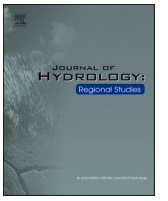
Journal of Hydrology: Regional Studies 64 (2026) 103118

Contents lists available at ScienceDirect

Journal of Hydrology: Regional Studies

journal homepage: www.elsevier.com/locate/ejrh



Spatial assessment of groundwater variability and drought impacts on ATES system suitability in Spain

Adela Ramos-Escudero [a,*], Carlos Toledo [b], Juan-de-Dios Gómez-Gómez [c], Martin Bloemendal [a,d], Antonio J. Collados-Lara [e], David Pulido-Velázquez [f]

[a] Department of Water Management, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, the Netherlands
[b] Murcian Institute of Agricultural and Environmental Research and Development (IMIDA), Sustainability and Quality Group of Fruit and Vegetable Products, Spain
[c] Spanish Geological Survey (IGME), CSIC, C/ La Calera 1, Tres Cantos, Madrid 28760, Spain
[d] KWR Water Research Institute, the Netherlands
[e] Department of Civil Engineering, University of Granada, Granada, Spain
[f] Water and Global Change Department, Geological and Mining Institute of Spain, Granada, Spain

ARTICLE INFO

Keywords:
ATES
Groundwater
Drought
Spain
Thermal energy storage
Quantitative suitability mapping
Temporal analysis
ATES site selection

ABSTRACT

Study region: This study is conducted across groundwater bodies within mainland Spain, as defined under the European Water Framework Directive.
Study focus: We conduct a preliminary, national-scale assessment of groundwater-body suitability for Aquifer Thermal Energy Storage (ATES) in Spain from a water-energy nexus perspective. The methodology is based on two complementary indicators derived from long-term piezometric records: (i) a Drought Stress Response Index (DSRI), reflecting aquifer reliability, resilience, and vulnerability over decadal time scales, and (ii) groundwater-level variability and long-term trends as proxies for hydraulic stability. Together, these indicators support a first-order screening of groundwater bodies from less suitable to more suitable conditions for ATES operation.
New hydrological insight of the region: The analysis of drought-response indicators reveals clear spatial patterns in aquifer vulnerability, resilience, and reliability across Spain, with only weak correlations with mean groundwater levels. Groundwater-level amplitude and trend analyses indicate that unstable conditions are concentrated in southern and eastern Spain, whereas northern regions generally exhibit more stable regimes. Building on these indicators, the results reveal pronounced spatial contrasts in ATES suitability, with generally more favorable conditions in northern regions and lower suitability in large parts of southeastern Spain, while extensive areas with intermediate suitability are also identified. Based on this national-scale screening, the study provides a preliminary assessment of ATES suitability for the main Spanish urban areas, offering an initial indication of where groundwater conditions are more or less favorable for ATES deployment.

1. Introduction

A sustainable energy transition requires not only expanding the use of renewable sources but also developing strategies that efficiently manage energy in space and time. Among these, Aquifer Thermal Energy Storage (ATES) systems have emerged as a robust

* Corresponding author:.

https://doi.org/10.1016/j.ejrh.2026.103118
Received 24 September 2025; Received in revised form 18 December 2025; Accepted 6 January 2026
Available online 26 February 2026

solution for seasonal heating and cooling, particularly in urban environments with high energy demand (Bloemendal et al., 2014; Heekeren and Bakema, 2015), with The Netherlands and Germany currently leading their large-scale implementation.

In Northern Europe, aquifers are generally characterized by relatively stable water tables,high and well-distributed annual precipitation, and modest groundwater withdrawals, creating favorable hydrogeological and regulatory conditions for long-term ATES operation (European Environment Agency, 2018; Pulido-Velazquez et al., 2020). In contrast, ATES deployment in Southern European Union Member States (SEUM) remains limited and understudied, largely due to stronger climatic variability and groundwater stress. SEUM regions, such as Spain the focus of this study experience prolonged droughts, irregular rainfall distribution, and hotter summers, resulting in episodic and spatially heterogeneous groundwater recharge, often confined to winter months. In addition, long-term, large-scale agricultural irrigation places substantial pressure on groundwater resources, accounting for up to 80% of total water withdrawals in some river basin districts (European Environment Agency, 2018). When groundwater abstraction exceeds natural recharge, aquifers become overexploited, leading to pronounced groundwater-level declines and enhanced seasonal and interannual fluctuations (Famiglietti et al., 2024; Saïd et al., 2020). These conditions critically constrain ATES feasibility, as declining water tables reduce effective storage volume and weaken thermal recovery performance. These factors result in a higher temporal variability of groundwater levels, which may affect the technical feasibility and long-term efficiency of ATES systems.

Reductions in groundwater availability can strongly affect ATES performance through geometric and hydraulic controls, particularly the hydraulic radius (R_h) and thermal radius (R_{th}) of the storage zone. While previous studies have examined ATES-induced hydrological disturbances (e.g., mounding or thermal plumes) (Bonte et al., 2011; Stemmle, et al., 2021), no work has evaluated how externally driven water-table declines, caused by droughts or long-term groundwater overexploitation, directly impair ATES efficiency. A falling water table reduces the saturated screen length, shrinking the cylindrical storage volume and increasing its surface-area-to-volume ratio (A/V). This ratio is fundamental in ATES systems because conductive heat losses occur across the boundary of the stored volume; thus, higher A/V geometries expose a larger fraction of the stored heat to the surrounding formation. As demonstrated analytically and numerically (Bloemendal and Hartog, 2018), thermal recovery efficiency decreases linearly with increasing A/V, making volume contraction a direct pathway for efficiency loss.

Beyond conductive leakage, groundwater-level decline also weakens hydraulic confinement and reduces the storage volume (V) and R_{th}, thereby lowering the R_{th}/u ratio that governs advective displacement by natural groundwater flow. When R_{th}/u approaches unity, displacement losses become dominant, sharply decreasing thermal recovery efficiency. As illustrated in Fig. 1, significant drawdowns can partially expose the screen, shrink the effective storage zone, and diminish the recoverable thermal capacity even under unchanged operational conditions. These geometric and hydraulic constraints show that water-table decline is an external forcing capable of degrading ATES performance independently of system design, representing a critical limitation for shallow, unconfined aquifers typical of southern Europe.

Previous ATES suitability research has focused on well-monitored, saturated aquifers in Northern Europe, where water levels remain relatively stable year-round. However, the impacts of climate-driven decline and fluctuations in groundwater level, as well as large- scale ongoing groundwater extractions with limited recharge, on ATES system suitability remain largely unexplored (Sommer et al., 2013). Building upon the initial national-scale assessment of ATES suitability in Spain by (Ramos-Escudero and Bloemendal, 2022), this study offers a more comprehensive spatial evaluation by incorporating updated hydrogeological data and explicitly analyzing long- term groundwater level variability and trends, which are critical for assessing the reliable operation of ATES systems under changing climatic conditions.

This paper addresses this knowledge gap by assessing the suitability of aquifers for ATES deployment in Spain through a novel analytical framework, introducing the Drought Stress Response Index (DSRI) based on the Standardized Groundwater Index (SGI) to evaluate aquifer stability and its implications for thermal energy storage.

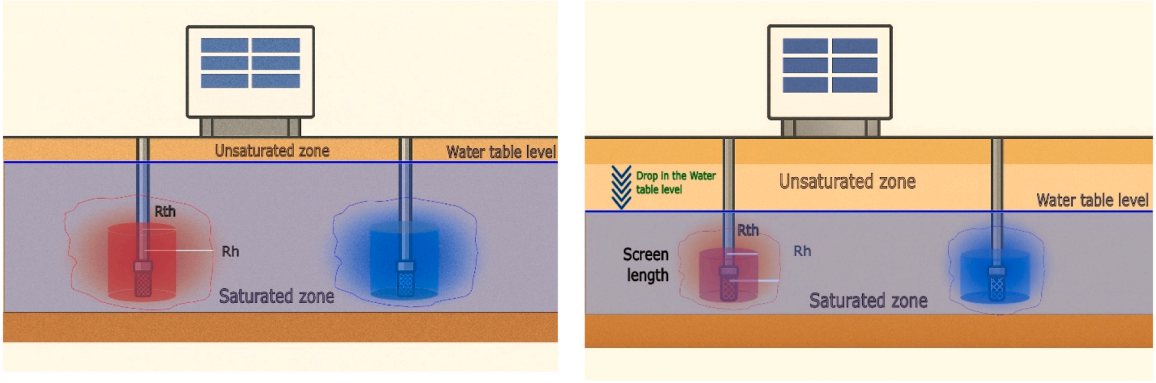


Fig. 1. Effect of groundwater table decline on ATES storage volume. Lowering of the groundwater table reduces the saturated screen length, shrinking the effective storage volume and the corresponding hydraulic and thermal radii (R_h and R_{th}). This contraction limits the groundwater volume available for heat injection and extraction and increases the surface-area-to-volume ratio (A/V), thereby reducing thermal storage efficiency.

2. Methodology

Fig. 2 presents the general framework we apply in this study. Two independent analyses are conducted to evaluate the suitability of groundwater bodies in Spain for Aquifer Thermal Energy Storage (ATES): a Drought Stress Response analysis and a Water Table Range and Trend analysis. Both approaches rely on the same input dataset historical groundwater level records from the national monitoring network and share a common initial stage of data selection and processing. Specifically, steps 1.1 (spatial vs temporal trade-off) and 1.2 (gap filling) are used exclusively for the Drought Stress Response path, while step 1.3 (correction of anomalies) feeds into the Range and Trend analysis. The workflows then diverge to follow distinct computational procedures. The Drought Stress Response analysis applies indices such as the Standardized Groundwater Index (SGI), the Drought Stress Response Index (DSRI), and indicators of resilience, reliability, and vulnerability (steps 2.1–2.3) to evaluate aquifer behavior under drought conditions, a critical factor in a country frequently affected by water scarcity. In contrast, the Range and Trend analysis focuses on intrinsic groundwater dynamics by analyzing the fluctuation range and long-term trends (steps 2.4.1 - 2.4.2), which are key to assessing the temporal stability of ATES performance. In both cases, groundwater levels are assessed at both the piezometer and water body levels. All calculations were carried out using QGIS 3.x (QGIS Development Team, 2025).

2.1. Data selection and processing

To detect robust drought or water scarcity patterns, scientific literature (Smakhtin, 2001; Tallaksen, and Van Lanen, 2004; Van Loon and Van Lanen, 2013) typically recommends a minimum of 10–30 years of hydrogeological data. Hence, a selection process was applied to the available groundwater level time series to ensure data quality and consistency. The objective was to identify piezometers with the most complete and reliable records over a defined period, thereby ensuring that the dataset was suitable for applying the proposed methodology. In this phase, the temporal coverage of each piezometer was assessed to determine the most suitable period for analysis. The time series of each station was scanned to identify the longest continuous interval with the highest number of monthly observations, thereby maximizing data availability and continuity. This adaptive selection ensured that, for each location, the most complete temporal dataset was retained for further processing. Only the temporal dimension was considered at this stage, without accounting for the spatial representativeness of the monitoring network. This approach is consistent with established practices in long-term groundwater trend assessments (Dai et al., 2019; Taylor et al., 2013), and it provided a robust foundation for both the drought response and groundwater dynamics analyses.

