

LC-SPACE-04-EO-2019-2020 “Copernicus evolution – Research activities in support of cross-cutting applications between Copernicus services”

Action acronym: **G3P**
Full title: Global Gravity-based Groundwater Product
Grant Agreement No: 870353



Global Gravity - based Groundwater Product

Deliverable D4.2: G3P Evaluation Report – Revision 1

Date: 10.03.2023

Authors:

Alexandra Urgilez Vinueza, IGRAC
Claudia Ruz Vargas, IGRAC
Sergio Contreras, FutureWater
Tania Imran, FutureWater

1. Change Record

Name	Author(s)	Date	File name
v1.	Urgilez Vinueza, A.; Ruz Vargas C., Contreras S., Imran T.	19.12.2022	G3P_D4_2_G3P_Evaluation_report
R1	Haas	10.03.2023	G3P_D4_2_G3P_Evaluation_report_R1

2. Table of Contents

1. Change Record	2
2. Table of Contents	3
3. List of Figures	4
4. Overview and summary	7
5. Methodology	8
5.1 In-situ data collection and formatting	8
5.2 G3P time series extraction	8
5.3 In-situ time series calculation	9
5.3.1 In-situ GWSA calculation	9
5.3.2 In-situ GWLA calculation	10
6. Results	10
6.1 Ogallala Aquifer	10
6.2 North China Aquifer System	14
6.3 Floridan Aquifer System	15
6.4 Basin and Range basin-fill Aquifers	17
6.5 Cambrian Ordovician aquifer system	19
6.6 Mississippi River Valley	21
6.7 Northern High Plains	23
6.8 Paris Basin	25
6.9 Murray Darling Basin	27
6.10 Great Artesian Basin	29
6.11 South of outer Himalayas aquifer	31
6.12 Guarani Aquifer System	33
6.13 Maranhão Aquifer System	35
7. Case study in Spain	38
7.1 Methodology	38
7.1.1 Study region	38
7.1.2 In-situ data collection	40
7.1.3 Preprocessing and retrieval of the GroundWater Index Anomaly (GWI)	41
7.1.4 Extreme conditions	42
7.1.5 Pearson correlations coefficients	43
7.1.6 Trends	44
7.1.7 Representation of extreme events	45
8. Discussion	46
9. Summary	48

10. Annex 1. Comparison of G3P GWSA_LA_hyd and GWSA_LA_grav to in-situ GWSA or GWLA 49

10.1 Ogallala aquifer 49

10.2 North China Aquifer system 50

10.3 Floridan Aquifer System 51

10.4 Basin and Range Aquifers 52

10.5 Cambrian-Ordovician Aquifer System 52

10.6 Mississippi River Valley 53

10.7 Northern High Plains 54

10.8 Paris Basin 54

10.9 Murray Darling Basin 55

10.10 Great Artesian Basin 56

10.11 South of outer Himalayas aquifer 56

10.12 Guarani Aquifer System 57

10.13 Maranhão Aquifer System 58

11. References 59

3. List of Figures

Figure 1. Location of the selected aquifers for validation 7

Figure 2. Thiessen polygons method example in the north-Ogallala aquifer system in the US. Each point in yellow represents a monitoring station. 9

Figure 3. North Ogallala aquifer, evaluated area, and in-situ boreholes for the period 2002-2016 11

Figure 4. North Ogallala aquifer, evaluated area, and in-situ boreholes for the period 2009-2016 12

Figure 5. In-situ and G3P derived GWSA for the north part of the Ogallala aquifer, using G3P v1.5 for the period 2002-20016. 13

Figure 6. In-situ and G3P derived GWSA for the north part of the Ogallala aquifer using G3P v1.5 for the period 2009-2016. 13

Figure 7. Spatial distribution of GWSA_G3P for the period 2002-2016 (a) and the period 2009-2016 (b) on August 1st, 2004. The area in (a) is contained in (b). 14

Figure 8. North China Aquifer System, evaluated area, and in-situ boreholes for the period 2005-2012 14

Figure 9. In-situ and G3P derived GWSA for the North China Aquifer System using G3P v1.5 for the period 2005-2012. 15

Figure 10. Floridan Aquifer system, in-situ boreholes, and evaluated area for the period 2002-2016 16

Figure 11. Area-average G3P GWSA for the blue region (Figure 9) within the Floridan Aquifer system and its uncertainty 16

Figure 12. Standardized in-situ GWLA and G3P GWSA in the Floridan Aquifer system for the period 2002-2016 17

Figure 13. Basin and Range basin-fill aquifers, in-situ boreholes and evaluated area for the period 2012-2016.....	18
Figure 14. Area average G3P GWSA of Basin and Range basin-fill aquifers for the period 2012-2016.....	18
Figure 15. Standardized in-situ GWLA and G3P GWSA in the Basin and Range basin-fill aquifers for the period 2012-2016.....	19
Figure 16. Cambrian-Ordovician aquifer system, in-situ boreholes and evaluated area for the period 2012-2016.....	20
Figure 17. Area average G3P GWSA of the Cambrian-Ordovician aquifer system for the period 2012-2016.....	20
Figure 18. Standardized in-situ GWLA and G3P GWSA in the Cambrian-Ordovician aquifer system for the period 2012-2016.....	21
Figure 19. Mississippi River Valley, in-situ boreholes and evaluated area for the period 2008-2016.....	22
Figure 20. Area average G3P GWSA of the Mississippi River Valley for the period 2012-2016.....	22
Figure 21. Standardized in-situ GWLA and G3P GWSA in the Mississippi River Valley for the period 2008-2016.....	23
Figure 22. Northern High Plains aquifer system, in-situ boreholes and evaluated area for the period 2008-2013.....	24
Figure 23. Area average G3P GWSA of the Northern High Plains aquifer system for the period 2008-2013.....	24
Figure 24. Standardized in-situ GWLA and G3P GWSA in the Northern High Plains aquifer system for the period 2008-2013.....	25
Figure 25. Paris Basin, in-situ boreholes and evaluated area for the period 2002-2016.....	26
Figure 26. Area average G3P GWSA of the Paris Basin for the period 2002-2016.....	26
Figure 27. Standardized in-situ GWLA and G3P GWSA in the Paris Basin for the period 2002-2016.....	27
Figure 28. Murray Darling Basin, in-situ boreholes and evaluated area for the period 2002-2016.....	28
Figure 29. Area average G3P GWSA of the Murray Darling Basin for the period 2002-2016.....	28
Figure 30. Standardized in-situ GWLA and G3P GWSA in the Murray Darling Basin for the period 2002-2016.....	29
Figure 31. Great Artesian Basin, in-situ boreholes and evaluated area for the period 2010-2016.....	30
Figure 32. Area average G3P GWSA of the Great Artesian Basin for the period 2010-2016.....	30
Figure 33. Standardized in-situ GWLA and G3P GWSA in the Great Artesian Basin for the period 2010-2016.....	31
Figure 34. South of outer Himalayas aquifer, in-situ boreholes and evaluated area for the period 2002-2007.....	32
Figure 35. Area average G3P GWSA of the South of outer Himalayas aquifer for the period 2002-2007.....	32
Figure 36. Standardized in-situ GWLA and G3P GWSA in the South of outer Himalayas aquifer for the period 2002-2007.....	33
Figure 37. Guarani Aquifer System, in-situ boreholes and evaluated area for the period 2012-2016.....	34
Figure 38. Area average G3P GWSA of the Guarani Aquifer System for the period 2012-2016.....	34

Figure 39. Standardized in-situ GWLA and G3P GWSA in the Guarani Aquifer System for the period 2012-2016.....	35
Figure 40. Maranhão Aquifer System, in-situ boreholes and evaluated area for the period 2014-2016	36
Figure 41. Area average G3P GWSA of the Maranhão Aquifer System for the period 2014-2016	36
Figure 42. Standardized in-situ GWLA and G3P GWSA in the Maranhão Aquifer System for the period 2014-2016.....	37
Figure 43. Distribution of Groundwater Bodies (GWB) in continental Spain. GWB are classified according to its dominant aquifer typology and groundwater productivity (HP: high productive, MP: Medium productive, LP: low productivity). Numbers correspond to the ID of each River Basin District.....	39
Figure 44. Percentage contribution of each GWB category in continental Spain	39
Figure 45. Division of the study area by GRACE quadrants of 1 deg and 0.5 deg. River Basin districts (from N-S): 14: Galicia-Costa, 18: Cantábrico Occidental, 17: Cantábrico Oriental, 11: Miño-Sil, 21: Duero, 91: Duero, 101: Cuencas Internas de Cataluña, 31: Guadiana, 81: Júcar, 40: Guadiana, 64: Tinto, Odiel y Piedras, 51: Guadalquivir, 71: Segura, 63: Guadalete y Barbate, 61: Cuenca Mediterránea Andaluza.....	40
Figure 46. Distribution of qualified boreholes classified according to their depths of penetration.....	41
Figure 47. Number of qualified monitoring wells at the 0.5 deg. pixel level. A qualified well is that one with at least 150 monthly observations in the 2002-2019 period.....	41
Figure 48. Right: Boxplots of pixel-by-pixel Pearson correlations coefficients between G3P-GWSA and GWI values retrieved using simple-arithmetic (mGWI) and weighted (wGWI) averaging methods. Left: Boxplot with Pearson correlation differences between mGWI and wGWI	43
Figure 49. Pearson cross correlations between mGWI and G3P-GWSA at 0.5 deg. resolution	43
Figure 50. Temporal variation of GWSA against SPI12 at Pixel ID: DF.....	44
Figure 51. Variation in SPI12, GRACE-GWSA and mGWI between reference and dry period.	45
Figure 52. Variation in SPI12, GRACE-GWSA and mGWI between reference and wet period	45

4. Overview and summary

Under the WP4: “Groundwater product development, evaluation and service preparation of the G3P project” of the G3P project, the G3P Groundwater Storage Anomaly (GWSA) products are evaluated against in-situ groundwater observations. For this purpose, 23 aquifers were selected in 2020 to proceed with the evaluation at a large scale. For this large-scale evaluation, **version v1.5 of G3P** was used. Additionally, a case study in Spain was set up to evaluate the G3P at the pixel scale. For this pixel-based evaluation, **version v1.3 of G3P** was used.

Groundwater level datasets as well as hydrogeological data from the 23 selected aquifers were collected from the national institutions or geological services of the countries to which each aquifer belongs to and complemented by data from the Global Groundwater Monitoring Network (GGMN) portal.

Out of the twenty-three aquifers, thirteen had enough in-situ groundwater observations and were evaluated. They are located in America, Europe, Asia, and Australia. Four aquifers had in-situ data, but the dataset was not enough to perform the evaluation. Six aquifers did not have an available dataset. See Figure 1.

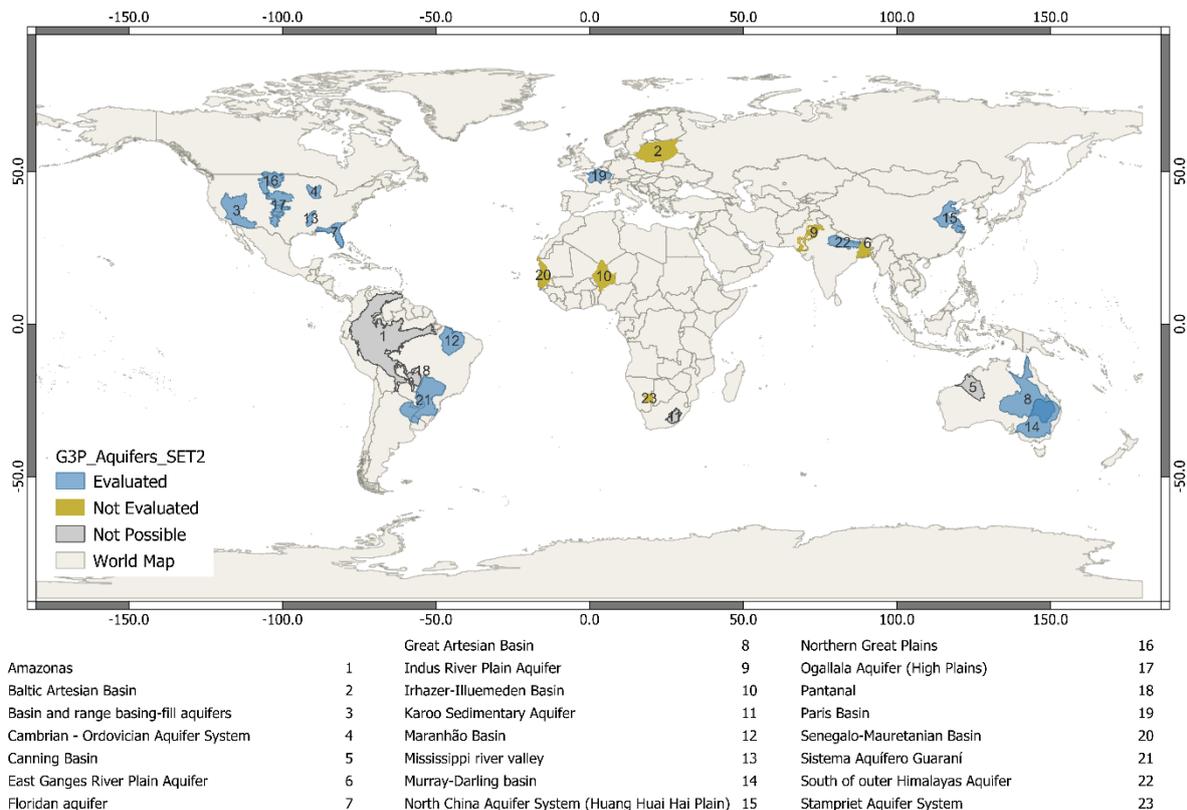


Figure 1. Location of the selected aquifers for validation

G3P GWSA is derived from the GRACE AND GRACE-FO satellites. It is calculated as the difference between the total water storage (TWS) and all superficial water compartments (i.e., snow, glaciers, lakes, and soil moisture). The result of this difference is interpreted as GWSA. In addition, a leakage approximation is performed to GWSA, resulting in two more products, namely: GWSA_LA_hyd and GWSA_LA_grav.

In-situ groundwater data is compared to the three products of G3P using the Pearson correlation. When hydrogeological information such as specific yield is available, in-situ GWSA is calculated and compared to G3P GWSA. On the contrary, when groundwater levels are the

only groundwater information available, groundwater level anomaly (GWLA) is calculated, and its signal is compared to G3P GWSA's signal.

For most of the aquifers, in-situ GWLA is compared to G3P GWSA, while few of the aquifers had specific yield information to calculate in-situ GWSA. Results show that most of the aquifers show a high correlation between in-situ groundwater data and G3P GWSA such as the Ogallala aquifer, Floridan aquifer, Paris Basin, South of Outer Himalayas aquifer; while other aquifers such as Murray Darling aquifer, the Great Artesian Basin, and the Basin and Range basin-fill aquifers show a low correlation. Low correlations are usually found in arid or semi-arid areas and in confined conditions. It is important to also consider that some of the validated aquifers were done using a limited number of in-situ observations.

At the pixel scale, the evaluation is done using in-situ groundwater level datasets from continental Spain. An in-situ Groundwater Index Anomaly (GWI) is calculated to be compared to G3P. In addition, extreme conditions are also analyzed and compared to G3P as well as to the SPI drought index. Results show that, at the pixel scale, correlations between G3P and GWI range from -0.4 to 0.6. The lowest correlations are found towards the south of the country, where there are drylands regions.

5. Methodology

5.1 In-situ data collection and formatting

The main source of data used for this project was found in national official sources to which the aquifers belong to, such as geological surveys, or national networks (e.g., the USGS in the United States, RIMAS in Brazil, GIN in Canada) (IGRAC, 2020). This data was complemented with data from the Global Groundwater Monitoring Network (GGMN) portal, available at <https://ggis.un-igrac.org/view/ggmn#/>. In addition, there were a couple of personal correspondence with researchers or national institutions who shared groundwater data, in the case of the North China aquifer and the Karoo sedimentary aquifer in South Africa.

The data that was collected, was run through a filtering process. First, the temporal resolution of the groundwater level data was adjusted to monthly, and in some cases, seasonal time scale. Second, the time series were evaluated to check for outliers. If these are caused by human errors in measurements, they would be discarded. On the other hand, if the cause for a deviant groundwater level value could not be explained and might be due to natural variations in groundwater, the value would be kept. In both cases, a critical evaluation of the situation needs to take place to make a final decision.

Finally, the data is arranged in tables from where the groundwater levels would be taken, and GWSA would be calculated when hydrogeological data is available.

Hydrogeological data, specifically the specific yield, was available for the Ogallala aquifer, and distributed by the USGS (McGuire et al., 2012). In the case of the North China aquifer system, the specific yield was included in the provided in-situ GWSA dataset.

5.2 G3P time series extraction

The G3P v1.5 from WP3 is provided as input data for the evaluation of G3P. Once the product is acquired, the area-average time series of GWSA of the area is extracted. There are four outputs from the product:

- GWSA: Groundwater storage anomalies.
- GWSA_LA_grav: Groundwater storage anomalies with GRACE-based leakage approximation.

- GWSA_LA_hyd: Groundwater storage anomalies with data-based leakage approximation.
- Uncertainty: Uncertainty related to GWSA, and it is calculated based on the error propagation formula from the associated uncertainties for terrestrial water storage from all the compartments. No uncertainty information is given for LA.

A detailed description of the GWSA products, associated uncertainty, and related products can be found in WP3 deliverables. All three products are used for the evaluation of G3P against in-situ groundwater data. Depending different circumstances (e.g., close/far from coast, close/far from glaciers, tropical areas), each G3P product may represent better the state of the aquifer.

5.3 In-situ time series calculation

With the purpose of evaluating the G3P against in-situ information, in-situ groundwater data is procured and filtered, as explained in Section 5.1. GWSA can be calculated from groundwater level data and hydrogeological data for the entire area and used for the evaluation. However, in the absence of hydrogeological data, GWLA were calculated instead. These were standardized and compared to standardized G3P GWSA. The purpose of this approach is to determine if the signals of the two time series are similar.

5.3.1 In-situ GWSA calculation

In-situ groundwater level time series and hydrogeological information (specific storage or specific yield), when available, are used to calculate in-situ area-averaged GWSA ($GWSA_{in-situ}$) in a specific aquifer. For this purpose, the Thiessen polygon method was used (Figure 2), where per each monitoring station i , there is an associated area $area_i$ and storage S_i that corresponds to the area $area_i$.

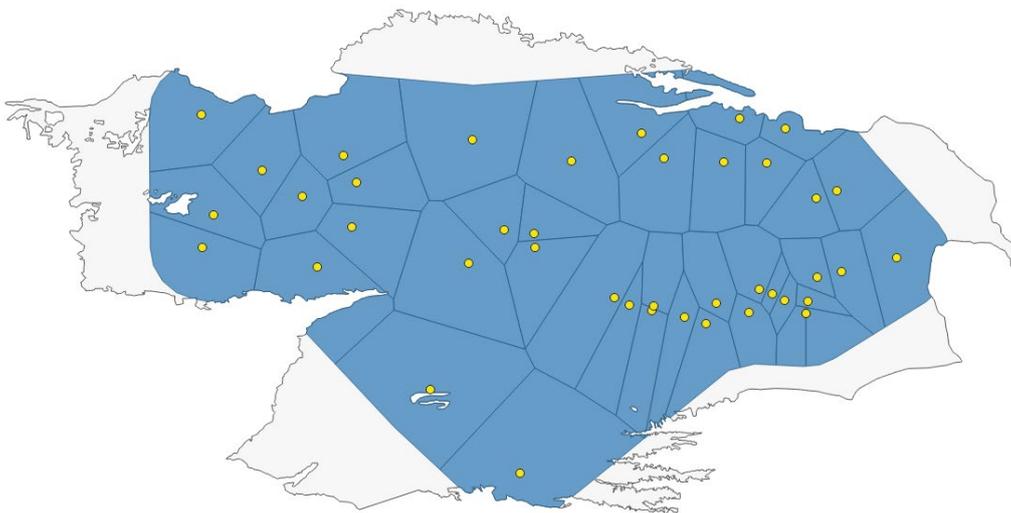


Figure 2. Thiessen polygons method example in the north-Ogallala aquifer system in the US. Each point in yellow represents a monitoring station.

Then, the $GWSA_{in-situ}$ is calculated as follows:

$$GWSA_{in-situ}[mm] = \sum_{i=1}^n \frac{h_{anomaly_i} \times area_i \times S_i}{area_T}$$

Eq. 1

$$h_{anomaly_i} = Mean\ GWL - GWL_i$$

Eq. 2

Where $h_{anomaly_i}$, per monitoring station i , refers to the difference between the average groundwater level for the whole period in consideration and the groundwater level at a given month (GWL_i), as shown in Eq. 2. The other two variables $area_i$ and S_i refer to the Thiessen area and storage coefficient per each monitoring station i . $area_T$ refers to the total area of the aquifer or evaluated site. n indicates the number of available monitoring stations.

Eq. 2 is used when the groundwater levels are given in terms of “depth below ground surface”, so the difference gives negative numbers when there is a depletion in groundwater, while positive numbers indicate a surplus of groundwater. When groundwater levels are given in terms of “meters above sea level”, $h_{anomaly}$ is calculated as shown in Eq. 3.

$$h_{anomaly_i} = GWL_i - Mean\ GWL$$

Eq. 3

Finally, when the $GWSA_{in-situ}$ is calculated, it is compared to the area-average G3P GWSA using the Pearson correlation coefficient. This coefficient measures the linear correlation between two sets of data and was chosen as a measure parameter since it can determine whether the signals of two time series are related.

5.3.2 In-situ GWLA calculation

In most cases during the evaluation process, hydrogeological data was not available. Therefore, in-situ GWSA could not be calculated. In this case, in-situ GWLA ($GWLA_{in-situ}$) is going to be used as a proxy for GWSA and is calculated as shown in Eq. 4.

$$GWLA_{in-situ} [mm] = \frac{1}{n} \sum_{i=1}^n h_{anomaly_i}$$

Eq. 4

Where $h_{anomaly}$ is calculated per each available monitoring station i following Eq. 2 or Eq. 3, and n is the number of available monitoring stations.

In this case, the area that is represented by the boreholes is the polygon that surrounds them with a buffer of 50 km on each side of the polygon.

Then, the evaluation is carried out by comparing the standardized time series of G3P GWSA and standardized time series of in-situ GWLA. The Pearson correlation coefficient is used as a measure parameter to determine if the signals of these time series are correlated. A similarity between them might indicate that G3P GWSA are behaving similarly to the in-situ ones.

6. Results

6.1 Ogallala Aquifer

The Ogallala Aquifer is one of the largest aquifers in the United States with an area of approx. 450.000 km², and extending over eight states: South Dakota, Nebraska, Wyoming, Colorado, Kansas, Oklahoma, New Mexico, and Texas. 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., 2020). This is a productive aquifer and consists of unconsolidated clay, silt, and sand, with some gravel and caliche. 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 (Miller et al., 1997).

Data for the evaluation was obtained from the National Ground-Water Monitoring Network provided by the USGS (U.S. Geological Survey, 2020). There are 310 available boreholes with groundwater level data from 2002 to 2020. Additionally, there is a detailed map of specific yield covering the entire area of the aquifer provided by the USGS. It was decided to use boreholes in the unconfined portion of the aquifer to accurately calculate GWSA.

Most of the available boreholes are in the northern part of the Ogallala aquifer. Therefore, it was decided to proceed with the evaluation process considering only the northern part of the aquifer. Two long periods that contain the most boreholes with complete time series are considered for the evaluation:

- 1) 2002 to 2016, which was the largest period with the most boreholes. Under this consideration, 21 boreholes were selected. A sub-area surrounding the selected boreholes is presented in Figure 3.
- 2) 2009 to 2016, which was the largest period with most boreholes covering a large portion of the aquifer. Under this consideration, 42 boreholes were selected. A sub-area surrounding the selected boreholes is presented in Figure 4.

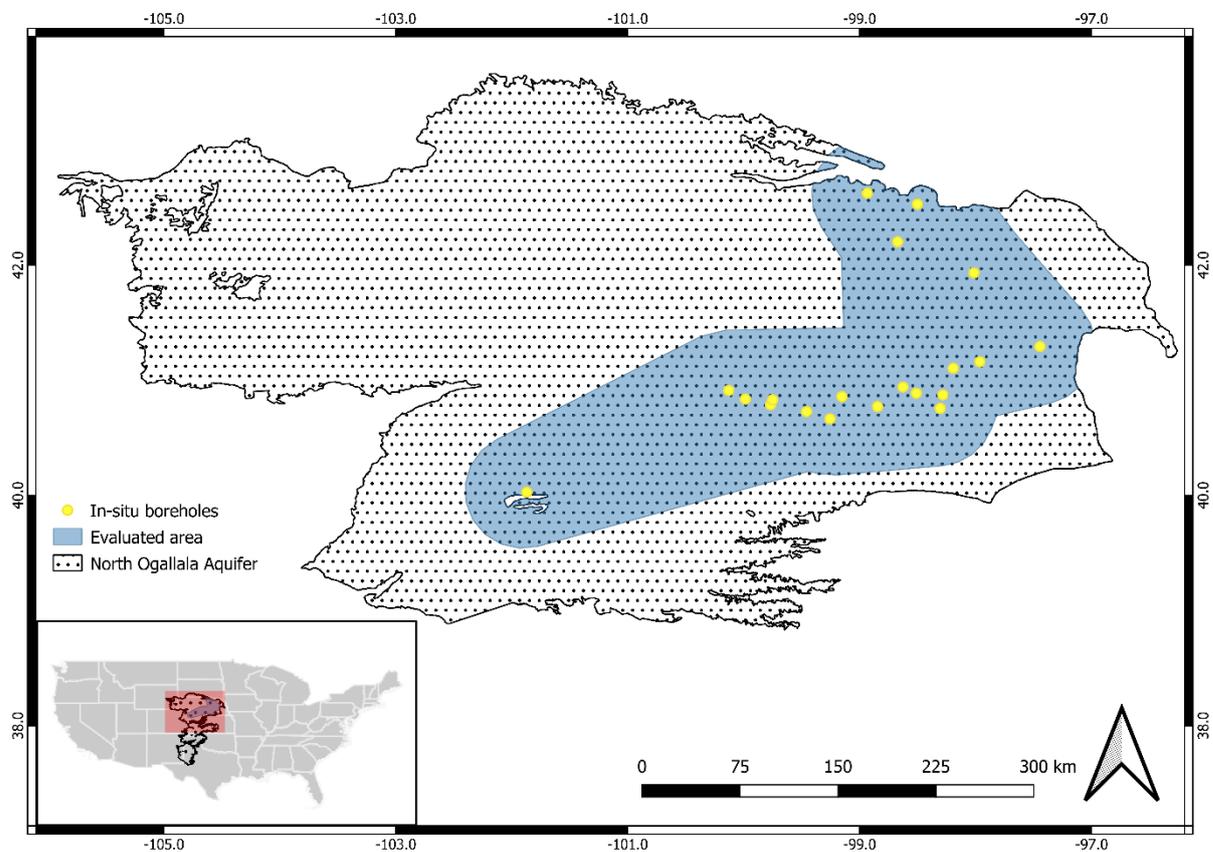


Figure 3. North Ogallala aquifer, evaluated area, and in-situ boreholes for the period 2002-2016

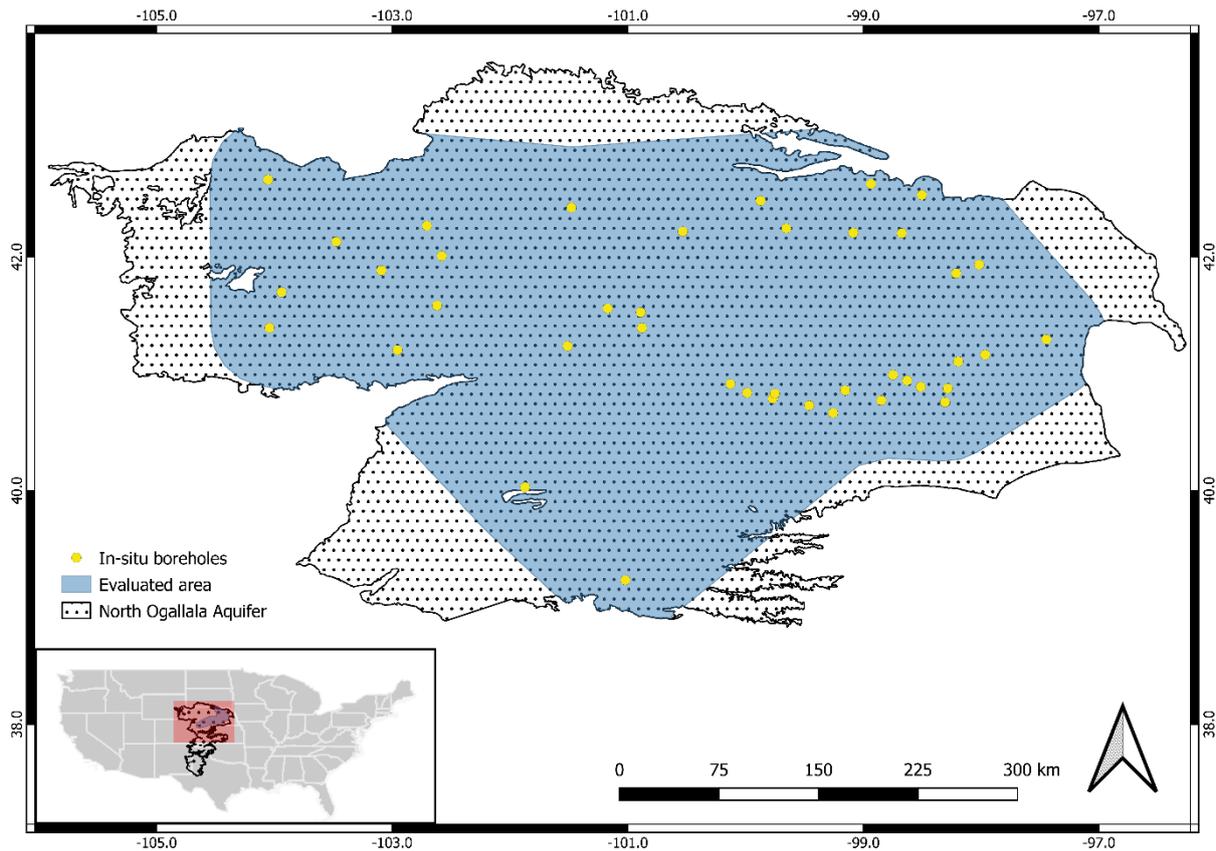


Figure 4. North Ogallala aquifer, evaluated area, and in-situ boreholes for the period 2009-2016

For the Ogallala aquifer, the G3P GWSA was calculated for both areas in blue from Figure 3 and Figure 4 following the procedure outlined in Section 5.2. The in-situ GWSA was calculated using groundwater level data and specific yield following the procedure outlined in Section 5.3.1 for both periods. Thereafter, the G3P GWSA and in-situ GWSA were compared to each other using the Pearson correlation. Figure 5 shows both GWSA time series (in-situ and G3P) and the uncertainty related to G3P GWSA, with a Pearson $r = 0.77$ for the period 2002-2016, while Figure 6 shows the comparison of both GWSA and the uncertainty related to G3P GWSA, with a Pearson $r = 0.51$ for the period 2009-2016. Table 1 shows the correlation coefficient between all G3P products and in-situ GWSA for both periods.

