THE APPLICATION OF NATIONAL SCALE REMOTELY SENSED EVAPOTRANSPIRATION (ET) ESTIMATES TO QUANTIFY WATER USE AND DIFFERENCES BETWEEN PLANTATIONS IN COMMERCIAL FORESTRY REGIONS OF SOUTH AFRICA

Report to the
WATER RESEARCH COMMISSION

by

A van Niekerk[1], C Jarmain[1,2], G Stephenson[1], J du Plessis[1], C Higgs[1],
V Ndyafi[1] and Z Mashimbye[1]

[1] Centre for Geographical Analysis, Geography and Environmental Studies, Stellenbosch University

[2] Department of Biological Sciences, University of Limerick, Republic of Ireland

WRC Report No. 2966/1/23
ISBN 978-0-6392-0388-1

April 2023

Obtainable from
Water Research Commission
Bloukrans Building, 2nd Floor
Lynnwood Bridge Office Park
4 Daventry Road
Lynnwood Manor
PRETORIA

orders@wrc.org.za or download from www.wrc.org.za

DISCLAIMER

This report has been reviewed by the Water Research Commission (WRC) and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the WRC, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

EXECUTIVE SUMMARY

BACKGROUND

Commercial plantations of introduced tree species provide most of the timber and fibre requirements of South Africa. However, not much is known about the spatial-temporal variations of water use in commercial forestry, how the water use varies among genera and regions, and how water used for commercial forestry compares with other competing land uses (e.g. agriculture) and the natural vegetation it often replaces.

South Africa has a semi-arid climate (mean annual precipitation of 500 mm), which means that most areas cannot sustain forestry. To manage the conflict for a limited water resource, policy introduced in 1972 initiated regulation of the commercial forestry industry. In 1998, new legislation declared the industry a streamflow reduction activity (SFRA), i.e. land use that may reduce the amount of water in rivers. Allocation is made for differences in consumptive water use between the principal commercial forestry genera (*Pinus*, *Eucalyptus* and *Acacia*) in the current SFRA water use licensing system. However, post-harvest changes from one genus to another (e.g. *Pinus* to *Eucalyptus*) constitute a change in water use and consequently imply a change in streamflow impacts. The question is whether any adjustment in plantation area (and hence adjustments to existing water use licences) is necessary to account for genus-specific consumptive water use differences.

Previous work on water use of forests has typically focussed on selected individual sites, which is difficult to extrapolate over regions. Emerging trends in the use of remotely sensed or satellite-derived data to estimate and compare the actual evapotranspiration (ET) of diverse commercial forestry plantations merit further investigation. An Earth observation (EO) approach is well suited to compare both current and historic consumptive water use of existing commercial forestry plantations (e.g. adjacent *Eucalyptus* and *Pinus* plantations growing in similar environments). Consequently, this project investigated the benefits of using spatially explicit remotely sensed (satellite) data to quantify the consumptive water use of commercial forest plantations in South Africa.

METHODOLOGY

The first aim was to *establish a geographical database of commercial forests in the main commercial forestry regions of South Africa*. In this study, a large and rich dataset of forest compartments was sourced from several commercial companies. This dataset, along with monthly ET data from 2009 to 2020 sourced from the FAO-funded WaPOR portal, was used to *determine consumptive water use (actual ET) of commercial forestry by means of RS data*, which was the second aim of this study. The third aim was to *validate (ground truth) RS-based consumptive water use of commercial forestry plantations using historical field-based measurements.* It was challenging to achieve this aim, as few field-based measurements have been carried out in commercial forests, and, in recent years, much of the research has been concentrated in KZN, which limited tree species. Nevertheless, data from several studies were sourced and used to validate the RS-based ET estimates. The fourth and final aim was to *describe, analyse and interpret location-specific differences in water use between and within the primary commercial forestry tree genera at specific locations in South Africa.* A series of analyses were carried out to compare the RS-based ET estimates of different genera, species and age groups grown under a range of environmental conditions.

FINDINGS

The median annual WaPOR ET data extracted and used in this study compared well with in situ measurements of previous studies. For instance, the ET estimates for *Acacia* (median ET = 1 096 mm/year) and maximum annual ET estimates to 1 600 mm/year) are in line with those of previous studies. Similarly, our median annual ET (1 123 mm/year) estimates for *Eucalyptus* is on par with the mean annual ET of 1 116 mm/year reported in previous studies. The reported ranges of ET from previous studies (500-1800 mm/year) also correspond well with the ET range of 575 to 1 618 mm/year estimated in this study. The relatively lower median annual ET estimated for *Pinus* (1 038 mm/year) in this study and the higher frequency of lower annual ET values of less than 900 mm/year also agree with previous work, although our *Pinus* estimates are generally higher than those reported in the literature.

Although the comparison of the WaPOR-based ET estimates of this study corresponds well with previous in situ measurements, some level of error (uncertainty) is inevitable. A comparison between different sources of ET data revealed substantial differences, particularly between the ET values of the WaPOR and MOD16 products. The WaPOR ET values corresponded relatively well with the WRC 2014/15 dataset produced in a previous WRC project. The WRC 2014/15 dataset is considered to be the most accurate available dataset covering South Africa, given that it was calibrated using seasonal climatic data captured by 239 weather stations around the country (which is considerably more than what is used in MOD16 and WaPOR). Some differences between the WaPOR and the WRC 2014/15 dataset were observed during winter months, which suggests that the WaPOR dataset is overestimating water use by about 10-30 mm per month during these months. However, this overestimation is unlikely to have a substantial effect on overall (e.g. annual) water use estimates, as most forests are located in the summer rainfall region. Also, the purpose of this project was not necessary to quantify the total water use of all forests, but rather to better understand how water use varies from one genus, species, age, region and type of site to another. The comparisons presented in this report are consequently valid, especially if one can assume that inaccuracies in the WaPOR data are consistent for different genera, species, age groups, regions, etc.

Our results show that water use varies significantly among genera, with *Eucalyptus* species using considerably more water than *Pinus* and *Acacia* species, particularly during the first five years of planting. Although this observation agrees with in situ measurements carried out in previous research, it is difficult to determine whether the relatively high water use of *Eucalyptus* species is biophysical in nature or if it is a factor of site selection/quality (or a combination of these two factors), given that *Eucalyptus* species are commonly established in sites with higher quality (e.g. deep soils, high rainfall) compared to *Pinus* species. In all likelihood, the relatively high water use of *Eucalyptus* species is a combination of these factors.

The water use of *Acacia* species varied considerably and depended on age and environmental conditions, and, since limited cases (compartments) were planted, it was difficult to draw concrete conclusions. Based on the available data, the *Acacia* species use less water than *Eucalyptus* species, but the ET of *Acacia* compartments is generally higher than *Pinus* compartments.

The biggest environmental drivers of ET variability are rainfall and slope gradient. This can be expected, given that these two factors determine (to a large extent) water availability. Precipitation is less likely to infiltrate soils on steep slopes (i.e. a larger proportion of rainfall will contribute to surface runoff on steep slopes), while on a level of moderately inclined terrain, a greater proportion of rainfall will permeate soils and become available for take-up by vegetation. Flatter areas are also typically associated with deeper soils and likely higher soil-water availability, which is conducive to tree growth.

A noteworthy finding of this study is the marked difference between the ET of *Eucalyptus* and *Pinus* compartments on moderately inclined and steep slopes, with the former having 10% higher values (on average). Overall, *Eucalyptus* species consistently used more water than *Pinus* species on moderately inclined and steep slopes. This is supported by the finding that the ET values of *Eucalyptus* compartments planted in mountainous regions are much higher compared to those planted on plains/flat and hilly/undulated terrain, while the opposite is true for *Acacia* and *Pinus* compartments. Whether this observation is the result of site selection or biophysical factors does not really matter in the context of the aim of this study, namely to identify and describe such variations rather than to determine the cause. Nevertheless, more insight into the factors driving these observations would greatly enhance our understanding of water use in the forestry sector and warrants further research.

INNOVATIONS AND CAPACTITY BUILDING

This study contributed significantly to new knowledge. Specifically, the temporal water use profiles of plantation forest planted with different species/genera and in sites with varying environmental (e.g. climatic and terrain) conditions are novel. To our knowledge, the use of remote sensing to quantify water used in commercial forestry has not been done previously. In addition, the use of deep learning for extracting forest patches from very high resolution (25 cm) colour aerial photography and the use of machine learning and satellite imagery for differentiating tree genera at regional scales is innovative. These novel techniques were developed and evaluated by the two MSc students working on the project and demonstrates the significant human capacity that was built in this project.

RECOMMENDATIONS

The findings of this study may be of value for the ongoing discussions on the principles and processes for GE regulation within the commercial forestry industry (being the only declared SFRA). The variations in water use highlighted in this study should also be considered in forest rotation planning. This study showed that there are marked differences among the water use of commercial plantation forestry genera and species/hybrids. For instance, exchanging *Pinus* species for *Eucalyptus* species may have a detrimental effect on stream flow if carried out over large areas within a catchment. The impact of environmental conditions on water use should also be taken into consideration.

Ideally, the ET estimations produced in this study should be compared to actual rainfall per compartment to assess the relationship between water availability and use. Unfortunately, we did not have access to weather station data from costly sources such as the ARC and SA Weather Services (SAWS). It is recommended that more be done to establish a dense network of weather stations throughout South Africa and that such data be made freely available for research purposes to support water use and accounting research. Although TerraClim is a step in the right direction, it requires more (financial) support from the research community.

This study made use of state-of-the-art remotely sensed satellite data and techniques to observe tree water use over an extensive area and period. Although it made a significant contribution to new knowledge, this research project took three years to complete. The South African forestry industry and regulators (e.g. government agencies) need such information to be updated on a regular basis. Ideally, operational solutions for calculating changes in water use associated with GE are required. These solutions should be based on scientifically sound techniques whereby GE regulations can be applied at the plantation stand/compartment level and across the country with the same statistical confidence. It is recommended that the techniques employed in this study are operationalised to produce water use estimations on an annual basis. This should be coupled with field-based measurements at strategic locations throughout South Africa to quantify the uncertainties in the resulting water use estimations.

The South African EO community is dwindling in size, as many senior scientists in the field retire or emigrate. It is critical that we continue to invest in building EO capacity to assist the private and public sectors to optimally use scarce resources such as water and fertile land. This is particularly important within the context of climate change, as the projected increases in temperatures and reduction in rainfall will have dire consequences for the forestry and agricultural industries. The central role that the WRC has played (and is playing) in building EO capacity is commendable; several students and young scientists were involved in this project (they are also co-authors of this report) and were exposed to advanced EO techniques. However, many of the students trained through WRC projects opt to emigrate and apply their skills abroad. More needs to be done to ensure that newly-trained scientists remain in South Africa. The only way to build a strong EO community is to establish employment opportunities that suitably incentivise scientists to remain on South Africa. Greater emphasis on the commercialisation and technology transfer of EO research is recommended, as such activities are more likely to generate sustainable employment opportunities. In addition, the commercialisation (and operationalisation) of EO technologies will substantially increase the impact of WRC-funded research and will ultimately lead to much needed economic growth and sustainable use of South Africa's limited water resources.

ACKNOWLEDGEMENTS

This report would not have been possible without the help of others. In particular, the project team would like to thank the following reference group members for their invaluable inputs and sage advice during the course of the project:

Dr L Nhamo	Water Research Commission (Chairman, 2020-2022)
Prof NS Mpandeli	Water Research Commission (Chairman, 2019)
Dr SN Hlophe-Ginindza	Water Research Commission
Dr R Heath	Forestry South Africa
Dr J Magidi	Tshwane University of Technology
Mr T Newby	Independent researcher
Dr S Dovey	SAPPI Shaw Research Centre
Prof A Ramoelo	University of Pretoria
Mr I Saloojee	South African National Space Agency
Dr M Toucher	South African Environment Observation Network
Dr N Fourie	Department of Water and Sanitation
Ms D Maluleke	Department of Water and Sanitation

In addition, the project team is grateful to:

- The Water Research Commission for funding this research.
- Dr I Germishuizen and the Institute for Commercial Forestry Research for supplying the data for the commercial forest database and assisting with its interpretation.
- Dr R Heath for facilitating the supply of commercial forestry data with industry.
- Dr Z Tshidi, Prof N Jovanovic, Dr S Dzikiti, Dr M Gush and Dr D le Maitre of the CSIR for assisting with the conceptualization of the project and for their assistance in the Inception Report (Deliverable 1) and first Progress Report (Deliverable 3).
- Students Ms C Higgs and Mr V Ndyafi of Stellenbosch University who contributed to this research project.
- Mr A Prins and Ms L McCarthy for their assistance in the forest mapping research.
- Dr J Kemp for his inputs on the use of Synthetic Aperture Radar (SAR).
- Western Cape Government Department of Agriculture and eLEAF, Netherlands for making available FruitLook data.
- Dr P Dye, Prof C Everson and Dr A Clulow for assisting with details on past forestry water use research.

Finally, the project team would like to extend their gratitude to both Prof NS Mpandeli and Dr L Nhamo for the professional way in which they managed this project. The project would not have materialised had it not been for their vision and understanding of the importance of water utilisation in the forestry sector.

This page was intentionally left blank

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

AETI	Actual evapotranspiration and interception
AIP	Alien invasive plant
ALEXI	Atmosphere-land exchange inverse
ANN	Artificial neural networks
ARC	Agricultural Research Council
BRDF	Bidirectional reflectance distribution function
CART	Classification and Regression Tree
DAAC	Distributed Active Archive Centres
DL	Deep learning
DT	Decision tree
ESA	European Space Agency
EVI	Enhanced vegetation index
GE	Genus exchange
GEE	Google Earth Engine
GEOBIA	Geographical object-based analysis
GIS	Geographical information system
HR	High resolution
ICFR	Institute for Commercial Forestry Research
KS	Kappa statistic
LAI	Leaf area index
LCLU	Land cover/land use
LDA	Linear discriminative analysis
LiDAR	Light detecting and ranging
LST	Land surface temperature
ML	Machine learning
NASA	National Aeronautics and Space Administration
NDVI	Normalised difference vegetation index
NN	Neural networks
OA	Overall accuracy
OLI	Operational land imager
RF	Random forest
RMSE	Root mean square error
ROCAT	Rapid object collection and analysis tool
RS	Remote sensing
SANLC	South African National land cover
SANSA	South African National Space Agency
SAR	Synthetic Aperture Radar
SAWS	SA Weather Services

SEBAL	Surface energy balance algorithm for land
SEBS	Surface energy balance system
SfM	Structure from motion
SUDEM	Stellenbosch University Digital Elevation Model
SVM	Support vector machine
UAV	Unmanned aerial vehicle
USGS	United States Geological Survey
VH	Vertical-horizontal
VHR	Very high resolution
VITT	Vegetation index/temperature trapezoid
WB	Water balance
WC	Western Cape
WCDOA	Western Cape Department of Agriculture

LIST OF UNITS

GHz	gigahertz
ha	hectare
km	kilometre
km^2	square-kilometre
m	meter
min	minute
mm	millimetre
µm	micrometre
%	percentage
°	degree

1 INTRODUCTION AND OBJECTIVES

1.1 Introduction

Commercial plantations of introduced tree species provide most of the timber and fibre requirements of South Africa. However, not much is known about the spatial-temporal variations of water use in commercial forestry, how the water use varies among genera and regions, and how water used for commercial forestry compares with other competing land uses (e.g. agriculture) and the natural vegetation it often replaces.

South Africa has a semi-arid climate (mean annual precipitation of 500 mm), which means that most areas cannot sustain forestry (Poynton 1971). The deep-rooted, tall, dense, evergreen physiology of tree plantations contrasts strongly with the typically short, seasonally dormant vegetation with shallow root systems (e.g. grassland/shrubs) that they typically replace. Differences in evapotranspiration rates and resultant impacts on water resources have been quantified through forest hydrology research using paired catchment experiments (Wicht 1948; Scott et al. 2000), in situ field measurements (Savage et al. 2010) and genus-specific modelling at a watershed scale (Van Wyk 1987), but very little is known about how much water is used for commercial forestry at regional and national scales.

To manage the conflict for a limited water resource, policy introduced in 1972 initiated regulation of the commercial forestry industry (Van der Zel 1995). In 1998, new legislation declared the industry a streamflow reduction activity (SFRA), i.e. land use that may reduce the amount of water in rivers (Gush et al. 2002). Allocation is made for differences in consumptive water use between the principal commercial forestry genera (*Pinus*, *Eucalyptus* and *Acacia*) in the current SFRA water use licensing system. However, post-harvest changes from one genus to another (e.g. *Pinus* to *Eucalyptus*) constitute a change in water use and consequently imply a change in streamflow impacts. The question is whether any adjustment in plantation area (and hence adjustments to existing water use licences) is necessary to account for genus-specific consumptive water use differences.

At present, the genus exchange (GE) regulation for SFRAs is being debated among government, private industry and academic/research institutes in South Africa. Discussions revolve around finalizing appropriate definitions, principles and processes for GE regulation within the commercial forestry industry (being the only declared SFRA). Consequently, there is a need for studies that explore differences between the water use of commercial plantation forestry genera vs species (even clones and hybrids where possible) using state-of-the-art approaches (e.g. remotely sensed satellite data) and observed tree water use data from field studies. Ideally, operational solutions for calculating changes in water use associated with GE are required. These solutions should be based on scientifically sound techniques whereby GE regulations can be applied at the plantation stand/compartment level to the extent that they can be applied across the country with the same level of statistical confidence.

Contemporary forest hydrology research has focussed on improved understanding of the water use processes associated with introduced and indigenous tree species (Everson et al. 2011). While a number of field studies on tree and plantation water use have taken place, streamflow reductions estimates associated with commercial forest plantations have relied on distributed modelling at quaternary and quinary scales across the entire country (Gush et al. 2002). These results are used to licence and regulate the commercial forestry industry. However, there are recognised shortcomings in the often-used ACRU-modelling approach (assumptions, generalisations, etc.). This approach assesses the water use differences between different plantations of commercial forestry genera growing under similar conditions (plantation age, spacing, soil type, tree health, etc.). However, specific growing conditions have an important bearing on the water use of a plantation of a particular genus and can thus not be ignored. Factors that need to be considered in terms of hydrological impacts include plantation age, species (clone/hybrid), stand densities, rotation lengths, stand management (weeding, pruning, thinning, etc.), canopy closure and changes in site/species preferences. The cost vs accuracy trade-off needs to be considered when obtaining these input parameters. Other model input considerations include deciding on the most representative tree age for rotation length water use of different genera and using appropriate model parameter values that most accurately represent that particular genus and tree age. The uncertainties inherent in measuring/obtaining these parameters – and in using them to model water use – are likely to be greater than species/clone/hybrid-specific differences, making efforts to differentiate water use very challenging.

Emerging trends in the use of remotely sensed or satellite-derived data to estimate and compare the actual evapotranspiration (ET) of diverse commercial forestry plantations merit further investigation. A remote sensing (RS) methodology to quantify water use of irrigated crops was developed in the recently completed and published in WRC report TT 745/17 (AN EARTH OBSERVATION APPROACH TOWARDS MAPPING IRRIGATED AREA AND QUANTIFYING WATER USE BY IRRIGATED CROPS IN SOUTH AFRICA). A similar approach was used in a follow-up study to estimate water use of different land covers/users (WRC project K5/2520). An Earth observation (EO) approach is well suited to compare both current and historic consumptive water use of existing commercial forestry plantations (e.g. adjacent *Eucalyptus* and *Pinus* plantations growing in similar environments). It could also be used to compare the water use of riparian vs upland stands of trees, as well as "before and after" scenarios, such as the effects of thinning, pruning or weeding operations on overall stand water use.

Consequently, this project investigated the benefits of using spatially explicit remotely sensed (satellite) data to quantify the consumptive water use of commercial forest plantations in South Africa.

1.2 Aims

The aims of this project (K5/2558//4) were to:

1. produce a geographical database of commercial forests in the main commercial forestry regions of South Africa;
2. determine consumptive water use (actual ET) of commercial forestry by means of RS data;
3. validate (ground truth) RS-based consumptive water use of commercial forestry plantations using historical field-based measurements; and

4. describe, analyse and interpret location-specific differences in water use between and within the primary commercial forestry tree genera at specific locations in South Africa.

1.3 Research and development activities and report structure

First, this report overviews the research activities and main findings of the project. Second, it reviews existing knowledge (Section 2) relating to the water use of commercial plantation forests in South Africa. The review also covers the data and techniques that are most appropriate for quantifying water use at regional scales. This is followed by an account of the data and methods (Section 3) used in this project. Section 4 reports on fundamental EO research carried out as part of this project, including the development of methods for mapping forests, three genus and age. Section 5 overviews the plantation forest water use quantifications and analyses the factors that had the biggest impact on water use variations. The report concludes with a summary of the main findings and recommendations (Section 6).

2 KNOWLEDGE REVIEW

2.1 Land and water used for forestry in South Africa

2.1.1 Forestry in South Africa

A shortage of wood for construction was identified as a problem during the development of the Cape Colony under Dutch rule (Zahn and Neetling 1929; Bennett and Kruger 2014). The problem was alleviated to a degree by the discovery of more extensive forests further east at Swellendam and later in the George-Plettenberg area. By the middle of the 19th century, the British colonial government recognised that they were going to need to establish plantations of fast-growing trees to meet demand. This led to the established of the first plantations of *Pinus radiata* at Tokai and elsewhere on the Cape Peninsula from the late 1800s onwards. These soon were followed by plantations at selected sites around South Africa.

Timber shortages during the 1st and 2nd World Wars fuelled the establishment of plantations in all the higher rainfall areas of the country (Beinart 1984; Bennett and Kruger 2014). There were also active campaigns aimed at stimulating private forestry, and most of the wattle (*Acacia mearnsii*) plantations were established by farmers and timber companies. Farmers and other land-owners soon began to raise concerns about the decreases in the volumes of water in streams and rivers in these afforested areas (Wicht 1948; Bennett and Kruger 2014, and many people were convinced that afforestation would increase the volume of water. These opposing views led to intensive discussions at the Empire Forestry Conference in 1930, which were resolved by the initiation of a programme of catchment experiments to determine whether or not the stream flows were decreased by afforestation.

The first study was initiated in the Jonkershoek valley near Stellenbosch in 1935 and was followed by further catchment experiments in KwaZulu-Natal, Mpumalanga and Limpopo. In 1961, the Interdepartmental Committee on the Conservation of Mountain Catchments published guidelines on the protection of water resources in mountain catchments, including principles for sustainable land uses such as forestry and agriculture (Department of Agricultural Technical Services, 1961).

The hydrological studies described in the next section demonstrated that plantations of introduced, fast-growing tree species reduce streamflow. The findings were used in the afforestation permit system, which limited the area that could be afforested in different catchments based on the impacts on river flows and low flows (Nänni 1970; Van der Zel 1995; Scott and Smith 1997). The permit system was later replaced by licences which were issued under the streamflow reduction provisions of the National Water Act (Gush et al. 2002; Bennett and Kruger 2014). The area under commercial forest plantations increased steadily to about 1.50 million ha but subsequently declined to 1.21 million ha due to the phasing out of plantations in areas that were no longer considered commercially viable.

2.1.2 Existing knowledge of water used for forestry

The long-term, catchment-based studies of the impacts of afforestation on streamflow were among the first of their kind of the world (Wicht 1948; Bennett and Kruger 2014). They generally followed a multiple catchment design, developed by and adapted by Dr Wicht to suit the circumstances at the various sites. At each site, a control catchment was maintained under natural vegetation as the baseline for estimating the reductions caused by afforestation. For the Jonkershoek catchments, the afforested catchments were progressively afforested, which allowed for the establishment of a pre-treatment rainfall-runoff relationship that could also be used to assess the impacts of afforestation.

The catchment experiments, summarised in Table 2-1, showed that the reductions in streamflow do not commence immediately after afforestation. A lag before clear evidence of streamflow reduction was observed (Van Wyk 1987; Bosch and Von Gadow 1990; Scott et al. 2000). The duration of the lag period is determined by the growth rate of the trees, which depended on the growing conditions and also the tree genus (Le Maitre and Versfeld 1997; Scott and Smith 1997; Gush et al. 2002). Afforestation with *Pinus* results in a longer lag period compared with *Eucalyptus*, and the lag period is shorter where the growing conditions are optimal compared to where they are sub-optimal.

Table 2-1 A summary of the findings of the long-term, multiple catchment studies of the effects of afforestation on streamflow.

Region	Site	Catchment	Soils	Species	Area (ha)	Mean annual rainfall (mm/year)	Pre-planting streamflow (mm/year)	Planted (%)	Riparian zone planted	Mean annual flow reduction when mature (mm/year)	Evaporation (mm/year)
Western Cape	Jonkershoek	Bosboukloof	Deeply weathered granite, sandstone talus	*Pinus radiata*	200.9	1449	527	57	No	172.9	1095
Western Cape	Jonkershoek	Biesievlei	Deeply weathered granite, sandstone talus	*Pinus radiata*	27.2	1297	476	98	Yes	274.4	1095
Western Cape	Jonkershoek	Lambrechtsbos A	Deeply weathered granite, sandstone talus	*Pinus radiata*	31.2	1477	433	84	No	207.2	1251
Western Cape	Jonkershoek	Lambrechtsbos B	Deeply weathered granite, sandstone talus	*Pinus radiata*	65.5	1477	410	82	No	250.3	1317
KZN Drakensberg	Cathedral Peak	II	Deeply weathered basalt	*Pinus patula*	190.7	1634	807	74	No	425.5	1253
KZN Drakensberg	Cathedral Peak	III	Deeply weathered basalt	*Pinus patula*	142	1528	683	84	No	395.1	1240
Mpumalanga Escarpment	Mokobulaan	A (euc)	Deeply weathered shales	*Eucalyptus grandis*	26.2	1164	244	100	Yes	128.7	1049
Mpumalanga Escarpment	Mokobulaan	B	Deeply weathered shales	*Pinus patula*	34.6	1171	217	100	Yes	151.2	1105
Limpopo Tzaneen	Westfalia	D (euc)	Deeply weathered granite-gneiss	*Eucalyptus grandis*	39.6	1477	190	83	Yes	267.6	1555
KZN Midlands	Mistley-Canema Estate	Two Streams	Deeply weathered shales	*Acacia mearnsii*	73	920	72	100	Yes#	50.3	898

Source: Scott et al. (2000) with data for Two Streams from Clulow et al. (2011)

2.2 Forest water use estimation methods

Numerous methods are available to estimate forestry water use (consumption). These methods capture the processes of transpiration and evaporation (from the soil surface and water intercepted by the canopy), as reflected by evapotranspiration (ET). The methods, developed locally and internationally, estimate water use at plant, field or catchment level and involve measurements or modelling.

2.2.1 Field-based methods

Field-based methods for measuring or estimating ET have been reviewed in previous work (Savage et al*.* 2010). These methods include:

- the soil water balance approach;
- pan methods (e.g. class A pan);
- the reference evaporation and crop factor approach;
- lysimetry;
- the eddy covariance method; and
- a range of aerodynamic methods that estimate sensible heat from which evaporation is estimated as a residual, using the shortened energy balance equation (Jarmain et al*.* 2009b).

The latter method includes the one sensor eddy covariance, Bowen ratio, surface renewal and scintillometry method. Some methods are used to estimate transpiration directly, e.g. the sapflow and heat pulse velocity methods (Jarmain et al*.* 2009b). Many of these methods have been used to estimate ET in SA as part of Water Research Commission (WRC) funded projects. See for example Bristow and De Jager (1981); Green and Clothier (1988); Dye et al*.* (1997); Savage et al*.* (1997); Everson et al*.* (1998); Everson (1999); Savage et al*.* (2004); Jarmain et al*.* (2009a); and Jarmain et al*.* (2014). The abovementioned methods are point- or field-based. As such, their ET estimates have a limited spatial "footprint".

Spatially explicit methods developed to estimate ET fill the need for geographical estimates of ET. Advances in remotely sensed data enable per-pixel assessments of ET at resolutions ranging from 1 000 m to as small as 20 m. Such EO methods for estimating ET over large areas are discussed in Section 2.2.2. These methods have been applied to study forestry water use by Gush and Dye (2009), Dye et al*.* (2008b), Jarmain and Everson (2002), Everson et al*.* (2005), Dye et al*.* (2008a) and others.

2.2.2 Earth observation methods

The increased availability of spatially referenced geographical information system (GIS) and RS data enable crop water use or ET estimation at pixel level and at high resolutions (e.g. 20 to 1 000 m). Such data can be aggregated and employed at different spatial scales and used over large areas. Because satellite data are frequently collected, estimates can be made regularly, and temporal trends studied. Such spatial and temporal coverage can contribute greatly towards improved water management at national and/or regional level down to individual farms or fields.

Estimates of evapotranspiration (ET) from a surface, including water consumption by vegetation, relate to the vapourisation of water from the land surface into the lower part of the atmospheric boundary layer. ET consists of evaporation of water from the soil, evaporation of intercepted water and transpiration losses by plants, and the sum of all these losses is often referred to as consumptive water use. The water volumes lost through the processes encompassed in ET form part of the hydrological cycle where no water is truly lost but merely changing in form.

Advances in the interpretation of RS information enable the spatial assessment of crop water use, biomass and yield production and associated WUE for each pixel of a satellite image, without having to rely on generalised crop coefficients. Different methods have been developed to provide information at a range of temporal and spatial scales and for various applications. A number of review papers describe methods used to spatially estimate ET, including Choudhury (1997), Courault et al*.* (2005), Kustas and Norman (1996), Verstraeten et al*.* (2005), Verstraeten et al*.* (2008) and Gibson et al*.* (2013). Numerous models have been developed for agricultural (field scale) applications. Examples include: surface energy balance algorithm for land (SEBAL); surface energy balance system (SEBS); mapping evapotranspiration with high resolution and internalised calibration (METRIC^{tm}); vegetation index/temperature trapezoid (VITT); two-source energy balance; the atmosphere-land exchange inverse (ALEXI); and normalised difference vegetation index diurnal surface temperature variation (NDVI-DSTV) triangle model. These methods either estimate ET as the residual of a shortened energy balance equation using land surface temperature (LST) estimates or use a WUE relationship to determine ET. Some of the models are used operationally for field scale agricultural water management[1], but most are used primarily in research applications. A selection of the models (SEBAL, SEBS, VITT and METRIC^{tm}) was reviewed by Jarmain et al*.* (2009b). The review included an assessment of each model's accuracy in estimating ET and their potential for operational applications in SA. It was found that some of the components of the energy balance (such as net radiation) were accurately simulated, but that the other energy balance components and ET were generally more complex. SEBAL and METRIC estimates of ET were generally lower than measured ET, while SEBS commonly overestimated ET. The VITT model yielded the least accurate evaporation estimates.

Other RS-based models have been developed and provide ET estimates at lower spatial resolutions (often ~1 to 3 km) but higher temporal resolutions (30 min to daily). The lower spatial resolution of these models makes them less suited for agricultural applications where information at field scale is required. A number of these models use Meteosat Second Generation satellite data and provide ET data at 30-minute intervals, at a resolution of 1-3 km[2]. ET data from HYLARSMET[3] and MODIS[4] are estimated daily for the entire globe at a 1 km resolution. The global water cycle monitor[5] from Princeton University also estimates ET at a daily time step, while the ALEXI model[6] can be used to estimate energy fluxes

[1] For instance, www.mijnakker.nl; fruitlook.co.za; www.idwr.idaho.gov/GeographicInfo/METRIC/et.htm
[2] http://landsaf.meteo.pt/ and http://www.ears.nl/
[3] http://sahg.ukzn.ac.za/soil_moisture/et/
[4] http://modis.gsfc.nasa.gov/data/dataprod/dataproducts.php?MOD_NUMBER=16
[5] http://hydrology.princeton.edu/~justin/research/project_global_monitor/
[6] http://alfi.soils.wisc.edu/cgi-bin/anderson/alexi_server.pl?region=SMEX02MOD

and other parameters daily, e.g. at a 10 km spatial resolution.

New approaches and models are continually being developed and tested. For instance, the ETLook model (Pelgrum et al. 2011) is used in the water productivity through open access of remotely sensed derived data (WaPOR)[7] initiative, which provides free access to satellite-based data on agricultural productivity in Africa and the Near East for the period 2009-2019. The purpose of the project is to allow for land and water productivity monitoring, using ET and biomass production data. Three levels of data products are available. Level 1 provides 250 m resolution data on a continental level. Level 2 provides 100 m data for a number of selected countries, including Morocco, Tunisia, Kenya and Mozambique. Additionally, Level 2 includes the Jordan/Litani River basin, the Nile basin, the Awash basin and the Niger inner delta. Level 3 provides 30 m resolution data on irrigation scheme level.

Van Niekerk et al. (2018) demonstrated how EO data can be employed to quantify water use of irrigated agriculture at national scale. A range of EO datasets was used to map irrigated fields, after which 250 m ET data, derived from the ETLook algorithm, was used to quantify water use at field level. Water use was also differentiated for different crop types and in different regions throughout South Africa. The results were verified and validated through an extensive stakeholder engagement exercise and by comparing the water use estimates with those quantified in previous studies.

2.3 Technologies and techniques beneficial to forestry land and water use estimation

It is clear from the previous sections that water use estimations require reliable, accurate and up-to-date data with good spatial coverage. This data are often needed for large areas, which necessitates the use of geospatial technologies such as GISs and satellite RS. This section provides an overview of the geospatial techniques and technologies that were used in project. The review starts with GIS and spatial modelling, as these are the fundamental technologies used to quantify water use. This is followed by an introductory overview of concepts such as optical, thermal and microwave RS. The section concludes with a short overview of image classification and object-based image analysis techniques.

2.3.1 GIS and spatial modelling

GIS is used to manage and analyse spatially referenced or geographical data and provides a unique platform capable of integrating large volumes of spatial data for analysis (Heywood et al. 2006). GIS offers a quick and easy way to monitor and manage resources, which is not possible with traditional (analogue) methods. Over the last 20 years, GIS has emerged as a mature technology with particular value in answering questions about spatial location, patterns, trends, conditions and their implications. Within a GIS, datasets of different formats at varying scales can be incorporated into a single database, which can be stored as vector and/or raster data.

Spatial modelling involves using such data to construct models to predict spatial outcomes that simulate

[7] http://www.fao.org/in-action/remote-sensing-for-water-productivity/database/database-dissemination-wapor/en/

the dynamics of natural processes (O'Sullivan and Unwin 2010). Spatial modelling in GIS embraces techniques and models that apply quantitative structures to systems in which the variables of interest vary across space. Spatiotemporal models simulate change over time using equations that represent real-world processes while taking spatial patterns and spatial interaction in the system into account (Karssenberg et al. 2008). Such spatial and temporal process models can be used for decision-making regarding spatial phenomena (also known as spatial decision support systems) but are also used to evaluate our understanding of complex spatial systems (Heywood et al. 2006). Models can be used to establish (*a priori*) theory or explore (*a posteriori*) theory (Hardisty et al. 1993). When modelling in GIS, the questions of validation and the roles of scale and accuracy need to be carefully considered (Goodchild 2005).

There are numerous examples of where GIS have been used in forest management. Recent examples include Akumu et al. (2019), who developed a GIS-based modelling procedure to predict and map relative soil moisture classes in a forested landscape. They used rule-based GIS (also known as cartographic modelling) to develop a technique to predict soil moisture classes (dry, fresh, moist and wet). Soil textural classes were derived from quaternary geology maps and water receiving areas, which were in turn derived from topographic attributes generated from a digital elevation model. Their approach yielded a soil moisture map with an overall accuracy of about 65% relative to 54% generated from soil wetness index reclassification approach.

2.3.2 Remote sensing and Earth observation

RS, or EO, is the practice of deriving information about the earth's surface using images acquired from an overhead perspective by sensing and recording electromagnetic (EM) radiation reflected or emitted from the features on the surface of the earth (Campbell 2007b). Due to its areal coverage of the earth's surface at a variety of scales (and common availability in near real-time), EO data form the foundation for many environmental spatial datasets. Additionally, the availability of archived satellite imagery over the past ~40 years provides an invaluable comprehensive database for environmental monitoring and modelling.

EO sensors detect and record incoming EM radiation in various regions of the EM spectrum using optical, thermal or microwave sensors (Figure 2-1). Optical sensors, also called multispectral sensors, measure energy reflected from the surface of the earth originally generated by an external source (usually the sun),and operate primarily in the visible and the infrared regions of the EM spectrum. The visible spectrum contains those wavelengths of radiation that can be perceived by human vision, i.e. from violet to red light. Wavelengths longer than those of the visible spectrum (but shorter than those of microwave radiation) are termed infrared, which can be further subdivided into near-, mid- and far-infrared. The visible, near- and mid-infrared wavebands are collectively referred to as the optical bands (Campbell 2002; Mather 2004). Unlike optical RS, thermal RS detects energy in the far-infrared portion of the EM spectrum, which is energy absorbed by the earth's surface and then emitted in the form of heat (or thermal radiation) (Campbell 2002; Mather 2004).

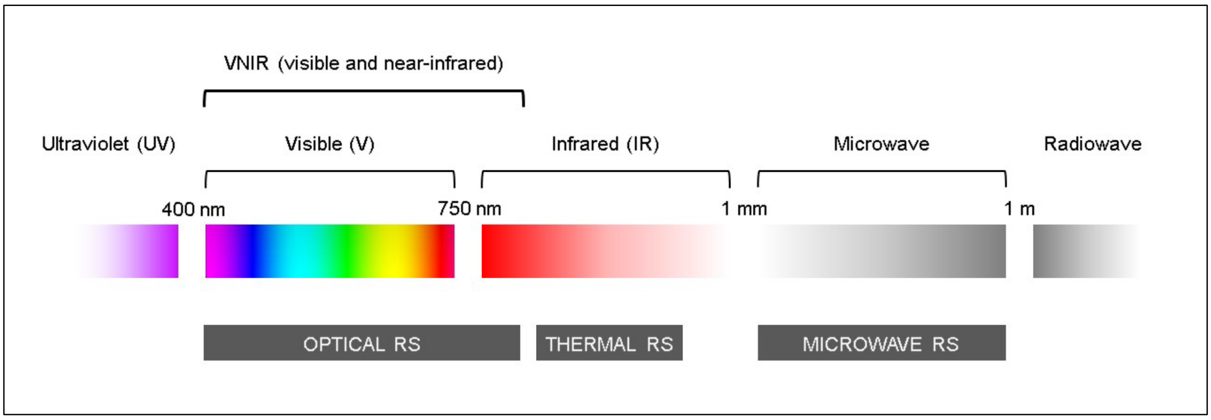


Figure 2-1 The electromagnetic spectrum and RS (not to scale)

The longest wavelengths commonly used in RS fall in the microwave spectrum. Although the earth itself emits some microwave energy, solar irradiance in this spectrum is negligible and is therefore rarely measured in RS. Instead, most microwave sensors are active radar (radio detection and ranging) sensors, which use their own energy to irradiate the ground and then measure the portion of energy reflected back to them (Campbell 2002; Mather 2004).

Optical, thermal and microwave RS techniques are overviewed in the following subsections. The aim of this overview is to provide a condensed account of the existing RS technologies and approaches that might be employed in this project.

2.3.2.1 Optical sensors

A large number of optical EO sensors are commercially or freely available. The choice of the appropriate sensor for a particular application is informed by factors relevant to the application. These include:

- spatial resolution (also known as the pixel size), which is a measure of the level of detail that can be recognised using the imagery;
- spectral resolution, referring to the number of spectral bands available;
- temporal resolution (also known as revisit cycle), which denotes the time interval between image acquisitions for the same area;
- swath width (also called image extent or scene footprint), which describes the square kilometre area covered by one scene; and
- cost per image.

Several optical satellite platforms were considered for use in this research. The low spatial resolutions of the freely available MODIS (250-1 000 m) and AVHRR (1 km) satellite imagery are unsuitable for detailed mapping exercises (e.g. mapping the boundaries of individual commercial forestry blocks). Conversely, although the sub-metre resolutions of very high resolution (VHR) sensors such as Ikonos, Quickbird, Worldview and GeoEye (in the panchromatic bands) are highly suitable for analysing the structural/spatial properties of commercial forestry, their use over large areas (e.g. at national scales) is limited as the sensors have small image footprints. This means that thousands of images will be required for one national coverage, which would be prohibitively expensive.

The imagery provided by high resolution (HR) optical sensors, such as those mounted on the Landsat, SPOT and Sentinel-2 satellites, are ideal for when very large mapping scales are not required. These images have large swaths, which means that not many images are required to cover a large area. The SPOT family of satellites have been recording HR satellite imagery for almost 30 years, with the latest addition being SPOT 7 launched on 30th June 2014. The South African National Space Agency (SANSA) and Airbus Defence, the owner of the SPOT satellite series, have in place a licence agreement which allows the use of SPOT imagery for government and research purposes.

The Sentinel-2 programme, developed by the European Space Agency, forms part of the European Union's comprehensive Copernicus EO programme aimed at performing terrestrial observations in support of services such as forest and agricultural monitoring, land cover change detection and natural disaster management. The platform is comprised of two identical HR multispectral satellites: Sentinel-2A (launched on 23 June 2015) and Sentinel-2B (launched on 7 March 2017). The spatial and spectral characteristics of the Sentinel-2 sensors are provided in Table 2-2.

Table 2-2 Sentinel-2 sensor characteristics

Sentinel-2 Bands	Central Wavelength (µm)	Resolution (m)
Band 1 - Coastal aerosol	0.443	60
Band 2 - Blue	0.490	10
Band 3 - Green	0.560	10
Band 4 - Red	0.665	10
Band 5 - Vegetation Red Edge	0.705	20
Band 6 - Vegetation Red Edge	0.740	20
Band 7 - Vegetation Red Edge	0.783	20
Band 8 - NIR (near infrared)	0.842	10
Band 8A - Vegetation Red Edge	0.865	20
Band 9 - Water vapour	0.945	60
Band 10 - SWIR - Cirrus	1.375	60
Band 11 - SWIR	1.610	20
Band 12 - SWIR	2.190	20

The Landsat sensor is the most commonly used HR data, with Landsat satellites continuously capturing images of the earth's surface since 1972. The Landsat Data Continuity Mission, run by the North American Space Agency and the United States Geological Survey, comprises ~40 years of imagery, all of which is freely available. To date, eight Landsat missions (Landsat 1-9) have been launched, with the latest being Landsat 9 (launched in September 2021). Landsat 9 is a twin of Landsat 8 and carries two instruments, the operational land imager (OLI) and thermal infrared (TIR) sensor. The OLI sensor detects seven multispectral bands at 30 m resolution and a panchromatic band at 15 m resolution. Landsat 5 TM was decommissioned in 2013, and the scan-line corrector of Landsat 7 ETM+ has been inoperative since 2003, resulting in large gaps in the imagery. However, the continuity and high spectral resolution among Landsat TM, ETM+ and OLI are highly beneficial for multitemporal analysis, which will be employed in this research.

2.3.2.2 Thermal remote sensing

Thermal RS deals with the acquisition, processing and interpretation of data acquired primarily in the thermal infrared (TIR) region of the EM spectrum (3 to 35 μm). In thermal RS, the radiation 'emitted' from the surface of the earth is considered, as opposed to optical RS where the radiation is 'reflected'. A commonly studied aspect in the domain of thermal RS is LST. LST provides information on the temporal and spatial variations of the surface equilibrium state (Li et al. 2013) and is an important observation particularly in the estimation of land surface atmospheric fluxes. However, the strong heterogeneity of land surface characteristics such as vegetation, topography and soil leads to a rapidly changing LST in both space and time, resulting in RS satellite data offering the only possibility for measuring LST over the entire globe with sufficiently high temporal resolution (Kalma et al. 2008; Li et al. 2013). For example, when using an energy balance approach to estimate evapotranspiration (Section 2.2.2), LST is used in the estimation of net radiation and to estimate the sensible heat flux. Although there were early doubts as to whether satellite-based radiometric temperature could be used in the estimation of evapotranspiration (Kalma et al. 2008), it has since been established that, to estimate evapotranspiration with an accuracy of more than 10%, LST must be retrieved at an accuracy of 1 K or more (Li et al. 2013). This reinforces the need to obtain accurate LST for critical observations in hydrology (e.g. ET).

2.3.2.3 Microwave remote sensing (RADAR)

Microwave RS functions by detecting energy backscattered from the earth's surface in the microwave region of the electromagnetic spectrum. This region ranges in frequency from 0.3 to 300 GHz, which corresponds to wavelengths of 3 mm to 30 cm. The long wavelengths of microwave radiation mean that it experiences very little atmospheric attenuation, making it possible for imaging radars to capture image scenes even in cloudy conditions. This makes radar imagery particularly powerful for obtaining unbroken time series of data, especially in tropical or cloudy regions. Most imaging radars are active sensors (in that they provide their own source of microwave illumination), which further means that they can capture imagery both day and night, thereby doubling the imaging capacity per orbit.

Synthetic aperture radar (SAR) is a type of microwave sensor that provides a dramatically improved spatial resolution over conventional real aperture radar systems. Both types of radars are currently operational as space-borne systems and will be briefly described in the following subsections.

Several radar satellites are currently in operation. Some of these satellites offer HR, fully-polarimetric (HH, HV, VH and VV) capabilities. Examples include RADARSAT-2, TerraSAR-X and COSMO-Skymed. The most reliable and commonly used SAR sensors currently active are RADARSAT-2, TerraSAR/TanDEM-X, ALOS PALSAR-2, Cosmo-SkyMed and Sentinel-1A. These sensors provide a mixture of spatial resolutions (1-100 m), wavelengths (X, C and L bands) and revisit periods (11-24 days). RADARSAT-2 is currently the only sensor that routinely provides reliable data in full polarisation, making it an important source of polarimetric data. However, TerraSAR-X and ALOS PALSAR-2 also have full polarimetric capacity. Sentinel-1 contains a dual-polarisation C-band SAR capturing data at three different resolutions (10, 25, or 40) and four band polarisations (VV, HH, VV + VH, and HH + HV). Sentinel-1 has a temporal resolution of six days and the data is made available to the public within 24 hours of observation.

2.3.3 Cloud-based remote sensing platforms

Hansen and Loveland (2012) speculated that the future of large-scale operational mapping lays with automated processing chains for multi-image classification, facilitated by the image standardisation of long-term, free satellite programmes such as Landsat (and more recently, Sentinel). This concept has been advanced by the recent development of open access data catalogues and cloud-based computing services for geospatial analysis, such as Google Earth Engine and Amazon Web Server (Koskinen et al. 2019; Mauya et al. 2019). These services, which are freely available for research purposes, provide access to decades of remotely sensed data, which can be queried, manipulated, analysed and visualised using a wide variety of RS toolboxes. As well as bypassing the need for data acquisition and on-site storage and processing, the architecture of these services allows for automated processing on a large scale in a scripting environment.

2.3.4 Image classification approaches

Digital image classification methodologies (image classifiers) involve a set of computer procedures that assign image pixels or objects to classes representing information categories relevant to the user, based on a diverse selection of inherent image features (Campbell 2007a). The development of image classifiers has been subject to ongoing research since the introduction of RS. A wide variety of classifier types and forms exists, each with its own strengths and weaknesses relative to applications to which they may be applied (Lawrence and Wright 2001; Mather 2004). When deciding on a classification method for an application, a user must weigh up the importance of several different factors. The efficacy of classification methods is usually assessed based on the accuracies of the final classification products using statistical metrics. However, the demand for human expertise, the time and expense of preparing and running the classifier, and the degree of automation required are aspects which must be taken into account (Pal and Mather 2003). It should also be noted that the accuracies of different classification methodologies are often specific to the application to which they are put (Lui et al. 2002). It is therefore important that the user is aware of the different types of available classifiers in order to judge which is better suited to the application at hand.

Conventional methods of image classification consist of supervised and unsupervised procedures, which rely strongly on a variety of statistical algorithms employed in spectral feature space. Although widely used in operational applications, these more traditional classifiers are not without their limitations. The progression of digital image analysis techniques, combined with the advancement of computer hardware and software, has led to the development and increased implementation of more advanced classifiers which utilise a greater degree of data mining for image pattern recognition (Tseng et al. 2008). This is done by incorporating techniques such as artificial intelligence, logical structures and expert knowledge into the classification procedures (Brown de Colstoun et al. 2003; Mather 2004). The following sections focus on the common methods used for discriminating land cover in remotely sensed imagery.

2.3.4.1 Unsupervised classification

Unsupervised classification is the clustering of image pixels into groups based on spectral information. This classification technique mainly entails two distinct steps, namely 1) the automatic classification of pixels into a user-specified number of image classes according to their spectral properties; and 2) the manual assignment of the classes into information classes (Campbell 2007b). Although the automated nature of the spectral delineation renders this classification method less user-intensive, it cannot be completely considered truly unsupervised in nature. According to Mather (2004:203), it is rather an "exploratory" technique where repeated unsupervised area delineations with different parameters allow a user to ascertain which real-world classes are spectrally distinct and which are spectrally similar. This understanding of image features can inform the construction of the set of real-world classes to be used in the classification, rendering unsupervised classification extremely useful where *a priori* information regarding the study area or the classification structure is unavailable or not pre-determined. Conversely, where a real-world class structure is already established, it is rare that it will correspond with the automatically delineated spectral classes, resulting in the lowering of the accuracy of the outcome (Campbell 2007b). This is especially true for HR imagery where features of interest commonly comprise multiple spectral classes shared by more than one information class. This is the primary disadvantage of unsupervised classification, and, for this reason, it often has limited use in operational applications.

2.3.4.2 Supervised classification

Supervised classification is defined by the application of *a priori* information of real-world classes to determine the identity of unknown image elements. Data for the real-world classes are acquired from an external source and entered into the classifier in the form of designated and labelled polygons called 'training areas' or 'training data'. These training areas contain statistical information regarding the spectral properties of each class, which is used by a classification algorithm to identify the class of unknown pixels (Mather 2004; Campbell 2007b). Classification algorithms varies but are all designed to compare the features of each of the classes with those of an unknown pixel in geometric space and assign a class based on the results of that comparison. Traditionally, the most widely used algorithm is the maximum likelihood classification (MaxL) algorithm due to its ready accessibility, robustness, strong theoretical foundation and high accuracies for a wide range of RS applications (Bolstad and Lillisand 1991; Albert 2002; Brown de Colstoun et al*.* 2003; Pal and Mather 2003; Tseng et al*.* 2008). Because of these traits, a number of studies have used MaxL as the benchmark with which to compare newly developed classification methods (Gumbricht et al*.* 1996) (Hepner et al*.* 1990; Lui et al*.* 2002; Neusch and Grussenmeyer 2003; Pal and Mather 2003; Hagner and Reese 2007; Nangendo et al. 2007).

Recent developments in image resolution (spatial and spectral), increases in data availability, as well as the integration of contextual and ancillary data have prompted the use of more powerful classifiers which incorporate elements of machine learning (ML) (Tseng et al. 2008). While more traditional classifiers (such as MaxL) estimate parameters to a data distribution, ML classifiers are non-parametric and therefore do not make assumptions about the distribution of data (Jain et al. 2000; Hubert-Moy et al. 2001). This is especially advantageous when working with geospatial data, which, in most cases, are not normally distributed. Additionally, ML techniques can easily be automated, allow for combinations of categorical and continuous input variables and have the ability to capture hierarchical and non-linear relationships (Hladik and Alber 2014). Several ML algorithms are available in RS, namely k-nearest neighbour (kNN) (Franco-Lopez et al. 2001; Ying and Bo 2009; Falkowski et al. 2010), support vector machine (SVM) (Lizarazo 2008; Li et al. 2010; Petropoulos et al. 2012), decision tree (DT) (Punia et al. 2011; Gómez et al. 2012; Hladik and Alber 2014) and random forest (RF) (Gislason et al. 2006; Chang et al. 2008; Rodriguez-Galiano et al. 2012). Artificial neural networks (ANNs) were one of the first classifiers to draw on the field of ML in RS (Hepner et al. 1990; Skidmore et al. 1997) but were not widely employed due to their non-intuitive usability and black-box nature. However, neural networks (NN) have recently benefited from developments in deep learning (DL) technology, and RS is seeing the increased application of NNs, including convolutional NNs, multi-layer perceptron NNs, autoencoders and deep belief networks (Heydari and Mountrakis, 2019).

kNN is a simple non-parametric, distance-based classifier that labels each unknown instance based on its k neighbouring known instances. A class is assigned to the unknown instance best represented by the training samples among the k neighbours (Cover and Hart 1967; Gibson and Power 2000). The kNN algorithm is effective in classifying data that are not normally distributed but has the disadvantage of assigning equal weight to all variables even though certain variables may have higher priority. This can result in incorrect class assignments and diffuse clusters (Cunningham and Delany 2007). To avoid this, only odd k-values (e.g. 1, 3 and 5) should be used, as suggested by Campbell (2007a).

The efficiency of SVM classifiers for RS applications has been demonstrated by Lizarazo (2008) and Petropoulos et al. (2012). Myburgh and Van Niekerk (2013) showed that SVM produces more accurate results than kNN and MaxL for land cover mapping using SPOT-5 imagery. SVM determines the optimal separating hyperplane between classes (Novack et al. 2011) by focussing on the training samples close to the edge (support vector) of the class descriptors (Tzotsos and Argialas 2008). In cases where the relationship between classes and features are non-linear, the radial basis function kernel is often applied. See Vapnik (1995) and Huang et al. (2002) for a detailed mathematical formulation of SVM.

A DT identifies relationships between a continuous response variable, known as the dependent variable, and multiple, continuous variables known as the independent variables. DTs hierarchically split a dataset into increasingly homogeneous subsets known as nodes (Lawrence and Wright 2001; Gómez et al. 2012). By recursively splitting the feature datasets, a leaf node is reached, with the class associated with the node assigned to the observation (Pal and Mather 2003). According to Novack et al. (2011), each node is limited to a split in feature space orthogonal to the axis of the selected feature. Each branch of the DT consists of divisions (or rules) of the most probable class. Applying these rules will assign the most likely class to an unknown instance (Lawrence and Wright 2001).

There has been a notable increase in the use of the RF classifier for RS applications (Gislason et al.

2006; Lawrence 2006; Duro et al. 2012; Immitzer et al. 2012), and it has been shown to be effective for many classification tasks (Lawrence and Wright 2001; Rodriguez-Galiano et al. 2012). RF, an enhancement of DTs (Immitzer et al. 2012), generates each DT by using a random vector sampled independently from the input vector. A vote is cast by each of the generated DTs (Breiman 2001; Pal 2005; Bosch et al. 2007). Each classifier contributes a single vote to the assignment of the most popular class of the input variable (Breiman 2001; Rodriguez-Galiano et al. 2012). RF makes use of bagging (Breiman 1996), a method which generates a training set for feature selection. This allows RF classifiers to have a low (even lower than DT classifiers) sensitivity to training set size (Rodriguez-Galiano et al. 2012). Two parameters are required to be set, namely the number of trees and the number of active (predictive) variables. Rodriguez-Galiano et al. (2012) showed that stability in accuracy is achieved at 100 trees and that a small number of split variables are optimal for reducing generalisation errors and correlations between trees. A more detailed discussion of the RF classifier can be found in Breiman (1996), Breiman (2001), Pal (2005) and Rodriguez-Galiano et al. (2012).

ANNs are more complex than traditional statistical classifiers as they can model non-linear relationships. They contain three elements: an input layer, hidden layers and an output layer. The input layer contains the source data (imagery), hidden layers represent weights of association between classes and pixel values, and there can be many hidden layers. The output layer represents the classes for the desired output, which is defined by the training data during model building. The input data are passed through the network and weights are adjusted until the expected classification (defined by the training data) is achieved. Once the NN is established, the input data can be replaced with other data. The disadvantages of ANNs are that they are complex and prone to overfitting (Han al. 2018). DL occurs when a multi-layered NN is formed, creating a deeper network than conventional NNs (Devi Mahalakshmi and Geethanjali 2019). A convolutional NN (CNN) contains convolutional layers, max-pooling layers and fully connected layers. Filters are applied to the convolutional layers, the dimensionality of the data is reduced in the max-pooling layers, and the fully connected layers ensure that all of the input data in one layer are connected to all of the units of the next layer (Devi Mahalakshmi and Geethanjali 2019). NNs are advantageous as they can accept various numerical data, even if the data does not have a statistical distribution, allowing them to process ancillary data to remotely sensed data (Mather 2004). A major disadvantage of NNs is that a large amount of training data and computing power are required (Han et al. 2018).

2.3.4.3 Object-based image analysis

The development of classification methodologies has been enhanced by the advent of geographical object-based analysis (GEOBIA). Traditional methods of image analysis consider each pixel as an individual unit, with little cognisance of its topological relations to its neighbours or the class structure it represents (Lira and Maletti 2002; Van Coillie et al. 2007). This individuality of pixels renders them susceptible to data noise, atmospheric effects and surface variation (Wicks et al. 2002) and limits the usability of spectral, textural and relational information (Rego and Koch 2003; Lennartz and Congalton 2004; Oruc et al. 2004). Considering these factors, Blaschke et al. (2000) argue that no form of per-pixel classification can really yield reliable, robust and accurate results. In contrast, object-orientated imagery analysis operates on pre-defined areas of the image, derived either from an external source or, more commonly, an internal region-partitioning process known as segmentation (Blaschke et al. 2000). The increased availability of fine spatial resolution satellite imagery has exposed further limitations of per-pixel techniques, as for many applications the pixels of these images are significantly smaller than the objects of interest. In such cases, pixels often show spatial autocorrelation – the concept that features nearby are more similar than features further away – and will, therefore, belong to the same classes as their neighbours (Blaschke et al. 2000; Lang 2008). The object-based nature of OBIA is inherently better suited to the analysis and classification of HR imagery. According to (Benz et al. 2004; Bock et al. 2005; Hay et al. 2005; Shiba and Itaya 2006), OBIA uniquely offers meaningful statistical calculation of spectral and textural qualities, availability of feature qualities such as shape and object topology, intuitive spatial relations between real-world objects and image objects, and the ease of integration between GIS and RS environments and flexibility among different software platforms.

2.3.4.4 Image classification for forests

The incorporation of satellite imagery in forest inventories has improved the cost efficiency, development speed, timeliness, accuracy and level of detail (McRoberts and Tomppo 2007). Consequently, there is a large body of research in this field. Nery et al. (2019) undertook a six-class land cover/land use (LCLU) classification in order to determine changes to natural and plantation eucalypt forests in Western Australia. An unsupervised ISODATA classification was undertaken to identify potential LCLU training data, which was then inputted with derived texture measures to classification and regression tree (CART), SVM and RF for seven dates of Landsat MSS, TM, ET+ and OLI imagery. SVM showed the highest overall accuracies for the six LCLU classes (>84%) and were also superior in differentiating between the spectrally similar native and plantation eucalypt forest classes. Ahammad et al. (2019) applied a MaxL supervised classification of composited Landsat imagery for 2003 and 2014 to analyse natural forest and plantation change over time. Accuracies of 79% and 83% were achieved respectively, with the higher than expected classification error being attributed primarily to the small field sizes and the similarities between young teak plantations and non-forested areas. ‘Audah et al. (2019) applied a MaxL supervised classification of three years of Landsat OLI and ETM+ imagery for nutmeg plantation mapping in Indonesia with accuracies of over 95% for all dates. Koskinen et al. (2019) undertook CART, SVM and RF classifications of Landsat OLI, ALOS PALSAR and Sentinel-1 imagery, as well as the 30 m SRTM elevation and slope data, in order to map forest plantations and plantation genera on a regional scale in Tanzania. Applied in a Google Earth Engine environment, they found RF to be the most accurate, with overall accuracies of 85% for plantation/non-plantation classification, and 65% for the further differentiation into Pines, Eucalyptus/Wattle, Natural Forest and Other classes. McMahon and Jackson (2019) also found RF classification of Landsat TM, ETM+ and OLI suitable for *Eucalyptus* plantation mapping in Minas Gerias state, Brazil, with accuracies of over 85%. Lira Melo de Oliveira Santos et al. (2019) achieved plantation accuracies of over 85% when applying RF classification to composited Landsat ETM+ and OLI imagery for land cover mapping in Sao Paulo state, Brazil. In an application using DL, Wagner et al. (2019)

applied a U-net convolutional network on segmented Worldview-3 imagery to differentiate between natural forest and *Eucalyptus* plantations in Sao Paulo state, Brazil, with accuracies of over 95%.

In addition to mapping the locations and extents of forest compartments, the structural characteristics of forests such as height, volume, and basal area (Parker 2020) have been successfully captured using RS technologies such as photogrammetry (Campbell and Wynne 2013), light detecting and ranging (LiDAR) (Holmgren and Thuresson 1998), and structure from motion (SfM) (Tang and Shao 2015). Parker (2020) used LiDAR RS measurements in an individual tree crown algorithm to estimate crown area and tree height. The accuracies obtained for tree height estimations were 97%, and crown area had a root mean square error (RMSE) of 8.8 cm. Hu and Hu (2020) used Landsat images in a RF model to classify forest cover yearly from 1998 to 2015 to subsequently use in a change detection analysis to determine damaged and recovered areas. They obtained an overall accuracy of 86% and showed an overall loss of 0.56 X 10^6 ha of forests in Primorky Krai, Russia. Genera, species and clone mapping have been done using RS and ML, which helps with biodiversity management. For example, Mngandi et al. (2019) used Sentinel 1 SAR with Sentinel 2 multispectral data in a linear discriminant analysis algorithm to classify forest plantation species and obtained an overall accuracy of 87%. Data captured through RS can also be transformed into vegetation indices (VIs) to determine the health status of trees. This is possible as the chlorophyll in plants absorbs red light and strongly reflects radiation in the near-infrared (NIR) region of the electromagnetic spectrum (EMS). Researchers have successfully identified unhealthy trees using this method, and the methodology has directed foresters to apply fertilisers to specific problem trees (Tang and Shao 2015).

The next section details how some of the technologies overviewed in this knowledge review were employed in this project.

3 METHODS AND MATERIALS

This project was carried out in two phases, namely:

1. forest plantation mapping; and
5. quantify and analyse consumptive water use (actual ET) of forest plantations.

Figure 3-1 diagrammatically overviews these phases and the various activities within each phase. The methods used to complete these activities are detailed in the following subsections.

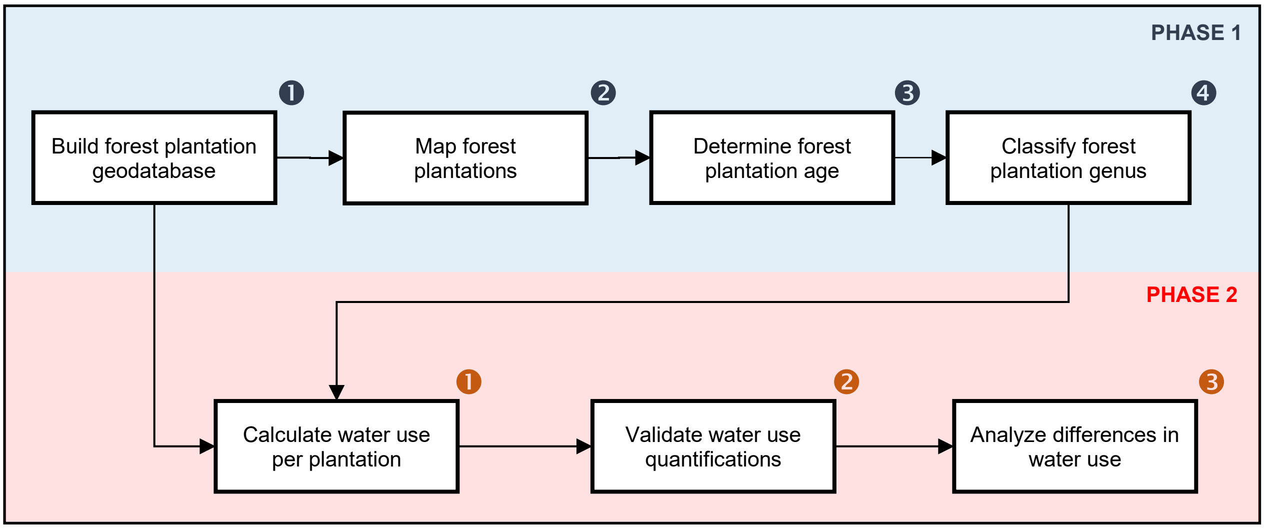


Figure 3-1 Project phases and research activities

3.1 Forest plantation geodatabase development

The first activity of Phase 1 was to obtain and collate as much spatially explicit plantation data as possible, preferably in GIS format. Detailed information about each plantation stand, including age, density, genus and clone information was targeted. This data had two purposes: 1) to be used as sample plantations in water use quantifications; and 2) to build and validate EO models for a range of forestry-related applications.

