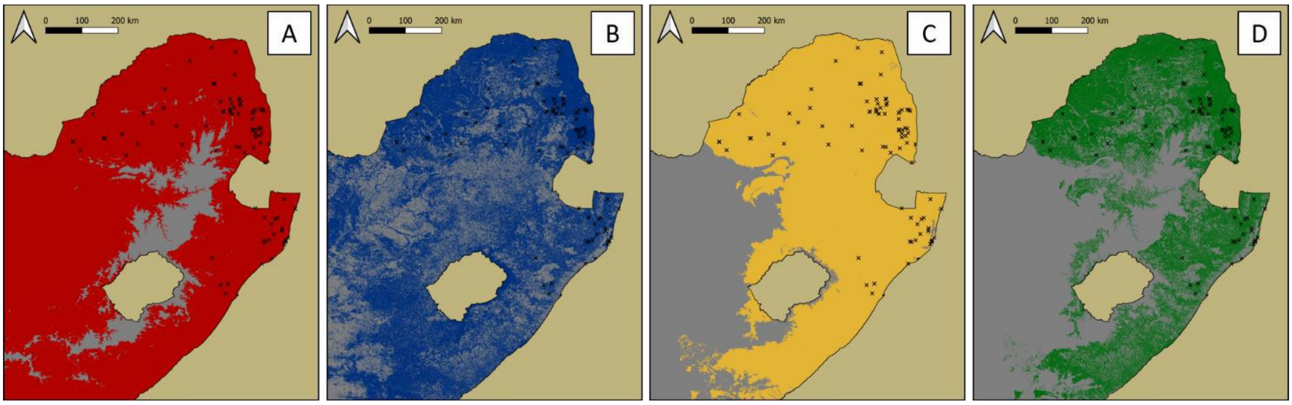

Figure 2: Map A represents all areas under 1600 metres above sea level. Maps B and C represent all national landcover classes and bioregions where the marula observations occur respectively. Map D is a combination of these three themes to show all areas where marula likely occurs.

2.4.2 Species distribution modelling (SDM) with MaxEnt

MaxEnt is a species distribution modelling software. It uses computer algorithms to correlate in situ environmental variables of known species occurrence points. It uses this to predict other areas where similar conditions are present in order to determine the potential distribution of the species. As illustrated in the cartographic model (Figure 3) for this approach, the input variables are the known localities of the species (accurate GBIF points) and the environmental variables (19 BIOCLIM derived temperature and rainfall variables). The output map shows a logistic gradation of areas where there is a very high probability that the species will occur, through to a zero likelihood of occurrence (Figure 4A).

The probability output can be converted into a binary map which can be created in various ways. Two of the most popular are the 10th percentile of training points as the binary threshold (P10), and the minimum training presence (MTP) which uses the value of the point with the lowest occurrence probability as the binary threshold. MTP is the less conservative estimation method and was chosen in this exercise as it shows the greater potential range and minimises the risk of excluding areas where marula could be present. Based on the MTP, the probability of distribution of marula in South Africa is 199 273 km^2 (Figure 4). The probability of distribution at a regional scale is 2.1 million km^2 (Figure 5). Note that these are potentially overestimates of the distribution. As indicated in Figure 4 it can include sites where marula is not known to occur (white markers), and might omit areas where the species is known to occur (red marker) as in south-western Angola in Figure 5. The model is limited in that it uses climatic variables only and therefore cannot account for non-climatic variables (e.g. anthropogenic influences, both historical and present) that may have potential influences on the distribution range.

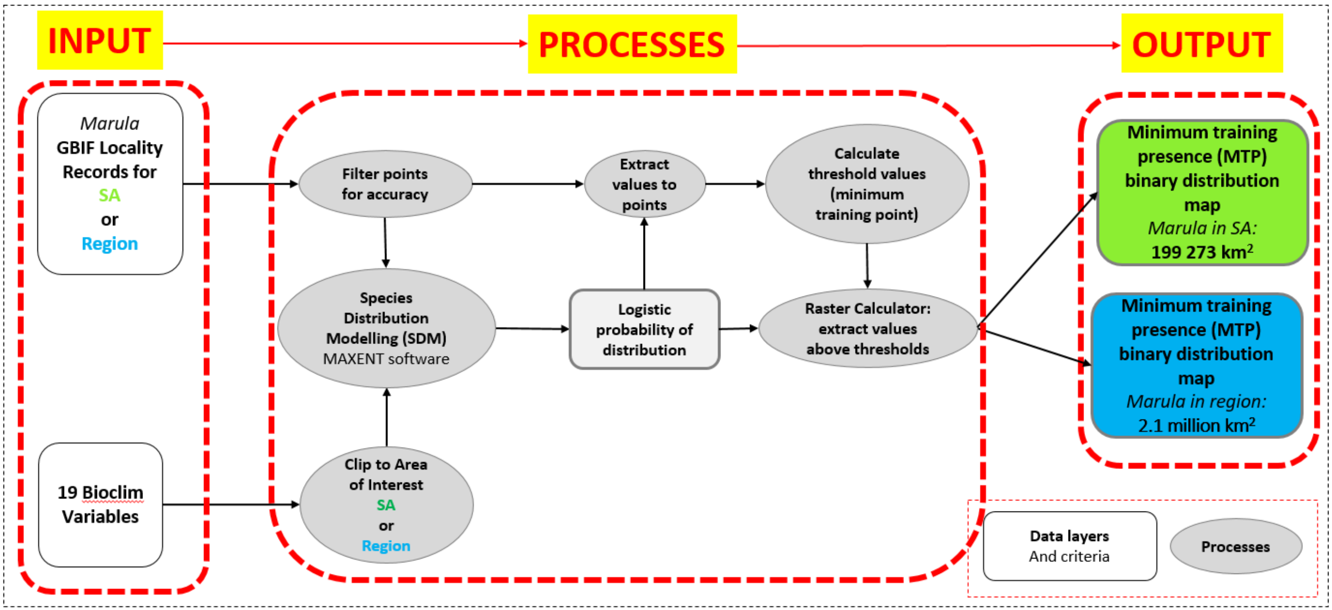

Figure 3: Illustration of typical steps in delineating the broad scale likely distribution of marula at a national and regional scale using SDM.

