UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS GEODÉSICAS E TECNOLOGIAS DA GEOINFORMAÇÃO

LUCAS DE SIQUEIRA SANTOS

MACHINE LEARNING INTEGRATING CLIMATE DATA WITH GRACE MISSION DATA FOR THE RECONSTRUCTION OF TERRESTRIAL WATER STORAGE ANOMALIES IN BRAZIL

Recife
2025

LUCAS DE SIQUEIRA SANTOS

MACHINE LEARNING INTEGRATING CLIMATE DATA WITH GRACE MISSION DATA FOR THE RECONSTRUCTION OF TERRESTRIAL WATER STORAGE ANOMALIES IN BRAZIL

Dissertation presented to the Graduate Program in Geodetic Sciences and Geoinformation Technologies, of the Federal University of Pernambuco, as part of the requirements for obtaining a Master's degree in Geodetic Sciences and Geoinformation Technologies.

Concentration area: Geodetic Sciences and Geoinformation Technologies

Advisor: Prof. Dr. Rodrigo Mikosz Gonçalves

Co-Advisor: Prof. Dr. Vagner Gonçalves Ferreira

Recife

202

Catalogação de Publicação na Fonte. UFPE - Biblioteca Central

Santos, Lucas de Siqueira.
Machine Learning Integrating Climate Data with GRACE Mission Data for the Reconstruction of Terrestrial Water Storage Anomalies in Brazil / Lucas de Siqueira Santos. - Recife, 2025.
48f.: il.

Dissertação (Mestrado) - Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, Programa de Pós-Graduação em Ciências Geodésicas e Tecnologias da Geoinformação, 2025.
Orientação: Rodrigo Mikosz Gonçalves.
Coorientação: Vagner Gonçalves Ferreira.
Inclui referências.

1. Terrestrial water storage; 2. Machine learning; 3. GRACE; 4. Climate change. I. Gonçalves, Rodrigo Mikosz. II. Ferreira, Vagner Gonçalves. III. Título.

UFPE-Biblioteca Central

LUCAS DE SIQUEIRA SANTOS

MACHINE LEARNING INTEGRATING CLIMATE DATA WITH GRACE MISSION DATA FOR THE RECONSTRUCTION OF TERRESTRIAL WATER STORAGE ANOMALIES IN BRAZIL

Dissertation presented to the Graduate Program in Geodetic Sciences and Geoinformation Technologies, of the Federal University of Pernambuco, as part of the requirements for obtaining a Master's degree in Geodetic Sciences and Geoinformation Technologies.

Approved in:

EXAMINATION COMMITTEE

_____

Prof. Dr. Rodrigo Mikosz Gonçalves
Advisor

_____

Prof. Dr. Henry Diverth Montecino Castro
Internal Examiner

_____

Prof. Dr. Pedro Rodrigues Mutti
External Examiner

SUMMARY

ABSTRACT

Terrestrial water storage (TWS) is a critical component of the hydrological cycle, directly influencing water security, energy production, and climate resilience in Brazil. Although the country has abundant freshwater resources, their uneven spatial distribution combined with the growing impacts of climate change exposes both the population and the economy to hydrological risks. The Gravity Recovery and Climate Experiment (GRACE) missions have provided valuable insights into TWS variability since 2002; however, their limited temporal coverage constrains long-term analyses. To overcome this limitation, this research reconstructed terrestrial water storage anomalies (TWSA) for Brazil's 12 major river basins from 1985 to 2002, integrating GRACE data with climatic variables (precipitation, soil moisture, temperature, and teleconnection indices) and anthropogenic indicators derived from land use and land cover (LULC) data. Two machine learning models—Random Forest (RF) and Long Short-Term Memory (LSTM)—were implemented and compared to assess performance, interpretability, and suitability for GRACE-TWSA reconstructions. Results indicate natural seasonality throughout the year, with vegetation and climate indices emerging as highly influential predictors of TWSA, while anthropogenic factors affect anomalies differently across basins, particularly in areas dominated by agriculture and livestock activities (such as cotton in the Amazon Basin and perennial crops in the São Francisco Basin). Both RF and LSTM achieved satisfactory performance, though LSTM was able to reconstruct the time series for only a few basins, while RF provided greater interpretability of variable contributions. The Mann-Kendall test applied to the RF-reconstructed TWSA series revealed significant long-term decreasing trends in the Uruguay, Parnaíba, São Francisco, and East Atlantic basins, underscoring Brazil's vulnerability to water stress under future climate scenarios. By extending GRACE-derived observations, this study advances understanding of how climate and LULC influence TWSA variability and provides evidence to support public policies for sustainable water resource management in Brazil.

Keywords: terrestrial water storage, machine learning, GRACE, climate change

1. Introduction

Terrestrial water storage (TWS), defined as the total water stored above and below the planet's surface (Girotto; Rodell, 2019), plays a fundamental role in the hydrological balance and socioeconomic security of a country, particularly in nations with a global prominence in renewable water resources, such as Brazil (FAO, 2024). Despite this abundance, the resource's distribution is notably uneven, and Brazil's dependence on its hydrographic network makes it particularly vulnerable to extreme climatic events, including droughts that have led to energy crises (Cuartas et al., 2022). Climate change (CC) projections indicate that, in the coming decades, the country may face significant reductions in water availability, a scenario that aligns with the global trend of decreasing TWS (Pokhrel et al., 2021; ANA, 2024). In this context, understanding the spatiotemporal dynamics of TWS and its fluctuations under different climatic conditions is essential to support efficient water resource management and mitigate the resulting socioeconomic impacts.

The Gravity Recovery and Climate Experiment (GRACE) missions have been pivotal in monitoring TWS in recent years. Their data has been used to study CC in various regions around the world (Tapley et al., 2019) including Brazil (Getirana, 2016). However, having started in 2002, the relatively short duration of these missions represents a limitation for more in-depth analyses, especially when it comes to climate variables, which ideally require a time series of at least 30 years. Faced with this challenge, many researchers have dedicated efforts to extending the GRACE database, reconstructing historical TWS anomalies (TWSA) values through data assimilation techniques, statistical reconstructions, and machine learning (Humphrey; Rodell; Eicker, 2023).

The machine learning (ML) approach offers advantages over conventional models, such as automation and efficiency in detecting climate patterns, as well as a high capacity to handle limited data samples (Karpatne et al., 2019). Among the models most commonly found in the scientific literature for TWSA reconstruction, two consistently stand out for their superior performance: Random Forest (RF) (Jing et al., 2020a; Jing et al., 2020b; Jing et al., 2020c; Tang et al., 2021; Xiong et al., 2022), and various types of Artificial Neural Networks (ANNs) (Long et al., 2014; Ferreira et al., 2019; Sun et al., 2019; Li et al., 2020; Li et al., 2021; Wang et al., 2023a).

Since TWS is heavily linked to land and atmosphere conditions, researchers commonly use climatic variables such as precipitation, soil moisture and temperature as input variables for the reconstruction or prediction of TWS fluctuations. However, there have been

only a few studies that considered the influence of anthropogenic activity in these variations. Furthermore, despite the extensive literature on the reconstruction of anomalies in other regions of the planet, no studies were found at the time of writing this project that had the primary objective of reconstructing and subsequently conducting long-term analyses of TWSA across the entirety of Brazil. Some studies have performed reconstructions for the Amazon Basin (Nie et al., 2016; Tian et al., 2021), and found significant seasonal variations, strongly correlated to the El Niño Southern Oscillation (ENOS). Others have conducted global reconstructions that included Brazil either entirely or partially (Li et al., 2020; Li et al., 2021; Yin et al., 2023), but did not include detailed discussions considering the country's geographic characteristics and historical data. Additionally, some models incorporated variables that are poorly representative of Brazil's reality, such as ice and snow water equivalents.

This research aimed to reconstruct TWSA values from the GRACE mission for each of the 12 main river basins of Brazil, utilizing both RF and Long-Short Term Memory (LSTM), a type of Recurrent Neural Network, for comparative analysis. The reconstruction extends back to 1985, leveraging predictive variables that include climate datasets (precipitation, temperature, soil moisture, and climate teleconnection indices) as well as anthropogenic factors such as land use and land cover. The expected outcomes of this reconstruction are to provide a more comprehensive database that facilitates a deeper understanding of TWS variability in Brazil within the context of CC. This work has the potential to foster more detailed scientific investigations and assist public agencies responsible for water resource management in making more informed decisions on how to prioritize their actions.

1.1 Objectives

1.1.1 General Objective

Reconstruction of historical (1984 – 2002) terrestrial water storage anomaly values across the territorial extent of Brazil, using the 12 main river basins as the spatial units of analysis, with machine learning and data from the GRACE mission combined with climatic and anthropogenic data.

1.1.2 Specific Objectives

- To compare the performance of a Neural Network and Random Forest models in reconstructing terrestrial water storage anomalies (TWSA)
- To describe the role of climatic variables in TWSA modeling

- To analyze the temporal variation of TWSA in Brazil for each basin for the reconstructed period

1.2 Research Hypothesis

1.2.1 Model performance

LSTM and RF models tend to exhibit similar performance, with LSTM being slightly superior in terms of accuracy. However, RF offers greater interpretability, which can be considered an advantage in understanding the interaction mechanisms between variables.

1.2.2 Role of climate variables

Precipitation and climate indices are the most influential climatic variables in determining TWSA anomalies in Brazil, reflecting the importance of the surface hydrological cycle in terrestrial water storage, particularly in a country predominantly located in a tropical zone.

1.2.3 Anthropogenic influence

The spatial variation of TWSA in Brazil is strongly influenced by land use and land cover, with basins predominantly characterized by agricultural activity showing greater vulnerability to positive TWSA anomalies due to infiltration caused by irrigation. In contrast, areas dominated by extensive livestock farming are more susceptible to negative anomalies due to vegetation removal and water extraction.

1.3 Justification

Brazil's abundance of water resources positions the country as a significant international exporter of virtual water (Suweis et al., 2013), while also enabling the generation of more than 61.9% of the nation's total electricity (EPE, 2023). Although the hydrological cycle exhibits natural spatial and temporal variations, global projections, such as those from the Intergovernmental Panel on Climate Change (IPCC), indicate that extreme events may become increasingly frequent in the coming years, with natural climate variability amplifying or attenuating their adverse consequences (Parmesan et al., 2022).

Some trend projections suggest that water availability in Brazil could decrease by up to 40% by 2040 due to CC (ANA, 2024). This reality has become increasingly evident in recent years, as shown by the severe drought event recorded in the Amazon in 2023, which was

attributed more to the impacts of CC than solely to the ENSO phenomenon, traditionally associated with drought episodes in the region (Clarke et al., 2024). Additionally, the impact of CC has been underscored by the recent and unprecedented identification of an arid climate area in the center of the Northeastern region of Brazil (INPE, 2024). This highlights the country's vulnerability to the risks posed by CC. If the projections from the IPCC and ANA are confirmed, Brazil may face a series of challenges related to the management and availability of its water resources. In this context, it is crucial to deepen the understanding of the spatial and temporal variations of terrestrial water storage across the national territory. This will equip regulatory agencies with essential information for public policy debates and provide the academic community with valuable data for more detailed investigations into climate variables and their influence on water availability.

Parallel to the climate threat to Brazil's water resources, the world is witnessing a rapid evolution in AI, which is revolutionizing various sectors, including the Geosciences (Karpatne et al., 2019). ML techniques are emerging as powerful tools in the fight against climate change (Rolnick et al., 2023), presenting immense potential to mitigate its impacts. The research presented here is precisely situated within this context, aiming to employ ML advancements to tackle climate challenges and contribute to the sustainable management of Brazil's water resources.

2 Literature review

2.1. Terrestrial water storage and GRACE missions

TWS is composed of six components: surface water, groundwater, soil moisture, ice, snow, and water stored in biomass, as shown in Equation 1 (Girotto; Rodell, 2019; Humphrey; Rodell; Eicker, 2023).

$$TWS = SW + GW + SM + SWE + LI + BW \quad (1)$$

Where SW represents surface water, GW represents groundwater, SM represents soil moisture, SWE represents snow water equivalent, LI represents land ice, and BW represents biomass water. The contribution of each component to TWS variations depends on the local geography (Zhang et al., 2022).

Considering the mentioned components, it is natural that TWS variation significantly depends on local climatic behavior. Therefore, climatic variables such as precipitation,

temperature, soil moisture, and climate indices are fundamental for understanding TWS behavior, with precipitation being the variable that contributes most to TWS changes in low-latitude regions such as Brazil. (Zhang et al., 2019).

It is possible to estimate TWS values using surface models such as the Global Land Data Assimilation System (GLDAS). However, thanks to the GRACE and GRACE-FO missions, TWS estimates can be made much more accurately and reliably. The concept behind these missions is based on observing variations in the distance between two satellites sharing the same orbit, leveraging the fact that Earth's gravitational field is heterogeneous due to the uneven distribution of mass. When a satellite approaches an area with greater gravitational force, its velocity increases, moving away from the trailing satellite; upon passing that area, its velocity decreases, reducing the distance between them (Figure 1). These distances are determined using a combination of highly precise instruments. Each satellite in the original GRACE mission was equipped with GNSS sensors, which enabled accurate estimation of their orbital positions and, consequently, the measurement of the distance between the two satellites. In the follow-up mission, GRACE-FO, a laser-ranging interferometer (LRI) has been introduced as a technology demonstrator to enhance the precision of these inter-satellite distance measurements. In addition, the mission employs the HAIRS (High Accuracy Inter Satellite Ranging System), which operates in the microwave spectrum, and corner-cube retroreflectors on both satellites, allowing ground-based laser tracking to determine the distance from each satellite to Earth.

The monthly analysis of these distance differences provides the temporal variations (or anomalies) of Earth's gravitational field (Girotto; Rodell, 2019; Humphrey; Rodell; Eicker, 2023). Considering that large geological formations and urban clusters have relatively static mass, mass variations on continents are primarily attributed to atmospheric and oceanic circulations and the redistribution of water on the surface. By using climate models, it is possible to identify and exclude the signals from these circulations, enabling the estimation of TWS variations in isolation. However, satellites cannot directly calculate storage values; instead, they calculate anomalies based on deviations from an average. Therefore, in addition to TWS, there are two similar concepts that frequently appear in the literature and must be differentiated: (1) TWS anomalies, or TWSA; and (2) TWS changes, or TWSC.

