QUANTIFYING YIELD GAPS FOR RAIN-FED MAIZE (*ZEA MAYS*) IN SOUTH AFRICA: A BOTTOM-UP APPROACH

By

Brian George Mandigora

Thesis presented in partial fulfilment of the requirements for the degree of Master of Science in Sustainable Agriculture in the faculty of AgriSciences at Stellenbosch University

Department of Agronomy

Supervisor: Dr. Marcellous Le Roux

Co-supervisor: Dr. Sander de Vries

March 2018

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Brian G. Mandigora

March 2018

Abstract

Maize (*Zea mays*) is an important crop that supports livelihoods in South Africa (RSA) and most of the Sub-Saharan Africa (SSA) region. Currently, RSA is the biggest maize producer in the region, exporting to countries like Zimbabwe, Botswana and Malawi. However, the anticipated increase of the world population to 9 billion people by 2050 presents new challenges for resource supply and management, including food supply. The bulk of this population increase is expected to occur in the SSA region, and feeding this bigger population without compromising land and water for other needs is a priority. This thesis presents a yield gap study that seeks to ascertain the levels to which South African yellow and white maize yields can be increased on more or less the same spatial scale by applying the Global Yield Gap and Water Productivity Atlas (GYGA) protocol. The South African maize producing regions were subdivided into five key agro-climatic zones, referred to as designated climate zones (DCZs), which are explained by the GYGA Extrapolation Domain (GYGA-ED) zonation scheme. Within these DCZs, eight Reference Weather Stations (RWS) were selected to represent the South African maize production zones. Using Geographic Information Systems (GIS), buffer zones of 100km radius were delineated around the RWS, and were clipped according to CZs to avoid overlap. Climatic, edaphic and crop management data were collected for each of the RWS buffer zones. These data were used to simulate water-limited potential yields (Yw) for yellow and white maize for the three main soil types within the buffer zone under specific crop management practices from 2000 to 2014. Data for actual yields achieved by farmers during the same 15-year period (Ya) were sourced from the GrainSA database, and was corrected for moisture content. The Hybrid-Maize model was used to simulate Yw for yellow and white maize for the period from 2000 to 2014. The national yield gap (Yg) was calculated as the difference between the weighted values of Yw and Ya (Yg = Yw – Ya). The relative yield percentage (Y%) was also calculated as Y% = Ya/Yw x 100. This study found that for South Africa, the yield gaps between average farmers' yields and simulated water limited potential yields were 2.70 tha^{-1} for yellow maize and 3.14 tha^{-1} for white maize, which represent 52.99% and 57.88% of the white and yellow maize yields, respectively that are achievable as simulated by Hybrid-Maize. The exploitable yield gap (Ye), which is calculated as Ye = 0.8Yw – Ya, is the yield that can realistically be achieved by farmers when economic and bio-physical limitations are considered. If the Ye is to be met by South African maize farmers, this would increase annual production by about 1.5 million tonnes (yellow maize) and 3 million tonnes (white maize), which indicate a total maize yield increase of 42%, translating to over R9 billion in additional gross income for farmers. Therefore, there is considerable potential for RSA to increase maize production on existing farmland using current crop management practices, which could provide national and regional food security.

Opsomming

Mielies (*Zea mays*) is ‘n belangrike gewas wat die lewensbestaan van verskeie mense in Suid-Afrika (RSA) en ‘n groot deel mense van die Sub-Sahara Afrika (SSA) streek onderhou. Huidiglik, is RSA die grootste mielie produsent in die streek en voer uit na lande soos Zimbabwe, Botswana en Malawi. Hoewel, die verwagte verhoging in die wêreld bevolking tot 9 miljard mense by 2050 hou nuwe uitdagings in vir voorsiening en bestuur van voedsel. Die grootste deel van die bevolkingsgroei word in SSA verwag en om die groter bevolking te voed sonder om grond- en waterbronne ten koste van ander moontlike gebruike te belemmer, is ‘n prioriteit. Die tesis het onderneem om ‘n opbrengste gaping studie te doen om die vlakke waartoe Suid-Afrikaanse geel- en wit mielie opbrengste verhoog kan word op min of meer dieselfde ‘ruimtelike’ skaal volgens die Globale Opbrengste Gaping en Water Produktiwiteits Atlas (‘GYGA’) protokol kon vasstel. Die mielie-produserende streke van RSA is onderverdeel in vyf sleutel agro-klimatologiese sones, wat na verwys word na as die gereserveerde klimaatsones (‘DCZs), en onderstreep word deur die GYGA-Ekstrapolasie Domein (GYGA-ED) sonasie skema. Agt verwysings weerstasies (‘RWS’) was gekies binne hierdie DCZs wat verteenwoordigend is van RSA se mielie-produserende streke. Deur gebruik te maak van Geografiese Informasie Stelsels (GIS), was buffer sones binne ‘n 100 km radius binne afgebakende ‘RWS’ opgestel, wat dan voorts gesnoei is volgens die klimaatsones om oorvleueling uit te skakel. Klimaats-, grond- en gewas bestuur data is geneem vir elke ‘RWS’ buffer sone. Die data was gebruik om water-beperkte potensiële opbrengste (Yw) vir geel- en wit mielies vir die drie hoof grondtipes binne die buffersones onder sekere spesifieke gewas bestuurspraktyke vanaf 2000 to 2014 te simuleer. Data vir die werklike opbrengste (Ya) soos deur boere vermag vir dieselfde 15-jaar tydperk is verkry vanaf die GraanSA databasis. Die ‘Hybrid-Maize’ model was gebruik om die Yw te simuleer vir geel- en wit mielies vir die periode tussen 2000 en 2014. Die nasionale opbrengste gaping (Yg) was bereken as die verskil tussen die vasgestelde waardes van Yw en Ya (Yg = Yw – Ya). Die relatiewe persentasie (Y%) is bereken volgens Y% = Ya/Yw x 100. Die studie het bevind dat die opbrengste gaping, dus die tussen die gemiddelde werklike opbrengste deur boere en gesimuleerde water-beperkte potensiële opbrengste, 2.70 $t.ha^{-1}$ vir geel mielies en 3.14 $t.ha^{-1}$ vir wit mielies was in Suid-Afrika, wat ‘n 57.88% en ‘n 52.99% van die wit- en geel mielies opbrengste, respektiewelik verteenwoordig soos gesimuleer deur ‘Hybrid-Maize’. Die uitbytbare opbrengste gaping (Ye), wat bereken was as Ye = 0.8Yw – Ya, verwys na die opbrengs wat realisties bereik kan word deur boere, wanneer ekonomiese en bio-fisiese beperkings in ag geneem word. Indien die Ye bereik kan word deur Suid-Afrikaanse boere, sal dit jaarlikse produksie verhoog tot 1.5 miljoen ton (geel mielies) en 3 miljoen ton (wit mielies), wat ‘n totale mielie opbrengs verhoging van 42% beteken, en indien omskryf word, tot oor R9 miljard in addisionele inkomste vir boere sal behaal. Dus, daar is aansienlike potensiaal vir RSA om mielie produksie te verhoog op bestaande boerdery landgoed indien bestaande gewas bestuurspraktyke gebruik word, wat sal lei tot nasionale en plaaslike voedsel sekuriteit.

Acknowledgements

Many thanks and appreciation to:

- First and most importantly the Lord, who gave me the strength and ability to complete this study.
- My supervisors Dr. Marcellous Le Roux (Stellenbosch University) and Dr. Sander de Vries (Wageningen University, Netherlands) for their dedicated guidance, motivation and constructive criticism throughout the course of the project. I will always cherish the unique skills they imparted on me.
- The Sustainable Agriculture South Africa Team at Stellenbosch University, particularly the programme leader Professor K. Dzama for making this project a reality.
- The Plant Production Systems (PPS) team at Wageningen University, particularly Dr. Gerrie Van de Ven and Dr. Marloes Van Loon for their professional and technical input especially with regards to programming and modelling throughout the research.
- SADC Climate Services Centre and the Animal Science Department at Stellenbosch University for funding this research. Opinions expressed and conclusions arrived at are those of the author and are not necessarily to be attributed to them.
- The South African Weather Services for supplying the weather data used in this research.
- Department of Agriculture, Forestry and Fisheries crop specialists in Delmas, Balfour, Wepener, Bloemhof, Potchefstroom and Mafikeng for the crop management information.
- Wiltrud Du Rand from the Agricultural Research Council – Grain Crops Institute who gave me a solid foundation on crop modelling and the South African maize crop landscape.
- My colleagues Tawanda Marandure, Robert Andrews, Obvious Mapiye, Flackson Tshuma and Vanessa van Niekerk who have been supportive with their advice and constructive criticism for the duration of the programme.
- My family, especially Joyline and Sandra for their support and encouragement during the course of this degree.
- Finally, and very importantly, I would like to thank my wife, Nyasha. Her support, encouragement, quiet patience and unwavering love were undeniably a bedrock upon which I leaned during the course of this thesis. I would also like to thank my son, Mukudzei, who has continually provided the requisite breaks from academia and the motivation to finish my degree with expediency.

Table of Contents

List of Abbreviations and Acronyms

AfSIS	Africa Soil Information Systems
CO2	Carbon dioxide
CZ	Climate Zone
DAFF	Department of Agriculture, Forestry and Fisheries
DCZ	Designated Climate Zone
ETO	Reference crop evapotranspiration
G x E x M	Genotype x Environment x Management
GDD	Growing Degree Days
GHG	Greenhouse Gas
GIS	Geographic Information Systems
GYGA	Global Yield Gap Atlas
H2O	Water
NASA	National Aeronautics and Space Administration
NASA-POWER	National Aeronautics and Space Administration -Prediction of Worldwide Energy Resource
PASW	Plant Available Soil Water
RSA	Republic of South Africa
RWS	Rainfall Weather Station
SAWS	South Africa Weather Service
SPAM	Spatial Production Allocation Model
SSA	Sub-Saharan Africa
Tmax	Maximum Temperature
Tmin	Minimum Temperature
USA	United States of America
Ya	Actual Yield
Ye	Exploitable Yield
Yg	Yield gap
Yp	Simulated potential yield
Yw	Simulated water-limited potential yield

List of Figures

.

List of Tables

CHAPTER 1

GENERAL INTRODUCTION

1.1 Background/Overview

The world's population is increasing at an estimated 1.18 percent annually, which translates to an additional 83 million people each year (United Nations Department of Economic and Social Affairs, 2015). This population increase presents an unprecedented global food security challenge. Researchers generally agree that in order to feed an anticipated 9 billion people by 2050, an estimated 60% increase over current primary agricultural productivity is required (Food and Agriculture Organization of the United Nations, 2015). However, demand for resources and services from ecosystems have also increased, and continues to do so, with land becoming more critical as a resource that sustains life. In this light, in most countries, but particularly in developing countries, there exists a conflict between nature conservation and agricultural area expansion. Over and above the anticipated rise in global population, climate change, diet changes and consumption pattern shifts will exacerbate global food security concerns. All these factors act against a background of crop yield plateaus and reductions in the rates of crop yield increase that have been experienced in some parts of the world in spite of technological advancement (Gobbett *et al.*, 2017). Combined with the fact that most of the prime quality agricultural land is already in use, the potential to produce more food by increasing the area under cultivation is limited.

One of the pathways towards satisfying the expected increase in demand for food crops identified by Keating *et al.* (2014) is closing yield gaps in agricultural production systems. This means that increased agricultural production will have to come from increased yields on more or less the existing cropland areas. However, the ability of some regions to support yield increments is higher than others due to more favourable climatic conditions, better soils and availability of irrigation; or because current yields are still far below their potential. This is particularly true of most developing countries in Sub-Saharan Africa (SSA) where large gaps exist between current yields and potential yields (Van Ittersum *et al.*, 2013); the so-called Yield Gap (Yg), which must be bridged to avoid potential food crises in the coming decades. Furthermore, understanding the underlying causes of Yg may assist in identifying strategies for narrowing the yield gaps as well as to make projections of future crop yields in different regions. Therefore, knowledge of factors that contribute to the Yg is essential for planning efforts to increase crop production.

Yield gap analyses provide a means of evaluating differences between actual farm yields (Ya) and potential crop yields under irrigation (Yp) or under rain-fed conditions (Yw) - the so-called water-limited potential yield (Van Wart, et al., 2013). For clarity, the Yg is defined as the difference between potential yield, denoted by "Yp" for irrigated systems and "Yw" for rain-fed systems, and average actual farm yield (Ya) (Grassini *et al.*, 2015), at certain pre-defined spatial and temporal scales. In Yg analyses, Ya can be defined as the average yield obtained by farmers in a certain area over a period of time. Relevant yield statistics can be obtained at various spatial levels, which include but are not limited to municipalities, districts and provinces (Van Bussel et al. 2015).

On the other hand, potential yield (or yield potential) - (Yp) - is the yield that is obtainable from a crop cultivar grown under ideal biotic conditions, and with sufficient water and nutrient supply (Evans & Fischer, 1999; Van-Bussel *et al.*, 2015). Yp and Yw can be determined in 3 main ways. They can be derived from maximum farmer yields, field experiments or crop modelling. Crop modelling has been the most widely used method for studies of Yg at large spatial scales (Van Ittersum *et al.*, 2013). However, the methods used for the collection of model input data for estimation of Yp and Yw are sensitive to errors and biases and often incoherent, thus the Global Yield Gap Atlas (GYGA) protocol for Yg estimation was developed to address these shortcomings.

The GYGA protocol was developed in such a way that it can be applied to cropping systems within both data-rich and data-poor contexts, therefore results can be compared amongst countries or regions. GYGA allows for upscaling and extrapolating yield gap results via a so-called technology extrapolation domain. Compared to other methods, the GYGA protocol has the advantage of using local agronomic knowledge and data as the starting point. Furthermore, consistently applying one protocol makes yields and yield gaps comparable, thereby addressing the challenge of comparing results from different studies. Methods employed in other studies may often be too coarse, lacking agronomic relevance (global studies) or too location-specific (Van Ittersum *et al.*, 2013), thus making it hard to validate estimates, or improve on them. Various Yg studies have been carried out around the world, with the focus on "key crops"; maize, wheat, rice, potatoes and soya-beans.

Very few studies to assess Yg of major crops have been conducted in the Republic of South Africa (RSA), and even fewer of them employed a bottom-up spatial approach. To date, only the yield gap for sugar cane has been quantified using the GYGA protocol (Global Yield Gap Atlas, 2015). However, no current study to quantify the yield gap for maize at a national scale, employing a reproducible method for RSA has been found in literature at the time of writing of this thesis, but yields obtained by farmers, especially under rain fed production, are almost certainly lower than those attainable using locally optimised agricultural best practices. Verdoodt *et al.*, (2003) conducted Yg studies in the Eastern Cape of RSA for maize and sunflower. However, the methods they used were coarse, and the studies would be difficult to validate and reproduce. Model input parameters were simplified, thus making it difficult to validate results and compare them with the results from other regions of the world.

This study therefore sought to quantify the Yg in rain-fed yellow and white maize in RSA using a method that is reliable and reproducible. The study will therefore provide a basis for planning and upscaling efforts to increase maize production.

1.2 Aim and objectives

The aim of this project was to quantify the yield gap for rain-fed yellow and white maize in South Africa using the GYGA protocol, thus providing a quantitative basis for further research on sustainable maize production.

1.2.1 Specific Objectives

1. To identify the study areas, which are defined by the rainfall weather station (RWS) buffer zones.
2. To quantify average farmers' yields (Ya) of rain-fed yellow and white maize at provincial and national scales in South Africa.
3. To determine potential yields of rain-fed yellow and white maize at RWS buffer zone level and upscale to South African provincial and national scales.
4. To quantify the water-limited yield gaps (Yg) for rain-fed yellow and white maize in South Africa at provincial and national scales.

1.3 Problem Statement

In the South African context, maize is the main staple crop consumed. White maize is the major source of calories for South Africa's (RSA) mostly black working class, and yellow maize is mainly used in the manufacture of animal feed. However, demand of white maize is not increasing as rapidly as that for yellow maize due to changing dietary preferences. Hence, RSA has become a net maize exporting country (Sihlobo & Kapuya, 2014). In 2014, South Africa exported 2.1 million tonnes of maize. Taiwan and Zimbabwe accounted for 26% and 14% of maize exports respectively; while Japan, Botswana and South Korea each accounted for 7% of total maize exports. Other countries to which maize was exported include Namibia, Mozambique, Lesotho, Swaziland and Italy; earning the country R6,5 billion in export revenue (Sihlobo & Kapuya, 2014).

This demand for South African maize is expected to rise further over the coming decades as the global population continues to increase, dietary preferences change and lifestyles continue to improve, particularly in SSA. In RSA, the RSA population is expected to reach 82 million by 2035, up from 49 million in 2009, with the middle class increasing by 30% between 2001 and 2004 (Goldblatt, 2011). Consumption of poultry and eggs has increased almost four-fold over the same period, which results in higher demand for yellow maize. Goldblatt (2011) further puts perspective into this by asserting that about 50% of RSA's maize is used for animal feed, of which 70% is used for poultry and egg production, both for local consumption and export. Therefore, there is need to ensure that not only the local, but indeed the foreign maize markets for South African maize are well supplied, albeit from more or less the same land size. One way of doing this is through bridging yield gaps (Koning & Van Ittersum, 2009; Rosegrant *et al.*, 2001). However, the question arises of how best to sustainably increase production, and two main pathways have been suggested; extensification and intensification. The intensification option seems plausible, considering its perceived lesser negative effects on biodiversity conservation among other advantages.

Although there has been some research conducted on maize yield gaps in RSA, none have employed a methodology that is reproducible, transparent and verifiable. Previous studies, most notably, Verdoodt *et al.*, (2003), that utilised crop growth modelling made use of coarse data. This data only included monthly instead of daily weather data, thus making the process susceptible to error, and difficult to reproduce and upscale. Yield gaps for maize at a national scale are not known, and this study will shed light on whether yield ceilings for maize in RSA have been reached, or whether there's still potential to sustainably increase maize production in RSA.

1.4 Research Question

The main research question of this project was: Is there potential to increase maize productivity on the already existing agricultural land used for rain-fed maize production in South Africa?

1.5 Justification

In light of the growing world population, anticipated higher quality of life and limited land resources, it is essential to quantify the productivity potential of the land in order to be able to feed this anticipated higher and more affluent population in the coming 3 decades. This productivity potential can be assessed using Yg analysis. However, a protocol that can be globally applied and verified needs to be used in order to compare results and propose interventions that can address the problem of feeding the world at a global scale.

Maize is an important staple food crop in South Africa and most parts of Sub-Saharan Africa (SSA) (Department of Agriculture Forestry and Fisheries, 2012, 2014). South Africa is a major exporter of maize in the SSA region, exporting to countries like Botswana, Zimbabwe, Zambia, Malawi and Mozambique among others. South Africa also supplies international markets like Japan with maize (Scheltema *et al.*, 2015). It is therefore important to ascertain the levels to which production can be increased on more or less the existing farmland in order to ensure food security. Added to this, a robust and spatially explicit methodology that ensures easy comparison of yield gaps is essential in order to inform researchers, economic models and policy makers on anticipated supply and demand dynamics for maize, which will help to optimise land use and to address the challenge of food distribution and natural resources management (Van Ittersum *et al.*, 2013).

Recurring droughts and low farm profits have led to a decline in the land area under agriculture in RSA. However, farmers have managed to maintain relatively constant levels of production over the last 20 years due to implementation of intensive production methods (Goldblatt, 2011). This yield gap analysis study will therefore help to ascertain whether South African maize yields are close to reaching or have reached a plateau, thus ascertaining whether there is scope to further increase land productivity. To achieve this, a methodology that facilitates such assessment at a large spatial scale while not compromising the robustness, reliability and quality of the results was needed.

The GYGA protocol was therefore developed to address some shortcomings of traditional methods of yield gap analysis, which make use of coarse data, for example gridded weather, soil, and crop data. This allows larger spatial coverage but might give unsubstantiated output. Hence, GYGA is based on methods that are transparent, reproducible and based on best available science (Van Ittersum, *et al.*, 2013; Van Wart *et al.*, 2015).

The assessment of maize yield gaps for South Africa using the GYGA protocol will make a significant contribution towards informing policy makers, as well other stakeholders in the agricultural sector, of current maize crop suitability across the country. The identification of maize yield gaps allows the formulation of appropriate strategies that will ensure efficient and sustainable agricultural production, and improve food security as well as livelihoods. A quantitative basis for identifying constraints to production is therefore established, and management options could be aligned accordingly.

Moreover, the results of Yg using this protocol are accessible to the public, and thus they can be compared with those from other regions/countries. In this way, the GYGA can be used to identify regions with greatest potential for agricultural development, technology sharing, monitoring, evaluation and investment. The suitability of crop intensification for food self-sufficiency can be assessed for each country, and further options can be considered if intensification proves to be unsuitable. GYGA provides a quantitative basis for further studies which aim to explain and may mitigate yield gaps, and associated phenomena like climate change in SA. The availability of yield gap data on a public platform allows improvement of Yg assessments as better data become available (Van Ittersum *et al.*, 2013).

1.6 Scope of project

This research falls under a larger project that seeks to compile yield gaps of major crops cultivated globally in an online open-source atlas. This so-called Global Yield Gap Atlas, is compiled using standard protocols and methodologies. It is unique in that data is collected following a hierarchy, in which it is iteratively selected and/or discarded depending on reliability and availability. The project entails the collaboration of locally based agronomists in various parts of the world, who have in-depth knowledge of crop management, biotic and agronomic practices in their countries. (Global Yield Gap Atlas, 2015).

1.7 Outline of the study

The remainder of this study is organised as follows: Chapter 2 focuses on the literature review, where the concepts underlying maize production and the methods for calculating Yg are discussed. Chapter 3 will discuss and explain the methods used in this study in detail, whereby focus will be on the site selection process, Ya collection, Yw simulation process and Yg calculation and upscaling. Added to this, the findings of the yield gap study will be presented in Chapter 4. Lastly, Chapter 5 articulates the yield gap findings, and the constraints that farmers face, whilst concluding with a summary and recommendations.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter reviews and contextualises the concepts employed in this study. The main concepts and issues reviewed include population dynamics and how they influence global food demand (in particular maize demand), South African maize production trends and statistics, crop growth modelling, the concepts and methods of yield gap analysis and finally the Global Yield Gap Atlas (GYGA) protocol.

2.2 Pressure on crop production

The world's population is currently increasing at an estimated rate of 1.18 percent annually, which translates to an additional 83 million people each year. Thus, a total population increase of over 1 billion people is expected over the next 15 years, reaching 8.5 billion in 2030, increasing further to 9.7 billion in 2050 and 11.2 billion by 2100. Moreover, results of recent studies seem to suggest that more than half of the global population growth that is anticipated up to the year 2050 will occur in Africa. (United Nations Department of Economic and Social Affairs, 2015).

In addition, high economic growth rates, which lead to an increase in income in developing and most populous countries, are expected to accompany this population growth. Furthermore, the demand for energy, grain and livestock products has not only increased in developing countries, but also worldwide (De Vries *et al*., 1997; Koning & Van Ittersum 2009; Godfray *et al.*, 2012). These factors, coupled with globally declining rates of crop yield increase have necessitated that ways of producing more food in order to feed future generations be found. Strategies to increase food crop production without compromising biodiversity or hindering urban development will be needed. Studies suggest that if global food demand is to be met by 2050, crop production needs to at least double from its 2005 levels (Tilman *et al.*, 2011, Godfray *et al.*, 2012).

Generally, two pathways can be followed to increase crop production; improving crop yields on existing farmland or expanding crop production areas. Expanding crop production area has been said to be economically, socially and environmentally unviable because of competition from human activities (De-Vries *et al.*, 1997), as well as from protection of biodiversity and public goods provided by natural ecosystems (for example storage of carbon in rainforests). The latter are now receiving high priority from policy makers (Tilman *et al.*, 2011).

Furthermore, land is becoming a scarce commodity worldwide. Urbanization has taken over previously productive agricultural land over the past few decades (United Nations Department of Economic and Social Affairs, 2015). South Africa, where only 12% of the country is suitable for rain-fed crops is no exception. Added to this, only 3% of the land is considered fertile; which falls short when compared to other countries like India, where 53% of the land is fertile (Goldblatt, 2011). Coupled with the decline in farming profitability and recurrent droughts which have resulted in the number of farms in RSA declining by about two-thirds between 1990 and 2012 (Goldblatt, 2011), there is indeed need to ensure that productivity is increased to ensure food security. Technological advancements and "green consciousness", which have given rise to an increase in production of bio-fuels, have also added to the pressures on land. Productive land in particular has also become limited because of unsustainable land management practices. These practices have resulted in phenomena like erosion, salinization and desertification amongst others, which render otherwise productive land to be barren and unproductive (Van Ittersum *et al.*, 2013; Tilman *et al.*, 2011). Land clearing and habitat fragmentation, which threaten biodiversity, are some of the environmental impacts of expanding the area under agriculture. Tilman *et al.*, 2011 argue that about 25% of global greenhouse gas (GHG) emissions result from land clearing, crop production, and fertilization. All the above make the area expansion option debatable, and raises the question whether intensification would be the more "sustainable" route to follow. According to Schoeman *et al.*, (2010), the cultivated area in RSA has decreased from 12.4% to 11.9% between 1994 and 2005. On the other hand, the area under urban areas, forestry and mining have increased. All this makes the extensification option debatable, thus raising questions whether extensification would be sustainable route to follow.

