FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 10 No. 17, September, 2026, pp 117 – 130
DOI: https://doi.org/10.33003/fjs-2026-1017-5815

A Review of Remote Sensing and GIS for Integrated Hydrogeological Characterization and Groundwater Dynamics

***Salwan S. Al-Hasnawi, Nada Farook Aboud, Shahad Adil Al-Qaraghuli, Hala Mohammed Al-Sekar**

Remote Sensing and Geophysics College, AlKarkh University of Science, Baghdad, Iraq

*Corresponding authors' email: salwan.alaiwi@kus.edu.iq

ABSTRACT

Remote sensing, GIS, field observations, geophysics, and process models are increasingly integrated in groundwater studies, but the hydrogeological significance of multi-sensor integration remains inconsistently defined. This systematic and critical review evaluates four domains: aquifer mapping and hydrogeological characterization (C1), groundwater recharge (C2), groundwater storage change (C3), and groundwater-related land subsidence (C4). A predefined framework separates integration level (I1-I4), sensor diversity (S1-S3+), and hydrogeological coupling (H0-H2), and assesses whether added data streams improve construct validity, uncertainty characterization, independent verification, scale representation, or process understanding. Four Scopus search modules yielded 12,125 records and 10,689 unique identifiers; consolidation produced a master dataset of 10,640 records. Title/abstract screening excluded 3,816 records, leaving 6,824 for eligibility assessment. Final reconciliation retained 4,014 report/study units and excluded 2,810. Detailed Q1-Q12 appraisal and critical inferential synthesis were restricted to 359 primary-source-verified studies; the remaining 3,655 retained records were not treated as fully verified or critically evaluated. The strongest designs moved beyond subjective overlay by integrating independent wells, tracers, hydrogeophysics, geodetic verification, cross-method estimation, uncertainty-aware fusion, and mechanistic coupling between groundwater and hydraulic pressure. Multi-sensor hydrogeology should therefore be judged by inferential gain rather than sensor count. Additional data are scientifically valuable when they provide non-redundant constraints, improve an explicit comparator, extend a key observation dimension, reduce uncertainty, strengthen causal attribution, or bridge a defined gap between process and inference.

Received: 25 Aug. 2026

Accepted: 08 Sept. 2026

Published: 21 Sept. 2026

Keywords: Remote Sensing-GIS Integration, Multi-Sensor Integration, Aquifer Mapping, Groundwater Recharge, Groundwater Storage Change, Land Subsidence

INTRODUCTION

Groundwater is an important underground resource and a spatially distributed hydrological condition, which makes comprehensive monitoring difficult because well measurements are direct but geographically dispersed. Groundwater reservoir testing and geophysical surveys can add important process-related information, although often local, whereas physical models rely on boundary conditions, visualization, calibration data, and scale. To overcome these limitations, remote sensing and geographic information systems (GIS) techniques are employed to help expand spatial and temporal coverage and integrate disparate evidence into a common analytical approach. However, most remote sensing techniques do not perform a direct groundwater detection, rather its condition is inferred from a different surface geological and hydrological indicator like the water storage changes, deformation of land surface, water balance components and topography. Therefore, the scientific value for this data do not relay on the satellite signal availability only, but on validity range of the hydrogeological inference chain which connect the detect signal by the underground property to be explained.

This problem clearly appears in two primary field, first is aquifer mapping and hydrogeological characterization where the data of geological, geomorphological, drainage network, soil, topography, land cover, climate, are integrated within GIS to produce a groundwater GIS suitability map. Remote sensing integration and GIS were used with Analytical Hierarchy Process (AHP) to detect groundwater potential area and it verification by wells data (Modibbo et al., 2025). Despite that the traditional overlay methods were developed to statistical methods and machine learning techniques, the results reliability stays relating with the variable target suitability, training data independence and verification, output compatibility range with field measurements and well data (Gumma and Pavelic, 2013; Naghibi et al., 2016). The second field represented by the groundwater recharge which it is difficult to direct measures on the regional level, therefore the satellite data are uses with water balance and hydrological samples to estimate that. Hoevere, it must be distinguished between recharge potential maps and quantitative recharge rates with consider the spatial and temporal compatibility, the quantity recharge rates, and uncertainly propagation (Gemitiz et al., 2017; Magnoni et al., 2020).

From the regional and continental scale, the gravity recovery and climate experiment (GRASE) system and GRASE Follow-On can be tracking the terrestrial water storage changes which requires splitting other storage components contribution and consider the spatial scale, filtration, leakage, and uncertainty (Rodell, et al., 2009; Castle, et al., 2014). On the other hand, the Interferometric Synthetic Aperture Radar (InSAR) provides a dense spatial measurement about earth surface deformation relates to the ground subsidence. But the deformation detecting alone not prove that is caused by groundwater depletion, because the causal attribution requires linking the deformation with level change, or hydraulic

pressures, or pumping, or aquifer system characteristics, or suitable geomechanically samples (Chaussard et at., 2013).
In the four areas, the integration of multiple sensors can refer to very different processes: stacking objective layers, integrating independent remote sensing products, multi-task time continuity, linking satellite monitoring with wells or geophysics, and incorporating them into groundwater models or compression models. These patterns did not represent a constant plane of scientific power, so merge a large number of connecting layers my lead to add a limited independence information, while it could for a just two source merging to provide large indicator value if the additional source can be restricted directly unconfirmed hydrological unit. Consequently, the integration evaluation depends in this review on the construct validity, verification independence, scale, uncertainty, and attribution, but not just on the number of sensors.
Accordingly, the present review aims to evaluate how to integration between the remote sensing and GIS with the hydrological evidences within the for field of C1-C4, and determine when the multi-sensor integration led to real improvement in forecasting and verification; and understanding operations and supporting; and when the data sources number increasing does lead only to greater complexity without reducing hydrogeological uncertainty. This research offers three contributions. First, four groundwater-related outcomes are evaluated within a validity framework specific to each outcome, rather than treating remote sensing as a single methodological category. Second, multi-sensor correlation is separated from hydrogeological correlation. Third, added value should be redefined as an inferential gain—evidence that the additional data stream improves an explicit comparator, provides an independent constraint, reduces uncertainty, expands a key dimension of observation, strengthens attribution, or bridges a gap between process and inference.
The final study-selection pathway is summarized in Figure 1.

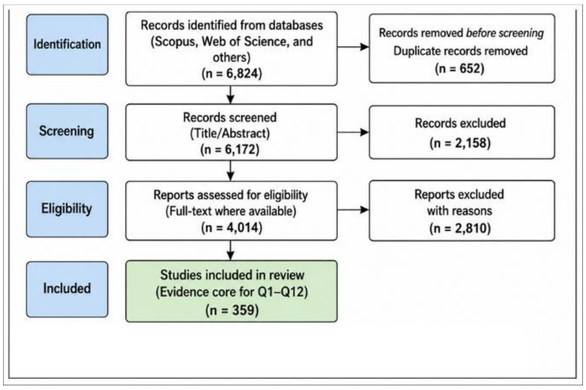

Figure 1: PRISMA 2020 Flow Diagram Showing the Study-Selection Process from 6,824 Records Entering Eligibility Assessment to 359 Studies Included in the Final Evidence Synthesis

MATERIALS AND METHODS

Review Design and Reporting Framework

We designed the study as a systematic and critical review. Reporting follows PRISMA 2020 and PRISMA-S, and protocol development was informed by PRISMA-P (Page et al., 2021; Rethlefsen et al., 2021; Shamseer et al., 2015). Search-strategy quality assurance followed PRESS principles (McGowan et al., 2016). An independent human PRESS peer review was not completed; this deviation is reported explicitly and is not presented as completed PRESS review.
The scientific unit of interest was an empirical or method-validation study in which remote sensing, Earth observation, or GIS contributed materially to the analysis and at least one of four groundwater domains was addressed in the four domains C1-C4 defined in the table 1. The review has global scope and is not restricted by aquifer type or geographical region.

Hydrogeological Domains and Integration Framework

The hydrogeological domain has been splits from the integration level to ensure applying an appropriate suitability criterion for each domain while maintaining the comparison capability across domains.
The four review domains and their primary hydrogeological constructs are summarized in Table 1.

