

Expanding groundwater drawdown funnels in North China Plain: A new insight from satellite gravimetry

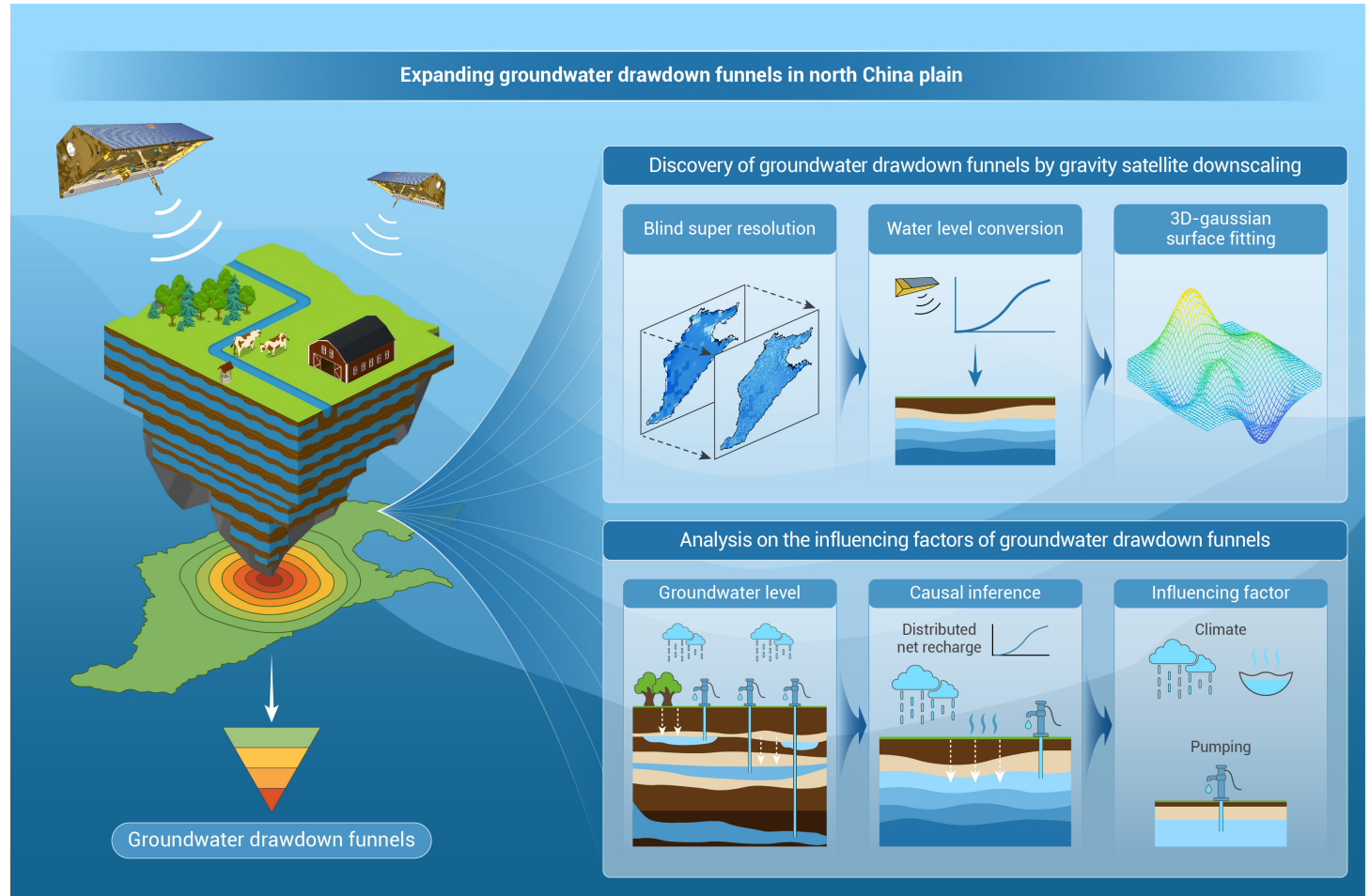
Jiaqi Yao,^{1,2,11} Huiqing Hao,^{2,3,4,11} Mengran Wang,^{1,2,5,11} Walter A. Illman,³ Guoqiang Wang,⁵ Tian-Chyi Jim Yeh,⁶ Peng Gong,⁷ Qingquan Li,^{8,9} Nan Xu,^{8,9,*} Yonghong Hao,^{1,2,4,*} and Hui Lu^{10,*}

*Correspondence: hhuxunan@gmail.com (N.X.); haoyh@sxu.edu.cn (Y.H.); luhui@tsinghua.edu.cn (H.L.)

Received: February 1, 2026; Accepted: July 18, 2026; Published Online: July 23, 2026; <https://doi.org/10.59717/j.tiw.2026.100010>

© 2026 The Author(s). This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

GRAPHICAL ABSTRACT



PUBLIC SUMMARY

- Blind downscaling reveals 5 km funnel expansion trends in North China Plain.
- Funnel morphology changed with > 30 m deepening and horizontal broadening.
- Response time rose from 20 to 60 months, indicating aquifer buffering decline.
- Causal analysis shows early climate control, later pumping, then partial rebound.

Expanding groundwater drawdown funnels in North China Plain: A new insight from satellite gravimetry

Jiaqi Yao,^{1,2,11} Huiqing Hao,^{2,3,4,11} Mengran Wang,^{1,2,5,11} Walter A. Illman,³ Guoqiang Wang,⁵ Tian-Chyi Jim Yeh,⁶ Peng Gong,⁷ Qingquan Li,^{8,9} Nan Xu,^{8,9,*} Yonghong Hao,^{1,2,4,*} and Hui Lu^{10,*}

¹Academy of Ecological Civilization Development for JING-JIN-JI Megalopolis, Tianjin Normal University, Tianjin 300387, China

²Faculty of Geography, Tianjin Normal University, Tianjin 300387, China

³Department of Earth and Environmental Sciences, University of Waterloo, Waterloo ON N2L 3G1, Canada

⁴Tianjin Key Lab of Water Resources and Environment, Tianjin Normal University, Tianjin 300387, China

⁵Advanced Interdisciplinary Institute of Satellite Applications, Beijing Normal University, Beijing 100875, China

⁶Department of Hydrology and Atmospheric Sciences, The University of Arizona, John W. Harshbarger Building, 1133 E. North Campus Drive, Tucson AZ 85721-0011, USA

⁷Department of Geography, Department of Earth Science, The University of Hong Kong, Pokfulam, Hong Kong SAR 999077, China

⁸MNR Key Laboratory for Geo-Environmental Monitoring of Great Bay Area & Guangdong Key Laboratory of Urban Informatics & Shenzhen Key Laboratory of Spatial Smart Sensing and Services, Shenzhen University, Shenzhen 518060, China

⁹School of Architecture & Urban Planning, Shenzhen University, Shenzhen 518060, China

¹⁰Department of Earth System Science, Institute for Global Change Studies, Tsinghua University, Beijing 100084, China

¹¹These authors contributed equally

*Correspondence: hhxunan@gmail.com (N.X.); haoyh@sxu.edu.cn (Y.H.); luhui@tsinghua.edu.cn (H.L.)

Received: February 1, 2026; Accepted: July 18, 2026; Published Online: July 23, 2026; <https://doi.org/10.59717/j.tiw.2026.100010>

© 2026 The Author(s). This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

Citation: Yao J., Hao H., Wang M., et al. (2026). Expanding groundwater drawdown funnels in North China Plain: A new insight from satellite gravimetry. *The Innovation Water* 1:100010.

Quantifying the evolution and recovery of groundwater drawdown funnels in the North China Plain (NCP) remains a grand challenge due to limited spatial resolution, vertical ambiguity, and attribution complexity in current satellite-based monitoring. Here, we present an integrative framework that addresses these bottlenecks by combining data-driven downscaling of GRACE gravimetric data, stratified signal separation using multi-depth well observations, and causality analysis driven by climate and groundwater abstraction indicators. This framework enables ~5 km resolution groundwater storage reconstruction, delineates 3D funnel morphology via Gaussian fitting, and attributes dynamic changes through wavelet-informed causality inference. Our key findings include: (1) The NCP groundwater system experienced a three-stage trajectory—initial slow decline (−0.05 cm/year), accelerated depletion (−0.19 cm/year), and partial recovery (+0.16 cm/year)—with funnel radius expanding from 209 km in 2002 to 340 km in 2022. (2) Funnel morphology transformed substantially, with the lowest groundwater level deepening from 20.93 m to 52.95 m and lateral half-width expanding from 0.56 m to 4.01 m over two decades. (3) The combined analysis of the transfer function noise (TFN) model and causality analysis based on information flow further revealed that early funnel changes were dominated by climate (~75% contribution), later shifts by intensive pumping (~90%), and recent recovery partially driven by precipitation (~16%), with aquifer response time increasing from 20 to 60 months, indicating a weakened response and delayed recovery. The lagged responses revealed by the causality analysis redefines our understanding of anthropogenic impacts on aquifer systems and offers a transferable path toward sustainable groundwater governance under changing climate and socioeconomic pressures.