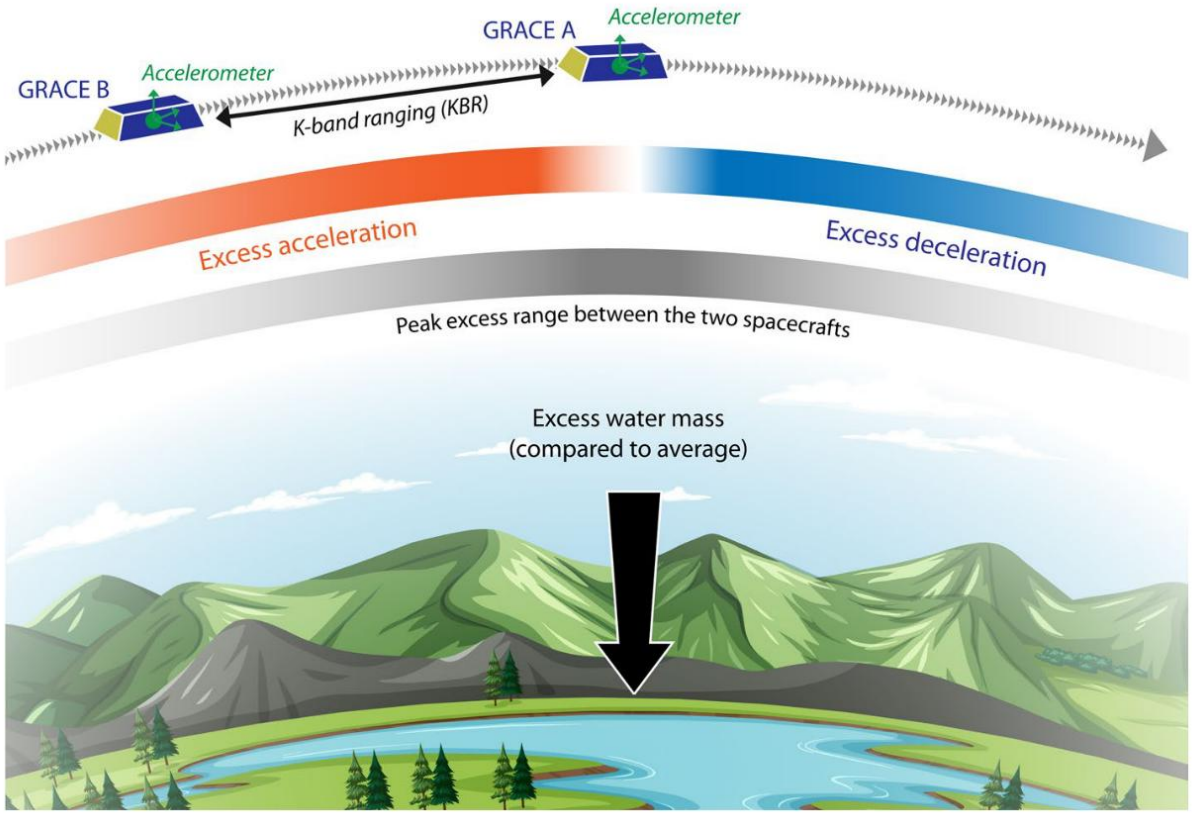


Figure 1. Principle of the GRACE missions. As a satellite passes over a region with a mass concentration, it increases its velocity toward the concentration due to the stronger gravitational pull. Upon moving past this region, the satellite remains under the influence of this anomaly, thereby reducing its velocity. (Humphrey; Rodell; Eicker, 2023).

The anomalies (TWSA) are calculated based on how much the gravitational attraction at a given point (excluding the influences of atmospheric and oceanic circulations) deviates from the mean over a reference period, as expressed in Equation 2 (Humphrey; Rodell; Eicker, 2023).

$$TWSA = TWS - \overline{TWS} \qquad (2)$$

Where $\overline{TWS}$ is the average TWS for a reference period. TWSC, in turn, refers to the variation of TWSA relative to the previous period, as expressed in Equation 3 (Lv et al., 2019):

$$TWSC_t = \frac{TWSA_t - TWSA_{t-1}}{\Delta t} \qquad (3)$$

Where t is the period (month), $t-1$ is the previous month, and Δt is the time difference between the two periods.

Currently, the mission's data processing is carried out by four centers: the Jet Propulsion Laboratory (JPL), the Center for Space Research (CSR), and the Goddard Space

Flight Center (GSFC) of NASA, as well as the GeoForschungsZentrum (GFZ) in Germany. The data can be processed in two ways to provide products to users: (1) spherical harmonics; and (2) mass concentration blocks (*mascons*).

Spherical harmonics are functions used to model variations in Earth's gravitational field, representing these variations continuously across the planet's surface based on the estimation of coefficients of degree *l* and order *m*. (Humphrey; Rodell; Eicker, 2023). Currently, spherical harmonic-based solutions are produced by the JPL, CSR, and GFZ centers. In contrast, mascon solutions do not represent TWSA anomalies as a continuous field but rather in blocks of specific dimensions, assigning a single TWSA value to each block. The JPL and CSR are the centers responsible for these solutions.

The GRACE mission products available to users, whether based on spherical harmonics or *mascons*, are presented in units of meters, centimeters, or millimeters of water equivalent. These units reflect the amount of water required on the ellipsoid's surface to account for the anomalies observed in the gravitational field (Humphrey; Rodell; Eicker, 2023).

Some studies have already explored the data provided by the GRACE mission to investigate the spatiotemporal distribution of TWS in regions of Brazil, particularly for monitoring hydrological droughts. (Oliveira et al., 2014; Getirana, 2016; Rebello et al., 2017; Rossi et al., 2023). However, no studies found have performed a reconstruction aimed at the long-term analysis of TWSA variations in Brazil. On the other hand, in other regions of the world, several studies have focused on TWSA reconstruction using ML techniques (as discussed in Section 2.3).

2.2. Machine Learning

Between the conception of the idea of machines performing intellectual work at a human level of performance (Turing, 1950) and the application of statistical learning models as tools to combat climate change (Rolnick et al., 2023), there is a massive number of scientific contributions aimed at theoretically defining these tools and guiding their applications. Defining ML, therefore, is an inherently subjective task. Simplified, ML, a subfield of AI, is dedicated to developing algorithms that enable computers to autonomously improve their performance through experience gained from data (Mitchel, 1997). Unlike operating under a rigid set of instructions, these systems learn to discern patterns and relationships in the data presented, allowing them to make predictions or decisions based on new data. These tasks are

divided into two main types: (1) classification, when the expected outcomes are categorical, and (2) regression, for outputs with numerical value (Molner, 2019)[1].

If the goal is to produce numerical estimates of TWSA for periods prior to the launch of the GRACE mission, associating observed anomaly values with other variables, including historical records of these variables in the model allows for the prediction of anomalies for the corresponding periods. Therefore, by inputting historical data as labeled entries, the model performs a regression task, estimating TWSA for past periods based on the relationships it has learned.

A recurring challenge in using ML models is the trade-off between performance and interpretability. Interpretability, in the context of ML, can be defined as the degree to which a human can understand the reasons behind a machine's decision in a given context (Miller, 2019). Highly complex and adaptive models, particularly those that capture the non-linearity of relationships between variables, often achieve higher levels of accuracy. However, due to this additional complexity, understanding the underlying processes can become obscured, making it difficult to discern how inputs are being transformed into outputs. (Molner, 2019). This trade-off becomes even more relevant in the context of Geosciences, where understanding the interaction mechanisms between different variables is an essential tool for scientific development. (Karpatne et al., 2019).

Therefore, the choice of ML models for this research considered both performance and interpretability. The next section provides a review of how different ML techniques have been used for the reconstruction of TWSA in other regions of the world.

2.3. Reconstruction of TWSA using Machine Learning

Research on the reconstruction of TWSA using ML techniques has grown remarkably in recent years. Given the diversity of algorithms and methodological approaches employed in the ML field, this review aims to address two fundamental questions: (1) What are the predominant ML models used in TWSA reconstructions derived from the GRACE missions? (2) What variables are most frequently adopted for this purpose? Through a search on the academic platforms Google Scholar and Scopus, using the search strategy specified in Figure

[1] The definition of ML used here and the division of tasks into classification and regression are associated with supervised learning, which is characterized by the use of pre-labeled data, learning to map input data to known outputs, enabling predictions or classifications for new datasets. In contrast, unsupervised learning deals with unlabeled data, aiming to uncover intrinsic patterns and relationships without the guidance of specific outputs. This enables the grouping of similar data into clusters or the identification of association rules, exploring the data structure without the need for prior labeling. (Singh, 2019; Sarker, 2021).

2, 19 studies were selected. An additional 8 studies were included through subsequent analysis of the references from the initial 19, totaling 27 peer-reviewed scientific articles published in international journals.

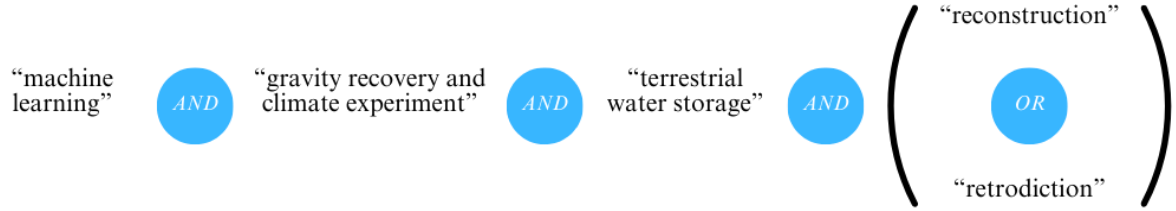


Figure 2. Search strategy used in the Google Scholar and Scopus databases to find the selected articles

The criteria for article inclusion were: (1) the use of ML as the primary tool for reconstructing TWSA data; and (2) a focus on extending the dataset to periods prior to the start of the GRACE mission, thereby excluding studies that focused exclusively on filling the gap between the GRACE and GRACE-FO missions. The justification for the second criterion lies in the need to address the reconstruction of anomalies for more distant periods, tackling the task of extending the dataset over a long timeframe, as opposed to merely addressing the 7-month data gap between the end of the GRACE mission and the start of the GRACE-FO mission.

The dataset extension was the primary focus of most analyzed articles, aiming to enable future climate research through an expanded dataset. However, some studies focused on specific topics, such as identifying and characterizing droughts and floods, or analyzing the impact of certain environmental variables on TWS variability. In these contexts, anomaly reconstruction emerged as a complementary objective, supporting the investigation of these more specific climatic and hydrological dynamics. Table 1 details the studies found, their year of publication, and the model(s) used.

Table 1. Articles included in the review

id	Authorship	Models used	Year
1	Long et al.	Neural Network (Multi-Layer Perceptron – MLP)	2014
2	Yang et al.	Neural Network (MLP)	2014
3	Nie et al.	simple linear regression	2016
4	Zhang et al.	Neural Network (MLP)	2016
5	Yang et al.	Neural Network (MLP), RF, SVM	2018

6	Ferreira et al.	Neural Network (NARX)	2019
7	Sun et al.	Neural Network (CNN)	2019
8	Chen et al.	Neural Network (MLP)	2019
9	Jing et al.	RF, XGB	2020
10	Jing et al.	RF, XGB	2020
11	Jing et al.	RF, multiple linear regression	2020
12	Li et al.	multiple linear regression, Neural Network (MLP and ARX)	2020
13	Wang et.	Neural Network (LSTM)	2021
14	Li et al.	multiple linear regression, Neural Network (MLP and ARX)	2021
15	Meng et al.	Neural Network (MLP), SVM	2021
16	Tang et al.	RF	2021
17	Tian et al.	multiple linear regression	2021
18	Yang et al.	RF, XGB	2022
19	Dannouf et al.	Boosted Regression Tree (BRT) and Neural Network (NARX)	2022
20	Xiong et al.	RF	2022
21	Kalu et al.	Neural Network (CNN) and SVM	2023
22	Yin et al.	RF and Neural Network	2023
23	Li et al.	multiple linear regression, RF	2023
24	Kumar et al.	Neural Network (MLP)	2023
25	Wang et al.	RecNet	2023
26	Zheng et al.	Neural Network (NARX)	2023
27	Zhu et al.	LightGBM	2023

The diversity of models employed in the studies reflects the adaptability of these techniques to different datasets and research objectives. The most used model was Neural Networks (NN) (18), encompassing various architectures such as MLP (9), NARX (3), CNN (3), ARX (2) and LSTM (1); followed by Random Forest (RF) (9), as shown in Figure 3. Other models included Linear Regression (6), Support Vector Machine (3), eXtreme Gradient Boost (3), Boosted Regression Tree (1), and LightGBM (1). The two most frequently used models (NN and RF) are discussed below. Subsequently, Section 2.6 presents the predictor variables used in the reviewed articles.

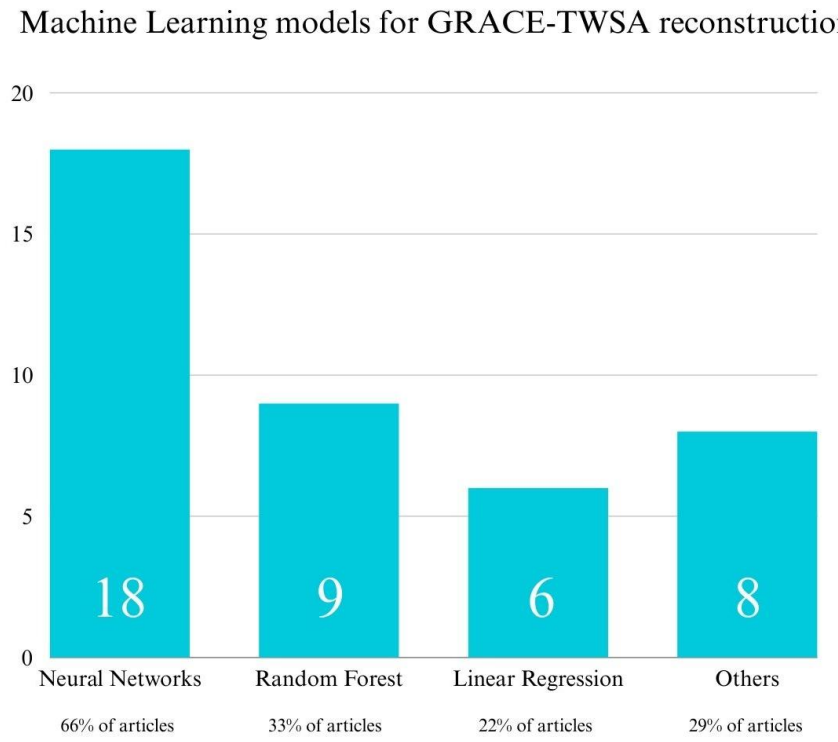


Figure 3. Frequency of ML models used in the reconstruction of TWSA. "Others" refers to: Support Vector Machine (3), eXtreme Gradient Boost (3), Boosted Regression Tree (1), and LightGBM (1).

2.4 Artificial Neural Network (ANN)

The functioning of an Artificial Neural Network (ANN) is inspired by the biological neural network model proposed by Mcculloch and Pitts (1943), which consists of a network of neurons that, by transmitting electrical signals among themselves, are able to perform learning tasks. ANNs operate through connections that transmit signals regulated by self-adjustable weights and processed based on activation functions, organized into layers of processing units (neurons), including an input layer, hidden layers, and an output layer for the result (Janiesch; Zschech; Heinrich, 2021). The number of layers and activation functions constitute the model's hyperparameters, defined externally by the user. The output value depends on how the input data interacts with the weights and the activation function used (Wu; Feng, 2018). Each connection between neurons transmits a signal, whose strength can be attenuated or amplified by a weight self-regulated by the model (Janiesch; Zschech; Heinrich, 2021). Figure 4 provides a visual representation of a simple ANN.