2.1.1. Spatial-temporal trade-off coverage

In the second phase, a standardized time window was selected across all piezometers (and its water bodies associated) to ensure spatial comparability of the results. While the initial step prioritized temporal completeness on a per-piezometer basis, this stage

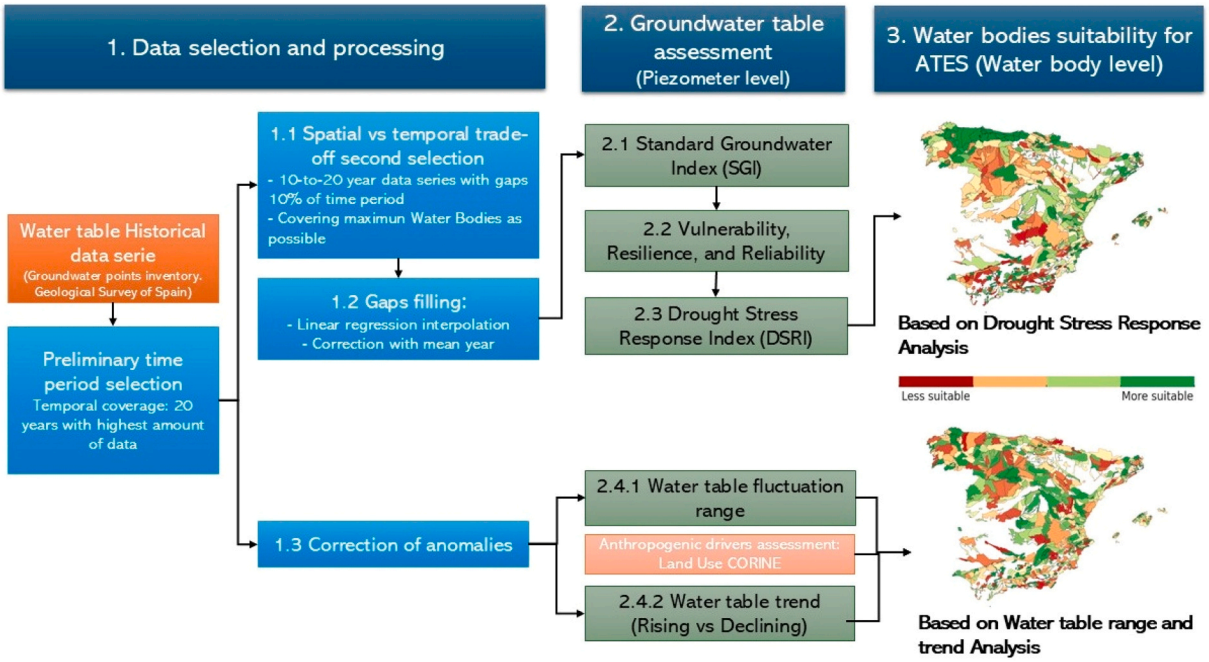


Fig. 2. Schematic representation of the methodological workflow for assessing the suitability of groundwater bodies for ATES systems. The approach combines two analyses: one based on long-term drought response (DSRI), and another on groundwater-level variability. Both are computed at the piezometer level and aggregated to evaluate suitability at the water body scale.

applies a common period to all selected stations, allowing for consistent computation of groundwater variability indicators at the national scale. The selection of this shared time frame involved a compromise: it had to be long enough to reflect meaningful groundwater trends, but also recent and data-rich enough to include as many piezometers as possible.

To ensure comprehensive spatial coverage of groundwater bodies, the time series used in this step was defined based on the temporal availability of data across the study area. The analysis includes multiple reference periods, ranging from 10 to 20 years, within an overall data availability range of 10–30 years. Groundwater datasets often contain gaps ranging from daily to yearly which affect the continuity and reliability of the signal. A trade-off was necessary to select piezometers with sufficient data availability, maximizing the number of groundwater bodies included in the analysis while maintaining minimal consecutive data gaps. A threshold of 10% missing values over the selected time period was defined to exclude piezometers with excessive data gaps, thereby minimizing the uncertainty associated with gap-filling procedures. This approach balances temporal robustness with spatial representativeness, a compromise that has been adopted in recent hydrogeological and drought studies when long-term records are incomplete or heterogeneously distributed. The 10% threshold and selection criteria follow approaches used in previous studies (Bloomfield and Marchant, 2013; Herrera-Pantoja and Hiscock, 2008), and ensure that only reliable time series are used for indicator calculation.

Only piezometers that met the minimum data quality criteria were retained for each selected period. For every case, the number of distinct groundwater bodies represented by valid piezometers was calculated. This method ensures that the final selection prioritizes spatial representativeness while maintaining the integrity of the time series, effectively balancing the trade-off between data quality and geographic coverage.

2.1.2. Gaps filling

In hydrological time series analysis, especially when calculating standardized drought indicators such as the Standardized Groundwater Index (SGI), data continuity is essential to ensure the validity of statistical metrics and comparability over time. Missing data can distort the underlying seasonal and long-term trends, introduce artificial variability, and bias the identification of drought episodes. Therefore, before deriving any performance metrics, it is critical to apply consistent and transparent gap-filling methods. This is particularly important for national-scale groundwater assessments where heterogeneity in monitoring practices often results in variable data completeness. Thus, the gap-filling process was conducted following these three steps:

2.1.2.1. Linear interpolation. The gap-filling process was carried out via linear temporal interpolation, which consists of estimating missing values by computing the average of the nearest available observations before and after the gap. Although this method preserves long-term trends and reduces the risk of introducing abrupt anomalies (Van Loon and Van Lanen, 2013), it can attenuate seasonal dynamics. To address this, a correction step was applied to restore intra-annual variability in the interpolated segments.

Missing values in the standardized groundwater level series (i, t) for each piezometer i were first interpolated linearly over time. This interpolation was applied across the full-time range, including gaps at the beginning and end of the series, as in Eq. 1:

$$z_{i,t}^{interp} = \text{Linearinterpolation}(z_{i,t}) \tag{1}$$

Since the analysis was based on one value per month, in cases where multiple measurements were available within the same month, their average was used to obtain a single representative monthly value before interpolation.

2.1.2.2. Estimation of the mean year and regression trend. To further reduce seasonal inconsistencies introduced during interpolation, a correction was applied based on the mean annual cycle of each piezometer (Bloomfield and Marchant, 2013; Zipper et al., 2019). This correction consisted of computing the average value for each calendar month (January to December) across all years with valid data and adjusting the interpolated series accordingly to restore intra-annual variability. Such detrending procedures are standard practices in groundwater time series analysis to isolate anomalies (Bloomfield and Marchant, 2013). For each piezometer i, the mean year $\overline{Y}_i$ was calculated as the average of all years with valid observations, as in Eq. 2, where Year(t) denotes the annual mean groundwater level for year t computed from all available monthly observations within that year, and n_i is the number of valid years with observations in the time series of piezometer i:

$$\overline{Y}_i = \frac{1}{n_1} \sum_{t=1}^{n_i} \text{Year}(t) \tag{2}$$

Monthly z-scores corresponding to year $\overline{Y}_i$ were then used to fit a linear regression capturing the mean temporal pattern, as in Eq. 3:

$$z_{i,t} = \alpha_i + \beta_i t \tag{3}$$

Where t represents the time index (in months from the beginning of the series), α_i is the intercept corresponding to the estimated groundwater level at the start of the series, and β_i is the slope representing the long-term trend in groundwater levels for piezometer i.

2.1.2.3. Correction of interpolated values. The final correction of the time series involved removing the estimated regression trend from the interpolated and measured groundwater levels. For each time step, the expected value based on the fitted regression line was subtracted from the interpolated value, producing a corrected series that reflects short-term variability while eliminating long-term systematic biases.

This approach ensures that the resulting time series reflects short-term variability around a stable baseline, which is required for the subsequent computation of standardized indices such as the SGI (Herrera-Pantoja and Hiscock, 2008).

The estimated slope β_i captures the long-term increase or decrease in groundwater levels for each piezometer. To correct the interpolated series for this trend, the fitted linear component $\alpha_i + \beta_i t$ was subtracted from the full time series (including interpolated values), effectively centering the data around a stationary baseline. This step ensures that subsequent analyses focus on anomalies and variability rather than long-term directional changes as in Eq. 4:

$$z_{i,t}^{detrended} = z_{i,t}^{interp} - \alpha_i + \beta_i t \quad (4)$$

Once the time series had been interpolated, detrended, and corrected for seasonal inconsistencies, the data were considered ready for drought index computation. At this stage, each piezometric series represented groundwater level anomalies that were free from long-term trends and retained meaningful intra-annual variability. These standardized and pre-processed time series served as the input for the drought index calculations described in the following section.

2.1.3. Correction of data anomalies

Historical groundwater level series often contain spurious values due to measurement errors, sensor failures, or transcription mistakes. These anomalies, although infrequent, can significantly distort key statistical metrics such as the mean, standard deviation, or extremes especially when computing standardized indicators. In particular, isolated outliers can artificially inflate the variability of a series and shift its minimum or maximum values far beyond the plausible physical range, leading to biased interpretations of groundwater trends or drought severity.

To address this, an anomaly correction step was implemented to identify and remove implausible values, as in Eq. 5.

$$x < \mu - 3\sigma \text{ or } x > \mu - 3\sigma \quad (5)$$

Where x denotes an individual groundwater level observation in the piezometer time series, μ represents the mean groundwater level of that piezometer, and σ is the corresponding standard deviation. Anomalous observations were flagged when lying outside the range $\mu \pm 3\sigma$ of each piezometers time series. This $\pm 3\sigma$ criterion corresponds to the 99.7% confidence interval of a normal distribution and has been widely applied in groundwater time series analyses to reduce the influence of measurement errors and extreme outliers (Bloomfield and Marchant, 2013; Van Loon and Laaha, 2017).

This rule, commonly used in environmental time series processing (Wilks, 2011), has also been applied in groundwater studies to improve data reliability and comparability (Gómez et al., 2022). The application of this filter reduces the influence of extreme errors and improves the accuracy of groundwater level amplitude calculations, thereby enhancing the robustness of subsequent variability and trend assessments.

Results of data selection and processing section are shown in Appendix.

2.2. Standard Groundwater Index (SGI)

The Standard Groundwater Index (SGI) is a groundwater drought indicator used to standardize groundwater level time series and to characterize groundwater drought conditions. It builds upon the methodology of the Standardized Precipitation Index (SPI), which is widely used for characterizing meteorological droughts, but it is adapted to account for the distinct characteristics of groundwater data. In this study, the SGI was computed as a standardized groundwater level anomaly using a z-score normalization, a conventional approach previously applied in groundwater drought and groundwater level anomaly studies (e.g (Mendicino et al., 2008a; McKee et al., 1993).).

The SGI is calculated for valid piezometers using Eq. 6:

$$SGI = \frac{x - \mu}{\sigma} \quad (6)$$

Where:

- X: Water table level (m.a.s.l.) measurement per month
- μ: average water level of the temporal series
- σ: Standard deviation of the series

Although the SGI in this study is computed using a z-score normalization rather than a non-parametric transformation, the resulting index is expressed in standardized units comparable to those used in previous groundwater drought studies, allowing the application of commonly adopted drought thresholds.