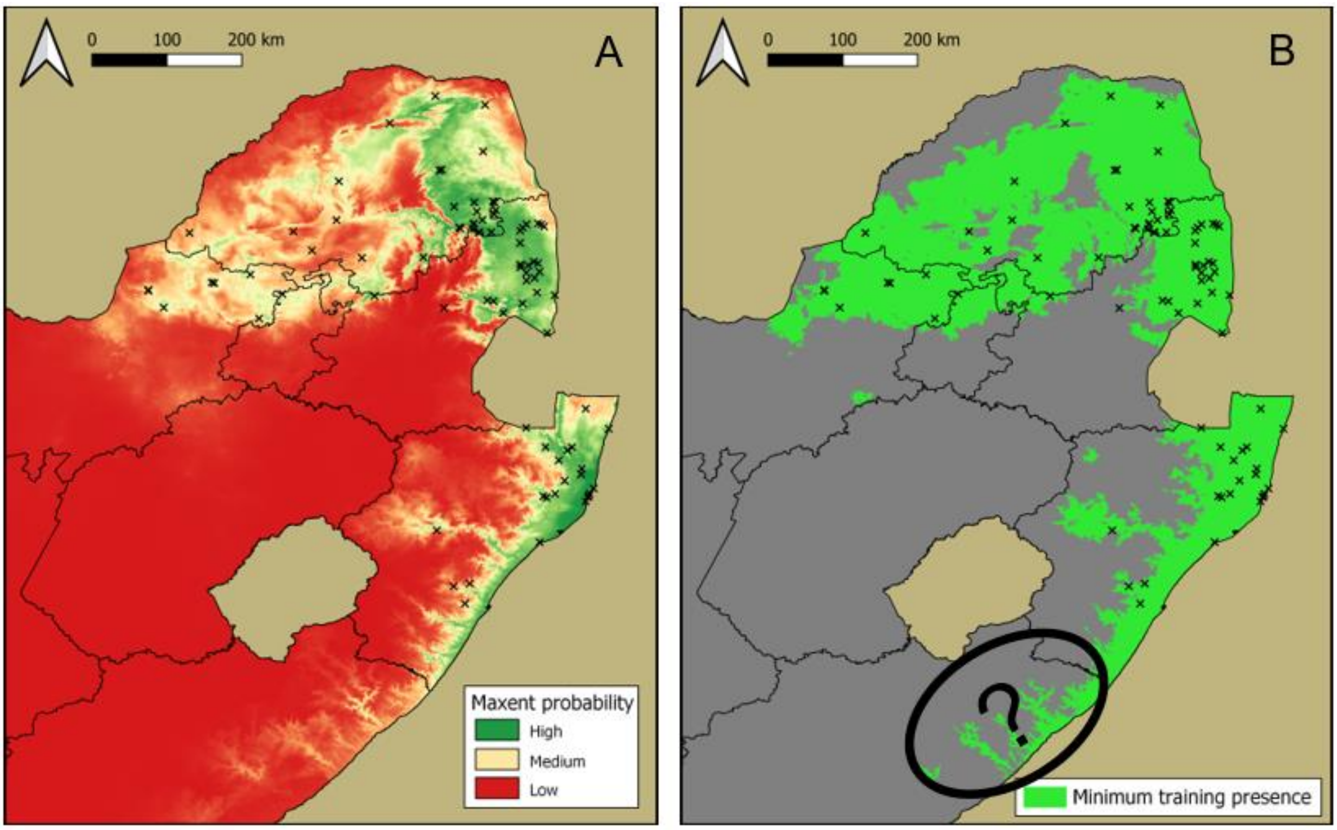

Figure 4: A - Full probability of distribution and B - Binary map showing areas where marula is likely/not likely to be present in SA.

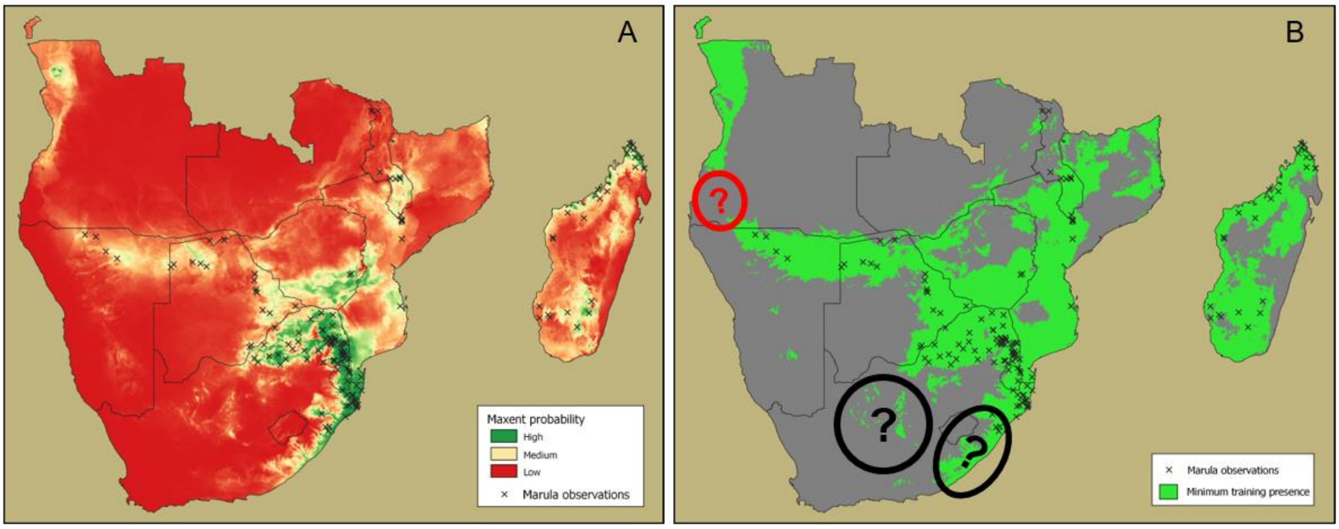

Figure 5: A - Full probability of distribution and B - Binary map showing areas where marula is likely/not likely to be present in the region.

2.4.3 MCS and SDM combined with expert input

The results of the MCS and SDM processes were combined to give a best estimate of marula distribution in South Africa and in the region. The process followed is illustrated in the cartographic model in Figure 6.