3.1.1 Data acquisition

Despite numerous requests, commercial forestry companies were reluctant to supply proprietary GIS data due to concern about confidentiality and the nature of the research (i.e. quantifying water use of forest plantations), which they saw as potentially damaging to the industry. Reference group member Dr R Heath of Forestry South Africa volunteered to act as intermediary between the research team and the forestry industry, and an agreement was finalised in October 2020. With the assistance of Dr Heath, datasets from four of the primary commercial forestry companies (Cape Pine, MTO, Mondi and Sappi) were provided to the research team by the Institute for Commercial Forestry Research (ICFR). The datasets comprised GIS shapefiles, including spatial (position, shape) and attribute (genus, species, age, etc.) data for 39 219 plantation compartments throughout South Africa. This dataset is a tremendous asset and is representative of much of the forestry activities in the country.

3.1.2 The commercial forestry geodatabase

The individual shapefiles from each commercial company were standardised to Albers Equal Area (25°E; -27°S and -31°S) projection and merged in one geodatabase (GDB). Each forestry company uses slightly different label formats for compartment attributes, of which the most important ones for this study were genus, species and planting date. Between the four companies, over 120 unique species codes were used to represent genus, species and hybrid, which were standardised with the assistance of Dr I Germishuizen of the ICFR (e.g. "*Enit"* = "*Eucalyptus nitens"*). Thirteen codes representing 226 compartments were not identifiable and were consequently disregarded. Dr Germishuizen also suggested that species with fewer than 20 compartments be discarded, as these were most likely planted for research purposes and were unlikely to be commercially active (a complete compartment species count is available in Appendix I). Additionally, for each compartment, the area in hectares was calculated using GIS, and the age in years was calculated using the planting date (not available for ~5% of cases). Other attribute labels e.g. plantation and compartment names, were also standardised as far as possible within the GDB. The final product was referred to as the commercial forestry database (CFDB), and the distribution of compartments according to genera can be seen in Figure 3-2 to Figure 3-5.

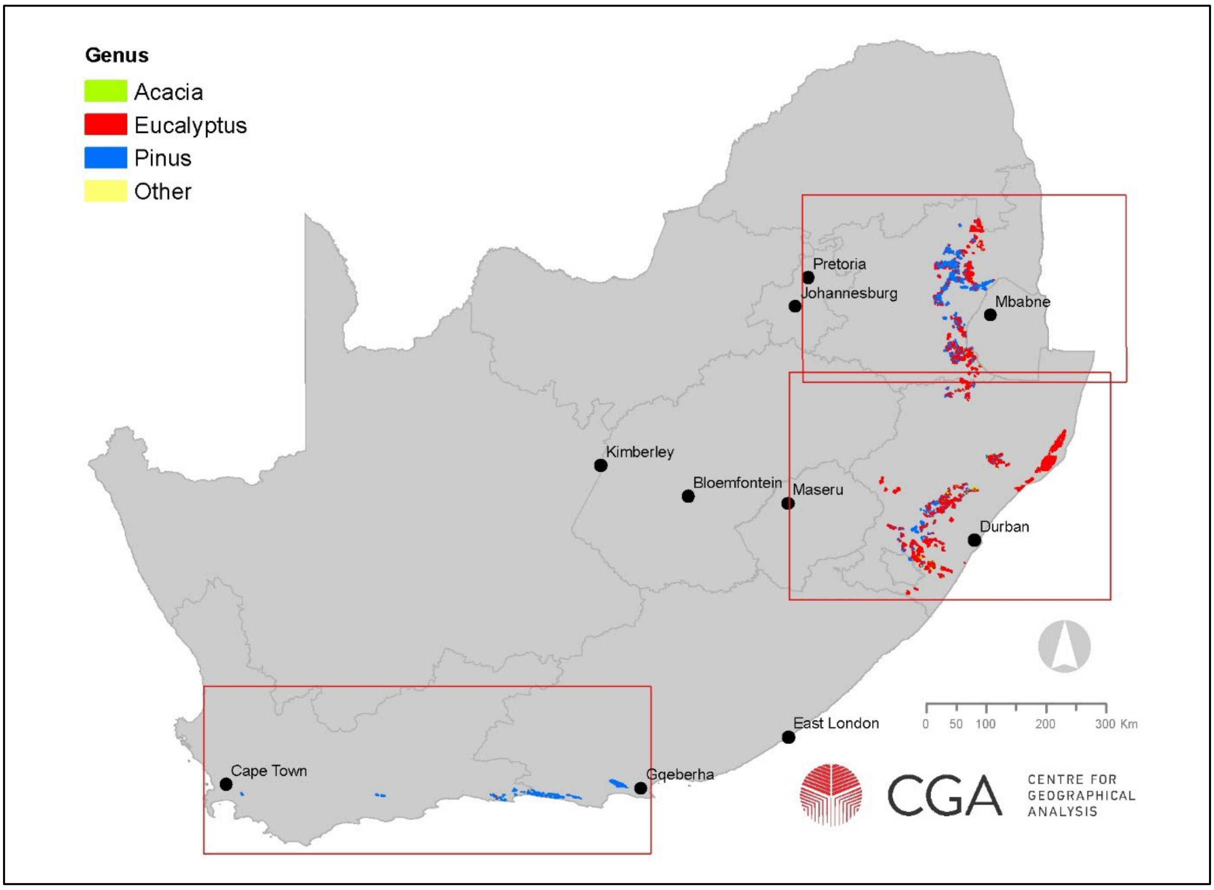


Figure 3-2 Distribution of compartments in the commercial forestry database in South Africa (inset maps below)

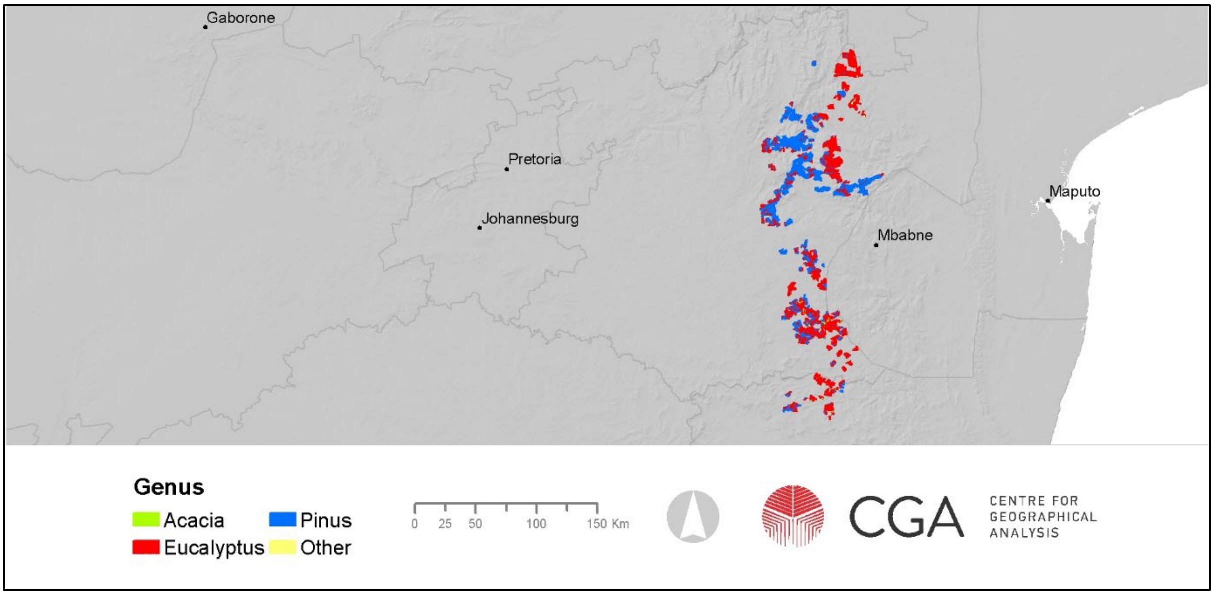


Figure 3-3 Distribution of compartments in of the commercial forestry database in Mpumalanga

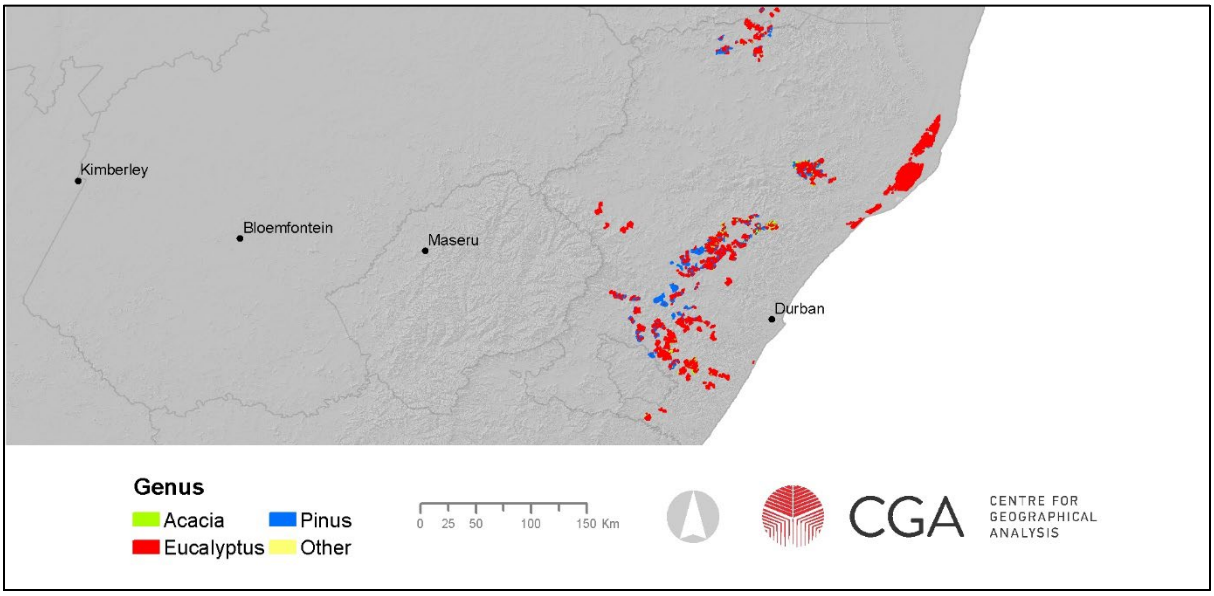


Figure 3-4 Distribution of compartments in the commercial forestry database in KwaZulu-Natal

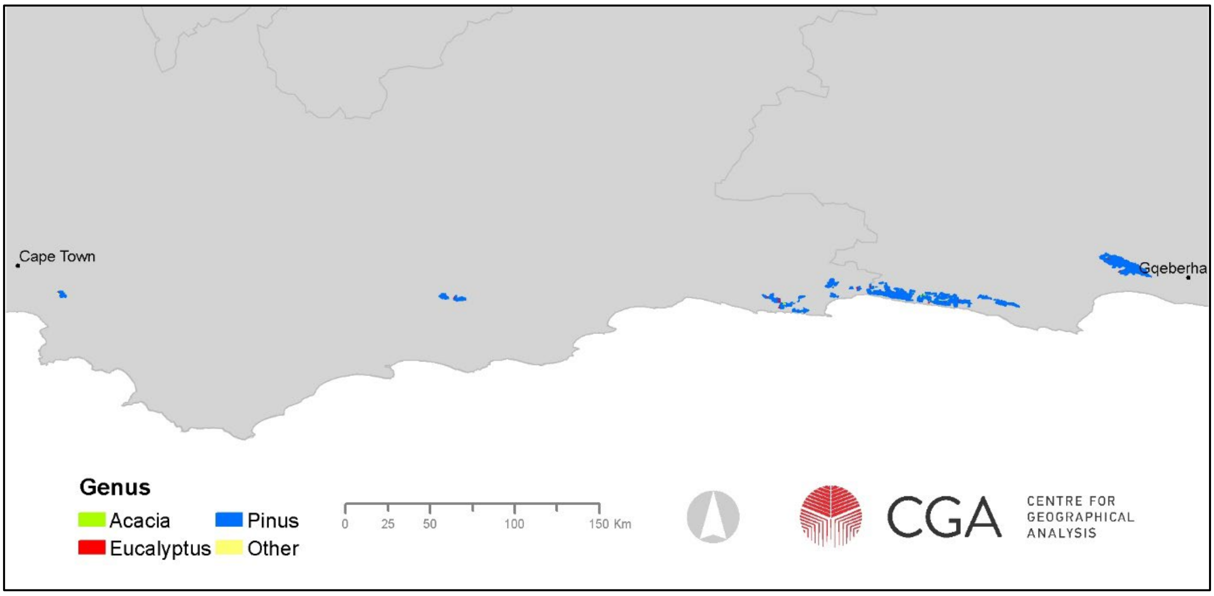


Figure 3-5 Distribution of the commercial forestry database in the Eastern and Western Cape

3.2 Remote sensing data collection

3.2.1 Optical imagery

While optical imagery was not directly used in the water use calculations, it was used for the supplementary analyses, including calculating normalised difference vegetation index (NDVI) profiles (Appendix II), and the plantation forest mapping (Section 4.1), genus classification (Section 4.2) and compartment age estimation (Section 4.3) research.

3.2.1.1 MODIS

The National Aeronautics and Space Administration (NASA) offers readily available MODIS-based products through the Distributed Active Archive Centres (DAACs). These centres process, archive, document and distribute data from NASA's past and current research satellites and field programmes. NDVI images are available as 250 m 16-day composites (MOD13Q1), generated using the two 8-day composite surface reflectance images from the MOD09A1 version 6 product in the 16-day period. For this project, the MOD13Q1 version 6 NDVI imagery from 2001 to 2020 for the whole of South Africa was sourced from the Application for Extracting and Exploring Analysis Ready Samples (AppEEARS) DAAC data portal[8].

3.2.1.2 Sentinel-2

The European Space Agency (ESA) offers freely available Sentinel-1, Sentinel-2, Sentinel-3 and Sentinel-5P user products through the Copernicus Open Access Hub[9]. Sentinel-2 multispectral imagery was used for the forest and genus mapping research. The spatial resolution ranges from 10 m to 60 m, depending on which bands it contains (see Table 2-2). These bands can then be used as is, and, as

[8] https://lpdaac.usgs.gov/tools/appeears/
[9] https://scihub.copernicus.eu/dhus/#/home

was the case for the mapping research, various vegetation indices can be derived from combining bands such as NDVI and enhanced vegetation index (EVI). The product package also includes the necessary metadata to do atmospheric corrections on the images, removing the effects of the atmosphere on the reflectance values of the images.

3.2.2 Evapotranspiration

The primary ET dataset used to calculate water use was the WaPOR actual ET data, although several other ET datasets were acquired for comparison purposes. Each are discussed in more detail in the following subsections.

3.2.2.1 WaPOR data

The Water Productivity Open-access Portal[10] (WaPOR), funded by the Food and Agricultural Organization of the United Nations (FAO), provides freely available water productivity data for Africa and the Middle East on an annual, monthly and decadal-scale. Included in these datasets are 250 m annual and monthly actual evapotranspiration and interception data (AETI; referred to as 'ET' in this study) dating back to 2009. Generated by a Dutch company, eLEAF, using the ETLook model, the dataset of 250 m ET data span over 11 years and covers the whole of South Africa. It is ideal for water use quantification studies. The monthly ET WaPOR rasters from January 2009 to October 2020 served as the primary data for water use calculations undertaken in this study.

3.2.2.2 WRC 2014/15 ET dataset

This dataset is similar to the WaPOR data, in that it was developed by Dutch company eLEAF using the SEBAL and ETLook models, though specifically for WRC Project TT 745/17 (Van Niekerk et al 2018). The dataset consists of ET data at 250 m spatial resolution at monthly intervals (mm/month), from August 2014 to July 2015, and was continuously calibrated throughout the project to account for the highly diverse climatic regions of South Africa. Further details can be found in WRC report TT 745/17 (Van Niekerk et al 2018).

3.2.2.3 FruitLook

FruitLook, an online RS analysis service funded by the Western Cape Department of Agriculture (WCDOA), assists farmers in agricultural decision-making at field level. A core component of this service is the provision of actual ET data. All FruitLook ET datasets are generated by eLEAF using the ETLook and METEOLOOK models (Bastiaanssen et al. 2012; Pelgrum et al. 2011) and satellite imagery from various platforms (Landsat 5, 7 and 8, MODIS, VIRRS, Deimos, UK-DMC-2, Sentinel 2). The satellite data and data processing are described in more detail by Goudriaan (2014) and Jarmain et al. (2011).

FruitLook ET datasets are available as a series of 20 m spatial resolution images provided at weekly intervals on a seasonal basis since 2011. Unlike WaPOR, temporal coverage varies from season to

[10] https://wapor.apps.fao.org/home/WAPOR_2/1

season, with coverage from October to April in the 2011/12 to 2015/16 seasons, August to April in the 2017/18 season, and August to July in the 2018/19 season. The spatial coverage of the images is also limited, covering only fruit and wine producing areas of the Western Cape (WC) in the 2011/12 to 2015/16 seasons, but expanded to cover the majority of the WC from the 2017/18 season onwards. For comparison purposes, this dataset was converted to a monthly dataset (mm/month).

3.2.2.4 MOD16 ET data

The MODIS Global Evapotranspiration Project (MOD16), developed and funded by NASA and the United States Geological Survey (USGS), provides freely available ET data since 2000 through the AppEEARS DAAC data portal[11]. This dataset is created using inputs of daily meteorological reanalysis data along with MODIS remotely sensed data products such as vegetation property dynamics, albedo and land cover (Running et al 2017). ET values are derived using an algorithm that is based on the logic of the Penman-Monteith equation (Running et al 2017) and is available at 500 m spatial resolution with 8-day cumulative ET values (mm/8-days). For comparison purposes, this dataset was converted to a monthly dataset (mm/month) for the months August 2014 to April 2015. This was done by reducing the 8-day composites to an ET value per day of the year and adding up the relevant daily values of each month.

3.3 Quantifying water use per compartment

The first step for the quantification of water use analysis was to extract the monthly WaPOR ET values for each plantation compartment. Statistically, the strongest results would be derived from analysing all 39 219 compartments in the CFDB; however, two factors limited the total number of viable samples. First, samples were reduced by temporal limitations, summarised by the age of the compartment relative to the water use before and after maturity, and relative to the 10-year limitation of the ET data. Second, compartment size was often much smaller relative to the size of the WaPOR ET image pixels (250 x 250 m), leading to the likely influence of land covers surrounding these compartments on the water use calculations. A sensitivity analysis was conducted to determine the effects of this second factor. Data extraction, temporal limitations and the results of the sensitivity analysis are discussed in further detail in the following sections.

3.3.1 Data extraction

The monthly WaPOR ET data were extracted for all compartments in the CFDB using a GIS process called zonal statistics. Here, a chosen statistical value (e.g. average) is calculated from the pixels of a given raster (e.g. ET for January 2009) intersecting a series of given zones (e.g. compartment polygons). A visual example of the application of zonal statistics for one compartment is illustrated in Figure 3-6.

[11] https://lpdaac.usgs.gov/tools/appeears/

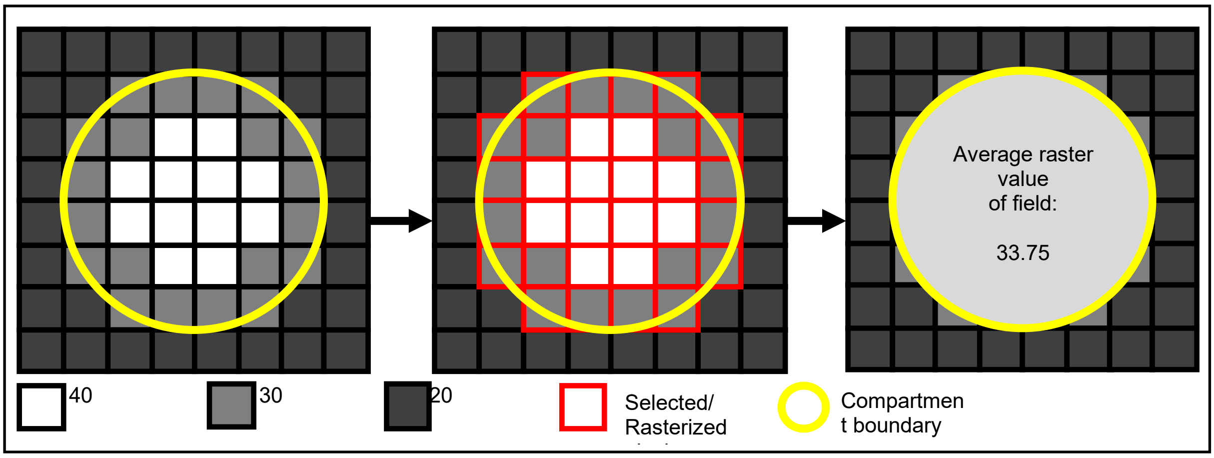


Figure 3-6 Conceptual overview of using zonal statistics to calculate the average raster value for one compartment

Zonal statistics were used to extract the WaPOR ET for every month of the year for the years 2009 to 2019 for each of the 39 219 compartments.

3.3.2 Temporal limitations

It is known that forest age affects ET (i.e. young forests generally use less water), so the samples should preferably be representative of compartments ranging from planting and increasing in age to maturity. However, since water use calculations rely on the ET data and the WaPOR ET dataset only extends to 2009, only ten years of compartment water use could be analysed over time. To remove the effects of seasonality (especially as areas of South Africa experienced drought over the last decade), only the compartments planted in 2009 were used in the analysis. This limited the number of viable compartments from 39 219 to 1 415 but ensured the water use calculations were applied to the first ten years of compartment growth (unaffected by the seasonal variation that would have been present with variable planting years). Figure 3-7 shows the ages of the compartments in the CFDB.

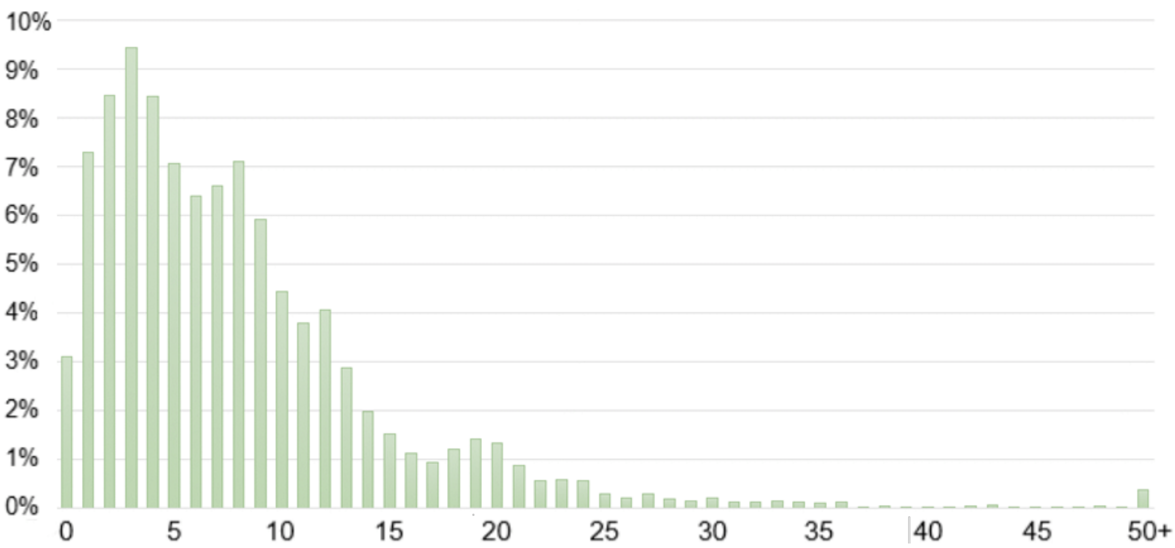


Figure 3-7 Histogram of compartment (n = 39 216) age in the Commercial Forestry GDB

3.3.3 Sensitivity analysis

The disparity between the average size of plantation compartments (12.6 ha) and ET pixels (6.3 ha), as well as the misalignment between compartments and pixel shape (Figure 3-8), raised concerns that the average ET values extracted per compartment would not be an accurate or 'pure' representation of ET in that compartment. This is especially detrimental where compartments are adjacent to water bodies, as was observed in Van Niekerk et al (2018). Consequently, a sensitivity analysis was undertaken prior to water use calculations to identify the compartments that are suitable for analysis.

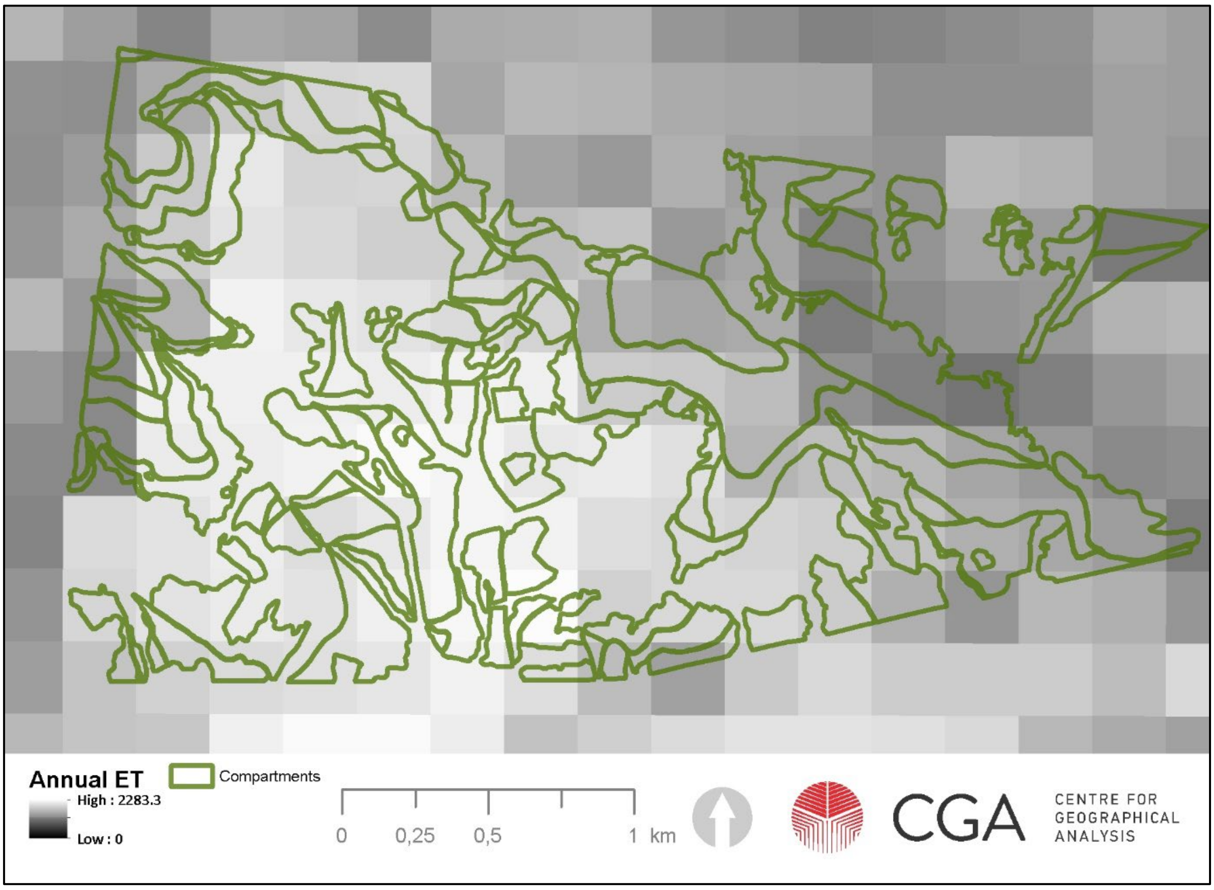


Figure 3-8 A comparison between compartments (green lines) and ET pixels (greyscale background)

Ideally, the effects of mixed pixels during ET extraction should have a negligible effect on the water use calculations. A reference dataset of 'pure pixels' was required to determine this.

3.3.3.1 Reference (pure pixel) ET dataset

The first ET dataset developed was in line with the theory that pixels that were not mixed (i.e. 'pure') would contain the true ET values of that compartment. Subsequently, all compartments in the CFDB that contained entire pixels were identified and the ET of each pixel was extracted (Figure 3-9). Due to the smaller size and irregular shapes of the compartments compared to the pixels, this resulted in a 'pure pixel' reference dataset of 3 035 out of the original 39 216 compartments. The majority of the

3 035 compartments contained only one entire pixel, but where there were more than one pixel, the ET values were averaged. In other words, no mixed pixels were considered in the extraction of ET values, resulting in a reference dataset of compartments of ‘rue’ ET values (called Sample A).

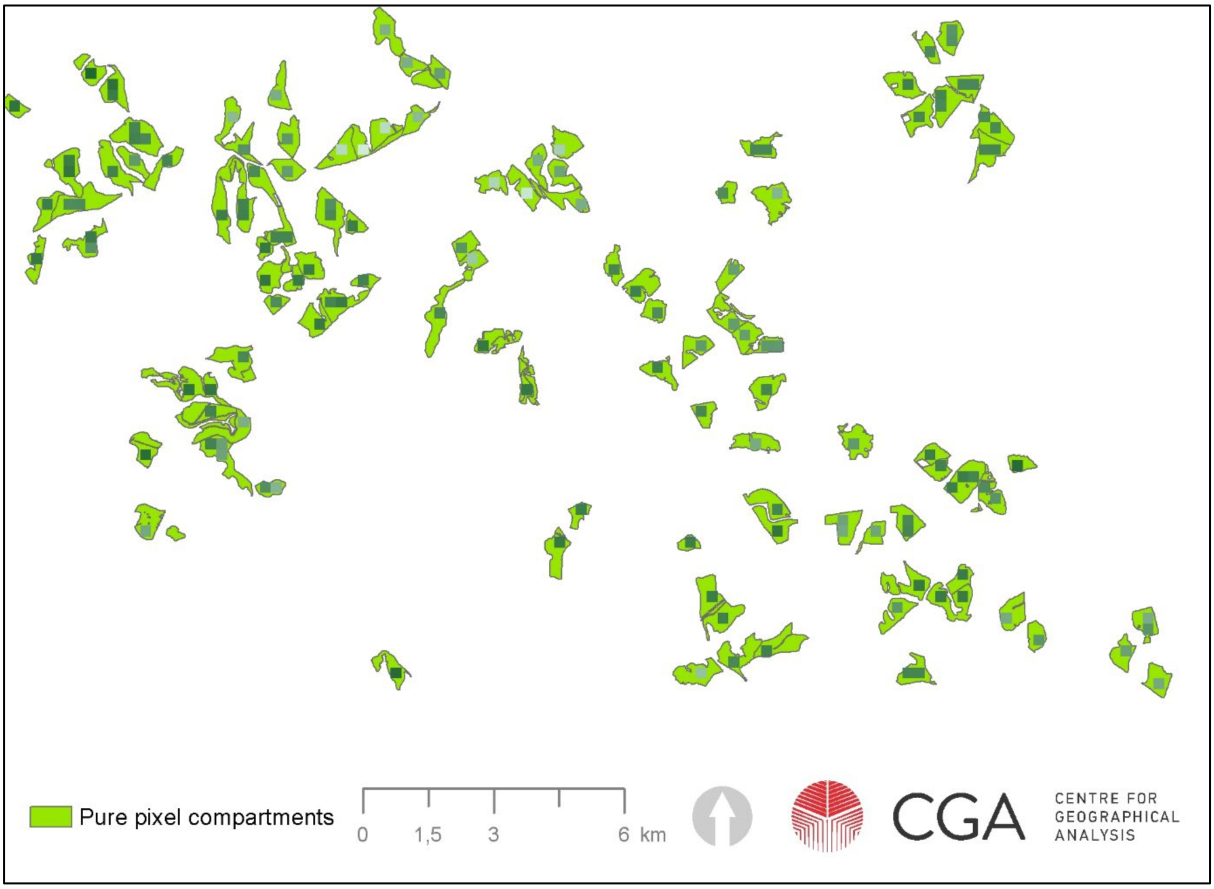


Figure 3-9 Example of compartments containing pure (unmixed) pixels used to develop the reference dataset

Only 82 compartments from the 3 035 out of 39 216 (7.7%) compartments containing pure pixels were planted in 2009. This was judged far too few to conduct meaningful statistical analysis on the water use among genera, species, climate and terrain. However, this dataset (called Sample B) was deemed suitable as a ‘pure pixel’ reference dataset to determine the effects of different approaches selecting viable compartments.

3.3.3.2 Rasterised ET dataset

The least conservative approach to ET extraction would be to make use of all pixels that intersect the compartments, even if only by the smallest fraction. This would entail the selection of all 39 219 compartments, which would also maximise the mixed pixel effect. Instead, a compromise between using all compartments and using only pure pixel compartments was attempted with the use of rasterisation.

Rasterisation involves the conversion of vector data (compartments) to a raster (ET pixels). In this instance, compartment polygons were rasterised to exactly match the cell grid of the 250 m ET images (Figure 3-10). Pixels were included in compartment conversion only if the majority (>50%) of the pixel fell within that compartment. The hypothesis of this approach was that the impact on ET by surrounding land covers would always be half or less than it would be if mixed pixels were included. The rasterisation of the 32 216 compartments resulted in 26 102 compartment polygons, with the reduction due to no pixel majority falling within a compartment (often very small or thin compartments). Of the 1 415 compartments planted in 2009, 876 rasterised compartments were retained. This number of compartments was judged suitable for the water use analysis, provided the mixed pixel effect on the ET values per compartment was not significant.

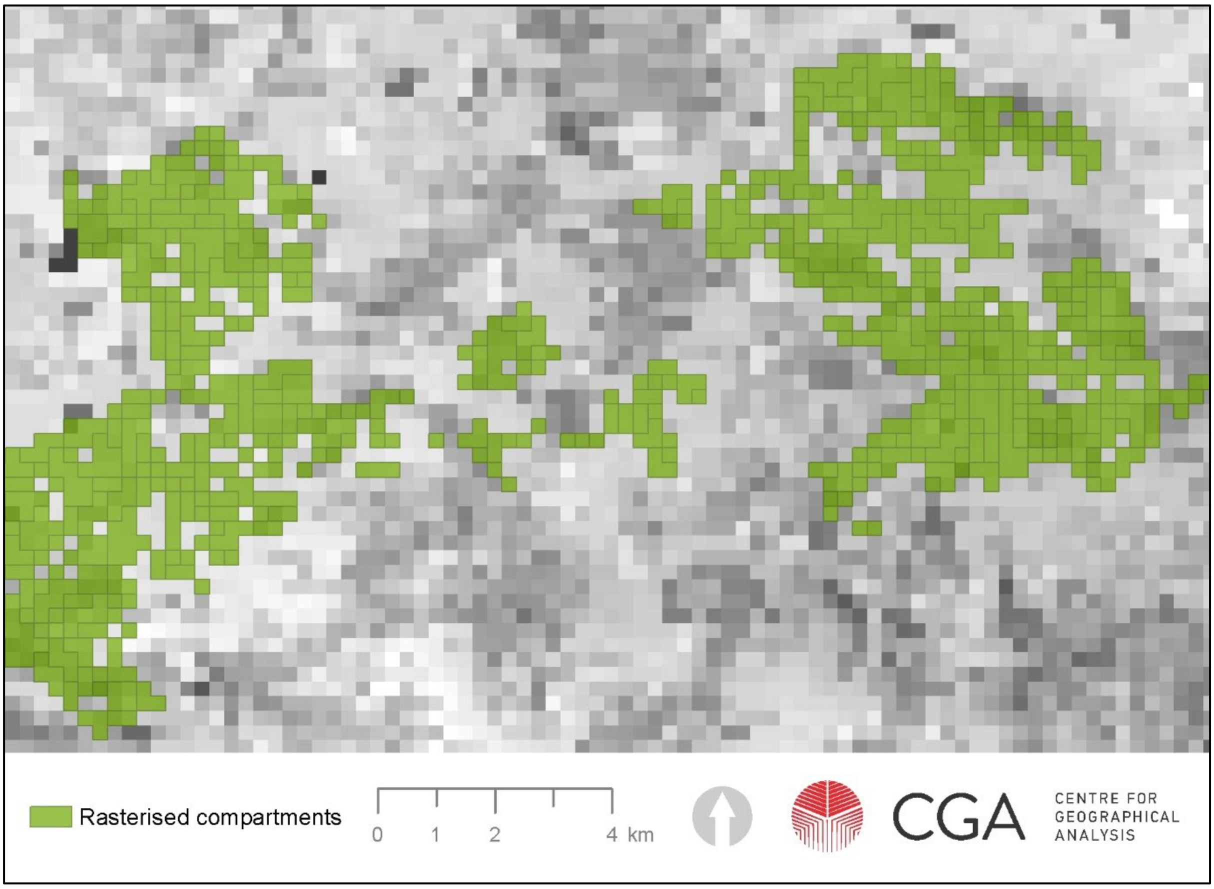


Figure 3-10 Example of rasterised compartments

3.3.3.3 Statistical analysis

The extracted ET values of the reference (pure pixel) and rasterised datasets were compared compartment-for-compartment using t-tests, time series graphs and histogram comparisons to quantify the impact of the mixed pixel effect during rastersiation. A t-test is a statistical hypothesis test used to determine whether two samples are significantly different from one another. The results of a t-test are returned as a probability value (p-value) that the means of the two datasets are significantly different, which is compared against a significance level specified by the tester. If the p-value is lower than the significance level, the null hypothesis is rejected and the difference is considered statistically significant. In this case, the null hypothesis was defined as: there is no significant difference between the reference (pure pixel) and rasterised datasets. A significance level of a conventional 5% (i.e. there was a 5% probability of rejecting the null hypothesis when it is true) was set.

Two t-tests were conducted. The first t-test compared the ET values of all 3 035 compartments of the reference dataset to that of the rasterised dataset (Sample A) on a month-by-month basis from 2009 to 2019. For example, the 3 035 compartments for January 2009 in the reference dataset were compared to the same 3 035 compartments for January 2009 in the rasterised dataset. The resulting p-values per month are shown in Table 3-1.

Table 3-1 P-values per month for reference vs rasterised dataset comparison (Sample A)

	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
January	0.66	0.69	0.72	0.76	0.78	0.67	0.84	0.50	0.94	0.78	0.41
February	0.85	0.91	0.83	0.58	0.74	0.57	0.62	0.31	0.85	0.94	0.64
March	0.62	0.89	0.82	0.96	0.87	0.57	0.68	0.45	0.97	0.60	0.99
April	0.84	0.82	0.91	0.91	0.59	0.74	0.65	0.64	0.85	0.57	0.65
May	1.00	0.90	0.88	0.90	0.53	0.80	0.54	0.64	0.60	0.98	0.42
June	0.69	0.65	0.77	0.89	0.82	0.68	0.52	0.69	0.55	0.82	0.50
July	0.63	0.36	0.49	0.42	0.72	0.45	0.33	0.58	0.81	0.87	0.69
August	0.23	0.12	0.14	0.07	0.27	0.25	0.09	0.22	0.42	0.33	0.07
September	0.14	0.08	0.10	0.07	0.15	0.18	0.12	0.11	0.09	0.16	0.11
October	0.18	0.11	0.13	0.13	0.13	0.18	0.24	0.16	0.23	0.20	0.08
November	0.22	0.18	0.13	0.23	0.23	0.32	0.28	0.25	0.35	0.15	0.10
December	0.52	0.36	0.22	0.69	0.64	0.49	0.26	0.62	0.51	0.11	0.20

The p-values were greater than the specified 0.05 significance level for all months, indicating that all tests failed to reject the null hypothesis. In other words, there was no significant difference between the ET values in the 3 035 compartments in the reference and rasterised datasets in any of the 137 months analysed.

To further confirm these findings, a second t-test was conducted specifically on the 82 pure pixel compartments planted in 2009 (Sample B). This sample consisted of 45 *Eucalyptus*, 36 *Pinus*, and one *Acacia* compartment. The two primary genera, *Eucalyptus* and *Pinus*, were compared separately to determine if genera affected the ET mixed pixel sensitivity (*Acacia* was discarded due to too few compartments). For all compartments in both datasets (Sample A and B), the ET values for each month from January 2009 to December 2009 were averaged. All 137 months of the reference datasets were then compared to the rasterised dataset in order to obtain a p-value for each genus. The p-values obtained for *Eucalyptus* and *Pinus* compartments were 0.56 and 0.91, respectively. In both cases the p-values were greater than the specified 0.05 significance level, indicating both tests failed to reject the null hypothesis. In other words, there was no significant difference between the pure pixel compartment ET values between reference and rasterised datasets for either *Eucalyptus* or *Pinus*.

These results are supported by a graphical analysis of the two datasets. Figure 3-11 and Figure 3-12 show the monthly time series of ET for the reference and rasterised datasets for *Eucalyptus* and *Pinus*, respectively.

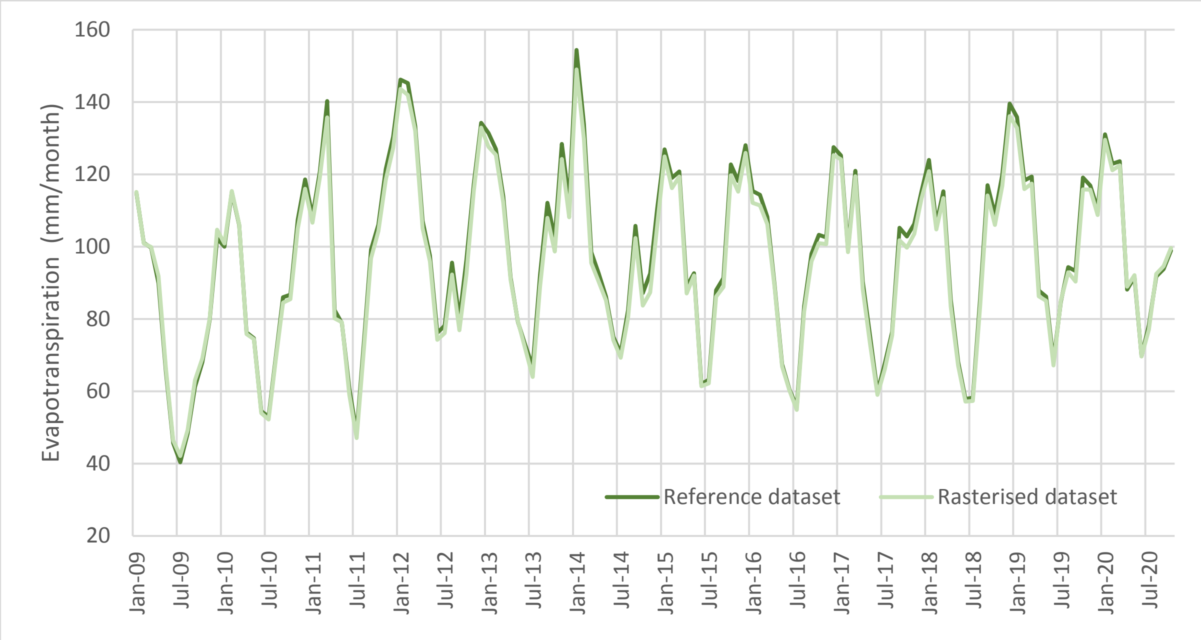


Figure 3-11 ET time series reference vs rasterised comparison for *Eucalyptus*

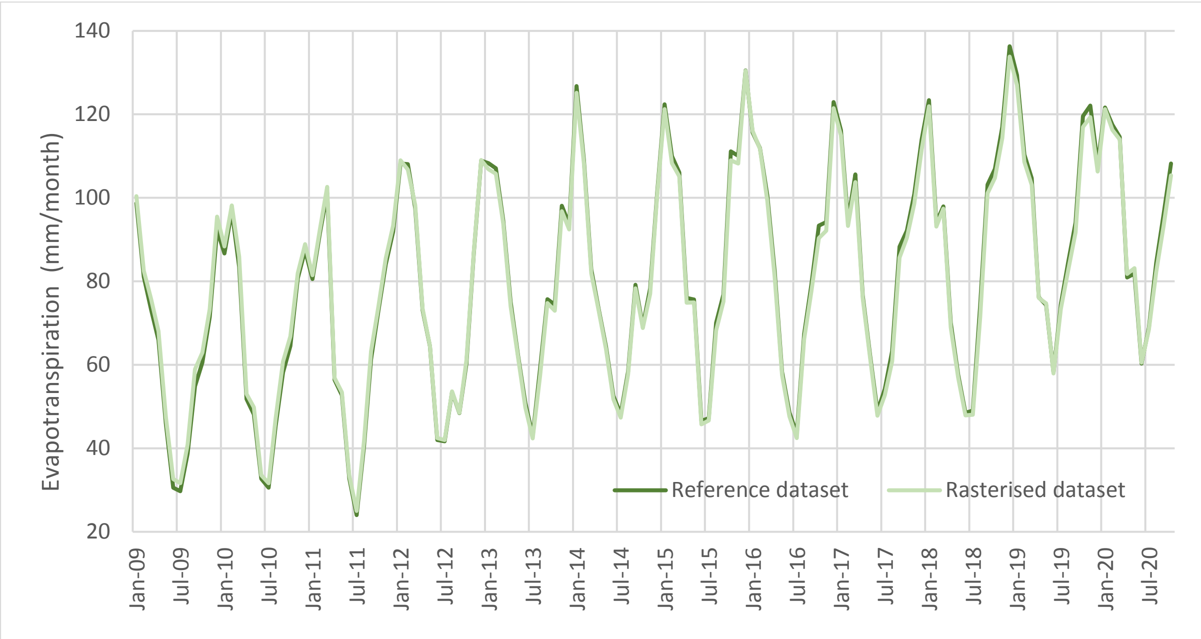


Figure 3-12 ET time series reference vs rasterised comparison for *Pinus*

Visual assessment of the ET time series of the genera reveals very little difference between the reference and rasterised datasets. This is further supported when assessing the histograms of the data. Figure 3-13 and Figure 3-14 show a histogram comparison of the monthly ET values between the reference and rasterised datasets for all compartments for June and December 2019.

As can be seen in Figure 3-13, the distribution of ET values between the two datasets matches fairly closely. The exception is of compartments with 20-30 mm/month actual ET, which were reduced during the rasterisation process. Despite this, overall actual ET averages between the two datasets during the winter months are minimal, with 62.4 mm/month and 62.8 mm/month for the reference and rasterised, respectively.

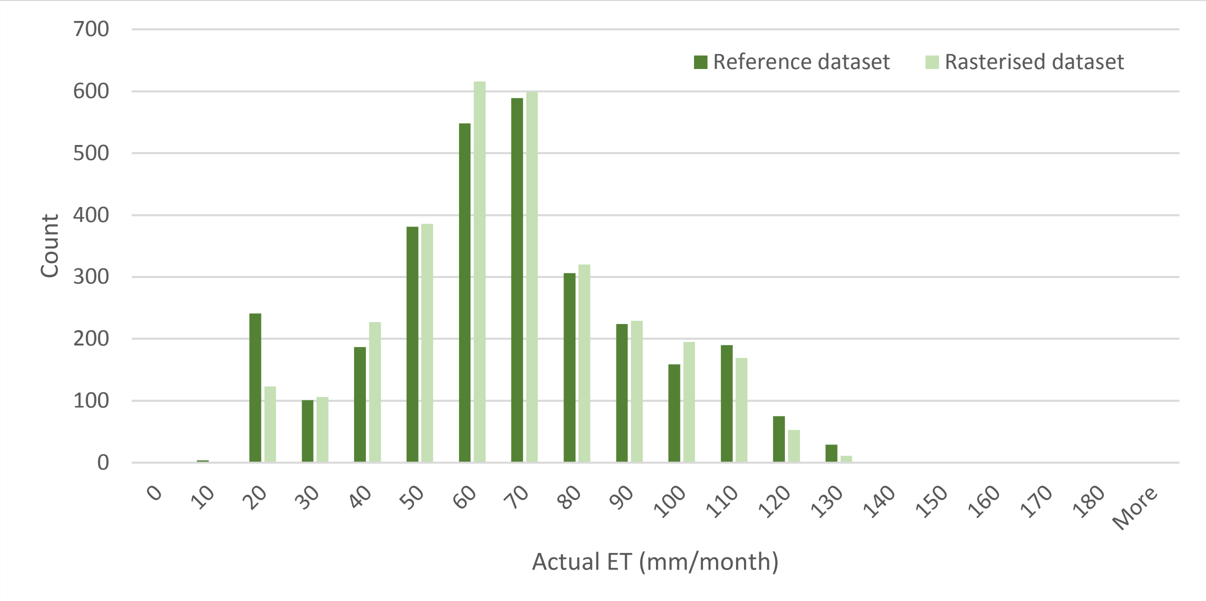


Figure 3-13 Actual ET histogram for June 2019

As can be seen in Figure 3-14, the distribution of compartment ET values between the two datasets matches closely. A similar trend was seen in June, where compartments with more extreme ET values are slightly reduced during the rasterisation, while compartments closer to the mode are slightly increased. Again, the effect on the overall actual ET averages is negligible, with 101.5 mm/month and 100.6 mm/month for the reference and rasterised datasets, respectively.

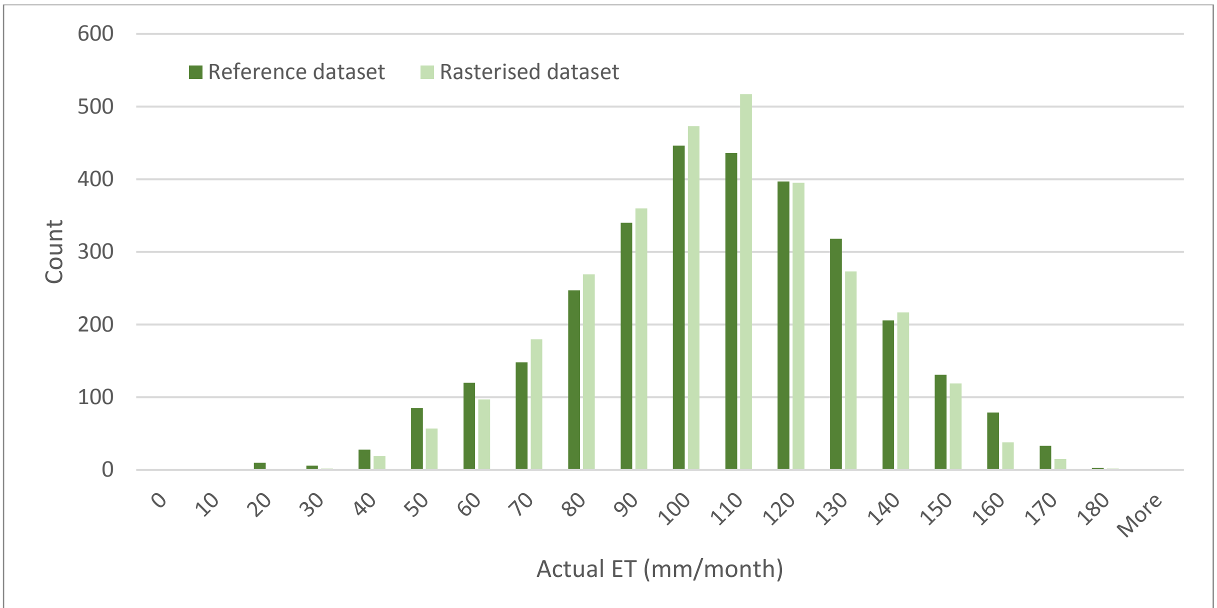


Figure 3-14 Actual ET histogram for December 2019

In conclusion, the sensitivity analysis definitively shows that the mixed pixel inclusion during the rasterisation of the compartments resulted in minor but statistically insignificant differences in the ET values that were extracted in each compartment. Additionally, the constraint imposed by rasterisation retained enough (876) compartments planted in 2009 to provide a statistical representation of the ET, even when analysed by genera, age, terrain or climate. The rasterised compartment dataset was therefore used in all the water quantification calculations described in the following sections.

3.3.4 ET dataset comparison

The WaPOR ET dataset was selected for the water use calculations in this study due to its national coverage, more than 10-year timespan, recognised ET modelling approach (ETLook) and the fact that it is freely available. Other ET datasets covering South Africa are the WRC 2014/15 ET dataset developed for WRC Project TT 745/17 (Van Niekerk et al 2018) and the MODIS Global Evapotranspiration (MOD16) dataset. Each dataset uses different empirical models and ancillary data to derive ET values and are available for different spatial extents, timespans and frequencies, and at varying spatial resolutions (see Section 3.2.2 for more details). Although the WaPOR (250 m), WRC 2014/15 (250 m) and MOD16 (500 m) data cover the entire South African territory, the analysis period was, however, limited by the WRC dataset, which is only available for one season. Using the 26 102 rasterised compartments (Section 3.3.3), zonal statistics was used to extract the average ET values from all three datasets on a monthly timestep from August 2014 to July 2015. The ET values extracted from the MOD16 product were in some cases unrealistic due to some anomalies (inconsistent 'nodata' values). The data were consequently cleaned by removing all negative values and all values above 200 mm/month.

Figure 3-15 compares a time series of the mean ET per month for each dataset. The WRC 2014/15 ET dataset is likely the most accurate of the three, as it was calibrated using seasonal climatic data captured by 239 weather stations around the country. The WaPOR values are higher than those of the WRC 2014/15 dataset during autumn, winter and spring (especially September), but the two datasets correspond well during summer. The MOD16 ET values are much higher than those of the other two datasets during summer but corresponds well with the WRC 2014/15 dataset during the autumn and winter. It is noticeable that the monthly ET values are always higher than 40 mm/month, even though the period was very dry. This is because these values represent the mean of multiple stands. It is possible that some of the stands had very low, or even zero, ET during this period.

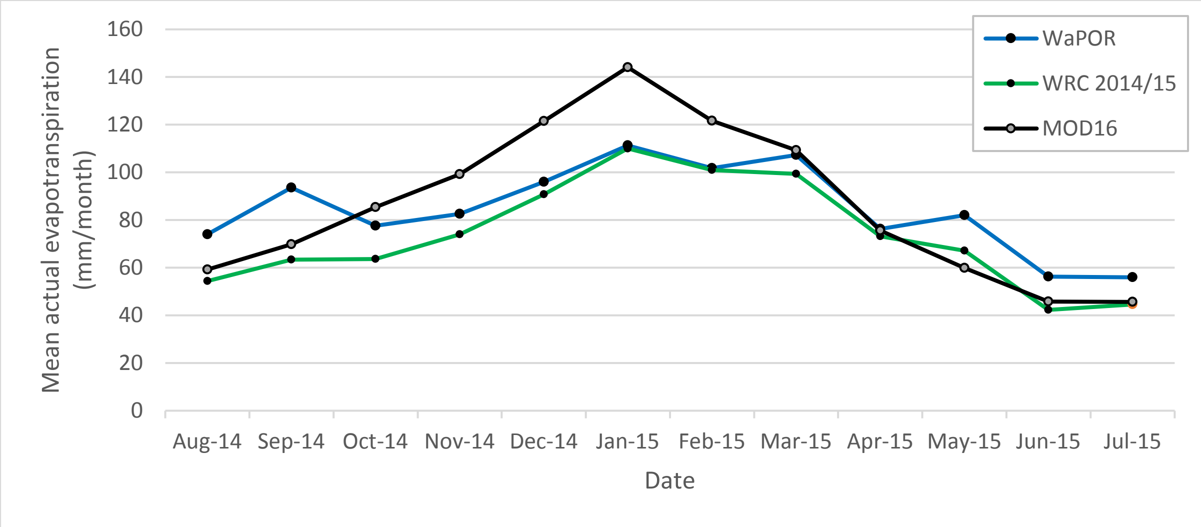


Figure 3-15 WaPOR compared to the WRC 2014/15 and MOD16 mean monthly ET

The ET histograms for selective months of each season (October, January, April and July) are shown in Figure 3-16 to Figure 3-19. It is clear that the MOD16 ET estimates for summer (January) was much higher than those of the other two datasets, with more than 40% of the compartments having monthly ET values exceeding 150 mm/month. MOD16 estimates for spring (October) were also higher than those of the other datasets (up to 150 mm/month), while most of the WRC 2014/15 ET values range between 40 and 90 mm/month.

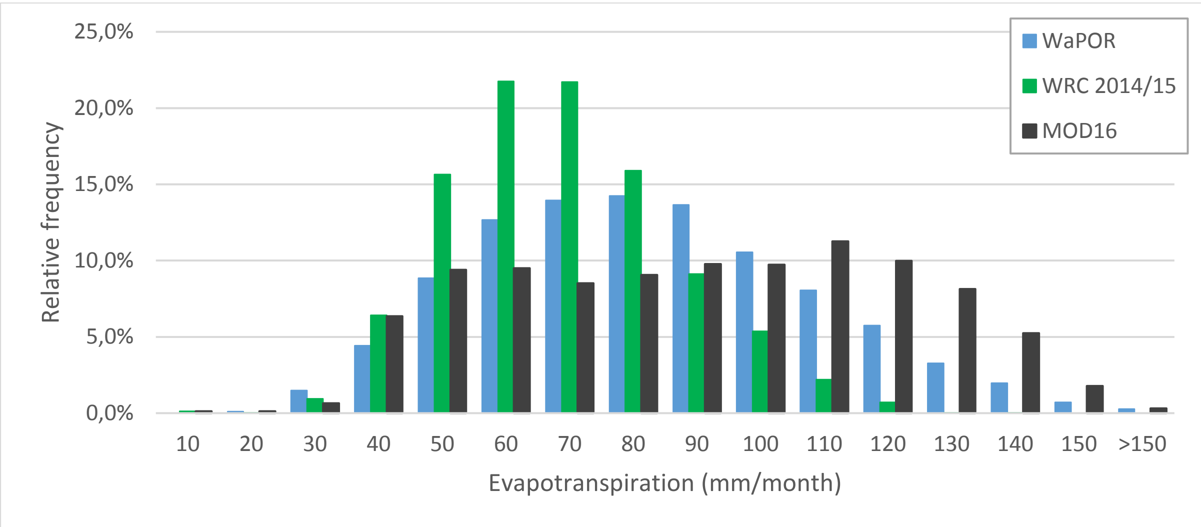


Figure 3-16 WaPOR, WRC 2014/15 and MOD16 ET histogram comparison for October 2014

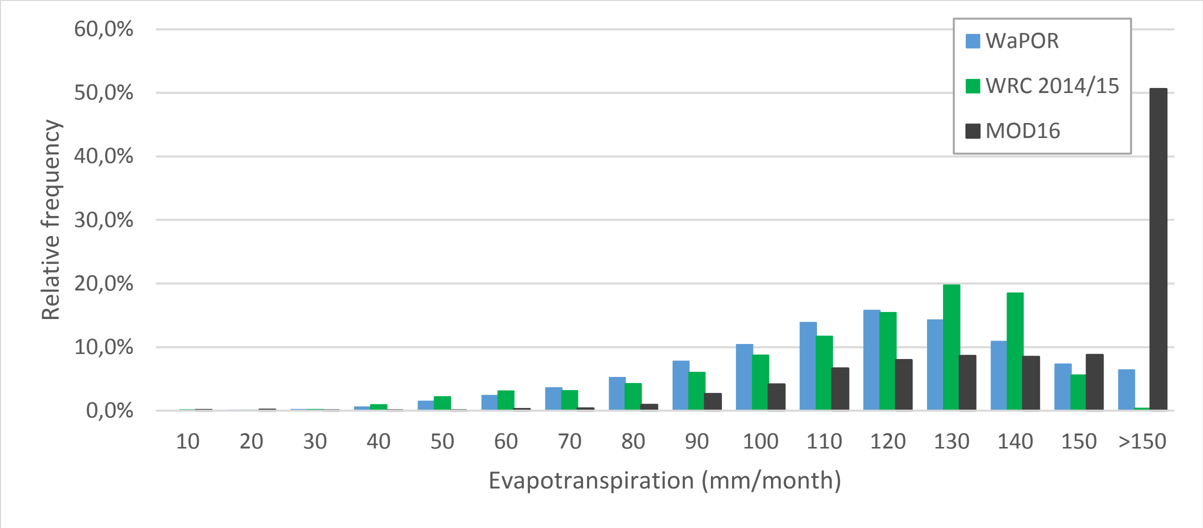


Figure 3-17 WaPOR, WRC 2014/15 and MOD16 ET histogram comparison for January 2015

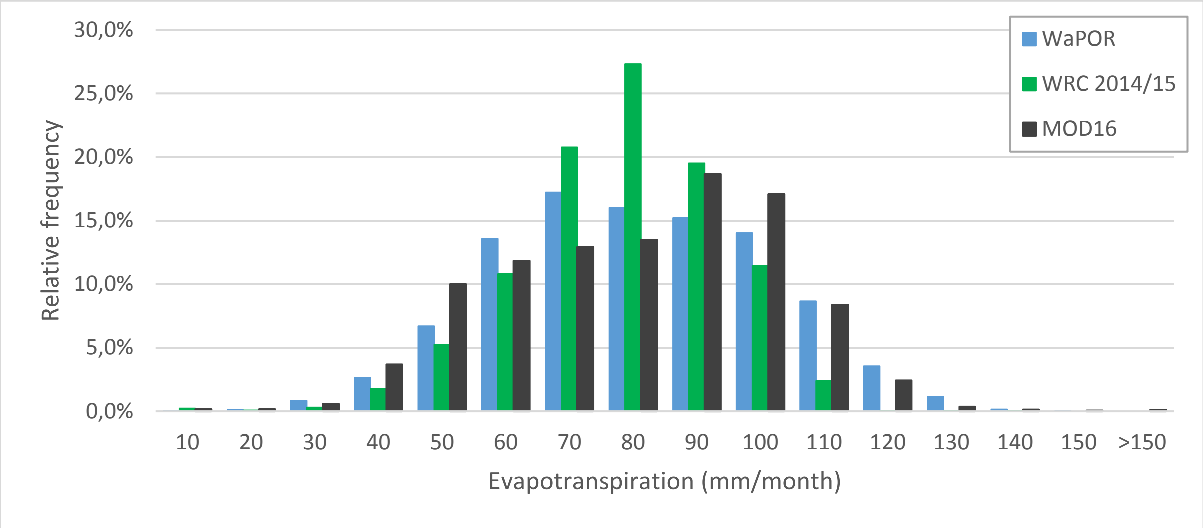


Figure 3-18 WaPOR, WRC 2014/15 and MOD16 ET histogram comparison for April 2015

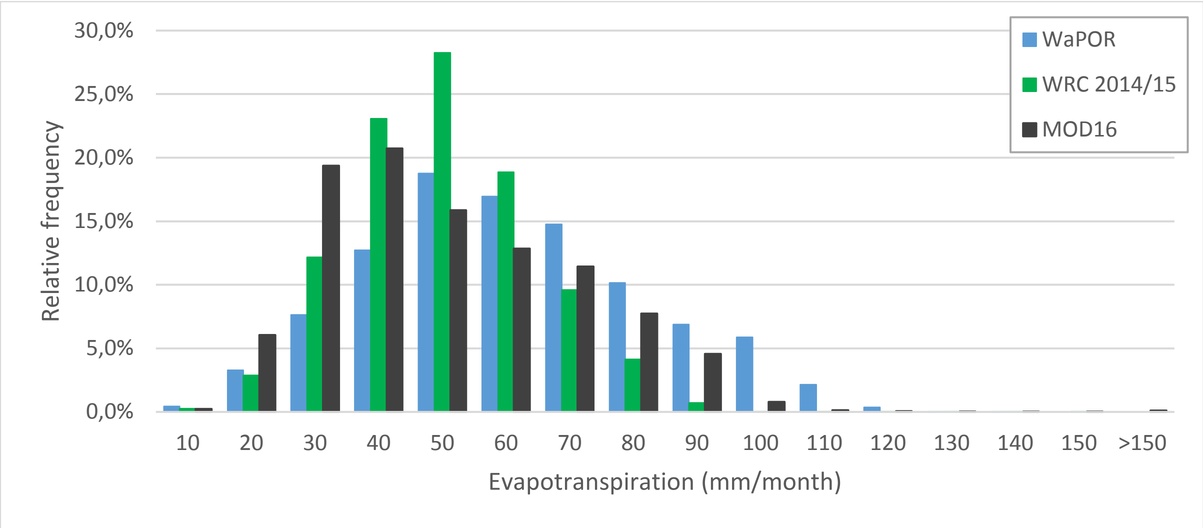


Figure 3-19 WaPOR, WRC 2014/15 and MOD16 ET histogram comparison for July 2015

The differences in ET between the three datasets are attributed to the different input datasets (weather, RS, vegetation, landcover) used in their modelling. However, the lower (500 m) resolution of the MOD16 data is likely a contributing factor as the mixed pixel effect (Section 3.3.3) will be much higher compared to those of WaPOR and WRC 2014/15, which both have a resolution of 250 m.