The analysis of drought conditions in hydrological systems has traditionally relied on standardized indices such as the SPI. However, groundwater level dynamics often deviate from parametric distributions due to aquifer storage effects, delayed responses, and anthropogenic pressures. To address these challenges, (Bloomfield and Marchant, 2013) proposed the use of a non-parametric normal score transformation for groundwater drought assessment, avoiding assumptions regarding the underlying data distribution (Everitt, 2002). In the present study, a simplified z-score standardization based on the long-term mean and standard deviation was adopted, enabling a consistent comparison of groundwater level anomalies across a large and spatially heterogeneous monitoring

network.

This methodological choice is justified by the primary objective of this work, which is not to perform a detailed probabilistic assessment of groundwater droughts, but rather to investigate the spatial heterogeneity of groundwater level dynamics and their implications for the potential deployment of efficient Aquifer Thermal Energy Storage (ATES) systems.

Various studies have employed the SGI to characterize groundwater droughts using standardized threshold values. Groundwater drought events are commonly identified more negative (e.g., (Bloomfield and Marchant, 2013); Fiorillo, 2010; Fiorillo and Guadagno, 2012; Mendicino et al., 2008b). In this study, a drought period is defined as the criteria applied by the Spanish Ministry for the Ecological Transition and other groundwater drought assessments (Pulido-Velazquez et al., 2015).

2.3. Drought Stress Response Index (DSRI)

The Drought Stress Response Index (DSRI) developed in this study is a composite metric used to quantify the temporal stability of groundwater levels, serving as a proxy for aquifer suitability for long-term Aquifer Thermal Energy Storage (ATES) deployment (Stemmle, et al., 2021; Loucks, 1997).

The DSRI evaluates aquifer resilience to drought using time series data that were rigorously pre-processed through quality control, gap filling, and anomaly correction to ensure the robustness of the indicators.

Anthropogenic drivers (e.g., agriculture, urban land use) and lithological characteristics were integrated to explore potential causes of the observed variability and to assess how hydrogeological properties influence aquifer behavior. Cross analysis with land-use and climate data revealed clear spatial gradients, with greater stability in the north and higher stress in the south.

The DSRI is calculated by first computing individual values for each piezometer, then aggregating them by groundwater body using the median of all local piezometers. The indicator combines three components reliability, resilience, and relative vulnerability to capture the degree of groundwater level variability over time, as in Eq. 7:

$$\text{DSRI} = \text{Reliability } x \text{ Resilience } x(1 - \text{Relative Vulnerability}) \tag{7}$$

The DSRI ranges from 0 to 1, with higher values indicating more stable systems, implicitly reflecting higher reliability and lower vulnerabilitykey aspects of aquifer suitability for ATES. In contrast, lower DSRI values reveal systems with greater temporal variability where fluctuating groundwater levels could reduce the thermal efficiency and operational reliability of ATES installations.

Detailed definitions and calculation procedures for each component are provided in the following section.

2.4. Vulnerability, Resilience, and Reliability

In order to comprehensively assess groundwater system performance under stress, it is critical to evaluate not only the magnitude and frequency of undesired conditions but also the capacity of the system to recover and maintain service (Loucks, 1997). established a foundational theoretical framework to quantify the sustainability of water resource systems, illustrating its application through various management scenarios, including surface storage and groundwater use. Within this framework, three performance criteria were introduced: vulnerability (VUL1), Resilience (RES1), and Reliability (REL), which can be applied to water resource indicators such as the SGI (Standardized Groundwater Index). These metrics allow for the quantification of how often a system meets desired thresholds (reliability), how quickly it recovers after a failure (resilience), and how severe the consequences are when it fails (vulnerability). Building on this work, (Mays, 2013) built on this framework and introduced an integrated Sustainability Index that incorporates these three criteria, tailoring it specifically for groundwater systems, reinforcing their relevance for long-term water resource planning. Applying these indicators to the Standardized Groundwater Index (SGI) at the piezometer level enables a spatially explicit evaluation of groundwater behavior under drought conditions. Aggregating these metrics at the water body (mass of water) scale allows for identifying aquifers with greater systemic vulnerability or lower resilience, supporting targeted management actions.

Vulnerability (VUL 1) is a probabilistic measure accounting for the extent and magnitude of an unsatisfactory condition. It is obtained as the summation of the unsatisfactory time series. It identifies periods of drought based on consecutive values of the SGI below a defined threshold, typically SGI < -1, up to SGI > 0, when the drought conditions are believed to end. It can be calculated for a time series as in Eq. 8.

$$\text{VUL}_{\text{piezo}} = \quad \text{SGI}_{\text{avg(drought)x}}\text{P}_{\text{SGI}} \tag{8}$$

Where

- $SGI_{avg(drought)xP_{SGI}}$ is the average of SGI of the piezometers during drought conditions.
- P_{SGI} is the probability of SGI to occur in the time series, at it can be calculated with Eq. 9:

$$\text{P}_{\text{SGI}} = \frac{\text{number of months with SGI} < \quad -1}{\text{Total number of months in the piezometer's time series}} \tag{9}$$

Resilience (RES) is assessed based on the average duration of drought events, following a methodology inspired by Zipper et al. (2019). In this context, drought events were periods during which the Standardized Groundwater Index (SGI) fell below −1. First, all drought events for each monitoring well are identified. Again, a drought event is considered to begin when the SGI dropped below −1 and to end when the SGI rose above −1. The duration (in months) of each drought event was calculated. The mean drought duration

was then computed by averaging the durations of all events observed throughout the record.

The resilience index (*RES*) was defined as the inverse of the mean drought duration, as in Eq. 10:

$$\mathrm{RES} = \frac{1}{\text{Mean drought duration(months)}} \tag{10}$$

In this formulation, wells exhibiting faster recovery (shorter drought durations) achieve higher resilience scores, while wells with prolonged drought periods exhibit lower key aspect of groundwater resilience as highlighted in previous studies (Zipper et al., 2019).

Reliability (REL) in a water resource system refers to the probability that the system will function satisfactorily over a given period. It measures how often a system meets performance criteria without failure, and can be calculated as in Eq. 11.

$$\mathrm{REL} = \frac{\text{Number of satisfactory outputs}}{\text{Total number of time periods observed}} \tag{11}$$

where satisfactory *output* was defined as a month in which groundwater levels remained above the drought threshold ($\mathrm{SGI} \geq -1$) indicating non-drought conditions.

2.5. Water table trend and range analysis

To evaluate the long-term viability of groundwater bodies for Aquifer Thermal Energy Storage (ATES), two key indicators of aquifer dynamics were considered: the amplitude of groundwater level fluctuations and their long-term trend. These metrics provide complementary insight into aquifer behavior: variability captures short-term stability, while trend reflects long-term sustainability.

Across Europe, typical annual groundwater level fluctuations range between 0.5 and 10 m in most regions, with lower variability observed in northern and Atlantic areas due to stable climatic conditions and effective water management (Döll and Fiedler, 2009). In contrast, southern European countries particularly Spain, Italy, and Greece often exhibit much greater variability, with ranges exceeding 20 m and, in extreme cases, surpassing 100 m, especially in semi-arid zones or overexploited aquifers (Taylor, et al., 2024). This highlights the increased sensitivity of southern regions to both climatic and anthropogenic pressures.

To quantify groundwater-level variability across Spain, time series from piezometers with monthly observations over the last two decades were analyzed. For each piezometer, the total phreatic range was calculated as the absolute difference between the maximum and minimum observed values (Total Variability), based on the entire available time series, regardless of continuity or data gaps. This approach allowed the inclusion of all piezometers with sufficient temporal coverage, without requiring interpolation or artificial gap filling. All time series had previously been screened and corrected for anomalous values following the procedure described in Section 2.1.3 (Correction of anomalies); consequently, the variability metrics were computed on data that were already free of outliers. Each observation point was then spatially linked to its corresponding groundwater body using the IGME hydrogeological database, which also provided key aquifer characteristics such as top and bottom elevations, total thickness, and saturated thickness.

In parallel, groundwater-level trends were analyzed to detect signs of long-term aquifer decline or recovery. These trends are relevant for ATES planning, as they signal whether the available storage space is increasing, stable, or shrinking over time. Reference values from across Europe were considered to contextualize results: in northern and central Europe, typical annual changes in groundwater levels remain within ±0.2 m/year, reflecting stable climatic regimes and regulated abstraction (Richter and Therrien, 2022). In contrast, southern countriesparticularly Spain, Italy, and Greece often report declining trends ranging from −0.5 to −2.0 m/year, especially in semi-arid regions subject to intensive agricultural withdrawals (Taylor, et al., 2024; Fowler et al., 2023). These values were used as a benchmark to assess the severity and spatial distribution of trends within the Spanish territory.

2.6. Water body suitability for ATES

ATES suitability was derived by converting the numerical values of both groundwater indicators (DSRII and the Variability Index) into ordinal suitability classes. Because neither index provides an absolute threshold to determine whether a groundwater body is intrinsically suitable or unsuitable, suitability was defined relative to hydrological favorability, ranking groundwater bodies from less to more suitable.

For the DSRII-based assessment, the index values computed at the piezometer level were aggregated at the groundwater-body scale using the median, which provides a robust representation of aquifer behavior while reducing the influence of local anomalies. The resulting median DSRII values were then classified into suitability classes based on their position within the empirical distribution: groundwater bodies with lower DSRII values indicating weaker drought-response behaviorwere assigned to the less suitable class, intermediate values to a medium suitability class, and those with higher DSRII values to the more suitable class.

A similar classification logic was applied to the Variability Index. The full range of index values was divided into ordinal classes, whereby groundwater bodies with higher variability (i.e., large or frequent water-level fluctuations) were assigned to the less suitable category, intermediate values to medium suitability, and groundwater bodies with low variabilityreflecting stable saturated conditions were classified as more suitable. This aligns with the operational requirement of ATES systems, which rely on consistent groundwater levels to maintain an effective saturated storage volume. When combining indicators, the highest suitability is therefore initially expected in areas that simultaneously exhibit a high Drought Stress Response Index (DSRI) indicating favorable and resilient behavior under drought conditions and a low Variability Index, which ensures stable water-level dynamics. Groundwater bodies meeting both criteria represent the most promising conditions for reliable ATES operation.

Together, these two classification approaches generate complementary hydrological suitability layers. The DSRII-based suitability reflects the long-term structural behavior of aquifers under drought stress, while the Variability Index captures the short-term and seasonal stability relevant to ATES operation. The resulting suitability classes should therefore be interpreted as relative favorability rankings, highlighting groundwater bodies where hydrological conditions are supportive of ATES deployment.

2.7. Application to the river basins in Spain

Spain is divided into 25 hydrographic demarcations under the Water Framework Directive (WFD), of which nine corresponding to the Canary Islands were excluded from this analysis due to the lack of long-term groundwater-level records. The primary source of groundwater information is the national Piezometric Monitoring Network operated by the Spanish Ministry for Ecological Transition (MITECO) (Pulido-Velazquez et al., 2015), comprising approximately 2500 piezometers with predominantly monthly manual measurements and, in some basins, daily automated observations. Spatial datasets for watersheds, groundwater bodies and monitoring points used in the GIS analysis were obtained from official repositories (Millán et al., 2004; Ministry for the Ecological Transition and the Demographic Challenge, 2025b).

