

RESPONSE LETTER

Manuscript Number: egusphere-2026-1378

Title: Managed Aquifer Recharge in Confined Multi-Layer Aquifers: A Scalable Framework for Drought Resilience in Central Europe

Corresponding Author: Abdelrahman Ahmed Ali Abdelrahman

All Authors: Abdelrahman Ahmed Ali Abdelrahman, Hagen Koch, Mobarok Hossain, Ronjon Heim, Clara Hauke, and Irina Engelhardt

Response to the Reviewer #1:

This is an interesting study addressing a real gap in MAR research for confined aquifers in Central Europe. The integration of three methodological components into a single transferable framework is a good contribution. The manuscript is generally well-written and clearly structured. However, several issues require attention before publication.

Reviewer's comments:

1) The authors report a high pearson correlation between SPEI-12 and SGI with "no lag". This is an interesting finding for a confined multi-layer aquifer system, where pressure propagation rather than direct recharge drives head changes. The manuscript does not sufficiently explain the hydrogeological mechanism behind this near-instantaneous coupling. In confined systems, piezometric responses can be rapid, but this is a different mechanism than moisture storage response observed in unconfined systems. This distinction must be clearly discussed and reconciled with the literature.

It would be more correct, rather to state that the absence of lag between SPEI 12 and SGI "indicates that groundwater levels in this confined aquifer respond almost immediately to climatic anomalies" (Discussion section, lines 487-488) to state that the time scale of the response process of groundwater levels to climatic anomalies is not detectable by the aggregated annual scale of SPEI 12 compared with SGI.

Due to the importance of this statement in discussion and conclusion section, to avoid misinterpreting of results and findings, a montly based analisys using SPEI 1 would be recommended to eventually detect and quantify lags or strongly confirm the "immediate response of GW to climate variability" statements of the whole manuscript.

Authors' response

We thank the reviewer for this important and constructive comment. We agree that the previous wording could be misinterpreted. We have changed the wording to avoid misinterpretation, clarifying that the absence of a detectable lag at the SPEI-12 scale does not demonstrate an immediate physical recharge response. Rather, it indicates that the response time of groundwater levels to climatic anomalies is not resolved by the 12-month accumulated SPEI timescale. Because SPEI-12 already integrates the climatic water balance over the preceding year, a zero-lag correlation means that SGI co-varies with annual-scale climatic anomalies, not that groundwater storage responds instantaneously.

Following the reviewer's suggestion, we performed a monthly lag-correlation analysis using SPEI-1, SPEI-3, SPEI-6, and SPEI-12 against SGI for lags of 0–24 months. The SPEI-1 analysis showed only weak correlations with SGI ($r = 0.108$ at 3-month lag), whereas correlations increased for longer accumulation periods and were strongest for SPEI-12 at zero lag ($r = 0.632$, $r^2 = 0.40$). This supports the interpretation that the groundwater system primarily reflects accumulated climatic anomalies over multi-month to annual timescales rather than individual monthly anomalies (please see new Table S2 in Supplementary Materials).

We have also added a hydrogeological explanation specific to the confined aquifer setting. In this system, SGI represents piezometric head variations. These may respond rapidly to pressure propagation and changes in boundary conditions, but this mechanism differs from direct recharge or soil-moisture storage responses typical of unconfined aquifers. We therefore revised the Discussion and Conclusion to distinguish between piezometric response and direct groundwater recharge/storage response.

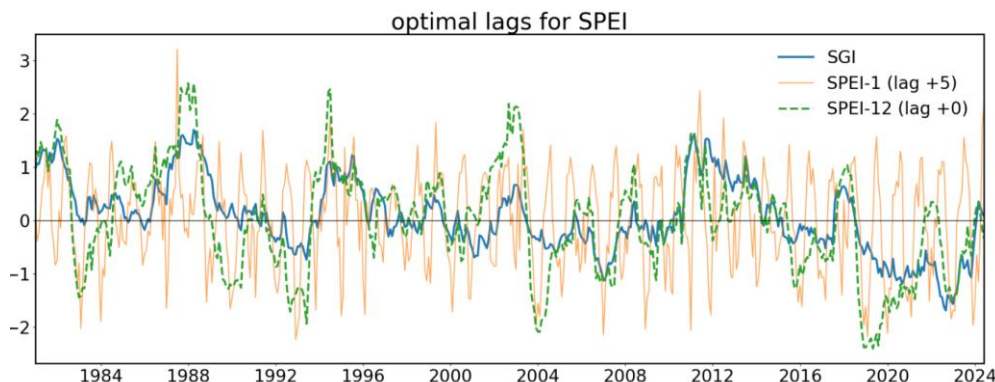
Changes made in the manuscript:

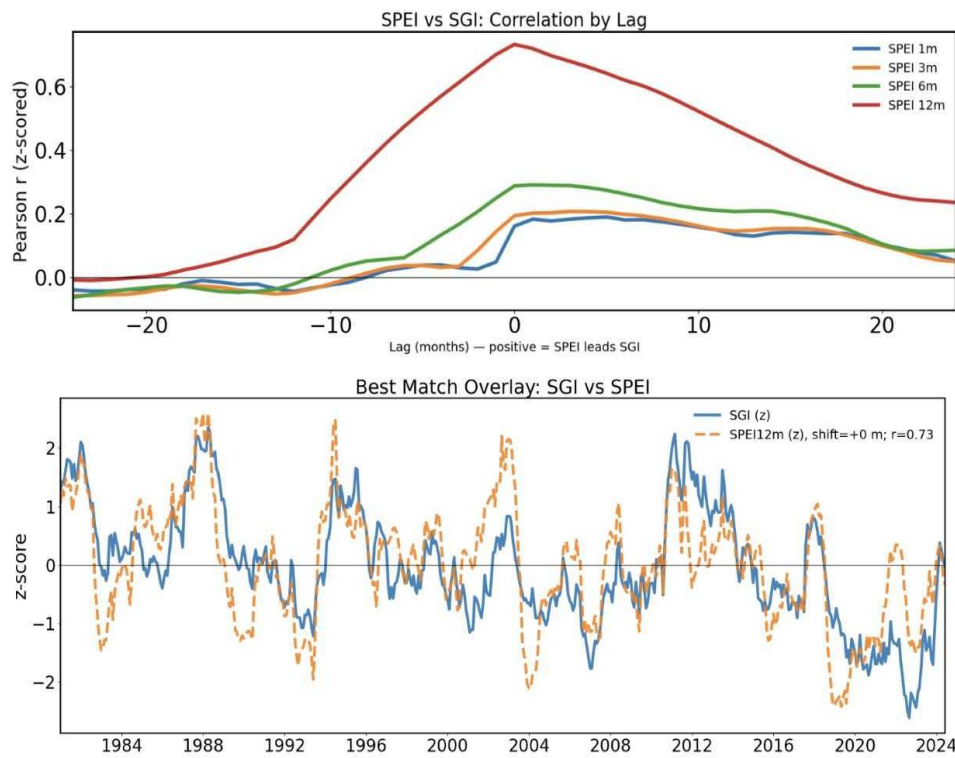
Discussion

The maximum SPEI-12–SGI correlation at zero lag does not indicate instantaneous recharge. Because SPEI-12 already represents the accumulated climatic water balance of the previous 12 months, the zero-lag correlation shows that groundwater anomalies co-vary with annual-scale climatic conditions. The weak SGI correlations with SPEI-1, which increased with longer accumulation periods and peaked for SPEI-12 (Table S2), further indicate that the confined multi-layer aquifer responds mainly to accumulated climatic deficits rather than individual monthly anomalies. Moreover, SGI in this confined aquifer primarily reflects piezometric head changes, which can respond rapidly to pressure propagation and changing boundary conditions without immediate movement of newly recharged water.

Table S2. Lag-correlation analysis for SPEI-1, SPEI-3, SPEI-6, and SPEI-12 vs SGI

Index	Best lag	Pearson r	r ²	n	Interpretation
SPEI-1 vs SGI	5 months	0.191	0.036	533	Very weak
SPEI-3 vs SGI	3 months	0.208	0.043	531	Very weak
SPEI-6 vs SGI	1 month	0.291	0.085	528	Weak
SPEI-12 vs SGI	0 month	0.734	0.539	522	Strongest





Conclusion

First, hydroclimatic and hydrogeologic indicators (SPEI-12 and SGI) show that groundwater-level anomalies in the confined multi-layer system are strongly coupled to annual-scale climatic water-balance anomalies, although the actual response time cannot be resolved from the SPEI-12 zero-lag relationship alone.

2) Section 3.2, Line 200-206: Application of SGI to this context should be explained properly and with more details. Some parts of this paragraph are unclear, for example you state that "a common preprocessing step is to remove non-stationarity" but it's not clear if you proceeded this way or not.

Authors' response: (M, H)

We thank the reviewer for pointing out that the original description of the SGI calculation was not sufficiently clear. We have revised Section 3.2 to explain the SGI method in more detail.

In the revised manuscript, we now specify that SGI was calculated following the non-parametric monthly method of Bloomfield and Marchant (2013). Groundwater levels were ranked separately for each calendar month and transformed to standard-normal scores using the inverse normal distribution. This month-wise standardization removes the recurring seasonal cycle from the groundwater-level records.