Table 1: Review Domains and Primary Hydrogeological Constructs

Code	Domain	Primary construct	Critical validity issue
C1	Aquifer mapping/ hydrogeological characterization	Groundwater occurrence, productivity, aquifer boundaries/properties, hydrostratigraphic or potential mapping	Whether spatial proxies and predictive maps are anchored to valid field/well reference data and independent validation.
C2	Groundwater recharge	Recharge rate, amount, flux, volume, spatial recharge zones, or recharge dynamics	Whether recharge is defined and quantitatively distinguished from recharge potential; temporal compatibility, mass balance, and validation.
C3	Groundwater storage change	Groundwater-storage anomaly, depletion, recovery, or trend	Whether groundwater is defensibly separated from total terrestrial water storage and uncertainty/scale effects are propagated.
C4	Groundwater-related land subsidence	Surface deformation or aquifer-system compaction attributable to groundwater processes	Whether deformation is referenced/corrected appropriately and groundwater causation is

Code	Domain	Primary construct	Critical validity issue
			supported rather than inferred from co-location alone.

The integration level has been classifying independently from the hydrogeological domain to I1-I4, also sensors diversity was classified to S1-S3 and the hydrogeological interconnectedness to H1-H2, according to definitions shown in Table 2.

Table 2: Integration, Sensor-diversity, and Coupling Coding

Code	Operational definition	Examples	Permitted interpretation
I1	One principal EO/sensor/product or one principal GIS-derived dataset; no multi-layer integration that changes inference.	Single Landsat analysis; single InSAR deformation product.	Comparator and field evolution; not evidence of multi-sensor benefit.
I2	Multiple thematic layers integrated in GIS/MCDA/statistical/ML analysis; may use one sensor family.	Weighted overlay; AHP; RF/BRT/CART with geospatial predictors.	Integrated-GIS conclusions; multi-sensor claims only if separately S2+.
I3	At least two distinct sensor/product families or physical measurement principles integrated in one inference chain.	Optical + SAR; GRACE + optical/thermal.	Primary evidence for explicit multi-sensor integration.
I4	Geospatial evidence explicitly coupled with ground or process evidence.	EO + wells; InSAR + heads; GRACE + modeled components; calibrated groundwater model.	Hydrogeological coupling; multi-sensor claims only when sensor-diversity criterion is also met.
S1/S2/S3+	One, two, or three-or-more distinct sensor/product families.	Orthogonal to I-level.	Describes sensor diversity, not evidence quality.
H0/H1/H2	Hydrogeological coupling absent / ground observations / calibrated process model or formal assimilation.	H1: wells; H2: calibrated groundwater or compaction model.	Describes coupling strength/type, not automatically study quality.

The temporal distribution of retrieved records across the review period is shown in Figure 2.

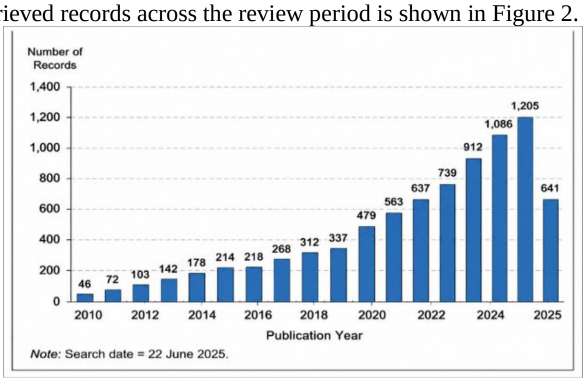

Figure 2: Annual Distribution of Records (2010-2025), Showing the Temporal Growth of the Retrieved Literature Across the Review Period

Eligibility Criteria

Eligibility was assessed through subject, methodology, results, type of publication, and the relationship between reports. Any restrictions did not apply for the language, or year, or document type in the preliminary research. Keeping the study requires represent groundwater is actual goal, and for remote sensing or GIS to play an important role in the analysis, and study addressed one of C1-C4 domains. Conference proceedings are eligible only if they are complete, archived research papers and are subject to documented review, and journal reports were linked at the study level to avoid data duplication. For report C4, the contextual reference to groundwater was insufficient, retention required a measurable or demonstrable, evidence of its contribution.

The eligibility and study-linkage decision rules are summarized in Table 3.

Table 3: Eligibility and Study-Linkage Checks

Item	Question	Decision rule
S1	Does the report study groundwater or an aquifer system?	No → exclude; unclear → retain for assessment.
S2	Is remote sensing/EO or GIS materially used in analysis or modeling?	No → exclude; map display alone is insufficient; unclear → retain.
S3	Does the report address at least one C1–C4 core outcome?	No → exclude; unclear → retain.
S4	Is it empirical or method-validation evidence with sufficient methods/results, or plausibly so?	No → exclude from primary evidence; reviews route to secondary set.

Item	Question	Decision rule
S5	Is the publication type potentially eligible and peer reviewed?	No → exclude; proceedings uncertainty → route-level verification.
S6	Is the groundwater contribution/data separable at report level?	No → exclude with primary reason; strict C4 requires measured/separable groundwater linkage.
P1	For proceedings, is the paper full archival, reviewed, and sufficiently reported?	No or unverifiable → FT-S5 exclusion.
L1	Is the record a duplicate or companion report?	Link at study level; retain the most complete report; count unique data once.

Information Sources and Search Strategy

We searched Scopus on 13 August 2026 using four outcome-specific modules. Each module combined a groundwater-system block, a remote-sensing/GIS/sensor block, and one domain block (C1–C4) within TITLE-ABS-KEY. The terms multi-sensor, multisensor, integration, and fusion were not required because relevant I1–I4 studies may not describe themselves using this vocabulary. No filters relating to date, language, or document type were applied during the initial retrieval.

The Scopus database was used as the primary bibliographic database, with citation sequencing and normative validation used as supplementary sensitivity checks.

The Verification of Search

The sensitivity of the search was assessed using eight pre-selected empirical reference studies, two for each C1-C4 domain. The initial verification resulted in the recovery of seven of them, the missing criterion, C3, was recovered after add full names of the GRACE and GRACE-FO missions, so that the final verification result becomes 8/8, and the test details are preserved in the Supplemental Review archive.

Records Management and Duplication Removal

The raw Scopus data was retained along with the module and query data. The process of removing duplicates thorough Scopus Electronic ID (EID) matching, DOI and uniform headings were then used to identify duplicates and accompanying reports. Conference and journal pairs were treated as potential companion reports, not as automatic duplicates.

Throughout the subsequent sorting processes, the data source was maintained at the identity level so that each selection decision could be traced back to the master dataset.

Data Extraction and Evidence Coding

From the primary-source verified synthesis set, bibliographic, spatial, and hydrostratigraphic data were extracted, and also C1-C4 domain; GIS/ML/modeling methods; I, s, and H levels; validation data, quantitative metrics, uncertainty, comparison, and explicit claims regarding added value. A study could contribute to more than one domain, but evidence was linked once at study level using domain-specific outcome fields.

Claims that multi-sensor integration improved performance was coded separately from the simple presence of multiple data sources, and considered such claims evidentially supported except when there is a baseline, single-sensor comparator, ablation, independent alternative capable of isolating the contribution of integration.

Critical Appraisal

Methodological quality was evaluated across 12 prespecified domains without calculating a summed quality score. Q1, Q6, Q7, Q8, Q12 were treated as critical. In Q12, a special rules for each domain were applied. Final study-level judgments were made only for the 359 primary-source verified studies. The 12-domain critical-appraisal framework is summarized in Table 4.

Table 4: Q1–Q12 Critical Appraisal Framework

Domain	Construct	Primary concern
Q1*	Hydrogeological construct validity	Is the target explicit, plausible, and distinct from a proxy or potential surface?
Q2	Setting and scale	Are aquifer setting and spatial/temporal support sufficient and matched to inference?
Q3	Data provenance and suitability	Are products, versions, dates, resolution, and suitability reported?
Q4	Pre-processing and alignment	Are corrections, resampling, reprojection, and temporal matching appropriate?
Q5	Integration transparency	Is the architecture reproducible, and is a baseline present when superiority is claimed?
Q6*	Reference data and validation independence	Are reference data representative and independent from model development/training?
Q7*	Spatial/temporal dependence and leakage	Does validation prevent neighbor, duplicate, temporal, or preprocessing leakage?
Q8*	Uncertainty and sensitivity	Are material uncertainties quantified/propagated and conclusions tested for robustness?
Q9	Confounding and attribution	Are alternative drivers evaluated and causal language proportional to design?
Q10	Transferability	Are claims bounded to tested settings or supported by external validation?
Q11	Reporting and reproducibility	Are products, parameters, software, thresholds, and workflows auditable?
Q12*	Outcome-specific validity	Do module-specific fatal-error checks pass?

