



**Politecnico
di Torino**

Politecnico di Torino

Master's of Science in Environmental and Land Engineering

A.a. 2021/2022

Graduation Session March 2022

**The spatio-temporal variability of
green and blue water footprint of
maize**

Supervisor:

Prof. Stefania Tamea

Co-supervisor:

Dr. Matteo Rolle

Candidate:

Andrea Borgo

Acknowledgments

I would like to warmly thank my thesis supervisor, Prof. Stefania Tamea, for giving me the possibility to work on such stimulating topic, for her constant support during the thesis elaboration and for transmitting me the knowledge and passion about the water footprint concept.

In addition, I would like to express my gratitude to Dr. Matteo Rolle for providing me all the needed materials to execute the work and for his continuous effort in solving my doubts.

Finally, my last appreciation goes to my mother, my father and Lucia, for believing in my dreams. Their endless support throughout my academic career, especially during the hardest times, allowed me to find the motivation to pursue my targets and achieve the hoped-for results.

Abstract

The management of water resources for human purposes is a central theme to the challenge of improving food security worldwide. The great majority of freshwater is used for agriculture; therefore, it is important to find a way to assess and quantify water consumption for the production of agricultural goods. The indicator of unit Water Footprint (uWF) identifies the volume of water required to produce a unit of product. This indicator may refer separately to green and blue water, which are respectively the rainwater component (green water footprint) and the irrigation contribution (blue water footprint). In this study, the spatial and temporal variability of green and blue water footprint of maize is investigated, exploring its evolution in the years 1970 – 2019 across the world, with the creation of high-resolution global maps (5 x 5 arc min resolution), and highlighting the spatial heterogeneities of this variable as dependent on crop yield and climatic conditions. To reach this result, the global temporal variability of harvested areas, irrigated areas and crop yield has been reconstructed, together with the hydrological soil water balance through an appropriate model, which joins vegetation growth with local soil and hydroclimatic conditions. From the results, it can be noticed that the green component is usually greater than the blue one and, in the majority of countries, green and blue uWF are progressively reducing through the years, mainly due to technological developments, better irrigation strategies and water management which allow to increase the crop yield. In fact, the main factor influencing the trend of uWF is crop yield, which along the time-interval of the analysis, exhibited significant changes. However, despite the increment of water use efficiency, the estimated overall volume of green and blue water shows a mild variation, mostly due to the effect of climate change and modification of agricultural and irrigated areas.

Contents

- 1 Introduction..... 8
- 2 Main concepts and references 12
 - 2.1 Global sensitivity of high-resolution estimates of crop water footprint..... 13
 - 2.2 Fast track approach..... 14
 - 2.3 Water Footprint Network..... 17
- 3 Data analysis and preparation 18
 - 3.1 National Yield..... 18
 - 3.1.1 Italy..... 18
 - 3.1.2 Rest of the world 19
 - 3.2 National yield and area in time..... 23
 - 3.2.1 Country variations in time..... 24
 - 3.3 Yield and area evolution..... 27
 - 3.4 Distributed yield area and production in time 29
 - 3.5 Maps of production..... 31
 - 3.6 Comparison between estimated map of year 2010 and GAEZ map of year 2010 33
 - 3.7 Temporal variations in GAEZ data (cell switches) 34
 - 3.7.1 Distribution in switched on and off cells 36
 - 3.7.2 Switches by country 38
 - 3.7.3 Empirical Cumulative Distribution Function (ECDF) of switches..... 43
 - 3.7.4 Compare national yield and total harvested area of the reconstructed and GAEZ maps 44
 - 3.7.5 Overall comparison of yield and area pixels of reconstructed and GAEZ maps46
- 4 Water footprint calculation 50
 - 4.1 Actual evapotranspiration..... 50
 - 4.2 Modelling the green and blue evapotranspiration 52
 - 4.3 Distributed unit water footprint at the global scale 54
 - 4.3.1 Global maps of green, blue and total uWF..... 54
 - 4.3.2 Historical series of mean national green, blue and total uWF 58
 - 4.4 Trend analysis..... 58
- 5 Results..... 60
 - 5.1 Local analysis: Pino Torinese..... 60
 - 5.1.1 Run the hydrological model on a single pixel: Pino Torinese 60

5.1.2	Calculation of uWF in Pino Torinese.....	66
5.2	Relation between blue fraction of uWF and climatology in Pino Torinese and at global scale	67
5.3	Display global maps of total, green and blue uWF	71
5.3.1	Total uWF, years 1970, 1980, 2000, 2019	72
5.3.2	Green uWF, years 1970, 1980, 2000, 2019.....	73
5.3.3	Blue uWF, years 1970, 1980, 2000, 2019.....	75
5.4	Evolution of global Area Equipped for Irrigation	77
5.5	Temporal evolution of total, green and blue uWF of some countries	79
5.5.1	Trend analysis of blue uWF and crop yield	84
5.6	Results verification	85
5.6.1	Comparison with Water Footprint Network.....	85
5.6.2	Verify coherence of results comparing bulk versus local-weighted data.....	89
5.6.3	Comparison with Fast Track approach.....	91
5.7	From unit water footprint [m ³ /ton] to water footprint [m ³].....	94
6	Conclusions.....	100
7	Annexes.....	102
7.1	Display of input variables of uWF in Pino Torinese	102
7.2	Display time series of harvested area, AEI, production and ET of the five countries 104	
7.2.1	Italy.....	104
7.2.2	US.....	105
7.2.3	Australia	105
7.2.4	Nigeria.....	106
7.2.5	Viet Nam	106
7.3	Global map of area equipped for irrigation, 2000	107
7.4	National green uWF table.....	107
7.5	National blue uWF table.....	111
8	References.....	114

1 Introduction

From the early stages of civilizations, humans have always relied on water consumption for their activities. Water security has always been of primary importance for world's population, especially nowadays, where the great number of people living on Earth, the improved life standards, but also the effects of climate change are challenging the management of water supplies, therefore requiring an enhancement of its consumption's efficiency. The global population is expected to continuously grow up to 9.6 billion people by 2050 (United Nations, 2013) and, together with the increasing wealth of people and the changing diets, they are putting pressure on the food supply chain, one of the sectors that mostly relies on water usage (Godfray & al., 2010). According to the FAO report of 2011, with an increase of population up to 9.6 billion by 2050, we will also face an increase of agricultural products' demand by 70% (FAO, 2011).

At a global scale, the great majority of water consumption occurs in the agricultural and industrial sectors; in particular, water use for food crops accounts for approximately 80-90 % of water use for human needs overall (Falkenmark & Rockstrom, 2004). The great part of total human freshwater withdrawals (about 70%) comes from the irrigation of crops (FAO, 2011). In the past 50 years, the world's cultivated area has grown by 12%, removing space to forests, wetlands and grasslands, and the global irrigated area has doubled. However, a further expansion of land for cultivation is getting increasingly limited, since most the easily accessible water resources are already under exploitation (FAO, 2011). Besides, climate change is seriously harming this already critical situation by limiting the occurrence of precipitation events, which allow the recharging of surface water bodies and groundwater aquifers, increasing the evapotranspiration of plants and by causing severe droughts in some parts of the worlds, such as the water crisis in Somalia (FGS, 2017). In addition, extreme weather events are likely to get increasingly frequent worldwide, thus leading to a general reduction of the crop yield (Powell & Reinhard, 2016).

Scientists, politicians, and decision makers are rising awareness that new strategies to ensure water security both to humans and to the environment must be developed and this will probably constitute one of the greatest challenges of the new millennium (Allamano, et al., 2013). It is therefore important to find a linkage between food products and water resources, in order to quantify water consumption according to the produced crop. Several indicators for the quantification of water consumption already exist, but one of those which gained greater interest

by the scientific community is Water Footprint (WF). This indicator expresses the volume of freshwater for the production of a specific good along its entire production chain and it looks not only at the direct water use of the producer or consumer, but it considers also the indirect consumption. Water Footprint is a multidimensional indicator, being the sum of three main contributes:

- Green Water Footprint: it refers to the consumption of rainwater, or, in general, the water that comes from precipitation
- Blue Water Footprint: this is the contribute provided by blue water resources, that is surface and groundwater
- Grey Water Footprint: it refers to the amount of water needed to assimilate the load of pollutants given natural background concentrations (Hoekstra et al., 2011).

The possibility of splitting Water Footprint into these components allow us to better understand how water consumption is partitioned over the production of a product, separating the green contribute from the irrigation inputs (in this thesis, grey WF will be omitted because it provides indication about the pollution rather than the effective water consumption).

Several studies have already conducted analysis of the historical evolution of water footprint, providing maps of its spatial heterogeneity worldwide. There have been studies dealing with water use of food crops accounting at global scale, but also at regional scale, such as at continental or country level, with the attempt of assessing water footprint at higher spatial resolution (0.5 x 0.5 arc deg resolution and 5 x5 arc min with time scales ranging from 1971 to 2005). All these studies depend on many assumptions, regarding the input parameters, the data sets and the modelling structure, while only few studies have developed a sensitivity analysis of the water footprint calculations to define the accuracy of the outcomes, being focused on specific regions (river basin scale) (Tuninetti et al., 2015).

In the present study, we focus on both unit Water Footprint (uWF) concept, that is the volume of green or blue water necessary to produce a unit of product [m^3/ton], and Water Footprint, the effective water volume consumed by a country in a given year [m^3]. The goal of this work is the reconstruction of the historical evolution of water footprint at global scale, separating the green from the blue contribute, thus reconstructing the historical world maps of green and blue unit water footprint. To this end, we produce 5 x 5 arc min resolution global maps in the years 1970

– 2019, relying on historical series of crop yield, precipitation, crop evapotranspiration and soil moisture.

This thesis chooses maize as testing crop for our model. In fact, maize belongs to the group of the four most cultivated grains (wheat, rice, maize and soybean) which, together, account for over 50% of the global human consumption in terms of caloric content (wheat: 20%, rice: 16%, maize: 13%, and soybean: 8%) (D'Odorico et al., 2014) and 50% of global crop water footprint (Mekonnen & Hoekstra, 2011). Moreover, this crop is rather easy to analyse because each growing period remains inside the one-year cycle (i.e., planting date always occur in Spring and harvesting in Autumn), while other crops, such as wheat, are generally planted in late Autumn and the harvesting day is Spring-Summer of the following year. Moreover, maize has a relatively simple way of cultivation, where water inputs only come from rain and/or irrigation (in many rice cultures, instead, a complete flooding of the field in some stages of the process is required).

2 Main concepts and references

First of all, as mentioned in the introduction, the water footprint corresponds to the water volume withdrawn for the production of a food crop and the unit Water Footprint is the wf per unit mass of produced goods, dimensionally associated to a volume over mass of product [m^3/ton].

In our analysis, we focus on the partition of Water Footprint into its green and blue components, analysing the spatio-temporal variability of these variables over 5 x 5 arc min resolution global maps, corresponding to pixels of about 9 km x 9 km at the equator (Tuninetti et al., 2015). In each pixel of such maps, unit Water Footprint is expressed as the ratio between the water evapotranspired by maize crop during its growing period, $ET_{a,(i,j),y}$ [mm/year] and the crop yield in that cell $Y_{(i,j),y}$ [ton/ha]. In this way, with the following equation

$$uWF_{(i,j),y} = 10 \cdot \frac{ET_{a,(i,j),y}}{Y_{(i,j),y}} \left(\frac{m^3}{ton} \right) \quad (2.1)$$

we obtain the cubic meters of water per ton of product which have been used in the cell (i,j) of the global grid at year y . The 10 factor is used to convert the millimetres of ET_a into a volume per unit of surface, [m^3/ha], so that we can simplify ha and obtain [m^3/ton]. Since ET_a can be split into green and blue contribute, we can say that

$$ET_a = ET_g + ET_b \quad (2.2)$$

where ET_g corresponds to Green Evapotranspiration, which keeps into account the evapotranspiration output referred to precipitation, and ET_b corresponds to Blue Evapotranspiration, that is the contribute which comes from irrigation.

This work comes from the need of expanding an analysis dealing with Water Footprint accounting that has been conducted in the past. More specifically, there have already been attempts to derive crop water consumption at high spatial resolution, at global or national scale, for example the work provided by Mekonnen and Hoekstra (Mekonnen & Hoekstra, 2011), such as the one of Hanasaki (Hanasaki et al., 2010) Plus, there have been studies which carried on sensitivity analysis of crop Water Footprint to input parameters (Tuninetti et al., 2015) and others which looked for data-based and simplified ways to derive time-varying unit water

footprint considering the sum of green and blue water (Tuninetti et al., 2015). There are also online platforms, such as the Water Footprint Network, which provide updates (by reporting papers and books on the topic), interactive tools, data and statistics for assessing Water Footprint at national and global scale, but mostly without a temporal variability and/or without a separation of the green and blue components. So, before the description of the steps of this study, it is necessary to briefly recall and highlight the key concepts of the main research articles which provided the basis for this work.

2.1 Global sensitivity of high-resolution estimates of crop water footprint

The study conducted by Tuninetti et al. (2015) investigates the water footprint of the four globally mostly produced crops, that are wheat, rice, maize and soybean, by producing 5 x 5 arc min resolution maps using the most recently available data for agricultural crop yield, precipitation, evapotranspiration and soil water content. From these maps, they evaluated the spatial heterogeneity of water footprint both at grid-cell and continental scale, to assess the main drivers of water footprint and to evaluate the relationship between water footprint and crop production. In addition, this study carries on a sensitivity analysis to understand which of the parameters directly manageable by farmers and land managers (e.g., planting date duration of the growing period, etc.) has a greater weight on the determination of crop Water Footprint. Final results come from an average of the input parameters over the time interval 1996 – 2005, in order to have input data independent from interannual fluctuations.

The crop evapotranspiration of a single growing season is derived by the sum of the daily evapotranspiration [mm/day] over the length of the growing period, becoming then [mm/year]. To estimate green and blue evapotranspiration in each cell, instead, they performed a weighted mean of the rainfed and irrigated evapotranspiration with the following equations

$$ET_{g,LGP} = \frac{ET_{g,LGP}^R * A^R + ET_{g,LGP}^I * A^I}{A^R + A^I}$$

(2. 3)

$$ET_{b,LGP} = \frac{ET_{b,LGP}^I * A^I}{A^R + A^I} \quad (2.4)$$

where $ET_{g,LGP}^R$ and $ET_{g,LGP}^I$ are the green evapotranspiration that respectively occur over the rainfed and irrigated areas for all the length of the growing period LGP, $ET_{b,LGP}^I$ is the blue evapotranspiration that only occurs in the irrigated areas and A^R and A^I are respectively the harvested rainfed and irrigated area in each pixel.

To determine crop Actual Yield, they resort to the results of the work provided by Monfreda et al. in 2008 and to data from FAOSTAT database, performing a weighted average over the period 1996 – 2005 with a combination of the two.

The sensitivity analysis is performed by the creation of the Sensitivity Index (SI_x), from a Taylor expansion of the functional dependence of Water Footprint.

Looking at the spatial heterogeneities of the reconstructed high-resolution maps of Water Footprint, it can be observed that this variability is mainly affected by yield pattern, with a correlation coefficient of 0.74, while the effect of climate drivers (included into the evapotranspiration term) shows a correlation coefficient of only 0.34. Therefore, according to this study, the agricultural technological development of a country is expected to be much more influent than its climate conditions regarding water consumption.

Regarding the sensitivity analysis, instead, the chosen parameters, that is the Available Water Content (AWC), Reference Evapotranspiration ET_0 , the length of the growing period LGP and the Planting Date PD, show different results according to the type of crop. More specifically, it has been observed that wheat is the most sensitive crop to the length of the growing period, rice instead is the one that is mostly influenced by changes of ET_0 , while maize and soybean to the planting date. The knowledge of these information by agricultural managers allow to refine the cultivation strategies to further reduce water requirements in food crops.

2.2 Fast track approach

The goal of the work performed by Tuninetti et al. in 2017 aims at finding a simple approach for the determination of crop Water Footprint. As we said earlier, Water Footprint is dependent from the evapotranspiration and crop yield. However, while the latter is easily available from

databases (FAOSTAT in particular, which provides the historical time series of mean national yield for each country from years 1961 – 2020), the former is in turn dependent on several parameters dealing with local climate and soil conditions (precipitation, reference evapotranspiration, soil moisture, etc.), which may not always be easily accessible, thus they require a longer time to be acquired. In addition, the calculation of crop evapotranspiration is rather complex and requires computationally demanding models. Therefore, this research work proposes a quick approach, named Fast Track, that considers the time variability of Water Footprint to be only dependent by crop yield patterns, neglecting the contribution of evapotranspiration variations, and it estimates the error between this approach and the detailed one.

According to the Fast Track (FT) approach, the crop water footprint of country c in year t is only influenced by yield variations, while the evapotranspiration is kept constant to an average value over a period T ,

$$uWF_{c,t}(Y) = 10 * \frac{\overline{ET}_{c,T}}{Y_{c,t}} \left(\frac{m^3}{ton} \right) \quad (2.5)$$

As for the previous study, the analysis is carried on by choosing wheat, rice, maize and soybean, since, together, they are globally the most impacting in terms of water use.

The time-averaged CWF can be scaled with yield variations, with the equation

$$uWF_{c,t}(Y) = \frac{\overline{CWF}_{c,T} * \overline{Y}_{c,T}}{Y_{c,t}} \left(\frac{m^3}{ton} \right) \quad (2.6)$$

where T has been selected to be the time interval from years 1996 – 2005, $\overline{Y}_{c,T}$ is the average crop yield over time T , $\overline{CWF}_{c,T}$ is the average crop water footprint over T and $Y_{c,t}$ is the yield of a specific year t .

By using the previous equation, it is possible to reconstruct the historical estimates of crop water footprint at country scale for the time interval 1961 – 2013. Indeed, the only input parameters are the averaged crop water footprint over T ($\overline{CWF}_{c,T}$), the average yield over T ($\overline{Y}_{c,T}$) and the historical series of crop yield from 1961 to 2013 ($Y_{c,t}$). The first variable is the result of the work provided by Tuninetti et al. (2015), where, from a 5 x 5 arc min resolution

global gridded map, national values have been obtained from a production-weighted mean of the pixel of each country. $\overline{Y_{c,T}}$, instead, has been derived by averaging FAOSTAT national yield series over T (that is, from 1996 to 2005), while $Y_{c,t}$ are the historical time series of mean national yield from 1961 to 2013.

At the same time the study provides an evaluation of the crop water footprint with a detailed approach, that is by keeping into account annual variations of both crop yield and evapotranspiration.

The annual crop evapotranspiration, which is the sum of all daily ETs over the crop growing period, is derived by the product of reference evapotranspiration ET_0 (driven by climate factors) and the crop coefficient k_c (which keeps into account the characteristics specific for that crop). While, for simplicity, crop properties are assumed to be constant over the years, reference evapotranspiration, such as precipitation data are taken from the CRU database, which provides 30 x 30 arc min resolution global maps from 1961 to 2013 on a monthly basis. Historical crop yield maps are instead reconstructed from the results provided by Monfreda et al (2008), adjusting the maps at 5 x 5 arc min resolution for year 2000 ($Y_{i,t=2000}^{M0}$).

Looking at the results of the study, it is possible to see that the coefficient of determination R^2 of the scatter plots of the national crop water footprint in year t within the period 1961 – 2013, comparing FT approach with detailed approach, is 0.977 for wheat, 0.965 for rice, 0.973 for maize and 0.914 for soybean. This confirms that crop yield pattern is the main factor influencing the temporal variability of crop water footprint. However, even with FT approach, climate variations are not neglected since they are incorporated in the yield time series. In fact, according to Ray et al. (2015), almost 30% of temporal yield variability is driven by climate factors (Ray, Gerber, MacDonald, & West, 2015).

Therefore, this study demonstrates that the application of the Fast Track approach to estimate temporal variability of crop water footprint is feasible, with a very limited error between this approach and the detailed one. In fact, by keeping constant the evapotranspiration, they found a rather low uncertainty of the water footprint estimates, with a standard deviation of the error around 0.1. Therefore, by applying this method, it is possible to acquire CWF estimates with an easier and fast applicability tool.

2.3 Water Footprint Network

The Water Footprint Network is a network of people, researchers and institutions, with a corresponding website from which Water Footprint data are taken to validate the results of this work. This is a platform that collaborates with companies, organisations and individuals which work in water-related businesses or simply interested in water footprint issues to help them in finding innovative strategies for a fair and smarter water use.

In 2002, Arjen Hoekstra proposed the concept of Water Footprint as a metric to measure water consumption for the production of goods over their supply chain. Over the years, the interest on this theme rapidly raised among the scientific community and in companies, especially those working in the field of food and beverages. Therefore in 2008, Hoekstra together with the global players from business, civil and academic society founded the Water Footprint Network, a network to create cohesion between people who believe in the power of water footprint as tool to promote more efficient water use.

This is a non-profit network which coordinates different activities, such as:

- Network and Exchange
- Awareness Raising
- Capacity Building
- Knowledge and Data Dissemination
- Influencing Policy and Practice

This platform is also a database that contains a set of resources which articulates in an archive of water footprint-related publications, interactive tools for the assessment of individual water consumption, global maps of water footprint and statistics of green, blue and grey water footprint at product and national scale.

From this database, we collect the values of green and blue water footprint of national production and consumption of maize, averaged over the period 1996 – 2005. These data are used to validate the results obtained with the method proposed in this study.

Moreover, from the Water Footprint Network we collect the book “Water Footprint Assessment Manual” (Hoekstra et al., 2011), which provide a complete and updated overview of the global standards on the assessment of this indicator.

3 Data analysis and preparation

First of all, we download the global 5 arc minutes resolution map of the yield of maize, expressed as a matrix of 2160 x 4320 pixels from GAEZ v4 Data Portal. As reported in the GAEZ webpage, this Portal is the result of the cooperation of the Food and Agriculture Organization of the United Nations (FAO) and the International Institute for Applied Systems Analysis (IIASA), which worked together to implement the Agro-Ecological Zones (AEZ) databases. These zones were built from a deep analysis of land evaluation principles to assess natural resources and estimate agricultural land use.

The Agro-Ecological Zones were initially implemented in the 1980s and have been continuously improved in time up to the first global AEZ assessment in 2000 (GAEZ v1), thanks to the technological development of geo-information systems. Then, other versions of the database were released (GAEZ v2 and v3) up to the today's assessment GAEZ v4, which comprises large amounts of agro-ecological indicators.

Results are represented in a raster format, with a resolution of 5 arc-minute, approximately 9 x 9 km for a single cell at the equator and provides the global maps of 2000 and 2010.

3.1 National Yield

3.1.1 Italy

The analysis starts from the global maps of yield [ton/ha] and harvested areas [ha] of maize in 2000, where a first calculation of Italian mean national yield is performed. In order to extract only the Italian cells from the global matrix we get helped by a sort of “mask matrix”, the *map_FAOcodes*, a matrix with the same dimensions of the one of GAEZ (2160 x 4320), where each cell contains the FAO code corresponding to the country in which it supposed to be part of.

With the *find()* command, choosing only the cells with yield and harvested area greater than zero and with the FAO code equal to 106 (code of Italy) we extract the Italian cells where maize

is cultivated. These cells are inserted in a 3760 rows array, which reports only the Italian cells which contains both positive values of yield and harvested surface.

The total Italian surface for maize crops is then calculated by summing the values of the vector of harvested areas:

$$Area_{tot}[ha] = A_1 + A_2 + A_3 + \dots + A_n \quad (3.1)$$

Later, the national maize production is obtained by the product of the yield and area vectors, obtaining the total tons of maize for each cell:

$$P[ton] = (Y_1 * A_1) + (Y_2 * A_2) + \dots + (Y_n * A_n) \quad (3.2)$$

At this point it is possible to derive the mean national Italian yield of 2000 by dividing total production with total harvested area:

$$\bar{Y}_{italy} \left[\frac{ton}{ha} \right] = \frac{P}{Area_{tot}} \quad (3.3)$$

Here the results for Italy:

Table 3. 1: results for Italy

Country	Mean National Yield (ton/ha)	Minimum Yield (ton/ha)	Maximum Yield (ton/ha)	Number of cells
Italy	9.6	1.03	11.7	3760

3.1.2 Rest of the world

At this point we can repeat the same procedure to all world's countries. The goal is to create a dataframe, *Summary_matrix*, with N rows, where N is the total number of countries where maize is grown and 10 columns, in which we have:

Column 1: FAO code,

Column 2: mean national yield

Columns 3, 4, 5: Min, Max and total number of cells

Column 6: total area

Columns 7, 8, 9: checking of empty countries

Column 10: mean national yield according to FAOSTAT

Data in the last column of the matrix refer to the mean national values of yield of world's countries according to FAOSTAT, the principal database of FAO, which contains a large amount of agro-food related indicators and rankings of all countries from 1960 to present days. By comparing processed data with FAOSTAT data we can understand if there are discrepancies between GAEZ and FAOSTAT datasets. The procedure is carried out with a *for* loop in Matlab, repeating the steps done for Italy. However, it must be noted that not all countries contain cells of yield and/or area, and there may be countries which actually contain both yields and areas, but these cells don't match each other, therefore columns 7, 8, 9 are dedicated to the check of empty countries, using a binary system. Inside the *for* loop, we insert an additional condition, which says that if there is at least one value of yield in the country, column 7 will reports "1", otherwise it will put "0". The same concept applies for harvested areas in column 8, while column 9 reports "1" if there is at least 1 pixel in the country which contains both a positive value of yield and area.

Once the matrix is built, we use a 3-dimension scatter plot to compare countries' mean national yields obtained from data processing with FAOSTAT data, checking for possible discrepancies. Each dot corresponds to a country (numbers are the FAO codes related to that country) and they have different dimensions depending on the number of hectares dedicated to maize crops. The plot is repeated also for harvested areas (this time, it is only a 2D scatter plot).

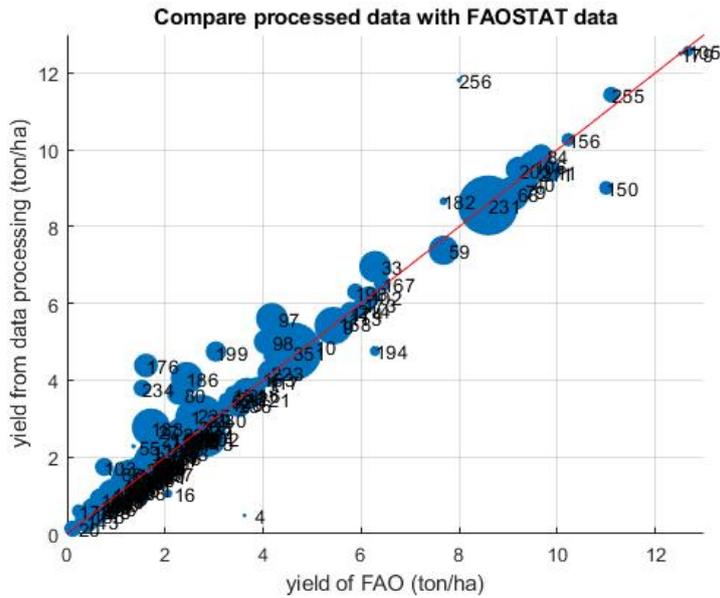


Figure 3. 1: scatter plot of yield data comparing between GAEZ and FAOSTAT datasets

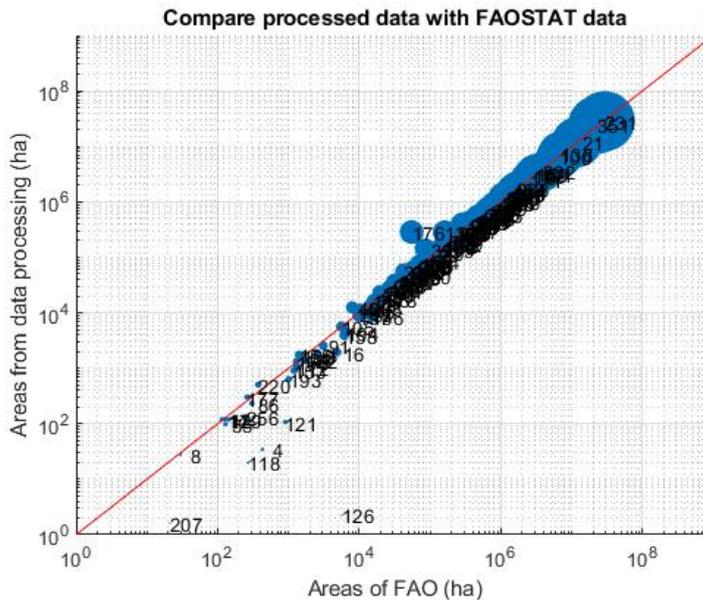


Figure 3. 2: scatter plot of harvested area data comparing between GAEZ and FAOSTAT datasets

At a first glance, the majority of countries are well aligned with the bisector, especially countries with vast harvested areas (China, US...), but there are other nations which significantly deviate from the bisector, in particular those with low yields and areas. To better see the misaligned countries the relative error between processed and FAOSTAT data is computed, obtained with the equation

$$Relative\ error = \frac{Y_{2000,GAEZ}^N - Y_{2000,FAOSTAT}^N}{Y_{2000,FAOSTAT}^N}$$

(3. 4)

where $Y_{2000,FAOSTAT}^N$ is the mean national yield of country N in year 2000 according to FAOSTAT, while $Y_{2000,GAEZ}^N$ is the yield obtained from data processing of GAEZ matrix. Then, countries with an error greater than 50% are isolated.

Table 3. 2: countries with error greater than 50%, yield

Country name	FAO code	National yield computed by GAEZ (ton/ha)	FAOSTAT yield (ton/ha)	Error (%)
Bulgaria	27	2.68	1.72	55.32
Dominica	55	2.28	1.36	67.93
Bosnia and Herzegovina	80	3.65	2.27	60.93
Grenada	86	1.58	1	56.99
Iraq	103	1.74	0.76	130.37
Timor-Leste	176	4.38	1.61	172.64
Eritrea	178	0.59	0.24	152.07
Romania	183	2.77	1.71	61.71
Serbia and Montenegro	186	4.06	2.44	66.52
Slovakia	199	4.75	3.04	56.32
Uruguay	234	3.79	1.53	148

Table 3. 3: countries with error greater than 50%, area

Country name	FAO code	Harvested Area computed by GAEZ (ha)	FAOSTAT area (ton/ha)	Error (%)
Algeria	4	34.4	430	-92
Bangladesh	16	1939.2	4855	-60.06
Chad	39	143192	85014	68.43
Congo	46	12517.1	7944	57.57
Kuwait	118	20	270	-92.59
Lebanon	121	107.4	900	-88.06
Lithuania	126	2.3	5659	-99.96
Mali	133	284500.5	161053	76.65
Timor-Leste	176	285207.9	55000	418.56

It can be seen that the values reported in Table 3. 2 and Table 3. 3 show very low values of mean national yield (they don't exceed 4.5 ton/ha) and rather low harvested areas. Anyway, most of codes in these two tables refer to developing countries in the Middle East, Africa and South America, where data collection is rather complex and not always reliable.

3.2 National yield and area in time

Now we create the global maps of yield and area from 1961 to 2019, to see the evolution of world's maize cultivations in the past 60 years.

To perform this calculation, it is first necessary to download from FAOSTAT the national yield of all countries from 1961 to 2019; these will be required to extend the previous analysis of year 2000 to the rest of the time period.

The procedure is again carried out in a double *for* loop in Matlab which iterates the calculation for every year from 1961 to 2019 and for all countries of the world. Thus, at each iteration, we identify the position of the cells corresponding to a certain country (according to its FAO code) and which contains positive values of yield and harvested areas, then, with historical mean national values for that country for every year we execute the following operation:

$$Y_{(i,j),yr} = Y_{(i,j),2000} * \frac{\bar{Y}_{c,yr}}{\bar{Y}_{c,2000}}$$

(3. 5)

In this equation, the value of yield of the pixel in position (i, j) at the year yr will be the result of the product between the value of yield of the pixel in position (i, j) in year 2000 and a correction factor, that is the ratio between the FAOSTAT national yield of the country in which the pixel is located at the i -th year and the national yield of 2000 according to the data processing performed in the section *National Yield*. This correction factor is non-dimensional and allows the extension of the year 2000 GAEZ map to the past and future years of the dataset, relating FAOSTAT data (the numerator, which incorporates all the history of maize cultivation in that country, such as technological developments, crises and famines, political assets etc.) with year 2000 GAEZ processed data.

We can repeat the same procedure for the creation of the historical global map of harvested areas. For every year (1961-2019) and for every country (Albania-Zimbabwe) we find the belonging pixels of area which have both yield and area greater than zero, then, using equation (3. 6) we multiply these terms with the correction factor, in this case the total area of a country in year (k) over the total area of the same country in 2000:

$$A_{(i,j) yr} = A_{(i,j)2000} * \frac{A_{tot,c,yr}}{A_{tot,c,2000}}$$

(3. 6)

3.2.1 Country variations in time

Nevertheless, the procedure may result incomplete if we leave it as it is. In fact, if we display the historical global maps prior to the 90s, we notice that some nations become blank before a particular year. This is simply because those states didn't exist prior to that year but were incorporated inside wider confederations. In particular, this is the case of:

- Belgium-Luxembourg, disaggregated in 2000
- Czechoslovakia, disaggregated in 1993
- Soviet Union (USSR), disaggregated in 1991
- People's Democratic Republic of Ethiopia (PDRE), disaggregated in 1992
- Yugoslavia, disaggregated in 1991

Therefore, in order to complete the missing spaces of the historical maps, we need to use the mean national yields of the previously listed federations from 1961 to the year in which they disaggregated. For example, to display the evolution of the yield of maize in Ukraine from 1961 to 2019, it is necessary to use in the equation (3. 6) the mean yield of USSR from 1961 to 1991, then the mean national yield of Ukraine from 1991 to 2019. Undoubtedly, yields in Ukrainian pixels prior to 1991 are supposed to be much less accurate than those subsequent to USSR disaggregation. In fact, the mean national yield of Soviet Union takes in account a much wider number of states, with countries having completely different cultivation techniques and climate (for example, the countries in Central Asia, Siberia, Baltic countries etc.), thus the pixels of Ukraine are multiplied by a weighted mean yield which does not correspond only to that geographic area.

The procedure starts by finding all today's countries which were part of the listed federations and their FAO codes, then they are grouped together therefore forming the ex-unions. In particular:

- Belgium-Luxembourg (FAO code 15), which is the sum of Belgium (FAO code 255) and Luxembourg (FAO code 256)
- Czechoslovakia (FAO code 51), which is the sum of Czechia (FAO code 167) and Slovakia (FAO code 199)
- USSR (FAO code 228), which is the sum of Armenia (FAO code 1), Azerbaijan (FAO code 52), Belarus (FAO code 57), Estonia (FAO code 63), Georgia (FAO code 73), Kazakhstan (FAO code 108), Kyrgyzstan (FAO code 113), Latvia (FAO code 119), Lithuania (FAO code 126), Moldova (FAO code 146), Russian Federation (FAO code 185), Tajikistan (FAO code 208), Turkmenistan (FAO code 213), Ukraine (FAO code 230), Uzbekistan (FAO code 235)
- People's Democratic Republic of Ethiopia (PDRE) (FAO code 62), which is the sum of Eritrea (FAO code 178) and Ethiopia (FAO code 238)
- Yugoslavia (FAO code 248), which is the sum of Bosnia and Herzegovina (FAO code 80), Croatia (FAO code 98), North Macedonia (FAO code 154), Serbia (FAO code 272), Montenegro (FAO code 273), Slovenia (FAO code 198)

From FAOSTAT, the list of historical annual mean yields of disaggregated countries is downloaded, converted from *.csv* format to *.mat* and imported in Matlab. Then, the matrix positions of the today's cells of previously listed countries are isolated by the use of the *find()* command, with the conditions of selecting pixels with positive yield and harvested area and FAO code corresponding to the country of interest. Later, these arrays of positions (*idx255*, *idx256* ...) are grouped in larger vectors, according to the federation which they were part of, in turn included in a structure named *index*:

```
index.idx15=[idx255; idx256];
index.idx51=[idx167; idx199];
index.idx62=[idx178; idx238];
index.idx228=[idx1; idx52; idx57; idx63; idx73; idx108; idx113; idx63; idx119; idx126;
idx146; idx185; idx208; idx213; idx230; idx235];
index.idx248=[idx186; idx98; idx198; idx154; idx80];
```

At this point the *for* loop can start. For every ex-federation and for every year (from 1961 to the year of disaggregation) the code identifies the pixels of yield and harvested area of that federation, thanks to the new indexes inside the structure, then, from these yields, the national yield of 2000 is calculated using the equation

$$\bar{Y}_{fdr} \left[\frac{ton}{ha} \right] = \frac{P}{Area_{tot}} \quad (3.7)$$

where $Area_{tot}$ is the sum of all harvested areas for that country and P is the sum of the products of yield and harvested area in positions (i, j) . Even though these federations didn't exist in 2000 anymore, we still need to exploit this expedient to obtain the average yield of such group of nations in 2000, which will constitute the denominator of the correction factor that will be used later.

Therefore, we can reproduce the historical maps of yield and area by the use of the correction factor:

$$Y_{(i,j) yr} = Y_{(i,j)2000} * \frac{\bar{Y}_{fdr(m),yr}}{\bar{Y}_{fdr(m),2000}} \quad (3.8)$$

$$A_{(i,j) yr} = A_{(i,j)2000} * \frac{A_{tot fdr(m),yr}}{A_{tot fdr(m),2000}} \quad (3.9)$$

Results are then saved in two structures, one for yield and the other for harvested areas, containing 59 elements and where each element is a 2160 x 4320 global map of yield or area for the year *yr*.

For the display, we rely on ArcGIS software. The maps of 1961, 1970, 2000 and 2019 are chosen for the purpose. Firstly, they need to be converted into a *.txt* file to be imported on the software. In Matlab, a specific function, *txt_per_QGIS*, converts the matrices from *.mat* into a *.txt* file that can be read by ArcGIS. The files are then imported on GIS and, after a little processing on the *Symbology* of the *Layer Properties* tab they can finally be represented. For a correct and satisfying display of the features inside the map, we choose the *Classified* option, we select 5 classes, and we use the *Manual* classification method for the intervals, which allows to manually select the break values of the classes.

3.3 Yield and area evolution

Now we want to understand how yield and harvested area evolved in the past 50 years in some specific countries, that is in the United States, Italy, Australia, Nigeria and Viet Nam. The choice about these nations is due the fact that we want one representative country per continent, with a sufficient number of harvested land hectares and reasonable values of yield. In addition, the dimensions of the chosen states, such as their climate, geomorphology, agricultural techniques and their relationship with maize crop allow us to have different conditions in the patterns of yield and harvested area. Then, in order to better see the evolution of the trend, we display the linear regression line, a linear model for the interpolation of the data.

Regarding yield, the yearly national yield for each country from 1961 to 2019 is plotted versus time. In Matlab, for the interpolation procedure, with the *polyfit* function we find the coefficients to be used for the linear model, applying the Least Square Method. Then, with *polyval* function, the independent variable (years), is used together with the derived coefficients to find the predicted y , that is the linear regression model.