We also clarified the issue of non-stationarity. No additional detrending was applied. Long-term groundwater-level trends were intentionally retained because they form part of the groundwater drought and depletion signal relevant to this study. The previous sentence referring generally to "removing non-stationarity" has therefore

Changes made in the manuscript:

Section 3.2

Groundwater drought was quantified using the non-parametric monthly Standardized Groundwater Index (SGI) of Bloomfield and Marchant (2013). This approach is suitable for groundwater-level data because it does not assume a specific probability distribution and accounts for seasonal variability. For each well, monthly groundwater-head records were arranged as a continuous time series, with missing values excluded only from ranking. SGI was calculated separately for each calendar month, so January values were compared

only with other January values, February values with February values, and so on. Within each month, groundwater levels were ranked, converted to plotting positions, and transformed to standard-normal scores:

For a given calendar month, the SGI was calculated as:

$$SGI = \Phi^{-1}(p_i)$$

$$p_i = \frac{i - 0.5}{n}$$

where Φ^{-1} is the inverse standard-normal cumulative distribution function and p_i is the plotting position i is the rank, and n is the number of observations for that month. Negative SGI values indicate below-normal groundwater levels, whereas positive values indicate above-normal conditions. Month-wise standardization removes the seasonal cycle, while no additional detrending was applied because long-term groundwater decline was considered part of the drought and depletion signal. SGI was first calculated for individual wells and then spatially aggregated for the entire catchment and its 21 sub-catchments

3) Globally, a flowchart image of the methodology would enrich methodological section.

Authors' response:

We thank the reviewer for this valuable suggestion. We agree that a flowchart significantly improves the clarity of our complex methodology. We have added a comprehensive methodological flowchart (Fig. 4) that illustrates the sequential structure of our framework, starting from data processing, through the four main analytical components (drought assessment, water balance, GIS-MCDA, and flow-threshold analysis), culminating in the final ASTR suitability and economic outputs. A reference to this figure has been added to Section 3.

Changes made in the manuscript:

Section 3.2

The overall structure and sequential progression of the methodological framework are illustrated in the flowchart (Fig. 4), which integrates hydroclimatic drought assessment, GIS-MCDA site suitability, surface water availability, and economic screening.

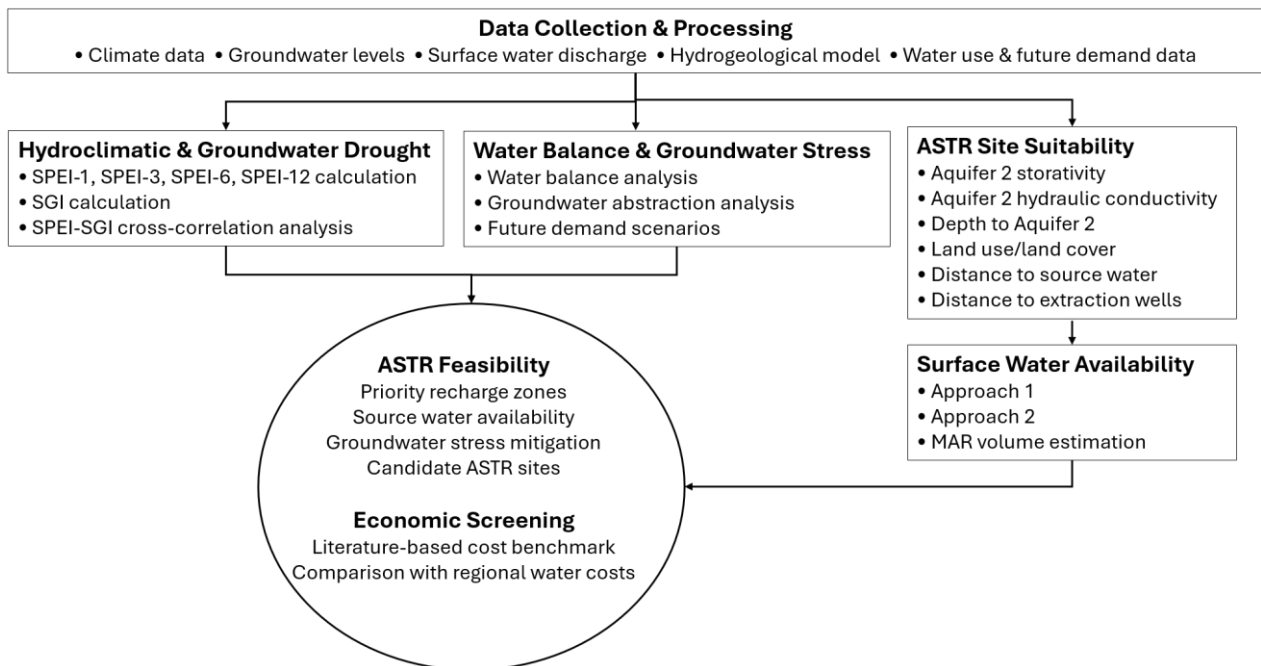


Figure 4. Methodological flowchart illustrating the sequential structure of the integrated ASTR feasibility assessment framework, from data processing and analytical components to final suitability and economic outputs.

4) In section 3.4, regarding the MCDA procedure for selecting site suitabilities for ASTR and MAR, you selected six site-selection criteria and classify their suitability scores. MCDA applications in groundwater related Issues, such as groundwater potential recharge evaluation and mapping, are very common in scientific literature of recent Years; often exploring and selecting different ranges of criteria due to case specific needs or constraints. In your case, criteria selection seems to be arbitrary or author-knowledge based, as well as the attributed scores in table 1 and their range values, which can be acceptable only if properly justified for the characteristics of your case study and strongly linked to scientific literature and to the purpose of your study. In any case, your choices have to be properly framed in the context of contemporary scientific literature on the topic. From both aspects, this methodological section of the manuscript is weak and needs improvements.

Authors' response:

We thank the reviewer for this valuable feedback. We have substantially revised Section 3.4 to explicitly ground all six criteria and their scoring thresholds in recent peer-reviewed MAR literature (e.g., Russo et al., 2015; Sallwey et al., 2019; Dillon et al., 2019; Seidl et al., 2024; Sitek et al., 2026), adding 13 new references.

To address concerns about arbitrary weighting, we performed a sensitivity analysis using three distinct weighting scenarios (equal weight, hydro-heavy, and distance-heavy). Results demonstrate high robustness: the "Very Low" suitability class remained stable at ~29% across all scenarios (variation <0.3%), and high-suitability zones stayed in the same geographic locations. This confirms that identified priority zones are controlled by actual physical constraints rather than subjective weight assignments (Hani et al., 2023; Vandala and Mahed, 2025). We have updated Sections 3.4 and 4.3, added Tables 2–3, and revised Figure 6 to present these findings.

Changes made in the manuscript:

Section 3.4

We applied a GIS-based multi-criteria decision analysis (MCDA) to map the suitability of the study catchment for ASTR implementation, following established MAR/ASTR frameworks (Khalil et al., 2022; Sallwey et al., 2019; Russo et al., 2015). We defined six key site-selection criteria (Table 1) based on their demonstrated influence on recharge efficiency in comparable hydrogeological settings (Gibson et al., 2018; Dillon et al., 2019; Seidl et al., 2024). These criteria included Aquifer 2 storativity (specific storage $\times$ thickness), hydraulic conductivity (K) derived from the 3D hydrogeological model, depth to Aquifer 2, land use, and the locations of rivers and extraction wells. We assigned suitability scores ranging from 1 (unsuitable) to 5 (highly suitable) using thresholds widely applied in MAR studies (Russo et al., 2015; Gibson et al., 2018; Abdo et al., 2024). We also applied exclusion masks to account for absolute constraints, including surface water bodies and Well Protection Zones 1-2.

3.4 ASTR/MAR Site Suitability (MCDA)

Aquifer 2 storativity: We calculated Aquifer 2 storativity as the product of specific storage (Ss) and layer thickness (b) ($S = Ss \times b$) to represent actual storage capacity, following methods established in MAR site assessments (Russo et al., 2015; Gibson et al., 2018). We derived specific storage values from lithology-based estimates within the 3D hydrogeological model and multiplied by the actual Aquifer 2 thickness for each cell. Suitability scores ranged from 5 ($>5 \times 10^{-3}$) to 1 ($<1 \times 10^{-4}$), consistent with storage capacity thresholds identified in successful MAR applications (Seidl et al., 2024; Vandala and Mahed, 2025).

Hydraulic conductivity (K): We prioritized areas with higher K values, as they allow water to infiltrate and spread more effectively, a critical factor for ASTR efficiency (Dillon et al., 2019; Meng et al., 2024). Suitability scores ranged from 5 (>30 m/d) to 1 (<5 m/d), aligning with thresholds commonly applied in MAR feasibility studies (Russo et al., 2015; Sitek et al., 2026).

Depth to Aquifer 2: We assessed the depth from land surface to the aquifer top, as greater depths increase injection costs and reduce economic feasibility (Gibson et al., 2018; Abdo et al., 2024). Scores ranged from 5

(highly suitable, <20 m) to 1 (unsuitable, >50 m), following recommendations from MAR implementation guidelines (Stefan and Ansems, 2018; Seidl et al., 2024).

Land Use/Land Cover: We excluded densely urban areas, water bodies, wetlands, and protected natural reserves, favoring forests, open land, and agricultural areas, as they are more likely to accommodate recharge facilities with minimal siting conflicts (Sallwey et al., 2019; Vandala and Mahed, 2025). Forest scored 5, shrubs/herbaceous 4, cropland 2, consistent with land-use suitability classifications in recent MAR studies (Sitek et al., 2026; Meng et al., 2024).

Distance to source water: We considered the distance to the river and its tributaries to minimize pipelines, canals, and infrastructure costs, while excluding flood-prone areas (Gibson et al., 2018; Russo et al., 2015). We estimated groundwater travel time from recharge areas to source water using particle tracking simulations within the hydrogeological model to prevent rapid return flow that could reduce recovery efficiency or affect the river ecosystem (Pollock, 1994; Harbaugh, 2005; Wang et al., 2020). Sites within 500-800 m scored 5, while those >1500 m scored 2, following distance-decay principles established in MAR site selection literature (Dillon et al., 2019; Seidl et al., 2024).

Distance to extraction wells: We assessed the proximity of groundwater pumping wells, excluding highly sensitive Protection Zones 1 and 2 to avoid well interference while ensuring effective recovery (Gibson et al., 2018). We relied on official LfU data on groundwater travel times to pumping wells, which are based on hydrogeological assessments and the delineation of water protection areas. For the remaining areas, we scored suitability according to travel time: <10 years = 5 (highly suitable), 10-30 years = 4, >30 years within the catchment = 3, and outside the catchment = 1 (unsuitable), consistent with well protection criteria in ASR applications (Russo et al., 2015; Stefan and Ansems, 2018).

