POLITECNICO DI TORINO

MASTER OF SCIENCE IN CIVIL ENGINEERING

TEMPORAL VARIABILITY OF IRRIGATION REQUIREMENTS

Advisors

Prof. Tamea Stefania

Dott. Ing. Rolle Matteo

Student

Fabrizio Gomes Acquaviva

Acquaviva, Fabrizio Gomes

Temporal Variabilty of Irrigation Requirements
São Paulo, Brasil - 10/2021.

86 p.

Politecnico di Torino.
Department of ENVIRONMENT, LAND AND INFRASTRUCTURE ENGINEERING.

Professor Tamea Stefania
Rolle Matteo

Life is nothing without family and friends

Acknowledgements

I want to thank all those who made my eleven years in Italy the best possible.

Firstly, to my parents Anna Leticia and Marcelo, none of this would have been possible without you. Leave your hometown at 14 years of age was not easy but you made it possible. I would not have become half of the man I am today without you. I am eternally grateful.

To my sister Tiziana, thank you for the friendship and companion. You gave me strength in the moments that even I did not know I needed. You have always been more than a sister.

To my grandfather, Celso, thank for being an academic role model, and supporting throughout this journey and to my grandmother Maria, you never stop taking care of me even 10.000 km distant.

To the Figueiras, my second father, mother and sisters. You made comebacks and goodbyes a lot easier. I am so happy to be able to share this new accomplishment with you.

To all my basketball teammates time spent on the court helped me blow out the academic stress.

To the friends who were geographically distant but were always with me in these two years. To the friendships that survived and even grew in these years of distance.

And finally, to Professor Stefania Tamea and Matteo Rolle, for the guidance and support, it was a long path but the results were very gratifying.

Abstract

In our day to day lives, we may not notice but agriculture is all around us. Not only on our meals and beverages, but on our clothes, shoes and a lot of daily objects. The need to have the most efficient usage of the cultivated land has been growing over the years. We all know that one of, if not the most, important players in this equation is water. Its correct usage in lands is crucial in order to get the maximum result of what we are cultivating. This implies that land needs to get the right amount of water in the right amount of time, and to get the best results irrigation may be absolutely necessary.

The objective of this work (or thesis) is to assess the temporal variability of irrigation requirements between 1981 and 2019 in two locations near Torino (Italy) and São Paulo (Brazil) That assessment was achieved with the calculations of irrigation requirements based on FAO (Food and Ag…) paper published on 1998. A previously created model was used to estimate the soil water balance and the irrigation volumes, taking precipitation data as a main driver. Precipitation data were extracted, manipulated and verified, to use them as an input in the model. The model results were analyzed in terms of spatial and temporal variability, considering different crops cultivated at the two sites. Finally, the interannual variability was interpreted using the tools of statistical inference.

Results showed that for São Paulo only maize did not show any signal of temporal trend. As for Torino, none of the crops analyzed presented temporal trends. Nonetheless, among the three studied distribution (Normal, LogNormal and Gumbel) the only one that fitted all crops for both cities was LogNormal

Outline

List of Figures

List of Tables

1. Introduction

1.1. Context

Water necessary for irrigation may represent an important strategic role in agriculture as 20% of irrigated harvested lands providing 40% of global food production (1)

The need to study and optimize water resources is growing at a fast pace all around the world. Irrigation techniques are evolving on daily basis, but the most important factor is when and how much irrigation should be provided.

Thus, the goal is to analyze and investigate whether irrigation requirements have a temporal variability over the years.

This thesis revisits precipitation data from 1981 to 2019 in two specific locations: Torino (Italy) and São Paulo (Brazil). The precipitation data was then uses as input to a previously created model in order to get irrigation requirements along the years. The focus is on a set of crops cultivated at both sites, wheat, citrus and maize.

One of the main organization behind the field of irrigation is FAO (Food and Agriculture Organization of the United Nations).It was created in 1945 and it is an international organization which the main purpose is to achieve food security for all and make sure that people have regular access to enough high-quality food to lead active, healthy lives. (2). Irrigation requirements is a crucial subject matter in order to always have sustainable agriculture and to win war against

1.2. Objective and Structure

The aim of this work is to analyze the dynamics and relations between some of the most important metrics involved in irrigation.

Prior to the analysis of the important metrics, key work was done.

Starting by the extraction, manipulation and analysis of precipitation data coming from the Copernicus Satellite.

Subsequently, based on Mirca dataset, a brief study on the most common crops cultivated in the two selected areas.

Using the precipitation data as input and the crops as filter to the results, the model provided two metrics: evapotranspiration and irrigation requirements.

At this point all needed information was collected, thus the temporal variability study proceeded with time series studies on evapotranspiration and irrigation requirements, correlation study and statistical inference

Three main metrics were studied: precipitation, evapotranspiration and irrigation requirements. Therefore, the structure of the thesis is:

- Literature Review
- Precipitation Data (extraction, manipulation, verification, comparison and analysis)
- São Paulo analysis on evapotranspiration and irrigation requirements
- Torino analysis on evapotranspiration and irrigation requirements
- Comparison and Conclusion

2. Literature Review

2.1. Introduction

This chapter introduces the concepts behind the quantification of irrigation requirements.

In 1998, FAO published a paper of extreme importance that is mainly used as primary literature when confronting irrigation. The paper presented an updated procedure for calculating reference and crop evapotranspiration from meteorological data and crop coefficients. (3).

2.2. Evapotranspiration

In order to introduce the parameters used to calculate the necessity of artificial irrigation, evapotranspiration needs to be studied.

It is the main factor that regulates how much water needs to inputed into the ground. As the name already gives up it is formed by two natural processes: Evaporation and Transpiration.

Evaporation is the natural process where water changes its status from liquid to vapor.

The difference between the water vapour pressure at the evaporating surface and that of the surrounding atmosphere is the force that removes water vapour from the evaporating surface. (3)

Some other factors play an important role when considering evaporation, such as solar radiation, wind speed, air humidity etc.

Evaporation tends to reach an equilibrium rate, when both pressures are equalized, this equilibrium is disturbed by the natural phenomes quoted above.

Transpiration is a plant related process. The same process occurs with crops where it predominately lose their water through stomata, small openings on the plant leaf through which gases and water vapour pass.” (3)

Transpiration depends on the same factors that influence evaporation.

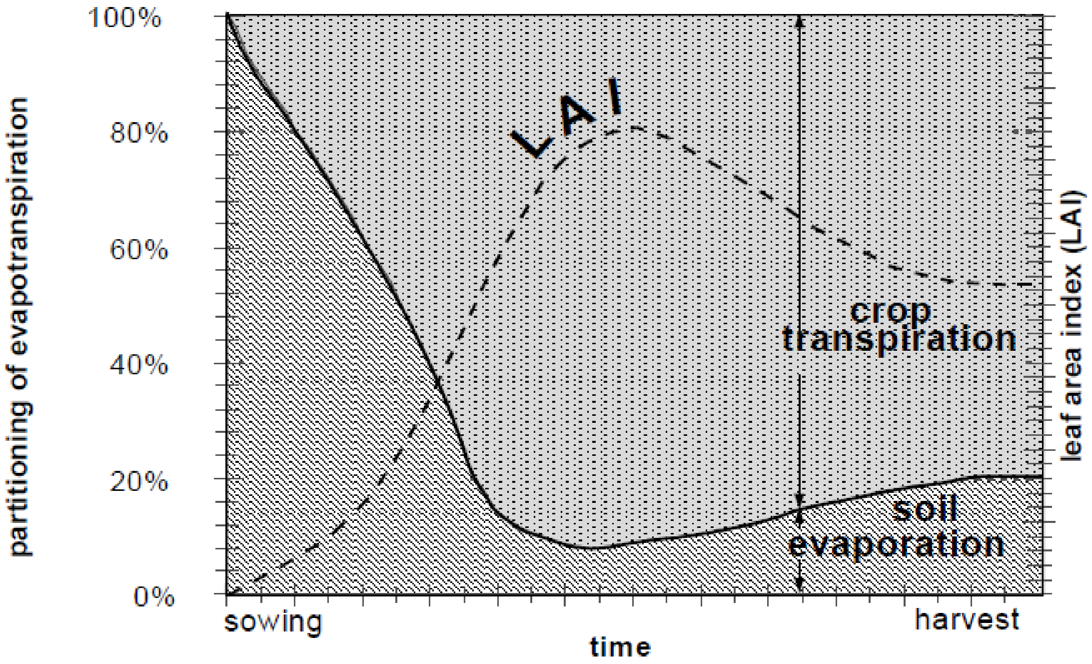


Figure 1 - Evapotranspiration partition of a specific crop (3)

Evapotranspiration is the combined process between evaporation and transpiration. It is usually calculated as a rate,in mm per time.

Evaporation and transpiration can occur simultaneously and there is no easy way of distinguishing between the two processes. (3)

As it is explained in figure 1, evaporation occurs mainly when the crops are still small, but, as time goes by, transpiration starts to play a much important role, mostly because of the leafs.

The process to calculate the evapotranspiration is quite complicated, having different ways to do it:

In order to calculate yearly ($ET_{a,y}$) evapotranspiration the FAO (3) method will be adopted, where:

$$ET_{a,y} = ET_g + ET_b \qquad (1)$$

Where:

$ET_g = green\ water\ evapotranspiration$

$ET_b = blue\ water\ evapotranspiration$

Literature refers as green water for water that comes from rainfall. Instead, blue water is water withdrawn from water bodies such as rivers or groundwater.

The major rationale for making a clear distinction between 'green' and 'blue' water is that the two sources of water have different storage and usage capabilities.

Green water is mostly stored inside the soil and has its main propose to proportionate the crop growth. Whereas blue water can be stored in different sites, such as lakes, natural aquifer, rivers and one of its mains examples is irrigation.

The irrigation water that supports the plants is blue, thus it is important to estimate it (4)

Equation 1 illustrates the total evapotranspiration of a crop during the growing season in a year. Thus, in order to the obtain the value during the whole growing period (GP), according to FAO's (3) approach:

$$ET_{a,GP} = \sum_{i=1}^{GP} ET_{a,i} = \sum_{i=1}^{GP} (k_{c,i} * ET_{0,i} * k_{s,i}) \quad (2)$$

Where:

$ET_{a,i} = daily\ values\ of\ evapotranspiration$

$k_c = crop\ coefficient$

$ET_0 = \text{reference evapotranspiration.}$

$k_s = water\ stress\ coefficient$

All the above metrics show some sort of mutual dependence as indicated in the following paragraphs.

2.3. ET_0

ET_0 is the reference evapotranspiration. The FAO Penman-Monteith (3) equation is responsible for better results.

The resistance factor are introduced $r_a\ (aerodynamic\ resistances)\ [s\ m^{-1}]$ and $r_s\ (surface\ resistance)\ [s\ m^{-1}]$and the expression reads

$$\lambda ET = \frac{\Delta\,(R_n - G) + \rho_a c_p \frac{(e_s - e_a)}{r_a}}{\Delta + \gamma\,(1 + \frac{r_s}{r_a})} \quad (3)$$

Also:

$e_s - e_a = Vapour\ Pressure\ Delta$

$\rho_a = Mean\ Air\ Density$

$c_p = Specific\ Heat\ of\ the\ Air$

$\Delta = Slope\ of\ the\ Saturation$

$\gamma = Psychrometric\ Constant$

ET_0 is the evapotranspiration in reference conditions: an hypothetical well-watered reference crop with an assumed crop height of 0.12 m, a fixed surface resistance of 70 s m-1 and an albedo of 0.23 (5)

By adopting the reference values and inserting them in the equation 3, the following ET_0 equation is achieved

$$ET_0 = \frac{0.408\,(R_n - G) + \gamma \frac{900}{T+273} u_2 (e_s - e_a)}{\Delta + \gamma\,(1 + 0.34\,u_2)}\ [mm\ day^{-1}] \quad (4)$$

Where:

$ET_0 = reference\ evapotraspiration$

$R_n = net\ radiation\ at\ the\ crop\ surface\ [MJ\ \ m^{-2} day^{-1}]$

$G = soil\ heat\ flux\ density\ [MJ\ \ m^{-2}day^{-1}]$

$T = mean\ daily\ air\ temperature\ at\ 2\ m\ height$ [°C]

$u_2 = wind\ speed\ at\ 2\ m\ height$

$e_s = saturation\ vapour\ pressure\ [kPa]$

$e_a = actual\ vapour\ pressure\ [kPa]$

$\Delta = slope\ vapour\ pressure\ curve\ [kPa$ °C$^{-1}]$

$\gamma = psychrometric\ constant\ [kPa$ °C$^{-1}]$

This final form of ET_0be used at different times of the year or other regions, or relate other crops to it.

The calculation of ET_0 of the FAO Penman-Monteith equation is a 4 step procedure where it starts by the wind speed, altitude and air temperature climatic parameters. After that we proceed to calculate $(e_s - e_a)$ and R_n and finally acquire reference evapotranspiration.

It should be noted that the FAO Penman-Monteith equation requires several parameters for its calculation. Also, it can be done for different time steps (monthly, ten days, 24 hours and hourly). Higher the resolution of the data, higher will be the accuracy of the results.

2.4. k_c

The crop coefficient allows to calculate evapotranspiration for specific crops in some regions of the globe.

As already mentioned, calculation of ET_0 is based on the FAO Penman-Monteith equation. As stated from the FAO manual, the principal difference between the two parameters is that k_C shows the capability of differentiating grass from other field crops.

The k_C of grass include four different characteristics regarding to crop height, reflectance of the crop surface, crop evaporation and resistance to vapor transfers.

The first thing to clarify is the temporal variability along the crop growing phases. From figure 2, four growth stages can be distinguished:

Initial Stage: it mainly goes from the crop planting to where it reaches 10% of the ground cover. During this phase we observe small leafs and mainly evaporation in the evapotranspiration process. K_C may vary if we have wet or dry soil (pause during wetting events)

Development Stage: it goes from 10% of the ground cover to the full crop growth. As the crop starts to grow, and to create shadow, evapotranspiration starts to migrate more from evaporation to transpiration. Thus, k_C will vary proportionally to the increase of transpiration, so increasing as the crop growths.

Mid-Season Stage: from full soil cover (end of phase 2) to start of maturation of the crop. k_C is at its max and it keeps constant.

Late Season Stage: maturity to harvest. We will observe that k_C starts to get lower behavior as there will be no more crop.

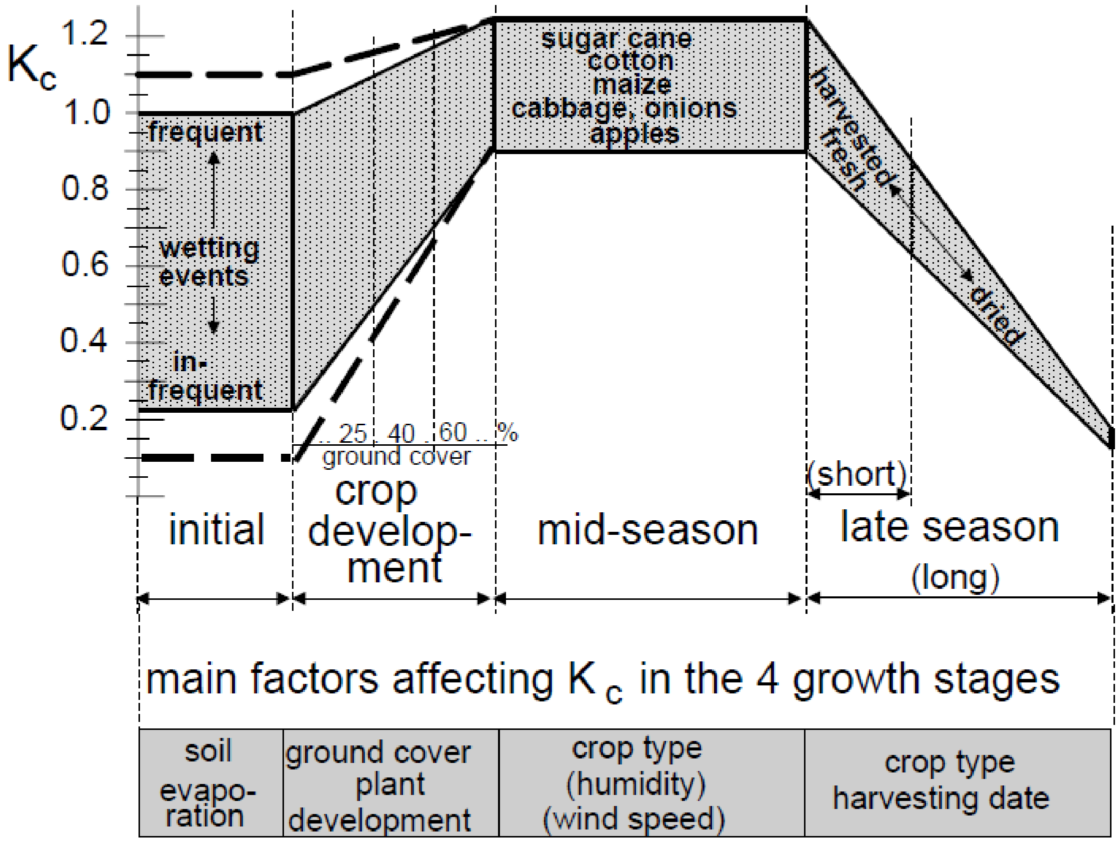


Figure 2 - Crop Growth (5)

According to FAO (3) methodology to The final form of k_c is:

$$k_{c,i} = \begin{cases} k_{c,initial} & i \in initial\ stage \\ i * \frac{k_{c,mid} - k_{c,in}}{\mathrm{i} - I_I} & i \in development\ stage \\ k_{c,mid} & i \in mid\ season \\ i * \frac{k_{c,mid} - k_{c,in}}{\mathrm{i} - I_I - I_{II} - I_{III}} & i \in late\ season \end{cases} \tag{5}$$

2.5. k_s

All the formulas listed so far take into account the standard conditions of the soil, that are mainly encountered in good managed and well-watered fields.

Water has a low potential energy in dry soils and it is strongly bounded to the soil matrix by capillary and absorptive forces, and it is not easily extracted by the crop. (3)

k_s is the coefficient that describes water stress, the condition where soil water is below a certain threshold. When $k_s < 1$ reflects the water in stress conditions.

Before introducing its equation, it is valid to state some other parameters that affect water stress condition

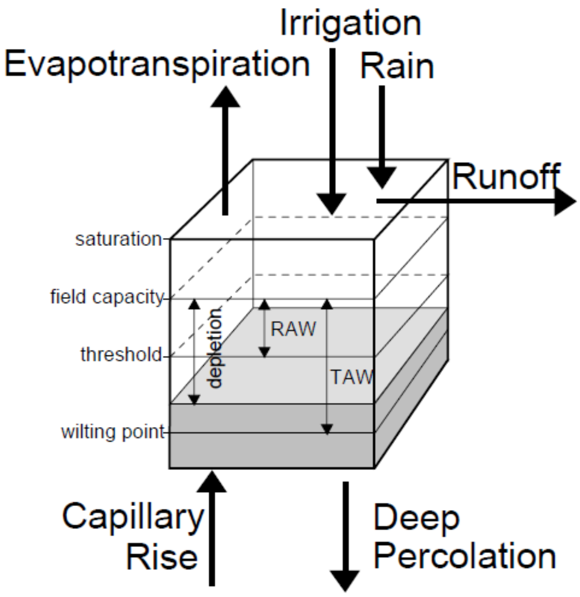


Figure 3 - Water balance on the root zone (3)

TAW is the total available water corresponding to the soil volume necessary to recollect water for the crops.

Instead, RAW, readily available water is the part of TAW that the crop can use without suffering water stress.

Following figure 3 the water stress coefficient equation can be introduced following FAO (3) approach:

$$k_s = \begin{cases} 1 \ \ if \ S_i \geq S_{fc}(1-\rho_i) \\ \frac{S_i - S_w}{(1-\rho_i)(S_{fc}-S_w)} \qquad if \ S_w < S_i < (1-\rho_i)(S_{fc} - S_w) \\ 0 \ \ if \ S_w \geq S_i \end{cases} \tag{6}$$

Where

$S_i = soil\ moisture\ at\ the\ i\ day\ [mm] = 1000\ (S_{fc} - S_w)\ Z_r = TAW$

$Z_r = rooting\ depth\ [mm]$

$S_{fc} = soil\ moisture\ at\ field\ capacity\ [mm]$

$S_w = soil\ moisture\ at\ wilting\ point\ [mm]$

$\rho_i = deplection\ fraction =$

These authors also recommend, in order to avoid crop water stress, irrigations should be applied before or at the moment when the readily available soil water is exhausted (3)

During the development of the thesis the parameters that were analyzed in depth were:

- Total Evapotranspiration (ET_a)
- Irrigation Requirements (ET_{blue})

These variables depend on the soil water balance through RAW and the main driver for the variability is precipitation.

3. Precipitation Data

3.1. ERA 5

Precipitation data are necessary to integrate the discussed model. However, the data collecting is a long process that require important information consisting of the data and the way the collect has been done.

As previously informed, the source of information is the Copernicus Programme, offered by the European Centre for Medium-Range Weather Forecasts, also kwon as ECMWF.

The European Centre for Medium-Range Weather Forecasts is an independent intergovernmental organization supported by 34 states. (6)

The important service provided by the ECMWF is the Copernicus Programme, that uses satellite Earth Observation and in-situ data to provide information services.

The Copernicus Programme have a lot of datasets but the one adopted in this thesis was the ERA 5. This dataset is available inside the Climate Data Service, also known as CDS.

The CDS provides a single point of access to a wide range of quality-assured climate datasets distributed in the cloud. (7)

There, it is allowed access different datasets and have multiple climates information on Earth, not only present but about the past and future estimations.

The selected font of information was the ERA 5, as earlier quoted, that is a reanalysis of global weather and climate during the last seventy years.

Reanalysis in the sense of among all the observational data, the dataset works with models to complete all kinds of climates information., such as:

Wind components

- Precipitation
- Vegetation cover
- Temperature
- Sea Level
- And many others.

Figure 4 has a brief description of how the dataset is structured:

DATA DESCRIPTION	
Data type	Gridded
Projection	Regular latitude-longitude grid
Horizontal coverage	Global
Horizontal resolution	Reanalysis: 0.25° x 0.25° (atmosphere), 0.5° x 0.5° (ocean waves) Mean, spread and members: 0.5° x 0.5° (atmosphere), 1° x 1° (ocean waves)
Temporal coverage	1979 to present
Temporal resolution	Hourly
File format	GRIB
Update frequency	Daily

Figure 4 - ERA 5 Description (8)

3.1.1 Data Extraction

For the data download, the user needs to create a login in the CDS page to generate the CDS API key. After completing the login and in possession of an API key, using Linux command window it can first need install python and then install the CDS AP (annex 1).

A specific python code was elaborated for the purpose of access ERA 5 and exctract a single variable from the dataset, total precipitation (annex 2).

All downloaded data was extracted in the 'netcdf' format beneficial to its capacity when using MATLAB and CDO (Climate Data Operator).

The size of the file is approximate 150 GB and cointaned worldly total precipitation from 1981 up to 2019, hour by hour.

Year of Data	Downloaded at Date
1981	06/09/2020 18:10
1982	07/09/2020 18:45
1983	07/09/2020 18:42
1984	20/10/2020 21:58
1985	09/09/2020 17:07
1986	12/09/2020 13:57
1987	10/09/2020 17:41
1988	12/09/2020 22:55
1989	15/09/2020 05:31
1990	15/09/2020 06:04
1991	16/09/2020 03:46
1992	18/09/2020 08:36
1993	23/09/2020 13:37
1994	19/09/2020 14:36
1995	21/09/2020 17:44
1996	23/09/2020 14:52
1997	24/09/2020 04:55

1998	24/09/2020 22:57
1999	25/09/2020 21:44
2000	28/04/2021 02:14
2001	15/08/2020 19:41
2002	18/08/2020 17:08
2003	20/08/2020 21:29
2004	23/08/2020 04:58
2005	24/08/2020 02:10
2006	24/08/2020 20:03
2007	25/08/2020 04:56
2008	25/08/2020 17:59
2009	27/08/2020 16:52
2010	29/08/2020 16:23
2011	29/08/2020 16:37
2012	30/08/2020 16:00
2013	31/08/2020 21:31
2014	01/09/2020 16:10
2015	01/09/2020 15:57
2016	02/09/2020 05:21
2017	05/09/2020 23:38
2018	03/09/2020 22:24
2019	06/09/2020 17:21

Table 1 - Downloaded ERA 5 Data

It is important to register all the downloading dates from the CDS because ERA 5 is in permanent stage of up-to-date motion (Table 1)

3.1.2. Data Manipulation

In chapter 3.2. it was stated that every file had a mean size of 150GB, multiplied it by the number of years 5.85 TB.

Thus, to reduce the size of each file is necessary to manipulate all the data. The downloaded data worked precipitation values for many hours but daily cumulated values are also in need for the final conclusions of the paper.

A tool named Climate Data Operator, also known as CDO, is employed. This software constitutes a collection of many operators for standard processing of climate and forecast model data. (9)

CDO is a very simple and easy to learn software that allows the user to do the transformation needed on every NetCDF files. The CDO is operated having in mind two functions, to collect data on a daily basis and to better manipulate the ERA 5 data.

cdo daymean –shifttime, -30 2006.nc day_2006.nc

CDO command ‘daymean’ provides a daily mean of the hourly data previously downloaded, inside the ‘2006.nc’ file.

The command ‘-shifttime, -30’ is used because precipitation that happened at midnight of the year before is actually allocated as the first value of the next year. Example: every mm of water that felt during 23:00 31/12/2005 is allocated as 1 string inside the 2006 value.

An example of why the usage of both commands is necessary is illustrated in annex 3.

The final file has 365 entries where it contains worldly mean values of precipitation for each day of the year. All these CDO manipulated files had an average values of 7GB, 96% less than the downloaded value.

Since the final file for each year had a mean daily value, the last manipulation done was to multiply each day by 24 in order to get cumulated daily precipitation values.

Table 2 lists all the work necessary for the extraction and manipulation of the precipitation data.

	Prior to the Manipulation	**After the Manipulation**
Scale	Global	Global
Frequency of data	Hourly values of precipitation	Cumulated daily values
Number of entries	24*365=8780	365
Size	150 GB	7 GB
Format	NetCDF	NetCDF
Unit of measure	m / hour	m / day

Table 2 - Precipitation data manipulation

3.1.3. ERA 5 Deep Dive

To better understand precipitation activity, a statistical analysis was performed and a specific analysis were made at two locations, São Paulo and Torino.

The idea is to take a look on figures 5 that describe the overall behavior of the precipitation in both places before examining the evapotranspiration.