This comparison highlights the complexity of these models and the varying results generated for the same area and land use. Considering only central values (e.g. median) (Figure 3-15) hides the differences in the frequency distribution of ET values. The comparisons suggest that the WaPOR dataset overestimates ET from forests during winter by around 10-30 mm/month. However, given that the majority of the 26 102 rasterised compartments are located in the summer rainfall area (KwaZulu-Natal and Mpumalanga), the impact of this overestimation is expected to be limited.

3.4 Water use validation

In this section, median monthly and annual evapotranspiration (ET) estimates are compared for *Acacia*, *Eucalyptus* and *Pinus* (Pine). These ET estimates were extracted from the CFDB using the rasterised compartments (see the previous section) for the first ten years of rotation growth and period 1 January 2009 to 31 December 2019. These results are compared to published ET estimates from previous water use studies carried out in South Africa.

3.4.1 Monthly and annual ET statistics

Table 3-2 summarises the 10-year mean, median and standard deviation for monthly and annual ET estimates. The mean annual ET for *Eucalyptus* over the 10-year period is higher than that of *Acacia* and *Pinus*. *Eucalyptus* recorded a mean annual ET of 1 131±256 mm/year, followed by *Pinus* at 1 035±260 mm/year (8% lower than *Eucalyptus*) and *Acacia* at 1 091±196 mm/year (2% lower than *Eucalyptus*). Median and mean ET estimates for the respective genera are similar (Table 3-2). Monthly median and mean ET estimates for the three genera did not vary much over the 10-year period. The winter monthly (July) median ET estimates for *Pinus*, *Acacia* and *Eucalyptus* were 52, 57 and 63 mm/month respectively. Regarding summer monthly (January) median ET estimates, *Pinus*, *Eucalyptus* and *Acacia* respectively recorded 115, 120 and 122 mm/month.

Table 3-2 Summary statistics of evapotranspiration (ET) for selected *Acacia*, *Eucalyptus* and *Pinus* compartments extracted in this study, according to month and year, based on the period 1 January 2009 to 31 December 2020

	ACACIA			*EUCALYPTUS*			*PINUS*		
MONTH	**MEDIAN**	**MEAN**	**SD**	**MEDIAN**	**MEAN**	**SD**	**MEDIAN**	**MEAN**	**SD**
JAN	122	123	25	120	119	28	115	116	30
FEB	109	111	23	109	110	26	104	104	26
MAR	105	107	21	106	108	27	99	100	26
APR	81	82	17	82	85	24	77	78	22
MAY	74	74	17	76	78	24	68	69	21
JUN	58	59	16	61	65	24	53	54	20
JUL	57	59	18	63	66	26	52	54	22
AUG	74	74	20	81	82	27	68	69	25
SEP	87	88	23	93	95	32	82	84	30
OCT	93	93	25	97	99	31	91	92	30
NOV	103	103	25	104	105	31	101	101	30
DEC	115	115	24	115	115	29	111	111	31
ANNUAL	1096	1091	196	1123	1131	256	1038	1035	260

Note: SD = Standard deviation; Units for all values are mm/month

The annual ET for all three genera (Figure 3-20) shows a normal (frequency) distribution with values ranging from about 400 to more than 1 600 mm/year. It is clear from Figure 3-20 that *Acacia* shows a higher frequency of ET estimates close to the mean in comparison to the other two genera. While *Pinus* has a higher frequency of lower ET values (less than 900 mm/year), *Eucalyptus* has a higher frequency of ET values exceeding 1 400 mm/year. It should be noted that the different number of compartments (samples) used to generate these statistics (n=416 for *Acacia*, n=9178 for *Pinus* and n=16 444 for *Eucalyptus*) (Figure 3-20) would have an impact on the ET statistics. The standard deviations of annual ET (expressed as a % of the mean annual ET) for all the genera is high. The mean annual standard deviation for *Eucalyptus*, *Pinus* and *Acacia* is 25% (256 mm/year), 23% (260 mm/year) and 18% (196 mm/year), respectively (Table 3-2). The mean annual standard deviation for *Eucalyptus*, *Pinus* and *Acacia* is 25% (256 mm/year), 23% (260 mm/year) and 18% (196 mm/year), respectively (Table 3-2). These large standard deviation values can be attributed to varying conditions, for example climate, age of plantations, soil depth and water availability.

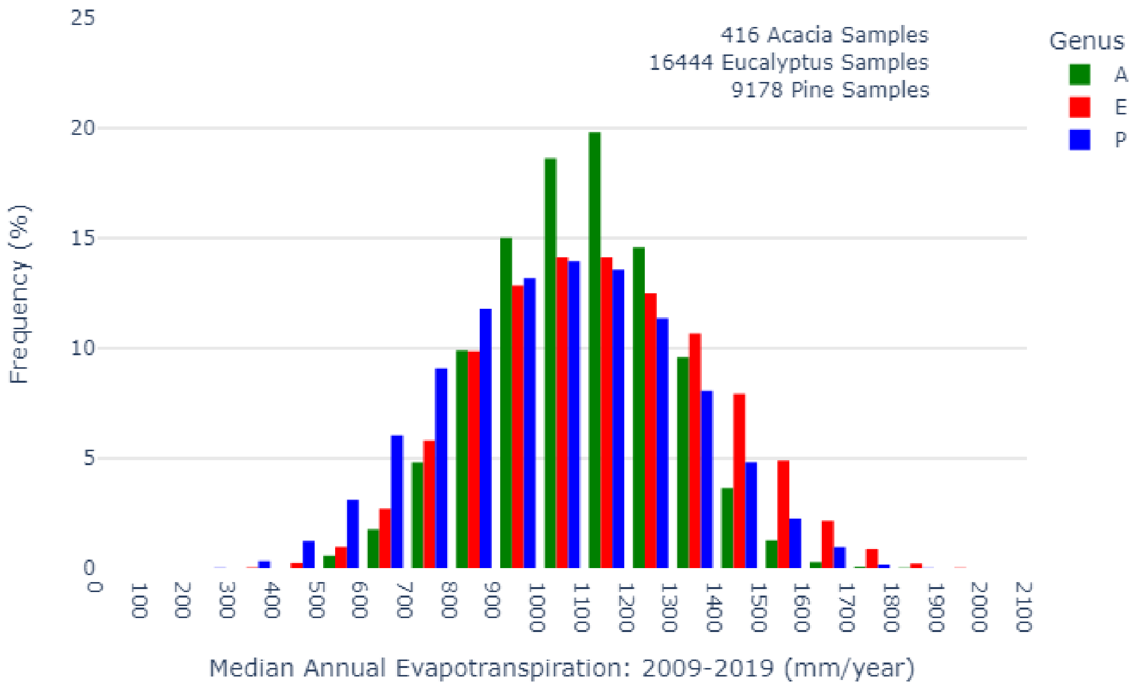


Figure 3-20 Frequency distribution of the median annual ET (mm) for selected *Acacia* (A), *Eucalyptus* (E) and *Pinus* (P) compartments, for the period 1 January 2009 to 31 December 2019

Frequency distributions of monthly median ET estimates are depicted in Figure 3-21 and Figure 3-22. Monthly summer median ET estimates occasionally exceed 200 mm/month (equivalent to 6.5 mm/day) and monthly winter median ET estimates exceed 150 mm (equivalent to 4.8 mm/day). ET estimates for the three genera seem to differ most in the winter months (June, July, August) when soil moisture availability is likely to differ and tree growth is slower (Table 3-2, Figure 3-21 and Figure 3-22).

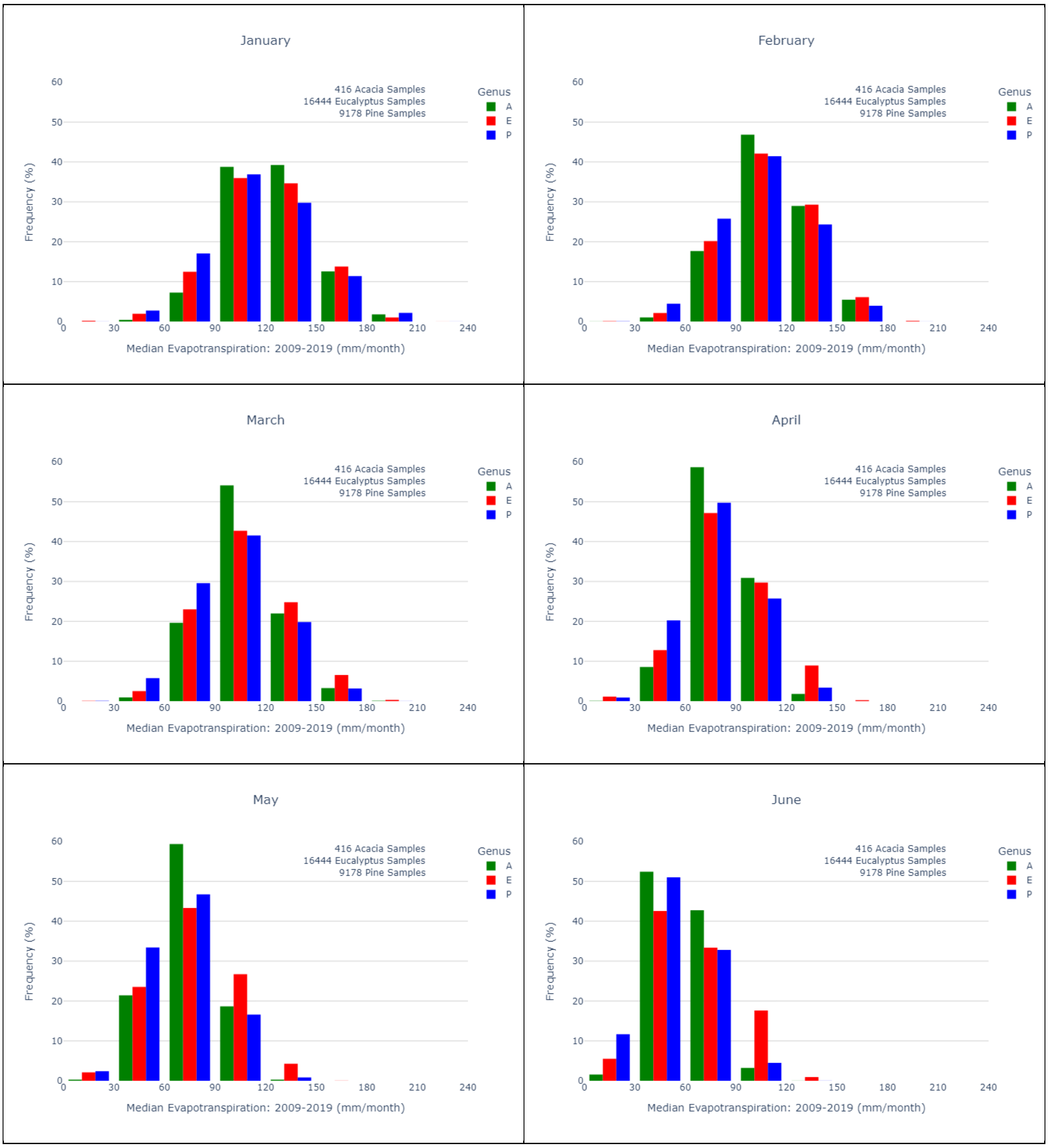


Figure 3-21 Frequency distribution of the median monthly ET (mm/month), months January to June, for selected *Acacia*, *Eucalyptus* and *Pinus* compartments, summarised for the period 1 January 2009 to 31 December 2019

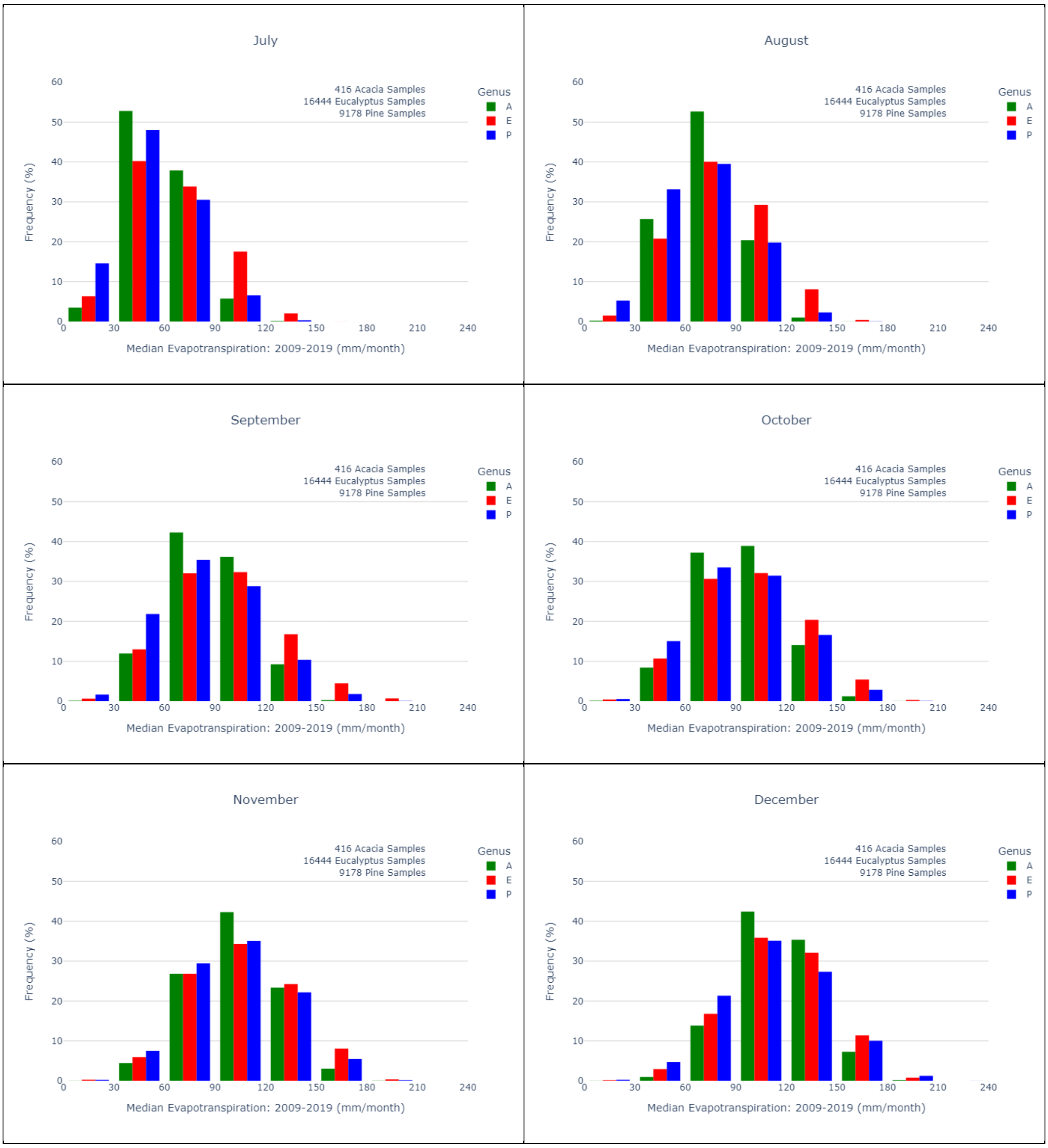
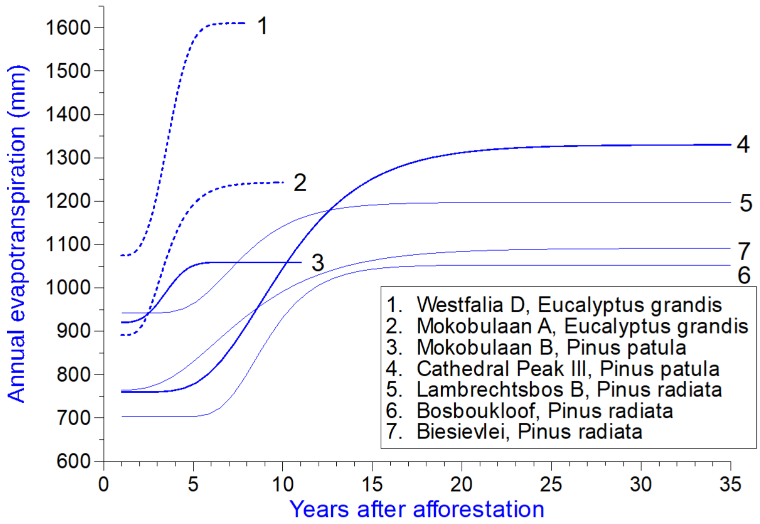


Figure 3-22 Frequency distribution of the median monthly ET (mm/month), months July to December, for selected *Acacia*, *Eucalyptus* and *Pinus* compartments, summarised for the period 1 January 2009 to 31 December 2019

3.4.2 Comparison of ET statistics with data from previous studies

Table 3-2 highlights the seasonality in the ET estimates of the different genera, from low median winter ET values to high summer median ET estimates for the period considered.

Figure 3-23 is based on previous research conducted at the catchment scale, considering the water balance where streamflow and precipitation were measured (Dye 1996). Dye (1996) noted differences in ET from *E. grandis* and *Pinus* spp. following afforestation (Figure 3-23), with ET of *E. grandis* showing a dramatic increase in ET during initial growth and canopy closure, compared to a relatively slow early growth for *Pinus* spp. Dye (1996) further highlights that the ET after canopy closure (flattening lines) varies substantially between the plantations illustrated in Figure 3-23 and attributes these variations to differences in mean annual precipitation. Nevertheless, the annual ET values shown in Figure 3-23 compare well with those in Table 3-2, where the median of *Eucalyptus* compartments was quantified as being 1 123 mm/year, with a standard deviation of 256 mm/year, and *Pinus* compartments having a median annual ET of 1 038 mm/year.



Taken from (Dye, 1996) and based on original research from Bosch and Gadow, 1990; Smith and Scott, 1992 and Versfeld, 1993)

Figure 3-23 Trends in post afforestation ET recorded from seven paired catchments in South Africa, representing different forestry regions

Clulow et al. (2011) captured the seasonality in ET for *A. mearnsii*. ET ranges between minimum monthly values in June of ~50 mm/month and maximum ET estimates in December/January of ~150 mm/month were noted (Figure 3-24). These ET estimates compare well with statistics extracted for *Acacia* (Figure 3-20*)*, where median ET values range between 57 and 122 mm/month and mean monthly ET estimates between 59 to 123 mm/month.

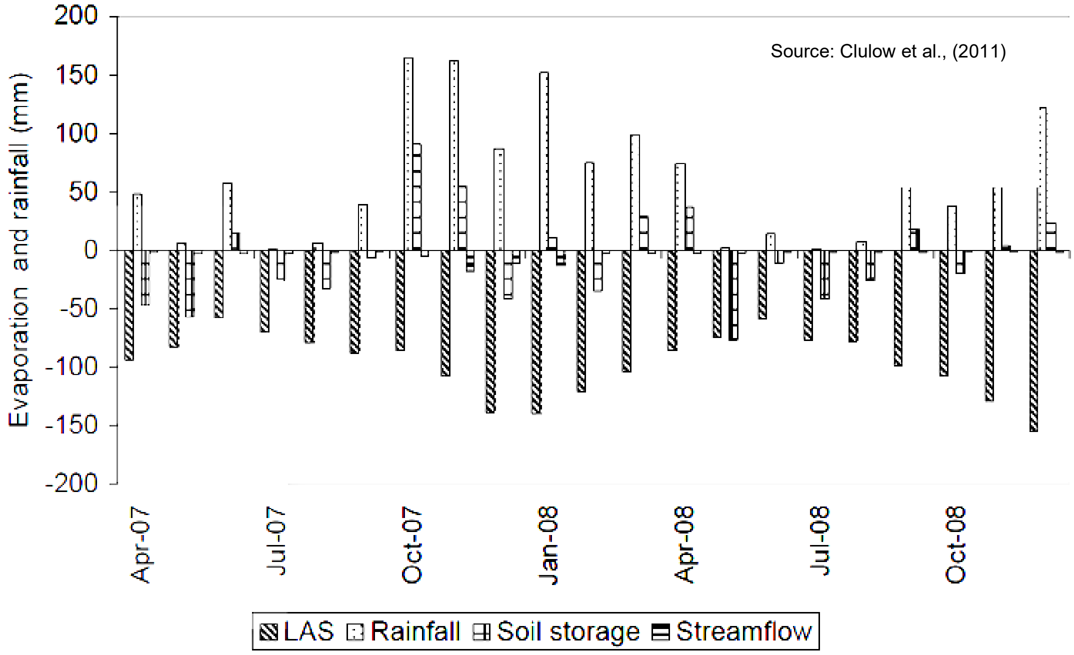


Figure 3-24 Monthly ET (shown here as negative evaporation) in mm/month from an *A. mearnsii* stand

Table 3-3 summarises ET and transpiration (T) estimates from nine publications containing annual ET (and T) estimates; in some cases, for the same forestry stand. The ET presented was estimated (mainly measured) using surface energy balance methods (e.g. scintillometry, Bowen Ratio), RS models (e.g. SEBAL) and the catchment water balance (WB). Data was available for the main forestry provinces and estimated for forestry compartments and alien invasive plant (AIP) stands of the same genera.

The annual ET estimates summarised in Table 3-3 show a large range: from 575 mm/year (precipitation = 865 mm/year) for *Eucalyptus* (Meijninger and Jarmain 2014) to 1 618 mm/year (precipitation 860 mm/year) also for *Eucalyptus* (Jarmain and Everson 2002) (Table 3-4). The latter was for a four- to five-year-old fast-growing forestry stand.

Table 3-3 Summary of Evapotranspiration, ET (and Transpiration, T) data from three main forestry Genera: *Acacia*, *Eucalyptus* and *Pinus*

Genus	*Species*	*ET/T*	*ET / T Mean**	*ET / T SD**	*Province*	*P Mean**	*P SD**	*Reference*	*Plan /AIP*	*Meas/ Mod*	*Other*
A	*A. mearnsii*	ET			KZN			Clulow et al. 2011	Plan	Meas	ET (Dec/Jan) = 150 mm/month; ET (Jun) = ~50 mm/month
A	*A. mearnsii*	ET	1240		KZN	874		Dye and Jarmain (2004) Jarmain and Everson (2002)	Plan	Meas	
A	*A. mearnsii*	ET	1364		KZN	616		Dye and Jarmain (2004) Jarmain and Everson (2002)	Plan	Meas	
A	*A. mearnsii*	ET	1239		KZN	1016		Dye and Jarmain (2004) Jarmain and Everson (2002)	Plan	Meas	
A	*A. mearnsii*	ET	1048		KZN	860		Dye and Jarmain (2004) Jarmain and Everson (2002)	Plan	Meas	
A	*A. mearnsii*	ET	951		KZN	1071		Everson et al. (2007)	Plan	Meas	
A	*A. mearnsii*	ET	793		KZN	897		Everson et al. (2007)	Plan	Meas	
A	*A. mearnsii*	ET	1121		KZN	1170		Everson et al. (2007)	Plan	Meas	
A	*A. mearnsii*	ET	705		KZN	659		Everson et al. (2007)	Plan	Meas	
A	*A. mearnsii*	ET	701		KZN	727		Everson et al. (2007)	Plan	Meas	
A	*A. mearnsii*	ET	1038		KZN	1139		Everson et al. (2007)	Plan	Meas	
A	*A. mearnsii*	ET	1242		KZN	869		Everson et al. (2014)	Plan	Meas	ET (Jun): 52 mm/month; ET (Dec):136 mm/month; ETmax: 8.5 mm/day
A	*A. mearnsii*	ET	1171		KZN	914		Everson et al. (2014)	Plan	Meas	
A	*A. mearnsii*	ET	1173		KZN	765		Everson et al. (2014)	Plan	Meas	
A	*A. mearnsii*	ET	1132		KZN	587		Everson et al. (2014)	Plan	Meas	
A	*A. mearnsii*	ET	1143		KZN	856		Everson et al. (2014)	Plan	Meas	
A	*A. mearnsii*	ET	1088		KZN	846		Everson et al. (2014)	Plan	Meas	
A	*A. mearnsii*	ET	1157		KZN	862		Everson et al. (2014)	Plan	Meas	
A	*A. mearnsii*	T						Dye and Jarmain (2004)	AIP	T meas; ET mod	Max:T 7 mm/day
A	*A. mearnsii*	ET	1503		WC			Dye and Jarmain, (2004)	AIP	ET mod	
A	*A. mearnsii*	ET	1260		KZN			Dye and Jarmain (2004)	AIP	ET mod	
A	*A. mearnsii*	ET			WC			Dye and Jarmain, (2004)	AIP	T meas	Tmax: 5-6 mm/day; Tmin:3 mm/day
A	*A. mearnsii*	ET	925	225	WC	650	140	Meijninger and Jarmain (2014)	AIP		RS mod
A	*A. mearnsii*	ET	740	145	KZN	870	135	Meijninger and Jarmain (2014)	AIP		RS mod
A	*A. mearnsii*	ET	615	140	KZN	900	125	Meijninger and Jarmain (2014)	Plan		RS mod
E	*E. dunnii*	T	673		MP	704		Dye (2013)	Plan	Meas	
E	*E. grandis*	T	1231		MP	1459		Dye (2013)	Plan	Meas	

Genus	Species	ET/T	ET / T Mean*	ET / T SD*	Province	P Mean*	P SD*	Reference	Plan /AIP	Meas/ Mod	Other
E	*E. grandis*	ET	1063		MP	1611		Dye (1996)	Plan		WB meas
E	*E. grandis*	ET	891		MP	1135		Dye (1996)	Plan		WB meas
E	*E. dunii, E. macarthurii*	ET	1246		KZN	874		Jarmain and Everson (2002)	Plan	Meas	
E	*E. dunii, E. macarthurii*	ET	1588			616		Jarmain and Everson (2002)	Plan	Meas	
E	*E. dunii, E. macarthurii*	ET	1428		KZN	1016		Jarmain and Everson (2002)	Plan	Meas	
E	*E. dunii, E. macarthurii*	ET	1618			860		Jarmain and Everson (2002)	Plan	Meas	
E	*E. grandis*	T	1347		MP			Dye et al. (2008)	Plan	Meas	Tmax (summer): 7 mm/day
E	*E. grandis*	T			MP			Dye et al. (2008)	Plan	Meas	T (winter): ~2.5 mm/day
E	*Eucalyptus spp.*	ET	945	230	WC	860	250	Meijninger and Jarmain (2014)	AIP		RS mod
E	*Eucalyptus spp.*	ET	575	195	KZN	865	65	Meijninger and Jarmain (2014)	AIP		RS mod
E	*Eucalyptus spp.*	ET	690	190	KZN	935	130	Meijninger and Jarmain (2014)	Plan		RS mod
P	*P.patula*	ET	918		MP	1135		Dye (1996)	Plan		WB meas
P	*P.patula*	ET	942		WC	1473		Dye (1996)	Plan		WB meas
P	*P.patula*	ET	818		KZN	1578		Dye (1996)	Plan		WB meas
P	*P.patula*	ET	703		WC	1296		Dye (1996)	Plan		WB meas
P	*P.patula*	ET	764		WC	1427		Dye (1996)	Plan		WB meas
P	*Pinus patula*		944		Swaziland	1357		Dye et al. (2008)	Plan	Mod	
P	*Pinus spp.*	ET	735	215	WC	790	205	Meijninger and Jarmain (2014)	Plan		RS mod
P	*Pinus spp.*	ET	650	155	KZN	935	130	Meijninger and Jarmain (2014)	Plan		RS mod

NOTES: ET and T are shown in mm/year, with occasional daily or monthly datasets. KZN refers to KwaZulu-Natal, WC to Western Cape, MP to Mpumalanga and NP to Northern Province. Data is either measured (Meas) or modelled (Mod), where RS refers to remote sensing techniques and WC to the catchment water balance. Data are shown for stands of Alien Invasive plants (AIP) or plantations (Plan). A = *Acacia*; E = Eucalyptus; P = *Pinus*; * mm/year

Considering *Acacia* only, a genus studied in detail over the past 20 years, the annual ET ranged between 615 mm/year (Meijninger and Jarmain 2014) (KZN) and 1 503 mm/year for an AIP stand in the WC (Dye and Jarmain 2004). This wide range in annual ET is due to differences in climatic conditions, the density of tree stands, and soil water availability (Table 3-3, Figure 3-25). Everson et al. (2007) and Everson et al. (2014) monitored the ET of an *Acacia* compartment for over a decade through different growth and management stages. They recorded annual ET estimates ranging between 701 and 1 242 mm/year (Table 3-3). ET estimates statistics for *Acacia* in this study (median ET = 1 096 mm/year, Table 3-2) and maximum annual ET estimates to 1 600 mm/year, Table 3-2) are in line with those from literature (Figure 3-25) and compare well to the ET statistics extracted for *Acacia* in this study (median ET = 1 096 mm/year, Table 3-2).

For *Eucalyptus,* data from previous studies show annual ET estimates between 575 and 1 618 mm/year, with a mean annual ET of 1 116 mm/year (Table 3-3, Figure 3-25). Comparably, the transpiration estimates from literature for *Eucalyptus* ranged between 673 and 1 347 mm/year (Table 3-3). The median annual ET and ET range (1 123 mm; 500-1 800 mm) estimated for *Eucalyptus* as part of this study (using RS data from a 10-year period and considering multiple forestry compartments across South Africa) fall within this annual ET range from literature.

For *Pinus,* data from past studies in SA considered here (Table 3-3) show a range in annual ET of 650 to 944 mm/year, which is lower than that recorded for *Acacia* and *Eucalyptus* in other studies (Table 3-3, Figure 3-25). These results from past studies are in agreement with the lower median annual ET estimate for *Pinus* found in this study (1 038 mm/year) (Table 3-2) and the higher frequency of lower annual ET values (Figure 3-25) of less than 900 mm/year, compared to *Acacia* and *Eucalyptus*. However, the results from the current study show higher ET estimates (Figure 3-20) for *Pinus* than reported in the literature (Table 3-3).

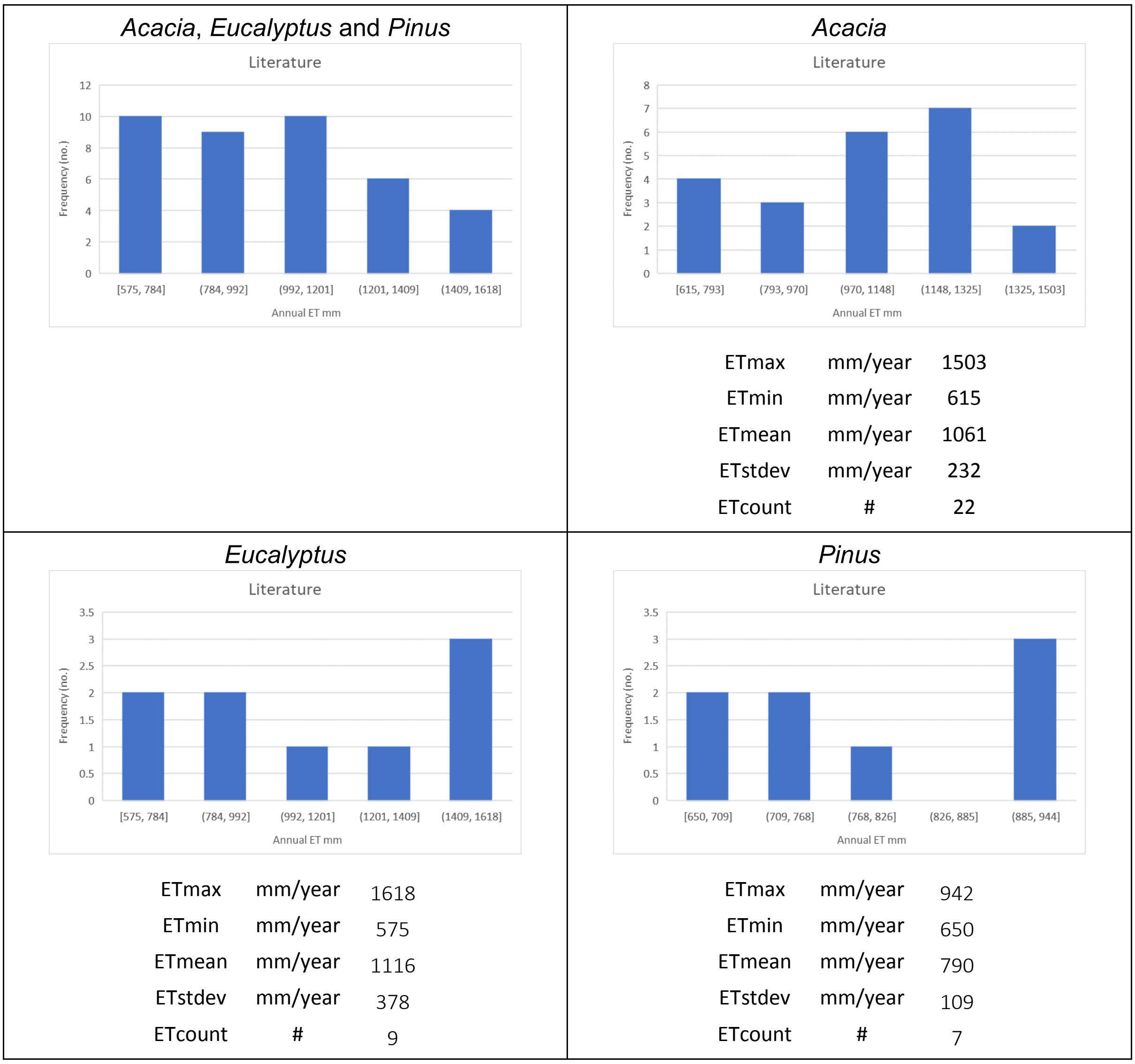


Figure 3-25 Frequency distribution of annual ET recorded from Acacia, Eucalyptus and Pinus in past studies and as summarised from a literature review (Table 3-3). Summary statistics are shown. ETmax, ETmn, ETmean, ETstdev and ETcount refer to maximum, minimum, mean standard deviation and number of ET samples respectively.

Rainfall data were not readily available for comparison to ET estimates. However, several South African studies compared annual ET or tree water use estimates together with annual rainfall and streamflow estimates. Dye and Bosch (2000) noted that the upper limit of ET is often set by rainfall, except where there is a significant storage of soil water from previous years (based on work by Scott and Lesch (1997)). Hence, where trees can access a soil water store, annual estimates of ET can exceed annual estimates of rainfall. The long-term catchment experiment at Two Streams, South Africa (Everson et al. 2014) shows the sustained impact of afforestation on streamflow in a catchment. Everson et al. (2014) recorded tree water use and ET estimates exceeding rainfall for sustained periods, ranging from an ET excess of 32% following planting of *A. mearnsii* trees to an ET excess of 22% from a maturing *A. mearnsii* stand; prior to planting the ET excess was as high as 46%. They monitored soil water content in the upper 2.4 m soil profile and found a steady decline in soil water content, though it cannot account for the full ET excess. Everson et al. (2014) found *A. mearnsii* roots to a depth of >8 m, suggesting the trees had access to soil water stores beyond the depth monitored.

It is important to note that the current study used average ET per compartment. The spatial variation within a compartment was not considered. Everson et al. (2014) confirmed data from a review on catchment streamflow by Scott and Lesch (1995). Both these publications show the impact of clearance of forest within and outside of the riparian zones and that streamflow was enhanced proportionately more following the removal of riparian trees than following removal of non-riparian trees. This is attributed to the greater rate of water use by riparian trees compared to higher lying trees. Everson et al. (2014) further showed that annual tree water use from south-facing slopes were lower than that from north-facing slopes due to differences in available solar and net radiation.

3.5 Environmental factor mapping and extraction

The per-compartment water use estimations were compared to a range of environmental (terrain and climatic) variables to determine whether these conditions impact water use. The variables and the data sources considered are listed Table 3-3. In many cases the original data were continuous, which required reclassification into meaningful categories. In some instances, the nominal classes were simplified.

Table 3-4 Environmental variables considered for explaining water use variations

Type	Variable	Source	Classification
Terrain	Slope gradient	Van Niekerk (2015)	Level/gently inclined (0-10%) Moderately inclined/steep (10-56%) Very steep (>56%)
	Slope aspect	Van Niekerk (2015)	North East South West
	Terrain morphology	Schulze (2007)	Plains/flat Hilly/undulated Mountainous
Climate	Rainfall seasonality	Schulze (2007)	All year Winter Early summer Mid summer Late summer Very late summer
	Climate zones (Köppen)	Schulze (2007)	Arid, hot and dry (BWh) Arid, cool and dry (BWk) Semi-arid, hot and dry (BSh) Semi-arid, cool and dry (BSk) Summers long dry and cool (Csb) Summers long, dry and hot (Csa) Wet all seasons, summers long and hot (Cfa) Wet all seasons, summers long and cool (Cfb) Winter long, dry and hot (Cwa) Winter long, dry and cool (Cwb) Tropical wet, dry and winter season (Aw)
	Annual rainfall	Schulze (2007)	Low (< 600 mm) Medium (600-800 mm) High (800-1 000 mm) Very high (> 1 000 mm)

Slope gradient and aspect were sourced from the SUDEM and reclassified as indicated in Table 3-3. The climate variables (annual rainfall, rainfall seasonality, climate zones), as well as the morphological units, were sourced from the *South African atlas of agrohydrology and -climatology* (Schulze, 2007). The morphological units were reclassified (simplified) as indicated in Table 3-5

Table 3-5 Reclassified terrain morphology units

Original units	Simplified units
Moderately undulating plains Plains	Plains/flat
Undulating hills and lowlands Undulating hills Irregular undulating lowlands and hills Strongly undulating irregular land	Hilly/undulated
Low mountains Highly dissected low undulating mountains Undulating mountains and lowlands Mountains and lowlands High mountains	Mountainous

The environmental conditions of each compartment were determined by extracting the dominant (mode) value from each of the variables listed in Table 3-3. These attributes were stored in the CFDB and used in the water use estimation analyses (Section 5).

4 PLANTATION FOREST MAPPING USING REMOTE SENSING

This project made use of existing databases of forest compartments supplied by commercial companies (Section 3.1.1). The data were ideal for the purposes of this project, as it represent a large proportion of the commercial forests in South Africa – which was the focus of this study. However, the data do not represent smaller, independent growers and communal forests. Smaller companies are unlikely to keep plantation records (forest inventories), making it difficult to obtain in situ data of their forests. Currently, forests are monitored every three years according to national indicators and criteria such as the development and maintenance of forest resources, biological diversity in forests, the health and vitality of forests, the productive functions of forests, the protective and environmental functions of forests, and the social functions of forests (Government of South Africa 1998). Forestry data are collected through questionnaires that assess the use of forest resources on communal land and through licences. These methods do not provide a comprehensive and up-to-date overview of forestry activities in South Africa. Mapping plantation forests over large areas and extracting pertinent information per compartment, such as planting date and genus, would support national reporting and regulatory activities. It will also allow for the water use quantification methods described in the previous section to be applied to all forests in South Africa and even in other countries. A number of student (capacity building) projects were consequently initiated as part of this larger WRC project to investigate how RS technologies can be used to cost-effectively extract forest compartment information at regional scales. These projects focussed on mapping plantation forests (e.g. identifying the location and extent of compartments) and extracting compartment characteristics (e.g. age and genus). An overview of these activities is provided in the following subsections.

4.1 Plantation forest mapping

Although the operationalisation of forest plantation mapping is not one of the aims of this project, the lack of publicly available forest data to support land- and water use decisions highlights the need for developing EO techniques that can potentially be used to map forests, amongst other land covers, at regional and national scales. Also, while the data obtained from the forestry industry represent the current status, EO techniques can potentially be used to generate historical forest plantation maps that will be useful for identifying trends in forestry-related land use. If successful, such techniques could also be used to generate forest plantation maps at regular intervals (e.g. annually) going forward, which will allow for updated water use estimations and water accounting activities. A third motivation for developing EO forest plantation mapping techniques is that there is a critical scarcity in EO human capacity in South Africa. This situation is worsened by the recent emigration of several key EO researchers. Research on the use of advanced EO techniques (e.g. ML) for forest monitoring will therefore contribute to building much needed capacity in RS and hopefully inspire the young researchers working on the project to become the next generation of EO specialists in South Africa.

Although the mapping of plantations was the primary goal, the EO methods tested often targeted more than just the 'commercial forestry' land cover class. Two approaches to land cover mapping were investigated: ML and DL, both in a GEOBIA environment. See Section 2.3.4 for more details on these concepts.

4.1.1 Machine learning

The use of ML for classifying satellite imagery has received much attention lately. The main reason for this is that traditional satellite image classification techniques (e.g. statistical and knowledge-based classification) are unsuitable for classifying the enormous volumes of EO data that are currently being generated. This section briefly overviews some experiments undertaken to evaluate ML techniques for mapping forest plantations.

Three 100x100 km areas (Sentinel-2 tiles), roughly corresponding to the Knysna, Pilgrims Rest and Richards Bay areas (Figure 4-1), were chosen for the initial experiments.

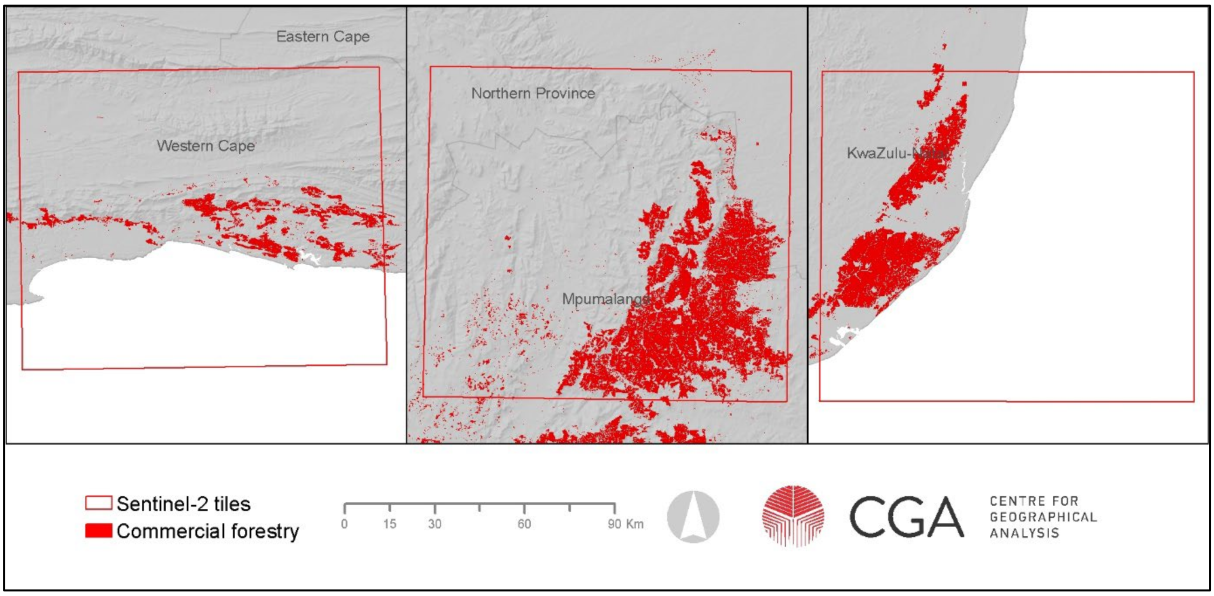


Figure 4-1 Regions in which machine learning experiments were carried out: Knysna, Pilgrims Rest and Richards Bay

Two cloud-free Sentinel-2 images of each area were acquired for April and October 2019, after which radiometric calibrations and atmospheric corrections were applied. The images were then topographically normalised to counter for variable incidence angles (shadow effects) and the effects of the bidirectional reflectance distribution function (BRDF). Several image transformations were applied to the normalised imagery, including NDVI, EVI, principal component analysis (PCA), a first principal component (PC1) on which was calculated Haralick GLDV entropy, and standard deviation texture measures. The 20 m Sentinel-2 bands were resampled to 10 m, and all the features were combined into a single analysis ready image stack. Additionally, the first principal component of the RGB (red, green, blue) bands of the National Geospatial Information (NGI) aerial photographs was used to calculate the same two texture measures as for the Sentinel-2 images. These three layers were also added to the stack along with two cross-polarised (VH) and the co-polarised (VV) Sentinel-1 SAR images.

Using the CGA's Rapid Object Collection and Analysis Tool (ROCAT) (Figure 4-2), GIS operators classified more than 14 000 randomly-selected land cover sample objects falling within the three study areas. Initially, the classification scheme consisted of 18 classes, which were later reduced to 8. This was done due to the overlap in many of the classes that shared similar physical and spectral properties such as *Grass* and *Grass crops* as well as *Woodland*, *Bushland* and *Orchards*.

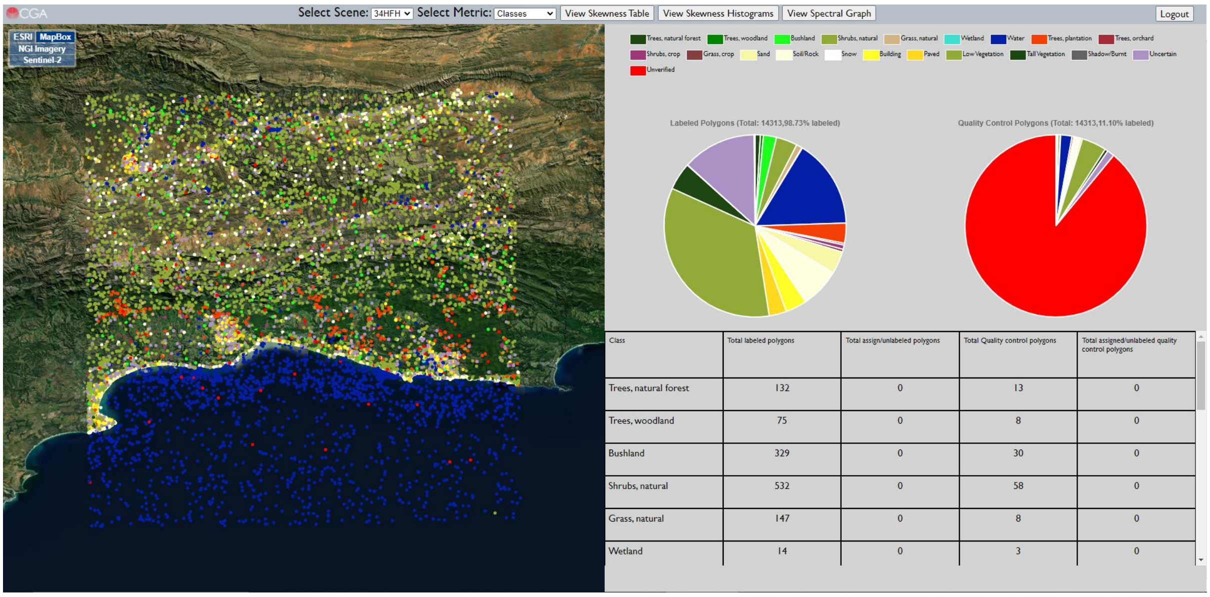


Figure 4-2 The CGA rapid object collection and analysis tool (ROCAT)

A series of 24 land cover mapping experiments were then performed using the samples to train RF and SVM classifiers. A variety of model adjustments were applied to the ML classifiers, including pre-modelling classification rules (e.g. using an NDVI threshold to reclassify *Plantation* training objects to *Soil/Rock* where harvest has occurred between analysis dates), different input features (e.g. image bands, vegetation indices, image texture, multitemporal Sentinel-2 data and SAR data), classification scheme variations (e.g. simplification of *Shrubs* and *Grass* to *Low vegetation*), different weighting of input training and validation data (e.g. equally balanced between classes vs distributed according to land cover coverage), and the application of post-classification rules (e.g. NDVI thresholds for vegetation class fine-tuning).

The results of the experiments were mixed, with the overall accuracy (calculated with a fairly distributed validation dataset) of the best experiment being 76% for an 8-class land cover classification. More pertinent to this study, however, was the persistent confusion between the *Plantation* class with both *Tall vegetation* and *Low vegetation*, with user accuracies for the *Plantation* class often being very low. This result was disappointing from a forest mapping perspective and explains why the South African National land cover (SANLC) products developed by Department of Forestry, Fisheries and the Environmental (DFFE) still make use of manually digitised masks for *Plantation* mapping.

When the 8-classes were aggregated into *Plantation* and *Non-plantation* classes, the overall accuracy increased to 97%. While this seems high, it is largely due to the more accurate mapping of the *Non-plantation* class, which was heavily weighted (>20 times more samples due to larger extent) compared to the *Plantation* class. A better indicator of accuracy is the kappa statistic (0.37) and the user's and producer's accuracies of the *Plantation* class, which were 65% and 28% respectively. From Figure 4-3, it is evident that the model overestimates plantation forests in some areas (shown in orange), particularly at the edges of compartments and indigenous forests, while in other areas plantations are underestimated (shown in blue).

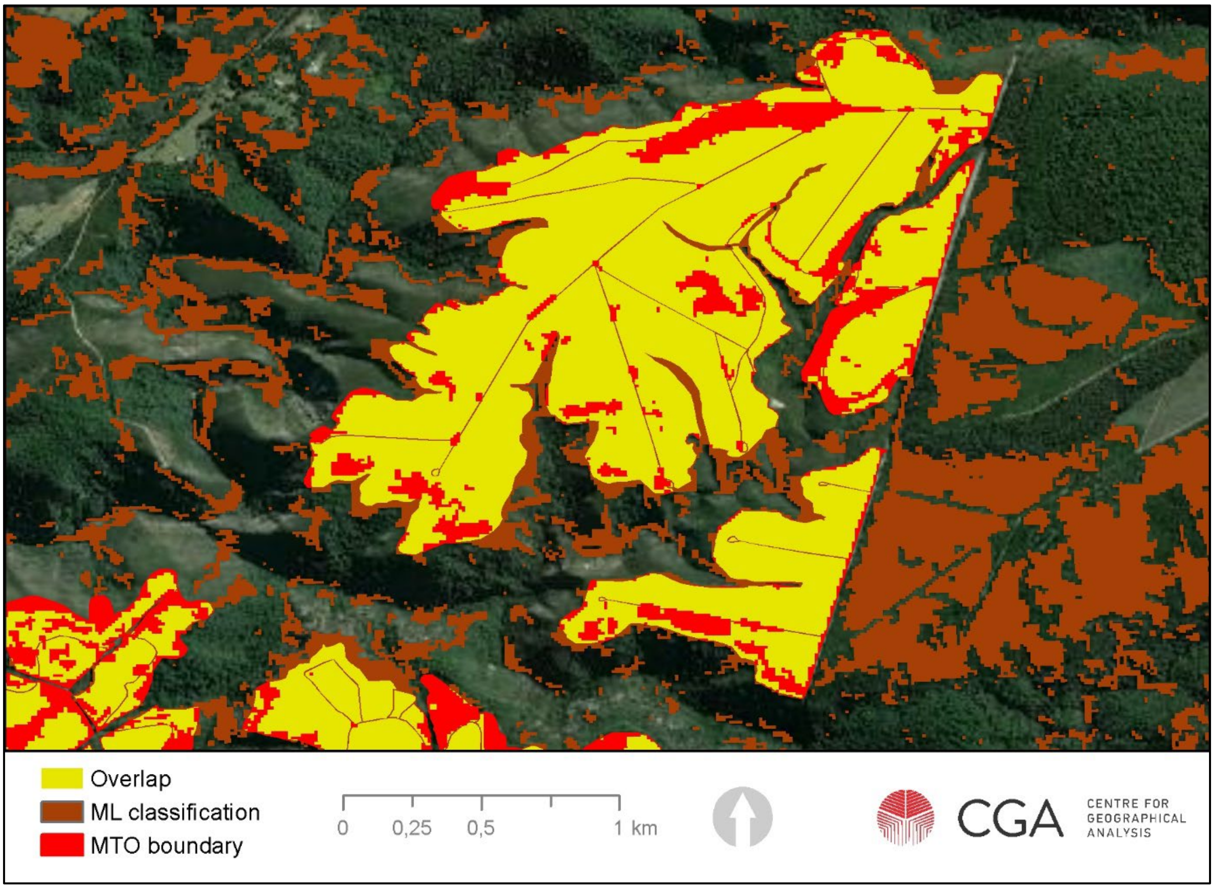


Figure 4-3 Machine learning classification result of Plantations compared to the MTO boundaries

From these results it was concluded that the spectral properties of forested land covers are too similar to accurately differentiate when using Sentinel-2 imagery. Even advanced ML algorithms such as RF and SVM, coupled with large quantities of training samples, were unable to produce acceptable (>80%) classifications. The main challenge is to differentiate between planted and indigenous forests. When the Sentinel-2 imagery was visually compared to VHR imagery (e.g. WorldView), it was clear that the structural differences/variations between planted and natural forests were lost at 10 m resolution. Although the incorporation of VHR aerial imagery and texture measures assisted the classifiers, this data were applied at object level (groupings of pixels), which diminished its impact. It was concluded that a computer vision approach – using DL – will likely produce better results.

4.1.2 Deep learning

DL and computer vision is routinely used in medicine, engineering and computer science. In principle, computer vision attempts to emulate human vision by developing procedures and techniques that can recognise patterns and colours in images. When RS operators interpret aerial imagery, they are often able to differentiate planted forests from other land covers as they often have very distinct spatial (textural, shape and size), spectral (lightness and hue), and contextual characteristics. CNNs (Section 2.3.4.2), a type of DL algorithms, attempt to identify and use these unique characteristics to label a particular image (or area within an image). In this project, MSc student Mr V Ndyafi carried out a range of DL experimentations, applied to VHR colour (RGB) aerial photography, to differentiate between indigenous and plantation forests (and other land covers). This subsection briefly overviews the approach used, as well as some preliminary results (the research is ongoing).

One of the main impediments to using DL in RS is the lack of adequate samples to train the models. DL algorithms require thousands, preferably millions, of samples per class in order to build robust models. Such data are invariably not available for remotely sensed data because it is expensive to collect/generate. The development of an automated method for generating a large sample set to train a CNN was consequently the first step in this MSc research. The 2020 South Africa National Land Cover (SANLC 2020) was used to automatically identify training and validation (reference) sample points. First, the SANLC 2020 classes were generalised (reclassified) from 173 to three classes namely, *Planted forest*, *Natural woody vegetation* and *Other*. The resulting classified layer was used in a stratified random sampling scheme to produce 100 samples per class. Each sampled point was used to extract (subset) a 32x32 m patch from 50 cm resolution colour aerial photographs. To artificially increase the number of samples per class, each sample patch was rotated in four directions, resulting in 100 original plus 300 modified sample patches per class. The 400 sample patches per class were used to train the Google TensorFlow CNN algorithm, as implemented in eCognition Developer 10.1.

The resulting CNN model was applied to a 1:50 000 (about 100x100 km area) tile of aerial imagery to produce class probability (membership) maps, also referred to as class heatmaps. The RGB image was subsequently segmented into homogenous objects, which were classified based on membership. Finally, the accuracy of the classification was assessed using a confusion matrix and derived metrics.

Experiments are still being carried out, but preliminary results of this methodology are very encouraging. Figure 4-4 shows an example of where the model was applied to a large forested area in KwaZulu-Natal, with an overall accuracy of 92% (Kappa = 0.87). This is remarkable given that only RGB imagery was used (i.e. imagery with a very limited spectral resolution) and considering that only 100 (unmodified) samples per class were used (compared to the thousands used in the ML – see Section 4.1.1). At closer inspection (Figure 4-5), it is clear that some misclassifications do occur. For instance, a number of clear-felled plantations were incorrectly classified as a planted forest (as seen in the centre of Figure 4-5). This is likely due to temporal differences between the aerial photographs (2016) and the 2020 imagery used to produce the SANLC 2020. With only 100 samples, sample labelling errors (due to temporal differences) can have a significant effect on classification accuracies.

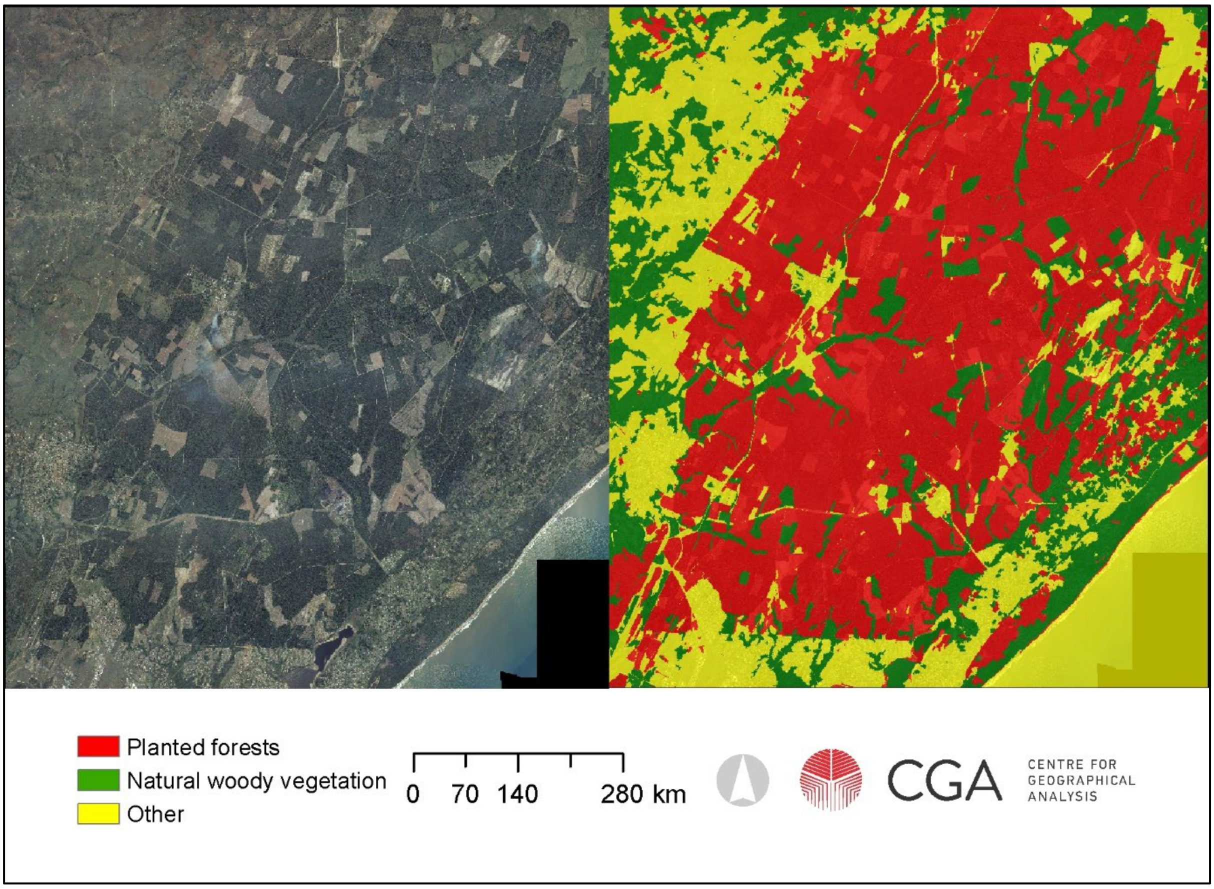


Figure 4-4 Convolutional neural network plantation and indigenous forest mapping results in Richards Bay, KwaZulu-Natal

Water bodies are also often confused with (misclassified as) planted forests (as seen in the south of Figure 4-5), likely because they are relatively dark and homogenous, similar to plantation forests. However, we believe that these issues can be circumvented by producing a larger number of samples, visually inspecting the samples prior to classification, and dropping samples that are clearly incorrect. Adding a separate class for water bodies is also being investigated. Various experiments are being carried out to improve the DL methodology. The intention is to incorporate Sentinel-2 imagery during the object-based classification step to further improve the accuracies of the resulting maps. The models will also be applied to various regions within South Africa to test its transferability.

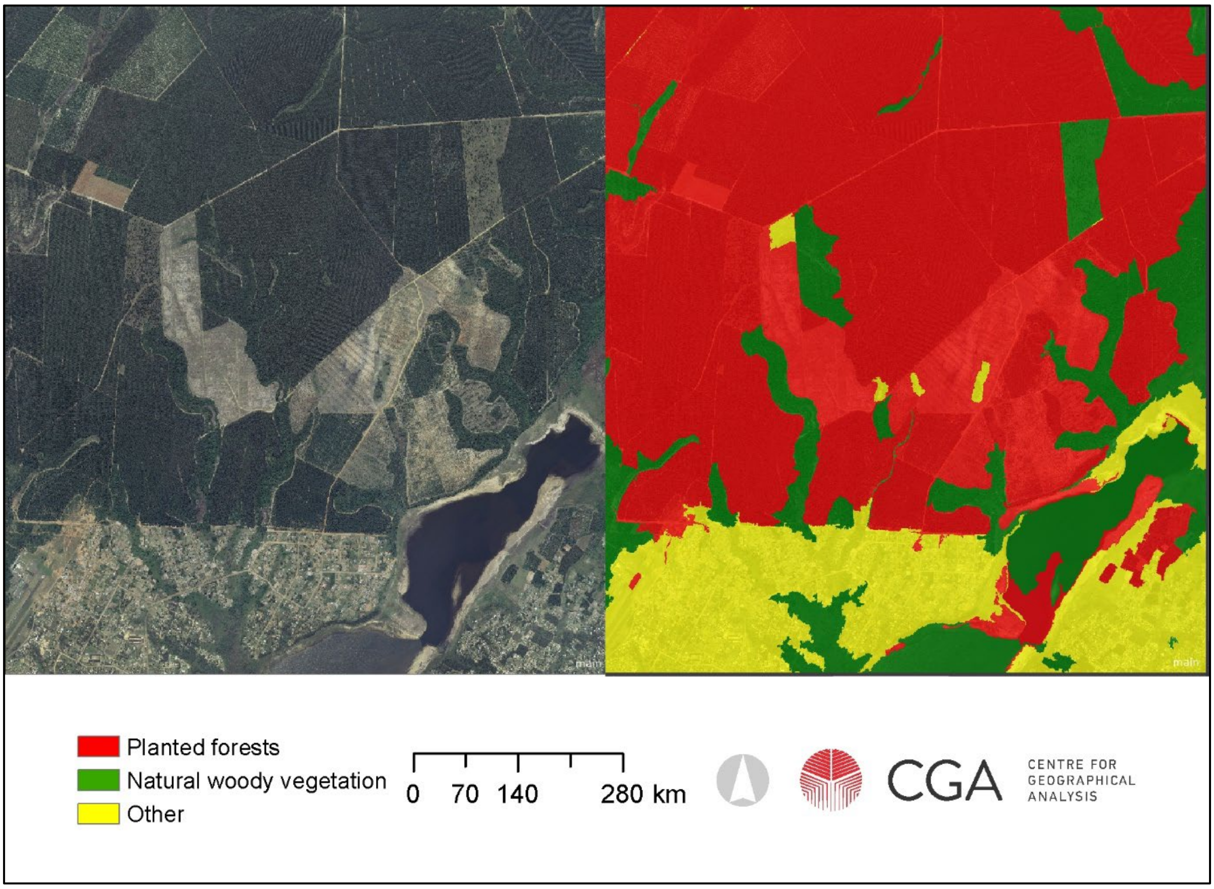


Figure 4-5 Detailed convolutional neural network plantation and indigenous forest mapping results in Richards Bay, KwaZulu-Natal

4.2 Genus mapping

Genus information is used in allometric equations for carbon stock estimates and changes (Maniatis et al. 2011) (to estimate the biomass and carbon stock levels) (Basuki et al. 2009) and in streamflow reduction models (Gush et al. 2002). Genus information is also used for planning, for land management, to maximise production by assessing the mean annual increments of growth rates, to assess water use, to monitor timber harvests and rotations, and to assist in forest management (Mati and Dawaki 2015).