Weighting and Sensitivity Analysis: To address potential subjectivity in criterion weighting (Hani et al., 2023; Abdo et al., 2024), we performed a sensitivity analysis comparing three scenarios: (S1) equal weighting (all criteria = 0.167), (S2) hydro-heavy weighting (storativity and K = 0.25 each), and (S3) distance-heavy weighting (distance to source and wells = 0.275 each) (Table 2). This follows best practices for validating MCDA robustness in groundwater studies (Hani et al., 2023; Vandala and Mahed, 2025).

Table 2: MCDA weighting scenarios for sensitivity analysis

Scenarios	Storativity	Hydraulic conductivity	Depth	Land Use	Distance to source	Distance to wells
S1 (equal weight)	0.167	0.167	0.167	0.167	0.167	0.167
S2 (hydro-heavy)	0.25	0.25	0.20	0.10	0.10	0.10
S3 (distance-heavy)	0.10	0.10	0.10	0.15	0.275	0.275

Section 4.3

We developed an ASTR suitability map using GIS-MCDA (Fig. 6). The analysis integrated six criteria (Fig. 4a), classified into five classes: very low (0-20), low (21-40), moderate (41-60), suitable (61-80), and highly suitable (81-100).

Sensitivity Analysis Results: We tested three weighting scenarios (Table 2). Results demonstrate high robustness (Fig. 6; Table 3). The "Very Low" suitability class remained stable at ~29% across all scenarios (variation <0.3%), indicating physical constraints consistently define unsuitable areas. The "Suitable" class showed moderate variation (36.7-43.3%; $\Delta = 6.6\%$), while "Highly Suitable" areas varied between 17.2% (S3) and 27.7% (S2) ($\Delta = 10.5\%$).

Despite percentage variations, the spatial distribution of high-suitability zones remained geographically consistent across all scenarios (Fig. 6a-c), confirming that site suitability is primarily controlled by underlying physical characteristics rather than subjective weight assignment (Hani et al., 2023; Vandala and Mahed, 2025).

Baseline Scenario (S1) Results: Under the equal-weight baseline scenario (S1), 19.5% of the area (670 km²) was highly suitable, and 43.3% (1,484 km²) suitable. Combined, 62.8% of the catchment (2,154 km²) is viable for ASTR. Highly suitable zones concentrate where Aquifer 2 has higher storativity ($>5 \times 10^{-3}$) and K (>30 m/d), optimal distances to rivers (500–800 m) and wells (<10 years travel time), and minimal land-use conflicts.

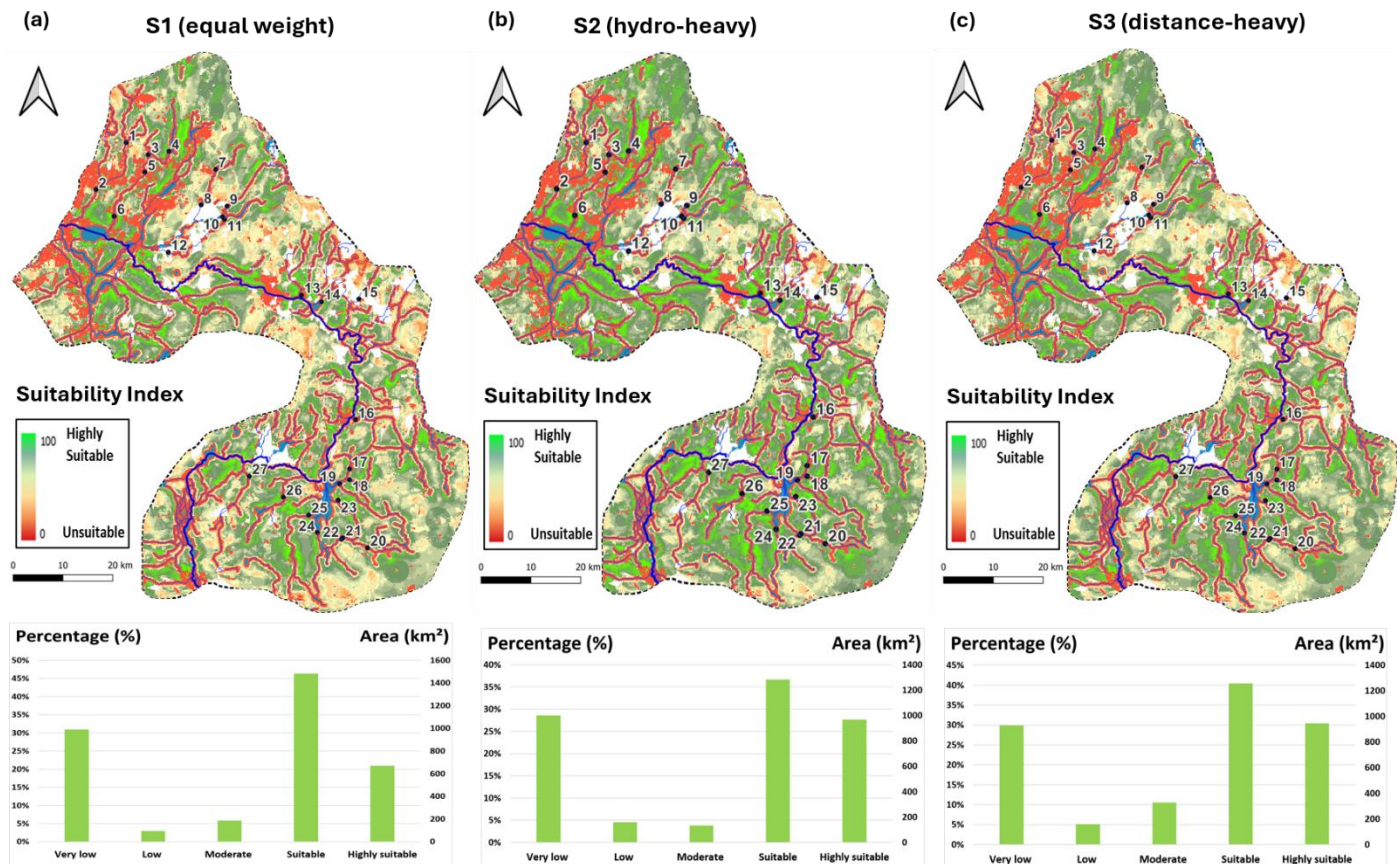


Figure 6. Sensitivity analysis of ASTR site suitability under three MCDA weighting scenarios: (a) S1 (equal weight), (b) S2 (hydro-heavy), and (c) S3 (distance-heavy). Bar charts show the percentage (%) and area (km²) per suitability class. The consistent spatial distribution of high-suitability zones across scenarios demonstrates the robustness of the MCDA results to criterion weighting. See Tables 2–3 for detailed weights and distributions.

Table 3: Suitability class distribution across weighting scenarios

Suitability class	S1 (%)	S2 (%)	S3 (%)	S1- S2	S1- S3
Very low	28.92	28.63	28.84	0.29	0.08
Low	2.82	3.21	2.92	-0.39	-0.1
Moderate	5.4	3.78	10.56	1.62	-5.16
Suitable	43.31	36.72	40.46	6.59	2.85
Highly suitable	19.56	27.66	17.22	-8.1	2.34

Discussion

..., suggesting the framework itself is highly transferable, moving the study beyond a single case towards a generalizable methodology for regional water security planning. Furthermore, a sensitivity analysis of the MCDA weighting schemes confirmed that these identified priority zones are driven by robust physical constraints (e.g., aquifer properties and depth) rather than subjective weight assignments, reinforcing the reliability and transferability of the framework for regional planning.

References

- Abdo, H. G., Vishwakarma, D. K., Alsafadi, K., Bindajam, A. A., Mallick, J., Mallick, S. K., Arun Kumar, K. C., Albanai, J. A., Kuriqi, A., & Hysa, A. 2024. GIS-based multi-criteria decision making for delineation of potential groundwater recharge zones for sustainable resource management in the Eastern Mediterranean: A case study. *Applied Water Science*, 14, Article 160. <https://doi.org/10.1007/s13201-024-02217-z>
- Dillon, P., Stuyfzand, P., Grischek, T., Lluria, M., Pyne, R. D. G., Jain, R. C., Bear, J., Schwarz, J., Wang, W., Fernandez, E., Stefan, C., Pettenati, M., van der Gun, J., Sprenger, C., Massmann, G., Scanlon, B. R., Xanke, J., Jokela, P., Zheng, Y., ... Sapiano, M. 2019. Sixty years of global progress in managed aquifer recharge. *Hydrogeology Journal*, 27(1), 1–30. <https://doi.org/10.1007/s10040-018-1841-z>
- Gibson, M.T., Campana, M.E., Nazy, D., 2018. Estimating Aquifer Storage and Recovery (ASR) Regional and Local Suitability: A Case Study in Washington State, USA. *Hydrology* 5(1), 7. <https://doi.org/10.3390/hydrology5010007>
- Hani, H.M., Nour El Din, M.M., Khalifa, A., Elalfy, E., 2023. Sensitivity Analysis for Multi-Criteria Decision Analysis Framework for Site Selection of Aquifer Recharge with Reclaimed Water. *Sustainability* 15(6), 5399. <https://doi.org/10.3390/su15065399>
- Lange, S., 2019. Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD (v1.0). *Geoscientific Model Development* 12(7), 3055–3070. <https://doi.org/10.5194/gmd-12-3055-2019>
- Meng, F., Khan, M. I., Naqvi, S. A. A., Sarwar, A., Islam, F., Ali, M., Tariq, A., Ullah, S., Soufan, W., & Faraj, T. K. 2024. Identification and mapping of groundwater recharge zones using multi influencing factor and analytical hierarchy process. *Scientific Reports*, 14, Article 19240. <https://doi.org/10.1038/s41598-024-70324-7>
- Russo, T.A., Fisher, A.T., Lockwood, B.S., 2015. Assessment of Managed Aquifer Recharge Site Suitability Using a GIS and Modeling. *Ground Water* 53(2), 269–287. <https://doi.org/10.1111/gwat.12213>
- Sallwey, J., Bonilla Valverde, J.P., Vázquez López, F., Junghanns, R., Stefan, C., 2019. Suitability maps for managed aquifer recharge: a review of multi-criteria decision analysis studies. *Environmental Reviews* 27(2), 138–150. <https://doi.org/10.1139/er-2018-0069>
- Seidl, C., Sprenger, C., Wang, W., 2024. Understanding the global success criteria for managed aquifer recharge. *Journal of Hydrology* 628, 130411. <https://doi.org/10.1016/j.jhydrol.2023.130469>
- Sitek, S., Janik, K., Piechota, A., Rubin, H., & Witkowski, A. J. 2026. Application of GIS-MCDA methodology for managed aquifer recharge suitability mapping in Poland. *Water*, 18(2), Article 219. <https://doi.org/10.3390/w18020219>
- Stefan, C., Ansems, N., 2018. Web-based global inventory of managed aquifer recharge applications. *Sustainable Water Resources Management* 4, 153–162. <https://doi.org/10.1007/s40899-017-0212-6>
- Vandala, B., Mahed, G., 2025. Managed aquifer recharge (MAR) site suitability in the Nelson Mandela Bay: the application of multi-criteria decision analysis techniques. *Sustainable Water Resources Management* 11, 78. <https://doi.org/10.1007/s40899-025-01241-4>