The watersheds included in this study exhibit substantial variability in size, ranging from ~5,000 km^2 to more than 85,000 km^2, and contain on average 80–100 groundwater bodies and 300–500 piezometers per demarcation. Monitoring density and temporal continuity vary significantly between basins: districts such as Ebro, Júcar, Segura and Guadiana maintain extensive,frequentlymeasured networks, whereasothers,including Guadalete-Barbate, have more limited datasets. Historical inconsistencies partly linked to institutional restructuring during the early implementation of the WFD (Peña and Mondragón, 2024). result in gaps in some time series, requiring additional data-quality screening and gap-handling procedures in this study.

Fig. 7 provides an overview of the spatial distribution of piezometers across Spain, including their hydrogeological classification (porous, fractured, confined) and the delineation of groundwater bodies. Water bodies currently classified at quantitative risk (shown in red) are also highlighted; many correspond to overexploited aquifers, particularly in Mediterranean and southeastern regions where groundwater demand is high and recharge is limited.

Climatic conditions further contribute to the spatial heterogeneity of groundwater-level dynamics. Mediterranean and semi-arid basins experience irregular precipitation, high evapotranspiration and frequent droughts, leading to strong interannual fluctuations and reduced natural recharge (Pulido-Velazquez et al., 2015; Nlend, et al., 2025). In contrast, the northern Atlantic region shows more regular rainfall and comparatively stable piezometric behavior, while continental interior basins present marked seasonality. These climatic gradients, combined with spatial differences in monitoring density and hydrogeological complexity, are central to interpreting groundwater-level variability and assessing aquifer suitability for ATES deployment across Spain (Millán et al., 2004).

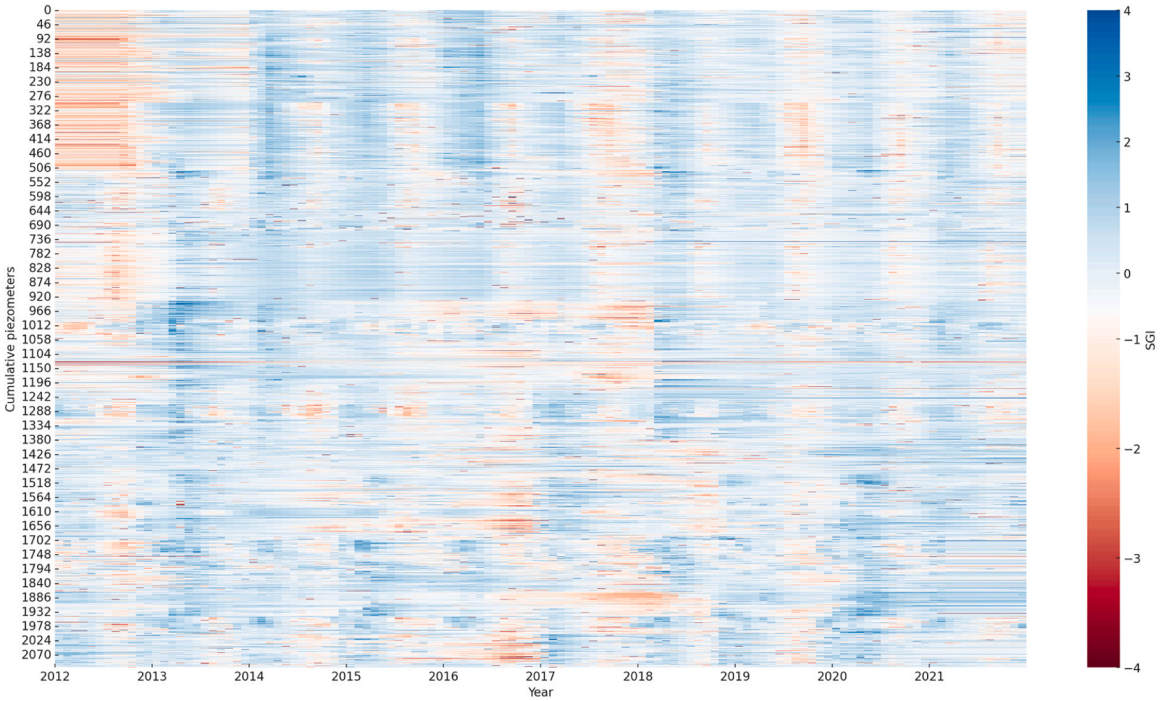


Fig. 3. Heatmap of the Standardized Groundwater Index (SGI) from 2012 to 2021. Each row represents an individual piezometer, and each column corresponds to a monthly observation. Colors indicate standardized groundwater level anomalies (SGI), with negative values representing below-average groundwater levels and positive values indicating above-average conditions. The figure illustrates the temporal variability and spatial heterogeneity of groundwater drought conditions across the monitoring network.

3. Results and discussion

3.1. Standard Groundwater Index (SGI)

The SGI results are depicted in a heatmap in the Fig. 3 based on interpolated and trend-corrected monthly data between 2012 and 2021. The heatmap provides a synoptic overview of the spatiotemporal variability of standardized groundwater level anomalies across the monitored piezometers. Several coherent temporal features can be identified. Periods characterized by predominantly negative SGI values affecting a large fraction of piezometers are visible as vertical bands notably in 2012, 2017, and 2019 suggesting episodes of widespread groundwater stress at the regional scale. In contrast, years such as 2013 and 2018 exhibit positive SGI values, indicating wetter-than-average conditions and potential aquifer recharge across the network, indicting relatively wetter-than-average groundwater conditions and enhanced recharge in multiple areas.

The heatmap also emphasizes the pronounced spatial heterogeneity in groundwater responses. While some piezometers consistently exhibit higher SGI values over the study period, others show persistently lower values, reflecting differences in hydrogeological settings, recharge conditions, and anthropogenic pressures such as groundwater abstraction. This variability persists despite the application of interpolation and trend- correction procedures designed to preserve the long-term characteristics of individual time series.

Overall, the SGI heatmap provides a compact overview of both short-term fluctuations and longer-term differences in groundwater dynamics at the national scale. The observed heterogeneity and intermittently synchronized drought periods underline the importance of spatially explicit analyses when evaluating groundwater-dependent applications, including the potential performance and resilience of Aquifer Thermal Energy Storage (ATES) systems.

3.1.1. Vulnerability, Resilience and Reliability

The analysis of the groundwater system using the vulnerability (VUL), reliability (REL), and resilience (RES) indicators revealed distinct patterns across the monitored piezometers. The results are depicted in Fig. 4. Vulnerability values ranged widely, with VUL1_piezo values spanning from approximately −0.01 to around −0.6, indicating that some piezometers were exposed to more frequent or severe drought conditions than others. However, a significant proportion of the piezometers exhibited low vulnerability (VUL values greater than −0.3), suggesting that severe drought episodes were relatively rare in much of the monitored area. Resilience

Fig. 4. Spatial distribution of piezometer-based drought-response indicators, showing Vulnerability (VUL; upper right), Resilience (RES; upper left), and Reliability (REL; bottom panel).

(RES) values also demonstrated notable differences across wells. The resilience index ranged from close to 0 up to approximately 0.7 averaging around 0.3 –0.4 for most piezometers. Reliability (REL) was generally high across the network, with most piezometers showing values above 0.8, and many closes to 0.9 or even higher. This suggests that, during the observation period, groundwater levels remained above the critical drought threshold most of the time for most monitoring wells.

When vulnerability (VUL), resilience (RES), and reliability (REL) are considered jointly, contrasting groundwater system behaviors become evident. While several groundwater bodies exhibit high reliability and low vulnerability, resilience often shows greater variability, indicating that systems maintaining acceptable groundwater levels under average conditions may nonetheless experience delayed recovery following drought events. Regional differences further highlight this pattern, with northern areas generally combining higher resilience and lower vulnerability, whereas central and southeastern regions tend to display increased vulnerability and reduced recovery capacity. These results underline the relevance of resilience for assessing groundwater system performance, particularly in the context of groundwater-dependent applications such as Aquifer Thermal Energy Storage (ATES).

3.1.2. *Vulnerability, Resilience and Reliability Correlations*

Regarding the relationships between indices, a strong negative correlation was observed between vulnerability and reliability (Pearson $r \approx -0.88$), confirming that piezometers with lower reliability tended to experience higher vulnerability. The correlation between vulnerability and resilience was weaker (Pearson $r \approx -0.59$). Finally, a moderate positive correlation between reliability and resilience (Pearson $r \approx -0.61$) was found.

The analysis of the relationship between vulnerability (VUL), resilience (RES), and reliability (REL) across groundwater bodies provides essential insights into the general behavior and heterogeneity of aquifer responses (Fig. 5). On average, most groundwater bodies exhibit relatively low vulnerability (VUL_{mean} ranging between approximately −0.01 and −0.30). Reliability (REL_{mean}) is generally high across most groundwater bodies, often close to 1.0. However, resilience (RES_{mean}) shows much greater variability among groundwater bodies. While some aquifers display moderate resilience values (around 0.3–0.4), others show considerably lower resilience.

The potential influence of groundwater levels on aquifer performance indicators was also assessed; the mean water level for each piezometer was calculated and compared to the vulnerability (VUL), resilience (RES), and reliability (REL) indices. However, the correlation analysis revealed very weak relationships between the mean water level and all three indices.

These results demonstrate that vulnerability, reliability, and resilience describe complementary but non-redundant aspects of groundwater system behavior. The identification of groundwater bodies with high reliability but only moderate resilience, as well as systems with moderate vulnerability yet strong recovery capacity, highlights the limitations of evaluating aquifer performance using individual indicators in isolation. For ATES applications, where both short-term operational stability and long-term recovery following drought events are critical, an integrated metric is therefore required to capture the combined effects of drought frequency, severity, and recovery dynamics. This conceptual need directly motivates the development of the Drought Stress Response Index (DSRI).

3.2. *Water body Drought Stress Response Index (DSRI)*

DSRI values across Spain show substantial spatial heterogeneity (Fig. 6), ranging from below 0.1, indicating unstable systems, to above 0.9, associated with highly stable groundwater conditions. A large proportion of piezometers exhibit low Drought Stress Response Index (DSRI) values below 0.4, suggesting that many groundwater systems across Spain are prone to frequent drought events, slow recovery after droughts, or a combination of both.