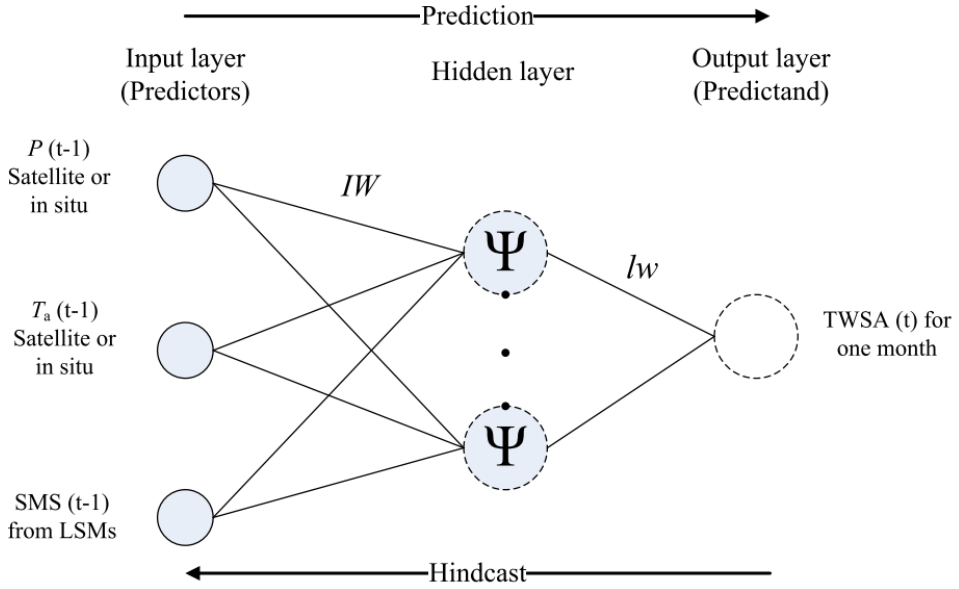


Figure 4. Neural network (multilayer perceptron, or MLP) used by LONG et al. (2014). Each of the three neurons in the input layer receives a dataset (a variable). *IW* and *lw* are self-regulated weights, and ψ is the activation function.

ANNs were the first ML model used for TWSA reconstruction (Long et al., 2014), and continue to be widely employed in more recent studies due to their ability to capture non-linear relationships among predictor variables (Wang et al., 2023), proving suitable for addressing the multifaceted nature of TWS. In all reviewed studies that utilized ANNs (as shown in Table 1), the model outperformed linear regression models and even TWSA estimates based on the sum of TWS components from Equation 1 derived from products like GLDAS. These reconstructions were also effective in identifying historical severe hydrological drought events (Long et al., 2014; Zhang et al., 2016; Wang et al., 2023; Zheng et al., 2023), demonstrating significant value for water resource management.

Although the MLP has demonstrated success in reconstructing TWSA values, a specific type of ANN is particularly well-suited for handling temporal and sequential data: recurrent neural networks (RNNs). This architecture has been utilized in five studies included in this review (6, 12, 14, 19, and 26), all of which reported satisfactory results.

An RNN is a type of neural network designed to handle sequential data by using feedback connections that allow information to persist over time (Rumelhart; Hinton; Williams, 1986). Unlike traditional feedforward networks (such as the MLP), RNNs have loops that enable the model to process one element of a sequence at a time while retaining information about previous inputs, as described in Figure 5. This is achieved by maintaining a hidden state that is updated at each time step, incorporating both the current input and the hidden state from

the previous step. While effective for many temporal tasks, RNNs often struggle with capturing long-term dependencies due to issues like vanishing gradients during training.

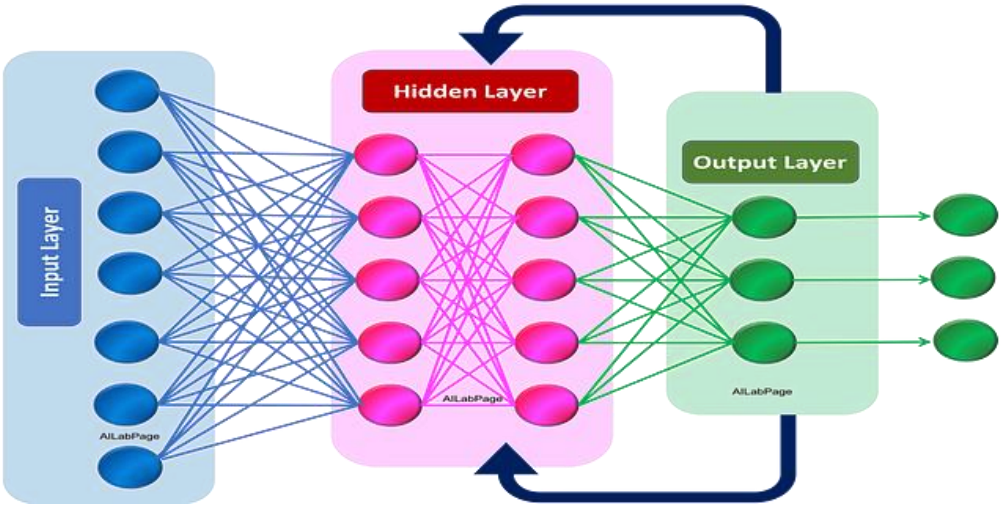


Figure 5. A recurrent neural network (RNN) architecture. The output of each neuron is fed back into itself at the subsequent time step, creating a recurrent loop that enables the model to learn from sequential data (adapted from SHARMA, 2019).

However, an advanced version of RNN, known as Long-Short Term Memory (LSTM), offers improved capabilities. It has been used only once for extending the GRACE dataset (Wang et al., 2021). It has also been used recently to fill gaps between GRACE missions and addressing missing months within the time series (Ferreira et al., 2024). The difference between an RNN and the LSTM model is that the latter incorporates a mechanism designed to overcome the limitations of traditional RNNs in handling long-term dependencies. The LSTM model achieves this through a unique architecture that includes memory cells, along with gates (input, output, and forget gates), which regulate the flow of information within the network. These gates enable LSTM to selectively remember or forget information over extended sequences, making it particularly effective in capturing temporal patterns in sequential data without the problem of vanishing or exploding gradients often encountered in standard RNNs (Hochreiter; Schmidhuber, 1997). Although the LSTM has only been used once in this context, its suitability for handling temporal data led to its selection as the ANN architecture for this research.

Regardless of the type of ANN employed, its inherent complexity can make it challenging to quantify the relationships it models (Sun et al., 2019; Yin et al., 2023). This often results in limited interpretability, which can constrain its application in the study of hydrological variables and TWS.

2.5 Random Forest (RF)

The Random Forest model consists of combining N decision trees that use different subsets of predictor variables. Each tree predicts a final value, and the average of the results from all trees is defined as the output of the RF model. Using a collection of decision trees instead of just one is a way to improve model performance, and this type of aggregated learning is referred to as ensemble learning (Sarker, 2021). The model is graphically represented in figure 6.

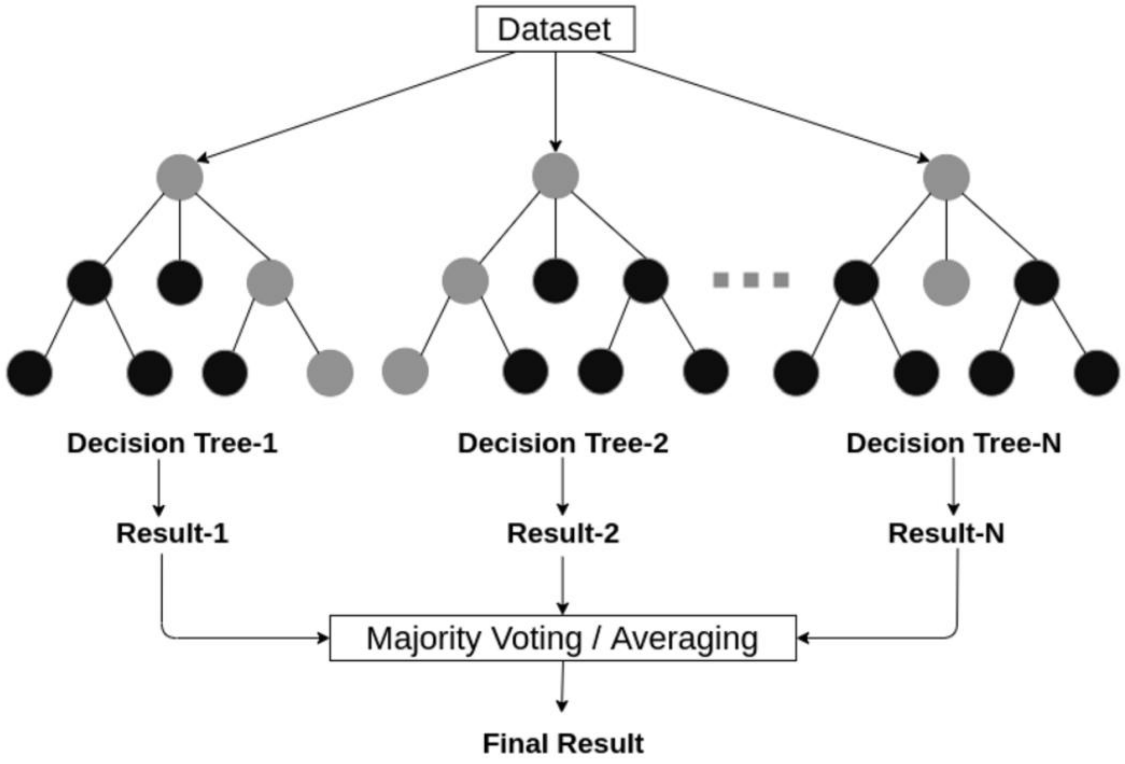


Figure 6. Random Forest model, consisting of NNN decision trees. The final result is the average of the outputs from each tree (Sarker, 2021).

Considering the interpretability limitations inherent to ANNs, RF offers a clear advantage. As mentioned earlier, understanding the interaction mechanisms among study variables is essential for scientific development in the geosciences (Karpatne et al., 2019), and RF addresses this need through its ability to quantify the contribution of each variable to the model's output (Breiman, 2001). Yang et al. (2022) applied RF and XGB algorithms to reconstruct TWSA in the Huang-Huai-Hai River basin in China, including both climatic and anthropogenic variables in their analysis. They concluded that climatic variables, particularly precipitation from the previous month, contributed up to 70.8% to the accuracy of TWSA reconstruction in the studied region. Similarly, Jing et al. (2020a) identified precipitation from

the past two months as the most significant variable in anomaly reconstruction for the Pearl River Basin in China.

Since the studies employing these models were conducted in basins with varying hydroclimatic regimes, sizes, and training data periods, comparing their results directly would be inappropriate. Therefore, the selection of models in this research was not based on superior performance, but rather on their widespread use and recognition in the literature as the most adopted for this purpose (despite LSTM architecture specifically being used only once, it was chosen for its conceptual suitability for temporal data).

2.6 Predictor variables

In any ML model, the selection of input data is crucial for generating accurate outputs. In TWSA reconstruction, predictor variables vary widely, reflecting the availability of data in the region of interest. However, one element stands out due to its prevalence: precipitation was included as a predictor variable in 24 out of the 27 analyzed articles (~92% of the studies), as shown in Figure 7. This choice is grounded in the intrinsic relationship between TWS and the hydrological cycle, where precipitation plays a central role by introducing water into the terrestrial system. (Girotto; Rodell, 2019). Moreover, the correlation between precipitation levels and TWSA is more significant when considering positive time lags, as precipitation precedes changes in terrestrial water storage (Jing et al., 2020a). The sources of precipitation data vary, including local meteorological stations, land surface models (LSMs) such as those provided by GLDAS, Earth observation (EO) products like NASA's Tropical Rainfall Measuring Mission (TRMM), and databases compiled primarily from in situ observations, such as the Climate Research Unit (CRU) and the Climate Prediction Center (CPC) precipitation datasets. In addition to precipitation, atmospheric temperature (T) and soil moisture (SM) are also among the most used variables, appearing as predictors in 69% and 62% of the studies, respectively. Again, the data sources vary among *in situ* measurements, LSMs, and EO products.

The preference for LSMs, particularly GLDAS, can be explained by their comprehensive temporal and spatial coverage. Besides providing estimates for several key hydrological and climatic variables, GLDAS extends its estimates back to 1948, significantly enhancing the possibilities for studying long-term hydrometeorological processes. EO techniques, on the other hand, while not covering extensive historical periods, can deliver more

precise measurements of precipitation and other variables, as these values are derived from direct observations rather than estimates, as in the case of GLDAS.

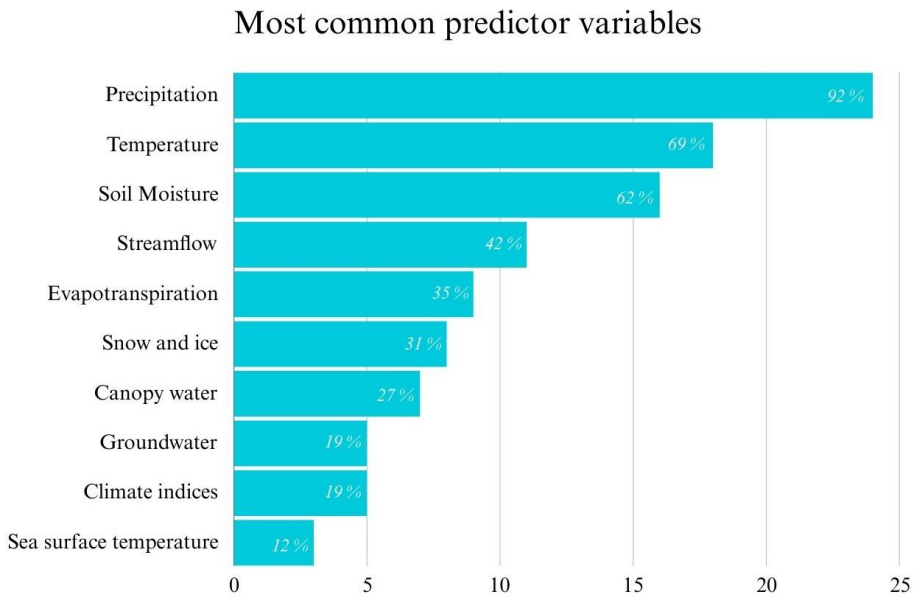


Figure 7. The 10 most frequently used predictor variables in TWSA reconstructions using ML. The percentages within each bar represent the proportion of articles that utilized each variable out of the total 27.

3 Methods

3.1. Study area

Brazil is the 5th largest country in the world and the largest in the Americas by territorial area, covering 8,510,417 km² (IBGE, 2023). The country has 12 main river basins (Figure 8), which serve as the administrative units of the national water regulatory agency (Agência Nacional de Águas, ANA). Approximately 81.4% of its territory has a tropical climate, with the highest concentration of precipitation in the Amazon region and the lowest in the Northeast (Alvares et al., 2013). All South America (SA) is strongly influenced by ENSO, with drier conditions commonly occurring in the North and Northeast regions of Brazil and wetter conditions in the South and Southeast during the positive phase (El Niño). Conversely, during the negative phase (La Niña), the opposite occurs: increased rainfall in the North and Northeast and reduced rainfall in the South and Southeast.

Regarding topography, Brazil's highest altitudes are found in the South and Southeast regions, decreasing toward the country's center.

The 12 Major River Basins in Brazil

Basin	Area (km²)
Amazon	3,844,917.76
Tocantins-Araguaia	918,273.16
Paraná	877,513.54
São Francisco	636,137.07
East Atlantic	386,068.13
Paraguay	362,263,92
Parnaíba	331,808.82
East Northeast Atlantic	285,281.21
West Northeast Atlantic	268,906.01
Southeast Atlantic	213,316.01
South Atlantic	186,079.86

Figure 8. The 12 major river basins in Brazil. Data: ANA, 2016

3.2. Data

The reconstruction of the TWSA time series was performed using climatic and anthropogenic variables as input. The climatic variables include precipitation (P), temperature

(T), soil moisture (SM), and climatic teleconnection indices. The first three variables are sourced from the Global Land Data Assimilation System (GLDAS), while the indices were obtained from the National Oceanic and Atmospheric Administration (NOAA) website. The anthropogenic variables were the land use and land cover (LULC) classes mapped by the MapBiomas.