INTRODUCTION

Groundwater systems are increasingly concerned worldwide, with depletion accelerating in over 30% of major aquifers due to prolonged overuse and climate variability. Recent assessments show groundwater levels are declining in most of the world's large aquifers, with rapid drops (> 0.5 m/year) now widespread in the 21st century.¹ Alarmingly, long-term records from hundreds of thousands of wells indicate that depletion rates have accelerated over the past few decades in about 30% of global aquifers.² Such over-exploitation has serious consequences – including land subsidence, drying of rivers and wetlands, and even measurable contributions to saline water intrusion – underscoring the urgent need for improved groundwater monitoring and management.^{3,4}

The North China Plain (NCP), a key grain-producing region, has experienced decades of intensive abstraction from both unconfined and confined

aquifers,⁵ resulting in one of the world's largest groundwater drawdown funnels (GDF). This has caused significant local groundwater declines and issues such as soil salinization, land subsidence, and wetland degradation.⁶ While recent interventions—such as managed aquifer recharge and inter-basin water transfers—have led to partial reversals of these declines, it remains unclear how these changes vary spatially, vertically, and causally across the NCP.^{1,7} Traditional well measurements and water balance models provide aggregate estimates that obscure local aquifer heterogeneity, limiting effective hotspot management.⁸ Moreover, although in-situ groundwater observations provide accurate local water-level information, monitoring wells are often spatially sparse and unevenly distributed, with inconsistencies in aquifer type and observation depth. These limitations make it difficult to continuously characterize the spatiotemporal evolution of the horizontal extent and vertical structure of groundwater drawdown funnels across the NCP. In contrast, satellite gravimetry provides spatially continuous and temporally consistent observations at regional scales, offering an important complement to ground-based monitoring for tracking long-term groundwater funnel evolution.

Satellite gravimetry has transformed global groundwater monitoring. The GRACE mission (2002–2017) and GRACE Follow-On (since 2018) enabled tracking of terrestrial water storage (TWS) via gravity field variations.^{9,10} Because GRACE measures total terrestrial water storage, separating groundwater requires removing signals from soil moisture, surface water, and snow, introducing uncertainty.¹¹ To improve reliability, researchers combine GRACE with ground observations and models.^{7,9} For example, assimilating data into land surface models or calibrating with well measurements.¹² GRACE-derived groundwater trends align well with in-situ well data, confirming its utility for large-scale spatiotemporal groundwater analysis.^{13,14}

Despite extensive research, satellite-based assessments—particularly those using GRACE,⁹ but remain constrained by coarse spatial resolution (3° or about 300 km), vertical signal aggregation, and limited attribution capacity.^{15,16} These limitations restrict their utility for local-scale management and for understanding stratified aquifer dynamics or driver-specific recoverability. This mismatch between the scale of GRACE observations and the local-scale management, observations, models, and our interests has motivated a push to downscale GRACE-based groundwater estimates to higher spatial (and temporal) resolutions.^{17,18} Various downscaling approaches for GRACE terrestrial water storage have been developed, broadly categorized into physical, statistical, and machine-learning methods.¹⁹ Physical model-based downscaling assimilates GRACE data into hydrological or land surface models to disaggregate water storage changes at finer scales.¹⁹ This approach enforces mass balance and leverages hydrogeological prior knowledge but remains limited by uncertainties in model parameters, forcing

functions, and scale mismatches between GRACE footprints and model grids.

Statistical downscaling builds empirical links between coarse GRACE signals and higher-resolution hydrologic variables (e.g., precipitation, evapotranspiration, groundwater proxies).²⁰ These data-driven methods, like linear regression, are simple and efficient but assume stationary relationships and lack physical constraints, limiting robustness.²¹ Machine learning approaches, including random forests and neural networks, have been applied to capture nonlinear dependencies and improve regional groundwater estimates using climate and remote sensing data.^{22,23} Deep learning models further enhance downscaling by learning complex spatiotemporal patterns but require large training datasets and careful validation to ensure physical consistency and avoid black box issues.²⁴

Another critical challenge is attributing observed groundwater trends to their drivers. Groundwater storage changes are driven by climate variability (natural recharge and discharge) and human activities (pumping, irrigation, artificial recharge), but disentangling these influences is complex. Traditional methods like correlation and water-balance models separate climate and anthropogenic impacts but lack rigor. Recently, advanced causal inference frameworks²⁵ have been applied to GRACE and hydroclimatic data to more precisely quantify groundwater changes caused by precipitation versus exploitation.^{6,26} In the NCP, independent component decomposition showed that climate largely drives short-term groundwater fluctuations.

In contrast, intensive groundwater abstraction drove the long-term decline.²⁷ Understanding these drivers is vital for assessing groundwater recoverability—the potential for aquifers to stabilize or rebound. In California's Central Valley, natural recharge during wet periods and pumping limits enabled partial recovery.²⁸ However, chronic overexploitation in regions like the NCP requires deliberate intervention and new analytical methods to evaluate and promote recoverability.

Against this backdrop, the NCP offers a unique lens to rethink groundwater analysis's technological and conceptual frameworks. This study integrates GRACE satellite gravimetry, blind super-resolution downscaling, well-informed vertical decomposition, and causal inference to investigate the evolution and recoverability of groundwater drawdown funnels in the NCP from 2002 to 2022. Specifically, this study aims to: (1) reconstruct high-resolution groundwater-level fields from coarse GRACE observations to capture the spatiotemporal evolution and lateral expansion of groundwater drawdown funnels, (2) quantify the three-dimensional morphology and volumetric evolution of the funnel, and (3) reveal the dynamic response and driving mechanisms of groundwater-funnel evolution by integrating TFN response functions with causal information flow analysis.

MATERIALS AND METHODS

Experimental dataset

The North China Plain (NCP), spanning ~130,000 km² in eastern China (112°30'E–119°30'E, 34°46'N–40°25'N; *Figure 1*), includes Beijing, Tianjin, Hebei, and parts of Henan and Shandong Provinces. As a major agricultural base, it relies heavily on groundwater, which supplies 69.4% of total water use.²⁹ Decades of overexploitation have caused severe groundwater storage (GWS) depletion and widespread land subsidence,^{30,31} forming the world's largest compound depression cone.³²

The NCP has a continental semi-arid climate, with ~80% of precipitation potentially recharging groundwater. Mean annual temperature is 12°C–13°C; pan evaporation ranges from 1,100 to 2,000 mm. Actual evapotranspiration is higher in the south, while both precipitation and evapotranspiration have declined in recent decades.³³ The groundwater system comprises unconsolidated porous aquifers, including Quaternary, Paleogene–Neogene, karst, and fractured bedrock aquifers.³⁴ In the last decade, many water diversion projects and other countermeasures have been conducted to alleviate the overexploitation of groundwater.³⁵

Blind super-resolution framework

To overcome the spatial resolution limitation of GRACE satellite gravimetry in groundwater monitoring, we developed a data-distribution-aware blind super-resolution method that enhances TWS from 1° to 0.0625° (~5 km) without requiring paired high-resolution training data. Based on unified preprocessing of GRACE and Global Land Data Assimilation System (GLDAS)

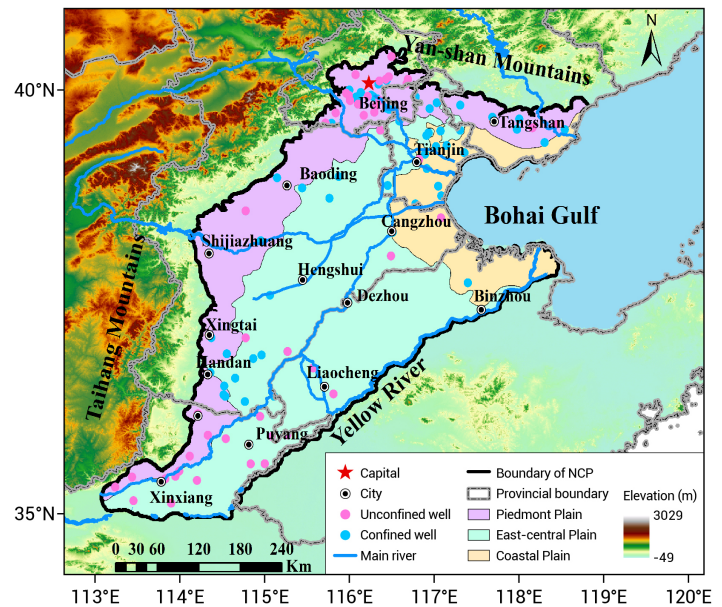


Figure 1. Overview of research area.

into physically consistent anomaly grids, the model reconstructs multisource water components via hierarchical learning that preserves spatial structure and numerical consistency.^{10,36} Structural similarity and peak signal-to-noise ratio were incorporated into the loss function to constrain spatial texture and accuracy jointly to ensure interpretability.^{16,22} This approach significantly enhances the applicability of satellite gravimetry in high-resolution groundwater assessments and supports advanced spatiotemporal and causal analysis of groundwater recoverability.

