ANALYSIS AND INTEGRATION OF REGIONAL SCALE TEMPERATURE DATASETS INTO A SEASONAL CROP MONITORING SYSTEM

by

TINOMUTENDA T. MAGADZIRE

Submitted in fulfilment of the academic requirements of

Doctor of Philosophy

in Agrometeorology

School of Agricultural, Earth and Environmental Sciences

College of Agriculture, Engineering and Science

University of KwaZulu-Natal

Pietermaritzburg

South Africa

December 2018

PREFACE

The research contained in this thesis was completed by the candidate while based in the Discipline of Agrometeorology, School of Agricultural, Earth and Environmental Sciences of the College of Agriculture, Engineering and Science, University of KwaZulu-Natal, Pietermaritzburg, South Africa. The research was financially supported by the United States Agency for International Development's Famine Early Warning System Network.

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

Signed: Professor M.J. Savage

Date: 10th December 2018

DECLARATION 1: PLAGIARISM

I, Tinomutenda Tamuka Magadzire, declare that:

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

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

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

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

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

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

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

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

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

Signed: Tinomutenda Tamuka Magadzire

Date: 10 December 2018

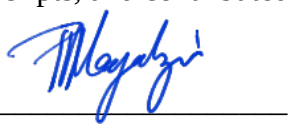
DECLARATION 2: PUBLICATIONS IN PREPARATION

- Magadzire TT, Savage MJ, Funk C. Detection of clouds and modelling of diurnal cloud-free brightness temperature variation using a multi-temporal technique on a single thermal infrared data stream. In preparation.

- Magadzire TT, Funk C, Savage MJ, Peterson S. Use of geostationary thermal infrared data for identifying temperature extremes. In preparation.

- Magadzire TT, Savage MJ, Funk C. Incorporating the effects of temperature into a simple, regional-scale, crop-specific water-balance model and seasonal monitoring system. In preparation.

For each of the research publications reported in this thesis, I designed the study, analysed the data, and drafted the manuscript. My supervisor Professor MJ Savage and co-supervisor, Dr C Funk provided helpful suggestions and guidance in manuscript development, revised the manuscripts, and contributed ideas to improve the study.



Signed: Tinomutenda Tamuka Magadzire

Date: 10 December 2018

ABSTRACT

Populations in excess of 20 million people in southern Africa annually face food insecurity. This number increases appreciably when detrimental seasonal climate conditions lead to widespread reductions in crop harvests. This situation has led to the development of regional-scale crop monitoring systems that incorporate crop-specific water balance (CSWB) models for early detection and warning of impending weather-related crop production shortfalls. Early warning of anticipated reductions in crop harvests facilitates early action in responding to potential crises. One such system, used by the Famine Early Warning Systems Network (FEWS NET) in southern Africa for crop monitoring, calculates the water requirements satisfaction index (WRSI) using a CSWB model. Operationally, CSWB models for calculating WRSI have used a static length of growing period (LGP) to bracket the period over which rainfall and evapotranspiration variations can affect crop yields. In the long term, concerns have been raised by some studies on the impact of rising air temperatures on crop production. There is therefore a need to incorporate the impacts of air temperature on crops directly into food security monitoring systems, in order to improve the accuracy of these monitoring systems in identifying and locating weather-related crop production shortfalls. This study sought to assess the potential improvements that can be introduced to the crop monitoring system in general and the WRSI in particular, by incorporating air temperature data into the CSWB model. To address this objective, daily maximum and minimum air temperature grids derived from a general circulation model reanalysis were used to generate thermal time estimates, expressed as growing degree day (GDD) grids for a maize crop. The GDDs were used to estimate the LGP of maize for each pixel of each summer season (which typically runs between around October and March) from 1982/1983 to 2016/2017 in southern Africa. The variable, temperature-driven LGP estimates compared favourably with LGP values obtained from literature for a few sample

locations. The variable LGP was used to calculate the WRSI for 35 seasons, and the resultant WRSI showed improved correlation with historical yield estimates compared to the static-LGP WRSI, particularly after the farming practice of planting on multiple dates was taken into consideration. Various expressions of WRSI were considered in the analysis, including WRSI calculated assuming planting at the onset of rains, WRSI aggregated from varying number of separate planting dates, including three and six planting dates as test examples, and WRSI calculated using a modified soil water holding capacity to better capture local soil management practices. Historical maize yields for sub-national administrative units from seven southern African countries were correlated with the various WRSI expressions. Gridded GDD data that reflect the accumulated severity of extreme warm temperatures experienced during the crop growth period, referred to as extreme growing degree days, or eGDD, were also noted to have significant correlations with historical maize yields in several southern African countries. A number of variants of the eGDDs were tested, including eGDDs accumulated throughout the crop's growth period, eGDDs that only occurred simultaneously with periods of crop water deficit, eGDDs that occurred during the crop flowering stage, and eGDDs scaled by the severity of crop water deficit. In several areas, the various eGDD expressions indicated higher correlations with yield than any of the WRSI variants indicated. The eGDD parameter showed strong correlations with WRSI, suggesting that the accumulated high temperatures were a reflection of the influence of low rainfall and low soil moisture during episodes of high temperature. More work is required to calibrate and refine the temperature-based monitoring parameters that were developed in this study, at local, sub-national scales. In particular, assumptions of the linearity of maize yield response for the various parameters should be tested. Potential improvements of the combined eGDD-water deficit parameter through the incorporation of prediction coefficients and constants should also be tested. A secondary aim of the study was to explore how readily available temperature-related datasets can be utilized

to derive air-temperature metrics. To this end, satellite-derived thermal infrared (TIR) brightness temperature data were analysed, and a method was developed for identifying cloud cover, while simultaneously estimating cloud-free diurnal brightness temperature curves, using a single TIR satellite channel. The diurnal brightness temperature curves were developed using a sinusoidal and exponential model for daytime and nighttime respectively, utilizing modifications that enabled the curves to be estimated from two known temperatures at any two given times with cloud-free brightness temperature scenes. Comparison of the cloud mask developed in this study with an existing operational cloud mask based on a methodology developed by the EUMETSAT Satellite Applications Facility for Nowcasting gave an accuracy of 85.4%, when the operational method was considered as truth in a confusion matrix analysis. Situations were identified in which the different cloud detection methods showed superior performance, and could therefore complement each other. A statistical method was also developed for calibrating the cloud-free brightness temperatures to station-observed 2-m air temperatures using relationships between the means and diurnal temperature ranges of the two datasets. This enabled the identification of periods of occurrence of extreme warm air temperatures with a coefficient of determination of 0.91, and demonstrated the potential for the usage of TIR data for generating estimates of useful air temperature metrics. The efficiency of the algorithms that were used for simultaneous cloud masking and generation of cloud-free brightness temperature should be improved, in order to enable the methodology to be scaled up to a regional or global gridded level of analysis. Further work for improving operational gridded air temperature datasets by combining station-observed temperature data, modelled data from global circulation models, satellite-derived modelled cloud-free brightness temperature data and cloud masks is recommended.

ACKNOWLEDGMENTS

I am full of gratitude as I think of all the people who invested a lot of time in assisting me start and finish this thesis. Family, friends and colleagues assisted me in various ways, and I am grateful to all of them. I would like to make special mention of the following:

To my supervisor Professor MJ Savage, I am deeply indebted for his tremendous support and guidance, from the time I arrived on campus, all throughout this PhD journey: the guidance, direction, mentorship, knowledge and patience are deeply appreciated. The experience I gleaned from him has been truly invaluable.

To my co-supervisor Dr Chris Funk for his guidance, direction and support, I am deeply grateful. His advisory role, going way back to my Masters days, also helped me maintain a healthy work-school-life balance and assisted me to see this thesis through to the end in a sustainable way.

I am *deeply* privileged to have had them both as my supervisor and co-supervisor. They held many long discussions with me that helped to stimulate ideas. Their comments and reviews tremendously improved my scientific writing and analytical skills. My deepest thanks.

I am very grateful for the funding assistance from UKZN and FEWS NET.

I am also very grateful to Jim Verdin, Jim Rowland, Mike Budde and Greg Husak for the professional support that enabled me to undertake this project.

I extend my sincere gratitude to Mrs. Meryl Savage for her assistance and support during this PhD.

My heartfelt thanks to the Funk family for their hospitality during my week at UCSB. I learnt a lot from them.

Thanks to Dr Alistair Clulow for the excellent tips, guidance and reference materials.

My gratitude is also expressed to Professor Desmond Manatsa and Dr Nsadisa Faka for the reviews, references and encouragement.

Thanks to Dr. Maxwell Mudhara for introducing me to UKZN.

My fellow graduate students, in particular Jonathan Pasi, Sheldon Strydom, Evidence Mazhawu and Nicholas Mbangiwa. I am grateful for their guidance as I learnt the ropes at UKZN.

Pete Peterson, Nick Novella, Laura Harrison, Farai Marumbwa, Seth Peterson, Libby White, Tembalami, I am grateful for the assistance with obtaining data, review of papers etc.

The late Dr Elijah Mukhala and Dr Bernard Manyena for the assistance with sourcing papers, for the recommendations, and for the encouragement. I am saddened that they could not see the fruits of the time investment they made in me. May their souls rest in peace.

I am grateful to an innumerable number of technical people who have made it a part of their lifestyle to share their knowledge on the internet: on blogs, youtube tutorials, research papers, in program code… not only did their generosity help me academically, but they taught me the value of freely sharing knowledge, I am a beneficiary… I will carry the torch and pass it on.

My *vazukuru*, Dr and Mrs Tsvuura, thank you for the tremendous support and hospitality while in KZN.

Baba na Mai Unesu… what can I say? Your hospitality on my way, on those long trips to UKZN, and your mentorship and encouragement… thank you.

To Sekuru and Gogo Nyadenga, thank you for the encouragement and support which is deeply appreciated, and transformational. You have taught us to value education, and we are grateful for the investment you have made in us and continue to do, through your continuous encouragement.

To the Plot 5 crew and all your families – a big thank you. Your encouragement was very encouraging (kikiki)! I learnt so much from you. Baba, because of watching you, I have learnt to try, no matter how insurmountable the task seems. Mhai, I continually remember the words you would say to me in high school admonishing and encouraging me to study: "mwanangu, TV haiperi" – I am realizing more and more the wisdom of your words. Thank you both for setting me on this academic journey 35 years ago, and for encouraging me not to quit high school!

To my wife and best friend, Petronella, and to my wonderful children Tambiraishe, Munashe and Joanna – you guys are awesome! Thank for your unwavering support and encouragement throughout this oft difficult process. Thank you for believing in me. Thank you for releasing me for a season so that I could undertake this degree. We endured many days with me away from home. Thank you for your understanding and patience and love. I love you guys. I could not have done this without your support. Thank you.

Jesus, my Lord, Saviour and Enabler. I now know that "with God all things are possible" (Matt 19:26). You gave me new strength when mine own was gone, hope when mine was lost; new ways of stretching my limits beyond what I dreamed possible. You sustained me. And You have connected me with these wonderful and amazing people, and many others I did not mention here. With all my heart, thank You.

TABLE OF CONTENTS

..... Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

1.1 Rationale

Food insecurity negatively affects large populations in southern Africa, with millions of people assessed to be food insecure each year. Between 2011 and 2017, the lowest reported number of food insecure people for thirteen countries in the Southern African Development Community (SADC) region was over 22 million (SADC 2017). This number peaked at 40 million in 2016 after an El Niño-induced drought caused extensive crop failure and water shortages across the region. The Food and Agriculture Organization (FAO) of the United Nations (UN) identified four pillars of food security, being availability of food, access to food, utilization, and stability (FAO 2009). The availability pillar is primarily related to food production, and in southern Africa, maize production, due to its importance as the staple crop in most southern African countries.

Several international and governmental organizations, including the Famine Early Warning Systems Network (FEWS NET), the World Food Programme (WFP) of the UN, the European Commission Joint Research Centre's Monitoring Agricultural Resources (EC JRC MARS) activity, and Ministries of Agriculture for many African governments, help to promote food security by routinely monitoring crop conditions during the agricultural season. These organizations provide early warning of potential or impending episodes of food insecurity, thereby facilitating pre-emptive early action to minimize future food shortages in affected areas. The seasonal monitoring activities undertaken often rely on field assessment reports and agrometeorological indicators derived from in-situ, satellite and modelled data. The process enables analysts to assess crop conditions and generate estimates of expected relative crop production outcomes. Areas where crop production shortfalls may occur can therefore be quickly

identified, ideally well in advance of rigorous crop surveys. The importance of this process was demonstrated during the 2015/2016 El Niño-induced drought, when several agencies utilized available data to issue alerts that culminated in the issuance of a $2.4 billion humanitarian appeal (Magadzire et al. 2017).

Global- and regional-scale meteorological datasets are required to run the agrometeorological models that are used for wide-area crop condition assessment. Due to the oftentimes limited or sparsely distributed in-situ meteorological data stations in sub-Saharan Africa (IEG 2013), analysts often use satellite data to operationally estimate meteorological parameters. These operational estimates have often been primarily focused on rainfall (e.g. Xie and Arkin 1996, Grimes et al. 1999, Funk et al. 2015), with less effort having been directed towards operationally estimating daily air temperature. Daily air temperature variation however has a direct impact on crop developmental rate (Cleland et al. 2007), and crop yields are negatively impacted by extreme temperatures (Lobell et al. 2011). Thermal time, usually expressed in units of growing degree days (GDD), is used to model the effect of air temperature on the crop development rate (Wang et al. 2002, Jones et al. 2003, Steduto et al. 2009). GDDs are calculated as the accumulated difference between the daily average air temperature and a base temperature below which no plant growth occurs (Moot et al. 2000). They are a good predictor for crop phenological development rate, as plant growth accelerates with increased air temperature, up to an optimum temperature. Temperature-related datasets with a long historical record are of particular interest in this field, due to the ability to analyse climatic trends, and determine their historical impacts on crop yields.

Geostationary satellite-based thermal infrared (TIR) brightness temperature datasets hold the potential to be processed and converted into moderate-resolution, gridded air temperature

estimates, which would improve air temperature data availability. Statistical methods have already been developed to estimate land surface temperature (LST) from brightness temperature (Duguay-Tetzlaff et al. 2015), and to estimate 2-m air temperature from LST (Janatian et al. 2017). In order for the satellite brightness temperature data to be thus used however, there is a need to screen out clouds and model the cloud-free equivalent brightness temperatures for pixels where cloud contamination occurred. The detection of clouds and subsequent generation of a cloud-dataset also promotes the possibility for incorporating cloud occurrence information into air temperature variability analysis. The relationship between air temperature and clouds has been demonstrated in previous studies (Karl et al. 1993, Dai et al. 1999). The development of methods for cloud detection and brightness temperature estimation using geostationary satellite TIR brightness temperature data is important for climatological analysis. Some of these geostationary satellite TIR datasets go as far back as 1980 (e.g. Knapp et al. 2011), and hold potential for enhancing the ability to understand long term trends in temperature variability.

Much of southern Africa experiences a hot, wet summer season and a cool dry winter season. With some location-specific variation and inter-seasonal variability, the summer season generally occurs between October of one year, and March of the following year, while winter typically falls between April and September (WMO 2015). High temperatures characterize the summer months, especially during periods of low cloud cover when high solar insolation is received. During wet spells, cloud cover can obscure the sky for extended periods, and reduced temperatures prevail. In winter, most areas typically experience cloud free conditions.

Extreme warm temperatures are known to have a negative effect on maize crop production. Maize yields have been noted to decrease inversely with increased duration and severity of

temperatures above 30 ℃, with greater impacts at the crop's flowering stage and during dry periods (Lobell et al. 2011). Efforts to improve the operational detection of extreme warm daily temperatures using commonly available datasets such as satellite-based TIR brightness temperature data will therefore contribute to food security early warning efforts.

One consideration that underlines the exigency of incorporating temperature analysis into operational crop monitoring is the climate change phenomenon, and in particular current and projected increases in temperature (Hansen and Lebedeff 1987, Stocker et al. 2013). In addition to increases in average temperature, the frequency and duration of extreme (warm) air temperature events have also been observed to increase over the last few decades (Panda et al. 2017), and are expected to continue increasing into the future (Stocker et al. 2013). The global increase in temperatures over the last few decades has had a negative effect on maize yields (Lobell and Field 2007). Operationally modelling the effects of observed or estimated temperatures on crop growth and production may further improve the accuracy of food security early warning systems.

The water requirements satisfaction index (WRSI) is one of the main agrometeorological indicators being operationally generated by several international agencies engaged in food security monitoring, including FEWS NET (Senay and Verdin 2003), and the Monitoring Agricultural Resources (MARS) Unit of the Joint Research Centre of the European Commission (Massart et al. 2010). The WRSI is calculated from a crop-specific water balance (CSWB) model that uses rainfall and short grass reference evapotranspiration data to estimate the extent to which a crop's water requirements are met. WRSI is correlated to maize yield in water-limited areas, providing potential for yield estimation in such areas. The inclusion of temperature information in the water balance

analysis, in particular its effect on phenological development and heat stress impacts could potentially increase the accuracy of the WRSI and its utility for food security monitoring.

One of the parameters used in the CSWB model is the length of the growing period (LGP), which is defined as the length of time taken by the crop through all its development stages from planting to physiological maturity. In the operational CSWB model run by FEWS NET, the LGP is treated as a constant (Senay and Verdin 2003), which can be approximated using the crop seasonal rainfall and evapotranspiration cycle (Senay and Verdin 2002). The rate of phenological development for maize and other crops is known to be related in part to the ambient air temperature regime (McMaster and Wilhelm 1997, Trudgill et al. 2005, Cleland et al. 2007). The dependency relationship between the rate of phenological development and temperature is referred to as thermal time, measured in growing degree days. An exploration of the feasibility for using available gridded air temperature data to derive operational thermal time grids could provide useful information for seasonal crop monitoring. Using such thermal time grids to drive the CSWB model could improve the accuracy of LGP estimates, and ultimately the WRSI. Likewise, the development of methods to operationally identify the occurrence of extreme warm temperatures during the crop growing season, as well as their potential implications for crop production, would further improve model-based crop monitoring systems. Such innovations would in turn improve the ability of these monitoring systems to provide early warning of potential weather-related maize yield reductions that could adversely affect food security conditions in southern Africa.

1.2 Research questions

A number of important research questions are posed in this study. Can air temperature data metrics be derived from globally available satellite-derived temperature datasets? Can temperature

information be integrated into a crop-specific water balance analysis to accurately capture the influence of temperature on crop development and production? If so, does inclusion of temperature information into an operational, regional-scale crop monitoring system improve the ability of the system to assess crop yields?

1.3 Aims and objectives

The main aim of this research is to assess the feasibility and potential benefit of incorporating air temperature data into an operational, regional-scale crop monitoring system in southern Africa, based on water balance modelling.

The specific objectives of the work to be undertaken are as follows:

- To develop a method of processing TIR brightness temperature data for cloud removal and generation of clear-sky brightness temperatures using techniques compatible with available long-term TIR brightness temperature data
- To investigate methods of processing clear-sky brightness temperature datasets for the retrieval of air temperature data variables
- To assess the utility of currently available, global-scale, modelled daily air temperature datasets for calculating regional-scale growing degree day grids
- To improve crop-specific water balance model calculations using variable LGP grids derived from air temperature data
- To determine the utility of the analysis of air temperature extremes for crop monitoring at a sub-national level in selected southern African countries

1.4 Thesis outline

The thesis is structured as a set of chapters, with each chapter, excluding Chapter 1 and Chapter 5, comprising a stand-alone peer reviewed publication that is either in preparation or has been submitted to a journal. Chapters 2 to 4 contain most of the standard sections of a journal article, including an introductory literature review, materials and methods, results, discussion, and conclusion sections. Chapter 5 provides some overarching conclusions for the thesis. References for all the chapters have been included as a single section at the end of the thesis. The processing or use of temperature data is a theme that runs through all the chapters, culminating in its usage for food security monitoring in Chapter 4. The remainder of this chapter briefly describes the content of Chapters 2 to 4:

Chapter 2 focuses on the processing of a readily available, medium resolution, satellite-based TIR brightness temperature to remove cloud contamination, and derive a spatially seamless, half-hourly set of clear-sky, cloud-free brightness temperatures. This sets up methods for developing a potentially global dataset that can be related to air-temperatures, which can in turn be directly related to crop phenological development, extreme temperature impacts on crop yields.

In Chapter 3, the cloud-free, clear sky brightness temperature estimates developed in Chapter 2 are analysed, and their relationship with station-observed air temperature data is assessed. A methodology for identifying episodes of extreme warm air temperature using brightness temperature is developed, thereby enhancing the utility of brightness temperature in food security monitoring, given to the adverse effect of extreme warm air temperatures on crop yields.

Chapter 4 focuses on the integration of air temperature data directly into CSWB models. The generation of GDD grids from air temperature data is described, and the use of the GDD grids in

a CSWB model to determine the LGP. The variation in LGP over a 35-year period is explored, and the performance of the resultant modified CSWB outputs is compared to the original WRSI values which do not use the improved LGP analysis. Parameters that provide an indication of the impact of extreme warm air temperature on crops are also developed and evaluated in this chapter.

Chapter 5 synthesizes the research work described in Chapters 2 through 4, highlighting how the findings from the different components contribute to the overarching research goal. Special emphasis is placed on the conclusions and recommendations, especially for future research work.

CHAPTER 2: DETECTION OF CLOUDS AND MODELLING OF DIURNAL CLOUD-FREE BRIGHTNESS TEMPERATURE VARIATION USING A MULTI-TEMPORAL TECHNIQUE ON A SINGLE THERMAL INFRARED DATA STREAM

2.1 Abstract

A system was developed for identifying clouds and modelling diurnal cloud-free brightness temperature variations using only a single thermal infrared (TIR) band. Clouds regulate both local and global temperature by altering the earth's radiative balance. Their identification at a high spatial and temporal resolution, together with sub-hourly cloud-free diurnal temperature variation estimates, can aid local scale temperature analysis. Using a sinusoidal and exponential function for daytime and nighttime respectively, cloud-free diurnal temperature curves (DTC) of half-hourly TIR brightness temperature were derived for four point locations in southern Africa. Clouds were identified by comparing the difference between the satellite-observed and cloud-free modelled temperatures with pre-defined thresholds. A number of rules were also used to further discriminate between clouds and cloud-free scenes. The process of DTC modelling and cloud identification was repeated iteratively to progressively refine the outputs, improving both the cloud mask and the cloud-free DTC estimates. Accuracy assessments identified a mean absolute error (MAE) of 1.1 ℃ for cloud-free observations, with 83.5% and 77.1% of all daytime and nighttime DTCs respectively having an MAE of less than 2 ℃. When compared with a cloud mask produced by the EUMETSAT Satellite Application Facility for Nowcasting (SAFNWC), the DTC-based cloud mask had an accuracy of 85.4%. The superiority of each method over the other varied depending on the situation, and combining the two techniques offers potential for improving the

DTC modelling. The DTC method's use of a single TIR satellite band provides a uniform day-night cloud masking scheme, and avoids the discontinuity between day and night associated with methods that utilize both visible and TIR satellite imagery for cloud masking. The new technique can be modified by the use of climatologically-based thresholds to make it more widely applicable to different climate regimes.

Keywords: *cloud mask; diurnal temperature; southern Africa; TIR; DTC*

2.2 Introduction

As clouds alter the earth's radiative balance, they play an important role in regulating climate, especially air and surface temperature at various scales from local to global. At a global scale, low-level clouds increase the planetary albedo and reduce the global average temperature (Ramanathan et al. 1989). Karl et al. (1993) indicated a strong inverse relationship between diurnal air temperature range (DTR) and cloud cover. Dai et al. (1999) demonstrated that clouds, and in particular low-base clouds, can cause large reductions in DTR of up to 50%. On a longer time-scale, there is evidence that global temperatures have risen over the last few decades, and are expected to continue rising (Stocker et al. 2013).

High temperatures are known to have various negative impacts on human health and agriculture (Du Preez et al. 1990, Lobell et al. 2011, Wang et al. 2013). Datasets that can provide improved monitoring and analysis of temperature-related phenomena, such as cloud occurrence, are therefore important for enhancing the understanding of temperature variability at a local scale. Such datasets include satellite-based thermal infrared (TIR) data. Satellite TIR data can be used to derive information on cloud occurrence and estimate the equivalent cloud-free brightness

temperature (BT), once the timing and location of cloud cover has been determined. This is a precursor to the retrieval of land surface temperature (LST) which has uses in ecological and agricultural applications. Although various products are available for cloud detection, some of these have a reduced spatial and temporal resolution at global level (Rossow and Garder 1993, Rossow and Schiffer 1999) and many are not operationally available. Some existing operational methods of cloud detection for geostationary satellites use both visible and infrared satellite bands for daytime and nighttime (Derrien and Le Gléau 2005). The use of visible channels in the cloud mask can lead to discontinuities and reduced accuracy in the cloud mask at dusk (Trepte et al. 2005, Derrien and Le Gléau 2007, Hocking et al. 2011). Zhuge and Zou (2016) used four infrared (IR) bands from the Advanced Himawari Imager sensor, which became operational in 2015, and is thus unavailable for historical analysis. An operational cloud mask at high resolution using a single consistent infrared satellite band with a long historical record, together with sub-hourly cloud-free diurnal temperature variation estimates, could therefore provide improved analytical inputs supporting climatological and local scale temperature analysis.

Several methods have been developed for cloud detection using satellite imagery (Minnis and Harrison 1984, Derrien et al. 1993, Rossow and Garder 1993, Feijt et al. 2000, Schädlich et al. 2001, Derrien and Le Gléau 2005, Frey et al. 2008). Rossow et al. (1985) compared six methods for cloud detection and noted that a major factor affecting the accuracy of the methods depended on their ability to accurately identify cloud-free conditions, which were the basis for determining cloudy conditions. They further noted that while most algorithms could correctly identify the majority of cloudy conditions, many methods displayed reduced accuracy in cloud identification at the edges of large clouds, low clouds, thin, high-level cirrus clouds, or rapidly varying land surfaces. Rossow and Garder (1993) concluded that many cloud detection algorithms had

increased error when clouds had low contrast from the surface they overlay, either due to optical thinness of the clouds, partial and broken cloud cover, or low-level clouds.

Minnis and Harrison (1984) used a combination of VIS (visible) and IR (infrared) band imagery from the Geostationary Operational Environmental Satellite (GOES) to detect clouds. As a first pass, the VIS band was used to detect areas with cloud-free daytime conditions using a minimum brightness threshold – all radiance values below the threshold were marked as cloud-free. This threshold, for land surfaces, was calculated for different geographic regions using models of reflectance based on the solar and viewing geometry, and calibrated using VIS observations presumed to be cloud-free. The corresponding temperature values for the pixels preliminarily defined as cloud-free (using the VIS approach) in each region were used to define the temperature threshold for cloudy pixels. Although the effect of emissivity and transmittance on measured radiances were acknowledged, they were not corrected for as they did not affect the results of the study. A truncated cosine curve was fitted by regression to the observed cloud-free temperatures to determine the sunset and sunrise temperatures, as well any missing daytime temperatures. Cloud-free nighttime temperatures were estimated from a power function or linear interpolation, using sunrise and sunset temperatures, depending on a predefined set of criteria. Further screening of clouds was done by using: (a) a minimum threshold temperature for land values, and (b) a maximum possible rate of change of temperature between consecutive time periods. They evaluated the accuracy of their technique using two data sources, namely surface observations and satellite photographs. For the surface observations, they used the mean monthly cloud amounts for each station in a region of interest, and calculated the average for the region. This value was compared with monthly mean cloud cover calculated for their method. Over land areas, they found

a mean percentage difference of 4.9% between the two methods, with a tendency for their satellite-based method to under-estimate compared to the ground observations.

Several studies have adopted similar approaches utilizing thresholding techniques on VIS and IR for cloud masking. Rossow et al. (1985) tested several cloud detection algorithms, one of which used a threshold method on both VIS and IR satellite imagery to distinguish cloud-free from cloudy pixels, with thresholds based on minimum reflectance and maximum BT calculated over a 15-day window. Rossow and Garder (1993) used both VIS and IR imagery to construct a series of tests for separation of clear from cloudy conditions. The tests included a spatial variability test which compared a pixel with nearby pixels in its neighbourhood defined by a moving window, and a temporal variability test which compared the BT at a pixel location on a given day with the BT for the corresponding time on the previous or the next day. The presence or absence of clouds was determined using predetermined thresholds. Schädlich et al. (2001) implemented a simple method for cloud detection that used thresholds based on monthly composite minimum VIS and maximum TIR for each 30-minute time slot. Monthly composites derived in this manner were assumed to be cloud-free. Pixels were considered cloudy in cases where the VIS value was above the VIS threshold, and where the TIR threshold fell below the TIR threshold.

Derrien and Le Gléau (2005) described a cloud masking methodology developed by the EUMETSAT Satellite Application Facility for supporting Nowcasting and very short range forecasting (SAFNWC) for use on Meteosat Second Generation (MSG) satellite data. The method required a minimum of four bands from the Spinning Enhanced Visible and Infrared Imager (SEVIRI) radiometer located on the MSG satellite. Several ancillary datasets were also required, including various climatological datasets. A series of tests were applied to the multiple satellite

image bands and combinations of bands, with various thresholds being used in each step of the series, and different tests using different thresholds were applied based on the time of day (daytime, nighttime or twilight) and the location of the pixel (land or sea, desert or not). The thresholds were based on pre-defined relationships of cloud/no-cloud thresholds with: (a) solar and viewing angles, as well as (b) surface temperature and total atmospheric water vapour content (TAWV) determined from numerical weather prediction models, as water vapour strongly attenuates thermal infrared radiation. Derrien and Le Gléau (2005) evaluated the system by comparing a 5 x 5 pixel window of the cloud/no-cloud classification to ground-observed cloud cover measured in octas, and found an accuracy of 91.7% for the method over areas in Europe and North Africa, particularly between 25 and 55° N.

Göttsche and Olesen (2001) described a method for modelling the diurnal temperature curve (DTC) of IR BTs from Meteosat infrared data, using a cosine function and an exponential term to model the daytime and nighttime temperatures respectively. Their algorithm assumed that a cloud screening mask had already been applied to the data, and that all observations used in fitting the model were cloud-free. Model fitting was done using a Levenberg–Marquardt method, which is generally used for fitting non-linear models. They developed maximum and median monthly composites of the BTs and fitted these composites to produce monthly composite DTCs. Their study was limited to the modelling of BTs rather than LST due to computational expense required for application of atmospheric corrections.

Ground-based meteorological data may be considered as ‘truth’ when comparing with satellite-based estimates, but this may not hold in the case of cloud cover. Kotarba (2009) noted that in most studies that compared satellite and ground-based cloud cover estimates, the former were

usually less than ground-based estimates, up to 15% less in the studies they cited. The differences between satellite and ground-based cloud cover estimates can be due to several reasons including differences in perspective between the ground observer and the satellite sensor, parallax error, and potential human observation error (Derrien and Le Gléau 2005). Many cloud cover estimation validation techniques are therefore based on a comparison between two techniques, in which one technique is considered as accurate.

This research explores a methodology for defining a height-sensitive cloud mask using a single TIR satellite image band. In our review of the existing literature on methods that used data from geostationary satellites to detect cloud cover, we did not encounter any methods that used only a single TIR band for cloud detection. Many methods used both visible bands and infrared bands. The visible band has the disadvantage that it is only usable during the daytime, and this creates a discontinuity between daytime and nighttime cloud estimates. The use of a single band for detecting cloud cover has an added advantage: the capability to use such an algorithm with the GridSat B1 infrared Climate Data Record (CDR), a dataset composited from several geostationary satellite sensors, with the TIR channel spectrally centred at approximately 11 μm (Knapp et al. 2011). The GridSat B1 data are used by the University of California Santa Barbara's (UCSB) Climate Hazards Group (CHG) for climatic data analysis over a global extent (Funk et al. 2015), and spans back to 1981. The CHG is developing a similar dataset for temperature, and this work contributes to that effort. As part of the process for cloud cover detection, a method of estimating the cloud-free BT DTC from satellite TIR BT observations was derived. Similar to Göttsche and Olesen (2001), our methodology ignored emissivity and atmospheric effects on thermal radiance, as it focused on BT.

2.3 Data and methods

2.3.1 Data

Modelling of diurnal temperature was done using half-hourly TIR BT satellite observations from NOAA CPC (Janowiak et al. 2001). These data are operationally corrected for viewing angle effects, which are primarily geometric and radiative in nature, and result in reduction of the derived temperature off-nadir. The NOAA CPC TIR data are composited from several geostationary satellites. Over Africa, the NOAA TIR data uses the Meteosat IR sensor, which, spectrally is centred at 11.5 μm (Janowiak et al. 2001). The TIR data, originally a quasi-global dataset (60° N to 60° S) were cropped to a southern Africa window (5 to 60° E, 6° N to 38° S) at the University of California Santa Barbara's Climate Hazards Group. The TIR data had a 0.036 degree resolution. This is the same data used by the UCSB CHG in the CHIRPS rainfall archive (Funk et al. 2015). The Meteosat cloud mask developed by the SAFNWC (Derrien and Le Gléau 2005) was downloaded from the EUMETSAT Data Centre website (EUMETSAT 2018) for comparison with the cloud masking algorithms that we developed. The cloud mask data were in the Meteosat cartographic projection, which was incompatible with our data analysis algorithm. The cloud mask data were therefore reprojected from the Meteosat projection to a latitude-longitude coordinate system at 0.05 degree resolution, to facilitate extraction of point data from the extraction algorithm. The data derived from this cloud mask were referred to interchangeably in this work as the SAFNWC cloud mask or the SAFNWC cloud classification. The pixel values of the TIR data that were co-located with each of four Automatic Weather Stations (AWS) managed by the Southern African Science Service Centre for Climate Change and Adaptive Land Management (SASSCAL) Weathernet observation network (Kaspar et al. 2015) were extracted. The geographic coordinates and landcover and climate of these station locations are shown in Table 2.1, while a location map

is shown in Figure 2.1. The climate classification was based on the Köppen–Geiger climate classification system (Beck et al. 2018). Data for January, April, July and October from 2014 to 2016 were extracted to enable the analysis to be done separately for the hot, cold, wet and dry seasons. Further research reported in a Chapter 3 will relate modelled TIR time series to observed air temperatures at the selected AWS locations.

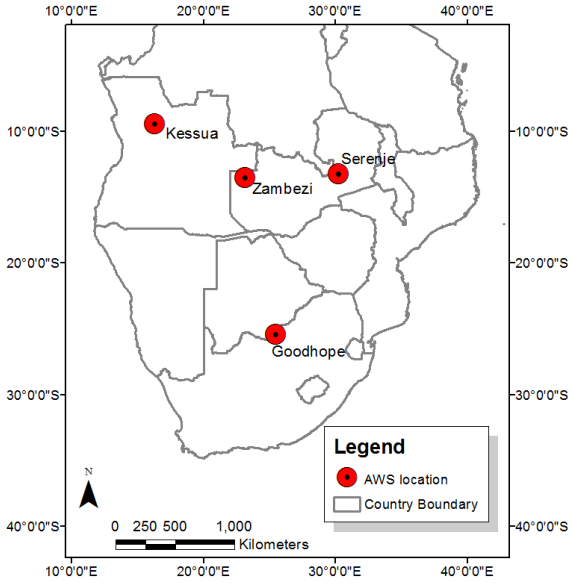


Figure 2.1 Location map for the AWS locations used in southern Africa

Table 2.1 AWS locations and descriptions. The land cover classification used was based on the USGS Global Land Cover Characterization (Loveland et al. 2000), while the local description is based on an analysis of the Google Earth interface

Station name	Country	Longitude, Latitude (°E, °S)	Altitude (m)	Local description	Land cover classification	Köppen–Geiger climate classification

Kessua	Angola	16.285,-9.464	1115	600 m away from a large (300 m width) seasonable river or floodplain. Surrounded by natural vegetation and fields, and a few buildings 60 m away	Deciduous broadleaf forest/ Savanna/ Water body	Aw: Tropical wet and dry or savanna climate
Goodhope	Botswana	25.427,-25.460	1275	Surrounded by sparse savanna grasslands and scattered buildings. Approximately 5 km away from large crop agriculture fields	Grassland	BSh: Hot semi-arid climate
Serenje	Zambia	30.215,-13.227	1395	Within 1000 m radius to large agricultural fields, savanna grasslands, urban settlement and small river	Dryland cropland and pasture/ Savanna	Cwb: Subtropical highland climate
Zambezi	Zambia	23.108,-13.534	1066	Located at airport, surrounded by large cleared area, small agricultural fields, housing settlement and large perennial river 1100 m away	Savanna	Cwa: Monsoon-influenced humid subtropical climate

2.3.2 Methods

We adapted the method developed by Minnis and Harrison (1984) to allow the identification of cloud-free observations using only IR data rather than VIS data in order to ensure consistency between daytime and nighttime cloud estimates. Cloud estimation was based on modelling of the diurnal brightness temperature cycle (DBTC), using a modification of equations developed by Parton and Logan (1981) and used by Savage (2016) for modelling the DBTC. The output DBTC from this process was not an estimate of land surface temperature, as emissivity and atmospheric corrections were not applied, but rather an approximation of what the satellite-observed BT would have been under the day's weather conditions, if cloud had not impeded the satellite observation

of the land surface directly. This quantity is alternately referred to as the cloud-free or the clear-sky brightness temperature. Among other decision-tree classification rules, the main premise was that a large difference between the observed BT and the modelled land-surface-equivalent temperature would be due to cloud cover. We tested the method using 30-minute time interval data.

The outputs of our DTC-based cloud masking methodology were validated using the SAFNWC cloud mask, which was considered as 'truth' for most purposes in this study, due to the results of validation work carried out by Derrien and Le Gléau (2005) on the SAFNWC algorithm. No comparison of the DTC-based cloud estimation with ground estimates was undertaken, due to potential limitations in compatibility between satellite- and ground-based cloud cover estimates (Derrien and Le Gléau 2005, Kotarba 2009).

2.3.3 Determination of the cloud-free diurnal brightness temperature curve

2.3.3.1 Theory

Equations and algorithms based on Parton and Logan (1981) were developed to model the cloud-free DBTC for each day where a minimum of two daytime and two nighttime cloud-free TIR satellite observations were available. Savage (2016) expressed three equations developed by Parton and Logan (1981) in terms of time intervals.

$$T(t) = T_{\mathrm{n}} + (T_{\mathrm{ss}} - T_{\mathrm{n}}) \exp\left[\frac{-b(t-t_{\mathrm{ss}})}{24-D}\right], \tag{2.1}$$

$$T(t) = T_{\mathrm{n}} + (T_{\mathrm{x}} - T_{\mathrm{n}}) \sin\left[\frac{\pi(t - t_{\mathrm{sr}} - c)}{D + 2a}\right], \tag{2.2}$$

$$D = t_{\mathrm{ss}} - t_{\mathrm{sr}}, \tag{2.3}$$

where $T(t)$ is the air temperature T at time t

T_{n} is the minimum air temperature for the current day

T_{x} is the maximum air temperature

T_{ss} is the air temperature at sunset

t_{ss} is the time of sunset

t_{sr} is the time of sunrise

D is the day length, in hours

b is a constant for modifying the exponential curve to change the rate of cooling

a is a constant, expressed in hours, that enables truncation of the sine curve

c is the amount of time in hours between occurrence of minimum temperature and sunrise.