Mapping forest plantation genera using RS and ML algorithms is a viable approach to reduce the time and cost compared to in situ data collection. Medium resolution imagery used in ML algorithms to differentiate between spectrally similar forest plantation species produced low accuracies <63% (Franco-Lopez et al. 2001; Stabach et al. 2009). Higher accuracies have been achieved using VHR satellite imagery (Francois and Leckie 2006; Immitzer et al 2012; Ke et al. Pu and Landry 2012), while unmanned aerial vehicle (UAV) data (Franklin, Ahmed and Williams 2017; Franklin and Ahmed 2018) and/or hyperspectral imagery (Buddenbaum et al.; Bujang and Baharum 2017; Fagan et al. 2015; Peerbhay et al. 2013; Voss and Sugumaran 2008) have also been used for this purpose. However, VHR, UAV and hyperspectral imagery are often too costly to acquire over large areas and require substantial computing power to process, making them less suitable for regional and/or national applications. The employment of high-resolution (10-60 m) imagery — such as those generated by the

Sentinel-2 constellation as part of the Copernicus Programme — is more viable, as it is freely available, frequently updated and can be cost-effectively processed on cloud computing platforms such as Google Earth Engine (GEE). Nomura and Mitchard (2018) used Sentinel-2 data with supervised classification to map seven commercial forest plantation species/genera in Myanmar. They achieved a 95% overall accuracy (OA) using an unbalanced training dataset. Mngadi et al. (2019) classified seven species in the Clan forest plantation (located in South Africa) using Sentinel-2 bands in a linear discriminative analysis (LDA) and achieved an 84% OA. When adding Sentinel-1 vertical-vertical (VV) and vertical-horizontal (VH) features, the accuracies increased to 87%.

Image classification is generally done by collecting training data in one image and applying the classifier to that same image scene (Knorn et al. 2009). This is problematic when classifying large areas that are covered by multiple image scenes, as the training data collection becomes time-consuming and costly. In ML modelling, signature extension or generalisation reduces the effort and cost of training data collection (Pax-Lenney et al. 2001). Signature extension is the process of training a model on one image scene and applying it to other image scenes (Laborte et al. 2010). However, many factors influence its effectiveness and it is not known whether signature extension is a viable strategy for mapping forest plantation genera over large regions. There is often a trade-off between extension distance and classification accuracy, while climatic variations can also negatively influence signature extension efficiency. These factors may outweigh the benefits of signature extension in complex regions such as South Africa, where forest plantations are sparsely distributed and where climate gradients are dramatic (range from subtropical summer rainfall to semi-arid winter rainfall). A further complicating factor is that genera are planted in unequal proportions throughout South Africa, with pine trees being favoured in the Western and Eastern Cape provinces, while most plantations in KwaZulu-Natal and Mpumalanga are planted with *Eucalyptus* trees.

To date, no published research has attempted to specifically map *Acacia*, *Eucalyptus*, and *Pinus* genera using RS in a South African context. Furthermore, no published research has attempted to use signature extension to map forest plantation genera over a large extent in South Africa. An MSc study, carried out by Ms C Higgs, that formed part of this project aimed to evaluate RS and ML methodologies to map forest plantation genera on a national scale. Specifically, it aimed to assess the importance of different sampling strategies on ML accuracies and to investigate the potential of signature extension for reducing training sample collection costs.

Table 4-1 A summary table showing the overall accuracy, the standard deviation of the overall accuracy, kappa statistic, the standard deviation of the kappa statistic, consumer's and user's accuracy and the maximum OA and KS of the 100 iterations per sample size of experiments A to G conducted on Study Area 1 (WC) and Study Area 2 (KZN)

		Study area 1 (WC)												Study Area 2 (KZN)											
						Consumers Accuracy			Producers Accuracy									Consumers Accuracy			Producers Accuracy				
Experiment	Iteration	OA	OA Std	KS	KS Std	Acacia	Eucalyptus	Pine	Acacia	Eucalyptus	Pine	Max OA	Max KS	OA	OA Std	KS	KS Std	Acacia	Eucalyptus	Pine	Acacia	Eucalyptus	Pine	Max OA	Max KS
	1	0.586	0.038	0.380	0.056	0.538	0.628	0.583	0.371	0.767	0.621	0.703	0.555	0.459	0.046	0.189	0.069	0.458	0.487	0.440	0.449	0.493	0.435	0.553	0.330
Exp. A (Even samples)	20	0.715	0.020	0.573	0.031	0.665	0.772	0.705	0.614	0.802	0.729	0.763	0.645	0.626	0.025	0.439	0.037	0.622	0.638	0.619	0.611	0.652	0.617	0.666	0.499
	40	0.744	0.020	0.616	0.031	0.697	0.803	0.731	0.655	0.828	0.748	0.800	0.700	0.656	0.015	0.483	0.022	0.660	0.664	0.644	0.645	0.674	0.649	0.683	* **0.525**
	60	* **0.763**	0.023	0.644	0.034	0.723	0.818	0.747	0.684	0.840	0.764	* **0.813**	0.720	* **0.674**	0.011	0.511	0.017	0.675	0.681	0.667	0.669	0.685	0.669	0.680	0.520
	Mean	* **0.725**	0.022	0.588	0.033	0.676	* **0.780**	0.718	0.625	* **0.816**	0.734	0.770	0.655	* **0.618**	0.020	0.427	0.030	0.622	* **0.629**	0.606	0.615	* **0.624**	0.616	0.670	0.505
Exp. B (50 acacia)	20	0.600	0.015	0.400	0.022	0.700	0.661	0.545	0.061	0.880	0.858	0.637	0.455	0.498	0.025	0.247	0.037	0.638	0.526	0.467	0.069	0.699	0.727	0.560	0.340
	40	0.610	0.012	0.415	0.018	0.716	0.677	0.550	0.043	0.907	0.879	0.650	0.475	0.506	0.020	0.259	0.030	0.641	0.542	0.470	0.039	0.733	0.745	0.557	0.335
	60	0.608	0.013	0.413	0.019	0.740	0.677	0.548	0.036	0.910	0.880	0.630	0.445	0.510	0.018	0.265	0.027	0.572	0.545	0.478	0.026	0.746	0.756	0.563	0.345
	Mean	0.606	0.017	0.408	0.025	0.683	0.661	0.556	0.073	0.895	0.849	0.639	0.458	0.499	0.024	0.249	0.036	0.585	0.531	0.468	0.076	0.699	0.723	0.560	0.340
Exp. C (50 eucalyptus)	20	0.586	0.037	0.379	0.055	0.523	0.832	0.585	0.702	0.296	0.759	0.657	0.485	0.525	0.026	0.287	0.039	0.536	0.709	0.493	0.723	0.131	0.720	0.573	0.360
	40	0.589	0.029	0.384	0.044	0.547	0.893	0.576	0.740	0.233	0.795	0.653	0.480	0.537	0.020	0.306	0.030	0.538	0.787	0.519	0.775	0.085	0.751	0.577	0.365
	60	0.583	0.030	0.374	0.045	0.553	0.908	0.563	0.757	0.197	0.793	0.663	0.495	0.534	0.018	0.301	0.027	0.531	0.810	0.521	0.779	0.068	0.754	0.607	0.410
	Mean	0.585	0.033	0.377	0.049	0.533	0.843	0.581	0.704	0.285	0.765	0.658	0.487	0.530	0.024	0.296	0.036	0.533	0.756	0.508	0.736	0.131	0.724	0.586	0.378
Exp. D (50 pine)	20	0.638	0.025	0.457	0.037	0.544	0.682	0.900	0.783	0.838	0.293	0.693	0.540	0.527	0.025	0.291	0.038	0.521	0.520	0.651	0.727	0.745	0.111	0.577	0.365
	40	0.651	0.021	0.476	0.031	0.557	0.701	0.946	0.822	0.877	0.253	0.703	0.555	0.531	0.023	0.297	0.034	0.531	0.523	0.672	0.765	0.759	0.070	0.590	0.385
	60	0.650	0.024	0.475	0.035	0.552	0.717	0.956	0.838	0.894	0.218	0.700	0.550	0.539	0.021	0.309	0.032	0.540	0.528	0.733	0.780	0.776	0.062	0.587	0.380
	Mean	0.639	0.025	0.459	0.037	0.550	0.686	0.890	0.781	0.851	0.287	0.699	0.548	0.526	0.024	0.290	0.036	0.521	0.521	0.655	0.731	0.734	0.113	0.584	0.377
Exp. E (50 acacia & 50 eucalyptus)	20	0.514	0.041	0.271	0.062	0.547	0.787	0.432	0.100	0.504	0.938	0.573	0.360	0.415	0.041	0.123	0.062	0.516	0.692	0.371	0.130	0.194	0.921	0.527	0.290
	40	0.487	0.036	0.231	0.054	0.529	0.815	0.413	0.067	0.424	0.970	0.550	0.325	0.402	0.031	0.103	0.046	0.543	0.751	0.364	0.102	0.151	0.953	0.493	0.240
	60	0.450	0.034	0.176	0.051	0.507	0.811	0.390	0.051	0.328	0.972	0.513	0.270	0.390	0.022	0.085	0.033	0.549	0.778	0.358	0.068	0.129	0.972	0.470	0.205
	Mean	0.506	0.035	0.260	0.052	0.527	0.781	0.436	0.110	0.481	0.928	0.546	0.318	0.414	0.034	0.122	0.050	0.538	0.702	0.374	0.143	0.189	0.911	0.497	0.245
Exp. F (50 acacia & 50 pine)	20	0.542	0.026	0.313	0.040	0.609	0.478	0.709	0.200	0.955	0.470	0.620	0.430	0.434	0.032	0.151	0.047	0.560	0.391	0.611	0.159	0.917	0.225	0.507	0.260
	40	0.513	0.023	0.269	0.035	0.584	0.444	0.768	0.144	0.976	0.419	0.570	0.355	0.411	0.029	0.116	0.043	0.579	0.374	0.635	0.107	0.957	0.168	0.490	0.235
	60	0.507	0.026	0.260	0.039	0.574	0.437	0.778	0.116	0.983	0.422	0.560	0.340	0.389	0.027	0.083	0.040	0.560	0.364	0.595	0.088	0.968	0.111	0.450	0.175
	Mean	0.531	0.028	0.297	0.042	0.596	0.471	0.726	0.177	0.956	0.460	0.583	0.375	0.423	0.030	0.135	0.045	0.558	0.388	0.614	0.154	0.912	0.205	0.482	0.223
Exp. G (50 eucalyptus & 50 pine)	20	0.572	0.036	0.358	0.054	0.450	0.829	0.832	0.913	0.442	0.360	0.650	0.475	0.459	0.031	0.188	0.046	0.415	0.588	0.582	0.915	0.219	0.243	0.553	0.330
	40	0.535	0.037	0.302	0.056	0.425	0.883	0.858	0.958	0.356	0.290	0.607	0.410	0.445	0.028	0.168	0.042	0.397	0.601	0.700	0.959	0.159	0.218	0.517	0.275
	60	0.529	0.034	0.293	0.051	0.420	0.895	0.900	0.969	0.333	0.285	0.623	0.435	0.442	0.029	0.163	0.043	0.394	0.622	0.727	0.976	0.147	0.204	0.550	0.325
	Mean	0.553	0.035	0.329	0.052	0.442	0.832	0.828	0.906	0.411	0.341	0.627	0.440	0.459	0.031	0.189	0.046	0.417	0.582	0.630	0.914	0.216	0.247	0.540	0.310
Exp. H (area-proportional samples)	20	0.531	0.034	0.297	0.052	0.599	0.761	0.448	0.124	0.545	0.925	0.572	0.345	0.626	0.025	0.439	0.037	0.622	0.638	0.619	0.611	0.652	0.617	0.666	0.499
	40	0.549	0.028	0.323	0.042	0.621	0.789	0.457	0.134	0.574	0.937	0.594	0.376	0.656	0.015	0.483	0.022	0.660	0.664	0.644	0.645	0.674	0.649	0.683	0.525
	60	0.560	0.025	0.340	0.037	0.624	0.807	0.464	0.137	0.601	0.940	0.632	0.410	0.674	0.011	0.511	0.017	0.675	0.681	0.667	0.669	0.685	0.669	0.680	0.520
	Mean	* **0.530**	0.028	0.295	0.043	0.573	0.789	0.443	0.108	0.545	0.938	0.587	0.365	0.618	0.020	0.427	0.030	0.622	0.629	0.606	0.615	0.624	0.616	0.670	0.505

Key: Low accuracies [colour scale] High accuracies; Low Std [colour scale] High Std; Overall accuracy (OA); Standard deviation (Std); Kappa statistic (KS); Maximum (Max); Important results (* Bold text)

The research set out two main experiments. The first experiment evaluated the impact of using an even, uneven or an area-proportionate training sample configuration and size in a RF ML model for classifying *Acacia*, *Eucalyptus*, and *Pinus* compartments. The model was built using Sentinel-2 bands, vegetation indices and textural measures (37 features in total). It was found that the study area that contained an uneven area planted with *Acacia*, *Eucalyptus*, and *Pinus* trees was classified more accurately, shown in Table 4-1, using a balanced training sample configuration (76%), compared to using an unbalanced (~ 55%) and area-proportionate training sample configuration (53%). It was also found that a saturation point exists where adding more training samples added little value to the OA. The saturation point was found to be ~ 57n, where n is the number of features used in the classification.

The second set of experiments was set out to test the viability of training data signature extension for constructing RF ML models to differentiate between *Acacia*, *Eucalyptus*, and *Pinus* trees using Sentinel-2 imagery as input. The study area was split into 19 Sentinel-2 tiles spanning the Mpumalanga, KwaZulu-Natal, Eastern Cape and WC provinces. Three separate RF models were built using training data collected in one tile located in Mpumalanga (Experiment 1), one tile located in KwaZulu-Natal (Experiment 2), and one tile located in the Eastern Cape (Experiment 3). A fourth model was built using training data from all three source tiles (Experiment 4). The four models were applied to all 19 Sentinel-2 tiles to map forest plantation genera. The overall accuracies from the first three models were correlated against rainfall seasonality, extension distance (i.e. distance between the training data and the area being classified), and temperature seasonality. Figure 4-6 shows that the OAs decrease as extension distance increases, with a general decline of 3%, 6% and 2% per 100 km for Experiment 1, 2 and 3 respectively.

The OAs are frequently below 50% when the extension distance exceeded 500 km. The statistical relationship between OA and extension distance was strong ($R^2 = 0.723$) in Experiment 2, but very weak (R2 = 0.18) in Experiment 3. In general, higher accuracies were obtained for tiles with a similar rainfall seasonality to the source tile. In Experiment 1, the statistical relationship between OA and rainfall seasonality was stronger ($R^2 = 0.5$) than extension distance ($R^2 = 0.403$), but weaker ($R^2 = 0.413$) in Experiment 2. As with extension distance, the relationship with rainfall seasonality was weak ($R^2 = 0.098$) in Experiment 3. The relationship between OA and temperature is weak ($R^2 < 0.11$) in all three experiments.

In summary, the OAs have the strongest relationship with rainfall seasonality in Experiment 1, whereas extension distance was the strongest driver of OAs in Experiments 2 and 3. Temperature has the weakest effect on the OA for all experiments.

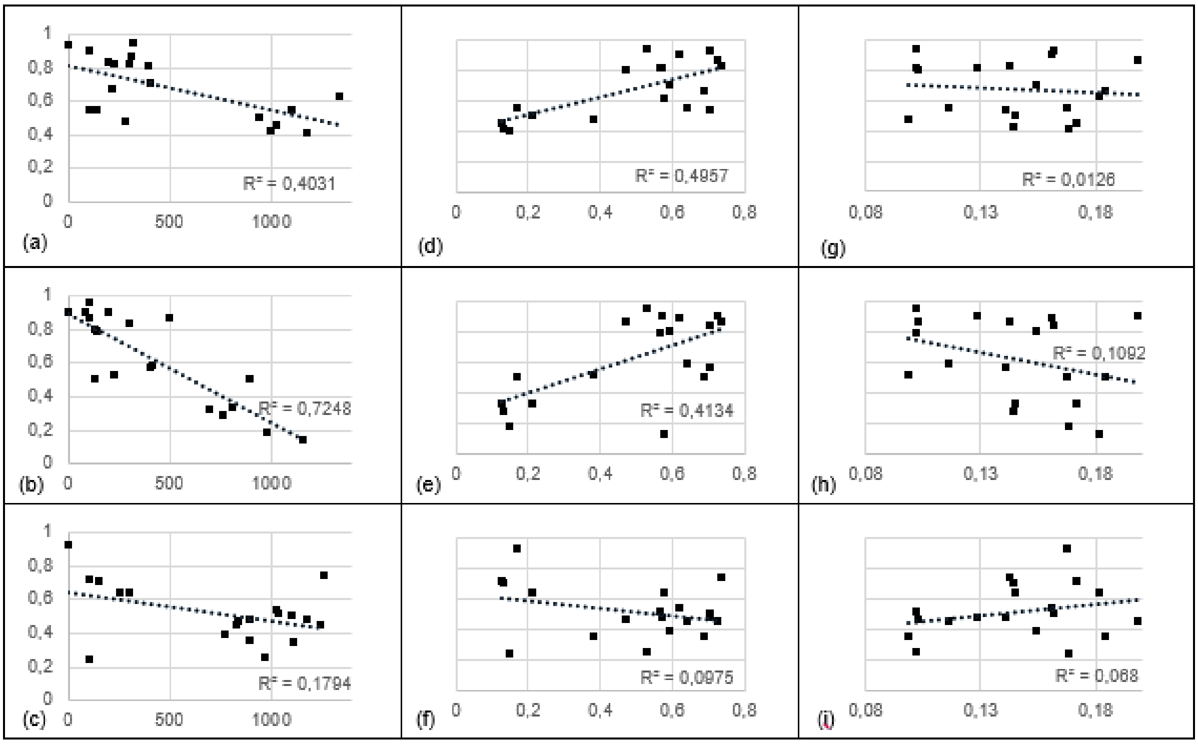


Figure 4-6 Overall accuracy vs distance from the source tile for exp 1 (Tile 4) (a). exp 2 (Tile 10) (b), and exp 3 (Tile 17) (c), overall accuracy vs rainfall seasonality index for exp 1 (d), exp 2 (e), and exp 3 (f), and overall accuracy vs temperature index for exp 1(g), exp 2 (h), exp 3 (i)

This is in agreement with Olthof et al. (2005), who generated spectral signatures in a source scene and applied them to other scenes. The OAs were strongly affected by extension distance. Similar conclusions were drawn by Knorn et al. (2009), who found a 1.9% average decrease in OA as extension distance increased by one Landsat scene (i.e. ~ 1% per 100 km), while we observed a mean decrease of ~ 4% in OA per 100 km increase in extension distance.

Our findings are also in agreement with Woodcock et al. (2001), who found that high accuracies can be achieved when the source and classified tiles are in the same climatic regions, but when the model is applied across different climatic regions, the accuracies decrease. They concluded that the accuracies are influenced by extension distance and climatic variation. Our findings show that lower accuracies are achieved when models are trained with data collected in areas with different rainfall seasonality to the area being mapped.

The results show that a ~ 70% OA can be achieved if the training data is collected in areas with similar climates (rainfall seasonality) to the areas that are being mapped. In addition, it was found that signature extension distance (i.e. distance between the training data and the area being classified) should not exceed 500 km.

The guidelines developed in this research can contribute towards regularly mapping forest plantation genera at regional scales and with minimal costs. This study purposefully used a composite image spanning a year to compare the results of each experiment. However, it is known that multitemporal variables can improve the separability between genera as phonological characteristics can be represented (Fagan et al. 2015). Mngadi et al. (2019) also showed that fusing active satellite imagery with optical imagery can help ML algorithms to differentiate between forest plantation genera/species. It is consequently recommended that multitemporal variables and SAR imagery be considered in future studies. However, this will significantly increase the dimensionality of the dataset, and feature selection methods to filter out the ineffective bands will likely have to be implemented, which in theory should further reduce the number of samples needed to differentiate between forest plantation genera.

4.3 Compartment age estimation

Knowing the age of a plantation (planting date) is critical for forest management (rotation scheduling and yield estimation). Tree age furthermore has a significant impact on water use (see Section 5.2). Compartment age can also be used in land cover mapping to differentiate between indigenous and plantation forests, given these two classes often have very similar spectral characteristics, which causes substantial confusion during image classification (Section 4.1). The age of (indigenous and plantation) forests can be used as an additional discriminator (i.e. forests older than 40 years are unlikely to be planted).

Using a time series of annual NDVI covering an area of 1 500 ha in southern Brazil, Le Maire et al. (2011) used MODIS imagery to estimate *Eucalyptus* plantation age, achieving a RMSE of 40 days. Chen et al. (2012) estimated the age of a rubber plantation in a multivariate regression analysis using normalised and ratio-based indices consisting of the red and NIR bands derived from Landsat imagery and achieved an RMSE of 5.96 years. Kou et al. (2015) achieved similar results by using NDVI, EVI and LSWI derived from Landsat imagery to identify the planting date of rubber plantations. McMahon and Jackson (2019) used Landsat images to derive a NIR vegetation index time series over a 2.2 million ha area in America to estimate the planting date of forest stands, and 86% of the stands were predicted to be within one year of its actual age.

To date, no published research has attempted to estimate forest compartment age using RS in a South African context. The aim of this BSc Honours research project, carried out by Mr M Mdwayi, was to evaluate how multitemporal Landsat-8 imagery can be used to estimate the plantation age of 2 804 compartments in the Richards Bay (KwaZulu-Natal) area. The study area was selected as it contained compartments of different ages and genera. Annual NDVI composites from 2013 to 2020 were generated using GEE. The resulting time series of annual NDVI values were then statistically analysed to determine the year when the NDVI values were lowest (i.e. minimum annual NDVI). The assumption was that the minimum annual NDVI would coincide with the planting date. The median ages (per pixel) were then extracted per compartment using zonal statistics (see Section 3.3.1) to estimate the compartment age.

The results (Table 4-2) showed that compartment age can be estimated to within one year of its actual age (MAE = 0.84, RMSE = 1.0 years). It was more difficult to estimate the age of *Acacia* (MAE = 1.36, RMSE = 1.31) compartments, compared to *Pinus* (MAE = 0.92, RMSE = 1.2) and *Eucalyptus* (MAE = 0.85, RMSE = 0.99).

Table 4-2 Forest compartment age estimation accuracies using multitemporal Landsat-8 imagery

Genus	Number of compartments	Mean absolute error (MAE)	Root mean square error (RMSE)
All	2804	0.84	1.00
Eucalyptus	2704	0.85	0.99
Pinus	79	0.92	1.20
Acacia	11	1.36	1.31

Closer inspection revealed that the closed canopies of *Eucalyptus*, combined with their rapid growth subsequent to planting, reduced the influence of the background soil signal, which is known to negatively impact the ability of NDVI to capture vegetation vigour. Although the sampled area was relatively small and dominated by *Eucalyptus* stands, the results are encouraging and warrant further investigation. It is recommended that future studies investigate the use of other vegetation indices (e.g. soil-adjusted vegetation index, SAVI and EVI) and other sources of imagery (e.g. Landsat-5 and Sentinel-2). It is also recommended that the methodology be tested in other areas in South Africa (and beyond) to evaluate its robustness (transferability).

5 PLANTATION FOREST WATER USE

Plantation water use is impacted by various factors including genus, species, tree age, litter and understorey vegetation; climatic conditions like rainfall, solar radiation and general climate zone; and site conditions like soil type, slope and terrain morphology. In the section below we describe the impact of selected factors on plantation water use for *Acacia*, *Eucalyptus* and *Pinus* grown in SA.

5.1 Water use per genus

The frequency distributions of the annual water use per genus were discussed in Section 3.4.1 and are not repeated here. Instead, Figure 5-1 shows the average annual ET per genus from 2009 to 2019, irrespective of plant date and rotations. The water use of *Eucalyptus* compartments is consistently higher than *Pinus* compartments throughout the 10-year period. The average water use of *Acacia* and *Eucalyptus* compartments was similar from 2009 to 2013, after which the mean ET of *Acacia* decreased substantially, reaching a low of 961 mm/year in 2017. It is clear from the graph that there is a downward trend in water use from 2013 to 2017, corresponding with the drought conditions in South Africa at the time. The water use of all three genera increased substantially from 2017 onwards, presumably as the rainfall increased. Unfortunately, this could not be confirmed due to lack of access to detailed climate data per compartment (see proposals for further research, Section 6.4).

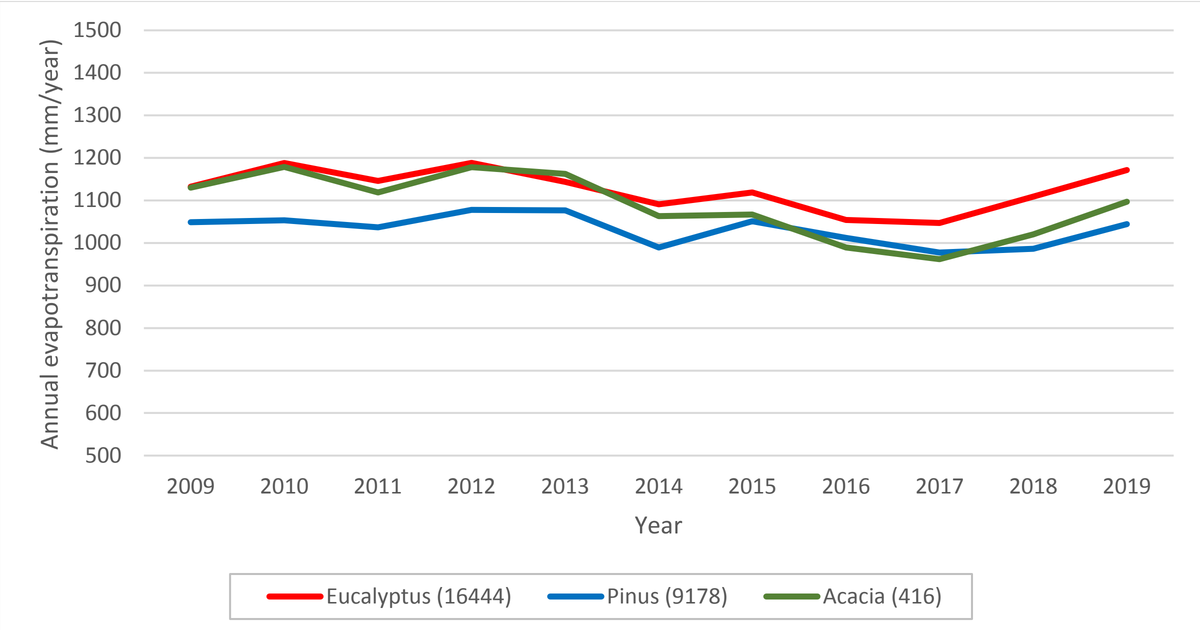


Figure 5-1 Average annual evapotranspiration per genus (mm/year), irrespective of plant date (compartment age)

Figure 5-1 was produced from all the compartments in the database, irrespective of when they were planted. The number of compartments (cases) consequently vary from year to year, with the majority (63%) being planted with *Eucalyptus* species and only a fraction (1.6%) being planted with *Acacia* species.

According to the literature (Section 3.4.2), there is a strong relationship between water use and stand age, with an exponential increase in ET during the first years after planting, whereafter ET rates reach a plateau. This trend is not visible in Figure 5-1 but the following subsections explore this relationship in more detail.

5.2 Water use per species

Selection of species for water use analysis was based on direction from the reference group, provided there was a significant number of compartments to return valid results. The species selected for further analysis were *A. mearnsii* (394), *E. benthamii* (782), *E. badjensis* (68), *E. dunnii* (6120), *E. grandis* (1421), *E. grandis x E. nitens hybrid* (2668), *E. grandis x E. urophylla hybrid* (3872), *P. elliotii* (3439), *P. greggii* (40), *P. patula* (3265), *P. radiata* (439) and *P. taeda* (523). A breakdown of all the species listed in the CFDB is available in Appendix I.

The annual evapotranspiration per species according to age is given in Table 5-1 and Figure 5-2.

Table 5-1 Mean annual evapotranspiration (mm/year) per species according to age

AGE	0	1	2	3	4	5	6	7	8	9	10
Acacia mearnsii	917	959	1059	1090	1091	1104	1099	1099	1127	1119	1119
Eucalyptus benthamii	798	857	989	1042	1043	1039	1020	1039	1064	1067	1093
Eucalyptus badjensis	762	825	997	1089	1038	983	985	1087	1154	-	-
Eucalyptus dunnii	924	1018	1115	1122	1116	1115	1114	1117	1134	1144	1145
Eucalyptus grandis	1200	1267	1322	1318	1308	1285	1289	1287	1273	1271	1271
Eucalyptus grandis x E. nitens hybrid	841	916	1061	1079	1081	1088	1092	1084	1075	1092	1149
Eucalyptus grandis x E. urophylla hybrid	1025	1219	1280	1291	1315	1305	1336	1365	1386	1372	1343
Pinus elliottii	808	806	851	898	954	1011	1057	1089	1117	1141	1161
Pinus greggii	779	867	1019	1039	1166	1198	1218	1228	1307	1216	1250
Pinus patula	724	739	807	893	968	1009	1033	1070	1076	1070	1078
Pinus radiata	880	871	865	844	884	926	949	973	1014	1039	1086
Pinus taeda	837	843	910	973	1033	1049	1068	1108	1131	1153	1181

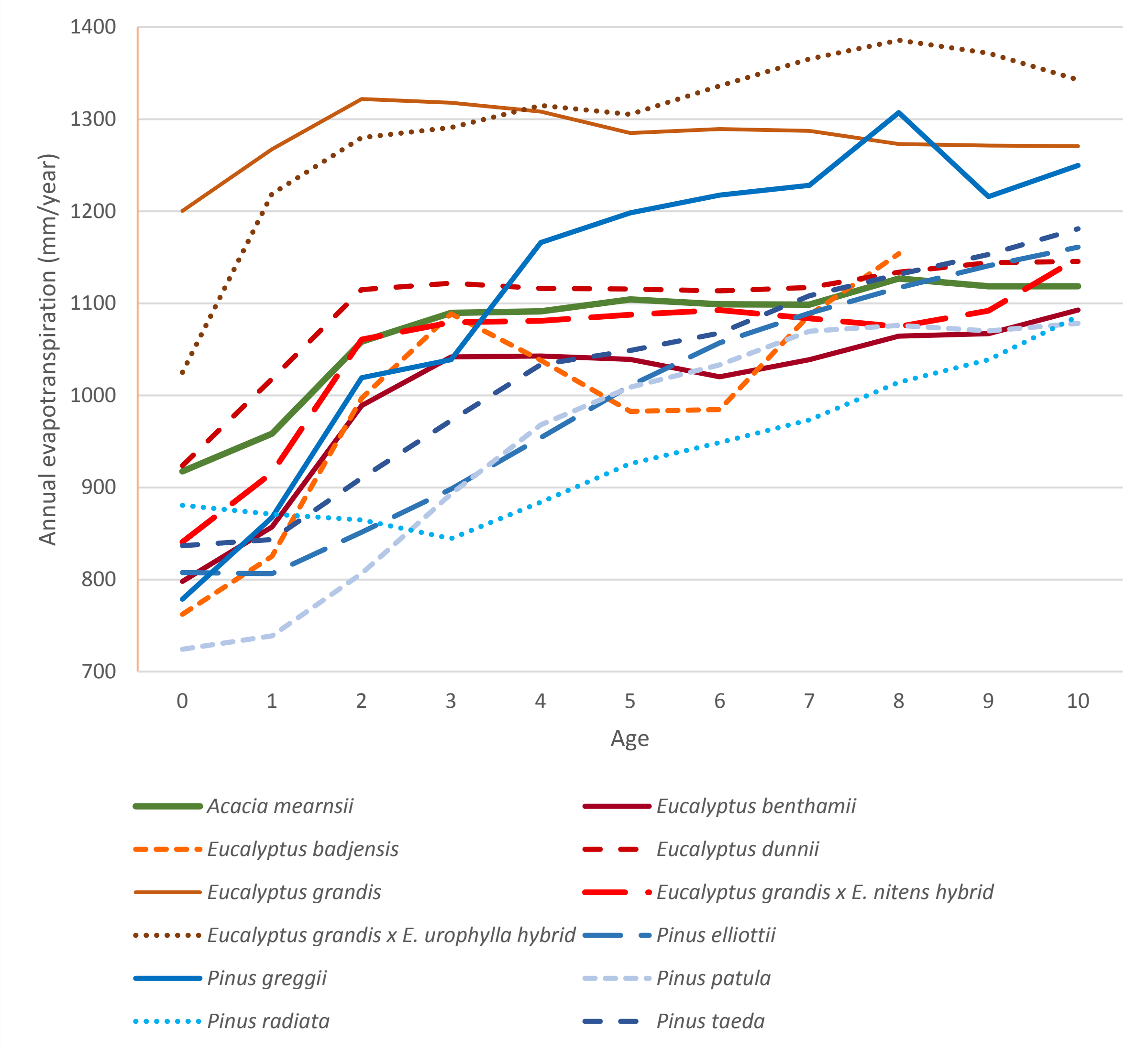


Figure 5-2 Average annual evapotranspiration (mm/year) per species according to age

Only one species of *Acacia* had enough compartments in the CFDB for the water use calculations: *A. mearnsii* (Figure 5-3). An increase in ET from planting to year 4 (917 to 1 090 mm/year) is noticeable, suggesting a high growth rate and subsequent ET at the time. Thereafter, the annual ET increased slightly from year 4 to 10 (1 090 to 1 119 mm/year), which suggests a reduced growth rate. This finding agrees with those of an unpublished field-based study carried out at the University of KwaZulu-Natal in collaboration with the Institute for Commercial Forestry Research and Sappi Forests. Information on management activities (e.g. thinning) was not available, but it is likely that thinning will impact the ET estimates.

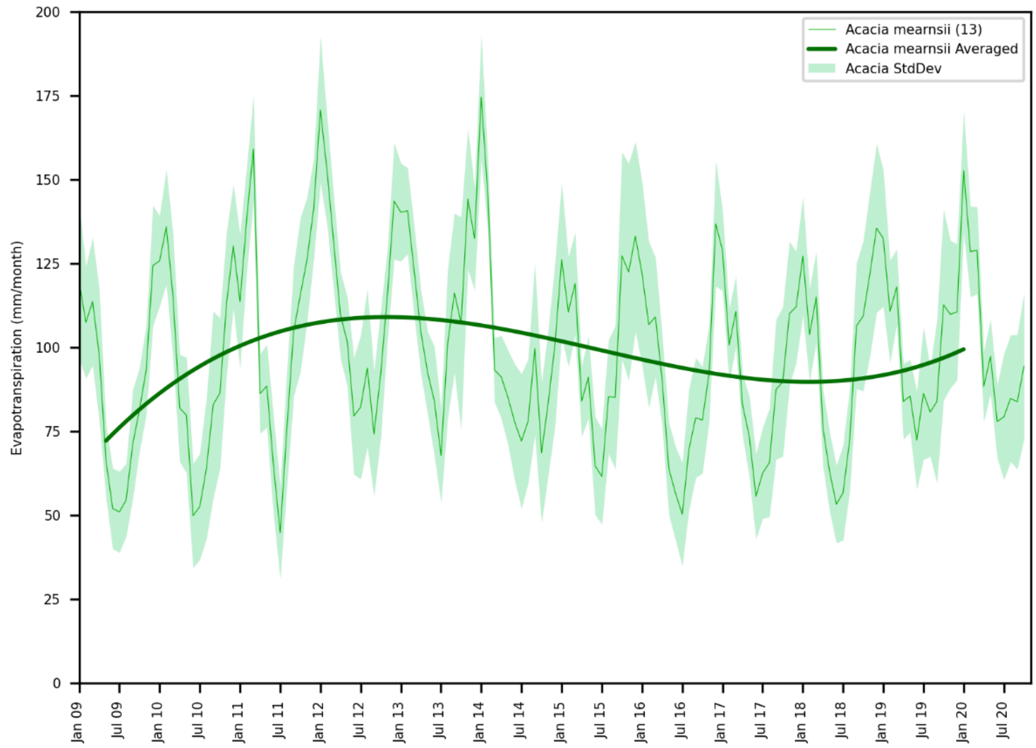


Figure 5-3 ET over time (mm/month) for ***A. mearnsii***

Five *Eucalyptus* species/hybrid species were considered for the water use comparisons: *E. benthamii*, *E. dunnii*, *E. grandis*, *E. grandis x E. nitens* hybrid and *E. grandis x E. urophylla* hybrid. The comparative annual ET according to age for each species is shown in Figure 5-4.

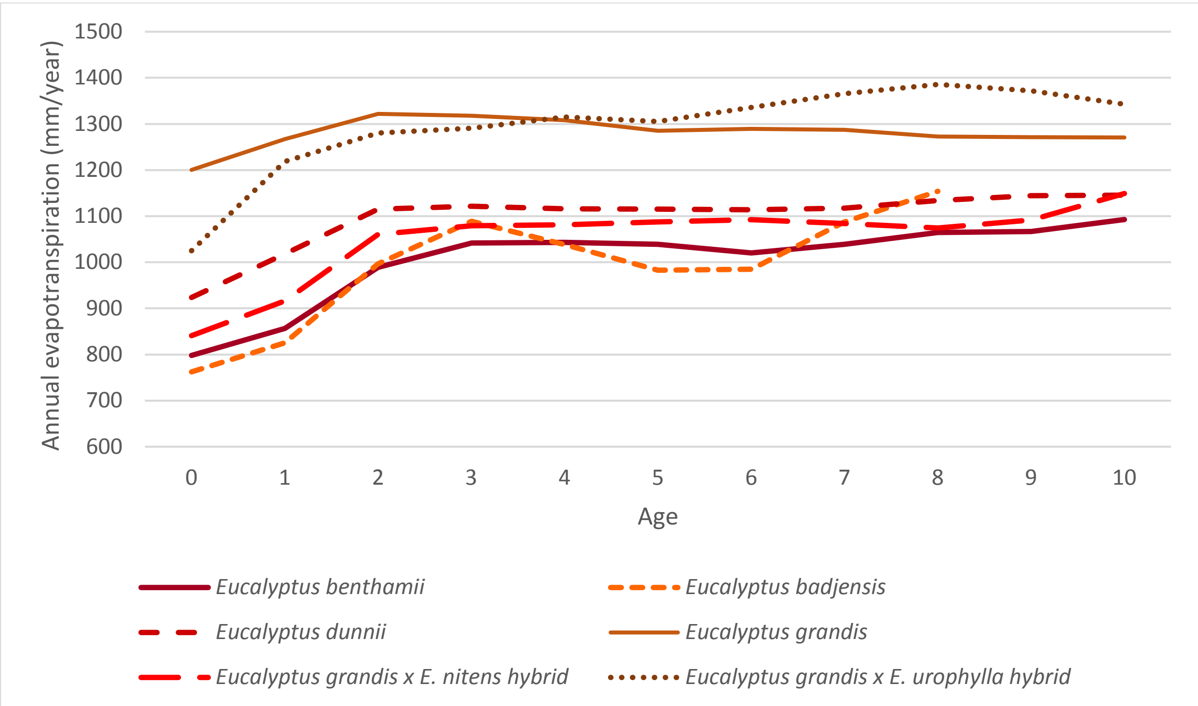


Figure 5-4 **Annual evapotranspiration (mm/year) according to age for all *Eucalyptus* species analysed**

There are marked differences in ET among compartments planted with different *Eucalyptus* species, especially during the first year after planting (i.e. year 0-1 on the graph). ET ranged from 762 mm/year (*E. badjensis*) to more than 1 386 mm/year (*E. grandis x E. urophylla hybrid*). *E. grandis x E. urophylla hybrid* showed the largest increase in ET over year 1 (from 1 025 mm/year to 1 219 mm/year), indicating a combination of vigorous growth and high transpiration rates, or evaporation from the soil and canopy intercepted water. The ET increases of the other species during the first two years after planting were more gradual. The ET estimates for most species appear to plateau (flattening of curve) after year 2 or 3. This appears to be on par with ET estimates of previous studies (Figure 3-23).

At 10 years of age, the highest ET was recorded for *E. grandis x E. urophylla hybrid* (1 343 mm/year) species and the lowest for *E. benthamii* (1 093 mm/year). The decrease in ET for *E. badjensis* from age 3-5 and the subsequent increase after year 6 suggest either drought conditions, water shortages, or some tree damage (possibly due to frost). It is not clear whether these differences are only attributed to species-specific differences in water use, as position in the landscape and environmental conditions would also have impacted ET values.

More details on the seasonal ET of each of these *Eucalyptus* species are given in Appendix III.

The *Pinus* species/hybrid considered for the water use comparisons were *P. elliotii, P greggii, P. patula, P. radiata* and *P. taeda.* The annual ET according to age for each species is shown Figure 5-5. *P. radiata* had the highest ET in year 1, although *P. greggii* showed the biggest (exponential) increase in ET in the first year after planting, exceeding all other species considered. This is followed by *P. taeda*. The high *P. radiata* ET after planting likely reflect a large component of soil evaporation and sufficient soil moisture (rainfall) to drive ET. The ET from *P. radiata* compartments remained similar for the first three years, whereafter it showed a steady increase in ET. However, over 10 years, this species had the lowest ET of those studied. In contrast to most of the *Eucalyptus* species, the slopes in ET of the *Pinus* species are gradual during the first few years and the flattening out is less pronounced, likely reflecting different tree growth curves. At 10 years of age, the annual ET for *P. greggii* exceeded that of *P. radiata* by more than 200 mm/year.

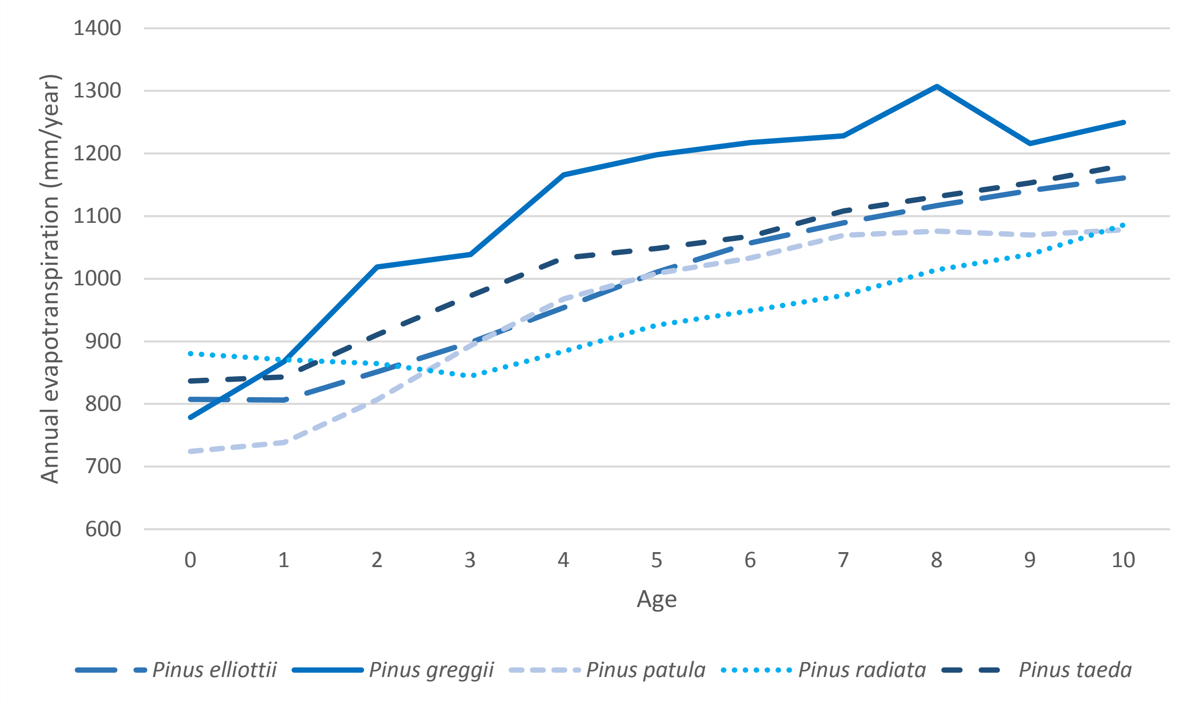


Figure 5-5 Annual evapotranspiration (mm/year) according to age for all *Pinus* species analysed

More details on the seasonal ET of each of these *Pinus* species are given in Appendix III.

5.3 Water use compared to tree age

It is clear from Figure 5-7 that water use increases during the first years after planting, and then stabilises. However, the increase in water use after planting is more dramatic for *Eucalyptus* compartments, which increase from an average of 941 mm/year to 1 151 mm/year within two years of planting, after which the water consumption remains constant. In contrast, the increase in water use of *Pinus* compartments is much more gradual during the first few years, starting at a mean of 805 mm/year and gradually increasing to 1 117 mm/year after ten years of growth. These observations match those of Dye (1996) (Figure 3-23) who noted a dramatic increase in the annual ET of *Eucalyptus* species over the first number of years after planting, while the initial increase for *Pinus* species is much lower. The water use curve of *Acacia* species is like that of *Eucalyptus* species, although ET was consistently lower, starting at 917 mm/year and reaching 1 102 mm/year after five years of growth, whereafter it stabilises.

It should be noted that the proportion of compartments planted with *Acacia* species is substantially (Appendix I) lower than those planted with *Eucalyptus* and *Pinus* and that the mean ET values of *Acacia* compartments may not be a good reflection of its true water use.

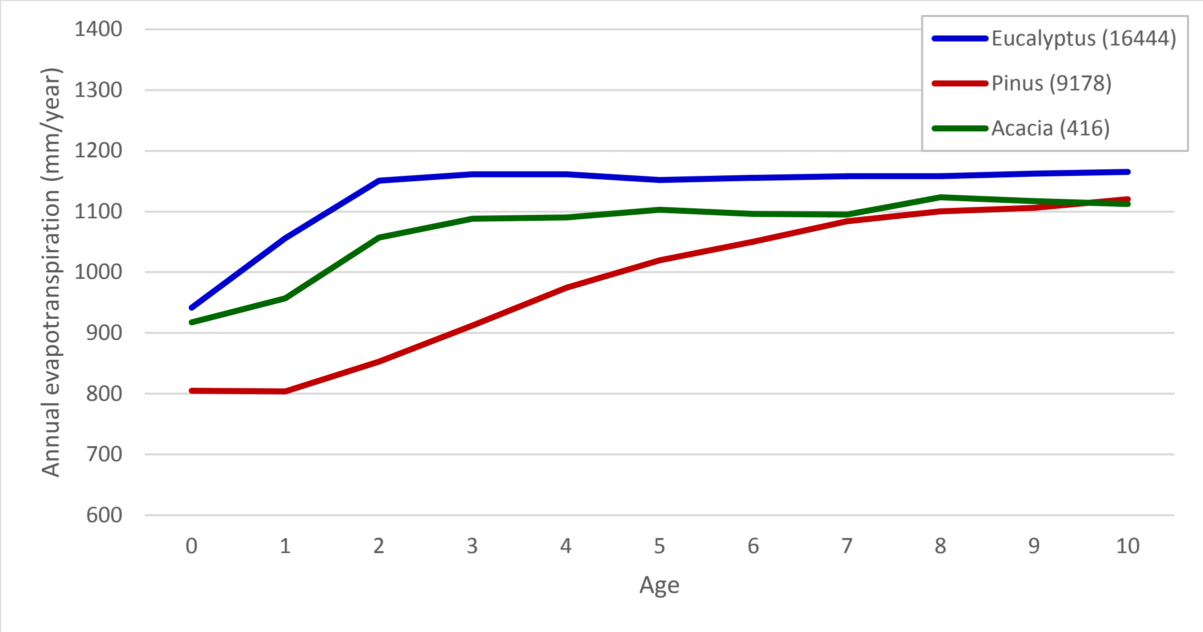


Figure 5-6 Average annual evapotranspiration (mm/year) per genus according to age

Similar trends are shown when the monthly ET values of compartments planted in 2009 are plotted against time (Figure 5-7). Based on 13 compartments planted with *Acacia* in 2009, the water use of *Acacia* compartments is higher than that of *Eucalyptus* compartments during the first six years of growth, but when considering a larger (40) set of compartments planted with *Acacia* in 2014, this genus uses less water compared to *Eucalyptus* during the first few years of growth. However, these results should be interpreted within the context of the drought conditions experienced from 2014-2017, the impact of which is clearly noticeable in Figure 5-7, where ET values dropped considerably for *Eucalyptus* and *Acacia* compartments in 2015. Interestingly, no drop in ET was observed for *Pinus* compartments, suggesting that the genus is less sensitive to drought conditions, or the impact of the drought (reduction on rainfall and water availability) on the areas planted with *Pinus* was less pronounced. Likewise, the water use estimations of *Acacia* may have been influenced by the fact that a larger proportion (19%) of *Acacia* compartments were planted during this (drought) time. The water use estimates for *Acacia* compartments should be interpreted within the context of the relatively few cases.

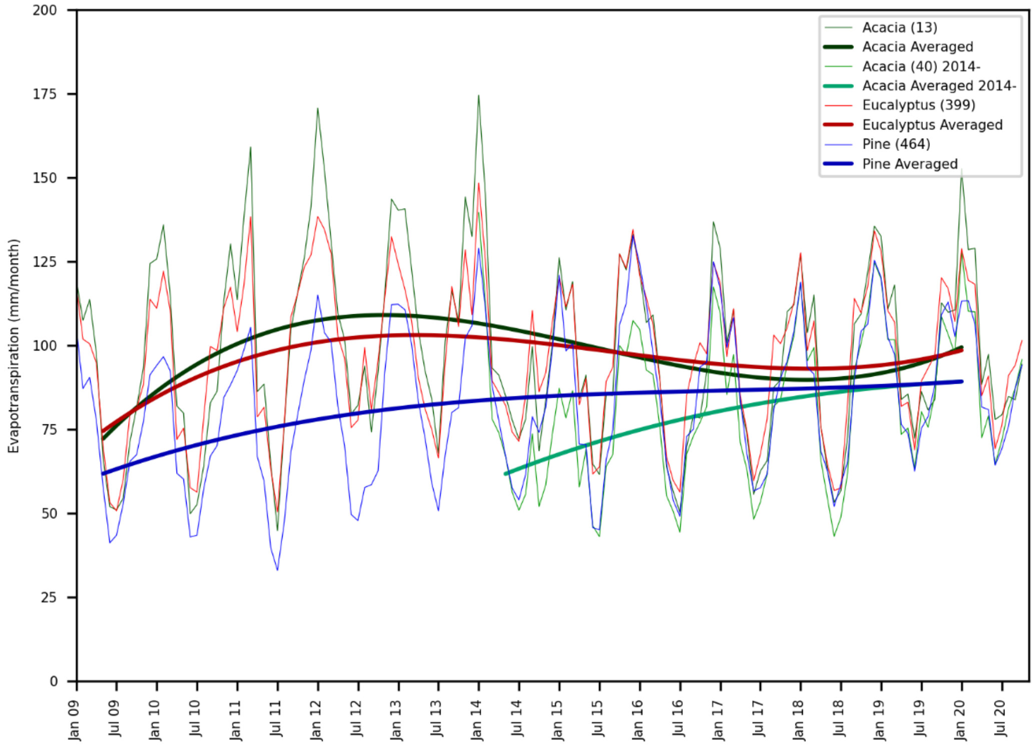


Figure 5-7 Monthly water use over time (mm/month) for *Acacia* (planted in 2009 and 2014), *Eucalyptus* and *Pinus*

The accumulated water use of each genus over a 10-year period is shown in Figure 5-8. The ET accumulation of *Eucalyptus* and *Acacia* during the initial years is much higher compared to *Pinus*. Ten years after planting, the accumulated ET is about 13 720 mm for *Eucalyptus*, compared to the less than 12 000 mm of *Pinus*. This constitutes a difference of about 13%.

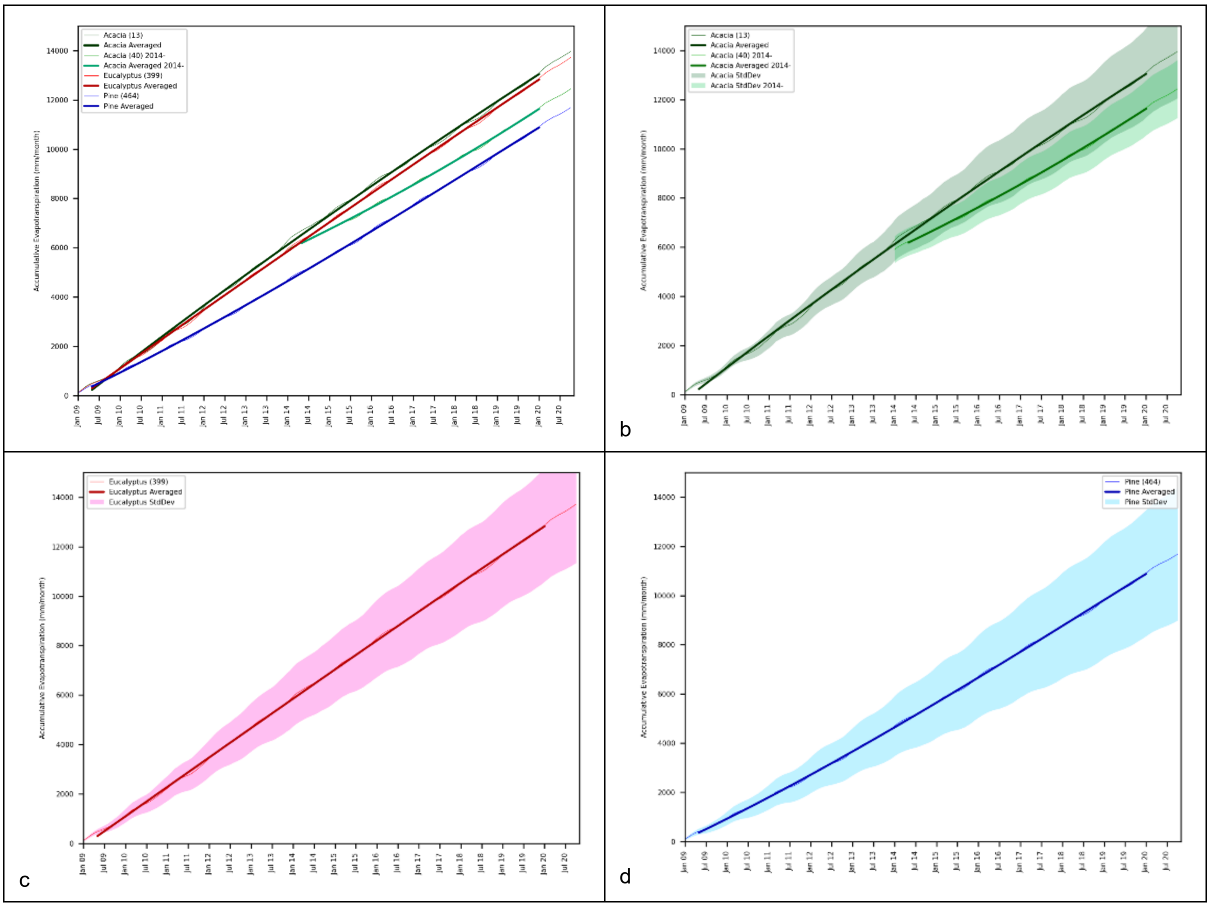


Figure 5-8 Accumulative water use over the first ten years of growth (a) in combination, as well as for (b) *Acacia* (planted in 2009 and 2014), (c) *Eucalyptus*, and (d) *Pinus* compartments separately (one standard deviation shown in lighter shades)

5.4 Water use compared to climatic conditions

This section compares the water use estimations to three climatic variables, namely mean annual rainfall, rainfall seasonality and climatic zones.

5.4.1 Mean annual rainfall

Figure 5-9 shows the mean annual rainfall, categorised in to low (< 600 mm), medium (600-800 mm), high (800-1000 mm) and very high (> 1 000 mm) ranges. Given that one compartment was located in the low rainfall region, the focus of this section is on the medium, high and very high rainfall categories.

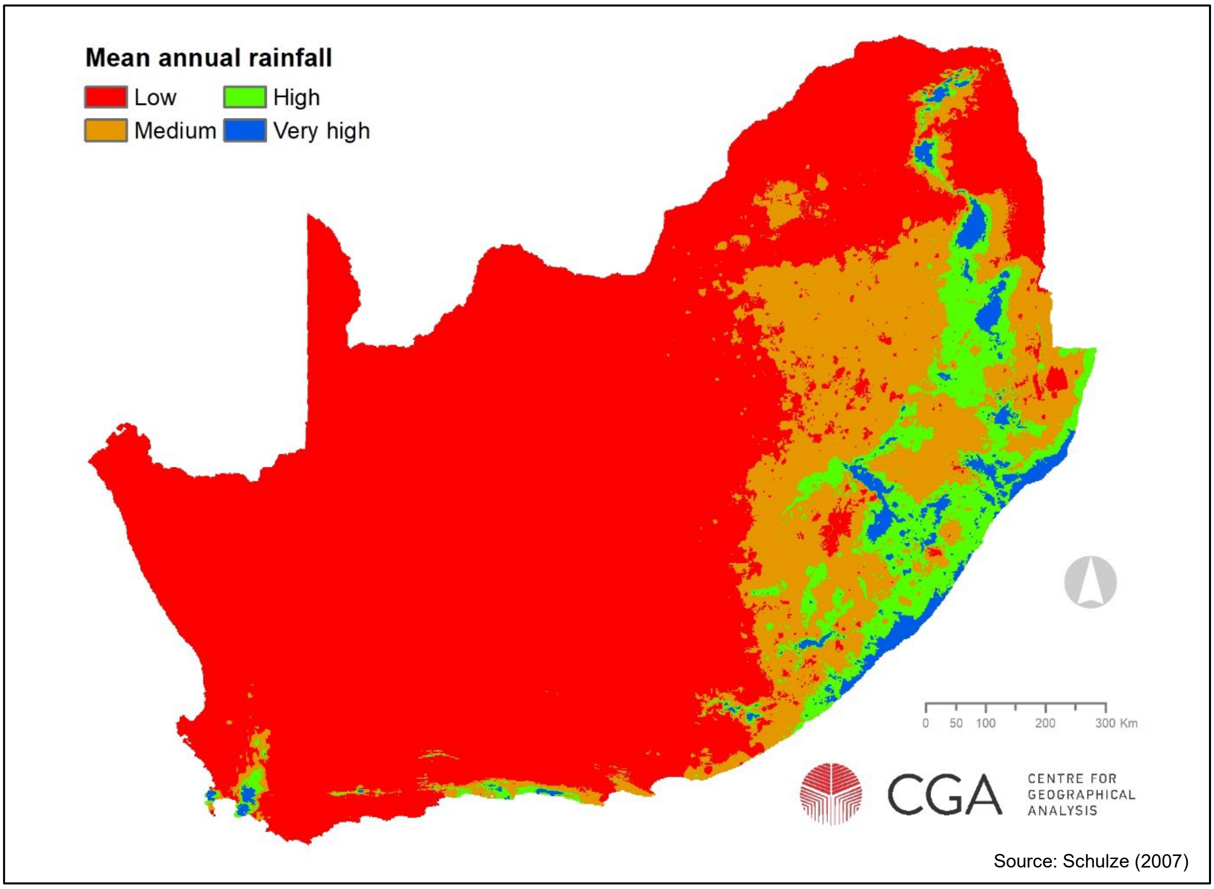


Figure 5-9 Mean annual rainfall

Table 5-2 summarizes the median ET values per rainfall category and genus. ET generally increases as rainfall (available water) increases. This ET likely includes a substantial fraction of evaporation of canopy intercepted rainfall, which may be even more pronounced in the fine-leafed *Acacia* trees. In areas with very high (> 1 000 mm/year) rainfall, the differences in water use among genera is negligible, while the difference in ET among genera is more dramatic in lower (medium and high) rainfall regions. This pattern is supported by the graphs in Figure 5-10 to Figure 5-12.

Table 5-2 Median annual water use (mm/year) per rainfall categrory and genus

Rainfall category	Evapotranspiration (mm/year)		
	Acacia	*Eucalyptus*	*Pinus*
Medium (600-800 mm/year)	-	1031 (n = 29)	805 (n = 129)
High (800-1000 mm/year)	1126 (n = 12)	1147 (n = 275)	993 (n = 277)
Very high (> 1000 mm/year)	1303 (n = 1)	1282 (n = 95)	1291 (n = 57)

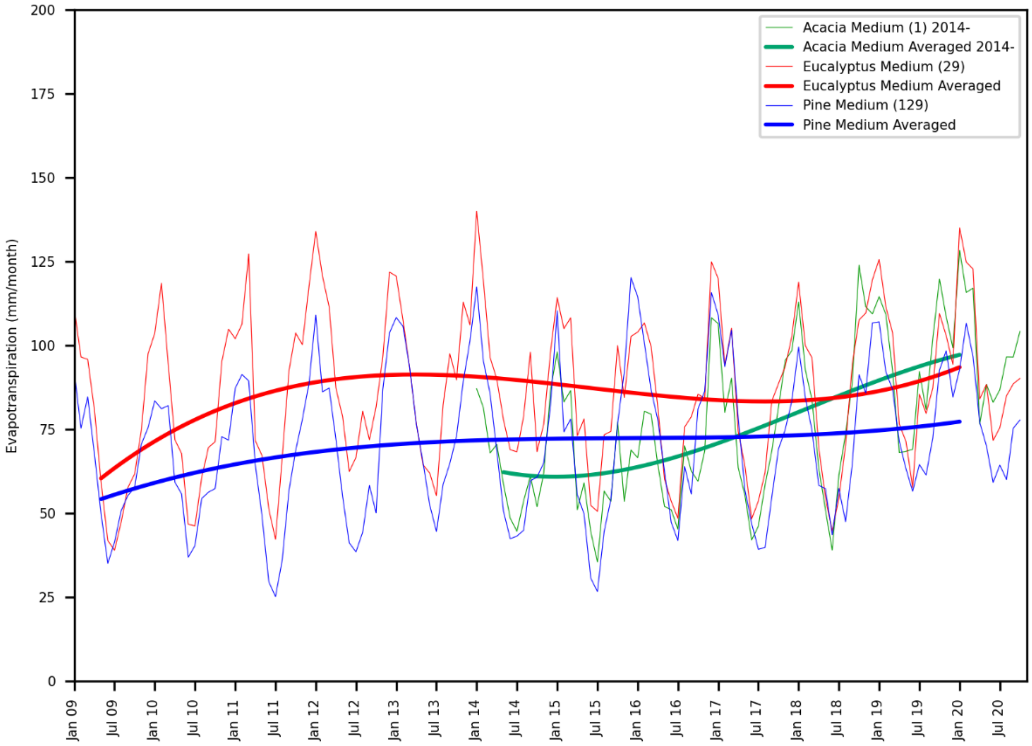


Figure 5-10 Water use over time (mm/month) for *Acacia* (planted in 2014), *Eucalyptus* and *Pinus* occurring in medium annual rainfall regions

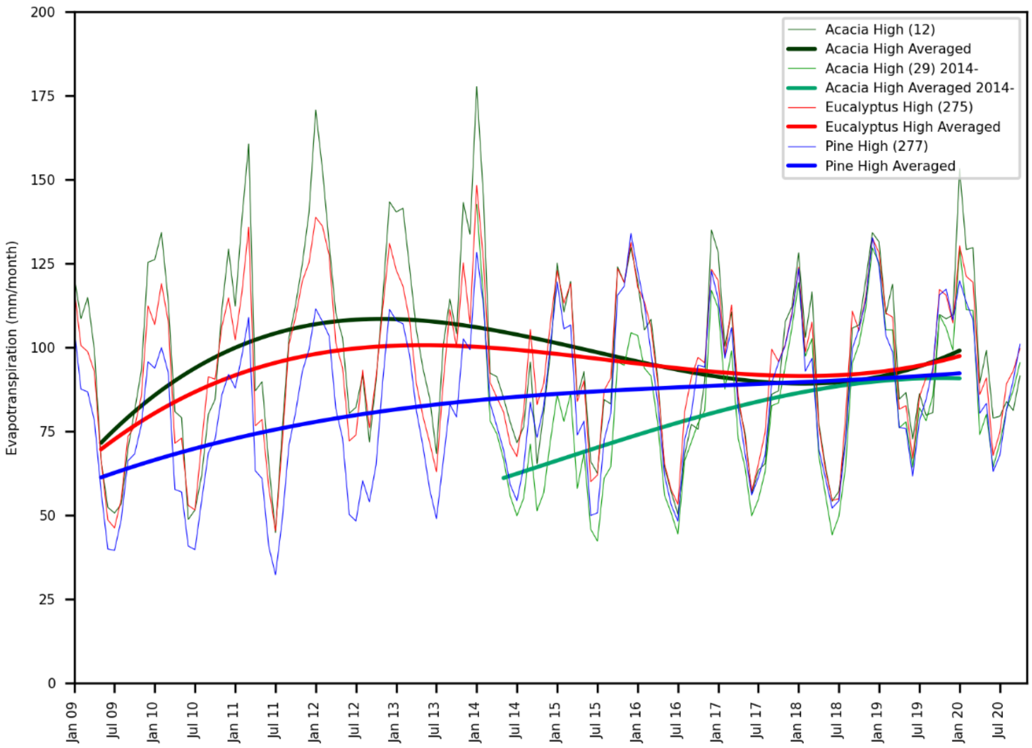


Figure 5-11 Water use over time (mm/month) for *Acacia* (planted in 2009 and 2014), *Eucalyptus* and *Pinus* occurring in high annual rainfall regions

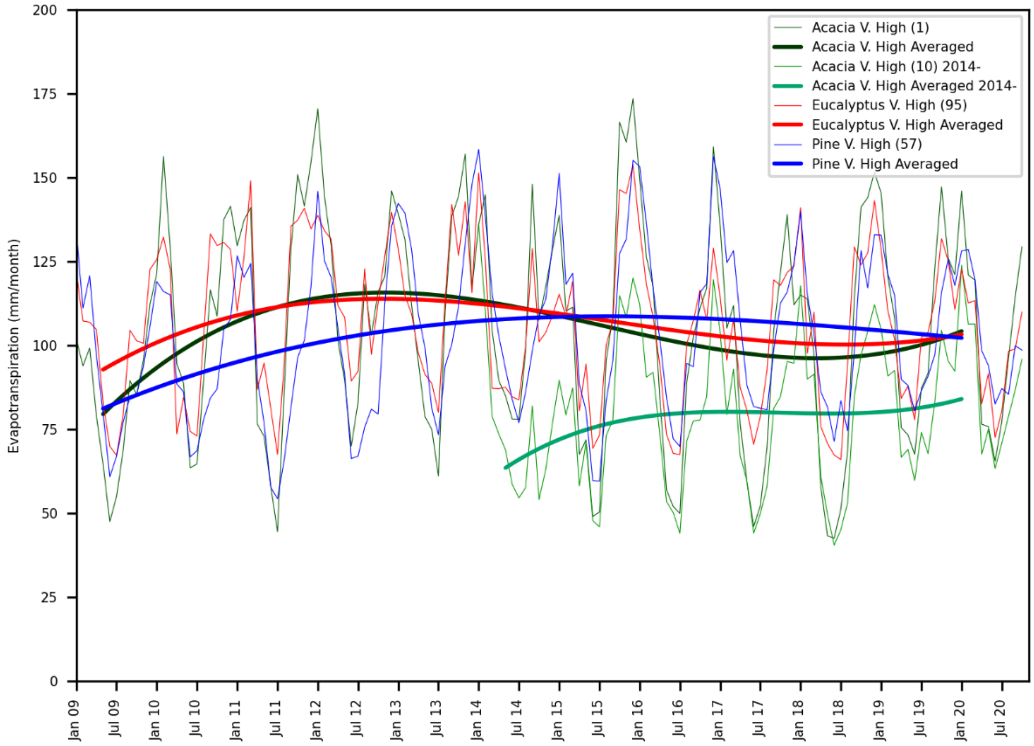


Figure 5-12 Water use over time (mm/month) for *Acacia* (planted in 2009 and 2014), *Eucalyptus* and *Pinus* occurring in very high annual rainfall regions

5.4.2 Rainfall seasonality

The climate of South Africa is highly diverse, ranging from subtropical conditions in the north-eastern parts to a Mediterranean climate in south-western parts of the country. These variations can impact the water use of forest plantations. The first experiment to investigate the impact of climate variation on forest plantation water use involved comparing the water use (ET) of *Acacia*, *Eucalyptus* and *Pinus* compartments to rainfall seasonality.

