Journal of Hydrology 681 (2026) 136710

Contents lists available at ScienceDirect

Journal of Hydrology

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

Human proximity to rivers in China

Xuejing Zhang [a,1], Wenfang Cao [b,1], Dandan Zhao [c], Yanlong Guan [d], Pengfei Wang [e,b], Jeffrey J. McDonnell [f,g,h], Junguo Liu [i,j,*]

[a] School of Water Conservancy, North China University of Water Resources and Electric Power, Zhengzhou 450046, China
[b] School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen, 518055, China
[c] Institute of Surface-Earth System Science, School of Earth System Science, Tianjin University, Tianjin 300072, China
[d] College of Resources and Environment, Fujian Agriculture and Forestry University, Fuzhou 350002, China
[e] School of Environment, Harbin Institute of Technology, Harbin 150090, China
[f] Global Institute for Water Security and School of Environment and Sustainability, University of Saskatchewan, Saskatoon S7N 2X8, Canada
[g] North China University of Water Resources and Electric Power, Zhengzhou 450046, China
[h] School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham B15 2TT, UK
[i] Yellow River Research Institute, North China University of Water Resources and Electric Power, Zhengzhou, 450046, China
[j] Henan Provincial Key Laboratory of Hydrosphere and Watershed Water Security, North China University of Water Resources and Electric Power, Zhengzhou, 450046, China

ARTICLE INFO

This manuscript was handled by Marco Borga, with the assistance of Hanbo Yang, Ph.D., Associate Editor

Keywords:
Human–river proximity
Proximity to Rivers Index
Population distribution
Random Forest
Urbanization
China

ABSTRACT

River proximity shapes human settlement patterns and water resource dependence, yet the long-term changes of human–river proximity and their drivers remain poorly understood. Using high-resolution population data for ten reference years, we developed a Proximity to Rivers Index (PRI) to quantify changes in population-density gradients relative to major rivers across China from 1953 to 2020. Random Forest models and counterfactual scenarios were used to examine socio-economic and natural associations with PRI. Although China's population more than doubled, the median population–river distance remained close to 10 km. However, PRI became progressively more negative from 1953 to 2020, indicating strengthening population concentration near rivers without a substantial shift in the median population location. Basin-scale trajectories were heterogeneous: human–river proximity generally strengthened in northern basins, whereas several southern basins exhibited weakening or more complex trends. Socio-economic variables were more strongly associated with the long-term PRI change, particularly in northern basins, while natural variability was more closely related to fluctuations among reference years in several southern basins. Urbanization rate and GDP had the highest relative importance nationally. Urban settlements were consistently closer to rivers than rural settlements, whereas the relationship between per-capita GDP and river proximity varied among Eastern, Central, and Western China. These findings highlight the need for basin-specific management that coordinates urban development, river-corridor protection, water-resource allocation, and climate adaptation.

1. Introduction

Rivers have served as a vital bond for human civilization for millennia (Macklin and Lewin, 2015; Pizaña and Alejandro, 2021). Early human societies emerged along major rivers, which provided reliable freshwater, fertile floodplains, and navigable waterways, thereby fostering the great civilizations (Gou et al., 2025; Kummu et al., 2011; Wang et al., 2022). However, modern water-supply systems, groundwater extraction, long-distance water transfers, and improved transportation networks may reduce direct dependence on nearby river channels (Zhang et al., 2021; Kuang et al., 2024; Wang and Xie, 2025). At the same time, settlements close to rivers may face greater exposure to flooding, especially under climate change and rapid urban expansion (Alifu et al., 2022; Wang et al., 2023; Masrur et al., 2022; Tang et al., 2024; Wang et al., 2024). Therefore, human proximity to rivers is dynamic rather than static, and is shaped by complex factors. Consequently, understanding the spatiotemporal patterns of human proximity to rivers is fundamental for socio-hydrology (Sivapalan et al., 2012), as

* Corresponding author at: Yellow River Research Institute, North China University of Water Resources and Electric Power, Zhengzhou, 450046, China.
E-mail address: junguo.liu@gmail.com (J. Liu).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.jhydrol.2026.136710
Received 5 April 2026; Received in revised form 25 September 2026; Accepted 27 October 2026
Available online 31 October 2026
0022-1694/© 2026 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

it provides key insights into the evolving human–water relationship under the influence of climate variability and socio-economic development (Di Baldassarre et al., 2013; Ridolfi et al., 2020).

Recent advancements in high-resolution population datasets and remote sensing technologies have made it possible to quantitatively assess human proximity to rivers (Mård et al., 2018). Global analyses have shown that human populations predominantly settle near rivers in most regions (Ceola et al., 2015), especially in downstream reaches and river mouths (Fang et al., 2018), with settlement density positively correlated with river density (Li et al., 2024a). Globally, over 50% of the population lives within 3 km of surface freshwater bodies (rivers or lakes), whereas only about 10% live farther than 10 km away (Kummu et al., 2011). However, a case study in the United States showed a significant evolution in the distance between populations and rivers over time. Before industrialization, humans tended to settle near rivers, while after industrialization, technological development and groundwater extraction led to a gradual shift away from rivers (Fang and Jawitz, 2019). Furthermore, historical evidence has indicated that topography determines the spatial variation of human reliance on rivers (Cao et al., 2023). The construction of flood control infrastructure, such as reservoirs and dams, has reduced flood risk, thus exacerbating the clustering of settlements near rivers (Di Baldassarre et al., 2013; Mård et al., 2018). These studies demonstrate that human proximity to rivers is dynamic and evolves with socio-economic development and environmental change.

Despite these advances, three critical gaps remain. First, most studies are global assessments, leading to limited exploration of the evolution of river-oriented settlement patterns in rapidly transforming countries. Second, although some studies suggest that natural changes or socio-economic development have altered the distance between humans and rivers, their relative contributions to these dynamics are rarely quantified. Third, research in China has examined ancient human reliance on rivers (Cao et al., 2023), the relationship between land-use conflicts and distance from coastlines (Zong et al., 2021), and the distribution characteristics of rural settlements along the Yellow River (Li et al., 2024a). However, these studies have not explicitly evaluated the dynamic evolution of population proximity to rivers in contemporary China and how socio-economic and natural factors reshape these patterns. Addressing these issues is essential for understanding human–water relationships and optimizing water resource allocation.

Over the past seven decades, China has undergone a rapid demographic and economic transformation (Chen and Liu, 2022; Chen et al., 2024; Deng et al., 2015), accompanied by substantial changes in regional hydrological and climatic conditions (Chi et al., 2023; Xu et al., 2022). These broad transformations provide important context for understanding the long-term evolution of human–river relationships. Here we investigate the spatiotemporal evolution and drivers of human–river proximity in China from 1953 to 2020. Specifically, we (1) quantified the long-term changes in population proximity to rivers at national and basin scales using high-resolution population data for ten reference years; (2) applied Random Forest models to assess the relative importance of selected natural and socio-economic factors; and (3) used counterfactual scenarios to distinguish model-predicted changes associated with socio-economic development and natural variability. By revealing the long-term evolution and regional heterogeneity of human–river proximity, this study advances understanding of human–river relationships and provides insights for basin-specific water management and sustainable river-basin planning.

2. Data and methods

2.1. Data sources and preprocessing

2.1.1. Data sources

Population data included national population statistics and 1-km gridded population datasets. This study used ten population reference years to ensure long-term temporal coverage and consistency across datasets: 1953, 1964, 1982, 1990, 1995, 2000, 2005, 2010, 2015, and 2020. The total population of China was obtained from the National Bureau of Statistics of China (NBS). The 1-km gridded population datasets for 1953, 1964 and 1982 were reconstructed from county-level census data following Wu et al. (2022), while gridded population data from 1990 to 2020 at 5-year intervals were obtained from the Resource and Environment Science Data Platform (RESDC) (https://www.resdc.cn/).

