Changing water availability in the

São Francisco River Basin, Brazil: exploring the role of

expanding agriculture and climate change

Maria Pereira Fleury

April 4, 2022

Submitted in partial fulfillment

of the requirements for the degree of

Bachelor of Science in Engineering

Department of Civil and Environmental Engineering

Princeton University

I hereby declare that I am the sole author of this thesis.

I authorize Princeton University to lend this thesis to other institutions or individuals for the purpose of scholarly research.
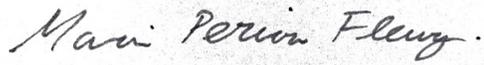
Maria Pereira Fleury

Maria Pereira Fleury

I further authorize Princeton University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.
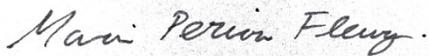
Maria Pereira Fleury

Maria Pereira Fleury

Acknowledgements

I would like to extend my gratitude to my thesis advisor Prof. Reed Maxwell for all his support, guidance and positive encouragement in the writing of this thesis. I appreciate how you pushed me to find a research topic that truly interested me and how you reminded me weekly that my slow progress was something to be proud of. I wanted to thank all of the Maxwell Research group for welcoming me into the world of research, and especially Jackson Swilley and Nick Jadallah for your consistent willingness to listen and help in whatever way possible. A special thanks also to Mary Michael O'Neill at NASA, for sharing her time, patience and invaluable expertise on the GRACE dataset.

In addition, I would like to thank Miqueias Mugge at Princeton's Brazil LAB for connecting me with the research team at MapBiomas, without whom this research would not have been possible. Carlos Souza and Juliano Schirmback at MapBiomas Água were essential in clarifying all my doubts and supporting me with my extra data requests. Thank you for seeing me as a collaborator and valuing my insight, even when I feel like I'm still learning myself.

To the CEE department faculty and staff, thank you for providing a welcoming learning environment. I appreciate all of the guidance and support I've received over the years. To my friends in CEE, especially my best friend and roommate Maya McHugh, I am grateful to have gone through these four years with you.

Finally, I am immensely grateful to my friends and family for supporting me through these four years. To all my roommates, current and past, and to all my friends at RFC, you make me feel loved and at home. Dad, you inspire me to learn and explore the world of intellectual ideas – I hope to be as accomplished as you some day. Mum, thank you for reminding me to breathe and to always 'take it easy', you make my world brighter and lighter.

Table of Contents

Abstract

Human activities not only depend on but can also substantially affect the availability of water. Therefore, it is a concern that water resources are quickly being depleted in regions which already experience drought, like the São Francisco River Basin in Brazil. This thesis explores how climate change and the growth of large-scale agriculture affects water availability at the São Francisco River Basin and associated sub-basins (Baixo, Submédio, Médio and Alto São Francisco). Remote-sensing observations and climate re-analysis data is used to quantify trends in agricultural land-use, climate parameters and water resources. Additionally, climate and land-use datasets are combined to estimate water used for irrigation by center pivot systems. Datasets used include MapBiomas data on land-use change, water surface area and center pivot irrigation, GRACE observations of total water storage and GLDAS climate re-analysis. Results show substantial depletion of water resources amidst growing agriculture and rising temperatures. Water surface area decreased at a rate of $5km^2$/month between 1985 and 2020, with significant correlation to depletion of total water storage (r=0.82, r^2=0.68). In the same period, 4.6% of the basin's area saw a transition from native vegetation to agriculture, and soy plantation increased by 40x in area. Center-pivot irrigation systems, which grew with agricultural expansion, are estimated to have used the equivalent of 8.7% of groundwater available for human consumption in 2020 alone. Through the sub-basin analysis, it is found that depletion of water resources is most pronounced in the Médio and Alto São Francisco sub-basins, where the largest growth in agriculture and center-pivot irrigation is also observed. These results suggest that commercial agriculture's impact on water resources in the São Francisco River Basin are significant and must be more carefully monitored and managed.

1. Introduction

1.1. Motivation

Water is essential in the production of everything we consume – food, energy, material goods – which makes it intimately connected to activities such as pasture, agriculture and energy production. Human activities not only depend on but can also substantially affect the availability of water stored as snow, soil moisture, groundwater, and surface water. In regions with arid or semi-arid climates, where water availability is already a concern, the influence of human activities, such as agriculture, is particularly important to examine. Agriculture often causes loss of local vegetation, coupled with overexploitation of water resources. In drought-prone regions these processes could lead to lack of water supply, threatening the sustainability and resiliency of essential food and water systems.

The land use changes (LUCs) associated with the expansion of agriculture are concerning in part because they can affect local hydrological processes. Cutting down native vegetation will likely disrupt infiltration processes, surface runoff, recharge rates and evapotranspiration at the local level. Recent studies have shown that land cover change linked to agriculture in Northeast Brazil and the Amazon have the potential to affect precipitation, particularly during the dry seasons (Bagley et al., 2014; Cunha et al., 2015). These changes in hydrological processes can in turn affect local water availability, enhancing drought processes.

Agricultural expansion also leads to increased water use for irrigation. In semi-arid or arid regions, this can create a concerning feedback loop, where water is overexploited

for irrigation creating a water deficit for future agricultural production, and further threatening both water and food systems. For example, a study on the U.S. High Plains found that, when extrapolating current groundwater depletion rates, 35% of the southern High Plains would be unable to continue irrigation within 30 years (Scanlon, Faunt, et al., 2012).

In the United States, where groundwater usage has grown in the past few decades, an overall depletion in water resources is observed, particularly where groundwater recharge rates are low (Scanlon, Faunt, et al., 2012; Thomas & Famiglietti, 2019). Groundwater helps stabilize soil moisture, evapotranspiration and irrigation needs making it a critical hydrological component to quantify and preserve in the face of climate change. In the São Francisco River basin, Brazil a similar overall depletion in water resources occurs, even though surface water is still used for irrigation. A concern across Brazil is that groundwater dynamics are very sparsely studied and as surface water is depleted, groundwater usage has been increasing, especially for irrigation (Oliveira et al., 2021).

A study on the Alto Grande watershed argues that recent decrease in rainfall and increase in center-pivot irrigation is to blame for considerable groundwater level decrease observed in local well data (Marques et al., 2020). In the nearby Urucuia Aquifer System studies show, through remote sensing data, a depletion rate of total water storage of 6.5 ± 2.6 mm yr^{-1} (Gonçalves et al., 2020). Both these studies cover the region of Western Bahia in Brazil, where agriculture and center pivot irrigation has grown consistently since 1985.

On the ground, depletion of water resources often translates to inequalities in water access. A recent news article on water disputes in Western Bahia reported that *geraizeiros*, local subsistence farmers recognized as traditional communities by Brazilian law, have found that nearby streams and springs are becoming drier. They blame large-scale commercial agriculture, which has expanded the production of soy in the region, for pumping large amounts of water through center-pivot irrigation systems (Daniel Grossman, 2021). Figure 1 illustrates the stark contrast between the local producer and large-scale agriculture. Considering the human impacts of water depletion, it is critical that we work towards drawing conclusions on how agricultural expansion may affect water availability so that water and land use planning can change accordingly.

Figure 1: (a) "Local farmer Liobino dos Santos on a dry canal near the Bota Bunda spring in Brejo Verde, a rural village in western Bahia." (b) "Pivot irrigation greens a field at Esperanto farm in western Bahia." Retrieved from: Daniel Grossman, 2021

The biggest challenge with correlating water depletion and agricultural expansion is that water used for irrigation, as well as the overall availability of water, are sparsely monitored and difficult to model. The lack of information on the usage of water resources creates barriers for more efficient water management – an essential aspect of resiliency building in the face of climate change. In most regions of Brazil, there is little publicly available data on streamflow and well water level, with only 409 wells currently registered in the national monitoring system (RIMAS, 2021). In comparison, the USGS National Water Information System in the United States has over 850,000 registered wells/springs/test-holes/etc. along with a wide network of streamflow measurements (USGS, 2021). However, even with a wide network of data, the spatial distribution of data points might lead to bias such that results are not representative of water availability in a region. This is particularly true in the case of groundwater, which is also very challenging to accurately model (Rateb et al., 2020). Connecting water availability changes with agricultural expansion is also not easy without current and updated data on land-use expansion.

Recent developments in remote sensing techniques allow us to address the described problems, providing current data for both land-use change and water availability for large areas. MapBiomas – a project that reconstructs annual land use in Brazil based on machine learning techniques applied to Landsat archive – has made a significant contribution to tracking LUCs in Brazil. Their data shows that agricultural coverage has increased by 172% since 1985 (Souza et al., 2020) and water surface area decreased by 15% since the early 90s (MapBiomas, 2021). They also recently developed

a deep learning technique that identifies center-pivot irrigation systems, made possible by the distinct circular shape of the agricultural field, as can be seen in Figure 1 (b) (Saraiva et al., 2020). This method has been validated and replicated in the Ogallala Aquifer, United States, where center-pivots have an established presence, and has proven valuable in more accurately estimating groundwater usage for irrigation (Daniel Cooley et al., 2021). Another dataset of particular interest is that collected by Gravity Recovery and Climate Experiment (GRACE), which allows us to quantify changes in total water storage. The dataset is particularly important when considering water resources which are hard to monitor, such as groundwater, and largely accepted as accurate (Gonçalves et al., 2020; Rateb et al., 2020; Scanlon, Longuevergne, et al., 2012; Thomas & Famiglietti, 2019). GRACE is a useful tool that not only estimates water depletion but is used to validate models and point observations.

My research aims to use MapBiomas and GRACE's remote-sensing datasets to correlate agricultural expansion and change in water availability. While numerical experiments, land-surface models and other hydrological models have successfully shown that LUCs associated with agriculture will affect hydrological processes in different biomes in Brazil (Andrade & Ribeiro, 2020; Bagley et al., 2014; Cunha et al., 2015), few direct comparisons seem to be available between current LUC and GRACE data on water availability. Even fewer bring together datasets that consider LUC, climate change and water availability. I hope to build on the work of scholars who have been studying the decrease in water availability in Western Bahia in Brazil, and its tie to growing agribusiness

in the region, (Gonçalves et al., 2020; Marques et al., 2020; Santos et al., 2020) with a more large scale analysis of the São Francisco River Basin and its sub-basins.

1.2. Questions and Goals

Considering that the interdependence of water and agriculture is critical for water planning in areas facing increased droughts, in this thesis I hope to draw a correlation between agricultural expansion and decrease in water availability, while also considering the impact of a changing climate. The relationship between these three key components of the research is summarized in Figure 2. With agricultural expansion often comes the loss of native vegetation, which disrupts natural mechanisms of water retention and increases water surface runoff. Disruptions to the hydrological cycle decrease water available for plants, which consequently decrease evapotranspiration and hence precipitation. In summary, increased deforestation for agriculture both decreases the water available and increases the need of water for irrigation. Climate change augments the effects of agricultural expansion on water availability by intensifying droughts and hence further increasing water demand for irrigation.
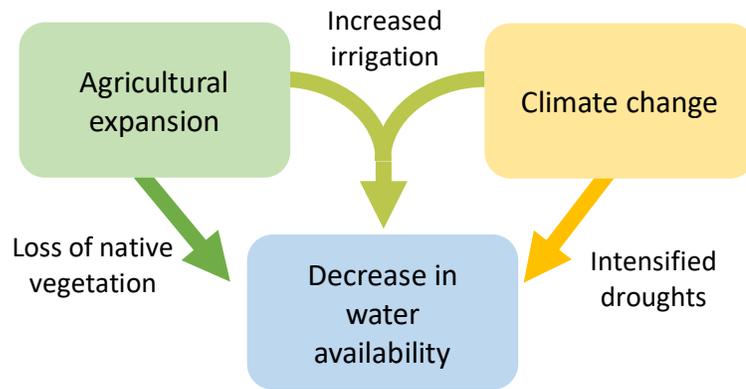
Figure 2: Relationship between the three key components of the research

The relationship between these three components – agricultural expansion, climate change and decrease in water availability – makes it essential that we analyze them in conjunction. This leads to the research question: **How does expansion in agriculture and climate variability affect the net availability of water in the São Francisco River Basin?** To answer this question the goals of this thesis include:

(1) Quantify expansion of agriculture and center-pivot irrigation

(2) Characterize natural climate variability and identify any changes in climate parameters

(3) Combine different datasets that quantify changes in water availability over time and robustly characterize these changes

(4) Estimate increased center-pivot irrigation water use

(5) Identify areas of the basin that are most affected by agricultural growth and changes in water availability

Given the limited scope of this thesis, I will be focusing my research on the São Francisco River basin and its sub-basins. The São Francisco River Basin is an important hydrological region in Brazil where large-scale, commercial agriculture, and hence, center pivot-irrigation has grown significantly. Its sub basins – Baixo São Francisco, Submédio São Francisco, Médio São Francisco and Alto São Francisco – cover different biomes, with climates ranging from arid to semi-arid to tropical, which lead to varying agricultural practices. This thesis will first provide a big-picture analysis of the changing agricultural, climate and water scenarios across the São Francisco River basin, and then turn to analyze the same parameters at each sub-basin. A deeper dive into the sub-basins will allow us to better understand the localized changes and its contribution to observations at the basin level (goal 5). Other questions that motivate this thesis include:

- How can we accurately and effectively quantify agricultural usage of surface water and groundwater?

