# A NOVEL APPROACH TO MAPPING AREAS SUITABLE FOR RAINFED PRODUCTION OF BAMBARA NUT AND COWPEA

**Simon Lake**

**Submitted in fulfilment of the academic requirements**
**of Master of Science in Hydrology**

in the
Centre for Water Resources Research
School of Agricultural, Earth and Environmental Sciences
College of Agriculture, Engineering and Science
University of KwaZulu-Natal

29 November 2024

# PREFACE

The research contained in this dissertation was completed by the candidate while based in the Discipline of Hydrology, School of Agricultural, Earth and Environmental Sciences in the College of Agriculture, Engineering and Science, University of KwaZulu-Natal, Pietermaritzburg Campus, South Africa. The research was supported with funding from the Water Research Commission (WRC) for a project titled "Developing a database and utility tool for underutilised indigenous crops for increased agricultural diversification in South Africa".

The contents of this work have not been submitted in any form to another university and, except where the work of others is acknowledged in the text, the results reported are due to investigations by the candidate.

[redacted]

Signed: Mr R Kunz

Date: 19 February 2025

# DECLARATION

I, Simon Vernon Lake, declare that:

(i) the research reported in this dissertation, except where otherwise indicated or acknowledged, is my original work;

(ii) this dissertation has not been submitted in full or in part for any degree or examination to any other university;

(iii) this dissertation does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons;

(iv) this dissertation does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:

a) their words have been re-written but the general information attributed to them has been referenced;

b) where their exact words have been used, their writing has been placed inside quotation marks, and referenced;

(v) where I have used material for which publications followed, I have indicated in detail my role in the work;

(vi) this dissertation is primarily a collection of material, prepared by myself, published as journal articles or presented as a poster and oral presentations at conferences. In some cases, additional material has been included;

(vii) this dissertation does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the dissertation and in the References sections.

[redacted]

Signed: Mr SV Lake

Date: 19 February 2025

# ACKNOWLEDGEMENTS

The author would like to thank the following individuals and organisations for their assistance and contributions to this research:

- First, thanks to my parents for their support and encouragement throughout the study.
- Mr R Kunz for his supervision and guidance throughout the study.
- Mr T Mthembu for his help and guidance.
- Prof T Mabhaudhi for sharing his knowledge on underutilised crops.
- The Water Research Commission for funding the research reported in this study.

# ABSTRACT

South Africa is a water scarce country, highlighting the need to cultivate water use efficient crops under rainfed conditions. Certain neglected and underutilised crops (NUCs) exhibit greater drought resistance, water use efficiency and nutrient content when compared to conventional crops (e.g. maize and soybean), and thus have potential to enhance food and nutrition security at the smallholder scale. To increase production of NUCs in southern Africa, small-scale farmers first need to know where these crops can be cultivated, which they typically obtain from existing land suitability maps. Various methods have been used to develop such maps, which are classified as either traditional or modern methods. Previous studies have used simulated data from a crop model to validate, but not develop, land suitability maps, especially in South Africa. Utilising modelled data to develop a land suitability map for rainfed production of taro was first conceptualised and tested in 2022, which was then modified and applied to orange flesh sweet potato (OFSP) and taro in 2024. This study represents a continuation of this work, where the methodology was further refined and improved, then applied to other NUCs (namely bambara nut and cowpea). Specific objectives include: (i) assessing the impact of four planting months (October, November, December and January) on simulated crop yield, (ii) determining the impact of frost occurrence on crop cycle length and yield, (iii) accounting for frost occurrence when mapping land suitability, and (iv) partially automating the mapping process using the Python programming language. The AquaCrop model was run to simulate 49 seasons of data (from 1950/51 to 1998/99) for over 5 800 relatively homogenous response zones across southern Africa. For each AquaCrop output variable (e.g. biomass, yield and crop water use), various statistics were then derived from the modelled data and stored in a crop database. For each NUC, the crop cycle length was calculated using thermal time, i.e. when sufficient heat units have accumulated for the crop to reach physiological maturity. Thereafter, the crop cycle length was reduced to the first frost date when the daily minimum air temperature dropped to 5.4°C or below. Hence, AquaCrop was run for two main scenarios, namely with and without frost impacting the crop's growing season. Simulations were performed for a single planting density (representing smallholder farming systems) for each planting month. In total, 16 national scale model runs were completed. Python code, together with a geographic information system, were then used to analyse and select AquaCrop output variables, which were then used to develop each land suitability map. The land suitability maps highlight the coastal regions of KwaZulu-Natal and the Eastern Cape as being highly suited to rainfed production of bambara nut and cowpea, whilst inland areas are moderately to marginally suitable. Findings also indicate that planting date affects crop production, with bambara nut and cowpea showing optimal suitability as well as higher yields at earlier and later planting dates, respectively. Frost occurrence reduced the crop cycle length, resulting in lower crop yield simulations, which was also influenced by planting date. Areas deemed suitable for crop production were substantially reduced by (i) frost occurrence, and by (ii) excluding unsuitable land uses (e.g. protected areas and urban areas). This study illustrates how planting date affects land suitability mapping, especially in frost-prone regions. This study assessed land suitability for crop production using an approach that was further enhanced by incorporating frost occurrence in crop simulation modelling, which was achieved for the first time, and thus is considered novel. Python automation expedited map development, producing numerous land suitability maps for each crop in reduced time.

Keywords: bambara nut; cowpea; land suitability; AquaCrop; frost occurrence

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

## LIST OF ABBREVIATIONS

| | |
|---|---|
| APSIM | Agricultural Production Systems Simulator |
| AZ | Altitude Zone |
| CR | Consistency Ratio |
| CSM | Crop Simulation Model |
| CWP | Crop Water Productivity |
| DAFF | Department of Agriculture, Forestry and Fisheries |
| DARDLA | Department of Agriculture, Rural Development and Land Administration |
| DFFE | Department of Forestry, Fisheries and the Environment |
| DSSAT | Decision Support System for Agrotechnology Transfer cropping system |
| FAO | Food and Agricultural Organisation of the United Nation |
| GIS | Geographic Information System |
| LSA | Land Suitability Analysis |
| LULC | Land Use Land Cover |
| MaxEnt | Maximum Entropy |
| MCDM | Multi-Criteria Decision Making |
| MicroLEIS DSS | Mediterranean Land Evaluation Information System Decision Support System |
| NLC | National Land Cover |
| NUC | Neglected and Underutilised Crops |
| SAEON | South African Environmental Observation Network |
| SDG | Sustainable Development Goals |
| STICS | Simulateur mulTIdisciplinaire pour les Cultures Standard |
| UKZN | University of KwaZulu-Natal |
| WRC | Water Research Commission |

# LIST OF SYMBOLS AND UNITS

| | |
|---|---|
| AWC | available water content (mm) |
| B | above-ground biomass production (dry t $ha^{-1}$) |
| $B_{AVE}$ | above-ground biomass production average (dry t $ha^{-1}$) |
| $B_{CV}$ | coefficient of variation of above-ground biomass production (%) |
| BREL | relative biomass potential (%) |
| $BREL_{AVE}$ | relative biomass potential average (%) |
| $BREL_{NORM}$ | normalised relative biomass potential (%) |
| CC | canopy cover (%) |
| CWP | crop water productivity (dry kg $m^{-3}$) |
| $CWP_{AVE}$ | crop water productivity average (dry kg $m^{-3}$) |
| $CWP_{CV}$ | coefficient of variation of crop water productivity (%) |
| CYCLE | crop cycle length (days) |
| $CYCLE_{AVE}$ | crop cycle length average (days) |
| E | soil water evaporation (mm) |
| E/Ex | relative soil water evaporation (%) |
| EC | electrical conductivity (dS $m^{-1}$) |
| $ET_A$ | actual evapotranspiration (mm) |
| $ET_O$ | reference crop evaporation (mm) |
| ExpStr | leaf expansion stress (%) |
| GDD | growing degree-day (°C day) |
| HI | harvest index (%) |
| $HI_{AVE}$ | harvest index average (%) |
| $HI_{NORM}$ | normalised harvest index (%) |
| $HI_O$ | reference harvest index (%) |
| LGP | length of growing period (days) |
| NRMSE | normalised root mean square error (%) |
| RCF | risk of crop failure (%) |
| RFL | seasonal rainfall (mm) |
| $RFL_{AVE}$ | average seasonal rainfall (mm) |
| r | Pearson correlation coefficient |
| $R^2$ | coefficient of determination |
| RH | relative humidity (%) |
| RMSE | root mean square error (% or t $ha^{-1}$) |
| StoStr | stomatal stress (%) |
| TmpStr | cold temperature stress (%) |

| | |
|---|---|
| Tr | total transpiration of crop and weeds (%) |
| Tr/Trx | relative crop transpiration of crop and weeds (%) |
| WF | water footprint ($m^3$ $t^{-1}$) |
| WP | water productivity (kg $m^{-3}$) |
| $Y_{AVE}$ | yield (dry t $ha^{-1}$) |
| $Y_{MISS}$ | number of missing yield simulations |
| $Y_{NUMB}$ | total number of yield simulations |
| $Y_{ZERO}$ | number of seasons with a zero yield |

# 1. INTRODUCTION

## 1.1 Background and Justification

Southern African countries (South Africa, Eswatini and Lesotho) are mostly classified as semi-arid to arid (Thopil and Pouris, 2016). This explains the need for crop irrigation, which can utilise up to 60% of South Africa's available water resources (Bonthuys, 2018; DWS, 2018). Expansion of irrigated areas is highly unlikely as many catchments are already water-stressed, i.e. water demand exceeds supply (Thopil and Pouris, 2016; DWS, 2018). Although the majority of South Africa's agriculture relies on rainfed production (Mabhaudhi et al., 2023), there is a growing need to produce higher crop yields to feed the rapidly growing population, whilst using less water. This can be achieved by growing more water use efficient crops, which is particularly important due to increasing water demands from other sectors (e.g. domestic and industrial use), especially as climate change intensifies.

The cultivation of Neglected and Underutilised Crops (NUCs) has been proposed as one possible solution, as these indigenous crops are adapted to low levels of water availability (Modi and Mabhaudhi, 2016). Certain NUCs have been identified as drought and/or heat-stress tolerant, and thus are ideal for rainfed cropping areas, especially in marginal areas (Mabhaudhi et al., 2017a; Mabhaudhi et al., 2023). In addition, Mabhaudhi et al. (2023) stated that these crops have the potential to contribute towards improving South Africa's agro-biodiversity, which can bolster the adaptability of agricultural systems to the changing climate and may result in increased food production (Renna et al., 2018). Wimalasiri et al. (2022) noted that NUCs can also contribute towards key Sustainable Development Goals (SDGs) that aim to reduce hunger and poverty.

Furthermore, southern Africa has experienced many problems related to poor land use change decisions, where arable land has been converted into urbanised and industrialised areas, resulting in reduced farmland for the production of, inter alia, food, feed, fibre and fuel (e.g. biofuel and bioenergy). Hence, it is important to understand the production potential of NUCs and how they may assist famers to become more resilient to increasing climate variability. In order to increase the production of NUCs in southern Africa, farmers first need to know where these crops can be successfully cultivated. Farmers typically obtain this knowledge from existing land suitability maps for crop production. Such maps are therefore needed to promote the local production of NUCs under rainfed conditions.

Land suitability is defined as the suitability of a particular land parcel for a specified use, whereas land evaluation measures the potential of land for alternative land uses (FAO, 1976).

Land suitability classification is the process of grouping specific regions according to suitability (FAO, 1976). Land suitability mapping is defined by Collins et al. (2001) as a tool that is used to determine areas that are most suitable for a certain land use, i.e. to produce crops under rainfed conditions. A land suitability analysis (LSA) is an important component in agricultural development as it informs a farmer about what crops can best be grown on their land. Land suitability maps provide farmers with additional crop choices to plant, which allows for increased agricultural diversification (Modi and Mabhaudhi, 2016). Land suitability studies are also undertaken to combat food insecurity, resulting in sustainable agricultural development (Akpoti et al., 2019).

Most land suitability studies use the FAO (1976) approach, which classifies land according to order and class. The order is defined as either suitable (S) or not suitable (N) and the class provides the level of suitability or unsuitability. For example, suitable areas are classified as highly (S1), moderately (S2) and marginally (S3) suitable. Similarly, areas classified as N1 and N2 are deemed currently and permanently unsuitable, respectively (FAO, 1976). Although subclasses and units are also defined in the FAO (1976) approach, they are seldom used in land suitability studies.

Based on a review of relevant literature, different methods have been used to develop land suitability maps, which are classified as either traditional or modern methods (Akpoti et al., 2019). The traditional (and simpler) methods are typically based on overlays of spatial climate and soil datasets to identify areas best suited to crop production (e.g. Holl et al., 2007; Nie et al., 2019). The modern (more complex) methods involve machine learning techniques (e.g. Estes et al., 2013; Mugiyo et al., 2022). Other modern approaches are based on multi-criteria decision analysis, in particular the analytical hierarchy process (e.g. Singha et al., 2020; Mugiyo et al., 2021b; Kau et al., 2023). It is worth noting that few land suitability studies (e.g. Estes et al., 2013 and Kunz et al., 2024) have used data from a crop simulation model. Although all these methods utilise climatic, edaphic, topographic and/or socio-economic data as input data, they adopt different approaches for mapping land suitability and utilise different variables, including different thresholds (i.e. cut-offs) and weightings that are applied to the variables.

Until recently, relatively little research has been done on mapping land suitability for NUCs in southern Africa. The following recent studies produced suitability maps for six underutilised indigenous crops: Mugiyo et al. (2021b) and Mugiyo et al. (2022) focused on production of NUCs (amaranth, cowpea, sorghum and taro) across South Africa and KwaZulu-Natal, respectively; Nhamo et al. (2022) considered bambara nut production in Limpopo; and Kunz et al. (2024) mapped orange flesh sweet potato (OFSP) and taro production in southern Africa. Therefore, only three land suitability maps have been developed for bambara nut and cowpea production across South Africa.

Lake (2022) utilised output from a water-driven crop model to assess land suitability for rainfed production of taro. The approach was considered novel (i.e. unique) since it has not been applied before in southern Africa. Thereafter, the methodology was modified by Kunz et al. (2024) and applied to both OFSP and taro. The analysis of modelled data was performed using Microsoft Excel spreadsheets, which was a time-consuming process, especially when applied to multiple crops. Previous approaches have not considered the impact of frost on crop cycle length and crop yield, as well as the impact on land suitability mapping, which further adds to the novelty in the current approach.

## 1.2 Main Aim and Objectives

The main aim of this study was to further improve the approach of utilising simulated data to identify and map land deemed suitable for rainfed crop production. Since the outcome is dependent on model input quality, this study focused on bambara nut and cowpea, due to the availability of well-calibrated crop parameters. The specific objectives were to:

(i) assess the impact of four planting dates on simulated crop yield,
(ii) determine the impact of frost occurrence on crop cycle length and yield,
(iii) account for planting date and frost occurrence when mapping land suitability,
(iv) further improve the methodology and apply it to other NUCs, and
(v) partially automate the methodology using the Python programming language.

Specific questions that this study addressed are as follows:

(i) How does planting date and frost occurrence influence land suitability mapping?
(ii) What are the advantages of using output from a crop simulation model when compared to existing land suitability mapping techniques?
(iii) How can output from a crop simulation model be used to identify suitable growing areas for both NUCs?
(iv) Where can bambara nut and cowpea be successfully cultivated under rainfed conditions in southern Africa?

## 1.3 Dissertation Outline

This dissertation consists of the following five chapters:

Chapter 1: provides the background information and justification for the study, as well as specific aims and objectives;

Chapter 2: describes different land suitability mapping techniques (especially those that used crop yield data);
Chapter 3: outlines the methodology developed and applied in this study;
Chapter 4: presents and discusses the results, focusing on the impact of planting date and frost occurrence on land suitability mapping; and
Chapter 5: conclusions of the study, including recommendations for future research.

# 2. LITERATURE REVIEW

Determining suitable areas for production of NUCs can help improve food and nutritional security as well as increase agricultural diversification, particularly at a smallholder farming scale. A review of available literature was conducted to identify different methods for mapping land suitability for crop production. A total of 14 case studies are presented in **Section 2.1** to illustrate the different methods. In addition, the cases studies were specifically chosen because they either produced land suitability maps for NUCs, or the methodology used some level of computational automation. More importantly, case studies that used measured and/or simulated crop yield data were also selected. **Section 2.2** provides an overview of the AquaCrop model, available crop parameter files as well as an overview of the national-scale crop modelling.

## 2.1 Land Suitability Mapping

Land suitability mapping techniques can be divided into two groups: traditional and modern methods. Traditional methods assess land suitability using qualitative, quantitative and parametric methods based on biophysical characteristics (Mugiyo et al., 2021a). The more complex modern techniques consider multi-criteria evaluation, remote sensing and machine learning approaches (Akpoti et al., 2019).

### 2.1.1 Traditional methods

In the past, most land suitability mapping studies used a Geographic Information System (GIS) to apply climatic criteria for crop growth (which are crop-specific), to spatial and temporal climate datasets. Several case studies are presented in chronological order, briefly describing the methodology used to develop the land suitability maps.

#### 2.1.1.1 Case study 1: Holl et al. (2007)

Holl et al. (2007) considered the production of *Jatropha curcas* in South Africa, assessing its water use and bio-physical potential. For the latter, a land suitability map was produced by applying thresholds (i.e. cut-off values) to five spatial variables, namely rainfall, temperature, soils, frost and slope. The following three-phase approach was used:

(i) elimination of areas that are unsuitable for crop production,
(ii) initial yield estimates were calculated from climate and other data using a weighted approach, and
(iii) a yield equation was then developed to produce estimates of potential yield.

The three-phase approach allowed unsuitable areas to be removed first, and thus, only remaining areas deemed suitable for crop production were further divided into land suitability classes. At the time of the study, no jatropha yield equation was available for South African growing conditions. Hence, selected variables (e.g. rainfall and temperature) were weighted according to their relative importance in determining crop yield (step 2). With the aid of GIS, the weightings were then used to calculate initial estimates of yield. For step 3, sunflower and tree growth yield equations were combined and then adjusted for jatropha (Holl et al., 2007). The yield equation was then used to estimate potential yield for the suitable areas. Yield categories were developed by assuming that yield follows a normal distribution. Hence, the final map identified areas suitable for jatropha production based on estimated yields.

#### 2.1.1.2 Case study 2: Jewitt et al. (2009)

Jewitt et al. (2009) developed land suitability maps identifying suitable areas for selected biofuel crops such as canola, cassava, jatropha, soybean, sugarbeet, sunflower and sweet sorghum. The study involved simple overlays for climatic (rainfall and temperature) data using a GIS that highlighted optimum growing areas. Climatic thresholds for optimum growth were sourced from the available literature. These thresholds were then applied to spatial datasets of rainfall and temperature to identify areas deemed optimally suited for crop growth.

#### 2.1.1.3 Case study 3: Kunz et al. (2015)

Kunz et al. (2015) also developed land suitability maps for canola, grain sorghum (**Figure 7.1** in **APPENDIX A**), soybean, sugarbeet and sugarcane. The maps were created by applying certain thresholds to spatial datasets of rainfall, temperature, relative humidity, soil depth and slope. The thresholds were based on a study undertaken by Khomo (2014) for soybean.

The approach followed the first two phases from Holl et al. (2007) (cf. **Section 2.1.1.1**). Firstly, unsuitable areas for production were eliminated. Thereafter, the variables were ranked according to their relative importance in determining crop growth. Rainfall was considered the most important for crop growth (weighted at 40%), followed by temperature and slope (each 20%), as well as relative humidity and soil depth (each 10%). These weightings were based on expert opinions (Bertling and Odindo, 2013; cited by Kunz et al., 2015), and thus were subjective.

Kunz et al. (2015) stated that rainfall distribution across the growing season can impact the crop's growth. A monthly crop coefficient (Kc) approach was used to determine the amount of rainfall required by the crop in different growth stages. A higher Kc value indicates that the crop requires more water in that month compared to other months (or growth stages). A suitability score was then determined and a score out of 3 was calculated for each month. The study then went further to classify unsuitable areas into two categories (or classes), namely functional “no-go” areas (N1) and absolute “no-go” areas (N2).

#### 2.1.1.4 Case study 4: Nie et al. (2019)

Nie et al. (2019) identified and mapped suitable sites in China for sweet sorghum at a 1 km resolution. The study adopted different approaches to assess the (i) distribution and (ii) potential of biomass production. Firstly, a GIS was used to identify areas suitable for sweet sorghum production when the following two criteria were met: (i) accumulated heat units exceed 2500 growing degree-days (i.e. area is sufficiently warm for crop growth), and (ii) the sand content of the soil is below 85% (i.e. drainage and water retention are suitable for the crop). Production potential was then divided into five categories, namely unsuitable, barely suitable, moderately suitable, highly suitable and perfectly suitable using (i) precipitation and (ii) slope. Secondly, FAO’s AquaCrop model (Steduto et al., 2009) was used to simulate yield using daily meteorological data to determine the potential biomass production for each region.

#### 2.1.1.5 Case study 5: Marraccini et al. (2020)

Marraccini et al. (2020) used an innovative approach to develop a land suitability map for soybean in France. The study uses not only bio-physical variables (e.g. climatic and edaphic) but also agronomical variables (e.g. yield) and economic variables (e.g. gross margin). The study first considered climatically suitable areas for production as well as unsuitable areas, e.g. non-agricultural areas, areas with permanent crops, fallow areas as well as areas with at least one forage crop (to avoid livestock farming areas). Soil suitability considered the carbonates percentage, stone/gravel percentage and the soil available water content. Agronomic criteria accounted for crop rotations which followed a set of rules explained in detail by Marraccini et al. (2020). Finally, economic criteria looked at the gross margins difference between the current rotation and the re-designed rotation for soybean. A difference greater than -50€ was considered unsuitable, with greater than +50€ being very suitable. The gross margins were calculated using the selling price, variable costs and yield simulated by the STICS model (Simulateur mulTIdisciplinaire pour les Cultures Standard; Brisson et al., 1998).

#### 2.1.1.6 Case study 6: Khalid et al. (2021)

Khalid et al. (2021) also identified and mapped suitable sites for bioenergy production using *Jatropha curcas* in the northern region of Pakistan. The study utilised the following two methods:

(i) The first applied selected thresholds to climate, elevation and slope data to determine three suitability classes (more, moderate and less suitable). This approach was similar to previous studies involving *Jatropha curcas* that predominately used climate (rainfall and temperature), elevation and slope data (e.g. Holl et al., 2007; cf. **Section 2.1.1.1)**. The study area was divided into smaller sub-regions using the Thiessen polygon method based on climate and soil data input.

(ii) The second method used AquaCrop model to simulate jatropha yield (Y in t $ha^{-1}$), due to the lack of reliable, measured data. The model was also used to estimate crop water use (WU in $m^3$) and water productivity (WP in t $m^{-3}$). From the modelled output, the water footprint (WF in $m^3$ $t^{-1}$) was calculated as the inverse of WP. The study undertook a simple overlay of Y and WF within each of the suitability classes, which showed the most suitable areas were associated with high Y and low WF (i.e. high WP), as expected.

#### 2.1.1.7 Case study 7: Nhamo et al. (2022)

Nhamo et al. (2022) developed a land suitability map for bambara nut production in the Limpopo province. The study applied six variables (rainfall, temperature, soil type, soil depth, topsoil clay content and slope) to classify marginal, moderate and highly suitable crop production areas. Each of the three suitability classes were further disaggregated into three land suitability sub-classes based on the potential (expressed as a percentage) of each land type, i.e. 10-30%, 30-50% and > 50%. This was done using land type data developed by the Agricultural Research Council, where land capability was assessed using soils, slope and climate data. Land types classified as less than 10% potential were considered unsuitable for crop production (**Figure 2.1**). The methodology was then repeated without considering rainfall to produce a land suitability map for irrigated production of bambara nut. The results showed the rainfall was a major limiting factor for crop production in the province.

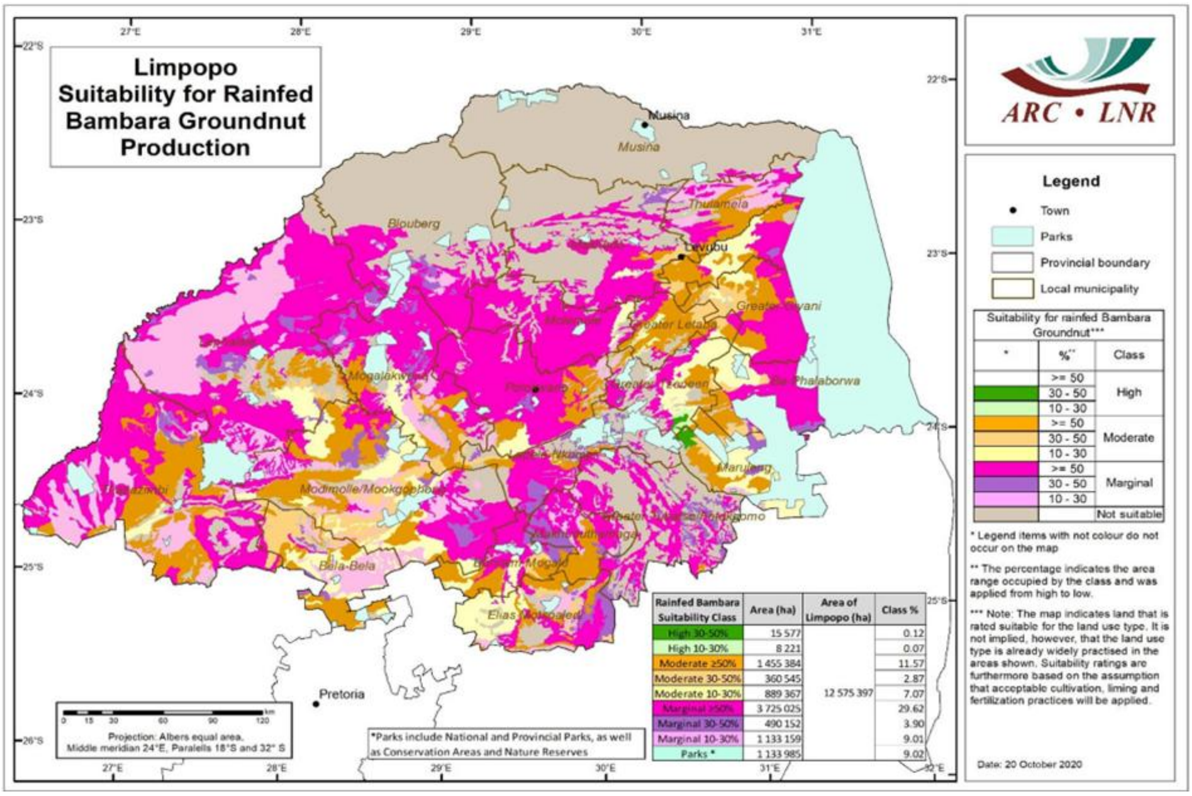


**Figure 2.1: Land suitability map for rainfed cultivation of bambara nut in the Limpopo province (Nhamo et al., 2022)**

### 2.1.2 Modern methods

Akpoti et al. (2019) described modern land suitability methods as those that combine GIS and machine learning models. Similarly, Mugiyo et al. (2021a) described modern methods as those utilising more complex and dynamic algorithms. Hence, modern methods are considered more complex than the simpler, traditional methods (cf. **Section 2.1.1**), as they require additional input datasets.

Mugiyo et al. (2021a) highlighted four main categories of modern methods as follows:

(i) computer assisted overlays: involves the use of GIS, remote sensing and cloud computing technologies;
(ii) machine learning: incorporates the use of, inter alia, species distribution models, fuzzy rule-based systems and artificial neural networks;
(iii) multi-criteria decision-making (MCDM): includes analytical hierarchy process (AHP), which is considered one of the most reliable MCDM methods; and
(iv) crop simulation models: mathematical models that describe crop growth as a function of climatic, edaphic and management conditions.

The analysis by Mugiyo et al. (2021a) of 101 journal publications showed that the most commonly used modern method for land suitability assessment involves machine learning

(25.7%), followed by AHP (14.9%), fuzzy logic (12.9%) and crop simulation models (9.9%). As mentioned at the start of this chapter, the case studies were carefully selected to illustrate the use of each method, especially those that either utilised crop yield data (measured and/or simulated), or considered NUCs, or the methodology was automated to some extent.

#### 2.1.2.1 Machine learning

The use of machine learning for mapping land suitability has recently gained popularity. For example, MaxEnt (Maximum Entropy) is a species distribution model based on a general-purpose machine learning approach with an intuitive and exact mathematical formulation (Philips et al., 2006). MaxEnt is trained with a presence dataset, i.e. a list of geographical coordinates (point locations) identifying where the target species was successfully grown in the past. These presence points, together with selected predictor variables (e.g. rainfall and temperature) are used by MaxEnt to predict the suitability of other locations for species growth. Hence, MaxEnt can create land suitability maps by predicting the probability of species growth at a specific location, ranging from 0 (unsuitable for growth) to 1 (ideally suited for growth). MaxEnt has also been applied to identify suitable crop cultivation areas, as illustrated by the following two case studies.

<u>Case study 8: Estes et al. (2013)</u>

Estes et al. (2013) used MaxEnt to create a land suitability map for maize cultivation in South Africa. The study used 11 390 presence (i.e. suitable maize areas) and 11 390 absence (i.e. unsuitable maize areas) data points. Five variables (temperature, precipitation, soil depth and topsoil organic carbon) were used to predict maize suitability. MaxEnt was also trained with high-productivity occurrence points, which improved the prediction considerably. Estes et al. (2013) used the DSSAT crop model to produce a land suitability map using climate and soil data available for 5 838 Altitude Zones across southern Africa (cf. **Section 2.2.3**). The study showed that MaxEnt and DSSAT (Decision Support System for Agrometeorology Transfer; Jones et al., 2003) were equally successful in predicting overall crop suitability compared to a land suitability map derived using observed maize yield data.

<u>Case study 9: Mugiyo et al. (2022)</u>

Mugiyo et al. (2022) also used MaxEnt to identify suitable growing areas in KwaZulu-Natal (KZN) for selected NUCs (amaranth, cowpea, sorghum and taro). The study only used 60 presence points in KZN, of which half were randomly selected for model training and the other half for model validation. The predictor variables were as follows:

(i) three climatic (precipitation, temperature & length of growing period),

(ii) seven soil (available soil water capacity, soil pH, soil depth, soil texture & fraction of clay, silt and sand),
(iii) two topographic (elevation & slope), and
(iv) two socio-economic (distance along the road network & distance to metro cities) variables.

In total, 1 000 model runs were completed to create land suitability maps. Four suitability classes were defined using MaxEnt probability values as follows: highly suitable (S1; suitability index > 0.80), moderately suitable (S2; 0.60 - 0.79), marginally suitable (S3; 0.20 - 0.59) and unsuitable (N1; < 0.19). The map produced for cowpea in KZN is shown in **Figure 2.2**.

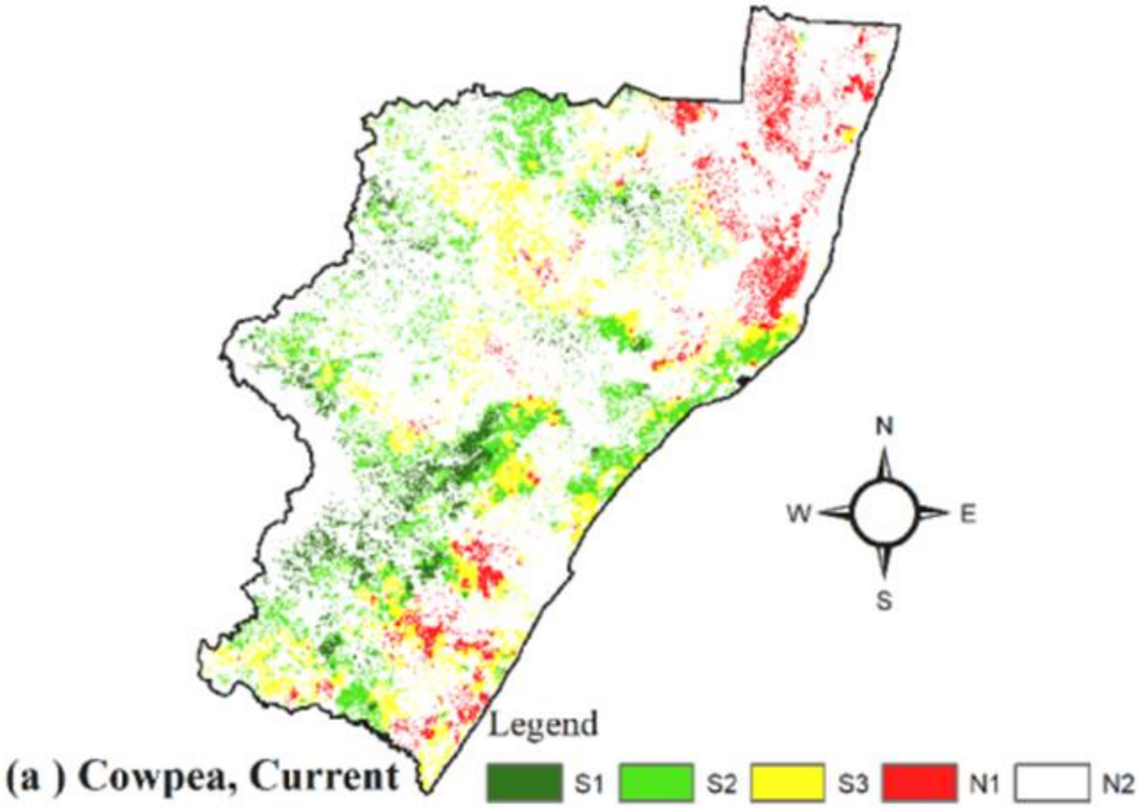


**Figure 2.2: Land suitability map for cowpea in KZN using MaxEnt (Mugiyo et al., 2022)**

Additionally, MaxEnt generates Jack-knife plots that quantify how each predictor variable contributes to the overall suitability index. The plots developed by Mugiyo et al. (2022) indicated that precipitation was the most influential variable in predicting cowpea suitability, followed by length of growing period (LGP) and temperature (both maximum and minimum). Soil depth, soil pH and slope provided the least contributions to overall suitability. Similarly, the two socio-economic variables (EUCDIST and ACCESS) did not contribute much to predictability **(Figure 2.3)**.

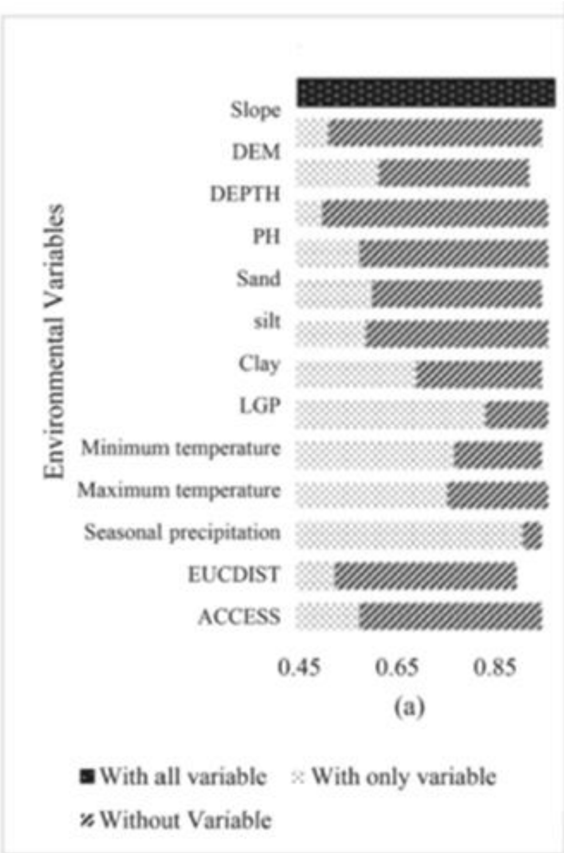


**Figure 2.3: Jack-knife plot evaluating the relative importance of environmental variables for predicting suitability of cowpea cultivation in KZN (Mugiyo et al., 2022)**

#### 2.1.2.2 Analytical hierarchy process

The second most widely used method for LSAs involves AHP, which was originally developed by Saaty (1987) to handle complex decision-making (Mugiyo et al., 2021a). The method uses pairwise comparisons to capture both subjective and objective components of decision-making and then synthesises the outcome into a single index (Saaty, 2008). AHP is based on a hierarchical structure (Saaty, 1987), which uses a nine-point scale to weight each variable according to its importance (Mugiyo et al., 2021a). A score of 1 means that each variable has equal importance, whereas 9 means one variable is far more important than the other (**Table 2.1**).

**Table 2.1: The fundamentals for pairwise comparison (Saaty, 2008)**

| Intensity of importance | Definition | Explanation |
|---|---|---|
| 1 | Equal importance | Two activities contribute equally to the objective |
| 3 | Moderate importance of one over another | Experience and judgement slightly favour one activity over another |
| 5 | The strong or essential importance | Experience and judgement strongly favour one activity over another |
| 7 | Very strong or demonstrated importance | Activity is strongly favoured, and its dominance showed in practice |
| 9 | Extreme importance | The evidence favouring one activity over another is of the highest possible order of affirmation |
| 2, 4, 6 & 8 | Even numbers represent intermediate values between the two adjacent judgements | When compromise is needed |

The AHP method utilises the pairwise comparison matrix to determine the weight of each variable by extracting the largest eigenvalue of the matrix and then normalising the sum of its components. To verify the weightings, a consistency ratio (CR) is calculated, which must be below 0.1, thus indicating that (i) the pairwise comparison matrix has acceptable consistency, and (ii) the weightings of each variable are valid and can therefore be utilised (Chen et al., 2010).

GIS-based AHP has become increasingly popular in recent years due to its ability to assimilate large quantities of data and to calculate weightings easily (Chen et al., 2010). The approach requires a set of (i) geographic boundaries (either polygons or grid cells) and (ii) criteria applied to selected spatial data layers. It then calculates the weightings associated with each data layer using a preference matrix, where each variable (or layer) is compared to the others with preference factors. Similar to the weighted combination method, the weightings can then be aggregated with the data layers (Chen et al., 2010). Three case studies are presented next that used AHP to develop land suitability maps.

Case study 10: Jayathilaka et al. (2012)

Jayathilaka et al. (2012) used the AHP method to develop a land suitability map for three crops (coconut, rubber and tea). The study used three climatic (rainfall, relative humidity and temperature) and two crop (evapotranspiration and yield) factors. Suitability classes and relative weightings for each crop were determined by experts. The plantation estates were randomly selected to determine historical annual yields and the spatial extent for each crop. The weighted average index was used to rank the five factors, and then the linear combination method was applied in computing the final suitability for each crop. Using the weighted average index, rainfall was ranked first for tea and coconut, whereas temperature ranked first for rubber. Crop yield and relative humidity were ranked second and least important for all three crops, respectively.

Case study 11: Singha et al. (2020)

Singha et al. (2020) used AHP to create a land suitability map for two crop rotations (rice-jute and potato-lentil), which are grown during the summer and winter seasons, respectively. The study collected 70 soil samples (soil pH, available nitrogen, phosphorous, potassium, organic carbon, electrical conductivity, available zinc and soil texture) collected across five sites. The weighting of each variable was based on expert opinions and existing literature. From the pairwise comparison, soil texture was identified as most important for rice and lentil, whereas potassium and pH were most important or potato and jute, respectively. Yield data for each plot was also collected to develop maps with four yield classes to visualise spatial variation. The yield maps were then aligned with the suitability classes, which identified areas with marginal suitability and low yields to be used for alternative crops. The yield maps were also used as validation for the land suitability maps.

Case study 12: Mugiyo et al. (2021b)

Mugiyo et al. (2021b) used AHP to create a land suitability map for selected NUCs (amaranth, cowpea, sorghum and taro). Nine variables were used in their pairwise comparison as follows: rainfall, temperature, reference evapotranspiration, length of growing period, elevation, slope, land use/cover, soil depth and distance from roads. Mugiyo et al. (2021b) assigned the highest pairwise weighting to rainfall, whereas distance from roads had the lowest as shown in **Table 7.2** (**APPENDIX A**). After the weightings were determined, the weighted linear combination method was applied. The final suitability index was determined using Liebig's law of the minimum (Mugiyo et al., 2021b). The study used the FAO land suitability classification (FAO, 1976; cf. **Section 1.1**) to develop the final maps. Each suitability class was assigned to a final suitability index range as follows: S1 (> 80), S2 (60-80), S3 (45-59), N1 (30-44) and N2 (< 30). The map produced for cowpea is shown in **Figure 2.4**, which identifies very few areas classified as highly suitable for cowpea cultivation, compared to large parts of the country (excluding the Northern Cape province) that are considered moderately and marginally suitable.

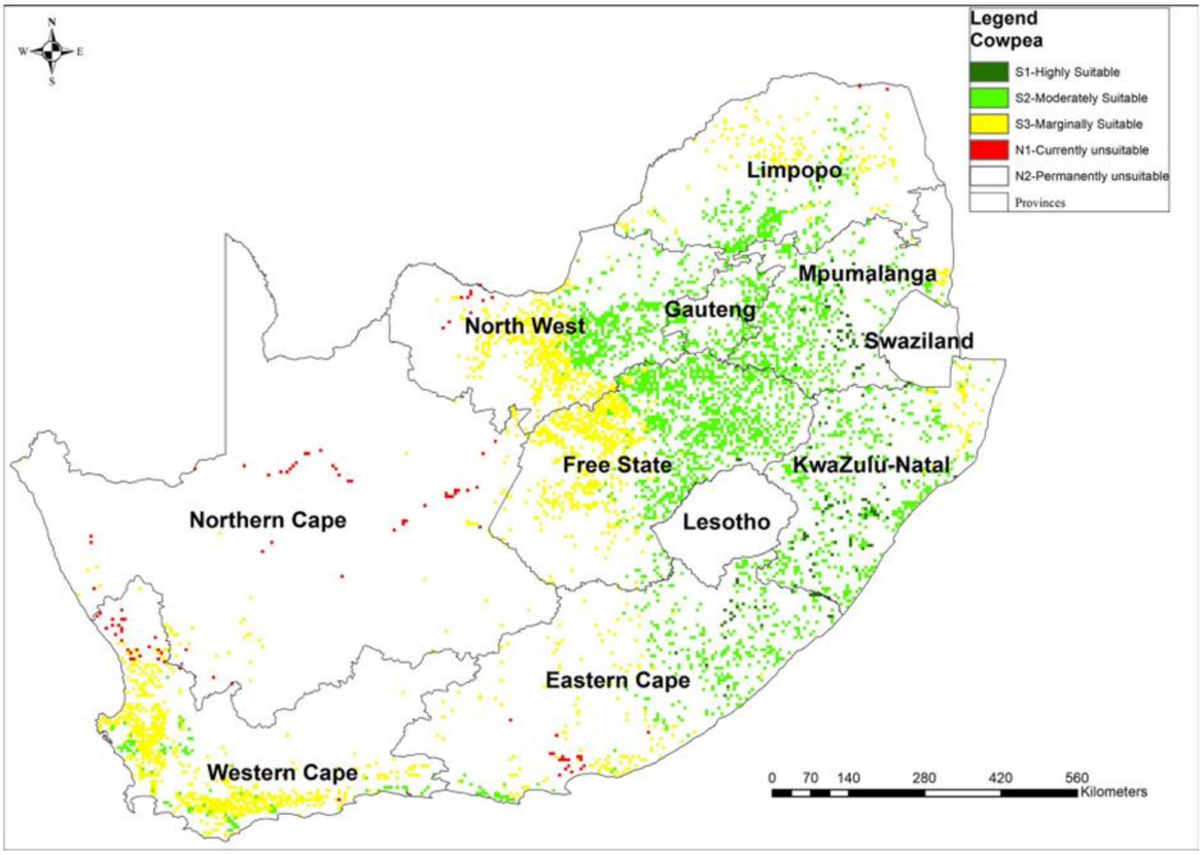


**Figure 2.4: Land suitability map for cowpea based on AHP (Mugiyo et al., 2021b)**

<u>Case study 13: Kau et al. (2023)</u>

Kau et al. (2023) also used AHP to create land suitability maps for three crops in Sudan (faba beans, sugarcane and sorghum). Based on recommendations by FAO (1976) and Sys et al. (1991), the maps were created by considering eight variables, namely rainfall, temperature, soil (texture, depth, pH and organic carbon), slope and land cover. The *ExtAHP20* extension available in ArcGIS version 10.6 (ESRI, 2017) was used to automate the process of calculating (i) the relative weighting of each variable, and (ii) the CR for the matrix. Kau et al. (2023) also used the *ModelBuilder* tool in ArcGIS to partially automate the process to create each land suitability map. This tool can be used to develop workflows using a sequence of geoprocessing tools, where the output of one process is used as input for another (ESRI, 2023a). Using the Reclassify tool, the classification of selected variables was incorporated into the model and assigned a suitability class corresponding to values from the input raster (i.e. grid) datasets. The limiting factors were then extracted using the *Extract by Attributes* tool to create a map of unsuitable areas, which was incorporated into the weighted overlay process to create the final suitability map.