For South Africa, at this stage of the analysis, expert input was included. This is seen as an acceptable way to refine models and maps of this nature at all scales. The map of South Africa was edited according to expert input from Prof. C. Shackleton. It was generalised to limit some of the spurious precision of its estimates. As noted, the SDM performed well in most areas with the main exception being the Eastern Cape Wild Coast area, where it does not naturally occur despite modelled predictions. The map below (Figure 7A) is the result of the expert input and the combination of the thematic MCS and MaxEnt SDM's showing our best estimate of marula distribution in South Africa at 143 595 km^2. When protected areas are taken in to consideration where no harvesting can occur (Figure 7B) and excluded from the area, the likely harvestable area of marula comes out at 103 976 km^2.

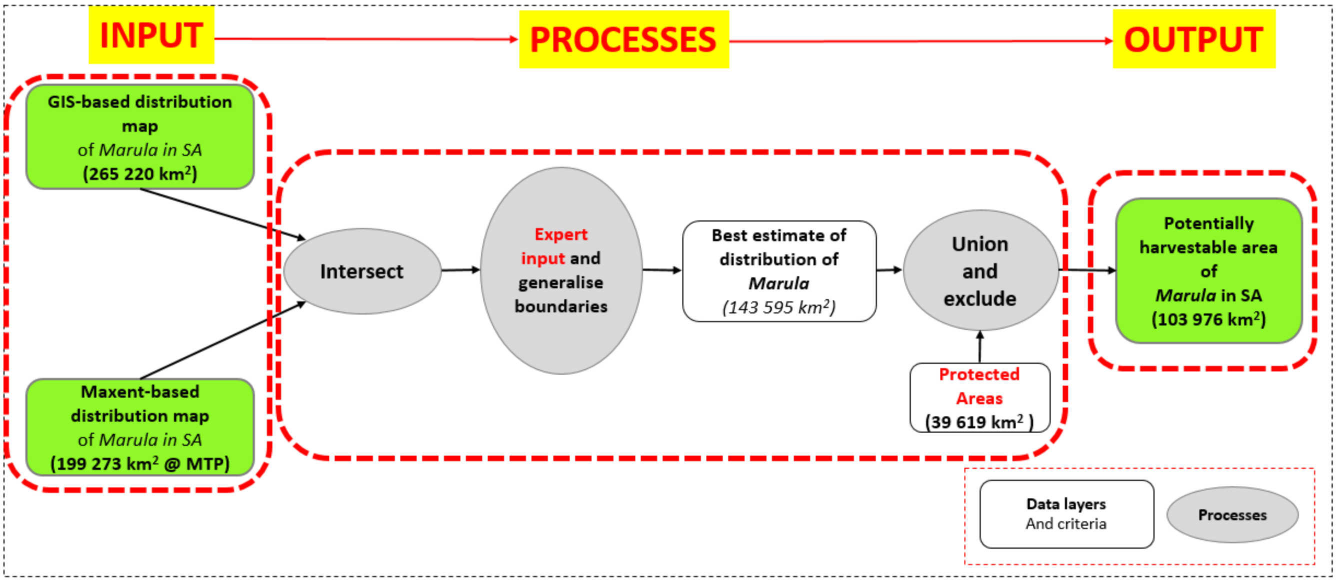

Figure 6: Combing the results of the MCS and SDM and refining with expert input and exclusion of protected areas.

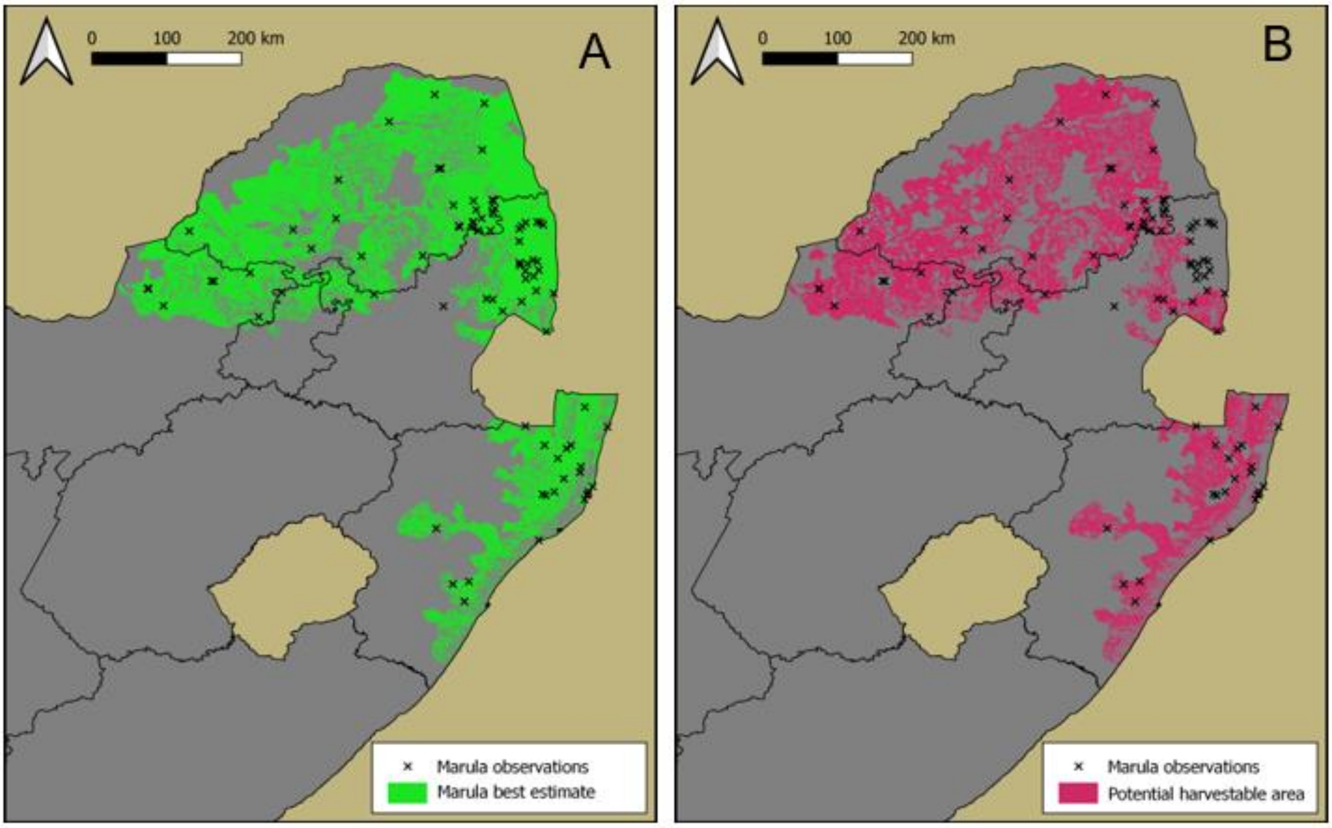

Figure 7: Best estimates of marula distribution in SA. A - Total area; B - Harvestable area (protected areas removed).

