Politecnico di Torino

Master of Science in Environmental and Land Engineering

Track: Climate Change

A.Y. 2024/2025

Graduation Session October 2025

Master's Degree Thesis

Evapotranspiration shed of key agricultural crops: combining agro-hydrological estimates with atmospheric moisture dynamics

Supervisors:

Prof. Marta Tuninetti
Dr. Lan Wang-Erlandsson
Dr. Elena De Petrillo

Candidate:

Giulia Cigna

Abstract

The increasing global demand for food, feed and flexible crops is exerting unprecedented pressure on the global hydrological cycle through landscape conversion and increasing irrigation demand, which altogether contribute to the alteration of land-atmosphere feedbacks. These feedbacks influence evaporation and precipitation patters through atmospheric flows. Atmospheric moisture flows connect sources of evaporation to sinks of precipitation, from local to regional and continental scale, up to thousands of kilometres away. Terrestrial sources of evaporation are crucial for global food production, regulating precipitation and climate patterns by redistributing water and latent heat. At the same time, the alteration of evapotranspiration dynamics from these sources is mainly driven by land-use conversion for pasture (cattle meat production), and feed crops (such as soy, and maize) and agricultural practises, such as irrigation.
Current water use assessment disregard these feedbacks and the role played by atmospheric moisture connection in redistributing evaporation from agricultural parcels to precipitation in downwind areas. This understanding is particularly key to better assess the water-related implication of pivotal crops such as soy, maize and wheat which account for 33% of global harvested land and the 30% of global water footprint of crop production. Addressing this gap, this thesis aims to advance the understanding of how evapotranspiration from agricultural areas contributes to precipitation whether or not to other agricultural area. It emblematically presents the cases of soy, maize and wheat.

The first part of this thesis updates actual evapotranspiration estimates for soy, maize and wheat production for the period 2008–2017 by means of the agro-hydrological model waterCROP, which solves the daily soil water balance on a global 5 arc-minute grid, with global coverage for both irrigated and rainfed conditions. In the present work, the model is updated to a newer version, made consistent with daily climatic data from ERA5 reanalysis.

In the second part, the evapotranspiration estimates are combined with atmospheric connections by means of the RECON dataset, a 4D matrix of annual moisture flow connections between any cell in the world at the spatial resolution of 0.5°. In the present work, each cultivated cell of soy, maize and wheat is linked to its blue and green evapotranspiration shed (i.e. the downwind area receiving precipitation from irrigated or rainfed crop production). Evaporation sheds are finally classified according to their land use category to analyse potential synergies and trade-off between land and water use between the sites at the origin of evaporation and at the fate of precipitation.

By characterizing these connections, the thesis sheds light on the hidden global links between cultivated land and downwind areas. Ultimately, this thesis contributes toward a more comprehensive evaluation of the interplay between water and land use at the site of production with atmospheric feedbacks with local and distant link in the global water cycle.

To all the girls who dream of studying.
To the social battles that made it possible for me to do so.

Acknowledgements

I would like to take a few lines to thank my supervisors and all the researchers I met along this journey.

A very special thank you goes to Marta Tuninetti, who from the very beginning entrusted me with this research project, listening to my interests and inclinations. I am deeply grateful for your trust, and even happier that we will continue working together in the future. Thank you for your time and support, and for patiently reading my long emails and endless lists of questions. I feel truly inspired by the many projects you are involved in and by your deep knowledge of your field.

Another special thank you goes to Elena De Petrillo, for both your technical support and your warm encouragement. Thank you for your time and suggestions. I am very grateful for all your help, and I hope I can return it one day. Your energy is both contagious and precious. I am also glad that our collaboration will continue.

I would also like to thank the entire Stockholm Resilience Centre community, who welcomed me so warmly during my three months there. In particular, I wish to thank Lan Wang-Erlandsson, who kindly agreed to supervise me without even knowing me beforehand. I truly appreciate the time you dedicated to me and your enthusiasm for my research. I greatly admire your work, the way you collaborate with others, and your inspiring research attitude. A huge thank you as well to everyone who spent part of their time giving me precious advice. I am especially grateful to the Water Resilience team, and in particular to Lucie Bakels and Agnes Pranindita.

Thank you all! I am very excited about the work I have done so far, and I cannot wait to continue this journey.

Table of Contents

Chapter 1

Introduction

1.1 Background and research gaps

Food, land use, and water are deeply interconnected. Agriculture stands at the centre of this nexus because it both depends on and reshapes freshwater resources, atmospheric flows, and land use. For this reason food production is among the largest driver of global environmental change. Agriculture occupies about 40% of the world's land, it accounts for up to 30% of global greenhouse gas emissions, and consumes around 70% of freshwater resources [1].

According to FAO, by 2050 agriculture will need to produce nearly 50% more food, fibre, and biofuel compared to 2012 levels [2]. Diets therefore represent a critical link between human health and environmental sustainability. Current dietary trends, combined with a projected global population of about 10 billion by 2050, are expected to intensify risks for both people and the planet [1].

Over the past 50 years, food production and dietary patterns have undergone profound transformations. While important progress has been achieved, such as reductions in hunger, improvements in life expectancy, and declines in infant and child mortality, new challenges have emerged. Diets increasingly feature high-calorie, heavily processed, and animal-source foods, placing mounting pressure on natural resources. These trends are driven by rapid urbanisation, rising incomes, and persistent barriers to accessing nutritious foods [3].

Projections indicate significant increases in global meat consumption: poultry consumption will increase by 21%, sheep by 16%, beef by 13%, and pig meat by 5% by 2034. Nearly 45% of this growth will occur in upper middle-income countries, driven by population and income expansion. For instance, Africa's population is expected to grow from 1.5 to 1.8 billion within the next decade, resulting in a 33% increase in regional meat consumption [3].

Numerous studies show that animal-source foods have far greater environmental impacts than plant-based foods across indicators such as greenhouse gas emissions [4], land use, energy consumption [1], and water use [5]. Figure 1.1 compares environmental effects by serving size, highlighting that ruminant meat has the highest impact among all food types. Future projections further suggest that environmental pressures will

intensify by 2050. While different food groups vary in their impacts, animal products consistently stand out as one of the most resource-intensive (Figure 1.2).

The EAT-Lancet Commission offers guidance on the necessary transformation of the global food system, calling for a substantial increase in the consumption of plant-based foods and a marked reduction in the consumption of animal-source foods. This transition is essential both to reduce environmental pressures and to improve human health outcomes [1].

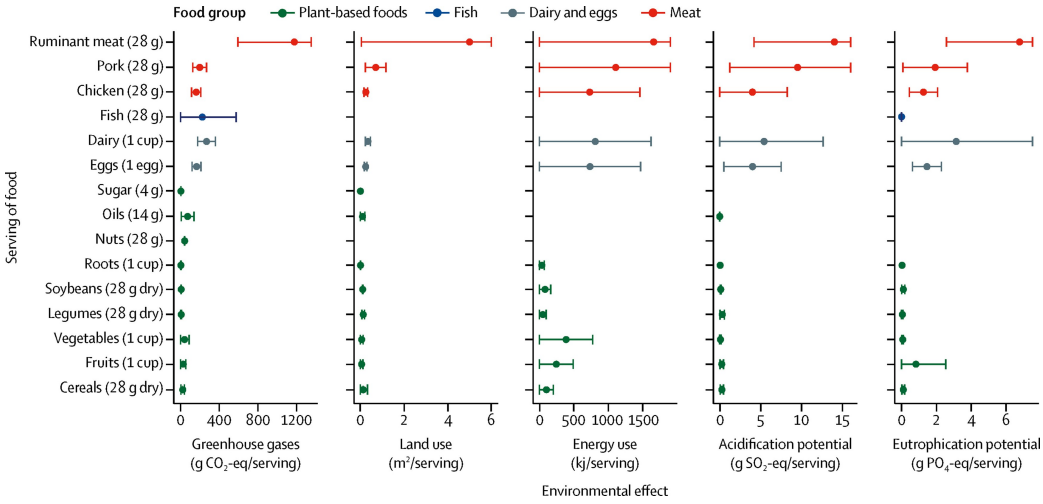


Figure 1.1: Environmental effects per serving of food produced in global food systems. Circles indicate mean values, and bars indicate standard deviations. Some results are missing for fish due to limited data in certain impact categories [1].

The growing global demand for food, feed, and flexible crops is placing unprecedented pressure on the hydrological cycle. Landscape conversion and rising irrigation needs are altering land-atmosphere feedbacks and reshaping moisture transport patterns [6]. Understanding these feedbacks between cropland water use and atmospheric processes is crucial for evaluating both agricultural sustainability and hydro-climatic risks.

Internal renewable water resources from rivers and aquifers, commonly referred to as *blue* water [7], are being extracted at rates that already exceed sustainable thresholds in several regions. Between 2000 and 2018, global per capita internal renewable water resources declined by about 20% [8]. These pressures are amplified by population growth, urban expansion, and the intensification of agriculture. In 2010, global water withdrawals were distributed as follows: 69% for agriculture, 12% for municipal use, and 19% for industry [2]. Therefore, agriculture is by far the largest user of freshwater, and within this share, a significant portion is devoted to feed crops that indirectly sustain livestock production. This adds complexity to the water-food nexus, since dietary choices strongly influence both the spatial distribution and intensity of water demand.

Yet most of the water used in food production is not blue water but *green* water [9], which includes terrestrial precipitation, evaporation, and soil moisture [10]. Between 1996 and 2005, the global water footprint of crop production was about 7,404 billion

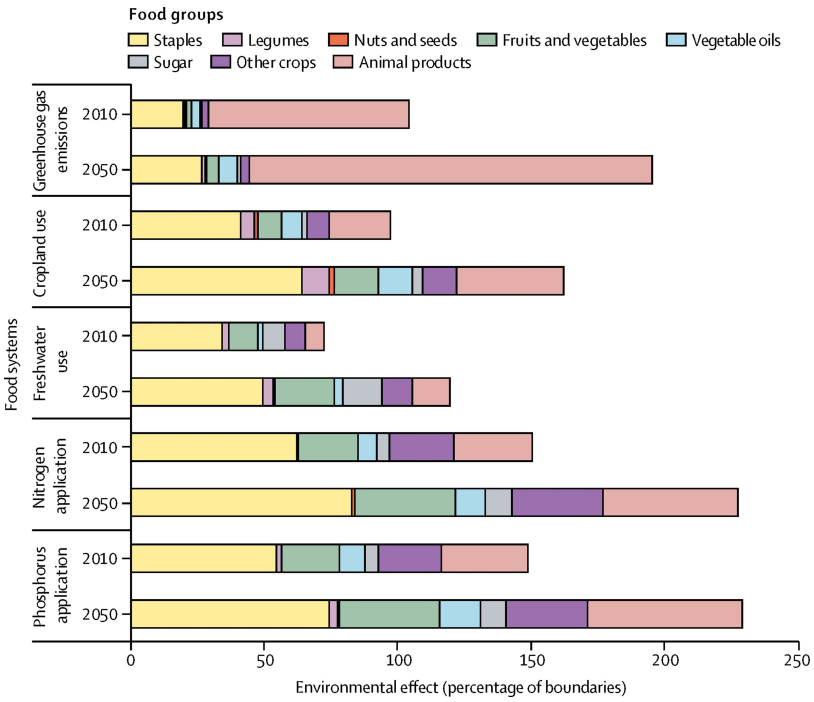


Figure 1.2: Environmental effects in 2010 and 2050 by food groups on various Earth systems based on business-as-usual projections for consumption and production [1].

cubic meters per year, of which 78% was green, 12% blue, and 10% grey [9]. While water footprint assessments quantify the volume of water consumed, they remain indicators of use rather than impact: they do not track how evapotranspired water re-enters the cycle, whether it returns as precipitation on the same crop, on other crops, or in entirely different regions.

Many regions experiencing chronic water scarcity relative to their populations rely heavily on agricultural commodity imports, effectively importing virtual water. Virtual Water Trade refers to the international or intra-national exchange of goods expressed in terms of the water embedded in their production: when goods are traded, the water physically consumed in the production area is virtually transferred to the region of consumption [11]. This line of research extends the water footprint concept to global supply chains, but both frameworks largely neglect the atmospheric component of water redistribution.

The origin and fate of precipitation are central to this perspective. Moisture tracking research has revealed that evapotranspiration sources and precipitation sinks are connected by atmospheric flows operating over thousands of kilometres [12, 13]. For example, the Amazon contributes 2-6% of rainfall to downwind regions in South America [14], highlighting the stabilizing role of intact ecosystems. Similar dynamics have been observed in Africa and Asia, where forests and wetlands act as rainfall sources. Globally, about 56% of terrestrial precipitation originates from the evapotranspiration of forests and other natural ecosystems, while rainfed agriculture contributes roughly 5% [2]. The remaining 39% of rainfall becomes surface runoff

(feeding rivers and lakes) or recharges groundwater aquifers [2]. These fluxes underpin the renewable freshwater resources on which societies depend. A portion of this water is withdrawn through infrastructure and diverted to different uses: much of it is eventually returned to rivers or aquifers, whereas water consumed through irrigation re-enters the atmosphere via evapotranspiration. The key question, therefore, is where and when this water returns as rainfall, and whether it supports agriculture again (green water) or replenishes blue water stocks such as aquifers and rivers.

Land-use change, also driven by food production, plays a critical role in shaping these fluxes. It has been demonstrated that land-use change can cause significant precipitation changes but only minor effects on runoff within the same basins [15], making it especially relevant for green water availability. Food-driven deforestation and land-use changes disrupt atmospheric fluxes by reducing evapotranspiration and inducing precipitation anomalies in downwind areas [14]. For instance, Amazonian deforestation, driven by soybean expansion, has caused precipitation declines in croplands in Argentina [16, 17]. Among agricultural commodities driving deforestation, cattle meat ranks first, while soybeans and maize, primary cattle feed components, rank fourth and fifth, respectively [18]. Consequently, dietary choices significantly influence not only freshwater use but also invisible atmospheric water fluxes.

Land-use change encompasses not only deforestation and cropland expansion, but also the transition from rainfed to irrigated agriculture. Irrigation withdraws large volumes of blue water, while also enhancing blue evapotranspiration fluxes into the atmosphere, thereby modifying local and downwind precipitation patterns [6]. While the total agricultural area has changed only modestly since 2000, the balance has shifted: land under permanent and irrigated crops has increased, while permanent meadows and pastures have declined substantially [8].

Current approaches often overlook the role of crops as distinct evaporative agents. Water footprint studies quantify consumption but not the atmospheric pathways through which evapotranspired water returns to precipitation. As presented in Figure 1.2, freshwater use and cropland use are treated as separate indicators, although they are deeply interconnected. Moisture tracking research maps these atmospheric flows, yet does not link them to specific crop types or to internal crop recycling. As a result, no global analysis has yet mapped crop-specific evapotranspiration sheds, nor traced how agricultural water use contributes to downwind rainfall and which land systems ultimately benefit from this recycled moisture.

1.2 Goal

This thesis addresses a critical gap in the literature by explicitly linking crop water use to atmospheric moisture recycling. It develops a framework that integrates crop-specific evapotranspiration, atmospheric moisture tracking, and land-use classification, with the aim of quantifying how agricultural water use recycles through the atmosphere and contributes to rainfall patterns worldwide.

Specifically, the work advances the understanding of how evapotranspiration from agricultural areas contributes to precipitation, both over agricultural land and other

surfaces. The analysis focuses on three pivotal crops, soy, maize, and wheat, which together account for about 33% of global harvested land [19] and roughly 30% of the global water footprint of crop production [9].

Beyond quantifying the magnitude of these feedbacks, the thesis maps their spatial distribution, identifying water donor and recipient regions of agricultural evapotranspired water. By integrating crop water accounting with moisture tracking, this approach bridges the gap between field-scale water use and atmospheric-scale precipitation dynamics, providing a global assessment of crop-specific evapotranspiration sheds and their implications for water resources, land-use planning, and sustainable agriculture.

1.3 Thesis structure and workflow

This thesis is structured as follows:

Chapter 2 lays out the theoretical foundations, describing the hydrological cycle, key agro-hydrological variables, and the concepts of moisture recycling, evaporation sheds, and precipitation sheds. It also reviews existing models for crop water estimation and atmospheric moisture tracking.

Chapter 3 presents the datasets employed, including global crop distribution maps, meteorological forcing data, and the RECON moisture connections dataset.

Chapter 4 details the methodological framework. The first part describes the waterCROP model, which updates crop water use accounting for 2008–2017 by solving the daily soil water balance on a global 5 arc-minute grid. A new version of the code was developed to incorporate daily meteorological data and enhance computational performance. The second part integrates these crop-specific water use estimates with RECON outputs to derive crop-specific evapotranspiration sheds, annual water balances, and land-use classifications of sinks.

Chapter 5 presents the key findings, including crop-specific evapotranspiration sheds, donor–recipient balances, and the classification of precipitation sinks. The results are visualized through maps and Sankey diagrams to illustrate the redistribution of agricultural evapotranspired water across various land types.

Finally, Chapter 6 synthesizes the results, highlights their relevance and potential applications, and outlines the limitations of the current analysis and directions for future research.

In summary, this thesis contributes a new perspective on agricultural water use by extending traditional water footprint assessments to include atmospheric feedbacks. By systematically quantifying how crop evapotranspiration contributes to rainfall, it provides both conceptual advances for the scientific community and practical insights for sustainable water and land management.

Chapter 2

Theoretical foundations

This chapter outlines the theoretical framework that underpins the present research. It begins with an overview of the hydrological cycle. Emphasis is placed on its fundamental role in sustaining human societies, with particular attention to agricultural production, where water availability directly conditions crop growth and productivity. Subsequently, a set of agro-hydrological variables that are recurrently employed in this thesis are introduced and rigorously defined, thereby establishing a consistent terminology for the subsequent analyses. Finally, the discussion focuses on two key components of the hydrological cycle that are central to this study: (i) the quantification of crop evapotranspiration, approached through agro-hydrological modelling frameworks, and (ii) the estimation of atmospheric moisture fluxes, which provide insight into large-scale water transport and its implications for regional and global hydrological balances.

2.1 Hydrological cycle

The global hydrological cycle regulates the functioning of the Earth system and provides the basis for all life. It regulates climate, enables the cycling of carbon through biomass production, and governs the transport of nutrients, chemicals, and pollutants across ecosystems [20, 21]. Conceptually, the hydrological cycle describes the continuous circulation of water on, above, and below the Earth's surface. This circulation is primarily driven by solar radiation and gravity, which together sustain the transfer of water across its different physical phases and reservoirs, including the atmosphere, oceans, terrestrial ecosystems, and groundwater.

Water enters the atmosphere via evaporation from oceans, inland water bodies, and soils, as well as through transpiration from vegetation. Once in the atmosphere, it is transported as vapour, undergoes condensation, forms clouds, and eventually returns to the surface as precipitation. Precipitation constitutes the ultimate source of freshwater and initiates a cascade of fluxes and storages that sustain terrestrial hydrology. Upon reaching land, precipitation may infiltrate into soils, generating soil moisture; flow across the surface as runoff, feeding streams, rivers, and wetlands; or evaporate directly from vegetation canopies, bare soils, and standing water.

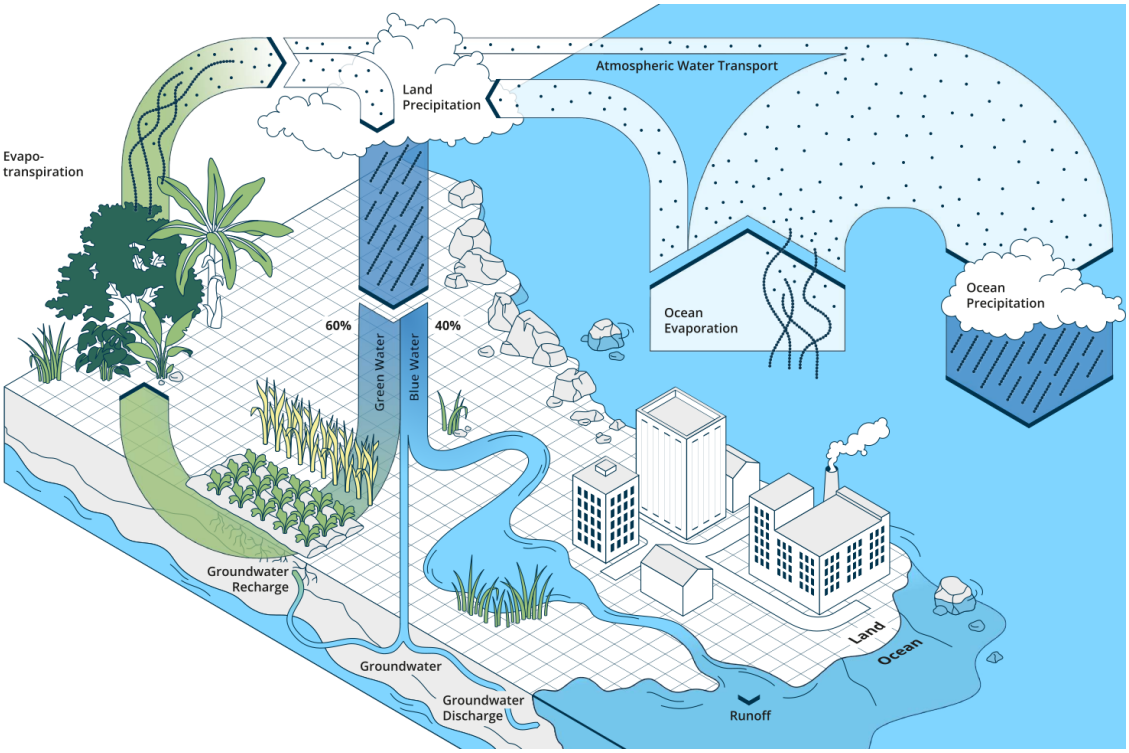


Figure 2.1: Main global hydrological flows distinguished into blue and green contributions, depicted with proportional arrows according to volume estimates in the latest IPCC AR6 Assessment [22]. [20]

The fate of infiltrated water can be further differentiated into green and blue water (Figure 2.1). *Green* water refers to the portion of soil moisture that is accessible to plants within the root zone and subsequently returned to the atmosphere through transpiration or direct evaporation from soil and vegetation surfaces [7]. *Blue* water, in contrast, corresponds to liquid water available in rivers, lakes, reservoirs, and aquifers. While part of the infiltrated water contributes to green water flows, a fraction percolates below the root zone, recharging groundwater and sustaining subsurface flows that eventually feed back into surface water bodies [7]. Distinguishing between stocks and flows is fundamental: blue water stocks encompass water stored in lakes, reservoirs, aquifers, glaciers, and snow, while blue water flows include river runoff and subsurface recharge. Similarly, green water stocks are defined as soil moisture and plant-held water, whereas green water flows comprise evapotranspiration fluxes. Furthermore, these categories are highly interconnected; for example, irrigation (blue water flow) applied to a field increases soil moisture (green water stock), which then sustains transpiration and evaporation (green water flows). On average, at the global annual scale, approximately 60% of precipitation over land is partitioned into green water and 40% into blue water, highlighting the predominance of green water as the main freshwater resource for ecosystems and agriculture [22].

A central component of the hydrological cycle is evapotranspiration, the combined flux of evaporation and transpiration. Evapotranspiration links the hydrological and

energy cycles: increasing air temperatures elevate both the atmosphere's capacity to hold water vapour and its evaporative demand. Consequently, as global warming intensifies, land and oceans evaporate more water, reinforcing the greenhouse effect through the positive feedback of water vapour on surface warming. Evapotranspiration can itself be partitioned into green and blue components. Green evapotranspiration is sustained by soil moisture originating from precipitation, whereas blue evapotranspiration refers to water consumed by plants that derives from irrigation or other managed withdrawals from blue water stocks [23, 9].

2.2 Important agro-hydrological variables

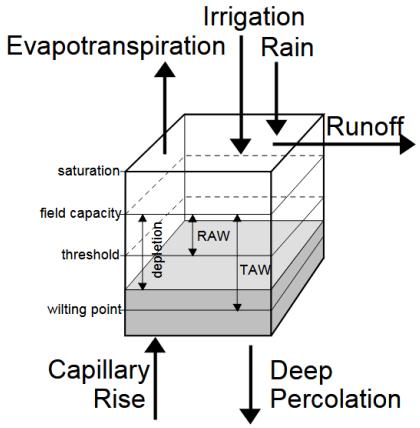


Figure 2.2: Water balance of the root zone [24].

Agro-hydrology involves multiple variables and coefficients to describe the interactions among soil, plants, and atmosphere. To ensure consistency, the Food and Agriculture Organization (FAO) provides widely adopted guidelines that establish a common language and standardized conventions. This thesis follows the FAO framework when dealing with agro-hydrological modelling, while integrating it with a few additional relevant concepts.

Figure 2.2 provides a schematic representation of the water balance in the root zone. It highlights the key fluxes, precipitation, evapotranspiration, irrigation, runoff, and deep percolation, as well as the variables related to soil moisture that are central to agro-hydrological modelling. The variables discussed below are directly linked to the processes shown in the figure, which serves as a useful visual reference for understanding how each component contributes to the overall soil–plant–atmosphere system.

The most important variables considered are listed and explained below:

- **Potential Evapotranspiration** (ET_0): The amount of evapotranspiration that would occur from a reference surface under optimal conditions. The reference surface is defined as a hypothetical crop with a height of 0.12 m, a surface

resistance of 70 s m^{-1}, and an albedo of 0.23, closely resembling the evaporation from an extensive, uniformly green grass surface that is actively growing and adequately watered. This reference crop allows estimation of evapotranspiration that represents the atmospheric demand for water. ET_0 is commonly calculated using the Penman–Monteith equation, which requires meteorological inputs such as radiation, air temperature, humidity, and wind speed [24].

- **Total Precipitation (P_{TOT})**: The total amount of rainfall and other forms of precipitation that reach the ground over a given period (e.g., daily, monthly).
- **Soil Moisture (θ)**: The volumetric water content of the soil, typically expressed as a fraction or percentage.
- **Drainage or Deep percolation (D)**: The downward movement of water beyond the root zone, contributing to groundwater recharge and loss of root zone water.
- **Runoff (R)**: The portion of precipitation or irrigation water that flows over the land surface without infiltrating into the soil.
- **Rooting Depth (Z_r)**: The depth of soil actively explored by crop roots, which affects water uptake capacity. It usually varies with crop type and growth stage.
- **Total Available Water (TAW)**: The amount of water available in the root zone between field capacity and permanent wilting point.

$$TAW = (\theta_{FC} - \theta_{WP}) \cdot Z_r$$

where θ_{FC} is the soil moisture at field capacity and θ_{WP} is the soil moisture at wilting point [24].

- **Readily Available Water (RAW)**: The portion of TAW that plants can extract without experiencing water stress. It is often a fraction of TAW, defined by a depletion fraction p:

$$RAW = p \cdot TAW$$

- **Crop Evapotranspiration (ET_c)**: The amount of water evapotranspired by the crop under optimal (no water stress) conditions, which depends on crop characteristics.

$$ET_c = k_c \cdot ET_0$$

where k_c is the crop coefficient.

- **Actual Crop Evapotranspiration (ET_a)**: The actual amount of water evapotranspired by the crop, which depends on water availability and crop characteristics.

$$ET_a = k_s \cdot ET_c = k_s \cdot k_c \cdot ET_0$$

where k_s is the water stress coefficient and k_c is the crop coefficient.

- **Root Zone Depletion (Dr)**: The water deficit relative to field capacity.

$$Dr(t) = Dr(t-1) - (P_{TOT}(t) - R(t)) - I(t) + ET_a(t) + D(t)$$

where $I(t)$ is the irrigation and $Dr(t-1)$ is the water deficit of the previous time step (usually the initial condition is to consider it zero, because at beginning the root zone is at field capacity).

- **Green Evapotranspiration (ET_{green})**: The portion of evapotranspiration sustained by soil moisture derived from precipitation stored in the root zone (rainwater). It represents the evapotranspiration of green water accessible to plants [23].

- **Blue Evapotranspiration (ET_{blue})**: The portion of evapotranspiration that relies on irrigation water sourced from surface or groundwater (blue water). It represents the consumptive use of irrigation water [23].

$$ET_{blue} = ET_a - ET_{green}$$

- **Crop Coefficient (k_c)**: A dimensionless factor that represents the ratio of crop evapotranspiration to reference evapotranspiration, accounting for crop type and growth stage under optimal (no water stress) conditions.

$$k_c = \frac{ET_c}{ET_0} = f(\text{crop height, albedo, canopy resistance, soil evaporation})$$

Since it depends on crop characteristics that change during the growing season, it is not constant over time [24].

- **Water Stress Coefficient (k_s)**: A dimensionless factor (between 0 and 1) that reduces evapotranspiration due to limited soil water availability [24].

$$k_s = \frac{TAW - Dr}{TAW - RAW}$$

2.3 Agro-Hydrological models

Agro-hydrological models are essential tools for understanding and quantifying the interactions between agricultural practices and the water cycle. By simulating processes such as soil water dynamics, crop growth, evapotranspiration, and irrigation management, these models provide valuable insights for water resource management, crop yield optimization, and sustainability assessments. They are widely applied to estimate crop water requirements and irrigation scheduling, evaluate the impacts of land use and climate change on water resources, support decision-making in agricultural planning and water management, and assess the sustainability of cropping systems in relation to water availability. Beyond practical applications, these models serve as a critical framework for scientific research and policy development, enabling

evidence-based strategies for sustainable agriculture and resilient water management under changing climatic and socio-economic conditions.

Several well-established agro-hydrological models have been developed specifically to simulate crop systems, rather than vegetation in general, and are widely used in research and practice. Considering the focus of this study on crop evapotranspiration, particular attention is given to crop-oriented agro-hydrological models.
The AquaCrop model, developed by FAO, focuses on simulating crop growth and yield under water-limited conditions, emphasizing the effects of water stress on productivity [25]. CROPWAT, also developed by FAO, is a simpler model primarily designed to estimate crop water requirements and irrigation scheduling based on climate, soil, and crop data [26]. CropSyst is a multi-year, multi-crop, daily time-step cropping systems simulation model developed to study the effects of climate, soils, and management on cropping system productivity and the environment. It simulates the soil water and nitrogen budgets, crop growth and development, crop yield, residue production and decomposition, soil erosion by water, and salinity [27]. All results obtained in this thesis relied on the waterCROP model, which is a physically based agro-hydrological model. It describes the main components of the soil–atmosphere–plant continuum (such as effective precipitation, leakage, and evapotranspiration) as functions of soil, crop, and growth stage during the season [28, 29]. While these examples illustrate some of the most widely used models, it is important to note that many others exist, varying in complexity, scale, and intended applications, and collectively they have contributed significantly to improving water management and supporting sustainable agriculture worldwide.

2.4 Atmospheric moisture recycling

Continental moisture recycling, the process through which terrestrial evapotranspiration returns as precipitation over land, is a fundamental component of the Earth system [12]. It shapes regional rainfall patterns, influences the spatial propagation of droughts, and determines whether continental interiors receive sufficient precipitation to sustain agriculture [14]. Because evapotranspiration flows can travel thousands of kilometres before re-precipitating, land-use changes such as deforestation or agricultural expansion may alter downwind precipitation regimes, drought severity, and hydrological dynamics. Although these teleconnections vary over time, they tend to follow consistent spatial patterns, making them a valuable indicator of land–atmosphere interactions [30].

Estimates indicate that about 45% of terrestrial evapotranspiration contributes to precipitation over land, underscoring the pivotal role of land surfaces in maintaining continental rainfall and agricultural productivity [31]. This highlights the need to explicitly account for atmospheric moisture flows in assessments of water resources and land management, as land-cover changes can trigger cascading impacts on regional and global hydrological stability. This thesis aim to further assess the crop-specific evapotranspiration which contributes to the precipitation over croplands.

The analyses addressed in this study are conducted on a yearly base. While the

hydrological cycle may not be balanced at local or regional scales within a single year, it is generally considered closed at the global annual scale [32, 33]. This implies that the total volume of water evapotranspired over the course of a year (in this thesis, from croplands) is expected to return to the Earth's surface as precipitation within the same year. Although spatial and temporal mismatches may occur, such as evapotranspiration in one region leading to precipitation in distant or delayed locations, on average, the global annual water balance remains conserved. The use of moisture tracking models and the analysis of evaporation sheds are used to investigate these mismatches.

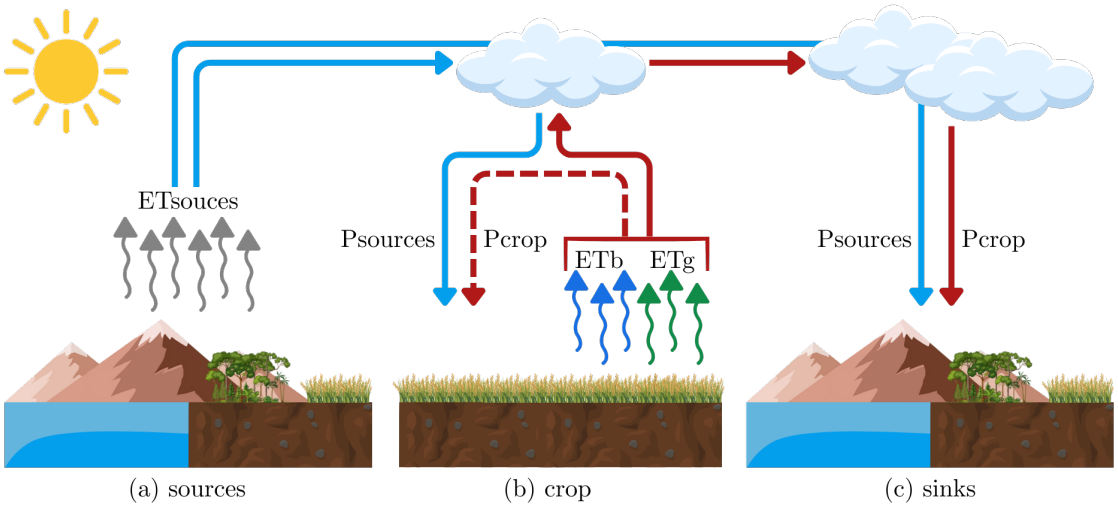


Figure 2.3: Schematic representation of the atmospheric branch of the water cycle, emphasizing two processes: moisture recycling that contributes to cropland precipitation, and moisture recycling resulting from crop evapotranspiration. Adapted from [14, 34]. Red arrows indicate the water pathway used to assess the contribution of cropland evapotranspiration to precipitation in downwind areas, with the dashed line illustrating local moisture recycling within the same region.

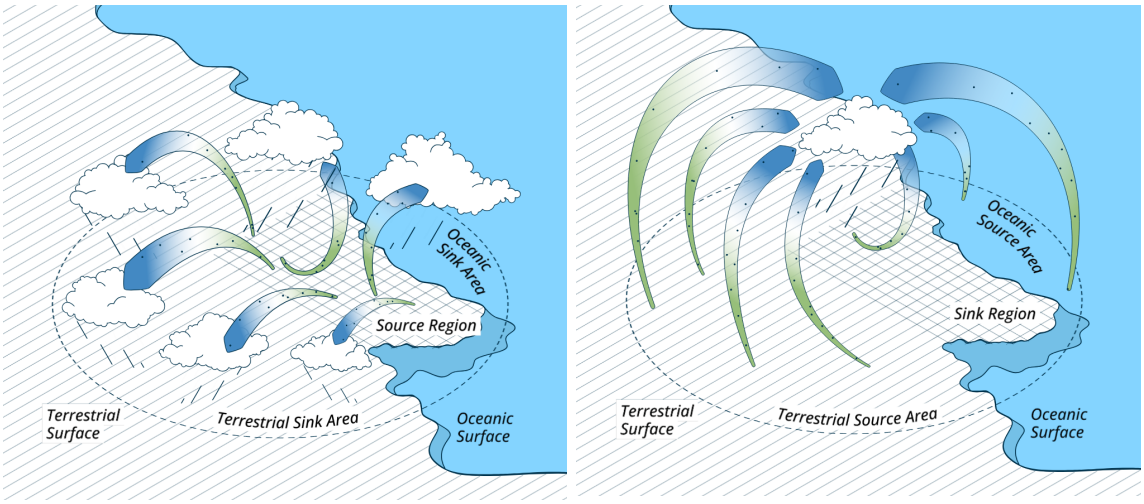
As shown in Figure 2.3, water is transported through the atmosphere, precipitates, and is subsequently re-evaporated from land surfaces. Panel (a) illustrates different moisture sources ($ET_{sources}$) that contribute to precipitation ($P_{sources}$), with a focus on agricultural land shown in panel (b). Cropland evapotranspiration ($ET_b + ET_g$) results in precipitation over croplands (P_{crop}), partly returning locally (local moisture recycling) and partly falling in downwind regions, as shown in panel (c).

2.4.1 Evaporation and Precipitation Shed

In the literature, two complementary concepts are commonly used to describe atmospheric moisture recycling: the *evaporation shed* and the *precipitation shed*. Together, they define the spatial links between evaporation and rainfall. The evaporation shed (Figure 2.4a) refers to the downwind regions (land or ocean) where precipitation

originating from evaporation in a given location is expected to fall. Conversely, the precipitation shed (Figure 2.4b) represents the upwind regions (land or ocean) that supply the evaporation contributing to precipitation in a given location.

In the context of croplands, these definitions can be adapted as follows: the evapotranspiration shed indicates the areas where crop evapotranspiration is expected to return as precipitation, while the precipitation shed identifies the areas from which the rainfall infiltrating into the soil of croplands originates.