Transfer function noise setup

This study applies the Transfer Function Noise (TFN) method to quantify how natural recharge and human abstraction affect groundwater levels in the NCP from 2002 to 2022.³⁷ Using Gamma-type impulse response functions, the model captures delayed and immediate responses to net recharge (precipitation minus evaporation) and pumping.^{37,38} Parameter inversion aligns simulated and observed data, revealing the temporal evolution of the groundwater depression cone. Discussion based on the lagged recharge characteristics at different stages further highlights the cone's formation mechanisms and the relative contributions of natural and anthropogenic drivers.

Causal inference setup

To assess lagged causal effects of key drivers on spring discharge, this study integrates the lag structure derived from the TFN model with the information flow theory proposed by Liang.^{25,26} This theory, rooted in stochastic dynamical systems, quantifies the causal strength between variables via sample covariances, including those involving time derivatives.³⁹ Precipitation, potential evapotranspiration, and groundwater use are introduced as external drivers. The method identifies dominant causal influences and reveals the groundwater system's dynamic response by comparing information flow magnitudes at different lag steps.

Overall technical setup

To capture the spatial and temporal evolution of GDFs in the NCP, we establish an integrated technical framework combining satellite gravimetry, blind super-resolution downscaling, and causal inference (*Figure 2*). First, GRACE and GLDAS anomalies are reprojected and aligned into consistent spatiotemporal grids. A blind downscaling approach enhances the resolution of GRACE-derived terrestrial water storage (TWS) to ~5 km, after which surface components are removed to estimate groundwater storage (GWS). Using 130 in-situ groundwater well observations, physical regression converts GWS anomalies into groundwater level (GWL) changes. Next, long-term GWL data are used to identify and characterize groundwater drawdown

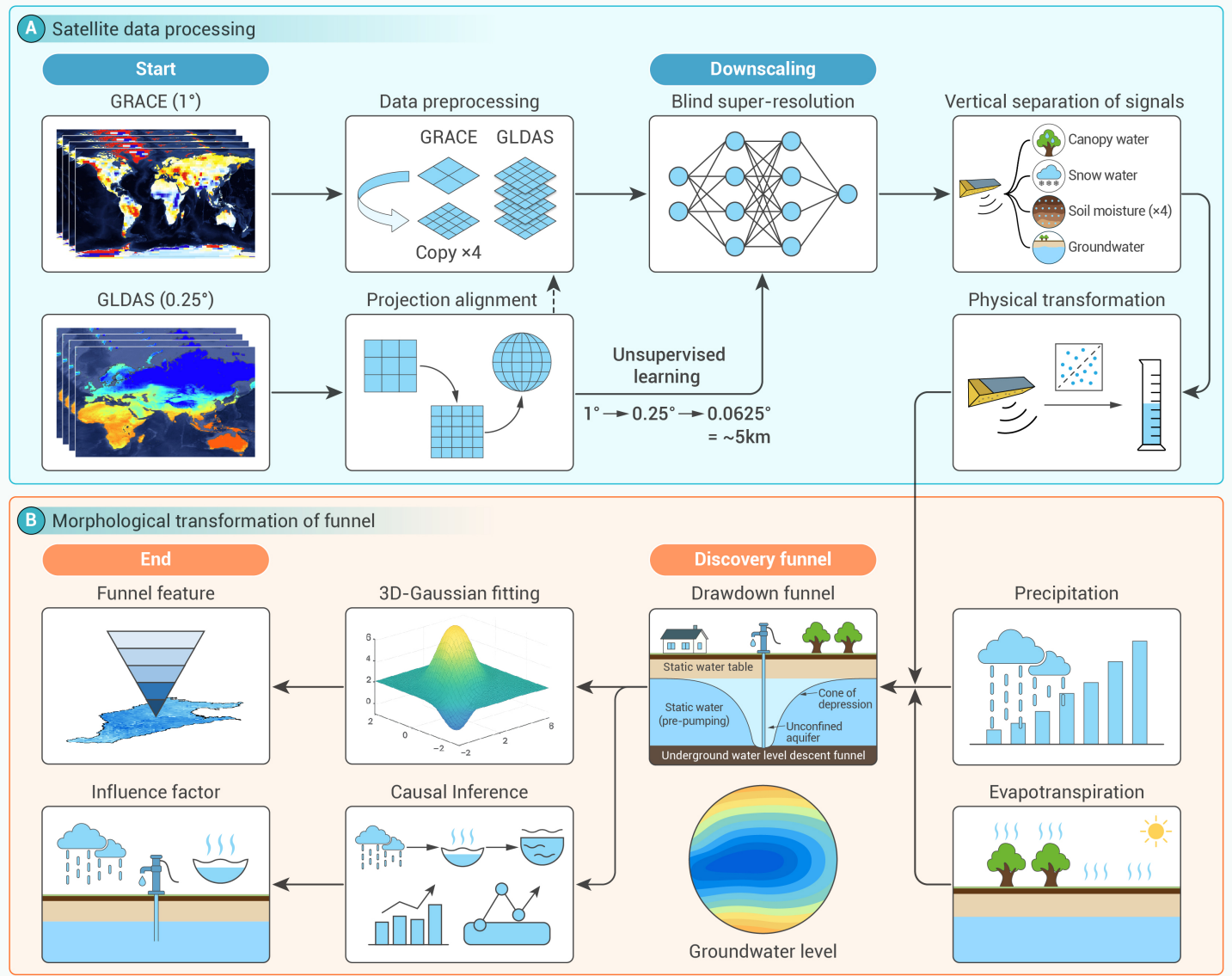


Figure 2. Integrated technical framework for groundwater funnel detection and attribution in the North China Plain (A) Satellite-based data processing. GRACE and GLDAS products are preprocessed and downscaled to ~5 km resolution. Groundwater storage (GWS) is isolated and converted into groundwater level (GWL) using 130 in situ well observations. (B) Morphological tracking of GDFs. High-resolution GWS time series are used to detect and quantify drawdown funnels via 3D Gaussian fitting, followed by causal attribution using precipitation (Pre) and potential evapotranspiration (PET) time series.

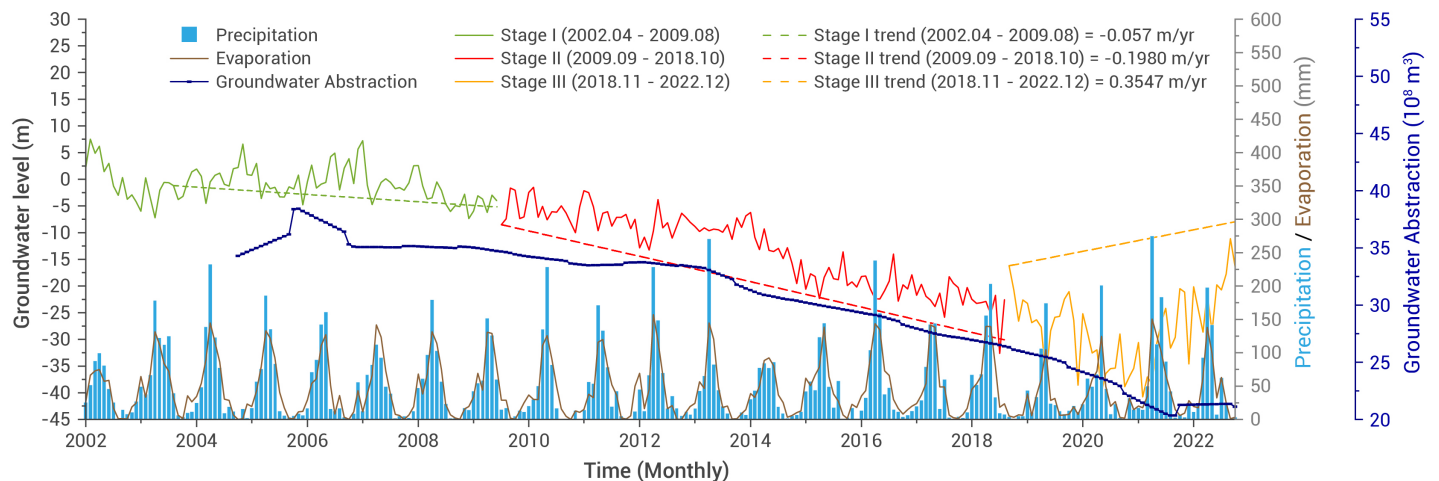


Figure 3. Time series of groundwater levels, precipitation, and evaporation in the North China Plain.