- How well do visible changes in surface water availability translate to changes in total water storage?

- To what extent do climate phenomena alone explain the observed changes in water availability?

While the above questions will not be directly answered in this thesis, the analysis presented will contribute to their discussion.

2. Methods

2.1. Data Requirements

To answer the research question, I analyze land-use change, water availability as well as climate parameters over the São Francisco River basin in Brazil. This includes the integration of remote-sensing data and climate re-analysis data. Key information on the datasets used is presented in Table 1, while more background information on each dataset, how the data is processed and potential limitations are discussed in the sections that follow.

As an overview, I use MapBiomas land use data to show the transitions in land cover, characterize the expansion in agriculture and growth in center-pivot irrigation. This includes identifying what kinds of crops have been growing and studying the drivers in LUC. I also use MapBiomas surface water area data and compare it to GRACE's changes in Total Water Storage (TWS). Through statistical analysis, GRACE is used to confirm the relationship between changes observed in surface water area and total water storage change in the region. The relationship helps validate the findings in the MapBiomas data, a dataset going back to 1985, compared to GRACE's 2002 start date. When analyzing changes in water availability long-term trends and cycles are identified. NASA's Global Land Data Assimilation System (GLDAS) climate re-analysis data is used to characterize local climate trends. Comparing these three classes of data – LUC, changes in water storage and climate data – I determine patterns in water storage, LUC and climate and draw potential correlations between them.

Table 1: Summary of datasets used for analysis at the São Francisco River Basin, Brazil

	Data type	*Spatial-Temporal Availability*	*Resolution*
MapBiomas v. 6.0	Land cover and LUC of 5 major classes: forest, non-forest natural formation, farming and non-vegetated areas; 20+ sub-classes available (including major crop-types)	Brazil 1985-2021, yearly	30m pixel
MapBiomas Água	Area of surface water	Brazil 1985-2020, monthly	30m pixel
MapBiomas	Area of center-pivot irrigation	Brazil 1985-2020, yearly	30m pixel
GRACE Global mascons (JPL RL06_v02)	Vertically integrated estimates of TWS anomalies	Global 2002-2017 2018-current, monthly	3-degree lat/lon measurements with 0.5-degree grid sampling
GLDAS	Climate re-analysis (Noah Land Surface Model); Includes modeled precipitation, evapotranspiration and temperature	Global 2000-2021 monthly	0.25-degree lat/lon grid

2.1.1. MapBiomas

The MapBiomas datasets come from processing of Landsat satellite images. Using machine learning techniques, the MapBiomas team classifies each 30m by 30m pixel into the appropriate category of land cover (MapBiomas, 2019). The MapBiomas v. 6.0 (collection 6) is a robust algorithm that has been improved over time to track LUC. The MapBiomas Água dataset of surface water as well as the center pivot irrigation tracking are in their first iteration.

2.1.2. GRACE

The Gravity Recovery and Climate Experiment (GRACE) twin satellites were launched in March 2002, collecting monthly measurements of anomaly in Earth's gravitational field up until June 2017. Their successors, part of the GRACE-FO mission, were launched in May 2018, resulting in a year-long gap in the dataset. The changes in gravitational pull can be processed to represent water mass anomalies by removing tidal, atmospheric and oceanic contributions, hence serving as a measure of the changing water storage (Scanlon, Longuevergne, et al., 2012). In this thesis we use the Jet Propulsion Lab (JPL) mass concentration blocks, or mascons, a Level 3 product that processes GRACE and GRACE-FO observations over areas of 3-degree by 3-degree, in their 6th release titled RL06_v2 (NASA Jet Propulsion Laboratory, n.d.). This corresponds to data for pixels of around 300km by 300km. The dataset is hence limited by its resolution, such that observations best and uniquely describe basins of around 100,000 km^2 or larger.

Measurements of total water storage can be broken down as (Scanlon, Longuevergne, et al., 2012):

Total Water Storage	=	Groundwater storage	+	Snow equivalent storage	+	Water reservoir storage	+	Soil moisture storage	(1)

At the São Francisco River Basin there is no snow equivalent storage so the GRACE total water storage data will be representative of a combination of groundwater, surface water reservoirs and soil moisture.

2.1.3. GLDAS

The Global Land Data Assimilation System (GLDAS) is a NASA dataset that takes in satellite and ground-based observations, and, using advanced land surface modeling and data assimilation techniques, generates land surface states and fluxes (NASA Land Data Assimilation System, 2022). For this thesis climate re-analysis data that was generated using a Noah Land Surface Model at a 0.25-degree x 0.25-degree grid.

2.2. Data Processing

Most of the data processing and analysis for this thesis was completed in Python's jupyter notebook. The following packages were used: *numpy, pandas, matplotlib, seaborn, geopandas, datetime, xarray* and *regionmask.* The complete scripts of code were uploaded to github and are accessible through the link included in Appendix 2. These include three scripts, one that analyzes agricultural land-use data, a second that processes

geospatial data and analyzes water availability datasets (GRACE and MapBiomas), and a third one that replicates some of the geospatial analysis for the climate re-analysis data (GLDAS) and eventually uses that data to estimate irrigation water use.

An important part of this analysis included processing geospatial data to get from pixel-level observation to relevant time-series that averaged the geospatially distributed information for a desired region of the São Francisco River Basin and its sub-basins. This processing was not necessary for the MapBiomas data, which was already clipped to the appropriate basin sizes, and for which relevant time-series were readily available. The technique was used on the GRACE and GLDAS datasets. First, we identify a rectangular frame in which the basin falls into. Then cut the dataset to the appropriate size. Using the *Regionmask* package in Python, and appropriate shapefiles, a mask is created to subset the data. The mask assumes a pixel is in or out based off the coordinate of its center. With a set of pixels assigned to the watershed an average can be computed for the whole region, at each time step. A limitation to this technique is that the pixels do not perfectly cover the entire watershed.

In addition, the GLDAS dataset is available to download as *netcdf* files which are geospatial datasets that stack different parameters over a single pixel location. Therefore, in order to perform a temporal analysis of the desired parameter, files for each time stamp were downloaded and re-stacked over time instead of by parameter.

2.3. Estimating center-pivot irrigation water use

Given the availability of data on center-pivot irrigation location, as part of determining the connection between climate, agriculture and water we will use components of the three to estimate center-pivot irrigation water use in the São Francisco River Basin. This estimate will be directly compared to the average river flow to provide perspective on how much more water is being used by large-scale agriculture today. A simple water budget estimate is used, following the example of Santos et al., 2020, which did a similar estimate in their case study of Western Bahia in Brazil. This calculation, described by Equation (2), assumes that if evapotranspiration is greater than the precipitation than the net loss of water must be compensated by irrigating the agricultural field. It scales this value by the efficiency of a center pivot system, which is approximately 0.85 (Santos et al., 2020). To compute this estimate evapotranspiration and precipitation GLDAS data will be combined with the remote-sensing observations of MapBiomas quantifying area of center-pivot irrigation.

$$D = \max\left(\frac{ET_a - P_a}{\varepsilon}, 0\right) \qquad (2)$$

Where

D = irrigation depth (mm/day)

ET_a = average evapotranspiration in region (mm/day)

P_a = average precipitation in region (mm/day)

ε = efficiency of center pivot (= 0.85)

From the estimate of irrigation depth, we can then calculate the water uptake for irrigation for the region, Q (m³/day) which is defined as:

$$Q = D \times A \hspace{6cm} (3)$$

Where

A = area covered by center pivot irrigation in the region (m³)

D = irrigation depth (m/day)

2.4. Statistical analysis

To better characterize the long-term trends in climate, changes in water availability, and changes in water used for irrigation, a simple linear regression y = mx + b was computed using *numpy.polyfit*. All of the calculated linear fits are included in Appendix 1, where, unless otherwise specified, y is the variable that titles each table and x is time, typically a monthly time-step. In most cases m, the slope, directly presents the long-term trend of interest. Measurements of rainfall and evapotranspiration are the exception, with units of mm/day but data points for each month. In these cases, the slope is multiplied by an average of 30.5 days so that the trendline has units of mm/month or mm/year. The intercept b of the trendlines is not of interest.

For the correlations computed between GRACE's total water storage and MapBiomas' surface water area, along with the linear regression, a Pearson correlation coefficient r was computed using *numpy.corrcoef*. The Pearson correlation coefficient is a measure of linear correlation between two datasets, comprised of a ratio of the covariance between the two variables and the product of their standard deviations

(NumPy Developers, 2022). From r, the coefficient of determination r^2 was also computed. The coefficient of determination defines the proportion of the dependent variable that can be predicted by the independent variable. While none of the trends are used to model future values, it is relevant to know r^2 as a further measure of agreement between the datasets.

3. São Francisco River Basin

The São Francisco River Basin in Brazil covers an area of around 630,000 km^2, draining part of six states – Minas Gerais, Goiás, Bahia, Pernambuco, Alagoas, and Sergipe – as well as part of the Federal District, Brasilia (Magalhães & Martins, 2021). The basin is divided into four sub-basins, Baixo São Francisco, Submédio São Francisco, Médio São Francisco and Alto São Francisco basin that can be seen in Figure 3. These sub-basins cover three different biomes: the semi-arid Caatinga, Brazil's tropical savannah the Cerrado, and the tropical forest Mata Atlântica. Figure 4 shows how the biomes overlap with the basin and sub-basins. The different biomes mean that there is a lot of diversity in the native vegetation of this basin. This diversity implies that removing vegetation could have varied effects on the basin's hydrology. It also means that there are different amounts of agricultural expansion within the sub-basins, as different soil types and climate will affect suitability for crop growth. With a large portion of the river crossing the Cerrado and Caatinga, the São Francisco River crosses some of the driest lands in the country, making it an invaluable resource for millions of people.

The São Francisco River is a large river, with an average discharge of 2769 m^3/s (Comitê da Bacia Hidrográfica do São Francisco, 2016), varying between 1077 m^3/s and 5290 m^3/s depending on if it is dry or wet season (Magalhães & Martins, 2021). The river has 168 tributaries, 99 of which are perennial rivers, meaning they continuously flow throughout the year. Below the ground there are 44 aquifer systems with a total annual recharge of 1828 m^3/s, 20% or 365 m^3/s of which is considered exploitable groundwater used for irrigation and water supply (Magalhães & Martins, 2021).

In this chapter we present data for the whole São Francisco river Basin to get a big picture understanding of the trends in agricultural land-use, climate and water availability in the basin. We begin by characterizing agricultural expansion, followed by climate, then changes in water availability and finally estimating irrigation water use across the basin.
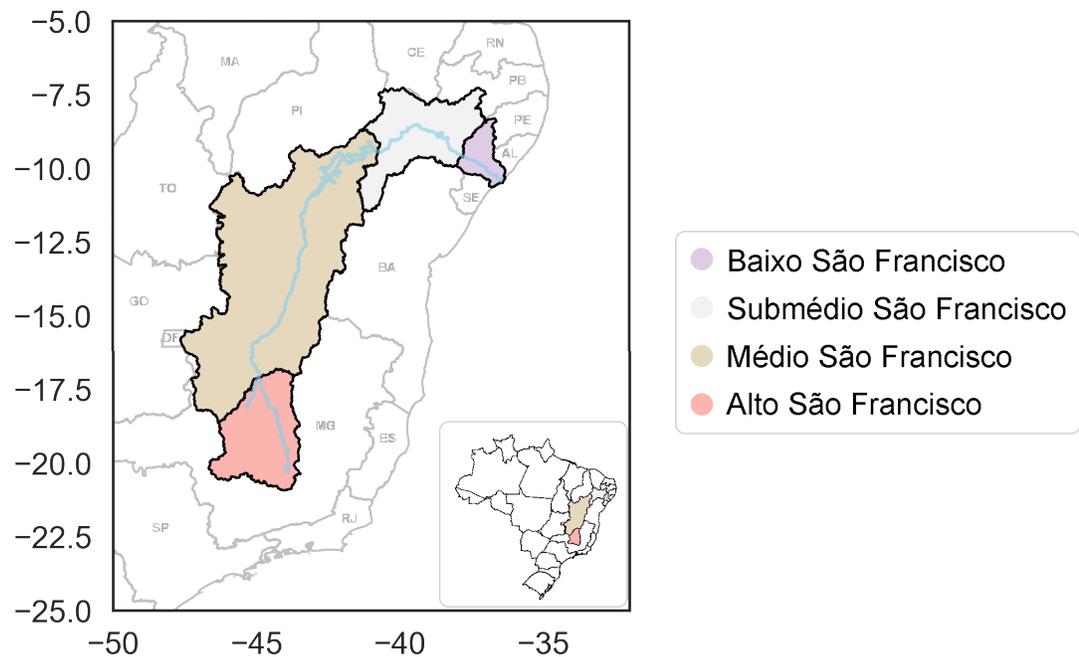
Figure 3: The São Francisco River Basin and its sub-hydrological regions in socio-political context. The main channel of the river is outlined in light blue. Right-bottom corner shows zoomed-out basin in the context of Brazil. Axis show latitude and longitude in degrees. Shapefile retrieved from (Resende, 2017) for the river, (IBGE, 2010) for the states, (ANA, 2022) for the basin.
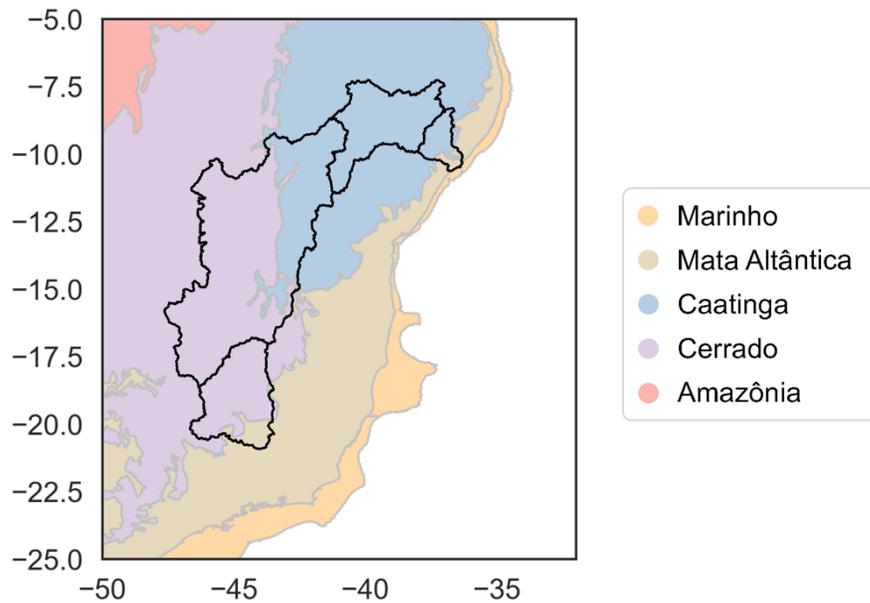
Figure 4: The São Francisco River Basin in the context of the biomes. Axis show latitude and longitude in degrees. Biome shapefile retrieved from (IBGE, 2014).