The critical category (DSRI < 0.1), indicating the least stable systems, is concentrated primarily in central Spain, particularly across the regions of Castilla-La Mancha parts of southern Aragón. This is consistent with previous observations of higher vulnerability and lower resilience in these areas. Intermediate DSRI values (between 0. 3 and 0.6) are widespread across southern and eastern Spain, including parts of Andalusia, the Murcia region, and the interior of Catalonia. These areas exhibit moderate stability but remain sensitive to prolonged droughts or intensive groundwater abstraction, which may progressively reduce their recovery capacity. In

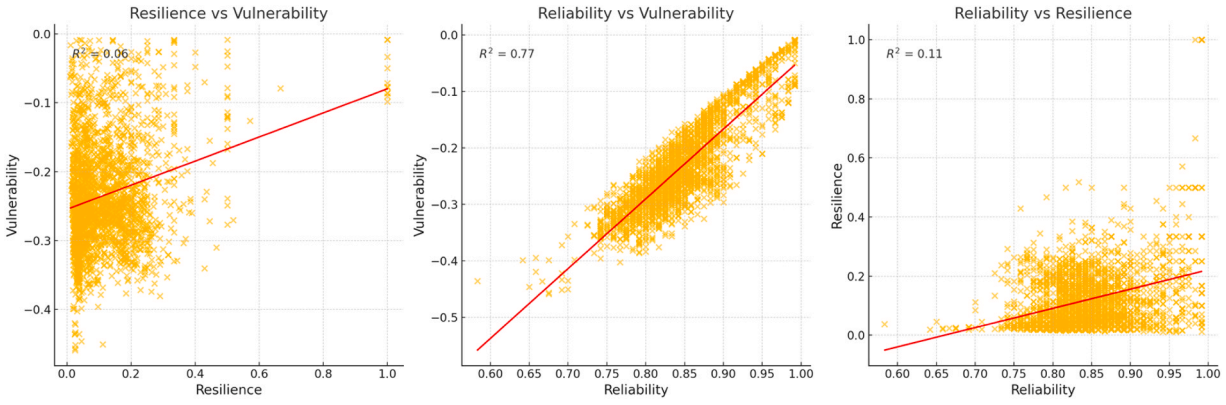


Fig. 5. Correlation between Vulnerability, Resilience and Reliability of the monitored and selected piezometers.

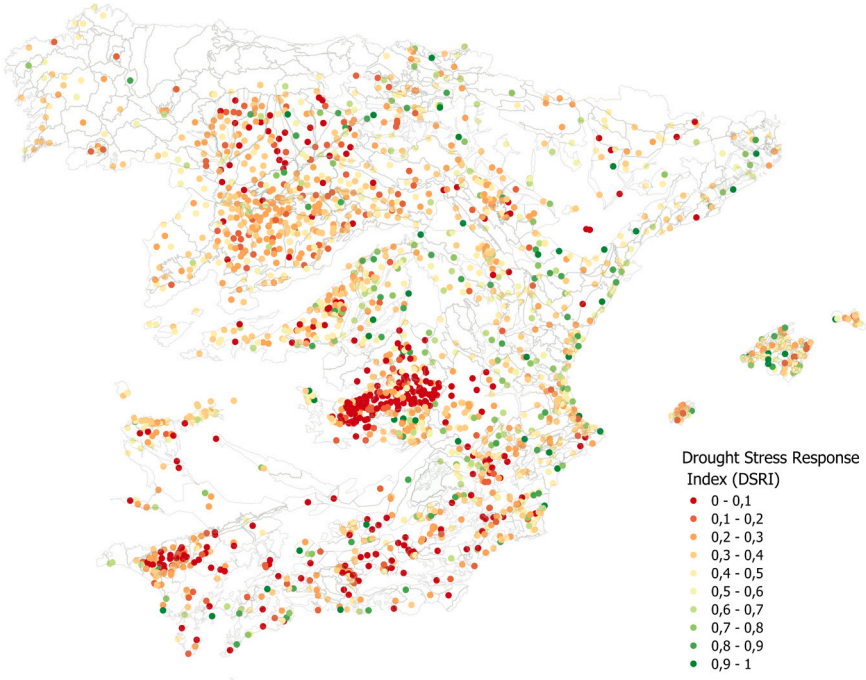


Fig. 6. Drought Stress Response Index (DSRI) for the piezometers as a result of the combination of Vulnerability, Resilience, and Reliability indexes.

contrast, higher DSRI values (above 0.7) are more spatially fragmented, with small clusters in parts of northeastern Spain, some localized areas in Castilla y León, and the Balearic Islands. These locations represent groundwater systems that currently display a higher capacity to withstand and recover from droughts, with greater capacity to withstand and recover from drought stress events.

However, the absence of extensive contiguous areas with high DSRI values highlights the fragmented nature of groundwater resilience across Spain. Overall, the spatial distribution of DSRI closely follows climatic gradients and patterns of groundwater exploitation, underscoring. This highlights the need for region-specific groundwater management and ATES planning strategies that explicitly account for both drought vulnerability and recovery capacity.

3.3. Water table amplitude range and trend analysis

The upper side of the Fig. 7 displays the absolute groundwater-level range recorded by piezometers over the last 20 years together with the long-term trend (declining or rising) of each water body. Piezometric amplitudes span 0–307 m, yet most observations (83%) fall below 20 m, 16% lie between 20 m and 80 m, and only 1.5% exceed 80 m. Groundwater-level trends partition the domain into 310 declining water bodies (rates from $^{-11.45 \text{ to } -0.01 \text{ myr}-1}$; $\sigma = 1.72$ myr-1); 280 rising water bodies ($_{+0.01}$ $^{\text{to}}$ $_{+19.33}$ $^{\text{m}}$ yr^{-1}; $\sigma = 1.36$ myr-1), and 42 hydraulically stable units. Although the mean trend is $^{\text{slightly positive } (+0.11 \text{ m yr-1}), \text{ the large standard deviation (1.64 m yr-1) and the}}$ presence of extreme declines and recoveries reveal pronounced spatial heterogeneity.

Areas that combine both wide groundwater-level ranges and strong negative trends cluster mainly in the south and east of the Iberian Peninsula, confirming that the most dynamic systems are also the most vulnerable to depletion. A noteworthy exception is the aquifers of Mancha Occidental I and II: despite their historic over-exploitation, recent data point to partial stabilization, attributable to stricter pumping controls, the Plan Especial Alto Guadiana, and the growing substitution of groundwater by surface-water supplies.

Conversely, aquifers such as Campo de Cartagena and Poniente Almeriense continue to decline, underscoring the importance of sustained, integrated management.

Bottom part of Fig. 7 superimposes the spatial distribution of groundwater level range and average trends on land-use classes (CORINE (Copernicus Land Monitoring Service, 2024)) and climatic zones (CTE (of Development of Spain, M. M., 2019)). The widest phreatic ranges (20–80 m) are strongly associated with intensive agricultural areas: permanently irrigated cropland (class 212) and complex cultivation patterns (242) dominate in southeastern Spain especially Almería and Murcia where declining trends reveal heavy groundwater abstraction. Similar patterns are observed in non-irrigated arable land (211) around Burgos, northern Málaga, and Huelva, as well as in extensive fruit-tree plantations (222) throughout the Valencian Community. By contrast, olive groves (223) tend to occur in zones with lower groundwater level variability and more stable or rising trends. A notable exception arises in the Mancha Occidental I and II aquifers, where vineyards (221) are the prevailing land use. Here the piezometers record some of the highest phreatic ranges in the country despite the absence of a strong declining trend, implying that large seasonal withdrawals for viti-cultureand subsequent winter recharge drive pronounced water-table fluctuations without sustained depletion. Urban areas (class 111 in black color) introduce additional stress where it overlaps very-hot climatic zones e.g. parts of coastal Andalusia, the Levante corridor

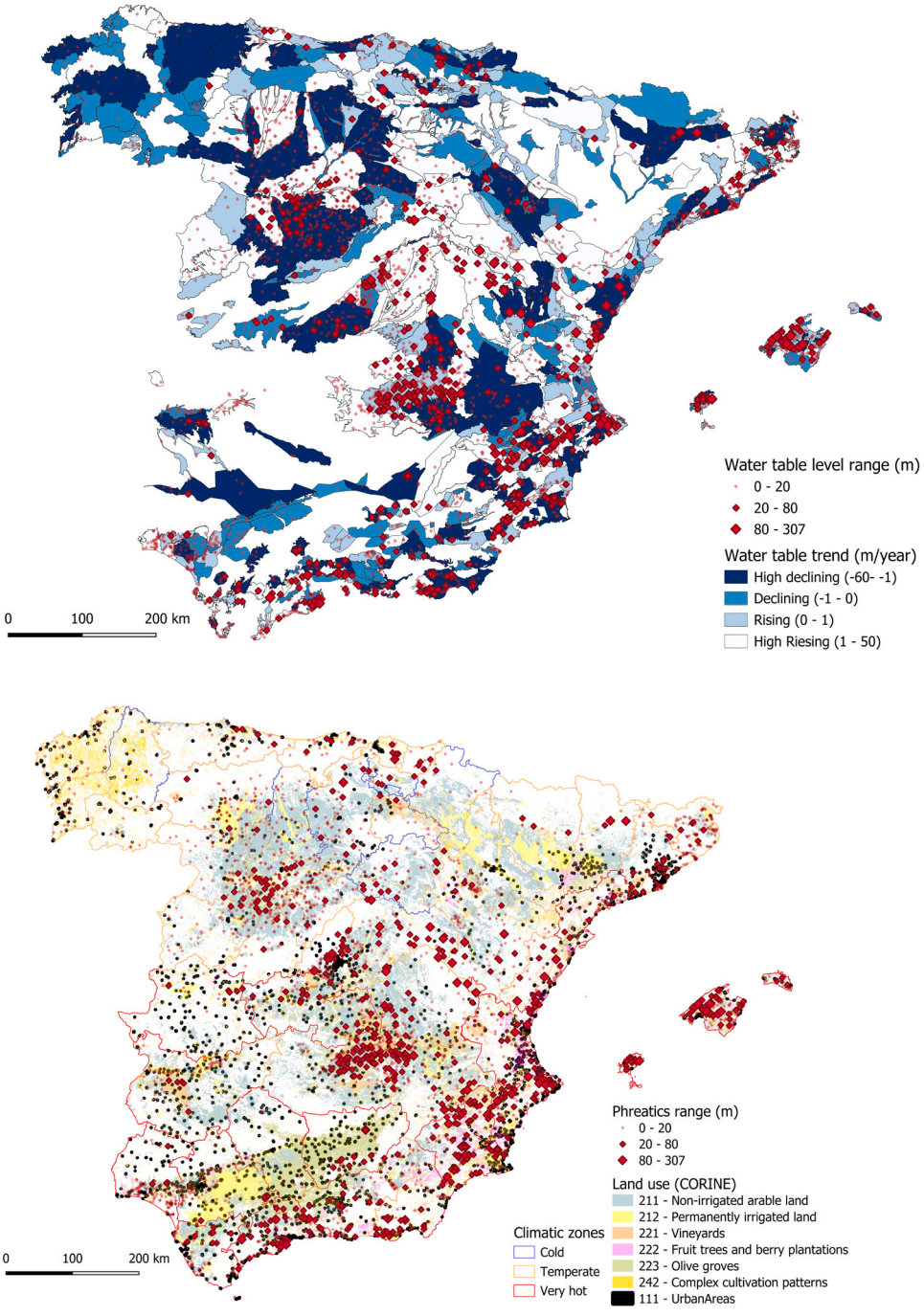


Fig. 7. Spatial trends in groundwater levels across Spain over the last 20 years, categorized as declining or rising (in m/year). The color of each water body represents the average trend, while red diamonds indicate the magnitude of the water table level range (in meters). Bottom. Distribution of water table level range values (same as in a.) overlaid with potential environmental and anthropogenic drivers, including land use types (CORINE classification) and simplified climatic zones (cold, temperate, very hot).

and the southern Madrid conurbation producing localized variability of 20-80 m. By comparison, urban centers in cold or temperate climates along the Cantabrian façade display much steadier water tables, reflecting lower climatic demand and reduced irrigation. It is important to note that land use was not quantitatively included in the suitability classification. Instead, CORINE land cover information was used qualitatively to contextualize observed groundwater variability patterns. This provided valuable interpretative insight into regional differences but did not form part of the classification procedure itself.