Equations (2.1) and (2.2) model the nighttime and daytime air temperature curves respectively. Nighttime cooling is modelled as an exponential decay; daytime warming as a sine function. These equations and their inherent concepts were modified for use with TIR-satellite data for modelling the DBTC. In this analysis, all temperatures referred to in equations (2.1) and (2.2) are adapted to refer to the TIR BT, instead of 2-m air temperature. The purpose of modelling the DBTC was to

remove the effects of cloud observations, which reduce the BT compared to the equivalent cloud-free temperatures, and capture the influence of the diurnal heating and cooling of the land surface on the BTs.

As equation (2.1) relies on the knowledge of T_{ss} and T_{n}, which data may not always be available due to cloud cover, equation (2.4) was developed as a new exponential function that could use any two cloud-free temperature measurements.

$$T(t) = T_{m2} + (T(t_1) - T_{m2}) \exp\left[\frac{-b_t(t-t_1)}{t_2-t_1}\right], \tag{2.4}$$

where T_{m2} is given by:

$$T_{m2} = \frac{T(t_2)-T(t_1)\exp(-b_t)}{1-\exp(-b_t)}, \tag{2.5}$$

and where $T(t_1)$ and $T(t_2)$ are temperatures at any two different nighttimes t_1 and t_2 on the same night. The constant b_t controls the rate of nighttime exponential decay, with values that can range between 0.1 for a near-linear decay, to 3 or more for a faster exponential decay.

The cloud-free nighttime temperature curve was determined using equations (2.4) and (2.5). The nighttime BT $T(t_1)$ used in equations (2.4) and (2.5) was determined daily as the highest observed BT at t_1 between t_{ss} + 0.5 h and 23h00 local time, while $T(t_2)$ was the highest BT observed at t_2 between 01h00 local time and t_{sr} + 0.5 h. Only BTs greater than a pre-defined lowest possible temperature were considered in the analysis. Cases in which too few non-cloud temperature observations were available were marked as missing.

Equations (2.6) through (2.10) were developed to enable determination of T_{n} and T_{x} for equation (2.2). Given any two daytime temperatures $T(t_1)$ and $T(t_2)$ that occur at times t_1 and t_2 on the same day, we can write:

$$\Delta T_{\mathrm{nx}} = \frac{T(t_1) - T(t_2)}{\sin \tau_1 - \sin \tau_2}, \tag{2.6}$$

$$T_{\mathrm{n}} = T(t_1) - \Delta T_{\mathrm{nx}} \sin \tau_1, \tag{2.7}$$

$$T_{\mathrm{x}} = T_{\mathrm{n}} + \frac{T(t_1) - T_{\mathrm{n}}}{\sin \tau_1}, \tag{2.8}$$

where

$$\Delta T_{\mathrm{nx}} = T_{\mathrm{x}} - T_{\mathrm{n}}, \tag{2.9}$$

and

$$\tau_i = \pi \times \frac{t_i - t_{\mathrm{sr}} - c}{D + 2a}, \tag{2.10}$$

The estimated daytime cloud-free BT curve was determined using equation (2.2), and equations (2.6) through (2.10) as required. The daytime BTs $T(t_1)$ and $T(t_2)$ used in equations (2.6) through (2.10), together with their corresponding times t_1 and t_2 were determined iteratively on a daily basis, as those temperatures that produced a DTC with the lowest mean absolute error (MAE)

when compared against all the observed potentially cloud-free BTs. The equations were fitted to cloud-free observations, which were also identified through an iterative process.

2.3.3.2 Modelling of DTCs and cloud cover

Diurnal temperature curves and cloudiness were modelled through an iterative process that progressively improved both the DTCs and the identification of cloudy conditions. The process was run at a monthly scale, for each month and year of available data. Firstly, the monthly maximum BT at each 30-minute time slot was determined using a simple statistical maximum value analysis. The resultant chronological series of maxima were fitted to a maximum monthly composite DTC using equations (2.2) to (2.10), as appropriate. After determining the monthly maximum DTC, the DTC for each day of the month was processed. This began with a preliminary identification of clouds on the day in question using the following criteria: A TIR observation was judged to be cloudy if:

- the BT was less than a predefined lowest acceptable cloud-free temperature threshold. This predefined temperature was set to -5 ℃ during the austral winter, and 0 ℃ at any other time of the year. These threshold values were observed to be lower than the lowest values observed on days with smooth curves typical of expected clear-sky DTCs, in the locations analysed
- any of the five observations made immediately before it, as well as any of the five observations made immediately after it, had higher BTs than the observation in question, while allowing for an uncertainty epsilon value of 2 ℃. This phenomenon was referred to as an eleven-point moving window temporal-thermal depression
- for nighttime observations, any later observation was at least at least 2 ℃ warmer than the observation in question

- BTs were rising at fast rate, greater than 10 ℃ per hour

The observations that had been preliminarily determined as cloudy were excluded from the next step, which was the first estimation of the daytime and nighttime DTCs. All possible combinations of the cloud-free observations were tried, using them as $T(t_1)$ and $T(t_2)$ in equations (2.5) and (2.6) for nighttime and daytime DTC estimation respectively, thus generating several candidate DTCs for each day. For both daytime and nighttime, t_2 was set to be always later than t_1, and for nighttime analysis, $T(t_2)$ could not be higher than $T(t_1)$, as the exponential nighttime function was expected to be monotonically decreasing. Combinations of $T(t_1)$ and $T(t_2)$ that did not meet this criteria were excluded from the analysis. Candidate DTCs had to meet a number of criteria for consideration, including:

- the DTC-calculated sunset and sunrise BTs could not be lower than the predefined lowest acceptable cloud-free BT
- the DTC-calculated minimum BT could not be more than 5 degrees lower than the lowest observed BT of the day
- the calculated sunset BT had to be lower than the highest observed BT of the day
- for nighttime DTCs, the calculated sunset BT had to be higher than the sunrise BT

For each day, the MAE for each candidate DTC generated was calculated using all the observations of the day that had been determined as cloud-free, and the DTC with the lowest MAE was selected as the best preliminary DTC of the day. The selected DTC was subjected to an MAE and cloudiness-based threshold test in order to determine if the DTC could be considered reliable. DTCs having a lower MAE and a higher number of cloud-free observations (i.e. less cloudy) were judged to be more reliable. Varying MAE thresholds, dependent on the number of cloud-free

observations, were used in this test. The thresholds were varied such that days with more cloud-free observations had higher MAE thresholds, while those with less cloud-free observations had more stringent MAE threshold requirements. For example, a DTC with more than 20 cloud-free observations needed an MAE below 3 ℃ to be considered reliable; while a DTC with between 15 and 20 cloud-free observations required an MAE of less than 2 ℃ to pass the reliability test. This approach was adopted after noting that days with a low number of cloud-free observations were prone to producing implausible DTCs while maintaining a moderately low MAE. All the days of the month with acceptably reliable DTCs, according to the criteria just outlined, were used to calculate a monthly median composite DTC.

The results of the DTC reliability test, together with the availability of the monthly median and maximum composite DTCs, enabled the application of a number of additional criteria to further improve the selection of daytime DTCs, including:

- if the previous night's DTC was determined to be reliable, the difference between the daytime-modelled sunrise BT and the previous night's modelled sunrise BT could not be greater than 1 ℃
- the DTC-modelled sunrise and sunset BTs could not be greater than the sunrise and sunset BTs of the monthly maximum composite DTC, respectively
- the modelled midday BT should be greater than modelled sunrise and sunset BTs

The preliminary DTCs that had passed the reliability test were used to further refine the cloud identification, based on DTC residuals that were calculated as observed BT minus DTC-modelled BT. All residuals less than -8 ℃ were classified as cloudy. Residuals between -4 and -8 ℃ were

classified as partially/possibly cloudy, while residuals greater than -4 ℃ were classified as cloud-free. Additional criteria used for the refined cloud identification algorithm included:

- all daytime observations before $t_{\mathrm{sr}} + c$ (as defined in equation (2.2)) were classified as cloudy
- when observations had been classified as possibly/partly cloudy, and at least 2 of the other surrounding observations within a 5-point moving window were also classified possibly/partially cloudy, then the observations were re-classified as cloud-free
- if an observation was classified as cloud-free, but its residual was more than 2 ℃ lower than the average of the residuals of the four surrounding time slots, then it was reclassified as possibly/partially cloudy

After reclassifying the cloudy observations based on these additional criteria, the MAE-based identification of the best DTC was re-done using the procedures previously described, and incorporating the newly-identified cloud-free observations.

In cases where the MAE or the number of cloud-free observations did not pass the reliability threshold tests, the model parameters of temporally close DTCs (i.e. DTCs estimated in days just before or after the day in question) were used in the second pass. For daytime, the set of DTC model parameters producing the best statistics (greater than pre-determined thresholds) within a seven-day window were used for selecting the model parameters a and c. If no DTCs within a 7-day window passed the acceptable thresholds, then the monthly median composite DTC's model parameters were used instead. This was referred to as a proxy DTC. A similar approach was used for the nighttime DTCs, for determining the b parameter. The proxy DTCs were constrained such that their maximum BTs could not be higher than the maximum BT of the monthly maximum

composite DTC, and on days with extensive cloudiness, the maximum BT could not be higher than the maximum BT of the monthly median composite DTC. Parameters used to supplement the proxy DTC included the sunset and sunrise BTs from preceding and following DTCs as appropriate, the BT range, and the difference between sunset and sunrise temperatures for selected DTCs.

Once the proxy DTCs had been established for each day, the cloud identification procedure previously described was undertaken again to produce a final cloud mask.

2.4 Results

2.4.1 Determination of clouds by modelling DTC of TIR brightness temperature

2.4.1.1 Diurnal brightness temperature curves

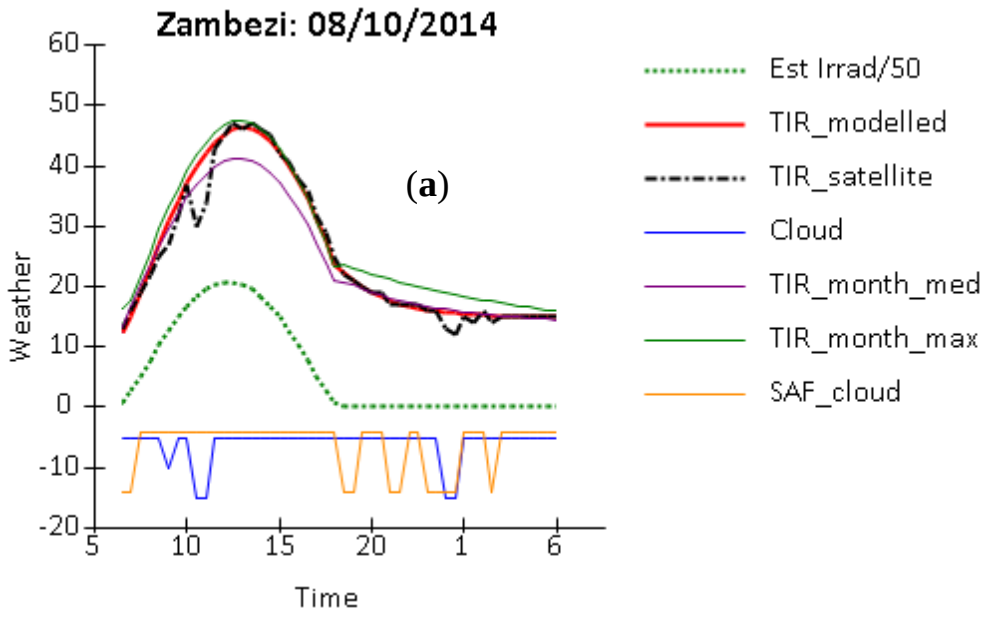
Visual inspection of graphed outputs (Figures 2.2 to 2.4) suggested that DTC models of cloud-free TIR BT, and the subsequent cloud classifications, identified comparatively high temperatures as cloud-free BTs, while most temporal-thermal depressions were generally identified as clouds. Figures 2.2 to 2.4 show the diurnal variation of several weather parameters for selected days and station locations, to illustrate some typical patterns observed in the model. The dashed black line is the observed TIR BT. The thick red solid line is the modelled cloud-free temperature for the day. The solid purple line is the monthly median temperature curve, calculated using observations for each 30-minute time slot, from all the days for which a DTC was successfully estimated using that day's observations. The solid green line is the monthly maximum temperature curve, calculated from using the maximum observed temperature of the month at each time slot. All temperatures are in ℃. The solid blue line is the cloud/clear-sky classifications for the TIR-based

model: values of -5 indicate clear-sky cloud-free conditions, values of -10 indicate possible/partial cloud, values of -15 indicate warm cloud, while -20 indicates cold cloud. The solid orange line is the SAFNWC cloud classification, for which values of -4 indicate cloud-free conditions, while values of -14 indicate cloudy conditions. The values used for the cloud classification are nominally selected, and unitless. The estimated solar irradiance in dashed green provides visual differentiation between daytime and nighttime, with a value of 0 indicating nighttime, while values above 0 indicate daytime. Solar irradiance is represented on the graph in Wm^{-2} divided by fifty in order for the solar irradiance curve to correctly display on the same graph as the temperature curves.

Days exhibiting a smooth variation of TIR BTs, and few, short-term temporal-thermal depressions, such as shown in Figure 2.2a, generally displayed close correspondence between the raw, cloud free TIR BTs (dashed black lines) and the modelled DTC (solid red lines). Even in cases where there were several short temporal-thermal depressions, the algorithm closely modelled the peaks between each depression (e.g. nighttime part of Figure 2.2b).

The modelled DTC also displayed close correspondence with cloud-free portions of the original TIR BT, even in cases with several hours of continuous cloud (e.g. Figures 2.3a and 2.3b), particularly when cloud-free observations were available at key times of the day. When the start of the TIR BTs daytime sinusoidal temperature increase was delayed by a few hours (e.g. in Figure 2.3b), potentially due to early morning fog or low-level cloud cover, the DTC algorithm similarly delayed the start of daytime temperature increase, and often showed close correspondence with ensuing cloud-free TIR BTs. In such instances, the modelled daytime DTC often started increasing before the observed BTs, as shown in Figure 2.3b.

Some days displayed noisy BT diurnal sequences, with several spikes and temporal-thermal depressions characterizing the day's observations, as shown in Figure 2.4a and 2.4b. In such cases, the ability of the model to estimate a DTC was influenced by the number of BT observations close to or higher than the monthly median temperature curve. When several observations met this criteria, the DTC was successfully estimated, as shown in Figure 2.4a. However, where much of the day's TIR BT were much lower than the monthly median temperature curve, then the monthly median curve was assumed (e.g. Figure 2.4b).



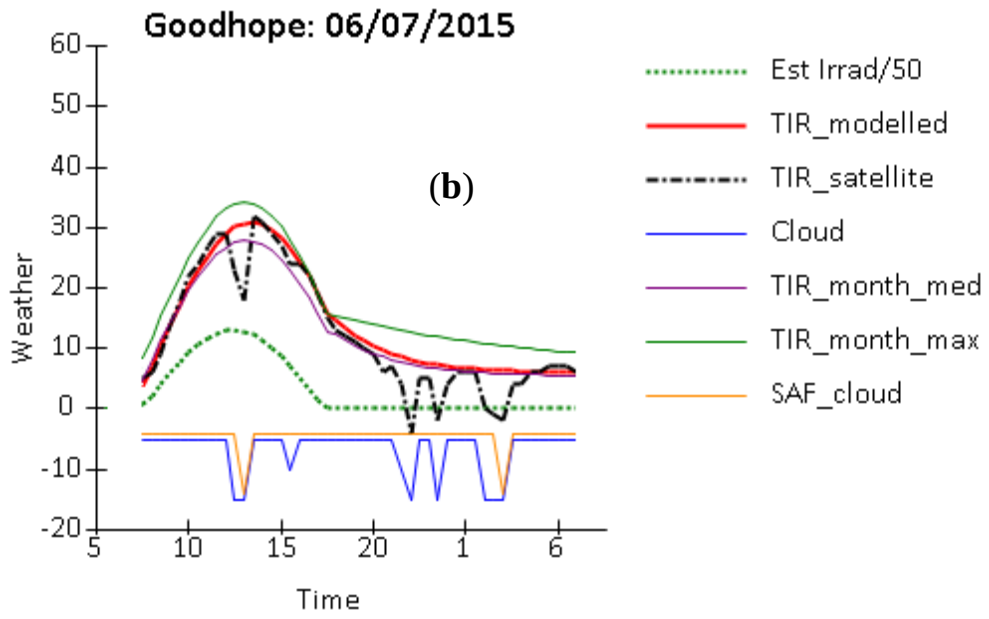


Figure 2.2 Observed and modelled DTCs, cloud/clear-sky classifications, and estimated solar irradiance divided by fifty for sample days and locations with low and short-duration cloud cover, namely (a) Zambezi on 8 October 2014 and (b) Goodhope on 6 July 2015

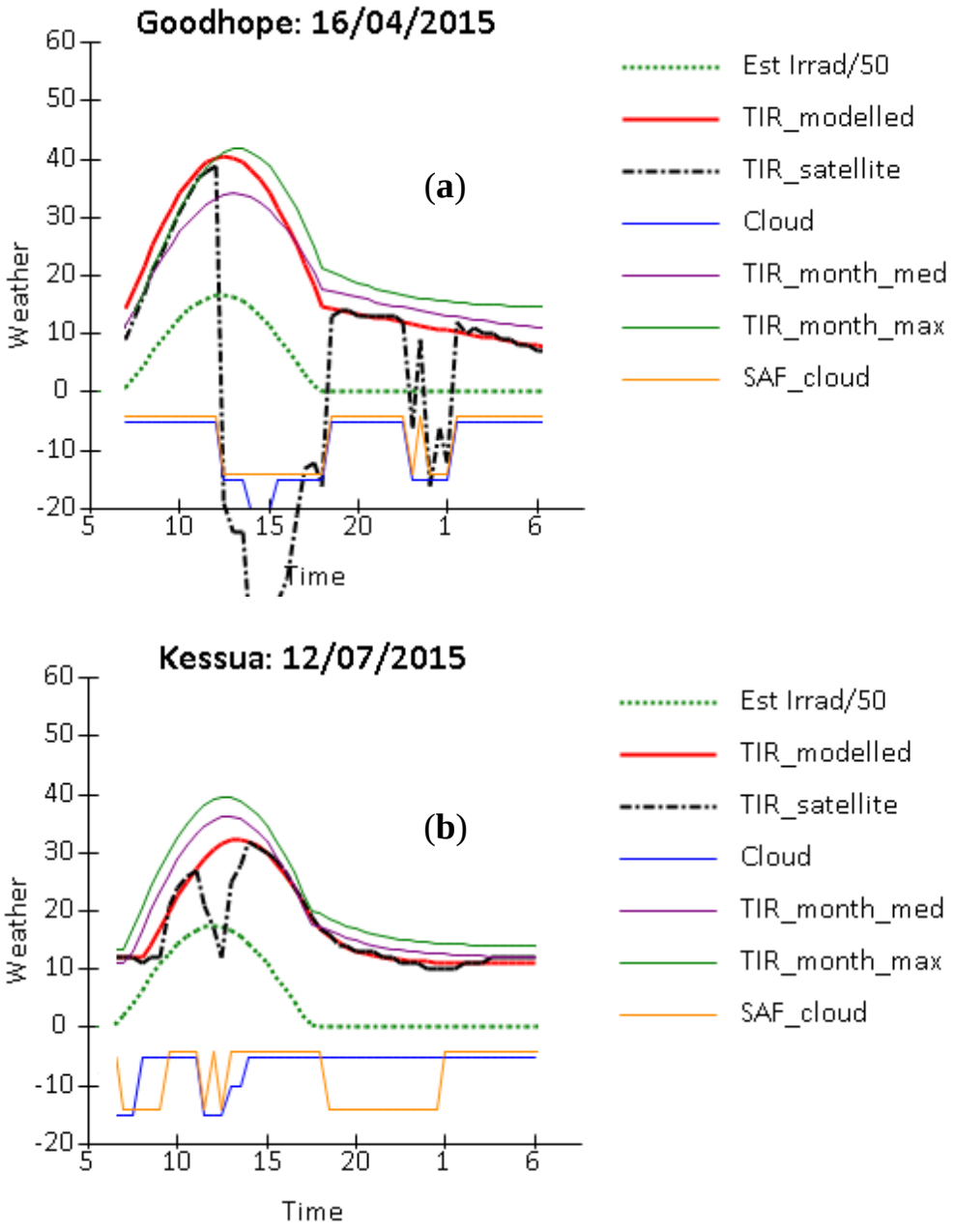


Figure 2.3 Observed and modelled DTCs, cloud/clear-sky classifications, and estimated solar irradiance divided by fifty for sample days and locations with consecutive hours of cloud cover, namely (a) Goodhope on 16 April 2015 and (b) Kessua on 12 July 2015

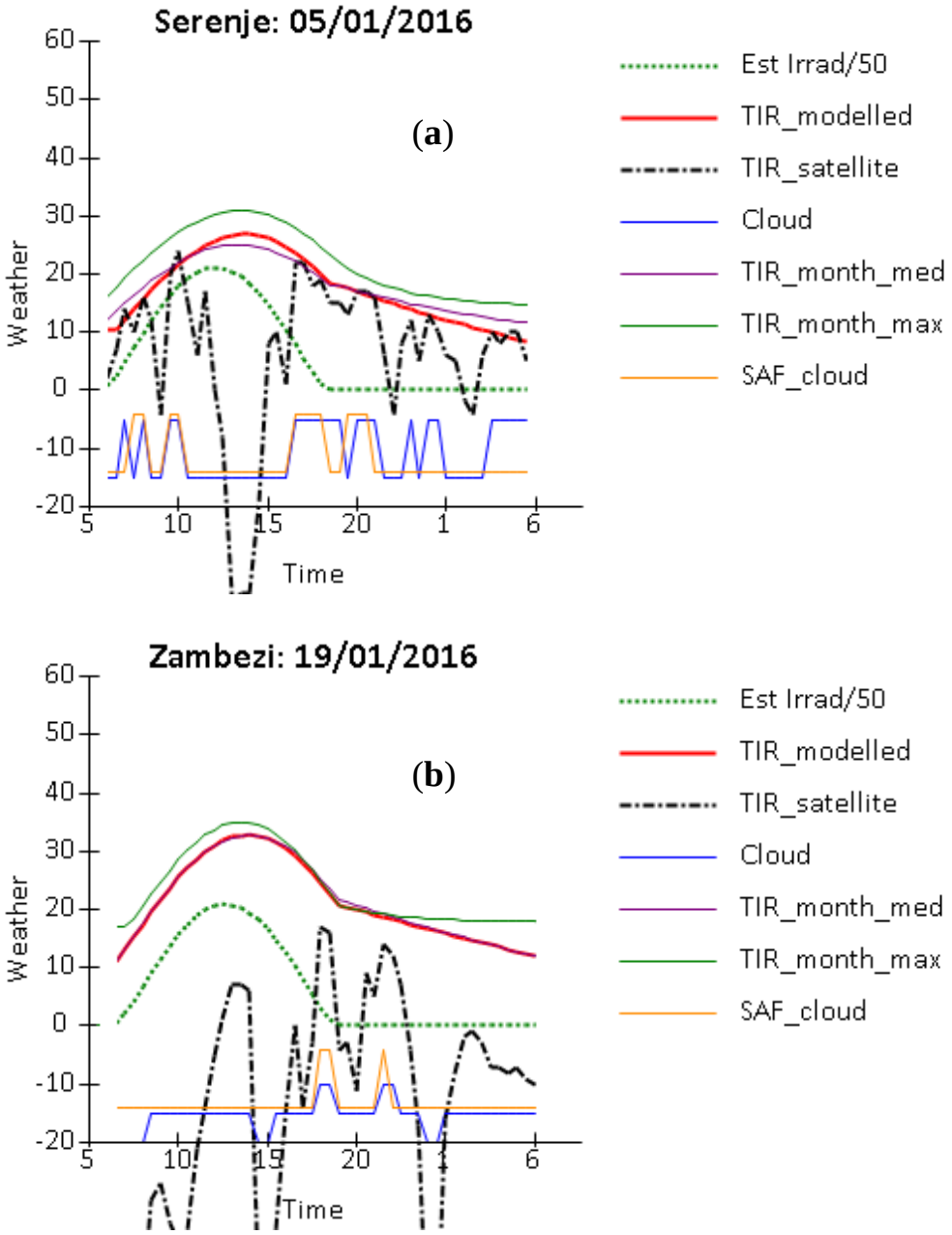


Figure 2.4 Observed and modelled DTCs, cloud/clear-sky classifications, and estimated solar irradiance divided by fifty for sample days and locations with noisy BT diurnal sequences, namely (a) Serenje on 5 January 2016 and (b) Zambezi on 19 January 2016

A total of 53,929 TIR observations across all years, months and stations were analysed. Of these, 33,636 observations were determined to be non-cloudy or possibly/partly cloudy. The MAE of the modelled non-cloudy and possible/partly cloudy temperatures was 0.9 ℃ (non-cloudy) and 5.4 ℃ (possibly/partly cloudy) and 1.1 ℃ overall.

Of the 1131 days (nights) analysed, 71.5% (75.4 %) produced daytime (nighttime) DTCs for TIR BT with a MAE of less than 2 ℃, 11.9% (7.0%) had a MAE of between 2 and 4 ℃, while 4.2% (1.9%) had MAE above 4 ℃. The remainder of the days (nights) did not have any cloud-free observations, and their MAE was therefore not calculated. The majority of DTCs passed the reliability test, 63% (67%) of all daytime (nighttime) DTCs. These DTCs were developed using concurrent daytime (nighttime) data, while the remainder of the DTCs were developed using the statistics of nearby DTCs or the monthly median as described in the methodology section. Of the reliable daytime (nighttime) DTCs, 95.3% (99.1%) had an MAE of less than 2 ℃. The daytime (nighttime) DTCs with the best statistics occurred in July, with 95.0% (99.7%) of all fitted July DTCs having a MAE of less than 2 ℃. For January, the worst performing month, 48.1% (39.6%) of January daytime (nighttime) DTCs had a MAE of less than 2 ℃, while 28.9% (44.4%) of January daytime (nighttime) DTCs had insufficient cloud-free observations to allow calculation of a MAE. January had an average daytime (nighttime) cloud cover of 8.9 (8.1) hours, while July had 0.9 (0.4) hours of daytime (nighttime) cloud cover on average.

For the daytime DTCs, the average a and c parameters as defined in equation (2.2) were calculated as 0.31 and 0.68 respectively, with coefficients of variation of 2.61 and 0.96 respectively. For the nighttime DTCs, the average b_t parameter defined in equation (2.10) and

(2.11) was 1.69, with a coefficient of variation of 1.33. Parton and Logan (1981) used 2.2 for the parameter *b* (as defined in equation (2.1)) for air temperature at 2 m.

2.4.1.2 TIR-based cloud mask

There was general correspondence observed between TIR-based cloud cover estimates (e.g. Figure 2.2 - 2.4, blue line) and SAFNWC-based cloud estimates (e.g. Figure 2.2 - 2.4, orange line). In particular, TIR BTs at low temperatures were identified by both methods as cloud (e.g. Figure 2.3a and Figure 2.4a). However, a number of differences were also noted in the two classifications. For example, observations classified as cloud by the TIR-based system due to the presence of a temporal-thermal depression were sometimes classified as cloud-free by the SAFNWC system (e.g. Figure 2.2a (daytime) and Figure 2.2b (nighttime)). In other cases, observations classified as cloud-free by the TIR-based system were classified as cloud by the SAFNWC system (e.g. Figure 2.2a (nighttime) and Figure 2.3b). The TIR-based analysis further differentiated between cold cloud, warm cloud, and possible cloud, although this separation of cloud levels was not subjected to further analysis. A confusion matrix analysis comparing TIR-based cloud cover estimates with SAFNWC-based cloud estimates (Table 2.2) was produced. The SAFNWC-based cloud estimates were regarded as the true values. The analysis showed an accuracy of 85.4%, a misclassification rate of 11.7%, a true positive rate of 78.6% and false positive rate of 7.6% (Table 2.3). The classification of possible cloud was not included in the calculations except as part of the totals.

Table 2.2 Confusion matrix comparing TIR-based cloud classifications with SAFNWC cloud mask for all available time slots at 30-minute intervals

	TIR-Based				
SAFNWC Cloud Mask		**Cloud**	**Non-Cloud**	**Possible Cloud**	**Total**
	Cloud	12530	2835	580	15945
	Non-Cloud	1808	21288	569	23665
	Total	14338	24123	1149	39610

Table 2.3 Statistical performance comparing TIR-based cloud classifications with SAFNWC cloud mask, based on results of confusion matrix (Table 2.2)

Statistical Performance	Percentage
Accuracy	85.4%
Misclassification rate (error rate)	11.7%
True positive rate (sensitivity)	78.6%
False positive rate	7.6%
Specificity	90.0%
Precision	87.4%
Prevalence	40.3%

An explanation on the interpretation of the statistics used in Table 2.3 is included in Appendix A.

A total of 815 daily graphs (of the type shown in Figure 2.2 - 2.4) were produced for days and station locations that contemporaneously had TIR data, SAFNWC mask data, and SASSCAL observations data. Forty-three of these graphs showed a particular difference in the cloud classification done using the two methods: in each of these cases, the TIR DTC showed at least one large temporary decline in temperature of over 5 ℃ from an otherwise smooth DTC in the temperature. This decline was interpreted as a cloud by the TIR-based cloud detection algorithm, while the SAFNWC-based cloud algorithm did not contemporaneously detect a cloud. Figure 2.2a highlights one such instance.

2.4.2 Monthly aggregations of DTC models and data

At each 30-minute time slot throughout the day, the monthly median cloud-free TIR BT was calculated for each month analysed between 2014 and 2016 inclusive, namely January, April, July and October, for each of the four locations. The monthly median DTC-modelled BT was similarly calculated, but using all the data available, rather than only that determined as cloud-free. A comparison between the median monthly DTC for cloud-free TIR temperature and modelled TIR temperature calculated as explained above suggests that the modelling, which was performed at a daily time-step, produced outputs that closely corresponded to the original cloud-free TIR temperatures, when the results were aggregated to a monthly time-step. The coefficient of determination between the monthly median-aggregated cloud-free observed TIR BT and the monthly median-aggregated modelled TIR temperatures was 0.98 (Figure 2.5). The slope of the fitted linear equation, at 1.01, was very close to the ideal value of 1, while the equation constant, at -0.45 ℃, was close to the ideal value of 0 ℃.

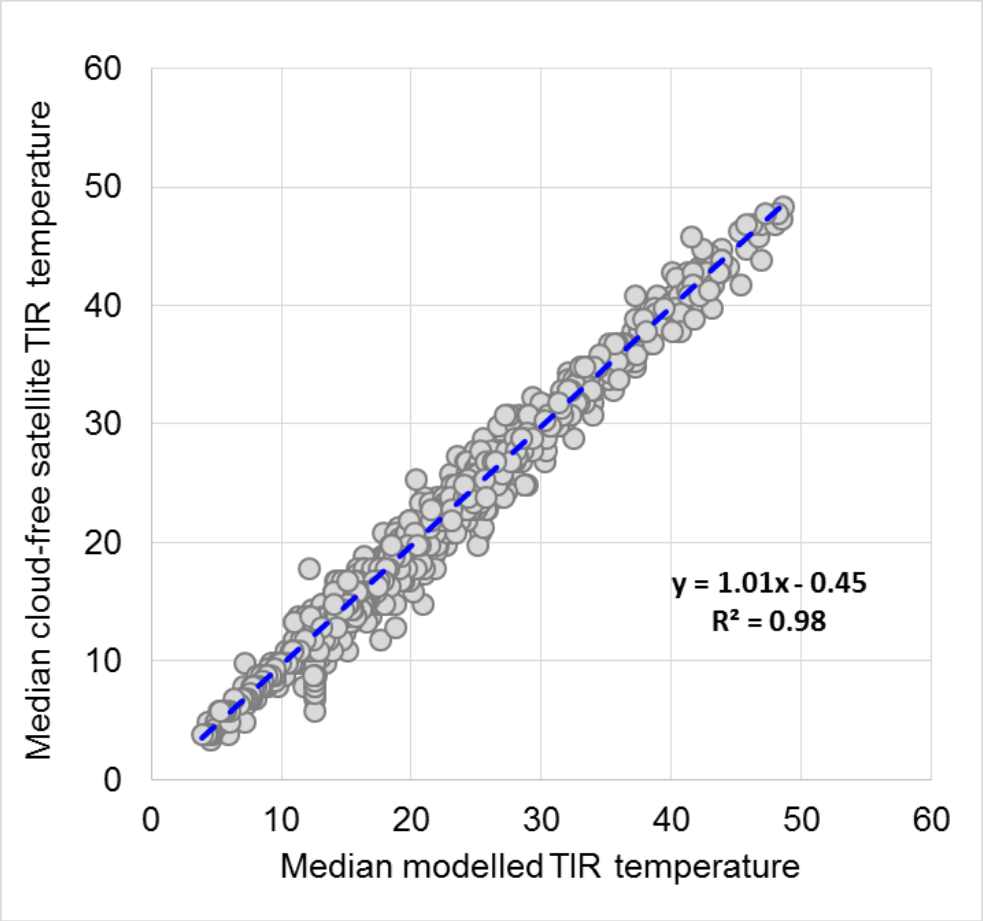


Figure 2.5 Scatterplot of the monthly median-aggregated modelled TIR temperatures against the monthly median-aggregated cloud-free observed TIR BT for all available data for each station, between 2014 and 2016 for January, April, July and October. Each dot on the graph represents a monthly median temperature at one of the 30-minute time slots of the day, for a given month, year and location. The dashed blue line is the least-squares fitted linear regression line

The monthly median-aggregated modelled temperatures, which were calculated using all the modelled TIR temperature estimates, were also compared to monthly median satellite TIR BTs that were calculated using all TIR temperature observations (without screening for clouds). This

provided greater comparability between the number of data points used to calculate the median for each dataset. In this case, the regression statistics are markedly worse than the case in which the cloud screening is applied. The coefficient of determination between the monthly median-aggregated observed TIR BT and the monthly median-aggregated modelled TIR temperatures was 0.42 (Figure 2.6). The slope of the fitted linear equation was 0.96, while the equation constant was -6.30 ℃. The scatterplot shows several data points that are below the regression line, while the top edge of the scatterplot, indicated by the red dashed line in Figure 2.6, has a relatively flat shape.

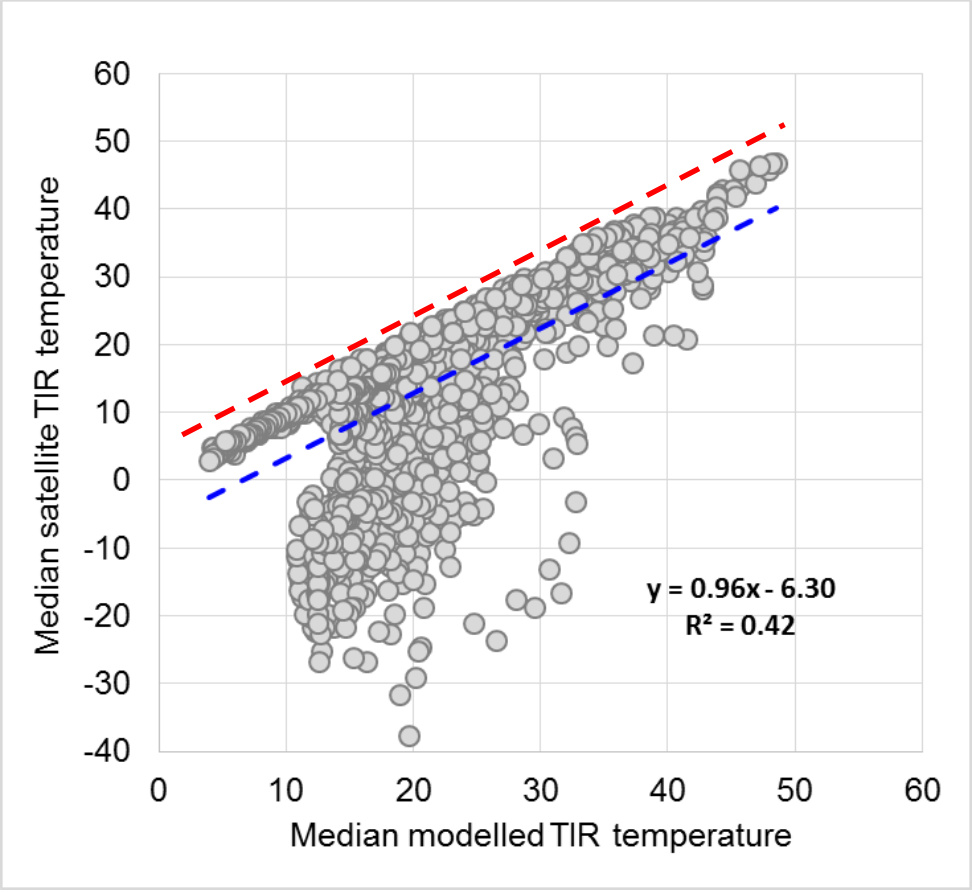


Figure 2.6 Scatterplot of the monthly median-aggregated modelled TIR temperatures against the monthly median-aggregated observed TIR BT for all available data for each

station, between 2014 and 2016 for January, April, July and October. Each dot on the graph represents a monthly median temperature at one of the 30-minute time slots of the day, for a given month, year and location. The dashed blue line is the least-squares fitted linear regression line. The red line is manually inserted to highlight the top edge of the scatterplot

2.5 Discussion

Despite the generally high correlation between the TIR and SAFNWC-based cloud detection methods, several errors were identified in the outputs from each method. For example, it was observed that the DTC-based cloud detection algorithm performed poorly on mornings with prolonged periods of low-level cloud cover or fog. TIR temperature profiles that exhibit near-constant temperatures for several observations immediately following sunrise, instead of the typical early daytime sinusoidal temperature rise, are likely to be due to the presence of low-level cloud cover. Figure 2.3b shows an example of this phenomenon. In most instances when this occurred, the SAFNWC-based cloud mask indicated a classification of cloud cover during these episodes, which typically ended when TIR temperatures showed a sudden increase in temperature, followed by a sinusoidal DTC until sunset. As the DTC-based cloud mask for such episodes was based primarily on the mathematical difference between observed and modelled temperature, the DTC-based cloud detection system tended to misclassify observations that occurred soon after the modelled time of start of daytime temperature increase. Despite this limitation, the remainder of the modelled daytime DTC followed an expected sinusoidal pattern. Temperature increase typically started even before the complete clearance of cloud, due to the diffuse radiation reaching the surface. The rapid increase in TIR temperature that occurred at the end of the period of near-

constant temperature, usually around mid-morning, typically coincided with the end of cloudy conditions, as determined from the SAFNWC-based cloud mask. The remainder of the daytime cloud-free TIR observations generally matched well with the modelled temperature. The error described above can be addressed in future work: as the DTC-based cloud classification framework used here combines the use of equations and decision-tree rules, a more rigorous set of rules can be introduced into the process to reduce the occurrence of this error.