#### 2.1.2.3 Crop simulation modelling

As noted earlier, crop simulation models (CSMs) are mathematical models that describe crop growth as a function of climatic, edaphic and management conditions. They were first developed in the late 1960s to describe how crops function by integrating existing knowledge on plant physiological processes (Bouman et al., 1996; Sinclair and Seligman, 1996).

According to Steduto (2006), CSM use either a carbon-, solar- or water-driven growth engine. A carbon-driven growth engine is based on carbon fixation during photosynthesis (Steduto, 2006; Salman et al., 2021). Canopy growth and development are influenced by temperature, radiation and atmospheric $CO_2$ levels, with water availability acting as a limiting factor. These models are often complex and require many input parameters (Todorovic et al., 2009). One of the most widely used models is WOFOST (WOrld FOod STudies; Van Diepen et al., 1989). A solar-driven growth engine estimates biomass directly from solar radiation through a single radiation use efficiency coefficient (Steduto, 2006; Salman et al., 2021). Examples of solar-driven models include APSIM (Agricultural Production Systems sIMulator; McCown et al., 1994), CropSyst (Stockle et al., 2003), DSSAT (cf. **Section 2.1.2.1**) and STICS (cf. **Section 2.1.1.5**). The water-driven growth engine is driven by a strong relationship between biomass and canopy transpiration through a coefficient called biomass water productivity (Steduto, 2006). Examples of models using a water-driven engine include AquaCrop (Steduto et al., 2009) and CropSyst, the latter being a dual growth engine model, which favours the solar-radiation driven engine.

CSMs have become useful tools for decision making, yield forecasting as well as assessing potential impacts of climate change on crops (Mabhaudhi, 2012), and thus their use has increased in recent years. They are regarded as one of the most reliable methods of determining land suitability at various scales (Mugiyo et al., 2021a). For example, the MicroLEIS DSS (Mediterranean Land Evaluation Information System Decision Support System; De la Rosa et al., 2004) model was used to predict land suitability (Mugiyo et al., 2021a). Estes et al. (2013; cf. **Section 2.1.2.1**) demonstrated the ability of DSSAT to predict land suitability. Simulated output from AquaCrop has also been used to develop land suitability maps, which is discussed next in the final case study.

Case study 14: Lake (2022)

Lake (2022) used AquaCrop output to estimate land suitability for taro. The study used AquaCrop output generated for 5 838 zones across southern Africa (cf. **Section 2.2.3**). For each zone, the model simulated 49 consecutive seasons from 1950/51 to 1998/99, from which long-term monthly and seasonal statistics were generated (e.g. averaged yield). If the zone is too dry or cold for viable crop growth, the model sometimes failed to simulate any output, and thus no statistics were produced.

The methodology developed by Lake (2022) followed a similar approach to Holl et al. (2007; cf. **Section 2.1.1.1**) and Kunz et al. (2015; cf. **Section 2.1.1.3**) where a three-step approach was used. Firstly, Altitude Zones that met any of the following criteria were deemed unsuitable for crop production, and thus were eliminated:

(i) First viable planting date - this was calculated by Kunz et al. (2020) using specific rainfall and temperature criteria, which identified 1 527 Altitude Zones with no viable planting date;

(ii) Risk of crop failure > 50% - this was calculated as the ratio between the number of seasons that produced a zero yield to the total number of simulated seasons, expressed as a percentage (e.g. 50% represents a crop failure every two years on average);

(iii) CWP < 0.1 kg $m^{-3}$ - crop water productivity (CWP) represents the ratio of crop yield (in kg) to water use (in $m^3$), which was used to eliminate zones with low productivity (either low yield or high water use);

(iv) Inter-seasonal variation in yield > 150% - zones with high variation in yield were considered unviable for crop cultivation due to high climate variability;

(v) Crop cycle length < 100 and > 300 days - crop cycle length was used to eliminate areas that are too hot or too cold for crop cultivation.

Given the limited information available on these crops, there is little consensus on what conditions are considered suitable or unsuitable. As a result, the thresholds for determining risks are subjective, varying from one farmer to another as well as differing for each crop. What one farmer considers risky, another might find acceptable, making it difficult to establish objective criteria without conducting surveys to gather more comprehensive data. The thresholds were derived through expert opinion or through testing (e.g. if too many areas were eliminated along the east coast, then the threshold was too high).

The second step involved classifying the remaining suitable areas into highly (S1), moderately (S2) and marginally suitable (S3), based on the FAO (1976) classification (cf. **Section 1.1**).

The mean CWP was calculated for the remaining zones and moderately suitable (S2) zones were defined as having a CWP within half a standard deviation of the mean value. The final step considered current land use, which identified areas that are currently (N1; e.g. sugarcane, orchards and forestry areas) and permanently unsuitable (N2; e.g. urban, protected, mining and water areas).

Case study 15: Kunz et al. (2024)

Kunz et al. (2024) modified the first step of the above approach for taro and OFSP as follows:

(i) If the crop cycle length exceeded 365 days for all seasons, the zone was eliminated since it is likely too cold for economically viable production of annual crops.
(ii) Zones that were considered too dry for rainfed crop production were eliminated if the zone's mean annual precipitation was below 400 mm (Brouwer and Heibloem, 1986).
(iii) If the risk of crop failure (RCF) was greater than 40%, the zone was also eliminated.
(iv) Zones were eliminated if CWP was < 0.1 kg $m^{-3}$ for taro and < 0.6 kg $m^{-3}$ for OFSP.
(v) The inter-seasonal variation in CWP was used to eliminate zones greater than 150% for taro and 100% for OFSP.

Kunz et al. (2024) considered collinearity amongst the variables used to develop each land suitability map. Tests showed that several variables outputted by AquaCrop are collinear. For example, stomatal and leaf expansion stress were highly collinear (coefficient of determination of 0.973) as both variables are indicative of water stress. CWP and yield exhibited a strong linear relationship with an $R^2$ value of 0.968. This was expected since CWP is calculated from yield, i.e. a high yield implies a high CWP. However, a study by Goudriaan and Bastiaanssen (2013) found a non-linear relationship between these two variables.

The land suitability map developed for OFSP is shown in **Figure 2.5**, which identifies coastal (and nearby inland) regions of KwaZulu-Natal and the Eastern Cape as highly suitable for OFSP cultivation. In contrast, parts of the North-West, Limpopo, Free-State and Mpumalanga provinces are considered marginally to moderately suitable. The AquaCrop model simulations undertaken to produce the land suitability map is based on a specific planting density (e.g. 31447 plants $ha^{-1}$) and planting date (01 November), which affects the model simulations. Kunz et al. (2024 produced four land suitability maps based on two planting densities and two planting dates, which differ slightly.

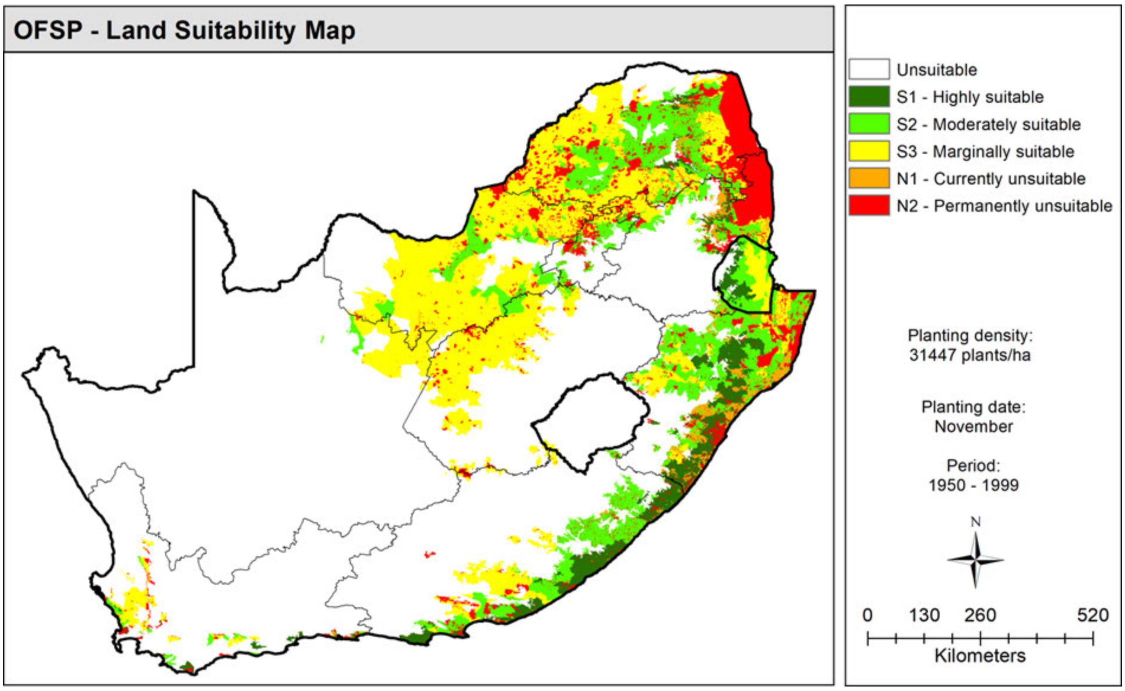


**Figure 2.5: Land suitability map for OFSP based on AquaCrop output (Kunz et al., 2024)**

### 2.1.3 Discussion and conclusions

The 14 case studies presented in **Section 2.1** highlighted various land suitability mapping approaches that range from simple traditional methods to more complex modern approaches. The case studies highlighted attempts to automate the methodology or how yield data was used in the mapping approach. The following sections discuss the (i) automation of the mapping approach, the (ii) importance of considering existing land use, as well as the (iii) selection of variables used in the methodology, with particular emphasis on collinearity amongst variables, and (iv) the use of observed or simulated yield.

#### 2.1.3.1 Automation of methodology

From the literature review, only one study (Kau et al., 2023; cf. **Section 2.1.2.2**) automated, to some extent, the land suitability mapping approach. The study utilised the *ModelBuilder* tool in ArcGIS (ESRI, 2023a) to automate the workflow to improve the efficiency of spatial data analyses. Other studies (e.g. Shaloo et al., 2022, Abebe et al., 2023; Saraswat et al., 2023) also utilised *ModelBuilder* to develop workflows for mapping land suitability. Ali et al. (2024) used a Python-based ArcGIS toolbox called *SDMtoolbox* to automate the pre-and post-processing of data for the MaxEnt model. Automating the land suitability mapping process enhances efficiency by reducing the time required for data processing and analysis, particularly when dealing with large datasets. By integrating automation tools such as

ModelBuilder and Python-based scripts, repetitive tasks can be streamlined, minimising human error and ensuring greater consistency in map production.

#### 2.1.3.2 Consideration of existing land use

Only seven studies considered unsuitable land uses as follows: (i) Estes et al. (2013) excluded national parks and nature reserves; (ii) Kunz et al. (2015) excluded both currently (N1; e.g. forest plantations and sugarcane areas) and permanently unsuitable areas (N2; e.g. urban and protected areas); (iii) Marraccini et al. (2020) eliminated non-agricultural areas such as forests, roads, cities, permanent crops (e.g. orchards and permanent pasture), fallow areas and areas that included at least one forage crop; (iv) Mugiyo et al. (2021b) excluded conservation areas, forest plantations, urban areas and water bodies; (v) Nhamo et al. (2022) excluded protected areas only; (vi) Although Mugiyo et al. (2022) did not mention the exclusion of certain land uses, the map showed protected areas adjacent to the Drakensberg Mountain range (KwaZulu-Natal) as well as urban areas around Durban were not included; (vii) Kunz et al. (2024) excluded both currently (N1) and permanently (N2) unsuitable areas, which reduced the suitable (S1-S3) areas by 22%. This highlights the importance of accounting for current land use as it avoids over-estimating land deemed suitable for crop cultivation, which results in more realistic land suitability maps.

#### 2.1.3.3 Selection of variables

<u>Type and number of variables</u>

From **Table 7.1**, **Table 7.3** and **Table 7.4** (cf. **APPENDIX A**), temperature is commonly used as a mapping variable in almost all studies (except Singha et al., 2020), with only two studies using both maximum and minimum temperature data. Using average temperatures should be avoided in LSAs since all crops require an upper and base temperature for development. Hence, both minimum and maximum temperatures are required to accurately calculate the crop cycle length and final yield. Frost was only considered in one study (cf. **Section 2.1.1.1**), which is surprising since many annual crops are frost sensitive. Length of growing period (LGP) is another important variable used in three case studies, of which only one (cf. **Section 2.1.2.3**) used an upper threshold to eliminate areas deemed too cold for economically viable crop production.

Intuitively, the number of variables considered in each case study influences the accuracy of land suitability assessment. For example, Jewitt et al. (2009; cf. **Section 2.1.1.2**) only used two climatic variables, resulting in a simple classification of either optimally suitable (S1) or currently unsuitable (N1) areas, potentially under-estimating suitable areas since moderate (S2) or marginal (S3) areas were classified as N1. Kunz et al. (2015; cf. **Section 2.1.1.3**) considered other variables (e.g. relative humidity, soil depth, and slope), and thus developed a more comprehensive and reliable land suitability assessment.

Ziadat and Sultan (2011) noted that classic (i.e. traditional) land suitability approaches have a common major limitation, in that they do not consider the farmer's management practices. This particularly impacts the machine learning approaches (e.g. MaxEnt), which require presence data that is typically obtained from remote sensing and/or manual surveys. However, presence data does not infer information regarding management practices, such as using supplemental irrigation and/or fertilisers, which can influence whether or not a particular crop is naturally suited to a particular location.

Collinearity between variables

Collinearity describes the relationship (i.e. correlation) between predictor variables (Enders, 2013) and is an important concept to consider when selecting variables for land suitability studies. Only one case study (Kunz et al., 2024) considered collinearity amongst the variables used to map land suitability. Three case studies (Khalid et al., 2021; Mugiyo et al., 2021b; 2022) used both temperature and elevation to assess land suitability. However, temperature is strongly correlated to altitude due to known lapse rates, i.e. cooler temperatures occur at higher elevations (Schulze and Maharaj, 2004). The use of both elevation and temperature in LSAs should therefore be avoided. However, if temperature is unavailable, then elevation can be used as a surrogate variable.

#### 2.1.3.4 Inclusion of yield data

As mentioned at the beginning of **Chapter 2**, eight of the 14 case studies were selected because they used measured and/or simulated crop yield data. Yield is an important variable for identifying suitable crop cultivation areas, despite it being difficult to obtain from farmers or estimate reliably using a CSM. Of the four case studies that incorporated crop yield data into their traditional method, only one (Holl et al., 2007; cf. **Section 2.1.1.1**) used estimated yield data to map land suitability. The authors used a simple yield equation to identify areas deemed suitable for jatropha cultivation. The main advantage of their approach is that it can be replicated for other crops. For example, crop growth criteria gleaned from the literature can be used to eliminate unsuitable growing areas. Thereafter, the simple rainfall and temperature-based yield equations developed by Smith (1994) could be used to develop spatial crop yield estimates as demonstrated by Schulze and Maharaj (2007) for sorghum. These yield estimates could then be used to classify the remaining suitable areas into S1, S2 and S3 classes.

Nie et al. (2019; cf. **Section 2.1.1.4**) used AquaCrop to estimate sweet sorghum's yield potential, which was not used to map land suitability, but rather to determine biomass production potential. Marraccini et al. (2020; cf. **Section 2.1.1.5**) used the STICs model to estimate soybean yield potential for calculating gross margins to determine economic suitability (not to assess land suitability). Khalid et al. (2021; cf. **Section 2.1.1.6**) performed a simple overlay of Y and WF estimates from AquaCrop, which as expected, showed the most suitable areas were associated with high Y and low WF. Sleiman et al. (2024) also used AquaCrop to determine Y and CWP but similar to other studies, model output was only used to validate (but not develop) the land suitability map.

Four case studies also utilised crop yield data in their modern method, with three utilising yield data directly to map land suitability. Estes et al. (2013; cf. **Section 2.1.2.1**) developed a maize suitability map using different methods. For example, the authors used MaxEnt, observed yield data and yield simulated by DSSAT to produce similar land suitability maps. Singha et al. (2020; cf. **Section 2.1.2.2**) utilised observed yield data to validate their land suitability maps, as was done by Dedeoğlu and Dengiz (2019) and Singha et al. (2023). Jayathilaka et al. (2012; cf. **Section 2.1.2.2**) showed that historical crop yield was ranked second of the five mapping variables, thus highlighting its relative importance in land suitability studies. Kunz et al. (2024; cf. **Section 2.1.2.3**) used selected output variables from AquaCrop to first eliminate unsuitable crop growing areas and then used the CWP variable (which is correlated to crop yield) to classify the remaining areas into three suitability classes.

The above-mentioned studies showed that high yield equates to high suitability. Yield data is also useful validating land suitability maps, and thus should be utilised in LSAs. Since neglected and underutilised crops have been under-researched in the past, measured yields are scarce, and thus modelled data can be used to fill this knowledge gap. Although three models (AquaCrop, DSSAT and STICS) were used to simulate yield, the approach taken by Kunz et al. (2024) with AquaCrop is unique as it used AquaCrop output for the first time to develop land suitability maps, and thus warrants further development. The following section describes the AquaCrop model and how it can be run at a national scale.

## 2.2 AquaCrop Model

The AquaCrop model was developed by the FAO (Steduto et al., 2009) and is based on the Doorenbos and Kassam (1979) equation that shows crop yield is correlated to crop evapotranspiration. When compared to WOFOST, CropSyst and DSSAT (cf. **Section 2.1.2.3**), AquaCrop is (i) more robust and less prone to errors (Steduto, 2006), (ii) simpler to use, and

(iii) requires less input parameters (Todorovic et al., 2009; Saab et al., 2015; Kephe et al., 2021).

Steduto et al. (2009) highlighted the applicability of AquaCrop in environments where water availability is a major limiting factor to crop growth. In South Africa, where mean annual rainfall is about 500 mm (Dennis and Dennis, 2012), crop production is mostly rainfed (Mabhaudhi et al., 2023), and thus AquaCrop is well suited to simulating crop production under water limited conditions. A brief description of the AquaCrop model is given next.

### 2.2.1 Model description

AquaCrop is structured on interactions between the soil, plant and atmosphere (**Figure 2.6**). The model runs at a daily time step, which adequately resembles the time scale at which crops respond to water stress. AquaCrop requires climatic inputs of rainfall, temperature (maximum and minimum) and reference evapotranspiration ($ET_O$). The model also requires annual atmospheric carbon dioxide ($CO_2$) levels.

AquaCrop separates actual evapotranspiration (ETa) into soil water evaporation (E) and crop transpiration (Tr). Biomass (B) is then calculated as the product of Tr and water productivity (WP). The latter is an important crop-specific parameter required by the model, which is normalised by the atmospheric $CO_2$ concentration (with a reference value of 369.41 ppm). In addition, Tr is divided by reference evapotranspiration ($ET_O$). The normalisation of both WP and Tr improves the model's robustness by correcting for varying climatic conditions across diverse locations, seasons as well as for future climate scenarios (Raes et al., 2018). The model then calculates Y as the product of B and the harvest index (HI) (Steduto et al., 2009).

Four studies (Geerts et al., 2009; Vanuytrecht et al., 2014; Ran et al., 2022; Akbari et al., 2024) related to AquaCrop's sensitivity to input soil parameters highlighted the importance of accurately determining rooting depth and soil water retention characteristics. They showed that yield simulations are sensitive to the soil water content (SWC) at field capacity and permanent wilting point, followed by effective rooting depth and saturation. Based on these findings, it is important to measure soil water retention parameters, rather than estimating values from texture using pedotransfer functions (e.g. Saxton and Rawls, 2006) or from depth-weighted values that represent a homogeneous, single layer soil.

The dashed lines in **Figure 2.6** account for cold temperature and soil water stress on various processes within the model. For example, cold temperature stress reduces leaf expansion and crop transpiration, as well as decreasing HI, thus resulting in lower biomass production and yield formation (Raes et al., 2018). The simulation of SWC is estimated using the soil water balance method, which accounts for water gains (rainfall, irrigation (I) and capillary rise) and losses (runoff, evapotranspiration and deep percolation). Soil water stress affects green canopy expansion, canopy senescence, stomatal conductance (which influences transpiration) as well as the harvest index. Hence, soil water stress also reduces biomass production and yield formation. Similarly, soil fertility and salinity stress affects canopy cover development, crop transpiration and biomass production (Raes et al., 2018).

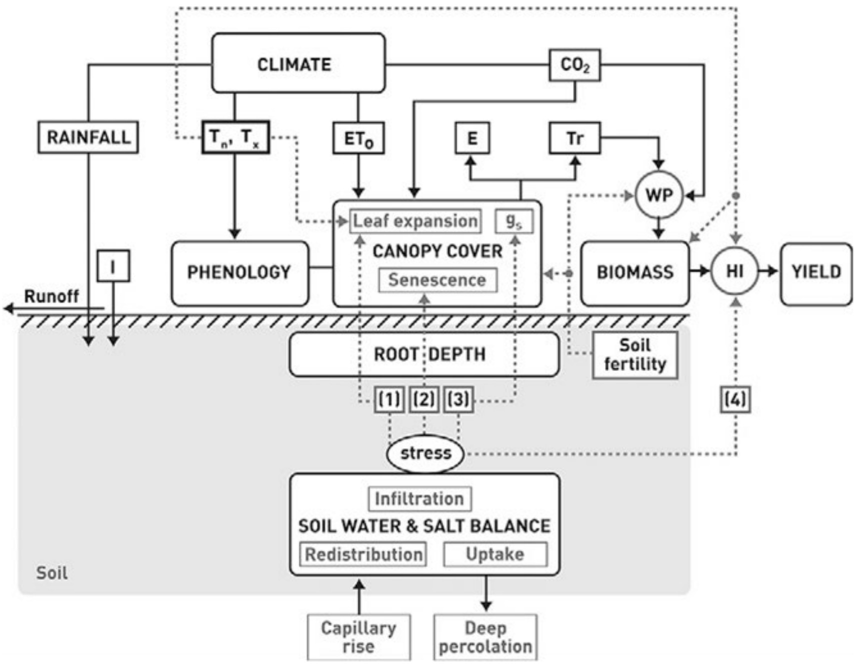


**Figure 2.6: Diagram showing AquaCrop's relationships between components (Steduto et al., 2009)**

The model outputs, inter alia, seasonal Y, ETa, crop water productivity (CWP = Y/ETa), crop cycle length and various stress factors related to cold temperature, water (stomatal conductance) and leaf expansion. A description of selected output variables is provided in **Table 8.1** in **APPENDIX B**.

Globally, AquaCrop is the third most used crop model after WOFOST and DSSAT (Lou et al., 2023). In South Africa, it is the most widely used crop model (Kephe et al., 2021), which is due to (i) the availability of crop parameters for most mainstream and underutilised crops, and (ii) the development of an automated procedure that runs the model at a national scale. Each aspect is discussed next in more detail.

### 2.2.2 Available crop parameters

Mabhaudhi et al. (2017b) identified 13 NUCs that should be prioritised for further research in South Africa. AquaCrop has been successfully calibrated and validated for seven of these 13 crops. These include: amaranth (Bello and Walker, 2017; Nyathi et al., 2018), bambara nut (Mabhaudhi et al., 2014a), cowpea (Kanda et al., 2021), sorghum (Hadebe et al., 2017), spider flower (Nyathi et al., 2018), OFSP (Beletse et al., 2013; Kunz et al., 2024) and taro (Mabhaudhi et al., 2014b; Kunz et al., 2024). In addition, AquaCrop has been calibrated for other underutilised crops, which include pearl millet (Bello and Walker, 2016) and Swiss chard (Nyathi et al., 2018). The model has also been calibrated for groundnut (Chibarabada et al., 2020a), although it is not considered an underutilised crop in South Africa (Mabhaudhi, 2023a; pers. comm.). Most of the above-mentioned research was funded by the Water Research Commission over the past decade.

### 2.2.3 National-scale crop modelling

AquaCrop can be run for 5 838 relatively homogenous response zones (called Altitude Zones) across southern Africa, i.e. South Africa, Lesotho and Eswatini. Each of the 1 946 quaternary catchments that were originally delineated by the Department of Water and Sanitation, have been disaggregated into three zones of relatively uniform altitude (Schulze and Horan, 2011). Since the altitudinal variation across each Altitude Zone (AZ) is minimal, the variation in climate and soils is also minimised (Kunz et al., 2024).

The daily rainfall and temperature datasets for each AZ were recently revised by Kunz et al. (2024). Rainfall data was obtained from 1 240 driver stations selected for all quaternary catchments. Monthly rainfall adjustment factors (ranging from 0.50 to 2.00) are used to improve the representativeness of average climate conditions across each AZ. Similarly, a driver temperature station was selected for each quaternary (543 stations in total) and a lapse rate adjustment was performed to account for the altitudinal difference between the temperature station selected for each quaternary and the average value for its three AZs. Daily $ET_O$ was calculated from the adjusted temperatures, assuming a default wind speed of 2 m $s^{-1}$ (Allen et al., 1998). Hence, each zone has representative climate data consisting of 50 years of daily rainfall, temperature (maximum and minimum) and $ET_O$ data from January 1950 to December 1999. This Altitude Zones climate database is sufficient to run the AquaCrop model (cf. **Section 2.2.1**).

Each zone also has representative soil data for both the A- and B-horizons. Clulow et al. (2023) updated the Altitude Zone soil database since each of the 7 082 land types was recently disaggregated into five terrain units. This produced 27 473 terrain unit polygons containing soil data for both soil horizons. For each AZ, these polygons were then used to provide area-weighted values for soil depth and soil water retention values, i.e. saturation (SAT), field capacity (FC) and permanent wilting point (PWP). Readily evaporable water (REW) was then calculated by Kunz et al. (2024) using the topsoil's FC and PWP values. Similarly, saturated hydraulic conductivity ($K_{SAT}$) was determined using soil water retention variables for both horizons using pedo-transfer function developed by Saxton and Rawls (2006). Finally, the curve number (CN) was derived using the estimated $K_{SAT}$ value for the topsoil (cf. Raes et al., 2018; Table 2.14d; p 2-165).

An automated procedure to run AquaCrop for all 5 838 AZs was initially developed by Kunz et al. (2015), which was subsequently improved by Kunz et al. (2020; 2024) to run faster and more efficiently. Despite these improvements, the model needs to be run on a high-end desktop PC (Kunz et al., 2024). The PC used for the national model runs has a Core i9 CPU with 10 cores, which facilitates 10 simultaneous model runs, each handling 538 AZs. Running the model in 10 parallel sessions is much faster than a single session involving 5 838 sequential simulations (Kunz et al., 2024).

## 2.3 Summary and Way Forward

The literature review identified various methods used to map land suitability that ranged from traditional methods involving simple overlays of spatial climate data using a GIS, to more complex modern methods involving machine learning, AHP and CSMs. The two main advantages of using traditional methods for land suitability assessment are as follows: (i) they are easier and less time consuming to apply, and (ii) they can provide relatively accurate results. The disadvantages are that the (i) thresholds for each predictor variable are not always known for a particular crop (especially for NUCs), (ii) the weightings of each variable are subjective, and (iii) some variables are not independent, as discussed previously (cf. **Section 2.1.3.3**).

The advantages of using more modern methods are that (i) they provide more accurate results, (ii) variable thresholds are not needed, and (iii) methods such as MaxEnt can provide objective (rather than subjective) weightings for each input variable. However, the disadvantages are that (i) they are more difficult and time consuming to implement, (ii) some methods (e.g. AHP) still require subjective weightings, and (iii) MaxEnt requires presence data that can be costly and time consuming to collect.

The case studies highlighted a number of important findings in land suitability assessments: (i) the type and number of predictor variables used can impact the accuracy of the land suitability map, (ii) no comprehensive attempt has been made to fully automate the different approaches (e.g. using *ModelBuilder* in ArcGIS) to develop a land suitability map, (iii) the issue of collinearity is not adequately addressed in most studies, (iv) the consideration of existing land use is important for developing realistic land suitability maps, (v) the inclusion of frost into land suitability maps is not accounted for in most studies, and (vi) yield data should be used to assess land suitability, not only to validate the final map.

**Section 2.1.2.3** highlighted the different crop growth engines used by CSMs and that water-driven models like AquaCrop are ideally suited for application in water-scarce environments where water availability strongly affects crop growth (cf. **Section 2.2**). AquaCrop has been calibrated and validated for seven prioritised NUCs for South Africa (cf. **Section 2.2.2**). More importantly, the model can be run at a national scale across southern Africa using a fully automated procedure (cf. **Section 2.2.3**). Lake (2022) was first to demonstrate the potential use of these national scale model runs for developing a land suitability map for taro (cf. **Section 2.1.2.3**). Kunz et al. (2024) then adopted and modified this approach to develop a land suitability map for OFSP and an updated map for taro. This study represents a continuation of this work, where the methodology was further improved and automated, before being applied to develop land suitability maps for another two NUCs.

# 3. MATERIALS AND METHODS

This chapter describes the approach taken to achieve each objective of this study, thus providing the following information: a description of the study area, an analysis of crop parameters files available for NUCs, which was then used to select two files. After the parameter files were verified, AquaCrop was run nationally to develop and extensive crop database. A preliminary analysis of the stored data was then completed to guide the approach followed to develop each land suitability map. Finally, the process to partially automate the methodology is explained.

## 3.1 Study Area Selection

This study covered the southern African region (South Africa, Eswatini and Lesotho), unlike Mugiyo et al. (2022) and Nhamo et al. (2022), who only considered the KwaZulu-Natal and Limpopo provinces, respectively. This was made possible since AquaCrop has been fully automated to run at a national scale (i.e. for all 5 838 AZs; cf. **Section 2.2.3**). Since South Africa is a water scarce country, expansion of irrigated areas is highly unlikely due to increasing water demands from other sectors (e.g. rising domestic use in response to the growing population). Therefore, there is a need to increase crop production under rainfed conditions. The region of particular interest is mainly the central and eastern parts of southern Africa where rainfed crop production is feasible, i.e. in AZs that receive more than 400 mm of mean annual precipitation (MAP) as shown in **Figure 3.1**.

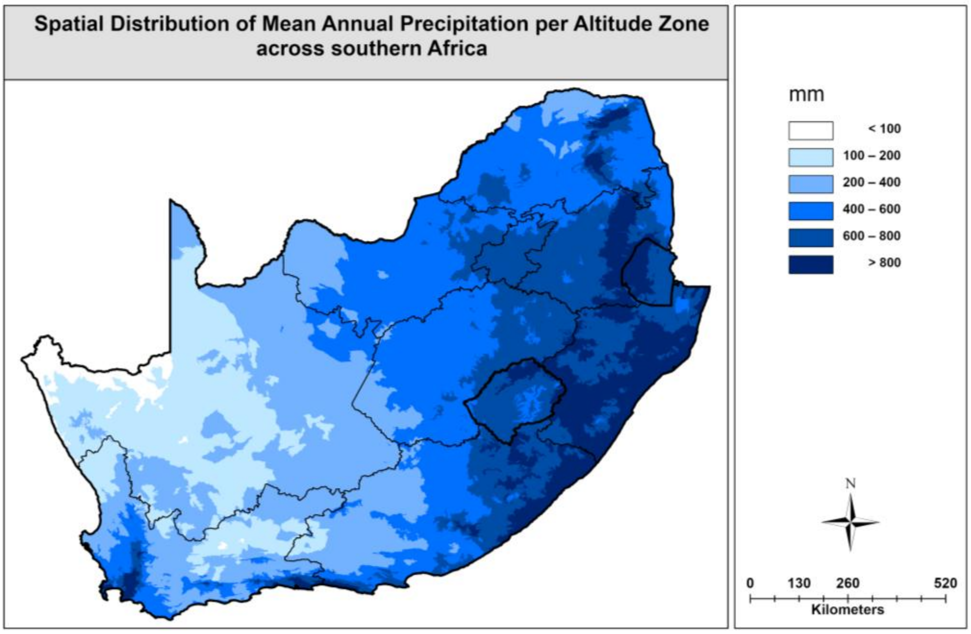


**Figure 3.1: Spatial distribution of mean annual precipitation per Altitude Zone across southern Africa (after Kunz et al., 2024)**

The annual rainfall threshold of 400 mm was proposed by Brouwer and Heibloem (1986) to identify areas where irrigation may be required for reliable crop production. The MAP data was obtained from the recently updated Altitude Zones climate database (Kunz et al., 2024) as described in **Sections 2.2.3** and **3.3.1**.

## 3.2 Crop Model Parameters

The ability of CSMs to adequately simulate canopy cover development, biomass accumulation and yield formation is crucial for accurate agricultural planning. Hence, model calibration and validation is evaluated using statistics such as root mean square error (RMSE) and its normalised variant (NRMSE), as well as coefficient of determination ($R^2$). RMSE quantifies the difference between predictions and observations of crop growth and yield. The index ranges from 0 to infinity, where 0 indicates perfect model performance (FAO, 2015). The following sections discuss how the various crop calibration and validation studies listed in **Section 2.2.2** were ranked, which was then used to select two NUCs that this study focused on. A verification was also done for the corresponding crop parameter files to ensure they contained the correct crop parameter values.

### 3.2.1 Ranking of calibration and validation studies

AquaCrop requires adequate calibration (and validation) of canopy cover (CC) development. **Table 9.1** (**APPENDIX C**) lists the above mentioned calibration statistics for CC development. Similarly, **Table 9.2** (**APPENDIX C**) provides the validation statistics of CC development for each crop. The Nash-Sutcliffe model efficiency coefficient and Wilmott's index of agreement were not considered since these statistics were not provided by all the calibration studies. Similarly, statistics for biomass production and final yield were also not provided by all the studies, and thus could not be used for ranking.

Assuming a well calibrated (and validated) model provides better estimates of crop growth and yield, the 14 studies were ranked from best to worst using NRMSE or RMSE, followed by $R^2$. The studies were also ranked by the number of water treatments used in the experiments, particularly for stressed conditions as these represent rainfed conditions, which the study is based on. A well calibrated and tested model across a wide range of water treatments is important, especially when simulating yields across an entire country with highly variable climatic conditions. Each study was also ranked on data availability and quality, as described next.

For the studies that only considered one water treatment, the calibration of OFSP by Beletse et al. (2013) was ranked last ($14^{th}$) since no RMSE statistics were provided and $R^2$ can be misleading with only four data points. Amaranth (Bello and Walker, 2017), pearl millet (Bello and Walker, 2016), OFSP (Nyathi et al. 2016) and sorghum (Hadebe et al., 2017) were all ranked 13$^{th}$ to 10$^{th}$ respectively, due to under-estimation of canopy cover as highlighted by high RMSE values (12.1-20.8%). Taro (Mabhaudhi et al., 2014b) and bambara nut (Mabhaudhi et al., 2014a) were ranked 9$^{th}$ and 8$^{th}$ respectively, due to the low RMSE values. In addition, $R^2$ for bambara nut was higher than that for taro.

For the studies with two water treatments, the calibration by Kunz et al. (2024) for OFSP and taro were ranked 7$^{th}$ and 6$^{th}$ respectively, both with high NRMSE and RMSE values due to under-estimation of canopy cover development during the initial growth stages. Amaranth, Swiss chard and spider flower (Nyathi et al., 2018) were ranked 5$^{th}$ to 3$^{rd}$ respectively, due to low RMSE values, especially for the severely water-stressed treatment, which better represents rainfed conditions (compared to the unstressed treatment).

The top two ranked crops were cowpea and groundnut, both having very good calibration statistics for four and three water treatments, respectively. Based on the validation of CC development, bambara nut and cowpea were the top two ranked crops, both exhibiting very good results under five and three water treatments, respectively.

### 3.2.2 Crop selection and description

Only the 13 underutilised crops that have been prioritised for further research in South Africa by Mabhaudhi et al. (2017a) were considered in this study. As a result, groundnut, Swiss chard and pearl millet were excluded. A limitation of AquaCrop (all versions up to 6) is its inability to accurately simulate leaf harvesting, which can be done up to four times in one season for leafy vegetables, (Nyathi et al., 2018). As a result, spider flower and amaranth were also excluded. Of the remaining crops, bambara nut and cowpea were selected since they were ranked the highest in terms of model validation (cf. **Section 3.2.1**). A brief description of each selected crop is given next.

Bambara nut and cowpea are legumes commonly grown across Africa and parts of Asia. These crops originated from West Africa and are typically grown in the tropics and subtropics. They are considered a dual-purpose crop since the both the pods and leaves can be consumed by humans and animals (DAFF, 2014; Kebede and Bekeko, 2020). Both crops are regarded as drought tolerant due to various water-stress response mechanisms. In addition, cowpea has a deep root system, which also limits water stress (Modi and Mabhaudhi, 2016), and thus is highly adaptable to marginal environments. Bambara nut produces pods beneath the soil surface, which is why it is often called bambara groundnut.

Bambara nut has a crop cycle ranging from 110 to 150 days (DAFF, 2016) and requires a minimum annual rainfall of 400-500 mm (Chibarabada, 2014), together with optimum temperatures of 20-28°C for crop development. Cowpea requires temperatures above 10°C for germination and between 21°C and 33°C for vegetative growth. For forage use, cowpea requires upwards of 750 mm of seasonal rainfall, whereas for human consumption, as little as 200 mm is required by the crop (Maundu et al., 2023). The typical growing period is around 100 days from sowing but depends on the planting date and climate.

### 3.2.3 Verification of crop parameter files

Crop parameter files for bambara nut and cowpea were obtained from the primary author of the respective calibration (and validation) study. For bambara nut, the climate, soil and irrigation files were also obtained for Roodeplaat (Pretoria). For cowpea, soil data was also obtained from the primary author, but the climate file was recreated using data measured by an automatic weather station at Ukulinga (Pietermaritzburg). The parameter files were then verified by running AquaCrop for optimum conditions using the climate and soil datasets obtained for each calibration site.

The verification run for bambara nut produced a simulated yield of 2.360 dry t $ha^{-1}$, which compared favourably with the simulated yield reported by Mabhaudhi et al. (2014a) of 2.370 dry t $ha^{-1}$ (i.e. yield deviation of 0.4%). Similarly, a simulated yield of 2.806 dry t $ha^{-1}$ was obtained for cowpea, which was almost identical (i.e. small deviation of only 0.07%) when compared to the yield reported by Kanda et al. (2021). Both deviations were well within 6%, which according to Dua et al. (2014), is within acceptable limits. This exercise was important since it confirmed the correct crop parameter files were used in this study for the national model runs.

## 3.3 Crop Model Inputs

To run AquaCrop for each AZ, key inputs are required such as climate and soil data, including plant density and planting date. A brief description of each input dataset and its source is given next.

### 3.3.1 Climate data

AquaCrop requires a climate (.CLI) file, which provides the name (and location) of the annual atmospheric $CO_2$ (.CO2) file, including the daily rainfall (.PLU), air temperature (.TNX), reference evapotranspiration (.ETO) files (cf. **Section 2.2.1**). The default .CO2 file packaged with AquaCrop version 6 was used, which contains annual values measured at the Mauna Loa station in Hawaii. Using the Altitudinal Zone climate database described in **Section 2.2.3**, Kunz et al. (2024) created the .PLU, .TNX and .ETO files in the exact format required by AquaCrop, allowing the model to be run for each AZ with 50 years of input data, thus providing 49 consecutive seasons of simulated data. The reader is referred to Kunz et al. (2024) for a more detailed description of the Altitude Zones climate database.

### 3.3.2 Soil data

Soil data for two soil horizons has been developed for each of the 5 838 AZs (cf. **Section 2.2.3**). Kunz et al. (2015; 2020) and Kunz and Mabhaudhi (2023) ran AquaCrop at a national scale using soil data available for both soil horizons, whereas Kunz et al. (2024) ran the model for a single (depth weighted) soil layer. This was done because (i) AquaCrop typically simulates yield more accurately when using a single soil profile (Mabhaudhi, 2023b; pers. comm.), and (ii) most of the underutilised crops have been calibrated and validated in South Africa using a one-layer soil profile. An analysis was therefore conducted to determine whether a two-layer soil profile would provide yield simulations different to those obtained using a one-layer (depth weighted) soil.

Groundnut was the only crop calibrated by Chibarabada et al. (2020a) using a two-layer soil at Umbumbulu. AquaCrop was run using climate data obtained from SASRI[1] (Station ID 459, Powerscourt-Roseleigh) from 30 November 2016 to 31 August 2017 and the results showed that there was no yield difference between the one- and two-layer soils. Climate and soil data for AZ number 4 780, which represents the Umbumbulu region, was also used, which again showed no difference in yield between a one- and two-layer soil profile.

Castañeda-Vera et al. (2015) simulated SWC using four crop simulation models for a soil with three horizons. The study found that model performance was attributed to the number of layers in which the soil profile is subdivided within the model: 1 layer in WOFOST, 5 in CERES, 12 in AquaCrop and 16 in CropSyst. The models with multiple soil layers resulted in more even changes in soil water content compared to WOFOST with only a single layer. When soil data from multiple layers is depth-weighted to a single layer, rooting distribution is no longer considered, which may explain the differences in yields simulated by the four models. Based on this evidence, using homogenous (one layer) soil data in crop models should be avoided. For this study, the national scale model runs (cf. **Section 3.4.1**) were therefore undertaken with two-layer soil data.

[1] SASRI: https://sasri.sasa.org.za/pls/sasri/r/weatherweb/home/5?session=9039703554804

### 3.3.3 Planting date and density

Kunz et al. (2020; 2024) showed that for crops such as sorghum, soybean, OFSP and taro, varying the planting date had a greater impact on simulated crop yield than different plant densities. For the national model runs, this study therefore considered four planting months for each crop (October, November, December and January), each with the same plant density that represents a smallholder farming system. The planting day was set to the first of each month as suggested by Kunz et al. (2020; 2024).