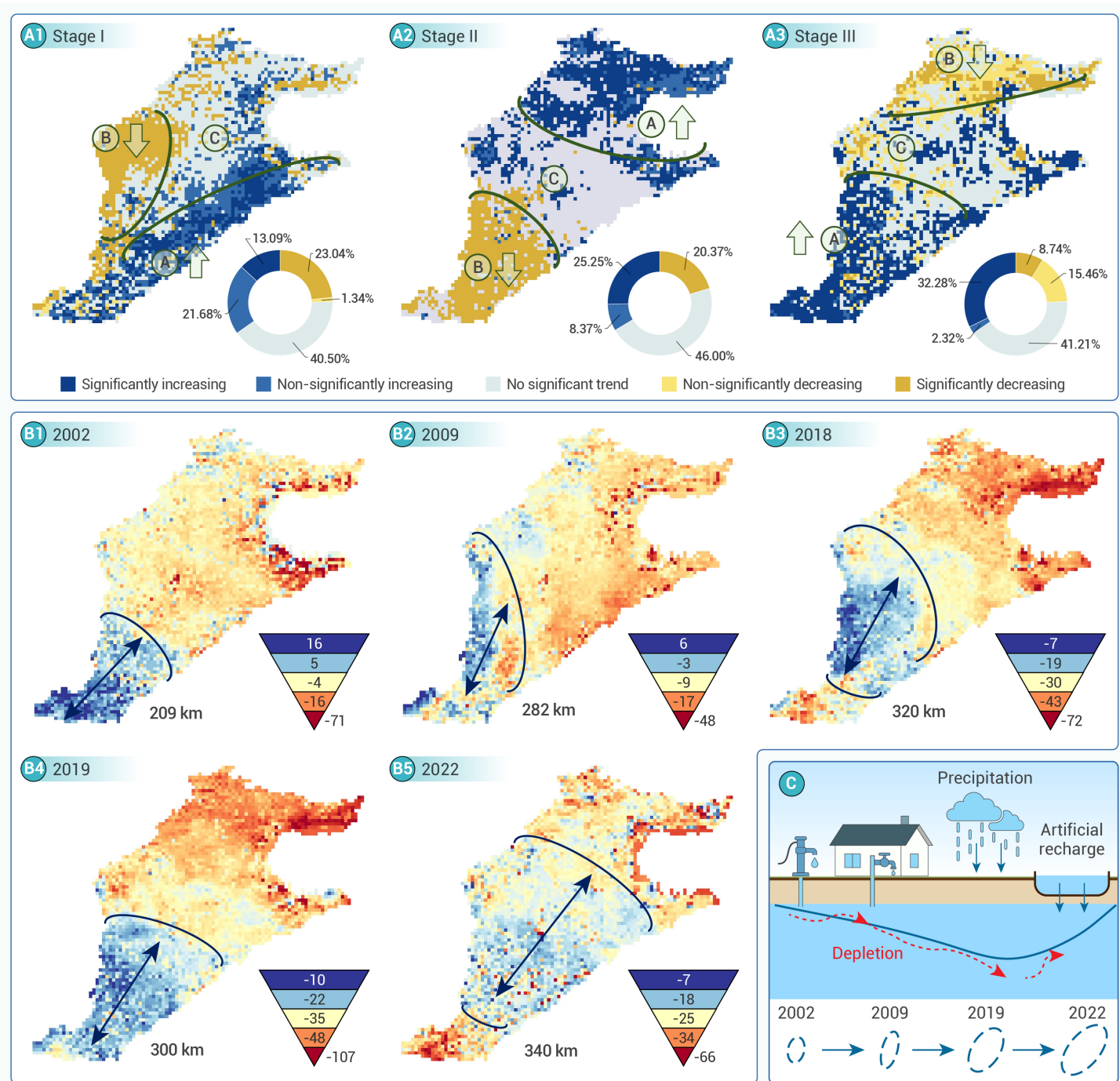


Figure 4. Drawdown funnel distribution of groundwater in the North China Plain (A1–3) Spatial variation trend of groundwater in Stage I–III by Mann-Kendall. A represents an upward trend, B represents a downward trend, and C represents no trend. (B1–5) Long-time series change of groundwater level from Satellite Gravimetry. (C) Concept map of groundwater drawdown funnel change. The funnel in the North China Plain has been expanding significantly and continuously.

funnels using 3D Gaussian fitting. Finally, the driving mechanisms are analyzed via causal inference techniques based on information flow and wavelet coherence, quantifying the delayed impacts of precipitation, evapotranspiration, and anthropogenic extraction.

RESULTS

Spatiotemporal heterogeneity of groundwater

Figure 3 shows time series of groundwater levels, precipitation, evaporation, and pumping volumes in the NCP (2002–2022). Groundwater levels have declined persistently, with expanding spatial impact, indicating that extraction surpassed recharge and led to a pronounced funnel effect (Figure 4). Trend analysis revealed three distinct stages: Stage I (2002–2009) with slow, stable decline (-0.0570 cm/year); Stage II (2009–2019) with rapid,

nonstationary decline (-0.1980 cm/year); and Stage III (2019–2022), showing nonstationary recovery at 0.1682 cm/year. Although the absolute minimum groundwater level occurred around 2021, the groundwater system exhibited a clear shift in dynamic behavior after 2019, including a reduced decline rate and an emerging recovery trend associated with strengthened groundwater management interventions. These stage transitions were mainly shaped by the changing balance between groundwater abstraction and management interventions.

Figures 4B1–B5 show groundwater level distributions from downscaled GRACE data corrected by well observations at five key time points (2002, 2009, 2018, 2019, 2022). The GDF expanded steadily, with its radius growing from 209 km in 2002 to 340 km in 2022, except for a temporary contraction to 300 km in 2019. These trends indicate ongoing drawdown but a recent

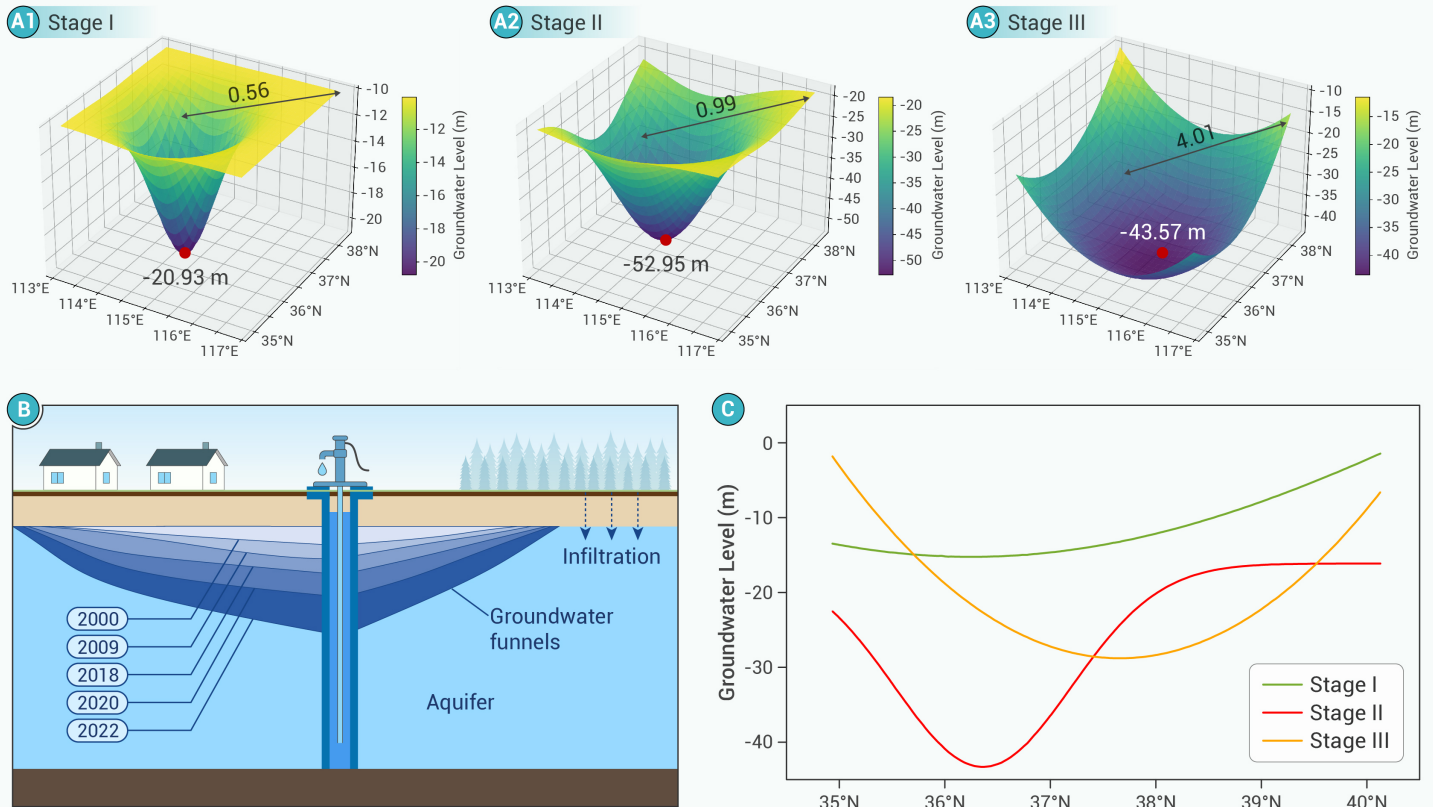


Figure 5. Temporal and spatial changes of morphological characteristics of groundwater drawdown funnel in the North China Plain (A1–3) Spatial variation trend of groundwater funnel 3D-Gaussian surfaces fitted in 2002–2022. (B) Concept map of vertical profile change of groundwater drawdown funnels. (C) A time series vertical profile is fitted to the groundwater drawdown funnels.

slowdown in decline, reflecting early impacts of groundwater management and recharge efforts.