3.1 Agricultural expansion

The MapBiomas land use change map for the São Francisco River Basin, Figure 5, shows that a significant area has transitioned from forest or natural non-forest areas to agriculture or non-forest areas (red pixels). More specifically, 4.6% of the total basin area underwent net transition from native vegetation to agriculture from 1985 to 2020.

The change in agricultural land-use area and crop-type is shown in Figure 6. Total agricultural land use increased by almost 7-times, from 5000 km^2 in 1985 to over 34000 km^2 in 2020. We see a large contribution of temporary crops, crops that are sown and harvested within the agricultural year, and a much smaller contribution of perennial crops, crops that do not need to be replanted after harvest. Perennial crops identified include citrus fruit plantation and coffee, which together contribute around 2000 km^2 in 2020 compared to a 1400 km^2 contribution in 1985. Temporary crops include sugar cane, and

soy. Soy is the main contributor to agricultural land use today, accounting for 66% of agricultural land use in the São Francisco River Basin and 3.5% of the total basin area. Soy plantation has increased 40 times in area, going from around 550 km^2 in 1985 to more than 22,500 km^2 in 2020. Scaling with the expansion of agricultural is the increased demand for water and the consequential growth in center-pivot irrigation area, which will be discussed more in Section 3.4.

The growth in soy production in Brazil comes in response to policies of development and international agreements, which encouraged the nation's integration into global grain markets, and transformed Brazil into a major food supplier for Asia (Russo Lopes et al., 2021). The specific growth observed in the São Francisco River Basin comes from its large overlap with the MATOPIBA region, which consists of areas in the states of Mato Grosso, Tocantins, Piaui and Bahia, where agriculture has increased significantly. Considering the international and national pressure to curb deforestation in the Amazon, which historically has been driven by the expansion of agriculture, soy production has been shifting to the Cerrado biome. In the Cerrado the Forest Code requires a 20-35% of land conservation of native vegetation, which contrasts with the zero-tolerance Soy Moratorium in the Amazon, which attempts to prevent deforestation from happening at all (Russo Lopes et al., 2021).
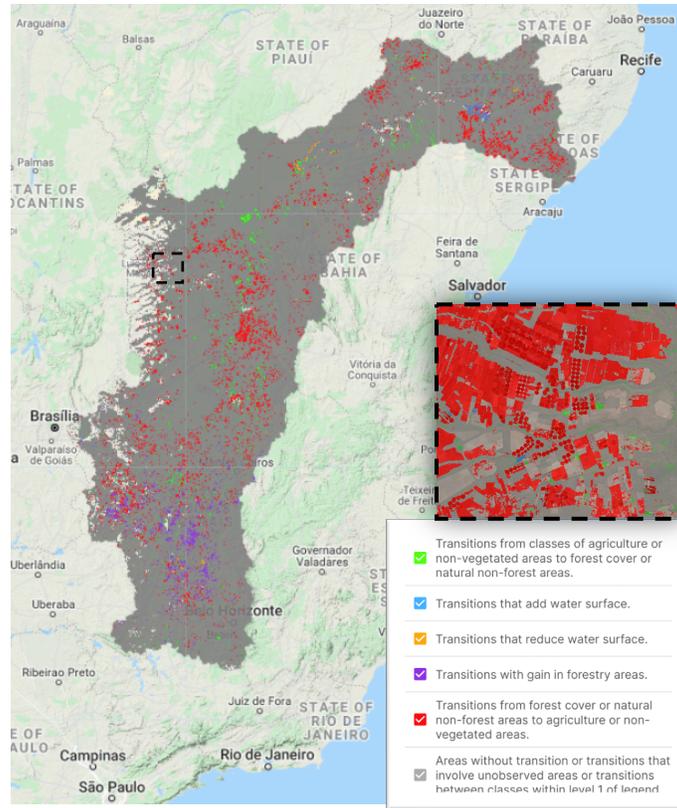
Figure 5: Land use change at the São Francisco River Basin (1985-2020). Dashed black square zooms into a section of Western Bahia that has undergone significant transition to agriculture. There we see a number of center-pivot irrigation systems. Retrieved from MapBiomas Project, 2021.
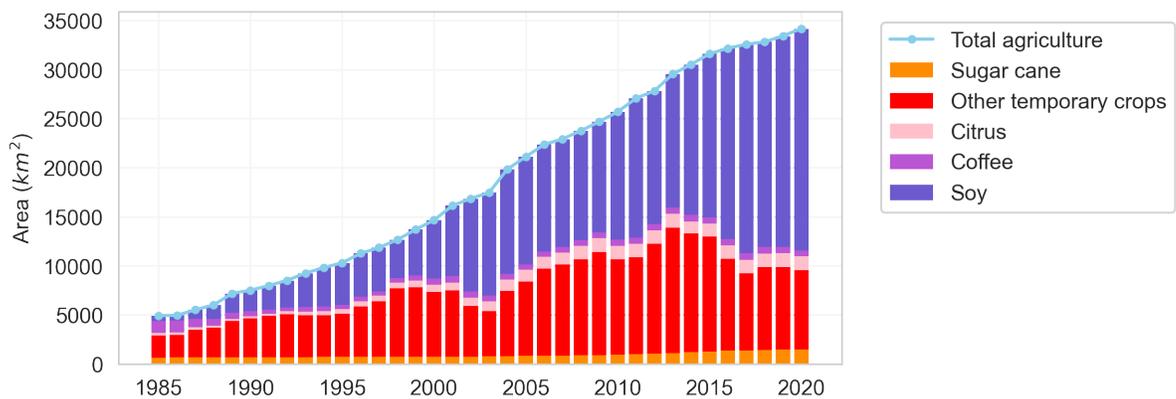

Figure 6: Agricultural land-use and expansion on the São Francisco River basin from 1985-2020.

3.2 Climate

The climate at the São Francisco river basin consists primarily of a dry season (April-September) and a rainy season (October-March), resulting in significant seasonal oscillations in climate parameters. Over the past twenty years changes in climate patterns have also been observed across the basin. Figure 7 illustrates both the seasonality and long-term trends for rainfall, temperature and evapotranspiration using monthly data from GLDAS, which, as discussed in Section 2.2, was geospatially averaged across the entire basin.

When looking at rainfall, Figure 7 (a) and (b), we see that the dry and rainy season are very pronounced. During the driest months, June-August, the average rainfall rate remains below 1 mm/day, with little variability in the data. In August, the driest month, close to no rain is observed each year. In contract, during the rainy season variability significantly increases, and two peaks can be observed in one season. The months of November through March have median rainfall rates of around 5 mm/day, with peaks each year reaching 6 or 7 mm/day. From 2020-2022, the trend was a slight decrease in rainfall rate with a slope of -0.057 mm/month or -0.69 mm/year.

For average air temperature, Figure 7 (c) and (d) show that temperatures in the São Francisco River basin falls within the range of 20-27°C with a yearly oscillation of
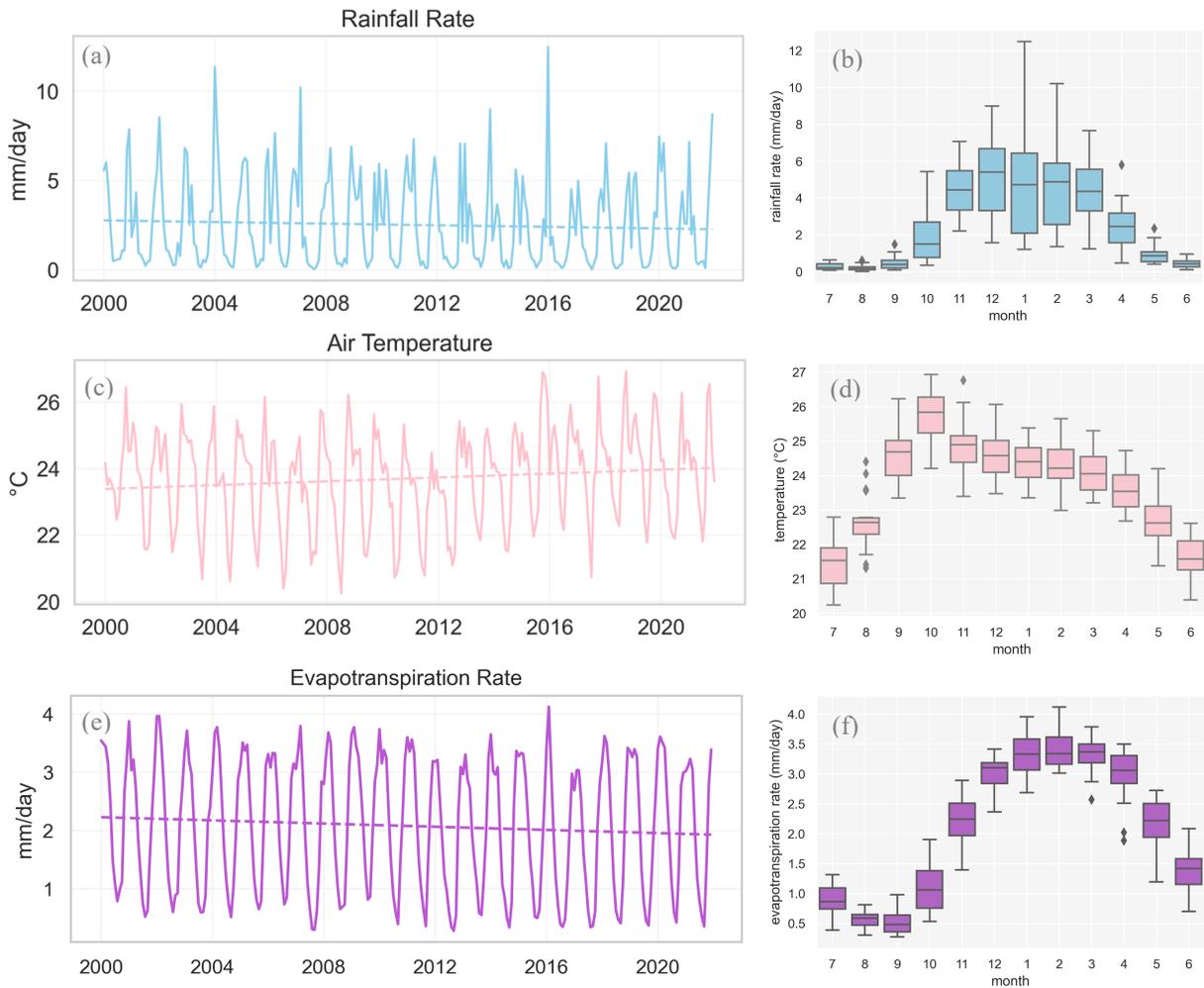
Figure 7: Modeled climate parameters across the São Francisco River Basin (a), (c), (e) monthly average timeseries and trends for rainfall rate, air temperature and evapotranspiration rate from 2000-2020. (b), (d), (f) box-and-whisker plots for each month highlights the seasonal pattern in a typical hemisphere water year (July 1-June 30). The center of the box indicates the median value, the edges the upper and lower quartile, the whiskers indicate the maximum and minimum values, with the exclusion of outliers, which are represented as diamonds.

around 4-5 °C. From the boxplot we see that October has the highest median temperature value across the years of about 26°C with the lowest median temperatures in June and July, at just below 22°C. Generally lower temperatures correspond to periods with lower rainfall rates, and the slightly higher temperatures correspond to the rainy period. A significant increase in average temperature is observed with a trend of 0.0024 °C/month

or 0.29°C/decade since 2000. In comparison the average global temperature increase is of 0.18°C/decade since 1981 (Lindsey & Dahlman, 2021).