In light of the prevailing competition for land from different uses, (sustainable) intensification of crop production is a pathway that could be followed in order to sustainably address rising crop demand as well as supply-side stressors. Ollenburger *et al.*, 2016 define agricultural intensification as “a process that results in increased output per unit of land as a consequence of intensive use of inputs and labour (per unit of land)”. The intensification option, whereby more food will be produced on more or less the same amount of land, is thus the most likely option to pursue, especially in Africa where significant population growth is expected in the coming decades (Nkamleu, 2011). Intensification generally involves innovation, especially through mechanisation and inputs that respond to rising population and land pressure.

Despite intensification seeming the “greener” route to follow in order to increase food production, it poses some environmental problems, just like extensification. However, some studies have shown that these are generally fewer than when new tracts of land are cleared for cropping. Research by Burney *et al.*, (2010) on GHG emissions resulting from the Green Revolution computed substantial extra greenhouse gas emissions. However, these were much less than would have occurred if production had been increased solely by clearing more land. Tilman *et al.*, (2001), calculated GHG emissions from converted land (area expansion) and from land on which yields were increased through efficient fertiliser application. Their results showed that land conversion was the poorer option in terms of greenhouse gas emission. These aforementioned studies both highlight (sustainable) intensification as being a more environmentally friendly way to increase crop production than extensification.

Identification of regions where sustainable intensification of food crop production can be implemented is critical. In sub-Saharan Africa (SSA), there are specific food crops that influence and affect the livelihoods of people more than other crops. It is therefore important to assess how SSA can meet the demand for staple food crops considering the low cereal self-sufficiency ratio (the ratio between domestic production and total consumption (or demand)), while the highest increase projections have been made for this region. While self-sufficiency alone is not an indicator of food security, in low-income countries of SSA it is important as it determines foreign exchange that is used on food imports.

A study carried out by Van-Ittersum *et al.*, 2016 concluded that the cereal yields in most SSA countries are growing more slowly than population and demand, and that cereal demand in

2050 is expected to be 335% higher than the 2010 level. Population growth alone makes up about 75% of this expected increase in demand. Maize is one such crop, and it is grown throughout most of the region. It is therefore crucial to be able to assess whether and where yields of maize can be increased sustainably. This can be achieved by carrying yield gap analyses in the areas where crops are produced.

2.3 Maize production overview

Maize (*Zea mays*) is a cereal crop that is grown widely under different agro-ecological conditions. Almost 50 cultivars with different colours, textures and grain shape and size exist, with white, yellow and red maize being the most common types (International Institute of Tropical Agriculture, 2009). In the developed countries of the world, maize is widely used as animal feed and/or as an industrial raw material, and thus consumed mainly as second-cycle produce in the form of meat, eggs and dairy produce. In the developing countries of SSA, maize is mainly used for human consumption, where it serves as a staple diet. Worldwide, more than 116 million tonnes of maize are consumed annually, with Africa consuming 30% and Sub-Saharan Africa 21% (International Institute of Tropical Agriculture, 2009). In Eastern and Southern Africa, 85% of the maize produced is consumed, unlike in other world regions where most of the maize is used as a primary product or raw material for other products like animal feed (International Institute of Tropical Agriculture, 2009).

Global maize demand by 2020 is expected to be 50% higher than the 1995 level of 558 million tonnes (Pingali, 2001). In the developing countries, maize demand is expected to rise from 282 million tons in 1995 to 504 million tons in 2020 (Rosegrant *et al.*, 2001). Maize is not native to Africa, but its nutritional qualities and ease of cultivation have made it an important dietary crop.

2.3.1 Maize production and consumption in South Africa

2.3.1.1 Production

There are two main maize production regions in South Africa. The Dryland Eastern Production Region (Figure 2.1) includes the "West Rand, Eastern Highveld, KwaZulu-Natal and the cooler parts of the Eastern Cape" (Pannar, 2017).

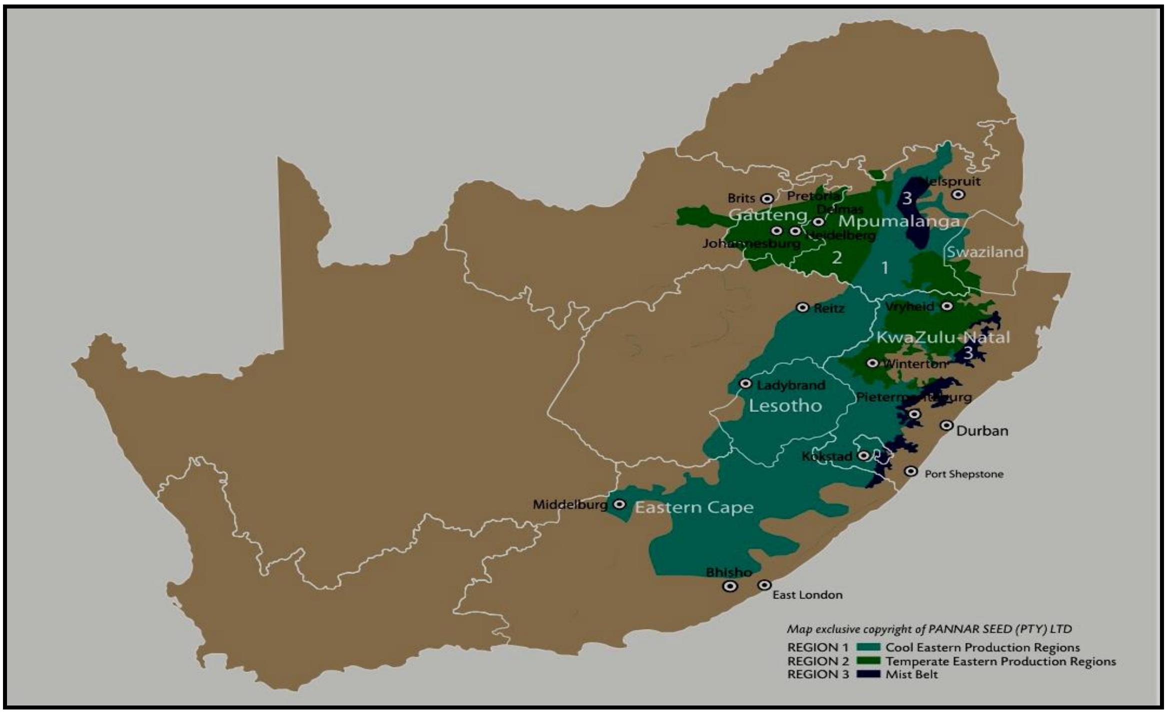


Figure 2.1: The South African Dryland Eastern Production Region *(Source: Pannar, 2017).*

The Dryland Western Production Region (Figure 2.2), which covers areas that include the North West, the Western Free State and Limpopo Provinces. The North West, Mpumalanga and Free State are the major production provinces, accounting for close to 73% of total national production (Department of Agriculture, Forestry and Fisheries, 2016).

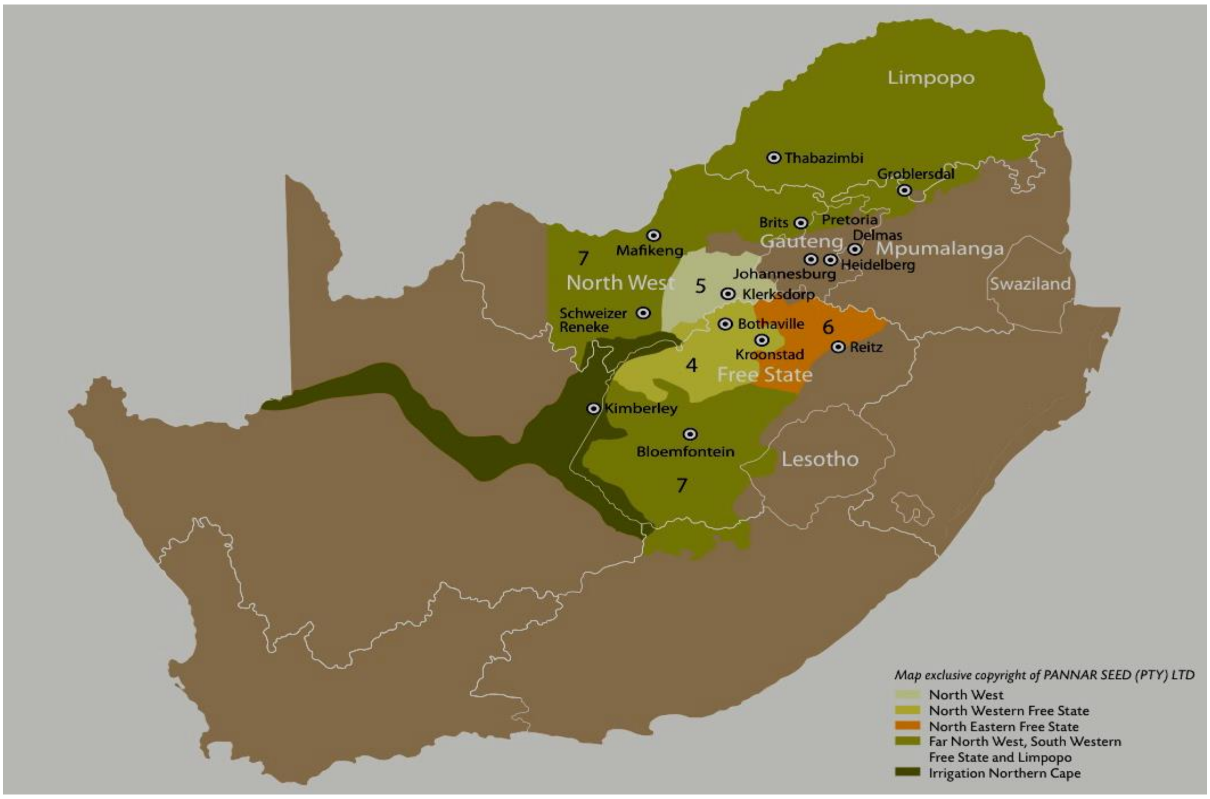


Figure 2.2: The South African Dryland Western Production Region *(Source: Pannar, 2017).*

Over the past decade, South Africa has produced an average of 10.4 million tonnes (Mt) of white and yellow maize annually on about 2.8 million hectares (ha) of land, thus an average commercial yield of 4.5 tonnes per hectare (tha^{-1}) is usually achieved (Agriculture Research Council-Grain Crops Institute, 2016). Commercial agriculture accounts for 98% of total maize production in South Africa, while the remaining 2% comes from smallholder or emerging producers.

Maize is predominantly produced under dry-land conditions, with less than 10% being produced under irrigation. On average, the Free State Province contributes 31% (64% white maize and 36% yellow maize), the North West Province contributes 13% (78 % white maize and 22% yellow maize) and Mpumalanga Province contributes 29% (59% yellow and 41 % of white) of the total maize output of South Africa (Department of Agriculture Forestry and Fisheries, 2016; The South African Grain Laboratory NPC, 2016). Table 2.1 shows the average provincial maize output for South Africa.

Table 2.1: Maize Production by Provinces in South Africa (Grain SA, 2017)

Season	2009		2010		2011		2012		2013		2014		2015		2016	
PROVINCE																
	Rain-fed	Irrigated	Rain-fed	Irrigated	Rain-fed	Irrigated	Rain-fed	Irrigated	Rain-fed	Irrigated	Rain-fed	Irrigated	Rain-fed	Irrigated	Rain-fed	Irrigated
Western Cape		35000		14000		12400		25000		30000		28500		34200		40000
Northern Cape		605000	13000	573000	6000	509200	6750	585250	6000	644000		638400		644000		675000
Free State	1490000	410000	1575000	327000	1185000	276500	1298000	475000	1720800	590000	2018000	469750	1312000	396500	755500	267500
Eastern Cape	15500	61000	15000	51000	13600	44000	22500	52500	24000	66000	32600	65000	19500	64500	17500	48500
KwaZulu-Natal	165000	108000	158000	94000	140000	95000	185000	85000	205000	110000	196000	96500	195000	88500	173500	133500
Mpumalanga	1525000	55000	1330000	45000	1245000	45000	1530000	95000	1892000	93000	1738500	136500	1488000	117300	1400000	167000
Limpopo	16500	58500	17000	63000	7000	41000	17000	82000	23000	115000	34000	90000	20000	104000	8400	123600
Gauteng	160500	15000	175800	16200	169700	10600	184000	16000	210000	20000	254000	37250	240000	52500	173500	61500
North-West	430000	125000	415000	103000	436000	72000	426000	132000	307000	148000	546000	159000	284000	160000	140000	185000

Proportionately, almost 48% of maize produced in South Africa is white and 52% is yellow. About 5.1% of the area under white maize is irrigated and 94.9% is dryland, while 13.8% of yellow maize is irrigated and 86.2% is rain-fed (Department of Agriculture, Forestry and Fisheries, 2012).

For rain-fed maize, 350 to 450mm of rainfall is required in the growing season in order to achieve commercial yields. Water deficiency is usually the most yield-limiting factor where efficient maize cultivation practices are applied. The maize plant flourishes where average daily temperature is 23ºC. (Du Plessis, 2003).

Maize grows best on deep, rich friable soils. However, in RSA, a considerable amount of maize is produced on sandy soils. Soils with a pH lower than 4.5 are not ideal for maize growth. Flowering of maize is optimum at temperatures between 19°C and 32°C. Temperatures above 32°C are detrimental to yields, while frost can damage maize at all growth stages and a frost-free period of 120 to 140 days is required to prevent damage. (Du Plessis, 2003).

2.3.1.2 Consumption

Maize is the most consumed food commodity in the Republic of South Africa (RSA) , and about two-thirds of maize produced domestically is locally consumed (Dube *et al.*, 2013). Human consumption alone accounts for half of this amount, and together with animal feed account for approximately 80% of the local maize consumption. Figure 2.3 shows human and feed maize consumption trends between 2000 and 2014. The remainder of maize produced in RSA is used for seed and other industrial uses (Department of Agriculture Forestry and Fisheries, 2012). Human consumption and animal feed manufacturing account for 39.8% and 37.4% respectively of the maize produced locally. Exports account for 17.9% of maize produced, while 4.8% is used for the production of starch and glucose (Scheltema *et al.*, 2015). During the 2013/14 season, 2.5 million tonnes of maize were exported mainly to Japan, China, Mexico, Namibia, Zimbabwe and Mozambique (Scheltema *et al.*, 2015). Thus, the importance of South African maize both domestically and beyond its borders cannot be overemphasised. Ways to maintain high production levels in order to cope with the anticipated higher demand both from within and outside RSA need exploration. Yield gap analysis provides a foundation for further studies into where and how production and productivity of maize can be sustainably increased in RSA, thus closing the current yield gaps.

Currently, South Africa's domestic maize production is adequate to meet its annual local consumption requirements. Self-sufficiency has been achieved through the adoption of efficient production technologies by farmers and the continuous improvement of cultivars.

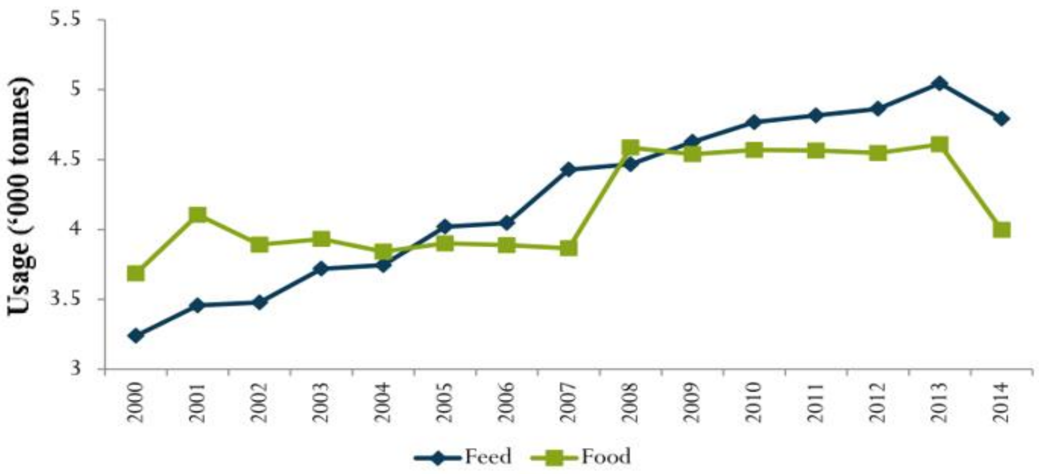


Figure 2.3: Maize Consumption Trends in South Africa 2000 – 2014. (Source: Scheltema *et al.*, 2015)

2.4 Yield gap analysis concepts

This section will review the essential concepts that underpin yield gaps and yield gap analyses. Of particular interest, are the three theoretical yield levels distinguished in literature: attainable yield, actual yield (Ya) and exploitable yield (Ye). Attainable yield can be further distinguished into two kinds; water-limited yields (Yw) (which are obtained in rain-fed systems) and nutrient limited yields. The former, together with Ya are relevant in this study.

2.4.1 Yield Gaps

The yield gap (Yg) of a crop is estimated for a specific location and year/period. Yield gap is defined according to the water regime under which the crop is grown. In irrigated systems, Yg is defined as the difference between potential yield (Yp) and actual average farm yield (Ya). In rain-fed systems, Yg is estimated by the difference between water limited yield potential (Yw), and Ya (Grassini *et al.*, 2015; Lobell *et al.*, 2009).

Accuracy in determining Ya, Yp and Yw is therefore crucial as the yield gap concept rests on the definition and measurement of yield potential.

2.4.2 Actual Yield (Ya)

Crop yield refers to the weight of grain or other economic product, at some agreed standard moisture content, per unit of land area harvested per crop (Fischer, 2015). For maize and other grains, the standard moisture content is 0.155 grams water per gram fresh weight (g H_20. g^{-1} FW), which translates to 15.5% moisture (Setiyono *et al.*, 2010). In yield gap analysis, average yields defined at specific spatial and temporal resolutions and achieved by farmers in a region under the most common management practices are used. Ya is usually smaller than Yp or Yw for since it is difficult to achieve the ideal biotic and abiotic factors for crop growth (Food and Agriculture Organization of the United Nations, 2015).

The number of years required to estimate Ya varies depending on the levels of management and technology applied to each field, and the climatic and edaphic conditions experienced in each particular region. In high yield environments where crops are irrigated, the average of the most recent 5 years will suffice to estimate Ya. Van Ittersum *et al.*, (2013) confirmed this in their studies of maize in Nebraska, USA and wheat in the Netherlands. For rain fed crops, longer time intervals of 15 to 20 years are usually needed to obtain reliable estimates of Ya (Van Ittersum *et al.*, 2013) which account for inter-annual variability in data resulting from factors like weather.

2.4.3 Yield Potential (Yp and Yw)

Yield potential (Yp), is the maximum crop yield that is achievable for a given cultivar when one of or combinations of water, nutrients and biotic stress do not limit growth (Van Ittersum and Rabbinge, 1997). Yp and Yw are determined by the crop genotype, environmental conditions in which the crop is grown and management practices under which the crop is grown. This makes them specific to particular sites. Therefore, yield-defining factors like climatic or weather parameters and crop cultivar characteristics determine the potential yield of a crop, given optimal supply of water and nutrients. Climatic factors that influence potential yield are radiation, ambient CO_2 concentration and temperature (Evans & Fischer, 1999; Van Ittersum *et al.*, 2013). Because these factors vary over the course of a year, crop management factors therefore also influence a crop's yield potential. In this regard, planting dates and maturity ratings influence crop yield potential at a given location. (Yang *et al.*, 2013) reported that in Lincoln, Nebraska, USA, simulated Yp for maize was 2Mgha^{-1} higher with a 7 day increase in hybrid maturity. The results of wheat trials in India also showed that early planted crop yields were 1 Mgha^{-1} higher than those planted a month later. Solar radiation, vapour pressure and temperature also determine photosynthesis and growth, and therefore Yp (Rodriguez & Sadras, 2007). Therefore, yield potential is largely influenced by planting date and cultivar maturity. The maximum value of Yp in irrigated cropping systems, where other than nutrients, pests and diseases; water is not a limitation to crop growth, will thus be the optimum combination of planting date and maturity at a given location (Fischer, 2015; Lobell et al. 2009). Potential yield for a particular crop defines the yield that is to be expected when a crop is grown free of constraints. It describes a crop grown using the best adapted variety, the best management of agronomic and other inputs and in the absence of manageable abiotic and biotic stresses (Fischer, 2015). A specific crop genotype is fully expressed for a particular climatic environment, in terms of ambient temperature and solar radiation (Pingali, 2001).

Water-limited yield potential (Yw) is similar to Yp, the only difference being that water supply is a yield-limiting factor. Yw therefore defines yield potentials in rain fed systems. Over and above the factors already discussed that limit yield in irrigated systems, the water holding capacity, effective rooting depth of soils and field topography also influence Yw in rain-fed systems (Food and Agriculture Organization of the United Nations, 2015; Van Ittersum *et al.*, 2013), as well as evaporative demand of the air (which is influenced by wind speed, vapour pressure and temperature). As highlighted, for a crop's potential yield to be attained, nutrients and water should be non-limiting; and pests, diseases, weeds, lodging, and other stresses are effectively controlled. Location is therefore an important factor that determines a crop's yield potential, making both Yp and Yw site-specific. Location relates directly to climatic and weather conditions that influence crop growth (Evans & Fischer 1999). Thus crops that experience weather damage, like frost or lodging, are excluded from Yp determination.

Although it is impossible to realise Yp or Yw, there are some methods of determining Yp and/or Yw, which vary widely. Data for Yp can be obtained from 3 main sources. These are i) maximum farmer yields, ii) field experiments and iii) crop modelling. The advantage of crop models is that, if calibrated and validated adequately, they are able to reproduce genotype × environment × management (G×E×M) interactions, and, therefore, capture spatial and temporal variations in Yp and Yw, while other methodologies fail to do so (Van Bussel *et al.*, 2015).

2.4.3.1 Maximum Farmer Yields

Maximum farmer yields involve the observation of the maximum yields achieved by farmers in a region of interest. This method determines maximum attainable yield in a specific location for each genotype x management x environment (GxExM) combination (Van Ittersum *et al.*, 2013). Large sample sizes are required to make plausible estimates of Yp. Individual farmers with reliable records are therefore crucial for this method. However, this method is largely appropriate for intensively managed cropping systems, where crop management practices are such that they eliminate yield limiting and reducing factors (Evans & Fischer 1999). This includes application of fertilisers, pest and disease controls that enable

Yp or Yw to be achieved. However, this is not feasible because of diminishing returns when inputs are applied beyond certain levels. It is therefore practically challenging for all surveyed farmers to reach Yp because it would be challenging to avoid all biotic and abiotic stresses.

The results of yields obtained using this method may also be a reflection of a single good year and do not necessarily represent long-term average yield potential for a location (Licker *et al.*, 2010). As observed by Van Ittersum *et al* (2013) in Australia, Kenya and USA, atypical years in farmers' observations can heavily bias Yw and Yp estimates. In Victoria, Australia, results of the study indicated average maximum yields over three years that were above the average simulated Yw over the same period (Van Ittersum *et al.*, 2013).

2.4.3.2 Field experiments

Unlike maximum farmer yields, field experiments try to control biotic and abiotic stresses. Crop management practices that are designed to eliminate yield limiting and yield reducing factors like nutrient deficiencies and diseases are required to be adopted when carrying out field experiments. Field experiments are convenient when there is a lack of data to calibrate and validate a robust crop model (Van Ittersum *et al.*, 2013). However, perfect growth conditions are difficult to achieve and maintain successfully. It is also practically impossible to eliminate all stresses, thus Yp or Yw of a genotype will be inaccurately determined.

Cultivar trials can be used as a measure of Yp and Yw in crops. Seed houses usually carry out trials over a wide area in order to cover a target region. As a result, minimum upscaling is necessary (Fischer, 2015). However, only specific traits of particular cultivars are considered and investigated through cultivar trials, thus Yp might not be accurately reflected from results of cultivar trials (Fischer, 2015). Experiments would also need to be replicated over many years in order to obtain robust estimates of Yp and/or Yw and to account for climatic variation (Lobell et. al 2009).

Real world cropping systems often present constraints to planting and harvesting dates for most crops. Field experiments must therefore represent planting and harvesting dates that reflect the prevailing cropping system in a particular region if they are to produce reliable results for Yp and Yw (Van Ittersum *et al.*, 2013). Measurement of yields in experimental stations often takes place on the most fertile soils with favourable topography, thus rendering them unrepresentative of surrounding production systems. This introduces bias in the Yw estimates when these results are taken to represent a region.

2.4.3.3 Simulation models

Simulation models are simplified mathematical representations of real world systems. The systems' characteristics are reproduced, and studied in an abbreviated time scale (Murthy, 2003). Thus, models represent objects, systems or ideas in a form that is not that of the entity or phenomenon itself, and aid researchers and decision makers to explain, understand or improve a system. However, because of the complexity of the real world, it is difficult to produce a representation that is accurate and captures all specific elements and mechanisms (Murthy, 2003). When used to model crop growth, simulation models use one or more sets of differential equations, and calculate both rate and state variables over time, normally from planting until harvest maturity or final harvest**.** One of the main goals of crop simulation models is to estimate agricultural production as a function of weather and soil conditions as well as crop management.

Models are classified depending on their purpose, and as such, various types of models have been applied for predicting crop yields, including statistical, deterministic, stochastic and empirical models. This research made use of Hybrid-Maize, a simulation model that is used to simulate crop growth of annuals under different levels of production, and thus will be ascertained in greater detail in this literature review.