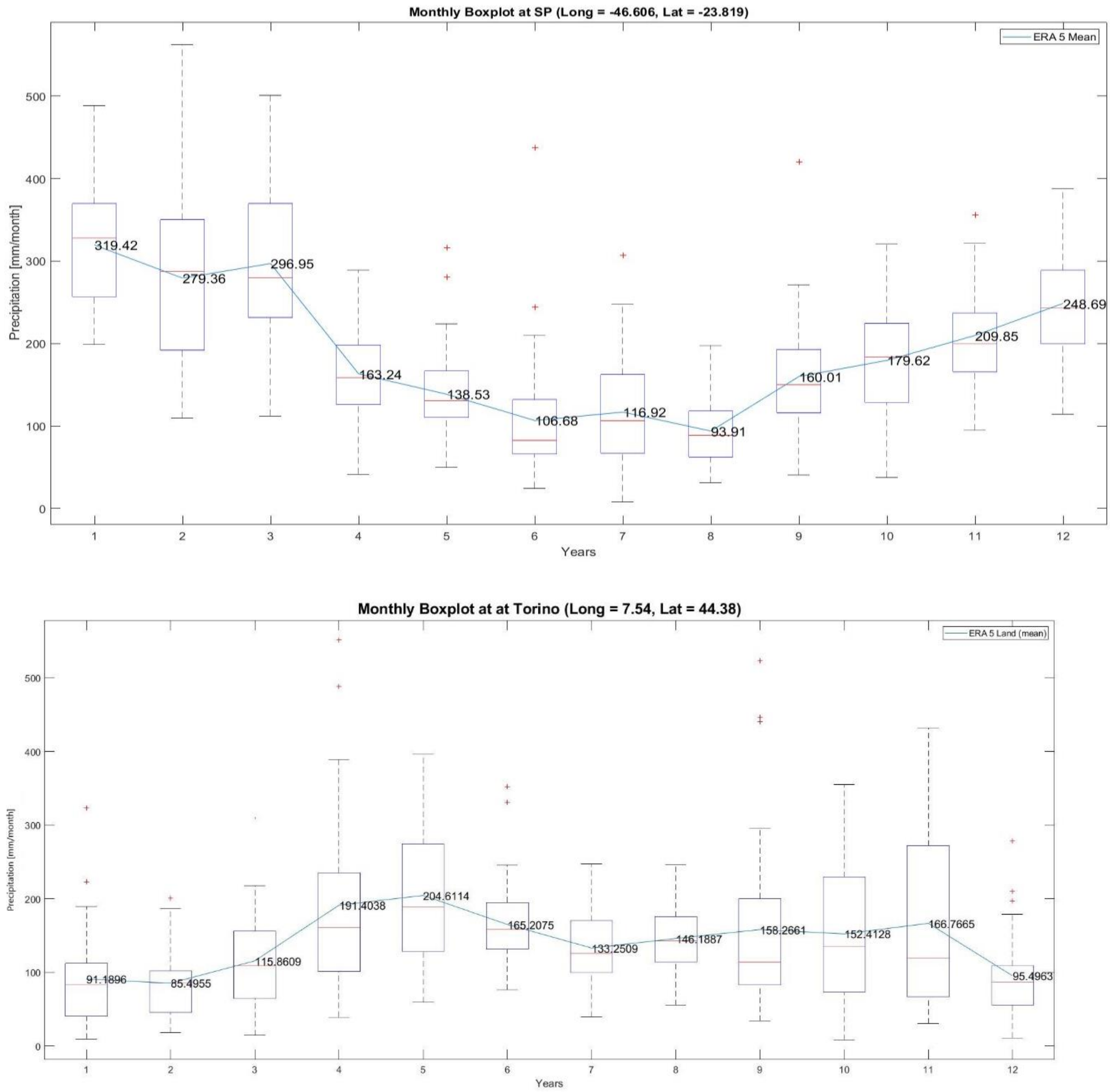


Figure 5 - Precipitation Boxplot for São Paulo and Torino

The rain season in São Paulo occurs during the summer time, from January to March, where mean values of precipitation may reach 319 mm per month.

From figure 5 one can see the gap that do exist between rainy and dry seasons, with the average precipitation value for August reaching 91 mm/month. The lowest value found in our data series pointed to exactly August of 2013 with 32 mm per month.

During the 2013–2015 drought in Southeastern Brazil, the anomalously low inflows to the Cantareira system caused reservoir storage to be reduced beyond minimum operational level (13)

Differently from São Paulo, precipitation behavior on Torino follow a distinct path. However, this latter region is also influenced by drastic rainfall events.

Although variable in seasons, Torino is a less wet area having maximum value of the mean precipitation of 204 mm per month.

The difference in precipitation for both locations can also be emphasized by the amount of rainy days. For this purpose, a rainy day is defined as a day having more than 0.1 mm of daily precipitation.

São Paulo has a higher amount of rainy days in comparison to Torino and, also, a more constant period of rain. A mean value of 334 days of precipitation per year is registered for the city and only of 275 for Torino.

The hemisphere position for both locations exerts an important influence on the precipitation

A common feature to both places is the so-called urban heat island (UHI), a known meteorological and climatic state where the air over urban areas is heated more than the normal pattern to due to the lack of cooling vegetation or to the high concentration of machines

São Paulo is prone to severe weather with major impacts on society given its steady urban growth in the past decades with microclimate changes (10)

Also Torino suffers from the presence of a UHI effect over the city (11)

São Paulo is located in an Atlantic forest biome and it is deeply affected by its proximity to the ocean, causing heavy rainfalls events. Not only, but São Paulo is characterized by extreme dry events.

Over the last years both events show clear tendency of presenting a positive trend

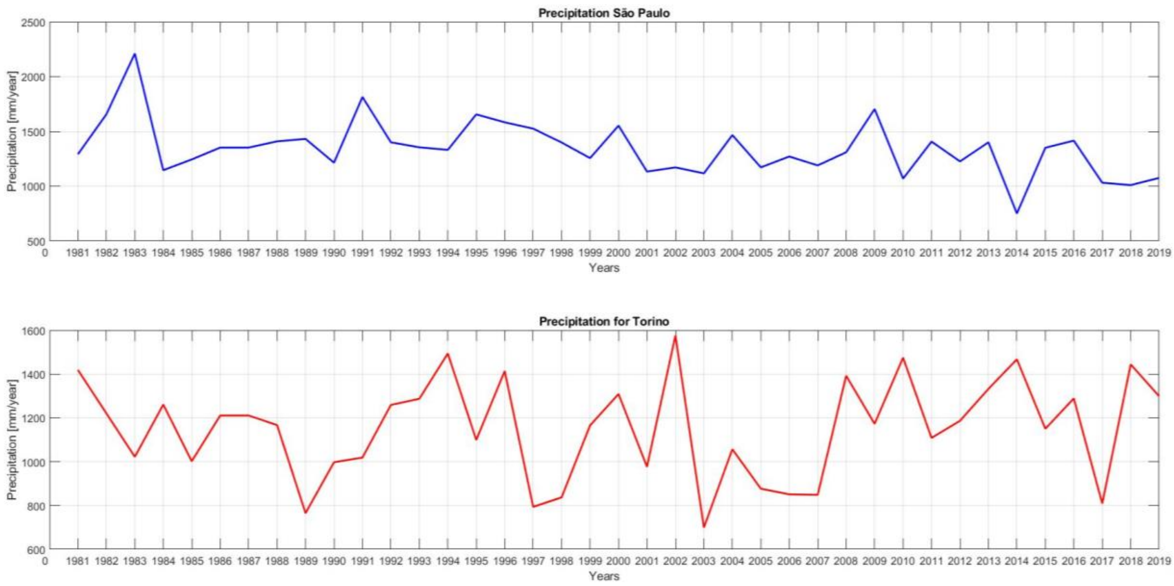


Figure 6 –Yearly Cumulated Precipitation for São Paulo and Torino

The increase in total precipitation is related to the increase in frequency of extreme precipitation, nevertheless also the consecutive dry days number have increased (12), observed in figure 6.

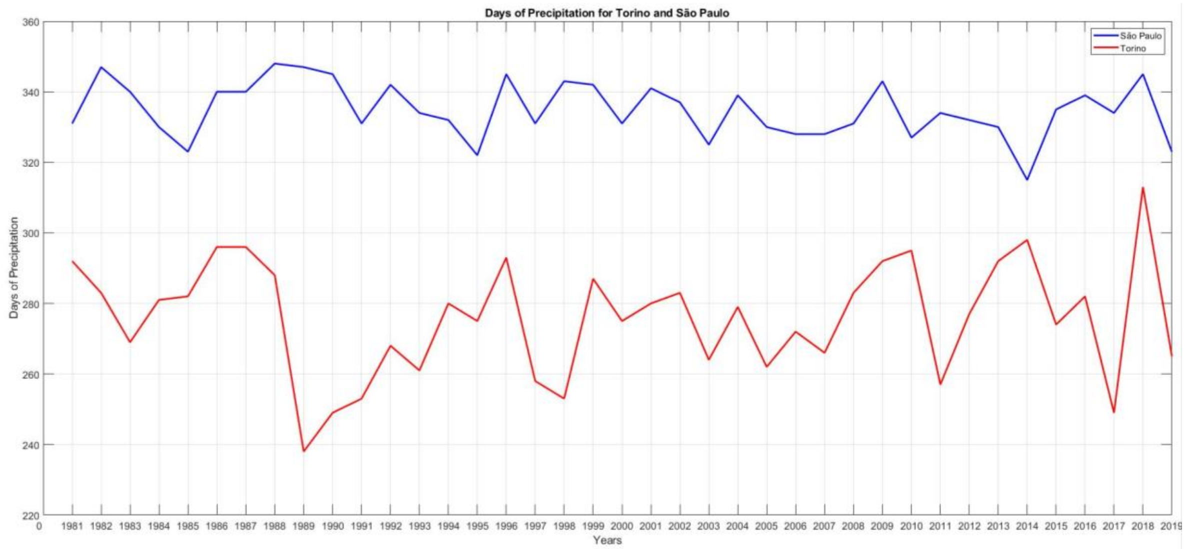


Figure 7 - Rainy Days for São Paulo and Torino

The selected locations have different patterns of behavior and that is the main reason for its choice.

Precipitation has a tremendous influence on irrigation requirements, as it will be highlighted further in the discussion.

Concluding, São Paulo is a more wet area with mean values of precipitation and rainy days higher than Torino. Also, São Paulo presents a higher gap separating rainy (January to March) of non-rain time, that also reflects the larger number of droughts among the years and the smaller values of precipitation during the drought seasons.

3.2. Data Comparison

3.2.1. CRU

In this section another dataset is going to be briefly introduced. This is provided by the Climatic Research Unit (CRU) of University of East Anglia and it is a commonly used dataset for many climate variables. It contains data collected from over 4000 weather stations.

In order to have a control parameter with the ERA 5 precipitation data, it was decided to compare and analyze the results from a ground dataset (CRU) with a satellite dataset (ERA 5).

Thus, the main goal of this analysis was to observe if differences existed and how much both datasets differ from each other.

3.2.2. ERA 5 vs CRU

CRU data was already downloaded as millimeters per month over a global grid with 360x720 cells

As for ERA 5 data the unit of measure was meters per day, as described in table 3. Thus, the necessary manipulation in order to compare with CRU dataset was to aggregate the data into m per month and then transforming it into mm per month.

Also, ERA 5 dataset have a 4320x2160 grid, besides the unit of measure manipulation, a grid reduction was necessary.

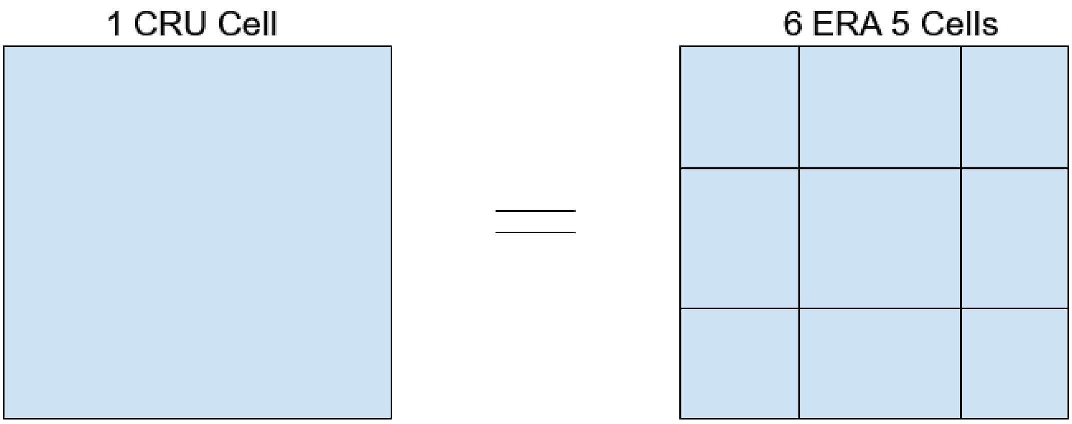


Figure 8 – CRU vs ERA 5 grid

In conclusion, in order to compare both satellites, another round of data manipulation was necessary on ERA 5 data.

Annexes 4 to 7 shows how, for each dataset, data is imported on Matlab and how it was manipulated in order to be able to compare both.

Table 3 sums up the necessary effort done in order to begin the comparison between both datasets.

	CRU	ERA 5	Action
Grid	360 x 720	4320 x 2160	Group 6x6 ERA5 cells
Orientation	360 x 720	4320 x 2160	Rotate anticlockwise and flip ERA 5
Unit of Measure	mm / month	m / day	Group daily values into monthly values for ERA 5 and transform into mm

Table 3 - ERA 5 vs CRU Manipulation

The comparison analysis was done in two steps. The first steps consider three different years were selected at random and for each year a specific location on both matrices also was selected. The objective was to prove that if randomly selecting point inside the dataset, CRU and ERA % would have similar results. The second step consider a 20-year period but focused only on São Paulo and Torino

The three location and years are shown in table 6:

Point on 360x720 grid	Year
90,375 (Torino)	1993
227,267 (São Paulo)	2003
109,639 (Tokyo)	2017

Table 4 ERA 5 vs CRU First Comparison Test

All delta variables inside this analysis considers CRU minus ERA 5.

ERA5 dataset always tends to give higher peaks than CRU, as highlighted in figure 9. Nonetheless, both datasets present the same pattern of behavior,

In the early months of 1993, January to April, there is a significant increase of precipitation phenomena. Both dataset explicitly demonstrate it but ERA 5 tends to have higher values, reaching its maximum delta from CRU on April.

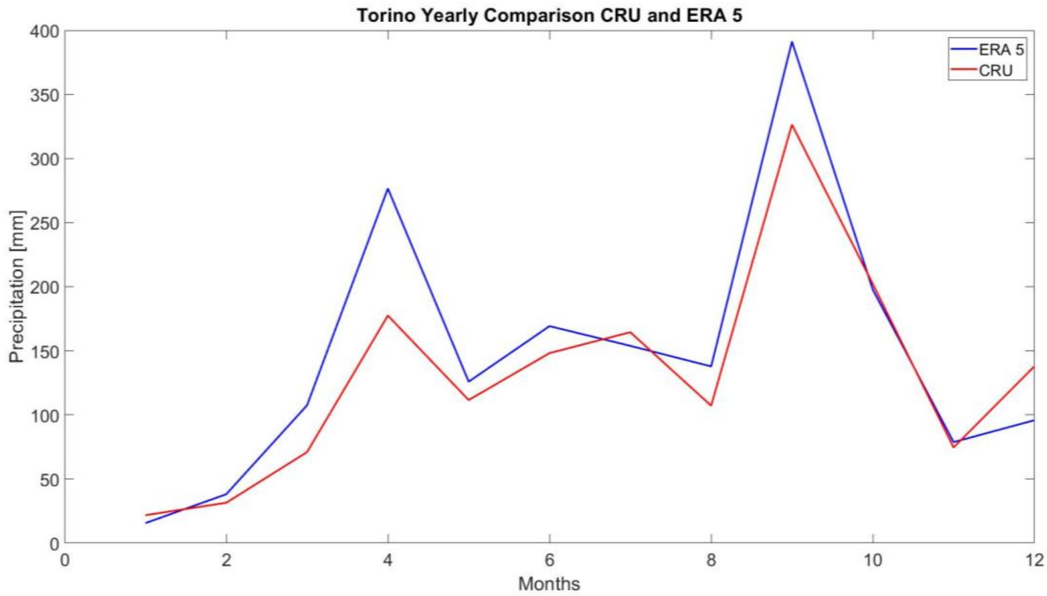


Figure 9 - ERA 5 vs CRU Comparison at Torino on 1993

Figure 10 highlights the last paragraph, the largest difference between the two datasets is on April, where 100mm of precipitation is perceived between both datasets.

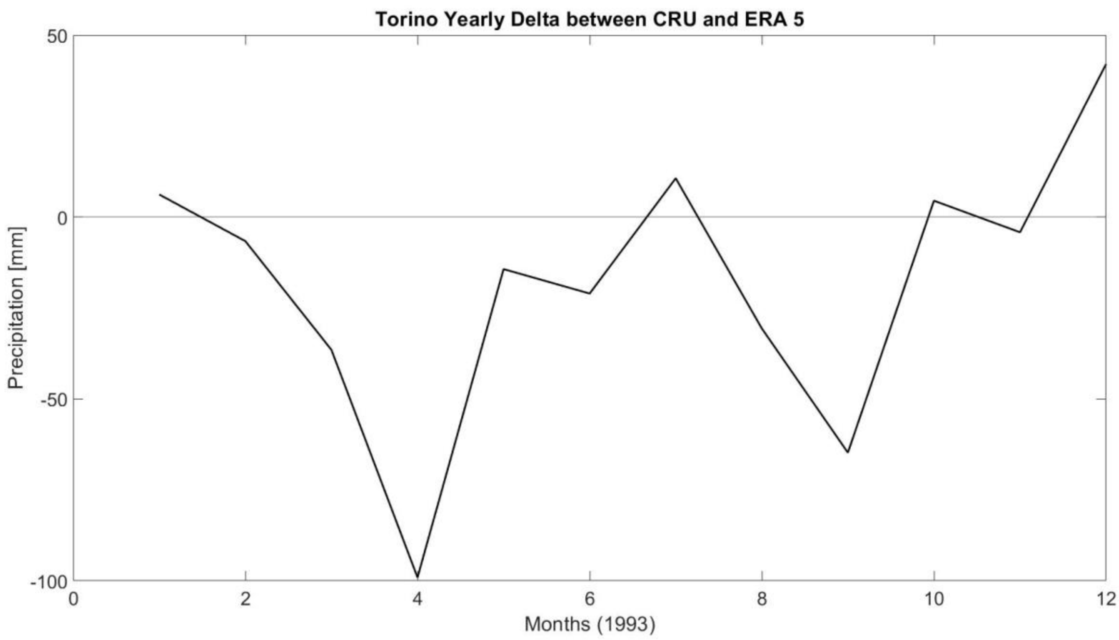


Figure 10 - ERA 5 vs CRU Delta at Torino on 1993

Although the difference between CRU and ERA 5 is always constant, table 7, shows that it very marked when considering single year in different points of the datasets. The maximum 5 of error was 13.09% with monthly mean errors of 25 mm per month.

Annexes 8 up to 12 presents the graphical results for São Paulo and Tokyo and a script example of how the comparison was made.

Location	CRU (mm/year)	ERA 5 (mm/year)	Monthly Mean Error	Overall Delta	Mean of both Dataset	Overall Delta % of Mean
Torino 1993	1,574.40	1,788.34	-17.83	-213.94	1,681.37	-12.72%
São Paulo 2003	1,201.10	1,278.56	-6.45	-77.46	1,239.83	-6.25%
Tokyo 2017	1,331.20	1,517.66	-25.47	-186.46	1,424.43	-13.09%

Table 5 - CRU vs ERA5 First Comparison Summary

Figure 11 illustrates what said in the earlier paragraph. 2017 have the lesser amount of mean error per month, up to -5.46 mm of precipitation. This means that ERA 5 have higher values than CRU.

By taking into account a mean precipitation monthly value for both datasets (by sampling summing all values and diving by 12 entries) and using the mean monthly error, the percentages of error start from 11.96% for 1993 and goes to 5.17% on 2017.

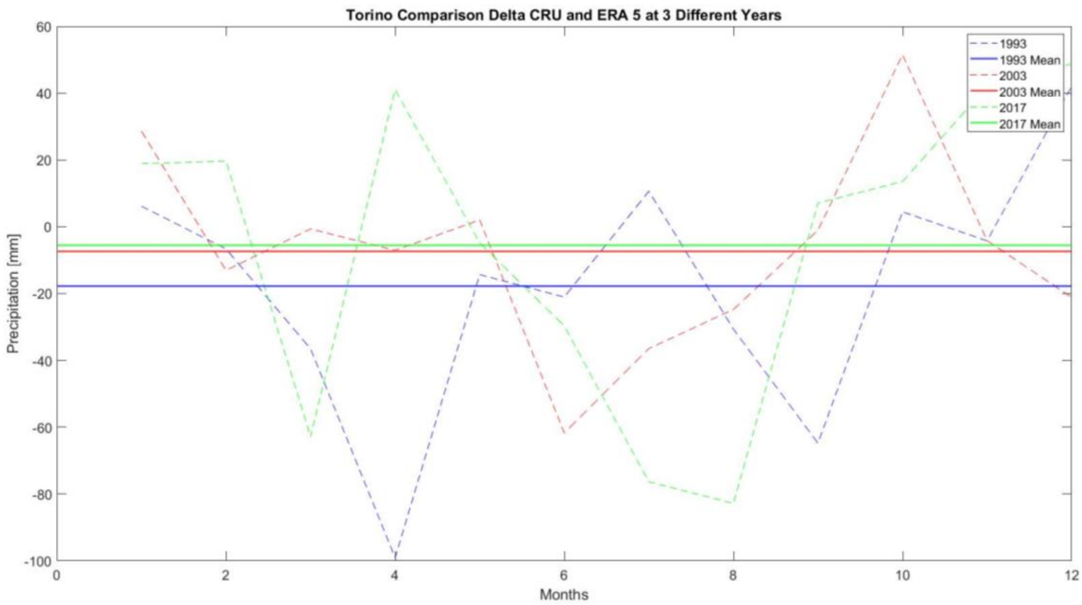


Figure 11 - CRU vs ERA Torino 3 Years Comparison

ECMWF is always improving their methods of calculations and providing reanalyzes on their datasets by optimally combining observations and models, providing representation of the main Earth system cycles (e.g. water, energy)" (14)

Figure 12 exemplifies how technology improved data quality, not only by enhancing the models but also adding new and innovative tools.

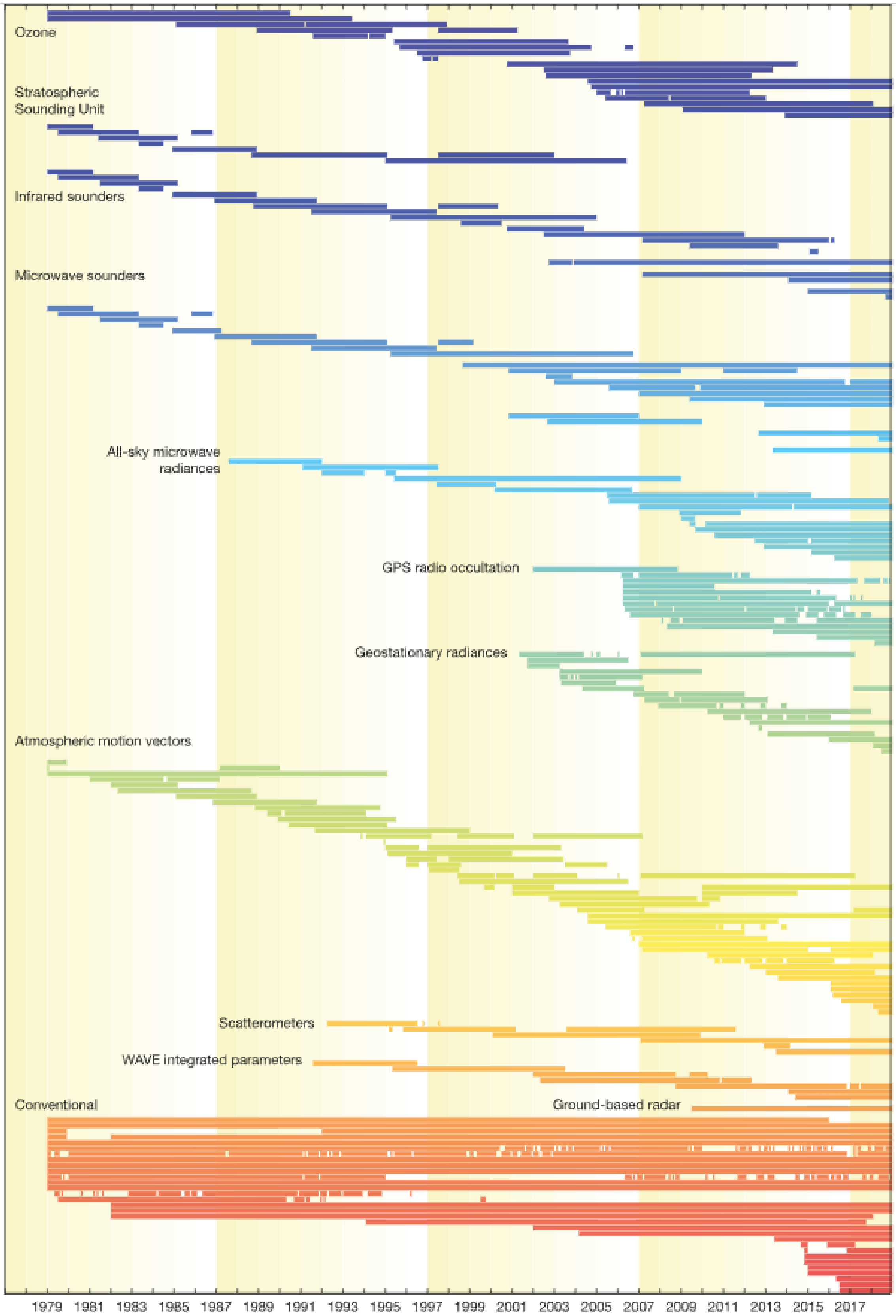


Figure 12 - Usage of Data on ERA 5 (14)

The usage of a particular satellite instrument, ground-based radar or a specific source of conventional data is represented by the horizontal bars represents.

For instance, 2017 data have a great deal more instruments, thus it is safe to conclude that ERA 5 and CRU differences tend to decrease with passing of time, as highlighted in figure 12.

As previously said, two comparison analysis were done, the first was already described, the second was done by enlarging the scale of time of the comparison.

A larger period of time was selected and then compared both CRU and ERA 5. Time frame chosen was from 1990 to 2010. This time only São Paulo and Torino were the decided locations.

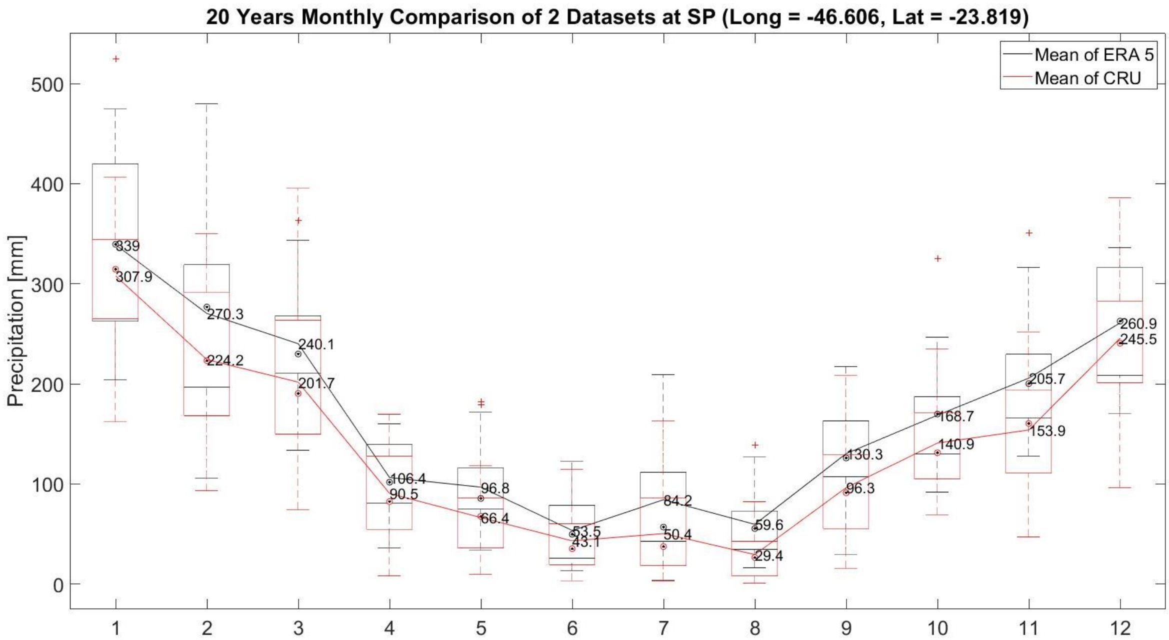


Figure 13 - Twenty Years Comparison Boxplot CRU and ERA 5 for São Paulo

The results on figure 13 confirm what previously stated. Differences do exist between both datasets but overall behavior follows the same trend.

Also, ERA 5 have higher values when compared to CRU. Figure 13 shows not only higher mean values but higher outliers and range of data.

Annexes x from x illustrate the script and other graphs resulting from the 20 years' analysis.

Although there are minor differences between ERA 5 and CRU, it was decided to proceed with full ERA 5 downloaded data.

3.3. Crop Selection

As described in table 4, downloaded data is on a global scale. In this section, crop and location of future analysis are going to be explained.

Evapotranspiration and irrigation requirements are process that are specific for each crop, as highlighted in section 2.2.