The correlations between rainfall and air temperature are in part responsible for the patterns we see in evapotranspiration rates, in Figure 7 (e) and (f). Oscillating between averages of 0.5 and 4.0 mm/day, evapotranspiration rates reach a low in August or September, at the end of the dry season. The peak median evapotranspiration corresponds to the end of the rainy season, in February or March, when there is the greatest water available in the soil for evapotranspiration by plants. Comparing the evapotranspiration and temperature timeseries it was found that, in most years, the peak in evapotranspiration is slightly offset from the peak in temperature by about two to three months each year, which suggests a greater dependence on availability of water than the energy available to evaporate it. Overall, evapotranspiration has been decreasing with a slope of -0.035 mm/month, which translates to -0.42 mm/year. The decrease reflects the overall fall in water available for evapotranspiration.

3.3 Change in water availability

From 1985 to 2020 the São Francisco River Basin experienced a decreasing trend in water surface area of -5km^2 water-area/month or -60km^2 water-area/year. While this long-term trend is only reflective of surface water, in Figure 8 it can be seen how changes in total water storage show a similarly decreasing trend from 2002 to 2020, with an equivalent water thickness loss of -1.8 cm/year. As described in Equation (1), changes in total water storage are representative of the cumulative surface water, groundwater and soil moisture. This means that on the ground observation by a local farmer of a drying stream (See Figure 1 (a)), are also being experienced underground, as soil moisture and groundwater are similarly being depleted.
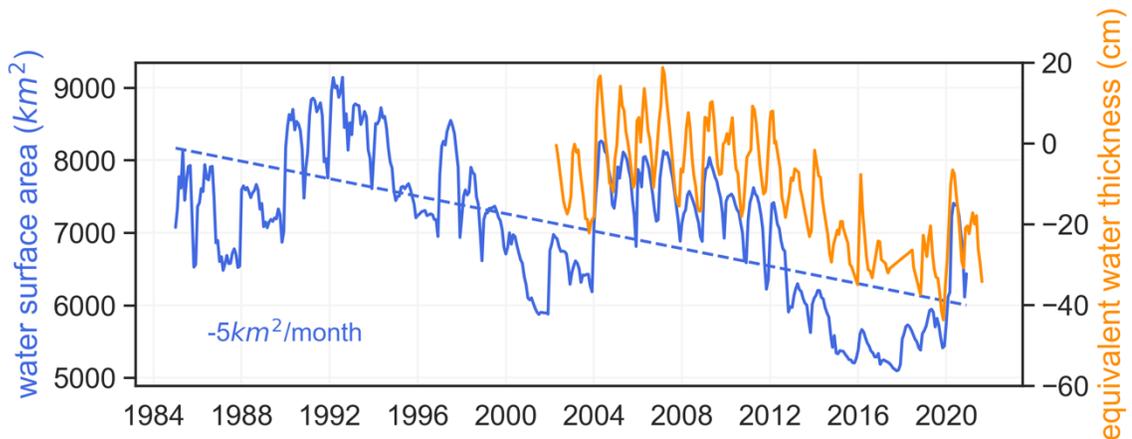
Figure 8: Compares changes in water surface area measurements by MapBiomas (blue) with water mass anomaly measurements by the GRACE satellites (orange). A long-term linear trend on the water surface area data shows an average decrease of 5km^2 of water surface area/month.

The yearly oscillations seen in the climate data are also reflected in water availability. In Figure 8 we see that there are seasonal cycles in surface water area and total water storage which correspond to the dry and rainy seasons characterized in Section 3.2. We also see a more long-term oscillation, particularly in water surface area,

with an overall low around 2002 and again in 2016. To better characterize these long-term cycles, as well as the correlation between the MapBiomas and GRACE data, I took a Fourier transform of both datasets presented in Figure 8 and plotted the resulting Fourier amplitude spectrum in Figure 9. Plotting the spectrum, we see that both graphs agree on the prominent yearly cycle, as well as a long-term 18-year cycle and a 2-year cycle. These cycles captured in both datasets again speak about the connection between surface water observations and changes in total water storage. They also suggest that there might be other long-term cycles, perhaps connected to climate, that will remain outside the scope of exploration of this thesis.
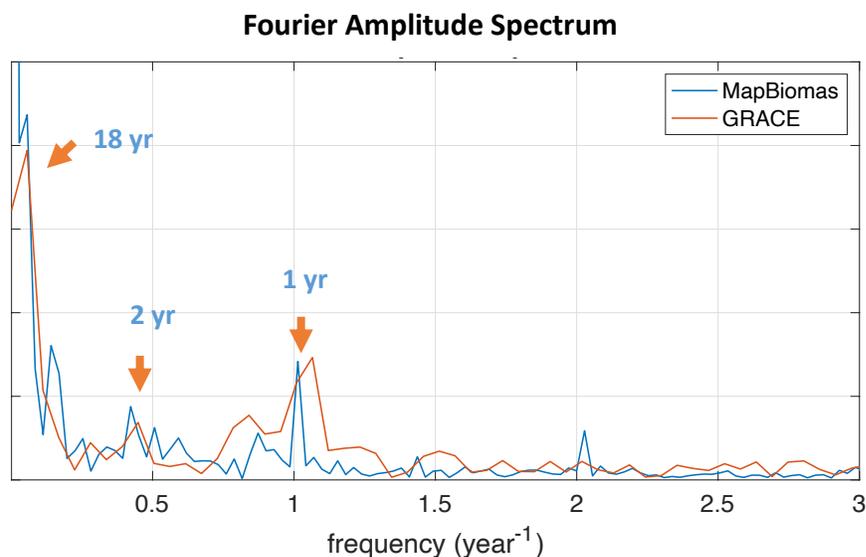
Figure 9: Fourier amplitude spectrum of the timeseries presented in Figure 3. To take the Fourier transform of the GRACE timeseries the data was interpolated to assign values for the missing time-steps, as a measurement is needed at each interval (1 month). The GRACE Fourier spectrum was then rescaled to overlay on top of the MapBiomas spectrum. Both analyses show a 1 year (seasonal), 2 year, and an 18-year cycle.

A direct correlation between total water storage and surface water area is shown in Figure 10. The scatter graph shows a positive trend with a Pearson correlation

coefficient, r of 0.82 and an r^2 of 0.68. This positive correlation confirms that a decrease in water surface area corresponds to a decrease in total water storage. While we only have GRACE observations starting in 2002, this correlation allows us to assume that total water storage has been decreasing since 1985 when the area of surface water measurements begin. The correlation is not strong enough to quantify changes in total water storage using surface area data however that is expected considering the different timescales of surface and groundwater recharge as well as other parameters, such as evapotranspiration, that would be affecting surface water availability but would not be significantly affecting total water storage.
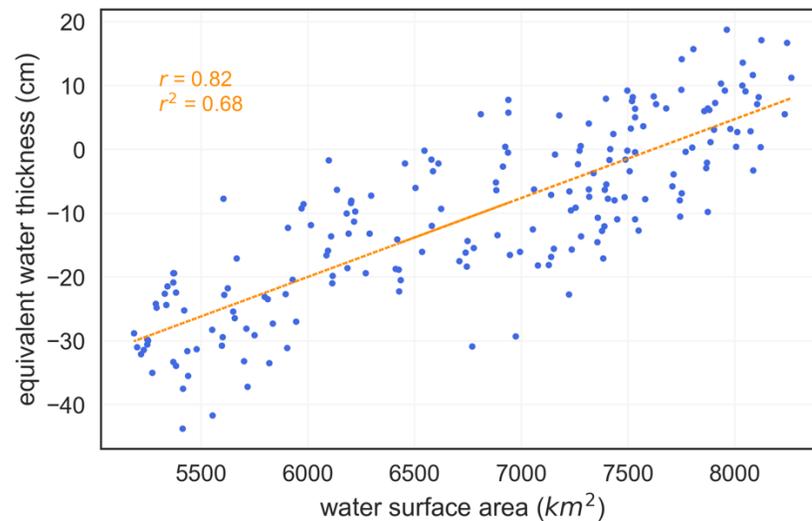

Figure 10: comparison between GRACE's equivalent water thickness and MapBiomas' water surface area. The data points included in this figure correspond to all of the months where GRACE and MapBiomas data availability overlapped from 2002-2020.

3.4 Irrigation water use

As discussed in Section 3.1, agricultural land-use has increased by 7-times since 1985, occupying around 5% of the total basin's area today. Growing agriculture increases need for irrigation water use. In the region, MapBiomas has detected a significant increase in use of center-pivot irrigation systems. As can be seen in Figure 11 (a) center-pivot have increased by 32 times in area since 1985, taking up more than 4,000 km^2 in 2020.

With the detection of center pivot area, we can estimate how much water is being pumped for irrigation each year. As discussed in the methods section, a simple water budget can be used to estimate volume of water needed for irrigation (see Equations (2) and (3)). The results of these calculations are illustrated in Figure 11. In Figure 11 (b) we see the estimated depth needed for irrigation each month, which was computed using GLDAS rainfall and evapotranspiration data. The oscillations clearly follow the seasonal trends: during the rainy season no irrigation water is needed, but during the dry season an average of up to 3mm/day might be needed. Combining estimated water depth needed with the yearly area of center-pivots (Figure 11(a)) we estimate the average volume of water needed per day Figure 11 (c). What is interesting to observe is that the seasonal oscillations persist but there is a general increase in the peaks, given the larger area being irrigated. Aggregating this by year Figure 11 (d) shows us that, on average, the increase in water used for irrigation is around 0.028km^3 water/year, peaking at around

1km^3 of water used in 2020 by center-pivot systems. We also see that the increase includes a lot of interannual variation depending on climatic conditions.
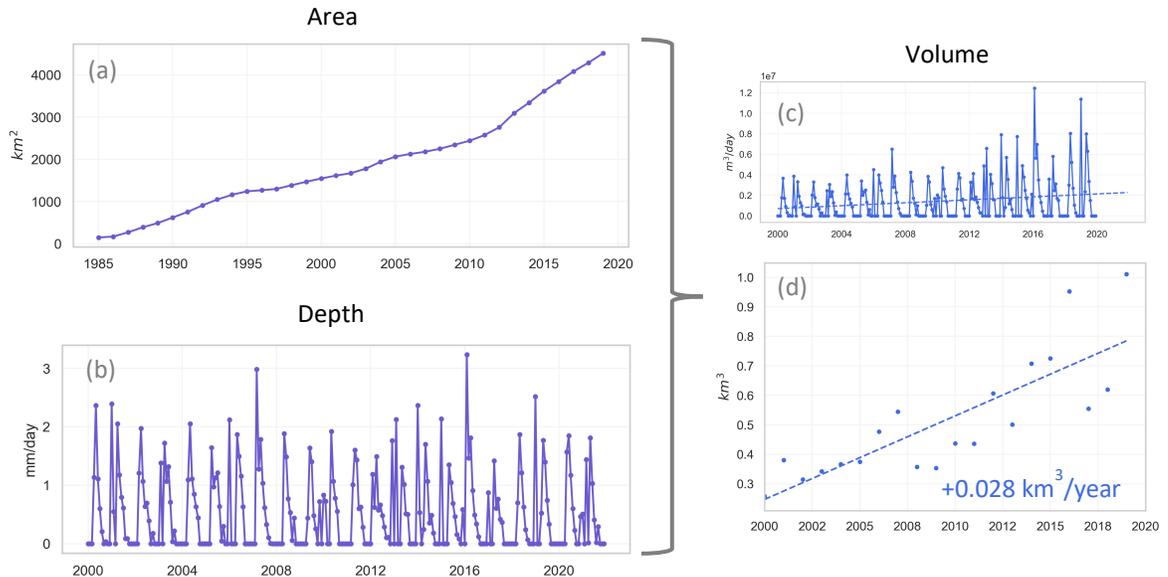
Figure 11: Center pivot irrigation (a) yearly surface area over the basin from 1985-2020 (b) average depth of water used per day for each month 2000-2020, computed from Equation (2) (c) combines data from (a) and (b) to get an average volume/day of water used (d) estimates the total water consumption of water by center pivots each year since 2000 and the rate of increase

The average discharge of the São Francisco River is 2769 m^3/s meaning that 87km^3 of water flows through the river each year (Comitê da Bacia Hidrográfica do São Francisco, 2016). The estimated water used for center-pivot irrigation in the region is around 1.1% of the total river discharge. Groundwater availability is reported as 365 m^3/s, which means that 12 km^3 of water flows in the subsurface each year and is accessible for human consumption (Comitê da Bacia Hidrográfica do São Francisco, 2016). Therefore, if water pumped for center-pivot irrigation was exclusively groundwater, around 8.7% of the total water storage would have been pumped in 2020. In reality, little data is available

determining what proportion of water pumped for center-pivot irrigation comes from surface versus groundwater.