Comparison between the SAFNWC-based cloud mask and the raw TIR BTs suggests some limitations in the accuracy of the SAFNWC mask. Two of these cases are illustrated in Figures 2.2a, 2.2b and 2.3b. One expression of potentially inaccurate cloud detection occurred when sudden large reductions in TIR BTs were classified as cloud-free by the SAFNWC method. Instances of such occurrences were highlighted in Figures 2.2a and 2.2b. Such large decreases in BT appear to be due to the movement of small, low-level clouds, and were identified as cloud cover by the DTC-based cloud detection algorithm. A second type of potential cloud misclassification by the SAFNWC cloud masking algorithm occurs when cloud-free observations are incorrectly classified as cloud. This happens when daytime observations that closely follow a sinusoidal curve, or nighttime observations that closely follow an exponential curve, are misclassified as cloudy conditions. The expectation is that under these conditions, the introduction of any cloud cover at 100 m or more above the ground surface would result in a reduction in BT of at least 1 °C, based on the dry adiabatic lapse rate, or approximately 0.5 °C for a typical moist adiabatic lapse rate (Minder et al. 2010). Such instances are highlighted in the nighttime curves of Figures 2.2a and 2.3b, when the SAFNWC algorithm detected cloud cover, although the BTs were closely following an exponential curve. As the SAFNWC cloud algorithm discriminates clouds from clear-sky conditions using only temperature thresholds at night, the error likely arises when

the TIR BTs are close to the algorithm's pre-determined temperature thresholds, more so during nighttime when the SAFNWC algorithm does not use reflectance data (Derrien and Le Gléau 2005).

The cloud masking algorithm used pre-defined temperature thresholds and threshold ranges based on a priori knowledge of the study area and seasonality. These thresholds may not apply for other areas with different climatology, for example where temperatures climatologically decrease below the thresholds used for the study area (Section 2.3.2). In order to successfully generalize the methodology to a wider range of geographic locations, climatology fields would need to be first developed for those areas where the masking algorithm is to be applied, and dynamically used in setting the thresholds required at different locations, rather than fixed thresholds.

The cloud classifier and DTC modelling showed higher inaccuracy under more extreme cloud conditions, with January, the cloudiest of the four months analysed, having the highest errors for both the DTC modelling and the cloud masking, while July, the least cloudy month had the lowest errors. The assumption that the monthly median DTC could represent the daily DTC during extensively cloudy conditions, such as on overcast days, was not always valid, as demonstrated by Figure 2.4b, which shows the temperature and cloud variations for the Zambezi station location on 19 January 2016. In this example, two observations that had relatively high TIR temperature (compared to temporally close TIR observations) fell well below the median temperature curve, and below the assumed daily temperature curve, but were identified as non-cloudy by the SAFNWC cloud mask. A combination of both SAFNWC-based and DTC-based cloud classifiers for identifying non-cloud conditions may improve the accuracy of the DTC modelling. Physics-

based modelling to simulate the potential impact of continuous cloud cover on BT may also provide an improvement over the assumption for median temperatures on overcast days.

The methodology used in this study was tested over four locations with a tropical, sub-tropical or semi-arid climate. These locations included some of the typical locations where summer-cropping is practiced in southern Africa. Many of these areas do not include areas where snow or ice occurs seasonally, or where persistent cloud cover may occur for consecutive days or weeks. As such, the performance of the DTC modelling and cloud masking algorithm has not been determined in such areas. Given the reliance of the methodology on the use a few cloud free observations every day, it is expected that the DTC estimation would exhibit low accuracy over areas where extended periods of cloud cover. The model in its current form is applicable to southern Africa summer-cropping areas and locations with similar climates. More testing and development work would be required for application to other climates. However, such work was beyond the scope of this study.

Part of the differences observed between the SAFNWC- and DTC-based cloud masks may be attributed to the difference in resolution of the two datasets (0.050 and 0.036 respectively). However, the consistency and explicability of many of the differences observed suggests, together with the small difference in resolution, suggests that this may be a minor factor.

The DTC modelling used in this study relied on an iterative procedure for finding the model fit with the lowest MAE. This proved computationally expensive, and alternative methods were considered. The Levenberg-Maquaart (LM) algorithm has been used in previous DTC modelling attempts (e.g. Göttsche and Olesen 2001) and can provide a quick solution with high convergence. This approach was found limiting in terms of (1) the requirement for a close approximation of the

initial values in order to produce an accurate model and (2) the complexity of incorporating decision-tree type rules in the classification in which the identification of clouds is an iterative part of the process.

The high level of correspondence observed between the median monthly DTC for cloud-free TIR temperature observations, and the DTC-modelled TIR temperature for all time slots (Figure 2.5), suggests that cloud-free temperature estimates generated by the DTC-estimation algorithm accurately represent the statistically calculated cloud-free DTC at a monthly time scale. This finding has more relevance given that the median for the modelled TIR DTCs was based on estimates for all days and all time slots, even in the presence of persistent cloud cover such as that observed in Figure 2.4. When all satellite BTs (with no cloud screening) are similarly compared with all DTC-modelled temperatures (Figure 2.6), a much lower correlation is observed. The large number of data points that are well below the regression line in Figure 2.6 are likely due to the influence of clouds when calculating the monthly median satellite temperatures. Cloud top temperatures are generally lower than co-located ground temperature due to the lapse rate, which causes the calculated monthly median temperature to be lower than the cloud-free median temperature, when any clouds are present. The dots close to the top edge of the scatterplot (dashed red line) likely represent observations that were cloud-free for most of the month (e.g. in July), an explanation supported by the relatively flat shape of that section of the scatterplot, as well as the general parallelism between the top edge and the regression slope which is very close to 1.0.

While the analysis used in this study only covered four distinct months of the year, for the years 2014 to 2016, the method has the potential to be applied to the entire archive of TIR data processed and maintained by NOAA CPC from September 1999 (Janowiak et al. 2001). The DTC cloud

masking method incorporates the use of BT thresholds to preliminarily distinguish cloudy from potentially clear-sky observations, and these thresholds vary from summer (e.g. October and January using thresholds of 0 °C) to winter (e.g. July using a threshold of -5 °C). In a scheme in which each month of the year from January to December is to be processed, decisions would need to be incorporated on which months to change from a summer BT threshold to a winter BT threshold, and the thresholds can differ from the two thresholds of 0 °C and –5 °C used in this study. Such monthly varying thresholds can be estimated at a pixel level using the data, for example as a function of the maximum BT calculated for each time step of each month. This time-step-based threshold has similar aspect to the method used by Schädlich et al. (2001). A detailed treatment of this proposed methodology is beyond the scope of this study.

While the NOAA CPC dataset used in this study extends back to 1999 (Janowiak et al. 2001), another TIR data, the GridSat B1 infrared Climate Data Record (Knapp 2008), extends further back to 1982 over Africa, potentially extending the application of the DTC-based cloud masking algorithm even further. The GridSat B1 data however comes with a temporal resolution of 3 hours, a significantly lower resolution than the 0.5 hour resolution on which the DTC methodology was developed. As such, additional tests would be required on the 3-hour resolution data to determine the applicability of the DTC cloud-masking methodology on this lower-resolution data.

2.6 Conclusions

A novel technique for cloud masking using a single thermal band on a geostationary satellite instrument has been developed. The technique utilizes a combination of a series of decision

classification rules, and fitted sinusoidal and exponential equations for daytime and nighttime analysis respectively. This method produces comparable results to the Satellite Application Facility for Nowcasting (SAFNWC) cloud mask product (Derrien and Le Gléau 2005), an existing, validated, operational technique. By using a dynamic thresholding technique, the cloud masking technique identifies likely cloud in some scenarios where the operational SAFNWC cloud mask misses them, thereby providing a supplemental technique which can potentially improve existing cloud-masking methods. Our method also uses a single satellite band in the thermal infrared part of the spectrum, thus providing a uniform day-night cloud masking scheme, and facilitating an alternate mechanism for cloud masking of long-term climate data records like the GridSat B1 IR dataset. In contrast, several past methods have relied on both thermal and visible satellite imagery for daytime cloud masking, and thermal only satellite imagery for nighttime masking, thereby creating a discontinuity between the two time periods.

The cloud masking method doubles as an algorithm for modelling the cloud-free diurnal temperature curve (DTC), using a feedback iterative approach. The preliminarily identified clouds were used as a cloud mask to model a cloud-free DTC, which was then used as a cloud mask to refine the cloud mask, and so on. A comparable method in the literature for producing satellite-based DTC relied on an existing cloud mask to model the DTC using identified cloud-free observations, without accounting for errors in the cloud mask. By coupling the development of the cloud mask and the DTC modelling in this innovative approach, increasingly accurate estimates of both the cloud mask and the cloud-free DTC can be produced.

Improvements in the methods developed in this study can be incorporated through several refinements that can be applied in future work. Firstly, additional rules can be applied to improve

the identification of cloud under conditions of early morning low-level cloud or fog. Secondly, the cloud masking algorithm can be combined with a priori information from alternative cloud masking schemes such as the SAFNWC cloud mask, to produce an improved mask, using findings from this study which suggested that the different cloud masks perform better under different conditions. Thirdly, the estimation of the DTC under prolonged cloudy conditions, which was done using the monthly median DTC as a proxy in this study, can also be improved upon by incorporating physics-based principles to simulate the potential impact of continuous cloud cover on BT, relative to cloud-free conditions that were modelled with high levels of accuracy. Further research will relate these modelled cloud-free temperature estimates to air temperature station observations.

2.7 Acknowledgements

This work was primarily supported by the USAID Famine Early Warning Systems Network Project. The authors are grateful to the UCSB CHG for supplying the TIR BT data, and to SASSCAL for providing the coordinates for the stations station locations that were used in this study. The authors also express their gratitude to independent reviewers who provided internal reviews for the manuscript before submission.

CHAPTER 3: USE OF GEOSTATIONARY THERMAL INFRARED DATA FOR IDENTIFYING TEMPERATURE EXTREMES

3.1 Abstract

Climate change has caused an increase in extreme warm air temperatures and is expected to continue to do so into the future. These high air temperatures can have an adverse effect on human health, livestock and agricultural yields. Air temperature data required for monitoring the occurrence of extreme air temperature events, however, are inadequate due to sparse meteorological station networks with inconsistent reporting, and opportunities exist for supplementing extreme air temperature monitoring using satellite data. Statistical relationships between modelled, cloud-free thermal infrared brightness temperature (BT) data and 2-m air temperatures were explored in this study, focusing on four months of the year, at four representative locations in southern Africa. Variations were observed in the relationship between the diurnal temperature range (DTR) of the modelled BTs and the 2-m air temperatures, with the BTs tending to show a much larger range under dry conditions. Regressions between the monthly modelled BTs and 2-m air temperatures indicated R-squared values of 0.73 and 0.61 for the mean and the DTR respectively. A linear function was developed using the BT and air temperature monthly statistics to calibrate the modelled BTs to 2-m air temperatures. The frequency of occurrence of temperature > 30 ℃ was calculated for the uncalibrated and calibrated BTs, and then compared with the frequency of occurrence for the original 2-m air temperature data. High temperature frequency estimates for the calibrated BTs had an R-squared of 0.91 when compared with the 2-m air temperature equivalent, and a Heidke skill score of 0.50. A Monte Carlo approach was used to examine the effect of using lower accuracy temperature data to calibrate the BTs. Modest reductions in accuracy (10%) still resulted in reasonably accurate frequency estimates. The

use of simple statistical techniques for calibration of BTs averts the high computational power required for radiative transfer models traditionally used for land surface temperature (LST) retrieval. The differential effects of atmospheric attenuation due to water vapour can be incorporated to reduce inter-seasonal variation in the air temperature/BT relationship.

Keywords: *air temperature extremes; brightness temperature; calibration; diurnal temperature range; southern Africa*

3.2 Introduction

Climate change will likely result in an increased frequency of extreme warm air temperatures (Stocker et al. 2013). Climate trend analysis at a global scale has shown an increase in the number of days in which maximum air temperatures exceed the 90th percentile (Seneviratne et al. 2012). Panda et al. (2017) observed an increasing trend in the frequency and duration of heat waves in India between 1951 and 2011. Climate projections by the IPCC (Stocker et al. 2013) estimate that at a global scale, the frequency and duration of such warm spells and heat waves are "very likely" to increase by the late 21st century.

These extremes can affect crops, livestock, and pregnant mothers, and increase mortality rates. Over 2000 excess deaths were recorded in England in August 2003 during a heat wave which lasted several days (Johnson et al. 2005). Wang et al. (2013) found an increase in the percentage of preterm births in Australia during heat waves. Heat stress, as expressed by temperature-humidity indices, negatively affects the performance of livestock such as dairy cattle (West 2003). Large parts of South Africa and Namibia have been estimated to be at risk of livestock heat stress (Du Preez et al. 1990). Excessive air temperature also affects maize production. Maize is a staple crop

for most countries in southern Africa. By comparing data from over 20,000 maize trials with concurrent daily weather data, Lobell et al. (2011) showed that maize experienced yield reductions of between 1 and 1.7% per each growing degree day above the optimal air temperature of 30 ℃. Examining health data, Grace et al. (2015) found significant relationships between an increased number of days with warm temperatures and increased frequencies of low birth weight deliveries, while Davenport et al. (2017) found that increased warm days were associated with increased rates of childhood stunting.

Unfortunately, monitoring extreme temperatures in sub-Saharan Africa remains a significant challenge. Good meteorological observation networks and high temporal resolution air temperature datasets are important for analysing the historical trends of air temperature extremes, thereby allowing the development of appropriate policy response options. Operational observation networks also enable the monitoring of these extreme events, facilitating timely emergency response. Much of Africa, however, has sparse networks of meteorological observation stations, many of which do not report regularly on the World Meteorological Organization's Global Telecommunications System, and the number of observing stations has been decreasing over the last few decades due to a diverse set of challenges (Snow et al. 2016). Accessibility and usage of the available data is sometimes further restricted by the sale of data (Vickery 2011), which many national meteorological agencies implement as part of a policy of cost recovery (Anderson et al. 2015).

Given the current poor state of in-situ observation networks, better satellite-based monitoring can assist in improving disaster risk reduction processes and mechanisms. For example, use of the CHIRPS global precipitation dataset (Funk et al. 2015) facilitated recognition that the 2015/2016

southern Africa drought was the driest that had been experienced in over 35 years across large swaths of the region, thus effectively contributing to a coordinated regional emergency appeal (Magadzire et al. 2017).

There are many products which use geostationary thermal infrared (TIR) data to estimate precipitation, but few estimate air temperature. The recently developed Climate Hazards Group Tmax (CHTmax) global temperature dataset (Funk et al. in internal review) shows that monthly maximum brightness temperatures (BT) in December 2015 were several standard deviations higher than the mean, indicating the occurrence of an extreme event (Figure 3.1). However, no datasets are available to show the number of days or hours for which BTs or air temperatures were above a critical threshold for various functions such as human safety and animal or plant productivity.

In this work, we explore applications of TIR data to map areas that experienced very warm conditions. We perform a detailed analysis of the relationships between the TIR data and automated weather stations (AWS) data at a few representative locations, and develop a methodology for estimating the frequency of high air temperature events (per month) that can negatively impact agricultural production.

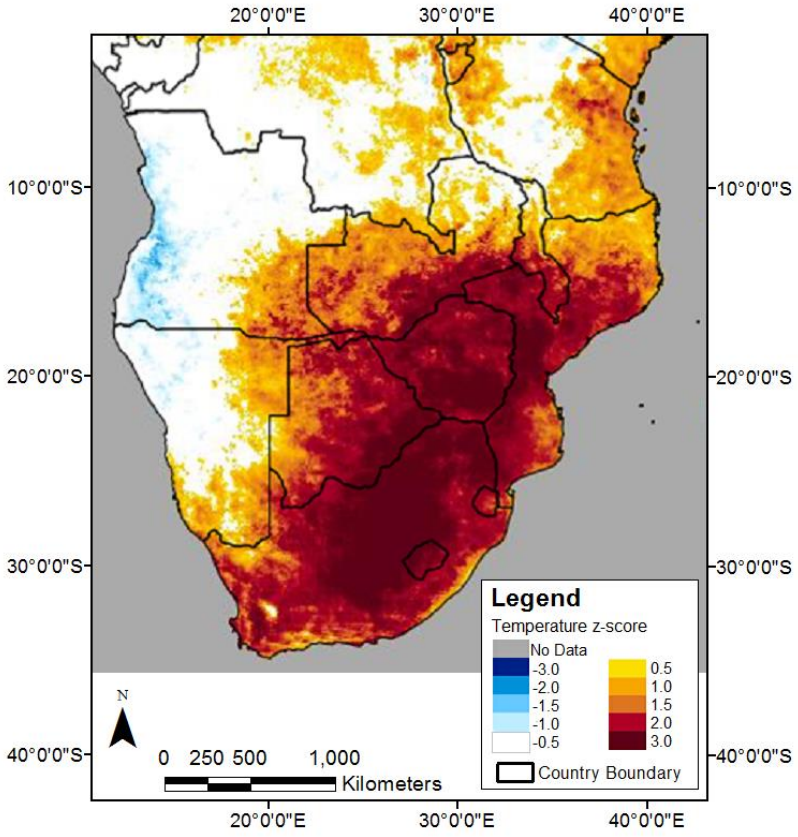


Figure 3.1 CHTmax estimate of maximum BT for December 2015, expressed as a z-score

3.2.1 Cloud-free TIR-based brightness temperature

The presence of clouds in satellite TIR data restricts the effective use of BTs for observing and analysing land-surface related variations in temperature. A system was developed in Chapter 2 for masking out clouds and modelling the cloud-free diurnal temperature curve (DTC) using half-hourly TIR BTs processed by the National Oceanic and Atmospheric Administration, or NOAA (Janowiak et al. 2001). This system used sinusoidal and exponential functions to model daytime and nighttime temperature variation respectively, and its validated results indicated a mean absolute error (MAE) of 1.1 ℃ when compared with cloud-free satellite BT data. The work undertaken in this chapter will build on the modelled DTC variations.

3.2.2 Effects of emissivity and atmospheric radiance on temperature

Satellite sensors that provide estimates of temperature measure scene radiance which is converted into temperature using the inverse of the Planck function (Weinreb and Han 2011). As the scene radiance measured by the sensor incorporates all the interactions from the surface and the atmosphere, it is referred to as top of atmosphere (TOA) radiance. The temperature thus derived is not the surface temperature, but the BT. The conversion from BT to land surface temperature (LST) using physical principles necessitates incorporating the effects of emissivity, as well as atmospheric transmittance, emission and scattering using radiative transfer functions. Li et al. (2013) provide a thorough treatment of this topic, including several different physics-based methods that have been used for LST retrieval. The atmospheric and emissivity effects are sizeable. Schädlich et al. (2001) found LST errors to increase with increasing LST and decreasing emissivity. Caselles et al. (1997) stated that atmospheric transmittance associated with water vapour can range from over 90 to under 40% for a dry and a humid atmosphere respectively. Radiative transfer modelling software such as MODTRAN (Berk et al. 1999) have traditionally been used to derive land surface temperature from TOA BT (Schädlich et al. 2001, French et al. 2003), using inputs such as the total atmospheric water vapour content (TAWV) to derive the atmospheric radiative influence. These methods are however computationally intensive, making them less practical for large scale operational work (French et al. 2003). More recently, statistical methods have been successfully used to derive LST (Duguay-Tetzlaff et al. 2015). Statistical methods have also been used to estimate 2-m air temperature from LST (Janatian et al. 2017).

This work continues on the current research trend of using statistical techniques to investigate relationships between BT, LST, and air temperature. In particular, we explore simplified statistical methods for deriving air temperature-related parameters directly from BT.

3.3 Data and methods

3.3.1 Data

Cloud-free modelled half-hourly TIR BT developed in Chapter 2 with NOAA CPC satellite TIR BT data (Janowiak et al. 2001) were used at the locations of four AWS from the Southern African Science Service Centre for Climate Change and Adaptive Land Management (SASSCAL) Weathernet station network (Kaspar et al. 2015). DTCs for TIR BT were derived in Chapter 2. Table 3.1 provides background information on these stations. Data for the months of January, April, July and October were used, thereby enabling the representation of the major seasons of the year. The data were for the years 2014 through 2016 inclusive, excluding January 2014 and October 2016 for some of the locations, in order to match contemporaneous availability of the SASSCAL data.

Table 3.1 Weather station locations and descriptions (used from Table 2.1)

Station name	Country	Longitude, Latitude (°E, °S)	Altitude (m)	Local description	Land cover classification
Kessua	Angola	16.285,-9.464	1115	600 m away from a large (300 m width) seasonable river or floodplain. Surrounded by natural vegetation and fields, and a few buildings 60 m away	Deciduous broadleaf forest/ Savanna/ Water body
Goodhope	Botswana	25.427,-25.460	1275	Surrounded by sparse savanna grasslands and scattered buildings. Approximately 5 km away from large crop agriculture fields	Grassland

Serenje	Zambia	30.215,-13.227	1395	Within 1000 m radius to large agricultural fields, savanna grasslands, urban settlement and small river	Dryland cropland and pasture/ Savanna
Zambezi	Zambia	23.108,-13.534	1066	Located at airport, surrounded by large cleared area, small agricultural fields, housing settlement and large perennial river 1100 m away	Savanna

3.3.2 Methods

3.3.2.1 Exploring statistical relationships between brightness temperature and air temperature

Monthly median values for each half-hourly time step were calculated for cloud-free satellite-observed TIR BTs, modelled TIR BTs, and observed SASSCAL 2-m air temperatures, for each month of each year, and each of the four station locations used in the study. Median values were used in preference to mean values in order to reduce the influence of extreme values which would be included in the mean. A total of 38 data points were available for the stations, months and years analysed, with a data point being represented by a station-month-year combination. Months which had observations missing in the SASSCAL data were excluded from the analysis. The SASSCAL data used was at a half-hourly time step, for comparability with the BT data. No quality control procedures were applied on the SASSCAL data, apart from checking the highest and lowest values in the data series, which were found to be within an acceptable range. Correlations between the SASSCAL 2-m air temperature and modelled TIR BTs were calculated, for both mean monthly temperature, and monthly temperature range, expressed as the difference between the 5th and 95th percentile temperatures.

3.3.2.2 Calculating frequency of extreme temperature events

For each month, year and station location the total number of half-hourly observations that were above a threshold of 30 ℃ were counted and calculated as a fraction of the total number of half-hourly observations. This frequency of warm temperature occurrences was calculated for 2-m air temperature data as well as modelled TIR BT. A calibration function was developed using the ratios of the monthly temperature ranges for the 2-m air temperature and the modelled TIR BTs, as well as the monthly mean 2-m air temperature and monthly mean TIR BTs.

$$k = \frac{T_{S95} - T_{S05}}{T_{M95} - T_{M05}}, \tag{3.1}$$

$$T_{MC} = \overline{T}_S + k \times (T_{MB} - \overline{T}_M), \tag{3.2}$$

where k is the scale factor for calibrating BT to air temperature, T_{S95} and T_{S05} are the 95th percentile and 5th percentile observed station temperatures respectively for the month, T_{M95} and T_{M05} are the 95th percentile and 5th percentile modelled TIR-based BT respectively for the month, T_{MC} is the modelled and calibrated TIR BT, $\overline{T}_S$ is the mean station temperature for the month, $\overline{T}_M$ is the mean modelled TIR BT for the month, and T_{MB} is the modelled BT. The calibration parameters were calculated separately for each station, month and year.

A Monte Carlo approach was used to examine sensitivities to uncertainties in the estimates of the 'true' local mean and standard deviation of the air temperatures. In this approach, the mean and

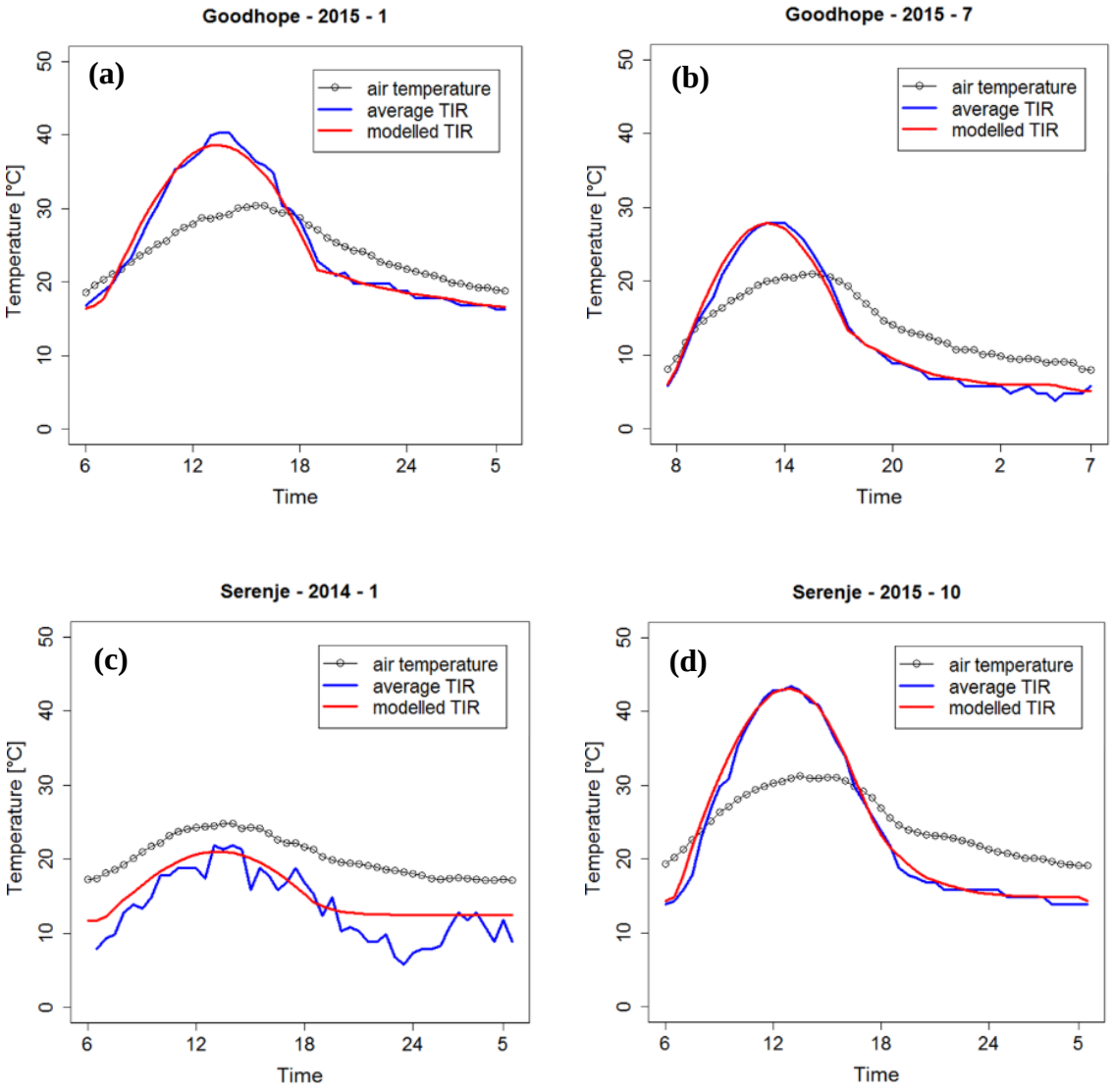
standard deviations of the underlying station observations were randomly perturbed by 10 or 20 percent, then the frequency of warm temperature occurrence was re-calculated using modelled TIR BT DTC temperatures based on the perturbed statistics. These perturbed estimates were developed to simulate the usage of air temperature means and standard deviations at points where local stations are not available, the normal scenario in remote sensing applications.

3.4 Results

Considerable variations were observed in the relationship between modelled TIR BTs and station 2-m air temperature, from month to month, and between stations.

For the Goodhope station, the monthly median TIR diurnal temperature range (DTR) was consistently larger than the 2-m air temperature range, the TIR maximum temperature was consistently higher than the 2-m air maximum temperature, and the TIR minimum temperature was lower than the minimum 2-m air temperature, for all the months for which Goodhope data were analysed, namely Jan, Apr, Jul and Oct. Figures 3.2a and 3.2b depict this pattern for Jan and Jul 2015 respectively. In contrast, Serenje, Kessua and Zambezi stations showed greater month-to-month variability, with a tendency for larger DTR for TIR BTs than for 2-m air temperature in July and October. TIR DTCs frequently had a higher maximum and a lower minimum temperature; while the TIR BT and 2-m air temperature DTRs for these stations were comparable in January and April. This is illustrated by Figures 3.2c and 3.2d. The January 2014 Serenje (Figure 3.2c) TIR DTR had a similar magnitude to the 2-m air temperature DTR, with a TIR DTC consistently lower than the 2-m air temperature DTC, while the TIR DTR for Serenje in October 2015 (Figure 3.2d) was much larger than that for the 2-m air temperature, with a lower minimum temperature and a higher maximum temperature. Large variations in DTR were also observed within the same month

at different locations, as illustrated by Figures 3.2e and 3.2f, which show the monthly average DTCs for April 2016 for Goodhope and Kessua respectively: in this instance, Goodhope's TIR DTC had a larger DTR than the 2-m air temperature DTC, while Kessua's DTRs are similar, with a TIR DTC consistently lower than the 2-m air temperature DTC. In other cases, the relationship between TIR and 2-m air temperature DTCs for a given station and given month exhibited inter-annual variation, as illustrated by Figures 3.2g and 3.2h, in which Zambezi's January 2015 TIR DTC was consistently lower than the equivalent 2-m air temperature DTC, with a comparable DTR, while the January 2016 TIR DTC was larger than the contemporaneous 2-m air temperature DTR, with a lower minimum temperature and a higher maximum temperature. In all cases analysed, at a monthly scale, the TIR minimum temperature was less than that for air temperature, except for July in Zambezi. The full set of median monthly DTC for observed station air temperature, cloud-free TIR temperature, and modelled TIR temperature that were developed for this study is presented in Appendix B.


Goodhope - 2015 - 1
(a)
air temperature
average TIR
modelled TIR
Temperature [°C]
Time
Goodhope - 2015 - 7
(b)
Serenje - 2014 - 1
(c)
Serenje - 2015 - 10
(d)

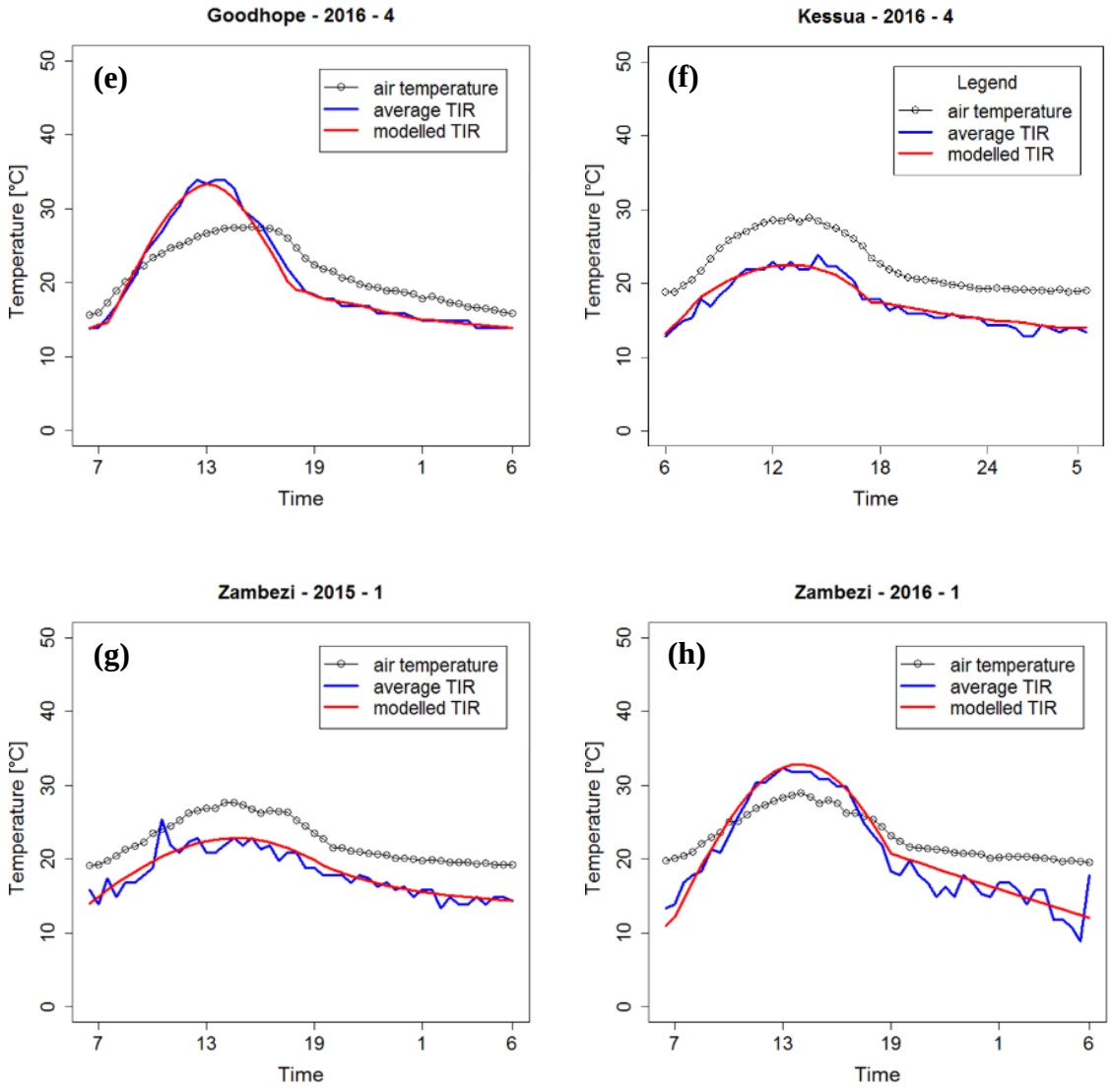


Figure 3.2 Comparison between median monthly DTC for observed station air temperature, cloud-free TIR temperature, and modelled TIR temperature for selected station locations and months, namely (a) Goodhope in January 2015 and (b) July 2015, (c) Serenje in January 2014 and (d) October 2015, (e) Goodhope in April 2016, (f) Kessua in April 2016, and Zambezi in (g) January 2015 and (h) January 2016

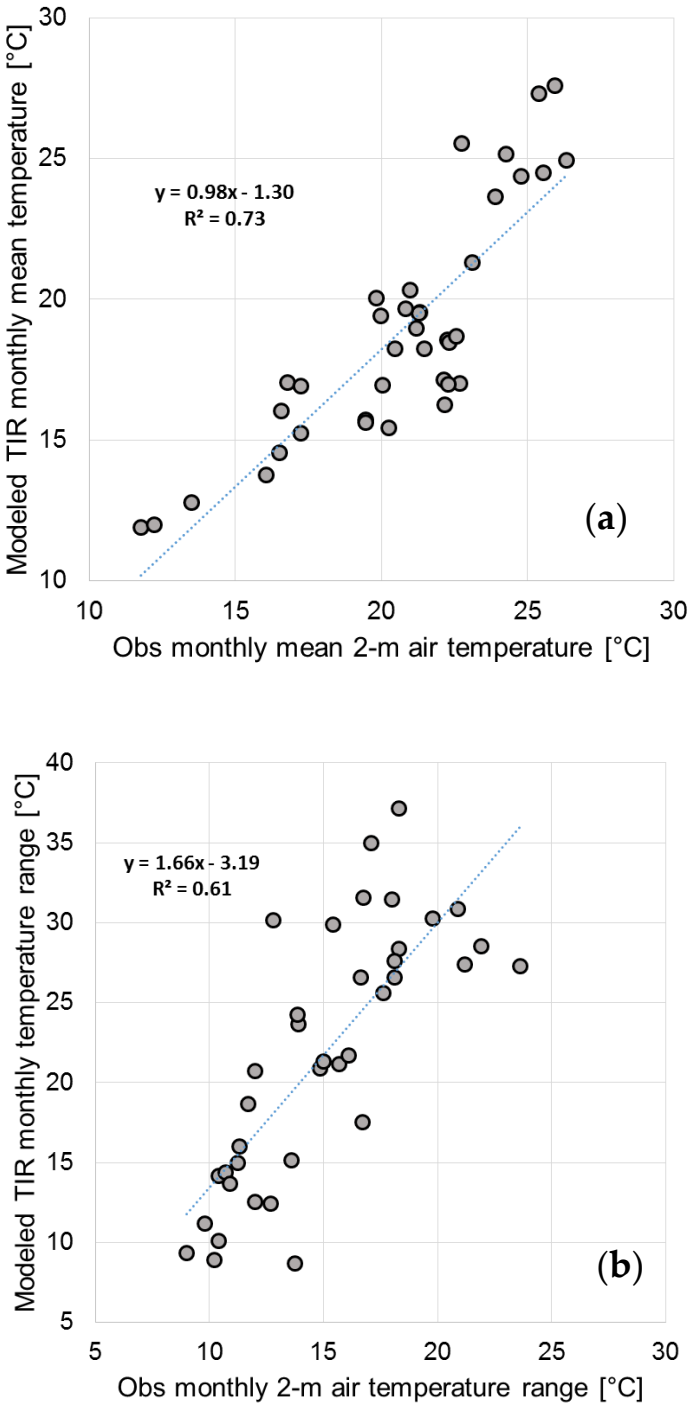


Figure 3.3 Scatterplot of monthly 2-m air temperature (observed) and modelled monthly TIR temperature for (a) monthly mean temperature and (b) monthly temperature range

The TIR-based modelled BTs were well correlated with 2-m air temperature. A comparison of the monthly modelled TIR BTs with monthly 2-m air temperatures, for both mean temperatures (Figure 3.3a) and temperature range (Figure 3.3b) produced R-squared values of 0.73 and 0.61 respectively, both of which were significant at $p < 0.01$.

Modelled BT overestimated the frequency of warm temperature occurrences when compared with station data (Figure 3.4a). The slope of the regression line was 1.86, with a statistically significant R-squared value of 0.70. After calibrating the modelled BTs as outlined in equations (3.1) and (3.2), the frequency of warm temperature occurrences estimated from calibrated BTs resulted in an R-squared of 0.91 against the station-derived frequencies (Figure 3.4b), with a regression slope of 1.20, indicating lower rates of overestimation. Figure 3.5 shows a comparison of frequencies of warm temperature occurrences for January, April, July and October, based on the different estimates of temperature. While the calibration was done using monthly aggregated statistics, the comparison of frequencies was based on half-hourly data.

