Final year internship report

Evaluation of the Global Gravity-based Groundwater Product (G3P) against *in-situ* groundwater observations

From the 08/04/2024 to the 28/06/2024

Autor: MORLIERE Sacha.
University: Sorbonne Université, Department SDUEE, Master STePE, Study-track HHGE, first year.
Published the: 13/06/2024.

Student: MORLIERE Sacha
Supervisors: LICTEVOUT Elisabeth and CAO Feifei
Tutor: JOST Anne

ABSTRACT

This study aims to evaluate the reliability of GRACE data for assessing groundwater storage variations in different climate and hydrogeological contexts. For that reason, the Paris Basin (France), Basin and Range (USA), and Ogallala (High Plains, USA) aquifers systems have been evaluated, from 2002 to 2023. The methodology involves validating GRACE data with in-situ measurement, processed with Specific Yield (Sy) to obtain groundwater storage anomalies (GWSA). The comparison uses Pearson correlation and Mann-Kendall tests to analyze linear correlation and trends. Results indicate a good correlation in the Paris Basin (0.65) and Ogallala aquifer (0.82), with slight discrepancies attributed to unconfined aquifers and Sy estimations. The Basin and Range aquifer shows a lower correlation (0.21), potentially due to climate contrasts and the impact of the Great Salt Lake. The study shows limitation of GRACE to perform in deep confined aquifers or arid regions and highlights the need for further investigations into the influence of hydrogeological characteristics and suggests improving in-situ data collection and analysis methods for better GRACE data validation.

ACKNOWLEDGMENTS

At the end of this internship, I would like to express my deep gratitude to all the people who contributed to the success of this professional experience.

First of all, I would like to warmly thank LICTEVOUT Elisabeth for her welcome within IGRAC and her supervision throughout this period. His availability and support were essential for my learning and the smooth running of my internship.

I would also like to express my particular gratitude to CAO Feifei, my tutor, for her support and the valuable help she gave me. His patience, wise advice and guidance throughout the project allowed me to take full advantage of this enriching experience.

Finally, a big thank you to the entire IGRAC team for their warm welcome and kindness. Their warmth and advice allowed me to progress and quickly integrate into the team.

TABLE OF CONTENTS

INTRODUCTION

1 M1 HHGE Internship

This internship is part of the 1st year of the HHGE Master's degree at Sorbonne University in Paris. It allows the student to discover the professional environment through a public or private organization, in connection with the student's professional project, and under the responsibility of a supervisor from this organization and a tutor from the home university.

The first year's internship is also an opportunity to train the student in the different stages of the scientific and technical process, such as: organizing and formalizing reasoning, processing raw data to obtain usable data, interpreting this information regarding the bibliography, and presenting the work in written and oral form.

The expectations of the internship are a report of thesis as well as an oral defense.

2 Presentation of IGRAC

The International Groundwater Resources Assessment Centre (IGRAC) facilitates and promotes global exchange of essential information and expertise necessary for the sustainable development and management of groundwater resources in the world. Since its establishment in 2003, IGRAC has been committed to contribute to systematic groundwater monitoring and assessment, to encourage information sharing, awareness raising and strategic partnerships.

Headquartered in Delft, the Netherlands, IGRAC operates in close collaboration with the IHE Delft Institute for Water Education. Recognized as the UNESCO Global Groundwater Centre and operating under the auspices of the World Meteorological Organization (WMO), IGRAC is also a valued corporate partner of the International Association of Hydrogeologists (IAH), with financial backing from the Government of the Netherlands.

IGRAC's main mission is to provide information and knowledge on groundwater worldwide to support decision-making and contributing to a sustainable planet. Aligned with its core mission of contributing to the improvement of groundwater management globally, IGRAC's main activities include developing and maintaining an open-access global database on groundwater, performing evidence-based groundwater assessment and research, supporting capacity development and awareness raising worldwide.

3 Background and description of the project

Groundwater is an essential factor for both ecosystems and humanity. It ensures ecosystem stability, energy, and food security, and promotes human health. Groundwater is the largest component of global liquid freshwater resources in the water cycle. It accounts for 33% of the global water withdrawals by mankind, with more than two billion people depending on groundwater as primary water resource. In many regions of the world, groundwater is the decisive factor for agricultural productivity, as it is intensively used for irrigation.

Despite its importance, groundwater has been constantly undervalued and mismanaged. Overexploitation of groundwater resources and groundwater contamination by pollutants have led to groundwater depletion in many parts of the world, causing current and foreseeable water crises. Climate change poses additional stress on groundwater resources, particularly in areas already experiencing water scarcity. Groundwater depletion impacts freshwater availability, causing potential serious consequences, including reduced agricultural productivity, increased food prices to food scarcity, land subsidence, seawater intrusion into estuaries and coastal aquifers, loss of springs and wetlands, ecosystems degradation, regional climate feedback following reduced evapotranspiration, and intra- and international political unrest.

Recognizing the importance of groundwater resources, the Global Climate Observing System (GCOS) defined groundwater as one of the Essential Climate Variables (ECV). While several other ECVs are accessible as operational products in the portfolio of the Copernicus Services, groundwater is not. The reason for this, but also for limited data availability on groundwater in general for many regions worldwide, is in part due to its obscure hidden nature below the Earth’s surface. Poor in-situ monitoring capabilities widely prevail, with sparse and un-representative groundwater monitoring networks, largely unknown storage capacities and specific yields of the aquifers, and inaccessibility of data. New methods of groundwater observation and resource quantification are thus urgently needed.

Thanks to the German American Gravity Recovery and Climate Experiment (GRACE) satellite mission (2002-2017) and its successor GRACE-FO (since 2018), it has been possible to collect monthly data on global Terrestrial Water Storage (TWS) variations for more than 20 years. The working principle involves measuring the time-variable gravity field of the Earth. At an orbital altitude of about 450 km, the distance between the two GRACE satellites was on average 220 km, but constantly changing due to the varying attraction of masses on the surface and inside the Earth. Repeated observations by this constellation ultimately yield changes in the mass distribution, and provide a basis to derive changes in TWS, which can then be converted into changes in groundwater storage.

The Global Gravity-based Groundwater Product (G3P) developed under the EU Horizon 2020 program, provides a gap-filled product of monthly groundwater storage variations with global coverage from 2002 until present, by the cross-cutting combination of GRACE and GRACE-FO satellites gravity data with water storage data that are based on other water storage compartments (WSCs) from various services that are already part of the Copernicus portfolio.

As a key participant of the G3P consortium (containing 12 partner organizations), IGRAC takes lead in the validation and evaluation of G3P data against *in-situ* observations.

4 *Missions and objectives of the internship*

Previous validation of G3P data was carried out by IGRAC in 2022, highlighting an overall good performance of the product in tracking long-term groundwater dynamics. New validation phase is underway in 2024 using the latest version of G3P times series data and updated methods for processing in-situ data and generating specific yield values, with the purpose of enhancing the utility and reliability of the G3P product.

The core mission of this internship is to contribute to the validation process by comparing G3P data with *in-situ* groundwater level data through correlation analysis and long-term trend assessments at three selected aquifers system across the world: Paris Basin (France), Basin and Range (USA) and Ogallala (High Plains) (USA).

The main objectives are (1) to evaluate the performance of the G3P product using the refined validation methodology and (2) to elucidate factors affecting consistency or disparity between in-situ and G3P data under various climate/hydrogeological contexts.

MATERIALS AND METHODS

1 *G3P time series extraction: 2022 vs 2024*

The G3P provides direct usable monthly dataset on GroundWater Storage Anomalies (GWSA) with a spatial resolution of 0.5°.

The working principle involves a subtraction process in which total Terrestrial Water Storage (TWS) variations measured from the satellites are reduced to GWSA by subtracting all other Water Storage Compartments (WSCs) *(Figure 1)*. Those 4 WSCs are Soil Moisture (SM), Glaciers, Snow Water Equivalent (SWE) and Surface Water Storage (SWS). Sometimes, some of those WSCs must be resampled monthly and to 0.5° spatial resolution.