Cowpea has various cultivar types (e.g. erect, semi-erect and prostrate/runner) that require different plant densities for optimising yield. Erect and semi-erect types can be planted more densely as they tend to grow upwards (DAFF, 2014). The brown mixed variety considered in this study is a prostrate/runner type (AGT Foods Africa, 2018), which requires a lower plant density as it grows horizontally, i.e. across the ground surface. For the runner type, plant density ranges from 60 000-133 333 plants ha$^{-1}$ (DAFF, 2014; Kanda et al., 2021). A density of 80 000 and 133 333 plants ha$^{-1}$ is recommended for smallholder and commercial farmers, respectively (Mabhaudhi, 2023b; pers. comm.). A plant density of 80 000 plants ha$^{-1}$ was selected in this study to represent smallholder farming systems, which matches the value used by Kanda et al. (2021) for the Ukulinga experiment.

Bambara nut is a short crop with lateral stems (Mabhaudhi, 2012), with recommended plant densities of 74 000-166 667 plants ha$^{-1}$ for smallholder farmers and up to 222 222 plants ha$^{-1}$ for commercial farmers (Mkandawire and Sibuga, 2002; DAFF, 2016; Mabhaudhi, 2023b; pers. comm.). A planting density of 166 667 plants ha$^{-1}$ was used by Mabhaudhi et al. (2014a) for the calibration and validation of bambara nut in South Africa. Thus, the same plant density was used in this study.

### 3.3.4 Maturity date

For some AZs, AquaCrop crashes when simulating a particular season due to a division-by-zero error (for example), and thus produces no output, which results in less than 49 simulated seasons. If less than four of the 49 seasons are simulated, there is insufficient data for calculating long-term statistics, and thus the AZ is flagged as -999 to indicate missing data.

AquaCrop calculates the end of the crop cycle using thermal time, i.e. when sufficient heat units have been accumulated for the crop to reach physiological maturity (Raes et al., 2018). If the crop cycle length (from emergence to physiological maturity) exceeds 365 days, AquaCrop is not run since the seasonal climate is deemed too cold for economically viable crop production (Kunz et al., 2024). Again, -999 is used to indicate no data is available for the

AZ. However, this approach does not consider the occurrence of frost, which can shorten the crop cycle length.

A study by Wimilasiri et al. (2022) showed that none of the 13 underutilised crops prioritised for further research in South Africa by Mabhaudhi et al. (2017a) have frost adaptive traits. This highlights the importance of considering the impact of frost occurrence when simulating crop yield of NUCs. Using data collected at the University of KwaZulu-Natal in Pietermaritzburg (29.628°S, 30.403°E; altitude 671 m), a relationship was developed between minimum surface temperature (measured using an infrared thermometer) and minimum air temperature at 2 m above ground level. The analysis considered daily data from December 2019 to August 2024 and focused on the late autumn and early winter months (April to August), when frost is most likely to occur in the summer rainfall region of southern Africa. The results showed that when the surface temperature reaches 0°C, the air temperature at 2 m is approximately 5.4°C (**Figure 3.2**).

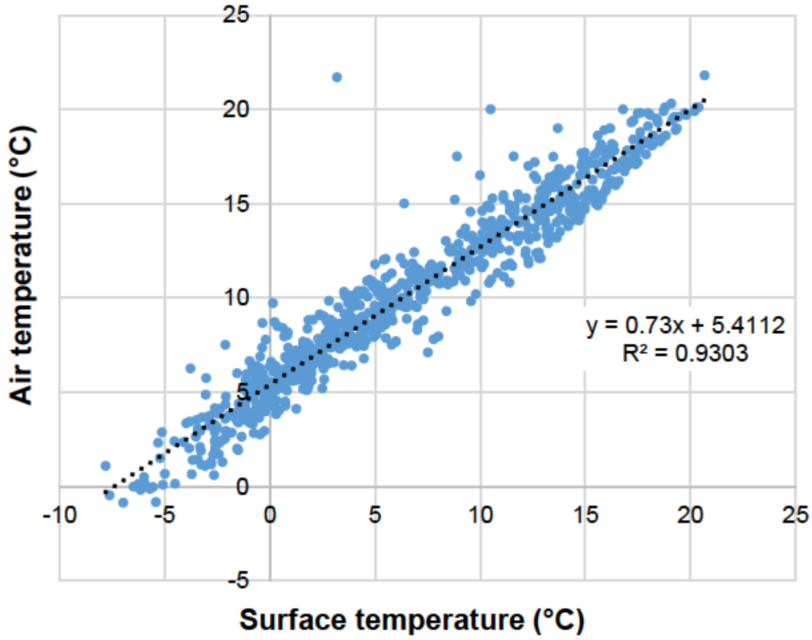


**Figure 3.2: Relationship between minimum air and surface temperature for the April to August period using data measured by an automatic weather station from 2020 to 2024 in Pietermaritzburg**

For this study, the assumption is made that when the daily minimum air temperature drops to 5.4°C or less, the crop is killed by frost (due to plant cell and tissue damage). Schulze (2008) noted that for frost to occur, there needs to be clear nights and low winds, as well as relatively dry air. Thus, the assumption was also made that these conditions would be met every day for frost. Thus, crop growth and yield formation is halted and the season is effectively ended, resulting in reduced yield or complete crop failure. However, if this temperature is reached before the crop has fully emerged, then frost occurrence has no effect and the crop cycle length is not ended prematurely. As noted earlier, if no frost occurs, then the crop cycle length is based on thermal time.

### 3.3.5 Management options

The management file in AquaCrop was set up to reflect: (i) no soil fertility stress, (ii) no weed growth, and (iii) no soil bunds to reduce surface runoff. Hence, model runs were undertaken for optimal field management conditions (i.e. stress-free growing conditions) where yield potential was mostly limited by soil water availability.

## 3.4 Development of the Crop Database

As outlined in **Section 2.2.3**, a fully automated procedure has been developed to simulate crop growth and yield using AquaCrop for 5 838 AZs across southern Africa. The reader is referred to Kunz et al. (2024) for further information regarding this achievement. For this study, it facilitated the development of a crop database containing long-term statistics for 25 variables outputted by the model, which was critical for identifying AZs deemed suitable for rainfed production of bambara nut and cowpea.

### 3.4.1 AquaCrop national runs

AquaCrop was run nationally on a high-end PC by completing two simple steps. Firstly, an ASCII text file called CRPVAL (CRoP VALues) was created that contained information regarding (i) which crop parameter file to use, (ii) the required simulation period (i.e. 1950-1999), (iii) the desired planting date and density, and (iv) the minimum air temperature at 2 m when frost could occur at the ground surface (e.g. 5.4°C). The second step involved starting a terminal window in Linux in which the RUN command was executed. This single command started the fully automated procedure, which ran AquaCrop for all 5 838 AZs as 10 parallel sessions, each with 538 zones, starting with cowpea and then bambara nut.

The above process was repeated for both bambara nut and cowpea, each with four planting dates. The model was also run without accounting for frost occurrence by setting the frost temperature to 99.0°C. This was done for each crop and four planting dates to assess the reduction in crop cycle length that may occur due to frost. In total, 16 national model runs were completed, which took approximately 80 hours to complete over eight days.

### 3.4.2 Generation of statistics

AquaCrop simulated 49 consecutive seasons of crop data for each AZ, from which long-term monthly and seasonal statistics (e.g. sum, mean and inter-seasonal variability) were automatically generated for 25 (of 30) output variables, e.g. mean relative biomass (BREL), yield and crop cycle length. Of these, 17 output variables were considered potentially useful for identifying AZs deemed suitable for crop production (cf. **APPENDIX B**). The number of seasons with a simulated yield of 0 dry t $ha^{-1}$ was also summed, which was then used to calculate another variable called the risk of crop failure (cf. **Section 3.6.1**). For each national run, numerous .csv files were generated by the automation procedure, i.e. each containing one statistic for a single output variable for all 5 838 AZs.

## 3.5 Data Analysis

A comprehensive preliminary assessment of the modelled data was conducted to identify potential issues, such as data anomalies that could affect the reliability of the land suitability analysis. The identification of spatial patterns and the collinearity tests between output variables are discussed next.

### 3.5.1 Identification of spatial patterns

Output from the national model runs were imported into ArcGIS Pro (v3.3; ESRI, 2023b) as a table and joined to the AZs shapefile (developed by Schulze and Horan, 2011) to generate individual maps for each output variable. The maps were visually inspected to identify spatial patterns or anomalies. For example, no cold temperature stress (TmpStr; cf. **APPENDIX B**) data was simulated for bambara nut during certain growth stages, and thus could not be used to assess land suitability for crop production.

For each AZ, an analysis was performed to identify the planting month with the highest yield. This was done using the Python programming language (v3.11.9; Python Software Foundation, 2023) and the Visual Studio Code editor (v1.94.2; Microsoft, 2024). 138 lines of code was written that extracted the AZ number and corresponding yield value for each planting date. If two months had the same (or very similar) yield, the earlier planting date was chosen to minimise the risk of frost occurring towards the end of the crop cycle.

An analysis was also conducted to determine the impact of frost occurrence on crop cycle length. For each planting month, the reduction in crop cycle length was determined by comparing simulations with and without frost consideration.

### 3.5.2 Testing for collinearity

As discussed in **Section 2.1.3.3**, collinearity is an important aspect to consider when developing land suitability maps. Pearson's correlation coefficient (r) is widely used as a statistical measure to evaluate the strength of the linear relationship between two variables, with values ranging from 0 (no linear relationship) to 1 (a perfect relationship) (Profillidis and Botzoris, 2019). The absolute value of r (|r|) was used to avoid any negative values. Approximately 200 lines of Python code were developed to test for collinearity amongst 17 variables, which automatically produced 136 relationships for each planting month, i.e. 1 088 tests for both crops. Nafis and Alvi (2023) stated that if the coefficient r is 0.8 or higher, then there is very strong collinearity between the two variables (**Table 3.1**), which concurs with Shrestha (2020).

**Table 3.1: Collinearity strength determined from the Pearson correlation coefficient (r) (Nafis and Alvi, 2023)**

| \|r\| | Collinearity strength |
|---|---|
| ≥ 0.80 | Very strong |
| 0.60 – 0.79 | Moderately strong |
| 0.30 – 0.59 | Fair |
| 0.01 – 0.29 | Poor |
| 0 | No correlation |

## 3.6 Land Suitability Assessment

As noted in **Section 2.1.2.3**, the three-step approach initially developed by Lake (2022) to map land suitability for taro production was modified by Kunz et al. (2024) and also applied to OFSP. To recap, the first step involved eliminating AZs that are too cold for economically viable crop cultivation. These zones typically have no simulated output (flagged as -999) because the crop cycle length exceeds 365 days, as explained in **Section 3.3.4**. Zones were also eliminated if they were considered too dry (i.e. low MAP), exhibited a high risk of crop failure, the climate was too variable or crop water use was too inefficient. The second step focused on the remaining suitable areas, which were classified from marginal to high suitability using CWP. The thresholds that were applied to each variable were obtained from available literature and expert opinion. The third step involved removing AZs that were classified as current "no-go" areas (i.e. N1) and permanent "no-go" areas (i.e. N2) using available national land cover data for South Africa.

It is important to note that for this study, the methodology adopted by Kunz et al. (2024) was further revised and improved, as described in the following sections. The same three-step approach used by Lake (2022) and Kunz et al. (2024) was again followed in this study.

### 3.6.1 Elimination of unsuitable areas

As described in **Section 3.5.2**, collinearity tests were conducted between all output variables to determine which pairs were strongly related. All collinear variables were then excluded from further analysis. The goal was to identify the fewest number of variables that successfully eliminated unsuitable zones for rainfed crop production. It involved a lengthy trial and error process that identified variables similar to those used by Kunz et al. (2024). For example, AZs were eliminated if (i) the crop cycle length exceeded 365 days, or (ii) less than four (of 49) seasons were simulated, and thus no statistics were calculated. These higher altitude zones are deemed too cold for crop production and are flagged as -999 in the output (.csv) files.

The first modification to the methodology used by Kunz et al. (2024) involved the elimination of AZs deemed too dry for rainfed crop production. Kunz et al. (2024) used an MAP threshold of 400 mm, which does not consider rainfall seasonality (i.e. summer vs winter vs year-round rainfall; Roffe et al., 2021). Instead, a threshold of 200 mm was applied to the $RFL_{AVE}$ variable (cf. **APPENDIX B**) to eliminate zones considered too dry for rainfed production (cf. **Section 3.2.2**). $RFL_{AVE}$ represents the average rainfall accumulated over the 49 growing seasons. For each crop, this was repeated for all four planting months.

The second modification involved the RCF calculation (cf. **Section 2.1.2.3**). Kunz et al. (2024) calculated this metric as the ratio of the number of seasons with a zero yield ($Y_{ZERO}$) to the total number of simulated yields ($Y_{NUMB}$), expressed as a percentage (**Table 3.2**). However, this definition does not consider seasons with no output because the model crashed (cf. **Section 3.3.4**). For this study, the assumption was made that if the model crashed, this was equivalent to a zero yield simulation. Hence, the number of missing yield simulations ($Y_{MISS}$) (i.e. with no data) was calculated as 49 – $Y_{NUMB}$ and added $Y_{ZERO}$ (cf. **APPENDIX B**), which was then divided by the maximum number of seasons (i.e. 49), instead of the number of successfully simulated seasons (**Table 3.3**). This prevented RCF from being under-estimated, especially for zone no. 3 289. A RCF threshold of 33%, which represents a 1 in 3-year risk of crop failure, was then used to eliminate unsuitable AZs. This threshold may not be acceptable to all smallholder farmers and is likely too high for commercial farmers.

**Table 3.2: Risk of crop failure (RCF) calculation used by Kunz et al. (2024)**

| Altitude Zone no. | $Y_{ZERO}$ | $Y_{NUMB}$ | RCF (%) |
|---|---|---|---|
| 3 242 | 19 | 42 | 45.2 |
| 3 289 | 1 | 18 | 5.6 |

**Table 3.3: Revised risk of crop failure (RCF) calculation used in this study**

| Altitude Zone no. | $Y_{ZERO}$ | $Y_{MISS}$ | Maximum no. of seasons | RCF (%) |
|---|---|---|---|---|
| 3 242 | 19 | 7 | 49 | 53.1 |
| 3 289 | 1 | 31 | 49 | 65.3 |

Similar to Kunz et al. (2024), AZs were eliminated if the coefficient of variation in CWP exceeded 150%, since a high variation in seasonal CWP indicates the climate is highly variable, and thus unsuitable for crop production. The inter-seasonal variation in yield was not used as it was found to be strongly correlated with seasonal rainfall ($RFL_{AVE}$).

The RCF and $CWP_{CV}$ thresholds were also determined using a trial and error approach. An iterative process was used to incrementally increase each threshold value to successfully eliminate unsuitable AZs in the western parts of the country, i.e. zones where rainfed crop production is highly unlikely. When viable production areas in the eastern parts of the country started being eliminated, the iterative process was halted and the threshold value was set.

All eliminated zones were classified as currently unsuitable (N1) for crop production according to the FAO (1976) method (cf. **Section 1.1**). The remaining (i.e. not eliminated) AZs were considered suitable for rainfed production of bambara nut or cowpea production and classified into three suitability classes as described next.

### 3.6.2 Classification of suitable growing areas

The FAO (1976) method defines three suitability classes, namely highly suitable (S1), moderately suitable (S2) and marginally suitable (S3). The average relative biomass ($BREL_{ave}$) was used to classify the remaining AZs into one of these suitability classes. This output variable represents the ratio of actual biomass to potential biomass for unstressed conditions, expressed as a percentage (cf. **APPENDIX B**). For each zone, $BREL_{AVE}$ was normalised using the highest value that represented the "best" zone in the country, where biomass production was least impacted by temperature and/or water stress. This allowed for the same thresholds to be used for all crops, such that AZs with $BREL_{NORM}$ above 75% were classified as S1, between 50% and 75% as S2, and between 25% and 50% as S3. For zones where $BREL_{NORM}$ was below 25%, they were classified as currently unsuitable (N1) as less than a quarter of the biomass potential was reached.

### 3.6.3 Verification of land suitability maps

To verify the accuracy of the generated land suitability maps, towns (as points) where cowpea is grown were obtained from the Department of Agriculture, Forestry and Fisheries (DAFF) production guidelines (DAFF, 2014). Similarly, districts (as polygons) in which bambara nut are produced were obtained from DAFF (2016). The centroid of each district polygon was determined using Google Earth Pro (v7.3.6; Google LLC, 2023) and exported as .KML files. These files were subsequently imported into ArcGIS Pro (v3.3; ESRI, 2023b), where they were converted into point shapefiles, then overlaid onto the land suitability maps to assess the spatial alignment.

### 3.6.4 Consideration of existing land use

**Section 2.1.3.2** highlighted the importance of incorporating existing land use data in land suitability assessments, considering suitable areas can be reduced by up to 22% (Kunz et al., 2024), thus preventing over-estimation of suitable cropping areas. Similar to Kunz et al. (2024), this study excluded land uses that are permanently unsuitable (N2) for crop production, i.e. mining, urban, wetlands, water bodies and protected areas. Currently unsuitable (N1) areas were also excluded, such as indigenous forests, commercial plantations, sugarcane production areas and orchards. Indigenous forests were excluded because the National Forests Act (no. 84 of 1998) prohibits the clearing of indigenous trees in any natural forest without a licence issued by the Department of Forestry, Fisheries and the Environment (DFFE). Commercial plantations and sugarcane areas were also excluded because these highly established industries are unlikely to convert to indigenous crop production. Similarly, orchards take years to bear fruit, and thus are also unlikely to become indigenous crop cultivation areas.

This study used the 2018 national land cover (NLC) dataset obtained from SAEON[2], together with the latest (2024) protected areas data from DFFE[3]. The latter dataset was imported into ArcGIS Pro and converted from vector to raster format using a grid cell size of 1 $km^2$ (to match the resolution of the NLC dataset). The extent of land cover within each grid cell (or pixel) was then determined before it was merged with the NLC raster. If the land cover exceeded 50% of the pixel's area, it was assumed the entire pixel had the same land use. These pixels were then reclassified as either 1 (N1) or 2 (N2). The suitable land areas were also converted from vector to raster format with a 1 $km^2$ resolution and reclassified a 3 (unsuitable production), 6 (S1), 9 (S2) or 12 (S3), before they were merged with the land cover raster. The final suitability maps were then produced by removing the unsuitable crop production areas and unsuitable land uses (N1 and N2).

---

[2]SAEON: https://catalogue.saeon.ac.za/records
[3]DFFE https://egis.environment.gov.za/data_egis/data_download/current

## 3.7 Automation of Methodology

From **Section 1.2**, an objective of this study was to partially automate the methodology to produce each land suitability map. The automation process is briefly outlined below:

(i) Folders were created to separate the input .csv files (containing statistics for each AquaCrop output variable; cf. **Section 3.4.2**), output files (e.g. collinear graphs and land suitability .csv files) and Python scripts.

(ii) As mentioned in **Section 3.5.1**, certain AquaCrop output variables such as RCF, $RFL_{AVE}$ and $Y_{AVE}$ were mapped individually to identify any anomalies. About 300 lines of Python code was written to automate the development of these .csv files, especially the calculation of RCF.

(iii) From **Section 3.5.2**, a total of 1 088 collinearity tests were performed for both crops (since each has four planting months). About 200 lines of Python code were developed to automate this repetitive, yet important process.

(iv) As mentioned in **Section 3.5.1**, 138 lines of code was written to determine which planting date gave the highest $Y_{AVE}$.

(v) For the elimination phase, 120 lines of code were written to exclude zones that were deemed unsuitable for crop production. Using all four variables highlighted in **Section 3.6.1**, the code produced an output .csv file containing the AZ number and a corresponding suitability index: 0 (no data), 1 (unsuitable) or 2 (suitable).

(vi) For classifying the remaining suitable areas (cf. **Section 3.6.2**), a further 100 lines of code were written to automate the process of normalising BREL, then using it to classify the suitable AZs as either currently unsuitable (0; N1), highly (1; S1), moderately (2; S2) or marginally (3; S3) suitable for rainfed crop production. The code therefore creates an output .csv file with values ranging from 0 to 3.

(vii) The final .csv file was imported into ArcGIS Pro for mapping.

This automation was considered very important as it facilitated the trial and error approach in identifying the most suitable AquaCrop output variables for eliminating and classifying AZs as well as for determining the thresholds for each of these variables. It also reduced the time required to produce each of the eight land suitability maps. Owing to time constraints, no attempt was made to automate the exclusion of unsuitable crop land based on existing land use, as described in **Section 3.6.4**. However, this should be considered in future work.

# 4. RESULTS AND DISCUSSION

As explained in **Section 3.5**, the preliminary data analysis on the planting date that produced the highest yield in each AZ, including the impact of frost occurrence on crop cycle length is presented first, followed by the collinearity tests. Thereafter, the results of the land suitability assessment for crop cultivation are shown.

## 4.1 Data Analysis

### 4.1.1 Planting date with highest yield

Determining the optimal planting date for farmers is essential for maximising crop growth and yield. For photosensitive crops, early sowing can result in delayed flowering, increased leaf production and reduced yield (Ilunga, 2014; Kamara et al., 2018). However, sowing late in the season can result in:

(i) early onset of podding and lower yield (Chibarabada et al., 2020b), or
(ii) crop failure due to (i) inadequate rainfall and high water stress (Kamara et al., 2018), and/or (ii) frost damage (Ilunga, 2014).

#### 4.1.1.1 Bambara nut

**Figure 4.1** shows that planting date has a substantial impact on bambara nut yield across the country. The white areas have no data (i.e. flagged as -999 in the statistics) as these zones are either too cold (crop cycle length > 365 days) or have less than four seasons of simulated yield, i.e. too few for calculating meaningful statistics (cf. **Section 3.3.4**). Excluding the drier western parts of the country where rainfed crop production is unviable (cf. **Section 3.1**), planting bambara nut in October should produce the highest yield across the majority of the central and eastern parts of the country. **Figure 10.1** (cf. **APPENDIX D**) shows the difference in $Y_{AVE}$ between October and November for each AZ. The October planting produces a higher $Y_{AVE}$ than November across 27.33% of southern Africa. However, $Y_{AVE}$ for October and November is identical across 9.49% of southern Africa. As stated in **Section 3.5.1**, when two planting months have the same yield, the earlier month was chosen. Further analysis showed that for the majority of these AZs, $Y_{AVE}$ was zero dry t ha$^{-1}$ for both planting dates, which explains why there was no yield difference in many AZs.

In contrast, **Figure 10.2** (cf. **APPENDIX D**) shows that although $Y_{AVE}$ in November is higher than October across 54.62% of southern Africa, the difference is less than 0.1 dry t ha$^{-1}$ for most (45.02%) of the area, i.e. the yield gain is marginal when planting one month later. However, the majority of these AZs are situated in the central to western half of the country, which is considered unviable for rainfed crop production (cf. **Section 3.1**).

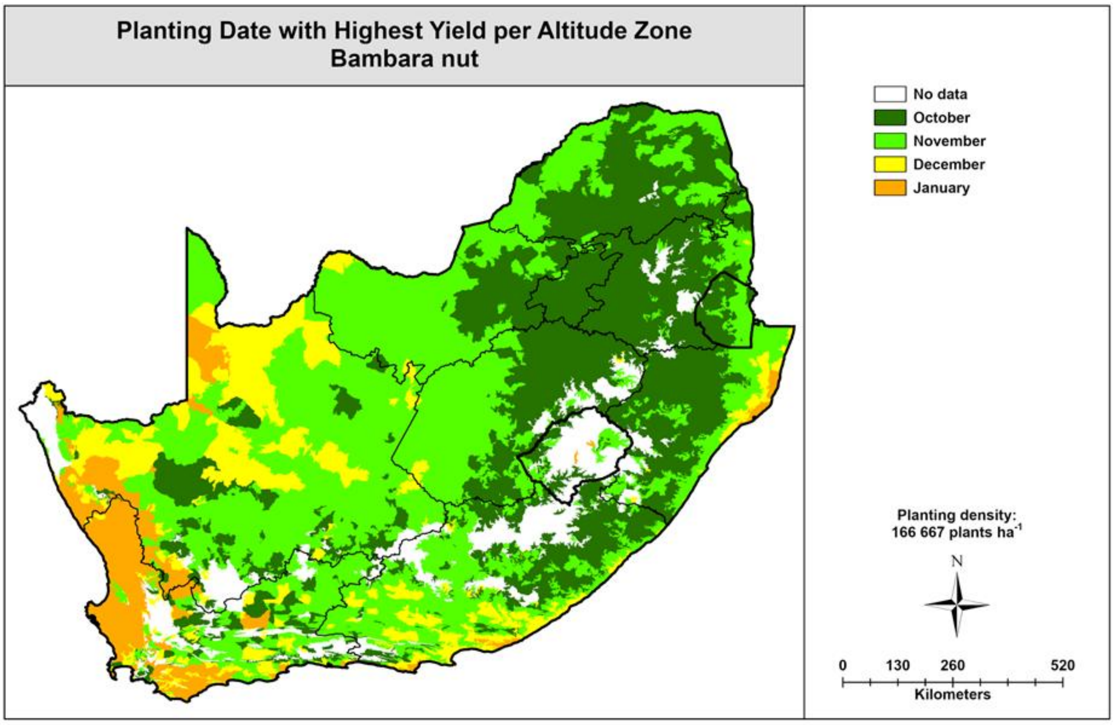


**Figure 4.1: Planting date that produced the highest simulated yield for bambara nut per Altitude Zone**

Although the difference in spatially averaged yields over the entire country between October and November is marginal (0.05 dry t ha$^{-1}$), values steadily decrease towards the January planting (**Table 4.1**). The results highlight that an October to November planting is preferable for optimising crop production, which concurs with the literature (DARDLA, 2012b; DAFF, 2016). Chibarabada et al. (2020b) showed that later plantings (i.e. December and January) resulted in decreased yield due to early onset of podding of bambara nut.

**Table 4.1: Spatially averaged yields for bambara nut across southern Africa for each planting month**

| Yield (dry t ha$^{-1}$) | | | |
|---|---|---|---|
| **Oct** | **Nov** | **Dec** | **Jan** |
| 0.60 | 0.55 | 0.38 | 0.21 |

One of the most important factors influencing the planting date of bambara nut is photoperiod response. As the photoperiod reduces, this triggers the change from vegetative to reproductive growth (Chibarabada et al., 2020b; Kendabie et al., 2020). Steduto et al. (2012) also highlighted that when yield is not limited by water stress, it is governed by daylength. Hence, it is important to determine the optimal sowing date for maximising pod growth. The maximum daylength for October to April ranges from 11.10 to 14.35 hours, with values increasing with latitude and from October to December (**Table 4.2**). Bambara nut is a short-day crop, requiring daylengths of 12 hours or less for maximum pod-set and seed yield. Daylengths longer than

14 hours do not produce pods and delay the onset of maturity (Steduto et al., 2012; Mayes et al., 2019), which results in reduced yields. Flowering typically starts about three months after planting when the bambara nut is most sensitive to daylength. For earlier plantings (October to November), flowering therefore occurs in January to February when daylengths are between 12.74 and 14.09 hours (**Table 4.2**), thus resulting in reduced pod formation.

Planting later in December to January means flowering occurs in March to April when daylengths are shorter (11.10-12.24 hours; **Table 4.2**). This results in more pods and higher yields. Thus, in terms of maximising yields based on daylength, the crop is best suited to a January planting when daylengths are shortest at flowering. However, the crop cycle length can be reduced by frost occurring from late autumn onwards. Since AquaCrop is a water-driven model, it does not consider daylength for photoperiod sensitive crops. This means that yield estimates for October, November and potentially December are likely to be over-estimated, due to the longer daylengths during flowering. Similarly, AquaCrop does not consider frost occurrence for frost sensitive crops, and thus yields can also be over-estimated due to unrealistically long cropping season lengths (cf. **Section 4.1.2**). When considering the impact of temperature in the model, cold temperature stress reduces crop growth whenever the temperature drops below the crop's base temperature, whereas the addition of frost kills the crop, thus halting crop growth altogether. The project (.PRM) file instructs AquaCrop to stop growth when the minimum air temperature drops below the frost threshold.

**Table 4.2: Maximum daylengths from October to April for the northern (Limpopo), central (Free State) and southern (Cape Town) regions of South Africa**

| Latitude | Province | Oct | Nov | Dec | Jan | Feb | March | April |
|---|---|---|---|---|---|---|---|---|
| -22 | Limpopo | 12.54 | 13.08 | 13.34 | 13.20 | 12.74 | 12.14 | 11.48 |
| -23 | (Musina) | 12.56 | 13.13 | 13.41 | 13.26 | 12.78 | 12.14 | 11.46 |
| -28 | Free State | 12.71 | 13.42 | 13.77 | 13.58 | 12.98 | 12.18 | 11.32 |
| -29 | (Bloemfontein) | 12.74 | 13.48 | 13.85 | 13.65 | 13.02 | 12.19 | 11.29 |
| -34 | Western Cape | 12.90 | 13.81 | 14.26 | 14.01 | 13.25 | 12.23 | 11.14 |
| -35 | (Cape Town) | 12.93 | 13.88 | 14.35 | 14.09 | 13.29 | 12.24 | 11.10 |

#### 4.1.1.2 Cowpea

From **Figure 4.2,** November and December are the most suitable planting dates for cowpea in the central and eastern parts of the country. This corresponds with the South African production guidelines for cowpea of late November to early December (DARDLA, 2012a; DAFF, 2014), with a later planting typically producing better yield quality (Kamara et al., 2018). From **Figure 10.3** and **Figure 10.4** (cf. **APPENDIX D**), there was a marginal difference in $Y_{AVE}$ (0-0.1 dry t ha$^{-1}$) between the November and December planting months for most of southern

Africa, which is similar to the small yield difference between October and November for bambara nut (cf. **Section 4.1.1.1**).

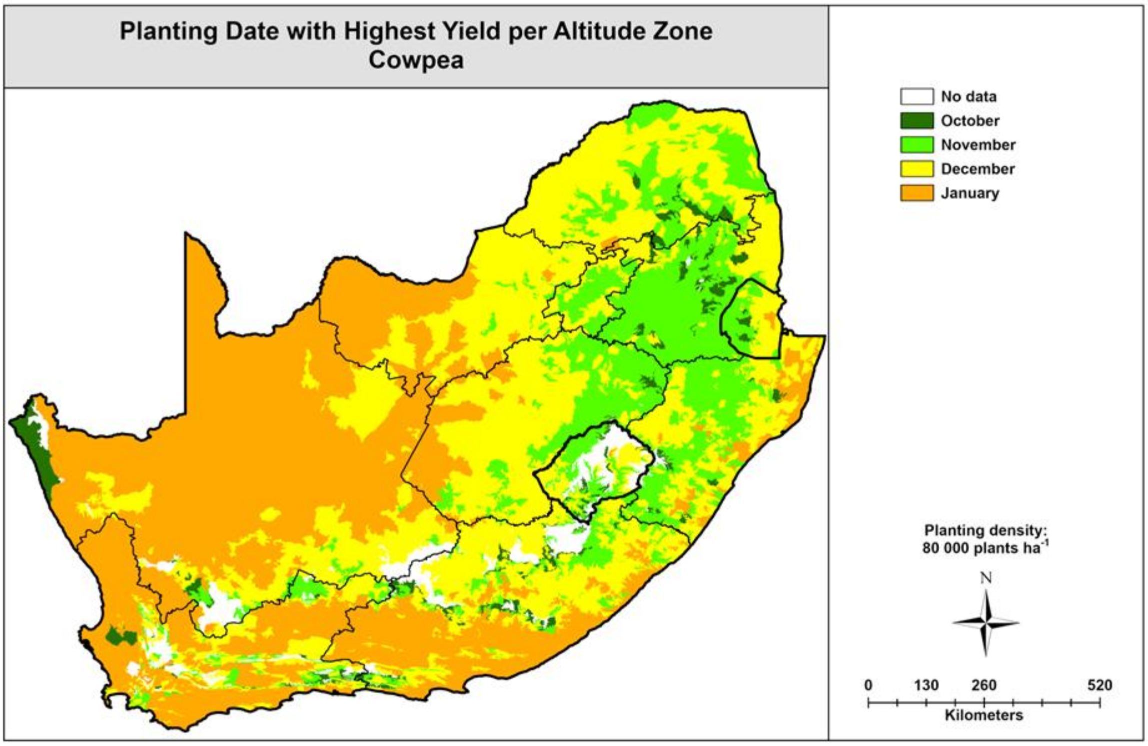


**Figure 4.2: Planting date that produced the highest simulated yield for cowpea per Altitude Zone**

The spatially averaged cowpea yields for each planting month again highlight little difference between the November and December plantings (**Table 4.3**). Cowpea sown at the onset of rains are predisposed to pests and diseases (Kamara et al., 2018). Mbong et al. (2010) showed that early sown cowpea (October) had higher incidences of scab infection and that yield quality and quantity was poor compared to cowpea sown at later dates (November/December). Since AquaCrop cannot account for pests or diseases, cowpea yields for the October planting are therefore likely to be over-estimated.

**Table 4.3: Spatially average yields for cowpea across southern Africa for each planting month**

| Yield (dry t ha$^{-1}$) | | | |
|---|---|---|---|
| **Oct** | **Nov** | **Dec** | **Jan** |
| 0.33 | 0.46 | 0.46 | 0.32 |

#### 4.1.1.3 Timing of crop stress

It is important to consider the impact of stress (water and temperature) during each phenological growth stage. Both NUCs are sensitive to stress during the flowering stage and pod filling stage (Ahmed and Suliman, 2010; Iheanacho et al., 2021; Agyeman et al., 2023).

When stress occurs during the flowering stage, it can lead to flower abortion and low fertilisation rates, directly impacting pod production and reducing yield. Stress during the pod filling stage can reduce the production of assimilates available for translocation to the pods, thus limiting the supply of essential carbohydrates and nutrients needed for seed development. This reduction can ultimately result in empty pods (Mabhaudhi, 2024; pers. comm.) and hence no yield.

**Table 4.4** identifies the AZ that produced the largest yield difference between the two best planting dates for each crop. For the later (November and December) planting date, AZ number 2 681 and 3 326 produced the higher yield for bambara nut and cowpea, respectively. Hence, planting one month later in these two zones produced higher yields despite $RFL_{AVE}$ being similar. This suggests that the later planting avoided potential stress occurring during the two critical growth stages (i.e. flowering and pod filling stages). In contrast, planting one month earlier in AZ 352 and 655 resulted in a better yield for bambara nut and cowpea, respectively.

**Table 4.4: Highest yield differences between the two best planting months for both crops**

| Crop | AZ no. | Date | $RFL_{AVE}$ | $B_{AVE}$ | HI | $Y_{AVE}$ |
|---|---|---|---|---|---|---|
| Bambara nut | 2 681 | Oct | 541 | 8.39 | 9.81 | 1.01 |
| | | **Nov** | 533 | 11.36 | 13.76 | **1.76** |
| Cowpea | 3 326 | Nov | 308 | 5.91 | 17.82 | 1.14 |
| | | **Dec** | 300 | 6.98 | 22.16 | **1.59** |
| Bambara nut | 352 | **Oct** | 744 | 10.75 | 16.46 | **1.87** |
| | | Nov | 711 | 9.17 | 8.87 | 0.84 |
| Cowpea | 655 | **Nov** | 815 | 7.17 | 15.09 | **1.51** |
| | | Dec | 832 | 6.82 | 4.03 | 0.35 |

This study has shown that planting date can impact crop yield, which concurs with the findings by Kunz et al. (2020; 2024). Of the 334 global studies between 2009 and 2019 involving AquaCrop, only 4% considered sowing date, with the majority focusing on climate change, model calibration and irrigation management (Salman et al., 2021). Modelling the impact of planting date on crop response is especially important for photosensitive crops such as bambara nut and cowpea. Although AquaCrop cannot account for photoperiod effects on crop yield, it does consider the timing of water and temperature stress in relation to crop phenology.

### 4.1.2 Impact of frost occurrence

AquaCrop can calculate unviable crop cycle lengths (Kunz et al., 2020; 2024), especially when simulating cooler areas. Crop cycle lengths determined using thermal time often exceed 365 days in AZs located in mountainous areas due to the colder climate. When this occurs, the model is not run, and thus no data is simulated for that season. In some cases, the model is not run for 46 or more seasons, and thus no statistics can be calculated for that AZ. Such AZs as shown in white and labelled as “No data” in **Figure 4.1** and **Figure 4.2**, as previously mentioned in **Section 4.1.1.1**.

More importantly, AquaCrop cannot account for the impact of frost occurrence on crop cycle length. One of the key contributions of this study was to reduce the crop cycle length using minimum air temperature, which is done in the .PRM file that AquaCrop requires to conduct the model runs. As discussed in **Section 3.3.4**, if the daily air temperature drops to 5.4°C or lower, the assumption is made that frost will occur, which kills the crop, i.e. crop growth is immediately halted. Hence, the crop does not reach physiological maturity, resulting in a shorter crop cycle length and lower yield. Schulze and Schütte, (2016) also stated that when the minimum air temperature drops to 5°C or below, many crops can die (before they reach physiological maturity), even without frost occurring.

#### 4.1.2.1 Bambara nut

When frost is not considered, long crop cycle lengths (> 240 days on average) can occur, especially when the crop is planted in January. Hence, physiological maturity is reached in August, well into winter when the first frost date should have already occurred. As expected, the impact of frost occurrence on the average crop cycle length ($CYCLE_{AVE}$) was more pronounced for the late season planting (i.e. January). From **Figure 4.3**, the crop cycle shortened by more than 210 days across 19.0% of southern Africa, which further highlights the importance of choosing the correct planting date for crop production, especially for crops with long growing seasons. $CYCLE_{AVE}$ was reduced by more than 90 days for the majority (61.4%) of the AZs.

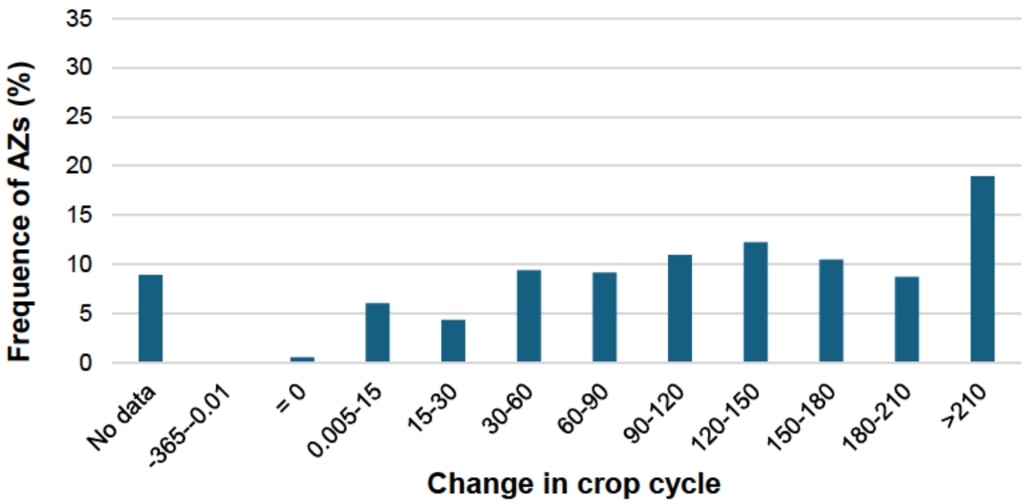


**Figure 4.3:** **Histogram of changes in $CYCLE_{AVE}$ for bambara nut planted in January due to frost occurrence**

Only 6.7% of the southern African region had a reduction in $CYCLE_{AVE}$ of less than 15 days. As shown in **Figure 4.4**, this mainly occurred in the warmer areas along the coastline of KwaZulu-Natal and the eastern parts of Limpopo and Mpumalanga. The areas shown in white in **Figure 4.4** represent 8.9% of southern Africa's total area (**Figure 4.3**), where no data exists because the model was not run for these cold AZs.

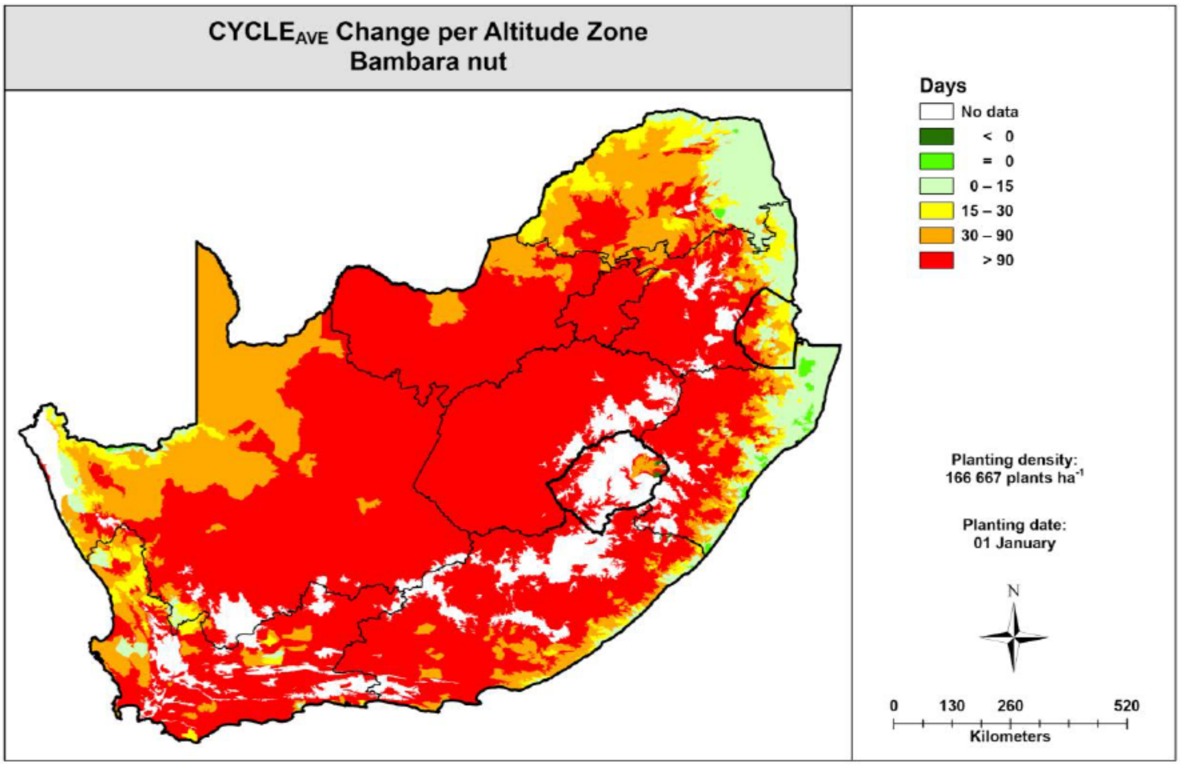


**Figure 4.4:** **Reduction in $CYCLE_{AVE}$ for bambara nut planted in January due to frost occurrence**

The reduction in $CYCLE_{AVE}$ due to frost occurrence is shown in **Figure 10.5** (cf. **APPENDIX D**) for bambara nut planted in October. In contrast, $CYCLE_{AVE}$ changed by less than 15 days across 39.9% of southern Africa. However, the reduction was 90 days or more for 19.5% of the study area, and even exceeded 210 days (i.e. seven months) in some areas (3.0%). **Figure 10.6** (cf. **APPENDIX D**) also shows that frost has a lower impact on $CYCLE_{AVE}$ when bambara nut is planted early in the season (October). The reduction in $CYCLE_{AVE}$ was less than 15 days along the Eastern Cape coastline and adjacent interior, including most of central and eastern KwaZulu-Natal, as well as eastern Mpumalanga and most of the Limpopo (and Eswatini). As expected, frost had the greatest impact on $CYCLE_{AVE}$ in the higher altitude areas of South Africa's plateau region. However, in a few zones (0.8% of the study area), the change in $CYCLE_{AVE}$ was negative (dark green areas in the maps), meaning it increased, which is discussed further in **Section 4.1.2.3**.