River network data were obtained from the HydroRIVERS dataset (https://www.hydrosheds.org/products/hydrorivers). River reaches with Strahler stream orders 4–9 were selected to represent the major river network used to calculate population–river distance, focusing on larger rivers more closely associated with settlement development, water supply, and navigation. To assess the sensitivity of PRI to the selected river-order threshold, we also recalculated PRI using an alternative river network that included Strahler stream orders 3–9 (Fig. S1). Basin boundaries for the nine major river basins in China and the regional boundaries delineating Eastern, Central, and Western China were obtained from the RESDC and used for basin- and regional-scale analyses, respectively (Fig. 1a and c).

Socio-economic and natural drivers were used to explain changes in PRI at both national and basin scales. Socio-economic drivers included urbanization rate, Gross Domestic Product (GDP), and secondary industry proportion, obtained from the National Bureau of Statistics of China and provincial statistical yearbooks. Because this study focuses on the redistribution of population relative to rivers, urbanization was represented by the urbanization rate, calculated as the proportion of the urban population in the total population. The proportion of secondary industry in GDP was included separately to characterize industrial structure and industrialization and was not used as a proxy for urbanization. Natural drivers included precipitation, temperature, arid index, runoff, climate extreme events, and Normalized Difference Vegetation Index (NDVI). Detailed information on these drivers, including data sources, temporal coverage, and relevance to PRI, is summarized in Table 1. In addition, urban and rural land-use data from 1980 to 2023 and 1-km gridded GDP data from 1995 to 2020 were obtained from the RESDC and used to compare human–river proximity across settlement types and per-capita GDP groups.

2.1.2. Data preprocessing

The 1-km gridded population data for 1953, 1964, and 1982 were reconstructed by allocating county-level population uniformly within each county. Gridded variables were processed at their original spatial resolutions and then aggregated to the analytical units used in this study, namely the national scale or the nine major river basins. For national-scale analysis, natural drivers derived from gridded datasets were aggregated to the national level using spatial averaging or summation, while socio-economic drivers were represented by national statistical records. For basin-scale analysis, gridded natural drivers were first clipped by basin boundaries and then aggregated within each basin using either spatial averaging or regional summation. Basin-scale socio-economic drivers were derived from provincial statistical data using population-density-weighted averages to account for the uneven spatial distribution of demographic and economic activities.

All variables were matched to the ten population reference years used for PRI calculation: 1953, 1964, 1982, 1990, 1995, 2000, 2005, 2010, 2015, and 2020. Because the temporal coverage differed among datasets, unavailable values in specific reference years were treated as missing values rather than being replaced by zeros or temporally interpolated. The temporal coverage and years used for each driving factor are summarized in Table 1.

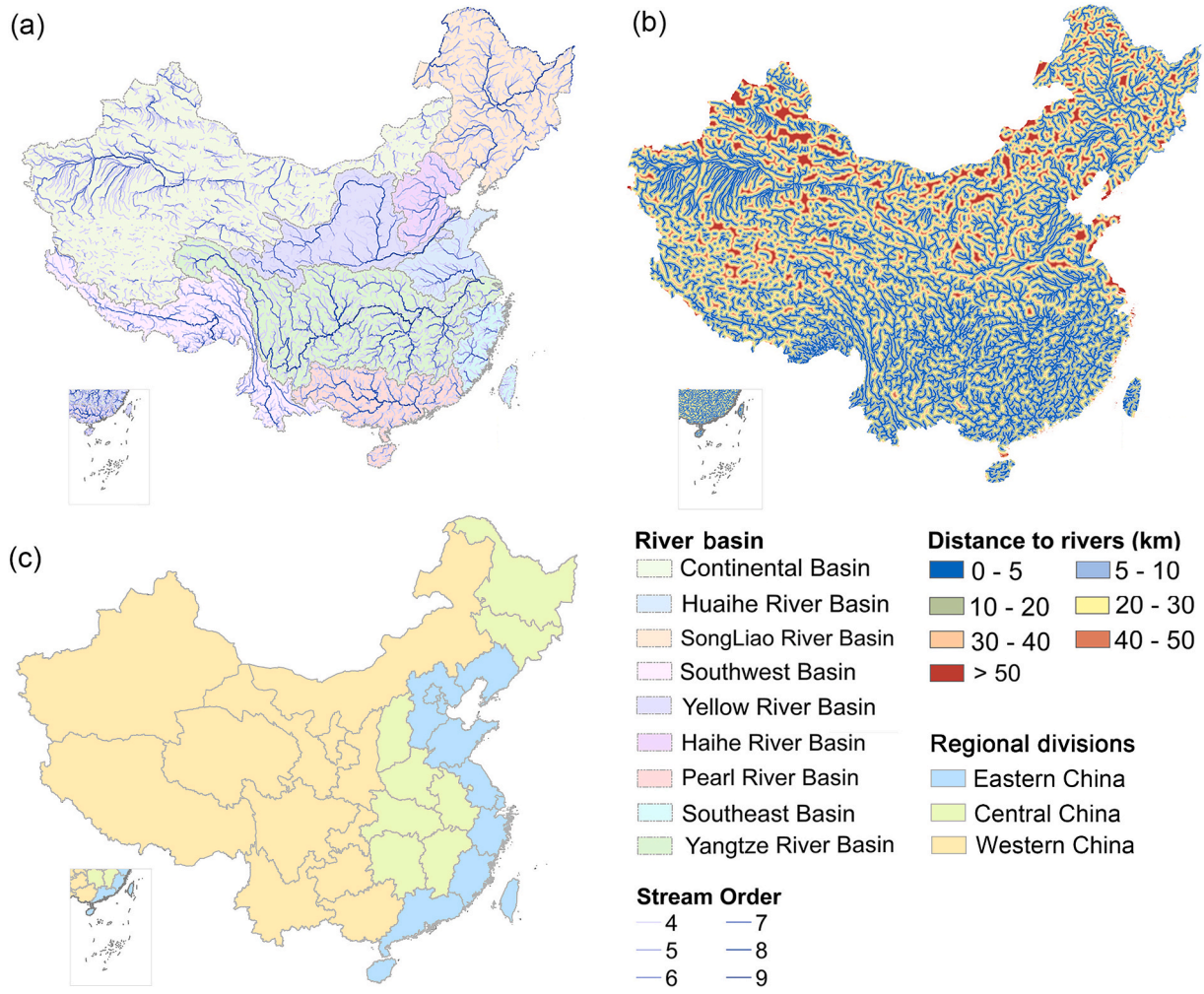


Fig. 1. Study area, river network, major river basins, distance to rivers, and regional divisions in China. (a) Nine major river basins and the selected river network with Strahler stream orders 4–9. (b) Spatial distribution of Euclidean distance to the nearest segment of the selected river network. (c) Division of China into Eastern, Central, and Western regions. Insets show the South China Sea islands.

2.2. Methods

2.2.1. Median population–river distance and normalized population density

The median population–river distance (D_{pop}) was defined as the distance within which 50% of the population is located relative to the river network, meaning that half of the population lives within D_{pop} km of a river, while the other half lives beyond this distance. To further characterize human–river proximity, we examined how normalized population density varies with distance to rivers. The calculation followed a four-step framework. First, we reconstructed 1-km population grids for 1953, 1964, and 1982 based on county-level census data using an average allocation method. These were integrated with post-1990 gridded population datasets at 5-year intervals to construct a long-term series (1953–2020). Second, the Euclidean distance (D) from each 1-km land grid cell to the nearest river segment was calculated (Fig. 1b). To limit the influence of extreme distances, the maximum distance was set at 50 km, which covers more than 95% of the study area and population (Fig. S2). Third, population data were overlaid with the Euclidean distance layer. For each 1-km distance zone ($d \in [0, 50]$), we summed the total population (P_d) and the total land area (A_d). Finally, population density for each distance zone was calculated as $\rho_{d} = P_d / A_d$. To ensure interannual comparability, ρ_d was normalized to the 0–1 range ($\widehat{\rho_d}$), using the maximum and minimum population densities recorded throughout the study period.