5) Moreover, regarding MCDA procedure, you apply equal weighting to all six MCDA criteria. It is well-known that weighting choices strongly influence MCDA outcomes in MAR suitability studies. The manuscript acknowledges this only in passing ("based on literature and expert judgment"). A sensitivity analysis testing alternative weighting schemes is strongly recommended.

Authors' response:

We fully agree that weighting significantly influences MCDA outcomes. As detailed in our response to Comment #4, We have performed a comprehensive sensitivity analysis testing three alternative weighting

schemes: (S1) equal weighting, (S2) hydro-heavy weighting (prioritizing storativity and hydraulic conductivity), and (S3) distance-heavy weighting (prioritizing proximity to source water and wells). The results (new Figure 6, Tables 2–3) confirm that while class proportions shift slightly, the spatial location of high-suitability zones remains consistent across all scenarios. This demonstrates the robustness of our findings and confirms that identified priority zones are controlled by actual physical constraints rather than subjective weight assignments. This sensitivity analysis has been explicitly added to Section 4.3.

Changes made in the manuscript:

Section 4.3

Text is in our response to Comment #4

6) Section 3.6 on Cost Benchmarking from Literature, the cost-analysis is just mentioned as evaluated "from a review of recent european MAR literature". An economic cost analysis would be very interesting and enrich your work, but it hasn't been investigated further since you state it's beyond the scope of your study. At least, you should properly report a summary table of specific informations (if available) for your case study from metioned literature in this paragraph.

It's important to underline that the economic analysis in your study uses a single literature-derived cost range of €0.37–0.51/m³. This range is derived from Ross and Hasnain (2018) and Sprenger et al. (2017), which are based on heterogeneous international datasets. You acknowledge that site-specific analysis was beyond scope, but the manuscript presents potential savings of €23–111 million/year with a level of confidence that is not warranted given the uncertainty in both cost and volume estimates. At minimum, a clear uncertainty paragraph should be introduced, and potential savings claims in the manuscript and in the abstract should be scaled back and framed explicitly only as order-of-magnitude estimates.

Authors' response:

We thank the reviewer for this critical feedback. We agree that presenting specific savings figures with high precision was unwarranted given the heterogeneity of the source data. We have: (1) added Table 4 summarizing four European cost sources with explicit uncertainty notes; (2) scaled back all savings claims to "order-of-magnitude estimates" (e.g., "potential for substantial cost savings" or "tens of millions of euros") in the Abstract, Discussion, and Conclusion; and (3) removed all specific monetary figures (€23–111M) from the manuscript to avoid false precision.

Changes made in the manuscript:

Abstract

...Literature-based cost estimates (€0.30–0.60/m³) are substantially lower than regional drinking-water production costs (~€1.80/m³), indicating potential for substantial cost savings relative to conventional water supply, though precise economic benefits require site-specific validation.

Section 3.6

To provide initial economic context, we derived indicative ASTR cost benchmarks from recent European MAR literature. While site-specific analysis was beyond the scope of this study, reported total costs for well-based MAR typically range from €0.30–0.60/m³ (CAPEX + OPEX) (Ross and Hasnain, 2018; Stefan and Ansems, 2018; Ward et al., 2022; Dillon et al., 2019). For comparison, regional drinking water production costs are ~€1.80/m³ (BWB, 2019). Table 4 summarizes specific benchmarks extracted from reviewed studies, highlighting their inherent limitations.

Table 4: Indicative ASTR/MAR Cost Benchmarks from European Literature

Study	Country	Technology	Total Cost (€/m ³)	Key Uncertainty & Limitations
-------	---------	------------	--------------------------------	-------------------------------

Ross & Hasnain (2018)	Germany	ASTR	0.37–0.51	Levelized cost; $\pm 30\%$ site-specific variation
Stefan & Ansems (2018)	Germany	ASTR	0.35–0.55	20–30 yr CAPEX amortization; rehabilitation excluded
Dillon et al. (2019)	Europe	ASTR	0.30–0.60	12 heterogeneous case studies
BWB (2019)	Germany	Water Production	~1.80	Conventional treatment benchmark

Note: All values are indicative order-of-magnitude estimates. Actual costs depend on local hydrogeology, energy tariffs, and regulatory requirements.

Discussion

Based on the indicative cost benchmarks in Table 4 and the technically viable recharge volumes identified, ASTR implementation at the candidate sites could yield annual cost savings on the order of tens of millions of euros relative to conventional drinking water production ($\sim \text{€}1.80/\text{m}^3$). However, given the $\pm 30\%$ uncertainty in literature-derived costs and variability in actual recharge performance, this figure should be treated as an order-of-magnitude estimate for strategic planning purposes only. Site-specific engineering and economic feasibility studies are necessary to refine these projections.

References

Ward, J.D., Simmons, C.T., Barnett, G.B., 2022. Economics of managed aquifer recharge: A European perspective. *Journal of Hydrology* 612, 128156. <https://doi.org/10.1016/j.jhydrol.2022.128156>

7) You note that the study covers only 33% of the total surface-water network. While this limitation is acknowledged in the Discussion, the scale of this gap is substantial and is used to argue that actual regional MAR potential is likely substantially higher. This extrapolation is speculative without further analysis since the unmonitored 67% could include tributaries with lower flow regimes, higher ecological sensitivity, or no overlap with suitable ASTR zones. This claim should be removed or framed as future research direction to investigate and verify.

Authors' response:

We agree with the reviewer that extrapolating the regional MAR potential to the unmonitored 67% is speculative. We have removed this claim and reframed it as a future research direction in the Discussion section (Lines 479-481):

Changes made in the manuscript:

Discussion

Finally, we note that the 27 gauges analysed represent only about 33% of the surface-water network. While this monitoring coverage provides a robust basis for the catchment-scale assessment, the remaining 67% of the network was not included in this analysis. Future research should investigate the hydrological characteristics, ecological constraints, and ASTR suitability of these unmonitored tributaries to verify the full extent of the regional MAR potential.

8) Approach 2 relies on hydroecological thresholds that are described generically. The manuscript does not clarify whether these thresholds were derived from local ecological assessments or transferred from literature.

Authors' response:

We thank the reviewer for highlighting this important clarification. We revised Section 3.5 to explicitly distinguish the sources of the thresholds applied in Approach 2. We clarified that we obtained the base hydroecological limit directly from the local guidelines of the State Office for the Environment Brandenburg (LfU), while we derived the additional hydrological thresholds (e.g., median flow and flood peak percentiles) from established scientific literature (e.g., Yarnell et al., 2020; CEFWG, 2021).

Changes made in the manuscript:

Section 3.5

Approach 2: A more restrictive method that limits abstraction to extreme flood peaks and prioritizes the preservation of environmental flows. We combined several hydrological and hydroecological thresholds to refine the conditions for abstraction. We obtained the base hydroecological limit directly from the local ecological flow guidelines provided by LfU, ensuring its applicability to the regional context. We derived the additional hydrological thresholds (e.g., median daily flow, 2-year and 5-year flood levels) from established scientific literature (e.g., Yarnell et al., 2020; CEFWG, 2021) to preserve channel maintenance flows and associated habitat functions. Under this approach, we do not divert water for ASTR when flows are at or below the hydroecological limit, at or below the median daily flow, or within the range defined by the 2-year and 5-year flood levels. Abstraction is only permitted when flows exceed the median daily flow but remain below the 2-year flood threshold, provided that post-abstraction discharge remains at or above the median daily flow, or when flows exceed the 5-year flood level up to the physical capacity of the intake, pretreatment, wells, and pond storage infrastructure (Schmidt et al., 2004; Vanham et al., 2022; Alley et al., 2022; Richter et al., 2011).

9) The abstract states results clearly but uses the phrase "offsetting 6–79% of local extractions" without specifying which sites this applies to. Readers may misinterpret this as a catchment-wide statement. Clarify that this refers to specific downstream sites.

Authors' response:

We agree with the reviewer that this phrase could be misinterpreted as a catchment-wide statement. We have revised the Abstract to explicitly clarify that this percentage applies only to specific high-potential downstream sites, preventing any overgeneralization.

Changes made in the manuscript:

Abstract

Flow-threshold analysis at 27 gauges showed that specific high-potential downstream sites could provide mean annual recharge volumes of 1.6–4.3 Mm³, offsetting 6–79% of local extractions at those specific locations.