For C1, high concern includes predictive/potential maps without field independent validation, or mapped proxies presented as occurrence/yield. For C2, the confusing between recharge potential and actual recharge rate, or using temporally incompatible balance components. For C3, terrestrial water storage is treated as groundwater storage without separating material components and uncertainty. For C4, deformation attributing causation to groundwater without pumping/head/hydrostratigraphic or appropriate modeling support.

Evidence Synthesis

Synthesis was organized by the domain (C1–C4) and then by integration level, sensor diversity, fusion stage, and hydrogeological coupling. Quantitative results are only valid when there is an explicit comparison or an independent reference to support the interpretation. Therefore, the main synthesis is structured and critical, focusing on construct validity, independent verification, uncertainty, transferability, and attribution.

RESULTS AND DISCUSSION

Results

Search Results and Verification

The four active Scopus p02 modules retrieved 3,590 C1 records, along with 3,523 C2 records, 2,239 C3 records, and 2,773 C4 records, for a total of 12,125 module-level records. Furthermore, the integration of precise EIDs between modules yielded 10,689 unique Scopus records, confirming significant, though non-cumulative, overlap between domains. Master dataset consolidation resulted in 10,640 records.

The number of direct retrievals at the unit level is shown in Table 5.

Table 5: Live search results in the Scopus database p02 (August 13, 2026).

Module	**Domain**	**Live records**	**Benchmark recall**
C1	Aquifer mapping / characterization	3,590	2/2
C2	Groundwater recharge	3,523	2/2
C3	Groundwater storage change	2,239	2/2
C4	Groundwater-related land subsidence	2,773	2/2
Total	—	12,125	8/8 overall

The validation rate of the criterion was 8/8 after modifying the GRACE/GRACE-FO vocabulary. The change had a negligible impact on the overall retrieval volume, adding six rows across modules with similar baselines before the modification, while also recovering previously undiscovered C3 criterion. The resulting strategy was adopted in the final search.

Table 6 summarizes the results of retrieving the criteria using the final search strategy.

Table 6: Retrieval of Experimental Parameters Through Final Search in Scopus

ID	Domain	Study	DOI	Final result of the search
B01	C1	Gumma & Pavelic (2013)	10.1007/s10661-012-2810-y	PASS
B02	C1	Naghibi et al. (2016)	10.1007/s10661-015-5049-6	PASS
B03	C2	Gemitzi et al. (2017)	10.1016/j.jhydrol.2017.01.005	PASS
B04	C2	Magnoni et al. (2020)	10.1007/s40899-020-00469-6	PASS
B05	C3	Rodell et al. (2009)	10.1038/nature08238	PASS
B06	C3	Castle et al. (2014)	10.1002/2014GL061055	PASS after amendment
B07	C4	Erban et al. (2014)	10.1088/1748-9326/9/8/084010	PASS
B08	C4	Chaussard et al. (2013)	10.1016/j.rse.2012.10.015	PASS

Study Selection and Final Reconciliation

The database-source composition of the 6,824-record eligibility inventory is shown in Figure 3.

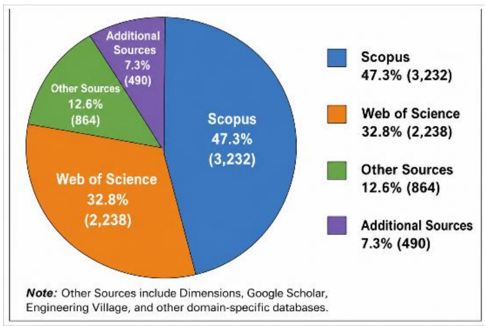

Figure 3: Records by Database Source (n = 6,824), Showing the Relative Contribution of Scopus, Web of Science, Other Sources, and Additional Sources

Of the 10,640 harmonized records, 3,816 were excluded during title/abstract screening and 6,824 entered eligibility assessment. Companion and outcome relationships were reconciled, and publication routes were verified. Five direct companion/outcome corrections were applied. Of 331 conference-paper candidates assessed against the publication-route criterion, 307 met the criterion and 24 were excluded. The final set comprised 4,014 retained report/study units and 2,810 exclusions, with no exact normalized DOI or normalized-title duplicate group remaining among retained records.

Two conventional PRISMA nodes could not be reconstructed numerically from the preserved records: reports not retrieved and reports assessed for eligibility through uniform full-text review. All 6,824 advanced records were designated for eligibility and retrieval assessment, but individual retrieval attempts were not logged completely. Therefore, these contracts are reported as unspecified rather than being assigned an unsupported zero value or considered as evidence of a uniform evaluation of the full text.

Retained Group Properties and the Verified Synthesis Core

There are 4,014 report/study units and 2,810 excluded units in the final archive. The majority of the archived records (3,698) are articles published in scientific journals, along with 307 conference papers that met the publication track criteria, five data sheets, three theses, and one doctoral dissertation. These bibliographic statistics describe the preserved inventory, but they are not used as a substitute for standardized full-text verification.

The critical inferential compilation and final evaluation process from Q1 to Q12 was limited to only 359 studies that were verified from primary sources (FV-001–FV-359). Domain memberships within this verified core are non-additive: C1 = 97, C2 = 52, C3 = 133, and C4 = 175. The core set consists of 333 articles published in scientific journals, 17 conference papers that have passed the first stage of publication (P1), five data papers, three letters, and one memorandum, covering the period from 2006 to 2026. The level of verification varied between the records, and this variation was taken into account during the evaluation rather than being combined into a hypothetical common whole-text level.

Table 7 summarizes the final selection and the limits of the verified composition.

Table 7: Final Selection and Verified Composition Limits

Dimension	Final value	Interpretation
Harmonized master dataset	10,640	Records that enter the sorting stage after coordination.
Title/abstract exclusions	3,816	It was excluded when checking the title/abstract.
Eligibility assessment set	6,824	The records submitted for eligibility assessment do not equal the 6824 retrieved full-text documents.
Retained report/study units	4,014	Final Report/Study Units Retained.
Excluded reports	2,810	This includes eligibility exceptions and publication path exceptions.
Verified synthesis / Q1–Q12 core	359	Documented studies from primary sources that include detailed assessments and inferential claims.
Additional retained report/study units	3,655	Not represented as uniformly full-text verified or Q1–Q12 appraised.
Verified-core C1 / C2 / C3 / C4	97 / 52 / 133 / 175	Non-additive domain memberships.
Verified-core publication types	333 articles; 17 conference; 5 data papers; 3 letters; 1 note	Conference papers passed route audit.
Conventional full-text PRISMA nodes	NR / not auditably determinable	No complete report-by-report retrieval-attempt ledger.

Aquifer Mapping and Hydrogeological Characterization (C1)

C1 evidence separates three claims that are often conflated. Model choice can improve predictive discrimination within the same general predictor environment. Rehman et al. (2024) reported frequency-ratio success/prediction rates of 89%/93% versus 82%/86% for weights of evidence. Arabameri et al. (2019) compared weights of evidence, binary logistic regression, random forest, and TOPSIS and found the highest prediction performance for binary logistic regression (prediction AUC 0.905; success AUC 0.918). These comparisons support a benefit from algorithm/model selection rather than superiority based on sensor count.

Independent hydrogeological information can also strengthen construct validity. Nainggolan et al. (2024) anchored AHP and fuzzy-extended AHP groundwater-potential indices to observed groundwater availability derived from aquifer thickness, groundwater depth, and porosity; correlation with that baseline increased from 0.56 for AHP to 0.67 for FE-AHP. Sangawi et al. (2023) validated a seven-factor GIS/RS/AHP potential map using eight two-dimensional electrical-resistivity tomography profiles, which confirmed correspondence between mapped classes and water-bearing versus non-water-bearing subsurface lithology. Gaber et al. (2020) extended this approach by combining optical/radar/terrain products with aeromagnetics, vertical electrical sounding, boreholes, hydrochemistry, and isotopes to characterize aquifer geometry and connectivity west of Qena.

The central C1 validity issue is whether validation matches the construct being claimed. Well presence does not automatically validate yield, storage, transmissivity, or aquifer geometry. Similarly, cross-validation of the model does not necessarily prove hydrogeological independence, and the increased area under the curve (AUC) resulting from a different algorithm does not provide sensor-level utility. Therefore, robust C1 designs define the reference target before model construction, maintain spatial independence between training and test data, and use geophysical or well information to limit subsurface interpretations when claims exceed surface suitability.

