An Evaluation of the Impact of Seasonal Land Cover Change on Evapotranspiration Estimates at the Catchment Scale in the Upper Gundar River Basin, Tamil Nadu, India

by

Akash Senthilkumaran

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Master of Science

in

Geography

Waterloo, Ontario, Canada, 2024

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Changes in the water cycle influence the energy balance of the Earth. The water cycle is represented using the water balance equation, in which Evapotranspiration (ET) is a vital parameter. One of the main drivers of the change in ET within a specific area is the change in land cover. This study focuses on estimating ET across the Upper Gundar River Basin located in the state of Tamil Nadu, India. Notable features of this landscape include agriculture throughout the year supported using an extensive network of tanks and borewells, and the presence of *Prosopis juliflora*, a widely prevalent invasive species known to consume groundwater and moisture. Due to the lack of spatial variability in point ET measurements, ET models using remote sensing imagery as the main forcing data have been widely used to assess the spatial variability and temporal variability based on the principle of surface energy balance. These models are collectively referred to as Surface Energy Balance (SEB) models. The model used in our study is the Surface Energy Balance Algorithm for Land (SEBAL) model to estimate ET for two periods of the year, indicating mid-summer and the end of the northeast monsoon for the years 2006, 2014 and 2021. Since land cover changes drive ET, land cover classification and seasonal change detection are also performed for the same time periods. Imagery from Landsat satellites is used, and one image is chosen to represent the specific season. The major land cover classes chosen in our study are water, pre-growth, agriculture, *Prosopis juliflora* (prosopis), barren land, and exposed soil. Along with the Landsat imagery, to run SEBAL, Aster DEM is used along with in-situ weather data and GLDAS data.

Over 90% levels of overall accuracy were achieved for all year-season combinations for the land cover classification. Using SEBAL, Actual Evapotranspiration (ET_a) for all the classes is calculated except the water classes. Due to the lack of in-situ measurements, an intermodal comparison was performed with the EEFlux product available at the same resolution derived using the METRIC algorithm using land cover classes as units of comparison. The comparisons are carried out using correlation coefficient (r), root mean squared error (RMSE), and mean values. Highest mean values were observed for either the agriculture or prosopis class, and the lowest mean value was exhibited by the exposed soil class on all occasions. Within all summers, considering all the years, the average correlation coefficient and RMSE were 0.8, 1.2 mm/day, and for monsoon, the averages were 0.5 and 0.85 mm/day, indicating increased proximity during the monsoon season between SEBAL and EEFlux. Similarly, the range of mean values between classes in summer is 2.12 mm/day, 1.36 mm/day in the monsoon. In terms of the energy fluxes used to determine ET_a, a decrease in monsoon is observed for soil heat flux (G), instantaneous net radiant energy (Rn_inst), and net radiation in a day (Rn_24). For sensible heat flux (H), classes with vegetation tend to have lower values in comparison to the classes without vegetation. Finally, average water outflux is calculated encompassing all classes by multiplying the area of a class with mean ET_a, and the values observed in summer and monsoon alternatively for the years 2006, 2014, and 2021 in m^3/day are 5,142,212, 3,534,906, 2,954,897, 4,046,322, 5,369,191, 4,512,596.

Acknowledgements

I would like to express immense gratitude to my supervisor, Dr. Richard Kelly, for giving me this opportunity, being a pillar of support, and always asking me the right questions. I feel blessed and grateful. I would like to thank my committee members, Dr. Grant Gunn and Dr. Peter Crank, for taking the time out of their busy schedules to be a part of my committee. I would like also to thank Dr. Chris Fletcher, Dr. Merrin Macrae, Dr. Linlin Xu for their support during the coursework. Many thanks to Vicky for her timely insights and valuable suggestions throughout the course of my program. A heartfelt thank you goes out to my friends in the cryo lab and the TAEOS group for their unwavering support. I express my gratitude to DHAN Foundation, our NGO partners, for their support during the field season in December 2022. Special appreciation is owed to the Water Resources Department, Groundwater Division, for providing me with the crucial data necessary for analysis. I extend my thanks to Dr. Sajid Pareeth, Bich Tran, and Amir Owlia from IHE Delft for their constant support and assistance throughout the project. I would like to acknowledge my undergraduate professors who played an integral role in helping me reach this point – Dr. Gandhimathi, Dr. Nisha Radhakrishan, Dr. Jeevan Joseph, and Dr. Lelitha Devi. Also, thanks to Dr. Mourad Ben Amor from University of Sherbrooke for providing me with my first opportunity in Canada. Additionally, I extend my thanks to Mitacs for their support since 2018. I want to express my deepest gratitude to my parents, for whom words don't suffice. I dedicate this thesis for humanity, which has endured a pandemic in the recent past. We will always win.

Table of Contents

List of Figures

List of Tables

1 Introduction

1.1 Motivation

Water is an important natural resource, driving the sustenance of all forms of life on earth. The flow of water through the earth systems is called the water cycle. Evaporation, condensation, groundwater flow, infiltration, percolation, precipitation, run-off, sublimation, and transpiration are some of the physicochemical components of the global water cycle along with anthropogenic interactions such as building water storage structures, or extractive processes like groundwater pumping etc (Allan et al., 2020; Abbott et al., 2019). To understand the global water cycle, precise quantification of global water fluxes and storage is required (Vargas Godoy et al., 2021). Traditionally, the global water balance has been represented through combining precipitation, runoff, evaporation, and transpiration to a change in water storage. Under a warming climate, the water cycle is expected to intensify in the form of increased precipitation and evaporation (Huntington, 2006; Lal, 2001).

When it comes to potable water, the global hydrological cycle's renewal processes control the amount of water that is available. Due to the predicted rise in temperatures in the future, countries in south Asia are likely to face acute water stress. Along with climate change, the increasing population will lead to increased per-capita demand and competition for water resources across industrial, agricultural, and domestic activities (Lal,2001).

Rivers have been the major sources of water for anthropogenic consumption since time immemorial. Historically, the birth and evolution of human civilizations occurred in regions close to rivers. India is a country with an extensive network of rivers with 37 major rivers spread across the landmass (Alagh, Pangare, & Gujja, 2006). This gives rise to a wide number of river basins, which house significant populations reliant on the resources present in the landscape for their daily livelihood. In the broader context, in the 21st century, in India, where there is a growing population, management of water resources plays a crucial role in ensuring adequate access to water for a variety of domestic and industrial activities. With growing industrialization, the water demand is set to increase in the forthcoming years. Along with industrialization, to feed the growing population, the agricultural water demand is expected to increase (Ahmad et al., 2006).

In this study, we focus on the Gundar River Basin located in the state of Tamil Nadu, India. Tamil Nadu has 17 major river basins. Most of the water resources prevalent in the river basins are diverted to agriculture, estimated to consume about 75% which is very similar to the global average of 70% making agriculture one of the top consumers of fresh water (Ahmad et al., 2006; Palanisami et al., 2011). But most of the river basins are water stressed and interlinking of rivers is a widely popular solution proposed to tackle the issue in hand (Dhawan, 2017). Amidst the water demand from agricultural, industrial, and domestic activities, there is also competition for surface water resources between the neighboring states (Palanisami et al., 2011). Thus, efficient planning and allocation of water resources are quintessential to maintaining the equilibrium between water demand and supply within river basins. In the broader context, India in the 21st century has a growing population and management of water resources plays a crucial role in

ensuring adequate access to water for a variety of domestic and industrial activities. With growing industrialization, the water demand is set to increase in the forthcoming years. Along with industrialization, to feed the growing population, the agricultural water demand is expected to increase (Ahmad et al., 2006).

Evapotranspiration (ET) defined at the plant-soil interface is the sum of evaporation from the land surface and transpiration from the stomatal pores present in the plant (Papadavid et al., 2017). Various atmospheric, soil, and biophysical processes influence ET at this interface, thereby playing a pivotal role in determining the land-atmosphere energy balance and boundary layer dynamics (Kundu et al., 2018; Schaller & Fan, 2009). In terrestrial water budgets, evapotranspiration (ET) is the second-largest term after precipitation (Glenn et al., 2017). Hence, long-term ET estimates can be considered as a reliable indicator for gauging the intensity of the water cycle (Huntington, 2006).

The absence of reliable ET estimates poses challenges in understanding the critical earth system interactions and developing effective adaptation strategies for food and water security (Soni & Syed, 2021). ET rates directly influence crop water requirements, thereby impacting the planning and allocation of water resources. Thus, anticipating changes in ET rates within the cropping period provides vital information for crop planning. In India, the contribution of surface water and groundwater resources for irrigation has been crucial to achieving food self-sufficiency. However, meeting future water demands necessitates not only exploring alternative water storage sources but also establishing robust frameworks to ensure judicious water usage. This

underscores the importance of obtaining large-scale, basin-wide estimates of ET to fulfill crop water requirements and conserve terrestrial water sources (Pathak et al., 2014). Accurate and basin-wide ET estimates are also essential for optimizing agricultural production, predicting floods and droughts, managing land use, and allocating water resources efficiently (Soni & Syed, 2021).

In recent years, it has been observed that climate change and human-induced Land Use and Land Cover (LULC) change have both had an increasing impact on ET (Li et al., 2017). Additionally, the influence of changing land use and land cover on other hydrological processes like soil moisture storage and groundwater recharge is widely acknowledged (Zhang & Schilling, 2006). In the context of this study, understanding the impacts of land use and land cover (LULC) change on evapotranspiration (ET) becomes crucial, as it is a key component that influences ecosystem services, water resource management leading to effective land use planning (Li et al., 2017). On a regional scale, LULC change primarily affects ET through changes in vegetation, agricultural activities, and urbanization (Li et al., 2017). Different land cover types exhibit distinct hydrological characteristics, with vegetative areas transpiring throughout the year, while agricultural lands transpire mainly during the cropping period. Barren land and urban areas generally have lower ET rates compared to land cover types with vegetation (Zhang & Schilling, 2006).

In India, ensuring food security post-independence saw the Green Revolution, which significantly contributed to self-sustenance. While this led to remarkable agricultural yield improvements, it also resulted in declining water tables in several parts of the country (Dhawan, 2017). Since this

process involved the conversion of forest land into agricultural land, considerable shifts in rainfall patterns and surface temperatures were observed. This emphasizes the potential impact of LULC change on ET rates, as higher rates are observed in warmer weather compared to cooler conditions (Singh & Singh, 2023).

Variations in ET across space, time, and land cover classes are critical indicators of water use sufficiency, dependability, and equity (Papadavid et al., 2017). ET is determined by meteorological variables, and alterations in these variables due to climate change will exert a notable influence on ET values (Darshana et al., 2013; Papadavid et al., 2017). These shifts could potentially affect agricultural water requirements (Papadavid et al., 2017).

Conventional estimates of ET typically rely on weather data collected at daily intervals. Some of the methods that are widely used in estimating ET are: FAO-56 Penman-Monteith, the FAO-24 Radiation, FAO-24 Blaney Criddle, 1985 Hargreaves, Priestley-Taylor, 1957 Makkink, and 1961 Turc (Gao et al., 2017). In-situ estimates of ET are primarily obtained through laborious and expensive methods such as lysimeters and eddy covariance measurements (Huntington, 2006; Kundu et al., 2018). However, these point estimates fall short in capturing the spatial variability, which is particularly significant in diverse landscapes like river basins encompassing various land cover classes (Ahmad et al., 2006; Li & Lyons, 1999).

The use of remotely sensed data for ET estimation has gained increased popularity in recent years, marking considerable progress in this field (Jana et al., 2016). Various hydrological, environmental, and agricultural studies demand continuous ET measurements at different scales,

as it serves as a crucial indicator of interactions at the plant-soil interface (Bala et al., 2016). Thus, accurate ET estimates are essential for conducting a comprehensive analysis of water balance components at a basin scale. Remote sensing data, available in various intervals (continuous, periodic), facilitates a deeper understanding of the hydrological attributes of the basin (Pal et al., 2018). In the realm of remote sensing and climate modeling, a 1° grid may encompass a range of land cover types, making a single point measurement within the grid less representative. Therefore, fine-scale measurements, such as those obtained at Landsat's 30-meter scale, are preferred. Since estimating ET for large areas involves multiple parameters, such as surface temperature, vegetation indices, and surface albedo, remote sensing emerges as the most reliable technique for obtaining these parameters consistently. Moreover, it is also cost-effective with many data sets available as, primarily relying on open-source data (Li & Lyons, 1999; Glenn et al., 2007). Accurate ET estimates are indispensable for various management tasks at local to regional scales, including weather forecasting, irrigation scheduling, watershed management, and anticipating the long-term impacts of land use change and global climate change (Glenn et al., 2007).

The National Water Mission report (2017) designates the Gundar river basin (GRB) as a major river basin in the state of Tamil Nadu, India, covering an area of 5690 sq. km. Supporting a population of 1.9 million, the basin plays a vital role in the livelihoods of the region's inhabitants. The river stretches approximately 150 km, supplying water for domestic, agricultural, and industrial purposes. As such, maintaining a continuous and ample water supply is crucial for

sustaining industries and agriculture, necessitating a comprehensive understanding of the basin's water balance.

Agriculture is pervasive across the region, with a diverse range of crops including paddy, chillies, coconut, cotton, sugarcane, and millets. Notably, paddy constitutes around 70% of the gross irrigated area in the basin. Along with the crop water requirements, the water demand of the basin is dependent on factors such as land holdings, land cover patterns, irrigation practices, and industrial employment (National Water Mission, 2017).

The basin is characterized by a network of water storage tank cascades that shape its hydrology. In arid and semi-arid regions, where rainfall is limited and significantly lower than evaporative demand, local water storage structures and groundwater are essential for irrigation (Ahmad et al., 2006). Consequently, the presence of these water storage structures is expected to alter the hydrological cycle in the basin, a pattern observed among farmers in the GRB.

A distinctive feature of the basin is the widespread presence of the invasive species, *Prosopis juliflora (prosopis*), known for its drought tolerance and rapid growth. Throughout this study, it will be referred to simply as "prosopis".

This species is known to consume groundwater and moisture, particularly in barren areas, and is found in both small and large patches (Shiferaw et al., 2021; Vanthof & Kelly, 2017). Its impact on the hydrology of the region is unknown and this study would serve as a starting step in that process.

Prosopis manifests itself in the form of trees. Such trees are observed growing inside the tanks and on the tank bunds, with larger patches spanning multiple acres seen in barren lands. The presence of tanks significantly influences agriculture and the sustenance of prosopis. Tanks serve as sources of surface water and recharge spots for groundwater, making them ideal for the growth of prosopis. Hence, examining ET rates in proximity to tanks and regions of high prosopis growth becomes imperative.

On one hand, prosopis is known to impact agriculture by inhibiting the flow of tank water to the fields and depleting the groundwater table by consuming water for its survival. On the other hand, people rely on firewood and charcoal generated from prosopis as an important source of income (Sato, 2013). This stresses the importance of the species in the landscape and the complexity of its interactions with the landscape. Understanding the spatio-temporal characteristics of propopis in the region would pave way for a deeper understanding of the behavior of the species.

With ET being a dominant component in the overall water balance, precise measurements are essential for water resource planning and allocation (Ahmad et al., 2006). Obtaining fine-scale ET estimates throughout the basin is paramount, as ET primarily dictates agricultural water demand. Combining the land cover maps with ET maps would allow us to quantify the changes that have happened between the years with respect to ET patterns which can be used as a reference to plan for optimal water resources management in the region and in identifying regions of

importance decided based on the magnitude of ET. This is along the lines of developing a robust framework resulting in optimal and sustainable practices in agriculture to minimize water losses.

1.2 Overall aim and objectives

The overall aim of this thesis is to quantify the representative net water transport via evapotranspiration across specific land cover classes during mid-summer and the end of monsoon in 2006, 2014, and 2021. To achieve this aim, three specific objectives are set:

- Conduct land cover classification and seasonal change detection analysis for the years 2006, 2014, and 2021
- Utilize the Surface Energy Balance Algorithm for Land (SEBAL) , a remote sensing-based method, to estimate actual evapotranspiration (ET_a) at a pixel level using selected images from the years 2006, 2014, and 2021
- Perform a comparative assessment between SEBAL and the EEFlux ET product available at the same resolution derived using the METRIC algorithm

2 Literature review

2.1 Gundar River Basin

2.1.1 Overview

Gundar River Basin (GRB) is one of the major river basins of Tamil Nadu, India, with a drainage area of 5690 sq. km. The location of the basin is between Latitude 9°05’ – 10° 03’ and Longitude 77°35’ – 78°35’. The basin is inundated seasonally by the Northeast Monsoon, impacting parts of Madurai, Sivagangai, Virudhunagar, Ramanathapuram, and Thoothukudi districts, where almost two million people live. The gross irrigated area of the GRB is 105,841 Ha, with paddy fields accounting for around 66 per cent of the land cover. The river has five tributaries consisting of forty-two check dams constructed across them. The check dams divert the water to irrigate an area of 14,388 ha through the small reservoirs (tanks) in the basin. 2,276 tanks serve an irrigation area of around 7,200 ha (National Water Mission, 2017).

Figure 2.1 Paddy fields in the Gundar river basin (Photo credit: Akash Senthilkumaran, 2022)

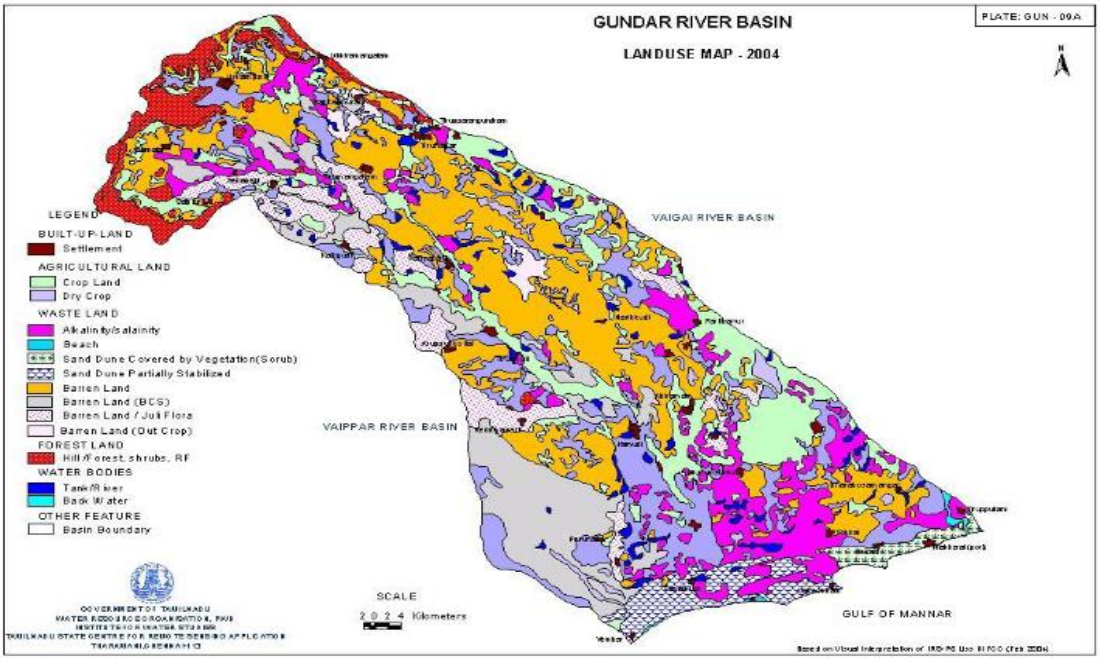

Figure 2.2 Land use map of the Gundar River Basin, 2004 (Source: National Water Mission, 2017)

The northwestern parts of the basin are occupied by the Varshanadu hills, with peaks exceeding 1000m from the mean sea level. The northeastern parts are home to Nagamalai hills, the elevation of which ranges between 200m to 400m above the mean sea level. Most parts of the basin are relatively low lying, and in an overall sense, the river basin has a gentle slope towards the Indian Ocean. The origin of the Gundar river is found at the Varshanadu hills, and it has five tributaries - Goundanadhi, Gridhamal river, Kanalodai, Paralaiaru, and Therkkar river. The river is non-perennial in nature (National Water Mission report, 2017).

2.1.2 Tanks in the Gundar River Basin and their Importance

In India’s semi-arid regions, approximately 120,000 small-scale tanks cover around 4.12 million hectares of land (Palanisami & Thangavel, 2020). These tanks serve various purposes, including irrigation, recreation, livestock, and domestic use. Primarily located in Tamil Nadu, Karnataka, and Andhra Pradesh, they account for roughly 60% of India's tank-irrigated area. Tamil Nadu alone houses approximately 38,949 tanks, classified as Public Works Department (PWD) tanks with a command area exceeding 40 hectares and Panchayat Union (PU) tanks with a command area of less than 40 hectares (National Water Mission report, 2017; Palanisami & Thangavel, 2020). Tanks are categorized based on their water sources: partially filled tanks, known as system tanks, rely on rivers and reservoirs, while those solely dependent on rainfall are termed non-system tanks. Typically, these tanks begin to fill after the Northeast Monsoon sets in (Palanisami & Thangavel, 2020).

In the GRB, a distinctive feature stands out among the tanks: tank cascades. These are interconnected water storage structures forming networks that allow water to flow from upstream tanks to downstream ones through gravity (Srivastava & Chinnasamy, 2021). Farmers often claim the historical presence of these structures, purportedly constructed to combat seasonal water scarcity by storing and aiding groundwater recharge. Consequently, these cascades wield significant influence over the river basin's hydrology.

Figure 2.3 Mudukkangulam tank located inside the river basin (Photo credit: Akash Senthilkumaran, 2022)

Storage of tank water occurs solely during the rainy season, with both tank management and climatic conditions exerting notable impacts on water volume (Sato & Duraiyappan, 2011). Over the years, the GRB has witnessed substantial changes in agricultural practices, leading to land use and land cover (LULC) changes that emphasize the critical need to optimize water usage and maintain tanks—upon which the region heavily relies (National Water Mission report, 2017; Twisa & Buchroithner, 2019).

2.1.3 Agricultural patterns

The river basin is a dry, semi-arid region largely reliant on monsoon rainfall to replenish water resources (National Water Mission report, 2017; Sato, 2013). As per the Koppen-Geiger scheme, the region can be classified as having an equatorial winter dry climate (Aw) (Kottek et al., 2006). Paddy is the predominant crop cultivated in this region. Other crops grown in this region include

chillies, sugarcane, coconut, and market garden vegetables. Irrigation heavily relied on tank cascades due to water stored in the tanks. A tank irrigation system involves agricultural fields and linked tanks, drawing water from rainfall in the catchment area and canals connected to rivers (Sato & Duraiyappan, 2011).

Since the 1990s, there has been a surge in cash crop cultivation in the upper basin, prompting the drilling of new wells and the deepening of existing ones. However, a similar trend in cultivation has not emerged in the lower basin, where farmers have not updated or installed new wells. Hence, the upper basin witnessed the introduction of new wells and the deepening of existing ones, leading to an income gap between farmers in the upper and lower basins (Sato & Duraiyappan, 2011).

2.1.4 Vegetation patterns in the basin

The income disparities between upper and lower regions within the basin resulted in the abandonment of arable lands in the latter due to diminished prospects for growing cash crops. Consequently, these lands became invaded and proliferated with Prosopis *juliflora* (Prosopis) (Sato, 2013). In response to meager monsoon rainfall affecting seasonal agriculture, Tamil Nadu introduced Prosopis in the early 1960s to provide an alternative income source, primarily through charcoal and firewood. *Prosopis juliflora*'s notable trait is its capacity to thrive in dry conditions, leading to its widespread growth across the basin. It appears in various forms: small plants on barren lands, alongside other vegetation in villages, and as extensive patches covering acres of adjacent dry barren lands.

Figure 2.4 Small prosopis trees and a pile of burnt prosopis used to produce charcoal (Photo credit: Akash Senthilkumaran, 2022)

Figure 2.5 Large prosopis patches found growing in vast swathes of barren land (Photo credit: Akash Senthilkumaran, 2022)

2.1.5 Monsoon precipitation

India relies heavily on monsoon-induced precipitation, experiencing two main monsoon seasons: the Northeast monsoon and the Southwest monsoon. The Southwest monsoon, also called the Indian Summer Monsoon (ISM), contributes to 75-80% of the country's annual rainfall, while the Northeast Monsoon covers the rest. The ISM largely affects eastern, northern, and western India, with the Northeast monsoon influencing the remaining areas (Misra & Bhardwaj, 2019; Maharana & Dimri, 2019). Typical monsoon climatic conditions, in comparison to summer, involve lower temperatures, higher precipitation rates, and fewer sunshine hours. Both monsoons significantly impact India's economy and hydro-climatological features. Research by Gadgil and Gadgil (2006) emphasized the monsoon's influence on GDP and agriculture from 1951 to 2000, noting that failed monsoons adversely affect food grain production, while surplus precipitation has a smaller impact on the nation's food grain production and GDP.

Tamil Nadu, the eleventh-largest state in India with a population of nearly 72 million, 80% literacy rate, and a population density of 555 people per square kilometer, experiences a tropical climate with distinct summer and winter temperatures. The summer months (April-June) witness temperatures above 40°C, whereas November-February marks lower temperatures, usually around 20°C (Varadan et al., 2017). The state's primary rainy season is the Northeast Monsoon, contributing significantly to its annual precipitation. Inland districts receive around 40% of their yearly rainfall from the Northeast Monsoon, while coastal districts receive approximately 60% of their overall share (Samui, Balasubramanian, & Kamble, 2013). Any deviation from average

rainfall during the southwest or Northeast Monsoons greatly affects agriculture and puts strain on Tamil Nadu's key irrigation sources—tanks, canals, and wells (Samui et al., 2013).

Tamil Nadu's physiography comprises seven agroclimatic zones: North-Eastern, North-Western, Western, High Altitude, Cauvery Delta, Southern, and High Rainfall Zones. The city of Madurai, situated closest to the GRB (Figure 3.2), falls under the Southern Zone, along with Thoothukudi, Virudhunagar, Tirunelveli, Ramanathapuram, Theni, Dindigul, and Sivagangai districts, five of which are part of the GRB. The region experiences four distinct seasons of precipitation annually: winter (January–February), Summer (March–May), Southwest monsoon (June–September), and Northeast monsoon (October–December). Based on weather records from 1969 to 2010, this zone stands as the driest in Tamil Nadu, receiving about half of its annual rainfall from the Northeast Monsoon (Varadan et al., 2017).

2.2 Evapotranspiration (ET)

2.2.1 Overview

The transfer of water from the land surface to the atmosphere constitutes terrestrial evapotranspiration (ET). ET also reflects the atmosphere's water demand for surface water sources which is fulfilled by the soil's transport of water obtained through precipitation, irrigation, and groundwater recharge processes. Throughout this process, water undergoes a phase change from liquid to gas, by absorbing energy, thus leading to the cooling of land surface. Evapotranspiration is a combination of evaporation and transpiration (Figure 2.6) (Krishna, 2019).