Further refinement of the harvestable area of marula was made with reference to land tenure. An explanation of possible land tenure classes is given under Section 4 of this report. The assumption was made that the harvestable area for marula in South Africa would most likely be in communal lands occurring in the marula distribution area as delineated by the former homeland boundaries. It was assumed that harvesting on private land (commercial farmland) was unlikely or at least not worth including in the estimate. State-owned land (not communal lands and not protected areas) was by necessity (data not available) combined with this private land class. Given that this discussion is about an approach to resource assessment, rather than the actual figures, the approach is potentially useful. Of the 103 976 km^2 of the potentially harvestable area of marula distribution, 47 150km^2 falls into this class of communal lands where marula is most likely to be valued as a harvestable resource. This figure is used in Section 3 for estimating the potential harvest yield of the resource.

2.4.4 Local scale mapping approaches

While regional and national resource mapping techniques share many methods and data sources, local mapping is more distinct in its approach and has its own set of specific data types and methods that are better suited to working at that scale. Outlined below are some data source options and mapping methods. Regardless of the approach used, all these methods are most accurate when combined with ground-truthing.

Data sources for local mapping

Table 7 outlines several data options for mapping at a local scale. In this exercise, samples of multispectral WorldView2 satellite imagery featuring central Kruger and north-central Namibia was obtained from DigitalGlobe to test mapping methods.

Table 7: Imagery options for local scale mapping

Type	Resolution	Advantages	Disadvantages
Google Earth Pro	High (<0.5 m)	Free	Not multi-spectral
WorldView2	High (<0.5 m)	Multi-spectral	Expensive
GSD 0,5m colour aerial photography 2015 to 2019	High (0.5 m – 0.25 m)	Free, can be multi-spectral on request	Difficult to acquire, long temporal frequency
Drone imagery	Very high (2-10 cm)	Cheaper than tasked aerial imagery, highest resolution	Legal permissions, constrained by accessibility and time
Tasked aerial imagery	High	Can use special sensors like LiDAR	Very expensive

Image classification techniques

At a local scale, image classification of a single species, and individuals within a species, is possible with high resolution imagery if the species is large enough to detect or occurs in dense communities (Geller et al., 2017). The local population structure is a key determinant of the data resolution and methods required. For example, dense and continuous plant communities can be mapped using lower resolution imagery since the need to discriminate from other species in close geographic association is less. However, in the case of marula, higher resolution imagery is required to a) identify separate individual trees in the landscape and b) detect spectral characteristics that will be able to distinguish it from other species.

Several methods of image classification were tested in this exercise on multi-spectral WorldView2 images. These included object-based image analysis segmentation, maximum likelihood classification and Iso cluster unsupervised classification. All these methods were limited either due to similar spectral signatures of marula trees and surrounding vegetation, or due to heavy shadows which made it difficult for the classifiers to determine where the canopy boundaries were.

Deep learning algorithms

Noteworthy (but unverified) results were achieved by using a deep learning model that detected marula trees from other trees in northern Namibia. Although only a scoping exercise, there is a lot of potential for this relatively new kind of geographic analysis, particularly on imagery with a resolution below $0.5m^2$. In the context of resource assessments and the availability of moderately high resolution RGB imagery, this has enormous potential for local and broader scale mapping. An example of the training set (100 images of known marula trees) for the analysis algorithm is shown in Figure 8 (left). An example of the output, identification per tree with a ‘certainty value’ is illustrated in Figure 8 (right). Notably, the algorithm only detected trees that it was trained to identify and ignored other trees in the image.

Figure 8: Examples of training set images (left) and the classifier in action (right) where it correctly identifies only marula trees in image.

The example in Figure 8 used the YOLOv4 deep learning classifier, written in the Python code language within the Google Colab environment. This method requires a relatively high level of competence and familiarity with coding and deep learning to apply it successfully to GIS. However, there are user-friendly commercial alternatives such Picterra which require minimal training and prerequisite knowledge to perform complex object detection to spatial images. Figure 9 illustrates an example of object detection performed in Picterra, which automatically detects and circles Marula trees.