Table 1. Pearson r between G3P products and in-situ GWSA for the periods 2002-2016 and 2009-2016

Product	Pearson r	
	2002-2016	2009-2016
GWSA	0.77	0.52
GWSA_LA_hyd	0.79	0.60
GWSA_LA_grav	0.52	0.27

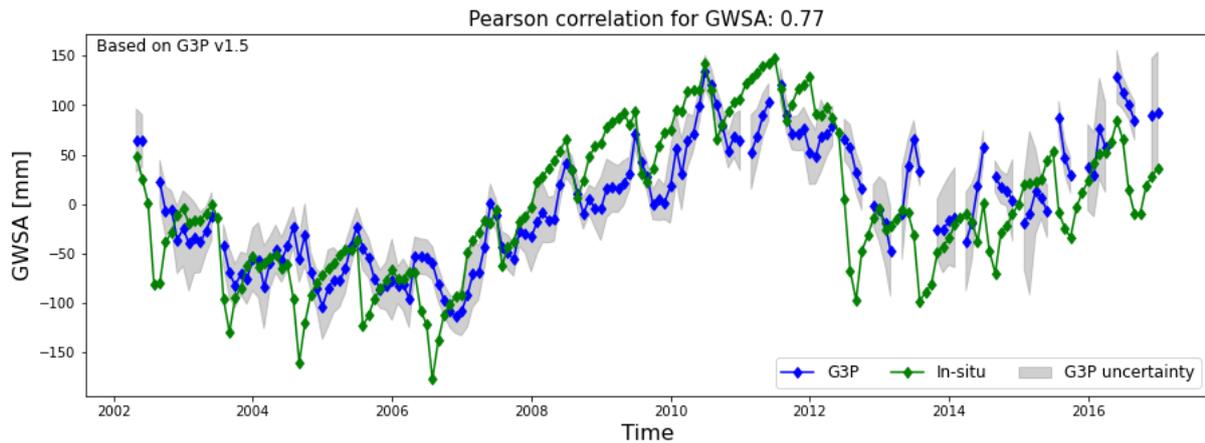


Figure 5. In-situ and G3P derived GWSA for the north part of the Ogallala aquifer, using G3P v1.5 for the period 2002-20016.

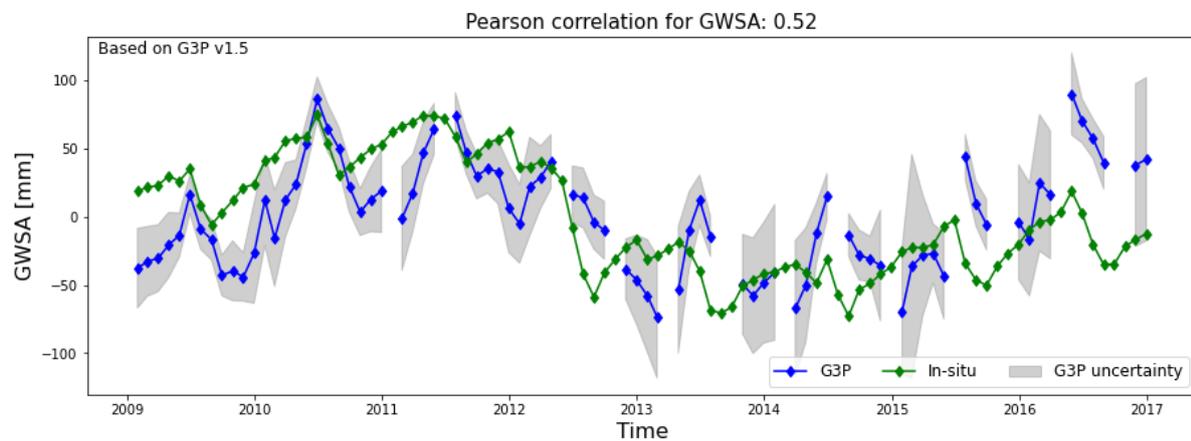


Figure 6. In-situ and G3P derived GWSA for the north part of the Ogallala aquifer using G3P v1.5 for the period 2009-2016.

The comparison to the other two products (GWSA_LA_hyd and GWSA_LA_grav) is presented in the annexes, Section 10.1.

The evaluation using the largest period shows a higher Pearson correlation than using the period between 2009-2016. This can happen due to several factors. In the case of the largest period, the borehole density is approximately 1 borehole per 900 km², whereas the borehole density for the period 2009-2016 is approximately 1 borehole per 1800 km². This means that, for the larger period, even though there is a smaller number of boreholes, these could replicate better the GWSA in a smaller area. This is also presented in Figure 7, where it is observed that the spatial distribution of GWSA is homogeneous over the smaller area (6a), while there are more extreme changes in GWSA in the larger area (6b). Since the area-average of GWSA from G3P is calculated to obtain monthly GWSA, these extreme values are included in the average, causing the GWSA values for specific months to be smoothed.

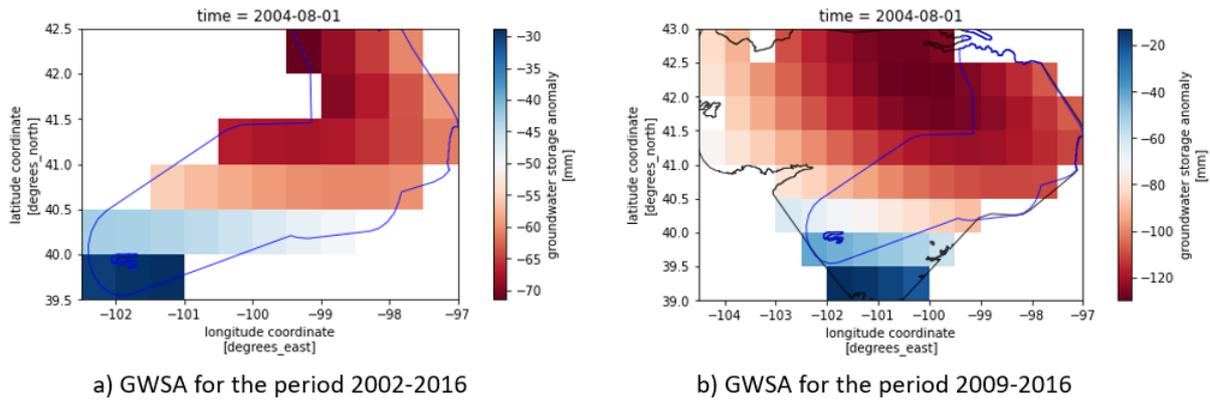


Figure 7. Spatial distribution of GWSA_G3P for the period 2002-2016 (a) and the period 2009-2016 (b) on August 1st, 2004. The area in (a) is contained in (b).

6.2 North China Aquifer System

Also known as North China Plain, the North China Aquifer system is located in the north-east of China, near the coast, with an area of approximately 440.330 km². It is characterized by cold, dry winters during December-March and hot, humid summers during July-September. Data for the evaluation was obtained from personal communication with Dr. Yun Pan, Capital Normal University, from his research on land subsidence in 2018 (Gong et al., 2018). The GWSA from 81 boreholes were provided for the period 2005-2012, from which 47 were within the boundaries of the aquifer and used for the evaluation. In this case, since the GWSA data were directly provided, an area-average of all the GWSA time series was calculated and a single time series for the area surrounding the boreholes was obtained. Figure 8 shows the evaluated area (in blue) within the aquifer and the boreholes used in the analysis.

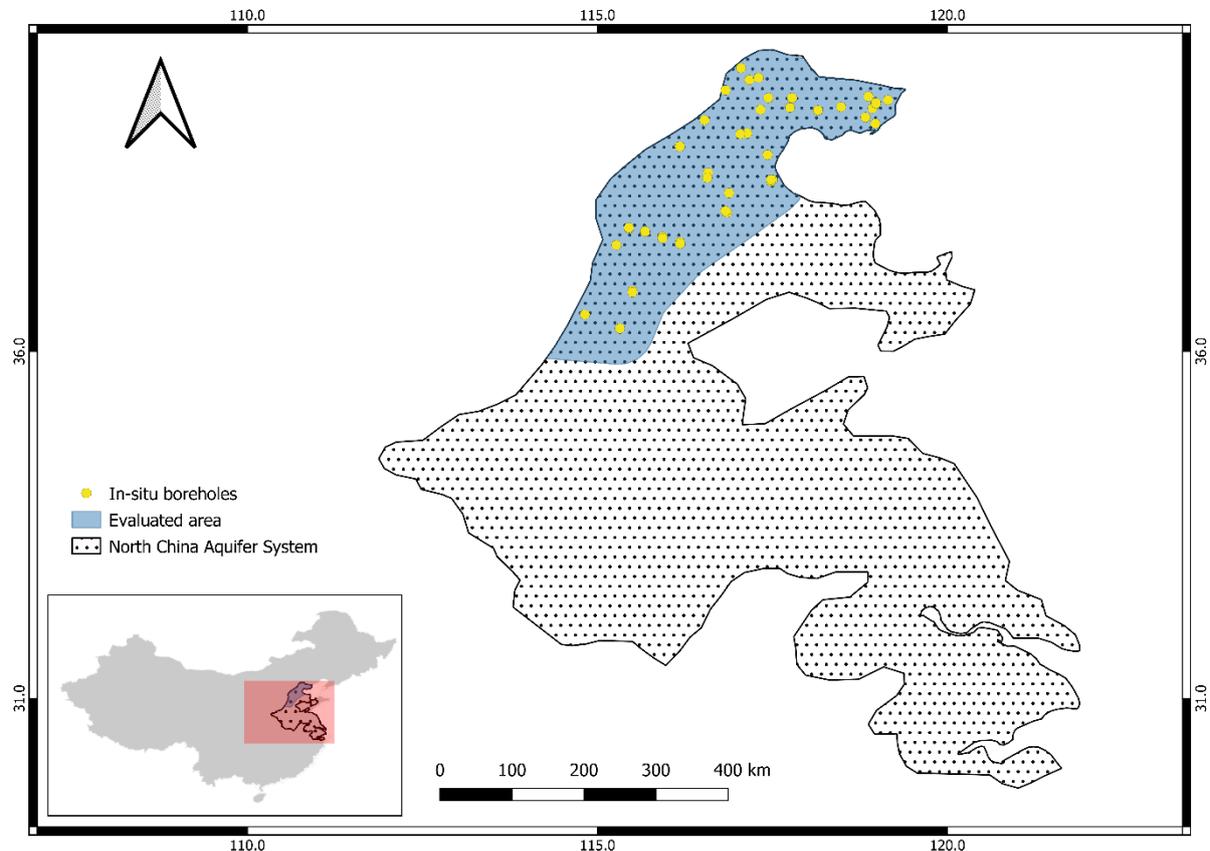


Figure 8. North China Aquifer System, evaluated area, and in-situ boreholes for the period 2005-2012

For the North China Aquifer system, the G3P GWSA was calculated for the area in blue, of approximately 83.690 km² (as shown in Figure 8) following the procedure outlined in Section 5.2. Thereafter, the G3P GWSA and in-situ GWSA were compared to each other using the Pearson correlation. Figure 9 shows both GWSA time series (in-situ and G3P), the uncertainty related to G3P GWSA, and Pearson $r = 0.57$ for the period 2005-2012. Table 2 shows the correlation coefficient between all G3P products and in-situ GWSA.

Table 2. Pearson r between G3P products and in-situ GWSA for the North China Aquifer system for v1.5

Product	Pearson r
GWSA	0.57
GWSA_LA_hyd	0.59
GWSA_LA_grav	0.38

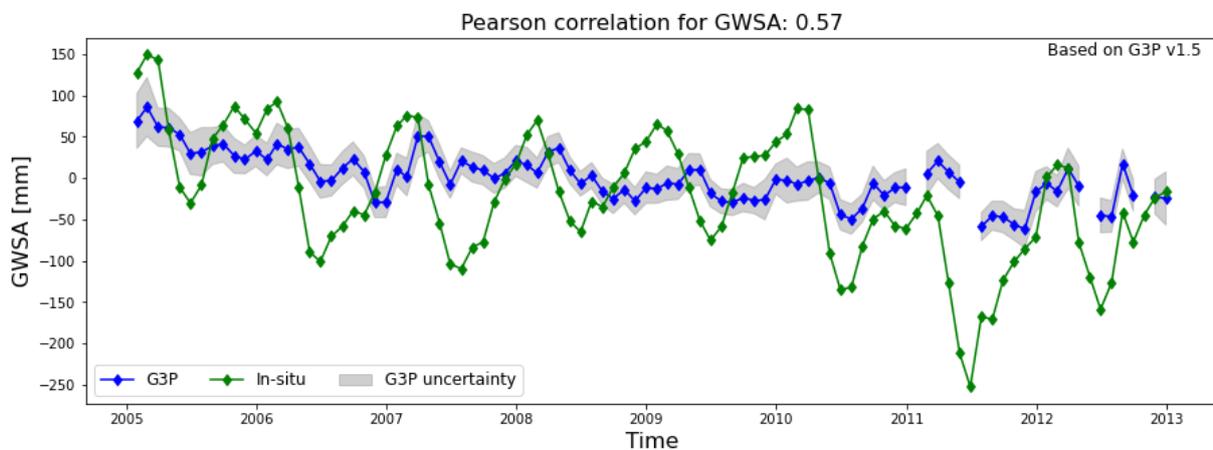


Figure 9. In-situ and G3P derived GWSA for the North China Aquifer System using G3P v1.5 for the period 2005-2012.

The comparison to the other two products (GWSA_LA_hyd and GWSA_LA_grav) is presented in the annexes, Section 10.2.

As the evaluation process advanced, it was notable that version 1.1 of G3P GWSA in this case had a better correlation with the in-situ GWSA than the later versions. Unfortunately, there is no version 1.1 for GWSA_LA_hyd and GWSA_LA_grav to compare to newer versions. In the discussion section, an evaluation of all versions and products of G3P for all aquifers is presented.

6.3 Floridan Aquifer System

The Floridan Aquifer System is a major source of groundwater in the southeast, being the only source of fresh water in some places of the United States. It has an area of approximately 262.000 km², covering the states of Florida, southeast Georgia, and small parts of Alabama and South Carolina (Bush & Johnston, 1988). The aquifer system is formed by limestone and dolomite, and, in terms of water yield, is the most productive of the aquifers in the area (Miller & Survey, 1990). See Figure 10.

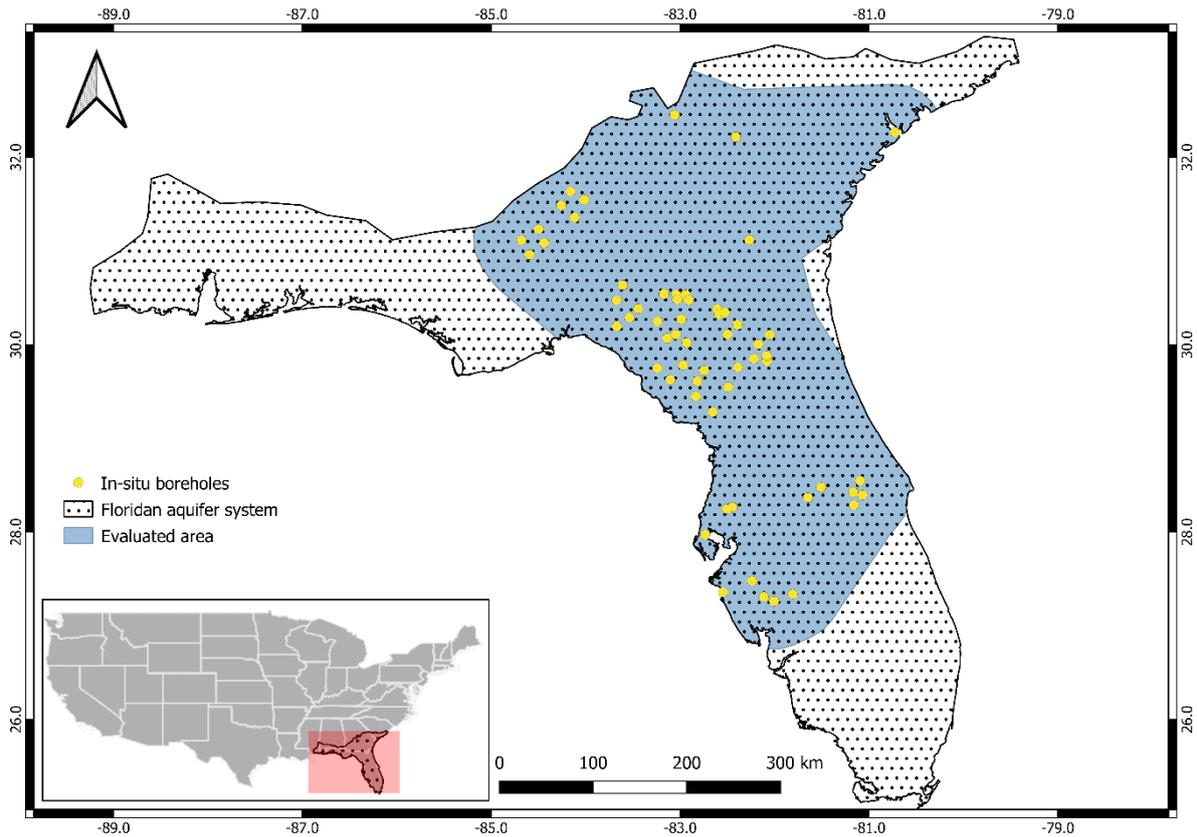


Figure 10. Floridan Aquifer system, in-situ boreholes, and evaluated area for the period 2002-2016

Data for the evaluation was obtained from the National Ground-Water Monitoring Network provided by the USGS (U.S. Geological Survey, 2020). There are 410 available boreholes with groundwater level data from 2002 to 2020. Due to the temporal sparsity of the data, 60 boreholes were selected to be further used in the evaluation. These boreholes had the longest period (from 2002 to 2016) with a complete groundwater level data set. These boreholes and the area to be evaluated can be observed in Figure 10. From this area, the area average G3P GWSA is calculated and presented in Figure 11. It is observed that GWSA is relatively stable, with positive peaks in 2006, 2010-2011, and 2014, and negative peaks in 2002, 2008, and 2012.

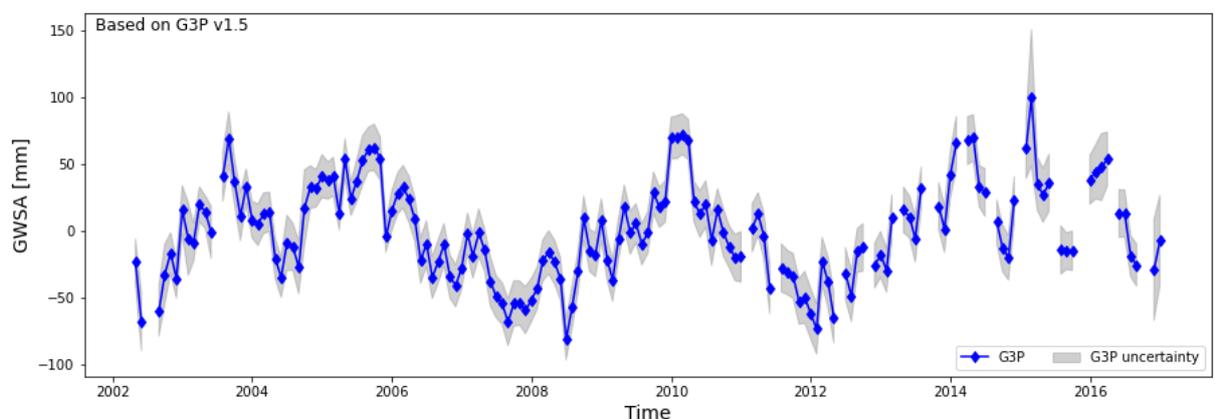


Figure 11. Area-average G3P GWSA for the blue region (Figure 10) within the Floridan Aquifer system and its uncertainty

Storage data was not available for the Floridan aquifer; therefore, the in-situ GWLA was calculated as detailed in Section 5.3.2 and compared to the G3P GWSA. Both time series were standardized and plotted as observed in Figure 12.

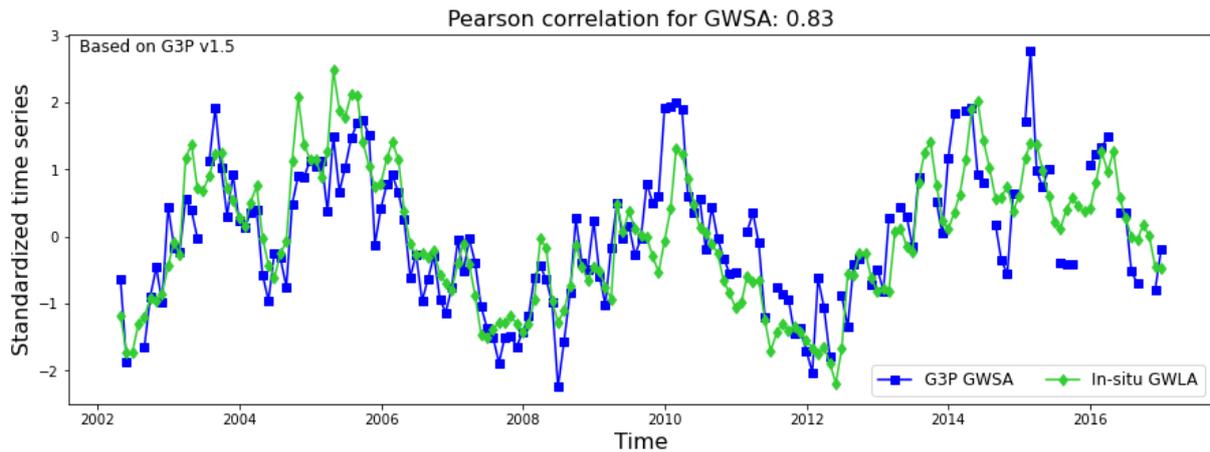


Figure 12. Standardized in-situ GWLA and G3P GWSA in the Floridan Aquifer system for the period 2002-2016

Table 3 shows the correlation coefficients between all 3 G3P products and in-situ GWLA. Even though the in-situ GWSA could not be calculated, it is observed that the GWLA signal replicates the G3P GWSA signal, with a Pearson correlation of 0.83 for G3P v1.5. In this case, the G3P product that is representing better the signal of in-situ GWLA is GWSA_LA_hyd, with a Pearson r of 0.89. These products are plotted against the in-situ GWLA and shown in the annexes, Section 10.3.

Table 3. Pearson r between in-situ GWLA and G3P GWSA for all 3 G3P products v1.5, for the period 2002-2016

Product	Pearson r
GWSA	0.83
GWSA_LA_hyd	0.89
GWSA_LA_grav	0.06

6.4 Basin and Range basin-fill Aquifers

In the west of the United States, there is a region that contains three principal aquifer types and is referred to as the Basin and Range aquifers. The aquifer system has an extension of approximately 635.000 km² and it extends 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 composed of volcanic and carbonate rocks and basin-fill deposits, the latter being the most productive. The area is the most arid in the US, where evapotranspiration is higher than the total amount of rainfall in a year (Planert et al., 1995). See Figure 13.

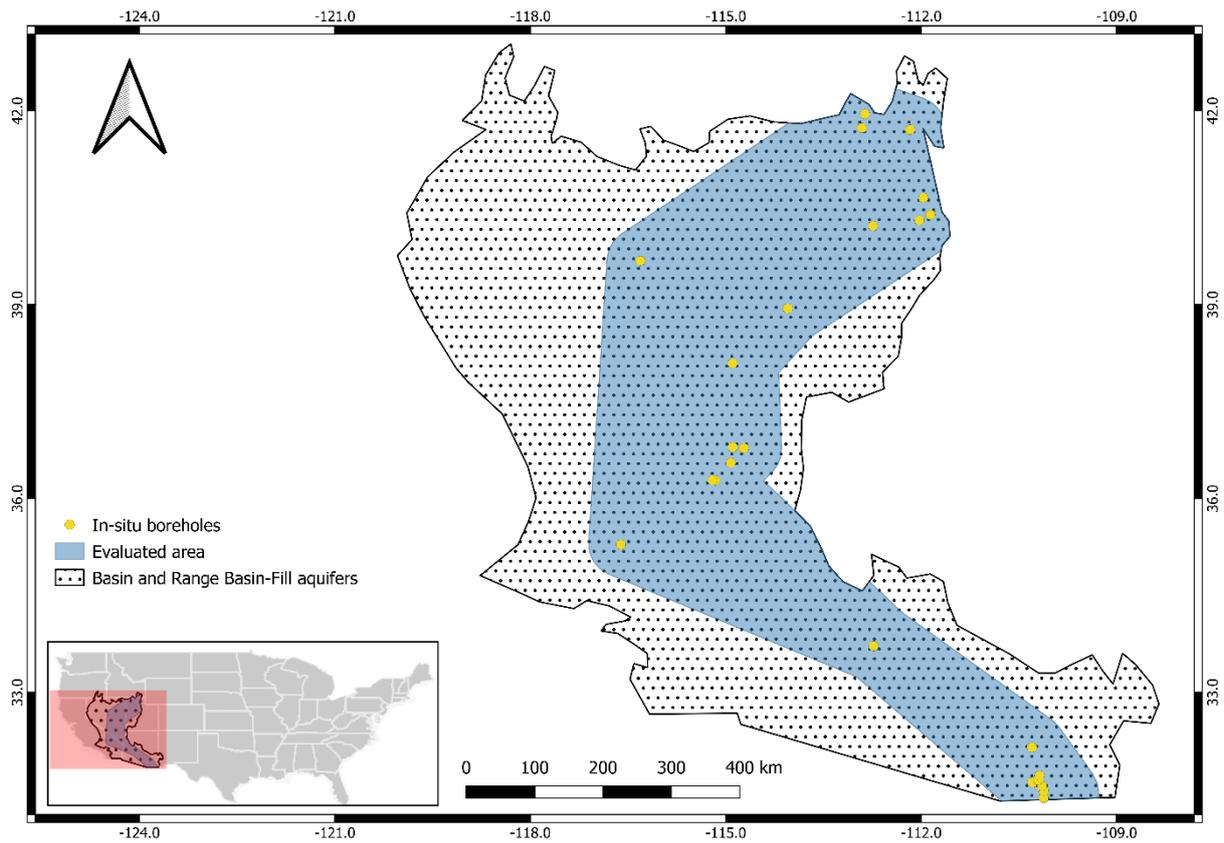


Figure 13. Basin and Range basin-fill aquifers, in-situ boreholes and evaluated area for the period 2012-2016

Data for the evaluation was obtained from the National Ground-Water Monitoring Network provided by the USGS (U.S. Geological Survey, 2020). There are 285 available boreholes with groundwater level data from 2005 to 2020. Due to the temporal sparsity of the data, 25 boreholes were selected to be further used in the evaluation. These boreholes had the longest period (from 2012 to 2016) with a complete groundwater level data set. These boreholes and the area to be evaluated can be observed in Figure 13. From this area, the area average G3P GWSA is calculated and presented in Figure 14. It is observed that there are a few missing data during this period; however, the trend is clearly indicating a negative trend towards groundwater depletion. The highest peak is at the beginning of the period in 2012, while the lowest peak occurs at the end of the period, by 2016.

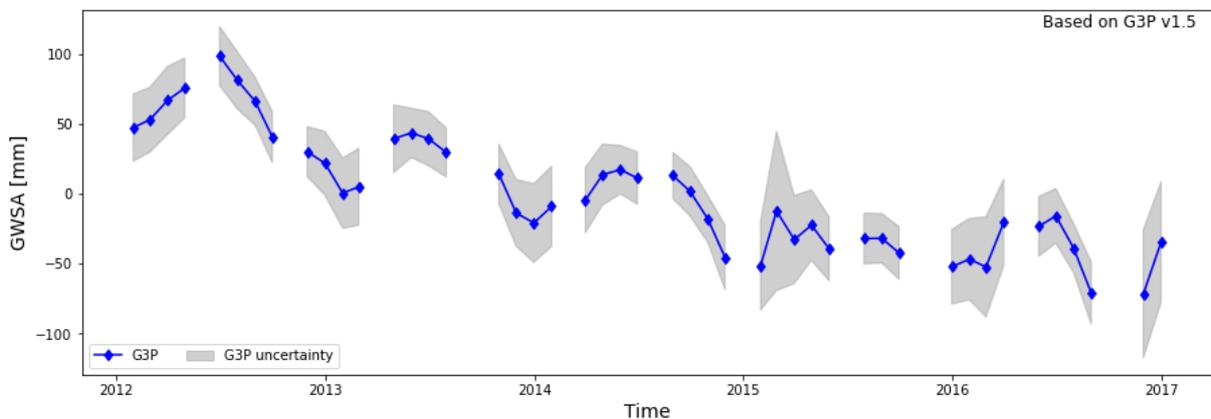


Figure 14. Area average G3P GWSA of Basin and Range basin-fill aquifers for the period 2012-2016

Storage data was not available in this area; therefore, the in-situ GWLA was calculated and compared to the G3P GWSA. Both time series were standardized and plotted as observed in Figure 15.