For harvested areas, we take the sum of all the surfaces for that country for each year and we plot the result versus time, then, for the interpolation procedure, we repeat the steps done for the yield.

Here the results for yield and area:

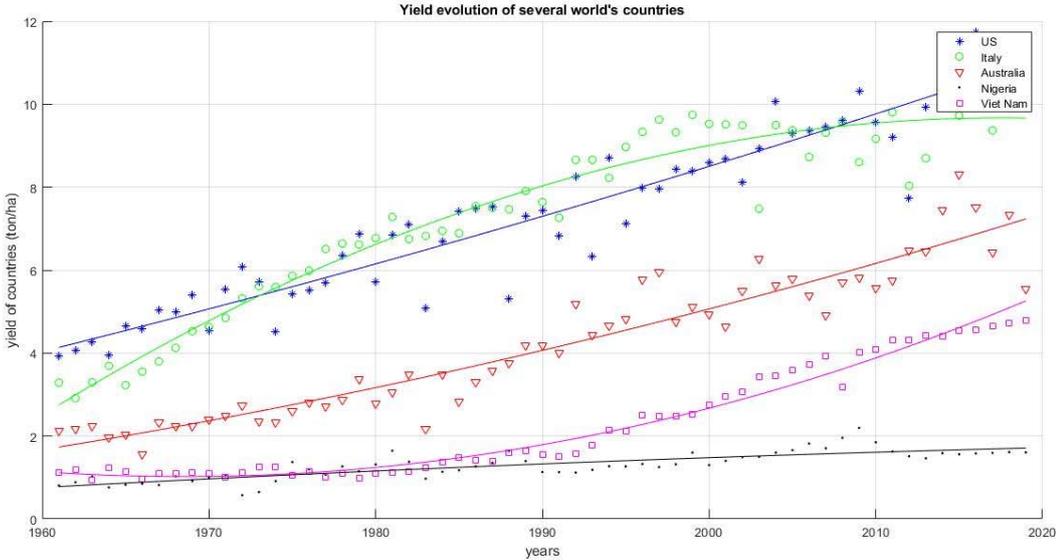


Figure 3. 3: time series of mean national yield of Italy, US, Australia, Nigeria and Viet Nam

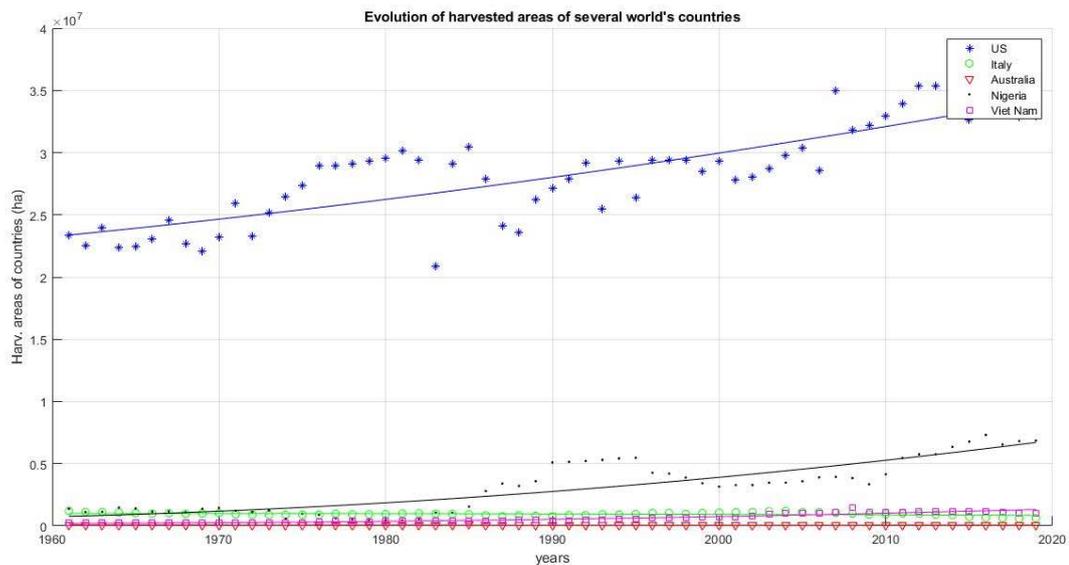


Figure 3. 4: time series of national harvested area of Italy, US, Australia, Nigeria and Viet Nam

Looking at the first illustration, it can be clearly observed an increasing trend in all countries. They all started in 1961 with mean national values lower than 4 ton/ha, then they gradually increased through time, reaching values which are twice or three times as much as their initial value in 2019.

For example, US had a mean national yield of approximately 4 ton/ha in 1961, then it continuously raised up to 11 ton/ha in 2019, with a trend that can be approximated as a straight line which has the intention to grow also in the future. Different topic for Italy, which generally follows similar values in terms of yield compared to US, even better in the years between 1975 and 2005, however it seems that its trend reached a sort of plateau from 1995, where values fluctuate between 7.5 and 10.5 ton/ha, but without a growth of the trendline. Australia, instead, follows a very similar pattern to US: a straight line from 2 to 7.5 ton/ha, that means that in 60 years the national yield of the country tripled. Furthermore, the tendency of Viet Nam shows a quadratic pattern, meaning that in the next years it will likely be the one with the highest rates of growth compared to all other countries. Nigeria, instead, has always been the nation with the lowest yield and with the lowest growth rate: up to the mid '80s, its values were comparable with those of Viet Nam, nevertheless the values of the former still remained low while those of

the latter significantly raised in the following decades. Even though Nigeria's yield is growing (it almost doubled in 60 years), it never exceeded 2 ton/ha.

Regarding harvested areas, it is clear that US is the nation that leads among the others: due to its land morphology (vast flatlands in the mid of the country) and to its dimensions (3rd largest country in the world), it owns a maize surface that has always exceeded 20 million hectares, while all other nations remained below 7 million (Italy, Australia and Viet Nam also below 2 million hectares).

Here, Nigeria reveals to be the second greatest country in terms of maize land cover extension: in fact, up to the mid '80s it showed comparable values to the other nations (US excluded), then the curve lifted to 7 million hectares, following a very similar rate to US. Throughout the last 60 years, Italy and Australia didn't show evident growth in terms of harvested area, while the curve of Viet Nam manifests a very mild increase, but always keeping values well below 5 million hectares.

3.4 Distributed yield area and production in time

At a first glance on the maps of Figure 3. 5 and Figure 3. 6, it can be noticed that both yield and harvested area globally increased; the first factor is mainly driven by the great technological improvements that have been put in place, which allow a maximisation of production with reduced use of water resources, while the increase of the latter is caused by the rise of food demand, determined by a continuously growing population and by a general improvement of the quality of life.

Regarding yield, it can be seen that US and Europe are the owners of the brightest cells, i.e., the highest values, followed by countries in the Austral Hemisphere (such as Brazil, Australia and New Zealand), while most of other nations don't show values greater than 8 ton/ha, especially in Africa, where also in 2019 many countries have yields which don't reach 3 ton/ha. In the harvested surface's maps, instead, we can clearly highlight four clusters with very high harvested area values: Central US (Nebraska, Minnesota, Wisconsin, Illinois...), Brazil, Eastern Europe (Romania, Moldova, Ukraine) and Northern China (around Yellow Sea and in

Manchuria). These pixels always had a high amount of harvested area, and, over time, they reached values between 1500 and 4000 ha (remembering that the side of each pixel is approximately 8.3 km at the equator). Also, Brazil and Nigeria increased their harvested surfaces over time, thus entering in the list of the greatest world's producers, however not reaching the same values per pixel of the three clusters.

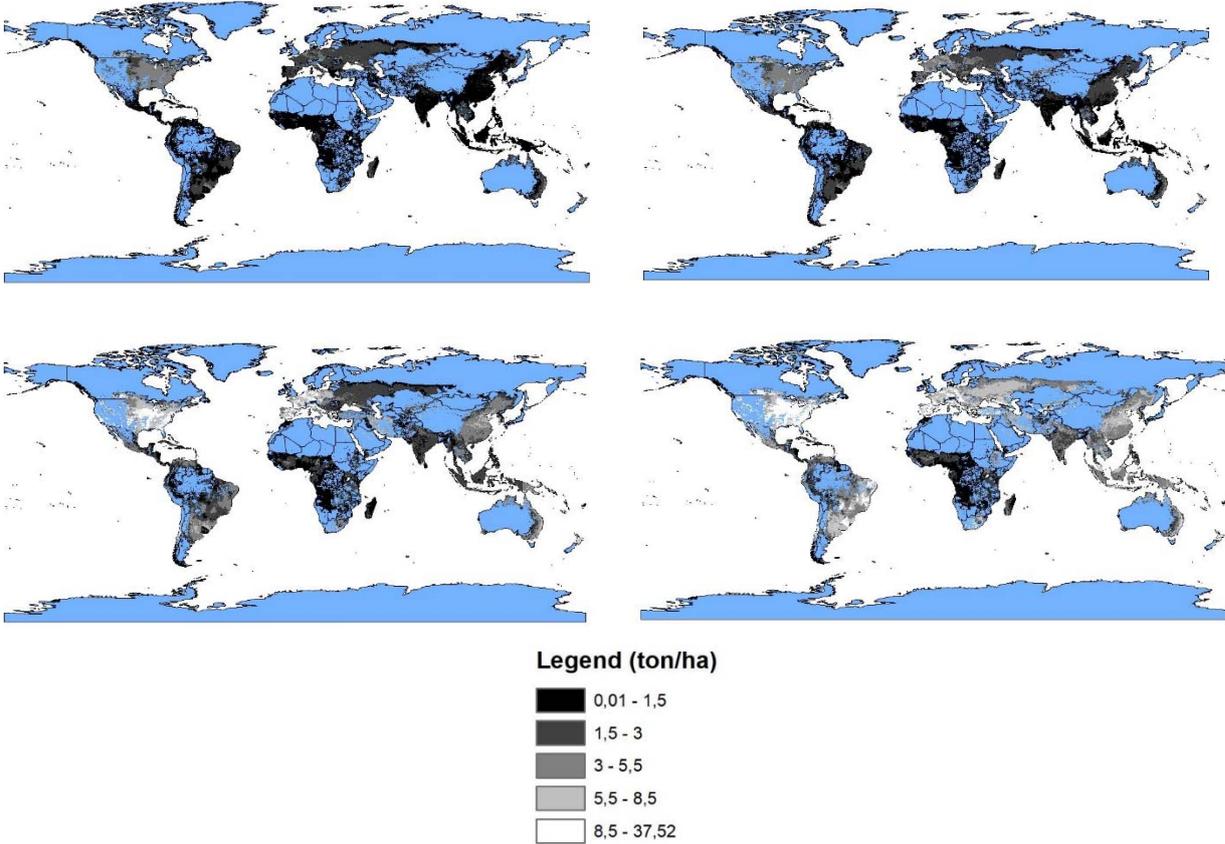
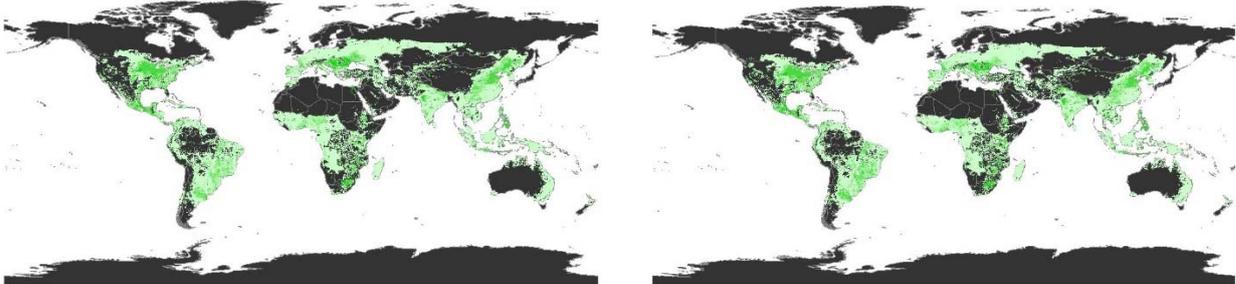


Figure 3. 5: maps of yield of years 1961, 1970, 2000, 2019



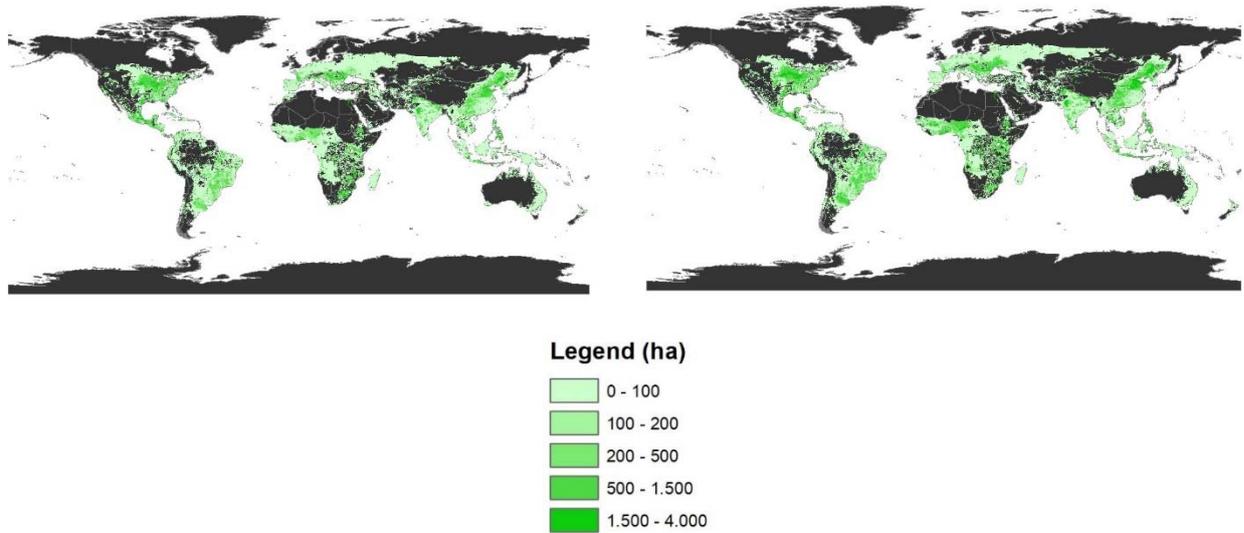


Figure 3. 6: maps of harvested area of years, 1961,1970, 2000, 2019

3.5 Maps of production

Having the maps of yield and harvested surfaces, the global maize production maps can be easily created.

Production is the product between yield $\left[\frac{ton}{ha}\right]$ and area $[ha]$, that is the number of tons produced in a given cell in a given year.

In the code, the procedure is carried on by a triple *for* loop, where for each year “a” the cells of yield at the position (i,j) is multiplied with the cells of area (i,j) . Inside these three *for* loops there is also an *if* condition, that says that the multiplication only occurs if both of cells contain values greater than zero, otherwise the result will be just -9999 (i.e., non-productive, or water-covered area).

Here the equation of the process:

$$P_{(i,j)year(a)} = Y_{(i,j)year(a)} * A_{(i,j)year(a)} \quad (3.10)$$

As for yield and harvested area, the maps of 1961, 1970, 2000 and 2019 are chosen to be displayed. They are exported as a *.txt* file and imported on ArcGIS, where, after few steps of processing on the *Layer Properties - Symbology* tab, the maps can finally be represented.

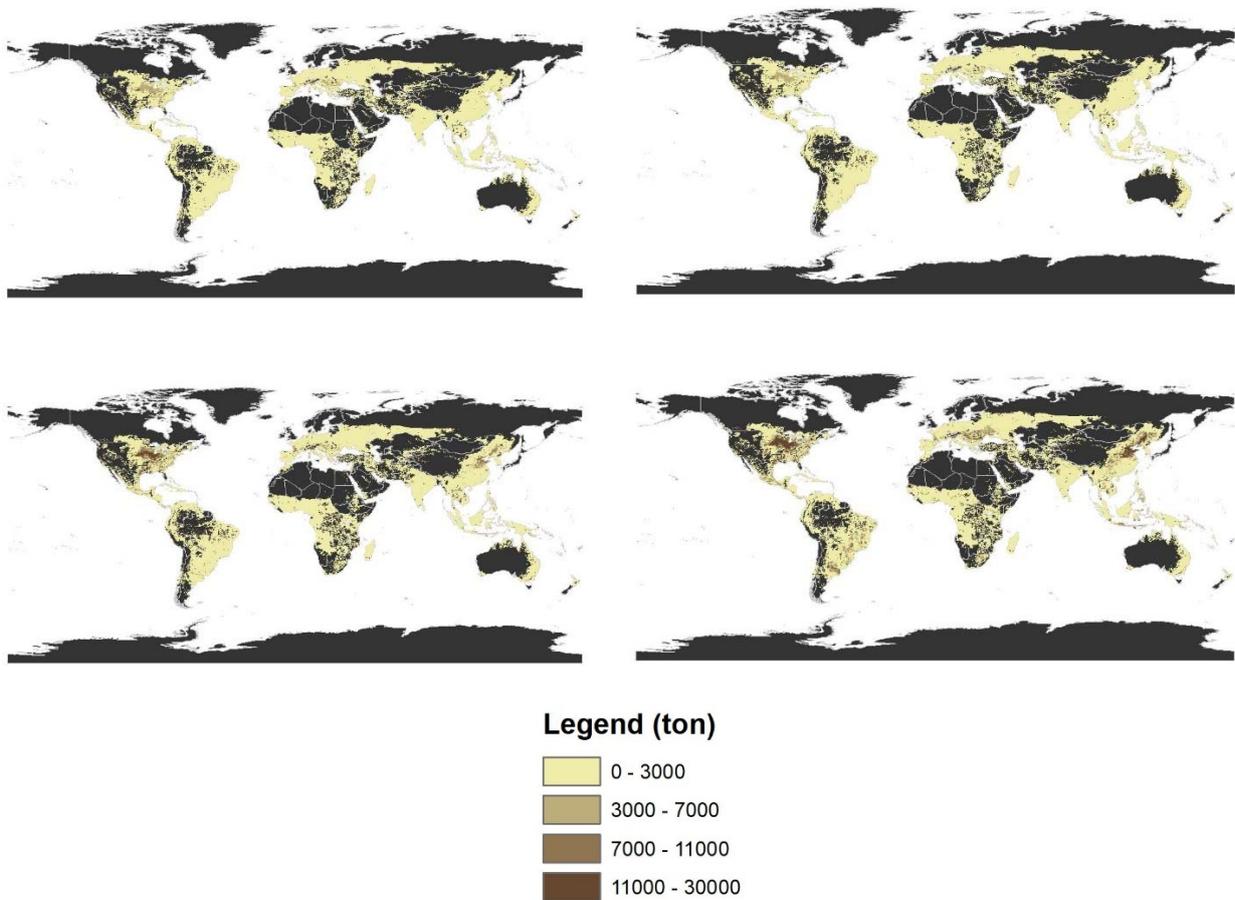


Figure 3. 7: maps of production of years 1961, 1970, 2000, 2019

Looking at the maps of production, the visibility of the clusters highlighted in the maps of harvested area is clearly evident. In fact, we can see that in US, Eastern Europe and China many pixels contain values of production that exceed 10^5 ton. The rest of the world, apart from few other areas in Brazil, India and Argentina, which contain little clusters of high production pixels, shows values which are lower than 3000-7000 tons per pixel. These results are coherent with what can be found in official statistics: Investopedia, such as STATISTA and FAOSTAT, says that the three largest maize producers are US, China and Brazil.

It is also worth to notice how much these clusters increased over time: from a monochromatic flat condition in 1961 map, we assisted to a continuous darkening of these areas during the time-interval.

3.6 Comparison between estimated map of year 2010 and GAEZ map of year 2010

The time series of yield and harvested area, derived from the data processing described in the previous parts, are both based on the GAEZ matrix of year 2000. In fact, all the maps of maize's yield and area from 1961 to 2019 are referred to how they were distributed in 2000, then the correction factor, which is based on the national time series of yield and harvested area, modifies the values of 2000 to those of the given year. However, even in the past and also in the following years, the distribution of pixels in the global maps remains the same as it was in 2000, the only things that change are the values inside each cell. However, between 2000 and 2010, it may have happened that, worldwide, new maize cultivations appeared in places which never had such crop, plus there may have been crop switches (change type of crop within a region), crop migrations (moving maize to more favourable areas) or just some political actions that changed the distribution of maize areas. On the global maps, in fact, we would expect not only a change in the pixel colour, but also the pop-up and shutdown of different pixels. We compare the statistics of matrix *Yield_maize_2010* (i.e., the GAEZ matrix of 2000 corrected with the correction factor) with the 2010 matrix directly downloaded from GAEZ Data Portal. Here the table with the results:

Table 3. 4: data and statistics of reconstructed vs GAEZ 2010, yield

YIELD	Reconstructed 2010	GAEZ 2010
Mean (ton/ha)	4.157	4.150
Min (ton/ha)	0.056	0.088
Max (ton/ha)	35.113	32.253
Mode (ton/ha)	2.680	2.592
Stdv	2.854	2.826
Variance	8.145	7.987
Number of cells	769450	764991

Table 3. 5: data and statistics of reconstructed vs GAEZ 2010, area

AREAS	Reconstructed 2010	GAEZ 2010
Mean (ha)	213.9	215.7
Min (ha)	0.0	0.0
Max (ha)	58201.4	8842.3
Mode	0.1	0.8
Stdv	424.3	417.0
Variance	180032.4	173854.5
Number of cells	769450	764991

In the first column of the two tables, we have the results regarding the reconstructed matrix of 2010, while in the second, the GAEZ map of 2010.

We notice that, generally, the values of the two matrices are similar: the global mean of yield and harvested areas are very similar, especially due to the very high number of cells which reduces the influence of potential outliers. Also, the differences between all other indices in the yield table are almost negligible, just the maximum differs of more than 2 ton/ha. In the harvested areas table, instead, we see that there are evident differences in the maximum, almost 50000 hectares of difference between the reconstructed 2010 and GAEZ 2010. This probably means that, in the global map of harvested areas in year 2000, there were one or more outliers that have then been corrected in the later version.

However, the matrix of the reconstructed 2010 contains more cells than the GAEZ 2010, 769450 and 764991, respectively, therefore a difference of 4459 cells. This doesn't mean that between 2000 (the reference matrix of the reconstructed 2010) and 2010 4459 cells just disappeared; these are the net result of the shutdown and pop-up of many more pixels across the map.

3.7 Temporal variations in GAEZ data (cell switches)

It is worth to check which and where are the switched off cells (cells that contain values in the GAEZ map of year 2000 and no data in the GAEZ map of year 2010) and switched on cells (cells that are empty in the GAEZ map of year 2000 and contain values in the GAEZ map of

year 2010) on the maps, to better understand which have been the effective changes between 2000 and 2010, which are the most affected countries, the geographical areas, the average latitudes etc.

The procedure starts from the creation of two logical matrices, one for *Yield_maize_2010* (the processed matrix) and one for *mze_2010_yld* (GAEZ), with the condition of finding the values greater than zero. In Matlab, *logical* is a function that converts numeric values into an array of logicals, that is 0 and 1, given a certain condition. As previously said, here the condition consists in finding all positive elements of *Yield_maize_2010* and *mze_2010_yld*, so three new files are created:

- *matrix1*: logical array of *Yield_maize_2010*
- *matrix2*: logical array of *mze_2010_yld*
- *matrix3*: *matrix1-matrix2*

Thus, the values displayed in the pixels of *matrix3* can only be +1, 0, -1, in particular:

- Pixels = +1 correspond to the switched off cells in 2010 (they were on in the GAEZ matrix of 2000, but they shut down in GAEZ 2010)
- Pixels = 0 correspond to the cells with no data or to those which contain both values in 2000 and 2010
- Pixels = -1 correspond to the switched-on cells in 2010 (they were off in the GAEZ matrix of 2000, but they pop-up in GAEZ 2010).

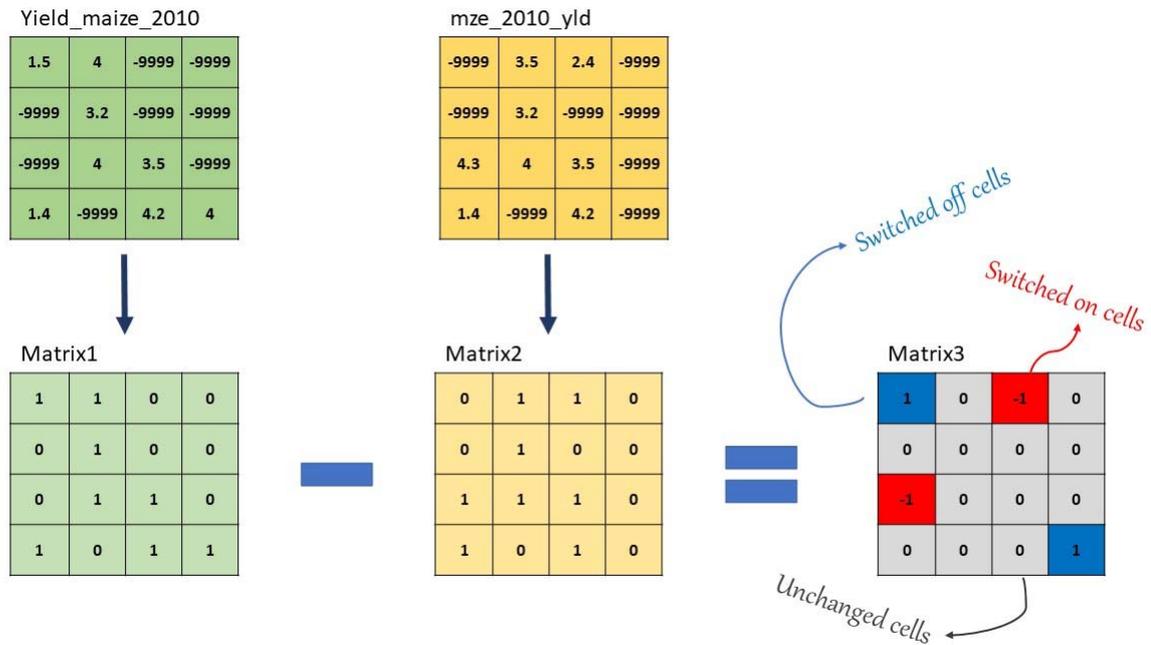


Figure 3. 8: identification schema of switched off and on cells between GAEZ 2000 and 2010

In the Matlab code, once created *matrix3*, we find the position of the cells equal to +1 (positive values) and those equal to -1 (negative values) with the *find* function:

$index3=find(matrix3(:,>0))$ (positions of switched off cells of 2000)

$index4=find(matrix3(:,<0))$ (positions of switched on cells of 2010 (GAEZ))

Therefore, the switched on and off cells in the original matrices will be:

$yields_off=Yield_maize.Yield_maize_2010(index3);$

$yields_on=mze_2010_yld(index4);$

3.7.1 Distribution in switched on and off cells

We are now able to visually analyse the differences between the reconstructed map of yield (*Yield_maize_2010*) and the GAEZ map (*mze_2010_yld*).

We start creating histograms of the frequency distribution of the pixel values, to compare which were the most frequent yields that shutdown and those which appeared, selecting a number of bins equal to 50 and bin width to 0.3.

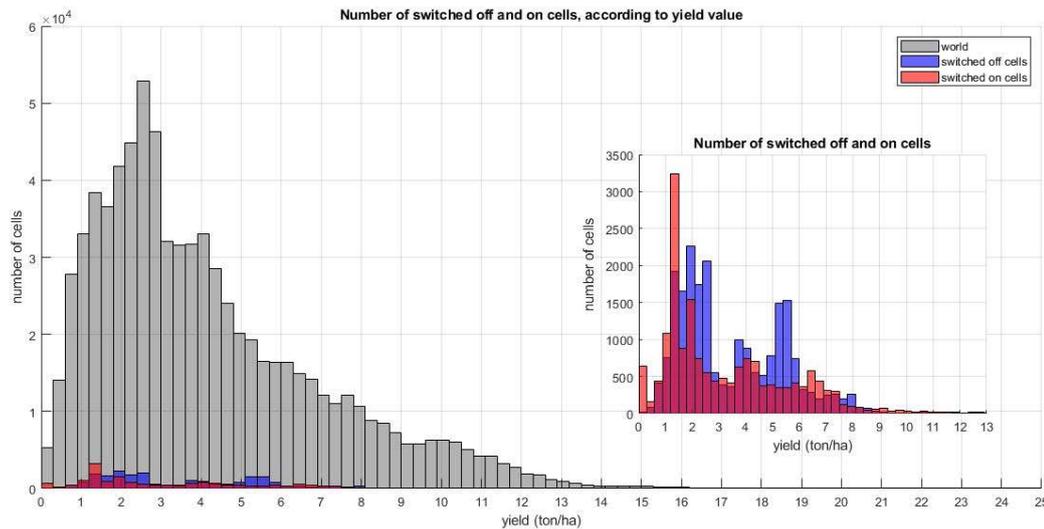


Figure 3. 9: frequency distribution of switched on and off cells

Figure 3. 9 shows the frequency distribution of the yields of the switched on and off cells together with the one of total world's pixels (in this case, the distribution of the values of *Yield_maize_2010*). At a first glance, we notice that the switched off and on cells are much less than the totality of the cells. The greatest peak is in the switched-on cells, more than 3000 cells around 1.5 ton/ha, while the peak of the switched off pixels is less sharp (2200 cells) and slightly shifted to the right (around 2-2.5 ton/ha). Since the greatest number of values of the switched-on cells is attested on low values, 1.5 ton/ha, this means that, according to GAEZ, between 2000 and 2010 one or more countries with rather low yields appeared in the map; it may be an undeveloped country or a nation with a less favourable climate (desertic, very cold etc.). Also, it is worth noticing that a considerable number of cells between 5 and 6 ton/ha turned off: this is unexpected, since these yields are rather high (greater than the global mean, approximately 4.1 ton/ha), so this may imply that one or more countries opted for a crop switch between 2000 and 2010.

3.7.2 Switches by country

Once understood the pattern of the frequency distribution of the switched on and switched off cells, it is interesting to see their spatial distribution.

We firstly start by identifying the countries containing switched on/off cells and counting how many of these cells each country contains, by using the Matlab commands *unique* and *accumarray*. The former selects the unique values in the array (i.e., the countries) and the latter returns the number of values for that country.

Then, we create a matrix where, in the first column, there is the list of all the countries which contain at least one turned on or off pixel, while in the second the amount of turned off cells and in the third the turned-on cells for that country. Since there are too many countries involved in the switch off or on of pixels (they are 107), we opt for filtering the dataset and select only the countries which contain at least more than 300 turned on or off cells inside their territory. Now, the number of countries drops to 22.

Table 3. 6: countries with high number of switched on and off cells

Country name	FAO code	Turned off cells	Turned on cells
[-]	0	0	3143
Argentina	9	1	454
Australia	10	313	106
Bangladesh	16	0	1714
Bolivia (Plurinational State of)	19	14	3499
Botswana	20	1	447
Myanmar	28	769	2
Central African Republic	37	3415	7
Denmark	54	0	847
Guyana	91	448	0
Iran (Islamic Republic of)	102	610	104
Kazakhstan	108	1669	4
Cambodia	115	0	453
Lao Peoples Democratic Republic	120	0	706
Lithuania	126	0	1303

Malaysia	131	3813	0
Papua New Guinea	168	1202	1
Peru	170	2219	3
United Republic of Tanzania	215	0	3227
Turkey	223	535	9
Venezuela (Bolivarian Republic of)	236	1884	3
China	351	3125	16

We are now able to create the bar chart: in the x axis we have the FAO code of the filtered countries, while in the y axis the number of switched-on/off cells that each of this country contain, plus, at the top of each bar, we add the percentage of these cells with respect to the total positive pixels in that country.

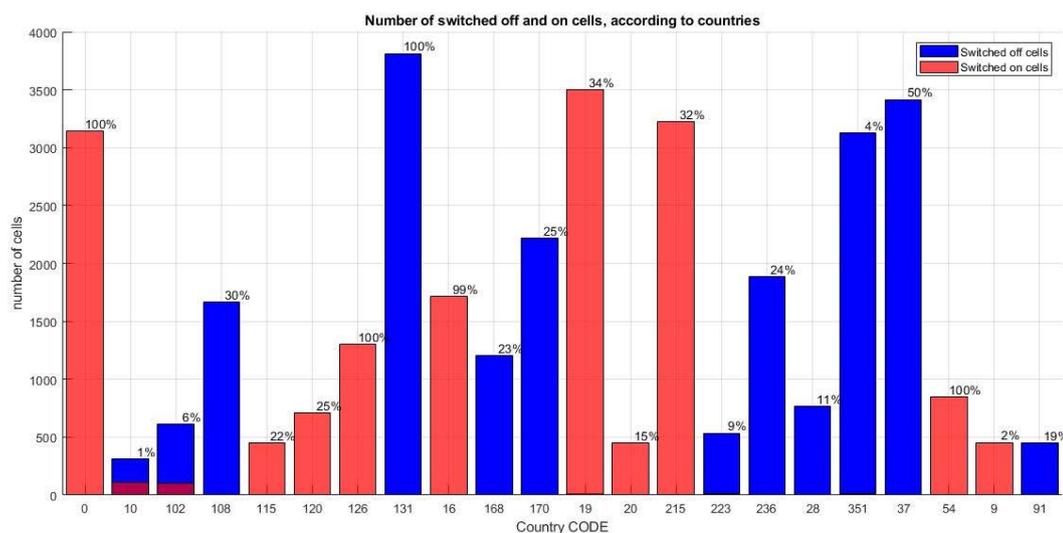


Figure 3. 10: bar chart of countries with high number of switched off and on cells

In some states, such as Bolivia (FAO code: 19) and Tanzania (FAO code: 215), a great number of pixels switched-on (more than 3000), however they remain a low portion with respect to the totality of the cells in these nations. They both own low mean national yields (2-2.5 ton/ha), so this demonstrates that many of the cells which switched on show low values (in Tanzania almost all new cells are below 1.5 ton/ha, while in Bolivia are between 0.7 and 6). Other nations, such as Bangladesh (FAO code: 16), Denmark (FAO code: 54), Lithuania (FAO code: 126) don't contain great numbers of switched-on cells (they are all below 1700), nevertheless these pixels

represent 100% of the cells in these countries: this means that, while in 2000 these nations didn't contain any cell, in GAEZ 2010 they entirely switched on, having rather high mean national yields (around 5.5 ton/ha).

In addition, it can be seen that, between 2000 and 2010, 3143 cells with FAO code equal to zero turned-on. Actually, FAO code 0 does not correspond to a country. These cells refer to the water bodies (seas, lakes, water streams), so all the pixels in the world which can't be considered land surface, thus which are not assigned to a specific FAO code. In fact, these pixels are usually located on the coastline or along the largest world's rivers (Amazon River, Nile, Volga...): according to the *map_FAOcode* matrix, these cells are supposed to be covered by water, but actually *mze_2010_yld* map indicates that they own maize production. There were cells with FAO code 0 containing values of yield also in 2000 (they were 3054) and in 2010 they became 3153, however the bar chart illustrates that they switched-on only in 2010. The reason is found on the fact that, during the processing of *Yield_maize_2010*, the cells with FAO code equal to zero have been excluded and they were assigned the value -9999 (no data), while in the data processing of *mze_2010_yld* these cells have been taken into consideration.

Regarding switched off cells, the countries which contain most of them are Malaysia (code 131), China (code 351) and Central African Republic (code 37). In the first, the switched off pixels are 100% of the totality for that country, so from 2000 to 2010 the nation opted to stop maize production. Yield distribution of switched off cells is very narrow and sharp, with rather high values with respect to China and Central African Republic yields (83% of values are between 5 and 6 ton/ha). According to FFTC Agricultural Policy Platform, maize is for Malaysia one of the most important commodities for livestock, however government has never supported local production, giving instead space to imports, believing that it's cheaper than self-production. In fact, almost 100% of grain corn is imported in Malaysia and importation growth is directly proportional to the increase of population. However, more than 95% of Malaysian maize's surface is dedicated to sweet corn harvesting, more convenient since it has lower costs of production and a higher consumers' demand.

In China, instead, despite being a large number (more than 3000), switched off pixels represent only 4% of the totality of national cells dedicated to maize crop. Here, most of switched off cells show values between 2 and 3 ton/ha, with a second peak between 3.8 – 4.2 ton/ha and, generally, a greater variance with respect to the other two countries.

Central African Republic's values show, instead, two evident and very sharp peaks between 1.5 and 2 ton/ha and between 2.3 and 2.6 ton/ha. However, most of switched off cells, particularly

those in the second peak, contain yields which are greater than the mean national yield of the country, around 2 ton/ha, so probably some kind of government strategies for a crop switch may have been also in this nation.