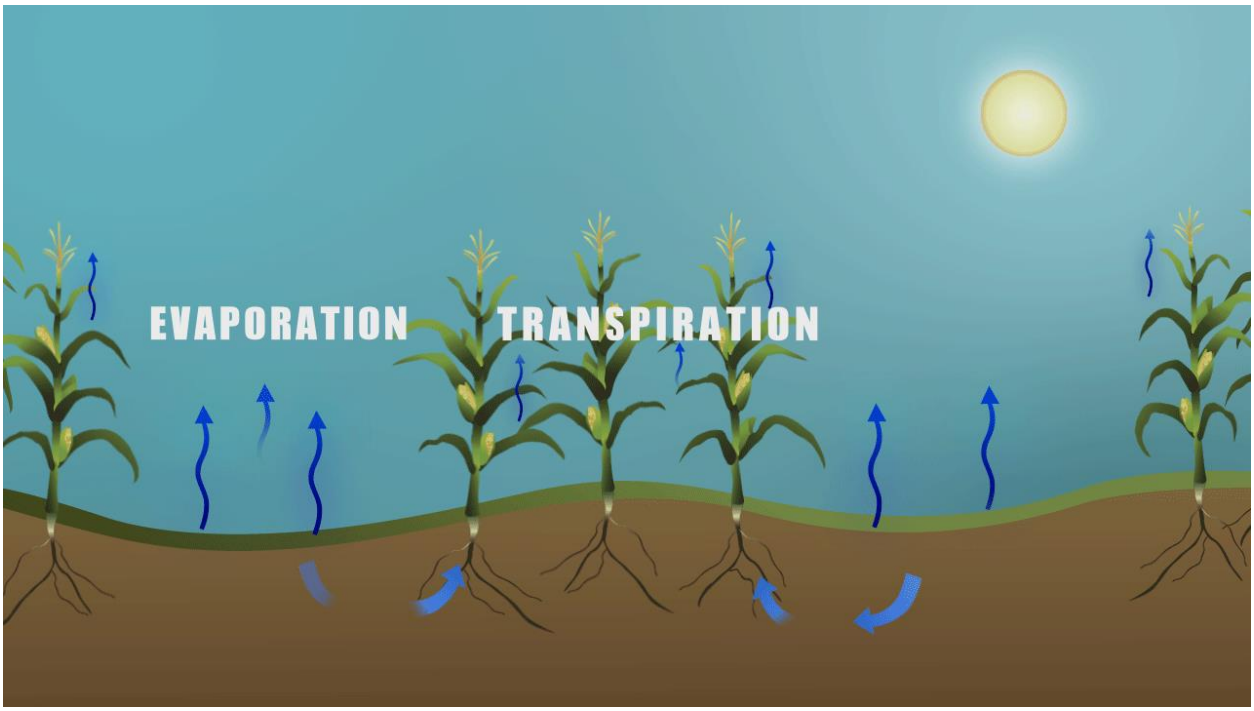

Figure 2.6 Schematic representation of Evapotranspiration (ET) at the plant-soil interface (Source: NASA SVS, 2012)

Evaporation is the process of conversion of liquid water into water vapor and primarily occurs land surfaces and transpiration involves the same happening in plants. Transpiration is the part of ET involving plant interaction with the atmospheric water cycle. Water moves from soil through plants to crop root cells based on potential gradients. It transpires into the atmosphere from various plant parts like leaves, stems, flowers, or roots. On the other hand, evaporation depends on solar radiation reaching the soil surface, influenced by canopy orientation. Initially, soil evaporation is the primary source of moisture loss in early crop growth. However, once the crop is fully grown, transpiration dominates. Different crops contribute differently to ET: nearly all is from evaporation at sowing, but over 90% is from transpiration at full crop cover (Krishna, 2019; Allen et al., 1998).

2.2.2 Types of Evapotranspiration

Reference evapotranspiration (ET_o) and potential evapotranspiration (PET) are both estimates of the atmospheric water demand. Thornthwaite (1948) introduced PET, defining it as the water that evaporates from areas with enough moisture-covered vegetation. It includes water evaporating from soil and plants, moving water from Earth back into the atmosphere, opposite to precipitation.

In 1963, the World Meteorological Organization defined PET as “the amount of water vapor emitted by a pure water surface, per unit area and time, under current atmospheric conditions” (World Meteorological Organization, 1963). Jensen (1968) later saw PET as “the rate of evapotranspiration for a given crop at a given stage of growth when water is not limiting and other factors such as insects, diseases, and nutrients have not materially restricted plant development”. However, this definition faced challenges due to the lack of global crop data, which restricted its applicability across different environments and regions worldwide (Xiang et al., 2020).

To alleviate the need to calibrate for different crops growing in different parts of the world, (Allen et al., 1998) defined the term “reference crop evapotranspiration” aka “reference evapotranspiration” and defined it as the “the rate of evapotranspiration from a hypothetic crop with an assumed crop height (12 cm) and a fixed surface resistance (70 s/m) and albedo (0.23) which would closely resemble evapotranspiration from an extensive surface of green grass cover of uniform height, actively growing, completely shading the ground and not short of water”. This

is termed as the FAO56 ET_o. Using a reference surface, such as with a constant surface resistance, helps stabilize fluctuating climatic variables. Meteorological factors like wind speed, humidity, air temperature, and radiation influence reference evapotranspiration (Allen et al., 1998). The American Society of Civil Engineers (ASCE) modified this definition, including alfalfa with a height of 0.5m as the reference crop, termed the ASCE ET_o (Xiang et al., 2020).

Since ET_o aims to differentiate plant ET across broad regions, the definition is clearer than that of PET. ET_o signifies the atmosphere's evaporative demand, varying by time and location and relying solely on climate parameters mentioned earlier. Soil and crop characteristics do not influence ET_o estimation (Allen et al., 1998).

ET_c, crop evapotranspiration under standard conditions, is derived by multiplying ET_o by K_c, the crop coefficient. K_c accounts for crop-specific traits distinguishing field crops from the reference crop. ET_c differs from ET_o in similar climates due to changes in stomatal characteristics, aerodynamic properties, surface albedo, and leaf anatomy (Allen et al., 1998).

The FAO56 modified PM ET_o equation (Allen et al., 1998) stands as the gold standard for estimating ET_o due to its strong theoretical base and detailed model description. The equation is given below (Equation 2.1):

$$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma u_2(e_S - e_a) \times \frac{900}{(T + 273)}}{\Delta + \gamma(1 + 0.34u_2)} \quad 2.1$$

Where, R_n stands for net radiation at the crop surface, G for soil heat flux, T for air temperature in ^oC at 2m height above the ground surface, u_2 wind speed at a height of 2m above the ground, e_s for saturation vapor pressure in kPa, e_a for actual vapor pressure in kPa, Δ for slope of vapour pressure curve in kPa $°C^{-1}$, γ psychrometric constant in kPa $°C^{-1}$. To compute ET_o values in data scarce regions, alternative methods like Priestly-Taylor, Turc, and Hargreaves can be used with minimal number of parameters (Bhimala, Patra, S, & Goroshi, 2023).

ET_a, actual evapotranspiration is the estimate that represents the true physical quantity of evapotranspiration at the surface (Allen et al., 1998). Rainfall, soil moisture, surface net solar radiation, temperature, humidity, wind speed, vegetation type, leaf area, stomatal conductance, and atmospheric CO_2 can be considered as the main factors that determine actual evapotranspiration. But the strength of influence of these factors are also dependent on regional conditions, seasonal variations in climate and large-scale atmospheric circulations too. (Bhimala et al., 2023). The uncertainty in estimating these factors precisely makes the estimation of ET_a extremely challenging. In terms of a river basin, basin wide estimates of ET_a would represent the actual water loss in the landscape which becomes more significant in cropped areas to estimate agricultural water losses helping in water resource planning.

2.2.3 Evapotranspiration Trends in India and Tamil Nadu

With extensive evidence of global warming, both reference evapotranspiration (ET_o) and actual evapotranspiration (ET_a) are expected to undergo significant impacts. The Indian subcontinent is highly susceptible to climate change. This eventually influences the water balance at varying

scales, leading to significant impacts on the agrarian economy (Bandyopadhyay et al., 2009). Therefore, analyzing the distinct impacts on river basins becomes essential, instead of generalizing the phenomenon. Evapotranspiration (ET), a nonlinear occurrence, relies on parameters like temperature, wind speed, radiation, humidity, crop type, and growth stage indicating a direct link between climate change and ET impacts (Bandyopadhyay et al., 2009).

Chattopadhyay & Hulme (1997) analyzed data from ten stations across India between 1976 and 1990 and reported a decreasing ET_o trend. Additionally, Goyal (2004) suggested a modest rise in ET_o due to climate change in Rajasthan, an arid region, based on their analysis between 1971 and 2002, supporting the impact of global warming. Bandyopadhyay et al. (2009) also examined ET_o trends between 1971 and 2002 and noted a substantial decrease attributed to reduced wind speed and increased humidity nationwide.

The process of ET involves phase transition of water and decreasing water vapor concentration in the atmosphere, involving both energy and mass transfer (Madhu et al., 2015). However, during drought years, ET could adversely affect plant growth due to existing water stress conditions. Analysis between 1901 and 2007 by (Madhu et al., 2015) showcased significant ET increases linked to temperature variation over the region, especially in drought years.

India's highly variable climate and vegetation dynamics have been primarily associated with temperature fluctuations and rainfall impacting ET variability. Enhancing understanding of ET dynamics and their correlation with climate variables remains a pivotal aspect for future projections (Goroshi et al., 2017). (Goroshi et al., 2017) indicated varying ET trends across seasons

and regions in India, reflecting increasing trends in certain seasons and locations while showing declines in others.

The estimation of crop water requirements is influenced by local conditions and crop characteristics. Furthermore, the scarcity and inadequate maintenance of weather stations in Tamil Nadu pose challenges, leading to data gaps affecting ET_o estimation, given its reliance on crucial variables (Mohan, 1991). Mohan's (1991) early study evaluating different ET_o estimation methods across weather stations noted varying but closely related values, with the Blaney-Criddle method exhibiting the highest correlation.

In the Kumbakonam region, crucial for rice cultivation in Tamil Nadu, agriculture relies significantly on the Cauvery River. Hari et al. (2021) compared crop ET estimates derived at a field scale using a linear regression model created using predefined crop coefficient values (K_c) and NDVI for paddy at five locations located inside the region from 2016-2019 with estimates from METRIC derived using satellite imagery. A good level of correspondence between the methods was observed with an R^2 value of 0.72 and an RMSE of 0.67 mm/day.

In continuation with the existing studies done in state of Tamil Nadu, this study which focusses on the Gundar River Basin, a unique landscape present in the state would further add to the steps taken to understand the spatial and temporal variations of ET within the state.

2.2.4 Remote sensing of Evapotranspiration

Over the past two to three decades, satellite retrieval of ET has gained popularity as a method and subject of study due to its high cost-effectiveness, large and repeatable coverage, and generally good accuracy (Zhang et al., 2016). Some of the popular methods for which remote sensing has been utilized to provide for forcing data are: Surface energy balance (SEB) methods, Penman-Monteith methods, Priestley-Taylor Methods, Water-Carbon linkage methods, Water balance methods, MEP methods, Empirical methods, Ts – VI methods (Glenn et al., 2007; Zhang et al., 2016) The evolution of these methods over the years is described in Figure 2.7.

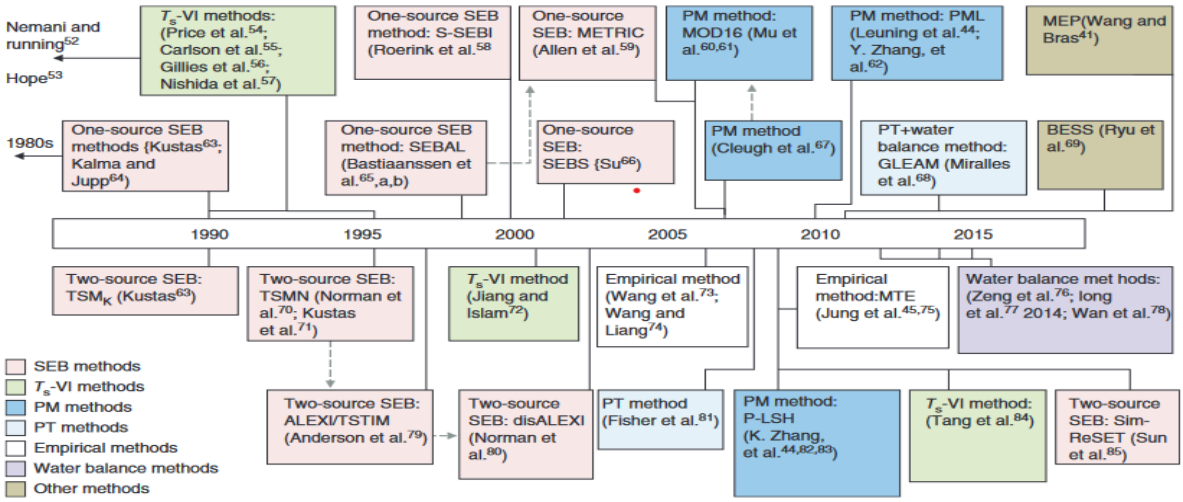

Figure 2.7 Evolution of ET methods involving the use of remote sensing data (Source: Zhang et al., 2016)

This study focusses on using methods based on the surface energy balance as they are more comprehensive in terms of involving multiple variables to estimate ET whereas the other methods work on using variables based on a few established relationships.

Since their development in the 1970s, surface energy balance (SEB) methods have been applied in various studies to estimate evapotranspiration (ET) in different global regions. These methods focus on estimating the latent heat of evapotranspiration (λET) after computing other flux components related to the surface energy balance, including the net radiation flux at the surface (R_n), soil heat flux (G), and sensible heat flux (H). These components are described in detail in Section 4.2.1 (Glenn et al., 2007; Zhang et al., 2016).

SEB models are divided into two classes – one source and two source models. The difference comes with respect to the treatment of the land surface chiefly parameterized using the term radiometric- convective resistance (r_r). In remote sensing-based ET models, r_r is estimated by treating the land surface like an electric circuit with the rate of heat transfer between two points above the land surface dependent on the potential difference (the temperature gradient) and resistance terms that are dependent on the internal properties of vegetation, land, and the atmosphere. Different approaches to the resistance networks allow us to derive various resistance network designs and to estimate r_r. As shown in Figure 2.8, the surface is treated as one single entity in one source models (2.8a) and is divided into two layers – soil surface and canopy in two source models (Zhang et al., 2016).

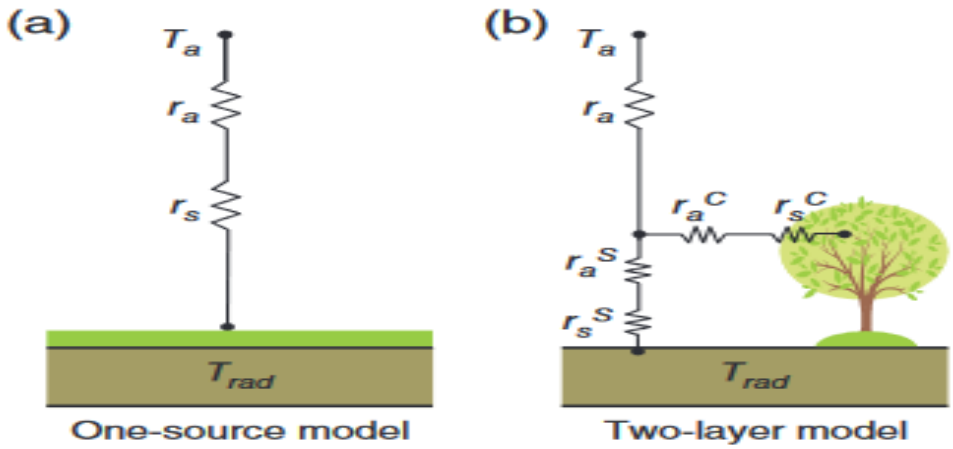

Figure 2.8 Aerodynamic resistance in one source (2.8a) and two source models (2.8b) (Source: Zhang et al., 2016)

In Figure 2.8, T_a refers to air temperature, r_s refers to the bulk surface resistance of the flow of water vapour through the plant-soil interface and r_a refers to aerodynamic resistance, the resistance term pertaining to the flow of heat and water vapour from the evaporating surface into the air. In a two-layer model, r_a^S, r_s^S refers to the same components pertaining to the soil surface and r_a^C, r_s^C relate to the canopy. r_a and r_s are combined to estimate r_r (Allen et al., 1998; Zhang et al., 2016).

In one-source models, independent contributions from the soil and canopy are not considered when computing r_r, while two-source models account for these contributions, as illustrated in Figure 2.8. Additionally, in one-source models, it is assumed that r_r equals r_a. This assumption makes one-source models relatively less complex and, as a result, they are widely used in remote sensing studies to estimate actual evapotranspiration (Zhang et al., 2016).

The early one-source SEB models were used for local and micro scale applications due to their dependence on local calibration, local reference surface fluxes or other data, lack of spatiotemporal scalability, and other factors (Zhang et al., 2016).Bastiaanssen et al. (1998) created the Surface Energy Balance Algorithm for Land (SEBAL) algorithm to overcome these restrictions. With only field data on shortwave atmospheric transmittance, surface temperature, and vegetation height needed, the SEBAL model can handle thermal infrared images at various spatial resolutions. It also estimates the spatial variation of most important hydrometeorological parameters by integrating empirical and physical parameterizations using inputs from satellite data and local weather data (Kamaraj & Rangarajan, 2022). Thus, SEBAL is a promising candidate to estimate large-scale ET_a estimates and is a commonly used parsimonious model to determine ET.

A variety of other one source models have been used in the literature to estimate ET from satellite imagery. Some of them are the Mapping Evapotranspiration at High Resolution with Internalized Calibration (METRIC), the Simplified Surface Energy Balance Index (SSEBI), and the Surface Energy Balance System Model (SEBSM), Simplified Surface Energy Balance (SSEB) (Glenn et al., 2007; Zhang et al., 2016). Numerous studies have successfully used SEBAL to understand the ET changes making it a widely used and validated technique to be applied (Kundu et al., 2018; Chemura et al., 2020; Munawir et al., 2022).

SEBAL has been applied to various Indian landscapes to estimate ET_a. Ahmad et al. (2006) utilized SEBAL with MODIS imagery to assess ET_a and map water usage patterns among vegetation in the

Krishna River basin, India. Jana et al. (2016) utilized Landsat 5 imagery and SEBAL to estimate ET_a in Doon Valley, Uttarakhand, India. They found acceptable ET_a ranges from SEBAL, correlating well with Penman-Monteith ET_o and ET_c obtained using data from the meteorological station located within the region of interest. Values were compared with surface temperature and vegetation indices to study land cover class variations. Bala et al. (2016) investigated SEBAL's application over an agricultural farm at the Indian Agricultural Research Institute (IARI) using Landsat7-ETM+ imagery. Validation against in-ground lysimeter measurements demonstrated consistency with SEBAL's ET_a values. The analysis of 7 different acquisition dates resulted in RMSE of 0.505 mm/day, MAE of 0.189 mm/day, and an R^2 of 0.91, highlighting SEBAL's utility in cropped regions.

Kamali & Nazari (2018) estimated maize water requirements in Iran's Mazandaran province using Landsat 8 imagery during the maize growth period. Evapotranspiration maps, with RMSE values of 0.74, 1.38, and 0.73mm/day, were compared with FAO56 Penman-Monteith, National Water Document (NWD), and Reference Book (RB) values used in Iran for crop characteristics. Good agreement between SEBAL and these references led to estimating K_c for maize crops, reaffirming SEBAL's reliability for obtaining crop K_c values. Rahimzadegan & Janani (2019) assessed SEBAL's efficiency in a 100-hectare pistachio farm in Semnan province, Iran, using 29 Landsat 8 images between 2013 and 2017. Validation against estimates from Intelligent Meteorological instrument (iMETos-Pessl) resulted in an R^2 of 0.8 and an RMSE of 2.5 mm per day, endorsing SEBAL's capacity in estimating ET_a over agricultural tracts.

SEBAL has not been used to estimate ET_a in the GRB: this study aims to be the first of its kind to do so. To help direct future model advancements, studies comparing various SEB models are required to determine the advantages and shortcomings of each model (Bhattarai et al., 2016). This study aims to conduct a multi-model comparison between SEBAL and METRIC across various land cover classes during mid-summer and the end of the monsoon season.

2.3 Land cover classification

2.3.1 Overview

Investigations into Land Use and Land Cover (LULC) changes have gained significance in recent years for monitoring environmental shifts and managing natural resources. These changes, stemming from human activities and natural phenomena, profoundly impact a region's landscape and its natural and human-made resources (Tian et al., 2014). LULC changes significantly influence food production, climate change, and biogeochemical cycles, recognized as key drivers of global environmental transformations since the mid-1970s (John et al., 2019). In addition, they play a major role in determining the evapotranspiration estimates, owing to the varying interactions between the surface and the atmosphere, leading to distinct differences in values among different land cover classes (Ayad Ali Faris Beg et al., 2016; Jana et al., 2016; Karishma et al., 2022; Li & Zhao, 2010; Ning et al., 2017). The differences become more pronounced when estimating different fluxes within the process of determining ET using surface energy balance (SEB) models. Especially, in the context of our study, LULC changes serve as drivers of seasonal evapotranspiration ET variations.

While often used interchangeably, "land use" and "land cover" represent distinct concepts in literature. Land use refers to human activities sustaining livelihoods, while land cover represents physical entities on the land surface, encompassing features like water bodies, forests, built-up areas, and wetlands. The classification of land cover involves delineating different classes by analyzing their spectral behavior, evolving from visual interpretation methods in the 1950s to modern computer-based algorithms (Alshari & Gawali, 2021).

Change detection in land cover classification identifies variations in a region's condition pertaining to different land cover types within a specified timeline. An accuracy assessment, critical for realistic change detection, involves comparing classified pixels with ground truth representations (Twisa & Buchroithner, 2019). The assessment results are interpreted by generating an error matrix, a cross-tabulation involving pixels from land cover classes of interest. To determine accuracy, common metrics such as user accuracy (UA), producer accuracy (PA), and overall accuracy (OA) are employed. UA measures the percentage of a category mistakenly included in another category on the ground (commission error), while PA measures the fraction of pixels excluded from a reference class (omission error). OA represents the likelihood that a pixel will be accurately categorized by the theme map (Chemura et al., 2020).

2.3.2 Land cover and surface hydrology

Over the past decades, significant changes due to deforestation, urbanization, and changes in agricultural practices have notably affected hydrological processes. These changes impact water resources by modifying evapotranspiration, interception, and infiltration rates. Particularly, the

expansion of built-up areas significantly alters hydrological processes such as stream flow, peak flow, runoff patterns, and water quality (Kamaraj & Rangarajan, 2022). Thus, when large landscapes like river basins are considered, examining the impact of LULC change on various water balance components becomes paramount but is challenging due to the scale and the subsequent availability of scientific tools necessary for the same (Samal & Gedam, 2021).

Various factors, including changes in water channels, land use changes, artificial river impoundments, and global climate change, contribute to variations in discharge patterns in major river basins. These modifications often occur in diverse combinations, making it complex to isolate their effects on hydrological components. Effective management of large-scale land use changes is crucial at the watershed level to prevent adverse impacts on regional ecological balance. Developing and applying a functional model is essential to comprehend the impact of land use changes on the hydrological system and predict potential outcomes (Samal & Gedam, 2021). Analyzing and monitoring regional and temporal shifts in LULC are vital for scientists, conservationists, agriculturalists, lawmakers, and urban planners (Kamaraj & Rangarajan, 2022).

2.3.3 Remote sensing and land cover classification using Landsat imagery

The first Landsat satellite was launched in 1972, and the mission has remained operational since then, except for a scanline error in Landsat 7 on May 21, 2003. The Landsat archive contains over 10.2 million images obtained using different satellites, starting from Landsat 1-5 MSS to Landsat 8-9 OLI/TIRS (Crawford et al., 2023). This vast collection of detailed and continuous data plays a crucial role in understanding LULC trends. Landsat images are organized into scenes, specifying

paths and rows relative to the satellite's position above Earth's surface. Notably, Landsat provides expansive coverage and historical data availability at no charge. The spatial resolution has progressed over time, from 60 meters in Landsat 1-3 Multispectral Scanner (MSS) to 30 meters (visible bands) and 15 meters (panchromatic) in Landsat 8 Operational Land Imager (OLI) (Phiri & Morgenroth, 2017). Table 2.1 describes the major Landsat satellites and their respective radiometric information:

Table 2.1 System characteristics of the major Landsat family of satellites (Source: Williams et al., 2017)

	Landsat 1		**Landsat 2**		**Landsat 5**		**Landsat 8**	
Year(s) used	1976		1980		1984-2008		2013	
Sensor	Multispectral Scanner (MSS)		Multispectral Scanner (MSS)		Thematic Mapper (TM)		Operational Land Imager (OLI)	
Spatial resolution	60 meters		60 meters		30 meters		30 meters	
Bands Used	Band	Wavelength (micrometers)	Band	Wavelength (micrometers)	Band	Wavelength (micrometers)	Band	Wavelength (micrometers)
	Green	0.5 – 0.6	Green	0.5 – 0.6	Blue	0.45 – 0.52	Blue	0.45 – 0.51
	Red	0.6 – 0.7	Red	0.6 – 0.7	Green	0.52 – 0.60	Green	0.53 – 0.59
	NIR	0.7 – 0.8	NIR	0.7 – 0.8	Red	0.63 – 0.69	Red	0.64 – 0.67
	NIR	0.8 – 1.1	NIR	0.8 – 1.1	VNIR	0.76 – 0.90	NIR	0.85 – 0.88
					SWIR	1.55 – 1.75	SWIR	1.57 – 1.65
					SWIR	2.08 – 2.35	SWIR	2.11 – 2.29

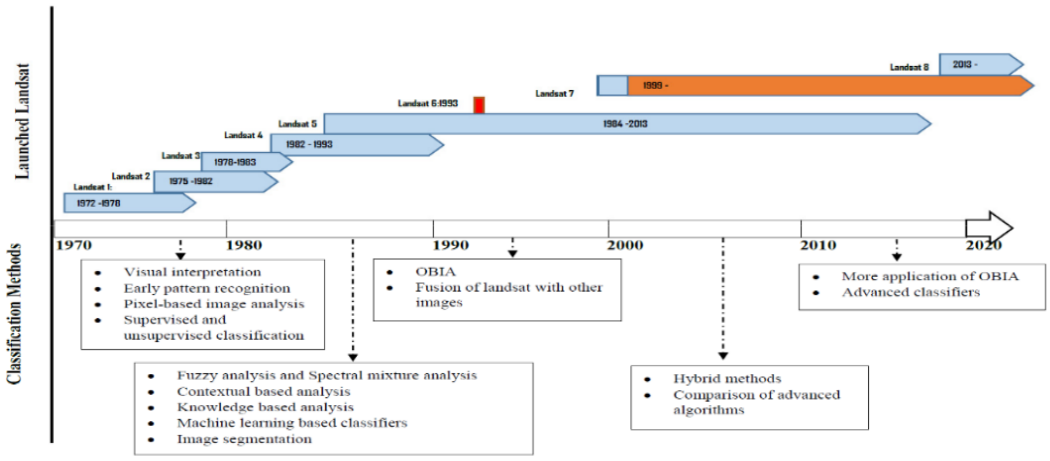

Figure 2.9 Evolution of Landsat classification methods over the years (Source: Phiri & Morgenroth, 2017)

The two primary pixel-based classification methods are supervised and unsupervised. Supervised algorithms involve training data to provide prior knowledge to the model. This is achieved by sampling pixels representing specific land cover classes and using their values to set thresholds for identifying similar pixels. Unsupervised methods, in contrast, do not rely on prior knowledge. They group similar pixels based on spectral values without predefined training (Phiri & Morgenroth, 2017). Some of the methods and their evolution over the years is described in Figure 2.10.

Various supervised algorithms (e.g., support vector machine (SVM), random forest (RF), spectral angle mapper (SAM), Mahalanobis distance (MD), radial basis function (RBF), decision tree (DT)) and unsupervised methods (e.g., cluster algorithm, fuzzy c-means algorithms, K-means algorithm, ISODATA) have been used for LULC classification. Accuracy levels vary among different machine learning techniques depending on the application. Notably, compared to other conventional classifiers, non-linear classifiers such as ANN, SVM, and RF often demonstrate higher accuracy (Talukdar et al., 2020).