All the data preprocessing described below was done using python libraries such as 'xarray', 'pandas', 'geopandas' and 'numpy', using the cloud computing service Google Colaboratory, and the LULC data was preprocessed using the cloud computing platform specialized in processing EO data, Google Earth Engine (GEE).

3.2.1. GRACE *mascons*

In this study, the target variable to be reconstructed (TWSA) was based on the monthly mascon set (RL06.3) processed by CSR. This dataset covers the period from April 2002 to May 2025 and is available for download with a 0.25° spatial resolution resampling. The reason for using mascon solutions instead of spherical harmonic solutions lies in the advantage that *mascons* do not require the application of smoothing filters or other preprocessing steps. (HUMPHREY; RODELL; EICKER, 2023).

After downloading the dataset, the TWSA values were first compiled into a single multiband raster file, where each band corresponded to a specific time step. For each basin and each month of the time series, descriptive statistics of the TWSA values were then computed, including the mean, median, variance, and standard deviation. These metrics were exported to a CSV file. The rationale for computing these statistics was to assess whether intra-basin variability could be related to the performance of the ML reconstructions and were not used as features (predictive variables). The resulting CSV file contained the following columns: `twsa_mean` (the target variable to be reconstructed), `twsa_median`, `twsa_var`, `twsa_std`, `basin`, `month`, and `year`. Lastly, the data was cropped to cover the period from April 2002 and December 2014, to be time-compatible with the GLDAS. This final, cropped dataset contained 16 missing months (10%) out of a total of 156 monthly observations for each basin, which were later imputed by the models themselves.

. The use of mean values for each variable aims to harmonize the datasets by compensating for the differences in spatial resolution between GRACE/GLDAS (0.25°) and MapBiomas (30 m), as well as addressing the non-overlapping grid cells present in the 0.25° GRACE/GLDAS framework.

3.2.2. Global Land Data Assimilation System (GLDAS)

The Global Land Data Assimilation System (GLDAS) is a land surface model (LSM) collaboratively developed by NASA and NOAA in the United States of America. It is a global land modeling system that integrates Earth Observation (EO) data with *in situ* data, using data assimilation techniques to generate fields of various variables related to land states and fluxes in near real-time. (RODELL et al., 2004).

Soil moisture from the GLDAS version 2 datasets have been compared to the Global Precipitation Climatology Centre (GPCC) observational precipitation dataset for the southeastern part of SA, and the Noah LSM (one of the GLDAS datasets) was regarded as the best due to its high correlation with the Standardized Precipitation Index (SPI) calculated with GPCC data (SPENNEMANN et al., 2015). Regarding precipitation, GLDAS has shown monthly values consistent with *in situ* measurements for the state of Mato Grosso, Brazil (PEDREIRA JUNIOR et al., 2021). Regarding air temperature, GLDAS values have shown consistent performance globally when compared to the global Historical Climatology Network (GHCN), although agreement indices were lower in SA compared to other regions (JI; SENAY; VERDIN, 2015). This difference is likely to be due to the significantly smaller number of weather stations available for validation in SA, which limits the reliability of the comparison and naturally results in lower quality assessments for this region.

The datasets are available at spatial resolutions of 0.25° and 1°, covering the period from 1948 to the present. For this research, the GLDAS Noah Land Surface Model L4 monthly 1.0 x 1.0-degree V2.0 (Beaudoing; Rodell, 2019) dataset was used (available from 1948 to 2014). Among the 36 available variables, only P, T, and SM were utilized, as these are the three most used variables in similar studies (as discussed in Section 2.3). Each variable underwent the same process of conversion into a raster file and mean aggregation by basin and time period, as described for the *mascons*.

The P variable in the GLDAS dataset is expressed in units of kg m^{-2} s^{-1}. To convert the precipitation values into millimeters (mm), the transformation was performed following the GLDAS guidelines outlined in the README Document for NASA GLDAS Version 2 Data Products. Since the monthly average files represent straightforward averages of 3-hourly data, each monthly average is expressed in units per 3 hours. Therefore, the conversion to mm was carried out using the formula presented in Equation 4.

$$P\left\{\frac{kg}{m^2}\right\} = P\left\{\frac{\frac{kg}{m^2}}{s}\right\} * 10800\left\{\frac{s}{3h}\right\} * 8\left\{\frac{3h}{day}\right\} * 30\{days\} \quad (4)$$

Since 1kg/m² of water is equivalent to 1 mm of water depth, the resulting value can already be interpreted as mm of precipitation.

The SM variable is expressed as kg m^{-2}, representing the water content in the first 10 cm of soil depth. To convert it to m^3/m^3, which is a standard measure for soil moisture, the value is divided by the depth of the soil layer in meters (0.1 m) and by the density of water (1000kg/m³), as shown in Equation 5.

$$SM\ \{m^3/m^3\} = \frac{SM\{kg/m^2\}}{0.1\ m * 1000kg/m^3} \quad (5)$$

The T variable is originally expressed in Kelvin and was further converted to Celsius, as described in Equation 6.

$$T\ \{°C\} = T\{K\} - 273.15 \quad (6)$$

3.2.3. MapBiomas

MapBiomas is a collaborative and independent initiative composed of Brazilian NGOs, universities, and technology startups that has been mapping LULC in Brazil since 2015, with data retrospectively extended back to 1985 (Souza et al., 2020). Collection 9 of the dataset includes 29 LULC classes with a spatial resolution of 30 meters. The use of LULC data as a predictor variable in TWSA reconstruction has also been reported by Yin et al. (2023) and serves to incorporate the effects of anthropogenic activity on TWS variation, as changes in this variable have been linked to TWS fluctuations (Chen et al., 2017; Wang et al., 2020). The existing classes in the dataset are Aquaculture, Beach, Dune and Sand Spot, Citrus, Coffee, Cotton, Floodable Forest, Forest Formation, Forest Plantation, Grassland, Herbaceous Sandbank Vegetation, Hypersaline Tidal Flat, Mangrove, Mining, Mosaic of Uses, Other non-Vegetated Areas, Other Perennial Crops, Other Temporary Crops, Palm Oil, Pasture, Rice, River, Lake and Ocean, Rocky Outcrop, Savanna Formation, Soybean, Sugar cane, Urban Area, Wetland, Wooded Sandbank Vegetation, Not Observed.

Using GEE, the proportion of LULC classes for each basin and at each time step was calculated, resulting in a dataset representing the percentage of each LULC type per basin, per

year. Later, the proportion of each class was used as a predictor variable for the TWSA reconstructions, except for the “Not Observed” class. This class accounted for less than 1% across all basins and was excluded to minimize noise in the model training.

3.2.4. Climate teleconnection indices

The climate of SA is strongly influenced by the El Niño-Southern Oscillation (ENSO) phenomenon, which triggers extreme events such as hydrological droughts and floods depending on its phase and the specific region of the continent (Cai et al., 2020). TWS variation has often been associated with ENSO fluctuations (Nie et al. 2016; YANG et al. 2018b; Li et al. 2020; Li et al. 2021; Tian et al. 2021; Wang et al. 2023a; Yin et al. 2023), indicating that incorporating indices that measure these fluctuations can enhance the accuracy of TWSA reconstructions.

In this study, the Oceanic Niño Index (ONI) was utilized. The ONI is calculated as a three-month moving average of sea surface temperature (SST) anomalies in the Niño 3.4 region (5°N–5°S, 120°W–170°W), with deviations from the 30-year climatological mean (updated every five years) determining the index values. An ONI value of ≥0.5°C signifies El Niño conditions, ≤−0.5°C indicates La Niña, and values in between represent a neutral phase.

Beyond ENSO, other climate teleconnections can also influence Brazilian climate, and although it is not possible to objectively rank the patterns that have the greatest impact on the country’s climate, several of them have well-documented and widely studied effects, such as the Pacific Decadal Oscillation (PDO), Southern Atlantic Ocean Dipole (SAOD) and Tropical Southern Atlantic (TSA) SST anomalies (REBOITA et al., 2021).

The PDO is a long-term climate pattern characterized by periodic SST changes in the North Pacific Ocean, primarily north of 20°N, with warm and cool phases that influence ENSO's effects on Brazil by intensifying droughts in the Northeast region and altering rainfall patterns in the South during its warm phase (Mantua; Hare, 2002; Kayano; Andreoli, 2007).

The SAOD is another significant SST pattern in the South Atlantic Ocean, marked by opposing anomalies between the northeastern Atlantic Niño region and the southwestern coast off Argentina, Uruguay, and Brazil. Lasting approximately eight months with peak impacts during austral winter, the SAOD drives climate variability and affects precipitation along Brazil’s southern coast (Nnamchi; Li; Anyadike, 2011). The Southern Atlantic Dipole Index (SAODI) quantifies this phenomenon by measuring the difference in average SST between two regions of intense warming and cooling associated with the SAOD.

Additionally, SST anomalies in the TSA influence Brazil's climate by interacting with the Intertropical Convergence Zone (ITCZ), a primary driver of precipitation in the northern and northeastern regions of Brazil (Ferreira; Giovanni; Mello, 2005). This interaction modulates regional climate variability and atmospheric circulation patterns (Enfield et al., 1999).

All indices used to account for these teleconnections were sourced from the Climate Prediction Center (CPC), from the NOAA website (NOAA, 2017). The indices were compiled into a dataset spanning monthly data from 1948 to 2014, with values recorded for each corresponding period.

3.3.TWSA reconstruction

3.3.1. Random Forest

After the data preprocessing mentioned in the previous sections, all the resulting preprocessed data was compiled into a single tabular dataset containing a column for the basin identifier, the month, year, the target variable (TWSA) and each of the predictor variables. Then, the implementation of the RF model was done using the python library scikit-learn, via the cloud computing platform Google Collaboratory.

The complete data from 2002 to 2014 was split into three parts: training (70%), validation (15%) and testing (15%). Then, the categorical variable representing the basin identifier was transformed using one-hot encoding, resulting in a binary (True/False) representation for each category. In this format, each basin was assigned a dedicated column; for instance, if the first row corresponded to the Amazon basin, the 'Amazon' column would be marked as True, while all other basin columns would be marked as False. This encoding scheme is commonly used to prepare unordered categorical data for RF algorithms.

The selection of the number of estimators (i.e. decision trees) was determined through an empirical trial-and-error approach. As varying this parameter resulted in only minor changes to the performance metrics, altering values solely at the third decimal place, the chosen number was set to 100 (the standard number for most ML applications), since increasing the number of trees would have led to higher computational costs without yielding significant improvements in accuracy.

The importance of each variable was assessed using the permutation importance method, which consists of randomly shuffling the values of each predictor several times and measuring the decrease in the model's R^2 caused by this perturbation (Breiman, 2001). This

approach is a key tool in the Random Forest framework, as it enhances the interpretability of the model's results.

3.3.2. Long Short-Term Memory

The performance of a neural network method is heavily dependent on the hyperparameters chosen by the user, as well as the nature of the data. The set of parameters used in this research was a mix between the parameters used by Wang et al., 2021 and other parameters selected by a trial-and-error approach: 4 hidden layers (with 50, 50, 60 and 10 neurons respectively and 20% dropout in each), window size of 3 months, 75 epochs and batch size of 2.

The LSTM model, due to its recurrent structure, was built independently for each basin to accurately represent sequential data. Given that the original dataset contained 12 monthly records per time step, the data had to be partitioned by basin. Consequently, a single performance metric applicable to all basins could not be established; instead, individual metrics were calculated for each basin.

The dataset was split into training, validation, and testing sets using a 70-15-15% ratio, as was done with the RF model. The implementation of the LSTM model was carried out using the TensorFlow Python library. The performance metrics for each model were calculated as described in the next section.

3.4. Metrics for performance assessment

The metrics used to assess model performance were the determination coefficient (R^2), mean absolute error (MAE) and root mean squared error (RMSE), as shown in Equations 7, 8 and 9, respectively.

$$R^2 = 1 - \frac{\sum(x_i - y_i)^2}{\sum(x_i - \bar{y})^2} \quad (7)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|x_i - y_i| \quad (8)$$

$$RMSE = \sqrt{\frac{\sum(x_i - y_i)^2}{n}} \quad (9)$$

Where x_i represents each observed value, y_i represents each predicted value, $\bar{y}$ is the mean of the predictions, and n is the number of observed-predicted value pairs (dataset size).

Each metric was calculated to evaluate the model's performance both across all basins collectively (for the RF model) and individually for each basin (RF and LSTM models).

3.5. Trend analysis

To find important increasing or decreasing trends in the reconstructed time series, the Mann-Kendall test was used. This non-parametric test is very commonly used for hydrometeorological data analysis (Abdullahi et al., 2023). It assesses the presence of a monotonic trend (either increasing or decreasing) over time without requiring the data to be normally distributed (hence, non-parametric). The test works by comparing each data point with all subsequent points in the series, counting the number of times a later value is higher or lower than an earlier one. The result is a test statistic that indicates the direction of the trend (increasing or decreasing), and a normalized Z value is used to determine its statistical significance. A positive Z value suggests an increasing trend, while a negative Z indicates a decreasing trend. If the absolute value of Z exceeds a certain threshold (based on the chosen significance level), the trend is considered statistically significant. Additionally, Kendall's Tau (τ) coefficient is reported to quantify the strength of the monotonic trend. Tau ranges from –1 to +1, where values close to zero indicate a weak trend and values farther from zero represent stronger increasing (positive τ) or decreasing (negative τ) tendencies. In this research, the test was employed using the 'pymannkendall' library in the python language.

3.6.Pearson's correlation

The correlation analysis throughout this research was employed with Pearson correlation, defined as:

$$r = \frac{\mathrm{Cov}(X, Y)}{\sigma_X \, \sigma_Y}$$

where Cov(X,Y) is the covariance between the variables X and Y, and $\sigma_X \; and \; \sigma_Y$ are their respective standard deviations. The Pearson correlation coefficient is a standardized measure of the linear association between two variables, ranging from –1 to +1. Values close to +1 indicate a strong positive linear relationship, values close to –1 indicate a strong negative linear

relationship, and values near 0 suggest little or no linear association. The values were interpreted as:

Table 2. Correlation interpretation intervals

Interval (±)	Level of correlation
0 – 0.20	No correlation
0.21 – 0.40	Weak correlation
0.41 – 0.70	Moderate correlation
0.71 – 1	Strong correlation

4 Results and Discussion

4.1. Random Forest reconstruction

The RF model's performance across all basins achieved an R^2 of 0.82, RMSE of 5.29, and MAE of 3.49. The performance metrics for each basin are presented in Table 3.

Table 3. Performance of the Random Forest model for each basin

Basin	R^2	RMSE	MAE	Basin size (km²)
Amazon	0.93	4.69	3.64	3,844,917.76
Paraguay	0.89	3.23	2.63	362,263.92
Tocantins-Araguaia	0.88	6.13	4.28	918,273.16
Paraná	0.85	3.75	2.49	877,513.54
Uruguay	0.83	2.38	1.85	174,127.78
Southeast Atlantic	0.80	3.99	3.28	213,316.01
South Atlantic	0.77	2.51	2.12	186,079.86
East Northeast Atlantic	0.72	3.33	2.70	285,281.21
Parnaíba	0.70	6.32	4.71	331,808.82
São Francisco	0.70	6.24	4.72	636,137.07

West Northeast Atlantic	0.69	11.71	7.42	268,906.01
East Atlantic	0.59	2.99	2.28	386,068.13

The basin that presented the best R^2 was the Amazon Basin (indicating that the model explained 93% of the variability in the data). In contrast, the East Atlantic Basin showed the lowest R^2 (0.59), meaning that the model explained only 59% of the data variability.