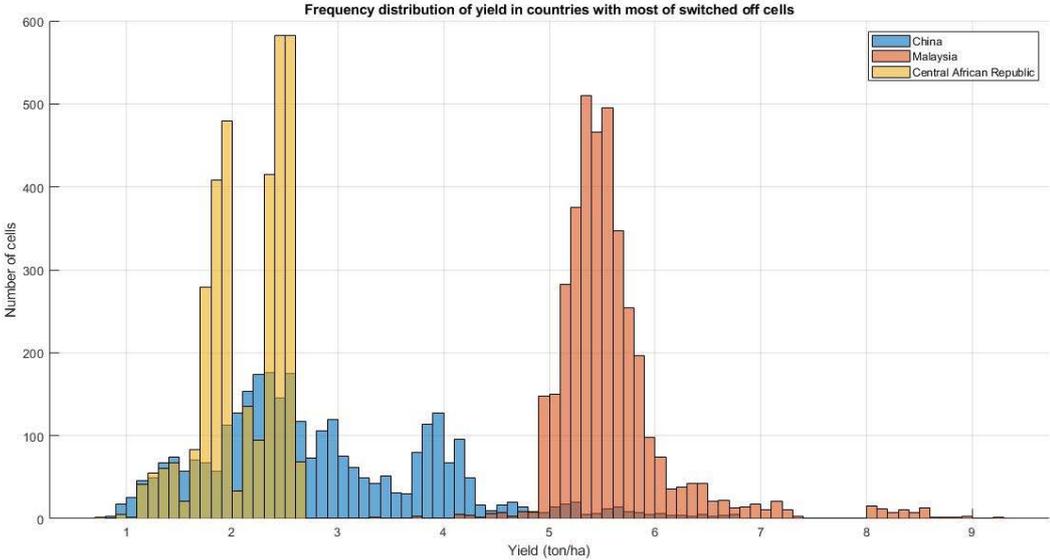


Figure 3. 11: frequency distribution of yield in countries with most of switched off cells

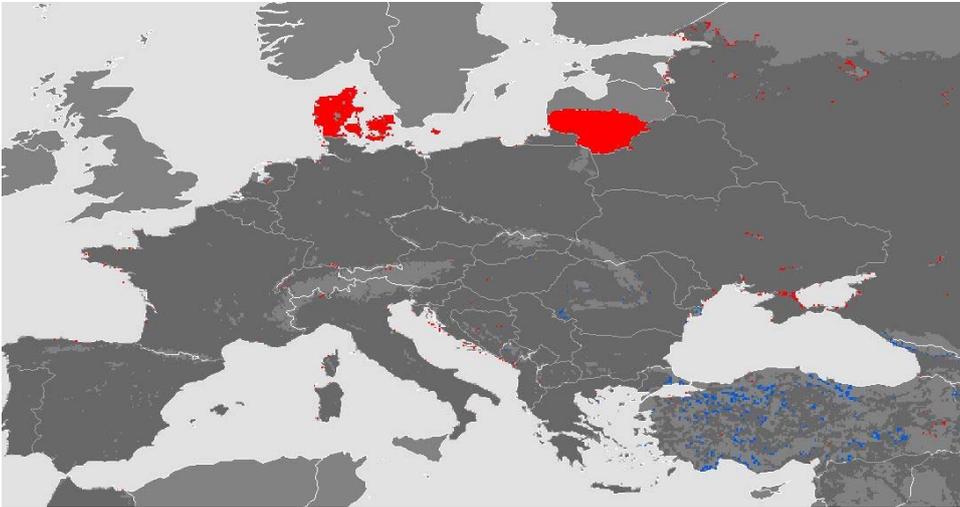


Figure 3. 12: switched on and off cells in Europe

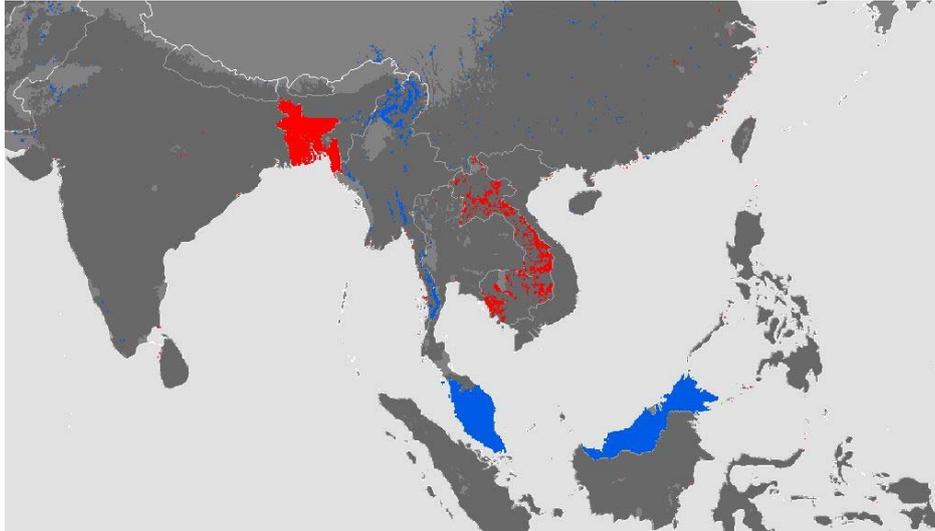


Figure 3. 13: switched on and off cells in South-East Asia

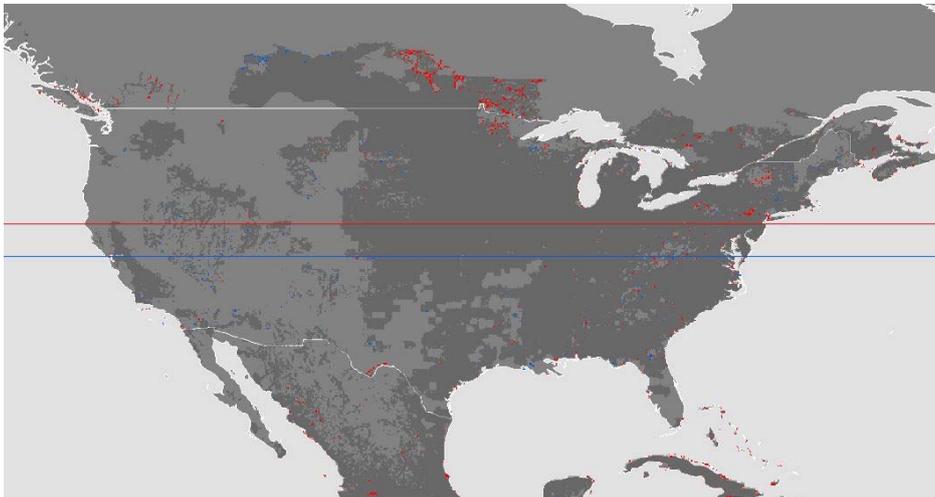


Figure 3. 14: switched on and off cells in US

The first two images represent a zoom of the map of the switched on and off cells in Europe and in Southeast Asia. In these maps it's evident the switch on and off of entire countries between 2000 and 2010, for example the switch on of Denmark, Lithuania in Europe, Bangladesh in Asia, and the switch off of Malaysia. The third map, instead, illustrates the distribution of switched off and on cells in North America; differently from the first two pictures, here the distribution of cells is more homogeneous, without the formation of big clusters, so in this case it's possible to calculate the mean latitude of the switched on and off cells. To do that, we firstly find the rows of the global map at which each cell is located, then we calculate the mean of these rows, and, in all the pixels of that row, we insert the number -1 for switched on cells and +1 for switched off. It is possible to see that the mean latitude of switched-on cells is higher than those which switched off. In this case, it could be due to the

effect of crop migration, that is the northward movement of crops as a response to a changing climate.

3.7.3 Empirical Cumulative Distribution Function (ECDF) of switches

It is worth also comparing the Empirical Cumulative Distribution Function (ECDF) of the switched off and on cells, to see how the probability of these values is distributed according to the yield.

To build the ECDF, we first need to sort all yield's values of the arrays of switched off and on cells in ascending order. Then we need to find the probability of each term i , that is the ratio between 1 and $N-1$, where N is the total number of elements in each array. We divide with $N-1$ because, theoretically, an empirical distribution is supposed to never reach 100% of probability:

$$p = \frac{1}{N - 1}$$

(3. 11)

Thus, with a *for* loop, we build a new array with the same length of the switched off and on arrays where, at each loop, we repeat the following procedure:

$$CDF(i) = p + p (i - 1)$$

(3. 12)

Which means that at each step we add the probability p plus the sum of the probabilities of all its previous steps. In this way we will obtain a graph with, at the x axis, the sorted values of the arrays and, on the y, the sequence of the ECDF. ECDF is a useful tool that enables to compare arrays with different lengths: in fact, the y axis has a relative scale ranging between 0 and 1, the lower and upper limits, thus it doesn't matter the number of elements of each array: the ECDF will just compare with a percentage the number of elements which are lower or greater than a given value of yield.

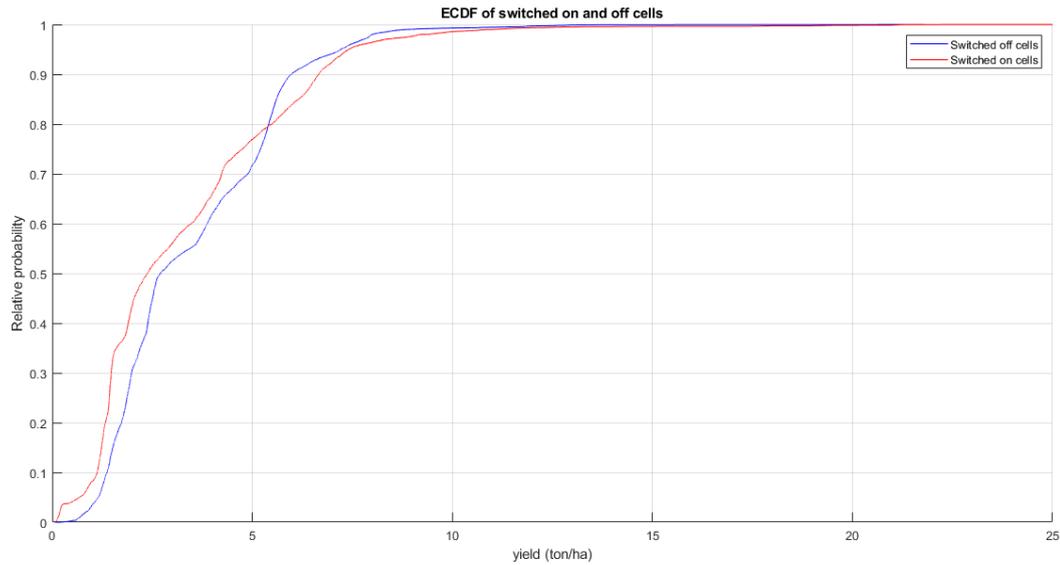


Figure 3. 15: ECDF of switched on and off cells

The plot illustrates the probability that an element is lower or greater than a certain yield. It is immediately evident that, between 0 and 5.5 ton/ha, switched-on cells are more likely than switched off cells. For example, below 2 ton/ha we have 20% of turned off cells, while regarding switched on cells, there is almost 40% of values. At 5.5 ton/ha we have an inversion of the tendency: the probability of switched off pixels is greater than the one for switched on.

3.7.4 Compare national yield and total harvested area of the reconstructed and GAEZ maps

Now it's necessary to compare the differences between *mze_2010_yld* and *Yield_maize_2010* maps at country scale. To do this, we repeat the procedure done in the section *National Yield*, during the creation of *Summary matrix*, but for the reconstructed and GAEZ matrices. This time, a new dataset named *Comparison_matrix* is built to contain the mean national yield, total harvested area, and the number of pixels in each country for both matrices.

In a *for* loop, that iterates the procedure for each country code, we firstly find the cells corresponding to that country which contain positive values of yield and area for the reconstructed matrix *Yield_maize_2010*, then we extract their values of yield and area from the maps, and we sum all the areas of each cell to get the total national area; we can repeat this

procedure also for *mze_2010_yld*. At this point, it is possible to calculate the total production (product between yield and area) and divide it with the total national area, in this way we get the mean national yield. The results of the *Comparison_matrix* are then saved as *.mat* file and are ready to be displayed.

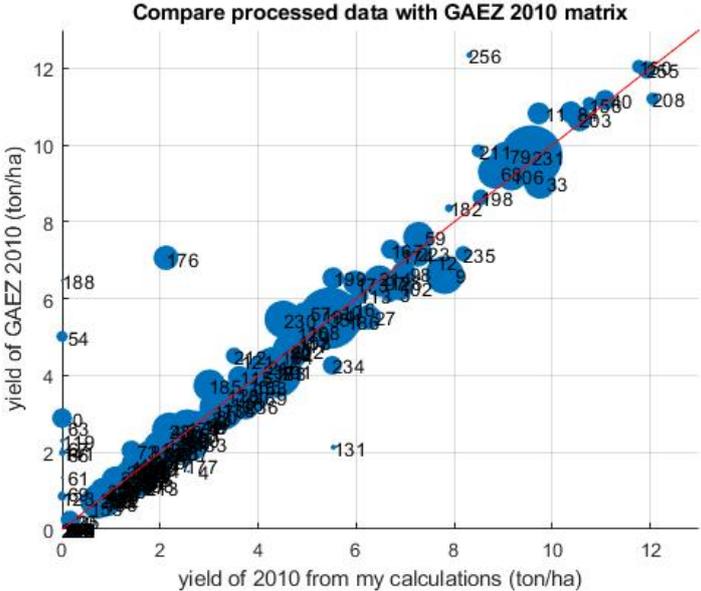


Figure 3. 16: compare reconstructed matrix vs GAEZ 2010 matrix, yield. Radius of the circles is proportional to national harvested area

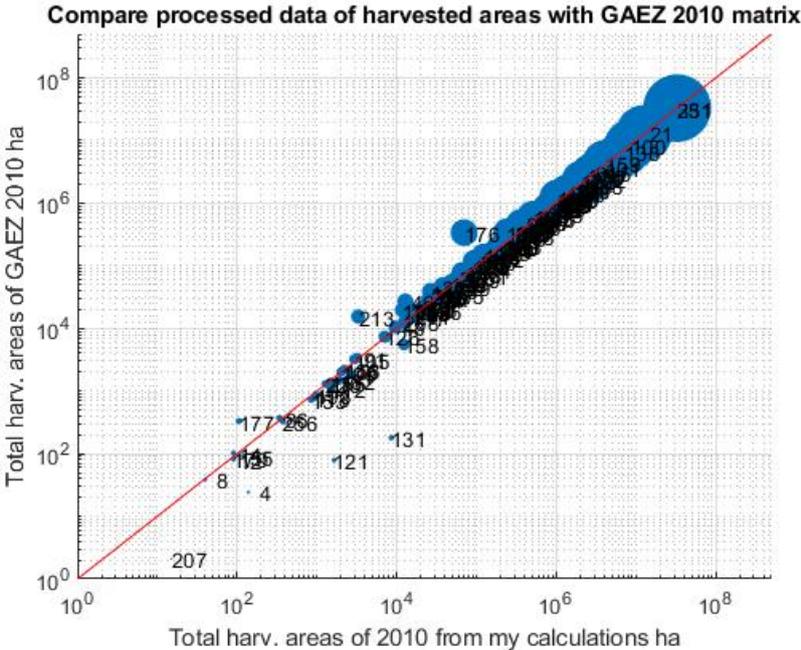


Figure 3. 17: compare reconstructed matrix vs GAEZ 2010 matrix, area. Radius of the circles is proportional to national harvested area

Looking at the scatter plot of yield, we notice that, generally, mean national yield of the countries of the reconstructed matrix are well aligned with those of GAEZ map, however there are some nations, such as Timor-Leste (FAO code: 176), Luxembourg (FAO code: 256) and Malaysia (FAO code: 131) that are not coherent between the two maps (regarding Malaysia, this is caused by the fact that, in GAEZ map, 4 cells remained on).

3.7.5 Overall comparison of yield and area pixels of reconstructed and GAEZ maps

We also create a third scatter plot, which displays the yield of all cells of both matrices, to check how they are correlated. We notice that, generally, yields are well aligned on the bisector, however there is an evident horizontal pseudo-correlation close to the x axis (circled in orange). It is worth to check where these pixels are located on the global map and if they belong to one or more countries. In addition, it may be interesting to also identify the position of the cells with greatest yields (green circle), again to check if they are concentrated in a single country or they if are spread around the world.

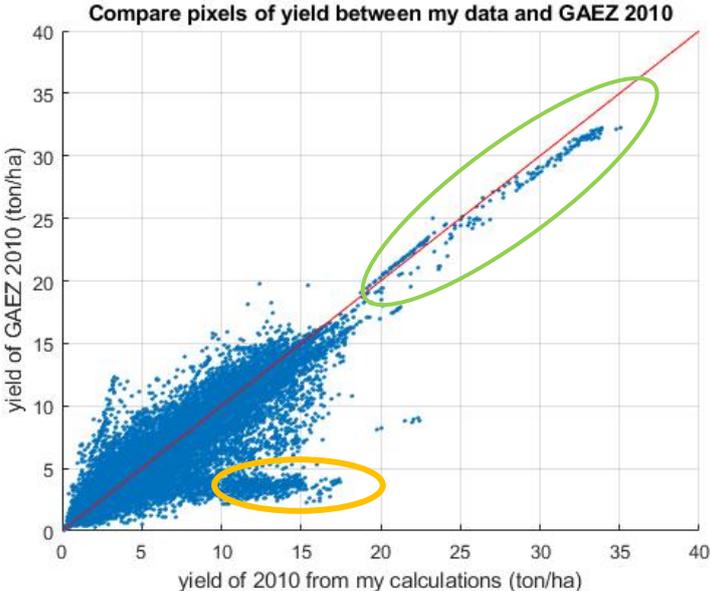


Figure 3. 18: scatter plot of all pixels of reconstructed and GAEZ 2010 matrices

We extract the cells in the orange circle by isolating the pixels by setting the condition $x > 10$ ton/ha, where x is the yield according to *Yield_maize_2010* and $y < 5$, where y is the yield according to *mze_2010_yld*. We notice that these cells are all concentrated in a region in South-Central Canada: this means that, in that area, the reconstructed matrix *Yield_maize_2010* has overestimated the yield with respect to GAEZ map.

Cells in the green circle are instead isolated by setting the condition $x > 18$ ton/ha and $y > 15$ ton/ha, then they are represented on GIS: these pixels are concentrated in some desertic countries of the Middle East, which however have very high technological standards and, consequently, very high production in a limited space. In particular, these countries are Israel, Jordan, Kuwait and Qatar.

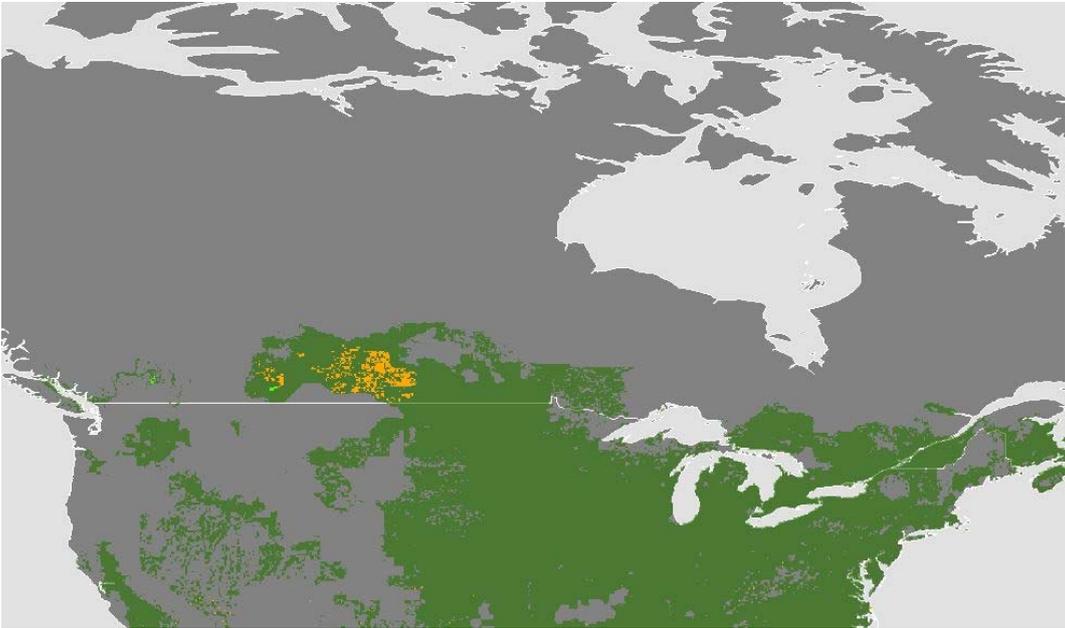


Figure 3. 19: identification of pseudo-correlation

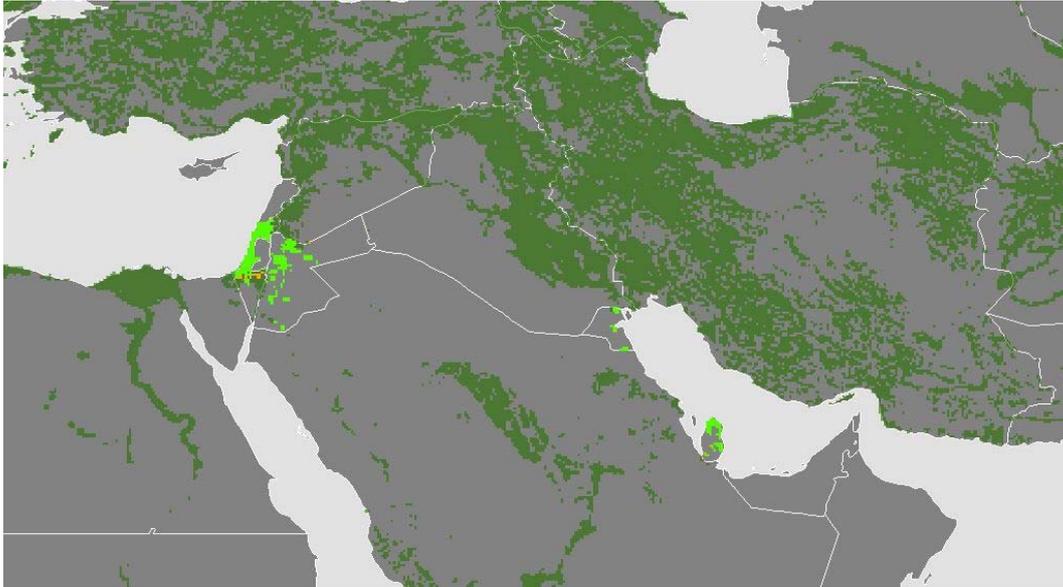


Figure 3. 20: identification of greatest yields

The scatter plot of all cells is repeated also for harvested areas, where the reconstructed map, *Area_maize_2010* and GAEZ matrix *mze_2010_area* are compared to each other. This time we don't see the presence of any pseudo correlation, all values are parallel to the bisector. Actually, some cells are not well aligned to the bisector, but there is a constant coefficient that separates them from the straight line. Probably these pixels belong to a country where the mean national yield of FAO differs from the mean national yield proposed by GAEZ.

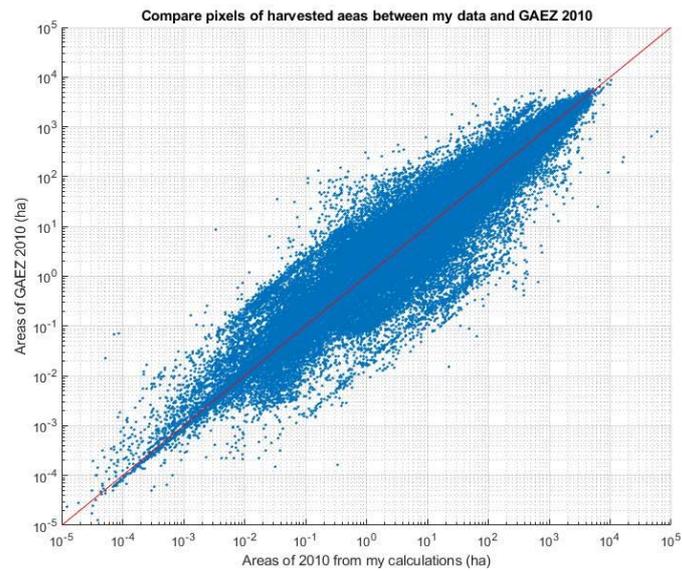


Figure 3. 21: scatter plot of all pixels of harvested area

4 Water footprint calculation

4.1 Actual evapotranspiration

After a deep analysis of the global evolution of yield and harvested areas of maize, we now pass to the data processing of crop and climatic data, in order to derive green and blue water footprint. The Water Footprint (WF) of a product is the volume of freshwater used to produce the product, measured over the full supply chain (Hoekstra et al., 2011). In our analysis, we are interested in the unit Water Footprint [m^3/ton], that expresses the amount of water necessary to produce a ton of maize, and the two parameters for the determination of the indicator are yield [ton/ha] and evapotranspiration [mm]. Now, the global evolution of yield has been derived from the data processing of the previous chapters, while the second parameter still needs to be obtained.

Under a physical point of view, evapotranspiration is the passage of water from liquid to vapor and it is the composition of two main processes, that is evaporation (from the soil or water surface) and transpiration (process that occurs inside the plant). Evapotranspiration is a parameter that depends on the local climate, so temperature, solar radiation, wind, latent heat of vaporization and then also on water availability.

There are three main types of evapotranspiration:

- Reference Evapotranspiration, ET_0 : it is the most general description of evapotranspiration, that is the evapotranspiration from a well-watered surface with a standard type of vegetation, with given dimensions and characteristics (0.12 m high, albedo equal to 0.23). Therefore, it doesn't refer to a specific crop, but only depends on weather parameters.
- Crop Evapotranspiration, ET_c : it introduces a coefficient that keeps into account crop characteristics; however, this crop is in standard conditions, with no pesticides, optimal amount of nutrients and a surface that is always well watered.

$$ET_c = k_c * ET_0$$

(4. 1)

- Actual Evapotranspiration, ET_a : this kind of evapotranspiration introduces a second coefficient that considers the presence of water stress in the crop

$$ET_a = k_s * ET_c$$

(4. 2)

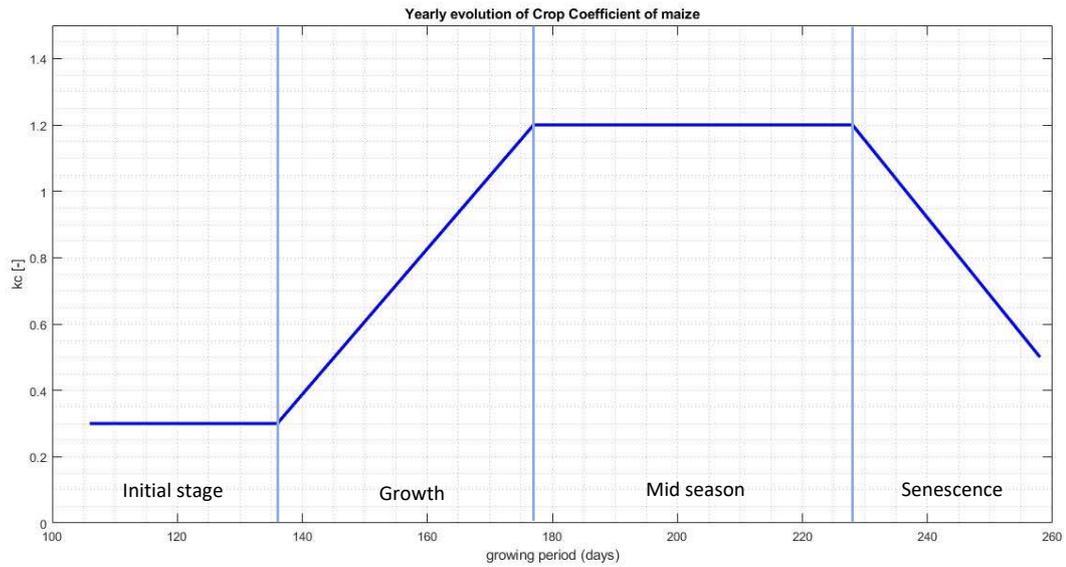


Figure 4. 1: evolution of kc of maize during the growing period

The parameter k_s varies from 0 to 1 according to the water content θ : if water content is high, between pore saturation n and θ^* (stomata closure), the crop is in “field capacity” condition, meaning that water availability is enough to guarantee a pattern equal to the standard one (ET_c). if water content is below θ^* , it means that the crop is suffering water stress, so the coefficient will start decreasing linearly to 0 since its performances will decrease. The wilting point θ_w corresponds then to the soil water content at which we have zero performance, therefore a k_s equal to zero.

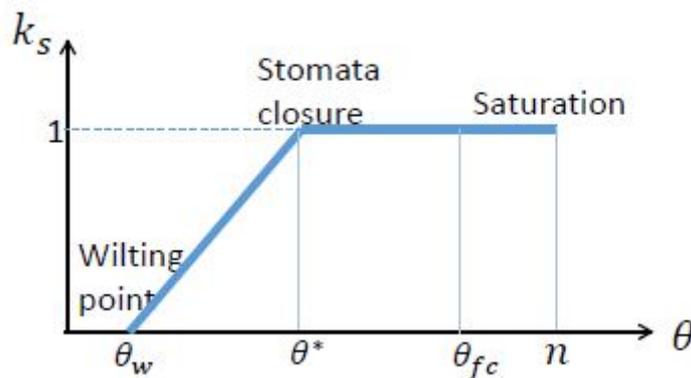


Figure 4. 2: pattern of k_s according to water content

4.2 Modelling the green and blue evapotranspiration

As it was said earlier, Actual Evapotranspiration (ET_a) depends on crop, weather, and soil parameters, so we need a model that, given these inputs, it generates temporal series of evapotranspiration. The provided model *MODEL_soil_water_balance_irrigated* fulfils this purpose: it is a Matlab function which requires information regarding

- Type of crop
- Climatic area
- The starting and ending dates of the culture
- Initial soil moisture
- Sowing soil moisture
- Available water capacity (AWC), i.e., the quantity of water that the soil can store after drainage
- Precipitation
- ET_0

From these inputs the model provides the crop water requirements, that is the amount of precipitation and irrigation water required to satisfy the needs of the plant and avoid water stress.

The choice of the crop is made by choosing its number from a list of possible crops that the model considers (in our case, maize is number 2) and the same thing applies also for the climatic region, where it's possible to choose between 10 regions which range from tropical to oceanic, continental, boreal and even arctic. Regarding the starting and ending date of the culture, the model asks the user to insert the day of the year based on a scale of 365 days (with 1 corresponding to the 1st of January and 365 to the 31st of December). For example, if in a pixel of the global map the cultivation starts on May 16th, we would indicate 136 as starting date.

For precipitation and ET_0 , instead, we need to provide to the model a matrix with the daily precipitation and reference evapotranspiration for that cell, with values expressed in [mm/day]. The initial soil moisture, then, corresponds to the initial condition of soil moisture before the soil water balance of the i-th day (this is actually the result of the soil water balance of the previous day). In the model, the function then converts the soil moisture from [m^3 of water/ m^3 of soil] to [mm], by considering the rooting depth of each day. The sowing soil moisture, instead, is the moisture of soil during sowing day, which is expressed with the same unit of the initial SM and AWC, that is in [m^3 of water/ m^3 of soil].

After data processing, the output of the model will then be the green evapotranspiration (ET_{green}), that is the contribute of evapotranspiration supplied by precipitation and, consequently, the irrigation requirements, that is the daily irrigation required to avoid crop water stress. For our case, we want to keep the k_s equal to 1 since we want to avoid water stress and maintain the crop at field capacity. It's important to remark that the irrigation requirements provided by the model do not correspond to the water quantity that has been effectively provided to the plant in that pixel in that year, while it is intended as the theoretical irrigation contribute that must be supplied to the plant to keep the cultivation at its maximum efficiency.

The soil water balance of the model works with a *for* loop that for each day of the year:

- 1st step: it calculates the daily crop coefficient, depending on the period of the growing season
- 2nd step: it derives the increment of soil moisture in the rooting zone, thanks to the information provided by precipitation matrix
- 3rd step: the model calculates TAW (Total Available Water in the rooting zone) and RAW (Readily Available Water), according to the following equations:

$$TAW = AWC * \text{rooting depth}_{day(i)} \quad (4.3)$$

$$RAW = TAW * \text{depletion factor} \quad (4.4)$$

- 4th step: calculation of the water stress coefficient, that is the minimum amount of irrigation necessary to avoid water stress. Keeping k_s equal to 1 means keeping the condition of field capacity in the cultivation, but actually k_s can also assume other values based on the final user needs (for example, k_s can be kept equal to 0.9 instead of 1, thus maintaining a mild but constant condition of water stress to limit irrigation inputs).
- 5th step: calculation of Actual Evapotranspiration ET_a, which is the result of

$$ET_a = k_s * k_c * ET_0 \quad (4.5)$$

- 6th step: the model calculates the reduction of soil moisture, that is the increase of deficit due to actual evapotranspiration, where deficit is the difference between TAW and SM

$$\text{deficit} = TAW - SM \quad (4.6)$$

- 7th step: it calculates then the evapotranspiration components, ET green and ET blue, i.e., the water input that comes from precipitation and irrigation respectively.

- 8th step: calculation of the total irrigation requirements, which is the sum of the evapotranspirative component and the water for soil component (which occurs only during sowing dates).
- 9th step: the model finally calculates the final soil moisture at the end of the i -th day, which is independent from the Rooting Depth, then this value is used as initial soil moisture for the next day.

4.3 Distributed unit water footprint at the global scale

4.3.1 Global maps of green, blue and total uWF

After a deep analysis of historical evolution of the water footprint pattern in Pino Torinese, it's now the time to expand this procedure to the rest of the world's pixels. To do this, it is necessary to collect the data of green and blue evapotranspiration such as irrigated and rainfed area for all the cells in the world. In fact, in each country, there is a certain number of harvested cells with a given crop. Each cell i is characterized by a known harvested surface, which may be decomposed in a fraction $A_{rf,i}$ that is only rainfed and another that is equipped for irrigation (so it is both irrigated and rainfed), $A_{irr,i}$. Therefore, the total harvested area can be expressed as

$$(A_{rf,i} + A_{irr,i}) = A_i \quad (4.7)$$

In the rainfed area of cell i the effective evapotranspiration corresponds to the green contribution only

$$ET_{a,i}^{rf} = ET_{g,i}^{rf} \quad (4.8)$$

while, in the irrigated fraction of the same cell, it is the sum of the green and blue component, namely

$$ET_{a,i}^{irr} = ET_{g,i}^{irr} + ET_{b,i}^{irr} \quad (4.9)$$

To obtain the mean green or blue national evapotranspiration for the nation N, it is necessary to perform a weighted mean of evapotranspiration values according to their rainfed or irrigated areas, that is

$$ET_g^N = \frac{\sum_{i \in N, rf \cup irr} (ET_{a,i}^{rf} \cdot A_{rf,i} + ET_{g,i}^{irr} \cdot A_{irr,i})}{\sum_{i \in N, rf \cup irr} (A_{rf,i} + A_{irr,i})} \quad (4.10)$$

$$ET_b^N = \frac{\sum_{i \in N, irr} (ET_{b,i}^{irr} \cdot A_{irr,i})}{\sum_{i \in N, rf \cup irr} (A_{rf,i} + A_{irr,i})} \quad (4.11)$$

where the sum of the products between ETs and their related areas of a country is then divided by the sum of all the rainfed and irrigated areas of the country.

Unit Water Footprint is then obtained by the ratio between the mean national evapotranspiration ET_g^N or ET_b^N and the mean national yield Y^N

$$uWF_g^N = 10 \cdot \frac{ET_g^N}{Y^N} \quad (4.12)$$

$$uWF_b^N = 10 \cdot \frac{ET_b^N}{Y^N} \quad (4.13)$$

From these data, it is possible to reconstruct the historical maps of total, green and blue unit water footprint.

As a first step, we recreate the historical maps of actual, green and blue evapotranspiration, creating a series of grid files with the same resolution of the maps of yield and harvested area (5 x 5 arc min). From .mat files *CROP_INFO_irr* and *CROP_INFO_irr* we extract the position in the maps, in terms of rows and columns, of all the cells in the rows of *ET_a*, *ET_blue*, *ET_green*, then they are converted from subscripts [row,col] to linear indices (idx1, idx2,...idxi). At this point, with a *for* loop running with all the 50 years, at each iteration the code extracts the column of the evapotranspiration matrices corresponding to the i-th year and correctly place the values of the array inside the ET maps, according to the linear indices.

Now it's the turn for rainfed and irrigated areas. In the previous chapters, we already described the steps to obtain the historical maps of maize harvested area in the time-interval 1961 – 2019, which were obtained by multiplying the harvested area of crop *c* in cell *i* in year 2000, $A_{i,c}(2000)$, with the ratio between the national harvested area of crop *c* in country *N* in year *t*, $A_{N,c}(t)$, over the national harvested area of the same crop in the same nation for year 2000,

$A_{N,c}(2000)$. This procedure assumes that all the cells of country N have proportionally grown in the years, with a uniform factor for all the country, therefore this implies that no cell switched off or on during the time interval, so that the distribution of the cells over the country remained constant.

Regarding the irrigated area of maize, instead, it is necessary to find a way to estimate its historical global evolution. Databases don't provide the time series of AEI for each crop, however FAOSTAT offers the historical time series of total AEI of each country, A_N^{irr} , considering all crops together, with values arranged on an annual scale, from 1961 to 2020. Therefore, the equation to reconstruct the maize irrigated area in cell i in year t can be expressed as

$$A_{i,c}^{irr}(t) = A_{i,c}^{irr}(2000) \cdot \frac{A_N^{irr}(t)}{A_N^{irr}(2000)} \quad (4.14)$$

where $A_{i,c}^{irr}(2000)$ is the irrigated area in cell i of crop c (maize) in year 2000, which is multiplied by the ratio between the total AEI in year t of country N , $A_N^{irr}(t)$, and the AEI of the same country in year 2000. This time, we add the assumption that, during the time interval, the agricultural area equipped for irrigation in a country grew proportionally for all crops.

At this point, the historical evolution of rainfed area in the cells that contain an irrigated fraction can be expressed as the difference between the maize harvested area in year t , $A_{i,c}(t)$, and the irrigated area in the same year, $A_{i,c}^{irr}(t)$, that is

$$A_{i,c}^{rf}(t) = A_{i,c}(t) - A_{i,c}^{irr}(t) \quad (4.15)$$

Instead, the cells which don't contain irrigated areas (thus are rainfed only) are simply considered as equal to the maize harvested area in year t , $A_{i,c}(t)$.

In the file *AEI_TOT_FAO*, arranged as a 178 x 50 matrix, where each row contains the time series of national AEI of a country from 1970 to 2019, we initially convert the absolute values in each cell into a ratio between the value of that cell over the value in year 2000, in order to already obtain the aforementioned factor $A_N^{irr}(t)/A_N^{irr}(2000)$. However, we also need to fill the gaps of the today's countries which were part of larger Unions in the past. More specifically, we keep into account the countries which were part of:

- Belgium – Luxembourg, up to 1999

- Czechoslovakia, up to 1992
- Ethiopia PDR, up to 1992
- URSS, up to 1991
- Yugoslavia, up to 1991

To do this, the rows of the today's countries composing these Unions are identified and the sum of their AEI in year 2000 is performed. For example, for Czechoslovakia it is the sum between the AEI of Czech Republic in 2000 and the AEI of Slovakia in 2000, thus obtaining $A_{Czechoslovakia}^{irr}(2000)$. Then, the values of the time series of national AEI of Czechoslovakia from 1961 to 1992 are divided by $A_{Czechoslovakia}^{irr}(2000)$, deriving the factor $A_N^{irr}(t)/A_N^{irr}(2000)$, which is then inserted in the gaps of Czech Republic and Slovakia from 1961 to 1992. This procedure is then repeated for the other counties involved in this process.

At this point it's possible to create the historical maps of Total, Green and Blue uWF. With a *for* loop running over the 50 years' time interval, the following equations are performed over each cell of the 2160 x 4320 matrix that contain positive values of ET and yield,

$$uWF_{tot,i} = 10 * \frac{ET_{a,i}^{rf} \cdot A_{rf,i} + ET_{g,i}^{irr} \cdot A_{irr,i} + ET_{b,i}^{irr} \cdot A_{irr,i}}{(A_{rf,i} + A_{irr,i}) * Y_i} \quad (4.16)$$

$$uWF_{g,i} = 10 * \frac{ET_{a,i}^{rf} \cdot A_{rf,i} + ET_{g,i}^{irr} \cdot A_{irr,i}}{(A_{rf,i} + A_{irr,i}) * Y_i} \quad (4.17)$$

$$uWF_{b,i} = 10 * \frac{ET_{b,i}^{irr} \cdot A_{irr,i}}{(A_{rf,i} + A_{irr,i}) * Y_i} \quad (4.18)$$

where the ET values are those got from the matrices of ETa, ETg, ETb, $A_{rf,i}$ and $A_{irr,i}$ are the values of area rainfed and irrigated in the i-th pixel from matrices of rainfed area and irrigated area and Y_i is the yield in the i-th pixel from matrix of yield.