Recharge of Groundwater (C2)

There is a clear hierarchy of hydrogeological power that has been illustrated by the evidence from category C2. Feed potential maps can be useful screening tools, but their value depends on the verification objective. The study of Ki et al. (2023) found strong agreement between spatial refeeding categories and tritium evidence, providing more direct support for the validity of constructing a refeeding zone than the thematic overlay alone. Maqsoom et al. (2022), in contrast, he produced a detailed, multi-factor map of rechargeability without independent validation of the final product at the outcome level. This discrepancy explains why additional layers should not be equated with greater hydrogeological precision, since the potential for water recharge remains merely a spatial prioritization concept unless it is linked to quantitative flow through water balance, indicators, groundwater level fluctuations, calibration process modeling, or another independent pathway.

Quantitative water recharge studies can provide a more robust test because the estimated flow can be compared independently. Coelho et al. (2017) combined rainfall data from TRMM/TMPA, as well as remotely derived surface runoff and evapotranspiration data from MODIS/SEBAL, into a water balance using GIS, and then compared water recharge using groundwater level fluctuations. Satellite-measured rainfall data matched data from 15 gauges with a correlation coefficient (r) of 0.93 and a mean absolute error (MAE) of 12.7 mm; in a 2011 comparison of alluvial soils, the recharge ratio was 154.6 mm based on groundwater level fluctuations versus 124.6 mm based on remotely sensed water balance. Alghafli et al. (2023) Al-Ghafli et al. (2023) compared the groundwater recharge rate derived from GRACE data with that of an in-situ multi-component groundwater recharge method, reporting a coefficient of determination (R^2) of 0.91, a Nash-Sutcliffe efficiency coefficient (NSE) of 0.7, and a mean root-squared error (RMSE) ranging from approximately 1.5 to 7.8 million cubic meters. These quantitative claims regarding added value have been confirmed by these studies, because satellite/GIS estimates are tested against distinct hydrogeological reference paths.

Integrating processes and models can facilitate a complementary approach. For example, Milevsky et al. (2009) used multiple satellite products with a SWAT framework calibrated for surface runoff to identify groundwater recharge zones in arid basins. In a separate study by Dekongmen et al. (2022), a SWAT model was calibrated and validated using runoff data, and separate maps of potential groundwater recharge zones were created using remote sensing, geographic information systems (GIS), well studies, and drilling techniques. These examples emphasize the need to treat recharge flow and recharge potential, as well as aquifer productivity, as separate structures even when they appear within the same integrated workflow.

Groundwater Storage Change (C3)

C3 is dominated by the challenge of separating groundwater storage from total terrestrial water storage while preserving uncertainty and scale. Su et al. (2022) provide a clear example of incremental performance. Their groundwater weighted fusion model combines GRACE-based groundwater-storage estimates derived through multiple hydrological-model pathways and evaluates performance against in situ groundwater-level measurements. Relative to the original estimates, the primary report showed correlation-coefficient increases of 9–40%, Nash–Sutcliffe efficiency increases of 23–657%, and RMSE reductions of 9–28%. Because both a simpler/original result and independent in situ observations are available, the study supports a method-level added-value statement. The remaining issue is how product dependence and component uncertainty propagate through the fusion.

Li and Ma (2024) combined GRACE/GRACE-FO, GLDAS, groundwater-monitoring wells, high-resolution land-cover data, ERA5-Land, and statistical information to estimate and interpret groundwater-storage change on the Loess Plateau. They reported a mean groundwater-storage decline of -6.9 ± 3.84 mm/yr from 2003 to 2022, corresponding to an 88 km³ loss, and used monitoring wells as a reliability check. Such designs strengthen regional interpretation by linking gravimetric change to ground observations and plausible drivers, although they remain constrained by GRACE support scale and component-separation assumptions.

Downscaling requires particular caution. Ali et al. (2024) reported an internal downscaling correlation near 0.98, whereas improvement against independent groundwater observations was more modest (approximately 0.75 to 0.77). Finer output grids therefore should not be interpreted as finer independent gravimetric information. Dai et al. (2024) show the value of combining gravity-based storage estimates with deformation and attribution information: reported shallow and deep storage trends differed substantially, and human contributions were estimated at high fractions of the observed decline. Throughout C3, the strongest designs combine defensible component separation and independent groundwater monitoring, data integration that takes uncertainty into account, and clear spatial interpretation limits.

The Land Subsidence Related to Groundwater (C4)

C4 clearly explains why observability must be separated from causality. InSAR provides dense deformation measurements, but a deformation map alone does not establish groundwater causation. Chen et al. (2016) combined satellite radar interferometry with GPS and groundwater information in Beijing, reporting GPS–InSAR agreement on the order of a 2.94 mm/yr RMS and correlation above 0.98. Zhou et al. (2017) used InSAR time series to evaluate land subsidence across land-use classes in the eastern Beijing Plain, while Tang et al. (2024) combined Sentinel-1 time-series analysis with hydrogeological and contextual evidence in Ningbo. These designs are stronger when geodetic measurements directly check deformation and groundwater variation is linked analytically rather than discussed only in context.

The most persuasive C4 architectures connect deformation to a physical aquifer-system model. Bockstiegel et al. (2024) calibrated groundwater-flow/compaction behavior with InSAR and leveling information in the Rafsanjan Plain. They reported up to approximately 21 cm/yr subsidence, about 8.8 km³ of irreversible storage-capacity loss, and a model RMSE of roughly 2.5 cm/yr; management scenarios reduced modeled rates by up to about 50%. However, Duffy et al. (2020) demonstrate the opposite, as high-quality PSI deformation maps may remain insufficiently hydrogeologically constrained when there is no direct correlation between pumping, hydraulic pressures, or aquifer mechanics with the measured displacement.

Long, multi-mission records can add temporal value, but this value is not necessarily causal. Kim et al. (2015) and Doo et al. (2019) demonstrated how decades-long ground subsidence records can be obtained by multi-aperture synthetic aperture radar missions. Their main contributions are temporal continuity and spatial coverage; however, groundwater source determination still depends on pressure/pump history, aquifer analysis, process modeling, and geodetic investigations. This distinction was rigorously applied in the eligibility screening process, where the contextual or assumed formulation of groundwater was insufficient for C4 eligibility when the contribution of groundwater was not measured or analytically separated.

Value-Added Mechanisms Across Fields

In all 359 documented studies, integrations add value through a limited set of iterative inference functions rather than through the number of sensors per se. These functions include improving prediction using an external reference, providing independent physical or hydrochemical screening, enhancing the identification of competing causes, linking metrics to explicit supporting matching, extending time continuity across missions or monitoring systems, and reducing or improving the characterization of uncertainty through model consolidation or closure. A single study can contribute to multiple functions. However, claims of superiority are stronger when the additional contribution is isolated through comparison, exclusion, or an independent verification pathway.

The ten most represented research domains in the screened literature are shown in Figure 4.

The subject records distribution shows a clear concentration in the key zones associated with groundwater, where hydrogeological domain receded the highest No. of 1,842 record, and then groundwater domain with No. of 1,521 record, then land subsidence with no. of 1,348 record. Also, remote sensing and interferometric Synthetic Aperture Radar show a notable representation with No. of 1.112 and 987 consecutive record. Conversely, No. of records was gradually decreased in the most specialized field, including GRASE/GRASE-FO, GIS, groundwater depletion, urban area, and water resources management. This pattern indicates that recovery literature is mainly focused about the hydrogeological description; groundwater and subsidence changing monitoring; and employing a diverse range of monitoring and geospatial analysis techniques.

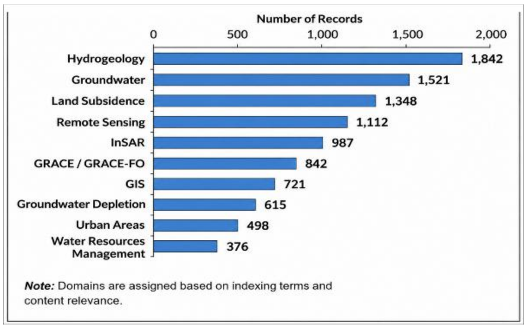

Figure 4: Top 10 Research Domains by Number of Records, Highlighting the Thematic Concentration of the Screened Literature

Representative primary-source examples supporting these inferential functions are summarized in Table 8. This Table represented examples shows that inferential value of integration did not link by increasing No. of sensors or the total sources themselves, but by the power source used for validation and its capability of independent testing hydrogeological claim. C1 show improvement through sample comparison, and link the potential maps with groundwater data geophysics measurements, while the powerful evidences of C2 were based on comparing satellite-derived feed estimates with independent assessment pathways. In C3, there is a benefit from merging GRASE with field measurements to improvement performance and checking for storage changes, while C3 shows that the deformation measurement precision by using InSAR becomes more hydrogeological valuable when it's linked to geodetic measurements, or groundwater data, or compression models. Generally, these results supporting inferential gain concept as a stronger criterion than simply having multiple data sources.