A table of contingency statistics comparing warm temperature occurrence frequency derived from station data analysis, with estimates based on the raw and calibrated versions of the TIR-based temperatures, showed an improvement in most statistical parameters as a result of the calibration of the TIR-based temperatures (Table 3.2). The definitions and formulae of statistics shown in Table 3.2 are included in Appendix C. The raw TIR-based frequency estimates had a Heidke skill score (HSS) of 0.44, while those developed from calibrated TIR had a HSS of 0.50. Frequency estimates produced using calibrated TIR based temperatures developed with perturbed station mean and standard deviations had HSS of 0.48 and 0.18 for simulated uncertainties of 10% and 20% respectively.

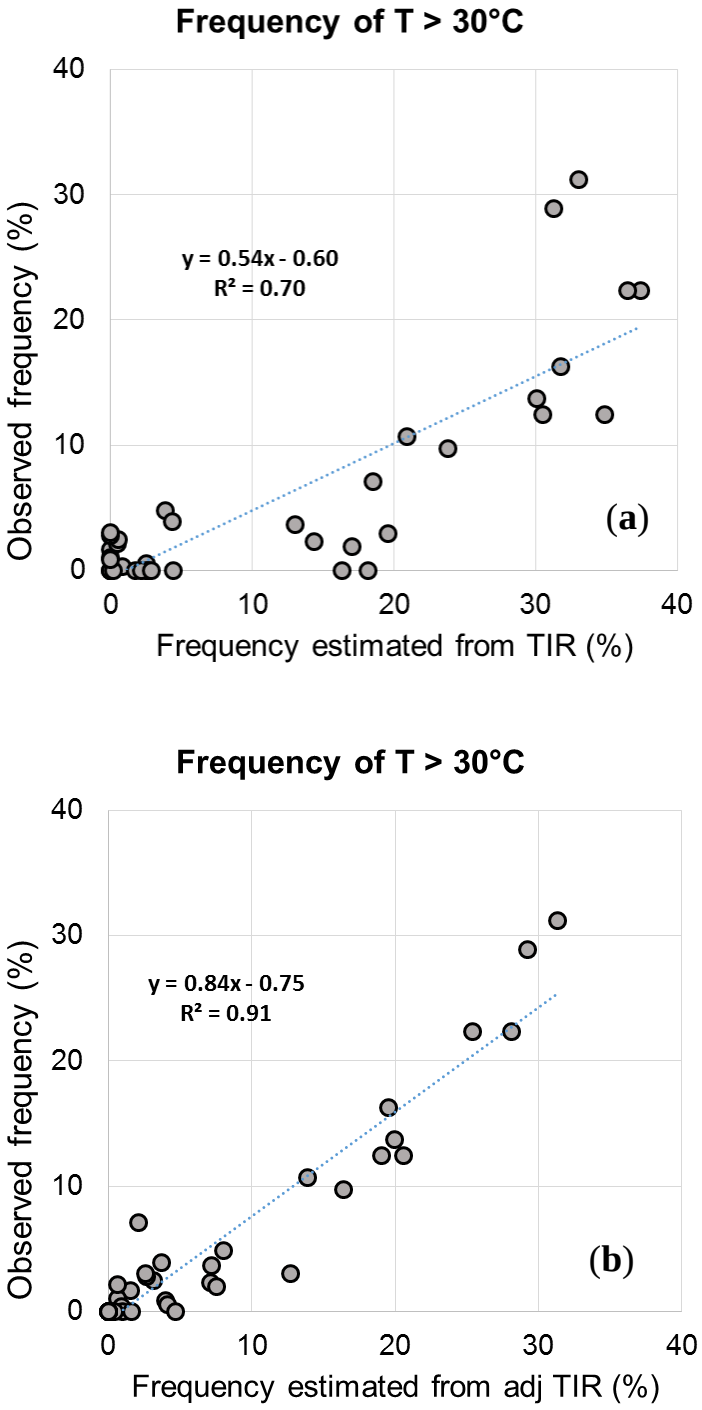


Figure 3.4 Scatter plots of observed fraction of temperature above 30 ℃ (a) before and (b) after calibrating TIR BT

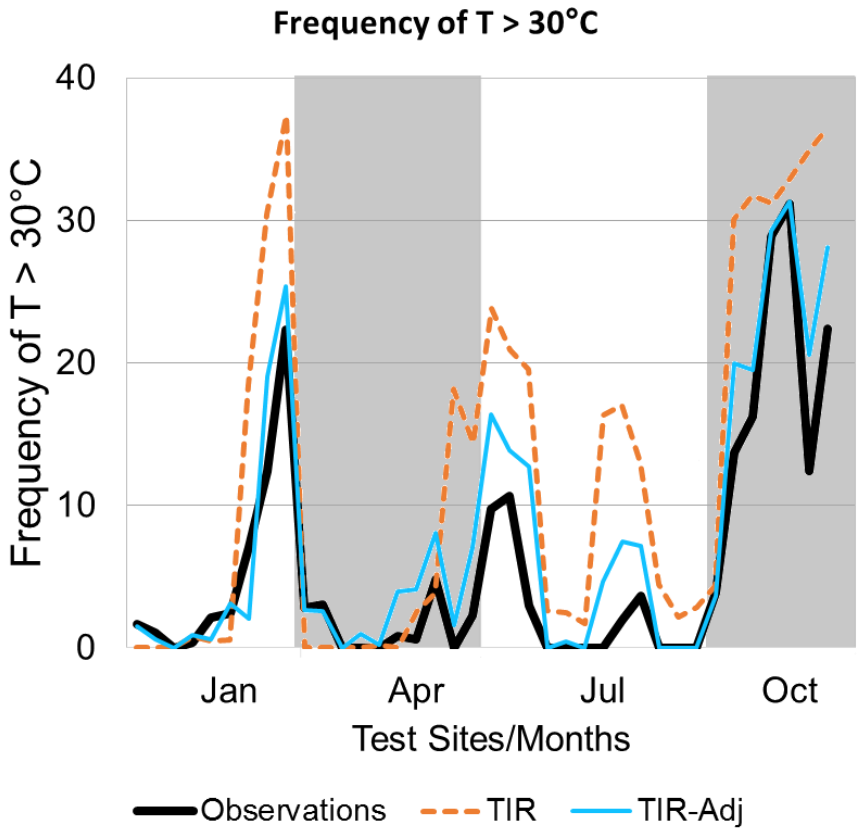


Figure 3.5 Frequency of temperature greater than 30 ℃ for the four different sites and months for observed air temperature in the solid black line, TIR BT in the orange dashed line, and calibrated TIR BT in the thin blue line. The different months are alternately colour-coded in the graph background as gray and white

Table 3.2 Contingency statistics for the amount of time (in half-hour periods) over 30 ℃ for raw TIR BT, adjusted (calibrated) TIR BT, and noise-adjusted TIR BT, using station-observed air temperature data as truth, for all months, years and locations studied

	TIR	TIR-Adj	TIR-Adj+10% noise	TIR-Adj+20% noise
Total	54184	54184	54184	54184
Hits	2357	2007	2205	1033
Misses	827	1177	979	2151
False Alarms	4189	2290	3082	4240
Correct Negatives	46811	48710	47918	46760
Accuracy	0.91	0.94	0.93	0.88
Hit Rate	0.74	0.63	0.69	0.32
Frequency Bias	2.06	1.35	1.66	1.66
False Alarm Rate	0.08	0.04	0.06	0.08
False Alarm Ratio	0.64	0.53	0.58	0.80
HSS	0.44	0.50	0.48	0.18

3.5 Discussion

The correlation of 0.73 between modelled TIR BTs and 2-m mean air temperature can be explained by the considerable variations observed in the relationship between the two parameters, from month to month, and between stations (Figure 3.2). One likely explanation for the variation in TIR temperature/2-m air temperature relationships is the difference in precipitable water content.

Analysis of NCEP Reanalysis sub-daily precipitable water content data (Kalnay et al. 1996) at pixels closest to the SASSCAL station locations was undertaken to explore this explanation. The average monthly precipitable water content (PWC) ranged from 9.2 to 33.6 kg m^{-2}. Using a MODTRAN (Berk et al. 1999) demonstration website (Spectral_Sciences_Incorporated) to plot the relationship between precipitable water content and atmospheric transmittance, unit conversion constants obtained in Liang (2005), and radiance-temperature equations developed for the Meteosat Second Generation (MSG) SEVIRI sensor (Tjemkes 2005), the PWC ranges observed in the NCEP data equated to atmospheric transmittances ranging from 0.61 to 0.91, and consequently, transmittance-induced temperature differences of between 6.4 and 36.8 °C given top-of-atmosphere radiances equivalent to 300 K. While this current study was only exploratory in nature, future work can improve the TIR – air temperature relationships by further developing the use of simple relationships that estimate the impact of atmospheric attenuation and radiance.

Another aspect explaining the variability between TIR and 2-m air temperatures relates to the air and surface properties. Seasonal swings between a dry and wet state affect the behaviour of surface heating properties due to the partitioning of incoming radiation between latent and sensible heat flux. Given the same amount of incoming solar radiation, the relative proportions of latent and sensible heat fluxes differ under varying soil moisture regimes, with lower levels of soil moisture resulting in an increase in sensible heat. The seasonal variation of vegetation abundance and greenness also attenuates the surface energy balance characteristics, including absorption and emittance of radiation (Arya 2001). The presence of vegetation and soil moisture both have the effect of reducing the diurnal surface temperature range (Arya 2001). These principles were observed in Figure 3.2, where a generally smaller TIR DTR was observed at higher rainfall stations such as Zambezi and Serenje in January, during the peak of the rainfall season, than in the dryer

months of July, April and October. The differences in specific heat capacities of air, wet soil and dry soil also play a role in the observed differences in DTCs: Farouki (1981) listed the specific heat capacities of various constituents of soil, indicating that soil mineral particles have the lowest specific heat capacity, followed by air, then soil organic matter, and finally water having the highest specific heat capacity. As a result, dry soil heats up faster than air, and very wet soils heat up the slowest. Wind and air turbulence also affect air temperature due to the influence of heat convective transfer. Land surface emissivity also affects the relationship between TIR BT and LST, which in turn influences near-surface air temperature primarily by conduction, and by convection under conditions of moderate to high wind speed. Land surfaces with lower emissivity values would have an underestimation of LST, if emissivity is not explicitly incorporated. Emissivity has been shown to increase with the normalized difference vegetation index, or NDVI (Van de Griend and Owe 1993), and therefore vegetation. This implies seasonal influences on the BT-air temperature relationship, as NDVI decreases during the winter season due to senescence. Parton and Logan (1981) also noted a larger DTR for the soil surface, compared to measurements taken at 1.5 m. This finding is similar to the observations made in several graphs in Figure 3.2. However, seasonal and inter-annual variations are apparent, as observed in Figure 3.2, due to the reasons explained above.

Despite the variable relationship between 2-m air temperature and TIR BT, the monthly means of the two datasets showed high levels of correlation (Figure 3.3a). The temperature ranges for the two datasets (Figure 3.3b) also showed significant correlation, although more scatter was apparent, and confirmed by the lower coefficient of determination. These high correlations occur despite the difference in the spatial measurement scale between the 2-m air temperature data, which is measured at a point on the order of magnitude of a few centimetres, and the TIR BT data, which

has a pixel size of 0.0363783, a spatial resolution of approximately 3.5 to 4.0 km in the study area. This difference is acceptable for temperature analysis, given that past studies have used distance lags of 25 km, and analysing air temperature semivariance for distances exceeding 500 km (e.g. Holdaway 1996). The monthly mean temperatures of the modelled TIR had an almost 1:1 relationship (Figure 3.3a) with monthly mean station temperature, with a regression slope of 0.98, suggesting that when observations are taken over the long term, modelled TIR BTs average out to similar values to those observed in air temperatures, which may be due to the conservation of energy, among other factors. The higher variability observed in the relationships between the temperature ranges of the two datasets (Figure 3.3b) is due to the factors discussed above such as vegetation and soil moisture, and is a key to the development of calibration relationships expressed in equations (3.1) and (3.2).

Of the 54,184 temperature observations that were analysed for the 4 stations, only 5.9 % of these observations had temperatures above 30 °C. This qualifies the occurrence of temperatures above 30 °C as a rare event, and the identification of these rare events is analogous to rare event forecasting evaluated by Doswell et al. (1990), for which they noted the HSS to be a good prediction skill assessment measure, with a value of 0 being no skill (chance) and a highest possible value of 1. In our study, the HSS for the detection of temperatures over 30 °C was calculated as 0.44, which improved to 0.50 after calibration using equations (3.1) and (3.2). Although we examined several measures of model performance, some, such as the accuracy (Table 3.2), overstated the performance due to the large number of temperatures below the threshold, while the objective of the analysis is the high temperatures, which are relatively rare events.

Our study used 2-m air temperature station data to calibrate the modelled gridded TIR temperatures at specific pixels for which co-located station data were available. While this approach produced accurate results, station data are not available to separately calibrate each pixel on a gridded dataset. Instead, proxy datasets such as monthly gridded interpolated station datasets, or downscaled reanalysis fields, could be used for a pixel-by-pixel analysis. These datasets typically have a lower accuracy than station data; uncertainty in interpolated data sets increases with distance from the nearest station, being dependent on the number and distance to nearby stations (Phillips and Marks 1996). Reanalysis fields use both station data and general circulation model fields, and the uncertainty in the fields is also dependent on the density of the input stations (Compo et al. 2011). As a sample indicative of reanalysis uncertainty, an ensemble spread (standard deviation) for 2-m air temperature for the NOAA-CIRES twentieth century reanalysis for 1st of January 2001 at 00 Z (Compo et al. 2011) accessed through the NOAA Earth Systems Research Laboratory Physical Sciences Division website (NOAA_ESRL), produced results in the range of 0.3 to 2 K over most parts of Africa. Barsi et al. (2003) reported errors of approximately 2 K for air temperatures in global forecasts produced by the National Centers for Environmental Prediction. With simulated uncertainties of 10%, the HSS decreased to 0.48, which was still an improvement on using the modelled TIR temperatures without any calibration. Simulated uncertainties of 20% resulted in a large HSS decrease to 0.18, which was worse than using the raw TIR, but still better than chance. In all cases of the actual and the simulated calibrations, the frequency bias was lower than that for the uncalibrated modelled TIR temperatures, suggesting that calibration, even with low accuracy datasets, provides some improvement to the analysis.

The statistical approach used in this study to estimate half-hourly air temperatures frequency effectively lumps together the conversion from BT to LST, then LST to air temperature for specific

temperature thresholds, into a statistical model based on monthly-aggregated parametrization. This approach introduces a few shortcomings into the study, when compared with dynamical approaches. First, sub-monthly variations in factors that influence the BT-LST relationship, such as atmospheric transmittance and emissivity, can reduce the accuracy of the monthly statistical models at sub-monthly level. Emissivity is affected by vegetation greenness (Van de Griend and Owe 1993), which changes significantly at season green-up and senescence. Emissivity is also affect by soil moisture (Mira et al. 2007), which changes constantly during the growing season as a function of temporal rainfall distribution and evapotranspiration. Atmospheric transmittance is affected by the total atmospheric water vapour content (TAWV), which can change over a few days, thus in turn potentially changing the BT-LST relationship over the course of the month. Secondly, a number of weather variables may affect the LST-air temperature relationship. Gallo et al. (2011) found that differences between LST and air temperature were much larger under clear-sky than in cloudy conditions. Wind speed, wind direction, and frontal systems are other factors that may affect the LST-air temperature relationship at a sub-monthly time scale which may not be detected by a monthly-scale statistical model. In contrast, dynamical models incorporating these factors, as well as some diurnal-scale statistical models could incorporate the factors listed above to derive an improved accuracy air temperature estimation model.

Several improvements, which were not implemented as they were beyond the scope of this exploratory study, can be made to the methodology in future work. One of these is a more in depth study of the influence of atmospheric attenuation and atmospheric radiance on the model, to determine if this can result in an improvement in the identification of extreme temperatures. Rudimentary analysis of reanalysis TAWV data at the study sites indicated large swings in the TAWV from month to month which results in large changes in the atmospheric transmittance and

hence the at-sensor BTs. Additionally, calibration of the TIR brightness temperatures at sub-monthly intervals such as 10-day or 5-day calibrations may also result in improved accuracy in extreme temperature identification, as the TAWV changes referred to above also occur at sub-monthly levels due to the movement of weather systems.

3.6 Conclusions

A method for identifying 2-m air temperature extremes was successfully developed and demonstrated using thermal infrared brightness temperatures (TIR BTs). This was despite various physics-based factors that convolute the relationship between TIR BTs and land surface temperature, and further still between land surface temperature and 2-m air temperature, which was the primary focus of this study. The methods used showed, at a monthly level of aggregation, a high level of correlation between modelled TIR BT means, and 2-m air temperature means. The monthly mean diurnal temperature ranges of the two datasets also exhibited strong correlations. The TIR brightness vs air temperature relationships that were determined in this study have important implications for the generation of new temperature data products that can be used to analyse historical temperature trends and events as well as for operational monitoring, at the moderate spatial resolution used by geostationary satellites. In particular, the ability to identify temperature extremes at a moderate spatial resolution will be useful for informing decision-making in agriculture and health, where high temperatures can have severe impacts. The ability to translate TIR BTs into measures of extreme temperature occurrence frequency using simple statistical techniques also means that the method would be easily applicable to many systems without requiring high computational power typical of radiative transfer models traditionally used to convert BT to land-surface temperature (LST).

3.7 Acknowledgements

This work was primarily supported by USGS cooperative agreement #G14AC00042, the U.S. Geological Survey Drivers of Drought Project, and the USAID Famine Early Warning Systems Network Project. The provision of the TIR BT data by the UCSB CHG and the air temperature data by SASSCAL is gratefully acknowledged.

CHAPTER 4: INCORPORATING THE EFFECTS OF TEMPERATURE INTO A SIMPLE, REGIONAL-SCALE, CROP-SPECIFIC WATER-BALANCE MODEL AND SEASONAL MONITORING SYSTEM

4.1 Abstract

Crop-specific water balance models are regularly used for seasonal monitoring of crop conditions to provide early warning of impending weather-induced yield reductions, and subsequent potential increase in food insecurity. One such model run by the Famine Early Warning Systems Network (FEWS NET) uses rainfall and evapotranspiration as meteorological inputs to produce the water requirements satisfaction index (WRSI), with a static length of growing period (LGP) approximating the period from planting to physiological maturity. In this study, gridded minimum and maximum air temperature data were used to estimate the growing degree days (GDD), and consequently, maps showing the growth duration (in dekads) of a maize crop from planting to maturity, for each summer season in southern Africa from 1982/1983 to 2016/2017. A dynamically variable LGP based on GDDs was thus created and integrated into the WRSI calculation. A comparison between the variable LGP generated by the analysis, and literature documenting LGP in a few southern Africa countries suggests that the new method produced reasonable LGP estimates. A number of variants of the WRSI were calculated, including WRSI based on planting at the onset of rains, and WRSI aggregated from several different planting dates. Correlations between historical maize yield in seven southern Africa countries and the different variations of WRSI suggest that the use of a variable, GDD-based LGP in the water balance calculation improves the ability of WRSI to estimate yield, especially after incorporating multiple planting dates into the model scheme, to better simulate local farming practices. Calibration of the soil water holding capacity was also found to improve the variable-LGP-based WRSI's performance

relative to historical yield in some instances. An accumulation of GDDs developed using extreme air temperatures of at least 30 ℃ was noted to have significant correlations with historical yields across several sub-national administrative units, and also indicated high correlations with WRSI. A compound parameter that integrates water stress and heat stress impacts was developed, and indicated high correlations with yield. Several of the temperature-based parameters indicated improved correlations with yield, compared to the original, static-LGP-based WRSI. More work is required to calibrate and refine the temperature-based monitoring parameters at local, sub-national scales.

Keywords: *calibration; extreme temperature; growing degree days; length of growing period; thermal time; water requirements satisfaction index; WRSI; CSWB; southern Africa*

4.2 Introduction

Simple, inexpensive and moderately accurate agrometeorological methods for monitoring crop growing conditions and their implications for crop yields, are an essential component of early warning systems for food insecurity assessments in developing countries. The water requirements satisfaction index (WRSI) calculation is one such method: based on a simple, crop-specific water balance (CSWB) model for assessing the effect of rainfall distribution and atmospheric moisture demand on crop performance, the index is calculated as the ratio between cumulative seasonal crop water supply and demand, expressed as a percentage (Senay and Verdin 2003). The Famine Early Warning Systems Network (FEWS NET) operationally runs the CSWB model and uses the WRSI for routine seasonal monitoring, enabling the detection of drought conditions that may potentially reduce crop yields (Senay and Verdin 2003). The CSWB model for calculating WRSI uses rainfall and short grass reference evapotranspiration (ETo) as the primary meteorological

inputs. The crop evapotranspiration (ETc) is deduced from ETo using crop coefficients that define the ratio between ETo and ETc at different stages of the crop's phenology (Allen et al. 1998). A fixed crop length of growing period (LGP) for the crop, such as 120 days is generally used in the calculation of the WRSI, as in Verdin and Klaver (2002), Senay and Verdin (2003), and Mukhala and Hoefsloot (2004). In their operational CSWB model for calculating the WRSI, USGS (2007) use a static LGP calculated for each pixel as the period for which the climatological rainfall is greater than a set fraction of the climatological potential evapotranspiration. The crop coefficients for the different crop growth stages are assigned to the model based on the fraction of the crop cycle length that has elapsed (Senay and Verdin 2003). WRSI has been used both as a qualitative indicator of yield (Rojas et al. 2005), as well as for quantitatively estimating crop yields (Manatsa et al. 2011). Reynolds et al. (2000b) noted the use of a fixed LGP as one limitation of the CSWB model, and suggested the use of temperature information from Meteosat satellite images to estimate crop cycle length.

The impacts of air temperature on plant phenological development rate and consequently crop cycle length are well documented (McMaster and Wilhelm 1997, Trudgill et al. 2005, Cleland et al. 2007). Thermal time, referred to in some literature as heat units (McMaster and Wilhelm 1997), and typically expressed in units of growing degree days (GDD), or just ℃ day (Soler et al. 2005) is the accumulated difference between the daily average air temperature and a base temperature below which no plant growth occurs (Moot et al. 2000). Plant growth accelerates as temperature increases between the base temperature and an optimum temperature (Cleland et al. 2007), making thermal time a good predictor for phenological development rate. As the plant development rate is retarded above the optimum temperature, the GDD value in some models decreases linearly with air temperature above the optimum temperature threshold, until it reaches zero at a critical

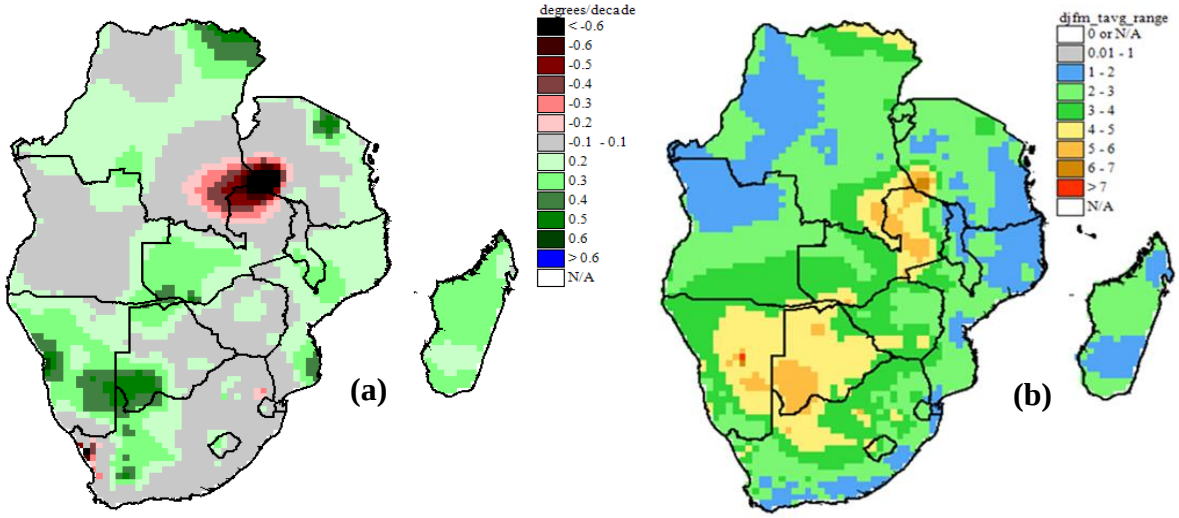
temperature above which no plant development occurs (Jones et al. 2003, Harrison et al. 2011). The effect of air temperature on crop phenological development is often used in deterministic crop simulation models such as AquaCrop (Steduto et al. 2009), APSIM (Wang et al. 2002) and DSSAT (Jones et al. 2003). A base temperatures of 10 ℃ and an optimum temperature of 30 ℃ are used in a number of applications (Reynolds et al. 2000b, Vina et al. 2004). Relationships that have been established between GDD accumulations at different crop stages, such as described by Yang et al. (2004), facilitate the modelling of crop growth through the different stages given only one published GDD parameter for a variety.

Using a simplified thermal time approach that utilizes daily maximum, minimum and mean air temperature averaged over each of the main crop phenological stages, Harrison et al. (2011) demonstrated potential reductions in crop growth period of up to 7 days in parts of Mozambique, from a 1.2 ℃ seasonal average increase in air temperature. Physical experiments by Hatfield and Prueger (2015) suggested that maize plants which experienced increased temperatures had a shortened crop cycle, with subsequent reductions in yields, particularly if extreme temperatures occurred during the grain-filling period. However, they also concluded that changes in air temperature that occurred during the vegetative period did not affect maize yields. By comparing data from over 20,000 maize trials in southern and eastern Africa, with concurrent daily weather data, Lobell et al. (2011) demonstrated that maize experienced yield reductions of between 1 to 1.7% per each GDD above the optimal temperature of 30 ℃, depending on the levels of moisture stress experienced during the flowering period. Their analysis showed little sensitivity of maize yield to temperature increase provided air temperatures remained below the optimal temperature and under conditions of optimal water availability; however, in trials where maize crops were

subjected to water stress during flowering and grain filling stage, any increase in air temperature resulted in a predictable reduction in yield.

Hansen and Lebedeff (1987) observed global increases in air temperature between the late 19th century and the late 20th century of 0.5 ℃ to 0.7 ℃. In this century-long analysis, they observed periods of both positive and negative fluctuations, but noted an increase of approximately 0.3 ℃ in the 20 years between 1965 and 1985. Stocker et al. (2013) quoted a global average air temperature rise between 1951 and 2012 of 0.12 ℃ per decade. An analysis of global monthly mean air temperature data compiled by Matsuura and Willmott (2015) for the period 1971 to 2014 showed an average increase of 0.12 ℃ per decade for the Dec-Mar (DJFM) period in mainland southern Africa, south of the 15 ° south latitude (Figure 4.1a). Climate projections by the IPCC (Stocker et al. 2013) suggest that global surface temperatures during the 2016-2035 period may be 0.3 to 0.7 ℃ warmer than the 1986-2005 reference period. Although these climate trend studies and the climate projections by the IPCC show slow changes in temperature over the long-term, the inter-seasonal variability is much higher; DJFM air temperature averages based on the Matsuura and Willmott (2015) dataset showed a range of 4.0 ℃ (Figure 4.1b) and a standard deviation of 1.0 ℃ in parts of southern Africa, between 1971-2014 period. The effect of this large inter-seasonal range in air-temperature on crop phenological development can be illustrated by considering two hypothetical seasons for a location X; season A with an average air temperature of 24 ℃, and season B with 28 ℃ average air temperature. A crop requiring 1600 GDD (base temperature 10 ℃) to reach physiological maturity would take 25 days less to reach maturity in season B than it would in season A. Late onset of seasonal rains sometimes delay planting of rainfed crops by as much as two months or more. When this happens, crops grow in a temperature regime different from the climatological conditions that occur when crops are planted closer to the average date of

planting. A short analysis of the Matsuura and Willmott (2015) dataset shows that in parts of southern and eastern Zambia, for example, crops growing in the October-November-December (OND) period on average experience seasonal average air temperatures 2 ℃ to 3 ℃ warmer than crops growing in the January-February-March (JFM) period (Figure 4.1c). Assuming corresponding 3-month average air temperatures of 24 ℃ and 26 to 27 ℃ in seasons C and D respectively, a 1600 GDD crop would take 14 to 20 days longer to reach maturity in season C than in scenario D. Using generated weather data run through a crop simulation model, Abraha and Savage (2006) showed that 2 to 4 ℃ increases in air temperature potentially reduced the crop cycle length by approximately 30 days and 50 days respectively, a result comparable with the rough calculations undertaken with the Matsuura and Willmott (2015) air temperature data.



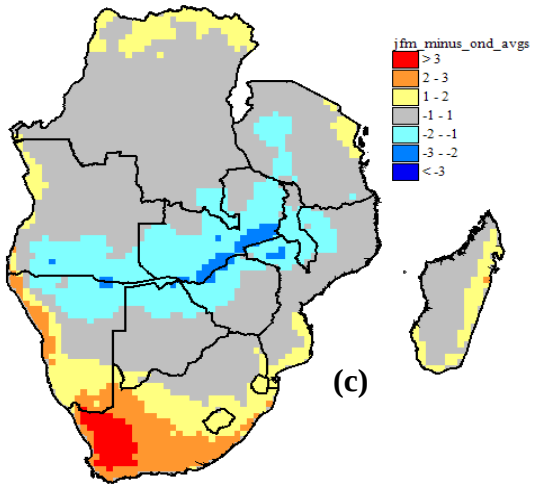


Figure 4.1 Potential causes for changes in the average air temperature conditions under which crops grow. (a) Rate of temperature change (per decade) between 1971/72 and 2013/14 for the DJFM period (b) Range of average DJFM temperatures between 1971/72 and 2013/14. (c) Changes in average seasonal temperature that occur from OND to JFM

The simplified scenarios presented above, based on the potential inter-seasonal variations in air temperature, illustrate the necessity for incorporating the impact of air temperature into the CSWB model for improved WRSI accuracy.

The CSWB model identifies four distinct stages of crop growth in which the crop coefficient, a constant defined as the ratio of a crop's water requirement to the reference evapotranspiration, can be calculated. The crop coefficient differs among crops and phenological stages. The four growth stages used in the CSWB model are the initial, vegetative (also referred to as crop development), reproductive (or mid-season) and ripening (or late season) stage (Allen et al. 1998). We refer to this schema as the descriptive crop growth schema. Allen et al. (1998) defines the initial stage as the time from planting until the crop covers approximately 10% of the ground area; the crop

development stage from 10% coverage until the crop reaches full coverage, which, for many crops coincides with the start of flowering; the mid-season stage from the start of full coverage to the onset of maturity (often signified by the onset of senescence); and the late season stage is through to harvest or complete drying of the crop, whichever comes first. An analysis of crop calendars for maize grain crops growing in six different parts of the world (Allen et al. 1998) suggests that for maize grain, the initial stage occurs over approximately the first 18% of the crop's full cycle, while the late season stage covers approximately the last 23% of the crop's full cycle.

An alternative method for describing the development of maize through various phenological stages is the leaf collar method, which is based on a count of the number of leaves on the plant with visible collars, (Nafziger 2009, Abendroth et al. 2011). The crop stages in the leaf collar method are denoted as various vegetative stages and reproductive stages. The vegetative stages described by Abendroth et al. (2011) and Nafziger (2009) were prefixed by the letter "V", followed by a number denoting the number of leaves with visible collars. The exceptions are VE, which indicates plant emergence from the soil, and VT which indicates the emergence of maize tassels. The reproductive stages were denoted by the letter "R" followed by a stage number one through six, and represent the silking (R1) through to grain filling and maturity (R6) stage. Nafziger (2009) listed the estimated GDDs required for each stage of a maize crop that needs 2700 GDD, calculated in degrees Fahrenheit (which translates to 1482 GDD in degrees Celsius), to reach maturity. Moeletsi (2017) noted that 1340 GDD were required for an early-maturing maize variety to reach physiological maturity, 1420 GDD for a medium variety, and 1480 GDD for a medium-to-late variety.

Kranz et al. (2008) noted that the amount of water used by the maize crop starts to increase from around the V8 crop stage, reaching a maximum near the VT stage, and starts to reduce again around R4. This description compares well to the changes that occur in the crop coefficient at various crop stages, as described by Allen et al. (1998).

This research integrates temperature effects into the CSWB by using a GDD-based variable crop-growing period, accumulating thermal time from the date of planting, in order to facilitate determination of the phenological stages and calculation of the associated crop water requirements dynamically. Correspondence between the leaf collar method and the descriptive crop growth schema is established to enable the use of GDD values in the CSWB model. Relationships between extreme temperature-based GDDs, yield and WRSI are explored in order to determine methods for integrating high temperature impacts into a seasonal crop monitoring system.

4.3 Data and methods

4.3.1 Data

Gridded maximum and minimum 2-m air temperature data from the Climate Forecast System Reanalysis (CFSR) dataset (Saha et al. 2010) for the period 1979 to 2010, and from the Climate Forecast System version 2 (CFSv2) dataset (Saha et al. 2014) for the period 2011 to 2017, were used to calculate GDDs for the period 1979 to 2017. The CFSR and CFSv2 datasets are jointly referred to as CFS data in this study. A number of differences in the two datasets exist as described by Saha et al. (2014). However, this study only used data from both CFSR and CFSv2 at time lags at or very close to the datasets' initial conditions, and minimal differences between the two are expected. The fact that CFSv2 retrospective forecasts use CFSR data as initial conditions further

underlines the similarity between the two datasets (Saha et al. 2014). The data were downloaded at a 6-hourly time-step and composited to a daily time-step. The CFS minimum and maximum 2-m air temperature data were compared with Global Surface Summary of the Day (GSOD) data (NOAA_NCEI 2018), a global dataset of daily weather observations compiled from hourly and synoptic data, archived by the National Oceanic and Atmospheric Administration (NOAA) National Centers for Environmental Information (NCEI). The GSOD data were interpolated using inverse distance weighting interpolation to facilitate comparison with gridded CFS data at the station location. Daily composites of 10-m wind speed, surface downward and upward longwave and shortwave surface radiative flux, pressure, and relative humidity were also processed from the 6-hourly CFS data, and were used in conjunction with the maximum and minimum 2-m air temperature CFS data to calculate daily reference evapotranspiration (ETo) for the period 1979 to 2017. The 10-m wind speed data were converted to 2-m wind speed estimates after Allen et al. (1998), before using the wind speed data in the ETo calculation. The Climate Hazards Infrared Precipitation with Stations (CHIRPS) rainfall data (Funk et al. 2015) for the 1982/1983 to 2016/2017 season was used in conjunction with ETo from the same period for driving the CSWB model. Data grids of soil water holding capacity (WHC) and LGP for running the CSWB model were provided by USGS. The soil WHC grid was originally derived from the Food and Agriculture Organization (FAO) digital soil map of the world (Senay and Verdin 2003), while the LGP grid was developed from a climatological analysis of rainfall and ETo data (USGS 2007). Historical sub-national maize yield data from the FEWS NET Data Warehouse (FDW) (www.fews.net) were used to compare the predictive power of different WRSI and temperature-based model variations. The yield data were de-trended before they were further used for analysis.

4.3.2 Methods

4.3.2.1. Calculation of GDDs

Based on a comparison of the leaf-collar method and descriptive crop growth schema, an approximate correspondence between the two methods was established as outlined by Table 4.1. This correspondence allowed the determination of GDD fractions at critical phenological stages used in the descriptive crop growth schema, based on crop-stage specific GDD information presented by Nafziger (2009) for the leaf-collar method. This approach enabled the calculation of cumulative GDD requirements for different crop varieties, for each phenological stage (as outlined in the descriptive crop growth schema). Crop coefficients for the calculation of ETc were subsequently estimated for each dekad of crop growth, by using cumulative GDDs and interpolating between the end-points of each phenological stage.

Table 4.1 Relationship between descriptive crop growth schema and the leaf-collar method

Descriptive crop growth schema (Allen et al. 1998)	**Leaf-collar method (Abendroth et al. 2011)**	**Percentage of GDD at maturity (Nafziger 2009)**
End of initial stage	V6-V8	20-26 (average=23)
End of crop development stage	V18-VT	45-49 (average=47)
End of mid-season stage	R4 (Dough Stage)	81
End of late-season stage	R6 (Maturity)	100

The GDD approach used was a combination of approaches described by Abendroth et al. (2011) and Harrison et al. (2011). Calculation of GDDs was primarily done using equation (4.1), as described by Abendroth et al. (2011).

$$GDD = \left(\frac{T_x + T_n}{2}\right) - T_b \tag{4.1}$$

where $T_n = T_b$ if $T_n < T_b$

and $T_x = T_o$ if $T_x > T_o$

and where T_x is the daily maximum air temperature, T_n is the daily minimum air temperature, T_b is the base temperature and T_o is the optimum temperature, which is the upper limit for accumulation of GDDs. In this study, the values of T_b and T_o for maize were set to 10 ℃ and 30 ℃ respectively, after McMaster and Wilhelm (1997)

In addition to the approach described by Abendroth et al. (2011), an extreme temperature limit T_e was defined, above which plant growth rate was assumed to linearly decrease to zero at a critical temperature T_c, after the method described by Harrison et al. (2011), and represented by equation (4.2).

$$GDD = [T_o - T_b] - \left([T_o - T_b] \times \left[\frac{T_a - T_e}{T_c - T_e}\right]\right) \tag{4.2}$$

where GDD is the thermal time in growing degree days

T_o is the optimum temperature, as defined in equation (4.1)

T_b is the base temperature for accumulation of GDDs

T_e is the extreme temperature limit above which photosynthesis starts to decrease

T_a is the average temperature, calculated as an average of T_x and T_n

T_c is the critical temperature above which no photosynthesis occurs

and $T_a = T_c$ if $T_a > T_c$

The values of T_e and T_c were set to 34 ℃ and 45 ℃ respectively. Equation (4.2) was only valid when $T_a > T_e$

GDDs were thus calculated on a daily time step using equations (4.1) and (4.2) for the CFSR and CFSv2 modelled data for 1979 to 2010 and 2011 to 2017 respectively. The calculated GDDs were summed to dekadal totals, for compatibility with the CSWB model, which is run on a dekadal time step. To assess the feasibility of attaining reduced data storage and improved computational efficiency in the calculation, growing degree dekads (GDDekads) were calculated using equations (4.1) and (4.2) on a dekadal time step. To this end, dekadal T_x and T_n were calculated as averages of all the daily maximum and daily minimum temperatures, respectively, in the dekad. The dekadal T_x and T_n values were then used in equations (4.1) and (4.2) to calculate the GDDekads. The calculated GDDekads were compared with GDDs summed to a dekadal time step. Data was extracted for each dekad from January 1979 to December 2017 at 25 selected point locations spread across varying elevations and ecologies, for a total of 35,100 data points. The locations of the points are shown in Appendix D

4.3.2.2 Calculation of the CSWB model output parameters and comparison with crop yield

The CSWB model was run in a gridded analysis environment using two primary variations that were subsequently compared. The first variation replicated the methodology used in the USGS operational CSWB model (USGS 2007). Characteristics of this methodology included the use of (1) dynamic calculation of the onset of rains, based on a 3-dekad rainfall analysis, for triggering planting of the crop in the model, (2) gridded rainfall and ETo grids as meteorological inputs to the model, (3) soil WHC derived from the Food and Agriculture Organization (FAO) digital soil map of the world, (4) FAO-defined crop coefficients for calculating crop water requirement, and (5) a static LGP determined from climatological analysis of rainfall and ETo data. The second variation used all the other aspects of the first variation except the LGP. In the second variation, a variable LGP was calculated by accumulating GDDs from the planting dekad until the dekad when the heat units required to reach crop maturity were met.