Furthermore, while these comparisons to the river discharge helps us put the estimated irrigation water usage value into greater context, it is important to remember that the analysis does not consider the location where center pivots are concentrated, and the specific aquifers or tributaries from which water is being drawn, for which such a value might be more reflective of local observations in water level. Other limitations of this calculation include not considering the specific water demand of crops being planted, as well as the assumption that planting of crops is happening throughout the year. As seen in Section 3.1, most of the agricultural land is used to plant temporary crops, such that water usage might be less than estimated. Finally, there is a difference in time-resolution in the data with yearly area of center pivots and monthly observations in rainfall and evapotranspiration. If monthly observations of center-pivots were available perhaps a better picture of when they are being used could be captured.

4   The sub-basins: Baixo, Submédio, Médio, Alto São Francisco

In its more than 630,000 km², the São Francisco River Basin encompasses a variety

of ecological and geo-political territories. As we saw in Figure 3 the basin is split into four

sub-basins, the Alto, Médio, Submédio and Baixo São Francisco. Table 2 details some of

the key geographical information of each sub-basin (Magalhães & Martins, 2021).

Table 2: Area and population of the sub-basins

	Area (km²)	% of basin	Biomes	Population
Alto São Francisco	100,076	16	Cerrado, Mata Atlântica	6.2 million
Médio São Francisco	402,351	63	Cerrado, Caatinga	3.2 million
Submédio São Francisco	110,446	17	Caatinga	1.9 million
Baixo São Francisco	25,523	4	Caatinga, Mata Atlântica	1.4 million

The Alto São Francisco, located in the state of Minas Gerais, makes up 16% of the

basin's area but holds most of the population, due to its overlap with the state's capital

Belo Horizonte. It lies on the transition between the tropical and wet Mata Atlântica to

the drier Cerrado and is where most of the river flow originates. The Médio São Francisco

makes up an overwhelming majority of the basin's area, 63%, and carries the second

largest population. As is seen in Figure 4 and 3, it falls on the transition between the

Cerrado and the Caatinga and encompasses a significant part of Minas Gerais and Bahia,

a small part of Goiás, and the Federal District of Brasília. The Submédio exclusively

occupies the Caatinga making it one of the driest regions in the basin. Taking up 17% of

the area it has a relatively small population of 1.9 million inhabitants. The Baixo São

Francisco makes up the remaining 4% of basin area, seeing the transition from the Caatinga back to the Mata Atlântica.

Considering all the differences that uniquely characterize the sub-basins, in this chapter we dive into a similar analysis of trends in agricultural expansion, climate, water availability and finally irrigation water use for the sub-basins. This analysis will allow for a better understanding of some of the similarities and differences across the São Francisco River Basin, allowing for more localized conclusions on the trends observed.
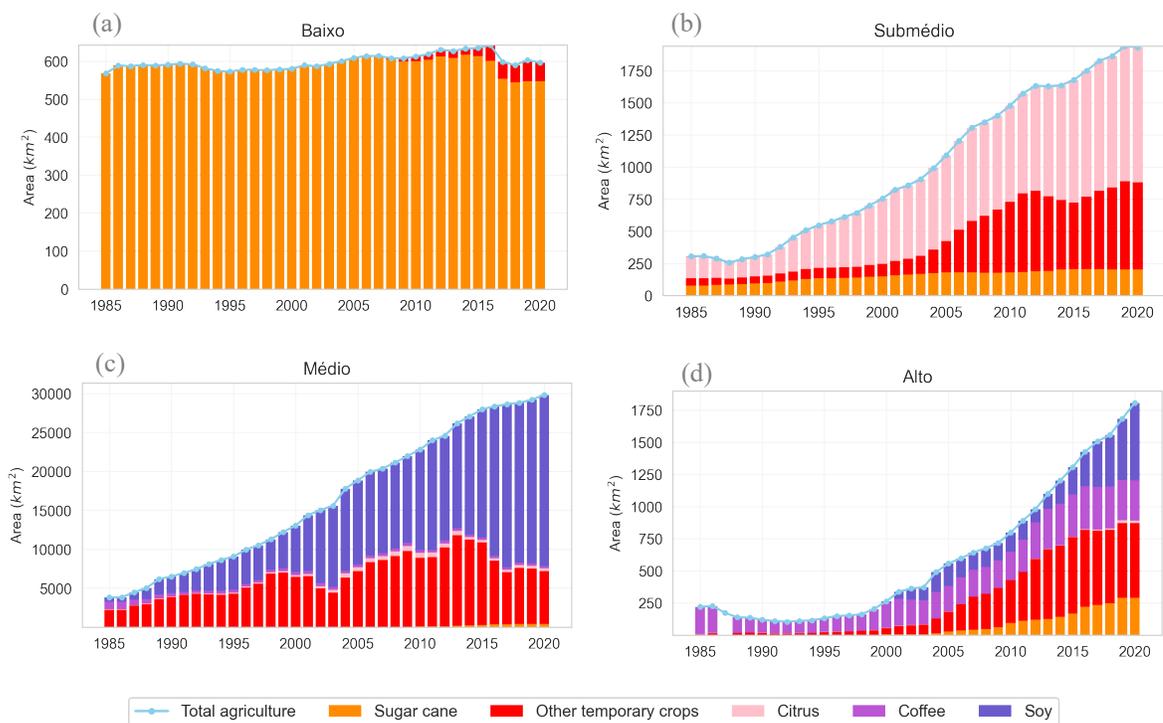
4.1 Agricultural expansion

Figure 12: Agricultural land-use and expansion at (a) Baixo São Francisco (b) Submédio São Francisco (c) Médio São Francisco and (c) Alto São Francisco from 1985-2020

The change in agricultural land-use across each sub-basin, as measured by MapBiomas, can be seen in Figure 12. Trends in agricultural expansion as well as types of crops planted vary significantly across the different sub-basins. First, it is interesting to note that there has been little growth in agricultural production in the Baixo São Francisco, with a relatively stable time-series of around 600 km^2 of primarily sugar cane plantations since the data started being collected in 1985 (Figure 12 (a)). In the Submédio the expansion in agriculture is significant, following the 7-time area growth observed in the whole basin (Figure 12 (b)). We see however, that unlike the whole basin, the transition to agriculture in this region has been led by citrus fruit plantations and other temporary crops. Close to no soy is produced in the region, likely due to the drier climate of the Caatinga. In the Médio we also see a significant increasing trend of agricultural area with the area occupied today 6 times that in 1985 (Figure 12 (c)). Considering the large % area coverage of this sub-region, it is expected that the trends observed in Figure 12 (c) are representative of trends for the whole basin (Figure 6). In fact, 85% of agricultural land-use in the São Francisco River basin is located in the Médio São Francisco. Leading agricultural expansion is soy plantation, along with other temporary crops. The Alto São Francisco shows the greatest overall increase in agricultural land-use, jumping from around 250 km^2 in 1985 to over 1750 km^2 today, also about a 7-time area increase. This sub-basin shows the greatest variety of crops including soy, coffee, other temporary crops and sugar cane. With the exception of coffee plantations, which have remained relatively steady, almost all other plantations have increased over time.

4.2 Climate

As previously discussed, the climate at the São Francisco river basin consists primarily of a dry season and a rainy season, resulting in significant seasonal oscillations in climate parameters, which is consistent throughout the basin. What is interesting is how long-term averages and trends vary across the sub-basins.

Figure 13 shows trends in air temperature from 2000-2020. For the Baixo and Submédio São Francisco sub-basins average air temperature is quite similar (Figure 13 (a) and (b)), oscillating between 22°C in colder months and 28°C in warmer months. In both locations average temperatures are around 25°C and that the increase is +0.012 °C/year at Baixo São Francisco and +0.013 °C/year at Submédio São Francisco. These long-term
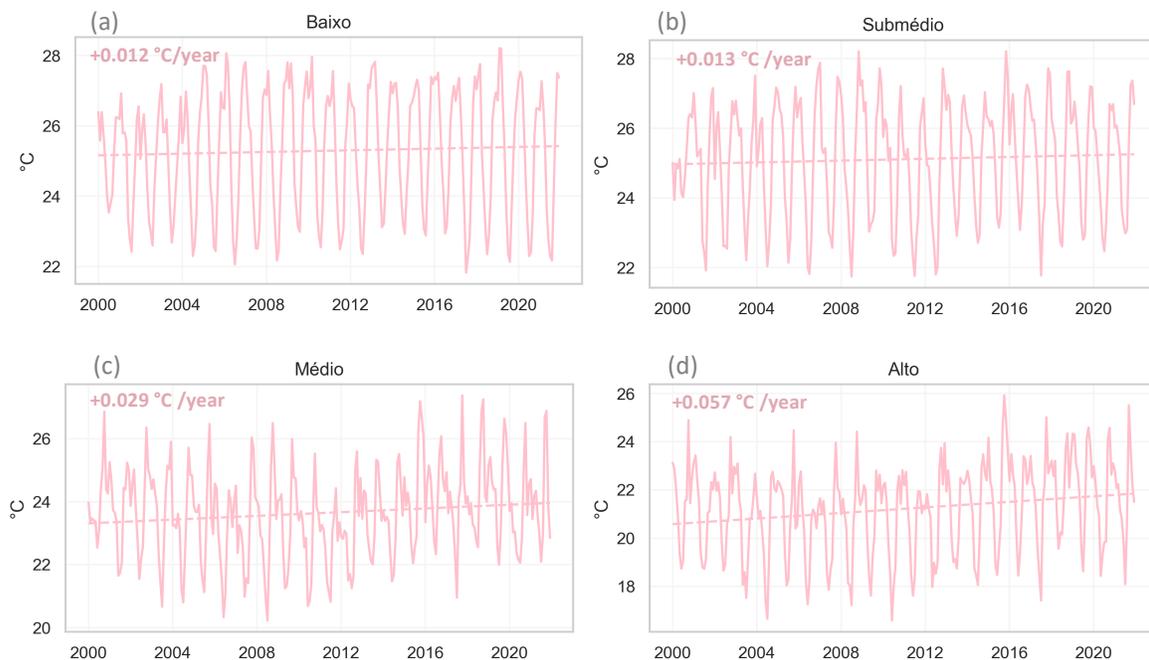
Figure 13: Average monthly air temperature and long-term trends at (a) Baixo São Francisco (b) Submédio São Francisco (c) Médio São Francisco and (c) Alto São Francisco from 2000-2020

trends are below the average increase of +0.029°C/year for the full basin and lower than the global average increase of +0.18°C/decade.

In contrast, at Médio and Alto São Francisco we see lower overall temperatures but more pronounced increases over time. At Médio São Francisco, average temperatures range between 20°C and 27°C, although in the last five years we see that generally the lowest months are above 22°C. The warmer low temperatures today are reflective of an increasing trend of +0.029°C/year, which is quite close to the average observed in the whole basin. This correlation makes sense considering the large area contribution of the Médio sub-basin. The Alto São Francisco sub-basin has temperatures ranging from around 17°C to 26°C in the past twenty years. It also has presented the greatest increase in average temperature of +0.057°C/year, which is more than double the average temperature increase observed over the whole basin, and more than three times the global average decadal increase. In summary, from 2000 to 2020, warming trends appear most pronounced in the Alto São Francisco sub-basin, and least pronounced in the Baixo and Submédio sub-basins.

Figure 14 offers details on rainfall rates across the sub-basins, which mostly oscillate between months of almost no rainfall, at the peak of the dry season, and up to 12.5 mm a day, at the peak of an especially strong rainy season. The Baixo São Francisco sub-basin is an exception. We see in Figure 14 (a) that its highest peak is close to 11 mm/day in early 2004. It is also clear that the sub-basin experiences the least average rainfall, with almost no peaks greater than 6 mm/day past 2012, and the greatest decreasing trend of -1.5 mm/year. Although the Submédio sub-basin shows a few higher

peaks in the 2000s, generally it also has a lower average rainfall with peaks mostly below 5 mm/day after 2010 (Figure 14 (b)). The decreasing trend is significant, with a linear fit of -0.95 mm/year. At the Médio São Francisco sub-basin rainfall has remained more stable in the past twenty years, but still with a decreasing long-term trend of -0.52 mm/year (Figure 14 (c)). On average the rainfall rate is about 2.5 mm/day and peaks are around 7.5mm/day. The Alto São Francisco sub-basin has the highest overall rainfall rate, with peaks generally higher than 7.5 mm/day and more consistently reaching above 10 mm/day (Figure 14 (d)). Still a general decrease in rainfall rates is also observed, with a trend of -0.89 mm/year.