Table 8: Representative Primary-Source Signals Illustrating Inferential Gain and its Limits

Domain / study	Integration and reference	Representative result	Interpretive value
C1 – Rehman et al. (2024)	GIS/RS predictors; FR versus WoE	FR success/prediction 89%/93% versus WoE 82%/86%	Supports model-selection gain; not sensor-count superiority.
C1 – Nainggolan et al. (2024)	Multiple RS inputs; observed groundwater availability	Correlation 0.56 (AHP) → 0.67 (FE-AHP)	External hydrogeological reference strengthens construct validity.
C1 – Sangawi et al. (2023)	GIS/RS/AHP + eight ERT profiles	ERT correspondence with water-bearing classes	Independent subsurface constraint on mapped potential.
C2 – Coelho et al. (2017)	Satellite water balance + WTF method	154.6 mm WTF versus 124.6 mm RS-WB in 2011	Comparing flows across different routes allows for a quantitative assessment.
C2 – Alghafli et al. (2023)	GRACE + in situ recharge pathway	R^2 0.91; NSE 0.7; RMSE ≈1.5–7.8 Mm^3	Testing an independent path for satellite-derived recharging.
C3 – Su et al. (2022)	GRACE/model fusion + in situ groundwater	CC +9–40%; NSE +23–657%; RMSE −9–28%	The benefit of gradual integration is clear compared to the reference..

Domain / study	Integration and reference	Representative result	Interpretive value
C3 – Ali et al. (2024)	Machine-learning GRACE downscaling + groundwater data	Internal r ≈0.98; external groundwater r ≈0.75–0.77	It warns that high output accuracy is not independent, accurate information.
C3/C4 – Raju et al. (2024)	Sentinel-1 + GRACE + water levels + rainfall	Cross-domain trend and deformation triangulation	It links storage, distortion, and ground observations.
C4 – Chen et al. (2016)	InSAR + GPS + groundwater information	GPS–InSAR RMS ≈2.94 mm/yr; r >0.98	Verifying the strong deformation; the causal link still requires groundwater.
C4 – Bockstiegel et al. (2024)	groundwater/compaction model + leveling calibrated by InSAR	Irreversible capacity loss of approximately 8.8 km³; mean root-squared error ≈ 2.5 cm/year	Automated interconnection transforms deformation into effects on the aquifer system.

The summary of quality assessment are presents in Table 9. Critical evaluation of the 359 verified studies shows that low concern category was dominated in all Q1-Q12 domains but its strength varied among evaluation elements. A high ratio of 64.3% was recorded for low concern in building suitability Q1, while other ratios recorded 63.5% for environmental and suitability Q2 standard, and 62.1% for integration transparency Q5. Conversely, the low value of 45.7% of low concern was for confounding and attribution Q9, followed by uncertainty and sensitivity Q8 with 47.6 ratio. Also, Q9 records high ration of high concern which was 9.7%, while it reached 8.9% in both reliability/data leakage Q7 and uncertainty Q8. These results indicate that essential sides of hydrogeological structure and study preparation generally was stronger than aspects related to causal attribution, uncertainty, and verification independence.

Table 9: The Summary of Quality Assessment Based on the Q1-Q12 Framework for Documented Evidence (n = 359)

Domain	Low Concern n (%)	Some Concerns n (%)	High Concern n (%)	Unclear / NA n (%)
Q1 Construct Validity	231 (64.3)	104 (29.0)	16 (4.5)	8 (2.2)
Q2 Setting and Scale	228 (63.5)	101 (28.1)	18 (5.0)	12 (3.3)
Q3 Data Provenance & Suitability	210 (58.5)	116 (32.3)	22 (6.1)	11 (3.1)
Q4 Preprocessing & Alignment	176 (49.0)	137 (38.2)	30 (8.4)	16 (4.4)
Q5 Integration Transparency	223 (62.1)	102 (28.4)	18 (5.0)	16 (4.5)
Q6 Validation Independence	193 (53.8)	111 (30.9)	28 (7.8)	27 (7.5)
Q7 Dependence / Leakage	186 (51.8)	114 (31.8)	32 (8.9)	27 (7.5)
Q8 Uncertainty & Sensitivity	171 (47.6)	128 (35.7)	32 (8.9)	28 (7.8)
Q9 Confounding & Attribution	164 (45.7)	126 (35.1)	35 (9.7)	34 (9.5)
Q10 Transferability	198 (55.2)	106 (29.5)	26 (7.2)	29 (8.1)
Q11 Reporting & Reproducibility	214 (59.6)	103 (28.7)	19 (5.3)	23 (6.4)
Q12 Outcome-specific Validity	207 (57.7)	107 (29.8)	22 (6.1)	23 (6.4)

Note: Percentages may not sum to 100 due to rounding.

Hierarchy of Evidence Strength in Integrated Groundwater Science

Documented evidence can support a practical five-level hierarchy of evidentiary strength. The first level is objective integrity without an independent groundwater reference; it can support spatial prioritization but does not support strong causal accuracy or strong claims. Level 2 adds field verification or well-based verification, which improves the credibility of the results, with the ability to verify the presence of the structure only, rather than verifying its structure itself. Level 3 introduces an independent physical or hydrochemical constraint—such as vertical electrical scanning, electrical resistivity imaging, isotope analysis, groundwater availability, or geodetic measurements—that directly tests a crucial part of the inference. Level 4 can add a clear path for independent comparison, exclusion, or estimation, which allows for defensible data with respect to performance improvement or quantitative consistency. Level 5 links observations to a groundwater mechanical model or aquifer pressure model, which includes uncertainty testing and scenario analysis. This hierarchy cannot represent a measure of quality nor does it replace levels 1 to 12, but rather aims to differentiate between increasing data volume and increasing hydrogeological information.

The resulting of these levels of evidence resulting in a coherent conceptual framework that connect satellite observation with ground-based data and environmental and anthropogenic factors, then these sources pass through data processing and integration before its used in the analysis and modelling. The series concluding with a connected output with changes in groundwater storage; sustenance patterns; risk assessment and managements support, with presence feedback loop which allows for improved integration and iterative analysis. This path reflects the main results of the review that the power of reasoning improves when the independence evidences are integrate across connected stages instead of simply compiling data sources. Figure 5 summarizes the resulting multi-source integration path.

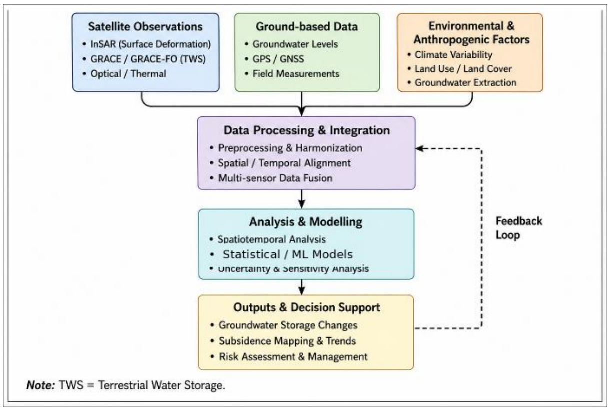

Figure 5: The Conceptual Framework for Multi-Source Integration in Groundwater Dynamics and Land Subsidence Assessment

The minimum proof requirements for each field are summarized in Table 10. The comparison of the four domains showed that the minimum evidence required for conclusion supporting varies depending on the hydrological structure to be proven. In C1, the claims related to the presence of groundwater, productivity, or reservoir engineering, with independence data from wells, production tests, or geophysics characteristics. In C2, recharge area maps require supporting from tracers, hydrochemistry, or wells, while feed flows estimates requires an independence quantity path like water balance or calibration modelling. C3 requires splitting the non-underground storage components, uncertainty estimation, and ground verification as possible, while C4 requires correction and appropriates retrospective for InSAR data, and linking deformation to levels, pumping, hydro stratigraphy, geodetic measurements, or compression models. So, proving requirements differs according to nature of claim and can not using a unified verification standard for all areas.