Van Ittersum *et al* (2013) carried out studies to compare the methods used to determine Yw in western Kenya and Victoria (Australia) for rain-fed maize and wheat respectively, and Yp in Nebraska, United States of America, for irrigated maize. They concluded that simulation modelling produced the most reliable results of Yp, Yw and Yg because interactions among weather, soils and management are accounted for. In cases where there were favourable growing conditions and Ya was high, for example irrigated maize in Nebraska, Yp and Yw estimations were close for model simulations, field experiments and farmer maximum yields. However, in Kenya, there was limited crop management capacity, thus little agreement among these estimates.

2.5 Input data requirements for simulation models, and their required spatial and temporal resolution

Simulation modelling enables the establishment of "consecutively decreasing predicted yield levels by stepwise increasing crop production limitations" (Taylor *et al.*, 2004). The quality of data that is entered into the model is therefore crucial to getting accurate output. Most crop simulation models make use of crop management data, cultivar data, soil data and weather data. Weather data can be at a range of spatial and temporal scales. These scales have a direct bearing on the accuracy of the simulated Yp or Yw, with data at a large spatial and temporal scale producing less reliable results from the simulation.

In instances where weather data is only available from sparsely distributed weather stations, weather data can be scaled up from point estimates at these weather stations to larger geospatial scales. This interpolated data provides global coverage of ecosystems (Van Wart, *et al.*, 2013). According to Baron *et al.* (2005), spatial interpolation of data presents two main weaknesses. Firstly, the variability in temperature, rainfall and solar radiation across a landscape are reduced. Secondly, results are not accurate since crop area is not uniformly distributed across a landscape. These issues can lead to the over or under estimation of yield by as much as 10 – 50% (Baron *et al.*, 2005). The quality of geospatially-interpolated data is also not uniform across the globe because of differences in density of weather stations in different regions. To overcome these shortcomings, the use of actual weather data over a large temporal scale (10 – 15 years) from ground stations that represent the spatial distribution of crop production is recommended (Van Wart *et al.*, 2013). As such, the GYGA protocol addresses these issues and this is discussed in section 2.6. The accurate estimation of Yp using simulation models is therefore dependent on input data availability and quality. However, it is also crucial to ascertain how the simulation responds to short-term unique input conditions, especially with regards to weather data (Lobell et. al, 2009).

Apart from the spatial resolution, the most appropriate time-step for weather data that is used to simulate crop yields has to be determined. Daily weather data provides the most accurate estimation of Yp and Yw. Daily data fully captures the impact of current crop management practices and how they are related to current crop management practices. Yp and Yw are also highly sensitive to planting date and cultivar selection. These determine the timing of key growth stages and length of crop growing season (Bai *et al.*, 2010; Van Wart *et al.*, 2013; Grassini *et al.*, 2015; Van Wart *et al.*, 2015).

Selection of an appropriate simulation model to estimate yield is essential if reliable results of Yp and Yw are to be obtained. Accurate calibration and validation of the selected model should be carried out in order to obtain reliable output. Models should be validated against yields of crops grown in fields where yield limiting factors have been eliminated (Van Wart *et al.*, 2013). Models used in previous studies, (De Vries et al. 1997), had shortcomings like the use of non-species specific relationships between solar radiation and plant biomass production. This meant that crop phenology could not be simulated, thus there was a need for methods that were robust, reproducible and transparent.

The Global Yield Gap and Water Productivity Atlas (GYGA) protocol was developed to attempt to account for these shortcomings. In its development, it attempted to address aforementioned issues related to minimum data requirements, required level of crop management data specificity and upscaling from local to national levels (Van Wart *et al*., 2013.

2.6 Global Yield Gap and Water Productivity Atlas (GYGA)

GYGA (www.yieldgap.org) is basically a GIS database and web-based mapping interface that contains yield gap information for various crops from all over the world. Yg analyses for major food crops are carried out in each country based on locally observed data, and the results are added to the database. GYGA "provides robust estimates of untapped crop production potential on existing farmland based on current climate and available soil and water resources," (Global Yield Gap Atlas, 2015). Here, yield gaps are estimated for major crops around the world with "local to global precision and relevance" (Van Bussel *et al.*, 2015).

Local knowledge of crop management and production environments is essential for the GYGA project. Therefore, agronomists with such knowledge work in collaboration to provide data that is used for Yw, Yp and Yg estimations (Global Yield Gap Atlas, 2015). Local results are then upscaled to regional and national levels. GYGA collates these Yg results into a database, with the aim of global coverage for all major food crops and countries that produce them. The project initially focused on cereal crops, and has recently broadened the list to include soyabean, sugarcane and potato (Global Yield Gap Atlas, 2015).

The methods used to estimate Yg for each of the crops in the different countries follow a similar protocol, thus they are comparable, reliable and reproducible. Unlike previous methodologies that would estimate Yg based on gridded weather data, the GYGA approach makes use of location-specific observed data; the so called "bottom-up" approach (Van Wart *et al.*, 2013).

2.6.1 The GYGA protocol

The GYGA protocol is a methodology that was developed to generate national yield gap estimates based on local data. This bottom-up approach is based on a climate zonation scheme, whereby agro-climatically similar areas are grouped for analysis (Van Ittersum *et al.*, 2013; Van Wart *et al.*, 2013; Global Yield Gap Atlas, 2015). Key climate zones (CZs), which are central to data aggregation within the protocol, are selected using guidelines outlined, and then reference weather stations (RWS) that represent these CZs are selected. Buffer zones that are 100 km in radius are created around each RWS, and Yw and Yp is determined for each of these buffer zones using crop models that are validated for local conditions and parameterised for local agronomic and soil conditions (Global Yield Gap Atlas, 2015; Van Bussel *et al.*, 2015).

Ya and Yw are then scaled up to regional and national scale using cropland area weighted averages (Van Bussel et al. 2015). For a crop to be evaluated in GYGA, its national harvested area needs to be more than 100 000 hectares (ha) (Global Yield Gap Atlas, 2015). The site selection process employed by GYGA facilitates focus on quality of data collected by limiting the number of locations for which site-specific weather, soils and cropping systems data is required.

The tiered data selection methodologies ensure data of the highest quality and relevance is used first. Furthermore, results of Yg, Yw and Yp can be validated and revised as data quality further improves and becomes available (Global Yield Gap Atlas, 2015). Utilising a global climate zonation scheme also facilitates comparisons and learning between cropping systems in different parts of the world (Gobbett *et al.*, 2017). It is however important to make use of the most locally relevant and highest quality data, which is substituted by lower-quality alternatives when necessary (Gobbett *et al.*, 2017).

2.6.2 Data collection in GYGA

As mentioned in previous sections, methods that were previously used to estimate yield gaps lacked transparency, were too empirical and were difficult to validate and reproduce. The GYGA protocol was therefore developed to try and address these issues (Van Bussel *et al.*, 2015; Van Ittersum *et al.*, 2013).

Crop simulation models are used to estimate Yw and Yp using the GYGA protocol. Weather data, crop management data and cropping system data are used to simulate Yw or Yp. Crop management data includes planting and harvesting dates, cultivar maturity and optimum plant populations. Soil data includes root zone water holding capacity, soil depth and slope. Weather data includes daily minimum and maximum temperatures, precipitation, solar radiation and relative humidity (Van Ittersum *et al.*, 2013; Global Yield Gap Atlas, 2015; Grassini *et al.*, 2015). The Hybrid-Maize model assumes optimal nutrient management and does not account for yield losses from weeds, pests, lodging or other stresses.

The sources of these data follow a tiered protocol, whereby the most desirable sources that give the most accurate estimation of Yp, Yw and Yg are preferred. However, in cases where these data are unavailable or of poor quality, GYGA dictates that proxy data be used.

2.6.2.1 Site Selection

The estimation of yield gaps occurs at varying spatial scales. The smallest spatial unit for Yg estimation is the point, which is a specific location within a region and is usually represented by a reference weather station (RWS), to buffer zones, climate zones, regions and countries. Climate zones (CZs) account for variation in climate, and best represent how a crop is produced in terms of weather, soils and cropping system are selected together with specific points and buffer zones within these CZs (Global Yield Gap Atlas, 2015). The GYGA climate zonation scheme is defined by growing degree days, temperature seasonality and aridity index (Van Wart *et al.*, 2013; Global Yield Gap Atlas, 2015). Other climate zonation schemes are defined by slightly different parameters. Van Wart *et al.*, (2013) compared global agro-climatic zonation schemes for suitability to upscale Yp and Yw from location scale up to regional, national and global scales. They compared the Global Agro-Ecological Zone modelling framework (GAEZ), The Centre for Sustainability and the Global Environment (SAGE) zonation scheme, The Global Environmental Stratification (GEnS) and GYGA. Of these zonation schemes, the authors concluded that one of two suitable methods for upscaling was the GYGA protocol because of its little within zone climatic heterogeneity. Unlike the other schemes that make use of gridded weather data, GYGA makes use of specific point weather data from the RWS. Furthermore, the crop specificity of the other CZ schemes limits their utilisation for scaling up Yg evaluations in regions where crop rotations are practiced as opposed to mono-cropping (Van Wart *et al.*, 2013).

After identification of CZs within a country, those with greater than 5% of harvested cropping area for the crop under consideration are selected. The Spatial production allocation model (SPAM) 2005 maps (You *et al.*, 2014) are used to determine cropped areas for maize. SPAM 2005 provides annual harvested area data at 10km x 10km resolution grids. SPAM databases are created using data from national statistics agencies for the different countries; as well as international sources, for example FAOSTAT (Van-Bussel *et al.*, 2015). The selected CZs are known as designated climate zones (DCZs), which should cover more than 50% of the national crop area (Global Yield Gap Atlas, 2015). Results of a study by (Van Wart *et al.*, 2013) indicated that for robust estimates of Yp and Yw to be achieved, 40 - 50% of total national production area should be covered by the DCZs (Van Wart *et al.*, 2013; Global Yield Gap Atlas, 2015).

The weather stations within the DCZs that contain more than 1% of total harvested area are then identified. These weather stations are further ranked based on harvested area, until 50% of national harvested area is achieved. The resultant stations are the so-called reference weather stations (RWS), for which weather data are collected. The selection protocol of the RWS ensures that highest quality stations in terms of crop harvested area covered are selected. The GYGA protocol specifies the creation and selection of buffer zones with a 100km radius around the RWS. (Van Wart *et al.*, 2013; Van Bussel *et al.*, 2015) go on to say that as prescribed by GYGA, a 50% coverage of national crop harvested area by the selected buffer zones is sufficient to produce robust Yg estimates where topography is relatively uniform.

However, Hochman *et al.* (2016) argue that the GYGA protocol still needs further evaluation and validation with regards to the size and zone of influence of the RWS. They further scrutinised the inadequate representation of soils in the creation of RWS. Van Bussel *et al.*, (2015) also highlighted in their study that scaling up Yg estimates in semi-arid rain-fed areas is prone to errors, especially where the number of weather stations are limited in each climate zone. Therefore, the need to validate the GYGA protocol remains key, especially in semi-arid areas with high spatial and temporal variability in rainfall (Hochman *et al.*, 2016).

2.6.2.2 Actual yield determination

The average annual crop yield obtained by farmers in a geographic area defines (Ya) (Grassini *et al.*, 2015). The collection of yield data is done at various spatial scales in different parts of the world. However, the lack of good quality, subnational scale data is identified as the weakest link in Yg analyses (Van Ittersum *et al.*, 2013). The GYGA protocol therefore attempts to minimise unreliability of actual yield data by employing a tiered approach for the collection of Ya data.

Grassini et al. (2015) identified 4 key aspects of the GYGA protocol that attempt to establish Ya data reliability. These are: (i) level of disaggregation by crop and water regime, (ii) number of available data-years, (iii) spatial resolution, and (iv) data quality. While traditionally Ya had been generally based on yields reported in publications like FAOSTAT which were spatially coarse and are likely to be a source of error and uncertainty in Yg estimation, GYGA makes use of reliable survey data at finer spatial scales (Van Ittersum *et al.*, 2013; Hochman *et al.*, 2016).

GYGA dictates that data on crop yields in each RWS should be at the finest level of spatial resolution, which is typically a district or municipality in the case of SA (Global Yield Gap Atlas, 2015). However, in sub-Saharan Africa (SSA), data at this level of spatial resolution are not always available. In this case, a tiered approach is followed to determine reliable values of Ya. Estimations can be made from values reported for larger administrative units, like provinces. However, these values may not be representative of the targeted area's Ya (Gobbett *et al.*, 2017). Progressively lower quality data may be incorporated into the Ya analysis (Hochman *et al.*, 2016).

When determining average Ya using the GYGA protocol, as many recent years of Ya data should be included to account for weather variability, while avoiding bias due to a technological time-trend (Van Ittersum *et al.*, 2013). The years for which Ya and Yp and/or Yw are determined should also be similar. In data rich countries, yield data for the 5 most recent years is recommended to calculate average yield if there is a steep yield trend. If there is no trend, data for 10 years can be used (Global Yield Gap Atlas, 2015; Grassini *et al.*, 2015). In data poor countries where rain-fed crop production is practiced, longer time frames can be considered, and a compromise must be found between adequately capturing variability on the one hand and avoiding the inclusion of technological change (possibly including climate change) on the other hand (Van Ittersum *et al.*, 2013).

2.6.2.3 Weather Data

Weather data selection using the GYGA protocol, like other parameters required for simulation, follows a tiered methodology whereby the most reliable and accurate data source gets priority for use. In rain fed environments, a minimum of 10 to 15 years of data is recommended, while 15 – 20 years may be necessary to account for weather variation (Van Wart *et al.*, 2013; Grassini *et al.*, 2015). Yg studies done in Argentina by Grassini *et al.* (2015) conclude that 10 years of weather data are sufficient to estimate average yields and a coefficient of variation that are within 10% of estimates obtained with a 30 year database. Van Ittersum et al. (2013) further argue that the period for which weather data is collected should be such that the analysis will not be affected by technological uptake or climate change.

The sources of weather data range from high quality, long term meteorological stations to mostly or fully generated or gridded weather data. The first preference would be an existing weather station within or as close as possible to the RWS with long term weather data. This is followed by a weather station close to the RWS as well, but with shorter term weather data, typically at least 10 years (Global Yield Gap Atlas, 2015).

The use of daily weather observations is recommended, as interpolated monthly observations may lead to overestimations of simulated yields especially in areas with high daily weather variability (Van Ittersum et al. 2013). The weather parameters that are required to simulate Yw and Yp are solar radiation, maximum and minimum temperature (Tmax and Tmin respectively), precipitation, reference evapotranspiration (ET_O), vapour pressure and sometimes wind speed; all at a daily time-step.

Measured solar radiation, which is usually not available from weather station data, can be derived from National Aeronautics and Space Administration (NASA) agro climatology weather data (Bai *et al.*, 2010). Studies done by Bai *et al.* (2010) across 39 sites in China showed that when NASA solar radiation data was used in combination with ground station temperature data, simulated Yp correlated well with simulations using only using weather station data with a R^2 value of 0.89 across all sites. In the absence of measured data, vapour pressure can be estimated from the measured Tmin. Local validation using observed data from a representative spatial and temporal scale is however necessary (Grassini *et al.*, 2015).

The GYGA protocol also suggests quality control measures for weather data, which can be compromised by missing values or sometimes suspicious values (Grassini *et al.*, 2015; Van-Wart *et al.*, 2015). For a particular year, at least 80% of precipitation, Tmax and Tmin data should be available. Furthermore, less than 20 consecutive days of Tmax and Tmin values should be missing, with the period decreasing to 10 consecutive days for precipitation.

2.6.2.4 Model calibration

The GYGA protocol requires that crop simulation models be calibrated to local conditions (Van Ittersum *et al.*, 2013; Van Wart *et al.*, 2013). Crop model calibration accounts for differences in crop phenology and growth related factors caused by the different crop cultivars and the different regions in which they are grown (Global Yield Gap Atlas, 2015). Therefore, models may differ based on crop or region as long as they have been validated.

In GYGA, either elaborate calibration or simple calibration can be done. Elaborate calibration makes use of data from field experiments where there are no biotic or abiotic stresses. All weather, soil and management data required for simulation are available (Global Yield Gap Atlas, 2015). Simple calibration on the other hand is used if elaborate calibration is not possible.

Models such as WOFOST (Boogaard *et al.*, 2014) and Hybrid-Maize (Yang *et al.*, 2013) are used to simulate maize Yp and Yw under different conditions. The robustness of a particular model is determined by the quality of data that is used to calibrate it (Grassini *et al.*, 2015).

2.6.2.5 Determining soil series

The GYGA protocol dictates that local relevance is important in characterising soil properties. Soil hydraulic properties which determine available water for plants, as well as landscape and soil properties that influence infiltration and runoff both influence Yw (Global Yield Gap Atlas 2015). Soil data should be based on observed data if it is available, and consideration should only be given to major soil types utilised in a cropping system (Gobbett *et al.*, 2017). However, this bias in soil type could result in an overestimation of Yw.

Different crop simulation models use different soil input data sets to simulate Yw. However, the common parameters are rooting depth, plant available soils water (PASW) and slope and drainage class (Global Yield Gap Atlas, 2015). The rooting depth is the soil depth within which a crop's root system can absorb water and nutrients without physical or chemical hindrance (Grassini *et al.*, 2015). Effective soil depth is also hindered by crop genotype and length of growing season. In most rain fed environments, the rooting depth for most grain crops is 1.5 m (Global Yield Gap Atlas, 2015; Grassini *et al.*, 2015).

PASW is determined by the upper and lower soil limits for water retention, as defined by the field capacity and wilting point respectively (Global Yield Gap Atlas, 2015). Soil water characteristics are typically estimated using pedo-transfer systems (PTF) based on soil texture since actual measurements are rarely available (Grassini *et al.*, 2015).Crop models estimate PASW differently, with some like WOFOST use the PASW values reported in soil databases directly. Hybrid-Maize makes use of soil texture as specified by the user to estimate PASW (Global Yield Gap Atlas, 2015). The incorrect specification of PASW and rootable depth can result in false values of Yw.

Terrain slope and drainage class are used to estimate water loss by runoff in the GYGA protocol. Since some rainfall will not be available for plant uptake due to runoff, locations with slopes greater than 2% and/or poor drainage should have runoff accounted for (Global Yield Gap Atlas, 2015).

2.6.2.6 Crop management data

For accurate estimation of Yw, Yp and Yg to be made, the cropping sequence followed in a particular area needs to be specified. Different crop sequences followed for the same crop under the same conditions have been shown to produce different estimates of Yw, Yp and Yg (Grassini *et al.*, 2015). Average sowing date, cultivar maturity and plant density define each cropping system.

The GYGA protocol is such that local agronomic management rules and practices are used for crop simulations (Global Yield Gap Atlas, 2015). However, cropping system details are usually not readily available especially in data deficient regions like SSA. For wheat cropping systems in Australia, Gobbett *et al.* (2016) surveyed consultants from 4 states for best practices. Generic sowing rules which varied in detail depending on the detail provided by the consultants in the different cropping regions were defined beforehand.

As shown in Figure 2.4, the dominant annual crop rotation is firstly considered. However, intercropping systems are not considered. Yp or Yw values are simulated for each crop cycle. Weighted averages based on proportion of harvested area of each cycle are used where there is more than one crop cycle (Grassini *et al.*, 2015). The target crop is then specified, as well as its water regime. The percentage total harvested area that represents the target crop is specified, as well as the planting and maturity dates. Plant populations and soil properties are also reported, as well as the rotation under that soil type (Global Yield Gap Atlas, 2015).

GYGA DATA FORM FOR A GIVEN CROPPING SYSTEM IN A REFERENCE WEATHER STATION BUFFER ZONE

Sign In

COUNTRY: WEATHER STATION NAME:

LATITUDE/LONGITUDE (decimal degrees): NEAREST LOCATION (district, county, municipality, city):

SOURCES OF THE REPORTED DATA BELOW (e.g. national soil maps, expert opinion, etc.): see Table page 2

Cropping system[1]	Crop [2]	Water regime[3]	% crop area under this system [4]	Sowing date / rule[5] (mm/dd)		Cultivar maturity[6]	Plant density [7]		Dominant soil types [8]	Soil type description [9]			% of system on each soil type [10]
				Actual	Optimal		Actual	Optimal		Texture	Soil depth	Slope	

Figure 2.4: Requirements for crop management data in GYGA *(Source: Global Yield Gap Atlas, 2015)*

Crop management data can be a source of error in simulations. Actual management practices in a particular area might fall far short of optimal practices required for the production of a particular crop. Sometimes sowing dates are delayed, or plant densities are sub-optimal especially in resource-scarce areas like SSA (Grassini *et al.*, 2015). In such cases, for example in South Africa where crop densities can be as low as 10,000 plants per hectare (Pannar, 2017), it is useful to distinguish between simulations using actual management and optimal management.

Crop cultivars are constantly being improved for yield increase and stability, therefore Yw and Yp simulations should be based on recently released cultivars grown with no biotic and abiotic stresses and used by the majority of farmers in a region (Grassini *et al.*, 2015). Growing degree days (GDD), which represent cultivar maturity, are usually not specified in the developing regions of the world. To address this, the period for physiological maturity are noted, and then GDD derived through simulations (Grassini *et al.*, 2015). This was the case in this research where the GDD to maturity were derived from period to physiological maturity as indicated in the Panaar Catalogue (Pannar, 2017). This approach however can still produce bias arising from uncertainties in simulated flowering dates or when harvest dates, which often differ according to location and socio-economic conditions, are used instead of physiological maturity (Grassini *et al.*, 2015).

2.7 Chapter summary

This chapter reviews and contextualises the concepts employed in this study. The main concepts and issues reviewed include population dynamics and how it influences global food demand (in particular maize demand), South African maize production trends and statistics, crop growth modelling, the concepts and methods of yield gap analysis and finally the Global Yield Gap Atlas (GYGA) protocol. It aslo served to highlight the importance of yield gap analyses and the need to conduct such studies. Methods that are used to review such studies were further reviewed, with an emphasis on the GYGA protocol.

Many researchers agree that significant changes to the agricultural and associated sectors will need to be implemented in order to feed the expected world population of 9 billion people by 2050. At the same time, the primary source of production, land, is becoming limiting due tio competing uses with other social and ecological needs. Any increase in agricultural production thus has to be achieved on more or less the same area of land as currently under production. However, environmental and socio-economic factors may be limiting. Therefore, there is need to quantify the levels up to which crop production can still be increased in different regions of the world.

Analyses of the production potential of land for the major food crops at a global scale using traditional methods is a cumbersome process, that might be prone to errors associated with scale or time. As discussed in the review, the GYGA protocol provides a methodology that seeks to eliminate bias, while achieving transparency, robustness and reproducability. Its strong agronomic foundation and the use of a globally consistent procedure allows validation of results of Yp, Yw and Yg. Most researchers seem to agree with the fact that the protocol produces reliable estimates of yield gaps from local to global scales. For maize in South Africa, the GYGA protocol gives simulated values of Yg that will enable policy formulation for the sustainable intensification of this important dietary crop. Insights into the constraints maize farmers are facing will provide a basis for further research aimed at achieving increased yields. Methods used to simulate maize yield gaps in South Africa are discussed in the next chapter.

CHAPTER 3

MATERIALS AND METHODS

3.1 Introduction

This study engaged a modified GYGA (www.yieldgap.org) protocol for yield gap analysis. As already explained in Chapter 1 and Chapter 2 of this thesis, the GYGA protocol is a bottom-up methodology that is based on a climate zonation scheme whereby areas that have similar agro-climatic characteristics are grouped for analyses (Van Ittersum *et al.*, 2013; Van Wart *et al.*, 2013; Global Yield Gap Atlas, 2015). The main concept underpinning GYGA is the use of relevant, high quality data at local level, which can be substituted with lower quality alternatives only when necessary (Grassini *et al.*, 2015). This section seeks to explain the methodology used in this study to estimate Yw and Ya for yellow and white maize in South Africa, and how these were used to calculate the maize Yg.

3.2 Study Area

The study area was selected according to the procedures prescribed by the GYGA protocol, which state that the minimum harvested area of any crop under study be 100 000 ha (Van Ittersum *et al.*, 2013; Global Yield Gap Atlas, 2015). Maize in South Africa is grown on an average area of 2.5 million ha annually (Scheltema *et al.*, 2015), thus satisfying the minimum harvested area requirement. Currently, the main commercial maize production areas in RSA are defined as the Eastern Production Region, and the Western Production Region (Figures 2.1 and 2.2). These areas generally receive a minimum of 400mm of summer rainfall annually. Good yields are dependent on even distribution of rainfall throughout the growing season. Soils vary from the lighter loams and sandy-loams in the west; to the heavier clay-loams and clays in the east. The Western Cape is unsuitable for rain-fed maize production as it is a winter rainfall area. A summer grown maize crop could be irrigated, but the high temperatures and strong winds experienced in this area mean that water consumption is high, and soil salinity might affect crops.

3.2.1 Study Site Selection

The site selection process followed a step-wise protocol that was adopted from the GYGA methodology. Site selection involved the use and manipulation of digital maps and R-programming scripts in a Geographic Information Systems (GIS) environment. Spatial production allocation model (SPAM) 2005 maps (You *et al.*, 2014) and the GYGA climate-zonation shape files were processed in ArcMap 10.2.1 (ESRI, 2013), and the end result was the delineation of designated climate zones (DCZs). Figure 3.1 gives an overview of the site selection procedure applied in this study.

