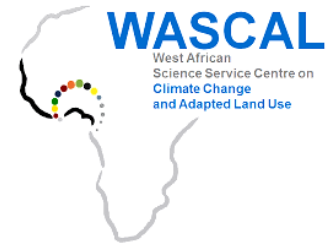**Université de Lomé**

West Africa Science Service Centre on Climate Change and Adapted Land Use

FACULTY OF ART AND HUMANITIES
DEPARTMENT OF GEOGRAPHY

MASTER RESEARCH PROGRAM
CLIMATE CHANGE AND HUMAN SECURITY

Potential of Irrigated Crop Water Productivity Mapping to Improving Food and Water Security around Lake Bam, Northwest Burkina Faso

Thesis N°…………

Thesis submitted in partial fulfilment of the requirement Master Research Dregree

DOMAIN:	HUMANITY AND SOCIAL SCIENCES
MENTION:	GEOGRAPHY
SPECIALTY:	CLIMATE CHANGE AND HUMAN SECURITY

Presented by: **Blou Boris OUATTARA**

Major supervisor: Prof. Saley BACHIR (Université Félix Houphouët Boigny, Côte d'Ivoire)
Co-supervisor: Dr. Gerald FORKUOR (WASCAL Competence Center, Burkina Faso)

Approved on January 30th 2018
Chair of the Committee: Prof. Kodjo TOUNOU (Université de Lomé, Togo)

Committee Members: Dr. Oble NEYA (WASCAL Competence Center, Burkina Faso)
Prof. Kodjovi EDJAME (Université de Lomé, Togo)
Dr. Gerald FORKUOR (WASCAL Competence Center, Burkina Faso)
Dr. Sidat YAFFA (University of the Gambia, The Gambia)

I would like to pay special thanks and appreciation to God, the Almighty, for his abundant grace and mercy showered on me and my family, my stay here and completion of all the work.

ACKNOWLEDGEMENTS

ABSTRACT

Pressure on water resources is increasing worldwide due to constantly growing population. This study aims to develop a spatial model and demonstrate its utility for water productivity mapping (WPM) in a smallholder informal irrigation system around Lake Bam (Burkina Faso) with emphasis on tomato. The study involved three major steps leading to WPM: (1) Sentinel-2 (S-2) data were used for crop productivity mapping involving crop classification with random forest algorithm, crop yield modelling with remote sensing (RS) indices, and yield model extrapolation to a larger area; (2) crop water use estimation by multiplying the irrigated surface area by the actual seasonal evapotranspiration developed through the surface energy balance model 'METRIC' with Landsat8 (L8) data; and (3) WPM produced by dividing raster layers of the two steps above. An overwhelming 89.46% (769.16 ha) of irrigated tomato area for the season 2016–2017 falls in low WP category of 2.5 kg.m^{-3} or less. Only 10.5% of the tomato cultivated area had a WP value of 2.5 kg.m^{-3} or higher. About 82.05% of the tomato area had values lower than 1.63 kg.m^{-3}. The results imply that there is significant scope for increasing WP without having to increase cultivated area or quantity of water utilised. The areas of low WP are spatially pin-pointed and can be used as focus for WP improvements through better land and water management practices.

Keywords: water productivity, evapotranspiration, remote sensing, food and water security, Burkina Faso

RESUME

Les ressources en eau subissent une pression croissante due à l'augmentation de la population. L'objectif de cette étude est de développer un modèle spatial et de démontrer son utilité pour la cartographie de la productivité de l'eau (PE) dans un système informel d'irrigation à petite échelle autour du Lac de Bam (Burkina Faso) ayant pour culture cible la tomate. La présente étude comprend trois étapes majeures conduisant à la production de carte de productivité de l'eau (CPE) : (1) à partir de données de Sentinel-2, l'algorithme de classification supervisée « Random Forest », puis une modélisation du rendement des cultures grâces aux indices de végétation/paramètres biophysiques, enfin une extrapolation à des zones plus grandes a permis d'établir la cartographie de productivité des cultures (kg.m^{-2}) ; (2) l'estimation de l'eau utilisée des cultures (m^3) en multipliant la surface irriguée par l'évapotranspiration saisonnière réelle obtenue à l'aide du modèle de bilan énergétique de surface « METRIC » grâce aux données de Landsat 8 ; (3) la CPE a été calculée en divisant les couches matricielles des deux étapes précédentes. Pour la saison de culture 2016-2017, la majorité des champs de tomates, 769,16 ha soit 89,46 % se situent dans la catégorie de faible PE avec des valeurs de 2,5 kg.m^{-3} ou moins. Seulement 10,5 % ont obtenu une valeur de PE de 2,5 kg.m^{-3} ou plus. Plus de 82,05 % des champs de tomate ont des valeurs de PE inférieures à 1,63 kg.m^{-3}. Ces résultats indiquent clairement qu'il existe des possibilités de stimuler considérablement la production alimentaire actuelle sans nécessairement augmenter la surface cultivée ou la quantité d'eau utilisée. Les zones de faible PE géographiquement localisées peuvent être utilisées comme point d'intérêt pour les améliorations de la PE grâce à de meilleures pratiques de gestion des terres et de l'eau.

Mots clés: productivité de l'eau, évapotranspiration, télédétection, sécurité alimentaire et de l'eau, Burkina Faso

CONTENTS

LIST OF FIGURES

LIST OF PHOTOS

LIST OF TABLES

I. INTRODUCTION

I.1. Background

Water is considered as the most critical resource for sustainable agricultural development in the world. Growing water scarcity is a major factor increasingly constraining food production, causing adverse impacts on the goals of food security and human well-being (Rosegrant, Ringler, and Zhu, 2009). In Burkina Faso, population growth rate averages three percent, which is among the highest in the world. Food insecurity and malnutrition rates are chronically high. The number of people undernourished rose from 3.8 million in 2008-10 to 4.4 million in 2011-13, corresponding to nearly a quarter of the total population (FADPA, 2014). These statistics call for an increase in food production to sustain the increasing population, but with the limited and increasingly diminishing water resources available. Considering the current scarcity of land and water for agriculture, there is an urgent need to pursue practices that achieve 'more crops per drop'. In other words, improving crop water productivity can boost current food production without necessarily increasing cultivated areas or quantity of water utilised. Water productivity (WP) has been defined as the ratio of the net benefits from crops, forestry, fisheries, livestock and mixed agricultural systems to the amount of water used to produce those benefits. In its broadest sense, it reflects the objectives of producing more food, income, livelihoods and ecological benefits at less social and environmental costs per unit of water consumed (Molden *et al*., 2010). Its data allows the ranking and labelling of agricultural fields, according to their water use efficiency (WUE), as an indicator of sustainability (Cai *et al*., 2010).

In order to derive accurate WP information for decision-making, adequate agricultural supervision systems in water use (WU) assessment, WUE evaluation and irrigation monitoring should be designed. Remote sensing (RS) data have, in the past, been useful in assessing WP of irrigation systems (Abdullaev and Molden, 2004; Yan and Wu, 2014). However, until recently, RS instruments have suffered from bad temporal and spatial coverage partly due to excessive cloud cover and other unfavourable environmental conditions. Additionally, the processing of the data required expert knowledge and advanced signal processing. With the launch of improved Earth observation (EO) sensors, the spatial resolution has been improved and the revisit time is much higher (e.g. Sentinel-2). At this time, the operational use of dense time series of multispectral imagery at high spatial resolution makes crop monitoring feasible (Forkuor *et al.* 2014; Dahms *et al*., 2016), capturing crop water use across the growing season, using RS data with suitable temporal and spatial resolutions (Calera *et al*., 2017). Despite these

1

improvements, there is been limited efforts at using of RS data to map and assess water productivity in West Africa (Zwart, 2010).

I.2. Problem Statement

Lake Bam, the largest natural freshwater lake in Burkina Faso, is drying. The depletion of the lake's water resources has resulted in the loss of 1/3 of its depth which has negatively impacted smallholder farmers' production performance and the capacity of the population to safeguard sustainable access to adequate quantities of water to support their livelihoods and well-being.

A possible cause of this problem is declining rainfall, high inter and intra-annual rainfall variability and climate change effects in the Sahel and semi-arid zones (West, Roncoli, and Ouattara, 2008; Salack *et al*., 2015). In addition, poor water management and unsustainable consumption patterns in irrigation systems in West Africa as a result of inadequate information on areas and extent of land irrigated, spatial distribution of crops irrigated, actual production and yield. In fact, in this informal irrigation system around Lake Bam, irrigation authorities do not monitor water extraction from the lake, but rather water is pumped out with motor pumps, which were estimated to be over 1,000 in 2014 (A. Ouedraogo, 2014). Water is extracted over a distance of up to two kilometers from the lake for irrigation. But there is limited information on water usage. The unavailability of these information makes WP studies challenging.

Enhancing WP is thus a critical response to growing water scarcity, including the need to ensure availability of water in rivers and lakes/reservoirs to sustain ecosystems to meet the growing demands of cities and industries (Molden *et al*., 2010). Arising knowledge on WP in informal irrigated systems like the one in Lake Bam might help stakeholders (decision-maker and small farmers) to manage water in an optimal manner in water scarce environments.

I.3. Objectives

The overall objective of this study is to develop a spatial model and demonstrate its utility for water productivity mapping (WPM) by using multi-temporal high resolution RS data from Sentinel-2 and Landsat 8. This is applied in conjunction with field-plot information in a smallholder informal irrigation system in Burkina Faso. The crop of interest here is tomato (*Solanum lycopersicum*).

In order to achieve the stated objective, the following specific objectives were pursued:

- map different irrigated crop types around the lake and model the spatial distribution of crop yield (biomass) of tomatoes;

- estimate the temporal WU for all irrigated fields as well as only for tomatoes; and

- investigate potentials/possibilities for increasing agricultural WP around the Lake.

The relevance of this study lies in exploration of operational capabilities of high to moderate spatial resolution EO data (Landsat and Sentinels) for a better cropping system characterisation and water management sustainability. The case of Lake Bam where water is scarce relative to land needs a way to increase productivity and focus on getting more production per unit of water.

I.4. Research Question and Hypotheses

Crop WP is a key element in longer-term and strategic water resource planning, but its actual and practically feasible values are hardly understood. Most of the WP studies with RS technologies deal with wheat, rice, cotton and maize (Zwart, 2010; Sadras, Grassini, and Steduto, 2012). Little is known about the case of tomatoes, especially in Africa. In the whole world, tomato is the second most valuable vegetable crop next to potato (FAO, 2017). In Burkina Faso, it is the most commonly used cultivated species due to its high economic value (Mertz, Lykke and Reenberg, 2001). So, the use of EO data as a base for establishing monitoring system is innovative in the context of Burkina Faso.

This research aims at investigating the use of satellite-based approaches to map crop WP in an informal irrigation system around Lake Bam in Burkina Faso and to recommend possible improvements in the utilisation of water resources for increased food and water security. The study focuses on estimating physical WP, i.e. the amount of agricultural production (harvestable yield) that can be attained per unit of water used, rather than on the economic WP where the production term is replaced by revenue or profit. The following hypotheses have been formulated:

- availability of high temporal and spatial resolution RS data can enhance the prediction of crop yield prior to harvest season;

- satellite-based evapotranspiration (ET) coupled with crop type information can improve the estimation of irrigation water use in informal irrigation systems in West Africa; and

- the determination of WP for irrigated tomatoes can help in evaluating current production performances and determine future cultivation trends.

I.5. Thesis Structure

The work flow and inter-linkages among the various aspects of the research is displayed in Figure 1.1.

An overview of the theory concerning crop WP, RS based yield (biomass) modelling and ET derivation will be reviewed in chapter II. Chapter III gives a brief description of the Bam watershed, including geography, climate, agriculture, hydrology, demography and water management.

In chapter IV, the methods used to retrieve primary productivity and actual evapotranspiration (ETa) will be given and the use of RS to extrapolate these calculations for WP analysis will be highlighted.

Chapter V presents the results on the crop productivity model, temporal crop water use and crop water productivity mapping.

Finally, chapter VI evaluates and discusses the results.
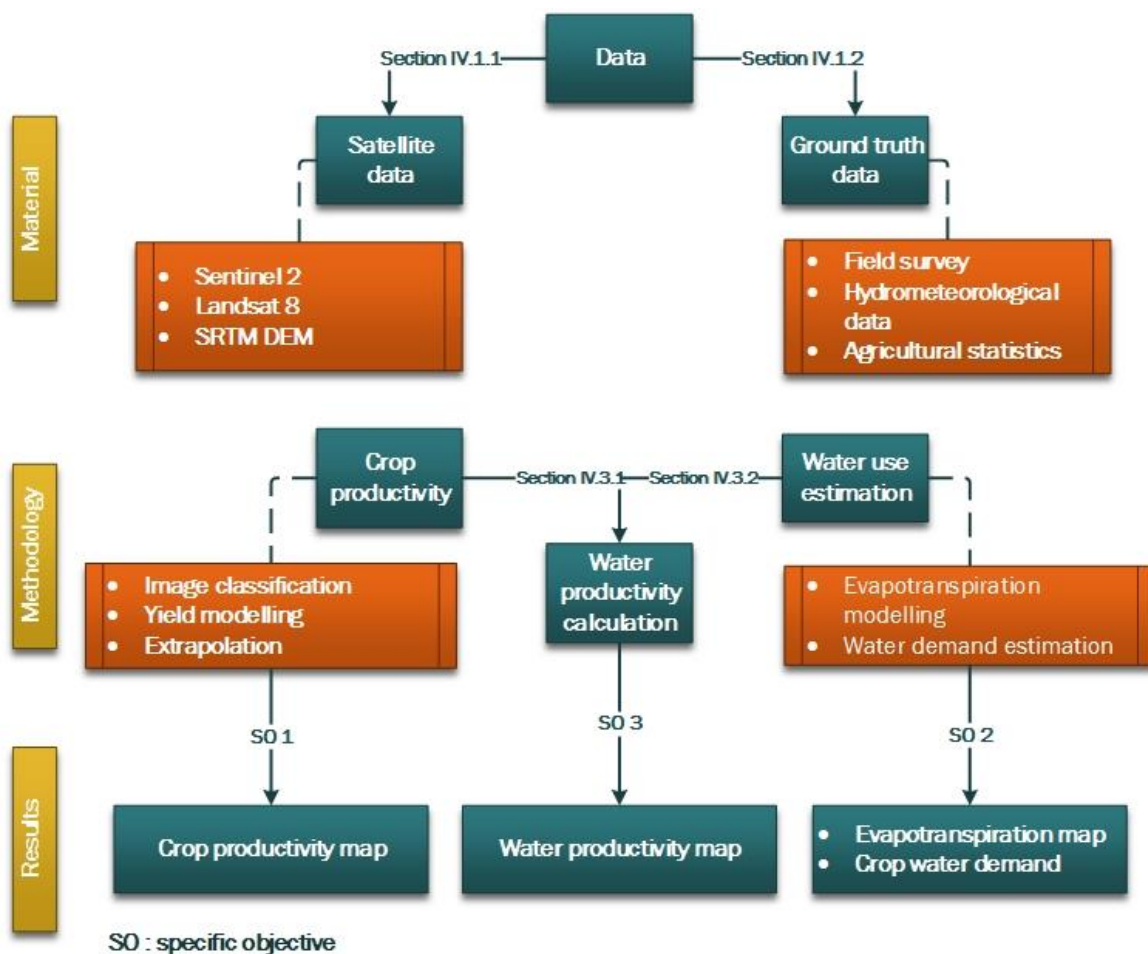
Figure 1.1: Overview of the data and methodological components corresponding to the objectives

II. LITERATURE REVIEW

The rapidly growing world population has intensified pressure on the limited fresh water resources. Irrigated agriculture is the largest water-consuming sector and it faces competing demands from the industrial sector. With an increasing population and less water available for agricultural production, the food security for future generations is at stake. Improving the efficiency with which water is used for agriculture can help reduce potential food insecurity. The use of EO data to model crop WP, especially in data poor regions like West Africa, is of great interest in the research community. In this domain, new methodological approaches are constantly being developed. Thus, to be able to properly apply them, some elementary knowledge is required.

The aim of this chapter is to shortly present the basic concepts of crop WP and thoroughly examine the tasks to be undertaken to model crop productivity and ET. The first sub-section will deal with terminology related to WP and its calculation. Second, yield prediction using RS data will be performed and discussed. Finally, approaches to estimated ET will be reviewed and implemented. The ease or difficulty in deriving ET via RS will be discussed.

II.1. Crop Water Productivity

Improvement in WP (Eq. 2.1) reflects the objectives of producing more food, income, livelihoods and ecological benefits at less social and environmental costs per unit of water consumed (Molden *et al*., 2010).

$$WP = \frac{Agricultural\ benefit}{Water\ use} \tag{2.1}$$

WP is water productivity ($kg.m^{-3}$ or $\$.m^{-3}$), agricultural benefit is crop yield productivity ($kg.m^{-2}$ or $tons.ha^{-1}$) or economic value ($\$.ha^{-1}$), water use is seasonal actual ET (mm, $m^3.m^2$ or $m^3.ha^{-1}$).

Physical quantification of WP is represented in units of kilograms per cubic metre. Another option is economic value of WP expressed in ($\$.m^{-3}$). Basically, WP aims at enabling rapid comparisons between water use systems in space and time. It can be used to set standards and as a baseline to monitor improvement. It can be derived for an individual farm, an irrigation system or even at national or continental scale.

The time period over which WP is estimated is determined by the cycle of agricultural production that drives the system (Cook, Gichuki, and Turral, 2006). Normally, this would include at least one complete crop cycle, extended over a complete year to account for

productive and non-productive water uses. Assessment may be extended over several years to derive estimates of average, minimum or maximum water productivity within each season.

For any WP estimation, both the production system (field-by-field, farm-scale, multiple administrative units) and the area water consumption (plot, field, sub-basin or basin) must be delineated. According to Molden *et al.* (2010), priority areas where substantive increases in water productivity are possible include:

- areas where poverty is high and water productivity is low;
- areas of physical water scarcity where competition for water is high;
- areas with little water resources development where high returns from a little extra water use can make a big difference; and areas of water-driven ecosystem degradation, such as falling groundwater tables, and river desiccation.

II.1.1. Physical and Economic Productivity

The beneficial outcome of agriculture is summarised in Table 2.1. This value can be displayed in a range of forms like yield (kg, Mg, t) or food equivalent (calories). Income ($) or others agreed measure of well-being derived from the goods and services coming from the agricultural system.

Table 2.1: Possible forms of numerator for estimating water productivity (Cook, Gichuki, and Turral, 2006)

Parameter	Numerator
Physical water productivity at field, farm or system level.	Yield (kg) of total biomass, or above ground biomass, or grain, or fodder.
Economic water productivity at farm level.	Gross value of the product, or net value of the product, or net benefit of irrigated production compared with rainfed production.
Economic water productivity at basin scale.	Any of the above valuations including those derived from raising livestock, fish or agroforestry.
Macroeconomic water productivity at regional or national scale.	Monetary value of all direct and indirect economic benefits minus the associated costs, for all uses of water in the domain of interest.

❖ **Physical Productivity**

Approaches to quantify the physical productivity vary from direct measurement of yield or by crop survey to usually secondary statistics as total tonnage for a given administrative area.

Traditionally, *in-situ* measurement cannot capture large-scale variations of above biomass. EO monitoring can help identify factors affecting output (Cai and Thenkabail, 2010). Simulation model provides one empirically derived value for the yield response to water. Yet, this value can vary greatly both temporally and spatially between crops, crop varieties and within single cultivars based on microclimates, soil environments and nutrient availability, thereby limiting accuracy of model outputs (Singh, Rao and Regar, 2010).

Crop yield forecasting at regional, national and continental levels requires reliable and near real time indicators of vegetation status able to cover extensive areas with high temporal frequency. Satellite EO systems have the potential to support operational activities in crops monitoring and yield forecasting as they provide spectral vegetation indices related to plant status or direct estimations of relevant biophysical parameters such as the leaf area index (LAI) or the fraction of absorbed photosynthetically active radiation (fAPAR) with global coverage on a daily basis (Baret *et al*., 2007).

❖ **Economic Productivity**

The simplest measure of economic productivity at a field scale is gross margin (GM) for a single product during a single phase of the crop rotation. The system may require estimates of GM from several seasons to cover all phases of a farming system. For areas that contain different production systems, a composite measure is standardised gross value of the product (SGVP) in which the price of the product is converted to the equivalent price of a standard crop, such as rice, then converted to the world market price. Expressed in a formula:

$$SGVP = \sum_{each\ crop} \left(Area \times Yield \times \left(\frac{local\ price}{base\ price} \right) World\ market\ price \right) \quad (2.2)$$

However, the utility of SGVP may be questioned since it includes no estimate of costs, and therefore attributes average total benefit of all farming inputs to water (Cook *et al*., 2006).

II.1.2. Consumed Water

Planning and monitoring of consumptive water use is necessary for sound management of scarce water resources. Consumptive use influences social, economic, agricultural, and environmental development. Water is consumed through ET which is the sum of evaporation from the land surface plus transpiration from plants. It transfers water from the land surface to the atmosphere in vapour form. Energy is required for ET to take place (for changing liquid water into vapour).

If precipitation over a specific land cover exceeds ET (e.g., forests), such a land cover class is a net producer of water resources. If, however, ET exceeds precipitation, such a land cover class will be a net consumer of water resources. Non-consumed water from precipitation feeds streams, rivers and aquifers. ET information can thus be used for water productivity (Zwart, 2010).

A number of techniques are used to derive ET, ranging from conventional point measurements to modelling and spatially distributed RS estimates.

At individual plant and field scales, lysimeters, heat pulse velocity, Bowen ratio, scintillometry, surface renewal, and eddy correlation systems are commonly used (Meijninger *et al*., 2002; Nagler *et al*., 2005). Although ground values of ET are generally considered accurate, those are point measurements and cannot capture spatial variability. The equipment cost, extensive labour, and coverage issues restrict use of these techniques at large scale.

It may be possible to represent the effect of climate variation on rainfed-crop WP by coupling a weather generator with crop simulation models for large areas (Nieto *et al*., 2012). However, more complex 2- and 3-dimensional modelling may be necessary to understand the consequences of land-use change on water availability and consumptive water use. Where the system is governed by surface water supply with limited groundwater, a simple node-link model like the Stockholm Environment Institute's water evaluation and planning (WEAP) system may be adequate to represent water budgets. If the system is dominated by rainfed agriculture, then soil and water assessment tool (SWAT) model, which integrates land use and hydrology may be preferred (Cook *et al*., 2006). Still, the data requirements may be daunting. A major lesson is that one needs to 'proceed with caution', since propagation of errors within data-hungry models can render complex results meaningless.

RS provides relatively frequent and spatially continuous measurement of biophysical variables used in estimating ET at different spatial scales. This includes radiation, land surface temperatures, vegetation coverage and density, precipitation, soil moisture, weather and climate variables.

RS based mapping of ET is a cost-effective way to estimate and monitor this flux. Definitely, satellite data are ideally suited for deriving spatially continuous fields of ET using energy balance techniques.

II.2. Phenological and Crop Productivity Remote Sensing Monitoring

The estimation of crop yields by space-borne instruments has been very well investigated (Bastiaanssen and Ali, 2003; Fortes *et al*., 2015) and the relationships have improved with the use of modern high spectral and spatial resolution sensors (Enclona *et al*., 2004). For example, the normalised difference vegetation index (NDVI) was found to correlate with net primary production, biomass, vegetation fractions, and yield (Teixeira, 2011).

The solar radiation reaching the surface on the 0.4-0.7μm spectral region is known as the photosynthetically active radiation (PAR). Fraction of absorbed photosynthetically active radiation (fAPAR) refers to the fraction of PAR that is absorbed by a vegetation canopy. fAPAR is difficult to measure directly, but is inferred from models describing the transfer of solar radiation in plant canopies, using RS observations as constraints (Gobron and Verstraete, 2009). Estimated through a radiation balance equation, the systematic record of fAPAR is suitable to reliably monitor the seasonal cycle and inter-annual variability of vegetation photosynthetic activity over terrestrial surfaces. Indeed, investigations showed that fAPAR increases rapidly with the day of year during the vegetative stage, remains relatively stable at the stage of reproduction, and finally decreases slowly during the senescence stage (F. Zhang, Zhou, and Nilsson, 2013).