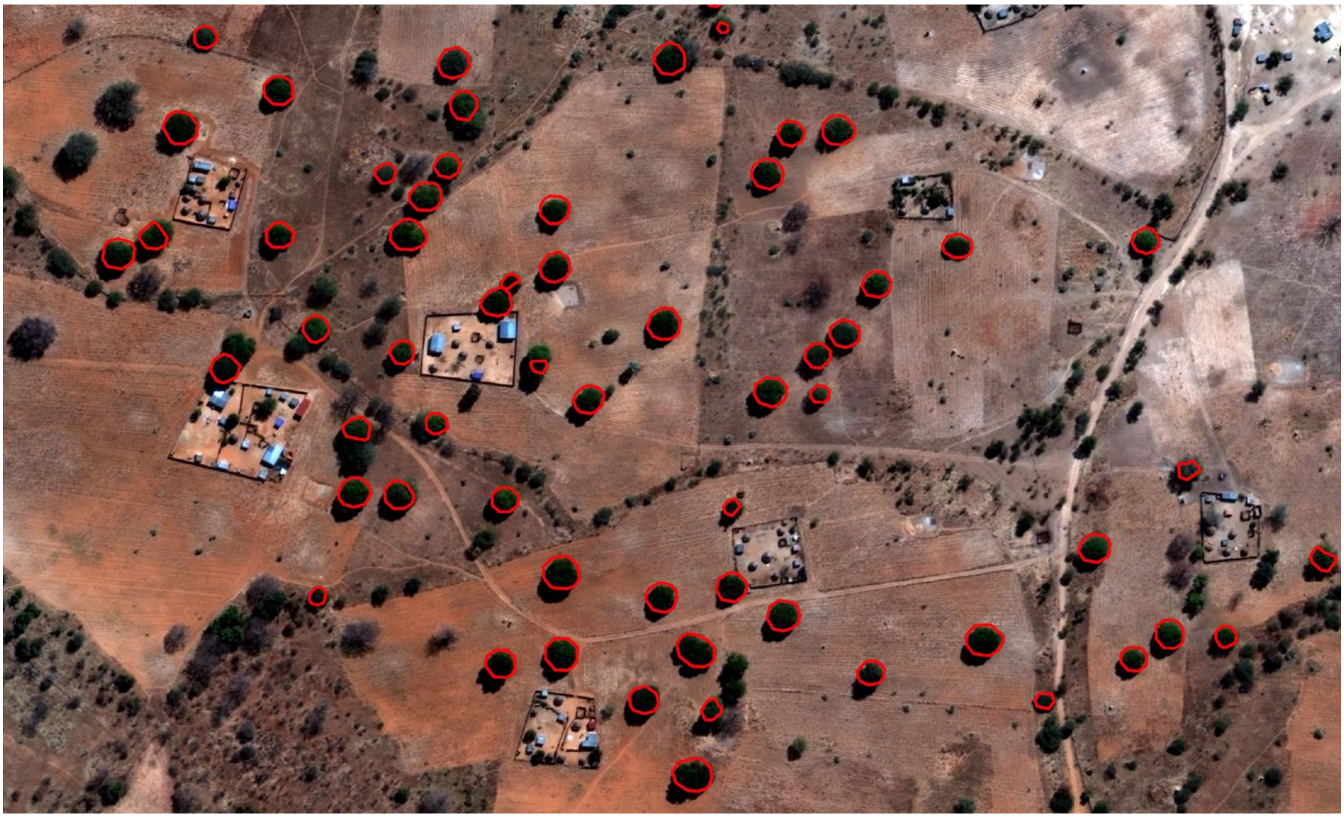

Figure 9: Automated object detection of marula trees with Picterra, a commercially available deep learning software.

Heads-up (manual) digitising
Heads-up digitising is the manual creation of vector GIS data from images by a GIS user. Despite advancements in image classification techniques, this is still one of the most reliable forms of mapping for local scale data. It is particularly useful when the image data is ambiguous and unclear to the computer, but still discernible to the human eye. In the case of marula, this is our recommended approach, as it is highly accurate and requires minimal user training or expertise to perform accurately. Examples of two kinds of digitising are shown in Figure 10. Map A represents point data, useful for calculating species density per unit area. In map B, each suspected marula tree (field verification required) has its canopy boundary drawn and classed into three size categories. This is useful for guiding fruit yield estimations, inspired by the published work of Botelle et al. (2002) who demonstrated a positive relationship between canopy size (w x h) and fruit yield.

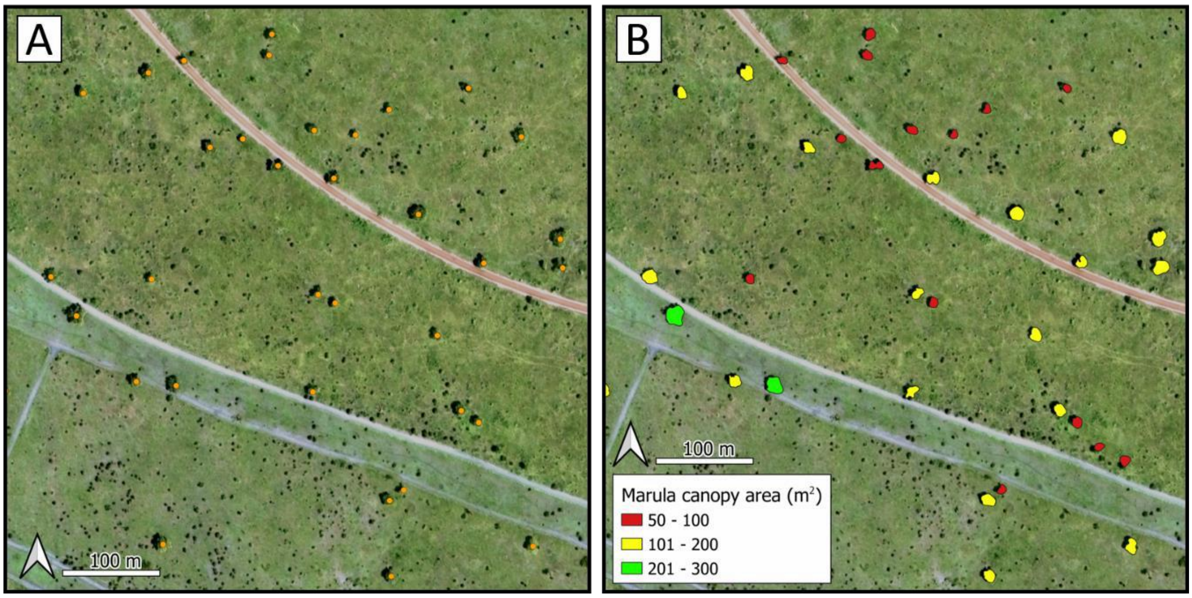

Figure 10: Examples of heads-up digitising of marula trees in an image. A – trees represented as points, and B as polygons, using WorldView2 imagery in Kruger National Park.

3. Combination of mapped results with existing harvest data to illustrate an approach to deriving 'potential harvestable volume' of marula for South Africa

Based on the limited number of resource surveys available, it is possible, but not advisable to extrapolate tree densities and yields across the area of distribution. The results of an exercise like this are illustrated in Table 6. This table presents potential marula fruit yields based on the modelled distribution in South Africa and the fruit yield and stem density data published by Shackleton (2002) and Shackleton et al. (2002) respectively. It must be noted that the value of this information is limited. It is based on the modelled distribution, not the actual distribution. In addition, it assumes variables such as constant stem density across the entire range of each rainfall zone and even male/female sex ratios. It excludes formally protected areas, many of which will not be accessible for harvesting, and state-owned and privately owned land. It therefore only considers marula distribution across areas of communal land tenure. These estimates can provide a starting place for further discussion on the potential availability of marula as a resource in South Africa.