Overall, we see lower rainfall rates at Baixo and Submédio São Francisco sub-basins which are characteristic of the Caatinga, where the climate is semi-arid. It is important to note that the lack of rainfall appears to be intensifying particularly in these drier regions. Higher rainfall rates are seen in the Médio and particularly the Alto sub-basin which cover more of the Cerrado and part of the Atlantic forest, where we would expect more intense rainfall. It is important to note how long-term trends in rainfall compare with the long-term trends in temperature. While temperature seems to be rising more slowly in the Baixo and Submédio basins, rainfall is decreasing more quickly. Rapidly decreasing rainfall rates, despite slower warming, may be attributed to the still overall higher temperatures in the region causing greater evapotranspiration. Higher rainfall rates might also be attributed to global climate oscillations which are beyond the scope of exploration of this thesis.
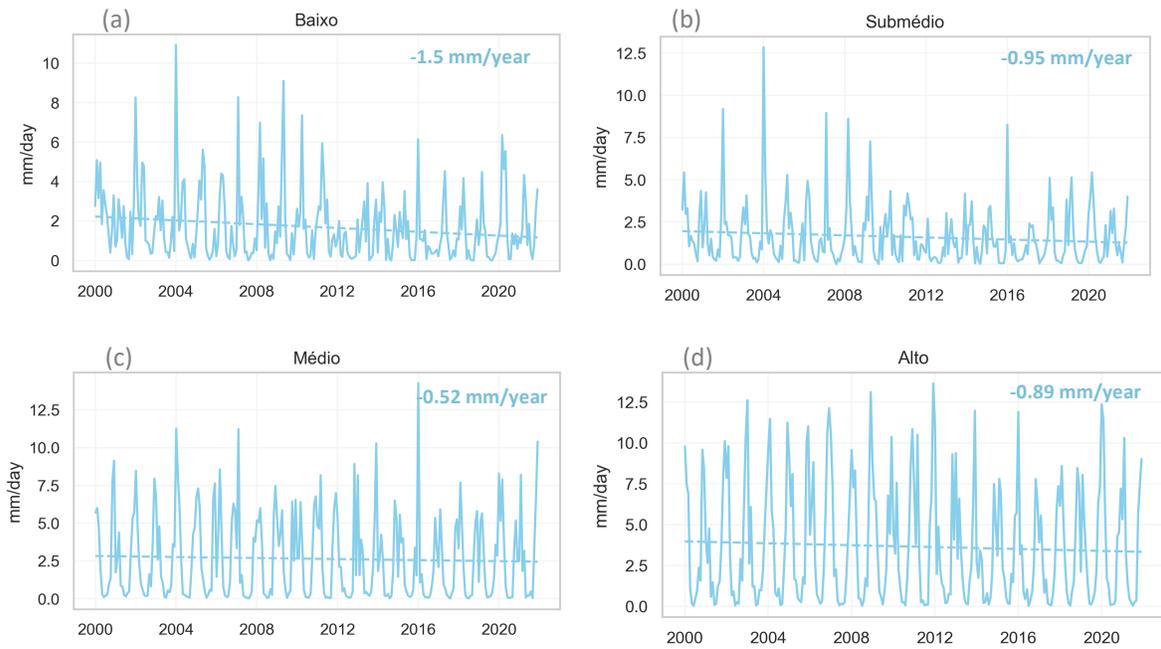
Figure 14: Average rainfall rate and long-term trends at (a) Baixo São Francisco (b) Submédio São Francisco (c) Médio São Francisco and (c) Alto São Francisco from 2000-2020

4.3 Change in water availability
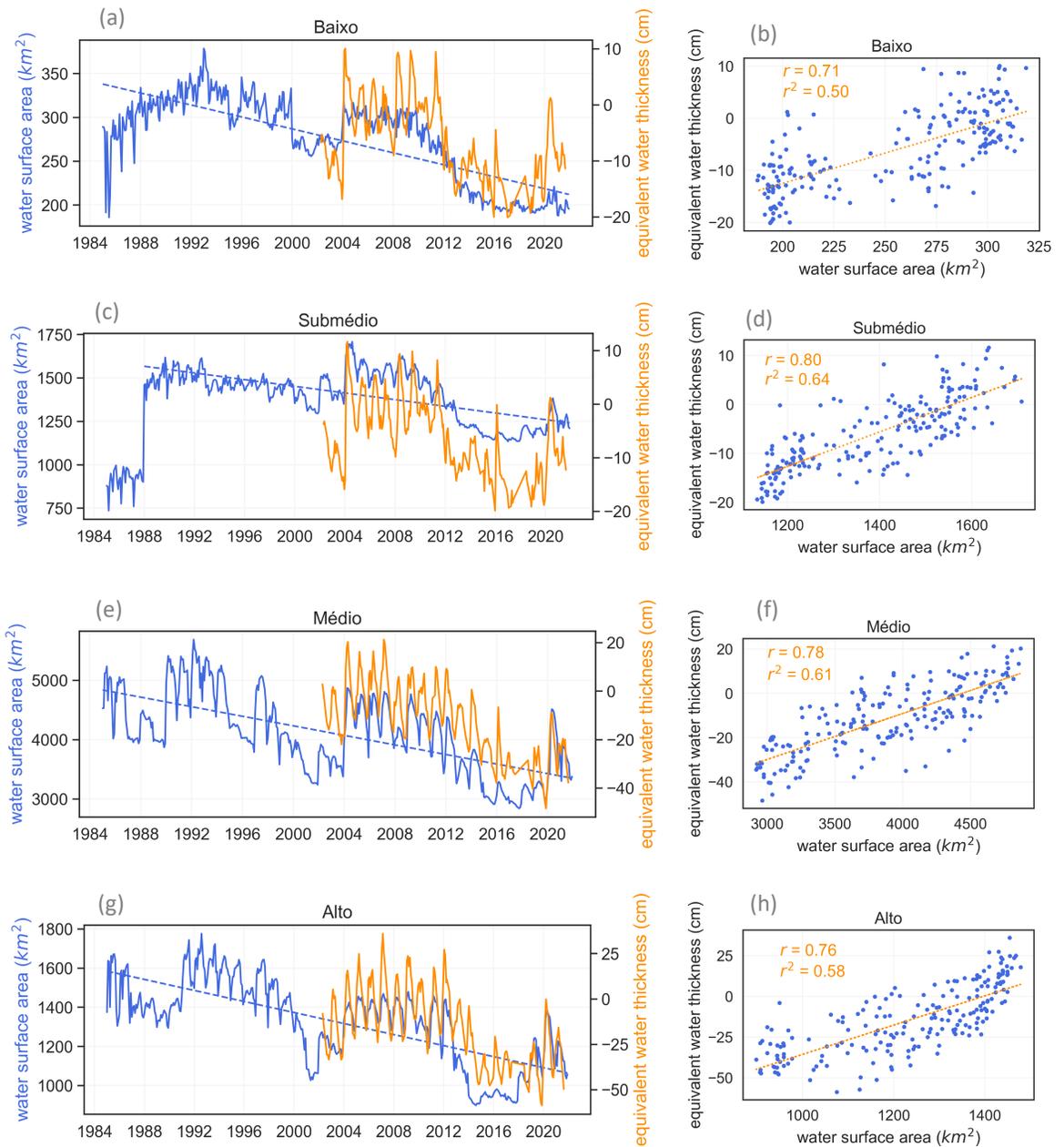

Figure 15: Changes in water surface area measurements by MapBiomas (blue) in comparison water mass anomaly measurements by the GRACE satellite (orange) for (a) (b) Baixo São Francisco, (c) (d) Submédio São Francisco, (e) (f) Médio São Francisco, (g) (h) Alto São Francisco. Left panels show trends in water storage, right panels show strength of correlation between the datasets.

In this section we characterize and compare long-term trends in water surface area and total water storage across the sub-basins. Figure 15 compiles long-term trends and correlation plots between the two datasets. In all sub-basins we see decreasing trends in water availability and a significant positive correlation between water surface area and total water storage. The correlation indicates that changes in water surface area, which is more often observed by locals, are reflective of overall total water storage changes, which likely indicate a decrease in soil moisture and groundwater, as well as surface water. We see a clear seasonal signal in response to the dry and rainy seasons. There are also clear long-term oscillations, similar throughout the sub-basins, with a general dip around between 2002 and 2004, and another major low in 2016.

For the Baixo São Francisco sub-basin, Figure 15 (a) plots water surface area trends since 1985, which decreases on average -3km^2/year. The equivalent water thickness, seen in the same figure, has been decreasing by -0.70 cm/year since 2002, less than 40% of the -1.8 cm/year decrease seen overall in the basin. In Figure 15 (b) we see a positive correlation with coefficient r = 0.71 and r^2 = 0.50 between GRACE observations and MapBiomas' water surface area, the lowest in the sub-basins. A reason for it, and potential limitation of the GRACE analysis of the Baixo São Francisco sub-basin, is the small area of the basin, 25,523 km^2. Given the resolution of the GRACE observations, of 3-degree by 3-degree (approximately 100,000 km^2), the area of this sub-basin is significantly smaller than the native resolution of a single pixel of data, such that the data plotted is likely capturing signals from nearby regions. In contrast the MapBiomas data has a resolution of 30 m x 30 m such that it precisely captures the changes in the specific

area. These varying resolutions create a greater discrepancy between the total water storage and surface water signals.

At the Submédio São Francisco sub-basin, Figure 15 (c) plots water surface area from 1985 to 2020. A significant jump in area from around 9,000 km^2 in 1987 to 14,500 km^2 in 1988 which significantly flattens the generally decreasing trend. This jump correlates with the construction of the third largest dam on the São Francisco river, the Itaparica dam, completed in 1988 to store 10,800 million m^3 of water (Magalhães & Martins, 2021). Considering this artificial increase in water surface area, the linear trend was computed starting in 1988 and a more pronounced decrease of -9.6 km^2/year found. The equivalent water thickness, seen in the same figure, has a decreasing trend of -0.72 cm/year since 2002, which is similar to the trend observed at the Baixo sub-basin, considering that observations for the two basins likely overlap given large pixel resolution. Interesting to note is that the correlation between the total water storage changes and surface water area is the highest at the Submédio São Francisco sub-basin, with a correlation coefficient r = 0.80 and r^2 = 0.64 (Figure 15 (d)).

The Médio São Francisco sub-basin has the largest surface area of water, since it is the largest sub-basin. Figure 15 (e) plots the linear trend in water surface area which shows a decrease of -40 km^2/year, as well as the decreasing total water storage, averaging at 2.1 cm/year across the sub-basin. The decrease in total water storage is slightly higher than the total basin average (1.8 cm/year), indicating a greater overall loss of water in the region. The correlation between the total water storage changes and surface water area is significant, although slightly lower than at the Submédio sub-basin, with a coefficient r

= 0.78 and r^2 = 0.61 (Figure 15 (f)). Considering that we see evidence of more large-scale agriculture in the region, indicated by the growing soy production (Section 4.1), it is more likely that groundwater is an important source of water for irrigation, which might create more of a discrepancy between surface water area observations and total water storage signals.

At the Alto São Francisco sub-basin, Figure 15 (g) shows that decreasing water availability is characterized by a trend of -14 km^2/year for surface water area and total water storage trend of -2.2 cm/year, the steepest in the sub-basins. The correlation between the datasets is a bit lower but still significant with r = 0.76 and r^2=0.58 (Figure 15 (f)).

Overall, the main difference between the sub-basins are the steeper trends in total water storage in the Alto and Médio sub-basins, compared to Baixo and Submédio São Francisco. Figure 16 visualizes the spatial variation in total water storage across the basin by plotting the standard deviation of each GRACE pixel from 2002-2020. We see that, as we move south along the basin, there is greater variation, which corresponds to a decrease in total water storage. At the Baixo and Submédio sub-basins, the standard deviation is less than 8 cm of equivalent water thickness. In contrast, in parts of the Alto and Médio sub-basins standard deviations rise up to 22 cm, which is a very significant oscillation. Comparing Figure 16 to Figure 4, which shows the overlap between the sub-basins and the biomes, we see that larger oscillations in total water storage are happening more in the Cerrado, and less in the Caatinga, with median values of around 14 to 16 cm in the transition between the biomes within the Médio São Francisco sub-basin.
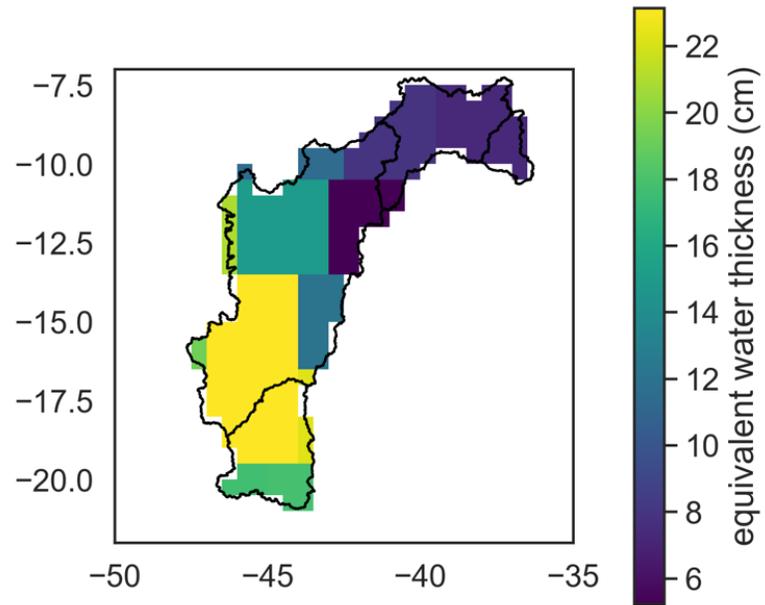
Figure 16: Variation in total water storage across the São Francisco River basin. Scale represents standard deviation of equivalent water thickness as determined by GRACE observations between 2002-2020. Map outlines the division of the sub-basins.