Figure 1: GWSA calculation of G3P (Source: Güntner and al, 2023)

The uncertainty of the resulting GWSA of the G3P subtraction process is calculated by propagating the uncertainties of the individual WSCs *(Equation 1)*. Each WSC in G3P provides a global layer that contains their monthly time-varying uncertainties σ at the scale of the 0.5 grid cells *(Güntner and al, 2023)*.

$$\sigma_{GWSA} = \sqrt{(\sigma_{TWS})^2 + (\sigma_{SWE})^2 + (\sigma_{RZSM})^2 + (\sigma_{SWS})^2 + (\sigma_{Glacier})^2}$$

Equation (1)

GRACE data are also subject to filtering and spatial leakage management processes. To reduce high-resolution spatial noise and ensure data consistency, a Gaussian filter with an optimal length of 200 km is applied to GRACE data. This filtering smooths the variations in WSCs to better align with the characteristics of TWS data. While large-scale trends are preserved, this process also results in the loss of some local information. Additionally, the smoothing of GRACE data reduces signal amplitudes and causes signal propagation between adjacent areas, known as the leakage effect *(Figure 2)*. These implications must be considered when interpreting our results. All G3P data sets are provided in NetCDF format.

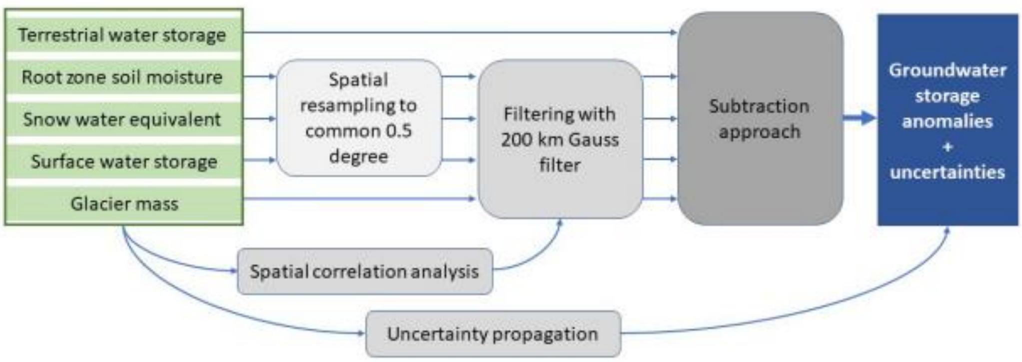

Figure 2: Flowchart of G3P working principle (Source: Güntner and al, 2023)

During the previous validation in 2022, G3P v1.5 version was used, offering data from 2002 to 2020. G3P data gaps due to instrumental errors on board the satellites, or during the transition from GRACE to GRACE-FO in 2017-2018 were left unfilled for the evaluation. The new evaluation phase in 2024 uses the latest v1.12 version of G3P, offering data from 2002 to 2023. In addition, a linear interpolation is applied to fill the remaining gaps due to instrumental errors, except the 2017-2018 satellites shift one. Continuous efforts are made to reduce uncertainties in the WSCs by updating G3P versions.

In summary, for the new validation phase, the latest version dataset (v1.12) was used, covering a larger period. By applying a linear interpolation, $GWSA_{G3P}$ gaps present in the data due to instrumental errors were filled, except the 2017-2018 satellites shift one *(Table 1)*.

Table 1: Synthesis between old and new G3P extraction method

	2022 VALIDATION (OLD ONE)	***2024 VALIDATION (NEW ONE)***
G3P VERSION	v1.5	v1.12
CONSIDERED PERIOD	From April 2002 to December 2020	From April 2002 to December 2023
INTERPOLATION	None	Yes (linear one to fill gaps)
USED SATELLITES	GRACE and GRACE-FO	GRACE and GRACE-FO
DATA FREQUENCY	Monthly	Monthly
DATA RESOLUTION	0.5°	0.5°

2 *In-situ time series extraction: 2022 vs 2024*

In order to validate $GWSA_{G3P}$ against in-situ groundwater observations, $GWSA_{in\text{-}situ}$ can be calculated as follows:

$$GWSA_{in-situ} = GWLA_{in-situ} * S_y \qquad \text{Equation (2)}$$

with $GWLA_{in\text{-}situ}$: GroundWater Level Anomaly, S_y: Specific Yield.

Depths to groundwater data are from the Global Groundwater Monitoring Network of IGRAC (GGMN: ggmn.un-igrac.org), which contains groundwater monitoring data shared from national groundwater monitoring networks[1]. In order to better represent the average change in the groundwater storage of an aquifer, the main objective was to include as much in-situ data as possible. Therefore, the decision was made to relax data restrictions initially. During the previous evaluation phase in 2022, only full datasets of monthly groundwater data were used.

[1] Paris Basin (France): data shared by French Geological Survey (BRGM) via the ADES platform (https://ades.eaufrance.fr/); Basin and Range aquifers and Ogallala aquifer: data shared by the United States Geological Survey (USGS) via the National Ground-Water Monitoring Network (NGWMN) Data Portal (https://cida.usgs.gov/ngwmn/index.jsp).

This limited calculations on restricted areas and short periods, leading to low representativity: it was not robust to use groundwater level data in a few wells to represent the situation of a whole aquifer. In 2024, the decision was made to filtered raw data by allowing to take boreholes with a maximum gap of 20% of their total dataset, without exceeding 6 consecutive months. The gaps were then filled with a linear interpolation and extrapolation before GWLA calculations. This allowed to not only increase the number and density of boreholes used, but also the period considered as well as the area evaluated. Another difference between 2022 and 2024 evaluation processes lies in the determination of the Specific Yield (S_y) parameter. In most scenarios during the 2022 evaluation process, hydrogeological data such as S_y was rarely available. Therefore, $GWSA_{in\text{-}situ}$ could not be calculated (*Equation 2*). In this case, $GWLA_{in\text{-}situ}$ was used as a proxy for $GWSA_{in\text{-}situ}$, and the evaluation was carried out by comparing the standardized time series of $GWSA_{G3P}$ and $GWLA_{in\text{-}situ}$ to evaluate the pattern correlation. In 2024, the S_y was determined using a surface-lithology-based approach. Those data were provided by GFZ German Research Centre for Geosciences. An S_y value has been attributed for each borehole depending on the associated surface-geological structure. Therefore, it allowed to determine $GWSA_{in\text{-}situ}$ for each borehole following the previous equation. Finally, a grid-based average method was applied: instead of simply averaging all wells together at each time step as it was done in 2022 validation, each well was assigned to a cell of a grid identic to the one used for the G3P data (0.5° resolution). The method consisted, in each cell, to calculate a monthly-average of each $GWSA_{in\text{-}situ}$ of each borehole, then average all the cells of an aquifer together. Contrary to the simple average method used in 2022, this new 2024 approach allowed us to better represent the aquifers, especially the ones having a less good boreholes distribution (such as boreholes clusters).

In summary, for the new validation phase, more groundwater level data was used, as well as S_y determined using a surface-lithology-based approach. By applying a linear interpolation, gaps present in the data were filled *(Table 2)*.

Table 2: Synthesis between old and new in-situ extraction method

	2022 VALIDATION (OLD ONE)	***2024 VALIDATION (NEW ONE)***
AUTHORIZED GW LEVEL GAP	0%	20% and < 6 consecutive months (filled with linear interpolation and extrapolation)
S_Y AVAILABILITY	Rarely	Always (via surface-lithology-based approach)
FINAL DATA USED FOR COMPARISONS WITH $GWSA_{G3P}$	$GWLA_{in\text{-}situ}$	GWLA and $GWSA_{in\text{-}situ}$
MONTHLY AVERAGE METHOD	Normal one (on all boreholes)	Grid-based approach one (on all cells, each one averaging the boreholes it contains)
DATA FREQUENCY	Monthly	Monthly
DATA RESOLUTION	0.5°	0.5°

3 Goodness-of-fit evaluation

In order to evaluate the correlation between $GWSA_{G3P}$ with GWLA/$GWSA_{in\text{-}data}$ time series, we decided to use two main different approaches. For the comparison with $GWSA_{in\text{-}data}$, we also compare the amplitude.

- Pearson correlation coefficient “r”:

The Pearson correlation coefficient, often denoted as “r”, is a statistical measure that calculates the strength and direction of the linear relationship between two time series. It is a widely used measure in statistics and can range from -1 to 1, where 1 indicates a perfect positive linear relationship and -1 indicates a perfect negative linear relationship. This correlation coefficient is not affected by the amplitude variations between both time series, but it has been chosen as a measure parameter since it can determine whether the signals of two time series are related.