To further evaluate whether the long-term PRI trend was influenced by changes in the overall magnitude of population density, we constructed a population-growth-only counterfactual scenario. Specifically, the spatial distribution of population was fixed at the 1953 pattern, while population density in each river-distance zone was proportionally scaled according to the ratio of national population size between each reference year and 1953. PRI was then recalculated using the same normalization and regression procedures as in the observed analysis (Fig. S3).

2.2.2. Proximity to Rivers Index (PRI) definition and trend detection

We developed the Proximity to Rivers Index (PRI) to quantify how population distribution changes with distance from rivers. The PRI was defined as the slope of a linear regression between the normalized population density ($\widehat{\rho_d}$) and the distance to rivers (D).

$$\widehat{\rho_d} = \alpha + \lambda D + \varepsilon \quad (1)$$

$$\mathrm{PRI} = \lambda \quad (2)$$

where α is the intercept, λ is the regression slope, and ε is the residual term. Because $\widehat{\rho}_d$ is dimensionless and D is measured in km, PRI has a unit of km^{-1}. Thus, PRI represents the change in normalized population density per kilometer increase in distance from rivers.

The sign and magnitude of PRI indicate the direction and strength of the population–river distance gradient, respectively. Negative values

Table 1
Natural and socio-economic driving factors used in the PRI framework.

Category	Indicator	Unit	Years used in this study	Data source	Relevance to PRI
Socio-economic	Urbanization rate	%	All PRI years	NBS; provincial statistical yearbooks	Reflects population urbanization and rural-to-urban population redistribution, which may reshape settlement distribution relative to rivers.
	GDP	10^8CNY	All PRI years	NBS; provincial statistical yearbooks	Represents overall economic development.
	Secondary Industry proportion	%	All PRI years	NBS; provincial statistical yearbooks	Represents industrialization and economic-structure change, defined as the share of secondary industry output in total GDP.
Natural	Runoff	mm/y	All except 1953 and 2020	CNRD v1.0	Represents surface water availability, directly related to the hydrological conditions of river systems.
	Precipitation	mm/y	All PRI years	TPDC	Represents regional water input, which may influence the attractiveness or necessity of living near rivers.
	Temperature	°C	All PRI years	TPDC	Represents thermal conditions that may affect habitability and regional settlement patterns.
	Climate extreme events	σ	All PRI years	NESSDC	Represents the severity of annual compound extreme climate events, including dry-hot, wet-hot, wet-windy, and dry-hot-windy events, which may influence regional settlement suitability and human–river proximity.
	NDVI	—	All except 1953 and 1964	RESDC	Represents vegetation and environmental conditions, which may reflect broader ecological suitability for settlement.
	Arid Index	—	All PRI years	TPDC	Represents regional dryness and water availability, which may influence human dependence on nearby rivers.

Note: PRI years refer to the ten discrete population reference years used for PRI calculation: 1953, 1964, 1982, 1990, 1995, 2000, 2005, 2010, 2015, and 2020. NBS, National Bureau of Statistics of China; TPDC, National Tibetan Plateau / Third Pole Environment Data Center; CNRD, China Natural Runoff Dataset; NESSDC, National Earth System Science Data Center; RESDC, Resource and Environment Science Data Platform. σ denotes the standardized severity unit provided in the source dataset metadata.

indicate greater population concentration near rivers, values near zero indicate a weak gradient, and positive values indicate increasing population density with distance; increasingly negative values over time indicate strengthening human–river proximity, whereas shifts toward zero or positive values indicate weakening proximity.

Potential breakpoints in basin-scale PRI trends were estimated using segmented regression with the Muggeo iterative procedure (Muggeo, 2003) and assessed using the Davies test (Davies, 1987). The P values across the nine basins were adjusted using the Benjamini–Hochberg procedure (Benjamini and Hochberg, 1995). A segmented model was retained only when the adjusted Davies $P < 0.05$ and its AICc was lower than that of the linear model; otherwise, the overall linear trend was reported.

To assess the sensitivity of PRI to the selected river-order threshold, we repeated the national- and basin-scale analyses using an alternative river network including Strahler stream orders 3–9. First- and second-order reaches were not included because they mainly represent small headwater streams. Their inclusion would substantially increase network density and computational demand and shift the analytical focus from major river corridors to local drainage channels. All calculation procedures were consistent with the main analysis, and the effective distance range was determined using the same criterion based on the distance distributions of population and non-river grid cells.

2.2.3. Driving factor analysis using Random Forest

To identify and quantify the primary drivers of human proximity to rivers, we employed Random Forest (RF) regression, a multi-decision tree model based on randomly sampled data and features (Salman et al., 2024), which is an effective tool for identifying complex nonlinear relationships among multiple factors (Deng et al., 2025). The optimal parameters for the model were determined by minimizing the out-of-bag (OOB) error (Table 2, Fig. S4), a method that ensures unbiased estimation of model generalization (Wang et al., 2021). The PRI was modeled as a nonlinear function of multiple natural and socio-economic variables:

$$\mathrm{PRI}_t = f\left(\mathrm{X}_{nat,t}, \mathrm{X}_{soc,t}\right) \tag{3}$$

where PRI_t represents the Proximity to Rivers Index at time t, and $\mathrm{X}_{nat,\ t}$ and $\mathrm{X}_{soc,\ t}$ represent the vectors of natural and socio-economic drivers at time t. In this study, $\mathrm{X}_{soc,\ t}$ included urbanization rate, GDP, and secondary industry proportion, representing population urbanization, economic development, and industrial structure, respectively. $\mathrm{X}_{nat,\ t}$ included precipitation, temperature, arid index, runoff, climate extreme events, and NDVI, representing hydroclimatic conditions, surface water availability, climate extremes, and vegetation-related environmental variability.

Table 2
Optimal Random Forest parameters and model performance.

Basin	n_{tree}	m_{try}	R^2	MSE
Continental Basin	80	2	0.89	9.36×10^{-7}
Haihe River Basin	130	3	0.93	7.49×10^{-7}
Huaihe River Basin	150	2	0.85	6.32×10^{-7}
Pearl River Basin	80	3	0.84	2.89×10^{-7}
SongLiao River Basin	90	3	0.89	2.51×10^{-7}
Southeast Basin	20	2	0.90	5.62×10^{-7}
Southwest Basin	70	4	0.86	5.99×10^{-7}
Yangtze River Basin	120	2	0.87	3.81×10^{-7}
Yellow River Basin	130	4	0.90	4.63×10^{-7}
National scale	10	3	0.94	2.31×10^{-7}

Note: n_{tree} denotes the number of trees in the Random Forest model, and m_{try} denotes the number of candidate predictor variables randomly selected at each split. R^2 represents the coefficient of determination, and MSE represents the mean squared error. MSE is expressed in km^{-2}. No P values are reported for R^2 and MSE because they are model-performance metrics rather than hypothesis-test statistics.

The contribution of each driving factor was quantified using the Mean Decrease in Accuracy (MDA), which is the increase in prediction error caused by random permutations of the variable's value. A higher MDA indicates a greater contribution of the factor to the model results (Han et al., 2016). For ease of comparison, we standardized the contribution values of multiple factors and expressed them as factor importance (%).

$$\textit{Factor importance}\,(\%) = \frac{\mathrm{MDA}_i}{\sum_{j=1}^{p}\mathrm{MDA}_j} \times 100 \tag{4}$$

where MDA_i represents the mean decrease in accuracy when the value of variable i is permuted, and p represents the total number of variables. It is important to note that factor importance (%) reflects the relative proportion of a variable in the increase in model prediction error, rather than its explained variance.