Results are then inserted into 2160 x 4320 matrices for each year and saved into Matlab structures.

4.3.2 Historical series of mean national green, blue and total uWF

To reproduce historical time series of mean national uWF, the procedure is very similar to the one for the historical maps, with the only difference that the equation is not performed cell by cell anymore, but along the cells of each nation for each year.

To obtain mean national green and blue evapotranspiration, we apply equations (3.33) and (3.34) over a double *for* loop that iterates the equations for all the countries of the world for all the 50 years. Then, Green and Blue uWF is derived applying equations (3.35) and (3.36), where the mean national yield is taken from the matrix Y_TOT_FAO , which contains the mean national yield of all countries from 1970 to 2019. Finally, the historical series are saved into 178 x 50 matrices (where the rows correspond to the number of countries and columns to the years of the time interval), named UWF_naz , $UWFg_naz$, $UWFb_naz$.

4.4 Trend analysis

In this thesis, the analysis of trends allows us to understand the pattern of a variable, if and how much it's increasing or decreasing, how sharp are the fluctuations etc. For example, there may be variables that increase monotonically, but at the same time there can be variables which are increasing at a minor growth rate, with strong fluctuations throughout the time interval. Consequently, it is worth to find an objective criterion to quantify the level of significance of a trend, a method that can tell us whether the trend of a certain dataset, according to how data are distributed, is significant or not. In this case, we rely on the application of the *t-Student* test, a statistical test that states whether there is or not dependency between the independent variable x and the dependent one y . We start from the definition of the null hypothesis H_0 , that is that the angular coefficient b_1 is a random variable with zero mean, thus x and y are independent between each other. First of all, we build a linear relationship between the variables x and y (in this case, years and precipitation/evapotranspiration) by defining a first-degree linear model,

$$y = \tilde{y} + \varepsilon = (b_0 + b_1 * x) + \varepsilon$$

(4. 19)

where b_0 and b_1 are respectively the known term and the angular coefficient of the linear regression model, while ε is the error between the true and the predicted estimation. The b_0 and

b_1 parameters are estimated by minimizing the sum of the squared errors, through the Least Square Method:

$$b_0 = \bar{y} - b_1 * \bar{x} \quad (4.20)$$

$$b_1 = \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{S_{xy}}{S_{xx}} \quad (4.21)$$

The errors must be independent and identically distributed in order to have a mean equal to zero and variance σ_ε^2 ,

$$\sigma_\varepsilon^2 = \frac{\sum_{i=1}^n \varepsilon^2}{n - 2} \quad (4.22)$$

At this point we can build the test variable T, namely

$$T = \frac{b_1}{\sqrt{\sigma_\varepsilon^2 / S_{xx}}} \quad (4.23)$$

which is the ratio between the angular coefficient of the regression model over a sort of standard deviation of the data.

Then we set a level of significance α , that is the probability of rejecting the null hypothesis, corresponding to the extremes of the distribution tails where H0 is not accepted. Therefore, the limits of these areas, t_{lim} ,

Correspond to the limits of acceptance of the t-Student test, where

$$t_{lim} = t\left(\frac{\alpha}{2}, n - 2\right) \quad (4.24)$$

If $|T| < t_{lim}$, the null hypothesis is accepted, therefore there is no dependency between x and y variables and the angular coefficient b_1 is not significantly different than zero. Conversely, if $|T| > t_{lim}$, this means that there H0 is rejected, thus we have a statistically significant dependence between x and y.

5 Results

In this chapter, the thesis results will be presented with the following organization. Initially, the hydrological model is performed on a single pixel, Pino Torinese, where we evaluate the time series of green and blue uWF together with further analysis on climatic data (precipitation, ET_0 , ET_a etc...). Later, we introduce the results of a multi-regression analysis that aims at finding a relation between ET_b/ET_a ratio and the climatic data (P , ET_0). Then the global maps of total, green and blue uWF are displayed, such as the time series of these variables for some countries (Italy, USA, Australia, Nigeria and Viet Nam) and the scatter plots of the comparison of our results with Water Footprint Network and with data obtained applying the Fast Track approach. In conclusion the last section is dedicated to the display of the time series of water footprint, at global and country level.

5.1 Local analysis: Pino Torinese

5.1.1 Run the hydrological model on a single pixel: Pino Torinese

To verify the model and have a deeper understanding of the results, we firstly run it for a single pixel of the global matrix, in this way we can perform further analysis which may result too complex if carried on for the whole globe. Therefore, we choose the town of Pino Torinese as representative area for our testing. The village is located in Piemonte (Italy), with coordinates $45^{\circ}02'36.56''N$ $7^{\circ}46'21.2''E$ UTM and [539, 2253] on the 5 arc-minute resolution global maps. This cell has a total agricultural area of 6058 hectares, where 248 of them are equipped for irrigation (Maize AEI is 98 ha). Pino Torinese belongs to the Climatic zone n° 5, that is Temperate sub-continental, the available water capacity (AWC) is estimated 0.2 m^3 of water/ m^3 of soil and maize cultivation period ranges between day 106 (16th April) to 258 (15th September). Then, from databases, we get the historical time series of precipitation, ET_0 and soil moisture of this pixel for 365 days a year for 50 years (from 1970 to 2019). These series are arranged as 50x365 matrices, where rows correspond to the years of the time interval and columns to the days of each year. To make the dimension of *season_start* and *season_end* (the beginning and ending of the culture) and AWC variables coherent with the other matrices, they are arranged as 50x1 arrays, then the model can finally be run.

The results of the model are as well arranged as 50x365 matrices, so now it's possible to create some plots to display the results from them.

We firstly display the yearly evolution of the main evapotranspirations (ET_a , ET_g and ET_b) of a single year, 2003, which has been characterized by particularly intense meteorological events in that pixel such as in the rest of Piemonte.

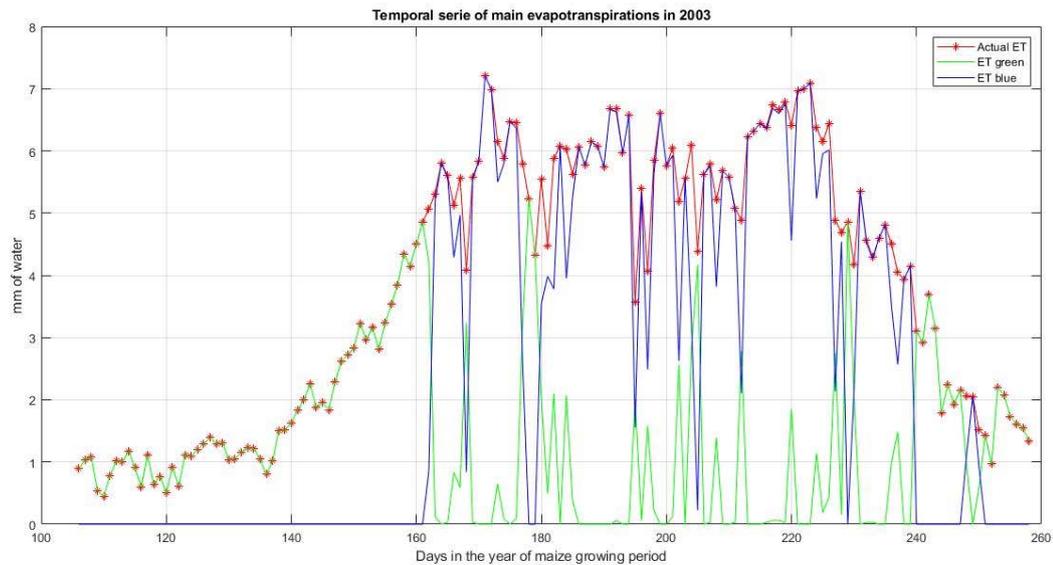


Figure 5. 1: temporal series of maize ETs in 2003 on Pino Torinese

Looking at the chart, it can be seen that the pattern of ET_a almost regularly follows the one of k_c (the crop coefficient), since, from equation 3.14, ET_a is simply the product between the crop coefficient and ET_0 (in this case the water stress coefficient is set to 1 so it doesn't influence the equation). The crop coefficient, in fact, during the initial period (planting) is supposed to be constant at a low value, then, during growth stage, it linearly grows up to its maximum, consequently it remains constant during the mid-season and finally it linearly decays during the last period, the senescence. A very similar behaviour occurs also for Actual Evapotranspiration.

In addition, we can see that, up to day 160 ET_g is able to satisfy the crop water requirements, therefore it corresponds to ET_a . After that day, irrigation requirements start to increase, balancing with ET_g .

However, this plot represents the pattern of a single year, obtained from the climatic conditions of that year, therefore we try now to repeat the previous plot for all the years of the time interval

adding also the mean Actual, Green and Blue Evapotranspiration, weighted over the entire period 1970-2019; in this way the yearly fluctuations should be attenuated and we may obtain a behaviour similar to the theoretical one.

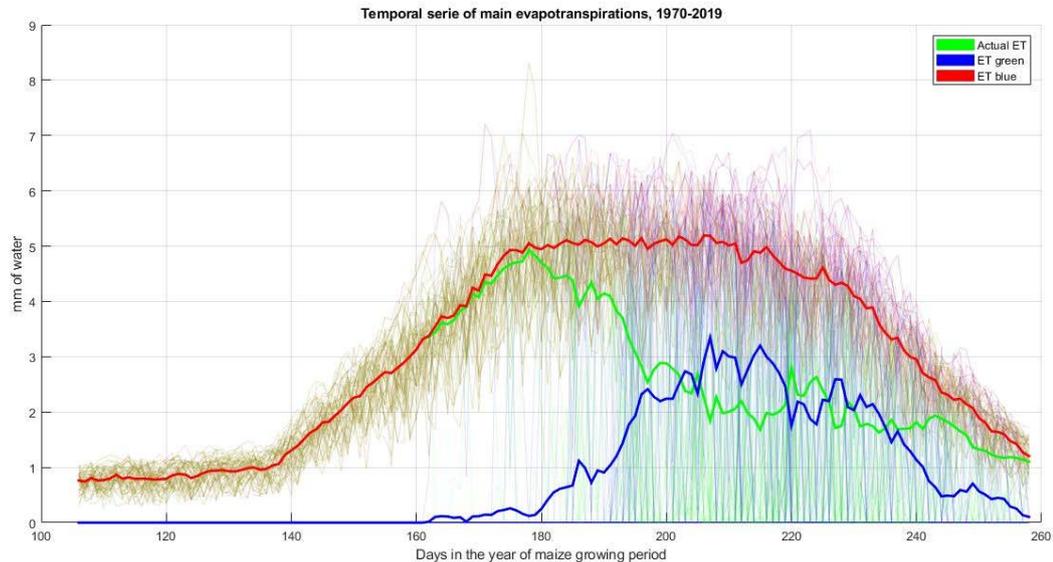


Figure 5. 2: time series of ET_a , ET_g and ET_b , 1970-2019 average on Pino Torinese

In Figure 5. 2 it is possible to see the thicker lines corresponding to the mean Actual, Green and Blue ET, while the thinner and more transparent lines refer to the evolution of the same variables of each year of the time-interval. It is evident that the average of the three variables attenuates all the fluctuations caused by yearly variability of the climatic conditions in that pixel.

Moreover, there is a clear similarity between the pattern of ET_a with the one of the crop coefficient k_c , that grows between days 140 – 180, then it remains constant during mid-season and finally linearly decreases up to the harvesting day (Figure 4. 1).

Looking at green evapotranspiration, we see that it perfectly follows the pattern of ET_a up to day 160, when irrigation starts to be practiced, then, after reaching its peak at day 180, ET_g 's trend begins to fall with some fluctuation up to the day of the harvest. Therefore, to balance the decrease of the green contribute, ET_b starts to rise from day 180 to approximately day 205-210, then it gradually decreases as well from day 220 since also Actual Evapotranspiration reduces in the senescence period. These trends can be explained by the fact that, during spring, the precipitation input is generally high, thus rainfall is able to be the only source of supply for the crop, but during summer this contribute usually reduces, hence it is necessary to resort to irrigation to keep maize culture at field capacity.

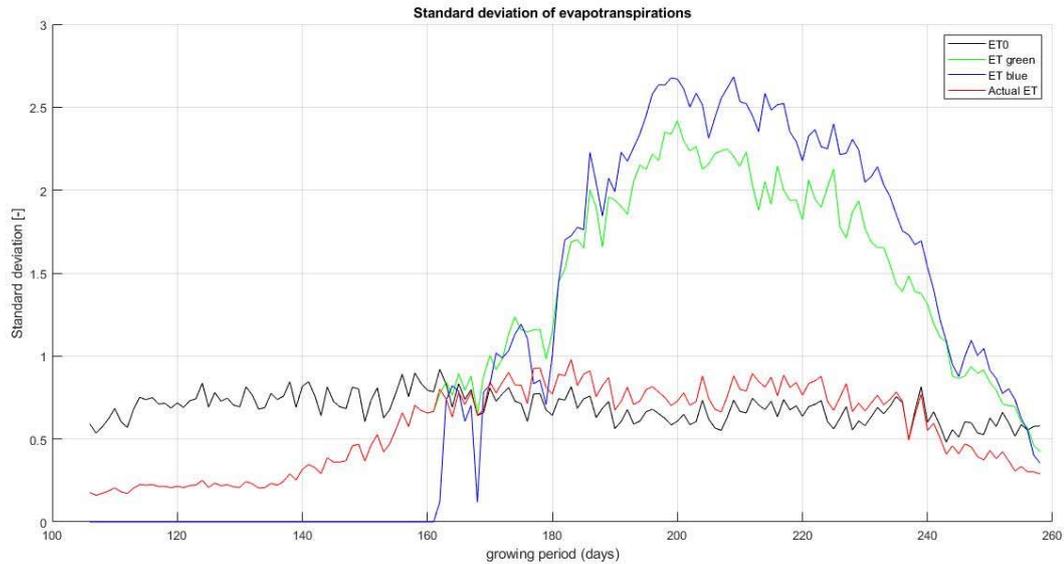


Figure 5. 3: standard deviation of ET_a , ET_g and ET_b , 2003 on Pino Torinese

We now want to look at the standard deviations (STD) of the contributes of evapotranspiration, remembering that standard deviation expresses the degree of dispersion of a set of values. In this case, the calculation of the standard deviation has been performed over the rows of the matrices, meaning that each day of the year contains 50 samples, as 50 are the years taken into analysis.

From the graph we can clearly see that the standard deviation of ET_0 is generally constant throughout the year with very little fluctuations; this means that the results of all the days of the year have data which have approximately the same degree of dispersion. The STD of ET_a is slightly different: initially its dispersion is very low (all observations show similar values) up to days 165-170, where it reaches the one of ET_0 and crosses it. Anyway, this pattern is very similar to ET_0 because it differs by the term kc , which is constant from day 177 to day 228.

Different situation applies to the STD of Green and Blue ET. As we said earlier, ET_g follows ET_a up day 160 then they separate, with ET_g a significantly increased dispersion throughout the rest of the growing period. ET_b , instead, is zero until day 160, then its dispersion suddenly rises crossing 2.5, thus being the variable with the greatest variability. The coefficient of variation (CV) of ET_b shows a particular pattern: it is very high at the beginning of irrigation period, then it gradually decreases, inversely proportional to the rise of its STD. We must remember, in fact, that

$$CV = \frac{\text{Standard Deviation}}{\text{Mean}}$$

(5. 1)

During first days of irrigation period, ET_b has higher standard deviation, but a very low mean, later mean Blue Evapotranspiration increases, so the fluctuations of STD are compensated by higher values of mean, which then result in a lower Coefficient of Variation (see Annexes, Figure 7. 1).

Let's now check the frequency distribution of Reference and Actual Evapotranspiration, to see how these data are arranged for Pino Torinese cell.

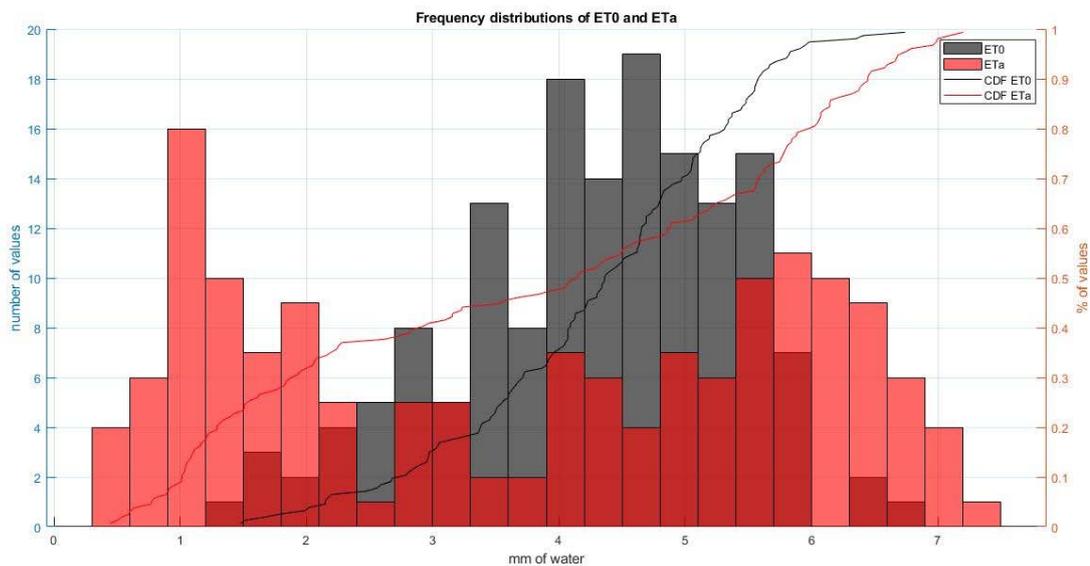


Figure 5. 4: frequency distribution of ET_0 and ET_a

This chart illustrates the histogram and the Cumulative Distribution Function of the two types of evapotranspiration in Pino Torinese in year 2003: while ET_a has a wider range, which goes from approximately 0.2 to 7.5 mm, Reference ET has a shorter variance, with most of data concentrated between 4 and 5.8 mm. ET_a data show peaks, one at 1 mm, corresponding to the initial period of the cultivation, and the other between 5.5 – 6.5 mm, that includes the values during the mid-season period.

The same characteristics can be seen also from CDF curves: the red one (corresponding to ET_a) is wider than the black curve and it shows a milder growth.

We now look at the histograms of ET_g and ET_b and compare them with those of precipitation and irrigation requirements along the growing period.

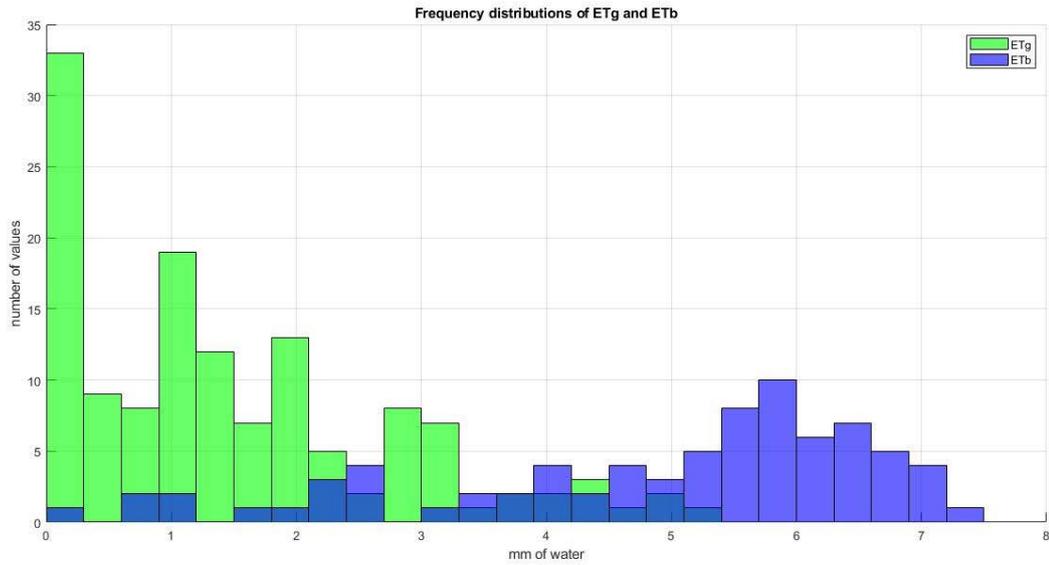


Figure 5. 5: frequency distribution of ETg and ETb

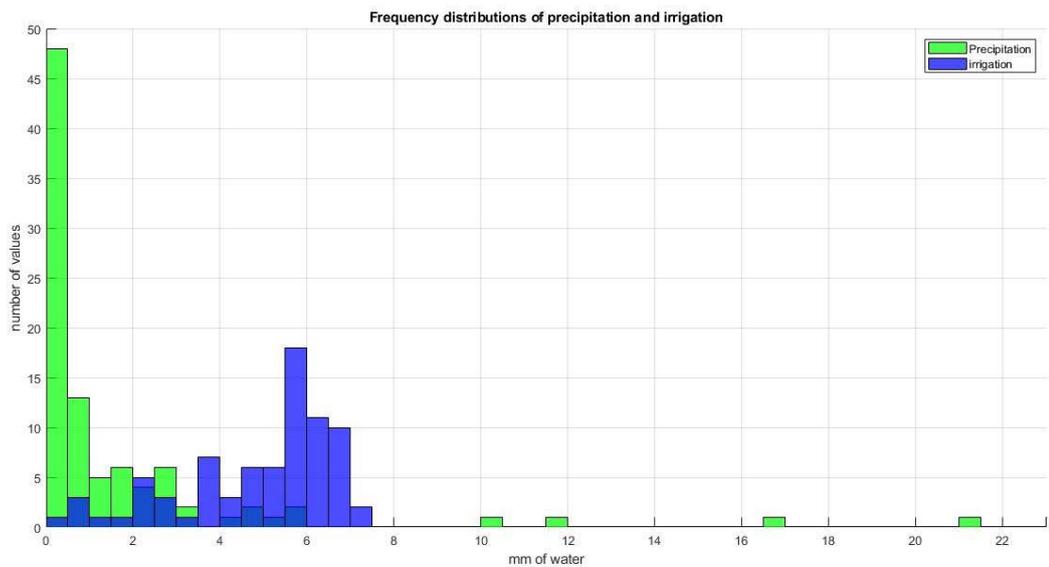


Figure 5. 6: frequency distribution of precipitation and irrigation

At a first glance, we notice that both graphs have in common the fact that the green variables (precipitation and ET_g) contain most of their data in the leftward columns, that is the bins of the lower values of mm of water. This is particularly evident for precipitation: almost 70% of data show values which are lower than 0.5 mm. Irrigation histogram, instead, is skewed to the left, with fewer data on the left side of the graph and the peak at 5.5 – 6 mm.

5.1.2 Calculation of uWF in Pino Torinese

At this point, we can calculate the unit water footprint and its green and blue components in the pixel of Pino Torinese. We already know that the two terms to determine uWF are the evapotranspiration and the crop yield, therefore, we need to identify the annual series of these two parameters for this pixel.

We start by identifying the pixel of Pino Torinese inside the matrix *Yield_maize*. Therefore, we extract from that matrix the yields from 1970 to 2019 at the coordinates [539,2253]. This series is already arranged with annual values; thus, we don't need further processing. Evapotranspiration is instead arranged as a 50 x 365 matrix, so we need to sum all the values in the rows of ET_g matrix, such as in ET_b , to obtain one single value per year.

Now, it's possible to calculate Total uWF, according to the equation

$$uWF = 10 * \frac{(ET_g + ET_b)}{Y_{Pino}} \left(\frac{m^3}{ton} \right) \quad (5.2)$$

while, for Green and Blue uWF, the equations become

$$Green\ UWF = 10 * \frac{ET_g}{Y_{Pino}} \left(\frac{m^3}{ton} \right) \quad (5.3)$$

$$Blue\ UWF = 10 * \frac{ET_b}{Y_{Pino}} \left(\frac{m^3}{ton} \right) \quad (5.4)$$

In this way we obtain three 50 x 1 arrays where each cell contains the total, green and blue unit water footprint estimates in Pino Torinese on a specific year. Now the time series are displayed on a plot to see their trend in time.

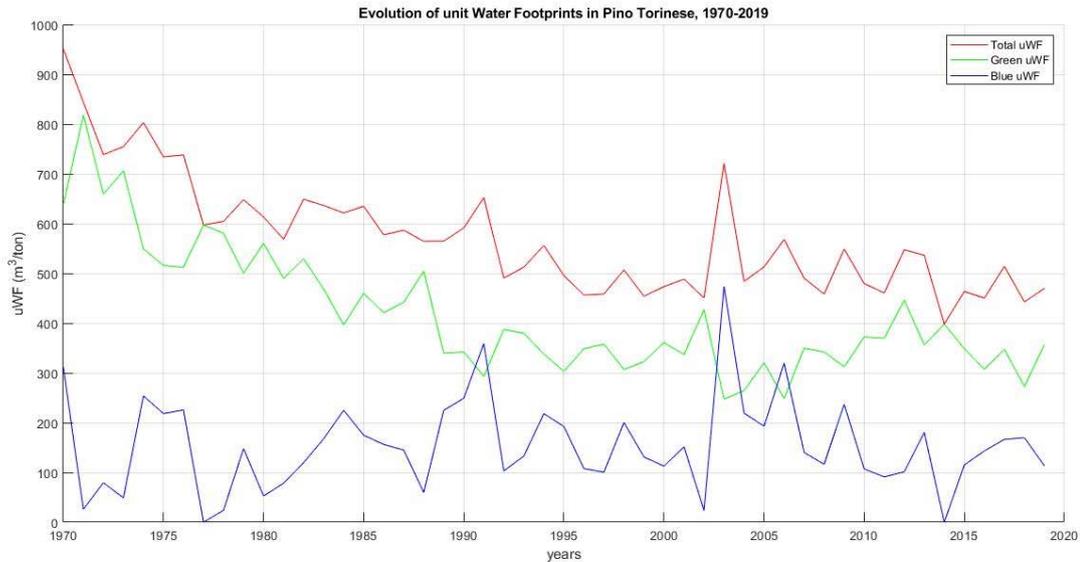


Figure 5. 7: evolution of total, green and blue uWF, 1970-2019

As we can see, total unit water footprint has gradually decreased over time, from approximately 950 m³/ton in 1970 to 500 in 2019. Also green water footprint shows a descending trend during the time-interval, while Blue uWF remained approximately constant, with no evident trend and with greater fluctuations. As $ET_a = ET_g + ET_b$, we clearly see that Green and Blue uWF compensate reciprocally: during rainy years, the total contribution has been fully satisfied by green unit water footprint, while in others (such as in 1991, 2003, 2006), where precipitation was not sufficient, irrigation became the greatest contribution to keep the crop at field capacity condition.

5.2 Relation between blue fraction of uWF and climatology in Pino Torinese and at global scale

The model *Model_soil_water_balwnce_irrigated* can be computationally demanding because it requires several input parameters to be able to run (climatic region, precipitation, ET_0 , soil conditions, etc.), also, it is not available to all the community of people interested in the investigation of crop water footprint. Therefore, it is worth to investigate an alternative way to obtain green and blue evapotranspiration without relying on all such input parameters. An

effective option can be the implementation of a multi regression analysis, where evapotranspiration is supposed to be dependent by a linear combination of some input parameters, for example precipitation P and reference evapotranspiration ET_0 , which are the parameters that can be mainly accessible to final users.

So, before jumping into the details of this multi regression analysis, we firstly analyse these two variables, by plotting their time series and checking how their evolution behaves in time.

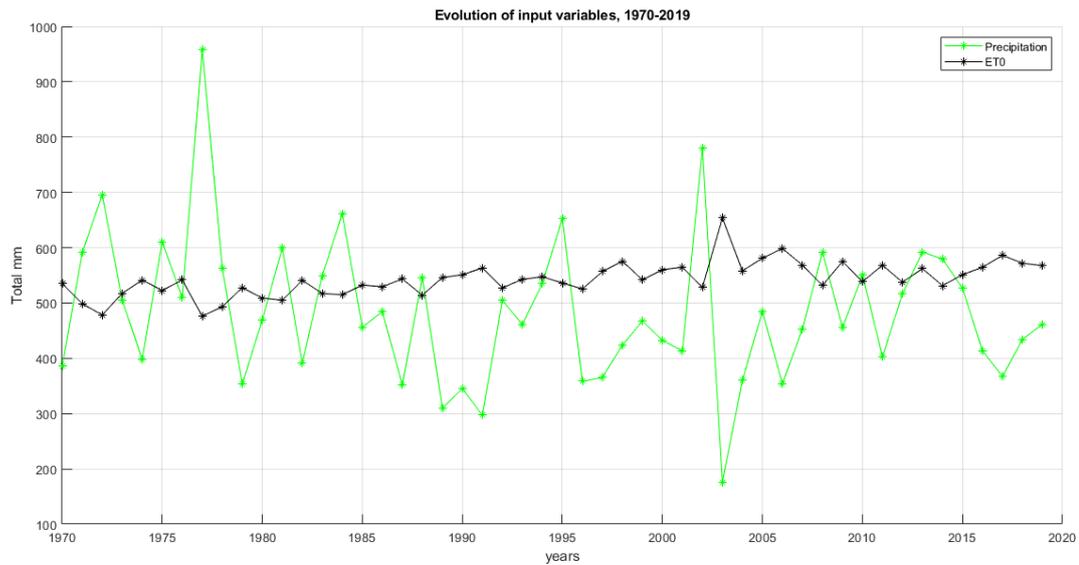


Figure 5. 8: time series of P and ET_0 , 1970-2019 on Pino Torinese

From this plot, we notice that, while ET_0 remains approximately constant over time (except for the changes induced by climate change, which generate a mild increasing trend), precipitation pattern looks irregular, with very sharp fluctuations, hence it's hard to see a clear trend. Therefore, we apply a statistical inference on their time series, checking the level of significance of their trends. The results we obtain are coherent from the previous expectations: for ET_0 , $|T| > t_{lim}$, thus we have a statistically significant trend, while for precipitation, the trend is not statistically significant. In fact, this variable has a too low angular coefficient of the linear regression model with respect to the high standard deviation of the oscillations, therefore it is not possible to state that this variable is characterised by a significant trend.

We now perform a multi regression analysis, trying to estimate evapotranspiration from the historical series of precipitation and ET_0 . More specifically, we suppose that the ratio of ET_b over ET_a can be described with a regressive law, namely

$$\frac{ET_b}{ET_a} = A_0 + A_1 * P + A_2 * ET_0$$

(5. 5)

where P and ET_0 are respectively the precipitation and reference evapotranspiration time series in that pixel in the time interval 1970-2019, while A_0 , A_1 , A_2 are a set of constant coefficients. In order to be able to determine the ratio of the two evapotranspiration parameters, we need to estimate the three coefficients applying the Least Square Method, to minimize the sum of the squared errors. We want to use ET_b/ET_a instead of ET_b/ET_g because, looking at the time series of the two ratios, we see that blue over green ET series shows much greater fluctuations with respect to ET_b/ET_a (see Annexes, Figure 7. 5); since it's too hard to reconstruct the annual variability of such an unpredictable variable, we focus of the ratio between blue and actual evapotranspiration, which instead exhibits weaker oscillations.

As mentioned before, the three coefficients are determined by minimising the sum of the squared prediction errors, that is by minimising the sum of the squared distances between the predicted and expected data points. To perform this operation, we need to identify a part of the dataset that will be used to “train” the model by finding the most appropriate coefficient of the linear regression and another one that performs the testing of the results. Therefore, the dataset is randomly split into training and testing data, where the training dataset contains 75% of data, while the testing dataset the remaining 25%. Since this operation is performed by a random selection of the data inside the series, we want to repeat this procedure for a certain number of times (in this case, we opt for 200 iterations), in order to have a stronger evaluation of the prediction error of the model. Then, the training set such as the testing set are identified: in this case, the training and testing X correspond to P and ET_0 , while the Y is associated to the ratio between ET_b and ET_a . Afterward, the Least Square Method is applied on the training dataset and the three coefficients A_0 , A_1 , A_2 are determined, having values of -1.5, $-8.7*10^{-5}$ and $3.3*10^{-3}$ respectively. At this point, the three coefficients are used on the testing dataset, to create an estimation of the ratio ET_b/ET_a , which is then compared with the true values (y_{test}) in a scatter plot, and the RMSE (Root Mean Square Error) between predicted and expected data is computed for each of the 200 iterations, with the equation

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}}$$

(5. 6)

where \hat{y}_i is the predicted ET_b/ET_a ratio applying the regression model and y_i is the expected ET_b/ET_a ratio.

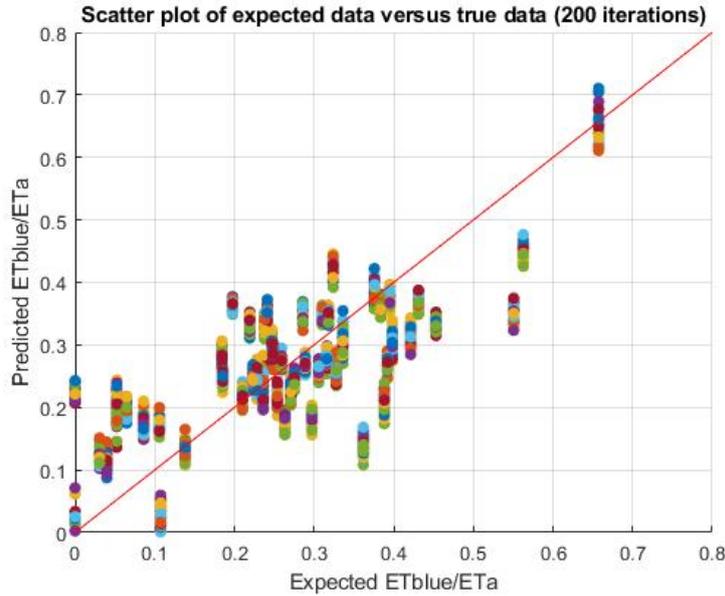


Figure 5. 9: scatter plot of predicted ET_b/ET_a ratio by the linear regression versus expected ET_b/ET_a (200 iterations)

From the scatter plot, we can see that the model provides rather well aligned results, considering that the dataset contains a limited number of data (only 50 data, as 50 are the years of the time-interval) and the only input parameters are P and ET_0 . By plotting the histogram of the error, we see that the pattern looks quite as a normal distribution, with the mean of RMSE corresponding almost to the peak of the curve, between 0.093 and 0.1, a mean coefficient of variation of RMSE that is around 0.34 and a coefficient of determination (R^2) which is 0.41. In addition, looking at the three coefficients of the parametric equation, it can be clearly observed that the coefficient A_1 , that refers to precipitation, has a lower weight with respect to A_2 , the one of ET_0 , in fact the first has an order of $10^{-5} - 10^{-4}$, while the second a dimension of 10^{-3} .

The analysis can be extended at the spatial scale, thus aiming at the estimation of the spatial heterogeneities of the ratio ET_b/ET_a over all the world's cells of a single year. We extract the values of ET_g , ET_b , ET_0 and precipitation of year 2000 in all the pixels containing irrigated areas, then the four arrays are grouped in a single matrix named *variables*. At this point, the procedure for the creation of the regression model becomes equal to the one done for Pino Torinese, therefore with the random splitting of the dataset into training and testing, the determination of the three coefficients A_0 , A_1 , A_2 (which account for $5.3 \cdot 10^{-3}$, $-1.8 \cdot 10^{-4}$, $4.13 \cdot 10^{-4}$ respectively) and the estimation of ET ratio. The scatter plot between true versus expected data is then represented; it can be seen that there is a less clear correspondence between expected versus predicted data, particularly due to the very high amount of data. This

time, the estimation provides worse results of mean cv_RMSE (0.80), while the R^2 is slightly higher, around 0.43. The analysis may be further deepened by performing a spatio-temporal regression, that is by estimating the temporal evolution of ET_b/ET_a ratio for all the pixels of the world, however, from the already obtained results, we can state that the best fitting results are found in the temporal variability rather than in the spatial one. In fact, the estimation of green and blue evapotranspiration at the spatial level must take into account different climatic regions and different crop growing periods, therefore the only information of annual precipitation and ET_0 is not able to describe the irrigation requirements. Instead, by performing the multi-regression analysis at the temporal scale, there is no variability of such variables, therefore we can have a slightly better estimation of the irrigation needs.

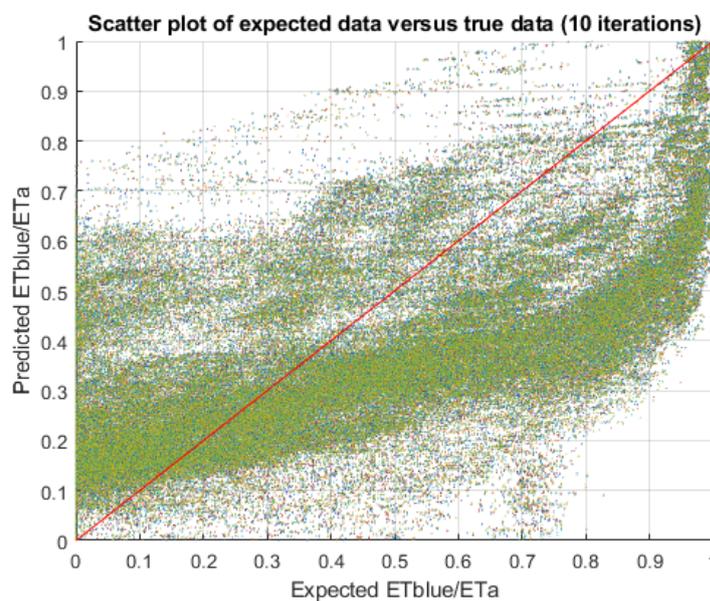
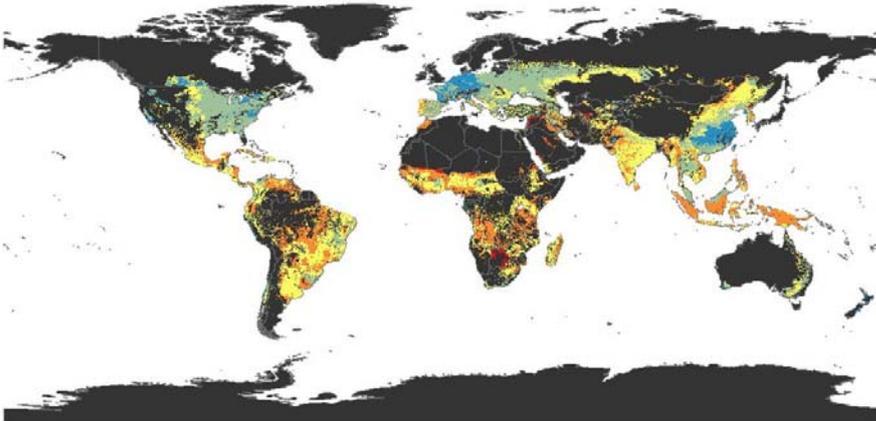


Figure 5. 10: scatter plot of expected versus true ET_b/ET_a ratio in all the pixels of the world, 2000

5.3 Display global maps of total, green and blue uWF

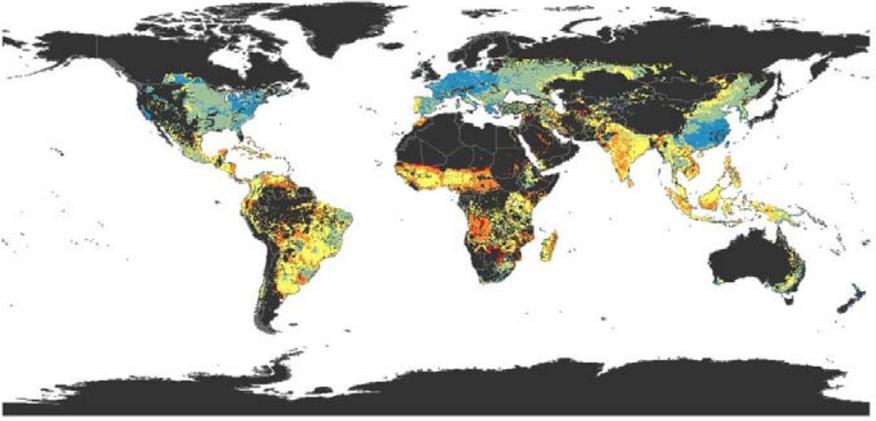
In this section, the global maps of total, green and blue uWF are represented, choosing years 1970, 1980, 2000 and 2019 for data display. To obtain these results, maps have been processed on ArcMap software, where the values of the cells have been split into 5 classes according to the range of data distribution.