Dong *et al*. (2015) and Dahms *et al*. (2016), after highlighting the relationship between field measure of fAPAR and vegetation indices (VIs), found that red-edge region is the most accurate spectral index for fAPAR estimates that are independent of crop types during the entire growing season. However, fAPAR retrieved from space-based instruments for that matter is an instantaneous value, i.e. the value at the time of the satellite overpass. So, the cumulative productivity of the vegetation can be assessed by the integral of fAPAR over time; such a number is usually well-correlated with the primary productivity of the environment (biomass production) or the yield of the observed crop (López-Lozano *et al*., 2015).

Moreover, most of the WP studies with satellite data have been done at regional level, using low-resolution images, resulting in a lot of generalisations. Sentinel-2 data contains multispectral information with a high frequency of revisiting for the systematic global coverage of land surfaces. This capability that forms the core of the Sentinel-2 sensors makes the classification of land use/land cover (LULC) more accurate than other moderate resolution sensors (Topaloğlu, Sertel and Musaoğlu, 2016).

II.3. Remote Sensing Modelling of Actual Evapotranspiration

Owing to its high cost-effectiveness, wide and repeatable coverage, and reasonably favourable accuracy, satellite RS retrieval of ET has become a popular tool and study area in the past two to three decades. RS based ET estimation started around the 1980s and has evolved into a variety of approaches and models (K. Zhang, Kimball, and Running, 2016).

II.3.1. Approaches of RS Based Evapotranspiration Estimation

Various algorithms have been developed utilising information from various types of space-borne sensors and often in conjunction with ancillary surface and atmospheric observations for the estimation of these parameters.

II.3.1.1. Surface Energy Balance Methods

The early one-source surface energy balance (SEB) models were usually used and suitable for local and micro-scale applications due to their dependence on local calibration, local reference surface fluxes or other data, lack of spatiotemporal scaleability, and other factors (Bastiaanssen *et al.*, 2002). To overcome these limitations, Bastiaanssen *et al*. (2002) developed the Surface Energy Balance Algorithm for Land (SEBAL) algorithm. The SEBAL model empirically estimates the spatial variation of most essential hydro-meteorological parameters. It requires only field information on shortwave atmospheric transmittance, surface temperature and vegetation height. Based on the SEBAL algorithm, Allen *et al.* (2011) designed the Mapping EvapoTranspiration with high Resolution and Internalised Calibration (METRIC) model that internally calibrates the SEB using ground-based reference ET to reduce computational biases inherent in the RS-based energy balance and to provide congruency with traditional ET estimation methods.

The SEB System (SEBS) developed by Su (2002) is another well-known one-source SEB model. SEBS consists of a set of tools to determinate the land surface's physical parameters from spectral reflectance to radiance measurements. A model for estimating the roughness length for heat transfer and a formulation to determine the evaporation fraction for limiting cases. In contrast to the one-source SEB, two-source SEB models account for the individual contributions of soil and vegetation to the total heat flux, and were designed for use with incomplete canopies (Norman *et al*., 2003).

II.3.1.2. Penman-Monteith and Priestley-Taylor Methods

The Penman-Monteith (PM) ET mapping models essentially compute ET/λE (λE= latent heat accompanying ET) directly and/or estimate sensible heat (H) in conjunction with the

energy balance equation. The PM models include one-, two- or multi-source models depending on how they account for the sources of ET. For two- or multi-source PM models, ET is usually partitioned into soil evaporation and canopy transpiration by partitioning available energy for ET using fractional vegetation cover (fCover) derived either from a satellite observed vegetation index or LAI.

The Priestley-Taylor model (Priestley and Taylor, 1972) is a simplification of the PM equation and was originally used to estimate ET under water unstressed conditions without computing aerodynamic and surface conductance.

II.3.1.3. Other Methods

❖ **T_s-VI Space Methods**

The T_s-VI space methods use spatial variation of T_s and the relationship between surface temperature (T_S) and VIs to partition Rn into λE and H.

❖ **MEP Method**

The MEP method formulates turbulent latent, sensible and ground heat fluxes over the land surface in terms of analytical functions of surface net radiation, temperature, and humidity (J. Wang and Bras, 2009).

❖ **Water Carbon Linkage Methods**

Some efforts have been devoted to building RS driven models that connect photosynthesis or vegetation productivity and ET. Some models adopt modules that have been implemented in land surface models or biosphere models to build process – based models (Long, Longuevergne, and Scanlon, 2014).

❖ **Water Balance Method**

Because E is also a component of the water budget, it can therefore be estimated as a residual of the water balance equation: $ET = P - R - \Delta S_w$, where P is precipitation, R is runoff, and ΔS_W is the change in water storage.

❖ **Empirical Models**

Almost all of the empirical models use statistical methods, ranging from simple single parameters or multivariate methods to complex machine learning approaches, linking ET to various explanatory variables that are more easily observed or obtained from ground and satellite RS (K. Wang, Li, and Cribb, 2006).

II.3.2. Inter-comparison Methods for Remote Sensing Evapotranspiration Estimation

Although there are a large number of RS ET models and approaches, there is no consensus on which one is the best because each method has both advantages and disadvantages relative to the other approaches (APPENDIX A). However, energy balance can detect reduced ET caused by water shortage, salinity or frost as well as increased ET caused by evaporation from bare soil or water intercepted by plant canopies following wetting events. In fact, using energy balance over vegetation-based methods means that actual ET rather than potential ET (based on amount of vegetation) is computed so that reductions in ET caused by disease, salinity or shortage of soil moisture are captured (Allen *et al*., 2011).

The ET of a well-irrigated crop is usually approached through an empirical equation using reference ET (ET_o) and crop coefficients (K_c) (Vanuytrecht *et al*., 2014). However, K_c incorporates and synthesises all the effects on the ET related to morpho-physiological characteristics of the different crops, phenological stage, degree of soil cover, soil and climate conditions, which make them different from the reference crop (Vanino *et al*., 2015). Water depletion is mainly caused by seasonal ETa from a range of crops and natural vegetation types, soil, built-up areas and water bodies.

A proper knowledge of ETa for various land cover is essential for managing scarce water resources and keeping long-term ETa in balance with precipitation. Indeed, the accurate determination of ETa significantly reduces uncertainties in the water balance of a (sub-) basin (Mu *et al.*, 2007), providing water managers with information on water resources being consumed and thus no longer available for downstream users, and water productivity, i.e., the consumption of water in terms of biomass production per unit of water (Molden *et al*., 2007).

The Concept and measurement of crop WP has been defined. It can be expressed either in physical or monetary value. Vegetation indices and biophysical parameters retrieved from satellite products have been investigated and satisfactory correlation with yield has been achieved. Definitely, in numerous RS ET models and approaches identified, there is no consensus on which one is the best considering that each method has pros and cons relative to other methods.

III. DESCRIPTION OF THE STUDY AREA

This part of the thesis describes the study area and is divided into four main sections. Overall, it gives a brief description of the watershed of Lake Bam, including its physical and human characteristic. The first section (III.1) marks location of the study area at the regional and local scales. It provides geophysical characteristics like climate, soils and vegetation of the region. Section III.2 displays the demographic setting of the population followed by the land use aspect in section III.3. Finally, the water resources and its management are discussed (section III.4).

III.1. Geographic and Climatic Settings

❖ **Localisation**

Lake Bam (13 ° 24' North, 1 ° 30' West), located in Burkina Faso to the north of the capital Ouagadougou (110 km), is the largest natural reservoir of water in the country (Figure 3.1). It is in the Bam District with Kongoussi as its capital. With an area of 25–22 km^2 at high water altitudes (Moser *et al*., 2014), this permanent freshwater situated along the Nakambé River (White Volta) is positioned 300 m above mean sea level.
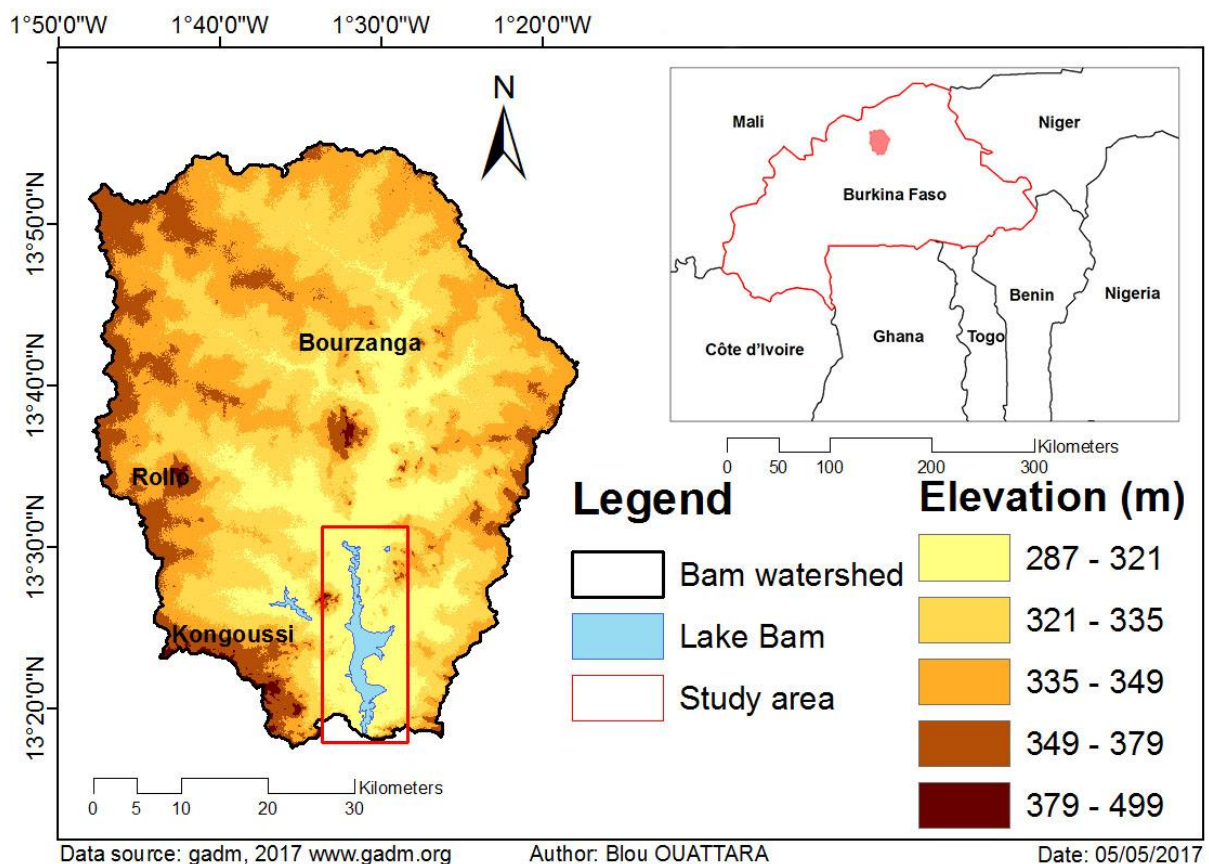
Figure 3.1: Localisation of the study area: Lake Bam (Burkina Faso, West Africa)

❖ **Climate**

The Bam region is considered as a part of the Sudano-Sahelian zone of West Africa because it shares biological and climatological characteristics of both the semi-arid Sahel and the wetter and more densely wooded Sudanian geographical area.

From the closest synoptic meteorological station (Ouahigouya: 13 ° 35' North, 2 ° 26' West) to the lake, the mean annual rainfall from 1970 to 2016 is 639 mm with a maximum summer temperatures of up to 41 °C in 2016 (Figures 3.2 and 3.3).
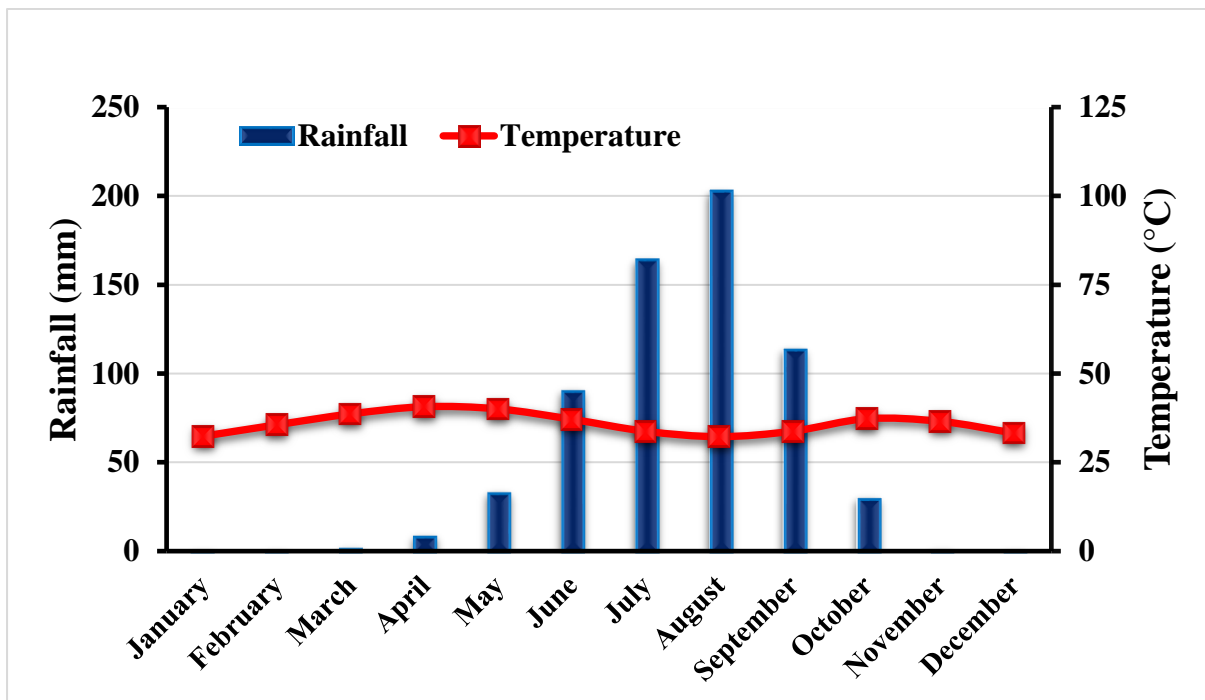
Figure 3.2: Average monthly rainfall and temperature at Ouahigouya (1970 to 2016)
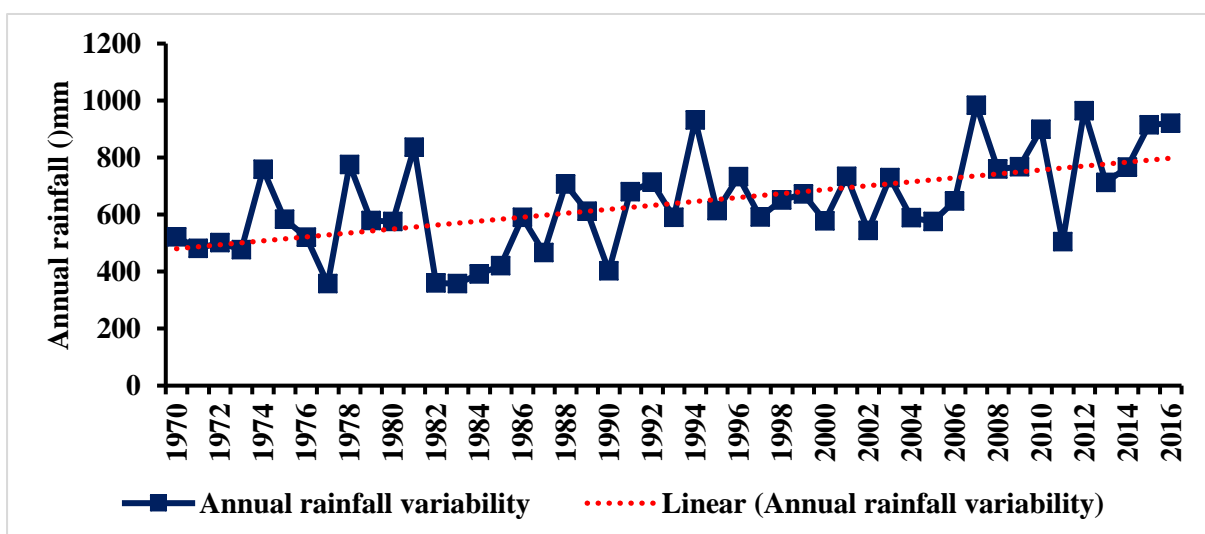
Figure 3.3: Annual rainfall from Ouahigouya

There are two distinct seasons in the study area – rainy and dry seasons. The rainy season spans from May/June to September/October. It is characterised by a high degree of temporal and spatial inter-variability which affects the lake's size. Figure 3.4 shows the inter/intra-annual changes in the lake's surface areas from 2000 to 2017.
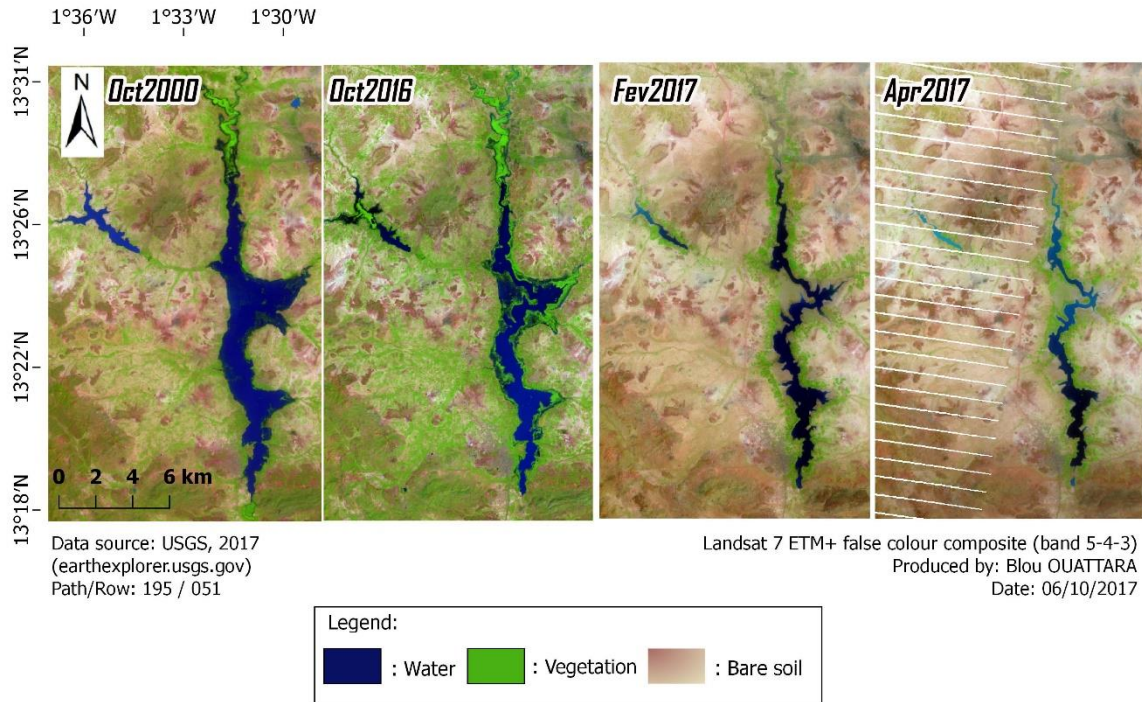
Figure 3.4: Inter-annual rainfall variability affecting the Lake Bam size

❖ **Relief, soil and vegetation**

The watershed, characterised by a low topographical variation, consists of an elevated plain with some scattered ranges of small mountains that reach 400 m (Figure 3.1). The poor laterite soils have little amount of organic matter and nutrients. The soils of the province are predominantly tropical ferruginous and have two variants:

- shallow tropical ferruginous soils and laundry on the glacis and plateaus; and
- deep tropical ferruginous soils in the lowlands.

Vegetation is primarily a mosaic of thorny scrub interspersed with savannah grasslands

III.2. Demographic Setting

In 2014, the population of the Bam District was estimated at 344,628 (Institut National des Statistiques de la Démographie, 2015). 90% of them live in rural area. Majority of the communities are Mossi and Fulani. Agriculture and human settlements are on the banks of the

lake. The inhabitants try as livelihood strategy to diversify their principal activity. So, apart from agriculture and cattle breeding, wood selling and artisanal mining activities are undertaken. However, the trend currently is the migration to the southern part of Burkina Faso or neighbouring countries (e.g. Côte d'Ivoire and Ghana) for arable land and jobs. It is reported that 25%-30% of the population of Bam have migrated (R. Ouedraogo, 2010).

III.3. Land Use

Rangelands, grasslands, forest, rainfed and irrigated agriculture are the main LULC in the Bam watershed. Around the lake, both rainfed and irrigated agriculture are practised. However, irrigated agriculture is applied only during long dry season at its banks from October to April.

Irrigation water is supplied by pumping the lake's water (Photo 3.1) using motorise pumps. In addition, Moser *et al.* (2016) found that irrigated land around the lake has increased during the last decade. Hence, the amplified numbers of those motor pump with over 1,000 in 2014 (A. Ouedraogo, 2014) extracting water along the lake up to two kilometers for irrigation. This has led to increased pollution of the scarce resource (e.g. from pesticide and fertiliser usage) and increased erosion and subsequent sedimentation of the lake.

The main crops cultivated during irrigated agriculture are tomato, spring onion, green beans, green peppers, cabbage, and lettuce. Production of the main irrigated crop for the last seven years is given in Table 3.1.

Table 3.1: Irrigated crop production from 2010 to 2017 in Kongoussi

Crop	Production/Year (in tons)							
	2010	2011	2012	2013	2014	2015	2016	2017
Tomato	16,738.75	15,242.3	13,953.31	15,977.18	17,642.24	17,568.02	19,796.2	19,302.27
Onion	13,261.00	14,145	13,209	14,359	14,504	16,515	17,588	17,000
Green beans	745.45	953.7	1,598.4	1,371.43	1,818.99	1,916.74	2,212.53	2,268.28

Data source : DPAAH/BAM (Direction Provinciale de l'Agriculture et des Aménagements Hydrauliques de Bam)

Source: Author

Photo 3.1: Spiders motor pumps at the shoreline.

III.4. Water resources and Management

The Bam Reservoir has a surface area of about 1200-1000 ha with a storage capacity of 45×10^6 m³ (R. Ouedraogo, 2010). Its catchment is around 2,600 Km² (Figure 3.1). The lake is part of the Nakambe (Volta) river system, which flows through the Central Plateau of Burkina Faso (dry southern-Sahel climate).

R. Ouedraogo (2010) reported that if the current human pressure, water extraction and siltation continues, Burkina Faso's largest natural freshwater lake might turn into a river in about 25 years. Indeed, between 1963 and 2006, siltation with the rate of 500,000 m³/year according to R. Ouedraogo (2010) has shrunk the lake to the 1/3 of its depth.

Some restoration initiatives (e.g., constructing dikes along rivers to reduce sediment transportation or planting trees on the banks of the lake) have been carried out by local fishermen and farmers. However, the inability of authorities to control water withdrawal has led to water insecurity. Thus, the growing competition among the various uses of water is now a major concern in the basin. Water from the basin is utilised for agricultural production, cattle breeding, domestic supplies, fish farming but also serves to sustain the environment. Moreover, the recent discovery of gold in the area, and the proliferation of small-scale miners, will put a further strain on the lake's resources and possibly aggravate the current level of pollution and environmental degradation.

Lake Bam, the biggest natural lake of Burkina Faso constitutes an important source of income for its inhabitants. However, the mismanagement of this precious resource tends to cloud the horizon of future generations. Despite the difficult climatic conditions during the dry season, the efficient management of this resource can ensure food and water security for the population and subsequently improve their welfare.

IV. DATA COLLECTION AND ANALYSIS

This chapter provides details of the data used to develop the methodology for this study. The development of the desired methodology requires different types of data. This includes satellite data for delineating irrigated areas and estimating crop water use (section IV.1.1.). Additionally, ground truth/reference data are required for accurate satellite data interpretation (section IV.2.1). Section IV.1 will thus present all the data that were used in this study. Section IV.2 will detail the various steps that were followed to accurately delineate irrigated fields around Lake Bam and develop a yield model that predicts the spatial distribution of yield (for tomatoes) around the lake. The section further explains how crop water usage was estimated through satellite based evapotranspiration calculations.