3 Study area and datasets

3.1 Study area description – Upper Gundar River Basin

The study area focuses on a segment of the Gundar river basin, specifically the upper Gundar river basin which covers around 2640 km². Its altitude ranges from 19 to 1184 meters above sea level, with mountain peaks in the northern region (Figure 3.1).

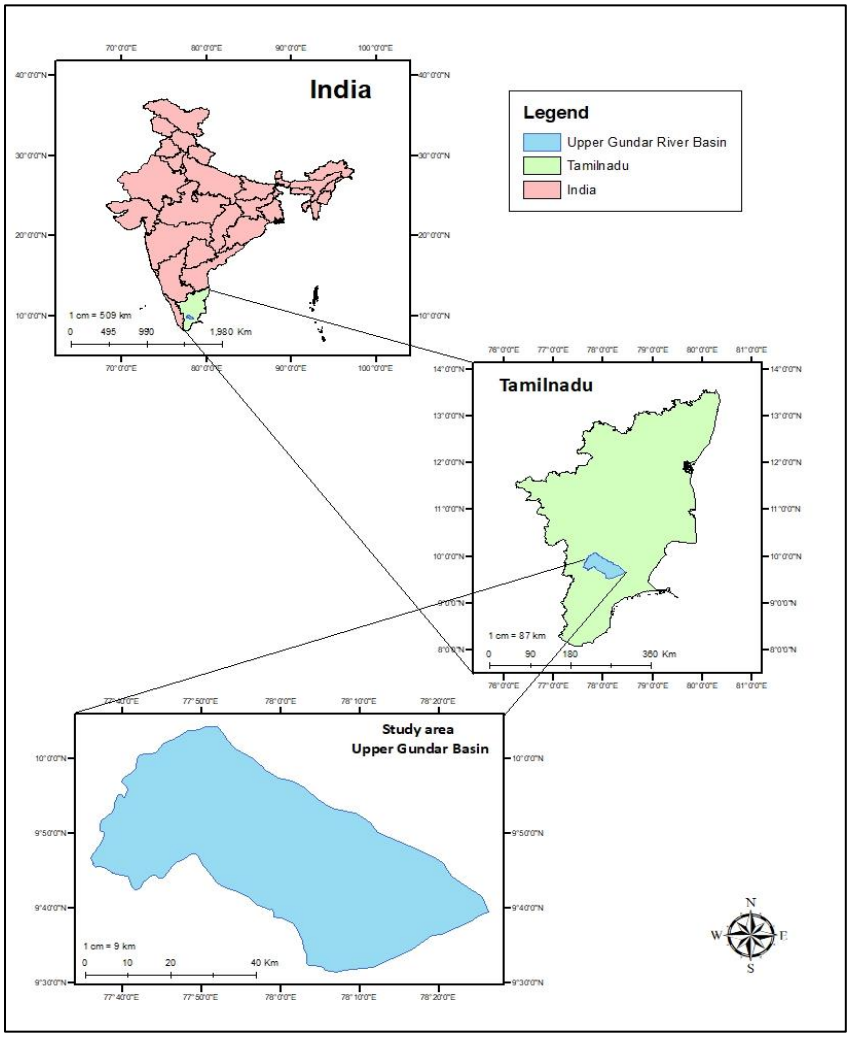

Figure 3.1 Study area map – Upper Gundar Basin

3.2 Datasets

3.2.1 Landsat data

The study utilizes various Landsat data sets, including Landsat 4-5 TM Collection 2 Level 1 data, Landsat 4-5 TM Collection 2 Level 2 data, Landsat 8-9 OLI/TIRS Collection 2 Level 1 data, and Landsat 8-9 OLI/TIRS Collection 2 Level 2 data. Collection 2 Level 1 data comprises unsigned 16-bit integer digital numbers (DNs), integral for obtaining spectral components relevant to the SEBAL methodology. Collection 2 Level 2 products offer surface reflectance data derived from the Level 1 datasets and are utilized for land cover classification analysis (Crawford et al., 2023).

The satellites have a 16-day repeat cycle and the length of the scenes are roughly 170 km by 183 km. Access to Landsat data is facilitated through the Earth Explorer archive (https://earthexplorer.usgs.gov/), organized via a path-row framework (Crawford et al., 2023). The study focuses on the scene with path number 143 and row number 53. For a given year, the initial image is chosen during the summer season (March to May) in this region, while another image is selected towards the conclusion of the northeast monsoon season (mid-December to early January), to examine seasonal variations within the year. These chosen images are considered representative of the mid-summer and the end of the northeast monsoon season. It is important to acknowledge that cloud cover significantly influenced the selection process, and the analysis chose the best available scenes with cloud cover below 20% (Table 3.1).

Table 3.1 Image details used for analysis

Image acquisition date	Season	Landsat number
April 03, 2006	Summer	Landsat 5
December 15, 2006	Northeast Monsoon	Landsat 5
April 09, 2014	Summer	Landsat 8
January 06, 2015	Northeast Monsoon	Landsat 8
April 28, 2021	Summer	Landsat 8
December 24, 2021	Northeast Monsoon	Landsat 8

3.2.2 Topographic Data – Aster DEM

Topographic data is integral for the execution of SEBAL. This study relies on version 3 of the Advanced Spaceborne Thermal Emission and Reflection Radiometer Global Digital Elevation model (ASTER GDEM), available from the Earth Science Data Systems (ESDS) Program's website: https://search.earthdata.nasa.gov/search. This dataset is presented in 1° tiles and comes in tiff format, offering a spatial resolution of 30m. It encompasses a single band providing elevation data relative to mean sea level. Version 3 has undergone rigorous quality checks, including reprocessing and manual inspection of image tiles, ensuring a highly accurate dataset. Any gaps in the data have been filled using SRTM data, while high latitude values have been supplemented with national datasets and interpolation techniques (Abrams et al., 2022). In this thesis, it is referred to as Aster DEM.

3.2.3 In-situ Weather Data

The weather data necessary for the analysis are acquired from the Water Resources Department, Groundwater Division, operating under the aegis of the Government of Tamil Nadu. Within the provided Landsat scene, weather information is available for five locations shown and tabulated below (Figure 3.5 and Table 3.2). While some stations are situated closer to the study area, it can be found that Kundrakudi and Vembakottai are located farther away. In terms of elevation, they range between 60m and 102m, indicating a relatively flat terrain when all stations are considered.

Figure 3.2 Upper Gundar basin and the neighboring weather stations

Table 3.2 Reference weather stations used in the study

Station	Coordinates (latitude, longitude)	Elevation (m)
Kavalur	9.57°N, 77.90°E	99
Savasapuram	9.48°N, 78.14°E	81
Parthibanur regulator (PBR)	9.63°N, 78.47°E	60
Vembakottai	9.33°N, 77.75°E	78
Kundrakudi	10.11°N, 78.70°E	102

Weather data is provided by the department twice a day, at 8:30 am and 5:30 pm. The 8:30 am data comprises observations from 5:30 pm the previous day to 8:30 am on the current day, while the 5:30 pm data encompasses records between 8:30 am and 5:30 pm of the same day. This dataset includes minimum and maximum temperatures, mean relative humidity, wind speed, rainfall, wet bulb temperature, and dry bulb temperature. The number of sunshine hours per day is recorded from 5 am to 7 pm at hourly intervals. However, there is often a lack of data for different variables due to instrument repairs. Wind speed data is excluded from the analysis due to the ambiguity in the measurements and potential uncertainties associated with the instruments.

3.2.4 Global Land Data Assimilation System (GLDAS)

Accurate estimations of terrestrial water and energy storages are crucial for diverse geoscience studies and for predicting climate change, weather patterns, biological/agricultural production, and flooding. The Global Land Data Assimilation System (GLDAS), developed collaboratively by

scientists at the National Oceanic and Atmospheric Administration (NOAA) National Centers for Environmental Prediction and the National Aeronautics and Space Administration (NASA) Goddard Space Flight Center (GSFC), employs data assimilation techniques to derive realistic land surface states. These states provide critical values like temperature, wind speed, pressure, and specific humidity (Rodell et al., 2004).

GLDAS has shown effectiveness in studies related to Tamil Nadu, notably in the study by Chinnasamy & Agoramoorthy (2015) aiming to quantify increased irrigation's impact on groundwater levels using soil moisture estimates from GLDAS and the GRACE dataset. It successfully determined fine-scale groundwater depletion rates and can help policymakers in water resource management strategies. Moreover, Janani et al., (2023) utilized GLDAS for a basin-scale study in the Lower Bhavani Basin, Tamil Nadu, between 2003 and 2022, estimating yearly and seasonal trends and variability of soil moisture. Similarly, KP et al., (2021) combined TRMM, GLDAS, and MODIS datasets to analyze agricultural drought intensities in Tamil Nadu. They proposed using a blend of remote sensing-based indices as effective proxies to gauge vegetative stress. Additionally, Chanu et al., (2020) studied groundwater trends in three districts of Tamil Nadu using TRMM, GRACE, and GLDAS datasets, observing good agreement in groundwater storage changes. Therefore, these studies provide some confidence in utilizing GLDAS data within our study area, located in the state of Tamil Nadu.

In terms of its applications involving SEBAL, Saboori et al., (2022) fused GLDAS with SEBAL for daily actual ET estimations and validated GLDAS effectively against the SEBAL model using

Landsat images and Eddy Covariance tower measurements. Similarly, Lima & Ribeiro (2018) utilized GLDAS with SEBAL to study the spatial and seasonal ET_a patterns in the Paraibuna watershed, southeast Brazil. By discussing a few applications of the usage of the GLDAS data, its effective utility in areas lacking ground meteorological data is highlighted.

For this study, we utilized the GLDAS Noah Land Surface Model L4 3-hourly 0.25 x 0.25-degree V2.1 (GLDAS_NOAH025_3H). This product offers global spatial coverage and a spatial resolution of 0.25° x 0.25°. Temporal coverage extends from 2000 to the present, with data available at 3-hour intervals. Consequently, eight files are generated daily, with timestamps referenced to Greenwich Mean Time (GMT). It is important to note that Indian Standard Time (IST) is 5 hours and 30 minutes ahead of GMT. Therefore, for a specific day, the first file corresponds to 12:00 AM GMT and 5:30 PM IST. The data is provided in the netCDF format (Beaudoing & Rodell 2020).

It is freely accessible from the NASA Goddard Earth Sciences Data and Information Services Center (GES DISC) website: https://disc.gsfc.nasa.gov/. The data was downloaded using the bash command from Ubuntu 20.04.6 LTS version. Throughout this document, the product is referred to as GLDAS or GLDAS data.

3.2.5 ESA WorldCover 10 m 2020 v100

The WorldCover 10m 2020 v100 is a high-resolution land cover product created at 10m resolution, utilizing data obtained from the Sentinel-1 and Sentinel-2 satellites. It comprises 11 classes derived according to the Land Cover Classification System prescribed by the UN-FAO. A

consortium, headed by VITO Remote Sensing, has developed the WorldCover product in collaboration with organizations like, CS SI, Brockmann Consult, IIASA, Gamma Remote Sensing AG and Wageningen University (Zanaga et al., 2021).

4 Methodology

4.1 Land cover classification - Random Forest algorithm

Random forest (RF) algorithms have been widely adopted in land cover classification studies (Amini et al., 2022; Nguyen et al., 2018; Phan et al., 2020). RF effectively handles outliers and noisy datasets and can perform well with high-dimensional and multi-source datasets, achieving higher accuracy than other classifiers like SVM, kNN, or MLC. According to a meta-analysis of 349 Google Earth Engine (GEE) peer-reviewed studies published in the last decade, the RF algorithm has been the most frequently used classification algorithm for satellite images (Phan et al., 2020).

RF is an extension of decision tree techniques and operates based on bootstrapping and bootstrap aggregation (or bagging). The concept involves using multiple iterations of a predictor or classifier to obtain a final choice through a plurality vote among the predictors. Bagging shows an increase in accuracy with more predictors until it begins to decline. Trees are developed by bootstrapping the sample, involving subsetting random samples with replacement in the training data. Not all samples are used to grow the tree, and the discarded samples within the bootstrapped sample are considered as out-of-bag (OOB) data. During tree construction, only a randomly chosen set of input features is evaluated in each node. The OOB data are used to calculate the classification error rate and the significance of the input variables (features) as new trees are added to the forest. Once the forest is complete, a majority vote among all trees in the forest is used to classify a case, like the concept of bootstrap aggregating (Kulkarni & Lowe, 2016; Nguyen et al., 2018).

The two main parameters required for running the algorithm are ntree (number of trees) and mtry (variable split at each node). The mtry parameter is typically set to the square root of the total number of input variables. Reducing the number of variables used in a split using the value suggested above not only decreases the algorithm's computational complexity but also reduces the correlation between the trees without compromising the performance of the algorithm. The RF split procedure aims to decorrelate the trees, making the classification outcome less unpredictable and more dependable (Kulkarni & Lowe, 2016; Nguyen et al., 2018).

With respect to our study, surface reflectance data from Collection 2 Level 2 Tier 1 images obtained for the six different dates are subjected to land cover classification using the random forest algorithm implemented through the google earth engine platform. The ntree parameter is set to 100 and mtry is set to 3. To train the classifier, reference polygons were generated to delineate distinct classes across the study area. The tools used for creating these polygons included various false color composites derived from Landsat imagery, Google Earth, and the images captured during a field visit in 2022. A total of 1000 points were sampled for the water classes, while 4000 points were sampled for the rest of the classes. A smaller number of points were chosen for the water classes due to their reduced coverage in the landscape compared to other land cover classes.

It was also made sure that points from all the polygons were chosen to represent the specific class. A 70-30 split was utilized to allocate training and validation samples, following a study conducted in the Munneru River Basin, located in the neighboring state of Andhra Pradesh, India

(Loukika, Keesara, & Sridhar, 2021). Later, the overall, producer and user accuracies are obtained validating the use of the map. All the bands in the visible and infrared ranges along with the band values from Aster DEM were used for the analysis. The Landsat spectral bands are processed by multiplying the scale factor and adding the offset as prescribed by (Crawford et al., 2023). Before classification, clouds, cloud shadow areas are masked out using the quality assessment (QA) band in the Landsat imagery and regions above 200m from MSL are masked out using Aster DEM. Also, we observed the built-up class interfering with the exposed soil class and barren land class, so we incorporated the built-up area layer from the Sentinel-2A from 2020 land cover product additionally to mask out the built-up areas present in the image (Zanaga et al., 2021).

Table 4.1 provides a comprehensive listing of the classes used in this study along with their descriptions. These classifications are established through a combination of user expertise and spectral characteristics.

Table 4.1 Land cover classes and description

Class	Description
Water	Bodies of water – mainly tanks
Exposed soil - type 1	Land areas with exposed and dry soil
Barren land	Land characterized by dry soil and presence of small prosopis plants
Pre growth	Land that has been either tilled or harvested
Agro	Green agricultural land
Prosopis	Prosopis juliflora vegetation
Turbid water	Bodies of turbid water – mainly tanks
Exposed soil - type 2	Like exposed soil - type 1

It is important to highlight that the class Exposed Soil - Type 2 is exclusively considered in the dataset for the month of April in 2006. Given the proximity of barren land and exposed soil in this geographical area, there exists the possibility of their mutual misclassification, constituting a potential caveat in the analysis.

The same caveat is also applicable to agro and prosopis classes but with the given tools at the disposal, best efforts are taken to delineate them. The different image acquisition dates and the cloud cover percentages within the Landsat scene are described in Table 4.2.

Table 4.2 Image acquisition details for the analysis

Image acquisition date	Season	Cloud Cover (%)
April 03, 2006	Summer	5
December 15, 2006	Northeast Monsoon	13
April 09, 2014	Summer	4
January 06, 2015	Northeast Monsoon	18
April 28, 2021	Summer	10
December 24, 2021	Northeast Monsoon	2

4.2 Actual Evapotranspiration estimation

4.2.1 Surface Energy Balance Algorithm for Land (SEBAL)

The Surface Energy Balance Algorithm for Land (SEBAL) operates on the fundamental principle of energy balance, analyzed at the pixel level. Instantaneous latent heat of evapotranspiration (λET) is obtained as a residue from the balance between net radiation flux at the surface (R_n), ground heat flux (G), and sensible heat flux to air (H) as shown in Equation 4.1 (Bastiaanssen et al., 1998)

$$\lambda ET = R_n - G - H \quad (4.1)$$

Subtracting all outgoing radiant fluxes from all incoming radiant fluxes yields R_n (Equation **Error! Reference source not found.**) (Waters et al., 2002):

$$R_n = (1-\alpha)R_{s\downarrow} + R_{L\downarrow} - R_{L\uparrow} - (1-\varepsilon_0)R_{L\downarrow} \quad (4.2)$$

Where; ε_0 represents surface thermal emissivity, α stands for the surface albedo, $R_{s\downarrow}$ is the incoming shortwave radiation, $R_{L\downarrow}$ is the incoming longwave radiation, $R_{L\uparrow}$ is the outgoing longwave radiation.

$R_{s\downarrow}$ which is the incoming solar radiation is calculated from:

$$R_{s\downarrow} = G_{sc} * cos\,\theta * d_r * \tau_{sw} \quad (4.3)$$

where G_{sc} is the solar constant 1367(W/m^2), θ is solar incidence angle, d_r is the inverse of the squared relative distance between the sun and the earth. τ_{sw} is the atmospheric transmissivity calculated using the elevation above the mean sea level represented by z (Waters et al., 2002):

$$\tau_{sw} = 0.75 + 2 * 10^{-5} * z \quad (4.4)$$

The outgoing longwave radiation, $R_{L\uparrow}$ in Equation 4.2 is calculated by determining select vegetation indices, two types of surface emissivities, corrected thermal radiance (R_c) and surface temperature (T_s). The vegetation indices, Normalized Difference Vegetation Index (NDVI), Soil Adjusted Vegetation Index (SAVI) and Leaf Area Index (LAI), are calculated from the following (Waters et al., 2002):

$$NDVI = \frac{(NIR - R)}{(NIR + R)} \quad 4.5$$

$$SAVI = \frac{(1 + L)\,(NIR - R)}{(L + NIR + R)} \quad 4.6$$

$$LAI = \frac{-\ln\left(\frac{0.69 - SAVI}{0.59}\right)}{0.91} \quad 4.7$$

where NIR and R denote the reflectivity in the Near Infrared and Red bands in Landsat imagery, respectively. The constant 'L' is integral in computing SAVI, commonly set to 0.5 as per (Allen et al., 1998). ε_{NB} is the first emissivity component denoting the thermal emission behavior of the surface within the narrow thermal band of Landsat imagery (10.4 to 12.5 µm). The next component, ε_0 represents the surface behavior across the broader thermal spectrum (6 to 14 µm), calculated using the following equations (Waters et al., 2002):

$$\varepsilon_{NB} = 0.97 + 0.0033\,LAI \quad 4.8$$

$$\varepsilon_0 = 0.95 + 0.01\,LAI \quad 4.9$$

If LAI is greater than 3, ε_{NB} and ε_0 are set to 0.98. R_c is obtained before calculating T_s through Equation 4.10 (Waters et al., 2002) :

$$R_C = \frac{L_6 - R_P}{\tau_{NB}} - \left(1 - \varepsilon_{NB}\right)R_{sky} \quad 4.10$$

where L_6 is the spectral radiance of the thermal band in the Landsat imagery, R_p is the path radiance in the narrow 10.4 – 12.5 µm band, R_{sky} downward thermal radiation for a clear sky pertaining to the narrow band and τ_{NB} is the transmissvity of air in the narrow 10.4 – 12.5 µm

band. It is recommended to set τ_{NB} to 1 in the absence of an accurate value (Waters et al., 2002). R_{sky} is estimated using the equation of Wukelic et al., 1989:

$$R_{sky} = 1.807 * 10^{-10} * T_a * (1 - 0.26\, e^{(-7.77 * 10^{-4} * (273.15 - T_a)^2)}) \quad 4.11$$

where T_a represents the near-surface air temperature during the time of Landsat overpass, measured in Kelvin (K). As the next step, T_s is obtained using Equation 4.12:

$$T_s = \frac{K_2}{ln\left(\frac{\varepsilon_{NB}\, K_1}{R_c} + 1\right)} \quad 4.12$$

K_1 and K_2 are constants pertaining to the Landsat imagery present as metadata. Finally, $R_{L\uparrow}$ is calculated using the Stefan – Boltzmann equation (Waters et al., 2002):

$$R_{L\uparrow} = \varepsilon_0 * \sigma * T_s^4 \quad 4.13$$

where ε_0 stands for the dimensionless 'broad-band' surface emissivity, σ denotes the Stefan-Boltzmann constant (5.67×10^{-8} W/m^2/K^4), and T_s represents the surface temperature measured in Kelvin (K).

The next major component in the radiation balance (equation 4.2), $R_{L\downarrow}$ the incoming longwave radiation is determined using the Stefan-Boltzmann equation again.

$$R_{L\downarrow} = \varepsilon_a * \sigma * T_a^4 \quad 4.14$$

where ε_a is the atmospheric emissivity (dimensionless), and T_a denotes the near-surface air temperature measured in Kelvin (K). ε_a is estimated through Equation 4.15.

$$\varepsilon_a = 0.85 * (-ln\,\tau_{sw})^{0.09} \quad 4.15$$

where τ_{sw} stands for the atmospheric transmissivity as discussed above.

Soil heat flux (G) is the next major component as a part of the radiation balance equation (equation 4.1). The rate at which heat is absorbed by plants and soil by conduction is known as the soil heat flux. SEBAL uses the empirical equation that Bastiaanssen (2000) established to obtain the ratio G/R_n (Equation 4.16):

$$\frac{G}{R_n} = \frac{T_s}{\alpha}(0.0038\alpha + 0.0074\alpha^2)(1 - 0.98NDVI^4) \quad 4.16$$

where, T_s represents the surface temperature (^oC), and α denotes the surface albedo. The calculation of G involves multiplying the derived G/R_n ratio by the computed value of R_n.

Finally, the sensible heat flux, (H in equation 4.1) is the most complicated component of the radiation balance to compute. The rate at which heat is lost to the air by convection and conduction as a result of a temperature differential is known as sensible heat flux (Waters et al., 2002; Zamani Losgedaragh & Rahimzadegan, 2018). It is estimated using from:

$$H = \frac{(\rho_{air} * C_p * dT)}{r_{ah}} \quad 4.17$$

where ρ_{air} represents air density (in kg/m^3), C_p denotes the specific heat of air (1004 J/kg/K), dT (K) signifies the temperature difference ($T_1 - T_2$) between two heights (z_1 and z_2), and r_{ah} stands for the aerodynamic resistance to heat transfer (measured in s/m) (Waters et al., 2002).

The presence of two unknowns - dT and r_{ah} makes it difficult to solve Equation 4.17 (Zamani Losgedaragh & Rahimzadegan, 2018). To simplify this process, the SEBAL makes use of two classes of anchor pixels (cold and hot) to fix the boundary conditions along with the wind speed at a defined height above the surface to determine H and dT. The anchor pixels must be situated within the study area to ensure reliable estimates of H and dT. Hot anchor pixels correspond to dry and uncultivated agricultural regions where the ET is close to 0.

Cold pixels represent agricultural regions that are well irrigated and fully vegetated. dT is estimated by acknowledging a major presumption of SEBAL which is the linear relationship between dT and T_s (Waters et al ., 2002).

When calculating sensible heat flux (H), atmospheric stability conditions are crucial since they greatly impact r_{ah} (Waters et al., 2002). The parameters r_{ah} and dT are iteratively calculated, incorporating corrections based on the Monin-Obukhov theory to account for the effects of buoyancy originating from surface heating. The Monin-Obukhov length (L) is utilized for this purpose, serving as an indicator of the stability conditions. The first step of this iterative process involves the usage of r_{ah} and frictional velocity (u_*) measured at the closest weather station as the baseline inputs (Zamani Losgedaragh & Rahimzadegan, 2018). Corrections are incorporated based on the neutrality condition determined by L, as described in Equation 4.18.

$$L = \frac{-\rho C_P u_*^3 T_s}{kgH} \qquad 4.18$$

Here, ρ represents the density of air (in kg/m³), C_p denotes the specific heat of air (1004 J/kg/K), u_* signifies the friction velocity (in m/s), T_s stands for the surface temperature (in Kelvin), g represents the gravitational constant (9.81 m/s²), and H represents the sensible heat flux.

In terms of atmospheric conditions, L < 0 represents unstable conditions, L = 0 represents neutral conditions, and L > 0 represents stable conditions. "With the inclusion of correction factors as outlined in (Waters et al., 2002) based on L values, the process continues until the convergence of rah and dT (Waters et al., 2002; Zamani Losgedaragh & Rahimzadegan, 2018). They are finally calculated thus:

$$r_{ah} = \frac{ln\frac{z_2}{z_1} - \psi_{h_{z_2}} + \psi_{h_{z_1}}}{u_* k} \quad 4.19$$

$$u_* = \frac{k\, u_{200}}{ln\frac{200}{z_{om}} - \psi_{m_{200}}} \quad 4.20$$

where, $\psi_{m(200)}$ represents the stability correction for momentum at a height of 200 meters, $\psi_{h(z1)}$ and $\psi_{h(z2)}$ denote the stability corrections for heat transport at heights of z_1 and z_2 meters, respectively. Additionally, k represents Von-Karman's constant which is set to 0.41 as per (Allen et al., 1998), u_{200} signifies the wind speed at a height of 200m, and z_{om} represents the momentum roughness length. z_{om} quantifies the combined impact of skin friction and form drag in the air layer interacting with the surface. After applying the stability corrections, the sensible heat flux (H) is estimated using Equation 4.17, and the latent heat flux (λET) is obtained using Equation 4.1. Figures 4.1 and 4.2 conceptualize the logic of the estimation of H and λET respectively using flowcharts. λET is depicted as LE in Figure 4.1.