Extreme temperature GDDs (eGDDs) for highlighting the cumulative impact of very high temperatures on crop productivity were also calculated using equation (4.1), as described by Lobell et al. (2011), setting T_b to 30 ℃ and T_o to ∞. The calculated eGDDs were categorized as follows for each season

- Cumulative eGDDs for the growing season
- Cumulative eGDDs occurring during periods of crop water deficit
- Total eGDDs for each crop stage, namely initial, vegetative, flowering and ripening stage of the crop growth
- Total eGDDs during periods of crop water deficit for each crop stage

In addition to the above-mentioned parameters, a new parameter was developed by multiplying the water deficit expressed as a percent of the crop water requirement, by the eGDD. For the purposes of this study, this parameter was referred to as deficit-adjusted eGDD, or DxeGDD. The DxeGDD were accumulated for each season from 1982/1983 to 2016/2017.

The historical crop yield data that was obtained from the FDW is aggregated at the level of administrative units such as districts and provinces, while the CSWB model outputs were gridded data. This necessitated further processing of the CSWB model output parameters, in order to make them directly comparable with the administrative-unit-wide crop yield data. The processing required entailed calculating the district or province-wide weighted spatial average value of a model parameter such as WRSI and GDD. The weighting was based on the fraction of cultivated area in each pixel in the administrative unit. The fraction of cultivated area per pixel of the analysis model was derived from the Globeland30 land cover map (Chen et al. 2015), a high resolution land cover map produced at 30 m pixel size. The Globeland30 map was first resampled to 500 m pixel resolution using a nearest neighbour resampling approach to improve data processing, then the number of 500 m pixels in each 0.1 degree grid cell was calculated. The Global Administrative Unit Layers (GAUL) maps (http://www.fao.org/geonetwork/srv/en/main.home) compiled by the Food and Agriculture Organization (FAO) of the United Nations (UN) were used to calculate the spatial averages for each administrative unit. The correlation of outputs from the different model variations with the FDW maize yields was assessed for Botswana, Lesotho, Malawi, South Africa, Swaziland, Zambia and Zimbabwe.

4.3.2.3 Incorporating variability in planting dates and soil management into the CSWB

The LGP is generally considered as the period over which there is sufficient moisture for crops to grow successfully, often defined as the number of days for which rainfall is higher than some fraction of evapotranspiration (De Pauw et al. 1996). This method, used by Senay and Verdin (2003) to run the CSWB model and calculate WRSI, results in length of growing periods of up to 180 days over most parts of Mozambique (Figure 4.4a). In contrast, Harrison et al. (2011) noted that some modelled late maturing maize varieties grown in Mozambique took between 100 and 120 days from sowing to maturity, while short-season varieties took approximately 90 days. The apparent disparity between these two approaches is resolved when considering that farmers in some parts of southern Africa tend to plant over a period of time that can span two months or more (Rurinda et al. 2013). To simulate this effect in the model, the CSWB modelling approach described by Senay and Verdin (2003) was modified to process a series of CSWB model runs using up to six different planting dekads each season: with the first planting coinciding with the onset of rains, and each subsequent planting dekad occurring when dekadal total rainfall exceeded 25 mm. The CSWB model runs for the different planting dates were then aggregated to calculate a single WRSI value per pixel, representative of the average field conditions over much of the humid period for the season of interest. Aggregations were tested over (a) a set of three different planting dates, and (b) a set of six different planting dates.

Soil compaction reduces the water available to crops by reducing infiltration as well as limiting root growth (Lipiec and Hatano 2003). Taylor and Brar (1991) noted a 50% reduction in water intake by maize in compacted fields compared to non-compacted fields, citing a study by Tardieu and Manichon (1987). Many conventional agricultural practices, including the use of draft animals,

implements, and tractors, cause soil compaction (Soane and Van Ouwerkerk 1994). These agricultural practices are commonly used in many parts of southern Africa. This implies that crop growing conditions under common conventional agriculture may have access to less soil moisture than implied by a CSWB model run using soil WHC estimates derived from pedo-transfer functions. WHC estimates obtained from the FAO soil map of the world (Verdin and Klaver 2002, Senay and Verdin 2003) were divided by 50, to simulate the 50% reduction in water availability for compacted soils noted by Taylor and Brar (1991). This approach facilitated a preliminary test of potential improvements that could be introduced into CSWB modelling by WHC calibration. Such improvements would allow better evaluation of the various CSWB-model-based yield estimation approaches.

4.3.2.4 Summary of CSWB-related parameters analyzed

Overall, the following parameters were analysed by comparison with yield statistics:

- $WRSI_{static\_LGP}$: WRSI calculated using a static LGP grid
- $WRSI_{variable\_LGP}$: WRSI for a single planting dekad calculated using a variable LGP based on GDD accumulation
- $WRSI_{variable\_LGP\_6onsets}$: WRSI aggregated from CSWB model runs for six different planting dekads, calculated using variable LGP based on GDD accumulation
- $WRSI_{variable\_LGP\_3onsets}$: WRSI aggregated from CSWB model runs for three planting dekads, calculated using variable LGP based on GDD accumulation
- $WRSI_{varLGP\_halfWHC\_3onsets}$: WRSI aggregated from CSWB model runs for three planting dekads, calculated using variable LGP based on GDD accumulation, and using a soil WHC equal to half the WHC of the USGS' operational model

- $eGDD_{3onsets}$: Total eGDD over the complete crop's growth period, for CSWB model runs based on three different planting dates
- $eGDD_{Dry\_3onsets}$: Total eGDD's occurring during periods of crop moisture deficit, over the complete crop's growth period, for CSWB model runs based on three different planting dates
- $eGDD_{FLW\_3onsets}$: Total eGDD's occurring during flowering period of the crop's phenology, for CSWB model runs based on three different planting dates
- $DxeGDD_{3onsets}$: Total DxeGDD over the complete crop's growth period, for CSWB model runs based on three different planting dates

Regional summaries of the performance of the different parameters were compiled for all the administrative units that had significant correlations between the de-trended yield index and WRSI values, for the different WRSI variations generated. Similar analyses were carried out for eGDD-yield correlations, and for DxeGDD-yield correlations. To allow regional level correlations to be calculated, the yield, WRSI, eGDD and DxeGDD parameters for each administrative unit were standardized by subtracting the mean and dividing by the standard deviation.

Correlation between WRSI and eGDD was also calculated in order to evaluate the covariance between the two variables, both of which had the potential to estimate yield.

4.4 Results

Daily Tmin (minimum air temperature) and Tmax (maximum air temperature) data derived from the CFSR (1979-2010) and CFSv2 (2011-2017) were compared with observed interpolated air temperature data from the GSOD stations. The results are shown in Table 4.2.

Table 4.2 Summary of validation of CFSR and CFSv2 minimum and maximum air temperature data using GSOD data

Parameter	Period [years]	R-squared	P-value	RMSE [℃]	MAE [℃]	Mean bias	Slope	Intercept [℃]	Number of points
Tmin	1979-1990	0.77	0	2.74	2.17	-0.57	0.95	0.32	568392
Tmin	1991-2000	0.78	0	2.76	2.21		0.95	0.10	566943
Tmin	2001-2010	0.82	0	2.60	2.04		0.93	0.71	708769
Tmin	2011-2017	0.81	0	2.74	2.37		0.94	-0.39	780795
Tmax	1979-1990	0.63	0	3.38	2.60	-0.07	0.99	0.24	568685
Tmax	1991-2000	0.66	0	3.46	2.64		1.07	-1.78	567090
Tmax	2001-2010	0.69	0	3.42	2.61		1.08	-2.71	709046
Tmax	2011-2017	0.72	0	3.18	2.68		1.04	-2.69	780987

The data were validated in groups of ten to twelve years for all but the last seven years in the dataset, which were validated for the period 2011-2017. The data was split into these groups of years due to limitations in Microsft Excel, the main software that was used for statistical analysis. Excel cannot handle more tha 1,048,576 rows of data at a time, and the groups of data analysed contained between 550,000 and 780,000 data points each. The coefficients of determination and mean absolute errors for 2-m minimum air temperature for each period of validation were all higher and lower respectively, than those for 2-m maximum air temperature. The CFSv2-based daily minimum and maximum air temperature captured 81% and 72% of the variation observed in the GSOD minimum and maximum air temperature data respectively. This result was comparable with the CFSR data verification, which gave R-squared values that ranged between 0.77 and 0.82

for minimum air temperature, and between 0.63 and 0.69 for maximum air temperature. The regression slope was close to 1 in most cases (Table 4.2). The regression intercept was close to 0 ℃ in most cases for minimum air temperature, while it ranged between -2 and -3 ℃ for the maximum air temperature.

A scatterplot of over 780,000 daily data points for Tmin and Tmax data comparing CFSv2 and GSOD data for the period are shown in Figure 4.2.

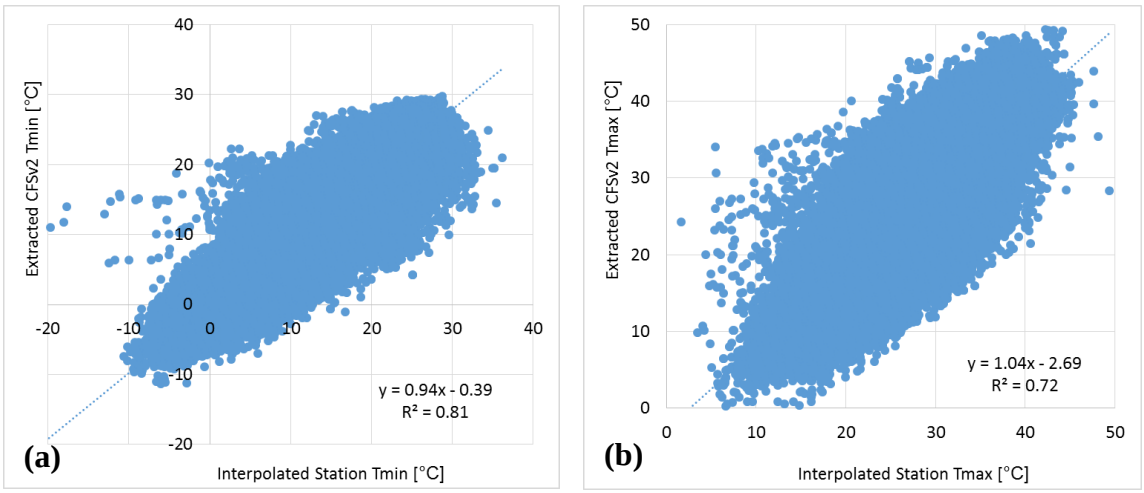


Figure 4.2 Scatterplots of (a) daily minimum and (b) daily maximum air temperature for the period 2011 to 2017, comparing interpolated GSOD station data and CFSv2 modelled data at co-located points

The comparison of dekadal total GDDs and GDDekads indicated a high coefficient of determination of 0.996 between the two parameters (Figure 4.3). The GDDekads indicated an overall mean bias of 0.92. The GDDekads tended to underestimate low GDDs, and overestimate

high GDDs. GDDekads between 0 and 100 had a mean bias of -0.52 when compared with dekadal accumulated GDDs, while those between 101 and 200 had a mean bias of 1.58.

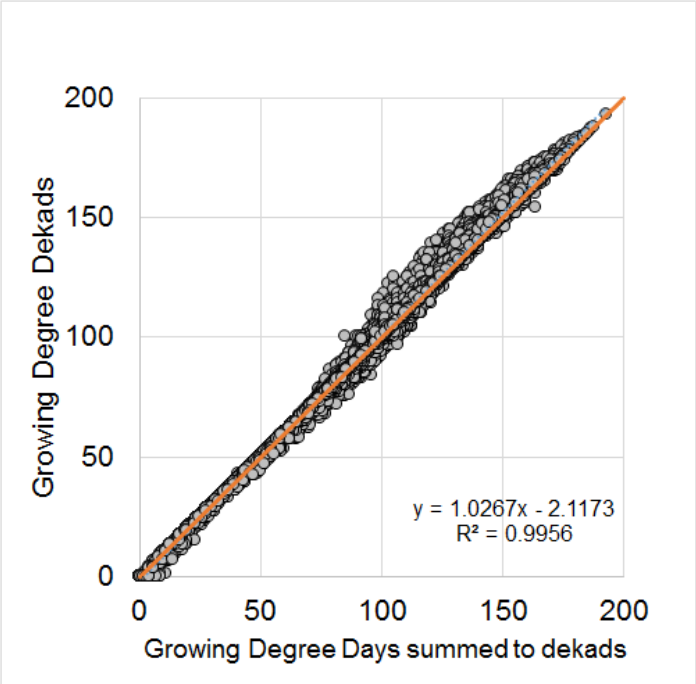


Figure 4.3 Scatterplot comparing dekadal total GDDs and GDDekads. The solid orange line is the 1:1 line

The LGP based on comparison between climatological rainfall and evapotranspiration (Figure 4.4a), which is the one used in the operational USGS WRSI version of the CSWB model, was found to be, in most areas, longer than that calculated based on accumulation of GDDs for both the early maturing (Figure 4.4b) and late maturing (Figure 4.4c) maize varieties. The operational CSWB's LGP was 18 dekads (approximately 180 days) in many parts of Angola, Democratic Republic of Congo, Lesotho, Madagascar, Malawi, Mozambique, South Africa, Swaziland, Tanzania and Zambia. In contrast, the early maturing model produced LGPs of nine to ten dekads in many areas. In nearly all the countries in the analysis areas, excluding Lesotho and South Africa,

most places had GDD-calculated LGP shorter than 14 dekads for the medium-to-late maturing varieties.

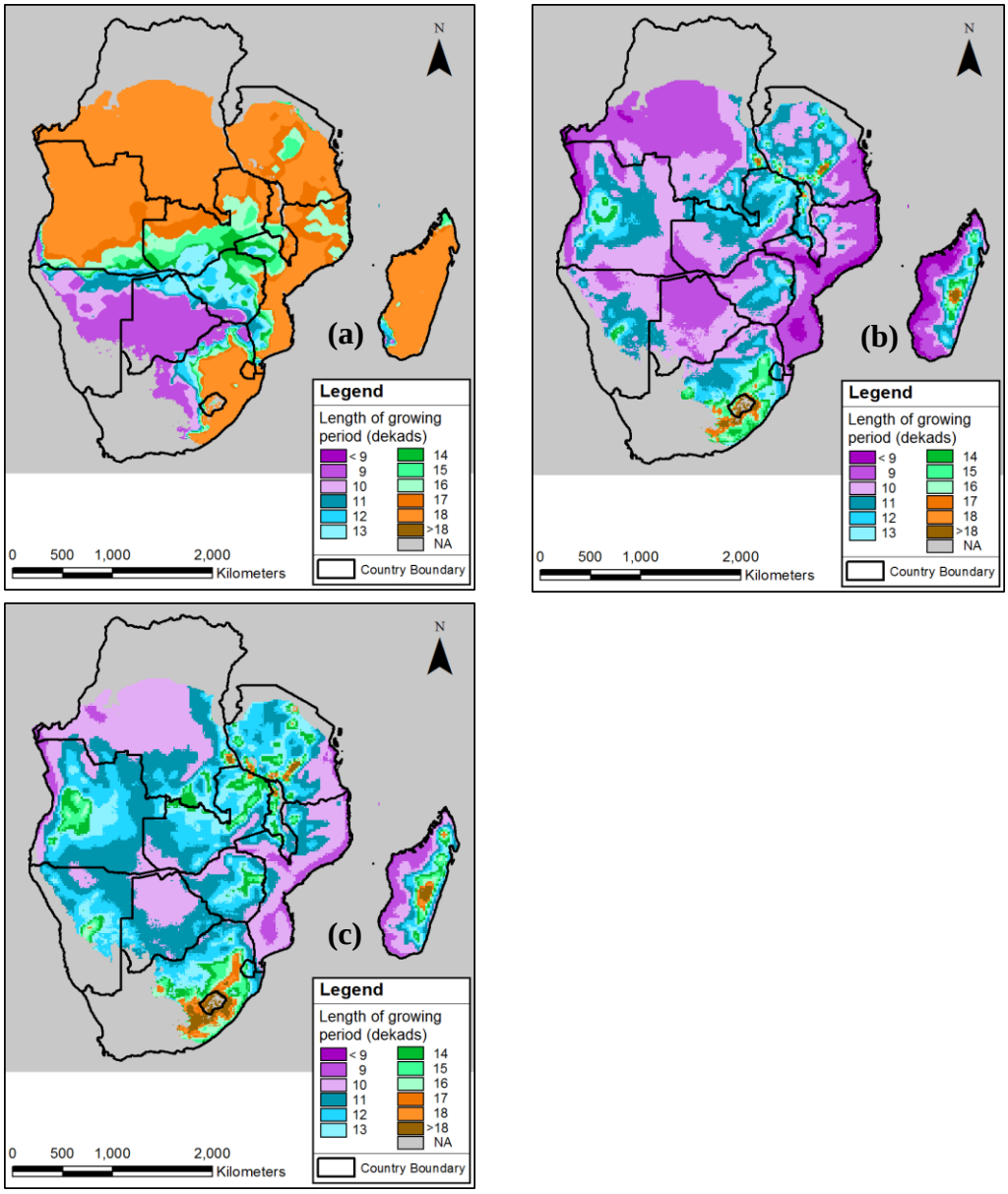


Figure 4.4 (a) Static LGP used in operational USGS CSWB model, compared with (b) average variable LGP for an early maturing maize variety with a total GDD requirement of 1340 and (c) average variable LGP for a medium-to-late maturing maize variety with a

total GDD requirement of 1482 GDD. The average variable LGP maps were calculated using air temperature data for 35 agricultural seasons from 1982/1983 to 2016/2017

Over the 35-year time series of LGP grids that was generated, some inter-seasonal variation was observed. The shortest LGP noted in the 1482 GDD analysis in most areas was between 9 and 11 dekads (Figure 4.5a), while the longest modelled LGP was at least 13 dekads in most areas, with many parts of Angola, Zambia and Zimbabwe having longest LGPs of over 14 dekads. Large parts of South Africa and Lesotho had longest LGPs of over 18 dekads (Figure 4.5b).

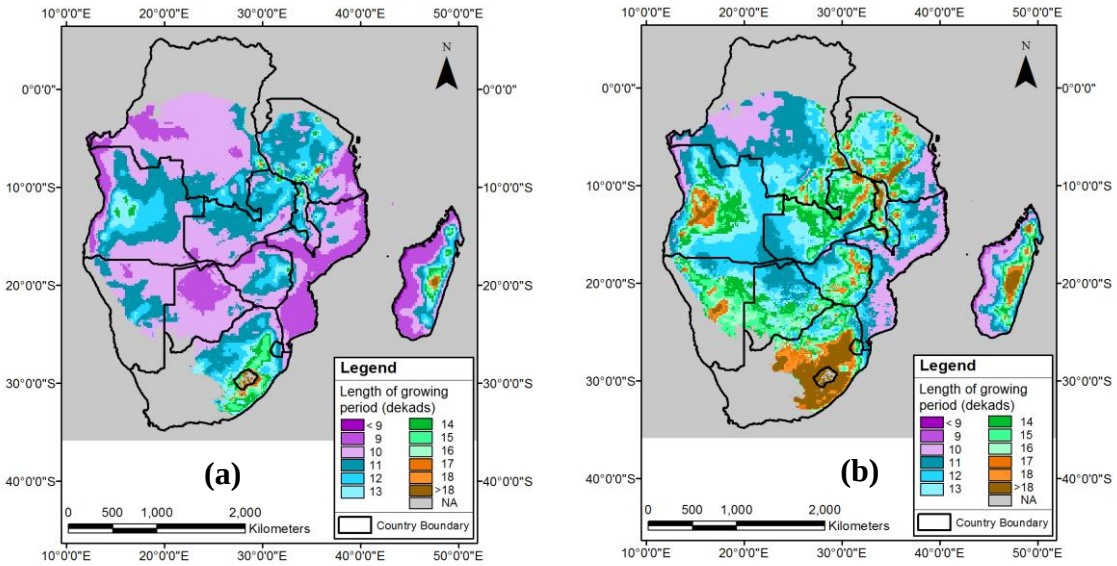


Figure 4.5 (a) Minimum and (b) maximum calculated LGP for 35 seasons, from 1982/1983 season to 2016/2017 main agricultural season. The LGP was calculated for a maize crop with a 1482 GDD requirement

Large differences in the LGP extremes were observed in some cases, with range in LGP (maximum minus minimum) of greater than 9 dekads in parts of South Africa, Lesotho, Malawi

and Tanzania, although it was less than 4 dekads in most areas (Figure 4.6a). The inter-seasonal variability as estimated by standard deviation (Figure 4.6b) was much lower, with most areas having a standard deviation of less than 2 dekads.

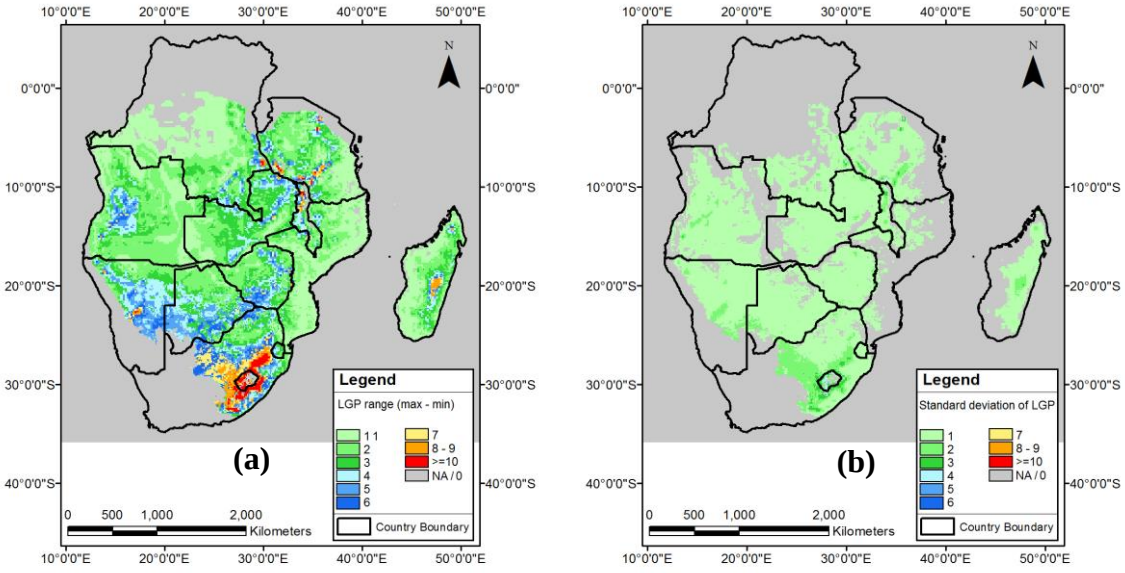


Figure 4.6 (a) Range and (b) standard deviation of LGP calculated for 35 seasons, from 1982/1983 season to 2016/2017 main agricultural season. The LGP was calculated for a 1482 GDD maize crop

Changes in LGP were noted to coincide with increased delay in planting dates (Figures 4.7 and 4.8). Earlier onset generally coincided with shorter LGP, while late planting resulted in longer LGP for many areas. The increase in LGP with delay in onset was particularly characteristic of specific areas, including Zimbabwe, parts of Zambia and Malawi, Botswana, South Africa, eastern Madagascar and Namibia. In other areas, including much of Mozambique, Angola, Tanzania and western Madagascar, little to no change in LGP occurred with delay in planting.

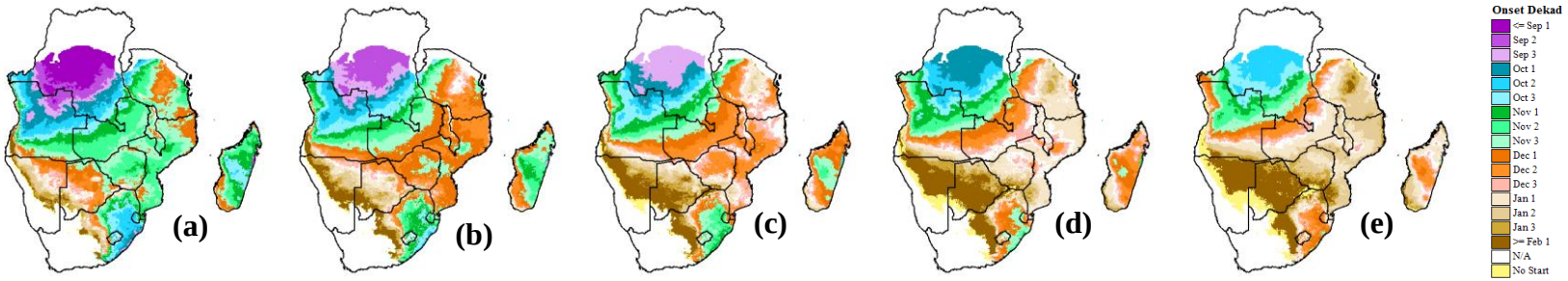


Figure 4.7 Potential consecutive planting dekads averaged over the period 1982/1983 to 2016/2017. Figure 4.7a shows the onset of rains, following the methodology described by Senay and Verdin (2003). Figure 4.7b to 4.7e show the first four dekads after the onset in which at least 25 mm of rainfall is received

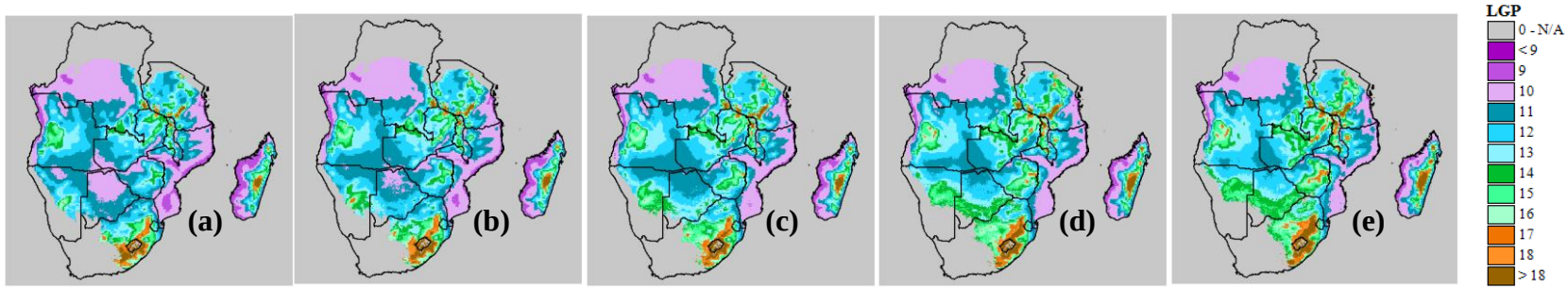


Figure 4.8 Average LGP associated with the average planting dates defined in Figure 4.7, based on a crop requiring 1482 GDD. Figure 4.8a is the average LGP calculated for planting dates summarized by Figure 4.7a; Figure 4.8b LGPs match Figure 4.7b planting dates, Figure 4.8c LGPs match Figure 4.7c planting dates, Figure 4.8d LGPs match Figure 4.7d planting dates, and Figure 4.8e LGPs match Figure 4.7e planting dates

The LGP was generally stable over the 35 year period between 1982 and 2016, as suggested by a trend analysis over this period (Figure 4.9). Most areas had an average change in LGP of less than 0.5 dekads per decade. A few areas of north-eastern Tanzania, eastern Madagascar, eastern South Africa and central Namibia had a shortening of the LGP of at least 0.5 dekads per decade. In contrast, areas around Botswana and central Zimbabwe showed an average increase in the LGP of greater than 0.5 dekads per decade.

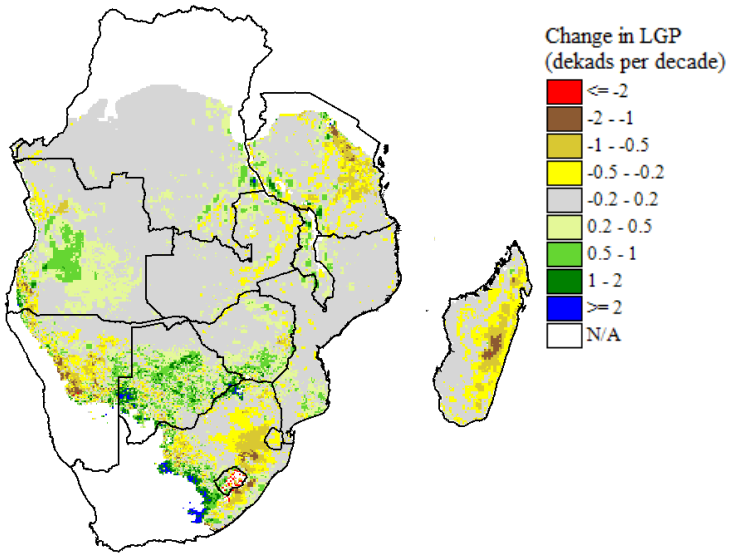


Figure 4.9 Average change in LGP in dekads per decade, calculated for the period from 1982 to 2016

The trend analysis of the eGDD for the 35-year period between 1982 and 2016 (Figure 4.10a) showed minor changes in the long term average eGDD. The change per decade was less than 10 GDD in most areas. Exceptions included eastern South Africa and eastern Madagascar where there was an average increase in the eGDD. An average decrease in the eGDD was also noted in southern Angola, northern Namibia and southern DRC. When compared to the 35-year average and standard deviation of the eGDD (Figures 4.10b and 4.10c respectively) the change in eGDD was generally low.

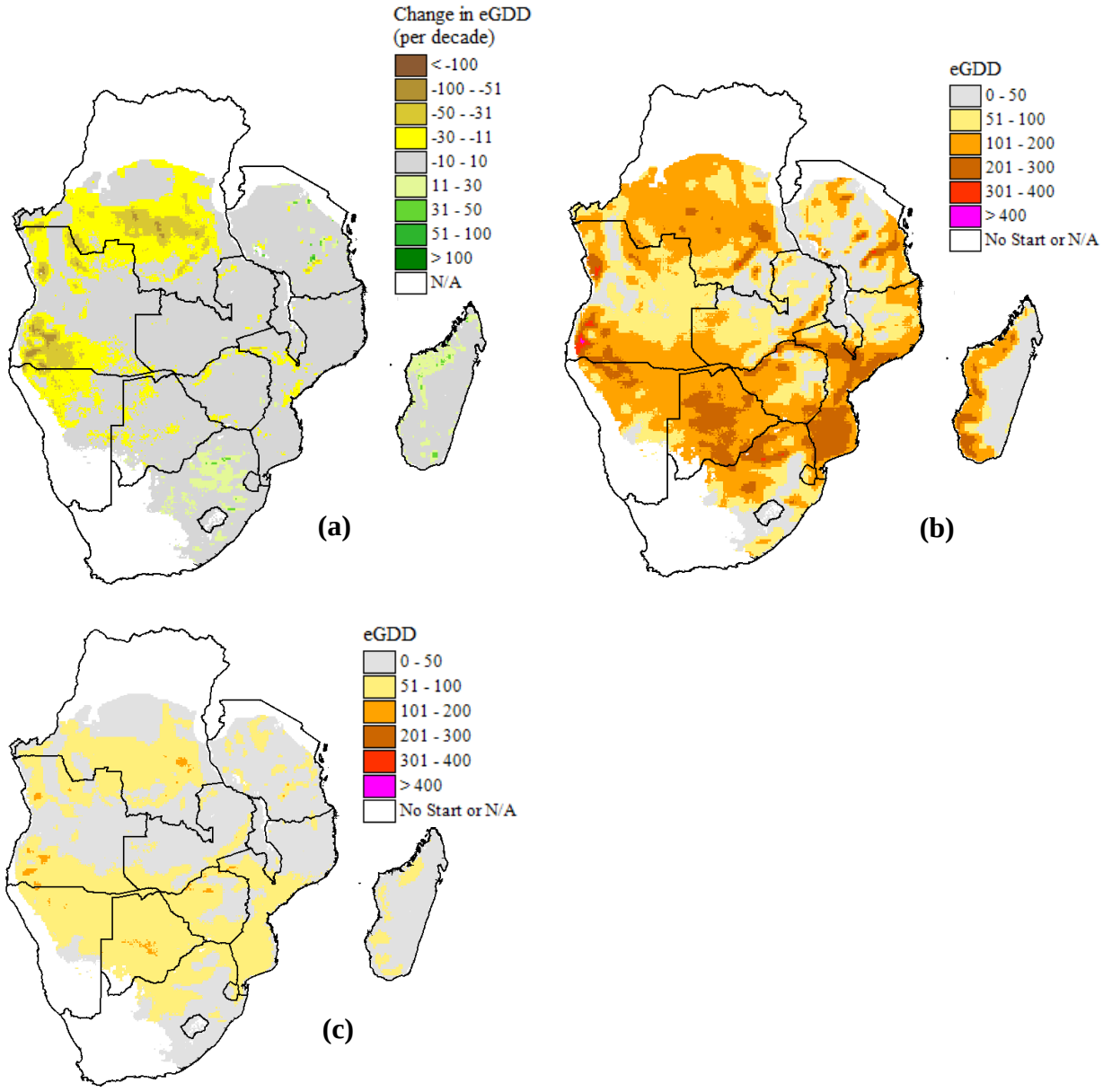
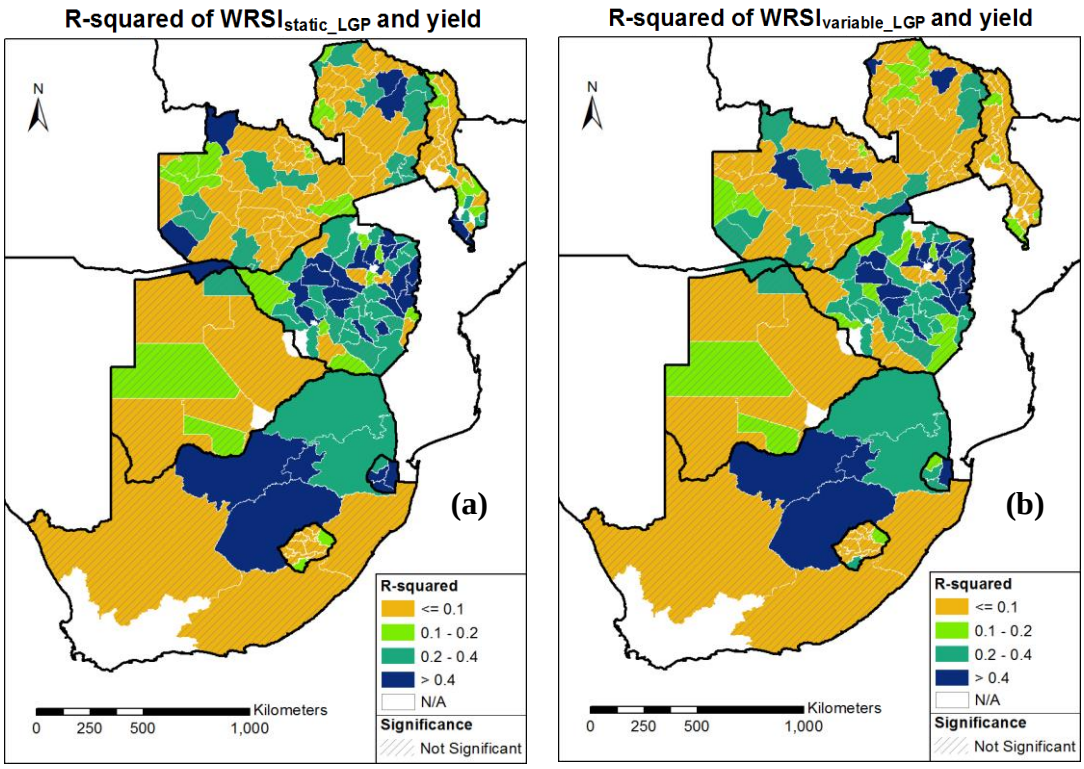


Figure 4.10 Analysis of eGDD data for the period 1982 to 2016, showing (a) average change in eGDD (trend), (b) average eGDD and (c) standard deviation of eGDD

Values of R-squared for various WRSI, eGDD and DxeGDD parameters with crop yield statistics were calculated. Nine different parameters were analysed for each of 183 sub-national administrative units, with correlations for eight of these shown in Figures 4.11 and 4.12. Out of

the 183 administrative units analysed, the number of regressions that were statistically significant for the different parameters ranged from 43 to 74 (for the eGDD and the DxeGDD parameters respectively), being 23% to 40% of the total number of administrative units (Table 4.3). Overall, the $DxeGDD_{3onsets}$, $WRSI_{varLGP\_halfWHC\_3onsets}$ and $WRSI_{static\_LGP}$ parameters had the highest number of statistically significant correlations with yield (Table 4.3). While most countries analysed had a larger number of WRSI-related parameters with a greater percentage of statistically significant correlations (Table 4.4), eGDD-related parameters for Botswana had stronger relationships with yield than WRSI parameters.