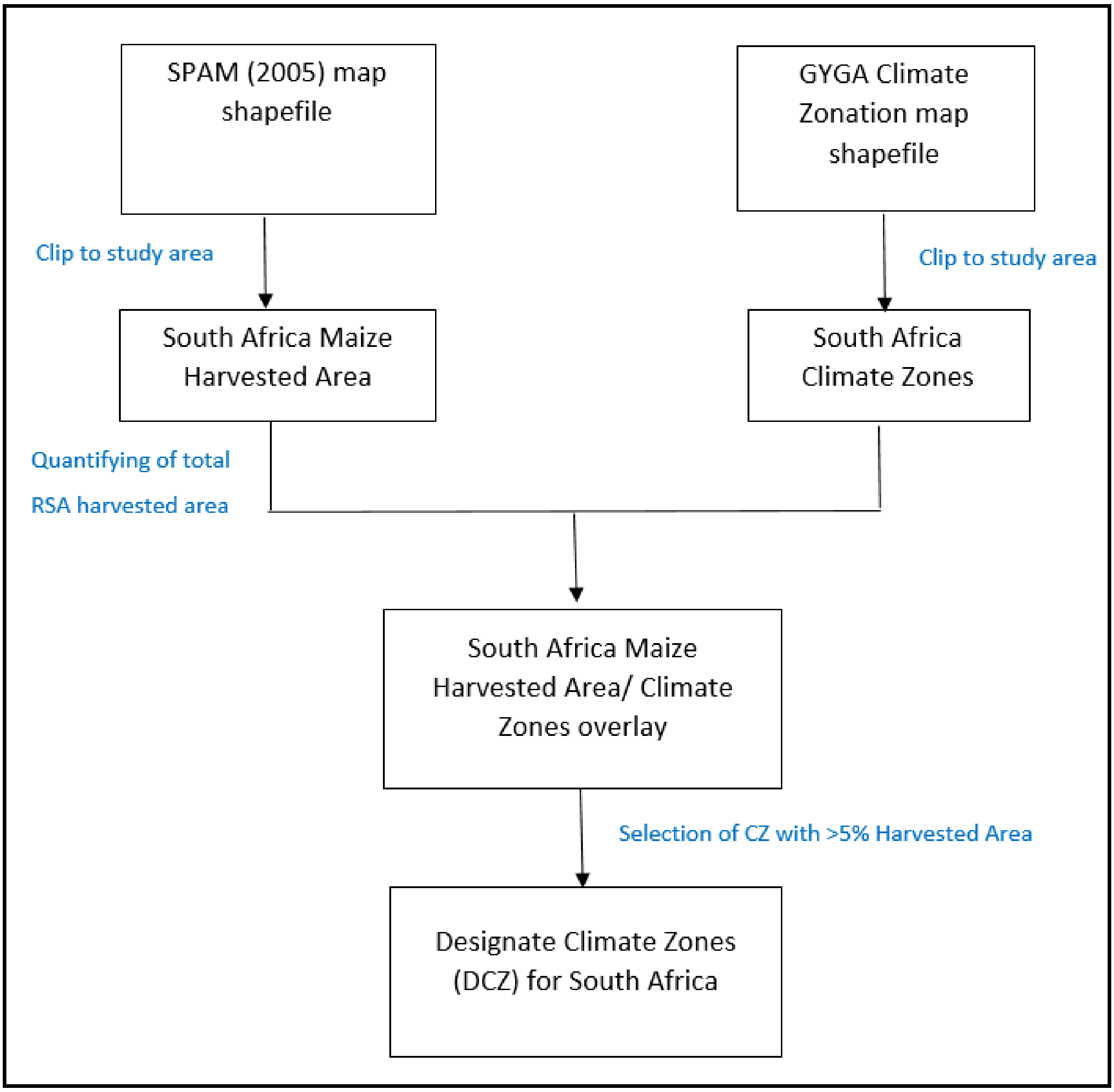


Figure 3.1: Process of deriving DCZs for South African maize in ArcMap 10.2.1.

The first step in the site selection procedure was to download the GYGA Climate Zonation shape file from the GYGA website (www.yieldgap.org). ArcMap 10.2.1 (ESRI, 2013) software was used to clip the South African borders from the global GYGA climate zones vector file and remain with just the South African Climate zones shape-file, which is shown in (Figure 3.2).

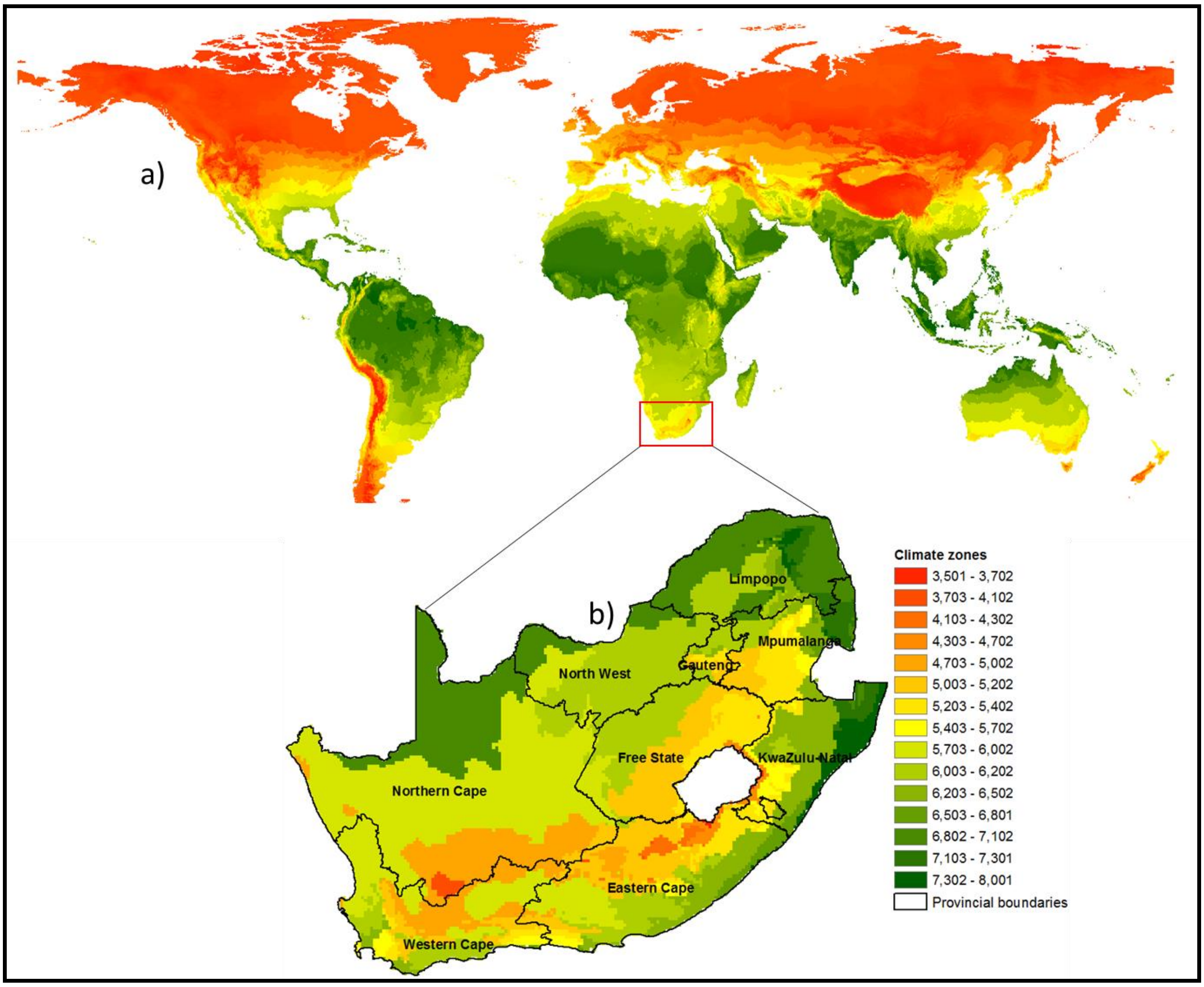


Figure 3.2: (a) GYGA Global Climate Zonation and (b) South African Climate Zones extracted from the GYGA Global Map.

The next step was to determine the average annual maize harvested area for South Africa. To achieve this, the SPAM (2005) (You *et al.*, 2014) data were used. The SPAM (2005) data were downloaded as a shape file, and clipped using the South Africa boundary shape-file, as shown in Figure 3.3.

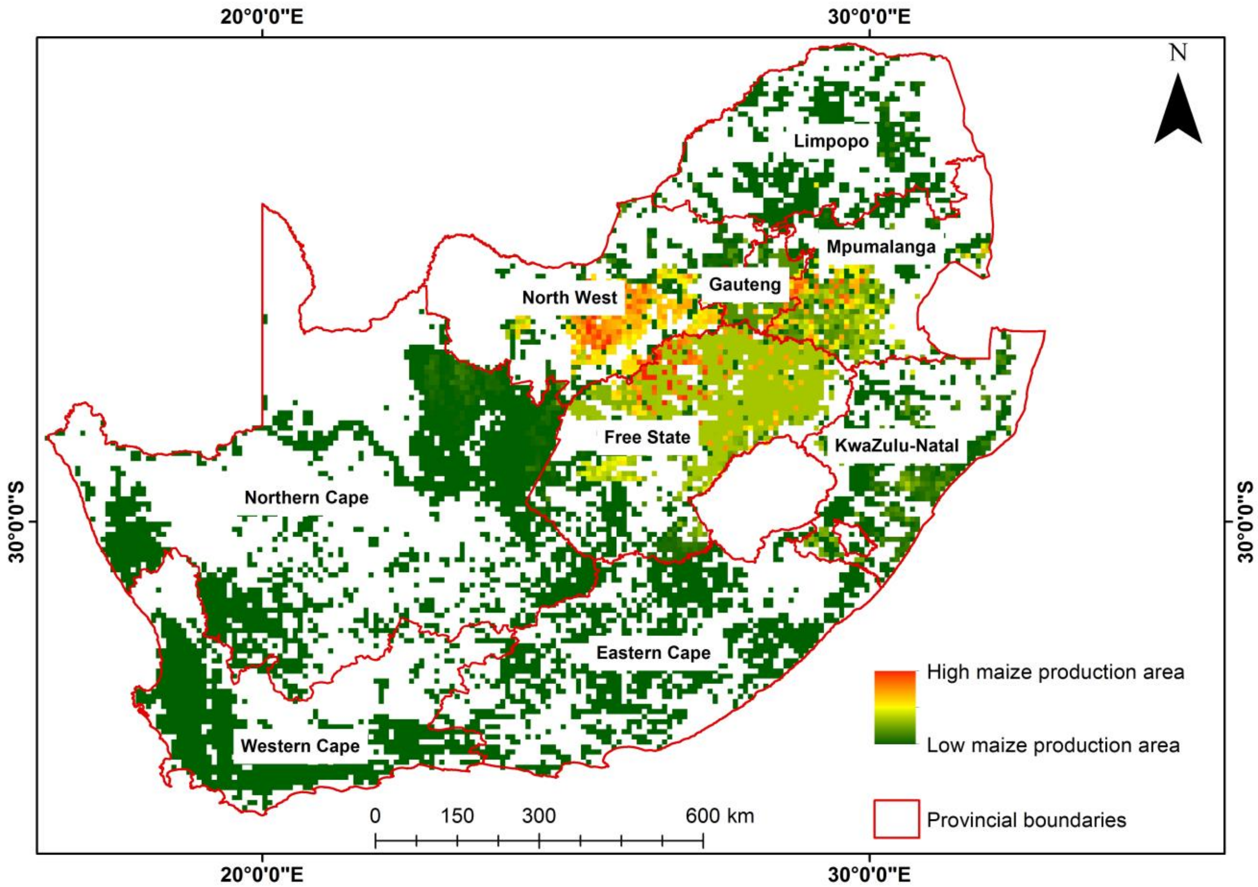


Figure 3.3: Potential rain-fed maize production areas extracted from SPAM (2005). Note: "High maize production area" and "Low maize production area" refer to only the production potential (tha^{-1}) of the said area; not the actual production.

SPAM (2005) maps provided data for average annual harvested area, as well as an indication of high and low production areas for dryland maize in South Africa. With the aid of ArcMap 10.2.1 (ESRI, 2013), the total harvested area according to SPAM (2005) (You *et al.*, 2014) was calculated, and was found to be 2 417 850 ha. This was in agreement with the statistics given in literature (Scheltema *et al.*, 2015). After the high potential production areas were identified and the corresponding total harvested area for maize quantified, the GYGA Climate Zone shape file for South Africa was overlaid with the downloaded SPAM (2005) map. A Python Script that was incorporated into ArcGIS was used to select climate zones, which contain more than 5% of total national harvested area of maize. The climate zones that resulted from this selection process were 5202, 5301, 5302, 6102 and 6202. These selected climate zones are known as the designated climate zones (DCZs), and they were found to contain 67% of the total national harvested area for maize, which is within the minimum prescribed GYGA proportion of 50%. They cover parts of the Mpumalanga, North West, Free State, Gauteng, Limpopo, Northern Cape and Eastern Cape provinces.

3.2.2 Selection of Reference Weather Stations (RWS) and Delineation of Rainfall Weather Station Buffer Zones

Following identification of the DCZs, the next step was to identify reference weather stations (RWS) and delineate buffer zones around the RWS. Data processing was done in ArcMap 10.2.1 (ESRI, 2013) and an overview of the process is shown in Figure 3.4.

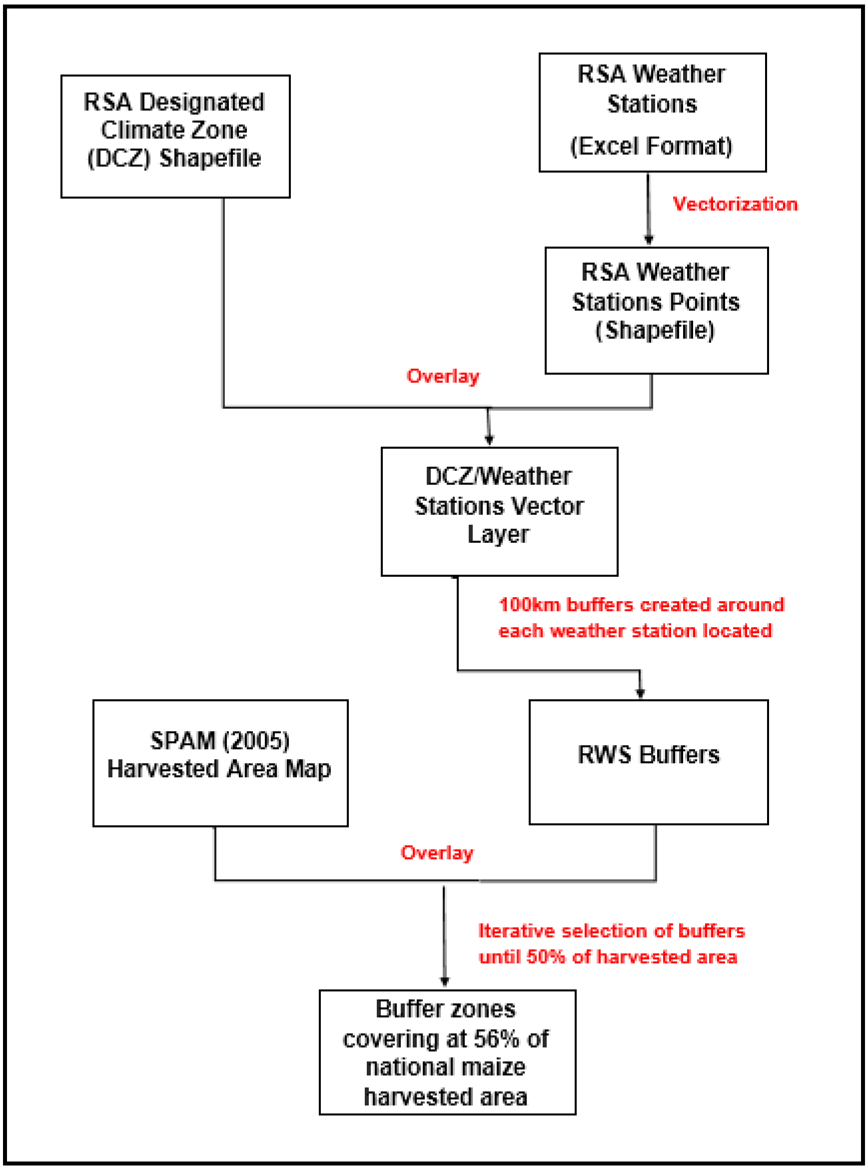


Figure 3.4: Process for selection of RWS buffer zones in ArcMap 10.2.1.

The South African Weather Service (SAWS) provided a file that contained weather data for the whole of South Africa. This Microsoft Excel file contained the location coordinates weather parameters for all weather stations (Section 3.4.1). The excel file was converted into a shape file of all the weather station points, which were then overlaid with the DCZ layer, and buffers with a radius of 100 km created around said weather stations. Subsequently, within those buffers, only those areas were kept where the climate zone is identical to the climate zone at the precise location of the weather station.

The buffer zones shape-file was then overlaid with the SPAM (2005) map for RSA. These buffers were iteratively ranked and selected for harvested area using a Python (Python Software Foundation, 2013) script until the area within the selected buffer zones reached at least 56% of the total maize harvested area for RSA, a total of 1 437 287 ha. The resultant RWSs were Bethal, Bethlehem, Bloemhof, Ermelo, Mafikeng, Potchefstroom, Wepener and Witbank (Figure 3.5). The minimum distance between the selected weather stations was 180 km as per GYGA guidelines. In addition, buffer zones that straddled 2 or more climate zones were clipped at the edges of the CZs, thus producing jagged boundaries. These buffer zones are the study area, or the smallest spatial units that were employed in this study, from which data were collected.

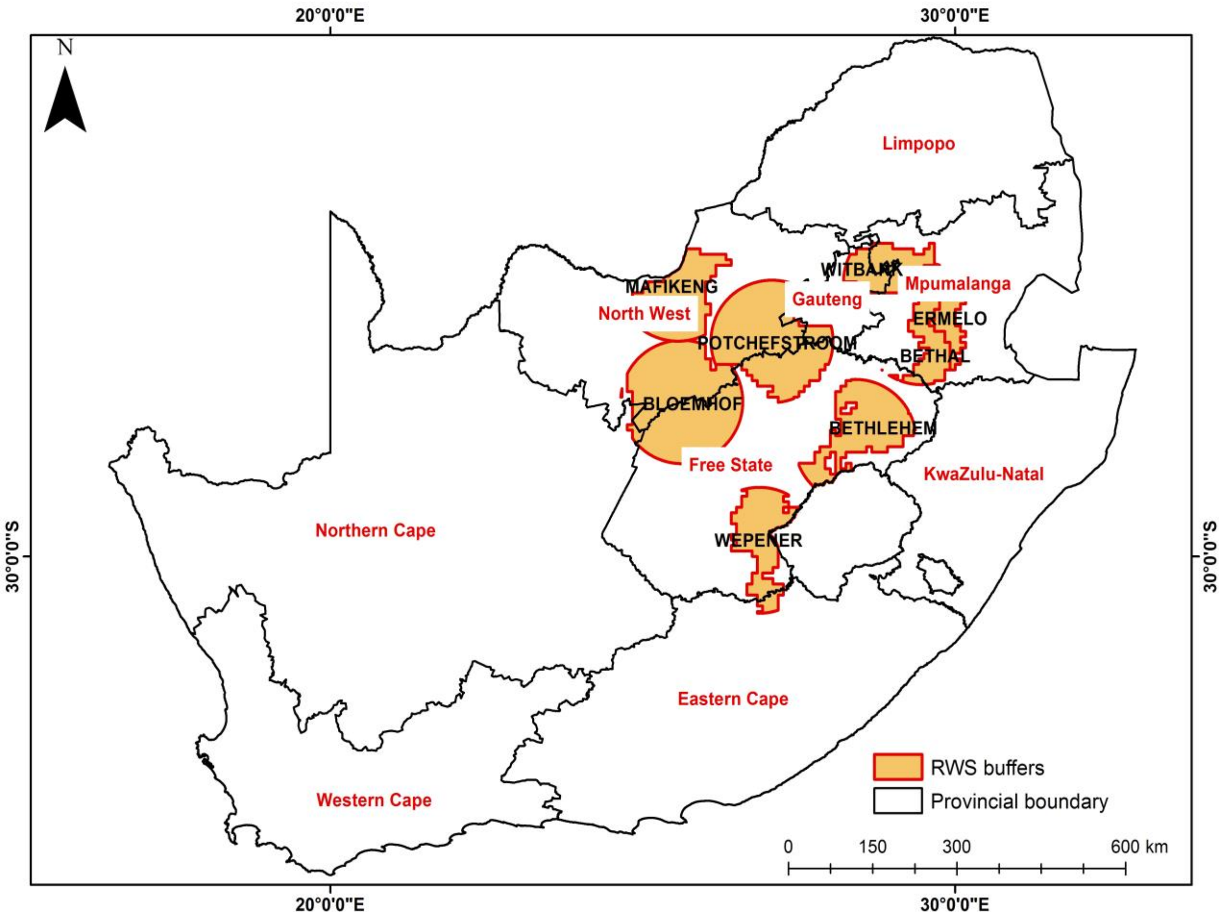


Figure 3.5: RWS buffer zones for rain-fed maize data collection sites.

3.3 Weather Data

A total of 8 RWSs were identified for data collection using the site selection procedure discussed in sections 3.1 to 3.2.2. For each of these stations, weather data from the SAWS for the period from January 2000 to December 2014 was used to simulate Yw. The weather parameters that were required for the simulations included:

i) Rainfall (mm)
ii) Minimum temperature (°C)
iii) Maximum temperature (°C)
iv) Humidity (%).
v) Evapotranspiration (mm)
vi) Solar radiation (MJ m^{-2})

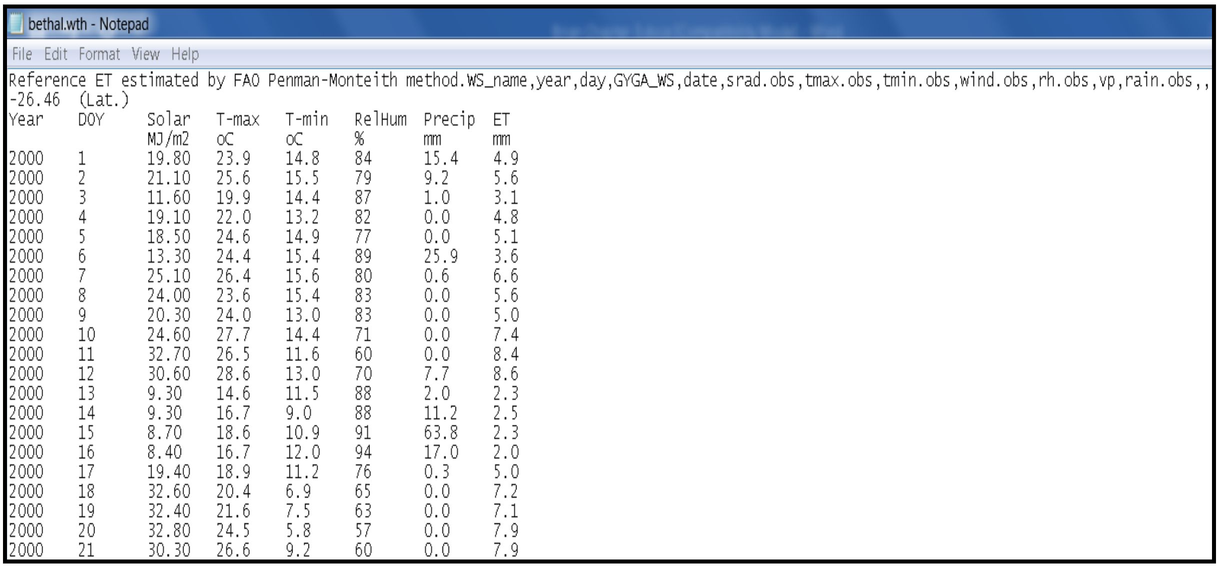
The solar radiation data was not provided by the SAWS and had to be sourced as interpolated data from the Prediction of Worldwide Energy Resource dataset from National Aeronautics and Space Administration Prediction of Worldwide Energy Resource (NASA–POWER) (Stackhouse *et al.*, 2015).

3.3.1 Solar Radiation

Field experiments done in China by Bai *et al.* (2010) showed that a combination of NASA solar radiation data with ground station temperature data could be used to fill geospatial weather data gaps when simulating maize yield potential. Therefore, in this maize yield gap study, data for solar radiation was retrieved from the NASA-POWER website (Stackhouse *et al.*, 2015). The coordinates for the RWSs were entered into an interface on the website, and the corresponding values for solar radiation were downloaded as text (.txt) files. They were then converted into Excel csv files, which were processed in ArcMap 10.2.1 (ESRI, 2013) using a Python script (Python Software Foundation, 2013) and saved as '.*wth*' format files.