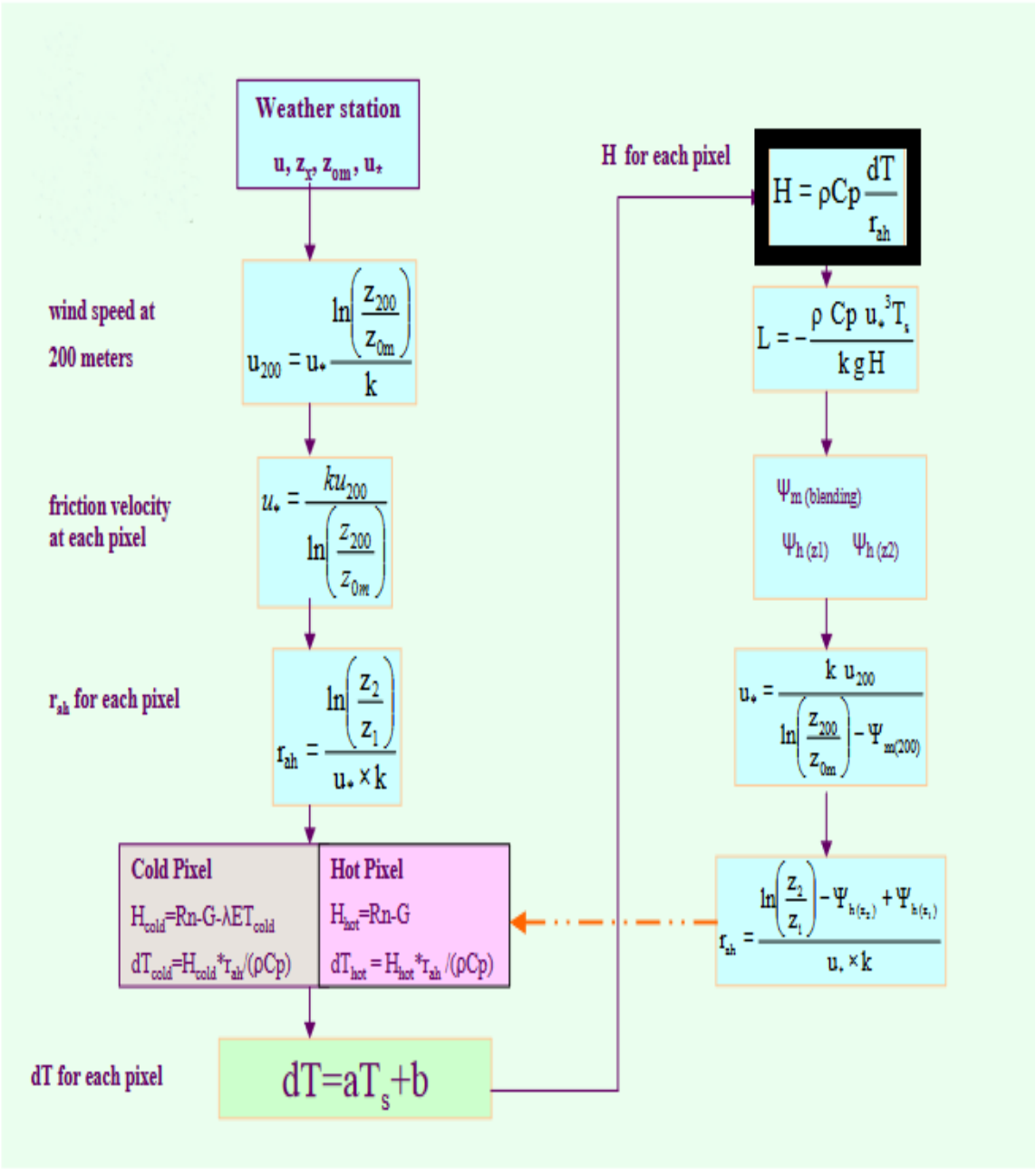

Figure 4.1 Iterative process to estimate H (Source: Waters et al., 2002).

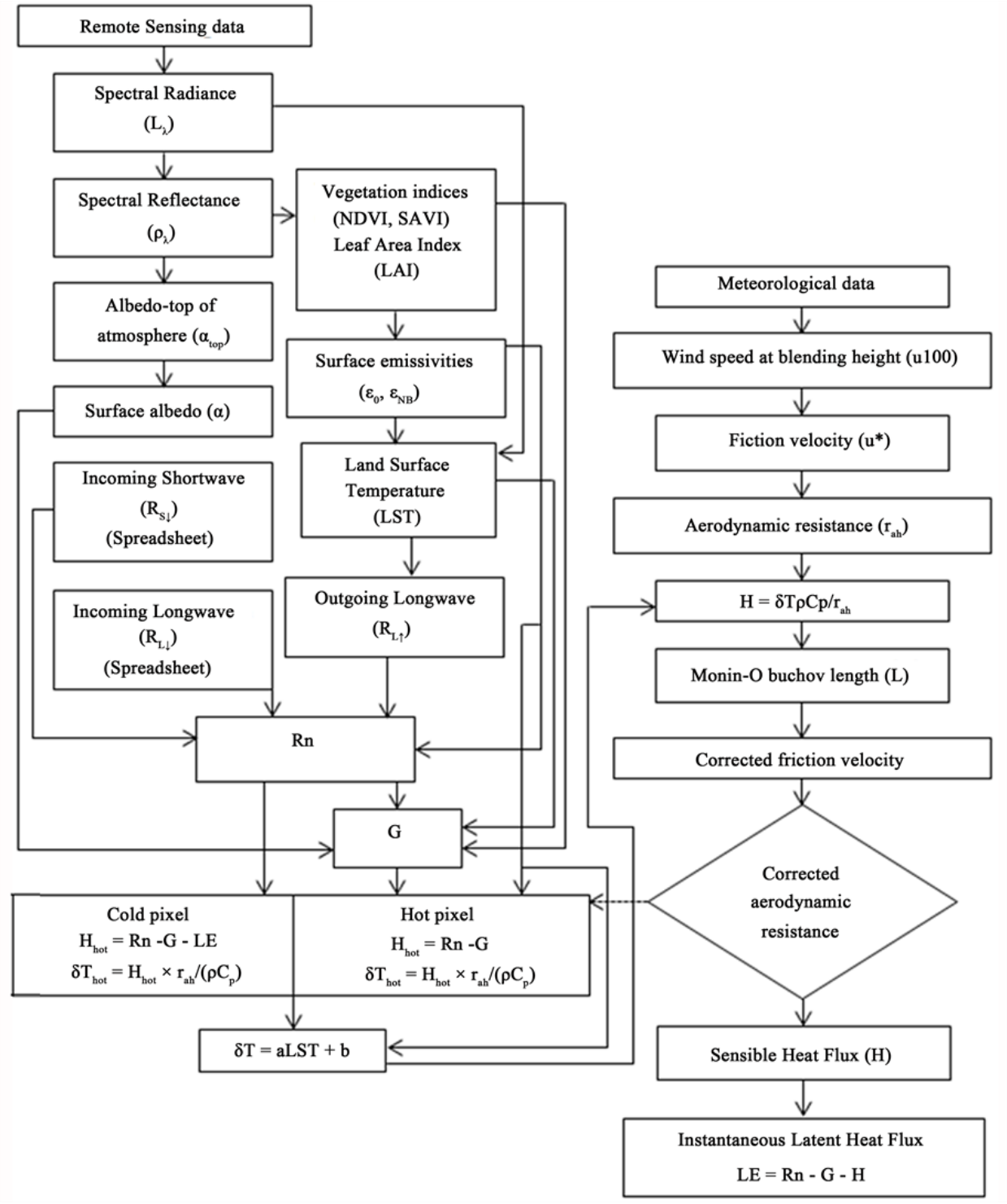

Figure 4.2 SEBAL Flowchart (Source: Bezerra et al., 2015)

4.2.1.1 pySEBAL - Model description

pySEBAL refers to the implementation of the SEBAL model in Python conceptualized and developed at IHE Delft by the Water Accounting group. The inputs required for the execution of the model to estimate actual evapotranspiration include spectral information, weather data, and topographic data. The code, developed as PySEBAL version 3.8, is open source and available on the GitHub repository. The estimation of hot and cold pixels is specific to one scene and is automated by incorporating thresholds based on percentiles using NDVI and T_s. While hot pixels are typically chosen from dry surfaces, cool pixels are chosen from water bodies or agricultural areas with well-developed vegetation. Hot pixels are identified based on the distribution of NDVI, and the 1st and 3rd percentiles are used as the respective lower and upper bounds. T_s is used to determine the cold pixels, and the lower and upper bounds are 2nd and 5th percentiles, respectively. The model is parsimonious in the sense that it requires minimal user input. The implementation currently supports Landsat, MODIS, Proba-V, and VIIRS satellites (Karimi et al., 2019; Pareeth & Karimi, 2023).

In pySEBAL, after calculating λET using Equation 4.1, the instantaneous evaporative fraction (EF_{inst}) is estimated using the equation of Pareeth & Karimi (2023):

$$EF_{inst} = \frac{\lambda ET}{R_n - G} \tag{4.25}$$

To account for the increase in the ET rates during the afternoon, pySEBAL incorporates an advection factor Ω derived from Pareeth & Karimi (2023):

$$\Omega = 1 + 0.985 * (exp[(e_{sat} - e_{act}) * 0.08] - 1) * EF_{inst} \quad 4.26$$

where, e_{sat} stands for saturated vapor pressure and e_{act} stands for actual vapor pressure, R_n stands for the instantaneous net radiation and G for instantaneous soil heat flux obtained in Equations 4.16 and 4.17 respectively. Finally, ET_{24}, the daily ET estimate in (mm/day) is obtained using Equation 4.27 (Pareeth & Karimi, 2023):

$$ET_{24} = \Omega * EF_{inst} * R_{n_{24}} \quad 4.27$$

where R_{n24} represents the net radiation in a day.

4.2.2 Mapping Evapotranspiration at High Resolution with Internalized Calibration (METRIC)

METRIC and SEBAL both operate on the same principle of estimating λET at the individual pixel level. They are similar in terms of computational aspects and differ only in certain relationships (Kamyab et al., 2022). Unlike SEBAL, which calculates τ_{sw} using the elevation above mean sea level, METRIC estimates τ_{sw} using air pressure and the water content of the atmosphere (Zamani Losgedaragh & Rahimzadegan, 2018).

$$\tau_{sw} = 0.35 + 0.627 * exp[\frac{-0.00146p}{k_t cos\theta_{hor}} - 0.75 * (\frac{w}{cos\theta_{hor}})^{0.4}] \quad 4.21$$

Here, p stands for atmospheric pressure (kPa), Θ_{hor} stands for solar zenith angle, k_t is a dimensionless quantity representing the turbidity coefficient of the atmosphere and w (in mm) stands for atmospheric water content. Through Equations 4.22 – 4.23, p and w is determined as:

$$p = 101.3 * (\frac{293 - 0.0065z}{293})^{5.26} \quad 4.22$$

$$w = 0.14 e_a P_{air} + 2.1 \quad (4.23)$$

Here, z is the elevation above the mean sea level, e_a is the vapor pressure close to surface (kPa) and P_{air} is the air pressure (kPa). In METRIC, the soil heat flux component, G/R_n is calculated using Equation 4.24.

$$\frac{G}{R_n} = \begin{cases} \frac{1.80(T_s - 273.15)}{R_n} + 0.084, & LAI < 0.5 \\ 0.18e^{-0.521LAI} + 0.05, & LAI \geq 0.5 \end{cases} \quad (4.24)$$

G is then calculated by multiplying it with the previously obtained R_n values. Unlike SEBAL, which assumes a zero value for H at the cold pixels, METRIC requires the alfalfa based reference evapotranspiration values (ET_r) for estimating H at these pixels. Additionally, to determine daily actual evapotranspiration (ET_{24}), SEBAL utilizes the instantaneous evaporative fraction (EF_inst), while METRIC uses the component known as the reference ET fraction (ET_rF) (Allen et al., 2011).

4.2.2.1 EEFlux model implementation

Earth Engine Evapotranspiration Flux (EEFlux) is a jointly developed by Google, the University of Idaho, Desert Research Institute, and the University of Nebraska – Lincoln. EEFlux is implemented using the METRIC algorithm within the Google Earth Engine platform, that contains a comprehensive historical record of Landsat imagery dating back to 1984. Weather data inputs are sourced from the North American Land Data Assimilation System (NLDAS) for the CONUS

region, while the Climate Forecast System Version 2 (CFSV2) gridded dataset is utilized for the rest of the world. The digital elevation map utilized is from the Shuttle Radar Topography Mission (SRTM), available at a resolution of 30 meters.

ET_a maps can be accessed via the EEFlux portal (https://EEFlux-level1.appspot.com/) by specifying the date and reference coordinates. The portal then generates a TIFF file corresponding to the Landsat scene identified by the provided coordinates. In this study, data for the Landsat scene with path number 143 and row number 53 are obtained (Allen et al., 2015).

4.2.3 Data Preprocessing and Model Implementation Scenarios

The primary input for the model is Landsat data. To implement pySEBAL, Landsat Collection 2 Level 1 data scenes for the specified date are downloaded from the Earth Explorer archive. In addition to Landsat imagery, the meteorological forcing data for pySEBAL is crucial. The required parameters for pySEBAL implementation include averages of surface solar radiation, temperature, wind speed, and relative humidity. The equivalent variables present in the GLDAS product, and their descriptions are outlined in the Table 4.3 given below:

Table 4.3 GLDAS forcing meteorological variables.

Parameter	Description	Unit
Swdown_f_tavg	Downward short-wave radiation	W/m^2
Qair_f_inst	Specific humidity	kg/kg
Psurf_f_inst	Pressure	Pa
Wind_f_inst	Wind speed	m/s
Tair_f_inst	Temperature	K

Temperature, wind speed, and downward shortwave radiation flux are directly obtained from the GLDAS dataset. However, relative humidity cannot be directly extracted. But, the model's outputs of air temperature, air pressure, and specific humidity can be used to calculate it described in Equations 4.1 to 4.3 given below (Pareeth & Karimi, 2023).

$$e_a = \frac{Q_{air} * P_{surf}}{(0.378 * Q_{air} + 0.622)} \quad 4.25$$

$$e_s = 6.112 * exp \frac{(17.67 * T_{air})}{(T_{air} + 243.5)} \quad 4.26$$

$$R_H = \frac{e_a}{e_s} * 100 \quad 4.27$$

Where, e_s represents the saturation vapor pressure, T_{air} denotes the air temperature in Celsius, e_a signifies actual vapor pressure, Q_{air} stands for specific humidity (dimensionless), P_{surf} indicates pressure in millibars, and R_H represents relative humidity as a percentage. Temperature is converted to ^oC by subtracting 273.15 from the absolute temperature (K).

The implementation of pySEBAL requires data at two levels – instantaneous (during the time of Landsat overpass – 11.30 AM) and daily. It is important to note that the input data is in a raster format rather than points. Due to the unavailability of instantaneous in-situ data, GLDAS data is utilized. As discussed in section 3.2.4, there are 8 GLDAS data points for each day, with the first file corresponding to 12:00 AM GMT (05:30 AM IST). Since the files are available at three-hour intervals, the third file of the day aligns directly with the time of satellite overpass at 11:30 AM. Hence, it is used to fulfill the instantaneous data requirement for all variables.

For meeting the daily data requirement, both GLDAS and in-situ data are utilized, and the model is executed alongside the Aster DEM, resulting in two scenarios in turn leading to two sets of outputs. In our study we consider two scenarios generated using SEBAL and the third scenario considered pertains to the EElux estimates of actual evapotranspiration.

Scenario 1 – GLDAS daily rasters + GLDAS instantaneous rasters (at overpass) + Aster DEM

Scenario 2 – In-situ interpolated rasters + GLDAS instantaneous rasters (at overpass) + Aster DEM

Scenario 3 - EEFlux ET_a estimates

In scenario 1, as detailed in section 3.2.3, GLDAS files corresponding to the timeline for average temperature and relative humidity (08:30 AM and 05:30 PM) and the number of sunshine hours (05:00 AM to 7:00 PM) are acquired and averaged. They are labeled as "GLDAS daily rasters" in the definition above. In scenario 2, in-situ data collected throughout the day, starting from 08:30 AM and obtained at 05:30 PM, are utilized for temperature and relative humidity. Surface solar radiation is derived from the number of sunshine hours using the conversions outlined in Allen et al. (1998). To ensure consistency with the GLDAS data, a simple inverse distance weighted approach is employed to interpolate the in-situ data, with the spatial resolution set to 0.25° x 0.25° for both scenarios. These rasters collectively are labeled as "In-situ interpolated rasters". The GLDAS rasters pertaining to 11.30 AM IST, which is the time of the satellite overpass, are termed as the "GLDAS instantaneous rasters".

4.2.4 Structure of the Results section

The results are divided into five sections:

Section 5.1 – Land cover classification and analysis: This section explores the seasonal variations within a year (2006, 2014, 2021) between different seasons while also presenting accuracy metrics. The methodology is described using a flowchart in Figure 4.3.

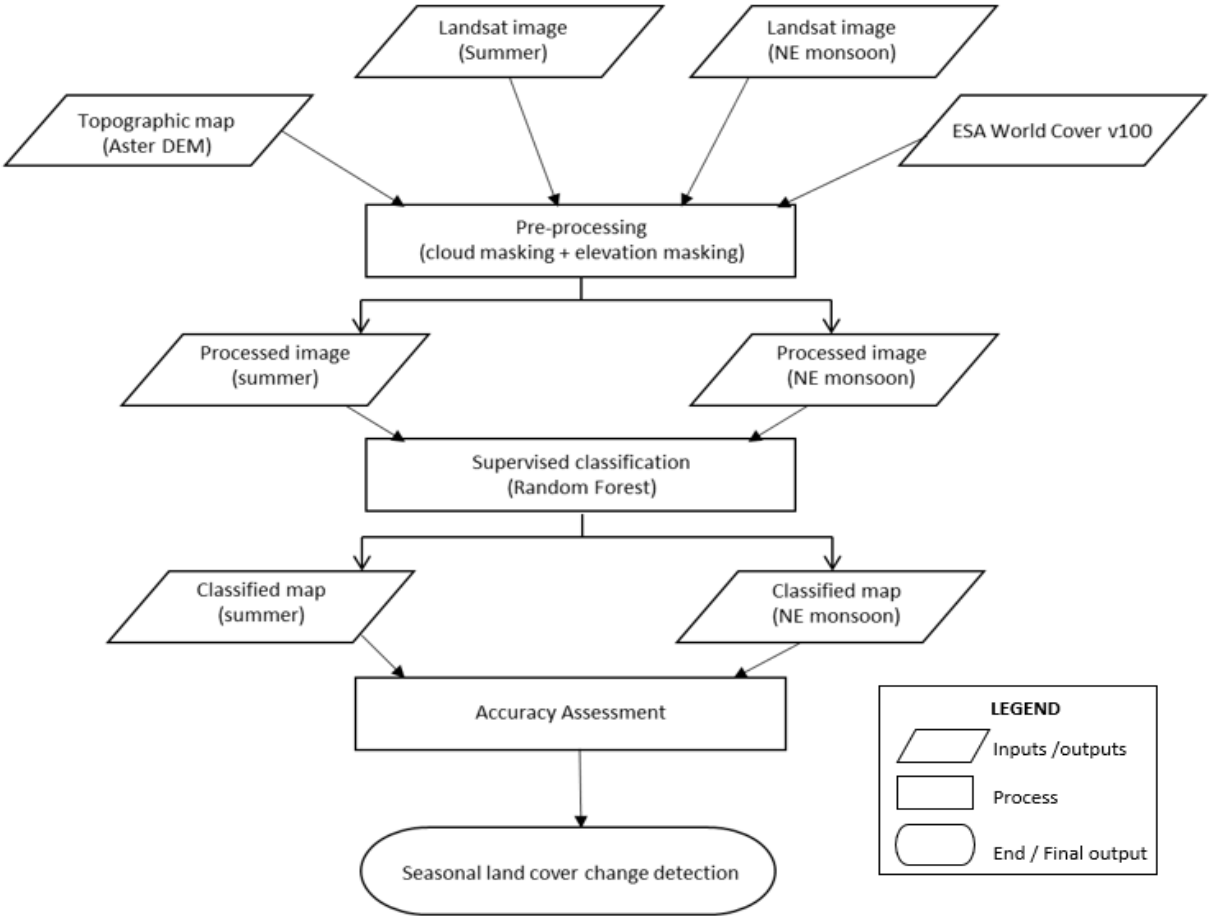

Figure 4.3 Methodology - Land cover classification and seasonal change detection

Section 5.2 – Comparative Analysis of Monthly Weather Data: In-Situ vs. GLDAS: This section presents a comparison between in-situ weather data and GLDAS weather data during the month of Landsat image acquisition across different years. Weather parameters such as average temperature, solar radiation, relative humidity, and reference evapotranspiration are compared.

The comparison is conducted for 6 months, with 2 months each in the years considered (2006, 2014, 2021). The comparison includes all weather stations, and the corresponding values in the GLDAS grid where the stations are located are utilized. The comparison is evaluated using the correlation coefficient, indicating the strength of agreement or association. It must be noted that the Penman-Monteith equation is utilized to calculate ET_o for the in-situ data, while the field 'Evap_tavg' representing evapotranspiration in the GLDAS data is employed. As the GLDAS data is available in $kg/m^2/s$, a conversion factor of 86,400 is multiplied to make it consistent with the in-situ ET_o estimate, which is presented in mm/day. The parameters utilized to estimate ET_o, namely average temperature, relative humidity, and solar radiation, are compared in this section to study their correspondence. Wind speed is excluded from the comparison due to a lack of data.

Section 5.3 - Interseason Comparison of Evapotranspiration Estimates Using SEBAL and EEFlux: ET_a maps generated using pySEBAL for all the scenarios are visualized. By overlaying classified land cover maps obtained earlier with scenario 1, scenario 2, scenario 3 maps, ET_a estimates pertaining to different land cover classes are arrived and plotted using box plots. The strength of the relationships between the three sets of outputs (scenario 1, scenario 2, scenario 3) taken two

at a time are quantified using the correlation coefficient, and the extent of the relationships is outlined using the root mean squared error metric (RMSE). This is followed by the analysis of mean estimates pertaining to the different land cover classes across the three scenarios. The metrics are outlined separately for mid-summer and the end of the Northeast Monsoon, followed by a comparison between the seasons within a year. The methodology for this section is described using a flowchart in Figure 4.4.

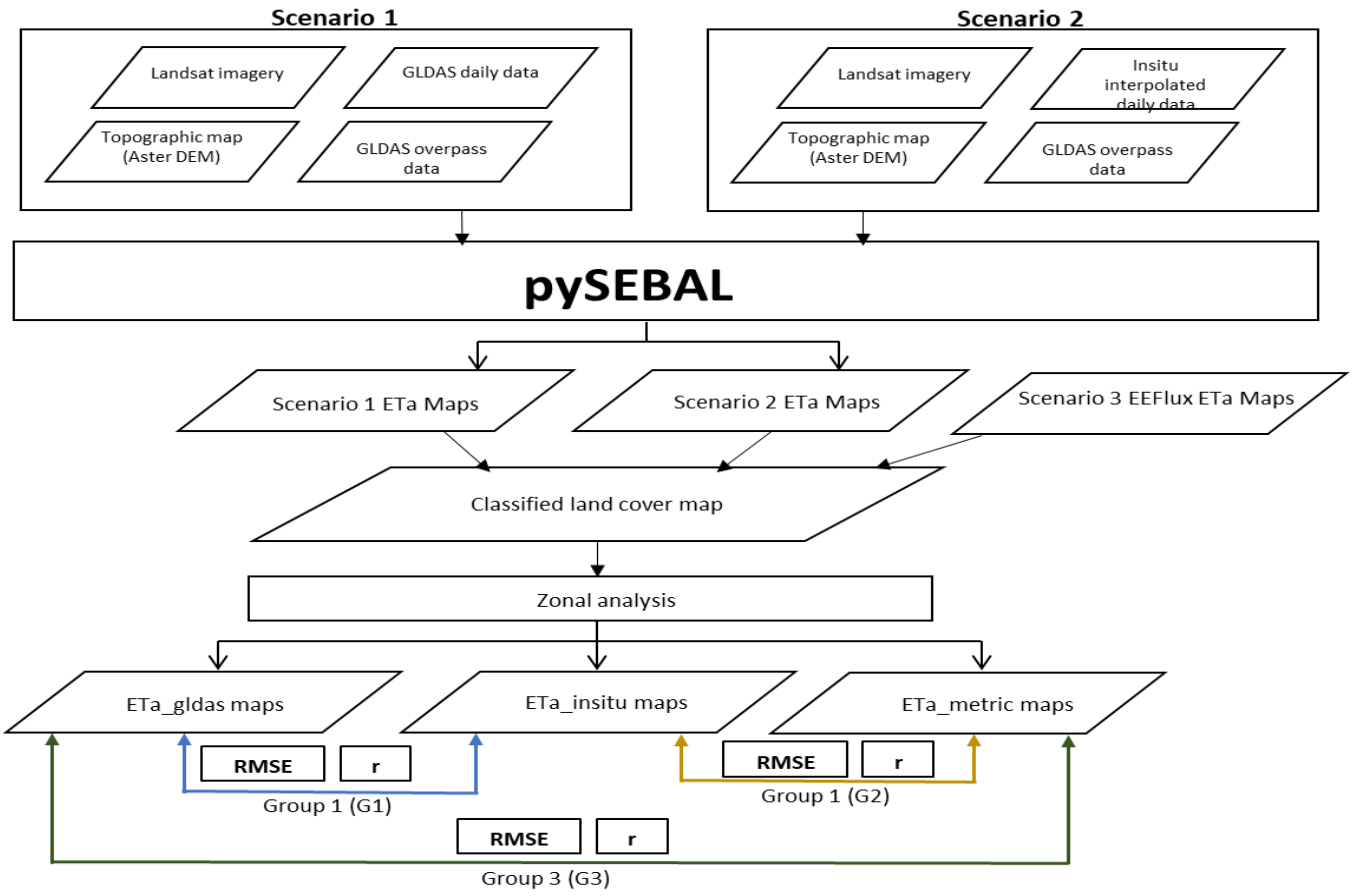

Figure 4.4 Methodology Interseason Comparison of Evapotranspiration Estimates Using SEBAL and EEFlux

Section 5.4 – Seasonal Variations in Actual Evapotranspiration Across Different Land Cover Classes: This section presents the seasonal variations of ET_a estimates across different land cover classes, using the components of the energy balance such as soil heat flux and sensible heat flux for explaining the seasonal variations. It also explores the reasons behind the seasonal variations and the causes for the uncertainty.

Section 5.5 – Estimating Water Outflux Variability Across Seasons and Years Across Different Land Cover Classes: The last step in the results section quantifies the water outflux (m^3/day) in the landscape across different year-season combinations. Patterns within the different years and between the seasons across different land cover classes will be discussed.

5 Results

5.1 Land cover classification and analysis

Land cover maps were generated using the supervised random forest algorithm. Seasonal change detection is performed within a year between the two selected seasons: summer and monsoon. This section is divided into three subsections corresponding to the three different years analyzed: 2006, 2014, and 2021.

Cloud mask layers were individually created for both images and applied to both the images corresponding to the summer and monsoon seasons within the year. These masks, along with the built-up area and elevation masks discussed in section 4.1, were incorporated into the analysis. This approach enables a change analysis of land cover classes within a year. In the accuracy metrics table, 'PA' denotes Producer Accuracy, 'UA' stands for User Accuracy, and the overall accuracy is provided at the end.

For each year, the percentage of the area covered by specific land cover classes is calculated during both seasons and compared to detect seasonal changes within a year. It is important to note that the percentage of the area is calculated based on the non-masked region within the study area, serving as the denominator.

Providing an overview of the land cover distribution, the upper regions of the study area serve as agricultural hubs, and they experience substantial growth during the monsoon season. Agriculture happens throughout the year across the study area, but with high levels of activity

during the monsoon. Despite the monsoon being a primary water source for irrigation, the presence of borewells and pumps in the study area enables cultivation cycles throughout the year.

The central portions as well as the areas in the west and northernmost regions of the study area, are characterized by lands with exposed soil. Some remain in their natural state, while others transition into agricultural lands during various months, predominantly during the monsoon. The barren land class can resemble the exposed soil in some cases and is found to have small prosopis plants growing within it.

Areas with prosopis and other dry species are often cleared with minimal vegetation remaining, as complete removal of those species is not possible. While scattered across the river basin, these barren lands are predominantly located in the western parts of the study area. Given the presence of agricultural lands in all regions of the basin, the pre-growth land cover class, indicating tilled or harvested lands, is prevalent across the entire basin but is concentrated primarily in the western parts of the study area. Prosopis patches are predominantly situated in the southern parts of the study area, growing in proximity to the tanks. Water storage primarily occurs through tanks, as discussed earlier, and spectral variations within the study area necessitate two distinct classes – water and turbid water.