(a) **(b)**

Figure 2.4: Schematic representations of (a) evaporation shed and (b) precipitation shed [20].

2.4.2 Moisture Tracking Models

Moisture recycling models are widely used to trace the movement of water through the atmosphere. Based on their spatial representation they can be classified as either Eulerian, which are grid-based, or Lagrangian, which are trajectory-based. In Eulerian models, moisture is exchanged between discrete grid cells at each time step, whereas in Lagrangian models, individual air parcels are tracked as their positions evolve over time [35].

Beyond this fundamental difference, all moisture-tracking approaches require assumptions concerning vertical mixing of moisture, the integration time step, interpolation methods, and the spatial and temporal resolution of the forcing dataset. Consequently, each study must adopt a set of assumptions that balances accuracy in representing evaporation sheds with computational demands, data availability, and simulation time [12].

Among the most widely adopted Eulerian approaches is the Water Accounting Model 2-layers (WAM-2layers), currently available in its third version [36]. This three-dimensional model simplifies the vertical dimension into two layers and tracks the transport of atmospheric moisture from sources (surface evaporation) to sinks

(precipitation), or vice versa. Its spatial resolution directly matches that of the input forcing data.

A more recent but already widely used alternative is UTrack [12]. UTrack employs a Lagrangian framework, tracking moisture flows globally at a resolution of 0.25°. It provides a comprehensive database of atmospheric moisture pathways, including monthly multi-annual means for the period 2008–2017. UTrack is forced with ERA5 hourly atmospheric reanalysis data (0.25° horizontal resolution) and uses three-dimensional fields on 25 tropospheric pressure levels, thereby capturing detailed patterns of moisture transport.

The dataset employed in this thesis is RECON [37], a global atmospheric moisture connections dataset in NetCDF format. RECON is a post-processed version of UTrack, providing annual moisture flow volumes (in cubic meters) between evaporation sources and precipitation sinks. It offers global coverage at 0.5° resolution, averaged over 2008–2017. While the monthly resolution of UTrack is lost, RECON enforces the closure of the hydrological cycle at the annual scale, ensuring consistency between global evapotranspiration and precipitation.

2.5 Land and water use changes effects on atmospheric moisture flows

Agricultural production, along with its hydrological consequences, expanded dramatically during the 20th century and is projected to continue rising throughout the 21st century. This expansion is clearly visible in the trends of areas dedicated to pasture, cultivated land, particularly irrigated cropland, the volumes of water withdrawn for irrigation, and the quantities of fertilizers applied. Agriculture represents both a key driver of land-use change and a critical component of the global water cycle [6]. In this thesis, within the broader interconnection between agriculture and the water cycle, particular emphasis is placed on crop-specific evapotranspiration and its role in moisture recycling.

Human-driven land-use changes affect nearly every component of the hydrological cycle [38]. Such changes influence how precipitation is distributed, how water infiltrates soils, moves through rivers and streams, or accumulates as surface flooding. They also alter the return of moisture to the atmosphere through evaporation and transpiration. Because the water cycle is highly interconnected, even localized changes can cascade through the system, ultimately reshaping freshwater availability at regional and global scales [22].

An emblematic example of land-use change related to the agricultural sector is deforestation. Clearing forests reduces soil moisture, evaporation, and local rainfall, while also triggering regional temperature shifts that influence precipitation regimes. Vegetation regulates these processes through transpiration, the uptake and release of water via stomata. As described in Section 2.4, about 45% of terrestrial precipitation originates from land evapotranspiration. Thus, changes in vegetation cover not only reshape infiltration and runoff but also directly affect atmospheric moisture recycling.

Agricultural expansion is a leading driver of deforestation: between 2001 and 2022,

an estimated 86% of global deforestation was linked to crop and cattle production [39]. Feed is the main connection between livestock and land use, both directly through grazing and indirectly through the consumption of cultivated grains and forage [40]. According to FAOSTAT, around 40% of global cereal production in 2022 was used for animal feed, a proportion that has been steadily increasing. This diversion of cropland to feed livestock amplifies pressures on both land and water resources.

Water use is reshaped not only by cropland expansion but also by the transition from rainfed to irrigated agriculture. Although irrigation may not drive a physical "land-use change," it substantially increases pressures on the water cycle. Irrigation withdraws large volumes of blue water, while also enhancing blue evapotranspiration fluxes into the atmosphere, thereby modifying local and downwind precipitation patterns [6].

Improved land and water management strategies, such as reforestation, sustainable irrigation, and conservation agriculture, offer pathways to reduce climate impacts while adapting to adverse changes already underway. Overall, evidence shows that changes in land use and land cover alter the water cycle at local, regional, and global scales, reshaping precipitation, evaporation, flooding, groundwater dynamics, and freshwater availability [22].

Chapter 3

Data source and processing

3.1 Climate Data

Originally, the waterCROP model was based on the CRU TS v. 2.0 high-resolution gridded dataset, which provides long-term averages, with a monthly temporal resolution and a spatial resolution of 5 arc-minutes [41]. The model simulates daily conditions by interpolating monthly data, as explained in detail in Subsection 4.1.2.

To ensure consistency with the RECON dataset, which will be combined with waterCROP results to identify sink regions, an update of the climatology was necessary. The RECON dataset is based on monthly ERA5 data from 2008 to 2017 with a spatial resolution of 0.5°. Since ERA5 data [42] are also available at daily resolution, the waterCROP model was modified to use actual daily data instead of relying on interpolation from monthly values.

The data were downloaded directly from the Copernicus Climate Data Store using their API. Downscaling to match the desired spatial resolution was performed using Climate Data Operators (CDO), which provides more than 600 operators to manipulate and analyse climate data. The processing employed the CDO function `remap`, selecting a conservative remapping method that preserves the integral of dataset values over the domain during interpolation.

While CRU TS data represent statistical averages accounting for the stochastic nature of precipitation, ERA5 data were used without any averaging. As explained later, the updated daily version of the waterCROP model runs over the entire period to derive an average yearly crop behavior, rather than calculating it directly from an averaged climatology.

Since the daily ERA5 data are heavy and the simulation which involved their use requires the use of HPC, monthly averages were also calculated starting from the previously downscale version of the daily data. The processing employed the CDO function `ymonmean` which calculate the mean for each month of each year.

3.2 Crop Data

In this thesis, all analyses were conducted on three major crops—wheat, maize, and soy—which together occupied about 33% of global harvested areas. Rainfed and irrigated harvested areas were sourced from the MapSPAM 2010 v2.0 dataset provided by the International Food Policy Research Institute (IFPRI) [19]. This dataset has global coverage, includes 42 crops, and has a spatial resolution of 5 arc-minutes.

To represent specific crops, the waterCROP model requires crop-specific data as input. The sowing dates and the length of the growing period (LGP) were obtained from the global dataset MIRCA2000 [43], which has a spatial resolution of 5 arc-minutes and distinguishes between rainfed and irrigated production.

The daily crop coefficient k_c is used in the calculation of actual evapotranspiration. The model computes k_c on a daily basis following predefined curves [28], which are divided into four stages: initial phase, development stage, midseason, and late season. Each stage has a specific length, and the corresponding constant values were adopted from Allen et al. (1998) [24]. Moreover, the length of each stage is expressed as a fraction of the LGP, with values defined according to Mekonnen and Hoekstra (2011) [9] for different climatic regions.

The daily water stress coefficient $k_{s,j}$ also contributes to the calculation of actual evapotranspiration. For irrigated production, $k_{s,j}$ is set equal to 1 throughout the growing period, while for rainfed production it is computed daily following Tuninetti et al. (2015) [28]. Its calculation uses 30 arc-second maps of available water content (AWC) from FAO/IIASA/ISRIC/ISSCAS/JRC (2012) [44], aggregated to match the spatial resolution of the model. Additional inputs to calculate $k_{s,j}$ include precipitation (as described above), rooting depth Z_r and the depletion fraction p, both from Allen et al. (1998) [24]. The rooting depth is assumed to be maximal in rainfed areas and minimal in irrigated areas.

3.3 RECON dataset

As presented in Subsection 2.4.2, this thesis adopted the global atmospheric moisture connections NetCDF dataset called RECON [31] to connect the crop-specific evapotranspiration to their sink areas. The dataset is open-source and can be easily download [37] but it easy heavy, so also its use requires the use of HPC. Since it offers global coverage at 0.5° resolution, while waterCROP has a finer resolution, all the output after the combination of this two will have a resolution of 0.5°.

3.4 Land use datasets

To maintain consistency with the data, the same harvested areas used to run the waterCROP model were also employed to characterize the regions identified as sinks. Since the original MapSPAM 2010 v2.0 dataset [19] has a spatial resolution of 5 arc-minutes, it was aggregated to match the output coarser resolution of 0.5°. This

aggregation was performed using the CDO function `gridboxsum`, which generates each output cell by summing the hectares of the corresponding 6×6 input cells.

The ISIMIP3 land-sea masks [45] were used to estimate the volume of water evaporated from crop cultivations that does not precipitate over land but instead falls over the oceans. The dataset has a spatial resolution of 0.5°, consistent with the RECON dataset, and therefore requires no additional processing. The mask assigns a value of NaN to ocean-covered cells and a value of 1 to land-covered cells.

Table 3.1: Data summary with scale, resolution, reference period and source.

Data	**Scale**	**Resolution**	**Period**	**Source**
Daily Potential Evapotranspiration ET_0	Global	0.5 deg	2008-2017	Hersbach et al. (2020) [42]
Daily Total Precipitation P_{TOT}	Global	0.5 deg	2008-2017	Hersbach et al. (2020) [42]
Rainfed and irrigated harvested areas	Global and crop specific	5 arc-min	2010	IFPRI (2019) [19]
Sowing dates	Global and crop specific	5 arc-min	-	Portmann et al. (2010) [43]
Length of the growing period LGP	Global and crop specific	5 arc-min	-	Portmann et al. (2010) [43]
Crop coefficient k_c	Crop specific	-	-	Allen et al. (1998) [24]
Length of the growing period LGP subdivisions	Global and crop specific	5 arc-min	-	Mekonnen and Hoekstra (2011) [9]
Climatic regions	Global	5 arc-min	-	Mekonnen and Hoekstra (2011) [9]
Available soil water capacity AWC	Global	30 arc-sec	-	FAO/IIASA/ISRIC/ISSCAS/JRC (2012) [44]
Rooting depth ranges	Crop specific	-	-	Allen et al. (1998) [24]
Depletion fraction p	Crop specific	-	-	Allen et al. (1998) [24]
RECON dataset	Global	0.5 deg	2008-2017	De Petrillo et al. (2024) [37]
Land-sea mask	Global	0.5 deg	2020	Lange, Buchner (2020) [45]

Chapter 4

Methods

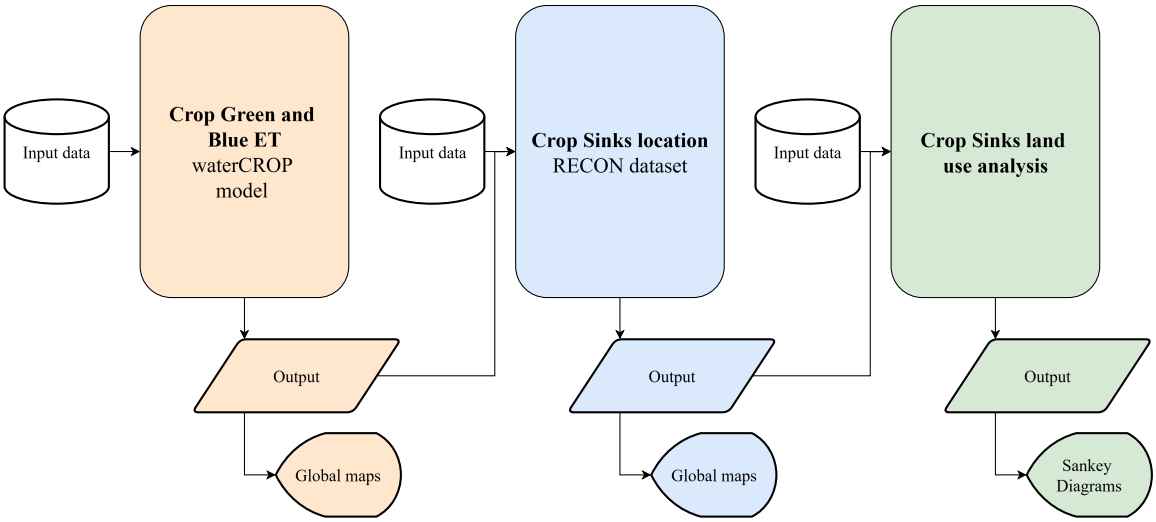


Figure 4.1: Flowchart describing the methodology adopted in this thesis.

The methodology developed in this work has been proposed to produce and investigate the location of water sinks of water which originate from specific crop cultivations. Figure 4.1 illustrate the overview of the methods followed to reach the desired results. Each section will be breaked down in the following paragraphs. Most of the analysis were performed using HPC, since the simulations deal with heavy input files.

4.1 Crop Green and Blue Evapotranspiration evaluation with the waterCROP model

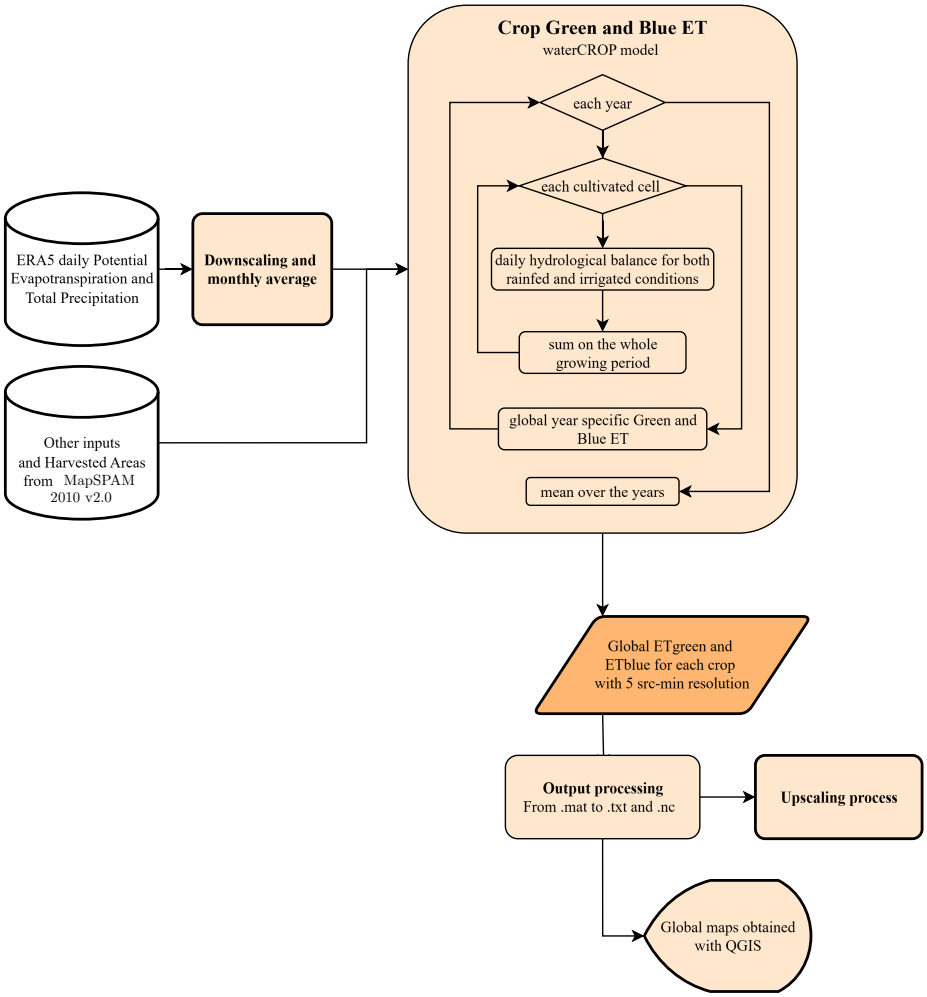


Figure 4.2: Flowchart describing the methods to obtain crop green and blue evapotranspiration adopting the waterCROP model.

The goal of this first part of the thesis is to update production-based water accounting for the period 2008-2017. To accomplish this goal was used the waterCROP model [28, 29], a physically-based Agro-Hydrological model which solves soil water balance on a daily basis running on a global grid of 5 arc-minute resolution. It describes the main components of the soil-atmosphere-plant continuum (such as effective precipitation, leakage, evapotranspiration, etc.) as a function of soil, crop, and period during the growing season. The model maps average annual water use for both food and feed crops. Each crop has maps of actual green and blue evapotranspiration, and irrigation demand evaluated to compensate for soil water stress over the whole growing period. This step enables the identification of how much water enters the atmosphere through evapotranspiration and where, providing the basis to assess the destination of that

water within the hydrological cycle.

4.1.1 Hydrological balance

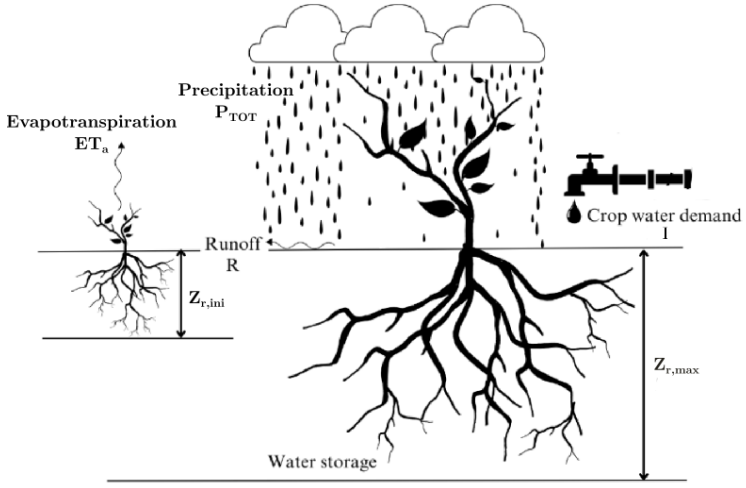


Figure 4.3: Water balance components of the waterCROP model: Total Precipitation P_{TOT}, Runoff R, Actual Evapotranspiration ET_a and Crop Water Demand (or Irrigation) I. $Z_{r,ini}$,$Z_{r,max}$ stand for initial rooting depth, maximum rooting depth respectively. Adapted from [29]

The waterCROP model was run to estimate crop evapotranspiration over a single growing season. It simulates the daily hydrological balance to estimate the daily actual evapotranspiration, $ET_{a,j}\left[\frac{\text{mm}}{\text{day}}\right]$, for each day of the growing period. All details of the daily estimation remain unchanged from the original model. Figure 4.3 summarizes all the variables considered in the daily water balance. Few assumption are made and must be taken into account while examinating the results. On the first day of the growing period the soil is considered to be at field capacity, the crop is far from water stress condition. In the irrigated areas, it is assumed that the field never suffer water stress, so each time the deficit at the end of the day is higher than the Readily Available Soil Water (RAW or θ^*) and the Total Precipitation P_{TOT} of that day is not enough to bring it back to a lower value, the difference is closed by Irrigation I. The daily values $ET_{a,j}$ are then summed over the growing season to obtain the annual evapotranspiration, $ET_{a,LGP}$ [mm], for a specific year. This value is subsequently averaged with results from other years to produce the annual average.

4.1.2 Daily model

The original model has a 360-days calendar and each month have the same length (30 days) and was runned with a average climatology computed on a 30-years time period. It evaluates the hydrological balance daily, but the Potential Evapotranspiration ET_0 and the Total Precipitation P_{TOT} are obtained from monthly data. Specifically the

ET_0 is obtained placing the monthly average value on the 15th day of the month and assuming a linear behaviour between them, while the monthly cumulative P_{tot} is evenly distributed on the 30 days for each month.

The new version of the model developed for this thesis uses actual daily data, as presented in the Section 3.1. The model run for all the years of the selected period, in this case 2008-2017, and then average all the years, in order to obtain a mean crop behaviour. This shift from monthly to daily data implicate that other changes need to be done. Since the sowing dates file is based on a 360-days calendar, the new code convert those information in a 365-days calendar, taking also into account the leap years.

This new version of the code handles large input datasets and therefore requires substantial computational time and memory. To evaluate whether this level of detail was necessary for the scope of the thesis, the daily and monthly versions of the model were compared at the annual scale, since evapotranspiration ET values are ultimately aggregated yearly. The results show that the discrepancy is negligible, with a correlation higher than 0.98 between the two versions (Figure 4.4). For this reason, the simulations presented in this thesis were carried out with the monthly model, with minor adjustments.

To further assess the differences in terms of spatial distribution, a series of global maps was produced (Figure 4.5), displaying the relative difference defined as:

$$\varepsilon_i[-] = \frac{ET_{daily,i}[m^3] - ET_{monthly,i}[m^3]}{ET_{daily,i}[m^3]}.$$

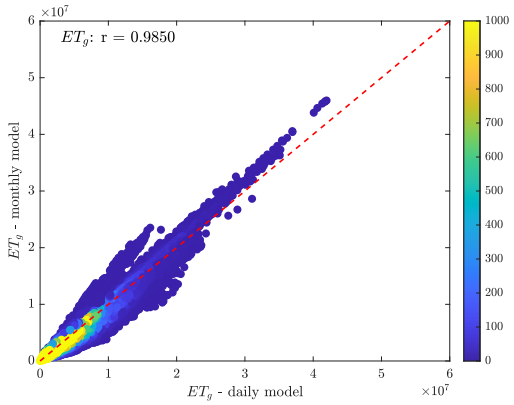
Assuming the daily version of the model to be more accurate, the monthly version appears to underestimate irrigation requirements in certain regions (e.g., the USA, Ukraine, France, and Spain), while overestimating them in Italy, Greece, India, North Korea, and Japan. The Green ET shows only minor differences, with notable exceptions such as the Virginia region in the USA. Overall, the differences in Total ET largely mirror the behaviour of Green ET, as Blue ET is considerably smaller.

Nevertheless, the daily version remains relevant, since its higher temporal resolution would significantly affect results aggregated at the monthly scale, even if the annual totals remain nearly unchanged.

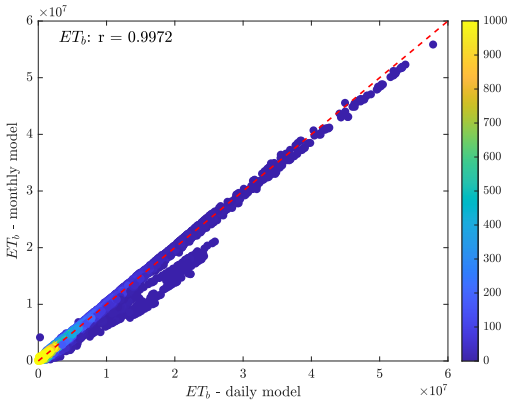
4.1.3 Cell insights

To better illustrate the behaviour of the model in daily simulations, a series of plots was generated. These plots visualize the daily evolution over the growing period for a single grid cell, under both rainfed and irrigated conditions. Specifically, they include:

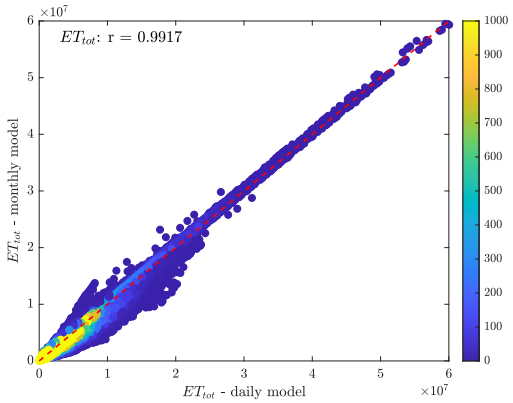
- the crop coefficients,
- the rooting depth, compared against the Readily Available Soil Water (RAW or θ^*) and the Total Available Soil Water (TAW or θ_{WP}),
- the daily hydrological balance in the soil (depletion D_r evolution) and in the atmosphere (evapotranspiration, ET, evolution).



(a) Green ET.



(b) Blue ET.



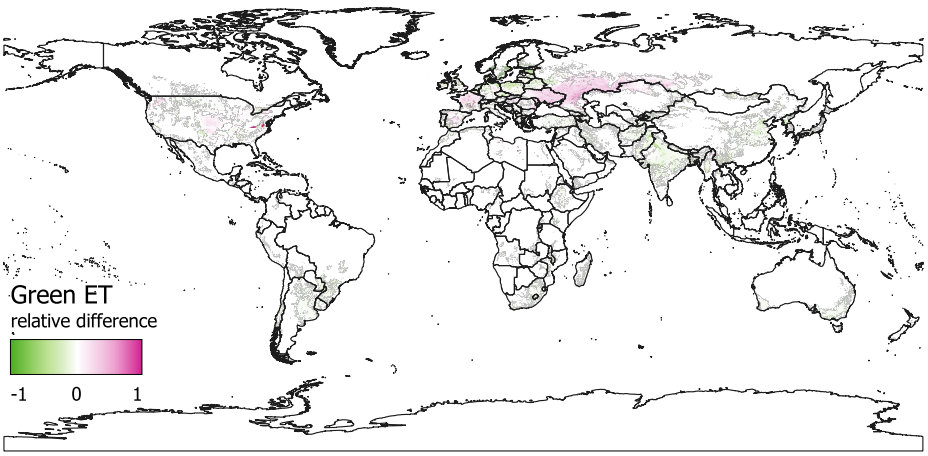
(c) Total ET.

Figure 4.4: Scatter plot and correlation of annual ET obtained with the daily and monthly models. Results are based on wheat cultivations averaged over 2008-2017.

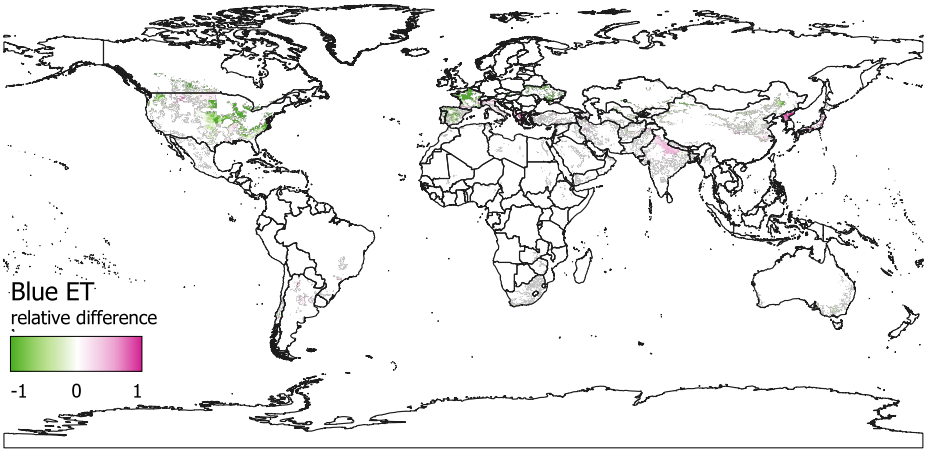
In addition, two further plots summarize the dynamics over an entire hydrological year. The first highlights the timing and potential overlap of rainfed and irrigated growing seasons, while the second presents the evolution of green and blue evapotranspiration (ET) from the total harvested area of the cell throughout the year. The annual plots are aligned with the hydrological calendar: in the Northern Hemisphere they begin on October 1st, whereas in the Southern Hemisphere they begin on July 1st.

Figures 4.6 to 4.9 show the plots described above, intended to demonstrate the model's functioning. The selected grid cells correspond to two distinct regions: Piedmont in Italy and Minas Gerais in Brazil. Further details about these case study cells are reported in Table 4.1. Simulations were performed withe the monthly version of the model for both cases in 2008, focusing on wheat cultivation.

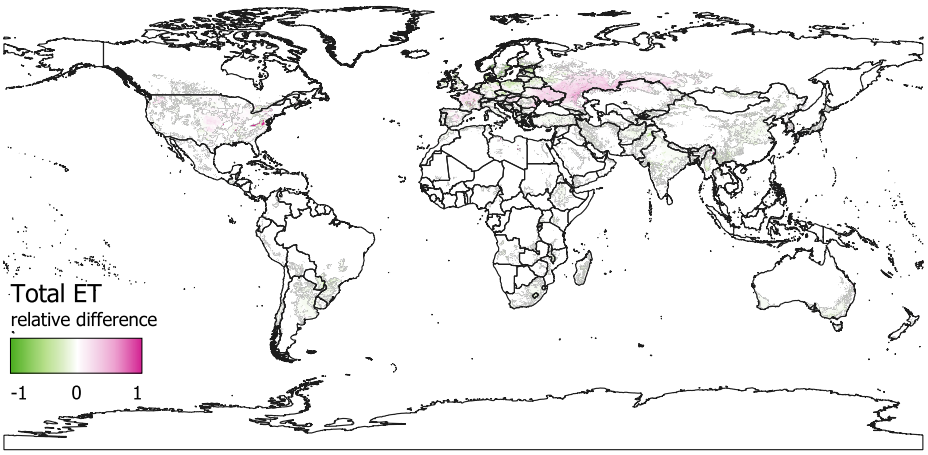
In Figures 4.6 and 4.8 (a) and (b), it can be observed that the crop coefficient dynamics are similar between the rainfed and irrigated cases, even when the two cultivations do not occur in the same period of the year, as in the Piedmont case, but instead span different lengths of the growing period. However, the coefficients



(a) Green *ET* relative difference.



(b) Blue *ET* relative difference.



(c) Total *ET* relative difference.

Figure 4.5: Relative difference between the annual *ET* obtained with the daily and monthly models. Results are based on wheat cultivations averaged over 2008-2017.

Table 4.1: Characteristics of cells selected as case study.

Data	**Piedmont**	**Minas Gerais**
Coordinates (WGS84)	44°45'36.0"N 7°30'36.0"E	16°10'00.1"S 46°34'59.9"W
Climatic region	Temperate (Oceanic)	Tropic
$AWC[mm/m]$	60	150
Rainfed Harvested Area $[ha]$	11.1	37.3
Irrigated Harvested Area $[ha]$	18.5	1.3
Rainfed Sowing Date	15 Oct	15 May
Irrigated Sowing Date	15 Apr	15 May
Rainfed $LGP[days]$	270	150
Irrigated $LGP[days]$	150	150

do not follow the same progression across the two locations. For example, the case studies exhibit different initial k_c values and durations of the initial stage, reflecting the distinct climatic regions to which they belong.

Figures 4.6 and 4.8 (c) and (d) show that only a small portion of the total rooting depth is actually relevant for plant evapotranspiration. Although maximum rooting depth is globally defined by [24], the RAW and TAW values also depend on the AWC, which explains why the maximum values reached in the two locations differ.

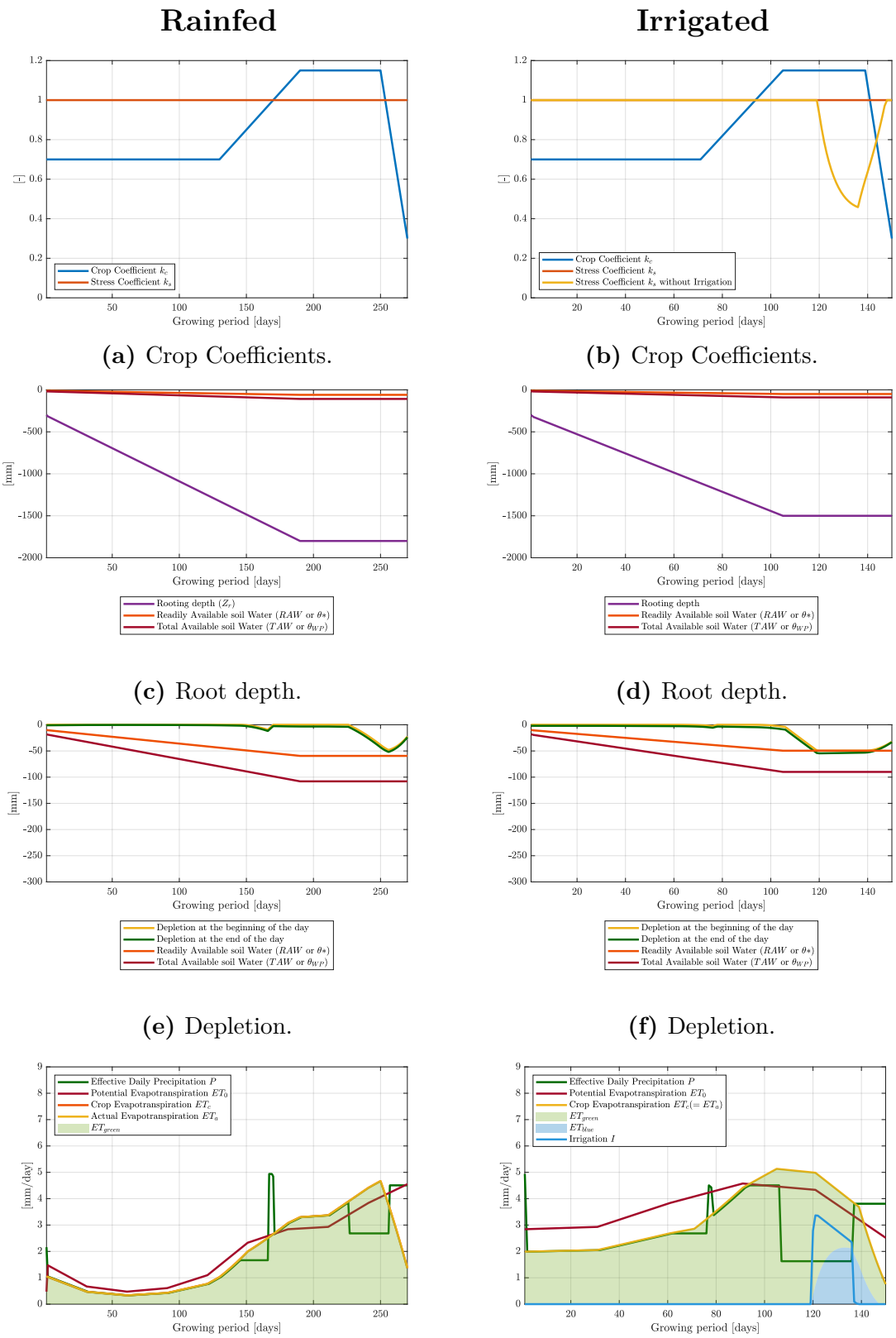
The depletion evolution, shown in Figures 4.6 and 4.8 (e) and (f), is closely tied to the local climate inputs. In Piedmont, the rainfed growing season coincides with the rainy period, preventing water stress, whereas in Minas Gerais wheat cultivation under rainfed conditions experiences water stress.

The atmospheric hydrological balance (Figures 4.6 and 4.8 (g) and (h)) reflects the processes occurring in the root zone, highlighting the partitioning of evapotranspiration between *green* and *blue* water.

Finally, Figures 4.7 and 4.9 illustrate the alternation and varying lengths of rainfed and irrigated growing seasons, as well as the additional water required to avoid stress conditions in the two different regions.

4.1.4 Other general updates

In the new version of the model it is possible to directly add the latitude and longitude of the area of interest as inputs, instead of the rows and columns of the grid. The code directly source from the input files the dimension of the grid, determining the spatial resolution of the results without any input from the user. It means that the code could run also with other resolution if all the input files are consistent in the dimensions. Finally a few lines were changed to make the code faster, shorter and more efficient.



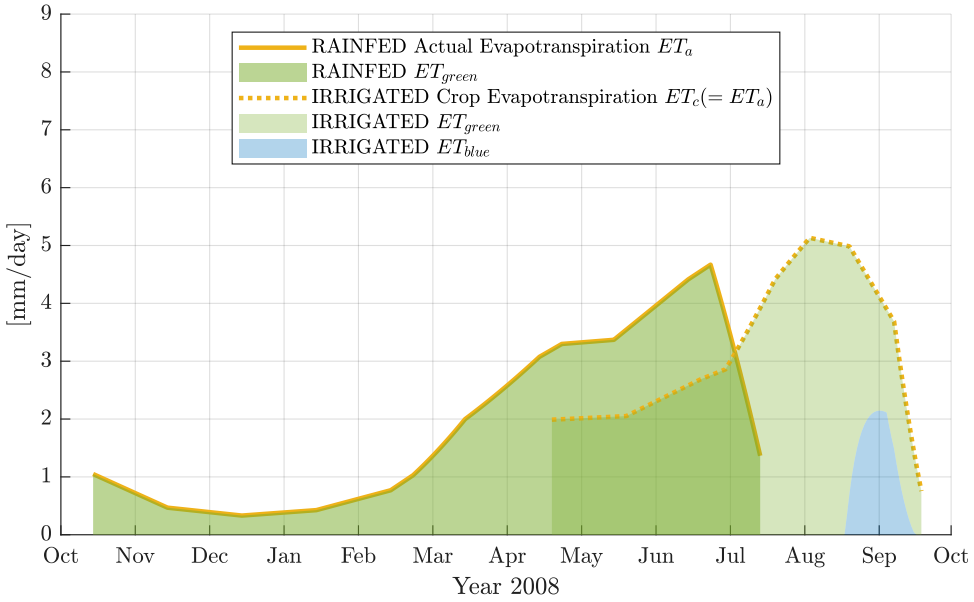
(a) Crop Coefficients. **(b)** Crop Coefficients.

(c) Root depth. **(d)** Root depth.