Table 10: Minimum Field-Specific Evidence Standards for Defensible Integrated Hydrogeological Claims

Domain	Claim to support	Minimum independent constraint	Claim to avoid without stronger evidence
C1	Groundwater occurrence / productivity / geometry	Representative wells, productivity tests, aquifer thickness/depth, or hydrogeophysical properties corresponding to the structure.	Labeling the thematic suitability map as a groundwater aquifer characteristics map, or claiming sensor superiority without a benchmark for comparison.
C2	Recharge zone or recharge flux	For areas: evidence from indicators/hydrochemistry/wells; for flow: independent water balance, water flow model, indicator, or calibrated process model	Equalizing the rechargeability with the actual recharge rate.
C3	Groundwater-storage anomaly/trend	A reasonable assumption regarding non-groundwater components, in addition to the uncertainty; independent groundwater testing where possible.	Treating TWS as groundwater or presenting downscaled pixels as independent fine-scale gravimetry.
C4	Groundwater-induced subsidence / compaction	InSAR correction/reference plus heads, pumping, hydrographic strata, geodesy, or compression model	Attributing deformation to groundwater from co-location or contextual wording alone.

Evidence-Based Priority Research Agenda

In the future, integrated groundwater science should move from collecting data from sensors to hypothesis-based information design. The evidence supports a research agenda built on clear concepts, independent verification and comparisons, matching the scale of support, the spread of uncertainty, ground observations with causal information, and reproducible processing. Verification across different sites and proactive monitoring experiments are of particular importance, because high accuracy within one basin may not transfer to different hydrogeological environments.

The results of the composition identified eight connected priorities to improve the future studies, begin by identifying the hydrogeological structure and the target to be tested before chose the sensors, then using the comparator and ablation to a true value isolation for each data source. Also, there is a needing to independence verification, uncertainty distribution, matching spatial and temporal scales, strengthening causal attribution, workflow documentation to ensure reapplication, and finally susceptibility test of results transporting between basins and different time periods. These priorities taken together indicate a required shift in design of studies based on data availability towards a design based on assumptions, information value, and hydrogeological inference.

Building on these standards, the priority research agenda is summarized in Table 11.

Table 11: Priority Research Agenda for Integrated Hydrogeological Remote Sensing and GIS

Priority	Why it matters	Recommended design / reporting action
Construct-first design	Prevents proxy substitution and overclaiming	Define the hydrogeological outcome and validation target before selecting sensors or predictors.
Comparator / ablation	Separates sensor gain from algorithm or data-volume gain	Report a single-source/simpler baseline and remove added streams one at a time where feasible.
Independent validation	Internal fit can conceal leakage or shared assumptions	Use spatially/temporally independent wells, tracers, geophysics, or geodesy; report independence explicitly.
Uncertainty propagation	Groundwater inference often depends on subtraction or fusion	Transfer the uncertainty in the key components, measurement, modeling, and quantification to the reported groundwater quantity.
Scale matching	Remote and ground observations have different supports	Report on original support, regrouping, compilation, and the scale at which claims remain valid.
Causal attribution	Correlated environmental trends are not causal evidence	Test competing drivers and use pumping/head/hydrostratigraphy or mechanistic models when making causal claims.
Reproducibility	Complex workflows are difficult to audit	Report product versions, code, parameters, thresholds, training/validation splits, and provenance.
Transferability	Performance may be basin-specific	Use cross-basin or cross-time validation and bound conclusions to tested hydrogeological settings.

Discussion

This analysis points to a crucial distinction in integrated groundwater science, where the difference here is not simply between single-sensor and multi-sensor studies, but rather between descriptive integration and constrained inference. Many workflows combine multiple layers, but they are only realized through internal map consistency or through an alternative model that does not match the alleged hydrogeological structure. The simplest design may be the most scientifically robust when a single, independent dataset directly tests the main inferential link. Therefore, the number of sensors, the level of integration, and the strength of the evidence should be considered separately. This conclusion agree with studies results of Sangawi et al., (2023) and Nainggolan et al., (2024), as groundwater potential became more powerful when linked with independence hydrogeological and geophysics information, and not just an increasing number of used layers. Also, agree with results of Arabameri et al., (2019) and Rehman et al., (2014), which shows that performance improved may be results from choice of model or algorithm, so it should not be automatically interpreted as a benefit resulting from the multiplicity of sensors.

The results show that the most important sources of common error between domains C1-C4, is replacing the hydrogeological structure to be measured with an indicator or substitute that does not fully represent it. In C1, the groundwater potential map did not in necessary the production, or aquifer thickness or structure, and in C2 the recharge possibility is not equal to actual recharge rate, also, in C3 the terrestrial water storage did not represent groundwater storage unless other components are splits. While, deformation measures in C4 do not proves that it is caused by groundwater without independent evidence of causation. These differences ensure that the verification method must be matched with the hydrogeological structure that the study claims to measure.

Hydrogeological added value can best be understood as a tangible improvement in the accuracy, reliability, and verifiability of monitoring, sourcing, or interpreting the process of groundwater inference, and can therefore be attributed to an additional data flow or integration step. Inferential improvement can take the form of a reduction in prediction error compared to an external reference, improved properties, a narrower uncertainty range, or confirmation through an independent physical measurement. This may also include better separation of groundwater from overlapping reservoir components, a wider temporal or spatial range of observation, stronger differentiation between competing causal explanations, or a more complete observational model of the process. This definition allows it to avoid rewarding complexity for its own sake.

This interpretation agrees with study of Su et al. (2022), where it can prove the addition value to merge operation through improved performance indicators as comparison with original estimates and based on independent field subsurface measurements. In contrast, study of Ali et al. (2024) that high internal correlation after decreasing GRASE data measurements do not reflect in the same power when compared with independence groundwater observations, which support the necessity of distinguish between the internal suitability quality and actual hydrogeological value.

A direct comparison can be particularly important when a study claims that combining multiple sensors improves performance. The effect of sensor diversity cannot be isolated if the algorithm, calibration data, and sensors are changed simultaneously. Therefore, ablation designs, alternative component separation pathways, single-sensor baselines, or independent estimation methods may be more informative than a high final AUC value or correlation alone.

Measurement mismatch continues to pose a significant threat to the validity of results across various fields. In domain C1, wells and geophysical sections are limited only by the extent of support they cover, and even intensive verification may still be spatially dependent on model development data. In C2, maps of potential areas should not be validated as if they were flow estimates, whereas satellite water balance or GRACE recharge requires comparison on compatible time and spatial scales. In C3, the correlation with internal expansion may be much stronger than the correlation with independent groundwater observations. In C4, GPS, leveling, and expansion measurements also directly confirm deformation more than they confirm causation, while regional GRACE signals cannot explain the localized deformation caused by InSAR without intermediate hydrogeological evidence.

This result agrees with studies of Coelho et al. (2017) and Alghafli et al. (2023), where the recharge estimates comparison with independence hydrogeological paths is more beneficial than depending on internal compatibility only. Also, the results of Chen et al. (2016) study, agree with this distinction in C4, so the strong agreement between GPS and InSAR confirms the reliability of the deformation measurements, but can not prove alone that groundwater is the reason, which is requires additional hydrogeological evidences.

Uncertainty not only improves reporting methods but also alters how integration is interpreted. Multiple hydrological models may provide a useful range of component separation estimates; however, these models may share impact data as well as structural assumptions. Remote sensing products may be statistically correlated. Therefore, combining data that treats correlated inputs as independent may produce overconfident results. Reliable studies point to factors such as product selection, covariance or sensitivity, spatial support, reference data limitations, and the transmission of uncertainty to the actual interpreted groundwater volume. That is agree with C3 results in present review, especially studies depending on GRASE, where the process of separating storage components, spatial scale, and model adoption remained major sources of uncertainty even when combining several products or models.

The most robust multi-domain designs utilize different data streams to constrain various links within the same hydrogeological inference chain. Wells and groundwater availability constrain their occurrence and storage; electrical resistivity/vertical electrical surveying and other geophysical techniques constrain subsurface geometry; isotopes constrain recharge origin or recency; satellite water balance parameters constrain recharge flow; the GRACE project constrains basin-scale mass change; InSAR analyzes deformation; and GPS/leveling/extensometer data confirm displacement. Calibrated groundwater/compression models can translate these observations into aquifer system mechanisms. The study by Raju et al. (2024) is particularly significant because it correlates GRACE, InSAR, precipitation, and piezometric data across the C3 and C4 layers. Bockstiegel et al. (2024) go a step further by incorporating deformation into a calibrated groundwater/compression model. So, these two studies agree with the general result which is the strongest integration is not the one that gathers the most data, rather, that is used in it different sources to restrict an independent rings within the same hydrogeological inference chain.