Although two locations were priory selected, São Paulo and Torino, specific crops needed to be selected. The aim of this section is to demonstrate how and why each crop was selected.

In order to proceed with the selection, MIRCA2000 dataset was used. It is a monthly dataset that offers rainfed and irrigated crop areas for 26 main crop classes with 5 arc minutes of spatial resolution.

Since the city of São Paulo cover a large area, 3 different points were selected. The goal was to select 3 crops are cultivated in both São Paulo and Torino, being 1 perennial crop (cultivated all year) and the remaining seasonal crops (cultivated on certain months)

The y axis on figure 16 represents the irrigated area on each location selected. The number inside the parenthesis on the x axis represent the point on the ERA 5 4320x2160 matrix.

Even though the obvious choice would be select São Paulo 1 because of its high amount of irrigated maize areas, other common crops with Torino had very low amounts of irrigated area.

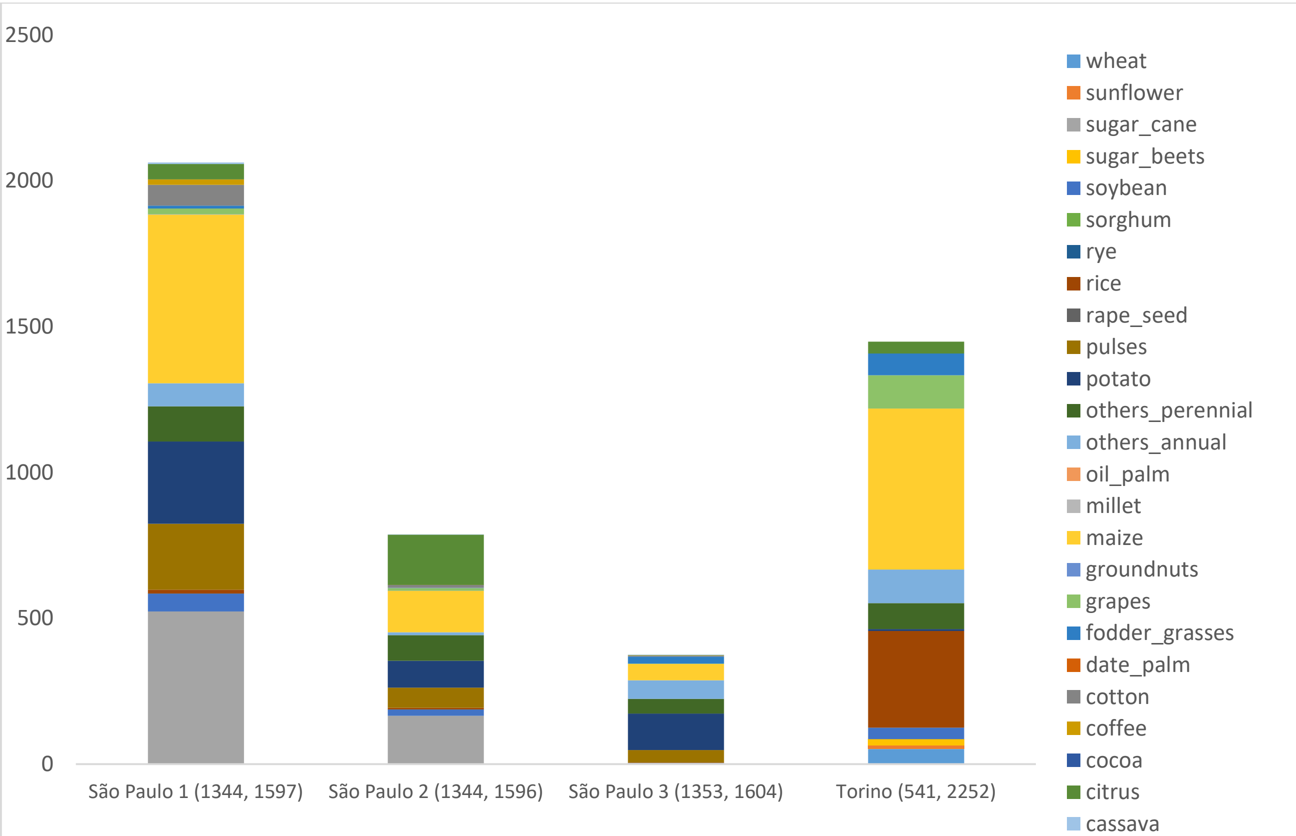


Figure 14 - MIRCA2000 Crop Selection

Thus, the selected location to was Sao Paulo 2 and the selected crops are listed in table 6. Each value represents acres of irrigated area.

	São Paulo 2 (1344, 1596)	Torino (541, 2252)
Wheat	0.75	51.43
Maize	142.22	551.61
Citrus	172.26	40.27

Table 6 - Selected Crops

4. Temporal Variability Analysis

4.1. Introduction

After collecting all the necessary precipitation data and selecting the specific crops and locations, the next steps is to feed all this information inside an already existing evapotranspiration model.

The model in question was previously developed by Matteo Rolle, PhD student at Politecnico di Torino.

The model calculates the irrigation requirement using a soil-water balance on land equipped for irrigation (15)

It takes daily cumulated precipitation values and generates the following results, taking into consideration the reference evapotranspiration, soil parameters and crop characteristics:

- Evapotranspiration (ET blue + ET green)
- Irrigation Requirements (ET blue)
- Soil Moisture

At section 3.3 it was detailed that the development of this thesis would only take into consideration four crops (table 6): Wheat, Maize and Citrus.

Thus, the outputs of the model used to generate every analysis in this section are daily cumulated values of each of the stated metrics.

The irrigation requirement for wheat at São Paulo is a 39x366 matrix, where 39 is the number of years analyzed (1981 up to 2019) and 366 is the maximum number of days in a year, counting bissextile years.

Annex 17 shows an example of data disposal by the model

The structure of the chapter is:

- Analyze results for São Paulo
- Analyze results for Torino

4.2. São Paulo Results

4.2.1. Evapotranspiration

In order to have a global visualization of the overall time series behavior of all crops, plots like figure 15 were constructed to have a brighter view of possible clear patterns.

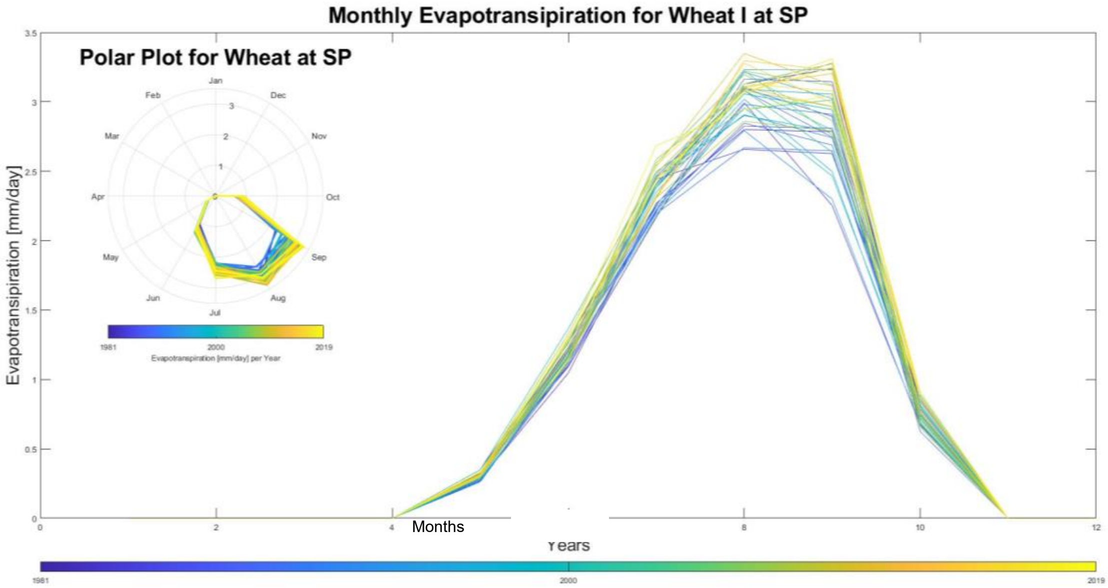


Figure 15 - Monthly Evapotranspiration for Wheat at São Paulo

As the color bar explicitly shows, older years have a blue line going up to the yellow line in more recent years. From this figure, a clear pattern can be observed: evapotranspiration for wheat in São Paulo has been increasing over the years.

The same increasing behavior can be observed in the plots for citrus and maize.

Taking a closer look at the data, table 8 illustrates better the overall behavior over the years for wheat.

It considers mean values of daily mean evapotranspiration for every nine years. Calculating the percentage of one group year to another, May presented the lowest mean of increase with 0.399% and September the highest with 4.119%

	1981-1990	1991-2000	2000-2009	2010-2019
May	0.310	0.300	0.318	0.313
June	1.156	1.171	1.228	1.211
July	2.280	2.342	2.429	2.440
August	2.931	3.015	3.120	3.120

September	2.795	2.796	2.934	3.150
October	0.752	0.753	0.763	0.814

Table 7 – Wheat Monthly Mean Evapotranspiration at São Paulo

Results for citrus and maize are presented in annexes 20 and 21.

Citrus is a perennial crop, the lowest values of the mean came from February -0.7029% and June obtained the highest percentage of increase with 0.8979% over the 9 years clusters.

Maize, on the other hand, is a seasonal crop with evapotranspiration values from November up to April. The highest mean value percentage attained was for December with 0.7270% and the lowest was on February with -0.6795%

Even though evapotranspiration and irrigation requirements have strong correlation, it is necessary to study irrigation requirements time series in order to determine its pattern over the years.

Since irrigation requirements depend on evapotranspiration, analyze the global behavior can give a glimpse of its overall conduct.

In order to prove this condition, a correlation analysis between evapotranspiration and irrigation requirement was done for all crops chosen that are present in the São Paulo region (wheat, citrus, and maize).

When data are correlated the change in the magnitude of 1 variable is associated with a change in the magnitude of another variable. (16)

Figure 16 is one of the three correlation analyses made. It explicitly shows that for wheat in São Paulo there is a strong and positive correlation between evapotranspiration and irrigation requirements.

Although already stated by equation 1, a strong and positive Pearson correlation coefficient means that as evapotranspiration values increase, also irrigation requirements will.

The Pearson correlation coefficient R is calculated following equation (17);

$$R = \frac{\sum_{i=1}^{n}(y_i - \mathrm{y}) * (x_i - \mathrm{x})}{\sqrt{[\sum_{i=1}^{n}(x_i - \overline{x})^2] * [\sum_{i=1}^{n}(y_i - \overline{y})^2]}}$$

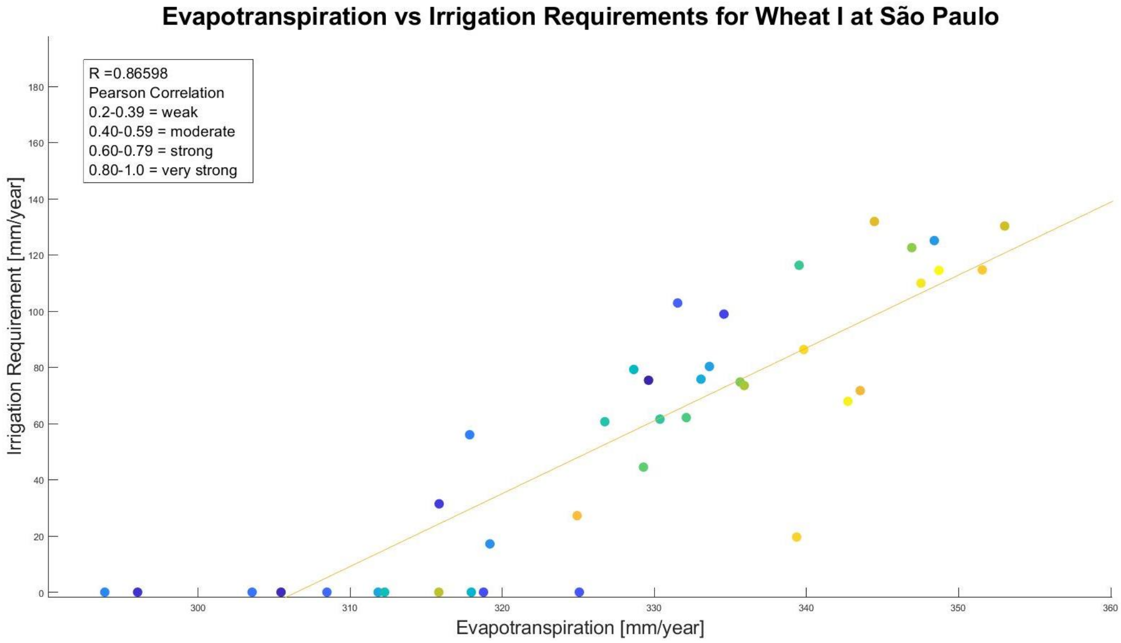


Figure 16 - Correlation between Evapotranspiration and Irrigation Requirements for Wheat at São Paulo

Table 7 sums up the correlation values for all crops at São Paulo

Crop	R	Correlation Strength
Wheat	0.8660	very strong
Citrus	0.6089	strong
Maize	0.8437	very strong

Table 8 – Pearson's Correlation Coefficients for Crops at São Paulo

.

4.2.2. Irrigation Requirements

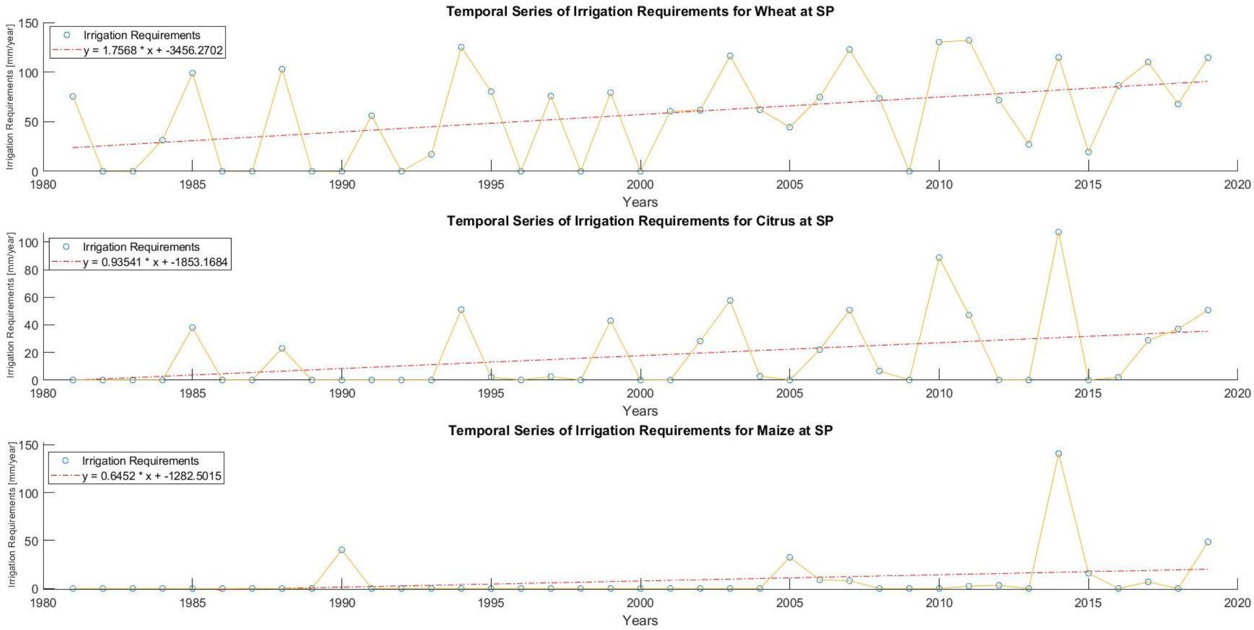


Figure 17 - Irrigation Requirements Time Series for São Paulo

As shown in figure 17, all three crops present positive values of the angular coefficient.

To verify that irrigation requirements present a temporal trend over the years, a statistical test was done, the t-student test.

It is a famous statistical hypothesis test that has the goal of determining if two variables are dependent on one another.

The null hypothesis (H_0) to test is that irrigation requirement and years are not linearly dependent.

Since the time series consists of 39 values and the significance chosen was 99.75%, $t_{lim} = 2.023$ selected based on the t-student critical table on annex 22.

	Wheat	Citrus	Maize
Mean	57.249	17.654	7.896
Sxy	8,678.392	4,620.930	3,187.282
b1	1.757	0.935	0.645
b0	-3,456.270	- 1,853.168	-1,282.502
sig2	1,770.425	635.097	561.977
T	2.935	2.609	1.913
H_0 Hypothesis ($t_{lim} < T$)	Rejected	Rejected	Accepted

Table 9 - T Student Results for Irrigation Requirements at São Paulo

Table 9 values were calculated accordingly to Student (18)

$$b_0 = \overline{y} - b_1 * \overline{x} \quad (8)$$

$$b_1 = \frac{\sum_{i=1}^{n}(y_i - \overline{y}) * (x_i - \overline{x})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} = \frac{S_{xy}}{S_{xx}} \quad (9)$$

$$S_\varepsilon = \sum_{i=1}^{n} (y_i - b_0 - b_1 - x_i)^2 = \sum_{i=1}^{n} \varepsilon^2 \quad (10)$$

$$\sigma_\varepsilon^2 = \frac{S_\varepsilon}{n-2} \quad (11)$$

$$T = \frac{b_1}{\sqrt{\left(\frac{\sigma_\varepsilon^2}{S_{xx}}\right)}} \quad (12)$$

The results of the t student test, table 9, highlight those irrigation requirements for wheat and citrus that have been growing over the analyzed years in a statistically significant way.

Although t student test shows important results, other topics need to be addressed.

Almost all the peaks years with maximum irrigation requirements for wheat are also maximum the for citrus: 1994, 2003, 2007, 2014.

As illustrated in chapter 3.1.3, 2014 was a very difficult drought year for São Paulo in terms of precipitation, this also may affect irrigation requirements.

Maize presents 0 values of irrigation requirements over the years for 75% of the time series (29 out of 39).

The number of years with zero irrigation requirements has changed over the analyzed period.

	1981-1990	1991-2000	2001-2010	2011-2019
Wheat	6	4	1	0
Citrus	8	6	3	3
Maize	9	10	7	3

Table 10 - Years with zero Irrigation Requirements for São Paulo

Every single crop shows a decrease of years with zero irrigation requirement over a nine-year period.

For wheat and maize, every nine years there may be a mean drop of 2 years that will need irrigation requirements. For citrus, the mean values are 1.67 years.

While an increase of irrigation requirement volume is observed along the years, for wheat and citrus, there is a need to observe if such pattern exists for days of necessary irrigation requirements

Thus, a correlation study between both metrics was done and results are shown in table 11.

	Angular Coefficient	R

Wheat	0.3243	0.9872
Citrus	0.4716	0.9950
Maize	0.2592	0.9942

Table 11 - Correlation between Irrigation Requirements and days of Necessary Irrigation Requirements at São Paulo

As previously defined, all three crops present very strong correlation coefficients. It means that while irrigation requirements grow also the days of necessary irrigation requirements increase.

4.2.3. Days of Necessary Irrigation

A day with necessary irrigation requirement is when irrigation requirement for that day is greater than zero.

Figure 18 illustrates the overall behavior of all crops. In order to verify the temporal trend for days of necessary irrigation, an additional t student test was performed, with results in table 12.

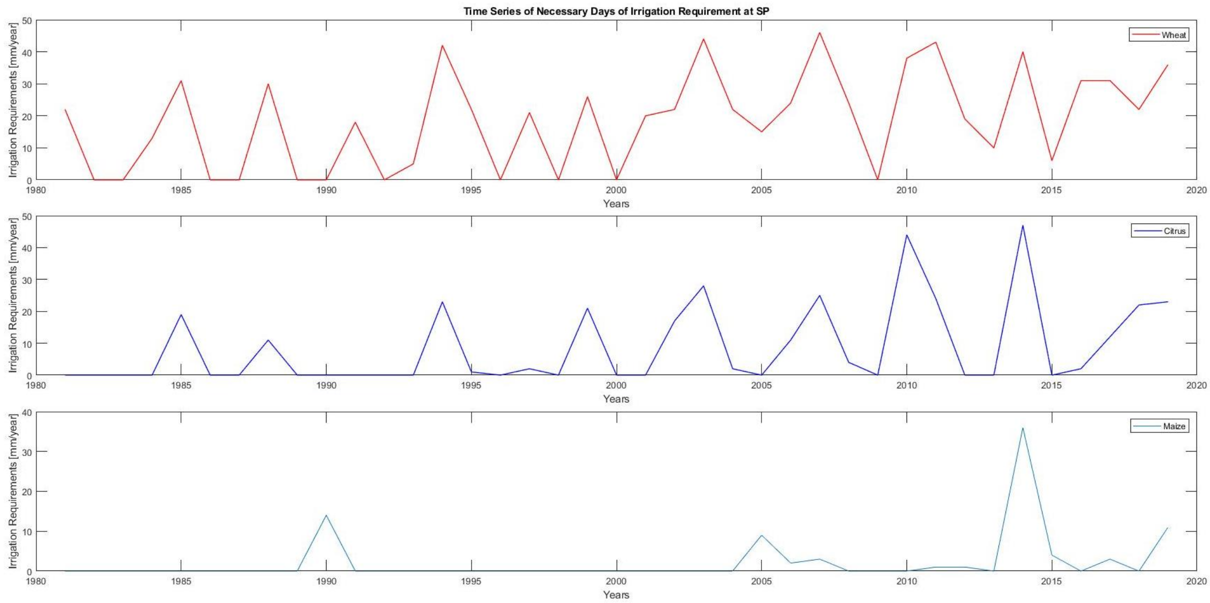


Figure 18 - Time Series of Days of Necessary Irrigation for São Paulo

	Wheat	Citrus	Maize
Mean	18.538	8.667	2.154
Sxy	2,882.000	2,243.000	785.000
b1	0.583	0.454	0.159
b0	- 1,148.263	-899.430	-315.660
sig2	190.063	141.358	38.604

T	2.974	2.684	1.798
$\boldsymbol{H_0}$ **Hypothesis** $(\boldsymbol{t_{lim}} < \boldsymbol{T})$	Rejected	Rejected	Accepted

Table 12 - T Student Results for Days of Necessary Irrigation at São Paulo

Days of necessary irrigation follow the same pattern of temporal trend as for Irrigation requirements, also explained by their strong correlation.

Thus, wheat and citrus present an increased value of day with necessary irrigation over the analyzed period.

By taking a closer look at each crop, since they do not have the same growing seasons, some particularities were found.

<u>Wheat</u>:

Has a growing season that runs between August and October, being a seasonal crop.

September is the month that needed most irrigation, not only in terms of volume (mm) but also in number of days, 39 mm/month, and 12 days of necessary irrigation on average. It represents an average of 70% of the total mm of irrigation and days of necessary irrigation.

The worst year in terms of mm of irrigation was 2017 with 97.95 mm/month over 28 days, almost 3,5 mm/day of irrigation.

While the worst year in terms of days of necessary irrigation was 2007 with 30 days of necessary irrigation and 90.65 mm/month, 3.02 mm/day on average.

The overall behavior of irrigation requirement and days of necessary irrigation both presented growth with mean values of 1.03 mm/month and 1.03 days respectively, year by year.

<u>Citrus:</u>

Although it is a full-year crop (or permanent crop), irrigation requirement is only needed between July and November.

Again, September is the critical month with, on average, 4 days of necessary irrigation and 9.10 mm/month. It represents 58.77% of total days of irrigation and 59.96% of total volumes of irrigation, meanly.

The worst month in terms of mm of irrigation was September 2010 with 49.05 mm/month over 26 days, almost 1.89 mm/day of irrigation. September 2010 is also the worst period when taking into consideration days of necessary irrigation

The overall behavior of irrigation requirement and days of necessary irrigation both presented growth with mean values of 0.33 mm/month and 0.61 days respectively, year by year.

<u>Maize:</u>

Maize is a seasonal crop, irrigation requirement is only needed between January and April.

February is the critical month with, on average, 1 day of necessary irrigation and 4.12 mm/month. It represents 42.68% of total days of irrigation and 44.31% of total volumes of irrigation, meanly.

February 2014 has the worst values both of irrigation requirements and days of necessary irrigation with 71.84 mm/month over 14 days, almost 5.05 mm/day of irrigation.

The overall behavior of irrigation requirement and days of necessary irrigation both presented growth with mean values of 0.28 mm/month and 0.29 days respectively, year by year.

4.2.4. Distribution Fitting

In order to fully understand the comportment of the time series of the present crops, a goodness of fit test was done using Pearson's Chi-Square Test (19)

It is used to understand which set of events are occurring at the same frequency or if they follow a predetermined known distribution.

It was done to verify whether the irrigation requirements data for wheat, citrus, and maize fit one of the following distributions: Normal, Lognormal, or Gumbel.

The first step is to determine the number of classes (k), probability inside each class (q) and the expected number of elements (E_i).

The test is called Pearson's because when calculating the number of classes k, Pearson's equation will be applied.

Thus,

$$k = 2 * i^{0.4} \tag{13}$$

i is the number of values inside the distribution. Since the analysis was done with cumulated yearly values of irrigation requirement, i is equal to 39 (years inside distribution)

$$q = \frac{1}{k} \tag{14}$$

$$E_i = \mathrm{q} * \mathrm{i} \tag{15}$$

In order to test the goodness of fit for each distribution, class limits need to calculated.

$$x_{\lim Normal} = [\theta_1 + \theta_2 * \Phi^{-1}(E_i)] \tag{16}$$

$$x_{\lim Log\ Normal} = e^{[\theta_1 + \theta_2 * \Phi^{-1}(E_i)]} \tag{17}$$

$$x_{\lim Gumbel} = [\theta_1 + \theta_2 * \Phi^{-1}(E_i)] \tag{18}$$

Where θ_1 and $\theta_2$are distribution parameters and Φ^{-1} is the norm inverse

Thus, table 13 sums up all the necessary parameters to proceed with the test application.

Parameter	Wheat	Citrus	Maize

k	9	7	6
q	0.1111	0.1429	0.1667
E_i	4.3333	2.7143	1.6667
$\theta_{1,normal} = \mu$	79.7395	36.2365	30.7944
$\theta_{2,normal} = \sigma$	33.6301	28.8934	41.8969
$\theta_{1,lognormal} = \ln(\overline{x}) - 0.5 * [\ln\left(s^2/\overline{x}^2\right)]$	4.2969	3.3440	2.9035
$\theta_{2,lognormal} = \sqrt[2]{\ln\left(1+ s^2/\overline{x}^2\right)}$	0.4046	0.7015	1.0236
$\theta_{1,Gumbel} = \overline{x} - 0.5772 * s * (\sqrt[2]{6} * \pi)$	64.6098	36.2365	30.7944
$\theta_{2,Gumbel} = s * (\sqrt[2]{6} * \pi)$	26.2213	22.5281	32.6668

Table 13 - Distribution Parameters for São Paulo

All the parameters In table 13 were calculated according to the Statistical Hydrology handbook (20):

$\mu, \overline{x}$ are the distribution and sampling mean.

σ, s are the distribution and sampling standard variation.

Defined the classes limits for each crop, it is necessary to attribute each value inside the distribution to its correspondent class.

At this point the chi-squared test was applied. It consist on calculating the squared difference between the observed values (O_i) and the expected number of elements (E_i) and divide it by E_i. The final result is a summation for all classes, as showed in equation 13:

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \qquad (13)$$

In order to reject or accept the null hypothesis (H_0), χ^2_{lim} is compared with χ^2 for each distribution.

χ^2_{lim} is calculated by taking into consideration:

- $dof = k - s - 1$
- $\alpha = 0.05$

Where dof are the degrees of freedom, s is the number of expected parameters (=2 in this test) and α is the level of significance of that H_0 is rejected.

Thus, these values are used as input to get χ^2_{lim} from the chi square distribution table, illustrated in annex 29.

If $\chi^2 \leq \chi^2_{lim}$, H_0 is accepted, therefore sample data (yearly volumes of irrigation requirement) fit well, within the level of significance, the known distribution.