2.2.4. Counterfactual scenarios for natural variability and socio-economic development

To distinguish model-predicted PRI changes associated with socio-economic development and natural variability, two counterfactual scenarios were constructed using the trained Random Forest model. For each predictor, the baseline value was defined using its earliest available reference year. All socio-economic variables were fixed at their 1953 values, whereas natural variables were fixed at the values from their respective earliest available reference years.

In the natural-driven scenario, socio-economic variables were held at their earliest-year values, while natural variables retained their observed year-specific values. In the socio-economic-driven scenario, natural variables were held at their earliest-year values, while socio-economic variables retained their observed year-specific values. For basin-scale simulations, the earliest-year values were determined separately for each basin.

$$PRI_{\text{Natural-driven},t} = f\left(X_{\text{nat},t}, X_{\text{soc},0}\right) \tag{5}$$

$$PRI_{\text{Socio-economic-driven},t} = f\left(X_{\text{nat},0}, X_{\text{soc},t}\right) \tag{6}$$

where $PRI_{\text{Natural-driven},\,t}$ and $PRI_{\text{Socio-economic-driven},\,t}$ denote the PRI values estimated under the natural-driven and socio-economic-driven scenarios in year t, respectively. $X_{\text{nat},\,t}$ and $X_{\text{soc},\,t}$ denote the observed natural and socio-economic drivers in year t, whereas $X_{\text{nat},0}$ and $X_{\text{soc},0}$ denote the values from the earliest available reference year for each variable. The function $f()$ represents the trained Random Forest model.

2.2.5. PRI comparison by settlement type and per-capita GDP level

Because gridded urbanization-rate data were unavailable, urban and rural settlement land-use data were used to distinguish settlement types. For each reference year, the areas of urban and rural settlements within successive 1-km distance bands were divided by the total land area of each band to calculate settlement-area density. The normalized densities were regressed against distance to derive urban and rural PRI using the same distance-gradient framework as the population-based PRI. D_U and D_R were defined as the distances at which the cumulative areas of urban and rural settlements reached 50%, respectively. Comparisons with overall human PRI provided an independent consistency assessment based on non-population data.

Per-capita GDP was calculated from the gridded GDP and population datasets. Within Eastern, Central, and Western China separately, grid cells were ranked by per-capita GDP and divided into approximately equal-area low-, medium-, and high-level groups for each reference year. Group-specific area density was calculated within each distance band, normalized, and regressed against distance to derive PRI. D_{Low}, D_{Mid}, and D_{High} represented the corresponding median distances. Regional human PRI was calculated independently from gridded population data for each region using the population-based distance-gradient framework.

3. Results

3.1. Overall patterns of human–river proximity

The total population in China increased from approximately 600 million in 1953 to 1.4 billion in 2020, with an average annual increase of approximately 11.9 million people (Fig. 2a). The median population–river distance (D_{pop}) was defined as the distance within which 50% of the population is distributed relative to the river network. The results showed that nearly 50% of the Chinese population lived within

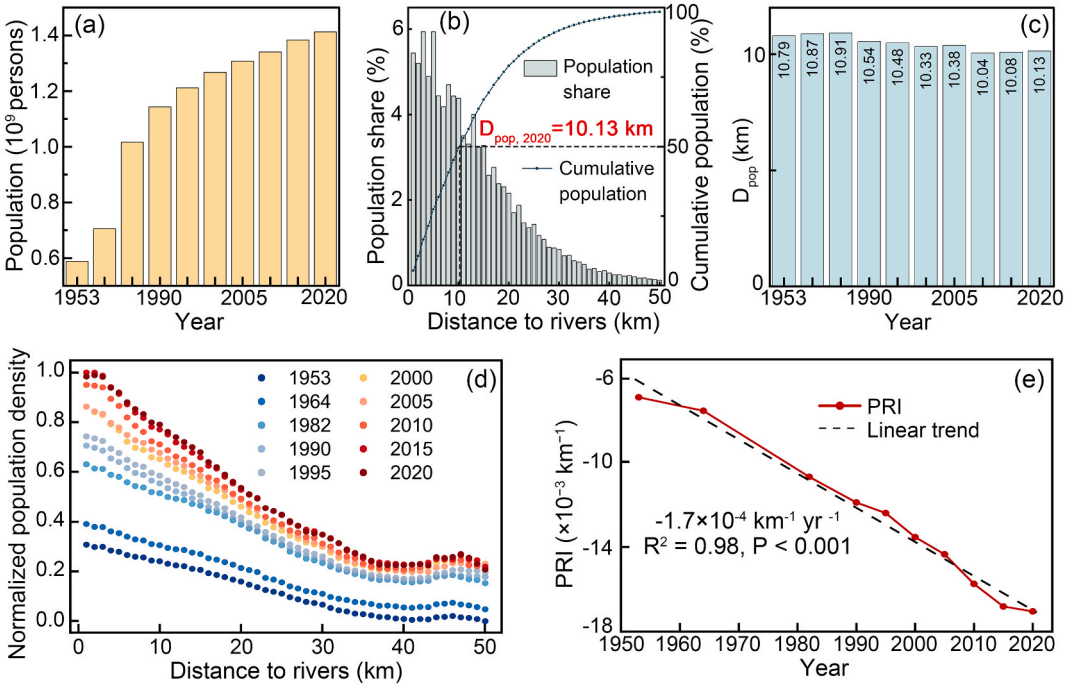


Fig. 2. Overall pattern of human proximity to rivers in China from 1953 to 2020. (a) Total population in China. (b) Schematic illustration of the calculation of the median population–river distance (D_{pop}) in 2020, defined as the distance within which 50% of the cumulative population is distributed relative to the river network. Results for 1953–2015 are shown in Fig. S5. Grey bars indicate the population share in each 1-km distance zone in 2020, the blue line indicates the cumulative population percentage, and dashed lines indicate the 50% cumulative population threshold used to define D_{pop}. (c) Temporal changes in D_{pop}. (d) Variation in normalized population density with distance to rivers at the national scale. (e) Temporal trend of the Proximity to Rivers Index (PRI) at the national scale. The red line indicates PRI values, and the black dashed line indicates the fitted linear trend. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

approximately 10 km of a river, while the remaining half lived beyond this distance (Fig. 2b and S5). D_{pop} remained relatively stable during 1953–2020, ranging from 10.04 km to 10.91 km (Fig. 2c), indicating that the 50th percentile location of the cumulative population distribution relative to rivers remained relatively stable at the national scale.

To further characterize human–river proximity, we examined how normalized population density varied with distance to rivers. From 1953 to 2020, the normalized population density consistently decreased with increasing distance from rivers, indicating that population density was higher in near-river zones than in farther zones (Fig. 2d). Within 1 km of rivers, normalized population density increased from approximately 0.3 in 1953 to nearly 1.0 in 2020, indicating a progressively stronger tendency for populations to concentrate near rivers. This tendency is quantified by the Proximity to Rivers Index (PRI), which remained negative throughout the study period and decreased from -6.9×10^{-3} km^{-1} in 1953 to -17.0×10^{-3} km^{-1} in 2020 (Fig. 2e), at a rate of -1.7×10^{-4} km^{-1} yr^{-1}. Although D_{pop} remained close to 10 km, PRI became progressively more negative, indicating that the population-density gradient toward rivers strengthened without a substantial shift in the median population location. Notably, the minor fluctuations in normalized population density within the 40–50 km range are likely related to sparse and spatially discontinuous population distributions in remote western and northwestern regions, as shown in Fig. 1b.