Table 6: Extrapolated tree density and yield for non-protected areas in South Africa, in communal lands.

Rainfall zone	Criteria/range	Extrapolated stem count	Extrapolated annual fruit yield* in tonnes per annum
Arid (500 mm)	7 557 km^2 (16%)	12.7 million	150 thousand
Semi-arid (670 mm)	22 991 km^2 (49%)	247 million	6.9 million
Mesic (>850 mm)	16 602 km^2 (35%)	62.5 million	1.05 million
TOTAL	**47 150 km^2**	**332.2 million**	**8.1 million**

*assumed 1:1 male/female ratio

4. Development of an approach to identifying sites for long term monitoring

This section deals with using GIS to prepare a map layout for potential monitoring sites. Further principles for a monitoring approach for all bio-traded species are outlined in Task 1: Section 1.3.

Despite a major focus in the South African arena on land tenure, there is no complete local or national scale land tenure dataset. (Relevant information about land tenure information can be created by combining land parcel data (Chief Surveyor General) with deeds data. Since land tenure affects use and management, certain assumptions can be made about the level of utilisation and care of a resource based on land tenure. For the purposes of this exercise (illustrating a possible approach) proxies for land tenure are described below in Table 5.

4.1 Points to consider for selection of monitoring sites for marula

The land tenure classes form the zones within which monitoring sites should be chosen. The process of identifying these sites would something like:

- Take a 10% sample of existing localities and allocate a relative proportion per land class (for marula, there are 3000 existing localities across the region, therefore 300 might become monitoring sites).
- Identify a number of super sites that cover all four classes: the site would always border on a protected area and be spread across classes 2 and 3. (This can be done with a GIS query).
- Possibly use ¼ degree squares, which are about 25x25km with *n* number of existing locality points as the border of a super site.
- Identify sites near existing areas of research with records (for marula, a good example would be the Wits Rural Research facility).

- In South Africa, align potential monitoring sites with existing SAEON LTER (Long Term Ecological Research) or EFTEON (Expanded Freshwater and Terrestrial Environmental Observation Network) sites.

Table 5: Possible land tenure classes for marula.

Classes	Description	Level of care
Class 1: Formally protected areas (PA)	This is the best protection class and represents a control site. It is a natural system, but has come under pressure with increasing elephant densities in many reserves. This is state-owned land managed by a conservation authority. Updated data is available on a quarterly basis (DEFF).	High
Class 2: Communal lands	In these areas marula are found in a semi-cultivated setting. Trees are valued and protected for cultural and resource use purposes. There is a preference for fruit-bearing female trees such that male trees are commonly removed, leading to an imbalance in the population (Male 20%, female 80%). In South Africa, former homeland areas are almost completely communal land and serve as a proxy for communal land in this exercise. This land is actually state-owned but is managed by/for the people by traditional authorities.	High
Class 3a: Privately-owned lands.	Land would be mostly commercial farms for stock and cultivation. Saplings are browsed heavily by stock, and trees are removed from lands by mechanical cultivation. Level of utilisation would be zero, and trees would be heavily impacted as they have little value.	Low
Class 3b: State lands	This would include land owned by for example: Transnet, Safcol, Eskom. There would likely be poor control/management of these lands. Some land would be commonage owned by local authorities. Some of it would be land owned by the state and allocated for agricultural development under various schemes. Ideally these kind of state lands should be separated from private lands, but for the purposes of this exercise these 'state lands' area treated as areas where marula is not utilised and is not specifically cared for.	Low

In terms of identifying potential sites for monitoring, the results of an exercise which overlays all 11 bio-traded species localities on ¼ degree squares of South Africa with existing SAEON LTER sites is shown in Figure 11. The higher the 'sum' of the number of species occurring in a ¼ degree square, the higher its potential as a monitoring site. In reality, there is only overlap of five of the 11 species. A similar approach is demonstrated for marula only on a regional scale: in this example (Figure 12), the number of marula observations per ½ degree square are used. Major town localities and protected areas are shown for reference and to give an idea of where the most suitable localities might be found.

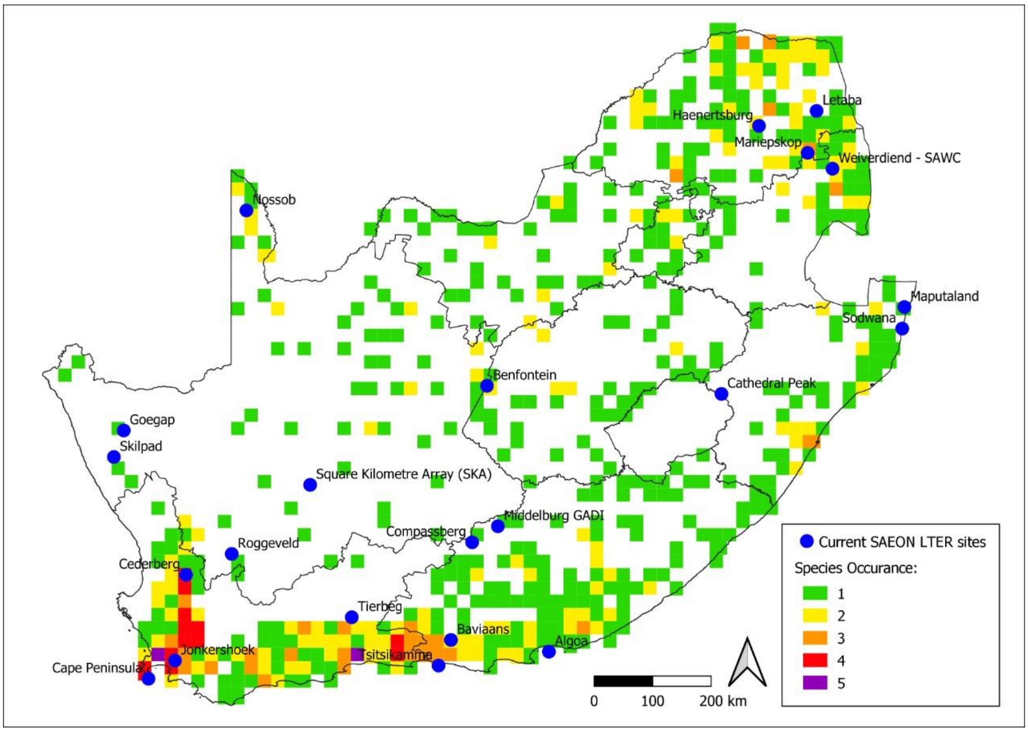

Figure 11: Illustration of a potential approach to identifying sites for monitoring of bio-trade species (including marula) in SA.