5.3.1 Total uWF, years 1970, 1980, 2000, 2019



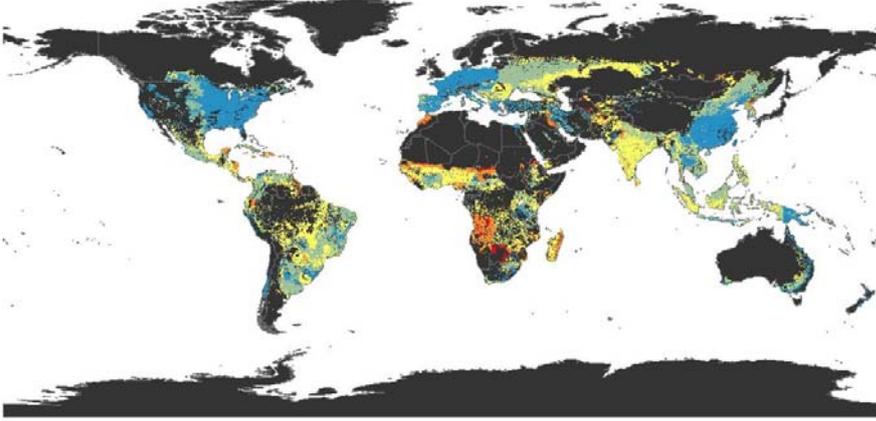
Legend (m³/ton)

- 0 - 1.000
- 1.000 - 2.000
- 2.000 - 4.000
- 4.000 - 16.000
- 16.000+



Legend (m³/ton)

- 0 - 1.000
- 1.000 - 2.000
- 2.000 - 4.000
- 4.000 - 16.000
- 16.000+



Legend (m³/ton)

- 0 - 1.000
- 1.000 - 2.000
- 2.000 - 4.000
- 4.000 - 16.000
- 16.000+

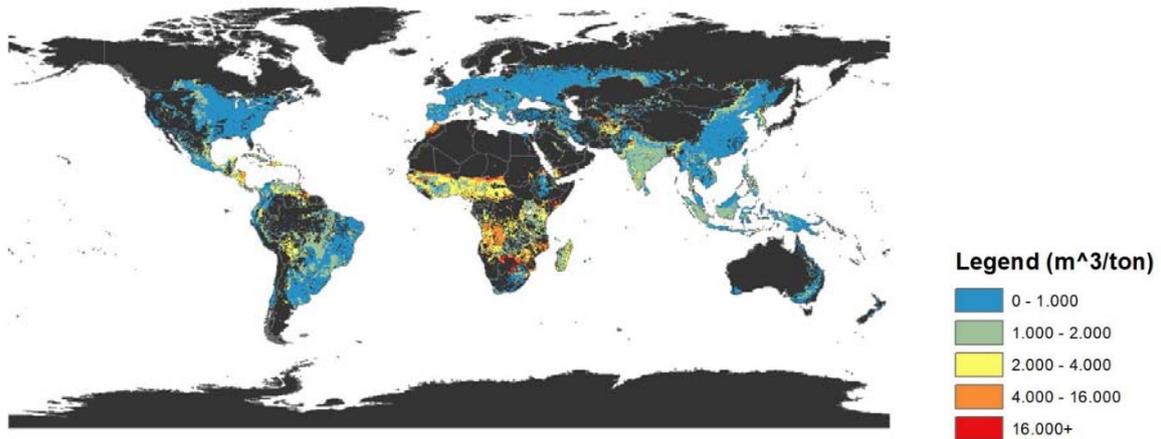
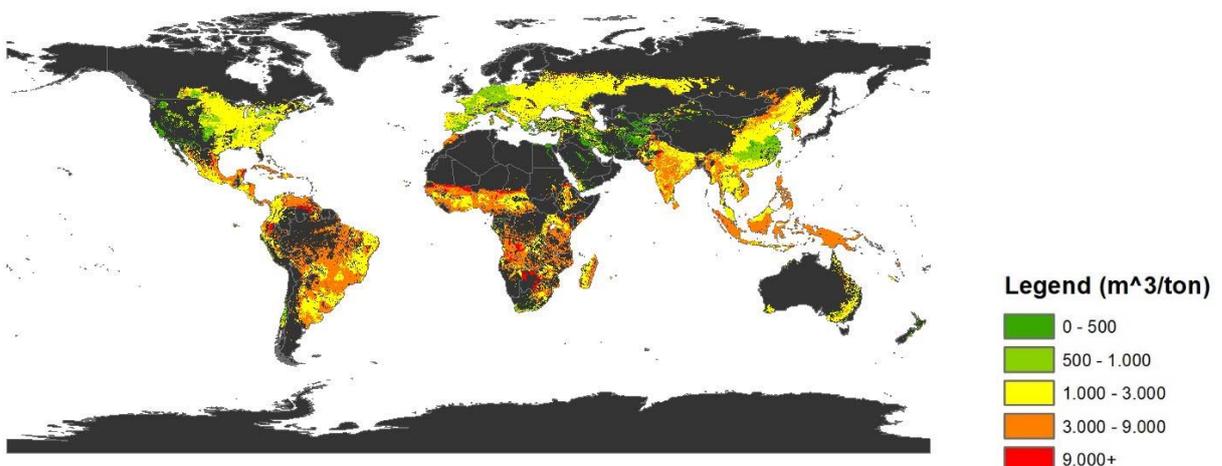
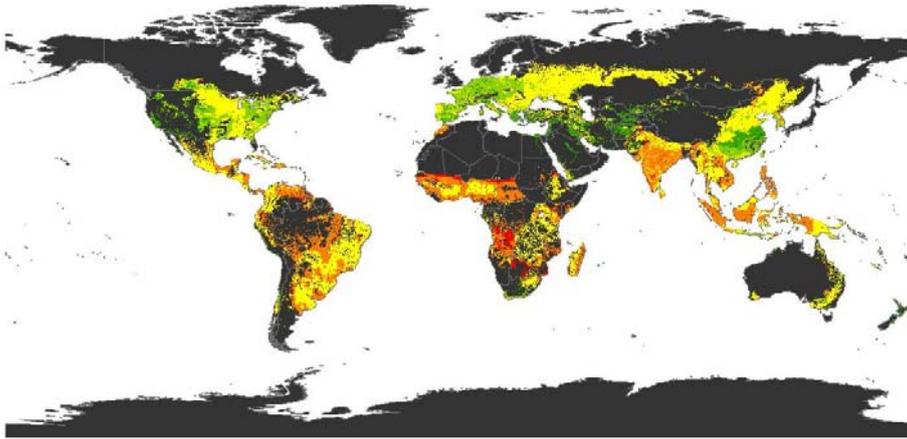


Figure 5. 11: total uWF, of years 1970, 1980, 2000, 2019

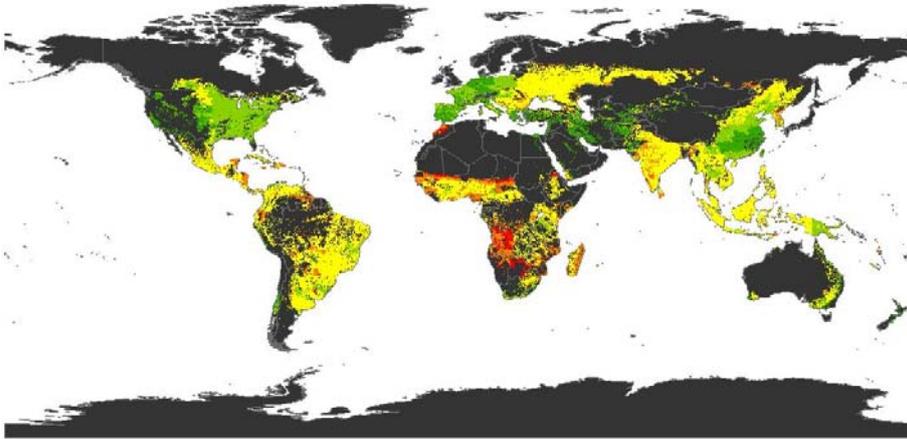
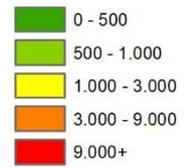
Looking at the four images, we can clearly observe that total uWF has generally decreased over time. In 1970, the “blue cells” (the cells with uWF values lower than 1000 m³/ton) were only in Central Europe and East China, then they progressively spread globally, with almost all continents reaching those values. African countries, instead, exhibit a minor decrease of water consumption per unit of product with respect to the other continents, with many areas in the southern part of the continent showing values above 5000 m³/ton in 2019.

5.3.2 Green uWF, years 1970, 1980, 2000, 2019

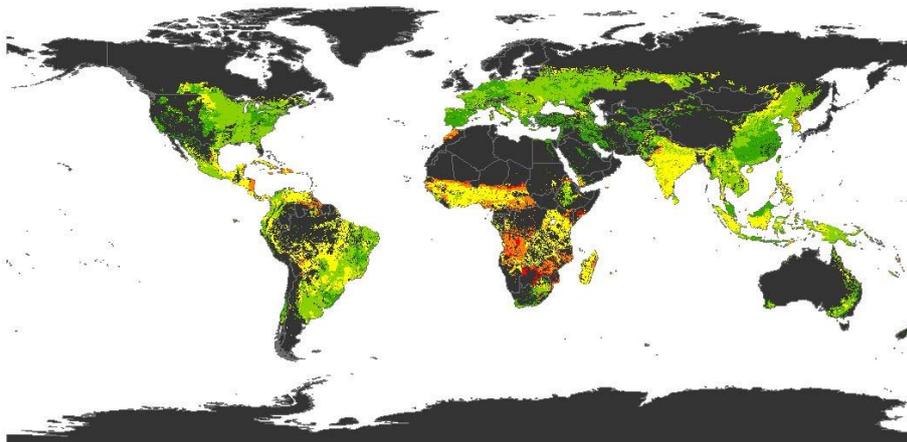
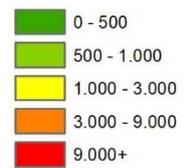




Legend (m³/ton)



Legend (m³/ton)



Legend (m³/ton)

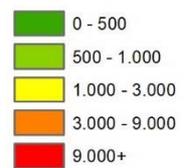
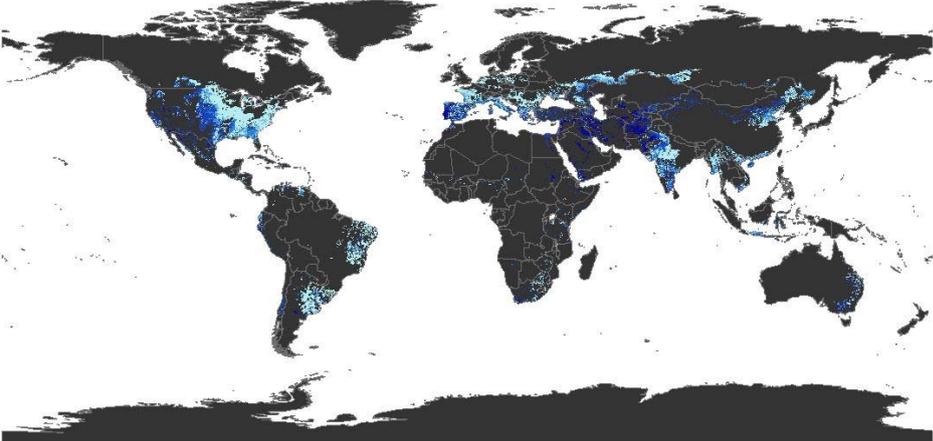
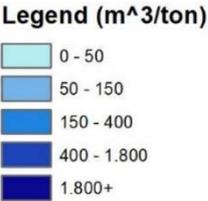
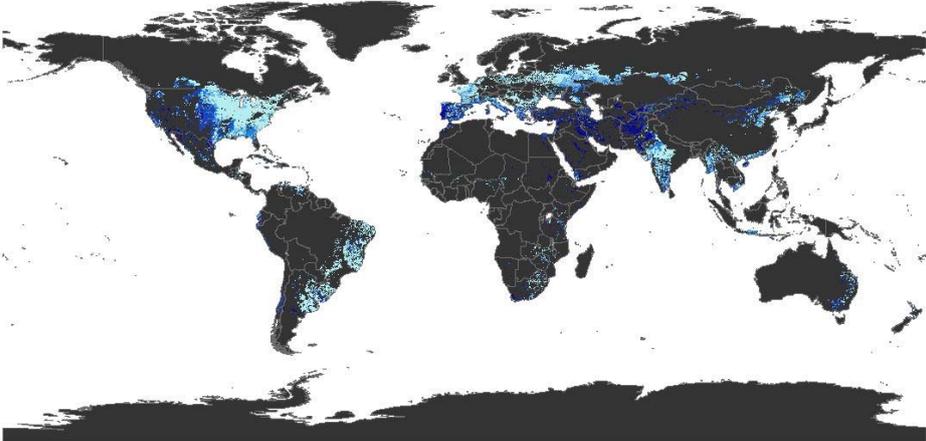


Figure 5. 12: green uWF of years 1970, 1980, 2000, 2019

5.3.3 Blue uWF, years 1970, 1980, 2000, 2019



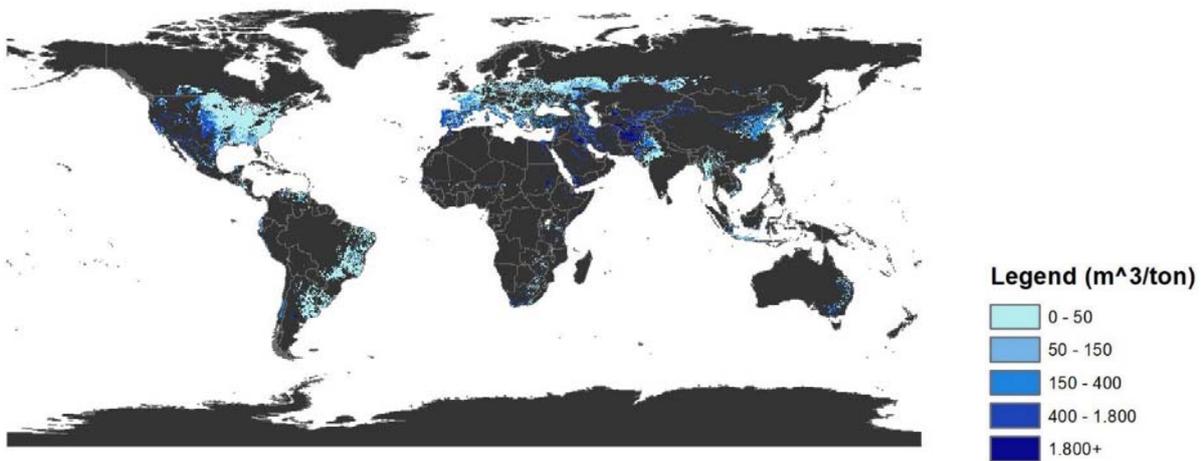
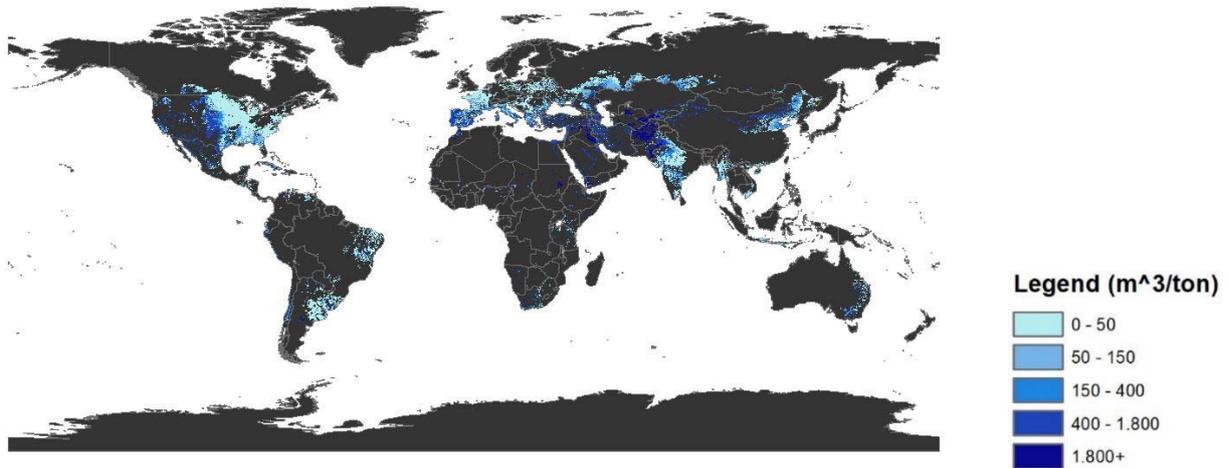


Figure 5. 13: blue uWF of years 1970, 1980, 2000, 2019

Comparing the historical maps of green and blue uWF, we can see that both green and blue uWF changed in the 1970 – 2019 time-interval. In addition, we can appreciate that the green contribute is much greater than the blue one: the former show values between approximately 200 – 10000 m³/ton, while the values of the latter range between 0 and 2000 m³/ton.

While green uWF exhibits an evident decreasing trend globally, the values of the blue uWF maps reduce with a smaller rate and, in some regions, they increase in time, meaning that it is harder to reduce blue water consumption per unit of product with respect to the green water.

5.4 Evolution of global Area Equipped for Irrigation

It is worth exploring the temporal evolution of global AEI, analysing how this variable globally and locally evolved throughout the time interval 1970 – 2019. In the section “Global maps of Green, Blue and Total uWF” we described the steps to reconstruct the global evolution of irrigated areas, correcting the pixel-scale value of year 2000 with the ratio between the national AEI of all crops of a certain year over the national AEI of all crops of 2000. The values, then, have been collected in matrix *Area_irr_hist*, which contains the evolution of all pixels for the 50 years’ time-interval.

From these data, we group them nation by nation, creating a new file *AEI_TOT*, with the AEI time series at national scale. Then, the values of all nation are summed together, thus obtaining the global AEI of a certain year.

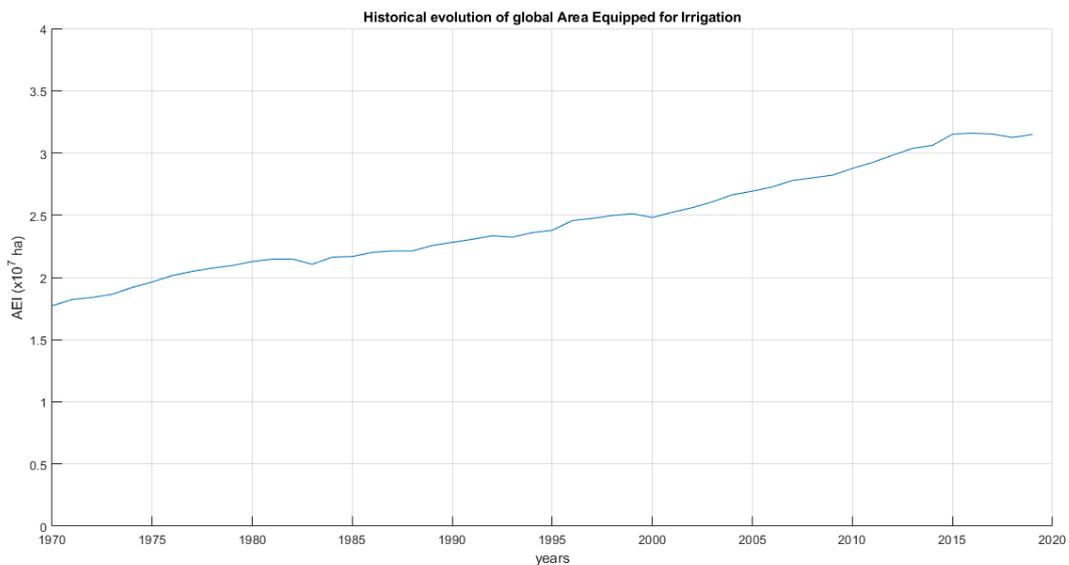


Figure 5. 14: time series of global Area Equipped for Irrigation (AEI), 1970-2019

The chart illustrates the temporal evolution of global AEI and it can be clearly observed that the areas significantly raised from 1970 to 2019. The variable, in fact, grows monotonically, behaving as a straight line. Calculating the growth rate between the beginning and end of the time interval applying the equation

$$GR (\%) = \frac{AEI_{end} - AEI_{start}}{AEI_{start}}$$

(5. 7)

where AEI_{end} is the global AEI averaged in the years 2015-2019 and AEI_{start} is the global AEI averaged in years 1970-1974, we notice that global AEI had a 70,7% growth during the 50 years' time interval, passing from 18 million hectares to 31 in 2019.

We extend the analysis at the national scale, checking the growth rate of each country. With a *for* loop, that iterates the operation for each nation, we identify the starting and ending date of the time series, and we perform the calculation of the growth rate, saving the result in a table, named *AEI_growth_table*. Then, the growth rates of each country are printed in a 2160 x 4320 resolution global map, where the surface of each country assumes the values of the national growth rate.

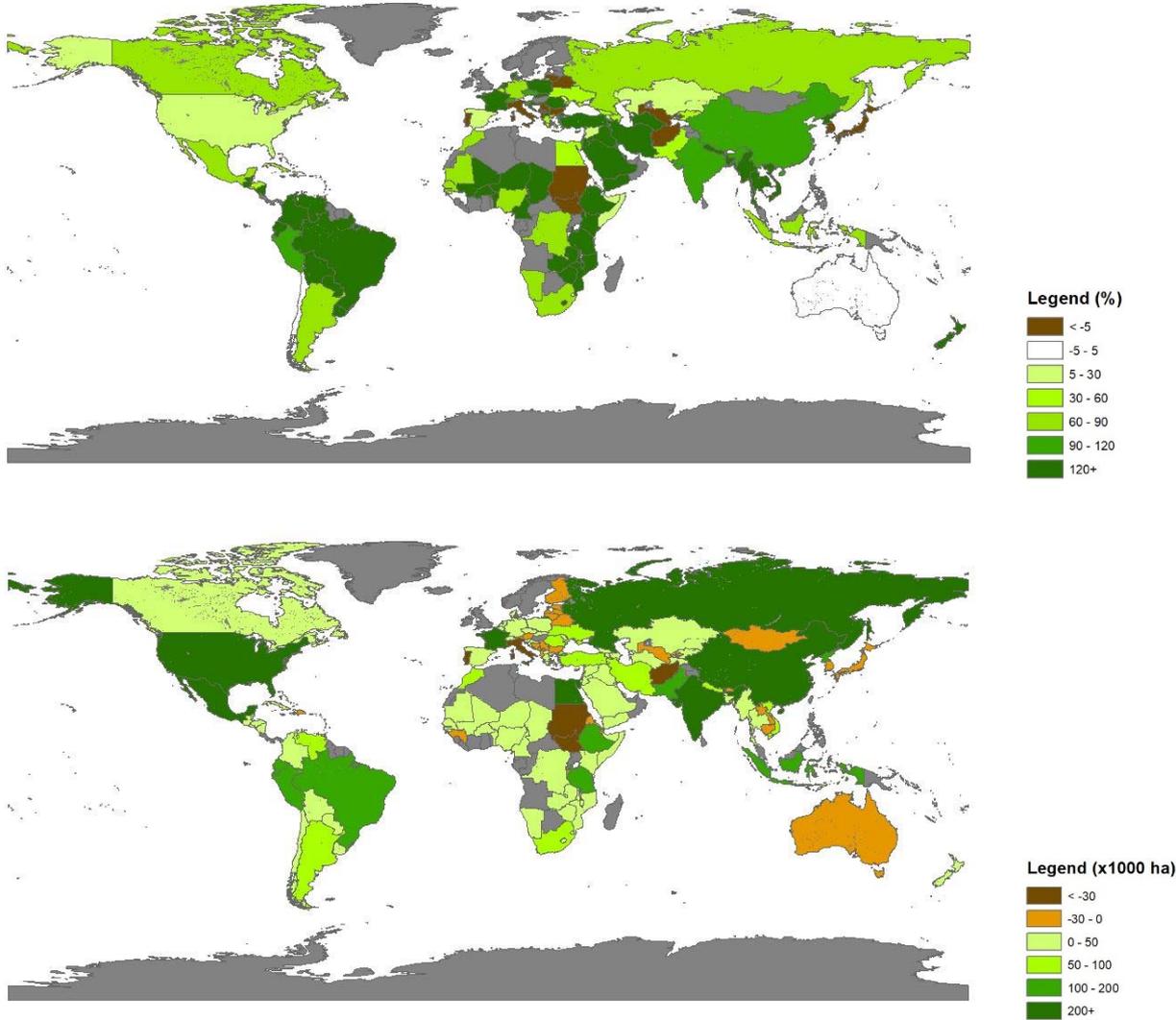


Figure 5. 15: relative and absolute national AEI growth rate, 1970-2019

By observing the maps, it can be seen that, as expected, national AEI significantly raised in most of countries; however, there are some exceptions in which the area equipped for irrigation reduced between 1970 and 2019 (such as Sudan, Portugal, Italy, Afghanistan, etc.) or it

remained rather constant (Australia and Chile). In US, the relative increment of maize AEI is not so high (< 30%), probably due to the already well-disposed irrigation system in the country, while in many African countries, such as in the Arabic peninsula and South America, the growth rate exceeds 120%. In fact, by looking at the Blue uWF maps, we see that many of these countries totally depend on irrigation to satisfy crop water needs (e.g., countries in Arabic Peninsula), with Blue uWF exceeding the green component, therefore it is expected that these regions invested on the expansion of the irrigation systems.

The map of the absolute increment of national AEI shows a different situation: there are countries with very little absolute AEI increase which instead had high rates in the relative map (e.g., several African and Arabic countries), while there are nations with a huge increment in the absolute AEI (such as US) even though the relative growth was very small. This difference depends on the total Area Equipped for Irrigation of the country, which gives different weight to the relative increment depending on its absolute dimensions. For a nation like US, which has a very extended irrigated surface, a difference of more than 200 thousand hectares is not so significant since its dimensions in 1970 were almost 4 million ha. In Yemen, instead, which had 9000 ha in the early '70s, a difference of 12000 ha constitutes a growth rate that exceeds 120%.

5.5 Temporal evolution of total, green and blue uWF of some countries

As we did in the previous parts, we choose Italy, US, Australia, Nigeria and Viet Nam as testing countries to visualize the historical evolution of Total, Green and Blue uWF. Firstly, we look at the time series of total uWF of the five countries and check the differences among them.

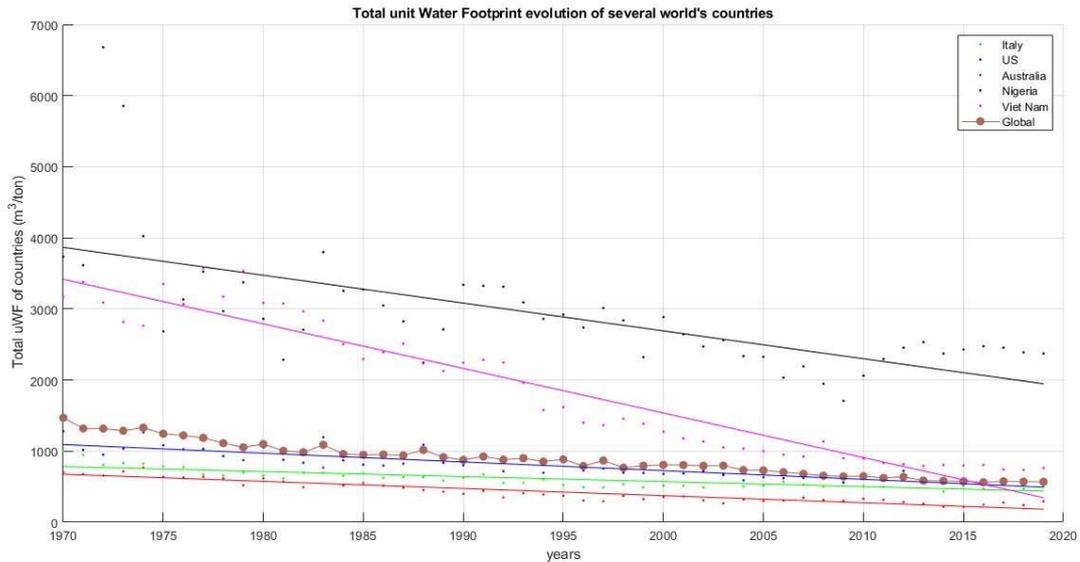


Figure 5. 16: historical evolution of total uWF of several nations

Looking at this graph, it can be clearly observed that, generally, green + blue uWF is decreasing in all the countries into analysis, such as for the global mean. Western countries show values lower or at least coincident with the global mean throughout the time interval, while developing countries, in this case Nigeria and Viet Nam, exhibit very high estimates of unit water footprint, which however are significantly decreasing in time. More specifically, among western countries, Italy and Australia are the countries that always showed the lowest values of uWF (always less than 1000 m³/ton), while data of US manifest a pattern which is very similar to the global mean. In fact, global uWF is calculated with a production-weighted mean: therefore, countries which produce more (such as US) are also more impacting on the global mean, while those with a limited production have a reduced effect. Regarding the trend, it is crystal clear that Viet Nam is the one with the sharpest decrease, passing from approximately 3500 m³/ton in 1970 to less than 800 m³/ton in 2019, reaching values comparable to those of the western countries. Nigeria is instead the nation with the greatest water consumption per unit of product; in the 70s its data were around 4000 m³/ton, but the trend didn't fall such rapidly as for Viet Nam, remaining above 2000 m³/ton also after year 2000. In addition, in the first decade of the time series, Nigerian uWF exhibits some very sharp fluctuations, with values crossing 6600 m³/ton in years 1972-73. Probably, this is due to the lack of reliability of the data, so they should be intended as outliers rather than effective values.

Now, turning into details, we display the time series of green and blue uWF country by country, comparing their values to the respective global mean of green and blue uWF and analysing the trends of the regression lines.

We start with the time series of Italy, one of the countries with the lowest water consumption per unit of product. It can be seen that values of green uWF are well below the global mean, while those of blue unit water footprint are fluctuating up and down the line of the global mean. In addition, while for the green line we notice a clear declining tendency, for blue uWF there is no evident trend, in fact the polynomial regression line looks flat, with an angular coefficient very close to zero.

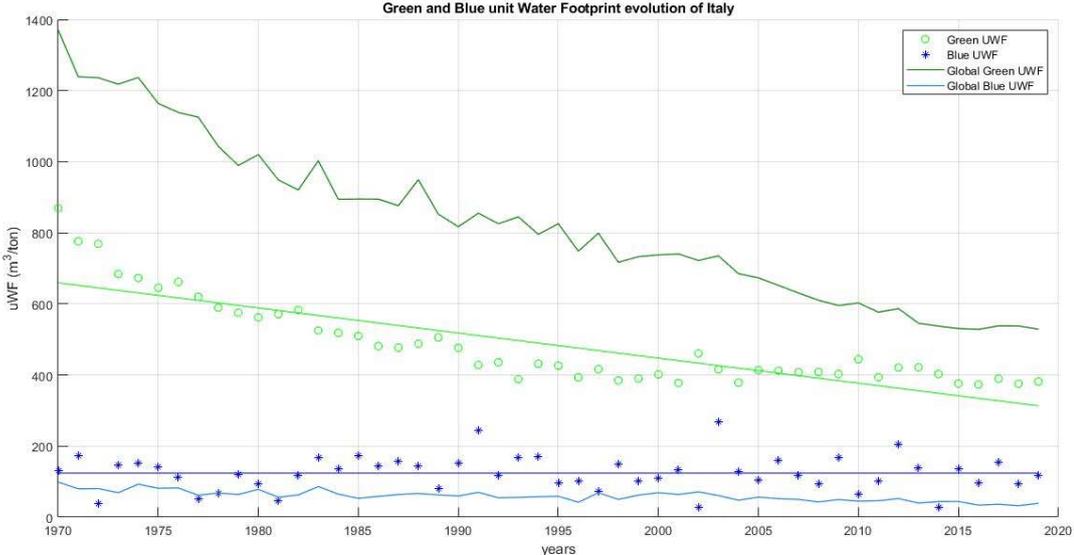


Figure 5. 17: time series of green and blue uWF, Italy

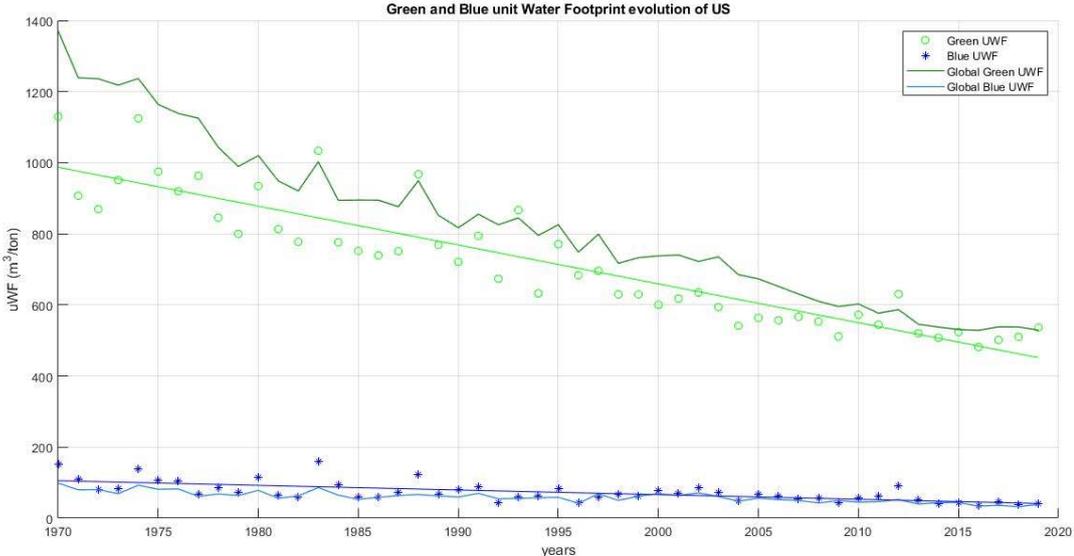


Figure 5. 18: time series of green and blue uWF, US

Also for USA, the behaviour is very similar to Italy: green uWF is significantly decreasing, while blue uWF presents a descending pattern which is rather imperceptible. It is worth noticing that the various spikes in the time series of global green uWF are associated with the fluctuations of US green uWF data, remarking the concept that US, being the nation with the greatest maize production, is the country that mostly affects global mean patterns. This phenomenon is clearly visible also in blue uWF regression line, which perfectly overlays the global mean blue unit water footprint.

Regarding Australia, it can be observed that, differently from the previous countries, green and blue uWF values are very close to each other, which means that in the country irrigation component displays a greater contribute. However, both trends are evidently declining toward values around 200 m³/ton.

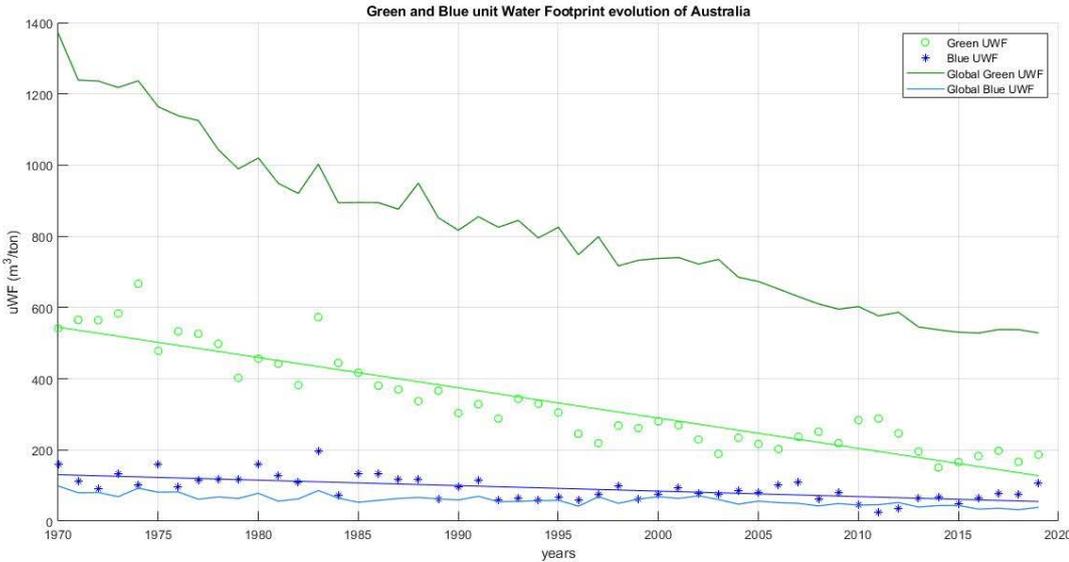


Figure 5. 19: time series of green and blue uWF, Australia

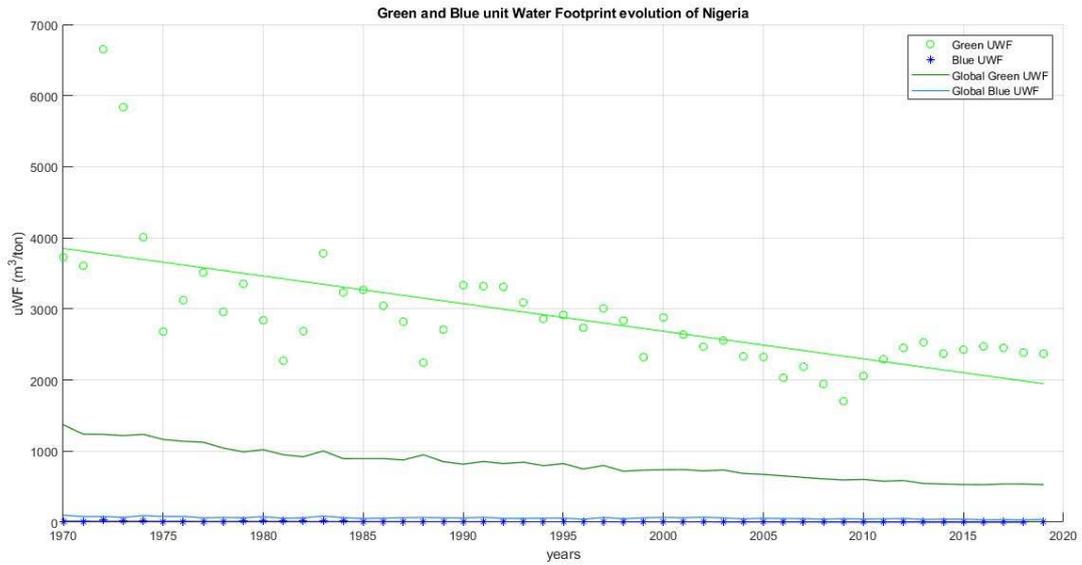


Figure 5. 20: time series of green and blue uWF, Nigeria

As said previously, Nigerian uWF is the one among the testing countries with the greatest water use per unit of product. The local crops require great water inputs to compensate the strong evapotranspiration of the plants which is due to the very hard climatic conditions. In addition, the country doesn't show values in the blue uWF time series, mainly because most of local crops are not equipped for irrigation, therefore all water input is provided by precipitation. The last country is Viet Nam. Also this nation almost fully relies on precipitation, nevertheless green uWF is drastically falling, especially thanks to the sharp increase of the crop yield, which grew from 1 ton/ha to almost 5 ton/ha in 50 years (see figure Figure 3. 3).