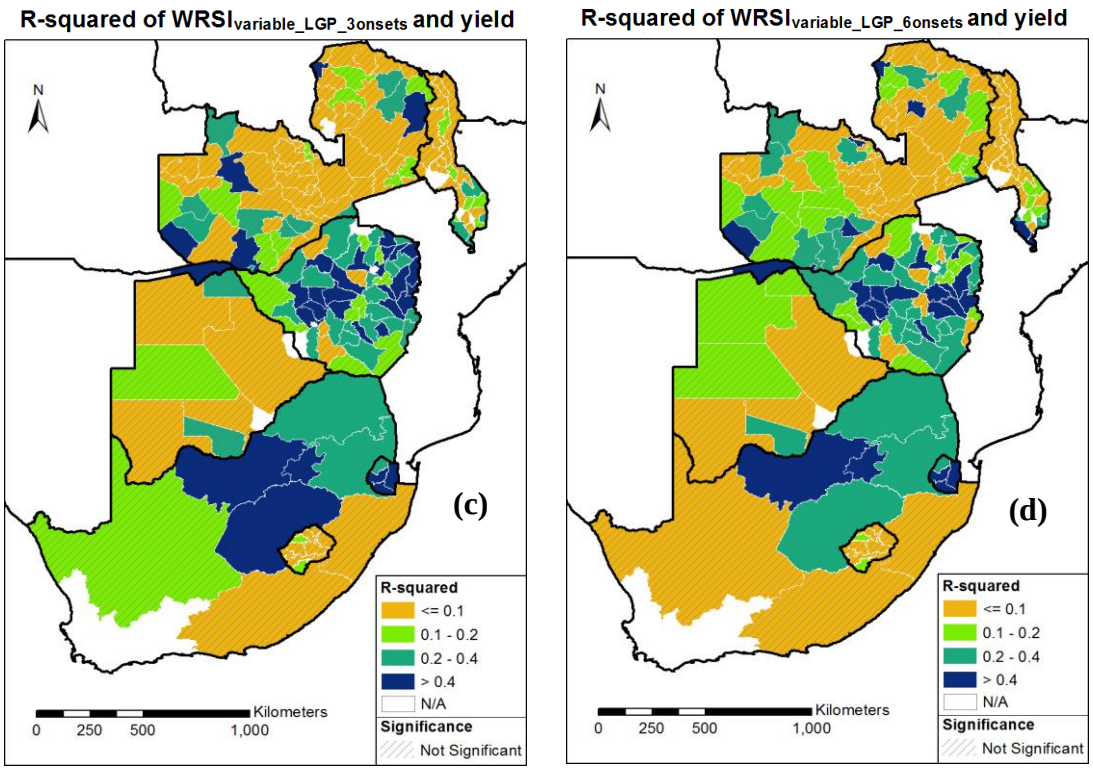


Figure 4.11 Correlation between historical sub-national maize yield and (a) WRSI based on a static LGP, (b) WRSI based on a variable, GDD-determined LGP, (c) average WRSI for three consecutive plantings, based on a variable, GDD-determined LGP, and (d) average WRSI for six consecutive plantings, based on a variable, GDD-determined LGP. All WRSI calculations were for a 1482 GDD maize crop. Regions in which the correlation coefficient is significant at the 0.05 level are shown without hatched lines. Hatched lines indicate those regions where the correlation coefficient is not significant at the 0.05 level

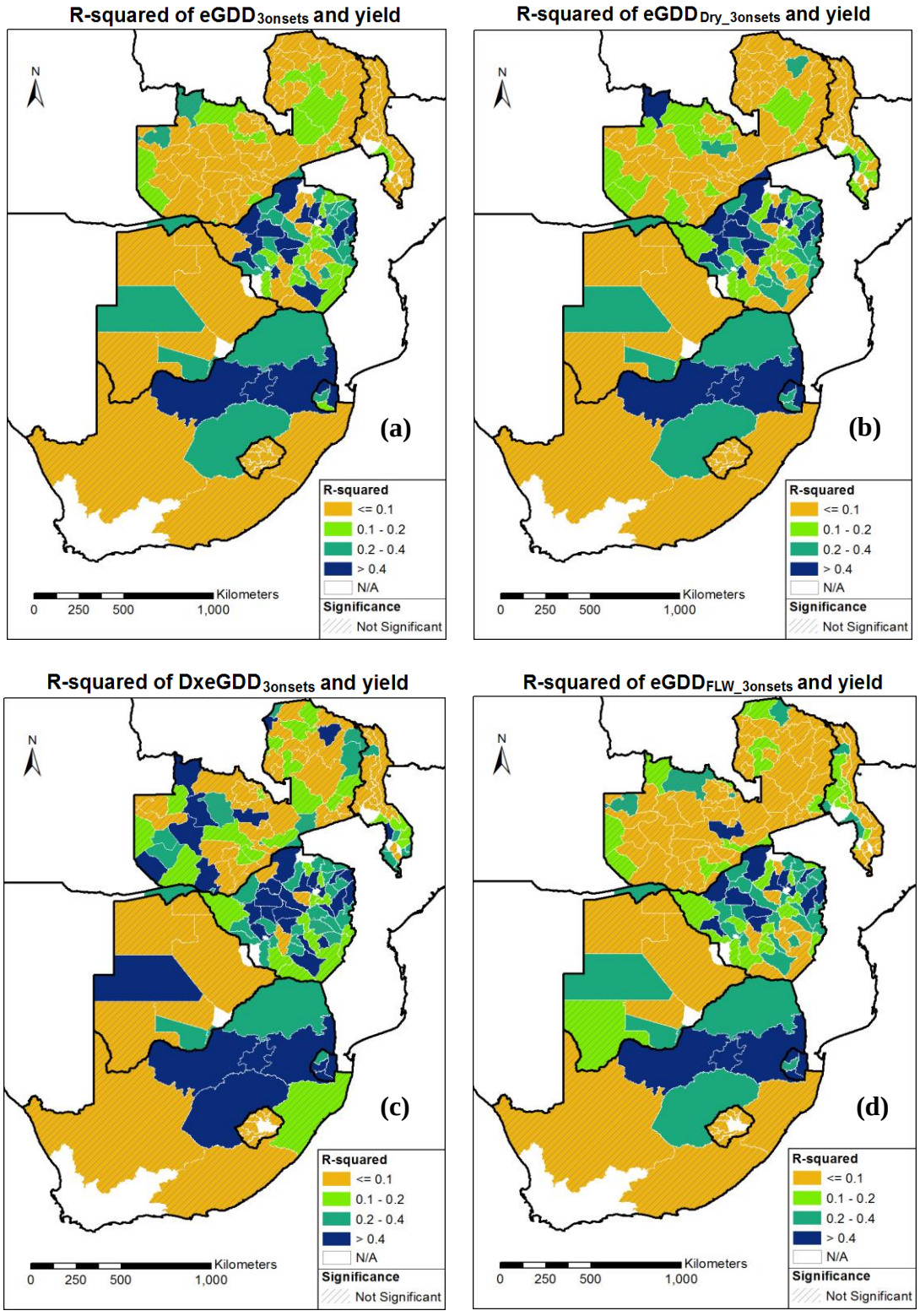


Figure 4.12 Correlation between historical sub-national maize yield and (a) total eGDD, (b) total eGDD occurring during dry periods of crop water deficit, (c) DxeGDD and (d) total eGDD during the flowering period. All GDD calculations were based on CSWB

model runs for a 1482 GDD maize crop, run for three consecutive planting dates. Regions in which the correlation coefficient is significant at the 0.05 level are shown without hatched lines, while hatched lines indicate those regions where the correlation coefficient is not significant at the 0.05 level

Table 4.3 Number of administrative units for which different WRSI and eGDD parameters indicated significant correlation with maize yield

Parameter	Number of admin units with significant correlation	Percent of admin units with significant correlation
$WRSI_{static\_LGP}$	67	37
$WRSI_{variable\_LGP}$	56	31
$WRSI_{variable\_LGP\_6onsets}$	62	34
$WRSI_{variable\_LGP\_3onsets}$	64	35
$WRSI_{varLGP\_halfWHC\_3onsets}$	70	38
$eGDD_{3onsets}$	43	23
$eGDD_{Dry\_3onsets}$	49	27
$eGDD_{FLW\_3onsets}$	52	28
$DxeGDD_{3onsets}$	74	40
Best parameter (combined)	103	56

Table 4.4 Percentage of administrative units in each country with statistically significant correlations between maize yield and each of the parameters analysed. Each value in the table is a percentage of the total number of administrative units in the country that were analysed

Parameter	Botswana	Lesotho	Malawi	South Africa	Swaziland	Zambia	Zimbabwe
$WRSI_{static\_LGP}$	0	20	39	63	100	7	72
$WRSI_{variable\_LGP}$	0	10	0	63	50	10	70
$WRSI_{variable\_LGP\_6onsets}$	13	20	22	63	100	11	63
$WRSI_{variable\_LGP\_3onsets}$	0	20	17	63	75	10	74
$WRSI_{varLGP\_halfWHC\_3onsets}$	13	30	43	63	100	10	68
$eGDD_{3onsets}$	25	0	4	63	50	1	54
$eGDD_{Dry\_3onsets}$	25	0	9	63	75	3	60
$eGDD_{FLW\_3onsets}$	25	0	22	63	75	3	60
$DxeGDD_{3onsets}$	25	0	43	63	100	17	70
Best Parameter	25	33	70	63	100	31	88

For each administrative unit, the parameter with the highest correlation to yield was identified and mapped (Figure 4.13). From the compilation of highest correlations, 103 of the 183 administrative units, or 56%, had statistically significant correlations between the best parameter and yield (Table 4.5). Of the 183 administrative units, 34 units, or 19% had the original WRSI parameter based on a static LGP as the best predictor of yield. A total of 79 administrative units, or 43%, had a parameter based on variable-LGP WRSI as the most predictive parameter, while 46 administrative units, or 25% had a parameter based on the eGDD as the best predictor, and 24 administrative units, or 13%, had the DxeGDD parameter as the best predictor.

At a country level, Lesotho, at 3 out of 9 administrative units, had the highest percentage of administrative units with the original static LGP-based WRSI as the best predictor of yield, while none of the Swaziland's 4 administrative units identified the original WRSI as the best predictor (Table 4.6). South Africa and Lesotho had only 1 administrative unit each (out of 8 and 9 districts respectively) identifying a eGDD-based parameter as the best predictor, while Swaziland had 3 out of 4 and Botswana had 4 out of 8 of their administrative units having their best yield predictors being one of the eGDD-based parameters.

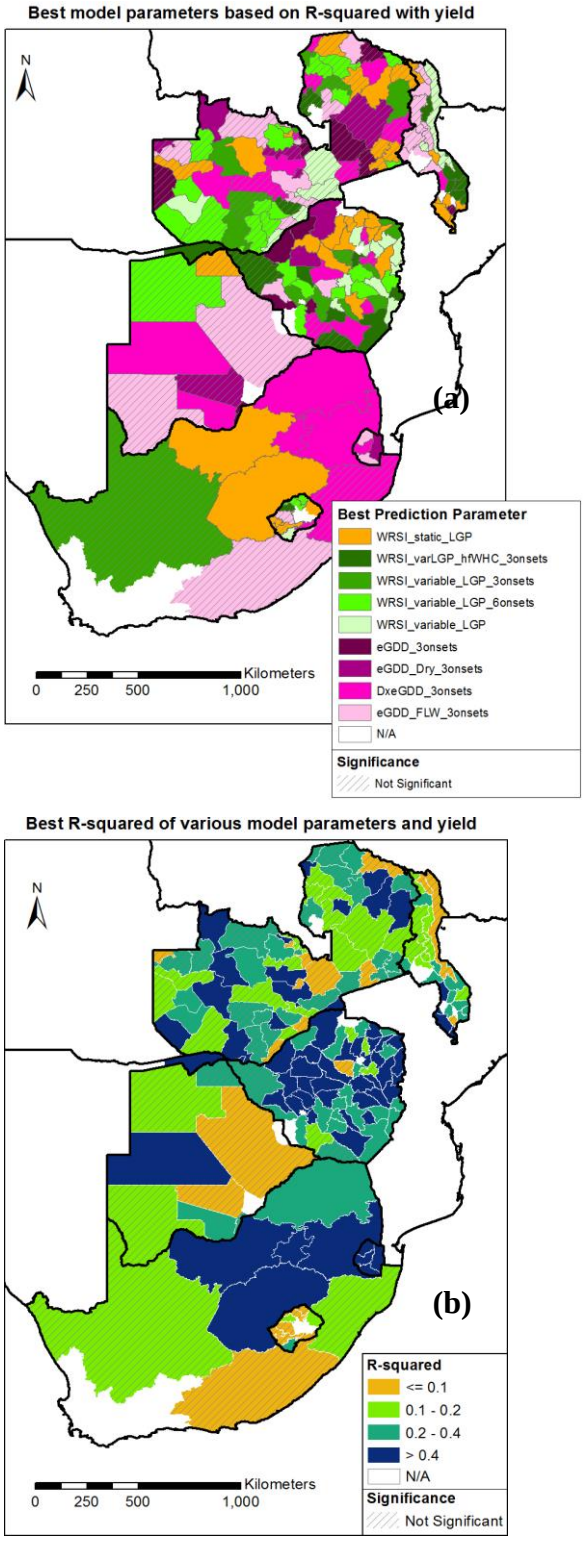


Figure 4.13 (a) Model parameters with the highest R-squared values against maize yields for each administrative unit and (b) the corresponding R-squared values

Table 4.5 Summary of performance of different model parameters showing the number of administrative units in which they displayed the highest correlations with maize yields

Parameter	Number of admin units with highest correlation	Number of statistically significant highest correlation admin units
$WRSI_{static\_LGP}$	34	21
$WRSI_{variable\_LGP}$	20	14
$WRSI_{variable\_LGP\_3onsets}$	16	11
$WRSI_{variable\_LGP\_6onsets}$	26	14
$WRSI_{varLGP\_halfWHC\_3onsets}$	17	12
$eGDD_{3onsets}$	11	4
$eGDD_{Dry\_3onsets}$	10	4
$eGDD_{FLW\_3onsets}$	25	10
$DxeGDD_{3onsets}$	24	13
Total	183	103

Table 4.6 Summary of a selected set of the best parameters for predicting yield by country

Country	Number of admin units analysed	Number of admin units with $WRSI_{static\_LGP}$ as best parameter	Number of admin units with eGDD-based parameters as best parameter	Number of admin units with DxeGDD as best parameter	Percent of admin units with $WRSI_{static\_LGP}$ as best parameter

Botswana	8	1	4	2	13
Lesotho	9	3	1	1	33
Malawi	24	7	9	1	29
South Africa	8	2	1	4	25
Swaziland	4	0	3	1	0
Zambia	71	10	20	9	14
Zimbabwe	57	11	8	6	19

Correlations between standardized de-trended yield index (hereafter referred to in this section as "standardized yield) and different version of various standardized WRSI- or eGDD-based parameters all indicated R-squared values of between 0.22 and 0.40 (Figure 4.14). The standardized, variable-LGP WRSI indicated an R-squared value of 0.30 against standardized yield (Figure 4.14b), but this improved substantially to 0.40 when optimal number of planting dekads (between 1 onset, 3 onsets or 6 onsets) were considered (Figure 4.14c), which was also the highest R-squared value indicated. The correlation between standardized yield and standardized static-LGP WRSI indicated the second-best R-squared value of 0.35 (Figure 4.14a), among the parameters analysed. The standardized eGDD had the lowest R-squared against standardized yield at 0.22 (Figure 4.14d), while the standardized DxeGDD had a moderate R-squared value of 0.33 against standardized yield. When only administrative units which indicated significant correlations were considered for each parameter, the WRSI-related correlations with yield all ranged between 0.38 and 0.40, however only 53 administrative units were considered for the single-onset, variable-LGP WRSI, followed by 67 administrative units for the static-LGP WRSI, and 75 administrative units for the multi-onset WRSI. A lower

number of administrative units compensated for the decrease in R-squared that occurred when more administrative units were considered. The eGDD-yield correlation indicated an R-squared of 0.35 when only administrative units indicating significant correlations were considered.

Linear regression analysis between standardized multi-onset variable-LGP WRSI and standardized yield gave equation (4.3)

$$Yield_{std} = WRSI_{vL\_multi\_std} \times 0.635 \quad (4.3)$$

where $Yield_{std}$ is the standardized yield and $WRSI_{vL\_multi\_std}$ is the standardized variable-LGP multi-onset WRSI

For all scatterplots of the different versions of WRSI against yield, the spread of the points, which indicates the variability in the data, was small for lower WRSI values, while higher WRSI values showed more spread. (Figure 4.14 a-c). In contrast, the eGDD-related scatterplots against yield had less spread at higher values of eGDD, and more spread at lower values of eGDD.

The eGDD indicated a moderately high R-squared of 0.37 when correlated against WRSI (Figure 4.14f). This was a comparable coefficient of determination to those observed when the various WRSI expressions were correlated against yield (Figure 4.14 a-c). Unlike with the eGDD-yield or WRSI-yield scatter plots, there was no notable change in variability from low to high values of either eGDD or WRSI.

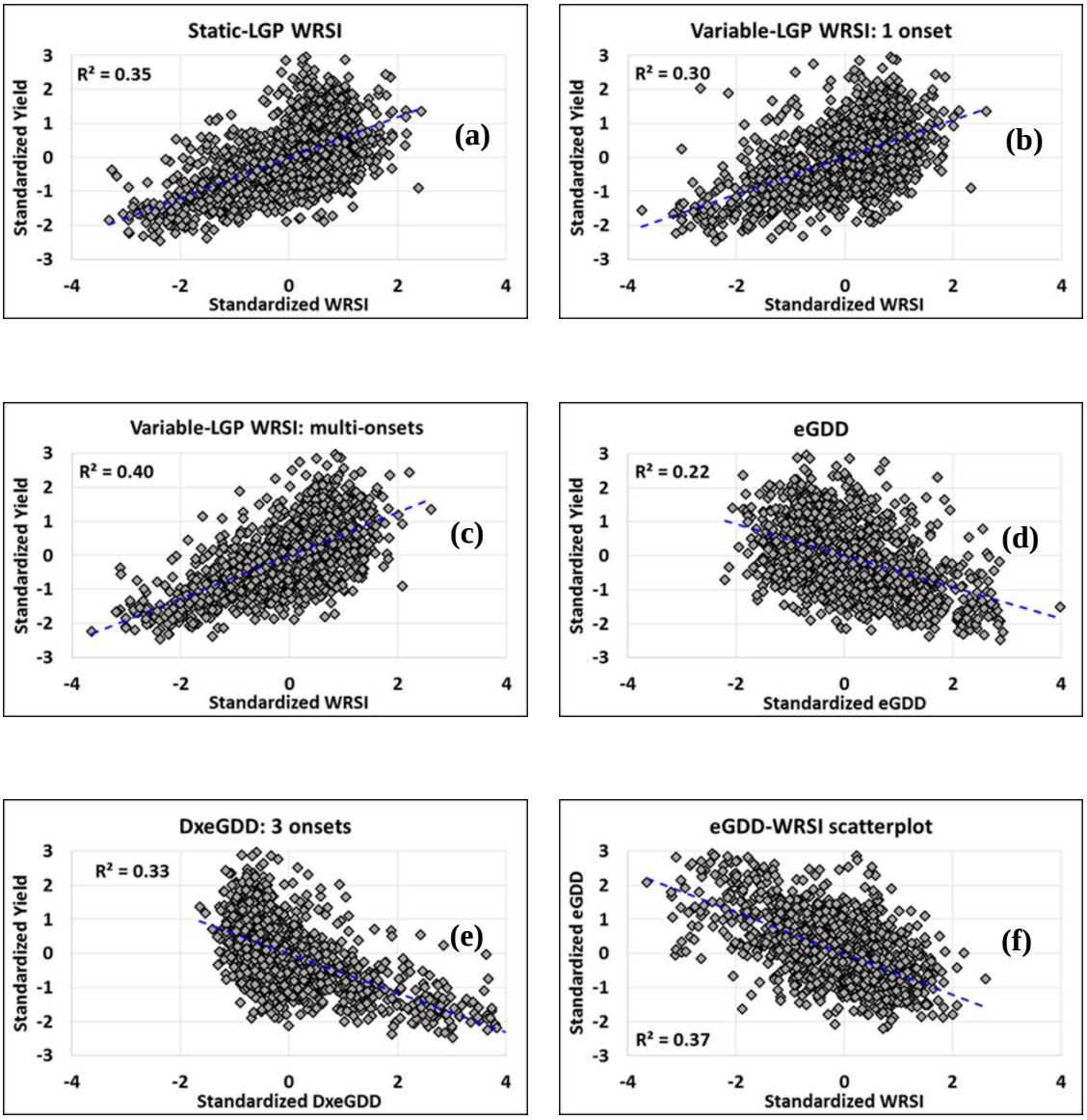


Figure 4.14 Region-wide scatter plots for all 1422 data points from 75 administrative units that indicated significant correlation between yield and the multi-onset WRSI parameter. Scatterplots are shown between standardized yield and (a) standardized WRSI calculated using static LGP, (b) standardized WRSI calculated using variable, GDD-based LGP for planting with the onset of rains (c) standardized WRSI calculated using variable, GDD-based LGP for optimal number of planting dates, (d) standardized

eGDD, and (e) standardized DxeGDD. Correlation between standardized eGDD and standardized WRSI is shown in (f)

A multiple linear regression model for estimating maize yield using standardized eGDD and standardized WRSI as predictors was tested to check the extra value that eGDD adds to crop monitoring analysis. Table 4.7 shows the R-squared values for each parameter and both parameters when correlated with yield. The inclusion of eGDD in the regression of WRSI with yield resulted in a marginal increase in R-squared from 0.40 to 0.41.

Table 4.7 R-squared values for regressions between standardized values of yield, WRSI and eGDD for 1422 data points from 75 administrative units

R-squared: WRSI vs yield	R-squared: eGDD vs yield	Multiple R-squared: [WRSI and eGDD] vs yield
0.40	0.22	0.41

4.5 Discussion

The CSWB modelling done in this study utilized a gridded approach, with water balance calculations being run for each pixel. As a result, gridded datasets such as the CFSR and CFSv2 were used for temperature analysis. The statistics obtained from the validation of the CFS air temperature data using GSOD air temperature, including high r-squared values of between 0.63 and 0.82, regression slopes close to 1, and regression intercepts between 0 and -3 ℃ data suggest that the use of CFS data for temperature analysis allows reasonably accurate tracking of the variations in temperature. The same validation statistics however also suggest potential for

improvement in the CFS data, given to the inherent error in the CFS, as suggested by the validation statistics (e.g. r-squared values of 0.63 in some cases). Such improvements can be implemented by merging the gridded CFS data with available observed daily station data such as the GSOD, using any number of existing techniques for combining raster and point data (e.g. Myers 1982, Nalder and Wein 1998, Budde et al. 2014, Li and Heap 2014). Future research can consider the use of existing techniques for combining these datasets, and additionally incorporating other data with strong influence on temperature, such as cloud masks, which may further improve the gridded temperature data.

Dekadal GDDs calculated from dekadal maximum and dekadal minimum air temperatures (GDDekads) showed a high correlation with daily GDDs summed to a dekadal timestep (Figure 4.3). The high correlation between the two versions of dekadal thermal time suggests that extreme temperatures, by definition, rarely occur during the dekad, and that weather systems generally do not experience sudden sub-dekadal changes. A large difference between dekadal summed GDDs and GDDekads would occur when several days within the dekad have $T_x > T_o$ or $T_n < T_b$ (as defined in Equation 4.1) but the dekadal average falls well between T_b and T_o. The difference between GDDekads and dekadal-summed GDDs is likely to be more pronounced when large temperature swings occur during the dekad, for example due to the slow passage of frontal systems through an area. The high correlation further suggests that dekadal temperature data sets can be used for calculating GDDs and running the variable-LGP CSWB model, which would reduce computer disk space and processing time required to run the model. However, the use of GDDs calculated using dekadal temperature data could sometimes result in an underestimation of eGDDs, especially for dekads in which maximum temperatures had high variability over the course of the dekad (e.g. due to the development of a weather system during the dekad that could cause maximum temperatures to change abruptly).

The extent to which the crop coefficient method and the thermal time method can be integrated has some limitations, in part due to differences in the definition of the different stages in the descriptive crop growth schema which utilizes crop coefficients, and the leaf collar method which uses GDDs. For example, the end of the initial stage, which is defined as the time the crop attains 10% ground coverage, will be to some extent dependent on the plant population, as the time to which 10% of ground coverage is attained will depend on the plant density as well as rate of vegetative development.

The LGPs calculated under various scenarios (Figures 4.4 and 4.8) were found to be comparable to those mentioned in various literature. For example, Harrison et al. (2011) noted LGP in Mozambique ranging from 90 to 120 days, similar to GDD-calculated LGP in this study. Pannar (2017) noted that the LGP for their early maturing maize varieties ranged from 110 to 150 days in warm and cool areas respectively. In comparison, the GDD-based methodology observed the early maturing variety to range from 100 to 160 days (Figure 4.4b) in the main maize growing areas of South Africa, located primarily in central parts of South Africa, north of Lesotho. Ngwira et al. (2014) observed the crop growing period for a medium-maturing maize variety at Chitedze, an agricultural research station in Malawi between 2007 and 2013, and noted it to range between 137 days and 144 days. Comparatively, our model calculated the average GDD-based LGP at the same location at 13 dekads. To further validate the accuracy of the LGP estimates made in this study, future studies can compare data measuring days to maturation, sourced from field trials at a more spatially representative distribution. As the determination of LGP was not the primary objective of this study, a rigorous validation of LGP calculations was not carried out.

The LGP standard deviation (Figure 4.6b) suggested that in most areas under study, the LGP is within 1 dekad of the mean LGP for 67% of the seasons. However, when there are extremes

in temperature, the LGP can vary by up to 3 dekads in many areas, and by 4 or more in some areas including parts of Botswana, Malawi, Namibia, South Africa, Zambia and Zimbabwe, as suggested by the LGP range (Figure 4.6a). Large variations in LGP were also noted to occur with changes in planting dates (Figures 4.7 and 4.8). The general increase in LGP associated with delay in planting was also noted by Moeletsi (2017) in Free State Province of South Africa; in our analysis of the model results over Free State, an average change of onset of 5 to 6 dekads (e.g. from dekad 2 of November to dekad 2 of January) resulted in an average increase in LGP by 3 to 5 dekads. This variability in LGP underlines the importance of consistent monitoring of temperatures and estimation of LGP, as these factors can affect the accuracy of WRSI due to timing of dry, wet and warm spells relative to the crop phenology, as well as the timing of harvest, which can have implications on food availability and food security. These findings align with the note made by Reynolds et al. (2000b) on the use of a static LGP as one limitation of the CSWB model, and highlights the importance of a variable LGP solution. The results also tie in with demonstrations by Harrison et al. (2011) of the potential reductions in crop growth period in Mozambique with increase in air temperature, and with experiments by Hatfield and Prueger (2015) that suggested maize plants had a shortened crop cycle when they experienced increased temperatures.

There were some administrative units noted where none of the WRSI-based or eGDD-based parameters showed significant correlations with historical yield. These included much of Zambia and northern Malawi. A study of climatological temperature and rainfall in these areas indicated that rainfall tended to be high, and maximum air temperature relatively low. For example, in Mpongwe, Zambia, the long-term average rainfall for the December to February (DJF) period, based on CHIRPS is 751 mm, while long-term average maximum air temperature, based on CFS data, is 26.9 ℃. Similarly, Mzimba in Malawi has average rainfall and maximum

air temperature of 617 mm and 25.9 ℃ respectively for the DJF period. This suggests that in both these areas, and many like them, neither rainfall nor extreme temperatures are factors that would tend to negatively impact maize yields, except during rare events. In contrast, Free State in South Africa had average rainfall and maximum air temperature of 247 mm and 31.3 ℃ respectively for the DJF period.

Given that WRSI works best as an indicator of crop yield in water-limited areas (Reynolds et al. 2000b, Verdin and Klaver 2002, Senay and Verdin 2003), it is unsurprising that the high rainfall areas such as those noted in countries like Malawi and Zambia did not have statistically significant WRSI correlations with yield. However the number of administrative units in which $WRSI_{varLGP\_halfWHC\_3onsets}$ was the most predictive, including some of the high rainfall areas in Malawi and Zambia, suggests that further calibration of the WRSI by adjusting the soil WHC and consequently changing water limitations, may improve the accuracy of the model. WHC values derived from soil maps can be generated using pedo-transfer functions (e.g. Reynolds et al. 2000a). While we arbitrarily tested a reduction of half the FAO-based WHC, assuming a 50% reduction in water availability under conditions of sub-surface compaction (Taylor and Brar 1991), various factors can increase or decrease the retention of water in soils to varying degrees, including soil organic matter, and management practices such as mulching, tillage practices, and use of cover crops (Minasny and McBratney 2018). A rigorous step-wise calibration to determine the effective WHC would improve the accuracy of the WRSI.

Changes in the LGP based on the temperature regime also improved the WRSI's predictive capability, given that 43% of all the administrative units had a parameter based on variable-LGP WRSI as the one most predictive of yield. Of these variable-LGP WRSI parameters, 75% were based on multiple onsets, effectively simulating the planting of crops by farmers on different dates over a planting window, which is a common occurrence in farming practice

(Rurinda et al. 2013). Only three variations in number of onsets were tested, namely 1, 3 and 6 onsets, and different number of onsets gave the best results in different areas. The optimum number of onsets may be a function of the timing of the onset of rains, the maturation period of the variety grown, and the climatological humid period over which crops can be grown. A full calibration can be undertaken with the first estimate based on calculating the above-mentioned function, then increasing or decreasing the number of onsets starting from the first estimate. Although the effect of multiple onsets was simulated by simply averaging the resultant WRSI values from the different planting dates, this result can be improved in future work by developing a probability distribution function of the fraction of total planting done as a function of time from the beginning of the planting window in different areas. Such a product would allow a weighted average of WRSI values from the different planting dates, based on the expected planting prevalence on each date. Additionally, reviewing the different WRSI values generated for each planting date enables the analysis of crop condition expectations for farmers who planted at different dates. In an operational crop monitoring system, information on planting dates may have been gathered early during the season, but various constraints may prevent regular visits to physically check the crop condition.

The eGDD shows the cumulative temperature stress experienced by a crop throughout the growing season. This can help to explain why it showed high, statistically significant correlation with crop yields in several administrative units. Analogous to the appropriateness of WRSI for water limited areas, the eGDD is expected to best predict yield in areas where extreme temperatures frequently occur during the crop-growing season, causing more frequent incidents of crop heat stress. This was borne out in Botswana, where, with a long-term average CFS-calculated maximum air temperature of 34 ℃, four out of eight of its admin units had a eGDD-based parameter as the most predictive of yield (Table 4.6), while another two admin units were

based on the DxeGDD parameter. Despite this limitation, the eGDD method for monitoring of extreme temperatures can nonetheless be used as an indicator for highlighting areas that may be negatively affected by high temperatures, whose impacts can be worsened as the duration of a hot spell increases (Lobell et al. 2011). It holds significant potential as a temperature-based crop monitoring indicator. The possibility of scaling it for different phenological stages to capture the differential effects described by Lobell et al. (2011) can also be explored. Finally, eGDD alert levels can be identified in order to help operational users recognize eGDD values that can cause serious damage to crops.

The DxeGDD, a rudimentary parameter combining both moisture stress and heat stress effects through a simple multiplication, gave the highest correlation with yield amongst all the parameters tested, in 13% of the admin units analysed. The approach of scaling eGDD by the percentage water deficit is in line with observations made by Lobell et al. (2011) that the impact of extreme temperatures on crop yield was higher under conditions of crop water stress. While a linear relationship between DxeGDD and crop yield was assumed for demonstration purposes, more work is required to determine if a linear relationship is optimal, and if any constants and coefficients should be applied in order to make the parameter more robust and predictive.

The methodology applied of standardizing the de-trended yield index and the various WRSI and eGDD parameters enabled development of a robust regional regression model based on over 1400 data points from 75 different administrative units (Figure 4.14c, Equation 14.3). To allow comparability between the different parameters, a single set of administrative units was used for all parameters, although not all the administrative units indicated significant yield/model-parameter correlations for all the different parameters tested. When applying a specific model, it would be necessary to develop it and apply it to only those administrative units where significant yield-model parameter correlations are indicated.

The scatterplots of the various WRSI and eGDD parameters with yield displayed a tendency for the yield spread to increase with increasing WRSI and with decreasing eGDD. This suggests that higher variability in yield can be expected for higher WRSI values and lower eGDD values. This can be explained by the fact that when WRSI is high, and in particular, consistently high for most seasons, rainfall is not a limiting factor and other factors unrelated to rainfall deficit (such as pests and diseases, availability of inputs, waterlogging etc.) can cause large changes in yield, but without reflecting similar changes in the WRSI. However, increasingly low values of WRSI will be associated with larger rainfall deficit-related yield reductions due to greater crop moisture stress. The reverse scenario holds for eGDD, as increased duration of higher temperatures (defining eGDD) have increasingly negative impacts on crop yields due to increased heat stress reflected by the high eGDD. At low temperatures however, other factors unrelated to temperature may have a comparatively larger influence on crop yields.

WRSI was noted to be inversely well correlated with eGDD. Three possible reasons for this correlation can be postulated. The first is that through the Penman-Monteith equation, air temperature is an input for the calculation of reference evapotranspiration (Allen et al. 1998), which in turn is one of the two main meteorological factors that affect WRSI. Thus an increase in temperature would result in an increase in reference evapotranspiration and hence a decrease in WRSI if the increased temperature event is not preceded by or concurrent with sufficient rainfall to provide adequate soil moisture to equal atmospheric demand. The second possible reason is that high temperatures could be caused by increased solar irradiance in turn caused by little cloud cover, given the relationships between cloud cover and temperature (Karl et al. 1993, Dai et al. 1999). The lack of cloud cover would result in reduced rainfall, and therefore lower WRSI values if the dry spell extends for an appreciable length of time. Trenberth and Shea (2005) noted that temperature and precipitation are strongly inversely correlated. A third reason

is that high temperature could be an indication of low soil moisture levels, which would cause a reduction in evaporative cooling and thus high temperatures (Dai et al. 1999). This discussion suggests that eGDD is not just a sign of heat stress, but is also likely an indicator for high evapotranspiration, low rainfall and low soil moisture, which all (especially the latter) cause crop water deficit. Its high correlation with WRSI also suggests that it would have limited utility as an additional parameter in multiple regression of WRSI with yield. However, it shows potential utility as an independent indicator of low rainfall and soil moisture conditions. Thus a high eGDD value may be used as a flag for closer inspection of crop conditions, as well as a cross-check for WRSI. An additional strength of the eGDD is its simplicity and limited assumptions, which represents lower opportunity for incorrect assumptions.

One challenge noted with the approach of testing various crop monitoring parameters by correlating them with official historical yield data is the uncertainty in the yield data quality due to various limitations in some of the crop yield forecasting and estimation methods used. Fermont and Benson (2011) noted a number of factors that can confound the accuracy of crop yield estimates for smallholder farming, including large spatial variations in yield that may be difficult to capture statistically, inter-cropping – with more than one crop planted in a field, and large differences between planted area and harvested area – resulting in large differences in calculated yield, depending on whether planted area or harvested area is used to calculate yield. Jayne and Rashid (2010) made mention of a number of weaknesses in the two main types of crop yield and production estimates used by many countries, of which one type is generally designed and led by ministries of agriculture, being reliant on the experience of their field staff, while the other type is based on crop surveys undertaken by national statistical agencies, and how these sometimes lead to inaccurate crop production estimates. The inherent errors in crop estimates are likely to vary amongst countries, based on the specific implementation used

therein. Evaluations of crop estimations parameters done using crop statistics should therefore be interpreted cautiously, and low correlations should not necessarily be dismissed as poor performance of the CSWB model-derived parameter. However, it is unlikely that errors in the historical crop production statistics would result in high and significant correlation with independent CSWB model-based parameters, and those administrative units where high, significant correlations between crop yields and the crop monitoring parameters were indicated provide supportive evidence for the utility of those parameters for crop monitoring.

4.6 Conclusions

The study demonstrated that the ability of a water requirements satisfaction index (WRSI) based on a crop-specific water balance model (CSWB) to estimate crop yield, can be improved through the use of simple temperature-based analytical techniques. In particular the use of growing degree days (GDD) for modelling the length of the growing period, improved the accuracy of the WRSI. Due to the tendency of farmers to plant on different planting dates and effectively cover the span of the humid period, which in many areas was much longer than the typical crop maturity period, it was found necessary to simulate multiple sowings and aggregate the seasonal WRSI from several planting dates. Different areas exhibited highest accuracy for differing numbers of plantings, and it is recommended to calibrate each area separately to the required number of plantings. Improved accuracy was also observed when the model was concurrently calibrated for soil water holding capacities, which can differ from those predicted by pedo-transfer functions due to local agricultural management practices.

The GDD methodology applied on extreme temperatures (eGDD) using a base temperature of 30 ℃ showed utility in identifying the cumulative impacts of extreme air temperatures on crop yields. High correlations were obtained in several administrative units. The use of multiple onsets due to local farmer practices and due to differences with the humid period improved

correlation between eGDD and maize yield in several administrative units. Statistically significant correlations between these temperature based products and crop yields were obtained, despite the use of a coarse modelled temperature dataset. Potential exists for improvement by incorporating accurately observed temperature into the analysis methodology, through the development of methods such as those proposed in Chapters 2 and 3 of this thesis. The eGDD methodology may be limited as a predictive parameter to areas where air temperatures regularly exceed extreme thresholds, but still retains the ability for highlighting areas where extreme temperatures may negatively impact crop yields, which would be a useful addition to the suite of monitoring tools used by the Famine Early Warning System Network (FEWS NET) for seasonal monitoring in support of food security analysis. The high correlation between eGDD and WRSI, as well as the limited additional value of including the eGDD in a multiple linear regression with WRSI and yield however suggests that eGDD may be expressing conflated effects of rainfall and soil moisture availability on temperature. While this reduces the potential utility of specifically measuring heat stress impacts on crops, it also presents an opportunity for the eGDD to be used as a cross-check for WRSI, due to the independent assumptions underpinning the eGDD-soil water balance relationship. However, more work is required in better understanding the eGDD. The DxeGDD, combining both water stress and heat stress effects, has potential as a crop monitoring tool in selected regions, but further analysis and calibration is required before it can be used as a regular monitoring tool.

4.7 Acknowledgements

The provision of the GSOD data for usage in the FEWS NET analysis by NOAA NCEI and the CFSR and CFSv2 data by the National Center for Atmospheric Research is gratefully acknowledged. FEWS NET is also gratefully acknowledged for providing the historical maize

yield data from the FEWS NET Data Warehouse (FDW). Special thanks to Gary Eilerts for facilitating access to the FDW data.

CHAPTER 5: CONCLUSIONS AND RECOMMENDATIONS

5.1 Introduction

Millions of people in southern Africa face conditions of food insecurity each year, and this number increases sharply when adverse weather conditions during the summer growing season affect crop conditions, resulting in reduced crop production. The ability for governments and relevant international agencies to provide a timely response to weather-induced crop production shortfalls is contingent on early and accurate identification of areas where crop production has been affected. The improvement of crop models, and generation of data parameters, that can provide accurate proxy information on crop conditions, is an important contribution in the fight against hunger and food insecurity. The motivation for this research was to explore ways of improving large-area, regional-scale food security early warning applications by developing models and indicators with improved accuracy and correlations with crop yields.

5.2 Aims and objectives revisited

The main aim of this research was to assess the feasibility and potential benefit of incorporating temperature data into an operational, regional-scale crop monitoring system in southern Africa, based on crop-specific water balance modelling, in order to improve the capability of operational crop condition monitoring systems to identify weather-related reductions in maize yields.