These results can be moderately attributed to the basin size, as indicated by a correlation of 0.52 between R^2 and basin area, meaning that the bigger the basin, the better the R^2. Moreover, no relationship was observed between the performance metrics and the level of basin anthropization, measured in this case by the proportion of anthropogenic LULC in each basin.

Regarding variable importance, Table 4 contains the top 5 most important variables for the reconstruction of each basin.

Table 4. The top 5 variables in Variable Importance for each basin

Basin	Variable	Importance
Amazon	Month	1.17
	Temperature	0.04
	Soil Moisture	0.02
	Rocky Outcrop	0.02
	Cotton	0.01
East Atlantic	Month	0.64
	Soil Moisture	0.11
	PDO	0.03
	Herbaceous Sandbank Vegetation	0.02
	SAODI	0.01
West Northeast Atlantic	Month	0.62
	Temperature	0.15
	Floodable Forest	0.02
	SAODI	0.01
	Mangrove	0.01
East Northeast Atlantic	Soil Moisture	0.20
	Temperature	0.09
	River, Lake and Ocean	0.05
	Precipitation	0.04
	Grassland	0.02
Southeast Atlantic	Month	0.97
	Soil Moisture	0.17

	River, Lake and Ocean	0.01
	Precipitation	0.01
	Wetland	0.01
South Atlantic	Month	0.51
	Soil Moisture	0.41
	Precipitation	0.04
	TSA	0.02
	SAODI	0.02
Paraguay	Month	1.28
	Soil Moisture	0.09
	Mosaic of Uses	0.03
	Temperature	0.01
	Grassland	0.01
Paraná	Month	0.62
	Soil Moisture	0.07
	River, Lake and Ocean	0.06
	PDO	0.03
	Savanna Formation	0.03
Parnaíba	Month	0.64
	Temperature	0.15
	Soil Moisture	0.13
	Mangrove	0.05
	Hypersaline Tidal Flat	0.02
São Francisco	Month	1.04
	Temperature	0.09
	Soil Moisture	0.07
	Other Perennial Crops	0.04
	River, Lake and Ocean	0.03
Tocantins-Araguaia	Month	1.05
	Temperature	0.19
	Soil Moisture	0.05
	ONI	0.02
	PDO	0.01
Uruguay	Soil Moisture	0.46
	Month	0.16
	Precipitation	0.09
	ONI	0.05
	PDO	0.04

The relevance of the month across all basins is expected, as seasonal variations are a fundamental factor in determining the climate of a region, as well as SM, since it is a component of TWS. Precipitation and temperature also emerged as influential variables, alongside the

teleconnection indices, whose effects are distributed unevenly across the different basins. The PDO emerges as a key driver in the East Atlantic, Paraná, Tocantins-Araguaia, and Uruguay basins, while the ONI shows relevance in Tocantins-Araguaia and Uruguay. TSA exerts influence exclusively in the South Atlantic basin, whereas the SAODI plays a role in the East Atlantic, West Northeast Atlantic, and South Atlantic. This spatial heterogeneity suggests that the sensitivity of TWSA to global climate variability differs across basins.

Regarding LULC, the model revealed a clear relationship between many different vegetation types and TWSA, with Floodable Forest, Savanna Formation, Grassland, Herbaceous Sandbank Vegetation and Mangrove emerging as relevant variables across different basins. Regions with humid vegetation typically exhibit higher TWSA values, while drier vegetation shows lower values, underscoring the connection between vegetation and TWSA. The presence of anthropic classes can be seen in the Paraguay basin (Mosaic of Uses), São Franciso (Other Perennial Crops) and Amazon (Cotton).

Mosaic of Uses encompasses different classes depending on the biome in question (Souza et al., 2020). In general, this class represents areas where it was not possible to differentiate between pasture and agriculture or, in the case of urban areas, other types of cultivated vegetation. However, since it represents anthropogenic activity, the results show that human influence contributes to the model's ability to predict TWSA values for the Paraguay basin.

The MapBiomas project does not specify which crops are included in the class “Other Perennial Crops”. Nevertheless, its classification as an anthropogenic feature indicates that, in the São Francisco Basin, TWSA variability is influenced by agricultural activity in general. This interpretation is consistent with previous findings for the Urucuia aquifer system (a major aquifer that covers 19.6% of the São Francisco Basin) where irrigation driven by diverse cropping systems has been shown to affect TWSA variability (Gonçalves et al., 2020).

For the Amazon Basin, cotton emerged among the top five most relevant variables. Its cultivation is highly dependent on irrigation and has undergone substantial expansion in recent years, particularly in the northern region of the state of Mato Grosso, located in the southern portion of the Amazon Basin (IPEA, 2022). The combination of this agricultural expansion and the prominence of cotton in the model suggests a potential correlation between cotton cultivation and TWSA, a relationship that has also been emphasized in previous studies in other parts of the world with this type of crop (Yin et al., 2023b).

Figure 9 presents the reconstructed time series for each basin, while Figure 10 summarizes the reconstructed series using a boxplot for all basins. When analyzing the complete time series (reconstructed + observed), the basin with the largest variation in TWSA is the West Northeast Atlantic, exhibiting an amplitude of 78.35 cm across the entire dataset. In contrast, the East Atlantic basin showed the smallest variation, with an amplitude of 20.69 cm.

The Mann-Kendall test, applied to the 30-year dataset, revealed statistically significant evidence of a decreasing trend in TWSA for 4 basins: Uruguay (p = 0.04), Parnaíba (p = 0.00), São Francisco (p = 0.00), East Atlantic (p = 0.02). The remaining basins did not exhibit any statistically significant trends ($p \geq 0.05$), indicating no evidence of a consistent increase or decrease over time. The test results for each basin are available in Table 5.

Table 5. Man-Kendall test statistics

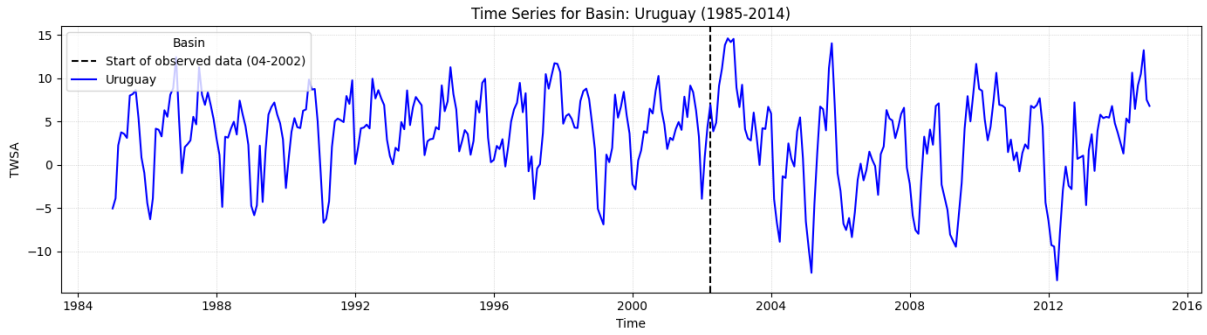
Basin	Trend	p value	Z score	τ
East Atlantic	Decreasing	0.0246	-2.2480	-0.0794
Parnaíba	Decreasing	0.0007	-3.3797	-0.1193
São Francisco	Decreasing	0.0004	-3.5577	-0.1256
Uruguay	Decreasing	0.0427	-2.0262	-0.0716
Amazon	No trend	0.3149	-1.0050	-0.0355
East Northeast Atlantic	No trend	0.4018	-1.7536	-0.0619
Paraguay	No trend	0.6375	-0.4712	-0.0167
Paraná	No trend	0.1211	1.5503	0.0548
South Atlantic	No trend	0.2823	-1.0751	-0.0380
Southeast Atlantic	No trend	0.3365	-0.9612	-0.0340
Tocantins-Araguaia	No trend	0.6871	-3.5577	-0.1256
West Northeast Atlantic	No trend	0.4018	0.8385	0.0296

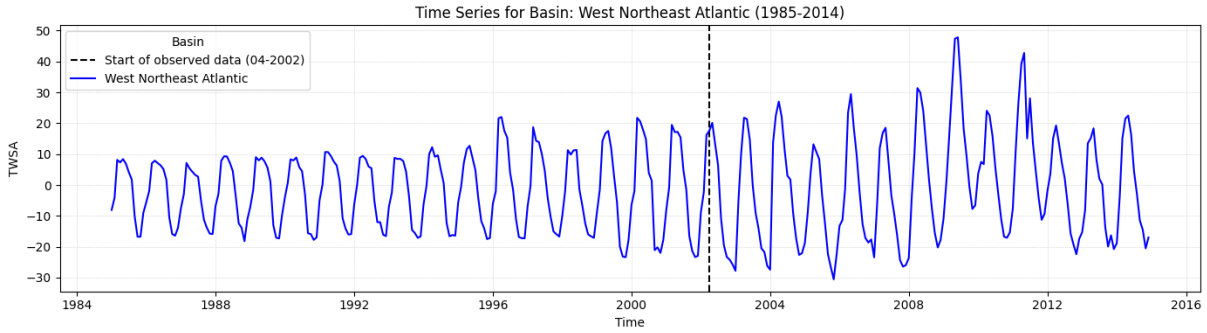
It is important to note that the reconstructed time series exhibits less amplitude compared to the observed time series. Since each basin has its own TWSA, P, T and SM spatial variability, the use of mean values could reduce part of the observed amplitude during processing, which may result in less accurate results. To evaluate whether intra-basin variability impacts performance, basin-level model R^2 was correlated with the observed variability of

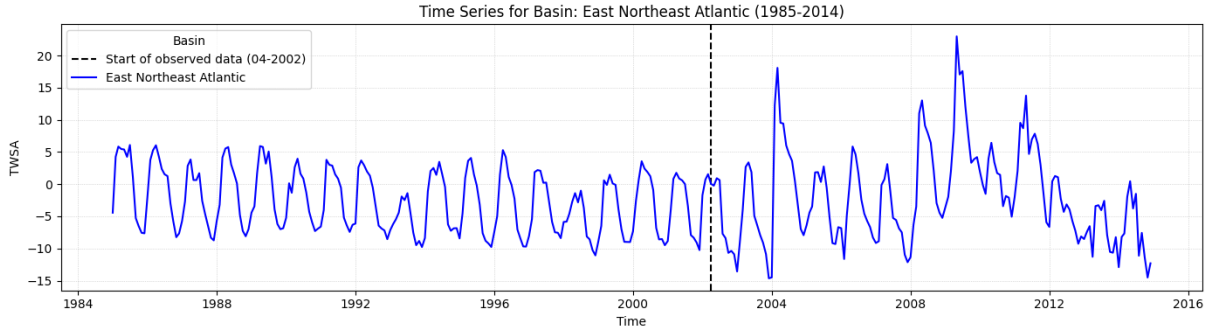
GRACE (standard deviation and variance). The correlations were r = 0.50 and r = 0.49, respectively, indicating a moderate positive association.

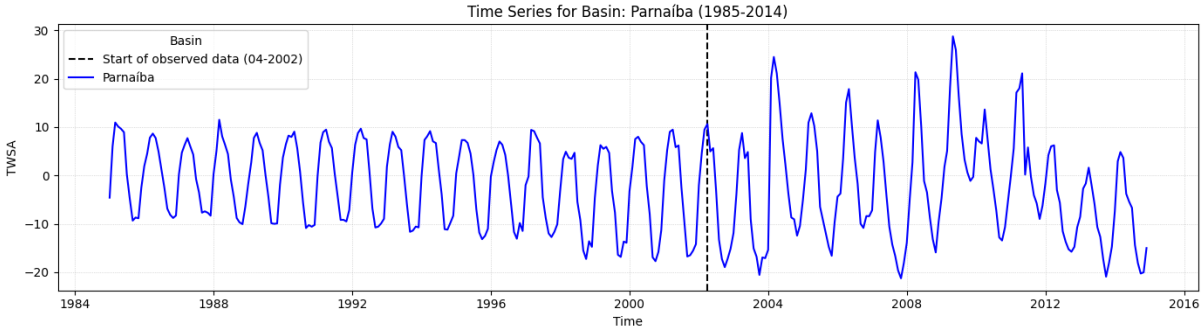
Another explanation that might contribute to this compression in amplitude is the fact that every time series contains a noise component that cannot be modeled. As a result, the observed time series shows greater variation, while the reconstructed series captures only the most significant part of the variation, excluding the noise, which is random and mostly unpredictable.

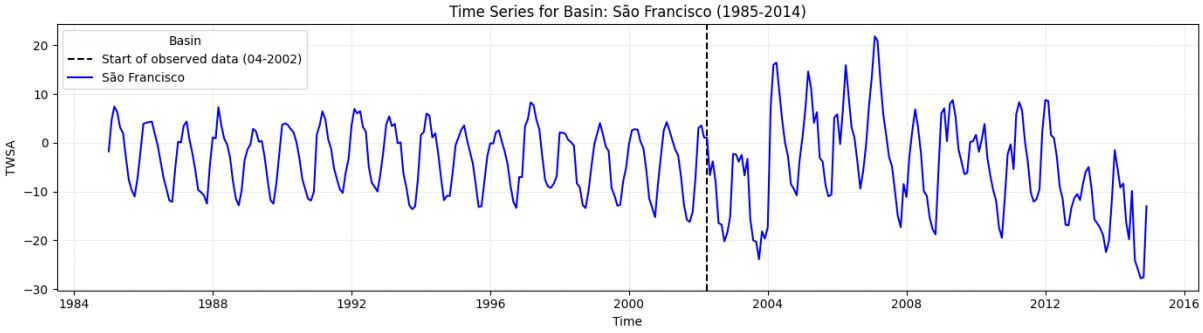
Additionally, Jing et al. (2020) caution that tree models can under-represent extremes outside the training range, which can also explain why the amplitude was not entirely represented in this model.