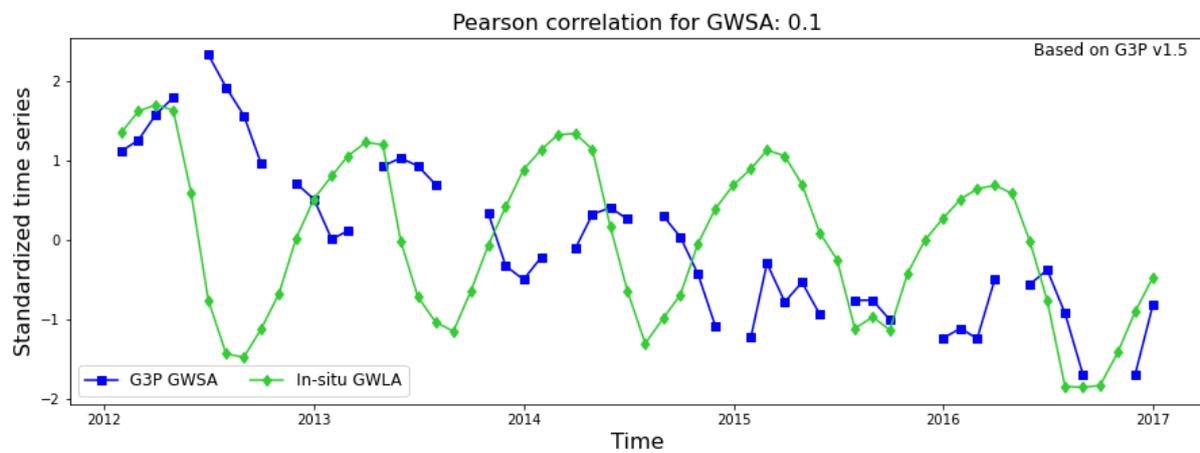


Figure 15. Standardized in-situ GWLA and G3P GWSA in the Basin and Range basin-fill aquifers for the period 2012-2016

Table 4 shows the correlation coefficients between all 3 G3P products and in-situ GWLA. In this case, it is observed that G3P could not replicate the signal that is observed in the in-situ groundwater level measurements. Pearson r is low, with a value of 0.1 for GWSA, and even lower values for the rest of the products with G3P version 1.5. This situation might occur due to the nature of the area. Since it is an arid area with higher rate of evapotranspiration than precipitation, the signal of groundwater might be responding to other water inputs. The in-situ GWLA might be inaccurately representing what happens in the ground due to the poor distribution of the wells over the area, the heterogeneity of the area, the short period that is considered for the evaluation, among other factors. In the annexes, Section 10.4, GWSA_LA_hyd and GWSA_LA_grav are plotted against the in-situ GWLA.

Table 4. Pearson r between in-situ GWLA and G3P GWSA for all 3 G3P products v1.5, for the Basin and Range basin-fill aquifers for the period 2012-2016

Product	Pearson r
GWSA	0.1
GWSA_LA_hyd	0.06
GWSA_LA_grav	-0.04

6.5 Cambrian Ordovician aquifer system

The Cambrian-Ordovician aquifer system is composed by individual aquifers, capped by a confining unit. There are outcrops of the Cambrian and Ordovician rocks in each of the states where they are located, namely: Iowa, Michigan, Minnesota, and Wisconsin. The aquifer system suffers from intensive groundwater pumping activities in southeastern Wisconsin, much of Iowa, and especially in Chicago. It is formed by a sandstone and dolomite aquifer and two sandstone aquifers, which are separated by less-permeable confining units (Olcott & Survey, 1992). The aquifer covers an area of approximately 130.000 km² in the northeast of the United States. See Figure 16.

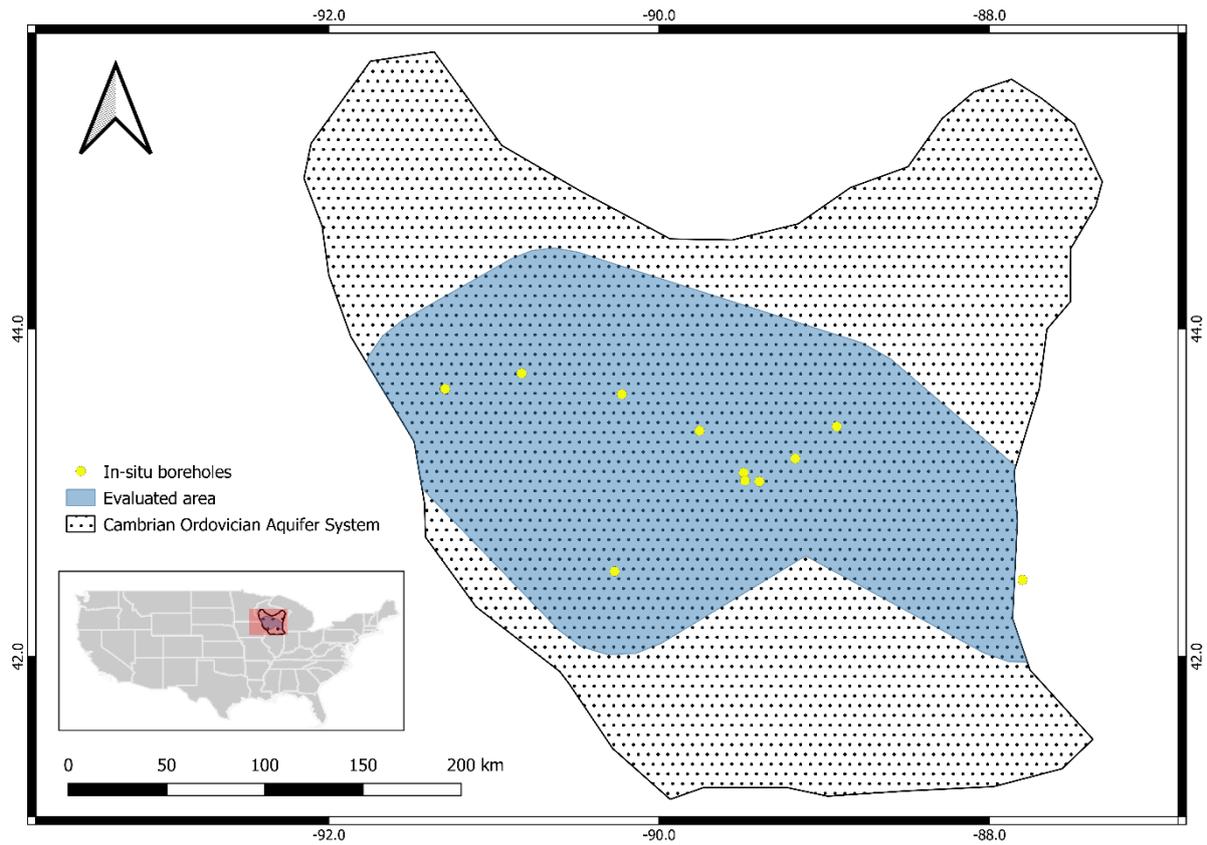


Figure 16. Cambrian-Ordovician aquifer system, in-situ boreholes and evaluated area for the period 2012-2016

Data for the evaluation was obtained from the National Ground-Water Monitoring Network provided by the USGS (U.S. Geological Survey, 2020). There are 242 available boreholes with groundwater level data from 1979 to 2020. Due to the temporal sparsity of the data, 16 boreholes were selected to be further used in the evaluation. These boreholes had the longest period (from 2012 to 2016) with a complete groundwater level data set. These boreholes and the area to be evaluated can be observed in Figure 16. From this area, the area average G3P GWSA is calculated and presented in Figure 17. It is observed that there are a few missing data during this period; however, the trend is clearly indicating a positive trend of GWSA. The highest peak occurs by the end of the period in 2016, while the lowest peak occurs in 2013.

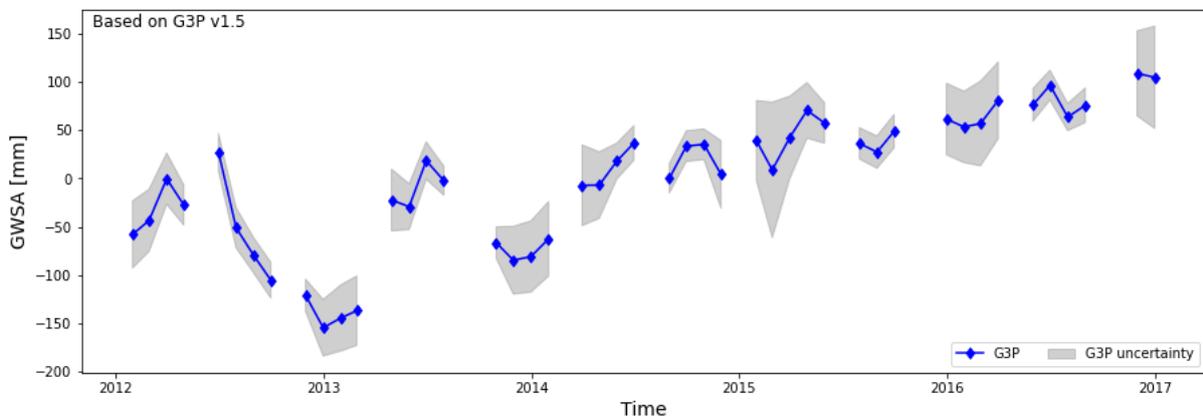


Figure 17. Area average G3P GWSA of the Cambrian-Ordovician aquifer system for the period 2012-2016

Storage data was not available in this area; therefore, the in-situ GWLA was calculated and compared to the G3P GWSA. Both time series were standardized and plotted as observed in Figure 18.

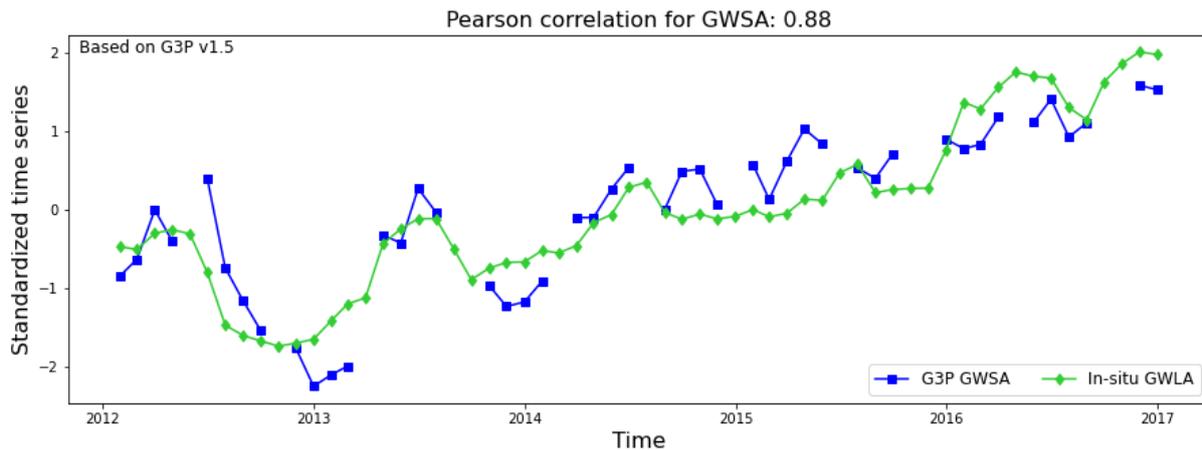


Figure 18. Standardized in-situ GWLA and G3P GWSA in the Cambrian-Ordovician aquifer system for the period 2012-2016

Table 5 shows the correlation coefficients between all 3 G3P products and in-situ GWLA for the Cambrian-Ordovician aquifer system. In this case, it is observed that G3P could replicate the signal that is observed in the in-situ groundwater level measurements, especially the positive trend of increasing GWSA and the seasonal changes throughout the period. Pearson r is high, with a value of 0.89 for GWSA, and for the rest of the products with G3P version 1.5 the value remains high (0.7-0.9). In the annexes, Section 10.5, GWSA_LA_hyd and GWSA_LA_grav are plotted against the in-situ GWLA.

Table 5. Pearson r between in-situ GWLA and G3P GWSA for all 3 G3P products v1.5 for the Cambrian-Ordovician aquifer system, for the period 2012-2016

Product	Pearson r
GWSA	0.89
GWSA_LA_hyd	0.9
GWSA_LA_grav	0.7

6.6 Mississippi River Valley

The Mississippi river valley is an extensive aquifer in the United States with an area of approximately 86.000 km² and underlies seven states near the East and West Gulf Coastal Plains, namely: Louisiana, Arkansas, Missouri, Illinois, Tennessee, Kentucky, and Mississippi. The area is drained by several rivers that end up in the Gulf of Mexico. Groundwater accounts for 38% of total water use in three of these states and is recharged mainly by rainfall, which is evidenced in its temporal and spatial variations (Renken & Survey, 1998). See Figure 19.

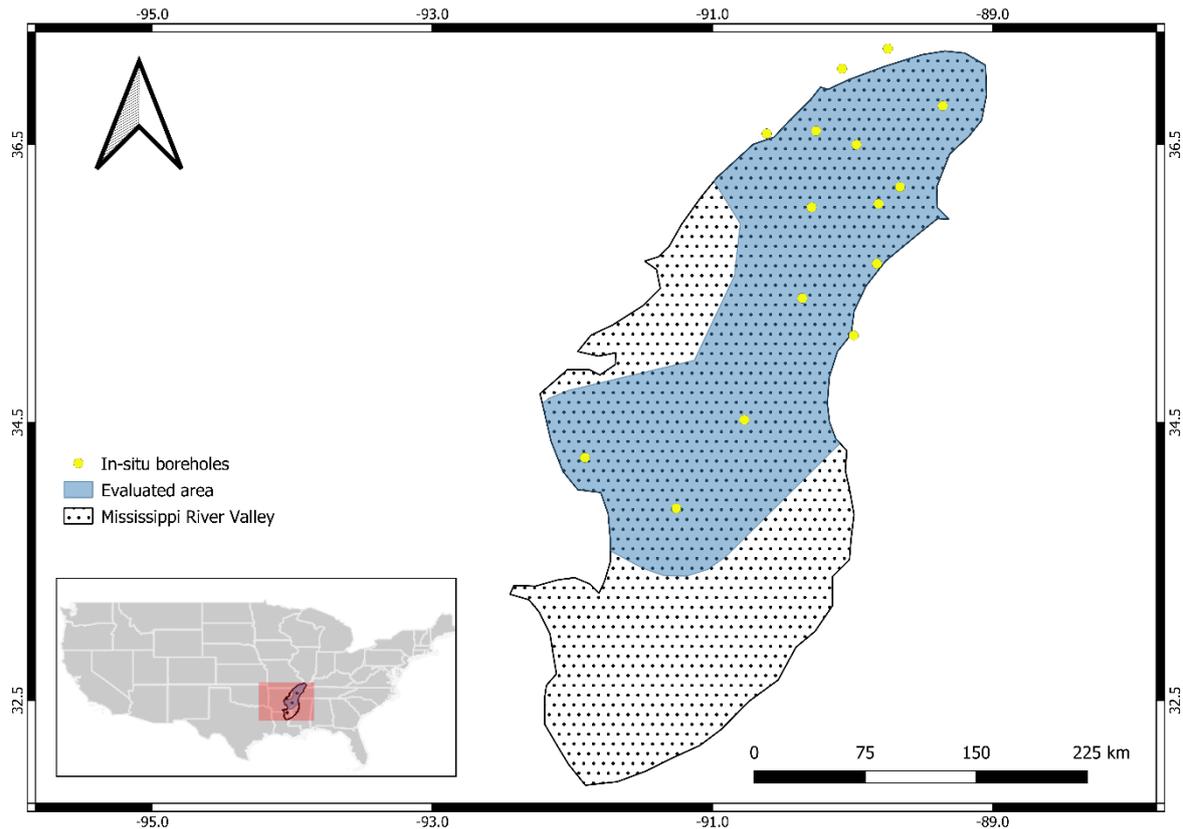


Figure 19. Mississippi River Valley, in-situ boreholes and evaluated area for the period 2008-2016

Data for the evaluation was obtained from the National Ground-Water Monitoring Network provided by the USGS (U.S. Geological Survey, 2020). There are 164 available boreholes with groundwater level data from 1960 to 2020. Due to the temporal sparsity of the data, 15 boreholes were selected to be further used in the evaluation. These boreholes had the longest period (from 2008 to 2016) with a complete groundwater level data set. These boreholes and the area to be evaluated can be observed in Figure 19. From this area, the area average G3P GWSA is calculated and presented in Figure 20. It is observed that the GWSA are stable until 2011 and after that, there is a decrease with most values near or below zero. The highest peak is observed at the beginning of the period in 2008 and the lowest value is observed in 2012. There are a few missing data during this period; however, the trend indicates lower values of GWSA for the second half of the period.

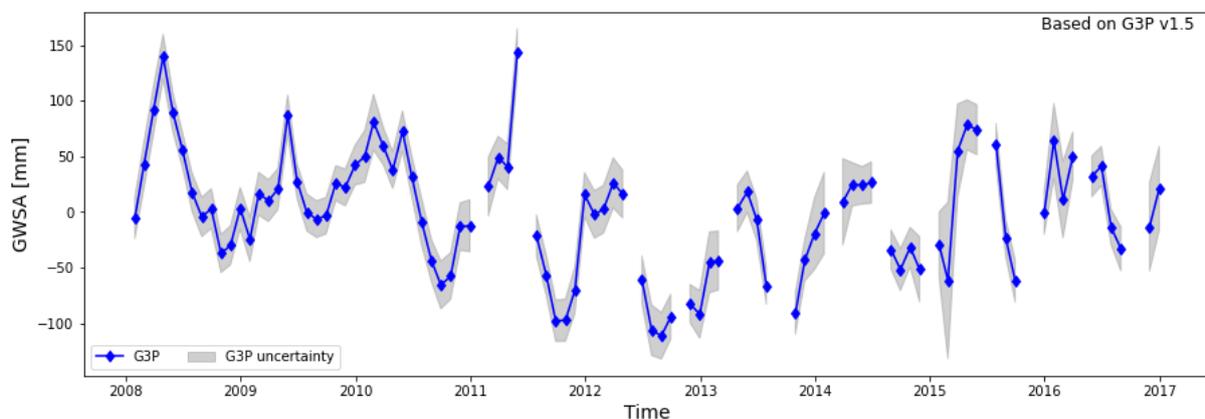


Figure 20. Area average G3P GWSA of the Mississippi River Valley for the period 2012-2016

Storage data was not available in this area; therefore, the in-situ GWLA was calculated and compared to the G3P GWSA. Both time series were standardized and plotted as observed in Figure 21.

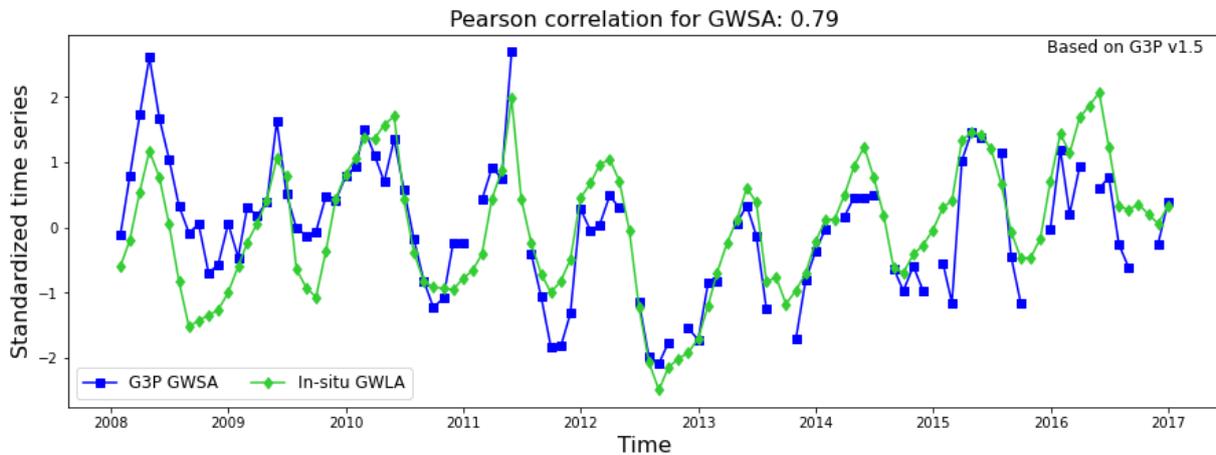


Figure 21. Standardized in-situ GWLA and G3P GWSA in the Mississippi River Valley for the period 2008-2016

Table 6 shows the correlation coefficients between all 3 G3P products and in-situ GWLA for the Mississippi River Valley. In this case, it is observed that G3P could replicate the signal that is observed through the in-situ groundwater level measurements. Pearson r is high, with a value of 0.79 for GWSA, but for the rest of the products with G3P version 1.5 the decreases (0.53-0.18). In the annexes, Section 10.6, GWSA_LA_hyd and GWSA_LA_grav are plotted against the in-situ GWLA.

Table 6. Pearson r between in-situ GWLA and G3P GWSA for all 3 G3P products v1.5 for the Mississippi River Valley, for the period 2008-2016

Product	Pearson r
GWSA	0.79
GWSA_LA_hyd	0.53
GWSA_LA_grav	0.18

6.7 Northern High Plains

The Northern High Plains aquifer system is a transboundary aquifer shared by the United States and Canada. It has an area of approximately 300.000 km², and on the US side covers most of North Dakota, and part of Montana, Wyoming, and South Dakota; while on the Canadian side, it covers part of Saskatchewan and Manitoba. See Figure 22.

The major aquifers of the Northern Great Plains aquifer system are sandstones of Tertiary and Cretaceous age and carbonate rocks of Paleozoic age. Intensive groundwater abstraction in the area have made groundwater levels and artesian pressures decline significantly. For this reason, steps have been taken from State governments to monitor these declines by limiting or stopping the installation or development of additional wells (Whitehead & Survey, 1996).

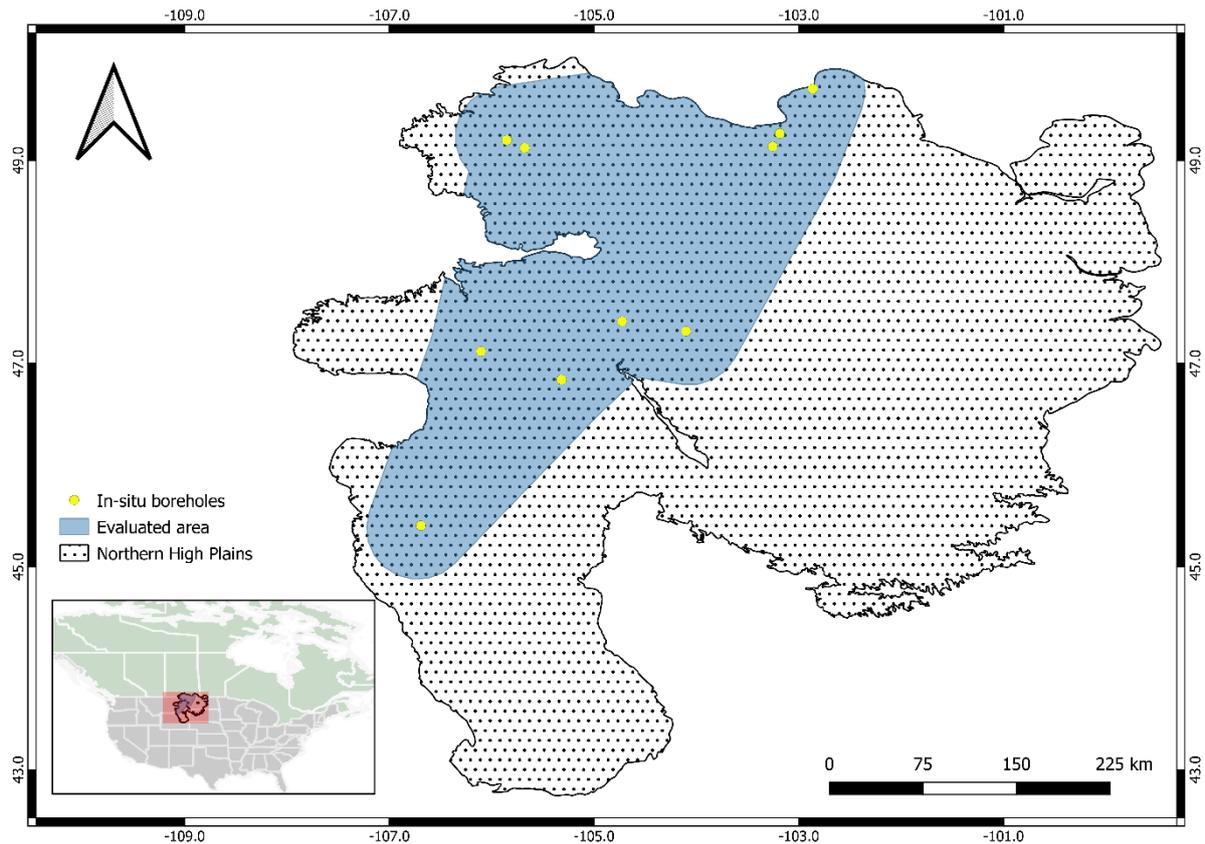


Figure 22. Northern High Plains aquifer system, in-situ boreholes and evaluated area for the period 2008-2013

Data for the evaluation was obtained from the National Ground-Water Monitoring Network provided by the USGS (U.S. Geological Survey, 2020). There are 84 available boreholes with groundwater level data from 1964 to 2020 in and around the limits of the aquifer system. Due to the temporal sparsity of the data, 10 boreholes were selected to be further used in the evaluation. These boreholes had the longest period (from 2008 to 2013) with a complete groundwater level data set. These boreholes and the area to be evaluated can be observed in Figure 22. From this area, the area average G3P GWSA is calculated and presented in Figure 23. It is observed that before 2010, GWSA values are below zero, indicating a depleting state of the aquifer; however, after 2010 GWSA have an increasing trend, being above zero for the rest of the period, until the end of 2013. The positive trend could be indicating the consequences of the limitation in groundwater abstraction from State control.

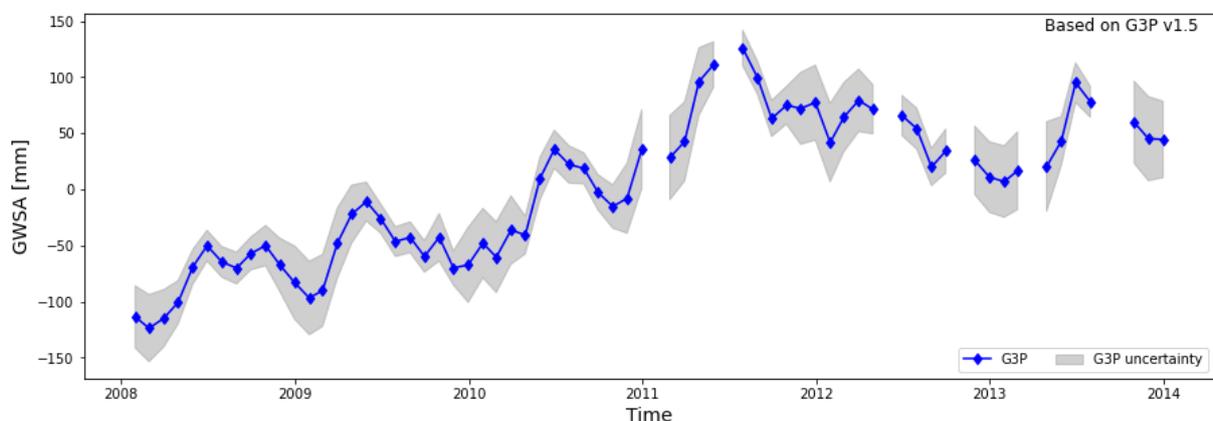


Figure 23. Area average G3P GWSA of the Northern High Plains aquifer system for the period 2008-2013

Storage data was not available in this area; therefore, the in-situ GWLA was calculated and compared to the G3P GWSA. Both time series were standardized and plotted as observed in Figure 24.

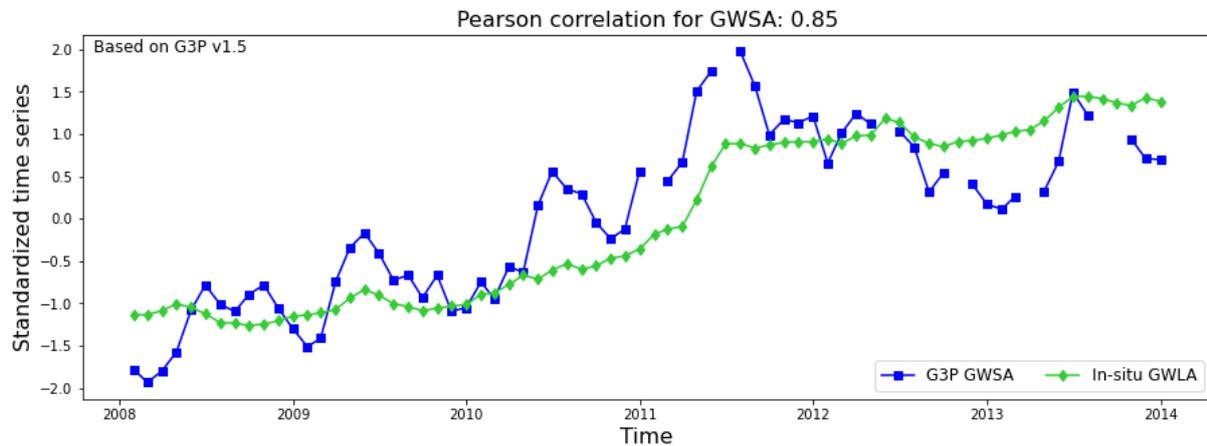


Figure 24. Standardized in-situ GWLA and G3P GWSA in the Northern High Plains aquifer system for the period 2008-2013

Table 7 shows the correlation coefficients between all 3 G3P products and in-situ GWLA for the Northern High Plains aquifer system. Pearson r is high, with a value of 0.85 for GWSA, and for the rest of the products with G3P version 1.5 it remains within a medium to high range (0.44-0.84). In this case, it is observed that G3P could replicate well the signal that is observed through the in-situ groundwater level measurements, but the extremes highs and lows are more pronounced in the G3P GWSA than in the in-situ GWLA. This might be due to the limited number of in-situ observations that might indicate the overall signal of the aquifer groundwater level fluctuations but fail to represent the specific behaviour of the evaluated area. In the annexes, Section 10.7, GWSA_LA_hyd and GWSA_LA_grav are plotted against the in-situ GWLA.