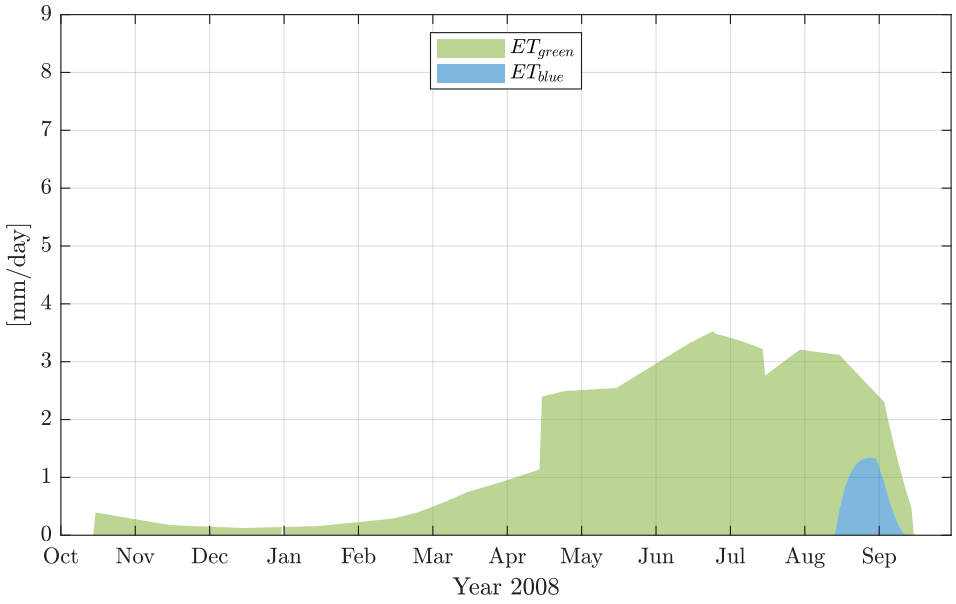
(e) Depletion. **(f)** Depletion.

(g) Hydrological balance in the atmosphere. **(h)** Hydrological balance in the atmosphere.

Figure 4.6: Cell Insights - Wheat Piedmont.

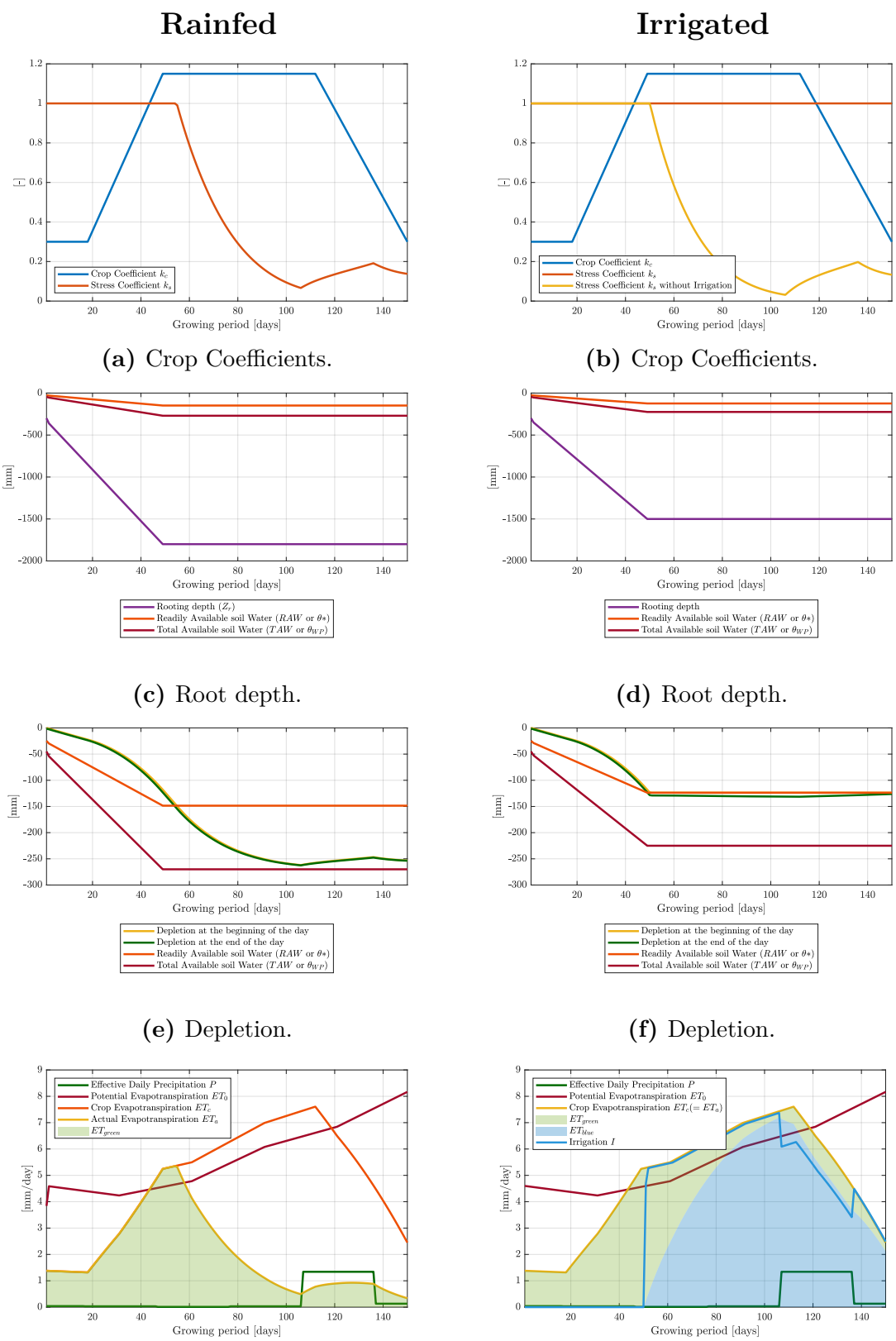


(a) Daily ET_a.



(b) Daily ET_a normalized with cultivated areas.

Figure 4.7: Cell Insights over an entire hydrological year - Wheat Piedmont.



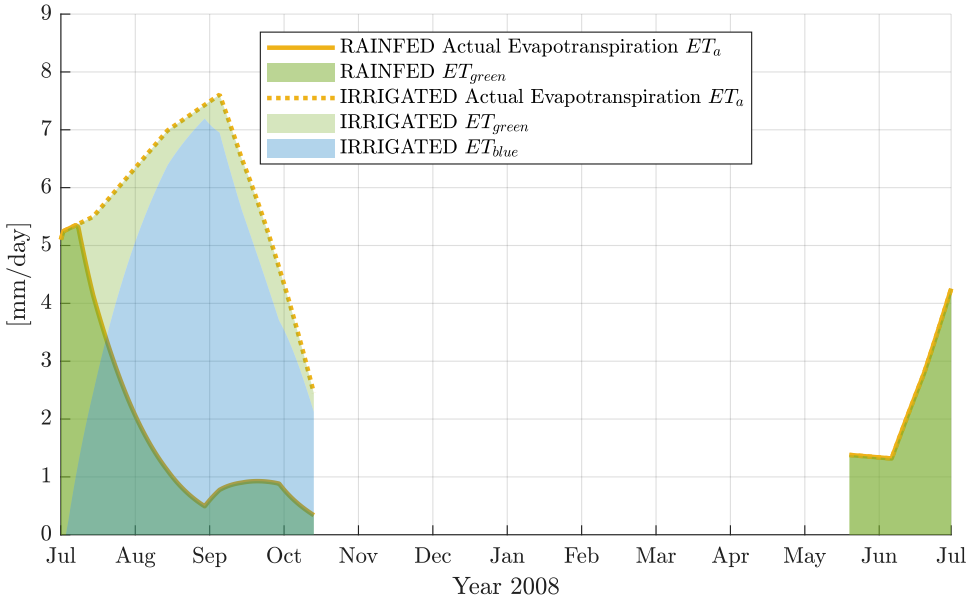
(a) Crop Coefficients. **(b)** Crop Coefficients.

(c) Root depth. **(d)** Root depth.

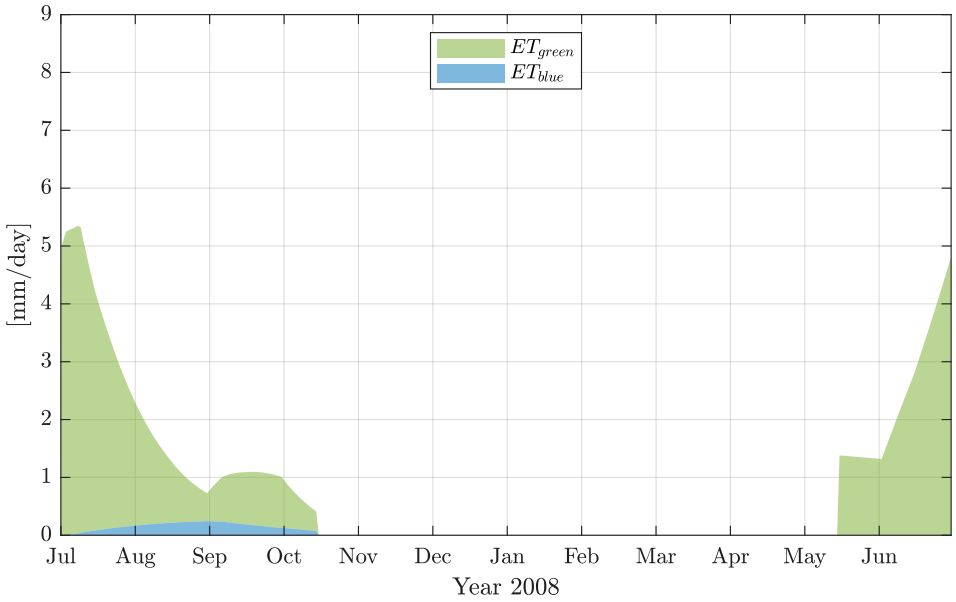
(e) Depletion. **(f)** Depletion.

(g) Hydrological balance in the atmosphere. **(h)** Hydrological balance in the atmosphere.

Figure 4.8: Cell Insights - Wheat Minas Gerais.



(a) Daily ET_a.



(b) Daily ET_a normalized with cultivated areas.

Figure 4.9: Cell Insights over an entire hydrological year - Wheat Minas Gerais.

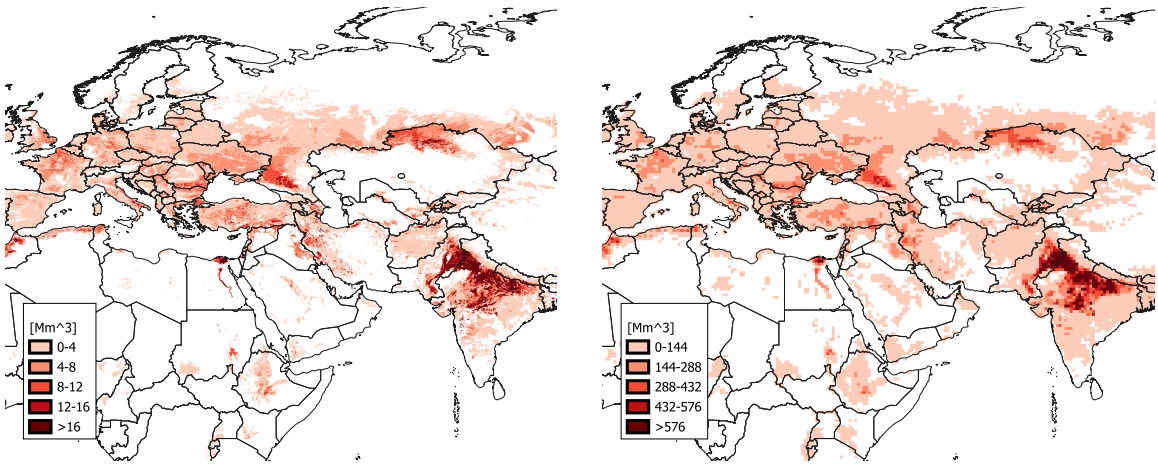
4.1.5 Output volumes in NetCDF format

To facilitate the combination with the RECON dataset, it is more convenient to have the results in term of volume instead of height per unit area. To convert the evapotranspiration the following formula was applied.

$$V_a[m^3] = ET_a \left[\frac{mm}{m^2}\right] \cdot A[ha] \cdot 10 \left[\frac{m^2}{mm \cdot ha}\right]$$

In the following steps, the obtained results need to be processed using CDO and python, for this reason the outputs of the waterCROP model must be saved in a NetCDF format. A Matlab function A.1 was written to achieve this goal. The function allow to choose the variable that must be saved, the name of the NetCDF file, the missing value, the cell size and the unit. By default the reference system is WGS84 with latitude from -90° to 90° and longitude from -180° to 180°.

4.1.6 Results upscaling process



(a) 5 arc-minutes resolution. **(b)** 0.5 deg resolution.

Figure 4.10: Visual difference between the waterCROP results (a) and the upscaled version (b).

The waterCROP model provides outputs at a finer spatial resolution compared to the datasets used later in this work. To combine the waterCROP results with the RECON dataset, an upscaling procedure is required. Since the outputs are expressed as volumes (cubic meters of water) per grid cell, the upscaling is performed by aggregating these volumes into larger cells. In line with the processing of harvested areas, the CDO function `gridboxsum` was applied, which generates each output cell by summing the cubic meters of water from the corresponding 6×6 input cells.

Figure 4.10 illustrates the difference between the original waterCROP outputs (a) and the post-processed upscaled version (b). Although the coarser resolution implies a loss of spatial detail, it ensures conservation of the total volume of water evapotranspired by crops (see details in Table 5.4). Because the data represent volumes

obtained by summation, a consistent visual comparison between the two maps requires scaling. Therefore, the colorbar of panel (b) is set to correspond to the colorbar of panel (a) multiplied by 36, which is the number of original cells aggregated into one upscaled cell.

4.2 Crop-specific evaporation sheds to map water sinks

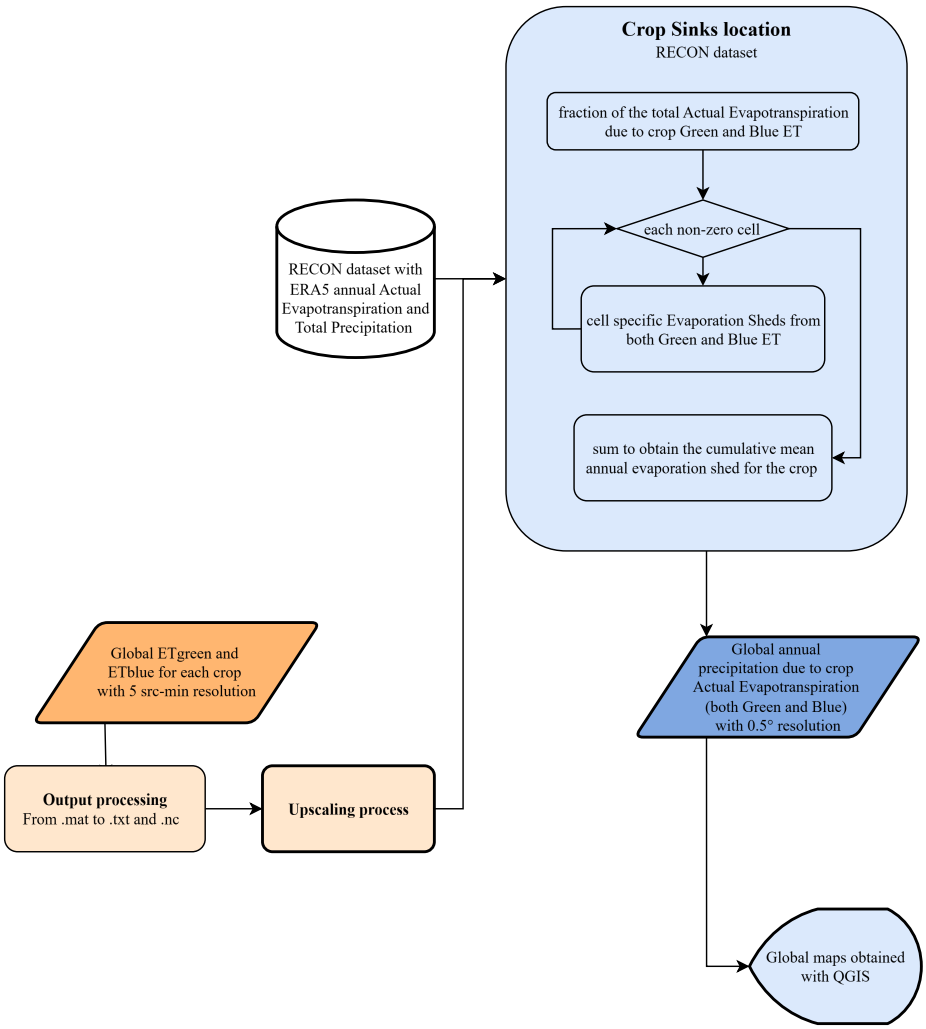


Figure 4.11: Flowchart describing the methods to obtain crop-specific evapotranspiration sheds adopting the RECON dataset.

The objective of this part of the thesis is to combine crop water use estimates with the outputs of moisture-tracking models, in order to disentangle the associated evapotranspiration (ET) sheds and identify the areas that receive the water evaporated or transpired by crops. The aim is to locate these sinks and then classify their land use, as will be explained in Section 4.3.

By adopting the RECON database, each cultivated pixel can be associated with its precipitation shed (i.e., the pixels where the rainfall infiltrating into the soil is expected to originate) and its evapotranspiration shed (i.e., the pixels where the crop evapotranspiration is expected to fall as precipitation). This approach opens new possibilities for quantifying the upstream–downstream implications of land-use changes (such as deforestation, crop switching, or conversion from rainfed to irrigated cultivation) on atmospheric water dynamics.
This thesis focuses on developing a methodology to derive the evapotranspiration shed of each cultivated pixel, quantifying the volumes of precipitation attributable to crop evapotranspiration, and then identify the types of land ultimately sustained by this recycled water.

4.2.1 Sink maps realization

To obtain cumulative sheds that allow the identification of crop-specific water sinks, a Python workflow was developed to process and analyse the outputs of the waterCROP model in combination with the RECON dataset. The code can be applied either to a selected region or at the global scale.

The first step is to harmonize the longitude coordinates across datasets. Both ERA5 and RECON use a longitude convention from 0° to 360°, meaning that the upper-left corner of the grid corresponds to latitude +90°, longitude 0°. Conversely, the waterCROP outputs (see Section 4.1.5) use a longitude convention from −180° to 180°, where the upper-left corner corresponds to latitude +90°, longitude −180°. To ensure consistency, the two vertical halves of the waterCROP matrices were swapped.

Since the RECON dataset provides the moisture flow associated with the entire vegetated surface of each grid cell, the calculation must isolate only the portion attributable to the specific cultivated crop under analysis. For this reason, the first part of the code computes a fraction matrix of actual evapotranspiration (ET_a) in volumetric terms, using data from waterCROP. This fraction scales the RECON moisture flow to the crop level:

$$ET_{\text{fraction},i}\ [-] = \frac{ET_{a,i}\ [m^3]}{ET_{\text{ERA5},i}\ [m^3]}$$

Two separate loops (for green and blue ET) iterate over the non-zero elements of the fraction matrices. For each grid cell, the code extracts the corresponding shed from RECON and accumulates it to build the global cumulative sheds. During this process, the code computes the geographic coordinates of the sinks, retrieves the moisture flow values (expressed as integer units in RECON), and applies the conversion to volumetric units according to De Petrillo et al. (2025) [31].

Because the calculations involve global-scale summations of small fractions of volumes per each cell, extended precision is required. The code uses `float128` arithmetic to minimize rounding errors and ensure mass conservation.

To illustrate the procedure, evaporation sheds were also computed for a single grid cell considering wheat cultivations (Figure 4.12). It is important to note that the Blue shed is plotted on a different scale to enhance visibility.

Then, the cumulative evaporation sheds for both green and blue ET were obtained by summing across all cultivated cells. These cumulative sheds provide the total ET shed for subsequent analyses. The results are saved as new NetCDF files, ensuring reproducibility and proper metadata storage.

4.2.2 Statistical relationship between crop evapotranspiration and sheds

A statistical comparison was carried out between crop evapotranspiration estimated with the waterCROP model and the corresponding crop-specific evapotranspiration shed. The objective is to assess how water redistributes itself as precipitation after being transferred to the atmosphere. As illustrated in Figure 4.12, evapotranspiration from a single grid cell is dispersed over a wider region, but a statistical comparison was necessary to evaluate whether this effect is balanced when all sheds are aggregated.

All analyses were performed by comparing values expressed in mm/m^2. The volumes from both outputs were converted into evapotranspiration depths by distributing the volume over the cell area. This conversion was carried out using the inverse of the equation presented in Section 4.1.5.

The scatter density plots in Figure 4.13 show that waterCROP output values span a wider range of evapotranspiration depths, while the evapotranspiration shed is characterized by generally lower values, with a strong concentration near zero. It is also evident that the two variables exhibit only a weak correlation, with coefficients ranging from 0.29 to 0.40. In other words, atmospheric processes largely dissolve the direct relationship between evapotranspiration in a given cell and the precipitation that can be attributed back to it. Following this analysis, we expect to observe in the Results Section (5) that high local evapotranspiration from crops will not be fully compensated by the precipitation returning to land. Instead, values are smoothed: lower in magnitude but distributed over a broader area, consistent with the behaviour observed for individual sheds.

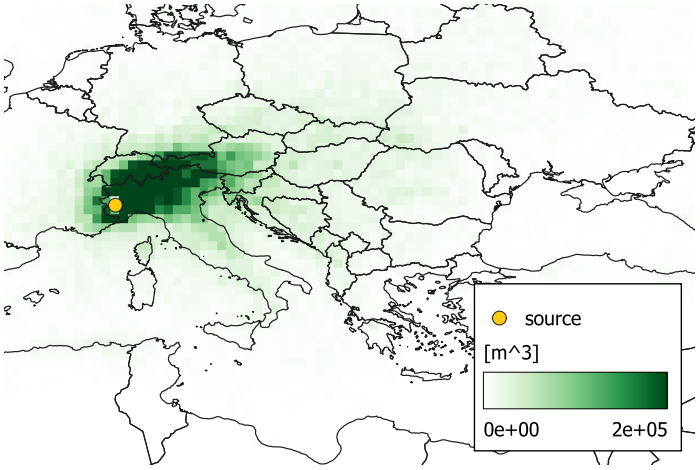
4.2.3 Post-processing

Further analyses were conducted on the crop-specific evapotranspiration sheds, combining these results with annual total precipitation volumes and crop evapotranspiration volumes previously assessed. The aim was to evaluate agriculture's contribution to overall precipitation and its role in recycling water within the same crop cultivations.

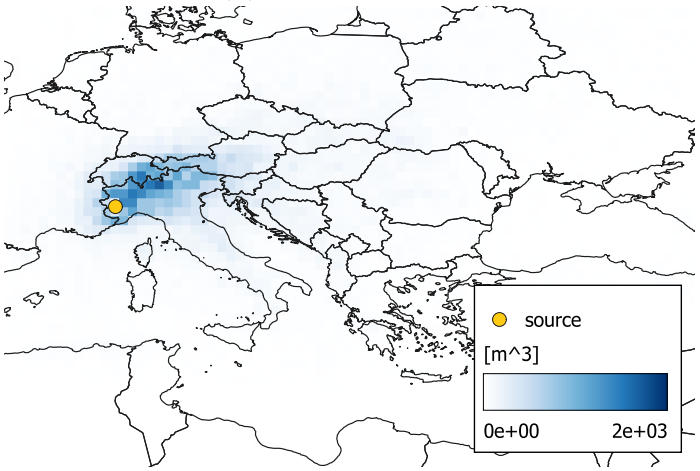
To better capture the relative importance of agriculture in the local water balance, a series of maps was produced showing, for each grid cell, the fraction of annual precipitation attributable to a specific crop:

$$P_{fraction,i}[-] = \frac{ET_{shed,i}[m^3]}{P_{tot,i}[m^3]}$$

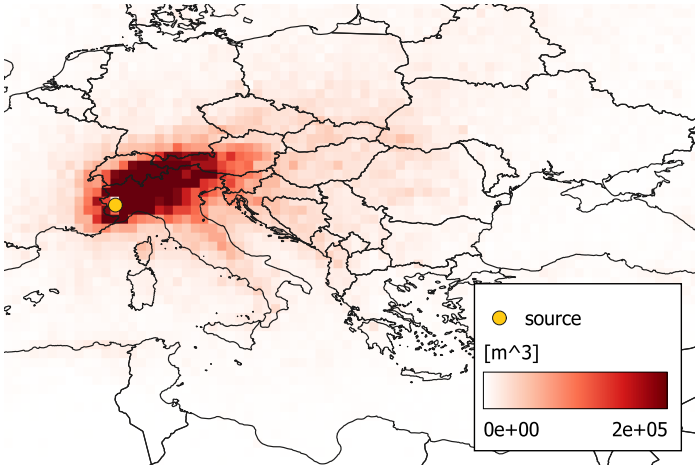
where $ET_{shed,i}$ is the precipitation in cell i attributable to evapotranspiration from the crop, and $P_{tot,i}$ is the total annual precipitation in that cell obtained from ERA5 data [42].



(a) Green *ET* shed.

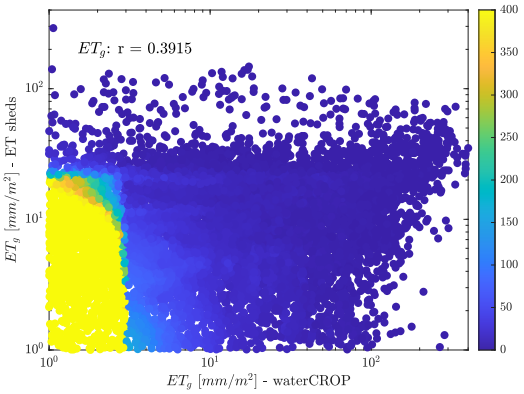


(b) Blue *ET* shed.

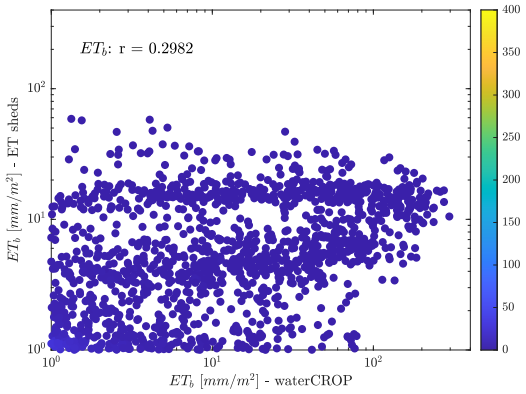


(c) Total *ET* shed.

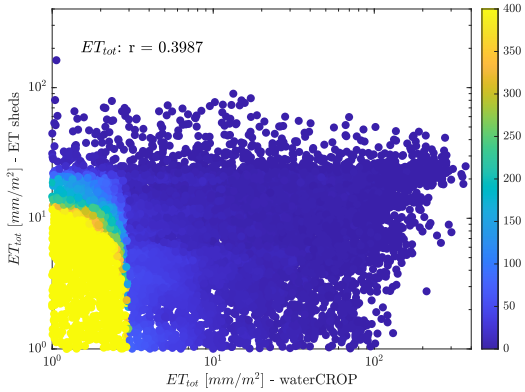
Figure 4.12: Crop-specific *ET* sheds of a single grid cell.



(a) Green ET scatter density plot.



(b) Blue ET scatter density plot.



(c) Total ET scatter density plot.

Figure 4.13: Density scatter plot comparing crop evapotranspiration estimates with crop-specific evapotranspiration sheds. The color bar represents point density, and both axes are shown on a logarithmic scale.

To assess whether a crop in a given cell contributes more water to the atmosphere than it receives back from its own global evapotranspiration volume, two metrics were defined: the annual Evapotranspiration Difference (ET_{diff}) and the Crop Recycling Ratio (CRR):

$$ET_{diff,i} = ET_{waterCROP,i} - ET_{shed,i} \cdot \frac{A_{crop}}{A_{cell}}$$

$$CRR_i = \begin{cases} \frac{ET_{waterCROP,i} - ET_{shed,i} \cdot \frac{A_{crop}}{A_{cell}}}{ET_{waterCROP,i}}, & ET_{waterCROP,i} > 0 \\ -1, & ET_{waterCROP,i} \leq 0 \end{cases}$$

where all ET terms are expressed in m^3, and areas are in hectares. The shed contribution is scaled by the fraction of the cell covered by the crop, ensuring that only the water returning to the same cultivated area is considered.

Finally, to verify mass conservation, global *ET* volumes derived from the sheds were compared against the total *ET* inputs. Since the global annual water cycle is closed, any mismatch quantifies the residual error in the workflow.

4.3 Land use classification of crop water sinks

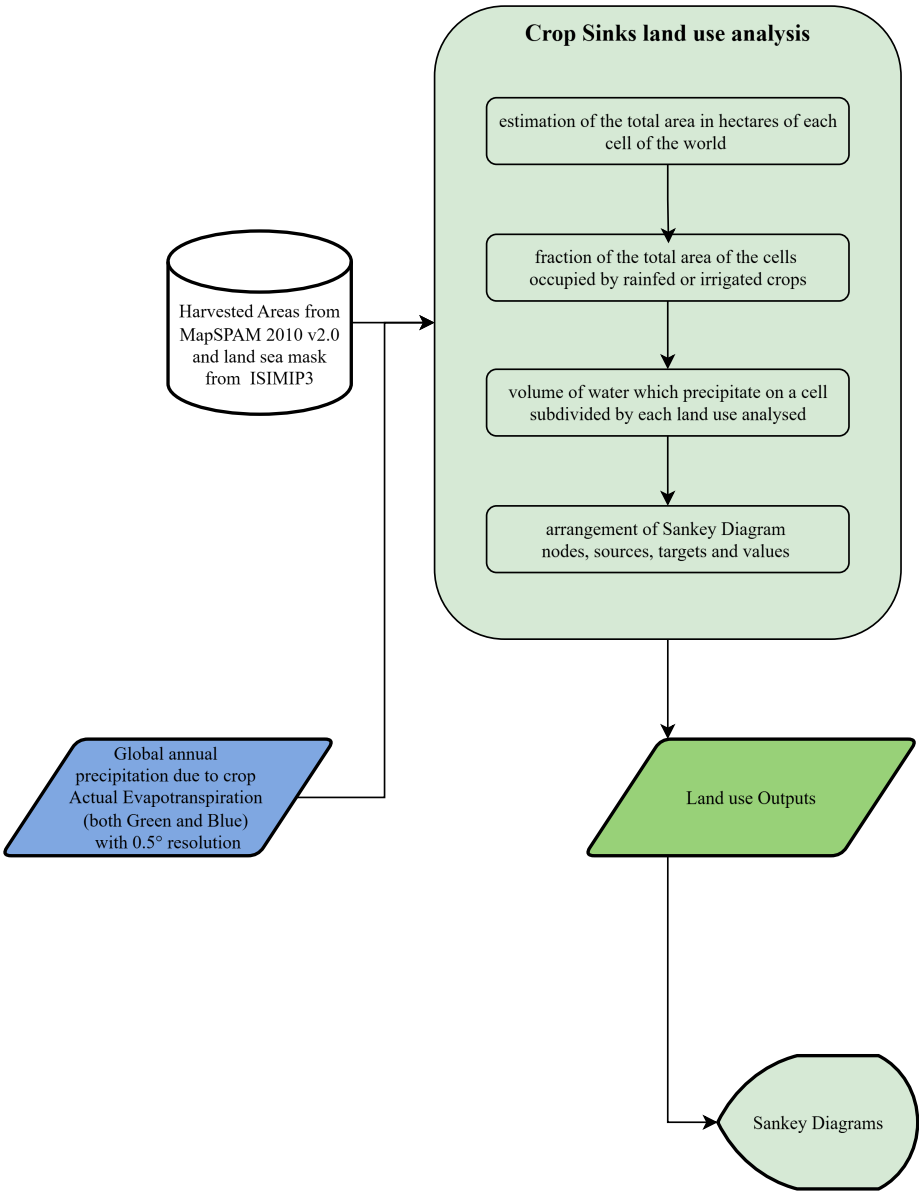


Figure 4.14: Flowchart describing the methods to obtain the land use classification of crop water sinks.

The land-use analysis of crop-specific sinks was performed with a dedicated Python code. Its purpose is to quantify how *ET* fluxes from each crop are redistributed across

different land types, namely other agricultural areas, non-agricultural land, and the ocean, and to represent these fluxes in the form of a Sankey diagram.

In the present implementation, the analysis focuses on agricultural areas of the three crops modelled with waterCROP. However, the method can be expanded in the future to classify all land-use types of the sinks by integrating more detailed land-use datasets, provided they remain consistent with the harvested areas used in waterCROP. As described in Chapter 3, a land–sea mask was applied to separate terrestrial surfaces from the ocean.

The first step in determining how much water precipitates on a specific land type is to compute the fractional coverage of each land type within every grid cell. This requires the calculation of the total cell area, approximating Earth as a sphere of radius $R = 6371\,\text{km}$. The surface area of a latitude–longitude grid cell is:

$$A_{cell,i} = R^2\,\Delta lon\,[\sin(lat_n) - \sin(lat_{n+1})]$$

where Δlon is the longitudinal resolution of the matrix in radians and lat_n, lat_{n+1} are the latitudinal bounds of the i-th grid cell. The result, computed in m^2, is converted into hectares to match the units of harvested area data.

The land-use fractions are then calculated as:

$$f_i[-] = \frac{A_{landtype,i}[ha]}{A_{cell}[ha]}$$

For each grid cell, the water volume that precipitates is multiplied by these fractions, thereby allocating the moisture fluxes to the different land uses. This procedure also allows the generation of maps showing the spatial distribution of volumes by land type.

Finally, the water fluxes are aggregated globally, yielding totals for green and blue ET precipitating on rainfed cropland, irrigated cropland, the ocean and other terrestrial surfaces.

These totals, expressed in m^3, form the basis of the Sankey diagram. In this diagram, nodes correspond to crops, land-use categories, and ET types, while links represent the volumetric flows between them. The output is both visual (the Sankey plot) and numerical, with the values stored in CSV files for further use.

Since the results of the Sankey diagram depend on the input evaporation sheds, if the sheds are regional rather than global, the land-use analysis is carried out for the same specific area.

Chapter 5

Results

This chapter presents the key findings of the thesis, highlighting both spatial patterns and hydrological implications of crop-specific evapotranspiration. The analysis begins with high-resolution maps of evapotranspiration from maize, wheat, and soybean, revealing how water use varies across regions and crops. These patterns provide the foundation for understanding how cultivated areas influence atmospheric moisture flows.

Next, crop-specific evapotranspiration sheds illustrate how water evaporated from cultivated areas is transported and eventually precipitates downwind. By comparing the local evapotranspiration from a given crop with the corresponding cumulative precipitation across the globe, it is possible to evaluate whether the annual water balance is maintained at the scale of individual cells. Maps of the relative differences between waterCROP outputs and the cumulative sheds identify regions where evapotranspiration either exceeds or falls short of the water ultimately returned as precipitation, offering insight into local versus teleconnected water dynamics.

The analysis then explores the land-use composition of the sink areas, providing a perspective on which landscapes benefit most from crop-generated moisture and how agricultural expansion or land management practices might influence these flows. Finally, the chapter assesses the effects of upscaling from high-resolution waterCROP outputs to coarser grids on the total annual water volume estimates, highlighting where aggregation may lead to under- or overestimation and discussing the implications for interpreting regional and global water budgets.

Overall, this chapter integrates spatial, hydrological, and land-use perspectives to provide a comprehensive view of how crop cultivation shapes the movement and redistribution of water in the atmosphere.

5.1 Crop Green and Blue Evapotranspiration evaluation with the waterCROP model

The waterCROP model estimates crop evapotranspiration (ET) over the entire growing season, both for rainfed and irrigated harvested areas. In previous studies, ET estimates have often been combined with crop yield data to assess the virtual water content of agricultural products [28]. In this thesis, however, the model is employed to quantify crop-specific ET at the global scale in volumetric terms.

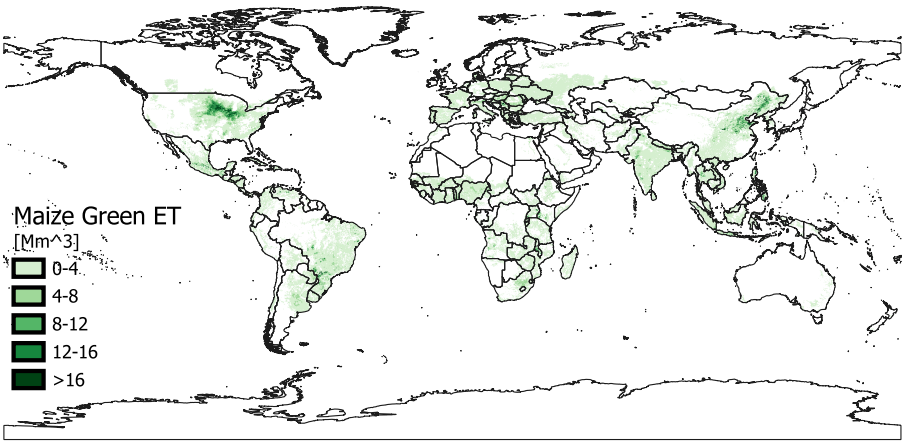
The simulations were carried out worldwide using ERA5 climatic data for the period 2008–2017, focusing on wheat, maize, and soybean. Figures 5.1 to 5.3 present the resulting annual ET volumes, expressed in millions of cubic meters, showing how much water enters the atmosphere through crop evapotranspiration. Because these volumes are computed by multiplying evapotranspiration depth (in mm) by the harvested area (as described in Section 4.1.5), the spatial patterns are strongly shaped by the distribution of cultivated land in each grid cell.