#### 4.1.2.2 Cowpea

If frost is not considered, long crop cycle lengths (> 160 days on average) can occur when the crop is planted in January, as is the case for bambara nut (cf. **Section 4.1.2.1**). Although the impact of frost occurrence on $CYCLE_{AVE}$ was one again more pronounced for the later planting, it was considerably less than that for bambara nut, due to cowpea's shorter growing season, i.e. the crop reaches physiological maturity before the winter months when frost is likely to occur. **Figure 4.5** clearly highlights the reduced impact of frost occurrence on cowpea compared to bambara nut (cf. **Figure 4.3** in **Section 4.1.2.1**) since the reduction was less than 15 days for more than half (57.1%) of southern Africa's total area. For cowpea planted in January, 13.8% of southern Africa had no change in $CYCLE_{AVE}$, compared to bambara nut which only had 0.6%.

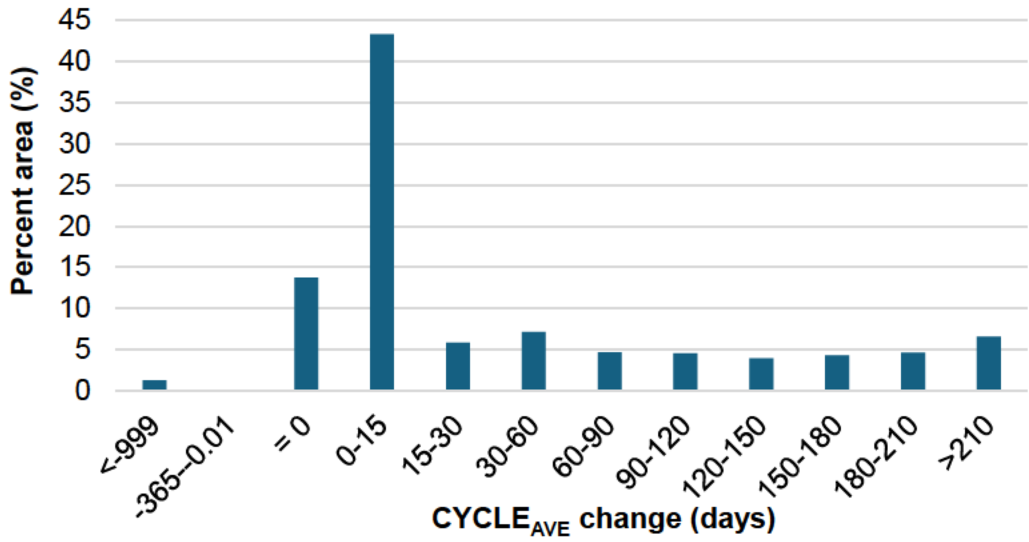


**Figure 4.5: Histogram of reduced $CYCLE_{AVE}$ for cowpea planted in January due to frost occurrence**

Although frost occurrence has a reduced impact on cowpea's $CYCLE_{AVE}$, 24% of the study area experienced changes of more than 90 days, including 6.7% of southern Africa where the reduction exceeded 210 days. As shown these changes occur in the higher altitude areas in southern Africa, as highlighted in red in **Figure 4.6**.

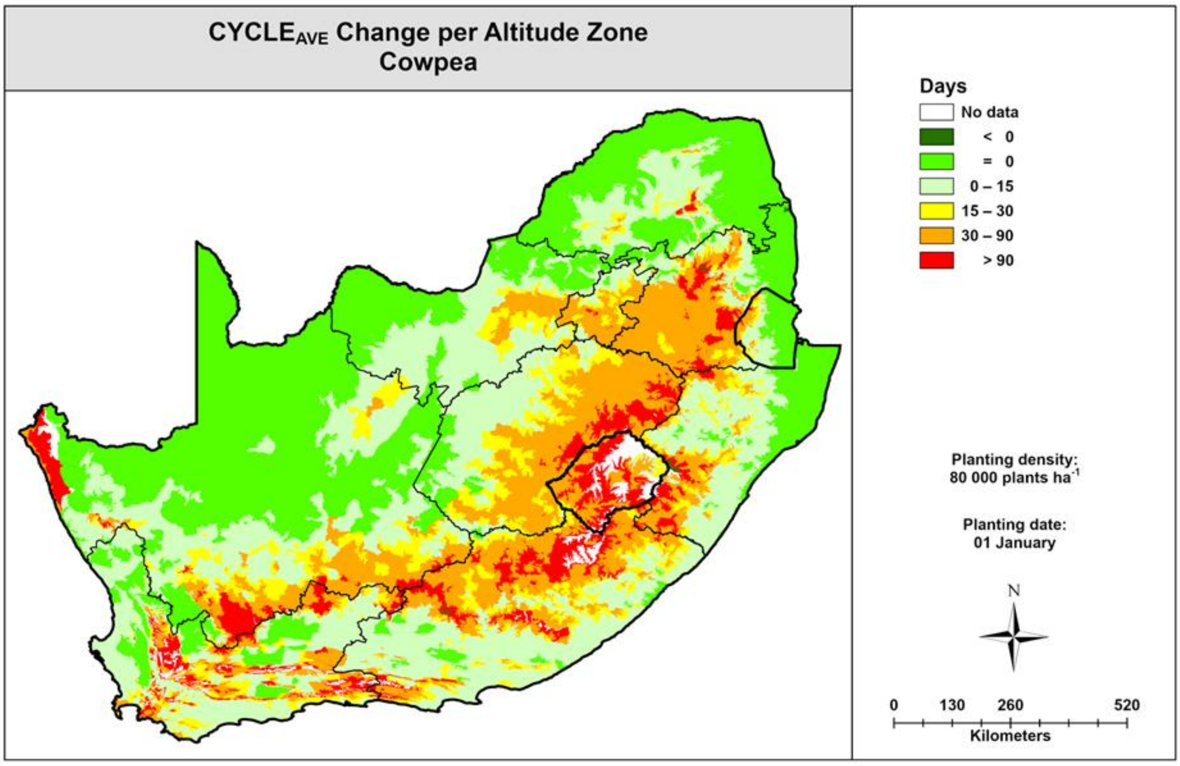


**Figure 4.6: Reduction in CYCLE$_{AVE}$ for cowpea planted in January due to frost occurrence**

The frost-induced reduction in CYCLE$_{AVE}$ for cowpea planted in November is shown in **Figure 10.7** (cf. **APPENDIX D**). As discussed, cowpea's growing season is shorter than bambara nut and so differences in CYCLE$_{AVE}$ between early (November) and late (January) planting are reduced. CYCLE$_{AVE}$ changed by less than 15 days across 56.2% of southern Africa, whereas the reduction exceeded 90 days in 6.9% of the study area, with only 1.9% exceeding 210 days. **Figure 10.8** (cf. **APPENDIX D**) further highlights the lower impact that frost had on CYCLE$_{AVE}$ when cowpea was planted early in November. Although **Figure 10.8** groups changes from 30-90 days (as shown in orange), the histogram (**Figure 10.7** in **APPENDIX D**) shows most of the reduction was 30-60 days (15.2%) vs 60-90 days (7.2%). This explains why the histograms were provided because of the extreme difficulty in developing a map displaying 12 individual colours.

#### 4.1.2.3 Increase in CYCLE$_{AVE}$

For each AZ, the crop cycle length (CYCLE) is calculated using thermal time (i.e. heat units in GDDs) just before AquaCrop is run for each of the 49 seasons. This is done in the project (.PRM) file which instructs AquaCrop to run for each season up to a specific date when the crop reaches physiological maturity. When CYCLE exceeds 365 days, the season is considered too cold to produce an economically viable yield. The model is not run, and thus no data exists for the season. This can occur for 46 or more seasons, meaning no simulated data or statistics exist for the AZ, and therefore $CYCLE_{AVE}$ is assumed to be zero days. However, with the inclusion of frost, CYCLE can decrease to below 365 days and the model is now run for that season. Statistics (i.e. $CYCLE_{AVE}$) are produced for the AZ if the model runs for four or more seasons. This explains why in some AZs, the change in $CYCLE_{AVE}$ was negative due to frost occurrence (i.e. change = 0 - $CYCLE_{AVE}$), as shown for bambara nut planted in October (cf. **Figure 10.6** in **APPENDIX D**). Further analysis showed that this occurred in 56 AZs, where RCF exceeded 93.8% since $Y_{AVE}$ was close to zero for the majority of the simulated seasons (**Table 4.5**). Therefore, these zones still represent areas that are too cold for crop cultivation. As the planting month "increases" (from October to January), the number of AZs that have data decreases. Upon further analysis, it was found that for certain seasons, the model did not simulate all months in the season, and typically crashed in the first month after planting, due mainly to division-by-zero errors. When this occurs, $CYCLE_{AVE}$ is typically a few days in length and $Y_{AVE}$ is zero dry t $ha^{-1}$. For such zones, the model should still not be run, even if CYCLE is reduced to less than 365 days due to frost occurrence.

**Table 4.5: AZs with $CYCLE_{AVE}$ data after accounting for frost**

| | Oct | Nov | Dec | Jan |
|---|---|---|---|---|
| **Bambara nut** | 56 AZs<br>RCF > 93.8% | 20 AZs<br>RCF > 93.8% | 9 AZs<br>RCF > 97.9% | 4 AZs<br>RCF = 100% |
| **Cowpea** | 12 AZs<br>RCF = 100% | 12 AZs<br>RCF > 97.9% | 4 AZs<br>RCF = 100% | 1 AZ<br>RCF = 100% |

#### 4.1.2.4 Reduction in yield

Accounting for frost occurrence reduced $Y_{AVE}$ by more than 1 dry t $ha^{-1}$ for both crops, due mainly to the shortened season length. This occurred mainly in the higher altitude areas of southern Africa, as shown in red in **Figure 10.9** (bambara nut ) and **Figure 10.10** (cowpea) in **APPENDIX D**. **Table 4.6** highlights the AZs with the largest decline in $Y_{AVE}$ for both crops and for the four planting dates. If frost occurs (i.e. minimum temperature drops below 5.4°C), the crop does not reach physiological maturity, which impacts the length of flowering and the duration of yield formation (pod filling stage). Thus, $Y_{AVE}$ decreased by 2.291-4.598 and 0.680-1.649 dry t $ha^{-1}$ for bambara nut and cowpea, respectively. For bambara nut, the reduction in $CYCLE_{AVE}$ ranged between 135-228 days, leading to a reduction in $Y_{AVE}$ of up to 4.60 dry t $ha^{-1}$. The smaller reduction in $CYCLE_{AVE}$ (82-179 days) for cowpea resulted in a lower impact on $Y_{AVE}$ (< 1.65 dry t $ha^{-1}$).

When comparing the reductions in $Y_{AVE}$ across the four planting dates for bambara nut, it was found that although the decrease in $CYCLE_{AVE}$ was greatest for October, this did not correlate to the largest $Y_{AVE}$ reduction in as seen for cowpea. One potential reason for this is the low

$Y_{AVE}$ for the October planting when frost was not considered (i.e. without frost), which is substantially less than the other three planting dates, yet $CYCLE_{AVE}$ was relatively similar. This is likely due to higher stress experienced during flowering or yield formation stages compared to the other planting dates (cf. **Section 4.1.1.3**), leading to lower $Y_{AVE}$. Accounting for the first frost date allows AquaCrop to simulate more realistic crop cycle lengths, which in turn results in more accurate yield estimates. This highlights the importance of taking frost into account when crop modelling, thus avoiding yields being over-estimated due to unrealistically long season lengths. The addition of frost as a limiting factor would thus require major changes to the model code as well as adding an additional parameter value.

**Table 4.6: AZs with the largest decrease in $Y_{AVE}$ due to frost occurrence**

| Crop | Planting date | AZ no. | Without frost | | With frost | | Change | |
|---|---|---|---|---|---|---|---|---|
| | | | $CYCLE_{AVE}$ | $Y_{AVE}$ | $CYCLE_{AVE}$ | $Y_{AVE}$ | $CYCLE_{AVE}$ | $Y_{AVE}$ |
| **Bambara nut** | Oct | 3 987 | 298 | 2.591 | 70 | 0.300 | 228 | 2.291 |
| | Nov | 3 332 | 303 | 4.020 | 168 | 1.000 | 135 | 3.020 |
| | Dec | 3 332 | 317 | 4.500 | 146 | 0.630 | 171 | 3.870 |
| | Jan | 2 685 | 322 | 4.908 | 124 | 0.310 | 198 | 4.598 |
| **Cowpea** | Oct | 3 328 | 267 | 2.039 | 88 | 0.390 | 179 | 1.649 |
| | Nov | 4 654 | 203 | 2.110 | 83 | 0.540 | 120 | 1.570 |
| | Dec | 3 311 | 244 | 1.870 | 148 | 1.190 | 96 | 0.680 |
| | Jan | 3 296 | 228 | 1.643 | 146 | 0.220 | 82 | 1.423 |

### 4.1.3 Testing for collinearity

As discussed in **Section 3.6.1**, it was important to keep the number of variables used to eliminate unsuitable AZs to a minimum, thus reducing the complexity involved in developing each land suitability map. In this study, four variables were used compared to five by Kunz et al. (2024; cf. **Section 2.1.2.3**). Any output variable that was deemed to be strongly correlated (i.e. collinear) with any other variable was not considered, even if it was only related for one planting date or crop. The average (across 49 simulated seasons) of 13 variables were tested, together with an additional statistic (coefficient of variation) for four variables, as shown in **Table 8.2** (**APPENDIX B**). Hence, 17 statistics in total were tested, which produced 136 relationships for each planting date (or 544 graphs for each crop). This would not have been possible if the collinearity test had not been fully automated using Python, as discussed in **Section 3.7**. This outcome ensured that the variables selected to eliminate or classify AZs were not related to one another.

From **Section 3.5.2**, the absolute value of Pearson's correlation coefficient (|r|) was used to test for a linear relationship between two variables. A threshold value of 0.8 was applied to determine whether two variables were correlated. Thus, one variable was excluded from further analysis if |r| > 0.8. However, an important issue to highlight is applying discrete thresholds with continuous datasets. As shown in **Table 4.7**, some variable pairs have correlation values just below 0.8, and thus were not excluded, but maybe should have been. Fortunately, none of the variables selected in the methodology to eliminate of classify AZs exhibited correlations close to the 0.8 threshold, and therefore were used with confidence.

**Table 4.7: Examples of variable pairs with a Pearson correlation coefficient close to the 0.8 threshold**

| Relationship | Crop | Planting date | \|r\| |
|---|---|---|---|
| $B_{AVE}$ vs $BREL_{AVE}$ | cowpea | Jan | 0.792 |
| $CYCLE_{AVE}$ vs $CWP_{AVE}$ | bambara nut | Nov | 0.792 |
| $B_{CV}$ vs $Y_{CV}$ | cowpea | Nov | 0.793 |
| $BREL_{AVE}$ vs $E/Ex_{AVE}$ | cowpea | Oct | 0.794 |

## 4.2 Land Suitability Assessment

The following section discusses the outcome of the land suitability mapping approach. The section highlights the importance of considering (i) planting date, (ii) frost occurrence, and (iii) existing land use when assessing land suitability for crop production.

### 4.2.1 Elimination of unsuitable areas

Kunz et al. (2024) used CYCLE, MAP, RCF, $CWP_{AVE}$ and $CWP_{CV}$ to eliminate AZs deemed unsuitable for OFSP and taro production (cf. **Section 2.1.2.3**). As discussed in **Section 3.6.1**, similar variables were used in this study, namely CYCLE, $RFL_{AVE}$, RCF and $CWP_{CV}$ (cf. **APPENDIX B**). Seasonal rainfall (RFL) accumulated from planting to the end-season date, which was then averaged across the 49 seasons ($RFL_{AVE}$), was selected to eliminate AZs considered too dry for rainfed production. $RFL_{AVE}$ was chosen instead of MAP since it represents the actual rainfall received by the crop, whereas MAP is the average rainfall accumulated over each hydrological year (October to September). To re-cap, the end-season date is either the crop's physiological maturity date (calculated using thermal time), or an earlier date if the minimum air temperature reaches 4.5°C or less.

$CWP_{AVE}$ was therefore excluded because it is strongly correlated to $RFL_{AVE}$ for bambara nut (**Figure 11.1**) and cowpea (**Figure 11.2**). Since $CWP_{AVE}$ is calculated from $Y_{AVE}$, it explains why both variables are strongly correlated. Similarly, $RFL_{AVE}$ and $Y_{AVE}$ are also highly collinear for both crops (cf. **Figure 11.3** and **Figure 11.4**). These four graphs in **APPENDIX E** show that seasonal (not annual) rainfall is one of the primary drivers of yield formation, further highlighting the importance of including this variable in land suitability assessments. **Table 4.8** highlights other AquaCrop output variables that are strongly correlated to $RFL_{AVE}$, thus resulting in their exclusion. The following section discusses the results from the elimination phase and focuses on the impact of planting date.

**Table 4.8: High correlations between RFL and other AquaCrop output variables**

| Variable | Crop | Planting date | \|r\| |
|---|---|---|---|
| B | cowpea | Oct | 0.949 |
| Tr | cowpea | Oct | 0.943 |
| E/Ex | bambara nut | Oct | 0.933 |
| BREL | bambara nut | Dec | 0.888 |
| E | bambara nut | Oct | 0.882 |
| ExpStr | bambara nut | Dec | 0.855 |

#### 4.2.1.1 Bambara nut

**Table 4.9** provides a summary of the threshold values and the number of AZs eliminated by applying each of the four criteria (i.e. variable and its threshold value). For the November planting, 2 757 zones were eliminated in total (i.e. unsuitable for bambara nut production), which corresponds to 47.22% of the 5 838 AZs or 50.46% of southern Africa's total area. The eliminated AZs are coloured white and red in **Figure 4.7**. Although these AZs are considered currently unsuitable for crop production, the changing climate could result in some of the zones becoming suitable in the future. The remaining 3 081 zones are considered suitable for crop production (shown in green in **Figure 4.7**), which are discussed further in **Section 4.2.2**. A total of 1 750 zones were eliminated using $RFL_{AVE}$, RCF and/or $CWP_{CV}$, as some zones are eliminated by more than one criterion (white areas in **Figure 4.7**). It is important to note that the order in which the variables are applied does not affect the outcome of the elimination process. In **Figure 4.7**, the red areas have no simulated data because the model was not run for these cold AZs (cf. **Section 3.6.1**). These 1 007 zones were identified using CYCLE, which when added to the eliminated zones (1 750), gives the total number of eliminated zones (2 757).

For the November planting, $RFL_{AVE}$ removed the most AZs (21%), followed by CYCLE (17%), RCF (16%) and $CWP_{CV}$ (15%), with the remaining 53% classified as suitable. For January, RCF removed the most AZs (32%), followed by $RFL_{AVE}$ and $CWP_{CV}$ (25%), with CYCLE having the least impact (18%). When compared to the November planting, RCF and $CWP_{CV}$ eliminated an additional 882 AZs for the January planting (**Table 4.9**). The higher crop failure rate and inter-seasonal variability in CWP could be a result of higher water stress during critical phenological growth stages (cf. **Section 4.1.1.3**) because the crop was planted late (i.e. January). Overall, November may be considered the most suitable planting month since the least number of AZs were eliminated. However, of the four planting months, October produced the highest bambara nut yields across most of the eastern seaboard of southern Africa, as highlighted in **Section 4.1.1.1**.

**Table 4.9: Criteria used to eliminate AZs deemed currently unsuitable for bambara nut production**

| Reasoning | Variable & threshold | Number of AZs | | | | Map colour |
|---|---|---|---|---|---|---|
| | | Oct | Nov | Dec | Jan | |
| Too dry | $RFL_{AVE}$ < 200 mm | 1 457 | 1 228 | 1 316 | 1 491 | |
| Too risky | RCF > 33% | 1 787 | 917 | 949 | 1 891 | |
| Too variable | $CWP_{CV}$ > 150% | 1 660 | 900 | 643 | 1 467 | |
| | **Currently unsuitable** | **2 334** | **1 750** | **1 846** | **2 600** | White |
| Too cold | CYCLE > 365 days | 961 | 1 007 | 1 007 | 1 039 | Red |
| | **Total eliminated** | **3 295** | **2 757** | **2 853** | **3 639** | White & red |
| | **Currently suitable** | 2 543 | 3 081 | 2 985 | 2 199 | Green |
| | | **% of southern Africa's total area** | | | | |
| | | **Oct** | **Nov** | **Dec** | **Jan** | |
| | Area eliminated | 62.24 | 50.46 | 51.82 | 66.37 | White & red |

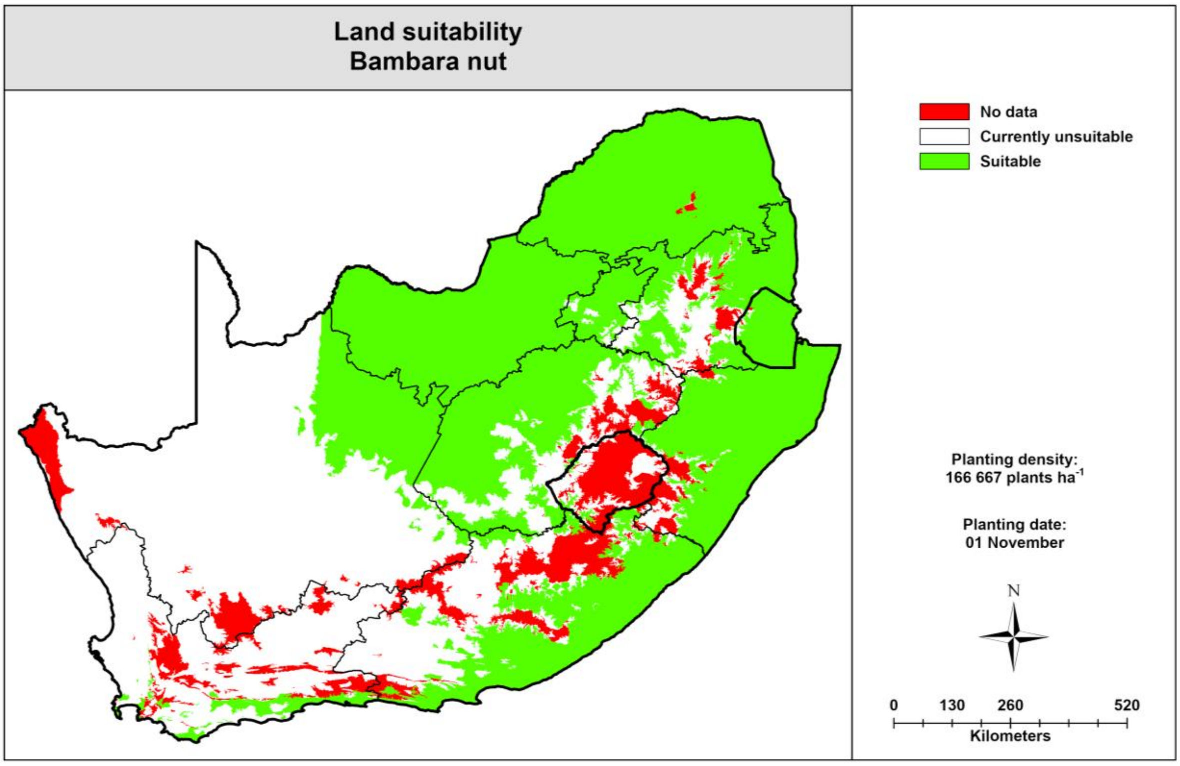


**Figure 4.7: Initial land suitability map for rainfed production of bambara nut in southern Africa, based on AquaCrop simulations**

#### 4.2.1.2 Cowpea

From **Table 4.10** , the least number of AZs were eliminated for the December planting (3 267), followed by November (3 605), January (3 768) and October (4 374). The total number of eliminated zones for the December planting corresponds to 55.96% of the 5 838 AZs or 67.07% of southern Africa's total area. As mentioned previously, these eliminated zones are coloured white and red in **Figure 4.8** and are considered currently (not permanently) unsuitable for crop production. The remaining 2 571 zones are considered suitable for crop production (shown in green in **Figure 4.8**), which were then classified as marginally, moderately and highly suitable (cf. **Section 4.2.2**). A total of 3 053 zones were eliminated using $RFL_{AVE}$, RCF and/or $CWP_{CV}$ (white areas in **Figure 4.8**). Only 214 zones have not modelled data (red areas in **Figure 4.8**), which when added to the eliminated zones (3 053), gives the total number of eliminated zones (3 267).

For the December planting, RCF removed the most AZs (42%), followed by $RFL_{AVE}$ (38%), $CWP_{CV}$ (28%) and CYCLE (4%), as some zones are eliminated by more than one criterion. The remaining 2 571 zones (44%) were classified as suitable. For October, RCF also eliminated the most zones (66%), followed by $RFL_{AVE}$ (60%), $CWP_{CV}$ (47%), with CYCLE again having the least impact (3%). When compared to the December planting, the October planting removed an additional 1 107 AZs, due mainly to the $RFL_{AVE}$, RCF and $CWP_{CV}$ variables (**Table 4.10**). Similar to bambara nut, the higher crop failure and inter-seasonal variability in CWP, together with the low $RFL_{AVE}$ in October, could be a result of high water stress during critical phenological growth stages (cf. **Section 4.1.1.3**). In contrast to bambara nut where CYCLE removed on average 1 004 AZs, CYCLE removed only 214 zones for cowpea, which is discussed in more detail in the next section. Overall, December may be considered the most suitable planting month since the least number of AZs were eliminated. However, of the four planting months, November and December produced the highest yields across most of the eastern seaboard of southern Africa, as highlighted in **Section 4.1.1.2**.

**Table 4.10: Criteria used to eliminate AZs deemed currently unsuitable for cowpea production**

| Reasoning | Variable & threshold | Number of AZs | | | | Map colour |
|---|---|---|---|---|---|---|
| | | Oct | Nov | Dec | Jan | |
| Too dry | $RFL_{AVE}$ < 200 mm | 3 527 | 2 606 | 2 213 | 2 234 | |
| Too risky | RCF > 33% | 3 827 | 3 033 | 2 437 | 2 573 | |
| Too variable | $CWP_{CV}$ > 150% | 2 737 | 2 094 | 1 643 | 1 625 | |
| | **Currently unsuitable** | **4 170** | **3 392** | **3 053** | **3 545** | White |
| Too cold | CYCLE > 365 days | 204 | 213 | 214 | 223 | Red |
| | **Total eliminated** | **4 374** | **3 605** | **3 267** | **3 768** | White & red |
| | **Currently suitable** | 1 464 | 2 233 | 2 571 | 2 070 | Green |
| | | **% of southern Africa's total area** | | | | |
| | | **Oct** | **Nov** | **Dec** | **Jan** | |
| | Area eliminated | 85.68 | 74.72 | 67.07 | 71.87 | White & red |

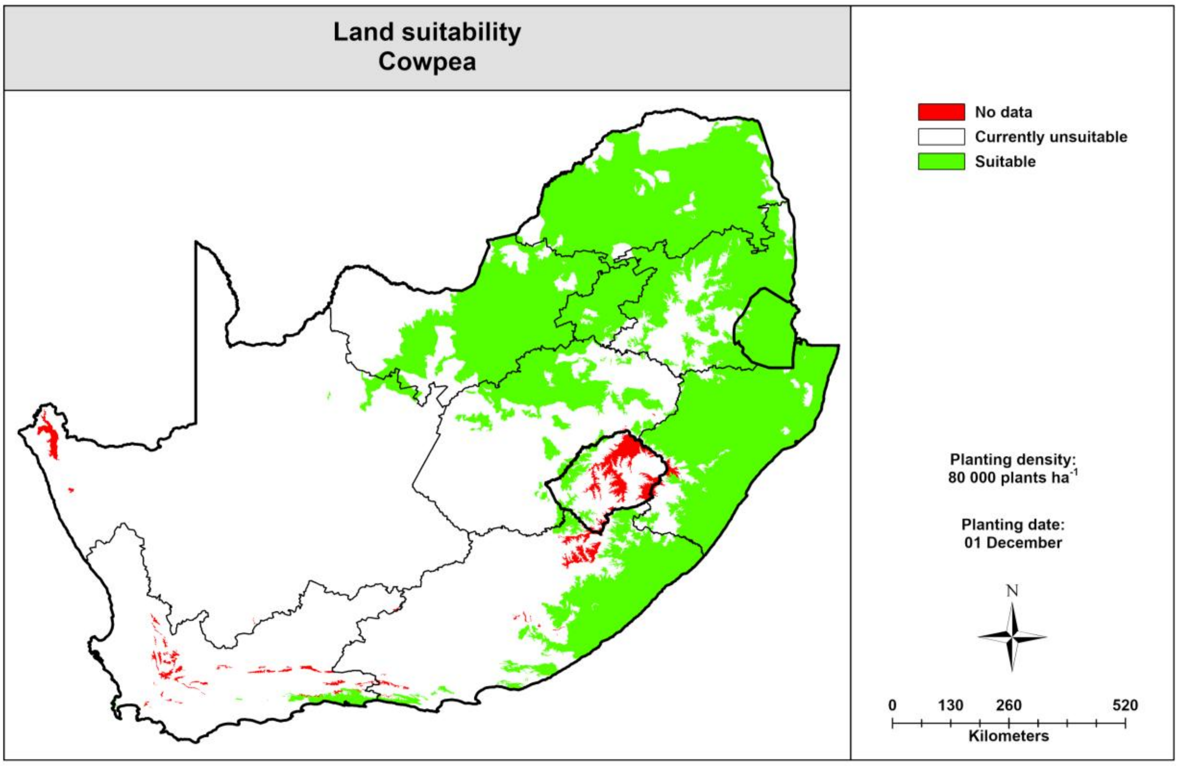


**Figure 4.8: Initial land suitability map for rainfed production of cowpea in southern Africa, based on AquaCrop simulations**

### 4.2.1.3 Crop cycle length

Results show that depending on the planting date, about 50-66% of southern Africa's total area is deemed unsuitable for bambara nut cultivation (cf. **Section 4.2.1.1**), compared to 67-86% for cowpea (cf. **Section 4.2.1.2**). This is due to $RFL_{AVE}$, RCF and $CWP_{CV}$ eliminating more zones for cowpea production compared to bambara nut. On average, about 1 004 AZs are considered too cold and economically unviable for bambara nut production (**Table 4.9**), compared to only 214 for cowpea (**Table 4.10**). This is related to bambara nut's longer time to reach physiological maturity of 2 227 GDDs, which results in an average crop cycle length of 194 days for the November planting (when frost is not considered). Cowpea on the other hand, has a much shorter time to physiological maturity of 1 259 GDDs, and thus the average crop cycle length is 130 days for the December planting. Although cowpea has a shorter CYCLE than bambara nut, meaning less AZs are impacted by frost occurrence, more AZs are eliminated. This suggests that cowpea experiences more water stress during the two critical phenological growth stages compared to bambara nut (cf. **Section 4.1.1.3**).

## 4.2.2 Classification of suitable growing areas

Kunz et al. (2024) used $CWP_{AVE}$ to classify the remaining suitable areas into three classes (cf. **Section 2.1.2.3**). For this study, $BREL_{AVE}$ was chosen as it better represents the suitability of each AZ. To re-cap (cf. **Section 3.6.2**), $BREL_{AVE}$ was normalised ($BREL_{NORM}$) and used to classify the suitable AZs into S3 (25% to 50%), S2 (50% to 75%) and S1 (> 75%). AZs with a $BREL_{NORM}$ < 25% were considered currently unsuitable (N1), as these zones produced less than a quarter of their potential biomass. Although these thresholds are subjective, the advantages of this approach are (i) the normalisation ensures that S1 areas always exist, and (ii) the thresholds are crop independent. This study did not use $Y_{AVE}$ as the classification variable because threshold values for separating S2 from S1 and S3 are (i) difficult to derive, (ii) crop dependent, (iii) different for smallholder and commercial farmers, (iv) not readily available in the literature, and thus (v) are very subjective.

#### 4.2.2.1 Bambara nut

For the October planting, 3 295 AZs were eliminated resulting in 2 543 AZs remaining as suitable for bambara nut production, with their spatial distribution shown in **Figure 4.9**. The corresponding map for the November planting, which has the most suitable areas, is provided in **APPENDIX F** (cf. **Figure 12.1**). The maps indicate that most of the central and eastern parts of southern Africa are considered marginally to moderately suitable for bambara nut production. The coastal regions of the Eastern Cape and KwaZulu-Natal provinces, as well as parts of Mpumalanga and neighbouring Eswatini are deemed highly suitable (S1), which corresponds to the higher rainfall regions (cf. **Figure 3.1** in **Section 3.1**). This is expected since the collinearity tests showed that $RFL_{AVE}$ was the primary driver of yield formation (cf. **APPENDIX E**).

Only 6.4% of southern Africa's total area is considered highly suitable (S1) for bambara nut planted in October, whereas 18.9% is moderately suitable (S2) and 12.5% is marginally suitable (S3) (**Figure 4.10**). Overall, southern Africa is marginally to moderately suitable for bambara nut cultivation. However, this outcome is sensitive to the threshold values selected for $BREL_{NORM}$. For bambara nut, $BREL_{NORM}$ was below 25% for 16 AZs (0.3% of southern Africa's area), and thus were reclassified as N1. This suggests that BREL could be used as an elimination variable.

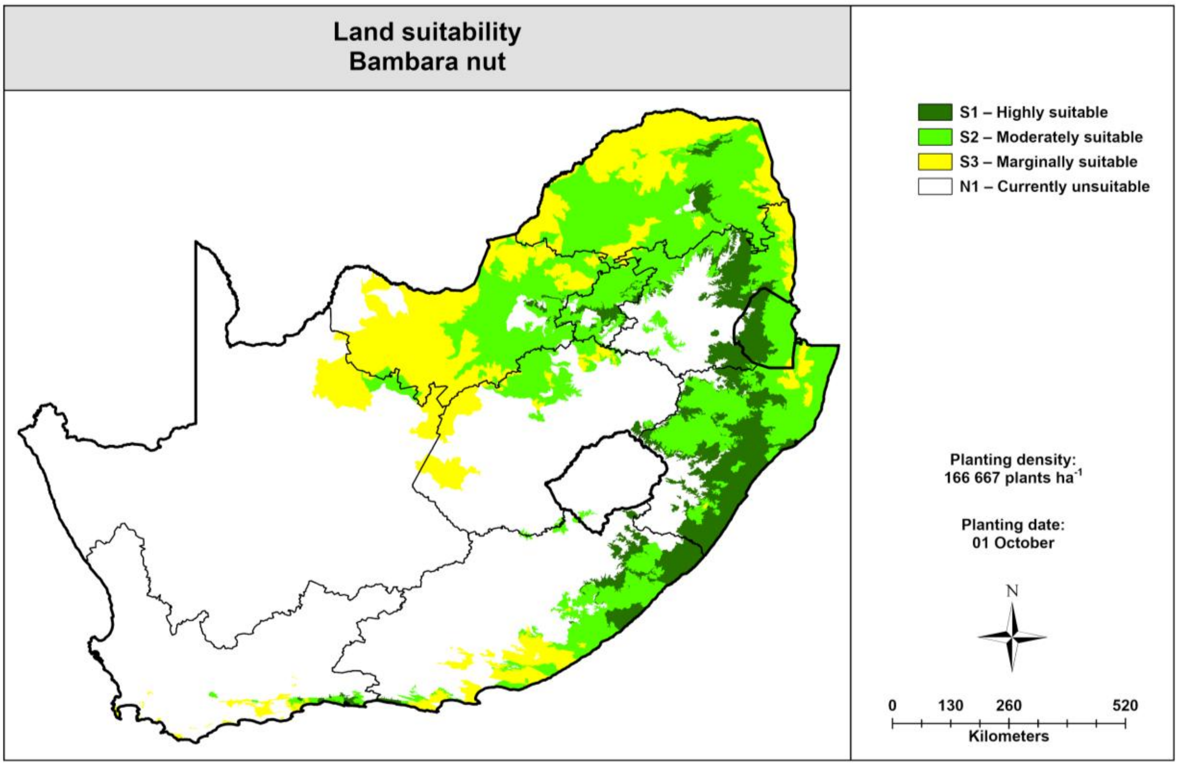


**Figure 4.9: Suitability classification map for rainfed production of bambara nut planted in October for southern Africa, based on $BREL_{NORM}$**

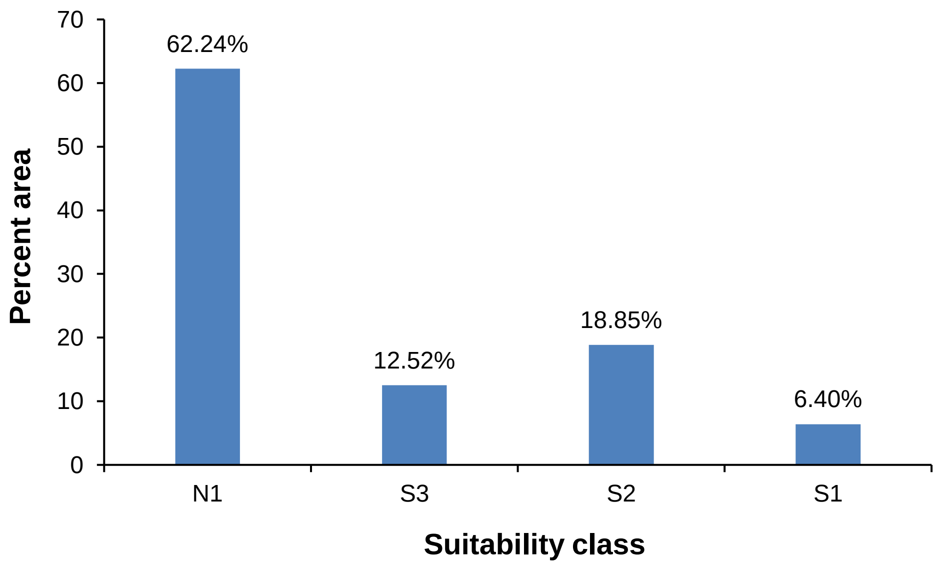


**Figure 4.10: Percent area suitability classes for bambara nut planted in October**

$Y_{AVE}$ values for each AZ were spatially averaged across each suitability class, as shown in **Table 4.11** for each planting month. As expected, $Y_{AVE}$ is highest for S1 and lowest for S3, except for the January planting. The likely reason for this is the large range in $Y_{AVE}$ associated with a particular $BREL_{NORM}$ value, which is discussed further in **Section 4.2.2.3**. In addition, the spatially averaged yields gradually decrease from the October to January planting, which is due to the reduction in $CYCLE_{AVE}$ caused by frost occurrence, since October experiences an average reduction of 60 days compared to January's 141 days (cf. **Section 4.1.2.1**). This reduction in $CYCLE_{AVE}$ shortens the length of the flowering and pod filling stages leading to reduction in yield formation. While the November planting produced the largest suitable area (**Table 4.9**), the October planting has the highest spatial $Y_{AVE}$ (**Table 4.11**). This suggests that October is the better planting date, as farmers are likely to favour higher yields rather than focusing on the extent of suitable areas.

**Table 4.11: Spatial $Y_{AVE}$ per suitability class for bambara nut**

| | **Oct** | **Nov** | **Dec** | **Jan** |
|---|---|---|---|---|
| **S1** | 1.55 | 1.24 | 0.76 | 0.42 |
| **S2** | 0.92 | 0.76 | 0.56 | 0.44 |
| **S3** | 0.39 | 0.36 | 0.35 | 0.30 |
| **ave** | 1.04 | 0.81 | 0.57 | 0.41 |

#### 4.2.2.2 Cowpea

Since 3 605 AZs were eliminated for the November planting, the remaining 2 233 zones are deemed suitable for cowpea production, as shown in **Figure 4.11**. The corresponding map for the December planting, which has the most suitable areas, is provided in **APPENDIX F** (cf. **Figure 12.2**). Similar to bambara nut (cf. **Figure 4.9**), most of the central and eastern parts of the study area are considered marginally to moderately suitable for cowpea production. The coastal region of Eastern Cape as well as parts of KwaZulu-Natal's interior, eastern Mpumalanga and western Eswatini are considered highly suitable (S1), which also corresponds with the higher rainfall region along the eastern seaboard of the country (cf. **Figure 3.1** in **Section 3.1**). This is expected since the collinearity tests showed that $RFL_{AVE}$ was the primary driver of yield formation (cf. **APPENDIX E**).

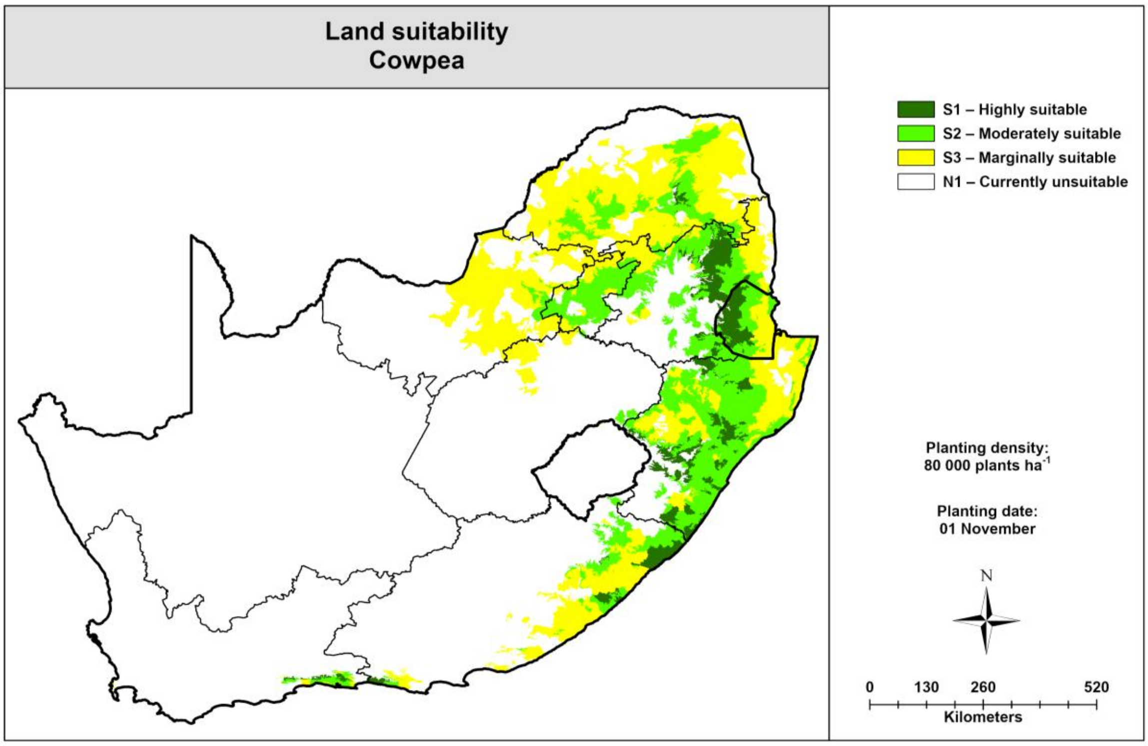


**Figure 4.11: Suitability classification map for rainfed production of cowpea planted in November for southern Africa, based on $BREL_{NORM}$**

For cowpea planted in November, only 2.4% of southern Africa's total area is considered highly suitable (S1), whereas 9.2% is moderately suitable (S2) and 13.7% is marginally suitable (S3) (**Figure 4.12**), resulting in southern Africa overall being considered marginally to moderately suitable for cowpea cultivation. Similar to bambara nut, 21 AZs (or 0.4% of southern Africa) exhibited a $BREL_{NORM}$ below 25%, and thus were classified as N1.