5.1.1 Seasonal land cover analysis - 2006

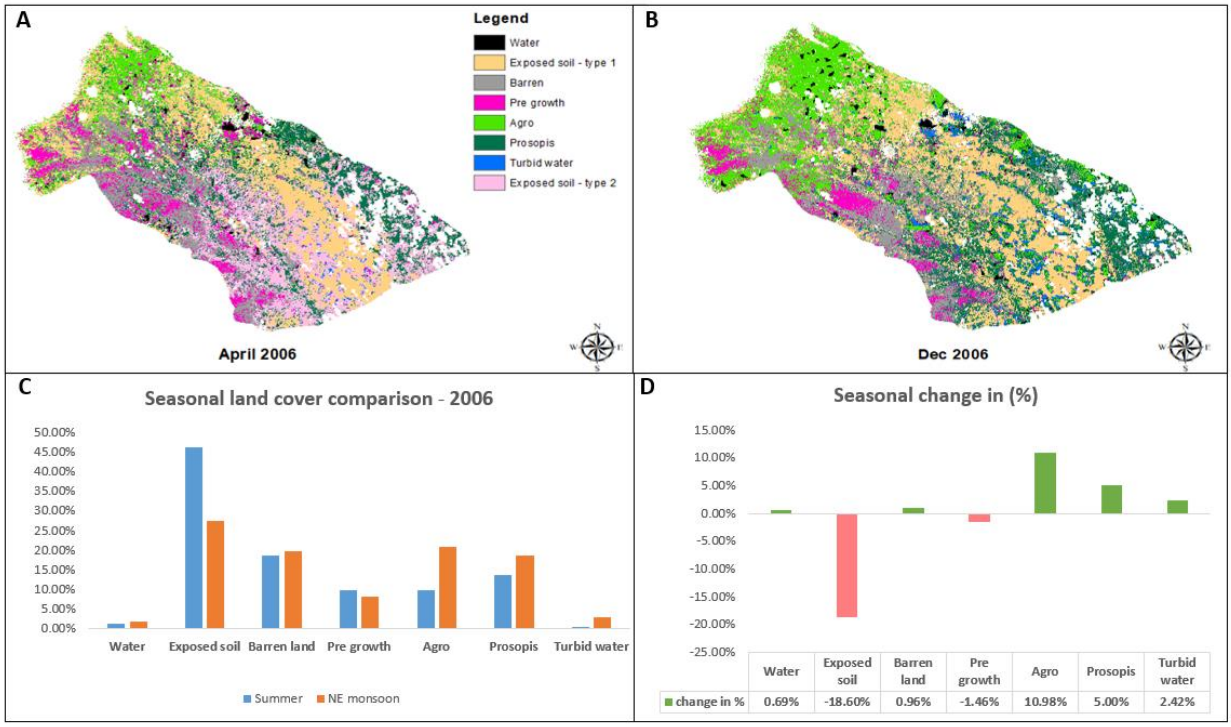

Figure 5.5.1 Land cover classification maps (a, b) for summer and Northeast Monsoon, seasonal comparison of land cover classes (c) in 2006, and percentage change (d) in the Upper Gundar River Basin, Tamil Nadu

Table 5.1 Accuracy metrics 2006

Season	**Summer**		**NE monsoon**	
Class	**PA**	**UA**	**PA**	**UA**
Water	99%	99%	99%	99%
Exposed soil - type 1	96%	96%	97%	98%
Barren land	96%	91%	96%	95%
Pre growth	93%	96%	96%	97%
Agro	98%	97%	97%	99%
Prosopis	97%	97%	99%	96%
Turbid water	98%	99%	99%	98%
Exposed soil - type 2	95%	95%	-	-
Overall accuracy	96%		97%	

The primary challenge encountered in the seasonal analysis within the year is the presence of clouds in various parts of the images, resulting in the exclusion of certain areas from the classification process. High values for both producer and user accuracy are observed across all classes, with instances of water classes achieving a near 100% accuracy level. This could be attributed to the ease of identification and distinct appearance of water classes in the various false-color composites used during analysis (Table 5.1).

The increase in agriculture is seen in the northern parts of the images where a lot of green areas seem to have popped up during the monsoon season. Small patches can also be seen across the basin (Figure 5.1a-b). Incremental changes of 0.69% and 2.42% are evident in the water classes, and the rise can be attributed to the Northeast Monsoon – the primary source of rainfall in the region that replenishes the tanks. Notably, blue regions signifying water become more prominent in the lower section of the study area, which is home to a substantial number of tanks, as discussed in previous sections (Figure 5.1b).

A substantial decrease of 18.60% is noticeable in the exposed soil class, corresponding with the increased agricultural activity during the monsoon season, where these lands are frequently utilized for cropping. The barren land class shows a slight increase of 0.96%, ensuring consistency in the analysis but the rise can be possibly indicating their misclassification due to their proximity to the exposed soil class and the pre-growth class. A marginal decline of 1.46% is evident in the pre-growth class (Figure 5.1c-d).

Most notably, a substantial surge of 10.98% is observed in the agriculture class, reflecting the heightened agricultural practices during this period. The increase in the prosopis class in this case can similarly be attributed to increased rainfall during the monsoon providing water for the growth of the plant (Figure 5.1c-d).

Mapping prosopis proves to be a challenging task, given its spectral similarity to agricultural classes. While discrimination could be more effective in summer, the surge in agriculture during the monsoon season can lead to the misclassification of prosopis pixels as representing agriculture or general vegetation (Figure 5.1c-d).

5.1.2 Seasonal land cover analysis - 2014

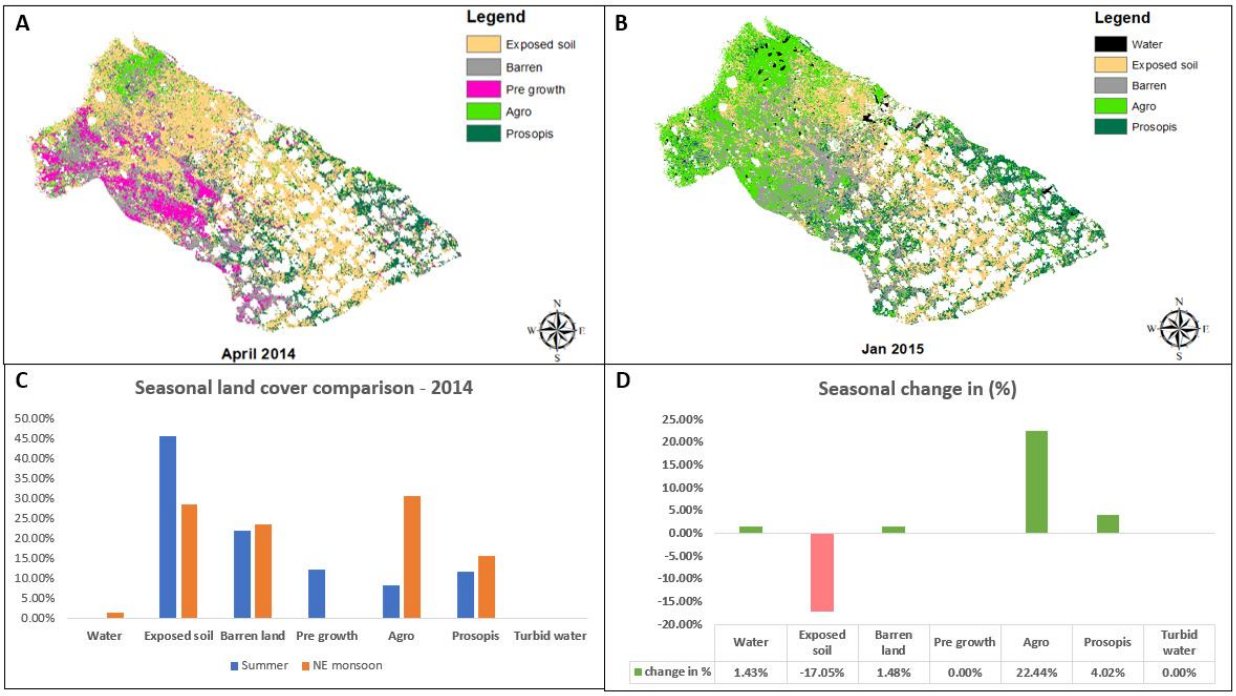

Figure 5.5.2 Land cover classification maps (a, b) for summer (2014 Apr) and Northeast Monsoon (2015 Jan), seasonal comparison of land cover classes (c), and percentage change (d) in the Upper Gundar River Basin, Tamil Nadu

Table 5.2 Accuracy metrics 2014

Season	**Summer**		**NE monsoon**	
Class	**PA**	**UA**	**PA**	**UA**
Water	-	-	98%	100%
Exposed soil - type 1	97%	98%	97%	98%
Barren land	95%	95%	98%	96%
Pre growth	95%	95%	-	-
Agro	94%	94%	94%	98%
Prosopis	98%	98%	98%	95%
Turbid water	-	-	-	-
Overall accuracy	96%		97%	

In the month of April this year, the images revealed minimal to no presence of water bodies, a phenomenon attributable to the shifting monsoon and groundwater flow patterns (Figure 5.2a). Also, in January, the pre-growth class was seldom observed, and the tanks in the lower reaches of the river basin remained unidentified due to the dense cloud cover within the image, in those parts (Figure 5.2b). Hence, these classes are excluded from our analysis for this particular year. Moreover, the images available during the end of the monsoon season exhibited high percentages of clouds and cloud shadows in the lower regions of the area, posing challenges for seasonal change detection.

In terms of spatial distribution, agriculture seems to be popping up more during January in the upper parts of the river basin, which is an expected outcome (Figure 5.2b). Consequently, this section of the analysis primarily focuses on investigating discernible patterns within the given images, considering the limitations posed by cloud cover. While a comprehensive change detection process for all classes is unfeasible, notable observations include a substantial increase in agriculture and a corresponding decrease in the exposed soil class (Figure 5.2c-d).

5.1.3 Seasonal land cover analysis - 2021

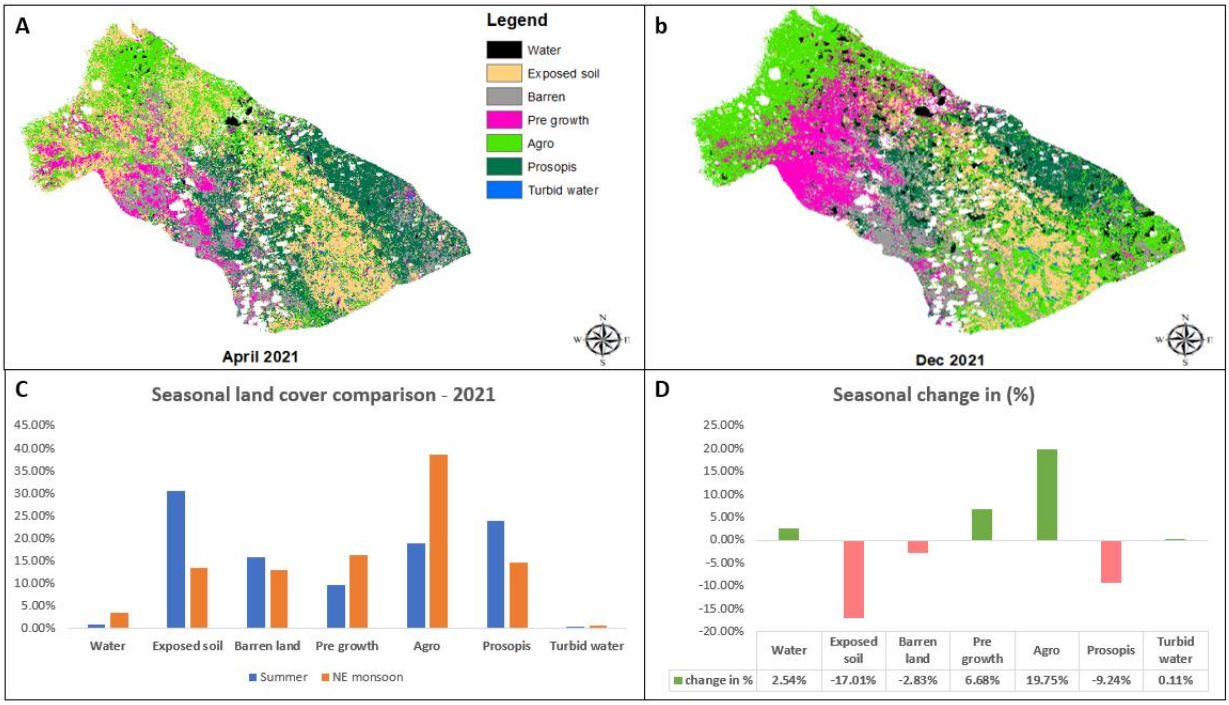

Figure 5.5.3 Land cover classification maps (a, b) for summer and Northeast Monsoon, seasonal comparison of land cover classes (c) in 2021, and percentage change (d) in the Upper Gundar River Basin, Tamil Nadu

Table 5.3 Accuracy metrics 2021

Season	Summer		NE monsoon	
Class	**PA**	**UA**	**PA**	**UA**
Water	99%	100%	99%	100%
Exposed soil - type 1	99%	100%	99%	98%
Barren land	97%	96%	99%	98%
Pre growth	98%	98%	98%	99%
Agro	99%	99%	98%	99%
Prosopis	98%	98%	100%	98%
Turbid water	99%	100%	99%	100%
Overall accuracy	99%		99%	

With these images displaying a relatively higher degree of freedom from cloud cover compared to previous datasets, a robust seasonal analysis becomes feasible. Observable trends include increases in water classes (2.54% and 0.11%) as seen in Figure 5.3d, indicating the filling of tanks in both upper and lower parts of the basin (Figure 5.3b). Consistent with the patterns presented above, a substantial decrease of 17.01% in exposed soil classes and a slight reduction in barren land classes are evident (Figure 5.3d).

These changes align with increases in pre-growth and agriculture classes (6.68% and 19.75%), again pointing towards the increased levels of agricultural activities during the monsoon season. The decrease in prosopis by 9.24% this year may be associated with its overlap with the agriculture class and the practice of cutting down prosopis trees for economic purposes and sustenance (refer to Figure 5.3d).

In conclusion, despite seasonal variations in the landscape and the presence of clouds and cloud shadow regions in different parts of the images during various seasons, significant increases in agriculture classes – 10.98%, 22.44%, and 19.75% – corresponding with notable decreases in the exposed soil class during the monsoons of all the years, substantiate the prevalence of high agricultural activity during this season.

5.2 Comparative Analysis of Monthly Weather Data: In-Situ vs. GLDAS

In this section, average temperature, solar radiation, relative humidity, and reference evapotranspiration are labeled as 'Avg temp', 'Solar rad', 'Rh', and 'ET_o', respectively, and are compared using scatter plots and correlation coefficients. In this section, it is noted that April corresponds to mid-summer, while December and January indicate the end of the Northeast Monsoon period. For convenience, one plot (Figure 5.4) corresponding to the April 2006 season and one table (Table 5.4) containing correlation coefficients for various weather parameters across the six year-month combinations are presented.

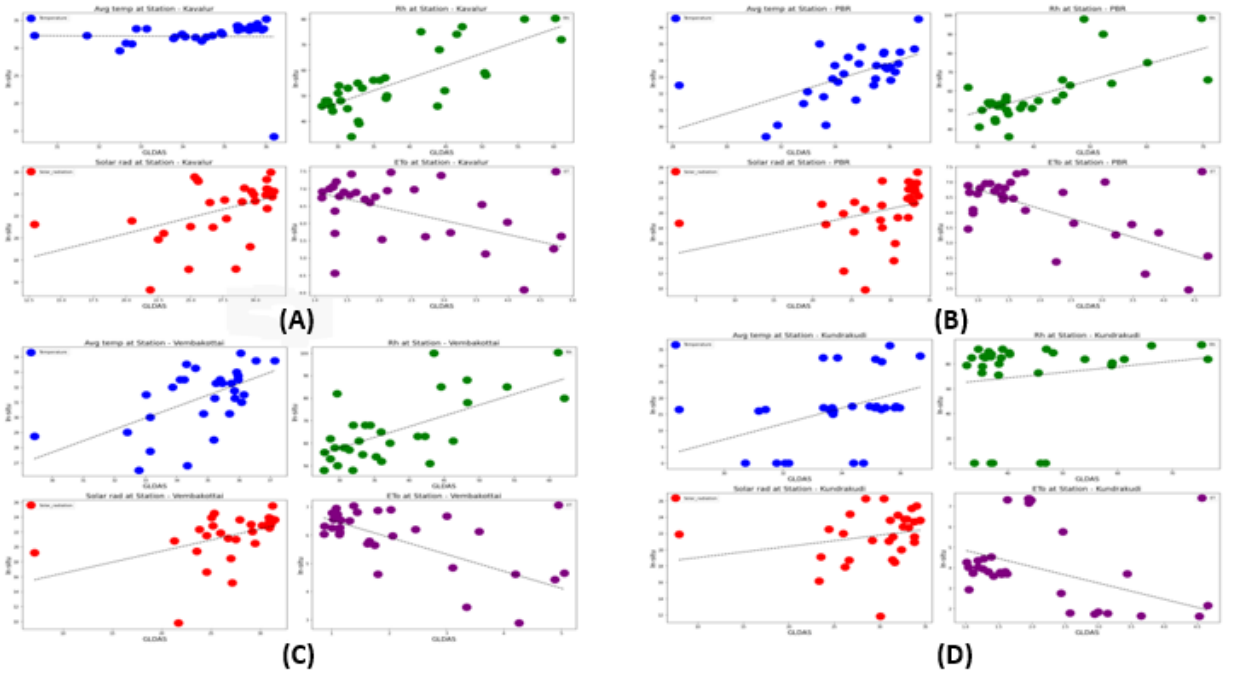

Figure 5.4 Scatter plots containing comparisons average temperature(°C), relative humidity, solar radiation (MJ m^{-2} day^{-1}), reference evapotranspiration (starting from left in a clockwise manner) for stations A) Kavalur B) PBR C) Vembakottai D) Kundrakudi

Table 5.4 **Correlation Coefficients Between In-Situ and GLDAS Data for Meteorological Variables**

Year - Season	Station	Avg temp	Solar rad	Rh	ET_o
2006 - April	Kavalur	-0.01	0.46	0.74	-0.53
2006 - April	PBR	0.61	0.36	0.64	-0.71
2006 - April	Kundrakudi	0.43	0.23	0.15	-0.45
2006 - April	Vembakottai	0.58	0.46	0.61	-0.72
2006 - December	Kavalur	0.19	0.83	0.73	0.48
2006 - December	PBR	0.2	0.74	0.44	0.43
2006 - December	Kundrakudi	0.18	0.58	-0.04	0.48
2006 - December	Vembakottai	0.28	0.77	0.48	0.27
2014 - April	Kavalur	0.28	0.71	0.69	-0.23
2014 - April	PBR	0.46	0.8	0.38	-0.47
2014 - April	Kundrakudi	0.43	0.68	0.26	-0.55
2014 - April	Vembakottai	0	0.03	0	-0.24
2014 - April	Savaspuram	0.38	0.78	0.57	-0.45
2015 - January	Kavalur	0.6	0.81	0.58	0.12
2015 - January	PBR	0.73	0.54	0.53	-0.04
2015 - January	Kundrakudi	-0.1	0.78	0.37	0.28
2015 - January	Vembakottai	0.26	0.1	-0.09	0.18
2015 - January	Savaspuram	0.37	0.76	0.57	0.06
2021 - April	Kavalur	0.88	0.85	0.57	-0.19
2021 - April	PBR	0.74	0.72	0.47	-0.01
2021 - April	Kundrakudi	0.77	0.76	0.04	-0.06
2021 - April	Vembakottai	0.68	0.72	0.29	0.15
2021 - April	Savaspuram	0.83	0.75	0.63	-0.17
2021 - December	Kavalur	0.62	0.9	0.84	0.34
2021 - December	PBR	0.63	0.37	0.64	-0.15
2021 - December	Kundrakudi	0.5	0.79	0.62	0.21
2021 - December	Vembakottai	0.36	0.56	0.79	0.22
2021 - December	Savaspuram	0.76	0.56	0.78	0.03

It has to be noted that cells with values more than 0.7 are marked in green across average temperature, solar radiation, relative humidity. Based on the plots (Figure 5.4) and Table 5.4 corresponding to April 2006, it is evident that all stations, except Kavalur, demonstrate positive relationships with the GLDAS data across the different variables.

However, outliers are observed within all stations and across all variables, which may be attributed to missing data or errors in data collection, particularly noticeable in the Kundrakudi station. For solar radiation, all stations exhibit positive relationships, albeit with correlation values below 0.5. In terms of relative humidity, Kundrakudi stands as an exception, while other stations exhibit higher correlation values. Additionally, all stations show strong negative correlations in reference evapotranspiration (ET_o) estimates. Similar observations are made for other year-month combinations. Most cases show positive correlations for relative humidity and temperature, with varying magnitudes ranging from values close to 0 and 1. It should also be noted that there is no data available for the Savaspuram station in April 2006 and December 2006, as well as for Vembakottai station in April 2014 for average temperature and relative humidity.

Solar radiation exhibits positive values across all scenarios. The cases pertaining to average temperature, relative humidity and solar radiation across years exhibiting values greater than 0.7 are highlighted in green. However, a pattern emerges with respect to reference evapotranspiration estimates, with a predominant negative relationship (marked in red) observed in April and a mostly positive relationship (marked in light green) in December/January for all years considered. It is important to reiterate that April signifies summer, while December or January refer to the end of the monsoon season. The key takeaways from this section suggest that the datasets are comparable, and GLDAS, based on climate models, is compatible in this region. However, outliers are present, indicating potential discrepancies in the data.

5.3 Interseason Comparison of Evapotranspiration Estimates Using SEBAL and EEFlux

In this section, we will conduct an interseason comparison of ET_a outputs within a year, focusing on different seasons and examining one year at a time. The primary unit for comparison will be the land cover class. During comparisons, the agro and prosopis classes are commonly referred to as the 'vegetation classes,' while the remaining classes are termed as the 'non-vegetation classes.' Within a given season, intermodal comparisons are conducted for each land cover class by considering two scenarios at a time. These comparisons are quantified using the correlation coefficient and root mean squared error (RMSE) in mm/day. With 3 scenarios in total, this analysis results in three groups, each consisting of two sets of values. The scenarios correspond to the ones discussed in section 4.2.4.

Later, the trends and patterns in ET_a mean values among the land cover classes per image across the three scenarios will be discussed. For the ease of the reader, descriptions of the scenarios and groups are tabulated below in Tables 5.5 and 5.6.

Table 5.5 Scenarios for ET_a mean comparisons

Scenario	Description
Scenario 1	GLDAS weather rasters + GLDAS raster (at overpass) + Aster DEM
Scenario 2	In-situ interpolated rasters + GLDAS raster (at overpass) + Aster DEM
Scenario 3	EEFlux portal ET_a estimates

Table 5.6 Groups for intermodal comparisons

Group	Description
Group 1 (G1)	Scenario 1 vs Scenario 2
Group 2 (G2)	Scenario 2 vs Scenario 3
Group 3 (G3)	Scenario 3 vs Scenario 1

To recap, scenario 1 involves ET_a outputs produced using SEBAL with GLDAS data as meteorological forcing. Scenario 2 comprises outputs generated using SEBAL with interpolated in-situ weather data as the meteorological forcing, while scenario 3 consists of outputs from the EEFlux portal generated using the METRIC algorithm. Box plots are generated for various land cover classes corresponding to all three scenarios mentioned above. The content in this section and the naming conventions used in the maps and figures presented within are detailed below.

It should be noted that in terms of the scenarios, scenario 2 is considered as the most important scenario as in-situ data is considered more representative of actual conditions compared to the GLDAS model data. Extending this to the groups, G2 is considered as the most important group because values obtained using in-situ data are compared with outputs from EEFlux, which represents a completely independent dataset. Scenario 1 and scenario 2 are closely related as they both utilize the same dataset to fulfill the instantaneous data requirement.

Maps: Every year is presented with a map containing 6 panels. It is split vertically, with the three panels on the left pertaining to the summer season and the three on the right pertaining to the monsoon season.

Within a season, the first panel corresponds to scenario 1 labeled ‘Gldas SEBAL’, the second panel to scenario 2 labeled 'Insitu-interp SEBAL’, and the third panel to scenario 3, labeled ‘METRIC’.

Figures: In this section, two figures are presented for every year, each consisting of a box plot where the unit of comparison is a land cover class. For each land cover class, three box plots are generated corresponding to the three scenarios mentioned above. They are labelled as ETa_gldas, ETa_insitu, and ETa_metric, respectively. The box plots are color-coded, and a legend is provided as a supplement.

Tables: Additionally, intermodal comparisons are conducted for each land cover class by considering two scenarios at a time, with the results outlined per group as per Table 5.5. These comparisons are presented in tables alongside the box plot for the reader’s convenience. They are color-coded, and a legend is provided for clarity.

First, the results for the year 2006 are presented, including maps, plots, and tables, along with corresponding observations. Subsequently, specific maps, figures, and tables for 2014 and 2021 are provided. Consolidated observations and conclusions are reported based on the combined analysis of all three years, considering the shared patterns observed between the seasons within the different years.