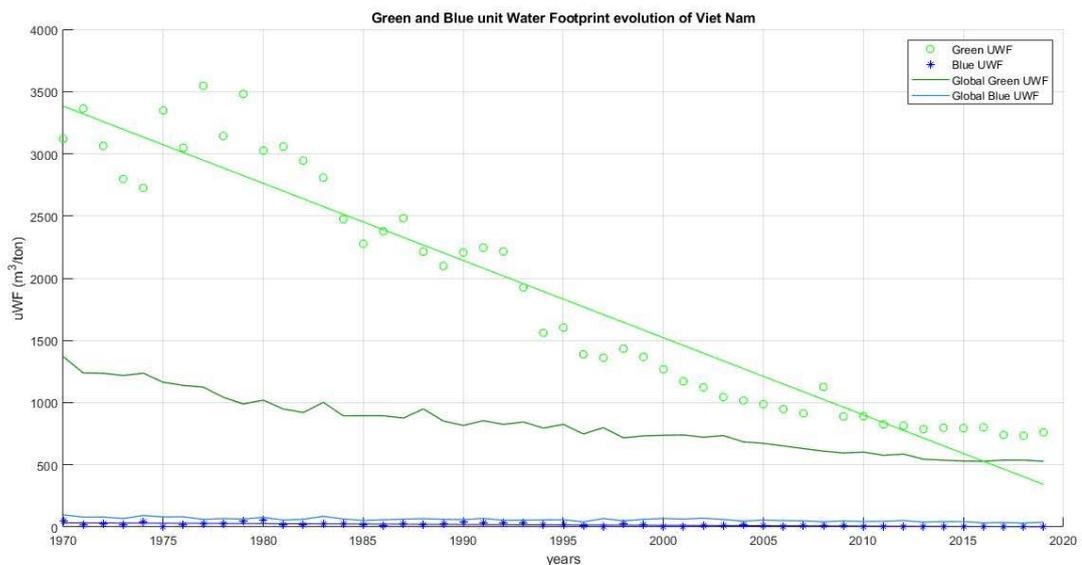


Figure 5. 21: time series of green and blue uWF, Viet Nam

5.5.1 Trend analysis of blue uWF and crop yield

The reason why Italian blue uWF trend is almost flat is explained by the fact that the numerator of blue uWF equation (ET blue) grew with the same rate of denominator (yield), resulting in a flat overall tendency.

Therefore, it is worth investigating this feature over the rest of the world's countries, using the level of significance of the blue uWF and crop yield. The best way to perform this operation is by applying the t-Student test over the time series of the two variables, reporting the results in a matrix that pulls them together for each country. More specifically, we create a matrix named *T-student_array*, that for each nation it reports the results of the t-Student test applied on blue uWF variable and crop yield, placing "0" if the test variable T is lower than t_{lim} , thus if there is no evident trend in the regression, "1" if there is a significant positive trend, "-1" if the trend is significant and negative.

Once the test is made running over all the countries and results are inserted in *T-student_array*, the outcomes can be displayed on a global map. With another iterative loop, the code assigns a number, X , to a country according to its condition (e.g., "0" if in country N there is no trend in blue uWF and positive in yield, "1" if there is no trend in both uWF and yield, "2" if there is no trend in uWF and negative in yield, etc.), then the results are represented in a 5 x 5 arc min resolution world map, where the resulting number X is inserted in all the pixels of the respective country. In this way the countries are coloured according to their condition.

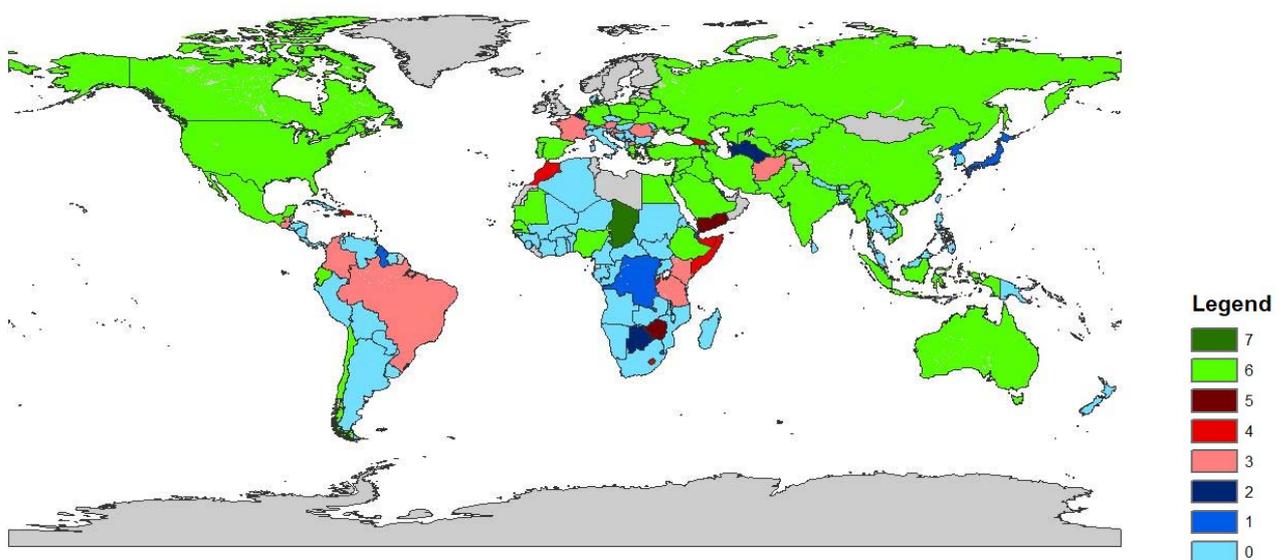


Figure 5. 22: t-Student test of uWFb and yield trend

Table 5. 1: description of legend numbers

Legend number	Trend in blue uWF	Trend in yield
7	Negative	No trend
6	Negative	Positive
5	Positive	Negative
4	Positive	No trend
3	Positive	Positive
2	No trend	Negative
1	No trend	No trend
0	No trend	Positive

5.6 Results verification

The two main ways to verify the goodness of results consist in applying a verification of internal and external coherence. In particular, the internal coherence aims at comparing the national uWF time series using the same input data, but two different approaches, while the external coherence checks if the obtained results are consistent with those obtained by other works. More specifically, we refer to the values provided by the Water Footprint Network, which provides green, blue and grey unit water footprint estimates at subnational scale, weighted along the time-interval 1996-2005.

5.6.1 Comparison with Water Footprint Network

In order to make data from the Water Footprint Network comparable with the results obtained from our work, we need to perform quick pre-processing operations. Firstly, data from Excel are imported into Matlab. They show up as two arrays of 3252 values, where each row of the first one contains the uWF of a specific region in a nation and in the second we have the name of the nation in which that region belongs to. With the *unique()* function, we extract the unique values of each nation and, with a *for* loop, we perform the mean of the regional unit water footprint values for each country. Then, we load the Excel file “Paesi_FAO”, that contains the names and FAO codes of all the countries which, according to our previous calculations, own

values of unit water footprint. In this way, with a second *for* loop, we associate a FAO code to the countries of the Water Footprint Network dataset. This operation is done for both green and blue unit water footprint.

Regarding data from our data processing, we perform the arithmetic mean of green and blue uWF values from 1996 to 2005, to obtain average values as closer as those contained in the Water Footprint Network dataset.

We are able now to create the scatter plot. As we did at the very beginning of this work, while comparing processed data of yield and harvested areas versus FAOSTAT data, we create a three-dimension scatter plot, where unit water footprint values are weighted with national maize production. In this way, we can understand whether a potential incoherence between the uWF according to the Network or to our calculations in a country is more or less significant.

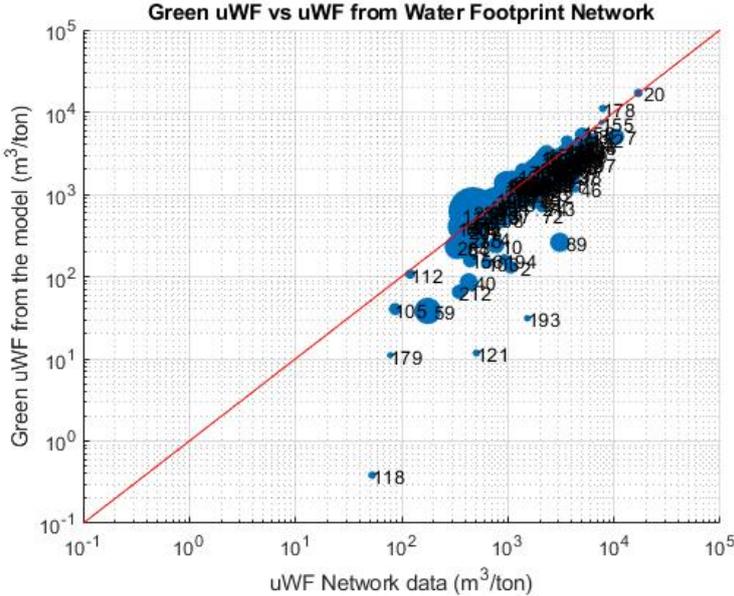


Figure 5. 23: green uWF compared with Water Footprint Network

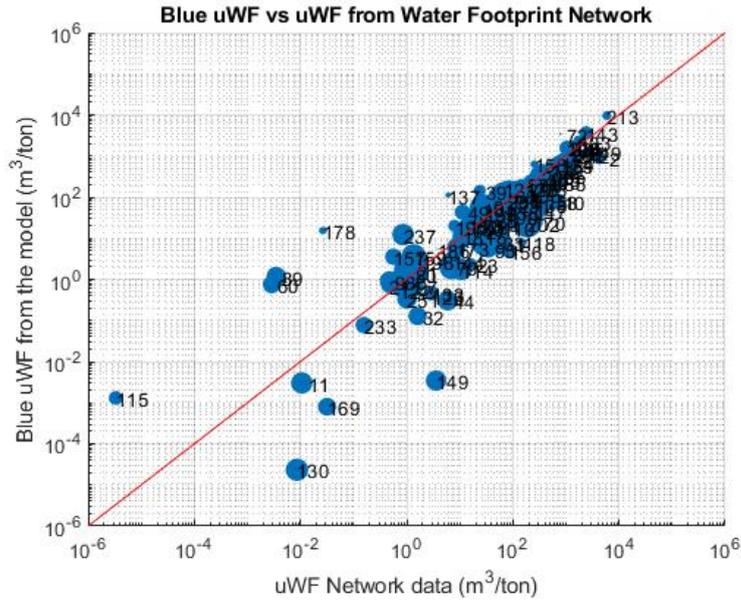


Figure 5. 24: blue uWF compared with Water Footprint Network

At a first glance, we can notice that green uWF scatter plot shows better alignment of the two arrays with respect to Blue uWF, where many large nations diverge from the bisector, especially those which exhibit low unit water footprint. In addition, while the range of green uWF data is shorter, between approximately 10^1 and 10^4 m^3/ton , national blue water footprints assume a wider range of values, between 10^{-5} and 10^4 m^3/ton .

To have a deeper look into details and quantify the incoherence between countries, we estimate the relative error between the national uWF values with our approach and those according to the Water Footprint Network, with the equation

$$Relative\ error = \frac{uWF_{this\ study} - uWF_{network}}{uWF_{network}} \quad (5.8)$$

At this point, we set a threshold in the relative error to identify the countries with the worst alignment. More specifically, we set 0.7 (70%) and 1 (100%) for green and blue unit water footprint respectively. The threshold of the second is greater than the one of the former because in blue water footprint the error between uWF estimations is generally higher.

Here the tables reporting the countries with estimation errors greater than the thresholds, with their corresponding green and blue uWF, and maize national productivity.

Table 5. 2: countries with error in blue uWF greater than 1

Country code	Country name	uWF Blue (m ³ /ton)	Production (kton)	Relative error (-)
115	Cambodia	3.32E-06	156.97	393.03
60	El Salvador	2.86E-03	582.68	259.90
89	Guatemala	3.50E-03	1053.55	334.49
178	Eritrea	2.64E-02	0.45	579.80
157	Nicaragua	0.56	181.00	5.30
237	Viet Nam	0.86	2005.90	13.37
159	Nigeria	1.36	3.78	1.66
137	Mauritius	5.99	0.62	17.78
195	Senegal	7.96	78.59	1.51
49	Cuba	11.26	273.20	2.82
39	Chad	23.80	64.01	5.01
121	Lebanon	68.86	3.50	1.44
158	Niger	250.91	412.20	1.52
72	Djibouti	783.90	0.01	3.34

Table 5. 3: countries with error in green uWF greater than 0.7

Country code	Country name	uWF Green (m ³ /ton)	Production (kton)	Relative error (-)
118	Kuwait	52.70	5.55	-0.99
179	Qatar	78.09	5.32	-0.86
59	Egypt	175.98	6474.45	-0.78
212	Syrian Arab Republic	347.29	190.50	-0.81
40	Chile	428.51	652.02	-0.80
121	Lebanon	505.16	3.50	-0.98
103	Iraq	629.95	55.00	-0.77
194	Saudi Arabia	930.95	40.61	-0.83
4	Algeria	975.94	1.56	-1.00
2	Afghanistan	1062.76	115.00	-0.88
193	Sao Tome and Principe	1526.92	2.23	-0.98
72	Djibouti	1880.25	0.01	-0.70
89	Guatemala	3081.68	1053.55	-0.92
55	Dominica	3738.63	0.18	-1.00
46	Congo	4314.57	6.36	-0.73

From these tables, we can clearly see that national blue water footprint shows much greater errors with respect to the green component. In fact, by setting a greater threshold in blue uWF

(1 versus 0.7), the table of blue unit water footprint contains 14 countries with bad estimations, while the table of uWF green contains 15 countries which, however, never exceed 1 in the error. Results are ordered from the country with lowest uWF value to the greatest.

It can be observed that there is a sort of correlation between the magnitude of blue uWF and the error. More specifically, the greatest errors are associated to very low values of unit water footprint, while, by increasing national uWF, we assist to a general decrease of the incoherence between Unit Water Footprint Network data and our estimations. This phenomenon is crystal clear in the blue uWF case since there are countries with very low uWF estimates: in fact, those with values lower than 10^{-2} m³/ton show errors in the order of 10^2 , while unit water footprint values greater than 10 m³/ton are associated to errors ranging between 1.5 – 5.

In green unit water footprint table, instead, this characteristic is not evident, mainly because green uWF estimates are generally greater than the blue ones (they range between 10^1 – 10^3) and the error is always below 1.

The reason why blue uWF shows greater errors, particularly in the lowest estimates, can be explained by the fact that historical series of variables for obtaining ET_b may not always be reliable worldwide. In fact, by looking at both tables, it can be noticed that all the countries involved are developing nations, belonging to African, Middle East, Asiatic and Central-South American regions. In fact, it is more likely to have greater data uncertainties in these countries rather than in European and North American nations, where it's usually easier to collect reliable agricultural data estimations. In addition, with such low estimates of national blue uWF, it is very hard to find a coherence between our values and those of the Water Footprint Network, since even a very small difference in the estimation reflects into a large relative error.

5.6.2 Verify coherence of results comparing bulk versus local-weighted data

As we said at the beginning of this chapter, it is also possible to check the consistency of the estimations by applying an internal coherence method, that verifies the results without comparing them with external data. In particular, national uWF time series obtained with a “bulk” calculation, that is by performing a weighted mean of ETs on their rainfed and irrigated areas, are compared with local-weighted data, obtained with the following equation,

$$uWF_{weighted}^N = \frac{\sum_{i \in N} (uWF_i \cdot P_i)}{\sum_{i \in N} (P_i)}$$

(5.9)

where uWF_i is the i -th cell of nation N obtained with the bulk approach and P_i is the crop production in the same cell. This approach performs a production-weighted mean of the uWFs of a country, where at each pixel of country N is assigned a weight according to its production, then they are summed, and the result is divided by the total national production in country N .

The time series of the results are then compared with the national green and blue uWF regressions calculated with bulk approach, choosing Italy, US, Australia and Viet Nam as testing countries for this verification analysis.

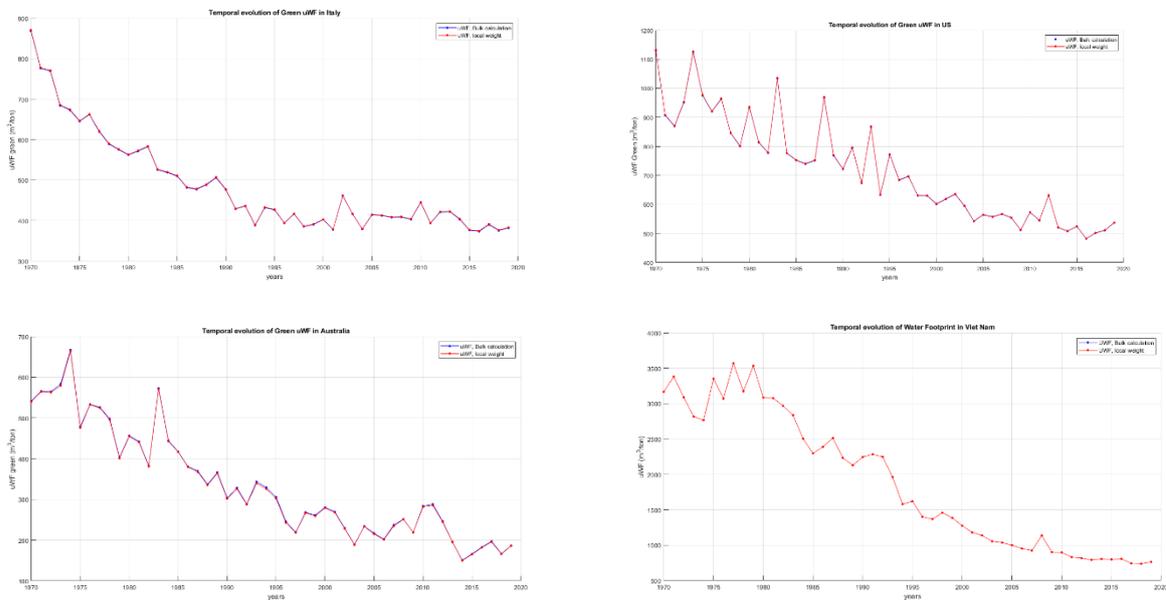
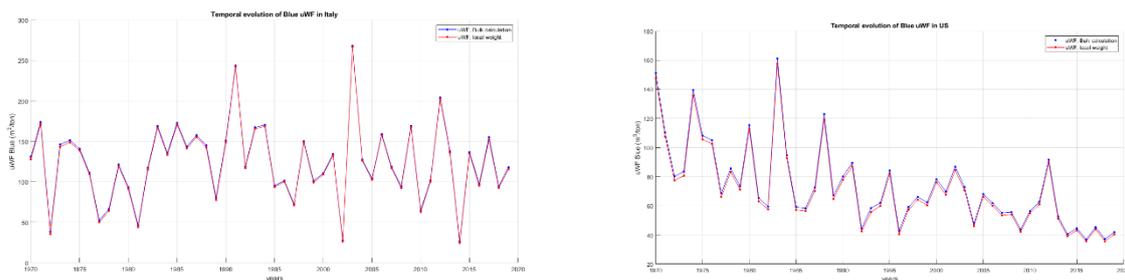


Figure 5. 25: green uWF time series calculated with bulk and local weight approaches



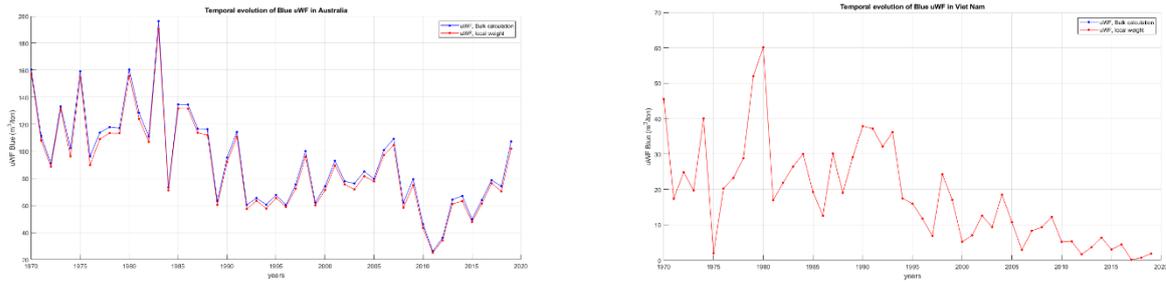


Figure 5. 26: blue uWF time series calculated with bulk and local weight approaches

Looking at the line plots, it is immediately visible that the internal coherence approach provides a satisfying fitting between bulk uWF time series and those obtained with the local weights. The time series of green uWF calculated with both approaches perfectly overlap in all countries, while, for blue uWF, we assist to very mild discrepancies between the two patterns in US and Australia. This is explained by the fact that there are countries with more homogeneity in the climate such as in the distribution of production, therefore many areas are supposed to contain values similar to the national mean, while other nations (like Australia) show greater heterogeneity in these two variables. In fact, there are warmer areas with greater water consumption but very low production and others with a more favourable climate that allows a greater production with lower uWF; this causes a misalignment of local weights of uWF from the original national values.

5.6.3 Comparison with Fast Track approach

As we said in the chapter “Methods and Data”, the Fast Track (FT) is a quicker and less computational demanding approach which allows the determination of unit water footprint without relying on complex hydrological models or an excessive amount of input data. More specifically, this method assumes that the variations of uWF are only dependent by crop yield patterns, while the evapotranspiration trend is supposed to be constant throughout the duration of the time-interval (Tuninetti et al., 2017). Comparing data which adopted a detailed approach versus those obtained with the Fast Track allow us to visualise the validity of this method, by looking at the historical evolution of both variables and the goodness of the estimation country by country.

Fast Track data are organised as a 255 x 56 matrix, where each row contains the historical time series of national uWF a country from 1961 to 2016. To be compared with uWF data with detailed approach, it is firstly necessary to execute a quick pre-processing to adjust the data and make the two matrices comparable (i.e., by having an equal time-interval and the list of countries ordered in the same way). Since FT data have a time-interval between 1961 to 2016 while uWF from rigorous approach are from 1970 to 2019, the chosen time interval will be the range 1970-2016.

At this point the uWF patterns of some countries are displayed on a line plot, choosing again Italy, US, Australia and Viet Nam for the test.

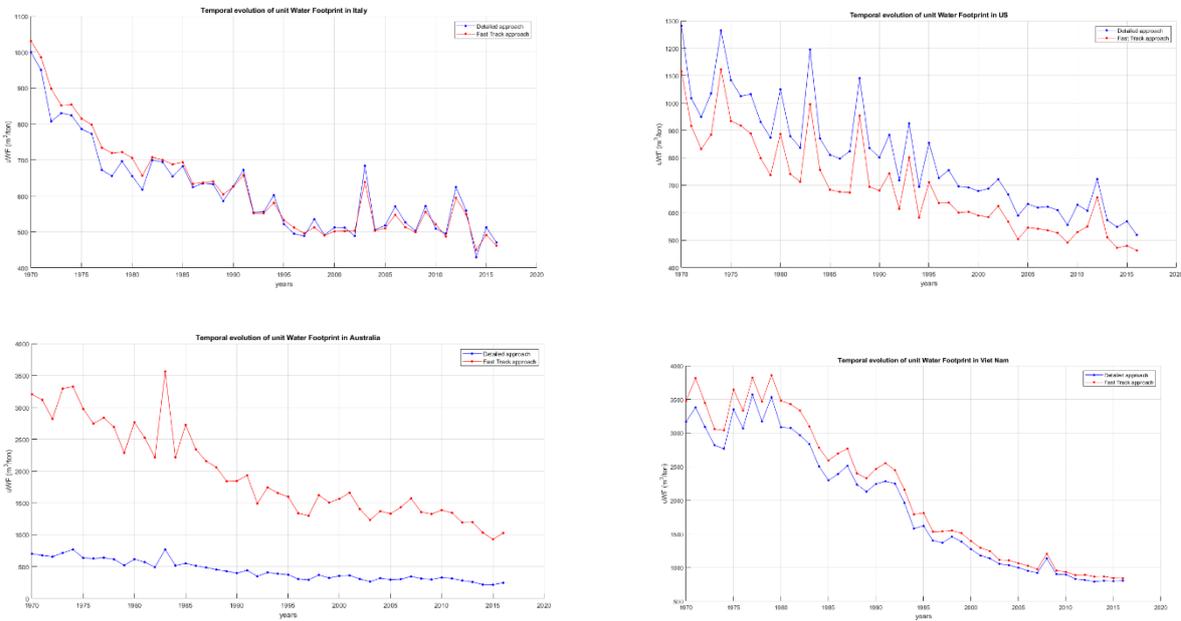


Figure 5. 27:time series of uWF calculated with detailed and Fast Track approaches

As we can see from the illustrations, apart from Italy, that manifests a good fitting between the two variables, in the other countries there is an evident distance between the regression obtained with the detailed approach and the one derived from the Fast Track. However, even if the curves do not coincide, the pattern of both regressions is very similar not only in the trend, but also in yearly fluctuations.

Moreover, it is worth reproducing a scatter plot displaying the coherence of the two estimations in all countries. To do this, we repeat the procedure adopted for the comparison of uWF results with Water Footprint Network data, that is by creating a 3D scatter plot with national values averaged in the period 1996-2005 and where each dot of the scatter has a size according to its national productivity.

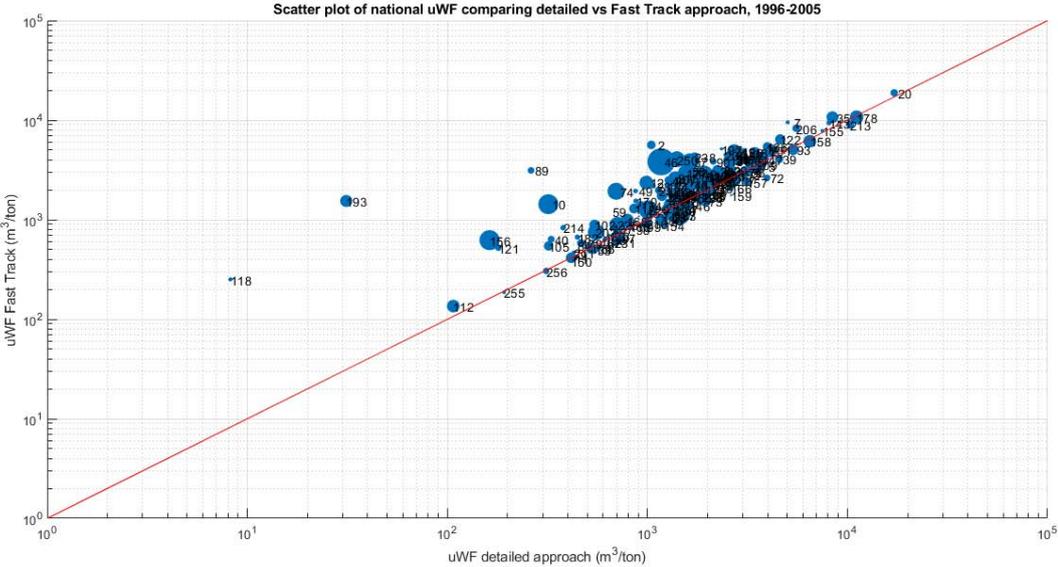


Figure 5. 28: uWF data between detailed approach vs Fast Track

At a first glance, we notice that most of countries show a good alignment, except for few nations where the estimate is completely out of the track (countries 193 and 118 in particular). Therefore, we calculate the relative error between the two variables, isolating the countries which have an error greater than 1 (100%) and reporting them in the following table:

Table 5. 4: countries with relative error greater than 1 (detailed vs FT approach)

Country code	Country name	UWF (m ³ /ton)	Production (kton)	Relative error (-)
118	Kuwait	8.23	1.77	29.55
193	Sao Tome and Principe	31.25	881.59	48.40
156	New Zealand	162.55	17556.90	2.84
121	Lebanon	179.81	18.83	1.90
89	Guatemala	262.07	25.90	10.94
10	Australia	319.87	16780.65	3.48
214	China, Taiwan Province of	379.12	8.90	1.19
74	Gabon	697.22	6474.45	1.77
49	Cuba	872.46	3.79	1.22
12	Bahamas	988.03	1851.73	1.39
2	Afghanistan	1045.94	115.00	4.40

23	Belize	1064.38	9.35	1.06
46	Congo	1171.30	106000.00	2.26
250	Democratic Republic of the Congo	1402.53	4472.90	1.89
37	Central African Republic	1631.28	6953.70	1.33
238	Ethiopia	1723.83	178.32	1.47
197	Sierra Leone	2346.30	0.45	1.19

To summarize, the table contains the list of countries where yield variable is not able enough to correctly estimate crop water footprint. In fact, most of these countries belong to desertic or tropical regions, where climate plays an important role in the determination of uWF, thus relying on yield variability only leads to a mismatch in the estimation. In addition, another reason for this misalignment may be due to a weak crop yield's trend, thus a time series showing sharp fluctuations and an unclear increasing or decreasing tendency, thus the climatic component gains importance in the ratio ET^N/Y^N .

5.7 From unit water footprint [m^3/ton] to water footprint [m^3]

So far, the work focused on the global and local time evolution of unit Water Footprint, uWF, that is the volume of water required to produce a unit of product. This is an indicator that expresses the agricultural efficiency in water use, however there is no mention about the effective amount of water which has been employed by a country to produce the annual quantity of maize. This indication is provided by Water Footprint, WF, which expresses the total cubic meters of water used to produce given units of crop and it is obtained with the equation

$$WF_c^N = uWF_c^N * P_c^N \quad (5.10)$$

where uWF_c^N is the unit water footprint of crop c in country N and P_c^N is the national production of the same product in nation N .

Having the matrices of national total, green and blue uWF already, such as those of national production, the calculation of green and blue WF is simply the product between the two. Once obtained these matrices, it is possible to display the annual series of some countries (Italy, US,

Australia, Nigeria and Viet Nam) such as the global evolution of this variable, to see if their trends are similar to those of uWF or if they follow a different shape.

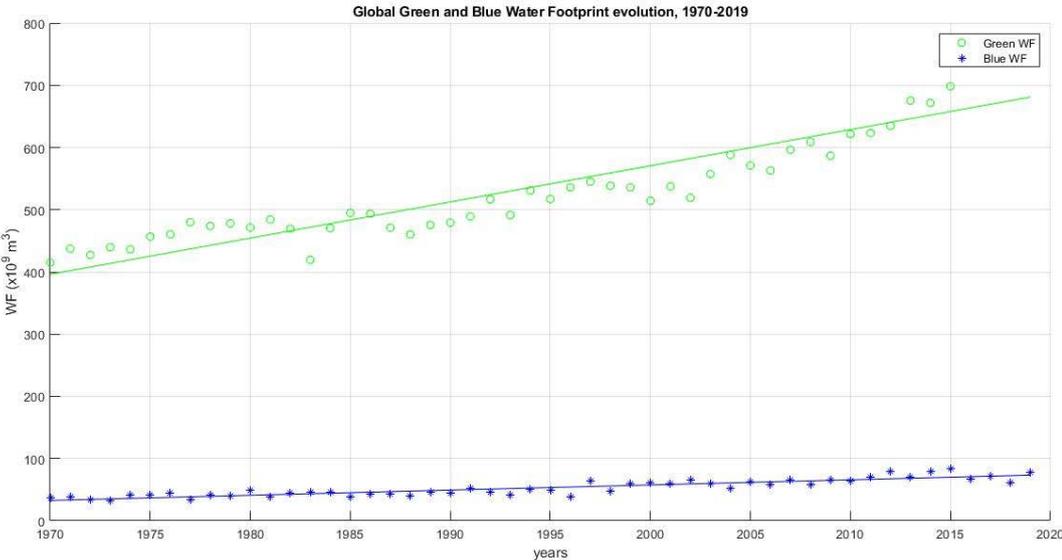


Figure 5. 29: global pattern of Green and Blue Water Footprint

By observing the historical time series of global green and blue WF, we immediately notice that the trend is reversed with respect to the global mean uWF regressions. In fact, even though water consumption per unit of product significantly reduced (green contribute) or at least remained quasi-constant (blue uWF), the global volume of water consumption for maize production increased both in the green and blue contribute. More specifically, green WF raised from 400 Gm³ to 700 Gm³ in 50 years, while the amount of water from irrigation has doubled, passing from 40 Gm³ to almost 80. Although water use efficiency increased in the past 50 years, the global maize production intensified with a greater rate, thus increasing the effective water volume required by the crop.

Here below the time series of Green and Blue WF of Italy, US, Australia, Nigeria and Viet Nam, with a remark on the angular coefficients of the linear interpolations of green blue data, reported in Table 5. 5.