IV.1. Data Collection

IV.1.1. Satellite Data and Pre-processing

This study aims at fulfilling the minimum data requirements for crop monitoring in all irrigation zones around Lake Bam. Only the dry season has been considered because it is the season that irrigated agriculture is practised. Satellite data from two sensors were used – Sentinel-2 (S-2, tile number 30PXV) and Landsat 8 (L8, path 195 and row 51). Overview of the sensor configurations and the acquisition dates of the satellite sensors used are summarised in Table 4.1. In order to understand the temporal evolution of irrigated crops, it is important to obtain and analyse time series data. Nowadays, RS sensors improve chances of obtaining time-series data due to shorter intervals between image acquisitions.

IV.1.1.1. Sentinel 2 data

Designed to perform continuous measurements for the next 20+ years, Sentinel sensors have been developed as part of the Copernicus program of the European Union Commission (ESA, 2017). The sensors have global coverage and provide free data in both the optical (Sentinels-2 and 3) and microwave (Sentinel-1) sections of the electromagnetic spectrum (Berger *et al*., 2012). This excellent initiative will greatly improve the scientific community's access to high-resolution satellite images, enhance scientific investigations, open up new application areas, and subsequently improve decision-making and policy formulation.

S-2 carries an innovative wide swath high-resolution multispectral imager (MSI) with 13 spectral bands. The combination of high resolution (up to 10 m), distinctive spectral capabilities (e.g., three bands in the red-edge plus two bands in the SWIR), wide coverage (swath width of 290 km) and at least five-day global revisit time (with twin satellites in orbit)

has improved information retrieval by S-2, whose data have been freely available since November 2015, compared to earlier sensors (Drusch *et al.*, 2012; ESA, 2017). S-2 data has been positively used for multi-temporal crop classification, land surface phenology and biophysical parameters illustrations (F. Zhang *et al.*, 2013; Majasalmi and Rautiainen, 2016; Topaloğlu *et al.*, 2016).

Eighteen S-2 images were downloaded from Oct-01-2016 to May-10-2017. Out of the 13 multi-spectral bands, 10 were used in this study for image analysis (Table 4.1). The reason to leave out band 1 (coastal aerosol) was that it was originally designed for coastal water and aerosol monitoring which does not match the study's purpose. Bands 9 and 10 (water vapour, cirrus) were taking out because they focus on cloud identification, which is not needed at the classification stage when the cloud masking had already been performed.

Dark object subtraction 1 (DOS1) atmospheric correction and bands set creation of all the data were performed using the open source Semi-Automatic Classification Plugin (SCP) in QGIS (Congedo, 2017).

First, the red-edge and short-wave infrared (SWIR) spectral bands of S-2 were resampled to 10 m. Afterwards, using layers stacking, the 10 multispectral band images were obtained for each acquisition date (Table 4.1). S-2 multi-spectral data were used to map the spatial distribution of irrigated crops and other LULC classes.

Table 4.1: Sensor configurations and the acquisition dates of Sentinel-2A

Acquisition date		Spectral bands	Range (µm)	Resolution (m)
October	23rd	Band 2 – Blue	0.458-0.523	10
November	11th, 21st	Band 3 – Green	0.543-0.578	10
		Band 4 – Red	0.650-0.680	10
December	1st, 11th, 21st, 31st	Band 5 – Vegetation Red Edge 1	0.698-0.713	20
January	10th; 30th	Band 6 – Vegetation Red Edge 2	0.733-0.748	20
February	9th; 19th	Band 7 – Vegetation Red Edge 3	0.765-0.785	20
March	1st; 11th; 21st, 31st	Band 8 – NIR	0.785-0.900	10
April	10th; 20th	Band 8A – Vegetation Red Edge 4	0.855-0.875	20
May	10th	Band 11 – SWIR 1	1.565-1.655	20
		Band 12 – SWIR 2	2.10-2.280	20

NIR: Near infrared; SWIR: Short-wave Infrared

IV.1.1.2. Landsat 8 data

Over the past 40 years, the Landsat mission has provided one of the most valuable datasets for mapping and monitoring the Earth's surface (Zhu, Wang, and Woodcock, 2015). Launched in 2013, L8 has increased capabilities such as new spectral bands in the blue part and cirrus cloud-detection portion of the spectrum, two new thermal bands, improved sensor signal to noise performance and several developments in radiometric resolution and duty cycle that allows a significant increase in collection of a number of images per day (Roy *et al*., 2014).

Composed of operational land imager (OLI) and thermal infrared Sensor (TIRS) images consist of nine spectral bands with a spatial resolution of 15 to 30 metres freely available for several applications (agriculture, urban planning, natural resource management, etc.). Thermal bands 10 and 11 are useful in providing more accurate surface temperatures for ET calculation. The thermal bands are acquired at 100-metre resolution, but are resampled to 30 metres in delivered data products (Table 4.2). Thermal information is essential for RS based mapping of ET. Nine L8 images were downloaded for evapotranspiration mapping (ETM) throughout the period of irrigated agriculture.

L8 surface reflectance (SR) product is generated at 30 m spatial resolution on a Universal Transverse Mercator or Polar Stereographic mapping grid. SR product is available through the ESPA Ordering Interface (ESPA, 2017). It supports land surface change studies. Delivered information for L8 bands 1 to 7 is given. This will be used during determination to have information of the surface atmosphere during the process of ET modelling.

Table 4.2: Sensor configurations and the acquisition dates of Landsat 8 data

Acquisition date		Spectral bands	Range (µm)	Resolution (m)
		Band 2 – Blue	0.452-0.512	30
November	20th	Band 3 – Green	0.533-0.590	30
December	06th, 22nd	Band 4 – Red	0.636-0.673	30
January	07th, 23rd	Band 5 – NIR	0.851-0.879	30
February	08th, 24th	Band 6 – SWIR 1	1.566-1.651	30
March	12th	Band 7 – SWIR 2	2.107-2.294	30
April	29th	Band 10 – TIRS 1	10.60-11.19	30
		Band 11 – TIRS 2	11.50-12.51	30

NIR: Near infrared; SWIR: Short-wave Infrared; TIRS: Thermal Infrared Sensor

IV.1.1.3. Digital Elevation Model

A digital elevation model (DEM) is a raster providing elevation data. A 30 m DEM from the Shuttle Radar Topography Mission (SRTM) was downloaded and used to delineate the Bam watershed and estimate topographic features for accurate retrieval of land surface parameters during the process of ET computation. SRTM was a joint project between the National Imagery and Mapping Agency (NIMA) and the National Aeronautics and Space Administration (NASA). It produced digital topographic data for 80% of the Earth's land surface (all land areas between 60º north and 56° south latitude), with data points located every 1-arc-second (approximately 30 metres) on a latitude/longitude grid. The absolute vertical accuracy of the elevation data is 16 metres (at 90% confidence).

IV.1.2. Ground Truth Data
IV.1.2.1. Field Survey

Field campaigns were organised in order to generate reference (or ground truth) data for satellite image interpretation. Training and validation data were collected for image classification and accuracy assessment, respectively.

IV.1.2.1.1. Sampling Method

The survey was purposive, samples are searched for and collected based on a desired need or outcome. A handheld global positioning system (GPS) was used to map representative plots of LULC classes, keeping the minimum mapping area at 30 m × 30 m. Each plot was mapped by walking around it and saving waypoints at its corners. As much as possible, trees were avoided in mapping agricultural fields. The GPS points were later processed into shapefile polygons using a standard GIS application.

IV.1.2.1.2. Reference Data

The reference data were divided into a training dataset for image classification and validation dataset to validate the results of the classification. The training and validation data are an important part of a classification process. The training data is used to train the classifier, i.e. to use areas with a known LULC type to statistically characterise it and subsequently use that information to classify unknown pixels. The validation data is used to validate the results, i.e. compare these areas with a known LULC type to the classification done with the developed methodology.

To reach a successful classification, it is important to have complete and representative training data. This means that all classes should be described with some statistics that separate

them from one another. A total of 842 polygons were mapped for this study. In all, reference data for thirteen LULC classes were collected. Out of the fourteen classes, six LULC classes were selected to constitute the first level of image classification (LULC classification) for the study (Table 4.3). They are:

- bare areas;
- rainfed land: area dedicated for agriculture only during the rainy season;
- irrigated land: area dedicated for agriculture only during the dry season;
- scrublands;
- forest; and
- water bodies.

All irrigated fields were treated as one class at the first level. From the irrigated land class, nine irrigated crops represented by tomato, spring onion, green beans and other minor crops (aubergine, carrot, cabbage, potato, pepper, tobacco) were used to map the spatial distribution of different irrigated crop types (crop classification). All non-tomato crop fields were finally considered as one class for the second level of classification. Spatial distribution and number of training sites used for both data sets and classification algorithms respectively LULC and crop classification is highlighted in Figure 4.1 and Table 4.3.

Table 4.3: Number of training and validation fields used in crop classification

Class	Training	Validation	Total
LULC classification			
Bare soil	44	14	58
Rainfed land	42	24	66
Irrigated land	186	115	301
Shrub land	39	14	53
Forest	27	12	39
Water	15	11	26
Total	353	190	543
Crop classification			
Tomato	73	58	131
Non-tomato crop	106	62	168
Total	179	120	299

Non-tomato crop: onions, green beans, aubergine, carrot, cabbage, pepper, tobacco
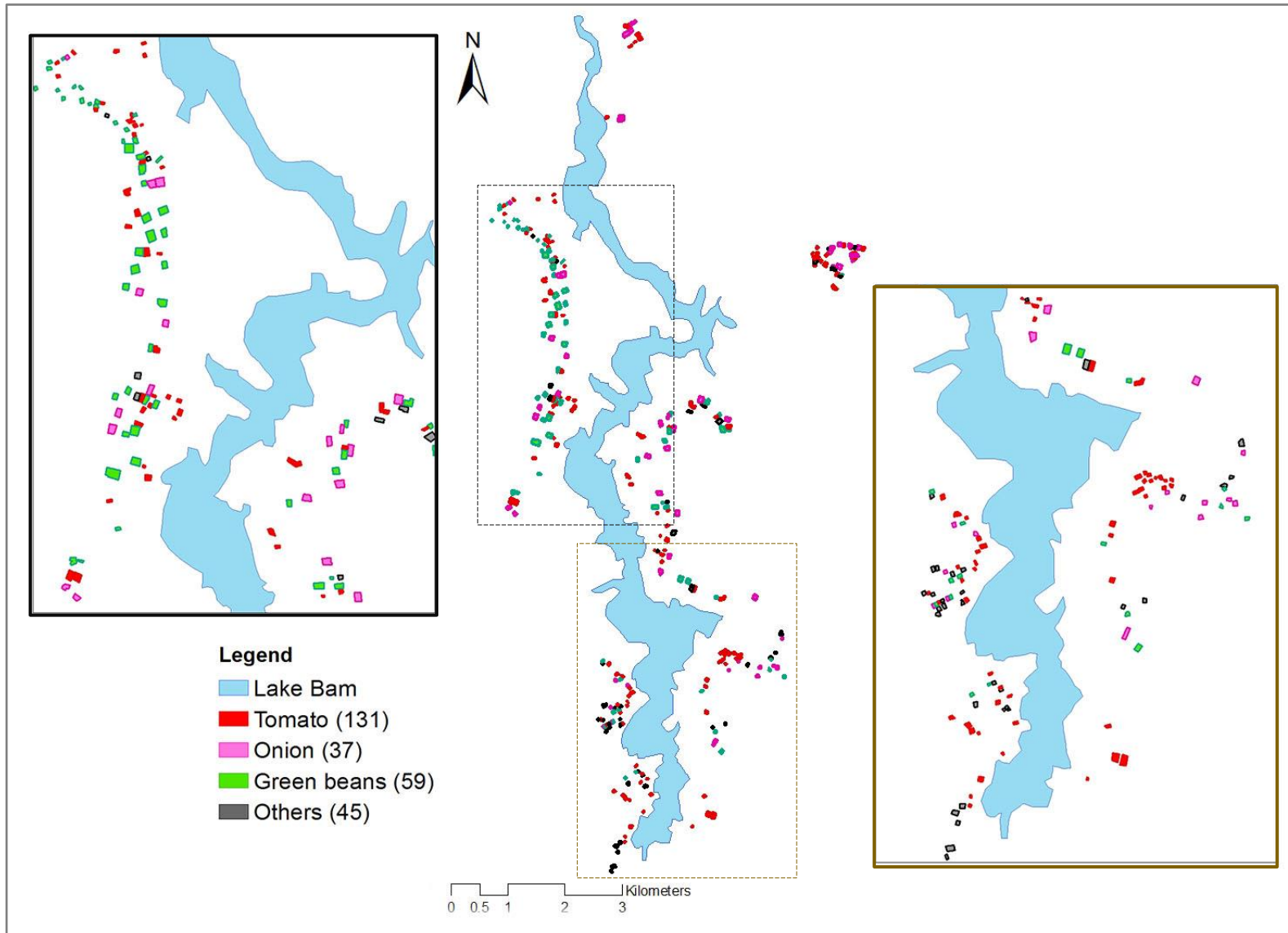
Figure 4.1: Spatial distribution and number of fields (crops only) surveyed

Legend

- Lake Bam
- Tomato (131)
- Onion (37)
- Green beans (59)
- Others (45)

IV.1.2.2. Hydrometeorological Data

Meteorological data for the period January 1980 to December 2016 were obtained from the national meteorological authorities (Table 4.4). The station of Ouahigouya (13 ° 35' North, 2 ° 26' West) was used for weather information because it is the closest synoptic station to the study area. As the meteorological station does not provide hourly radiation information essential for the ET modelling, this has been downloaded from Copernicus Atmosphere Monitoring Service (CAMS, 2017). It provides time series of global, direct, and diffuse irradiation on horizontal surface, and direct irradiation on normal plane for the actual weather conditions as well as for clear-sky conditions. The hydrometeorological data comprised of time series of daily temperature, precipitation, radiation, wind speed and relative humidity (Table 4.4).

Table 4.4: Meteorological data used in this study

Parameter	Units	Frequency	Time period
Temperature	Celsius	hourly	
Relative humidity	%	hourly	
Wind speed	m.s-1	hourly	1980–2017
Precipitation	mm	daily	
Radiation[*]	$W.m^{-2}$	hourly	

Data source : DGM (Direction Générale de la Météorologie)
Radiation data is from CAMS

IV.1.2.3. Agricultural Statistics

Provincial level (2nd administrative level) agricultural statistics for Kongoussi were obtained from the office of Agricultural Provincial Direction of BAM (DPAAH/BAM: Direction Provinciale de l'Agriculture et des Aménagements Hydrauliques de Bam). This data, together with those collected during fieldwork and questionnaire administration, contain information on crop calendar (i.e., planting, growing stages, harvesting) and production constitute agricultural census data.

The data are for the 2016–2017 cropping season. The statistics are derived from surveys that were conducted by agricultural extension officers employed in Kongoussi, whereas the questionnaire administration was based on purposive approach (APPENDIX B).

Tomato, *Solanum lycopersicum*, is a herbaceous annual crop in the family Solanaceae grown for its edible fruit. Most farmers around Lake Bam plough their fields between late October and early November, but not all fields in this area are seeded at the beginning of dry

season (Table 4.5). Field interviews conducted revealed that tomato has three planting dates with an interval of 2 weeks maximum for some fields. The earlier is in late November, the second group of tomatoes have their peak in late January whereas the last set in early May.

Table 4.5: Cropping calendar for tomato in the study based on field survey results.

The start or the harvest period indicated may differ by up to two weeks or more

Crop	2016			2017				
	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May
Tomato 1								
Tomato 2								
Tomato 3								

▬ : land preparation; ▬ : crop development stage; ▬ : harvesting period

Production by farmers is expressed in terms of boxes. It is done through three main boxes called by local farmers Ghana box (big), Ouaga box (moderate) and Togo box (small). The weight of these boxes filled with tomato balances approximately 275 kg, 135.5 kg and 68.75 kg respectively for the box of Ghana, Ouaga and Togo (field survey). The yield is derived by dividing production by the area under cultivation.

IV.2. Methodology

WPM requires two main inputs: (1) estimation of crop productivity and (2) estimation of the quantity of water used to produce the benefit (Eq. 2.1). Therefore, the methodology section comprises two main sub-sections that explain how each input was derived.

Sub-sections IV.2.1 will describe all the steps to accurately delineate and extrapolate the yield model from initially random forest classification. In sub-sections IV.2.2, water used estimation through evapotranspiration modelling to retrieve crop WU will be detailed. The last sub-sections (IV.2.3) will explain how the main inputs from both steps above could lead to deduction of WPM.

IV.2.1. Crop Productivity Mapping

Estimating crop productivity is one of the major components of WPM. In this study, crop productivity is represented by the crop yield, which is the quantity of crops produced per unit area. In order to obtain a spatial distribution of tomato yield in the whole study area, a yield model was developed based on satellite and survey data. Figure 4.2 presents the work flow to model crop productivity. First, satellite images were analysed to map the spatial distribution of

tomato fields in the study area. Then, a yield model was developed using samples yield data collected from the field and satellite data spectral information. Finally, the developed model was used to extrapolate yield for the whole study area.
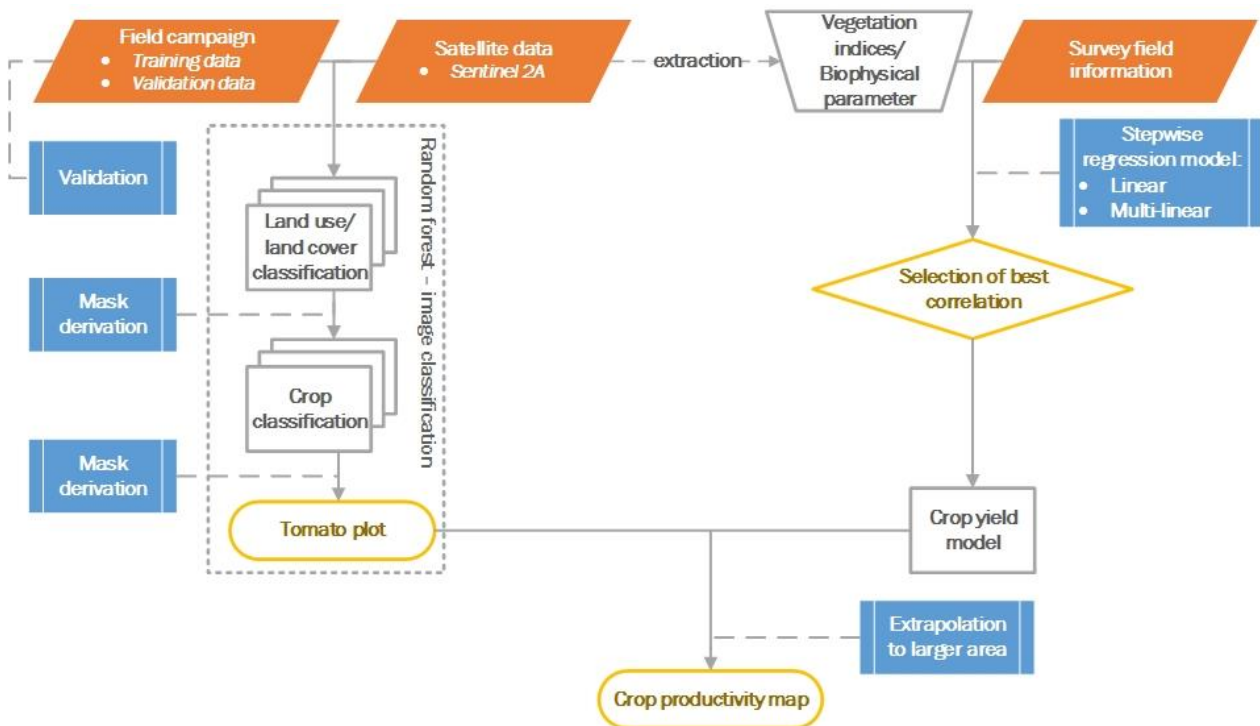
Figure 4.2: Flow chart of yield analysis process

IV.2.1.1. Image Classification

A classification technique (or classifier) is a systematic approach to building classification models from an input data set (Tan, Steinbach, and Kumar, 2006). Examples include decision trees (DTs) classifiers, rule-based classifiers, neural networks, support vector machines, naïve Bayes classifiers. Each technique employs a learning algorithm to identify a model that best fits the relationship between the attribute set and class label of the input data – accurately predict the class labels of previously unknown records.

Evaluation of the performance of a classification model is based on the counts of test records (pixels) correctly and incorrectly predicted by the model. These counts are tabulated in a table known as a confusion matrix.

IV.2.1.1.1. Random Forests

Image classification was performed using random forests (RF) (Breiman, 2001). RF is a machine learning algorithm that builds an ensemble (or *forest*) of DTs from a randomised set of training data. As argued by Friedl and Brodley (1997), DT classifiers have advantages over

traditional supervised classification methods such as maximum likelihood classification. In particular, DTs classifiers are non-parametric and can efficiently deal with large, complicated datasets without requiring assumptions regarding distributions of input data. This makes them ideal candidates for ensemble methods since they usually have low bias and high variance, making them very likely to benefit from the averaging process. Each tree is composed of nodes:

- decision nodes (root node): represents a choice that will result in the subdivision of all records into two or more mutually exclusive subsets;
- chance nodes (internal nodes): represent one of the possible choices available at that point in the tree structure;
- end nodes (leaf nodes): represent the final result of a combination of decisions or events.

DTs's analysis aims to identify the best model for subdividing all records into different segments. When the sample size is large enough, study data can be divided into training and validation datasets. Using the training dataset to build a decision tree model and a validation dataset to decide on the appropriate tree size needed to achieve the optimal final model. The training data consisting of N records whose class labels are known is typically a table of values obtained by overlaying field reference data (e.g. polygons) on a set of satellite images (e.g. spectral bands and indices) and extracting the corresponding values. Columns in the training data represent variables/predictors while rows represent training samples for the various classes/response variables. This set data is subsequently applied to the test set (validation dataset) which consists of records with unknown class labels.

As indicated in the Figure 4.3, the RF algorithm uses a defined N number for training data from M number of variables/predictors in the classifier. In this study, the input is divided into 60% and 40% training and validation data respectively. The number of trees to be constructed (Ntree) is fixed so that m numbers (Mtry) of input variables randomly selected are used to determine the decision at a node of the tree (m should be much less than M). This is to allow prediction of every training sample at least a few times (Liaw and Wiener, 2002).

Then, by choosing each training cases N times with replacement, training set is determined for the tree. The rest of the cases is used to estimate the error of the tree, by predicting their classes. For each node of the tree, m variables are randomly chosen in order to base the decision on that node. Finally, the best split is calculated based on the m variables in the training set. Each tree is fully grown and not pruned.