	Wheat	Citrus	Maize
dof	6	4	3
χ^2_{Normal}	6.538	5.684	11.599
χ^2_{lim}	12.591	9.488	7.815
$H_{0\ Normal}$	accepted	accepted	rejected
$\chi^2_{LogNormal}$	9.308	9.368	4.400
χ^2_{lim}	12.591	9.488	7.815
$H_{0\ LogNormal}$	accepted	accepted	accepted
χ^2_{Gumbel}	8.385	6.421	9.200
χ^2_{lim}	12.591	9.488	7.815
$H_{0\ Gumbel}$	accepted	accepted	rejected

Table 14 - Chi Squared Test Results São Paulo

Table 14 lists the results of the tests for all crops.

4.2.5. Conclusion

In conclusion, it was observed in this chapter that wheat and citrus have a similar behavior. Both crops show temporal trends for yearly volumes of irrigation requirements and days of necessary irrigation. Also, both fitted well for all three analyzed distribution

On the other hand, maize presented different pattern. Did not have a temporal trend for yearly volumes of irrigation requirements or days of necessary irrigation and the only distribution that fitted its samples was the Lognormal

4.3. Torino Results

4.3.1. Evapotranspiration

In this chapter the results of the analysis made for Torino will be presented and commentated.

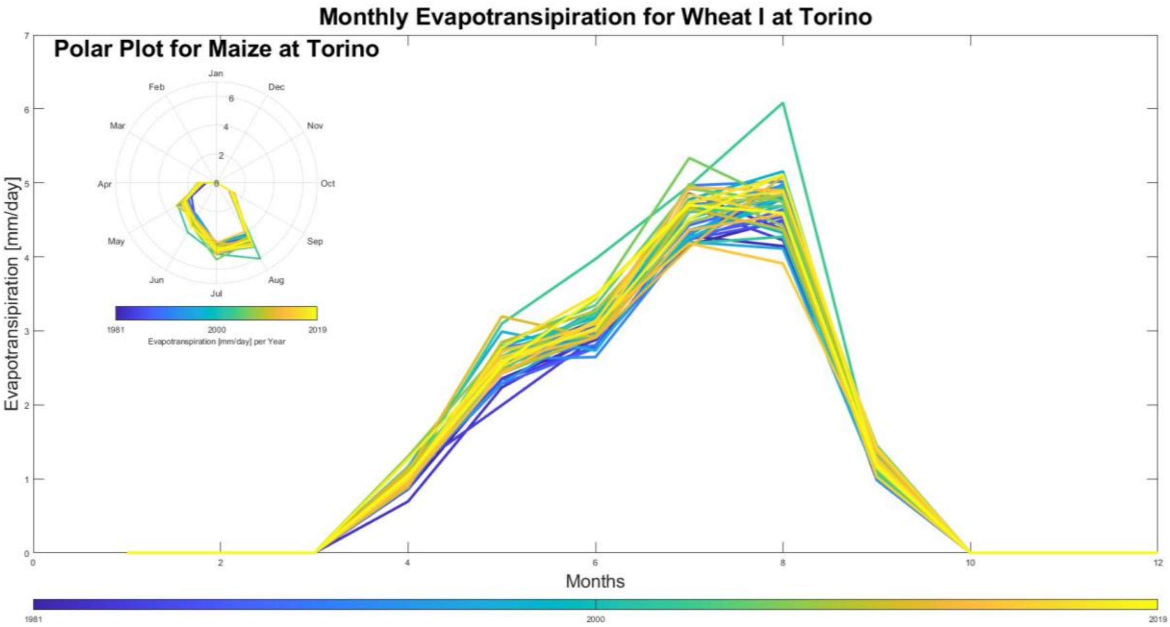


Figure 19 - Monthly Evapotranspiration for Wheat at Torino

From figure 19 there is no clearly visual pattern of increasing evapotranspiration. Although for April to May, some increasing values may be observed, the same pattern does not repeat itself over the next months.

Citrus and maize, show similar behavior where in the first months of crop evapotranspiration it seems there is an increasing value, but in the final months, this pattern cease to exists.

Important to denote that middle 2000s (green lines) assume the highest values, that may be a response of the heat wave already commentated.

The same 9-year monthly analysis done for São Paulo, with results on table 16, was performed for Torino

	1981-1990	1991-2000	2001-2010	2011-2019	2003
April	0.957	0.963	0.970	0.996	1.040
May	2.439	2.459	2.457	2.499	3.098
June	2.922	2.931	2.882	2.900	3.971
July	4.420	4.501	4.496	4.440	4.958
August	4.527	4.579	4.643	4.705	6.083
September	1.168	1.191	1.180	1.166	1.139

Table 15 - Wheat Monthly Mean Evapotranspiration at Torino

Results in table 16 are mm/day of evapotranspiration.

In order to be able to compare crops behavior in both location, the same skeleton of analysis made for São Paulo were applied to Torino.

Starting by calculating the correlation between evapotranspiration and irrigation requirement for crop at Torino (Figure 20)

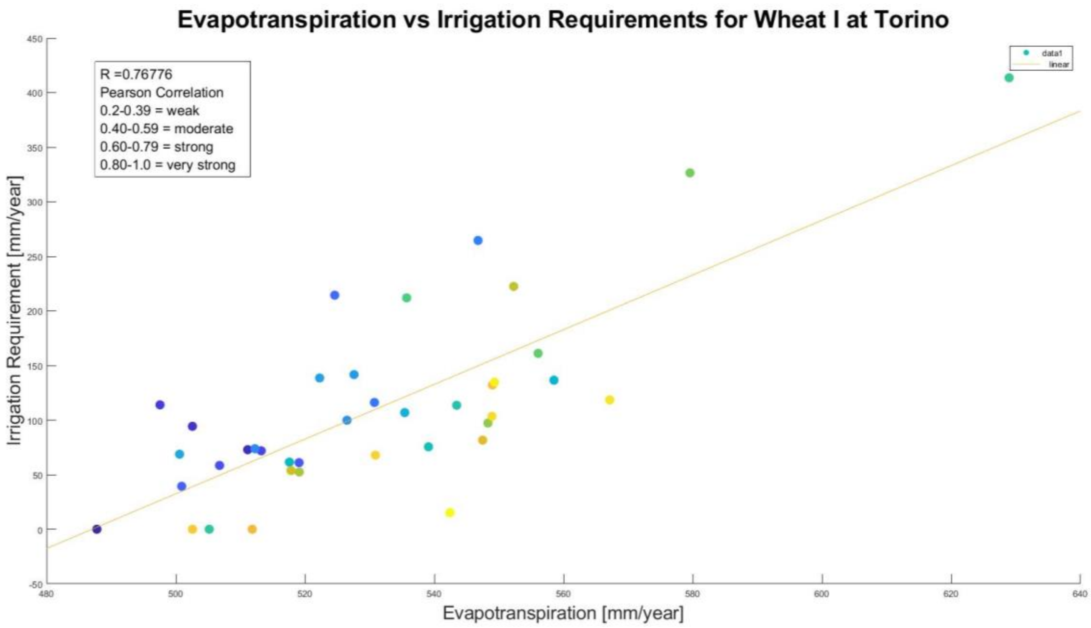


Figure 20 - Correlation between Evapotranspiration and Irrigation Requirements for Wheat at Torino

As for São Paulo, wheat at Torino also presents high correlation between both metrics.

The outlier year in figure 21, is 2003. Important to highlight that this specific year because it was when an extraordinary high temperature wave hit Europe. Piedmont, region where Torino is situated, reported highest mean summer temperatures since the beginning of the data collection (21).

As seen in chapter 1, very high temperatures also affect evapotranspiration and consequently, irrigation requirements.

Table 15 sums up the Pearson correlation values for each crop.

Crops	R	Correlation Strength
Wheat	0.7678	strong
Citrus	0.5707	moderate
Maize	0.8254	very strong

Table 16 - Pearson Correlation Coefficients for Torino

There is no clear increasing behavior of evapotranspiration over a 9-year period by noting table 16.

Accordingly, to what previously written, April presents the highest increase percentage, with a mean value of 1.315%. June had the highest decrease with mean values of -0.252% for wheat.

The same pattern of increase values on the beginning of crop season and decrease reaching the end of the season is observed for citrus

It lowest mean of evapotranspiration over a 9-year period was on October -1.503%, eleven months into its season since it is a full year crop. On the other hand, the highest value was observed on February, right at the start of the year, with 2.283%.

Maize season start on April and finish on September. It highest mean percentage value over a 9-year period is on April (1.315%) and the lowest on September (-0.140%)

Graphs and 9-year mean values table for citrus and maize are listed in the annex.

4.3.2. Irrigation Requirements

Following São Paulo’s analysis, figure 21 has the time series of irrigation requirements for all 3 crops.

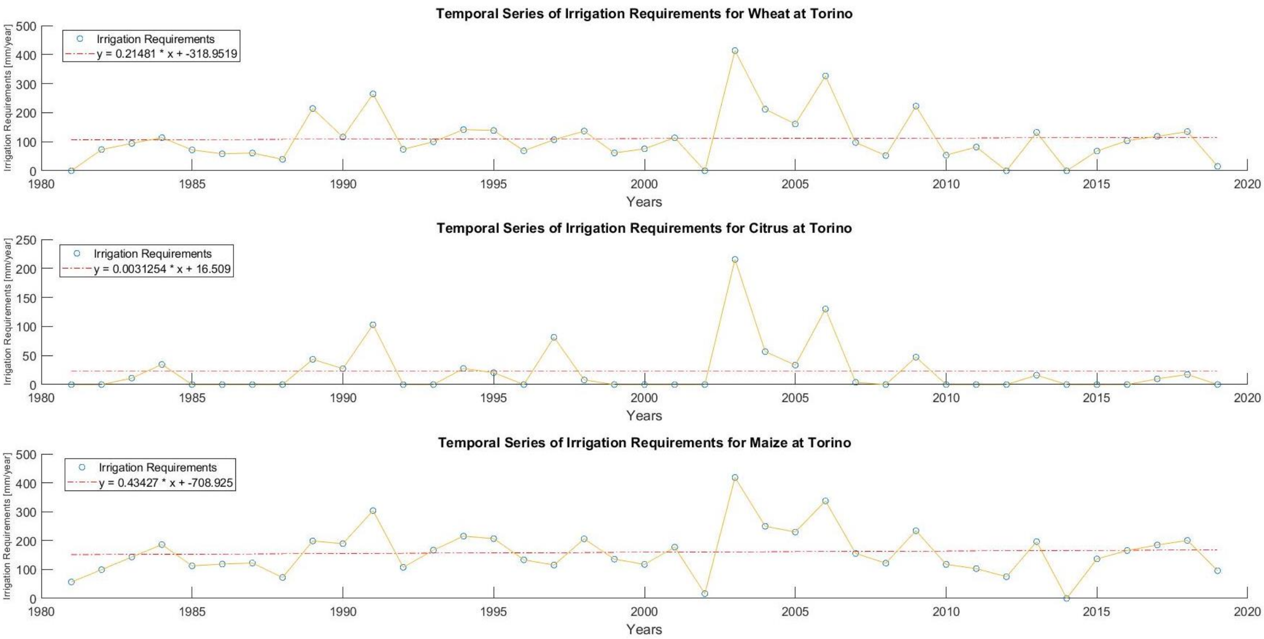


Figure 21 - Irrigation Requirements Time Series for Torino

Recalling what stated for the 2003 spike in temperature, there is a significant change in the time series after it.

Mean values of cumulated irrigation requirements for wheat before 2003 are around 96 mm/year. Mean values considering only from 2003 up to 2019 is 128 mm/year. It means a 33.62% difference in total cumulated irrigation requirements if the time series is split around that year.

For citrus and maize, the percentage increase is 92.22% and 22.10% respectively.

Cultivation in 2003 suffered from various problems, caused by the heat wave, not only due to the high temperatures but also from reduction on water availability (21)

This could mean that 2003 had an impact not only on that year, but years to follow.

By only analyzing the values of the angular coefficient of the regression line for all three plots, it seems that there is no temporal trend.

Also, the regression line itself do not follow any upwards or downwards trend, meaning that irrigation requirements does not appear to show any increase, or decrease, over the analyzed period.

In order to test linear dependency between time and irrigation requirements, a t-student test was applied, results are in table 17.

Since the parameters are the same as for the São Paulo test, $t_{lim} = 2.023$.

	Wheat	**Citrus**	**Maize**
Mean	110.675	22.760	159.601
Sxy	1061.200	15.440	2145.280
b1	0.003	0.434	0.215
b0	-318.952	16.508	-708.925
sig2	7823.700	1969.200	6773.500
T	0.171	0.005	0.371
$\boldsymbol{H_0}$ **Hypothesis** $(\boldsymbol{t_{lim} < T})$	Accepted	Accepted	Accepted

Table 17 - T Student Results for Irrigation Requirements at Torino

As expected, all crops passed the test. As to say that irrigation requirements, under the here stated conditions, does not show grow or decrease trends between 1981 and 2019.

Nevertheless, it is still important to take a closer look on the behavior of the days of necessary irrigation.

4.3.3. Days of Necessary Irrigation

A day with necessary irrigation requirement is when irrigation equipment for that day is greater than zero.

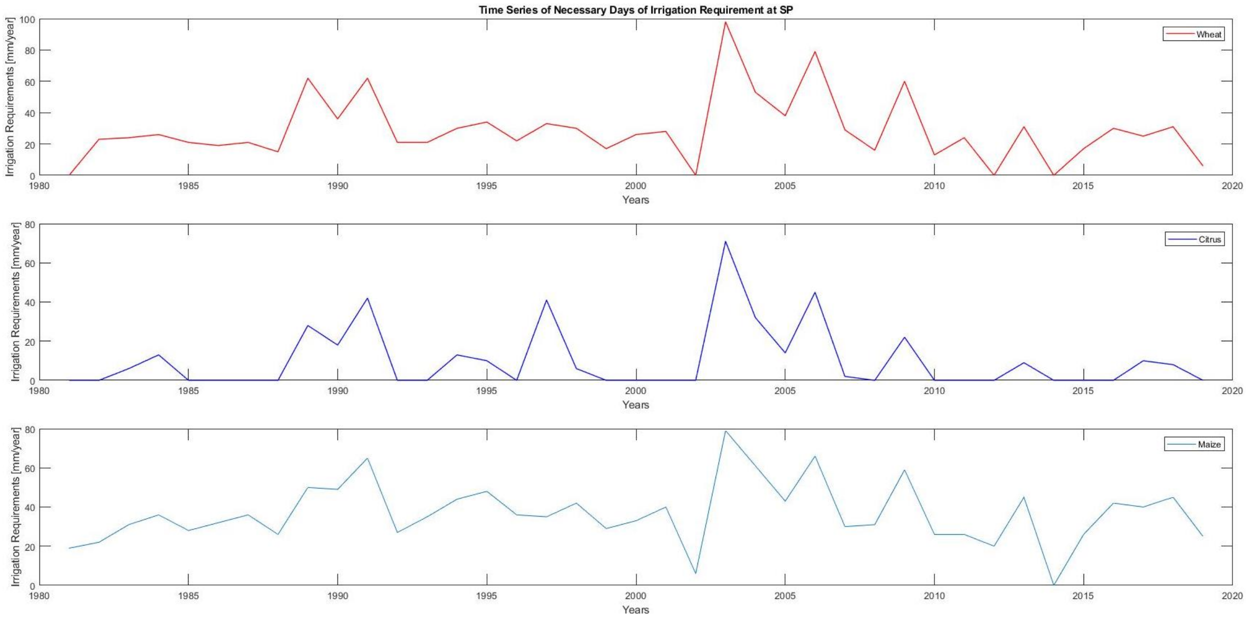


Figure 22 - Time Series of Days of Necessary Irrigation for Torino

2003 also represented a peak in days of necessary irrigation. The mean number of days of necessary irrigation, without 2003, for wheat is approximately 26, that year it was necessary 98 days of irrigation

Wheat season goes from April to September, 183 days. 26 days of irrigation represent 14.20% of the total season. In 2003 that percentage spiked at 53.56%, 39% more.

For citrus and maize this delta is 17.15% and 23.07% respectively.

Angular coefficients of time series for days of necessary irrigation do not show a temporal increase of it but, in order to have a statistical tool to verify it, another t-student test was applied.

	Wheat	Citrus	Maize
Mean	28.744	10	36.743
Sxy	71	-115	71
b1	-0.021	-0.023	0.0144
b0	70.444	56.558	7.998
sig2	445.440	275.117	253.363
T	-0.069	-0.098	0.063
$\boldsymbol{H_0}$ **Hypothesis** $(\boldsymbol{t_{lim} < T})$	Accepted	Accepted	Accepted

Table 18 - T Student Results for Days of Necessary Irrigation at Torino

As anticipated, none of the crops show a significant temporal trend for days of necessary irrigation.

It is also important to highlight some other aspects about each crop.

<u>Wheat</u>:

Only have days of necessary irrigation between April and September, being a seasonal crop.

August is the month that needed most irrigation, not only in mm but also in days, 61 mm/month, and 14 days of necessary irrigation on average. It represents an average of 58.57% of the total mm of irrigation and days of necessary irrigation.

The worst period in terms of mm of irrigation was August 2003 with 151.669 mm/month over 29 days, almost 5.3 mm/day of irrigation.

It also the worst period in terms of days of necessary irrigation, only two days did not need irrigation.

<u>Citrus:</u>

Although it is a full-year crop, irrigation requirement is needed between March and November.

Again, August is the critical month with, on average, 3.71 days of necessary irrigation and 8.86 mm/month. It represents 46.70% of total days of irrigation and 48.38% of total volumes of irrigation, meanly.

Differently from wheat, the worst period in terms of mm of irrigation and days of necessary irrigation was July 2006, with 93.01 mm/month over 28 days, almost 3.32 mm/day of irrigation.

<u>Maize:</u>

Maize is a seasonal crop, irrigation requirement is only needed between June and September.

If considering mm/month of irrigation, historically, July is the worst year with 73.50 mm/month of irrigation requirement. While looking at days of necessary irrigation, August have a higher value of 17.5 days of necessary irrigation against 14.4 of July.

July 2006 has the worst values for irrigation requirements with 188.28 mm/month and 29 days of necessary irrigation. It is not the worst period overall only because in July 2003 30 days of necessary irrigation were required.

While July 2006 had 6.49 mm /day of irrigation requirement, July 2003 presented 5.45 mm/day.

4.3.4. Distribution Fitting

With the same intentions of chapter 4.3.4, a chi squared test was applied in order to evaluate if cumulated yearly values of irrigation requirements fit one of three known distributions: normal, lognormal or Gumbel.

The procedure adopted was the same used for São Paulo and the distribution parameters are showed in table 19

Parameter	**Wheat**	**Citrus**	**Maize**
k	9	7	9
q	0.1111	0.1429	0.1111
E_i	4.3333	2.5714	4.3333
$\theta_{1,normal} = \mu$	110.6754	49.3129	159.6092
$\theta_{2,normal} = \sigma$	87.3143	53.8477	87.3143
$\theta_{1,lognormal} = \ln(\overline{x}) - 0.5 * [\ln\left(s^2/\overline{x}^2\right)]$	4.4646	3.5057	4.9572
$\theta_{2,lognormal} = \sqrt[2]{\ln\left(1 + s^2/\overline{x}^2\right)}$	0.6956	0.8860	0.4806
$\theta_{1,Gumbel} = \overline{x} - 0.5772 * s * (\sqrt[2]{6} * \pi)$	71.3491	25.0876	123.005
$\theta_{2,Gumbel} = s * (\sqrt[2]{6} * \pi)$	68.0790	41.9849	63.4379

Table 19 - Distribution Parameters for Torino

Thus, after diving and attributing the data sample into the calculated classes, the chi squared test was applied with results in table 20.

	Wheat	Citrus	Maize
dof	6	4	6
χ^2_{Normal}	12.002	12.333	10.153
χ^2_{lim}	12.591	9.488	12.591
$H_{0\ Normal}$	Accepted	Rejected	Accepted
$\chi^2_{LogNormal}$	10.153	1.445	6.002
χ^2_{lim}	12.591	9.488	12.591
$H_{0\ LogNormal}$	Accepted	Accepted	Accepted
χ^2_{Gumbel}	6.002	7.667	7.384
χ^2_{lim}	12.591	9.488	12.591

$H_{0\ Gumbel}$	Accepted	Accepted	Accepted

Table 20 - Chi Squared Results for Torino

4.3.5. Conclusion

In summary, none of the crops here analyzed presented any temporal trend when applying the t-student test.

Wheat and maize had positive results when applying the chi-squared test for normal, lognormal and Gumbel. Citrus, on the other hand, did not perceived a positive result only for the normal distribution.

5. Comparison and Conclusion

5.1. Comparison

This chapter will focus on the differences and similarities between crops at São Paulo and Torino.

The comparison analysis will evaluate evapotranspiration and irrigation requirements.

5.1.1. Wheat

Although situated in opposite side of the planet, wheat crop behavior for São Paulo and Torino have some similarities.

Both crops have 6 months seasons and they start with one month delay, Torino startin and finishing earlier (figure 23).

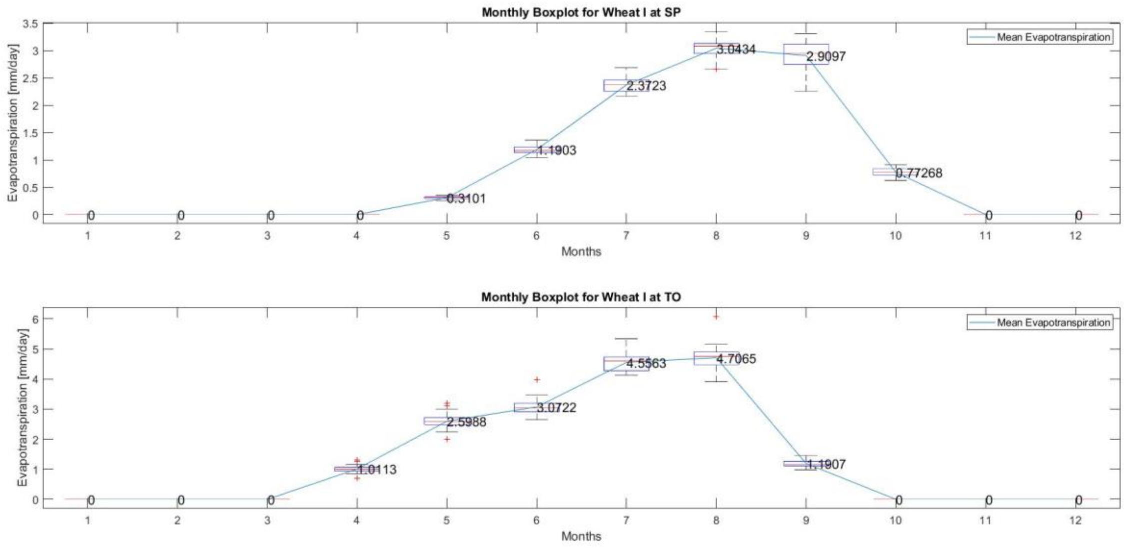


Figure 23 – Wheat Evapotranspiration Boxplot Comparison between Torino and São Paulo

Important to notice that wheat season for São Paulo happens during the winter, may be one of the reasons why mena daily evapotrasnpiration is lower.

Maximum mean values for both time series happen to be on August.

Torino presents higher number of outliers, as we previously stated 2003 was the heat wave year, thus evapotranspiration values spiked.

Table 21 lists overall cumulated yearly evapotranspiration

	Torino	São Paulo	Delta
Min	487.8	293.9	193.9
Max	629	353	276
Mean	531.2	328.6	202.6
Median	527.6	330.4	197.2
Mode	487.8	293.9	193.9

Standard Deviation	26.75	15.4	11.35

Table 21 - Wheat Evapotranspiration Comparison

Torino max value is on 2003 while São Paulo's on 2011.

Torino min value is on 1981 while São Paulo's on 1992.

Both cities presented strong correlation coefficients between evapotranspiration and irrigation requirements, being R equal to 0.866 for São Paulo and 0.768 for Torino.

Irrigation is not needed during the whole season as illustrated on chapter 2. Nevertheless, Torino has recorded values of irrigation from April to September, although the significant values are presented only during July to September.

On the other hand, São Paulo only needs 3 month of necessary irrigation, August to October, although October values are much lower.

Figure 24 illustrate the temporal variability of irrigation requirement for the two locations.

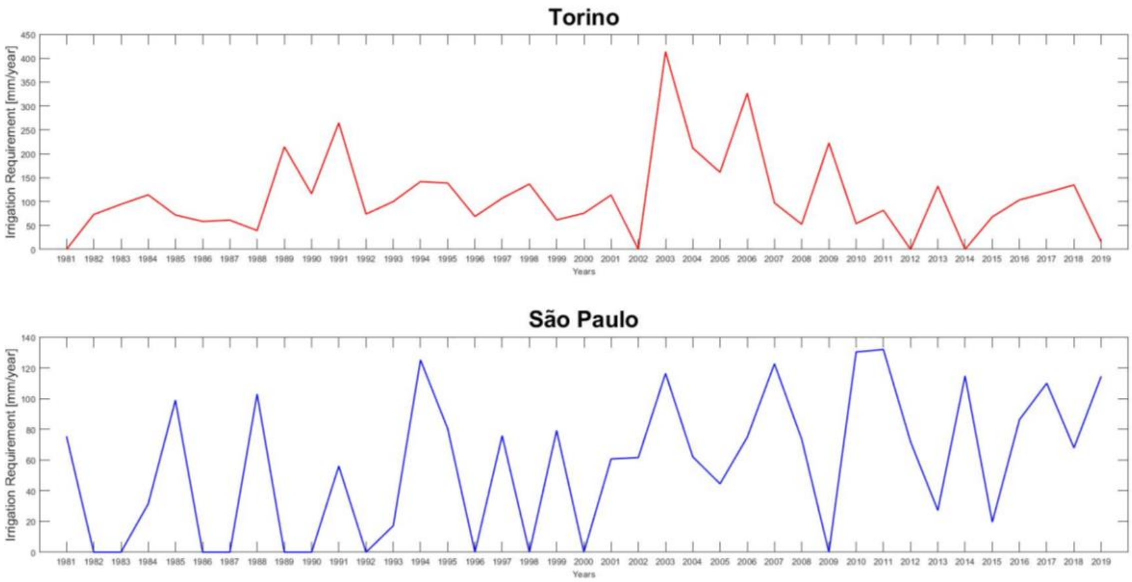


Figure 24 - Wheat Irrigation Requirement Time Series Comparison

As observed for evapotranspiration, yearly values of irrigation requirement for Torino are higher than on São Paulo. Mean values for Torino are around 110.7 mm/year as for São Paulo 57.25, 93% higher.

Although lower values, São Paulo demonstrate higher variability of its time series as the standard deviation is 46.1 with maximum value of 131.9. Torino has a standard deviation and maximum of 87.31 and 413.6 respectively.

Figure 25 shows that São Paulo has a higher variability than Torino, not only on cumulated yearly values but also on daily mean values. Torino data series shows a higher difference between the quartile ant its maximum values. By taking a closer look to data, the maximum values are related to 2003, the heatwave year.

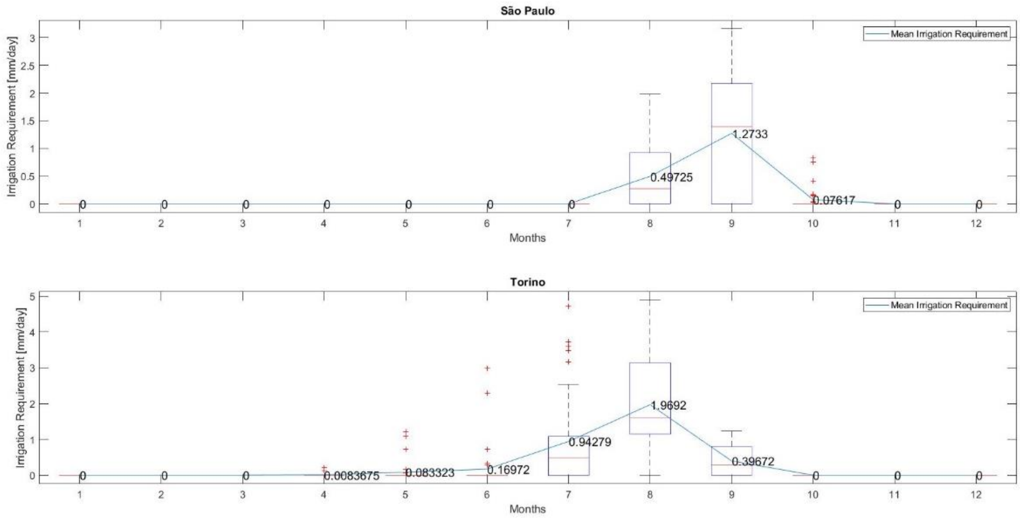


Figure 25 - Wheat Irrigation Requirement Boxplot

This behavior is also confirmed by the T-Student test performed on the trend of both series on chapters 4.2. and 4.3

	Torino	São Paulo
Irrigation Requirements	Accepted	Rejected
Days of Necessary Irrigation	Accepted	Rejected

Table 22 - Wheat T-Student Results

São Paulo rejected the results, meaning that irrigation requirements and days of necessary irrigation, show a significant linear increase. As years went by, both metrics show higher values.