For each aquifer, we also calculated the Pearson correlation coefficient for each cell within our grid. This means that the time series of $GWLA_{in\text{-}situ}$ and $GWSA_{G3P}$ for each cell were used to compute this coefficient. This allowed us to capture the spatial variability of correlation within the aquifer.

- Mann-Kendall analysis (comparison of long-term trend):

The Mann-Kendall Trend Test (sometimes called the MK test) is used to analyze time series data for consistently increasing or decreasing trends (monotonic trends). It is a non-parametric test, which means it works for all distributions (i.e. data doesn't have to meet the assumption of normality), but data should have no serial correlation. Mann-Kendall Test is a powerful trend test, so several others modified Mann-Kendall tests like Multivariate MK Test, Regional MK Test, Correlated MK test, Partial MK Test, etc. Seasonal Mann-Kendall test also developed to remove the effect of seasonality. Currently, the package used in that report has 11 Mann-Kendall Tests and 2 Sen’s slope estimator function, and the Seasonal Mann-Kendall test was used here *(Hussain et al., 2019)*.

In summary, for all following aquifers, all $GWSA_{G3P}$, $GWLA_{in\text{-}situ}$ and $GWSA_{in\text{-}situ}$ were calculated for the blue grid. Thereafter, normalized time series from $GWLA_{in\text{-}situ}$ and $GWSA_{G3P}$ were compared using the Pearson correlation coefficient. A grid-based correlation map was thereafter calculated between $GWLA_{in\text{-}situ}$ and $GWSA_{G3P}$. Then $GWSA_{in\text{-}situ}$ and $GWSA_{G3P}$ were compared using the Pearson correlation coefficient, and MK tests have been run on both time series with the purpose to compare long-term trend in groundwater storage.

RESULTS AND DISCUSSIONS

1 *Paris Basin*

The Paris Basin is located in the north of France, covering an area of approximately 150.000 km^2. It is composed of seven main sandstones-aquifer layers with carbonated-semi-permeable layers in-between *(Contoux et al., 2013; Megnien, 1980)*. Its thickness can reach up to 3.000 m. With a tempered climate, the mean annual precipitation is around 650 mm, and temperatures oscillate between four and sixteen degrees Celsius during winter and summer, respectively *(Güntner and al, 2023)*. The primary uses of groundwater from the Paris Basin aquifer include supplying drinking water, industrial applications, thermal uses, and a bit of irrigation. In-situ data include 440 boreholes, compared to 310 in the 2022 validation phase (+42%) *(Figure 3)*.

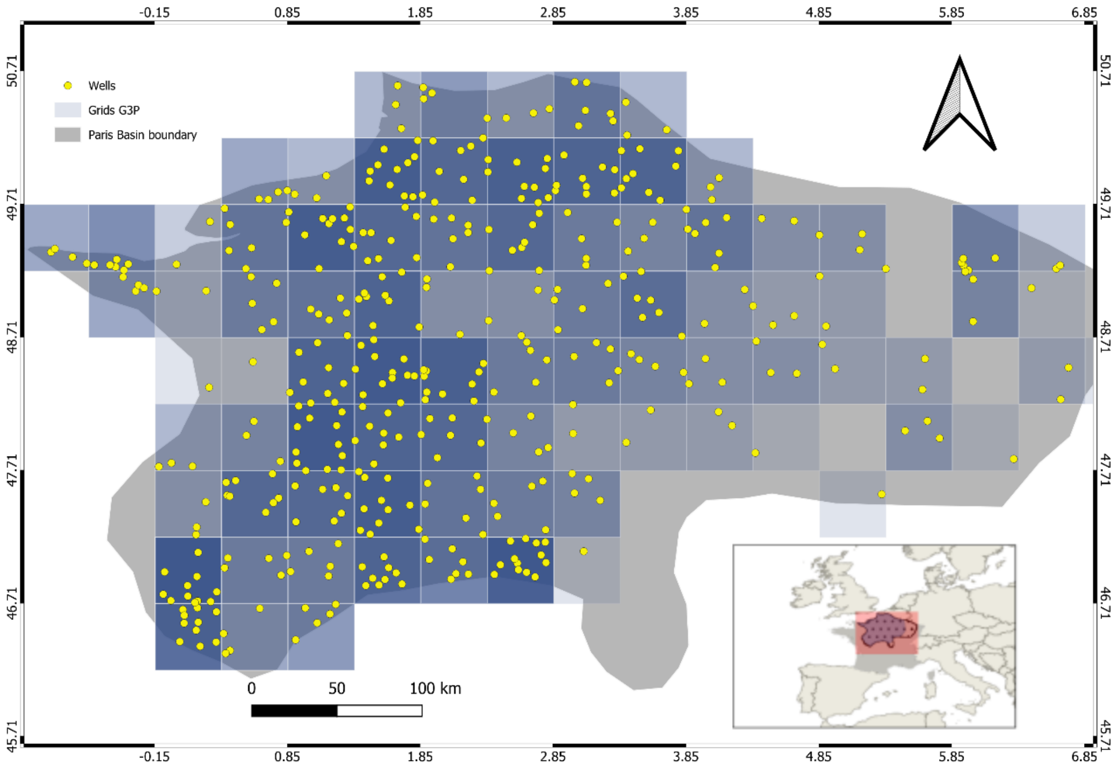

Figure 3: Paris Basin. Blue color indicates the wells density.

According to the results, the Pearson correlation coefficient of normalized data is high (0.63), indicating an overall good linear positive correlation with in-situ data *(Figure 4)*. The Pearson correlation coefficient of GWSA data is higher (0.65), presenting an even better linear positive correlation with in-situ data with the addition of Sy values *(Figure 5)*. However, in both, G3P data is higher than in-situ data before the satellite switch, and lower afterward. G3P data also tends to overestimate the amplitude of seasonal variations over the whole period. Finally, both MK trends are decreasing, even if G3P's slope is higher that the in-situ one *(Figure 5)*. Both are statistically significant.

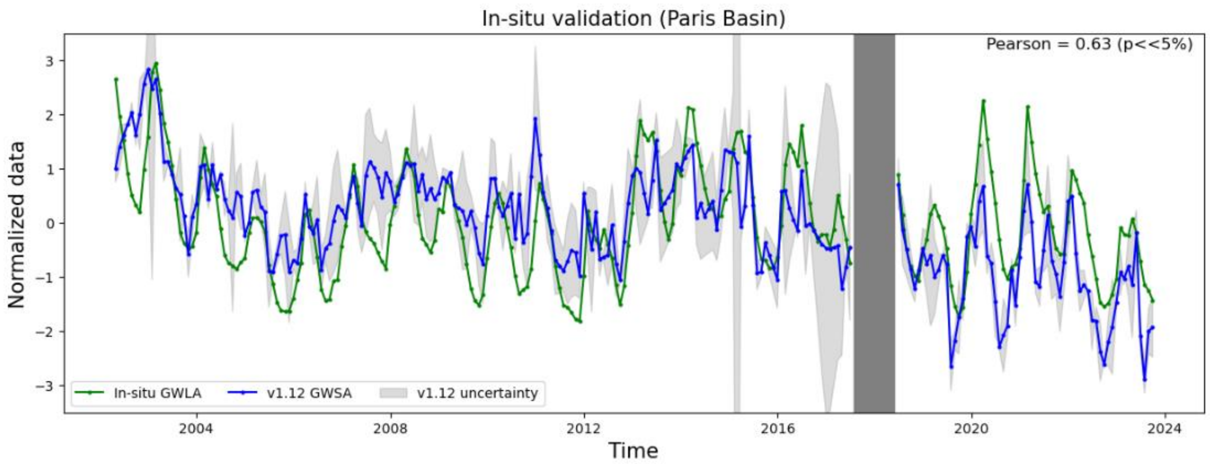

Figure 4: $GWSA_{G3P}$ vs $GWLA_{in\text{-}situ}$ (Paris Basin)