Morphology of the groundwater funnel

Identifying GDFs is key to assessing aquifer depletion's extent and severity. GDFs form in low-elevation zones due to sustained pumping, with their shape indicating hydrological stress and system resilience. This study used 3D-Gaussian surface fitting on downscaled GRACE-derived groundwater fields to characterize GDF morphology—depth, extent, and curvature. The fitted surfaces provide intuitive, continuous representations, capturing both the minimum groundwater level (depth) and horizontal spread (sigma x).

The morphology of GDFs in the NCP evolved significantly over the study period (Figures 5A1–A4). The lowest groundwater level declined from 20.93 m in 2002 to 52.95 m in 2019, then slightly rebounded to 45.57 m in 2022. Meanwhile, the Gaussian-derived half-width coefficient, defined as the half-width at half maximum of the fitted groundwater depression, increased from 0.56 to 4.01, indicating major lateral expansion of the funnel. These trends are consistent with the three-stage dynamics described at the beginning of the Results section. Figure 5B depicts vertical changes, and Figure 5C traces the shifting GDF axis. The results show that the NCP faces an increasingly severe groundwater funnel crisis. During Stage II and Stage III, the rate of deepening slowed, the GDF's radius continued to expand, and the center of the GDF shifted northward, reflecting the spatial redistribution of pressure heads possibly due to managed aquifer recharge and the South-to-North Water Diversion Project. This spatial migration suggests that policy interventions not only reduced decline rates but also altered the locus of groundwater stress. However, these interventions have not fundamentally reversed the overall declining trend.

These observed morphological shifts mark a transition from concentrated, rapidly declining GDFs to more fragmented, stabilizing ones, echoing early signs of aquifer recoverability. We applied causal information flow analysis in the following section to unravel the driving forces behind these transitions further. This enables a quantitative decomposition of anthropogenic and climatic influences on GDF dynamics, providing mechanistic insight into why

such morphological changes have occurred and under what conditions they may continue.

Simulation of groundwater funnel

To explore temporal changes in groundwater dynamics, the study period was divided into three phases—Stage I (2002.04–2009.08), Stage II (2009.09–2018.10), and Stage III (2018.11–2022.12) - based on trend shifts identified by the Mann-Kendall and Sen's slope methods.¹² Subsequently, the TFN model was applied to each stage. The model was calibrated using monthly precipitation and evaporation data, supplemented by gravity satellite data, to simulate groundwater level dynamics via a nonlinear recharge model.³⁸ To reduce uncertainty in the results, the model structure incorporates a noise response function that more effectively captures the system's random perturbations. Despite slight mismatches, the model reproduced key spatial patterns of groundwater change (Figure 5A), achieving an R^2 of 0.84. Simulated trends closely followed observations.

Figures 6B–C show the step and block response functions (i.e., the system's response to an input that suddenly jumps from zero to a constant value), revealing the groundwater system's memory and responsiveness over three stages. Over time, the system's response duration progressively lengthens, indicating a continuous decline in the aquifer's regulatory and response capacity. The final amplitude of the step response also varies across stages. In stage I, the system stabilizes within approximately 20 months, with a maximum amplitude of about 0.5 m/mm, reflecting a rapid but limited cumulative recharge—suggesting that the aquifer was responsive to precipitation but moderately saturated. In contrast, during II, the response duration extends significantly to about 60 months, and the final amplitude rises to around 0.9 m/mm. This increase may be attributed to a deeper groundwater funnel and enhanced vertical hydraulic gradients, which allowed prolonged recharge and increased storage space, even as the overall system condition deteriorated due to overexploitation. In stage III, groundwater levels showed signs of stabilization, with a slightly shortened response duration of about 40 months and a reduced amplitude of approximately 0.65 m/mm, suggesting partial recovery—possibly due to policy measures like the South-to-North

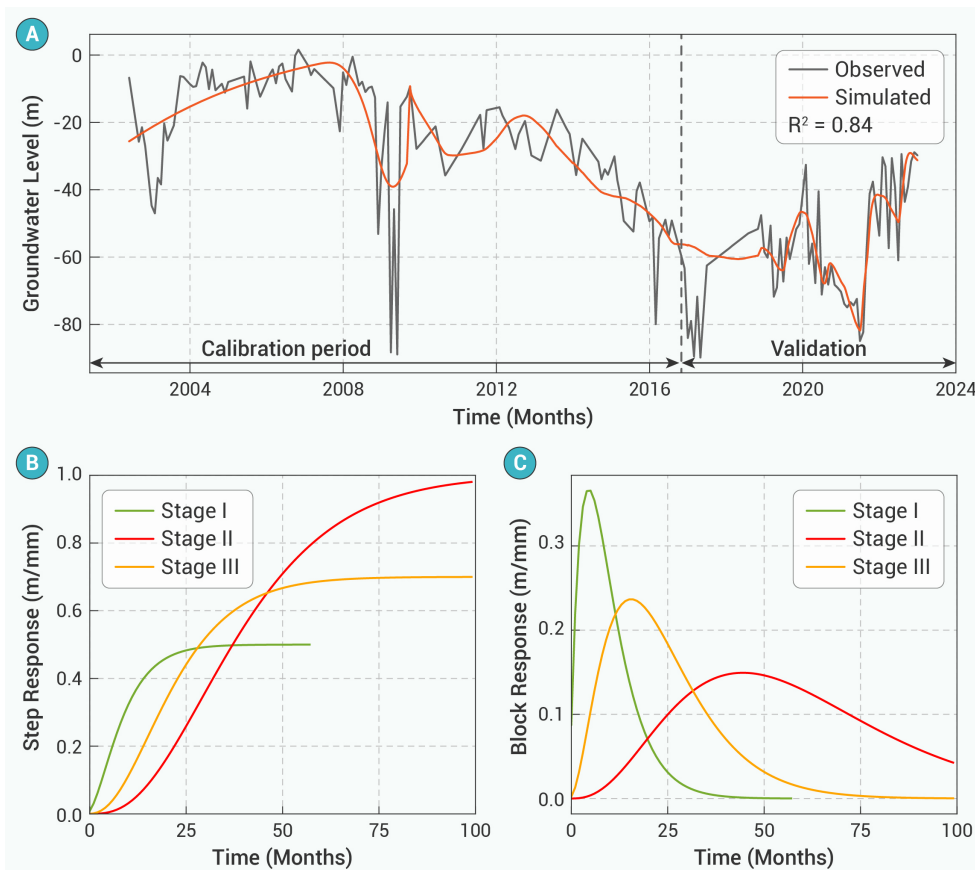


Figure 6. TFN-Based simulation and dynamic response analysis of the groundwater funnels (A) Observed and simulated average groundwater levels in the cone area. (B) The step response functions show the response of the groundwater level to a sudden unit increase in recharge that extends infinitely in time. (C) The block response shows how the groundwater levels react to a 1-month recharge event of 1 mm.

results show extraction at 78.0%, with precipitation and evaporation dropping to 12.2% and 9.8%, respectively.

In stage III (2018.10–2022.12), groundwater extraction once again exhibited a significantly enhanced causal influence, with 28 significant lags identified within the 1–40-month range and a peak causal value reaching 0.63, indicating a more pronounced dominance. Precipitation and evaporation regained partial significance at 30–40 months, reflecting delayed recovery of natural influences due to policies like the South-to-North Water Diversion Project. CSI results indicate extraction at 59.2%, while precipitation and evaporation contributions rose to 23.3% and 17.5%, respectively, suggesting a gradual move toward a more balanced groundwater system.

Overall, groundwater extraction has been the key driving factor in forming and maintaining the groundwater funnel across all three stages, reaching its peak dominance during the second

stage. Water Diversion Project and extraction restrictions. Pulse response analysis (Figure 6C) further confirms a stepwise decline in magnitude alongside increased duration, quantitatively tracking the evolution from high responsiveness (Stage I) to low responsiveness (Stage II) and finally to a fragile recovery phase (Stage III), thereby highlighting the phased impacts of human activities and policy interventions.