10) The study focuses on ASTR (Aquifer Storage, Transfer and Recovery), but the framework description and some results sections frequently use ASR and ASTR interchangeably. Since ASTR involves an additional spatial separation between injection and recovery wells compared to ASR, this distinction carries operational significance and should be used consistently throughout.

Authors' response:

We thank the reviewer for this important clarification. We agree that ASTR differs from ASR due to the spatial separation between injection and recovery wells, which has operational significance. We have standardized the terminology throughout the manuscript by replacing ASR/ASTR and standalone ASR terms with ASTR where appropriate. We also added a clarification in the Introduction (Lines 60–65) to distinguish ASTR from ASR.

Changes made in the manuscript:

Introduction

The adoption of Aquifer Storage and Recovery (ASR) and Aquifer Storage, Transfer and Recovery (ASTR) in these settings is limited by data scarcity and operational uncertainty. While ASR typically involves co-located injection and recovery wells, this study focuses specifically on ASTR, which utilizes spatially separated injection and recovery sites. This spatial separation carries distinct operational, hydrogeological, and regulatory significance, which our framework explicitly addresses.

Throughout the manuscript

We replaced any standalone use of “ASR” or “ASR/ASTR” with “ASTR” to ensure consistent terminology, except when referring to general concepts or citing studies that specifically address ASR.

Response to the Reviewer #2:

General:

The authors present a conceptual, comprehensive and quantitative approach for the identification of aquifer sections most suitable for MAR – measures to combat groundwater resources shortages. The approach presented not only considers the actual suitability of aquifers based on the hydraulic characteristics but also the dynamics of boundary conditions, such as the temporal variability of predicted drought patterns and the availability of infiltration water from surface water bodies by considering hydroecological criteria. The authors build their research on a wealth of data on a number of subcatchments in the Lower Spree area, close to Berlin. This allows the consideration of the large variability of the complex structure of a terrain, formed by glacial action. This approach allows, in my opinion, a generalization of the approach because of the number of catchments considered in the study. Furthermore, the authors present cost estimates as decision making criterion for water resources managers.

In sum: the approach is new, quantitative and also considers the temporal dynamics of water resources as well as that of the availability of recharge water and I recommend publication of the manuscript with minor revision, providing the comments below are being taken care of.

Reviewer's comments:

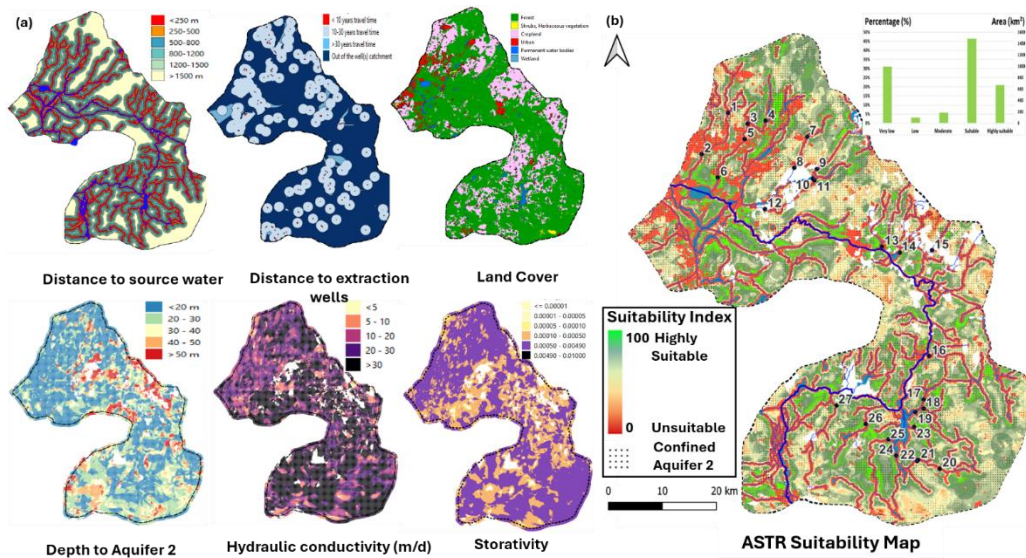
There is one main aspect that requires some more attention by the authors: they claim that their technique applies to confined aquifer systems such as Aquifer 2. In Chapter 3.4 they present the different criteria applied for the ASTR suitability assessment (L239 onwards). One criteria is storativity, a critical property within the context of MAR, especially if the aquifer system is confined, implying feasibility or efficiency of MAR measures. I do recommend that the areas with the aquifer 2 being confined to be outlined in Figure 6f and figure 10.

Authors' response:

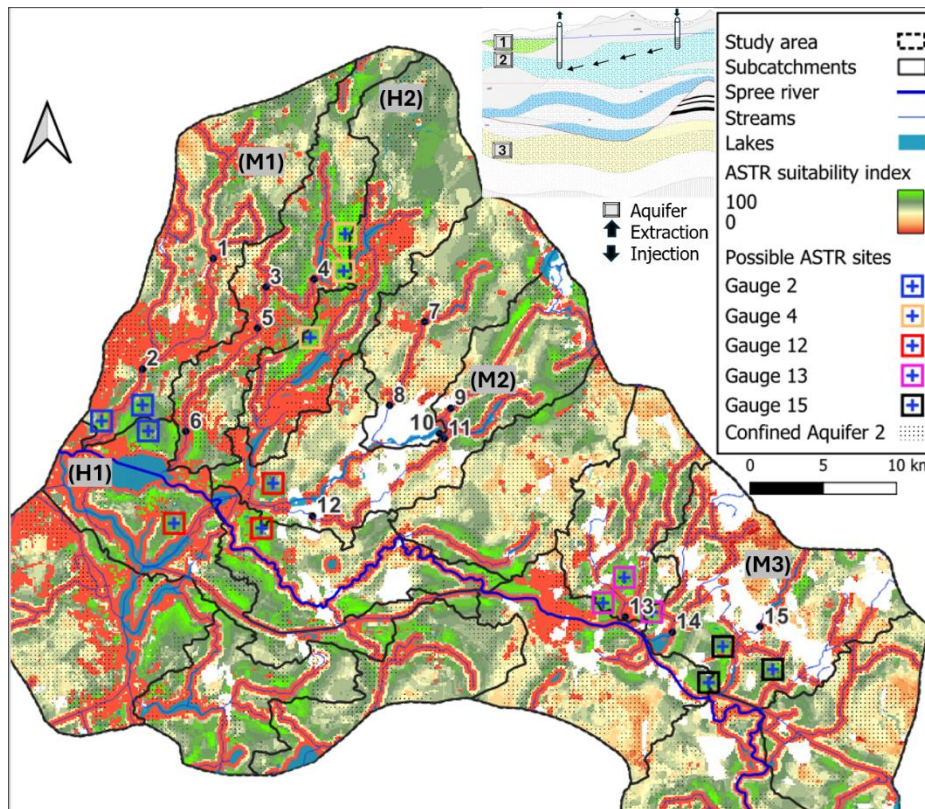
We thank the reviewer for this valuable suggestion. We have updated Figure 6 and Figure 10 to explicitly outline the areas where Aquifer 2 is confined (indicated by the dotted pattern in the legend). Furthermore, we have added a schematic diagram in the top-right corner of Figure 10 to illustrate the ASTR process, demonstrating direct injection into and recovery from the confined Aquifer 2. These visual additions clarify the feasibility of substantial storage volumes and the rationale behind our suitability criteria.

Changes made in the manuscript:

Updated Figure 6



Updated Figure 10



Should artificial recharge be accomplished via the unconfined Aquifer 1 and aquifer 2 be recharged via hydraulic windows, then, different criteria need to be selected on top. In order to clarify the concept for the reader, I recommend to the authors that they present a schematic diagram for different ASTR settings to demonstrate the feasibility of substantial volumes of storage, as well as the expansion of the suitability criteria. (schematic diagram in the Updated Figure 10)

Authors' response:

We thank the reviewer for raising this important conceptual point. As clarified in the updated schematic diagram (Figure 10), our framework specifically evaluates direct injection and recovery within the confined Aquifer 2. Because the study focuses on this direct pathway, the current suitability criteria (e.g., Aquifer 2 storativity and hydraulic conductivity) are fully appropriate. We have added a clarifying sentence in Section 3.4 to explicitly distinguish the evaluated direct ASTR setting from the indirect pathway via Aquifer 1, explaining why the criteria were not expanded to cover the latter.

Changes made in the manuscript:

Section 3.4

To clarify the conceptual framework for the reader, the schematic diagram in Figure 10 illustrates that this study specifically evaluates direct injection and recovery within the confined Aquifer 2. While indirect recharge via the unconfined Aquifer 1 and subsequent leakage through hydraulic windows represents an alternative ASTR setting, it requires distinct hydrogeological criteria and is therefore excluded from the current suitability assessment.

In order to make the results better “digestible” and in order to provide maximum use of the study for the reader, I recommend the authors to provide a table, listing all 24 catchments with individual columns specifying catchment characteristics (material properties, groundwater regime (confined/unconfined), ASTR suitability and in the final column the reason why it was selected suitable /unsuitable). This analysis should then be picked up, grouped into catchment classes and conclusions discussed after L 508.

Authors' response:

We thank the reviewer for this constructive suggestion. We have added a new table (Table S2 in the Supplementary Material) that lists all assessed catchments. The table includes individual columns specifying material properties (Mean Hydraulic Conductivity and Mean Storativity), the percentage distribution of the groundwater regime (confined, unconfined, and no aquifer), the mean MCDA suitability score, the final ASTR suitability classification, and a brief rationale for each classification. Furthermore, we have added a paragraph after Line 508 grouping the sub-catchments into three classes (Highly Suitable, Suitable, and Moderately/Low Suitable) based on their hydrogeological properties and MCDA scores, as detailed in the new Table S2. This classification provides a clear, prioritized roadmap for water resources managers to focus pilot projects on the most viable catchments.