Climatic controls reinforce these land-use patterns. Water bodies in very-hot zones, marked by low precipitation and high

evapotranspiration, show the greatest variability and the steepest declines, whereas cold or temperate zones in northern and north-western Spain exhibit modest fluctuations and near-stable trends.

Overall, observed trends in the study area range from −2.5–1.2 m yr^{-1}, far exceeding the typical ±0.2 $^{m\ yr\text{-}1}$ reported for northern and central Europe and matching the −0.5 to −2.0 $^{m\ yr\text{-}1}$ typical of semi-arid aquifers in Spain, Italy and Greece. The comparison highlights the critical situation of large parts of the Iberian Peninsula, where intense agriculture and limited recharge drive some of the fastest groundwater declines recorded in Europe.

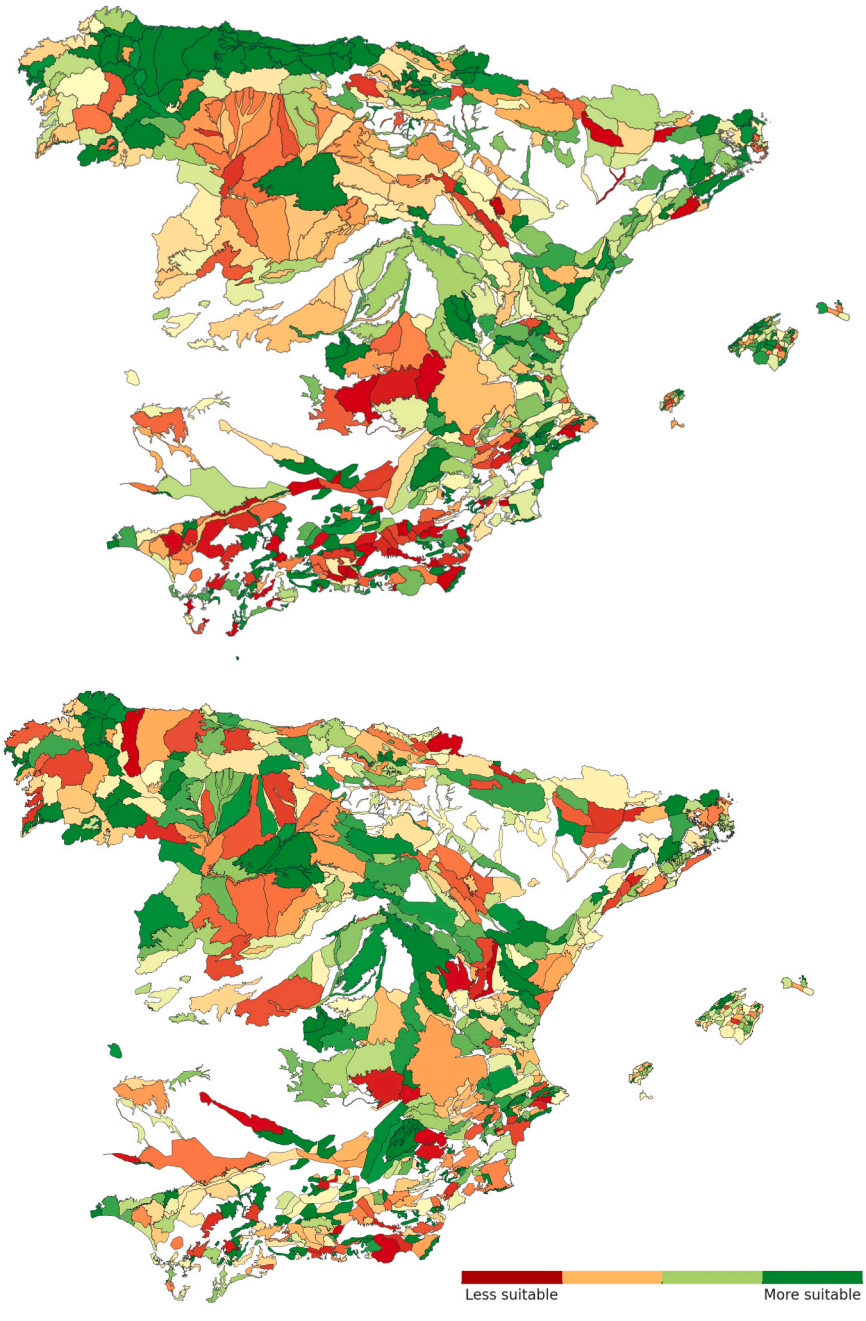


Fig. 8. Spatial distribution of suitability for Aquifer Thermal Energy Storage (ATES) across Spanish groundwater bodies, based on: (Up) the Drought Stress Response Index (DSRI), which reflects the persistence of drought stress over time, and (bottom) groundwater-level variability, measured as the amplitude of seasonal and interannual fluctuations.

3.4. Water bodies suitability for ATES

Groundwater level dynamics play a central role in the feasibility and performance of Aquifer Thermal Energy Storage (ATES) systems, as they directly control both hydraulic and thermal operating conditions. Persistent groundwater depletion reduces the effective saturated thickness of the aquifer, and may limit available storage volume and well screen length, while high abstraction rates increase groundwater velocities, leading to enhanced thermal dispersion and reduced recovery efficiency. As a result, long-term groundwater stress directly constrains ATES operation by reducing storage capacity and system efficiency.

Fig. 8 depicts the results of the suitability analysis for ATES systems based on (a) the Drought Response Index (DSRI), and (b) the range of phreatic level fluctuations. Both maps reveal a clear and consistent north south gradient across Spain. The Atlantic and Cantabrian regions, including cities such as Bilbao, Santander, Oviedo, and Vitoria, are systematically identified as more suitable in both layers. These zones benefit from abundant recharge, limited groundwater abstraction, and low-to-moderate variability, making them ideal candidates for ATES deployment.

Conversely, the southeastern quadrant of the peninsula especially Campo de Cartagena, Bajo Vinalopó, Poniente Almeriense, and Murcia is consistently flagged as least suitable due to a combination of semi-arid climate, intense agricultural pressure, and long-term groundwater declines. These areas are dominated by sharp seasonal drawdowns and low recharge, that reduces the effective saturated screen length and storage capacity, and also by high abstraction rates increase groundwater velocities, leading to enhanced thermal dispersion and reduced recovery efficiency.

Between these two extremes, many inland zones such as Madrid, Zaragoza, and Valencia are rated as less suitable due to strong urban-driven phreatic fluctuations. Other areas, including Barcelona (north), León, and Burgos, appear as moderately suitable, suggesting that ATES systems might operate under specific control measures depending on local hydrogeological context. In such urban environments, large phreatic oscillations are typically associated with pumping cycles and recharge variability, which can result in unstable hydraulic conditions and increased uncertainty in thermal plume behavior.

The comparison between the two approaches highlights that ATES suitability may be constrained by different groundwater dynamics. In groundwater bodies where low DSRI values indicate persistent drought stress, ATES operation is mainly limited by reduced long-term water availability and slow recovery after drought events. In contrast, in areas with higher DSRI but strong seasonal or short-term groundwater-level fluctuations, ATES feasibility is more likely constrained by hydraulic instability and thermal plume dispersion.

Overall, these results confirm that both long-term drought persistence and groundwater-level variability must be jointly considered when assessing ATES suitability. Transitional areas where the two indicators provide contrasting classifications should be interpreted as zones of higher uncertainty, where detailed site-specific hydrogeological analysis and thermo-hydraulic modelling are required prior to ATES deployment.

4. Conclusion and future research

This study developed a national-scale framework to evaluate the suitability of groundwater bodies for Aquifer Thermal Energy Storage (ATES) under conditions of drought and groundwater variability. By combining two complementary approaches a Drought Stress Response Index (DSRI) based on reliability, resilience, and vulnerability metrics, and a range trend analysis of groundwater level fluctuations our methodology captures both chronic and dynamic constraints on aquifer stability. This dual approach provides a systematic screening tool that can be replicated in other regions where groundwater stress and climate variability are critical factors for renewable energy planning.

Application of the framework across Spain revealed a clear spatial pattern. Northern regions, such as Galicia, Asturias, and the Basque Country, exhibited high suitability for ATES due to stable groundwater levels and limited drought impacts. In contrast, southeastern basins, including Segura, Júcar, and the Levante corridor, displayed wide seasonal fluctuations and long-term declines, indicating reduced suitability. These patterns align with climatic gradients and land-use pressures: temperate zones with balanced recharge conditions tend to be more stable, whereas semi-arid regions with intensive irrigation are more vulnerable to drought-induced variability.

The present analysis is subject to several limitations related to data availability and observational depth. Many piezometric time series remain short or contain data gaps, constraining the detection of long-term groundwater trends. In addition, groundwater-level variability was assessed using shallow monitoring wells, whereas ATES systems may operate at variable depths depending on local aquifer conditions. This introduces uncertainty when extrapolating shallow groundwater dynamics to ATES operational performance, particularly in confined or multilayered aquifer systems. The characterization of groundwater related droughts also remains methodologically unsettled; therefore, the groundwater-level metrics applied in this study are intended to support spatial comparison of groundwater dynamics relevant to ATES suitability rather than to provide a formal groundwater drought classification.

Building on these findings, future research will focus on translating the spatial diagnostics presented here into process-based performance assessments. A second research phase will involve detailed thermo-hydraulic modelling of representative groundwater bodies to evaluate whether aquifers classified as drought-stressed or hydraulically unstable exhibit reduced thermal recovery, lower reliability, or decreased operational efficiency under realistic ATES operating conditions. A key research question concerns the extent to which observed phreatic-level variability propagates to the depths typically used for ATES, and under which hydrogeological settings vertical decoupling may occur.

Overall, this work provides an operationally relevant, spatially explicit screening tool to guide the identification of promising aquifers for ATES deployment in drought-prone regions. While the present framework already delivers actionable insights for strategic planning, future studies can build on these results by integrating thermo-hydraulic modelling to assess site-specific performance. This

staged approach combining national-scale screening with local-scale validationensures that renewable energy infrastructure is designed with both climatic resilience and hydrogeological feasibility in mind.