RF uses the gini index for determining the usefulness of a variable in splitting a node (Breiman, 2001). If a dataset T is split into two subsets T_1 and T_2 with sizes N_1 and N_2 respectively, the gini index of the split data contains examples from n classes, the gini index (T) is defined as:

$$Gini_{split}(T) = \frac{N_1}{N} gini(T_1) + \frac{N_2}{N} gini(T_2) \qquad (4.1)$$

The attribute value that provides the smallest Gini$_{split}$ (T) is chosen to split the node. The final class of each tree is aggregated and voted by weighted values to construct the final classifier. Finally, the decision tree model is used on the test data and prediction is done.
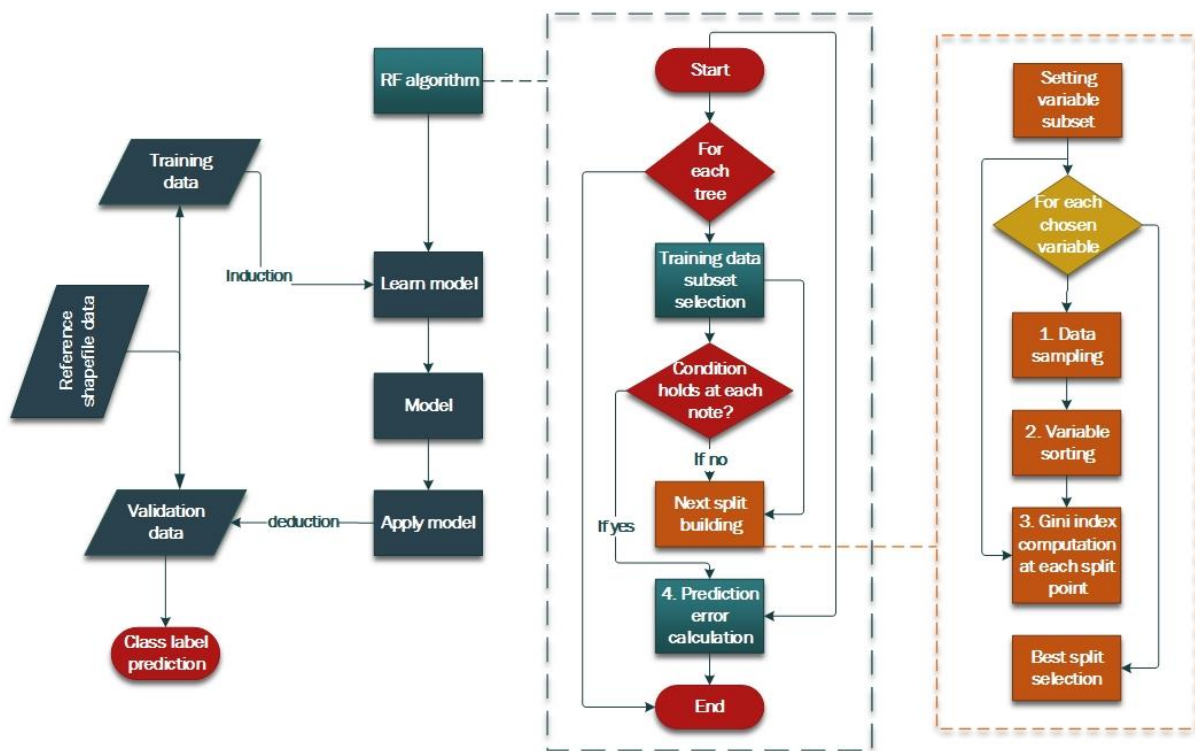
Figure 4.3: Flow chart of the image classification model with random forests

RF was applied to the 10 m resampled S-2 raster layers to create LULC maps of the study area in R software environment. This uses prior knowledge of the classes to classify pixels to the required classes. Two successive supervised classifications were performed to accurately identify different LULC classes as follows:

- image classification 1: in this classification, a general LULC mapping was performed including 6 classes (bare soil, rainfed cultivated area, irrigated areas, shrub lands, forest and water). It must be noted that at this level, all irrigated crops were combined into one class – irrigated areas;

- image classification 2: at this level of classification, only two classes were considered. These are tomatoes and other (non-tomato) crops. All irrigated crops apart from tomatoes (i.e., aubergine, cabbage, potato, carrot, pepper, tobacco) were combined into the non-tomato class. It must be noted that this classification was performed on only the irrigated areas as identified in level 1 above.

IV.2.1.1.2. Validation Method: Classification Accuracy

A map is a model or generalisation of reality and, therefore, usually has some errors. Consequently, it is important to assess the quality of the derived map and express it in a meaningful way. The value of thematic maps constructed from remotely sensed data is clearly a function of classification accuracy (Foody, 2002). Validating LULC products provides critical data quality information to users and producers of these maps.

To determine the accuracy of each classification, thematic accuracy assessment was performed. The referenced data set were used to build a confusion matrix. It basically tells to which class the pixels of known class have been classified by showing the known classes in the columns and the predicted classes on the rows of a matrix (Jin, Stehman, and Mountrakis, 2014).

Although a confusion matrix provides the information needed to determine how well a classification model performs, summarising this information with a single number would make it more convenient to compare the performance of different models. This can be done using a performance metric such as overall accuracy, which is defined as follows:

$$Overall\ Accuracy = \frac{Number\ of\ correct\ predictions}{Total\ number\ of\ predictions} \times 100 \qquad (4.2)$$

Classification algorithms seek models that attain the highest accuracy, this varies from 0 to 100%.

Beside this, the producer's accuracy, the ratio of pixels classified to the right class and the total number of pixels belonging to that class is calculated. It is the probability of a record in a certain class being classified correctly – a measure of omission error. The user's accuracy is the ratio of pixels classified to the right class and the total number of pixels that are classified to that class. This is the probability that a record classified as a certain class is correct – commission error.

Additionally, the F_1 score (Eq. 4.3), which combines producer's and user's accuracy into a composite measure, was computed for each class. This measure enables a better assessment of

class-wise accuracies. The score has a theoretical range between '0' and '1', where '0' represents the worst results, and '1' the best.

$$F_1 score = 2 \times \frac{precision \times recall}{precision + recall}$$
$$= 2 \times \frac{user's\ accuracy \times producer's\ accuracy}{user's\ accuracy + producer's\ accuracy}$$

(4.3)

Finally, as formulated by Cohen (1960), KAPPA, a statistic that removes chance as a factor of probability is computed. It considers the rate of agreement between the actual class and the classifier's decision. Since agreements are on the diagonal in the confusion matrix, agreement is the same as overall accuracy, but this factors in chance. KAPPA is defined as:

$$K = \frac{n_a - n_s}{n - n_s}$$

(4.4)

n is the number of records, n_a is the number of agreements, and n_s is the number of agreements due to chance.

IV.2.1.1.3. Mask Derivation

One obstacle to successful modelling and prediction of crop yields using remotely sensed imagery is the identification of image masks. Image masking involves restricting an analysis to a subset of an image data rather than using the whole. Kastens *et al*. (2005) found that cropland masking, where all sufficiently cropped pixels are included in the mask regardless of crop type, generally leads to improvement in crop yield forecasting ability. This, prior to classification, improves crop classification accuracy (Wardlow and Egbert, 2008). For instance, heterogeneous fields like the ones in West Africa, crop mapping on full-image scene results in considerable confusion between crop/non-crop areas. Most often, farming is done around hamlets and in bushes.

After the image classification 1, only pixels belonging to the irrigated area class were selected for the crop type classification at level 2. This means all the other LULC classes identified at level 1 (i.e. bare areas, rainfed areas, shrubs, forest, water) were masked out from the time-series of S-2 images. Then, classification of tomatoes and non-tomato fields were done on only the irrigated areas.

IV.2.1.2. Spectral Indices and Biophysical Parameters Extraction

IV.2.1.2.1. Vegetation Indices

One of the goals of agricultural production is to achieve maximum crop yield at a minimum cost. Early detection and management of problems associated with crop yield indicators can help increase yield and subsequent profit. VIs are mathematical combinations of mainly red, green and infrared spectral bands. They are designed to find functional relationships between crop characteristics and RS observations (Wiegand *et al.*, 1990).

Spectral analysis was conducted for each image to extract normalised and enhanced vegetation indices. The four common VIs used to estimate biomass and their mathematical formula are listed in Table 4.6 (Cammarano *et al.*, 2014). These are EVI (Enhanced Vegetation Index), EVI2 (Enhanced Vegetation Index 2), NDRE (Normalised Difference Red Edge) and NDVI.

Table 4.6: Vegetation Indices used in this study and their mathematical formula

Vegetation index	Equation	Reference
EVI	$G \dfrac{NIR - Red}{NIR + C_1 \times Red - C_2 \times Blue + L}$	(Huete *et al.*, 2002)
EVI2	$G \dfrac{NIR - Red}{NIR + 2.4 \times Red + 1}$	(Jiang *et al.*, 2008)
NDRE	$\dfrac{NIR - Red\ Edge}{NIR + Red\ Edge}$	(Gitelson and Merzlyak, 1996)
NDVI	$\dfrac{NIR - Red}{NIR + Red}$	(Rouse *et al.*, 1973)

C_1=6, C_2=7.5 (coefficients of the aerosol resistance), G=2.5 (gain factor), L=1 (soil-adjustment factor), NIR: near-infrared band, Red: red band

IV.2.1.2.2. Biophysical Parameters

Biophysical parameters provide an absolute description of the physical characteristics of crops such as biomass and chlorophyll content of plants. Nowadays, RS methods are able to accurately estimate those parameters (Dong *et al.*, 2015; Dahms *et al.*, 2016).

In our study, these parameters were retrieved from the sentinel application platform (SNAP) toolbox. Among them, fAPAR, fCover and LAI were used for this investigation. fAPAR, fCover and LAI are acute variables for understanding vegetation growth rate and predicting crop productivity (Fensholt *et al.*, 2004; Baret *et al.*, 2007; Dahms *et al.*, 2016). It mainly consists of generating a comprehensive database of vegetation characteristics and the associated S-2 top of canopy (TOC) reflectances. The actual algorithm running in SNAP runs

the prediction step of neural network from a set of precomputed coefficients computed during the training phase.

IV.2.1.3. Stepwise Regression Model and Extrapolation

Stepwise regression was utilised to build the crop yield model. This selects only features – VIs or/and biophysical parameters that strongly and significantly explain yield variability. The simulation adds variables to a regression model for the purpose of identifying a useful subset of predictors. Successful application of this technique has been reported in many studies, such as development of indices for global estimation of crop parameters that are sensitive to high biomass levels and less influenced by soil reflectance and atmospheric effects (Jiang *et al*., 2008). The subset selection in regression has been done in R environment with the 'leaps' package. It performs an exhaustive search for the best size of the independent variables for predicting the dependent variable in linear regression, using an efficient branch-and-bound algorithm (Lumley, 2017). One of these advantages is that it takes into consideration the combinations of features. Indeed, since the algorithm returns an optimum model of each set, the results do not depend on a penalty model for model size – it doesn't make any difference whether you want to use akaique, Bayesian, deviance information criertion, etc.

The statistical significance is done by calculating the p-value of an F-statistic for the various features using diverse feature combinations (Draper and Smith, 1998). Terms are added or rejected with the help of a null hypothesis, depending on if the term is already in the model or not. If a term is not currently in the model, the null hypothesis is that adding the term to the model would lead to a zero coefficient for the term. The term is added to the model if the null hypothesis can be rejected. In practice, the terms with p – values less than the defined entrance tolerance are compared and the one with the smallest value is added to the model. This step is repeated as long as the p-values are low enough. If the term is already in the model, the null hypothesis is that it has a zero coefficient.

Computed VIs and biophysical parameters have been used as independent variables. Accounted like described in Table 4.7, a total of 266 independent variables has been tested against yield value (dependent variable) of a dimension n= 52 values.

Table 4.7: Independent variables description

Parameter formulation	Description	Example	Total
P_{date}	value of the independent variable extracted at the specific date by mean	LAI_{apr20}	17×7
sP_{date}	value of the independent variable extracted at the specific date by sum	$sLAI_{apr20}$	17×7
P_s	value of the independent variable integral over time by mean extracted with the function sum	LAI_s	7
P_m	value of the independent variable integral over time by mean extracted with the function mean	LAI_m	7
sP_s	value of the independent variable integral over time by sum extracted with the function sum	$sLAI_s$	7
sP_m	value of the independent variable integral over time by sum extracted with the function mean	$sLAI_m$	7

P: name of the independent variable (NDVI, NDRE, EVI, EVI2, fAPAR, LAI, fCover)

All the regression models were weighted by their coefficient of determination (R^2), adjusted coefficient of determination (R_{adj}^2) and the root mean square error (RMSE) values. R_{adj}^2 varies from 0 to 1, the closest value to 1 will be selected for the model equation. The extrapolation of yield value to the largest area will be done through raster calculation following Eq. 4.5:

$$Yield = R_1.\alpha + R_2.\beta + \cdots + R_n.\omega + C \qquad (4.5)$$

R_1, R_2, R_n are the raster file of the parameter 1, 2 until n (n=number of maximum of subsets); α, β and ω are the coefficient of correlation for the respective parameters 1, 2, n; C is the constant of the equation.

IV.2.2. Water Use Estimation: Actual Evapotranspiration with Surface Energy Balance models

This section explains the processes followed to determine the amount of water used to produce the modelled agricultural benefit (crop productivity). The determination is mainly based on evapotranspiration calculation.

The amount of water required by various crops to grow optimally (or to meet the water loss through ET) is defined as the crop water needs which results mainly from the ETa. A relationship between the crop water needs and the crops surface area gives information about crop water use. The influence of the climate on crop water needs is given by the reference crop evapotranspiration (ET_{ref}) discussed in sub-section IV.2.2.3.

Water used by crops (water demand) will be determined from RS data by calculating ETa based on the following steps discussed in detail in:

- determining ET_{frac} from L8 thermal data;
- calculating the ET_{ref} by applying PM equations;
- computing the ETa by multiplying ET_{frac} with ET_{ref}; and
- retrieving the WU by multiplying ETa with irrigated surface area.

IV.2.2.1. Surface Energy Balance Principle and METRIC Model

Satellite image provides information for the overpass time only. Energy balance based ET algorithms compute an instantaneous ET flux for the image time. The ET flux is calculated for each pixel of the image as a 'residual' of the surface energy budget equation. This requires the computation of net radiation (Rn), soil heat flux (G), and sensible heat flux (H) to solve for latent heat flux (LE) or ET as a residual. The basic truth is that evaporation consumes energy. In doing so, the energy balance includes all major sources (Rn) and consumers (ET, G, and H) (Figure 4.4).

$$LE = Rn - G - H \qquad (4.6)$$

LE is latent heat flux consumed by ETa ($W.m^{-2}$), Rn is net radiation ($W.m^{-2}$), G is soil heat flux ($W.m^{-2}$), and H is the sensible heat flux convected to the air ($W.m^{-2}$).
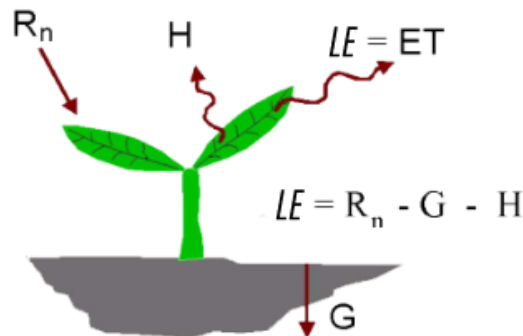
Figure 4.4: Surface energy balance (Bastiaanssen *et al.*, 2002)

Several methods exist incorporating multispectral and thermal bands from satellite imagery to estimate actual crop ET. The one used in this research is the METRIC algorithm as published by Allen *et al.* (2007). It estimates three of the four surface energy balance terms, and calculates latent heat flux (i.e., ET) as a residual.

METRIC applies an internal calibration using extreme pixels of well-irrigated and dry soil agricultural fields to calculate the ET_{inst}. This calibration approach requires less additional atmospheric data for running the model and is therefore widely used for estimating large scale

crop ET in agricultural areas. Moreover, it is crop classification free. For this reason, it is well-suited for this study to estimate the seasonal ETa for a large agricultural irrigated area. According to Allen *et al*. (2007), ETa calculated by METRIC has a very high correlation with ET, measured by lysimeters. This widely used model has been validated in many parts of the world (Rodemaker and Driese, 2008; Allen *et al*., 2011; Bhattarai *et al.*, 2017).

However, it should be noted that RS based ET is not a measurement of ET, but makes approximations from observed radiation at different wavelengths to describe the atmospheric and surface interactions (Allen *et al.,* 2011). These estimations need to be validated with field measurements of ET such as eddy covariance data, which is a direct measurement of ET or reference evapotranspiration (ET_{ref}) from historical meteorological data.

IV.2.2.2. Land Surface Energy Balance

This section explains how the various components of equation 4.6 is calculated. The sub-sections below provide details of each.

IV.2.2.2.1. Surface Radiation Balance Equation: Net Radiation (R_n)

R_n represents the actual radiant energy available at the surface. It is computed by subtracting all outgoing radiant fluxes from all incoming radiant fluxes (Figure 4.5). This is given in the surface radiation balance equation:

$$R_n = R_{s\downarrow} - \alpha R_{s\downarrow} + R_{L\downarrow} - R_{L\uparrow} - (1 - \varepsilon_0)R_{L\downarrow} \tag{4.7}$$

Where: $R_{S\downarrow}$ is the incoming shortwave radiation ($W.m^{-2}$), α is the surface albedo (dimensionless), $R_{L\downarrow}$ is the incoming longwave radiation ($W.m^{-2}$), $R_{L\uparrow}$ is the outgoing longwave radiation ($W.m^{-2}$), and ε_o is the surface thermal emissivity (dimensionless).
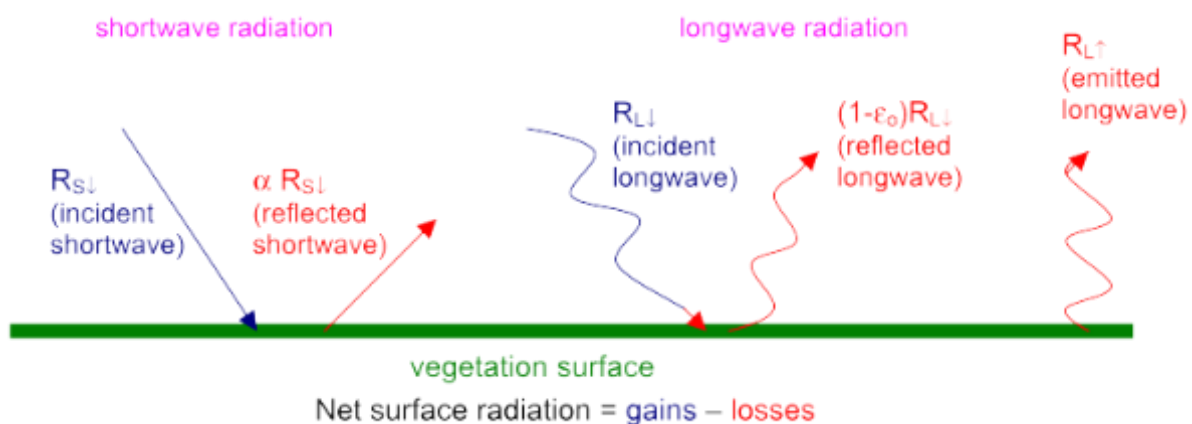
Figure 4.5 : Surface Radiation Balance (Bastiaanssen *et al*., 2002)

❖ **Incoming Solar Radiation**

Incoming solar radiation, direct and diffuse at the Earth's surface, represents the principal energy source for ET. It is computed from a L8 image through a set of steps. Firstly, in order to avoid differences in the surface temperature (and finally ET) caused by different incidence angles and/or elevations, a surface model − a stacked raster with DEM, slope and aspect is calculated. Then, L8 metadata file (containing sun elevation and azimuth) combined with aspect and slope maps estimate solar angles to calculate the incoming solar radiation for each pixel of the scene based on equation Eq. 4.8.

$$R_{S\downarrow} = \frac{G_{sc} cos\theta_{rel}\,\tau_{SW}}{d^2} \qquad \text{(4.8)}$$

Where G_{sc} = solar constant (1367 W.m^{-2}); θ_{rel} = solar incidence angle; d^2 = square of the relative Earth–Sun distance; and τ_{sw} = broad-band atmospheric transmissivity.

❖ **Surface Albedo**

Representing the integrated reflectance across the short-wave spectrum (0.2 to 3.2 µm), surface albedo is the ratio of reflected solar radiation to the incident solar (short-wave) radiation at the surface. In this study, coefficients adjusted by Tasumi, Allen, and Trezza (2008) are used as follow:

$$\boldsymbol{albedo = \rho_{s,B} \times 0.254 + \rho_{s,G} \times 0.149 + \rho_{s,R} \times 0.147 + \rho_{s,NIR} \times 0.311}$$
$$\boldsymbol{+ \rho_{s,SWIR1} \times 0.103 + \rho_{s,SWIR2} \times 0.036} \qquad \text{(4.9)}$$

$\rho_{s,b}$ is the surface reflectance for band b.

❖ **Outgoing and Incoming Long-Wave Radiation**

Outgoing long-wave radiation, $R_{L\uparrow}$, emitted from the surface, is determined by land surface temperature (T_s) and surface emissivity (ε_0). Landsat's thermal band gives information on surface emissivity and brightness temperature. However, T_s requires LAI values to be computed. This value has been retrieved through the method described by (Pôças *et al.*, 2014):

$$\boldsymbol{LAI = 11 \times SAVI^3} \qquad \text{(4.10)}$$

$$\boldsymbol{SAVI = \frac{(1 + L)(NIR - Red)}{NIR + Red + L}}$$

SAVI is the soil adjusted vegetation index, L=0.5.

T_s is computed based on split-window (SW) algorithm developed by Jiménez-Muñoz *et al.* (2014) (Eq. 4.11).

$$T_s = T_{10} + 1.378(T_{10} - T_{11}) + 0.183(T_{10} - T_{11})^2 - 0.268$$
$$+ (54.30 - 2.238\omega)(1 - \varepsilon) + (-129.20 + 16.40\omega)\Delta\varepsilon \tag{4.11}$$

T_s is the land surface temperature (K), T_{10} and T_{11} are the at-sensor brightness temperatures for bands 10 and 11 of Landsat 8 (K), ε is the mean emissivity, ω is the total atmospheric water vapour content (in g.cm^{-2}) and $\Delta\varepsilon$ is the emissivity difference.

Using Stefan-Boltzmann (Eq. 4.12) in complement with empirical equation of Bastiaanssen (1995) with coefficients developed by Allen (2000) *in* Olmedo *et al.* (2017), long wave incoming radiation is computed. Outgoing long-wave radiation has been computed through Stefan-Boltzmann equation (Eq. 4.13).

$$R_{L\downarrow} = \varepsilon_a \times \sigma \times T_{s\_cold}^4 \tag{4.12}$$

$$R_{L\uparrow} = \varepsilon_0 \times \sigma \times T_s^4 \tag{4.13}$$

$$\varepsilon_a = 0.85 \times (-ln\tau_{SW})^{0.09} \tag{4.14}$$

$$\varepsilon_0 = 0.95 + 0.01 \times LAI \text{ for LAI} \leq 3 \text{ and } \varepsilon_0 = 0.98 \text{ when LAI} > 3 \tag{4.15}$$

ε_a is effective atmospheric emissivity (-); σ is the Stefan-Boltzmann (W.m^{-2}.K^{-4}) constant; T_{s_cold} is the surface temperature at a reference point (cold pixel); ε_0 is broad-band surface emissivity; τ_{sw} is one-way shortwave transmittance for the atmosphere.

Considering the surface radiation balance, the net radiation is thus estimated. It uses information from the image, in addition to measurements of actual vapour pressure and altitude.

IV.2.2.2.2. Soil Heat Flux (G)

Soil heat flux is the rate of heat storage into the soil and vegetation due to conduction. The following Eq. 4.16 based on Chávez *et al.*, (2005) findings has been used:

$$G = R_n \times (0.3324 - 0.024 \times LAI) \times (0.8155 - 0.3032 \times ln(LAI)) \tag{4.16}$$

IV.2.2.2.3. Sensible Heat Flux (H)

Sensible heat flux is the rate of heat loss to the air by convection and conduction, due to a temperature difference. METRIC differs from previous applications of SEBAL principally in how the 'H function' is calibrated for each specific satellite image. In both METRIC and SEBAL, H is estimated from an aerodynamic function as follows:

$$H = \rho_{air} \times C_p \times \frac{dT}{r_{ah,1.2}} \qquad (4.17)$$

ρ_{air} is air density ($kg.m^{-3}$), C_p is specific heat of air at constant pressure ($J.kg^{-1}.K^{-1}$) and $r_{ah,1,2}$ is aerodynamic resistance (s/m) between two near-surface heights, z_1 and z_2 (generally 0.1 and 2 m above the zero-plane displacement height) computed as a function of estimated aerodynamic roughness of the particular pixel and an iterative stability correction scheme for atmospheric heat transfer based on the Monin-Obhukov stability length scale (Allen, 1996).