The specific objectives of the work to be undertaken were as follows:

- To develop a method of processing thermal infrared (TIR) brightness temperature (BT) data for cloud removal and generation of clear-sky BTs using algorithms compatible with long-term TIR data

- To investigate methods of processing clear-sky brightness temperature datasets for the retrieval of air temperature data variables
- To assess the utility of currently available, global-scale, modelled daily temperature datasets for calculating regional-scale growing degree day grids and derivative length of growing periods (LGP) for selected crops
- To improve crop-specific water balance (CSWB) model calculations using variable LGP derived from air temperature data
- To determine the utility of parameters based on analysis of air temperature extremes for crop monitoring at a sub-national level in seven southern African countries

5.3 Contributions to new knowledge

This study successfully demonstrated the improvement that can be attained in a commonly used CSWB model by integrating temperature data into the model's calculation procedures. Correlations between water requirements satisfaction index (WRSI) and maize yield data improved in several administrative units of selected southern African countries when the LGP was calculated dynamically each season using daily minimum and maximum air temperature data. The study identified that calibration of the number of plantings that occurred during the season was required in order to account for the difference between: (a) the number of days the crop grew from planting to maturity, and (b) the total number of days the crop could potentially be grown from onset to cessation of rains. Additionally, calibration of the soil water holding capacity (WHC) to better reflect ground conditions due to soil management practices offered potential for further improving the WRSI's CSWB model.

A new parameter was developed for highlighting the severity of extreme warm temperatures that could negatively affect crop yields. This parameter, an accumulation of thermal time associated with extreme warm temperatures known to negatively impact crop yields, was

referred to as extreme growing degree days (eGDD). The eGDD parameter was noted to have high correlations with WRSI, suggesting that the high temperature accumulations were conflating the effects of rainfall and soil moisture availability on crop productivity. The eGDD is unlikely to be related only to the cumulative heat stress experienced by a crop, as it adds limited value to WRSI in a multiple linear regression involving eGDD, WRSI and maize yield. The eGDD was indicated to have high correlations with crop yield in administrative units in several southern African countries. A number of variants of the eGDD were developed, including the DxeGDD, which scaled the eGDD by the water deficit experienced by a crop contemporaneously with extreme warm temperature events. The DxeGDD was found to have the highest correlation with crop yields among all the parameters considered in several administrative units. In many cases, the DxeGDD's correlation with maize yield were higher than those of WRSI, which is one of the standard parameters used for crop monitoring.

The potential for development of useful air temperature metrics from a readily available dataset was also successfully demonstrated. The frequency of exceedance of air temperature above an extreme value was calculated by calibrating a satellite-based TIR BT dataset using monthly diurnal air temperature range and monthly mean air temperature from station observations at four representative locations in southern Africa. The calibrated TIR-based air temperature metrics were made possible by the generation of diurnal temperature curves of clear-sky, cloud-free BT estimates using a sinusoidal and exponential model for daytime and nighttime respectively. The generation of sinusoidal and exponential curves was based on new techniques developed during this study, which modified the approach originally proposed by Parton and Logan (1981), so that any two times during daytime (nighttime) with known temperatures could be used to fit the sinusoidal (exponential) curve.

5.4 Challenges

One of the main challenges encountered in the study was that significant computational resources were required for generating cloud-free TIR BT estimates, due to the complexity and number of iterations in the algorithm developed. As no high performance computing resources were available, processing the data on a normal laptop computer took large amounts of time. The iterations thus typically required in a research and development process of this nature, to produce a draft solution, review it, and refine the process, was curtailed. In particular, the heavy computing requirements hindered the generation of regional-scale grids for testing the algorithm at a large scale, spatially and temporally, but instead testing was done for selected representative locations and selected months.

A second challenge met was the lack of availability of phenological data for undertaking a rigorous evaluation of the temperature-based LGP. Most data providers contacted, who had data from agricultural test sites, only had phenology data up to the date of silking, and not up to the date of physiological maturity. This prevented the research from providing an authoritative conclusion on the accuracy of the LGP products, but rather relied on comparison with documentation that had been made in the literature for a few locations within the study area.

5.5 Future research possibilities

A number of improvements can be made on the work presented in this research. Key among these is the further development of the processing of BTs into a fully gridded, computationally efficient algorithm. Such gridding and algorithmic streamlining could cover the full spectrum of BT-related processes implemented in this study, including the identification of clouds, generation of cloud-free BT estimates, and estimation of occurrence of extreme, warm air temperatures. With the development of gridding operations for estimating cloud-free BTs, the

seasonal and geographic limitations of the technique need to be established, as there are likely to be areas and seasons where continuous cloud cover may constrain the capability to estimate cloud-free diurnal brightness temperatures curves. The gridding concept is particularly important, as the final result of this study deals with gridded datasets. The gridding of the BT processes described in Chapters 2 and 3 would therefore further open up opportunities for their integration into gridded crop monitoring operations.

An accurate and reliable dataset is one of the underpinnings in a modelling exercise, and in this study, temperature is key. Various temperature-related datasets from various sources were used in this study, but the primary variables of interest are daily minimum and maximum air temperature. Research can be undertaken to develop ways of operationally producing improved minimum and maximum air temperature data by combining the various available datasets, including accurate, in-situ air temperature data from weather stations, readily available but coarse gridded modelled air temperature data such as that available from reanalysis datasets, cloud-free satellite BT data, and cloud data. The demonstrated capability for using BT for ultimately identifying occurrence of extreme air temperature using simple statistical techniques suggests that cloud-free, TIR-based BT estimates can contribute to the improvement of air temperature estimates. Other studies have also shown strong relationships between temperature and cloud cover, and such relationships can be explored further to determine how to incorporate cloud cover into interpolation of temperature in conjunction with correlated datasets such as reanalysis, using techniques such as co-kriging.

The study demonstrated ways in which estimates of high air temperature occurrence, similar in concept to the eGDD approach, can be derived. The correlation between the satellite-based high temperature occurrence parameter and the eGDD should be further analysed, as this

presents opportunities for exploring the direct usage of the former in crop monitoring, in similar ways to those identified for the eGDD.

The application of WRSI and eGDD parameters in yield estimation may be improved by testing the validity of a number of assumptions made in this study. For example, the relationship between the various parameters tested and crop yields was assumed to be linear. Although this approach produced good correlations, further work can be undertaken to test this assumption, and identify whether non-linear approaches may not provide better correlations. Additionally, while this research focused on a uniformly composed seasonal accumulation of eGDD parameters, future work can assess whether weights can be assigned to different crop stages, similar to the crop coefficient approach used in the CSWB model. Likewise, the DxeGDD was developed by simply multiplying moisture stress and heat stress indicators. This approach can be improved upon by testing the effect of constants and coefficients in the composition of the DxeGDD, in particular, whether this results in improvement in the robustness and predictive ability of the DxeGDD for yields.

Some preliminary calibration of the input parameters used in the CSWB modelling system, particularly the number of planting dates, and the soil WHC was done during this study. A full calibration of these CSWB model parameters will facilitate greater operational benefits of the work presented here. This includes calibration for the effective soil WHC, which is influenced by various soil management practices. The optimum number of planting dates for each area should be calibrated locally, and the composite WRSI derived can be based on weights dependent on the temporal clustering of planting dates, rather than the simple average that was used in this study. Weighted composites such as these, however, rely on the collection of data on planting practices, and a determination will be required on whether significant inter-seasonal variability occurs, or whether a single value can be used. Field data on crop phenology from

planting to physiological maturity should also be collected, in order to allow a full validation of the variable LGP product that was developed during this study.

5.6 Final comments

The incorporation of the methodologies that were designed during this research into operational seasonal monitoring procedures represents an opportunity for improvement of food security early warning systems. A new method for identifying 2-m air temperature extremes from thermal infrared brightness temperatures using simple statistical techniques was proposed and developed. The ability to identify temperature extremes at a moderate spatial resolution will be useful for informing decision-making in agriculture and health, where high temperatures can have severe impacts. The use of simple statistical techniques for these temperature estimates also means that the method would be easily applicable to many systems without requiring high computational power typical of radiative transfer models traditionally used to convert BT to land-surface temperature (LST).

A modified method for calculating WRSI using a variable-LGP was developed as part of this research. The use of this method within an operational crop monitoring system will improve the accuracy of the calculated WRSI values, especially when implemented within a framework of multiple planting dates. The improved accuracy of this important index will potentially enhance the capacity of early warning institutions to identify impending incidents of food insecurity.

Additionally, a number of new parameters were developed for this research that can be incorporated into operational food security early warning systems. In particular, the eGDD parameter and all its variants are useful temperature-based crop monitoring indicators. These parameters can be used not only as predictors of crop yield, but also for monitoring and

highlighting periods of high air temperature, which may be related to moisture and heat stress during the crop's growth cycle. This will be particularly useful given the increases in temperature that have been observed globally since the 1970s, as well as the increase in incidence of warm spells noted in some areas. From an operational perspective, the temperature-related seasonal monitoring grids that were developed in this study can be generated every ten days and made available to analysts and other researchers for testing and assessing their value. The system will thus be able to, in new ways, contribute to the provision of early warning regarding potential, weather-related causes of food insecurity.

REFERENCES

Abendroth LJ, Elmore RW, Boyer MJ, Marlay SK. 2011. Corn growth and development. Ames, Iowa, USA: PMR, Iowa State University Extension.

Abraha MG, Savage MJ. 2006. Potential impacts of climate change on the grain yield of maize for the midlands of KwaZulu-Natal, South Africa. Agriculture, Ecosystems and Environment 115:150-160.

Allen RG, Pereira LS, Raes D, Smith M. 1998. Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. FAO, Rome 300:D05109.

Anderson G, Kootval H, Kull D, Clements J, Fleming G, Frei T, Lazo J, Letson D, Mills B, Perrels A, Rogers D, Vaughan C, Zillman J. 2015. Valuing Weather and Climate: Economic Assessment of Meteorological and Hydrological Services. Geneva, Switzerland: World Meteorological Organization.

Arya PS. 2001. Introduction to Micrometeorology. San Diego, California, USA: Academic Press.

Barsi JA, Barker JL, Schott JR. 2003. An atmospheric correction parameter calculator for a single thermal band earth-sensing instrument. Pages 3014-3016. Geoscience and Remote Sensing Symposium. Toulouse, France: IEEE.

Beck HE, Zimmermann NE, McVicar TR, Vergopolan N, Berg A, Wood EF. 2018. Present and future Köppen-Geiger climate classification maps at 1-km resolution. Scientific data 5:180214.

Berk A, Anderson GP, Bernstein LS, Acharya PK, Dothe H, Matthew MW, Adler-Golden SM, Chetwynd Jr JH, Richtsmeier SC, Pukall B. 1999. MODTRAN 4 radiative transfer modeling for atmospheric correction. Pages 348-353. SPIE's International Symposium on Optical Science, Engineering, and Instrumentation. Denver, CO, United States.

Budde M, Verdin J, Galu G, Magadzire T, Pedreros D, Funk C, Husak G, Peterson P, Landsfeld M, White L. 2014. Building capacity for production of gridded precipitation products in the east africa community. AGU Fall Meeting Abstracts.

Caselles V, Coll C, Valor E. 1997. Land surface emissivity and temperature determination in the whole HAPEX-Sahel area from AVHRR data. International Journal of Remote Sensing 18:1009-1027.

Chen J, Chen J, Liao A, Cao X, Chen L, Chen X, He C, Han G, Peng S, Lu M. 2015. Global land cover mapping at 30 m resolution: A POK-based operational approach. ISPRS Journal of Photogrammetry and Remote Sensing 103:7-27.

Cleland EE, Chuine I, Menzel A, Mooney HA, Schwartz MD. 2007. Shifting plant phenology in response to global change. Trends in Ecology and Evolution 22:357-365.

Compo GP, Whitaker JS, Sardeshmukh PD, Matsui N, Allan RJ, Yin X, Gleason BE, Vose RS, Rutledge G, Bessemoulin P. 2011. The twentieth century reanalysis project. Quarterly Journal of the Royal Meteorological Society 137:1-28.

Dai A, Trenberth KE, Karl TR. 1999. Effects of clouds, soil moisture, precipitation, and water vapor on diurnal temperature range. Journal of Climate 12:2451-2473.

Davenport F, Grace K, Funk C, Shraddhanand S. 2017. Child health outcomes in sub-Saharan Africa: A comparison of changes in climate and socio-economic factors. Global Environmental Change 46:72-87.

De Pauw E, Nachtergaele F, Antoine J, Fisher G, Velthuizen H. 1996. A provisional world climatic resource inventory based on the length-of-growing-period concept. National soil reference collections and databases (NASREC):30-43.

Derrien M, Farki B, Harang L, Le Gléau H, Noyalet A, Pochic D, Sairouni A. 1993. Automatic cloud detection applied to NOAA-11/AVHRR imagery. Remote Sensing of Environment 46:246-267.

Derrien M, Le Gléau H. 2005. MSG/SEVIRI cloud mask and type from SAFNWC. International Journal of Remote Sensing 26:4707-4732.

Derrien M, Le Gléau H. 2007. Temporal-differencing and region-growing techniques to improve twilight low cloud detection from SEVIRI data. Pages 2428. Proceedings of the Joint 2007 EUMETSAT Meteorological Satellite Conference and the 15th Satellite Meteorology and

Oceanography Conference of the American Meteorological Society, Amsterdam, The Netherlands: Citeseer.

Doswell CA, Davies-Jones R, Keller DL. 1990. On summary measures of skill in rare event forecasting based on contingency tables. Weather and Forecasting 5:576-585.

Du Preez JH, Giesecke WH, Hattingh PJ. 1990. Heat stress in dairy cattle and other livestock under southern African conditions. I. Temperature-humidity index mean values during the four main seasons. The Onderstepoort Journal of Veterinary Research 57:77-87.

Duguay-Tetzlaff A, Bento VA, Göttsche FM, Stöckli R, Martins J, Trigo I, Olesen F, Bojanowski JS, da Camara C, Kunz H. 2015. Meteosat land surface temperature climate data record: achievable accuracy and potential uncertainties. Remote Sensing 7:13139-13156.

EUMETSAT. 2018. EUMETSAT Data Centre - EUMETSAT. (20 January 2018; https://www.eumetsat.int/website/home/Data/DataDelivery/EUMETSATDataCentre/index.html)

FAO. 2009. Declaration of the World Summit on Food Security. World Summit on Food Security 2009. Rome: FAO.

Farouki OT. 1981. Thermal properties of soils. Hanover, NH, USA: U.S. Army Corps of Engineers, Cold Regions Research and Engineering Laboratory.

Feijt A, De Valk P, Van Der Veen S. 2000. Cloud detection using Meteosat imagery and numerical weather prediction model data. Journal of Applied Meteorology 39:1017-1030.

Fermont A, Benson T. 2011. Estimating yield of food crops grown by smallholder farmers. International Food Policy Research Institute, Washington DC:1-68.

French A, Norman J, Anderson M. 2003. A simple and fast atmospheric correction for spaceborne remote sensing of surface temperature. Remote Sensing of Environment 87:326-333.

Frey RA, Ackerman SA, Liu Y, Strabala KI, Zhang H, Key JR, Wang X. 2008. Cloud detection with MODIS. Part I: Improvements in the MODIS cloud mask for collection 5. Journal of Atmospheric and Oceanic Technology 25:1057-1072.

Funk C, Peterson P, Landsfeld M, Pedreros D, Verdin J, Shukla S, Husak G, Rowland J, Harrison L, Hoell A. 2015. The climate hazards infrared precipitation with stations--a new environmental record for monitoring extremes. Scientific Data 2.

Funk C, Peterson P, Peterson S, Shukla S, Davenport F, Michaelsen J, Landsfeld M, Husak G, Harrison L, Pomposi C, Nicholson S, Pedreros D, Magadzire T, Knapp K, Dinku T. in internal review. A new 1983-present satellite-plus-in situ observation high resolution Tmax environmental data record.

Gallo K, Hale R, Tarpley D, Yu Y. 2011. Evaluation of the Relationship between Air and Land Surface Temperature under Clear- and Cloudy-Sky Conditions. Journal of Applied Meteorology and Climatology 50:767-775.

Göttsche F-M, Olesen FS. 2001. Modelling of diurnal cycles of brightness temperature extracted from METEOSAT data. Remote Sensing of Environment 76:337-348.

Grace K, Davenport F, Hanson H, Funk C, Shukla S. 2015. Linking climate change and health outcomes: Examining the relationship between temperature, precipitation and birth weight in Africa. Global Environmental Change 35:125-137.

Grimes D, Pardo-Iguzquiza E, Bonifacio R. 1999. Optimal areal rainfall estimation using raingauges and satellite data. Journal of Hydrology 222:93-108.

Hansen J, Lebedeff S. 1987. Global trends of measured surface air temperature. Journal of Geophysical Research 92:345-313.

Harrison L, Michaelsen J, Funk C, Husak G. 2011. Effects of temperature changes on maize production in Mozambique. Climate Research 46:211.

Hatfield JL, Prueger JH. 2015. Temperature extremes: Effect on plant growth and development. Weather and Climate Extremes 10:4-10.

Hocking J, Francis PN, Saunders R. 2011. Cloud detection in Meteosat second generation imagery at the Met Office. Meteorological Applications 18:307-323.

Holdaway MR. 1996. Spatial modeling and interpolation of monthly temperature using kriging. Climate Research 6:215-225.

IEG. 2013. Adapting to climate change: Assessing the World Bank Group experience. Washington, DC, USA: World Bank.

Janatian N, Sadeghi M, Sanaeinejad SH, Bakhshian E, Farid A, Hasheminia SM, Ghazanfari S. 2017. A statistical framework for estimating air temperature using MODIS land surface temperature data. International Journal of Climatology 37:1181-1194.

Janowiak JE, Joyce RJ, Yarosh Y. 2001. A real–time global half–hourly pixel–resolution infrared dataset and its applications. Bulletin of the American Meteorological Society 82:205-217.

Jayne TS, Rashid S. 2010. The value of accurate crop production forecasts. Agricultural Risks Management in Africa: Taking Stock of What Has and Hasn't Worked. Lilongwe, Malawi.

Johnson H, Kovats RS, McGregor G, Stedman J, Gibbs M, Walton H, Cook L, Black E. 2005. The impact of the 2003 heat wave on mortality and hospital admissions in England. Health Statistics Quarterly 25:6-11.

Jones JW, Hoogenboom G, Porter CH, Boote KJ, Batchelor WD, Hunt L, Wilkens PW, Singh U, Gijsman AJ, Ritchie JT. 2003. The DSSAT cropping system model. European Journal of Agronomy 18:235-265.

Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J. 1996. The NCEP/NCAR 40-year reanalysis project. Bulletin of the American Meteorological Society 77:437-471.

Karl TR, Knight RW, Gallo KP, Peterson TC, Jones PD, Kukla G, Plummer N, Razuvayev V, Lindseay J, Charlson RJ. 1993. A new perspective on recent global warming: asymmetric trends of daily

maximum and minimum temperature. Bulletin of the American Meteorological Society 74:1007-1023.

Kaspar F, Helmschrot J, Mhanda A, Butale M, de Clercq W, Kanyanga J, Neto F, Kruger S, Castro Matsheka M, Muche G. 2015. The SASSCAL contribution to climate observation, climate data management and data rescue in Southern Africa. Advances in Science and Research 12:171-177.

Knapp KR. 2008. Scientific data stewardship of International Satellite Cloud Climatology Project B1 global geostationary observations. Journal of Applied Remote Sensing 2:023548.

Knapp KR, Ansari S, Bain CL, Bourassa MA, Dickinson MJ, Funk C, Helms CN, Hennon CC, Holmes CD, Huffman GJ. 2011. Globally gridded satellite observations for climate studies. Bulletin of the American Meteorological Society 92:893-907.

Kotarba AZ. 2009. A comparison of MODIS-derived cloud amount with visual surface observations. Atmospheric Research 92:522-530.

Kranz W, Irmak S, van Donk S, Yonts C, Martin D. 2008. Irrigation Management for Corn. NebGuide G1850. Lincoln, Neb.: University of Nebraska–Lincoln. Lincoln, Nebraska, USA: University of Nebraska–Lincoln.

Li J, Heap AD. 2014. Spatial interpolation methods applied in the environmental sciences: A review. Environmental Modelling and Software 53:173-189.

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

Liang S. 2005. Quantitative remote sensing of land surfaces. Hoboken, New Jersey, USA: John Wiley & Sons.

Lipiec J, Hatano R. 2003. Quantification of compaction effects on soil physical properties and crop growth. Geoderma 116:107-136.

Lobell DB, Bänziger M, Magorokosho C, Vivek B. 2011. Nonlinear heat effects on African maize as evidenced by historical yield trials. Nature Climate Change 1:42-45.

Lobell DB, Field CB. 2007. Global scale climate–crop yield relationships and the impacts of recent warming. Environmental Research Letters 2:014002.

Loveland TR, Reed BC, Brown JF, Ohlen DO, Zhu Z, Yang L, Merchant JW. 2000. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. International Journal of Remote Sensing 21:1303-1330.

Magadzire T, Galu G, Verdin J. 2017. How climate forecasts strengthen food security. WMO Bulletin 66:10-15.

Manatsa D, Nyakudya IW, Mukwada G, Matsikwa H. 2011. Maize yield forecasting for Zimbabwe farming sectors using satellite rainfall estimates. Natural Hazards 59:447-463.

Massart M, Rembold F, Rojas O, Leo O. 2010. The use of remote sensing data and meteorological information for food security monitoring, examples in East Africa. Advances in Earth Observation of Global Change, Springer. pp 201-216.

Matsuura K, Willmott CJ. 2015. Terrestrial air temperature: 1900-2014 gridded monthly time series. (http://climate.geog.udel.edu/~climate/html_pages/Global2014/README.GlobalTsT2014.html)

McMaster GS, Wilhelm W. 1997. Growing degree-days: one equation, two interpretations. Agricultural and Forest Meteorology 87:291-300.

Minasny B, McBratney A. 2018. Limited effect of organic matter on soil available water capacity. European Journal of Soil Science 69:39-47.

Minder JR, Mote PW, Lundquist JD. 2010. Surface temperature lapse rates over complex terrain: Lessons from the Cascade Mountains. Journal of Geophysical Research: Atmospheres 115:1-13.

Minnis P, Harrison EF. 1984. Diurnal Variability of Regional Cloud and Clear-Sky Radiative Parameters Derived from GOES Data. Part I: Analysis Method. Journal of Climate and Applied Meteorology 23:993-1011.

Mira M, Valor E, Boluda R, Caselles V, Coll C. 2007. Influence of soil water content on the thermal infrared emissivity of bare soils: Implication for land surface temperature determination. Journal of Geophysical Research: Earth Surface 112.

Moeletsi ME. 2017. Mapping of maize growing period over the Free State Province of South Africa: heat units approach. Advances in Meteorology 2017:1-11.

Moot D, Scott W, Roy A, Nicholls A. 2000. Base temperature and thermal time requirements for germination and emergence of temperate pasture species. New Zealand Journal of Agricultural Research 43:15-25.

Mukhala E, Hoefsloot P. 2004. AgroMetShell manual. Rome: FAO.

Myers DE. 1982. Matrix formulation of co-kriging. Journal of the International Association for Mathematical Geology 14:249-257.

Nafziger E. 2009. Corn. Illinois Agronomy Handbook, vol. 1394. Champaign and Urbana, USA: University of Illinois at Urbana-Champaign, College of Agriculture, Cooperative Extension Service. pp 13-26.

Nalder IA, Wein RW. 1998. Spatial interpolation of climatic normals: test of a new method in the Canadian boreal forest. Agricultural and Forest Meteorology 92:211-225.

Ngwira A, Aune JB, Thierfelder C. 2014. DSSAT modelling of conservation agriculture maize response to climate change in Malawi. Soil and Tillage Research 143:85-94.

NOAA_ESRL. Twentieth Century Reanalysis (V2): 4-times Daily and Daily Average Monolevel. (23 January 2018; https://www.esrl.noaa.gov/psd/data/gridded/data.20thC_ReanV2.monolevel.html)

NOAA_NCEI. 2018. Global Surface Summary of the Day - GSOD. (19 February 2018 https://data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.ncdc:C00516)

Panda DK, AghaKouchak A, Ambast SK. 2017. Increasing heat waves and warm spells in India, observed from a multiaspect framework. Journal of Geophysical Research: Atmospheres 122:3837-3858.

Pannar. 2017. Pannar 2017 Product Catalogue in Pannar, ed: Pannar.

Parton WJ, Logan JA. 1981. A model for diurnal variation in soil and air temperature. Agricultural Meteorology 23:205-216.

Phillips DL, Marks DG. 1996. Spatial uncertainty analysis: propagation of interpolation errors in spatially distributed models. Ecological Modelling 91:213-229.

Ramanathan VLRD, Cess RD, Harrison EF, Minnis P, Barkstrom BR, Ahmad E, Hartmann D. 1989. Cloud-radiative forcing and climate: Results from the Earth Radiation Budget Experiment. Science 243:57-63.

Reynolds C, Jackson T, Rawls W. 2000a. Estimating soil water-holding capacities by linking the Food and Agriculture Organization soil map of the world with global pedon databases and continuous pedotransfer functions. Water Resources Research 36:3653-3662.

Reynolds C, Yitayew M, Slack D, Hutchinson C, Huete A, Petersen M. 2000b. Estimating crop yields and production by integrating the FAO Crop Specific Water Balance model with real-time satellite data and ground-based ancillary data. International Journal of Remote Sensing 21:3487-3508.

Rojas O, Rembold F, Royer A, Negre T. 2005. Real-time agrometeorological crop yield monitoring in Eastern Africa. Agronomy for Sustainable Development 25:63-77.

Rossow WB, Garder LC. 1993. Cloud detection using satellite measurements of infrared and visible radiances for ISCCP. Journal of Climate 6:2341-2369.

Rossow WB, Mosher F, Kinsella E, Arking A, Desbois M, Harrison E, Minnis P, Ruprecht E, Seze G, Simmer C. 1985. ISCCP cloud algorithm intercomparison. Journal of Climate and Applied Meteorology 24:877-903.

Rossow WB, Schiffer RA. 1999. Advances in understanding clouds from ISCCP. Bulletin of the American Meteorological Society 80:2261-2287.

Rurinda J, Mapfumo P, van Wijk MT, Mtambanengwe F, Rufino MC, Chikowo R, Giller KE. 2013. Managing soil fertility to adapt to rainfall variability in smallholder cropping systems in Zimbabwe. Field Crops Research 154:211-225.

SADC. 2017. SADC regional vulnerability assessment and analysis synthesis report 2017 - State of food security and vulnerability in the Southern African Development Community. Gaborone, Botswana: SADC.

Saha S, Moorthi S, Pan H-L, Wu X, Wang J, Nadiga S, Tripp P, Kistler R, Woollen J, Behringer D. 2010. The NCEP climate forecast system reanalysis. Bulletin of the American Meteorological Society 91:1015-1058.

Saha S, Moorthi S, Wu X, Wang J, Nadiga S, Tripp P, Behringer D, Hou Y-T, Chuang H-y, Iredell M. 2014. The NCEP climate forecast system version 2. Journal of Climate 27:2185-2208.

Savage MJ. 2016. Nowcasting daily minimum air and grass temperature. International Journal of Biometeorology 60:183-194.

Schädlich S, Göttsche F-M, Olesen FS. 2001. Influence of land surface parameters and atmosphere on METEOSAT brightness temperatures and generation of land surface temperature maps by temporally and spatially interpolating atmospheric correction. Remote Sensing of Environment 75:39-46.

Senay GB, Verdin J. 2002. Evaluating the performance of a crop water balance model in estimating regional crop production. Proceedings of the Pecora.

Senay GB, Verdin J. 2003. Characterization of yield reduction in Ethiopia using a GIS-based crop water balance model. Canadian Journal of Remote Sensing 29:687-692.

Seneviratne SI, Nicholls N, Easterling D, Goodess CM, Kanae S, Kossin J, Luo Y, Marengo J, McInnes K, Rahimi M, Reichstein M, Sorteberg A, Vera C, Zhang X. 2012. Changes in climate extremes and their impacts on the natural physical environment. In: Field CB, et al., eds. Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation: A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC). Cambridge, UK, and New York, NY, USA: Cambridge University Press. pp 109-230.

Snow JT, Biagini B, Benchwick G, George G, Hoedjes J, Miller A, Usher J. 2016. A New Vision for Weather and Climate Services in Africa. New York, USA: UNDP.

Soane B, Van Ouwerkerk C. 1994. Soil compaction problems in world agriculture. In: Soane B, Van Ouwerkerk C, eds. Soil Compaction in Crop Production, vol. 11. Amsterdam, The Netherlands: Elsevier. pp 1-21.

Soler CT, Sentelhas P, Hoogenboom G. 2005. Thermal time for phenological development of four maize hybrids grown off-season in a subtropical environment. The Journal of Agricultural Science 143:169-182.

Spectral_Sciences_Incorporated. MODTRAN Demo. (23 January 2018; http://modtran.spectral.com/modtran_home#plot)

Steduto P, Hsiao TC, Raes D, Fereres E. 2009. AquaCrop—The FAO crop model to simulate yield response to water: I. Concepts and underlying principles. Agronomy Journal 101:426-437.

Stocker T, Qin D, Plattner G, Tignor M, Allen S, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM. 2013. IPCC, 2013: Summary for Policymakers. Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.

Tardieu F, Manichon H. 1987. Etat structural, enracinement et alimentation hydrique du maïs. II.--Croissance et disposition spatiale du système racinaire. Agronomie 7:201-211.

Taylor H, Brar G. 1991. Effect of soil compaction on root development. Soil and Tillage Research 19:111-119.

Tjemkes SA. 2005. On the conversion from radiances to equivalent brightness temperatures. (24 January 2017; https://www.eumetsat.int/website/wcm/idc/idcplg?IdcService=GET_FILE&dDocName=PDF_MSG_SEVIRI_RAD2BRIGHT&RevisionSelectionMethod=LatestReleased)

Trenberth KE, Shea DJ. 2005. Relationships between precipitation and surface temperature. Geophysical Research Letters 32.

Trepte Q, Minnis P, Heck R, Palikonda R. 2005. Improvements in Near-Terminator and Nocturnal Cloud Masks using Satellite Image Data over the Atmospheric Radiation Measurement Sites. Fifteenth ARM Science Team Meeting Proceedings, Daytona Beach, Florida , March 14-18, 2005 Florida: NASA.

Trudgill D, Honek A, Li D, Straalen N. 2005. Thermal time–concepts and utility. Annals of Applied Biology 146:1-14.

USGS. 2007. Croplands Water Requirement Satisfaction Index (WRSI) product documentation. (15 March 2018; https://earlywarning.usgs.gov/fews/product/128#documentation)

Van de Griend A, Owe M. 1993. On the relationship between thermal emissivity and the normalized difference vegetation index for natural surfaces. International Journal of Remote Sensing 14:1119-1131.

Verdin J, Klaver R. 2002. Grid-cell-based crop water accounting for the famine early warning system. Hydrological Processes 16:1617-1630.

Vickery G. 2011. Review of recent studies on PSI re-use and related market developments. Information Economics. Paris: European Commission.

Vina A, Gitelson AA, Rundquist DC, Keydan G, Leavitt B, Schepers J. 2004. Monitoring maize (Zea mays L.) phenology with remote sensing. Agronomy Journal 96:1139-1147.

Wang E, Robertson M, Hammer G, Carberry P, Holzworth D, Meinke H, Chapman S, Hargreaves J, Huth N, McLean G. 2002. Development of a generic crop model template in the cropping system model APSIM. European Journal of Agronomy 18:121-140.

Wang J, Williams G, Guo Y, Pan X, Tong S. 2013. Maternal exposure to heatwave and preterm birth in Brisbane, Australia. BJOG: An International Journal of Obstetrics and Gynaecology 120:1631-1641.

Weinreb M, Han D. 2011. Conversion of GVAR infrared data to scene radiance or temperature. (26 May 2018; https://www.ospo.noaa.gov/Operations/GOES/calibration/gvar-conversion.html)

West J. 2003. Effects of heat-stress on production in dairy cattle. Journal of Dairy Science 86:2131-2144.

WMO. 2014. Forecast Verification for the African Severe Weather Forecasting Demonstration Projects. Geneva: World Meteorological Organization.

WMO. 2015. The Climate in Africa: 2013 World Meteorological Organization.

Xie P, Arkin PA. 1996. Analyses of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions. Journal of Climate 9:840-858.

Yang H, Dobermann A, Lindquist JL, Walters DT, Arkebauer TJ, Cassman KG. 2004. Hybrid-maize—a maize simulation model that combines two crop modeling approaches. Field Crops Research 87:131-154.

Zhuge X, Zou X. 2016. Test of a modified infrared-only ABI cloud mask algorithm for AHI radiance observations. Journal of Applied Meteorology and Climatology 55:2529-2546.

APPENDIX A: DEFINITION OF CONFUSION MATRIX STATISTICS

Table 2.3 of Chapter 2 lists a number of statistics for evaluating the performance of the TIR-based cloud classifications, using the SAFNWC cloud mask as truth. The statistics used in the evaluation are explained in this appendix. For quick reference, Table 2.3 has been included below.

Table 2.3 Statistical performance comparing TIR-based cloud classifications with SAFNWC cloud mask, based on results of confusion matrix

Accuracy	85.4%
Misclassification rate (error rate)	11.7%
True positive rate (sensitivity)	78.6%
False positive rate	7.6%
Specificity	90.0%
Precision	87.4%
Prevalence	40.3%

The statistics show in Table 2.3 are calculated based on the outcome of a confusion matrix comparing a model-estimated classification with an observed classification. In the case of cloud classification, a positive classification is defined as the presence of cloud, while a negative classification is the absence of cloud. Any one of four possible scenarios is possible. :

1. True positive (TP): the model correctly identified a positive classification (in this case, cloud was correctly identified)
2. True negative (TN): the model correctly identified a negative classification (in this case, absence of cloud was correctly identified)

3. False positive (FP): the model made a positive classification, but observations made a negative classification
4. False negative (FN): the model made a negative classification, but observations made a positive classification

Positively classified observations, or cloudy observations are denoted as OP, while negatively classified observations, or clear-sky conditions are denoted as ON. Modelled positive classifications are denoted as MP, while modelled negative classifications are denoted as MN

Using this criteria, the statistics used in Table 2.3 as follows:

Accuracy: the number of all correct classifications, both positive and negative (i.e., *TP* and *TN*, in this case, both cloudy conditions and clear-sky conditions respectively) as a fraction of the total number of classified observations, *TO*.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}, \quad \text{(A1)}$$

Misclassification rate: The total number of incorrectly classified observations (i.e., *FP* and *FN*) as a faction of the total number of classified observations

$$Missclassification\ Rate = \frac{FP+FN}{TP+TN+FP+FN}, \quad \text{(A2)}$$

True positive rate: The fraction of all positively classified observations (i.e. observed cloudy conditions) that were correctly classified by the model

$$True\ Positive\ Rate = \frac{TP}{OP}, \quad \text{(A3)}$$

False positive rate: The fraction of all negatively classified observations (i.e. observed clear-sky conditions) that were incorrectly classified as positive (i.e. cloudy) by the model

$$False\ Positive\ Rate = \frac{FP}{ON}, \quad \text{(A4)}$$

Specificity: The fraction of all negatively classified observations (i.e. observed clear-sky conditions) that were correctly classified by the model

$$Specificity = \frac{TN}{ON}, \quad \text{(A5)}$$

Precision: The fraction of all modelled positive classifications that were correctly classified

$$Precision = \frac{TP}{MP}, \quad \text{(A6)}$$

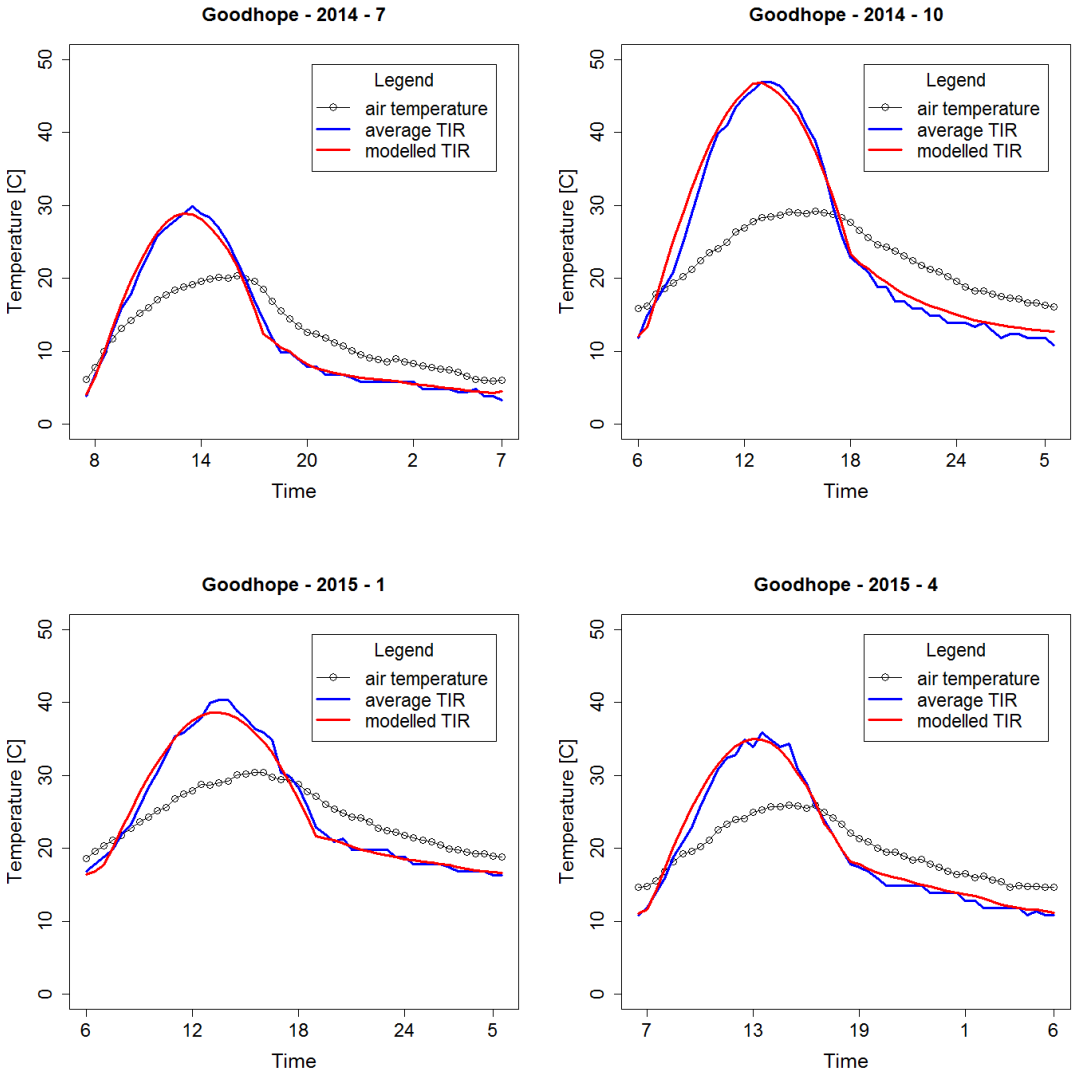
Prevalence: The fraction of all observations that are positively classified observations

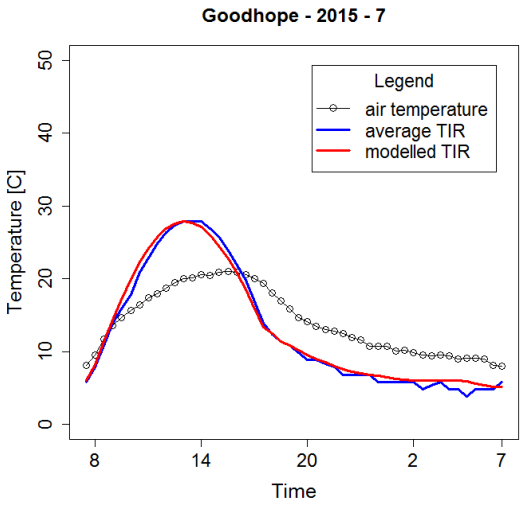
$$Prevalence = \frac{OP}{TP+TN+FP+FN}, \quad \text{(A7)}$$

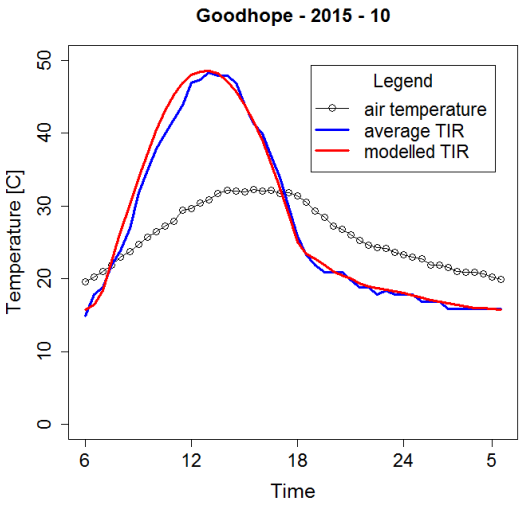
The different statistics are used to ascertain various aspects of the quality of the modelled classification. For example, the accuracy gives an indication of the model's overall ability to correctly classify both positive and negative conditions, while the false positive rate and the specificity both analyse the model's capability to identify negative, or in this case, clear-sky conditions.