Table S2: Catchment characteristics and ASTR suitability assessment.

Sub catchment	Aquifer Properties (Mean K m/d & Mean S)	Groundwater Regime % (Confined & Unconfined & No Aquifer)	Mean S1 Suitability (%)	ASTR Suitability	Reason
1	18.7 & 6.6e-4	67.6 & 30.0 & 2.4	56	Moderately Suitable	Moderate permeability; confined dominates
2	16.8 & 5.0e-4	57.3 & 31.2 & 11.5	69	Suitable	Good S1, but notable no-aquifer area
3	16.5 & 1.2e-3	56.3 & 43.2 & 0.5	69	Suitable	Highest storativity; aquifer almost fully present
4	16.7 & 6.1e-4	34.9 & 50.4 & 14.8	61	Moderately Suitable	Unconfined dominates but significant no-aquifer areas
5	19.3 & 1.3e-3	34.8 & 52.5 & 12.7	60	Moderately Suitable	High Kz & storativity, but reduced by no-aquifer areas
6	16.6 & 9.6e-4	71.9 & 28.1 & 0.0	69	Suitable	Aquifer fully present; good S1 despite confinement
7	15.3 & 7.5e-4	63.0 & 36.0 & 1.0	62	Suitable	Lower Kz but acceptable, aquifer mostly present
8	17.3 & 6.7e-4	77.5 & 20.9 & 1.7	74	Highly Suitable	Highest S1; excellent properties despite confinement
9	21.0 & 3.8e-4	74.0 & 20.2 & 5.8	71	Highly Suitable	Highest permeability; excellent for injection
10	19.9 & 6.5e-4	68.1 & 31.7 & 0.2	67	Suitable	High Kz, aquifer almost fully present
11	19.2 & 4.9e-4	42.3 & 48.0 & 9.8	56	Moderately Suitable	Balanced regime but low storativity & significant NA
12	18.8 & 5.2e-4	69.5 & 18.7 & 11.8	60	Moderately Suitable	Confined + large no-aquifer area limit potential
13	18.9 & 1.2e-3	47.2 & 52.7 & 0.1	64	Suitable	Excellent storativity, unconfined dominates
14	19.2 & 6.1e-4	50.3 & 43.5 & 6.3	62	Suitable	Balanced regime, good overall properties
15	17.6 & 4.1e-4	46.5 & 25.6 & 27.9	54	Low Suitability	Largest no-aquifer area, lowest S1. Not recommended
16	19.0 & 8.8e-4	36.2 & 60.5 & 3.3	62	Suitable	Unconfined dominates, good storativity
17	20.3 & 6.2e-4	49.8 & 44.6 & 5.6	63	Suitable	High Kz, balanced regime
18	19.6 & 8.5e-4	36.5 & 62.4 & 1.0	64	Suitable	Unconfined dominates, aquifer almost fully present
19	16.1 & 6.8e-4	52.7 & 46.3 & 1.1	65	Suitable	Moderate Kz but good S1; aquifer mostly present
20	20.8 & 6.9e-4	49.1 & 38.6 & 12.3	68	Suitable	High Kz, good S1, but notable no-aquifer area
21	18.1 & 8.1e-4	60.4 & 36.6 & 3.0	58	Moderately Suitable	Confined dominates, moderate S1

Changes made in the manuscript:

Discussion

To enhance the practical application of these results, we classified the assessed sub-catchments based on their hydrogeological characteristics and ASTR suitability (Table S2). Highly Suitable includes sub-catchments 8 and 9, which show excellent hydraulic properties and minimal "no aquifer" areas, making them priority candidates for ASTR implementation. Suitable comprises most sub-catchments (e.g., 2, 3, 6, 7, 10, 13, 14, 16–20) with good overall conditions and MCDA scores of 60–70%, while Moderate to Low Suitability includes areas limited by lower storativity or extensive "no aquifer" zones. This classification supports prioritization of pilot projects in highly and suitable catchments to maximize drought resilience benefits.

L227: RCP8.5 is highly unlikely; needs a statement, why it is considered here.

Authors' response:

We thank the reviewer for raising this important point. We acknowledge that the reference to RCP8.5 was incorrect. In this study, we do not use RCP-based results; instead, we consider socio-economic development pathways based on the SSP framework (specifically SSP1-2.6 and SSP2-4.5), consistent with IPCC AR6. We have entirely removed RCP8.5 from all calculations, figures, and text in the manuscript. To address the specific points raised:

- (i) Climate projections are based on regionalized datasets (UBA) derived from established CMIP ensembles.
- (ii) Socio-economic assumptions are represented through SSP pathways and integrated using sector-specific water demand projections from local stakeholders (BWB, Adelphi Research, ZALF).
- (iii) Yes, the projected climate variables were bias-corrected and calibrated against historical observations within the hydrological model to ensure reliable simulations.

Changes made in the manuscript:

Section 3.3

Projections were developed using SSP pathways (SSP1-2.6 and SSP2-4.5), consistent with IPCC AR6 and UBA regional climate projections (Kreienkamp et al., 2022). These pathways represent different socio-economic trajectories and were linked to sector-specific water demand projections from local stakeholders (BWB, Adelphi Research, ZALF). Climate variables were bias-corrected and calibrated against historical meteorological and hydrological records within the SWIM model to ensure reliable water-balance estimates.

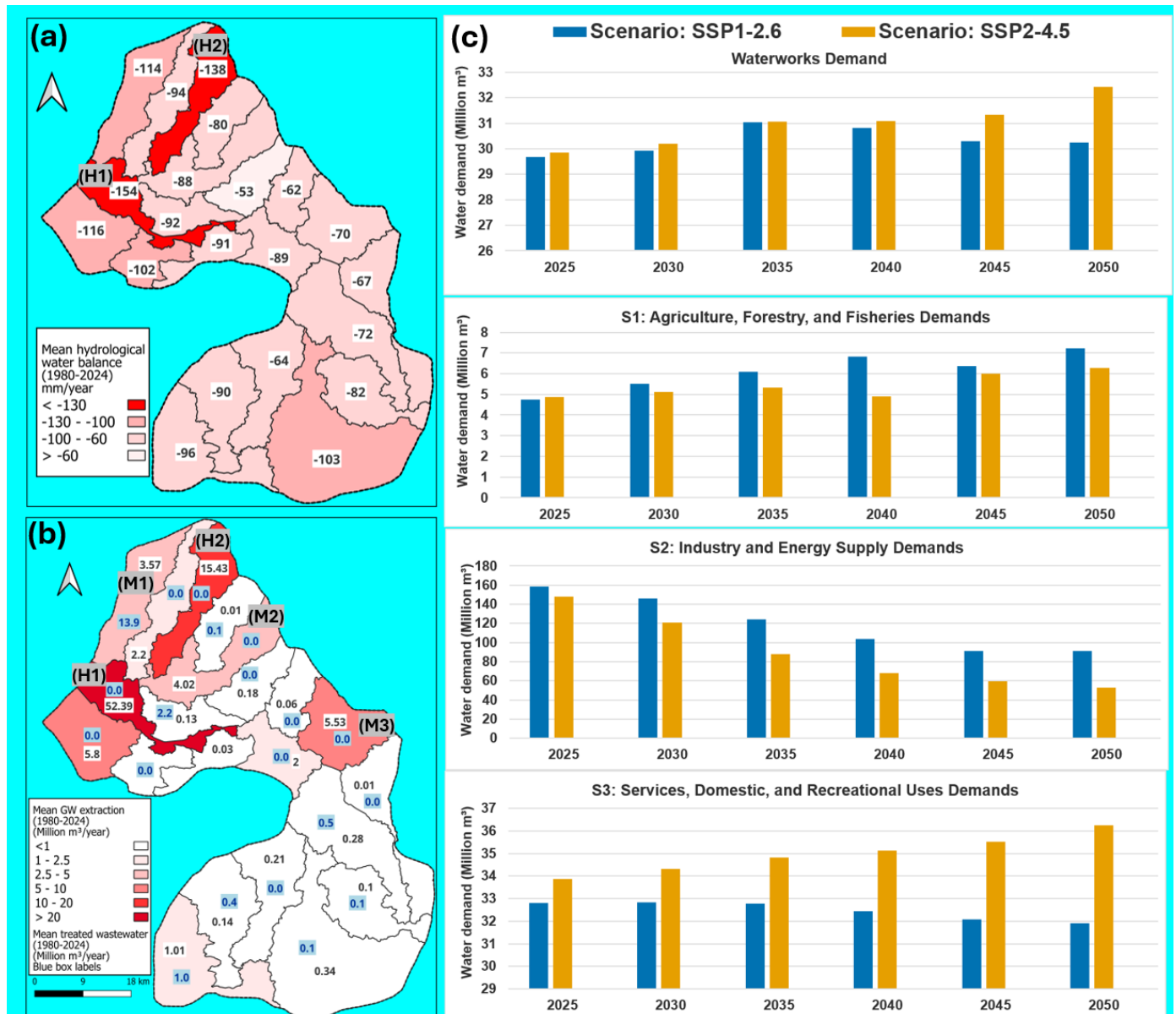
Section 4.2

We analysed projected water-demand scenarios for 2025-2050 and found a moderate but steady increase in total waterworks demand across the considered scenarios (Fig. 5c). Under SSP1-2.6, total waterworks demand fluctuates around 30-31 Mm³, while the most-likely SSP2-4.5 scenario shows a modest increase from ~30 to ~32 Mm³ by 2050. This spread (~1-2 Mm³) defines the stress-test range for storage and recovery capacity under realistic future conditions.

Sectoral patterns exhibit contrasting dynamics. S1 (Agriculture, Forestry, and Fisheries) shows a consistent upward trend, reflecting higher irrigation needs under warmer conditions. S2 (Industry and Energy) declines markedly due to improvements in water-use efficiency and structural transitions. The decline is most pronounced under SSP2-4.5 (from ~150 to ~55 Mm³). Despite the decrease, S2 remains the dominant contributor to total variance and peak loads. S3 (Services, Domestic, and Recreational Uses) diverges across scenarios: slightly decreasing under SSP1-2.6 (~33 to 31 Mm³) and gradually increasing under SSP2-4.5 (~34 to 37 Mm³), with pronounced seasonal peaks during summer months.