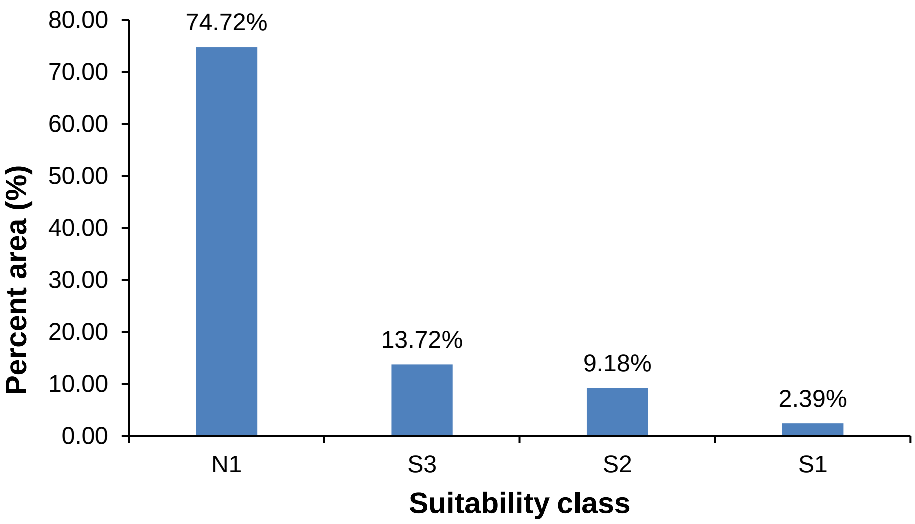


**Figure 4.12: Percent area suitability classes for cowpea planted in November**

Similar to bambara nut (cf. **Section 4.2.2.1**), $Y_{AVE}$ values for each AZ were spatially averaged across each suitability class, as shown in **Table 4.12**. As expected, $Y_{AVE}$ is highest for S1 and lowest for S3. In general, the spatially averaged yields gradually decreased from the October to January planting. This is due mainly to the reduction in $CYCLE_{AVE}$ caused by frost occurrence, with October experiencing an average reduction of 59 days compared to January's 79 days (cf. **Section 4.1.2.2**). However, $Y_{AVE}$ increased from the October to November planting for both S2 and S3 areas, and due to their spatial extent being much greater than the S1 area, the average yield across all three suitability classes also showed November's yield was highest (**Table 4.12**). This can be attributed to the higher crop failure in October (more zero yields across the 49 seasons), which resulted in lower averaged yields (cf. **Section 4.2.1.2**). While December planting produced the most suitable areas (**Table 4.10**), the November planting produced the highest spatial $Y_{AVE}$ (**Table 4.12**).

**Table 4.12: Spatial YAVE per suitability class for cowpea**

| | Oct | Nov | Dec | Jan |
|---|---|---|---|---|
| **S1** | 1.69 | 1.68 | 1.55 | 1.30 |
| **S2** | 0.95 | 1.02 | 0.94 | 0.77 |
| **S3** | 0.47 | 0.51 | 0.51 | 0.41 |
| **ave** | 0.88 | 0.92 | 0.88 | 0.71 |

#### 4.2.2.3 Limitations of BREL

**Figure 4.13** illustrates a curvilinear relationship exists between $BREL_{AVE}$ and $Y_{AVE}$, indicating the Pearson correlation coefficient is not appropriate for collinearity testing. For each threshold value of $BREL_{AVE}$, there is a corresponding upper limit of $Y_{AVE}$ as shown in **Table 4.13**. This means $Y_{AVE}$ should exceed 2 dry t ha$^{-1}$ for S1 areas. Similarly, zones that were classified as N1 (not S3) because BREL was low (< 25%) were associated with a $Y_{AVE}$ value of 0.0-0.4 dry t ha$^{-1}$. Although $Y_{AVE}$ increases with increasing $BREL_{AVE}$ as expected, the range in $Y_{AVE}$ widens considerably, and thus $BREL_{AVE}$ might not be the best variable for classifying the remaining suitable areas.

**Table 4.13: Upper limit of $Y_{AVE}$ for each threshold of $BREL_{AVE}$ for cowpea planted in November**

| $BREL_{AVE}$ | $Y_{AVE}$ |
|---|---|
| 25% | 0.4 (dry t ha$^{-1}$) |
| 50% | 1.2 (dry t ha$^{-1}$) |
| 75% | 2.0 (dry t ha$^{-1}$) |

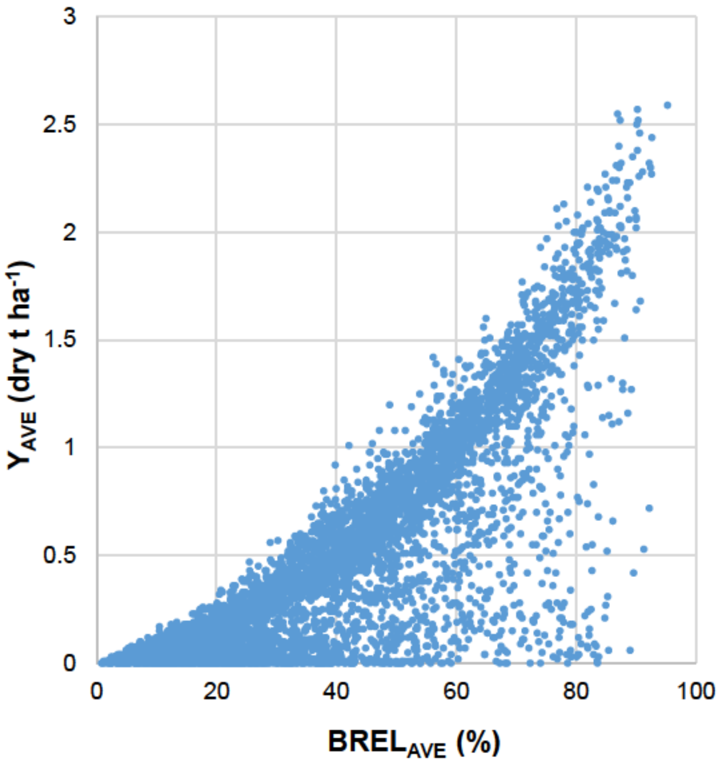


**Figure 4.13: Relationship between relative biomass potential ($BREL_{AVE}$) and yield ($Y_{AVE}$) simulated by AquaCrop for cowpea planted in November**

The figure above identified a limitation of using $BREL_{NORM}$ to classify suitable areas into three classes since AZs can have a high $BREL_{NORM}$, yet a low $Y_{AVE}$. This means some AZs were classified as highly suitable, yet they should have been classified as moderate to marginal due to low $Y_{AVE}$. This is further illustrated in **Table 4.14**, which compares AZs with a similar $BREL_{NORM}$ but different $Y_{AVE}$ values. For bambara nut, AZ number 4 295 and 4 465 have similar $BREL_{NORM}$ values of ~97%, yet the yield values differ by 1.98 dry t ha$^{-1}$. Similarly for cowpea, a $BREL_{NORM}$ of ~83% was simulated for AZ number 679 and 802 produce, yet $Y_{AVE}$ differs by 1.13 dry t ha$^{-1}$. This suggests that highly suitable areas shown in each land suitability map are likely to be over-estimated.

**Table 4.14: AZs considered highly suitable with different $Y_{AVE}$ values**

| Crop | Planting date | AZ no. | $BREL_{NORM}$ | $Y_{AVE}$ |
|---|---|---|---|---|
| **Bambara nut** | Oct | 4 295 | 96.79 | 0.68 |
| | | 4 465 | 97.17 | 2.60 |
| **Cowpea** | Nov | 679 | 82.91 | 0.60 |
| | | 802 | 82.74 | 1.73 |

When comparing the data from model runs when frost occurrence was not considered, further analysis showed that for both AZs, the low $Y_{AVE}$ resulted from:

(i) $CYCLE_{AVE}$ reducing by 170 (bambara nut) and 112 (cowpea) days, leading to less time for flowering or pod-filling;
(ii) $Y_{ZERO}$ increasing from 0 to 14 (bambara nut) and 11 (cowpea), resulting in a lower $Y_{AVE}$; and
(iii) $HI_{AVE}$ decreasing by 9.21% (bambara nut) and 7.82% (cowpea), which also caused a lower $Y_{AVE}$ (**Table 4.15**).

**Table 4.15: AZ for bambara nut and cowpea with low $Y_{AVE}$**

| Crop | AZ | Without frost | | | | With frost | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | $CYCLE_{AVE}$ | $Y_{ZERO}$ | $HI_{AVE}$ | $Y_{AVE}$ | $CYCLE_{AVE}$ | $Y_{ZERO}$ | $HI_{AVE}$ | $Y_{AVE}$ |
| **Bambara nut** | **4 295** | 286 | 0 | 15.72 | 2.22 | 116 | 14 | 6.51 | 0.62 |
| **Cowpea** | **679** | 247 | 0 | 15.34 | 1.31 | 135 | 11 | 7.52 | 0.60 |

The above analysis identified issues when using $BREL_{NORM}$ for classifying suitable areas since it revealed that some S1 areas can have unexpectedly low yields. It is important to note that this analysis was undertaken after the land suitability maps were finalised. It was not feasible to revise the methodology, since there was insufficient time to alter the Python code to regenerate (and re-analyse) all eight land suitability maps. Instead, effort focused on identifying an alternative (and better) variable, as discussed next.

#### 4.2.2.4 Harvest index

$HI_{AVE}$ was tested for classifying suitable areas for bambara nut (October planting) and cowpea (November planting). HI provides a reliable indication of how effectively the crop converts biomass into utilisable yield. Normalised HI ($HI_{NORM}$) was calculated by dividing each AZ's $HI_{AVE}$ by the reference harvest index ($HI_O$) for each crop, then expressing the ratio as a percentage. Hence, for unstressed growing conditions, HI at harvest should be similar to $HI_O$, and thus the ratio approaches unity (i.e. 100%). Similarly, when HI deviates from $HI_O$ due to water (and temperature) stress, the ratio approaches 0 (i.e. 0%). The thresholds used to classify S1, S2 and S3 areas were the same as $BREL_{NORM}$ (cf. **Section 4.2.2.3**) at 75%, 50% and 25% respectively. AZs with a $HI_{NORM}$ less than 25% were also reclassified as currently unsuitable (N1).

The results for these two planting dates revealed that, while the spatial averaged $Y_{AVE}$ for each suitability class remained relatively unchanged, the use of $HI_{NORM}$ led to an increase in minimum values of $Y_{AVE}$ across all three suitability classes, except for cowpea's S2 class (**Table 4.16**). The ranges in $Y_{AVE}$ using $HI_{NORM}$ are considered to better represent each suitability class, with a reduced likelihood of over-estimating the number of zones in each class, particularly for the S1 areas. $HI_{NORM}$ should therefore be considered in future work for classifying suitable areas instead of $BREL_{NORM}$.

**Table 4.16: Minimum, maximum and average yields per suitability class for bambara nut (October) and cowpea (November) based on two normalised classification variables (BREL and HI)**

| Crop | Planting date | Class | Classification using $BREL_{NORM}$ | | | Classification using $HI_{NORM}$ | | |
|---|---|---|---|---|---|---|---|---|
| | | | **Min** | **Max** | **Ave** | **Min** | **Max** | **Ave** |
| **Bambara nut** | **Oct** | **S1** | 0.56 | 2.79 | 1.55 | 0.90 | 2.79 | 1.59 |
| | | **S2** | 0.16 | 1.99 | 0.92 | 0.46 | 1.86 | 0.92 |
| | | **S3** | 0.07 | 1.01 | 0.39 | 0.19 | 1.16 | 0.50 |
| **Cowpea** | **Nov** | **S1** | 0.60 | 2.59 | 1.68 | 1.02 | 2.59 | 1.63 |
| | | **S2** | 0.41 | 1.77 | 1.02 | 0.32 | 1.87 | 0.88 |
| | | **S3** | 0.16 | 1.08 | 0.51 | 0.16 | 1.15 | 0.46 |

### 4.2.3 Verification of land suitability maps

As discussed in **Section 3.6.3**, a verification of the land suitability maps was done to determine the accuracy of the methodology developed and used in this study. Data obtained from the DAFF (2014) and DAFF (2016) production guidelines were used to verify the land suitability maps of cowpea and bambara nut, respectively. This important step is not always considered in land suitability assessments. From the literature review, only one study (Estes et al., 2013; cf. **Section 2.1.2.1**) verified their land suitability map using observed (presence) data.

#### 4.2.3.1 Bambara nut

The production guidelines identified 11 local municipal districts as the main cultivation areas for bambara nut, where the centroid of each district is marked with an 'X' in **Figure 4.14**. DAFF (2016) identified the Ehlanzeni, Greytown and Nkandla as districts where bambara nut is grown extensively. These three regions align with the highly suitable areas shown in the figure below. However, it is important to note that the DAFF production guidelines are outdated and no longer fully reflect current cultivation areas.

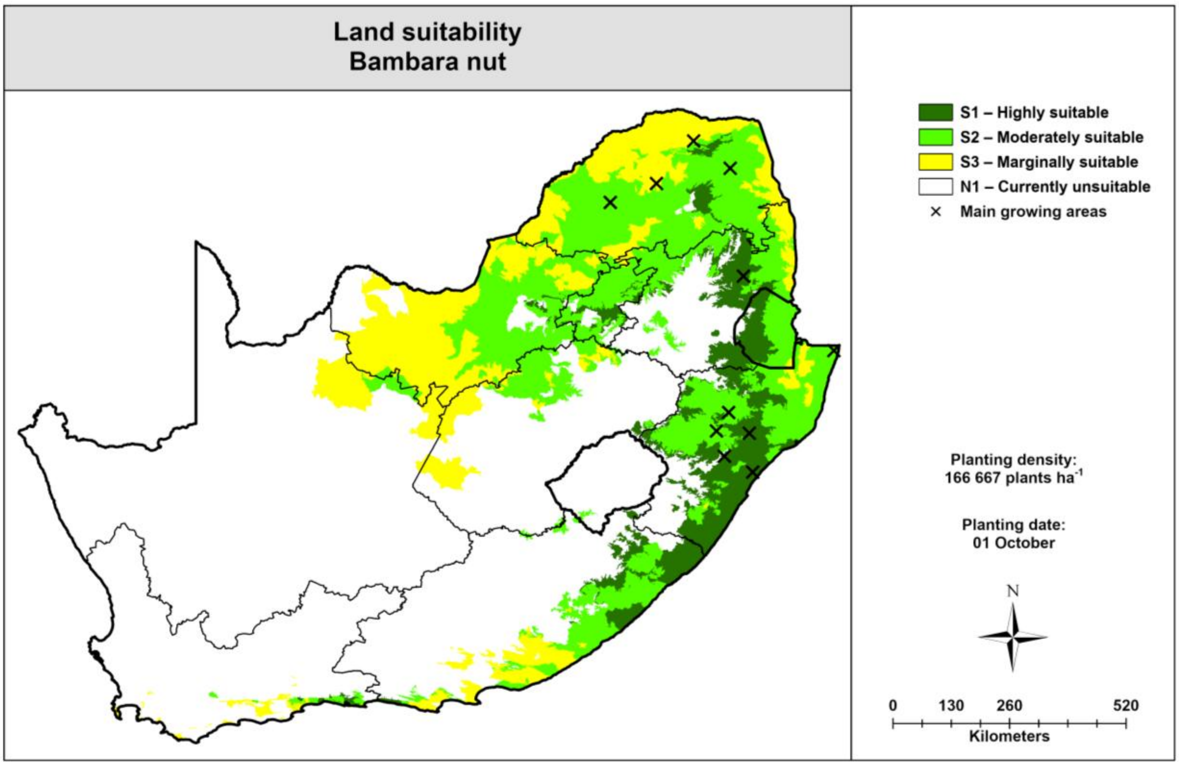


**Figure 4.14: Verification of suitable growing areas for bambara nut planted in October using the main producing areas listed by DAFF (2016)**

#### 4.2.3.2 Cowpea

DAFF (2014) identified 33 towns that were considered the main cowpea producing areas, which are shown in **Figure 4.15** with an 'X'. However, three towns (in northern Limpopo, central Mpumalanga and western North-West) did not align with the suitable production areas identified in this study. The AZ corresponding to the town in Limpopo and North-West was eliminated due to low $RFL_{AVE}$ (< 200 mm) whereas AZ no. 869 (town in Mpumalanga) has eliminated due to high RCF (> 33%), as highlighted in **Table 4.17**. These results suggest that either: (i) the land suitability approach is under-estimating suitable areas because the thresholds are set too high, (ii) the town locations represent irrigated cowpea production, or (iii) the climate and/or soil data used for each zone is not representative of the entire zone, i.e. micro-climates suitable for crop cultivation are excluded.

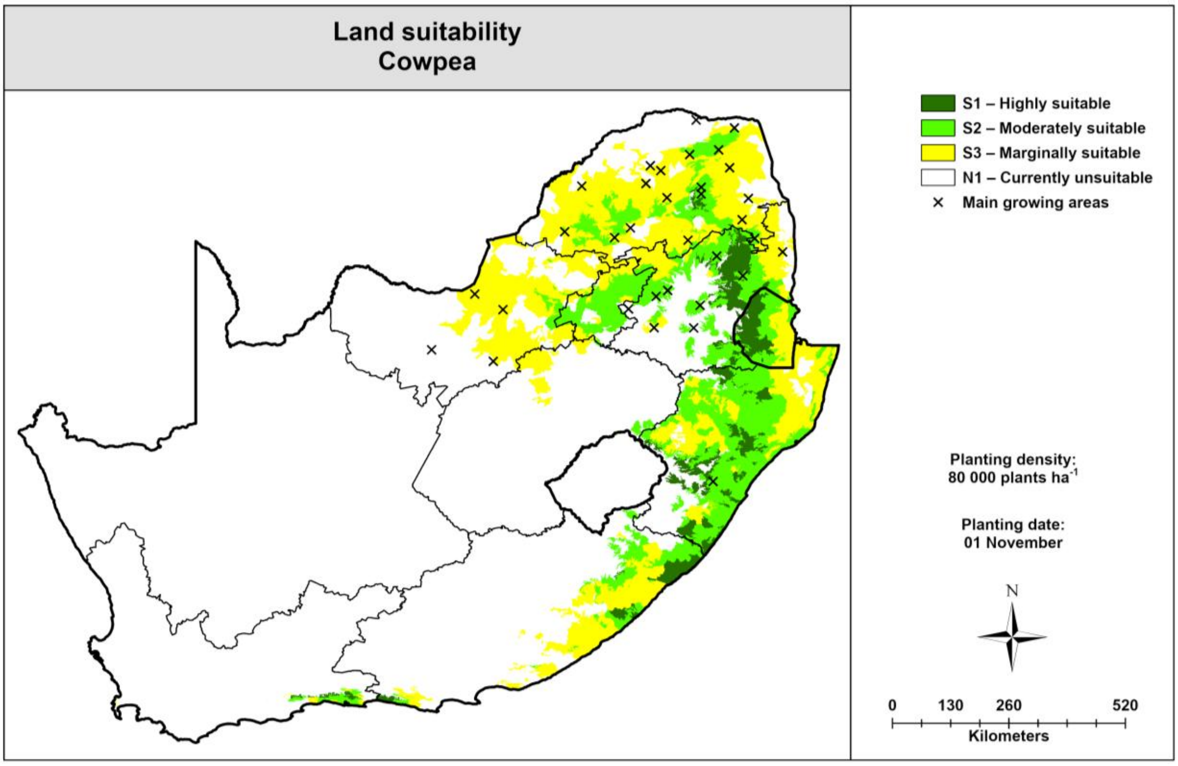


**Figure 4.15: Verification of suitable growing areas for cowpea planted in November using the main producing areas listed by DAFF (2014)**

**Table 4.17: AZs considered currently unsuitable (N1) for cowpea cultivation for three main producing areas**

| Province | AZ no. | $RFL_{AVE}$ (mm) | RCF (%) | $CWP_{CV}$ (%) |
|---|---|---|---|---|
| Limpopo | 345 | **158** | 12 | 101 |
| Mpumalanga | 869 | 365 | **39** | 132 |
| North-West | 990 | **165** | 43 | 143 |

### 4.2.4 Consideration of existing land use

To re-cap the approach described in **Section 3.6.4**, land uses representing mining, urban, water bodies, wetlands and protected areas were excluded as permanently unsuitable areas (N2), whilst sugarcane production areas, commercial forest plantations, indigenous forests and orchards were excluded as currently unsuitable areas (N1). The final land suitability map was produced by integrating the land use map produced in this study (**Figure 4.16**) with the suitability classification maps shown in **Section 4.2.2**. Since land cover data was unavailable for the neighbouring countries, suitable areas for crop production are likely to be over-estimated, especially in Eswatini.

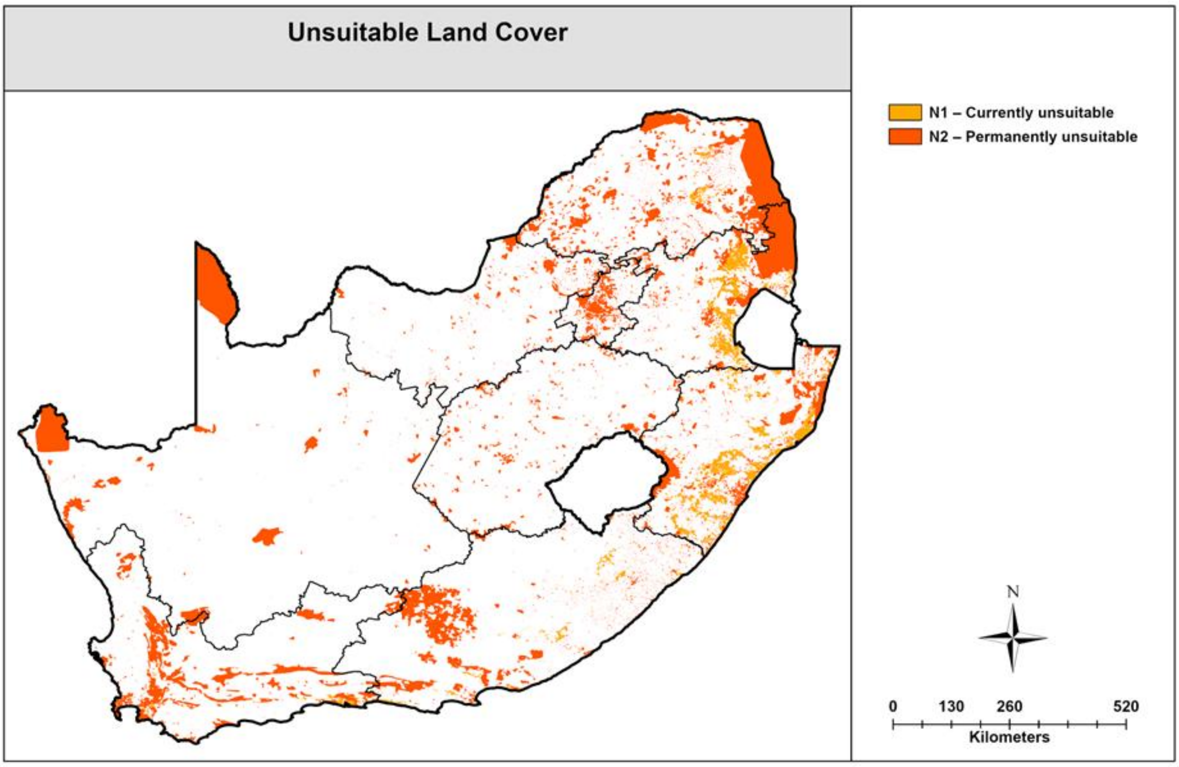


**Figure 4.16: Currently (N1) and permanently (N2) unsuitable areas based on the 2018 national land cover dataset**

#### 4.2.4.1 Bambara nut

**Table 4.18** illustrates the reduction in suitable areas for bambara nut production after existing land use was considered. For the October planting, marginally (S3) and moderately (S2) suitable areas decreased by 15.88% and 18.56% respectively, whereas highly suitable (S1) areas decreased by 28.98%. This resulted in a combined reduction in suitable areas of 19.45%. The table also highlights an overall decrease in suitable areas of 16.20%, 16.56% and 19.02% for the respective November, December and January plantings. These results highlight the importance of taking land use into consideration when developing land suitability maps to avoid land suitability for crop production from being over-estimated.

**Table 4.18: Reduction in suitability areas after consideration of land use for bambara nut**

| Class | Consideration of land use | Oct | Nov | Dec | Jan |
|---|---|---|---|---|---|
| S3 | Before (km$^2$) | 144 031 | 226 433 | 198 297 | 115 765 |
| | After (km$^2$) | 121 165 | 199 848 | 164 901 | 83 177 |
| | Reduction (km$^2$) | 22 866 | 26 858 | 33 396 | 32 588 |
| | Reduction (%) | **15.88** | 11.74 | 16.84 | 28.15 |
| S2 | Before (km$^2$) | 216 345 | 270 402 | 282 315 | 224 306 |
| | After (km$^2$) | 176 198 | 226 731 | 244 314 | 195 690 |
| | Reduction (km$^2$) | 40 147 | 43 671 | 38 001 | 28 616 |
| | Reduction (%) | **18.56** | 16.15 | 13.46 | 12.76 |
| S1 | Before (km$^2$) | 74 397 | 76 505 | 76 619 | 47 478 |
| | After (km$^2$) | 52 839 | 53 861 | 55 749 | 34 986 |
| | Reduction (km$^2$) | 21 558 | 22 644 | 20 870 | 12 492 |
| | Reduction (%) | **28.98** | 29.60 | 27.24 | 26.31 |
| Total | Before (km$^2$) | 434 773 | 573 340 | 557 231 | 387 549 |
| | After (km$^2$) | 350 202 | 480 440 | 464 964 | 318 853 |
| | **Reduction (km$^2$)** | **84 571** | **92 900** | **92 267** | **73 696** |
| | **Reduction (%)** | **19.45** | **16.20** | **16.56** | **19.02** |

#### 4.2.4.2 Cowpea

**Table 4.19** shows the reduction in suitable areas for cowpea after considering existing land use. For the November planting, highly suitable (S1) areas decreased by 31.84%, followed by S2 (23.07%) and S3 (21.12%). This equated to an overall reduction in suitable areas of 22.84%, with notable decreases of 24.46%, 19.77% and 19.72% for the October, December and January planting, respectively. Land use had a greater impact on bambara nut compared to cowpea, with an additional 17 933 km² of suitable land being eliminated. This difference was observed for the most suitable planting dates, with similar differences found across the other planting dates. This is due to the suitable area before land use is considered, with bambara nut having substantially more area to eliminate land use from.

**Table 4.19: Reduction in suitability areas after consideration of land use for cowpea**

| Class | Consideration of land use | Oct | Nov | Dec | Jan |
|---|---|---|---|---|---|
| S3 | Before (km$^2$) | 83 219 | 158 230 | 203 256 | 187 319 |
| | After (km$^2$) | 66 149 | 124 819 | 168 066 | 156 127 |
| | Reduction (km$^2$) | 17 070 | 33 411 | 35 190 | 31 192 |
| | Reduction (%) | 20.51 | **21.12** | 17.31 | 16.65 |
| S2 | Before (km$^2$) | 64 507 | 105 783 | 145 147 | 112 170 |
| | After (km$^2$) | 46 577 | 81 376 | 115 300 | 87 913 |
| | Reduction (km$^2$) | 17 930 | 24 407 | 29 847 | 24 257 |
| | Reduction (%) | 27.80 | **23.07** | 20.56 | 21.63 |
| S1 | Before (km$^2$) | 17 262 | 27 702 | 30 652 | 24 625 |
| | After (km$^2$) | 11 901 | 18 882 | 20 760 | 16 158 |
| | Reduction (km$^2$) | 5 361 | 8 820 | 9 892 | 8 467 |
| | Reduction (%) | 31.06 | **31.84** | 32.27 | 34.38 |
| Total | Before (km$^2$) | 164 988 | 291 715 | 379 055 | 324 114 |
| | After (km$^2$) | 124 627 | 225 077 | 304 126 | 260 198 |
| | **Reduction (km$^2$)** | **40 361** | **66 638** | **74 929** | **63 916** |
| | **Reduction (%)** | **24.46** | **22.84** | **19.77** | **19.72** |

### 4.2.5 Final land suitability map

As noted previously, the final land suitability map was produced by integrating the suitability classification maps with the land use map. Hence, the currently unsuitable (N1) areas in each suitability classification map (shown in white) were combined with the N1 areas from the land use map (shown in light orange). The combined N1 areas are also shown in white in the maps below, which could become suitable for crop production in the near future due to the changing climate and/or changes in land use dynamics. For example, commercial forestry areas might become available for crop cultivation if forestry companies adopt an agro-forestry approach to improve sustainability. The N2 areas (in dark orange) in the land use map (**Figure 4.16**) represent potentially suitable areas that are now deemed permanently unsuitable for crop production. Bambara nut and cowpea were selected in this study so that the final land suitability maps (for each planting month) could be compared to existing maps.

#### 4.2.5.1 Bambara nut

**Figure 4.17** shows the final land suitability map for bambara nut planted in October, combined with the land use data that identifies the N1 and N2 regions. Only one of the 14 case studies reviewed in this study (namely Nhamo et al., 2022) produced a land suitability map for bambara nut production for the Limpopo province only (cf. **Figure 2.1** in **Section 2.1.1.7**). A visual comparison of this map with **Figure 4.17** highlighted the northern Limpopo region, which is classified as not suitable by Nhamo et al. (2022), yet marginally suitable in this study. This discrepancy could be due to differences in the minimum seasonal rainfall thresholds of 350 mm and 200 mm used by Nhamo et al. (2022) and this study respectively. Hence, more research should be conducted on minimum climatic thresholds required for bambara nut cultivation in southern Africa.

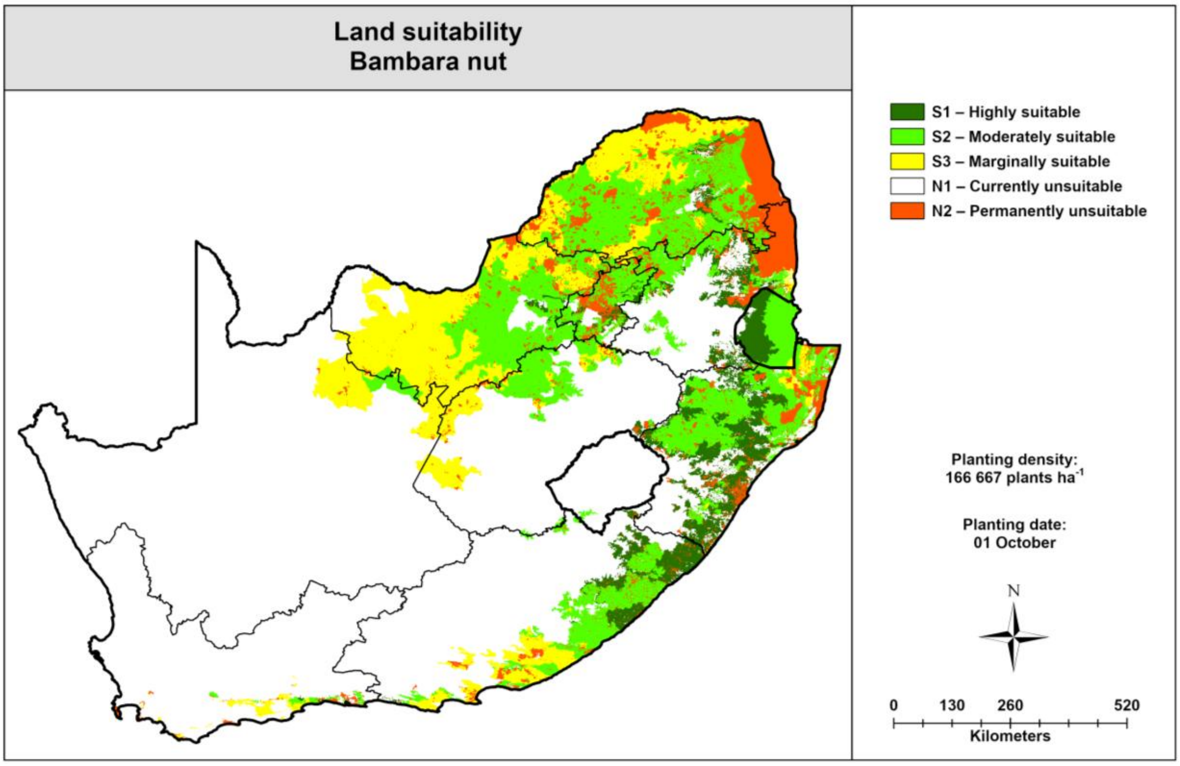


**Figure 4.17: Final land suitability map for bambara nut planted in October**

#### 4.2.5.2 Cowpea

**Figure 4.18** shows the equivalent final land suitability map for cowpea planted in November. Only two of the reviewed case studies produced a land suitability map for cowpea production. Mugiyo et al. (2022) used the MaxEnt model to develop a land suitability map for the KwaZulu-Natal province (cf. **Figure 2.2** in **Section 2.1.2.1**). A visual comparison of this map with **Figure 4.18** below highlights the north-eastern region of KwaZulu-Natal, which Mugiyo et al. (2022) classified as currently unsuitable for cowpea production, yet is mostly classified as marginal in this study and by Mugiyo et al. (2021b).

Mugiyo et al. (2021b) used the AHP method to develop a cowpea land suitability map for South Africa (cf. **Figure 2.4** in **Section 2.1.2.2**). A visual comparison of this map with **Figure 4.18** below highlighted the largest discrepancy in the Western Cape province, particularly along the western and southern coastlines. This region is characterised by predominantly winter rainfall and is unsuitable for cowpea production in the warmer summer season, unless under irrigation. Hence, it is unclear how Mugiyo et al. (2021b) classified this region as marginally suited to cowpea production. Similarly, the drier western regions of the North-West and Free

State provinces were also classified by Mugiyo et al. (2021b) as marginal. These regions were excluded in the study due to seasonally average rainfall ($RFL_{AVE}$) from November onwards being less than the 200 mm threshold for rainfed crop production (cf. **Figure 12.4** in **APPENDIX F**). In contrast, Mugiyo et al. (2021b) used a minimum rainfall threshold of 400 mm. This further highlights that more research should be conducted on minimum climatic thresholds required for cowpea cultivation in southern Africa.

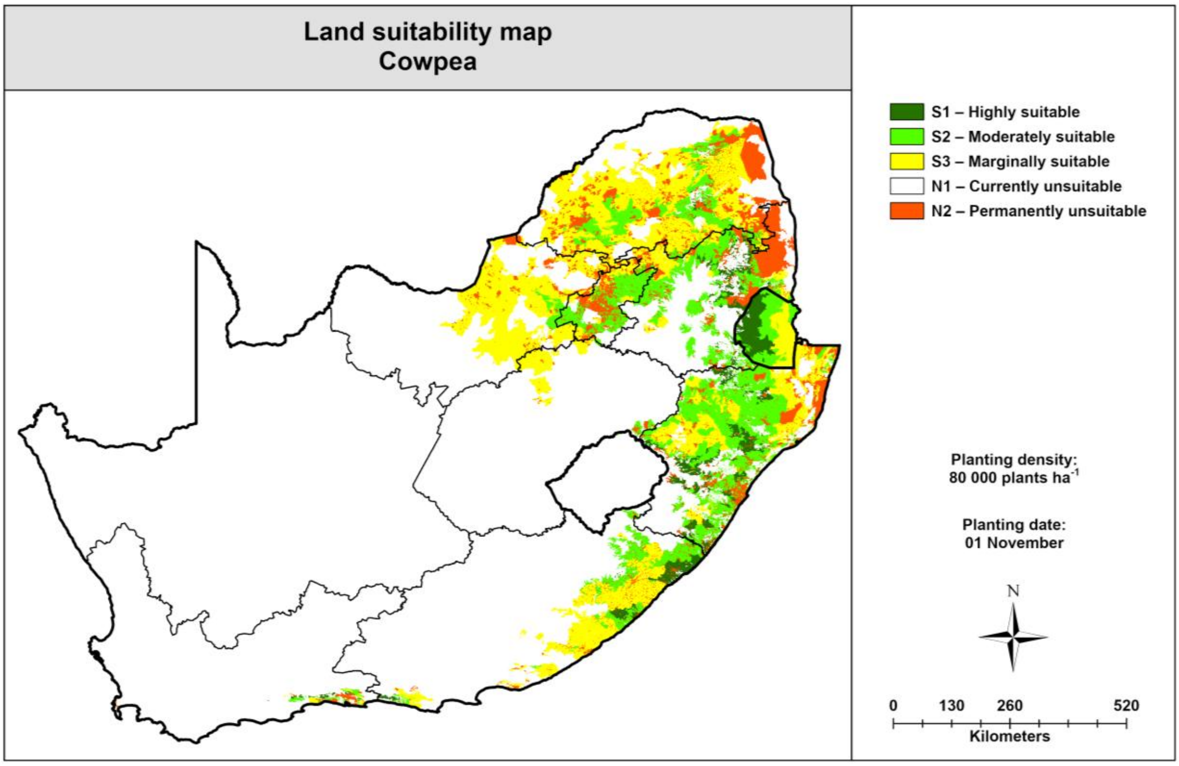


**Figure 4.18: Final land suitability map for cowpea planted in November**

The maps developed Mugiyo et al. (2021b) and Mugiyo et al. (2022) tend to over-estimate land suitability for cowpea production due to (i) how existing land use was considered, and (ii) the temporal scale of climate variables used. These two important topics are discussed next.

#### 4.2.5.3 Accounting for land use

Mugiyo et al. (2021b) and Mugiyo et al. (2022) excluded protected areas, indigenous forests, urban areas and water bodies from their land suitability maps, but not sugarcane production areas and commercial plantations. The latter two land uses occupy large areas along KwaZulu-Natal's coastline and adjacent interior, with forestry areas extending into the central and eastern parts of Mpumalanga. These are large, well-established industries that are highly unlikely to convert to NUC production, unless the latter becomes more profitable in the future.

This study found that suitable areas for cowpea cultivation decreased by 22.84% due to existing land use considerations.

#### 4.2.5.4 Accounting for climatic variability

Most traditional methods and modern methods tend to over-estimate land suitability for crop production since they mostly use long-term monthly and annual means of climate variables (e.g. rainfall and temperature). As a result, they do not consider inter-seasonal climate variability, nor the impact of droughts or frost that typically result in low yields or even total crop failure. In contrast, the approach used in this study utilised a crop simulation model that runs at a daily time step, which considers the effects of hot and dry periods on crop survival and final yield. In addition, this study utilised statistics that were generated from 49 consecutive seasons of AquaCrop simulations, which allowed for inter-seasonal variability to be considered. This explains the importance of using RCF and $CWP_{CV}$ in this study to eliminate unsuitable crop production areas. From **Section 4.2.1**, RCF had the greatest impact on eliminating unsuitable areas since this variable removed up to 66% of southern Africa's total area. Additionally, $CWP_{CV}$ was the third most influential factor removing up to 47% of unsuitable areas. These two variables highlight the importance of considering inter-seasonal variability in land suitability studies.

Mugiyo et al. (2021b) highlighted areas in the Western Cape, Free State as well as parts of the Eastern Cape and Mpumalanga as marginally to moderately suitable for cowpea production (cf. **Figure 2.4** in **Section 2.1.2.2**). This study, however, found these regions currently unsuitable due to a high RCF (**Figure 4.19**) and $CWP_{CV}$ (cf. **Figure 12.3** in **APPENDIX F**) as well as low $RFL_{AVE}$ (cf. **Figure 12.4**). Similarly, a visual comparison of **Figure 2.2** (cf. **Section 2.1.2.1**) developed by Mugiyo et al. (2022) with **Figure 4.18** highlights the south-western KwaZulu-Natal region that borders Lesotho and the Eastern Cape, which is classified as marginally to moderately suitable for cowpea production by Mugiyo et al. (2022). This region was eliminated in this study (i.e. currently unsuitable) due to the high RCF (cf. **Figure 4.19**) and high $CWP_{CV}$ (cf. **Figure 12.3** in **APPENDIX F**). This again highlights the importance of considering inter-seasonal variability in developing land suitability maps to avoid the over-estimation of areas deemed suitable for crop cultivation.

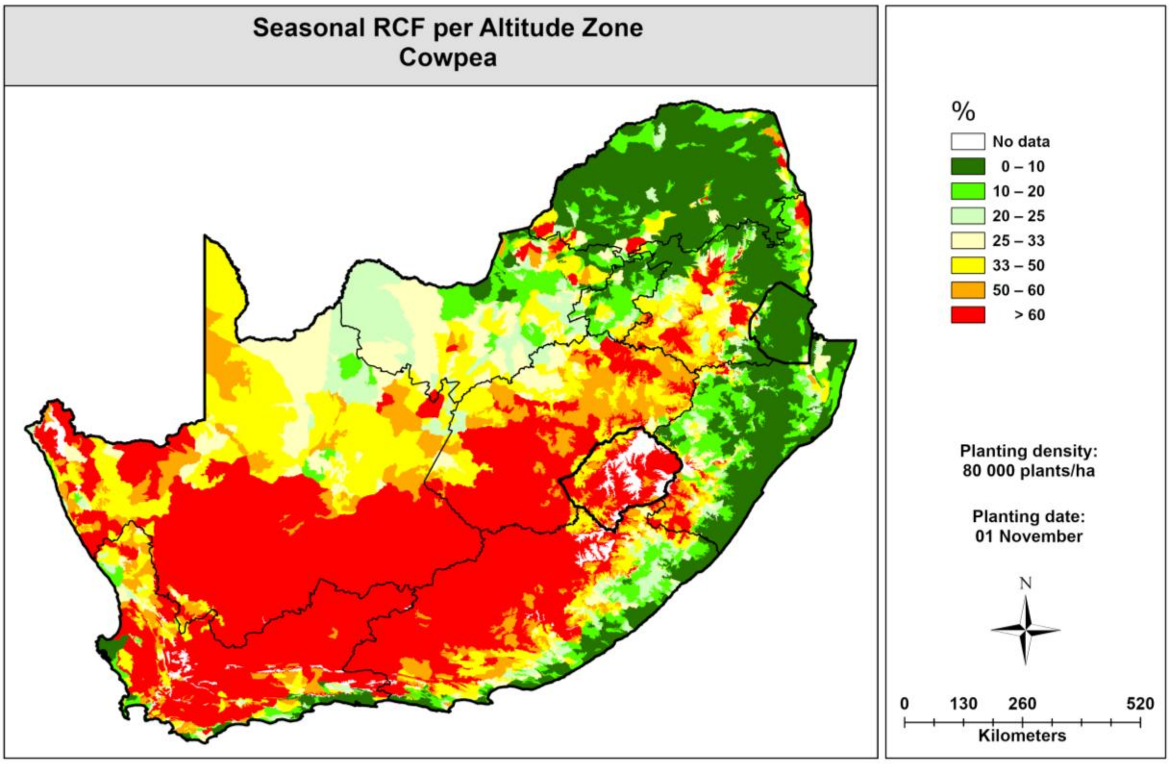


**Figure 4.19: Seasonal risk of crop failure (RCF) per Altitude Zone for cowpea planted in November**

#### 4.2.5.5 Impact of frost occurrence

Bambara nut: **Figure 4.20** shows the spatial distribution of potential bambara nut production areas in southern Africa for the October planting before frost was considered. A visual comparison with the land suitability in **Figure 4.9** (cf. **Section 4.2.2.1**) highlights notable differences in the Free State as well as parts of Mpumalanga, Northern Cape and Eastern Cape provinces. These areas are classified as marginally suitable in **Figure 4.20** (without frost), yet are deemed currently unsuitable in **Figure 4.9** (with frost). This difference is mostly due to the high RCF (cf. **Figure 12.5** in **APPENDIX F**) experienced in these areas due to frost occurrences. **Table 4.20** summarises the reduction in suitability areas after the consideration of frost, which shows a 49% and a 0.6% decrease in marginally and moderately suitable areas, respectively. In contrast, highly suitable areas showed an increase by 47.5%. This increase in S1 areas is a result of S2 areas (no frost) being re-classified as S1 (with frost). Similarly, S3 areas (no frost) are being re-classified as S2 (with frost). This increase in BREL due to frost, resulting in the suitability increasing is unknown and therefore warrants further research.