Figure 5-13 maps the rainfall seasonality in terms of winter, all year and summer (early, mid, late and very late) summer rainfall. When this map is compared to Figure 3-2, one can see that most of the commercial forestry compartments occur in the summer rainfall regions and that the winter and all year regions mainly contain *Pinus* plantations.

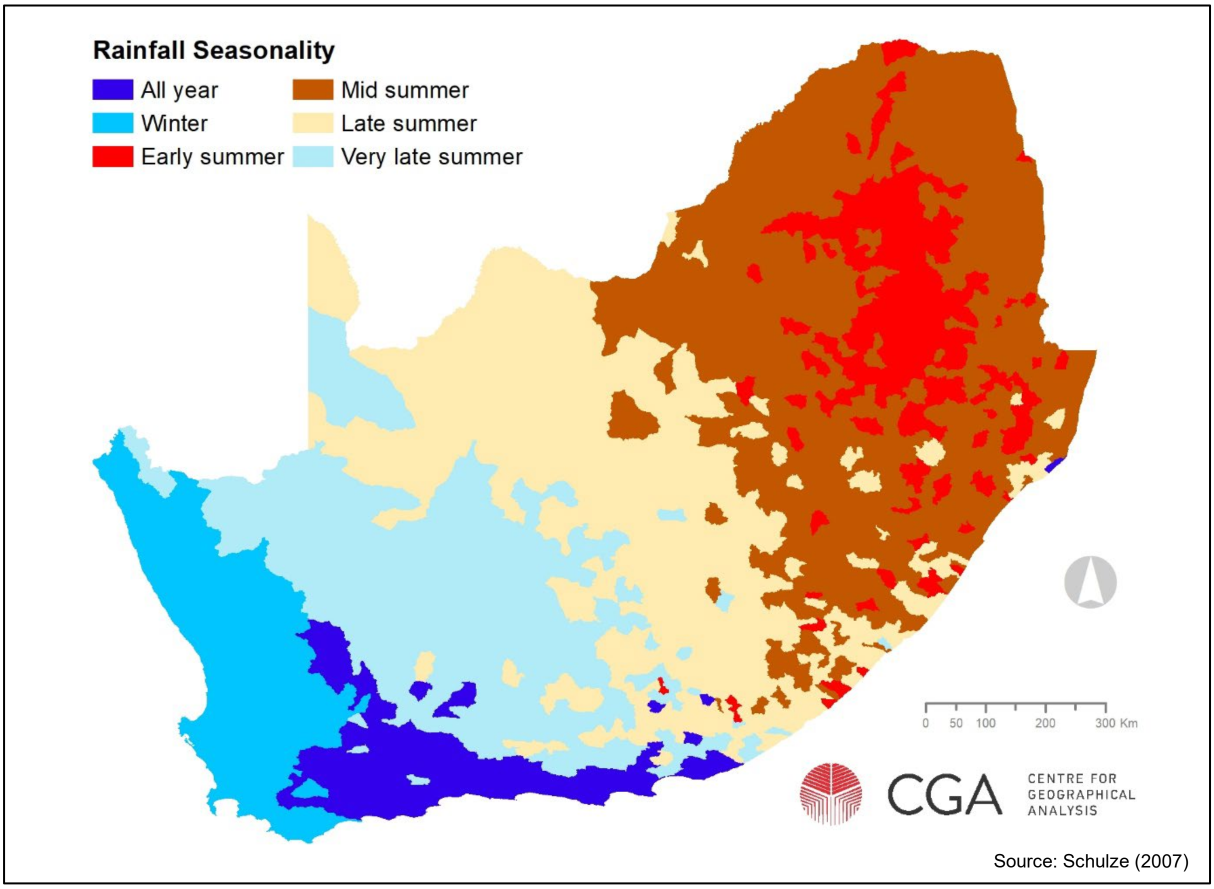


Figure 5-13 Rainfall seasonality regions

Figure 3-14 shows that, in the early summer rainfall regions, *Eucalyptus* compartments use on average more water than *Pinus* compartments. The differences in water use are larger for the first eight or so years, which seem to be linked to the abrupt increases in water use during the early growing stages of *Eucalyptus* compartments, in comparison with a more gradual increase in water use for *Pinus* in this rainfall region.

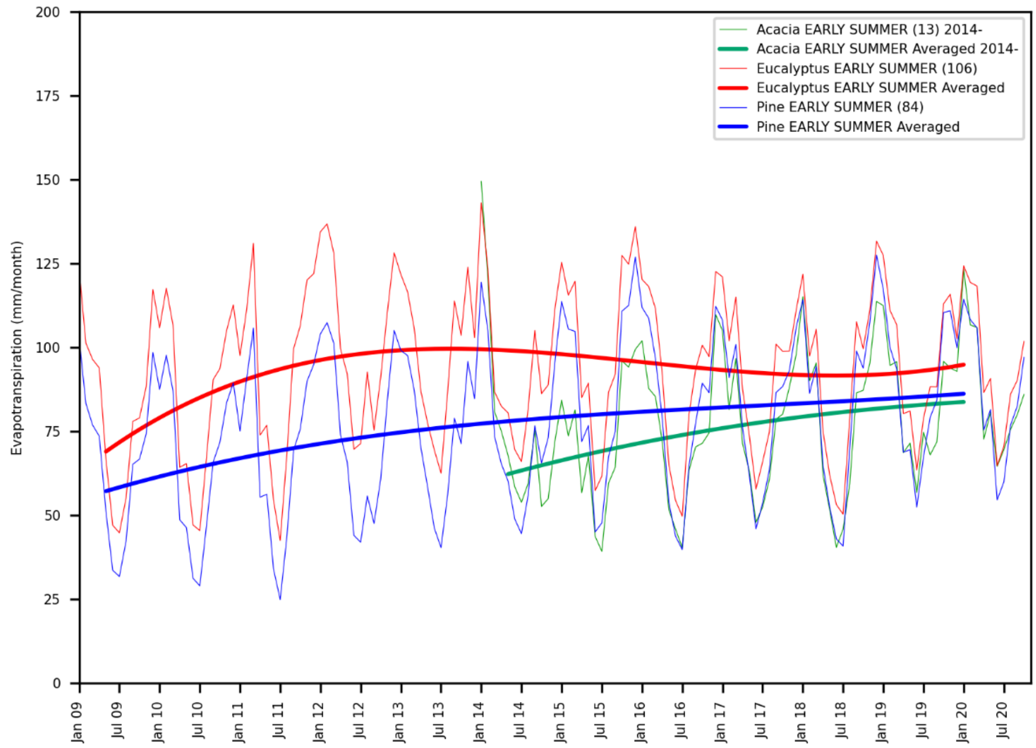


Figure 5-14 Water use over time (mm/month) for *Acacia* (planted in 2009 and 2014), *Eucalyptus* and *Pinus* occurring in the early summer rainfall region

In contrast to the early summer rainfall region, water use of *Eucalyptus* and *Pinus* compartments is similar in the late summer rainfall regions (Figure 5-15). However, the number of cases within this region was generally low (33 compartments in total), which might not be representative.

The water use of compartments in the mid-summer region seems to mimic those of the early summer region (compare Figure 5-16), with *Eucalyptus* compartments using significantly more water than *Pinus* in the first seven years or so after planting, after which the difference in water use of the two genera appear insignificant. Although it is difficult to make a direct comparison between *Acacia* compartments and the other two genera, it seems that the water use of *Acacia* is initially higher than *Pinus* compartments (but substantially lower than *Eucalyptus* compartments), but then it flattens out quite quickly to about 90 mm/month.

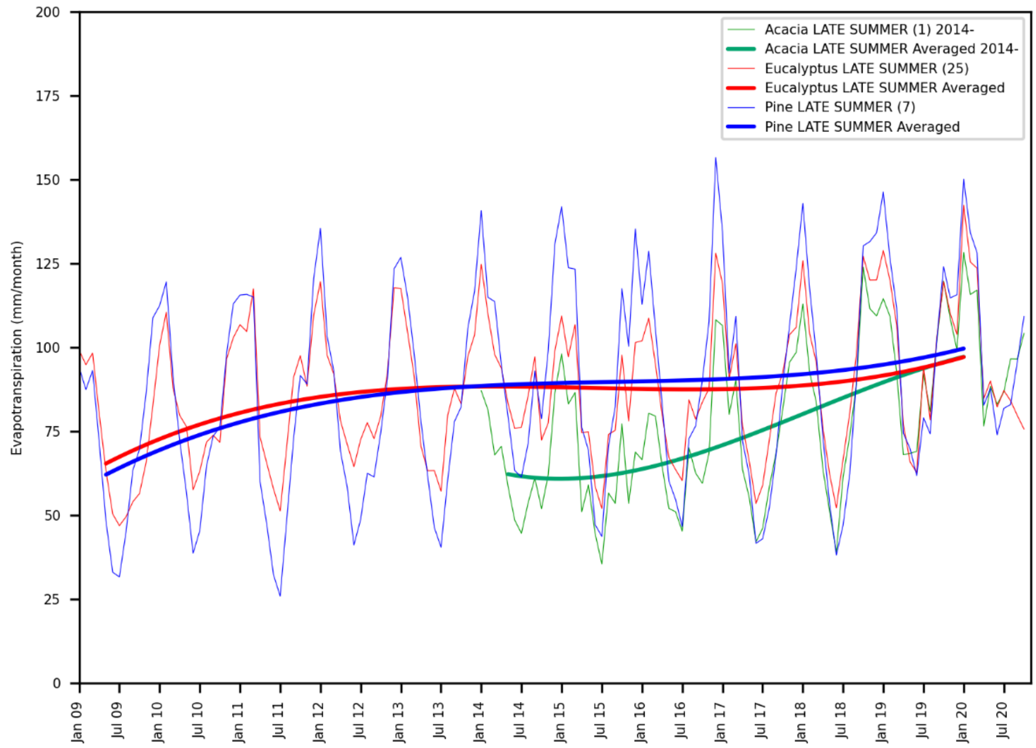


Figure 5-15 Water use over time (mm/month) for *Acacia* (planted in 2009 and 2014), *Eucalyptus* and *Pinus* occurring in the late summer rainfall region

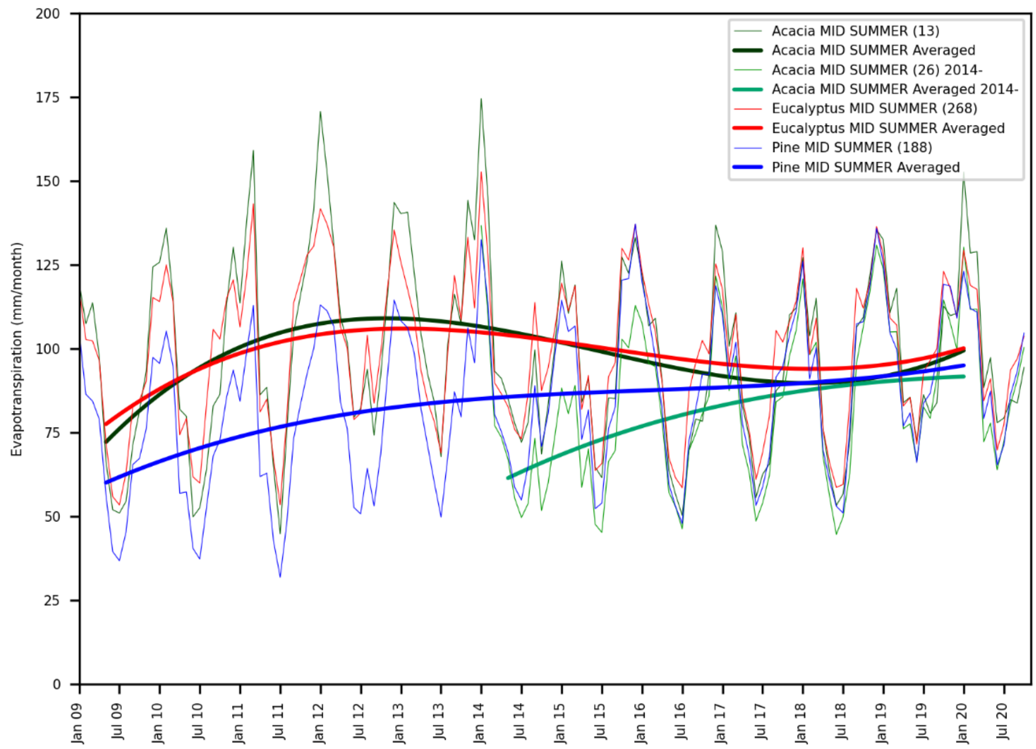


Figure 5-16 Water user over time (mm/month) for *Acacia* (planted in 2009 and 2014), *Eucalyptus* and *Pinus* occurring in the mid-summer rainfall region

As indicated earlier, the all year rainfall region is planted with *Pinus* compartments only and the difference in water use between newly planted *Acacia* compartments and those that have matured is less dramatic compared to the other rainfall regions (Figure 5-17).

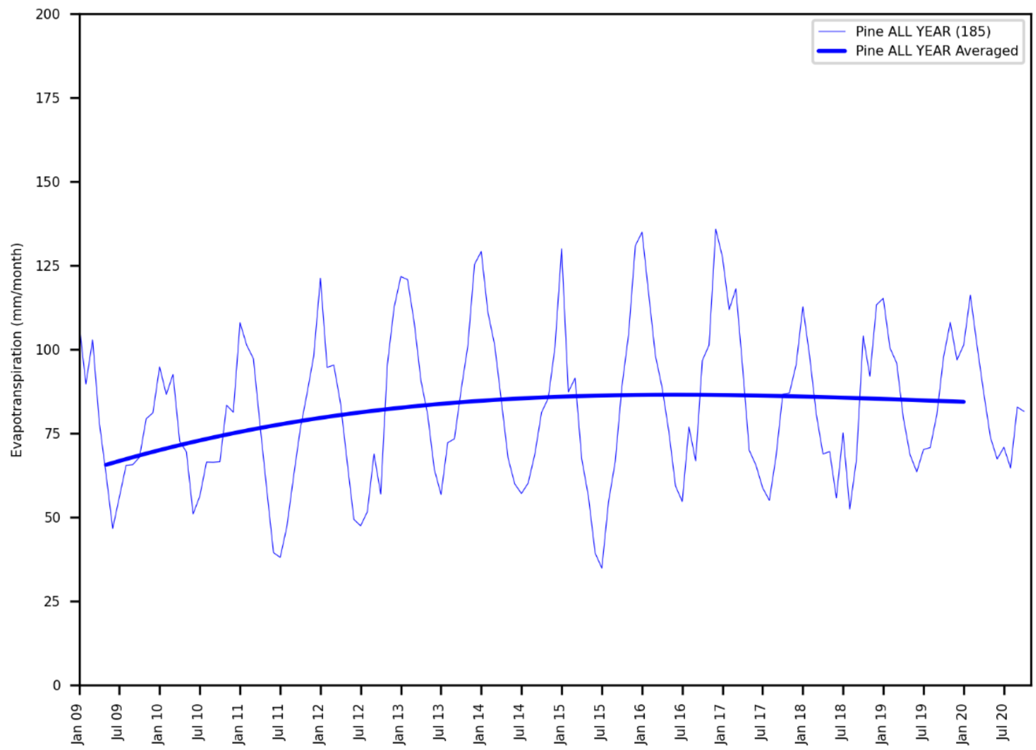


Figure 5-17 Water use over time (mm/month) for Pinus species in the all year rainfall region

Table 5-3 Median water use values (mm/year) per rainfall region and genus

Rainfall region	**Evapotranspiration (mm/year)**		
	Acacia	***Eucalyptus***	***Pinus***
All year	-	-	909 mm/year (n = 185)
Winter	-	-	-
Early summer	-	1109 mm/year (n = 106)	875 mm/year (n = 84)
Mid summer	1149 mm/year (n = 13)	1194 mm/year (n = 268)	1015 mm/year (n = 188)
Late summer	-	1013 mm/year (n = 25)	1021 mm/year (n = 7)

5.4.3 Climate zones

From the previous section, it seems that rainfall seasonality has an impact on genus water use, but that the differences are not dramatic. This section takes a closer look at how water use varies with climate. Specifically, the Köppen climatic zones (Figure 5-18) are used to disaggregate water use (ET) per genus. Köppen zones divide climates into five main climate groups, namely A (tropical), B (dry), C (temperate), D (continental), and E (polar), with each group being further divided based on seasonal precipitation and temperature patterns, which are indicated by the lowercase second and third characters in the notation respectively. Of these, only Aw (Tropical savanna climate with dry-winter characteristics), Cfa (Humid subtropical climates), Cfb (Oceanic climate), Cwa (Dry-winter humid subtropical climate) and Cwb (Dry-winter subtropical highland climate) intersect with the compartments in the geodatabase.

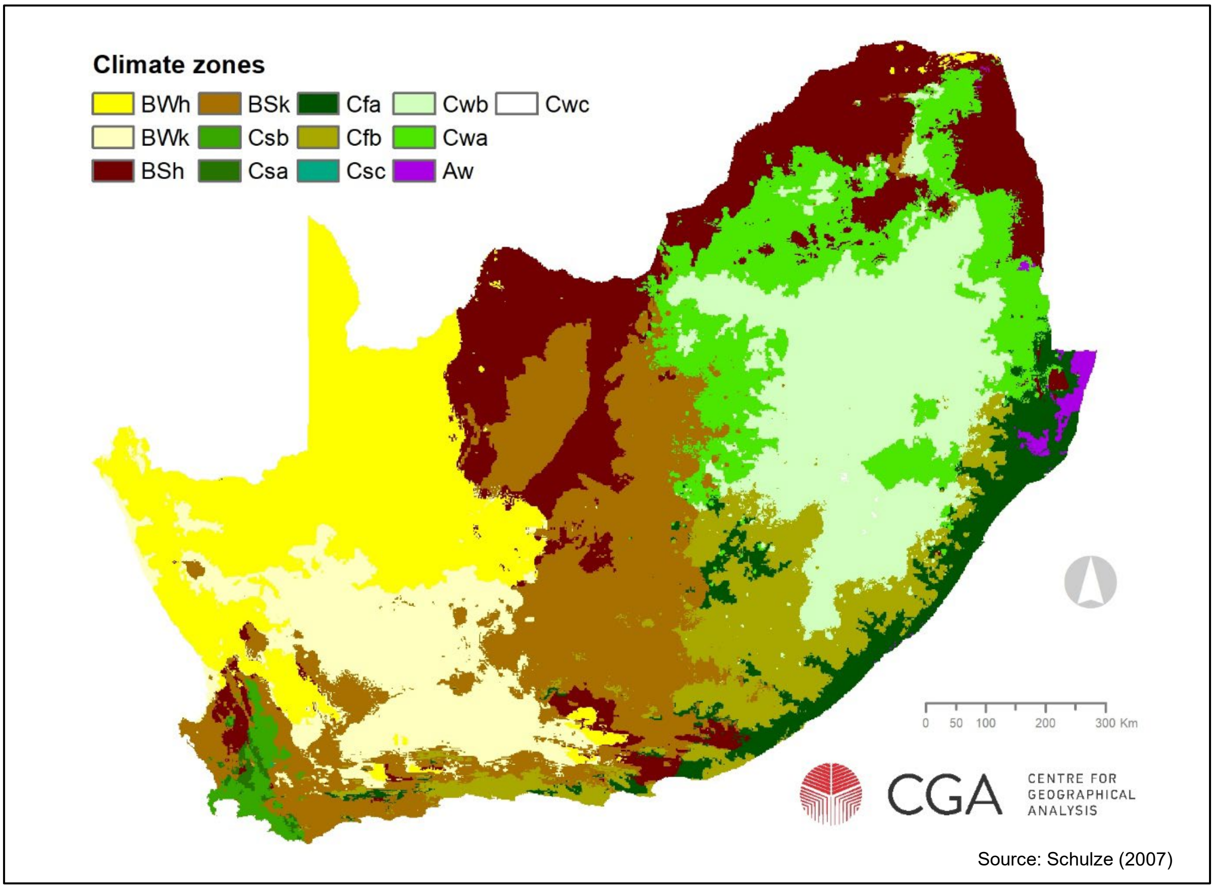


Figure 5-18 Köppen climate zones of South Africa

Figure 5-19 shows the water use of compartments within the different Köppen climate zones. The number of compartments in the Aw climate zone (Figure 5-19) are limited (5) and the water use quantifications are likely not reliable. However, the relatively low water use of the *Eucalyptus* compartments from 2014 to 2017 is noteworthy and is attributed to the drought conditions that were experienced during this period[12]. A similar albeit less dramatic trend for *Eucalyptus* compartments is noted in the neighbouring Cfa climate zone (Figure 5-20). In contrast, the effect of the drought is not noticeable for *Pinus* compartments. In fact, the water use of *Pinus* compartments peaked (at about 165 mm/month) during this period. Although *Eucalyptus* compartments used more water than those planted with *Pinus* species in the first five years after planting, the difference is not as stark as in the other climatic zones (see later). The water use of mature *Pinus* compartments exceeded those planted with *Eucalyptus* trees, but this is likely due to the drought conditions.

[12] According to Ndlovo and Demlie (2020), 2015 was the driest year on record (1970-2017) for KwaZulu-Natal.

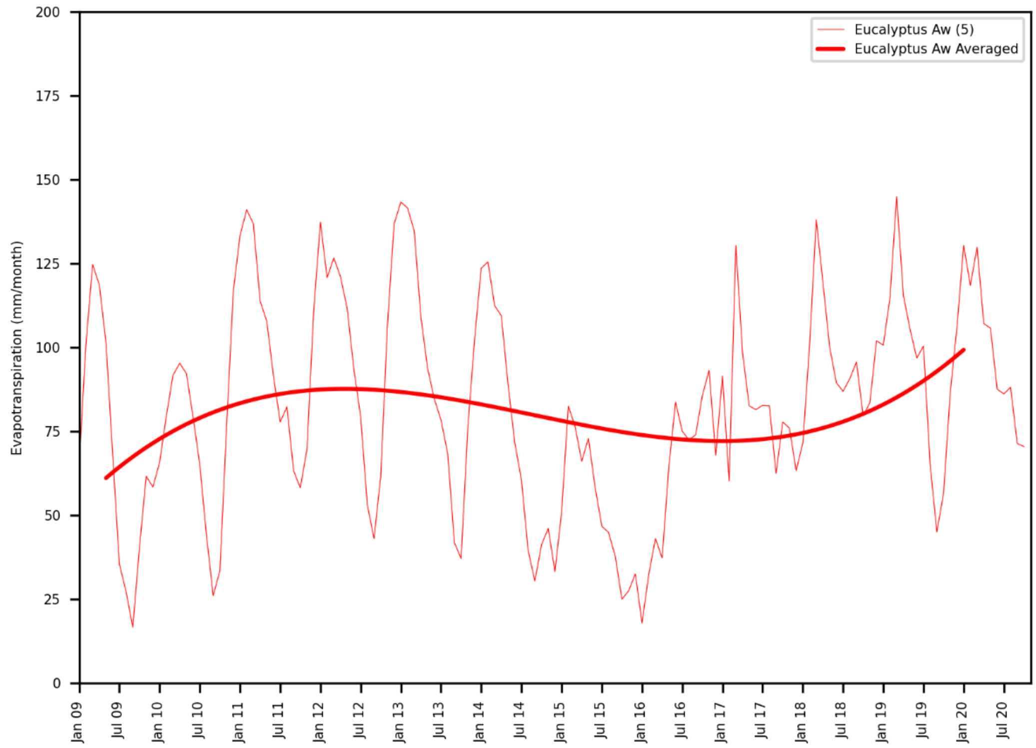


Figure 5-19 Water use over time (mm/month) for *Eucalyptus* occurring in the Aw (Tropical savanna climate with dry-winter characteristics) climatic zone

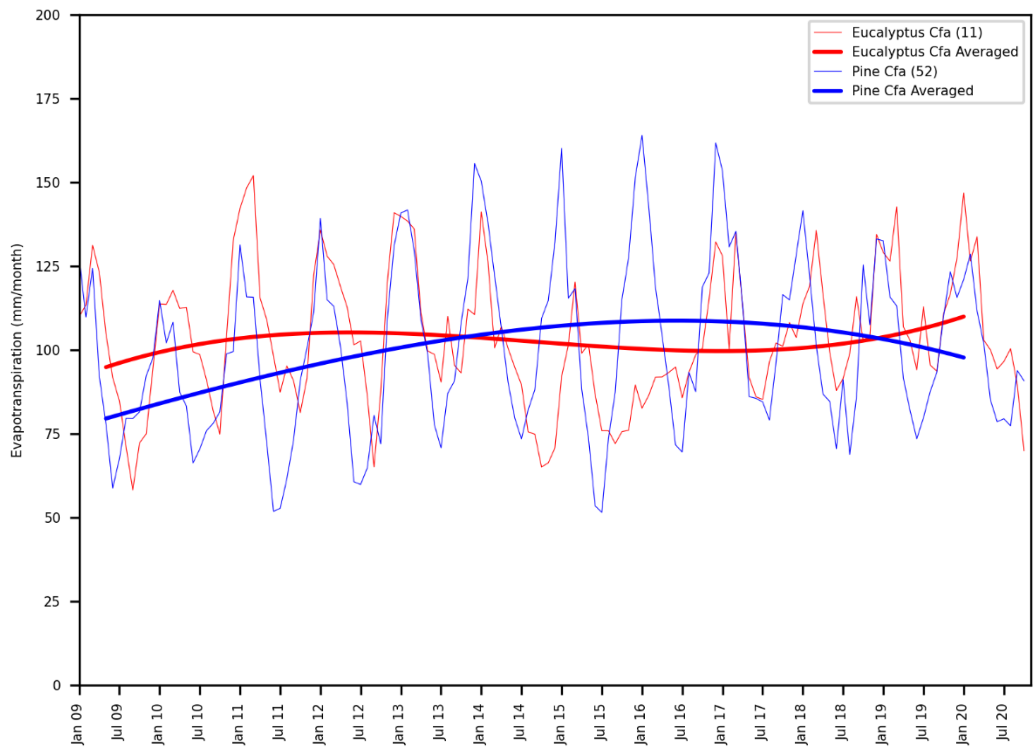


Figure 5-20 Water use over time (mm/month) for *Eucalyptus* and *Pinus* occurring in the Cfa (Humid subtropical climates) climatic zone

In contrast to the Cfa zone, *Eucalyptus* compartments consistently used more water than *Pinus* compartments in the Cfb climatic zone (Figure 5-21). The difference is more pronounced during the first seven years of planting, after which the water use of the two genera seem to converge. The number of *Acacia* compartments within this region is limited, with only nine planted in 2009, but based on the available data, it seems that the water use of *Acacia* compartments is dramatically higher than those planted with the other two genera, especially during the first four years. As in Cfa, it seems that *Pinus* compartments in Cfb were less affected by the drought conditions (2014-2017), while water use of *Acacia* and *Eucalyptus* plantations were significantly lower during this time.

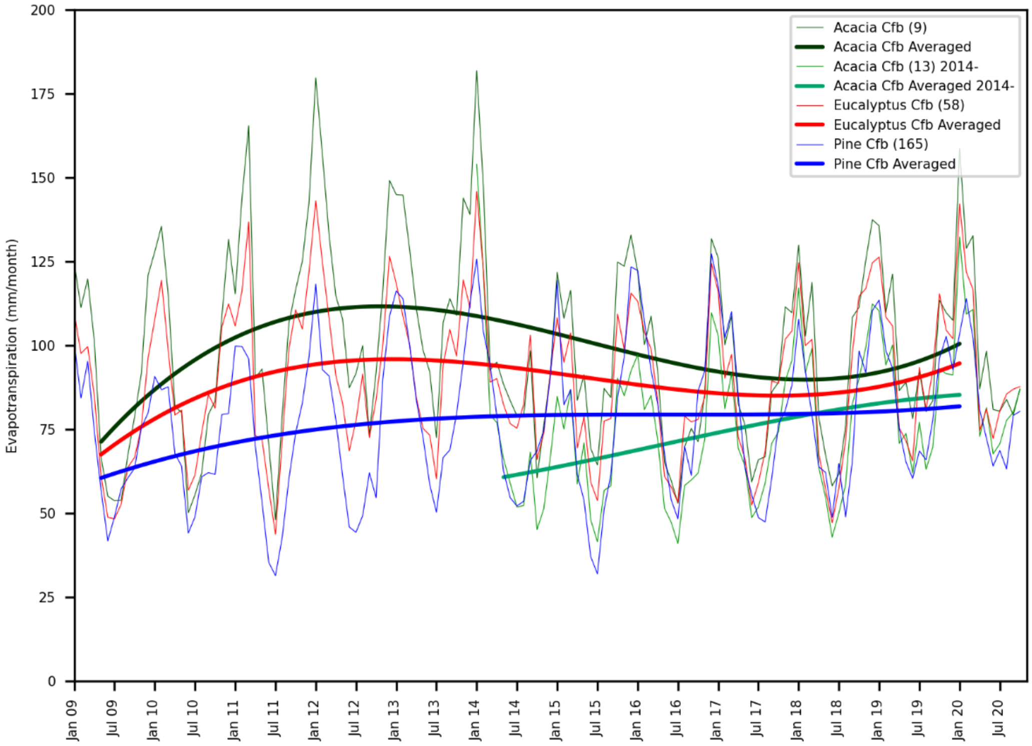


Figure 5-21 Water use over time (mm/month) for *Acacia* (planted in 2009 and 2014), *Eucalyptus* and *Pinus* occurring in the Cfb (Oceanic climate) climatic zone

The patterns of water use by *Pinus* and *Eucalyptus* compartments in the Cwa zone (Figure 5-22) are similar to those in Cfb, although the differences in water use by these genera are larger. Unfortunately, no conclusions can be drawn from the water use of *Acacia* compartments in this region, as only one compartment was planted in 2009.

Eucalyptus compartments in the Cwb zone used substantially more water than *Pinus* compartments from planting up to about seven years, after which the differences in water use become negligible (Figure 5-23). Based on the limited number (3) of *Acacia* compartments in this zone, it seems that the water use of *Acacia* and *Eucalyptus* compartments are on par.

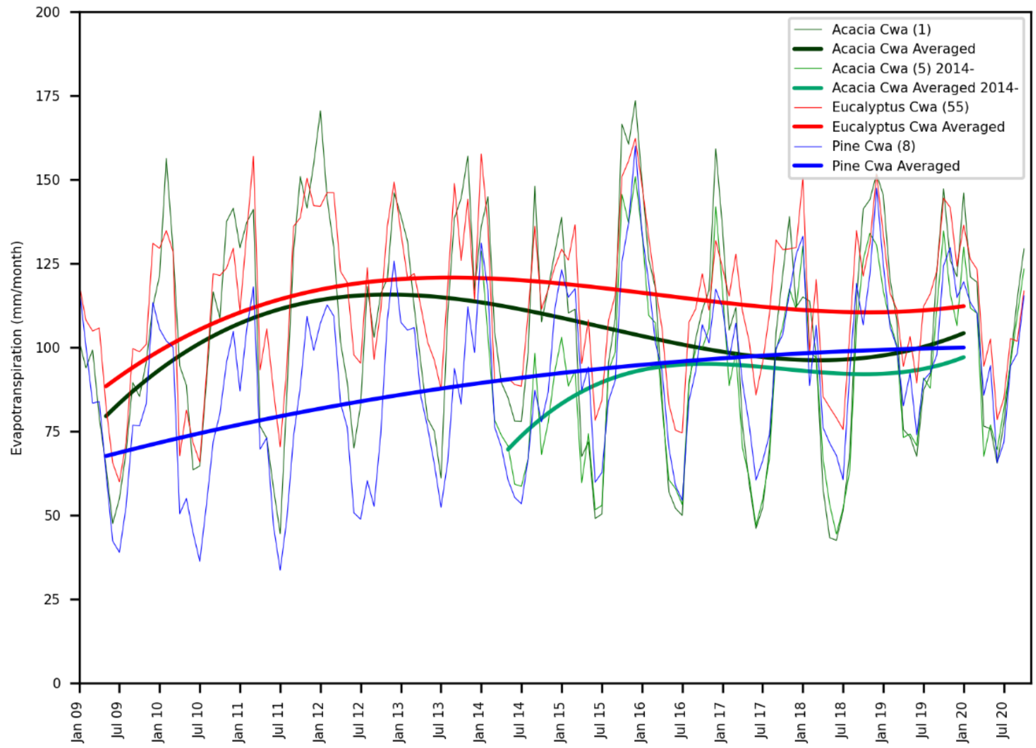


Figure 5-22 Water use over time (mm/month) for *Acacia* (planted in 2009 and 2014), *Eucalyptus* and *Pinus* occurring in the Cwa (Dry-winter humid subtropical climate) climatic zone

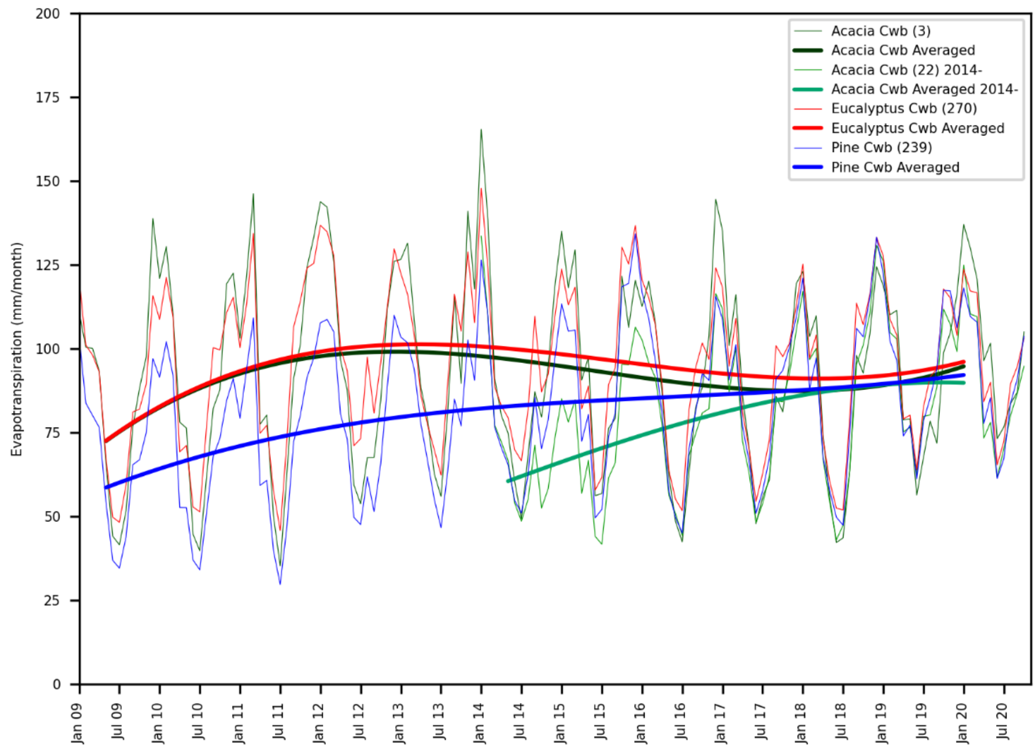


Figure 5-23 Water use over time (mm/month) for *Acacia* (planted in 2009 and 2014), *Eucalyptus* and *Pinus* occurring in the Cwb (Dry-winter subtropical highland climate) climatic zone

Based on the results of this section, it is clear that *Eucalyptus* compartments use significantly more water than *Pinus* compartments for about seven years from planting. The difference in water use by these two genera narrows from seven years onwards, especially in the Cwb (Dry-winter subtropical highland climate) zone, which is characterized by lower rainfall. The narrowing of water use by *Eucalyptus* and *Pinus* seems to be linked to a drastic reduction of water use by *Eucalyptus* compartments seven years after planting. The increase in water use by *Pinus* during the first seven years of planting is gentle, as is the slight decrease from seven years onwards. In contrast, *Eucalyptus* compartments exhibit a drastic increase in water use in the initial period of planting followed by a substantial decrease in water use from seven years after planting onwards.

Table 5-4 Median water use values (mm/year) per climate zone and genus

Climate zone	Evapotranspiration (mm/year)		
	Acacia	***Eucalyptus***	***Pinus***
Aw (Tropical savanna climate with dry-winter characteristics)	-	992 (n = 5)	-
Cfa (Humid subtropical climates)	-	1274 (n = 11)	1285 (n = 52)
Cfb (Oceanic climate)	1147 (n = 9)	1096 (n = 58)	853 (n = 165)
Cwa (Dry-winter humid subtropical climate)	1303 (n = 1)	1417 (n = 55)	1078 (n = 8)
Cwb (Dry-winter subtropical highland climate)	1119 (n = 3)	1135 (n = 270)	952 (n = 239)

5.5 Water use compared to terrain characteristics

5.5.1 Slope gradient

The previous section showed that there are substantial ET variations among plantations, even for those planted with the same genus. This section explores the impact that slope gradient has on compartment water use. Figure 5-24 is a slope gradient map, classified into level/gently inclined (0-10%), moderately inclined/steep (10-56%) and very steep (>56%) slopes.

Figure 5-25 reveals that there is relatively little difference among the water use of genera planted on level and gently inclined slopes. *Eucalyptus* compartments use more water initially (first three years after planting), but from six years onwards, *Pinus* plantations use more water. Thereafter they appear to converge. Due to the relatively small sample of *Acacia* (11 compartments), it is difficult to draw any conclusions about its water use on level and gently inclined slopes.

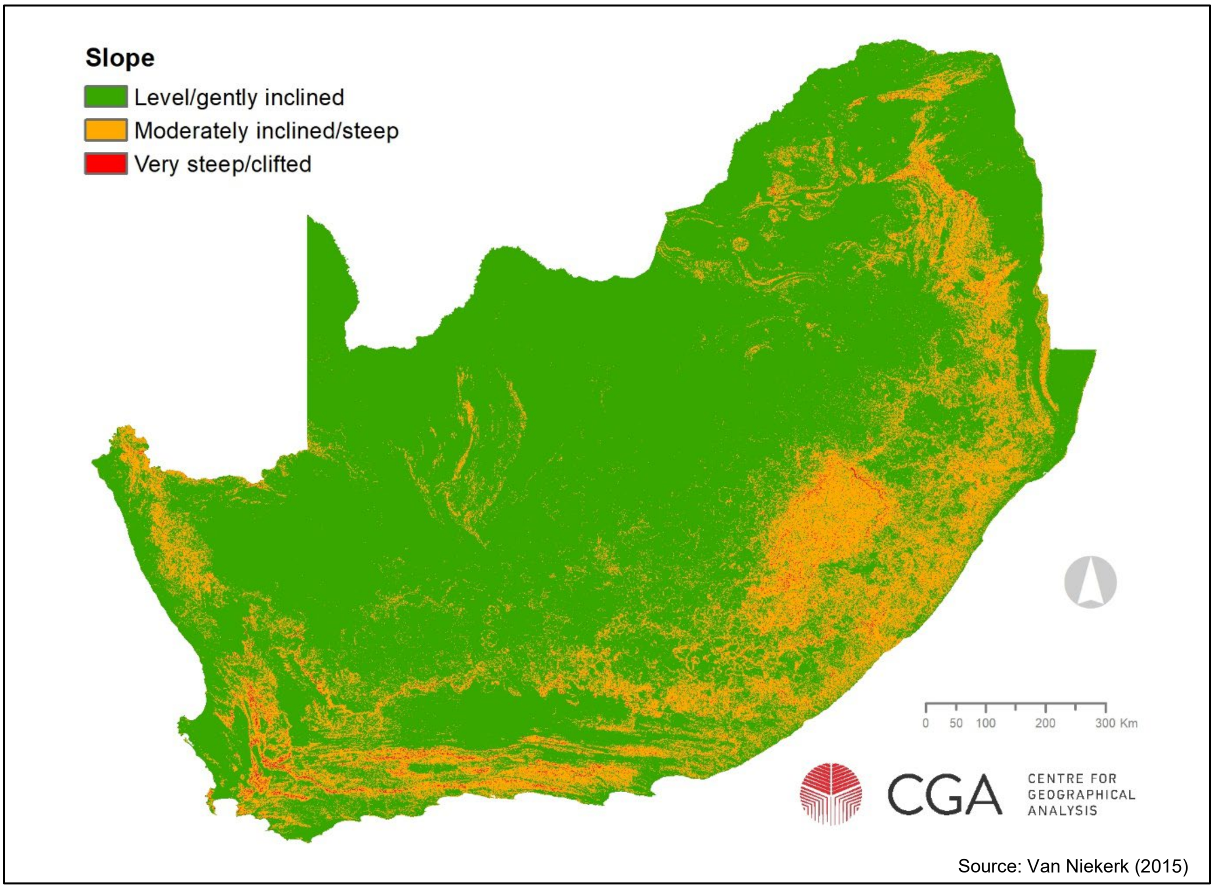


Figure 5-24 Slope gradient map of South Africa

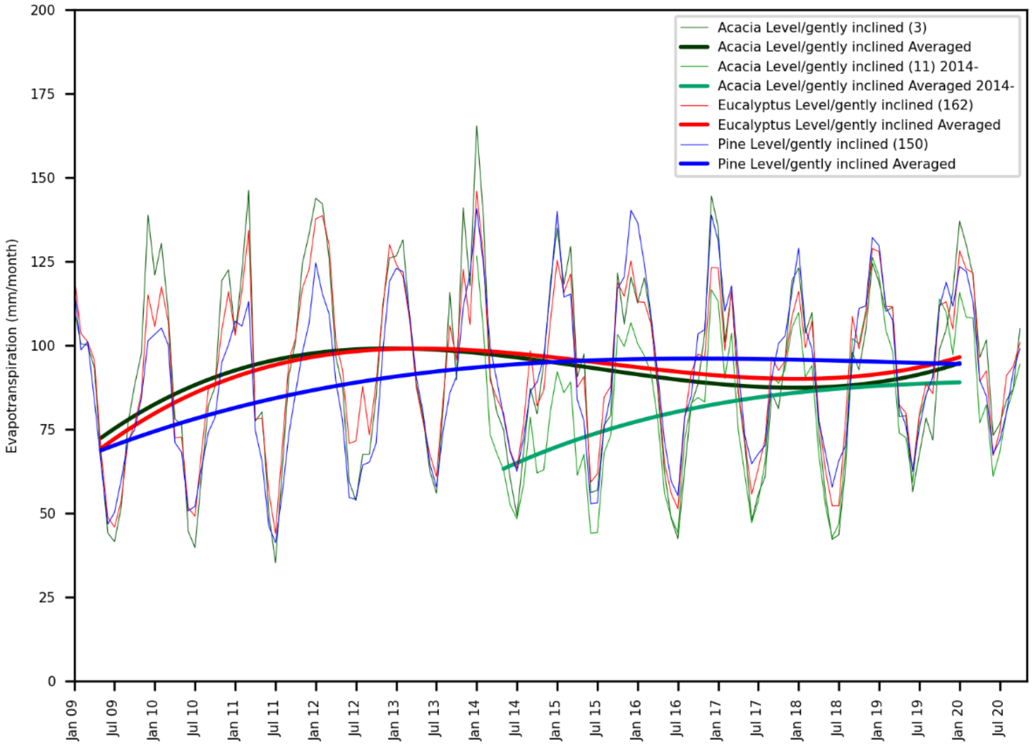


Figure 5-25 Water use over time (mm/month) for *Acacia* (planted in 2009 and 2014), *Eucalyptus* and *Pinus* on level/gently inclined slopes

In contrast to Figure 5-25, Figure 5-26 shows that there are substantial differences in the water use among genera on moderately inclined and steep slopes (Table 5-5). There is a marked difference between the ET of *Eucalyptus* and *Pinus* compartments, with the former having 10% higher values (on average) over the period being considered. Overall, *Eucalyptus* consistently uses more water than *Pinus* on moderately inclined and steep slopes.

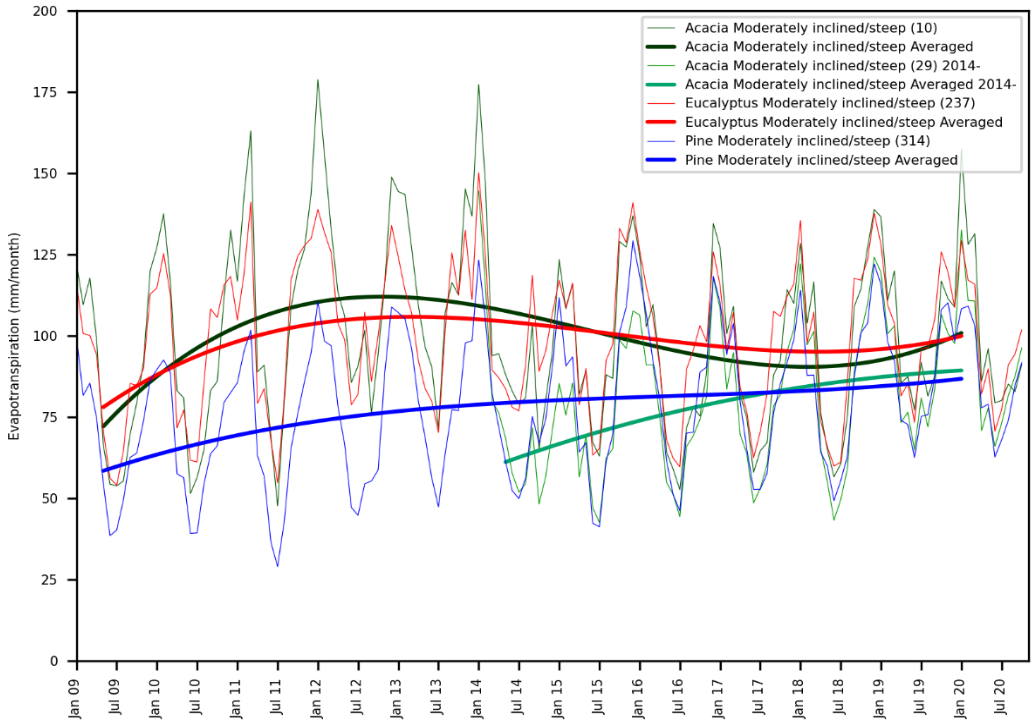


Figure 5-26 Water use over time (mm/month) for *Acacia* (planted in 2009 and 2014), *Eucalyptus* and *Pinus* on moderately inclined/steep slopes

A comparison between water use of compartments on very steep slopes was not carried out as no such cases exist in the compartment database.

Table 5-5 Median ET values (mm/year) per slope gradient category

Slope category	Evapotranspiration (mm/year)		
	Acacia	***Eucalyptus***	***Pinus***
Level/gently inclined (0-10%)	1119 (n = 3)	1109 (n = 162)	1138 (n = 150)
Moderately inclined/steep (10-56%)	1190 (n = 10)	1208 (n = 237)	952 (n = 314)
Very steep slopes (>56%)	-	-	-

As slope increases, the water use of *Pinus* decreases, while the water use of *Eucalyptus* increases. Steeper slopes could result in increased runoff and subsequently lower soil water content, leading to lower ET. Conversely, steeper slopes (especially in high lying areas) could be associated with mist and hence rainfall/mist interception. Evaporation of this intercepted water can subsequently result in higher ET. An increased steepness, elevation and the likelihood of more mist will also result in less solar radiation to drive ET. However, these are theories and require verification.

5.5.2 Slope aspect

From the previous section it is clear that slope gradient has a significant effect on water use among the primary genera planted in South Africa. Figure 5-27 shows slope aspect (i.e. direction of slope) at national scale, as generated from the Stellenbosch University Digital Elevation Model (SUDEM). Aspect ranges from 0 to 360 degrees (measured from north in a clockwise direction). For the sake of this analysis, aspect values were reclassified into four categories, namely north, east, south and west.

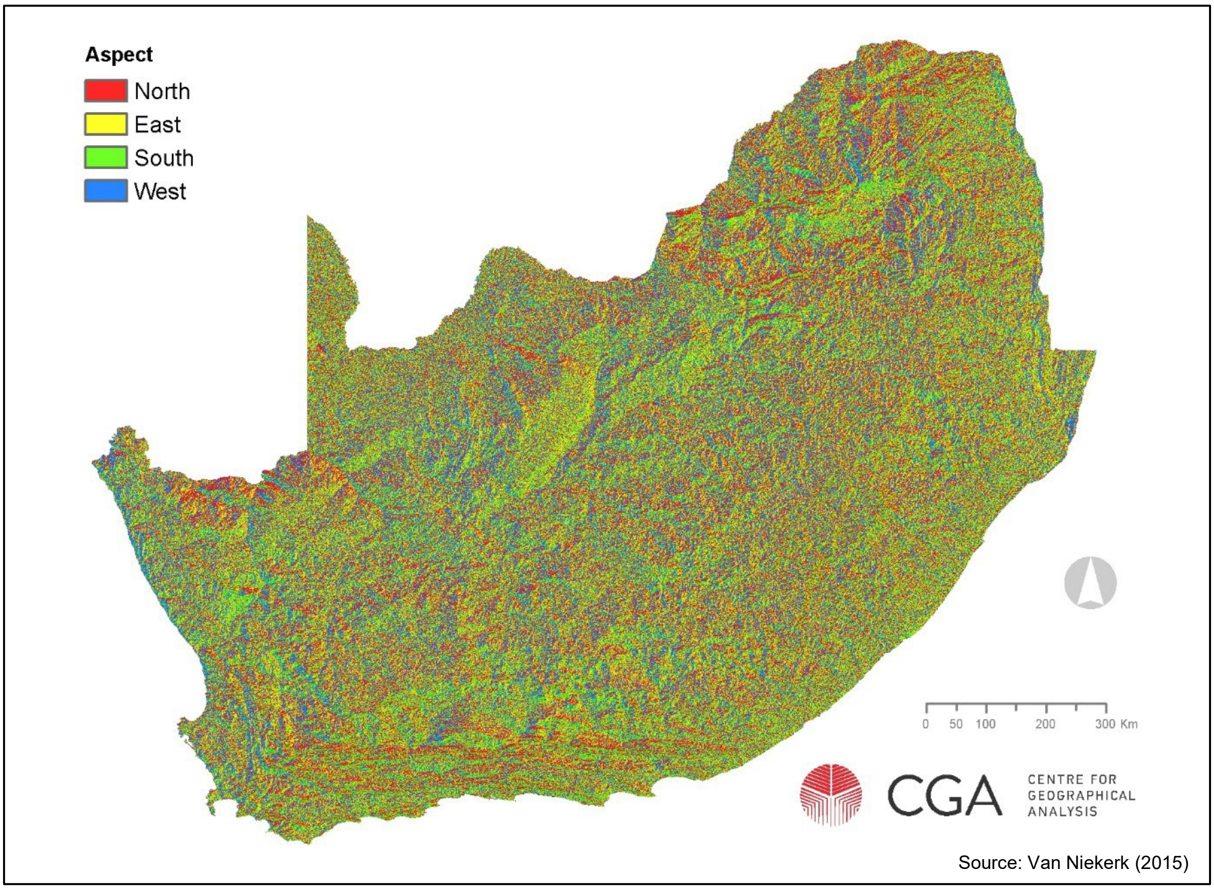


Figure 5-27 Slope aspect for South Africa

The water use of the genera per slope aspect category are shown in Figure 5-28 to Figure 5-31. Water use of *Eucalyptus* compartments is generally higher on northern slopes compared to southern slopes, which can be expected given compartments with northern aspects would receive more solar radiation. ET from southern slopes is lower than that from northern slopes due to increased periods of shade and hence lower solar radiation. Interestingly, the impact of slope was less on *Pinus* species. The water use profiles of *Eucalyptus* and *Pinus* compartments are relatively consistent among slope directions, while those for *Acacia* are erratic. The large differences among the *Acacia* water use profiles are attributed to the relatively few cases.

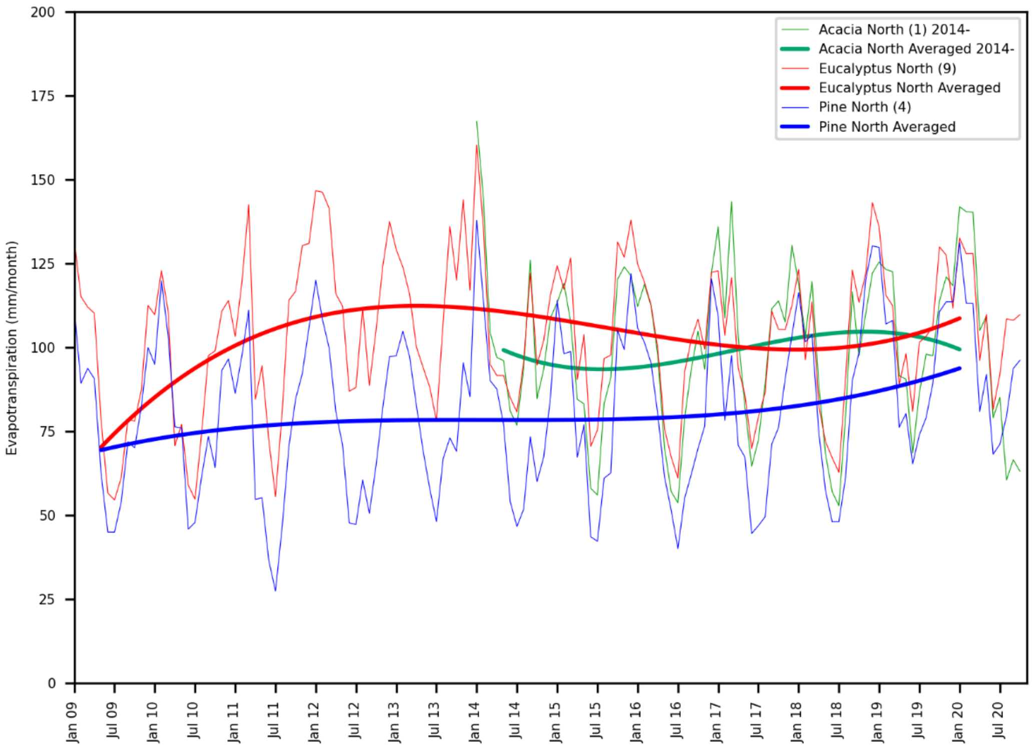


Figure 5-28 Water use over time (mm/month) for *Acacia* (planted in 2009 and 2014), *Eucalyptus* and *Pinus* occurring on North facing slopes

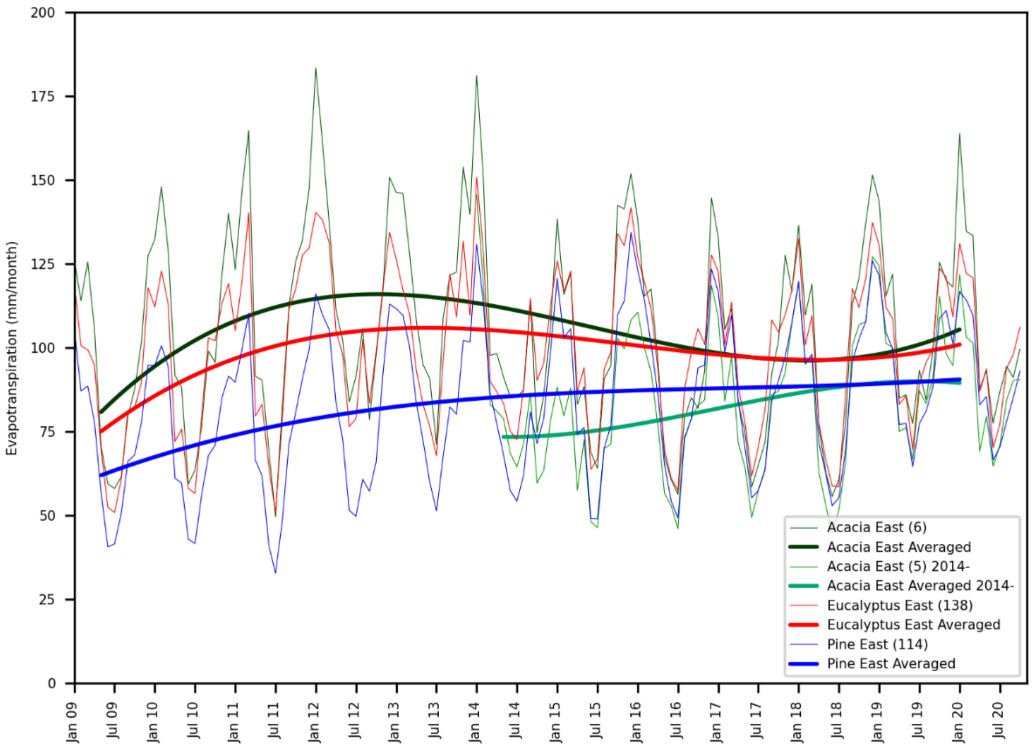


Figure 5-29 Water use over time (mm/month) for *Acacia* (planted in 2009 and 2014), *Eucalyptus* and *Pinus* occurring on East facing slopes

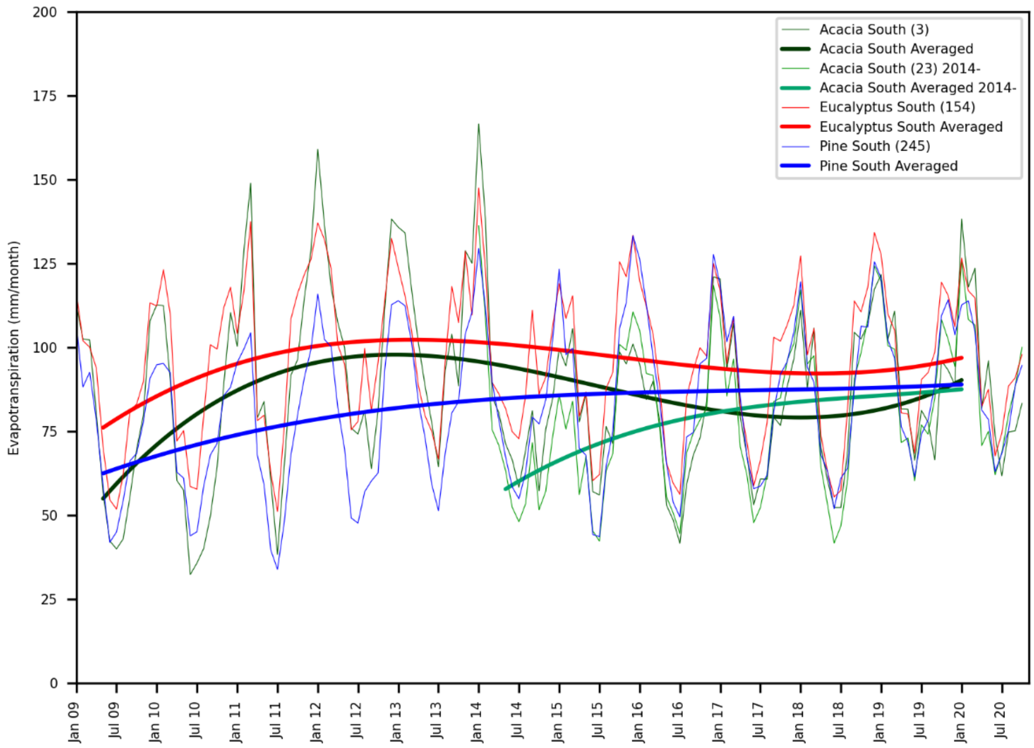


Figure 5-30 Water use over time (mm/month) for *Acacia* (planted in 2009 and 2014), *Eucalyptus* and *Pinus* occurring on South facing slopes

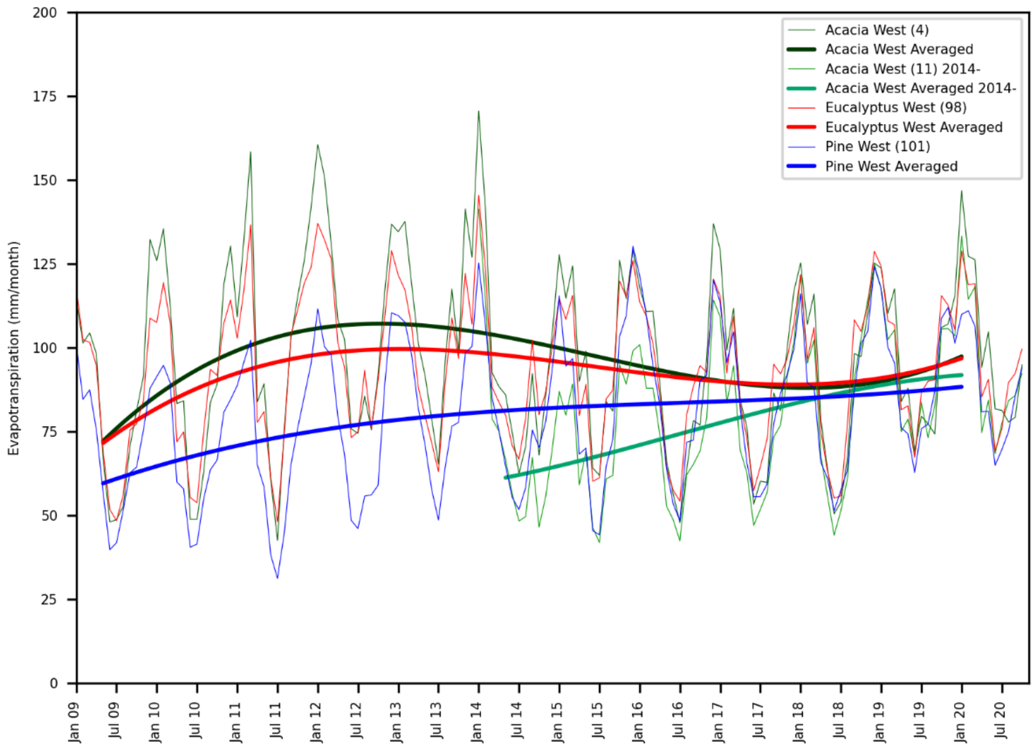


Figure 5-31 Water use over time (mm/month) for *Acacia* (planted in 2009 and 2014), *Eucalyptus* and *Pinus* occurring on West facing slopes

Table 5-6 Median ET values (mm/year) per slope aspect category

Aspect category	Evapotranspiration (mm/year)		
	Acacia	***Eucalyptus***	***Pinus***
North	-	1240 (n = 9)	1006 (n = 4)
East	1254 (n = 6)	1212 (n = 138)	982 (n = 114)
South	990 (n = 3)	1152 (n = 154)	1005 (n = 245)
West	1121 (n = 4)	1128 (n = 98)	959 (n = 101)

5.5.3 Terrain morphology

A set of analyses were carried out to determine whether other terrain characteristics, such as morphology, has an impact on water use. Figure 5-32 shows the 27 morphology units covering South Africa, taken from Schulze (2007). This map was simplified by grouping similar morphology regions, as overviewed in Table 3-5. Only regions that intersect plantation compartments were considered (others were renamed "other" in Figure 5-33).