The integration of different fields has direct implications for the design of the study. In the C1 phase, the reference target must be defined before the model is built because the existence of the well, its productivity, saturation thickness, hydraulic pressure, and reservoir geometry are not interchangeable results, so the verification process must be consistent with the defined structure. In C2, the possibility of recharge should remain an element for determining spatial priorities unless it is associated with a quantitative flow through water balancing, groundwater level fluctuations, indicators, calibration process modeling, or another independent pathway. In C3, GRACE-based groundwater estimates require defensible segregation of components and consideration of uncertainty, whereas microproducts should be evaluated against independent groundwater observations rather than internal fit alone. In C4, InSAR measures deformation; the cause of groundwater only becomes convincing when the deformation is related to pumping, pressures, water layers, process-based compression modeling, or geodetic investigations.

These conclusions are compatible with examples of studies examined, study of Sangawi et al. (2023) was supported the C1 maps with measurements of electrical resistance, also, study of Coelho et al. (2017) used an independent path to comparing the recharge estimates in C2, while Bockstiegal et al. (2024) study links the deformation with flow model and tank compression in C4. These examples illustrate that is type of addition required evidence is different according to uncertainty position in the chain of reasoning.

These differences suggest that monitoring networks should be designed around inferential bottlenecks rather than sensor availability. When there is prevailing uncertainty regarding the presence of groundwater, wells or geophysics may add more information than any other associated surface indicator. For recharge flow, having an independent flow constraint or tracer is more useful than having additional objective layers. For groundwater storage separation, uncertainty in the composition and in-situ groundwater records is critical. Regarding the causes of land subsidence, studying the water layers and pumping/pressure history may add more value than a second deformation product. Therefore, the design of multiple sensors can be viewed as an information value problem: each additional data set must be justified by the ambiguity it resolves, the uncertainty it reduces, or the resolution it improves.

Strengths and Limitations

One of the key strengths of this review is the clear separation between the hydrogeological field, the integration architecture, the diversity of sensors, and their interrelationships. This distinction prevents confusion between the complexity of remote sensing and the strength of hydrogeological evidence. The sensitivity of the search was verified against eight predefined empirical criteria, all of which were recovered by the final strategy. Furthermore, the selection process preserved the identity of the records, applied rigorous rules for results, linked accompanying reports, and verified procedures according to the publication pathway rather than disqualifying them entirely.

The second strength lies in the distinction between the extensive archive and the core of the meticulously researched synthesis. The final archive contains 4,014 report/study units, while the detailed inferential claims and Q1-Q12 assessments are limited to 359 meticulously researched primary-source studies. This limitation reduces the risk of turning decisions regarding title/abstract selection or metadata into unsupported statements about the quality of the study.

A complete report-by-report retrieval-attempt ledger is not available for the full 6,824-record eligibility set. Standard PRISMA counts for reports not retrieved or uniformly assessed at full text therefore cannot be reconstructed, and the retained records are not treated as though they had all undergone equivalent full-text verification. The 4,014 retained inventory should be interpreted as the final selection set rather than as a conventional uniform full-text denominator.

The primary bibliographic search was limited to Scopus. Although benchmark validation, broad vocabulary, and citation-based diagnostics improved sensitivity, eligible studies not indexed in Scopus may have been missed. Verification depth also varies within the 359-study critical-synthesis core, ranging across full primary text, repository manuscripts, detailed publisher content, and report-level methods/results evidence. This variation is carried into study-level appraisal and limits simple prevalence interpretations of methodological strengths or weaknesses.

The review also features a deliberate diversity in findings and methodology. Groundwater potential mapping, recharge rate estimation, GRACE residual storage analysis, and groundwater-related subsidence do not share a single significant effect size. Therefore, the absence of comprehensive aggregate analysis is a design choice, not a failure of aggregates. Quantitative signals are interpreted within comparable methodological frameworks and in light of the validity of the underlying hydrogeological structure.

CONCLUSION

The present review shows that the multi-sensor hydrogeology evaluation should depend on inferential gain, not the number of sensors or data sources. In C1, aquifer maps is strengthens when related to independent evidences about presence of groundwater, or productivity, or aquifer structure; in C2, the concept of recharge potential must be distinguished from the concept of actual recharge flow; in C3, the GRACE-based groundwater storage inference requires defensible aquifer component separation, with uncertainty-conscious and independent verification; while in C4, the InSAR measurement provides effective monitoring of deformation, but attributing it to groundwater requires evidence of pumping, levels, aquifer characteristics, or suitable mechanical modeling.

Generally, additional data flow is of scientific value when it contributes to adding a non-redundant constraint, improves an explicit comparator, reduces uncertainty, expands a key dimension of observation, enhances causal attribution, or bridges a defined gap between process and inference. The future studies should be focus on comparison, baselines, or resections, independent verification, matching spatial and temporal scales, disseminate uncertainty, and reapplication capability. Out of 6,824 records that underwent eligibility assessment, 4,014 retained report/study units, while the critical structure and details evaluation Q1-Q12 were based to 359 studies that were verified from primary sources; which is necessary distinction to maintain the verifiability and scientific strength of conclusions.

REFERENCES

Alghafli, K., Shi, X., Sloan, W., Shamsudduha, M., Tang, Q., Sefelnasr, A., & Ebraheem, A. A. (2023). Groundwater recharge estimation using in situ and GRACE observations in the eastern region of the United Arab Emirates. Science of the Total Environment, 867, 161489. https://doi.org/10.1016/j.scitotenv.2023.161489

Ali, S., Ran, J., Khorrami, B., Wu, H., Tariq, A., Jehanzaib, M., Khan, M. M., & Faisal, M. (2024). Downscaled GRACE/GRACE-FO observations for spatial and temporal monitoring of groundwater storage variations at the local scale using machine learning. Groundwater for Sustainable Development, 25, 101100. https://doi.org/10.1016/j.gsd.2024.101100

Arabameri, A., Rezaei, K., Cerdà, A., Lombardo, L., & Rodrigo-Comino, J. (2019). GIS-based groundwater potential mapping in Shahroud plain, Iran: A comparison among statistical, data-mining and MCDM approaches. Science of the Total Environment, 658, 160–177. https://doi.org/10.1016/j.scitotenv.2018.12.115

Bockstiegel, A., Heße, F., Doummar, J., & Kolditz, O. (2024). Simulation of present and future land subsidence in the Rafsanjan plain, Iran, due to groundwater overexploitation using numerical modeling and InSAR data analysis. Hydrogeology Journal, 32(1), 289–305. https://doi.org/10.1007/s10040-023-02657-y

Castle, S. L., Thomas, B. F., Reager, J. T., Rodell, M., Swenson, S. C., & Famiglietti, J. S. (2014). Groundwater depletion during drought threatens future water security of the Colorado River Basin. Geophysical Research Letters, 41, 5904–5911. https://doi.org/10.1002/2014GL061055

Chaussard, E., Amelung, F., Abidin, H., & Hong, S.-H. (2013). Sinking cities in Indonesia: ALOS PALSAR detects rapid subsidence due to groundwater and gas extraction. Remote Sensing of Environment, 128, 150–161. https://doi.org/10.1016/j.rse.2012.10.015

Chen, B., Gong, H., Li, X., Lei, K., Ke, Y., Duan, G., & Zhou, C. (2016). Imaging land subsidence induced by groundwater extraction in Beijing (China) using satellite radar interferometry. Remote Sensing, 8(6), 468. https://doi.org/10.3390/rs8060468

Coelho, V. H. R., Montenegro, S. M. G. L., Almeida, C. N., Silva, B. B., Oliveira, L. M., Gusmão, A. C. V., Freitas, E. S., & Montenegro, A. A. A. (2017). Alluvial groundwater recharge estimation in semi-arid environment using remotely sensed data. Journal of Hydrology, 548, 1–15. https://doi.org/10.1016/j.jhydrol.2017.02.054

Dai, Y., et al. (2024). Tracking shallow and deep groundwater storage changes in North China Plain with improved fusion method and hybrid spectral analysis approach. Journal of Hydrology, 633, 131001. https://doi.org/10.1016/j.jhydrol.2024.131001

Dehghani, M., Valadan Zoej, M. J., Hooper, A., Hanssen, R. F., Entezam, I., & Saatchi, S. (2013). Hybrid conventional and Persistent Scatterer SAR interferometry for land subsidence monitoring in the Tehran Basin, Iran. ISPRS Journal of Photogrammetry and Remote Sensing, 79, 157–170. https://doi.org/10.1016/j.isprsjprs.2013.02.012