4.4 Irrigation water use

To close this Chapter, Section 4.4 analyzes the role of center-pivot irrigation in each sub-basin, estimating irrigation water use. This analysis follows the same calculations present for the whole basin in Sections 3.4, combining land-use data, climate re-analysis data for estimating a depth of water needed for irrigation and combining both to estimate a volume of water used. In Figure 17 we see the breakdown of center-pivot irrigation area compared across the basins. The breakdown highlights that almost all of the center-pivots are located in the Médio São Francisco sub-basin, with the Alto also making a noticeable contribution, but almost negligible area of center-pivots in the Submédio and Baixo São Francisco in comparison to the rest of the basin.
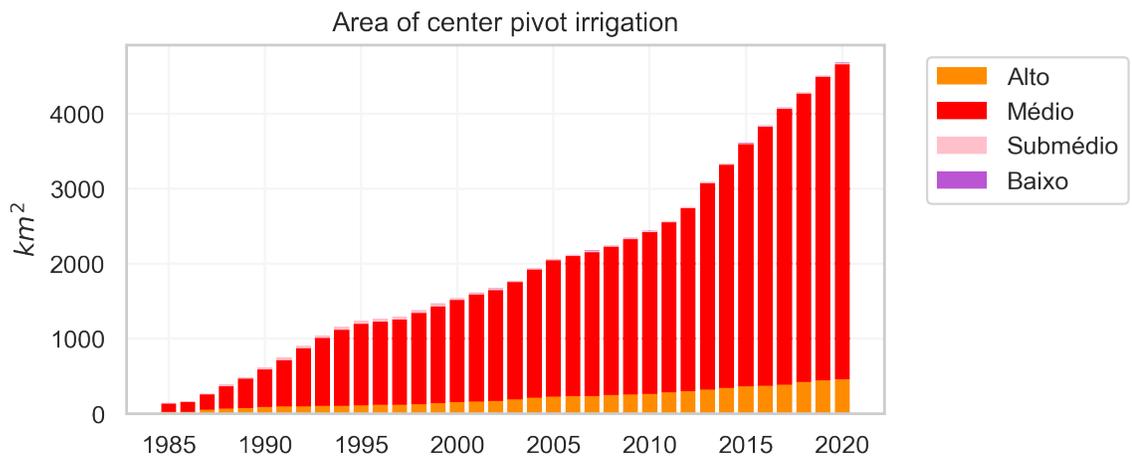
Figure 17: Comparison of center-pivot irrigation area across the sub-basins, 1985-2020

Figure 18 breaks down the trend in center-pivot irrigation area for each sub-basin, coupling them with the estimated trends in total water used for irrigation each year. At the Baixo São Francisco sub-basin we see a lack of data for the period between 1986 and 1988, with a generally increasing trend after that (Figure 18 (a)). The area goes from less than 1 km² in 1988 to around 5.6 km² in 2020, which is still quite small for a 25,523 km²

sub-basin. This increase happens in jumps, with a particularly big one between 2018 and 2020, suggesting coordinated efforts, like for example a large investment in center-pivots by agribusiness. This translates to an overall increase in volume of water pumped for center-pivot irrigation over the years. In Figure 18 (b) we see that the yearly increase in volume of water pumped is estimated around 0.00001 km^3/year or 10,000 m^3/year. However, perhaps because of the recent jump in center pivot irrigation area, this correlation is very weak, with an r = 0.29. If the center-pivot irrigation area remains at 2020 levels or continues to increase the volume of irrigation could become a greater cause for concern. However, considering how stable agricultural land-use has been in the region (Figure 12 (a)) that is perhaps not to be expected at the moment.

At the Submédio São Francisco sub-basin, center-pivot irrigation area shows a more unique trend, with a peak in 1995 of around 45 km^2 followed by a significant decrease to around 20 km^2 today (Figure 18 (c)). Again, these are very small areas compared to the size of the sub-basin of 110,446 km^2. The decrease in center-pivot irrigation areas leads to a general decrease in the estimated volume of water pumped for irrigation, with a trend of -0.00012 km^3/year as seen in Figure 18 (d). What is important to remember however, is that agricultural production, particularly production of citrus fruits has been growing in the region. This suggests that that there might be other forms of irrigation that could be important to account for but are not being considered in these calculations. The exclusion of alternative forms of irrigation presents a significant limitation to the estimates.

At the Médio and Alto São Francisco sub-basins center-pivot irrigation area and associated water usage follow a more similar trend to that seen in the whole basin, with increases of different magnitudes. At the Médio sub-basin center-pivots have increased by more than 37 times since 1985, and currently occupy more than 4,000 km^2 or around 1% of the total basin area (Figure 18 (e)), an increase associated with the growth of commercial soy production. This leads to around 1km^3 of water used for irrigation per year today, with variation depending on need. The sub-basin also sees the greatest increase in estimated volume of water used for irrigation, with a trend of + 0.033 km^3/year (Figure 18 (f)), a volume even greater than what was estimated for the whole basin, likely because the average water deficiency (estimated depth of water needed for irrigation) across the whole basin was lower. This finding presents another limitation of this calculation, which is the spatial averaging of climate parameters. A more accurate estimation of water usage would consider geo-location of the center pivots and use more localized rainfall and evapotranspiration values to estimate how much water is needed in the specific location of each center pivot.

At the Alto São Francisco sub-basin center-pivots have increased by more than 18 times since 1985, and currently occupy more than 400 km^2 (Figure 18 (g)). This leads to around 0.1km^3 of water used for irrigation per year today, an order of magnitude less water than used at the Médio sub-basin. The sub-basin also sees an increase in estimated volume of water used for irrigation, with a trend of +0.0031 km^3/year (Figure 18 (h)).
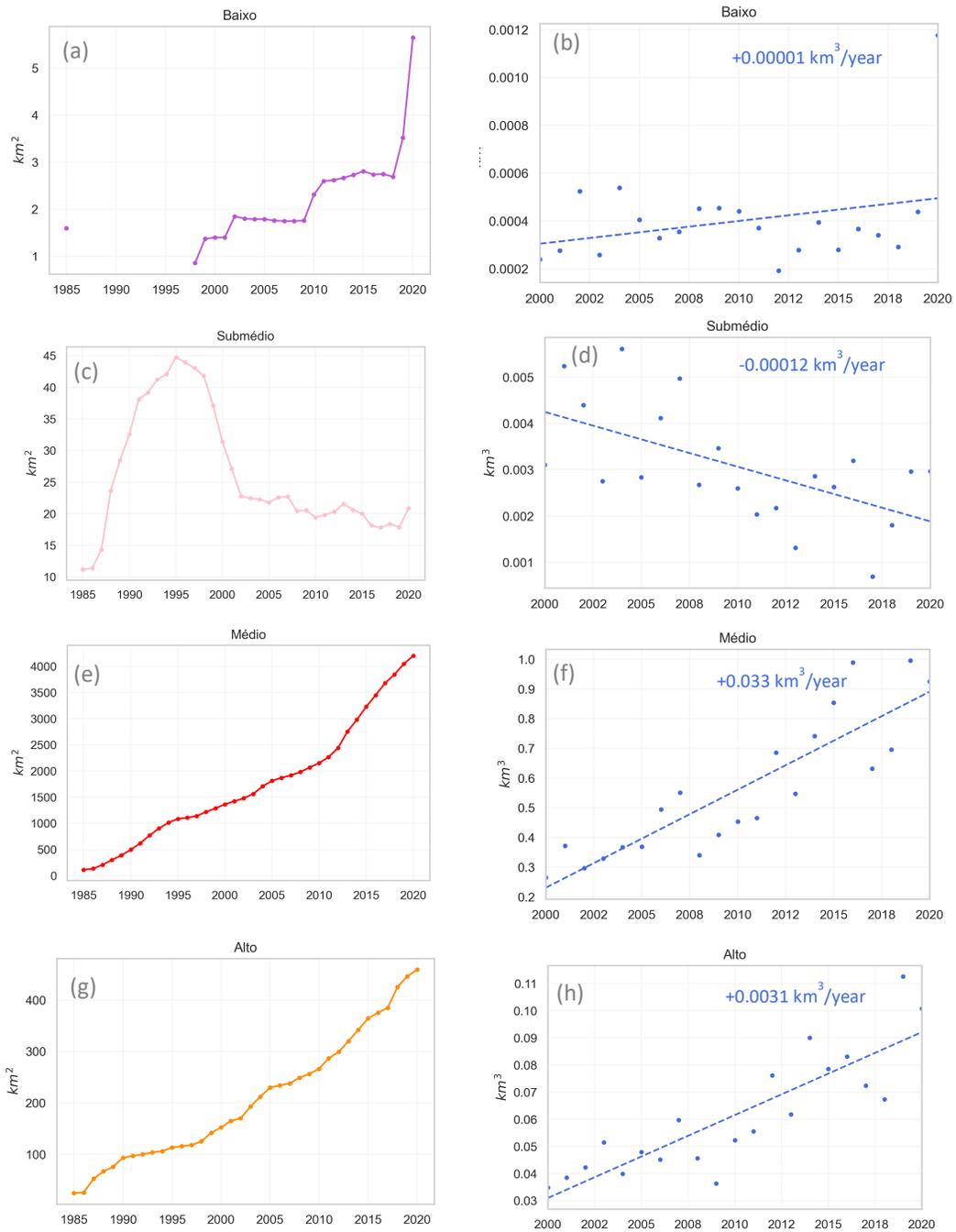
Figure 18: Center-pivot irrigation surface area, 1985-2020 (left) compared to estimated total volume of water pumped for center-pivot irrigation each year and rate of change, 2000-2020 (right) for (a) (b) Baixo São Francisco, (c) (d) Submédio São Francisco, (e) (f) Médio São Francisco, (g) (h) Alto São Francisco.

Overall, we see that the estimated water used for irrigation using center-pivot irrigations is greater at the Médio and Alto sub-basins, where center-pivots have become more popular since 1985. The increased use of water for irrigation can also be associated with large-scale soy and other temporary crops in the lower parts of the basin, overlapping with the Cerrado. Findings suggest that this might not be an accurate analysis or even proportional measure of irrigation water use at the Baixo and Submédio sub-basins since center-pivot irrigation does not occupy a large area, but we still see an increase in agriculture in the Submédio sub-basin. Future research should consider other irrigation types in the region.

5   Conclusion and Outlook

To reiterate, the goals of this thesis were to:

    (1) Quantify expansion of agriculture and center-pivot irrigation

    (2) Characterize natural climate variability and identify any changes in climate patterns

    (3) Combine different datasets that quantify changes in water availability over time and robustly characterize these changes

    (4) Estimate increased center-pivot irrigation water use

    (5) Identify areas of the basin that are most affected by agricultural growth and changes in water availability

The research was successful in achieving these goals in its analysis of data for the São Francisco River Basin and its sub-basins. Combining remote-sensing datasets and climate re-analysis data this thesis shows how agricultural expansion and climate change correlate to a decrease in water availability in the region.

In summary it was found that, in all sub-basins, with the exception of the Baixo São Francisco sub-basin, agricultural land use is expanding. In the Médio and Alto sub-basins the agricultural expansion is primarily driven by large-scale soy plantation. At the same time, across the whole basin there is an increase in temperature and decrease in rainfall rate. The increase in temperature is particularly significant in the Alto and Médio São Francisco River sub-basins, while the decreasing rainfall is most concerning in the Baixo and Submédio sub-basins. As consequence, water availability is decreasing all across the basin. Decrease in surface water area, observed remotely and locally, is

correlated to a decrease in total water storage. The concern lies especially in the Médio and Alto sub-basins, within the Cerrado biome, where we see the greatest change in total water storage. These regions are also where we observed the greatest growth in center-pivot irrigation which, combined with climate data, allowed us to estimate a significant increase in water used for irrigation. While quite simple and carrying several limitations discussed throughout the analysis, irrigation water use calculations served to close the agriculture-climate-water nexus, underscoring how the sub-basins with greatest agricultural expansion and increased temperatures are experiencing a heightened decrease in water availability.

The findings of this thesis show that the growth of large-scale agriculture comes at the price of increased use of freshwater resources. This work calls for better monitoring of water resources and policies that take the often-ignored cost of agriculture, water, into account. It is also important to remember that depleting water resources typically lead to inequalities in water access. While out of the scope of this thesis to analyze the socio-political relations between agricultural expansion and water availability, to better understand the implications of this research, related work should consider the following:

- What motivates agricultural expansion? Who is benefiting from it?

- Who is most impacted by LUCs and changes in water availability?

- What is the role of the state in both agricultural expansion and water management?

- How can scientific findings, like the ones presented in this thesis, inform water management and land-use practices and/or policies?

Within the field of this thesis there are also a number of avenues for future research. For example, the use of more local datasets, such as groundwater well data from the Rede Integrada de Monitoramento das Águas Subterrâneas (RIMAS) could be incorporated into the water availability analysis. Similarly, land-grid scaling could be applied to the GRACE dataset to provide higher resolution estimates of total water storage changes. These research avenues would complement the sub-basin level findings of this research by adding a more local understanding of water depletion. As previously discussed there a number of limitations to the estimates of irrigation water use in this research. A better estimate would consider the specific location of center pivots and use more local climate data, or soil moisture data, to model specific crops water needs. Furthermore, the framework and datasets used in this thesis, innovatively combining data for climate, agriculture and water could be replicated in other relevant locations.