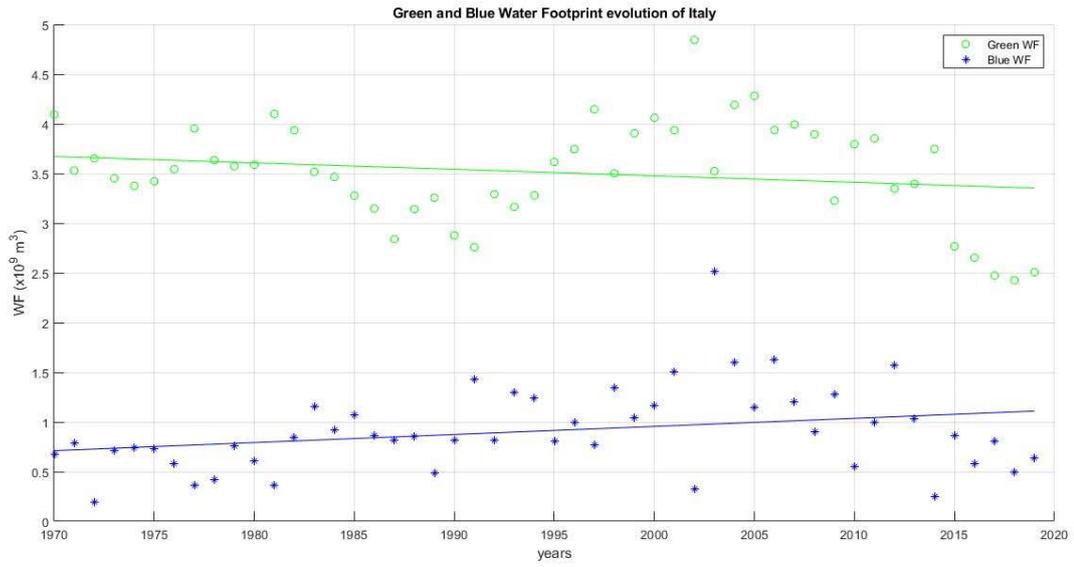


Figure 5. 30: time series of Green and Blue WF, Italy

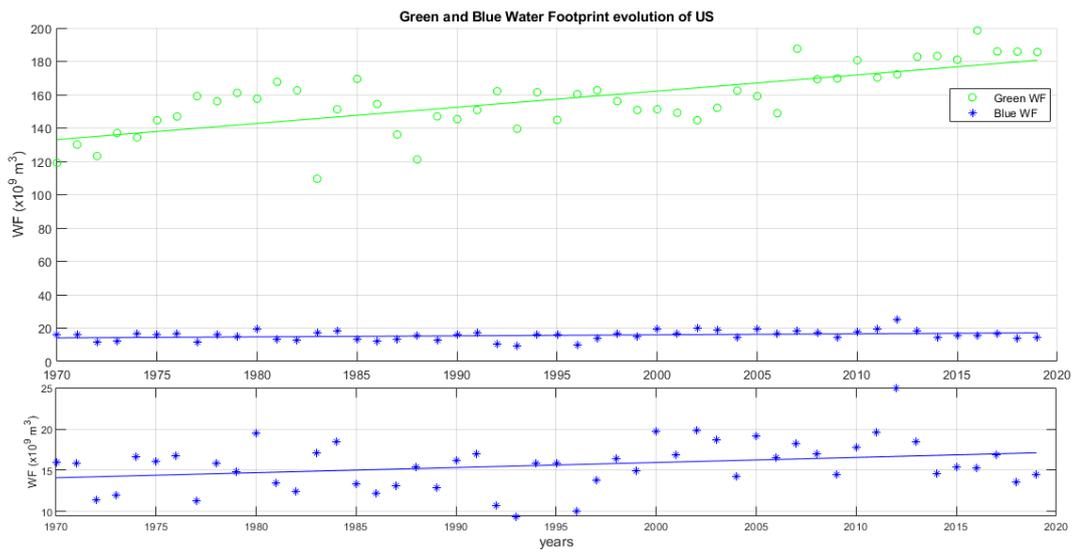


Figure 5. 31: time series of Green and Blue WF, US

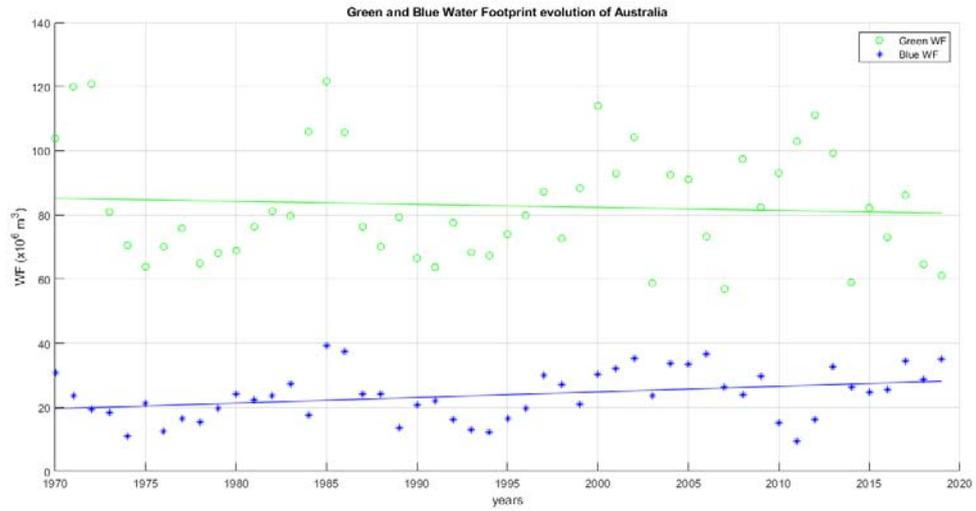


Figure 5.32: time series of Green and Blue WF, Australia

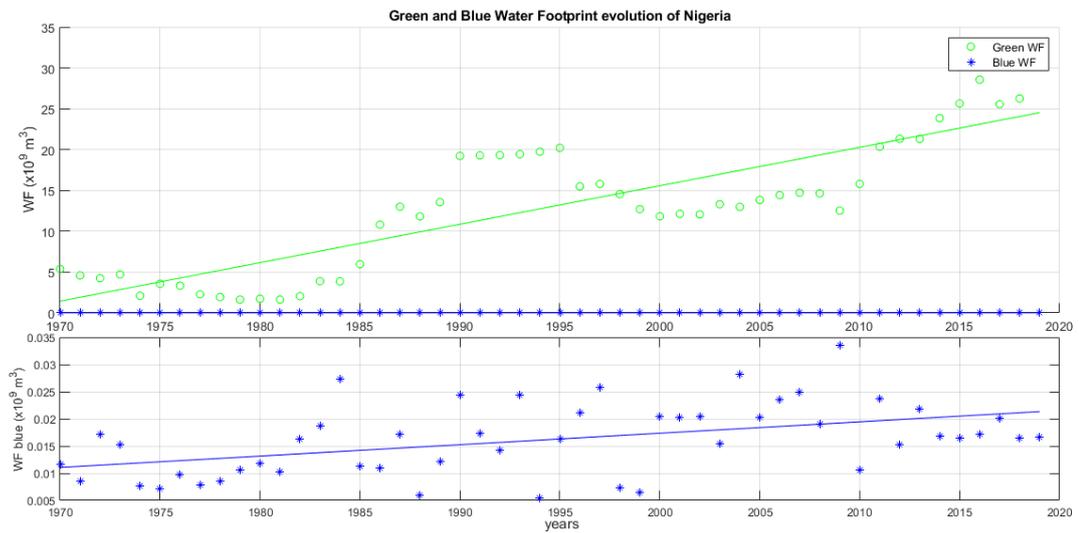


Figure 5.33: time series of Green and Blue WF, Nigeria

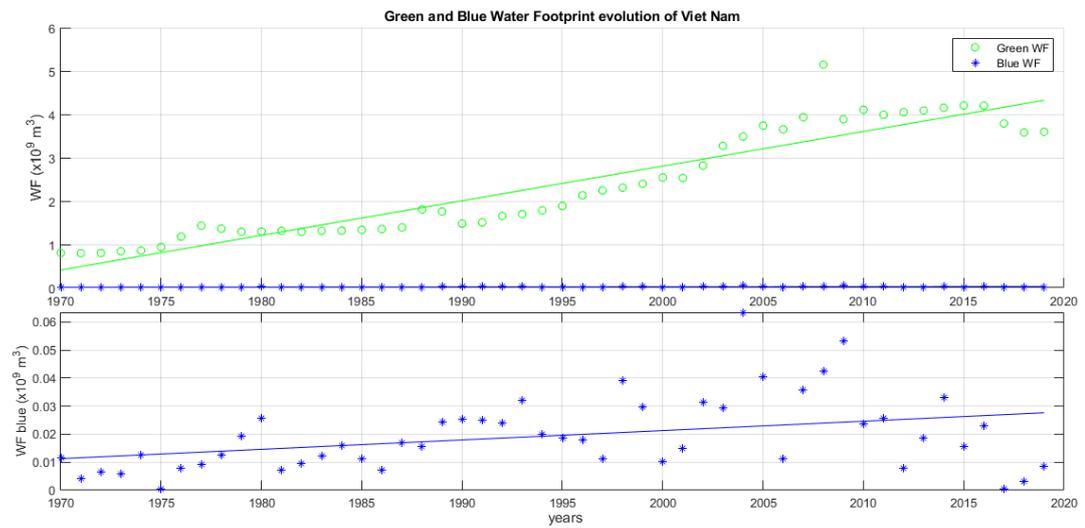


Figure 5.34: time series of Green and Blue WF, Viet Nam

Table 5. 5: slope coefficient (m) of green and blue water footprint

Country	$m_{\text{blue}} (\times 10^6 \text{ m}^3/\text{y})$	$m_{\text{green}} (\times 10^6 \text{ m}^3/\text{y})$
Italy	8.84	-7.28
US	61.69	970.83
Australia	0.17	-0.09
Nigeria	0.20	470.78
Viet Nam	0.33	79.87

Looking at the time series of Italy, it is possible to observe that the green WF has a decreasing tendency, while the trend of the blue contribute is instead increasing at a greater rate. Comparing these regressions to the time series of the input variables of Italy (see Annexes, Figure 7. 6), we notice that green and WF fluctuations follow the pattern of harvested area and production.

In the US time series, instead, there is a growth on both green and blue components. An interesting detail of this graph is that, in 1983, there is a sudden drop of the green contribute, suggesting that a serious drought may have happened. However, blue WF value of that year didn't exhibit a sharp increase to balance the lack of precipitation, but it remained very close to the mean. The paper provided by W. M. Wendland (1984) in the Bulletin of the American Meteorological Society reports that, in 1983, an exceptional dry summer involved the Upper Mid-West of US, with mean July and August temperatures above $+2^\circ\text{C}$ the monthly averages of the period 1951-80. However, most of US irrigated areas are concentrated in central US, so probably they haven't been affected so much by the drought, which instead involved more the rainfed-only areas that couldn't balance with the blue contribute to keep the crop at field capacity.

However, the previous method is not the only way to calculate green and blue water footprint, since it can be derived also from evapotranspiration [mm] and area [ha], using the equations

$$WF_g^N = \sum_{i \in N, rf \cup irr} (ET_{a,i}^{rf} \cdot A_{rf,i} + ET_{g,i}^{irr} \cdot A_{irr,i}) \quad (5. 11)$$

$$WF_b^N = \sum_{i \in N, irr} (ET_{b,i}^{irr} \cdot A_{irr,i}) \quad (5. 12)$$

that is, by multiplying ET_g or ET_b times the rainfed and irrigated area of the i -th cell and then making the sum of the products for all the cells of country N . This is a procedure that avoids

the calculation of uWF and national production, since it calculates WF from evapotranspiration and area data, thus it is quicker and more direct. To verify if data obtained with both approaches are coherent, we compare the time series of Italy obtained in both ways.

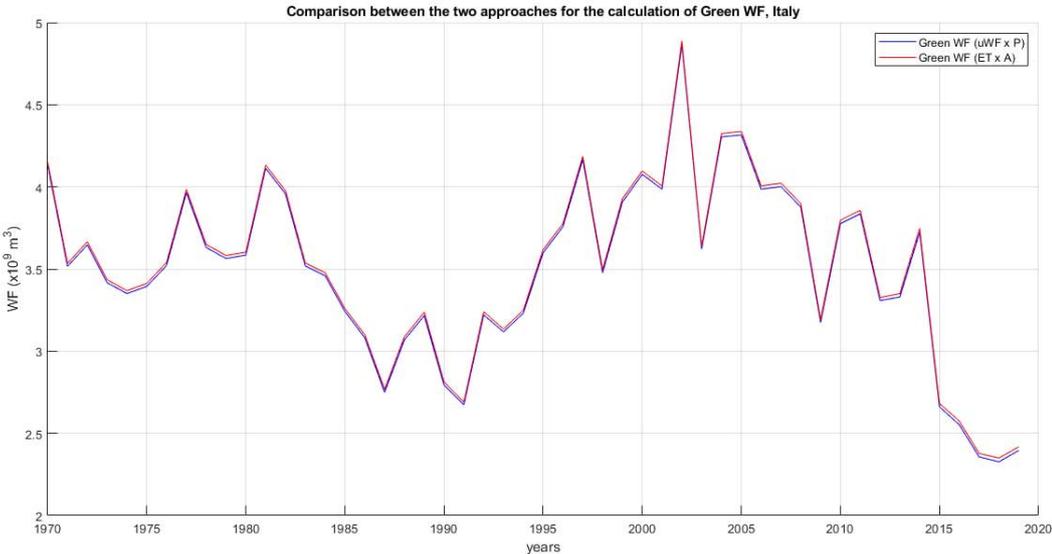


Figure 5. 35: comparison between the two approaches for the calculation of green WF, Italy

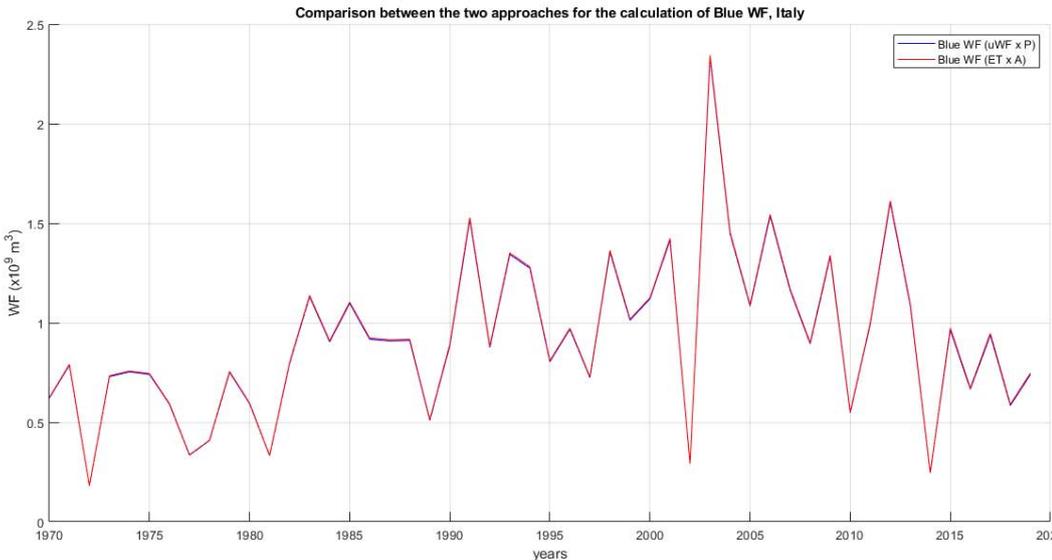


Figure 5. 36: comparison between the two approaches for the calculation of blue WF, Italy

From both illustrations we can appreciate that the regressions of both approaches well fit each other, with a cv_RMSE (Coefficient of Variation of RMSE) equal to $5.6 \cdot 10^{-3}$ and $6.2 \cdot 10^{-3}$ for green and blue WF respectively. Thus, we can state that both methods are perfectly suitable for the description of water footprint at national scale, both for Italy and all other world's nations.

6 Conclusions

In the present thesis, the temporal variability of green and blue unit water footprint of maize, has been analysed. Unit Water Footprint [m^3/ton] is the indicator that expresses the water volume employed to produce a unit of product, and it can be partitioned into green uWF, the water that comes from precipitation, and blue uWF, the contribute coming from irrigation. Values were obtained from high-resolution modelling and intense data analysis, relying on national time series and gridded data, which account for spatio-temporal variations of this crop worldwide. We noticed that there is a strong spatial heterogeneity of this indicator around the world, where countries with higher technological development and more favourable climatic conditions are those with the lowest values of green and blue unit water footprint, thus a more efficient use of water resources. Along the considered time-interval, we observe a global decrease of unit water footprint. While for the green contribution this phenomenon is widespread in all the analysed countries, the blue component has a less remarked decreasing trend, that for some countries shows an angular coefficient very close to zero.

Looking at the historical evolution of green uWF on the global maps, we see that from 1970, where the only areas with values lower $1000 \text{ m}^3/\text{ha}$ were in East China, Central Europe and some regions in the Middle East, the values drastically reduced worldwide, especially in the US, in the rest of Europe, in China and in South America. Africa, instead, shows a rather negligible change with respect to the rest of the world.

Blue uWF maps, instead, exhibit a different pattern: in 1970, the lowest values belong to the countries with an already efficient irrigation system (US and Europe) or which do not rely much on irrigation (e.g., India, Brazil and some African countries), while nations in desertic environments, which can't count on precipitation as main source of water supply, show very high rates of blue uWF. Over time, some nations significantly reduced their blue unit water footprint contribute (US and some European countries), while other nations kept on maintaining high irrigation volumes per unit of product.

Tuninetti et al. (2015) demonstrated that, for all crops, the spatial heterogeneity of crop water footprint is mainly driven by yield patterns, suggesting that agricultural practices are more effective on water footprint than climatic conditions. With this work, we prove that the dependence of WF from crop yield is evident also on a temporal scale. In fact, between 1970 and 2019 the national yield significantly raised in the countries into analysis, with different growth rates according to the nation; meanwhile, the water footprint of these nations reduced proportionally to the trend of yield. More specifically, the yield of Viet Nam is one of those

with the greatest growth rates, therefore this explains why its unit water footprint has drastically reduced over the years. On the contrary, crop yield in Nigeria followed a minor growth with respect to the other countries, thus water footprint reduction was less remarked.

In addition, we cross-compared our national uWF data with those obtained with the Fast Track approach (Tuninetti et al., 2017), with uWF data averaged over time-interval ranging from 1996 to 2005. We noticed that data show a good alignment, with only 17 countries over 149 having a relative error greater than 1. In addition, the comparison between the time series of uWF calculated with our approach and with Fast Track of some nations results very similar in the pattern; in some countries, such as Italy, the two curves are almost overlapping throughout the time interval, while in other countries they just differ by a quasi-constant coefficient. This fact proves the potential differences from an accurate estimation of the uWF (this thesis) and a valid but simplified method of calculation (Fast Track).

By looking at the historical evolution of green and blue Water Footprint, it is clearly observed that the trend is generally increasing, both at global and country level. In fact, even though the efficiency in water use improved over the time-interval, the huge increase of maize harvested area and production throughout the 1970 – 2019 time-interval determined an increment of water volumes. Among the analysed countries, United States is the country with the greatest yearly growth of green WF (more than 900 million cubic meters), followed by Nigeria (400 million m³), while the trend of Australia and Italy is mildly decreasing, mainly due to the decrease of the harvested areas. Regarding blue WF, the trend is rising in all the five countries, meaning that the irrigation requirements are getting more and more important due the effect of climate change.

This thesis contributes to extend the knowledge on the spatio-temporal variability of green and blue uWF of crops. However, for the seek of simplicity, the analysis has been developed only for a single crop, maize, one of the most cultivated worldwide, thus one of the most impacting in terms of water consumption. To enrich the survey on this field, the analysis can be further extended on other crops, starting from wheat, rice and soybean, which, together with maize, represent the four most cultivated grains, and later to other irrigation-demanding products. In addition, it is worth to deepen the study about the relation between the blue fraction of uWF and climatology, thus the possibility of estimating the ratio ET_b/ET_a on the base of climate variables, that are precipitation and reference evapotranspiration. This would allow to achieve an alternative and less computational demanding method which does not require the use of complex hydrological soil water balance models.

7 Annexes

7.1 Display of input variables of uWF in Pino Torinese

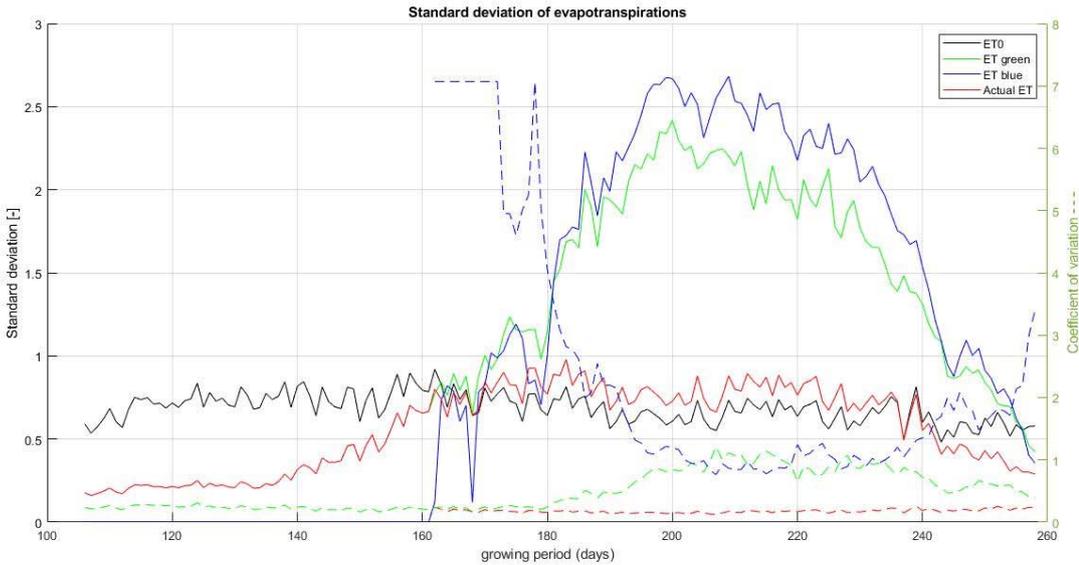


Figure 7. 1: Standard deviation and coefficient of variation of ETs

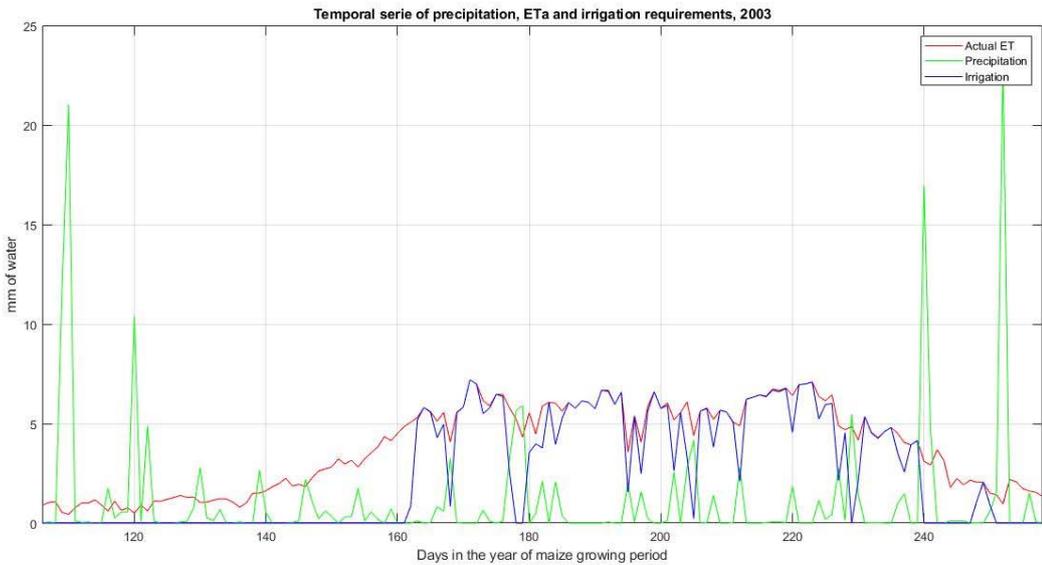


Figure 7. 2: temporal series of P, ETa and irrigation requirements, 2003

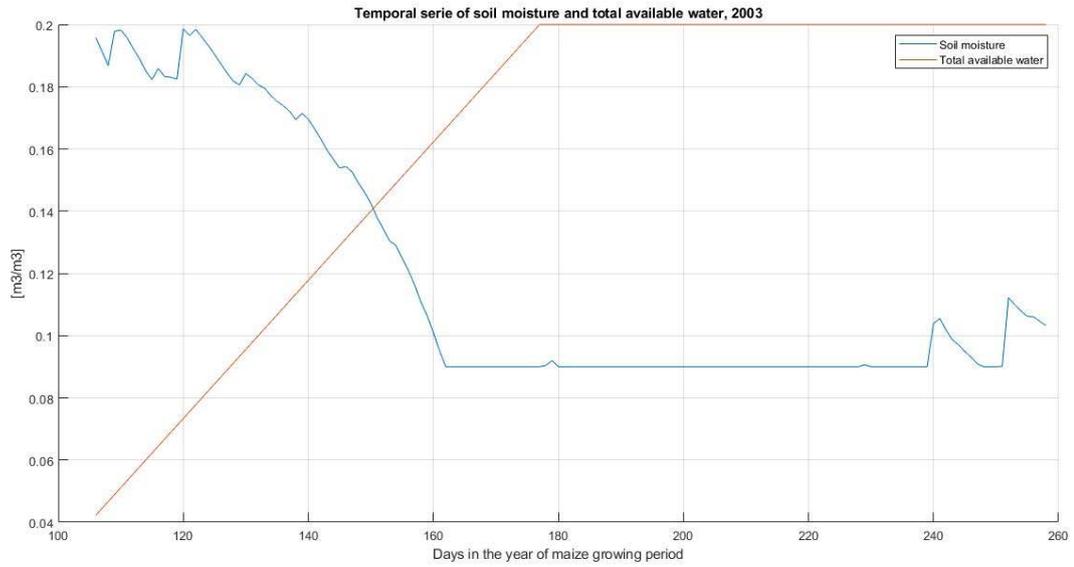


Figure 7. 3: temporal series of soil moisture and TAW, 2003

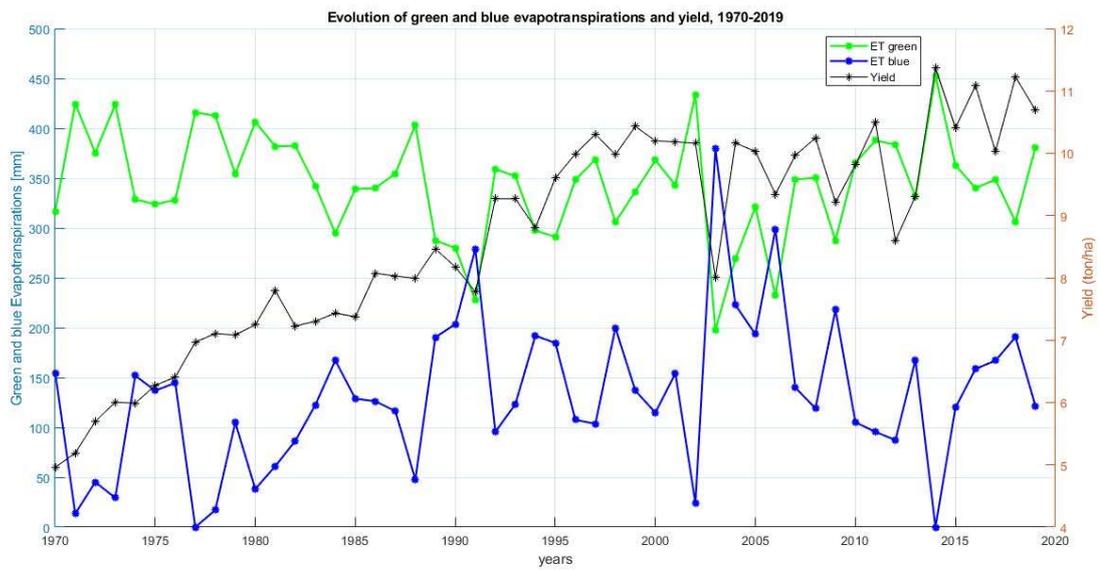


Figure 7. 4: evolution of green and blue evapotranspiration and yield, 1970-2019

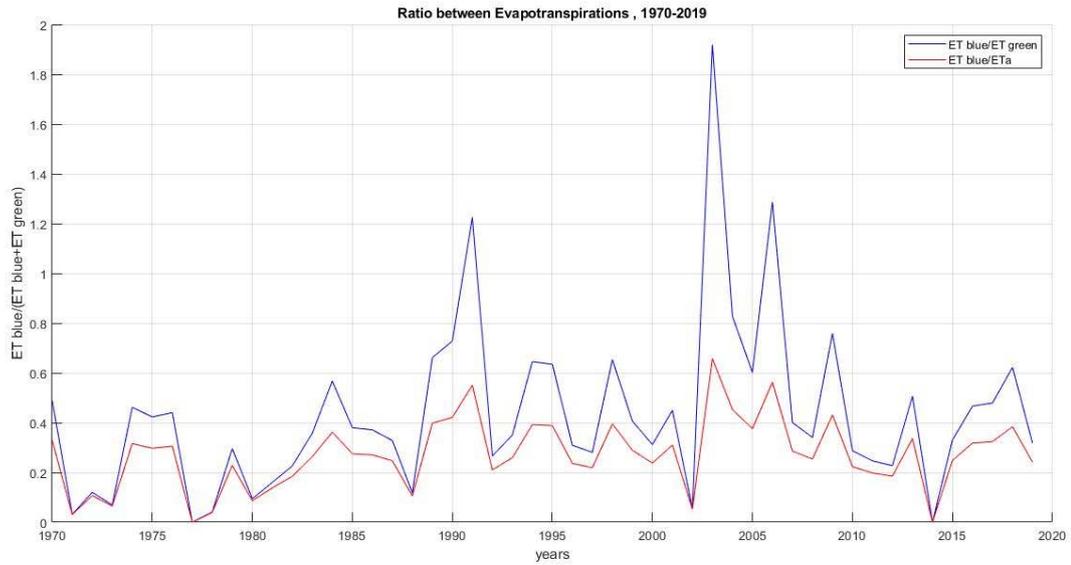


Figure 7. 5: ratio between evapotranspirations, 1970-2019

7.2 Display time series of harvested area, AEI, production and ET of the five countries

7.2.1 Italy

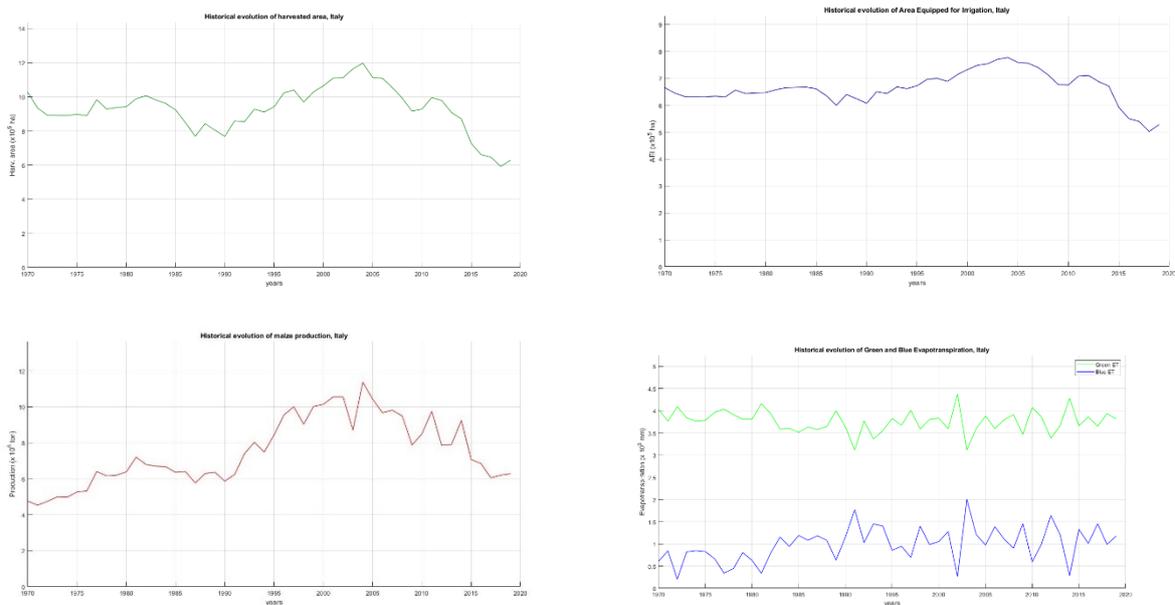


Figure 7. 6: plots of uWF input variables, Italy

7.2.2 US

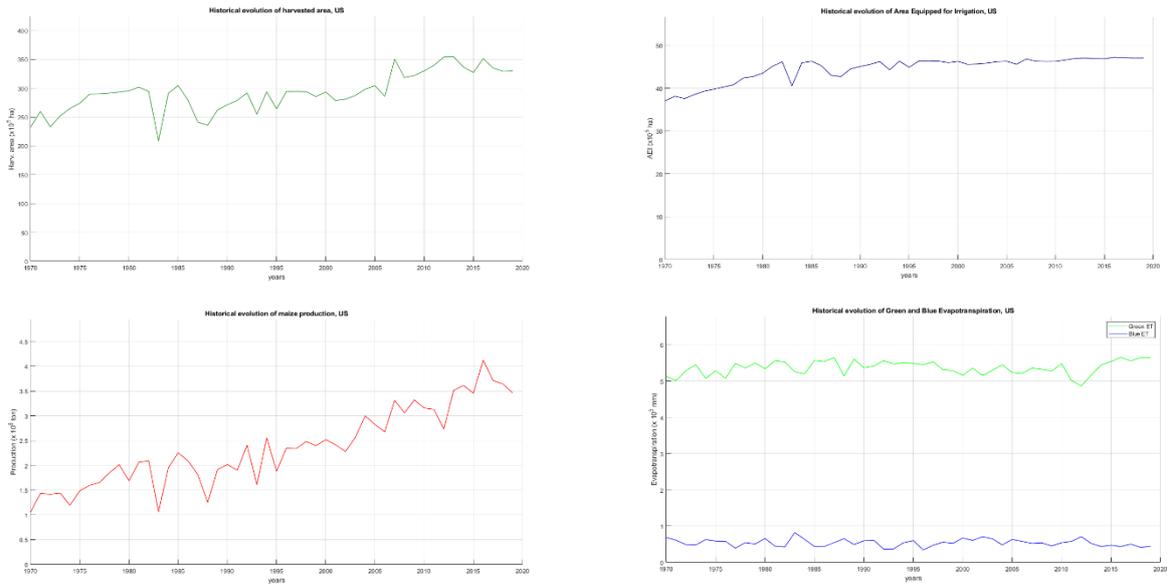


Figure 7. 7: plots of uWF input variables, US

7.2.3 Australia

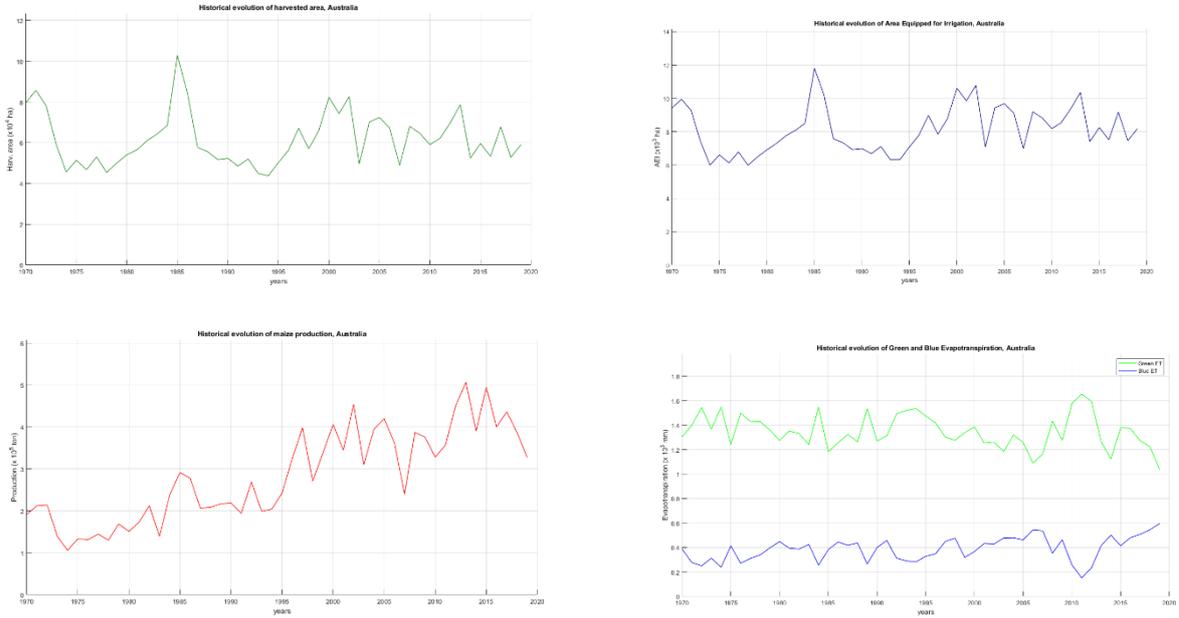


Figure 7. 8: plots of uWF input variables, Australia

7.2.4 Nigeria



Figure 7. 9: plots of uWF input variables, Nigeria

7.2.5 Viet Nam

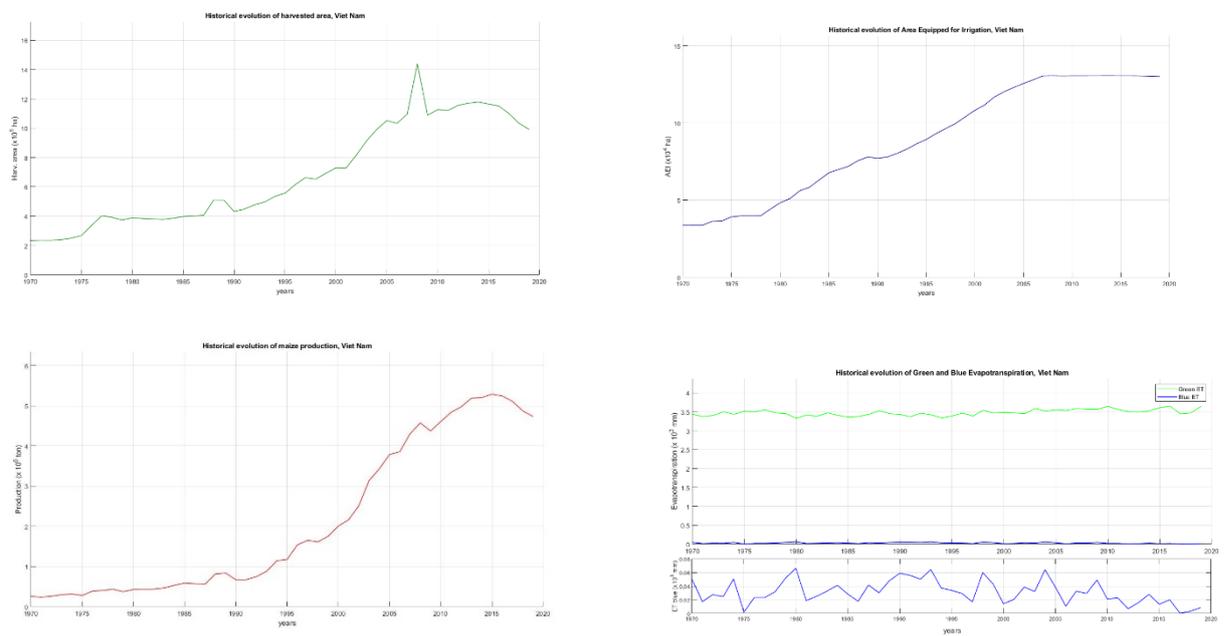


Figure 7. 10: plots of uWF input variables, Viet Nam

7.3 Global map of area equipped for irrigation, 2000

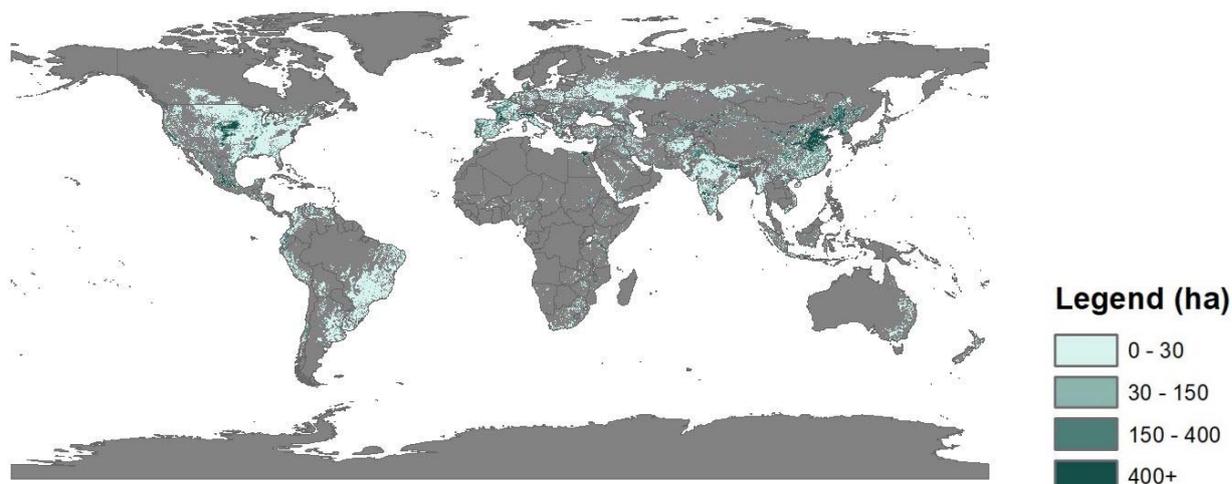


Figure 7. 11: global Area Equipped for Irrigation (AEI), 2000

7.4 National green uWF table

Country name	FAO code	1970	1975	1980	1985	1990	1995	2000	2005	2010	2015	2019
Afghanistan	2	65.7	60.1	59.5	60.6	77.3	99.4	224.3	120.1	168.7	125.6	131.4
Albania	3	909.9	790.8	716.3	601.9	588.3	865.3	593.1	681.9	425.9	369.6	378.8
Angola	7	3384.9	4137.0	4630.2	7618.1	11864.5	8963.0	5524.5	4721.9	4485.3	2708.9	3001.6
Antigua and Barbuda	8	1486.6	1967.5	2084.6	1886.7	1523.3	2233.0	2194.0	2475.7	2474.9	2033.4	2331.1
Argentina	9	2244.4	2131.2	1977.6	1410.6	1424.5	1111.8	953.2	698.1	665.7	689.8	635.5
Armenia	1	0.0	0.0	0.0	0.0	0.0	2310.6	1827.0	940.6	934.0	634.4	1036.7
Australia	10	541.7	478.4	456.3	417.2	303.7	305.4	280.6	216.9	283.7	166.2	186.6
Austria	11	792.1	562.6	551.9	475.3	481.9	466.9	400.0	370.7	407.6	463.2	404.4
Azerbaijan	52	0.0	0.0	0.0	0.0	0.0	1516.4	550.6	366.9	414.1	252.5	194.2
Bahamas	12	2561.9	2509.3	2235.6	1878.4	1695.9	1336.7	996.1	718.0	329.0	299.3	264.8
Bangladesh	16	3202.7	3278.5	4417.7	3583.9	3005.5	2828.0	1436.0	594.7	522.4	422.8	389.8
Barbados	14	1585.0	1477.9	1488.1	1462.8	1405.6	1322.0	1423.7	1462.1	1407.6	1256.1	1367.2
Belarus	57	0.0	0.0	0.0	0.0	0.0	1391.1	1603.0	882.6	774.3	764.4	641.4
Belgium	255	0.0	0.0	0.0	0.0	0.0	0.0	311.6	317.8	297.8	365.8	389.2
Belize	23	1595.2	2405.7	1368.0	1884.6	1320.9	1474.0	1166.6	1022.0	943.0	900.0	1048.9
Benin	53	4858.0	4186.4	4013.5	3532.5	3489.5	2616.1	2672.1	2729.2	2759.3	2417.9	2048.7
Bhutan	18	2168.7	2108.5	2209.7	2219.2	3102.6	2151.5	1873.4	1025.8	1238.1	812.3	862.0
Bolivia (Plurinational State of)	19	3338.8	3315.6	3141.4	2676.3	2688.7	2271.1	2000.0	2049.7	1568.0	1835.5	2016.2
Bosnia and Herzegovina	80	0.0	0.0	0.0	0.0	0.0	1178.8	2099.8	915.0	1045.1	1191.1	772.5
Botswana	20	38569.8	4761.7	12578.6	14721.1	9107.5	3149.5	30437.8	12443.7	18925.5	8723.9	12418.3
Brazil	21	2921.1	2752.8	2268.1	2166.1	2238.2	1544.7	1496.9	1350.0	952.2	750.5	746.5
Bulgaria	27	1171.2	1026.7	1109.0	1407.9	1503.2	1172.5	2223.9	823.3	717.1	838.2	640.5