Figure 5.1 highlights the widespread extent of maize cultivation at the global level. The largest ET volumes are estimated in the central United States, southern Brazil, the Po Valley in Italy, the Nile River basin, and northern China.

In general, Blue ET values are considerably lower than Green ET, which is why Total ET tends to mirror the spatial behaviour of Green ET. Nonetheless, exceptions emerge where irrigation plays a major role. For instance, in the case of wheat (Figure 5.2), Green ET (a) is dominant in northern India, while Blue ET (b) is more pronounced in the western regions. Here, the two components are of comparable magnitude, so the spatial distribution of Total ET (c) clearly reflects their combined influence.

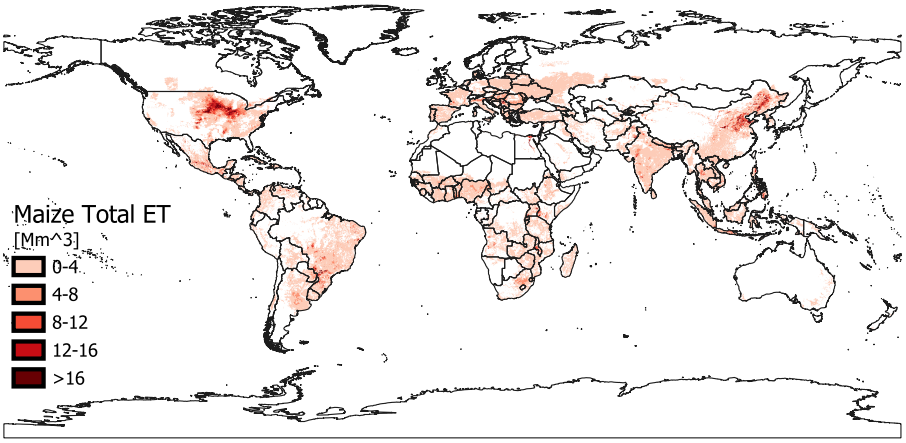
Soybean cultivation is concentrated mainly in the United States, central India, Brazil, and South America more broadly (Figure 5.3). It should be noted, however, that irrigated soybean areas are likely underestimated.

Overall, these results provide the baseline for the subsequent analysis of evapotranspiration sheds, where the spatial distribution of crop-specific ET will be linked to downwind precipitation and land-use dynamics.



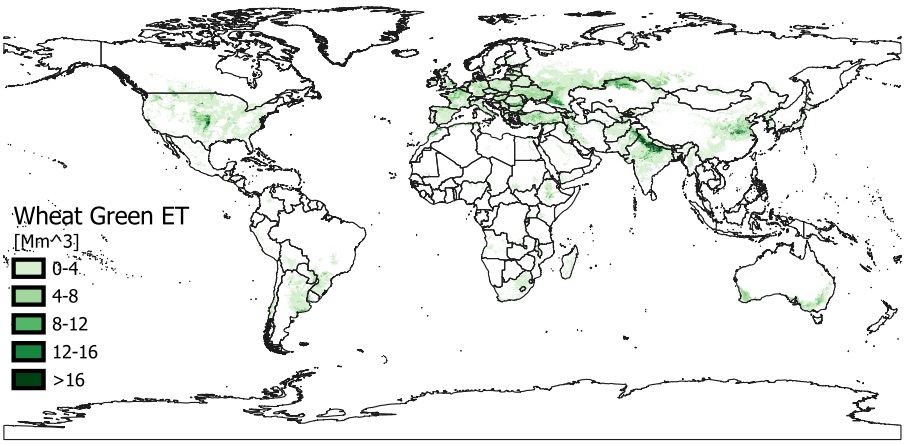
(a) Green *ET*.

(b) Blue *ET*.

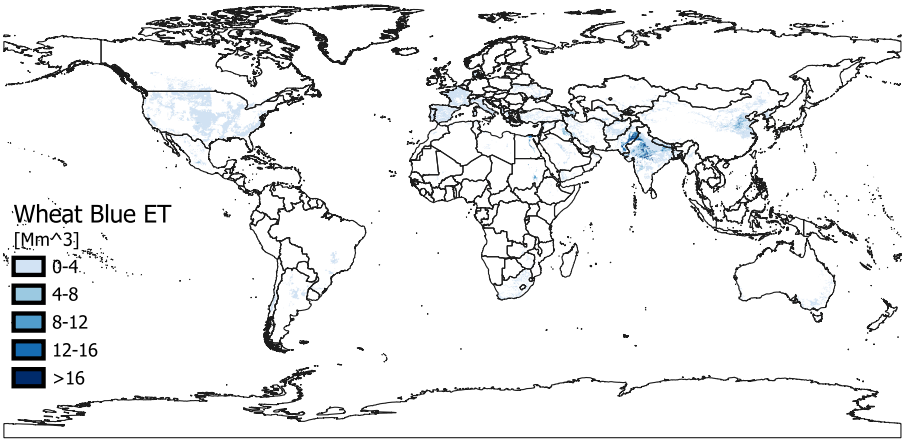


(c) Total *ET*.

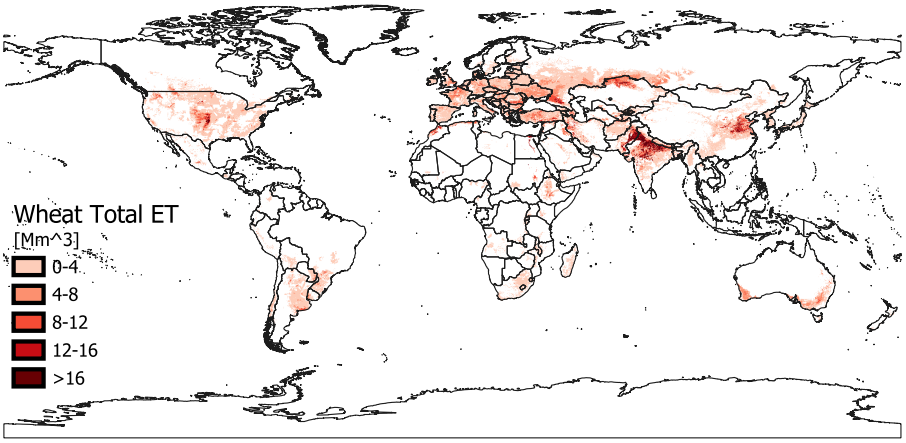
Figure 5.1: Annual evapotranspiration from Maize cultivations.



(a) Green *ET*.

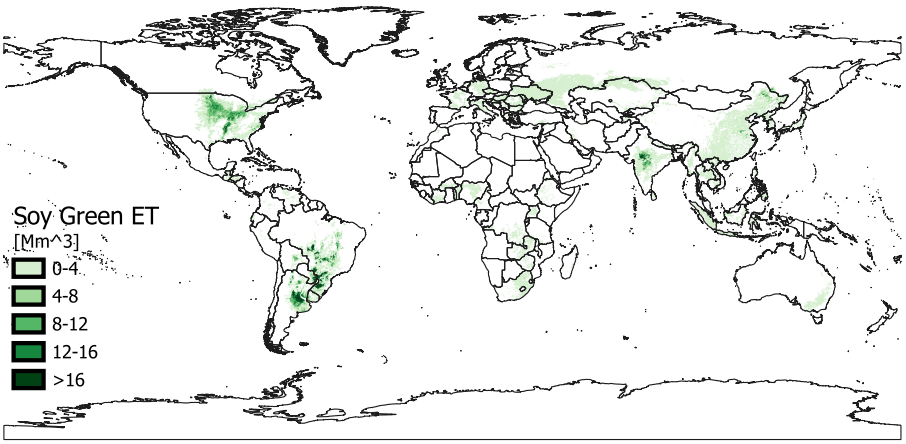


(b) Blue *ET*.



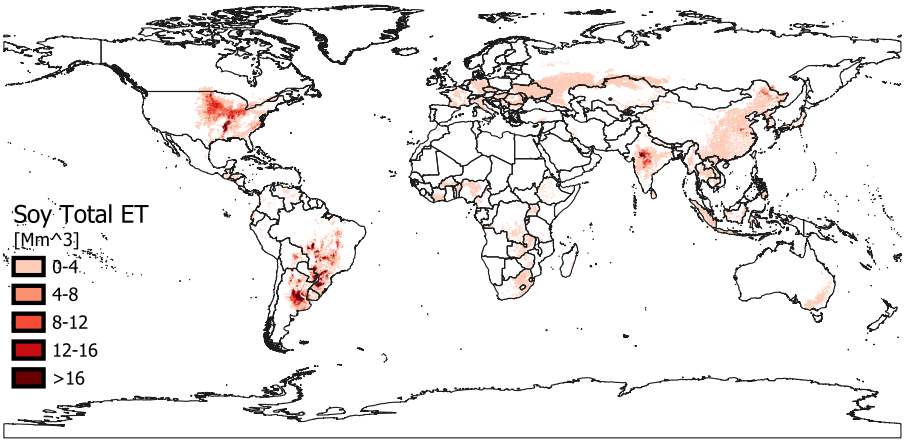
(c) Total *ET*.

Figure 5.2: Annual evapotranspiration from Wheat cultivations.



(a) Green *ET*.

(b) Blue *ET*.



(c) Total *ET*.

Figure 5.3: Annual evapotranspiration from Soy cultivations.

5.2 Crop specific cumulative Evaporation Sheds to map water sinks

It is well established in the literature that agriculture is sustained not only by direct water inputs but also by vegetation evapotranspiration, and that land-use changes can significantly influence rainfall in downwind regions [46, 34, 47]. The innovative contribution of this work lies in the construction of crop-specific evapotranspiration sheds, obtained by proportionally summing the contributions of each harvested-area cell, as detailed in Section 4.2.

Figures 5.4 to 5.6 present the resulting sheds for maize, wheat, and soybean. To enhance readability, the Blue sheds are plotted on a separate scale: if shown on the same scale as the Green sheds, their much smaller values would result in poor visibility and the shed structures would effectively disappear.

Across all sheds, major mountain ranges emerge clearly. Orographic lifting and rain shadow effects strongly shape precipitation, making the Andes, the Alps, the Urals, the Himalayas, and the Taihang Mountains stand out in the maps. Beyond topographic features, atmospheric circulation patterns are also evident: for instance, cultivations in southern South America contribute to precipitation fluxes directed toward the South Atlantic Ocean.

The atmosphere also introduces a smoothing effect on evapotranspiration distributions, as already explained in Section 4.2.2. The distribution of waterCROP *ET* are sharply peaked with a steep progression, whereas those of the sheds are smoother, with lower means and higher standard deviations. This reflects how atmospheric transport redistributes water over broader areas, dampening localized extremes while maintaining overall consistency.

The sheds make evident several regionally distinct behaviours. In Europe, maize (Figure 5.4) and wheat (Figure 5.5) sheds show a marked eastward displacement relative to the waterCROP source regions (Figures 5.1, 5.2). This implies that reduced evapotranspiration from Western European crops, for example, due to droughts or land-use change, could translate into rainfall deficits in Eastern Europe. Local shocks therefore propagate into downwind regions, highlighting the interconnectedness of agro-hydrological systems.

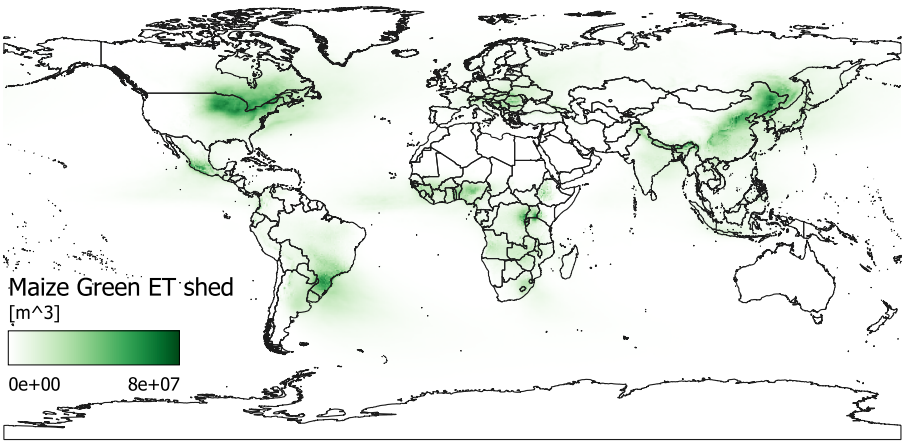
In India, although crops are grown throughout the country, much of the precipitation originating from evapotranspiration converges in the north-east and Nepal. Rivers such as the Ganges help redistribute this water back to downstream cultivations, and infiltration recharges aquifers. However, the timescales and sustainability of this recycling remain open questions.

The comparison between the Po Valley and the Nile Basin illustrates further contrasts. Both regions host dense cultivation, but while the Po Valley receives significant recycled moisture, enhanced by the Alps, the Nile Basin shows low precipitation feedback from its croplands.

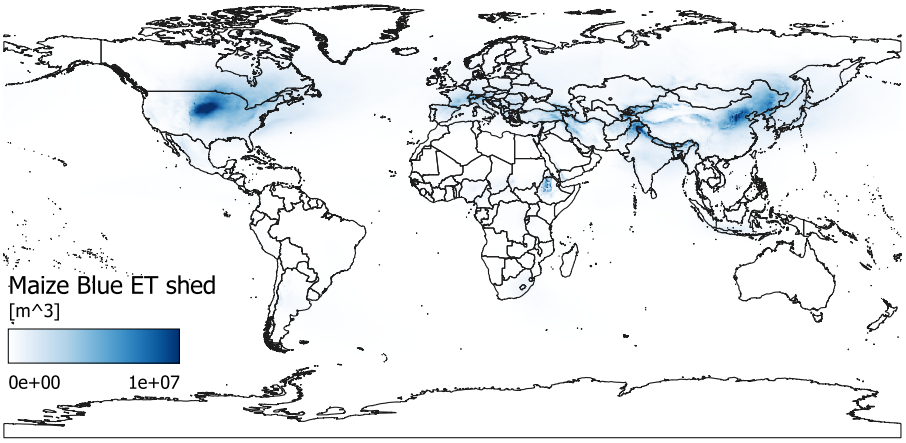
To better capture the relative importance of agriculture in the local water balance, a series of maps (Figures 5.7 to 5.9) was produced showing, for each cell, the fraction of annual precipitation attributable to a specific crop.

These maps provide a powerful diagnostic of agricultural contributions to rainfall. In some hotspots, maize contributes up to 5% of total annual precipitation: for example, 3.8% in Minnesota (USA) and nearly 5% in Shanxi and Heilongjiang (China). Irrigation plays a decisive role in these dynamics. Without irrigation, the Blue water volumes would follow other pathways; when diverted to fields, they re-enter the atmosphere at specific locations. For example, in Pakistan, wheat irrigation alone accounts for around 3.5% of total annual precipitation.

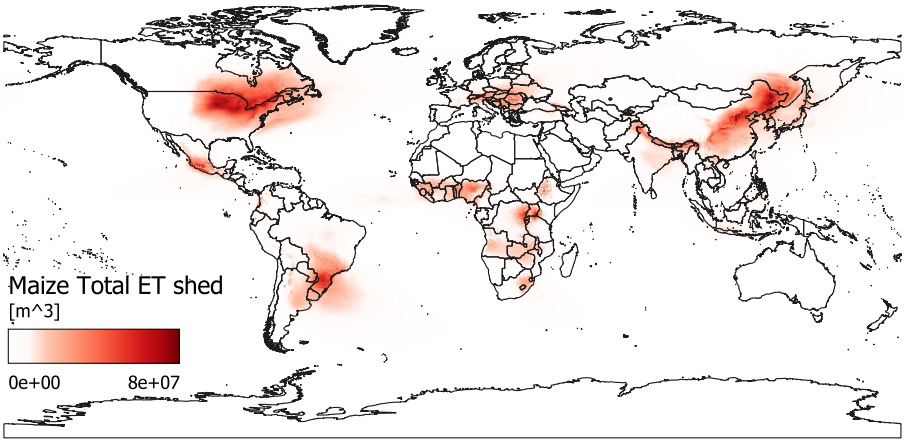
Taken together, these results demonstrate how crop-specific evapotranspiration sheds connect local cultivation practices to both local and remote rainfall, revealing critical dependencies between agricultural regions and the atmosphere. They underscore how changes in land use, irrigation, or climate can reverberate across regions, reshaping not only local but also downwind hydro-climatic systems.



(a) Green *ET* shed.

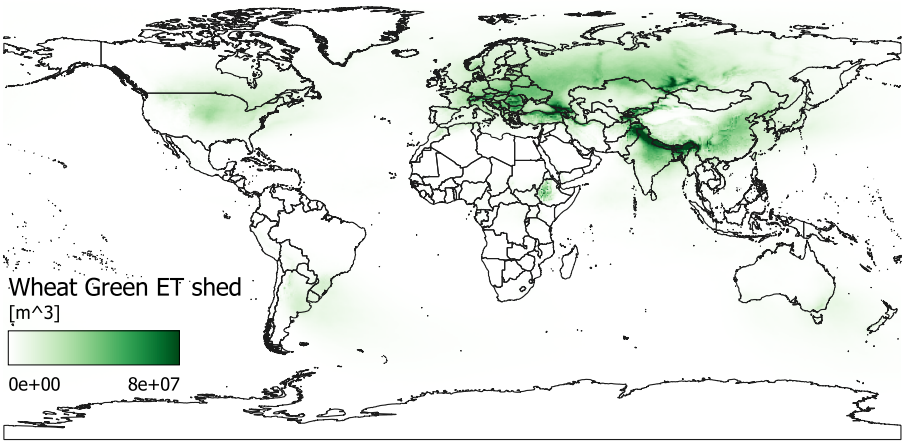


(b) Blue *ET* shed.

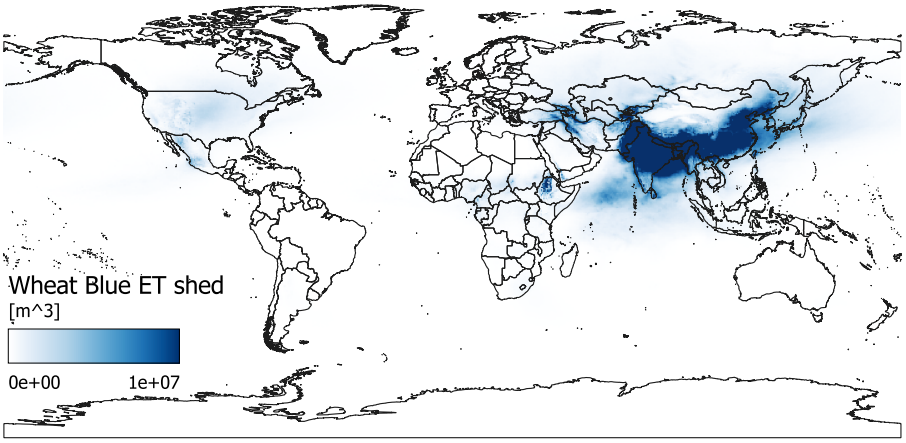


(c) Total *ET* shed.

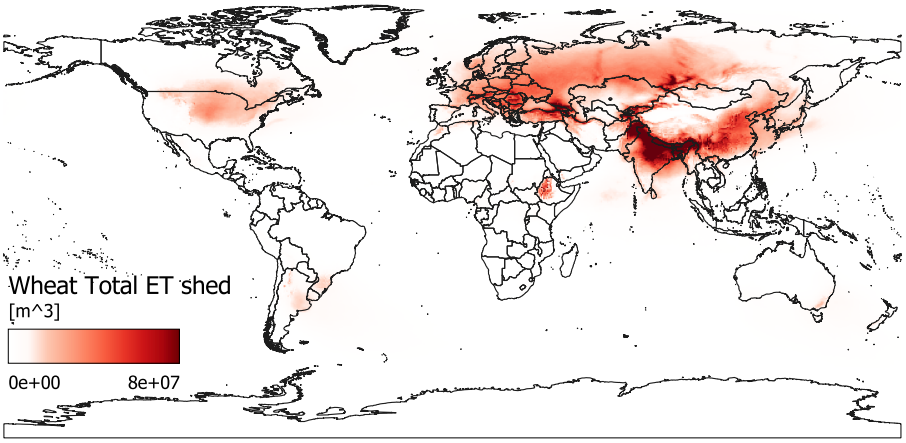
Figure 5.4: Annual evapotranspiration shed from Maize cultivations.



(a) Green *ET* shed.

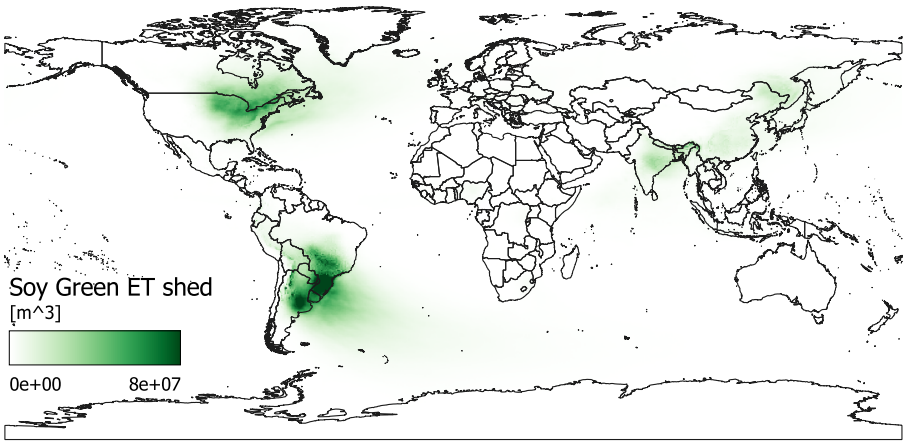


(b) Blue *ET* shed.

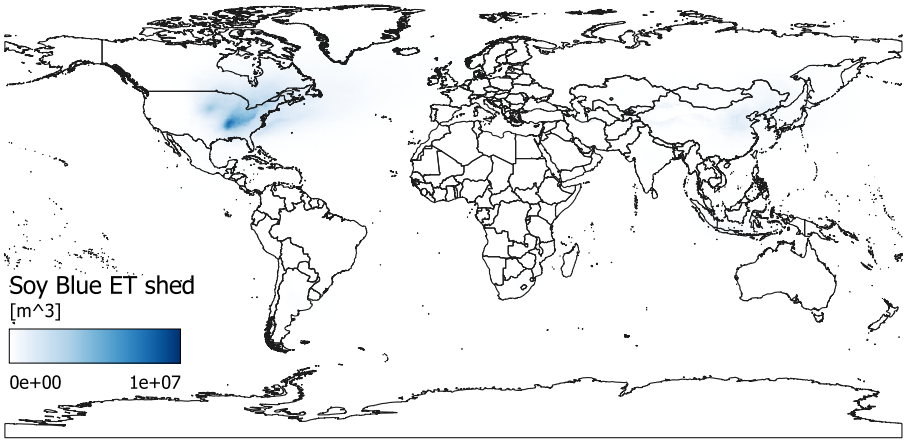


(c) Total *ET* shed.

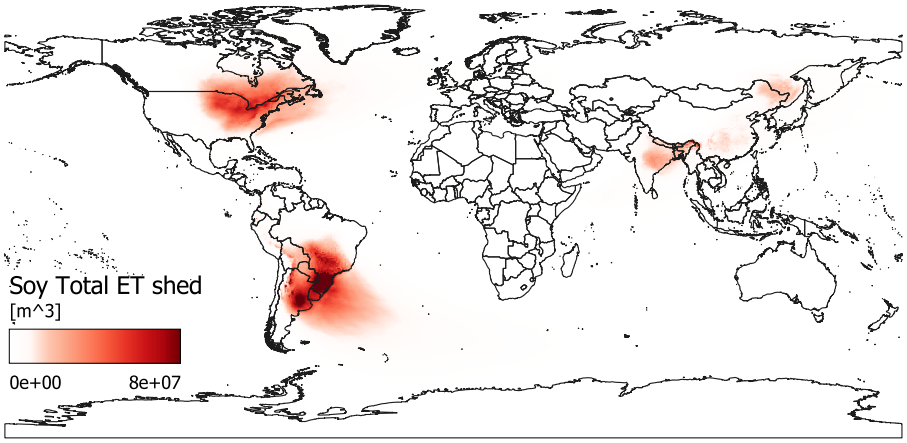
Figure 5.5: Annual evapotranspiration shed from Wheat cultivations.



(a) Green *ET* shed.

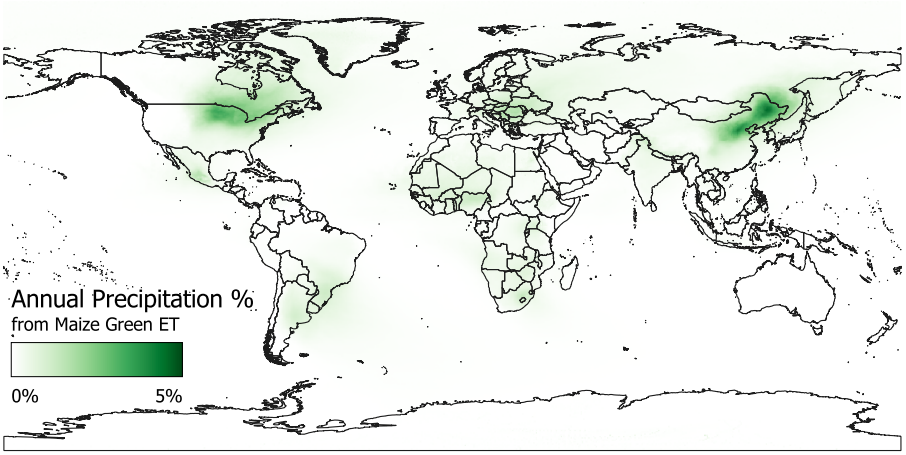


(b) Blue *ET* shed.

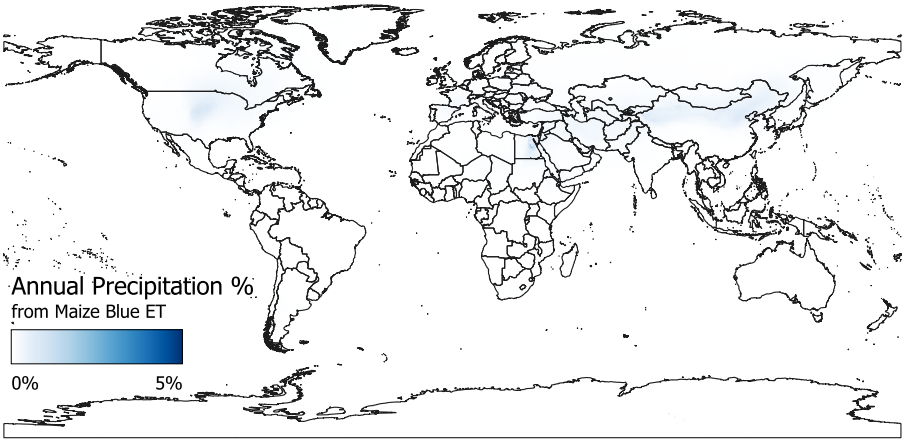


(c) Total *ET* shed.

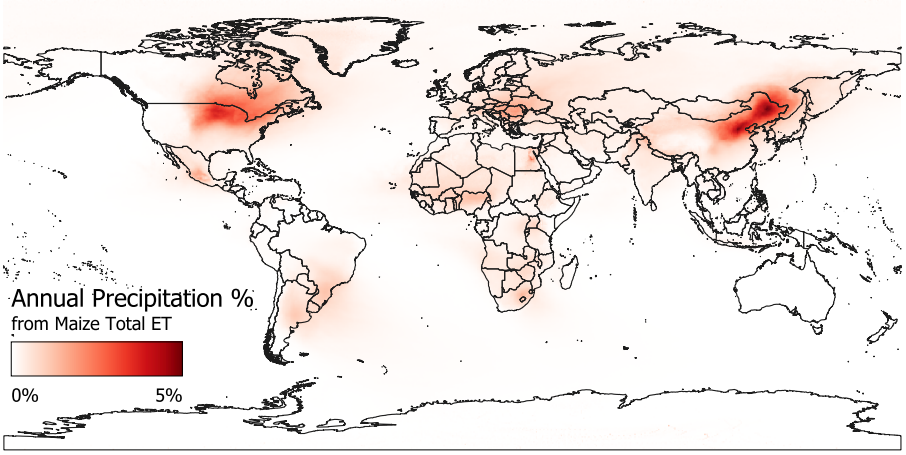
Figure 5.6: Annual evapotranspiration shed from Soy cultivations.



(a) Green *ET* contribution.

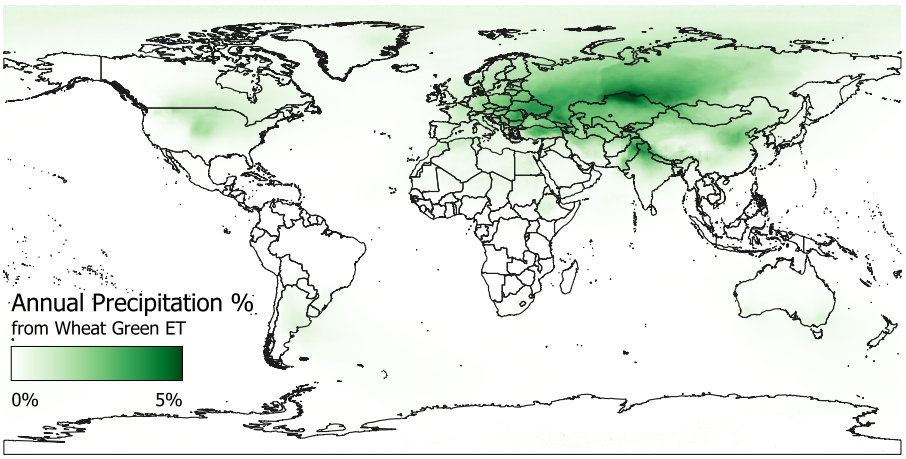


(b) Blue *ET* contribution.

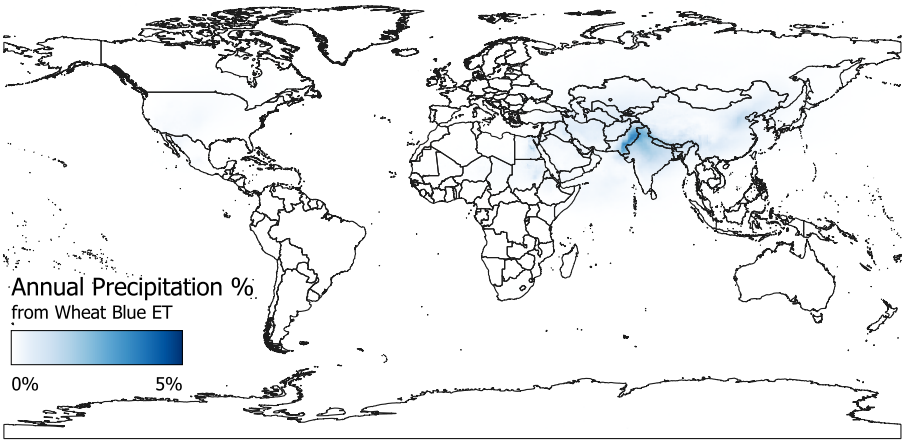


(c) Total *ET* contribution.

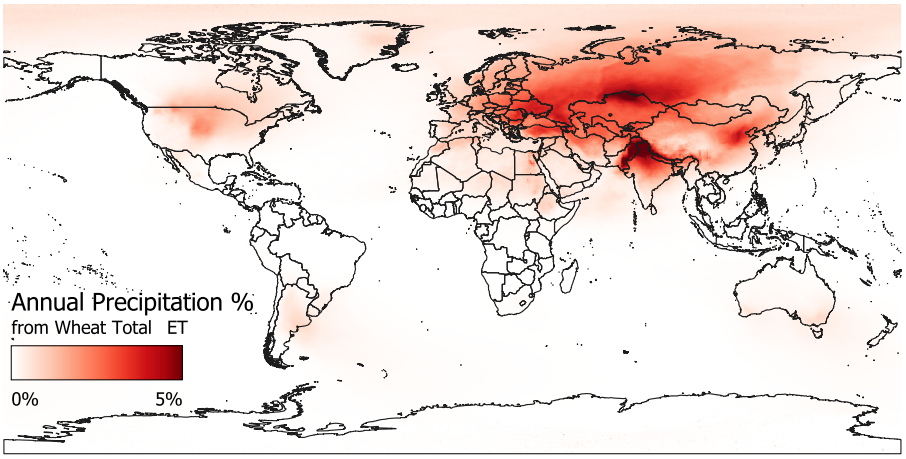
Figure 5.7: Fraction of annual precipitation originated from global Maize cultivations.



(a) Green *ET* contribution.

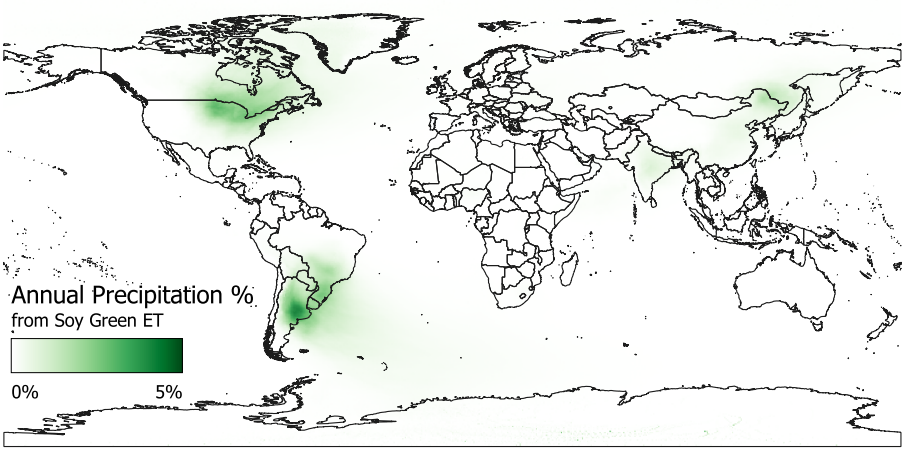


(b) Blue *ET* contribution.

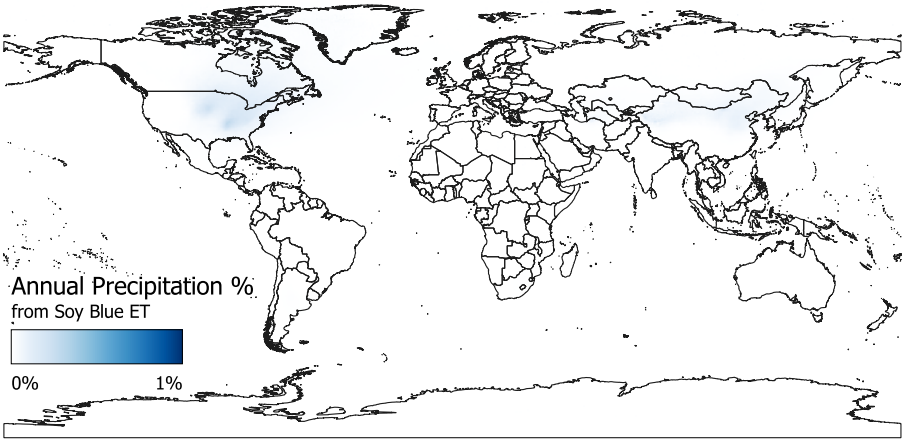


(c) Total *ET* contribution.

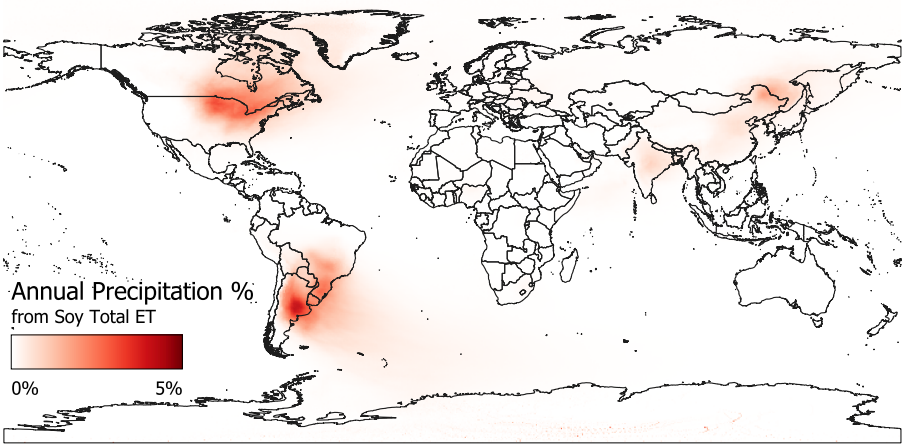
Figure 5.8: Fraction of annual precipitation originated from global Wheat cultivations.



(a) Green *ET* contribution.



(b) Blue *ET* contribution.



(c) Total *ET* contribution.

Figure 5.9: Fraction of annual precipitation originated from global Soy cultivations.