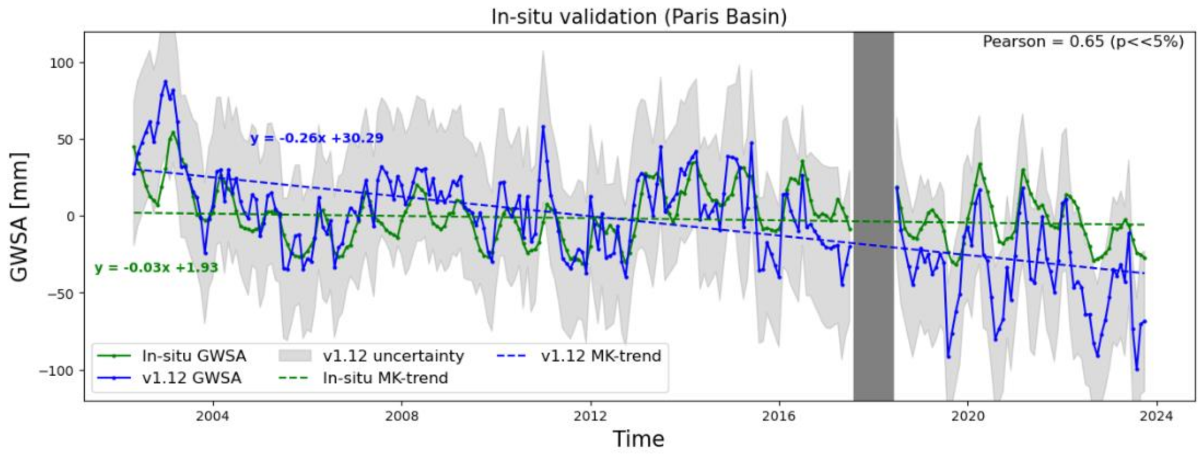

Figure 5: $GWSA_{G3P}$ vs $GWSA_{in\text{-}situ}$ (Paris Basin)

The Pearson coefficient correlation map demonstrates a good correlation across the entire Paris Basin aquifer (median = 0.46). However, an anomaly emerges in the central region where the Pearson coefficient becomes negative (-0.41) *(Figure 6).*

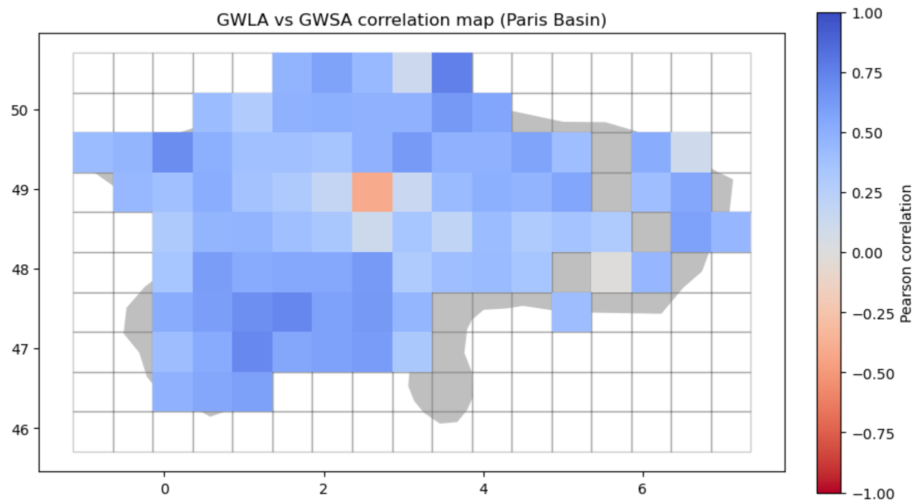

Figure 6: Correlation map (Paris Basin)

A negative Pearson coefficient indicates a negative correlation in the pattern of the in-situ and G3P time series: when one increases, the other decreases. This anomaly might be related to the hydrogeological context of the Paris Basin. The Paris Basin, primarily formed during the Mesozoic era, rests on an older Hercynian basement, and features a succession of sedimentary layers from the Triassic, Jurassic, and Cretaceous. Its particularity is its elliptical structure resulting from the geotectonic processes typical of intracratonic sedimentary basins. These processes caused sediment layers to stack concentrically around a slightly depressed center, where deposits reach several kilometers in depth. Key aquifers, such as those in the Albien's, are crucial for potable water supply *(Contoux et al., 2013) (Figure 7).*

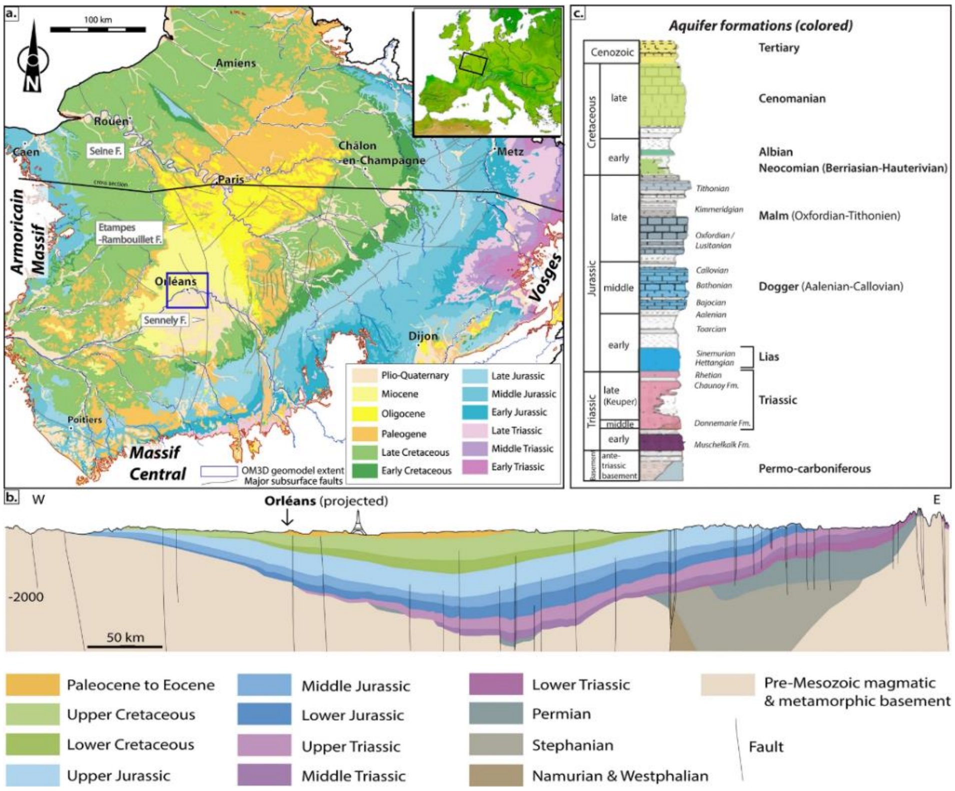

Figure 7: Geomorphological context of Paris Basin

Due to the geological context, the aquifers are often confined beneath the Tertiary formation in the center of the basin. Storage coefficients can be several orders of magnitude smaller in confined aquifers than those in unconfined aquifers. Consequently, groundwater level variations in confined aquifers would result in much smaller groundwater storage changes than similar level variations in unconfined aquifers. Calculating the average GWLA with potential presence of wells in confined aquifer layers could therefore introduce bias and uncertainties in groundwater storage estimates, leading to inconsistency between G3P data and in-situ observations. On the other hand, the filtering process of G3P could result in loss of information in local scales while preserving large-scale patterns. As a result, localized heterogeneities reflected by in-situ data may not be well captured by G3P data, potentially explaining the spatial variability of correlation coefficients in Paris Basin as shown in Figure 6.