3.3.2 Processing of Weather Data

The weather data that was received from SAWS was in Microsoft Excel format, and had to be converted to '.*wth*' format, which is readable by Hybrid-Maize (Yang *et al.*, 2013). In order to achieve this, the weather files were first checked for missing data using a Python Script. Following the GYGA guidelines, no more than 20 consecutive missing days for T_{max} and T_{min}, and no more than 10 consecutive missing days for precipitation would be permissible. Thereafter, a Python script was used to arrange the data into rows and columns that are specific to the requirements of the simulation software. Figure 3.6 illustrates this required format where there are 8 columns for year, day of the year, solar radiation ($MJ.m^{-2}$), maximum temperature (T_{max}, ^{o}C), minimum temperature (T_{min}, ^{o}C), relative humidity (%), precipitation (mm) and evapotranspiration (mm).



```
bethal.wth - Notepad
File  Edit  Format  View  Help
Reference ET estimated by FAO Penman-Monteith method.WS_name,year,day,GYGA_WS,date,srad.obs,tmax.obs,tmin.obs,wind.obs,rh.obs,vp,rain.obs,,
-26.46  (Lat.)
Year    DOY     Solar   T-max   T-min   RelHum  Precip  ET
                MJ/m2   oC      oC      %       mm      mm
2000    1       19.80   23.9    14.8    84      15.4    4.9
2000    2       21.10   25.6    15.5    79      9.2     5.6
2000    3       11.60   19.9    14.4    87      1.0     3.1
2000    4       19.10   22.0    13.2    82      0.0     4.8
2000    5       18.50   24.6    14.9    77      0.0     5.1
2000    6       13.30   24.4    15.4    89      25.9    3.6
2000    7       25.10   26.4    15.6    80      0.6     6.6
2000    8       24.00   23.6    15.4    83      0.0     5.6
2000    9       20.30   24.0    13.0    83      0.0     5.0
2000    10      24.60   27.7    14.4    71      0.0     7.4
2000    11      32.70   26.5    11.6    60      0.0     8.4
2000    12      30.60   28.6    13.0    70      7.7     8.6
2000    13      9.30    14.6    11.5    88      2.0     2.3
2000    14      9.30    16.7    9.0     88      11.2    2.5
2000    15      8.70    18.6    10.9    91      63.8    2.3
2000    16      8.40    16.7    12.0    94      17.0    2.0
2000    17      19.40   18.9    11.2    76      0.3     5.0
2000    18      32.60   20.4    6.9     65      0.0     7.2
2000    19      32.40   21.6    7.5     63      0.0     7.1
2000    20      32.80   24.5    5.8     57      0.0     7.9
2000    21      30.30   26.6    9.2     60      0.0     7.9
```

Figure 3.6: Weather Data Input Format for Hybrid-Maize.

3.4 Selection of soils

The GYGA protocol requires that soil characterizations providing parameters for modeling should be as locally relevant as possible. For this Yg study for South Africa, the best available soil data for use in the crop growth model was obtained from the Africa Soil Information Service (AfSIS) (ISRIC—World Soil Information, 2013) database. This data was processed to determine the 3 dominant soil types per buffer zone. The process for determining soil types is shown in Figure 3.7.

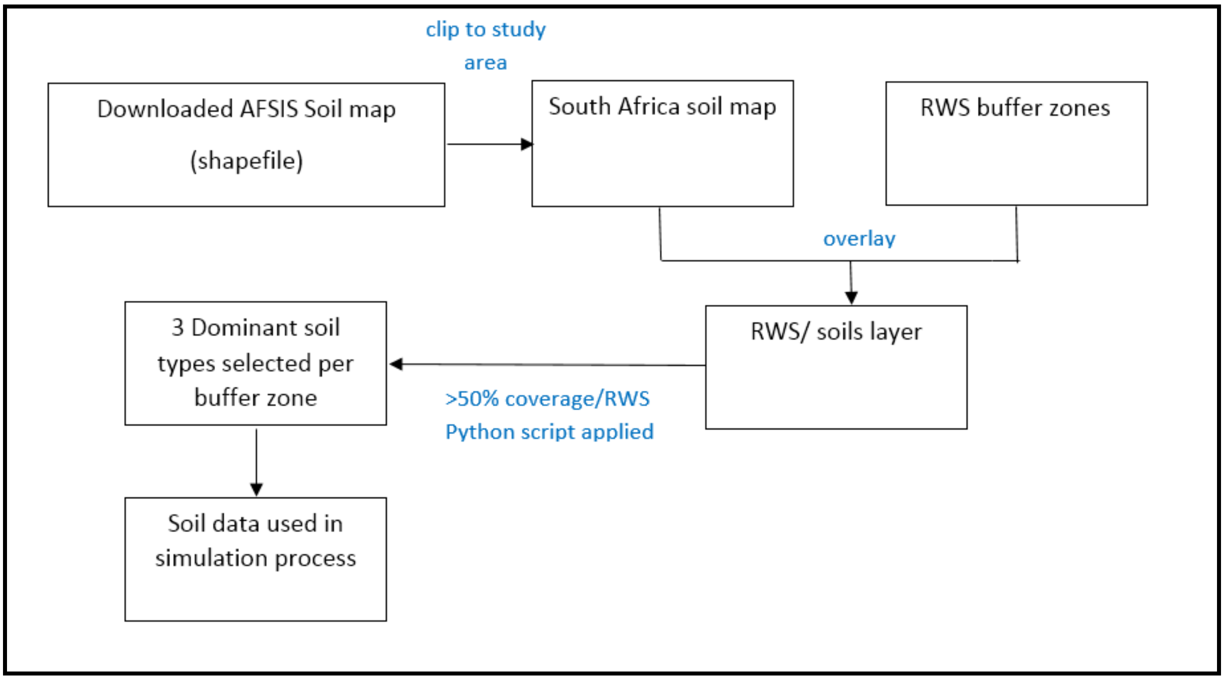


Figure 3.7: Process of selecting soils and obtaining parameters for each RWS buffer zone.

The soil selection process was done in a GIS environment, using ArcMap 10.2.1 (ESRI, 2013). The AfSIS soil data were downloaded as a shape-file. The file was clipped to remain with just the soil data for RSA, which was then overlaid with the RWS buffer zones shape file, and the relative proportions of the 3 dominant soil types within each buffer zone were determined (Table 3.1) using a Python script. The rooting depth, soil texture and bulk density for each of these soil types, which were used in the simulation of Yg, were extracted.

Hybrid-Maize requires data for the rooting depth, plant available soil water (PASW) and the slope (drainage class) of the selected soils in order to simulate water-limited yields. The rooting depth and drainage class were obtained from the meta-data that accompanied the soil maps. The PASW was calculated internally by the Hybrid-M based on the soil texture for each layer.

Table 3.1: Relative proportions of the 3 dominant soils per buffer zone as per the AfSIS soil characterization.

RWS Buffer	Soil Type	Relative Share (%)
Bethal	Sandy loam	42
	Silt loam	29
	Loam	28
Bethlehem	Silt loam	49
	Loam	29
	Sandy loam	22
Bloemhof	Sandy	61
	Sandy loam	31
	Loam	8
Ermelo	Clay	41
	Clay loam	30
	Silt loam	29
Mafikeng	Clay loam	47
	Sandy loam	29
	Silt loam	24
Potchefstroom	Sandy loam	52
	Clay loam	28
	Silt loam	20
Wepener	Silt loam	47
	Loam	30
	Sandy loam	23
Witbank	Sandy loam	59
	Silt loam	28
	Loam	13

3.5 Crop management data

The Hybrid-Maize simulation model requires appropriate combinations of planting date, crop density and plant variety in order to give reliable estimates of Yw. For this Yg study, crop management data were mainly obtained from the Pannar product catalogue for 2017 (Pannar, 2017). Crop specialists and agronomists from the local municipal Department of Agriculture Forestry and Fisheries (DAFF) offices were also consulted to confirm and/or ascertain best agronomic practices for maize production.

Planting rules were defined based on information obtained from the Pannar 2017 catalogue. Sowing of maize in all the RWS buffers is driven mainly by the onset of rainfall, of which dates varied annually, and according to production region. In the cool and temperate eastern production regions, where the Bethlehem, Bethal, Ermelo and Witbank RWSs are located, planting usually commences at the beginning of October to the end of November. In the dryland western production region (Bloemhof, Mafikeng, Potchefstroom and Wepener), planting takes place from mid-November to mid-December. For this Yg study, the planting rule was based on the amount of available water in the soil. The planting date was set to annually coincide with the onset of rainfall in each buffer zone, and the accumulation of soil moisture to levels that allow sowing to take place.

Most of the rain-fed maize crop in RSA is produced at a single cropping intensity. Therefore, for this study, maize-winter fallow maize-winter fallow rotations were assumed across all RWSs as the dominant sequence. For each buffer zone, an ultra-early, early, medium-early, or medium maturing cultivar was selected for use in the simulations for each of the 15 years from 2000 to 2014 depending on the relative proportion of the area on which that cultivar was grown.

The genetic characteristics of the dominant maize cultivars that are grown in each RWS buffer determine the number of days to physiological maturity (black layer), and consequently the growing degree days (GDD). For this yield gap study, GDDs for the cultivars grown in each buffer zone were estimated based on the daily maximum and minimum temperatures for each RWS. Specifically, GDD was calculated using Equation 1:

$$GDD = (Tmax + Tmin/2) - 10 \qquad [1]$$

Where: GDD=>Growing degree days

Tmax=>Daily maximum temperature (ºC)

Tmin=>Daily minimum temperature (ºC)

10=> the base temperature for maize (ºC).

To determine GDD for the growing season in each buffer zone, GDD for all the days of the growing season calculated using Equation 1 were added.

3.6 Simulation

Yield potentials in this study were determined using the Hybrid-Maize simulation model. This modeling software was selected because it combines strengths of existing models pertaining to growth and development functions with the mechanistic formulation of photosynthesis and respiration (Yang *et al.*, 2006), and has been shown to be accurate at estimating water-limited yield potential for maize (Aramburu-Merlos *et al.*, 2015; Meng *et al.*, 2013; Schulthess *et al.*, 2013; Yang *et al.*, 2013). Simulations were carried out for each soil type x cultivar combination for the 15 years from 2000 to 2014.

Model input parameters for each of the 8 buffer zones were entered into separate Microsoft Excel sheets following the template, which is provided by the Hybrid-Maize software developers. Before the simulations were run, test-runs were done by entering the simulation input parameters directly into the Hybrid-M interface. Crop management parameters were edited for each run in order to accommodate those that were at variance with the default settings. Of note were the GDD, which have a default range of between 975 and 1730. For this study, GDD ranged from a minimum of 1100 to a maximum of 2350 and the model was adjusted accordingly.

3.6.1 Model Calibration

The lack of high quality experimental data (field studies in which crops were grown under near-optimal conditions) meant that for this yield gap study, calibration of the Hybrid-Maize model could not be done using independently generated data. The GYGA protocol recommends 2 methods for calibrating crop simulation models. Firstly, there is *elaborate calibration*, whereby data from crops that are grown in experimental stations without influence from yield reducing or limiting factors are used to calibrate.

However, for this study, the second method, which is *"simple"* or *"Phenology only"* calibration was used since the data from experimental stations was not readily available. Most parameters relating to crop growth and cultivar characteristics were obtained from the Pannar 2017 maize production guide (Pannar, 2017). However, the GDDs at all eight RWSs were modified as described under section 3.5.

After calibration, the excel files containing the input parameters were uploaded individually for each RWS buffer zone into the Hybrid-M interface, and the model was run. Simulations were run separately for yellow and white maize for the 15-year period under investigation, within each RWS buffer zone. Results were obtained for each soil type x cultivar combination for each buffer zone.

3.6.2 Quality validation

Once Yw was simulated for each RWS buffer, the simulated values were screened for inconsistencies and inaccuracies by checking for, and eliminating and re-running Yw values that were less than zero.

3.7 Scaling Up

Annual Yw was scaled up from weather station buffer zone, then subsequently to climate zone, provincial and then national level for both yellow and white maize. Within each RWS buffer zone, the 3 dominant soil types were selected for Yw simulation. This means that there were 45 Yw simulations per weather station (15 years x 3 soil types x 2 maize types). The fraction of cropped area that falls under each soil type within each RWS buffer was determined using GIS analysis in ArcMap, and Yw was then weighted by the cropped area proportion per soil type for each year; and this was the weighted average per crop cycle, which was calculated using the equation:

$$Yw\ crop\ cyle = \frac{\sum_{i=1}^{n} Yw\ simulationi\ x\ Soil\ wieighti}{\sum_{i=1}^{n} Soil\ weighti} \qquad [2]$$

Where: n => the number of soil types and

Soil weight *i* => the harvested area of soil type *i*.

The next step was to upscale to DCZ level. This was done by aggregating the Yw at RWS level by weighting its fraction of the crop harvested area within each DCZ. To achieve this, equation 2, shown below, was used.

$$Yw\ climate\ zone = \frac{\sum_{i=1}^{p} Yw\ stationi\ x\ Area\ RWS\ buffer\ zonei}{\sum_{i=1}^{p} Area\ RWS\ buffer\ zonei} \qquad [3]$$

Where: p => the number of RWSs within the climate zone and

Area RWS buffer zone i => the harvested area in buffer zone i.

From the DCZ level, the Yw values were up-scaled to provincial level. This was done by aggregating Yw at DCZ level by weighting its fraction of harvested area of each climate zone per province.

$$Yw\ province = \frac{\sum_{i=1}^{r} Yw\ climate\ zonei\ x\ Area\ climate\ zonei}{\sum_{i=1}^{r} Area\ climate\ zonei} \qquad [4]$$

Where: r => the number of climate zones within each province

Area climate zone i => the harvested area of each climate zone within province i

The last stage of aggregation involved deriving Yw at national level. Each DCZ was weighted by its fraction of national harvested crop area. To achieve this, equation 3 was used.

$$Yw\ country = \frac{\sum_{i=1}^{q} Yw\ climate\ zonei \times Area\ climate\ zonei}{\sum_{i=1}^{q} Area\ climate\ zonei} \qquad [5]$$

Where: q => the number of climate zones within the country and

Area climate zone i => the harvested area per climate zone i.

3.8 Actual yields determination

The GYGA protocol requires that actual yield data for crops under consideration in yield gap studies should be collected at the finest spatial scales that are congruent with the RWS buffers (Van Ittersum *et al.*, 2013; Van Wart *et al.*, 2013; Grassini *et al.*, 2015; Van Bussel *et al.*, 2015). In South Africa, actual maize yield data at the very fine district and municipal scales were not available; therefore, the next best option was considered as prescribed by GYGA. The smallest scale at which actual maize yield data were available for South Africa were at the provincial scale, and was provided by Grain SA (Grain SA, 2017) for the 15 years from 2000 to 2014. Moisture levels were adjusted from 12.5% to 15.5% using simple proportion. The actual yields were therefore defined at 15.5% moisture content.

3.9 Yield gap calculation

The yield gap was calculated as the difference between Yw and Ya at the RWS provincial and national scales (Equation 6).

$$Yg = Yw - Ya \qquad [6]$$

Furthermore, the exploitable yield (Ye), which is the additional yield that is realistically harvestable after considering economic, edaphic and climatic constraints (Lobell *et al.*, 2009) was calculated as:

$$Ye = (Yw \times 0{\cdot}8) - Ya \qquad [7]$$

Lastly, the relative yield (Y %), was calculated as;

$$Y\% = (Ya/Yw) \times 100 \qquad [8]$$

3.10 Synopsis

This chapter discussed the methods that were adopted and applied in this yield gap study. The GYGA approach, which is a bottom-up methodology, is applied to the rain-fed maize crop for RSA. The site selection procedure was explained with regards to selection of RWS buffers. Ya were collected at provincial level, which is a coarser scale than what would have been ideal. Yw was simulated using Hybrid-Maize, and Yg, Ye and Y% calculations were discussed.

CHAPTER 4

RESULTS AND ANALYSIS

4.1 Introduction

This chapter presents the results obtained from the yield gap analysis, and provides some interpretation of these results. Firstly, the study area that was identified through the procedure explained in section 3.2.1, as well as the rainfall weather station (RWS) buffer zones that were delineated following the procedure outlined in section 3.2.2, will be presented. In addition, the soil and crop management data that were collected for use in the simulation model will be presented and discussed. Thereafter, the actual and potential yields that were calculated and simulated at different spatial scales for both yellow and white maize will be presented. Lastly, the yield gaps for both yellow and white maize will be calculated and analysed.

4.2 Study area

The selected study area (RWS buffer zones, see Section 3.2.2), from which data for this research were collected, covered parts of the Mpumalanga (MP), Free State (FS), North West (NW), Northern Cape (NC), Eastern Cape (EC) and Gauteng (GP) provinces. These provinces are highlighted in literature as the main maize producing areas in South Africa (Department of Agriculture, Forestry and Fisheries, 2012). The Spatial Production Allocation Model data (SPAM, 2005) (You *et al.*, 2014), which were an important component of the site selection procedure (Section 3.2.1), also describe these provinces as the main production areas for maize in South Africa (Goldblatt, 2011; Sihlobo & Kapuya, 2014). A visual representation of these harvested areas from which the study area selection process is premised, is given in Figure 3.3.

According to the SPAM (2005) data, the total harvested maize area in the South African designated climate zones (DCZs, Figure 1) added up to approximately 2 336 184 ha, which is above the minimum of 100,000 ha required for a crop to be included in the Global Yield Gap Atlas (Van Ittersum *et al.*, 2013; Van Wart *et al.*, 2013). Furthermore, this area represents 89.6% (Table 4.1) of the total national harvested area for maize in South Africa

(2 656 200 ha), which is above the 50% pre-requisite recommended by the GYGA protocol (www.yieldgap.org).

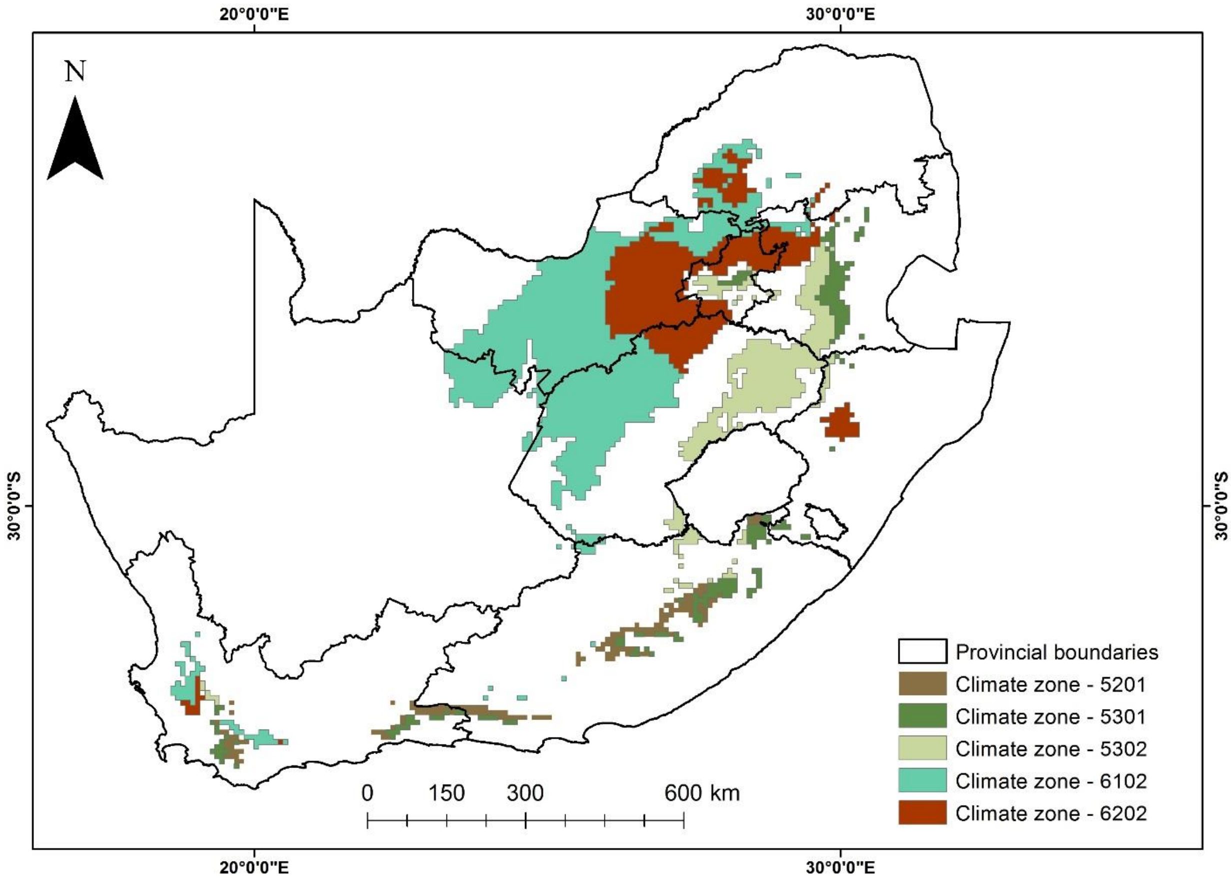


Figure 4.1: Designated climate zones (DCZ) derived from the Global Yield Gap Atlas Extrapolation Domain (GYGA-ED) climate zones (CZ), covering the South African maize production zone and broad cropping regions.

As emphasized in Chapter 3, the GYGA climate zonation scheme provided the foundation for the site selection procedure that was used for this research. According to the output obtained, the main climate zones where maize is grown in South Africa are 5201, 5301, 5302, 6102 and 6202. Table 4.1 shows the harvested maize area within these climate zones. DCZ 6102 accounts for the largest proportion of the DCZs, covering an area of 836 698 ha (35.81% of the total area of DCZs), while DCZ 5301, which is the smallest in size, covers 76 917 ha, or 3.3% of the total area of DCZs.

Table 4.1: Climate Zones of maize production areas of RSA and their total and harvested areas.

DCZ	Maize harvested area (ha)	Total harvested area represented by DCZ (%)
5201	514,951.27	19.75
5301	76,916.77	2.95
5302	430,212.45	16.50
6102	836,698.04	32.09
6202	477,405.46	18.31
TOTAL	**2,336,183.99**	**89.60**

4.2.1 RWS Buffers

The buffer selection procedure explained in section 3.2 was applied to identify RWS buffers from which data was collected. A total of eight RWS buffers were selected within the climate zones where maize production was highest in terms of both harvested area and productivity (Figure 4.2). These buffers spanned over 4 provinces; namely Mpumalanga, North West, Free State and Gauteng. These provinces typically make up the maize production "triangle" in which rain-fed maize is primarily produced in South Africa (Dube *et al.*, 2013; Goldblatt, 2011; Gouse *et al.*, 2006; Pannar, 2017).

An underpinning principle of the GYGA protocol is that the selected climate zones and RWS buffers should represent the best conditions (i.e. weather, soils and cropping system) under which a given crop can be produced (Van Ittersum *et al.*, 2013; Van Wart *et al.*, 2013). With this in mind, parts of the Western Cape province which lie within some of the DCZs were not considered because of the highly saline soils and winter rainfall, which create unsuitable conditions for dryland maize production (Du Plessis, 2003; Scheltema *et al.*, 2015).

Spatial analysis showed that the total area covered by the buffer zones is 1 437 287 ha, which is 55.12% of the total maize harvested area in South Africa. This is above the GYGA recommended minimum of 50% of the total national harvested crop that is required to obtain reliable results (Food and Agriculture Organization of the United Nations, 2015; Gobbett *et al.*, 2017; Grassini *et al.*, 2015).

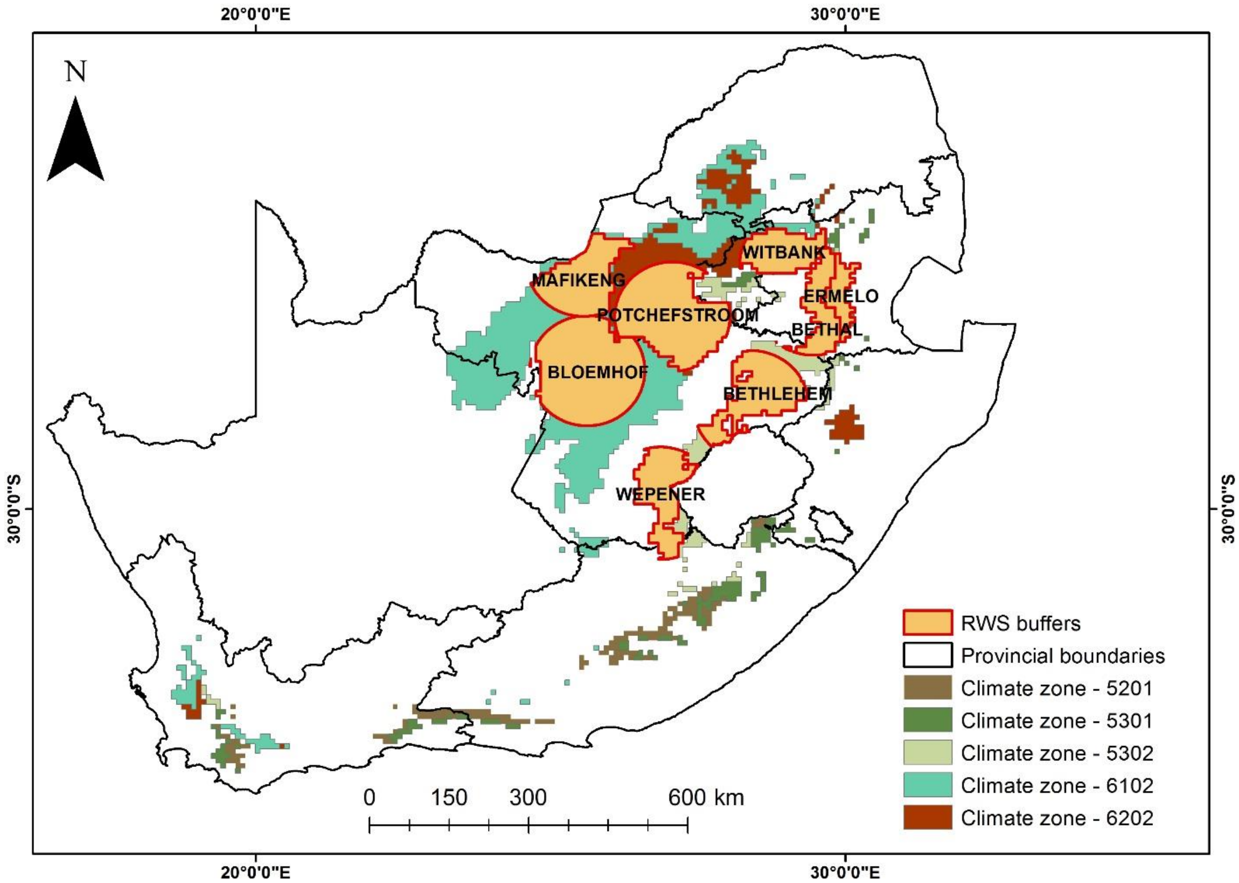


Figure 4.2: Eight selected RWS buffer zones clipped to climate zones as per GYGA protocol. The eight RWS buffers were the smallest spatial units for data collection used in this yield gap study.