5.2.1 Annual crop water balance: donor and recipient areas

To assess whether a crop in a given cell contributes more water to the atmosphere than it receives back from its own global evapotranspiration volume, two metrics were defined in Section 4.2.3: the annual Evapotranspiration Difference (ET_{diff}) and the Crop Recycling Ratio (CRR).

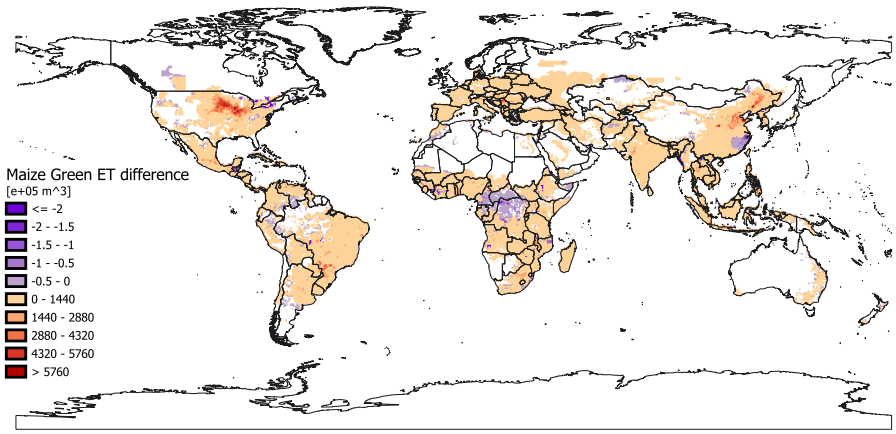
Cells with positive CRR values are classified as *donor areas*, meaning that crops in these regions export more water to the atmosphere than they regain through precipitation linked to their own evapotranspiration. Negative values identify *recipient areas*, where crops receive more water via atmospheric transport than they locally release. This effect is particularly pronounced in maps considering only Blue ET, because rainfed cultivated areas act as recipients of water originating from irrigated regions.

The scales for positive and negative values are intentionally asymmetric, reflecting the very different magnitudes of water export and import. A symmetric scale would obscure these differences and reduce the interpretability of the results.

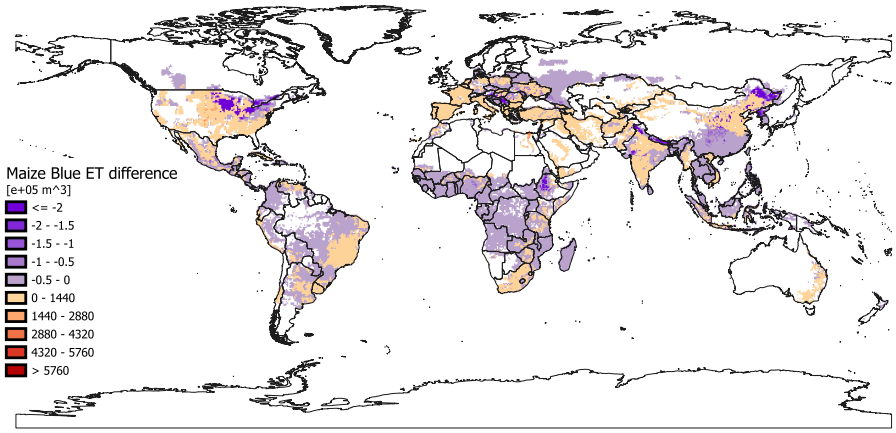
Notable examples are evident in the maps. In Figure 5.10, the Congo Rainforest benefits from precipitation originating from maize cultivations. Wheat cultivations in southern Canada are sustained by moisture from U.S. wheat fields, while wheat fields in southern China receive precipitation originating from northern China (Figure 5.12). Large rainfed regions appear as recipients of Blue ET, whereas intensive irrigation in India and Pakistan emerges as a net water export, with minimal compensation from atmospheric recycling.

CRR maps not only illustrate the donors and recipients of water within the croplands of a given crop but also indicate the extent to which a crop attenuates its own local evapotranspiration through moisture recycling. Light blue areas exhibit very low crop recycling, meaning that the contribution of global cultivations to rainfall in that cell is small relative to its evapotranspiration. Purple areas correspond to crop recycling higher than 8%, while magenta areas receive more water than they evapotranspire, classifying them as recipient areas. For instance, wheat cultivations in southern Canada, eastern Europe, Russia, eastern India, Pakistan, and China benefit from rainfall originating from other wheat cultivations (Figure 5.13).

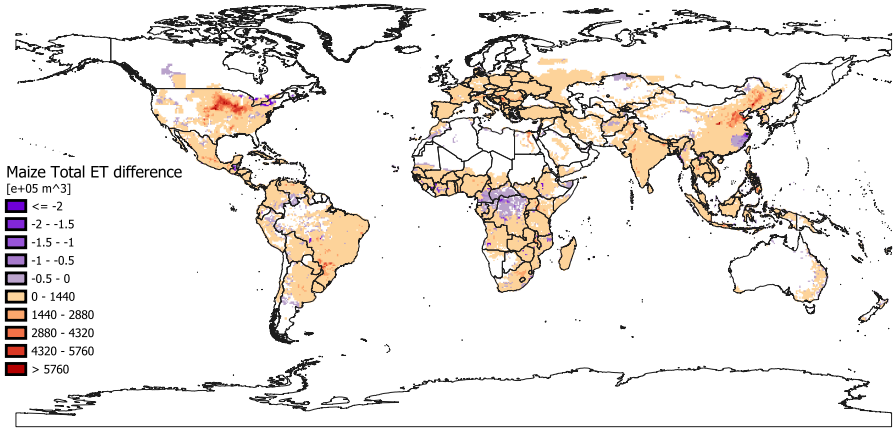
This approach enables the identification of regions that act as net exporters or importers of crop-specific water, highlighting the interdependence between cultivated areas and the precipitation they help generate downwind. Such insights are crucial for understanding regional water dynamics and informing irrigation planning and water management strategies.



(a) Green *ET* difference.

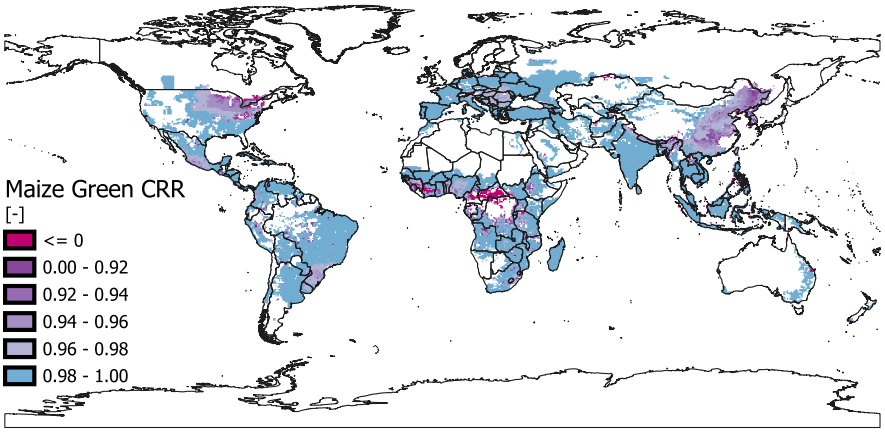


(b) Blue *ET* difference.

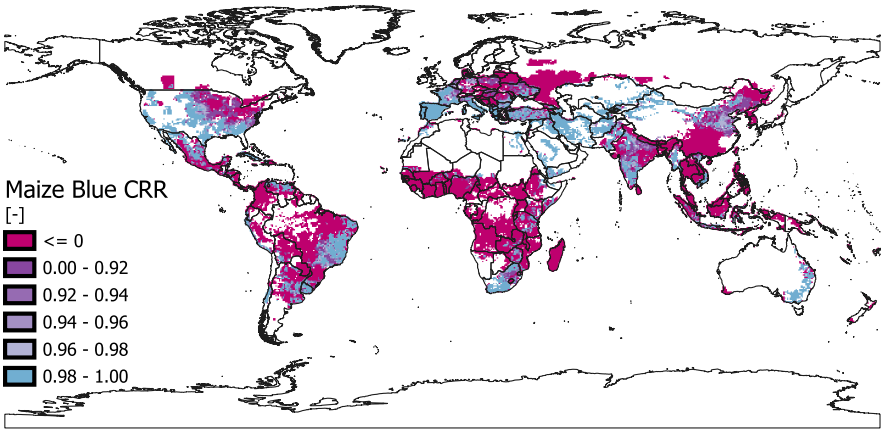


(c) Total *ET* difference.

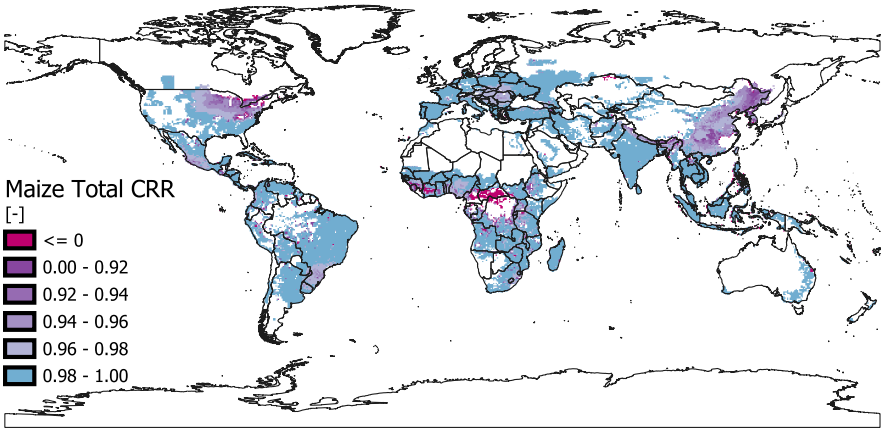
Figure 5.10: Annual difference between maize crop evapotranspiration and the volume of water redistributed as precipitation originated from the same crop (evapotranspiration shed).



(a) Green Crop Recycling Ratio.

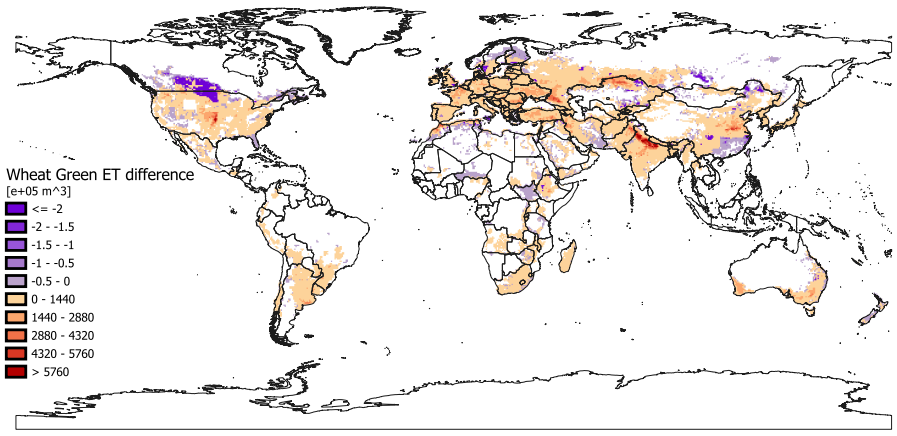


(b) Blue Crop Recycling Ratio.

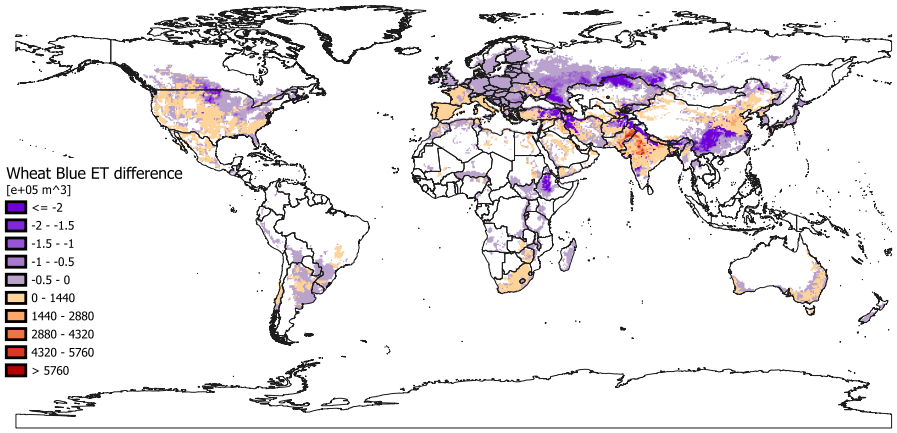


(c) Total Crop Recycling Ratio.

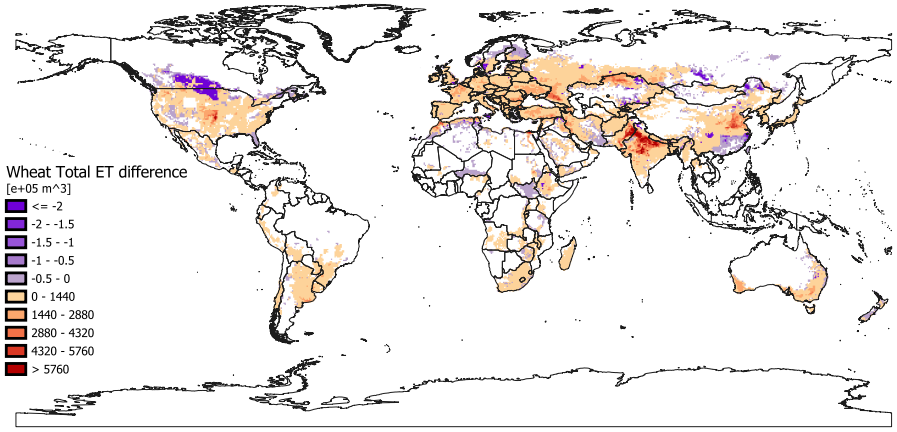
Figure 5.11: Annual Crop Recycling Ratio between maize crop evapotranspiration and the volume of water redistributed as precipitation originated from the same crop (evapotranspiration shed).



(a) Green.

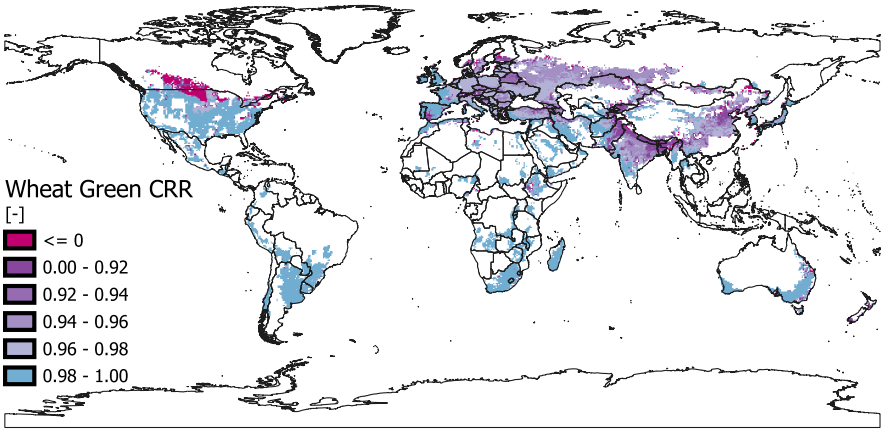


(b) Blue.

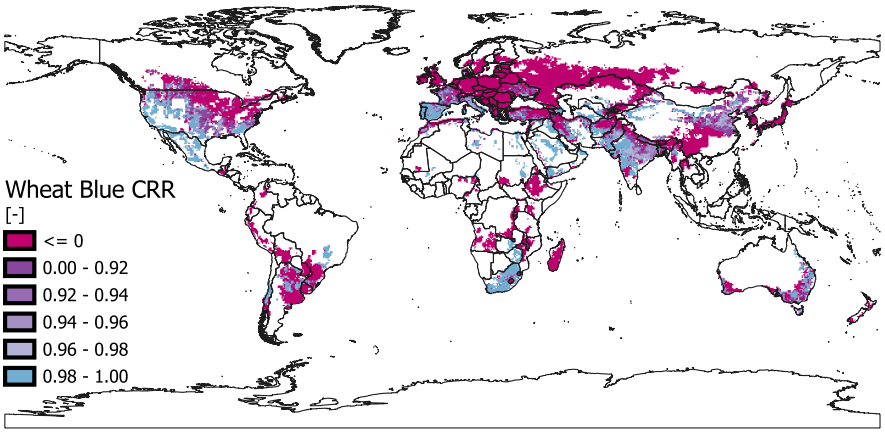


(c) Total.

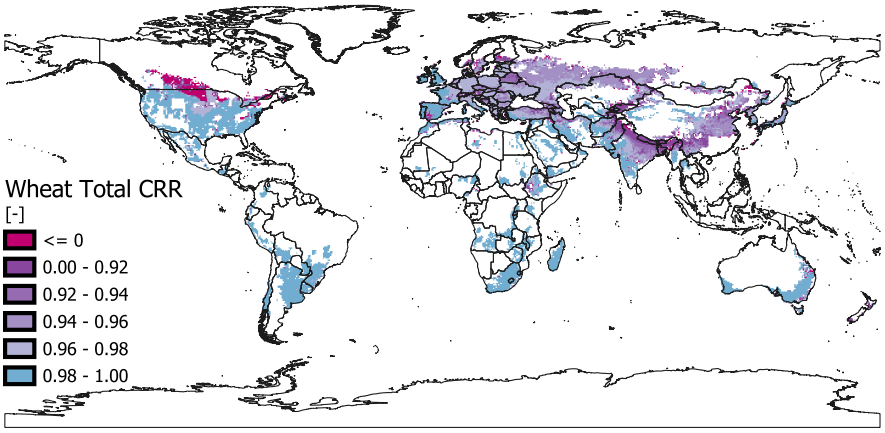
Figure 5.12: Annual difference between wheat crop evapotranspiration and the volume of water redistributed as precipitation originated from the same crop (evapotranspiration shed).



(a) Green Crop Recycling Ratio.

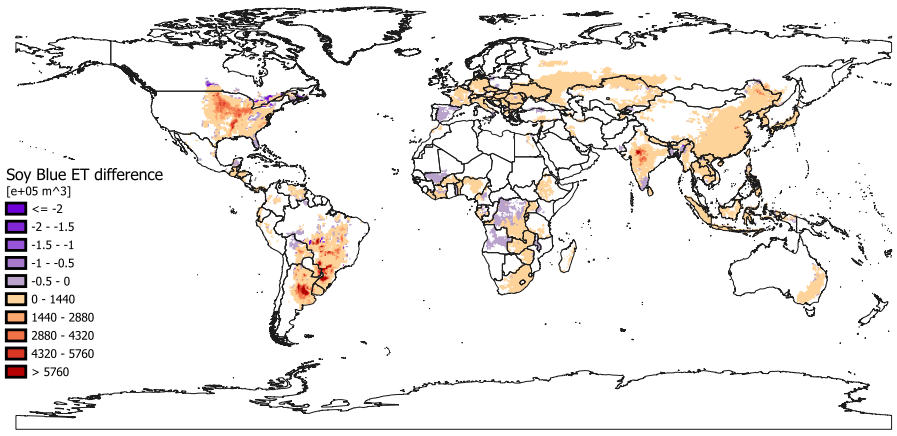


(b) Blue Crop Recycling Ratio.

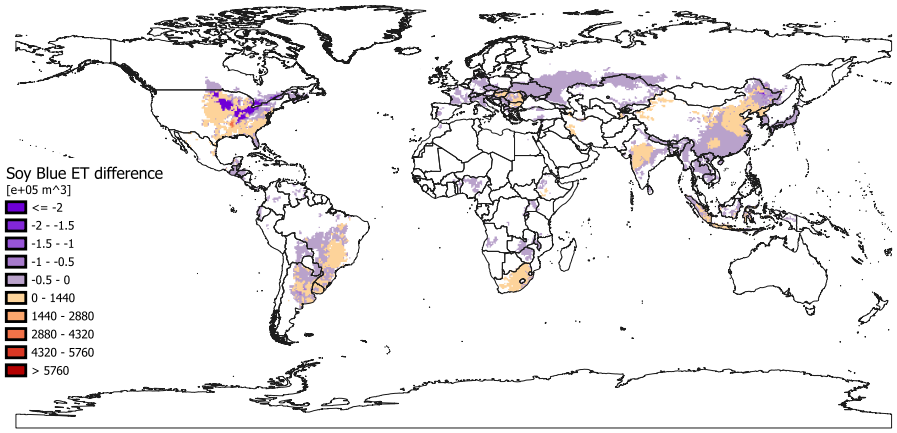


(c) Total Crop Recycling Ratio.

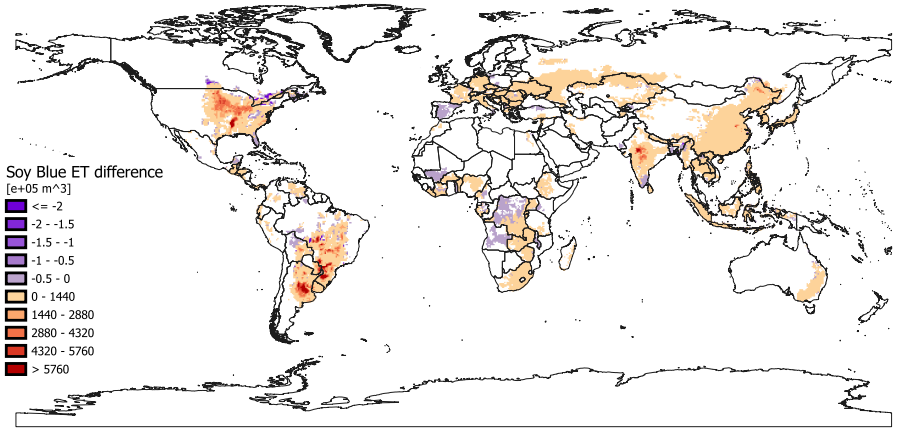
Figure 5.13: Annual Crop Recycling Ratio between wheat crop evapotranspiration and the volume of water redistributed as precipitation originated from the same crop (evapotranspiration shed).



(a) Green.

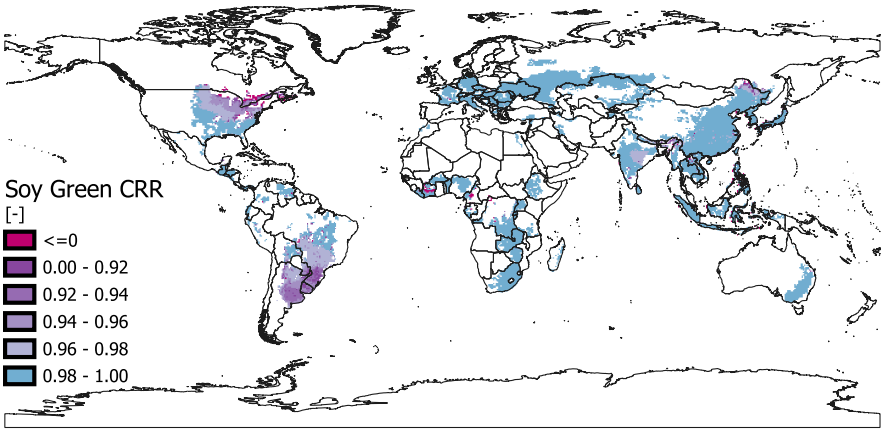


(b) Blue.

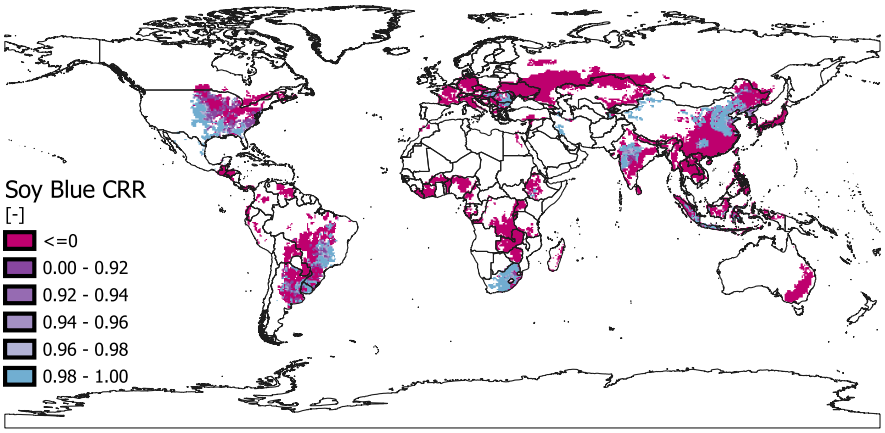


(c) Total.

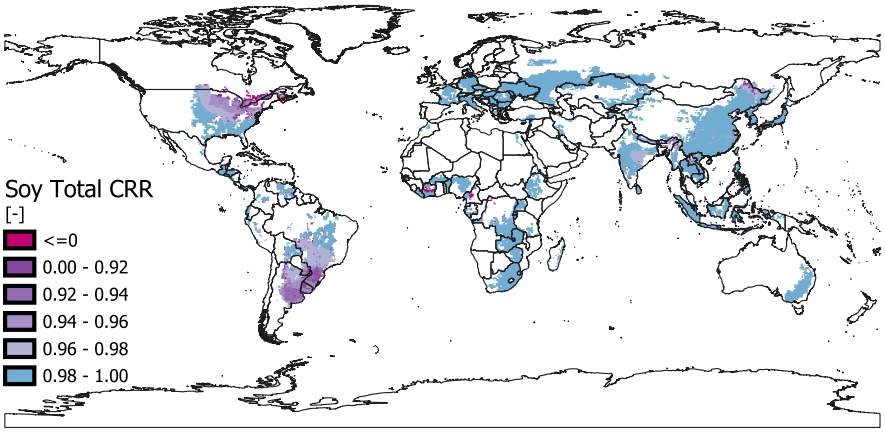
Figure 5.14: Annual difference between soy crop evapotranspiration and the volume of water redistributed as precipitation originated from the same crop (evapotranspiration shed).



(a) Green Crop Recycling Ratio.



(b) Blue Crop Recycling Ratio.



(c) Total Crop Recycling Ratio.

Figure 5.15: Annual Crop Recycling Ratio between soy crop evapotranspiration and the volume of water redistributed as precipitation originated from the same crop (evapotranspiration shed).

5.3 Land use classification of crop water sinks

The objective of this section is not only to determine where the water originating from a specific crop re-precipitates, but also to identify the type of land on which it falls. This analysis helps clarify how crops sustain themselves and other agricultural systems, revealing the degree to which agricultural moisture is recycled within farming landscapes. It also highlights which land types may experience increased precipitation if crop evapotranspiration rises, or reduced precipitation and potential drought if it declines.

To quantify these flows, the volumes of water precipitating over different land-use types are aggregated and visualized with Sankey diagrams. These diagrams (Figures 5.16 to 5.18) display how crop-specific evapotranspired water is redistributed across rainfed and irrigated agricultural areas, other terrestrial surfaces, and the ocean, providing an intuitive overview of the path of agricultural water in the hydrological cycle.

The Sankey diagrams reveal that most of the evapotranspired water from crops re-precipitates over land. This percentage is even higher than the global average moisture recycling ratio [31], since oceanic evapotranspiration is excluded here and only cropland sources are considered.

However, due to the smoothing effect of atmospheric transport on evapotranspiration described in Section 5.2, the redistributed volumes are dispersed over much larger areas than their source regions. As a result, only a small fraction of the precipitation returns to the same crop type. Among the crops analysed, soy shows the highest level of self-recycling within its cultivation areas, although the overall percentage remains low. Furthermore, blue evapotranspiration appears to be more effectively recycled into agricultural land compared to green evapotranspiration. This pattern could be further examined by analysing the relative distance and spatial distribution of irrigated versus rainfed areas.

A more detailed assessment of the “other land” category will require consistent land-use datasets to further disaggregate this class and better capture the diversity of terrestrial sinks.

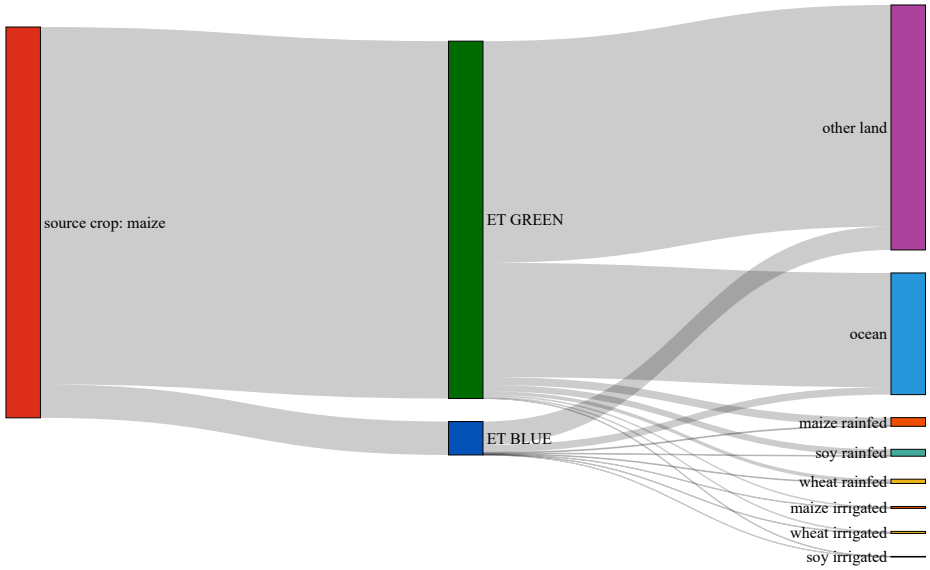


Figure 5.16: Sankey diagram showing the distribution of evapotranspired water from maize cultivation across different land-use types.

Table 5.1: Tabulated values of evapotranspired water redistribution from maize to land-use types. These values correspond to the flows illustrated in Figure 5.16.

	Sink land use	**Volume** $[m^3]$	**Percentage** [%]		
ET Green	maize rainfed	1.36E+10	2.22		
	wheat rainfed	6.28E+09	1.03	5.02	
	soy rainfed	1.08E+10	1.77		6.04
	maize irrigated	2.78E+09	0.46		
	wheat irrigated	2.97E+09	0.49	1.02	
	soy irrigated	4.16E+08	0.07		
	ocean	1.95E+11		31.96	
	other land	3.79E+11		62.01	
ET Blue	maize rainfed	1.32E+09	2.31		
	wheat rainfed	1.07E+09	1.87	5.83	
	soy rainfed	9.47E+08	1.65		7.82
	maize irrigated	5.07E+08	0.89		
	wheat irrigated	5.41E+08	0.94	1.99	
	soy irrigated	9.07E+07	0.16		
	ocean	1.28E+10		22.31	
	other land	4.00E+10		69.87	

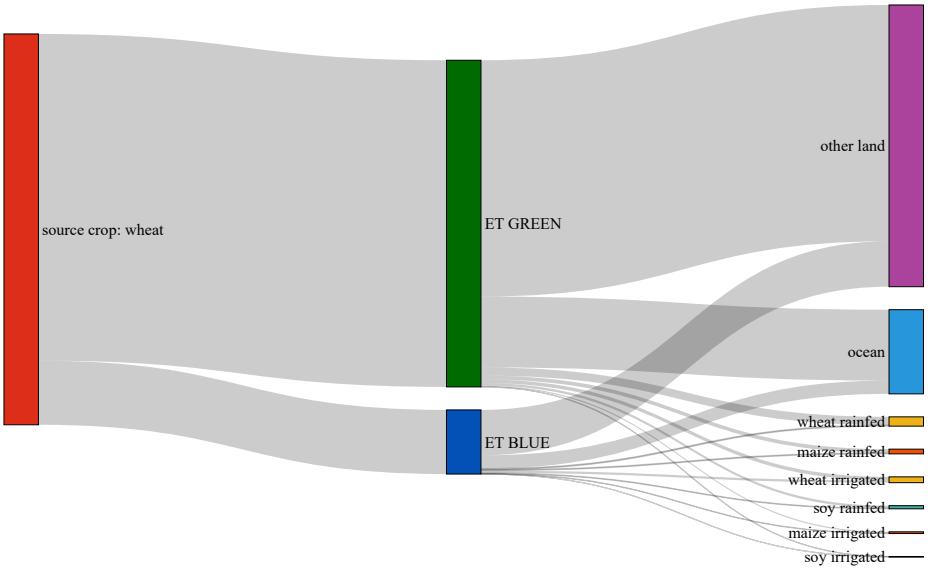


Figure 5.17: Sankey diagram showing the distribution of evapotranspired water from wheat cultivation across different land-use types.

Table 5.2: Tabulated values of evapotranspired water redistribution from wheat to land-use types. These values correspond to the flows illustrated in Figure 5.17.

	Sink land use	**Volume** $[m^3]$	**Percentage** [%]		
ET Green	maize rainfed	9.06E+09	1.16		5.96
	wheat rainfed	1.96E+10	2.51	4.43	
	soy rainfed	5.88E+09	0.76		
	maize irrigated	3.03E+09	0.39		
	wheat irrigated	8.50E+09	1.09	1.53	
	soy irrigated	3.93E+08	0.05		
	ocean	1.69E+11		21.70	
	other land	5.63E+11		72.34	
ET Blue	maize rainfed	2.20E+09	1.44		8.49
	wheat rainfed	2.59E+09	1.69	4.20	
	soy rainfed	1.63E+09	1.07		
	maize irrigated	1.11E+09	0.72		
	wheat irrigated	5.36E+09	3.51	4.29	
	soy irrigated	8.49E+07	0.06		
	ocean	3.17E+10		20.73	
	other land	1.08E+11		70.79	

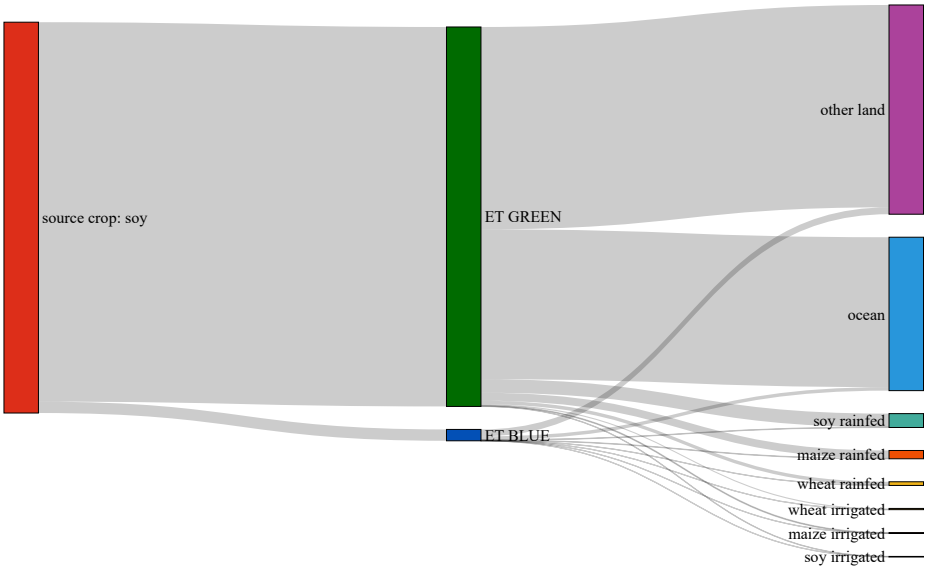


Figure 5.18: Sankey diagram showing the distribution of evapotranspired water from soy cultivation across different land-use types.

Table 5.3: Tabulated values of evapotranspired water redistribution from soy to land-use types. These values correspond to the flows illustrated in Figure 5.18.

	Sink land use	**Volume** $[m^3]$	**Percentage** [%]		
ET Green	maize rainfed	9.85E+09	2.10		
	wheat rainfed	4.40E+09	0.94	6.61	
	soy rainfed	1.68E+10	3.57		7.18
	maize irrigated	8.94E+08	0.19		
	wheat irrigated	1.44E+09	0.31	0.57	
	soy irrigated	3.21E+08	0.07		
	ocean	1.86E+11		39.51	
	other land	2.50E+11		53.33	
ET Blue	maize rainfed	4.25E+08	3.02		
	wheat rainfed	1.54E+08	1.10	7.06	
	soy rainfed	4.13E+08	2.94		8.52
	maize irrigated	9.46E+07	0.67		
	wheat irrigated	5.96E+07	0.42	1.46	
	soy irrigated	5.20E+07	0.37		
	ocean	4.35E+09		30.90	
	other land	8.52E+09		60.58	

5.4 Consistency of total water volumes

An evaluation of the consistency between the waterCROP outputs and the crop-specific evapotranspiration sheds was conducted. The waterCROP results are considered reliable, as the model has already been validated through comparisons with the literature [28]. Assuming a closed annual hydrological cycle, this assessment verifies whether the shed-based calculations systematically under- or overestimate the total evapotranspired volumes.

As shown in Table 5.4, the cumulative shed volumes are slightly higher than those obtained directly from the waterCROP model. However, the discrepancy is very small, less then 0.05%, indicating that the methodology preserves the overall water balance with high fidelity.

Table 5.4: Total global water volumes across the different calculation steps, with the relative difference between them expressed as a percentage. Results indicate that the cumulative sheds slightly overestimate total volumes compared to the direct waterCROP estimates, but the discrepancy is minimal, confirming overall conservation of the water balance.