Torino, on the other hand, did not showed any temporal trend, as intuited from figure 24.

On previous chapters, correlation between irrigation requirements and days of necessary irrigation was analyzed and presented high values for both Torino and São Paulo (0.9851 and 0.9872 respectively)

Figure 26 highlights that not only a correlation exists but temporal variability follows the same behavior.

It is interesting to notice how São Paulo's lines are more aligned, meaning that both metrics have a more equal comportment.

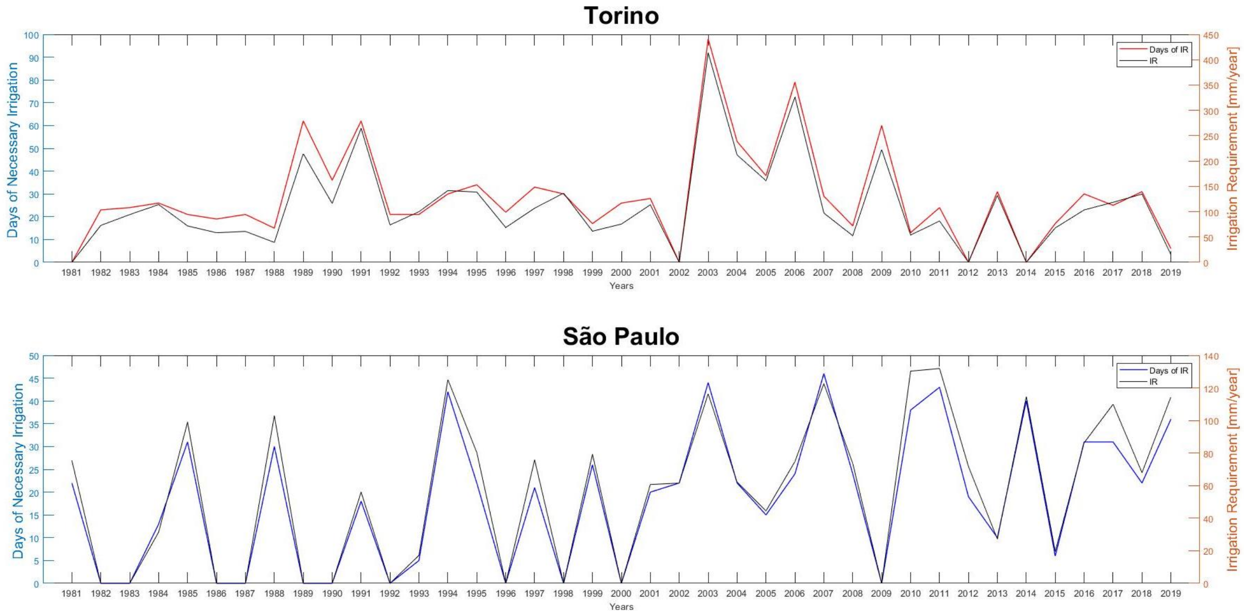


Figure 26 - IR and Days of IR Comparison for Wheat

Although both crops have the same length of season, Torino needs, on average, 55.02% more days of necessary irrigation (28.7 against 18.4)

The last step of the analysis for each city was the fitting of the irrigation statistics with appropriate probability distribution. Results for the chi-squared test of adaptation for both crops are listed in table 23

	Torino	São Paulo
Gumbell	rejected	accepted
Normal	accepted	accepted
LogNormal	accepted	accepted

Table 23 - Wheat Chi-Squared Results

5.1.2. Citrus

As previously commentted, citrus is a yearly crop, as shown in figure

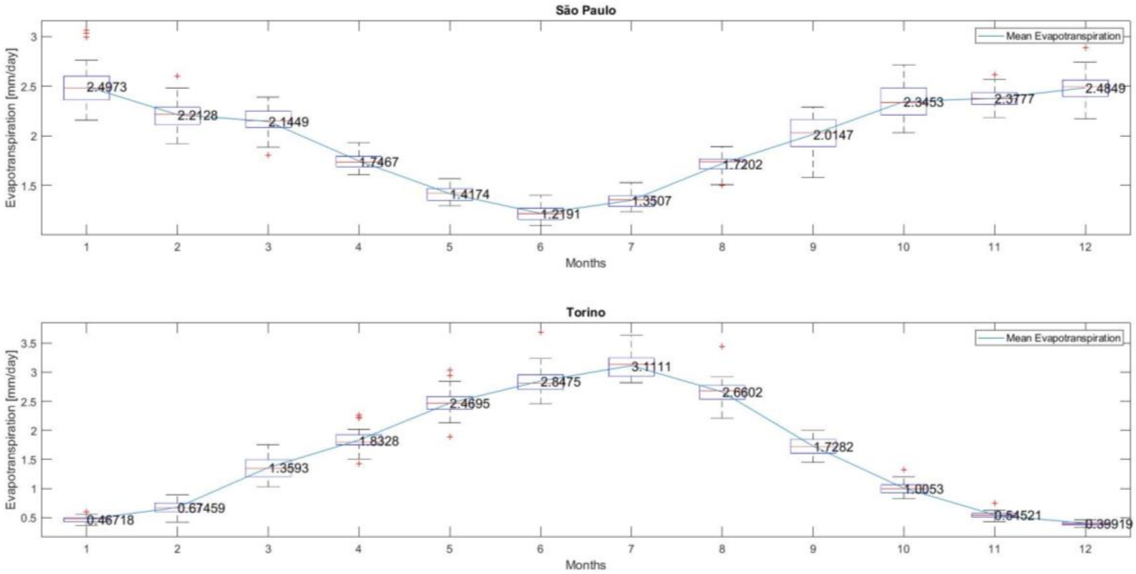


Figure 27 - Citrus Evapotranspiration Boxplot Comparison between Torino and São Paulo

Since location are situated on different hemispheres, yearly crops have the opposite behavior looking at months in a year.

Nonetheless, citrus demands higher values of evapotranspiration during summer for both cities. Torino’s evapotranspiration rises from June up to August, while São Paulo has its peaks from November to January.

Maximum and minimum pattern differ for both locations. Torino have higher values of mean daily evapotranspiration on July and lower on December. It means that Evapotranspiration have a higher annual variability at Torino, going from 3.11 mm/day to 0.39 mm/day.

While São Paulo have a softer comportment where its maximum and minimum values go from 2.49 mm/day to 1.22 mm/day.

São Paulo has a lower number of outliers since it mea and median are almost equal for each month. On the other hand, Torino present some months where media and mean do not have similar values (April: mean = 1.83 and median = 1.78).

Important to highlight also that the lowest months of evapotranspiration for Torino have very low variability (December and January). The same pattern is not observed for São Paulo as June and July boxes are quite large.

Table 24 lists overall cumulated yearly evapotranspiration

	Torino	**São Paulo**	**Delta**
Min	550.4	666.7	-116.3

Max	673.5	796.7	-123.2
Mean	592.1	729.5	-137.4
Median	590.5	726.3	-135.8
Mode	550.4	666.7	-116.3
Standard Deviation	26.49	25.79	0.7

Table 24 - Maize Evapotranspiration Comparison

Torino max value is on 2003 while São Paulo's on 2015.

Torino min value is on 1997 while São Paulo's on 1984.

Citrus was the crop that presented lowest values of correlation between evapotranspiration and irrigation requirements, being R equal to 0.609 for São Paulo and 0.571 for Torino.

Irrigation is not required throughout the season, as shown in Chapter 2. However, Turin recorded irrigation values from April to September, although significant values are only reported from July to September.

On the other hand, São Paulo only needs 3 months of necessary irrigation, from August to October, although the October values are significantly lower.

Figure 28 illustrate the temporal variability of irrigation requirement for the two locations.

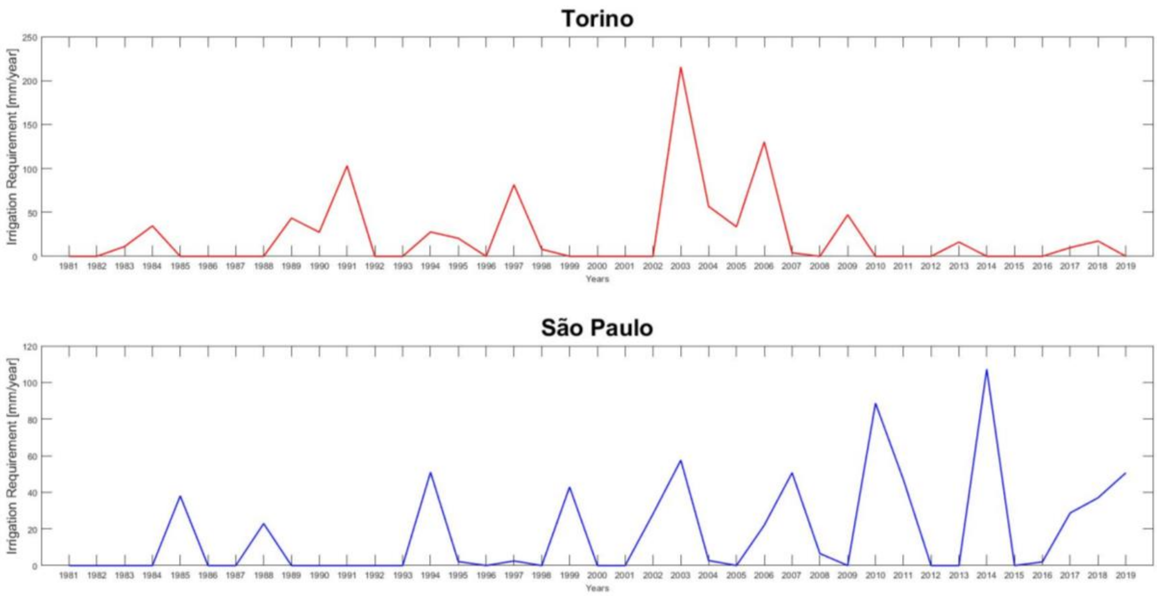


Figure 28 - Citrus Irrigation Requirements Time Series

Differently from evapotranspiration, yearly values of irrigation requirement for Torino are higher than on São Paulo. Mean values for Torino are around 22.76 mm/year as for São Paulo 17.65

Both series have similar values of years with zero mm of irrigation requirement, 20 for São Paulo and 21 for Torino. The first have zeros more distanced from one another, on the other hand the second have groups of years where irrigation is not needed.

Figure 29 shows how irrigation requirement behaves for both cities. São Paulo indicate a more concentrated need of irrigation around a couple of months (August to September). While Torino presents a less concentrated conduct over the year.

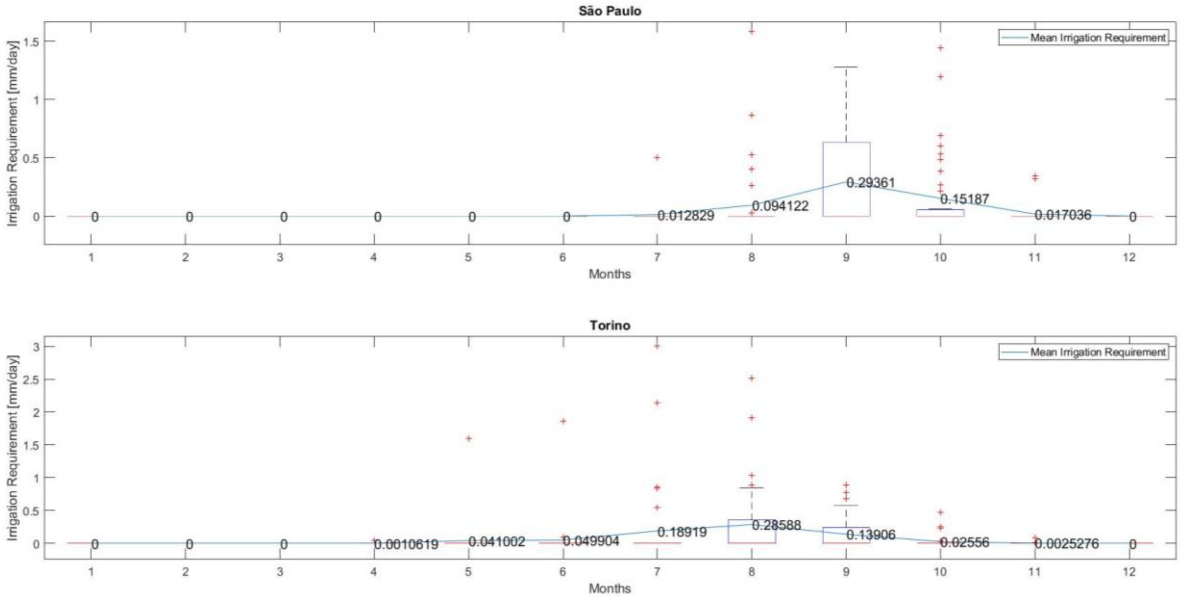


Figure 29 - Citrus Irrigation Requirement Boxplot

September is a critical month for São Paulo where most of its irrigation is accumulated.

	Torino	São Paulo
Irrigation Requirements	Accepted	Rejected
Days of Necessary Irrigation	Accepted	Rejected

Table 25 - Citrus T-Student Test Results

Following the same behavior as for wheat, São Paulo rejected the results, of the t-student test. Meaning it presents a temporal trend over the analyzed years.

By looking at figure 30, the intuition that Torino did not presented any temporal trend existed and the t-student test confirmed it as both metrics accepted the null hypothesis.

Also for citrus, correlation between irrigation requirements and days of necessary irrigation presented high values of Pearson coefficient R, meaning strong correlation between both metrics (R=0.9950 São Paulo and R=0.9729 for Torino).

Figure 30 highlights exactly what said for the correlation coefficients. São Paulo’s line are more overlaid than the ones on Torino’s.

To better undestand what effects this may bring, in 1989 there was 28 days of necessary irrigation for 43 mm/year. If the same proportion was to be applied for 1994, knowing the days of necessary irrigation (13), the predicted behavior would be of 19mm/year, instead in that year there was 27 mm/year of irrigation

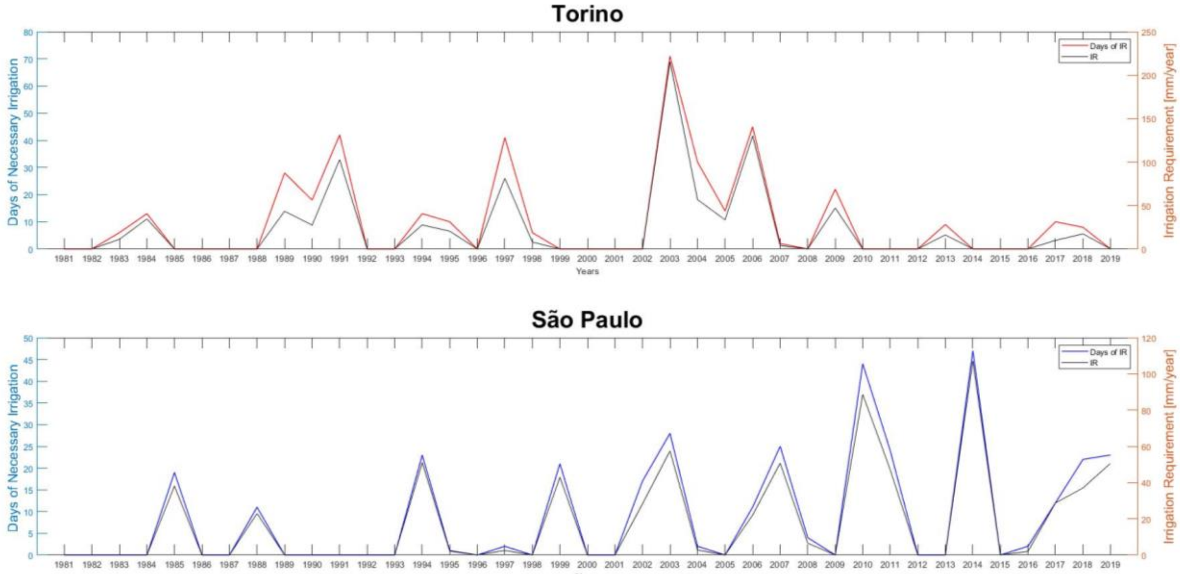


Figure 30 - IR and Days of IR Comparison for Citrus

Although both crops have the same length of season, Torino needs, on average, 55.02% more days of necessary irrigation (28.7 against 18.4)

Table 26 lists the Chi-Squared test results for citrus.

	Torino	São Paulo
Gumbell	rejected	accepted
Normal	rejected	accepted
LogNormal	accepted	accepted

Table 26 - Citrus Chi-Squared Test Results

5.1.3. Maize

Although diverse season than Wheat, also Maize is a seasonal crop.

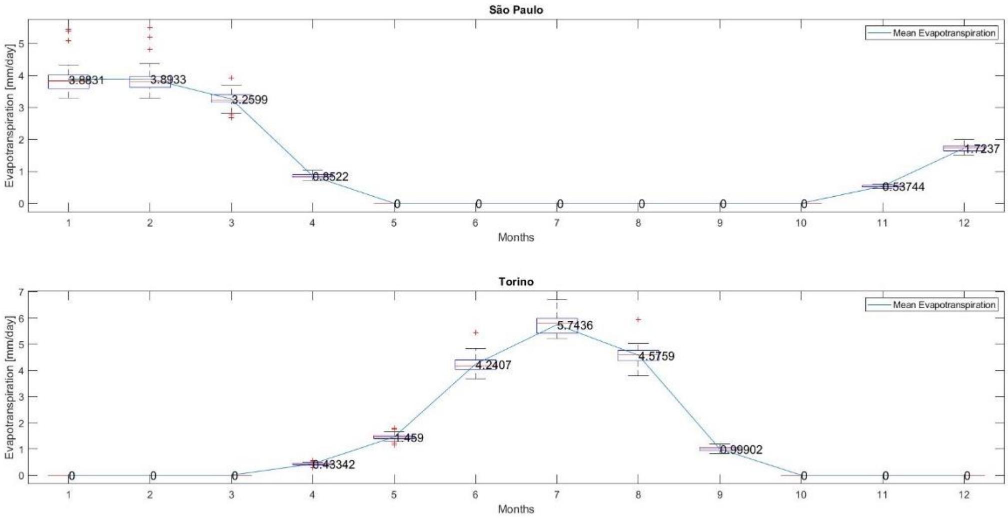


Figure 31 - Maize Evapotranspiration Boxplot Comparison between Torino and São Paulo

Six months' seasons for maize, cultivated during the summer

If a normalization of season is done considering the first month of the season as one and the last six. Torino's evapotranspiration has higher mean daily values for all months excluded the second (May).

Maximum and minimum pattern differ for both locations. Torino presents a more abrupt behavior going, with peak in the fourth month. It means that Evapotranspiration have a higher annual variability at Torino, with maximum and minimum of 5.74 mm/day and 0.43 mm/day respectively.

São Paulo also reaches its peak on the fourth month but the difference from the preceding and foreseeing month are much softer than the ones seen for Torino.

Both time series have the same number of outliers when analyzing mean daily evaporation.

Mid-season variability seems to be a pattern for both cities since 1st and 3rd quartiles for the 3 middle months are bigger than the ones ate the end of the season

Table 27 lists overall cumulated yearly evapotranspiration

	Torino	São Paulo	Delta
Min	496.5	400	96.5
Max	646.2	554.3	91.9
Mean	541	438.6	102.4
Median	535.2	434.4	100.8
Mode	496.5	400.5	96

Standard Deviation	28.71	30.3	-1.59

Table 27 - Maize Evapotranspiration Comparison

Torino max value is on 2003 while São Paulo’s on 2015.

Torino min value is on 1981 while São Paulo’s on 1984.

Maize is the crop that showed the highest values of correlation for evaporation and irrigation requirements, with 0.844 for São Paulo and 0.825 for Torino. Both classified as very strong correlated

Nevertheless, Torino has recorded values of irrigation from June to September, although the significant values are presented only during July to September.

On the other hand, São Paulo only needs 4 month of necessary irrigation, although very low amounts of irrigation were recorded.

Figure 35 illustrate the temporal variability of irrigation requirement for the two locations.

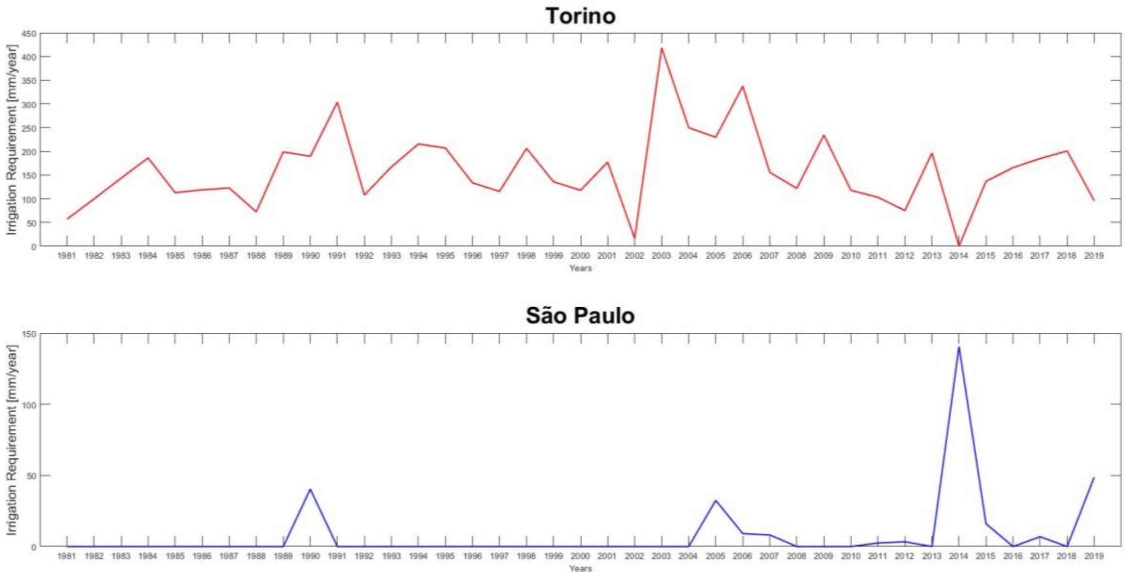


Figure 32 - Maize Irrigation Requirements Time Series

74% of the years inside São Paulo distribution did not need any mm of irrigation. While for Torino only 1 year did not need irrigation, 2014.

Torino have a mean values of 159.61 mm/year of irrigation, with maximum and minimum (excluding zero) going from 418.25 mm/year on 2003 to 16 mm/year on 2002

Figure 36 shows how irrigation requirement behaves for both cities. São Paulo boxplot only points out the low need of irrigation and how its distributed over a few months.

Nevertheless, Torino have peaks months of irrigation requirements on July and August.

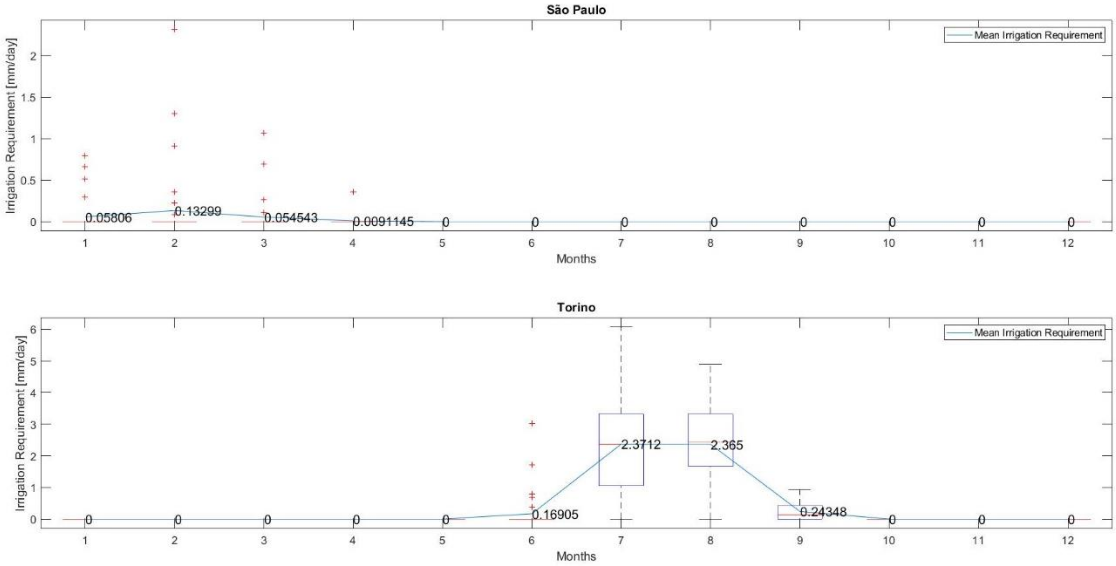


Figure 33 - Citrus Irrigation Requirement Boxplot

Results of the T-Student test are shown in table 28.

	Torino	São Paulo
Irrigation Requirements	Accepted	Accepted
Days of Necessary Irrigation	Accepted	Accepted

Table 28 - Maize T-Student Test Results

Contrarily to previous crops, maize was the only one that did not show any temporal trend.

Correlation between irrigation requirements and days of necessary irrigation for maize express great values for both Torino and São Paulo (0.960 and 0.994 respectively)

São Paulo almost perfect overlaid is due to its strong correlation and high amount of zeros.

Torino behavior is far from overlaid, days of necessary irrigation seem to have a poorer reaction to increase values of irrigation requirement.

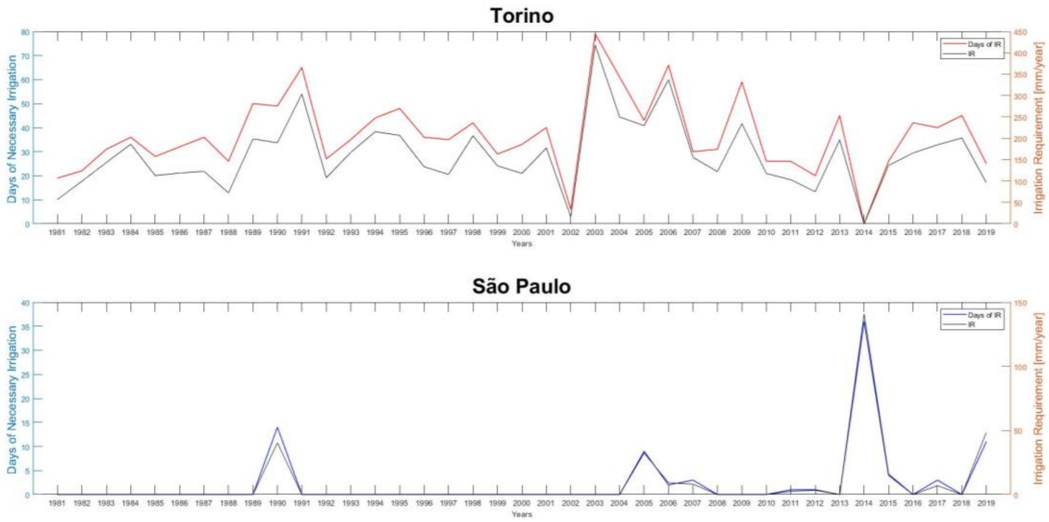


Figure 34 - R and Days of IR Comparison for Maize

Torino needs an average of 36 days of necessary irrigation, with maximum of 79 on 2003 and minimum (without zero) of 6 on 2002.

Table 29 lists the Chi-Squared test results for citrus

	Torino	**São Paulo**
Gumbell	rejected	rejected
Normal	accepted	rejected
LogNormal	accepted	accepted

Table 29 - Maize Chi-Squared Test Results

5.2. Conclusion

The only crops that showed temporal trend between the analyzed years were wheat and citrus for São Paulo, as summarized in table 30.