Changes made in the manuscript:

Updated figure 5c



L243: Hydraulic conductivity – Kvertical?

Authors' response:

We thank the reviewer for this precise question. We have clarified in Section 3.4 (Line 243) that the hydraulic conductivity values used in our regional screening refer to the horizontal hydraulic conductivity (HK), as it governs the lateral spread of the injected water and recovery well yields at the regional scale. We also added a clear statement acknowledging that while vertical hydraulic conductivity is critical for local injection efficiency, its detailed evaluation is beyond the scope of this regional screening and is reserved for future site-specific pilot designs.

Changes made in the manuscript:

Section 3.4

Horizontal hydraulic conductivity (HK): We prioritized areas with higher HK values, as they facilitate the lateral spread of the injected water plume and support sustainable recovery well yields at the regional scale. Suitability scores ranged from 5 (>30 m/d) to 1 (<5 m/d). We acknowledge that vertical hydraulic conductivity

is also critical for local injection efficiency; however, its detailed evaluation is beyond the scope of this regional screening and is reserved for future site-specific pilot designs.

- Is there any weighting considered between the different criteria

Authors' response:

We thank the reviewer for raising this critical point regarding MCDA weighting. We have now explicitly detailed a sensitivity analysis in Section 3.4 (and added Table S3 in the Supplementary Material) comparing three weighting scenarios: equal weighting (S1), hydrogeology-heavy (S2), and distance-heavy (S3). We retained S1 as our primary model because it provides a robust, unbiased baseline that ensures the framework's transferability. However, the analysis demonstrates that local decision-makers can easily adapt the framework using S2 or S3 based on specific regional priorities.

Changes made in the manuscript:

Section 3.4

We combined all criteria using a weighted overlay approach. To address potential subjectivity in criterion weighting, we conducted a sensitivity analysis comparing three scenarios: S1 (equal weighting), S2 (hydrogeology-focused, prioritizing storativity and conductivity), and S3 (distance-focused, prioritizing proximity to source water and wells). This analysis demonstrates the framework's robustness and transferability without relying on single-factor prioritization (Table S3 in the Supplementary Material).

Section 4.3

To validate that the suitability map is not driven by arbitrary weighting, we evaluated the three scenarios (Table S3). The results demonstrate high robustness: the "Very Low" suitability class remained remarkably stable at ~29% across all scenarios (variation <0.3%), indicating that physical constraints consistently define unsuitable areas. While the "Suitable" and "Highly Suitable" classes showed moderate percentage variations (e.g., Highly Suitable ranged from 17.2% in S3 to 27.7% in S2), the spatial distribution of these high-priority zones remained geographically consistent across all scenarios (Fig. 6). This confirms that site suitability is primarily controlled by underlying physical characteristics rather than subjective weight assignments.

Supplementary Material

Table S3: MCDA weighting scenarios for sensitivity analysis

Scenarios	Storativity	Hydraulic conductivity	Depth	Land Use	Distance to source	Distance to wells
S1 (equal weight)	0.167	0.167	0.167	0.167	0.167	0.167
S2 (hydro-heavy)	0.25	0.25	0.20	0.10	0.10	0.10
S3 (distance-heavy)	0.10	0.10	0.10	0.15	0.275	0.275

L239 or L268: A statement is required comparing the method presented here with others already published..

Authors' response:

We thank the reviewer for this important point. We have added a statement at the end of Section 3.4 comparing our method to existing published frameworks. We highlighted that while conventional MCDA studies rely primarily on static hydrogeological criteria, our framework uniquely integrates dynamic hydroclimatic (SPEI/SGI) and ecological flow thresholds, providing a more robust spatiotemporal assessment for drought resilience.

Changes made in the manuscript:

Section 3.4

Unlike conventional MAR suitability studies that primarily rely on static hydrogeological and spatial criteria (e.g., Khalil et al., 2022), the framework presented here integrates dynamic hydroclimatic and ecological constraints. By coupling the static GIS-MCDA with time-series drought indices (SPEI/SGI) and ecological flow thresholds, our approach transitions from a purely spatial assessment to a spatiotemporal evaluation. This ensures that identified suitable sites are not only hydrogeologically viable but also climatically resilient and ecologically sustainable, addressing a critical gap in existing MAR planning methodologies.

L347: What time intervals are used for the computation? Year/month/day?

Authors' response:

We thank the reviewer for pointing out the need for clarity regarding temporal resolutions. We have added a summary in Section 3.1 explicitly stating the time intervals used: daily data for SPEI and flow-threshold analysis, monthly records for SGI, and monthly-to-annual aggregation for the water balance.

Changes made in the manuscript:

Section 3.1

To ensure reproducibility, computations used the following temporal resolutions: daily data for SPEI (aggregated to 12-month periods) and flow-threshold analysis; monthly records for SGI; and monthly-to-annual aggregation for the water balance.

L231: Schematic Diagram of ASTR – Process needs to be included

Authors' response:

We have added a direct reference to the ASTR process schematic diagram (now included in Figure 10). This ensures the reader is immediately guided to the visual explanation of the direct injection and recovery process in the confined Aquifer 2.

Changes made in the manuscript:

Figure 10 caption

Figure 10. Streamflow-driven ASTR suitability across the catchment, with candidate recharge sites at Gauge 2 (blue), Gauge 4 (orange), Gauge 12 (red), Gauge 13 (pink), and Gauge 15 (dark blue); stressed sub-catchments (H1, H2, M1–M3) are labelled, highlighting where monitored surplus streamflow aligns with favorable hydrogeology for ASTR. The inset schematic diagram illustrates the conceptual ASTR process involving direct injection and recovery in the confined Aquifer 2.

Response to the Reviewer #3:

1- The manuscript presents an integrated framework combining drought indices, MCDA, and environmental flow analysis. While the integration itself is useful, it remains unclear which component constitutes the principal scientific innovation relative to existing MAR suitability assessment frameworks.

Authors' response:

We have added a statement in the Introduction (Lines 65-75) highlighting that the principal scientific innovation of this framework is its shift from static, unconfined-aquifer screening to a dynamic, spatiotemporal assessment specifically designed for confined multi-layer aquifers. This is achieved by integrating time-series drought indices for operational timing and ecological flow thresholds for source-water sustainability.

Changes made in the manuscript:

Introduction Section

Unlike conventional MAR suitability frameworks that primarily focus on static hydrogeological criteria in unconfined aquifers, the principal scientific innovation of this study lies in its dynamic, multi-component integration specifically designed for confined multi-layer aquifers. By coupling GIS-MCDA with time-series drought indices (SPEI/SGI) to trigger recharge operations, and ecological flow thresholds to safeguard source water, the framework transitions from a purely spatial screening tool to a spatiotemporal, climate-resilient operational blueprint for ASTR.

2-The suitability assessment represents one of the core outputs of the study, yet no uncertainty analysis appears to be conducted.

Authors' response:

We agree that addressing uncertainty is critical. We have added a paragraph in Section 4.3 detailing our uncertainty approach. We used the sensitivity analysis on criterion weighting (Table S3) to demonstrate the robustness of the suitability map across different scenarios. Additionally, we clarified that the strict application of absolute constraint masks (e.g., Well Protection Zones) significantly mitigates spatial uncertainty and prevents false-positive site selections.

Changes made in the manuscript:

Section 4.3

To address uncertainty in the MCDA suitability assessment, we conducted a sensitivity analysis on criterion weighting (Table S3), which demonstrated that the spatial distribution of highly suitable zones remains robust across different weighting scenarios (S1, S2, S3). Furthermore, while uncertainties exist in the 3D hydrogeological model inputs (e.g., K and storativity), the application of strict absolute constraint masks (e.g., Well Protection Zones 1-2, water bodies) significantly reduces the risk of false-positive site selection, ensuring a conservative and reliable final suitability map.

3-The manuscript should better explain how key hydrogeological controls on ASTR performance are represented. Factors such as aquifer heterogeneity, vertical hydraulic conductivity, leakage through confining units, storage properties, groundwater age, and geochemical compatibility can significantly influence recharge efficiency and recovery rates. The current framework appears to evaluate site suitability largely from a screening perspective and may therefore overestimate practical feasibility. The authors should discuss these limitations more explicitly.

Authors' response:

We thank the reviewer for this crucial point. We added a paragraph in the Discussion explicitly acknowledging that as a regional screening tool, the framework may overestimate practical feasibility. We listed the key local hydrogeological controls that must be evaluated in future site-specific pilot studies.

Changes made in the manuscript:

Discussion

As a first-tier regional screening tool, our framework may overestimate practical feasibility if applied directly to operational design. Critical local controls—such as vertical hydraulic conductivity, leakage through confining units, groundwater age, and geochemical compatibility—are not explicitly resolved at this scale and must be evaluated in subsequent site-specific pilot studies.

4-The manuscript would benefit from independent validation of the suitability results. For example, the identified high-potential zones could be compared with existing MAR facilities, groundwater management schemes, hydrogeological investigations, or known recharge-favorable areas. Without external validation, it remains difficult to evaluate whether the suitability map reflects actual field conditions or methodological artefacts.

Authors' response:

We thank the reviewer for this valuable comment. We have added a clarification in the Discussion regarding validation. Although direct validation is limited by the absence of operational ASTR schemes in the region,

the identified high-potential zones show agreement with known favorable hydrogeological areas, providing indirect validation. Future work will apply transient groundwater flow modeling (e.g., MODFLOW) to further evaluate storage efficiency and recovery performance.