To avoid the buffer zones from overlapping, the buffer zones were clipped according to climate zones, so that each buffer zone only represented a single climate zone. However, there can be more than one buffer zone per climate zone, as in the case of DCZs 5301, 6102 and 6202 (Table 4.2). Analyses of the actual harvested area of maize within the buffer zones showed that the RWS buffer with the largest harvested area is Potchefstroom, which has an area of 398 030 ha, while Witbank has the smallest harvested area of 36 820 ha (Table 4.2).

Table 4.2: Maize harvested area and locations of the RWS buffer zones for rain-fed maize in South Africa

DCZ	RWS buffer	Weather station location (Lat/Long º)	Maize harvested area within buffer zone (ha)
5201	Wepener	26.866944/-29.9	37 836
5301	Bethal	29.460556/-26.462222	119 501
	Ermelo	29.983889/-26.497778	58 070
5302	Bethlehem	28.4/-28.249167	232 810
6102	Bloemhof	25.621944/-27.651111	380 447
	Mafikeng	25.560556/-25.808333	173 773
6202	Potchefstroom	27.066944/-26.733056	398 030
	Witbank	29.188611/-25.836944	36 820

4.3 Weather data

Robust simulations of crop growth and yield depend on weather data of sufficient quality (Grassini *et al.*, 2015). To test the robustness of the weather data, the number of missing days of data were checked using a script in R program (R Core Team, 2015). No data for the weather parameters that were used in this study were missing, as shown by the time series plots in Figures 4.3 a to f.

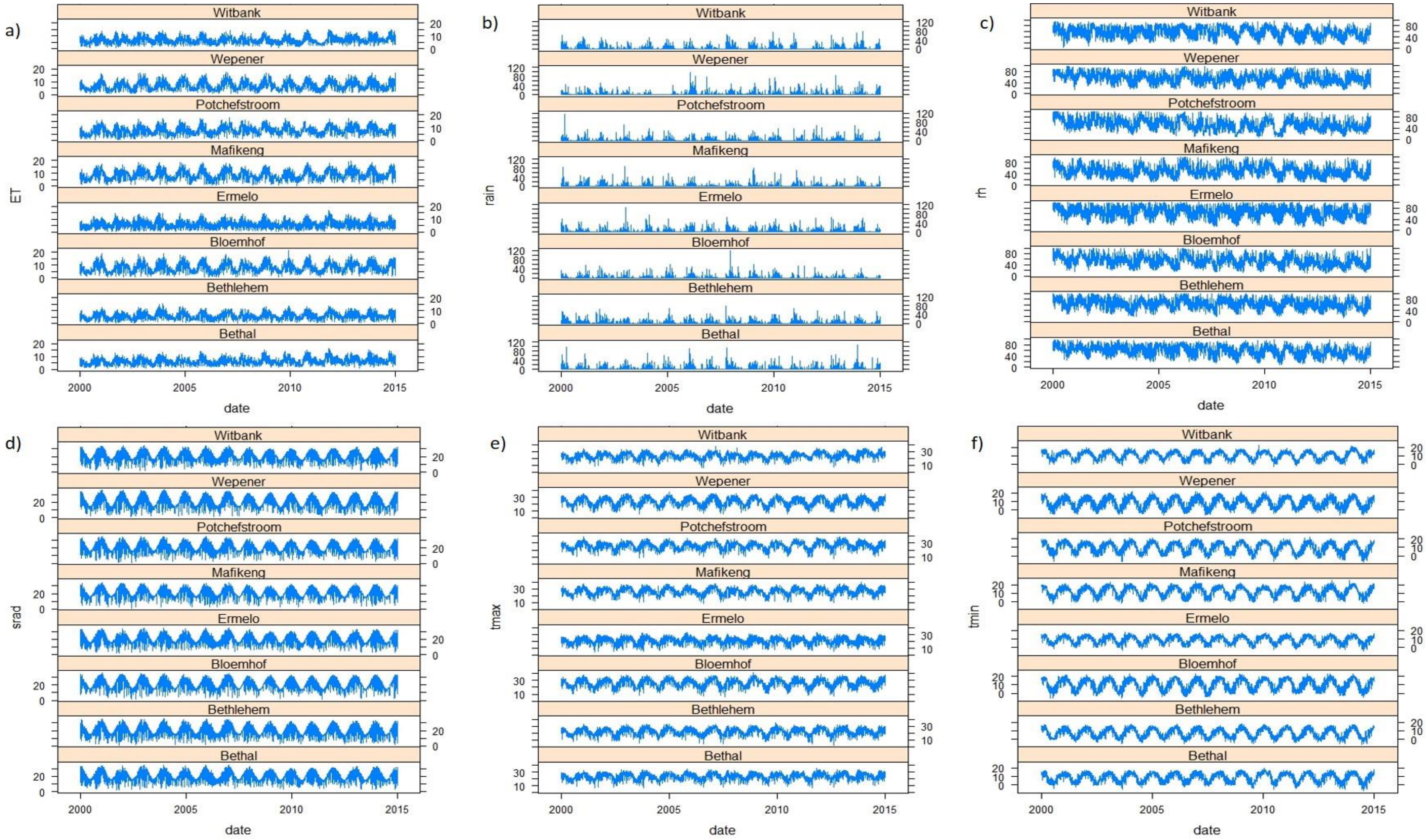


Figure 4.3: Time series plots of the weather variables that were collected at each of the eight reference weather stations selected for this study. The plots give a visual impression of missing data for evapotranspiration (a), rainfall (b), relative humidity (c), solar radiation (d), maximum temperature (e), and minimum temperature (f).

4.4 Soil data

The three most dominant soil types per buffer zone were determined using the AfSIS soil map (www.africasoils.net). The selection of these three dominant soil types was guided by the GYGA protocol as described in Section 3.4, and the results are shown in Table 4.3.

As indicated in Table 4.3, sandy loam and silt loam soils were the most common soil types across the eight buffer zones, accounting for a combined 50% of the soil types. Loam soils accounted for 21% of soils across all RWS buffers, while clay loam soil made up 13%. The effective rooting depth varied from a minimum of 75 cm to a maximum of 150 cm across all soil types and buffer zones. Due to a lack of accurate data, bulk density was pegged at 1.3 gcm^{-3} for all soil types across the RWS buffers, following (Yang *et al.*, 2017) who estimate bulk density for top soil to range between 1.1 g/cm^3 and 1.4 g/cm^3.

4.5 Crop management data

The specific maize varieties that were planted annually per soil type in each of the buffer zones during the period under investigation (2000 to 2014) are unknown. In light of this, the Pannar (2017) guide was used as a reference to select the most dominant maize variety planted in each of the RWS buffers as input for the simulation model. Here, the dominant maize variety planted in each buffer zone was determined by identifying the production region (dryland western or cool / temperate eastern) where each RWS buffer is located (Figure 4.3). DAFF agronomists from the relevant maize growing regions were also consulted to confirm some of the cultivar information. The maize variety from the dryland western region was identified as the medium season variety, while the variety from the cool-temperate eastern region was classified as the short season variety. Parameters that were required from these varieties for the Hybrid-maize model were cultivar maturity and growing degree days to silking and to maturity. The characteristics of these varieties and their corresponding growing degree days (GDDs) are shown in Table 4.4.

Table 4.3: Relative proportions of the three dominant soils per buffer zone following the AfSIS soil characterization

RWS buffer	Soil type	Relative share in RWS buffer (%)	Rooting depth (cm)	Bulk density (g/cm^3)
Bethal	Sandy loam	42	75	1.3
	Sandy loam	29	115	1.3
	Silt loam	28	75	1.3
Bethlehem	Silt loam	49	75	1.3
	Sandy loam	29	75	1.3
	Sandy loam	22	115	1.3
Bloemhof	Clay loam	61	115	1.3
	Sandy loam	31	115	1.3
	Clay loam	8	150	1.3
Ermelo	Sandy loam	41	115	1.3
	Sandy loam	30	75	1.3
	Silt loam	29	75	1.3
Mafikeng	Clay loam	47	115	1.3
	Sandy loam	29	115	1.3
	Clay loam	24	150	1.3
Potchefstroom	Sandy loam	52	115	1.3
	Clay loam	28	115	1.3
	Sandy loam	20	75	1.3
Wepener	Silt loam	47	75	1.3
	Silt loam	30	115	1.3
	Sandy loam	23	115	1.3
Witbank	Sandy loam	59	115	1.3
	Silt loam	28	115	1.3
	Silt loam	13	75	1.3

Table 4.4: Cultivar characteristics for rain-fed maize per buffer zone used in the Hybrid-Maize Model

		Variety Grown		GDD 10°C		GDD to Silking	
Production Region	**Buffer Zone**	**White**	**Yellow**	**White**	**Yellow**	**White**	**Yellow**
Temperate Eastern	Witbank	Short	Short	1375	1375	709	709
Cool Eastern	Bethlehem			1170	1250	625	658
	Ermelo			1100	1100	955	596
	Bethal	Short	Short	1300	1300	678	678
	Wepener			1500	1600	760	801
Western	Bloemhof	Medium	Medium	1975	2350	955	1109
	Mafikeng			1850	1900	904	924
	Potchefstroom			1925	2200	935	1047

As shown in Table 4.4, for both yellow and white maize, the short season variety was the dominant cultivar planted in both the temperate- and cool eastern production regions, while the medium season variety was mainly planted in the dryland western production region. The GDD (10°C) for yellow maize ranged from 1100 in Ermelo to 2350 in Bloemhof. For white maize, GDD (10°C) ranged from 1100 (Ermelo) to 1975 (Bloemhof). Important to note is that the cultivars planted in the selected areas were assumed to remain the same over the 15 years considered for this study, therefore change in hardiness is not accounted for.

4.5.1 Planting dates

In South Africa, planting of rain-fed maize begins with the onset of the first rains, and attainment of sufficient soil water. Optimal planting dates therefore differ according to maize production region. Recommended planting dates for the dry western region are from 15 November to 31 December. For the temperate eastern region, optimum dates range from 1 November to 10 December, while planting in the cool eastern region is recommended to take place from 1 October to 30 December annually (DuPlessis, 2003; Pannar, 2017).

For this study, specific planting dates were determined by the onset of rainfall and the achievement of adequate soil moisture within the RWS buffers. Hybrid-maize was used to identify the specific dates of planting, and these are shown in Figure 4.4. For the cool eastern production region, Ermelo had both the earliest and latest planting dates; 3 October and 28 December, respectively. For the temperate eastern production region, Witbank had the earliest and latest planting dates; 14 October and 13 November, respectively. Planting in the dryland western production regions commenced as early as 13 October (Potchefstroom) until as late as 24 December (Bloemfontein).

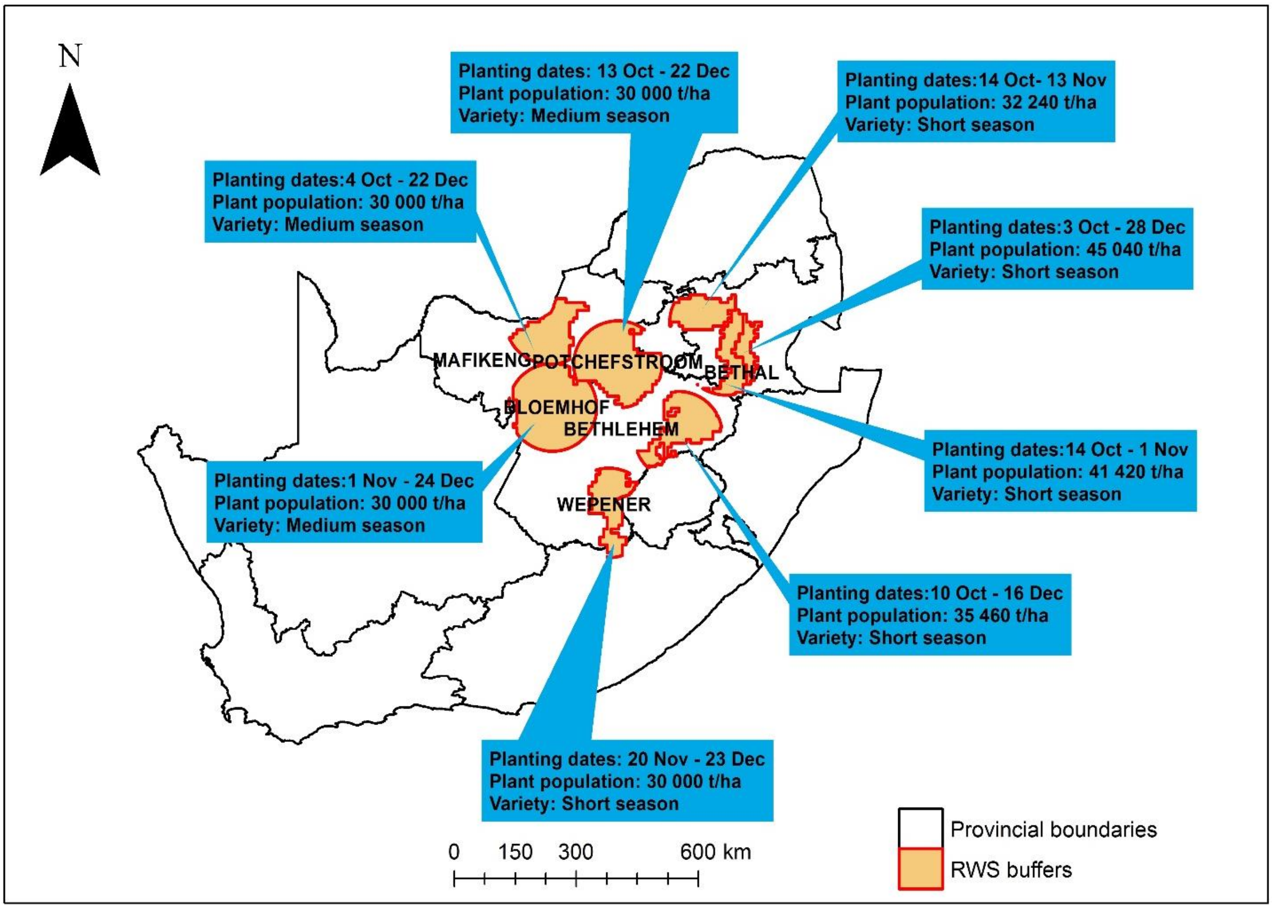


Figure 4.4: Crop management data for yellow and white maize used in the Hybrid-Maize model to simulate water-limited yield potentials for the crop in South Africa.

4.5.2 Plant Population

The planting density, i.e. population of plants planted per hectare, has an effect on yield. Planting should be done in a way that minimises competition between plants while also ensuring that each plant has access to adequate soil water. In the dryland western production region of South Africa, over the 15 years under consideration for this study, plant population averaged 30000 plants per hectare for both yellow and white maize (Figure 4.4). This is in agreement with the Panaar-recommended plant populations of between 25000 and 40000 plants per hectare for yellow maize, and 20000 to 36000 plants per hectare for white maize (Pannar, 2017). On the other hand, in the eastern production region, plant populations ranged from 35000 plants per hectare to 48000 plants per hectare for both yellow and white maize. Figure 4.4 shows the plant populations for each buffer zone in greater detail.

4.6 Actual yields (Ya)

South African maize yield data are available from Grain SA (Grain SA, 2017), who have a dataset of provincial and national annual maize yield figures for the period from 1980 to date. Since maize is harvested at a moisture content of 12.5% in RSA, a correction factor was applied to the GrainSA yields in order to get the yields at 15.5% moisture content, which is the moisture content at which Hybrid-Maize simulates yields.

4.6.1 Provincial actual yields

For this study, Ya data for both yellow and white maize were obtained from Grain SA (Grain SA, 2017) at provincial scale. The averages of the provincial yields and harvested areas for rain-fed maize during the period from 2000 to 2014 are shown in Table 4.5. Based on the Grain SA data (Grain SA, 2017), during the period from 2000 to 2014, white maize was produced on 58.9% more land than yellow maize. The average area under which yellow and white maize is grown in South Africa is 1026000 ha and 1630200 ha respectively. KwaZulu-Natal province achieved the highest yield of 5.70 tha^{-1} for yellow maize amongst the provinces over the 15-year period which were under consideration for this study. This was over an average cultivated area of 42600 ha. Mpumalanga province produced the highest white maize yields of 5.65 tha^{-1} amongst the provinces for the period from 2000 to 2014. Over the same period, Limpopo province produced the lowest average yields of 2.05 tha^{-1} for yellow maize and 2.22 tha^{-1} for white maize.

Table 4.5: Average provincial yields for rain-fed maize for the years 2000 to 2014, corrected to weight at 15.5% moisture (Source: Grain SA, 2017).

Province	**Yellow maize average yield (tha^{-1})**	**White maize average yield (tha^{-1})**	**Average annual harvested area (000ha)**	
			Yellow maize	**White maize**
Western Cape	0.96	0.33	2.5	3.6
Northern Cape	2.77	0.84	44.5	44.6
Free State	4.00	4.59	401.3	654.9
Eastern Cape	4.31	4.16	12	33.6
KwaZulu-Natal	5.70	5.44	42.6	389.1
Mpumalanga	5.63	5.65	281.3	206.7
Limpopo	2.05	2.22	13.9	32.3
Gauteng	5.01	5.05	45.9	67.1
North-West	3.00	3.60	181.9	622.1
TOTAL			**1 026**	**1 630.2**

In terms of harvested area, the Free State province recorded an average of 401 300 ha for yellow maize and 654 870 ha for white maize over the 15 years considered in this study, accounting for the largest harvested areas amongst the provinces. The Western Cape had the lowest average harvested areas of 2 500 ha for yellow maize and 3 600 ha for white maize.

4.6.2 National actual yields

Yield data from Grain SA (2017) indicate that yellow maize was produced on an average area of 1 026 000 ha compared to 1 630 200 ha for white maize (Table 4.5). From this harvested area, mean national yields for yellow and white maize at a moisture content of 15.5%, for the 15-year period under investigation, were 3.71 tha^{-1} and 3.54 tha^{-1} respectively. Analysis of the yield trends shows that the coefficient of variation (CV) for white maize was 16.9%, with a standard deviation of 0.60, against a CV of 19.5% for yellow maize, and a standard deviation of 0.73, thus the yields of white maize were more stable than those of yellow maize during the period from 2000 to 2014 (Figure 4.5).

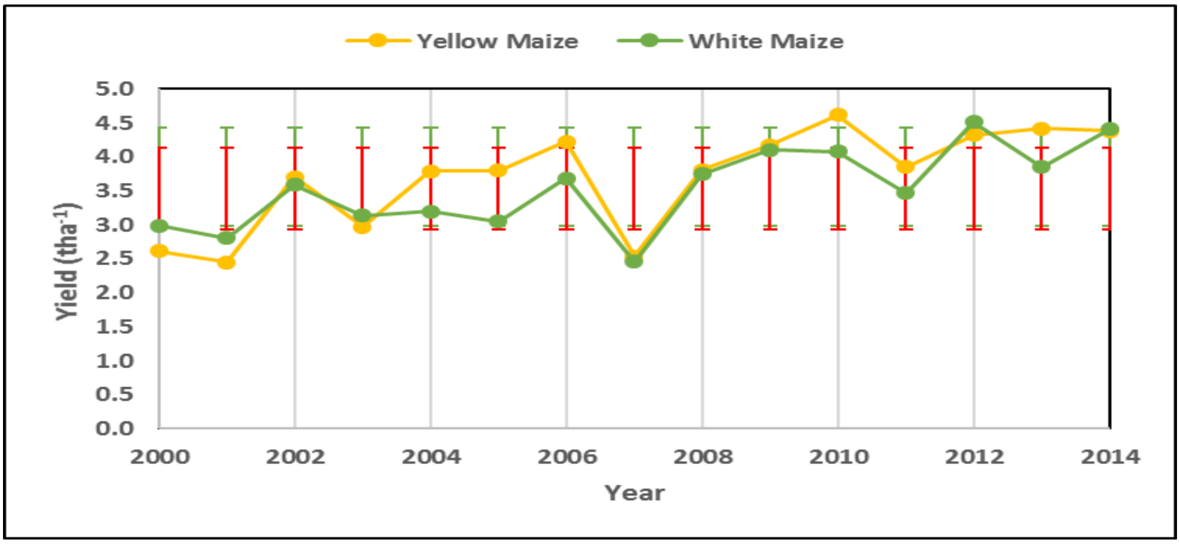


Figure 4.5: Mean annual yields of rain-fed yellow and white maize, corrected to a moisture content of 15.5%. Also included are the standard errors for yellow maize (red error bars) and white maize (green error bars).

For yellow maize, the highest yield obtained in the 15 years under consideration in this study was 4.62 tha^{-1} which was obtained in 2010, and the highest white maize yield was 4.52 tha^{-1} which was harvested in 2012. The lowest yields for yellow and white maize were 2.45 tha^{-1} (2001) and 2.46 tha^{-1} (2007) respectively, which were both drought years (Dube *et al.*, 2013). Detailed annual yields are shown in Table 4.6.

Interestingly, in 2013 the yields for white maize dropped by 14.6% to 3.85 tha^{-1} from 4.52 tha^{-1} achieved in 2012. On the other hand, over the same period, yellow maize yields increased by 2.3% from 4.32 tha^{-1} to 4.42 tha^{-1}. The favourable average gross margin per hectare of yellow maize in 2012 could have driven farmers to commit more resources to the 2013 yellow maize crop in order to maximize yields (Department of Agriculture Forestry and Fisheries, 2012). This resulted in a 1.16 % reduction in planted area for white maize from 1 636 200 ha in 2012, to 1 617 200 ha in 2013, while the area under yellow maize increased by 9.5 % from 1 063 000 ha in 2012 to 1 164 000 ha in 2013.

Table 4.6: Average annual yields for yellow and white maize (corrected to 15.5% moisture content) in South Africa (Source: Grain SA, 2017)

Year	**Yellow maize yield (tha^{-1})**		**White maize yield (tha^{-1})**		**Area planted - yellow maize (000 ha)**	**Area planted - white maize (ha)**
2000	2.11	2.62	2.41	2.99	1280.9	2149
2001	1.98	2.45	2.26	2.80	1111.9	1562
2002	2.99	3.71	2.90	3.59	1174.3	1842.6
2003	2.40	2.97	2.53	3.14	952.5	2232.5
2004	3.06	3.79	2.58	3.20	1001.3	1842
2005	3.07	3.80	2.46	3.05	1110	1700
2006	3.41	4.23	2.97	3.68	567.2	1033
2007	2.05	2.54	1.99	2.46	927	1624.8
2008	3.07	3.81	3.02	3.75	1062	1737
2009	3.37	4.18	3.31	4.10	938.5	1489
2010	3.73	4.62	3.29	4.08	1022.7	1719.7
2011	3.11	3.85	2.80	3.47	954	1418.3
2012	3.48	4.32	3.64	4.52	1063	1636.2
2013	3.56	4.42	3.11	3.85	1164	1617.2
2014	3.53	4.38	3.56	4.41	1137	1551.2
Average	**2.99**	**3.71**	**2.86**	**3.54**	**1026**	**1630.2**

4.7 Water-limited potential yields (Yw)

The water-limited potential yields for both yellow and white maize for the 15 years considered for this investigation were simulated using Hybrid-Maize (Yang *et al.*, 2004) software following the procedure outlined in section 3.6. Potential yields were simulated at RWS buffer zone level, and then up-scaled to climate zone, provincial and national levels.

4.7.1 RWS buffer zone level

Annual water-limited potential yields (Yw) for the years 2000 to 2014 were simulated for the three dominant soil types in each RWS buffer zone. To calculate the mean Yw at the RWS buffer zone level, Equation 2 (Section 3.6) was used to weight the simulated Yw values using the three dominant soil types.

The highest mean Yw for yellow maize at the RWS buffer level was 9.52 tha^{-1} which was recorded at Ermelo, while the highest Yw for white maize was 9.81 tha^{-1} which was recorded at Wepener. On the other hand, the lowest mean Yw values achieved for both yellow and white maize were 4.13 tha^{-1} and 4.70 tha^{-1} respectively, and they were both recorded for Potchefstroom. The mean Yw across all the RWS buffers were 7.22 tha^{-1} for yellow maize and 7.78 tha^{-1} for white maize. The mean Yw for each buffer zone are depicted in Figure 4.6.