São Paulo	Angular Coeff	T	H0 Hypothesis
Wheat	1.7568	2.9345	Rejected
Citrus	0.9354	2.6088	Rejected
Maize	0.6452	1.9129	Accepted

Torino	Angular Coeff	T	H0 Hypothesis
Wheat	0.2148	0.1707	Accepted
Citrus	0.0031	0.0050	Accepted
Maize	0.4343	0.3709	Accepted

Table 30 - Final T-Student Results

Besides, the only distribution that fitted well all crops for both cities was LogNormal as listed in table 31

São Paulo	Gumbel	Normal	LogNormal
Wheat	Accepted	Accepted	Accepted
Citrus	Accepted	Accepted	Accepted
Maize	Rejected	Rejected	Accepted

Torino	Gumbel	Normal	LogNormal
Wheat	Rejected	Accepted	Accepted
Citrus	Rejected	Rejected	Accepted
Maize	Rejected	Accepted	Accepted

Table 31 - Final Chi-Squared Results

Comparing locations, Torino showed higher results of irrigation requirements for all three crops.

Figures 35 and 36 are examples of the Empirical Distribution Function and probability plots for wheat, at São Paulo and Torino respectively

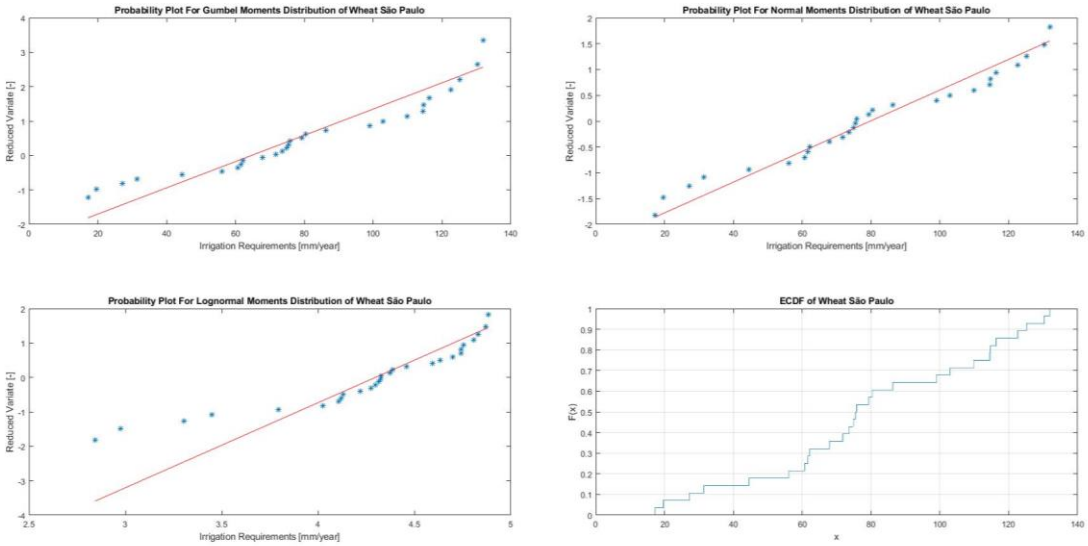


Figure 35 - Probability Plots and ECDF for Wheat at São Paulo

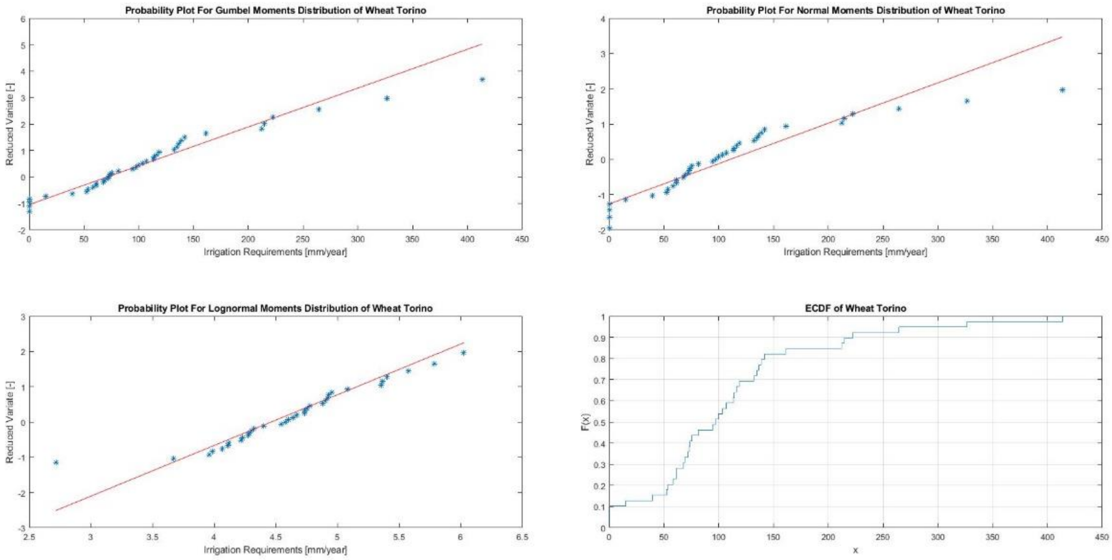


Figure 36 - Probability Plots and ECDF for Wheat at Torino

Results for citrus and maize are illustrated in annex

The conclusion is that the statistical analysis seen for floods can give useful results in assessing the frequency (or rarity) of drought events, which are becoming more relevant in current days.

6. References

1. **Bank, The World.** The World Bank. [Online] 05 2020. [Cited: 10 16, 2021.] https://www.worldbank.org/en/topic/water-in-agriculture..

2. **FAO.** FAO. *FAO.* [Online] 2021. [Cited: 10 16, 2021.] http://www.fao.org/about/en/.

3. **FAO, Richard G. ALLEN, Luis S. PEREIRA, Dirk RAES, Martin SMITH.** *FAO Irrigation and Drainage Paper No. 56.* Rome : FAO, 1998.

4. *Green-blue water accounting in a soil water balance.* **Y.Hoekstra, Arjen.** 2019, Science Direct, pp. 112-117.

5. **FAO.** *FAO Expert Consultation on Revision of FAO Methodologies for Crop Water.* 1998.

6. **ECMWF.** *ECMWF.* [Online] 2021. [Cited: 09 12, 2021.] https://www.ecmwf.int/en/about/who-we-are.

7. **Climate Copernicus.** *Climate Data Store.* [Online] 2021. [Cited: 10 5, 2021.] https://climate.copernicus.eu/climate-data-store.

8. **ECMWF.** Climate Copernicus. *ERA 5.* [Online] 2021. [Cited: 08 14, 2021.] https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=overview.

9. **Schulzweida, Uwe.** *CDO User Guide.* s.l. : MPI for Meteorology, 2020.

10. *Severe Weather Caused by Heat Island and Sea Breeze Effects in the Metropolitan Area of São Paulo, Brazil.* **Felipe Vemado, Augusto José Pereira Filho.** 2016, Advances in Meteorology, p. 13.

11. *Urban heat islands effects on Torino.* **Massimo, Milelli.** 2019, COSMO.

12. *Changing Trends in Rainfall Extremes in the Metropolitan Area of São Paulo: Causes and Impacts.* **Marengo Jose A., Ambrizzi Tercio, Alves Lincoln M., Barreto Naurinete J. C., Simões Reboita Michelle, Ramos Andrea M.** 2020, Frontiers in Climate , p. 3.

13. *São Paulo drought: trends in streamflow and their relationship to climate and human-induced change in Cantareira watershed, Southeast Brazil.* **Freitas, Gabriela Nicoleti de.** 2020, Hydrology Research, p. 750.

14. **Hans Hersbach, Bill Bell, Paul Berrisford, András Horányi, Joaquín Muñoz Sabater, Julien Nicolas, Raluca Radu, Dinand Schepers, Adrian Simmons, Cornel Soci, Dick Dee.** *Global reanalysis: goodbye ERA-Interim, hello ERA5.* s.l. : ECWMF, 2019. Newsletter.

15. *ERA5-based global assessment of irrigation.* **Matteo Rolle, Stefania Tamea, Pierluigi Claps.** 2021, PLOS ONE, p. 21.

16. *Correlation Coefficients: Appropriate Use and Interpretation.* **Schober Patrick, Boer Christa, Schwarte, Lothar.** 02 01, 2018, Anesthesia & Analgesia.

17. **TDV, Swinscow.** *Statistics at square one.* s.l. : BMJ Publishing Group, 1997. Vol. 9. 9.

18. *The probable error of a mean.* **Student.** s.l. : JSTOR, 1908, Biometrika, p. 25.

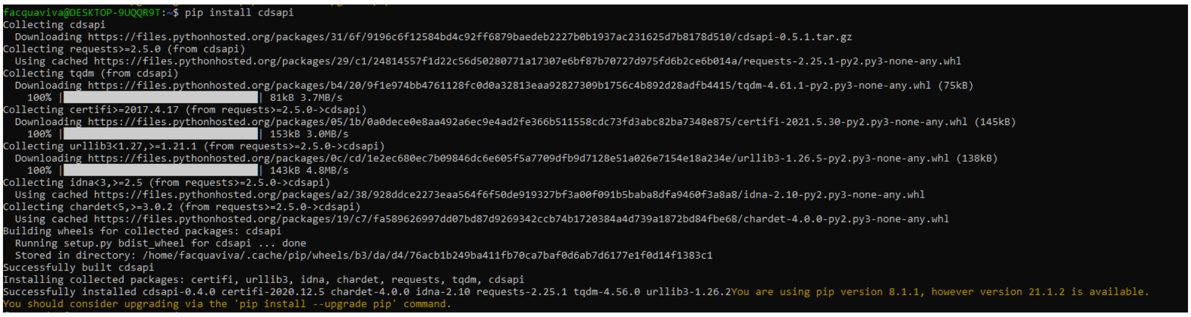
19. *On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling.* **Peason, Karl.** 1900, he London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, pp. 157-75.

20. **Grimaldi, Salvatore, Kao Shih-Chieh, Castellarin Attilio, Papalexiou Simon Michael, Viglione Alberto, Laio Francesco, Aksoy Hafzullah, Gedikli, Abdullah.** Statistical Hydrology. 2011.

21. *Characteristics of the Summer 2003 Heat Wave in Piedmont, Italy, and its Effects on Water Resources.* **Cassardo Claudio, Mercalli Luca, Berro Daniele.** 43, 08 31, 2007, Journal of the Korean Meteorological Society, p. 195.

22. ***MIcrobeNotes.*** **[Online] [Cited: 10 19, 2021.] https://microbenotes.com/chi-square-test/.**

7. Annexes



```
facquaviva@DESKTOP-9UQQR9T:~$ pip install cdsapi
Collecting cdsapi
  Downloading https://files.pythonhosted.org/packages/31/6f/9196c6f12584bd4c92ff6879baedeb2227b0b1937ac231625d7b8178d510/cdsapi-0.5.1.tar.gz
Collecting requests>=2.5.0 (from cdsapi)
  Using cached https://files.pythonhosted.org/packages/29/c1/24814557f1d22c56d50280771a17307e6bf87b70727d975fd6b2ce6b014a/requests-2.25.1-py2.py3-none-any.whl
Collecting tqdm (from cdsapi)
  Downloading https://files.pythonhosted.org/packages/b4/20/9f1e974bb4761128fc0d0a32813eaa92827309b1756c4b892d28adfb4415/tqdm-4.61.1-py2.py3-none-any.whl (75kB)
    100% |████████████████████████████████| 81kB 3.7MB/s
Collecting certifi>=2017.4.17 (from requests>=2.5.0->cdsapi)
  Downloading https://files.pythonhosted.org/packages/05/1b/0a0dece0e8aa492a6ec9e4ad2fe366b511558cdc73fd3abc82ba7348e875/certifi-2021.5.30-py2.py3-none-any.whl (145kB)
    100% |████████████████████████████████| 153kB 3.0MB/s
Collecting urllib3<1.27,>=1.21.1 (from requests>=2.5.0->cdsapi)
  Downloading https://files.pythonhosted.org/packages/0c/cd/1e2ec680ec7b09846dc6e605f5a7709dfb9d7128e51a026e7154e18a234e/urllib3-1.26.5-py2.py3-none-any.whl (138kB)
    100% |████████████████████████████████| 143kB 4.8MB/s
Collecting idna<3,>=2.5 (from requests>=2.5.0->cdsapi)
  Using cached https://files.pythonhosted.org/packages/a2/38/928ddce2273eaa564f6f50de919327bf3a00f091b5baba8dfa9460f3a8a8/idna-2.10-py2.py3-none-any.whl
Collecting chardet<5,>=3.0.2 (from requests>=2.5.0->cdsapi)
  Using cached https://files.pythonhosted.org/packages/19/c7/fa589626997dd07bd87d9269342ccb74b1720384a4d739a1872bd84fbe68/chardet-4.0.0-py2.py3-none-any.whl
Building wheels for collected packages: cdsapi
  Running setup.py bdist_wheel for cdsapi ... done
  Stored in directory: /home/facquaviva/.cache/pip/wheels/b3/da/d4/76acb1b249ba411fb70ca7baf0d6ab7d6177e1f0d14f1383c1
Successfully built cdsapi
Installing collected packages: certifi, urllib3, idna, chardet, requests, tqdm, cdsapi
Successfully installed cdsapi-0.4.0 certifi-2020.12.5 chardet-4.0.0 idna-2.10 requests-2.25.1 tqdm-4.56.0 urllib3-1.26.2You are using pip version 8.1.1, however version 21.1.2 is available.
You should consider upgrading via the 'pip install --upgrade pip' command.
```

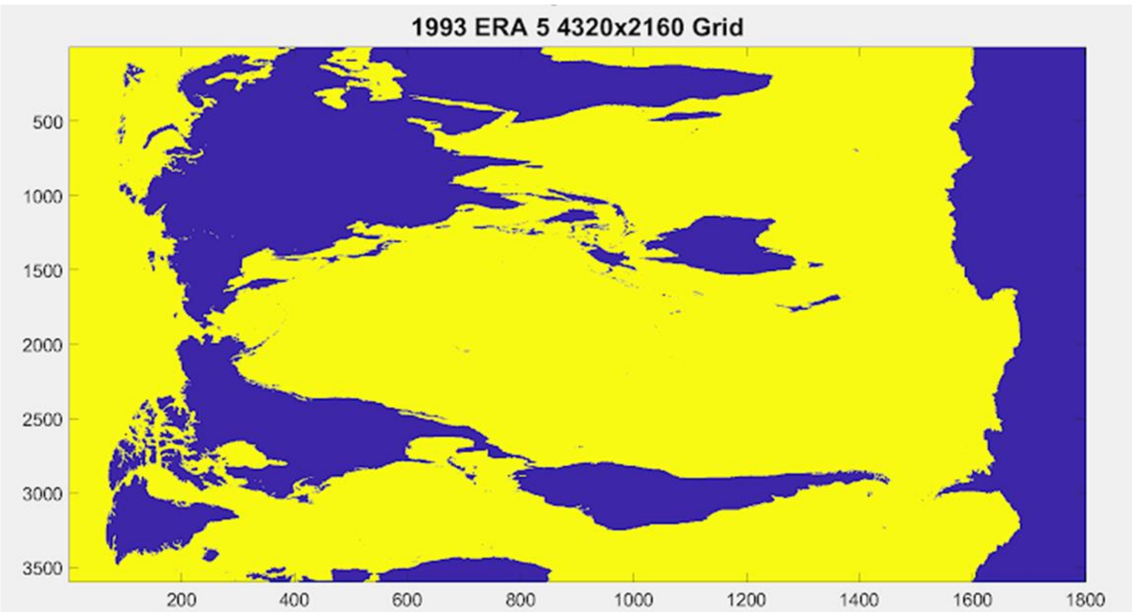
Annex 1 - CDS API Client Install

```
    import cdsapi

c = cdsapi.Client()

c.retrieve(
    'reanalysis-era5-single-levels',
    {
        'product_type': 'reanalysis',
        'variable': 'total_precipitation',
        'year': ['2006'],
        'month': [
            '01', '02', '03',
            '04', '05', '06',
            '07', '08', '09',
            '10', '11', '12',],
        'day': ['01', '02', '03',
            '04', '05', '06',
            '07', '08', '09',
            '10', '11', '12',
            '13', '14', '15',
            '16', '17', '18',
            '19', '20', '21',
            '22', '23', '24',
            '25', '26', '27',
            '28', '29', '30',
            '31',
        ],
        'time': [
            '00:00','01:00','02:00','03:00','04:00','05:00',
            '06:00','07:00','08:00','09:00','10:00','11:00',
            '12:00','13:00','14:00','15:00','16:00','17:00',
            '18:00','19:00','20:00','21:00','22:00','23:00'
        ],
        'grid': '0.08333/0.08333',
        'area': '89.959/-179.95833/-89.959/179.95833',
        'format': 'netcdf',
    },
    '2006.nc')
```

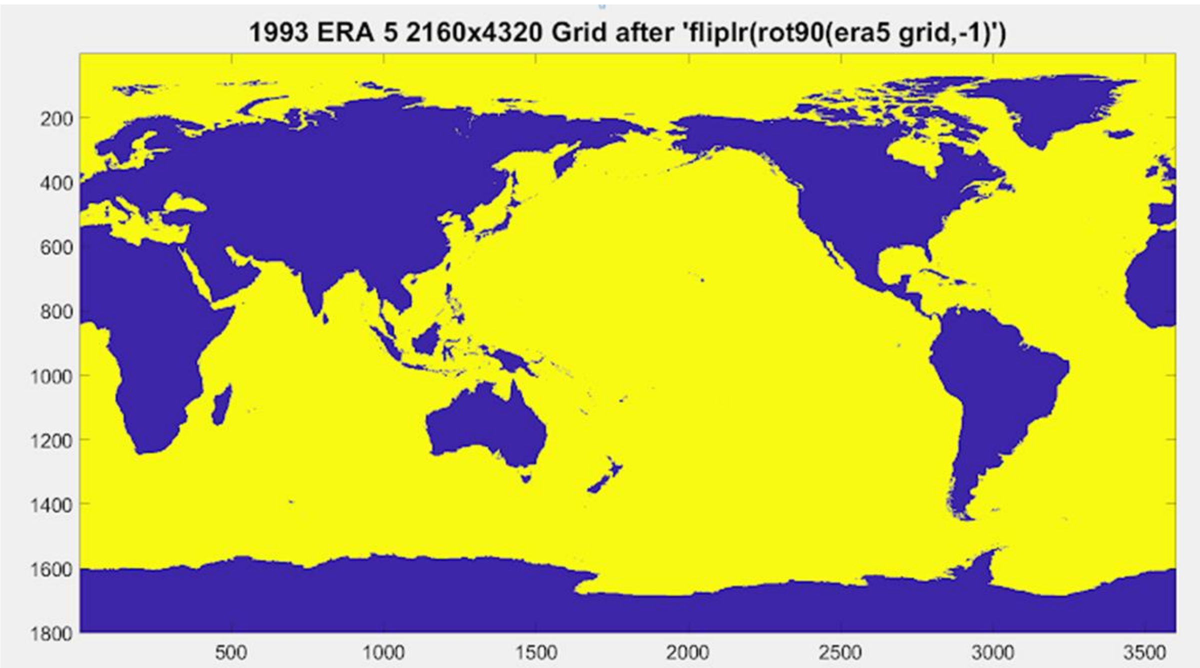
Annex 2 - Python Script Example for Data Precipitation Download

```
-1 :       Date     Time   Level Gridsize    Miss :     Minimum       Mean     Maximum : Parameter ID
 1 : 2000-01-01 00:00:00       0  9331200       0 :      0.0000 8.5702e-05    0.020354 : -1
 2 : 2000-01-01 01:00:00       0  9331200       0 :      0.0000 8.5202e-05    0.019808 : -1
 3 : 2000-01-01 02:00:00       0  9331200       0 :      0.0000 8.5392e-05    0.016553 : -1
 4 : 2000-01-01 03:00:00       0  9331200       0 :      0.0000 8.6504e-05    0.015731 : -1
 5 : 2000-01-01 04:00:00       0  9331200       0 :      0.0000 8.7146e-05    0.016738 : -1
 6 : 2000-01-01 05:00:00       0  9331200       0 :      0.0000 8.8831e-05    0.020072 : -1
 7 : 2000-01-01 06:00:00       0  9331200       0 :      0.0000 9.0158e-05    0.020459 : -1
```

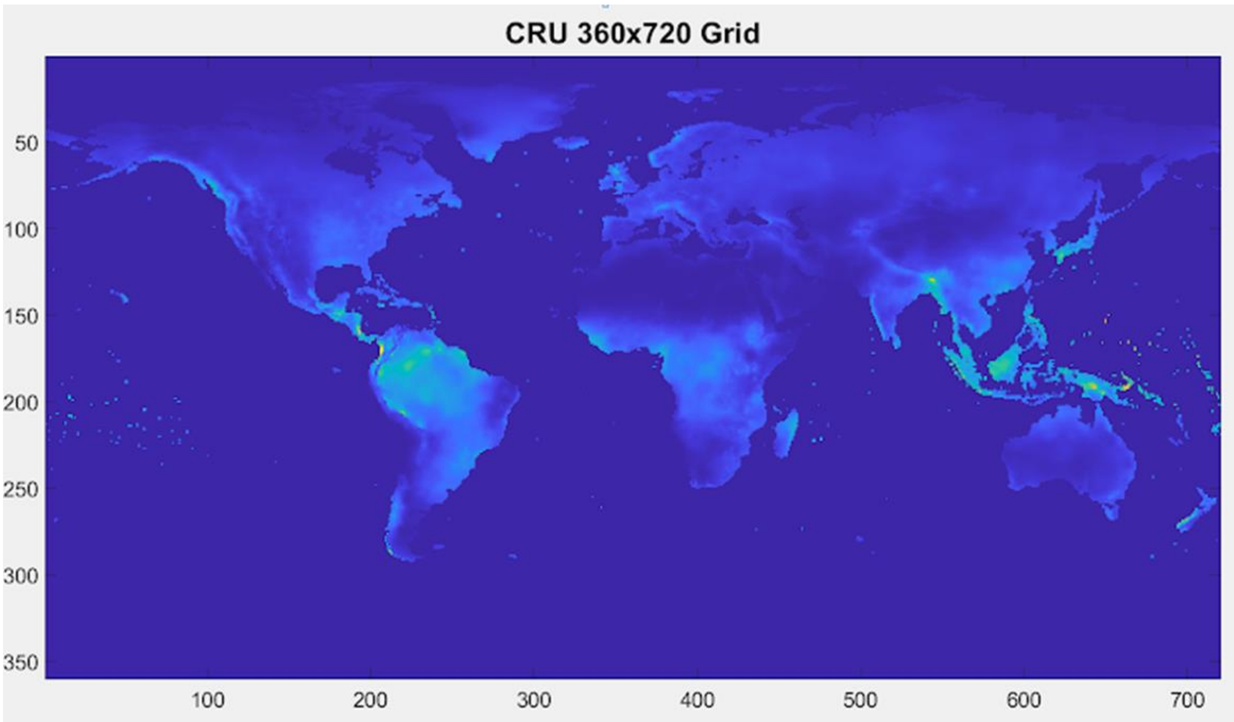
Annex 3 - CDO Command Usage



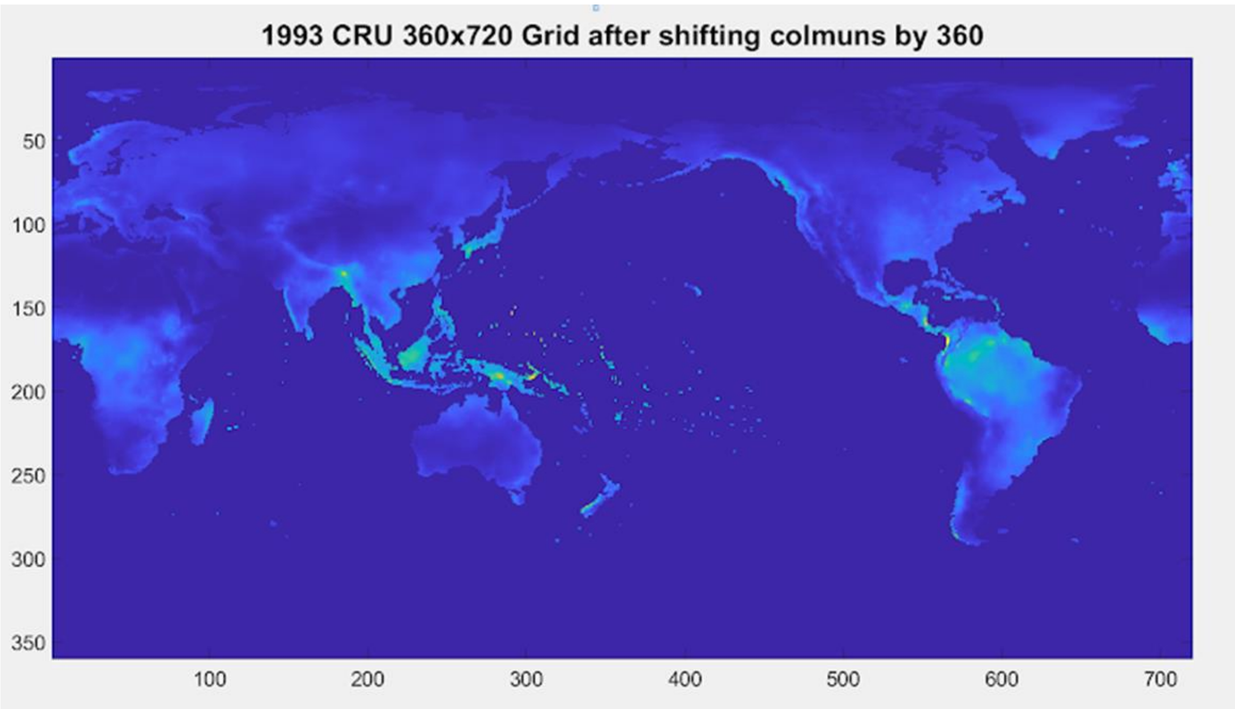
Annex 4 - ERA5 Dataset prior to Manipulation



Annex 5 - ERA5 Dataset after Necessary Manipulation



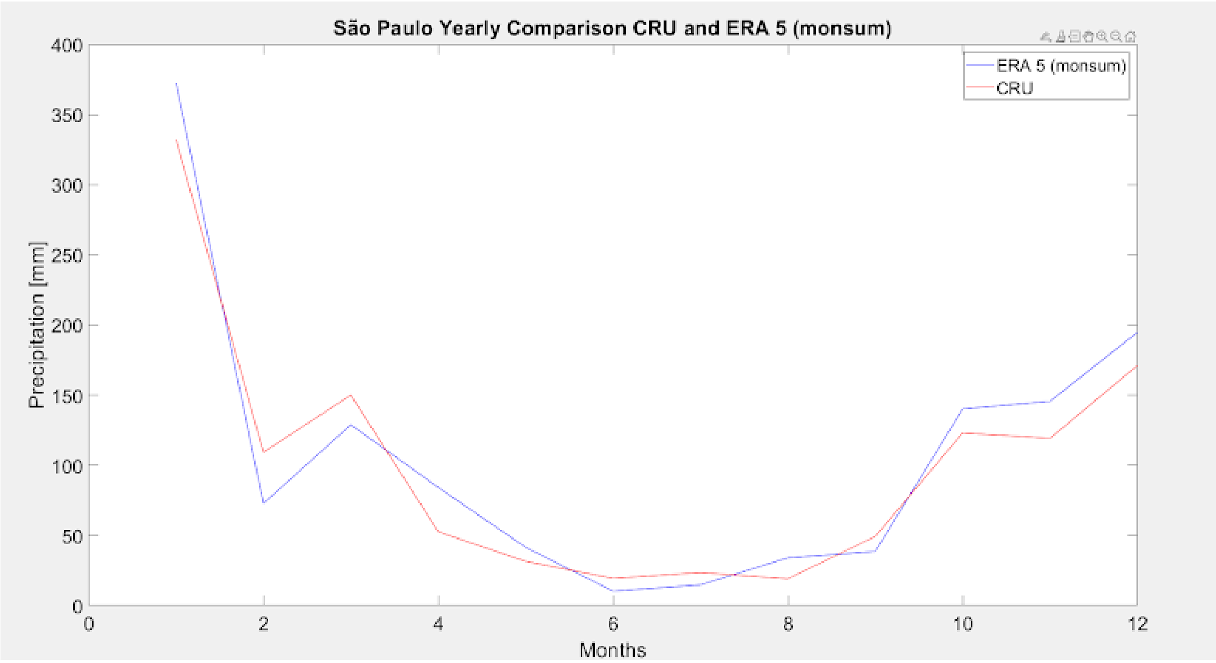
Annex 6 -CRU Dataset prior to Manipulation