2 *Basin and Range*

The Basin and Range aquifer is located in the west of United States, covering an area of approximately 635.000 km^2, extending over most of Nevada, and in parts of south California, west of Utah, south Arizona, southwest New Mexico, and south Oregon and Idaho. The aquifers are formed of volcanic and carbonate rocks and unconsolidated to consolidated basin-fill deposits. Its thickness can reach up to around 500 m. The area is the most arid in the US, especially in the south where evapotranspiration is higher than the total amount of rainfall in a year. The primary uses of groundwater from the Basin and Range aquifer include intensive irrigation (*Planert et al., 1995; USGS.gov*). In-situ data include 53 boreholes, compared to 25 in the 2022 validation phase (+112%) *(Figure 8).*

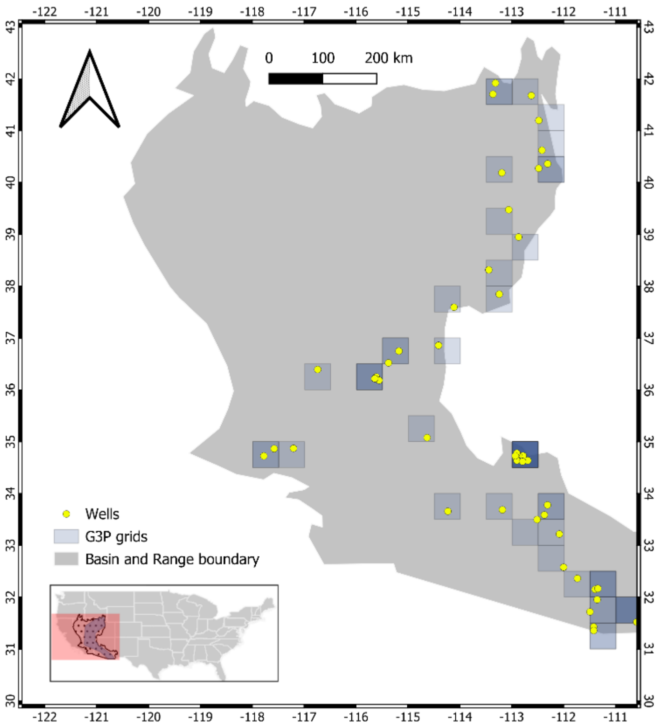

Figure 8: Basin and Range. Blue color indicates the wells density.

According to the results, the Pearson correlation coefficient of normalized data is low (0.21), indicating an overall weak correlation with in-situ data *(Figure 9)*. Moreover, the Pearson correlation coefficient of GWSA data is lower (0.14), presenting an even lower linear positive correlation with in-situ data with the addition of Sy values *(Figure 10)*. In both, in-situ and G3P data present a strong signal dephasing between them. Nevertheless, the MK trends show similar slopes: both decreases at around -0.4 mm per year for the period evaluated; implying that the long-term trend is well captured by GRACE satellites *(Figure 10)*. Both are statistically significant.

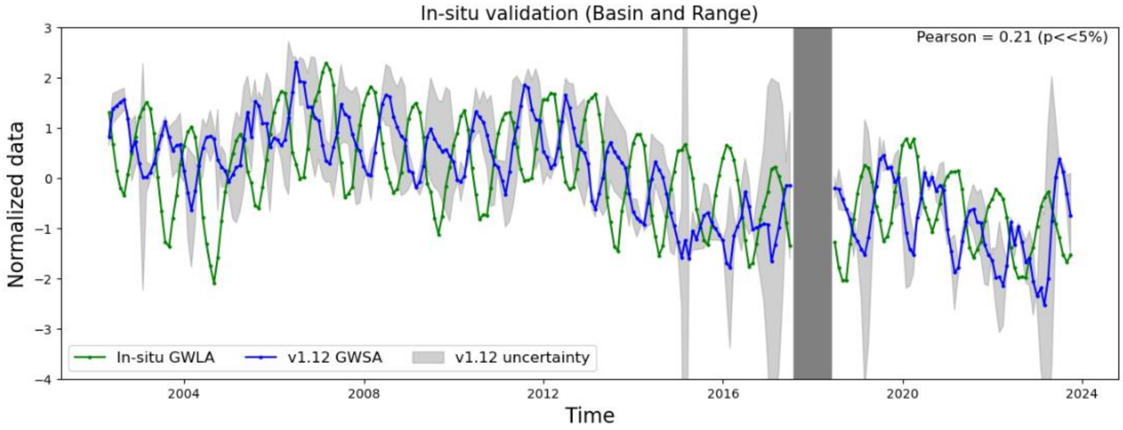

Figure 9: $GWSA_{G3P}$ vs $GWLA_{in\text{-}situ}$ (Basin and Range)

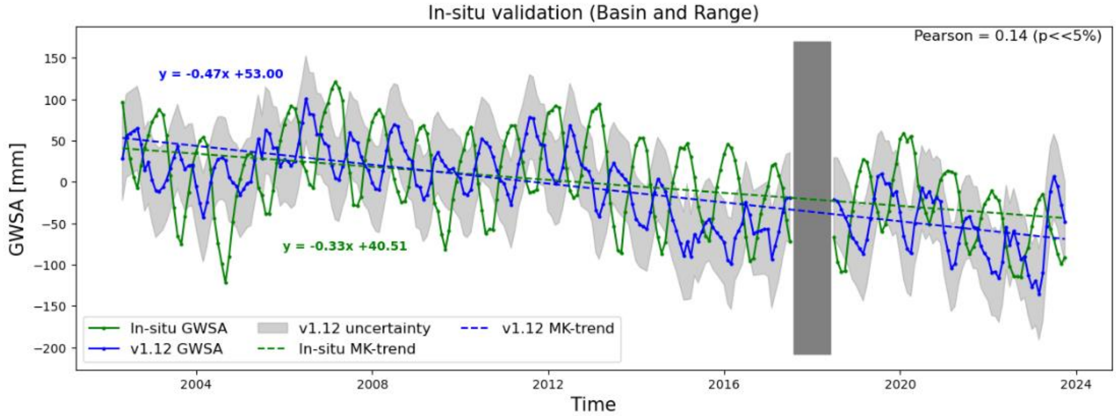

Figure 10: $GWSA_{G3P}$ vs $GWSA_{in\text{-}situ}$ (Basin and Range)

The Pearson coefficient correlation map demonstrates a relatively strong correlation across the entire Basin and Range aquifer (median = 0.49). Higher correlations are found in the southern part of the aquifer (max = 0.87), while lower correlations are observed in the northern part (0.18) *(Figure 11)*.

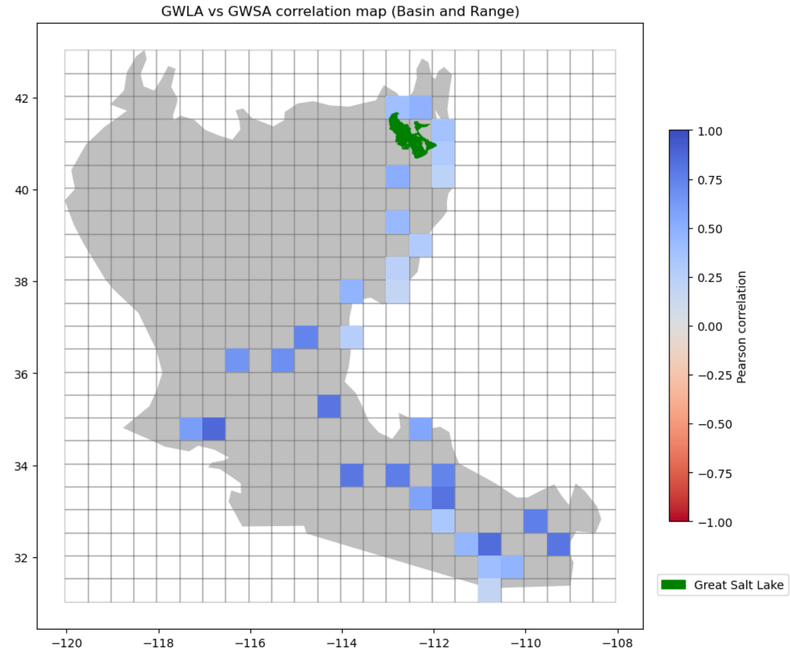

Figure 11: Correlation map (Basin and Range)

These observations could probably be explained by the stark climate contrast in the region. From the Cascade Mountains and the Sierra Nevada in northern California, where precipitation is abundant, to the Great Basin in Nevada and the deserts of southern California, which have some of the most arid environments in the United States, it is clear that few regions in the USA exhibit such a diversity of topography and climate (*USGS, Ground Water Atlas of the United States, 2000) (Figure 12).*