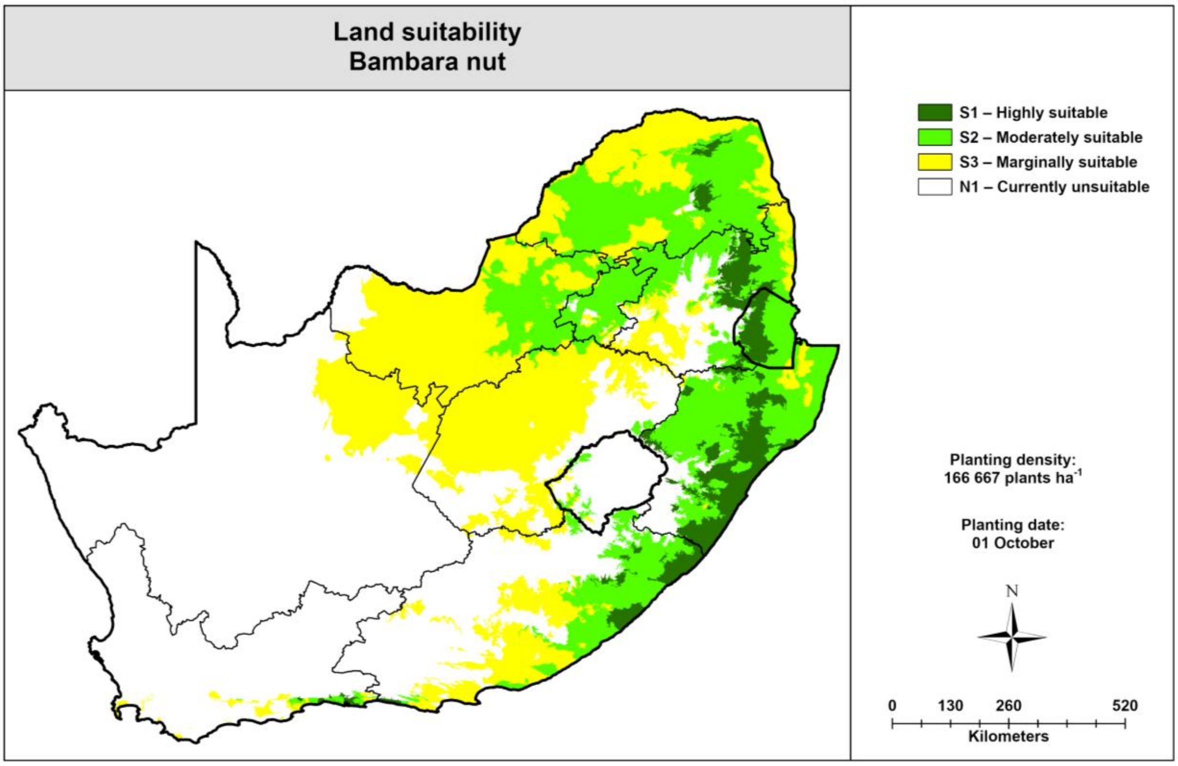


**Figure 4.20: Suitability classification map for rainfed production of bambara nut planted in October, based on $BREL_{NORM}$ without frost occurrence**

Cowpea: **Figure 4.21** shows that spatial distribution of potential cowpea production areas in southern Africa for November planting without the consideration of frost. A visual comparison with the land suitability map in **Figure 4.11** (cf. **Section 4.2.2.2**) reveals distinct differences. In the figure below, parts of the Free State, Eastern Cape and Mpumalanga provinces, along with areas in Lesotho, are classified as marginally and moderately suitable for cowpea production. However, the same regions are considered currently unsuitable in **Figure 4.11** (after frost was considered). Similar to bambara nut, these differences are due to the high RCF (cf. **Figure 4.19** in **Section 4.2.5.4**) and $CWP_{CV}$ (cf. **Figure 12.3** in **APPENDIX F**) when frost is considered. **Table 4.20** shows the ~28% decline in suitable areas across all three classes due to frost occurrence.

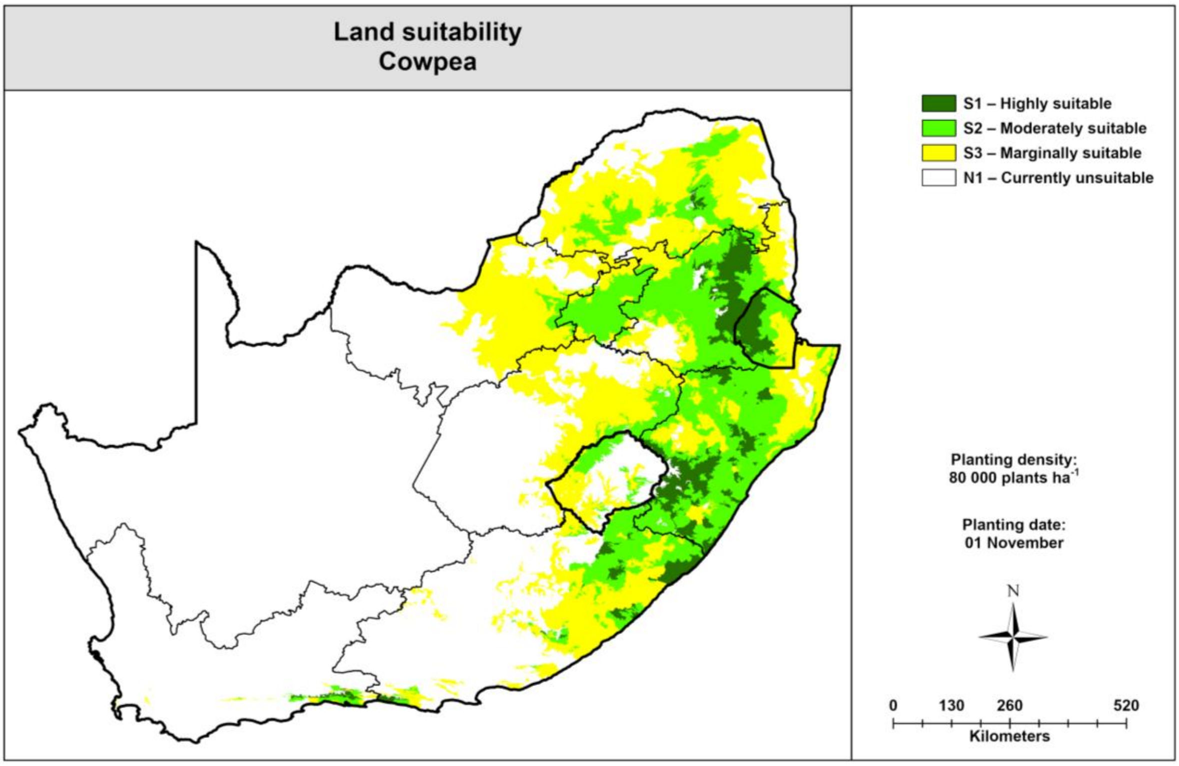


**Figure 4.21: Suitability classification map for rainfed production of cowpea planted in November for southern Africa, based on $BREL_{NORM}$ without frost occurrence**

The total reduction in suitable areas due to frost occurrence for bambara nut planted in October was 115 834 km$^2$ (or 9.11% of southern Africa). For cowpea planted in November, the total reduction in suitable areas was 116 106 km$^2$ (or 9.13%). This further highlights why considering frost in land suitability assessments for crop production is so important. Although frost has been shown to substantially reduce suitable areas for crop production (9.1%), the impact of existing land use was considerably greater, resulting in a 19.45% reduction for bambara nut and 22.84% for cowpea. This suggests that, while frost should still be considered in land suitability mapping, existing land use has a more pronounced effect and should be regarded as imperative in land suitability mapping approaches.

**Table 4.20: Reduction in suitability areas for bambara nut (October planting) and cowpea (November planting) after consideration of frost**

| Crop | Planting date | Class | Without frost ($km^2$) | With frost ($km^2$) | Change ($km^2$) | Change (%) |
|---|---|---|---|---|---|---|
| Bambara nut | Oct | S1 | 50 446 | 74 397 | 23 951 | 47.48 |
| | | S2 | 217 712 | 216 345 | -1 367 | -0.63 |
| | | S3 | 282 449 | 144 031 | -138 418 | -49.01 |
| Cowpea | Nov | S1 | 38 153 | 27 702 | -10 451 | -27.39 |
| | | S2 | 147 508 | 105 783 | -41 725 | -28.29 |
| | | S3 | 222 160 | 158 230 | -63 930 | -28.78 |

#### 4.2.5.6 Impact of management practices

In **Section 2.1.3.3**, it was noted that one of the major limitations to conventional land suitability mapping approaches was they did not consider management practices. However, CSMs like AquaCrop require both planting date and plant density as inputs, as well as being able to simulate the impacts of fertility, salinity and weed stress on crop yield. Since this study utilised output from AquaCrop to develop each land suitability map, the approach can incorporate management practices. For example, this study has clearly demonstrated the impact of planting date on land suitability mapping.

## 4.3 Automation of Methodology

As mentioned in **Section 3.7**, Python code was written to automate parts of the preliminary data analysis, such as mapping individual AquaCrop output variables and determining the which planting date produced the highest $Y_{AVE}$. For these two analyses, the average run time to develop the required .csv files was 17.49 seconds per planting month, or 139.92 seconds in total (i.e. four planting months for each of the two crops). The automation resulted in a more efficient workflow for checking individual variables and determining the highest yield in each AZ. The latter would require nested IF functions in MS Excel to perform the same process manually, which is more time consuming.

To efficiently test for collinearity between two variables, a Python script comprising of ~200 lines of code was developed (cf. **Section 4.1.3**). This script automated the calculation of the absolute value of the Pearson correlation coefficient, enabling the analysis of 17 statistics for 13 variables (cf. **Table 8.2** in **APPENDIX B**). The code produced 136 relationships per planting date, with an average run time of 10.24 seconds. This automation allowed for 1 088 relationships for both crops to be analysed relatively quickly and efficiently. This process ensured that all selected AquaCrop output variables using in this study for eliminating unsuitable AZs and classifying remaining zones were independent. One limitation of this approach is that not all relationships between two variables are linear (e.g. $BREL_{AVE}$ vs $Y_{AVE}$; cf. **Section 4.2.2.3**). Consequently, the calculated Pearson coefficient may not identify collinearity between two variables that exhibit a curvilinear relationship.

A partial automation of the land suitability mapping approach was implemented, consisting of two key phases: (i) generating the elimination .csv file, and (ii) creating the suitability classification .csv file (cf. **Section 3.7**). For the elimination phase, approximately 120 lines of Python code was written to exclude AZs deemed unsuitable based on the variables highlighted in **Section 4.2.1**, resulting in an average run time of 2.09 seconds for each planting date. For the suitability classification phase, an additional 100 lines of Python code was written to automate the normalisation of BREL and classify suitable areas into S1, S2 and S3 classes, which took 2.38 seconds on average to run. The total time taken to generate both .csv files required to map the elimination and classification areas was therefore 4.47 seconds. This automation allowed the study to map four planting dates and two crops, resulting in a total of eight land suitability maps being created. The automation of this section allowed for a lengthy trial and error approach to identify the fewest number of variables that successfully eliminated unsuitable zones. Furthermore, it allowed for a trial and error approach to determine the thresholds for RCF and $CWP_{CV}$ as well as determining the thresholds for classifying suitable areas.

However, due to time constraints, the full automation of the mapping process could not be achieved. Fully automating the mapping process in future studies could further enhance efficiency, allowing for additional crops, planting dates or planting densities to be mapped.

# 5. SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

The following chapter summarises the approach used to achieve this study's main aim and objectives. A summary of the main findings related to each objective is given. The limitations of this study are discussed, including recommendations for future work.

## 5.1 Summary of Approach

This main aim of this study was to determine where two NUCs can be grown in southern Africa under rainfed conditions using simulated output from FAO's AquaCrop model. It therefore provides knowledge on where and when these two NUCs should be grown in southern Africa. This knowledge contribution can assist in revising the outdated South African crop production guidelines for each NUC.

The literature review identified 14 case studies that produced land suitability maps using both traditional and modern methods, with only two previous studies (Estes et al., 2013 and Kunz et al., 2024) utilising data from a CSM. Prior to Kunz et al. (2024), no other study had applied thresholds to AquaCrop output to develop land suitability maps. Kunz et al. (2024) produced land suitability maps for OFSP and taro, after modifying an approach originally developed by Lake (2022). This study represents a continuation of this work, which was further improved in this study by, inter alia, accounting for frost occurrence in crop modelling, which was achieved for the first time (i.e. a novel approach).

The literature review also highlighted several gaps in land suitability assessments: (i) management practices (e.g. planting date and density) are seldom considered, (ii) frost occurrence has only been considered in one study (but not as a factor to restrict crop growth), (iii) crop cycle length has only been used in two case studies (with only one study applying an upper threshold to limit unrealistically long season lengths), (iv) collinearity amongst predictor variables remains largely unaddressed, (v) only three case studies used crop yield data to map land suitability, with another five using yield data solely for validation purposes, and (vi) no attempt has been made to fully automate the assessment of land suitability crop production.

The literature review identified various calibration and validation studies that developed AquaCrop parameters for 10 crops, of which nine are considered NUCs. The parameters represent local cultivars/landraces and growing conditions in South Africa. Using two statistics (RMSE and $R^2$) for the calibration and validation of CC development, these studies were ranked, which showed that cowpea and groundnut are considered the two best calibrated crops, whilst bambara nut and cowpea are the two best validated crops. Since the outcome of this study is dependent on the accuracy of modelled output, which in turn is strongly influenced by the quality of model input, bambara nut and cowpea were selected for land suitability mapping, since groundnut is not considered a NUC.

National model runs were conducted for both selected crops, each with four planting months (October to January) representing early to late planting. This task was achieved using an automated system developed by Kunz et al. (2024), which runs AquaCrop with input climate and soil data for 5 838 Altitude Zones. Statistics (e.g. mean and coefficient of variation) were automatically produced from simulations across 49 consecutive seasons. The length of the crop cycle (i.e. time from sowing to physiological maturity) for each AZ was calculated using thermal time in GDDs (not calendar days).

A first preliminary data analysis was conducted to determine the impact of planting date on simulated crop yield. Python code was written to determine which planting date produced the highest simulated yield averaged across all 49 seasons ($Y_{AVE}$) in each AZ. A second preliminary analysis was done to assess the impact of frost occurrence on $Y_{AVE}$ and the averaged crop cycle length ($CYCLE_{AVE}$). Histograms were generated using MS Excel to assess changes in $Y_{AVE}$ and $CYCLE_{AVE}$ due to frost occurrence. Maps were developed to illustrate spatial reductions in both variables. The AZs with the largest reductions in $Y_{AVE}$ and $CYCLE_{AVE}$ were also highlighted. A third preliminary data analysis tested for collinearity, which used Pearson's correlation coefficient to determine the strength of the linear relationship between 17 variables that could be used to assess land suitability for rainfed crop production. Python code was written to automate this process, which produced 136 relationships for each planting date, i.e. 1 088 tests for both crops.

An important objective of this study was to further improve the methodology used by Kunz et al. (2024) to map areas deemed suitable for rainfed crop production. The first major improvement was the inclusion of the first frost date to halt crop growth. The crop cycle length is reduced if the minimum air temperature drops to 5.4°C or less, before the crop reaches physiological maturity. This represents the first date after planting when the surface temperature reaches 0°C (or less) and the crop is likely to be killed by frost. Five years of hourly temperature data measured during the winter period by an automatic weather station in Pietermaritzburg was analysed to determine the relationship between air temperature (at 2 m) and ground (0 m) temperature. The national model runs were then repeated to determine the impact of frost occurrence on crop response. Hence, AquaCrop was run for two scenarios, namely with and without frost impacting the crop's growing season, i.e. 16 national model runs in total.

The second improvement was utilising average seasonal rainfall ($RFL_{AVE}$) rather than mean annual rainfall (MAP) to eliminate AZs deemed too dry to sustain rainfed crop production. AquaCrop calculated the daily rainfall accumulated from planting to the end-season date for up to 49 consecutive seasons, which were then averaged. A threshold of 200 mm was obtained from the literature as the minimum seasonal rainfall required for rainfed crop production of bambara nut and cowpea.

The third improvement was made to the risk of crop failure (RCF) calculation to correct this metric from being under-estimated. This was achieved by accounting for seasons not simulated by the model (i.e. $Y_{MISS}$), which occur when the model crashes unexpectedly during a seasonal simulation. The threshold applied to RCF values was also changed to 33%, representing a crop failure every third year (on average). The $CWP_{CV}$ variable was again used in this study to account for inter-seasonal variability, with a threshold of 150%. The combination of these three variables ($RFL_{AVE}$, RCF, $CWP_{CV}$) eliminated AZs deemed unsuitable for rainfed crop production. The model was not run for zones that are too cold for viable crop production, and thus were also eliminated, especially since these zones have no modelled data.

The final improvement was made to the classification of the remaining suitable AZs. This study utilised $BREL_{NORM}$ to classify suitable areas into S1 (> 75%), S2 (50-75%) and S3 (25-50%) classes. Zones where $BREL_{NORM}$ < 25% were reclassified as currently unsuitable, since less than a quarter of the biomass potential was reached. To verify the land suitability maps, districts of known bambara nut production areas, and towns where cowpea is grown, were obtained from existing (but outdated) production guidelines, which were then overlaid onto the land suitability maps.

Utilising the 2018 national land cover dataset for South Africa, this study also accounted for existing land use. Protected, urban, water bodies, wetlands and mining areas were eliminated as permanently unsuitable (N2) areas for crop production, whilst sugarcane production areas, commercial forestry plantations, indigenous forests and orchards were eliminated as currently unsuitable (N1) areas. The final land suitability maps were then produced by combining the suitable areas with the land use map using GIS software.

## 5.2 Summary of Findings

Findings from the preliminary data analysis revealed that planting date has a notable impact on $Y_{AVE}$. These findings correlated with literature that states bambara nut responds better to earlier plantings (October to November), compared to cowpea, which typically produces better yields when planted later in the season (November to December). Planting date is especially important for photosensitive crops such as bambara nut. Although AquaCrop cannot account for photoperiod effects on crop yield, it does consider the timing of water and temperature stress in relation to crop phenology, which explains the differences in $Y_{AVE}$ across the four planting dates.

Accounting for frost occurrence had a substantial impact on $CYCLE_{AVE}$, especially for bambara nut due to its longer growing season when compared to cowpea. For example, $CYCLE_{AVE}$ for bambara nut planted in January was reduced by more than 90 days across 61.4% of southern Africa's total area, compared to only 19.5% when planted in October. This also highlighted the impact that planting date had on frost occurrence and average crop cycle length. Frost had less impact on cowpea due to its shorter growing season, since only 24% of southern Africa experienced a reduction in $CYCLE_{AVE}$ > 90 days for a January planting, compared to only 6.9% when cowpea was planted in November. When the crop cycle length was shortened by frost occurrence, $Y_{AVE}$ was reduced as expected, thus highlighting the importance of accounting for frost in crop modelling, so that more accurate and reliable yields are simulated. Similarly, the decision to not run the model for AZs with cold climates is justified, since the risk of crop failure is typically high and it avoids unrealistic crop cycle lengths from being simulated, which results in over-estimated crop yields.

The collinearity tests revealed several AquaCrop output variables are strongly correlated to $RFL_{AVE}$, most notably $CWP_{AVE}$, $BREL_{AVE}$ and $Y_{AVE,}$ which indicates that rainfall is a primary driver of yield formation for both NUCs. Automation of the collinearity tests ensured that the relationship between all variable pairs was examined, confirming that the variables used for elimination of AZs are independent. It also revealed that the relationship between two variables is sometimes curvilinear.

The elimination of unsuitable AZs for rainfed crop production was mostly achieved using three variables, namely $RFL_{AVE}$, RCF and $CWP_{CV}$. $RFL_{AVE}$ eliminated the most zones for the November planting, compared to RCF for January. In contrast, RCF removed the most AZs for cowpea across all four planting dates. This highlights the importance of inter-seasonal variability, which is seldom accounted for in land suitability assessments. AquaCrop not only outputs data for assessing inter-seasonal variability, but it also models the impact of stress during critical phenological growth stages, which is strongly influenced by planting date.

The elimination of unsuitable AZs for rainfed crop production highlighted the importance of considering management practices (e.g. planting date) when assessing land suitability. For bambara nut, half and two-thirds of southern Africa's total area was eliminated for the November and January planting, respectively. Similarly for cowpea, the least (67%) and most (86%) area of southern Africa was eliminated for the December and October plantings, respectively. Based on these findings, November is the best time to plant bambara nut, while December is best for cowpea, i.e. since the least area was eliminated.

The classification of the remaining suitable areas highlighted that the most suitable planting dates (as mentioned above) do not necessarily produce the highest spatially averaged yields, which occur for the October and November plantings of bambara nut and cowpea, respectively. For bambara nut in particular, spatially averaged yield (across all suitable zones) decreased from the October to January planting due to the shortening of $CYCLE_{AVE}$, which again highlights the impact of planting date.

Although the histograms showed the majority of southern Africa is marginally to moderately suited to rainfed production of bambara nut and cowpea cultivation, the land suitability maps identified regions along the eastern seaboard of the country as being highly suitable, which aligns with the high rainfall regions of the country. However, using $BREL_{NORM}$ to classify suitable areas resulted in the over-estimation of S1 areas since high $BREL_{AVE}$ does not necessarily imply high $Y_{AVE}$. Instead of using $BREL_{NORM}$ for classifying land suitability, $HI_{NORM}$ was evaluated as a better alternative, with initial results highlighting its potential use.

A verification of the land suitability maps showed good agreement with known bambara nut and cowpea growing areas, despite the location data being limited and outdated. For example, 30 of the 33 towns where cowpea is grown aligned with the suitable growing areas identified in this study (for the November planting). Similarly, the centroid of all 11 bambara nut producing districts matched the suitable areas shown in the map (October planting), with five located in highly suitable areas.

The consideration of existing land use resulted in substantial reductions in suitable areas, highlighting the importance of integrating land use when assessing land suitability map for crop production. For both bambara nut (October planting) and cowpea (November planting), land use in 2018 resulted in a reduction of highly suitable areas by 29% and ~32%, respectively. The total reduction in suitable areas was 19% and 24% for bambara nut and cowpea respectively, with similar reductions for other planting dates. However, suitable crop production areas in Eswatini are over-estimated because the 2018 land cover dataset did not have data for this neighbouring country.

A visual comparison of the final land suitability maps with maps developed in previous studies highlighted some important differences in the methodologies used. For example, a minimum seasonal rainfall threshold of 200 mm was adopted in this study for both crops, whereas Nhamo et al. (2022) used 350 mm for bambara nut production in the Limpopo province. Similarly, Mugiyo et al. (2021b) applied a minimum rainfall threshold of 400 mm for rainfed production of cowpea across South Africa. Since rainfall is considered the primary driver of crop growth, using different thresholds affected the outcome of each land suitability assessment.

Most land suitability assessments use long-term means of monthly (or annual) rainfall, which do not account for inter-seasonal climate variability. This study showed that the majority of unsuitable production areas were eliminated using RCF and $CWP_{CV}$, thus highlighting the significance of accounting for inter-seasonal climate variability. The maps developed by Mugiyo et al. (2021b; 2022) tend to over-estimate suitable areas for cowpea production when compared to areas eliminated in this study, due mainly to high RCF and $CWP_{CV}$.

The comparison of the final land suitability maps with those developed in previous studies also highlighted the importance of removing land uses not suited to crop production, thus preventing over-estimation of land suitability. For bambara nut, the previous study overlooked mining and urban areas that are prevalent in Limpopo. For cowpea, both previous studies overlooked sugarcane production areas and commercial plantations. These well-established industries, particularly in KwaZulu-Natal, are unlikely to consider production of NUCs.

More importantly, most of the 14 case studies that were reviewed did not account for frost occurrence. This study showed substantial reductions in $CYCLE_{AVE}$ and $Y_{AVE}$, which resulted in suitable production areas being reduced by 21% and 28% for bambara nut and cowpea, respectively. This highlights why considering frost is so important since it improves the accuracy of the land suitability assessment.

This study demonstrated numerous advantages of using a water-driven CSM like AquaCrop for land suitability assessments. Other land suitability mapping techniques such as MaxEnt and AHP, cannot easily account for various factors that strongly influence crop response, such as (i) frost occurrence, and (ii) management practices (e.g. planting date), nor can they consider the (iii) timing of water stress in relation to critical phenological growth stages, or (iv) inter-seasonal variability in rainfall. In contrast, AquaCrop accounts for the latter three factors and was adapted in this study to handle frost impacts on, inter alia, crop cycle length and yield. Overall, the novel approach developed and used in this study provides a more realistic assessment of land suitability for rainfed crop cultivation, with future research potentially enhancing the approach by incorporating photoperiod sensitivity into the simulations.

## 5.3 Revisiting Aims and Objectives

The main aim of this study was to further improve a recently developed approach that utilised AquaCrop simulations to identify and map land deemed suitable for rainfed crop production of bambara nut and cowpea. Each specific objective was achieved as follows:

(i) *Assessing the impact of four planting dates on simulated crop yield*: this objective was achieved by determining the planting date that achieved the highest yield across all 5 838 AZs for both crops (cf. **Section 3.5.1**). The results showed that planting date has a large impact on simulated crop yield (cf. **Section 4.1.1**).

(ii) *Determining the impact of frost occurrence on crop cycle length and simulated yield*: this objective was met by (i) developing a novel method to reduce the crop cycle length when frost is likely to occur, which then kills the crop, and (ii) comparing $CYCLE_{AVE}$ and $Y_{AVE}$ values affected by frost, to those when frost was not considered (cf. **Section 3.5.1**). The results showed that the inclusion of frost led to substantial reductions in $CYCLE_{AVE}$ across all planting dates, with consequential decreases in $Y_{AVE}$ (cf. **Section 4.1.2**).

(iii) *Accounting for planting date and frost occurrence when mapping land suitability*: for each crop, this objective was achieved by running AquaCrop with four planting dates, both with and without frost occurrence (cf. **Section 3.6**). The results highlighted the importance of considering planting date (cf. **Section 4.2**) as well as frost (cf. **Section 4.2.5.4**), since both factors strongly influenced the final land suitability maps.

(iv) *Further improve the methodology and apply to other NUCs*: this objective was met by (i) considering four planting dates, (ii) restricting the crop cycle length by the first frost date, (iii) eliminating areas too dry for rainfed crop production using seasonal (not annual) rainfall, (iv) correcting the calculation of crop failure risk, and (v) using a different variable (BREL) for classification (cf. **Section 3.6**). For both crops, a more realistic land suitability map for rainfed crop production was produced, especially since frost occurrence was considered (cf. **Section 4.2.5.5**).

(v) *Partially automate the methodology using the Python programming language*: the last objective was achieved by writing code to automate the (i) preliminary data analysis, (ii) testing for collinearity, and (iii) parts of the land suitability mapping process (cf. **Section 3.7**). This automation facilitated the use of a trial and error approach for identifying the most suitable AquaCrop output variables for eliminating and classifying AZs, as well as for determining the thresholds used for each of the variables. It also reduced the time required to produce each of land suitability map (cf. **Section 4.3**).

### 5.4 Limitations of the Study

Bambara nut and cowpea were calibrated and validated at only one location (or agro-ecological zone). The assumption was made that the crop parameter files are representative of all crop growing areas, yet this has not been validated. The reliance on a single location for calibration may result in less reliable predictions for other locations with differing climatic and soil characteristics, thus reducing the overall accuracy and robustness of the land suitability assessments. This represents the main limitation of this study.

Another important limitation of this study is the temperature threshold (5.4°C) used to indicate frost occurrence. The threshold was derived from data measured at a single location but was then applied at a national scale. This assumes the threshold can be extrapolated to other areas, which is unlikely to be the case for all areas.

As noted in **Section 4.1.1.1**, AquaCrop potentially over-estimates yield for legumes (especially bambara nut) since it does not consider daylength. Using a water-driven CSM for simulating yields of photoperiod sensitive crops could therefore result is an over-estimation of land deemed suitable for rainfed crop production.

Another limitation is the subjectivity of thresholds used for each mapping variable, which is true for most land suitability assessments. Ideally, thresholds should be gleaned from literature. However, such information is either limited or fragmented, especially for under-researched crops. Instead, the thresholds for RCF, $CWP_{CV}$ and $BREL_{NORM}$ were determined through a trial and error approach.

As noted in **Section 4.2.2.3**, one of the challenges encountered while classifying suitable areas was the overestimation of highly suitable areas when using $BREL_{NORM}$. This issue arose from the wide range of $Y_{AVE}$ values observed as $BREL_{AVE}$ increased, which resulted in low $Y_{AVE}$ values with high $BREL_{AVE}$.

One of the issues regarding the AZs, is that each zone has a different area. Hence, 50% of AZs does not represent 50% of southern Africa's total area. For this reason, results were reported as a percentage of total area, instead of the total number of zones. Another assumptions made is that the modelled data is deemed representative of the entire AZ. An entire zone may be considered unsuitable for crop production, yet due to micro-climatic effects, parts of the zone may be able to sustain crop cultivation.

Due to computational constraints, the study utilised the 2018 NLC dataset for South Africa with a 1 $km^2$ spatial resolution. Ideally, the more recent 2022 NLC dataset at a higher resolution (20 m) should have been used to achieve a more precise land cover classification. This land cover dataset has no data for both Lesotho and Eswatini. This was not an issue for Lesotho since majority of the country was considered currently unsuitable for rainfed crop production. However, land cover for Eswatini was obtained late in the study, and thus could not be included due to time constraints.

## 5.5 Recommendations for Future Research

Based on the findings of this study, the following recommendations for future research are proposed:

(i) The crop parameter files should be validated against measured data from different agro-ecological zones, thus improvement their robustness.
(ii) The air to surface temperature relationship needs to be determined for different climates across the country and compared to the threshold value obtained in this study.
(iii) Future studies should account for daylength effects on crop growth, especially on photosensitive crops.
(iv) AquaCrop did not simulate any temperature stress for bambara nut across the entire southern African region, which is difficult to understand. Thus, future research needs to determine why this is the case, or if there is a problem with certain crop parameter values.
(v) Statistics for temperature over the growing period should also be derived, as this may help to better understand temperature related stress effects on crop yield.
(vi) The reason why AquaCrop sometimes crashes due to a division-by-zero error, especially when simulating cold seasons not well suited to crop production, warrants further investigation.
(vii) The study showed that running the model for cold seasons, even if CYCLE is reduced to less than 365 days by frost occurrence, still simulates mostly zero crop yields. Thus, future studies should rather not consider these zones, which will also speed up the model run times.
(viii) Future research should consider using another more suitable statistic to test for curvilinear relationships between two variables.
(ix) More research is required for NUCs, especially for determining thresholds such as minimum RFL required.

(x) The study showed that BREL could potentially be used as another elimination variable.

(xi) Since BREL resulted in overestimation of highly suitable areas, the use of HI as a more suitable variable for classifying suitability should be considered in the future.

(xii) Future studies should consider developing a gridded approach to running the AquaCrop model, as this would provide a uniform mapping unit.

(xiii) Remote sensing should be used to determine where both crops are currently grown, which will assist in verifying land suitability maps and for updating outdated crop production guidelines.

(xiv) Future studies should utilise more recent land cover data at a finer spatial resolution as well as include land cover data for Eswatini, thus resulting in more realistic land suitability maps.

(xv) Conducting a break-even yield calculation for each crop could help in determining appropriate yield thresholds for economically viable crop production, which could also be applied to eliminate unsuitable cropping areas.

(xvi) Fully automating the mapping process (e.g. using ModelBuilder) is also recommended, especially the exclusion of unsuitable cropland based on existing land use. This will reduce the overall time required to develop each land suitability map, thus facilitating additional scenarios to be analysed (e.g. other crops, planting dates and plant densities).

# 6. REFERENCES


ABEBE S, DERIBEW KT, ALEMU G, MOISA MB (2023) Modeling *Eragrostis tef Zucc* and *Hordeum vulgare* L. cropland in response to food insecurity in the Southwestern parts of Ethiopia. Heliyon **9**.

AGT FOODS AFRICA (2018) *Cowpeas – Vigna unguiculata*. AGT Foods Africa (Pty) Ltd., Krugersdorp, South Africa. https://agtfoods.co.za/wp-content/uploads/2018/07/15.-Cowpeas.pdf

AGYEMAN K, FRIMPONG F, AMANKWAA-YEBOAH P, OSEI-BONSU, YEBOAH S, KETEKU A, MARNO P, BREMPONG MB, DANQUAH EO, ADJEI AE, QUAYE MO, BERCHIE JN (2023) Physiological resilience of bambara groundnut (*Vigna subterranean* L. Verc) genotypes to intermittent periods of drought stress at different growth stages. Agricultural Sciences **14 (11)** 1573-1592.

AHMED FE, SULIMAN AS (2010) Effect of water stress applied at different stages of growth on seed yield and water-use efficiency of cowpea. Agriculture and Biology Journal of North America **1 (4)** 534-540.

AKBARI E, BOLOORANI AD, VERRELST J, PIGNATTI S, SAMANY NN, SOUFIZADEH S, HAMZEH S (2024) How global sensitive is the AquaCrop model to input parameters? A case study of silage maize yield on a regional scale. Frontiers in Agronomy **6**.

AKPOTI K, KABO-BAH AT, ZWART SJ (2019) Agricultural land suitability analysis: State-of-the-art and outlooks for integration of climate change analysis. Agricultural Systems **173** 172-208.

ALI S, UMAIR M, MAKANDA TA, SHI S, HUSSAIN SA, NI J (2024) Modeling current and future potential land distribution dynamics of wheat, rice, and maize under climate change scenarios using MaxEnt. Land **13**.

ALLEN RG, PEREIRA LS, RAES D, SMITH M (1998) *Crop evapotranspiration – guidelines for computing crop water requirements*. Irrigation and Drainage Paper No. 56, Chapter 6, Food and Agricultural Organisation, Rome, Italy. https://www.fao.org/4/X0490E/x0490e00.htm

BELETSE YG, LAURIE R, DU PLOOY CP, LAURIE SM, VAN DEN BERG A (2013) Simulating the yield response of orange fleshed sweet potato 'Isondlo' to water stress using the FAO AquaCrop model. Acta Horticulture **1007** 935-942.

BELLO ZA, WALKER S (2016) Calibration and validation of AquaCrop for pearl millet (*Pennisetum glaucum*). Crop and pasture science **67** 948-960.

BELLO ZA, WALKER S (2017) Evaluating AquaCrop model for simulating production of amaranthus (*Amaranthus cruentus*) a leavy vegetable, under irrigation and rainfed conditions. Agricultural and Forest Meteorology **247** 300-310.

BERTLING I, ODINDO AO (2013) Personal communication. University of KwaZulu-Natal, Pietermaritzburg, South Africa, 20 April 2013.

BONTHUYS J (2018) *Irrigation Water use*. The Water Wheel **17 (4)** 26-29. https://www.wrc.org.za/wp-content/uploads/mdocs/WW%20July_Aug%202018%20IRRIGATION%20WATER%20USE.pdf

BOUMAN BAM, VAN KEULEN H, VAN LAAR HH, RABBINGE R (1996) The 'school of de Wit' crop growth simulation models: A pedigree and historical overview. Agricultural Systems **52 (2-3)** 171-198.

BROUWER C, HEIBLOEM M (1986) *Irrigation water management: Irrigation water needs training manual no. 3*. Food and Agricultural Organisation, Rome, Italy. https://www.fao.org/3/s2022e/s2022e01.htm

CASTAÑEDA-VERA A, LEFFELAAR PA, ÁLVARO-FUENTES J, CANTERO-MARTÍNEZ C, MÍNGUEZ MI (2015) Selecting crop models for decision making in wheat insurance. European Journal of Agronomy **68** 97-116.

CHEN Y, YU J, KHAN S (2010) Spatial sensitivity analysis of multi-criteria weights in GIS-based land suitability evaluation. Environmental Modelling & Software **25 (11)** 1582-1591.

CHIBARABADA TP (2014) *Seed quality and water use characteristics of a bambara groundnut (Vigna subterranean L.) landrace differing in seed coat colour*. MSc dissertation, faculty of crop science, University of KwaZulu-Natal, KwaZulu-Natal, South Africa.

CHIBARABADA TP, MODI AT, MABHAUDHI T (2020a) Calibration and evaluation of AquaCrop for groundnut (*Arachis hypogaea*) under water deficit conditions. Agricultural and Forest Meteorology **281**.

CHIBARABADA TP, MODI AT, MABHAUDHI T (2020b) Options for improving water productivity: A case study of bambara groundnut and groundnut. Physics and Chemistry of the Earth **115**.

COLLINS MG, STEINER FR, RUSHMAN MJ (2001) Land-use suitability analysis in the United States: historical development and promising technological achievements. Environmental Management **28 (5)** 611-621.

DE LA ROSA D, MAYOL F, DIAZ-PEREIRA E, FENANDEZ M, DE LA ROSA JR. M (2004) A and evaluation decision support system (MicroLEIS DSS) for agricultural soil protection with special reference to the Mediterranean region. Environmental Modelling and Software **19** 929-942.

DEDEOĞLU M, DENGIZ O (2019) Generating of land suitability index for wheat with hybrid system approach using AHP and GIS. Computers and Electronics in Agriculture **167**.

DENNIS I, DENNIS R (2012) Climate change vulnerability index for South Africa aquifers. Water SA **38 (3)** 417-426.

DEPARTMENT OF AGRICULTURE, FORESTRY AND FISHERIES (DAFF) (2014) *Production guidelines for cowpea*. DAFF, Pretoria, South Africa. http://www.dalrrd.gov.za/phocadownloadpap/Brochures_and_Production_Guidelines/COWPEA%20PRODUCTION%20GUIDELINES%202014.pdf

DEPARTMENT OF AGRICULTURE, FORESTRY AND FISHERIES (DAFF) (2016) *Bambara groundnuts (Vigna subterranea)*. DAFF, Pretoria, South Africa. http://www.dalrrd.gov.za/phocadownloadpap/Brochures_and_Production_Guidelines/Bambara.pdf

DEPARTMENT OF AGRICTURE, RURAL DEVELOPMENT AND LAND ADMINISTRATION (DARDLA) (2012a) *Cowpea (Vigna unguiculata L.).* DARDLA, Pretoria, South Africa. https://dardlea.mpg.gov.za/documents/publications/tips/cowpea.pdf

DEPARTMENT OF AGRICTURE, RURAL DEVELOPMENT AND LAND ADMINISTRATION (DARDLA) (2012b) *Bambara groundnut (Vigna subterrnea (L.) verdc).* DARDLA, Pretoria, South Africa. https://dardlea.mpg.gov.za/documents/publications/tips/bambara_nuts.pdf

DEPARTMENT OF WATER AND SANITATION (DWS) (2018) *National Water and Sanitation Master Plan. Volume 1: Call to action* (final draft). https://www.dws.gov.za/National%20Water%20and%20Sanitation%20Master%20Plan/Documents/NWSMP(Master%20Plan)%20Call%20to%20Action%20v10.1.pdf

DOORENBOS J, KASSAM AH (1979) Yield response to water. Irrigation and Drainage Paper No. 33. Food and Agricultural Organisation, Rome, Italy. https://www.researchgate.net/publication/284800975_Yield_response_to_water_Irrigation_and_Drainage

DUA VK, GOVINDAKRISHNAN PM, SINGH BP (2014) Calibration of WOFOST model for potato in India. Potato Journal **41 (2)** 105-112.

ENDERS FB (2013) Collinearity, Encyclopedia Britannica. 04 August 2023. https://www.britannica.com/topic/collinearity-statistics

ESRI (2017) extAhp20 – Analytic Hierarchy Process for ArcGIS. https://www.arcgis.com/home/item.html?id=bb3521d775c94b28b69a10cd184b7c1f

ESRI (2023a) What is ModelBuilder? https://pro.arcgis.com/en/pro-app/latest/help/analysis/geoprocessing/modelbuilder/what-is-modelbuilder-.htm

ESRI (2023b) ArcGIS Pro (Version 3.3) [Software]. https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview

ESTES LD, BRADLEY BA, BEUKES H, HOLE DG, LAU M, OPPENHEIMER MG, SCHULZE R, TADROSS MA, TURNER WR (2013) Comparing mechanistic and empirical model projections of crop suitability and productivity: implications for ecological forecasting. Global Ecology and Biogeography **22** 1007-1018.

FOOD AND AGRICULTURAL ORGANISATION (FAO) (1976) *A framework for land evaluation*. FAO Soils bulletin 32. Food and Agricultural Organisation, Rome, Italy. https://www.fao.org/3/x5310e/x5310e00.htm

FOOD AND AGRICULTURAL ORGANISATION (FAO) (2015) *AquaCrop new features and updates Version 5.0*. Food and agricultural organisation, Rome, Italy. https://www.fao.org/3/bc087e/bc087e.pdf

GEERTS S, RAES D, GARCIA M, MIRANDO R, CUSICANQUI JA, TABOADA C, MENDOZA J, HUANCA R, MAMANI A, CONDORI O, MAMANI J, MORALES B, OSCO V, STEDUTO P (2009) Simulating yield response of quinoa to water availability with AquaCrop. Agronomy Journal **101 (3)** 499-508.

GOOGLE LLC (2023) Google Earth Pro (Version 7.3.6) [Software]. https://www.google.com/earth/versions/

GOUDRIAAN R, BASTIAANSSEN WGM (2013) *Land and water productivity in the Doukkala irrigation scheme, Morocco*. Internal Report to FAO Land and Water Division, Wegeningen, Netherlands. 89 pp.

HADEBE AT, MODI AT, MABHAUDHI T (2017) Calibration and testing of AquaCrop for selected sorghum genotypes. Water SA **43 (2)** 209-221.

HOLL MA, GUSH MB, HALLOWES J, VERSFELD DB (2007) *Jatropha curcas in South Africa: An Assessment of its Water Use and Bio-Physical Potential.* WRC Report No. 1497/1/07 Water Research Commission, Pretoria, South Africa.

IHEANACHO NS, LAWRENCE FS, OLUWOLE AB (2021) Effects of induced water stress at the flowering stage of bambara groundnut (*Vigna subterranean* L. Verdc) on its yield. American Journal of Science **9 (6)** 192-198.

ILUNGA K (2014) *Physiological responses of cowpea (Vigna unguiculata) to water stress under varying water regimes*. MSc dissertation, faculty of crop science, University of KwaZulu-Natal, KwaZulu-Natal, South Africa.

JAYATHILAKA PMS, SONI P, PERRET SR, JAYASURIYA HPW, SALOKHE VM (2012) Spatial assessment of climate change effects on crop suitability for major plantation crops in Sri Lanka. Regional Environmental Change **12 (1)** 55-68.