Table 7. Pearson r between in-situ GWLA and G3P GWSA for all 3 G3P products v1.5 for the Northern High Plains aquifer system, for the period 2008-2013

Product	Pearson r
GWSA	0.85
GWSA_LA_hyd	0.84
GWSA_LA_grav	0.44

6.8 Paris Basin

The Paris Basin is located in the north of France, covering an area of approximately 150.000 km². It is composed of seven main aquifer layers with semi-permeable layers in-between (Contoux et al., 2013). Its thickness can reach up to 3.000 m. and it consists of sandstones and carbonated permeable horizons separated by low permeability formations (Megnien, 1980). The mean annual precipitation is around 650 mm. and temperatures oscillate between four and sixteen degrees Celsius during winter and summer, respectively. See Figure 25.

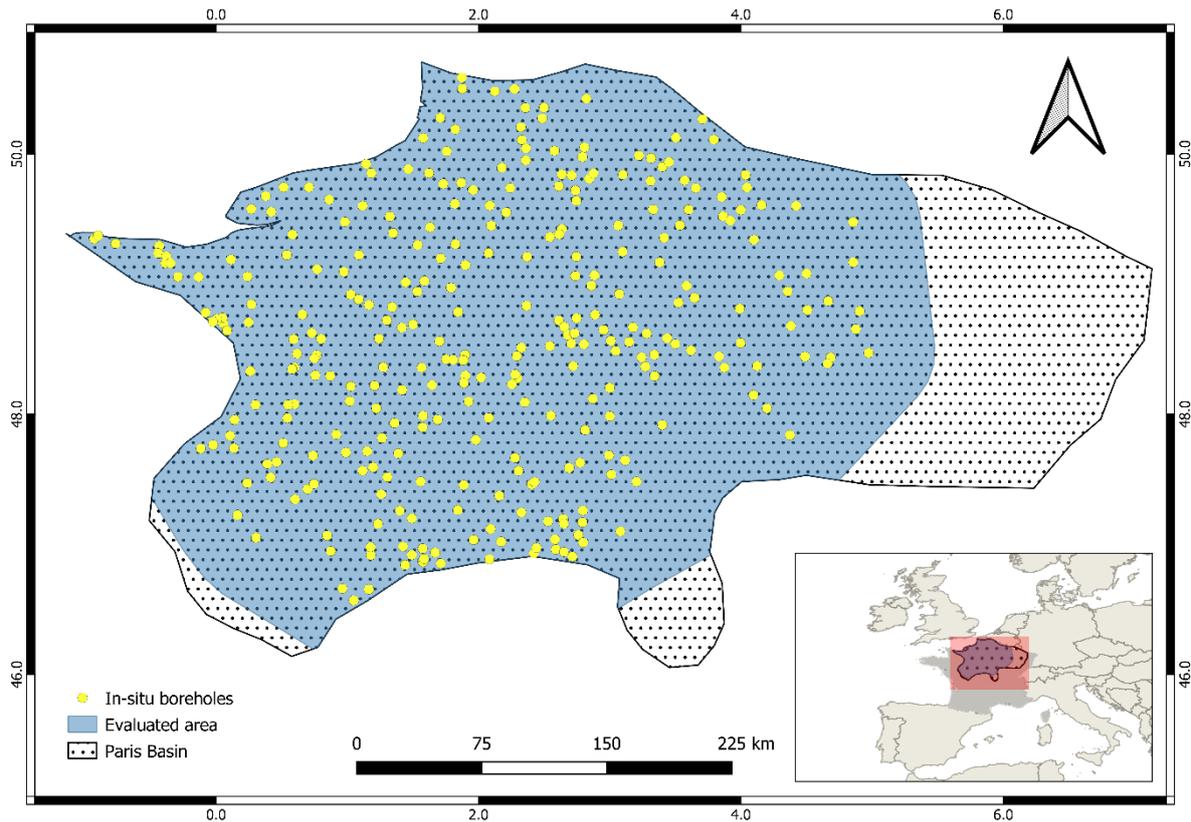


Figure 25. Paris Basin, in-situ boreholes and evaluated area for the period 2002-2016

Data for the evaluation was obtained from the National Portal for Access to Groundwater Data (ADES), which is maintained by the French Geological Survey (BRGM, 2022). There are 936 available boreholes with groundwater level data from 1970 to 2021 in and around the limits of the aquifer system. Due to the temporal sparsity of the data, 310 boreholes were selected to be further used in the evaluation. These boreholes had the longest period (from 2002 to 2016) with a complete groundwater level data set. These boreholes and the area to be evaluated can be observed in Figure 25. It is notable that in this study this is the aquifer with most borehole density, it being 1 borehole per 500 km² approximately. From this area, the area average G3P GWSA is calculated and presented in Figure 26. It is observed that the GWSA are relatively stable and fluctuating around zero.

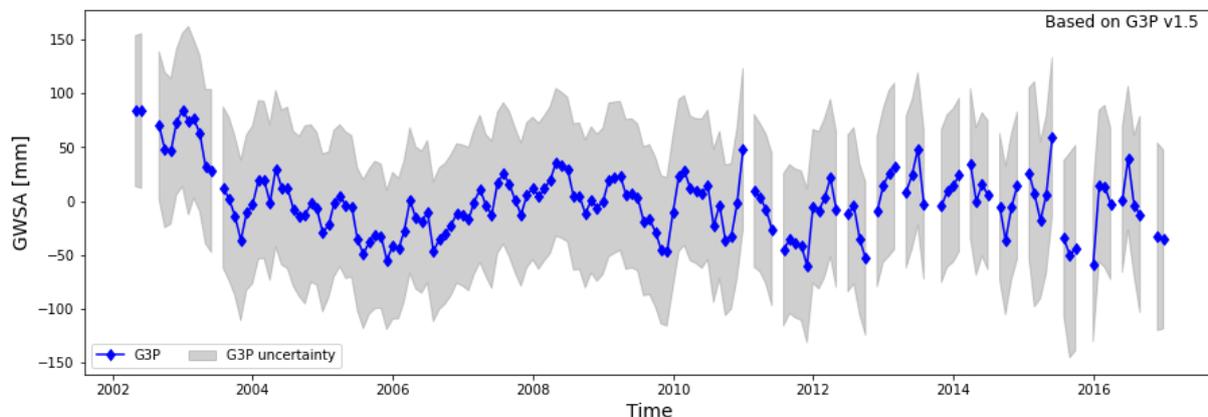


Figure 26. Area average G3P GWSA of the Paris Basin for the period 2002-2016

In this case, it is important to note that the uncertainty related to the G3P GWSA is greater than in the previously described aquifers. And as in previous aquifer evaluations, it is not

possible to access aquifer storage data to calculate in-situ GWSA. Therefore, the evaluation that is carried out in this case using GWLA needs to be critically interpreted. Figure 27 shows both standardized in-situ GWLA and G3P GWSA time series.

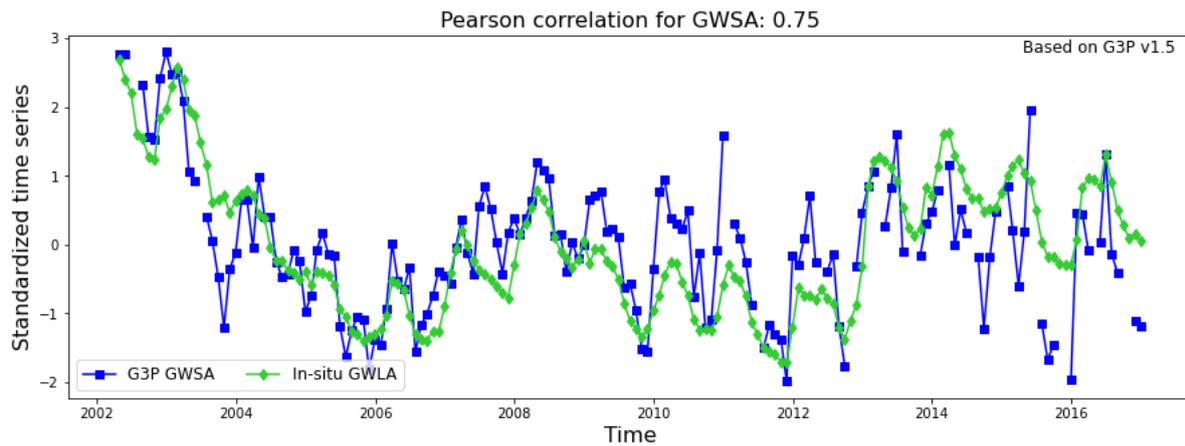


Figure 27. Standardized in-situ GWLA and G3P GWSA in the Paris Basin for the period 2002-2016

Table 8 shows the correlation coefficients between all 3 G3P products and in-situ GWLA for the Paris Basin. Pearson r is high, with a value of 0.75 for GWSA, and for the rest of the products with G3P version 1.5 it changes, being 0.72 for GWSA_LA_hyd and 0.02 for GWSA_LA_grav. It is observed that G3P could replicate well the signal that is calculated using the available in-situ groundwater level measurements. The declining trend at the beginning of the period and the slight increase by 2013 is also captured by G3P. Despite the similarities between the in-situ GWLA and G3P GWSA, the uncertainty needs to be addressed every time the data set in this area is to be used. In the annexes, Section 10.8, GWSA_LA_hyd and GWSA_LA_grav are plotted against the in-situ GWLA.

Table 8. Pearson r between in-situ GWLA and G3P GWSA for all 3 G3P products v1.5 for the Paris Basin, for the period 2002-2016

Product	Pearson r
GWSA	0.75
GWSA_LA_hyd	0.72
GWSA_LA_grav	0.02

6.9 Murray Darling Basin

The Murray Darling Basin in Australia covers an area of around 1.1 million of km² over Queensland, New South Wales, Victoria, and South Australia, and holds major groundwater systems of fractured rocks, alluvial deposits, and tertiary limestone. Around 2 million people live in the area, and it provides 75% of water to the agriculture sector. Depending on the location, the climate can vary from tropical in the north, Mediterranean and Alpine in the south, and semi-arid in the west (Fu et al., 2022). See Figure 28.

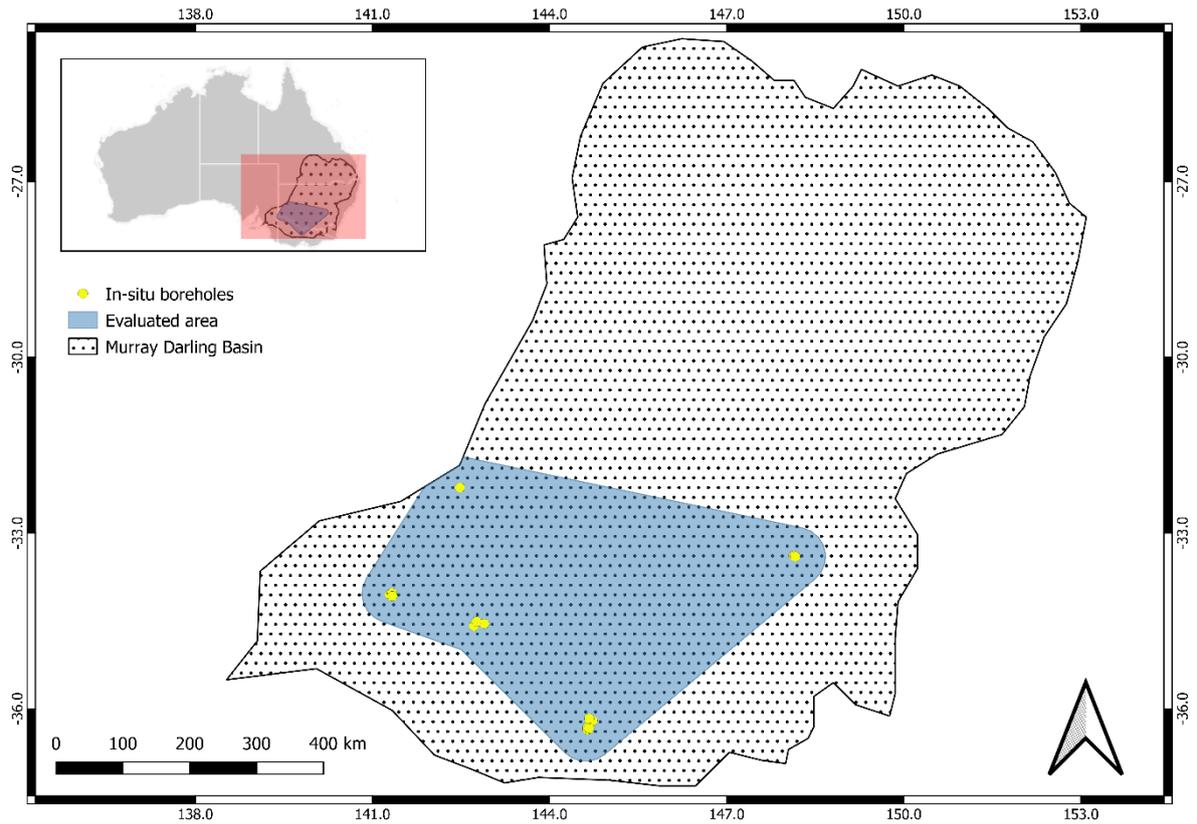


Figure 28. Murray Darling Basin, in-situ boreholes and evaluated area for the period 2002-2016

Data for the evaluation was obtained from the Australian Groundwater Explorer (Bureau of Meteorology, 2022). There are around 36.000 available boreholes with groundwater level data from around 1900 onwards. Due to the temporal sparsity of the data, and to comply with the temporal G3P data availability, only 36 boreholes could be selected to be further used in the evaluation. These boreholes had the longest period (from 2002 to 2016) with a complete groundwater level data set. These boreholes and the area to be evaluated can be observed in Figure 28. From this area, the area average G3P GWSA is calculated and presented in Figure 29. It is observed that the GWSA has a declining trend at the beginning of the period until 2010. After 2010, there is an increase in GWSA, and it stabilizes with values above zero until the end of the period. This is in accordance with previous studies in the area, where declining trends have also been found during this period, likely due to groundwater abstractions (Fu et al., 2022).

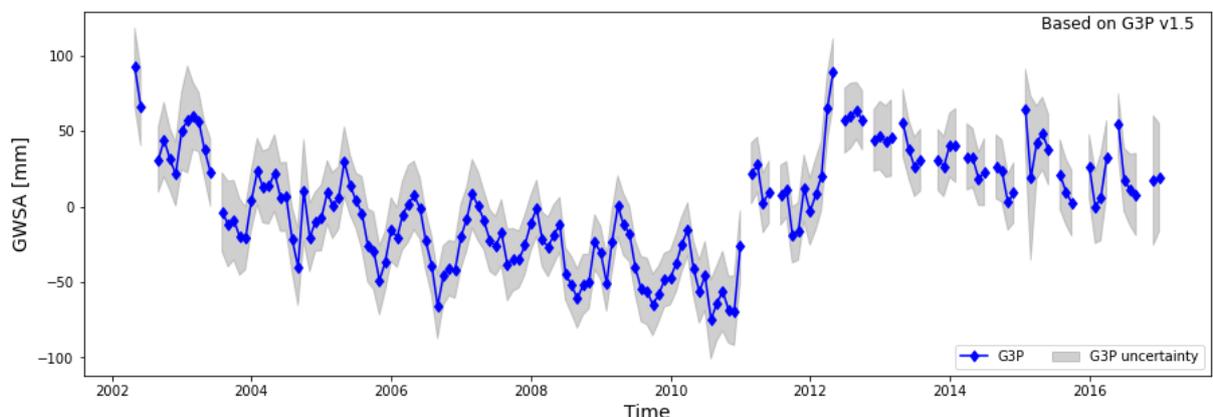


Figure 29. Area average G3P GWSA of the Murray Darling Basin for the period 2002-2016

Storage data was not available in this area; therefore, the in-situ GWLA was calculated and compared to the G3P GWSA. Both time series were standardized and plotted as observed in Figure 30.

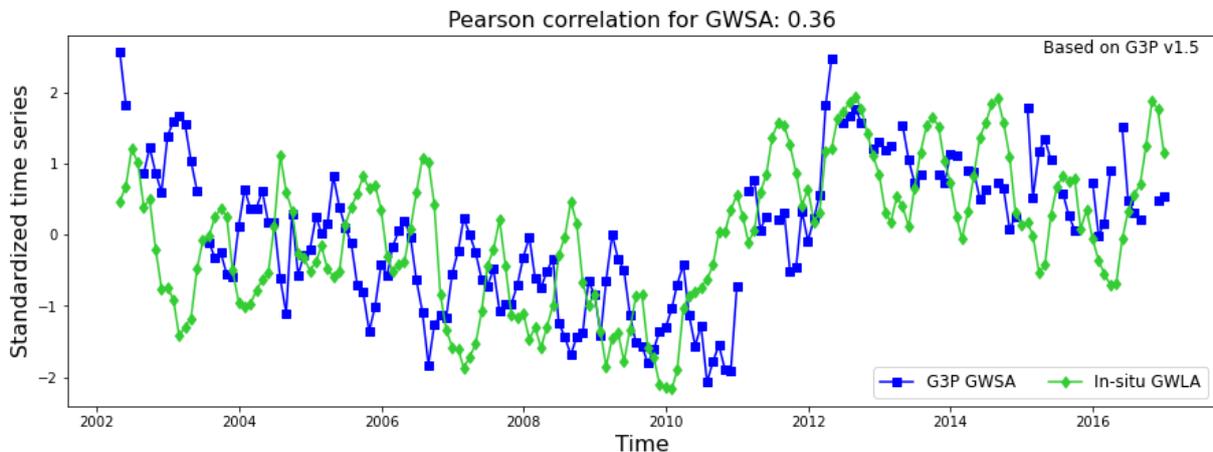


Figure 30. Standardized in-situ GWLA and G3P GWSA in the Murray Darling Basin for the period 2002-2016

Table 9 shows the correlation coefficients between all 3 G3P products and in-situ GWLA for the Murray Darling Basin. Pearson r is low, with a value of 0.36 for GWSA, and for the rest of the products with G3P version 1.5 it remains low, being 0.54 for GWSA_LA_hyd and -0.28 for GWSA_LA_grav. In this case, it is observed that G3P follows the initial declining trend and the increasing trend by 2010; however, the seasonal changes that are detected with in-situ groundwater levels are not replicated by the G3P product. The limited number of in-situ monitoring stations that were used in this case might not be representing the behaviour of the groundwater in the area that was selected for evaluation and might miss specific groundwater storage changes in this heterogenous aquifer. The presence of confined units within the aquifer might also be altering the GWLA calculation and cannot be directly compared to GWSA. The more data is available, the better the area can be represented, and an improved evaluation could be carried out. In the annexes, Section 10.9, GWSA_LA_hyd and GWSA_LA_grav are plotted against the in-situ GWLA.

Table 9. Pearson r between in-situ GWLA and G3P GWSA for all 3 G3P products v1.5 for the Murray Darling Basin, for the period 2002-2016

Product	Pearson r
GWSA	0.36
GWSA_LA_hyd	0.54
GWSA_LA_grav	-0.28

6.10 Great Artesian Basin

The Great Artesian Basin is a confined groundwater basin in Australia and is one of the largest artesian basins in the world. It has an area of approximately 1.7 million of km², covering most of Queensland, and parts of New South Wales, South Australia, and the Northern. Most of the basin is covering arid and semi-arid regions, and only some areas in the north have tropical seasonal rainfall. Since its discovery in 1878, it has been exploited in benefit of pastoral industry (Habermehl, 2020). See Figure 31.

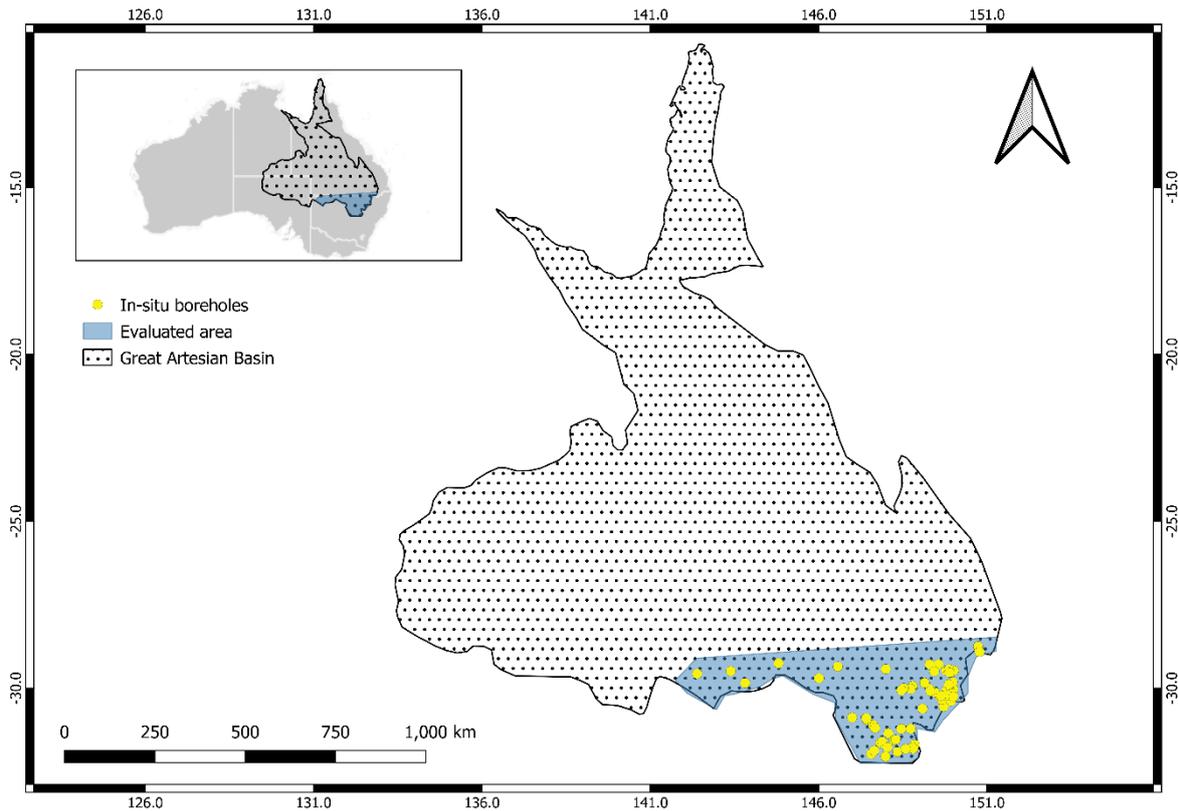


Figure 31. Great Artesian Basin, in-situ boreholes and evaluated area for the period 2010-2016

Data for the evaluation was obtained from the Australian Groundwater Explorer (Bureau of Meteorology, 2022). There are around 5120 available boreholes with groundwater level data from around 1970 onwards. Due to the temporal sparsity of the data, and to comply with the temporal G3P data availability, 62 boreholes were selected to be further used in the evaluation. These boreholes had the longest period (from 2010 to 2016) with a complete groundwater level data set. These boreholes and the area to be evaluated can be observed in Figure 31. From this area, the area average G3P GWSA is calculated and presented in Figure 32. It is observed that the GWSA starts with a declining tendency for 2010, but after 2010 it has an increasing tendency until 2012, where it stabilizes and remains above zero until 2015, where it remains stable around zero until the end of the period.

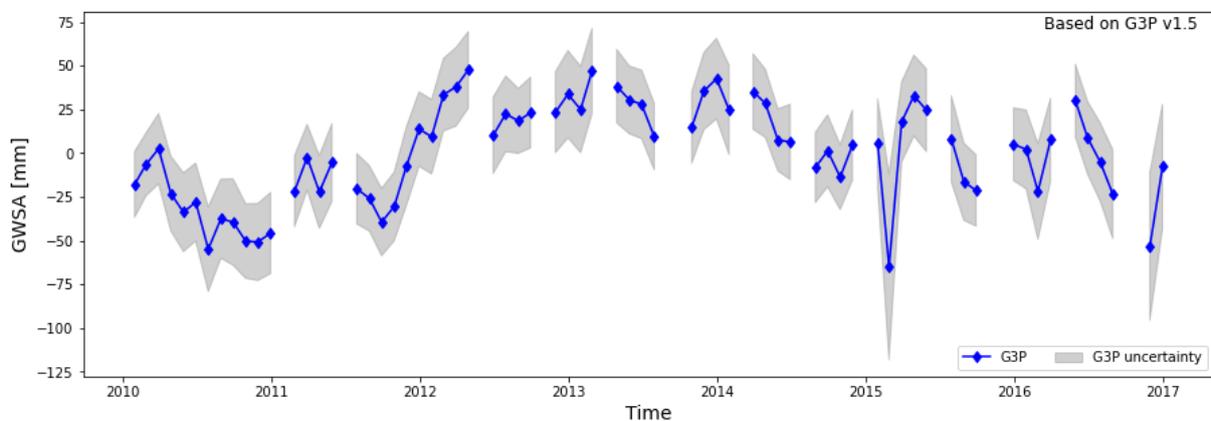


Figure 32. Area average G3P GWSA of the Great Artesian Basin for the period 2010-2016

Storage data was not available in this area; therefore, the in-situ GWLA were calculated and compared to the G3P GWSA. Both time series were standardized and plotted as observed in Figure 33.

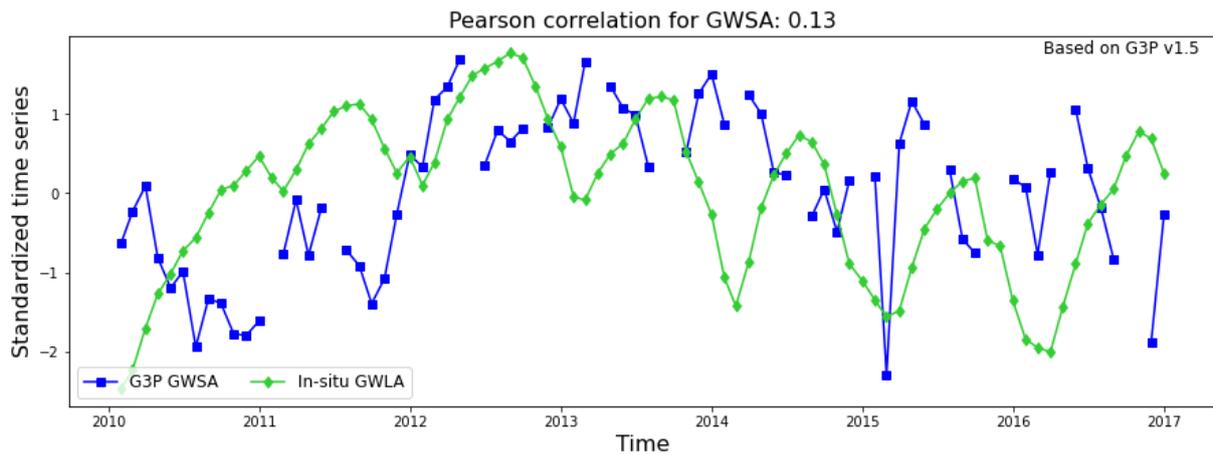


Figure 33. Standardized in-situ GWLA and G3P GWSA in the Great Artesian Basin for the period 2010-2016

Table 10 shows the correlation coefficients between all 3 G3P products and in-situ GWLA for the Great Artesian Basin. Pearson r is low, with a value of 0.13 for GWSA, and for the rest of the products with G3P version 1.5 it remains low, being 0.3 for GWSA_LA_hyd and -0.04 for GWSA_LA_grav. In this case it is observed that the in-situ GWLA have a different behaviour than the G3P GWSA. In-situ GWLA indicates an increasing trend from 2010 to 2013 and a mild decreasing trend after 2013. GWP GWSA is not able to replicate this overall trend, nor the seasonal trends that are observed every year. The fact that this is a confined aquifer, might interfere with the interpretation of the in-situ GWLA since its signals are not influenced by the specific storage variable. Additionally, the arid to semi-arid climate at the location might indicate local recharge points within the aquifer that is not reflected by the limited number of in-situ observations. At this point, a correlation between these two variables could not be conveyed; however, if storage data is available, a direct comparison between in-situ and G3P GWSA needs to be made and analysed. In the annexes, Section 10.10, GWSA_LA_hyd and GWSA_LA_grav are plotted against the in-situ GWLA.

Table 10. Pearson r between in-situ GWLA and G3P GWSA for all 3 G3P products v1.5 for the Great Artesian Basin, for the period 2010-2016

Product	Pearson r
GWSA	0.13
GWSA_LA_hyd	0.3
GWSA_LA_grav	-0.04

6.11 South of outer Himalayas aquifer

The south of outer Himalayas aquifer is a transboundary aquifer located in the north of India in the states of Uttar Pradesh and Bihar, and the southern part of Nepal. It has an area of approximately 310.000 km². The area is underlain by crystalline rocks, schistose, meta-sedimentaries, and volcanic rocks with little porosity; therefore, groundwater recharge and flow occur within the areas where secondary porosity like fractures, weaker planes, among others exists. The main source of groundwater in the region is glaciers and overland flow from rainfall in foothill areas (Gupta, 2014). See Figure 34.