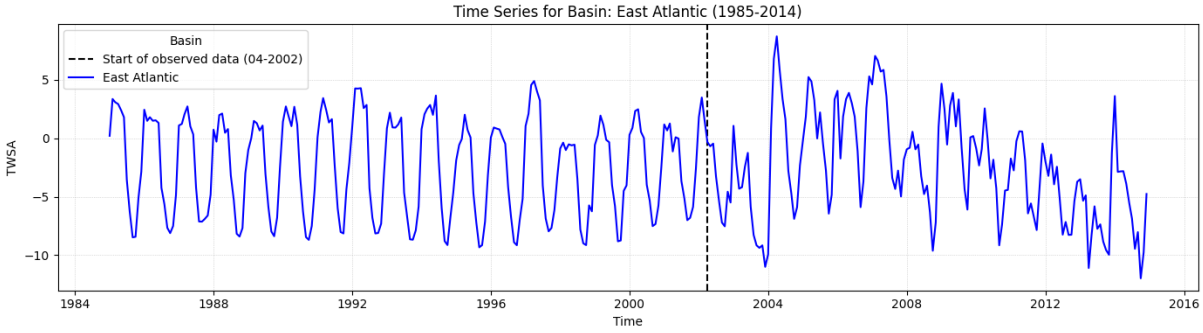


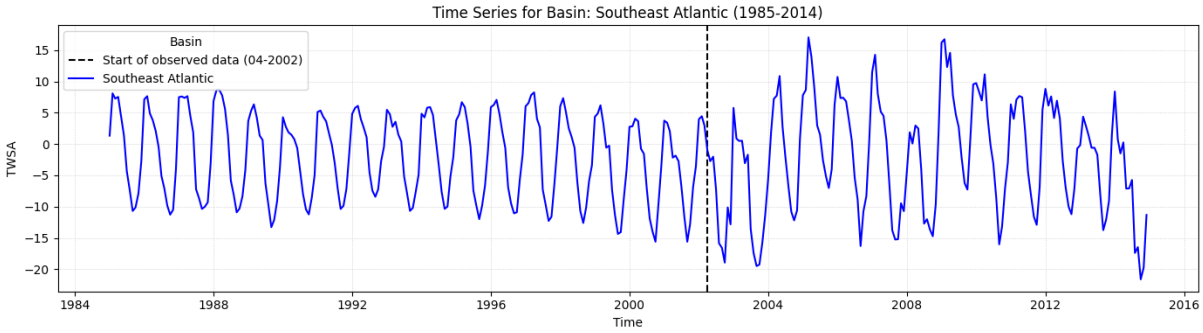


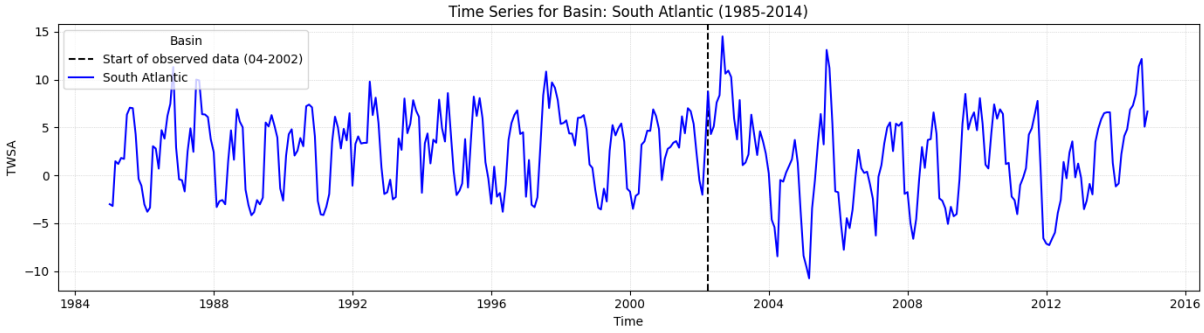


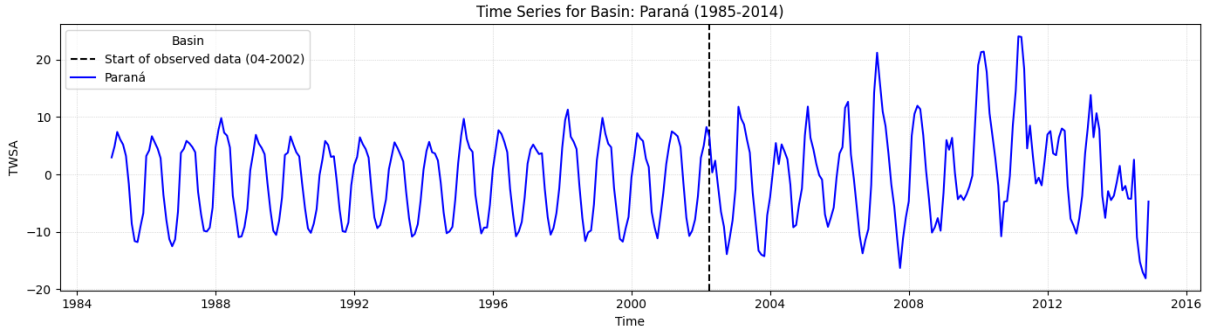

Time Series for Basin: Parnaíba (1985-2014)
Basin
Start of observed data (04-2002)
Parnaíba
TWSA
Time

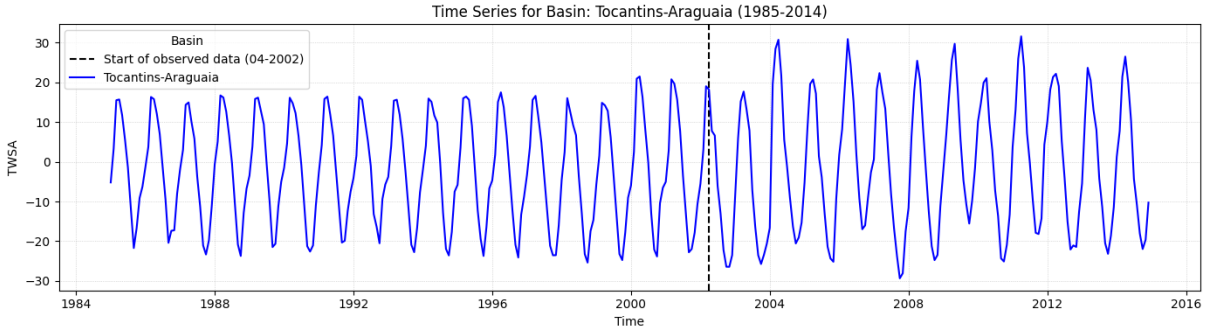

Time Series for Basin: São Francisco (1985-2014)
Basin
Start of observed data (04-2002)
São Francisco
TWSA
Time

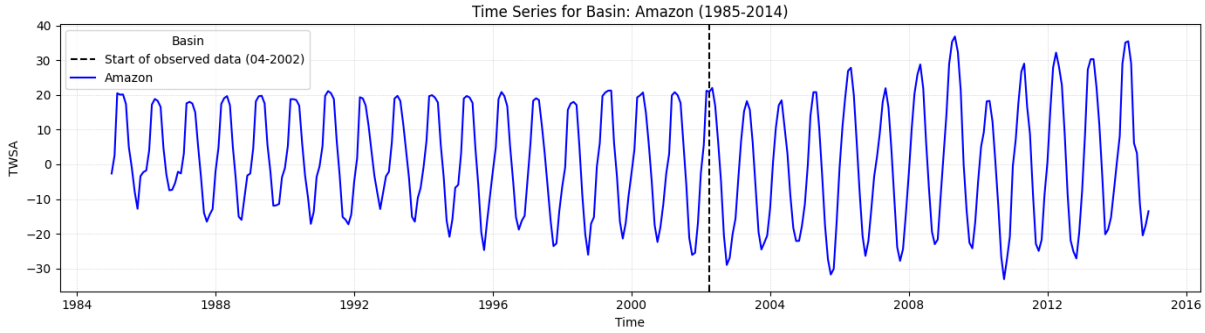

Time Series for Basin: East Atlantic (1985-2014)
Basin
Start of observed data (04-2002)
East Atlantic
TWSA
Time


Time Series for Basin: Southeast Atlantic (1985-2014)
Basin
Start of observed data (04-2002)
Southeast Atlantic
TWSA
Time


Time Series for Basin: South Atlantic (1985-2014)
Basin
Start of observed data (04-2002)
South Atlantic
TWSA
Time


Time Series for Basin: Paraná (1985-2014)
Basin
Start of observed data (04-2002)
Paraná
TWSA
Time


Time Series for Basin: Tocantins-Araguaia (1985-2014)
Basin
Start of observed data (04-2002)
Tocantins-Araguaia
TWSA
Time


Time Series for Basin: Amazon (1985-2014)
Basin
Start of observed data (04-2002)
Amazon
TWSA
Time

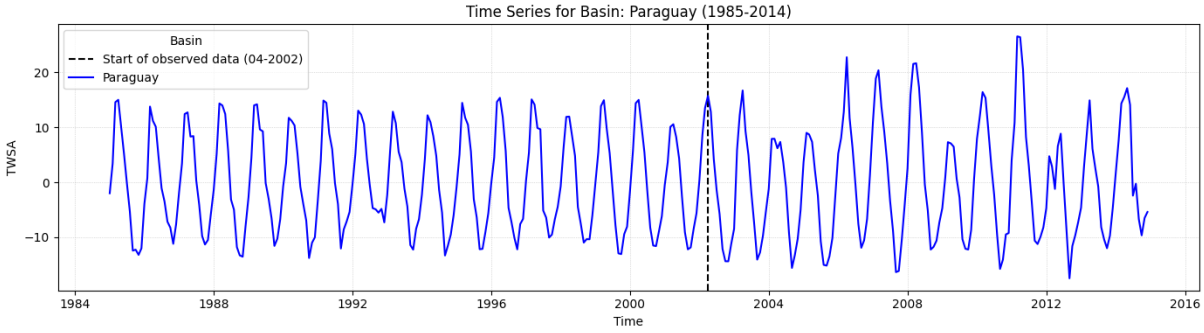


Figure 9. The complete time series (reconstructed (RF) + observed) for each basin.

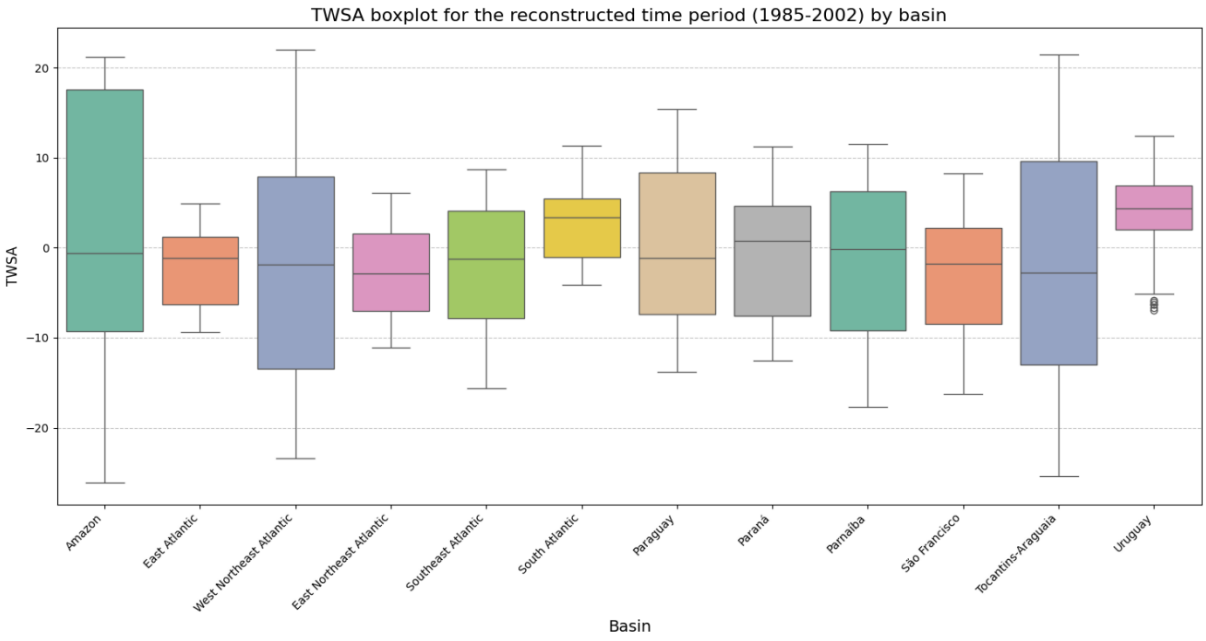


Figure 10. Boxplot for the reconstructed time series (RF)

By examining the boxplot, it becomes clear that the basins with the highest standard deviations (and therefore greater heterogeneity) are also the largest in size. This is expected, since larger basins are influenced by multiple hydroclimatic regimes and LULC patterns. This relationship is further supported by the strong correlation between basin size and standard deviation ($r = 0.95$).

The correlations between the reconstructed TWSA and the climate variables are presented in Table 6. The analysis revealed a weak negative correlation between TWSA and temperature ($r = -0.20$), as well as a weak positive correlation between TWSA and precipitation ($r = 0.30$). A moderate positive correlation was observed between TWSA and soil moisture ($r = 0.50$). However, no significant correlations were found between TWSA and the climate indices examined in this study.

Table 6. Correlations between reconstructed TWSA and climate data

Variable	Correlation with reconstructed TWSA
T	-0.20
SM	0.50
P	0.30
ONI	0.01
PDO	0.13
SAODI	0.03
TSA	0.11

Regarding land use classes, no significant correlations were observed between most classes and the reconstructed TWSA values. However, weak positive correlations were identified for Herbaceous Sandbank Vegetation (r = 0.29), Wooded Sandbank Vegetation (r = 0.27) and Citrus (r = 0.21).

4.2. Long Short-Term Memory Reconstruction

The LSTM model, due to its recurrent structure, was built independently for each basin to accurately represent sequential data. Given that the original dataset contained 12 monthly records per time step, the data had to be partitioned by basin. Consequently, a single performance metric applicable to all basins could not be established; instead, individual metrics were calculated for each basin. In this model, only 4 basins presented usable results (Table 7). For the remaining basins, the R^2 was close to 0 or negative, meaning that the model did not perform better than the simple use of the mean TWSA value for making the reconstruction.

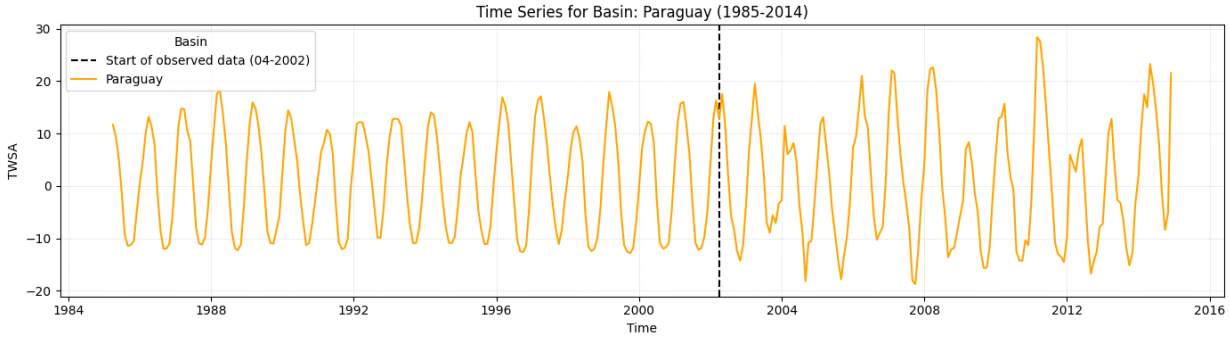
Table 7. Performance of the LSTM model for each basin

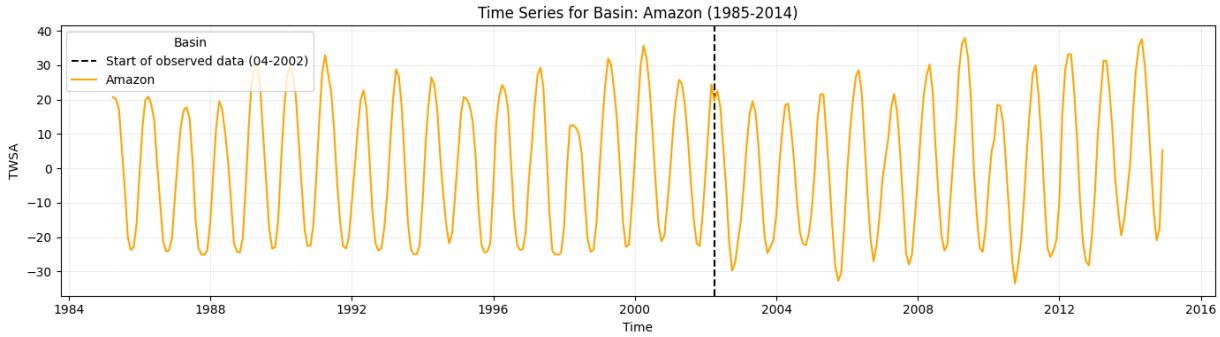
Basin	R^2	RMSE	MAE
Tocantins-Araguaia	0.96	3.17	2.53
Amazon	0.77	9.34	7.10
West-Northeast Atlantic	0.78	7.59	6.17
Paraguay	0.60	7.38	5.04