The temperature gradient dT is designed to 'float' above the surface, beyond the height for sensible heat roughness (z_{oh}) and zero plane displacement, and can be approximated as a relatively simple linear function of T_s, as pioneered by Bastiaanssen (1995).

$$dT = a + bT_{s\,datum} \qquad (4.18)$$

'a' and 'b' are empirically determined constants for a given satellite image, $T_{s\,datum}$ is surface temperature adjusted to a common elevation data for each image pixel using a DEM and customized lapse rate. $T_{s\,datum}$ corrects for cooling impacts on T_s due to increasing elevation within an image that is not related to dT and H.

Determination of 'a' and 'b' in Eq. 4.18 involves locating a hot (dry) pixel in a fallow agricultural field with large T_s and a cold (wet) pixel with a small T_s (irrigated field) in the RS image. Extreme wet and dry pixels has been automatically searched within the satellite scene based on the method described by Owusu (2017). The cold condition is typically a well-irrigated alfalfa field where ET= ET_{ref}. The hot condition is typically a dry, bare agricultural field where ET=0. Then, the energy of Eq. 4.6 can be solved for H_{cold} and H_{hot}, as respectively following Eq. 4.19 and 4.20.

$$H_{cold} = (R_n - G)_{cold} - LE_{cold} \qquad (4.19)$$

$$H_{hot} = (R_n - G)_{hot} - LE_{hot} \qquad (4.20)$$

IV.2.2.3. ASCE Standardized Penman-Monteith Equation

ET_{ref} is defined as the rate at which readily available soil water is vaporized from specified vegetated surfaces (Jensen, Burman, and Allen, 1990). Its purpose is to separate the influence of the weather conditions on the ET. The ET_{ref} provides a standard to which ET at different periods of the year or in other regions can be compared and ET of other crops can be related.

The relationship between the ET_{ref} and the actually grown crop is given by the crop factor, Kc, as shown in the following formula:

$$ET_{ref} \times K_c = ETa \hspace{3cm} \text{(4.21)}$$

Where ET_{ref} is the reference crop evapotranspiration (mm/day), K_c is the crop factor, ETa is the actual crop evapotranspiration (mm/day).

American Society of Civil Engineers (ASCE, 2000) establishes a uniform ET estimates, transferable crop coefficients and two standardized reference evapotranspiration surfaces ET (ET_{sz}). Based on the FAO 56 PM equation, two equations with appropriate constants for daily and hourly time periods were developed:

- ET_{os}: for a short crop with an approximate height of 0.12 m (similar to clipped, cool-season grass); and
- ET_{rs}: for a tall crop with an approximate height of 0.50 m (similar to full-cover alfalfa).

The ASCE standardized the PM method for reference ET calculation by Eq. 4.22

$$ET_{sz} = \frac{0.48 \times \Delta \times (R_n - G) + \gamma \times \dfrac{C_n}{T + 273} \times u_2 \times (e_s - e_a)}{\Delta + \gamma \times (1 + C_d \times u_2)} \hspace{1.5cm} \text{(4.22)}$$

Where ET_{sz} is the standardized reference evapotranspiration for grass (ET_{os}) or alfalfa (ET_{rs}) in units based on the time step of $mm.d^{-1}$ for a 24-h day or $mm.h^{-1}$ for an hourly time step, Rn is the net radiation at the crop surface ($MJ.m^{-2}.day^{-1}$), G is the soil heat flux ($MJ.m^{-2}.day^{-1}$), T is the mean daily air temperature at 2 m height (°C), u_2 is the wind speed at 2 m height ($m.s^{-1}$), e_s is the saturation vapour pressure (kPa), e_a is the actual vapour pressure (kPa), ($e_s - e_a$) is the saturation vapour pressure deficit (kPa), Δ is the slope vapour pressure curve ($kPa.°C^{-1}$), γ is the psychrometric constant ($kPa.C^{-1}$), Cn is the numerator constant and Cd is the denominator constant that changes with reference type and calculation time step.

Note that ET_{os} equation is identical to the FAO-56 Penman-Monteith. For this study, ET_{os} has been selected because the major irrigated crops (tomato, onions and green beans) have their average height lower than 0.5 m around Lake Bam (field survey). For the forthcoming sections, hourly ET_{os} will be represented by ET_o whereas daily ET_{os} by ET_r.

Except in narrow mountain valleys, ET_{ref} is relatively stable and often representative for areas of 50–100 km distance with relatively small changes (Allen, Tasumi, and Trezza, 2007). Thus, if the weather station is in the center of a Landsat image, then a single station may be appropriate to use in computing ET_o and ET_r.

IV.2.2.4. Crop Water Use

RS based estimation of ET allows the computation of K_c equivalent called reference evapotranspiration fraction (ET_{frac}) where knowledge of the type of crop and the growth stage are not a requirement. ET_{frac} value is given by the ratio between the instantaneous actual evapotranspiration (ET_{inst}) at the satellite overpass and the hourly reference evapotranspiration at this time (ET_o) (Eq. 4.23).

$$ET_{frac} = \frac{ET_{inst}}{ET_o} \tag{4.23}$$

$$\text{With} \quad ET_{inst} = 3600 \times \frac{LE}{\lambda \rho_w} \tag{4.24}$$

ET_{inst} is the instantaneous ETa at the satellite overpass (mm.h^{-1}), 3600 is the conversion factor from seconds to hours, ρ_w is the density of water = 1000 kg.m^{-3}, λ is the water latent heat of vapourisation = 2.26 10^6 (J.kg^{-1}) and ET_o is the hourly ET_{ref} at the time of satellite overpass (mm.h^{-1}).

The daily actual evapotranspiration pixel by pixel (ET_{24}) is finally computed by using ET_{frac} and daily reference evapotranspiration (ET_r) as described in Eq. 4.25.

$$ET_{24} = ET_{frac} \times ET_r \tag{4.25}$$

Then, by assuming that ET_{24} is constant during the month for the given crop, ET for a specific month ($ET_{month\,i}$) is determined. This leads to accumulation of water evapotranspired for the pixel over a span of 1 month by multiplying its ET_{24} by the number of day in the month.

$$ET_{month} = ET_{24} \times d \tag{4.26}$$

d is the number of day for the specific month.

The seasonal crop water use ($WU_{seasonal}$), is developed by the integral of ET_{month} of the specific month multiplied by irrigated surface as follows:

$$WU_{seasonal} = \sum_{i=n}^{m} ET_{month_i} \times S \tag{4.27}$$

$WU_{seasonal}$ is the cumulative WU for a period beginning from month 'm' and ending at month 'n' (in the case of this study from November to April) (m^3), $ET_{month\,i}$ is the actual evapotranspiration (m^3.ha^{-1}) for one specific month of the farming season, S is the surface area (m^2) of irrigated tomatoes.

IV.2.3. Crop Water Productivity Calculation

WP is a concept to express the value or benefit derived from the use of water and includes essential aspects of water management such as production for arid and semi-arid regions. Increasing WP means either to produce the same yield with less water resources or to obtain higher crop yields with the same water resources (Zwart, 2010).

In this study, the computation of WP will be done through a simple division of the yield and ETa output raster layers as indicated in Figure 4.6. The calculation of WP will be based on the Eq. 2.1.

Crop productivity   Water use Estimation   Water productivity
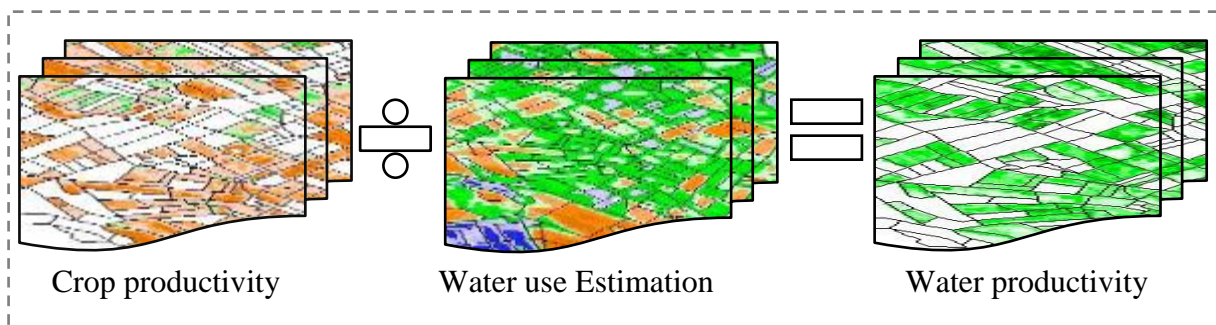
Figure 4.6: Flow chart of water productivity mapping

V. RESULTS

This chapter presents results of the steps conducted to derive WPM in this study. Divided into three main parts, the first two sections highlight the result of the main input described in the methodology section. The sub-section V.1 provides results of the crop productivity modelling. The accuracy obtained in the LULC and crop classifications are first given and briefly discussed. Then, results of the yield modelling to determine the spatial distribution of tomato yield in the study area is presented. Next, the sub-section V.2 presents findings on water used by crops. Initially, calculated ET_{frac} from L8 thermal data and ET_{ref} for each month of the growing season are displayed and briefly discussed. Afterwards, seasonal ETa and $WU_{seasonal}$ is shown.

V.1. Crop Productivity Map

V.1.1. Spatial Distribution of Crops

V.1.1.1. Irrigated Area Mapping

Table 5.1 presents the confusion matrix of the LULC classification (level 1) to determine the extent of irrigated and other general LULC classes. An overall accuracy of 92.8% was achieved. The kappa value (0.92) indicates that fewer than 1% of pixels classified is due to chance. The producer's accuracy shows that 95% of the record in irrigated land class is classified correctly with a probability of 0.86 that this classified record is correct. This is confirmed by its F_1 score of 0.9. It was found that images acquired during the early period of the dry season (October-December) produce the best results.

Table 5.1 shows that despite the good accuracy obtained for irrigated area class, there exists some confusion between this class and both rainfed areas and forest classes. This can be explained by the fact that at the onset of the dry season, forest is green, while rainfed cultivated areas and irrigated areas may be bare (harvested and ploughed). However, at the peak of the dry season, forest and irrigated areas can all be green, thereby causing spectral confusion whereas rainfed areas have minimal confusion with these two classes.

Table 5.1: Accuracy estimates for the derived irrigated crop mask

Class	Bare soil	Rainfed area	Irrigated area	Shrub land	Forest	Water	Row Total	Producer's Accuracy	User's Accuracy	F₁
Bare soil	995	5	0	0	0	0	1000	1.00	0.98	0.99
Rainfed area	21	933	30	2	14	0	1000	0.93	0.95	0.94
Irrigated area	0	26	**946**	8	20	0	1000	0.95	**0.86**	**0.90**
Shrub land	2	13	9	847	129	0	1000	0.85	0.95	0.90
Forest	0	7	110	33	850	0	1000	0.85	0.84	0.84
Water	0	0	1	0	0	999	1000	1.0	1.0	1.0
Column Total	1018	984	1096	890	1013	999	6000			

Overall Accuracy = 92.83%; Kappa = 0.92

V.1.1.2. Tomato Plot Mapping

Table 5.2 shows the accuracy estimates obtained in the crop classification which aimed at mapping tomato cultivated areas (tomato mask) against other irrigated crops. An overall classification accuracy of 74.3% and kappa coefficient of 59% were achieved. Pixels classified as tomatoes have 70% likelihood of being correctly classified. User's accuracy value of the tomato class indicates that 76% of the samples could be correctly classified. Statistics from producer's and user's accuracy allowed the determination of an F_1 score for tomato crop above 70%.

In general, there was substantial confusion between tomatoes and other irrigated crops (non-tomatoes). Reasons for this confusion are discussed in section VI.1.

Table 5.2: Accuracy estimates for the tomato crop mask

Class	Tomato	Non-Tomato crop	Row Total	Producer's Accuracy	User's Accuracy	F₁
Tomato crop	**678**	297	975	**0.70**	**0.76**	**0.73**
Non Tomato crop	210	790	1000	0.79	0.73	0.76
Column Total	888	1087	1975			

Overall Accuracy = 74.33%; Kappa = 0.59

Figure 5.1 provides a detailed picture of the classification results obtained for the two classifications. First, irrigated areas were mapped by classifying six general LULC classes, including irrigated areas. Then, based on the mapped irrigated areas, two crop classes – tomato areas and non-tomato areas – were mapped.
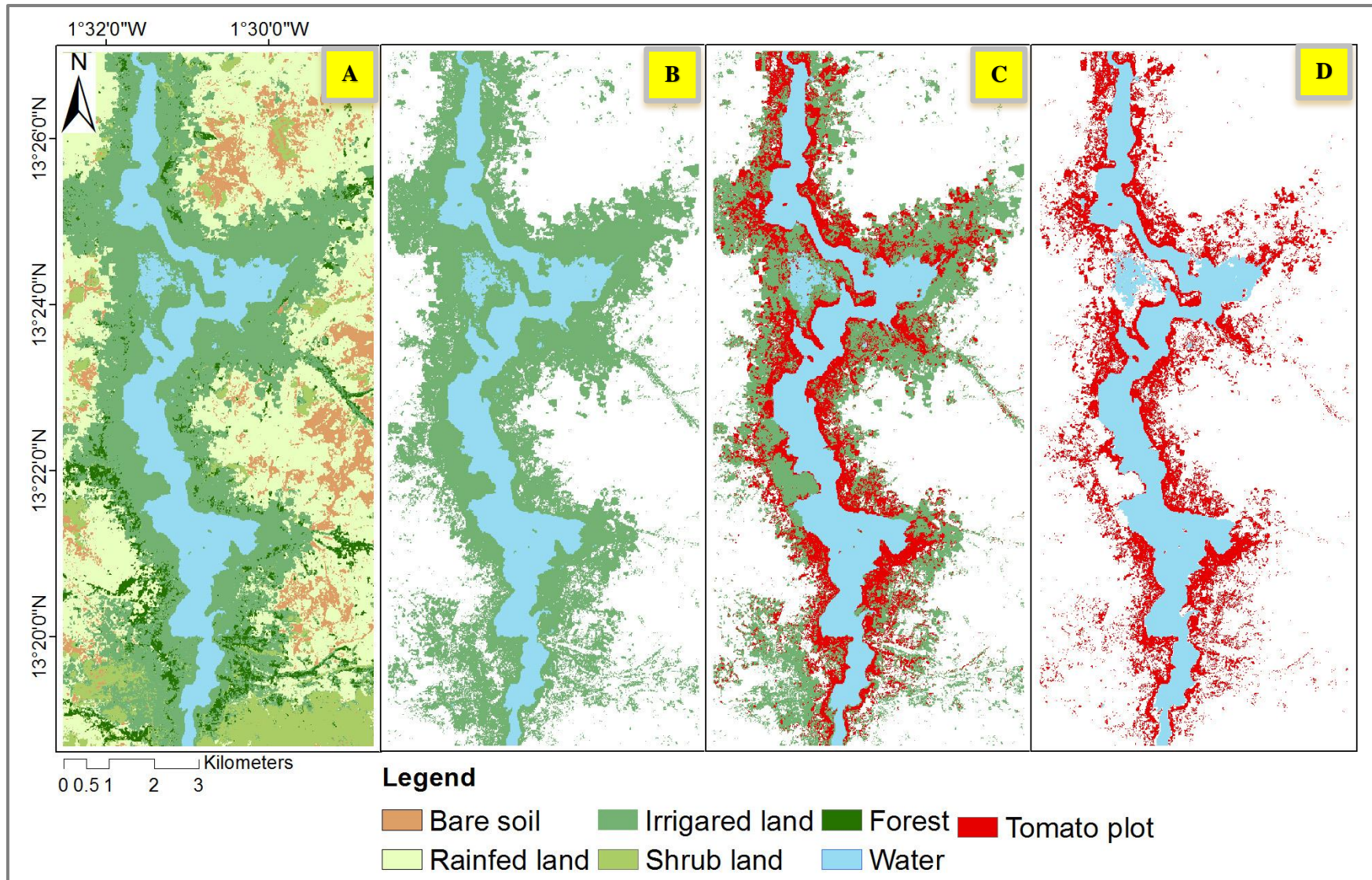
Figure 5.1: Mask derivation A: LULC map; B: Irrigated land mask; C: Tomato crop under irrigated land map; D: Extracted tomato plot

V.1.2. Yield Modelling

The tomato yield (kg.m^{-2}) values (52 plot values in total), obtained from the field survey, were correlated with VIs and biophysical parameters derived from S-2 data. The yield prediction was conducted through a stepwise regression model to select the best predictor subset that significantly explains yield variability at field level. The model selected five variables/predictors, explaining 54% (R^2_{adj}) of the variability (Table 5.3). These five predictors provided better results than any other combination.

This regression analysis revealed that for a simple linear model, only the sum of the temporal integral of fAPAR (fAPAR$_s$) value explain more than 20% of yield variability. The best fit subset model is composed of:

- LAI$_{apr20}$: LAI value of April 20 extracted by the function mean;
- LAI$_m$: mean of LAI integral over time extracted by the function mean;
- sEVI$_{may10}$: EVI value of May 10 extracted by the function sum;
- sLAI$_{apr20}$: LAI value of April 20 extracted by the function sum;
- EVI2$_{apr20}$: EVI2 value of April 20 extracted by the function mean.

Table 5.3: Best subset and image combination for tomato yield variability

Subset size	Best model	Image combination	R	R^2	R$^2_{adj}$	RMSE
1	Linear	fAPAR$_s$	0.48	0.23	0.21	4.23
2	multi-linear	LAI$_{may10}$ sLAI$_{apr20}$	0.58	0.34	0.31	3.95
3	multi-linear	LAI$_{dec11}$ sEVI$_{dec11}$ Sndvi$_{All\_s}$	0.66	0.44	0.40	3.69
4	multi-linear	LAI$_{dec11}$ LAI$_s$ sEVI$_{dec11}$ sfAPAR$_m$	0.71	0.50	0.46	3.50
5	multi-linear	LAI$_{apr20}$ LAI$_m$ sEVI$_{may10}$ sLAI$_{apr20}$ EVI2$_{apr20}$	**0.76**	**0.58**	**0.54**	**3.24**

V.1.3. Spatial Extrapolation and Crop Productivity Map

The most significant model (Table 5.4) was used on the relevant images to extrapolate and derive tomato yield map for irrigated fields around Lake Bam. The regression relationship of the best model (R^2_{adj} =0.54) is given by:

$$Yield = LAI_{apr20} \times 68.94 + LAI_m \times 33.71 - sEVI_{may10} \times 33951.20 \qquad (5.1)$$
$$- sLAI_{apr20} \times 1.23 + EVI2_{apr20} \times 1.31 - 34.14$$

The fit is statistically significant since the p-value is 8.22×10^{-08} (<0.05).

Table 5.4: Best regression analysis model

	Estimate	Standard Error	Standardised	t-value	p-value
Intercept	-34.14	7.27		-4.70	<.001
LAI$_{apr20}$	68.94	10.64	1.31	6.48	<.001
LAI$_m$	33.71	8.60	0.42	3.92	<.001
sEVI$_{may10}$	-33,951.20	7515.06	-0.95	-4.52	<.001
sLAI$_{apr20}$	-1.23	0.18	-1.87	-7.02	<.001
EVI2$_{apr20}$	1.31	0.25	0.80	5.25	<.001

F-statistic: 12.79 on 5 and 46 DF, p-value: 8.22×10^{-08}

The result of extrapolation shows spatial variability within tomato yield (kg.m^{-2}) using S-2 data (Figure 5.2). This extrapolation was possible as a result of the good understanding we had through model development (Table 5.4) and delineation of tomato plot. This CPM is one of the key parameters in understanding within and between field variability.

The yield value around the lake mostly varies from 0.01 to 16 kg.m^{-2} and these are parcels close to the bank of the lake. Descriptive statistics show that 0.01, 100 and 10.35 are respectively minimum, maximum and mean yield value (kg.m^{-2}). Higher values are located in the northern part of the lake (Figure 5.2). Extracted yield from twenty-two tomatoes plot statistics analysis shows that maximum frequencies of yield ranged from 5–20 kg.m^{-2}.
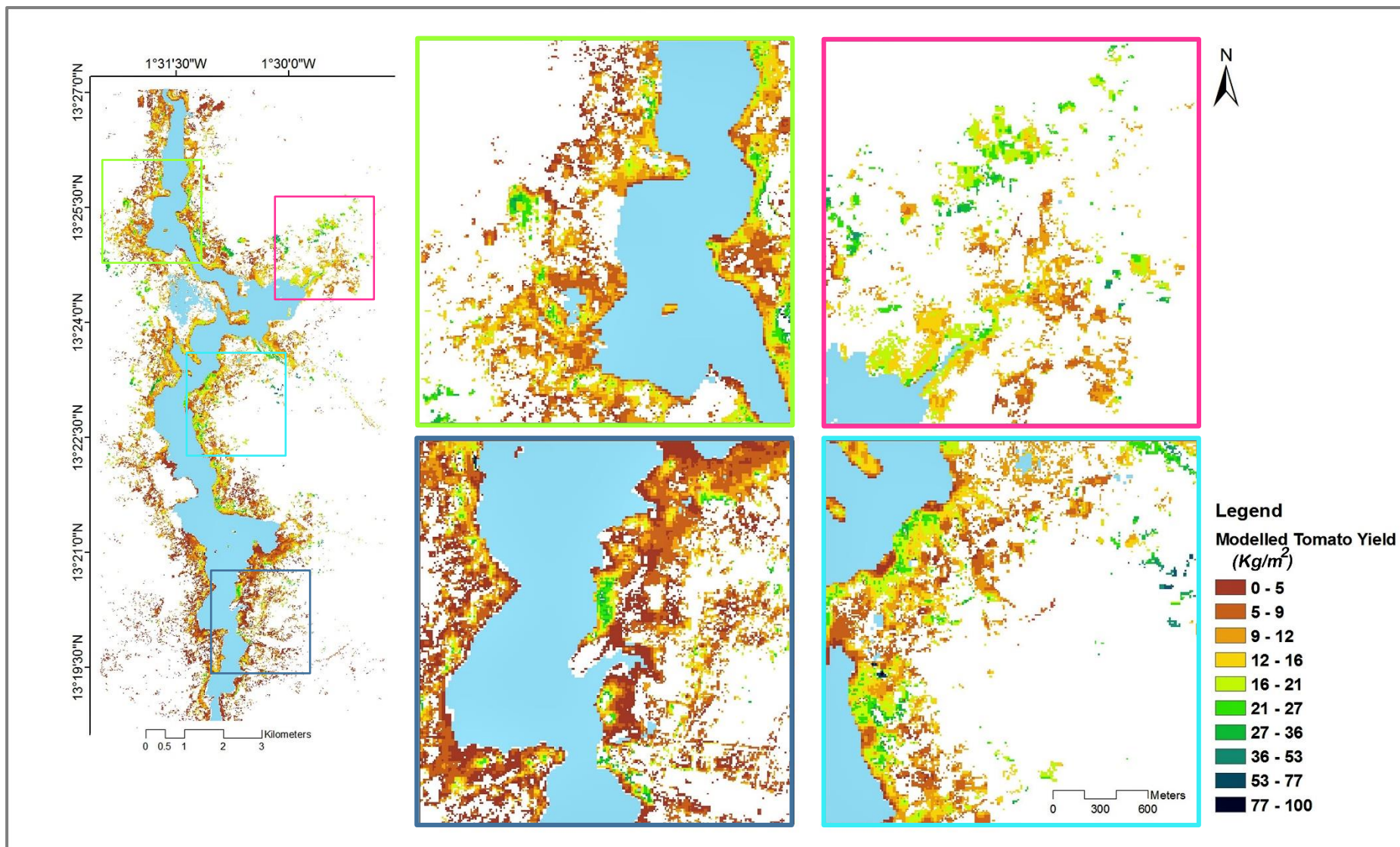
Figure 5.2: Tomato crop productivity map

V.2. Water Use Estimation

V.2.1. Evapotranspiration Fraction Modelling (Crop Coefficient)

ET_{frac} raster layers (Figure 5.3) were derived using the ETa at the satellite overpass from the latent heat flux (i.e., ET) as a residual in METRIC model and ET_o.