Causal Inference of groundwater funnel

To quantify the causal influences driving groundwater funnel formation and their temporal evolution across different stages, this study employs the dynamic characteristics of the impulse response function (IRF, the system's response to a brief input) and step response function (SRF, the system's response to a sudden sustained input). Maximum information flow lags are set at 20, 60, and 40 months for the three stages. A causal strength index (CSI) evaluates the relative contributions (Table S1) of precipitation (Figure 7A), evaporation (Figure 7B), and groundwater extraction (Figure 7C). The CSI is calculated as the ratio of the causal statistic (T-value), which measures the significance of the causal effect, to the confidence interval threshold, and is normalized within each stage to express the relative driving strength of the three factors in percentage terms.

In Stage I (2004.12–2008.12), average precipitation and groundwater extraction had rapid, significant causal effects. Precipitation showed 10 significant lags (0–12 months) peaking at month 3, while extraction showed 11 significant lags (0–10 months), driving rapid funnel formation. Evaporation had weak significance (months 2–5). CSI results attributed 45.6% to extraction, 34.6% to precipitation, and 19.8% to evaporation, highlighting extraction as the dominant disturbance alongside natural factors.

In Stage II (2009.9–2018.10), the causal significance of all factors declined sharply. Precipitation and evaporation exerted only weak effects, with information flow peaking at lags of 30–40 months, indicating a slow and delayed groundwater response. The IRF peaked near 40 months, and the SRF stabilized around 60 months, reflecting reduced system sensitivity and regulation. Groundwater abstraction's causal influence persisted but lacked significance, suggesting dominance of anthropogenic impacts and a dulled response. CSI

stage. In contrast, the impacts of precipitation and evaporation exhibit clear stage-dependent variations: strong in the first stage, diminished in the second, and gradually recovering in the third alongside policy interventions. These results indicate a quantitative shift in groundwater system behavior under anthropogenic disturbance, from rapid response and strong recovery to delayed response and weak recovery.

DISCUSSION

Although GRACE enables continuous monthly monitoring of regional groundwater changes, its coarse spatial resolution limits its ability to characterize the spatial heterogeneity of groundwater storage variations across the NCP. For instance, the standard 1° resolution mascon solution from NASA (Figure 8A1) covers only 24 grid cells over the NCP, which is insufficient to reflect the region's internal hydrological complexity. This limitation is particularly important in regions where groundwater abstraction intensity, aquifer structure, and management interventions vary substantially in space. Coarse-resolution GRACE products may smooth local groundwater depletion signals, thereby affecting the identification of groundwater drawdown funnels and the evaluation of groundwater management effects. Although some machine learning-based "pseudo-downscaled" products at 0.25° resolution (Figure 8A2) increase the grid count to 504, their strong reliance on external reference data often results in artifacts such as uniform pixel values across adjacent cells, lacking genuine spatial variability.

In this study, we propose a downscaling approach that does not rely on high-resolution training data. This method increases the spatial resolution to 0.0625° (approximately 5 km, Figure 8A3), improving the ability to identify localized depletion centers and lateral expansion patterns and providing a more suitable basis for diagnosing aquifer stress in heavily exploited plains. More importantly, unlike traditional statistical or interpolation-based downscaling methods, this approach achieves a physically meaningful transformation from TWS to groundwater level, overcoming a longstanding barrier in hydrological modeling related to the unobservability of subsurface processes.

To validate the accuracy of the downscaled products, we incorporated observational data from 130 monitoring wells across the NCP (Figure 8B),

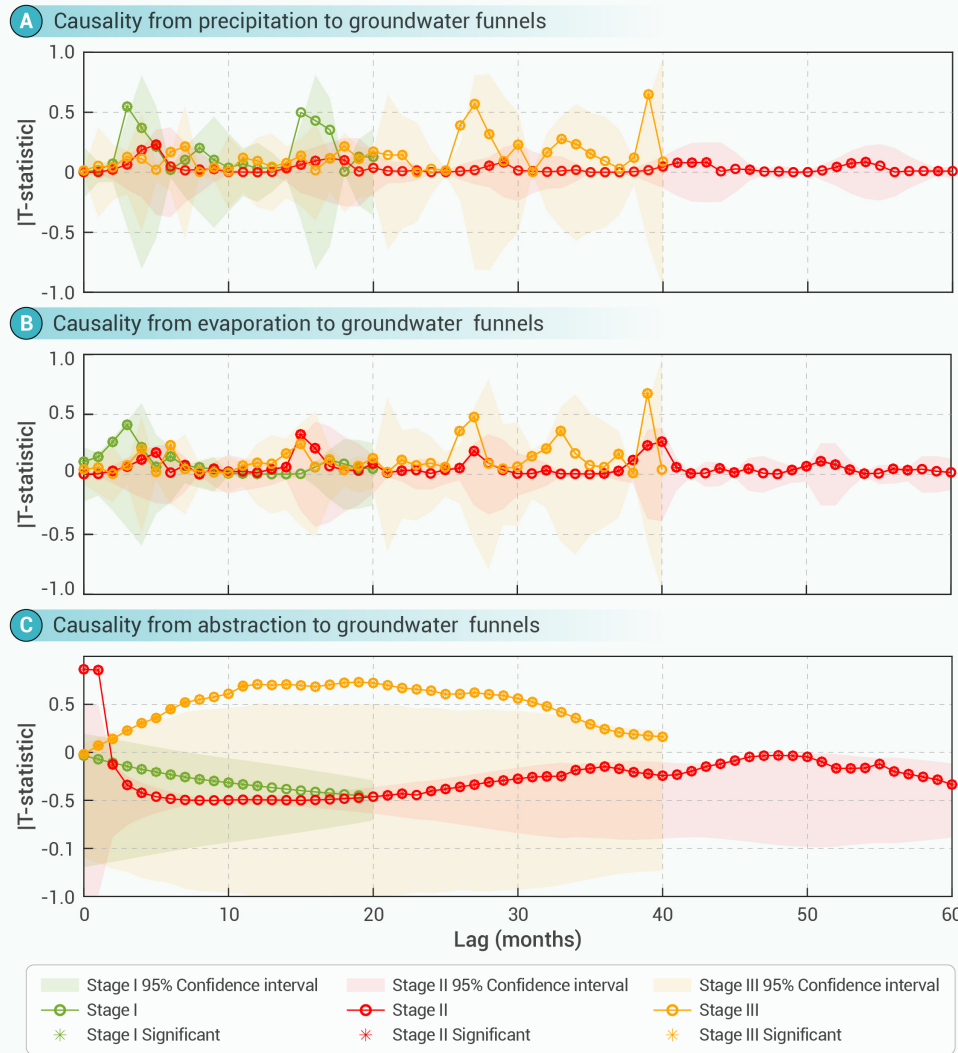


Figure 7. Lagged causal information flow of the groundwater funnel in the North China Plain (A) From precipitation to the groundwater funnel. (B) From evaporation to the groundwater funnel. (C) From groundwater abstraction to the groundwater funnel.

representative case. This study develops an integrative framework combining blind super-resolution downscaling, vertical signal decomposition, and causal inference to analyze groundwater drawdown funnel (GDF) dynamics in the NCP from 2002 to 2022, enabling three key advances:

(1) The proposed downscaling framework transformed coarse-resolution GRACE signals into 0.0625° groundwater-level fields, providing a reliable spatial basis for identifying localized depletion centers and lateral expansion patterns of the groundwater drawdown funnel. Validation against 130 monitoring wells showed an overall accuracy of approximately 89.5%. The groundwater dynamics across the NCP from 2002 to 2022 demonstrated a distinct temporal evolution: an initial gentle decline (-0.0570 cm/year) transitioning to rapid depletion (-0.1980 cm/year), and subsequently shifting towards partial recovery ($+0.1682$ cm/year) in recent years. Concurrently, the GDF markedly expanded spatially, with its radius increasing progressively from 209 km in 2002 to 340 km by 2022, signifying an ongoing regional groundwater depletion. However, the recent reduction in depletion rate and stabilization of groundwater levels indicate positive early responses to implemented management strategies such as aquifer recharge and inter-basin water transfer projects.

(2) The NCP's GDF underwent significant

including 83 confined and 47 unconfined wells, covering the region's major hydrogeological units. Comparisons with well measurements indicate a strong agreement, with an overall accuracy of approximately 89.5%, representing a 40% improvement over the original 1° GRACE product. The error distribution (Figure 8D) further reveals that for confined aquifers, ~83.0% of the errors fall within the -10 m to 0 m range, while moderate errors (-20 m to -10 m) account for ~16.4%. Extreme errors below -20 m are rare ($< 0.6\%$). Unconfined wells show a similar but slightly more dispersed pattern, reflecting their greater sensitivity to short-term disturbances such as precipitation and pumping.