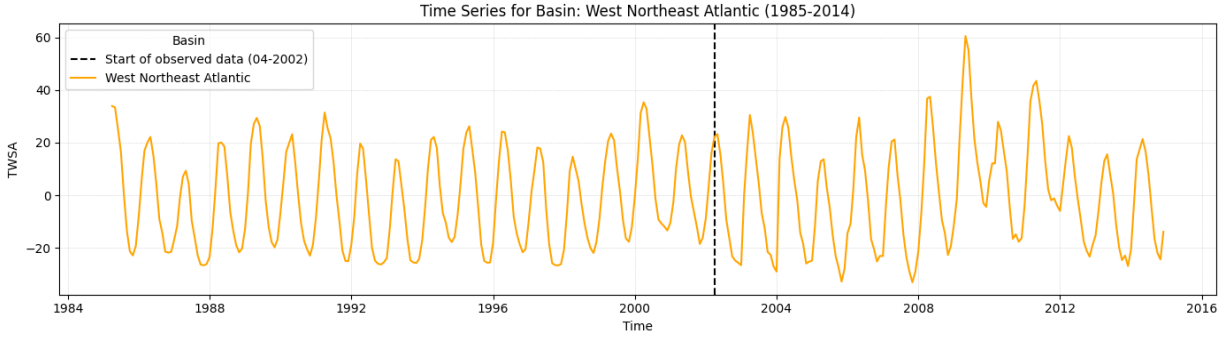
As was the case with RF, the model performance might be moderately attributed to basin size, since the correlation between RMSE and MAE was 0.53 and 0.59, respectively. There was no significant correlation between the size of the basin and the R^2 (0.08).

The observation that only four basins yielded good results in contrast with the successful application of LSTM models in previous studies (Wang et al., 2021) suggests that using identical parameters or predictive variables across a wide range of basins of different sizes with diverse climates and hydrological regimes is suboptimal. Ideally, the parameters or sets of predictive variables should be tailored to the specific characteristics of each basin being studied.

The reconstructed datasets for each basin are presented in Figure 11, while Figure 12 provides a summary of these datasets in the form of a boxplot. Among the basins analyzed, the West Northeast basin exhibited the greatest variability in the reconstructed dataset, with an amplitude of 93.60 cm. In contrast, the Paraguay basin showed the least variability, with an amplitude of 47.13 cm. The Mann-Kendall test applied to the entire time series indicated no significant trends for any of the reconstructed basins.







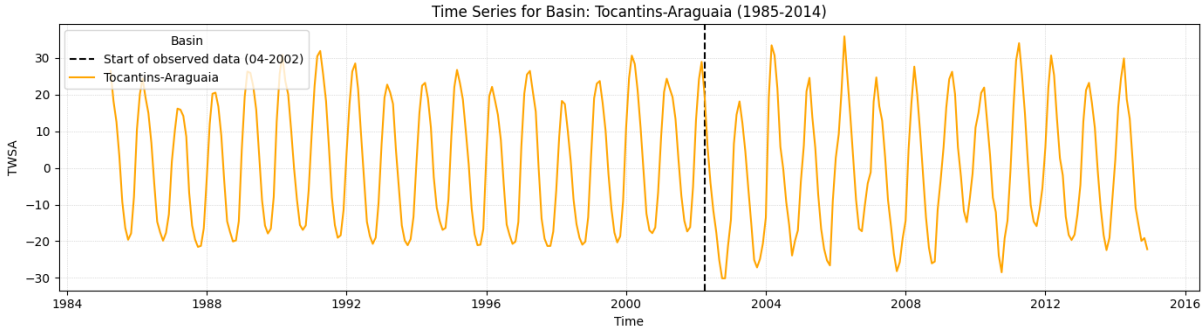


Figure 11. The complete time series (reconstructed (LSTM) + observed) for each basin.

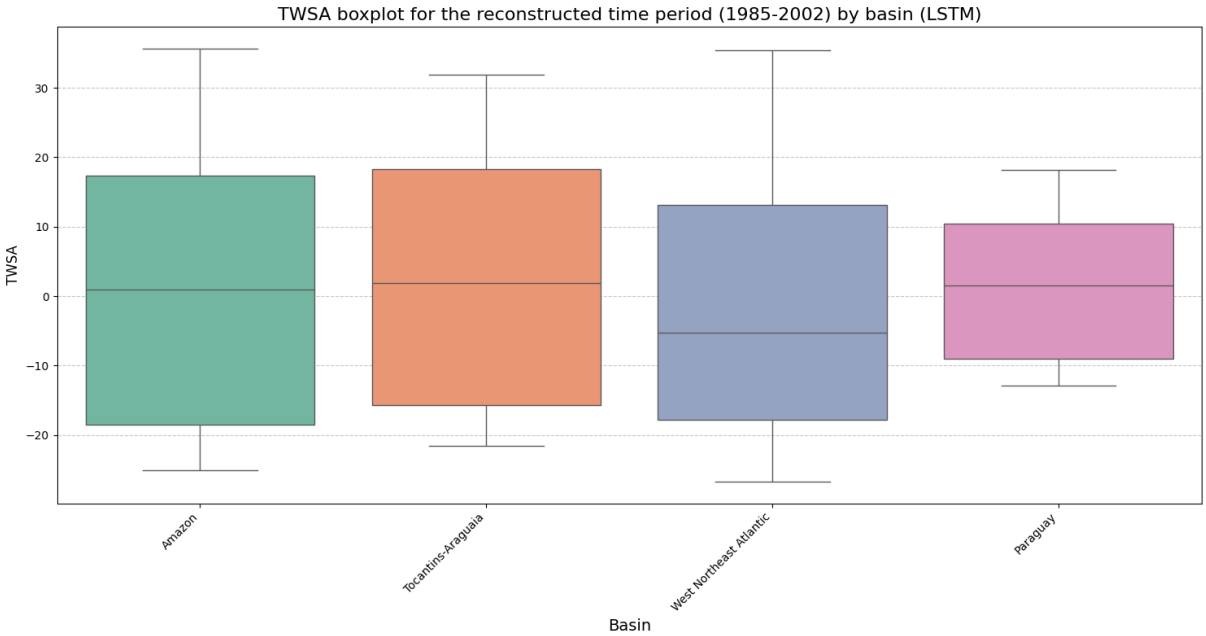
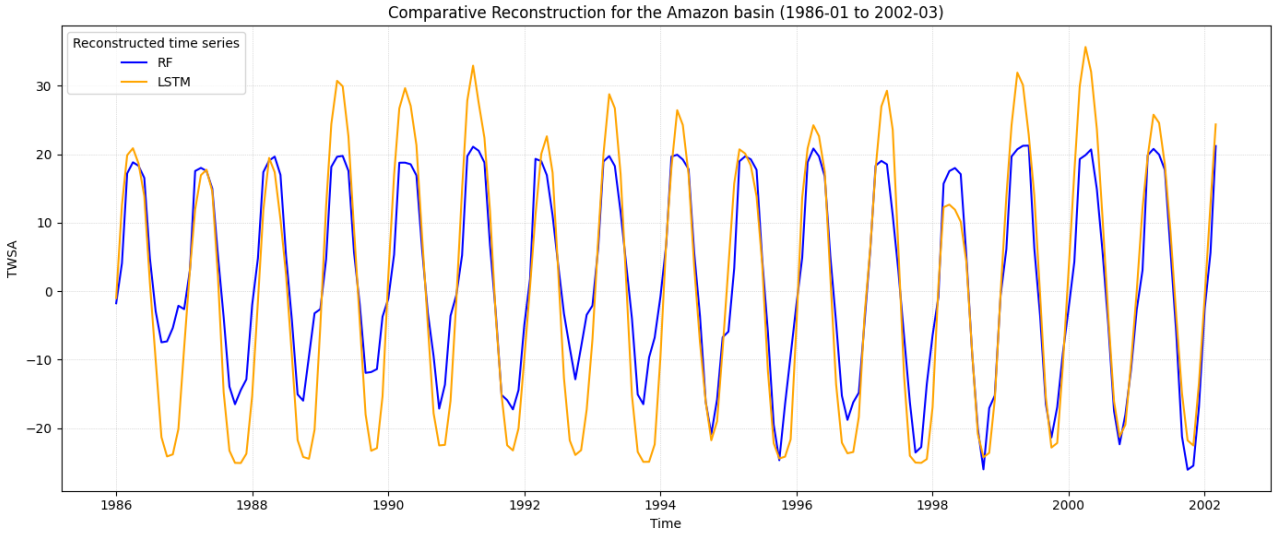
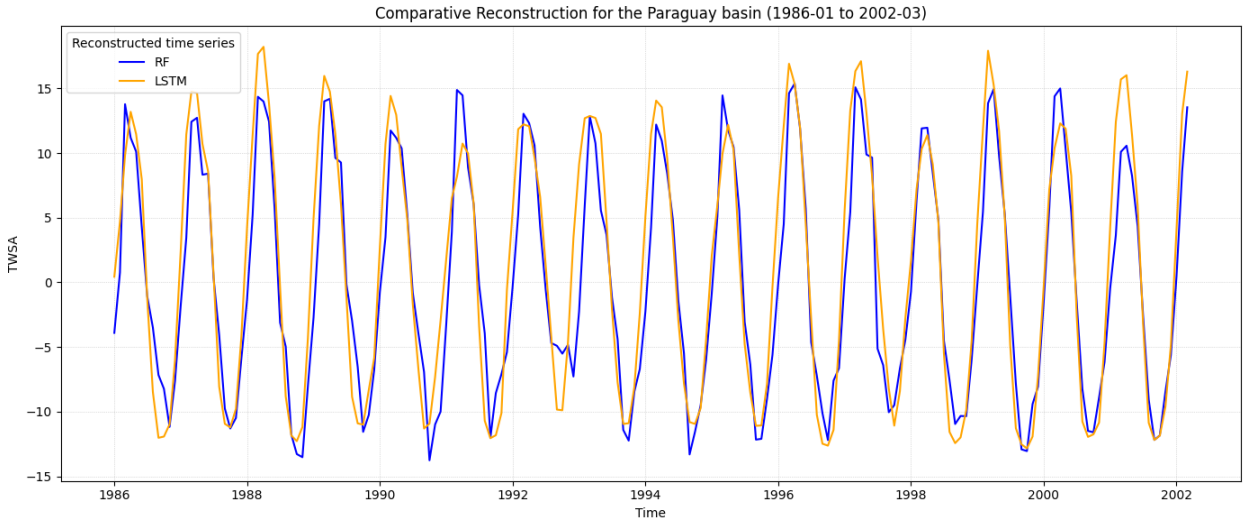
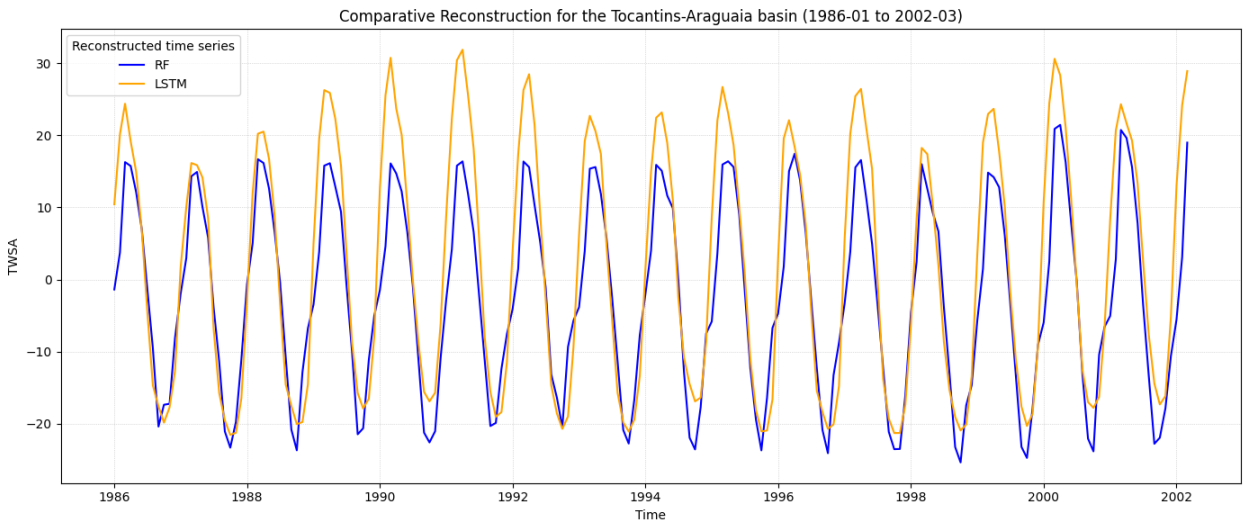


Figure 12. Boxplot for the reconstructed time series (LSTM)

When comparing the two reconstructions for the commonly analyzed basins (represented in Figure 13), strong positive correlations were observed across all cases. The correlations were 0.88 for the two West Northeast Atlantic basins and 0.93 for the remaining ones (Amazon, Tocantins-Araguaia, and Paraguay). These results indicate consistency between the reconstructions for these regions.


Comparative Reconstruction for the Amazon basin (1986-01 to 2002-03)
Reconstructed time series
RF
LSTM
TWSA
Time


Comparative Reconstruction for the Paraguay basin (1986-01 to 2002-03)
Reconstructed time series
RF
LSTM
TWSA
Time


Comparative Reconstruction for the Tocantins-Araguaia basin (1986-01 to 2002-03)
Reconstructed time series
RF
LSTM
TWSA
Time

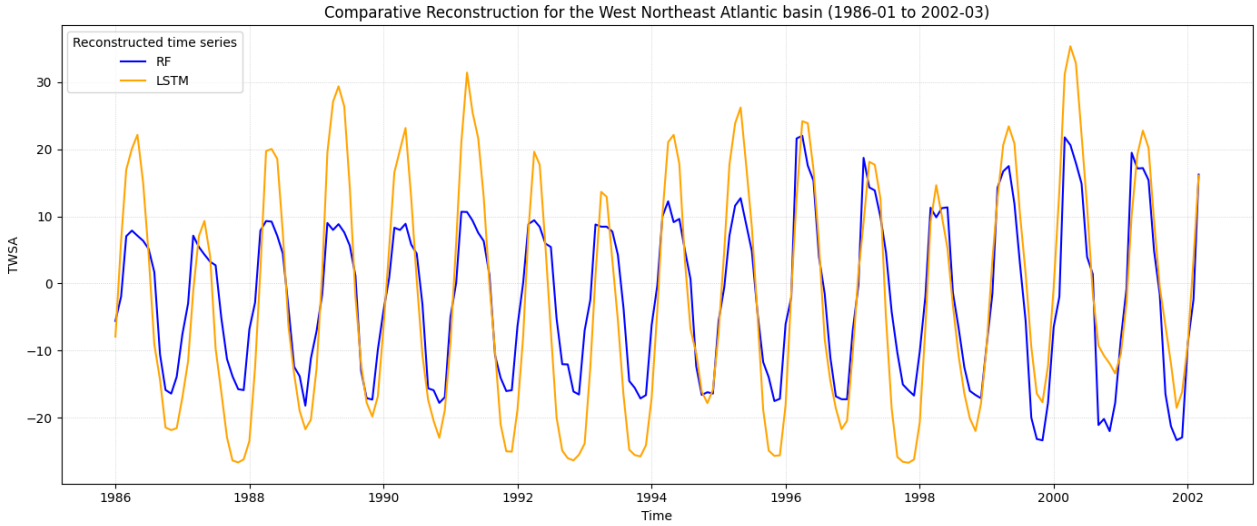


Figure 13. Comparison among the different reconstructions

Conclusions

The findings of this study underscore some conclusions regarding the performance and applicability of machine learning models for TWSA reconstruction for the major Brazilian basins. Firstly, the performance of both the RF and LSTM models shows a moderate dependence on basin size, with larger basins generally yielding better results. This suggests that basin size might play a role in the models' ability to explain variability, likely due to the reduced influence of localized noise in larger hydrological systems.