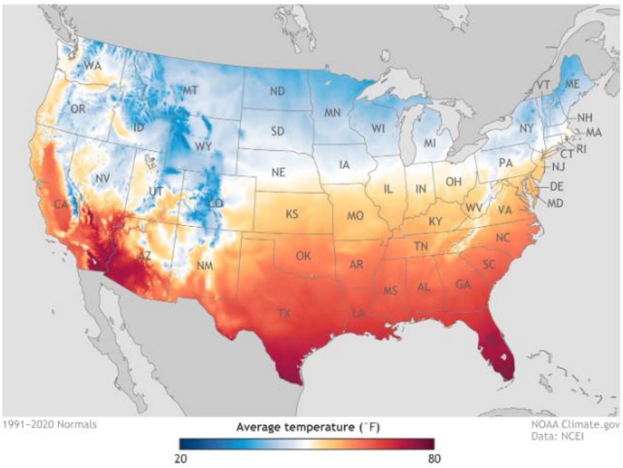

Figure 12: Annual average temperature in Basin and Range region (Source: NOAA)

As noted earlier, $GWSA_{G3P}$ is calculated by subtracting all Water Storage compartments (WSCs) from Total Water Storage (TWS), as outlined in *Equation 2*, and its uncertainty is calculated by propagating the uncertainties of the individual WSCs *(Equation 1).* However, this strong climate contrast between the northern and southern parts of the Basin and Range aquifer might imply different implications of WSCs variations and uncertainties in both regions. As the northern region is wetter and present a more pronounced topography that the southern one, resulting $GWSA_{G3P}$ calculated in that region could be more influenced by variations and uncertainties in Snow Water Equivalent (SWE), Surface Water Storage (SWS) or Soil Moisture (SM) water compartments.

Weaker Pearson correlations to the north compared to the south may also be influenced by the presence of the Great Salt Lake. This eighth-largest terminal lake in the world has fluctuating water levels due to imbalances between inputs from rivers and precipitation and outputs through only evaporation. The lake's shallow depth means even small changes in water level cause significant variations in its shoreline, contributing to high coverage variability throughout the year *(USGS, Great Salt Lake, Utah).* This seasonal surface variability could have a notable impact on GRACE data, especially due to the leakage effect of G3P product. As presented earlier, the leakage effect implies that the actual net WSC for a grid cell of interest depends on the signal characteristics in its surroundings. Therefore, Surface Water Storage (SWS) compartment near the lake could have a very high seasonal variability, thereby affecting $GWSA_{G3P}$ calculation.

Then those explanations cannot confirm the observable results and additional studies must be carried out, especially by using other WSCs time series from our specific grid and around the lake in order to measure the leakage effect and to estimate its real impact on $GWSA_{G3P}$ calculations.

3 *Ogallala (High Plains)*

The Ogallala Aquifer is one of the largest aquifers in the United States with an area of approximately 450.000 km^2, extending over eight states: South Dakota, Nebraska, Wyoming, Colorado, Kansas, Oklahoma, New Mexico, and Texas. This is a productive aquifer and consists of unconsolidated clay, silt, and sand, with some gravel and caliche. The aquifer is mostly unconfined with saturated thickness from 0 to 300 m (mean 60 m) and water-table depth from 0 to 150 m (mean 30 m) *(Dennehy, 2000*). The climate in northwest part of the aquifer is considered as semiarid, since the average annual precipitation is usually less than 500 mm per year and mean annual evaporation from 1,500 to 2,700 mm *(Miller et al., 1997, Strassberg and al, 2007)*. It serves as the main source of water supply (i.e., agricultural, public), sustaining the economic development for the entire region for more than 80 years. It accounts for 30% of total crop and animal production and 90% of all the pumped water is used for irrigation in agriculture *(Cano et al., 2018; Deines et al., 202*0). Extensive irrigation (30% of GW used for irrigation in the US *(Dennehy, 2000)* decreased GW levels by an average of 4 m over the entire aquifer since predevelopment (1950) and by up to 68 m in some regions *(McGuire, 2004)*. In-situ data include 40 boreholes, compared to 21 in the 2022 evaluation phase (+90%) *(Figure 13)*.

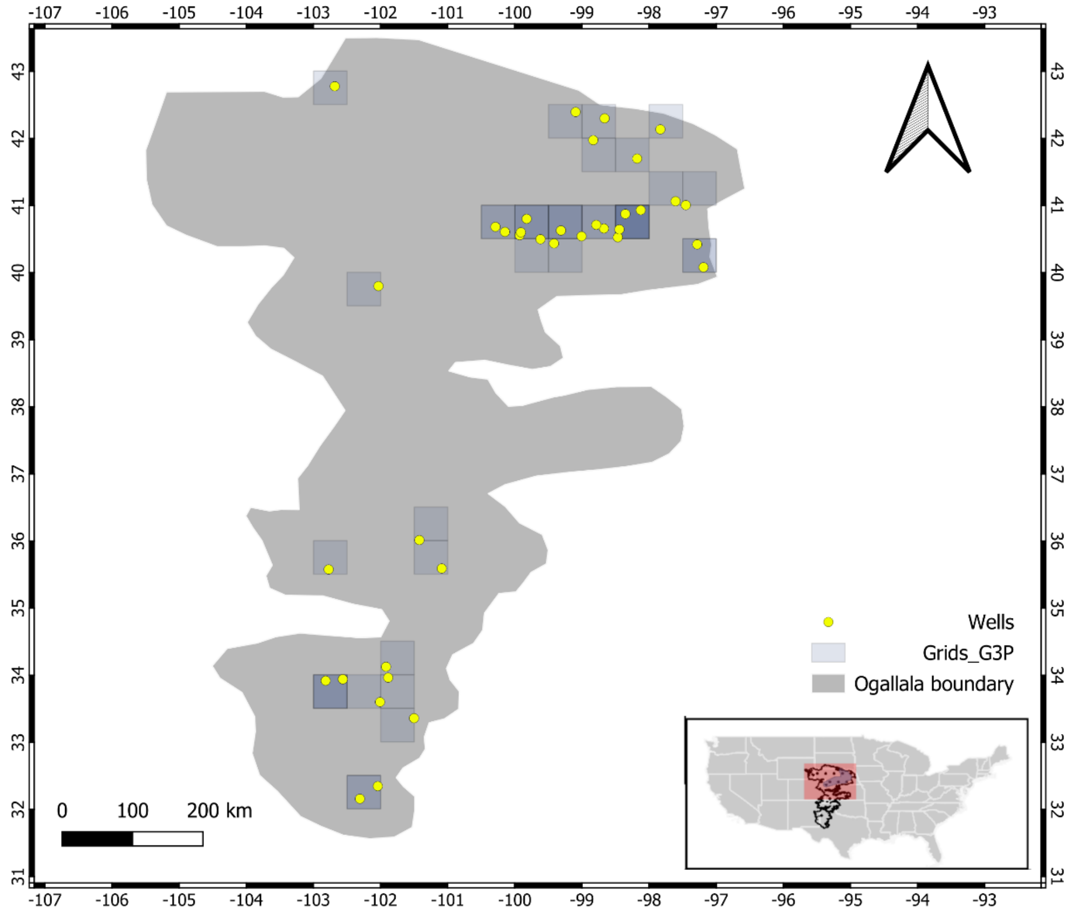

Figure 13: Ogallala (High Plains) aquifer

According to the results, the Pearson correlation coefficient of normalized data is high (0.82), indicating an overall good linear positive correlation with in-situ data *(Figure 13)*. Moreover, the Pearson correlation coefficient of GWSA data is almost identical (0.82), presenting a good linear positive correlation too with in-situ data due to the addition of Sy values *(Figure 16)*. For GWSA data, G3P is lower than in-situ before 2012, while higher afterward. Also, G3P shows lower seasonal amplitude variations than in-situ over the whole period, especially during 2012 drop that appears much less pronounced than that of in-situ. Indeed, in-situ data usually shows lower minimum local values. Finally, both MK trends are decreasing, even if G3P shows a lower slope. Both are statistically significant.