Dekongmen, B. W., et al. (2022). Groundwater recharge estimation and potential recharge mapping in the Afram Plains of Ghana using SWAT and remote sensing techniques. Groundwater for Sustainable Development, 17, 100741. https://doi.org/10.1016/j.gsd.2022.100741

Du, Z., Ge, L., Ng, A. H.-M., Zhu, Q., Zhang, Q., Kuang, J., & Dong, Y. (2019). Long-term subsidence in Mexico City from 2004 to 2018 revealed by five synthetic aperture radar sensors. Land Degradation & Development, 30(15), 1785–1801. https://doi.org/10.1002/ldr.3347

Duffy, C. E., Braun, A., & Hochschild, V. (2020). Surface subsidence in urbanized coastal areas: PSI methods based on Sentinel-1 for Ho Chi Minh City. Remote Sensing, 12(24), 4130. https://doi.org/10.3390/rs12244130

El-Bagoury, K., et al. (2025). Potentiality delineation of groundwater recharge in arid regions using multi-criteria analysis. Water, 17(5), 766. https://doi.org/10.3390/w17050766

El-Hadidy, M. (2024). Towards sustainable development goals: Leveraging multi-data remote sensing fusion for monitoring groundwater-induced bedrock subsidence dynamics in Egypt's Nile Valley. Groundwater for Sustainable Development, 27, 101353. https://doi.org/10.1016/j.gsd.2024.101353

Erban, L. E., Gorelick, S. M., & Zebker, H. A. (2014). Groundwater extraction, land subsidence, and sea-level rise in the Mekong Delta, Vietnam. Environmental Research Letters, 9, 084010. https://doi.org/10.1088/1748-9326/9/8/084010

Gaber, A., Mohamed, A. K., ElGalladi, A., Abdelkareem, M., Beshr, A. M., & Koch, M. (2020). Mapping the groundwater potentiality of West Qena Area, Egypt, using integrated remote sensing and hydro-geophysical techniques. Remote Sensing, 12(10), 1559. https://doi.org/10.3390/rs12101559

Gemitzi, A., Ajami, H., & Richnow, H.-H. (2017). Developing empirical monthly groundwater recharge equations based on modeling and remote sensing data—Modeling future groundwater recharge to predict potential climate-change impacts. Journal of Hydrology, 546, 1–13. https://doi.org/10.1016/j.jhydrol.2017.01.005

Gumma, M. K., & Pavelic, P. (2013). Mapping of groundwater potential zones across Ghana using remote sensing, geographic information systems, and spatial modeling. Environmental Monitoring and Assessment, 185, 3561–3579. https://doi.org/10.1007/s10661-012-2810-y

Ki, I., Chakroun, H., Koussoube, Y., & Zouari, K. (2023). Assessment of aquifer recharge potential using remote sensing, GIS and the analytical hierarchy process combined with hydrochemical and isotope data (Tamassari Basin, Burkina Faso). Water, 15(4), 650. https://doi.org/10.3390/w15040650

Kim, J.-W., Lu, Z., Jia, Y., & Shum, C. K. (2015). Ground subsidence in Tucson, Arizona, monitored by time-series analysis using multi-sensor InSAR datasets from 1993 to 2011. ISPRS Journal of Photogrammetry and Remote Sensing, 107, 126–141. https://doi.org/10.1016/j.isprsjprs.2015.03.013

Li, J., & Ma, J. (2024). Evaluating the dynamics of groundwater storage and its sustainability in the Loess Plateau: The integrated impacts of climate change and human activities. Remote Sensing, 16(23), 4375. https://doi.org/10.3390/rs16234375

Magnoni, P. H. J., Silva, C. O. F., & Manzione, R. L. (2020). Groundwater recharge and water table levels modelling using remotely sensed data and cloud-computing. Sustainable Water Resources Management, 6, 113. https://doi.org/10.1007/s40899-020-00469-6

Maqsoom, A., Aslam, B., Khalid, N., Ullah, F., Anysz, H., Almaliki, A. H., Almaliki, A. A., & Hussein, E. E. (2022). Delineating groundwater recharge potential through remote sensing and geographical information systems. Water, 14(11), 1824. https://doi.org/10.3390/w14111824

McGowan, J., Sampson, M., Salzwedel, D. M., Cogo, E., Foerster, V., & Lefebvre, C. (2016). PRESS peer review of electronic search strategies: 2015 guideline statement. Journal of Clinical Epidemiology, 75, 40–46. https://doi.org/10.1016/j.jclinepi.2016.01.021

Milewski, A., Sultan, M., Yan, E., Becker, R., Abdeldayem, A., Soliman, F., & Abdel Gelil, K. (2009). A remote sensing solution for estimating runoff and recharge in arid environments. Journal of Hydrology, 373(1–2), 1–14. https://doi.org/10.1016/j.jhydrol.2009.04.002

Modibbo, M. A., Kana, A. A., Bello, I. E., & Eya, A. I. (2025). An integrated remote sensing, geographic information system and analytical hierarchy process for determination of groundwater potential in Keffi-GRA and environs. FUDMA Journal of Sciences, 9(1), 16–28. https://doi.org/10.33003/fjs-2025-0901-2769

Naghibi, S. A., Pourghasemi, H. R., & Dixon, B. (2016). GIS-based groundwater potential mapping using boosted regression tree, classification and regression tree, and random forest machine learning models in Iran. Environmental Monitoring and Assessment, 188, 44. https://doi.org/10.1007/s10661-015-5049-6

Nainggolan, L., Ni, C.-F., Darmawan, Y., Lo, W.-C., Lee, I.-H., Lin, C.-P., & Hiep, N. H. (2024). Cost-effective groundwater potential mapping by integrating multiple remote sensing data and the index–overlay method. Remote Sensing, 16(3), 502. https://doi.org/10.3390/rs16030502

Page, M. J., McKenzie, J. E., Bossuyt, P. M., et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ, 372, n71. https://doi.org/10.1136/bmj.n71

Raju, A., Singh, R. P., Kannojiya, P. K., Patel, A., Singh, S., & Sinha, M. (2024). Declining groundwater and its impacts along Ganga riverfronts using combined Sentinel-1, GRACE, water levels, and rainfall data. Science of the Total Environment, 920, 170932. https://doi.org/10.1016/j.scitotenv.2024.170932

Rehman, A., Xue, L., Islam, F., Ahmed, N., Qaysi, S., Liu, S., Alarifi, N., Youssef, Y. M., & Abd-Elmaboud, M. E. (2024). Unveiling groundwater potential in Hangu District, Pakistan: A GIS-driven bivariate modeling and remote sensing approach for achieving SDGs. Water, 16(22), 3317. https://doi.org/10.3390/w16223317

Rethlefsen, M. L., Kirtley, S., Waffenschmidt, S., et al. (2021). PRISMA-S: an extension to the PRISMA Statement for reporting literature searches in systematic reviews. Systematic Reviews, 10, 39. https://doi.org/10.1186/s13643-020-01542-z

Rodell, M., Velicogna, I., & Famiglietti, J. S. (2009). Satellite-based estimates of groundwater depletion in India. Nature, 460, 999–1002. https://doi.org/10.1038/nature08238

Sangawi, A., Al-Manmi, D. A. M., & Aziz, B. Q. (2023). Integrated GIS, remote sensing, and electrical resistivity tomography methods for the delineation of groundwater potential zones in Sangaw Sub-Basin, Sulaymaniyah, KRG-Iraq. Water, 15(6), 1055. https://doi.org/10.3390/w15061055

Shamseer, L., Moher, D., Clarke, M., et al. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ, 349, g7647. https://doi.org/10.1136/bmj.g7647

Su, K., Zheng, W., Yin, W., Hu, L., & Shen, Y. (2022). Improving the accuracy of groundwater storage estimates based on groundwater weighted fusion model. Remote Sensing, 14(1), 202. https://doi.org/10.3390/rs14010202

Tang, W., Ng, A. H.-M., Wang, H., Kuang, J., & Du, Z. (2024). Surface subsidence characteristics and causes analysis in Ningbo Plain by Sentinel-1A TS-InSAR. Remote Sensing, 16(13), 2438. https://doi.org/10.3390/rs16132438

Zhou, C., Gong, H., Chen, B., Li, J., Gao, M., Zhu, F., Chen, W., & Liang, Y. (2017). InSAR time-series analysis of land subsidence under different land use types in the eastern Beijing Plain, China. Remote Sensing, 9(4), 380. https://doi.org/10.3390/rs9040380