In addition, Figure 8C further highlights the differences in groundwater level change patterns between confined and unconfined aquifers. Figure 8C1 shows a denser concentration of water level decline values for confined wells, indicating a more centralized and uniform response to long-term stress. In contrast, Figure 8C2 demonstrates a broader distribution with extended tails for unconfined wells, reflecting their greater sensitivity to short-term meteorological fluctuations. This vertical difference underscores the stratified structure of the aquifer system and the need for differentiated recovery strategies. A more detailed analysis of the three-dimensional morphology and volumetric evolution of groundwater drawdown funnels is provided in Text S1 and Figure S1.

CONCLUSION

Despite advances in satellite-based groundwater assessments, GRACE studies are limited by coarse resolution, vertical signal mixing, and weak attribution, hindering local management and causal insights. The North China Plain (NCP), a key site of long-term aquifer depletion and recovery, offers a

morphological transformations from 2002 to 2022, quantitatively evidenced by 3D-Gaussian fitting analyses. The lowest groundwater elevation deepened from 20.93 m in 2002 to a maximum of 52.95 m in 2019, followed by partial recovery to 45.57 m by 2022. These morphological changes reveal a paradox of fragile recovery: although the funnel center has begun to rebound after 2019, the continued lateral expansion indicates that the accumulated regional groundwater deficit has not yet been fully released. Concurrently, the funnel's lateral half-width notably increased from an initial 0.56 in 2002 to 4.01 in 2022, reflecting extensive horizontal broadening and intensified aquifer stress. These morphological shifts correspond closely with the identified three-stage groundwater dynamics, indicating that intensive groundwater extraction substantially reshaped aquifer structure, while recent management interventions have begun to reverse these trends. Thus, precise monitoring of morphological parameters like depth and width effectively captures groundwater system health and informs resource management strategies.

(3) Causal inference analysis revealed distinct drivers underpinning the observed morphological changes of the GDF in the NCP. Early-stage (2002–2009) groundwater depletion was predominantly influenced by climatic variability, with precipitation driving ~75% of short-term groundwater fluctuations. However, intensive groundwater abstraction became dominant from 2009 onward, contributing to rapid funnel expansion and groundwater level decline (~90% anthropogenic control). From 2019, the influence of climate reemerged, explaining approximately 16% of groundwater recovery, indicating a partial response to improved recharge conditions and reduced extraction due to policy interventions. Furthermore, the system's response duration to recharge increased significantly from about 20 to 60 months, reflecting diminished aquifer buffering and greater dependency on prolonged

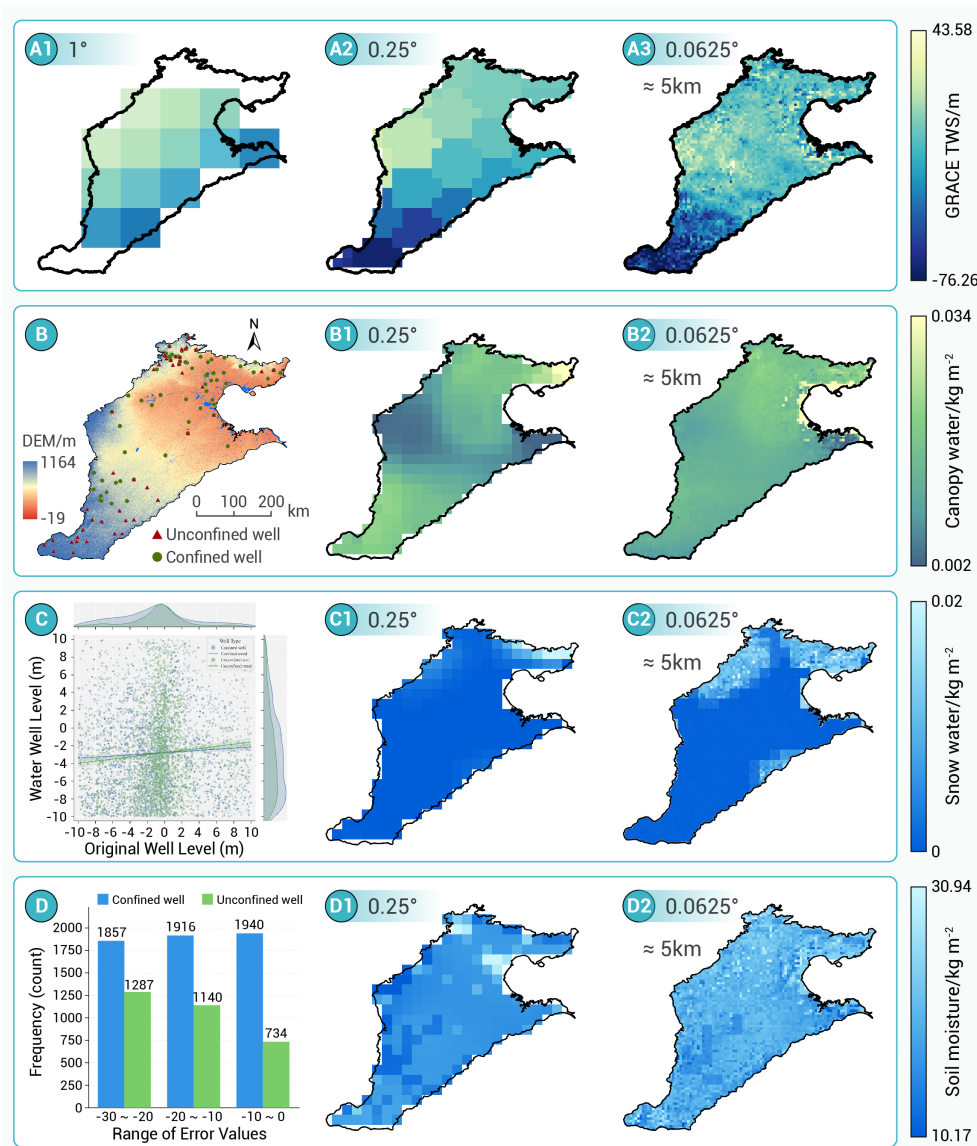


Figure 8. Accuracy verification of downscaled GRACE terrestrial water storage products A1–3. TWS products from GRACE satellite at different spatial resolutions. b. Distribution of groundwater monitoring wells in northern China. B1–2. Canopy water products from GLDAS at three different spatial resolutions. c. Comparison of water level change density between confined and unconfined aquifers. C1–2. Snow water products from GLDAS at different spatial resolutions. d. Error statistics between GRACE satellite data and groundwater monitoring well data. D1–2. soil water products from GLDAS at different spatial resolutions.

large depletion in water storage in response to the 2011 drought in Texas. *Geophys. Res. Lett.* **40**:3395–3401. DOI:10.1002/grl.50655

9. Rodell M., Famiglietti J. S., Wiese D. N., et al. (2018). Emerging trends in global freshwater availability. *Nature* **557**:651–659. DOI:10.1038/s41586-018-0123-1
10. Gou J. and Soja B. (2024). Global high-resolution total water storage anomalies from self-supervised data assimilation using deep learning algorithms. *Nat. Water* **2**:139–150. DOI:10.1038/s44221-024-00194-w
11. Chen Y., Wang Y., Li L., et al. (2024). Monthly monitoring of inundated areas and water storage dynamics in China's large reservoirs using multisource remote sensing. *Water Resour. Res.* **60**:e2023WR036450. DOI:10.1029/2023WR036450
12. Zhao J., Li G., Zhu Z., et al. (2024). Analysis of the spatiotemporal variation of groundwater storage in Ordos Basin based on GRACE gravity satellite data. *J. Hydrol.* **632**:130931. DOI:10.1016/j.jhydrol.2024.130931
13. Rateb A., Scanlon B. R., Pool D. R., et al. (2020). Comparison of groundwater storage changes from GRACE satellites with monitoring and modeling of major US aquifers. *Water Resour. Res.* **56**:e2020WR027556. DOI:10.1029/2020WR027556
14. Bhanja S. N., Mukherjee A., Saha D., et al. (2016). Validation of GRACE based groundwater storage anomaly using in-situ groundwater level measurements in India. *J. Hydrol.* **543**:729–738. DOI:10.1016/j.jhydrol.2016.10.042
15. Feng W., Zhong M., Lemoine J. M., et al. (2013). Evaluation of groundwater depletion in North China

recharge periods for stabilization. These quantitative insights highlight critical transitions in groundwater dynamics, emphasizing the efficacy of management strategies alongside climatic influences.