A sensitivity analysis using an expanded river network comprising Strahler orders 3–9 yielded the same direction of national PRI change and broadly similar basin-scale temporal patterns. Although the national trend became more negative, changing from -1.7×10^{-4} km^{-1} yr^{-1} for orders 4–9 to -4.1×10^{-4} km^{-1} yr^{-1} for orders 3–9, the principal conclusion of strengthening human–river proximity remained unchanged (Fig. S1). A sensitivity analysis was further conducted to evaluate the influence of overall population growth on PRI changes. The population-growth-only counterfactual scenario also showed a declining PRI trajectory, indicating that overall population growth contributed to part of the observed PRI decline. However, the observed PRI became more negative than the counterfactual trajectory after approximately 1995, suggesting that spatial redistribution of population density relative to rivers further strengthened the population-density gradient (Fig. S3).

3.2. Spatiotemporal dynamics of human–river proximity across basins

The PRI exhibited significant spatial heterogeneity across nine river basins. In the Continental, SongLiao, Haihe, and Yellow River basins, PRI remained consistently negative, indicating that population density generally decreased with increasing distance from rivers and that human settlements were therefore concentrated near rivers. The SongLiao, Haihe, and Yellow River basins showed significant overall decreasing trends, with slopes of −1.05, −1.95, and -1.58×10^{-4} km^{-1} yr^{-1}, respectively, but no robust breakpoint was retained (Fig. 3b–d). In the Continental Basin, formal breakpoint detection identified a statistically supported change around 2002. The decreasing rate of PRI accelerated from -0.48×10^{-4} km^{-1} yr^{-1} before 2002 to -5.18×10^{-4} km^{-1} yr^{-1} thereafter (Fig. 3a), indicating a marked strengthening of population concentration near rivers.

Statistically supported changes in trend were also identified in the Southwest and Yangtze River basins. In the Southwest Basin, PRI decreased at a rate of -2.08×10^{-4} km^{-1} yr^{-1} before the breakpoint around 2009, followed by a slight increasing trend of 0.44×10^{-4} km^{-1} yr^{-1} (Fig. 3f). This suggests that the earlier strengthening of river proximity subsequently weakened and approached stabilization. In the

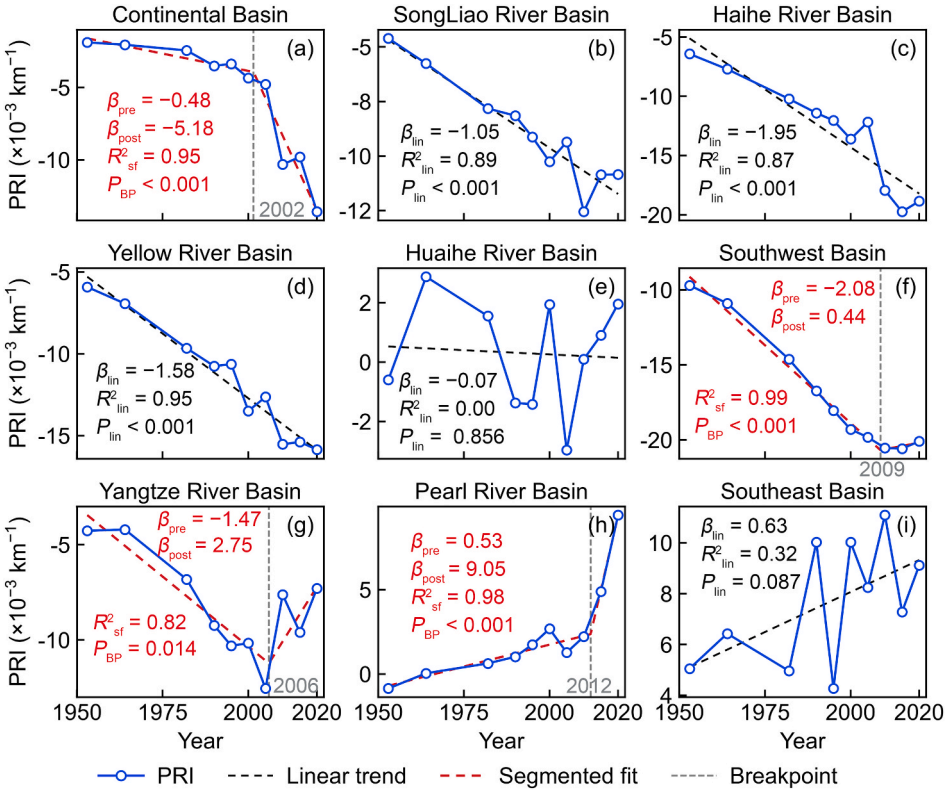


Fig. 3. Temporal dynamics of the PRI and formally detected breakpoints across the nine major river basins of China. Segmented fits are shown only when the Benjamini–Hochberg-adjusted Davies *P* value was below 0.05 and the segmented regression had a lower AICc than the corresponding linear regression; otherwise, the overall linear trend is shown. β_{pre} and β_{post} denote the temporal slopes before and after the breakpoint, respectively, whereas β_{lin} denotes the overall linear slope. P_{BP} denotes the Davies-test *P* value adjusted across the nine basins using the Benjamini–Hochberg procedure, whereas P_{lin} denotes the *P* value for the overall linear trend. R^2_{sf} and R^2_{lin} denote the coefficients of determination of the segmented and linear fits, respectively. PRI values are displayed in units of 10^{-3} km^{-1}, and all temporal slopes are expressed in units of 10^{-4} km^{-1} yr^{-1}.

Yangtze River Basin, PRI decreased at -1.47×10^{-4} km^{-1} yr^{-1} before the breakpoint around 2006 and then increased at 2.75×10^{-4} km^{-1} yr^{-1} (Fig. 3g), indicating a reversal from strengthening to weakening human–river proximity.

PRI values in the Huaihe, Pearl, and Southeast basins were predominantly positive, indicating that normalized population density tended to increase with distance from rivers and that river proximity was comparatively weak (Fig. 3e, h, i). The Pearl River Basin exhibited a statistically supported breakpoint around 2012, with the increasing rate of PRI accelerating from 0.53×10^{-4} km^{-1} yr^{-1} before the breakpoint to 9.05×10^{-4} km^{-1} yr^{-1} thereafter (Fig. 3h). By contrast, neither the Huaihe River Basin nor the Southeast Basin showed a statistically significant overall linear trend, with slopes of -0.07 and 0.63×10^{-4} km^{-1} yr^{-1}, respectively (Fig. 3e, i).

Overall, the northern basins were characterized by persistently negative and generally decreasing PRI values, indicating strong and increasing concentration of human settlements near rivers. In contrast, the southern basins generally exhibited positive or increasing PRI values, suggesting weaker human–river proximity. The formally detected trend reversals in the Yangtze and Southwest basins and the accelerated decline in the Continental Basin and accelerated increase in the Pearl River Basin further demonstrate that the temporal evolution of human–river proximity differed substantially among basin regions.

3.3. Driving mechanism of human proximity to rivers

The Random Forest model showed good performance (Table 2). At the national level, urbanization rate and GDP had the highest relative importance, contributing 30.5% and 24.5%, respectively (Fig. 4), suggesting that national changes in human–river proximity were more strongly associated with population urbanization and economic development. At the basin scale, socio-economic factors showed greater relative importance in six basins, particularly in several water-scarce northern basins, including the Continental, SongLiao, Haihe, and Yellow River basins, whereas natural factors showed greater importance in most southern basins (Fig. 4). For example, temperature contributed 22.1% in the Continental Basin, while arid index and runoff contributed 21.8% and 18.5% in the Yangtze River Basin, respectively.

However, the Southwest Basin was a notable exception among the southern basins. GDP accounted for 49.2% of the total Random Forest variable importance, substantially exceeding that of any individual natural predictor (Fig. 4). This result indicates that temporal changes in PRI in the Southwest Basin were more strongly associated with GDP than with the other predictors included in the model. The possible mechanism is discussed in Section 4.3.