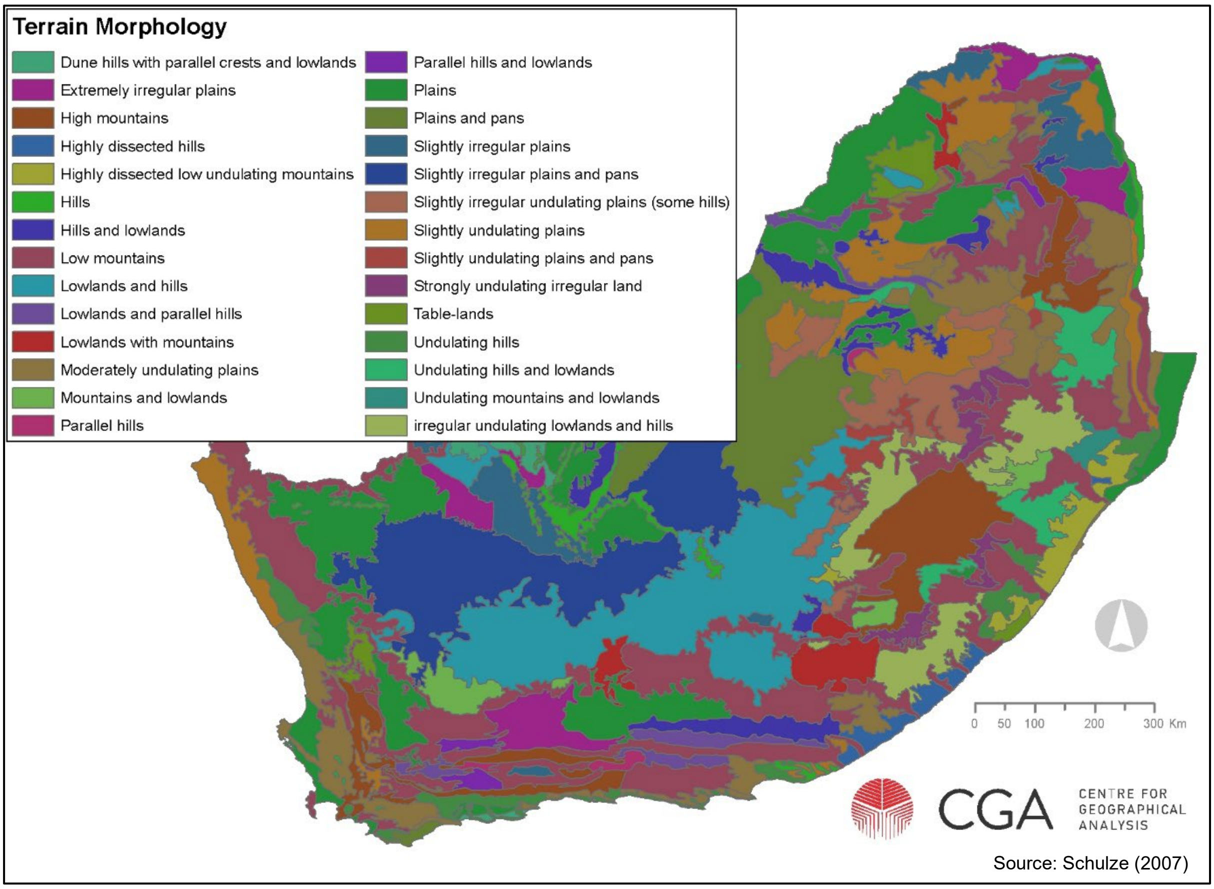


Figure 5-32 Terrain morphology units

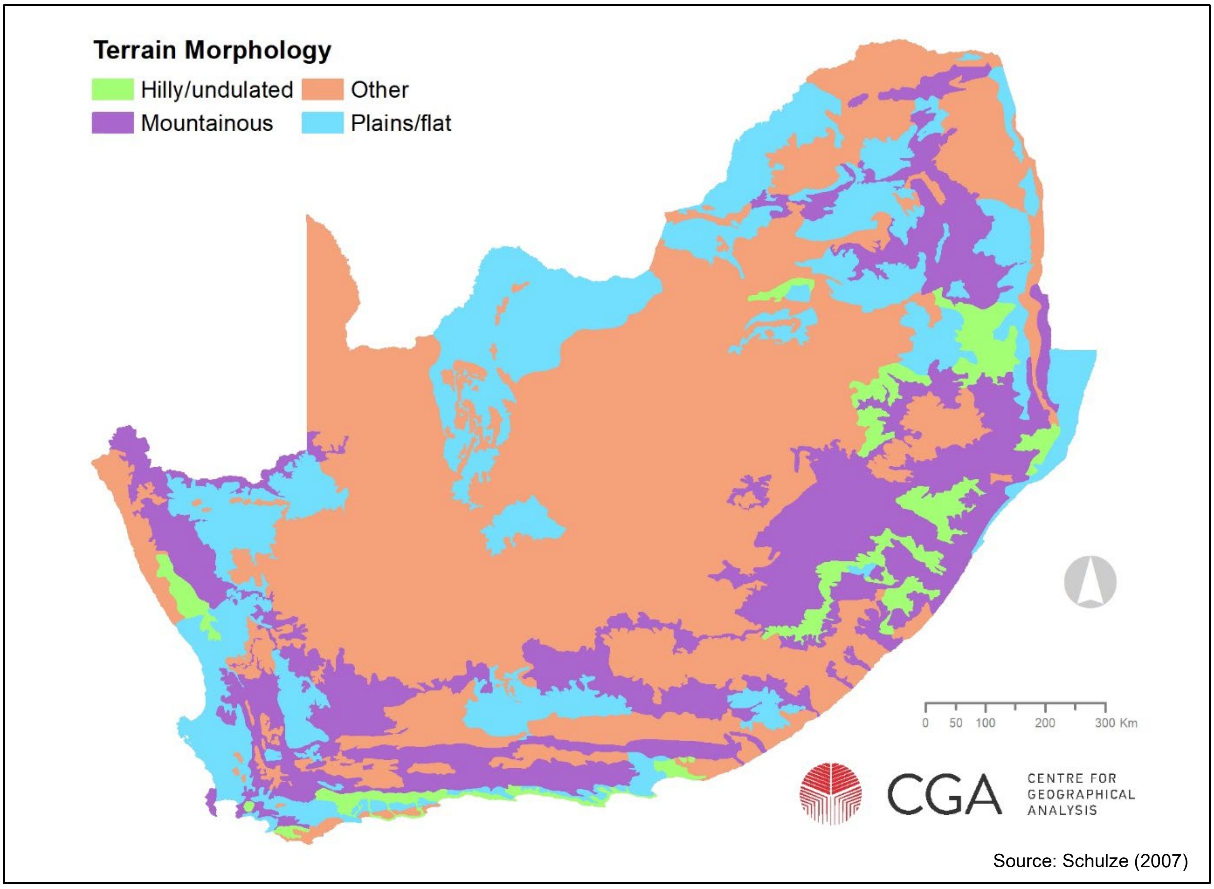


Figure 5-33 Simplified terrain morphology regions

When Figure 5-34, Figure 5-35 and Figure 5-36 are compared, it is evident that terrain morphology has a significant impact on the ET values of compartments. The ET values of *Eucalyptus* compartments planted in mountainous regions are much higher (Figure 5-36) compared to plains/flat terrain (Figure 5-34) and hilly/undulated terrain (Figure 5-35), while the opposite is true for *Acacia* and *Pinus* compartments. For instance, at a mature age of 5 years (Jan 2014), the average ET of *Eucalyptus* compartments in mountainous regions was 149.9 mm/month, compared to 135.5 mm/month for those occurring on plains/flat terrain – a difference of 14.4 mm/month (9.6%). In contrast, the mean ET of *Pinus* (in Jan 2014) was 120.2 mm/month in mountainous areas, while on plains/flat terrain it was 154.5 mm/month at that time; amounting to a 34.3 mm/month (28.6%) difference. The impact of terrain morphology is less for mature (5-year-old) *Acacia* compartments. For instance, in Jan 2019, the average monthly ET for *Acacia* compartments on plains/flat terrain was 133.1 mm/month, while in mountains regions it was 114.5 mm/month (i.e. a 13.9% difference). In general, while *Pinus* uses more water than *Eucalyptus* on plain/flat terrain, *Eucalyptus* uses more water than *Pinus* in mountainous areas. On hilly/undulated areas, *Eucalyptus* uses more water than *Pinus* for the first six years after planting, after which *Pinus* uses slightly more water than *Eucalyptus*.

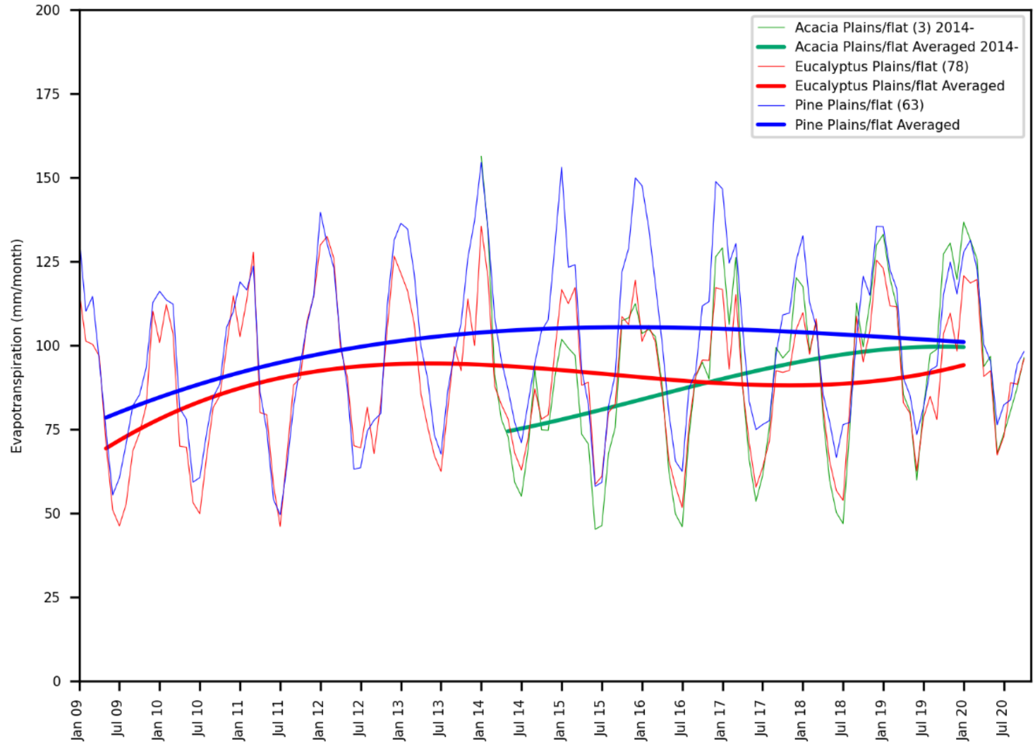


Figure 5-34 Water use over time (mm/month) for *Acacia* (planted in 2009 and 2014), *Eucalyptus* and *Pinus* on plains/flat terrain

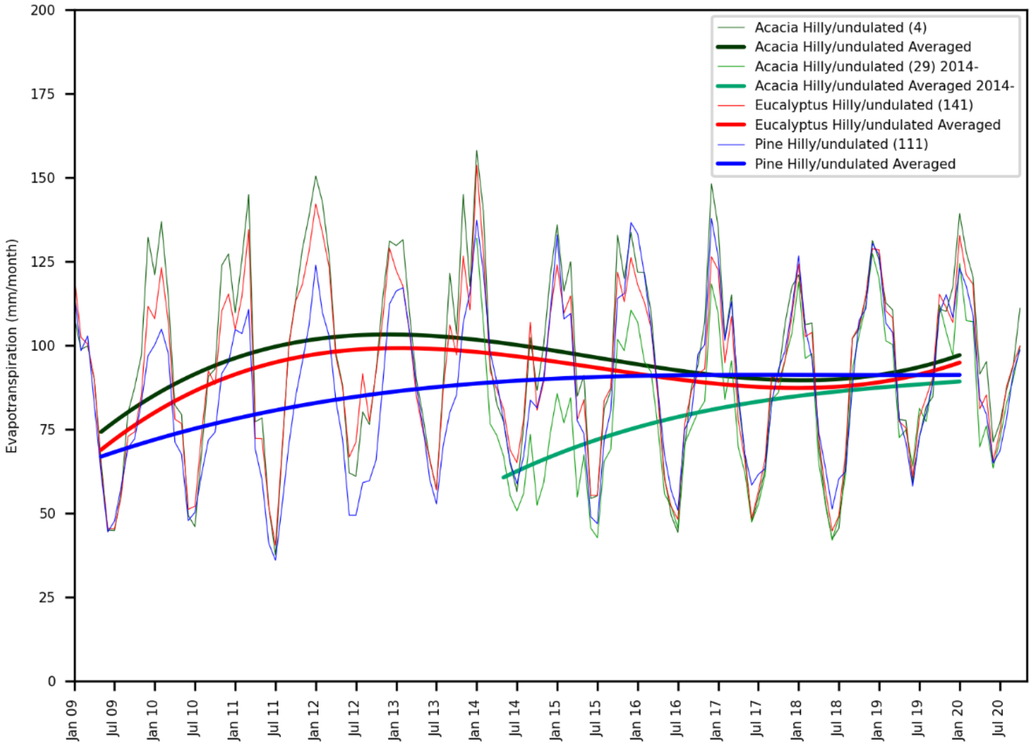


Figure 5-35 Water use over time (mm/month) for *Acacia* (planted in 2009 and 2014), *Eucalyptus* and *Pinus on* hilly/undulated terrain

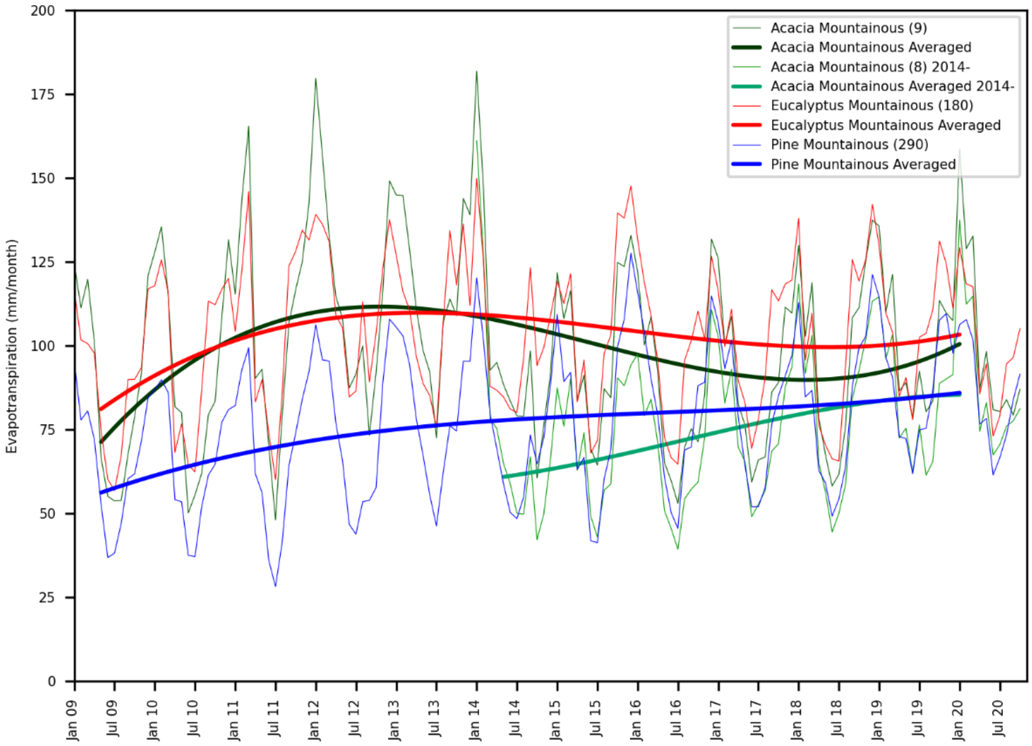


Figure 5-36 Water use over time (mm/month) for *Acacia* (planted in 2009 and 2014), *Eucalyptus* and *Pinus on* mountainous terrain

Table 5-7 Median ET values (mm/year) per morphological unit and genus

Morphological unit	Evapotranspiration (mm/year)		
	Acacia	***Eucalyptus***	***Pinus***
Plains/flat		1070 mm/year (n = 78)	1229 mm/year (n = 63)
Hilly/undulated	1147 mm/year (n = 4)	1142 mm/year (n = 141)	1044 mm/year (n = 111)
Mountainous	1149 mm/year (n = 9)	1253 mm/year (n = 180)	926 mm/year (n = 290)

6 DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS

Previous work on water use of forests has typically focussed on selected individual sites, which is difficult to extrapolate over regions. The purpose of this study was to improve our understanding of how water use (ET) varies at regional scales, focussing on commercial forestry. The first aim was to *establish a geographical database of commercial forests in the main commercial forestry regions of South Africa*. In this study, a large and rich dataset of forest compartments was sourced from several commercial companies. This dataset, along with monthly ET data from 2009 to 2020 sourced from the FAO-funded WaPOR portal, was used to *determine consumptive water use (actual ET) of commercial forestry by means of RS data*, which was the second aim of this study. The third aim was to *validate (ground truth) RS-based consumptive water use of commercial forestry plantations using historical field-based measurements.* It was challenging to achieve this aim, as few field-based measurements have been carried out in commercial forests, and in recent year much of the research has been concentrated in KZN and on limited tree species. Nevertheless, data from several studies were sourced and used to validate the RS-based ET estimates. The fourth and final aim was to *describe, analyse and interpret location-specific differences in water use between and within the primary commercial forestry tree genera at specific locations in South Africa.* A series of analyses were carried out to compare the RS-based ET estimates of different genera, species and age groups, grown under a range of environmental conditions. The methods and results of these experiments were detailed in Sections 3 and 5 respectively. Some of the main findings are summarised in the next subsection. The section concludes with proposals for future research, and a number of recommendations are made.

6.1 Main findings

The median WaPOR ET data extracted and used in this study compared well with in situ measurements of previous studies (Section 3.4.2). For instance, the ET estimates for *Acacia* (median ET = 1 096 mm/year, Table 3-2) and maximum annual ET estimates (1 600 mm/year, Table 3-2) are in line with those of previous studies (Table 3-3, Figure 3-25). Similarly, our median annual ET (1 123 mm/year) estimates for *Eucalyptus* is on par with the mean annual ET of 1 116 mm/year reported in previous studies (Table 3-3, Figure 3-25). The reported ranges of ET from previous studies (500-1 800 mm/year) also correspond well with the ET range of 575 to 1 618 mm/year estimated in this study. The relatively lower median annual ET estimated for *Pinus* (1 038 mm/year) in this study (Table 3-2, Figure 3-25) and the higher frequency of lower annual ET values (Figure 3-25) of less than 900 mm/year also agree with previous work, although our *Pinus* estimates are generally higher (Figure 3-20) than those reported in the literature (Table 3-3).

Although the comparison of the WaPOR-based ET estimates of this study corresponds well with previous in situ measurements, some level of error (uncertainty) is inevitable. A comparison between different sources of ET data (Section 3.3.4) revealed substantial differences, particularly between the ET values of the WaPOR and MOD16 products. The WaPOR ET values corresponded relatively well with the WRC 2014/15 dataset produced in a previous WRC project (Van Niekerk et al. 2018). The WRC 2014/15 dataset is considered to be the most accurate available dataset of South Africa, given that it was calibrated using seasonal climatic data captured by 239 weather stations around the country (which is considerably more than what is used in MOD16 and WaPOR). Some deviations between the WaPOR and the WRC 2014/15 dataset were observed during winter months, which suggests that the WaPOR dataset is overestimating water use by about 10-30 mm per month during these months. However, this overestimation is unlikely to have a substantial effect on overall (e.g. annual) water use estimates, as most forests are located in the summer rainfall region (Table 5-3). Also, the purpose of this project was not necessary to quantify the total water use of all forests, but rather to better understand how water use varies from one genus, species, age, region and type of site to another. The comparisons presented in this report are consequently valid, especially if one can assume that inaccuracies in the WaPOR data are consistent for different genera, species, age groups, regions, etc.

Our results showed that water use varies significantly among genera, with *Eucalyptus* species using considerably more water than *Pinus* and *Acacia* species, particularly during the first five years of planting (Figure 5-1). Although this observation agrees with in situ measurements carried out in previous research (Section 3.4.2), it is difficult to determine whether the relatively high water use of *Eucalyptus* species is biophysical in nature, or if it is a factor of site selection/quality (or a combination of these two factors) given that *Eucalyptus* species are commonly established in sites with higher quality (e.g. deep soils, high rainfall) compared to *Pinus* species. In all likelihood, the relatively high water use of *Eucalyptus* species is a combination of these factors.

Due to limited cases (compartments) planted, the water use of *Acacia* species varied considerably according to age and environmental conditions, and it was difficult to draw concrete conclusions. Based on the available data, the *Acacia* species use less water than *Eucalyptus* species, but the ET of *Acacia* compartments is generally higher than *Pinus* compartments.

The biggest environmental drivers of ET variability are rainfall and slope gradient. This can be expected, given that these two factors determine (to a large extent) water availability. Precipitation is less likely to infiltrate soils on steep slopes (i.e. a larger proportion of rainfall will contribute to surface runoff on steep slopes), while on level or moderately inclined terrain a greater proportion of rainfall will permeate soils and become available for take-up by vegetation. Flatter areas are also typically associated with deeper soils and likely higher soil water availability, which is conducive to tree growth.

A noteworthy finding of this study is the marked difference between the ET of *Eucalyptus* and *Pinus* compartments on moderately inclined and steep slopes, with the former having 10% higher values (on average). Overall, *Eucalyptus* species consistently used more water than *Pinus* species on moderately inclined and steep slopes. This is supported by the finding that the ET values of *Eucalyptus* compartments planted in mountainous regions are much higher (Figure 5-36) compared to those planted on plains/flat (Figure 5-34) and hilly/undulated (Figure 5-35) terrain, while the opposite is true for *Acacia* and *Pinus* compartments. Whether this observation is the result of site selection instead of biophysical factors does not really matter in the context of this study, which sought to identify and describe such variations rather than to determine the cause. Nevertheless, a better understanding of the factors driving these observations would greatly enhance our understanding of water use in the forestry sector and warrants further research.

6.2 Innovations and capacity building

This study contributed significantly to new knowledge. Specifically, the temporal water use profiles of plantation forest planted with different species/genera and in sites with varying environmental (e.g. climatic and terrain) conditions are novel. To our knowledge, the use of remote sensing to quantify water used in commercial forestry has not been done previously. The range of maps published in this report is also new. In addition, the use of deep learning for extracting forest patches from very high resolution (25 cm) colour aerial photography and the use of machine learning and satellite imagery for differentiating among tree genera at regional scales is innovative. These novel techniques were developed and evaluated by the two MSc students working on the project and demonstrates the significant human capacity that was built in this project.

6.3 Recommendations

The findings of this study may be of value to the ongoing discussions on the principles and processes for GE regulation within the commercial forestry industry (being the only declared SFRA). The variations in water use highlighted in this study should also be considered in forest rotation planning. This study showed that there are marked differences among the water use of commercial plantation forestry genera and species/hybrids. For instance, exchanging *Pinus* species for *Eucalyptus* species may have a detrimental effect on stream flow if carried out over large areas within a catchment. The impact of environmental conditions on water use should also be taken into consideration.

Ideally, the ET estimations produced in this study should be compared to actual rainfall per compartment to assess the relationship between water availability and use. Unfortunately, we did not have access to (did not budget for) weather station data from (costly) sources such as the ARC and SA Weather Services (SAWS). It is recommended that more be done to establish a dense network of weather stations throughout South Africa and that such data be made freely available for research purposes to support water use and accounting research. Although TerraClim is a step in the right direction, it requires more (financial) support from the research community.

This study made use of state-of-the-art remotely sensed satellite data and techniques to observe tree water use over an extensive area and period. Although this research project made a significant contribution to new knowledge, it took three years to complete. The South African forestry industry and regulators (e.g. government agencies) need such information to be updated on a regular basis. Ideally, operational solutions for calculating changes in water use associated with GE are required. These solutions should be based on scientifically sound techniques whereby GE regulations can be applied at the plantation stand/compartment level and across the country with the same statistical confidence. It is recommended that the techniques employed in this study be operationalised to produce water use estimations on an annual basis. This should be coupled with field-based measurements at strategic locations throughout South Africa to quantify the uncertainties in the resulting water use estimations.

The South African EO community is dwindling as many senior scientists in the field retire or emigrate. It is critical that we continue to invest in building EO capacity to assist the private and public sectors to optimally use scarce resources, such as water and fertile land. This is particularly important within the context of climate change, as the projected increases in temperatures and reduction in rainfall will have dire consequences for the forestry and agricultural industries. The central role that the WRC has played (and is playing) in building EO capacity is commendable; several students and young scientists were involved in this project (they are also co-authors to this report) and were exposed to advanced EO techniques. However, many of the students trained through WRC projects opt to emigrate and apply their skills abroad. More needs to be done to ensure that newly-trained scientists remain in South Africa. The only way to build a strong EO community is to establish employment opportunities that suitably incentivise scientists to remain in South Africa. Greater emphasis on the commercialisation and technology transfer of EO research is recommended, as such activities are more likely to generate sustainable employment opportunities. In addition, the commercialisation (and operationalisation) of EO technologies will substantially increase the impact of WRC-funded research and will ultimately lead to much needed economic growth and sustainable use of South Africa's limited water resources.

6.4 Proposals for future research

The scope of this study was limited to commercial forests. Although this focus provided invaluable insights into water use variations within the forestry sector, relatively little is known about the water use of state-owned forests (e.g. SAFCOL), small private growers, communal forests and forestry species that have spread and established as invasive alien plants. Future work should thus consider including a wider range of forestry data. Although the methods used in this study can be employed in such studies, the (250 m resolution) WaPOR dataset is not suitable for application on compartments smaller than 300 x 300 m (90 ha), which limits its use for estimating the water use in small stands.

It is possible to extract RS-based ET at higher resolutions (e.g. 20 m provided by FruitLook) using higher resolution (commercial) satellite data, but the cost may be prohibitively expensive for regional (national) implementations. Techniques for improving the spatial and temporal resolution of ET products using freely available HR satellites imagery (e.g. Sentinel-2) are also available and should be investigated.

Alternatively, a sample-based approach can be used whereby 250 m ET data can be used to extract water use in suitably large compartments, and these water use estimations can be extrapolated (per area unit) to surrounding, smaller compartments. This proposed sample-based approach can potentially be applied at a national scale to quantify the total water use by the forestry sector, similar to what was done for irrigated agriculture in a previous WRC project (Van Niekerk et al. 2018). Given that water use varies significantly per genus, it is recommended that a sample-based approach takes the genus of each compartment into consideration. The genus classification methodology developed in this study (Section 4.2) can be employed to classify each forest stand (even each pixel) into a genus category.

However, a major challenge will be to accurately map each individual forest compartment (stand). Manual mapping approaches (e.g. visual interpretation of satellite/aerial imagery and on-screen digitising) will likely be too time-consuming and costly. National land cover maps can be used as a sampling framework, as they contain the location and extent of planted forests and indigenous forests. Based on our assessments, the accuracy of such maps is high enough to be used at regional scales but are generally not detailed and accurate enough to be used at local or sub-catchment scales. Based on our assessments, small woodlots are often excluded or misclassified. At such sub-catchment scales, an alternative would be to develop and implement a mapping methodology that focuses on tree-based classes.

As was demonstrated in this study (Section 4.1), automated forest mapping approaches are viable, but more work is needed to translate site-specific (case study) methods to regional and national scales. The biggest challenge is to accurately separate plantation forests from indigenous forests – particularly in areas where these two land uses are intermixed – because the spectral characteristics of these two land uses/covers are very similar. Based on preliminary findings, CNN applied to VHR (e.g. 25-50 cm) colour imagery seems to be effective for this purpose, but more work is required to assess the transferability of the technique. It is likely that the most effective and robust approach would be a combination of methods (e.g. CNN on VHR imagery, combined by multitemporal analyses of Landsat/Sentinel optical and SAR imagery to determine stand age, tree structure and phenology). If necessary, multitemporal image analyses can also be used to estimate the age (planting date) of each stand (Section 4.3). This data can be used to improve the land cover classifications (e.g. reclassify plantation forest stands older than 40 years to indigenous forest) and to refine the ET estimations.

Extending the forestry water use results to forest water productivity and forest carbon sequestration of plantations by incorporating growth or production data and linking the water and carbon balances will be of great value to the industry. Such investigations will yield improved understating of the value of these products for improving water use estimations that transcend different land uses, including agriculture and rangelands. It is also critical to improve the local skills base and to build capacity in the use of EO methods and data for monitoring of water use.

This study only considered the water use of one land use (commercial forestry). Future research should consider all land uses, preferably within a water accounting framework, to assess the true impact of different land uses on stream flow and determine whether current and future land uses are sustainable from a water availability perspective.

Due to lack of data, this study did not investigate the impact of soil properties on water use. More work on the development of detailed soil maps for South Africa is needed. The initiative to develop a 30 m soil map for South Africa, started by a number of scientists from the Agricultural Research Council (ARC), Stellenbosch University and others, is a step in the right direction and requires (financial) support from the scientific community.

The increased availability of products such as WaPOR has brought ET data into the mainstream and has opened up many avenues for research. Such data are increasingly being utilised in decision-making, particularly in the agricultural sector. Currently, most of these products are being developed abroad. As was demonstrated in this study, there are substantial differences in the ET estimations among products, likely due to inadequate calibration to local climatic conditions (due to lack of weather station data). However, high density weather station networks are being established through initiatives such as Climate Smart Agriculture (climatesmartagri.co.za) and used by TerraClim (www.terraclim.co.za) to produce highly detailed climate surfaces that can be utilised to develop highly accurate and detailed ET products. The timing is right to start investing in establishing an ET modelling capacity within South Africa and to investigate the influence of landforms and soil characteristics on plantation forest (and agricultural) water use.

REFERENCES

'Audah S, Nazliyati, Bakruddin, Saputra E, Wathan S and Mina Rizky M (2019) Visual Analysis of Satellite Landsat Images Multitemporal and GPS as a Geographic Information System for Mapping of Nugmet Plantations in Tapaktuan. *IOP Conference Series: Materials Science and Engineering* 506: 012037.

Ahammad R, Stacey N, Eddy IMS, Tomscha SA and Sunderland TCH (2019) Recent trends of forest cover change and ecosystem services in eastern upland region of Bangladesh. *Science of The Total Environment* 647: 379-389.

Akumu CE, Baldwin K and Dennis S (2019) GIS-based modeling of forest soil moisture regime classes: Using Rinker Lake in northwestern Ontario, Canada as a case study. *Geoderma* 351: 25-35.

Albert T (2002) Evaluation of remote sensing techniques for ice-area classification applied to the tropical Quelccaya Ice Cap, Peru. *Polar Geography* 26: 210-226.

Bastiaanssen WGM, Cheema MJM, Immerzeel WW, Miltenburg I and Pelgrum H (2012) Surface energy balance and actual evapotranspiration of the transboundary Indus Basin estimated from satellite measurements and the ETLook model. *Water Resources Research* 48: W11512

Beinart W (1984) Soil erosion, conservationism and ideas about development: a Southern African exploration, 1900-1960. *Journal of Southern African Studies* 11: 52-83.

Bennett BM and Kruger FJ (2014) *Forestry and water conservation in South Africa: History, science and policy*. Canberra: The Australian National University.

Benz UC, Hofmann P, Willhauck G, Lingenfelder I and Heynen M (2004) Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing* 58: 239.

Blaschke T, Lang S, Lorup E, Strobl J and Zeil P (2000) Object-oriented image processing in an integrated GIS/remote sensing environment and perspectives for environmental applications. In Cremers A and Greve K (eds) *Environmental information for planning, politics and the public* 555-570. Marburg: Metropolis Verlag.

Bock M, Xofis P, Mitchley J, Rossner G and Wissen M (2005) Object-oriented methods for habitat mapping at multiple scales – Case studies from Northern Germany and Wye Downs, UK. *Journal for Nature Conservation* 13: 75-89.

Bolstad PV and Lillisand TM (1991) Rapid maximum likelihood classification. *Photogrammetric Engineering and Remote Sensing* 57: 67-74.

Bosch A, Zisserman A and Muoz X (2007) Image Classification using Random Forests and Ferns. Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference. Rio de Janeiro Brazil.

Bosch JM and Von Gadow K (1990) Regulating Afforestation for Water Conservation in South Africa. *South African Forestry Journal* 153: 41-54.

Breiman L (1996) Bagging Predictors. *Machine Learning* 24: 123-140.

Breiman L (2001) Random forests. *Machine learning* 45: 5-32.

Bristow KL and De Jager JM (1981) A proposed technique to measure evapotranspiration using micrometeorological methods. *Water SA* 7: 49-53.

Brown de Colstoun EC, Story MH, Thompson C, Commisso K, Smith TG and Irons JR (2003) National park vegetation mapping using multitemporal Landsat 7 data and a decision tree classifier. *Remote Sensing of Environment* 85: 316-327.

Buddenbaum H, Schlerf M and Hill J (2005) Classification of coniferous tree species and age classes using hyperspectral data and geostatistical methods. *International Journal of Remote Sensing* 26, 24: 5453-5465.

Bujang MA and Baharum N (2017) Guidelines of the minimum sample size requirements for Cohen 's Kappa. *Epidemiology Biostatistics and Public Health* 14, 2.

Campbell JB (2002) *Introduction to remote sensing*. 3rd edition. New York: The Guildford Press.

Campbell JB (2007a) *Introduction to remote sensing*. 4th Edition. London: Taylor and Francis.

Campbell JB (2007b) *Introduction to remote sensing*. London: Taylor and Francis.

Campbell JB and Wynne R (2013) *Introduction to remote sensing*. Remote Sensing 5,1

Chang J, Paelinckx D and Alterra-Centrum L (2008) Evaluation of Random Forest and Adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery. *Remote Sensing of Environment* 112: 2999-3011.

Chen B, Li X, Xiao X, Zhao B, Dong J, Kou W, Qin Y, Yang C, Wu Z, Sun R, Lan G and Xie G (2016) Mapping tropical forests and deciduous rubber plantations in Hainan Island, China by integrating PALSAR 25-m and multi-temporal Landsat images. *International Journal of Applied Earth Observation and Geoinformation* 50: 117-130.

Cho MA, Malahlela O and Ramoelo A (2015) Assessing the utility WorldView-2 imagery for tree species mapping in South African subtropical humid forest and the conservation implications: Dukuduku forest patch as case study. *International Journal of Applied Earth Observation and Geoinformation* 38: 349-357.

Choudhury BJ (1997) *Estimating areal evaporation using multispectral satellite observations.* NATO ASI Series, 364-381. Heidelberg: Springer Verlag

Clulow AD, Everson C and Gush M (2011) The long-term impact of *Acacia mearnsii* trees on evaporation, streamflow and groundwater resources. WRC Report no. TT 505/11. Pretoria: Water Research Commission.

Courault D, Seguin B and Olioso A (2005) Review on estimation of evapotranspiration from remote sensing data: From empirical to numerical modeling approaches. *Irrigation and Drainage Systems* 19: 223-249.

Cover T and Hart P (1967) Nearest neighbor pattern classification. *IEEE Transactions on Information Theory* 13: 21-27.

Cunningham PA and Delany SJ (2007) K-Nearest Neighbour Classifiers. Technical Report UCD-CSI-2007-4, Dublin.

Department of Agricultural Technical Services (1961) Report on the interdepartmental committee on the conservation of mountain catchments in South Africa. Pretoria: Department of Agricultural Technical Services.

Devi Mahalakshmi S and Geethanjali V (2019) Classification of CT Scan Images for Detection of Tuberculosis Using Convolutional Neural Network. *Journal of Advanced Research in Dynamical and Control Systems* 11, 7: 611-617.

Duro DC, Franklin SE and Dubé MG (2012) Multi-scale object-based image analysis and feature selection of multi-sensor earth observation imagery using random forests. International *Journal of Remote Sensing* 33: 4502-4526.

Dye PJ and Bosch JM (2000) Sustained water yield in afforested catchments: the South African experience. In: von Gadow L, Pukkala T and Tomé M (eds.) *Sustainable forest management*. Kluwer Academic Publishers, Dordrecht. 99-120.

Dye PJ, Gush MB, Everson CS, Jarmain C, Clulow AD, Mengistu M, Geldenhuys CJ, Wise R, Scholes RJ, Archibald S and Savage MJ (2008a) Water-use in relation to biomass of indigenous tree species in woodland, forest and/or plantation conditions. WRC Report no. K5/1462/1/08. Pretoria: Water Research Commission.

Dye PJ and Jarmain C (2004) Water use by black wattle (*Acacia mearnsii*): implications for the link between removal of invading trees and catchment streamflow response. *South African Journal of Science* 100.

Dye PJ, Jarmain C, Le Maitre DC, Everson CS, Gush MB and Clulow AD (2008b). Modelling Vegetation Water Use for General Application in Different Categories of Vegetation. WRC Report no. 1319/1/08. Pretoria: Water Research Commission.

Dye PJ, Poulter AG, Soko S and Maphanga D (1997) The determination of the relationship between transpiration rate and declining available water for *Eucalyptus* Grandis. WRC Report no. 441/1/97. Pretoria: Water Research Commission.

Dye PJ (2013) A review of changing perspectives on *Eucalyptus* water-use in South Africa. *Forest Ecology and Management* 301: 51-57.

Dye PJ (1996). Climate, forest and streamflow relationships in South African afforested catchments, Source: The Commonwealth Forestry Review.

Dye PJ, Jarmain C, le Maitre DM, Everson CS, Gush MB and Clulow AD (2008). Modelling vegetation water use for general application in different categories of vegetation. Pretoria: Water Research Commission.

Everson CS (1999) Evaporation from the Orange River: Quantifying open water resources. WRC Report No 683/1/99. Pretoria: Water Research Commission.

Everson CS, Clulow AD, Becker M, Watson A, Ngubo C, Bulcock H, Mengistu M, Lorentz S and Demlie M (2014) The long term impact of *Acacia mearnsii* trees on evaporation, streamflow, low flows and ground water resources. Phase II: Understanding the controlling environmental variables and soil water processes over a full crop rotation. WRC Report no. 2022/1/13. ISBN 978-1-4312-0516-5.

Everson CS, Dye PJ, Gush MB and Everson TM (2011) Water use of grasslands, agroforestry systems and indigenous forests. *Water SA* 37: 781-788.

Everson CS, Molefe GL and Everson TM (1998) Monitoring and modelling components of the water balance in a grassland catchment in the summer rainfall area of South Africa. WRC Report no. 493/1/98. Pretoria: Water Research Commission.

Everson CS, Moodley M, Gush MB, Jarmain C, Govender MA and Dye PJ (2005). Effective management of the riparian zone vegetation to significantly reduce the cost of catchment management and enable greater productivity of land resources. WRC Report no. 1284/1/07. Pretoria: Water Research Commission.

Everson CS, Clulow AD, Becker M, Ngubo C, Bulcock H, Mengistu M, Lorentz S and Demlie M (2014). The long-term impact of *Acacia* mearnsii trees on evaporation, streamflow and groundwater resources. Phase II, Understanding the controlling environmental variables and soil water processes over a full crop rotation. Pretoria.

Everson CS, Gush MB, Moodley M, Jarmain C, Govender M, Dye P (2007) Effective management of the riparian zone vegetation to significantly reduce the cost of catchment management and enable greater productivity of land resources. Pretoria: Water Research Commission.

Fagan ME, DeFries RS, Sesnie SE, Arroyo-Mora JP, Soto C, Singh A, Townsend PA and Chazdon RL (2015) Mapping species composition of forests and tree plantations in northeastern Costa Rica with an integration of hyperspectral and multitemporal landsat imagery. *Remote Sensing* 7, 5: 5660-5696.

Falkowski MJ, Hudak AT, Crookston NL, Gessler PE, Uebler EH and Smith AMS (2010) Landscape-scale parameterization of a tree-level forest growth model: a k-nearest neighbor imputation approach incorporating LiDAR data. *Canadian Journal of Forest Research* 40: 184-199.

Franco-Lopez H, Ek AR and Bauer ME (2001) Estimation and mapping of forest stand density, volume, and cover type using the k-nearest neighbors method. *Remote Sensing of Environment* 77: 251-274.

Francois A and Leckie DG (2006) The individual tree crown approach to Ikonos images. *Photogrammetric Engineering & Remote Sensing* 72, 11: .1287-1297.

Franklin SE and Ahmed OS (2018) Deciduous tree species classification using object-based analysis and machine learning with unmanned aerial vehicle multispectral data. *International Journal of Remote Sensing* 39, 15-16: 5236-5245.

Franklin SE, Ahmed OS and Williams G (2017) Northern conifer forest species classification using multispectral data acquired from an unmanned aerial vehicle. *Photogrammetry Engineering & Remote Sensing* 83, 7: 501-507.

Gibson LA, Jarmain C, Su Z and Eckardt FE (2013) Estimating evapotranspiration using remote sensing and the Surface Energy Balance System – A South African perspective. *Water SA* 39: 477-483.

Gibson PJ and Power CH (2000) *Introduction to remote sensing: Digital image processing and applications*. London: Taylor and Francis.

Gislason PO, Benediktsson JA and Sveinsson JR (2006) Random forests for land cover classification. *Pattern Recognition Letters* 27: 294-300.

Gómez C, Wulder MA, Montes F and Delgado JA (2012) Modeling Forest Structural Parameters in the Mediterranean Pines of Central Spain using QuickBird-2 Imagery and Classification and Regression Tree Analysis (CART). *Remote Sensing* 4: 135-159.

Goodchild MF (2005) *GIS and modeling overview: GIS, spatial analysis, and modeling*. Redlands: ESRI Press.

Goudriaan R (2014) An operational service to improve crop water and nitrogen management in grapes and other deciduous fruit trees using satellite technology for the season 2013-14. Cape Town: Western Cape Province Department of Agriculture.

Green S and Clothier B (1988) Water use of kiwifruit vines and apple trees by the heat-pulse technique. *Journal of Experimental Botany* 39.

Gumbricht T, McCarthy J and Mahlander C (1996) Digital interpretation and management of land-cover: A case study of Cypres. *Ecological Engineering* 6: 273-279.

Gush MB and Dye PJ (2009) Water-use efficiency within a selection of indigenous and exotic tree species in South Africa as determined using sap flow and biomass measurements. *Acta Horticulturae* 846: 323-330.

Gush MB, Scott DF, Jewitt GPW, Schulze RE, Hallowes LA and Görgens AHM (2002) A new approach to modelling streamflow reductions resulting from commercial afforestation in South Africa. *Southern African Forestry Journal* 196: 27-36.

Hagner O and Reese H (2007) A method for calibrated maximum likelihood classification of forest types. *Remote Sensing of Environment* 110: 438-444.

Han D, Liu Q and Fan W (2018) A new image classification method using CNN transfer learning and web data augmentation. *Expert Systems with Applications* 95: 43-56.

Hansen MC and Loveland TR (2012) A review of large area monitoring of land cover change using Landsat data. *Remote Sensing of Environment* 122: 66-74.

Hardisty J, Taylor DM and Metcalfe SE (1993) *Computerised environmental modelling: a practical introduction using Excel*. Chichester, West Sussex: John Wiley and Sons.

Hay GJ, Castilla G, Wulder MA and Ruiz JR (2005) An automated object-based approach for the multiscale image segmentation of forest scenes. *International Journal of Applied Earth Observation and Geoinformation* 7: 339-359.

Hepner F, Logan T, Ritter N and Bryant N (1990) Artificial neural network classification using a minimal training set: Comparison to conventional supervised classification. *Photogrammetric Engineering and Remote Sensing* 56: 469-473.

Heydari SS and Mountrakis G (2019) Meta-analysis of deep neural networks in remote sensing: A comparative study of mono-temporal classification to support vector machines. *ISPRS Journal of Photogrammetry and Remote Sensing* 156: 192-210

Heywood I, Cornelius S and Carver S (2006) *An introduction to geographical information systems, 3rd edition.* Harlow, England: Pearson Education Limited.

Hladik C and Alber M (2014) Classification of salt marsh vegetation using edaphic and remote sensing-derived variables. *Estuarine, Coastal and Shelf Science* 141: 47-57.

Holmgren P and Thuresson T (1998) Satellite remote sensing for forestry planning – A review. *Scandinavian Journal of Forest Research* 13, 1-4: 90-110.

Hu Y and Hu Y (2020) Detecting forest disturbance and recovery in Primorsky Krai, Russia, using annual Landsat time series and multi-source land cover products. *Remote Sensing* 12, 21:

Huang C, Davis LS and Townshend JRG (2002) An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing* 23: 725-749.

Hubert-Moy L, Cotonnec A, Le Du L, Chardin A and Perez P (2001) A Comparison of Parametric Classification Procedures of Remotely Sensed Data Applied on Different Landscape Units. *Remote Sensing of Environment* 75: 174-187.

Immitzer M, Atzberger C and Koukal T (2012) Tree Species Classification with Random Forest Using Very High Spatial Resolution 8-Band WorldView-2 Satellite Data. *Remote Sensing* 4: 2661-2693.

Jain AK, Duin RPW and Jianchang M (2000) Statistical pattern recognition: a review. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22: 4-37.

Jarmain C and Everson CS (2002) Comparative evaporation measurements above commercial forestry and sugarcane canopies in the KwaZulu-Natal Midlands. CSIR Report no. ENV-C-S 2002-005. Pretoria: CSIR.

Jarmain C, Bastiaanssen WGM, Mengistu M, Jewitt G and Kongo V (2009a) A methodology for near-real spatial estimation of evaporation. WRC Report no. K5/1751:2009. Pretoria: Water Research Commission.

Jarmain C, Everson CS, Savage MJ, Mengistu M and Clulow AD (2009b). Refining tools for evaporation monitoring in support of water resources management, WRC Report no. K5/1567/08. Pretoria: Water Research Commission.

Jarmain C, Klaasse A, Basson FC, Meijninger W, Wilmink S and Bastiaanssen WGM (2011) Developing an operational remote sensing system for monitoring of efficient crop water and nitrogen use of grapes. Stellenbosch: Western Cape Province Department of Agriculture.

Jarmain C, Klaasse A, Basson F, Meijninger W, Wilmink S and Bastiaanssen WGM (2014) Water use efficiency of selected irrigated crops determined with satellite imagery. WRC Report no. TT 602/14. Pretoria: Water Research Commission.

Kalma JD, McVicar TR and McCabe MF (2008) Estimating land surface evaporation: A review of methods using remotely sensed surface temperature data. *Surveys in Geophysics* 29: 421-469.

Karssenberg D, Schmitz O, De Vries LM and De Jong K (2008) A tool for construction of stochastic spatio-temporal models assimilated with observational data. 11th AGILE International Conference on Geographic Information Science, Girona, Spain.

Ke Y, Quackenbush LJ and Im J (2010) Remote sensing of environment synergistic use of QuickBird multispectral imagery and LIDAR data for object-based forest species classification. *Remote Sensing of Environment* 114, 6: 1141-1154.

Kou W, Xiao X, Dong J, Gan S, Zhai D, Zhang G, Qin Y and Li L (2015) Mapping Deciduous Rubber Plantation Areas and Stand Ages with PALSAR and Landsat Images. *Remote Sensing* 7:1048-1073.

Knorn J, Rabe A, Radeloff VC, Kuemmerle T, Kozak J and Hostert P (2009) Land cover mapping of large areas using chain classification of neighboring Landsat satellite images. *Remote Sensing of Environment* 113, 5: 957-964.

Koskinen J, Leinonen U, Vollrath A, Ortmann A, Lindquist E, d'Annunzio R, Pekkarinen A and Käyhkö N (2019) Participatory mapping of forest plantations with Open Foris and Google Earth Engine. *ISPRS Journal of Photogrammetry and Remote Sensing* 148: 63-74.

Kustas WP and Norman JM (1996) Use of remote sensing for evapotranspiration monitoring over land surfaces. *Hydrological Sciences Journal* 41: 495-516.

Laborte AG, Maunahan AA and Hijmans (2010) Spectral Signature Generalization and Expansion Can Improve the Accuracy of Satellite Image Classification. *PLoS ONE* 5, 5:e10516.

Lang S (2008) Object-based image analysis for remote sensing applications: Modelling reality – dealing with complexity. In Blaschke T, Lang S and Hay G (eds) *Object-based image analysis: Spatial concepts for knowledge-driven remote sensing applications*, 3-26. Berlin/Heidelberg: Springer Verlag.

Lawrence L and Wright A (2001) Rule-based classification systems using classification and regression tree (CART) analysis. *Photogrammetric Engineering and Remote Sensing* 67: 1137-1142.

Lawrence RL (2006) Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (RandomForest). *Remote Sensing of Environment* 100: 356-362.

Le Maitre DC and Versfeld DB (1997) Forest evaporation models: relationships between stand growth and evaporation. *Journal of Hydrology* 193: 240-257.

Lennartz SP and Congalton RG (2004) Classifying and mapping forest cover types using IKONOS imagery in the Northeastern United States. Proceedings of the ASPRS Annual Conference, Denver, Colorado.

Li H, Gu H, Han Y and Yang J (2010) Object-oriented classification of high-resolution remote sensing imagery based on an improved colour structure code and a support vector machine. *International Journal of Remote Sensing* 31: 1453-1470.

Li ZL, Tang BH, Wu H, Ren H, Yan G, Wan Z, Trigo IF and Sobrino JA (2013) Satellite-derived land surface temperature: Current status and perspectives. *Remote Sensing of Environment* 131: 14-37.

Lira J and Maletti G (2002) A supervised contextual classifier based on a region-growth algorithm. *Computers and Geosciences* 25: 951-959.

Lira Melo de Oliveira Santos C, Augusto Camargo Lamparelli R, Figueiredo G, Dupuy S, Boury J, Luciano A, Torres R and le Maire G (2019) Classification of Crops, Pastures, and Tree Plantations along the Season with Multi-Sensor Image Time Series in a Subtropical Agricultural Region. *Remote Sensing* 11: 334.

Lizarazo I (2008) SVM-based segmentation and classification of remotely sensed data. *International Journal of Remote Sensing* 29: 7277-7283.

Lui X, Skidmore AK and Van Oosten H (2002) Integration of classification methods for improvement of land-cover map accuracy. *ISPRS Journal of Photogrammetry and Remote Sensing* 56: 257-268.

Maniatis D, Malhi Y, Saint André L, Mollicone D, Barbier N, Saatchi S, Henry M, Tellier L, Schwartzenberg M and White L (2011) Evaluating the potential of commercial forest inventory data to report on forest carbon stock and forest carbon stock changes for REDD+ under the UNFCCC. *International Journal of Forestry Research* 2011: 1-13.

Mather PM (2004) *Computer Processing of Remotely-Sensed Images. 3rd*. Chicester, West Sussex: Wiley.

Mati A and Dawaki SA (2015) Role of forest inventory in sustainable forest management: A review. *International Journal of Forestry and Horticulture* 1, 2: 33-40.

Mauya EW, Koskinen J, Tegel K, Hämäläinen J, Kauranne T and Käyhkö N (2019) Modelling and Predicting the Growing Stock Volume in Small-Scale Plantation Forests of Tanzania Using Multi-Sensor Image Synergy. *Forests* 10: 279.

McMahon DE and Jackson RB (2019) Management intensification maintains wood production over multiple harvests in tropical *Eucalyptus* plantations. *Ecological Applications* 29.

McRoberts RE and Tomppo EO (2007) Remote sensing support for national forest inventories. *Remote Sensing of Environment* 110: 412-419.

Meijninger WML and Jarmain C (2014) Satellite-based annual evaporation estimates of invasive alien plant species and native vegetation in South Africa. *Water SA* 40:95-107.

Mngadi M, Odindi J, Peerbhay K and Mutanga O (2019) Examining the effectiveness of Sentinel-1 and 2 imagery for commercial forest species mapping. *Geocarto International* 0, 0: 1-12.

Myburgh G and Van Niekerk A (2013) Effect of feature dimensionality on object-based land cover classification: A Comparison of Three Classifiers. *South African Journal of Geomatics* 2(1): 13-27.

Nangendo G, Skidmore AK and van Oosten H (2007) Mapping east African tropical forests and woodlands – A comparison of classifiers. *ISPRS Journal of Photogrammetry and Remote Sensing* 61: 393-404.

Nänni UW (1970) Trees, Water and Perspective. *South African Forestry Journal* 75: 9-17.

Nery T, Sadler R, Solis Aulestia M, White B and Polyakov M (2019) Discriminating native and plantation forests in a Landsat time-series for land use policy design. *International Journal of Remote Sensing* 40: 4059-4082.

Neusch T and Grussenmeyer P (2003) Remote sensing object-oriented image analysis applied to half-timbered houses. Proceedings of the CIPA XIX[th] International symposium, Antalya, Turkey.

Nomura K and Mitchard ETA (2018) More Than Meets the Eye: Using Sentinel-2 to Map Small Plantations in Complex Forest Landscapes. *Remote Sensing* 10, 11:1693

Novack T, Esch T, Kux H and Stilla U (2011) Machine learning comparison between WorldView-2 and QuickBird-2-simulated imagery regarding object-based urban land cover classification. *Remote Sensing* 3: 2263-2282.

Olthof I, Butson C and Fraser R (2005) Signature extension through space for northern landcover classification: A comparison of radiometric correction methods. *Remote Sensing of Environment* 95, 3: 290-302.

O'Sullivan D and Unwin DJ (2010) *Geographic data analysis, 2nd edition*. Hoboken: John Wiley and Sons, Inc.

Oruc M, Marangoz AM and Buyuksalih G (2004) Comparison of pixel-based and object-oriented classification approaches using Landsat-7 ETM spectral bands. Proceedings of the ISRPS 2004 Annual Conference, Istanbul, Turkey.

Pal M and Mather PM (2003) An assessment of the effectiveness of decision tree methods for land cover classification. *Remote Sensing of Environment* 86: 554-565.

Pal M (2005) Random forest classifier for remote sensing classification. *International Journal of Remote Sensing* 26: 217-222.

Parker G (2020) Tamm review: leaf area index (LAI) is both determinant and consequence of important processes in vegetation canopies. *Forest Ecology and Management* 477: 1-15.

Pax-Lenney M, Woodcock CE, Macomber SA, Gopal S & Song C (2001) Forest mapping with a generalized classifier and Landsat TM data. *Remote Sensing of Environment.* 77, 3: 241-250.

Peerbhay KY, Mutanga O and Ismail R (2013) Commercial tree species discrimination using airborne AISA Eagle hyperspectral imagery and partial least squares discriminant analysis (PLS-DA) in KwaZulu-Natal, South Africa. *ISPRS Journal of Photogrammetry and Remote Sensing* 79: 19-28.

Pelgrum H, Miltenburg IJ, Cheema MJM, Klaasse A and Bastiaanssen WGM (2011) ETLook: A novel continental evapotranspiration algorithm. *Remote Sensing and Hydrology*, Jackson Hole, Wyoming, USA.

Petropoulos GP, Kalaitzidis C and Prasad Vadrevu K (2012) Support vector machines and object-based classification for obtaining land-use/cover cartography from Hyperion hyperspectral imagery. *Computers and Geosciences* 41: 99-107.

Poynton RJ (1971) A silvicultural map of southern Africa. *South African Journal of Science* 67: 58-60.

Pu R and Landry S (2012) A comparative analysis of high spatial resolution IKONOS and WorldView-2 imagery for mapping urban tree species. *Remote Sensing of Environment* 124: 516-533. [online]. Available from: http://dx.doi.org/10.1016/j.rse.2012.06.011

Punia M, Joshi PK and Porwal MC (2011) Decision tree classification of land use land cover for Delhi, India using IRS-P6 AWiFS data. *Expert Systems with Applications* 38: 5577-5583.

Rego FL and Koch B (2003) Automatic classification of land cover with high resolution data of the Rio de Janeiro city Brazil comparison between pixel and object classification. In Carstens J (ed) *Remote Sensing of Urban Areas* 153-157.

Running SW, Mu Q, Zhao M and Moreno A (2017) User's Guide: MODIS Global Terrestrial Evapotranspiration (ET) Product (NASA MOD16A2/A3). NASA Earth Observing System MODIS Land Algorithm: Version 1.5. Available online at https://ladsweb.modaps.eosdis.nasa.gov/missions-and-measurements/modis/MOD16_ET_User-Guide_2017.pdf

Rodriguez-Galiano VF, Chica-Olmo M, Abarca-Hernandez F, Atkinson PM and Jeganathan C (2012) Random Forest classification of Mediterranean land cover using multi-seasonal imagery and multi-seasonal texture. *Remote Sensing of Environment* 121: 93-107.

Savage MJ, Everson CS and Metelerkamp BR (1997) Evaporation measurement above vegetated surfaces using micrometeorological techniques. WRC Report no. 349/1/97. Pretoria: Water Research Commission.

Savage MJ, Everson CS, Odhiambo GO, Mengistu MJ and Jarmain C (2004) Theory and practice of evaporation measurement, with special focus on SLS AS AN operational tool for the estimation of the spatially-averaged evaporation. WRC Report no. 1335/1/04. Pretoria: Water Research Commission.

Savage MJ, Odhiambo GO, Mengistu MG, Everson CS and Jarmain C (2010) Measurement of grassland evaporation using a surface-layer scintillometer. *Water SA* 36: 1-8.

Schulze RE (1997) South African atlas of agrohydrology and -climatology. WRC Report no. TT82/96. Pretoria: Water Research Commission.

Scott DF and Lesch W (1995) The water yield gains obtained from clearfelling riparian zone vegetation. Proceedings of the 7th South African National Hydrological Symposium, Grahamstown, September 1995.

Scott DF and Lesch W (1997) Streamflow responses to afforestation with *Eucalyptus grandis* and *Pinus patula* and to felling in the Mokobulaan experimental catchments, Mpumalanga Province, South Africa. *Journal of Hydrology* 199:360

Scott DF and Smith RE (1997) Preliminary empirical models to predict reductions in total and low flows resulting from afforestation. *Water SA* 23: 135-140.

Scott DF, Prinsloo FW, Moses G, Mehlomakulu M and Simmers ADA (2000) A re-analysis of the South African catchment afforestation experimental data. WRC Report no. 810/1/00. Pretoria: Water Research Commission.

Shiba M and Itaya A (2006) Using eCognition for improved forest management and monitoring systems in precision forestry. Proceedings of the International Precision Forestry Symposium, Stellenbosch, South Africa.

Skidmore AK, Turner BJ, Brinkhof W and Knowles E (1997) Performance of a neural network: Mapping forests using GIS and remotely sensed data. *Photogrammetric Engineering and Remote Sensing* 63, 5: 501-514.

Stabach JA, Dabek L, Jensen R and Wang YQ (2009) Discrimination of dominant forest types for Matschie's tree kangaroo conservation in Papua New Guinea using high-resolution remote sensing data. *International Journal of Remote Sensing* 30, 2: 405-422.

Tang L and Shao G (2015) Drone remote sensing for forestry research and practices. *Journal of Forestry Research* 26, 4: 791-797.

Tseng M-H, Chen S-J, Hwang G-H and Shen M-Y (2008) A genetic algorithm rule-based approach for land-cover classification. ISPRS Journal of Photogrammetry and Remote Sensing 63: 202-212.

Tzotsos A and Argialas D (2008) A support vector machine approach for object based image analysis. *Object-Based Image Analysis* 663-667.

Van Coillie FMB, Verbeke LPC and De Wulf RR (2007) Feature selection by genetic algorithms in object-based classification of IKONOS imagery for forest mapping in Flanders, Belgium. *Remote Sensing of Environment*. 110: 476-487.

Van der Zel WD (1995) Accomplishments and Dynamics of the South African Afforestation Permit System. *South African Forestry Journal* 172: 49-58.

Van Niekerk A, Jarmain C, Goudriaan R, Muller J, Ferreira F, Munch Z, Pauw T, Stephenson GR and Gibson LA (2018) An earth observation approach towards mapping irrigated areas and quantifying water use by irrigated crops in South Africa. WRC Report no. TT 745/17. Pretoria: Water Research Commission.

Van Niekerk A (2015) Stellenbosch University Digital Elevation Model (SUDEM) Product Description v15.15. 10.13140/RG.2.1.1795.6324.

Van Wyk DB (1987) Some effects of afforestation on streamflow in the Western Cape Province, South Africa. *Water SA* 13: 31-36.

Vapnik VN (1995) *The nature of statistical learning theory*. Springer-Verlag.

Verstraeten WW, Veroustraete F and Feyen J (2005) Estimating evapotranspiration of European forests from NOAA-imagery at satellite overpass time: Towards an operational processing chain for integrated optical and thermal sensor data products. *Remote Sensing of Environment* 96: 256-276.

Verstraeten WW, Veroustraete F and Feyen J (2008) Assessment of evapotranspiration and soil moisture content across different scales of observation. *Sensors* 8: 70-117.

Voss M and Sugumaran R (2008) Seasonal effect on tree species classification in an urban environment using hyperspectral data, LiDAR, and an object-oriented approach. *Sensors* 8, 5: 3020-3036.

Wagner F, H, Sanchez A, Tarabalka Y, Lotte R, Ferreira M, Aidar PM, Gloor M, Phillips O and Aragão L (2019) Using the U-net convolutional network to map forest types and disturbance in the Atlantic rainforest with very high resolution images. *Remote Sensing in Ecology and Conservation.*

Wicht CL (1948) Hydrological research in South African forestry. *Journal of the South African Forestry Association* 16: 4-21.

Wicks TE, Smith GM and Curran PJ (2002) Polygon-based aggregation of remotely sensed data for regional ecological analyses. *International Journal of Applied Earth Observation and Geoinformation* 4: 161-173.

Woodcock CE, Macomber SA, Pax-Lenney M and Cohen WB (2001) Monitoring large areas for forest change using Landsat: Generalization across space, time and Landsat sensors. *Remote Sensing of Environment* 78, 1-2: 194-203.

Ying L and Bo C (2009) An improved k-nearest neighbor algorithm and its application to high resolution remote sensing image classification. 2009 17th International Conference on Geoinformatics.