		A **Annual ET** [m^3] waterCROP 5 arc-min	***B*** **Annual ET upscaled** [m^3] gridboxsum 0.5 deg	***C*** **Annual ET shed** [m^3] RECON 0.5 deg	$A - B$ [%]	$B - C$ [%]	$A - C$ [%]
Maize	Green	6.12E+11	6.12E+11	6.12E+11	0	-0.0431	-0.0431
	Blue	5.75E+10	5.75E+10	5.76E+10	1.33E-14	-0.0437	-0.0437
	Total	6.70E+11	6.70E+11	6.70E+11	0	-0.0431	-0.0431
Wheat	Green	7.81E+11	7.81E+11	7.81E+11	-3.13E-14	-0.0438	-0.0438
	Blue	1.53E+11	1.53E+11	1.53E+11	-4.00E-14	-0.0449	-0.0449
	Total	2.80E+12	2.80E+12	2.80E+12	-1.74E-14	-0.0440	-0.0440
Soy	Green	4.69E+11	4.69E+11	4.70E+11	-6.50E-14	-0.0398	-0.0398
	Blue	1.41E+10	1.41E+10	1.41E+10	0	-0.0357	-0.0357
	Total	4.84E+11	4.84E+11	4.84E+11	-6.31E-14	-0.0397	-0.0397

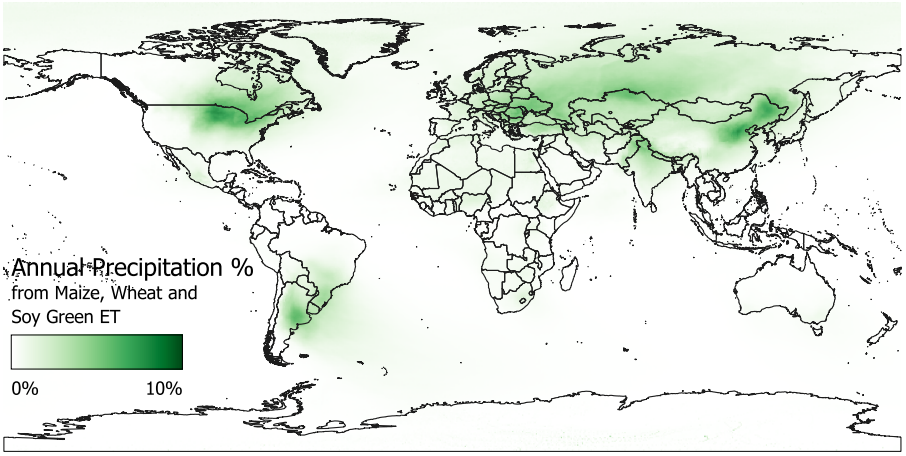
Chapter 6

Conclusions

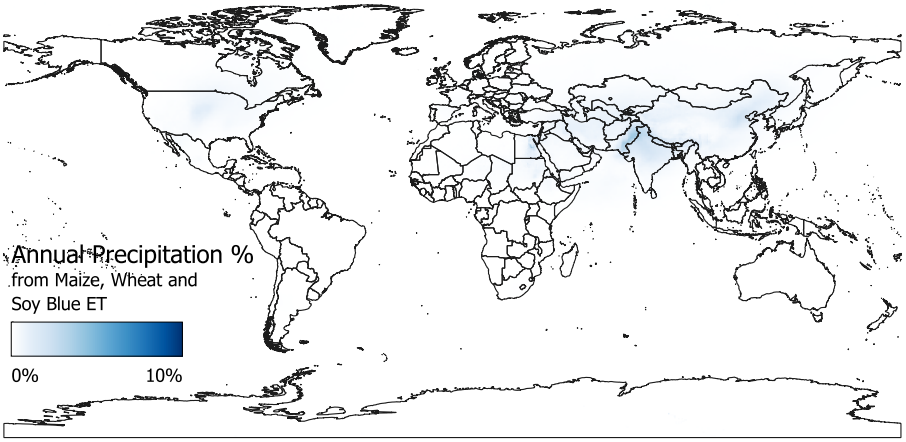
Agriculture is deeply intertwined with the hydrological cycle, not only through direct water inputs such as precipitation and irrigation, but also through the recycling of water via crop evapotranspiration. Previous studies have shown that land-use changes and vegetation feedbacks can significantly influence rainfall patterns in downwind regions, but a crop-specific and globally consistent quantification of these processes has remained limited. To provide a broader perspective on what happens to the green water accounted for in the water footprint of crops, and how this water re-enters the cycle and can be reused, this thesis developed and applied the waterCROP model in combination with the RECON dataset. The aim was to explore how major crops contribute to atmospheric moisture recycling and to map where this recycled water subsequently precipitates.

The approach integrated global crop distribution datasets, actual evapotranspiration (distinguishing between green and blue), and a Lagrangian moisture tracking framework. This combination enabled the construction of crop-specific evapotranspiration sheds, crop-specific water balances, and land-use classifications of water sinks on an annual basis. By systematically linking source areas of crop evapotranspiration to downwind precipitation, the method advances the understanding of how agricultural water use contributes to the global water cycle.

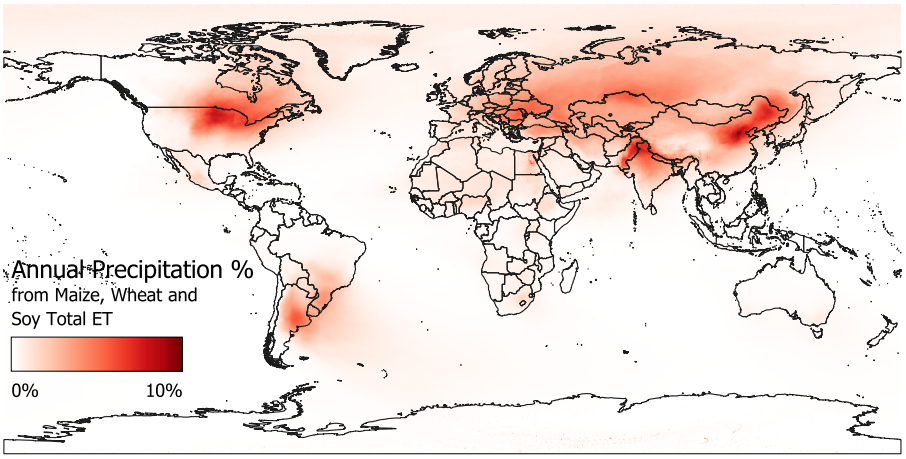
Crop-specific evapotranspiration sheds were mapped for maize, wheat, and soybean. These revealed clear geographical patterns shaped by topography and large-scale atmospheric circulation, showing how agricultural water vapour released in one region supports rainfall in distant areas. Notably, atmospheric transport smooths out the intensity of evapotranspiration inputs, redistributing water more evenly across larger regions. Cumulative precipitation fractions attributable to crop evapotranspiration were also quantified. In some regions, a single crop contributes up to 5% of annual precipitation (e.g., maize in Minnesota and in parts of China). Irrigation (Blue ET) emerged as a major source of precipitation, with examples such as wheat irrigation in Pakistan contributing more than 3% of annual rainfall locally. By summing contributions from all crops, the overall importance of agriculture becomes even clearer. Figures 6.1 show that 5% of contribution is frequently exceeded: with peaks of 8.2% in North America, 5.8% in Argentina, 5.2% in Ukraine, 5.7% in Romania and Moldova, 6% in Egypt, 5.7% in northern Kazakhstan, 7.8% in India and Pakistan,



(a) Green *ET* contribution.

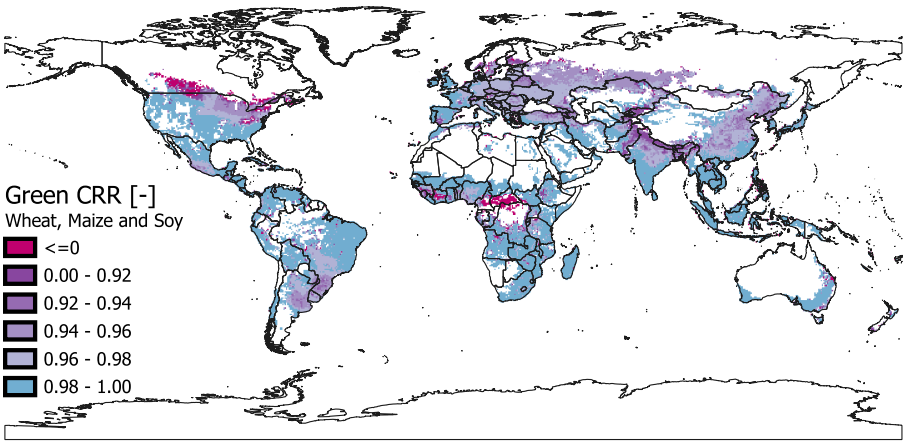


(b) Blue *ET* contribution.

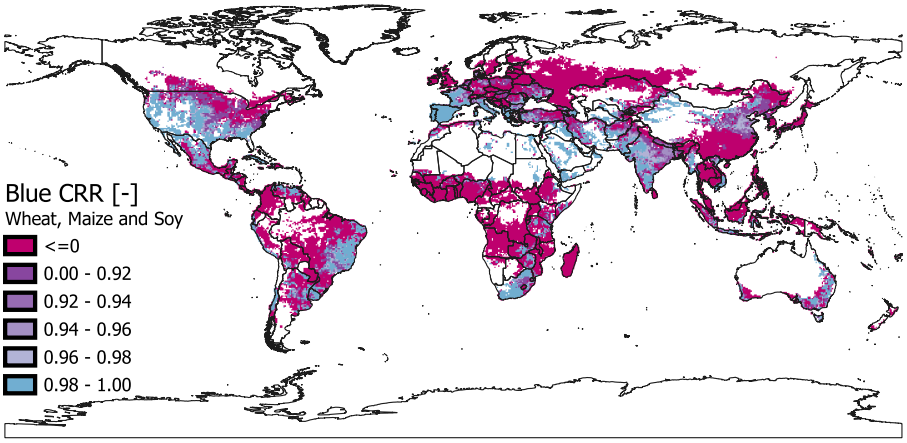


(c) Total *ET* contribution.

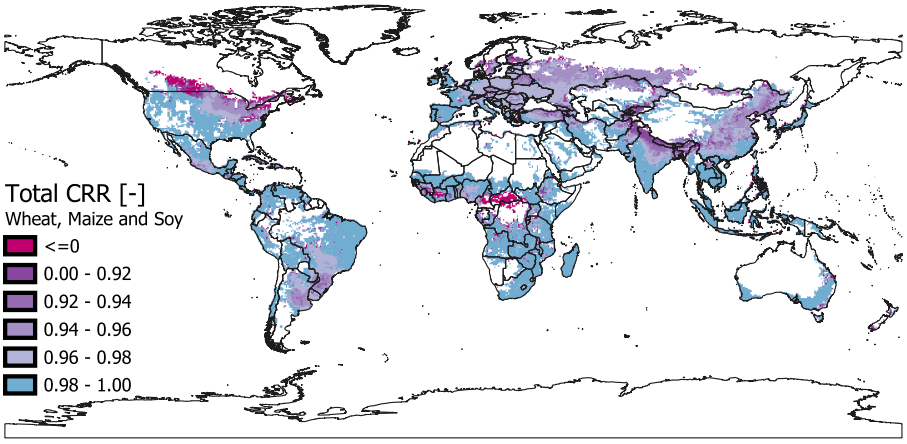
Figure 6.1: Fraction of annual precipitation originated from global Maize, Wheat and Soy cultivations.



(a) Green Crop Recycling Ratio.



(b) Blue Crop Recycling Ratio.



(c) Total Crop Recycling Ratio.

Figure 6.2: Crop Recycling Ratio of maize, wheat and soy.

and up to 9.8% in China.

Donor and recipient areas (Figure 6.2) were identified for each crop through an annual water balance analysis. Rainfed croplands often act as recipients, benefitting from irrigation-fed donor regions. For instance, cultivations in southern Canada receives water recycled from cultivations in the USA, while intensive irrigation in South Asia emerges as a net exporter of water vapour, not fully compensated, but with a relevant contribution from crop recycling.

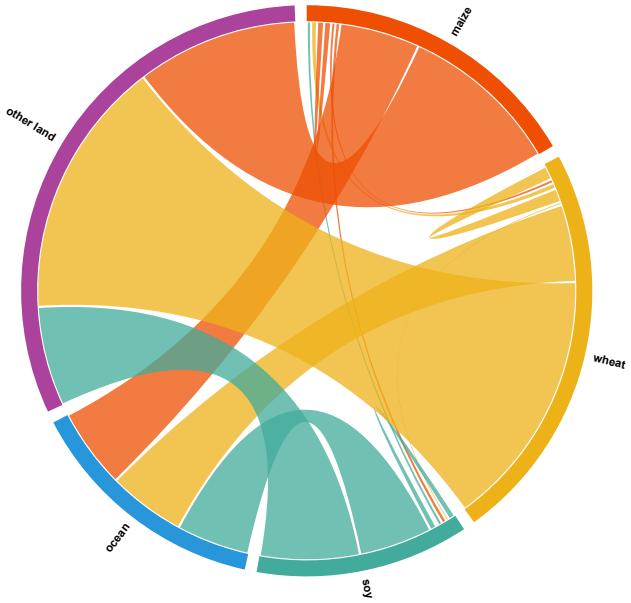


Figure 6.3: Chord diagram showing the distribution of annual evapotranspired water from maize, wheat and soy cultivation across different land-use types. 6.51% falls again on croplands of these three crops, 65.74% on other land and 28.75% on oceans.

Land-use classification of sinks demonstrated that most crop evapotranspired water precipitates back on land rather than the ocean. However, due to the atmospheric smoothing effect, only a small fraction falls back on the same crop type, with soy showing the highest, though still limited, self-recycling potential. Sankey diagrams provided a visual representation of how agricultural water is redistributed across rainfed cropland, irrigated cropland, other land types, and the ocean. By combining these land-use classifications, a chord diagram was created (Figure 6.3), summarizing all the Sankey diagrams presented in Section 5.3. This visualization highlights the limited extent of water recycling within the same crop type and, more generally, the small fraction of evapotranspired water that returns to cropland. Overall, only 6.51% of water evapotranspired from maize, wheat, and soy precipitates again on croplands of these three crops, while 65.74% falls on other land (including croplands of different crops), and 28.75% ends up in the oceans.

These findings offer several important insights. First, they underscore that agriculture is not only a consumer of water but also a driver of regional and global

precipitation dynamics. Second, they highlight how local changes in land use or irrigation practices can propagate far beyond the field scale, altering water availability in distant regions. Finally, they demonstrate the need to integrate atmospheric moisture recycling into assessments of agricultural sustainability and water management.

At the scientific level, the framework and datasets developed in this thesis provide a basis for refining global water-use estimates by incorporating crop-specific feedbacks. For policy and management, the identification of donor and recipient areas can inform transboundary water governance and help evaluate the unintended consequences of irrigation expansion or land-use change. More broadly, the results contribute to ongoing discussions on food–water–climate interdependencies, emphasizing the importance of considering atmospheric water flows in sustainable agriculture strategies.

In conclusion, this thesis advances the understanding of crop-specific moisture recycling by quantifying how agricultural evapotranspiration contributes to rainfall across the globe. By filling the gap between field-scale water use and atmospheric-scale precipitation dynamics, it provides a novel perspective on the role of agriculture in shaping the global water cycle, offering both conceptual insights and practical tools for future research and water management.

6.1 Limitations

While the present work provides new insights into agriculture as a driver of regional and global precipitation dynamics, several limitations need to be acknowledged.

First, the analysis relies on annual averages of evapotranspiration and precipitation. As described in Section 4.1, evapotranspiration values were aggregated at the yearly scale to ensure consistency with the RECON dataset. This choice smooths out seasonal dynamics, which are particularly relevant for crops whose evapotranspiration peaks may or may not coincide with rainfall seasons. As a result, water recycled from one crop and precipitating onto another could fall outside its cultivation period. Capturing seasonal differences would therefore be important to assess whether crop water use reinforces or counteracts local rainfall regimes in downwind cultivated areas.

Finally, the propagation of uncertainties affects all stages of the analysis. Both the derivation of crop-specific evapotranspiration sheds (Section 4.2) and the classification of precipitation sinks by land-use type (Section 4.3) rely on multiple datasets and modelling steps, each introducing potential errors. Systematically quantifying and propagating these uncertainties remains a necessary step for strengthening the robustness of the results.

6.2 Future developments

Building on the results of this thesis, several directions for future research emerge.

First, the classification of precipitation sinks could be refined by including a broader set of land-use categories. In Section 4.3, the current analysis distinguishes between irrigated and rainfed croplands of the three main crops analysed, other terrestrial areas, and the ocean. Expanding this classification to additional land types would

allow a more detailed assessment of how agricultural evapotranspiration supports different ecosystems.

Second, it would be valuable to disentangle the contribution of local moisture recycling from that of long-range atmospheric transport. Section 5.2 highlighted the smoothing effect of atmospheric transport on evapotranspiration signals, but a clearer separation between local and transported components would help to clarify the relative importance of nearby versus remote feedbacks.

Third, future work should investigate not only the sinks of crop evapotranspiration, but also the sources of precipitation sustaining agricultural areas. The Lagrangian atmospheric tracking model adopted in this study already provides the necessary data to map precipitation origins. This could be extended to build crop-specific precipitation sheds, thereby identifying what water sources are feeding agriculture.

Finally, a broader integration with land-use change driven by dietary habits would represent a significant research advancement. As introduced in the background, deforestation, often linked to global dietary patterns and demand for specific crops, plays a crucial role in shaping both water availability and atmospheric moisture flows. Extending the present framework to consider how dietary choices, from single crops to animal feed and protein sources, influence precipitation patterns would enable a direct connection between food systems and hydro-climatic impacts.

Understanding how everyday human choices, such as dietary preferences, shape precipitation patterns and trigger consequences thousands of kilometres away would be crucial for informing both individuals and decision-makers.

Appendix A

waterCrop model

```
% Note: when you start using the script you should modify the path
% to the folder, both for input and for output data.

clear all
clc
close all

tic
%% Inputs

% simulation performed over the following years
period = [2008 2017];
% Folders to reach input files

% folder with the tif file which define climatic regions
cartella_climate = 'path/to/climatic/regions';

% folder with monthly global ET0
cartella_ET0 = 'path/to/ET0';

% folder with monthly global Precipitation
cartella_Pre = 'path/to/Pre';

% folder with global crop coefficients
cartella_kc = 'path/to/kc';

% folder with available water contents
cartella_awc = 'path/to/awc';


% Input climatic regions

cd(cartella_climate)
climatic_zone_file = 'thcli1.tif';
```

```
35 climate = imread(climatic_zone_file);
36 % Defines climatic regions: each cell has a value between 1 and 10 which
37 % corresponds to a different climatic region
38
39 ROWS=[1,imfinfo(climatic_zone_file).Height];
40 COLS=[1,imfinfo(climatic_zone_file).Width];
41
42
43 % Coordinates input
44 % Latitude and longitude should be written in degrees. The code works both
45 % with two scalar values or with two vectors contatining the start and
46 % the end of the region.
47
48
49 % input
50
51 lat = [44.76 44.76];       % cell in Piedmont
52 lon = [7.51 7.51];
53
54 lat = [90 -89.99];         % WORLD
55 lon = [-180 179.99];
56
57 % grid resolution
58 lat_res = 180 / ROWS(2);  % 5 arcmin in degrees
59 lon_res = 360 / COLS(2);
60
61 % transform the lat and lon into indeces
62 ind_row = floor(-(1/lat_res)*lat + ROWS(2)/2) +1
63 ind_col = floor((1/lon_res)*lon + COLS(2)/2) +1
64
65 %% Crop selection
66
67 raccolto=1:40;
68 for r = 3:3 %
69     r
70     switch raccolto(r)
71
72     %%%%%%%%%%%%%%%%%%----------------    First growing season
73
74         case 1 %Maize
75             cartella='path/to/grow_season_I';
76             cartella_risultati='Output';
77
78             mkdir(cartella_risultati)
79
80             area_irrigata='area_ir.mat';
81             area_rainfed='area_rf.mat';
82
83             data_semina='semina_rf.mat';
```

```
84              data_semina_irr='semina_ir.mat';
85              lgp='lgp_rf.mat';
86              lgp_irr='lgp_ir.mat';
87
88              coeff_colturale='kc_maize';
89
90              rd_ini=0.3;  % [m]
91              rd_max_rainfed=1.7; %fao 56 tab 22 pag 190
92              rd_max_irrigated=1;
93              depl_fraction=0.55;
94
95          case 2 %Rice
96              cartella='path/to/grow_season_I';
97              cartella_risultati='Output';
98
99              mkdir(cartella_risultati)
100
101             area_irrigata='area_ir.mat';
102             area_rainfed='area_rf.mat';
103
104             data_semina='semina_rf.mat';
105             data_semina_irr='semina_ir.mat';
106             lgp='lgp_rf.mat';
107             lgp_irr='lgp_ir.mat';
108
109             coeff_colturale='rice';
110
111             rd_ini=0.3;
112             rd_max_rainfed=1; %fao 56 tab 22 pag 190
113             rd_max_irrigated=0.5;
114             depl_fraction=0.2;
115
116         case 3 %Wheat
117             cartella='path/to/grow_season_I';
118             cartella_risultati='Output';
119
120             mkdir(cartella_risultati)
121
122             area_irrigata='area_ir.mat'; % [ha]
123             area_rainfed='area_rf.mat';
124
125             coeff_colturale='wheat'; % name of the Excel sheet
126
127             data_semina='semina_rf.mat';
128             data_semina_irr='semina_ir.mat';
129             lgp='lgp_rf.mat';
130             lgp_irr='lgp_ir.mat';
131
132             rd_ini=0.3; % [m]
```

```
133             rd_max_rainfed=1.8; % [m] %Siebert and doll
134             rd_max_irrigated=1.5; % [m]
135             depl_fraction=0.55;
136 
137         case 4 %Soy
138             cartella='path/to/grow_season_I';
139             cartella_risultati='Output';
140 
141             mkdir(cartella_risultati)
142 
143             coeff_colturale='kc_soybean';
144 
145             area_irrigata='area_ir.mat';
146             area_rainfed='area_rf.mat';
147 
148             data_semina='semina_rf.mat';
149             data_semina_irr='semina_ir.mat';
150             lgp='lgp_rf.mat';
151             lgp_irr='lgp_ir.mat';
152 
153             rd_ini=0.3;
154             rd_max_rainfed=1.30; %Siebert and doll
155             rd_max_irrigated=0.60;
156             depl_fraction=0.50;
157 
158     end
159 
160 
161 % Crop characteristics
162 
163 cd(cartella)
164 % cultivated area with 5x5 minutes of arc resolution
165 % two rasters, area rainfed and irrigated area
166 area_irr=importdata(area_irrigata); % [ha]
167 area_rain=importdata(area_rainfed); % [ha]
168 area_tot=area_irr+area_rain; % [ha]
169 
170 % sowing date with 5x5 minutes of arc resolution in a 360days
171 % calendar
172 day_plant_modified_0=importdata(data_semina);
173 day_plant_modified_irr_0=importdata(data_semina_irr);
174 
175 % length of growing period of the crop
176 % calculated as data_raccolta-data_semina (harvest date - sowing date)
177 lgp_ini=importdata(lgp);
178 lgp_ini_irr=importdata(lgp_irr);
179 
180 % crop coefficient and length of growing phases in percentages
181 % Hoekstra values, function of the 10 climatic regions
```

```
cd(cartella_kc)
kc=xlsread('kc_global_NEWCROPS_def.xlsx',coeff_colturale,'C17:I26');

% soil water capacity
cd(cartella_awc)
awc_final=importdata('awc_mmalm.mat');

% Initialization

% initialization of matrices of final results
ETc_tot_rain=zeros(ROWS(2),COLS(2)); % Crop Evapotranspiration
ETc_tot_irr=zeros(ROWS(2),COLS(2));

% RAINFED: only precipitation
% Total Evapotranspiration is Green Evapotranspiration
% effective evapotranspiration: ETc*ks
ETa_tot=zeros(ROWS(2),COLS(2));
% precipitation totally infiltrated in the soil
Ptot_rf=zeros(ROWS(2),COLS(2));
Ptot_gr_seas_rf=zeros(ROWS(2),COLS(2));

% IRRIGATED
I_tot=zeros(ROWS(2),COLS(2));
ETgreen_tot=zeros(ROWS(2),COLS(2));
ETblue_tot=zeros(ROWS(2),COLS(2));
% coincide with ETa because k_s=1
CWU_tot=zeros(ROWS(2),COLS(2));
% precipitation totally infiltrated in the soil
Ptot_ir=zeros(ROWS(2),COLS(2));
Ptot_gr_seas_ir=zeros(ROWS(2),COLS(2));

ET0_tot_rf=zeros(ROWS(2),COLS(2));
ET0_tot_ir=zeros(ROWS(2),COLS(2));


%% Here starts the code that run for each year

for y=period(1,1):period(1,2)     % run for each year
yy = y-period(1,1)+1;        % index which goes from 1 to number of years

% is it a leap year?
if mod(y,4) == 0
    ndays = 366;  % Leap year
    days_in_months365 = [31, 29, 31, 30, 31, 30, 31, 31, 30, 31, 30, 31];
else
    ndays = 365;  % Non-leap year
    days_in_months365 = [31, 28, 31, 30, 31, 30, 31, 31, 30, 31, 30, 31];
end