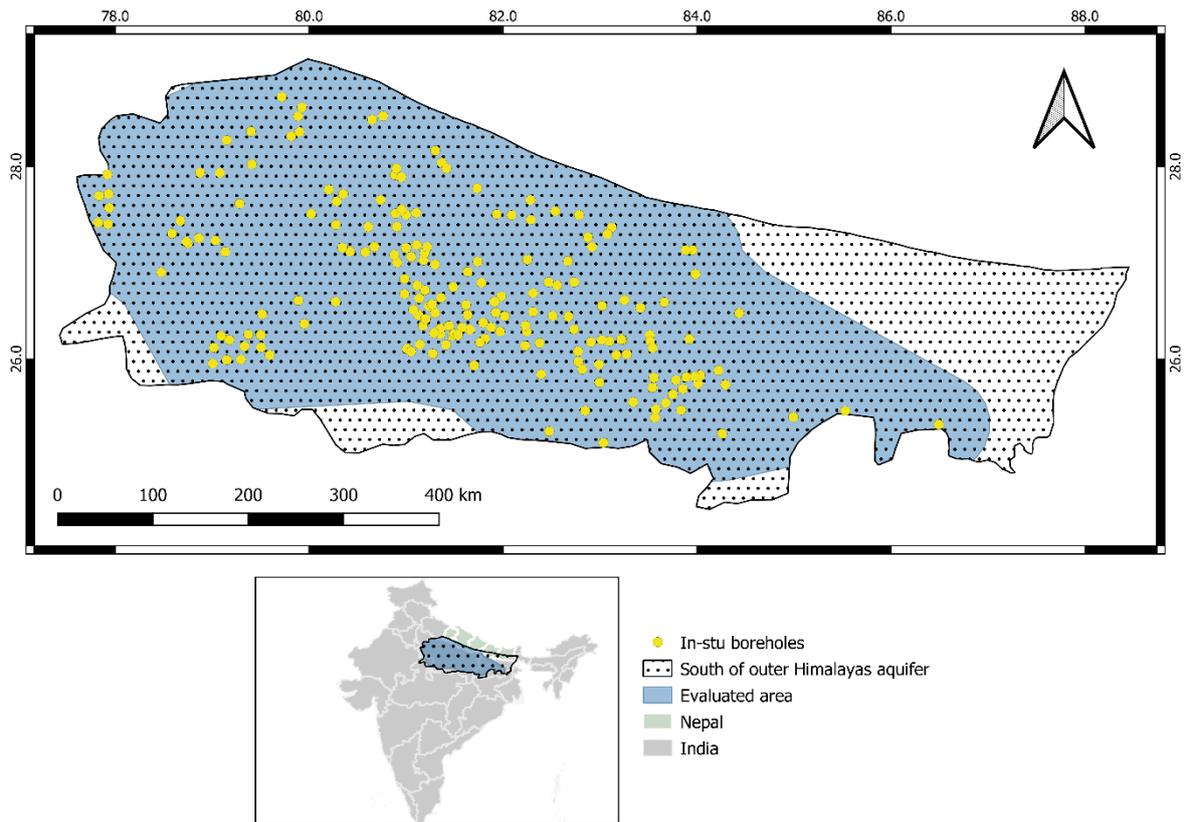


Figure 34. South of outer Himalayas aquifer, in-situ boreholes and evaluated area for the period 2002-2007

Data for the evaluation was obtained from the India Water Resources Information System (National Water Informatics Centre, 2022), managed by the Central Ground Water Board (CGWB). There are around 2758 available boreholes with groundwater level data from around 1994 onwards. The data is collected four times every year, namely: January, May, August, and November. Once the seasonal data is evaluated, and to comply with the temporal G3P data availability, 276 boreholes were selected to be further used in the evaluation. These boreholes had the longest period (from 2002 to 2007) with a complete seasonal groundwater level data set. These boreholes and the area to be evaluated can be observed in Figure 34. From this area, the area average G3P GWSA is calculated and presented in Figure 35. It is observed that the uncertainty values associated to GWSA are large. In this area, it might be worth reviewing the water compartments that contribute to the large uncertainty that is presented, such as snow or glaciers.

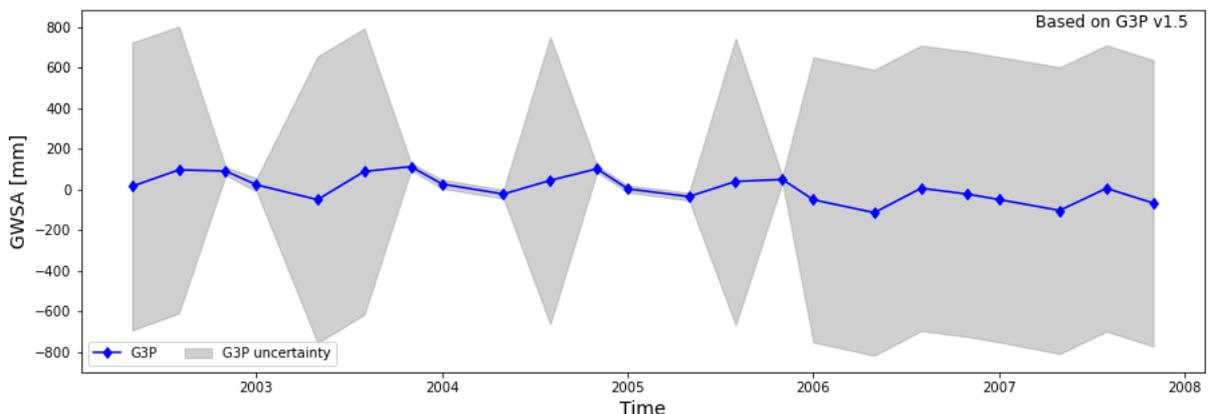


Figure 35. Area average G3P GWSA of the South of outer Himalayas aquifer for the period 2002-2007

It is important to note that the uncertainty is greater than in the previously described aquifers. And as in previous aquifer evaluations, it is not possible to access aquifer storage data to calculate in-situ GWSA. Therefore, the evaluation that is carried out in this case using GWLA needs to be critically interpreted. Figure 36 shows both standardized in-situ GWLA and G3P GWSA time series.

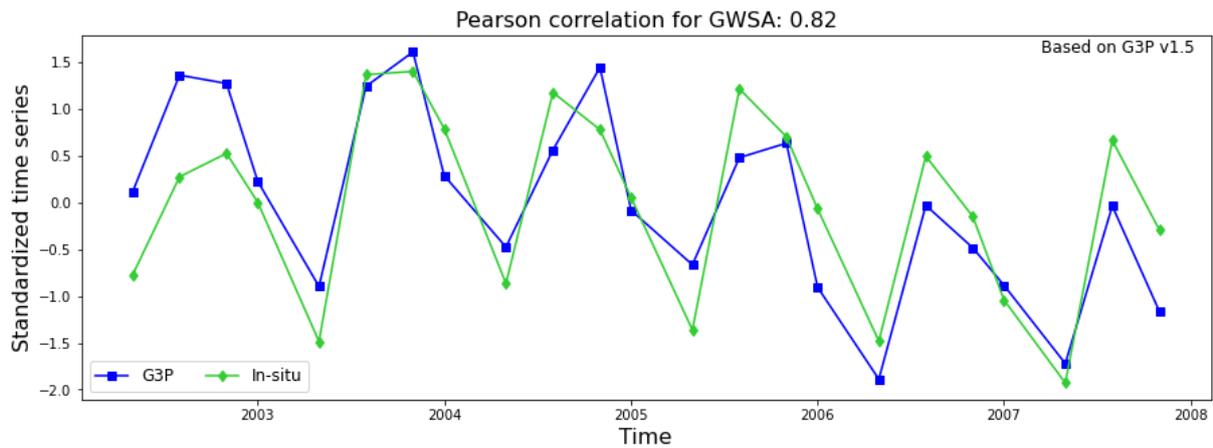


Figure 36. Standardized in-situ GWLA and G3P GWSA in the South of outer Himalayas aquifer for the period 2002-2007

Table 11 shows the correlation coefficients between all 3 G3P products and in-situ GWLA for the south of outer Himalayas aquifer. Pearson r is high, with a value of 0.82 for GWSA, and for the rest of the products with G3P version 1.5 it remains relatively high, being 0.92 for GWSA_LA_hyd and 0.56 for GWSA_LA_grav. It is observed that G3P could replicate well the signal that is calculated using the available in-situ groundwater level measurements. The trend is relatively stable around zero and the seasonal changes observed with in-situ measurements are well replicated by G3P. Despite the similarities between the in-situ GWLA and G3P GWSA, the uncertainty needs to be addressed every time G3P in this area is used. In the annexes, Section 10.8, GWSA_LA_hyd and GWSA_LA_grav are plotted against the in-situ GWLA.

Table 11. Pearson r between in-situ GWLA and G3P GWSA for all 3 G3P products v1.5 for the south of outer Himalayas aquifer, for the period 2002-2007

Product	Pearson r
GWSA	0.82
GWSA_LA_hyd	0.92
GWSA_LA_grav	0.56

6.12 Guarani Aquifer System

The Guarani Aquifer System is a transboundary aquifer of approximately 1.2 million of km². It is one of the largest in the world and it covers four countries in Latin America: Argentina, Brazil, Uruguay and Paraguay. It is formed by a sequence of sandy layers deposited in continental, aeolian, fluvial and lagoon environments, above a regional erosional surface and below an extensive layer of Cretaceous basalts. Part of the Guarani Aquifer System is confined by underlying and overlying deposits. There are also outcropping areas where the aquifer becomes confined or semi-confined (Gonçalves et al., 2020). See Figure 37.

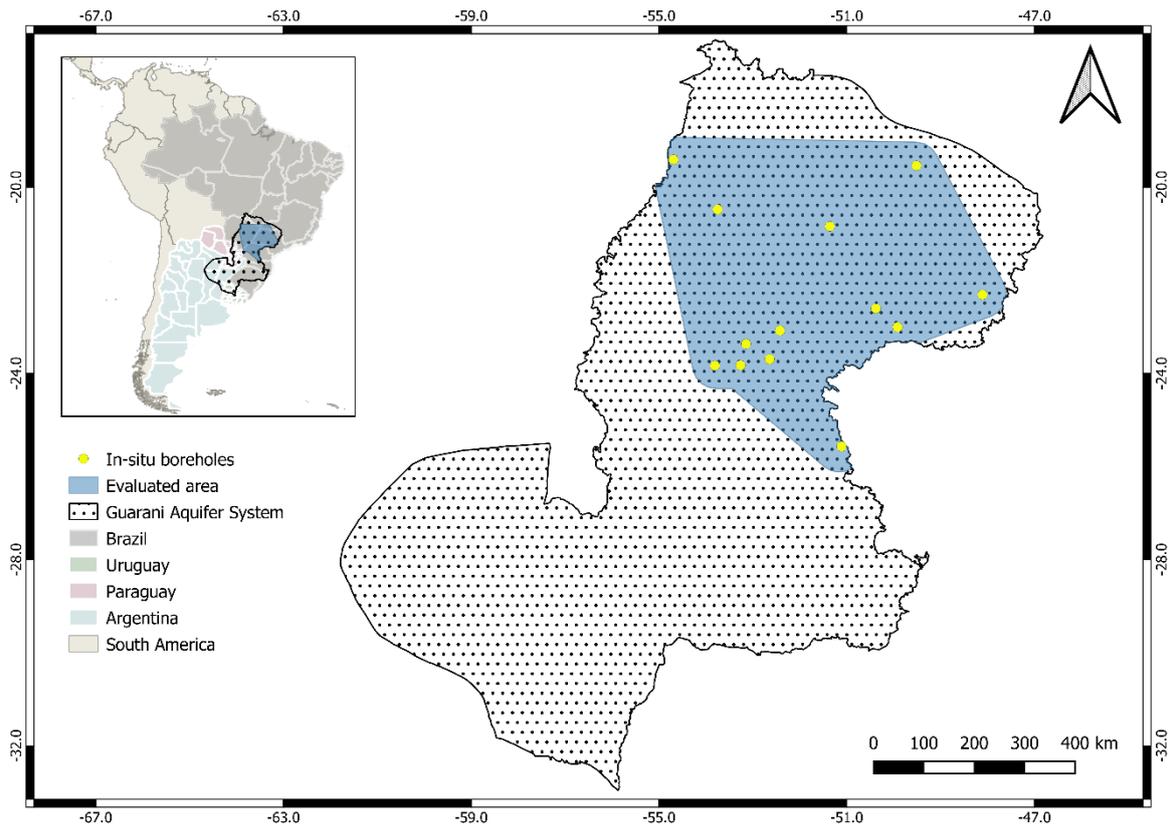


Figure 37. Guarani Aquifer System, in-situ boreholes and evaluated area for the period 2012-2016

Data for the evaluation was obtained from the Integrated Groundwater Monitoring Network (RIMAS), managed by the Geological Service of Brazil (CPRM or SGB) (Geological Service of Brazil, 2022). Within the aquifer, there are 76 available boreholes with groundwater level data from around 2010 to 2019. Due to the temporal sparsity of the data, and to comply with the temporal G3P data availability, 13 boreholes were selected to be further used in the evaluation. These boreholes had the longest period (from 2012 to 2016) with a complete groundwater level data set. These boreholes and the area to be evaluated can be observed in Figure 37. From this area, the area average G3P GWSA is calculated and presented in Figure 38. It is observed that for this specific period there are some missing values. However, it can be observed that at the beginning of the period the GWSA are stable around zero until 2015. After 2015 there is a sudden increase in GWSA, with an equally rapid decrease by the end of 2016.

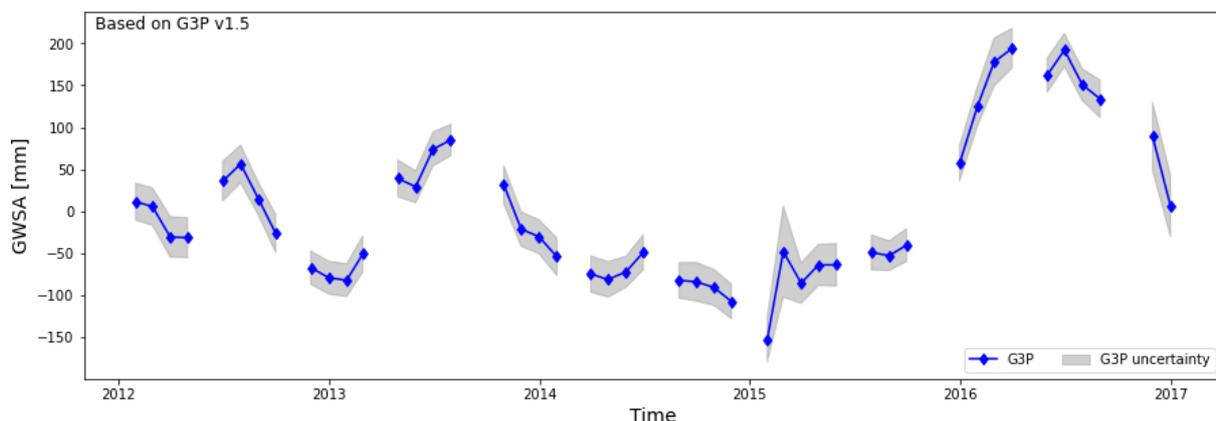


Figure 38. Area average G3P GWSA of the Guarani Aquifer System for the period 2012-2016

Storage data was not available in this area; therefore, the in-situ GWLA was calculated and compared to the G3P GWSA. Both time series were standardized and plotted as observed in Figure 39.

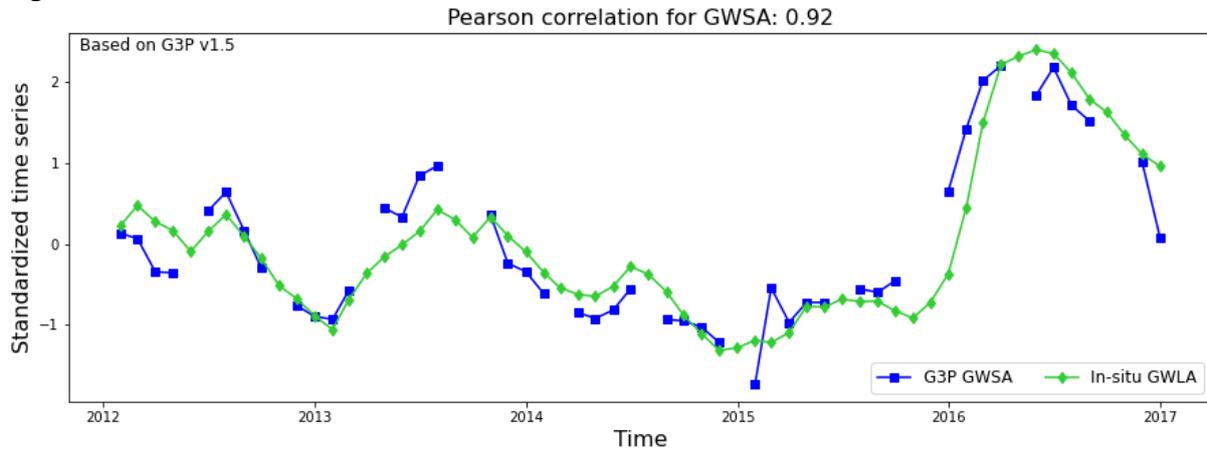


Figure 39. Standardized in-situ GWLA and G3P GWSA in the Guarani Aquifer System for the period 2012-2016

As it is also evident in Figure 39, Table 12 shows the correlation coefficients between all 3 G3P products and in-situ GWLA for the Guarani Aquifer system. Pearson r is high, with a value of 0.92 for GWSA, and for the rest of the products with G3P version 1.5 it remains relatively high, being 0.90 for GWSA_LA_hyd and 0.68 for GWSA_LA_grav. It is observed that G3P could replicate well the signal that is observed from in-situ groundwater level measurements. The stable trend around zero until 2015 and the sudden increase and decrease in by the end of 2015 and 2016, respectively, is also captured by G3P. It is important, however, to address the limited number of in-situ boreholes that were used to calculate GWLA as well as the number of missing G3P GWSA values within the evaluated period. The trend seems to be replicated by G3P, but more data is needed eliminate the uncertainties related to the limitations. In the annexes, Section 10.8, GWSA_LA_hyd and GWSA_LA_grav are plotted against the in-situ GWLA.

Table 12. Pearson r between in-situ GWLA and G3P GWSA for all 3 G3P products v1.5 for the Guarani Aquifer system, for the period 2012-2016

Product	Pearson r
GWSA	0.92
GWSA_LA_hyd	0.90
GWSA_LA_grav	0.68

6.13 Maranhão Aquifer System

The Maranhão Aquifer System is a large aquifer located in the states of Pará, Maranhão, Piauí, Tocantins, and Ceará, in the northeast of Brazil. It has an area of approximately 600.000 km² and it is formed by unconsolidated sedimentary deposits and contains confined and unconfined aquifers (de Sousa, 2000). See Figure 40.

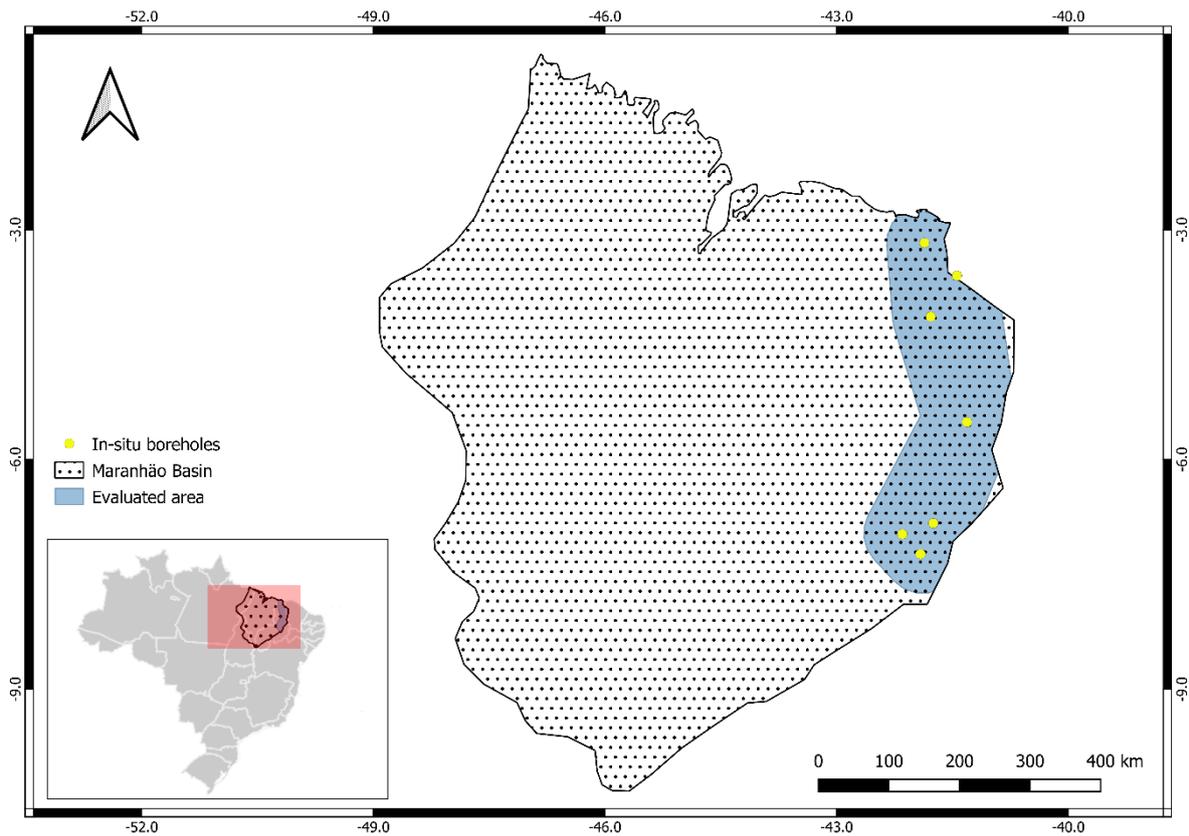


Figure 40. Maranhão Aquifer System, in-situ boreholes and evaluated area for the period 2014-2016

Data for the evaluation was obtained from the Integrated Groundwater Monitoring Network (RIMAS), managed by the Geological Service of Brazil (CPRM or SGB) (Geological Service of Brazil, 2022). Within the aquifer, there are 30 available boreholes with groundwater level data from around 2013 to 2019. Due to the temporal sparsity of the data, and to comply with the temporal G3P data availability, 7 boreholes were selected to be further used in the evaluation. These boreholes had the longest period (from 2014 to 2016) with a complete groundwater level data set. These boreholes and the area to be evaluated can be observed in Figure 40. From this area, the area average G3P GWSA is calculated and presented in Figure 41. It is observed that for this specific period there are several missing values. However, it can be observed that at the beginning of the period the GWSA is at surplus, being above zero until mid-2015. After by 09-2015 there is a decrease in GWSA, and it remains below zero until the end of 2016.

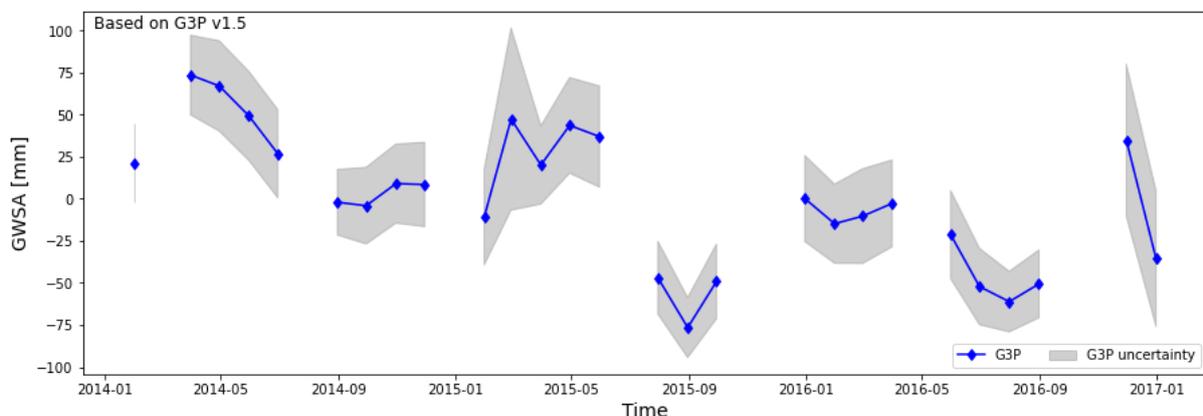


Figure 41. Area average G3P GWSA of the Maranhão Aquifer System for the period 2014-2016

Storage data was not available in this area; therefore, the in-situ GWLA were calculated and compared to the G3P GWSA. Both time series were standardized and plotted as observed in Figure 42.

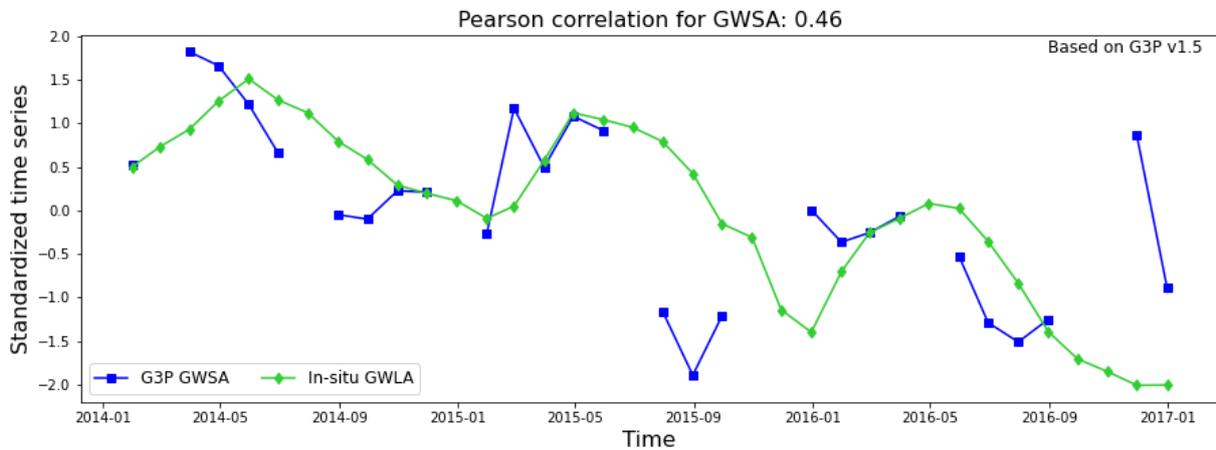


Figure 42. Standardized in-situ GWLA and G3P GWSA in the Maranhão Aquifer System for the period 2014-2016

As it is also evident in Figure 42, Table 13 shows the correlation coefficients between all 3 G3P products and in-situ GWLA for the Maranhão Aquifer system. Pearson r is 0.46 for GWSA, and for the rest of the products with G3P version 1.5 it remains around the same range, being 0.63 for GWSA_LA_hyd and 0.54 for GWSA_LA_grav. In this case it is observed that G3P follows the initial declining trend throughout the period; however, the seasonal changes that are detected with in-situ groundwater levels are not replicated by the G3P product. It is also important to address the fact that a limited number of in-situ boreholes were used to calculate GWLA, the period of analysis is short, and there are a relatively high number of missing G3P GWSA values within the evaluated period. The overall trend seems to be detected by G3P, but more data is needed to eliminate the uncertainties related to the mentioned limitations. In the annexes, Section 10.13, GWSA_LA_hyd and GWSA_LA_grav are plotted against the in-situ GWLA.

Table 13. Pearson r between in-situ GWLA and G3P GWSA for all 3 G3P products v1.5 for the Maranhão Aquifer System, for the period 2014-2016

Product	Pearson r
GWSA	0.46
GWSA_LA_hyd	0.63
GWSA_LA_grav	0.54

7. Case study in Spain

This case study aims to evaluate the overall performance of the G3P product at the pixel level by comparing spatiotemporal patterns of G3P groundwater storage anomalies against observed ones retrieved from ground-based measurements in continental Spain. The performance was also assessed with respect to extreme hydrometeorological events that have occurred in the region during the period covered by the GRACE Follow and Follow-On missions.

7.1 Methodology

7.1.1 Study region

The study area covers continental Spain, which spans over an area of approximately 500,000 km². The region is characterized by its large climate diversity, where up to 5 climate Köppen-Geiger zones exist. Hot-summer (*Csa*) and warm-summer (*Csb*) Mediterranean climates mostly dominate the entire country along the South and Eastern Mediterranean coastline as well as the central plains located in the inlands, whereas Oceanic climate (*Cfb*) prevails in the North and North-West coastline, and warm-summer Continental (*Dfb*) governs the North-Eastern sectors. The mean annual precipitation is around 680 mm, but it varies strongly in time and space¹. Meteorological drought periods are very frequent, while extreme floods usually happen during the autumn period along the Mediterranean coastline fringe as result of convective rainfall events.

There are 14 major river basins in continental Spain, out of which three are shared with Portugal: Duero, Tajo, and Guadiana. Management of surface waters and groundwater are committed to the River Basin Authorities.

A total of 638 groundwater bodies (GWB) have been established in continental Spain (Figure 43, Figure 44). GWB is concept established at the European level only for management purposes. Usually, these GWBs refer to a unique hydrogeological unit or aquifer. However, in other cases especially in complex hydrogeological contexts, a GWB can include multilayer aquifers with different hydrogeological properties.

¹ https://en.wikipedia.org/wiki/Climate_of_Spain

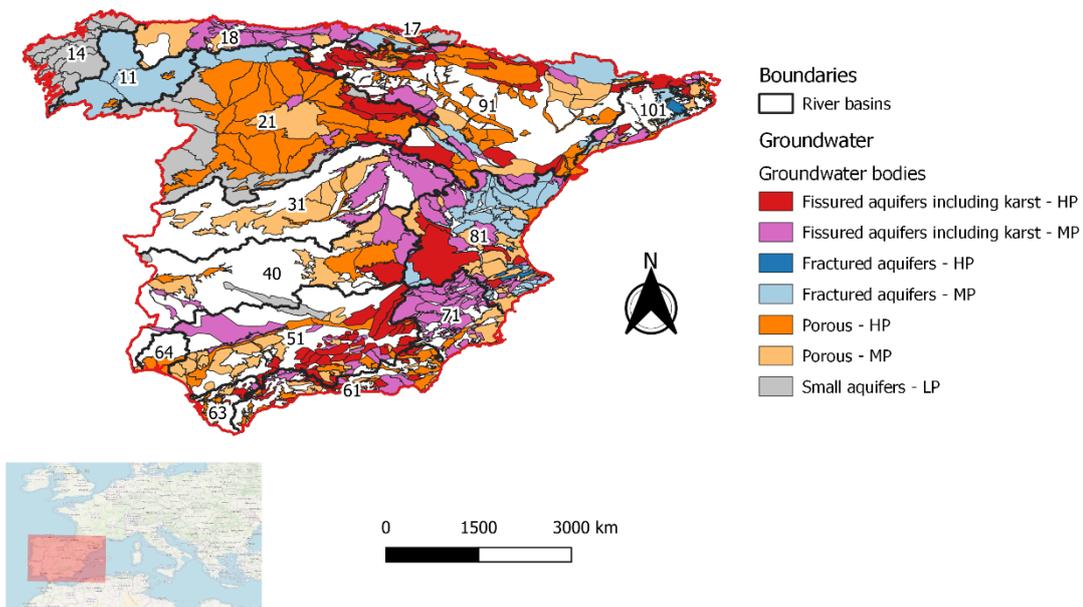


Figure 43. Distribution of Groundwater Bodies (GWB) in continental Spain. GWB are classified according to its dominant aquifer typology and groundwater productivity (HP: high productive, MP: Medium productive, LP: low productivity). Numbers correspond to the ID of each River Basin District.