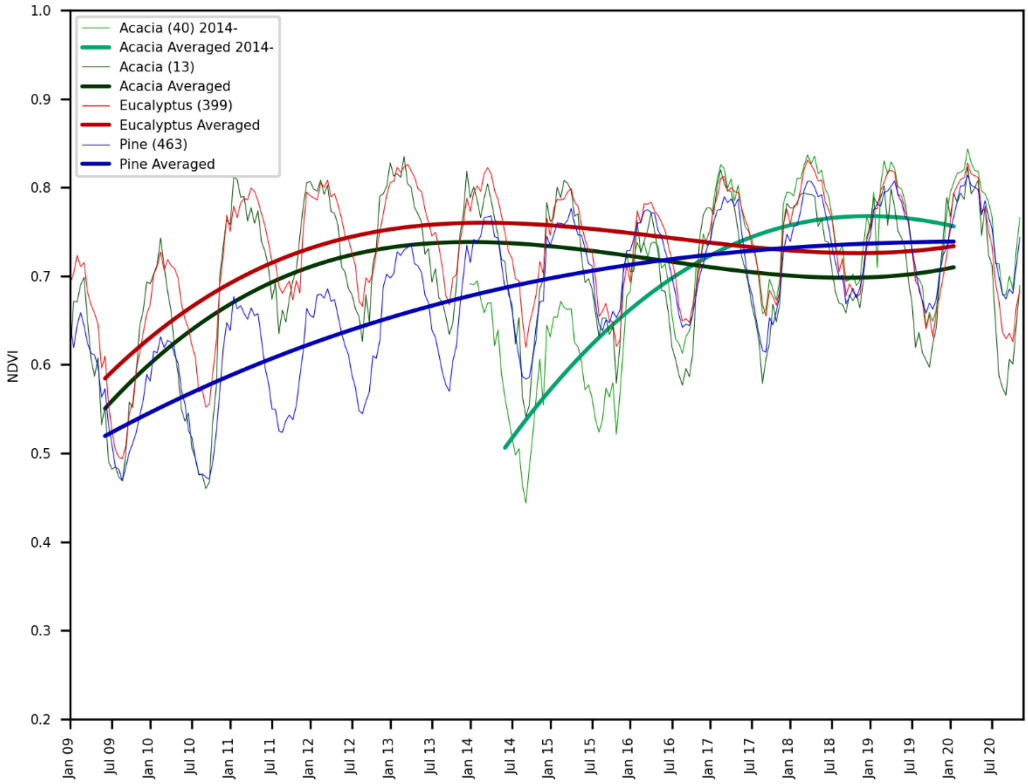
Zahn GA and Neethling EJ (1929) Notes on the exotic trees in the Cape Peninsula. *South African Journal of Science* 26: 211-234.

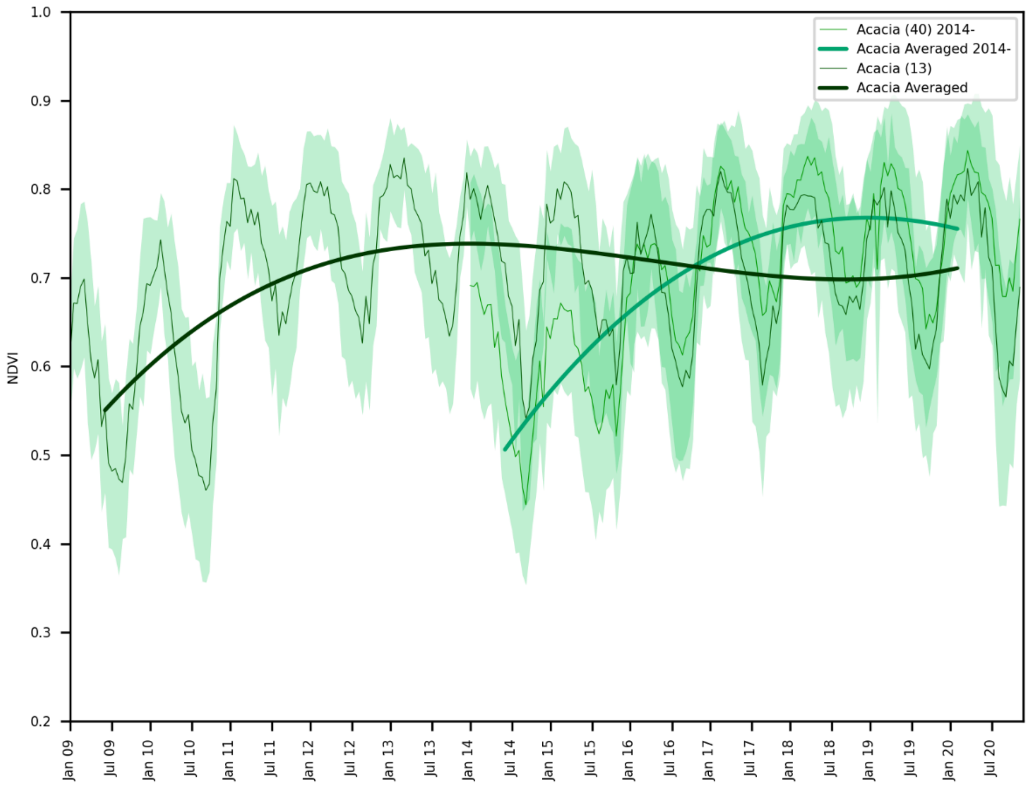
APPENDIX I: CFDB SPECIES DATA

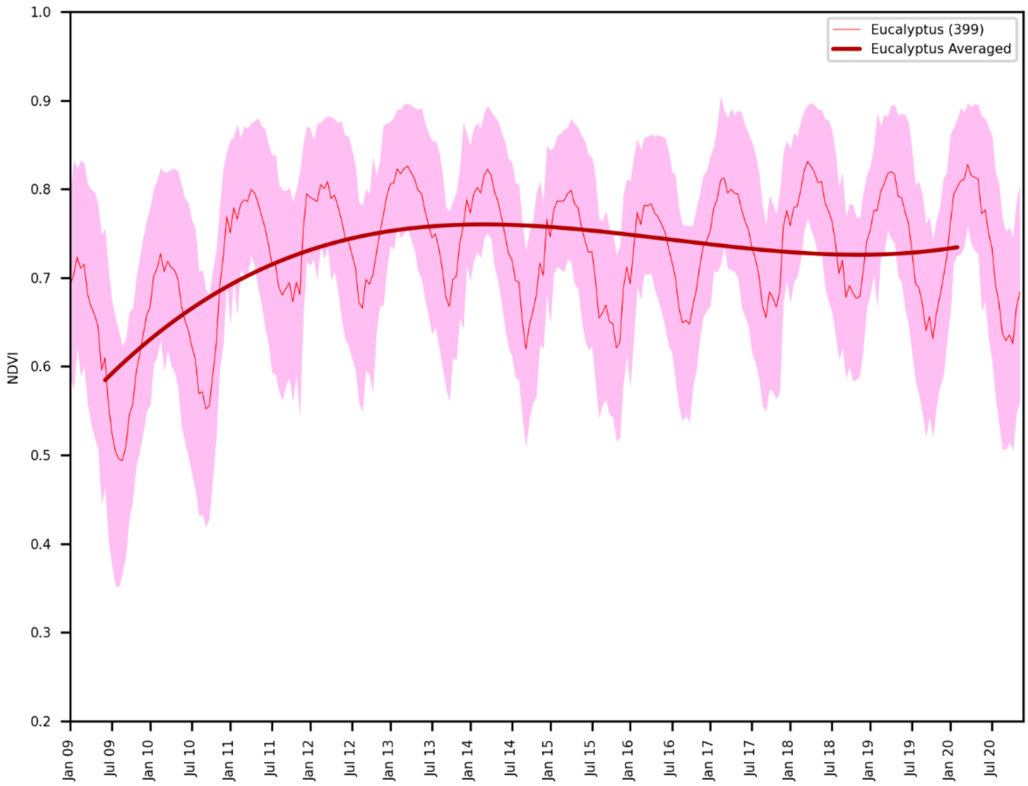
Species	**#**	***Species***	**#**
Acacia juncifolia	1	*Pinus* (mixed)	50
Acacia mearnsii	394	*Pinus* (unknown)	4
Acacia melanoxylan	20	*Pinus elliottii*	3439
Corymbia henryi	35	*Pinus elliottii + P. radiata*	12
Eucalyptus (mixed)	61	*Pinus elliottii x P. caribaea hybrid*	675
Eucalyptus (other)	1	*Pinus greggii*	40
Eucalyptus (unknown)	5	*Pinus greggii northern population*	1
Eucalyptus badjensis	68	*Pinus greggii southern population*	30
Eucalyptus benthamii	782	*Pinus hybrids*	3
Eucalyptus camaldulensis	1	*Pinus kesiya*	3
Eucalyptus cladocalyx	3	*Pinus oocarpa*	2
Eucalyptus cloeziana	6	*Pinus patula*	3265
Eucalyptus diversicolor	20	*Pinus patula + P. elliotii*	5
Eucalyptus dorrigoensis	6	*Pinus patula + P. greggii*	2
Eucalyptus dunnii	6120	*Pinus patula + P. radiata*	1
Eucalyptus fastigata	3	*Pinus patula x P. oocarpa hybrid*	13
Eucalyptus globulus	2	*Pinus patula x P. taeda hybrid*	19
Eucalyptus grandis	1421	*Pinus patula x P. tecunumanii hybrid*	28
Eucalyptus grandis x E. camaldulensis hybrid	166	*Pinus patula x P. tecunumanii lower altitude variety*	458
Eucalyptus grandis x E. macarthurii hybrid	1	*Pinus pinaster*	109
Eucalyptus grandis x E. nitens hybrid	2668	*Pinus pseudostrobus*	4
Eucalyptus grandis x E. saligna hybrid	1	*Pinus radiata*	439
Eucalyptus grandis x E. tereticornis hybrid	2	*Pinus radiata + P elliottii*	9
Eucalyptus grandis x E. urophylla hybrid	3872	*Pinus radiata + P. patula*	4
Eucalyptus macarthurii	463	*Pinus radiata + P. patula x P. tecunimanii hybrid*	1
Eucalyptus nitens	368	*Pinus radiata + P. tecunumanii*	2
Eucalyptus paniculata	1	*Pinus taeda*	523
Eucalyptus paniculata x E. grandis hybrid	24	*Pinus taeda + P. radiata*	1
Eucalyptus saligna	14	*Pinus tecunumanii*	6
Eucalyptus saligna x E. urophylla hybrid	3		
Eucalyptus smithii	324		
Eucalyptus urophylla	3		
Eucalyptus viminalis	4		

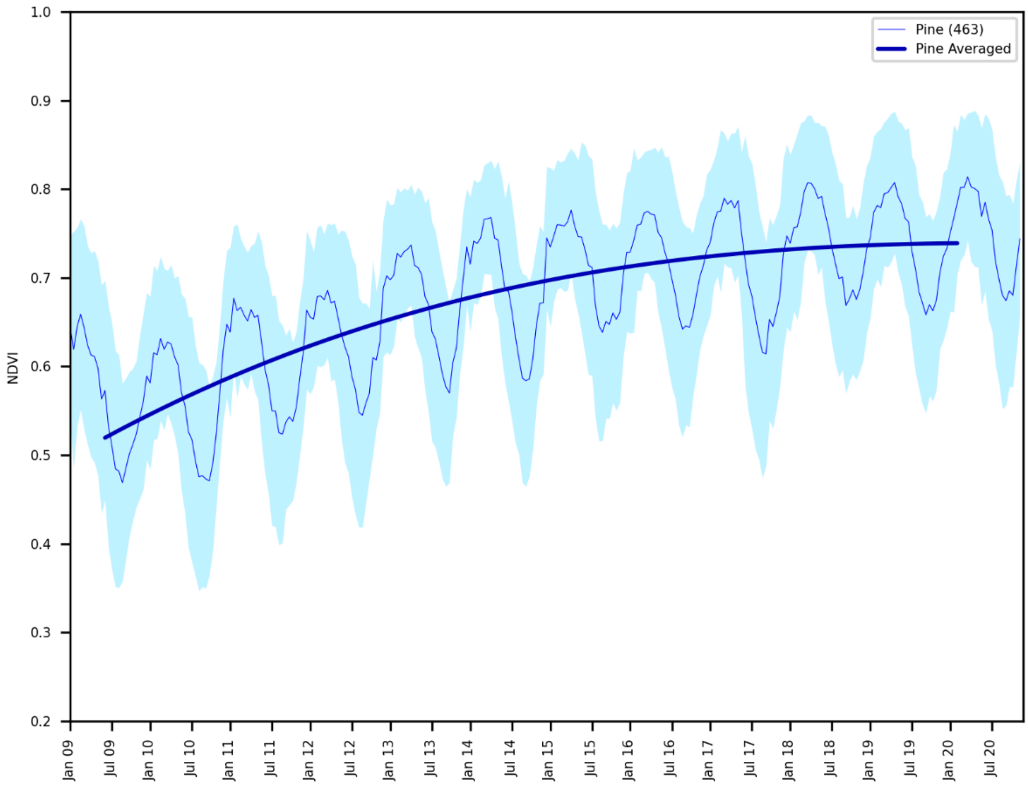
APPENDIX II: NDVI PROFILES

Compared to Compartment Age

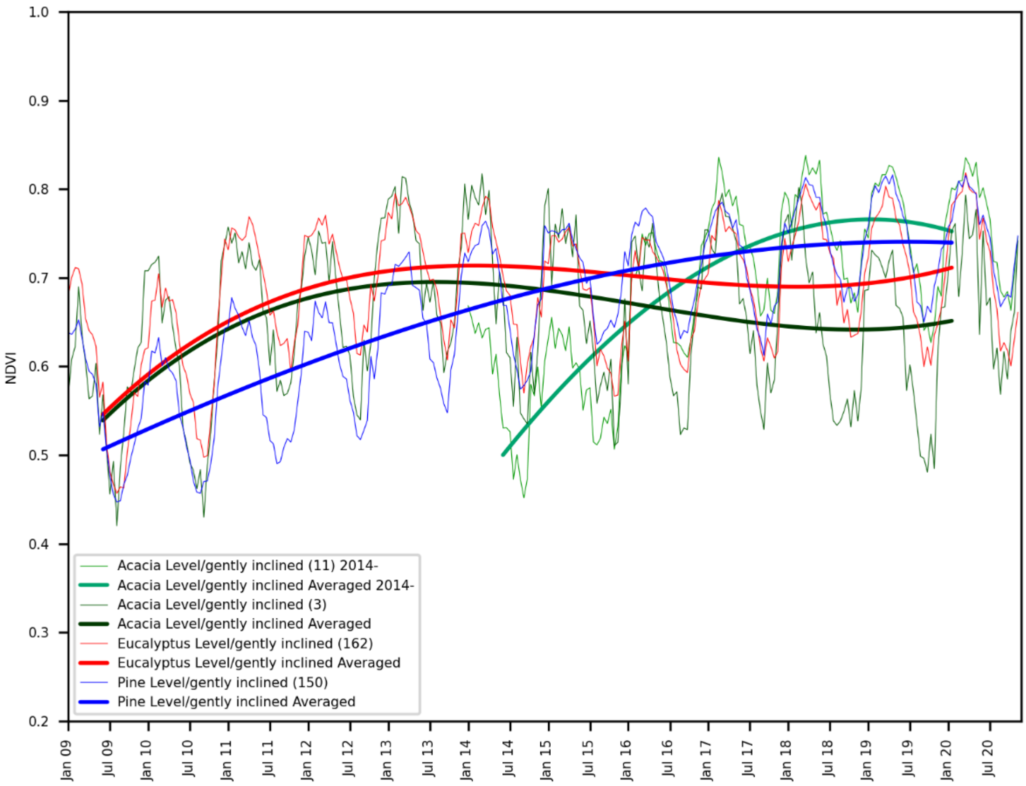


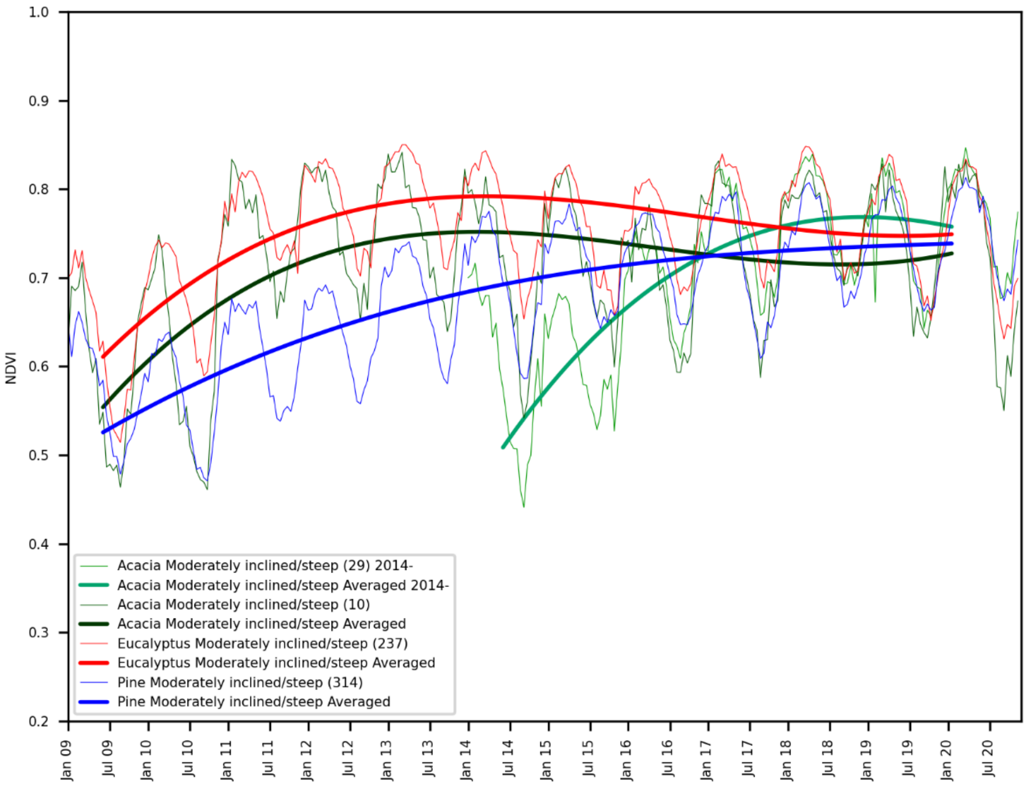



Eucalyptus (399)
Eucalyptus Averaged
NDVI
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
Jan 09
Jul 09
Jan 10
Jul 10
Jan 11
Jul 11
Jan 12
Jul 12
Jan 13
Jul 13
Jan 14
Jul 14
Jan 15
Jul 15
Jan 16
Jul 16
Jan 17
Jul 17
Jan 18
Jul 18
Jan 19
Jul 19
Jan 20
Jul 20

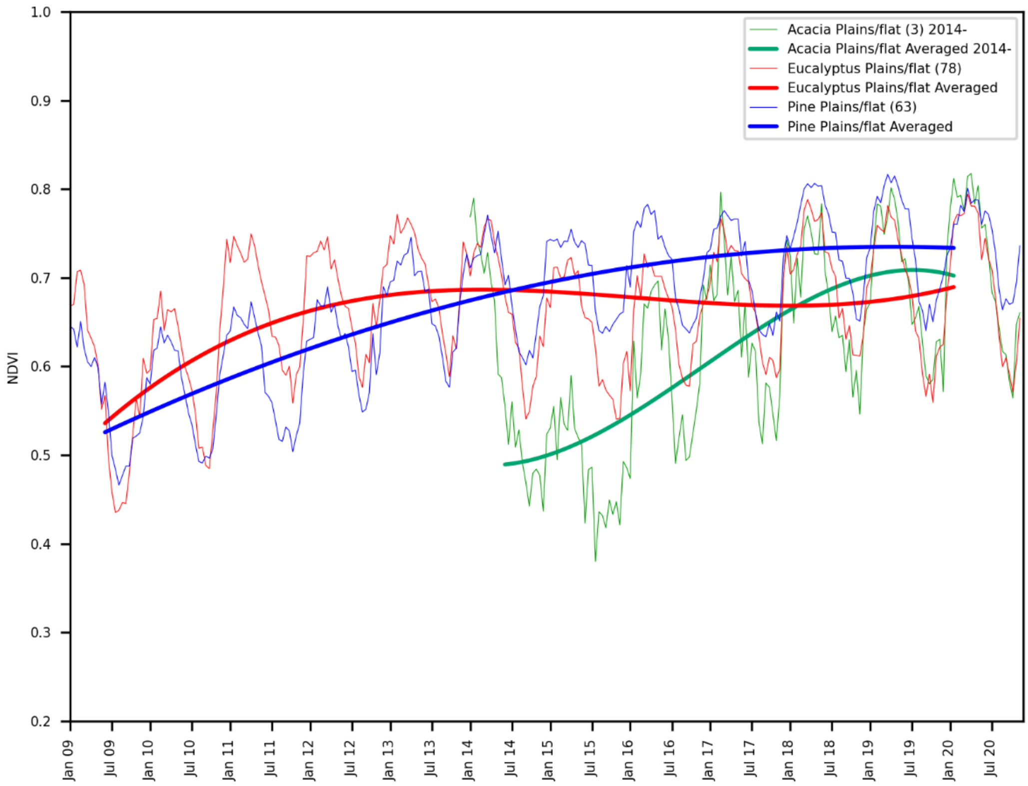

Pine (463)
Pine Averaged
NDVI
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
Jan 09
Jul 09
Jan 10
Jul 10
Jan 11
Jul 11
Jan 12
Jul 12
Jan 13
Jul 13
Jan 14
Jul 14
Jan 15
Jul 15
Jan 16
Jul 16
Jan 17
Jul 17
Jan 18
Jul 18
Jan 19
Jul 19
Jan 20
Jul 20

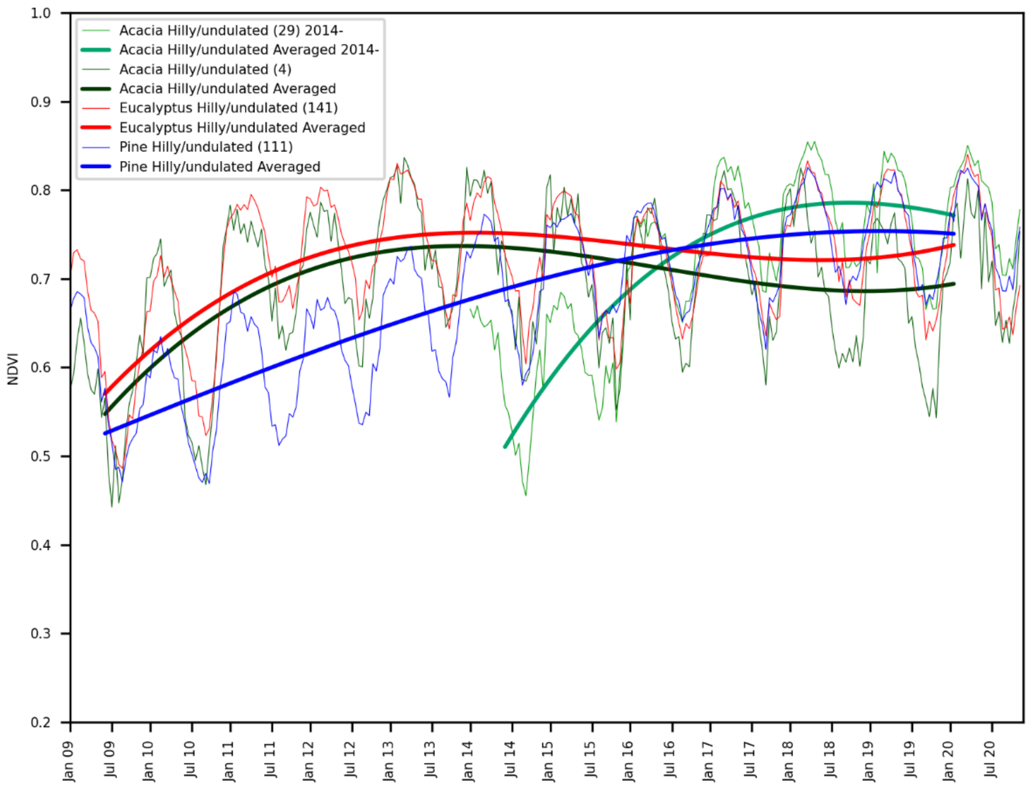
Compared to Slope Gradient

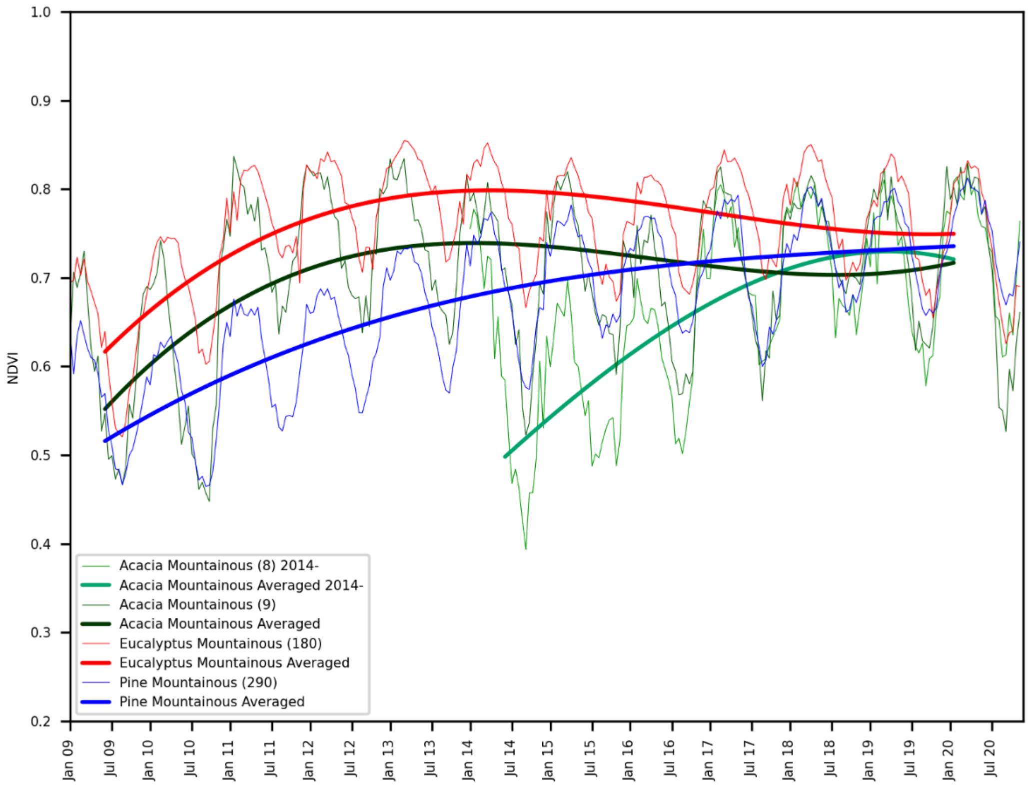




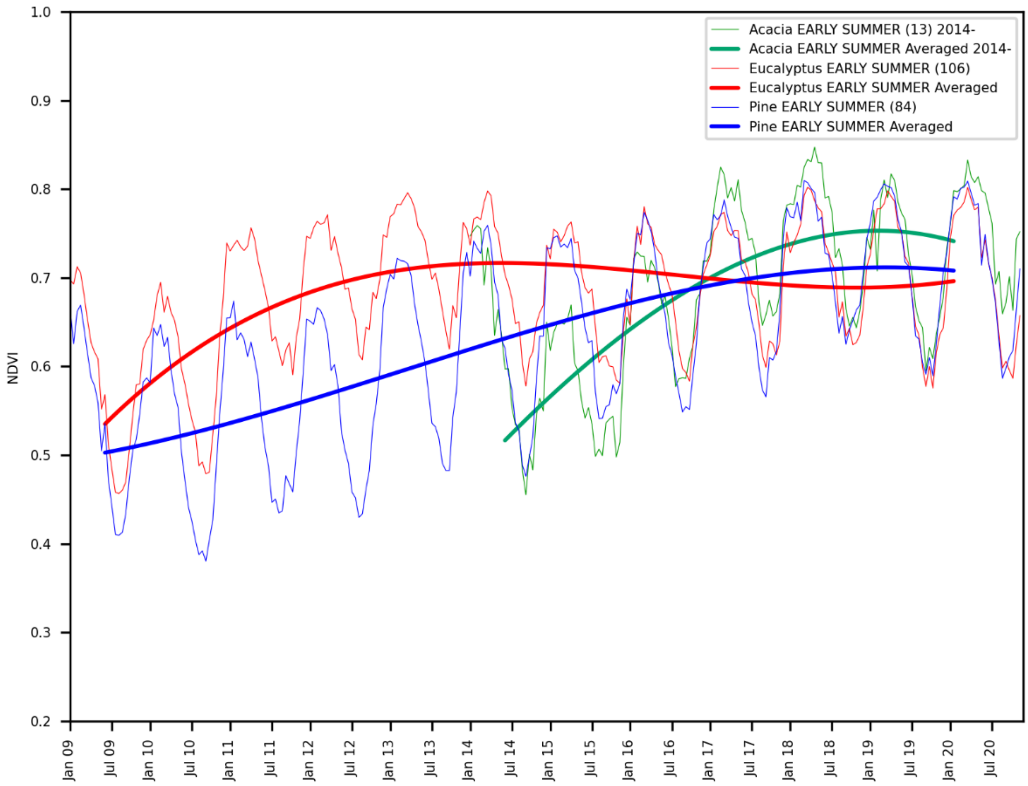
Compared to Terrain Morphology

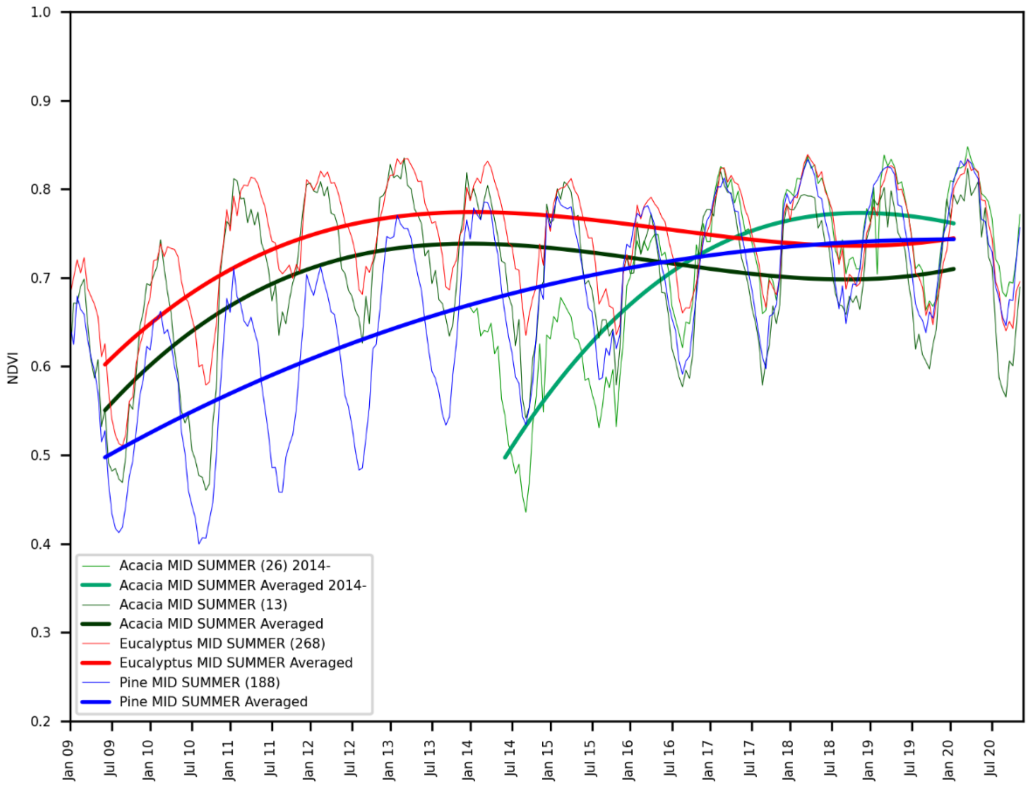


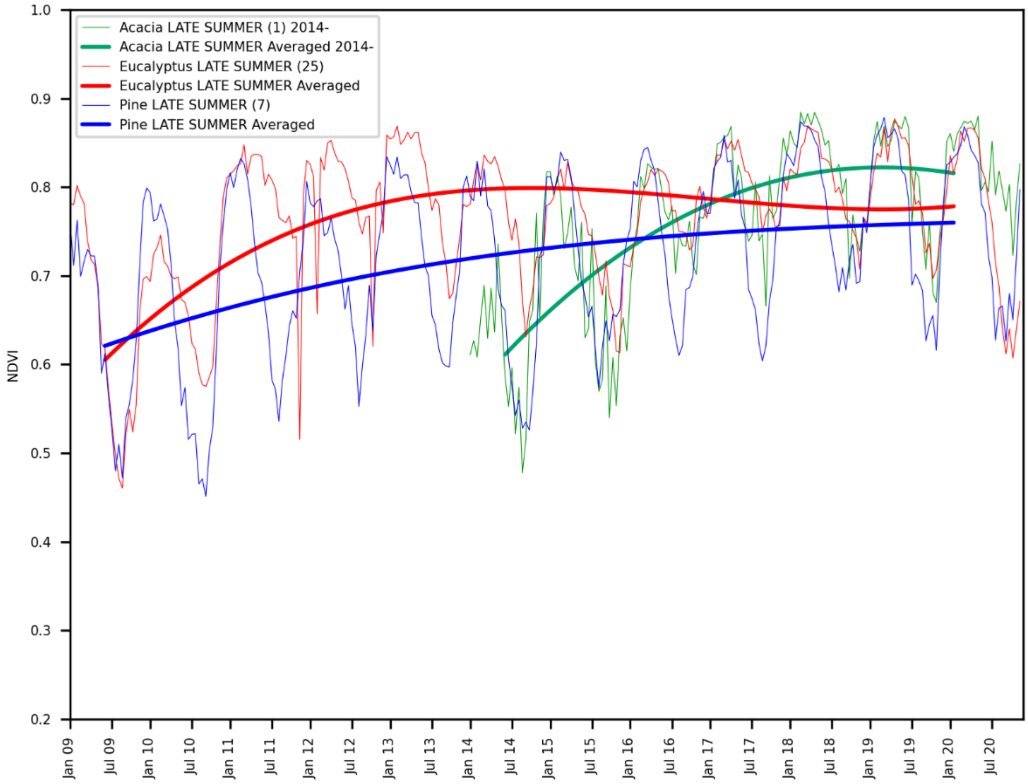



1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
NDVI
Acacia Mountainous (8) 2014-
Acacia Mountainous Averaged 2014-
Acacia Mountainous (9)
Acacia Mountainous Averaged
Eucalyptus Mountainous (180)
Eucalyptus Mountainous Averaged
Pine Mountainous (290)
Pine Mountainous Averaged
Jan 09
Jul 09
Jan 10
Jul 10
Jan 11
Jul 11
Jan 12
Jul 12
Jan 13
Jul 13
Jan 14
Jul 14
Jan 15
Jul 15
Jan 16
Jul 16
Jan 17
Jul 17
Jan 18
Jul 18
Jan 19
Jul 19
Jan 20
Jul 20

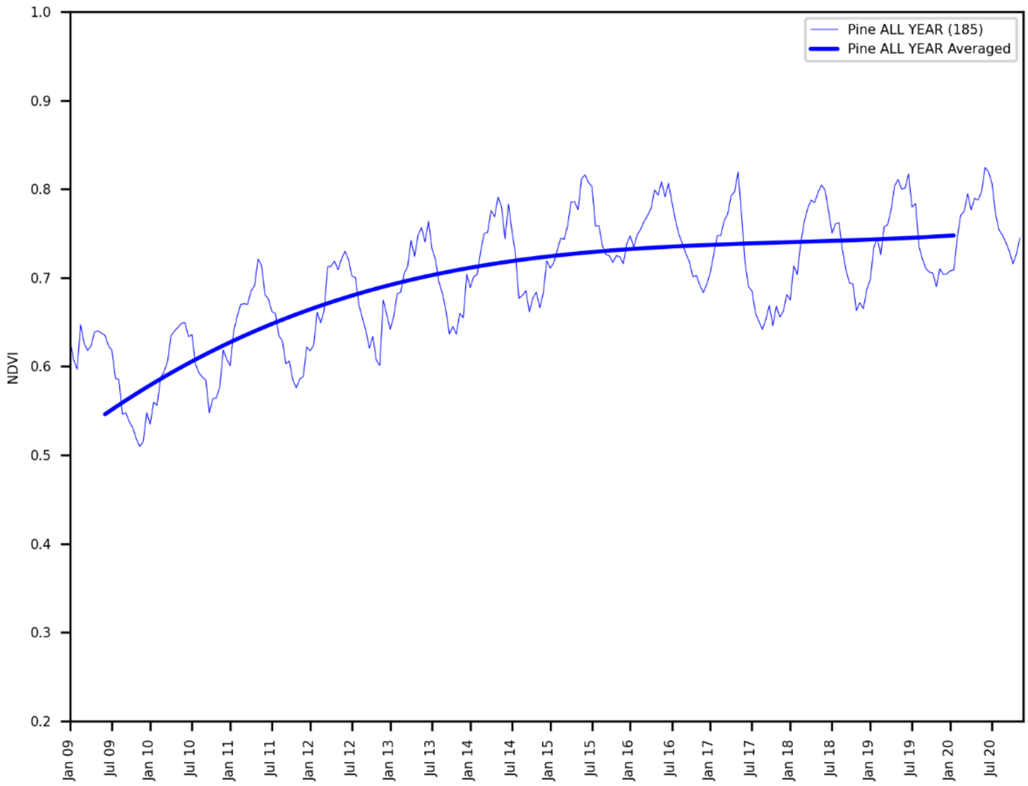
Compared to Rainfall Seasonality





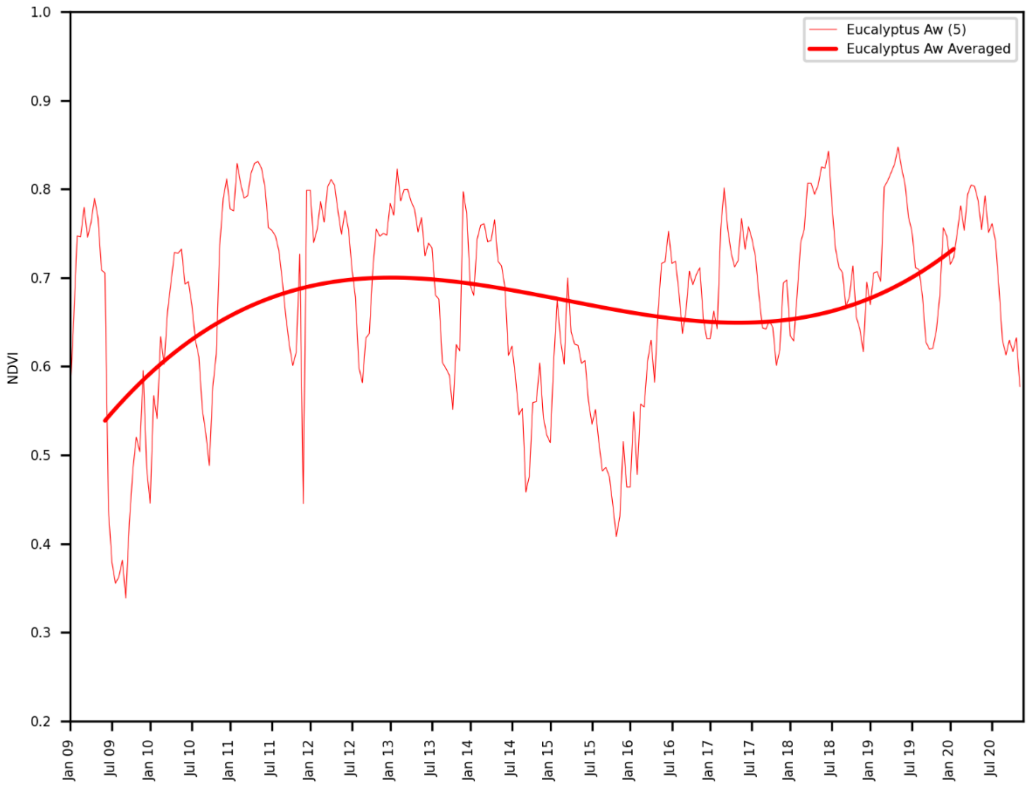


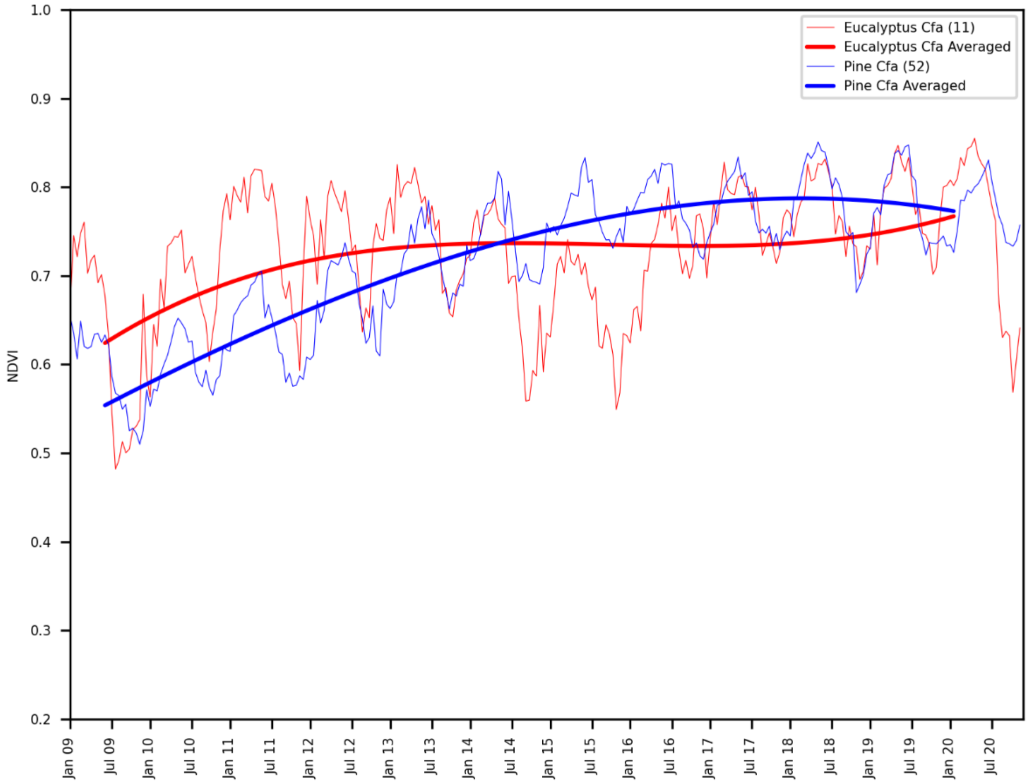
Acacia LATE SUMMER (1) 2014-
Acacia LATE SUMMER Averaged 2014-
Eucalyptus LATE SUMMER (25)
Eucalyptus LATE SUMMER Averaged
Pine LATE SUMMER (7)
Pine LATE SUMMER Averaged
NDVI

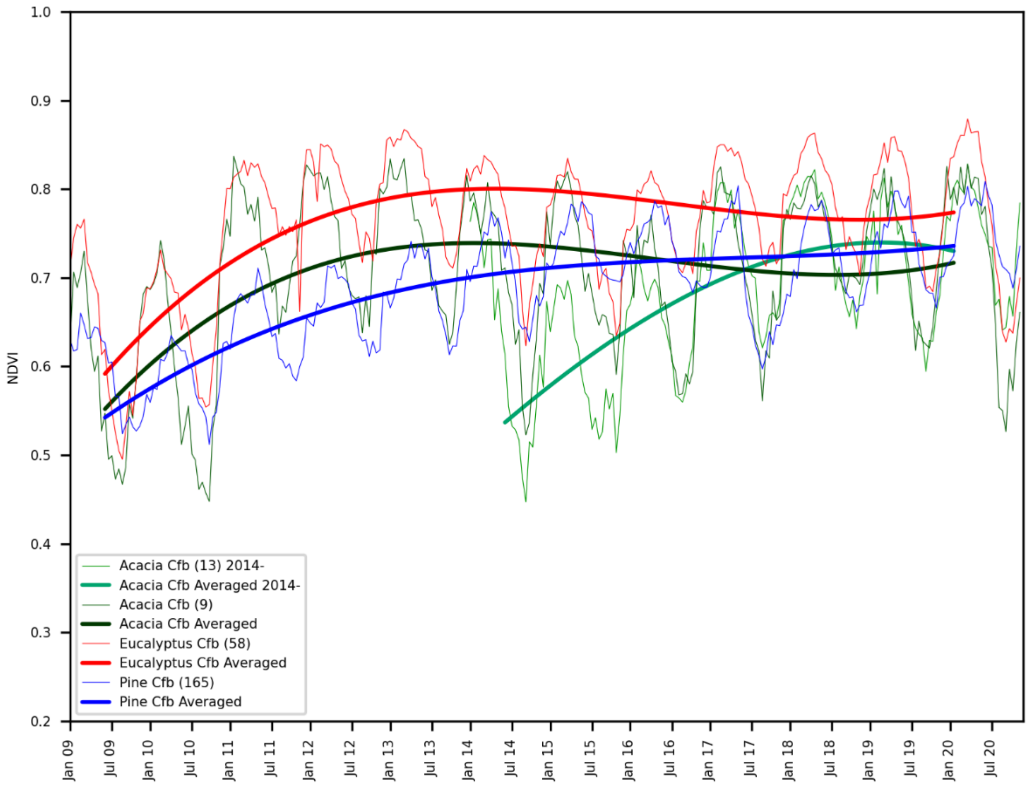


Pine ALL YEAR (185)
Pine ALL YEAR Averaged
NDVI

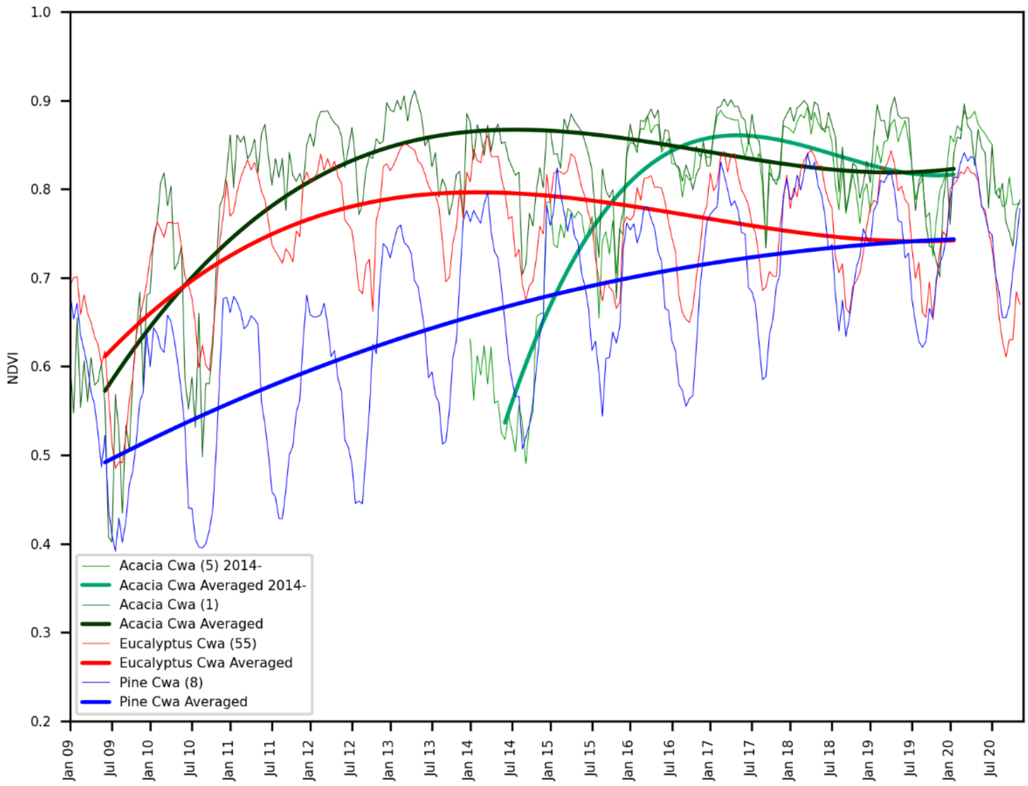
Compared to Climate Zones



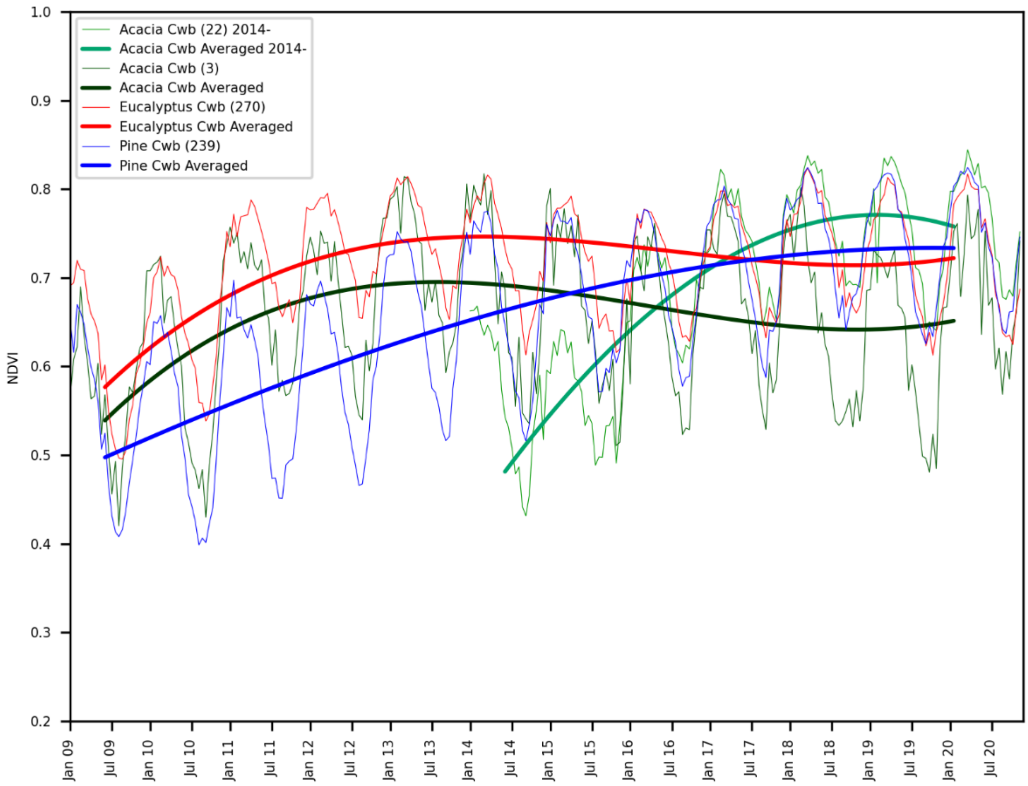




NDVI
Acacia Cfb (13) 2014-
Acacia Cfb Averaged 2014-
Acacia Cfb (9)
Acacia Cfb Averaged
Eucalyptus Cfb (58)
Eucalyptus Cfb Averaged
Pine Cfb (165)
Pine Cfb Averaged

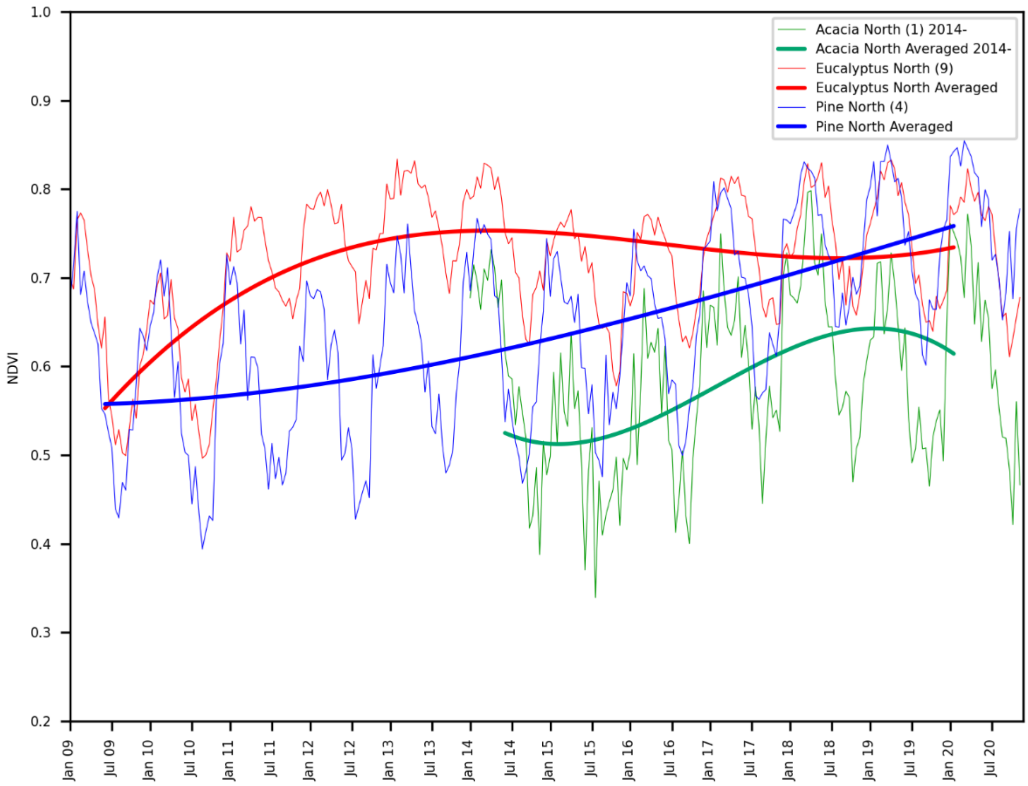


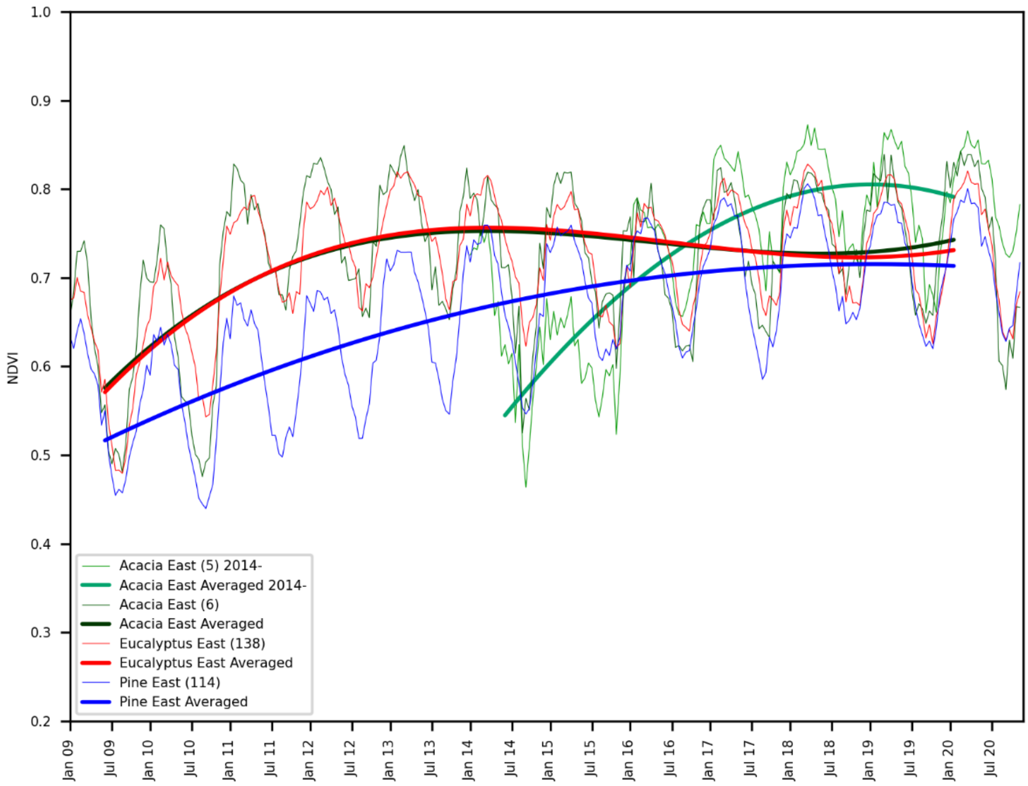
NDVI
Acacia Cwa (5) 2014-
Acacia Cwa Averaged 2014-
Acacia Cwa (1)
Acacia Cwa Averaged
Eucalyptus Cwa (55)
Eucalyptus Cwa Averaged
Pine Cwa (8)
Pine Cwa Averaged

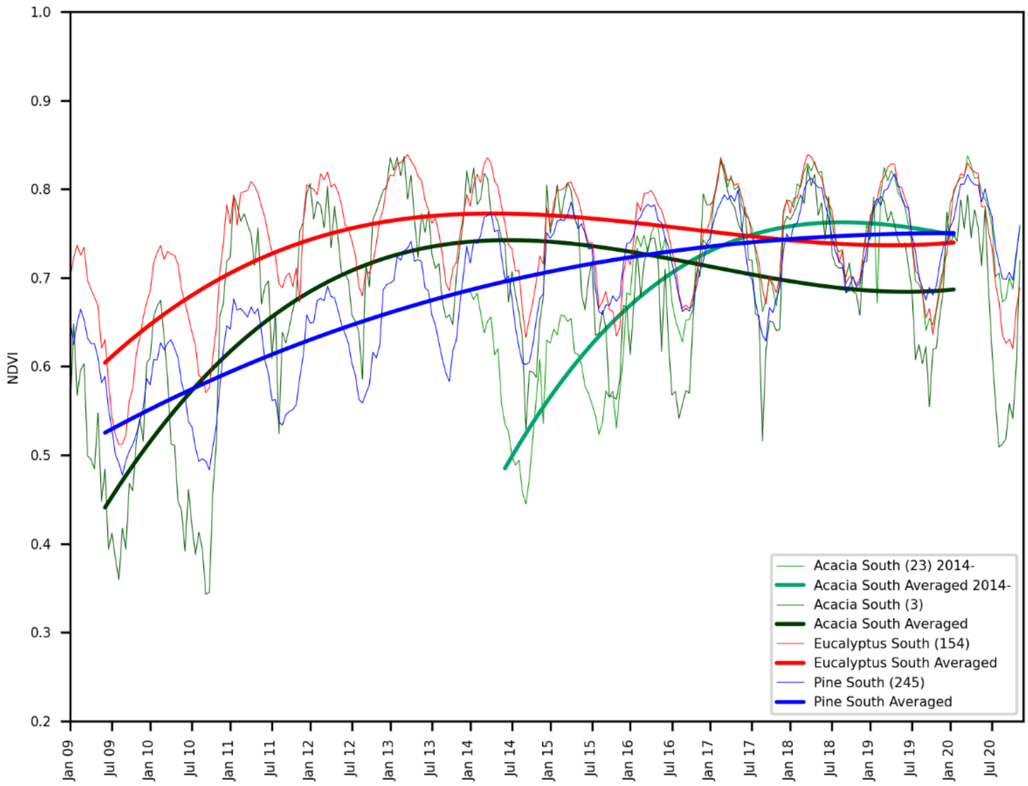


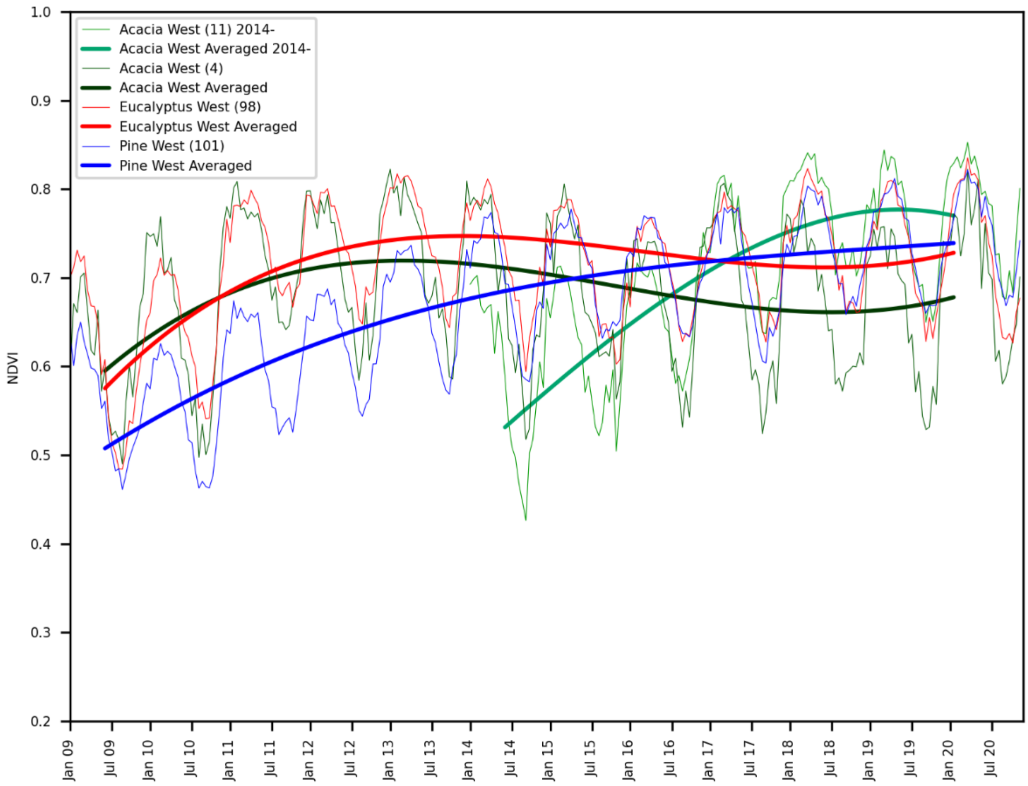
Acacia Cwb (22) 2014-
Acacia Cwb Averaged 2014-
Acacia Cwb (3)
Acacia Cwb Averaged
Eucalyptus Cwb (270)
Eucalyptus Cwb Averaged
Pine Cwb (239)
Pine Cwb Averaged
NDVI
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
Jan 09
Jul 09
Jan 10
Jul 10
Jan 11
Jul 11
Jan 12
Jul 12
Jan 13
Jul 13
Jan 14
Jul 14
Jan 15
Jul 15
Jan 16
Jul 16
Jan 17
Jul 17
Jan 18
Jul 18
Jan 19
Jul 19
Jan 20
Jul 20