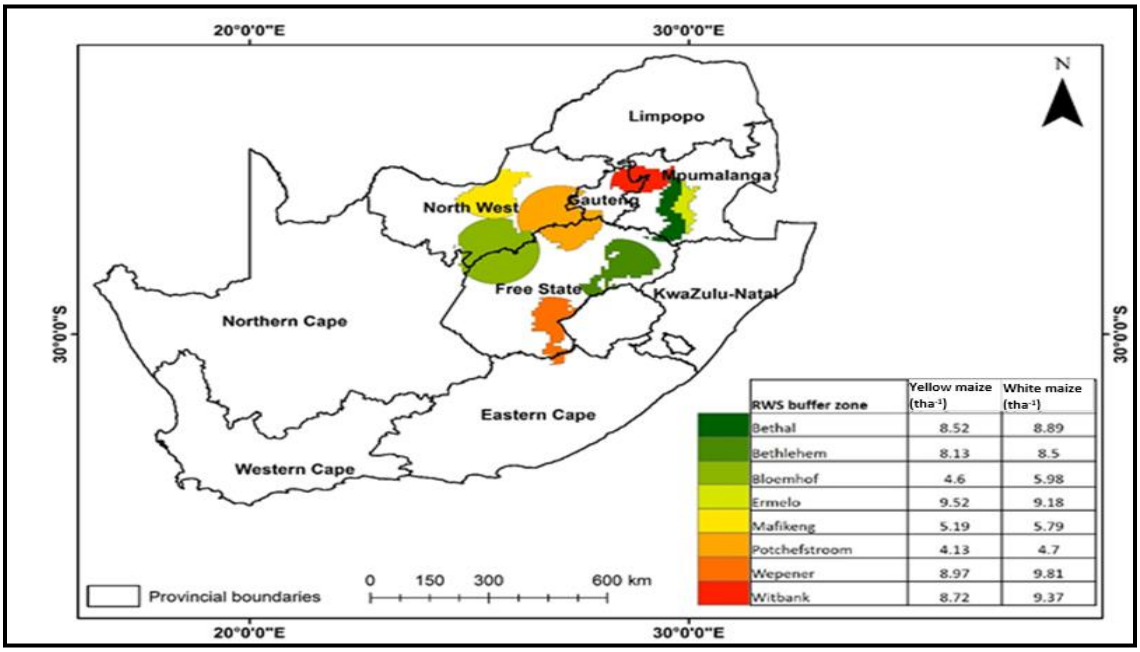


Figure 4.6: Mean Yw values simulated per buffer zone in Hybrid-maize software for yellow and white maize in South Africa.

The simulated Yw values showed the same trends with the yields obtained in the various agro-climatic regions of South Africa, where the highest yields are usually obtained in the cool and temperate eastern production region (Department of Agriculture Forestry and Fisheries, 2012; Grain SA, 2017; Pannar, 2017), which is where the RWS at Ermelo (highest Yw value for yellow maize) and Wepener (highest Yw value for white maize) are located. On the other hand, the lowest yields for rain-fed maize in South Africa are usually obtained in dry western production regions (Department of Agriculture Forestry and Fisheries, 2012; Pannar, 2017), which is where Potchefstroom (lowest Yw values for both yellow and white maize) is located.

4.7.1.1 Stability of simulated yields

Do determine the stability of simulated yields at the RWS buffer zone level over the 15-year study period, the annual trends of Yw values were assessed. The mean Yw for yellow and white maize for the 15-year period under investigation were 6.20 tha^{-1} and 7.31 tha^{-1}, respectively. The coefficient of variation (CV) for yellow maize was 22.1%, with a standard deviation of 1.36, against a CV of 17.16% for white maize, and a standard deviation of 1.25. Therefore, the simulated yields of white maize were more stable than those of yellow maize during the period from 2000 to 2014 (Figure 4.7).

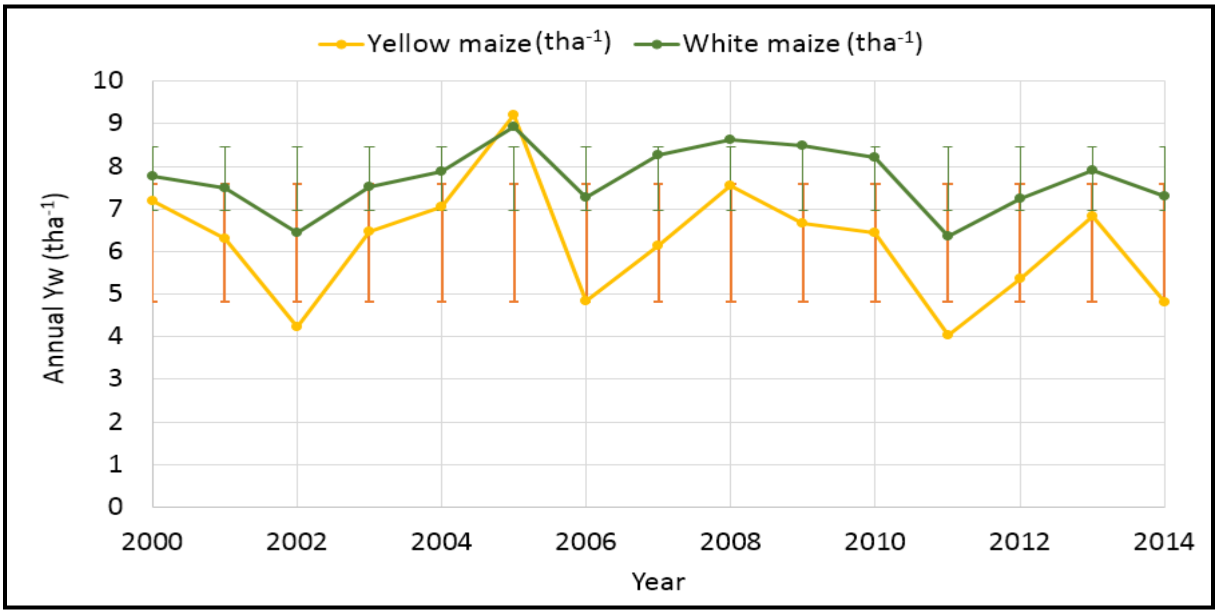
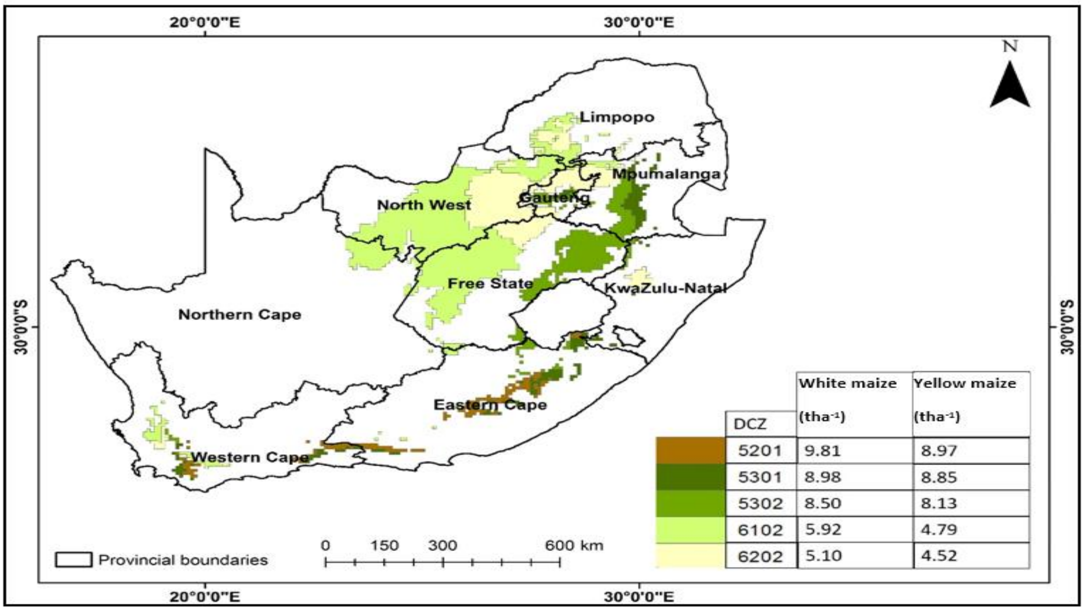


Figure 4.7: Simulated water-limited potential yield (Yw) trends for yellow and white maize for the years 2000 to 2014. Also included are the standard errors for yellow maize Yw (orange error bars) and white maize Yw (green error bars).

4.7.2 Climate zone level

The water-limited potential yields for each climate zone were obtained by upscaling the RWS buffer Yw values by the harvested area using Equation 3 (Section 3.6). The resultant Yw are shown in Figure 4.8.



DCZ	White maize (tha-1)	Yellow maize (tha-1)
5201	9.81	8.97
5301	8.98	8.85
5302	8.50	8.13
6102	5.92	4.79
6202	5.10	4.52

Figure 4.8: Water-limited potential yields per climate zone for white and yellow maize.

DCZ 5201 was found to have the highest Yw values of 8.97 tha^{-1} (yellow maize) and 9.81 tha^{-1} (white maize). The lowest simulated Yw values across the climate zones were 4.52 tha^{-1} (yellow maize) and 5.10 tha^{-1} (white maize) both in DCZ 6202.

4.7.3 Provincial level

The Yw values at provincial level were obtained by calculating the proportion of harvested area per climate zone that falls within each province, then computing the weighted averages of these Yw values using Equation 4 (Section 3.7), and the results are shown in Table 4.8.

Table 4.7: Provincial mean Yw values for yellow maize.

Province	SPAM (2005) harvested maize area (ha)	Weighted mean provincial Yw (tha^{-1}) – yellow maize	Weighted mean provincial Yw (tha^{-1}) – white maize
Eastern Cape	13550	8.19	8.95
Free State	928 867	5.44	6.37
Gauteng	99 521	5.30	5.87
Kwazulu-Natal	71 721	5.50	6.02
Limpopo	40 650	4.72	5.72
Mpumalanga	429 157	6.66	7.14
North West	707 162	4.73	5.74
Northern Cape	42 985	4.79	5.92
Western Cape	5 373	6.22	7.16

As shown in Table 4.7, the Eastern Cape Province has the highest Yw for yellow and white maize, being 8.19tha^{-1} and 8.95 tha^{-1} respectively. However, the maize harvested areas within the Eastern Cape (0.58%) and Western Cape (0.23%) Provinces only accounted for 1% the total maize harvested area in RSA combined; therefore no RWS buffers were selected in these provinces (Van Bussel *et al.*, 2015; Van Wart *et al.*, 2013).

Of the provinces where RWS buffers were identified for this study, the Mpumalanga Province had the highest Yw value for both yellow (6.66 tha^{-1}) and white (7.14 tha^{-1}). Over the 15-year period, Limpopo Province had the lowest mean Yw values of 4.72 tha^{-1} for yellow maize and 5.72 tha^{-1} for white maize.

For the simulated yields, the SPAM (2005) data (You *et al.*, 2014) was used to estimate the harvested area for maize in South Africa. According to the data, the Free State province has the largest maize harvested area of 928 867 ha, while the Western Cape Province has the smallest maize harvested area of 5 373 ha. However, for the areas of interest in this study, Limpopo Province had the smallest average harvested areas for both yellow and white maize, with a total area of 40 650 ha. It is important to note that there is no separation between the rain-fed yellow and white maize for the harvested areas derived from the SPAM (2005) data. Thus, more accurate predictions of Yw at all spatial scales could be achieved with a more specific map that separates harvested areas under yellow and white maize.

4.7.4 National level

To obtain the Yw values at national level, Equation 5 (Section 3.7) was used to upscale from the climate zone level. For yellow maize, the Yw value at national level was 6.41 tha^{-1}, while that for white maize was 7.19 tha^{-1}. The water-limited potential yields for yellow maize at the RWS, CZ and country level are summarized in Table 4.8.

Table 4.8: Summary of simulated Yw values for yellow and white maize at RWS buffer zone, climate zone and national levels, with 'Y' and 'W' representing yellow and white maize respectively.

DCZ	RWS	% of National harvested area covered by RWS buffer	% of National harvested area covered by climate zone	Y_W (tha^{-1})					
				RWS buffer		Climate zone		National	
				Y	W	Y	W	Y	W
5201	Wepener	1.62	19.75	8.97	9.81	8.97	9.81		
5301	Bethal	5.12	2.95	8.52	8.89	8.85	8.98		
	Ermelo	2.49		9.52	9.18			6.41	7.19
5302	Bethlehem	9.97	16.50	8.13	8.50	8.13	8.50		
6102	Bloemhof	16.28	32.09	4.60	5.98	4.79	5.92		
	Mafikeng	7.44		5.19	5.79				
6202	Potchefstroom	17.04	18.31	4.13	4.70	4.52	5.10		
	Witbank	1.58		8.72	9.37				

4.8 Water-Limited Yield Gaps For Rain-fed Maize

The water-limited yield gaps (Yg) for yellow and white maize were calculated at two spatial levels, namely provincial and national level.

4.8.1 Provincial yield gaps (Yg)

At the provincial level, Yg was calculated by subtracting the Ya values from the Yw values for each of the nine provinces (Table 4.9). The largest yield gap (Yg) for both yellow and white maize were observed in the Western Cape Province (5.26 tha^{-1} and 6.83 tha^{-1} respectively). The Western Cape therefore had the highest exploitable yield (Ye) values of 4.02 tha^{-1} and 5.40 tha^{-1} for yellow and white maize respectively. Furthermore, the Y % values of 15.42% and 4.60% for yellow and white maize respectively were the lowest amongst the provinces.

KwaZulu-Natal Province recorded the smallest yield gaps of 1.30 tha^{-1} for yellow maize, and 1.72 tha^{-1} for white maize. The exploitable yields in the province of 0.20 tha^{-1} and 0.52 tha^{-1} were also the lowest amongst the provinces.

When considering the main commercial maize producing provinces of RSA, the North West province has the largest yield gaps of 1.73 tha^{-1} for yellow maize and 2.84 tha^{-1} for white maize. These yield gaps translate to Ye values of 0.79 tha^{-1} and 1.69 tha^{-1} for yellow and white maize respectively. The Y% values of 63.32% and 50.52% for yellow and white maize respectively were the lowest amongst the provinces. On the other hand, the major commercial maize producing province with the lowest yield gaps was Mpumalanga Province. Here, the yield gaps for yellow and white maize were 1.54 tha^{-1} and 1.49 tha^{-1}. Similarly, the exploitable yields were the lowest at 0.21 t/ha and 0.06 t/ha for yellow and white maize respectively. In addition, Mpumalanga province recorded the highest Y% values of 76.88% and 79.11% for yellow and white maize respectively.

Table 4.9: Water limited yield gaps for rain-fed maize (Yg), exploitable yields (Ye) and relative yields for yellow and white maize at provincial level.

Province	Provincial Yw (tha^{-1})		Provincial Ya (tha^{-1})		Yg (tha^{-1})		Ye		Relative Yield (%)	
	Yellow	White	Yellow	White	Yellow	White	Yellow	White	Yellow	White
Western Cape	6.22	7.16	0.96	0.33	5.26	6.83	4.02	5.40	15.42	4.60
Northern Cape	4.79	5.92	2.77	0.84	2.02	5.08	1.06	3.90	57.79	14.16
Free State	5.44	6.37	4.00	4.59	1.44	1.78	0.35	0.51	73.55	72.04
Eastern Cape	8.19	8.95	4.31	4.16	3.88	4.79	2.24	3.00	52.67	46.48
KwaZulu-Natal	5.50	6.02	4.20	4.30	1.30	1.72	0.20	0.52	76.36	71.43
Mpumalanga	6.66	7.14	5.12	5.65	1.54	1.49	0.21	0.06	76.88	79.11
Limpopo	4.72	5.72	2.05	2.22	2.67	3.50	1.73	2.36	43.33	38.76
Gauteng	5.30	5.87	3.47	3.11	1.83	2.76	0.77	1.59	65.47	52.98
North West	4.73	5.74	3.00	3.60	1.73	2.84	0.79	1.69	63.32	50.52

4.8.2 National Yg

The national yield gaps were calculated by subtracting the simulated Yw values for both yellow and white maize from the Ya values. The national Yg values obtained were 2.70 tha^{-1} for yellow maize and 3.14 tha^{-1} for white maize (Table 4.10).

Table 4.10: National yield gaps for yellow and white maize in South Africa.

	National Yw (tha^{-1})	National Ya (tha^{-1})	National Yg (tha^{-1})	Exploitable Yg (tha^{-1})	Relative Yield %
Yellow Maize	6.41	3.71	2.70	1.42	57.88
White Maize	6.68	3.54	3.14	1.80	52.99

Furthermore, the exploitable yield gaps (Ye) and relative yield % at national level were also calculated using Equation 7 and Equation 8 respectively. The relative yield for yellow maize was 57.88%, while that for white maize was 52.99%. For yellow maize, the Ye was 1.42 tha^{-1} and that for white maize was 1.80 tha^{-1}. Assuming that the average harvested areas for yellow and white maize remain at 1 026 000 ha and 1 630 200 ha respectively, this translates to an extra yield of 1 456 920 t of yellow maize and 2 934 360 t of white maize. Therefore, assuming an average maize producer price of R2000 per tonne (Grain SA, 2017; Sihlobo & Kapuya, 2014), this translates to additional income of almost three billion Rands for yellow maize, and almost six billion Rands for white maize annually.

4.9 Synopsis

The purpose of this chapter was to describe and summarise the main results that were obtained for the yield gap study. From the GYGA climate zonation data, five designated climate zones (DCZs) were identified to have the ideal climate regime for maize production in South Africa, and they make up 89.6% of the total maize harvested area. GIS manipulation in ArcGIS (ESRI, 2013) and R programming (R Core Team, 2014) resulted in the identification and selection of eight rainfall weather station (RWS) buffer zones, which were the smallest spatial units from which data for Yw simulation was collected. It was interesting to note that the Western Cape Province, which is not a maize producing region, has areas that lie within high potential DCZs, and could potentially be utilized for commercial maize production. However, limitations stemming from the unsuitable soils and dry summers will need to be addressed. In terms of the GYGA protocol, the climate zonation methodology could be refined to cater for similar scenarios.

The average area under which yellow maize was planted between 2000 and 2014 was 58.9% smaller than that of white maize. The Free State province contributed the largest harvested areas for both yellow and white maize; accounting for 39.1% and 40.2% of the total respective areas.

Hybrid-maize model (www.hybridmaize.unl.edu) was used to simulate Yw across the buffer zones. The largest Yw for both yellow and white maize were obtained in Ermelo, while Potchefstroom had the lowest. The coefficients of variation for both yellow (22.1%) and white (17.2%) maize, indicate that simulated Yw values for the 15-year period under investigation were relatively stable. Water-limited yield gaps (Yw) were up-scaled to climate zone, provincial and national levels using the protocol described in Section 3.7. Interestingly, results show that the highest potential to increase yields would be in the Eastern Cape Province, where potential yields were found to be 8.19 tha^{-1} (yellow) and 8.95 tha^{-1} (white). Finally, results indicate a higher yield gap (Yg) for white maize (3.14 tha^{-1}) than for yellow maize (2.70 tha^{-1}). These results presented in this chapter are scrutinized in greater detail in Chapter 5 of this thesis.

CHAPTER 5

DISCUSSION AND CONCLUSION

5.1 Overview

The aim of this study was to quantify the yield gaps for yellow and white maize in South Africa by adopting the GYGA approach, which is a bottom-up approach that allows global comparisons of results. This study has revealed yield gaps (Yg) of 3.14 tha^{-1} for white maize and 2.70 tha^{-1} for yellow maize, which translated to a relative yield percentage (Y%) of 52.99% and 57.88% respectively. The average annual yields, adjusted for moisture, for the period from 2000 to 2014 were 3.71 tha^{-1} and 3.54 tha^{-1} for yellow and white maize respectively, while simulated water-limited yields for the same period were 6.41 tha^{-1} (yellow) and 6.68 tha^{-1} (white). Although no other studies of a similar nature and scale have been conducted in RSA, the results are consistent with other international studies for example Aramburu-Merlos *et al.*, (2015) and Meng *et al.*, (2013); which gives credence to the methodology applied.

The Y% indicate that there is potential for farmers to increase maize yields from their current level. This could be achieved through the adoption of best management practices. However, one aspect that this analysis does not cover is the socio-economic barriers that hinder maize yields in RSA. This could form part of future research, or might lead to building of models that incorporate these factors when simulating yield potential. Quantitative analysis of food security is important in the formulation of national, regional and international policies. Therefore, these studies should be guided by the principles of accuracy, reproducibility and transparency, which allow scrutiny of the results. The GYGA protocol attempts to incorporate these attributes, though data limitations might be compromising. This study therefore justifies investing in research to investigate the causes and narrow the yield gaps, more-so under future climate change and climate variability. Provinces with high yield gaps, for example the Northern Cape, Eastern Cape and Limpopo (Table 4.10) could be the focus of research to increase the yields on the existing croplands.

Furthermore, judging from the 15 years that were considered for this study, South African yellow and white maize farmers have been producing more than 50% of the water-limited yields (Yw) that could be achieved under optimum management conditions. A few other studies that investigated potential yields as driven by various factors were identified. A study by Verdoodt *et al.* (2003) that simulated radiation-thermal, water-limited and natural production potentials of maize and sunflowers in Guquka, Eastern Cape Province, gave water-limited potential yields of about 4.5 tha^{-1}. Using the actual data from GrainSA (Grain SA, 2017) for Eastern Cape actual yields, this translates to a Y% of 50.3%. However, their simulations used a 10-day weather time step, which has been shown to introduce inaccuracies in the modeling process (Grassini *et al.*, 2015; Van-Wart *et al.*, 2015). The estimated Yw of 6.41 tha^{-1} and 6.68 tha^{-1} for yellow and white maize, respectively, are slightly lower than those that were found in literature as determined by the Agriculture Research Council (ARC) which reported 7.55 tha^{-1} and 8.13 tha^{-1} for yellow and white maize (Agriculture Research Council-Grain Crops Institute, 2016). Yet, in terms of the yellow: white Yw ratios, the ARC trials had a ratio of 7%, while for this study it was 5%.

Despite there being no yield gap studies for maize in RSA found in literature at the time this research was conducted, the Y% estimates in this study are broadly supported by other Sub-Saharan African (SSA) studies. Yg investigations carried out in Kenya reported Y% values of about 31.4% (Van Ittersum *et al.*, 2013b). Bott (2014) also reports Y% of about 40% for maize in SSA, which are attributed to chronic soil infertility. These farming systems are typically intensive in terms of N fertilizer requirements. However, despite the poor agricultural soils that are found in RSA, South African commercial maize production systems are input-intensive, and technologically advanced. Hence poor soils are ameliorated using relevant interventions for example fertilizer and lime. Bott (2014) in his research highlighted a CV of 23% for farm maize yields in RSA. This compares favorably with the 16.9% and 19.5% for yellow and white maize calculated in this study. Indeed, discrepancies amongst these studies (Verdoodt *et al*., 2003; Bott, 2014), could be attributed to the fact that, unlike this study, no distinction between yellow and white maize was made.

On a global scale, the Y% for rain-fed maize in South Africa compares reasonably well with the rain-fed yield potentials of other countries. In Argentina, the Y% for maize was estimated to be 41% (Aramburu-Merlos *et al.*, 2015) while that in China was about 50% (Meng *et al.*, 2013). However, Y% values of up to 80% have been reported in the input intensive, irrigated maize production systems of Nebraska (Van Ittersum *et al.*, 2013b; Grassini *et al.*, 2015; Van Bussel *et al.*, 2015). The Nebraska Y% highlights the potential that exists in the RSA production systems if there was capacity for supplementary irrigation during critical stages like flowering, which is a heat- and water- sensitive growth stage that usually coincides with midsummer dry spells (Du Plessis, 2003; ARC-Grain Crops Institute, 2016). However, there is little capacity in RSA for irrigation of maize at large scales and further research and innovation is needed to be able to supplement water for rain-fed crops in RSA.

The midsummer dry spells point towards the erratic rainfall patterns that are experienced in RSA and most of SSA as a result of climate change and climatic variability (Dube *et al.*, 2013). One way of overcoming this challenge is to invest in irrigation infrastructure. However, installing irrigation in RSA is impractical considering the limited water resources, thus other smart and innovative solutions will be required in order to make efficient use of the scarce water resources, especially in the dryland western maize production region.. The erratic rainfall patterns result in reduced fertilizer use efficiencies (Mueller *et al.*, 2012), thus further research and innovation is needed to overcome this challenge..

The impacts of rainfall variability on maize yields are often related to soil texture (MacCarthy *et al.*, 2017; Mueller *et al.*, 2012; Ray *et al.*, 2013). Agricultural activities in RSA are largely based in the arid and semi-arid regions of the country, where droughts are common (Du Plessis, 2003) and soil acidity compromises grain yields (Agriculture Research Council-Grain Crops Institute, 2016). Coarse-textured and shallow soils normally have low water holding capacity. In light of this, there is a need to reduce runoff and evaporative losses from soil. Conservation agriculture and zero tillage are two practices that are gaining traction amongst South African farmers, and in the long-run, this may help to increase yields while conserving soil water and soil health (Agriculture Research Council-Grain Crops Institute, 2016).

The heavier soils in the eastern production region generally have slow infiltration rates but a good water retention capacity, while the lighter soils in the western production regions have a poor water retention capacity emanating from a rapid infiltration rate. In addition, soils in the major intensive cropping areas have been deteriorating due to continuous cultivation (Dube *et al.*, 2013; Du Plessis, 2003), hence the adoption of soil management practices that promote soil health and quality is pivotal. Annual application of synthetic fertilizers by large-scale commercial farmers has helped to ameliorate the effects of poor soil quality on yields. Compared to the rest of SSA, South African farmers apply in excess of the average doses of elemental nitrogen, phosphorus and potassium (NPK) fertilizer, which gives comparatively higher output (Mueller *et al.*, 2012; Ray *et al.*, 2013). However, assessing the fertilizer use, efficiencies (FUE), particularly N, in order to build sustainable and resilient maize production systems and attain higher yields, thus closing the yield gap, is imperative.