The RF model revealed critical insights into the relationships between TWSA and environmental and anthropogenic factors. Notably, the influence of vegetation on TWSA became particularly evident, as variables representing different types of vegetation emerged as key predictors. Additionally, anthropogenic factors, such as cotton, other crops and mosaic of uses class, exhibited a stronger relationship with TWSA than traditional climatic variables like precipitation and temperature. This finding highlights the impact of human activity on hydrological processes, reinforcing the importance of integrating land-use variables in hydrological models for certain basins. Furthermore, the RF model demonstrated its ability to detect significant decreasing trends in TWSA for some basins, highlighting its potential for identifying long-term changes in water availability.

In contrast, the LSTM model displayed significant variability in performance across basins, indicating that the use of uniform hyperparameters is not suitable for diverse hydrological contexts. The behavior of the hyperparameters was inconsistent, further

emphasizing the need for tailored approaches when applying LSTM models to basins with differing sizes, climates, and hydrological regimes. Furthermore, it is important to note as a limitation that the fact of the basin average was used limits the conclusions, since each basin has its own TWSA variability, even if the performance does not seem to be significantly correlated to basin TWSA variability. Hence, the data reconstructed in this research make more sense when analyzed in the long-term and for spatially bigger events, since it is only the basin average used for analysis. For future research, it is recommended to explore alternative neural network architectures and diverse sets of hyperparameters, in order to enhance model adaptability and robustness across the heterogeneous hydroclimatic contexts of Brazilian basins.

Despite these differences, both models demonstrated high consistency between themselves, capturing similar patterns and trends in the reconstructed data. This suggests that either model can be useful for understanding TWSA dynamics, with their applicability depending on the specific context and objectives of the analysis.

References

ABDULLAHI, N. I. et al. A NON-PARAMETRIC MANN-KENDALL AND SEN'S SLOPE ESTIMATE AS A METHOD FOR DETECTING TREND WITHIN HYDRO-METEOROLOGICAL TIME SERIES: A REVIEW. [S.l.: S.n.]. Disponível em: <www.academyjsekad.edu.ng>.

ALVARES, Clayton Alcarde et al. Köppen's climate classification map for Brazil. Meteorologische Zeitschrift, v. 22, n. 6, p. 711–728, 2013.

BREIMAN, L. Random Forests. Machine Learning, [S.L.], v. 45, n. 1, p. 5-32, 2001. Springer Science and Business Media LLC. http://dx.doi.org/10.1023/a:1010933404324

CAI, Wenju et al. Climate impacts of the El Niño–Southern Oscillation on South America. Nature Reviews Earth & Environment, v. 1, n. 4, p. 215–231, 10 abr. 2020.

CHEN, Zheng et al. The Impact of Precipitation Deficit and Urbanization on Variations in Water Storage in the Beijing-Tianjin-Hebei Urban Agglomeration. Remote Sensing, v. 10, n. 2, p. 4, 22 dez. 2017.

CUARTAS, Luz Adriana et al. Recent Hydrological Droughts in Brazil and Their Impact on Hydropower Generation. Water (Switzerland), v. 14, n. 4, 1 fev. 2022.

ENFIELD, David B. et al. How ubiquitous is the dipole relationship in tropical Atlantic sea surface temperatures? Journal of Geophysical Research: Oceans, v. 104, n. C4, p. 7841–7848, 15 abr. 1999.

FERREIRA, Antonio Geraldo; GIOVANNI, Namir; MELLO, Silva. PRINCIPAIS SISTEMAS ATMOSFÉRICOS ATUANTES SOBRE A REGIÃO NORDESTE DO BRASIL E A INFLUÊNCIA DOS OCEANOS PACÍFICO E ATLÂNTICO NO CLIMA DA REGIÃO 1 MAIN ATMOSPHERIC SYSTEMS ACTING OVER THE NORTHEASTERN REGION OF BRAZIL AND THE INFLUENCE OF PACIFIC ANS ATLANTIC OCEANS ON THE REGION CLIMATERevista Brasileira de Climatologia. [S.l.: S.n.].

FERREIRA, Vagner G. et al. A multi-sourced data retrodiction of remotely sensed terrestrial water storage changes for West Africa. Water (Switzerland), v. 11, n. 2, 1 fev. 2019.

FERREIRA, Vagner G. et al. Estimating groundwater recharge across Africa during 2003–2023 using GRACE-derived groundwater storage changes. Journal of Hydrology: Regional Studies, v. 56, p. 102046, dez. 2024.

GETIRANA, Augusto. Extreme water deficit in Brazil detected from space. Journal of Hydrometeorology, v. 17, n. 2, p. 591–599, 2016.

GIROTTO, Manuela; RODELL, Matthew. Terrestrial water storage. In: Extreme Hydroclimatic Events and Multivariate Hazards in a Changing Environment: A Remote Sensing Approach. [S.l.]: Elsevier, 2019. p. 41–64.

GONÇALVES, Roger D. et al. Using GRACE to quantify the depletion of terrestrial water storage in Northeastern Brazil: The Urucuia Aquifer System. Science of the Total Environment, v. 705, 25 fev. 2020.

HOCHREITER, Sepp; SCHMIDHUBER, Jürgen. Long Short-Term Memory. Neural Computation, v. 9, n. 8, p. 1735–1780, 1 nov. 1997.

HUMPHREY, Vincent; RODELL, Matthew; EICKER, Annette. Using Satellite-Based Terrestrial Water Storage Data: A Review. Surveys in Geophysics, v. 44, n. 5, p. 1489–1517, 13 out. 2023.

JANIESCH, Christian; ZSCHECH, Patrick; HEINRICH, Kai. Machine learning and deep learning. Electronic Markets, v. 31, n. 3, p. 685–695, 8 set. 2021.

JI, Lei; SENAY, Gabriel B.; VERDIN, James P. Evaluation of the Global Land Data Assimilation System (GLDAS) air temperature data products. Journal of Hydrometeorology, v. 16, n. 6, p. 2463–2480, 2015.

JING, Wenlong et al. Can Terrestrial Water Storage Dynamics be Estimated From Climate Anomalies? Earth and Space Science, v. 7, n. 3, 1 mar. 2020a.

JING, Wenlong et al. A data-driven approach to generate past GRACE-like terrestrial water storage solution by calibrating the land surface model simulations. Advances in Water Resources, v. 143, 1 set. 2020b.

JING, Wenlong et al. Extending GRACE terrestrial water storage anomalies by combining the random forest regression and a spatially moving window structure. Journal of Hydrology, v. 590, 1 nov. 2020c.

KARPATNE, Anuj et al. Machine Learning for the Geosciences: Challenges and Opportunities. IEEE Transactions on Knowledge and Data Engineering, v. 31, n. 8, p. 1544–1554, 1 ago. 2019.

KAYANO, Mary T.; ANDREOLI, Rita V. Relations of South American summer rainfall interannual variations with the Pacific Decadal Oscillation. International Journal of Climatology, v. 27, n. 4, p. 531–540, 30 mar. 2007.

LI, Fupeng et al. Comparison of Data-Driven Techniques to Reconstruct (1992–2002) and Predict (2017–2018) GRACE-Like Gridded Total Water Storage Changes Using Climate Inputs. Water Resources Research, v. 56, n. 5, 8 maio 2020.

LI, Fupeng et al. Long-Term (1979-Present) Total Water Storage Anomalies Over the Global Land Derived by Reconstructing GRACE Data. Geophysical Research Letters, v. 48, n. 8, 28 abr. 2021.

LONG, Di et al. Drought and flood monitoring for a large karst plateau in Southwest China using extended GRACE data. Remote Sensing of Environment, v. 155, p. 145–160, 1 dez. 2014.

LV, Meixia et al. Quantitative Analysis of Terrestrial Water Storage Changes Under the Grain for Green Program in the Yellow River Basin. Journal of Geophysical Research: Atmospheres, v. 124, n. 3, p. 1336–1351, 16 fev. 2019.

MCCULLOCH, Warren S.; PITTS, Walter. A logical calculus of the ideas immanent in nervous activity. The Bulletin of Mathematical Biophysics, v. 5, n. 4, p. 115–133, dez. 1943.

NIE, Ning et al. Reconstructed Terrestrial Water Storage Change (ΔTWS) from 1948 to 2012 over the Amazon Basin with the Latest GRACE and GLDAS Products. Water Resources Management, v. 30, n. 1, p. 279–294, 1 jan. 2016.

NNAMCHI, Hyacinth C.; LI, Jianping; ANYADIKE, Raymond N. C. Does a dipole mode really exist in the South Atlantic Ocean? Journal of Geophysical Research, v. 116, n. D15, p. D15104, 4 ago. 2011.

OLIVEIRA, Paulo Tarso S. et al. Trends in water balance components across the Brazilian Cerrado. Water Resources Research, v. 50, n. 9, p. 7100–7114, 2014.

PAIVA ALCOFORADO REBELLO, Vitor et al. Monitoring Drought in Brazil by Remote Sensing. In: [S.l.: S.n.]. p. 197–218.

REBOITA, Michelle Simões et al. Impacts of teleconnection patterns on South America climate. Annals of the New York Academy of Sciences, v. 1504, n. 1, p. 116–153, 2021.

RODELL, BY M. et al. THE GLOBAL LAND DATA ASSIMILATION SYSTEM This powerful new land surface modeling system integrates data from advanced observing systems to support improved forecast model initialization and hydrometeorological investigations. 2004.

ROLNICK, David et al. Tackling Climate Change with Machine Learning. ACM Computing Surveys, v. 55, n. 2, p. 1–96, 28 fev. 2023.

ROSSI, Júlia Brusso et al. Drought Propagation in Brazilian Biomes Revealed by Remote Sensing. Remote Sensing, v. 15, n. 2, 1 jan. 2023.

RUMELHART, David E.; HINTON, Geoffrey E.; WILLIAMS, Ronald J. Learning representations by back-propagating errors. Nature, v. 323, n. 6088, p. 533–536, out. 1986.

SARKER, Iqbal H. Machine Learning: Algorithms, Real-World Applications and Research Directions. SN Computer ScienceSpringer, , 1 maio 2021.

SOUZA, Carlos M. et al. Reconstructing three decades of land use and land cover changes in brazilian biomes with landsat archive and earth engine. Remote Sensing, v. 12, n. 17, 1 set. 2020.

SPENNEMANN, Pablo C. et al. A comparison of GLDAS soil moisture anomalies against standardized precipitation index and multisatellite estimations over South America. Journal of Hydrometeorology, v. 16, n. 1, p. 158–171, 2015.

SUN, Alexander Y. et al. Combining Physically Based Modeling and Deep Learning for Fusing GRACE Satellite Data: Can We Learn From Mismatch? Water Resources Research, 2019.

SUWEIS, Samir et al. Water-controlled wealth of nations. Proceedings of the National Academy of Sciences of the United States of America, v. 110, n. 11, p. 4230–4233, 12 mar. 2013.

TANG, Senlin et al. Random Forest-Based Reconstruction and Application of the GRACE Terrestrial Water Storage Estimates for the Lancang-Mekong River Basin. Remote Sensing, v. 13, n. 23, p. 4831, 28 nov. 2021.

TAPLEY, Byron D. et al. Contributions of GRACE to understanding climate change. Nature Climate ChangeNature Publishing Group, , 1 maio 2019.

TIAN, Kunjun et al. Drought events over the amazon river basin (1993–2019) as detected by the climate-driven total water storage change. Remote Sensing, v. 13, n. 6, 2 mar. 2021.

TURING, A. M. I.—COMPUTING MACHINERY AND INTELLIGENCE. Mind, v. LIX, n. 236, p. 433–460, 1 out. 1950.

WANG, Fei et al. Developing a long short-term memory (LSTM)-based model for reconstructing terrestrial water storage variations from 1982 to 2016 in the tarim river basin, Northwest China. Remote Sensing, v. 13, n. 5, p. 1–18, 1 mar. 2021.

WANG, Jielong et al. A deep learning model for reconstructing centenary water storage changes in the Yangtze River Basin. Science of the Total Environment, v. 905, 20 dez. 2023a.

WANG, Jielong et al. A deep learning model for reconstructing centenary water storage changes in the Yangtze River Basin. Science of the Total Environment, v. 905, 20 dez. 2023b.

WANG, Xinxin et al. Gainers and losers of surface and terrestrial water resources in China during 1989–2016. Nature Communications, v. 11, n. 1, p. 3471, 10 jul. 2020.

WU, Yu chen; FENG, Jun wen. Development and Application of Artificial Neural Network. Wireless Personal Communications, v. 102, n. 2, p. 1645–1656, 1 set. 2018.

XIONG, Jinghua et al. A Novel Standardized Drought and Flood Potential Index Based on Reconstructed Daily GRACE Data. Journal of Hydrometeorology, v. 23, n. 9, p. 1419–1438, set. 2022.

YANG, Chuanxun et al. Reconstructing the Historical Terrestrial Water Storage Variations in the Huang–Huai–Hai River Basin With Consideration of Water Withdrawals. Frontiers in Environmental Science, v. 10, 25 ago. 2022.

YANG, Peng et al. Reconstruction of terrestrial water storage anomalies in Northwest China during 1948–2002 using GRACE and GLDAS products. Hydrology Research, v. 49, n. 5, p. 1594–1607, 1 out. 2018.

YIN, Jiabo et al. GTWS-MLrec: global terrestrial water storage reconstruction by machine learning from 1940 to present. Earth System Science Data, v. 15, n. 12, p. 5597–5615, 8 dez. 2023a.

YIN, Jianguo et al. Regional Characteristics and Impact Factors of Change in Terrestrial Water Storage in Northwestern China From 2002 to 2020. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, v. 16, p. 386–398, 2023b.

ZHANG, Dan et al. GRACE-based hydrological drought evaluation of the Yangtze River basin, China. Journal of Hydrometeorology, v. 17, n. 3, p. 811–828, 2016.

ZHANG, Xu et al. Global hydroclimatic drivers of terrestrial water storage changes in different climates. Catena, v. 219, 1 dez. 2022.

ZHANG, Yafeng et al. The relative contributions of precipitation, evapotranspiration, and runoff to terrestrial water storage changes across 168 river basins. Journal of Hydrology, v. 579, 1 dez. 2019..