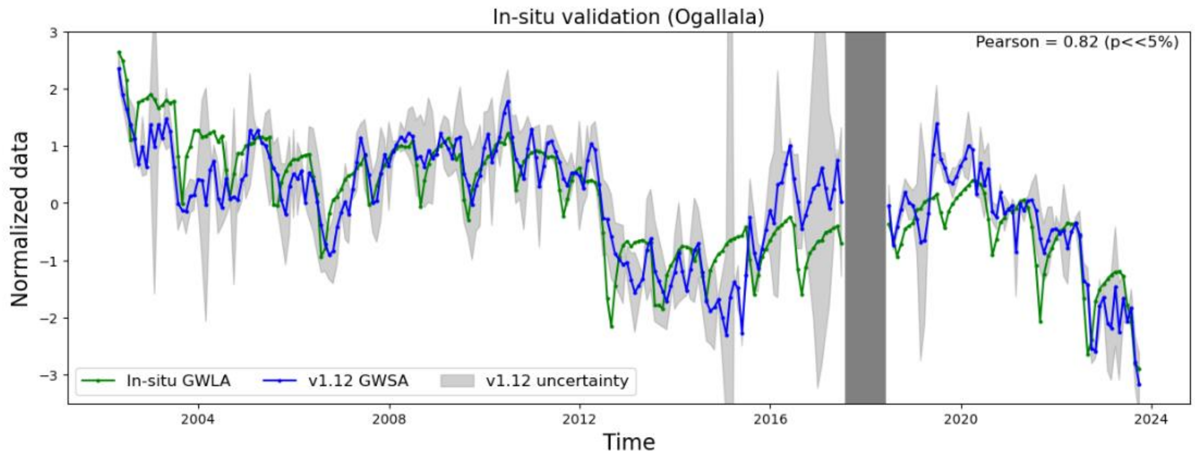

Figure 13: $GWSA_{G3P}$ vs $GWLA_{in\text{-}situ}$ (Ogallala)

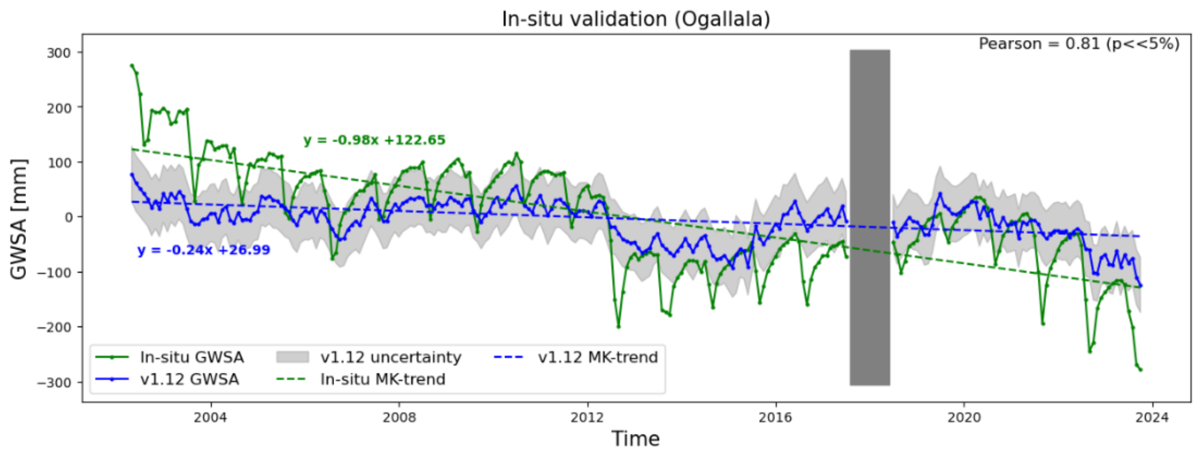

Figure 14: GWSAG3P vs GWSAin-situ (Ogallala)

These observations might be related to the overestimation of Specific Yield (Sy) values used during the calculation of $GWSA_{in\text{-}situ}$. It's likely that employing lower Sy values would have resulted in a better alignment with GWSA derived from G3P data. This overestimation of Sy values may stem from the methodology, which relies on a surface-geological-layer approach. In this approach, Sy values tend to be higher compared to those at greater depths. If some of the wells used to assess this aquifer are located in confined layers, it becomes imperative to consider this factor when determining Sy values. Such adjustments can lead to more accurate estimations of groundwater storage anomalies, allowing to better catch the long-term trend and improve the overall reliability of the data analysis. In-situ data showing higher amplitudes variations with more pronounced minimum local values can be explained by human activities and withdrawals in the region, impacting directly the water levels in some wells.

The Pearson coefficient correlation map demonstrates a good correlation across the entire Ogallala aquifer (mean = 0.72). Higher spatial correlations are found in the eastern part of the aquifer (max = 0.95), while lower correlations are observed in the western part (min = 0.03) *(Figure 15).*

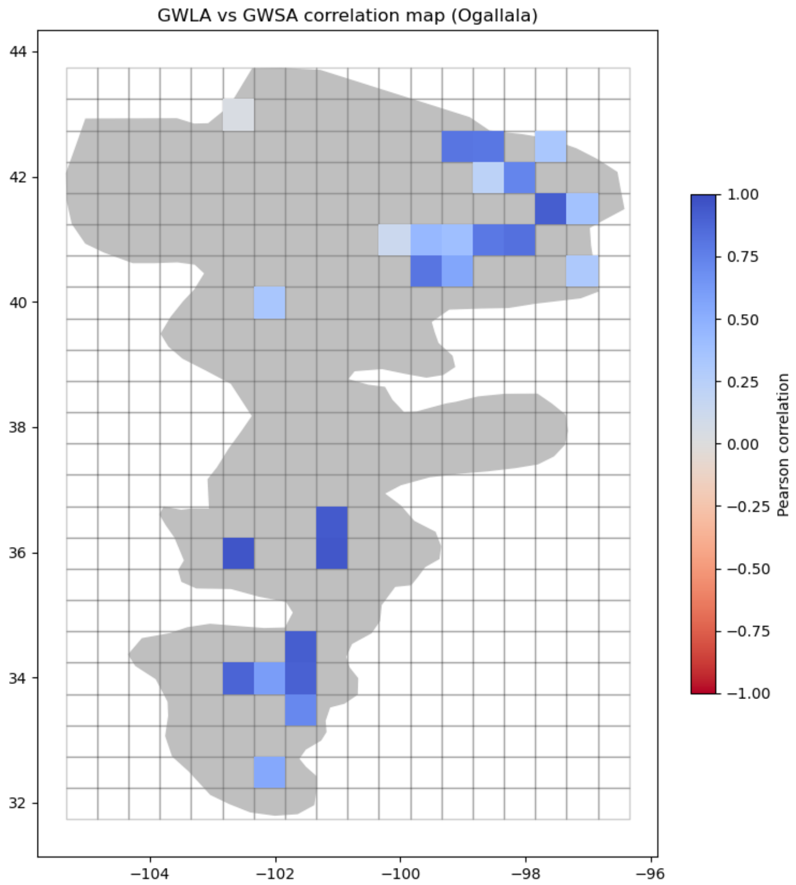

Figure 15: Correlation map (Ogallala)

Although the overall good correlation can be explained by a less contrasted climate in the High Plains region than the one in Basin and Range, the lower correlation in the west hasn't clearly found an explanation yet. Nevertheless, it could be related to some heterogeneities due to human withdrawals, easily captured by in-situ data but not by G3P. Further studies have to be carried out in this part of the aquifer.

CONCLUSION

This study assesses the reliability of GRACE data for tracking groundwater storage changes in the Paris Basin (France), Basin and Range (USA), and Ogallala (High Plains, USA) aquifers systems from 2002 to 2023.

To validate GRACE data, the approach was based on using in-situ measurements obtained from national geological organizations. These measurements consist of monthly water depth data from wells across the study area, which were converted into groundwater level anomalies (GWLA) and then to groundwater storage anomalies (GWSA) using specific yield (Sy) values provided by the GFZ. The GRACE data, with a monthly resolution and 0.5° spatial resolution, were directly obtained from the GFZ. These data underwent a subtraction process to isolate GWSA by removing variations of other water storage compartments (WSCs) from Total Water Storage (TWS).

To evaluate the performance of GRACE, we used the Pearson correlation coefficient to analyze linear correlation and observed differences in seasonal amplitudes. Additionally, we employed the Mann-Kendall (MK) statistical test to compare trends over the evaluated period.

The major improvements over the 2022 evaluation lie in the broadening of criteria for utilizing in-situ data, access to Sy data to compute in-situ GWSA, the grid/pixel approach, the analysis of Mann-Kendall trends, and finally, the latest version of GRACE (v1.12).