Ranging from 0 (dry) to 1 (wet), unitless, ET_{frac} represents the fraction of net radiant energy being used in transpiration by plants or in evaporation of moisture from the surface. In calculation of ET_{frac} (Eq. 4.23), each pixel retains a unique value for ET_{inst} that is derived from a common value for ET_o, and the representative weather station data.
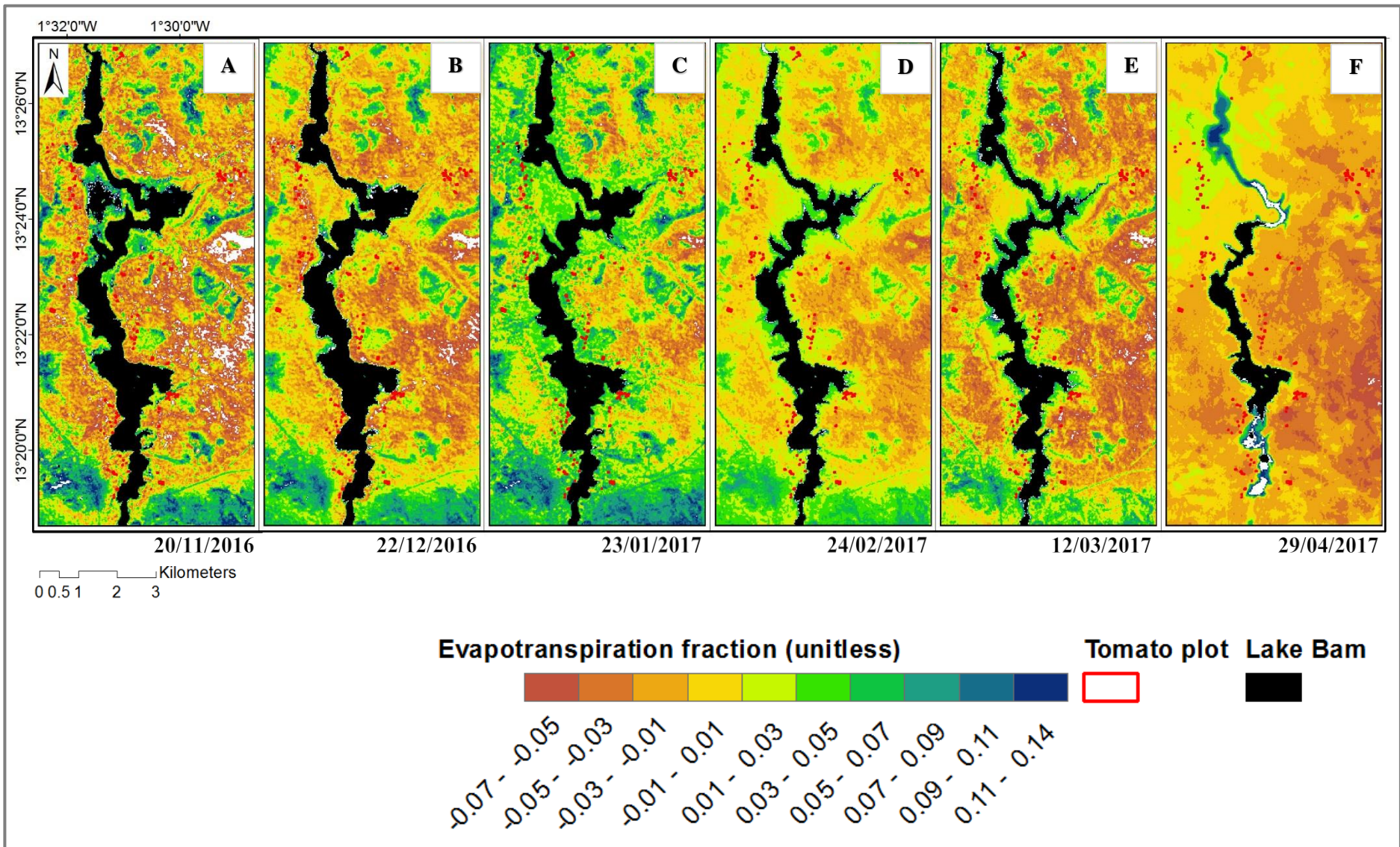
Figure 5.3: Spatio-temporal seasonal changes of evapotranspiration fraction

The first and second group of tomato (Tomato 1, Tomato 2) are respectively planted in November and December. In January, Tomato 1 is at the maturity stage whereas Tomato 2 in the development stage; hence we see significant values of ET_{frac} (Figure 5.3. C, Figure 5.4). This value is even accentuated by the fact that in late January, the third group of tomato (Tomato 3) is planted. ET_{frac} reaches its maximum in January (green colour spreads spatially and increases in intensity). The ET_{frac} starts decreasing in February and stays stable until early March (Figure 5.4) because Tomato 1 has been harvested and Tomato 2 is in mid-season growth stage. In late March and April, ET_{frac} drastically decreases because there is almost no crop on the field. Especially in late April, most of the tomato surface is totally dry apart from the north-western part of the lake where wetland is dominated at that time of the season (Figure 5.3.F).
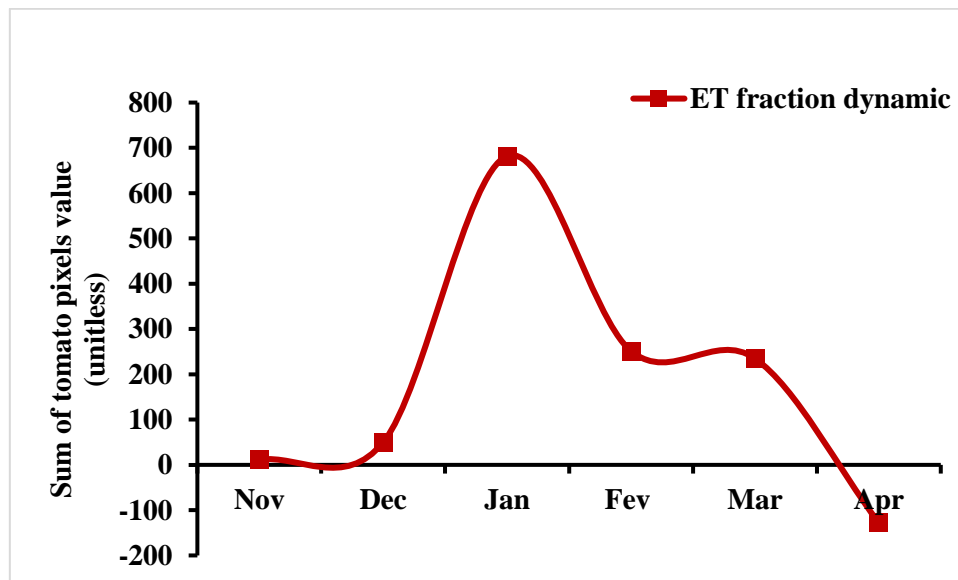
Figure 5.4: Tomato plot evapotranspiration fraction dynamic curve

V.2.2. Reference Evapotranspiration

Hourly meteorological data (section IV.1.2.2) from the closest (Ouahigouya) synoptic meteorological station was used to apply the ASCE standardized PM method to compute hourly and daily ET_{ref} (Table 5.5).

Table 5.5: ASCE-EWRI standardized reference evapotranspiration

Reference ET	2016			2017		
	Nov	Dec	Jan	Feb	Mar	Apr
ET_o (mm.hour^{-1})	5.71	4.10	5.01	5.76	7.17	12.86
ET_r (mm.day^{-1})	50.44	38.60	45.04	53.75	64.39	110.79

ET: Evapotranspiration; ET_r: short reference crop ET at daily time step
ET_o: short reference crop ET at hourly (daytime) time step – time of the satellite overpass

The time of satellite overpass ranges from 10:25 AM to 10:26 AM. ET_{ref} for both hourly and daily time scale increase gradually from December to March. However, that of April, which is the maximum, is almost twice the value of March. Allen *et al.* (2005) recommend the use of the standardized equation and procedure when possible to represent reference ET for the establishment of reproducible and universally transferable ET estimates, climatic description, and derived crop and landscape coefficients.

V.2.3. Seasonal Water Use

The daily ETa (ET_{24}) was computed by using ET_{frac} and the daily reference evapotranspiration (ET_r). METRIC allows production of maps that can express ETa at both the proportion and the spatial distribution of ETa. It has some significant advantages over many traditional applications of satellite-based energy balance in that its calibration is made using ET_{ref} rather than evaporative fraction. The use of ET_{ref} for the extrapolation of ET_{inst} from periods of 24 hours and longer compensates for regional advection effects by not tying the evaporative fraction to net radiation, since ET can exceed daily net radiation in many arid or semi-arid locations. The quantification of ETa on a field by field for each the irrigated cropping season 2016-2017 around Lake Bam is highlighted on Figure 5.5. The water loss through evaporation and transpiration are expressed in thousands $m^3.ha^{-1}$ for the entire study area.
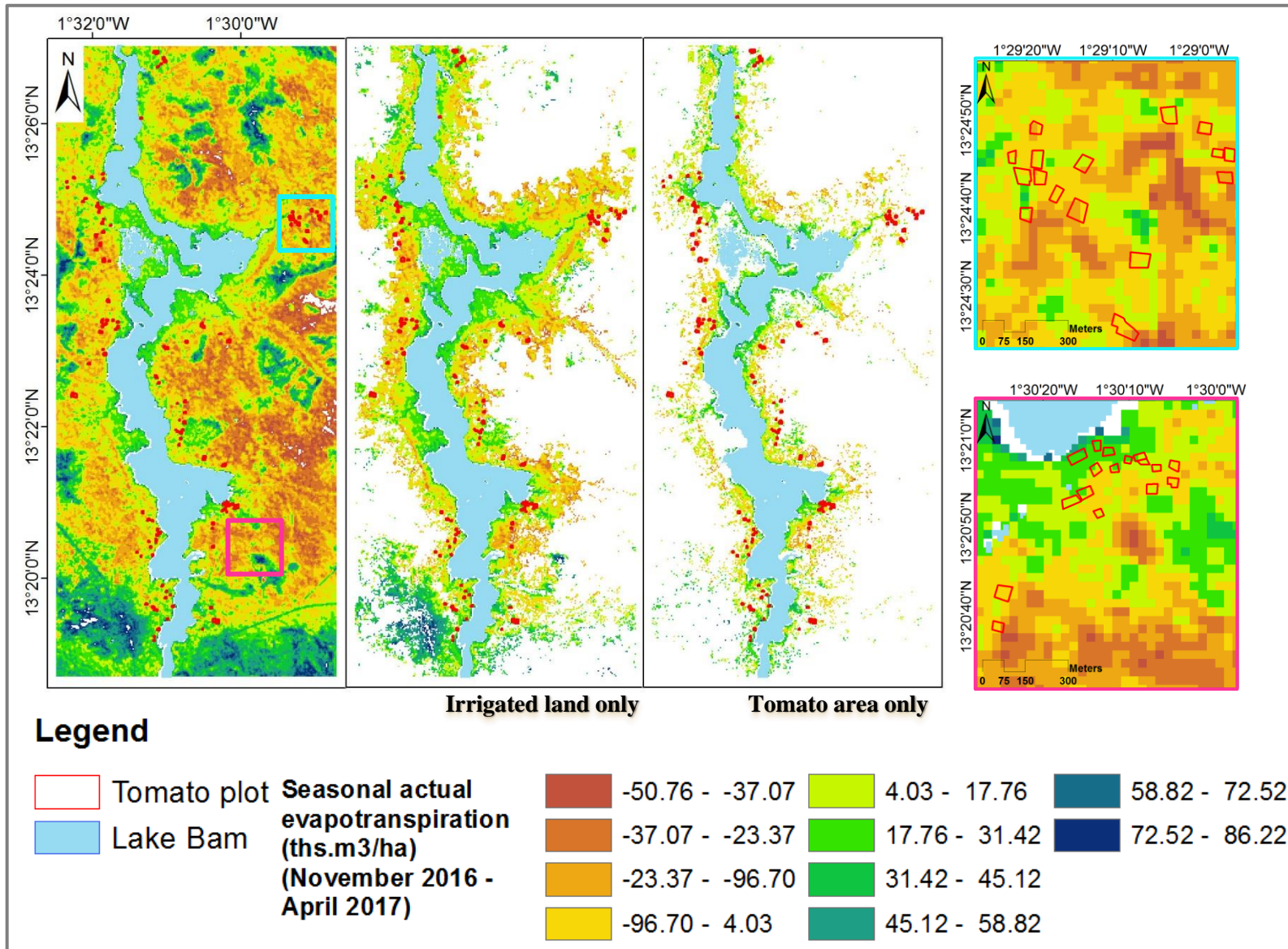
Figure 5.5: Actual seasonal evapotranspiration around Lake Bam for the growing period of 2016–2017

The mean value of ETa was used to give indications about ETa characteristics of irrigated areas for each month during the tomato cropping season. Negative values of ETa has not been treated as 0 as recommended by Wilcox and Sly (1976). Twenty-two selected tomato plots were used for inter-comparison of ETa over the tomato cropping season at the satellite overpass (Figure 5.6). Extracted values for January, February and March have the highest ETa with daily average quantity of respectivily 34.5, 11.63 and 14.35 mm.day^{-1}. This is because it is at that time most of the irrigated crops are either in development or maturity growth stage as confirmed by the crop calendar from the field survey (Table 4.5). Daily ETa for December 22 got the lowest mean and negative value, this suggests that the moisture content of the soil increased before the image acquisition.
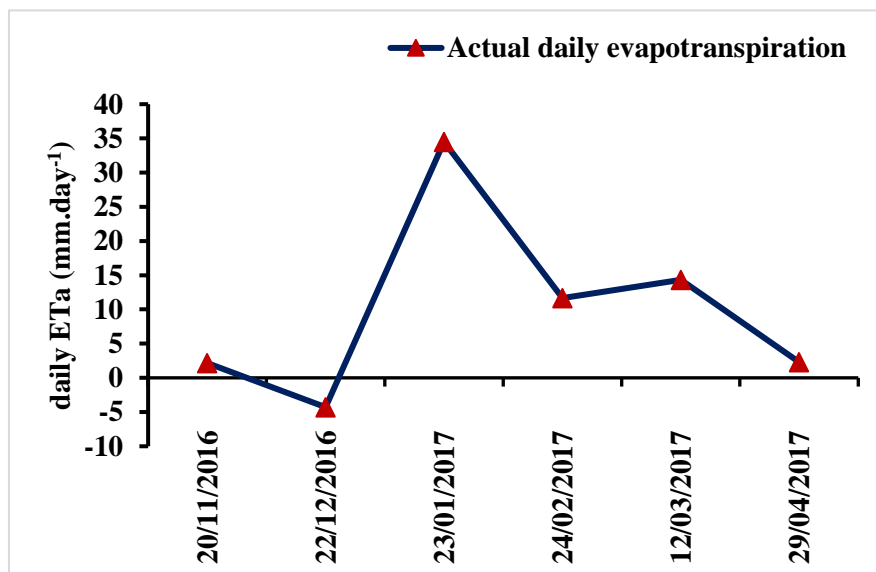
Figure 5.6: Daily evapotranspiration extracted from 22 tomato plot at the time of the satellite overpass

Estimation of ETa at different temporal scales allow the assessment of crop water usage over the season. The water use value has been retrieved by multiplying the ETa of each pixel by its surface area. As each L8 pixel has been resample to 10 m, one pixel surface is equal to 0.01 ha. The overall crop water use for the 3162.68 ha of irrigated area for the 2016/2017 cropping season was estimated at about 4.82×10^6 m^3. Tomato plots (859.75 ha) constitutes out of that 37% of irrigation water requirements with 1.48×10^6 m^3 (Figure 5.7).
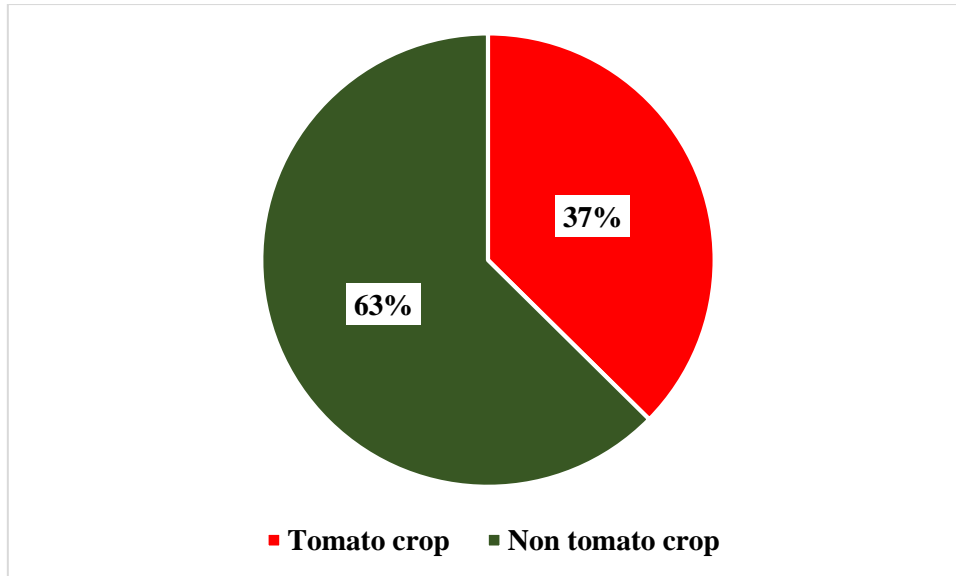
Figure 5.7: Seasonal water demand of crop under irrigated land

V.3. Crop Water Productivity

The Tomato CPM was derived by dividing the crop productivity map by the seasonal water use layer. The result, expressed in $kg.m^{-3}$, is illustrated in Figure 5.8. The WPM shows within and between field variability in tomato WP. The best bet scenario is to continue to produce more (increase WP) food from existing croplands and water. In the literature, a review based on 11 publications, from 1998 to 2008, with a number of 181 experimental point worldwide (Rashidi and Gholami, 2008) found that tomato WP from between 2.58 to 11.88 $kg.m^{-3}$. In order to benchmark to a global standard, the maximum pixel value has been set to 12 $kg.m^{-3}$. Thus, WP of the irrigated tomato crop (the most dominant crop in and around Lake Bam) varied between 0–12 $kg.m^{-3}$.

The result in Figure 5.8 shows that 600 ha out of 859.75 ha have crop WP range of 0.01-1 $kg.m^{-3}$. That is not concentrated in a particular area but dispersed in the study area. Crop WP for 22 tomato plots ranges between 0.33 and 3.5 $kg.m^{-3}$. With a standard deviation of 0.95 $kg.m^{-3}$, the mean value of crop WP for those plots is 1.24 $kg.m^{-3}$,which is largely below the international reference for tomatoes defined by 10-12 $kg.m^{-3}$ by FAO (2017).

On Figure 5.9, only 10.5% of the tomato crop area had a WP value of 2.5 $kg.m^{-3}$ or higher. About 82.05% of the tomato area had values lower than 1.63 $kg.m^{-3}$. In other words, the map shows that 89.46% (769.16 ha) tomato area for the season 2016–2017 falls in the category of low WP of 2.5 $kg.m^{-3}$or less. This clearly implies the opportunity to grow more food on the existing lands through better land and water management practices.

55
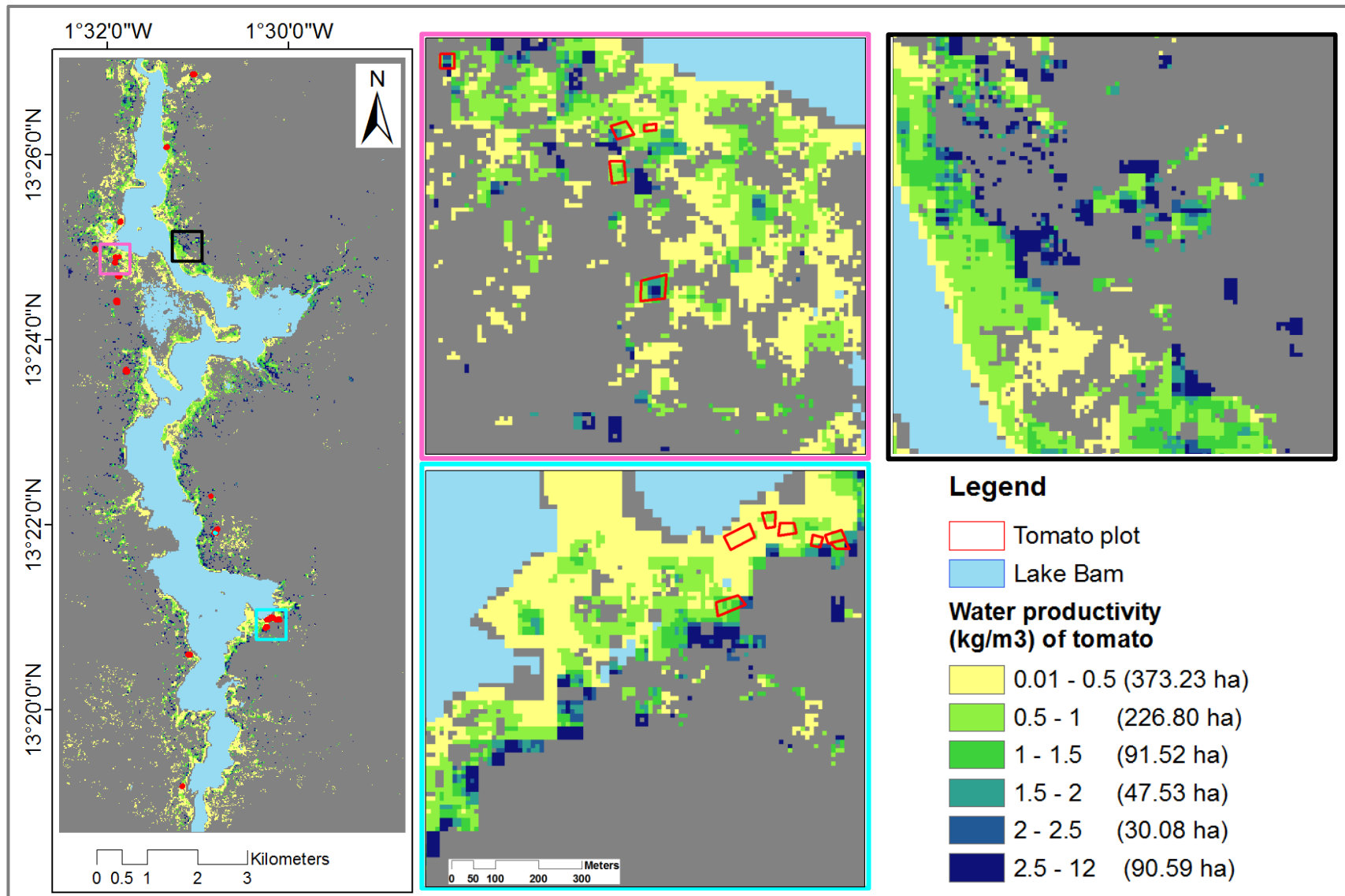
Figure 5.8: Tomato water productivity map (WPM)

Legend

Tomato plot

Lake Bam

Water productivity (kg/m3) of tomato

0.01 - 0.5 (373.23 ha)

0.5 - 1 (226.80 ha)

1 - 1.5 (91.52 ha)

1.5 - 2 (47.53 ha)

2 - 2.5 (30.08 ha)

2.5 - 12 (90.59 ha)
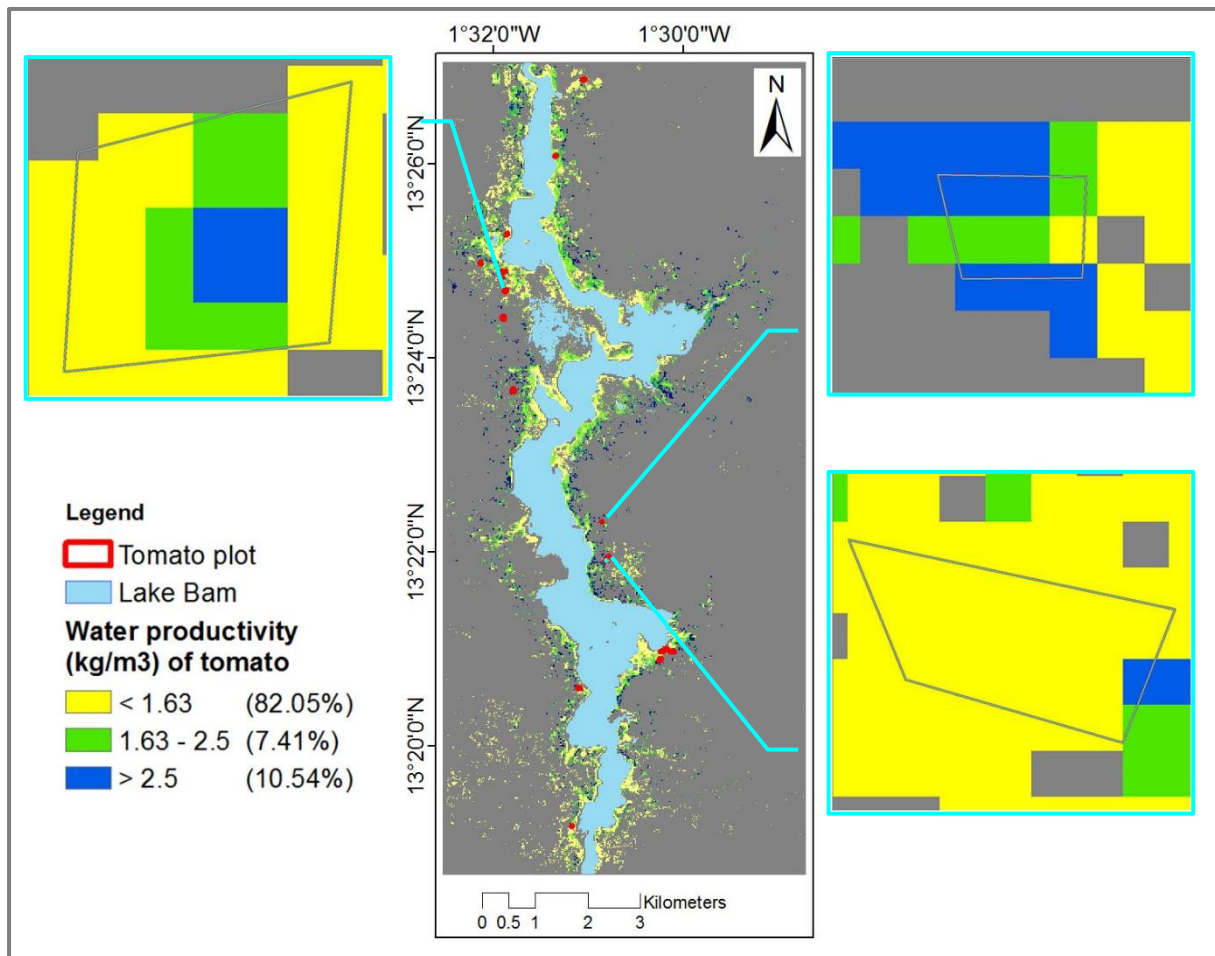
Figure 5.9: Pin-pointing areas of low and high tomato water productivity

VI. DISCUSSION OF RESULTS

This section discusses the results presented in Chapter V. It is structured around three main sub-sections. The first (VI.1) deals with results obtained from tomato crops mapping process. It discusses challenges in crop mapping in West Africa and points out the advantages of implemented classification approach. RS approach that helps in predicting yield is further discussed. The second sub-section (VI.2) discusses results obtained from water use estimation. The meaning of values found in section V.2 is argued. The last sub-section explains the practical value of WP obtained during this research.