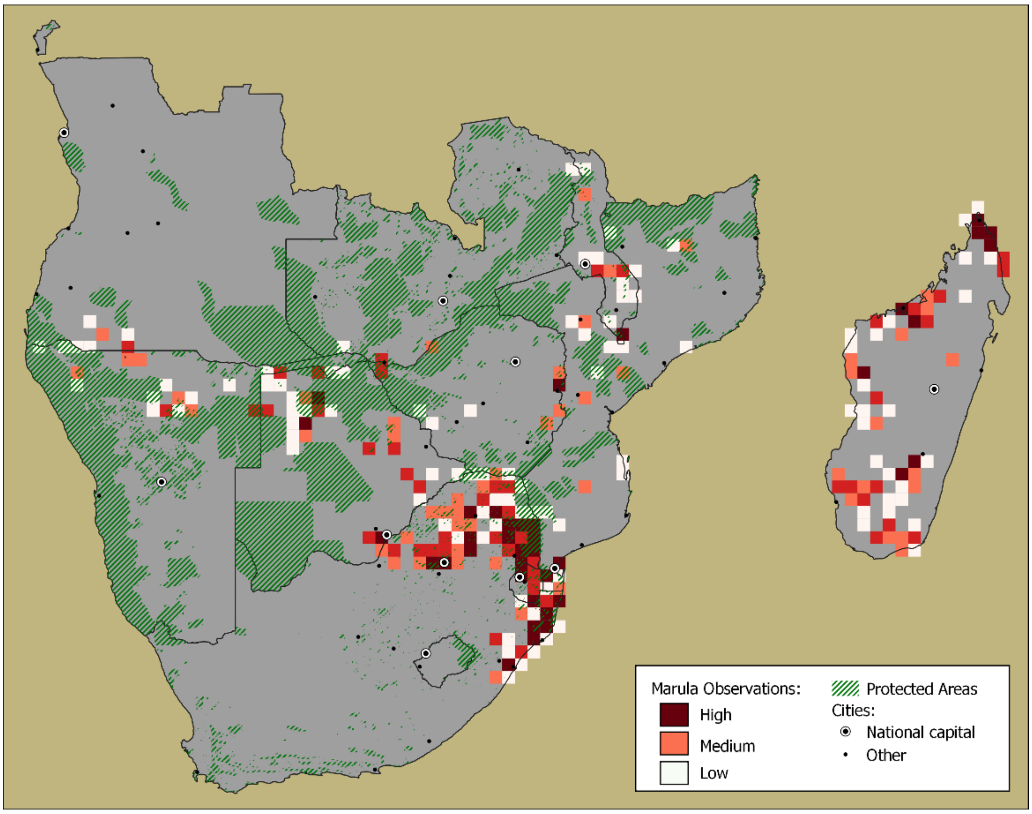

Figure 12: Illustration of a potential approach to identifying sites for monitoring marula in the region.

5. Concluding comments

5.1 Principles of a GIS-based approach to resource assessment and monitoring

There are many suitable resources available that can contribute towards a methodical and rigorous GIS-based desktop approach to resource assessment and monitoring. Some of the key principles that have been dealt with in this report are summarised below.

Availability of software and data
Suitable, freely available software and spatial data of good resolution exist. These can be used to produce repeatable and reputable outputs at a national and regional scale as a starting point for distribution mapping. National scale assessments can be carried out in greater detail depending on availability of suitable and accurate spatial data and expertise.

Local scale mapping potential
There is tremendous potential for desktop method development at a local scale which may yield very detailed, highly accurate results. Applied across a broader scale, with automated approaches, this data may feed in to national/regional scale mapping to produce highly accurate information. The use of these kind of approaches may be dependent on suitable expertise and funding being available.

Use of expert knowledge
Experts with knowledge of the resource can provide valuable input to refinement of modelled results as well as original input at all scales. They can make a significant contribution to the value and accuracy of the output.

Monitoring site selection
Various mapping approaches can be helpful in narrowing down site selection. The choice of monitoring sites will be influenced by the purpose of the monitoring and the nature of the species and its associated use as a resource. Logistics and the practicalities of running a monitoring programme will influence choice of sites, as will the presence of existing focus/research sites.

The value of field data collection
Any resource assessment or monitoring programme must be supported by improved, co-ordinated field data collection across various areas of survey.

Complexity and resource assessment
In a field as complex as conservation (and resource assessment), one cannot hope to produce a 'recipe book' of methods, applicable to every situation. "The only method is that there is no method"(Cunningham, 2002).

5.2 Limitations of regional/national distribution mapping and modelling for resource assessments and future recommendations.

The best available data has been used to address the objectives of this task. Datasets for South Africa are available at a good local scale (1: 50 000 for vector features, 0.5m resolution for imagery) as well as various medium resolution thematic datasets. At a regional scale there are fewer thematic datasets, but medium resolution datasets (eg: Sayres et al., 2020 and the WWF ecoregions) are available. High and moderate resolution imagery (freely available or at a price) is available for local-level mapping. What is lacking is data on a local scale of resource assessments with a common method describing density of plants, history of resource use and associated yields. This kind of information is needed to produce figures on potential abundance, yields and sustainability which can be extrapolated to represent larger areas. Without this data the mapping approaches cannot not yield more information, and more importantly, the accuracy of the mapping cannot be verified.