Changes made in the manuscript:

Discussion

While this framework identifies potentially suitable recharge volumes at the regional scale, it does not simulate local hydraulic consequences. Future site-specific validation using transient groundwater flow modeling (e.g., MODFLOW) will be required to assess hydraulic responses, storage efficiency, recovery performance, and potential impacts on neighboring users, thereby bridging the gap between regional screening and operational implementation.

5-ASTR systems are generally most valuable during prolonged drought periods, which are precisely the periods when surface-water availability becomes most constrained. Consequently, mean annual recharge volumes may not accurately reflect operational feasibility. The authors should evaluate recharge availability during critical low-flow years and discuss how climate variability and drought persistence affect the reliability of source-water supplies.

Authors' response:

We thank the reviewer for this crucial point. We have added an analysis of the severe drought year 2018 (new Table S3), which demonstrates MAR volume reductions of 40–100% compared to long-term means (Results show that under Approach 1, MAR volumes dropped by 6–100% (mean 41%), and under Approach 2 by 20–100% (mean 77%) compared to long-term means (see new Table S3)). We clarified in Section 4.4, the Abstract, and the Conclusion that mean annual figures may overestimate drought-period reliability. Furthermore, we explicitly emphasized that the primary operational value of ASTR lies in capturing surplus wet-period flows for later recovery, rather than relying on constrained surface-water availability during droughts.

Changes made in the manuscript:

Section 4.4

While mean annual MAR volumes provide a useful baseline, analysis of the severe drought year 2018 reveals substantial reductions (40–100% depending on site and threshold, Table S3). This confirms that mean annual figures can overestimate operational reliability during prolonged dry periods. However, the core operational value of ASTR lies precisely in capturing surplus flood flows during wet periods for storage and subsequent recovery during droughts, rather than relying on constrained surface-water availability during the drought itself.

Abstract

Under mean annual conditions, site-specific recharge volumes range from 1.6–4.3 Mm³, though these volumes drop substantially (40–100%) during severe drought years (e.g., 2018), emphasizing the necessity of capturing wet-period surpluses.

Conclusion

While mean annual estimates indicate significant potential, drought-year constraints highlight the importance of capturing wet-period surplus—the fundamental rationale for ASTR implementation.

Table S3

Table S3. MAR-eligible volumes (Mm³) during drought year 2018 vs. mean annual values.

Site	Approach 1 (2018/Mean)	Approach 2 (2018/Mean)
1	0.3 / 0.6 (50%)	0.0 / 0.4 (0%)
2	1.5 / 1.6 (94%)	1.2 / 2.0 (60%)
4	0.6 / 1.3 (46%)	0.0 / 0.9 (0%)
12	3.2 / 4.4 (73%)	0.0 / 3.0 (0%)
13	0.0 / 1.1 (0%)	0.0 / 0.9 (0%)
15	0.9 / 1.0 (90%)	0.4 / 0.5 (80%)

6-The framework identifies potentially available recharge volumes but does not evaluate their hydraulic consequences within the receiving aquifer system. Groundwater flow modelling could provide valuable information regarding water-level recovery, storage efficiency, recovery ratios, residence times, and potential impacts on neighbouring users. The absence of such analysis limits the ability to translate suitability results into operational recommendations.

Authors' response:

We thank the reviewer for this valuable comment. We agree that groundwater flow modeling represents an important next step for translating regional suitability results into site-specific operational recommendations. We have clarified in the Discussion that the current framework provides a regional screening assessment, while transient groundwater flow modeling will be applied in future work to evaluate hydraulic responses, storage efficiency, recovery performance, and potential impacts on neighboring users.

Changes made in the manuscript:

Discussion

While this framework identifies potentially suitable recharge volumes at the regional scale, future site-specific validation using transient groundwater flow modeling will assess hydraulic responses, storage efficiency, recovery performance, and potential impacts on neighboring users, bridging the gap between regional screening and operational implementation.

7-The economic evaluation appears to rely on generalized literature-derived unit costs. However, ASTR economics are strongly site-specific and influenced by infrastructure requirements, treatment processes, monitoring obligations, energy costs, regulatory constraints, and long-term maintenance expenditures. The authors should provide a more detailed discussion of the uncertainties associated with the estimated savings and assess the sensitivity of the results to variations in cost assumptions.

Authors' response:

We thank the reviewer for highlighting this critical issue. We have already addressed this concern in our response to Reviewer 1 (Comment 6) by: (1) adding Table 4 summarizing European cost benchmarks with explicit uncertainty notes; (2) inserting a dedicated paragraph in Section 3.6 detailing $\pm 30\%$ site-specific cost variability; (3) scaling back all savings claims to "order-of-magnitude estimates" in the Abstract and Discussion; and (4) removing specific monetary figures (€23–111M) to avoid false precision. These revisions comprehensively address the uncertainties in literature-derived costs and emphasize that detailed techno-economic assessments are required before investment decisions.

8-Since the proposed framework is intended to enhance drought resilience, its application under future climate conditions should be explored. Incorporating climate projections or at least discussing the potential impacts of changing precipitation patterns, evapotranspiration rates, and streamflow regimes would considerably strengthen the long-term relevance of the study.

Authors' response:

We thank the reviewer for this valuable point. We have added a discussion paragraph addressing the application of our framework under future conditions. We clarified that while our demand projections incorporate future socio-economic pathways (SSP1 and SSP2), the supply-side assessment relies on historical flow regimes. We explicitly noted that projected increases in evapotranspiration and precipitation variability will likely make future surface-water availability more intermittent, reinforcing the necessity of our restrictive Approach 2 to ensure operational viability and ecological safety.

Changes made in the manuscript:

Discussion

While our demand projections explicitly incorporate future socio-economic pathways (SSP1 and SSP2), the supply-side assessment of surface-water availability in this study relies on historical flow regimes. Given projected increases in evapotranspiration and precipitation variability in Central Europe, future surface-water availability for ASTR will likely become more intermittent and constrained. This anticipated shift reinforces the necessity of the restrictive Approach 2, which prioritizes the capture of extreme flood peaks, ensuring that ASTR operations remain ecologically safe and operationally viable even as baseline flows decline under future climate conditions.

9-The future water-demand and climate assessment is based on RCP 2.6, RCP 4.5, and RCP 8.5 scenarios. However, the manuscript does not justify the use of the RCP framework instead of the more recent SSP-RCP scenario framework adopted in CMIP6 and the IPCC AR6. Given that socioeconomic development constitutes a central component of future water demand, the use of standalone RCPs may not adequately represent the interactions between climate forcing and socioeconomic pathways. The authors should clarify:

- (i) whether the climate projections originate from CMIP5 or CMIP6 datasets;
- (ii) why SSP-based scenarios (e.g., SSP1-2.6, SSP2-4.5, SSP5-8.5) were not considered;
- (iii) how socioeconomic assumptions were incorporated into the demand projections; and
- (iv) whether the projected climate variables were bias-corrected prior to their use in the hydrological assessment.

Authors' response:

We thank the reviewer for this crucial correction. We have updated the manuscript to explicitly use the combined SSP-RCP framework (SSP1-2.6 and SSP2-4.5), aligning with IPCC AR6, as this was the framework applied by our project partners for assessing socio-economic dimensions. To address the specific points raised: (i) The climate projections are based on regionalized datasets (UBA) aligned with established CMIP ensembles. (ii) The SSP-RCP framework was used to properly capture the interactions between climate forcing and socio-economic pathways. (iii) Socioeconomic assumptions were incorporated using detailed, sector-specific water demand projections from local stakeholders (BWB, Adelphi Research, ZALF) mapped to SSP narratives. (iv) Yes, the projected climate variables were bias-corrected and rigorously calibrated against observed historical records (DWD, LfU) within the hydrological model.

(Note: SSP5-8.5 was excluded per recent scientific consensus on its unrealistic trajectory).

Changes made in the manuscript:

Section 3.3

Projections were developed using the combined SSP-RCP scenario framework (specifically SSP1-2.6 and SSP2-4.5), aligning with the IPCC AR6 and recent German Environment Agency (UBA) regional climate projections (Kreienkamp et al., 2022). This framework explicitly integrates climate forcing with socio-economic pathways, providing a more robust representation of future water demand interactions. SSP1-2.6 represents a sustainability-oriented pathway, while SSP2-4.5 represents a middle-of-the-road pathway. (Note: The high-emission SSP5-8.5 scenario was excluded, as recent scientific consensus considers it an unrealistic trajectory). Socioeconomic assumptions for water demand were incorporated via detailed, sector-specific projections from local utilities and research institutes (e.g., BWB, Adelphi Research, ZALF). Furthermore, the projected climate variables were bias-corrected and calibrated against observed historical meteorological and hydrological records within the SWIM hydrological model to ensure the reliability of the water-balance estimates.

10- Since hydrological simulations are highly sensitive to biases in precipitation and temperature, the absence of information regarding bias correction raises concerns about the reliability of the projected water-balance estimates. A detailed description of the downscaling and bias-correction procedures is therefore necessary.

Authors' response:

We thank the reviewer for highlighting this important methodological detail. We clarify that no standalone RCP climate scenarios were used in this study; the previous RCP terminology was incorrect and has been replaced with the SSP framework throughout the manuscript. Climate inputs were derived from bias-adjusted and spatially downscaled datasets (W5E5v2.0 and ISIMIP Phase 3). Furthermore, the reliability of these bias-corrected inputs was verified by comparing simulated historical streamflows against observed discharge records, ensuring robust water-balance estimates.

Changes made in the manuscript:

Section 3.3

To ensure the reliability of the water-balance estimates, we derived climate inputs from bias-adjusted and spatially downscaled datasets. Historical model calibration and validation were performed using the W5E5v2.0 dataset, which is bias-adjusted against observational records (Lange, 2019). The reliability of these inputs was further evaluated by comparing simulated historical streamflows with observed discharge records, ensuring a robust representation of the regional water balance.

References

Lange, S., & Büchner, M. 2021. ISIMIP3b bias-adjusted atmospheric climate input data. ISIMIP Repository. <https://doi.org/10.48364/ISIMIP.842396.1>