Overall, PRI changes in the northern basins were more strongly associated with socio-economic factors, particularly urbanization rate and GDP, whereas natural factors showed greater relative importance in several southern and coastal basins. Urbanization rate had the highest relative importance at the national scale, while GDP ranked higher in most basin-scale models. These differences indicate that the relative importance of PRI drivers was scale dependent and are further discussed in Section 4.5.

To distinguish the influences of socio-economic development and natural environmental variability, PRI was simulated under two counterfactual scenarios. The socio-economic development more closely reproduced the long-term decline in PRI, whereas environmental variability contributed more to fluctuations in PRI among reference years

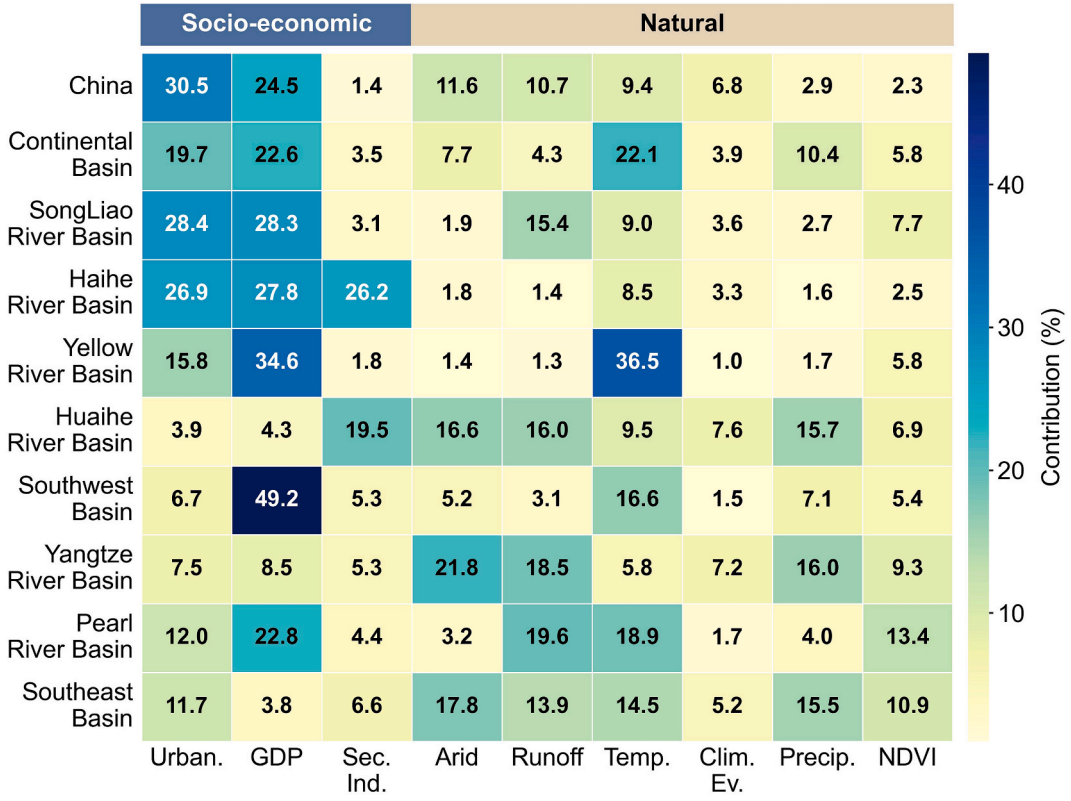


Fig. 4. Relative contributions of socio-economic and natural drivers to PRI. The heatmap presents the normalized relative contribution (%) of each predictor, derived from the mean decrease in accuracy (MDA) of the Random Forest models. Predictors are grouped into socio-economic factors (blue header) and natural factors (tan header). Cell values indicate their relative contributions at the national scale and across the nine major river basins. ***Abbreviations:*** *Urban., urbanization rate; GDP, gross domestic product; Sec. Ind., secondary industry proportion; Arid, arid index; Temp., temperature; Clim. Ev., climate extreme events; Precip., annual precipitation; NDVI, normalized difference vegetation index.* (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

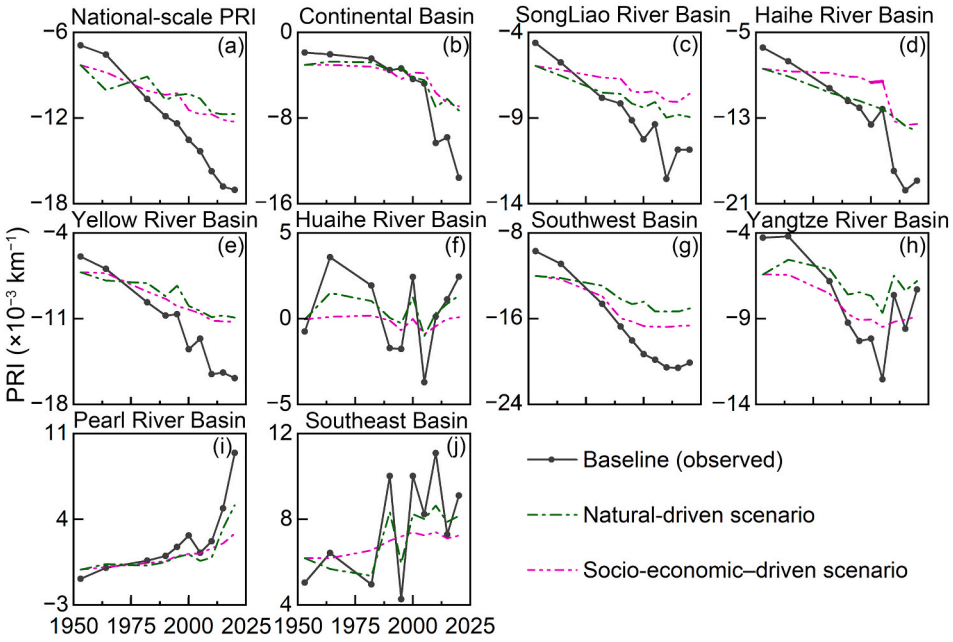


Fig. 5. PRI dynamics under the observed, natural-driven and socio-economic-driven scenarios at the national and basin scales. The black solid line denotes the observed PRI. The green dashed line represents the natural-driven scenario, in which socio-economic drivers were fixed at their values from the earliest available reference year, while natural drivers retained their observed year-specific values. The pink dash-dotted line represents the socio–economic–driven scenario, in which natural drivers were held at their values from the earliest available reference year, while socio-economic drivers retained their observed year-specific values. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(Fig. 5). The observed PRI trajectories in northern basins were more consistent with the socio-economic-driven scenario, especially after the 1980s, while those in the Yangtze, the Pearl, and the Southeast basins were more consistent with the natural-driven scenario (Fig. 5h-j). These patterns support stronger associations of human–river proximity with socio-economic development in northern China and with natural variability in several southern and coastal basins. At the national scale, the two groups of drivers sometimes produced changes in opposite