Compared to Slope Aspect

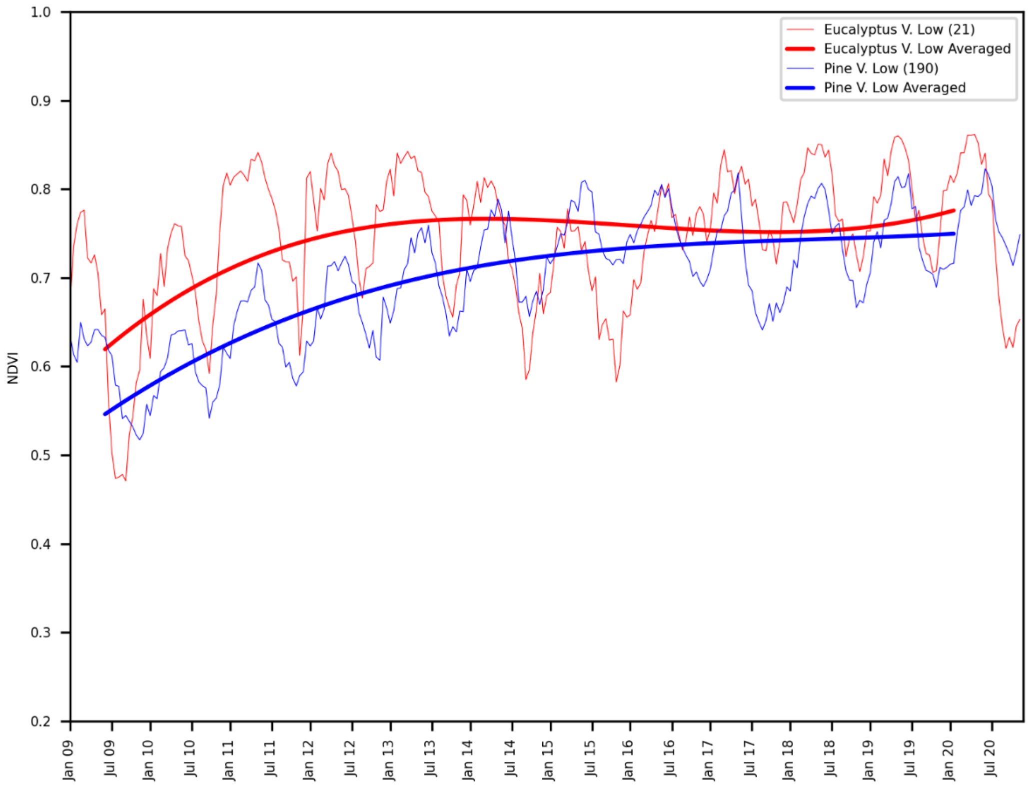


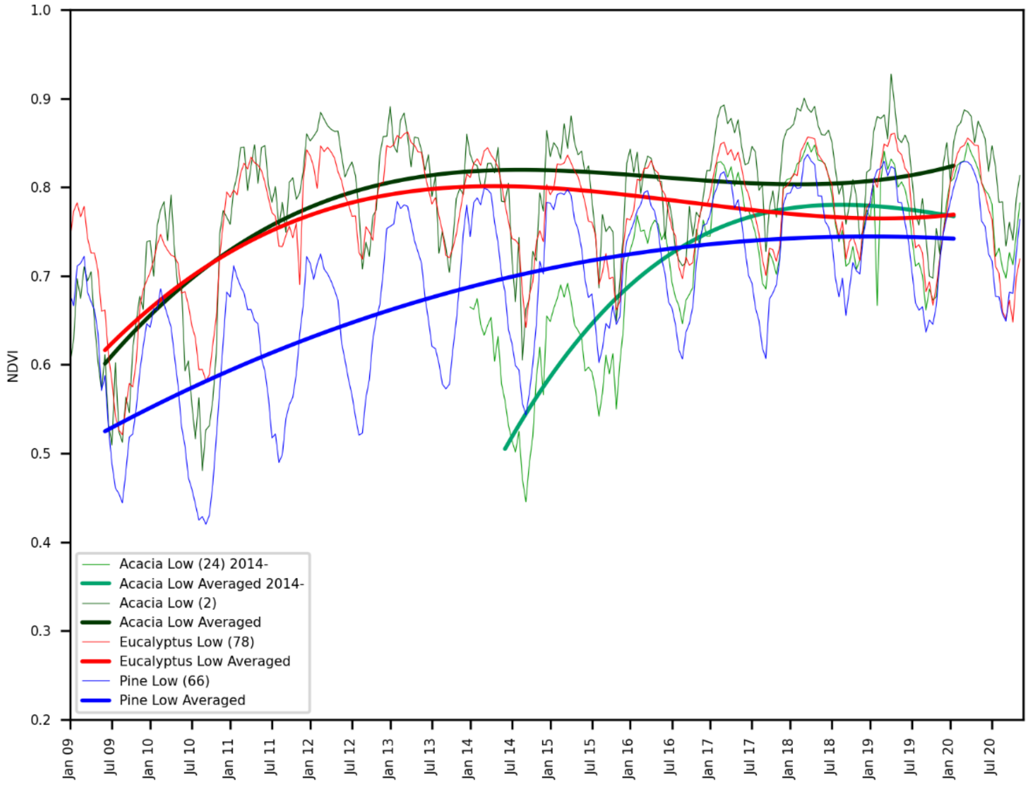


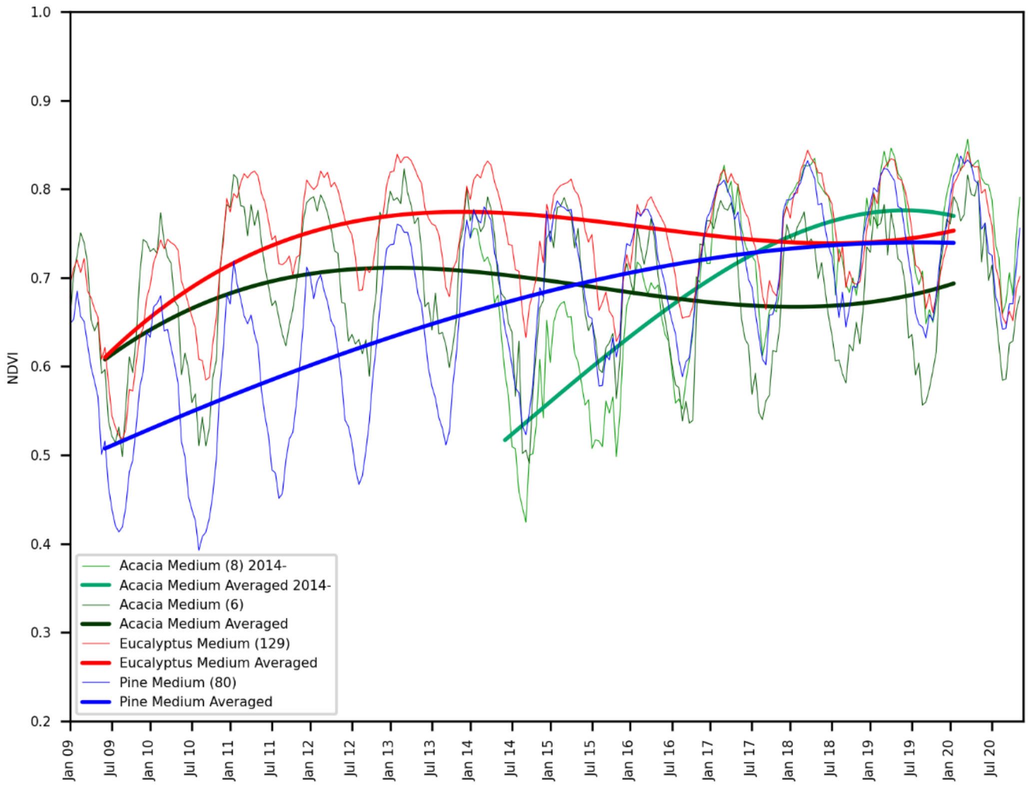

1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
NDVI
Jan 09
Jul 09
Jan 10
Jul 10
Jan 11
Jul 11
Jan 12
Jul 12
Jan 13
Jul 13
Jan 14
Jul 14
Jan 15
Jul 15
Jan 16
Jul 16
Jan 17
Jul 17
Jan 18
Jul 18
Jan 19
Jul 19
Jan 20
Jul 20
Acacia South (23) 2014-
Acacia South Averaged 2014-
Acacia South (3)
Acacia South Averaged
Eucalyptus South (154)
Eucalyptus South Averaged
Pine South (245)
Pine South Averaged


1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
NDVI
Jan 09
Jul 09
Jan 10
Jul 10
Jan 11
Jul 11
Jan 12
Jul 12
Jan 13
Jul 13
Jan 14
Jul 14
Jan 15
Jul 15
Jan 16
Jul 16
Jan 17
Jul 17
Jan 18
Jul 18
Jan 19
Jul 19
Jan 20
Jul 20
Acacia West (11) 2014-
Acacia West Averaged 2014-
Acacia West (4)
Acacia West Averaged
Eucalyptus West (98)
Eucalyptus West Averaged
Pine West (101)
Pine West Averaged

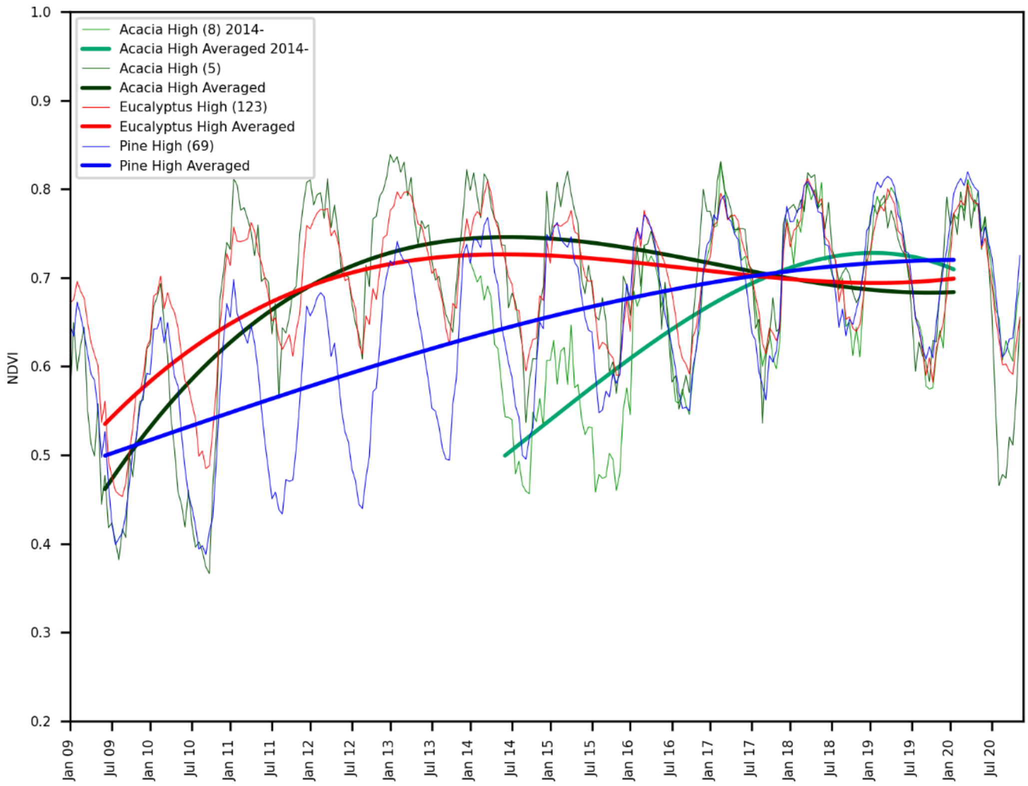
Compared to Annual Solar Radiation



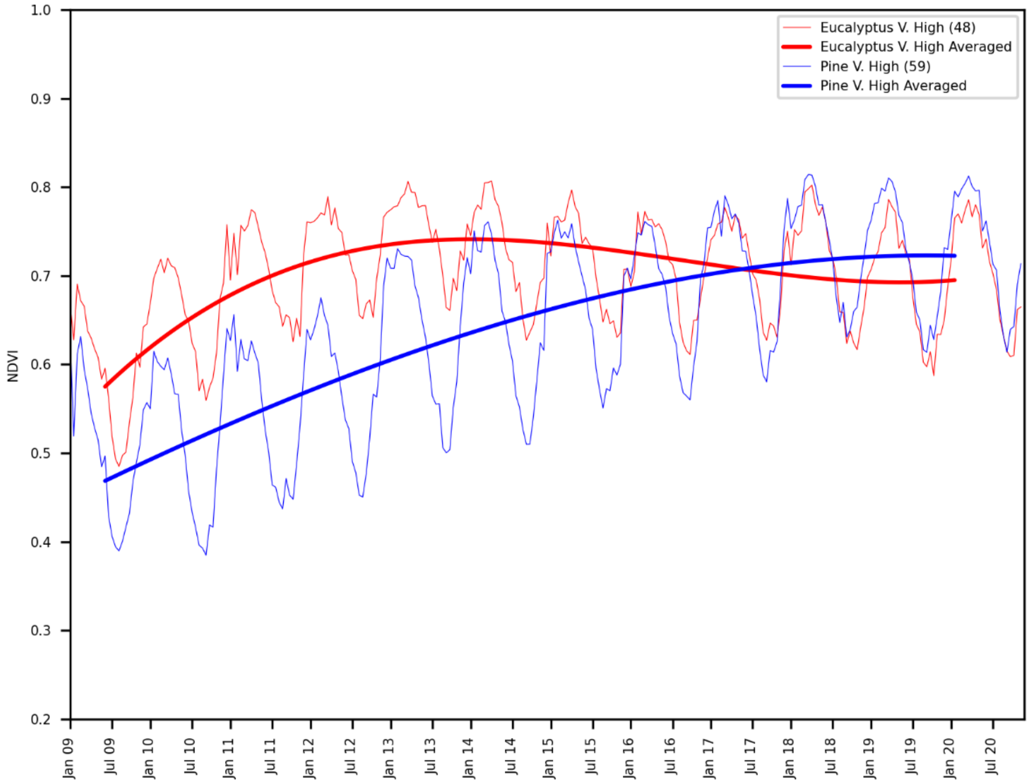




NDVI
Acacia Medium (8) 2014-
Acacia Medium Averaged 2014-
Acacia Medium (6)
Acacia Medium Averaged
Eucalyptus Medium (129)
Eucalyptus Medium Averaged
Pine Medium (80)
Pine Medium Averaged
Jan 09
Jul 09
Jan 10
Jul 10
Jan 11
Jul 11
Jan 12
Jul 12
Jan 13
Jul 13
Jan 14
Jul 14
Jan 15
Jul 15
Jan 16
Jul 16
Jan 17
Jul 17
Jan 18
Jul 18
Jan 19
Jul 19
Jan 20
Jul 20

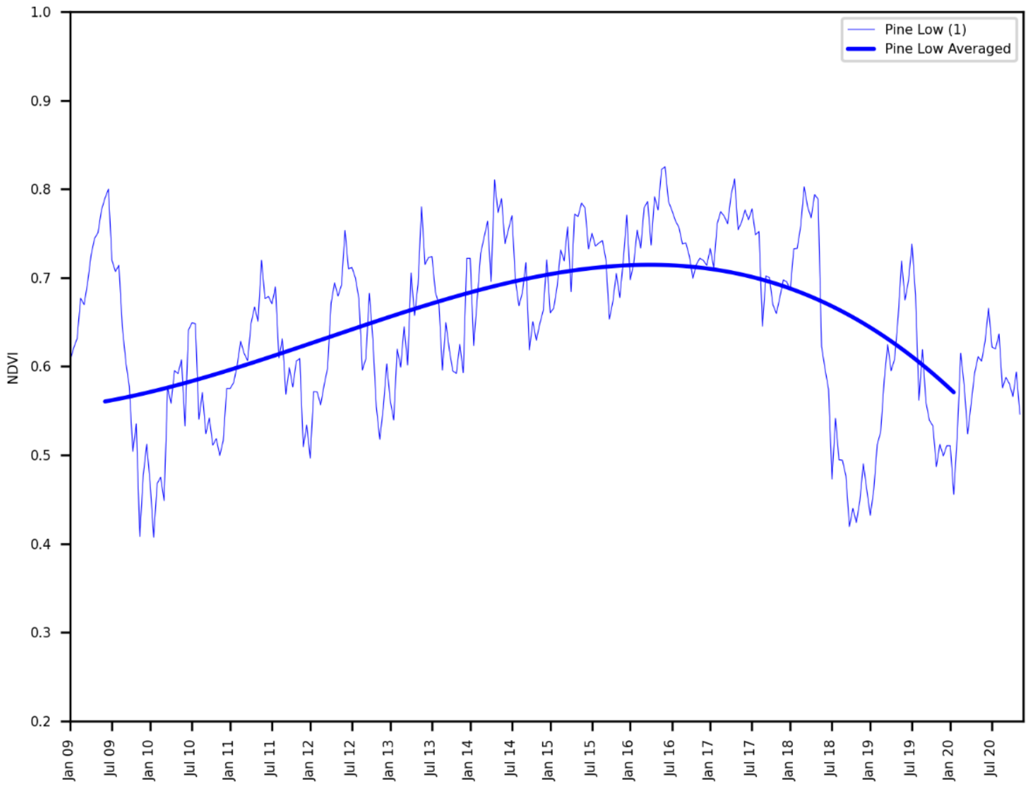


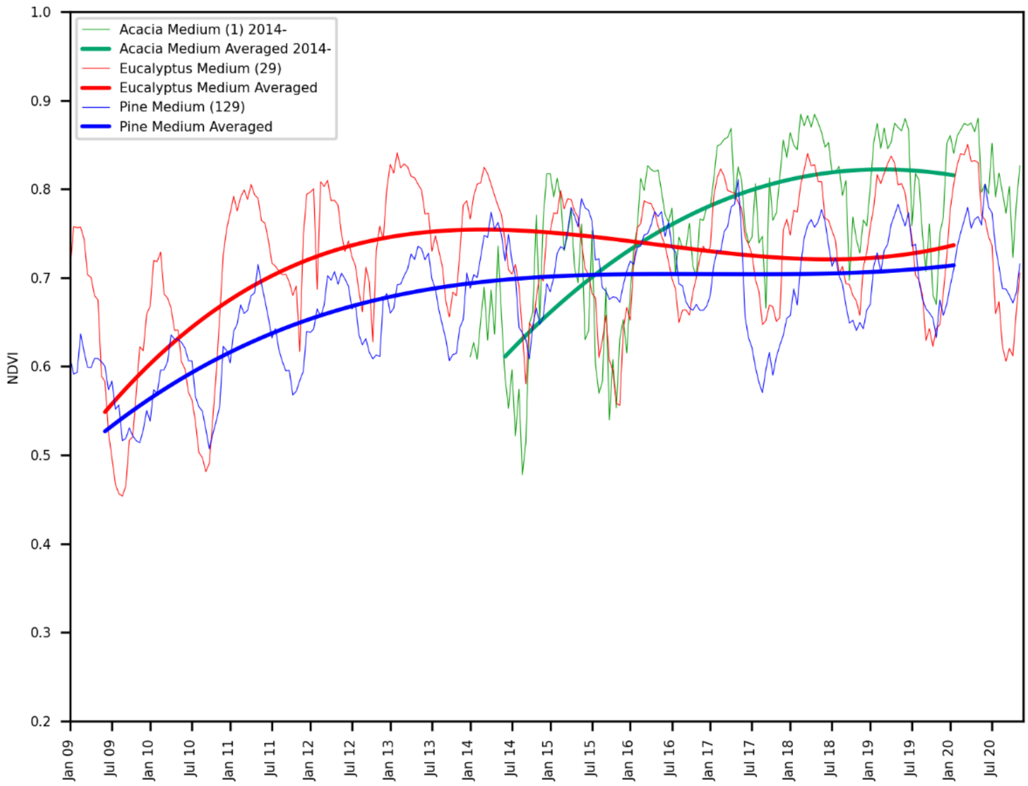
NDVI
Acacia High (8) 2014-
Acacia High Averaged 2014-
Acacia High (5)
Acacia High Averaged
Eucalyptus High (123)
Eucalyptus High Averaged
Pine High (69)
Pine High Averaged
Jan 09
Jul 09
Jan 10
Jul 10
Jan 11
Jul 11
Jan 12
Jul 12
Jan 13
Jul 13
Jan 14
Jul 14
Jan 15
Jul 15
Jan 16
Jul 16
Jan 17
Jul 17
Jan 18
Jul 18
Jan 19
Jul 19
Jan 20
Jul 20

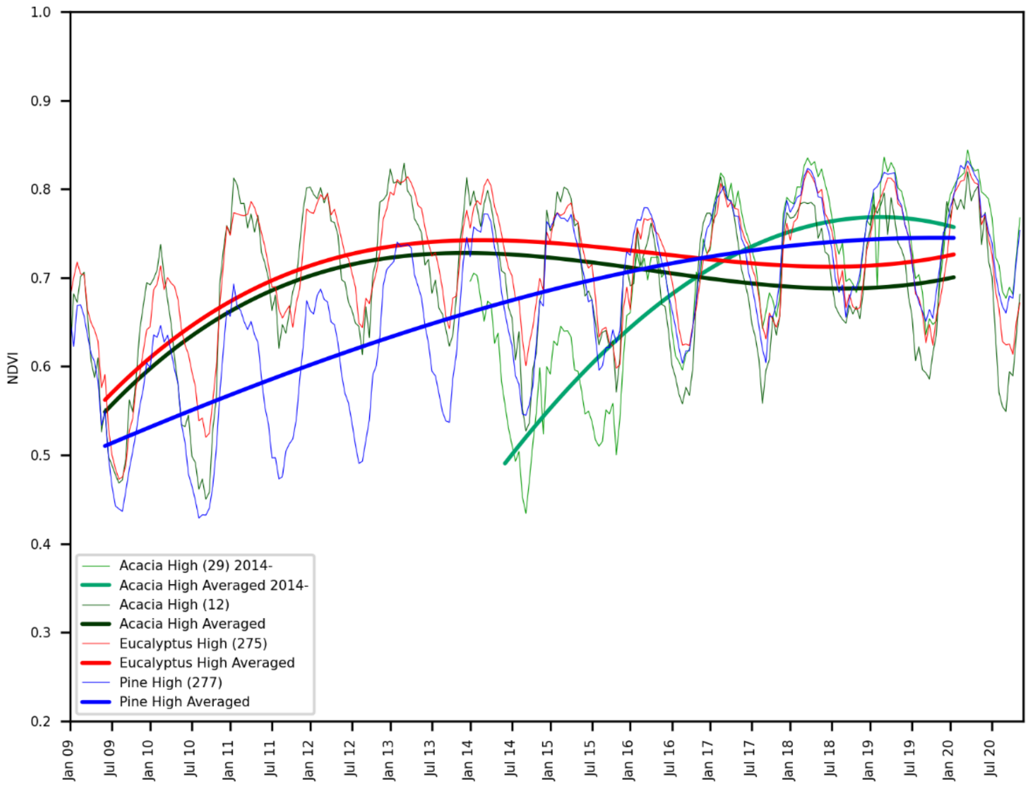


Eucalyptus V. High (48)
Eucalyptus V. High Averaged
Pine V. High (59)
Pine V. High Averaged
NDVI
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
Jan 09
Jul 09
Jan 10
Jul 10
Jan 11
Jul 11
Jan 12
Jul 12
Jan 13
Jul 13
Jan 14
Jul 14
Jan 15
Jul 15
Jan 16
Jul 16
Jan 17
Jul 17
Jan 18
Jul 18
Jan 19
Jul 19
Jan 20
Jul 20

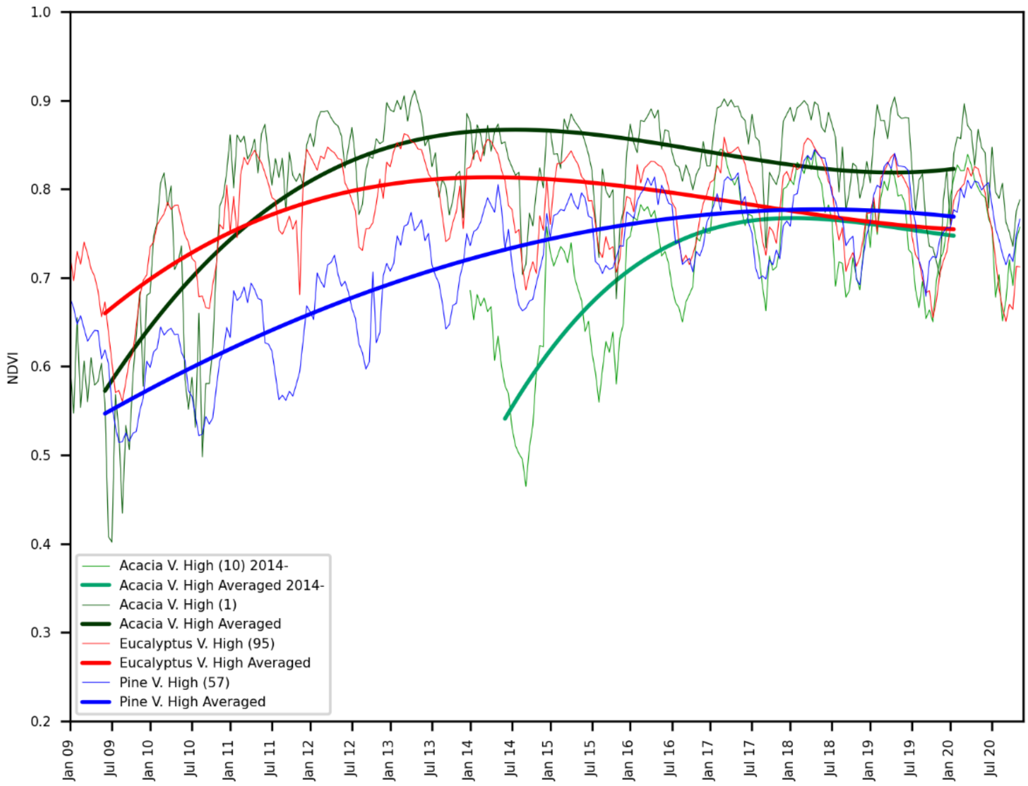
Compared to Mean Annual Rainfall







NDVI
Acacia High (29) 2014-
Acacia High Averaged 2014-
Acacia High (12)
Acacia High Averaged
Eucalyptus High (275)
Eucalyptus High Averaged
Pine High (277)
Pine High Averaged
Jan 09
Jul 09
Jan 10
Jul 10
Jan 11
Jul 11
Jan 12
Jul 12
Jan 13
Jul 13
Jan 14
Jul 14
Jan 15
Jul 15
Jan 16
Jul 16
Jan 17
Jul 17
Jan 18
Jul 18
Jan 19
Jul 19
Jan 20
Jul 20



NDVI
Acacia V. High (10) 2014-
Acacia V. High Averaged 2014-
Acacia V. High (1)
Acacia V. High Averaged
Eucalyptus V. High (95)
Eucalyptus V. High Averaged
Pine V. High (57)
Pine V. High Averaged
Jan 09
Jul 09
Jan 10
Jul 10
Jan 11
Jul 11
Jan 12
Jul 12
Jan 13
Jul 13
Jan 14
Jul 14
Jan 15
Jul 15
Jan 16
Jul 16
Jan 17
Jul 17
Jan 18
Jul 18
Jan 19
Jul 19
Jan 20
Jul 20

APPENDIX III: ADDITIONAL WATER USE PROFILES

For the sake of brevity, some of the water use profiles generated in this project were excluded from the report and are provided here.

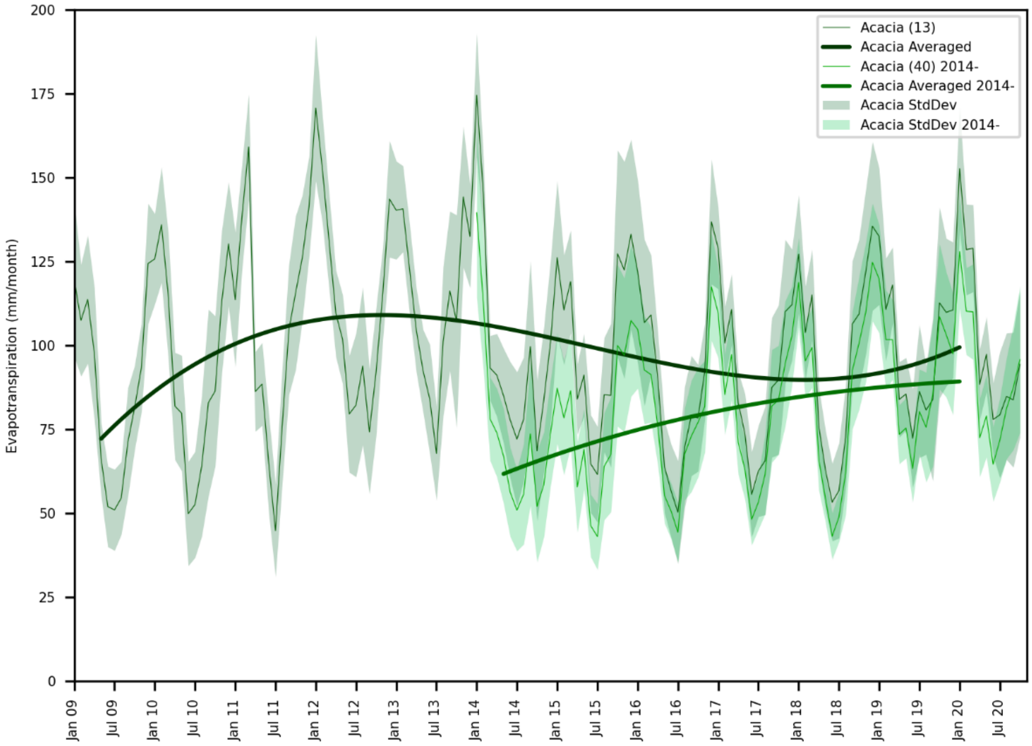


Figure A-1 Monthly water use over time for Acacia planted in 2009 and in 2014

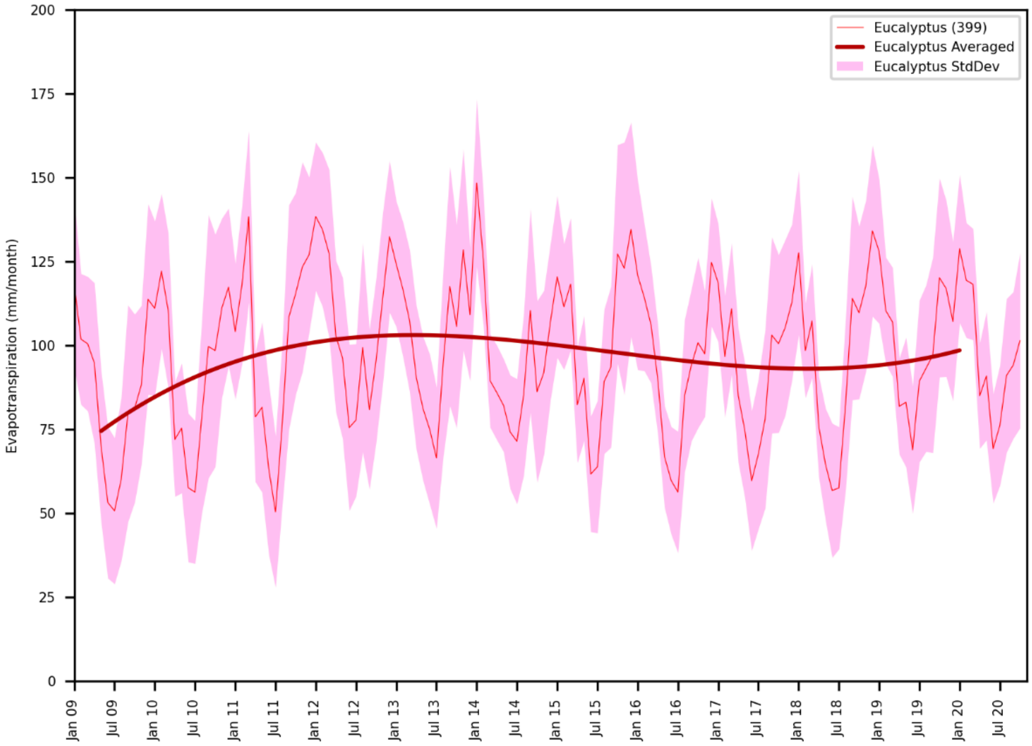


Figure A-2 Monthly water use over time for *Eucalyptus*

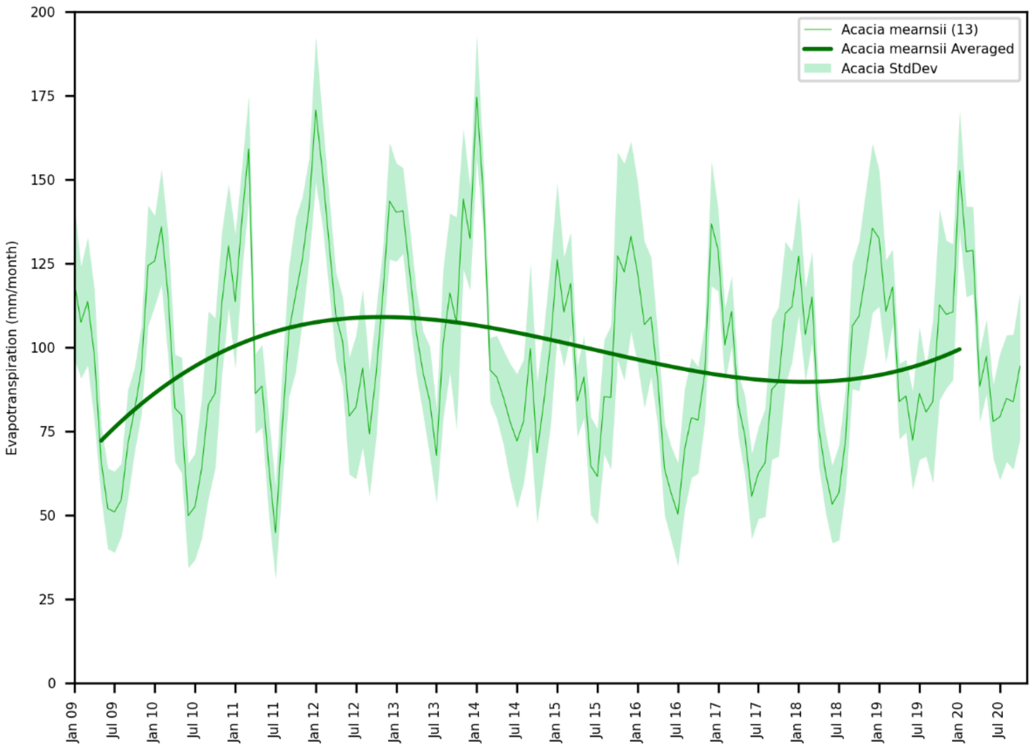


Figure A-3 ET over time for *A. mearnsii*

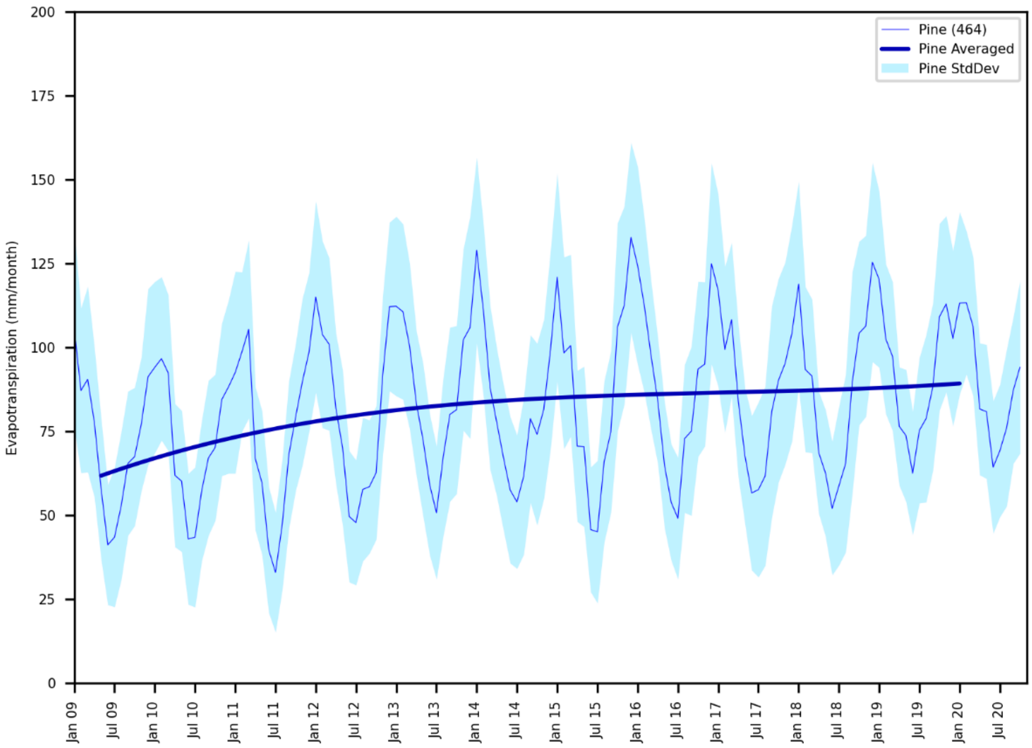


Figure A-4 Monthly water use over time for *Pinus*

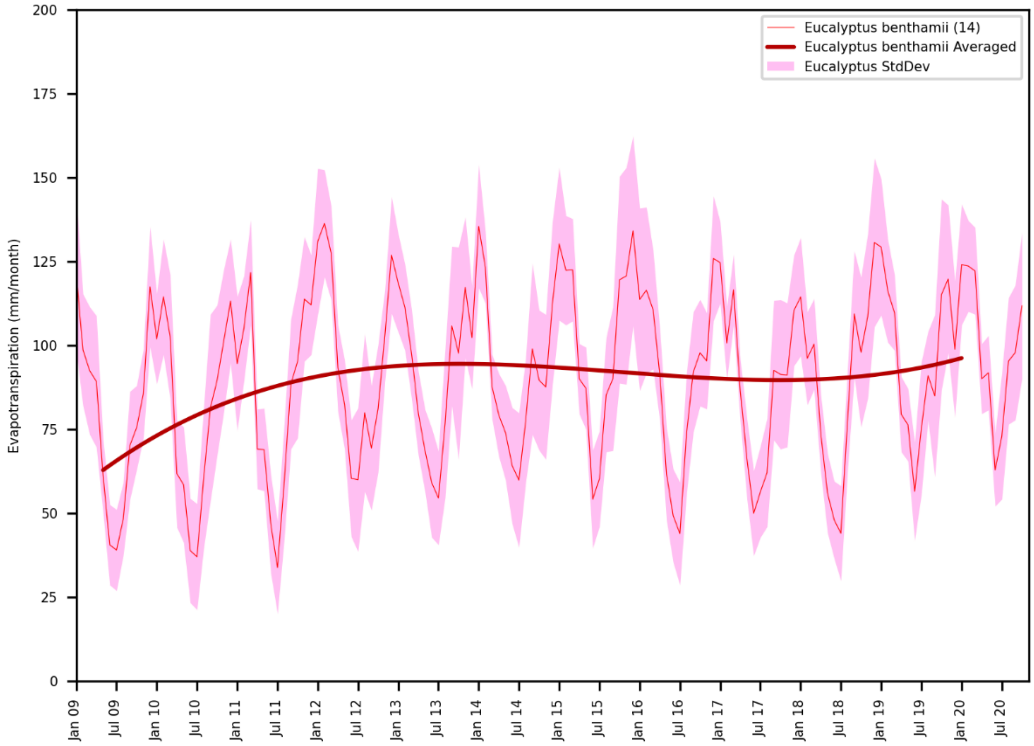


Figure A-5 ET over time for *E. benthamii*

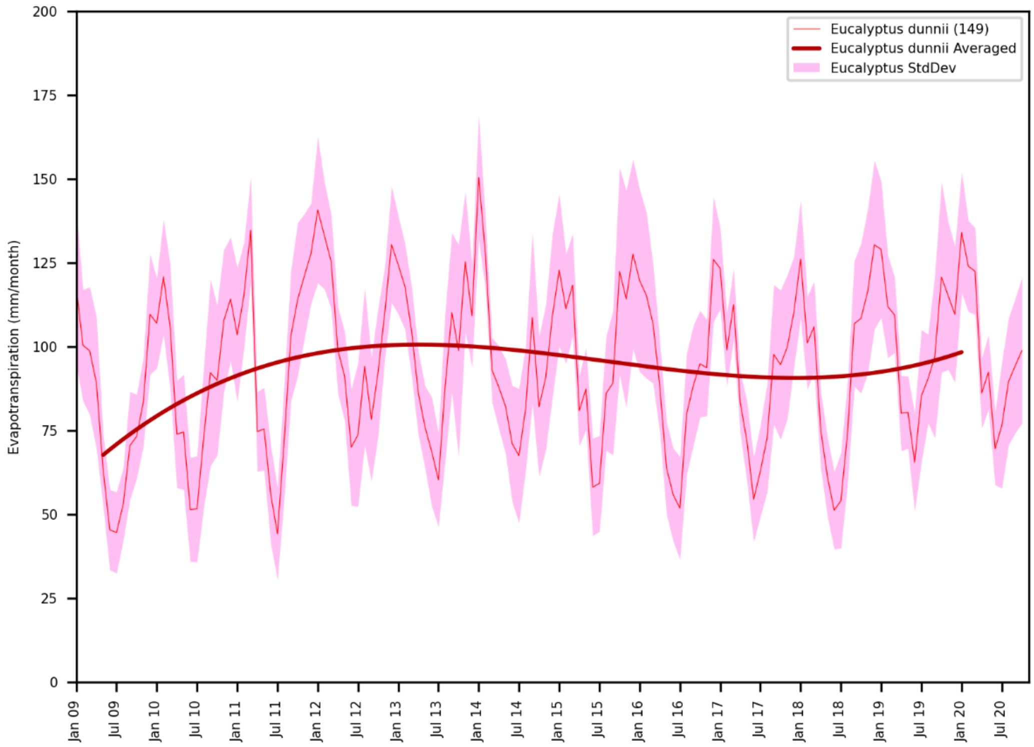


Figure A-6 ET over time for *E. dunnii*

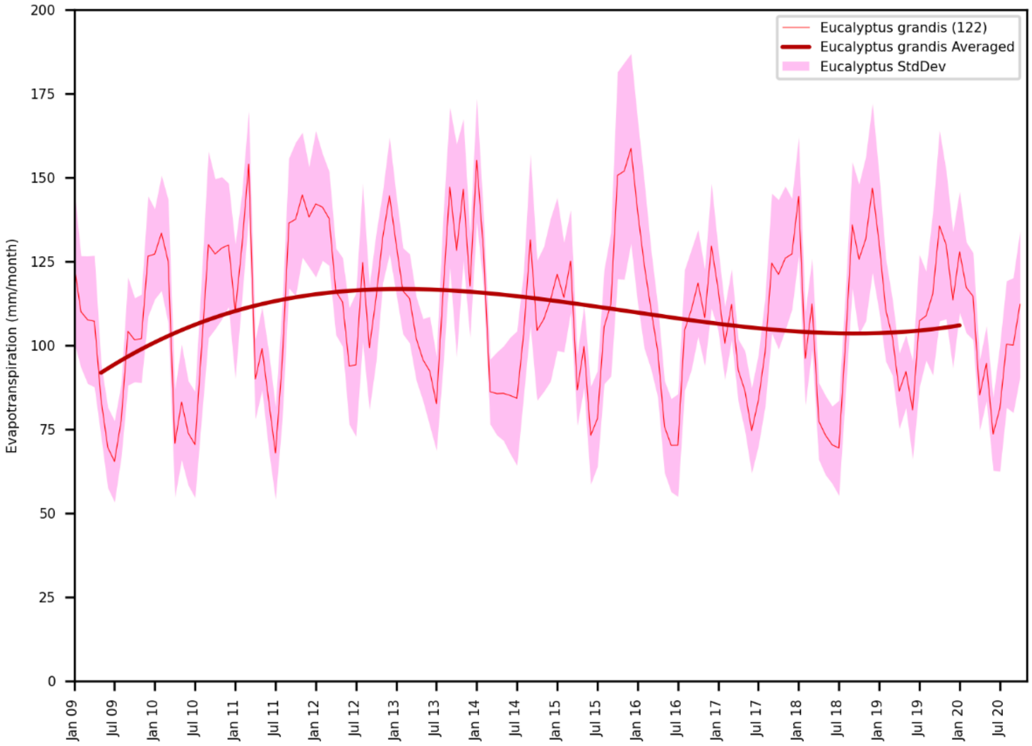


Figure A-7 ET over time for *E. grandis*

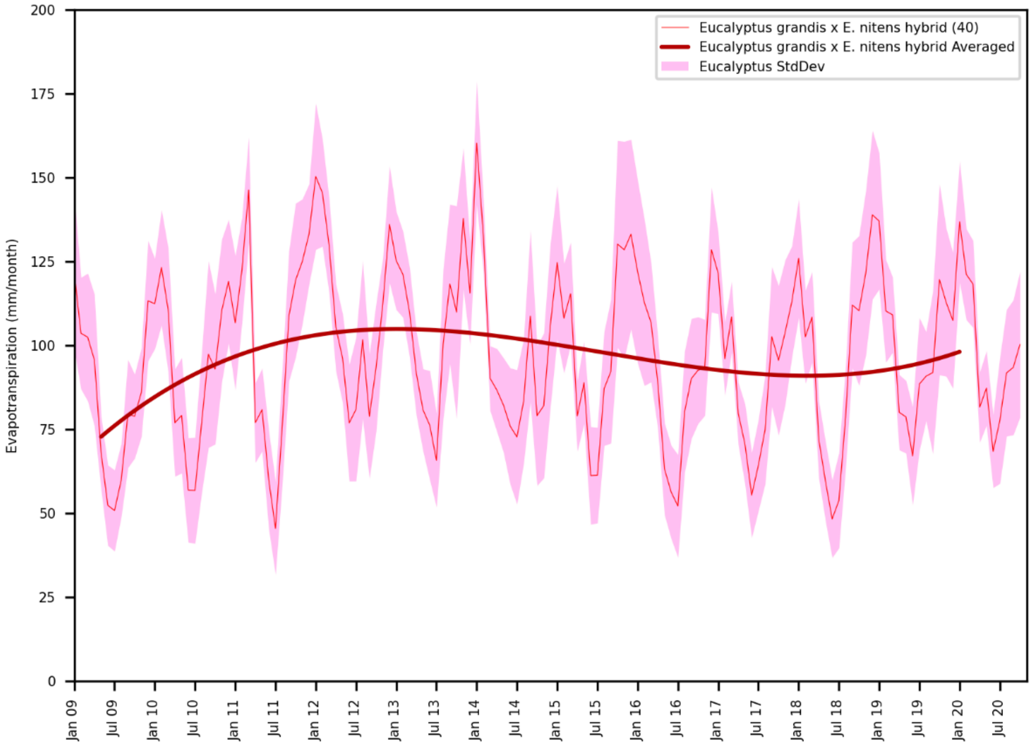


Figure A-8 ET over time for *E. grandis* x *E. nitens* hybrid

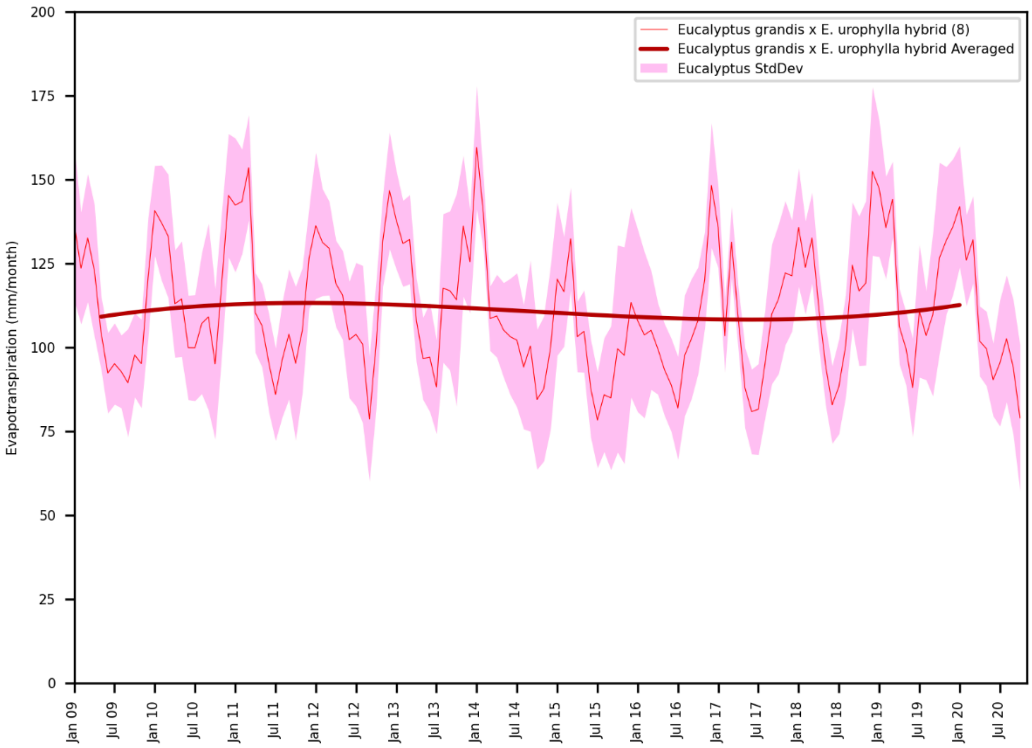


Figure A-9 ET over time for *E. grandis* x *E. urophylla*

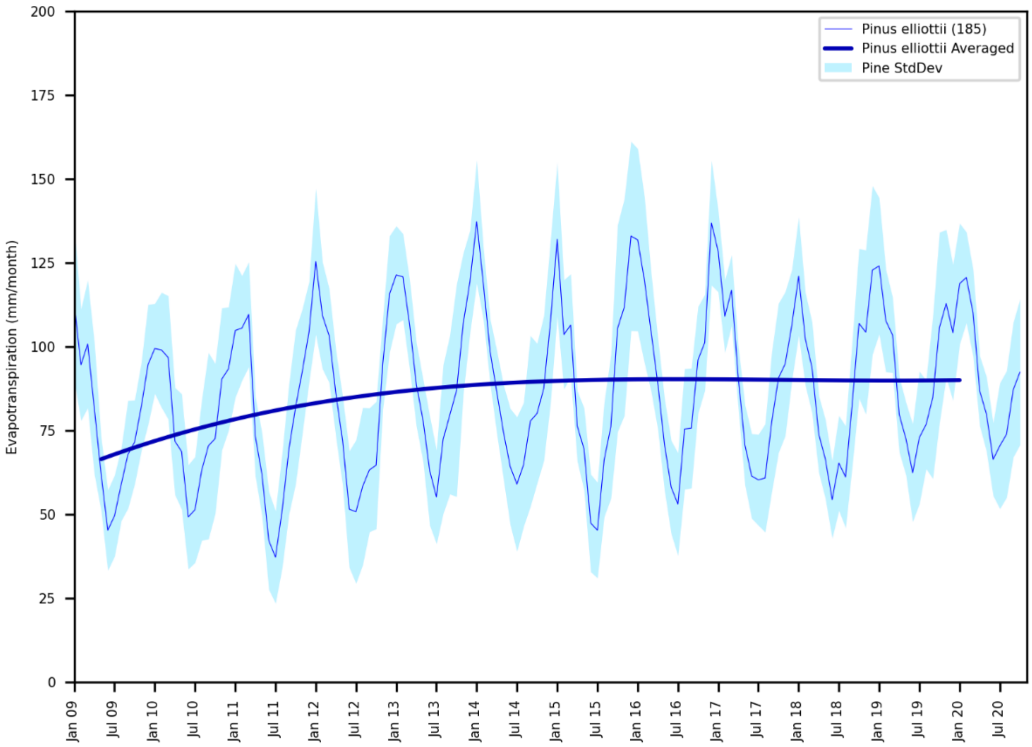


Figure A-10 ET over time for *P. elliottii*

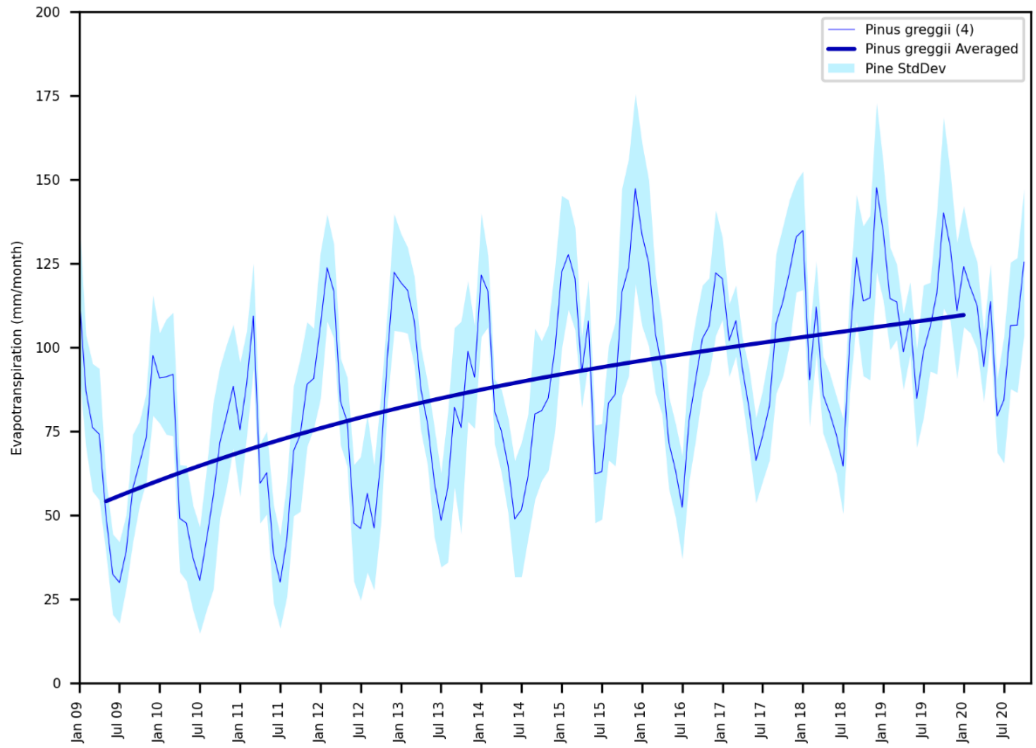


Figure A-11 ET over time for *P. greggii*

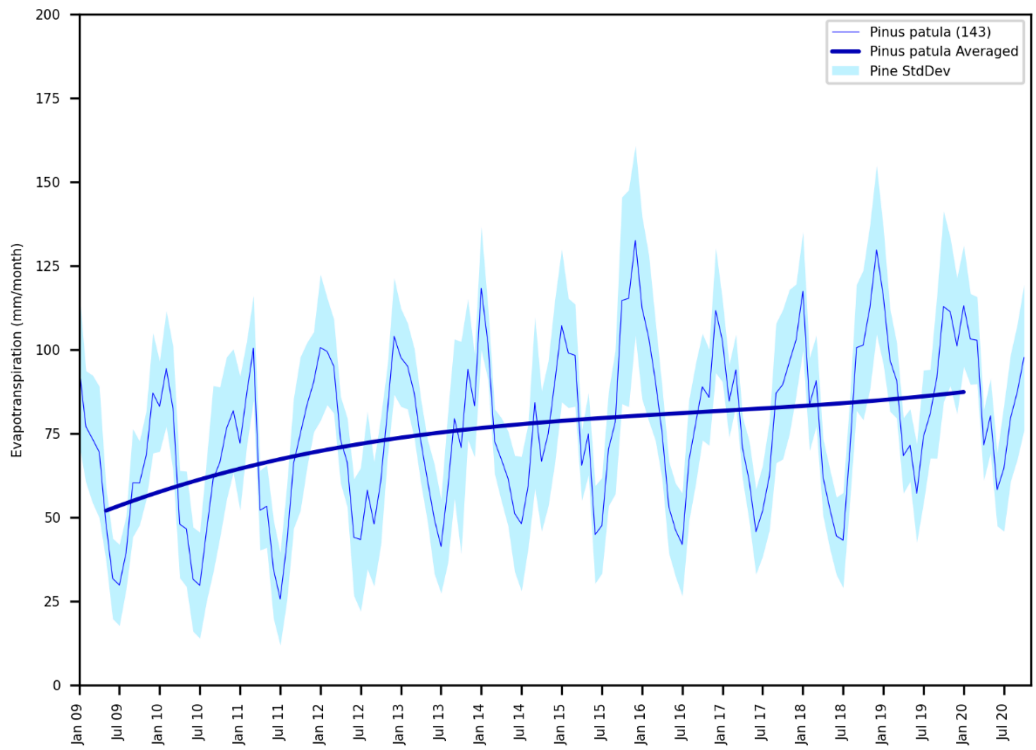


Figure A-12 ET over time for *P. patula*

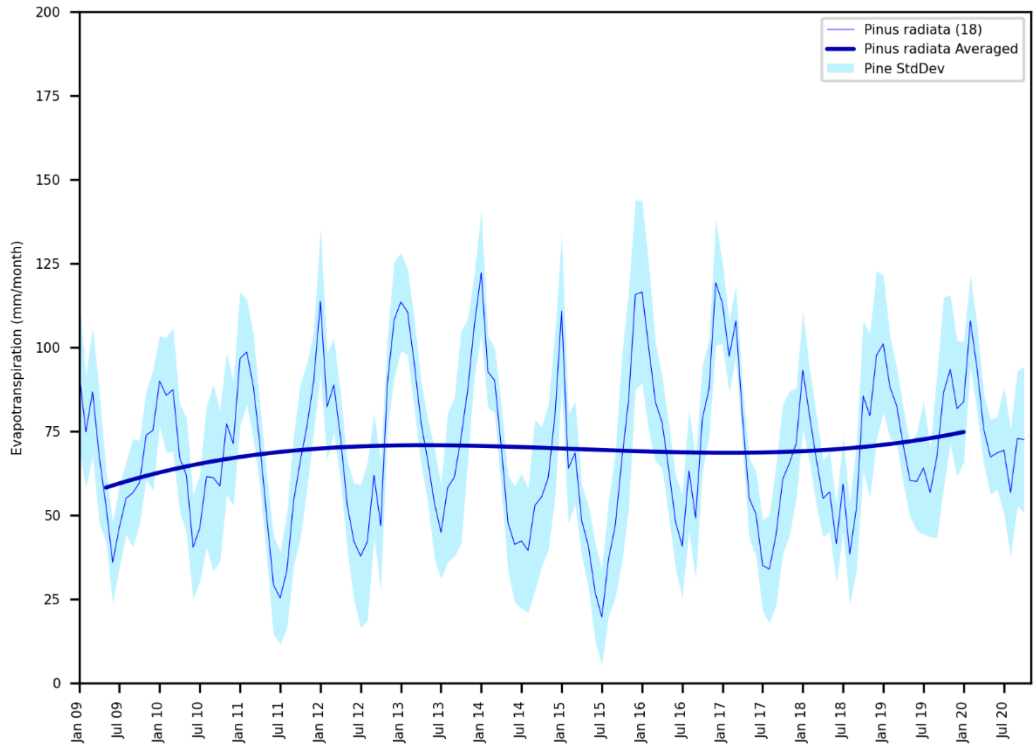


Figure A-13 ET over time for *P. radiata*

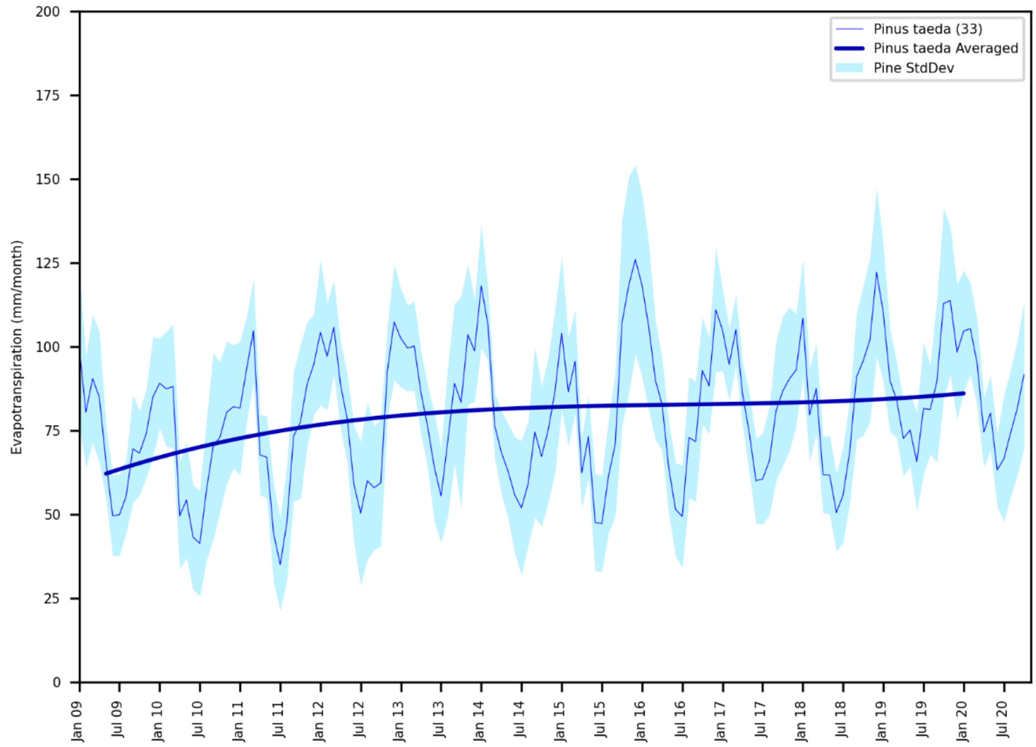


Figure A-14 ET over time for *P. taeda*

APPENDIX IV: CAPACITY BUILDING

The project made provision for two MSc students, namely Mr V Ndyafi and Miss C Higgs. Both students registered for MSc in Geoinformatics in February 2020. Mr Ndyafi's research is on automated forest (plantation and indigenous) mapping, while Miss Higgs' research is on the identification of plantation tree genus. Both studies made use of machine and/or deep learning techniques and were carried out within the context of regional (national) scale mapping. The following subsections provide brief overviews of the studies.

CALEY HIGGS

Miss Higgs, supervised by Prof Van Niekerk, completed her thesis in October 2021 and graduated in December 2021 with *cum laude*. Below is a summary of the thesis.

Forest inventories are constructed on a compartmental level and contain information such as forest age, species/genus, location, and extent. An up-to-date forest inventory is critical for monitoring harvests, assessing the production of timber, planning, maximising production, assessing water use, and assessing timber quality. On a national scale, forest inventories are used for monitoring the impact forests have on the climate and stream flow, assessing the contribution forests have on alleviating poverty, monitoring forest trends, and supporting policy and trade decisions. Conventional methods for obtaining forest inventory information, such as plantation genus/species, is done in-field, which is time-consuming and costly. Remote sensing is a more efficient way to capture forest genus information. Very high resolution, hyperspectral, and unmanned aerial vehicle (UAV) imagery have been shown to contain suitable spectral and spatial information for machine learning algorithms to differentiate between forest species. However, such data requires extensive processing and is expensive to acquire, making it unsuitable for mapping over larger areas. High-resolution imagery, such as Sentinel-2, combined with textural measures and vegetation indices as features in machine learning algorithms, have shown potential to differentiate between spectrally similar classes. However, it is not known what impact training sample scheme and size have on classification accuracies when classifying *Acacia*, *Eucalyptus*, and *Pinus* (pine) genera. It is also not known whether signature extension is a viable method for reducing the time and effort spent on obtaining in situ training data when mapping forest plantations over a large area.

This research set out two main experiments. The first experiment evaluated the impact of using an even, uneven, or an area-proportionate training sample configuration and size in a random forest machine learning model for classifying acacia, eucalyptus and pine plantations. It was found that the study area that contained an uneven coverage of the three genera was classified more accurately using a balanced training sample configuration, compared to using an unbalanced and coverage-proportionate training sample configuration. It was also found that a saturation point exists where adding more training samples adds little value to the OA. The saturation point was found to be ~ 57n, where n is the number of features used in the classification.

The second set of experiments was set out to test the viability of training data signature extension for constructing random forest machine learning models to differentiate between acacia, eucalyptus and pine trees using Sentinel-2 imagery as input. The study area was split into 19 Sentinel-2 tiles spanning the Mpumalanga, KwaZulu-Natal, Eastern Cape and Western Cape provinces. Three separate random forest models were built using training data collected in one tile located in Mpumalanga, one tile located in KwaZulu-Natal, and one tile located in the Eastern Cape. A fourth model was built using training data from all three source tiles. The four models were applied to all 19 Sentinel-2 tiles to map forest genera. The results show that a ~70% OA can be achieved if the training data is collected in areas with similar climates (rainfall seasonality) to the areas that are being mapped. In addition, it was found that signature extension distance (i.e. distance between the training data and the area being classified) should not exceed 500 km.

More details of the research are provided in Section 4.2 of this report.

VINCE NDYAFI

Mr Ndyafi, supervised by Prof Van Niekerk and Mr Stephenson, is studying part-time and was still busy with his thesis at the time of writing this report. His research is focussing on the use of deep learning and very high resolution colour (RGB) imagery to differentiate between indigenous and plantation forests. Some of the methods and preliminary findings were overviewed in Section 4.1 of this report.

APPENDIX V: PUBLICATIONS

No publications have emanated from this project at the time of writing, but two articles are currently in preparation for submission to scientific journals.

Title	Authors
Impact of training set configurations for differentiating between plantation forest genera with Sentinel-2 imagery and machine learning	Higgs, C Van Niekerk, A
Signature extension as a machine learning strategy for mapping plantation forest genera with Sentinel-2 imagery	Higgs, C Van Niekerk, A

Two other scientific articles relating to Mr Ndyafi's MSc research on the use of convolutional neural networks and object-based image analysis for forest plantation mapping are being planned. Three popular articles in the Tip Mag, Forestry Focus and Water Wheel are planned.

The WRC's role in these publications will be suitably acknowledged.

APPENDIX VI: ACCESS TO DATA GENERATED THROUGH THIS PROJECT

The commercial forestry data used in this study (see Section 3.1) were provided by a number of commercial forestry companies under strict conditions of confidentiality. The data are, as such, not available to any third party.

The satellite imagery used this his study can be freely downloaded from the respective sources (see Section 3.2 for details).

All of the data used in this project will be archived and stored for at least five years.