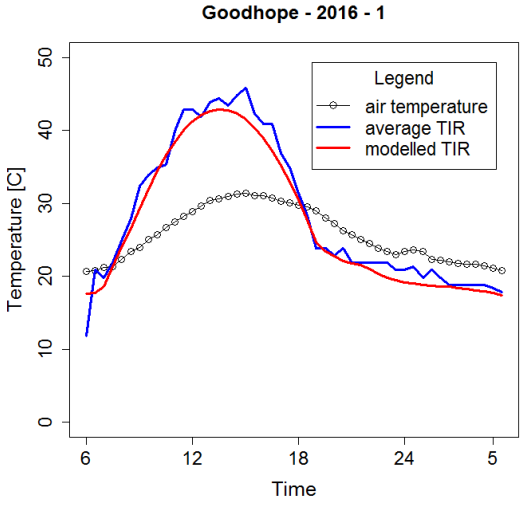
APPENDIX B: MEDIAN MONTHLY DTC

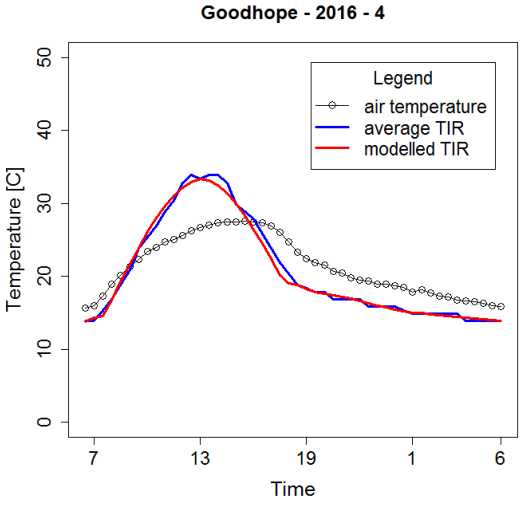
Figure 3.2 provided a subset of median monthly DTC for observed station air temperature, cloud-free TIR temperature, and modelled TIR temperature. The full set of median monthly DTCs that were developed for this study is presented in this appendix.

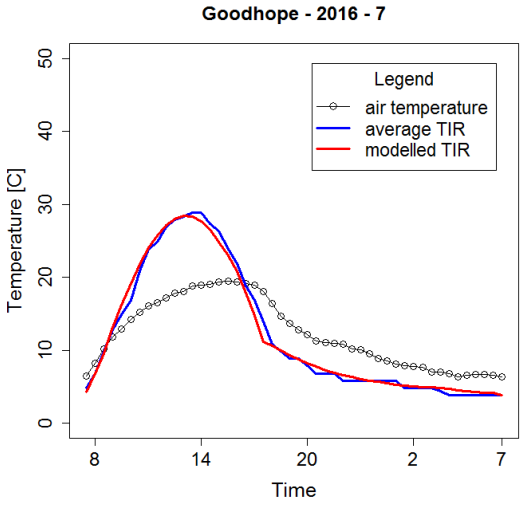


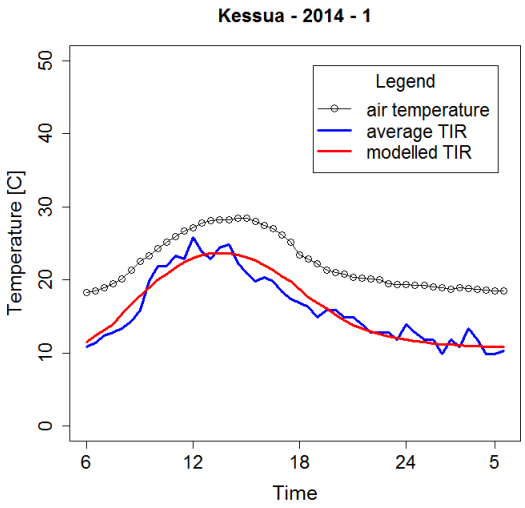

Goodhope - 2015 - 7
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
8
14
20
2
7
Time

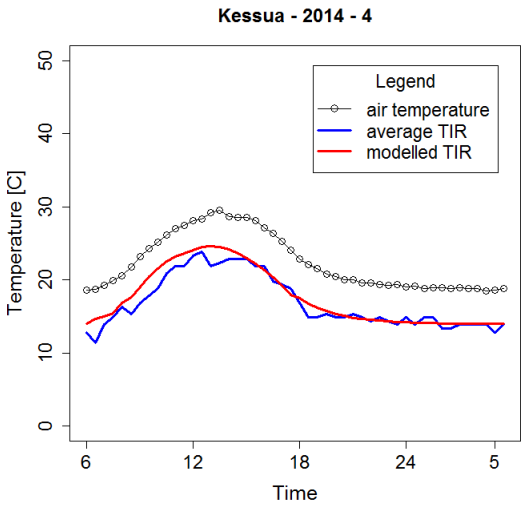

Goodhope - 2015 - 10
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
6
12
18
24
5
Time

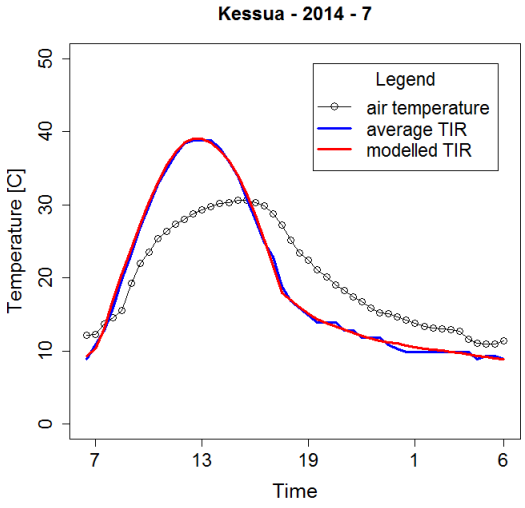

Goodhope - 2016 - 1
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
6
12
18
24
5
Time

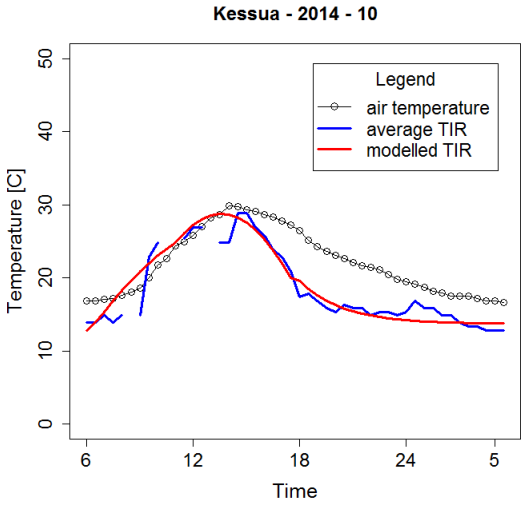

Goodhope - 2016 - 4
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
7
13
19
1
6
Time

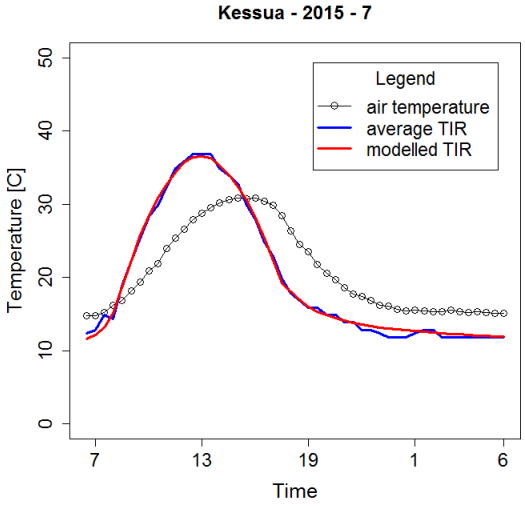

Goodhope - 2016 - 7
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
8
14
20
2
7
Time

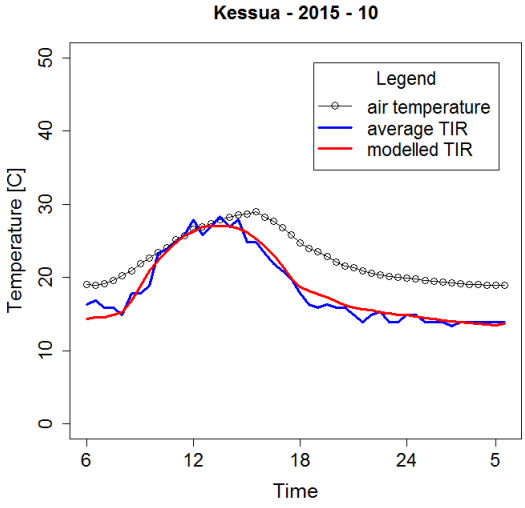

Kessua - 2014 - 1
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
6
12
18
24
5
Time

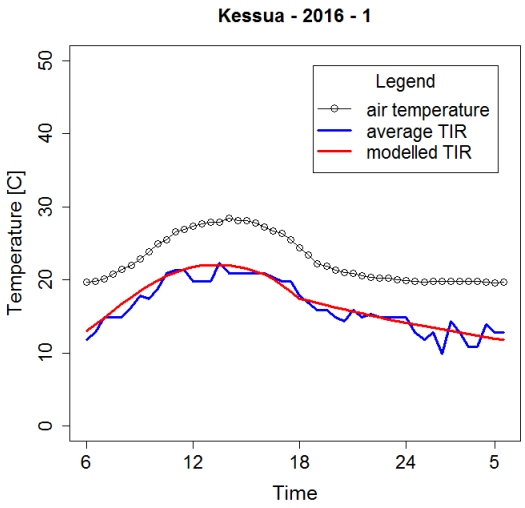

Kessua - 2014 - 4
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
6
12
18
24
5
Time

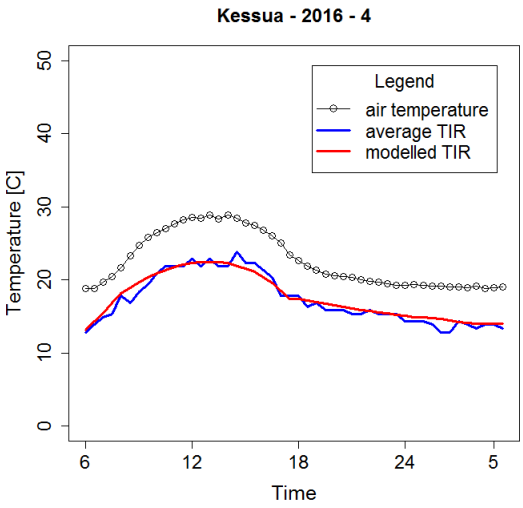

Kessua - 2014 - 7
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
7
13
19
1
6
Time

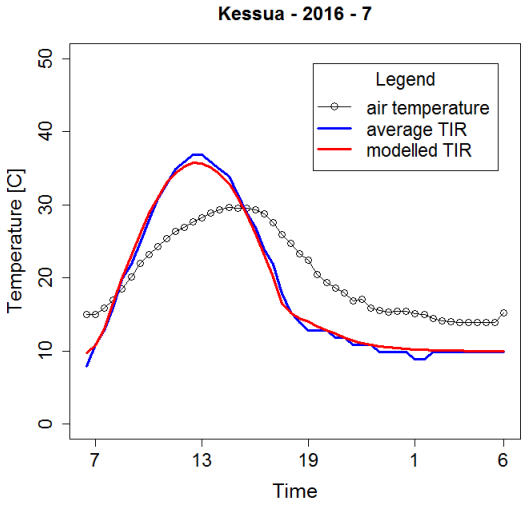

Kessua - 2014 - 10
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
6
12
18
24
5
Time

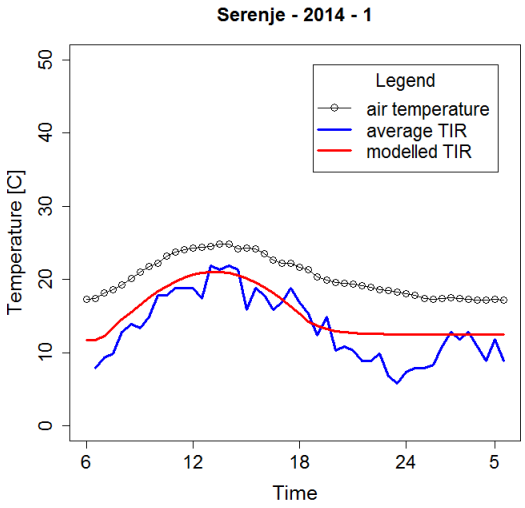

Kessua - 2015 - 7
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
7
13
19
1
6
Time

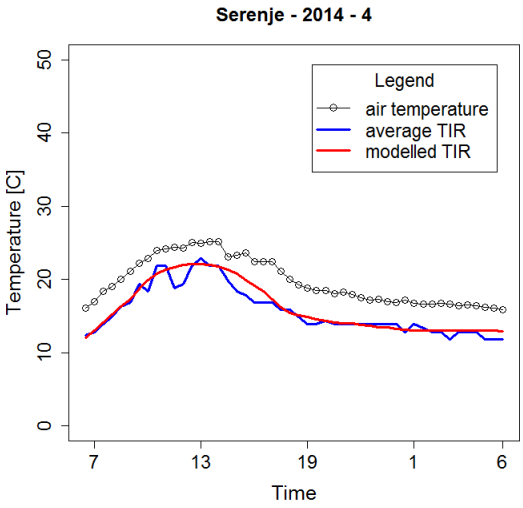

Kessua - 2015 - 10
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
6
12
18
24
5
Time

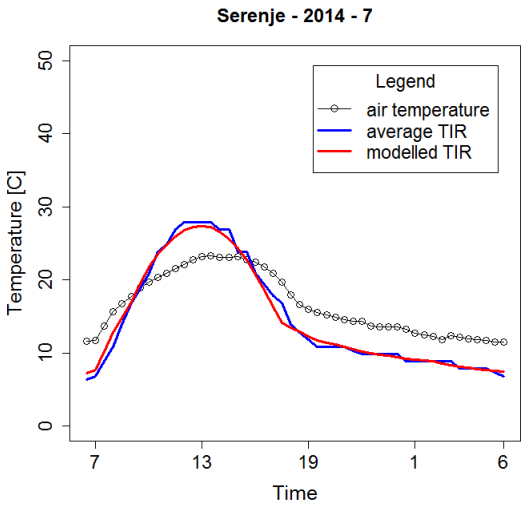

Kessua - 2016 - 1
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
6
12
18
24
5
Time

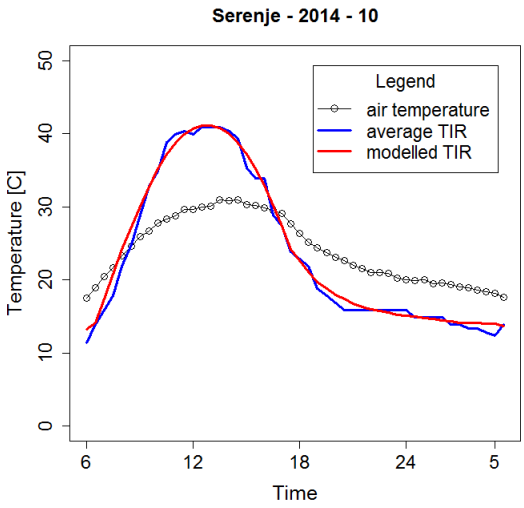

Kessua - 2016 - 4
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
6
12
18
24
5
Time

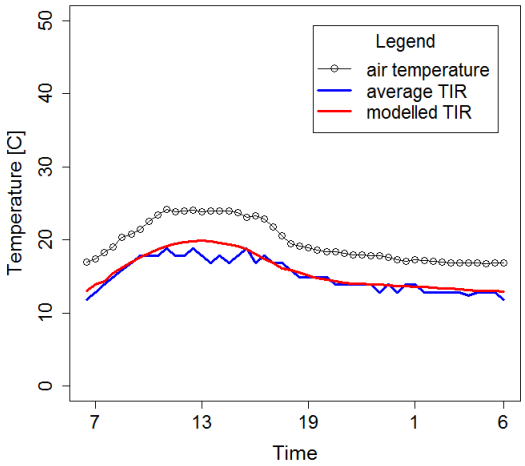

Kessua - 2016 - 7
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
7
13
19
1
6
Time

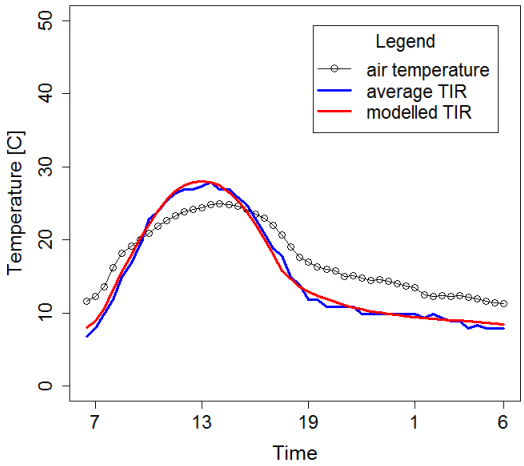

Serenje - 2014 - 1
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
6
12
18
24
5
Time

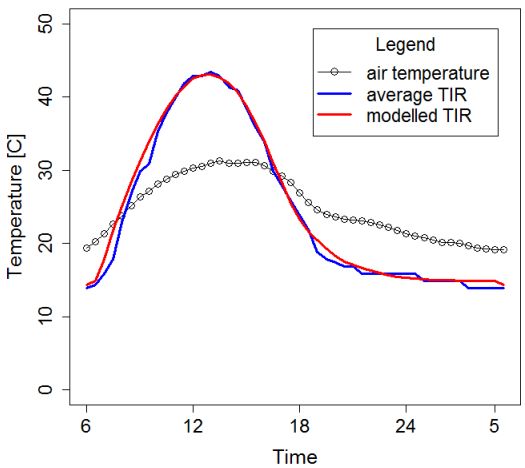

Serenje - 2014 - 4
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
7
13
19
1
6
Time

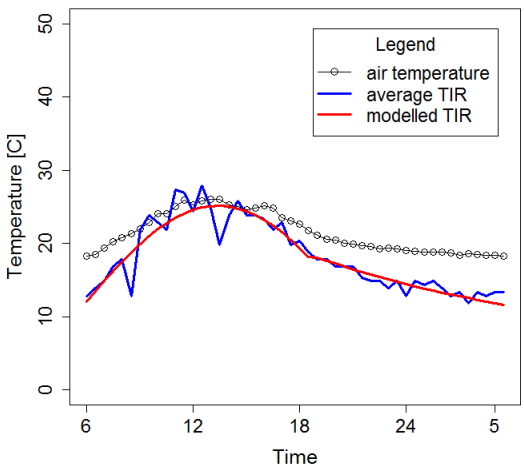

Serenje - 2014 - 7
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
7
13
19
1
6
Time

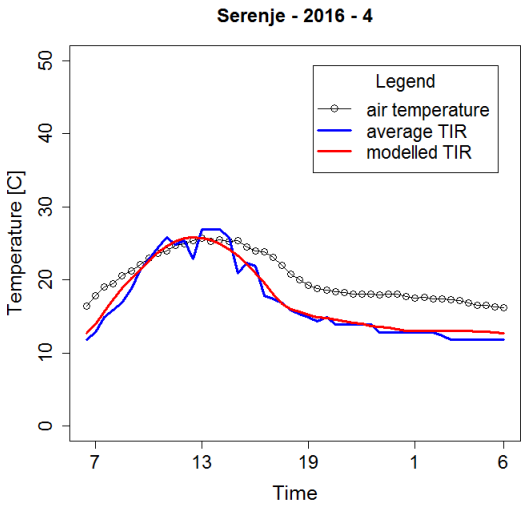

Serenje - 2014 - 10
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
6
12
18
24
5
Time

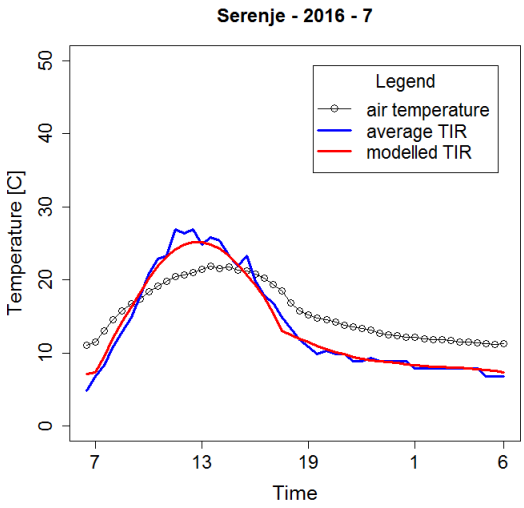

Serenje - 2015 - 4
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
7
13
19
1
6
Time

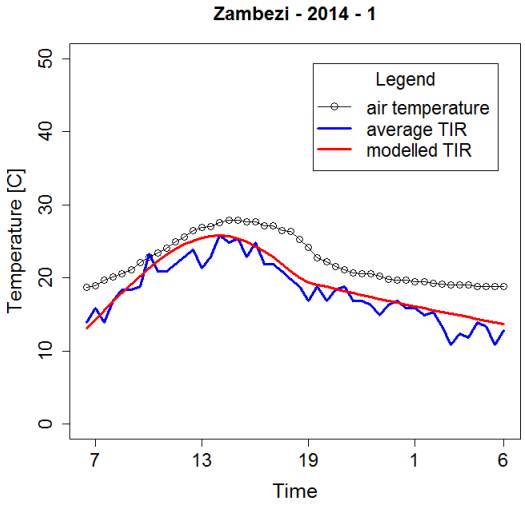

Serenje - 2015 - 7
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
7
13
19
1
6
Time

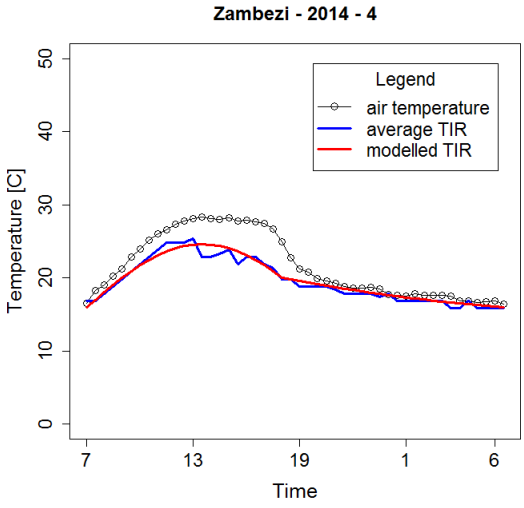

Serenje - 2015 - 10
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
6
12
18
24
5
Time

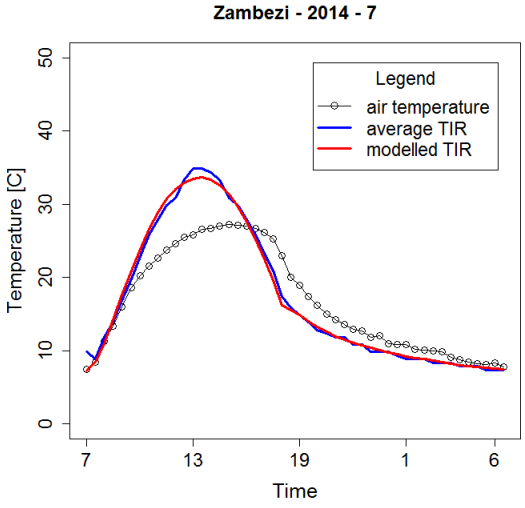

Serenje - 2016 - 1
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
6
12
18
24
5
Time

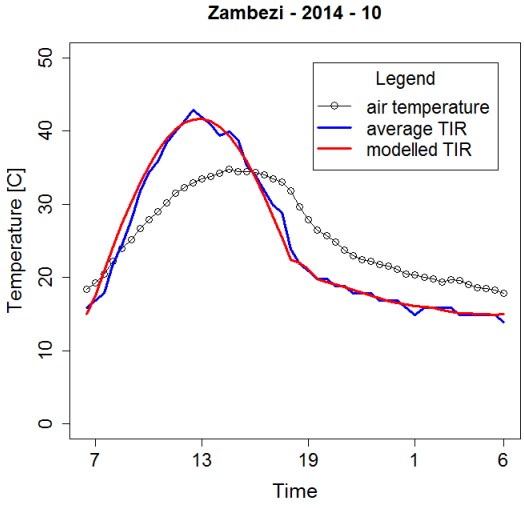

Serenje - 2016 - 4
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
7
13
19
1
6
Time

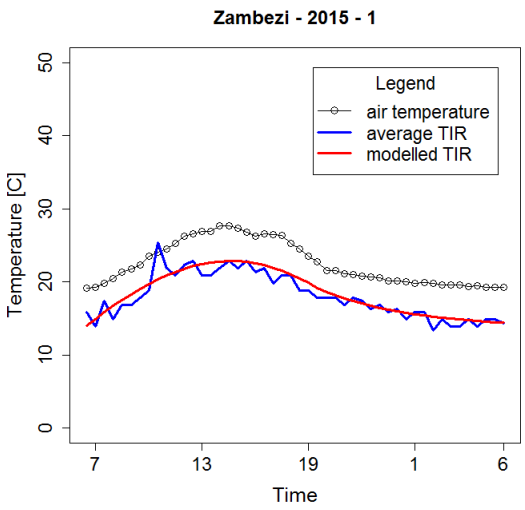

Serenje - 2016 - 7
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
7
13
19
1
6
Time

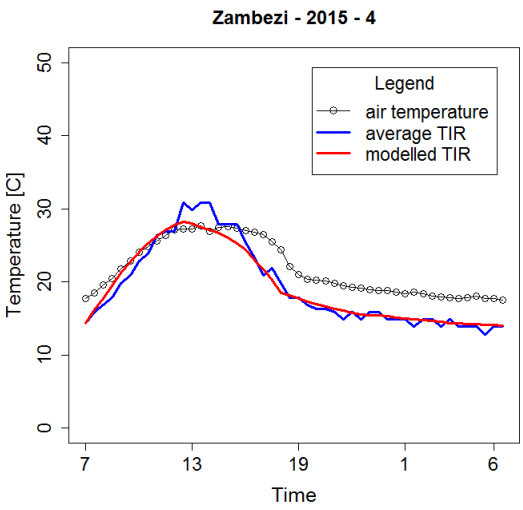

Zambezi - 2014 - 1
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
7
13
19
1
6
Time

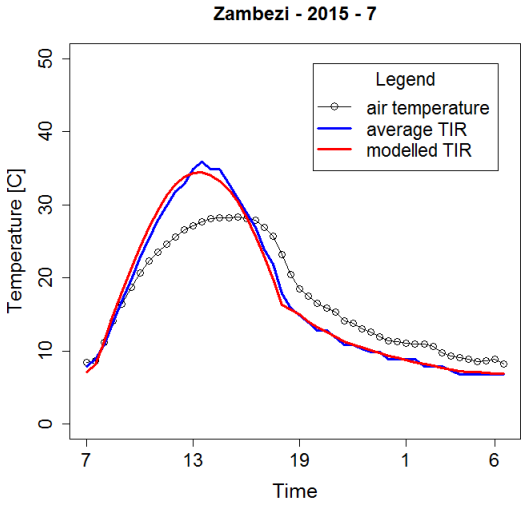

Zambezi - 2014 - 4
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
7
13
19
1
6
Time

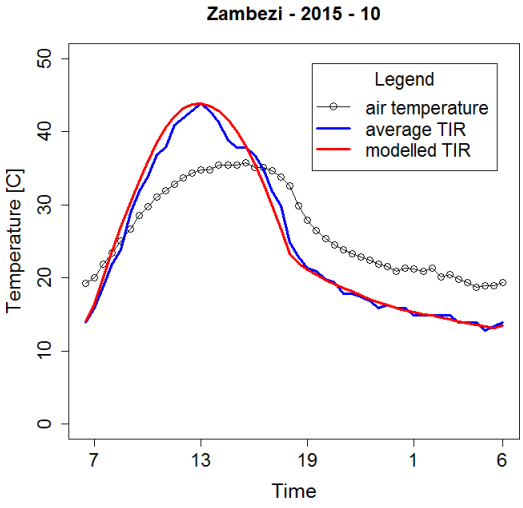

Zambezi - 2014 - 7
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
7
13
19
1
6
Time

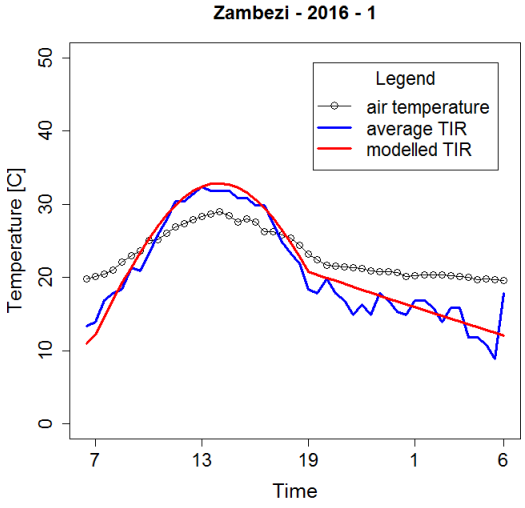

Zambezi - 2014 - 10
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
7
13
19
1
6
Time

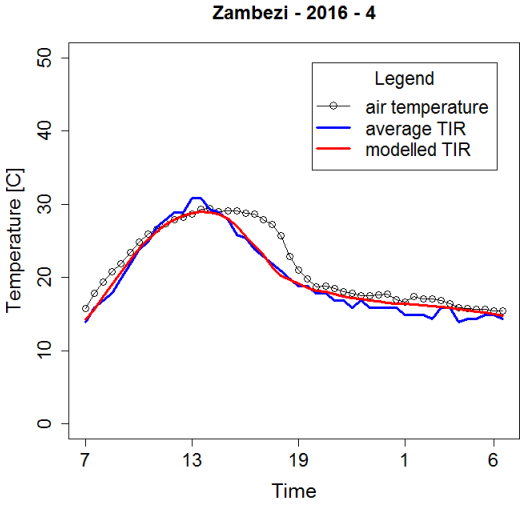

Zambezi - 2015 - 1
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
7
13
19
1
6
Time

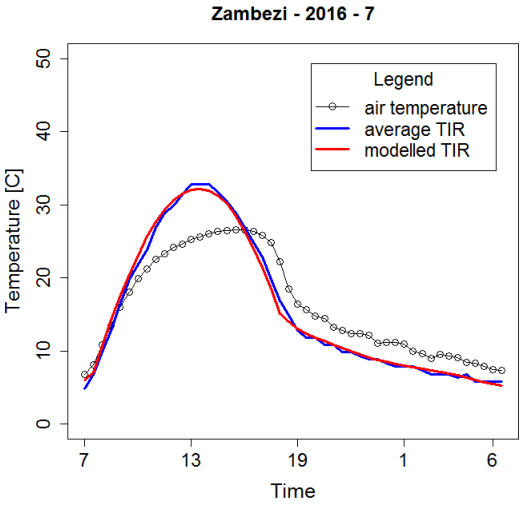

Zambezi - 2015 - 4
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
7
13
19
1
6
Time


Zambezi - 2015 - 7
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
7
13
19
1
6
Time


Zambezi - 2015 - 10
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
7
13
19
1
6
Time


Zambezi - 2016 - 1
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
7
13
19
1
6
Time


Zambezi - 2016 - 4
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
7
13
19
1
6
Time


Zambezi - 2016 - 7
Legend
air temperature
average TIR
modelled TIR
Temperature [C]
0
10
20
30
40
50
7
13
19
1
6
Time

APPENDIX C: DEFINITION OF CONTINGENCY STATISTICS USED IN TABLE 3.2

Table 3.2 of Chapter 3 lists a number of statistics for evaluating the performance of various versions of TIR-based temperatures in estimating the occurrence of temperatures over 30 ℃, using air temperature observations as truth. The following statistics were used in evaluating the different estimators: hit, miss, false alarm, correct negative, accuracy, hit rate, frequency bias, false alarm rate, false alarm ratio, Heidke skill score (HSS). Several of these statistics can be cross-referenced to those used in Appendix A, but using a different nomenclature. The nomenclature used here is typically used in categorical forecasting. The definitions and equations used in this appendix are adapted from WMO (2014) and modified to fit this study. The statistics generally differentiate between a positive and negative classification. In this instance, a positive classification is one of temperatures over 30 ℃, while a negative classification is one of temperatures under 30 ℃. These terms are defined below:

Hit: the estimator correctly identified a positive classification (temperatures above 30 ℃). A hit is the same as a "True Positive" (*TP*) in Appendix A

Miss: the estimator made a negative classification, but observations made a positive classification. A miss is the same as a "False Negative" (*FN*) in Appendix A

False alarm: the model made a positive classification, but observations made a negative classification: A false alarm is the same as a "False Positive" (*FP*) in Appendix A

Correct negative: the model correctly identified a negative classification (temperatures below 30 ℃). A correct negative is the same as a "True Negative" (*TN*) in Appendix A

Accuracy: the number of all correct classifications, both positive and negative (i.e., *TP* and *TN*) as a fraction of the total number of classified observations, *TO*.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}, \quad \text{(C1)}$$

Hit rate: The fraction of all positively classified observations (i.e observed temperatures above 30 ℃) that were correctly classified by the estimator. Defined as the *True positive rate* in Appendix A.

$$Hit\ Rate = \frac{TP}{TP+FN}, \quad \text{(C2)}$$

Frequency bias: A ratio of the number of times that an event is estimated (in this case, positive classification by the estimator) to the number of times that the event actually occurs.

$$Frequency\ Bias = \frac{TP+FP}{TP+FN}, \quad \text{(C3)}$$

False alarm rate: The fraction of all negatively classified observations that were incorrectly classified by the estimator as false alarms.

$$False\ Alarm\ Rate = \frac{FP}{TN+FP}, \quad \text{(C4)}$$

False alarm ratio: The fraction of all estimated positive classifications that were false alarms

$$False\ Alarm\ Ratio = \frac{FP}{FP+TP}, \tag{C5}$$

Heidke Skill Score (HSS): An indicator that provides an indicator of the quality or skill of a forecast (or estimator). A HSS of 0 indicates a forecast with no skill, meaning the forecast is similar to a random selection. A HSS value less than 0 indicates a forecast that performs worse than random selection, while an HSS of 1 is a perfectly accurate forecast. The formulae for HSS are shown in equation C6 and C7

$$Z = \frac{(TP+FP)\times(TP+FN)+(FN+TN)\times(FP+TN)}{TP+FP+FN+TN}, \tag{C6}$$

$$HSS = \frac{TP+TN-Z}{TP+FP+FN+TN-Z}, \tag{C7}$$

APPENDIX D: LOCATION MAP FOR POINTS SELECTED IN COMPARISON OF GROWING DEGREE DAYS AND GROWING DEGREE DEKADS

For section 4.3.2.1, a new parameter, growing degree dekads (GDDekads), was developed to facilitate reduced data storage and computation required for calculation of variable length of crop growing period based on thermal time. The location points shown in Figure D1 were developed to facilitate comparison of growing degree days (GDDs) summed to dekadal totals with GDDekads. The points were selected to cover a large area of southern and central Africa and Madagascar, as well as a variety of elevations and ecologies, based on a visual assessment of a digital elevation

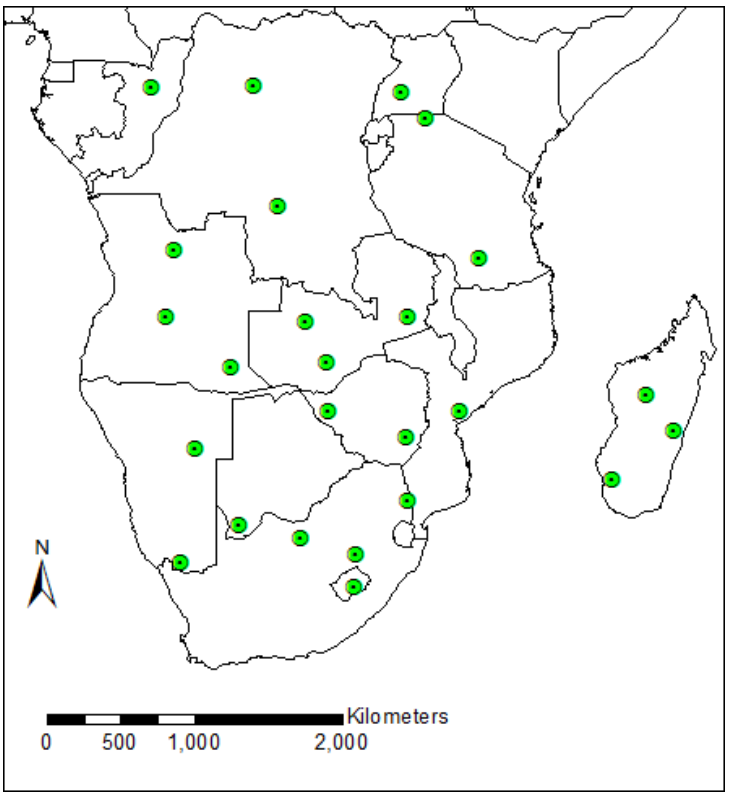


Figure D1. Location map for points selected to facilitate comparison of growing degree days and growing degree dekads