REFERENCES

1. Jasechko S., Seybold H., Perrone D., et al. (2024). Rapid groundwater decline and some cases of recovery in aquifers globally. *Nature* **625**:715–721. DOI:10.1038/s41586-023-06879-8
2. Liu P.-W., Famiglietti J. S., Purdy A. J., et al. (2022). Groundwater depletion in California's Central Valley accelerates during megadrought. *Nat. Commun.* **13**:7825. DOI:10.1038/s41467-022-35582-x
3. Cuthbert M., Gleeson T., Bierkens M., et al. (2024). Concerns regarding proposed groundwater Earth system boundary. *Nature* **635**:E4–E5. DOI:10.1038/s41586-023-06083-8
4. Hasan M. F., Smith R., Vajedian S., et al. (2023). Global land subsidence mapping reveals widespread loss of aquifer storage capacity. *Nat. Commun.* **14**:6180. DOI:10.1038/s41467-023-41933-z
5. Zhong Y., Zhong M., Feng W., et al. (2018). Groundwater depletion in the West Liaohe River Basin, China and its implications revealed by GRACE and in situ measurements. *Remote Sens.* **10**:493. DOI:10.3390/rs10040493
6. Feng T., Shen Y., Chen Q., et al. (2022). Groundwater storage change and driving factor analysis in north china using independent component decomposition. *J. Hydrol.* **609**:127708. DOI:10.1016/j.jhydrol.2022.127708
7. Yang W., Long D., Scanlon B. R., et al. (2022). Human intervention will stabilize groundwater storage across the North China Plain. *Water Resour. Res.* **58**:e2021WR030884. DOI:10.1029/2021WR030884
8. Long D., Scanlon B. R., Longuevergne L., et al. (2013). GRACE satellite monitoring of

using the gravity recovery and climate experiment (GRACE) data and ground - based measurements. *Water Resour. Res.* **49**:2110–2118. DOI:10.1002/wrcr.20192

16. Frappart F. and Ramillien G. (2018). Monitoring groundwater storage changes using the Gravity Recovery and Climate Experiment (GRACE) satellite mission: A review. *Remote Sens.* **10**:829. DOI:10.3390/rs10060829
17. Li H., Pan Y., Huang Z., et al. (2023). A new GRACE downscaling approach for deriving high-resolution groundwater storage changes using ground-based scaling factors. *Water Resour. Res.* **60**:e2023WR035210. DOI:10.1029/2023WR035210
18. Wang Y., Li C., Cui Y., et al. (2024). Spatial downscaling of GRACE-derived groundwater storage changes across diverse climates and human interventions with Random Forests. *J. Hydrol.* **640**:131708. DOI:10.1016/j.jhydrol.2024.131708
19. Chen L., He Q., Liu K., et al. (2019). Downscaling of GRACE-derived groundwater storage based on the random forest model. *Remote Sens.* **11**:2979. DOI:10.3390/rs11242979
20. Yin W., Hu L., Zhang M., et al. (2018). Statistical downscaling of GRACE - derived groundwater storage using ET data in the North China plain. *J. Geophys. Res. Atmos.* **123**:5973–5987. DOI:10.1029/2017JD027468
21. Long D., Pan Y., Zhou J., et al. (2017). Global analysis of spatiotemporal variability in merged total water storage changes using multiple GRACE products and global hydrological models. *Remote Sens. Environ.* **192**:198–216. DOI:10.1016/j.rse.2017.02.011
22. Reichstein M., Camps-Valls G., Stevens B., et al. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature* **566**:195–204. DOI:10.1038/s41586-019-0912-1
23. Humphrey V. and Gudmundsson L. (2019). GRACE-REC: A reconstruction of climate-driven water storage changes over the last century. *Earth Syst. Sci. Data* **11**:1153–1170. DOI:10.5194/essd-11-1153-2019
24. Kalu I., Ndehedehe C. E., Ferreira V. G., et al. (2024). Statistical downscaling of GRACE terrestrial water storage changes based on the Australian Water Outlook model. *Sci.*

- Rep. **14**:10113. DOI:10.1038/s41598-024-60366-2
25. Liang X. S. (2014). Unraveling the cause-effect relation between time series. *Phys. Rev. E* **90**:052150. DOI:10.1103/PhysRevE.90.052150
 26. Thomas B. F. and Famiglietti J. S. (2019). Identifying climate-induced groundwater depletion in GRACE observations. *Sci. Rep.* **9**:4124. DOI:10.1038/s41598-019-40155-y
 27. Li P., Zha Y., Shi L., et al. (2021). Identification of the terrestrial water storage change features in the North China Plain via independent component analysis. *J. Hydrol. Res. Stud.* **38**:100955. DOI:10.1016/j.ejrh.2021.100955
 28. Alam S., Gebremichael M., Ban Z., et al. (2021). Post - drought groundwater storage recovery in California's Central Valley. *Water Resour. Res.* **57**:e2021WR030352. DOI:10.1029/2021WR030352
 29. Li M., Sun J., Xue L., et al. (2025). Characterizing aquifer properties and groundwater storage at North China Plain using geodetic and hydrological measurements. *Water Resour. Res.* **61**:e2024WR037425. DOI: 10.1029/2024WR037425
 30. Gong H., Pan Y., Zheng L., et al. (2018). Long-term groundwater storage changes and land subsidence development in the North China Plain (1971–2015). *Hydrogeol. J.* **26**:1417–1427. DOI:10.1007/s10040-018-1768-4
 31. Cai X. (2008). Water stress, water transfer and social equity in Northern China-implications for policy reforms. *J. Environ. Manage.* **87**:14–25. DOI:10.1016/j.jenvman.2006.12.046
 32. Chen X., Wang P., Muhammad T., et al. (2020). Subsystem-level groundwater footprint assessment in North China Plain – The world's largest groundwater depression cone. *Ecol. Indicators* **117**:106662. DOI: 10.1016/j.ecolind.2020.106662
 33. Yang J. and Huang X. (2021). The 30 m annual land cover dataset and its dynamics in China from 1990 to 2019. *Earth Syst. Sci. Data* **13**:3907–3925. DOI:10.5194/essd-13-3907-2021
 34. Zongyu C., Yuhong F., Jinxiang M., et al. (2009). Analysis on evolution of groundwater depression cones and its leading factors in North China Plain. *Resour. Sci.* **31**:394–399. <https://www.resci.cn/EN/Y2009/V31/I3/394>
 35. Zhao X., Yu Y., Cheng J., et al. (2024). A novel framework for heterogeneity decomposition and mechanism inference in spatiotemporal evolution of groundwater storage: Case study in the North China Plain. *Water Resour. Res.* **60**:e2023WR036102. DOI:10.1029/2023wr036102
 36. He H., Yang K., Wang S., et al. (2021). Deep learning approaches to spatial downscaling of GRACE terrestrial water storage products using EALCO model over Canada. *Canadian J. Remote Sens.* **47**:657–675. DOI:10.1080/07038992.2021.1954498
 37. von Asmuth J. R., Bierkens M. F. and Maas K. (2002). Transfer function - noise modeling in continuous time using predefined impulse response functions. *Water Resour. Res.* **38**:23–21–23–12. DOI: 10.1029/2001WR001136
 38. Collenteur R. A., Bakker M., Klammler G., et al. (2021). Estimation of groundwater

recharge from groundwater levels using nonlinear transfer function noise models and comparison to lysimeter data. *Hydrol. Earth Syst. Sci.* **25**:2931–2949. DOI:10.5194/hess-25-2931-2021

39. Liang X. S., Xu F., Rong Y., et al. (2021). El Niño Modoki can be mostly predicted more than 10 years ahead of time. *Sci. Rep.* **11**:17860. DOI:10.1038/s41598-021-97111-y

FUNDING AND ACKNOWLEDGMENTS

The work is partially supported by the National Key Research and Development Program of China (2024YFE0106700, 2024YFC3212200), the National Natural Science Foundation of China (42301501), the Open Research Fund of Ministry of Education Key Laboratory for Earth System Modeling (Tsinghua University), the Open Research Fund of Key Laboratory of Water Safety for Beijing-Tianjin-Hebei Region of Ministry of Water Resources (IWHR-JJJ-202401), the Natural Science Foundation of Jiangsu Province (BK20240258), Fundamental and Interdisciplinary Disciplines Breakthrough Plan of the Ministry of Education of China (JYB2025XDXM111), and Shenzhen Key Laboratory Program (SYSPG20241211173845013). Yeh is partially supported by US NSF grant 000316729. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

AUTHOR CONTRIBUTIONS

Jiaqi Yao, Huiqing Hao, and Mengran Wang contributed equally to this work. Jiaqi Yao wrote the original draft. Huiqing Hao and Mengran Wang were responsible for the technical implementation, model development, and data analysis. Nan Xu, Yonghong Hao, and Hui Lu conceived and supervised the study. Walter A. Illman, Guoqiang Wang, Tian-Chyi Jim Yeh, Peng Gong, and Qingquan Li contributed to result interpretation and manuscript revision. All authors contributed to the manuscript and approved the final version.

DECLARATION OF INTERESTS

The authors declare no competing interests.

DATA AND CODE AVAILABILITY

Data are available from the corresponding author upon reasonable request.

SUPPLEMENTAL INFORMATION

It can be found online at <https://doi.org/10.59717/j.tiw.2026.100010>.