VI.1. Crop Productivity

VI.1.1. Classification Accuracy

In a RF classification, each node is split using the best predictor out of a subset of predictors randomly chosen at that node. This somewhat counterintuitive strategy turns out to perform very well compared to many other classifiers, including discriminant analysis, support vector machines and neural networks, and is robust against overfitting (Breiman, 2001). In addition, it is very user-friendly in the sense that it has only two parameters (the number of variables in the random subset at each node and the number of trees in the forest), and is usually not very sensitive to their values. The method used in our study executed well since the overall accuracy of the consecutive classification was respectively 92.83% and 74.33%.

VI.1.2. Crop Calendar Overlaps and Variable Planting Dates

An underlying challenge to crop separability that was identified in classifications step is the overlap of the cropping calendar. Farmers in the study area do not seed and/or plow at the same time. As identified by Hassan (1996), farmers who have plans of planting multiple crops on the same land may start cultivation earlier than those who practise mono-cropping. However, identification of unique crop spectral profiles is essential for the success of crop mapping (Odenweller and Johnson, 1984). That practice, like around Lake Bam, reduces classification accuracy by creating similarities in the spectral profiles of different crops and differences in the spectral profiles of same crops (Forkuor *et al*., 2014). This is the case for onions and tomato in January due to analogous response captured by the satellite. Peña-Barragán *et al*. (2011 also attributed the major confusion between some summer crops (safflower, sunflower and tomatoes) to similarities in their development stages and cropping calendar.

Despite the challenge of cropping calendar overlaps, the implemented approach, and availability of satellite imagery for the whole cropping season (18 S-2 images) achieved acceptable separability for tomato crop class.

VI.1.3. Landscape Heterogeneity and Per-pixel Classification

Traditional pixel-based analysis of remotely sensed data results in inaccurate identification of some crops due to pixel heterogeneity, mixed pixels, spectral similarity, and crop pattern variability. High spectral within-field heterogeneity, which could be caused by variations in soil fertility, soil moisture conditions and pest or diseases, often results in a classifier assigning different classes to pixels in the same field (Smith and Fuller, 2001). In West Africa, the heterogeneity of the landscape (Cord *et al.*, 2010) and frequent intercropping aggravates this situation.

The accuracy of the classifications in this study could be improved by adopting a per-field classification approach as reported by Ban (2003). But, per-field approaches require availability of a topographic vector database of field boundaries. This is a major limitation for implementation of such approaches in most parts of West Africa, since vector databases of field boundaries do not exist.

VI.1.4. Yield Predictions

The regression model developed to estimate the crop productivity can be used to predict crop yield in subsequent years provided no significant changes are made to the existing cropping system. Derivation of the dependent RS variables (Eq 5.1) ahead of the harvest season can provide agricultural managers with some yield estimates at the end of the season. However, there could be high uncertainties in such yield predictions. This is particularly the case in early season. Generally, the uncertainty declines towards the end of the season (Hansen, 2005). Although VIs like NDVI is widely used for yield modelling (Kumar, 2015), it is a very poor indicator of vegetation properties when quantitative, reliable, accurate estimates are required. Indeed, the problem with NDVI is that while high values of this index do suggest the presence of vegetation in the scene, its actual numerical value is quite sensitive to a number of other factors, including some which have nothing to do with the vegetation itself. For instance, cloud contamination, soil moisture, or angular configuration of illumination and observation can significantly affect the index value, even if the surface type and properties did not change. By contrast, physically based parameters like fAPAR, LAI and fCover are more reliable variables to describe the instantaneous productivity of plant canopies. They are a critical parameter for

assessing vegetation growth rates and predicting crop productivity (Fensholt, Sandholt, and Rasmussen, 2004; Claverie *et al.*, 2012).

This explains why most of the highly correlated independent variables in the regression model turned out to be biophysical parameters, especially those related to the late season. A study in Argentina by Lopresti, Di Bella, and Degioanni (2015) demonstrated wheat yield prediction 30 days before the harvest, after stages of heading and anthesis. This has been done at a regional scale combining yield data of past nine seasons. The relationship between moderate-resolution imaging spectroradiometer (MODIS) - 250 m of spatial resolution image and NDVI was explained by an R^2 value of 0.75. Our case used data from one season and a high resolution sensor (10 m).

Regular monitoring of biomass production can be used to timely identify areas of lagging production which could indicate pests, diseases or other potential threats. Timely detection supports timely action to mitigate crop damage. Additionally, money can be saved due to a more localised application of pesticides. On a larger scale, biomass production of different fields within an irrigation system, region or country can be compared, giving insight in spatial variations in production.

VI.2. Evapotranspiration

VI.2.1. Water Use Maps

The accuracy of water use maps depends on the validity of ET_{frac} maps and ET_{ref}. One should generally expect ET_{frac} values to range from 0 to about 1.0 (Jensen, Burman, and Allen, 1990). At a completely dry pixel, ET=0 and therefore ET_{frac} =0. ET_{ref} generally represents an upper bound on ET for large expanses of well-watered vegetation. Negative values for ET_{frac} can occur in METRIC due to systematic errors caused by various assumptions made earlier in the energy balance process and due to random error components. So for the dry pixels (non-transpiring vegetation or dry soil), 0 should be considered as the final ET_{frac} value. That was the case in this study.

ET_r should be permanently greater than the estimated daily ETa for a validation of the model. As shown in Table 6.1, for all the dates of ETa modelling, ET_r is constantly greater than ETa. Platonov *et al.* (2008) suggested that to get the same ETa values, applying different methods for ET_r calculation requires use of different crop coefficients. According to FAO (2017), maximum K_c (ET_{frac}) for tomato is around 1.2-1.4, meaning that seasonal ET can be underestimated and WP values overestimated. Similar observation was made by Biradar *et al.*

(2008) in ET modelling of wheat, maize and cotton based on simplified surface energy balance model (SSEB).

Table 6.1: Reference evapotranspiration and 24 actual evapotranspiration comparison

ET value	2016			2017		
	Nov	**Dec**	**Jan**	**Feb**	**Mar**	**Apr**
max ETa $(mm.day^{-1})$	26	30.88	18.48	28.4	30.91	15.3
ETr $(mm.day^{-1})$	50.44	38.6	45.04	53.75	64.39	110.79

ET: Evapotranspiration; max ETa: maximum actual ET pixels value for the day
ETr: short reference crop ET at daily time step

VI.2.2. Water Requirement

Monitoring ET variation on a regular basis gives a direct insight in crop water status. Areas of low crop water use can be identified, either within field status or within an entire irrigation system. Large differences in ET within fields might indicate points of water logging or faulty irrigation equipment or prompt better field levelling. The L8 pixel 30m×30m (0.09 ha) is lower than the mean surface of tomato plot which is 0.2 ha (the maximum is 0.62 ha, field survey). This can explain why spatial variation of ETa within tomatoes plot is so constant. In Asia and Europe, where land for irrigation is very big, the variation within the field could be apparent and obvious. Irrigation managers can monitor the crop water consumption within their area on a real-time basis, or look at accumulated figures from an entire season and analyse the spatial distribution of crop water consumption from year to year.

VI.3. Crop Water Productivity Standard

VI.3.1. Tomato Water Productivity Benchmarking

The extensive literature review of WP values in the world for main irrigated crops by Zwart and Bastiaanssen (2004) covers wheat, rice, cotton, and maize. A plausible crop WP range under farm management conditions for tomato has been investigated by Rashidi and Gholami (2008) in Iran. The crop WP range is determined by taking the 5 and 95 percentiles of the cumulative frequency distribution. The range of WP values for tomatoes is very large (2.58-11.88 kg.m^{-3}) with a maximum frequency of crop WP ranged from 3.5 to 5.5 yields among all experiments. Around Lake Bam, only about 10% of the area under tomatoes has WP values within this range. FAO (2017) established that the water utilisation efficiency for harvested yield for fresh tomatoes is 10 to 12 kg.m^{-3}. Our results imply that there is significant

scope for increasing WP for growing more food to feed the ballooning populations in the coming decades without having to increase allocations of croplands and\or water.

VI.3.2. Improving Crop Water Productivity

In order to understand why some plots in the study area have higher WP values than others, a field visit on the lower WP plots needs to be undertaken to know different land and water management practices – amount of pesticides, soil properties, etc. Platonov *et al.*, (2008) found that in Galaba farms (Uzbekistan), lower WP values occurred mainly because of the high percentage of soil salinity (43% of farms) and water logging (31% of farms). The other factors that influenced WP variations were land levelling (14%), water deficit (7%), and others (5%).

The highest values of tomato WP (9.26 to 13.3 kg.m^{-3}) are reported in Iran by Sadreghaen *et al*. (2002) in a combination of alternate micro irrigation and deficit irrigation. In South Africa, Pienaar (2014) utilised and tested various innovation technologies aimed at increasing tomato WP. Improved management practices, such as soil nutrient management and mulching, were introduced in the 2nd and 3rd seasons of tomato trials at field level at Mzilela farm. Results showed tomato yield increased from an average of 26.5 t.ha^{-1} to 120.9 t.ha^{-1} and WP increased from 4.61 kg.m^{-3} to 17.69 kg.m^{-3} after experimentation. Thus, specific practice to specify place can increase tomato WP around Lake Bam.

The next and last chapter of this thesis briefly explained the implications of the findings for all the stakeholders in line with the Lake Bam water resource and proffer possible solutions for the limitation of my study.

VII. CONCLUSION

VII.1. Principal Findings

WP spatial model has been developed using multi-temporal high resolution RS data (S-2 and L8). Examination of rudimentary irrigation system around Lake Bam (Burkina Faso) to map crop WP through the use of satellite-based approaches highlights areas of high and low WP. During the determination of the two main inputs of WP in the crop productivity part, RF algorithm gave satisfactory result to the two-step approach for image classification as well as the formulated model prediction. Then, for the second input, spatio-temporal water use by crop for all irrigated fields as well as for tomatoes was succinctly quantified.

Low WP areas dominate around Lake Bam. This result will help us focus on areas of low WP and establish causes for same. Once this is accomplished, strategies can be developed to increase WP of these areas. If this is achieved, food security of future generations can be secured without having to increase croplands and\or greater water use.

VII.2. Policy and Practice Recommendations

VII.2.1. Crop Monitoring and Biomass Production

Actual biomass production data is valuable for individual farmers. Such information can help farmers to identify potential weak spots in their fields that require urgent attention. Moreover, efficient application of costly agricultural inputs such as fungicides and pesticides can mean substantial financial savings. Extrapolated crop yield maps can thus allow improvement on areas of low productivity with appropriate management practices.

Production gap analyses can give governments and policymakers valuable insight into the extent to which the productivity capacity can improve within their borders to be able to feed future generations. This will significantly influence the development of food security policies and agricultural investments. In other word, the current production can be compared to actual and realistically attainable production for a certain region with similar circumstances (variety crop, soil moisture and weather conditions) so that with good farming practices and better use of agricultural inputs the same high production can be attainable.

VII.2.2. Water Resources Management

Evapotranspiration data is valuable for a lot of people dealing with water resources, from farmers themselves to policymakers. Individual farmers can use the data to optimise their crop water status, and thus their production. This is in collaboration with agricultural advisors who can oversee bottlenecks on a larger scale and support farmers to improve their practices to

achieve the best possible yield. Water managers can use information to improve water distribution within their irrigation systems. However, informal irrigation system in many places in West Africa can restrict the proper usage of that spatial statistic.

For Policymakers ETa information will mean factual information on which to to base sustainable water management decisions. Indeed, combining ET data with land use maps provides quantified information on water consumption per land use categories such as agricultural land, forests, rangeland and wetlands. This could support the monitoring of water quotas per category as defined by (inter-) national policy and justify actions to be taken to reach such quota. Up-to-date and basin-wide ET information is a crucial input in hydrological models capable of simulating the future impact of water management decisions made today. Being able to compare water availability scenarios when deciding on future water management policies will support water management authorities tasked with the development of long-term river basin management plans to take the most sustainable course.

A policy mechanism for water extraction from the Lake using motor pumping could be planned. A field interview around Lake Bam revealed that the three main motor pumps used around the lake draw water at a rate of (power) 60 to 100 $m^3.h^{-1}$. Thus, for each month, a timing of pumping water based on the type of motor pump can be set to achieve crop water requirements. This can improve WUE which will subsequently increase WP around the lake.

VII.3. Further Studies

The need for conducting WPM studies using high spatial resolution remote sensing data from Landsat type sensors is critical so that crop level water use and WP can be studied. This is a significant advantage over coarser resolution imagery like the one from MODIS. Existence of high resolution images like S-2 is now overcoming the problem of resolution for crop delineation. METRIC model requires rigorous meteorological data which may discourage WPM studies using remote sensing in data scare region.

Due to inability of optical systems to acquire useful images during cloudy weather, contribution of synthetic-aperture radar (SAR) data in both crop mapping and ET modelling is of paramount interest. Indeed, SAR sensors are independent of weather conditions, as they have their own source of energy and therefore do not depend on the sun's energy. Its image acquired either during the day or night has high resolution capability. A combination of very high image resolution like the one of rapid eye sensors (6.5 m) and SAR may increase

variability of ET within field and ensure accurate discrimination of different crop types. This will finally reduce uncertainty of the modelling.

This study was limited to physical part of WP, but a monetary value on WP will put more attention on both farmers and policy makers on the importance of improving WP. However, the need to understand factors affecting WP is very important. Be able to come out with the degree of influence of various factors on WP variations within and between fields may necessarily lead to finding appropriate solutions.

REFERENCES

Abdullaev, I., & Molden, D. (2004). Spatial and temporal variability of water productivity in the Syr Darya Basin, central Asia. *Water Resources Research*, *40*(8), 1–6. http://doi.org/10.1029/2003WR002364

Allen, R. (1996). Assessing integrity of weather data for reference evapotranspiration estimation. *Journal of Irrigation and Drainage Engineering*, *122*(April), 97–106. http://doi.org/10.1061/(ASCE)0733-9437(1996)122:2(97)

Allen, R. (2000). *RAPID long-wave radiation calculations and model comparisons Internal report*. Kimberly, Idaho.

Allen, R., Irmak, A., Trezza, R., Hendrickx, J. M. H., Bastiaanssen, W., & Kjaersgaard, J. (2011). Satellite-based ET estimation in agriculture using SEBAL and METRIC. *Hydrological Processes*, *25*(26), 4011–4027. http://doi.org/10.1002/hyp.8408

Allen, R., Tasumi, M., Morse, A., Trezza, R., Wright, J. L., Bastiaanssen, W., … Robison, C. W. (2007). Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC) - Applications. *Journal of Irrigation and Drainage Engineering-Asce*, *133*(4), 395–406. http://doi.org/10.1061/(asce)0733-9437(2007)133:4(395)

Allen, R., Tasumi, M., & Trezza, R. (2007). Satellite-Based Energy Balance for Mapping Evapotranspiration with Internalized Calibration (METRIC)— Model. *Journal of Irrigation and Drainage Engineering*, *133*(4), 395–406. http://doi.org/10.1061/(ASCE)0733-9437(2007)133

Allen, R., Walter, I., Elliott, R., Terry, H., Itenfisu, D., & Jensen, M. (2005). *The ASCE Standardized Reference Evapotranspiration Equation*. Retrieved from https://www.kimberly.uidaho.edu/water/asceewri/ascestzdetmain2005.pdf

ASCE. (2000). ASCE's Standardized Reference Evapotranspiration Equation Authors: ASCE Standardization of Reference Evapotranspiration Task Committee. In *Conference Arizona* (pp. 1–11). Phoenix.

Ban, Y. (2003). Synergy of multitemporal ERS-1 SAR and Landsat TM data for classification of agricultural crops. *Canadian Journal of Remote Sensing*, *29*(4), 518–526. http://doi.org/10.5589/m03-014

Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., … Leroy, M. (2007). LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION. Part 1: Principles of the algorithm. *Remote Sensing of Environment*, *110*(3), 275–286. http://doi.org/10.1016/j.rse.2007.02.018

Bastiaanssen, W. (1995). *Regionalization of surface flux densities and moisture indicators in composite terrain: a remote sensing approach under clear skies in Mediterranean climates. Doctoral thesis, Wageningen Agricultural University, Wageningen The Netherlands.* http://doi.org/90-5485-465-0

Bastiaanssen, W., & Ali, S. (2003). A new crop yield forecasting model based on satellite measurements applied across the Indus Basin, Pakistan. *Agriculture, Ecosystems & Environment*, *94*(3), 321–340. http://doi.org/10.1016/S0167-8809(02)00034-8

Bastiaanssen, W., Waters, R., Allen, R., Tasumi, M., & Trezza, R. (2002). *Manual for Surface Energy Balance Algorithms for Land*.

Berger, M., Moreno, J., Johannessen, J. A., Levelt, P. F., & Hanssen, R. F. (2012). ESA's sentinel missions in support of Earth system science. *Remote Sensing of Environment*, *120*, 84–90. http://doi.org/10.1016/j.rse.2011.07.023

Bhattarai, N., Quackenbush, L. J., Im, J., & Shaw, S. B. (2017). A new optimized algorithm for automating endmember pixel selection in the SEBAL and METRIC models. *Remote Sensing of Environment*, *196*, 178–192. http://doi.org/10.1016/j.rse.2017.05.009

Biradar, C. M., Thenkabail, P. S., Platonov, A., Xiao, X. M., Geerken, R., Noojipady, P., … Vithanage, J. (2008). Water productivity mapping methods using remote sensing. *Journal of Applied Remote Sensing*, *2*(November), 22. http://doi.org/Artn 023544\rDoi 10.1117/1.3033753

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. http://doi.org/10.1023/A:1010933404324

Cai, X., Sharma, B. R., Matin, M. A., Sharma, D., & Gunasinghe, S. (2010). *An Assessment of Crop Water Productivity in the Indus and Ganges River Basins: Current Status and Scope for Improvement. IWMI Research Report* (Vol. 140). Colombo, Sri Lanka. http://doi.org/10.5337/2010.232

Cai, X., & Thenkabail, P. (2010). *Using remote sensing to assess crop water productivity*. Colombo, Sri Lanka.

Calera, A., Campos, I., Osann, A., D'Urso, G., & Menenti, M. (2017). Remote Sensing for Crop Water Management: From ET Modelling to Services for the End Users. *Sensors*, *17*(5), 1104. http://doi.org/10.3390/s17051104

Cammarano, D., Fitzgerald, G. J., Casa, R., & Basso, B. (2014). Assessing the robustness of vegetation indices to estimate wheat N in mediterranean environments. *Remote Sensing*, *6*(4), 2827–2844. http://doi.org/10.3390/rs6042827

CAMS (Copernicus Atmosphere Monitoring Service). (2017). CAMS radiation service - www.soda-pro.com. Retrieved November 21, 2017, from http://www.soda-pro.com/web-services/radiation/cams-radiation-service?p_p_lifecycle=0&p_p_id=58

Chávez, J. L., Neale, C., Hipps, L., Prueger, J., & Kustas, W. (2005). Comparing aircraft-based remotely sensed energy balance fluxes with eddy covariance tower data using heat flux source area functions. *Journal of Hydrometeorology*, *6*(6), 923–940. http://doi.org/10.1175/JHM467.1

Claverie, M., Demarez, V., Duchemin, B., Hagolle, O., Ducrot, D., Marais-Sicre, C., … Dedieu, G. (2012). Maize and sunflower biomass estimation in southwest France using high spatial and temporal resolution remote sensing data. *Remote Sensing of Environment*, *124*, 844–857. http://doi.org/10.1016/j.rse.2012.04.005

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, *20*(1), 37–46. http://doi.org/10.1177/001316446002000104

Congedo, L. (2017). *Semi-Automatic Classification Plugin Documentation*. http://doi.org/10.13140/RG.2.1.2137.4884

Cook, S., Gichuki, F., & Turral, H. (2006). *Agricultural Water Productivity: Estimation at Plot , Farm and Basin Scale. Draft working paper, CGIAR Challenge program on Water & food.*

Cord, A., Conrad, C., Schmidt, M., & Dech, S. (2010). Standardized FAO-LCCS land cover mapping in heterogeneous tree savannas of West Africa. *Journal of Arid Environments*, *74*(9), 1083–1091. http://doi.org/10.1016/j.jaridenv.2010.03.012

Dahms, T., Seissiger, S., Borg, E., Vajen, H., Fichtelmann, B., & Conrad, C. (2016). Important Variables of a RapidEye Time Series for Modelling Biophysical Parameters of Winter Wheat. *Photogrammetrie - Fernerkundung - Geoinformation*, *2016*(5), 285–299. http://doi.org/10.1127/pfg/2016/0303

Dahms, T., Seissiger, S., Conrad, C., & Borg, E. (2016). MODELLING BIOPHYSICAL PARAMETERS OF MAIZE USING LANDSAT 8 TIME SERIES. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLI-B2*(July), 171–175. http://doi.org/10.5194/isprs-archives-XLI-B2-171-2016

Dong, T., Meng, J., & Shang, J. (2015). Evaluation of Chlorophyll-Related Vegetation Indices Using Simulated Sentinel-2 Data for Estimation of Crop Fraction of Absorbed Photosynthetically Active Radiation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *8*(8), 4049–4059. http://doi.org/10.1080/01431161.2015.1042122

Draper, N. R., & Smith, H. (1998). *Applied Regression Analysis (Wiley Series in Probability and Statistics). Time*. http://doi.org/10.1002/0471722235

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., … Bargellini, P. (2012). Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment*, *120*, 25–36. http://doi.org/10.1016/j.rse.2011.11.026

Enclona, E. A., Thenkabail, P. S., Celis, D., & Diekmann, J. (2004). Within-field wheat yield prediction from IKONOS data: a new matrix approach. *International Journal of Remote Sensing*, *25*(2), 377–388. http://doi.org/10.1080/0143116031000102485

ESA. (2017). Introducing Sentinel-2. Retrieved September 7, 2017, from http://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus/Sentinel-2/Introducing_Sentinel-2

ESPA. (2017). Landsat 8 Surface Reflectance data. Retrieved December 16, 2017, from https://espa.cr.usgs.gov/

FADPA (Food and Agriculture Policy Decision Analysis). (2014). *Country Fact Sheet on Food and Agriculture Policy Trends: Burkina Faso*.