CRediT authorship contribution statement

Ramos Escudero Adela: Writing – review & editing, Writing – original draft, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Carlos Toledo:** Data curation. **Martin Bloemendal:** Methodology, Formal analysis, Conceptualization. **Juan-de-Dios Gómez-Gómez:** Data curation. **Collados-Lara Antonio:** Validation, Supervision, Resources, Methodology. **David Pulido-Velazquez:** Supervision, Methodology, Formal analysis, Conceptualization.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Adela Ramos Escudero reports financial support was provided by European Commission. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

Acknowledgements

This project has received funding from the European Union's Horizon Europe research and innovation program under the Marie Sklodowska-Curie agreement No. 101149399. IN addition, this research was partially funded by the project: SIGLO-PRO (PID2021 – 128021OB-IOO/AEI/10.13039/501100011033/FEDER,UE)

Appendix

Data selection and processing

Water table historical series data selection and processing

A minimum of 20 years of data was needed to calculate the indexes above. The complete time series for the piezometers spans from 1965 to 2023, though the temporal coverage varies across different watershed authorities. To start with, the 20-years period with the highest amount of data was selected, from 2001 to 2021. Fig. 9 depicts an histogram with the amount of measurements for all the piezometers along the historical data series, and in blue color, those years selected. This way, the years covered of the selected piezometer goes from 1 to 21 years, and the average data years of all the piezometers is 11.7 years. Even within the period selection, the dataset exhibits significant temporal variability regarding the number of years covered, the number of active piezometers, and the distribution across river basins. As for annual records, the number of phreatic level measurements increases along the data series, starting at approximately 15000 data points in 2001 peaking at more than 50,000 in 2021. There is a slight decreasing in 2020 due to the pandemic limitations. In addition, as aforementioned, there is a decreasing in the number of measurements and unique piezometers due to the implementation of the Water Management Directive in the River basin Authorities in Spain, which can be seen both in Fig. 10. After this phenomenon, the number of measurements recovers around 2010.

This data selection was used both for DSRI and water table range analyses.

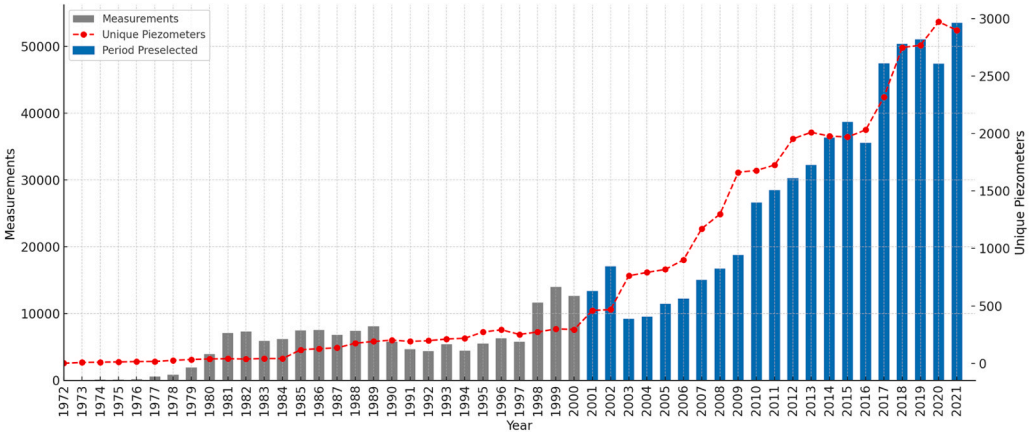


Fig. 9. Phreatic measurements along the historic data series and the selected period

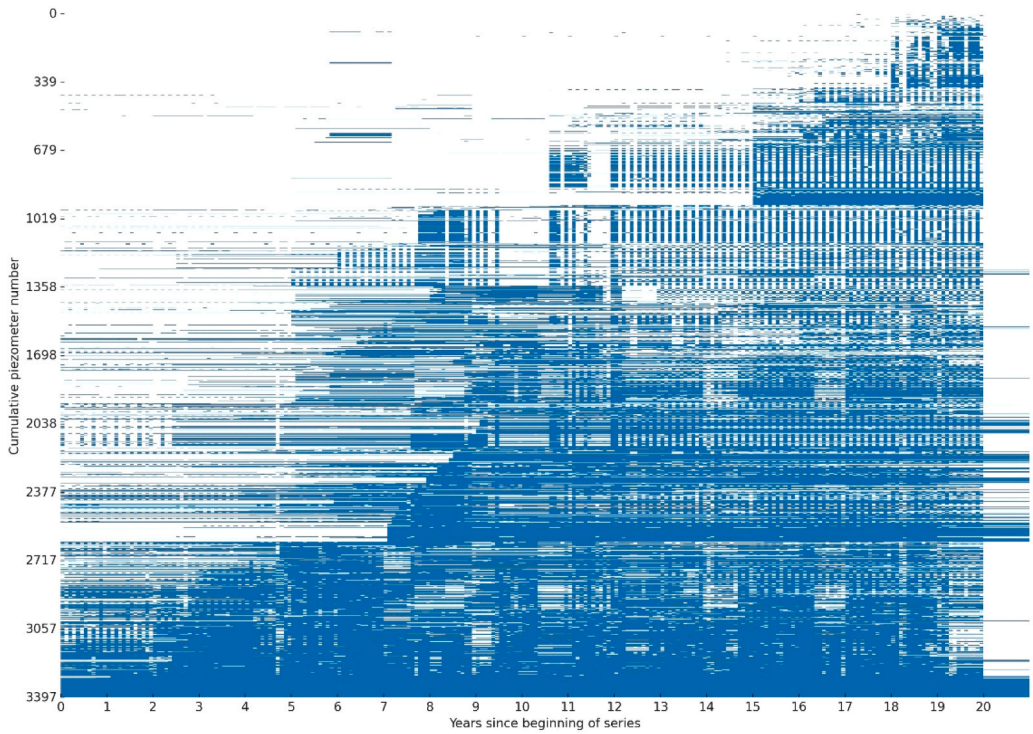


Fig. 10. Temporal data availability by piezometer (monthly resolution)

Year's period and gap allowed selection

In addition, the selected time frame displays substantial temporal variability in the number of active piezometers, as well as significant gaps in the data series (Fig. 11). To systematically evaluate the trade-off between temporal and spatial coverage in the monitoring of groundwater bodies, while prioritizing the inclusion of the highest possible number of piezometers and water bodies with acceptable data continuity, a sensitivity analysis was performed. In this analysis, temporal coverage is defined through a set of scenarios characterized by both the duration of the data series (ranging from 10 to 20 years) and their specific positioning within the full observational record. In all cases, the maximum allowable gap length is constrained to 10% of the total duration. The results are illustrated in Fig. 11, where each horizontal bar represents one temporal scenario and its corresponding time span. Dots indicate the number of piezometers included in each scenario, with red dots highlighting those with the highest coverage. The color gradient of each bar reflects the number of groundwater bodies included, with darker green indicating higher spatial coverage. The optimal scenariodefined by a dark green bar with a red dot achieves the best balance between spatial and temporal criteria, covering 586 groundwater bodies over a 10-year period (2011–2021) while maintaining robust data continuity, including a maximum allowable gap of one year. As expected, increasing the length of the time series results in a reduction in spatial coverage, both in terms of piezometers and groundwater bodies, due to stricter continuity constraints. Conversely, shorter time windows allow the inclusion of more monitoring points, enhancing spatial representativeness. From a hydrogeological perspective, however, a 10-year data window may be insufficient to fully capture the long-term dynamics of groundwater systems, particularly with regard to drought assessment. Groundwater responses to climatic variability often exhibit significant time lags and memory effects due to the slow movement of water through the subsurface and the buffering capacity of aquifers. As such, droughts may unfold over multi-decadal scales, and their full impact especially in deep or low-productivity aquifers might not be observable within a single decade. Shorter time series may therefore overlook important slow-onset trends or fail to contextualize anomalous recharge or abstraction events. Consequently, while a 10-year period offers a practical compromise between data completeness and spatial representativeness, it may limit the robustness of analyses aiming to evaluate long-term groundwater sustainability or the full magnitude and recurrence of drought episodes.

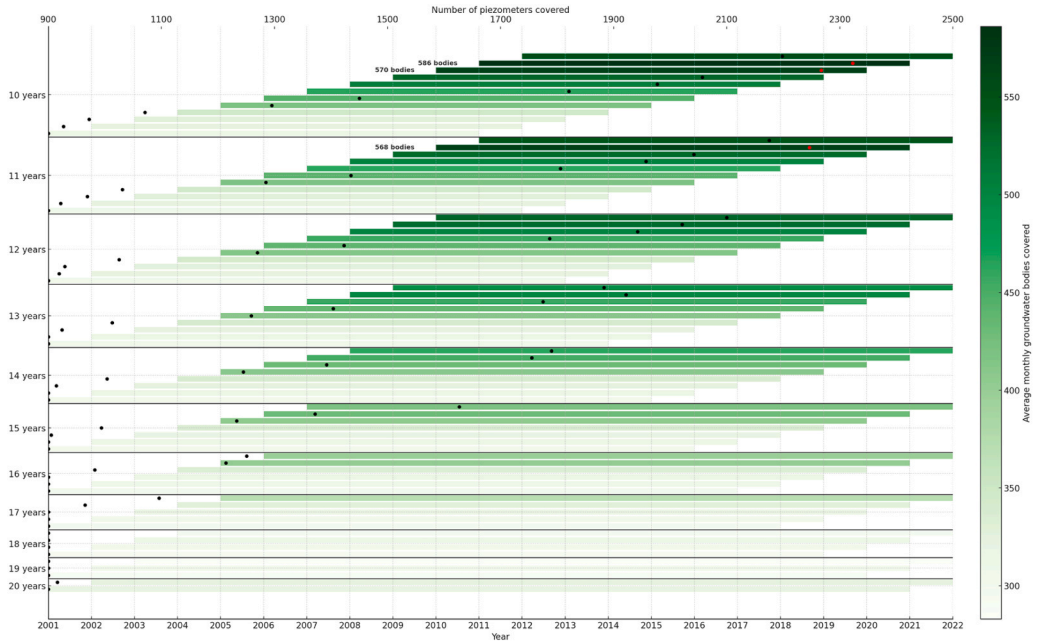


Fig. 11. Spatial coverage vs temporal coverage sensitive analysis. Average monthly number of groundwater bodies covered by piezometric data for different minimum time series lengths (10–20 years) between 2001 and 2021

Gaps interpolation

Although the trustworthiness of corrections depends on the availability of original measurements and the stability of the aquifer system, the results obtained from the dataset demonstrate a clear improvement in estimation accuracy when applying the Mean Year Correction method.

Specifically, while interpolated values may follow a simple flat or linear pattern disconnected from seasonal dynamics, the corrected series integrates temporal structure drawn from years with complete data, producing more realistic curves. This is particularly evident at the start of 2012 (Fig. 12), where no original measurements were available. The initial interpolation filled the gap with constant values, failing to reflect any seasonal variation. However, the corrected values successfully reproduced a plausible seasonal pattern, closely aligned with what would be expected based on the behavior of the aquifer in subsequent years.