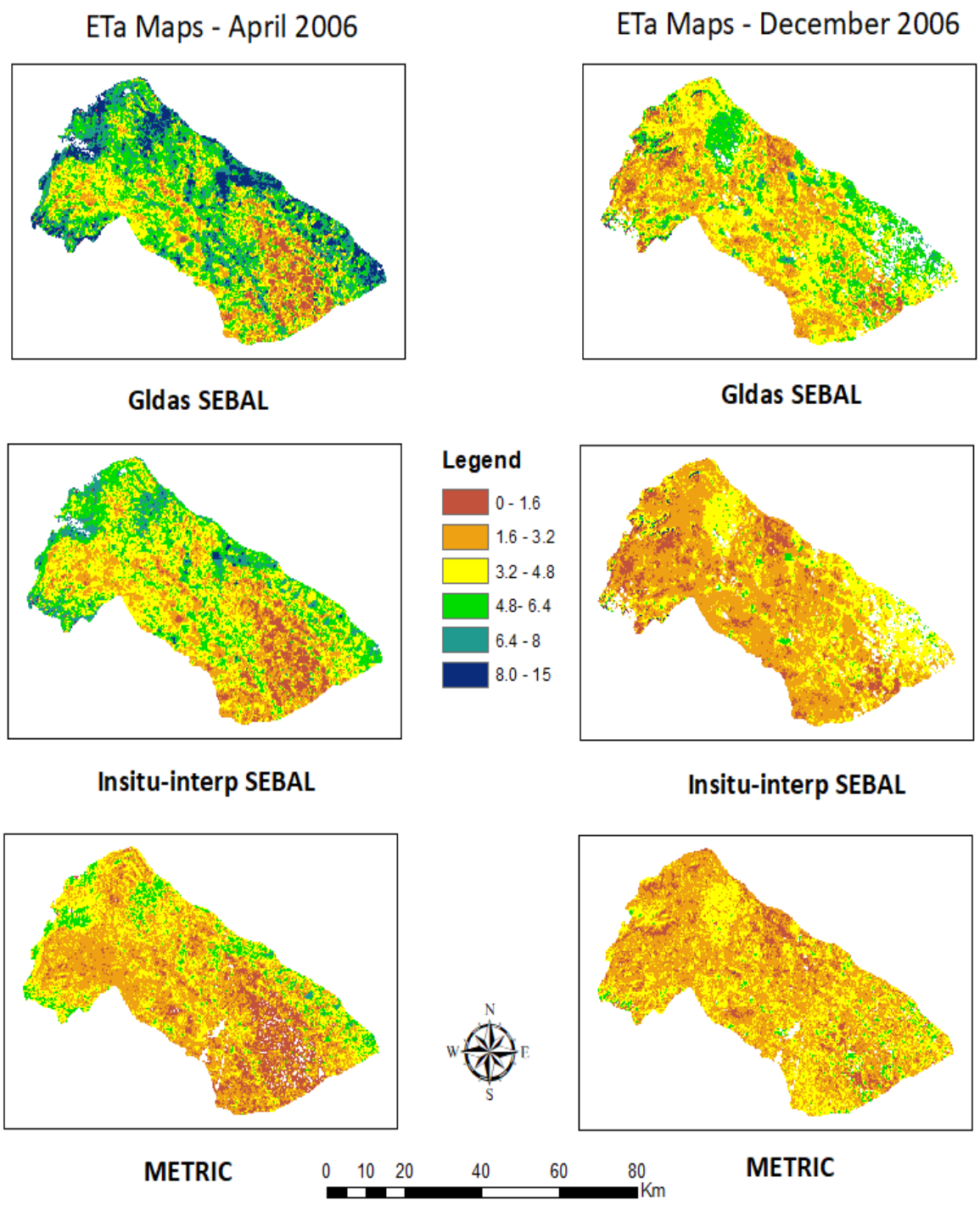

Figure 5.5 Actual Evapotranspiration maps (mm/day), 2006

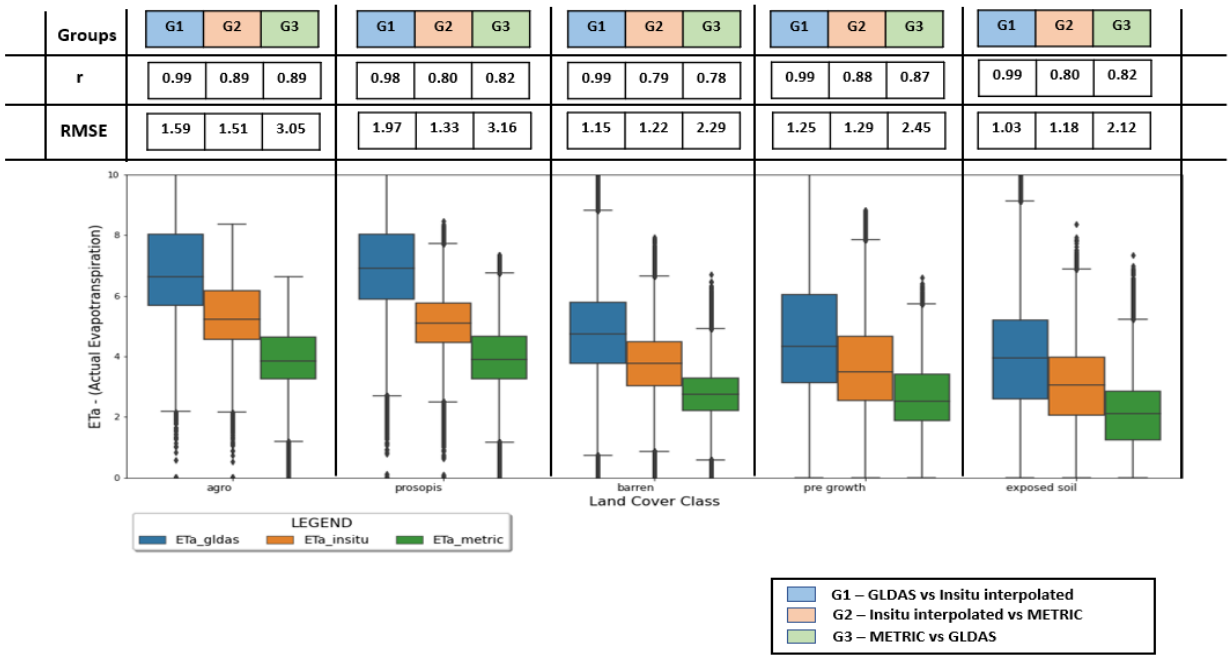

Figure 5.6 ET_a comparison across scenarios and land cover classes, 2006 Summer

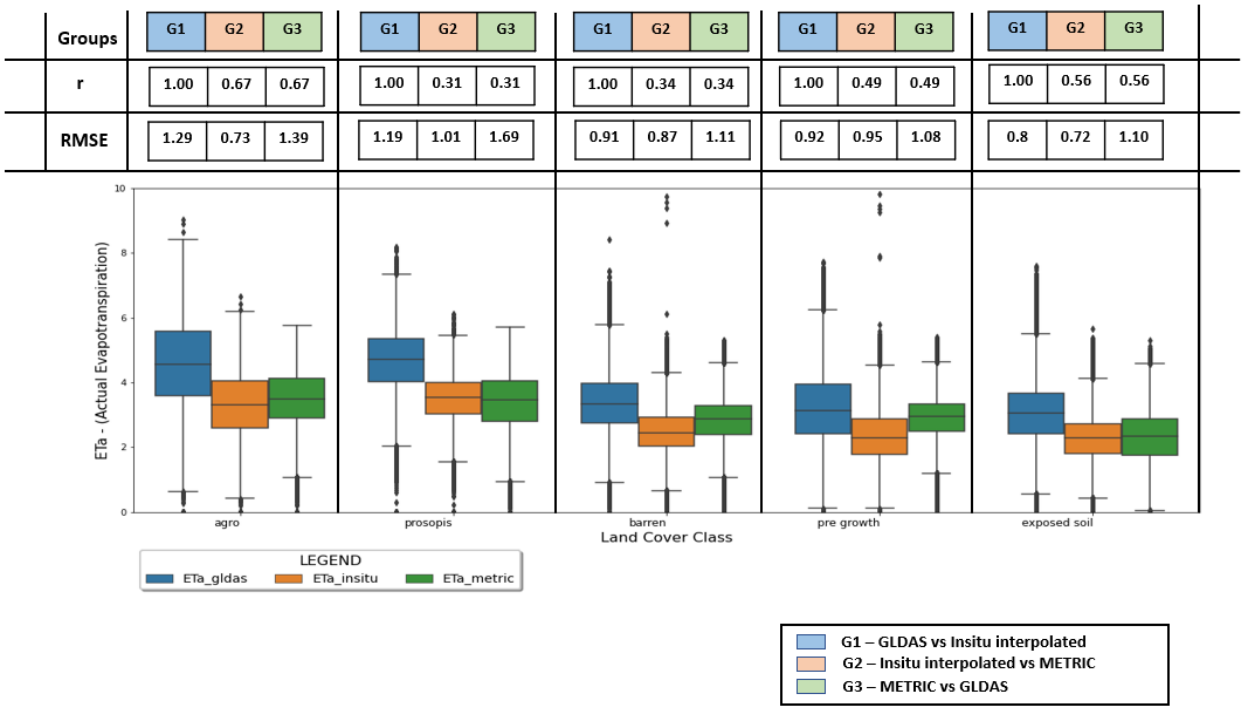

Figure 5.7 ET_a comparison across scenarios and land cover classes, 2006 Monsoon

Strong levels of agreement, as indicated by the correlation coefficient, are found in G1 for all land cover classes. This high correlation can be attributed to the determination of instantaneous heat fluxes and subsequently instantaneous evaporative fraction (EF_inst) described in Equations 4.25-4.27 using instantaneous data, which is common between scenario 1 and scenario 2. The difference between the two scenarios is primarily reflected in the final computation stage where net radiation in a day (Rn_24) is multiplied with EF_inst, making Rn_24 the most important meteorological parameter in the daily data, as confirmed earlier. This explains why the correlation value is close to 1 and highlighting the importance of Rn_24 in determining the final ET_a estimate.

Similar levels are observed between the other two groups, G2 and G3, in both seasons (Figures 5.6-5.7). During summer, the lowest RMSE values for the non-vegetation classes (barren, exposed soil, pre-growth) are observed in G1, while for vegetation classes (agro and prosopis), they are observed in G2 (Figure 5.6). For the monsoon season, except for the pre-growth class, the lowest RMSE values for the classes are observed in G2, with the RMSE value for pre-growth in G2 exceeding that of G1 by 0.03 mm/day (Figure 5.7). For all classes, the highest RMSE values are observed in G3 in both seasons (Figures 5.6-5.7).

Lower RMSE values imply a higher degree of concordance between the datasets, while the highest value signifies a greater magnitude of divergence. It is shown that the in-situ interpolated SEBAL values lie in proximity with the outputs from EEflux seen in G2 (Figures 5.6-5.7).

In this study, we consider G2 as the most important group for analysis as the in-situ-interpolated values are compared with the values generated from METRIC. In G2, in summer, the lowest value pertains to the exposed soil class with 1.18 mm/day, while the highest corresponds to the agro class with 1.51 mm/day (Figure 5.6). In the monsoon, the lowest value pertains to the exposed soil class, and the highest to the prosopis class (Figure 5.7).

When RMSE is averaged across all classes in G2, we get a value of 1.30 mm/day in summer and 0.86 mm/day in monsoon. When the correlation coefficient is averaged, the value is 0.84 in summer and 0.47 in the monsoon season (Figure 5.6-5.7).

The decrease in correlation coefficient in the monsoon indicates a decrease in the linear relationship between the datasets. However, a decrease in RMSE indicates the convergence of the predictions during the wet season between the datasets. From Figures 5.6-5.7, in terms of mean values across the different scenarios, both in summer and monsoon, the highest values are observed in scenario 1 (colored in blue), pertaining to SEBAL outputs from the GLDAS data. In summer, scenario 2 (colored in brown) exhibits higher values compared to scenario 3 (colored in green) whereas in the monsoon season, the values are closer.

Focusing on scenario 2, our focus group, in both seasons, it can be observed that the non-vegetation classes collectively display lesser values compared to vegetation classes, and the lowest value is always exhibited by the exposed soil class, consistent with the class definition as it is expected to be the driest class among those selected for our analysis.

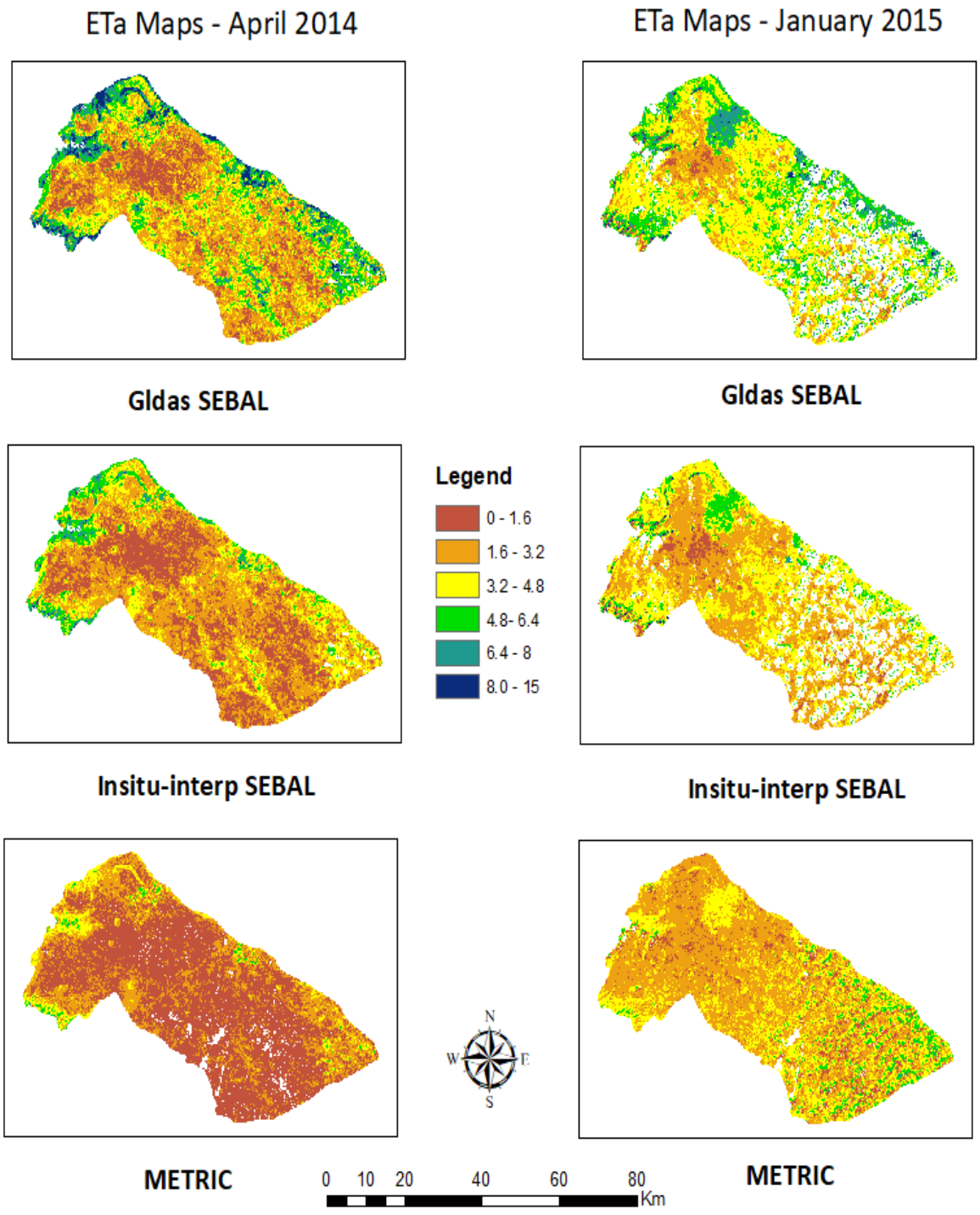

Figure 5.8 **Actual Evapotranspiration maps (mm/day), 2014**

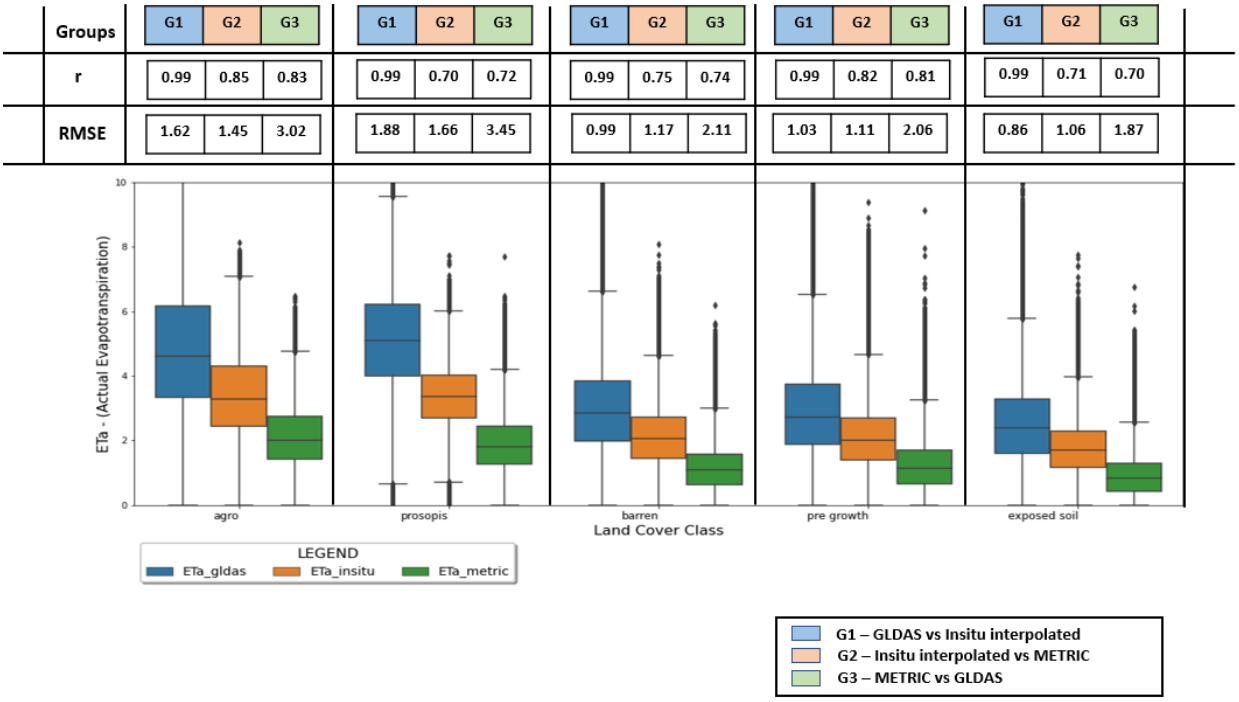

Figure 5.9 ET_a comparison metrics across scenarios and land cover classes, 2014 Summer

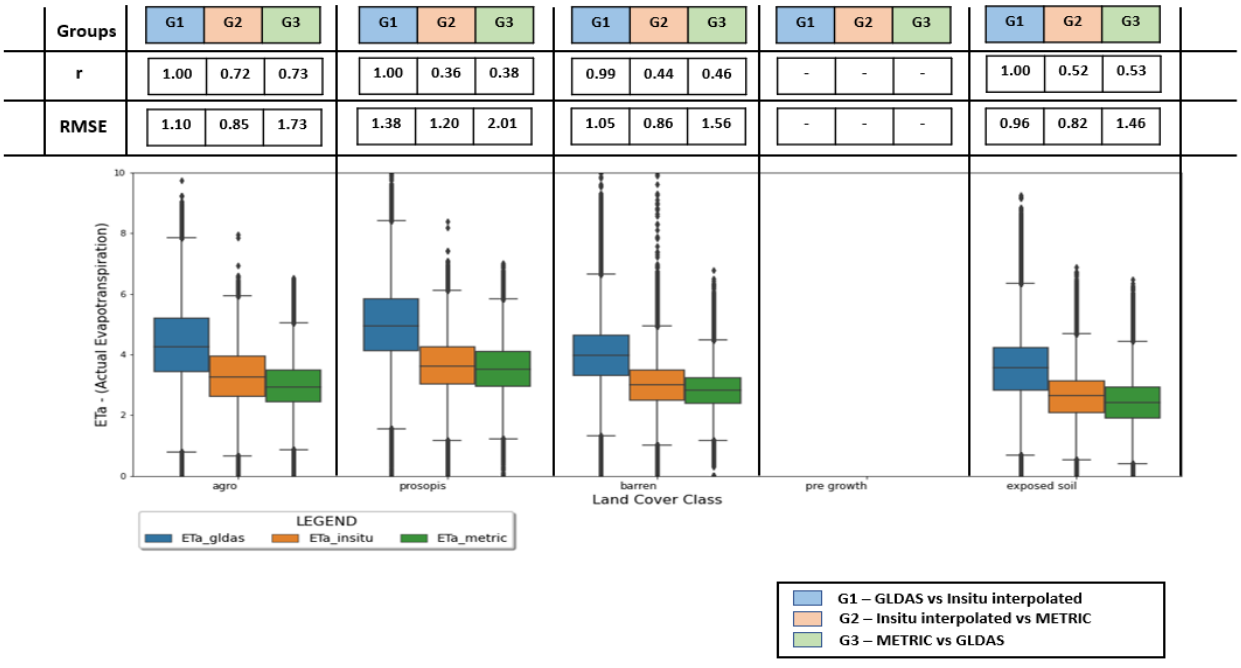

Figure 5.10 ET_a comparison metrics across scenarios and land cover classes, 2014 Monsoon

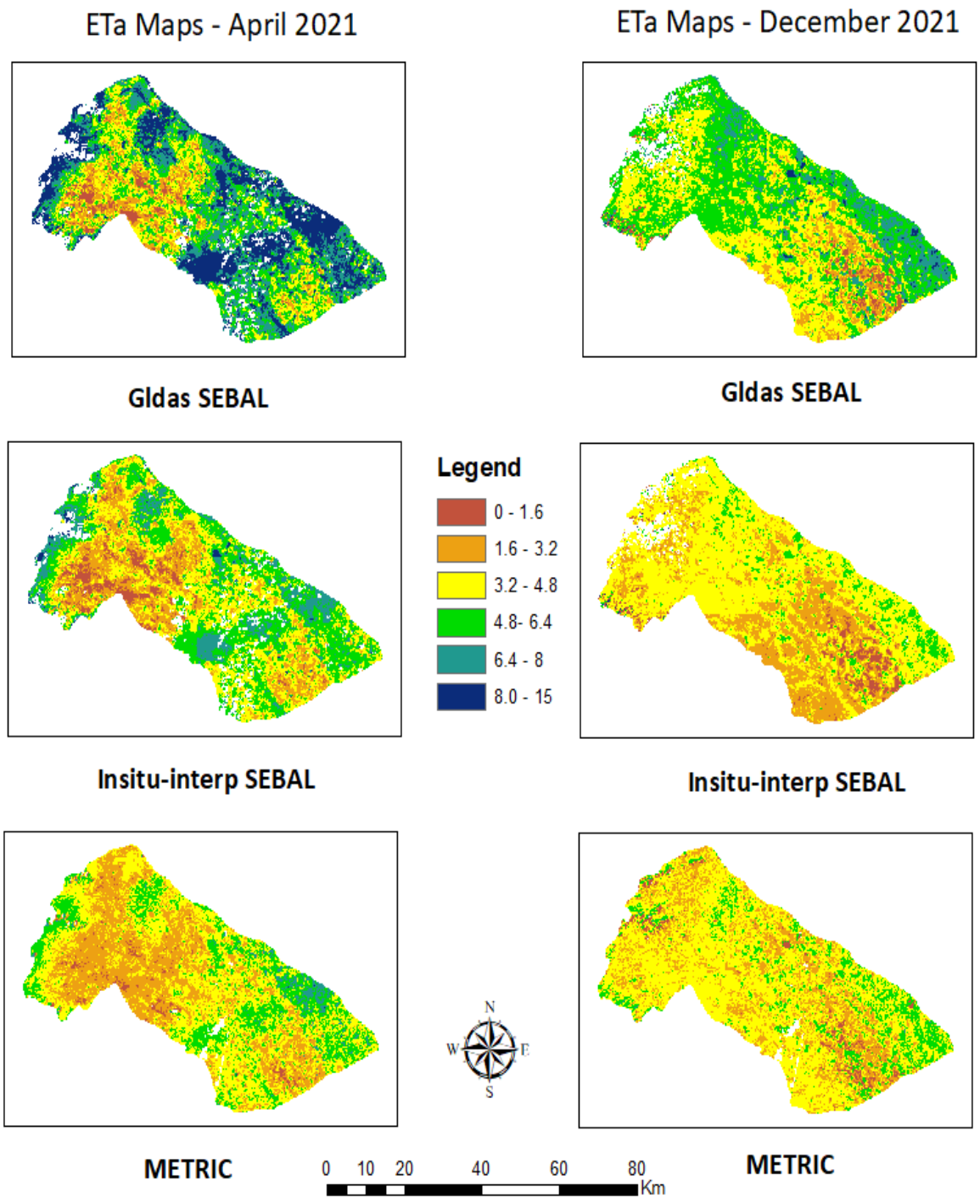

Figure 5.11 Actual Evapotranspiration maps (mm/day), 2021

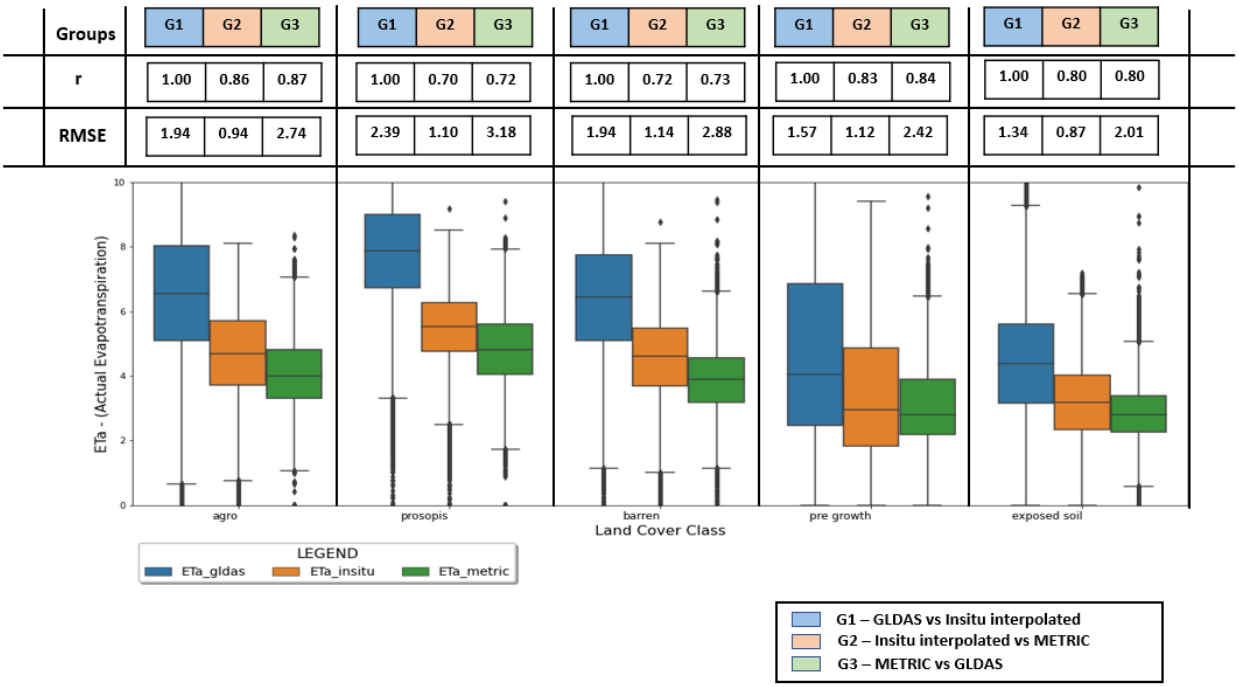

Figure 5.12 ET_a comparison across scenarios and classes, 2021 Summer

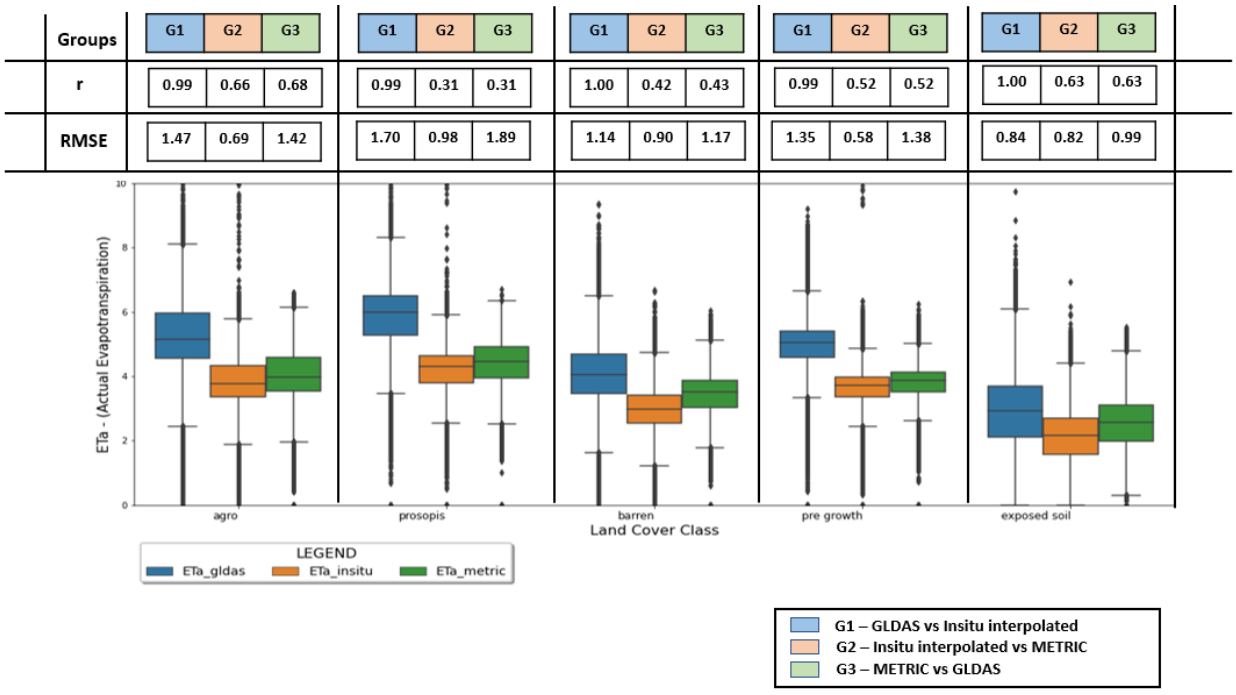

Figure 5.13 ET_a comparison across scenarios and classes, 2021 Monsoon

In 2014, the summer season follows the same patterns as the 2006 summer, with the non-vegetation classes exhibiting the least values in G1 and the vegetation classes in G2 when considering RMSE (Figure 5.6 & Figure 5.9). The remaining combinations (2014 Monsoon, 2021 Summer, and 2021 Monsoon) exhibit similar RMSE patterns, with all classes showing the least RMSE in G2 and the highest in G3 (Figure 5.10, Figure 5.12, and Figure 5.13).