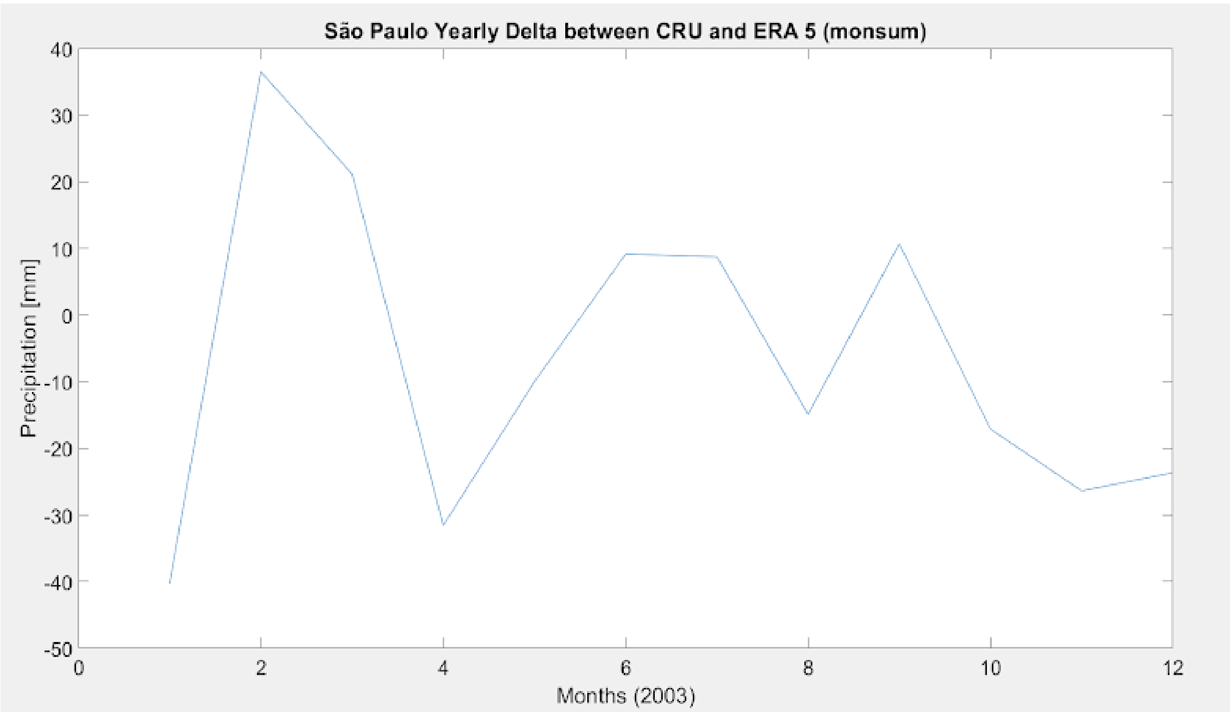
Annex 7 - CRU Dataset after Necessary Manipulation

```
%% Torino Verifications

torino_cru_1993 = sum(cru_1993(90,375,:),[1,2]); %%Select Torino Cell in New CRU Matrix (90,375 old one)
S = size(torino_cru_1993);
final_torino_cru_1993 = reshape(torino_cru_1993,[S(1)*S(2),S(3)]);

torino_era5_1993 = sum(era_5_360x720(90,375,:).*1000,[1,2]); %%Select Torino Cell in ERA 5 Matrix
S = size(torino_era5_1993);
final_torino_era5_1993 = reshape(torino_era5_1993,[S(1)*S(2),S(3)]);

delta_torino = final_torino_cru_1993 - final_torino_era5_1993;
```

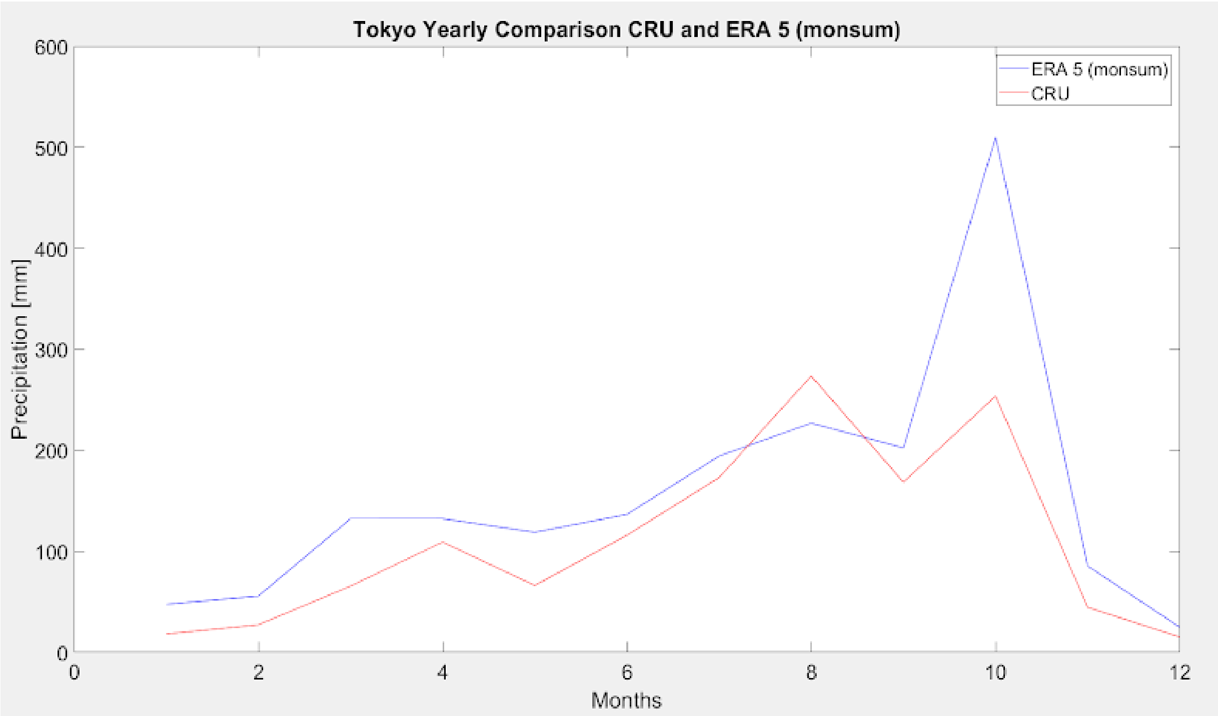
Annex 8 - ERA5 and CRU Script Example for Comparison



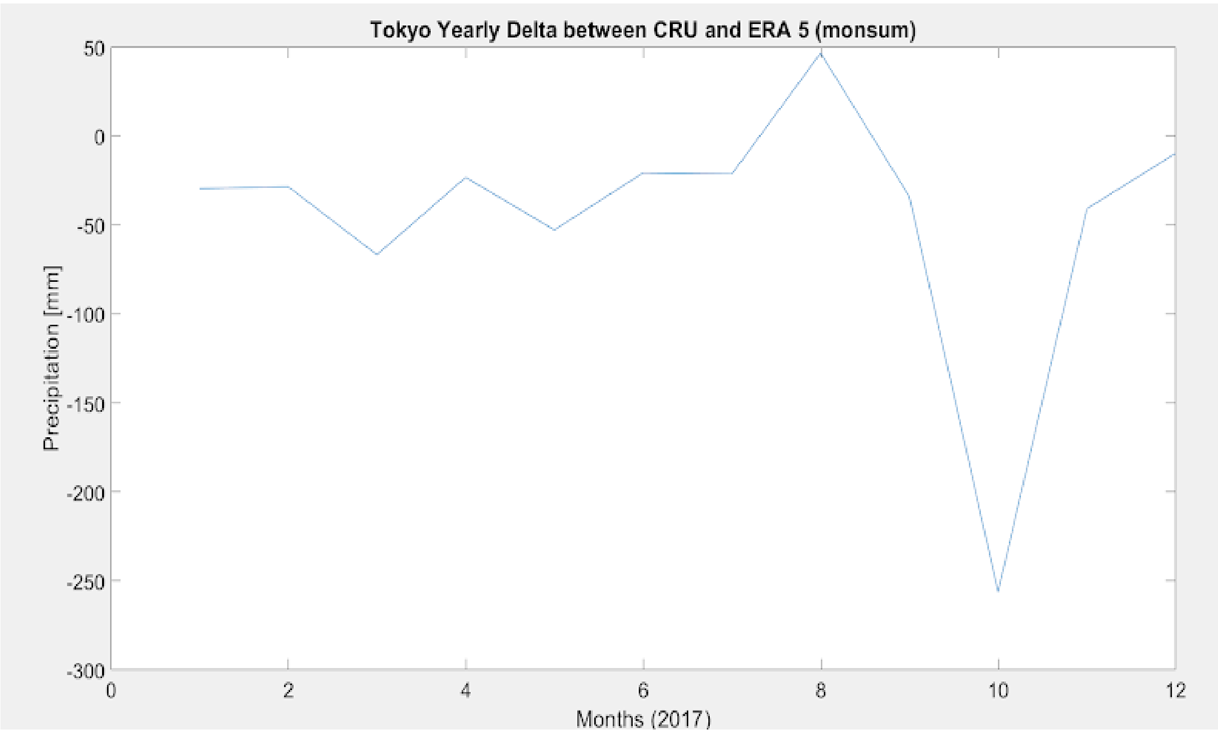
Annex 9 - ERA 5 vs CRU Comparison at São Paulo on 2003



Annex 10 - ERA 5 vs CRU Delta at São Paulo on 2003



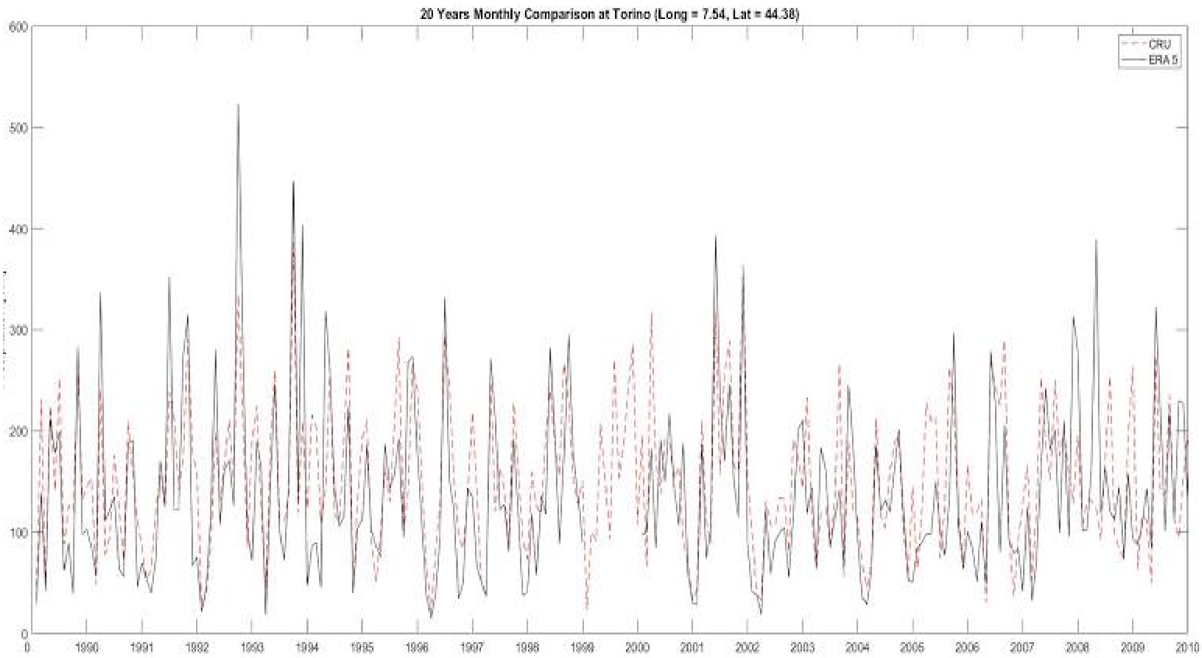
Annex 11 – ERA 5 vs CRU Comparison at Tokyo on 2017



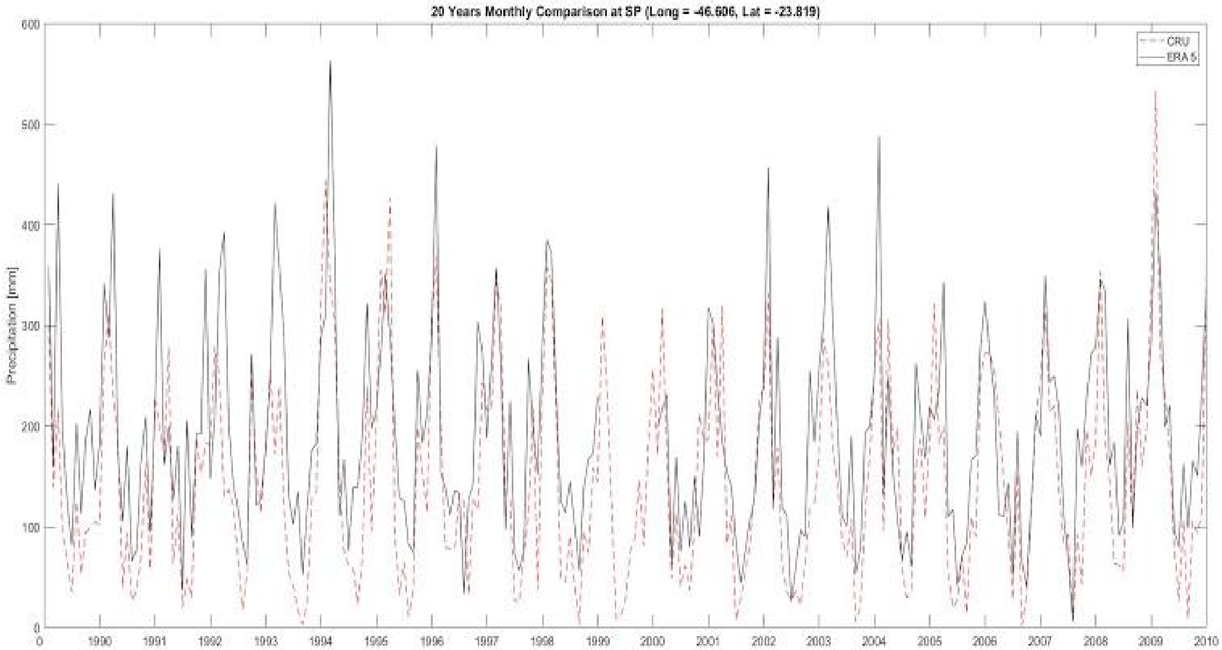
Annex 12 - ERA 5 vs CRU Delta at Tokyo on 2017

```matlab
clear all
%% Declaration
torino_long = 7.54;
torino_lat = 44.38;
A_TO = 180 + torino_long; %Pass from a -180/+180 Grid to a 0/360
Y_ERA5_TO =  round(A_TO * 4320/360); %360x180 Grido to 4320x2160
X_ERA5_TO = round(torino_lat *2160/180); %360x180 Grid to 4320x2160 (Don't need the extra step because we are lat don't change)
Y_CRU_TO = round(A_TO * 720/360); %360x180 Grid to 4320x2160
X_CRU_TO = round(torino_lat *360/180); %360x180 Grid to 4320x2160

months = [1:31 , 1:28, 1:31, 1:30, 1:31, 1:30, 1:31, 1:31, 1:30, 1:31, 1:30, 1:31];
kk=0;

%% Load CRU FIles
for i = 1:21
    for j =1:12
        years = 1989 + i;
        year = num2str(years);
        month = num2str(j);
        import_data(:,:,j) = rot90(fliplr(importdata(['C:\Users\User\Documents\MATLAB\torino_all_years\',year,'_',month,'_P_mmmonth.mat'])),1);
        torino_select_CRU(j,i) = import_data(Y_CRU_TO,X_CRU_TO,j);
    end
end
cru = reshape(torino_select_CRU,1,[]);

%% Import ERA 5 Values and Select Torino Cell (6x6)
for j = 1990:2010
    days = 2;
    year = num2str(j);
    open_day = (['day_' year '.nc']);
                day_torino(:,:,:) = ncread(open_day,'tp',[Y_ERA5_TO X_ERA5_TO 1],[6 6 366])*24*1000;
        kk = kk + 1;
    all_year_mean(kk,:) = sum(day_torino,[1:2]); %Step necessary because of the CRU grid
    jj = 0;
 for ii = months
     if ii == 1
        jj = jj + 1;
     else
     end
       if days <= 366
       day_torino_divided(ii,jj,kk) = all_year_mean(kk,days); %Discard the first day because it is related to the year before
       days = days + 1;
       else
       end
 end
era5_monthly_mean = mean(day_torino_divided,1);
S = size(era5_monthly_mean);
torino_select_era5 = reshape(era5_monthly_mean,[S(1)*S(2)],[]);
era5 = reshape(torino_select_era5,1,[]);
end
```

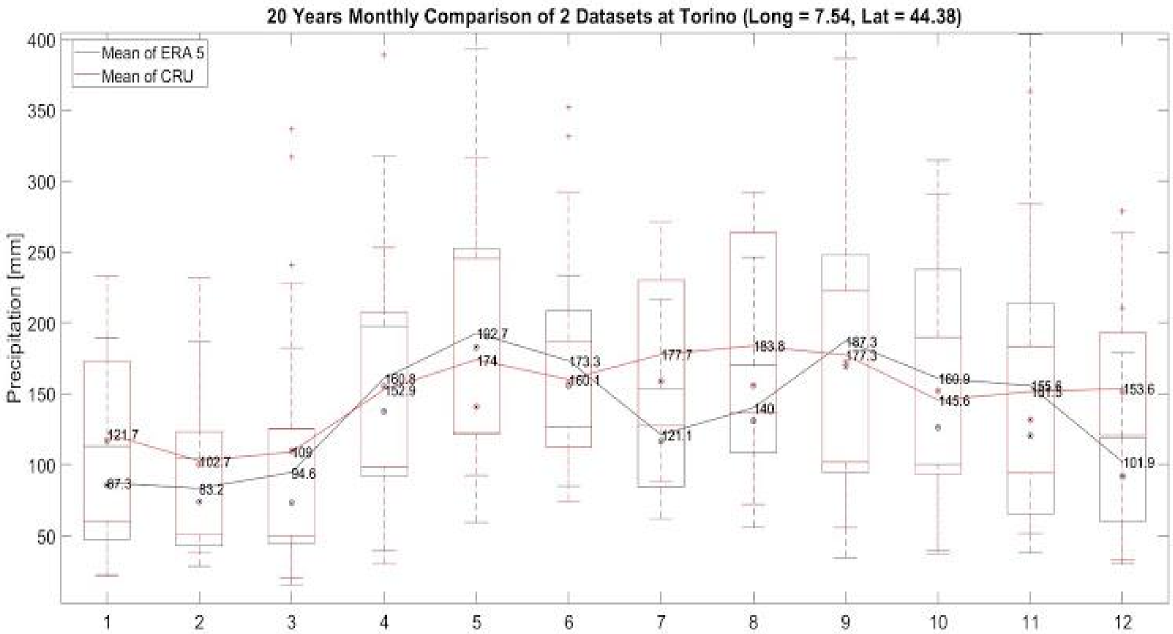
Annex 13 - 20 Years ERA5 and CRU Comparison Script



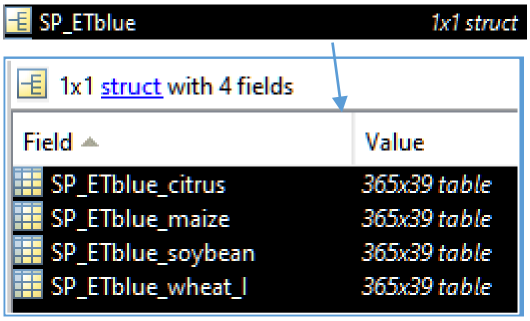
Annex 14 - Torino ERA5 vs CRU 20 Years Comparison



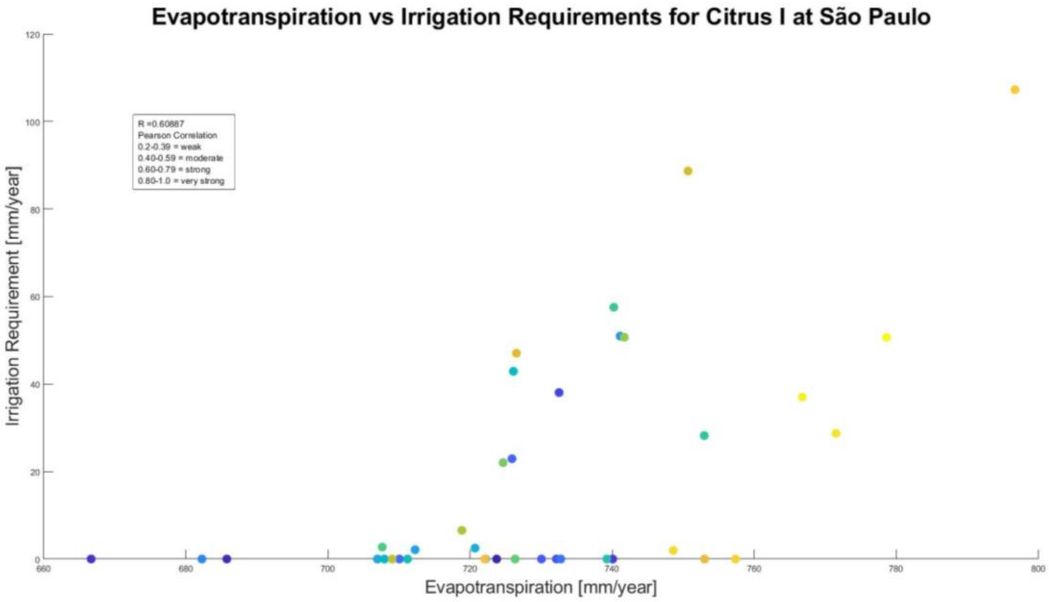
Annex 15 - Torino ERA5 vs CRU 20 Years Comparison



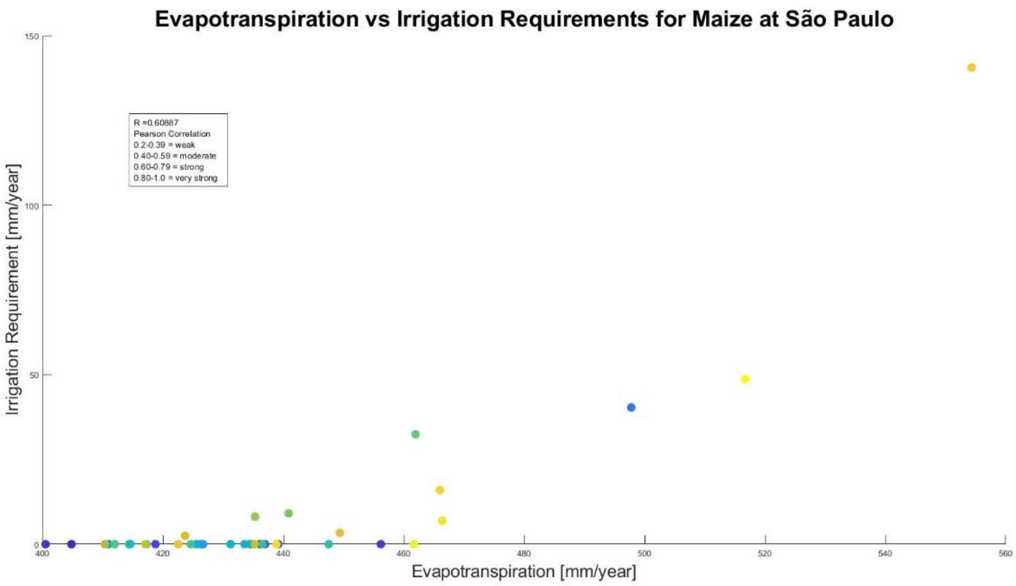
Annex 16 - Twenty Years Comparison Boxplot CRU and ERA 5 for São Paulo



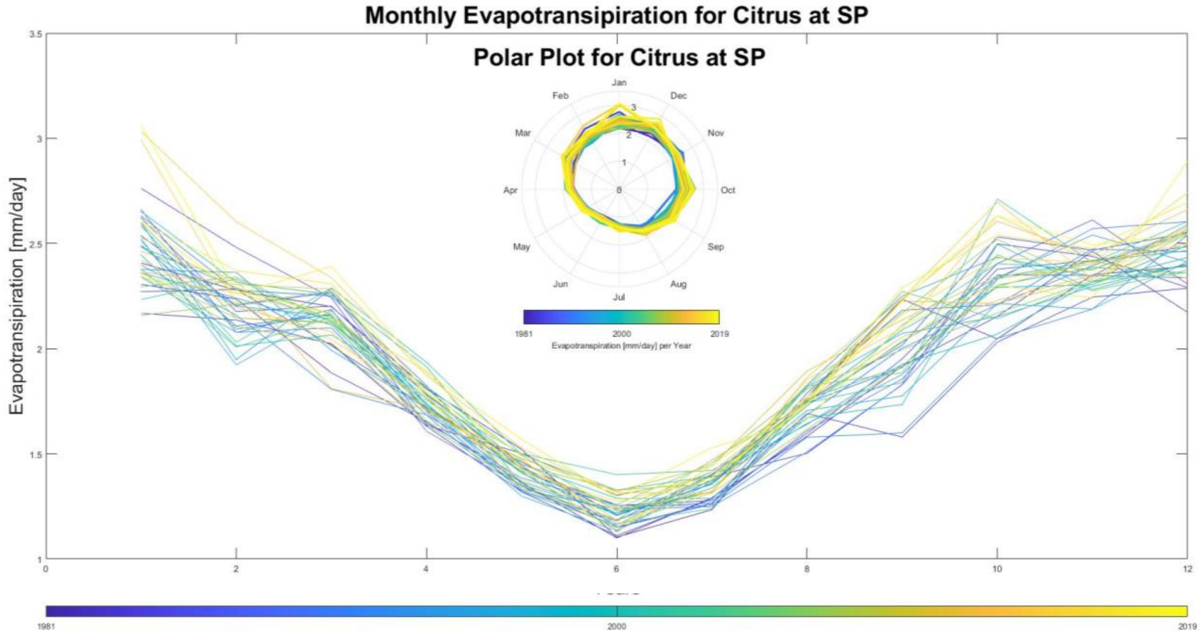
Annex 17 - Model Output Example



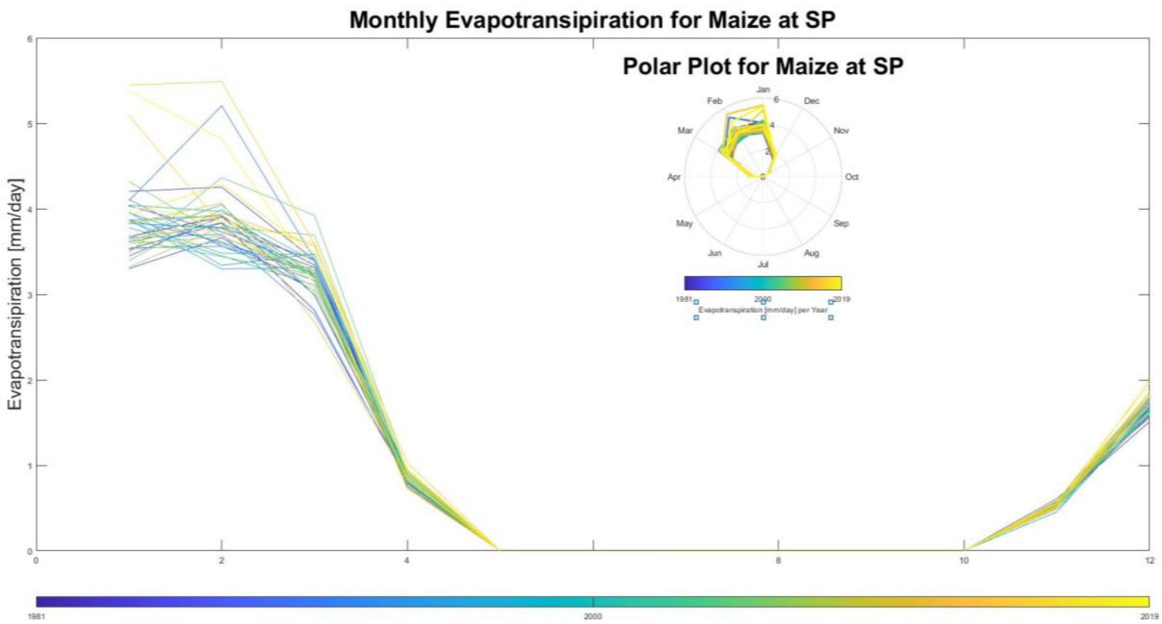
Annex 18 - Correlation between Evapotranspiration and Irrigation Requirements for Citrus at São Paulo



Annex 19 - Correlation between Evapotranspiration and Irrigation Requirements for Maize at São Paulo



Annex 20 - Monthly Evapotranspiration for Citrus at São Paulo



Annex 21 - Monthly Evapotranspiration for Maize at São Paulo

	1981-1990	1991-2000	2001-2010	2011-2019
January	2.486	2.459	2.406	2.654
February	2.220	2.130	2.192	2.314
March	2.112	2.099	2.167	2.173
April	1.712	1.722	1.762	1.793

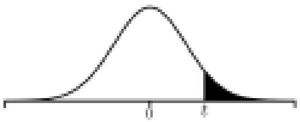
May	1.405	1.383	1.436	1.450
June	1.185	1.204	1.259	1.235
July	1.301	1.333	1.383	1.389
August	1.654	1.704	1.764	1.763
September	1.930	1.936	2.027	2.182
October	2.293	2.259	2.367	2.476
November	2.431	2.334	2.359	2.387
December	2.398	2.461	2.463	2.632

Annex 22 – Citrus Monthly Mean Evapotranspiration at São Paulo

	1981-1990	**1991-2000**	**2001-2010**	**2011-2019**
January	3.810	3.774	3.702	4.287
February	3.923	3.652	3.792	4.215
March	3.180	3.170	3.364	3.287
April	0.846	0.839	0.836	0.890
May	0.000	0.000	0.000	0.000
June	0.000	0.000	0.000	0.000
July	0.000	0.000	0.000	0.000
August	0.000	0.000	0.000	0.000
September	0.000	0.000	0.000	0.000
October	0.000	0.000	0.000	0.000
November	0.557	0.520	0.537	0.536
December	1.656	1.706	1.708	1.837

Annex 23 - Maize Monthly Mean Evapotranspiration at São Paulo

Critical Values for Student's t-Distribution.