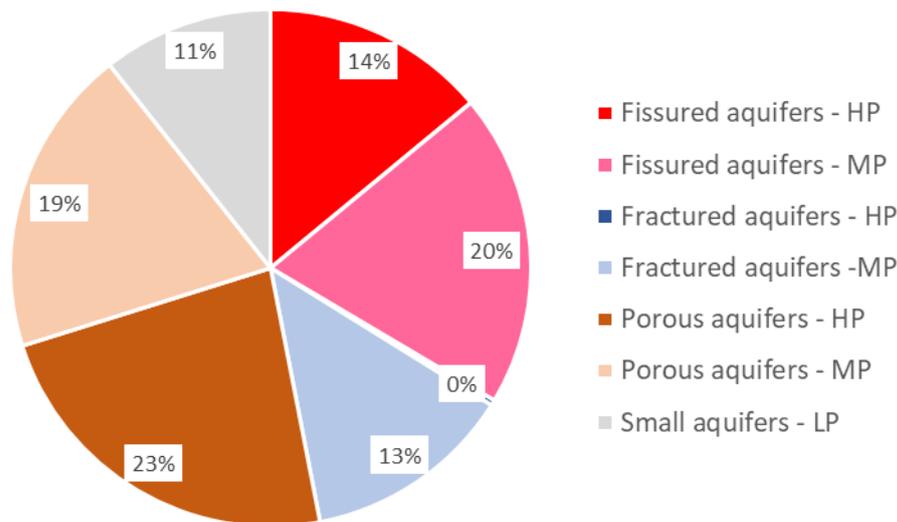


Figure 44. Percentage contribution of each GWB category in continental Spain

The study region covers the continental Spain. In order to simplify and ease the quick location of the pixels, the region was divided into quadrants of 1 deg. and 0.5 deg. following the grid fishnet used for GRACE products (Figure 45). Each 1 deg. quadrant was named following a two capital letters which symbolize the row and column position in the grid, while the 0.05 quadrants are named with a 2-number digits according to its relative position inside the 1 deg. quadrant (11 for the upper-left, and 22 for the lower-right positions).

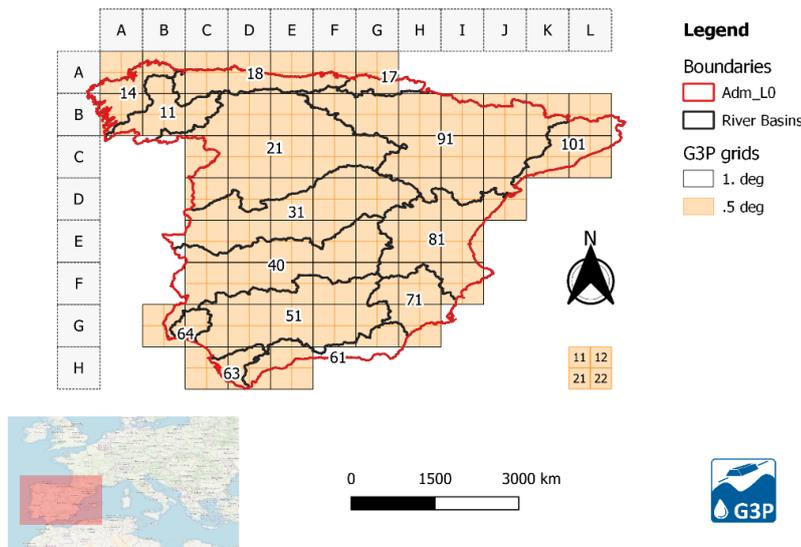


Figure 45. Division of the study area by GRACE quadrants of 1 deg and 0.5 deg. River Basin districts (from N-S): 14: Galicia-Costa, 18: Cantábrico Occidental, 17: Cantábrico Oriental, 11: Miño-Sil, 21: Duero, 91: Duero, 101: Cuencas Internas de Cataluña, 31: Guadiana, 81: Júcar, 40: Guadiana, 64: Tinto, Odiel y Piedras, 51: Guadalquivir, 71: Segura, 63: Guadalete y Barbate, 61: Cuenca Mediterránea Andaluza.

7.1.2 In-situ data collection

A large dataset of groundwater level (GWL) measurements has been collected from the Spanish National Water Monitoring System². The dataset includes 679,000 daily measurements in the 1965-2019 period collected in 3285 boreholes. For the purposes of this study, a monthly sub dataset was generated covering the period from Jan 2002 to Dec 2019. Monthly GWL values were computed for each borehole as the average of all the daily measurements available. Only boreholes with more than 150 computed monthly observations were qualified for the study. In total, 707 boreholes were finally qualified, most of them located in the Ebro (27%), Guadiana (18%) and Guadalquivir (18%) river basins.

Table 14. Size of the native groundwater level dataset and subset for this analysis. Only basins with boreholes qualified for this study are shown

Basin	Native dataset		Sub dataset	
	Nº boreholes	Meas. in 2002-2019	Nº boreholes*	Meas. in 2002-2019
Ebro	418	280629	195	35480
Guadiana	471	64348	129	24460
Guadalquivir	322	95100	126	22364
Júcar	358	88145	102	18335
Tajo	222	33099	81	13564
Segura	274	29439	31	4961
Cuencas Mediterráneas Andaluzas	147	16833	19	2873
Duero	549	40452	11	1663
Miño-Sil	23	2583	7	1122
Cantábrico oriental	31	2998	6	1061
TOTAL	2815	307,645	707	125,883

* only boreholes with more than 150 monthly measurements in the 2020-2019 period.

² <https://sig.mapama.gob.es/redes-seguimiento/>

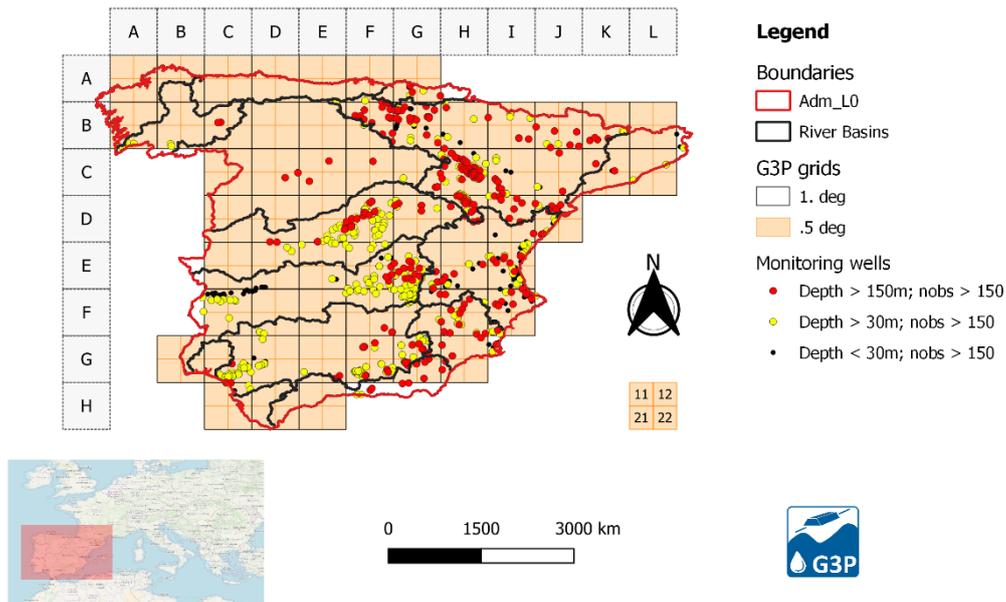


Figure 46. Distribution of qualified boreholes classified according to their depths of penetration

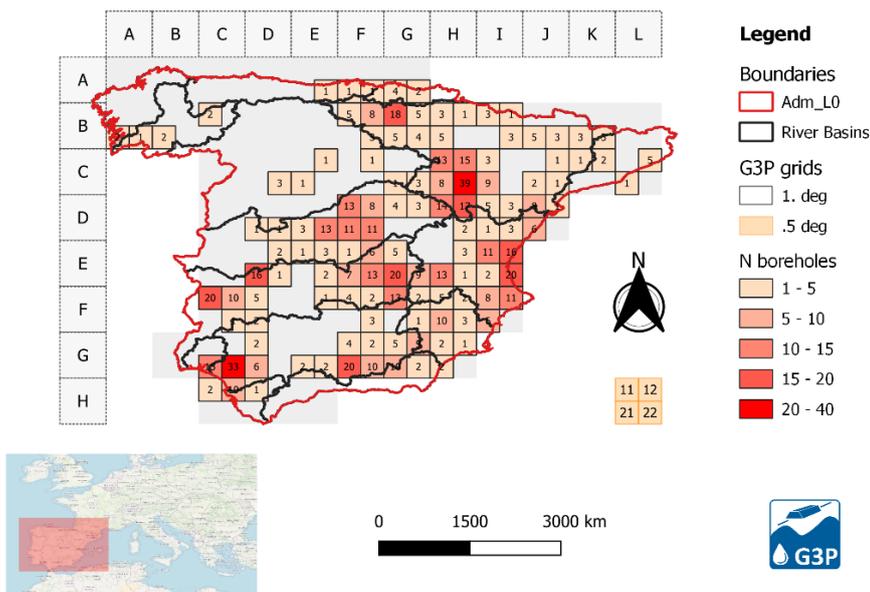


Figure 47. Number of qualified monitoring wells at the 0.5 deg. pixel level. A qualified well is that one with at least 150 monthly observations in the 2002-2019 period.

7.1.3 Preprocessing and retrieval of the GroundWater Index Anomaly (GWI)

Raw GWL values at each borehole were processed in order to retrieve an anomaly index called the GroundWater Index (GWI). GWI was computed as:

$$GWD_{bid,t} = GWL_{bid,t} - GWL_{bid,t-1} \tag{Eq. 5}$$

$$GWI_{bid,t} = \frac{GWD_{bid,t} - \text{mean}(GWD_{bid})}{\text{stdev}(GWD_{bid})} \tag{Eq. 6}$$

where, $GWD_{bid,t}$ is the deviation of the groundwater level measured at borehole *bid* at the time *t* in relation with its value at time *t-1*, and *GWI* is the Z-score value of *GWD* at time *t* relative to the mean and standard deviation observed in the period of analysis (2002-2019).

In a second step, values of *GWI* at the pixel level and each timestep were computed by averaging all the *GWI* values previously retrieved in all the boreholes located in each pixel. Two averaging methods were adopted. The first method applied a simple arithmetic average from all borehole values inside the pixel. Resulting *GWI* is termed as *mGWI*. The second method consisted in a two-step procedure where firstly simple arithmetic averages were retrieved for each groundwater body located in a pixel, and then these were weight-averaged according to the relative areal contribution of each *GWB* to the total area of the pixel. Resulting *GWI* was termed as *wGWI*.

Before being used in further analyses, gaps in resulting *GWI* timeseries were filled using a linear interpolation method.

7.1.4 Extreme conditions

The increasing severity and frequency of climate-induced disasters in the last two decades are also reflected in different remotely sensed products. Datasets providing satellite-based indicators of hydro-meteorological variables should, in principle, capture better the occurrence of such extreme events in comparison with non-extreme conditions. For example, very high temperatures may be related with heatwaves, long dry spells with no or very low precipitation may be related with severe droughts, while the combination of high precipitation and the presence of water on the landscape may be related with floods. In this analysis, the performance of the G3P product was also tested in its ability to detect very wet and dry spell conditions.

It is hypothesized that groundwater anomalies over a 12-month period should be higher and lower during the driest and wettest years of a period than the ones observed during normal-rainfall years.

The analysis was performed at one 1-deg. pixel, in that one for which the Pearson correlation between *mGWI*, *GWSA* and Standardized Precipitation Index (*SPI12*) values was found to be the highest³. Then, the extreme years were defined by those time windows where the highest and lowest *SPI12* values were observed. In order to capture the potential delays in groundwater response to precipitation, the 1 year period was set up for starting 9 months before and ending 3 months after the date when *SPI12* was highest/lowest. To establish the reference or normal-rainfall period, two conditions were set up: i) *SPI12* must be higher than -1.0 and lower than 1.0 (thresholds that determines the range of precipitation values that are around the average), and ii) the time window must be the same as the one used for the extreme period. Differences between normal and extreme conditions were visualized using boxplots for the *mGWI* and the *G3P-GWSA*.

³ The *SPI* is a well-known meteorological drought index that informs about the anomaly of precipitation observed in a over a 12-month aggregated timescale (WMO and GWP, 2016). *SPI* ranges between -3.1 (driest condition) to 3.1 (wettest conditions). "Normal" rainfall conditions range from -1.0 to 1.0.

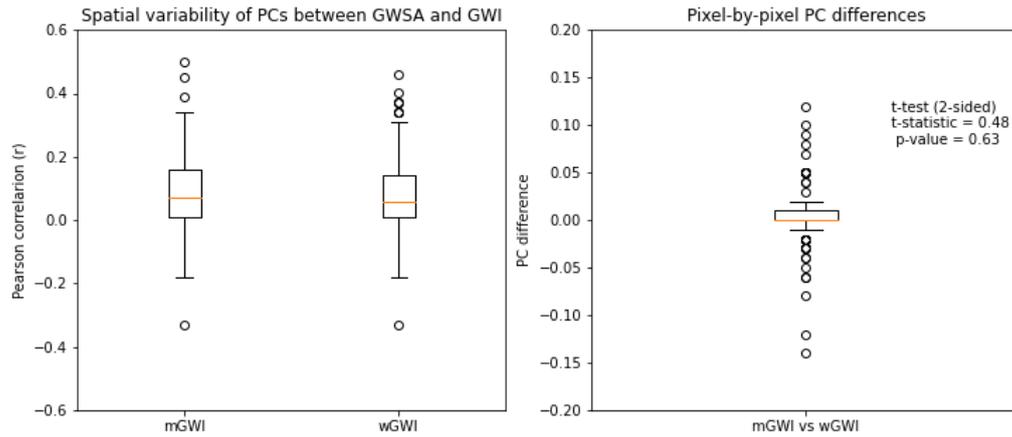


Figure 48. Right: Boxplots of pixel-by-pixel Pearson correlations coefficients between G3P-GWSA and GWI values retrieved using simple-arithmetic (mGWI) and weighted (wGWI) averaging methods. Left: Boxplot with Pearson correlation differences between mGWI and wGWI

7.1.5 Pearson correlations coefficients

To test how closely the G3P-GWSA and in-situ GWI were linked, Pearson correlations coefficients -PCC- were derived using the timeseries (2002 – 2019) for all pixels with qualified monitoring wells. As shown in Figure 47, the higher PCC were found along the Ebro valley and headwaters of Tajo basin. Highest PCC, with a value of 0.6, was found at pixel BH21. In general, PCC seems to decrease towards the southern Spain. This spatial pattern may be justified by the hydrogeological properties of the region, with a major dominance of shallow detritic aquifers in the Ebro basin, and deeper fissured and karst aquifers in the S-SE of Spain. Higher correlations in the Ebro may be also explained by its nival regime which would contribute to imprint a higher inertia in the surface-hydrogeological system. This would not be the case in drylands regions at S-SE Spain where groundwater dynamics may be affected by other disturbance effects that break the potential connection between the surface and groundwater compartments.

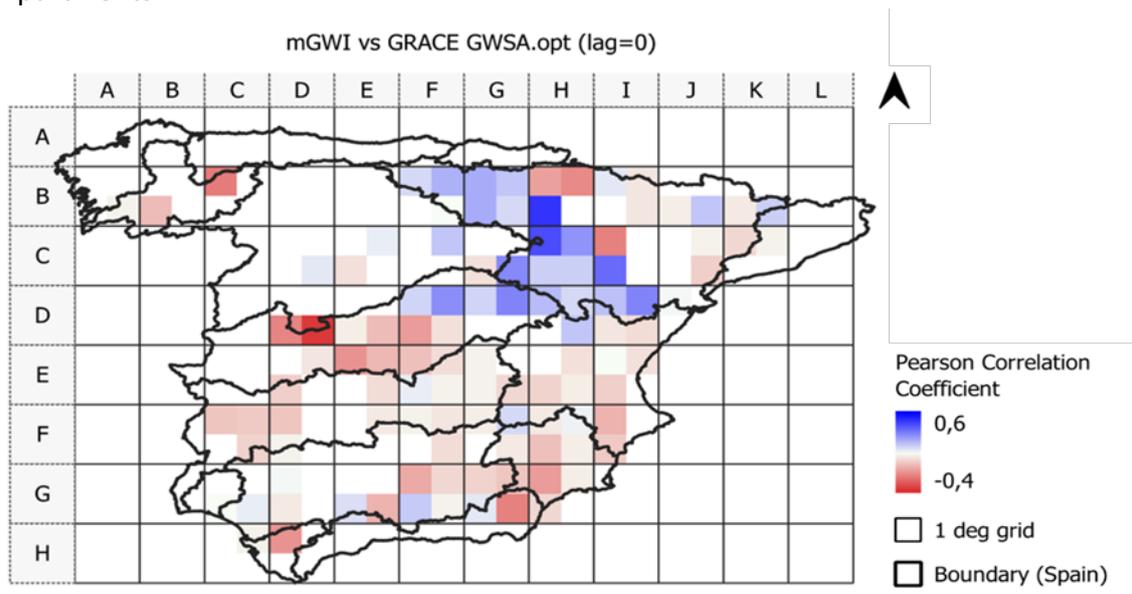


Figure 49. Pearson cross correlations between mGWI and G3P-GWSA at 0.5 deg. resolution

Variation in response time also impacts the correlations between GWSA and mGWI. The response time is strongly governed by the local hydrogeological conditions such as hydraulic conductivity, porosity, piezometric levels and so on. An analysis at 0.5 degree, therefore,

cannot account for the spatial heterogeneity in the region and comes with certain limitations, foremost being the density and distribution of groundwater monitoring network. While G3P-GWSA are derived from remote observations and hence offer global coverage, the GWI 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.

The inverse relationship between in-situ observations and GRACE-G3P groundwater storage anomalies could also be a result of anthropogenic activities as abrupt declines in groundwater levels were observed in some of the monitoring wells. Since the G3P data only provides estimates at a monthly resolution, these changes in groundwater levels are not accounted for – unlike in the in-situ measurements which were aggregated based on daily observations to compute the groundwater index.

7.1.6 Trends

In addition to determining the correlations between in-situ ground observations and GRACE-groundwater storage anomalies, the temporal variation in G3P-GWSA was also analyzed. For this, the GWSA available at a monthly resolution and 1-deg. scale were aggregated for a 12-month timescale and compared with SPI12. As shown in **Error! Reference source not found.**, the GWSA did track reasonably well the SPI12 trajectory, and transitions between the interannual phases of dry and wet conditions were captured. In some periods, e.g., from February 2019 till December 2019, the trend aligned strongly. However, the magnitude was not always fully synchronized; this was especially relevant during some wetter periods with positive SPI values, when the observed anomalies in groundwaters react in the opposite way than the expected after positive gains in precipitation.

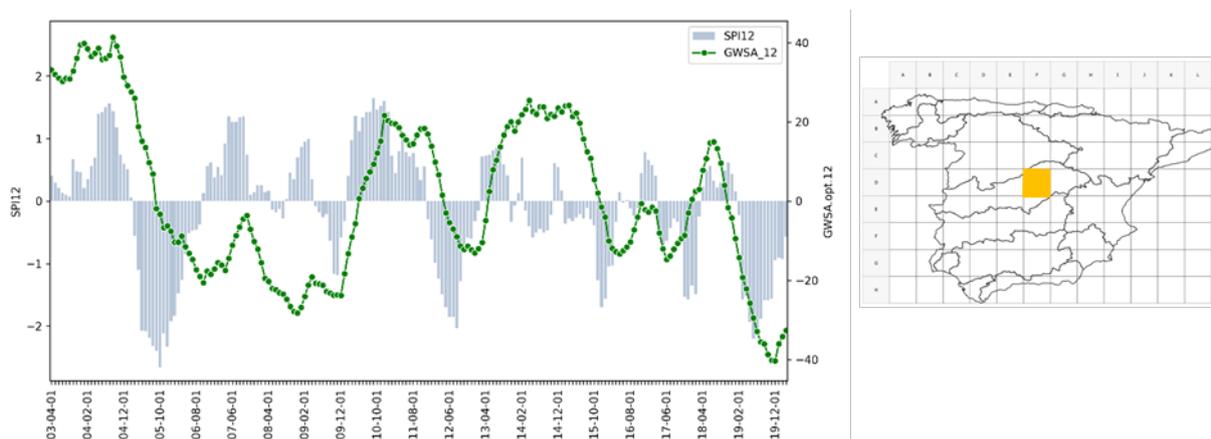


Figure 50. Temporal variation of GWSA against SPI12 at Pixel ID: DF

Given the extent of the study region, as well as the variability within, the trend shown in Figure 47 cannot be considered representative. The trends obtained for other pixels had high variability, and as discussed in detail in the next section, groundwater response times greatly influence the correlations with SPI12

7.1.7 Representation of extreme events

To determine whether in-situ groundwater observations and G3P groundwater storage anomalies were able to reflect extreme events, the following analyses (Figure 48) was generated at a selected pixel.

The longest continuous dry spell was identified based upon SPI12 values. Groundwater anomalies during the dry period were compared against the observed ones during ‘normal-rainfall’ periods, which were taken as a benchmark, when SPI12 values ranged from -1 to 1. Since the highest correlations between mGWI and SPI12, and mGWI and GWSA were found for PixelID CI21, this was selected for further analysis. The longest spell of dry conditions in this region was observed in year 2016, where the mean annual SPI12 value was -1. The ‘normal-rainfall’ benchmark for this pixel was composed of three years: 2006, 2010 and 2011.

As shown in Figure 6, during the longest dry period, the maximum, median and minimum values for GWSA decreased. This reflects the response of groundwater storage to changes in meteorological conditions. However, a similar contrast between the reference and driest period for mGWI was not found. This could be due to the influence of anthropogenic activities on the piezometric levels during this period since the GWI is derived from daily observations.

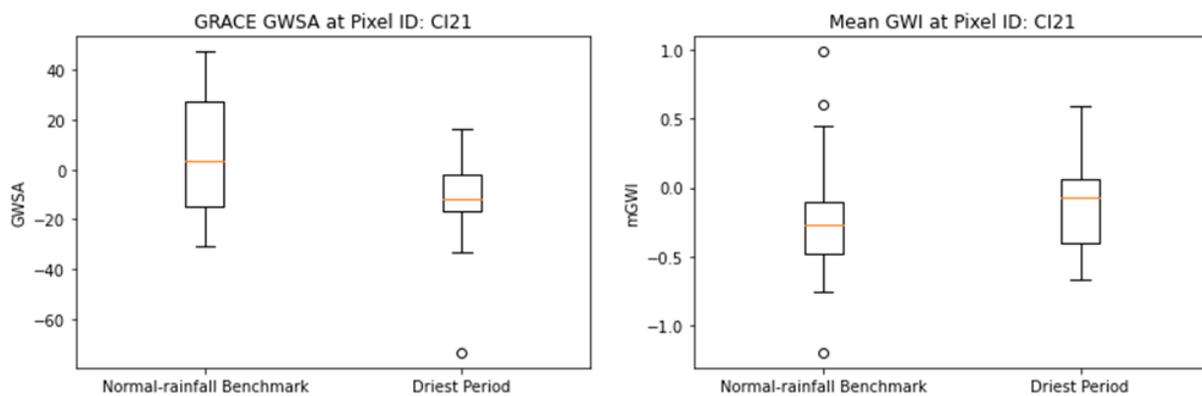


Figure 51. Variation in SPI12, GRACE-GWSA and mGWI between reference and dry period

Similarly, the behavior of GWSA and mGWI was also analyzed under wet conditions which was identified based on highest observed SPI12 values (2.5). The mean SPI12 for the wet period (year 2003) was 1.9. The normal-rainfall benchmark comprised of the same periods as stated earlier.

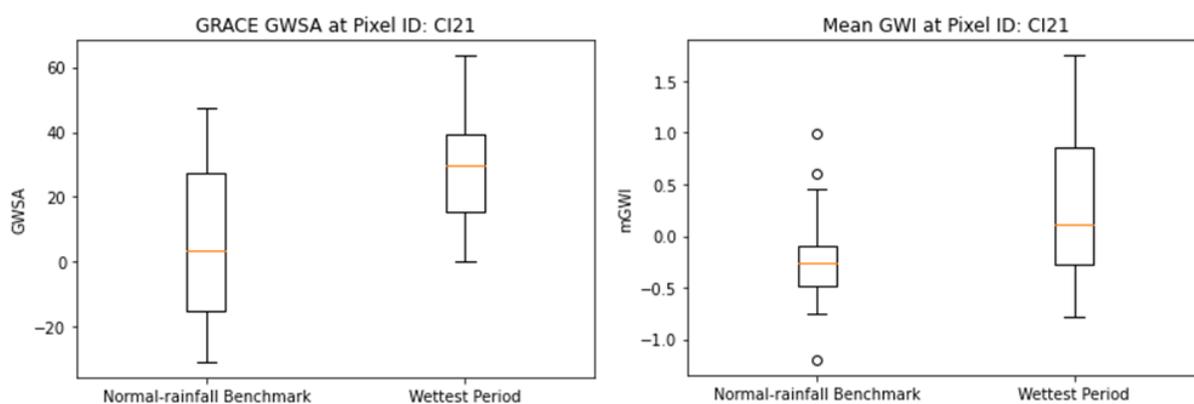


Figure 52. Variation in SPI12, GRACE-GWSA and mGWI between reference and wet period

As shown in Figure 52 above, both GWSA and mGWI exhibited an increase. The median for GWSA increased significantly, with almost no negative anomalies recorded in this period while for mGWI, there was a larger variation, with a slightly higher median than the reference period. This suggests potential groundwater recharge during the wet period, which is reflected in both GRACE-GWSA and mGWI datasets.

8. Discussion

The evaluation of the G3P product was carried out for a previously defined set of aquifers around the world. Originally, 23 aquifers were selected to be evaluated, but due to the limited availability of groundwater data in these areas, 10 aquifers could not be evaluated and included in this report. Therefore, the results that are shown and discussed here are based on the 13 aquifers that could be evaluated. The pixel-based assessment was carried out using groundwater level data distributed across continental Spain.

For the large-scale evaluation, the data availability varied from aquifer to aquifer. Specific yield values were available for the Ogallala aquifer and the North China aquifer system. These were the only two cases where in-situ GWSA was available and compared directly to G3P GWSA. In both cases, the correlation was relatively high. For the rest of the aquifers, specific yield data was not available, and the comparisons were made using in-situ GWLA. Although it is not possible to make a direct comparison between GWSA and GWLA, the signal of both variables was evaluated and served as a proxy to determine the similarities between in-situ observations and the G3P products.

Additionally, the number of boreholes per aquifer was also varied. The maximum number of boreholes used for an evaluation is 310 (Paris Basin), while the minimum number is 7 (Maranhão Basin). This is a critical factor to consider when taken the results into context. Some of the aquifers have a good correlation to the G3P products, but in cases where the number of used boreholes for the evaluation is low, these good correlations need to be carefully taken, and if possible, complemented with other data. The same applies to the distribution of boreholes within the evaluated area. The evaluated area was selected based on the location of the boreholes, but the correlation needs to be carefully assessed in cases where the boreholes are clustered in specific locations within the polygon.

The evaluation of the aquifers resulted in varied correlation coefficients. For G3P v1.5, there are aquifers with high correlation (>0.5) between in-situ observations against the G3P products, namely: Ogallala aquifer, Floridan aquifer, Northern Great Plains, Cambrian-Ordovician aquifer system, Mississippi river valley, Maranhão basin, Guarani aquifer system, North China aquifer system, Paris Basin, and South of Outer Himalayas aquifer. Most are characterized by having seasonal rainfall that recharges the aquifer and by being productive aquifers. Some of them suffer from intensive abstraction (e.g., Northern High Plains) but measures have been taken to overcome this issue.

On the other hand, there are aquifers with low correlation, namely: Basin and range basin-fill aquifers, Great Artesian basin, and the Murray-Darling basin. These aquifers are located in arid to semi-arid zones, and some have dominant confined conditions (e.g., Murray Darling basin). In this case, it is difficult to convey that the G3P products are not representing what is observed in the field, because the low correlation could be associated to uncertainties related to the in-situ data. In addition, in arid areas there is an absence of other surface water compartments such as soil moisture, glaciers, or snow. The uncertainties related to these compartments were not evaluated in this report, but these water compartments might be under or overestimated in these specific areas, and therefore, the contribution of each

compartment to GWSA and their respective uncertainties need to be checked. At the same time, there are a limited number of boreholes that are representing a large aquifer area. In a heterogenous aquifer, in-situ boreholes that are low in number and poorly spatially distributed, yield in-situ GWSA and GWLA with high uncertainties, which have a limited capacity to capture the intricate dynamics of such aquifers. Given the uncertainty from both in-situ and G3P sides, it is encouraged to evaluate G3P in these type of areas with more data to obtain more reliable results, if possible.

In the case of the Ogallala aquifer, two periods were studied. A longer period but with fewer boreholes covering a smaller area yielded a greater correlation than a shorter period but with more boreholes covering a more extensive area. The aquifer covers a great extension and therefore, the more area is covered, the more heterogeneity is encountered (e.g., confined, and unconfined conditions). This could explain why a smaller area, represented by fewer boreholes might capture the smaller scale changes that are well captured by G3P.

Throughout the G3P project, there have been four versions of the G3P products, namely: v1.1, v1.3, v1.4, and v1.5. An evaluation of the evolution of the G3P product related to the available in-situ observations was made in order to compare the differences between versions of G3P. The results are observed in Table 15.

Table 15: Pearson correlation between in-situ observations and all G3P products for all aquifers and all versions of G3P. Values in yellow are the highest Pearson r per aquifer.