Within each year between the seasons, in 2014 and 2021, similar to 2006, a decrease in both correlation coefficient and RMSE is observed for all classes across all groups (Figures 5.6-5.7, Figures 5.9-5.10, and Figures 5.12-5.13).

When considering RMSE in G2, the most important group across classes, the exposed soil tends to exhibit the least values in 2014 summer and 2021 summer, similar to the patterns observed in 2006 summer and 2006 monsoon. In terms of the highest RMSE in G2, the prosopis class exhibits the highest variability in all occasions except for 2006 summer and 2021 summer. Therefore, it can be concluded that in terms of variability between the in-situ interpolated SEBAL values and EEFlux values, the exposed soil and prosopis classes demonstrate the least and highest variability, respectively (Figures 5.6-5.7, Figures 5.9-5.10, and Figures 5.12-5.13).

When G2 is considered and after averaging RMSE and correlation coefficient, as observed in 2006, higher values are observed for both metrics in summer compared to the monsoon season, as tabulated in Table 5.7. This again demonstrates the increased proximity and reduced variability between the in-situ interpolated outputs and EEFlux outputs in the monsoon compared to summer.

Table 5.7 Averages of correlation coefficient (r) and RMSE across land cover classes in different seasons and years

Year - Season	Average r	Average RMSE (mm/day)
2006 Summer	0.84	1.30
2006 Monsoon	0.47	0.86
2014 Summer	0.77	1.29
2014 Monsoon	0.51	0.90
2021 Summer	0.78	1.03
2021 Monsoon	0.51	0.79

In terms of mean values across all years and seasons, scenario 1 consistently exhibits the highest value among all scenarios. Throughout the summer season in all years, the means of scenario 3 tend to be lower than those of scenario 2, while in the monsoon, they remain similar. This observation aligns with the findings presented in Table 5.7, where RMSE values are lower in the monsoon compared to summer in G2. Regarding the classes, once again, the pattern of the exposed soil class exhibiting the lowest mean value and either of the vegetation classes exhibiting the highest value remains consistent for the remaining years as well (Figures 5.6-5.7, Figures 5.9-5.10, and Figures 5.12-5.13).

Since EEFlux is a widely used and well-developed dataset created through a joint effort spanning multiple organizations, this analysis highlights that in-situ interpolated SEBAL values exhibit good consistency, with RMSE values close to 1 mm/day during summer and even better performance in the monsoon season with values less than 1 mm/day (Table 5.7). This indicates the usefulness of this project. However, the absence of site measurements using lysimeters and flux towers presents a limitation. Validation with proper field data would provide further insight into the utility of these datasets (Allen et al., 2015).

5.4 Seasonal Variations in Actual Evapotranspiration Across Different Land Cover Classes

This section describes the seasonal variations using the mean value as a reference for all land cover classes, focusing on scenario 2 containing ET_a estimates from SEBAL forced using in-situ interpolated data. Through this section, we also explore the influence of the different components of the energy balance equation that are used to determine instantaneous latent heat flux (LE) and, in turn, actual evapotranspiration, as described in Equations 4.1, 4.25, and 4.27. Two figures are included: Figure 5.14, consisting of five panels with each panel containing box plots corresponding to one land cover class across all year-season combinations; and Figure 5.15, containing six tables corresponding to the year-season combinations analyzed in the previous section.

Each table includes data on instantaneous soil heat flux (G), instantaneous sensible heat flux (H), instantaneous latent heat flux (LE), instantaneous net radiant energy (Rn_inst), net radiation in a day (Rn_24), and instantaneous evaporative fraction (EF_inst). The units for all energy fluxes are (W/m^2), and EF_inst is a dimensionless quantity. The significance of Rn_24 is observed in Equation 4.27, where it is multiplied with EF_inst and the advection factor (Ω) to determine daily actual evapotranspiration (ET_{24}) in mm/day. In this section, the region considered for analysis is the part of study area that is masked by 6 cloud and cloud shadow layers pertaining to all images chosen in this study, along with built-up area masks and digital elevation masks.

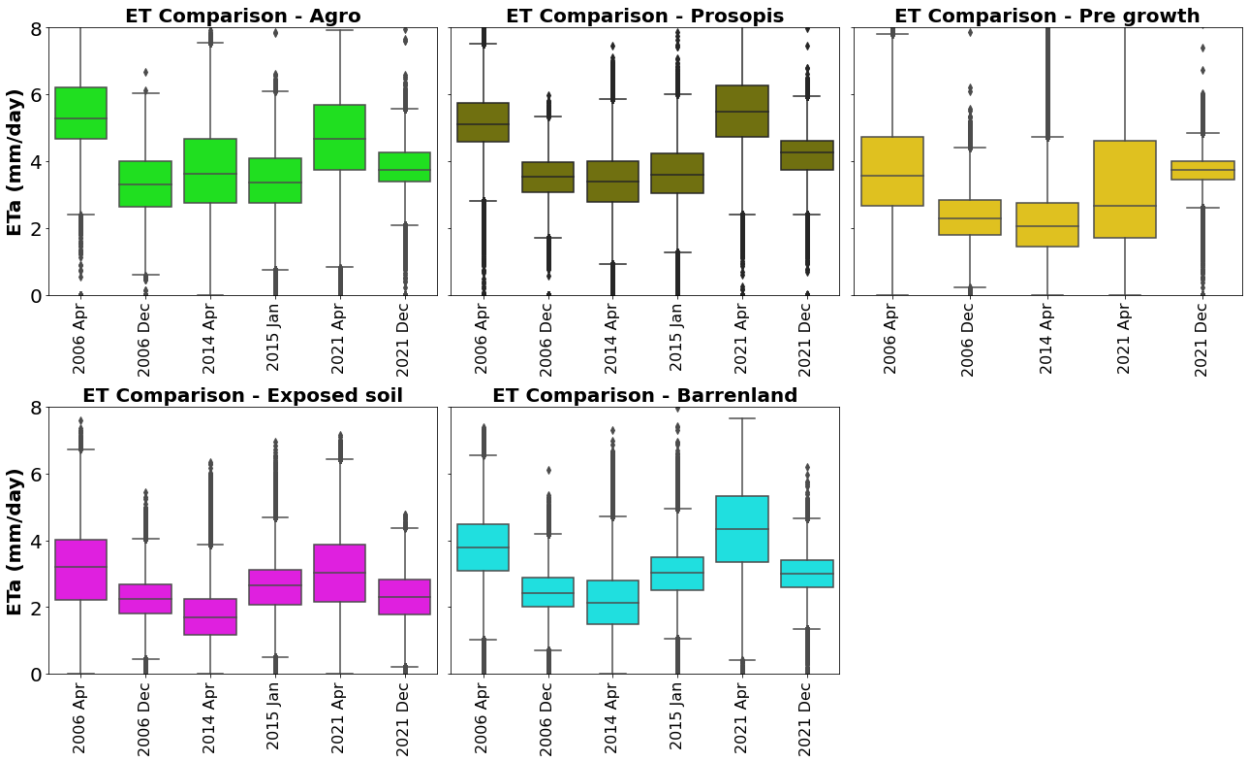

Figure 5.14 Seasonal variations of actual evapotranspiration across land cover classes

Table 5.8 Seasonal variations of energy fluxes in (W/m²) across land cover classes

2006 Summer

Class	G	H	LE	Rn_inst	Rn_24	EF_inst
Agro	83	93	278	458	180	0.74
Barren	81	148	179	432	169	0.51
Exposed soil	88	166	151	424	162	0.45
Pre growth	83	154	182	441	174	0.51
Prosopis	85	103	275	475	173	0.70

2006 Monsoon

Class	G	H	LE	Rn_inst	Rn_24	EF_inst
Agro	54	127	253	441	129	0.65
Barren	49	163	172	410	116	0.48
Exposed soil	53	162	159	390	115	0.47
Pre growth	51	178	169	421	121	0.46
Prosopis	46	95	256	417	119	0.69

2014 Summer

Class	G	H	LE	Rn_inst	Rn_24	EF_inst
Agro	101	155	193	451	156	0.55
Barren	95	195	103	418	144	0.32
Exposed soil	100	197	88	399	140	0.30
Pre growth	94	208	103	435	151	0.30
Prosopis	100	156	184	455	149	0.52

2014 Monsoon

Class	G	H	LE	Rn_inst	Rn_24	EF_inst
Agro	55	132	212	413	138	0.59
Barren	58	150	182	412	135	0.51
Exposed soil	62	147	161	384	126	0.50
Prosopis	51	107	239	420	129	0.65

2021 Summer

Class	G	H	LE	Rn_inst	Rn_24	EF_inst
Agro	83	118	261	473	165	0.67
Barren	78	119	230	463	153	0.60
Exposed soil	91	174	163	440	157	0.47
Pre growth	83	182	159	452	158	0.43
Prosopis	87	80	311	490	163	0.77

2021 Monsoon

Class	G	H	LE	Rn_inst	Rn_24	EF_inst
Agro	56	72	253	394	121	0.75
Barren	50	112	183	382	111	0.55
Exposed soil	63	149	142	371	115	0.46
Pre growth	52	76	233	385	117	0.70
Prosopis	60	74	298	436	132	0.79

From Figure 5.14, a decrease in mean ET_a values during the monsoon season compared to the summer is observed for the years 2006 and 2021 when considering vegetation classes (agro and prosopis). Mean values remain similar in 2014. In non-vegetation classes, a comparison of the pre-growth class cannot be made for all years as it is absent during the monsoon season of 2014. In other years, a decrease in the mean is observed in 2006 within the year, while a slight increase is noted in 2014. Regarding other non-vegetation classes (exposed soil and barren), a similar pattern of a decrease in mean values is observed in 2006 and 2021, but a reversal is seen in 2014.

These changes can be explained by examining the different energy fluxes. Between vegetation and non-vegetation classes, across all year-season combinations, the difference observed in soil heat flux (G) values within a season is close to 10W/m², indicating that there is not a remarkable difference between the classes concerning G. This is because soil heat flux is primarily influenced by surface temperature (T_s) more than albedo and NDVI, which are part of the equation used to determine G (Equation 4.16). In the monsoon season, a reduction in G is observed due to lower surface temperatures and increased soil moisture levels resulting from more precipitation compared to summer (Figure 5.15).

With respect to the radiation components Rn_inst and Rn_24, notable differences are observed between the vegetation and non-vegetation classes. Within a year-season, when studying the range (maximum value – minimum value) of Rn_inst, values greater than 50 W/m² are observed in all cases except the 2014 monsoon season. When Rn_24 is considered, the values are lower between the vegetation and non-vegetation classes since it is averaged throughout the day, with

a difference range between 12-21W/m² across the year seasons. This difference is attributed to the increased absorption of radiation fluxes by surfaces with vegetation compared to surfaces devoid of any vegetation, as observed in (Li & Zhao, 2010). A decline in both components is seen in the monsoon season compared to summer, which is typical during the monsoon season in this part of India.

The most complicated and challenging energy flux in SEBAL is the sensible heat flux (H), reliant on various parameters, corrections, and assumptions as discussed in the theoretical section 4.2.1 above. pySEBAL involves the automated selection of hot and cold pixels using percentiles of T_s and NDVI, which play a crucial role in the final estimate of H as described in section 4.2.2. Unlike soil heat flux (G), net radiation components Rn_inst and Rn_24, where a decrease is observed during the monsoon season, seasonal patterns or variations are not observed in sensible heat flux (H) (Figure 5.15). This is because H is dependent on the combined influence of the near-surface temperature difference between two vertical levels (dT) and aerodynamic resistance to heat transport (r_{ah}) (Equation 4.17), with both components involving an elaborate calculation procedure as shown in Figures 4.1-4.2 (Bezerra et al., 2015).

But within the different land cover classes, non-vegetation classes consistently showcase higher H values compared to vegetation classes. This tendency can be attributed to the relatively cooler plant-soil interface in agriculture, obtained because of watering. In the case of prosopis, its utilization of groundwater and atmospheric moisture could lead to a more humid plant-soil interface, thereby minimizing disparities between surface and air temperature (dT) (Shiferaw et

al., 2021) Consequently, this phenomenon leads to lower estimates of H for the vegetation classes compared to the non-vegetation classes (Figure 5.15).

Ultimately, the combination of all fluxes decides EF_inst, which along with Rn_24 and the advection factor, dictates the final ET estimate (Equations 4.26 - 4.27). Across all the year-season combinations, the vegetation classes have higher EF_inst values compared to the non-vegetation classes within a chosen year-season (Figure 5.15). However, the patterns of mean ET_a values do not necessarily reflect in EF_inst as Rn_24 is also needed to calculate the final ET_a estimate.

When comparing the summer season mean ET_a estimates across all land cover classes between 2006, 2014, and 2021, a sharp drop is observed in 2014 following a rise in 2021, starting from 2006. For the monsoon season, though, the estimates across all classes do not exhibit such a sharp variation as seen in the summer season. Mild increasing trends are observed for the agro and prosopis classes (Figure 5.14).

Through the findings, we can conclude that the seasonal changes in the climatology of the landscape during the monsoon, reflected by lower surface temperatures, reduced radiation, higher precipitation, and the difference in land cover types, influence the actual evapotranspiration estimates.

5.5 Estimating Water Outflux Variability Across Seasons and Years Across Different Land Cover Classes

This section presents the average net water outflux (in m^3/day) across different land cover classes throughout the landscape, as depicted in Figure 5.18. The water outflux estimate is derived by multiplying the area covered by each land cover class within a day by the corresponding mean ET_a value for that class on the same date. The normalization procedures applied to the landscape, including masking various layers as outlined in the previous section, remain consistent in this context. Figures 5.16 to 5.18 below illustrate variations in mean ET_a values, changes in the study area, and the resultant water outflux between seasons and across different years for all land cover classes obtained by combining area with ET_a estimates pertaining to one class. In the figures, the following labels are used: AG for agriculture class (green), BA for barren land (red), ES for exposed soil (gold), PG for pre-growth (purple), and PS for prosopis (blue).

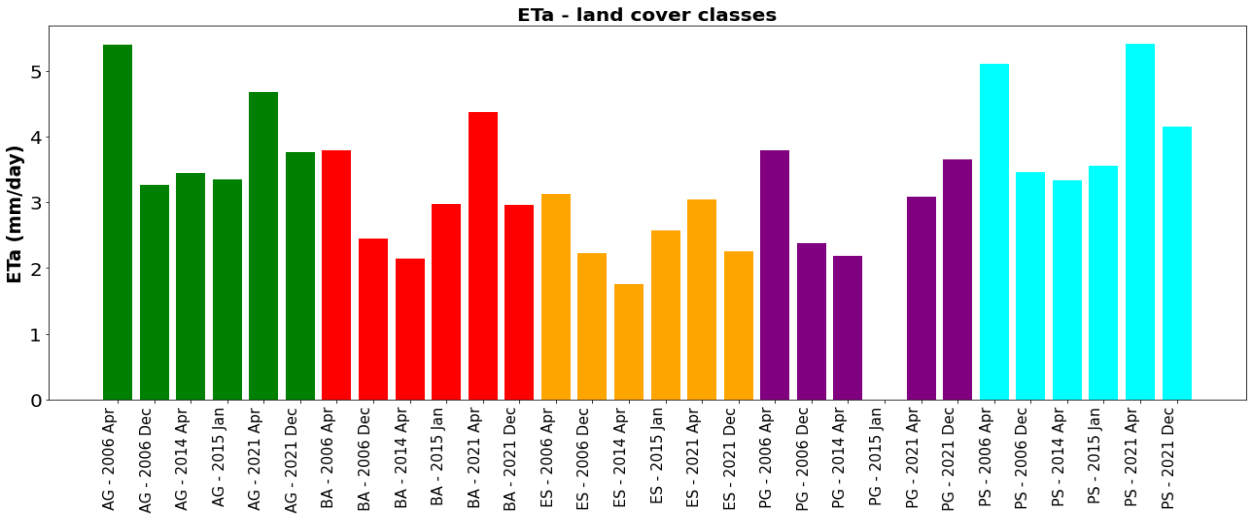

Figure 5.15 ET_a in (mm/day) across land cover classes over seasons and years

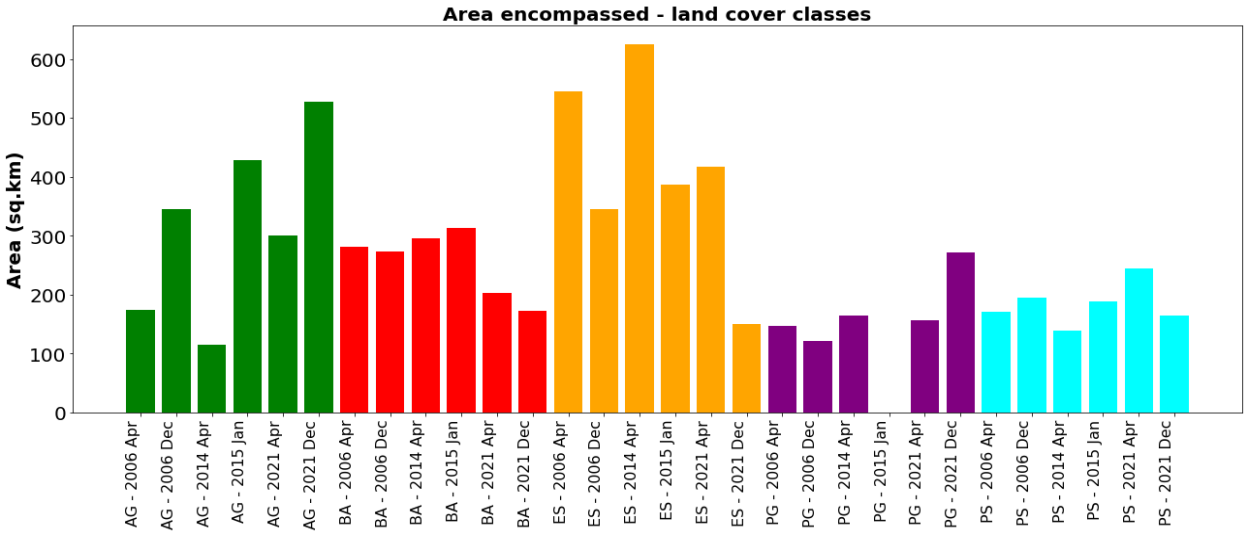

Figure 5.16 Area encompassed in (km^2) across land cover classes over seasons and years

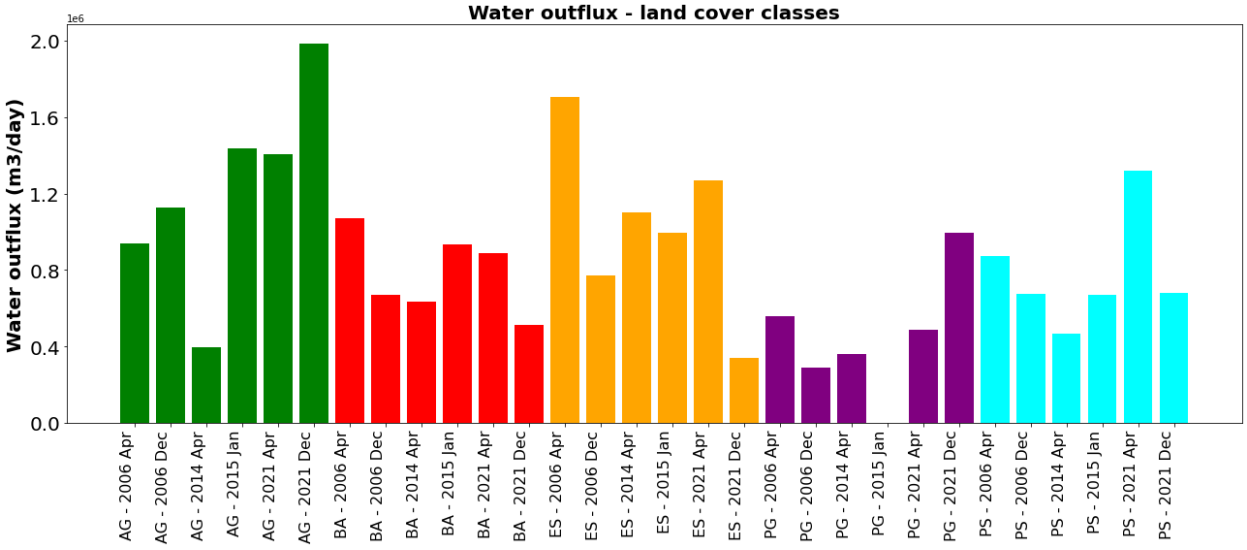

Figure 5.17 Average water outflux in (m^3/day) across land cover classes over seasons and years. Note: To obtain the actual values on the y-axis, multiply the values by 10^6

The most important pattern observed in the land cover analysis (section 5.1 and Figure 5.17), where there is a concurrent decrease in the exposed soil class along with an increase in the agriculture class during the monsoon, is reflected here even after aggregating the evapotranspiration component to obtain the water outflux, over 2006 and 2021 (Figures 5.16-5.18). However, the difference (decline) in the value in the exposed soil class is less apparent in 2014 due to the combined effects of ET_a means and the decline in area (Figure 5.18).

A clear decline in the outflux rates in monsoon is observed in 2006 and 2021, along with an increase in 2014, for both the barren land and prosopis classes. For the barren land class, it is inferred that the decline and the increase in corresponding years is related to the ET_a estimates to a large extent, as there is not a significant change in area observed between seasons within a year in the years considered for analysis. The decrease in monsoon for the pre growth class in 2006 relates to the corresponding decreases in ET_a and the area and the increase in 2021 can be attributed to increased mean ET_a and the increase in area during 2021 (Figures 5.16-5.18).

In the case of prosopis, despite an increase in area in monsoon, the decrease in ET_a has resulted in lesser outflux values during 2006 whereas a decline in area and ET_a in 2021 is associated with a decline in the net water outflux in 2021. Increase in both the values in 2014 has resulted in an increase in water outflux in 2014 monsoon (Figures 5.16-5.18).

In an overall sense, when summer seasons are exclusively compared, for all the land cover classes, starting at 2006, a consistent pattern of a decline 2014 followed by an increase in 2021 is seen (Figure 5.18).

When monsoon is considered, an increasing pattern is observed for agriculture between 2006 and 2014. For both barren land and exposed soil classes, starting at 2006, an increase in 2014 and a decrease in 2021 is seen. For prosopis, another major class of concern, a stable range is found between the monsoon seasons in 2006, 2014 and 2021 which is an important finding in our study (Figure 5.18).

Table 5.9 Total Water Outflux in m^3/day across the years and seasons

Season	Total Water Outflux in (m3/day)
2006 Summer	5142212
2006 Monsoon	3534906
2014 Summer	2954897
2014 Monsoon	4046322
2021 Summer	5369161
2021 Monsoon	4512596

Table 5.8 presents the total water outflux occurring in the landscape, obtained by adding uo the water outflux from all individual land cover classes per season and year (Figure 5.18). From the table, an increase in the monsoon season is noted in 2014, whereas a decline in the monsoon season is observed for the rest of the years.

It should be noted that these estimates represent a single day during the satellite overpass. They could be regarded as reasonable estimates for water budgeting purposes during mid-summer and the end of the northeast monsoon. As discussed earlier in the thesis, one image represents mid-summer and one the end of the northeast monsoon. However, to quantify the ET_a outflux

for a whole season, more images would be needed to account for variations in meteorology and landcover changes which control actual evapotranspiration.

6 Discussion

The Gundar River Basin is a unique landscape, home to a diverse set of land cover types. Evapotranspiration (ET) is an important hydrological variable and to the best of our knowledge, this study represents the first comprehensive effort to measure basin-wide ET and assess its variability across different land cover classes and seasons within a year in this landscape, providing insights into the seasonal behaviour of ET and the corresponding influence of different land cover classes on the final estimate. Given the interconnectedness of various land cover types in the water balance of the region, assessing ET over different land cover types becomes highly important. By combining results from the land cover classification analysis and ET estimations using SEBAL, a comprehensive understanding of water loss through ET within the landscape is obtained. In regions like these, where agriculture is prevalent throughout the year, ET estimates are crucial for managing water resources in an effective manner which involves a wide variety of stakeholders.

Despite seasonal variations and the presence of clouds and cloud shadow regions, noticeable patterns were observed across all land cover classes, particularly water classes, exposed soil, and agriculture. Consistent patterns were observed in the seasonal change detection, indicating the conversion of dry agricultural land (referred to as the exposed soil class in our analysis) from summer into wet agricultural land during the monsoon season for all the years. The increase in area and appearance of pixels indicating water classes during the monsoon season further indicates the effectiveness of the land cover maps in simulating real-time conditions.

It has to be noted that, despite the presence of numerous tanks, not all of them get filled during the monsoon season due to poor maintenance and also possibly due to the hydrological dynamics present in the region, influenced by activities like agriculture, groundwater extraction etc. Utilizing the ESA World Cover 10m, 2020 global land cover product to mask out built-up areas proved effective for seasonal land cover change detection since built-up areas often contain diverse components and can be challenging to differentiate from other classes. The use of a consolidated two-layer cloud mask for seasonal analysis, along with a six-layer cloud mask, has proven beneficial for a comprehensive analysis within seasons and across years, particularly for assessing the variability of ET_a.

GLDAS data are used currently to fulfill the instantaneous data requirement due to the lack of availability of instantaneous in-situ data during satellite overpasses. Exploring and comparing the use of other gridded products could provide insights into the difference in performances when used as a forcing parameter in pySEBAL.

Correlation coefficients and RMSE values across all the year-season combinations in G2 show lower values during the monsoon season compared to the summer season. Additionally, our findings regarding the range of ET_a mean values with the different land cover classes align with those reported in previous studies (Ayad Ali Faris Beg et al., 2016; Jana et al., 2016; Karishma et al., 2022; Li & Zhao, 2010; Ning et al., 2017) exhibiting pySEBAL's ability to demarcate classes where the non-vegetation classes exhibit lower ET_a mean when compared to the vegetation class.

The differences between the extreme mean values (between the exposed soil and either of the vegetation classes) range from 0.97 mm/day during the 2014 monsoon to 2.4 mm/day in the 2006 summer, with four year-season combinations crossing 2 mm/day overall. Regarding heat flux components, T_s dictates soil heat flux values, with reductions during the monsoon season due to lower temperatures. Sensible heat flux (H), influenced by r_{ah} and dT, introduces uncertainties in ET_a estimation, particularly notable as it involves multiple parameterizations and corrections. Vegetation classes generally exhibit lower H values compared to non-vegetation classes. A similar pattern is also observed for the components - instantaneous net radiant energy (Rn_inst) and net radiation in a day (Rn_24).