JEWITT G, WEN H, KUNZ R, CAN ROOYEN A (2009) *Scoping study on water use of crops/tree for biofuels in South Africa*. WRC Report No. 1772/1/09, Water Research Commission, Pretoria, South Africa.

JONES JW, HOOGENBOOM G, PORTER CH, BOOTE KJ, BACTHELOR WD, HUNT LA, WILKENS PW, SINGH U, GIJSMAN AJ, RITCHIE JT (2003) The DSSAT cropping system model. European Journal of Agronomy **18** 235-265.

KAMARA AY, OMOIGUI LO, KAMAI N, EWANISHA SU, AJEIGBE HA (2018) Improving cultivation of cowpea in West Africa. In: *Achieving sustainable cultivation of grain legumes Volume 2: Improving cultivation of particular grain legumes*, SIVASANKAR S, BERGVINSON D, GAUR P, KUMAR S, BEEBE S, TAMO M, 1-18. Burleigh Dodds Science Publishing, Cambridge, UK.

KANDA EK, SENANJE A, MABHAUDHI T (2021) Calibration and validation of the AquaCrop model for full and deficit irrigated cowpea (*Vigna unguiculata* (L.) Walp). Physics and Chemistry of the Earth **124**.

KAU AS, GRAMLICH R, SEWILAM H (2023) Modelling land suitability to evaluate the potential for irrigated agriculture in the Nile region in Sudan. Sustainable Water Resources Management **9 (10)**.

KEBEDE E, BEKEKO Z (2020) Expounding the production and importance of cowpea (*Vigna unguiculata* (L.) Walp.) in Ethiopia. Cogent Food and Agriculture **6 (1)**.

KENDABIE P, JØRGENSEN ST, MASSAWE F, FERNANDEZ J, AZAM-ALI S, MAYES S (2020) Photoperiod control of yield and sink capacity in bambara groundnut (*Vigna subterranean*) genotypes. Food and Energy Security **9 (4)**.

KEPHE PN, AYISI KK, PETJA BM (2021) Challenges and opportunities in crop simulation modelling under seasonal and projected climate change scenarios for crop production in South Africa. Agriculture and Food Security **10 (10)**.

KHALID F, ULLAH S, REHMAN F, HADI R, KHAN N, IBRAHIM F, KHAN T, AZIZ F, FEROZ DA, NADEEM SG, HUSSAIN M (2021) Identification of suitable sites for *Jatropha curcas* L. Bioenergy plantation using the AquaCrop model. Forests **12 (12)** 1772.

KHOMO TL (2014) *Spatial assessment of optimum and sub-optimum growing areas for selected biofuel feedstocks in South Africa*. MSc dissertation, Faculty of Hydrology, University of KwaZulu-Natal, KwaZulu-Natal, South Africa.

KUNZ R, MABHAUDHI T (2023) *Development of production guidelines and green water use estimation of underutilised indigenous crops using available models within selected bio-climatic regions of South Africa: Volume 2*. WRC Report No. 2717/2/23, Water Research Commission, Pretoria, South Africa.

KUNZ R, MASANGANISE J, REDDY K, MABHAUDHI T, LEMBEDE L, NAIKEN V, FERRER S (2020) *Water use and yield and soybean and grain sorghum for biofuel production*. WRC Report No. 2491/1/20, Water Research Commission, Pretoria, South Africa.

KUNZ R, DAVIS NS, THORNTON-DIBB SLC, STEYN JM, DU TOIT ES, JEWITT GPW (2015) *Assessment of biofuel feedstock production in South Africa: Atlas of water use and* *yield of biofuel crops in suitable growing areas (Volume 3)*. WRC Report No. 652/15, Water Research Commission, Pretoria, South Africa.

KUNZ R, REDDY K, MTHEMBU T, LAKE S, MABHAUDHI T, CHIMONYO V (2024) *Crop and nutritional water productivity of sweet potato and taro*. WRC Report No. 3124/1/24, Water Research Commission, Pretoria, South Africa.

LAKE SV (2022) *A novel approach for mapping areas suitable for rainfed production of taro*. Hydrology Honours project report, Centre for Water Resources Research, University of KwaZulu-Natal, KwaZulu-Natal, South Africa.

LOU L, SUN S, XUE J, GAO Z, ZHAO J, YIN Y, GAO F, LUAN X (2023) Crop yield estimation based on assimilation of crop models and remote sensing data: A systematic evaluation. Agricultural Systems **210**.

MABHAUDHI T (2012) *Drought tolerance and water-use of selected south African landraces of taro (Colocasia esculenta L. schott) and bambara groundnut (Vigna subterranea L. verdc)*. PhD, Faculty of crop science, University of KwaZulu-Natal, KwaZulu-Natal, South Africa.

MABHAUDHI T (2023a) Personal communication. Professor in Department of Crop Science, University of KwaZulu-Natal, South Africa, 29 June 2023.

MABHAUDHI T (2023b) Personal communication. Professor in Department of Crop Science, University of KwaZulu-Natal, South Africa, 17 October 2023.

MABHAUDHI T (2024) Personal communication. Professor in Department of Crop Science, University of KwaZulu-Natal, South Africa, 7 November 2024.

MABHAUDHI T, CHIMONYO VGP, CHIBARABADA TP, MODI AT (2017a) Developing a roadmap for improving neglected and underutilised crops: A case study of South Africa. Frontiers in Plant Science **8 2143**.

MABHAUDHI T, CHIMONYO VGP, MODI AT (2017b) Status of underutilised crops in South Africa: Opportunities for developing research capacity. Sustainability **9 (9)**.

MABHAUDHI T, CHIMONYO V, KUNZ R, MODI A (2023) *Volume 1: Developing a guideline for rainfed production of underutilised indigenous crops and estimating green water use of indigenous crops based on available models within selected bio-climatic regions of South Africa.* WRC Report No. 2717/1/23, Water Research Commission, Pretoria, South Africa.

MABHAUDHI T, MODI AT, BELETSE YG (2014a) Parameterization and testing of AquaCrop for a South African Bambara Groundnut landrace. Agronomy **106 (1)** 243-251.

MABHAUDHI T, MODI AT, BELETSE YG (2014b) Parameterisation and evaluation of FAO-AquaCrop model for a South African taro (*Colocasia esculenta* L. Schott) landrace. Agricultural and Forest Meteorology **192-193** 132-139.

MARRACCINI E, GOTOR AA, SCHEUER O, LECLERCQ C (2020) An innovative land suitability method to assess the potential for the introduction of a new crop at a regional level. Agronomy **10 (3)** 300.

MAUNDU P, KIOKO J, MUNENE C, HUNZIKER M (2023) *Cowpea (revised).* Infonet biovision. https://infonet-biovision.org/indigenous-plants/cowpea-revised

MAYES S, HO WK, CHAI HH, GAO X, KUNDY AC, MATEVA KI, ZAHRULAKMAL M, HAHIREE MKIM, KENDABIE P, LICEA LCS, MASSAWE F, MABHAUDHI T, MODI AT, BERCHIE JN, AMOAH S, FALOYE B, ABBERTON M, OLANIYI O, AZAM-ALI SN (2019) Bambara groundnut: an examplar underutilised legume for resilience under climate change. Planta **250** 803-820.

MBONG GA, AKEM CN, OLABI O, EMECHEBE AM, ALEGBIJO (2010) Effect of sowing dates on yield and yield components of cowpea infected with scab. Asian Journal of Agricultural Sciences **2 (2)** 57-62.

MCCOWN RL, HAMMER GL, HARGREAVES JNG, HOLZWORTH DP, FREEBAIRN DM (1994) APSIM: a novel software for model development, model testing and simulation in agricultural systems research. Agricultural systems **50** 255-271.

MICROSOFT (2024) Visual Studio Code, Version 1.94.2 [Software]. https://code.visualstudio.com/

MKANDAWIRE FL, SIBUGA KP (2002) Yield response on bambara groundnut to plant population and seedbed type. African Crop Science Journal **10 (1)** 39-49.

MODI AT, MABHAUDHI T (2016) *Developing a research agenda for promoting underutilised, indigenous and traditional crops*. WRC Report No. KV 362/16, Water Research Commission, Pretoria, South Africa.

MUGIYO H, CHIMONYO VGP, KUNZ R, SIBANDA M, HHAMO L, MASEMOLA R, MODI AT, MABHAUDHI T (2022) Mapping the spatial distribution of underutilised crop species under climate change using the MaxEnt model: A case of KwaZulu-Natal, South Africa. Climate Services **28** 100330.

MUGIYO H, CHIMONYO VGP, SIBANDA M, KUNZ R, MASEMOLA CR, MODI AT, MABHAUDHI T (2021a) Evaluation of land suitability methods with reference to neglected and underutilised crop species: A scoping review. Land **10 (2)** 125.

MUGIYO H, CHIMONYO VGP, SIBANDA M, KUNZ R, NHAMO L, MASEMOLA CR, DALIN C, Modi AT, Mabhaudhi T (2021b) Multi-criteria suitability analysis for neglected and underutilised crop species in South Africa. PLOS ONE **16 (1)** e0244734.

NHAMO L, PATERSON G, VAN DER WALT M, MOELETSI M, MODI A, KUNZ R, CHIMONYO V, MASUPHA T, MPANDELI S, LIPHADZI S, MOLWANTWA J, MABHAUDHI T (2022) Optimal production areas of underutilized indigenous crops and their role under climate change: Focus of Bambara groundnut. Frontiers in Sustainable Food Systems **6**.

NIE Y, CAI W, WANG C, HUANG G, DING Q, YU L, LI H, JI D (2019) Assessment of the potential and distribution of an energy crop at 1-km resolution from 2010 to 2100 in China – The case of sweet sorghum. Applied Energy **239** 395-407.

NYATHI MK, ANNANDALE JG, BELETSE YG, BEUKES DJ, DU PLOOY CP, PRETORIUS B, VAN HALSEMA GE (2016) Nutritional water productivity of traditional vegetable crops. WRC report No.2171/1/16, Water Resources Commission, Pretoria, South Africa.

NYATHI MK, VAN HALSEMA GE, ANNANDALE JG, STRUIK PC (2018) Calibration and validation of the AquaCrop model for repeatedly harvested leafy vegetables grown under different irrigation regimes. Agricultural Water Management **208** 107-119.

PHILIPS SJ, ANDERSON RP, SCHAPIRE RE (2006) Maximum entropy modelling of species geographic distributions. Ecological Modelling **190 (3-4)** 231-259.

PROFILLIDIS VA, BOTZORIS GN (2019) Statistical methods for transport demand modeling. In: *Modeling of transport demand: Analyzing, calculating, and forecasting transport demand*, Chapter 5, 163-224. Elsevier.

PYTHON SOFTWARE FOUNDATION (2023) Python (Version 3.11.9) [Software]. https://www.python.org/

RAES D, STEDUTO P, HSIAO TC, FERERES E (2018) *Reference Manual AquaCrop (Version 6.0-6.1)*. Food and Agricultural Organisation, Rome, Italy. https://openknowledge.fao.org/server/api/core/bitstreams/579aa71b-8dcf-4aa7-91fa-79ca0c52a5de/content

RAN H, KANG S, HU X, YAO N, WANG W, GALDO MV, CHALLINOR AJ (2022) A framework to quantify uncertainty of crop model parameters and its application in arid Northwest China. Agricultural and Forest Meteorology **316**.

RENNA M, MONTESANO FF, SIGNORE A, GONNELLA M, SANTAMARIA P (2018) BiodiverSO: A case study of integrated project to preserve the biodiversity of vegetable crops in Puglia (Southern Italy). Agriculture **8 (8)** 128.

ROFFE SJ, FITCHETT JM, CURTIS CJ (2021) Quantifying rainfall seasonality across South Africa on the basis of the relationship between rainfall and temperature. Climate Dynamics **56** 2431-2450.

SAAB MTA, TODOROVIC M, ALBRIZIO R () Comparing AquaCrop and CropSyst models in simulating barely growth and yield under different water and nitrogen regimes. Does calibration year influence the performance of crop growth models? Agricultural Water Management **147** 21-33.

SAATY RW (1987) The Analytic Hierarchy Process-What it is and how it is used. Mathematical Modelling **9 (3-5)** 161-176.

SAATY TL (2008) Decision making with the analytic hierarchy process. International Journal of Services Sciences **1 (1)** 83-98.

SALMAN M, Garcia-VILA M, FERERES E, RAES D, STEDUTO P (2021) *The AquaCrop model: enhancing crop water productivity. Ten years of development, dissemination and implementation 2009-2019*. FAO Water Report No. 47, Rome, Italy. https://openknowledge.fao.org/items/890f6033-554e-4e48-8dbb-5f1d809536e2

SARASWAT A, RAM S, ABDELRAHMAN MAE, RAZA MB, GOLUI D, HOMBEGOWDA HC, LAWATE P, SHARMA S, DASH AK, SCOPA A, RAHMAN MM (2023) Combining Fuzzy, Multicriteria and mapping techniques to assess soil fertility for agricultural development: A case study of Firozabad district, Uttar Pradesh, India. Land **12 (4)**.

SAXTON KE, RAWLS WJ (2006) Soil water characteristic estimates by texture and organic matter for hydrologic solutions. Soil Science Society of American Journal **70 (5)** 1569-1578).

SCHULZE RE (2008) Frost: Background Information. In: *South African atlas of climatology and agrohydrology*, Section 9.1. WRC Report No. 1489/1/06, Water Research Commission, Pretoria, South Africa.

SCHULZE RE, SCHÜTTE S (2016) Frost. In: SCHULZE RE (eds), *Handbook on adaption to climate for farmer, officials and others in the agricultural sector of South Africa*, Chapter B3.7, 144-148. Department of Agriculture, Forestry and Fisheries (DAFF), Pretoria, South Africa.

SCHULZE R, HORAN M (2011) Methods 1: Delineation of South Africa, Lesotho and Swaziland into quinary catchments. In: *Methodological approaches to assessing eco-hydrological responses to climate change in South Africa*, SCHULZE R, HEWITSON B, BARICHIEVY K, TADROSS M, KUNZ R, HORAN M, LUMSDEN T. WRC Report No. 1562/1/10, Chapter 6, 55-62, Water Research Commission, Pretoria, South Africa.

SCHULZE RE, MAHARAJ M (2004) Development of database of gridded daily temperature for Southern Africa. WRC Report No. 1156/2/04. Water Research Commission, Pretoria, South Africa.

SCHULZE RE, MAHARAJ M (2007) Sorghum Yield Estimation. In: *South African atlas of climatology and agrohydrology*, Section 16.4, SCHULZE, RE. WRC Report No. 1489/1/06, Water Research Commission, Pretoria, South Africa.

SHALOO, SINGH RP, BISHT H, JAIN R, SUNA T, BANA RS, GODARA S, SHIVAY YS, SINGH N, BEDI J, BEGAM S, TAMTA M, GAUTAM S (2022) Crop-suitability analysis using the analytic hierarchy process and geospatial techniques for cereal production in North India. Sustainability **14 (9)**.

SHRESTHA N (2020) Detecting multicollinearity in regression analysis. American Journal of Applied Mathematics and Statistics **8 (2)** 39-42.

SINCLAIR TR, SELIGMAN NAG (1996) Crop modelling: from infancy to maturity. Agronomy Journal **85 (5)** 698-704.

SINGHA C, SWAIN KC, SWAIN SK (2020) Best crop rotation selection with GIS-AHP technique using soil nutrient variability. Agriculture **10 (6)** 213.

SINGHA C, SWAIN KC, SAHOO S, ABDO HG, ALMOHAMAD H, DUGHAIRI AAA, ALBANAI J (2023) Total land suitability analysis for rice and potato crops through FuzzyAHP technique in West Bengal, India. Cogent food and agriculture **9 (1)**.

SLEIMAN R, GERARD JA, FAHED S, TODOROVIC M, SELLAMI MH, ROSSELLA A, SAAB MTA (2024) Irrigation and agricultural opportunities: Evaluating Hemp (*Cannabis sativa* L.) suitability and productivity in Lebanon. Water **16 (13)**.

SMITH JMB (1994) *Crop, Pasture and Timber Yield Index*. Cedara Report, N/A/94/4, Natal Agricultural Research Institute, Cedara, South Africa. 82 pp.

STEDUTO P (2006) *Biomass water-productivity: Comparing growth-engines of crop models*. WUEMED training course: Integrated approaches to improve drought tolerance in crops. Faculty of Agriculture, Bologna, Italy.

STEDUTO P, HSIAO TC, RAES D, FERERES E (2009) AquaCrop – The FAO Crop Model to Simulate Yield Response to Water: I. Concepts and Underlying Principles. Agronomy Journal **101 (3)** 426-435.

STEDUTO P, HSIAO TC, FERERES E, RAES D (2012) *Crop yield response to water*. Irrigation and Drainage Paper No. 66. Food and Agricultural Organisation, Rome, Italy. https://www.fao.org/4/i2800e/i2800e00.htm

STÖCKLE CO, DONATELLI M, NELSON R (2003) CropSyst, a cropping systems simulation model. European Journal of Agronomy **18** 289-307.

SYS IC, VAN RANST E, DEBAVEYE RJ (1991) Land evaluation part 1: Principals in land evaluation and crop production calculations. Agricultural publications No. 7, General Administration for Development Cooperation, Brussels, Belgium.

THOPIL GE, POURIS A (2016) A 20 year forecast of water usage in electricity generation for South Africa amidst water scarce conditions. Renewable and Sustainable Energy Reviews **62** 1106-1121.

TODOROVIC M, ALBRIZO R, ZIVTIC L, SAAB, MTA, STÖCKLE C, STEDUTO P (2009) Assessment of AquaCrop, CropSyst, and WOFOST models in the simulation of sunflower growth under different water regimes. Agronomy Journal **101 (3)** 509-521.

VAN DIEPEN CA, WOLF J, VAN KEULEN H, RAPPOLDT C (1989) WOFOST: a simulation model for crop production. Soil Use and Management **5 (1)** 16-24.

VANUYTRECHT E, RAES D, WILLEMS P (2014) Global sensitivity analysis of yield output from the water productivity model. Environmental Modelling and Software **51** 323-332.

WIMALASIRI EM, JAHANSHIRI E, PEREGO A, AZAM-ALI SN (2022) A novel crop shortlisting method for sustainable agricultural diversification across Italy. Agronomy **12 (7)**.

ZIADAT FM, SULTAN KA (2011) Combining current land use and famers' knowledge to design land-use requirements and improve land suitability evaluation. Renewable Agriculture and Food Systems **26** 287-296.

# 7. APPENDIX A

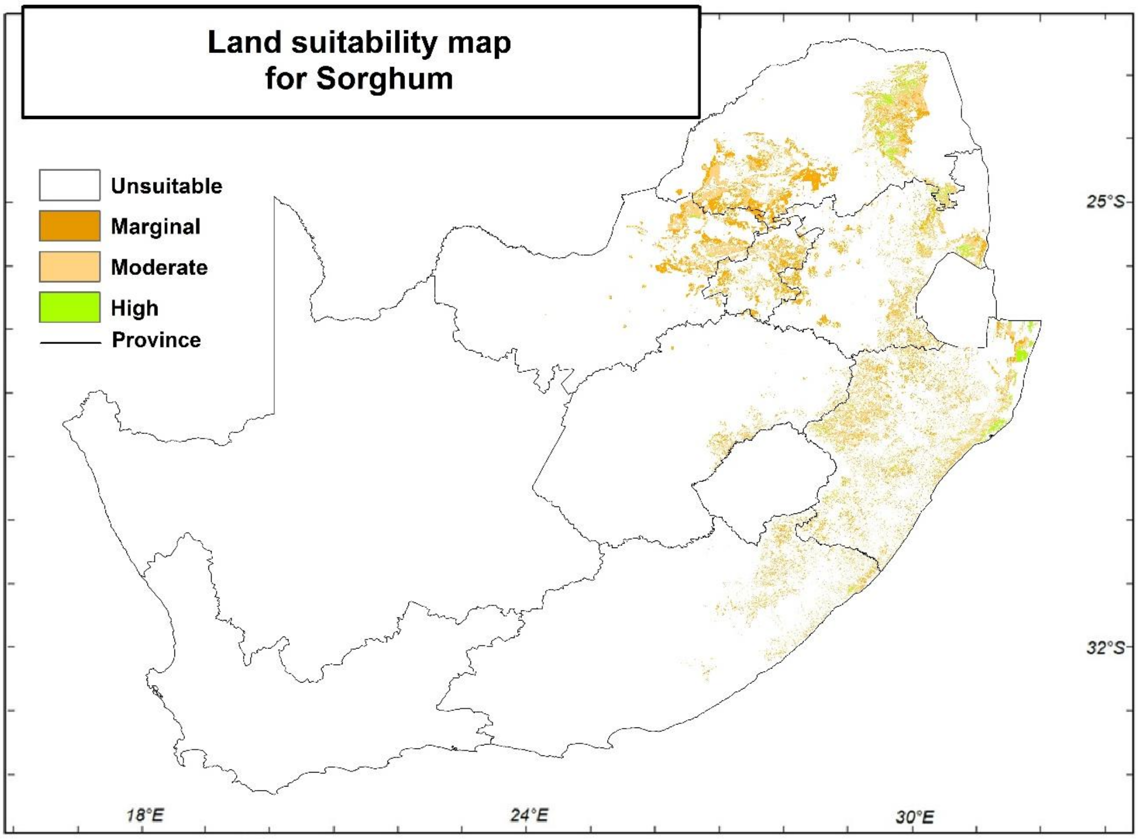


**Figure 7.1: Land suitability map for grain sorghum production (Kunz et al., 2015)**

**Table 7.1: Summary of growth criteria used in each traditional case study**

| Reference | Growth criteria | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Rain | Temp | RH | Frost | Soil depth | Soil fertility | Soil texture | Slope | Elev-ation | Topsoil clay content | Carbon-ates | AWC | Crop rotation | Gross margin |
| Holl et al. (2007) | | | | | | | | | | | | | | |
| Jewitt et al. (2009) | | | | | | | | | | | | | | |
| Kunz et al. (2015) | | | | | | | | | | | | | | |
| Nie et al. (2019) | | | | | | | | | | | | | | |
| Marraccini et al. (2020) | | | | | | | | | | | | | | |
| Khalid et al. (2021) | | | | | | | | | | | | | | |
| Nhamo et al. (2022) | | | | | | | | | | | | | | |

**Table 7.2:** **Pairwise comparison matrix (Mugiyo et al., 2021b)**

| Factors | Rainfall | Temp | $ET_O$ | LGP | Elevation | Slope | LULC | Soil depth | Distance to road | Weighting |
|---|---|---|---|---|---|---|---|---|---|---|
| **Rainfall** | 1 | 2 | 2 | 2 | 5 | 5 | 3 | 2 | 9 | 0.24 |
| **Temp** | 1/2 | 1 | 2 | 3 | 3 | 3 | 3 | 2 | 8 | 0.18 |
| **LGP** | 1/2 | 1/3 | 3 | 1 | 5 | 3 | 3 | 3 | 5 | 0.17 |
| **$ET_O$** | 1/2 | 1/2 | 1 | 1/3 | 5 | 3 | 3 | 2 | 5 | 0.13 |
| **Slope** | 1/5 | 1/3 | 1/3 | 1/3 | 2 | 1 | 2 | 2 | 5 | 0.08 |
| **Soil depth** | 1/2 | 1/2 | 1/2 | 1/3 | 2 | 1/2 | 2 | 1 | 5 | 0.08 |
| **LULC** | 1/3 | 1/3 | 1/3 | 1/3 | 2 | 1/2 | 1 | 1/2 | 3 | 0.06 |
| **Elevation** | 1/5 | 1/3 | 1/5 | 1/5 | 1 | 2 | 1/2 | 1/2 | 2 | 0.04 |
| **Distance to road** | 1/9 | 1/8 | 1/5 | 1/5 | 1/2 | 1/5 | 1/5 | 1/3 | 1 | 0.02 |

**Table 7.3:** **Growth criteria used in case studies involving machine learning techniques**

| Reference | Growth criteria | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Rain** | **Max. Temp** | **Min. Temp** | **LGP** | **Slope** | **Elev-ation** | **Soil pH** | **Soil depth** | **Soil texture** | **SOC** | **Distance to road** | **Distance to metros** |
| Estes et al. (2013) | | | | | | | | | | | | |
| Mugiyo et al. (2022) | | | | | | | | | | | | |

**Table 7.4: Growth criteria used in case studies involving AHP techniques**

| Reference | Growth criteria | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Rain | Temp | RH | LGP | $ET_O$ | Yield | Slope | Elev-ation | Soil pH | Soil depth | Soil texture | EC | SOC | Soil fertility | LULC | Distance to road |
| Jayathilaka et al. (2012) | | | | | | | | | | | | | | | | |
| Singha et al. (2020) | | | | | | | | | | | | | | | | |
| Mugiyo et al. (2021b) | | | | | | | | | | | | | | | | |
| Kau et al. (2023) | | | | | | | | | | | | | | | | |

**Table 7.5: Growth criteria used in case studies involving CSM techniques**

| Reference | Growth criteria | | | | |
|---|---|---|---|---|---|
| | Crop cycle length | $CWP_{AVE}$ | $CWP_{CV}$ | MAP | No. of yield simulations |
| Kunz et al. (2024) | | | | | |

# 8. APPENDIX B

AquaCrop simulated up to 49 seasonal values for each output variable shown in **Table 8.1**. Tests were then conducted to check for collinearity between each variable. A description of each variable is given next.

**Table 8.1: Output variables simulated by FAO's AquaCrop model (Raes et al., 2018)**

| AquaCrop variable | Description |
|---|---|
| RFL | Seasonal rainfall (mm) |
| CYCLE | Crop cycle length (days) |
| E | Soil water evaporation (mm) |
| E/Ex | Relative soil water evaporation (%) |
| Tr | Total transpiration of crop and weeds (mm) |
| Tr/Trx | Relative crop transpiration of crop and weeds (%) |
| TmpStr | Cold temperature stress (%) |
| ExpStr | Leaf expansion stress (%) |
| StoStr | Stomatal stress (%) |
| B | Above-ground biomass production (dry t $ha^{-1}$) |
| BREL | Relative biomass potential (%) |
| Y | Crop yield (dry t $ha^{-1}$) |
| CWP | Crop water productivity (dry kg $m^{-3}$) |

**RFL** Accumulated seasonal rainfall from planting date to end-season date (either the physiological maturity date or the first frost date).

**CYCLE** Length of the crop cycle from emergence to (i) physiological maturity date, or to (ii) frost occurrence date.

**E** Amount of water evaporated from the soil surface during the growing season.

**E/Ex** Relative soil water evaporation, i.e. ratio of actual to maximum evaporation.

**Tr** Amount of accumulated water transpired from the crop surface over the growing season.

**Tr/Trx** Relative transpiration, i.e. ratio of actual to maximum transpiration.

**TmpStr** When cold temperature stress occurs, crop growth is slowed or even halted. Cold stress occurs when the minimum air temperature is below the lower threshold (i.e. temperatures are too cold). The lower and upper temperature thresholds are crop dependent and obtained from the crop parameter file. The air temperature stress coefficient (Ks) value ranges from 0 (full stress) to 1 (i.e. no stress). A logistic shape curve is used in AquaCrop when considering Ks. The curve plots Ks against relative stress ($S_{rel}$).

| | |
|---|---|
| **ExpStr** | Leaf expansion is integral for photosynthesis and transpiration to occur. Larger leaves result in higher transpiration rates, and thus more growth. When the leaves are under stress, they stop growing and under prolonged stress will wither and die, thus reducing transpiration. The leaf expansion stress coefficient ($Ks_{exp,w}$) is related to total available soil water (TAW). The upper threshold is set slightly below field capacity and the lower threshold is set just above the permanent wilting point. Ks ranges from 0 (i.e. full stress) to 1 (i.e. no stress). A linear or concave Ks curve is used to determine $S_{rel}$ ($Ks_{exp,w}$ as a percentage). |
| **StoStr** | Under water stress, the stomata can close to reduce the amount of water lost through transpiration, which allows the plant to survive for a longer period. Hence, stomatal stress ($Ks_{sto}$) is related to soil water depletion. At field capacity, $Ks_{sto}$ is 1 (i.e. no stress; stomata close to prevent transpiration). At permanent wilting point, the crop becomes fully water stressed ($Ks_{sto}$ is 0). |
| **B** | Above-ground biomass is calculated by multiplying the normalised biomass water productivity (WP) parameter by the crop transpiration (Tr) to reference evapotranspiration ($ET_O$) ratio. WP is calculated as the above-ground dry matter (kg) that is produced per unit area ($m^2$), per unit of water transpired (mm). |
| **BREL** | Relative biomass is the biomass expected in a stressed field compared to that of a reference field with no stress, i.e. $100 \cdot B_{STR}/B_{REF}$. |
| **Y** | Final dry yield is calculated by multiplying B by the harvest index. |
| **CWP** | *C*rop water productivity (CWP), which is often referred to as water use efficiency (WUE); it is the yield produced in dry kg per $m^3$ of water evapotranspired over the growing season. CWP is used to evaluate how efficient a crop is in converting water into yield, i.e. “crop per drop” concept. |

Statistics (e.g. average and inter-seasonal variation) were then calculated for each output variable. Other variables were also calculated and used to eliminate AZs considered unsuitable for rainfed crop production and for classifying the remaining zones, as shown in **Table 8.3**.

**Table 8.2: Statistics generated for AquaCrop output variables used in the preliminary data analysis**

| Statistic generated | Description |
|---|---|
| $RFL_{AVE}$ | Seasonal rainfall average (mm) |
| $E_{AVE}$ | Soil water evaporation average (mm) |
| $E/Ex_{AVE}$ | Relative soil water evaporation average (%) |
| $Tr_{AVE}$ | Total transpiration of crop and weeds average (mm) |
| $Tr/Trx_{AVE}$ | Relative transpiration of crop and weeds average (%) |
| $TmpStr_{AVE}$ | Cold temperature stress average (%) |
| $ExpStr_{AVE}$ | Leaf expansion stress average (%) |
| $StoStr_{AVE}$ | Stomatal stress average (%) |
| $B_{AVE}$ | Above-ground biomass production average (dry t $ha^{-1}$) |
| $B_{CV}$ | Above-ground biomass production variability (%) |
| $BREL_{AVE}$ | Relative biomass potential average (%) |
| $BREL_{CV}$ | Variability in relative biomass potential (%) |
| $Y_{AVE}$ | Crop yield average (dry t $ha^{-1}$) |
| $Y_{CV}$ | Variability in crop yield (%) |
| $CWP_{AVE}$ | Crop water productivity average (dry kg $m^{-3}$) |
| $CWP_{CY}$ | Crop water productivity variability (%) |
| RCF | Risk of crop failure (%) |

**Table 8.3: Variables used for eliminating currently unsuitable areas and for classifying the remaining suitable areas**

| Variable | Description | Purpose |
|---|---|---|
| CYCLE | Crop cycle length (days) | Elimination |
| $RFL_{AVE}$ | Average seasonal rainfall (mm) | Elimination |
| $CWP_{CV}$ | Coefficient of variation (CV) in CWP (%) | Elimination |
| RCF | Risk of crop failure (%) | Elimination |
| $Y_{ZERO}$ | Number of zero yields | Calculation of RCF |
| $Y_{MISS}$ | Number of missing seasons | Calculation of RCF |
| $Y_{NUMB}$ | Number of yield simulations | Calculation of RCF |
| $BREL_{NORM}$ | Normalised average BREL (%) | Classification |
| $BREL_{AVE}$ | Average BREL (%) | Calculation of $BREL_{NORM}$ |

**CYCLE** Crop cycle length for each season.

**$RFL_{AVE}$** Average rainfall accumulated across each season.

**$CWP_{CV}$** Inter-seasonal variation in CWP.

**RCF** Risk of crop failure = $100 \cdot (Y_{ZERO} + Y_{MISS})/49$.

**$Y_{ZERO}$** Number of zero yields simulated by AquaCrop.

**$Y_{MISS}$** Number of missing yields = $49 - Y_{NUMB}$, i.e. number of seasons that were not simulated by AquaCrop, and thus have no data because the model crashed.

**$Y_{NUMB}$** Total number of yield simulations (maximum is 49).

**$BREL_{NORM}$** Normalised $BREL_{AVE}$ (using the maximum value of $BREL_{AVE}$).

**$BREL_{AVE}$** Average relative biomass potential.

# 9. APPENDIX C

**Table 9.1: AquaCrop calibration statistics for selected NUCs**

| Crop | Treatment | Cultivar/landrace | Type, location of trial | Sample size (n) | Canopy cover CC (%) | | | Source |
|---|---|---|---|---|---|---|---|---|
| | | | | | NRMSE | RMSE | $R^2$ | |
| **Cowpea** | Unstressed<br>Moderately stressed<br>Severely stressed<br>Rainfed (with irrigation) | Brown mix | Pot trial, UKZN<br>Field trial, Ukulinga | 14<br>15 | 12.2<br>15.7<br>19.6<br>18.5 | | 0.98<br>0.96<br>0.96<br>0.90 | Kanda et al. (2021) |
| **Groundnut** | Unstressed<br>Moderately stressed<br>Rainfed | *Arachis hypogaea L.* | Field trial, Ukulinga | 6 | | 4.8<br>6.3<br>7.4 | 0.98<br>0.84<br>0.92 | Chibarabada et al. (2020a) |
| **Spider flower** | Unstressed<br>Severely stressed | *Beta vulgaris* | Rain shelter, Roodeplaat | 13<br>15 | | 0.15<br>0.08 | 0.98<br>0.98 | Nyathi et al. (2018) |
| **Swiss chard** | Unstressed<br>Severely stressed | *Beta vulgaris* | Rain shelter, Roodeplaat | 15 | | 1.22<br>0.81 | 0.98<br>0.98 | Nyathi et al. (2018) |
| **Amaranth** | Unstressed<br>Severely stressed | *Amaranthus cruentus* | Rain shelter, Roodeplaat | 12 | | 0.06<br>1.24 | 0.99<br>0.98 | Nyathi et al. (2018) |
| **Taro** | Unstressed<br>Stressed | Dumbe dumbe L. | Greenhouse, UKZN | 22 | 27.8<br>8.6 | 14.0<br>3.4 | 0.71<br>0.98 | Kunz et al. (2024) |
| **Orange flesh sweet potato** | Unstressed<br>Stressed | Orange flesh (199062.1) | Greenhouse, UKZN | 19 | 12.1<br>17.8 | 8.8<br>12.0 | 0.92<br>0.76 | Kunz et al. (2024) |
| **Bambara nut** | Rainfed (with irrigation) | Red L. | Field trial, Roodeplaat | 6 | | 3.37 | 0.94 | Mabhaudhi et al. (2014a) |
| **Taro** | Rainfed (with irrigation) | Umbumbulu L. | Field trial, Ukulinga | 8 | | 2.38 | 0.79 | Mabhaudhi et al. (2014b) |
| **Sorghum** | Rainfed (with irrigation) | Macia<br>Ujiba<br>Pan8816 | Field trial, Ukulinga | 15 | | 8.4<br>9.2<br>14.4 | 0.87<br>0.84<br>0.66 | Hadebe et al. (2017) |
| **Orange flesh sweet potato** | Unstressed | Orange flesh (*Bophelo)* | Rain shelter, Roodeplaat | 15 | | 12.1 | 0.77 | Nyathi et al. (2016) |
| **Pearl millet** | Rainfed (with irrigation) | GCI17<br>Monyaloti | Field trial, UFS | 5<br>6 | | 13.0<br>17.9 | 0.91<br>0.91 | Bello and Walker (2016) |
| **Amaranth** | Rainfed (with irrigation) | *Amaranthus cruentus*<br>ex Arusha | Field trial, UFS | 5 | | 20.8 | 0.12 | Bello and Walker (2017) |
| **Orange flesh sweet potato** | Unstressed | Orange flesh (Isondlo*)* | Rain shelter, Roodeplaat | 4 | | | 0.92 | Beletse et al. (2013) |

**Table 9.2: AquaCrop validation statistics for selected NUCs**

| Crop | Treatment | Cultivar/Landrace | Type, location of trial | Sample size (n) | Canopy cover CC (%) | | | Source |
|---|---|---|---|---|---|---|---|---|
| | | | | | NRMSE | RMSE | $R^2$ | |
| **Bambara nut** | Unstressed<br>Moderately stressed<br>Severely stressed<br>Irrigated<br>Rainfed | Red L. | Rain shelter, Roodeplaat<br>Rain shelter, Roodeplaat<br>Rain shelter, Roodeplaat<br>Field trial, Roodeplaat<br>Field trial, Roodeplaat | 10<br><br><br><br>6 | | 14.1<br>18.3<br>18.1<br>9.7<br>6.2 | 0.95<br>0.90<br>0.81<br>0.86<br>0.95 | Mabhaudhi et al. (2014a) |
| **Cowpea** | Unstressed<br>Moderately stressed<br>Severely stressed | Brown mix | Pot trial, UKZN | 14 | 15.4<br>17.8<br>37.5 | | 0.94<br>0.96<br>0.85 | Kanda et al. (2021) |
| **Swiss chard** | Unstressed<br>Severely stressed | *Beta vulgaris* | Rain shelter, Roodeplaat | 15 | | 1.58<br>0.94 | 0.98<br>0.98 | Nyathi et al. (2018) |
| **Spider flower** | Unstressed<br>Severely stressed | *Beta vulgaris* | Rain shelter, Roodeplaat | 16 | | 1.45<br>1.28 | 0.94<br>0.97 | Nyathi et al. (2018) |
| **Amaranth** | Unstressed<br>Severely stressed | *Amaranthus cruentus* | Rain shelter, Roodeplaat | 12 | | 1.26<br>1.65 | 0.98<br>0.64 | Nyathi et al. (2018) |
| **Taro** | Rainfed (with irrigation)<br>Rainfed (with irrigation) | Umbumbulu landrace | Field trial, Ukulinga<br>Mare and Modi (2009) | 8 | | 1.85<br>20.2 | 0.84<br>0.02 | Mabhaudhi et al. (2014b) |
| **Groundnut** | Rainfed | *Arachis hypogaea* L. | Field trial, Fountainhill<br>Field trial, Umbumbulu | 6<br>5 | <br>10.30 | 11.8<br>3.3 | 0.83<br>0.97 | Chibarabada et al. (2020a) |
| **Orange flesh sweet potato** | Unstressed | Orange flesh (Bophelo) | Rain shelter, Roodeplaat | 15 | | 4.98 | 0.99 | Nyathi et al. (2016) |
| **Pearl millet** | Rainfed (with irrigation) | GCI17<br>Monyaloti | Field trial, UFS | 11 | | 8.6<br>13.6 | 0.81<br>0.82 | Bello and Walker (2016) |
| **Sorghum** | Rainfed (with irrigation) | Pan8816<br>Macia<br>Ujiba | Field trial, Ukulinga | 15 | | 12.4<br>19.5<br>22.7 | 0.86<br>0.79<br>0.711 | Hadebe et al. (2017) |
| **Taro** | Rainfed | Dumbe dumbe L. | Field trial, Fountainhill | 10 | | | | Kunz et al. (2024) |
| **Orange flesh sweet potato** | Rainfed | Orange flesh (199062.1) | Field trial, Fountainhill | 10 | | | | Kunz et al. (2024) |
| **Amaranth** | Rainfed (with irrigation) | *Amaranthus cruentus* ex Arusha | Field trial, UFS | 7 | | | | Bello and Walker (2017) |
| **Orange flesh sweet potato** | Unstressed | Orange flesh (Isondlo) | Rain shelter, Roodeplaat | 4 | | | | Beletse et al. (2013) |