List of evaluated aquifers in Set 2	Pearson r									
	v1.5			v1.4			v1.3			v1.1
	GWSA	LA_hyd	LA_grav	GWSA	LA_hyd	LA_grav	GWSA	LA_hyd	LA_grav	GWSA
Ogallala Aquifer (High Plains) - 2002-2016	0.77	0.79	0.53	0.77	0.55	0.53	0.77	0.78	0.55	0.78
Ogallala Aquifer (High Plains) - 2009-2016	0.52	0.6	0.27	0.52	0.07	0.27	0.52	0.59	0.3	0.6
Basin and range basin-fill aquifers	0.1	0.06	-0.04	0.11	-0.43	-0.04	0.08	0.03	-0.07	0.43
Floridan aquifer	0.83	0.89	0.06	0.84	0.71	0.08	0.84	0.89	0.14	0.88
Northern great plains	0.85	0.84	0.44	0.84	0.77	0.44	0.82	0.82	0.43	0.79
Cambrian-Ordovician aquifer system	0.89	0.9	0.7	0.89	0.69	0.71	0.88	0.88	0.69	0.88
Mississippi river valley	0.79	0.53	0.18	0.79	-0.29	0.18	0.87	0.84	0.67	0.81
Maranhão Basin	0.46	0.63	0.54	0.47	-0.58	0.55	0.47	0.64	0.55	0.8
Guarani aquifer system	0.92	0.9	0.68	0.92	0.68	0.68	0.91	0.9	0.66	0.89
Great Artesian basin	0.139	0.3	-0.04	0.14	-0.32	-0.04	0.15	0.3	-0.04	0.64
Murray-Darling Basin	0.365	0.54	-0.28	0.368	-0.06	-0.27	0.39	0.54	-0.27	0.64
North China Aquifer System (Huang Huai Hai Plain)	0.57	0.59	0.38	0.58	-0.02	0.39	0.52	0.51	0.32	0.67
Paris Basin	0.75	0.72	0.02	0.75	0.18	0.04	0.75	0.72	0.04	0.7
South of outer Himalayas Aquifer	0.82	0.93	0.56	0.82	-0.24	0.56	0.85	0.91	0.55	0.93

It is observed that some of the aquifers have high correlations to the G3P products in both the newer version (v1.5) and the first version (v1.1). The confined aquifers that are located in arid to semi-arid areas show higher correlations in version v1.1 (Basin and range basin-fill aquifers, Great Artesian basin, and the Murray-Darling basin), while correlations using version v1.5 are higher in aquifers that have seasonal rainfall and are considered productive aquifers (Ogallala aquifer, Floridan aquifer, Northern Great Planes, Cambrian-Ordovician aquifer system, Mississippi river valley, Maranhão basin, Guarani aquifer system, North China aquifer system, Paris Basin, and South of Outer Himalayas aquifer). Newer versions of the product with new

versions of the superficial water compartments with their own uncertainties, might have under or overestimated values in areas where they have little to no influence (e.g., arid areas). In these cases, it is recommended to consider the uncertainty related to each compartment and revise the one(s) contributing the most to the large discrepancy. This way, a better understanding of the evaluation results could be reached. Table 15 also shows that GWSA_LA_grav does not have a high correlation coefficient to in-situ data in any aquifer in any version, while v1.4 of GWSA_LA_hyd had the worst correlation coefficients. Most of the aquifers have a high correlation coefficient with GWSA_LA_hyd (except GWSA_LA_hyd v1.4), and the rest with GWSA.

The pixel-based evaluation, showed that the correlation coefficients between in-situ GWI and G3P GWSA are mild to low, being 0.6 the largest correlation coefficient along the Ebro valley and decreasing towards the S-SE where dryland regions are present. As also discussed in the large-scale evaluation, the uncertainties related to the different surface water compartments need to be considered in places where these might have a greater influence. There might be specific pixels that are affected by a particular component (e.g., soil moisture in arid or dry areas). A pixel-based evaluation is limited in terms of representativity. The local variability of hydrogeological factors such as hydraulic conductivity, porosity, piezometric levels are more difficult to be captured at the 0.5-degree scale. Additionally, and as also observed at the large-scale evaluation, the limited number of observation wells and their limited spatial distribution might misrepresent the situation in the ground. Nonetheless, it was found that G3P could capture changes in groundwater storage influenced by meteorological conditions at the pixel scale. An overall increase of GWSA was found under a wet period, while a decrease was found during the longest dry period.

9. Summary

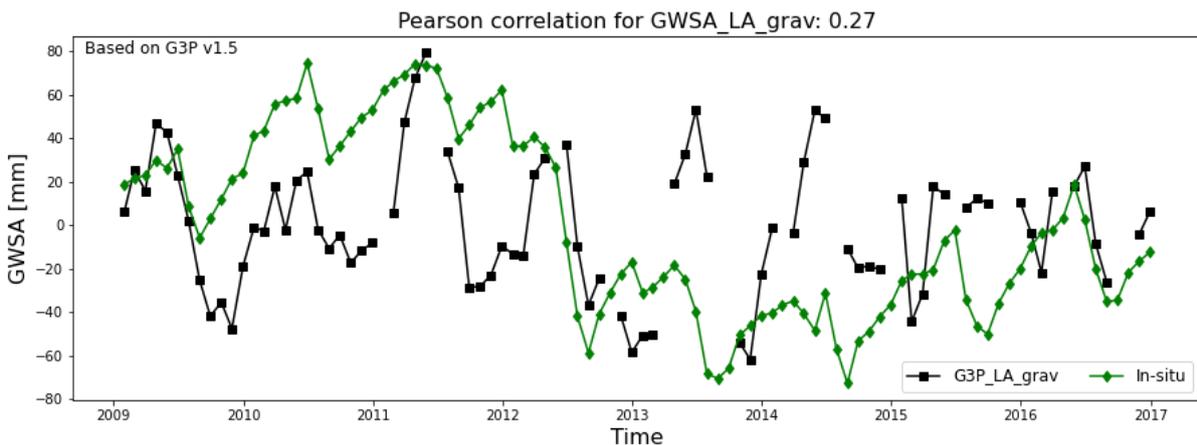
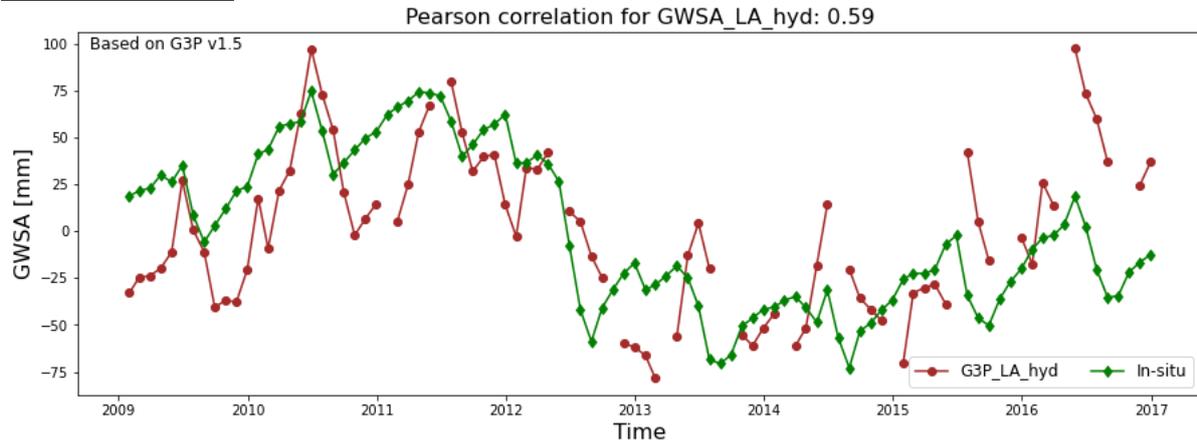
The evaluation process has been done using thirteen pre-selected aquifers around the world and a separate case-study in Spain using a different methodology. It has been observed that most of the evaluated aquifers' in-situ data show a good correlation with the G3P products for version v1.5. These aquifers are mainly productive aquifers and located in areas where seasonal rainfall is common and mostly are the main source of recharge of the aquifer. Three evaluated aquifers did not have in-situ data with a high correlation coefficient with the G3P products. These are located in arid or semi-arid areas where the uncertainties related to the superficial water compartments might be playing an important role. This is also reflected in the pixel-based evaluation, where low correlation coefficients were found towards the south of the country where dryland regions are present. Additionally, the limited number of in-situ observations might be hindering a proper representation of the local heterogeneities of the aquifer recharge and dynamics. Even though the evaluation process found relatively high correlation coefficients at the large scale for most of the evaluated aquifers, these results need to be taken with caution, since the uncertainties related to the calculated in-situ GWSA and GWLA were not quantified at this stage. For the pixel-based evaluation, the correlation between in-situ groundwater data and G3P GWSA was lower than the evaluation performed at the large scale. Ultimately, it is observed that, at this stage, G3P is better representing large aquifer areas and their dynamics in favorable conditions (as discussed above), and a more detailed, pixel-based assessment needs to be accompanied by a thorough evaluation of each of the water compartments and their uncertainties to convey more reliable results.

10. Annex 1. Comparison of G3P GWSA_LA_hyd and GWSA_LA_grav to in-situ GWSA or GWLA

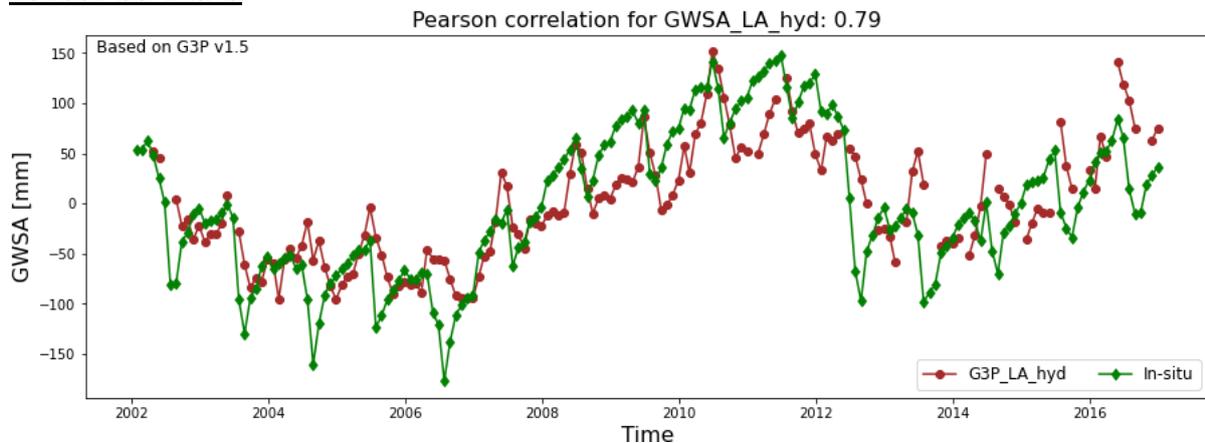
Note: the impact of leakage approximation on uncertainty is not yet clear. Accordingly uncertainties are not displayed in this section.

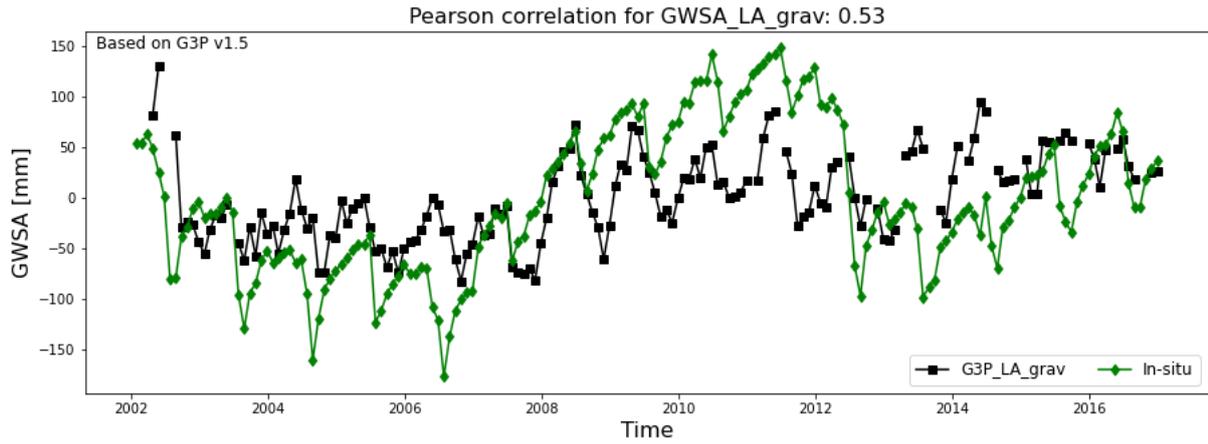
10.1 Ogallala aquifer

Period 2009-2016

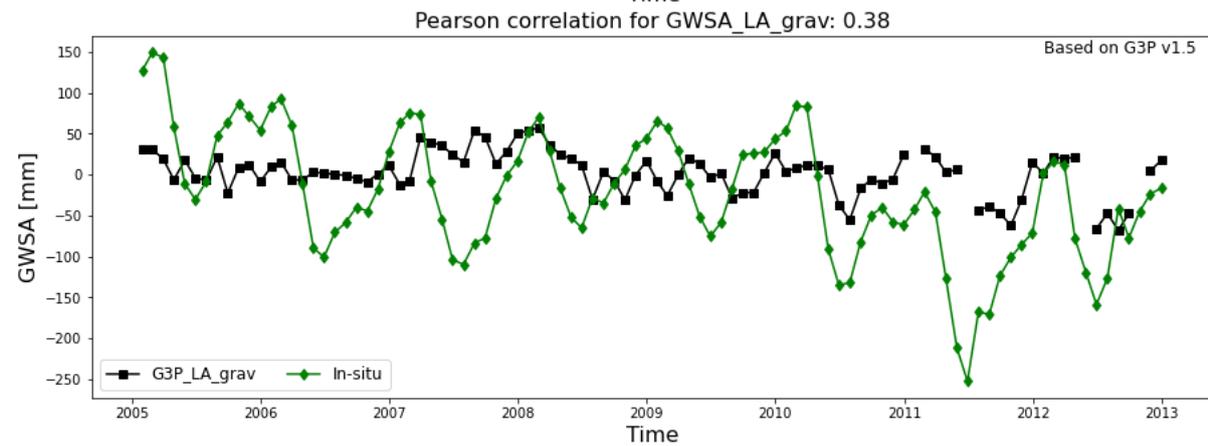
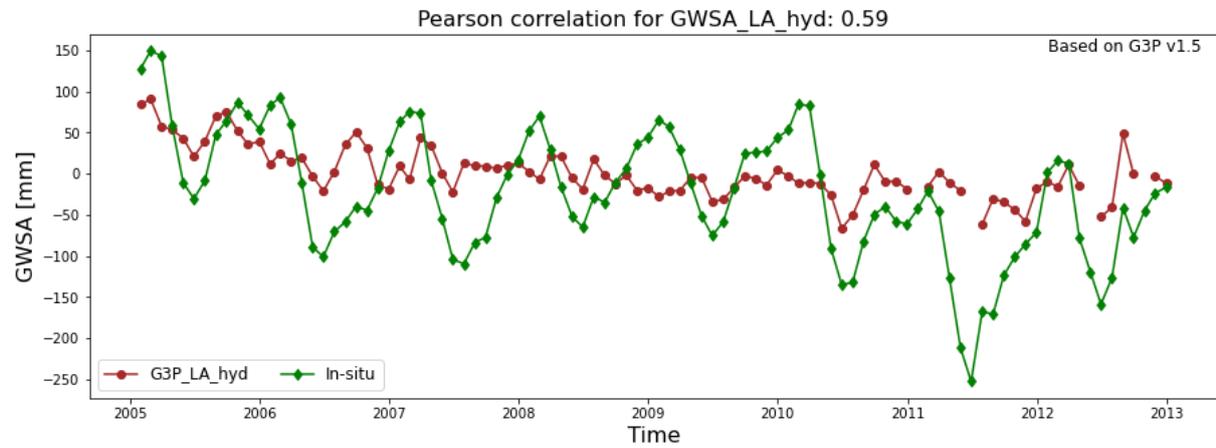


Period 2002-2016

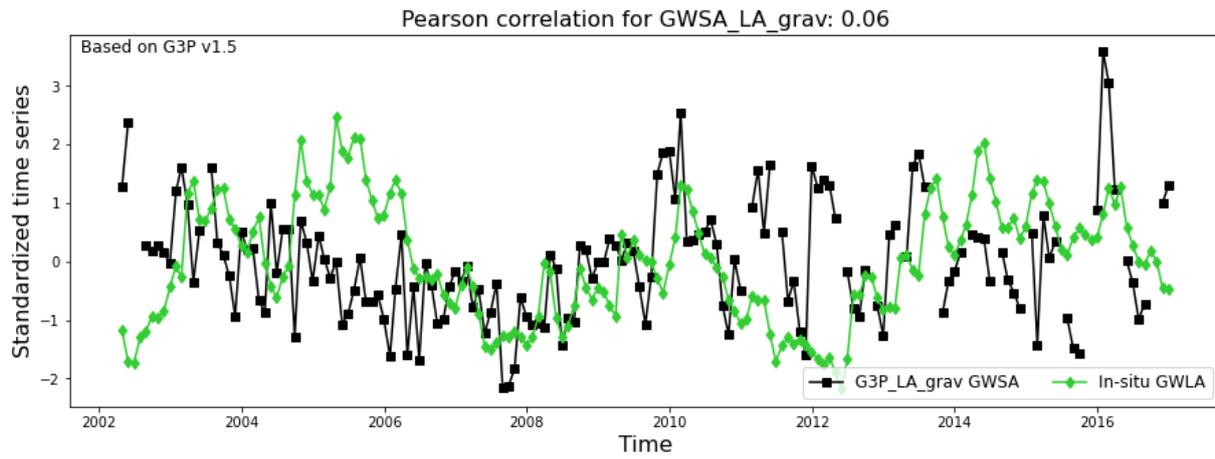
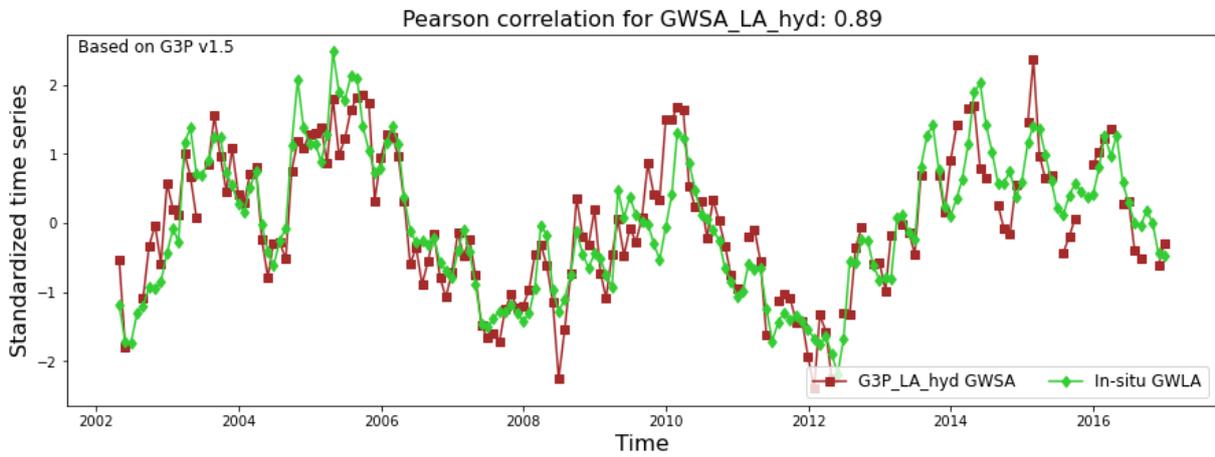




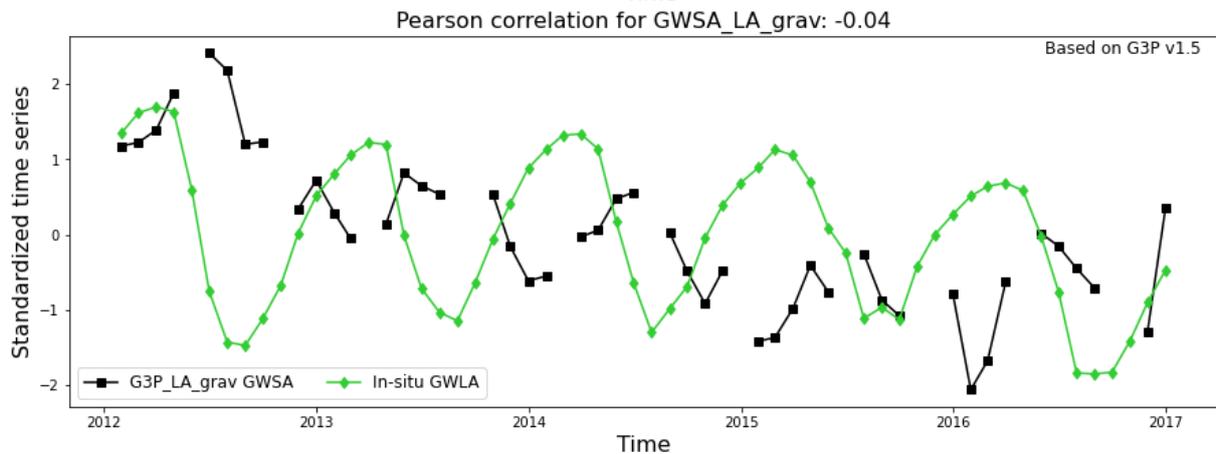
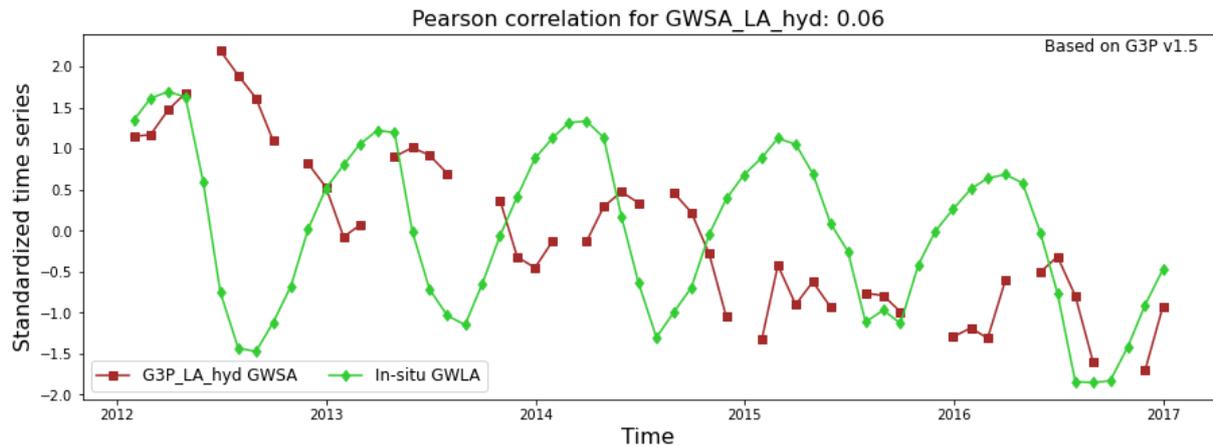
10.2 North China Aquifer system



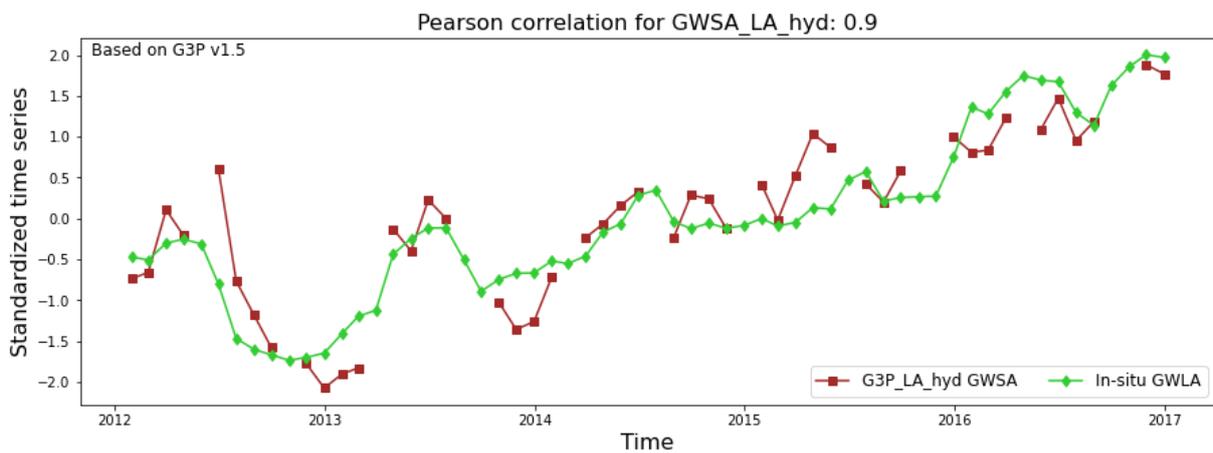
10.3 Floridan Aquifer System

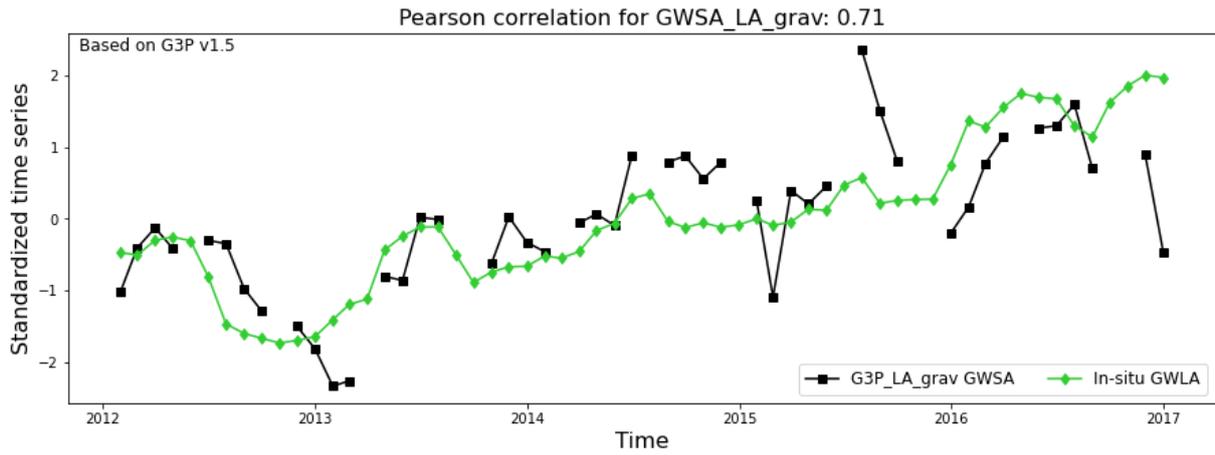


10.4 Basin and Range Aquifers

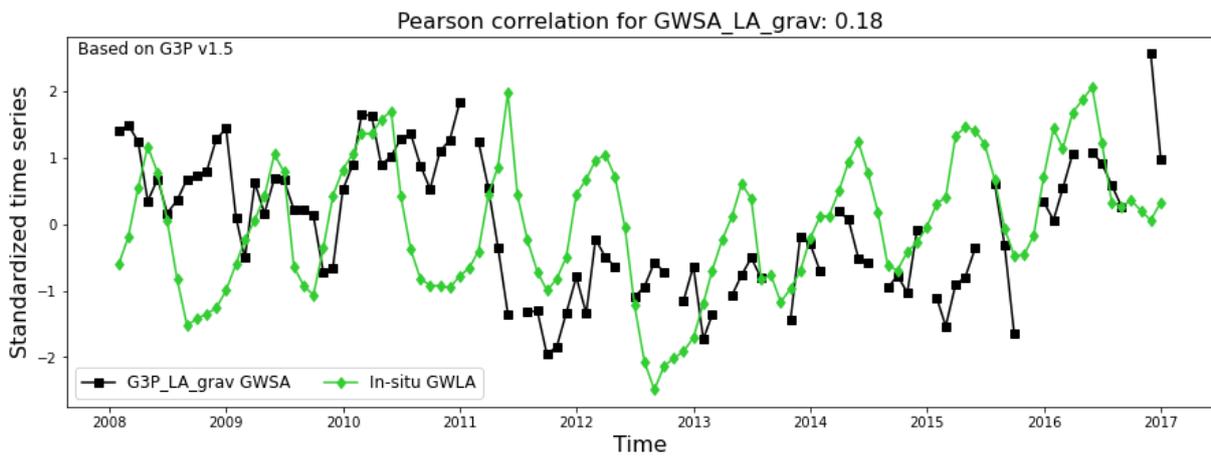
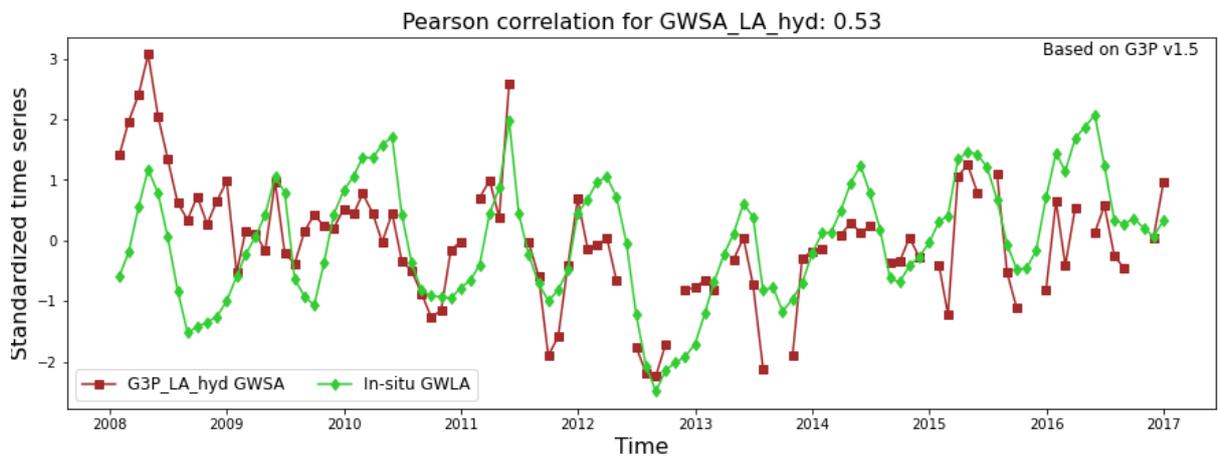