Fertiliser use efficiency is also associated to plant population. In this yield gap study, plant population in the eastern production region were higher than those in the western production region. Plant density determines competition for nutrients and plant growth through light and heat competition (Pannar, 2017). The simulated plant populations ranged from 30 000 plants per hectare in the western production regions to almost 50 000 plants per hectare in the eastern production region of South Africa. However, depending on rainfall availability, plant populations in parts of the western production region could go as low as 19 000 plants per hectare during low rainfall years (Pannar, 2017). This is done to compensate for the low levels of plant available soil water, especially in regions where soil management practices do not encourage reduction of soil water loss.. In irrigated environments, plant populations can be much higher; up to 90 000 plants per hectare (Van Ittersum *et al.*, 2013b), indicating the importance of water availability to achieving high grain yields.

5.2 Assessing Suitability of the GYGA Protocol

According to Grassini *et al.* (2015), the spatial and temporal variation in environmental conditions determine simulated and actual crop yields. They therefore recommend the use of primary data for crop simulations over aggregated or interpolated data. This yield gap study therefore applied the Global Yield Gap (GYGA) protocol (www.yieldgap.org), which employs a “bottom-up” approach, whereby location-specific data are used to estimate yields at larger spatial scales. The GYGA protocol also allows use of the best existing data, which can be improved upon over time. However, in line with the underpinning principles for this protocol, the methodology that is followed should be transparent, reproducible and consistent in data selection. Furthermore, there is a preference for publicly accessible data and the incorporation of local experts to corroborate and/or input data (Van Ittersum *et al.*, 2013b; Grassini *et al.*, 2015; Sadras *et al.*, 2015).

For this South African maize yield gap study, the GYGA climate zonation scheme and SPAM (2005) maps premised the site selection process. Designated climate zones were selected through the process described in section 3.2. The resultant DCZs were 5201, 5301, 5302, 6102 and 6202, which together made up over 89% of the total maize harvested area in South Africa. These DCZs are mainly located in the “maize triangle” of South Africa; being parts of Mpumalanga Province (MP), Free State Province (FS) and North West Province (NW). These areas are described in literature as being the main maize producing areas of RSA, accounting for about 83% of total annual production (Department of Agriculture, 2012). Interestingly, the site selection protocol identified high production potential areas in the Western Cape Province, where DCZs 5201, 5301, 6102 and 6202 were selected. However, because of unsuitable soils (predominantly saline and shallow) and winter rainfall patterns in the WC province, the area was deemed unsuitable for rain-fed maize production and discarded. In light of this, the GYGA site-selection protocol could be refined to cater for rainfall seasonality.

To identify the main harvested areas for maize SPAM (2005) data were used. A point of interest is that SPAM (2005) data for South African rain-fed maize do not distinguish between areas grown under yellow maize and white maize. This might be a source of inaccuracy in estimates of simulated yield. However, SPAM (2005) maps estimate the total maize harvested area to be 2 336 184 ha, which correlates closely with the average cropped area given by GrainSA of 2 656 000 ha. Results obtained from this study indicate that the CZ with the largest maize harvested area was 6102, while DCZ 5301 had the smallest harvested area.

The 8 RWS buffer zones that were selected for this study were at Mafikeng, Potchefstroom, Bloemhof (all western production region), Witbank, Ermelo, Bethal, Bethlehem and Wepener (all eastern production region). The total maize harvested area within the selected buffer zones was 1 427 287 ha, or 55.12% of the total harvested area in RSA, which satisfied the requirement of at least 50% of total harvested area to obtain reliable estimates of Yw (Van Ittersum *et al.*, 2013b; Grassini *et al.*, 2015). The buffer zones with the largest harvested areas were Potchefstroom (DCZ 6202) and Bloemhof (DCZ 6102). These RWS buffers are located in the high-potential western production region of RSA. This area is characterized by large, commercial maize production sites, which generally experience warmer temperatures and receive less rainfall than the eastern production region.

Data that was used for the simulation of rain-fed yield potential was collected at the RWS buffer zone level. The tiered GYGA data selection protocol recommends that “high quality” data be prioritized. Daily weather data of acceptable quality and quantity are required for robust Yw simulations (Grassini *et al.*, 2015). For this study, 15 years of daily weather data, as recommended by Van Wart *et al.* (2013) and Grassini *et al* (2015), were obtained from the South African Weather Service (SAWS) before being processed into formats that were useable in Hybrid-Maize using Python (Python Software Foundation, 2013) and R-programming (R Core Team, 2015). Quality checks applied to the weather data that were used in this yield gap study indicate that the data met the requirements to be considered “reliable” (See Section 4.3). Notably, however, the SAWS weather data do not include data for solar radiation. According to the tiered GYGA protocol, the next best source would have to be considered, which in this case is the National Aeronautics and Space Administration Prediction of Worldwide Energy Resource (NASA-POWER) (NASA-POWER, 2015) database (Grassini *et al.*, 2015). NASA-POWER solar radiation data has been shown to have accuracy of up to 90% when simulating maize yields (Bai *et al.* 2010), thus were adequate for the purposes of this research.

For this yield gap study, the Africa Soil Information Service (AfSIS) database (www.africasoils.net) maps were used to identify the soil properties of the 3 dominant soils in each buffer zone. Due to limited time and resources, and also difficulty in obtaining soil maps that gave the required soil parameters for the specific buffer zones used in this study, soil maps of a finer scale could not be used. Studies done by Gobbett *et al.*, (2016) on Australian wheat yield gaps show that using soil maps of a finer spatial resolution gives more accurate estimates of Yw, and therefore Yg. This is also advisable and applicable to RSA, especially in the eastern production region, where soils are highly variable.

The AfSIS maps present soil data at spatial resolutions of up to 250 m.. The rooting depth and plant-available soil water are two soil parameters that were required by Hybrid-Maize to simulate Yw (Grassini *et al.*, 2015; Van-Bussel *et al.*, 2015) and should be included or derived from the AfSIS data. According to this data, the soil texture in the RWS buffers in the western production region are mainly light (sandy and sandy loam), while those in the eastern production regions are mainly heavy (clay loam and clay). The soil texture influences soil moisture content, which determines plant populations.

The common maize varieties that were planted on the identified soil types in each of the RWS buffers were obtained from the Pannar maize production guide (Pannar, 2017) together with local agronomists from the Department of Agriculture, Fisheries and Forestry (DAFF). As in the case of soil data, better quality crop management data could not be obtained from local commercial maize farmers, or other institutions despite the researcher making efforts to do so. The crop management data used for simulations can therefore be further refined; but for the purposes of this study, the Panaar production catalogue was seen to be sufficient.

The GYGA protocol makes use of the most current maize variety grown. At the time of this study, the short- and medium season varieties were dominant in the eastern and western production regions respectively. The GDD is an important parameter that is determined by the specific maize cultivar and differs amongst seed companies. For this study, neither the specific GDD nor the seed brand that is planted in the RWS buffers was known. The GDDs were then computed according to the varieties selected. More accurate at finer spatial scales about maize cultivars grown in the RWS buffers would greatly improve accuracy of the simulations.

According to GYGA, actual yields data should be collected at the finest spatial scale possible, for example municipalities or districts. For this yield gap study, Ya data was only available at the provincial scale as provided by GrainSA (Grain SA, 2017). Although this was a coarse scale for data collection, it was still acceptable until such a time when data at finer spatial resolutions is released or has been collected.

The GYGA protocol provides appropriate alternatives to challenges due to lack of data. The analysis discussed in this thesis indicates that for the methodology to be relevant, each simulation unit needed to be as representative of the actual agronomic situation as possible, while all underpinning data should be based as much as possible on observed data. However, where data were lacking (for example soil data and national harvested area data in this research), other open-access data sources were used. For this study, parameters like sowing date had large variations in planting dates, which might be due to large annual variations in onset of rains. Therefore, the sowing dates were derived from dynamic simulation, based on amount of rainfall and soil water.

5.3 Contribution of the Study to Sustainable Agriculture

The concept of sustainability broadly refers to the use of resources in a manner that ensures that future generations will be able to benefit from those resources. This forms the premise for sustainable agriculture, which involves the production of crops, livestock and other agricultural goods and services in a manner that is environmentally friendly, socially sound and economically viable. However, this task is made more difficult by the expected world population of 9 billion people by 2050, which presents increased demand and competition for natural resources, including land and water (Nelson *et al.*, 2010a). Simultaneously, there is a need to combat and mitigate the effects of climate change by adopting farming practices that reduce greenhouse gas emissions, and conserve biodiversity.

A pathway that has been suggested to achieve increased yields on more or less the same spatial area of land is agricultural intensification. This yield gap analysis therefore established the levels to which maize yields in RSA could be increased on the existing cropped lands. This means that there will be minimal agricultural encroachment into natural areas and habitats, and maintaining biodiversity. However, the high levels of inputs under this intensification dogma necessitate further investigation into their anticipated harmful effects on the environment.

Economically, farmers would be able to benefit from increased yields. This would only be possible if the producer price of maize doesn't fall below production cost thresholds. South Africa is the biggest exporter of maize in SSA, and increased production would be beneficial to the gross domestic product (GDP). Annually, RSA exports about 2 million tonnes of maize, mainly to other Southern African countries (Sihlobo & Kapuya, 2014). Closing the yield gaps therefore will benefit the farmers and RSA's economy at large.

Socially, maize is South Africa's staple, and over 85% of white maize is used for human consumption (Southern African Grain Laboratory, 2011). Increasing maize production through narrowing the yield gap therefore could ensure a food-secure country. Furthermore, avoiding clearing of large tracts of land for agriculture helps to combat climate change and its effects like droughts and floods, which promotes a secure society. In addition, the anticipated population increase in SSA will need new towns and cities to be developed since Africa is also urbanizing at a fast rate (Nelson *et al.*, 2010a; Nelson *et al.*, 2010b). Closing the yield gaps would relieve competition for land use, and enable sustainable social development to take place.

In light of the above, it is clear that yield gap studies play an important informative role to policymakers and actors within the agricultural, social and economic spheres of society. The information generated could help to shape the development plans of RSA.

REFERENCES

Agriculture Research Council-Grain Crops Institute. 2016. *Maize Information Guide 2016*. Potchefstroom.

Aramburu-Merlos, F., Monzon, J.P., Mercau, J.L., Taboada, M., Andrade, F.H., Hall, A.J., Jobbagy, E., Cassman, K.G., et al. 2015. Potential for crop production increase in Argentina through closure of existing yield gaps. *Field Crops Research*. 184:145–154.

Bai, J., Chen, X., Dobermann, A., Yang, H., Cassman, K.G. & Zhang, F. 2010. Evaluation of nasa satellite-and model-derived weather data for simulation of maize yield potential in China. *Agronomy Journal*. 102(1):9–16.

Baron, C., Sultan, B., Balme, M., Sarr, B., Traore, S., Lebel, T., Janicot, S. & Dingkuhn, M. 2005. From GCM grid cell to agricultural plot: scale issues affecting modelling of climate impact. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences*. 360(1463):2095–2108.

Boogaard, H.L., Van Diepen, C.A., Rötter, R.P., Cabrera, J.M.C.. & Van Laar, H.H. 2014. WOFOST Control Centre 2.1 and WOFOST 7.1.7. 1–133.

Bott, R. 2014. World maize and Its Mega-Environments. *Igarss 2014*. (1):1–5.

Burney, J.A., Davis, S.J. & Lobell, D.B. 2010. Greenhouse gas mitigation by agricultural intensification. *Proceedings of the National Academy of Sciences of the United States of America*. 107(26):12052–7.

Classens, L., Van Ittersum, M., Bussel, L., Wolf, J., Cassman, K., Wart, J., Grassini, P. & Yang, H. 2014. *GYGA Workshop Presentation Notes*. Wageningen.

De-Vries, F., Rabbinge, R. & Groot, J. 1997. Potential and attainable food production and food security in different regions. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*. 352:917–928.

Department of Agriculture, F. and F. 2012. MAIZE MARKET VALUE CHAIN PROFILE Department of Agriculture , Forestry and Fisheries Directorate Marketing Email : MogalaM@daff.gov.za. *Maize Market Value Chain Profile*. 1–43.

Department of Agriculture Forestry and Fisheries. 2012. *MAIZE MARKET VALUE CHAIN*

PROFILE.

Department of Agriculture Forestry and Fisheries. 2014. *Trends in the Agricultural Sector 2013*.

Department of Agriculture Forestry and Fisheries. 2016. A PROFILE OF THE SOUTH AFRICAN MAIZE MARKET VALUE CHAIN: 2016. 1–69.

Dube, S., Scholes, R.J., Nelson, G.C., Mason-D'Croz, D. & Palazzo, A. 2013. South African Food Security and Climate Change: Agriculture Futures. *Economics*. 7(2013–35).

DuPlessis, J. 2003. Maize production. *Resource Centre Directorate Agricultural Information Services*. 1–38.

ESRI. 2013.

Evans, L.T. & Fischer, R.A. 1999. Yield Potential: Its Definition, Measurement, and Significance. *Crop Science*. 39:1544–1551.

Fermont, A.M., van Asten, P.J.A., Tittonell, P., van Wijk, M.T. & Giller, K.E. 2009. Closing the cassava yield gap: An analysis from smallholder farms in East Africa. *Field Crops Research*. 112(1):24–36.

Fischer, R.A. 2015. Definitions and determination of crop yield, yield gaps, and of rates of change. *Field Crops Research*. 182:9–18.

Food and Agriculture Organization of the United Nations. 2015. *Yield gap analysis of field crops: Methods and case studies*. Rome.

Global Yield Gap Atlas. 2015. *Global Yield Gap Atlas*. [Online], Available: http://www.yieldgap.org/web/ [2016, June 10].

Gobbett, D.L., Hochman, Z., Horan, H., Navarro Garcia, J., Grassint, P. & Cassman, K.G. (in press). Yield gap analysis of rainfed wheat demonstrates local to global relevance. *Journal of Agricultural Science*. accepted.

Gobbett, D.L., Hochman, Z., Horan, H., Navarro Garcia, J., Grassini, P. & Cassman, K.G. 2017. Yield gap analysis of rainfed wheat demonstrates local to global relevance. *The Journal of Agricultural Science*. 155(2):282–299.

Godfray, H.C.J., Beddington, J.R., Crute, I.R., Haddad, L., Lawrence, D., Muir, J.F., Pretty, J., Robinson, S., et al. 2010. Food Security: The Challenge of Feeding 9 Billion People. *Science*. 327(February):812.

Goldblatt, A. 2011. Agriculture: Facts and Trends South Africa. *World Wide Fund for Nature (WWF)*. 2–26.

Gouse, M., Pray, C., Schimmelpfennig, D. & Kirsten, J. 2006. Three seasons of subsistence insect-resistant maize in South Africa: Have smallholders benefited? *AgBioForum*. 9(1):15–22.

Grain SA. 2017. *Maize Production Report: 2016*. [Online], Available: http://www.grainsa.co.za/pages/industry-reports/production-reports [2017, May 07].

Grassini, P., van Bussel, L.G.J., Van Wart, J., Wolf, J., Claessens, L., Yang, H., Boogaard, H., de Groot, H., et al. 2015. How good is good enough? Data requirements for reliable crop yield simulations and yield-gap analysis. *Field Crops Research*. 177(JUNE):49–63.

Hochman, Z., Gobbett, D., Horan, H. & Navarro Garcia, J. 2016. Data rich yield gap analysis of wheat in Australia. *Field Crops Research*. 197:97–106.

International Institute of Tropical Agriculture. 2009. *Maize*. Ibadan. [Online], Available: http://www.iita.org/maize [2016, September 27].

ISRIC—World Soil Information. 2013. *Soil property maps of Africa at 1 km*. [Online], Available: http://www.isric.org/data/soil-property-maps-africa-1-km [2016, May 15].

Keating, B.A., Herrero, M., Carberry, P.S., Gardner, J. & Cole, M.B. 2014. Food wedges: Framing the global food demand and supply challenge towards 2050. *Global Food Security*. 3(3–4):125–132.

Koning, N. & van Ittersum, M.K. 2009. Will the world have enough to eat? *Current Opinion in Environmental Sustainability*. 1(1):77–82.

Licker, R., Johnston, M., Foley, J.A., Barford, C., Kucharik, C.J., Monfreda, C. & Ramankutty, N. 2010. Mind the gap: How do climate and agricultural management explain the “yield gap” of croplands around the world? *Global Ecology and Biogeography*. 19(6):769–782.

Lobell, D.B., Cassman, K.G. & Field, C.B. 2009. Crop Yield Gaps: Their Importance, Magnitudes, and Causes. *Annual Review of Environment and Resources*. 34:179–204.

MacCarthy, D.S., Adiku, S.G., Freduah, B.S., Kamara, A.Y., Narh, S. & Abdulai, A.L. 2017. Evaluating maize yield variability and gaps in two agroecologies in northern Ghana using a crop simulation model. *South African Journal of Plant and Soil*. 1862(December):1–11.

Meng, Q., Hou, P., Wu, L., Chen, X., Cui, Z. & Zhang, F. 2013. Understanding production potentials and yield gaps in intensive maize production in China. *Field Crops Research*. 143:91–97.

Mueller, N.D., Gerber, J.S., Johnston, M., Ray, D.K., Ramankutty, N. & Foley, J.A. 2012. Closing yield gaps through nutrient and water management. *Nature*. 490(7419):254–257.

Murthy, V.R.K. 2003. Crop Growth Modeling and Its Applications in Agricultural Meteorology. *Satellite Remote Sensing and GIS Applications in Agricultural Meteorology*. 235–261.

NASA-POWER. 2015.

Nelson, G.C., Rosegrant, M.W., Palazzo, A., Gray, I., Ingersoll, C., Robertson, R., Tokgoz, S. & Zhu, T. 2010a. *Gerald C. Nelson, Mark W. Rosegrant, Amanda Palazzo, Ian Gray, Christina Ingersoll, Richard Robertson, Simla Tokgoz, Tingju Zhu, Timothy B. Sulser, Claudia Ringler, Siwa Msangi, and Liangzhi You*.

Nelson, G.C., Rosegrant, M.W., Palazzo, A., Gray, I., Ingersoll, C., Robertson, R., Tokgoz, S. & Zhu, T. 2010b. Food security and climate change. *Change*.

Nkamleu, G.B. 2011. EXTENSIFICATION versus INTENSIFICATION : REVISITING THE ROLE OF LAND IN AFRICAN AGRICULTURAL GROWTH EXTENSIFICATION versus INTENSIFICATION : REVISITING THE ROLE OF LAND IN AFRICAN AGRICULTURAL. In Tunis.

Ollenburger, M.H., Descheemaeker, K., Crane, T.A., Sanogo, O.M. & Giller, K.E. 2016. Waking the Sleeping Giant: Agricultural intensification, extensification or stagnation in Mali’s Guinea Savannah. *Agricultural Systems*. 148:58–70.

Pannar. 2017. *Pannar: 2017 Product Catalogue*. Greytown.

Pingali, P.L. (ed). 2001. *CIMMYT 1999 - 2000 World Maize Facts and Trends. Meeting*

World Maize Needs: Technological Opportunities and Priorities for the Public Sector. Mexico City.

du Plessis, J. 2003. Maize production. *Resource Centre Directorate Agricultural Information Services*. 1–38.

Python Software Foundation. 2013. *Python Language Reference, version 2.7*. [Online], Available: http://www.python.org.

R Core Team. 2015. *R: A language and environment for statistical computing*. [Online], Available: http://www.r-project.org/.

Ray, D.K., Mueller, N.D., West, P.C. & Foley, J.A. 2013. Yield Trends Are Insufficient to Double Global Crop Production by 2050. *PLoS ONE*. 8(6).

Rodriguez, D. & Sadras, V.O. 2007. The limit to wheat water-use efficiency in eastern Australia. I. Gradients in the radiation environment and atmospheric demand. *Australian Journal of Agricultural Research*. 58(4):287–302.

Rosegrant, M.W., Paisner, M.S., Siet, M. & Witcover, J. 2001. *2020 Global Food Outlook: Trends, Alternatives, and Choices*. Washington, D.C.

Sadras, V.O., Cassman, K.G.G., Grassini, P., Hall, A.J., Bastiaanssen, W.G.M., Laborte, A.G., Milne, A.E., Sileshi, G., et al. 2015. *Yield gap analysis of field crops Methods and case studies*.

Scheltema, N., Meyer, F., Delport, M. & Gxotiwe, V. 2015. *Adding Value in the South African Maize Value Chain*.

Schoeman, F., Newby, T.S., Thompson, M.W. & Van den Berg, E.C. 2010. *FINAL REPORT to the Food and Agriculture Organisation of the United Nations: SOUTH AFRICAN NATIONAL LAND-COVER CHANGE MAP*. Pretoria.

Schulthess, U., Timsina, J., Herrera, J.M. & McDonald, A. 2013. Mapping field-scale yield gaps for maize: An example from Bangladesh. *Field Crops Research*. 143:151–156.

Setiyono, T.D.D., Walters, D.T.T., Cassman, K.G.G., Witt, C. & Dobermann, A. 2010. Estimating maize nutrient uptake requirements. *Field Crops Research*. 118(2):158–168.

Sihlobo, W. & Kapuya, T. 2014. *Supply and demand scenarios for South Africa's maize market: Looking into the 2015/2016 and 2016/2017 marketing years*.

Southern African Grain Laboratory. 2011. *South African Maize Crop: Quality Report 2010/2011 Season*. Pretoria.

Stackhouse, P.W., Westberg, D., Hoell, J.M., Chandler, W.S. & Zhang, T. 2015. *Prediction Of Worldwide Energy Resource (POWER) - Agroclimatology Methodology - (1.0° Latitude by 1.0° Longitude Spatial Resolution)*. [Online], Available: https://power.larc.nasa.gov/documents/Agroclimatology_Methodology.pdf [2017, April 12].

Taylor, P., Zinck, J.A., Berroteràn, J.L., Farshad, A., Moameni, A., Wokabi, S. & Ranst, E. Van. 2004. Approaches to Assessing Sustainable Agriculture. *Journal of Sustainable Agriculture*. (November 2012):37–41.

The South African Grain Laboratory NPC. 2016. *South African Maize Crop Quality Report-2015/2016 Season*. Pretoria.

Tilman, D., Fargione, J., Wolff, B., D'Antonio, C., Dobson, A., Howarth, R., Schindler, D., Schlesinger, W.H., et al. 2001. Forecasting Agriculturally Driven Global Environmental Change. *Science (New York, N.Y.)*. 292(5515):281–284.

Tilman, D., Balzer, C., Hill, J. & Befort, B.L. 2011. Global food demand and the sustainable intensification of agriculture. *Proceedings of the National Academy of Sciences*. 108(50):20260.

United Nations Department of Economic and Social Affairs. 2015. *World Population Prospects: The 2015 Revision, Key Findings and Advance Tables.* (ESA/P/WP.241). New York.

Van-Bussel, L.G.J., Grassini, P., Van Wart, J., Wolf, J., Claessens, L., Yang, H., Boogaard, H., de Groot, H., et al. 2015. From field to atlas: Upscaling of location-specific yield gap estimates. *Field Crops Research*. 177:98–108.

Van-Ittersum, M.K. & Rabbinge, R. 1997. Concept of production ecology for analysis and quantification of agricultural imput-output combination. *Field Crops Research*. 52:197–208.

Van-Ittersum, M.K., Cassman, K.G., Grassini, P., Wolf, J., Tittonell, P. & Hochman, Z. 2013. Yield gap analysis with local to global relevance — A review. *Field Crops Research*. 143:4–17.

Van-Ittersum, M.K., van Bussel, L.G.J., Wolf, J., Grassini, P., van Wart, J., Guilpart, N., Claessens, L., de Groot, H., et al. 2016. Can sub-Saharan Africa feed itself? *Proceedings of the National Academy of Sciences*. 113(52):14964–14969.

Van-Wart, J., van Bussel, L.G.J., Wolf, J., Licker, R., Grassini, P., Nelson, A., Boogaard, H., Gerber, J., et al. 2013. Use of agro-climatic zones to upscale simulated crop yield potential. *Field Crops Research*. 143:44–55.

Van-Wart, J., Kersebaum, K.C., Peng, S., Milner, M. & Cassman, K.G. 2013. Estimating crop yield potential at regional to national scales. *Field Crops Research*. 143:34–43.

Van-Wart, J., Grassini, P., Yang, H., Claessens, L., Jarvis, A. & Cassman, K.G. 2015. Creating long-term weather data from thin air for crop simulation modeling. *Agricultural and Forest Meteorology*. 209:49–58.

Verdoodt, A., Ranst, E. Van & Averbeke, W. Van. 2003. Modelling crop production potentials for yield gap analysis under semiarid conditions in Guquka, South Africa. *Soil Use and Management*. 19:372–380.

Yang, H., Dobermann, A., Cassman, K.G. & Walters, D.T. 2006. Features, applications, and limitations of the hybrid-maize simulation model. *Agronomy Journal*. 98(3):737–748.

Yang, H., Grassini, P., Cassman, K.G., Aiken, R.M. & Coyne, P.I. 2017. Improvements to the Hybrid-Maize model for simulating maize yields in harsh rainfed environments. *Field Crops Research*. 204:180–190.

Yang, H.S., Dobermann, A., Lindquist, J.L., Walters, D.T., Arkebauer, T.J. & Cassman, K.G. 2004. Hybrid-maize - A maize simulation model that combines two crop modeling approaches. *Field Crops Research*. 87(2–3):131–154.

Yang, H.S., Dobermann, A., Cassman, K.G., Walters, D.T. & Grassini, P. 2013. Hybrid-Maize (ver 2013.4). A Simulation Model for Corn Growth and Yield. *Nebraska Cooperative Extension, University of Nebraska-Lincoln*.

You, L., Wood-Sichra, U., Fritz, S., Guo, Z., See, L. & Koo, J. 2014. *Spatial Production Allocation Model (SPAM) 2005 v3.2*. [Online], Available: http://mapspam.info [2016, May 07].