# 10. APPENDIX D

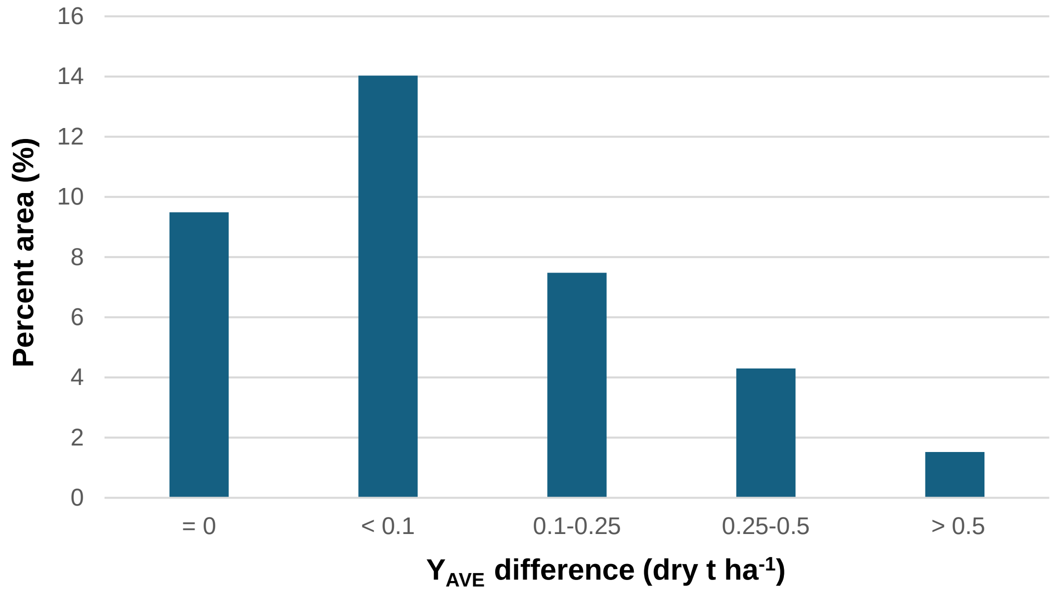


**Figure 10.1: Difference in $Y_{AVE}$ between the October and November planting dates for bambara nut**

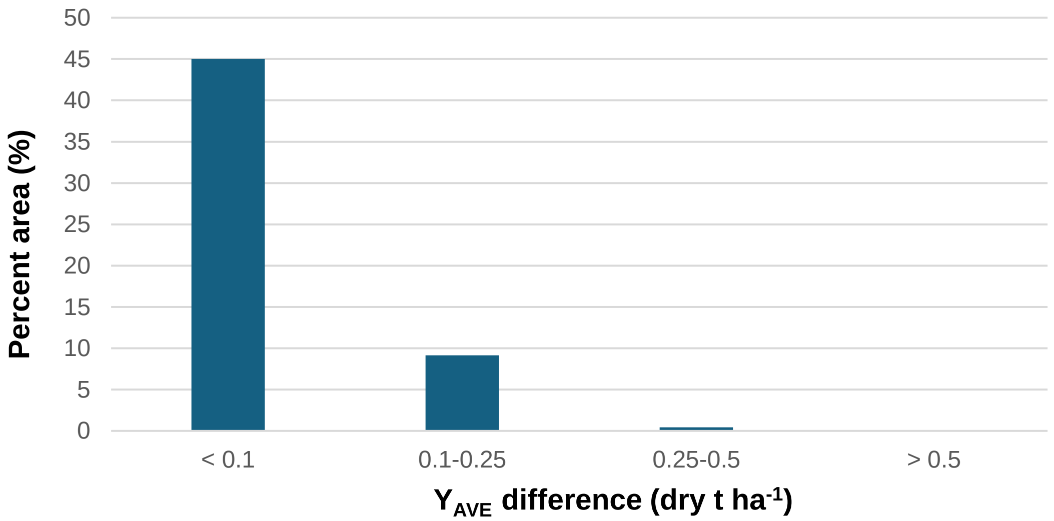


**Figure 10.2: Difference in $Y_{AVE}$ between the November and October plantings date for bambara nut**

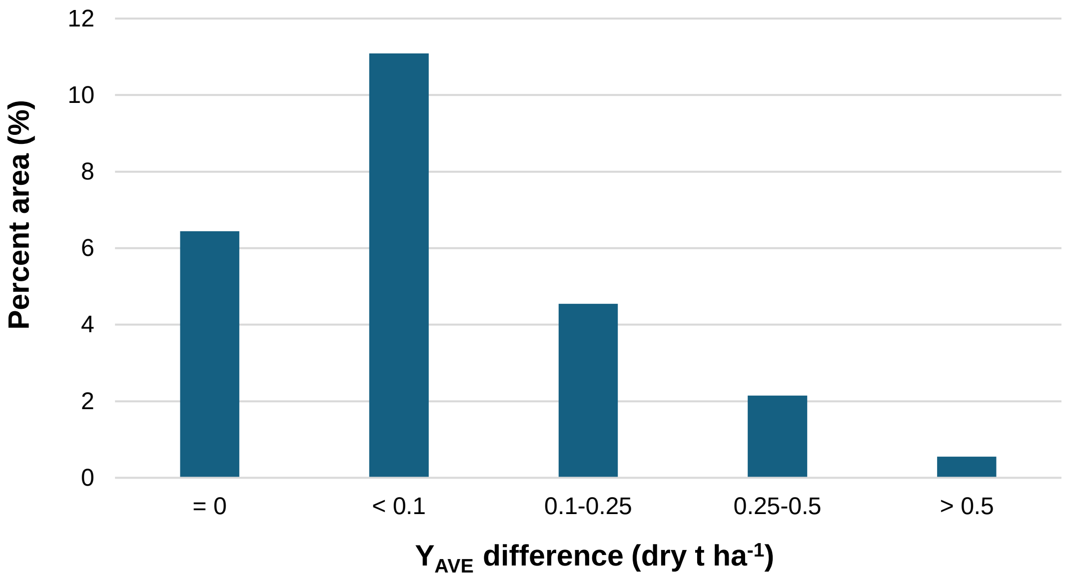


**Figure 10.3: Difference in $Y_{AVE}$ between the November and December planting dates for cowpea**

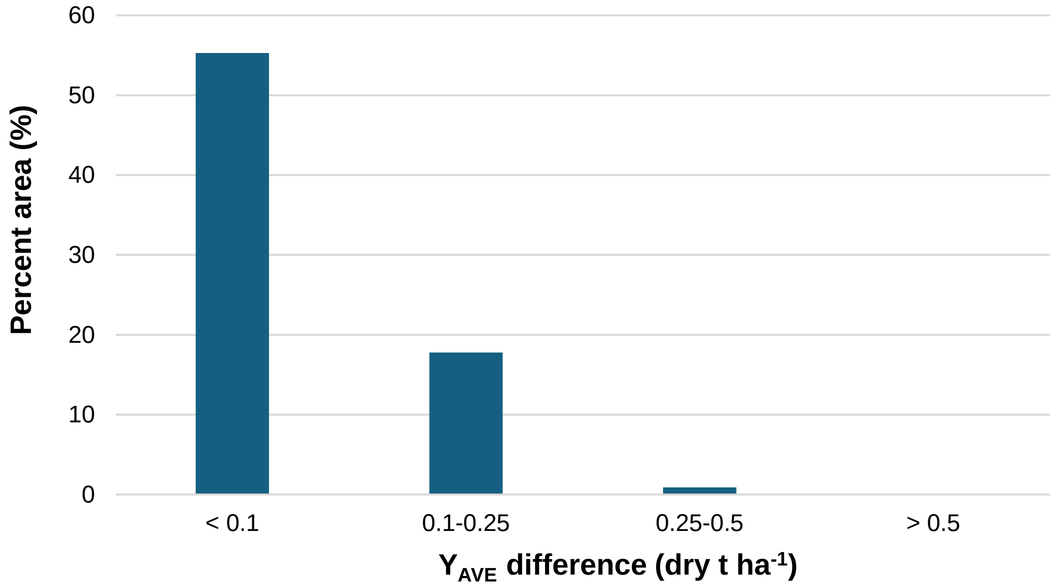


**Figure 10.4: Difference in $Y_{AVE}$ between the December and November planting dates for cowpea**

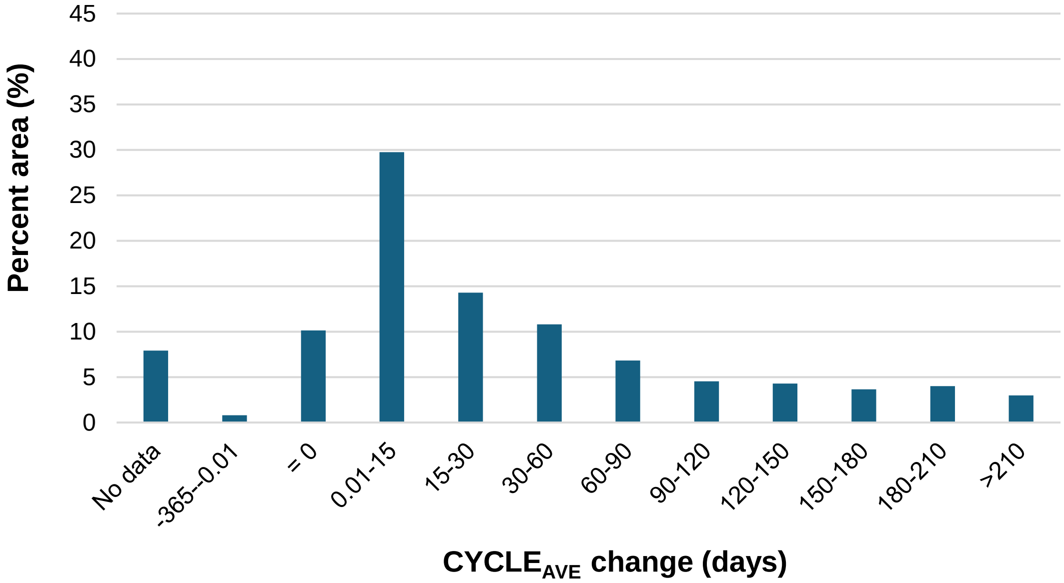


**Figure 10.5: Histogram of the change in CYCLE$_{AVE}$ for bambara nut planted in October due to frost occurrence**

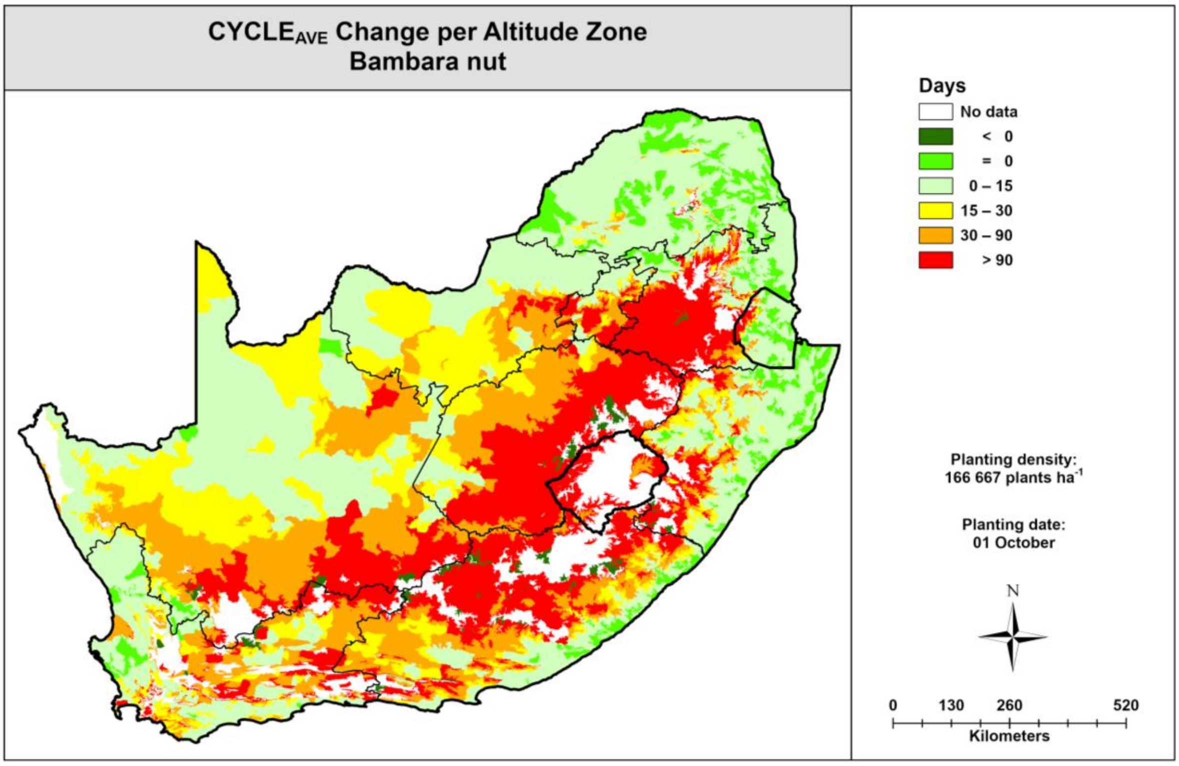


**Figure 10.6: Change in CYCLE$_{AVE}$ for bambara nut planted in October due to frost occurrence**

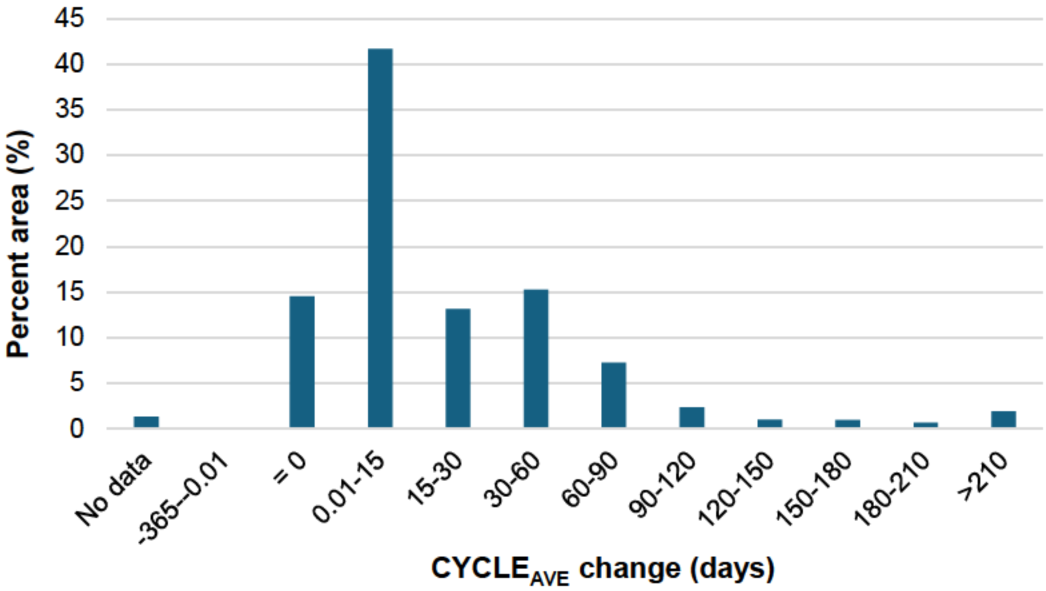


Figure 10.7: CYCLE$_{AVE}$ change for cowpea planted in November due to frost occurrence

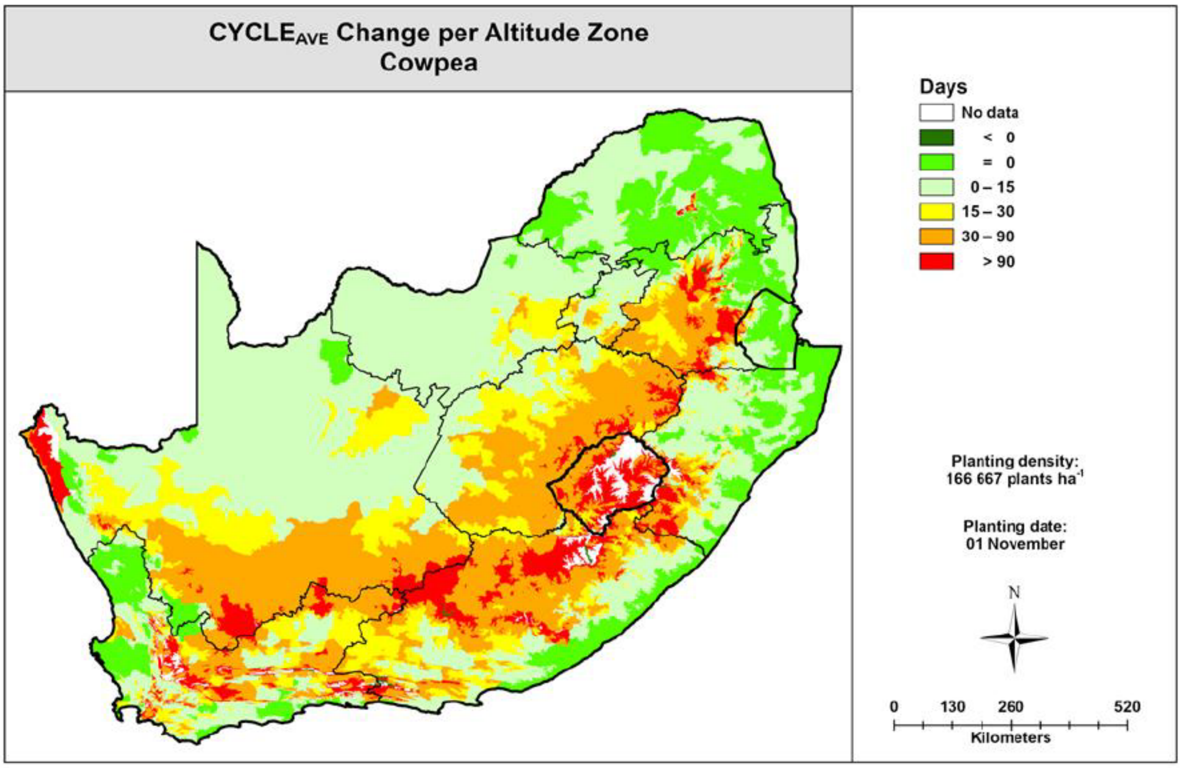


Figure 10.8: CYCLE$_{AVE}$ change for cowpea planted in November due to frost occurrence

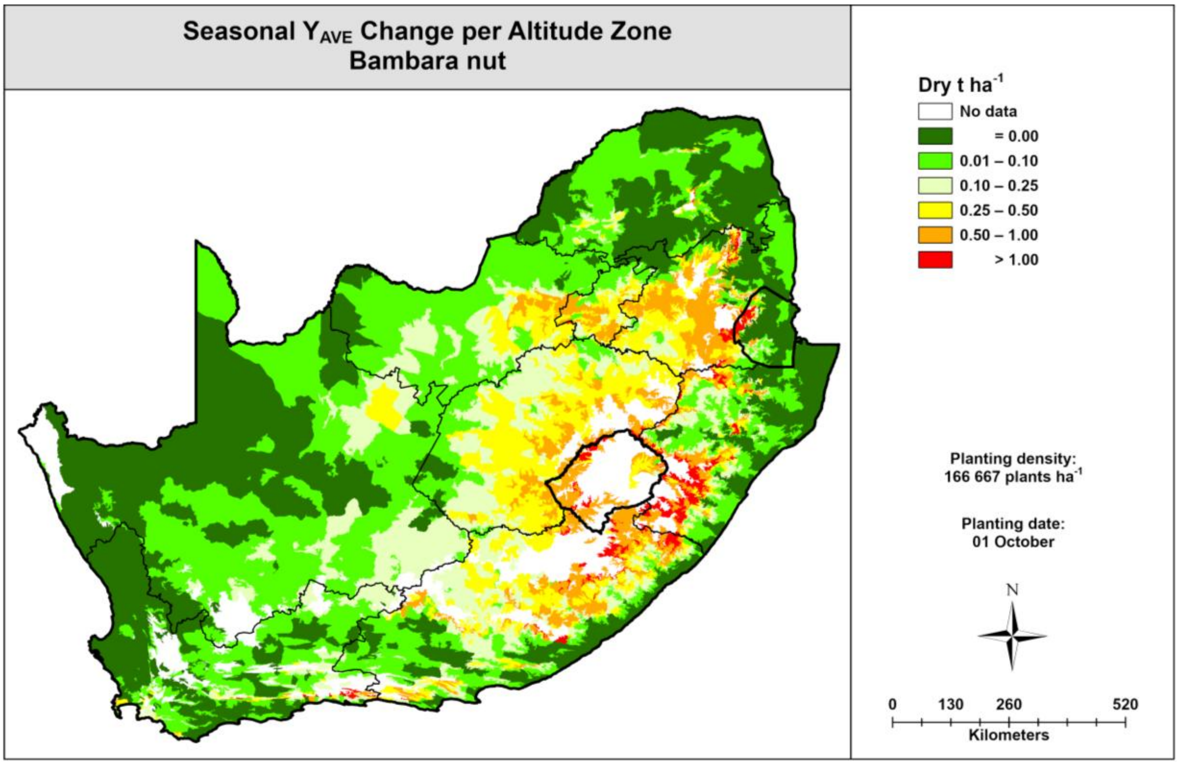


**Figure 10.9: Decrease in $Y_{AVE}$ for bambara nut planted in October due to frost occurrence**

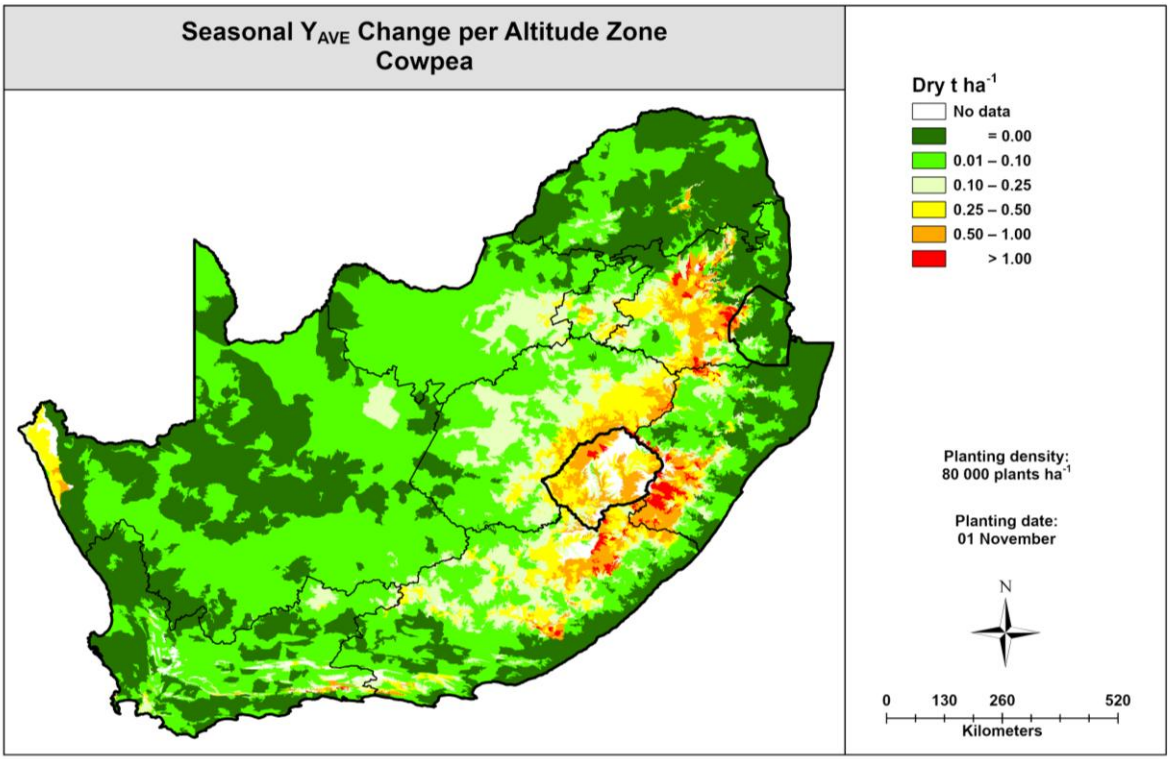


**Figure 10.10: Decreased in $Y_{AVE}$ for cowpea planted in November due to frost occurrence**

# 11. APPENDIX E

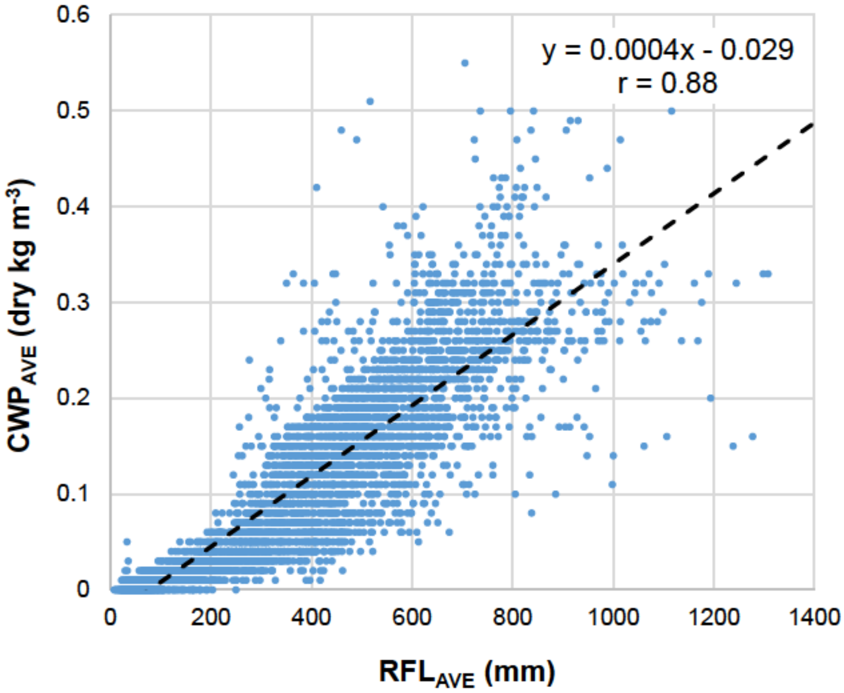


**Figure 11.1: Relationship between RFL$_{AVE}$ and CWP$_{AVE}$ for bambara nut planted in October**

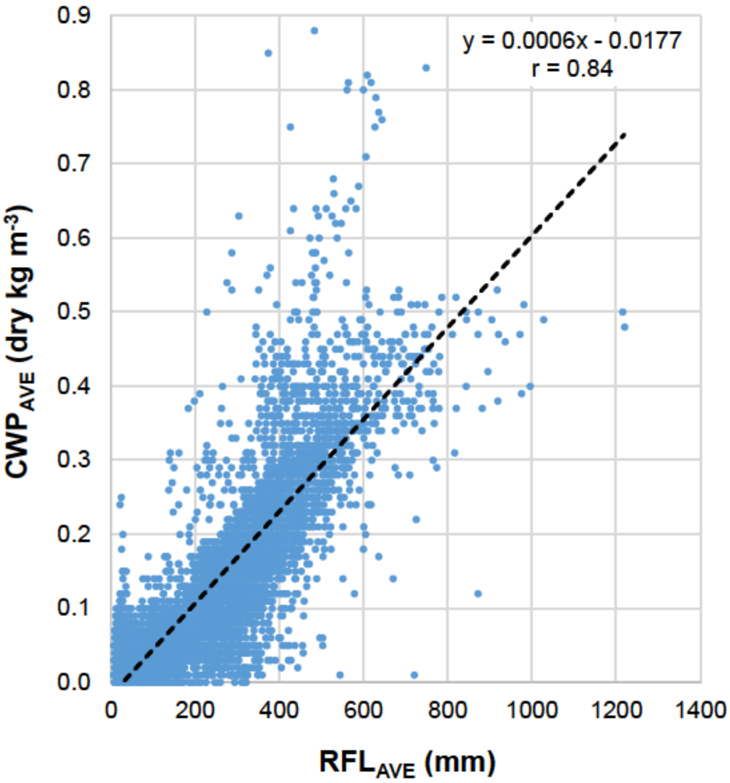


**Figure 11.2: Relationship between RFL$_{AVE}$ and CWP$_{AVE}$ for cowpea planted in November**

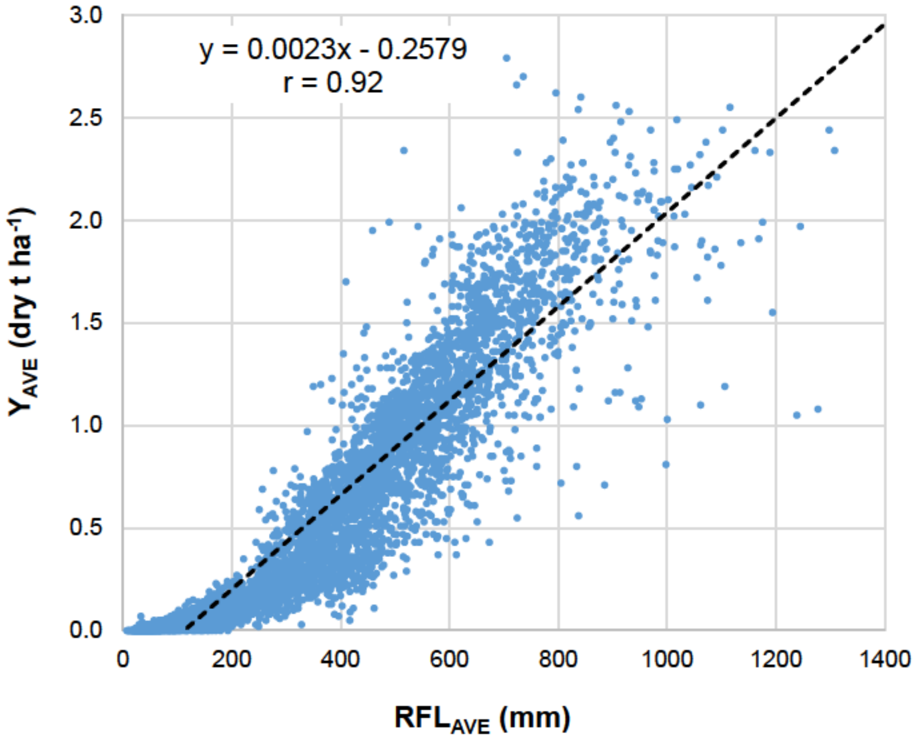


**Figure 11.3: Relationship between $RFL_{AVE}$ and $Y_{AVE}$ for bambara nut planted in October**

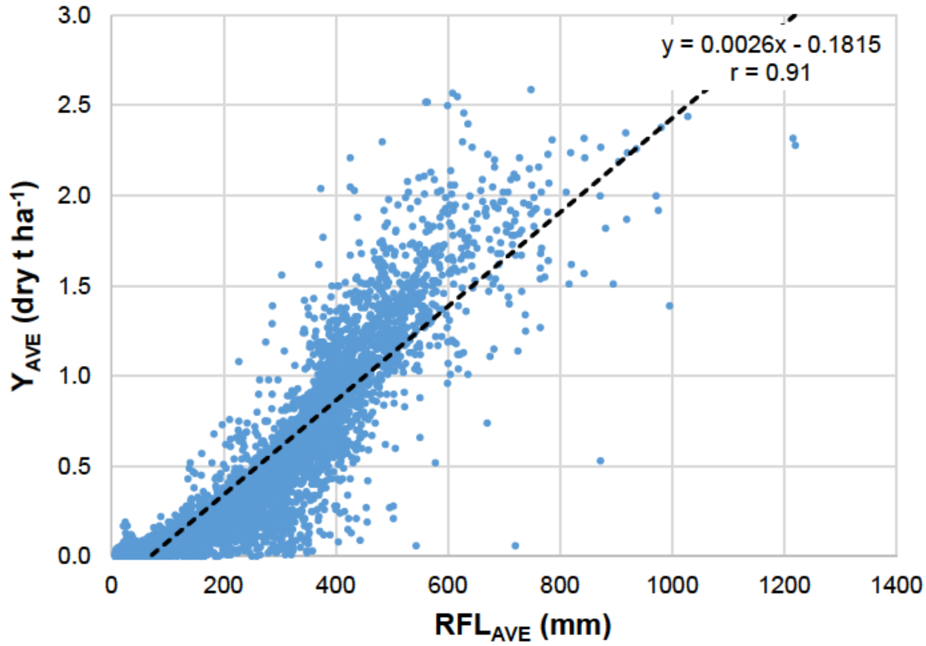


**Figure 11.4: Relationship $RFL_{AVE}$ and $Y_{AVE}$ for cowpea planted in November**

# 12. APPENDIX F

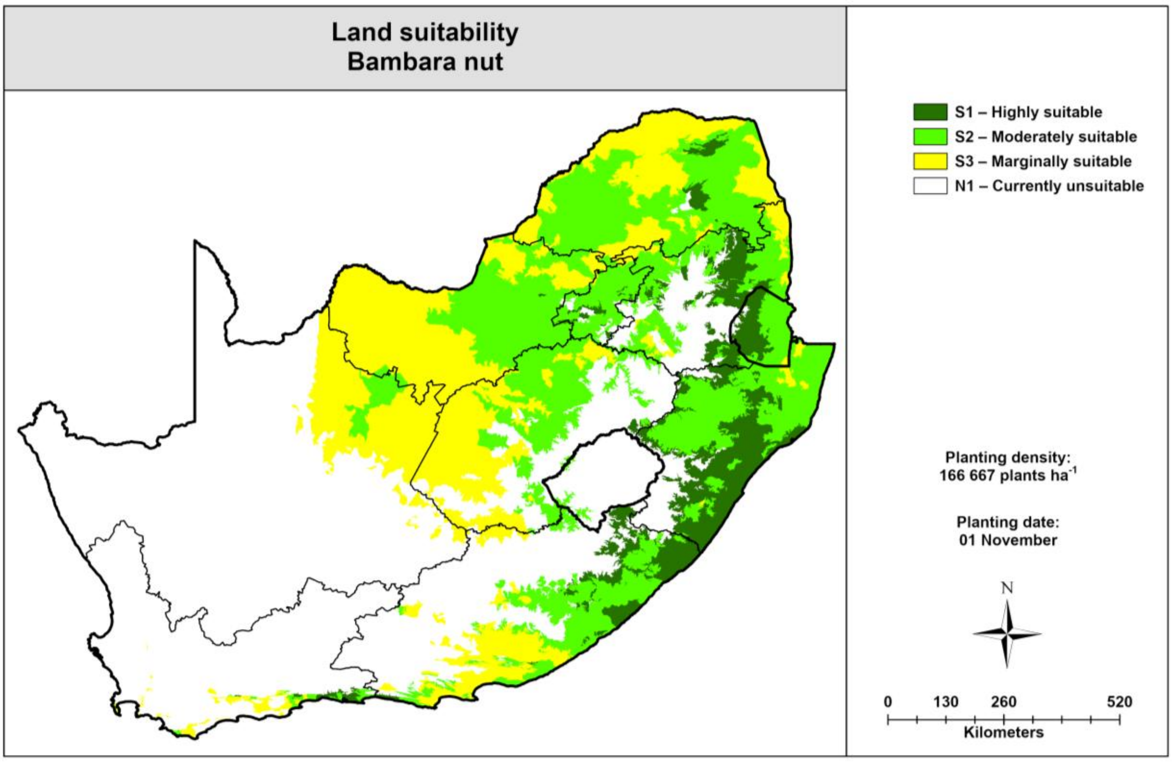


**Figure 12.1: Suitability classification map for rainfed production of bambara nut planted in November for southern Africa**

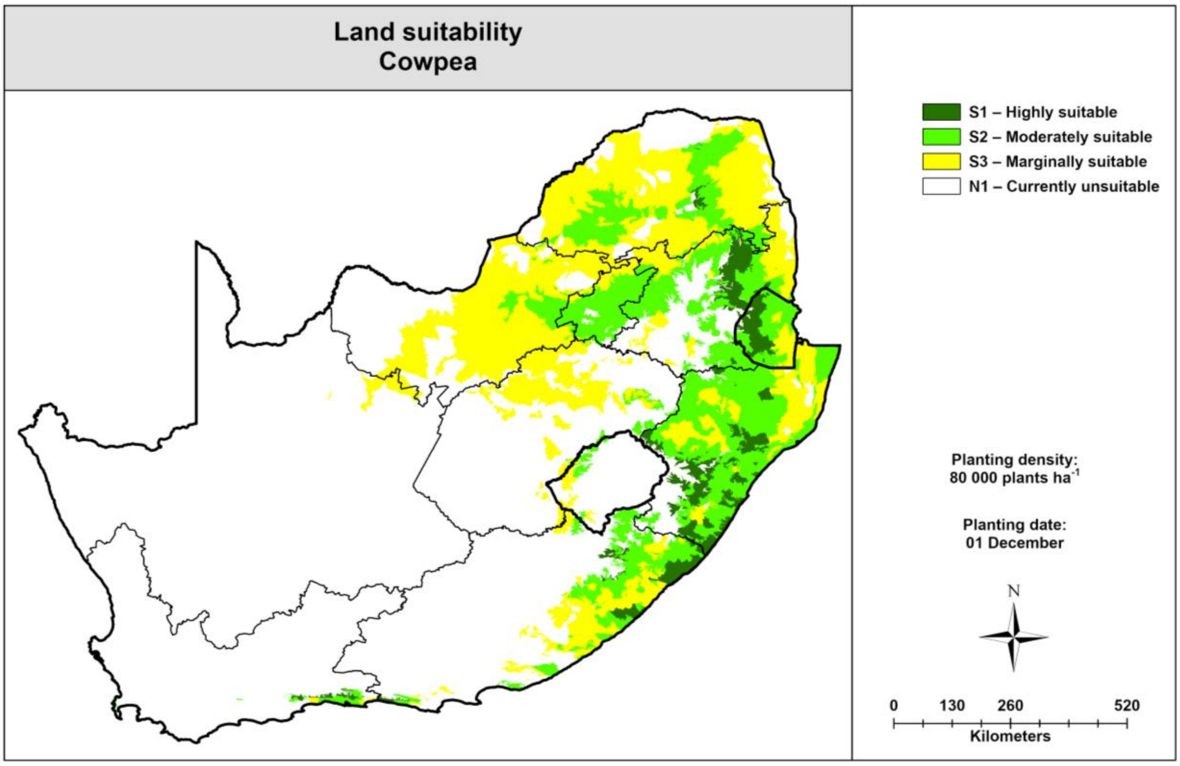


**Figure 12.2: Suitability classification map for rainfed production of cowpea planted in December for southern Africa**

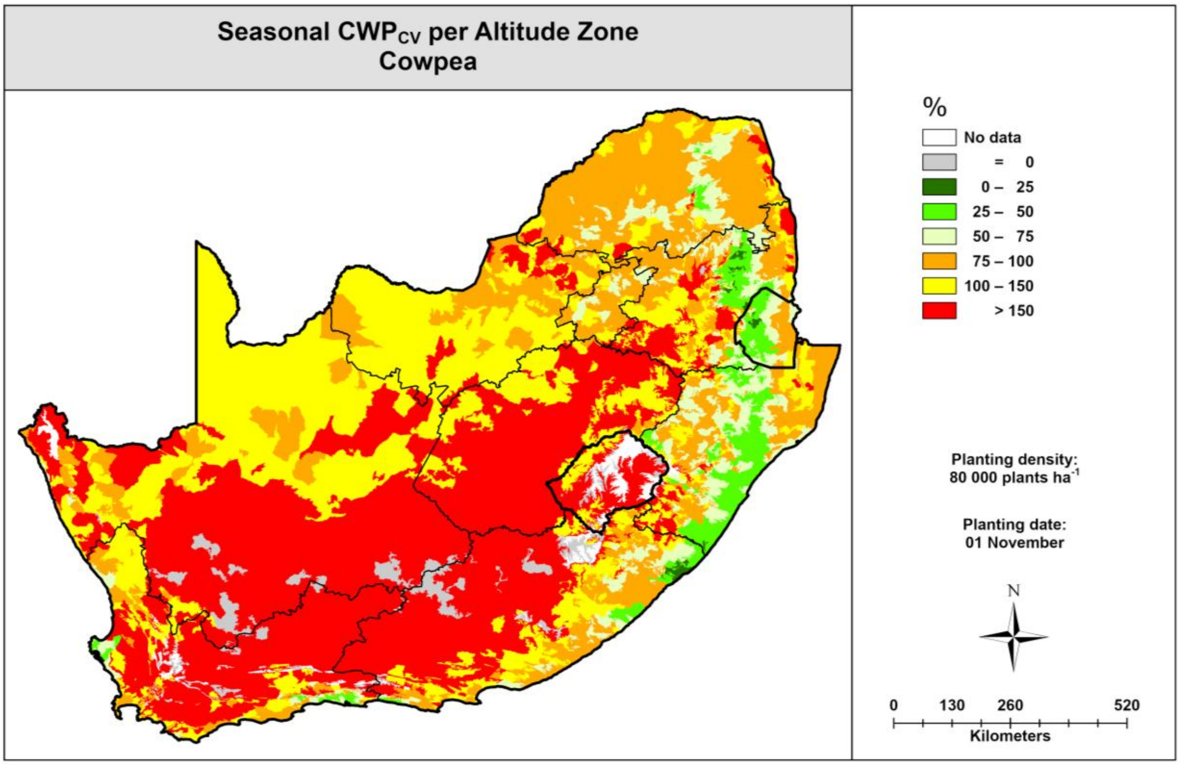


**Figure 12.3: Coefficient of variation in crop water productivity ($CWP_{CV}$) per Altitude Zone for cowpea planted in November**

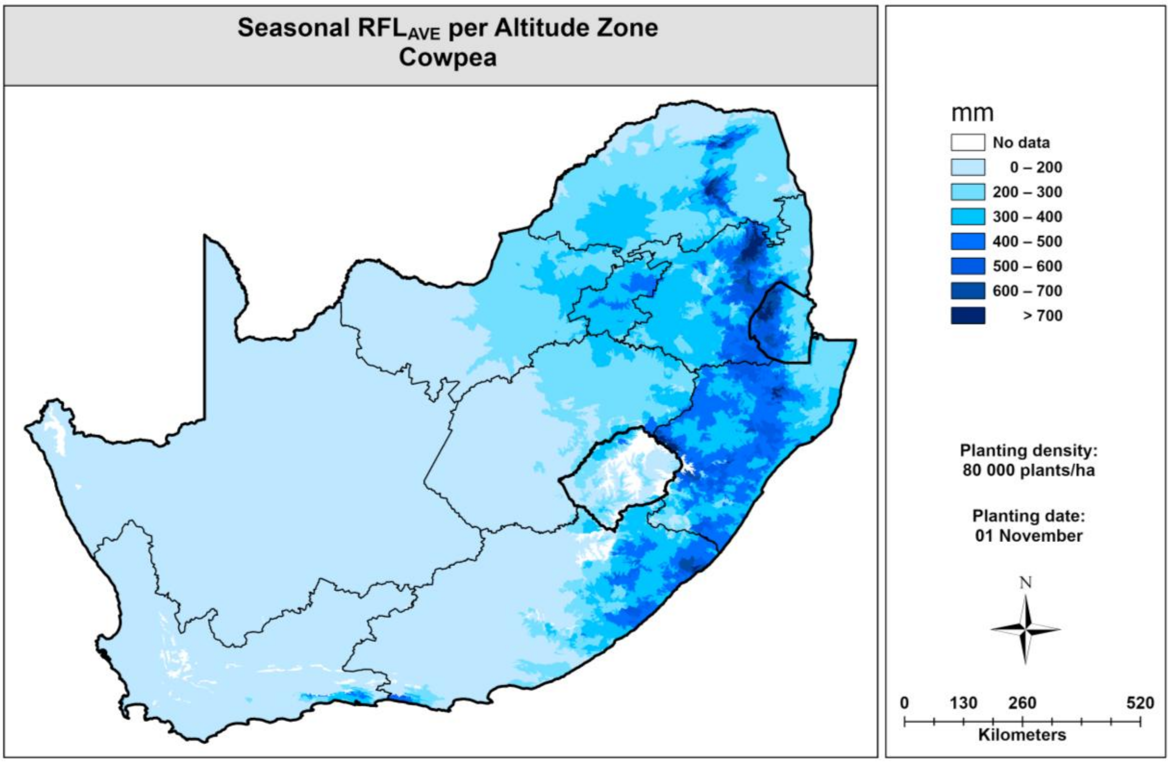


**Figure 12.4: Seasonally averaged rainfall ($RFL_{AVE}$) per Altitude Zone for cowpea planted in November**

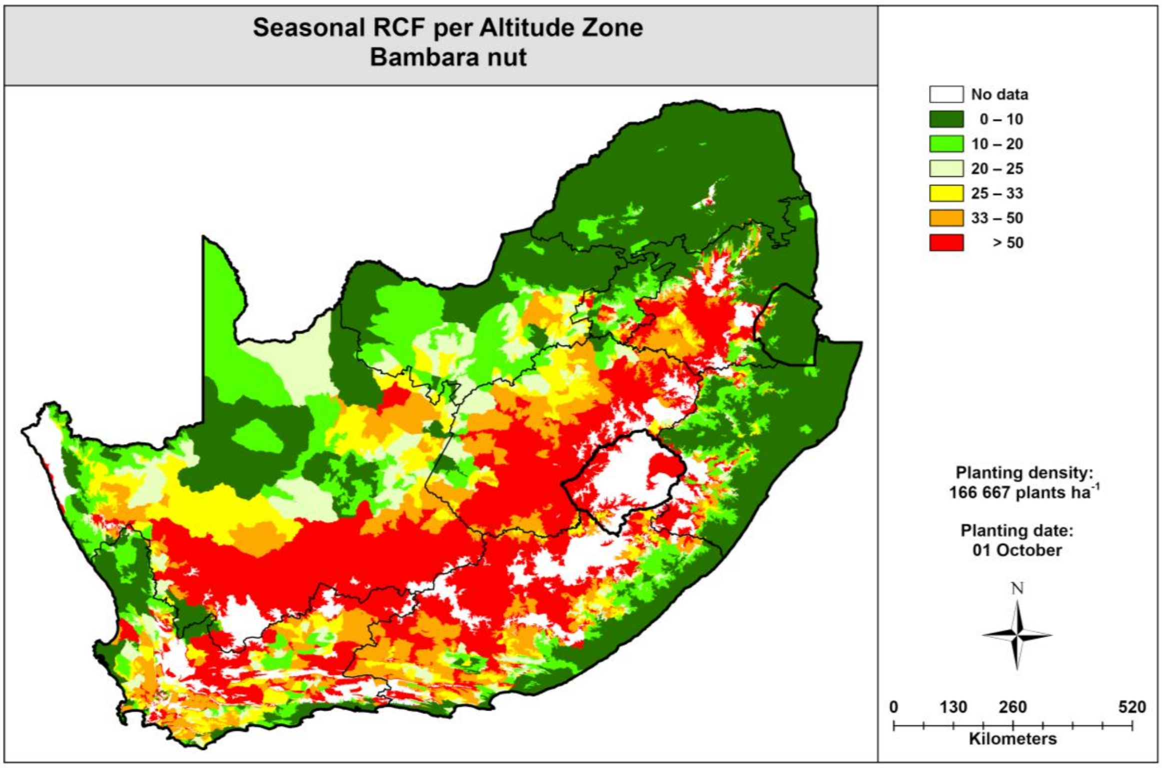


**Figure 12.5: Risk of crop failure (RCF) per Altitude Zone for bambara nut planted in October**