10.5 Cambrian-Ordovician Aquifer System

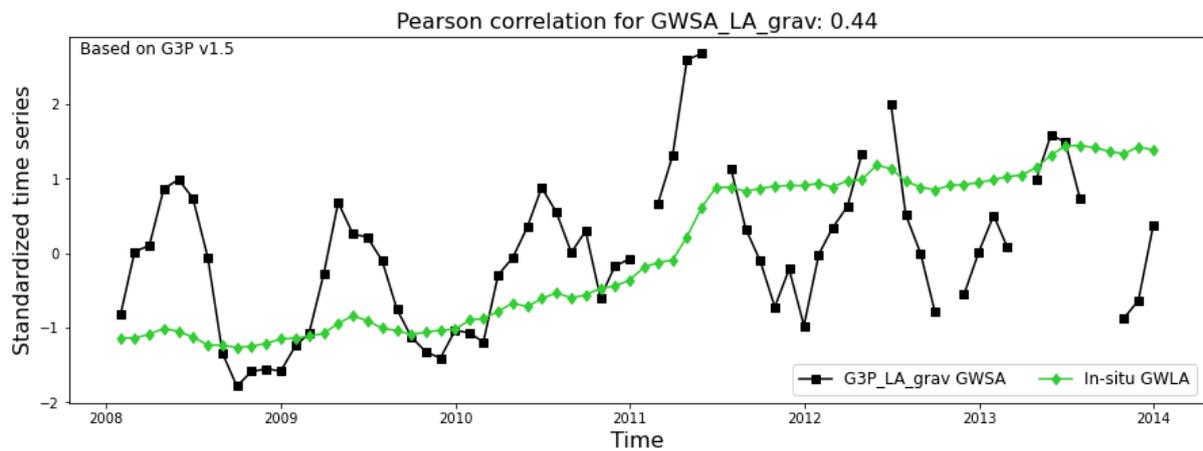
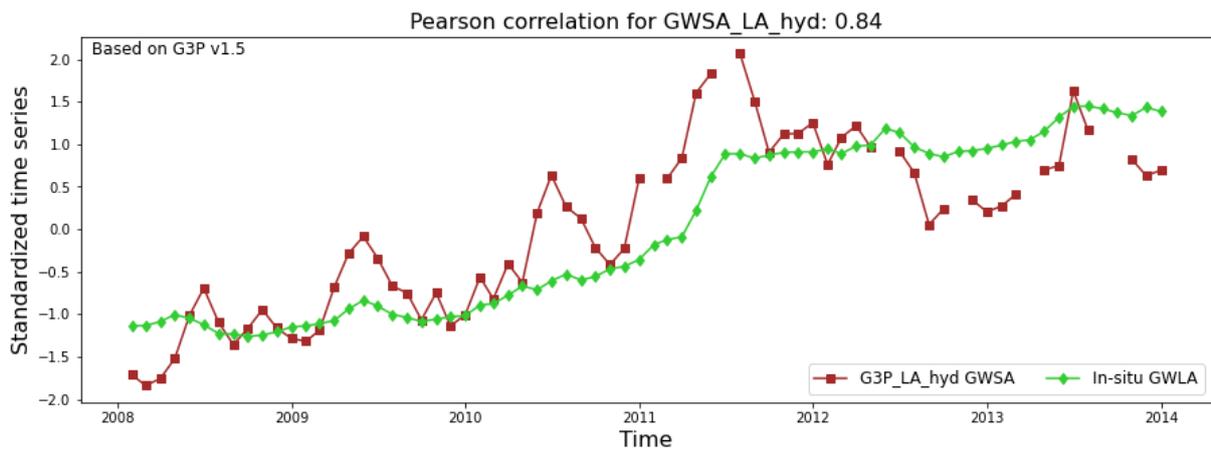




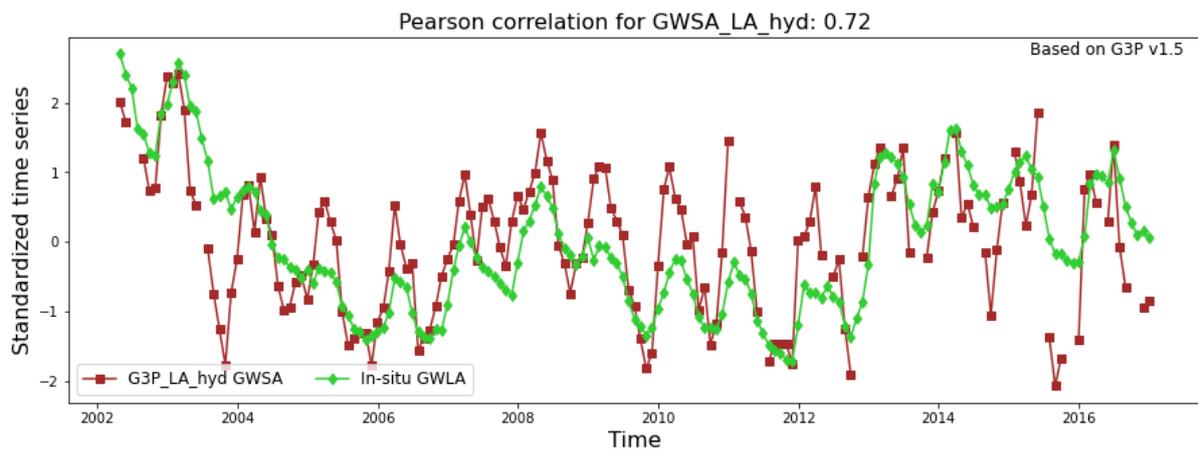
10.6 Mississippi River Valley

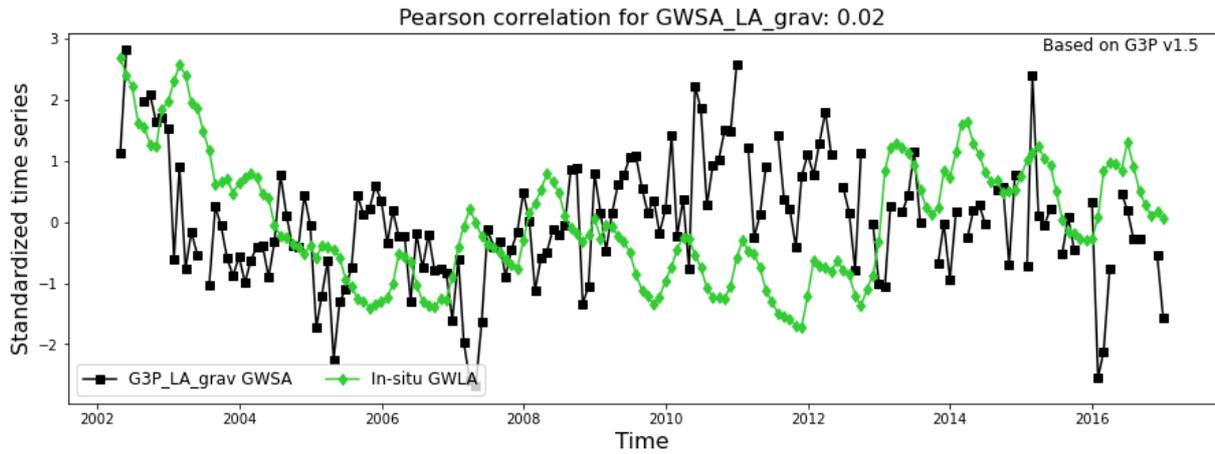


10.7 Northern High Plains

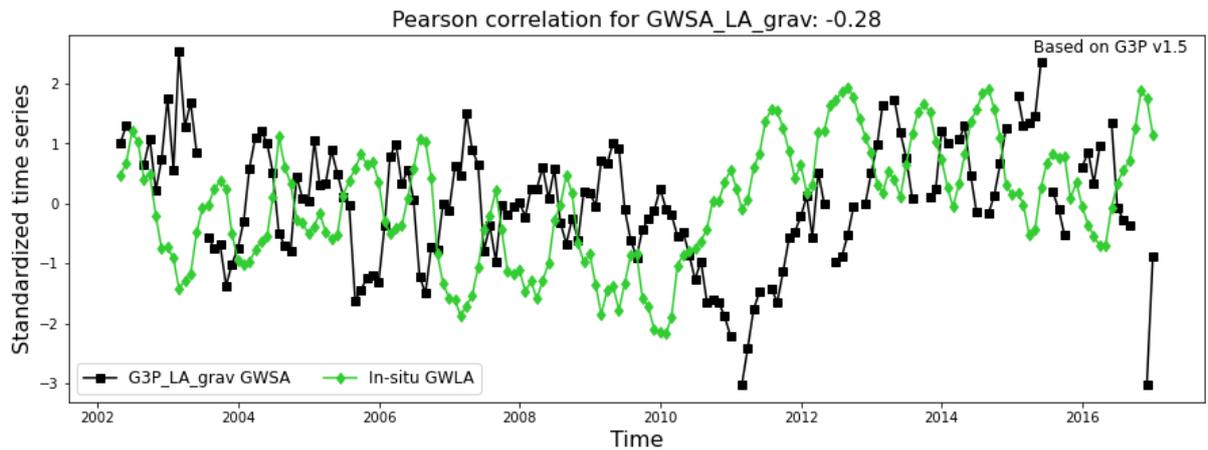
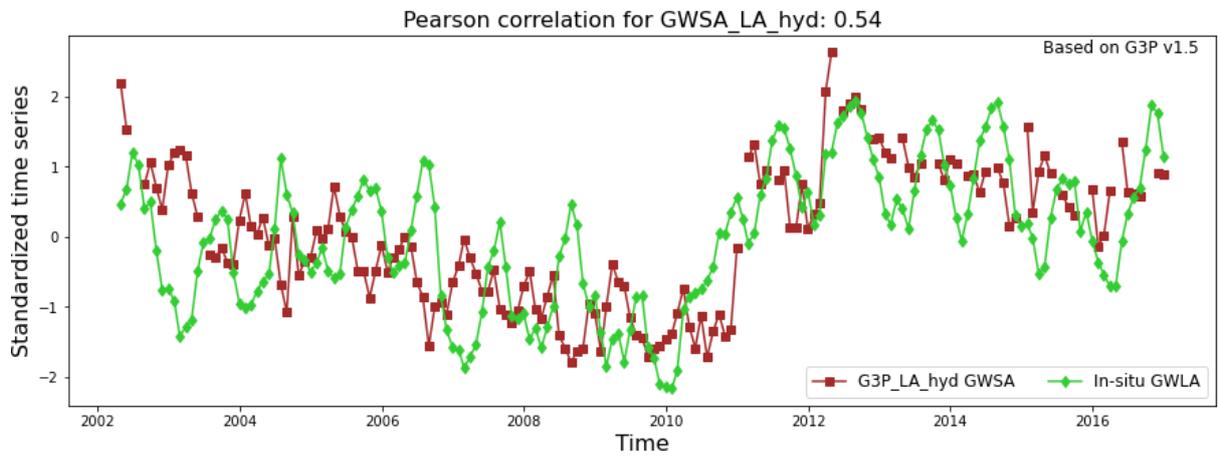


10.8 Paris Basin

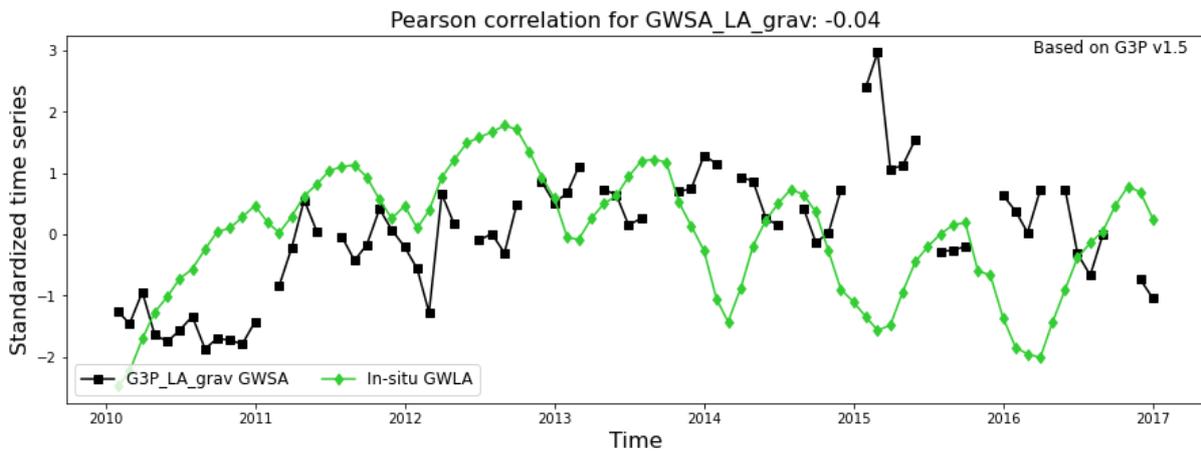
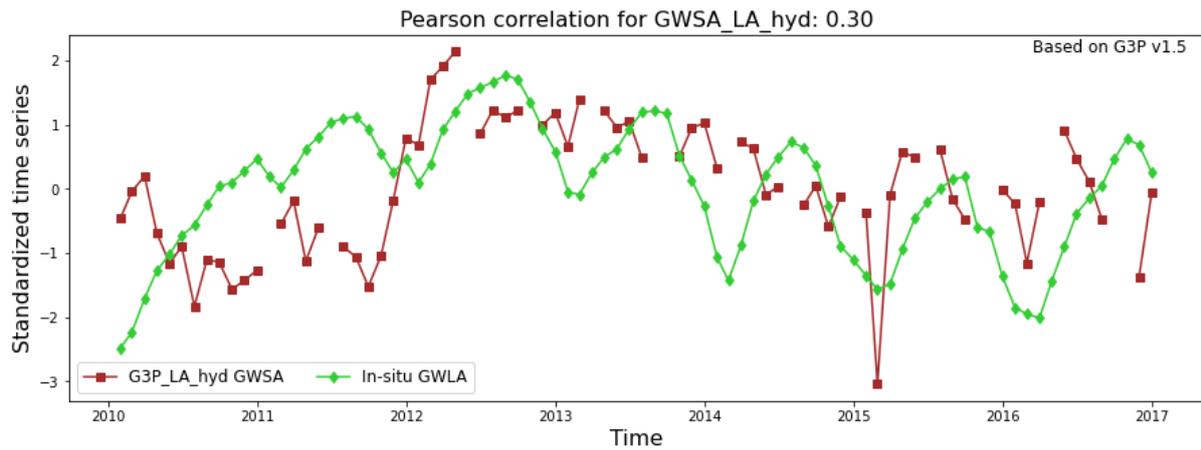




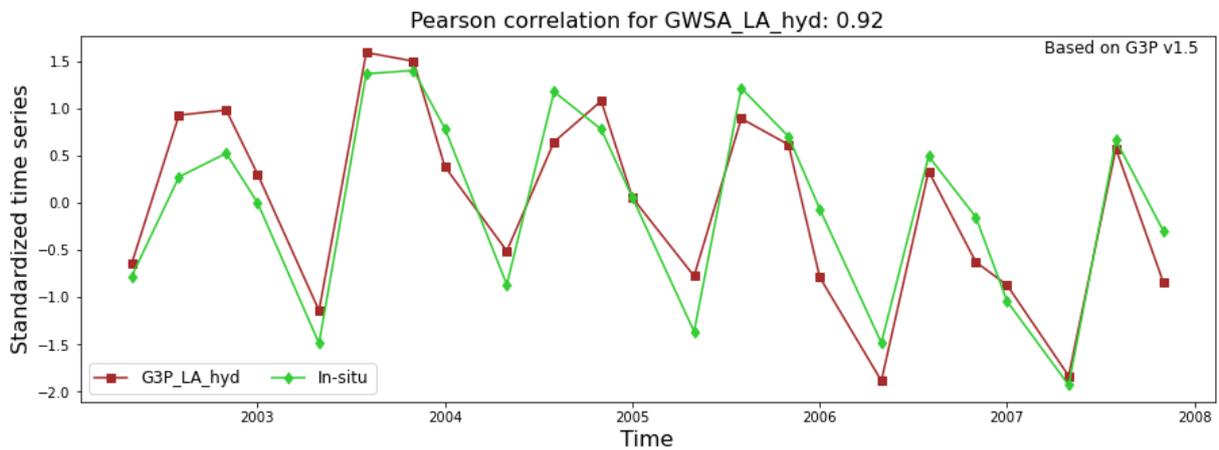
10.9 Murray Darling Basin

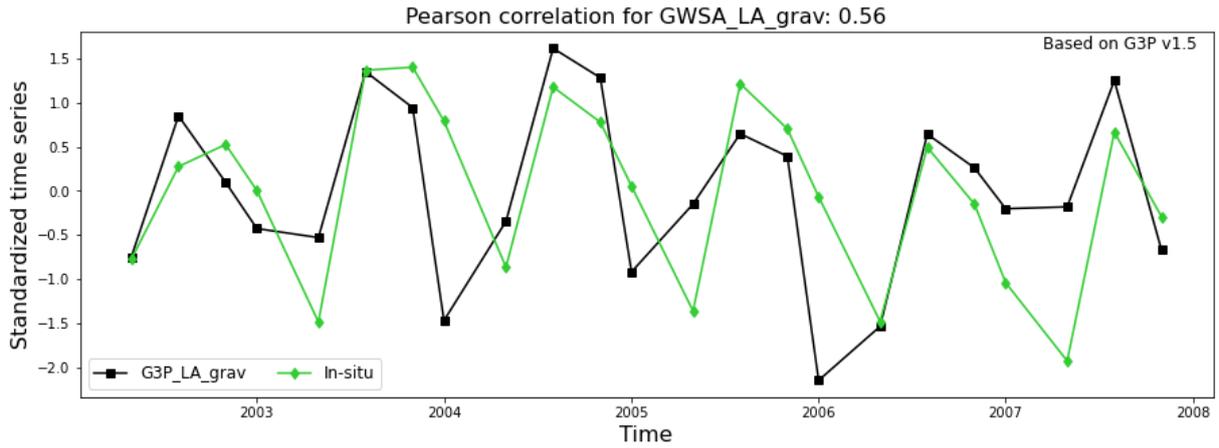


10.10 Great Artesian Basin

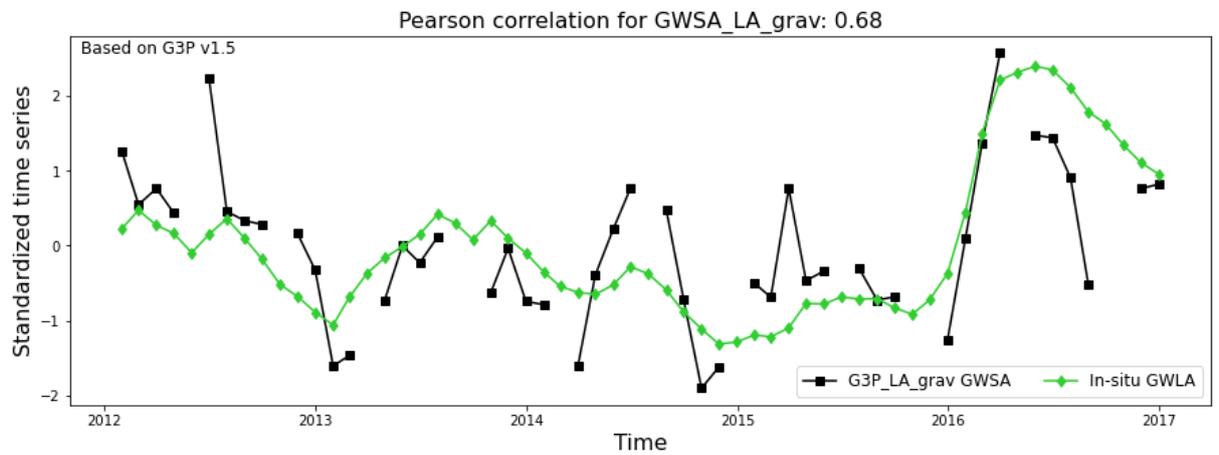
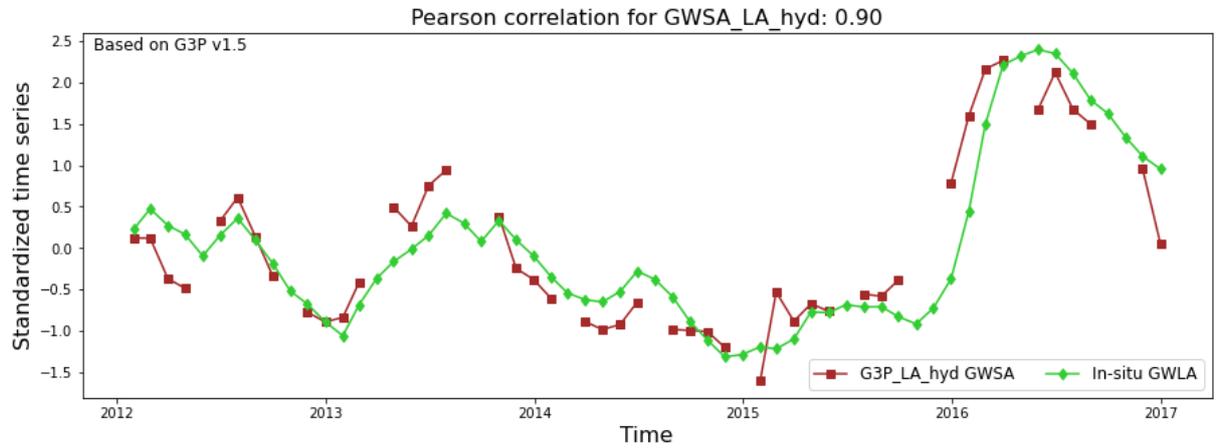


10.11 South of outer Himalayas aquifer

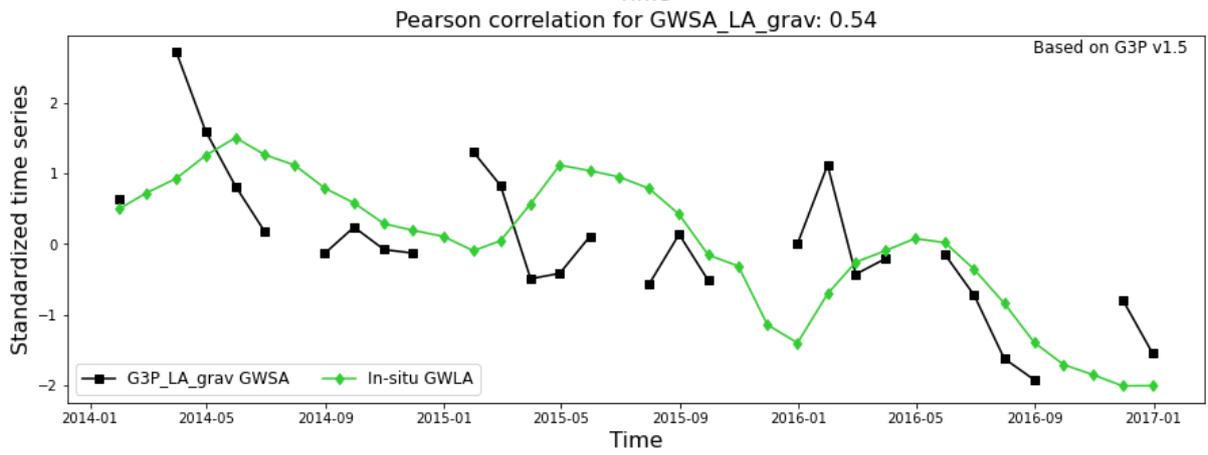
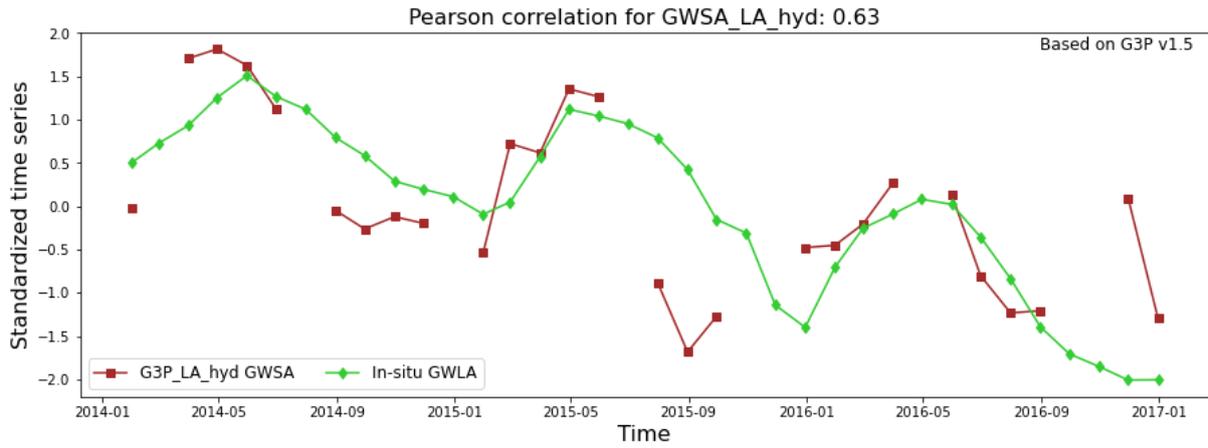




10.12 Guarani Aquifer System



10.13 Maranhão Aquifer System



11. References

- BRGM. (2022, June). *ADES, National Portal for access to groundwater data*. <https://Ades.Eaufrance.Fr/>.
- Bureau of Meteorology. (2022, May). *Australian Groundwater Explorer*. <http://www.bom.gov.au/water/groundwater/explorer/map.shtml>.
- Bush, P. W., & Johnston, R. H. (1988). *Ground-water hydraulics, regional flow, and ground-water development of the Floridan aquifer system in Florida and in parts of Georgia, South Carolina, and Alabama*. US Government Printing Office.
- Cano, A., Núñez, A., Acosta-Martinez, V., Schipanski, M., Ghimire, R., Rice, C., & West, C. (2018). Current knowledge and future research directions to link soil health and water conservation in the Ogallala Aquifer region. *Geoderma*, 328, 109–118.
- 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.
- de Sousa, S. B. (2000). Sistema aquífero da Ilha do Maranhão (MA)* (Aquifer system of the Maranhão Island, Maranhão, Brazil). *Águas Subterrâneas*.
- Deines, J. M., Schipanski, M. E., Golden, B., Zipper, S. C., Nozari, S., Rottler, C., Guerrero, B., & Sharda, V. (2020). Transitions from irrigated to dryland agriculture in the Ogallala Aquifer: Land use suitability and regional economic impacts. *Agricultural Water Management*, 233, 106061.
- Fu, G., Rojas, R., & Gonzalez, D. (2022). Trends in Groundwater Levels in Alluvial Aquifers of the Murray–Darling Basin and Their Attributions. *Water*, 14(11), 1808.
- Geological Service of Brazil. (2022, May). *Integrated Groundwater Monitoring Network (RIMAS)*. <http://rimasweb.cprm.gov.br/layout/index.php>.
- Gonçalves, R. D., Teramoto, E. H., & Chang, H. K. (2020). Regional groundwater modeling of the guarani aquifer system. *Water*, 12(9), 2323.
- Gong, H., Pan, Y., Zheng, L., Li, X., Zhu, L., Zhang, C., Huang, Z., Li, Z., Wang, H., & Zhou, C. (2018). Long-term groundwater storage changes and land subsidence development in the North China Plain (1971–2015). *Hydrogeology Journal*, 26(5), 1417–1427.
- Gupta, S. (2014). Ground Water Scenario of Himalayan Region, India. *Central Ground Water Board*.
- Habermehl, M. A. (2020). The evolving understanding of the Great Artesian Basin (Australia), from discovery to current hydrogeological interpretations. *Hydrogeology Journal*, 28(1), 13–36.
- IGRAC. (2020). *Groundwater monitoring programmes: A global overview of quantitative groundwater monitoring networks*.
- McGuire, V. L., Lund, K. D., & Densmore, B. K. (2012). Saturated thickness and water in storage in the High Plains aquifer, 2009, and water-level changes and changes in water in storage in the High Plains aquifer, 1980 to 1995, 1995 to 2000, 2000 to 2005, and 2005 to 2009. *US Geological Survey Scientific Investigations Report*, 5177, 28.
- Megnien, C. (1980). Tectogenese du Bassin de Paris; etapes de l'évolution du bassin. *Bulletin de La Société Géologique de France*, 7(4), 669–680.
- Miller, J. A., Appel, C. L., & Survey, U. S. G. (1997). Ground Water Atlas of the United States: Segment 3, Kansas, Missouri, Nebraska. In *Hydrologic Atlas*. <https://doi.org/10.3133/ha730D>

- Miller, J. A., & Survey, U. S. G. (1990). Ground Water Atlas of the United States: Segment 6, Alabama, Florida, Georgia, South Carolina. In *Hydrologic Atlas*. <https://doi.org/10.3133/ha730G>
- National Water Informatics Centre. (2022, May). *India Water Resources Information System*. <https://indiawris.gov.in/wris/#/DataDownload>.
- Olcott, P. G., & Survey, U. S. G. (1992). Ground Water Atlas of the United States: Segment 9, Iowa, Michigan, Minnesota, Wisconsin. In *Hydrologic Atlas*. <https://doi.org/10.3133/ha730J>
- Planert, M., Williams, J. S., & Survey, U. S. G. (1995). Ground Water Atlas of the United States: Segment 1, California, Nevada. In *Hydrologic Atlas*. <https://doi.org/10.3133/ha730B>
- Renken, R. A., & Survey, U. S. G. (1998). Ground Water Atlas of the United States: Segment 5, Arkansas, Louisiana, Mississippi. In *Hydrologic Atlas*. <https://doi.org/10.3133/ha730F>
- U.S. Geological Survey. (2020). *National Ground-Water Monitoring Network*. <https://cida.usgs.gov/ngwmn/index.jsp>.
- Whitehead, R. L., & Survey, U. S. G. (1996). Ground Water Atlas of the United States: Segment 8, Montana, North Dakota, South Dakota, Wyoming. In *Hydrologic Atlas*. <https://doi.org/10.3133/ha730I>