```

```
231
232 % Input global daily ET0 and Precipitation
233
234 % Input raster of Global Daily Mean ET0.
235 % Spatial resolution: 5 minutes of arc
236 cd(cartella_ET0)
237 fileET0 =  ['dailyET0_' num2str(y) '_WC.nc'];
238 ET0_orig = ncread(fileET0, 'pev');
239 ET0 = pagetranspose(ET0_orig);   % to have the array with rows,cols,days
240 ET0 = ET0*(-1000)*24;            % from m/h to mm/day
241
242 ET0(ET0 < 0) = NaN;
243
244 % Input raster of Global Cumulative Daily Precipitation [mm].
245 % Spatial resolution: 5 minutes of arc
246 cd(cartella_Pre)
247 filePre =  ['dailyPre_' num2str(y) '_WC.nc'];
248 Pre_orig = ncread(filePre, 'tp');
249 Pre = pagetranspose(Pre_orig);
250 Pre = Pre*1000*24;               % from m to mm
251
252 % Conversion from 360days calendar to 365days
253 % (or 366 for leap years) for SOWING DATES
254
255 day_plant_modified = day_plant_modified_0;
256 day_plant_modified_irr = day_plant_modified_irr_0;
257
258 % Define the number of days in each month for a year
259 days_in_months360 = 30*ones(1,12);
260
261 % Create a cumulative sum of days in the months
262 cum_days365 = cumsum(days_in_months365);
263 cum_days360 = cumsum(days_in_months360);
264
265 % Initialize an array to store converted days (same size as dates_360)
266 dates_365 = zeros(ROWS(2), COLS(2));
267 dates_365_irr = zeros(ROWS(2), COLS(2));
268
269 for i = 1:ROWS(2)
270     for j = 1:COLS(2)
271         day_360 = day_plant_modified(i, j);
272         day_360_irr = day_plant_modified_irr(i, j);
273
274         % Find the month for the given day
275         mon = find(cum_days365 >= day_360, 1);
276         mon_irr = find(cum_days365 >= day_360_irr, 1);
277
278         if mon == 1
279             day_of_year = day_360; % If it's the first month
```

```
280         else
281             day_of_year = day_360 +
282                 cum_days365(mon-1) - cum_days360(mon-1);
283         end
284
285         if mon_irr == 1
286             day_of_year_irr = day_360_irr; % If it's the first month
287         else
288             day_of_year_irr = day_360_irr +
289                 cum_days365(mon_irr-1) - cum_days360(mon_irr-1);
290         end
291
292         % Store the converted day
293         dates_365(i, j) = day_of_year;
294         dates_365_irr(i, j) = day_of_year_irr;
295     end
296 end
297
298 day_plant_modified = dates_365;
299 day_plant_modified_irr = dates_365_irr;
300
301 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
302
303 %% Calcultations start here
304
305
306 for m = ind_row(1):1:ind_row(2)          % read along the rows
307     for n = ind_col(1):1:ind_col(2)      % read along the columns
308
309         if area_tot(m,n)>0 && climate(m,n)>0 &&
310         awc_final(m,n) && ET0(m,n,1)>=0 && Pre(m,n,1)>=0
311
312     ET0_daily=squeeze(ET0(m,n,:));
313     Pre_tot_daily=squeeze(Pre(m,n,:));
314
315
316 %% build kc and lgp based on the climatic zone
317
318     sel_climate = climate(m,n);
319
320 if sel_climate ~= 0
321
322         kc_ini=kc(sel_climate,1);
323         kc_mid=kc(sel_climate,2);
324         kc_end=kc(sel_climate,3);
325
326         lgp_2=round(lgp_ini(m,n).*kc(sel_climate,5));
327         lgp_3=round(lgp_ini(m,n).*kc(sel_climate,6));
328         lgp_4=round(lgp_ini(m,n).*kc(sel_climate,7));
```

```
329         lgp_1=lgp_ini(m,n)-lgp_2-lgp_3-lgp_4;
330
331
332         lgp_2_irr=round(lgp_ini_irr(m,n).*kc(sel_climate,5));
333         lgp_3_irr=round(lgp_ini_irr(m,n).*kc(sel_climate,6));
334         lgp_4_irr=round(lgp_ini_irr(m,n).*kc(sel_climate,7));
335         lgp_1_irr=lgp_ini_irr(m,n)-lgp_2_irr-lgp_3_irr-lgp_4_irr;
336
337 else
338         kc_ini=0;
339         kc_mid=0;
340         kc_end=0;
341         lgp_1=0;
342         lgp_2=0;
343         lgp_3=0;
344         lgp_4=0;
345         lgp_1_irr=0;
346         lgp_2_irr=0;
347         lgp_3_irr=0;
348         lgp_4_irr=0;
349 end
350
351 %% build kc in the rainfed and in the irrigated case
352 lgp=lgp_1+lgp_2+lgp_3+lgp_4;
353 lgp_irr=lgp_1_irr+lgp_2_irr+lgp_3_irr+lgp_4_irr;
354 kc_crop=zeros(lgp,1);
355 kc_crop_irr=zeros(lgp_irr,1);
356
357 % Rainfed
358 kc_crop(1:lgp_1)=kc_ini;
359 for i=lgp_1+1:lgp_1+lgp_2
360     kc_crop(i)=(kc_mid-kc_ini)/lgp_2*(i-lgp_1)+kc_ini;
361 end
362 kc_crop(lgp_2+lgp_1+1:lgp_2+lgp_1+lgp_3)=kc_mid;
363 for i=lgp_2+lgp_1+1+lgp_3:lgp
364     kc_crop(i)=(kc_end-kc_mid)/lgp_4*(i-lgp_3-lgp_2-lgp_1)+kc_mid;
365 end
366
367
368 % Irrigated
369 kc_crop_irr(1:lgp_1_irr)=kc_ini;
370 for i=lgp_1_irr+1:lgp_1_irr+lgp_2_irr
371     kc_crop_irr(i)=(kc_mid-kc_ini)/lgp_2_irr*(i-lgp_1_irr)+kc_ini;
372 end
373 kc_crop_irr(lgp_2_irr+lgp_1_irr+1:lgp_2_irr+lgp_1_irr+lgp_3_irr)=kc_mid;
374 for i=lgp_2_irr+lgp_1_irr+1+lgp_3_irr:lgp_irr
375     kc_crop_irr(i)=(kc_end-kc_mid)/lgp_4_irr*
376         (i-lgp_3_irr-lgp_2_irr-lgp_1_irr)+kc_mid;
377 end
```

```
378
379 % function to calculate the temporal variability of the crop rooting depth
380 % values from FAO 56 tabella 22, pag 163. Two values are given.
381
382
383 % in the RAINFED scenario the maximum depth was selected
384 rd=zeros(lgp,1);
385 rd(1)=rd_ini;
386 for i=2:lgp_1+lgp_2
387     rd(i)=rd_ini+(rd_max_rainfed-rd_ini)/(lgp_1+lgp_2)*i;
388 end
389 rd(lgp_1+lgp_2+1:lgp)=rd_max_rainfed;
390
391 % in the IRRIGATED scenario the minimum depth was selected
392 rd_irrigated=zeros(lgp_irr,1);
393 rd_irrigated(1)=rd_ini;
394 for i=2:lgp_1_irr+lgp_2_irr
395     rd_irrigated(i)=rd_ini+(rd_max_irrigated-rd_ini)./
396         (lgp_1_irr+lgp_2_irr)*i;
397 end
398 rd_irrigated(lgp_1_irr+lgp_2_irr+1:lgp_irr)=rd_max_irrigated;
399
400 % calculate TAWC in the RAINFED scenario
401 tawc=zeros(lgp,1);
402 for i=1:lgp
403     % taw [mm(water)/m(soildepth)]
404     tawc(i,1)=awc_final(m,n).*rd(i,1);
405 end
406
407 % calculate TAWC in the IRRIGATED scenario
408 tawc_irrigated=zeros(lgp_irr,1);
409 for i=1:lgp_irr
410     %[mm(water)/m(soildepth)]
411     tawc_irrigated(i,1)=awc_final(m,n).*rd_irrigated(i,1);
412 end
413
414 % definition of the deplietion fraction vector
415 % values from FAO 56
416 % this parameter was considered constant throughtout the whole growing
417 % period
418
419 % Rainfed
420 f=zeros(lgp,1);
421 f(1:lgp)=depl_fraction;
422
423 % Irrigated
424 f_irr=zeros(lgp_irr,1);
425 f_irr(1:lgp_irr)=depl_fraction;
426
```

```
427 % estiamtion of RAWC in the RAINFED scenario
428 % rawc=tawc*fraction
429 rawc=zeros(lgp,1);
430 for t=1:lgp
431         rawc(t,1)=tawc(t,1).*f(t);
432 end
433
434 % estiamtion of RAWC in the IRRIGATED scenario
435 rawc_irrigated=zeros(lgp_irr,1);
436 for t=1:lgp_irr
437         rawc_irrigated(t,1)=tawc_irrigated(t,1).*f_irr(t);
438 end
439
440
441 % initialization of hydrological balance matrices
442 deficit_start=zeros(lgp,1);
443 deficit_start_i=zeros(lgp_irr,1);           % I balance
444 deficit_st_irrigated = zeros(lgp_irr,1);    % II balance
445
446 deficit_end=zeros(lgp,1);
447 deficit_end_i=zeros(lgp_irr,1);             % I balance
448 deficit_end_irrigated = zeros(lgp_irr,1);   % II balance
449
450 surplus=zeros(lgp,1);
451 surplus_i=zeros(lgp_irr,1);
452
453 ETc_daily=zeros(lgp,1);
454 ETc_daily_irr=zeros(lgp_irr,1);
455
456 % rainfed
457 ETa_daily=zeros(lgp,1);
458
459 % irrigated
460 ETgreen_daily=zeros(lgp_irr,1);
461 CWU_daily=zeros(lgp_irr,1);
462 ETblue_daily=zeros(lgp_irr,1);
463
464 ks_rain=zeros(lgp,1);
465 ks=zeros(lgp_irr,1);
466 ks_irrigated = zeros(lgp_irr,1);
467
468 Pre_eff_daily=zeros(lgp,1);
469 Pre_eff_daily_i=zeros(lgp_irr,1);
470 Pre_tot_daily_growing_season = zeros(lgp,1);
471 Pre_tot_daily_growing_season_IR = zeros(lgp_irr,1);
472
473 I=zeros(lgp_irr,1);
474 % Initialization of vectors for daily ETO during the crop growing phases.
475 ETO_rf=zeros(lgp,1);
```

```
476 ET0_ir=zeros(lgp_irr,1);
477
478 ETgreen_year=zeros(ndays,1);
479 ETblue_year=zeros(ndays,1);
480
481 %% Hydrological balances
482
483 % RAINFED
484 % For each cell estimation of daily ETc, ETa, ETgreen, CWU and ETblue.
485 % Sum each day to have the total corresponding to the whole growing period.
486   if area_rain(m,n)>0 && day_plant_modified(m,n)>0
487         day_start=day_plant_modified(m,n);
488         day_rain=day_plant_modified(m,n)+1; % 2nd day of the gs
489
490         ETc_daily(1,1)=kc_crop(1,1)*ET0_daily(day_start,1);
491
492         % HYDROLOGICAL BALANCE to compute ks
493         deficit_start(1,1)=0;
494         ETa_daily(1,1)=ETc_daily(1,1);
495         deficit_end(1,1)=ETa_daily(1,1)+deficit_start(1,1);
496         ks_rain(1,1)=1;
497         Pre_eff_daily(1,1)=Pre_tot_daily(day_start,1);
498         Pre_tot_daily_growing_season(1,1)= Pre_tot_daily(day_start,1);
499         surplus(1,1)=0;
500         ET0_rf(1,1)=ET0_daily(1,1);
501    % first balance rainfed
502
503         for i=2:lgp
504         ET0_rf(i,1)=ET0_daily(day_rain,1);
505         ETc_daily(i,1)=kc_crop(i,1).*ET0_daily(day_rain,1);
506         Pre_tot_daily_growing_season(i,1)= Pre_tot_daily(day_rain,1);
507
508
509             if deficit_end(i-1,1)-Pre_tot_daily(day_rain,1)<0
510                 deficit_start(i,1)=0;
511                 surplus(i,1)=Pre_tot_daily(day_rain,1)-deficit_end(i-1,1);
512                 % surplus is runoff, what does not infiltrate
513             else
514             deficit_start(i,1)=deficit_end(i-1,1)-Pre_tot_daily(day_rain,1);
515             surplus(i,1)=0;
516             end
517
518             if Pre_tot_daily(day_rain,1)>0
519                 Pre_eff_daily(i,1)=Pre_tot_daily(day_rain,1)-surplus(i,1);
520             else
521                 Pre_eff_daily(i,1)=0;
522             end
523
524
```

```
525             if deficit_start(i,1)>=rawc(i,1)
526                 ks_rain(i,1)=(tawc(i,1)-deficit_start(i,1))/
527                 (tawc(i,1)-rawc(i,1));
528             else    ks_rain(i,1)=1;
529             end
530
531             if ks_rain(i,1)>1
532                 ks_rain(i,1)=1;
533             end
534
535             if  ks_rain(i,1)<0
536                 ks_rain(i,1)=0;
537             end
538
539             ETa_daily(i,1)=ks_rain(i,1)*ETc_daily(i,1);
540             deficit_end(i,1)=deficit_start(i,1)+ETa_daily(i,1);
541
542             day_rain=day_rain+1;
543
544             if day_rain>=ndays      % last day of the year
545                     day_rain=1;     % first day of the year
546             end
547
548
549         end
550
551
552         ETa_tot(m,n) = ETa_tot(m,n) + sum(ETa_daily,'omitnan');
553
554         ETc_tot_rain(m,n) = ETc_tot_rain(m,n) + sum(ETc_daily,'omitnan');
555         ET0_tot_rf(m,n) = ET0_tot_rf(m,n) + sum(ET0_rf,'omitnan');
556         Ptot_rf(m,n) = Ptot_rf(m,n) + sum(Pre_eff_daily,'omitnan');
557         Ptot_gr_seas_rf(m,n) = Ptot_gr_seas_rf(m,n) +
558             sum(Pre_tot_daily_growing_season,'omitnan');
559
560
561   end
562
563
564 %% Cell Insights Plots RAINFED
565
566 if isempty(lat) == 0 && isempty(lon) == 0
567
568     % coefficients
569     figure (1)
570     plot(1:1:lgp,kc_crop,"LineWidth",2)
571     hold on
572     plot(1:1:lgp,ks_rain,"LineWidth",2)
573     grid on
```

```
574     xlabel({'Growing period [days]', },'Interpreter','latex')
575     xlim([1 lgp])
576     ylabel('[-]','Interpreter','latex')
577     ylim([0 1.2])
578     title(['(a) Coefficients RAINFED - ','Lat: ',num2str(lat(1)),
579         ' Lon: ',num2str(lon(1)), ' - Year ', num2str(y)],
580         'Interpreter','latex')
581     legend('Crop Coefficient $k_c$','Stress Coefficient $k_s$',
582         'Location','sw','Interpreter','latex')
583     set(gcf, 'Units', 'centimeters', 'Position', [1 1 16 10]);
584     ax = gca;
585     ax.TickLabelInterpreter = 'latex';
586     print(gcf, 'coefficients_rf.svg', '-dsvg', '-r300');
587     hold off
588
589     % rooting depth
590     figure (2)
591     plot(1:1:lgp,-rd*10^3,'Color','#7E2F8E',"LineWidth",2,
592         'DisplayName','Rooting depth ($Z_r$)')
593     hold on
594     plot(1:1:lgp,-rawc,'Color','#EE510E',"LineWidth",2,
595         'DisplayName','Readily Available soil Water ($RAW$ or $\theta*$)')
596     plot(1:1:lgp,-tawc,'Color','#A2142F',"LineWidth",2,
597         'DisplayName','Total Available soil Water ($TAW$ or $\theta_{WP}$)')
598     grid on
599     xlabel('Growing period [days]','Interpreter','latex')
600     xlim([1 lgp])
601     ylabel('[mm]','Interpreter','latex')
602     ylim([-2000 0])
603     title(['(c) Rooting depth RAINFED - ','Lat: ',num2str(lat(1)),
604         ' Lon: ',num2str(lon(1)), ' - Year ', num2str(y)],
605         'Interpreter','latex')
606     legend('Location','southoutside','Interpreter','latex')
607     set(gcf, 'Units', 'centimeters', 'Position', [1 1 16 10]);
608     ax = gca;
609     ax.TickLabelInterpreter = 'latex';
610     print(gcf, 'root_rf.svg', '-dsvg', '-r300');
611     hold off
612
613     % deficit
614     figure (3)
615     plot(1:1:lgp,-deficit_start,'Color','#EDB120',"LineWidth",2,
616         'DisplayName','Depletion at the beginning of the day')
617     hold on
618     plot(1:1:lgp,-deficit_end,'Color','#006B08',"LineWidth",2,
619         'DisplayName','Depletion at the end of the day')
620     plot(1:1:lgp,-rawc,'Color','#EE510E',"LineWidth",2,
621         'DisplayName','Readily Available soil Water ($RAW$ or $\theta*$)')
622     plot(1:1:lgp,-tawc,'Color','#A2142F',"LineWidth",2,
```

```
623         'DisplayName','Total Available soil Water ($TAW$ or $\theta_{WP}$)')
624     grid on
625     xlabel('Growing period [days]','Interpreter','latex')
626     xlim([1 lgp])
627     ylabel('[mm]','Interpreter','latex')
628     ylim([-300 0])
629     title(['(e) Hydrological balance in the soil RAINFED - ',
630         'Lat: ',num2str(lat(1)), ' Lon: ',num2str(lon(1)), ' - Year ',
631         num2str(y)],'Interpreter','latex')
632     legend('Location','southoutside','Interpreter','latex')
633     set(gcf, 'Units', 'centimeters', 'Position', [1 1 16 10]);
634     ax = gca;
635     ax.TickLabelInterpreter = 'latex';
636     print(gcf, 'depletion_rf.svg', '-dsvg', '-r300');
637     hold off
638
639     figure (4)
640     plot(1:1:lgp,Pre_eff_daily,'Color', '#006B08',"LineWidth",2,
641         'DisplayName','Effective Daily Precipitation $P$') % dark green
642     hold on
643     plot(1:1:lgp,ET0_rf, 'Color','#A2142F',"LineWidth",2,
644         'DisplayName','Potential Evapotranspiration $ET_0$')
645     plot(1:1:lgp,ETc_daily, 'Color','#EE510E',"LineWidth",2,
646         'DisplayName','Crop Evapotranspiration $ET_c$')
647     plot(1:1:lgp,ETa_daily, 'Color','#EDB120',"LineWidth",2,
648         'DisplayName','Actual Evapotranspiration $ET_a$')
649     fill([1:1:lgp lgp:-1:1],[(ETa_daily); zeros(length(ETa_daily),1)],
650         [0.4660 0.6740 0.1880],'FaceAlpha',0.3,
651         'EdgeColor',"none",'DisplayName','$ET_{green}$')
652     grid on
653     xlabel('Growing period [days]','Interpreter','latex')
654     xlim([1 lgp])
655     ylabel('[mm/day]','Interpreter','latex')
656     ylim([0 9])
657     title(['(g) Hydrological balance in the atmosphere RAINFED - ',
658         'Lat: ',num2str(lat(1)),'Lon: ',num2str(lon(1)), ' - Year ',
659         num2str(y)],'Interpreter','latex')
660     legend('Location','nw','Interpreter','latex')
661     set(gcf, 'Units', 'centimeters', 'Position', [1 1 16 10]);
662     ax = gca;
663     ax.TickLabelInterpreter = 'latex';
664     print(gcf, 'balance_rf.svg', '-dsvg', '-r300');
665     hold off
666
667 end
668
669  %IRRIGATED
670
671   if area_irr (m,n)>0 && day_plant_modified_irr(m,n)>0
```

```
672
673         day_start_irr=day_plant_modified_irr(m,n);
674         day_irr=day_plant_modified_irr(m,n)+1;
675
676         deficit_start_i(1,1)=0;
677         ks(1,1)=1;
678         ks_irrigated(1,1)=1;
679         ETc_daily_irr(1,1)=kc_crop_irr(1,1).*ET0_daily(day_start_irr,1);
680
681         Pre_eff_daily_i(1,1)=Pre_tot_daily(day_start_irr,1);
682         Pre_tot_daily_growing_season_IR(1,1)=Pre_tot_daily(day_start_irr,1);
683
684         surplus_i(1,1)=0;
685         CWU_daily(1,1)=ETc_daily_irr(1,1); % no irrigation on the first day
686         ETgreen_daily(1,1)=CWU_daily(1,1); % only green ET
687         deficit_end_i(1,1)=CWU_daily(1,1)+deficit_start_i(1,1);
688         ETblue_daily(1,1)=0;
689         I(1,1) = 0;
690         deficit_st_irrigated(1,1) = 0;
691         deficit_end_irrigated(1,1) =
692             deficit_st_irrigated(1,1)+CWU_daily(1,1)-I(1,1);
693
694         ET0_ir(1,1)=ET0_daily(day_start_irr,1);
695
696         for i=2:lgp_irr
697         ET0_ir(i,1)=ET0_daily(day_irr,1);
698         ETc_daily_irr(i,1)=kc_crop_irr(i,1).*ET0_daily(day_irr,1);
699         Pre_tot_daily_growing_season_IR(i,1)= Pre_tot_daily(day_irr,1);
700
701             if deficit_end_i(i-1,1)-Pre_tot_daily(day_irr,1)<0
702                 deficit_start_i(i,1)=0;
703                 surplus_i(i,1)=Pre_tot_daily(day_irr,1)-
704                     deficit_end_i(i-1,1);
705                Pre_eff_daily_i(i,1)=deficit_end_i(i-1,1);
706             else
707                 deficit_start_i(i,1)=deficit_end_i(i-1,1)-
708                     Pre_tot_daily(day_irr,1);
709                 surplus_i(i,1)=0;
710                 Pre_eff_daily_i(i,1)=Pre_tot_daily(day_irr,1);
711             end
712
713             if deficit_start_i(i,1)>=rawc_irrigated(i,1)
714                 % water stress, so irrigation
715                 ks(i,1)=(tawc_irrigated(i,1)-deficit_start_i(i,1))/
716                     (tawc_irrigated(i,1)-rawc_irrigated(i,1));
717
718             else
719                 ks(i,1)=1;
720             end
```

```
721
722
723             if ks(i,1)>1
724                 ks(i,1)=1;
725             end
726
727             if  ks(i,1)<0
728                 ks(i,1)=0;
729             end
730             ETgreen_daily(i,1)=ks(i,1)*ETc_daily_irr(i,1);
731             deficit_end_i(i,1)=deficit_start_i(i,1)+ETgreen_daily(i,1);
732
733 %  balance with irrigation
734 deficit_st_irrigated(i,1)=deficit_end_irrigated(i-1,1)
735     -Pre_tot_daily(day_irr,1);
736 if deficit_st_irrigated(i,1)<0
737     deficit_st_irrigated(i,1)=0;
738 end
739
740 if deficit_st_irrigated(i,1)>=rawc_irrigated(i,1)
741
742
743     %-------------------------VERSIONE 1.1-------------------------------
744     % irrigation to have ks=1
745     I(i,1)=deficit_st_irrigated(i,1)-rawc_irrigated(i,1);
746     %--------------------------------------------------------------------
747
748 % ks_irrigated(i,1) = 1;
749 deficit_st_irrigated(i,1) = deficit_st_irrigated(i,1) - I(i,1);
750 ks_irrigated(i,1)=(tawc_irrigated(i,1)-deficit_st_irrigated(i,1))/
751     (tawc_irrigated(i,1)-rawc_irrigated(i,1));
752
753 else
754     I(i,1)=0;
755     ks_irrigated(i,1)=(tawc_irrigated(i,1)-deficit_st_irrigated(i,1))/
756         (tawc_irrigated(i,1)-rawc_irrigated(i,1));
757 end
758
759 if ks_irrigated(i,1)>1
760     ks_irrigated(i,1)=1;
761 end
762
763 if ks_irrigated(i,1)<0
764     ks_irrigated(i,1)=0;
765 end
766
767     %-------------------------VERSIONE 1.1-------------------------------
768         CWU_daily(i,1)=ks_irrigated(i,1)*ETc_daily_irr(i,1);
769         deficit_end_irrigated(i,1)=deficit_st_irrigated(i,1)+CWU_daily(i,1);
```

```
770         ETblue_daily(i,1)=CWU_daily(i,1)-ETgreen_daily(i,1);
771     %-------------------------------------------------------------------
772
773
774                 day_irr=day_irr+1;
775
776                 if day_irr>=ndays     % last day of the year
777                     day_irr=1;        % first day of the year
778                 end
779         end
780
781
782         ETgreen_tot(m,n) = ETgreen_tot(m,n) + sum(ETgreen_daily,'omitnan');
783         CWU_tot(m,n) = CWU_tot(m,n) + sum(CWU_daily,'omitnan');
784         ETblue_tot(m,n) = ETblue_tot(m,n) + sum(ETblue_daily,'omitnan');
785         I_tot(m,n) = I_tot(m,n) + sum(I,'omitnan');
786         ETc_tot_irr(m,n) = ETc_tot_irr(m,n) + sum(ETc_daily_irr,'omitnan');
787         ET0_tot_ir(m,n) = ET0_tot_ir(m,n) + sum(ET0_ir,'omitnan');
788         Ptot_ir(m,n) = Ptot_ir(m,n) + sum(Pre_eff_daily_i,'omitnan');
789         Ptot_gr_seas_ir(m,n) = Ptot_gr_seas_ir(m,n) +
790             sum(Pre_tot_daily_growing_season_IR,'omitnan');
791
792
793
794   end
795
796
797 %% Cell Insights Plots IRRIGATED
798
799 if isempty(lat) == 0 && isempty(lon) == 0
800
801     % coefficients
802     figure (5)
803     plot(1:1:lgp_irr,kc_crop_irr,"LineWidth",2)
804     hold on
805     plot(1:1:lgp_irr,ks_irrigated,"LineWidth",2)
806     plot(1:1:lgp_irr,ks,"LineWidth",2)
807     grid on
808     xlabel('Growing period [days]','Interpreter','latex')
809     xlim([1 lgp_irr])
810     ylabel('[-]','Interpreter','latex')
811     ylim([0 1.2])
812     title(['(b) Coefficients IRRIGATED - ','Lat: ',num2str(lat(1)),
813         'Lon: ',num2str(lon(1)), ' - Year ',
814         num2str(y)],'Interpreter','latex')
815     legend('Crop Coefficient $k_c$','Stress Coefficient $k_s$',
816         'Stress Coefficient $k_s$ without Irrigation',
817         'Location','sw','Interpreter','latex')
818     set(gcf, 'Units', 'centimeters', 'Position', [1 1 16 10]);
```

```
819     ax = gca;
820     ax.TickLabelInterpreter = 'latex';
821     print(gcf, 'coefficients_ir.svg', '-dsvg', '-r300');
822     hold off
823 
824     % rooting depth
825     figure (6)
826     plot(1:1:lgp_irr,-rd_irrigated*10^3,'Color','#7E2F8E',"LineWidth",2,
827         'DisplayName','Rooting depth')
828     hold on
829     plot(1:1:lgp_irr,-rawc_irrigated,'Color','#EE510E',"LineWidth",2,
830         'DisplayName','Readily Available soil Water ($RAW$ or $\theta*$)')
831     plot(1:1:lgp_irr,-tawc_irrigated,'Color','#A2142F',"LineWidth",2,
832         'DisplayName','Total Available soil Water ($TAW$ or $\theta_{WP}$)')
833     grid on
834     xlabel('Growing period [days]','Interpreter','latex')
835     xlim([1 lgp_irr])
836     ylabel('[mm]','Interpreter','latex')
837     ylim([-2000 0])
838     title(['(d) Rooting depth  IRRIGATED - ','Lat: ',num2str(lat(1)),'
839         Lon: ',num2str(lon(1)), ' - Year ',
840         num2str(y)],'Interpreter','latex')
841     legend('Location','southoutside','Interpreter','latex')
842     set(gcf, 'Units', 'centimeters', 'Position', [1 1 16 10]);
843     ax = gca;
844     ax.TickLabelInterpreter = 'latex';
845     print(gcf, 'root_ir.svg', '-dsvg', '-r300');
846     hold off
847 
848     % deficit
849     figure (7)
850     plot(1:1:lgp_irr,-deficit_st_irrigated,'Color','#EDB120',"LineWidth",2,
851         'DisplayName','Depletion at the beginning of the day')
852     hold on
853     plot(1:1:lgp_irr,-deficit_end_irrigated,'Color','#006B08',"LineWidth",2,
854         'DisplayName','Depletion at the end of the day')
855     plot(1:1:lgp_irr,-rawc_irrigated,'Color','#EE510E',"LineWidth",2,
856         'DisplayName','Readily Available soil Water ($RAW$ or $\theta*$)')
857     plot(1:1:lgp_irr,-tawc_irrigated,'Color','#A2142F',"LineWidth",2,
858         'DisplayName','Total Available soil Water ($TAW$ or $\theta_{WP}$)')
859     grid on
860     xlabel('Growing period [days]','Interpreter','latex')
861     xlim([1 lgp_irr])
862     ylabel('[mm]','Interpreter','latex')
863     ylim([-300 0])
864     %title(['(f) Hydrological balance in the soil IRRIGATED - ',
865         'Lat: ',num2str(lat(1)),' Lon: ',num2str(lon(1)), ' - Year ',
866         num2str(y)],'Interpreter','latex')
867     legend('Location','southoutside','Interpreter','latex')
```

```
868     set(gcf, 'Units', 'centimeters', 'Position', [1 1 16 10]);
869     ax = gca;
870     ax.TickLabelInterpreter = 'latex';
871     print(gcf, 'depletion_ir.svg', '-dsvg', '-r300');
872     hold off
873 
874     figure (8)
875     plot(1:1:lgp_irr,Pre_eff_daily_i,'color','#006B08',"LineWidth",2,
876         'DisplayName','Effective Daily Precipitation $P$')
877     hold on
878     plot(1:1:lgp_irr,ET0_ir,'color','#A2142F',"LineWidth",2,
879         'DisplayName','Potential Evapotranspiration $ET_0$')
880     plot(1:1:lgp_irr,ETc_daily_irr,'color','#EDB120',"LineWidth",2,'
881         DisplayName','Crop Evapotranspiration $ET_c (= ET_a)$')
882     fill([1:1:lgp_irr lgp_irr:-1:1],[(ETgreen_daily+ETblue_daily);
883         flip(ETblue_daily)], [0.4660 0.6740 0.1880],
884         'FaceAlpha',0.3,'EdgeColor',"none",
885         'DisplayName','$ET_{green}$')
886     fill([1:1:lgp_irr lgp_irr:-1:1],[ETblue_daily;
887         zeros(length(ETblue_daily),1)], [0 0.4470 0.7410],
888         'FaceAlpha',0.3,'EdgeColor',"none",
889         'DisplayName','$ET_{blue}$')
890     plot(1:1:lgp_irr,I,'color','#2D95DA',"LineWidth",2,
891     'DisplayName','Irrigation $I$')
892     grid on
893     xlabel('Growing period [days]','Interpreter','latex')
894     xlim([1 lgp_irr])
895     ylabel('[mm/day]','Interpreter','latex')
896     ylim([0 9])
897     %title(['(h) Hydrological balance in the atmosphere IRRIGATED - ',
898         'Lat: ',num2str(lat(1)),'Lon: ',num2str(lon(1)), ' - Year ',
899         num2str(y)],'Interpreter','latex')
900     legend('Location','nw','Interpreter','latex')
901     set(gcf, 'Units', 'centimeters', 'Position', [1 1 16 10]);
902     ax = gca;
903     ax.TickLabelInterpreter = 'latex';
904     print(gcf, 'balance_ir.svg', '-dsvg', '-r300');
905     hold off
906 
907 end
908 %% Annual Plot
909 
910 if isempty(lat) == 0 && isempty(lon) == 0
911 
912     % beginning of the hydrological year
913     if m<1080
914         % north hemisphere --> 1st of October (274th day of the year)
915         water_year = datetime(y, 10, 1);
916 
```

```
917     else
918         % south hemisphere --> 1st of July (182th day of the year)
919         water_year = datetime(y, 7, 1);
920     end
921
922     dates = [(water_year + caldays(0:ndays-1)) (water_year +
923         caldays(0:ndays-1))];
924
925     time_rain = day_plant_modified(m,n):1:(day_plant_modified(m,n)+lgp-1);
926     date_rain = dates(time_rain + ndays - day(water_year,'dayofyear'));
927     time_irr = day_plant_modified_irr(m,n):1:
928         (day_plant_modified_irr(m,n)+lgp_irr-1);
929     date_irr = dates(time_irr + ndays - day(water_year,'dayofyear'));
930
931
932     figure (9)
933
934     % if the growing period reach the end of the water year, it starts
935     % again from the beginning
936
937     % rainfed
938     if ndays - find(dates==date_rain(1),1) < lgp
939
940         % Split the dates and data at the end of the year
941         split_index = find(date_rain == dates(end));
942
943         % Data for the end of the year
944         date_rain_end = date_rain(1:split_index);
945         Pre_eff_daily_end = Pre_eff_daily(1:split_index);
946         ET0_rf_end = ET0_rf(1:split_index);
947         ETc_daily_end = ETc_daily(1:split_index);
948         ETa_daily_end = ETa_daily(1:split_index);
949
950         % Data for the beginning of the year
951         date_rain_start = date_rain(split_index+1:end);
952         Pre_eff_daily_start = Pre_eff_daily(split_index+1:end);
953         ET0_rf_start = ET0_rf(split_index+1:end);
954         ETc_daily_start = ETc_daily(split_index+1:end);
955         ETa_daily_start = ETa_daily(split_index+1:end);
956
957         % Plot end of the year data
958         plot(date_rain_end, Pre_eff_daily_end, 'Color', '#006B08',
959             "LineWidth", 2, 'DisplayName',
960             'RAINFED Effective Daily Precipitation P')
961         hold on
962         plot(date_rain_end, ET0_rf_end, 'Color', '#A2142F',
963             "LineWidth", 2, 'DisplayName',
964             'RAINFED Potential Evapotranspiration ET_0')
965         plot(date_rain_end, ETc_daily_end, 'Color', '#EE510E',
```

```
966             "LineWidth", 2, 'DisplayName',
967             'RAINFED Crop Evapotranspiration ET_c')
968         plot(date_rain_end, ETa_daily_end, 'Color', '#EDB120',
969             "LineWidth", 2, 'DisplayName',
970             'RAINFED Actual Evapotranspiration ET_a')
971         fill([date_rain_end flip(date_rain_end)], [ETa_daily_end;
972             zeros(length(ETa_daily_end), 1)]', [0.4660 0.6740 0.1880],
973             'FaceAlpha', 0.5, 'EdgeColor', "none",
974             'DisplayName', 'RAINFED ET green')
975
976         % Plot beginning of the year data
977         plot(date_rain_start, Pre_eff_daily_start, 'Color', '#006B08',
978             "LineWidth", 2, 'HandleVisibility', 'off')
979         plot(date_rain_start, ET0_rf_start, 'Color', '#A2142F',
980             "LineWidth", 2, 'HandleVisibility', 'off')
981         plot(date_rain_start, ETc_daily_start, 'Color', '#EE510E',
982             "LineWidth", 2, 'HandleVisibility', 'off')
983         plot(date_rain_start, ETa_daily_start, 'Color', '#EDB120',
984             "LineWidth", 2, 'HandleVisibility', 'off')
985         fill([date_rain_start flip(date_rain_start)],
986             [ETa_daily_start; zeros(length(ETa_daily_start), 1)]',
987             [0.4660 0.6740 0.1880], 'FaceAlpha', 0.5,
988             'EdgeColor', "none", 'HandleVisibility', 'off')
989
990     else
991
992         plot(date_rain,Pre_eff_daily,'Color', '#006B08',
993             "LineWidth",2,'DisplayName',
994             'RAINFED Effective Daily Precipitation P')
995         hold on
996         plot(date_rain,ET0_rf, 'Color','#A2142F',
997             "LineWidth",2,'DisplayName',
998             'RAINFED Potential Evapotranspiration ET_0')
999         plot(date_rain,ETc_daily, 'Color','#EE510E',
1000            "LineWidth",2,'DisplayName',
1001            'RAINFED Crop Evapotranspiration ET_c')
1002        plot(date_rain,ETa_daily, 'Color','#EDB120',
1003            "LineWidth",2,'DisplayName',
1004            'RAINFED Actual Evapotranspiration ET_a')
1005        fill([date_rain flip(date_rain)],[(ETa_daily);
1006            zeros(length(ETa_daily),1)], [0.4660 0.6740 0.1880],
1007            'FaceAlpha',0.5,'EdgeColor',"none",
1008            'DisplayName','RAINFED ET green')
1009
1010    end
1011
1012    % irrigated
1013    if ndays - find(dates==date_irr(1),1) < lgp_irr
1014
```

```
% Split the dates and data at the end of the year
split_index_irr = find(date_irr == dates(end));

% Data for the end of the year
date_irr_end = date_irr(1:split_index_irr);
Pre_eff_daily_i_end = Pre_eff_daily_i(1:split_index_irr);
ET0_ir_end = ET0_ir(1:split_index_irr);
ETc_daily_irr_end = ETc_daily_irr(1:split_index_irr);
I_end = I(1:split_index_irr);
ETgreen_daily_end = ETgreen_daily(1:split_index_irr);
ETblue_daily_end = ETblue_daily(1:split_index_irr);

% Data for the beginning of the year
date_irr_start = date_irr(split_index_irr+1:end);
Pre_eff_daily_i_start = Pre_eff_daily(split_index_irr+1:end);
ET0_ir_start = ET0_ir(split_index_irr+1:end);
ETc_daily_irr_start = ETc_daily_irr(split_index_irr+1:end);
I_start = I(split_index_irr+1:end);
ETgreen_daily_start = ETgreen_daily(split_index_irr+1:end);
ETblue_daily_start = ETblue_daily(split_index_irr+1:end);

% Plot end of the year data
plot(date_irr_end, Pre_eff_daily_i_end,':', 'Color', '#006B08',
    "LineWidth", 2, 'DisplayName',
    'IRRIGATED Effective Daily Precipitation P')
hold on
plot(date_irr_end, ET0_ir_end,':', 'Color', '#A2142F',
    "LineWidth", 2, 'DisplayName',
    'IRRIGATED Potential Evapotranspiration ET_0')
plot(date_irr_end, I,':', 'Color', '#2D95DA',
    "LineWidth", 2, 'DisplayName',
    'IRRIGATED Crop Evapotranspiration ET_c')
plot(date_irr_end, ETc_daily_irr_end,':', 'Color', '#EDB120',
    "LineWidth", 2, 'DisplayName',
    'IRRIGATED Actual Evapotranspiration ET_a')
fill([date_irr_end flip(date_irr_end)],
    [(ETgreen_daily_end+ETblue_daily_end); flip(ETblue_daily_end)],
    [0.4660 0.6740 0.1880], 'FaceAlpha',0.3,
    'EdgeColor',"none",'DisplayName','IRRIGATED ET green')
fill([date_irr_end flip(date_irr_end)],[ETblue_daily_end;
    zeros(length(ETblue_daily_end),1)], [0 0.4470 0.7410],
    'FaceAlpha',0.3, 'EdgeColor',"none",
    'DisplayName','IRRIGATED ET blue')

% Plot beginning of the year data
plot(date_irr_start, Pre_eff_daily_i_start,':', 'Color', '#006B08',
    "LineWidth", 2, 'HandleVisibility', 'off')
plot(date_irr_start, ET0_ir_start,':', 'Color', '#A2142F',
    "LineWidth", 2, 'HandleVisibility', 'off')
```

```
        plot(date_irr_start, I,':', 'Color', '#2D95DA',
            "LineWidth", 2, 'HandleVisibility', 'off')
        plot(date_irr_start, ETc_daily_irr_start,':', 'Color', '#EDB120',
            "LineWidth", 2, 'HandleVisibility', 'off')
        fill([date_irr_start flip(date_irr_start)],
            [(ETgreen_daily_start+ETblue_daily_start);
            flip(ETblue_daily_start)], [0.4660 0.6740 0.1880],
            'FaceAlpha',0.3,'EdgeColor',"none", 'HandleVisibility', 'off')
        fill([date_irr_start flip(date_irr_start)],
            [ETblue_daily_start; zeros(length(ETblue_daily_start),1)],
            [0 0.4470 0.7410],'FaceAlpha',0.3, 'EdgeColor',"none",
            'HandleVisibility', 'off')

    else

        plot(date_irr,Pre_eff_daily_i,':','color','#006B08',
            "LineWidth",2,'DisplayName',
            'IRRIGATED Effective Daily Precipitation P')
        plot(date_irr,ET0_ir,':','color','#A2142F',
            "LineWidth",2,'DisplayName',
            'IRRIGATED Potential Evapotranspiration ET_0')
        plot(date_irr,ETc_daily_irr,':','color','#EDB120',
            "LineWidth",2,'DisplayName',
            'IRRIGATED Crop Evapotranspiration ET_c (= ET_a)')
        fill([date_irr flip(date_irr)],[(ETgreen_daily+ETblue_daily);
            flip(ETblue_daily)], [0.4660 0.6740 0.1880],'FaceAlpha',0.3,
            'EdgeColor',"none",'DisplayName','IRRIGATED ET green')
        fill([date_irr flip(date_irr)],[ETblue_daily;
            zeros(length(ETblue_daily),1)], [0 0.4470 0.7410],
            'FaceAlpha',0.3,'EdgeColor',"none",
            'DisplayName','IRRIGATED ET blue')
        plot(date_irr,I,':','color','#2D95DA',"LineWidth",2,
            'DisplayName','Irrigation')

    end


    grid on
    xlabel(['Year ', num2str(y)],'Interpreter','latex')
    xlim([water_year water_year+years(1)-days(1)])
    xticks([1 31 61 91 121 151 181 211 241 271 301 331 360])
    xticklabels(tickmonths);
    ylabel('[mm/day]','Interpreter','latex')
    ylim([0 9])
    title(['(a) Daily $ETa$ - ','Lat: ',num2str(lat(1)),'
        Lon: ',num2str(lon(1))],'Interpreter','latex')
    legend('Location','north','Interpreter','latex')
    set(gcf, 'Units', 'centimeters', 'Position', [1 1 18 10]);
    ax = gca;
```

```
    ax.TickLabelInterpreter = 'latex';
    print(gcf, 'annual.svg', '-dsvg', '-r300');
    hold off


% Obtain daily year normalized values


    ETgreen_year(time_rain,1) = ETa_daily(1:end,1) *
        area_rain(m,n) / area_tot(m,n);
    ETgreen_year(time_irr,1) = ETgreen_year(time_irr,1) +
        (ETgreen_daily(1:end,1) * area_irr(m,n) / area_tot(m,n));
    ETblue_year(time_irr,1) = (ETblue_daily(1:end,1) *
        area_irr(m,n) / area_tot(m,n));

    last = water_year:datetime(y,12,31);
    first = datetime(y+1,1,1):(water_year+years(1)-1);

    % plot
    figure(10)

    fill([last flip(last)], [(ETgreen_year(day(water_year,'dayofyear')
        :end,1)+ETblue_year(day(water_year,'dayofyear'):end,1));
        flip(ETblue_year(day(water_year,'dayofyear'):end,1))],
        [0.4660 0.6740 0.1880], 'FaceAlpha',0.5,'EdgeColor',"none",
        'DisplayName','ET green')
    hold on
    fill([last flip(last)], [ETblue_year(day(water_year,'dayofyear')
        :end,1); zeros(ndays-day(water_year,'dayofyear')+1,1)],
        [0 0.4470 0.7410], 'FaceAlpha',0.5,'EdgeColor',"none",
        'DisplayName','ET blue')
    fill([first flip(first)], [(ETgreen_year(1:day(water_year,'dayofyear')
        -1,1)+ETblue_year(1:day(water_year,'dayofyear')-1,1));
        flip(ETblue_year(1:day(water_year,'dayofyear')-1,1))],
        [0.4660 0.6740 0.1880],'FaceAlpha',0.5,
        'EdgeColor',"none",'HandleVisibility', 'off')
    fill([first flip(first)], [ETblue_year(1:day(water_year,'dayofyear')
        -1,1); zeros(day(water_year,'dayofyear')-1,1)],
        [0 0.4470 0.7410],'FaceAlpha',0.5,
        'EdgeColor',"none",'HandleVisibility', 'off')
    grid on
    xlabel(['Year ', num2str(y)],'Interpreter','latex')
    xlim([water_year water_year+years(1)-days(1)])
    xticks([water_year:30:water_year+360-1])
    xticklabels(tickmonths);
    ylabel('[mm/day]','Interpreter','latex')
    ylim([0 9])
    title(['(b) Daily $ETa$ normalized with cultivated areas - ',
        'Lat: ',num2str(lat(1)),'Lon: ',num2str(lon(1))],
```

```
        'Interpreter','latex')
    legend('Location','north','Interpreter','latex')
    set(gcf, 'Units', 'centimeters', 'Position', [1 1 18 10]);
    ax = gca;
    ax.TickLabelInterpreter = 'latex';
    print(gcf, 'annual_vol.svg', '-dsvg', '-r300');
    hold off

end

        end
    end
end

clear day_plant_modified
clear day_plant_modified_irr
end

% Averaging on the number of years
% Since the variables were summed all together, to obtain the mean it
% is sufficient to divide them by the number of years.

% rainfed
ETa_tot = ETa_tot /yy;
ETc_tot_rain = ETc_tot_rain /yy;
ET0_tot_rf = ET0_tot_rf /yy;
Ptot_rf = Ptot_rf /yy;
Ptot_gr_seas_rf = Ptot_gr_seas_rf /yy;

% irrigated
ETgreen_tot = ETgreen_tot /yy;
CWU_tot = CWU_tot /yy;
ETblue_tot = ETblue_tot /yy;
I_tot = I_tot /yy;
ETc_tot_irr = ETc_tot_irr /yy;
ET0_tot_ir = ET0_tot_ir /yy;
Ptot_ir = Ptot_ir /yy;
Ptot_gr_seas_ir = Ptot_gr_seas_ir /yy;

% Final ET
VOL_GREEN = ETa_tot .* area_rain .* 10 + ETgreen_tot .* area_irr .* 10;
VOL_BLUE = ETblue_tot .* area_irr .* 10;
VOL_TOT = VOL_GREEN + VOL_BLUE;

%% Download ET results
mkdir(cartella_risultati)
cd(cartella_risultati)
copyfile("path/to/nc_from_watercrop.m",
    fullfile(cartella_risultati,"nc_from_watercrop.m"));
```

```
copyfile("path/to/txt_per_QGis.m",
    fullfile(cartella_risultati,"txt_per_QGis.m"));

save('ETa_rain.mat','ETa_tot')

save('VOL_GREEN.mat','VOL_GREEN')

% Save into txt files to open on QGIS
txt_per_QGis(ETa_tot,'ETa_rain','-9999','0.0833333','2')
txt_per_QGis(VOL_GREEN,'VOL_GREEN','-9999','0.0833333','2')
% nan, cell size, digits

% Save into NetCDF files
nc_from_watercrop(VOL_GREEN,'VOL_GREEN',-9999, 0.0833333,"m^3")


end

elapsed_time = toc          % seconds
```

A.1 From waterCrop results to NetCDF files

```
function nc_from_watercrop(var, name, nodata, cellsize, unit)
    % var: 2D matrix
    % name: base filename, add as a string (without extension)
    % nodata: missing value, add as a number (e.g., -9999)
    % cellsize: grid resolution, add as a number (e.g., 0.0833333 degrees)
    % unit: string with unit

    [nrows, ncols] = size(var);

    % Define lower-left corner
    xllcorner = -180;
    yllcorner = -90;

    % Compute 1D coordinate vectors (cell centers)
    lon = xllcorner + (0:ncols-1) * cellsize + cellsize/2;
    lat = yllcorner + (0:nrows-1) * cellsize + cellsize/2;

    % Flip data to match lat orientation
    var = flipud(var);

    % Create NetCDF
    ncFile = [name, '.nc'];
    ncid = netcdf.create(ncFile, 'CLOBBER');

    % Define dimensions
```

```
    dimid_lon = netcdf.defDim(ncid, 'lon', ncols);
    dimid_lat = netcdf.defDim(ncid, 'lat', nrows);

    % Define coordinate variables
    lon_id = netcdf.defVar(ncid, 'lon', 'double', dimid_lon);
    lat_id = netcdf.defVar(ncid, 'lat', 'double', dimid_lat);

    % Add attributes to coordinates
    netcdf.putAtt(ncid, lon_id, 'standard_name', 'longitude');
    netcdf.putAtt(ncid, lon_id, 'units', 'degrees_east');

    netcdf.putAtt(ncid, lat_id, 'standard_name', 'latitude');
    netcdf.putAtt(ncid, lat_id, 'units', 'degrees_north');

    % Define CRS variable
    crs_id = netcdf.defVar(ncid, 'crs', 'int', []);
    netcdf.putAtt(ncid, crs_id, 'grid_mapping_name', 'latitude_longitude');
    netcdf.putAtt(ncid, crs_id, 'epsg_code', int32(4326));
    netcdf.putAtt(ncid, crs_id, 'semi_major_axis', 6378137.0);
    netcdf.putAtt(ncid, crs_id, 'inverse_flattening', 298.257223563);

    % Define main data variable
    var_id = netcdf.defVar(ncid, name, 'double', [dimid_lon, dimid_lat]);
    netcdf.defVarFill(ncid, var_id, false, nodata);

    netcdf.putAtt(ncid, var_id, 'long_name', name);
    netcdf.putAtt(ncid, var_id, 'units', unit);
    netcdf.putAtt(ncid, var_id, '_FillValue', nodata);
    netcdf.putAtt(ncid, var_id, 'coordinates', 'lon lat');
    netcdf.putAtt(ncid, var_id, 'grid_mapping', 'crs');

    % Global attributes
    netcdf.putAtt(ncid, netcdf.getConstant('NC_GLOBAL'),
        'title', [name ' raster']);
    netcdf.putAtt(ncid, netcdf.getConstant('NC_GLOBAL'),
        'Conventions', 'CF-1.6');

    netcdf.endDef(ncid);

    % Write coordinate data
    netcdf.putVar(ncid, lon_id, lon);
    netcdf.putVar(ncid, lat_id, lat);
    netcdf.putVar(ncid, var_id, var');

    netcdf.close(ncid);
    fprintf('Wrote CF-compliant geo2d NetCDF: %s\n', ncFile);
end
```