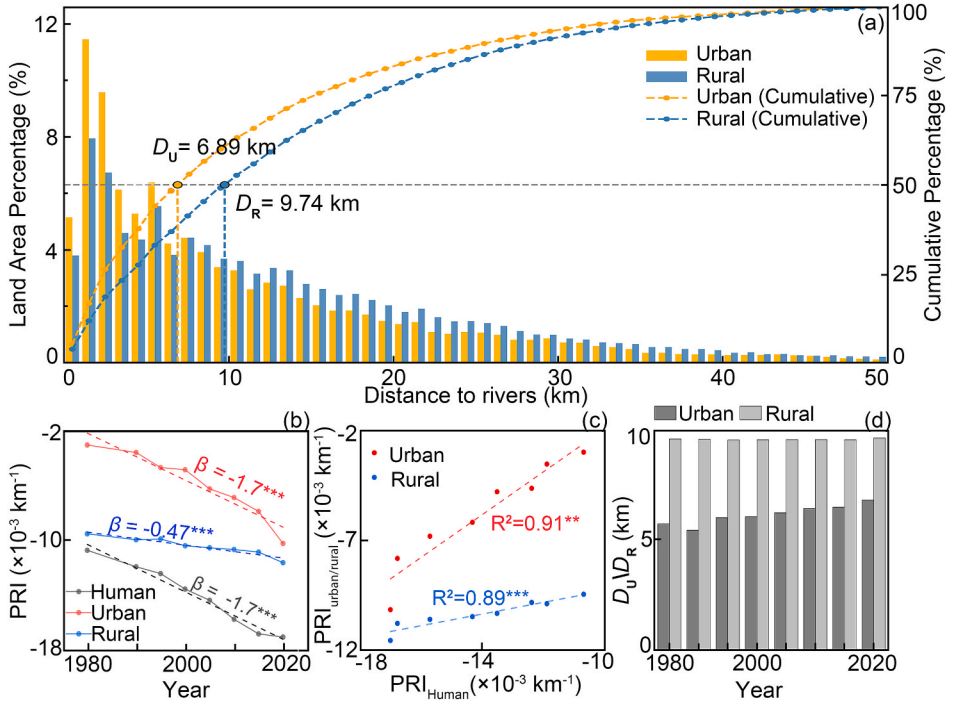


Fig. 6. Contrasting PRI patterns between urban and rural settlements. (a) Distributions of urban and rural land area with distance to rivers and their cumulative percentages. D_U and D_R denote the distances at which the cumulative urban and rural land areas reach 50%, respectively. (b) Temporal variations in overall human PRI and urban- and rural-specific PRI from 1980 to 2020. (c) Covariation of urban- and rural-specific PRI with overall human PRI. (d) Temporal variations in D_U and D_R from 1980 to 2020. Dashed lines indicate linear fits. The temporal trend coefficient β is expressed in units of 10^{-4} km^{-1} yr^{-1}. Asterisks indicate statistical significance: * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

directions, moderating the observed decline in PRI (Fig. 5a).

3.4. Impacts of urban–rural settlements and per-capita GDP levels on PRI

Urbanization rate and GDP were identified as two major drivers of human proximity to rivers in most basins. We therefore examined whether urban and rural settlements differed in their spatial proximity to rivers. The median distance to rivers was consistently shorter for urban settlements than for rural settlements. For urban settlements, D_U increased from 5.50 km in 1980 to 6.89 km in 2020, with a mean of 6.21 km, whereas D_R for rural settlements ranged from 9.65 to 9.74 km, with a mean of 9.68 km (Fig. 6a, d and S6). These results show that urban settlements were consistently located closer to rivers than rural settlements.

PRI was subsequently recalculated separately for urban and rural settlements using observed land-use data. Both PRI series remained negative throughout 1980–2020 and became progressively more negative, indicating strengthening settlement concentration near rivers (Fig. 6b). Urban PRI decreased faster than rural PRI, with temporal trends of -1.7×10^{-4} and -0.47×10^{-4} km^{-1} yr^{-1}, respectively (both $P < 0.001$; Fig. 6b). Both PRI series also covaried closely with overall human PRI ($R^2 = 0.91$, $P < 0.01$ for urban; $R^2 = 0.89$, $P < 0.001$ for rural; Fig. 6c). These results indicate that changes in both settlement types were consistent with the long-term evolution of population–river proximity, with a more pronounced strengthening in urban areas.

We next examined whether river proximity differed among areas grouped by per-capita GDP. Regional human PRI became significantly more negative in Eastern and Central China (-4.16 and -1.2×10^{-4} km^{-1} yr^{-1}, respectively; both $P < 0.05$), whereas no significant temporal trend was observed in Western China (Fig. 7a, d, g). Most GDP-group-specific trends were not significant; the only significant change was an increase in PRI for low-per-capita-GDP areas in Western China (3.06×10^{-4} km^{-1} yr^{-1}, $P < 0.05$; Fig. 7g), indicating weakening proximity to rivers. Covariation with regional human PRI was generally nonsignificant, except for the medium-per-capita-GDP group in Central China ($R^2 = 0.70$, $P < 0.05$; Fig. 7b, e, h). Median distances were broadly similar among the three groups in Eastern China. In Central China, high-per-capita-GDP areas generally had greater median distances than the low- and medium-per-capita-GDP groups, except in 2020, whereas in Western China they consistently had substantially greater distances throughout the study period (Fig. 7c, f, i; Table S1). Overall, the relationship between economic development and river proximity varied geographically, with no consistent tendency for areas with higher per-capita GDP to be located either closer to or farther from rivers.

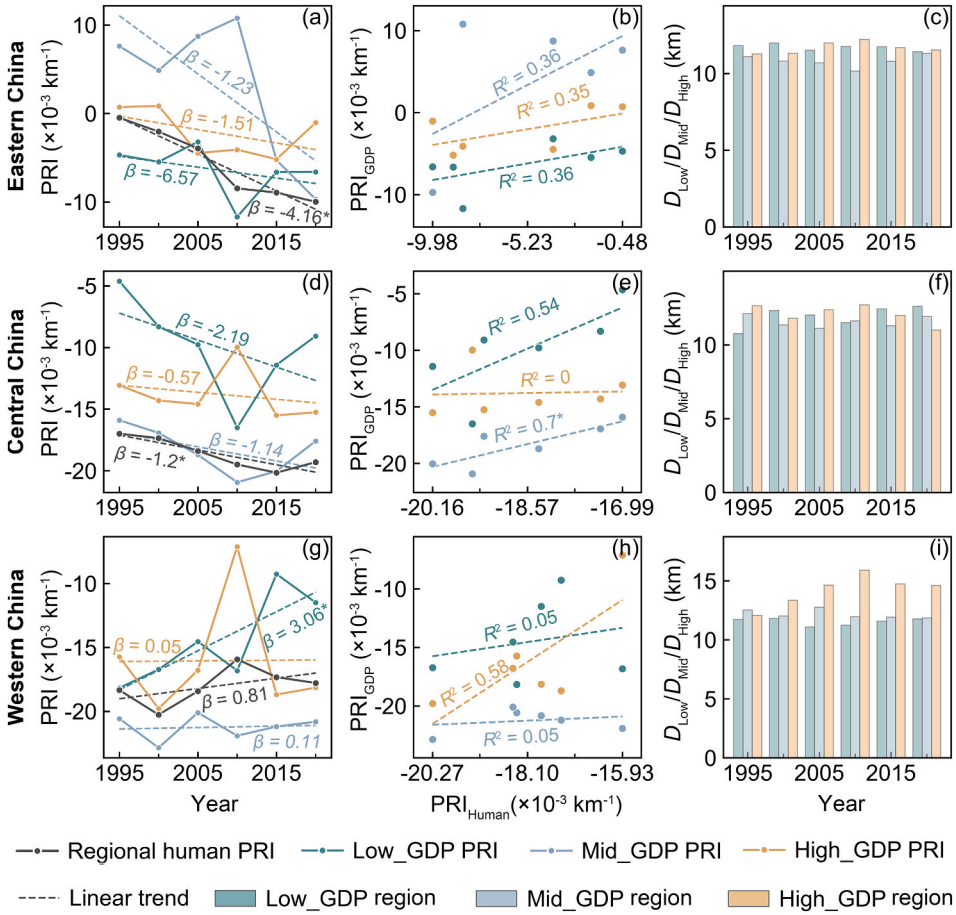


Fig. 7. Regional heterogeneity in PRI across regions grouped by per-capita GDP. (a, d, g) Temporal variations in regional human PRI and PRI for low-, mid-, and high-per-capita-GDP groups in Eastern, Central, and Western China, respectively, from 1995 to 2020. (b, e, h) Covariation between regional human PRI and the PRI values for the three per-capita-GDP groups in Eastern, Central, and Western China, respectively. (c, f, i) Temporal variations in the median distances to rivers for the three per-capita-GDP groups, shown separately for Eastern, Central, and Western China, respectively. These distances are denoted by D_{Low}, D_{Mid}, and D_{High}. Dashed lines indicate linear fits. The temporal trend coefficient β is expressed in units of 10^{-4} km^{-1} yr^{-1}. Asterisks indicate statistical significance: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