- The Paris Basin exhibits a positive linear correlation (0.65). This correlation is generally strong across the aquifer, except for a negative correlation in the central region, possibly due to some wells located in confined aquifers introducing bias and uncertainties in GWSA estimates, or due to filtering process leading to lower representativity of local heterogeneities in G3P data.
- The Basin and Range aquifer exhibits the lowest correlation (0.21) and an out-of-phase signal, more pronounced in the north. This discrepancy may be due to a strong climate contrast, considering different WSCs influences in both regions, affecting the overall correlation. Additionally, the potential presence of the Great Salt Lake in the north of the aquifer, combined with the leakage effect of G3P, might also impact GRACE data.
- The Ogallala aquifer shows the best correlation (0.82). The difference between the signals may be related to an overestimation of Sy values used for $GWSA_{in\text{-}situ}$ calculation due to the surface-based methodology used. If wells in confined layers are included, adjusting Sy values is crucial. However, this strong correlation is consistent throughout the aquifer, except for one unexplained cell in the west, warranting further investigation.

In essence, within the framework of this study, we can conclude that our in-situ evaluation of the GRACE data is rather good. In all scenarios, G3P mostly managed to catch the long-term trend and validate the patterns of the time series. Correlation maps seem to be an interesting and promising approach, having allowed us to push our thinking and our analysis towards a more local approach.

The major limitations lie in:

- The consideration of confined aquifers: confined aquifers present a double problem: the first is that, due to their hydrogeological characteristics, variations in GWLA don't match variations in GWSA. The second is due to the methodology, which tends to over-estimate the in-situ GWSA values calculated due to Sy values based on the surface lithological layers. In both cases, in-situ and GRACE signals present different amplitudes, despite positive linear correlations, then impacting the overall trend.
- Contrasted climate: in contrasted climate regions, all WSCs could impact $GWSA_{G3P}$ differently according to the area, leading to different uncertainty evaluations.
- The pixel-based approach: this approach is limited in terms of representativity, due to the filtering process and leakage effect. An analysis at 0.5 degree cannot account for the spatial heterogeneity in the region (such as withdrawals) and comes with certain limitations. While $GWSA_{G3P}$ are derived from remote observations and hence offer global coverage, the $GWLA_{in\text{-}situ}$ can only be computed where groundwater observations wells are present. Multiple aquifers can also be present within one pixel, therefore understanding the overall groundwater dynamics at this grade scale can be misrepresentative if it is solely based on a limited number of boreholes.

Given the uncertainty from both in-situ and G3P sides, it is encouraged to evaluate G3P with more data to obtain more reliable results, if possible. Avenues for improvement can emerge, such as by reducing the restrictions in order to expand the in-situ data and increasing the surface area evaluated, selecting only wells located in unconfined aquifers, or refining the evaluation of the Sy coefficients (via pumping tests). Comparing in-situ data with various WSCs, rather than solely $GWSA_{G3P}$, could help measure the leakage effect and then identify factors affecting the results, especially in contrasted climates. It is also advisable to consider the uncertainty of each compartment and revise those contributing most to the discrepancies.

We must keep in mind that in-situ and GRACE data represent two different things, which also leads us to question the very purpose of this validation. For example, the overall trend of an aquifer system is of no use to a small community water agency, whereas it is essential for governments or international institutions to take necessary environmental guidelines. On the other hand, the pixel-based approach could be more interesting for a local area but has to be carefully interpreted.

Finally, although the idea of global monitoring of transboundary aquifers via an efficient satellite network on a global scale is rather attractive, our work justifies the importance of investing today in piezometric networks on a global scale, in order to supply hydrogeological data banks that are still too light.

REFERENCES

– GFZ Potsdam, IGRAC, – EODC GmbH, Güntner, Haas, Sharifi, Ruz, Vargas, & Kidd. (2022). *G3P Product Report*. Horizon2020.

Contoux, C., Violette, S., Vivona, R., Goblet, P., & Patriarche, D. (2013). How basin model results enable the study of multi-layer aquifer response to pumping: the Paris Basin, France. *Hydrogeology Journal*, *21*(3), 545-557. https://doi.org/10.1007/s10040-013-0955-6

Deines, J. M., Kendall, A. D., Butler, J. J., Basso, B., & Hyndman, D. W. (2021). Combining Remote Sensing and Crop Models to Assess the Sustainability of Stakeholder-Driven Groundwater Management in the US High Plains Aquifer. *Water Resources Research*, *57*(3). https://doi.org/10.1029/2020wr027756

Dennehy, K. F. (2000). High Plains regional ground-water study. *Fact Sheet*. https://doi.org/10.3133/fs09100

G3P. (2022). Work Package 3 QUANTIFICATION OF STORAGE COMPARTMENTS. Dans *G3P*.

GRAVIS - GFZ - Gravity Information Service. (s. d.). https://gravis.gfz-potsdam.de/gws

Groat, C. (2000). *Ground Water Atlas of the United States*. https://doi.org/10.3133/ha730

International Groundwater Resources Assessment Centre. (s. d.). *IGRAC, Your Partner in Groundwater | International Groundwater Resources Assessment Centre*. https://www.un-igrac.org/

Mas, P., Calcagno, P., Caritg-Monnot, S., Beccaletto, L., Capar, L., & Hamm, V. (2022). A 3D geomodel of the deep aquifers in the Orléans area of the southern Paris Basin (France). *Scientific Data*, *9*(1). https://doi.org/10.1038/s41597-022-01876-4

McGuire, V. L. (2004). Water-level changes in the high plains aquifer, predevelopment to 2003 and 2002 to 2003. *Fact Sheet*. https://doi.org/10.3133/fs20043097

Mitchell-McCallister, D., Cano, A., & West, C. (2020). Meta-analysis of crop water use efficiency by irrigation system in the Texas High Plains. *Irrigation Science*, *38*(5-6), 535-546. https://doi.org/10.1007/s00271-020-00696-x

Rateb, A., Scanlon, B. R., Pool, D. R., Sun, A., Zhang, Z., Chen, J., Clark, B., Faunt, C. C., Haugh, C. J., Hill, M., Hobza, C., McGuire, V. L., Reitz, M., Schmied, H. M., Sutanudjaja, E. H., Swenson, S., Wiese, D., Xia, Y., & Zell, W. (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

Rodell, M., & Reager, J. T. (2023). Water cycle science enabled by the GRACE and GRACE-FO satellite missions. *Nature Water*, *1*(1), 47-59. https://doi.org/10.1038/s44221-022-00005-0

Strassberg, G., Scanlon, B. R., & Rodell, M. (2007a). Comparison of seasonal terrestrial water storage variations from GRACE with groundwater-level measurements from the High Plains Aquifer (USA). *Geophysical Research Letters*, *34*(14). https://doi.org/10.1029/2007gl030139

Strassberg, G., Scanlon, B. R., & Rodell, M. (2007b). Comparison of seasonal terrestrial water storage variations from GRACE with groundwater-level measurements from the High Plains Aquifer (USA). *Geophysical Research Letters*, *34*(14). https://doi.org/10.1029/2007gl030139

Tangdamrongsub, N., Han, S. C., Tian, S., Schmied, H. M., Sutanudjaja, E. H., Ran, J., & Feng, W. (2018). Evaluation of Groundwater Storage Variations Estimated from GRACE Data Assimilation and State-of-the-Art Land Surface Models in Australia and the North China Plain. *Remote Sensing*, *10*(3), 483. https://doi.org/10.3390/rs10030483

Tierney, K., Bevc, C., & Kuligowski, E. (2006). Metaphors Matter: Disaster Myths, Media Frames, and Their Consequences in Hurricane Katrina. ˜ *The œ Annals Of The American Academy Of Political And Social Science/˜ The œ Annals*, *604*(1), 57-81. https://doi.org/10.1177/0002716205285589

Urgilez Vinueza, Ruz, Vargas, Contreras, Imran, IGRAC, & FutureWater. (2022). *G3P Evaluation Report*. Horizon2020.

pymannkendall. (2023, 14 janvier). PyPI. https://pypi.org/project/pymannkendall/

New maps of annual average temperature and precipitation from the U.S. Climate Normals. (2021, 11 octobre). NOAA Climate.gov. https://www.climate.gov/news-features/featured-images/new-maps-annual-average-temperature-and-precipitation-us-climate