Accurately mapping species occurrence over a large area usually requires high resolution data, or time-consuming traditional field surveying methods. Either approach tends to be expensive, and therefore an SDM is a common, cost-effective way to determine where a species is likely to occur over a large area (Geller et al., 2017). There are some commonly recognised limitations of this kind of desktop analysis approach:

- The occurrence data for species is inherently incomplete, limited, biased and inaccurate over the entire area of interest
- It does not account for biotic interactions such as inter-species competition
- It assumes the species is in equilibrium with the environment
- It is difficult to account for anthropogenically-related distribution influences
- Generalist species are not necessarily closely tied to specific autecological variables

Except for adding improved (regional) thematic datasets and more accurate marula observation points, the SDM approach for mapping marula distribution has largely been exhausted in this exercise. Further work will provide diminishing returns without addressing some of the fundamental limitations of the SDM method as discussed above. However, there remains significant potential for local mapping approaches that have not been explored here. The use of low-cost mapping drones presents one such approach. Drones have made high resolution imagery easier and cheaper to obtain than before. Furthermore, advancements in the last few years in image classification techniques and deep learning algorithms, particularly in computer vision, have made the processing and analysis of large, high resolution datasets easier and more efficient. Despite this, however, it is still often easier and more accurate to manually digitise small and medium size areas as it requires significantly less expertise and set-up time that automated workflows require.

5.3 Resource assessment methodology workshop: Participant input

In the virtual resource assessment methodology workshop held on 3 December 2020, participants gave input in discussion sessions around the use of technology for resource assessments. A summary of the most relevant information from these sessions is presented here.

Value of field surveys

The value of field surveys must never be forgotten. It is a key source of information and serves to verify all desk-top work. There are many tried and tested methods which are widely in use. Improvements in terms of the value or integrity of field measurements would contribute to better resource assessment. This might be achieved through the use of modern technology for managing data collection.

Variability of conditions

The conditions under which marula is used and under which it occurs varies tremendously from one country to another and even from one community to another. For some it is a precious and essential resource, for others it is one of many livelihoods. In a communal land context, the tree is valued and conserved. In a commercial agricultural setting it poses an obstruction to mechanical cultivation. This variety can only be considered on a case-by-case basis and broad scale generalisations may have limited value.

Nature of information required by different stakeholders

For a resource assessment, the regional, national and local needs in terms of what information is required, and what information is useful, will vary. The information required to meet the needs of policy makers, industry, resource users, ecologists and 'watchdog' organisations will be different and catering for all these needs is difficult to achieve. For example, 'watchdog' organisations would appreciate an understanding of how to interpret some of the information that is available to them. Industry on the other hand would like a checklist of the most important things to see if a resource inventory has value.

The role of modern technology

Modern technologies such as remote sensing offer a lot. However, some basic on the ground inventories by local users with local knowledge and experience, using appropriate survey methods, yield satisfactory information. These basic but adequate methods should not be overlooked in favour of what may be very costly approaches that do not yield better results. A combination of the best and most affordable technology, plus traditional methods, plus use of local experts in deploying some of the technology (ie: as citizen scientists) has great potential.

Value of data/information

There is a need to include a measure of certainty/error in data that is acquired so that the validity of decisions based on that data can be measured by the users.

References

Botelle, A., du Plessis, P., Pate, K. and Laamanen, R., 2002. *A survey of marula fruit yields in North-Central Namibia.* CRIAA SA-DC.

Busby, J. R. 2002. Biodiversity mapping and modelling. In: Skidmore, A.K., 2002. (Ed.) *Environmental modelling with GIS and Remote sensing*. Taylor & Francis, New York.

Cunningham, T., 2002. *People, Wild Plant Use and Conservation*. Earthscan, London. *

Diaz, S., et al., 2019. Pervasive human-driven decline of life on Earth points to the need for transformative change. *Science*, 366, 1327. Available: ttps://science.sciencemag.org/content/366/6471/eaax3100

De Castro, A., Vlok J., Mcllelan W. 2010. *Field survey of the distribution of Pelargonium sidoides and size of selected sub-populations.* Resource Assessment: study conducted for the South African National Biodiversity Institute.

Geller, G.N., et al., 2017. Remote Sensing for Biodiversity. In: Walters, M., and Scholes, R. J., 2017. (Ed's). *The GEO Handbook on Biodiversity Observation Networks.* Springer International, Switzerland.

Hall, J.B., O'Brien, E.M., Sinclair, F.L. 2002. *Sclerocarya birrea: a monograph*. School of Agricultural and Forest Sciences, Publication Number 19, University of Wales, Bangor. *

McGregor, G.K, 2020. *Unpublished PhD notes*. Rhodes University, Makhanda.

Palmer, T. & Weidemann, C., 2020. *Aloe ferox Resource Assessment.* Biotrack South Africa (Pty) Ltd February 2020. (Project Q6546/2019).

Sayre, R., Karagulle, D., Frye, C., Boucher, T., Wolff, N.H., Breyer, S., Wright, D., Martin, M., Butler, K., Van Graafeiland, K. and Touval, J., 2020. An assessment of the representation of ecosystems in global protected areas using new maps of World Climate Regions and World Ecosystems. *Global Ecology and Conservation, 21.*

Schill, S. R., Williams K. 2017. Remote sensing for biodiversity. In: Walters M., and Scholes, R. J., (Ed's). *The GEO Handbook on Biodiversity Observation Networks.* Springer International, Switzerland.

Schippmann, U., Leaman D.J., and Cunningham, A.B., 2002. Impact of Cultivation and Gathering of Medicinal Plants on Biodiversity: Global Trends and Issues. In: FAO, 2002. *Biodiversity and the Ecosystem Approach in Agriculture, Forestry and Fisheries.* Satellite event on the occasion of the Ninth Regular Session of the Commission on Genetic Resources for Food and Agriculture. Rome, 12-13 October 2002. Inter-Departmental Working Group on Biological Diversity for Food and Agriculture. Rome. Available: http://www.fao.org/docrep/005/AA010E/AA010E00.HTML

Shackleton, C., 2002. Growth and fruit production of *Sclerocarya birrea* in the South African lowveld. *Agroforestry systems*, 55(3), pp.175-180.*

Shackleton, S.E., Shackleton, C.M., Cunningham, T., Lombard, C., Sullivan, C.A. and Netshiluvhi, T.R., 2002. Knowledge on *Sclerocarya birrea subsp. caffra* with emphasis on its importance as a non-timber forest product in south and southern Africa: a summary. Part 1: taxonomy, ecology and role in rural livelihoods. *The Southern African Forestry Journal*, 194(1), pp.27-41. *

Tierney, D.A., Powell, T., and Eriksson, C.E., 2019. *Vegetation Mapping*. Oxford Bibliographies. Available: https://doi.org/10.1093/obo/9780199830060-0176

(* Key references)