Country name	FAO code	1970	1975	1980	1985	1990	1995	2000	2005	2010	2015	2019
Burkina Faso	233	5551.5	6188.2	4023.2	3686.4	2568.5	3188.6	2112.2	2067.0	2475.1	2045.1	2152.0
Burundi	29	2377.7	2460.5	2618.6	2259.4	2106.6	2237.4	2853.9	2741.6	2711.8	2154.4	2885.4
Cabo Verde	35	7727.0	5778.7	5077.2	6145.4	7344.5	9779.9	3466.6	9833.0	14150.2	10289.2	93692.5
Cambodia	115	2362.8	3084.5	3624.6	4029.0	1993.5	3133.0	1306.9	1058.3	1061.8	805.2	729.1
Cameroon	32	2707.6	1909.3	2935.7	1485.6	1412.0	1616.0	1019.1	1185.9	1250.0	1442.7	1532.3
Canada	33	776.7	708.9	688.6	630.5	574.5	563.3	589.5	477.0	408.6	385.2	433.8
Central African Republic	37	2287.1	3680.3	3998.4	1572.5	2155.9	1820.5	1505.6	1754.2	970.6	1991.8	2550.2
Chad	39	2097.4	2275.6	4385.3	4330.3	3907.2	4148.3	4901.2	3358.8	4412.9	3158.2	2840.5
Chile	40	230.3	242.4	315.3	106.7	87.6	70.9	100.9	78.1	65.9	63.0	61.5
China	351	1706.3	1338.3	1110.9	961.4	773.7	699.2	719.5	654.6	620.2	549.8	515.7
China, Taiwan Province of	214	883.7	774.3	786.2	501.2	477.6	427.0	382.4	360.6	369.0	390.0	400.6
Colombia	44	1892.0	2010.6	1940.8	1821.5	1783.6	1668.0	1221.1	954.3	947.8	729.8	717.9
Comoros	45	4102.6	4163.0	4387.2	2978.1	1907.3	1858.0	2026.7	2071.1	2166.5	2029.2	2200.4
Congo	46	1607.4	1496.5	1144.5	1218.7	1590.6	1205.1	1137.8	1176.0	1140.1	1088.2	926.4
Costa Rica	48	2478.1	2146.7	1495.8	1889.7	1809.9	1626.6	1634.8	1414.9	1475.8	1670.3	1722.9
Côte d'Ivoire	107	3766.0	5100.3	3215.8	2995.4	2276.4	1789.5	1365.9	1277.0	1387.7	1331.1	1338.3
Croatia	98	0.0	0.0	0.0	0.0	0.0	935.4	1079.7	659.9	657.2	710.3	530.8
Cuba	49	1655.7	1463.1	1490.8	1547.1	2243.3	1656.0	705.1	586.0	751.4	481.0	616.4
Czechia	167	0.0	0.0	0.0	0.0	0.0	897.8	573.1	519.5	553.8	721.8	491.0
Democratic Peoples Republic of Korea	116	755.4	822.9	661.4	570.0	473.1	1348.7	1604.2	1076.0	956.6	807.1	676.6
Democratic Republic of the Congo	250	1509.7	1463.1	1372.7	1374.8	1372.9	1417.3	1370.9	1478.8	1454.2	1460.9	1434.4
Denmark	54	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	638.2	510.3	421.0
Djibouti	72	0.0	0.0	0.0	874.6	594.8	958.1	614.2	929.0	685.9	346.6	545.2
Dominican Republic	56	2446.5	2174.4	3323.2	2633.6	2512.5	3042.7	3298.3	2879.1	2480.8	2817.8	2459.5
Ecuador	58	2076.1	1916.9	2023.5	1342.0	1779.3	1637.8	1363.8	983.3	1019.8	534.6	530.5
Egypt	59	77.9	75.7	60.1	55.2	44.7	46.9	37.5	34.7	37.3	37.3	36.7
El Salvador	60	2596.5	2584.3	2596.9	2408.1	2245.7	2002.9	2117.8	1497.2	1399.7	1984.7	1800.6
Eritrea	178	0.0	0.0	0.0	0.0	0.0	11684.3	15584.4	8148.8	4233.9	3616.6	3949.5
Eswatini	209	5881.7	2274.2	2629.8	1703.4	2711.5	1588.3	1816.5	2647.9	2893.6	2897.4	2923.5
Ethiopia	238	0.0	0.0	0.0	0.0	0.0	2102.1	1741.7	1442.8	1298.8	1095.3	672.5
France	68	855.7	1014.9	791.9	635.3	638.9	484.5	452.8	431.9	430.0	447.1	449.4
Gabon	74	730.3	745.7	707.2	741.2	679.2	690.3	652.5	705.4	686.2	676.8	683.5
Gambia	75	3803.7	3278.2	3805.3	2541.4	3350.8	3066.1	2536.9	3836.3	3195.4	5075.3	8190.9
Georgia	73	0.0	0.0	0.0	0.0	0.0	1436.1	2612.3	1875.7	3139.1	2499.2	1489.4
Germany	79	684.8	629.6	581.7	521.2	514.4	473.5	365.3	388.1	385.7	419.5	406.8
Ghana	81	2582.0	2625.8	3132.6	2838.9	2401.0	1908.6	1967.2	1816.6	1525.1	1496.6	1460.3
Greece	84	706.7	711.5	306.0	217.0	215.2	243.3	201.8	237.8	242.8	226.2	236.1
Grenada	86	3535.2	3493.8	3286.8	3191.0	2904.1	2982.3	3075.2	3239.2	3241.1	2783.7	3016.3
Guadeloupe	87	2143.9	3718.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Guatemala	89	394.6	274.6	343.3	290.5	236.1	248.3	280.9	217.8	254.3	234.8	238.9
Guinea	90	2496.6	2502.2	2925.9	2690.5	3005.4	2494.6	2123.3	1955.8	1790.9	2370.0	2046.0
Guinea-Bissau	175	5305.8	5355.6	3580.9	5434.6	3478.1	3628.9	3684.5	3395.8	3396.2	4157.5	3557.7
Guyana	91	2507.2	1861.5	2539.5	3639.6	2378.3	2799.5	2949.5	3060.8	2674.1	2654.2	2688.0
Haiti	93	3636.4	4295.0	4741.4	5302.3	5097.1	5109.8	4651.9	5107.1	5134.0	5183.4	5646.1
Honduras	95	2624.2	2814.8	2730.8	2290.5	2089.8	2012.3	2292.9	2052.2	1972.1	1881.6	1913.9

Country name	FAO code	1970	1975	1980	1985	1990	1995	2000	2005	2010	2015	2019
Hungary	97	1235.5	839.9	745.6	660.2	911.0	933.9	836.4	557.1	631.3	703.4	537.8
India	100	3059.3	3160.4	3370.4	3463.9	2495.1	2466.6	2164.8	2075.5	1540.9	1514.1	1222.3
Indonesia	101	3880.4	3212.2	2466.8	2050.9	1714.1	1588.0	1329.4	1050.9	857.1	642.0	618.5
Iran (Islamic Republic of)	102	508.5	348.7	505.7	214.6	158.9	153.9	95.5	93.4	91.7	91.0	116.7
Iraq	103	315.4	179.3	156.3	193.0	128.3	274.0	282.5	114.1	115.8	100.4	79.4
Israel	105	111.2	112.9	106.7	34.9	30.1	52.9	32.3	41.1	18.1	26.2	25.6
Italy	106	868.9	645.4	562.1	509.9	476.2	426.3	402.1	413.9	444.5	376.1	381.5
Jamaica	109	3127.2	2308.0	2807.6	3369.1	4057.1	3213.3	3742.2	3122.6	3348.1	3799.4	3879.3
Japan	110	1251.0	1298.4	1620.4	1394.7	1378.5	1317.4	1442.2	1409.4	1416.8	1348.1	1314.3
Jordan	112	5087.9	1598.5	212.5	1503.2	179.7	127.6	99.7	72.4	55.2	49.7	58.0
Kazakhstan	108	0.0	0.0	0.0	0.0	0.0	813.9	453.1	299.1	247.4	211.2	197.4
Kenya	114	2660.0	2248.0	2721.4	1843.2	1965.1	1744.4	2057.7	1979.6	1895.5	1828.5	1854.4
Kuwait	118	0.0	61.9	6.9	0.2	0.9	0.3	0.2	0.2	0.1	0.8	0.4
Kyrgyzstan	113	0.0	0.0	0.0	0.0	0.0	963.1	514.9	450.3	471.0	427.5	380.5
Lao Peoples Democratic Republic	120	1921.6	2092.9	3175.0	2561.3	1809.7	1810.8	1341.0	752.9	704.7	550.9	701.5
Lebanon	121	20.6	27.0	15.6	26.3	8.8	8.2	13.6	11.5	8.3	19.6	10.2
Lesotho	122	7729.6	5363.2	4152.3	4893.9	3453.0	4910.1	5361.4	5472.8	4338.8	6233.6	5696.4
Lithuania	126	0.0	0.0	0.0	0.0	0.0	2748.7	1735.1	1074.0	484.3	695.3	446.7
Luxembourg	256	0.0	0.0	0.0	0.0	0.0	0.0	457.8	417.2	471.2	626.4	886.2
Madagascar	129	3034.0	2796.1	3113.8	3030.9	3253.3	3146.4	3451.1	2068.2	1953.8	1739.6	1803.9
Malawi	130	3621.9	3159.2	2665.8	2563.1	3261.2	2373.9	1779.6	4106.2	1535.3	1857.6	1720.3
Malaysia	131	1488.7	2494.2	2800.1	1973.0	1909.9	1647.0	1338.0	1104.9	552.3	510.3	392.1
Mali	133	3093.0	3834.2	3049.9	2631.0	3005.2	2668.5	2600.2	2322.5	1237.7	1356.2	1278.7
Mauritania	136	5559.9	12234.8	5218.9	11077.2	4118.9	4570.2	3718.6	3330.0	4405.0	6845.1	3908.7
Mauritius	137	1855.8	1690.5	1704.4	794.1	882.3	952.2	406.8	508.5	577.7	625.1	523.8
Mexico	138	2726.3	2627.7	1847.7	1817.5	1729.4	1451.1	1387.5	1176.5	1013.0	1047.0	898.4
Morocco	143	3678.0	3452.0	2965.7	2983.1	2309.1	17817.8	5717.2	7701.5	1583.9	2493.3	2262.9
Mozambique	144	2799.8	7813.2	6056.2	7334.7	7491.9	4679.7	3402.8	6505.6	2687.3	3913.3	4167.2
Myanmar	28	3462.8	3193.9	2254.0	1401.5	1669.5	1528.1	1460.8	896.6	726.5	671.2	673.6
Namibia	147	2837.6	2695.3	2853.7	4602.7	2552.7	2927.3	5300.2	1640.2	1941.6	1320.7	1918.1
Nepal	149	1642.8	1820.9	1933.1	2156.4	1879.7	1896.9	1799.8	1562.0	1431.1	1310.0	1111.3
Netherlands	150	781.0	556.8	787.2	281.4	305.2	382.2	291.0	390.7	266.9	319.6	399.5
New Caledonia	153	1620.3	1306.2	1378.0	999.7	1582.8	949.3	956.0	1038.3	1011.8	486.3	464.1
New Zealand	156	217.5	183.1	170.6	161.3	176.7	163.3	156.2	157.1	167.7	155.4	162.0
Nicaragua	157	4242.3	4234.7	3429.7	3586.2	3112.9	3353.1	3160.3	2833.4	2941.4	3685.7	3150.0
Niger	158	5403.5	9563.6	5495.5	10906.8	9347.6	2445.1	4999.8	4421.8	5873.7	3112.4	2476.9
Nigeria	159	3725.7	2680.0	2840.5	3267.1	3333.6	2916.4	2880.4	2322.8	2058.5	2427.4	2370.5
North Macedonia	154	0.0	0.0	0.0	0.0	0.0	1228.9	1150.3	985.2	932.9	910.2	920.2
Pakistan	165	2910.5	3100.9	2760.2	2769.5	2418.0	2215.1	1897.5	1175.3	1020.0	852.2	707.5
Panama	166	4209.4	4163.3	4016.1	3259.7	2988.9	2540.7	2511.4	2564.9	2635.5	2479.2	1763.5
Papua New Guinea	168	4628.0	3650.6	1663.0	1824.1	1552.4	1128.6	740.7	826.3	816.2	728.3	654.0
Paraguay	169	3572.2	3583.7	3069.9	3007.1	2236.6	1887.6	2550.9	2360.1	1251.9	897.0	959.0
Peru	170	1221.2	1161.9	1264.8	1049.8	1035.1	1052.3	850.4	796.9	669.1	612.8	607.0
Philippines	171	4322.6	4212.1	3578.9	3089.1	2592.9	2307.7	1927.3	1620.3	1396.0	1162.8	1088.4
Poland	173	1433.2	660.1	937.1	827.8	715.3	718.7	597.1	632.1	628.0	749.6	663.7

Country name	FAO code	1970	1975	1980	1985	1990	1995	2000	2005	2010	2015	2019
Portugal	174	986.9	990.4	803.9	682.7	404.3	366.0	255.0	265.7	229.9	144.4	152.1
Puerto Rico	177	955.0	1562.1	3304.0	2873.4	1497.4	1065.1	1251.2	1565.0	1483.6	822.7	835.0
Qatar	179	0.0	0.0	0.0	0.0	13.5	16.5	11.2	9.5	3.7	20.3	16.3
Republic of Korea	117	2573.7	2095.7	762.9	703.9	777.8	838.1	886.5	753.9	759.7	736.2	726.0
Republic of Moldova	146	0.0	0.0	0.0	0.0	0.0	1513.2	1808.7	1446.7	1387.5	1603.6	1008.8
Réunion	182	1039.2	789.1	795.6	754.1	769.4	432.5	456.8	431.9	0.0	0.0	0.0
Romania	183	2002.0	1553.5	1285.4	1103.3	1528.7	1356.7	2298.8	1084.7	1001.3	1246.6	676.1
Russian Federation	185	0.0	0.0	0.0	0.0	0.0	1195.7	1639.4	848.4	989.3	662.6	563.9
Rwanda	184	1919.5	2184.9	2007.9	1859.4	2339.8	2155.1	3537.5	2800.4	1017.3	1561.6	1670.6
Sao Tome and Principe	193	44.5	44.1	43.9	46.2	35.1	30.3	30.1	29.3	41.4	28.9	34.1
Saudi Arabia	194	824.3	798.9	400.1	324.5	87.0	234.5	74.1	71.4	62.3	64.7	74.0
Senegal	195	4390.3	4086.3	4735.0	2510.6	2947.2	3270.9	3190.3	1260.4	2358.3	2288.5	1555.0
Serbia and Montenegro	186	0.0	0.0	0.0	0.0	0.0	1171.4	1807.0	860.0	0.0	0.0	0.0
Sierra Leone	197	2183.4	2244.9	2403.7	2141.6	2135.8	2330.6	2481.3	2354.3	1564.9	3671.5	671.2
Slovakia	199	0.0	0.0	0.0	0.0	0.0	903.9	1385.6	683.2	843.7	919.3	673.7
Slovenia	198	0.0	0.0	0.0	0.0	0.0	631.4	721.0	477.1	472.0	472.8	447.4
Somalia	201	2947.6	3231.9	2987.9	2485.7	2159.5	3666.1	1880.6	2386.1	1960.9	4584.4	5173.7
South Africa	202	2389.9	1641.6	1341.6	1754.3	1443.5	2179.0	1100.4	927.8	705.9	835.3	615.7
Spain	203	551.3	664.3	458.3	332.5	323.5	281.7	229.3	166.7	214.1	173.9	164.7
Sri Lanka	38	4896.3	5480.3	3092.1	4035.1	3247.7	3670.6	3365.5	2327.0	1362.6	1077.7	924.3
Sudan (former)	206	2577.8	3743.9	3854.2	6371.1	2936.1	5617.3	1702.5	1843.4	1536.8	0.0	0.0
Suriname	207	3379.7	2690.7	3329.2	3163.2	2423.5	2331.1	2693.4	2908.6	2207.7	2150.5	1947.4
Switzerland	211	580.6	515.9	502.8	495.1	422.8	411.7	354.8	356.1	399.8	409.6	343.9
Syrian Arab Republic	212	939.2	322.2	195.6	124.3	52.0	78.4	71.2	74.3	76.7	97.4	67.9
Tajikistan	208	0.0	0.0	0.0	0.0	0.0	578.5	311.8	249.2	112.1	76.8	72.7
Thailand	216	1374.1	1474.9	1603.2	1378.6	1536.8	1078.0	935.8	935.4	865.4	871.5	798.0
Timor-Leste	176	2645.2	3818.8	2579.8	2678.3	1892.2	1424.8	1903.4	1709.0	2108.5	966.3	855.6
Togo	217	2436.9	2447.8	3192.9	3449.4	3234.1	3595.1	2568.1	2690.0	2553.7	2234.6	2510.7
Trinidad and Tobago	220	590.5	671.0	922.9	908.6	602.4	879.9	1011.4	1133.2	1066.8	1190.6	1616.6
Turkey	223	1550.7	1355.6	1135.4	717.5	612.8	770.8	545.9	357.1	358.7	300.8	299.9
Turkmenistan	213	0.0	0.0	0.0	0.0	0.0	158.3	433.3	501.2	270.4	332.5	436.1
Uganda	226	2486.5	2560.5	2904.3	2527.9	2211.7	2056.9	1916.3	2066.2	1433.5	1340.3	1188.1
Ukraine	230	0.0	0.0	0.0	0.0	0.0	1421.6	1375.3	990.2	966.7	720.8	584.5
United Republic of Tanzania	215	4294.2	1697.7	1757.7	1604.9	1410.1	1031.9	958.1	2102.3	1349.4	1386.7	1341.9
United States of America	231	1129.9	974.9	934.3	751.8	721.3	770.8	600.5	563.8	572.3	523.9	536.5
Uruguay	234	5201.9	4217.0	4416.7	3406.3	2125.2	1603.7	2407.8	919.0	739.3	707.5	526.0
Uzbekistan	235	0.0	0.0	0.0	0.0	0.0	229.2	273.2	185.3	142.3	57.7	96.1
Vanuatu	155	7216.5	7712.0	7476.1	7367.4	7430.7	7860.1	7466.7	7380.8	6719.9	6385.2	6112.6
Venezuela (Bolivarian Republic of)	236	3343.4	3143.7	3010.2	2341.4	2002.7	1666.6	1292.7	1489.8	1288.5	1554.9	1278.3
Viet Nam	237	3120.4	3348.4	3025.5	2276.5	2206.1	1603.1	1268.4	988.0	891.3	795.4	760.5
Yemen	249	1057.4	1049.4	841.9	1421.9	786.4	1272.1	1241.5	2170.7	1095.2	1177.0	1299.8
Zambia	251	5441.3	2332.4	2086.9	1729.7	2469.7	2501.3	1887.3	1914.8	1280.5	1129.3	1473.5
Zimbabwe	181	2476.9	1832.7	2535.3	1481.2	1915.6	5433.8	2626.3	4986.1	3713.0	5312.5	3893.7

7.5 National blue uWF table

Country name	FAO code	1970	1975	1980	1985	1990	1995	2000	2005	2010	2015	2019
Afghanistan	2	505.3	434.9	442.4	504.2	591.5	575.4	1461.8	913.8	699.4	622.1	951.7
Albania	3	412.1	384.0	543.4	637.2	726.2	440.0	673.0	378.9	263.8	341.8	314.0
Argentina	9	8.7	10.4	26.4	9.1	15.5	13.2	5.5	6.3	5.2	5.6	3.4
Armenia	1	0.0	0.0	0.0	0.0	0.0	9.4	19.8	3.1	5.6	3.6	14.6
Australia	10	160.4	159.1	160.5	134.7	95.4	67.8	74.2	79.6	46.2	49.8	107.3
Austria	11	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Azerbaijan	52	0.0	0.0	0.0	0.0	0.0	2633.1	1002.8	625.5	693.6	554.2	388.7
Belarus	57	0.0	0.0	0.0	0.0	0.0	370.2	126.8	123.5	33.1	78.6	9.2
Belgium	255	0.0	0.0	0.0	0.0	0.0	0.0	0.8	0.5	0.8	0.7	3.1
Belize	23	2.1	29.0	0.0	0.6	0.1	0.1	1.7	5.9	0.0	0.9	14.1
Benin	53	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Bolivia (Plurinational State of)	19	0.0	2.2	0.9	0.6	4.4	1.0	0.2	0.8	1.0	0.3	0.9
Bosnia and Herzegovina	80	0.0	0.0	0.0	0.0	0.0	0.8	2.9	0.6	1.1	1.6	0.9
Brazil	21	0.9	1.0	1.0	1.0	2.4	1.1	1.3	2.0	1.6	1.4	1.5
Bulgaria	27	9.3	3.8	14.1	34.8	40.8	8.4	48.1	0.4	1.9	2.0	1.2
Burkina Faso	233	0.0	0.1	0.1	0.1	0.3	0.2	0.1	0.0	0.0	0.0	0.0
Burundi	29	0.5	0.4	0.5	0.4	0.5	0.6	1.3	0.8	1.0	0.7	0.4
Cambodia	115	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Cameroon	32	0.0	0.0	0.2	0.0	2.7	0.0	0.0	0.0	0.0	0.1	0.0
Canada	33	15.9	11.5	7.1	12.4	8.5	6.6	11.0	6.3	4.6	7.2	5.9
Chad	39	269.3	283.8	247.4	190.2	343.0	170.7	232.5	80.7	70.7	55.2	38.5
Chile	40	892.3	713.2	589.4	364.2	310.2	272.3	289.9	152.9	183.1	169.3	239.9
China	351	112.7	135.2	96.4	53.2	34.1	39.2	91.4	54.4	66.4	75.3	73.6
Colombia	44	0.0	0.1	0.5	0.1	0.0	0.2	0.6	0.2	0.0	0.7	0.3
Croatia	98	0.0	0.0	0.0	0.0	0.0	0.3	1.1	0.7	2.1	5.6	3.5
Cuba	49	7.2	15.0	5.1	15.6	49.3	8.0	68.2	2.2	0.4	120.5	16.2
Czechia	167	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Democratic Peoples Republic of Korea	116	0.5	0.0	6.1	1.9	0.0	0.0	17.6	0.2	0.4	9.1	55.9
Democratic Republic of the Congo	250	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Denmark	54	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	49.5	1.8	18.1
Djibouti	72	0.0	0.0	0.0	2787.3	3579.8	2865.3	3196.7	3437.6	2978.7	3047.8	2488.7
Dominican Republic	56	0.2	0.5	0.4	0.3	0.6	0.7	1.9	0.6	0.0	2.6	0.9
Ecuador	58	194.1	127.7	239.6	122.8	171.0	59.0	35.0	137.4	73.3	42.4	55.9
Egypt	59	1198.2	1162.9	949.9	844.4	676.0	747.5	573.1	549.9	611.9	605.9	596.4
El Salvador	60	0.8	1.7	0.1	1.2	0.7	0.0	0.8	0.0	0.0	2.6	1.0
Eritrea	178	0.0	0.0	0.0	0.0	0.0	8.4	21.9	6.3	3.6	6.6	2.9
Eswatini	209	7.7	0.0	1.5	0.1	0.4	3.5	0.0	1.4	1.9	5.0	1.8
Ethiopia	238	0.0	0.0	0.0	0.0	0.0	55.5	38.7	44.6	38.3	52.5	30.2
France	68	41.5	74.5	34.6	68.1	163.9	139.2	75.9	157.9	121.5	150.8	158.7
Gambia	75	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Georgia	73	0.0	0.0	0.0	0.0	0.0	12.7	50.6	13.9	98.7	76.7	57.6
Germany	79	4.3	15.5	2.0	3.2	6.4	5.4	0.4	1.4	2.2	2.5	6.1

Greece	84	498.4	407.3	226.7	226.8	195.3	185.4	242.6	176.0	183.2	211.1	268.1
Guadeloupe	87	0.0	140.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Guatemala	89	0.4	0.7	0.2	0.7	0.5	0.0	0.8	0.3	0.2	2.4	1.5
Haiti	93	47.5	46.2	29.6	53.1	68.9	61.6	100.1	37.3	22.6	85.8	36.7
Honduras	95	2.2	15.9	2.3	1.0	0.9	0.4	5.3	0.9	0.0	9.0	13.6
Hungary	97	0.2	0.3	0.4	0.8	1.8	1.0	2.0	0.0	0.2	1.3	0.9
India	100	25.8	7.0	78.6	80.8	26.9	29.4	60.8	19.8	9.4	38.9	2.0
Indonesia	101	107.4	5.8	104.1	69.5	71.1	46.6	37.1	36.1	0.1	64.2	39.5
Iran (Islamic Republic of)	102	3039.9	1964.9	2791.3	1288.0	999.7	536.0	485.9	329.8	406.0	483.8	444.4
Iraq	103	5551.4	2495.0	2061.5	2819.8	1809.3	3187.4	6211.4	1086.7	1532.5	1540.0	779.2
Israel	105	794.1	637.2	825.6	199.7	254.8	363.3	305.4	296.7	139.5	191.0	181.1
Italy	106	130.8	140.8	93.1	172.7	150.9	95.3	110.5	104.0	64.6	136.4	117.8
Japan	110	2.1	11.7	0.3	105.9	79.0	86.1	28.1	11.7	25.4	3.2	1.4
Jordan	112	5.2	0.8	1.3	3.7	0.3	0.1	0.1	0.0	0.1	0.1	0.1
Kazakhstan	108	0.0	0.0	0.0	0.0	0.0	2387.4	1387.0	1156.6	1022.2	855.1	761.5
Kenya	114	0.9	0.6	0.9	0.6	1.1	1.1	2.7	1.6	1.9	1.6	1.3
Kuwait	118	0.0	1775.3	197.4	6.5	29.0	21.9	7.1	6.0	4.7	15.3	6.8
Kyrgyzstan	113	0.0	0.0	0.0	0.0	0.0	417.5	234.7	255.8	219.0	228.6	247.5
Lao Peoples Democratic Republic	120	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Lebanon	121	330.1	313.0	236.5	342.3	168.5	112.7	166.3	184.2	103.7	158.7	176.9
Lesotho	122	0.8	0.0	0.7	0.0	0.0	3.5	0.0	0.0	0.8	2.8	6.7
Lithuania	126	0.0	0.0	0.0	0.0	0.0	0.9	0.0	0.4	0.0	0.1	0.0
Malawi	130	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Mali	133	0.1	0.2	0.5	0.1	0.5	0.1	0.5	0.5	0.0	0.0	0.1
Mauritania	136	225.1	264.6	195.9	772.3	472.2	268.2	90.9	123.5	48.5	198.2	104.6
Mauritius	137	11.0	54.4	31.2	41.1	73.1	59.2	46.1	16.5	23.3	5.2	61.5
Mexico	138	86.2	97.4	85.8	69.3	32.8	55.7	98.0	69.6	43.4	56.1	60.2
Morocco	143	1043.8	697.5	1125.5	1025.7	742.5	7547.1	4114.9	9899.0	1494.4	3491.0	5317.2
Mozambique	144	5.1	3.5	8.9	5.1	6.0	9.3	0.0	8.5	1.9	5.4	1.7
Myanmar	28	130.4	63.2	46.9	29.0	36.5	38.2	35.5	19.2	14.6	8.2	6.1
Namibia	147	40.1	29.8	27.3	71.4	26.2	72.5	33.4	26.4	19.5	30.7	50.2
Nepal	149	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0
Netherlands	150	42.2	116.7	35.5	2.4	84.9	114.5	2.0	8.3	25.3	7.3	57.9
New Zealand	156	0.6	0.2	0.0	3.4	5.1	2.5	1.2	3.0	6.2	11.7	11.1
Nicaragua	157	0.0	0.2	0.2	2.7	3.2	0.0	1.9	0.1	0.0	42.0	17.4
Niger	158	247.7	78.0	49.1	253.3	1034.7	660.0	443.9	1249.7	149.9	45.1	46.2
Nigeria	159	8.1	5.4	19.2	6.2	4.2	2.4	5.0	3.4	1.4	1.6	1.3
North Macedonia	154	0.0	0.0	0.0	0.0	0.0	49.7	126.0	147.5	193.5	232.9	213.6
Pakistan	165	724.8	150.8	517.7	589.2	404.6	241.1	471.1	219.0	93.9	106.7	86.9
Paraguay	169	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Peru	170	39.9	9.7	64.5	35.1	55.0	42.2	18.4	34.8	27.5	17.6	24.1
Poland	173	86.4	50.3	8.8	10.0	25.3	44.9	1.7	7.7	1.4	10.1	9.4
Portugal	174	1274.7	1730.6	1510.5	1323.6	1219.2	783.5	621.4	900.2	569.8	489.5	438.5
Puerto Rico	177	50.2	657.0	133.2	559.6	400.2	441.6	548.2	0.0	53.4	586.4	302.6
Qatar	179	0.0	0.0	0.0	0.0	498.5	510.4	339.6	534.8	324.7	536.3	542.2
Republic of Korea	117	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.0	0.0	1.7	1.7

Republic of Moldova	146	0.0	0.0	0.0	0.0	0.0	99.3	141.8	65.8	64.2	216.3	77.7
Romania	183	4.6	6.2	7.4	15.7	34.2	10.7	55.3	6.9	17.8	25.0	13.9
Russian Federation	185	0.0	0.0	0.0	0.0	0.0	197.5	172.6	141.4	347.0	72.4	57.1
Saudi Arabia	194	1452.9	2499.7	6027.4	3770.4	2551.0	1884.0	401.7	277.5	301.8	834.2	374.7
Senegal	195	36.8	19.7	21.1	6.2	12.1	11.0	24.0	6.2	7.1	7.4	5.2
Serbia and Montenegro	186	0.0	0.0	0.0	0.0	0.0	2.5	15.1	2.3	0.0	0.0	0.0
Slovakia	199	0.0	0.0	0.0	0.0	0.0	62.4	102.7	13.7	5.9	16.5	7.0
Slovenia	198	0.0	0.0	0.0	0.0	0.0	0.9	4.0	0.4	2.7	4.7	4.1
Somalia	201	1180.2	649.9	937.5	673.0	962.3	1140.6	816.6	904.6	766.6	2036.3	2280.2
South Africa	202	30.3	12.5	15.5	14.4	20.2	51.5	12.1	19.0	18.5	34.1	27.0
Spain	203	516.1	516.6	448.4	369.5	434.1	475.6	350.2	405.1	349.9	325.3	337.5
Sudan (former)	206	2630.0	1682.9	2632.6	2986.0	7009.5	2710.0	3972.4	2508.7	2099.2	0.0	0.0
Switzerland	211	0.0	0.0	0.0	0.0	0.1	0.1	0.0	0.0	0.0	1.9	0.3
Syrian Arab Republic	212	16536.8	5355.2	2987.8	2159.2	968.2	1091.3	1255.4	1247.6	1462.1	1615.2	1163.1
Tajikistan	208	0.0	0.0	0.0	0.0	0.0	2740.3	2018.8	1254.9	407.7	518.7	410.6
Thailand	216	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Timor-Leste	176	0.0	0.0	0.0	0.0	111.9	118.5	158.5	174.1	12.4	145.2	206.8
Turkey	223	152.9	161.5	185.6	168.8	163.7	146.5	179.9	102.3	107.3	68.7	70.1
Turkmenistan	213	0.0	0.0	0.0	0.0	0.0	1979.8	7258.1	7528.1	4317.6	5559.3	5955.2
Ukraine	230	0.0	0.0	0.0	0.0	0.0	92.9	62.4	49.1	31.5	18.7	11.2
United Republic of Tanzania	215	44.7	15.1	27.9	24.6	28.5	22.1	34.7	33.9	31.1	27.3	26.9
United States of America	231	151.1	108.2	115.4	59.2	80.0	84.2	78.2	68.0	56.5	44.5	41.9
Uruguay	234	5.7	5.5	8.4	15.3	8.5	26.6	99.6	17.5	0.1	14.7	5.5
Uzbekistan	235	0.0	0.0	0.0	0.0	0.0	1668.6	2255.4	1237.2	702.3	536.2	623.1
Venezuela (Bolivarian Republic of)	236	16.2	28.3	48.8	26.2	23.3	20.3	22.8	16.3	8.3	87.7	57.3
Viet Nam	237	45.5	1.9	60.1	19.3	37.8	16.0	5.2	10.7	5.2	3.0	1.8
Yemen	249	836.1	352.4	802.8	722.4	1010.7	868.1	924.3	2016.9	674.2	1936.5	1263.8
Zambia	251	0.2	0.0	0.0	0.0	0.1	0.4	0.0	1.8	0.0	0.1	0.6
Zimbabwe	181	10.5	0.8	5.2	1.3	5.1	22.7	0.0	23.4	10.3	42.0	21.4

8 References

- Allamano, P., Claps, P., D'Odorico, P., Laio, F., Ridolfi, L., & Tamea, S. (2013). globalizzazione del cibo e geografia dell'acqua. il caso italiano. In *L'acqua che mangiamo* (pp. 163-179). Torino: DIATI Politecnico di Torino.
- D'Odorico, P., Carr, J. A., Laio, F., Ridolfi, L., & Vandoni, S. (2014). Feeding humanity through global food trade. *Earth Future*, 458–469.
- Falkenmark, M., & Rockstrom, J. (2004). Balancing Water for Humans and Nature: The New Approach in Ecohydrology. *Earthscan*.
- FAO. (2011). *The State of the World's Land and Water Resources for Food and Agriculture: Managing Systems at Risk*. Rome.
- FAO. (2021). *FAOSTAT*. Retrieved from <https://www.fao.org/faostat/en/>
- FAO. (2021). *GAEZ v4 Data Portal*. Retrieved from <https://gaez.fao.org/>
- FGS. (2017). *Somalia Drought Impact & Needs Assessment*.
- Godfray, H. C., & al., e. (2010). Food security: The challenge of feeding 9 billion people. *Science*, 327(5967), 812–818.
- Hanasaki, N., Inuzuka, T., Kanae, S., & Oki, T. (2010). An estimation of global virtual water flow and sources of water withdrawal for major crops and livestock products using a global hydrological model. *J. Hydrol*, 232–244.
- Hoekstra, A. Y., Chapagain, A. K., Aldaya, M. M., & Mekonnen, M. M. (2011). *The Water Footprint Assessment Manual*. London: Earthscan.
- Kopp, C. M. (2021, 7 30). Retrieved from Investopedia: <https://www.investopedia.com/articles/markets-economy/090316/6-countries-produce-most-corn.asp>
- Mekonnen, M. M., & Hoekstra, A. Y. (2011). The green, blue and grey water footprint of crops and derived crop products. *Hydrol. Earth Syst. Sci*, 1577-1600.
- Monfreda, C., Ramankutty, N., & Foley, A. J. (2008). Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Glob. Biogeochem. Cycles* 22 GB1022. Retrieved from <http://dx.doi.org/10.1029/2007GB002947>
- Nazmi, M. S., Dardak, R. A., Rani, R. A., & Rabu, M. R. (2021). Benchmarking Indonesia for the Development of grain corn industry in Malaysia. *Food and Fertilizer Technology Center for the Asian and Pacific Region Platform (FFTC-AP Platform)*.
- Powell, J., & Reinhard, S. (2016). Measuring the effects of extreme weather events on yields. *Weather and Climate Extremes*, 67-79. Retrieved from <https://doi.org/10.1016/j.wace.2016.02.003>
- Ray, K. D., Gerber, S. J., MacDonald, K. G., & West, C. P. (2015). Climate variation explains a third of global crop yield variability. *Nat. Commun*.

- Tuninetti, M., S., T., F., L., & L., R. (2017). A Fast Track approach to deal with the temporal dimension of crop water footprint. *Environ. Res. Lett.* Retrieved from <https://doi.org/10.1088/1748-9326/aa6b09>
- Tuninetti, M., Tamea, S., D'Odorico, P., Laio, F., & Ridolfi, L. (2015). Global sensitivity of high-resolution estimates of crop. *Water Resources Research*(51), 8257–8272. doi:10.1002/2015WR017148
- United Nations. (2013). *World population prospects, Highlights and advance tables*. Department of Economic and Social Affairs, Population Division. New York: Working Pap. ESA/P/WP.228.
- Water Footprint Network. (2021). *Water Footprint Network*. Retrieved from <https://waterfootprint.org/en/>
- Wendland, W. M. (1984). A Climatic Review of Summer 1983 in the Upper Midwest. *Bull. Amer. Meteor. Soc.*, 1068–1072. Retrieved from [https://doi.org/10.1175/1520-0477\(1984\)065%3C1068:ACROSI%3E2.0.CO;2](https://doi.org/10.1175/1520-0477(1984)065%3C1068:ACROSI%3E2.0.CO;2)