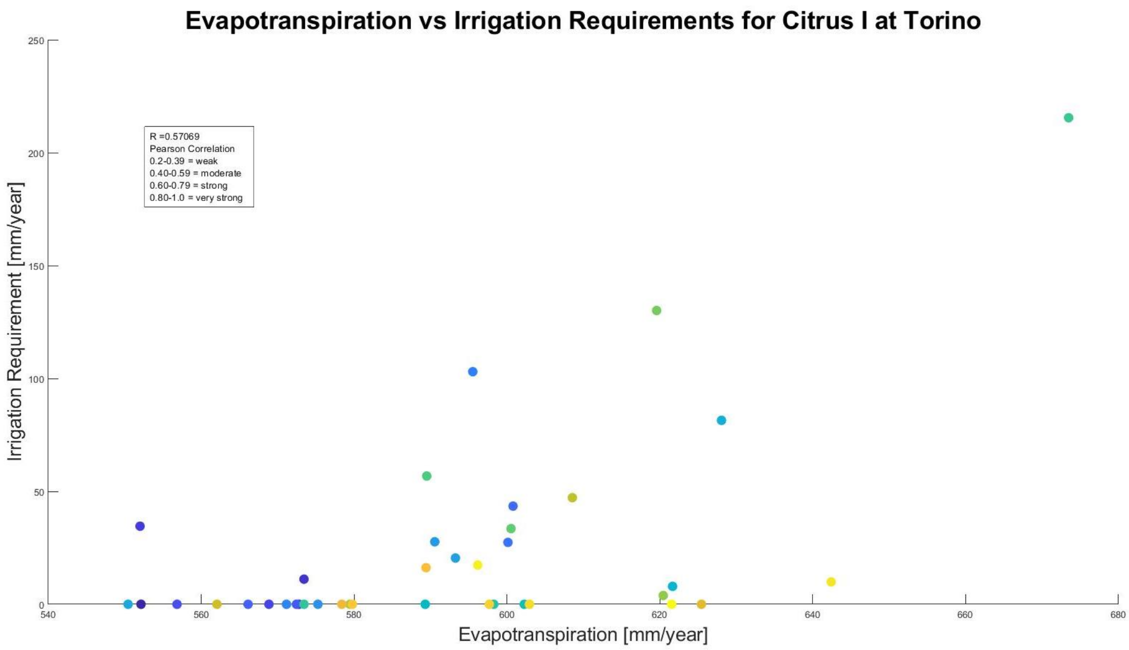


df	0.2	0.1	0.05	0.04	0.03	0.025	0.02	0.01	0.005	0.0005
	Upper Tail Probability: $\Pr(T > t)$									
1	1.376	3.078	6.314	7.916	10.579	12.706	15.895	31.821	63.657	636.619
2	1.061	1.886	2.920	3.320	3.896	4.303	4.849	6.965	9.925	31.599
3	0.978	1.638	2.353	2.605	2.951	3.182	3.482	4.541	5.841	12.924
4	0.941	1.533	2.132	2.333	2.601	2.776	2.999	3.747	4.604	8.610
5	0.920	1.476	2.015	2.191	2.422	2.571	2.757	3.365	4.032	6.869
6	0.906	1.440	1.943	2.104	2.313	2.447	2.612	3.143	3.707	5.959
7	0.896	1.415	1.895	2.046	2.241	2.365	2.517	2.998	3.499	5.408
8	0.889	1.397	1.860	2.004	2.189	2.306	2.449	2.896	3.355	5.041
9	0.883	1.383	1.833	1.973	2.150	2.262	2.398	2.821	3.250	4.781
10	0.879	1.372	1.812	1.948	2.120	2.228	2.359	2.764	3.169	4.587
11	0.876	1.363	1.796	1.928	2.096	2.201	2.328	2.718	3.106	4.437
12	0.873	1.356	1.782	1.912	2.076	2.179	2.303	2.681	3.055	4.318
13	0.870	1.350	1.771	1.899	2.060	2.160	2.282	2.650	3.012	4.221
14	0.868	1.345	1.761	1.887	2.046	2.145	2.264	2.624	2.977	4.140
15	0.866	1.341	1.753	1.878	2.034	2.131	2.249	2.602	2.947	4.073
16	0.865	1.337	1.746	1.869	2.024	2.120	2.235	2.583	2.921	4.015
17	0.863	1.333	1.740	1.862	2.015	2.110	2.224	2.567	2.898	3.965
18	0.862	1.330	1.734	1.855	2.007	2.101	2.214	2.552	2.878	3.922
19	0.861	1.328	1.729	1.850	2.000	2.093	2.205	2.539	2.861	3.883
20	0.860	1.325	1.725	1.844	1.994	2.086	2.197	2.528	2.845	3.850
21	0.859	1.323	1.721	1.840	1.988	2.080	2.189	2.518	2.831	3.819
22	0.858	1.321	1.717	1.835	1.983	2.074	2.183	2.508	2.819	3.792
23	0.858	1.319	1.714	1.832	1.978	2.069	2.177	2.500	2.807	3.768
24	0.857	1.318	1.711	1.828	1.974	2.064	2.172	2.492	2.797	3.745
25	0.856	1.316	1.708	1.825	1.970	2.060	2.167	2.485	2.787	3.725
26	0.856	1.315	1.706	1.822	1.967	2.056	2.162	2.479	2.779	3.707
27	0.855	1.314	1.703	1.819	1.963	2.052	2.158	2.473	2.771	3.690
28	0.855	1.313	1.701	1.817	1.960	2.048	2.154	2.467	2.763	3.674
29	0.854	1.311	1.699	1.814	1.957	2.045	2.150	2.462	2.756	3.659
30	0.854	1.310	1.697	1.812	1.955	2.042	2.147	2.457	2.750	3.646
31	0.853	1.309	1.696	1.810	1.952	2.040	2.144	2.453	2.744	3.633
32	0.853	1.309	1.694	1.808	1.950	2.037	2.141	2.449	2.738	3.622
33	0.853	1.308	1.692	1.806	1.948	2.035	2.138	2.445	2.733	3.611
34	0.852	1.307	1.691	1.805	1.946	2.032	2.136	2.441	2.728	3.601
35	0.852	1.306	1.690	1.803	1.944	2.030	2.133	2.438	2.724	3.591
36	0.852	1.306	1.688	1.802	1.942	2.028	2.131	2.434	2.719	3.582
37	0.851	1.305	1.687	1.800	1.940	2.026	2.129	2.431	2.715	3.574
38	0.851	1.304	1.686	1.799	1.939	2.024	2.127	2.429	2.712	3.566
39	0.851	1.304	1.685	1.798	1.937	2.023	2.125	2.426	2.708	3.558
40	0.851	1.303	1.684	1.796	1.936	2.021	2.123	2.423	2.704	3.551

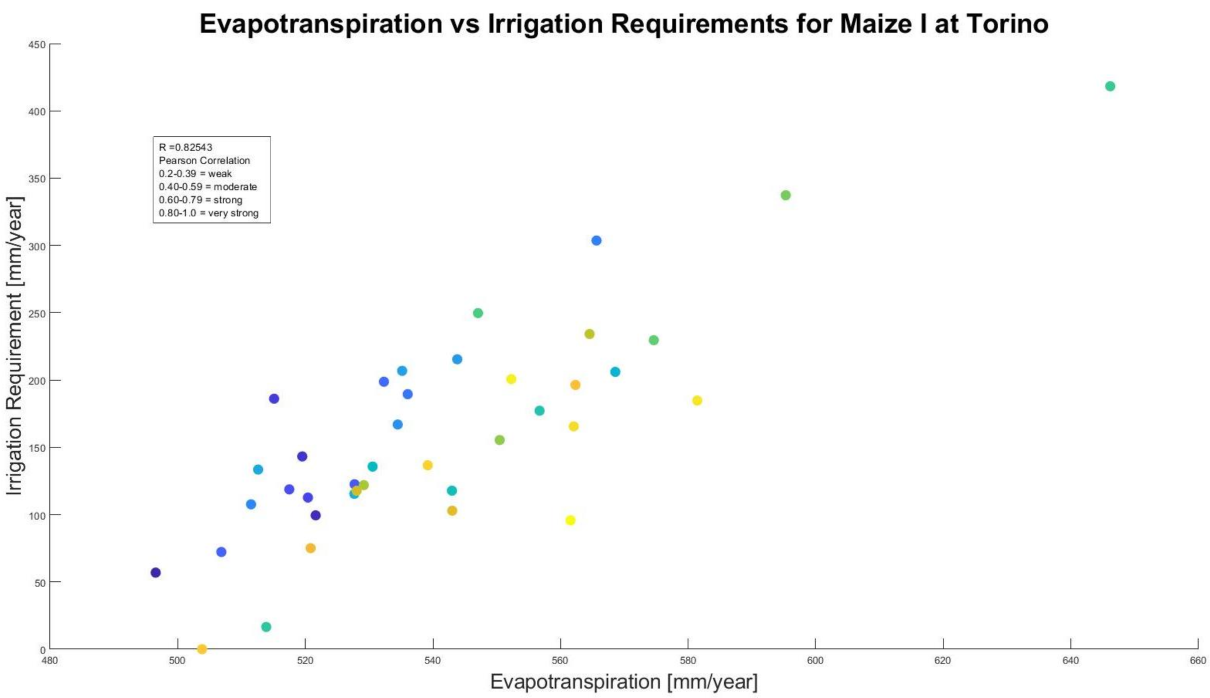
Annex 24 - T-Student Table Values (18)

Degree of Freedom	Probability of Exceeding the Critical Value								
	0.99	0.95	0.90	0.75	0.50	0.25	0.10	0.05	0.01
1	0.000	0.004	0.016	0.102	0.455	1.32	2.71	3.84	6.63
2	0.020	0.103	0.211	0.575	1.386	2.77	4.61	5.99	9.21
3	0.115	0.352	0.584	1.212	2.366	4.11	6.25	7.81	11.34
4	0.297	0.711	1.064	1.923	3.357	5.39	7.78	9.49	13.28
5	0.554	1.145	1.610	2.675	4.351	6.63	9.24	11.07	15.09
6	0.872	1.635	2.204	3.455	5.348	7.84	10.64	12.59	16.81
7	1.239	2.167	2.833	4.255	6.346	9.04	12.02	14.07	18.48
8	1.647	2.733	3.490	5.071	7.344	10.22	13.36	15.51	20.09
9	2.088	3.325	4.168	5.899	8.343	11.39	14.68	16.92	21.67
10	2.558	3.940	4.865	6.737	9.342	12.55	15.99	18.31	23.21
11	3.053	4.575	5.578	7.584	10.341	13.70	17.28	19.68	24.72
12	3.571	5.226	6.304	8.438	11.340	14.85	18.55	21.03	26.22
13	4.107	5.892	7.042	9.299	12.340	15.98	19.81	22.36	27.69
14	4.660	6.571	7.790	10.165	13.339	17.12	21.06	23.68	29.14
15	5.229	7.261	8.547	11.037	14.339	18.25	22.31	25.00	30.58
16	5.812	7.962	9.312	11.912	15.338	19.37	23.54	26.30	32.00
17	6.408	8.672	10.085	12.792	16.338	20.49	24.77	27.59	33.41
18	7.015	9.390	10.865	13.675	17.338	21.60	25.99	28.87	34.80
19	7.633	10.117	11.651	14.562	18.338	22.72	27.20	30.14	36.19
20	8.260	10.851	12.443	15.452	19.337	23.83	28.41	31.41	37.57
22	9.542	12.338	14.041	17.240	21.337	26.04	30.81	33.92	40.29
24	10.856	13.848	15.659	19.037	23.337	28.24	33.20	36.42	42.98
26	12.198	15.379	17.292	20.843	25.336	30.43	35.56	38.89	45.64
28	13.565	16.928	18.939	22.657	27.336	32.62	37.92	41.34	48.28
30	14.953	18.493	20.599	24.478	29.336	34.80	40.26	43.77	50.89
40	22.164	26.509	29.051	33.660	39.335	45.62	51.80	55.76	63.69
50	27.707	34.764	37.689	42.942	49.335	56.33	63.17	67.50	76.15
60	37.485	43.188	46.459	52.294	59.335	66.98	74.40	79.08	88.38
	Not Significant							***Significant***	

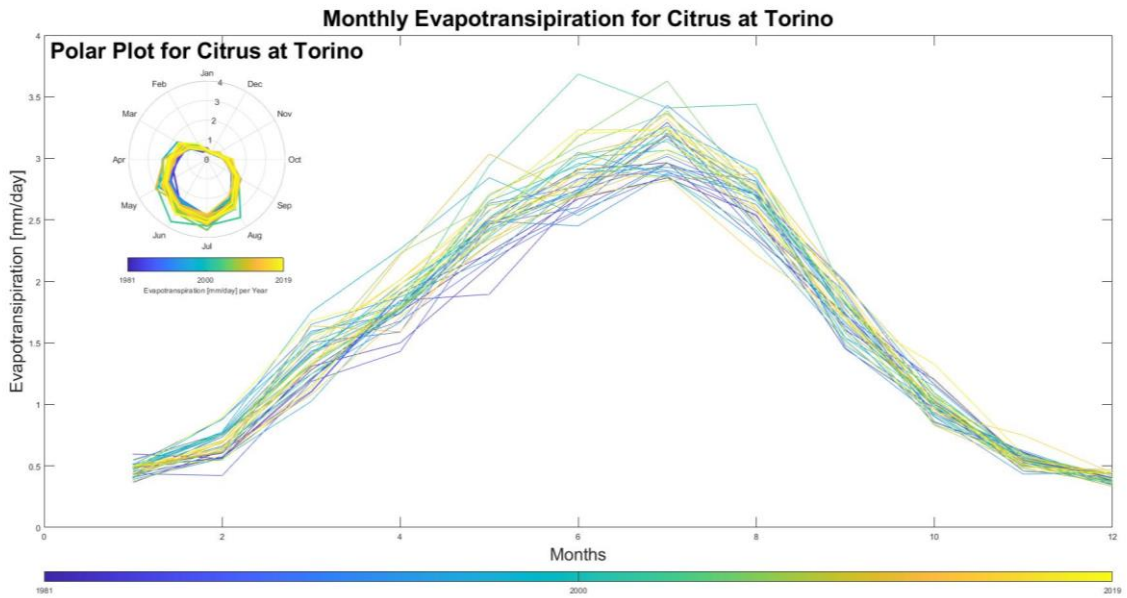
Annex 25 - Chi Squared Table (22)



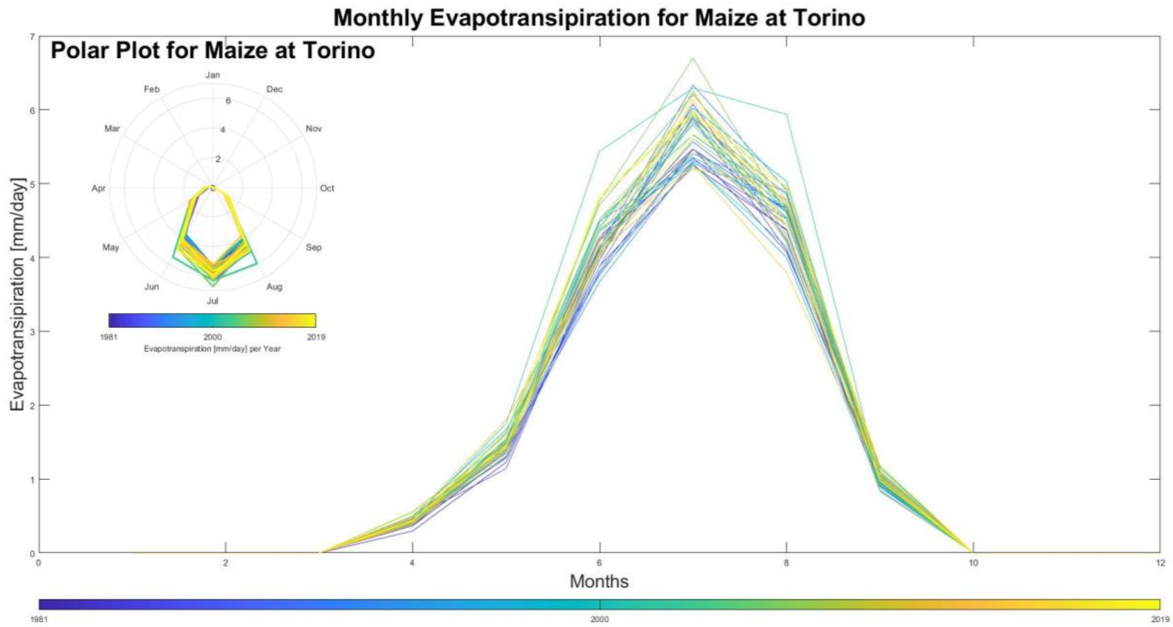
Annex 26 - Correlation between Evapotranspiration and Irrigation Requirements for Citrus at Torino



Annex 27 - Correlation between Evapotranspiration and Irrigation Requirements for Maize at Torino



Annex 28 - Monthly Evapotranspiration for Citrus at Torino



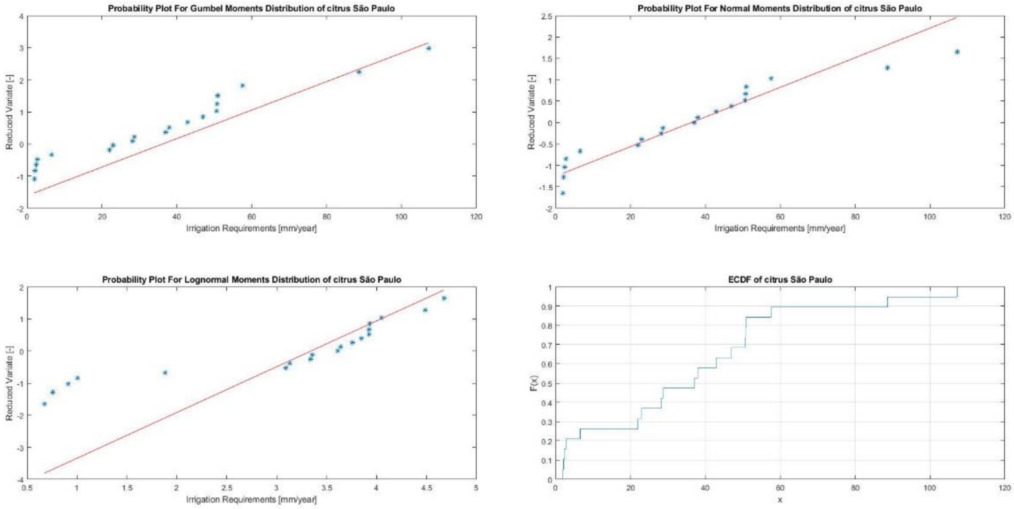
Annex 29 - Monthly Evapotranspiration for Maize at Torino

	1981-1990	1991-2000	2001-2010	2011-2019
January	0.470	0.474	0.453	0.473
February	0.632	0.723	0.674	0.671
March	1.284	1.418	1.342	1.388
April	1.718	1.869	1.850	1.907
May	2.328	2.499	2.533	2.535
June	2.712	2.782	3.007	2.897
July	3.018	3.080	3.228	3.120
August	2.559	2.691	2.709	2.684
September	1.747	1.653	1.736	1.782
October	1.034	0.973	0.978	1.040
November	0.546	0.538	0.545	0.553
December	0.405	0.388	0.382	0.425

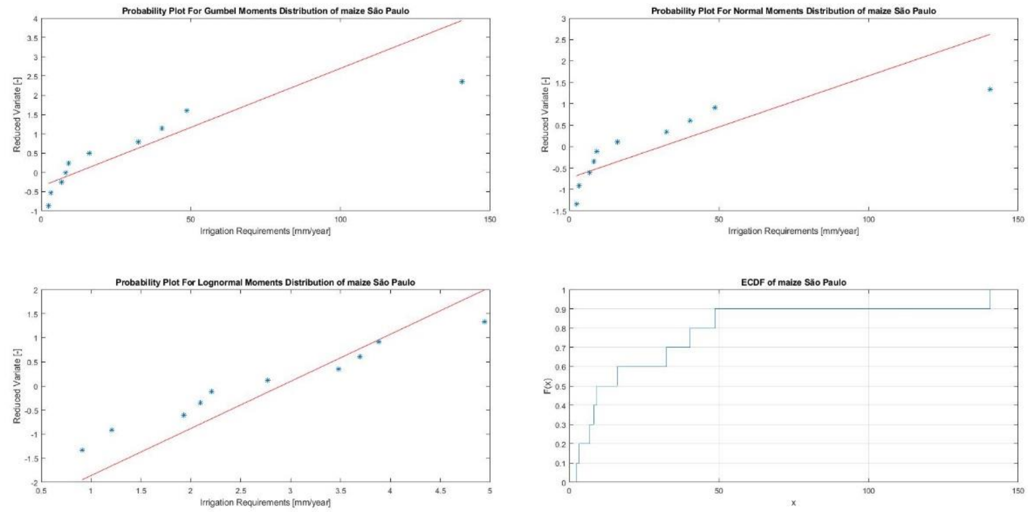
Annex 30 - Citrus Monthly Mean Evapotranspiration at Torino

	1981-1990	1991-2000	2001-2010	2011-2019
January	0.000	0.000	0.000	0.000
February	0.000	0.000	0.000	0.000
March	0.000	0.000	0.000	0.000
April	0.409	0.430	0.448	0.447
May	1.373	1.469	1.508	1.497
June	4.033	4.142	4.493	4.314
July	5.572	5.685	5.959	5.760
August	4.405	4.639	4.655	4.607
September	0.979	0.968	1.020	1.033
October	0.000	0.000	0.000	0.000
November	0.000	0.000	0.000	0.000
December	0.000	0.000	0.000	0.000

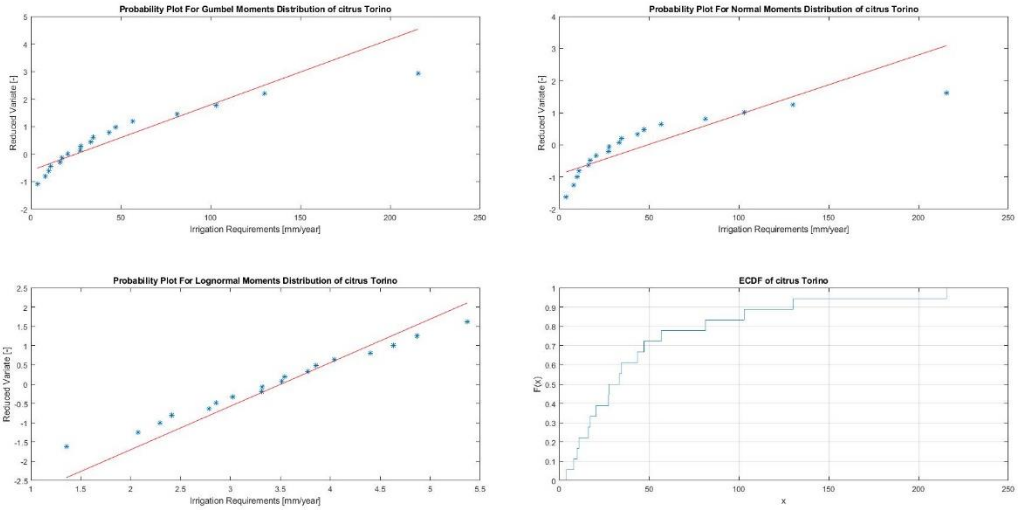
Annex 31 - Maize Monthly Mean Evapotranspiration at Torino



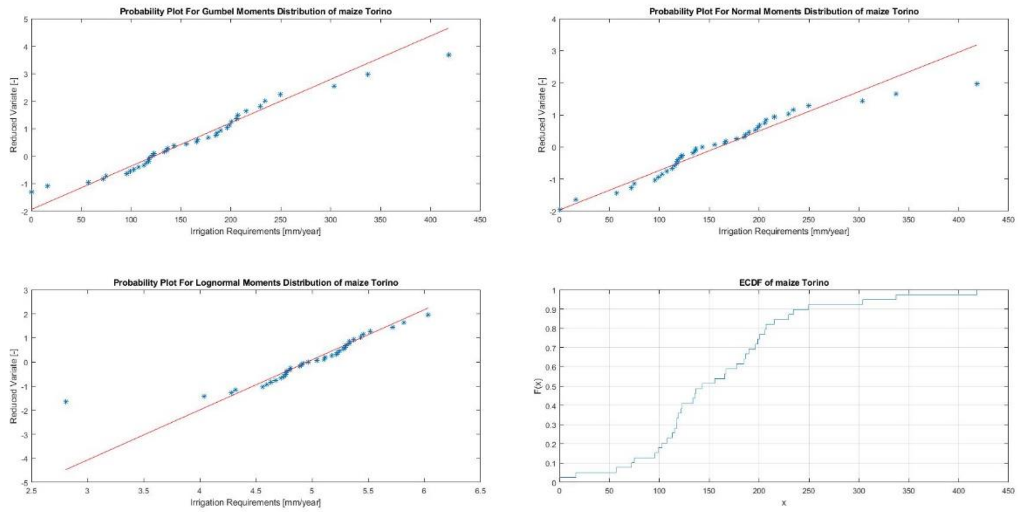
Annex 32 -Probability Plots and ECDF for Citrus at São Paulo



Annex 33 - Probability Plots and ECDF for Maize at São Paulo



Annex 34 - Probability Plots and ECDF for Citrus at Torino



Annex 35 - Probability Plots and ECDF for Maize at Torino