Appendix B

Cumulative ET sheds code

```
import numpy as np
import xarray as xr
import matplotlib.pyplot as plt
import cartopy.crs as ccrs
import cartopy.feature as cfeature
from netCDF4 import Dataset
import os
import subprocess

## GETTING READY

# Input Configuration
crop = "wheat"

# Input files
# Everything must have the same spatial resolution
# and the same reference system
moist_input_file = "RECON_moisture_flows_0.5.nc"
ERA5_eta_input_file = "RECON_ERA5_avgYear_0.5_volumes.nc"

# waterCrop results files upscaled
WC_etg_input_file = "VOL_GREEN_sum.nc"
WC_etb_input_file = "VOL_BLUE_sum.nc"


# Case study coordinates in a system lon=0,360 lat=90,-90

#cs_lat = np.array([44.76])      #cell in Piedmont
#cs_lon = np.array([7.51])

#cs_lat = np.array([-16.17])     #cell in Minas Gerais
#cs_lon = np.array([313.42])

#cs_lat = np.array([])
```

```
#cs_lon = np.array([])

# To run the global analysis assign np.nan to the case study coordinates
cs_lat = np.nan
cs_lon = np.nan


# Constants
ymax = 122079329.40990189
ymin = 10**-3

# Load dataset
dataset = xr.open_dataset(moist_input_file)
dataset_ERA5_eta = xr.open_dataset(ERA5_eta_input_file)
dataset_etg = xr.open_dataset(WC_etg_input_file)
dataset_etb = xr.open_dataset(WC_etb_input_file)

# Extract lat/lon values from RECON
sinklats = dataset["sinklat"].values
sinklons = dataset["sinklon"].values
sourcelats = dataset["sourcelat"].values
sourcelons = dataset["sourcelon"].values

# Matrix initialization
ROWS = dataset["sourcelat"].shape[0]    # dataset dimensions
COLS = dataset["sourcelon"].shape[0]
# those will be the cumulative evaporation shed for the crop
ET_shed_g = np.zeros((ROWS, COLS), dtype=np.float128)
ET_shed_b = np.zeros((ROWS, COLS), dtype=np.float128)
ET_shed = np.zeros((ROWS, COLS), dtype=np.float128)

# Case study cell index
if not np.any(np.isnan(cs_lat)) or not np.any(np.isnan(cs_lon)):
    # Compute coordinates element-wise
    ind_lat = -(ROWS/180) * cs_lat + ROWS/2
    ind_lon =  (COLS/360) * cs_lon
    cs = np.round(np.array([ind_lat,ind_lon])).astype(int)


############################################################

## PART 1 - OBTAIN FRACTION MATRIX OF ETa

# CHANGE REFERENCE SYSTEM OF WC RESULTS

# Flip latitude (reverse the lat coordinate)
# In the .nc file from WC the lat range is -90,90 deg
dataset_etb = dataset_etb.reindex(lat=dataset_etb.lat[::-1])
dataset_etg = dataset_etg.reindex(lat=dataset_etg.lat[::-1])
```

```
84
85  # Reshape the matrix from waterCrop to have the origin in lon=0deg
86  # Important to have it consistent with ERA5 matrix
87  # green
88  temp = np.zeros((ROWS, COLS))
89  temp = dataset_etg['VOL_GREEN'].values      # m^3
90  ETg_fromWC = np.zeros((ROWS, COLS))
91  ETg_fromWC[:,0:360] = temp[:,360:720]
92  ETg_fromWC[:,360:720] = temp[:,0:360]
93  # blue
94  temp = np.zeros((ROWS, COLS))
95  temp = dataset_etb['VOL_BLUE'].values       # m^3
96  ETb_fromWC = np.zeros((ROWS, COLS))
97  ETb_fromWC[:,0:360] = temp[:,360:720]
98  ETb_fromWC[:,360:720] = temp[:,0:360]
99
100 # Extract ERA5 values
101 ETa_fromERA5 = dataset_ERA5_eta['ERA5_ET_averageyear'].values   # m^3
102
103
104 # Calculating the % of that evaporation which then
105 # contribute to Green ET of the specific crop
106 # ETgreen fraction of the total ETa of each cell
107 ETa_gweight = ETg_fromWC / ETa_fromERA5
108 # ETblue fraction of the total ETa of each cell
109 ETa_bweight = ETb_fromWC / ETa_fromERA5
110
111 ############################################################
112
113 ## PART 2 - LOOPS TO OBTAIN CUMULATIVE SHEDS
114
115 # Get the indices of non-zero elements if no case study
116 if np.any(np.isnan(cs_lat)) or np.any(np.isnan(cs_lon)):
117     green_no_zero = np.argwhere(ETa_gweight != 0)
118     blue_no_zero = np.argwhere(ETa_bweight != 0)
119 else:
120     green_no_zero = np.array([1])
121     blue_no_zero = np.array([1])
122
123 # Process each non-zero element (or only the case study)
124
125 # 1) GREEN ET SHED
126 for index in green_no_zero:
127     if np.any(np.isnan(cs_lat)) or np.any(np.isnan(cs_lon)):
128         i, j = index  # Index pair for row and column
129     else:
130         # case study coordinates
131         i = cs[0]
132         j = cs[1]
```

```
133
134     # Get moisture flow
135     ms_4d = dataset["moisture_flow"].isel(sourcelat=i,sourcelon=j).values
136     # reshape moisture flow array
137     moisture_flow = np.squeeze(ms_4d)
138
139     # Convert to cubic meters
140     evaporation_shed = np.where(
141         moisture_flow == 0,
142         0,
143         10 ** (((moisture_flow - 1) / 254) *
144                (np.log10(ymax) - np.log10(ymin)) + np.log10(ymin))
145     )
146
147     # To obtain the cumulative shed
148     ET_shed_g = ET_shed_g + evaporation_shed*ETa_gweight[i,j]
149
150 print("Green loop completed.")
151
152 # 2) BLUE ET SHED
153 for index in blue_no_zero:
154     if np.any(np.isnan(cs_lat)) or np.any(np.isnan(cs_lon)):
155         i, j = index  # Index pair for row and column
156     else:
157         i = cs[0]
158         j = cs[1]
159
160     # Get moisture flow
161     ms_4d = dataset["moisture_flow"].isel(sourcelat=i,sourcelon=j).values
162     moisture_flow = np.squeeze(ms_4d)
163
164     # Convert to cubic meters
165     evaporation_shed = np.where(
166         moisture_flow == 0,
167         0,
168         10 ** (((moisture_flow - 1) / 254) *
169                (np.log10(ymax) - np.log10(ymin)) + np.log10(ymin))
170     )
171
172     # To obtain the cumulative shed
173     ET_shed_b = ET_shed_b + evaporation_shed*ETa_bweight[i,j]
174
175 print("Blue loop completed.")
176
177 # Total ET SHED
178 ET_shed = ET_shed_g + ET_shed_b
179
180 ##########################################################
181
```

```
## PART 3 - SAVE RESULTS

# Folder where you want to save the new NetCDF files
if np.any(np.isnan(cs_lat)) or np.any(np.isnan(cs_lon)):
    output_folder = f'/RECON/output/{crop}/'
else:
    output_folder = f'/RECON/output/{crop}_{cs_lat}_{cs_lon}/'

# Create the output directory if it does not exist
os.makedirs(output_folder, exist_ok=True)

# Open the RECON NetCDF file to read grid dimensions
with Dataset(ERA5_eta_input_file, 'r') as ERA5_nc:
    # Read grid dimensions (latitude & longitude)
    latitudes = ERA5_nc.variables['lat'][:]
    longitudes = ERA5_nc.variables['lon'][:]
    # Number of latitude points
    n_lat = len(latitudes)
    # Number of longitude points
    n_lon = len(longitudes)

# Specify results to save and the output file names
results = [ET_shed_g, ET_shed_b, ET_shed]
file_names_final = ['ET_shed_g.nc', 'ET_shed_b.nc', 'ET_shed.nc']

# Transform the results array of arrays in a single 3d array
results_3d = np.stack(results, axis=0)
# Change again the system to lon -180,180
results_180_3d = np.zeros((len(results),ROWS,COLS), dtype=np.float128)
results_180_3d[:,:,0:360] = results_3d[:,:,360:720]
results_180_3d[:,:,360:720] = results_3d[:,:,0:360]
# Reshape the 3d array in an array of arrays
results_180 = [results_180_3d[i] for i in range(results_180_3d.shape[0])]

# Loop through each matrix shifted and save it to a
# NetCDF file with the -180,180 reference system
for k, (matrix, file_name) in enumerate(zip(results_180, file_names_final)):
    full_path = os.path.join(output_folder, file_name)

    with Dataset(full_path, 'w', format='NETCDF4') as new_nc:
        # Create dimensions
        new_nc.createDimension('lat', n_lat)
        new_nc.createDimension('lon', n_lon)

        # Create coordinate variables
        lat_var = new_nc.createVariable('lat', latitudes.dtype, ('lat',))
        lon_var = new_nc.createVariable('lon', longitudes.dtype, ('lon',))

        # Write coordinates data
```

```
231         lat_var[:] = latitudes
232         lon_var[:] = np.ma.arange(-180,180,0.5)
233 
234 
235         # Add CF-compliant attributes to coordinate variables
236         lat_var.units = 'degrees_north'
237         lat_var.standard_name = 'latitude'
238         lat_var.long_name = 'Latitude'
239 
240         lon_var.units = 'degrees_east'
241         lon_var.standard_name = 'longitude'
242         lon_var.long_name = 'Longitude'
243 
244         # Create data variable
245         data_var = new_nc.createVariable('volume',np.float64,('lat','lon'))
246         data_var[:, :] = matrix
247 
248         # Add attributes to data variable
249         data_var.units = 'm^3'
250         data_var.description = f'Cumulative {file_name} for {crop}'
251 
252     print(f"Cumulative shed {k+1} saved to {file_name} with WGS84 grid.")
253 
254 
255 ############################################################
256 
257 ## PART 4 - PLOTS
258 
259 # Define lat/lon grid
260 lats = np.arange(90, -90, -0.5)
261 lons = np.arange(0, 360, 0.5)
262 
263 # Colormap with NaN as white
264 blues_cmap = plt.cm.get_cmap("Blues").copy()
265 blues_cmap.set_bad(color='white')
266 reds_cmap = plt.cm.get_cmap("Reds").copy()
267 reds_cmap.set_bad(color='white')
268 
269 # Actual plots
270 
271 # ET-SHED PLOT
272 plt.figure(figsize=(10, 8))
273 ax = plt.axes(projection=ccrs.PlateCarree())
274 
275 vmin = np.percentile(ET_shed[ET_shed > 0], 10)
276 vmax = np.nanmax(ET_shed)
277 
278 # convert to float to support np.nan
279 ET_shed_nan = ET_shed.astype(float)
```

```
# Set values below 10th percentile to NaN
ET_shed_nan[ET_shed_nan < vmin] = np.nan


im = ax.imshow(
    ET_shed_nan,
    extent=[lons.min(), lons.max(), lats.min(), lats.max()],
    origin='upper',
    vmin=vmin,
    vmax=vmax,
    cmap=reds_cmap,
    transform=ccrs.PlateCarree()
)

# Add features
ax.coastlines()
ax.add_feature(cfeature.BORDERS, linestyle=':')
if not np.any(np.isnan(cs_lat)) or not np.any(np.isnan(cs_lon)):
    ax.plot(cs_lon, cs_lat, marker='o', color='red', markersize=3,
            transform=ccrs.PlateCarree(), label='Source Location')


# Title and colorbar
ax.set_title(f"Annual precipitation originating from {crop} ET")
plt.colorbar(im, label="moisture flow [m$^3$]", orientation="vertical")
plt.xlabel("Longitude")
plt.ylabel("Latitude")
ax.legend(loc="lower left")

# Save plot to specified folder as SVG
plt.savefig(os.path.join(output_folder, 'ET_shed_plot.svg'),
            format='svg', bbox_inches='tight')

plt.show()
```

Appendix C

Sinks land-use classification code

```
import numpy as np
import xarray as xr
import matplotlib.pyplot as plt
import cartopy.crs as ccrs
import cartopy.feature as cfeature
from netCDF4 import Dataset
import os
import plotly.graph_objects as go
import pandas as pd



## GETTING READY

# Input Configuration
cropWC = "soy"       # the one analysed
crops_tot = ["maize", "wheat", "soy"]

# Input files
input_files = {
    "land": "landseamask_water-global.nc",
    "maize_rf": "maize_rf_05deg.nc",
    "wheat_rf": "wheat_rf_05deg.nc",
    "soy_rf":   "soybean_rf_05deg.nc",
    "maize_ir": "maize_ir_05deg.nc",
    "wheat_ir": "wheat_ir_05deg.nc",
    "soy_ir":   "soybean_ir_05deg.nc",
    "ET_shed_g": f"/RECON/output/{cropWC}/ET_shed_g.nc",
    "ET_shed_b": f"/RECON/output/{cropWC}/ET_shed_b.nc"
}

```

```
32  # Folder where you want to save the new files
33  output_folder = f'/land_use_classification/output/{cropWC}/'
34  if not os.path.exists(output_folder):
35      os.makedirs(output_folder)
36
37
38  # Open all datasets in a loop and store in a dictionary
39  datasets = {}
40  for key, filepath in input_files.items():
41      datasets[key] = xr.open_dataset(filepath)
42
43
44  # Dataset dimensions (spatial resolution)
45  ROWS = datasets["ET_shed_g"]["lat"].shape[0]
46  COLS = datasets["ET_shed_g"]["lon"].shape[0]
47
48  # Load land mask to obtain the oceans
49  land = datasets["land"]["mask"].values
50  ocean = np.where(land == 1, 0, 1)
51  # Replace '1' with 0, because where there is land, there is 0 ocean
52  # Replace '-' with 1, because where there is no land, there is the ocean
53  fraction_ocean = ocean
54
55  # Load area arrays into lists, stack into np array (n_crops x ROWS x COLS)
56  area_rf_list, area_ir_list = [], []
57  for crop in crops_tot:
58      area_rf_list.append(datasets[f"{crop}_rf"]["area_rf"].values)
59      area_ir_list.append(datasets[f"{crop}_ir"]["area_ir"].values)
60
61  area_rf = np.array(area_rf_list)
62  area_ir = np.array(area_ir_list)
63
64  # Flip vertically (reverse rows) for all crops in area_rf and area_ir
65  area_rf = np.flip(area_rf, axis=1)
66  area_ir = np.flip(area_ir, axis=1)
67
68
69  # Open the cumaltive ET sheds
70  ET_shed_g = datasets["ET_shed_g"]["volume"].values
71  ET_shed_b = datasets["ET_shed_b"]["volume"].values
72  ET_shed = ET_shed_g + ET_shed_b
73
74  ############################################################
75
76  ## PART 1 - CALCULATE ha PER EACH CELL OF THE WORLD
77
78  # Earth radius (meters)
79  R = 6371e3
80
```

```
# Grid resolution in degrees
lat_res = 180 / ROWS   # latitude resolution (e.g., 0.5deg)
lon_res = 360 / COLS   # longitude resolution

# Latitude and longitude edges
lat_edges = np.linspace(90, -90, ROWS + 1)
lon_edges = np.linspace(-180, 180, COLS + 1)

# Preallocate area matrix (ROWS x COLS)
area_cell = np.zeros((ROWS, COLS))

# Loop over latitude bands (rows)
for i in range(ROWS):
    # Latitude edges in radians for this band
    lat1 = np.deg2rad(lat_edges[i])
    lat2 = np.deg2rad(lat_edges[i + 1])

    # Longitudinal width in radians
    dlon = np.deg2rad(lon_res)

    # Area of the cell (same for all longitudes at this latitude)
    # area in m^2
    cell_area_m2 = R**2 * dlon * (np.sin(lat1) - np.sin(lat2))
    # convert to ha
    cell_area_ha = cell_area_m2 / 1e4

    # Fill the entire row (all longitudes at this latitude)
    area_cell[i, :] = cell_area_ha


############################################################

## PART 2 - OBTAIN FRACTION MATRIX OF CULTIVATED AREAS

# Fraction matrices element-wise division with broadcasting:
fraction_rf = area_rf / area_cell     # shape: n_crops x ROWS x COLS
fraction_ir = area_ir / area_cell


############################################################

## PART 3 - WATER WHICH WILL FALL ON SPECIFIC CROPS

# Map crop water volumes:
map_crop_g_rf = ET_shed_g * fraction_rf   # broadcast over spatial dims
map_crop_b_rf = ET_shed_b * fraction_rf

map_crop_g_ir = ET_shed_g * fraction_ir
map_crop_b_ir = ET_shed_b * fraction_ir
```

```
130
131 # Map water that go into the ocean
132 map_ocean_g = ET_shed_g * fraction_ocean
133 map_ocean_b = ET_shed_b * fraction_ocean
134
135 ##########################################################
136
137 ## PART 4 - VOLUMES OF WATER PER CROP
138
139 # Total volumes of water involved
140 ETVOL_shed_g = np.nansum(ET_shed_g)
141 ETVOL_shed_b = np.nansum(ET_shed_b)
142 ETVOL_crop_g = [np.nansum(map_crop_g_rf, axis=(1,2)),
143                 np.nansum(map_crop_g_ir, axis=(1,2))]
144 ETVOL_crop_b = [np.nansum(map_crop_b_rf, axis=(1,2)),
145                 np.nansum(map_crop_b_ir, axis=(1,2))]
146 ETVOL_ocean = [np.nansum(map_ocean_g), np.nansum(map_ocean_b)]
147
148
149 ETVOL_other_land = [
150     ETVOL_shed_g - np.nansum(ETVOL_crop_g) - ETVOL_ocean[0],
151     ETVOL_shed_b - np.nansum(ETVOL_crop_b) - ETVOL_ocean[1]
152 ]
153
154
155 # Necessary to build the Sankey Diagram
156 labels_1 = [f'source crop: {cropWC}'] + ["ET GREEN", "ET BLUE"] +
157             [f"{c} rainfed" for c in crops_tot] +
158             [f"{c} irrigated" for c in crops_tot] +
159             ["ocean" , "other land"]
160 sources_1 = [0, 0,
161              1 * np.ones(2 * len(crops_tot) + 2),
162              2 * np.ones(2 * len(crops_tot) + 2)]
163 targets_1 = [1, 2,
164              np.arange(3, 2 * len(crops_tot) + 5),
165              np.arange(3, 2 * len(crops_tot) + 5)]
166 values_1 = [
167     ETVOL_shed_g,
168     ETVOL_shed_b,
169     ETVOL_crop_g,
170     ETVOL_ocean[0],
171     ETVOL_other_land[0],
172     ETVOL_crop_b,
173     ETVOL_ocean[1],
174     ETVOL_other_land[1]
175 ]
176
177 labels = []
178 sources = []
```

```
179 targets = []
180 flat_list = []
181 values = []
182
183 # Convert array to list and extend
184 for item in labels_1:
185     if isinstance(item, np.ndarray):
186         labels.extend(item.tolist())
187     else:
188         labels.append(item)
189
190 for item in sources_1:
191     if isinstance(item, np.ndarray):
192         sources.extend(item.tolist())
193     else:
194         sources.append(item)
195
196 for item in targets_1:
197     if isinstance(item, np.ndarray):
198         targets.extend(item.tolist())
199     else:
200         targets.append(item)
201
202 # Flatten numpy array into list and extend
203 for item in values_1:
204     if isinstance(item, np.ndarray):
205         values.extend(item.tolist())
206     elif isinstance(item, list):
207         values.extend(item)
208     else:
209         values.append(item)
210
211     if isinstance(item, (np.float64, float)):
212         flat_list.append(float(item))
213     elif isinstance(item, list):
214         for arr in item:
215             flat_list.extend(arr.flatten().tolist())
216     else:
217         raise TypeError(f"Unexpected type: {type(item)}")
218
219 # Convert to numpy array
220 values = np.array(flat_list)
221
222
223 sources_int = [int(x) for x in sources]
224
225 ############################################################
226
227 ## PART 5 - SAVE IN A .CSV FILE
```

```
228
229 # Create a DataFrame with each vector as a column
230 df = pd.DataFrame({
231     'Source': sources_int,
232     'Target': targets,
233     'Value': values
234 })
235
236 # Save DataFrame to CSV (without index)
237 df.to_csv(os.path.join(output_folder,
238                        f'{cropWC}_table.csv'),index=False)
239
240 # Save the same CSV file, but with labels instead of numbers
241 source_label = np.array(labels)[sources_int]
242 targets_label = np.array(labels)[targets]
243
244 # Create a DataFrame with each vector as a column
245 df_l = pd.DataFrame({
246     'Source': source_label,
247     'Target': targets_label,
248     'Value': values
249 })
250
251 # Save DataFrame to CSV (without index)
252 df_l.to_csv(os.path.join(output_folder,
253                          f'{cropWC}_table_labels.csv'),index=False)
254
255 ############################################################
256
257 ## PART 6 - SANKEY DIAGRAMS
258
259 node_colors = [
260     "#DC3220",   # red
261     "#006B08",   # green
262     "#0C51B5",   # blue
263     "#EE510E",   # maize
264     "#EDB120",   # wheat
265     "#44AA99",   # soy
266     "#EE510E",   # maize
267     "#EDB120",   # wheat
268     "#44AA99",   # soy
269     "#2D95DA",   # ocean
270     "#AA4499"    # other land
271 ]
272
273 fig = go.Figure(data=[go.Sankey(
274     node=dict(
275         pad=20,
276         thickness=30,
```

```
        line=dict(color="black", width=0.5),
        label=labels,
        color=node_colors
    ),
    link=dict(
        source=sources_int,
        target=targets,
        value=values
    )
)])

fig.update_layout(
    font=dict(family="Times New Roman, serif", size=14, color="black"),
    width=900,
    height=600,
    margin=dict(l=50, r=50, t=70, b=50)
)

fig.write_image(os.path.join(output_folder,
                             f'{cropWC}_sankey.svg'), scale=2)
```

Bibliography

[1] Walter Willett et al. «Food in the Anthropocene: the EAT–Lancet Commission on healthy diets from sustainable food systems». English. In: *The Lancet* 393.10170 (Feb. 2019), pp. 447–492. ISSN: 0140-6736, 1474-547X. DOI: 10.1016/S0140-6736(18)31788-4 (cit. on pp. 1–3).

[2] Food and Agriculture Organization (FAO). *AQUASTAT – Water Use Methodology*. 2010 (cit. on pp. 1–4).

[3] OECD-FAO. *Agricultural Outlook 2025-2034: Full Report | OECD*. 2025 (cit. on p. 1).

[4] David Tilman and Michael Clark. «Global diets link environmental sustainability and human health». en. In: *Nature* 515.7528 (Nov. 2014), pp. 518–522. ISSN: 1476-4687. DOI: 10.1038/nature13959 (cit. on p. 1).

[5] Arjen Y. Hoekstra and Mesfin M. Mekonnen. «The water footprint of humanity». In: *Proceedings of the National Academy of Sciences* 109.9 (Feb. 2012), pp. 3232–3237. DOI: 10.1073/pnas.1109936109 (cit. on p. 1).

[6] Line J. Gordon, Garry D. Peterson, and Elena M. Bennett. «Agricultural modifications of hydrological flows create ecological surprises». In: *Trends in Ecology & Evolution* 23.4 (Apr. 2008), pp. 211–219. ISSN: 0169-5347. DOI: 10.1016/j.tree.2007.11.011 (cit. on pp. 2, 4, 14, 15).

[7] M. Falkenmark and J. Rockström. «The New Blue and Green Water Paradigm: Breaking New Ground for Water Resources Planning and Management». EN. In: *Journal of Water Resources Planning and Management* 132.3 (May 2006), pp. 129–132. ISSN: 0733-9496. DOI: 10.1061/(ASCE)0733-9496(2006)132:3(129) (cit. on pp. 2, 7).

[8] FAO. *The State of the World's Land and Water Resources for Food and Agriculture – Systems at breaking point (SOLAW 2021)*. en. FAO, Dec. 2021. ISBN: 978-92-5-135327-1. DOI: 10.4060/cb7654en (cit. on pp. 2, 4).

[9] Mesfin Mekonnen and Arjen Hoekstra. «The green, blue and grey water footprint of crops and derived crop products». In: *Hydrology and Earth System Sciences Discussions* 8 (Jan. 2011). DOI: 10.5194/hessd-8-763-2011 (cit. on pp. 2, 3, 5, 8, 17, 18).

[10] Lan Wang-Erlandsson et al. «A planetary boundary for green water». en. In: *Nature Reviews Earth & Environment* 3.6 (June 2022), pp. 380–392. ISSN: 2662-138X. DOI: 10.1038/s43017-022-00287-8 (cit. on p. 2).

[11] Paolo D'Odorico et al. «Global virtual water trade and the hydrological cycle: patterns, drivers, and socio-environmental impacts». en. In: *Environmental Research Letters* 14.5 (Apr. 2019), p. 053001. ISSN: 1748-9326. DOI: 10.1088/1748-9326/ab05f4 (cit. on p. 3).

[12] Obbe A. Tuinenburg and Arie Staal. «Tracking the global flows of atmospheric moisture and associated uncertainties». English. In: *Hydrology and Earth System Sciences* 24.5 (May 2020), pp. 2419–2435. ISSN: 1027-5606. DOI: 10.5194/hess-24-2419-2020 (cit. on pp. 3, 11, 13, 14).

[13] Rudi J. van der Ent, Hubert H. G. Savenije, Bettina Schaefli, and Susan C. Steele-Dunne. «Origin and fate of atmospheric moisture over continents». en. In: *Water Resources Research* 46.9 (2010). ISSN: 1944-7973. DOI: 10.1029/2010WR009127 (cit. on p. 3).

[14] Patrick W. Keys, Lan Wang-Erlandsson, and Line J. Gordon. «Revealing Invisible Water: Moisture Recycling as an Ecosystem Service». en. In: *PLOS ONE* 11.3 (Mar. 2016), e0151993. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0151993 (cit. on pp. 3, 4, 11, 12).

[15] Simon Felix Fahrländer, Lan Wang-Erlandsson, Agnes Pranindita, and Fernando Jaramillo. «Hydroclimatic Vulnerability of Wetlands to Upwind Land Use Changes». en. In: *Earth's Future* 12.3 (2024), e2023EF003837. ISSN: 2328-4277. DOI: 10.1029/2023EF003837 (cit. on p. 4).

[16] John C O'Connor, Maria J Santos, Stefan C Dekker, Karin T Rebel, and Obbe A Tuinenburg. «Atmospheric moisture contribution to the growing season in the Amazon arc of deforestation». en. In: *Environmental Research Letters* 16.8 (July 2021), p. 084026. ISSN: 1748-9326. DOI: 10.1088/1748-9326/ac12f0 (cit. on p. 4).

[17] Delphine Clara Zemp, Carl-Friedrich Schleussner, Henrique M. J. Barbosa, Marina Hirota, Vincent Montade, Gilvan Sampaio, Arie Staal, Lan Wang-Erlandsson, and Anja Rammig. «Self-amplified Amazon forest loss due to vegetation-atmosphere feedbacks». In: *Nature Communications* 8 (Mar. 2017), p. 14681. ISSN: 2041-1723. DOI: 10.1038/ncomms14681 (cit. on p. 4).

[18] Chandrakant Singh and U. Persson. *Global patterns of commodity-driven deforestation and associated carbon emissions*. Feb. 2024. DOI: 10.31223/X5T69B (cit. on p. 4).

[19] International Food Policy Research Institute (IFPRI). *Global Spatially- Disaggregated Crop Production Statistics Data for 2010 Version 2.0*. en. Feb. 2025. DOI: 10.7910/DVN/PRFF8V (cit. on pp. 5, 17, 18).

[20] Commission on the Economics of Water. *The Economics of Water: Valuing the Hydrological Cycle as a Global Common Good*. en-US. 2024 (cit. on pp. 6, 7, 13).

[21] Tom Gleeson et al. «Illuminating water cycle modifications and Earth system resilience in the Anthropocene». en. In: *Water Resources Research* 56.4 (2020), e2019WR024957. ISSN: 1944-7973. DOI: 10.1029/2019WR024957 (cit. on p. 6).

[22] Intergovernmental Panel On Climate Change (Ipcc) and Hervé Douville. *Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. en. 1st ed. Cambridge University Press, July 2023. ISBN: 978-1-009-15789-6. DOI: 10.1017/9781009157896 (cit. on pp. 7, 14, 15).

[23] Stefanie Rost, Dieter Gerten, Alberte Bondeau, Wolfgang Lucht, Janine Rohwer, and Sibyll Schaphoff. «Agricultural green and blue water consumption and its influence on the global water system». en. In: *Water Resources Research* 44.9 (2008). ISSN: 1944-7973. DOI: 10.1029/2007WR006331 (cit. on pp. 8, 10).

[24] Richard Allen, L. Pereira, Dirk Raes, and M. Smith. «Crop evapotranspiration guidelines for computing crop requirements. FAO Irrig. Drain. Report modeling and application». In: *J. Hydrol.* 285 (Jan. 1998), pp. 19–40 (cit. on pp. 8–10, 17, 18, 25).

[25] M. Salman. *The AquaCrop model – Enhancing crop water productivity*. English. FAO ; 2021. ISBN: 978-92-5-135222-9 (cit. on p. 11).

[26] Martin Smith and Food and Agriculture Organization of the United Nations. *CROPWAT: A Computer Program for Irrigation Planning and Management*. en. Food & Agriculture Org., Jan. 1992. ISBN: 978-92-5-103106-3 (cit. on p. 11).

[27] Claudio O. Stöckle, Marcello Donatelli, and Roger Nelson. «CropSyst, a cropping systems simulation model». In: *European Journal of Agronomy*. Modelling Cropping Systems: Science, Software and Applications 18.3 (Jan. 2003), pp. 289–307. ISSN: 1161-0301. DOI: 10.1016/S1161-0301(02)00109-0 (cit. on p. 11).

[28] Marta Tuninetti, Stefania Tamea, Paolo D'Odorico, Francesco Laio, and Luca Ridolfi. «Global sensitivity of high-resolution estimates of crop water footprint». en. In: *Water Resources Research* 51.10 (2015), pp. 8257–8272. ISSN: 1944-7973. DOI: 10.1002/2015WR017148 (cit. on pp. 11, 17, 20, 39, 62).

[29] Nike Chiesa Turiano, Marta Tuninetti, Francesco Laio, and Luca Ridolfi. «Evaluating Country-Scale Irrigation Demand Through Parsimonious Agro- Hydrological Modeling». en. In: *Hydrology* 12.9 (Sept. 2025), p. 240. ISSN: 2306-5338. DOI: 10.3390/hydrology12090240 (cit. on pp. 11, 20, 21).

[30] Elena De Petrillo, Luca Monaco, Marta Tuninetti, Arie Staal, and Francesco Laio. «Cell-scale atmospheric moisture flows dataset reconciled with ERA5 reanalysis». In: *Scientific Data* 12 (Apr. 2025). DOI: 10.1038/s41597-025-04964-3 (cit. on p. 11).

[31] Elena De Petrillo, Simon Fahrländer, Marta Tuninetti, Lauren Andersen, Luca Monaco, Luca Ridolfi, and Francesco Laio. «Reconciling tracked atmospheric water flows to close the global freshwater cycle». In: *Communications Earth & Environment* 6 (May 2025). DOI: 10.1038/s43247-025-02289-y (cit. on pp. 11, 17, 32, 58).

[32] Michael G. Bosilovich, Franklin R. Robertson, Lawrence Takacs, Andrea Molod, and David Mocko. «Atmospheric Water Balance and Variability in the MERRA-2 Reanalysis». en. In: (Feb. 2017). DOI: 10.1175/JCLI-D-16-0338.1 (cit. on p. 12).

[33] Fanny Lehmann, Bramha Dutt Vishwakarma, and Jonathan Bamber. «How well are we able to close the water budget at the global scale?» English. In: *Hydrology and Earth System Sciences* 26.1 (Jan. 2022), pp. 35–54. ISSN: 1027-5606. DOI: 10.5194/hess-26-35-2022 (cit. on p. 12).

[34] Maganizo Kruger Nyasulu, Ingo Fetzer, Lan Wang-Erlandsson, Fabian Stenzel, Dieter Gerten, Johan Rockström, and Malin Falkenmark. «African rainforest moisture contribution to continental agricultural water consumption». In: *Agricultural and Forest Meteorology* 346 (Mar. 2024), p. 109867. ISSN: 0168-1923. DOI: 10.1016/j.agrformet.2023.109867 (cit. on pp. 12, 43).

[35] Luis Gimeno et al. «Recent progress on the sources of continental precipitation as revealed by moisture transport analysis». In: *Earth-Science Reviews* 201 (Feb. 2020), p. 103070. ISSN: 0012-8252. DOI: 10.1016/j.earscirev.2019.103070 (cit. on p. 13).

[36] Peter Kalverla, Imme Benedict, Chris Weijenborg, and Ruud J. van der Ent. «Atmospheric moisture tracking with WAM2layers v3». English. In: *Geoscientific Model Development* 18.14 (July 2025), pp. 4335–4352. ISSN: 1991-959X. DOI: 10.5194/gmd-18-4335-2025 (cit. on p. 13).

[37] Elena De Petrillo, Luca Monaco, Marta Tuninetti, Arie Staal, and Francesco Laio. *RECON - Cell-scale atmospheric moisture flows dataset reconciled with ERA5 reanalysis*. Nov. 2024. DOI: 10.5281/ZENODO.14191920 (cit. on pp. 14, 17, 18).

[38] Lan Wang-Erlandsson, Ingo Fetzer, Patrick W. Keys, Ruud J. van der Ent, Hubert H. G. Savenije, and Line J. Gordon. «Remote land use impacts on river flows through atmospheric teleconnections». English. In: *Hydrology and Earth System Sciences* 22.8 (Aug. 2018), pp. 4311–4328. ISSN: 1027-5606. DOI: 10.5194/hess-22-4311-2018 (cit. on p. 14).

[39] Chris West, Gabriela Rabeschini, Chandrakant Singh, Thomas Kastner, Mairon Bastos Lima, Ahmad Dermawan, Simon Croft, and U. Martin Persson. «The global deforestation footprint of agriculture and forestry». en. In: *Nature Reviews Earth & Environment* 6.5 (May 2025), pp. 325–341. ISSN: 2662-138X. DOI: 10.1038/s43017-025-00660-3 (cit. on p. 15).

[40] Mario Herrero et al. «Biomass use, production, feed efficiencies, and greenhouse gas emissions from global livestock systems». In: *Proceedings of the National Academy of Sciences* 110.52 (Dec. 2013), pp. 20888–20893. DOI: 10.1073/pnas.1308149110 (cit. on p. 15).

[41] Mark New, David Lister, Mike Hulme, and Ian Makin. «A high-resolution data set of surface climate over global land areas». en. In: *Climate Research* 21 (May 2002), p. 1. ISSN: 0936-577X, 1616-1572. DOI: 10.3354/cr021001 (cit. on p. 16).

[42] Hans Hersbach et al. «The ERA5 global reanalysis». en. In: *Quarterly Journal of the Royal Meteorological Society* 146.730 (2020), pp. 1999–2049. ISSN: 1477-870X. DOI: 10.1002/qj.3803 (cit. on pp. 16, 18, 33).

[43] Felix T. Portmann, Stefan Siebert, and Petra Döll. «MIRCA2000—Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling». en. In: *Global Biogeochemical Cycles* 24.1 (2010). ISSN: 1944-9224. DOI: 10.1029/2008GB003435 (cit. on pp. 17, 18).

[44] FAO/IIASA/ISRIC/ISSCAS/JRC. *Harmonized World Soil Database (version 1.2)*. 2012 (cit. on pp. 17, 18).

[45] Stefan Lange and Matthias Büchner. *ISIMIP3 land-sea masks*. 2020. DOI: 10.48364/ISIMIP.822294 (cit. on p. 18).

[46] P. W. Keys, R. J. van der Ent, L. J. Gordon, H. Hoff, R. Nikoli, and H. H. G. Savenije. «Analyzing precipitationsheds to understand the vulnerability of rainfall dependent regions». English. In: *Biogeosciences* 9.2 (Feb. 2012), pp. 733–746. ISSN: 1726-4170. DOI: 10.5194/bg-9-733-2012 (cit. on p. 43).

[47] Justin E Bagley, Ankur R Desai, Paul A Dirmeyer, and Jonathan A Foley. «Effects of land cover change on moisture availability and potential crop yield in the world's breadbaskets». en. In: *Environmental Research Letters* 7.1 (Jan. 2012), p. 014009. ISSN: 1748-9326. DOI: 10.1088/1748-9326/7/1/014009 (cit. on p. 43).