Regarding the estimation of water outflux for the entire landscape, a decrease was observed during the monsoon seasons of 2006 and 2021, while an increase was noted in 2014. One plausible explanation could be a drought year, as indicated by the reduced area encompassed by water classes compared to previous years. Comparing the water outflux values during the summer seasons across the years, a consistent pattern emerged: a decrease in 2014 followed by an increase in 2021, starting from 2006, across all land cover classes. Interestingly, when comparing the monsoon seasons between the years, the most important finding was that the water outflux remained relatively stable for the prosopis class, with minimal variability observed between the years.

6.1 Limitations

Limitations include the 16-day Landsat image cycle, which may result in data loss due to cloud cover. Additionally, the use of image composites introduces uncertainties in accuracy, as composites may not fully represent the actual spectral behavior of the study area. However, evaluating the sensitivity of different composites presents an opportunity to assess their usefulness. As a workaround, consolidated cloud masks were used in this study, although this approach may result in loss of area coverage.

Lack of instantaneous in-situ data is another constraint. Assumptions in the SEBAL model, such as constant EF_inst throughout the day and a linear relationship between T_s and dT for estimating the sensible heat flux (H), need to be acknowledged. In terms of processing, pySEBAL relies on a single Landsat scene. If the study area spans multiple scenes, each scene must be processed separately, and then the values for the study area are combined. This approach may present challenges, particularly when estimating H, which requires the selection of anchor pixels.

6.2 Future Work

Future work could involve obtaining seasonal and yearly ET estimates using multiple images, utilizing datasets with fused synthetic images to avoid cloud interference like planet fusion data (Planet Fusion Team, 2021), mapping prosopis more thoroughly using higher resolution imagery, and localized calibration of parameters for determining sensible heat flux.

As mentioned in section 2.2.4, usage of other SEB models like SSEBI, SEBSM and SSEB can also be explored. Additionally, validation through field lysimeters and flux towers would enhance the study's value, especially when examining seasonal changes and behavior.

7 Conclusion

Section 5.1 discusses the land cover analysis, which achieved high overall accuracy levels exceeding 90% using the random forest algorithm across all six images. This emphasizes the capability of the random forest algorithm in differentiating land cover classes, as discussed in the literature review. Section 5.2 compares in-situ data with GLDAS data ensuring the compatibility of the GLDAS dataset, and depicting the usefulness of the GLDAS data in the analysis. Section 5.3 presents a comprehensive analysis that involves estimation of ET_a using two scenarios using pySEBAL and the comparison with the EEFlux data which is an independent dataset using correlation coefficient and RMSE. Section 5.4 elucidates the role of different energy fluxes and their influence on final ET estimates, while Section 5.5 highlights the combined role of areas encompassed by different land cover classes and mean ET_a values between seasons.

The seasonal land cover analysis highlighted the predominant patterns present in the region, including an increase in agriculture and the filling up of water bodies during the monsoon season. The proximity observed between SEBAL and EEFlux based outputs, especially during the monsoon season, demonstrates the successful implementation of the SEBAL model in our study area and its comparability with equivalent products. The next step would involve validation with in-situ measurements. A more comprehensive understanding of energy fluxes can be obtained through a time-series analysis involving multiple images, for which this study is a valuable reference. Estimating water outflux could help in zonal planning, considering that different zones

have varying resources and budgets. For instance, a district could be considered as a zone in the Indian context.

This research presents a framework for analyzing evapotranspiration in large, heterogeneous landscapes like river basins, encompassing various land cover classes undergoing seasonal changes. In conclusion, this study provides a multi-step scientific framework for estimating ET in data-scarce, heterogeneous, and semi-arid regions where water management and planning efforts are crucial. In the study area, the importance of tank maintenance, strategic handling of prosopis due to its implications on the local economy, and zonal planning becomes crucial to support the variety of activities, especially year-round agriculture.

References

1. Abbott, B. W., Bishop, K., Zarnetske, J. P., Minaudo, C., Chapin, F. S., Krause, S., ... & Pinay, G. (2019). Human domination of the global water cycle absent from depictions and perceptions. Nature Geoscience, 12(7), 533-540.
2. Abrams, M., Yamaguchi, Y., & Crippen, R. (2022). Aster Global Dem (GDEM) Version 3. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 43, 593-598.
3. Ahmad, M.-D., Biggs, T., Turral, H., & Scott, C. A. (2006). Application of SEBAL approach and MODIS time-series to map vegetation water use patterns in the data scarce Krishna River basin of India. Water Science and Technology, 53(10), 83–90. https://doi.org/10.2166/wst.2006.301
4. Allan, R. P., Barlow, M., Byrne, M. P., Cherchi, A., Douville, H., Fowler, H. J., ... & Zolina, O. (2020). Advances in understanding large-scale responses of the water cycle to climate change. Annals of the New York Academy of Sciences, 1472(1), 49-75
5. Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. Fao, Rome, 300(9), D05109.
6. Allen, R., Irmak, A., Trezza, R., Hendrickx, J. M., Bastiaanssen, W., & Kjaersgaard, J. (2011). Satellite-based ET estimation in agriculture using SEBAL and METRIC. Hydrological processes, 25(26), 4011-4027.

7. Allen, R., Morton, C., Kamble, B., Kilic, A., Huntington, J., Thau, D., Erickson, T., Moore, R., Trezza, R., Ratcliffe, I., & Robison, C. (2015). EEFlux: A Landsat-based Evapotranspiration mapping tool on the Google Earth Engine.
8. Amini, S., Saber, M., Rabiei-Dastjerdi, H., & Homayouni, S. (2022). Urban Land Use and Land Cover Change Analysis Using Random Forest Classification of Landsat Time Series. Remote Sensing, 14(11), 2654. https://doi.org/10.3390/rs14112654
9. Ayad Ali Faris Beg, Ahmed H. Al-Sulttani, Adrian Ochtyra, Anna Jarocińska, & Adriana Marcinkowska. (2016). Estimation of Evapotranspiration Using SEBAL Algorithm and Landsat-8 Data—A Case Study: Tatra Mountains Region. Journal of Geological Resource and Engineering, 4(6). https://doi.org/10.17265/2328-2193/2016.06.002
10. Bala, A., Rawat, K. S., Misra, A. K., & Srivastava, A. (2016). Assessment and validation of evapotranspiration using SEBAL algorithm and Lysimeter data of IARI agricultural farm, India. Geocarto International, 31(7), 739–764. https://doi.org/10.1080/10106049.2015.1076062
11. Bandyopadhyay, A., Bhadra, A., Raghuwanshi, N. S., & Singh, R. (2009). Temporal trends in estimates of reference evapotranspiration over India. Journal of Hydrologic Engineering, 14(5), 508-515.
12. Bastiaanssen, W. G. M., Menenti, M., Feddes, R. A., & Holtslag, A. A. M. (1998). A remote sensing surface energy balance algorithm for land (SEBAL). 1. Formulation. Journal of Hydrology, 212–213, 198–212. https://doi.org/10.1016/S0022-1694(98)00253-4

13. Beaudoing, H. and M. Rodell, NASA/GSFC/HSL (2020), GLDAS Noah Land Surface Model L4 3 hourly 0.25 x 0.25 degree V2.1, Greenbelt, Maryland, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: [Data Access Date], 10.5067/E7TYRXPJKWOQ
14. Bezerra, B. G., da Silva, B. B., dos Santos, C. A., & Bezerra, J. R. (2015). Actual evapotranspiration estimation using remote sensing: comparison of SEBAL and SSEB approaches. Advances in Remote Sensing, 4(03), 234.
15. Bhattarai, N., Shaw, S. B., Quackenbush, L. J., Im, J., & Niraula, R. (2016). Evaluating five remote sensing based single-source surface energy balance models for estimating daily evapotranspiration in a humid subtropical climate. International Journal of Applied Earth Observation and Geoinformation, 49, 75–86. https://doi.org/10.1016/j.jag.2016.01.010
16. Bhimala, K. R., Patra, G. K., & Goroshi, S. (2023). Annual and seasonal trends in actual evapotranspiration over different meteorological sub-divisions in India using satellite-based data. Theoretical and Applied Climatology, 152(3-4), 999-1017.
17. Chanu, C. S., Munagapati, H., Tiwari, V. M., Kumar, A., & Elango, L. (2020). Use of GRACE time-series data for estimating groundwater storage at small scale. Journal of Earth System Science, 129(1), 215. https://doi.org/10.1007/s12040-020-01465-2
18. Chattopadhyay, N., & Hulme, M. (1997). Evaporation and potential evapotranspiration in India under conditions of recent and future climate change. Agricultural and Forest Meteorology, 87(1), 55–73. https://doi.org/10.1016/S0168-1923(97)00006-3

19. Chemura, A., Rwasoka, D., Mutanga, O., Dube, T., & Mushore, T. (2020). The impact of land-use/land cover changes on water balance of the heterogeneous Buzi sub-catchment, Zimbabwe. Remote Sensing Applications: Society and Environment, 18, 100292. https://doi.org/10.1016/j.rsase.2020.100292
20. Chinnasamy, P., & Agoramoorthy, G. (2015). Groundwater Storage and Depletion Trends in Tamil Nadu State, India. Water Resources Management, 29(7), 2139–2152. https://doi.org/10.1007/s11269-015-0932-z
21. Crawford, C.J., Roy, D.P., Arab, S., Barnes, C., Vermote, E., Hulley, G., Gerace, A., Choate, M., Engebretson, C., et al. 2023. "The 50-Year Landsat Collection 2 Archive." Science of Remote Sensing (2023) https://doi.org/10.1016/j.srs.2023.100103.
22. Darshana, Pandey, A., & Pandey, R. P. (2013). Analysing trends in reference evapotranspiration and weather variables in the Tons River Basin in Central India. Stochastic Environmental Research and Risk Assessment, 27(6), 1407–1421. https://doi.org/10.1007/s00477-012-0677-7
23. Dhawan, V. (2017). Water and agriculture in India. In Background paper for the South Asia expert panel during the Global Forum for Food and Agriculture (Vol. 28, pp. 80-85) Fr. 5.00. WMO No. 156,Technical Note No. 63. .
24. Gadgil, S., & Gadgil, S. (2006). The Indian monsoon, GDP and agriculture. Economic and political weekly, 4887-4895.

25. Gao, F., Feng, G., Ouyang, Y., Wang, H., Fisher, D., Adeli, A., & Jenkins, J. (2017). Evaluation of reference evapotranspiration methods in arid, semiarid, and humid regions. JAWRA Journal of the American Water Resources Association, 53(4), 791-808..
26. Glenn, E. P., Huete, A. R., Nagler, P. L., Hirschboeck, K. K., & Brown, P. (2007). Integrating Remote Sensing and Ground Methods to Estimate Evapotranspiration. Critical Reviews in Plant Sciences, 26(3), 139–168. https://doi.org/10.1080/07352680701402503
27. Goroshi, S., Pradhan, R., Singh, R. P., Singh, K. K., & Parihar, J. S. (2017). Trend analysis of evapotranspiration over India: Observed from long-term satellite measurements. Journal of Earth System Science, 126(8), 113. https://doi.org/10.1007/s12040-017-08912
28. Goyal, R. K. (2004). Sensitivity of evapotranspiration to global warming: a case study of arid zone of Rajasthan (India). Agricultural water management, 69(1), 1-11.
29. Gujja, B., Alagh, Y. K., Pangare, G., & Gujja, B. (2006). Interlinking of Rivers in India. New Delhi: Academic Foundation, in collaboration with the National Civil Society Committee on Interlinking of Rivers in India (NCSCILR).
30. Hari, M., Tyagi, B., Huddar, M. S. K., & Harish, A. (2021). Satellite-based regional-scale evapotranspiration estimation mapping of the rice bowl of Tamil Nadu: A little water to spare *. Irrigation and Drainage, 70(4), 958–975. https://doi.org/10.1002/ird.2553
31. Huntington, T. G. (2006). Evidence for intensification of the global water cycle: Review and synthesis. Journal of Hydrology, 319(1–4), 83–95. https://doi.org/10.1016/j.jhydrol.2005.07.003

32. Jana, C., Rawat, M., Sena, D. R., Alam, N. M., Mandal, U., Kaushal, R., & Mishra, P. K. (2016). Application of SEBAL model to estimate Evapotranspiration in Doon Valley, India.
33. Janani, N., Kannan, B., Nagarajan, K., Thiyagarajan, G., & Duraisamy, M. R. (2023). Trend analysis and variability of satellite-based soil moisture data for the Lower Bhavani basin, Tamil Nadu using Google Earth Engine. Journal of Applied and Natural Science, 15(2), 555-559.
34. Jensen, M. E. (1967). Empirical methods of estimating or predicting evapotranspiration using radiation.
35. Jensen, M. E. (1968). Water consumption by agricultural plants (Chapter 1).
36. John, J., Chithra, N. R., & Thampi, S. G. (2019). Prediction of land use/cover change in the Bharathapuzha river basin, India using geospatial techniques. Environmental Monitoring and Assessment, 191(6), 354. https://doi.org/10.1007/s10661-019-7482-4
37. Kamali, M. I., & Nazari, R. (2018). Determination of maize water requirement using remote sensing data and SEBAL algorithm. Agricultural Water Management, 209, 197–205. https://doi.org/10.1016/j.agwat.2018.07.035
38. Kamaraj, M., & Rangarajan, S. (2022). Predicting the future land use and land cover changes for Bhavani basin, Tamil Nadu, India, using QGIS MOLUSCE plugin. Environmental Science and Pollution Research, 29(57), 86337–86348. https://doi.org/10.1007/s11356-021-17904-6
39. Kamyab, A. D., Mokhtari, S., & Jafarinia, R. (2022). A comparative study in quantification of maize evapotranspiration for Iranian maize farm using SEBAL and METRIC-1 EEFLux

algorithms. Acta Geophysica, 70(1), 319–332. https://doi.org/10.1007/s11600-021-00704-4

40. Karimi, P., Pareeth, S., & Michailovsky, C. (2019). Rapid assessment of the water accounts in Urmia Lake basin. Project Report.
41. Karishma, C. G., Kannan, B., Nagarajan, K., Panneerselvam, S., & Pazhanivelan, S. (2022). Spatial and temporal estimation of actual evapotranspiration of lower Bhavani basin, Tamil Nadu using Surface Energy Balance Algorithm for Land Model. Journal of Applied and Natural Science, 14(2), 566-574.
42. Kottek, M., Grieser, J., Beck, C., Rudolf, B., & Rubel, F. (2006). World map of the Köppen-Geiger climate classification updated.
43. KP, R., Kannan, B., PJ, P., & GR, M. (2021). Agricultural drought monitoring in Tamil Nadu in India using Satellite-based multi vegetation indices. Journal of Applied and Natural Science, 13(2), 414-423.
44. Krishna, P. R. A. (2019). Evapotranspiration and agriculture-A review. AGRICULTURAL REVIEWS, 40(1).
45. Kulkarni, A. D., & Lowe, B. (n.d.). Random Forest Algorithm for Land Cover Classification. International Journal on Recent and Innovation Trends in Computing and Communication, 4(3).
46. Kundu, S., Mondal, A., Khare, D., Hain, C., & Lakshmi, V. (2018). Projecting Climate and Land Use Change Impacts on Actual Evapotranspiration for the Narmada River Basin in

Central India in the Future. Remote Sensing, 10(4), 578. https://doi.org/10.3390/rs10040578

47. Lal, M. (2001). Climatic Change—Implications for India's Water Resources. Journal of Social and Economic Development
48. Li, F., & Lyons, T. J. (1999). Estimation of regional evapotranspiration through remote sensing. Journal of Applied Meteorology and Climatology, 38(11), 1644-1654.
49. Li, G., Zhang, F., Jing, Y., Liu, Y., & Sun, G. (2017). Response of evapotranspiration to changes in land use and land cover and climate in China during 2001–2013. Science of the Total Environment, 596, 256-265.
50. Li, S., & Zhao, W. (2010). Satellite-based actual evapotranspiration estimation in the middle reach of the Heihe River Basin using the SEBAL method. Hydrological Processes, 24(23), 3337–3344. https://doi.org/10.1002/hyp.7748
51. Loukika, K. N., Keesara, V. R., & Sridhar, V. (2021). Analysis of land use and land cover using machine learning algorithms on google earth engine for Munneru River Basin, India. Sustainability, 13(24), 13758.
52. Madhu, S., Kumar, T. V. L., Barbosa, H., Rao, K. K., & Bhaskar, V. V. (2015). Trend analysis of evapotranspiration and its response to droughts over India. Theoretical and Applied Climatology, 121(1–2), 41–51. https://doi.org/10.1007/s00704-014-1210-3
53. Maharana, P., & Dimri, A. P. (2019). The Indian Monsoon: past, present and future. Proceedings of the Indian National Science Academy, 85(2), 403-420

54. Misra, & Bhardwaj, A. (2019). Defining the Northeast Monsoon of India. Monthly Weather Review, 147(3), 791–807. https://doi.org/10.1175/MWR-D-18-0287.1

55. Mohan, S. (1991). Intercomparison of evapotranspiration estimates. Hydrological Sciences Journal, 36(5), 447–460. https://doi.org/10.1080/02626669109492530

56. Munawir, A., June, T., Kusmana, C., & Setiawan, Y. (2022). SEBAL Model to Estimate Biophysics and Energy Flux Variable: Availability of Evapotranspiration Distribution Using Remote Sensing in Lore Lindu National Park. IOP Conference Series: Earth and Environmental Science, 950(1), 012022. https://doi.org/10.1088/1755-1315/950/1/012022

57. National Water Mission. (2017, July 21). Gundar Basin. National Water Mission. http://nwm.gov.in/sites/default/files/Gundar_Basin-21-07-2017.pdf

58. Nguyen, H. T. T., Doan, T. M., & Radeloff, V. (2018). APPLYING RANDOM FOREST CLASSIFICATION TO MAP LAND USE/LAND COVER USING LANDSAT 8 OLI. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-3/W4, 363–367. https://doi.org/10.5194/isprs-archives-XLII-3-W4-363-2018

59. Ning, J., Gao, Z., & Xu, F. (2017). Effects of land cover change on evapotranspiration in the Yellow River Delta analyzed with the SEBAL model. Journal of Applied Remote Sensing, 11(1), 016009. https://doi.org/10.1117/1.JRS.11.016009

60. Pal, L., Kumar, A., Ojha, C. S. P., & Chandniha, S. K. (2018). Sebal Based Evapotranspiration Estimation for Upper Tapi Basin (India). IGARSS 2018 - 2018 IEEE

International Geoscience and Remote Sensing Symposium, 7380–7382. https://doi.org/10.1109/IGARSS.2018.8518982

61. Palanisami, K., Ranganathan, C. R., Vidhyavathi, A., Rajkumar, M., & Ajjan, N. (2011). Performance of agriculture in river basins of Tamil Nadu in the last three decades–A total factor productivity approach. A Project Sponsored by Planning Commission, Government of India. Centre for Agricultural and Rural Development Studies, Tamil Nadu Agricultural University, 1-171.
62. Papadavid, G., Neocleous, D., Kountios, G., Markou, M., Michailidis, A., Ragkos, A., & Hadjimitsis, D. (2017). Using SEBAL to Investigate How Variations in Climate Impact on Crop Evapotranspiration. Journal of Imaging, 3(3), 30. https://doi.org/10.3390/jimaging3030030
63. Pareeth, S., & Karimi, P. (2023). Evapotranspiration estimation using Surface Energy Balance Model and medium resolution satellite data: An operational approach for continuous monitoring. Scientific Reports, 13(1), 12026. https://doi.org/10.1038/s41598-023-38563-2
64. Pathak, H., Pramanik, P., Khanna, M., & Kumar, A. (2014). Climate change and water availability in Indian agriculture: Impacts and adaptation. The Indian Journal of Agricultural Sciences, 84(6). https://doi.org/10.56093/ijas.v84i6.41421
65. Phan, T. N., Kuch, V., & Lehnert, L. W. (2020). Land Cover Classification using Google Earth Engine and Random Forest Classifier—The Role of Image Composition. Remote Sensing, 12(15), 2411. https://doi.org/10.3390/rs12152411

66. Phiri, D., & Morgenroth, J. (2017). Developments in Landsat Land Cover Classification Methods: A Review. Remote Sensing, 9(9), 967. https://doi.org/10.3390/rs9090967

67. Rahimzadegan, M., & Janani, A. (2019). Estimating evapotranspiration of pistachio crop based on SEBAL algorithm using Landsat 8 satellite imagery. Agricultural Water Management, 217, 383–390. https://doi.org/10.1016/j.agwat.2019.03.018

68. Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D., & Toll, D. (2004). The Global Land Data Assimilation System. Bulletin of the American Meteorological Society, 85(3), 381–394. https://doi.org/10.1175/BAMS-85-3-381

69. Saboori, M., Mousivand, Y., Cristóbal, J., Shah-Hosseini, R., & Mokhtari, A. (2022). An Automated and Improved Methodology to Retrieve Long-time Series of Evapotranspiration Based on Remote Sensing and Reanalysis Data. Remote Sensing, 14(24), 6253. https://doi.org/10.3390/rs14246253

70. Samui, R. P., Balasubramanian, R., & Kamble, M. V. (2013). Northeast monsoon rainfall and agricultural production in Tamil Nadu and Andhra Pradesh: II - Dry and wet spell and its impact on cropping pattern. MAUSAM, 64(3), 489–500. https://doi.org/10.54302/mausam.v64i3.731

71. Sato, T. (2013). Beyond water-intensive agriculture: Expansion of Prosopis juliflora and its growing economic use in Tamil Nadu, India. Land use policy, 35, 283-292.

72. Sato, T., & Duraiyappan, P. R. (2011). The effects of expansion of private wells on rural livelihood in tank intensive watersheds: A case study in upper Gundar River Basin, Tamil Nadu. Japanese Journal of Southeast Asian Studies, 49(1), 124-150.

73. Schaller, M. F., & Fan, Y. (2009). River basins as groundwater exporters and importers: Implications for water cycle and climate modeling. Journal of Geophysical Research, 114(D4), D04103. https://doi.org/10.1029/2008JD010636

74. Shiferaw, H., Alamirew, T., Dzikiti, S., Bewket, W., Zeleke, G., & Schaffner, U. (2021). Water use of Prosopis juliflora and its impacts on catchment water budget and rural livelihoods in Afar Region, Ethiopia. Scientific reports, 11(1), 2688.

75. Singh, G., & Singh, S. K. (2023). Evapotranspiration Over the Indian Region: Implications of Climate Change and Land Use/Land Cover Change. Nature Environment and Pollution Technology, 22(1), 211-219.

76. Soni, A., & Syed, T. H. (2021). Analysis of variations and controls of evapotranspiration over major Indian River Basins (1982–2014). Science of The Total Environment, 754, 141892. https://doi.org/10.1016/j.scitotenv.2020.141892

77. Srivastava, A., & Chinnasamy, P. (2021). Water management using traditional tank cascade systems: A case study of semi-arid region of Southern India. *SN Applied Sciences*, *3*(3), 281. https://doi.org/10.1007/s42452-021-04232-0

78. Talukdar, S., Singha, P., Mahato, S., Shahfahad, Pal, S., Liou, Y.-A., & Rahman, A. (2020). Land-Use Land-Cover Classification by Machine Learning Classifiers for Satellite

Observations—A Review. Remote Sensing, 12(7), 1135. https://doi.org/10.3390/rs12071135

79. Tan, L., Zheng, K., Zhao, Q., & Wu, Y. (2021). Evapotranspiration Estimation Using Remote Sensing Technology Based on a SEBAL Model in the Upper Reaches of the Huaihe River Basin. Atmosphere, 12(12), 1599. https://doi.org/10.3390/atmos12121599
80. Team, P. F. (2021). Planet Fusion Monitoring Technical Specification, Version 1.0. 0-beta. 3, San Francisco, CA.
81. Thangavel, M. (2020). Kuppannan Palanisami. Sustainable Development in India: Groundwater Irrigation, Energy Use, and Food Production.
82. Thornthwaite, C. W. (1948). An approach toward a rational classification of climate. Geographical review, 38(1), 55-94.
83. Tian, H., Banger, K., Bo, T., & Dadhwal, V. K. (2014). History of land use in India during 1880–2010: Large-scale land transformations reconstructed from satellite data and historical archives. Global and Planetary Change, 121, 78–88. https://doi.org/10.1016/j.gloplacha.2014.07.005
84. Twisa, S., & Buchroithner, M. F. (2019). Land-Use and Land-Cover (LULC) Change Detection in Wami River Basin, Tanzania. Land, 8(9), 136. https://doi.org/10.3390/land8090136
85. V. R. Vanthof and R. E. J. Kelly, "Mapping prosopis juliflora invasion within rainwater harvesting structures in India using Google Earth Engine," 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 2017, pp. 1115-1118, doi: 10.1109/IGARSS.2017.8127152.

86. Varadan, R. J., Kumar, P., Jha, G. K., Pal, S., & Singh, R. (2017). An exploratory study on occurrence and impact of climate change on agriculture in Tamil Nadu, India. Theoretical and Applied Climatology, 127(3–4), 993–1010. https://doi.org/10.1007/s00704-015-1682-9

87. Vargas Godoy, Markonis, Y., Hanel, M., Kyselý, J., & Papalexiou, S. M. (2021). The Global Water Cycle Budget: A Chronological Review. Surveys in Geophysics, 42(5), 1075–1107. https://doi.org/10.1007/s10712-021-09652-6

88. Waters, R., Allen, R., Bastiaanssen, W., Tasumi, M., Trezza, R., 2002. Surface Energy Balance Algorithms for land, Idaho Implementation, Advanced Training and User's Manual. NASA, USA.

89. Williams, M. D., Hawley, C., Madden, M., & Shepherd, J. M. (2017). Mapping the spatio-temporal evolution of irrigation in the Coastal Plain of Georgia, USA. Photogrammetric Engineering & Remote Sensing, 83(1), 57-67.

90. WMO, 1963. Sites for wind-power installations. Geneva, 1964, 28 illus., 4 tables, Swiss

91. Wukelic, G. E., Gibbons, D. E., Martucci, L. M., & Foote, H. P. (1989). Radiometric calibration of Landsat Thematic Mapper thermal band. Remote Sensing of Environment, 28, 339–347. https://doi.org/10.1016/0034-4257(89)90125-9

92. Zamani Losgedaragh, S., & Rahimzadegan, M. (2018). Evaluation of SEBS, SEBAL, and METRIC models in estimation of the evaporation from the freshwater lakes (Case study: Amirkabir dam, Iran). Journal of Hydrology, 561, 523–531. https://doi.org/10.1016/j.jhydrol.2018.04.025

93. Zanaga, D., Van De Kerchove, R., De Keersmaecker, W., Souverijns, N., Brockmann, C., Quast, R., Wevers, J., Grosu, A., Paccini, A., Vergnaud, S., Cartus, O., Santoro, M., Fritz, S., Georgieva, I., Lesiv, M., Carter, S., Herold, M., Li, Linlin, Tsendbazar, N.E., Ramoino, F., Arino, O., 2021. ESA WorldCover 10 m 2020 v100. (doi:10.5281/zenodo.5571936)

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

95. Zhang, Y.-K., & Schilling, K. E. (2006). Effects of land cover on water table, soil moisture, evapotranspiration, and groundwater recharge: A Field observation and analysis. Journal of Hydrology, 319(1–4), 328–338. https://doi.org/10.1016/j.jhydrol.2005.06.044