FAO. (2017). FAOSTAT online database- Crop information: tomato. Retrieved December 13, 2017, from http://www.fao.org/land-water/databases-and-software/crop-information/tomato/en/

Fensholt, R., Sandholt, I., & Rasmussen, M. S. (2004). Evaluation of MODIS LAI, fAPAR and the relation between fAPAR and NDVI in a semi-arid environment using in situ measurements. *Remote Sensing of Environment*, *91*(3–4), 490–507. http://doi.org/10.1016/j.rse.2004.04.009

Foody, G. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, *80*(1), 185–201. http://doi.org/10.1016/S0034-4257(01)00295-4

Forkuor, G., Conrad, C., Thiel, M., Ullmann, T., & Zoungrana, E. (2014). Integration of

optical and synthetic aperture radar imagery for improving crop mapping in northwestern Benin, West Africa. *Remote Sensing*, *6*(7), 6472–6499. http://doi.org/10.3390/rs6076472

Fortes, R., Prieto, M. H., García-Martín, A., Córdoba, A., Martínez, L., & Campillo, C. (2015). Using NDVI and guided sampling to develop yield prediction maps of processing tomato crop. *Spanish Journal of Agricultural Research*, *13*(1), 1–9. http://doi.org/10.5424/sjar/2015131-6532

Friedl, M. A., & Brodley, C. E. (1997). Decision tree classification of land cover from remotely sensed data. *Remote Sensing of Environment*, *61*(3), 399–409. http://doi.org/10.1016/S0034-4257(97)00049-7

Gitelson, A. A., & Merzlyak, M. N. (1996). Signature Analysis of Leaf Reflectance Spectra: Algorithm Development for Remote Sensing of Chlorophyll. *Journal of Plant Physiology*, *148*(3–4), 494–500. http://doi.org/10.1016/S0176-1617(96)80284-7

Gobron, N., & Verstraete, M. M. (2009). *Fraction of Absorbed Photosynthetically Active Radiation (FAPAR). Assessment of the status of the development of the standards for the Terrestrial Essential Climate Variables*. Rome, Italy. Retrieved from http://www.fao.org/gtos/ECV-T10.html

Hansen, J. W. (2005). Integrating seasonal climate prediction and agricultural models for insights into agricultural practice. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1463), 2037–2047. http://doi.org/10.1098/rstb.2005.1747

Hassan, R. M. (1996). Planting strategies of maize farmers in Kenya: A simultaneous equations analysis in the presence of discrete dependent variables. *Agricultural Economics*, *15*(2), 137–149. http://doi.org/10.1016/S0169-5150(96)01194-2

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, *83*(1–2), 195–213. http://doi.org/10.1016/S0034-4257(02)00096-2

Institut National des Statistiques de la Démographie. (2015). *Annuaire statistiques 2014*.

Jensen, M., Burman, R., & Allen, R. (1990). *Evapotranspiration and irrigation water requirements. ASCE Manuals and Reports on Engineering Practice No. 70* (Vol. 1).

Jiang, Z., Huete, A. R., Didan, K., & Miura, T. (2008). Development of a two-band enhanced vegetation index without a blue band. *Remote Sensing of Environment*, *112*(10), 3833–3845. http://doi.org/10.1016/j.rse.2008.06.006

Jiménez-Muñoz, J. C., Sobrino, J. A., Skokoviˊ, D., Mattar, C., & Cristóbal, J. (2014). Land surface temperature retrieval methods from Landsat-8 thermal infrared sensor data. *Geoscience and Remote Sensing Letters, IEEE*, *11*(10), 1840–1843. http://doi.org/10.1109/LGRS.2014.2312032

Jin, H., Stehman, S. V, & Mountrakis, G. (2014). Assessing the impact of training sample selection on accuracy of an urban classification : a case study in Denver , Colorado. *International Journal of Remote Sensing*, *35*(6), 2067–2081. http://doi.org/10.1080/01431161.2014.885152

Kastens, J. H., Kastens, T. L., Kastens, D. L. A., Price, K. P., Martinko, E. A., & Lee, R. Y. (2005). Image masking for crop yield forecasting using AVHRR NDVI time series

imagery. *Remote Sensing of Environment*, *99*(3), 341–356. http://doi.org/10.1016/j.rse.2005.09.010

Kumar, D. (2015). Remote Sensing Based Vegetation Indices Analysis to Improve Water Resources Management in Urban Environment. *Aquatic Procedia*, *4*(Icwrcoe), 1374–1380. http://doi.org/10.1016/j.aqpro.2015.02.178

Liaw, A., & Wiener, M. (2002). Classification and Regression by randomForest. *R News*, *2*(December), 18–22. http://doi.org/10.1177/154405910408300516

Long, D., Longuevergne, L., & Scanlon, B. R. (2014). Uncertainty in evapotranspiration from land surface modeling, remote sensing, and GRACE satellites. *Water Resources Research*, *50*(2), 1131–1151. http://doi.org/10.1002/2013WR014581

López-Lozano, R., Duveiller, G., Seguini, L., Meroni, M., García-Condado, S., Hooker, J., … Baruth, B. (2015). Towards regional grain yield forecasting with 1km-resolution EO biophysical products: Strengths and limitations at pan-European level. *Agricultural and Forest Meteorology*, *206*, 12–32. http://doi.org/10.1016/j.agrformet.2015.02.021

Lopresti, M. F., Di Bella, C. M., & Degioanni, A. J. (2015). Relationship between MODIS-NDVI data and wheat yield: A case study in Northern Buenos Aires province, Argentina. *Information Processing in Agriculture*, *2*(2), 73–84. http://doi.org/10.1016/j.inpa.2015.06.001

Lumley, T. (2017). *leaps: Regression subset selection*. R package version 3.0. Retrieved from http://cran.r-project.org/web/packages/leaps/leaps.pdf

Majasalmi, T., & Rautiainen, M. (2016). The potential of Sentinel-2 data for estimating biophysical variables in a boreal forest: a simulation study. *Remote Sensing Letters*, *7*(5), 427–436. http://doi.org/10.1080/2150704X.2016.1149251

Meijninger, W. M. L., Green, A. E., Hartogensis, O. K., Kohsiek, W., Hoedjes, J. C. B., Zuurbier, R. M., & De Bruin, H. A. R. (2002). Determination of area-averaged water vapour fluxes with large aperture and radio wave scintillometers over a heterogeneous surface - Flevoland field experiment. *Boundary-Layer Meteorology*, *105*(1), 63–83. http://doi.org/10.1023/A:1019683616097

Mertz, O., Lykke, A., & Reenberg, A. (2001). Importance and seasonality of vegetable consumption and marketing in Burkina Faso. *Economic Botany*, *55*(2), 276–289. http://doi.org/10.1007/BF02864565

Molden, D., Frenken, K., Barker, R., Fraiture, C. De, Attapatu, S., Giordano, M., & Inocencio, A. (2007). Trends in water and agricultural development. *Water for Food, Water for Life: A Comprehensive Assessment of Water Management in Agriculture*, 57–89. http://doi.org/10.4324/9781849773799

Molden, D., Oweis, T., Steduto, P., Bindraban, P., Hanjra, M. A., & Kijne, J. (2010). Improving agricultural water productivity: Between optimism and caution. *Agricultural Water Management*, *97*(4), 528–535. http://doi.org/10.1016/j.agwat.2009.03.023

Möller, M., Gerstmann, H., Gao, F., Dahms, T. C., & Förster, M. (2017). Coupling of phenological information and simulated vegetation index time series: Limitations and potentials for the assessment and monitoring of soil erosion risk. *Catena*, *150*, 192–205. http://doi.org/10.1016/j.catena.2016.11.016

Moser, L., Schmitt, A., Wendleder, A., & Roth, A. (2016). Monitoring of the Lac Bam

Wetland Extent Using. *Remote Sensing*, *8*(4). http://doi.org/10.3390/rs8040302

Moser, L., Voigt, S., Schoepfer, E., & Palmer, S. (2014). Multitemporal wetland monitoring in sub-Saharan West-Africa using medium resolution optical satellite data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *7*(8), 3402–3415. http://doi.org/10.1109/JSTARS.2014.2336875

Mu, Q., Heinsch, F. A., Zhao, M., & Running, S. W. (2007). Development of a global evapotranspiration algorithm based on MODIS and global meteorology data. *Remote Sensing of Environment*, *111*(4), 519–536. http://doi.org/10.1016/j.rse.2007.04.015

Nagler, P. L., Scott, R. L., Westenburg, C., Cleverly, J. R., Glenn, E. P., & Huete, A. R. (2005). Evapotranspiration on western U.S. rivers estimated using the Enhanced Vegetation Index from MODIS and data from eddy covariance and Bowen ratio flux towers. *Remote Sensing of Environment*, *97*(3), 337–351. http://doi.org/10.1016/j.rse.2005.05.011

Nieto, J. D., Fisher, M., Cook, S., Laderach, P., & Lundy, M. (2012). Weather indices for designing micro-insurance products for small-holder farmers in the tropics. *PLoS ONE*, *7*(6), 1–11. http://doi.org/10.1371/journal.pone.0038281

Norman, J. M., Anderson, M. C., Kustas, W. P., French, A. N., Mecikalski, J., Torn, R., … Tanner, B. C. W. (2003). Remote sensing of surface energy fluxes at 10 1 -m pixel resolutions. *Water Resources Research*, *39*(8), 9–17. http://doi.org/10.1029/2002WR001775

Odenweller, J. B., & Johnson, K. I. (1984). Crop identification using Landsat temporal-spectral profiles. *Remote Sensing of Environment*, *14*(1–3), 39–54. http://doi.org/10.1016/0034-4257(84)90006-3

Olmedo, G. F., Ortega-Farias, S., de la Fuente-Sáiz, D., Fonseca-Luego, D., & Fuentes-Penaililo, F. (2017). *water: Actual Evapotranspiration with Energy Balance Models*. R package version 0.7. Retrieved from http://midraed.github.io/water

Ouedraogo, A. (2014). Ambroise Ouédraogo, directeur du projet de restauration du Lac Bam : « le Lac a perdu actuellement les ¾ de ses ressources en eau ». Retrieved May 28, 2017, from http://lefaso.net/spip.php?article60373

Ouedraogo, R. (2010). The dying Lake Bam in Burkina Faso: Save people or save the lake? (p. 43). Vienna, Austria. Retrieved from www.oeaw.ac.at/vid/download/col100601ro.pdf

Owusu, G. (2017). *Surface Energy Balance and Crop Coefficient Estimation with R*. R package version 3.4.1.

Peña-Barragán, J. M., Ngugi, M. K., Plant, R. E., & Six, J. (2011). Object-based crop identification using multiple vegetation indices, textural features and crop phenology. *Remote Sensing of Environment*, *115*(6), 1301–1316. http://doi.org/10.1016/j.rse.2011.01.009

Pienaar, C. J. (2014). *Increased Water Productivity in irrigated Tomato production in the smallholder farming community of Giyani*. Stellenbosch University.

Platonov, A., Thenkabail, P., Biradar, C. M., Cai, X., Gumma, M., Dheeravath, V., … Isaev, S. (2008). Water productivity mapping (WPM) using landsat ETM+ data for the irrigated croplands of the Syrdarya river basin in Central Asia. *Sensors*.

http://doi.org/10.3390/s8128156

Pôças, I., Paço, T. A., Cunha, M., Andrade, J. A., Silvestre, J., Sousa, A., … Allen, R. (2014). Satellite-based evapotranspiration of a super-intensive olive orchard: Application of METRIC algorithms. *Biosystems Engineering*, *128*, 69–81. http://doi.org/10.1016/j.biosystemseng.2014.06.019

Priestley, C., & Taylor, R. (1972). On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly Weather Review*, (February), 81–92. http://doi.org/10.1175/1520-0493(1972)100<0081:OTAOSH>2.3.CO;2

Rashidi, M., & Gholami, M. (2008). Review of Crop Water Productivity Values for Tomato , Potato , Melon , Watermelon and Cantaloupe in Iran. *International Journal of Agriculture & Biology*, (4), 432–436.

Rodemaker, E. J., & Driese, L. K. (2008). *Modeling Consumptive Water Use in the Green River Basin, Wyoming Using Remotely Sensed Data and the METRIC/SEBAL Model*. University of Wyoming. Laramie, Wyoming.

Rosegrant, M. W., Ringler, C., & Zhu, T. (2009). Water for Agriculture: Maintaining Food Security under Growing Scarcity. *Annual Review of Environment and Resources*, *34*(1), 205–222. http://doi.org/10.1146/annurev.environ.030308.090351

Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1973). Monitoring Vegetation Systems in the Great Okains with ERTS. *Third Earth Resources Technology Satellite-1 Symposium*, *1*, 325–333.

Roy, D. P., Wulder, M. A., Loveland, T. R., C.E., W., Allen, R., Anderson, M. C., … Zhu, Z. (2014). Landsat-8: Science and product vision for terrestrial global change research. *Remote Sensing of Environment*, *145*, 154–172. http://doi.org/10.1016/j.rse.2014.02.001

Sadras, V. O., Grassini, P., & Steduto, P. (2012). Status of water use efficiency of main crops . *SOLAW Background Thematic Report - TR07*, *2013*(October). Retrieved from http://www.fao.org/fileadmin/templates/solaw/files/thematic_reports/TR_07_web.pdf

Sadreghaen, S., Shafieajbishe, R., Rafati, M., & Shahriari, D. (2002). *Technical Evaluations of Micro Irrigation Systems and Assessing their Application in Deficit Irrigation and Compared with Surface Irrigation on Tomato*.

Salack, S., Sarr, B., Sangare, S. K., Ly, M., Sanda, I. S., & Kunstmann, H. (2015). Crop-climate ensemble scenarios to improve risk assessment and resilience in the semi-arid regions of West Africa. *Climate Research*, *65*, 107–121. http://doi.org/10.3354/cr01282

Singh, Y., Rao, S. S., & Regar, P. L. (2010). Deficit irrigation and nitrogen effects on seed cotton yield, water productivity and yield response factor in shallow soils of semi-arid environment. *Agricultural Water Management*, *97*(7), 965–970. http://doi.org/10.1016/j.agwat.2010.01.028

Smith, G. M., & Fuller, R. M. (2001). An integrated approach to land cover classification: An example in the Island of Jersey. *International Journal of Remote Sensing*, *22*(16), 3123–3142. http://doi.org/10.1080/01431160152558288

Su, Z. (2002). The Surface Energy Balance System(SEBS) for estimation of turbulent heat fluxes. *Hydrology and Earth System Sciences*, *6*, 85–99. http://doi.org/10.5194/hess-6-85-2002

Tan, P.-N., Steinbach, M., & Kumar, V. (2006). Classification : Basic Concepts , Decision Trees , and. *Introduction to Data Mining*, *67*(17), 145–205. http://doi.org/10.1016/0022-4405(81)90007-8

Tasumi, M., Allen, R., & Trezza, R. (2008). At-Surface Reflectance and Albedo from Satellite for Operational Calculation of Land Surface Energy Balance. *Journal of Hydrologic Engineering*, *13*(February), 51–63. http://doi.org/10.1061/(ASCE)1084-0699(2008)13:2(51)

Teixeira, A. H. de C. (2011). *Measurements and modelling of evapotranspiration to assess agricultural water productivity in basins with changing land use patterns: a case study in the São Francisco River basin, Brazil. Water.* Wageningen University. http://doi.org/10.1017/CBO9781107415324.004

Topaloğlu, R. H., Sertel, E., & Musaoğlu, N. (2016). Assessment of classification accuracies of Sentinel-2 and Landsat-8 data for land cover/use mapping. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *41*(July), 1055–1059. http://doi.org/10.5194/isprsarchives-XLI-B8-1055-2016

Vanino, S., Nino, P., De Michele, C., Bolognesi, S. F., & Pulighe, G. (2015). Earth Observation for Improving Irrigation Water Management: A Case-study from Apulia Region in Italy. *Agriculture and Agricultural Science Procedia*, *4*, 99–107. http://doi.org/10.1016/j.aaspro.2015.03.012

Vanuytrecht, E., Raes, D., Steduto, P., Hsiao, T. C., Fereres, E., Heng, L. K., … Mejias Moreno, P. (2014). AquaCrop: FAO's crop water productivity and yield response model. *Environmental Modelling and Software*, *62*, 351–360. http://doi.org/10.1016/j.envsoft.2014.08.005

Wang, J., & Bras, R. L. (2009). A model of surface heat fluxes based on the theory of maximum entropy production. *Water Resources Research*, *45*(11). http://doi.org/10.1029/2009WR007900

Wang, K., Li, Z., & Cribb, M. (2006). Estimation of evaporative fraction from a combination of day and night land surface temperatures and NDVI: A new method to determine the Priestley-Taylor parameter. *Remote Sensing of Environment*, *102*(3–4), 293–305. http://doi.org/10.1016/j.rse.2006.02.007

Wardlow, B. D., & Egbert, S. L. (2008). Large-area crop mapping using time-series MODIS 250 m NDVI data: An assessment for the U.S. Central Great Plains. *Remote Sensing of Environment*, *112*(3), 1096–1116. http://doi.org/10.1016/j.rse.2007.07.019

West, C. T., Roncoli, C., & Ouattara, F. (2008). Local perceptions and regional climate trends on the central plateau of Burkina Faso. *Land Degradation and Development*, *19*(January), 242–256. http://doi.org/10.1002/ldr.842

Wiegand, C. L., Gerbermann, A. H., Gallo, K. P., Blad, B. L., & Dusek, D. (1990). Multisite analysis of spectral-biophysical data for corn. *Rem Sens Env*, *33*, 1–16.

Wilcox, J. C., & Sly, W. K. (1976). Use of negative values of potential evapotranspiration in estimation of annual irrigation requirements. *Canadian Journal of Soil Science*, *56*(Nov), 507–509.

Yan, N., & Wu, B. (2014). Integrated spatial-temporal analysis of crop water productivity of winter wheat in Hai Basin. *Agricultural Water Management*, *133*, 24–33.

http://doi.org/10.1016/j.agwat.2013.11.001

Zhang, F., Zhou, G., & Nilsson, C. (2013). Remote estimation of the fraction of absorbed photosynthetically active radiation for a maize canopy in Northeast China. *Journal of Plant Ecology*, *8*(4), 429–435. http://doi.org/10.1093/jpe/rtu027

Zhang, K., Kimball, J. S., & Running, S. W. (2016). A review of remote sensing based actual evapotranspiration estimation. *Wiley Interdisciplinary Reviews: Water*, *3*(6), 834–853. http://doi.org/10.1002/wat2.1168

Zhu, Z., Wang, S., & Woodcock, C. E. (2015). Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4-7, 8, and Sentinel 2 images. *Remote Sensing of Environment*, *159*, 269–277. http://doi.org/10.1016/j.rse.2014.12.014

Zwart, S. J. (2010). *Benchmarking water productivity in agriculture and the scope for improvement: remote sensing modelling from field to global scale*. VSSD. http://doi.org/ISBN: 978-90-6562-237-2

Zwart, S. J., & Bastiaanssen, W. (2004). Review of measured crop water productivity values for irrigated wheat, rice, cotton and maize. *Agricultural Water Management*, *69*(2), 115–133. http://doi.org/10.1016/j.agwat.2004.04.007

APPENDICES

APPENDIX A

Summary of existing major satellite-based ET mapping methods (K. Zhang, Kimball, and Running, 2016)

Model	Advantages	Assumptions/Limitations	ET partition	Input parameters
One-source SEB; One-source spatial variability SEB	Simple, low requirement for metrological data	• Only available for clear-sky; • requires parameterisation of excessive resistance and local calibration; • susceptible to T_s and T_a errors; • requires scaling of instantaneous to daily values	No partition	Rn, Ts. Ta
Two-source SEB	Low requirement for meteorological forcing	• Only available for clear-sky; • high sensitivity to surface temperature errors; • requires scaling of instantaneous to daily values	Soil and vegetation components	Rn, Ts. Ta
Two-source time differencing SEB	Low requirements for meteorological forcing; reduced sensitivity to absolute $T_s - T_a$ difference; does not need local calibration	• Only available for clear-sky; • high sensitivity to surface temperature errors; • requires scaling of instantaneous to daily values	Soil and vegetation components	Rn, Ts
Ts – VI methods	Low sensitivity to T_s errors; low metrological data requirement	• Only available for clear-sky; • relationships derived from $T_s - VI$ space is oversimplified; • requires scaling of instantaneous to daily values	Soil and vegetation components	Rn, Ts, VI
PM models	Process-based; physically sound; temporally continuous coverage; flexible time step; no or low requirements for surface temperature	• High meteorological forcing requirements; • simplified or semi-empirical estimate of canopy conductance	Soil, vegetation, and/or open water components	Rn/radiations, Ta, air vapour pressureand LAI/VI

Model	Advantages	Assumptions/Limitations	ET partition	Input parameters
PT models	Simple; moderate requirements for meteorological forcing	• Many simplifications of physical processes; • requires ground heat flux as an input or assumes that it is negligible; • applied on a monthly time scale	Soil and vegetation components	Rn, Ta, albedo, water vapour pressure
MEP model	Low requirement for meteorological forcing	Requires continuous land surface temperature to produce continuous ET record	Soil, vegetation, and/or open water components	Rn, Ts, qs
Water balance methods	Simple and easy to be applied	• Cannot directly derive gridded evaporation values; • has coarse spatiotemporal resolution; • sensitive to precipitation data error	No partition	P, runoff and ΔS_w
Other water carbon linkage methods	Considering the linkage between carbon and water fluxes	• Can have high requirements for forcing data and be impacted by data gaps and errors but differ in different models; • empirical carbon – water relationships may be used	Usually partition into soil and vegetation components	Varies among different models
Empirical models	Simple, easy to apply	• Requires calibration; • degraded capabilities outside of calibration area; oversimplification of physical processes; • subject to weather conditions if land surface temperature is required	Usually do not partition E	Varies among different models.

R_n: net radiation, T_s: surface temperature, T_a: air temperature, VI: vegetation indices, LAI: leaf area index, ΔS_w: change in water storage

APPENDIX B

| TigerBridge Project: |
| Questionnaire for Irrigated Agriculture |

WASCAL
West African
Science Service Centre on
Climate Change
and Adapted Land Use

A. GENERAL INFORMATION

1. Name of village: --

2. Date/time of survey: --

3. Name of village: --

B. CULTIVATED TERRESTRIAL AREA AND

4. Waypoint Numbers --

	Name	Height (cm)	Age (weeks/months)
5. Which crops have been cultivated?	------------------	------------------	------------------
	------------------	------------------	------------------
	------------------	------------------	------------------

6. When was the land Ploughed? --

7. Has any crop been harvested? Yes ☐   No ☐

8. If yes, what & when?   What   When

------------------ ------------------

------------------ ------------------

9. If no, anticipated harvest --

10. What crop(s) did farmer cultivate last --

11. If yes, what was the production volume (tons) or yield   Crop   Production/yield

------------------ ------------------

------------------ ------------------

NOTE --

XIII