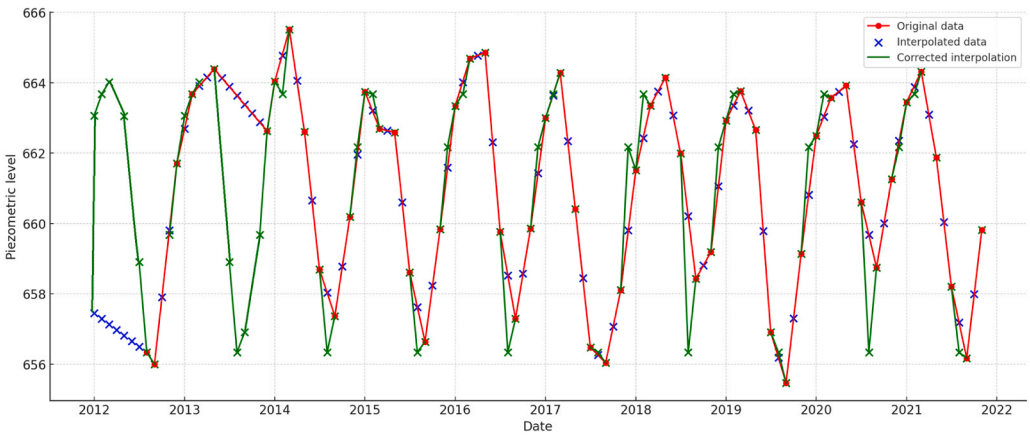


Fig. 12. Original and interpolated piezometric levels for one piezometer from 2012 to 2022. Red circles and lines represent the original measurements. Blue crosses indicate interpolated values to fill temporal gaps. Green stars and lines represent corrected interpolations using the Mean Year Correction method

Correction of data anomalies

To improve the accuracy of groundwater dynamics characterization, a preliminary correction was applied to refine both phreatic range estimates and long-term trend calculations. Several piezometers exhibited artificially high variability due to anomalous values

likely caused by sensor errors or isolated data issues. Outlier detection was performed using a standard statistical threshold, whereby values falling outside the interval defined in Eq. 5 were classified as anomalous. This analysis revealed that approximately 12% of the piezometers (224 in total) contained such outliers, which were subsequently removed. A new value of phreatic range was then calculated for each affected series, reflecting only physically plausible groundwater level variations. On average, the original phreatic range was 203.9 m, with some cases exceeding 300 m. After outlier correction, the mean range was reduced to 51.3 m, highlighting an average outlier impact of 152.6 m. Although the number of outliers per time series was low (mean of 0.67), their influence on minimum and maximum values and therefore on range and slope estimates was substantial. This correction step was essential to ensure that both range-based indicators and long-term groundwater trends reflected true aquifer behavior rather than artefacts. A summary of these results is presented in Table 1.

Table 1
Summary statistics of phreatic range correction across selected piezometers

Metric	Mean (m)	Std Dev (m)	Min (m)	Max (m)
Original Minimum Level	80.75	125.29	1.38	225.18
Original Maximum Level	284.65	40.12	238.33	308.68
Original Phreatic Range	203.9	165.39	13.15	307.3
Corrected Minimum Level	80.75	125.29	1.38	225.18
Corrected Maximum Level	132.07	95.27	54.28	238.33
Corrected Phreatic Range	51.32	37.41	13.15	87.91
Outlier Impact on Range	152.58	134.58	0	254.4
Number of Outliers per Series	0.67	0.58	0	1

Data availability

Data will be made available on request.

References

Bloemendal, M., Hartog, N., 2018. Analysis of the impact of storage conditions on the thermal recovery efficiency of low-temperature ATES systems. Geothermics 71, 306–319.

Bloemendal, M., Olsthoorn, T., van de Ven, F., 2014. Combining climatic and geo-hydrological preconditions with spatial demand patterns for aquifer thermal energy storage (ATES). Energy Build. 68, 203–887, 212.

Bloomfield, J.P., Marchant, B.P., 2013. Analysis of groundwater drought building on the standardized precipitation index approach. Hydrol. Earth Syst. Sci. 17 (12), 4769–4787.

Bonte, M., Stuyfzand, P.J., Van den Berg, G.A., Hulscher, T.E.M., 2011. Effects of aquifer thermal energy storage on groundwater quality: A review. Hydrogeol. J. 19, 1141–1156.

Copernicus Land Monitoring Service (2024). Corine land cover (clc) - Copernicus land monitoring service. ⟨https://land.copernicus.eu/en/products/corine-land-cover⟩. Accessed: 2025-06-18.

Dai, Z., Wang, Y., Zhang, Z., Zhang, L., Chen, B., 2019. Quantifying groundwater recharge and discharge processes in the heihe river basin based on multi-objective calibration. J. Hydrol. 570, 503–513.

Döll, P., Fiedler, K., 2009. Global-scale modeling of groundwater recharge. Hydrol. Earth Syst. Sci. 13 (6), 885–899.

European Environment Agency (2018). European Waters: Assessment of Status and Pressures 2018. European Environment Agency. Accessed: May 2025.

Everitt, B.S., 2002. The Cambridge Dictionary of Statistics, 2nd ed.,. Cambridge University Press, Cambridge.

Famiglietti, J.S., Ferguson, G., Gleeson, T., Taylor, R.G., Wada, Y., 2024. Groundwater is vanishing: a global synthesis of groundwater trends, drivers, and socioecological impacts. Nat. Synth. 3, 60–75.

Fiorillo, F., 2010. Groundwater resource decrease in southern Italy during drought periods: Natural and anthropogenic effects. Hydrogeol. J. 18 (5), 1309–1320.

Fiorillo, F., Guadagno, F.M., 2012. Long-term trend and drought proneness of karst groundwater resources in a mediterranean region of southern Italy. Hydrogeol. J. 20 (5), 1009–1021, 52.

Fowler, K.J., et al., 2023. Rapid groundwater decline in dryland cropland regions globally. Nature 617, 823–829.

Gómez, L., Vargas, A., Rodríguez, M., Herrera, M., Custodio, E., 2022. Rescue of groundwater level time series: How to visually identify and treat errors. J. Hydrol. 605, 127310.

Heekeren, V. v, Bakema, G., 2015. The Netherlands country update for 2015. World Geothermal Congress 2015 (Melbourne, Australia).

Herrera-Pantoja, M., Hiscock, B.R., 2008. The effects of climate change on potential groundwater recharge in great Britain. Hydrol. Process. 22 (1), 73–86.

Loucks, D.P., 1997. Quantifying trends in system sustainability. Hydrol. Sci. J. 42 (4), 513–530. Published online: 25 December 2009.

Mays, L.W., 2013. Groundwater resources sustainability: Past, present, and future. Water Resour. Manag. 27 (13), 4409–4424. Published online: 20 August 2013.

McKee, T.B., Doesken, N.J., Kleist, J., 1993a. The relationship of drought frequency and duration to time scales. Proceedings of the 8th Conference on Applied Climatology. American Meteorological Society, Anaheim, California, pp. 179–184.

Mendicino, G., Senatore, A., Versace, P., 2008a. A groundwater resource index for drought monitoring and forecasting in a Mediterranean basin. J. Hydrol. 357, 282–302.

Mendicino, G., Senatore, A., Versace, P., 2008b. A groundwater resource index (gri) for drought monitoring and forecasting in a mediterranean climate. J. Hydrol. 357 (3-4), 282–302.

Millán, M.M., Estrela, M.J., Miró, J., 2004. Rainfall components: Variability and spatial distribution in a Mediterranean area. J. Clim. 18 (14), 2424–2443.

Ministry for the Ecological Transition and the Demographic Challenge (2025b). Water Bodies - Hydrological Planning Cycle 2022-2027. Last accessed: March 20, 2025.

Nlend, B., Huneau, F., Garel, E., Santoni, S., Leydier, T., 2025. Rainfall and groundwater isotopes to investigate aquifer recharge mechanisms and seasonality in the Mediterranean region: Implications for long-term groundwater sustainability in the climate change context. J. Hydrol. Reg. Stud. 61, 102671.

of Development of Spain, M. M (2019). Basic document he: Energy saving. Technical Building Code (CTE).

Peña, A., Mondragón, J., 2024. La gobernanza de las cuencas hidrográficas a partir de la directiva marco del agua. incremento de funciones y pérdida de policy capacity de las confederaciones. Gestión y Análisis de Políticas Públicas.
Pulido-Velazquez, D., Romero, J., Collados-Lara, A.-J., Alcalá, F.J., Fernández- Chacón, F., Baena-Ruiz, L., 2020. Using the turnover time index to identify potential strategic groundwater resources to manage droughts within continental Spain. Water 12 (11), 3281.
Pulido-Velazquez, M., Peña-Haro, S., García-Prats, A., Mocholi-Almudever, A.F., Henriquez-Dole, L., Macian-Sorribes, H., 2015. Integrated assessment of groundwater use under climate change. J. Hydrol. 528, 605–618.
QGIS Development Team (2025). QGIS Geographic Information System. Open Source Geospatial Foundation. Version 3.x.
Ramos-Escudero, A., Bloemendal, M., 2022. Assessment of potential for aquifer thermal energy storage systems for spain. Sustain. Cities Soc. 81, 103849.
Richter, D., Therrien, R., 2022. Influence of low-frequency variability on high and low groundwater levels in european aquifers. Hydrol. Earth Syst. Sci. 26, 4779–4795.
Saïd, M., Stadnyk, T., Koenig, K., Demuth, M., 2020. Changes in seasonality of groundwater level fluctuations in a temperate-cold climate transition zone. J. Hydrol. Reg. Stud. 30, 100707.
Smakhtin, V.U., 2001. Low flow hydrology: a review. J. Hydrol. 240 (3- 4), 147–186.
Sommer, W.T., Bakr, M., Zuurbier, K.G., Leusbrock, I., Grotenhuis, T., Rijnaarts, H.H., 2013. Efficiency of and interference among multiple aquifer thermal energy storage systems: A Dutch case study. Renew. Energy 60, 53–62.
Stemmle, M., Bayer, P., Blum, P., 2021. Influence of groundwater flow on the thermal performance of aquifer thermal energy storage systems. J. Hydrol. 603, 126931.
Tallaksen, L.M., Van Lanen, H.A., 2004. Hydrological Drought: Processes and Estimation Methods for Streamflow and Groundwater. Elsevier, Amsterdam.
Taylor, R.G., et al., 2024. Multi-decadal groundwater observations reveal declining trends in southwestern Europe. Nat. Geosci.
Taylor, R.G., Bonsor, C.D., Scanlon, M.A., MacDonald, K.M., Schröder, L.M., Koussis, Y., Treidel, J.C., Eckhardt, A.M., 2013. Ground water and climate change. Nat. Clim. Change 3, 322–329.
Van Loon, A.F., Laaha, G., 2017. Drought in a human-modified world: Reframing drought definitions, understanding, and analysis approaches. Hydrol. Earth Syst. Sci. 21 (9), 3631–3650.
Van Loon, A.F., Van Lanen, H.A.J., 2013. Making the distinction between water scarcity and drought using an observation-modeling framework. Hydrol. Earth Syst. Sci. 17, 4769–4786.
Wilks, D.S., 2011. Statistical Methods in the Atmospheric Sciences. volume 100 of International Geophysics Series, 3rd edition. Academic Press.
Zipper, S.C., Jasechko, S., Perrone, D., 2019. Groundwater pumping impacts on terrestrial ecosystems across the united states. Nat. Sustain. 2 (10), 1–8 (Zhu, K.,).