Bibliography

ANA. (2022, March 29). Regiões Hidrográficas (CNRH). ANA. Retrieved from
https://dadosabertos.ana.gov.br/datasets/5e520cacd27a4d0888d185c74c5b762
b_4/about

Andrade, M. P. de, & Ribeiro, C. B. de M. (2020). Impacts of land use and cover change
on Paraíba do Sul whatershed streamflow using the SWAT model. *RBRH*, *25*, e12.
https://doi.org/10.1590/2318-0331.252020190034

Bagley, J. E., Desai, A. R., Harding, K. J., Snyder, P. K., & Foley, J. A. (2014). Drought and
Deforestation: Has Land Cover Change Influenced Recent Precipitation Extremes
in the Amazon? *Journal of Climate*, *27*(1), 345–361. https://doi.org/10.1175/JCLI-
D-12-00369.1

Comitê da Bacia Hidrográfica do São Francisco. (2016). *Plano de Recursos Hídricos da
Bacia Hidrográfica do Rio São Francisco 2016-2025*. Retrieved from
https://2017.cbhsaofrancisco.org.br/wp-content/uploads/2016/08/PRH-
SF_Apresentacao_26ago16.pdf

Cunha, A. P. M. A., Alvalá, R. C. S., Kubota, P. Y., & Vieira, R. M. S. P. (2015). Impacts of
land use and land cover changes on the climate over Northeast Brazil: Impacts of
land use and land cover changes on climate. *Atmospheric Science Letters*, *16*(3),
219–227. https://doi.org/10.1002/asl2.543

Daniel Cooley, Reed M. Maxwell, & Steven M. Smith. (2021, November). *Center Pivot
Irrigation Systems and Where to Find Them: A Deep Learning Approach to
Provide Inputs to Hydrologic and Economic Models*.

Daniel Grossman. (2021, November 10). Water War: Is Big Agriculture Killing Brazil's
Traditional Farms? [Yale Environment 360]. Retrieved December 12, 2021, from
https://e360.yale.edu/features/with-traditional-farms-withering-why-is-brazil-
running-dry

Gonçalves, R. D., Stollberg, R., Weiss, H., & Chang, H. K. (2020). Using GRACE to quantify
the depletion of terrestrial water storage in Northeastern Brazil: The Urucuia
Aquifer System. *Science of The Total Environment*, *705*, 135845.
https://doi.org/10.1016/j.scitotenv.2019.135845

IBGE. (2010). Estados do Brasil. Retrieved from
http://www.codegeo.com.br/2013/04/shapefiles-do-brasil-para-download.html

IBGE. (2014). Biomas do Brasil. Retrieved from
https://hub.arcgis.com/maps/inea::biomas-do-brasil-15-000-000/about

Lindsey, R., & Dahlman, L. (2021, March 15). Climate Change: Global Temperature
[Climate.gov]. Retrieved March 29, 2022, from https://www.climate.gov/news-
features/understanding-climate/climate-change-global-temperature

Lunetta, R. S., Knight, J. F., Ediriwickrema, J., Lyon, J. G., & Worthy, L. D. (2006). Land-
cover change detection using multi-temporal MODIS NDVI data. *Remote Sensing
of Environment*, *105*(2), 142–154. https://doi.org/10.1016/j.rse.2006.06.018

Magalhães, A. R., & Martins, E. S. P. R. (2021). The São Francisco River Basin. In J.
Schmandt, A. Kibaroglu, R. Buono, & S. Thomas (Eds.), *Sustainability of*

Engineered Rivers In Arid Lands (1st ed., pp. 132–151). Cambridge University Press. https://doi.org/10.1017/9781108261142.011

MapBiomas. (2021, August). Water surface in Brazil reduces 15% since the early 1990s. Retrieved October 6, 2021, from https://mapbiomas.org/en/superficie-de-agua-no-brasil-reduz-15-desde-o-inicio-dos-anos-90?cama_set_language=en

MapBiomas Project. (2021). Collection 6 of the Annual Land Use Land Cover Maps of Brazil. Retrieved from https://code.earthengine.google.com/?accept_repo=users%2Fmapbiomas%2Fuser-toolkit&scriptPath=users%2Fmapbiomas%2Fuser-toolkit%3Amapbiomas-user-toolkit-lulc.js

Marques, E. A. G., Silva Junior, G. C., Eger, G. Z. S., Ilambwetsi, A. M., Raphael, P., Generoso, T. N., et al. (2020). Analysis of groundwater and river stage fluctuations and their relationship with water use and climate variation effects on Alto Grande watershed, Northeastern Brazil. *Journal of South American Earth Sciences*, *103*, 102723. https://doi.org/10.1016/j.jsames.2020.102723

NASA Jet Propulsion Laboratory. (n.d.). Monthly Mass Grids - Global mascons (JPL RL06_v02). Retrieved April 1, 2022, from https://grace.jpl.nasa.gov/data/get-data/jpl_global_mascons/

NASA Land Data Assimilation System. (2022, March 10). GLDAS: Project Goals. Retrieved April 3, 2022, from https://ldas.gsfc.nasa.gov/gldas

NumPy Developers. (2022). numpy.corrcoef. Retrieved April 3, 2022, from https://numpy.org/doc/stable/reference/generated/numpy.corrcoef.html

Oliveira, P. T. S., Lucas, M. C., de Faria Godoi, R., & Wendland, E. (2021). Groundwater recharge and sustainability in Brazil. In *Global Groundwater* (pp. 393–407). Elsevier. https://doi.org/10.1016/B978-0-12-818172-0.00029-3

Rateb, A., Scanlon, B. R., Pool, D. R., Sun, A., Zhang, Z., Chen, J., et al. (2020). Comparison of Groundwater Storage Changes From GRACE Satellites With Monitoring and Modeling of Major U.S. Aquifers. *Water Resources Research*, *56*(12). https://doi.org/10.1029/2020WR027556

Resende, T. C. (2017, January 29). World Rivers. Retrieved from http://ihp-wins.unesco.org/layers/geonode:world_rivers

RIMAS. (2021, October 7). RIMAS: Rede Integrada de Monitoramento das Águas Subterrâneas. Retrieved October 18, 2021, from http://rimasweb.cprm.gov.br/layout/

Russo Lopes, G., Bastos Lima, M. G., & Reis, T. N. P. dos. (2021). Maldevelopment revisited: Inclusiveness and social impacts of soy expansion over Brazil's Cerrado in Matopiba. *World Development*, *139*, 105316. https://doi.org/10.1016/j.worlddev.2020.105316

Santos, A. B., Heil Costa, M., Chartuni Mantovani, E., Boninsenha, I., & Castro, M. (2020). A Remote Sensing Diagnosis of Water Use and Water Stress in a Region with Intense Irrigation Growth in Brazil. *Remote Sensing*, *12*(22), 3725. https://doi.org/10.3390/rs12223725

Saraiva, M., Protas, É., Salgado, M., & Souza, C. (2020). Automatic Mapping of Center Pivot Irrigation Systems from Satellite Images Using Deep Learning. *Remote Sensing*, *12*(3), 558. https://doi.org/10.3390/rs12030558

Scanlon, B. R., Longuevergne, L., & Long, D. (2012). Ground referencing GRACE satellite estimates of groundwater storage changes in the California Central Valley, USA: GRACE GROUNDWATER DEPLETION CENTRAL VALLEY CA. *Water Resources Research*, *48*(4). https://doi.org/10.1029/2011WR011312

Scanlon, B. R., Faunt, C. C., Longuevergne, L., Reedy, R. C., Alley, W. M., McGuire, V. L., & McMahon, P. B. (2012). Groundwater depletion and sustainability of irrigation in the US High Plains and Central Valley. *Proceedings of the National Academy of Sciences*, *109*(24), 9320–9325. https://doi.org/10.1073/pnas.1200311109

Souza, C. M., Z. Shimbo, J., Rosa, M. R., Parente, L. L., A. Alencar, A., Rudorff, B. F. T., et al. (2020). Reconstructing Three Decades of Land Use and Land Cover Changes in Brazilian Biomes with Landsat Archive and Earth Engine. *Remote Sensing*, *12*(17), 2735. https://doi.org/10.3390/rs12172735

Thomas, B. F., & Famiglietti, J. S. (2019). Identifying Climate-Induced Groundwater Depletion in GRACE Observations. *Scientific Reports*, *9*(1), 4124. https://doi.org/10.1038/s41598-019-40155-y

USGS. (2021, October 8). USGS Groundwater Data for the Nation. Retrieved October 8, 2021, from https://waterdata.usgs.gov/nwis/gw

Appendix 1

Details all trendlines and correlation coefficients. Note: to find rainfall and evapotranspiration trends, the slope was multiplied by 30.5, the average number of days in a month, considering that one average, per day, data point was available for each month.

Rainfall (mm/day)

	Linear Trend	mm/month	mm/year	mm/decade
SF	y=-0.001873x+(2.771199)	-0.0571265	-0.685518	-6.85518
Baixo	y=-0.004025x+(2.226997)	-0.1227625	-1.47315	-14.7315
Submédio	y=-0.002588x+(1.955995)	-0.078934	-0.947208	-9.47208
Médio	y=-0.001425x+(2.832178)	-0.0434625	-0.52155	-5.2155
Alto	y=-0.002428x+(3.975890)	-0.074054	-0.888648	-8.88648

Temperature (°C)

	Linear Trend	°C/month	°C/year	°C/decade
SF	y=0.002414x+(23.387848)	0.002414	0.028968	0.28968
Baixo	y=0.000992x+(25.163947)	0.000992	0.011904	0.11904
Submédio	y=0.001085x+(24.970168)	0.001085	0.01302	0.1302
Médio	y=0.002467x+(23.311219)	0.002467	0.029604	0.29604
Alto	y=0.004817x+(20.580077)	0.004817	0.057804	0.57804

Evapotranspiration (mm/day)

	Linear Trend	mm/month	mm/year	mm/decade
SF	y=-0.001138x+(2.225384)	-0.034709	-0.416508	-4.16508
Baixo	y=-0.003652x+(2.093909)	-0.111386	-1.336632	-13.36632
Submédio	y=-0.002331x+(1.842500)	-0.0710955	-0.853146	-8.53146
Médio	y=-0.000864x+(2.303637)	-0.026352	-0.316224	-3.16224
Alto	y=0.000184x+(2.476955)	0.005612	0.067344	0.67344

Water surface area (km^2)

	Linear Trend	km^2/month	km^2/year	km^2/decade
SF	y=-5.022093x+(8165.712354)	-5.022093	-60.265116	-602.65116
Baixo	y=-0.284265x+(337.947327)	-0.284265	-3.41118	-34.1118
Submédio (all)	y=-0.128774x+(1392.767940)	-0.128774	-1.545288	-15.45288
Submédio (1988-current)	y=-0.800279x+(1595.493879)	-0.800279	-9.603348	-96.03348

Médio	y=-3.354032x+(4835.315633)	-3.354032	-40.248384	-402.48384
Alto	y=-1.173177x+(1583.873084)	-1.173177	-14.078124	-140.78124

Equivalent Water Thickness or TWS (cm)

	Linear Trend	cm/month	cm/year	cm/decade
SF	y=-0.152482x+(4.350678)	-0.152482	-1.829784	-18.29784
Baixo	y=-0.057968x+(-0.422778)	-0.057968	-0.695616	-6.95616
Submédio	y=-0.060459x+(-0.066175)	-0.060459	-0.725508	-7.25508
Médio	y=-0.175073x+(5.928528)	-0.175073	-2.100876	-21.00876
Alto	y=-0.185900x+(4.072322)	-0.1859	-2.2308	-22.308

Correlation: y = GRACE (cm)   x = MapBiomas (km^2)

	Linear Trend	r	r^2
SF	y=0.000124x+(-94.146360)	0.82475716	0.68022438
Baixo	y=0.001162x+(-35.754756)	0.70622551	0.49875447
Submédio	y=0.000364x+(-56.622662)	0.79756942	0.63611698
Médio	y=0.000208x+(-92.202030)	0.78166477	0.61099981
Alto	y=0.000898x+(-125.520839)	0.7620271	0.58068531

Depth of water used for center-pivot irrigation (mm/day)

	Linear Trend	mm/month	mm/year	mm/decade
SF	y=-0.000302x+(0.559737)	-0.009211	-0.110532	-1.10532
Baixo	y=-0.001269x+(0.637073)	-0.0387045	-0.464454	-4.64454
Submédio	y=-0.000858x+(0.492989)	-0.026169	-0.314028	-3.14028
Médio	y=-0.000114x+(0.629455)	-0.003477	-0.041724	-0.41724
Alto	y=-0.000279x+(0.626639)	-0.0085095	-0.102114	-1.02114

Volume of irrigated water per year (km^3)

		km^3/year	r	r^2
SF	y=0.028291x+(-56.334458)	0.028291	0.81560628	0.6652136
Baixo	y=0.000010x+(-0.018711)	0.00001	0.29435019	0.08664204
Submédio	y=-0.000118x+(0.240641)	-0.000118	-0.5906097	0.34881982
Médio	y=0.032991x+(-65.751377)	0.032991	0.88101489	0.77618724
Alto	y=0.003056x+(-6.080025)	0.003056	0.85936213	0.73850328

Appendix 2

The full code used for the data analysis presented in this thesis can be accessed at:
https://github.com/mariafleury/sao-francisco-water