4. Discussion

4.1. Stable median population–river distance despite a strengthening proximity gradient

Previous work has shown that river network density is strongly associated with habitability and population density (Wang et al., 2022), indicating that human populations tend to concentrate in river-rich regions. Our results extend this understanding by showing that normalized population density consistently decreased with increasing distance from rivers over the study period (Fig. 2d), which is also consistent with the global findings (Ceola et al., 2015). The median population–river distance (D_{pop}) remained close to 10 km in China. This distance is larger than that reported in previous work, which showed that more than 50% of the world's population lives within 3 km of a surface freshwater body (Kummu et al., 2011). This difference likely reflects our focus on higher-order rivers, namely Strahler orders 4–9, and the exclusion of headwaters, lower-order streams, and lakes. Despite the relatively stable D_{pop}, PRI became increasingly negative over time, indicating a strengthening population-density gradient toward rivers. Our counterfactual analysis further suggests that this pattern reflects both the effect of overall population growth and the spatial redistribution of population relative to rivers, particularly after 1995.

This apparent contrast arose because D_{pop} and PRI measured different dimensions of human–river proximity. D_{pop} is a percentile-based metric that represents the distance within which 50% of the population is located relative to the river network. In contrast, PRI is a gradient-based metric that describes how normalized population density changes across all distance zones. Therefore, a stable D_{pop} does not imply an unchanged spatial relationship between population and rivers. Population can become more concentrated in near-river zones relative to farther zones, leading to a more negative PRI, even when the median population–river distance changes only slightly. Thus, the stable median distance and the decreasing PRI are not contradictory; rather, they reflect different but complementary aspects of human–river proximity.

4.2. Urban–rural settlement patterns and regional economic differences in human–river proximity

Historical path dependence provides a plausible explanation for the stronger river orientation of urban settlements. Many Chinese cities originated along rivers because access to freshwater, navigation, and waste disposal was essential to their early development (Mishra and Saxena, 2024). Subsequent urban expansion often occurred around these established centers, allowing the initial river-oriented spatial pattern to persist even after modern infrastructure reduced direct dependence on nearby rivers (Barrington-Leigh and Millard-Ball, 2015). Rural settlements, by contrast, are generally more dispersed, and their spatial distribution is shaped by the close coupling of residential location with agricultural production, topography, water availability, and transport accessibility (Gong et al., 2022).

However, economic development did not produce a uniform gradient in river proximity. Its effects likely depend on regional industrial structures, infrastructure, and development pathways. Financial and service activities may be less constrained by immediate access to rivers and more strongly attracted to transport hubs, metropolitan regions, and policy-supported development zones (He et al., 2008; Zhang et al., 2013). Conversely, resource-based, manufacturing, and labor-intensive activities may retain stronger links to rivers through water demand, industrial processing, and established transport corridors (Long et al., 2009; Song et al., 2012). The balance among these mechanisms varies across Eastern, Central, and Western China, explaining why per-capita GDP was not associated with a consistent movement either toward or away from rivers.

The urban–rural and per-capita-GDP analyses also represent different dimensions of human–river relationships. Urban and rural settlements were identified from detailed land-use classes and occupy only a limited proportion of the landscape, whereas the GDP classification divided the entire study area into broad economic groups using 1-km gridded data. Therefore, urban settlements constitute only a spatial subset of economically developed areas, and the two classifications should not be expected to produce identical patterns. Water-supply systems, groundwater extraction, inter-basin transfers, flood-control facilities, and diversified transportation networks may further mediate the relationship between development and river proximity. These factors were not quantified because consistent, spatially explicit long-term data were unavailable, and their roles should be examined when suitable infrastructure datasets become available.

4.3. North–south differences in PRI across basins

Most northern basins exhibited persistently negative and generally decreasing PRI values, whereas most southern basins showed positive, increasing, or statistically stable PRI values. This broad contrast may partly reflect differences in water-resource availability. Northern China is generally water-scarce, whereas southern China is relatively water-abundant (Zheng et al., 2022, 2025). Populations in northern basins may therefore remain more dependent on surface water bodies, consistent with evidence that population distribution is more strongly associated with proximity to water in arid regions than in humid regions (Kummu et al., 2011). Differences in industrialization and urbanization may further reinforce this pattern (Bosker and Buringh, 2017; Chen et al., 2024). In contrast, abundant water resources, humid conditions, and coastal development may reduce dependence on major rivers in southern basins (He et al., 2014; Li et al., 2024b; Varis et al., 2026).

However, this north–south contrast was not uniform. PRI shifted from decreasing to increasing around 2006 in the Yangtze River Basin, while its increase accelerated after approximately 2012 in the Pearl River Basin. No significant overall trend was detected in either the Huaihe River Basin or the Southeast Basin. The Southwest Basin was another exception, with GDP showing particularly high relative importance (Fig. 4). PRI declined markedly before an estimated breakpoint around 2009 and then approached stabilization. This pattern may reflect the complex topography of the Southwest Basin, where mountainous terrain and localized karst landscapes constrain habitable and accessible land, concentrating settlements, infrastructure, and economic activities in river valleys and other accessible corridors (Wang et al., 2022; Wang et al., 2024). Consequently, the high relative importance of GDP may partly capture this terrain-constrained concentration of human activities.

These basin-level differences support differentiated water management. Northern basins may prioritize coordination among urban development, river-corridor protection and water-resource allocation, whereas southern and coastal basins may place greater emphasis on runoff variability and climate adaptation. Management strategies should therefore reflect basin-specific environmental conditions and development pathways.

4.4. Changes in PRI trends since the early 21st century

Since the early 21st century, changes in human–river proximity have become increasingly differentiated among river basins. In northern China, the continued concentration of population near rivers may be linked to rapid industrialization, urban growth, and rising water demand under persistent water scarcity (Zheng et al., 2025). In such regions, major rivers remain important for water supply, industrial development, and urban expansion. This interpretation is consistent with the relatively strong influence of socio-economic factors in northern basins and with previous evidence that population distribution, economic development, and industrial expansion are closely associated with river pollution in northern China (Wang et al., 2019).

By contrast, the weakening or stabilization of human–river proximity

in several southern basins may reflect rapid urban expansion beyond traditional river corridors. Abundant water resources, improved water-supply infrastructure, transport development, and strong demand for urban land may have reduced the dependence of settlement growth on proximity to major rivers. Expansion toward inland and coastal areas may also have redistributed population toward locations with lower land costs and greater development space (He et al., 2014; Zong et al., 2021). Overall, the contrasting changes since the early 21st century suggest that water scarcity continued to reinforce river-oriented settlement in northern China, whereas urban expansion and improved infrastructure promoted more dispersed settlement patterns in parts of southern China.

4.5. Uncertainty and limitations

This study acknowledges the following limitations. First, early-period population data introduce uncertainty. Because fine-scale historical population distribution data were unavailable, county-level populations for 1953, 1964, and 1982 were uniformly allocated within each county to generate 1-km gridded data. This assumption may smooth intra-county population heterogeneity and attenuate distance-related population gradients, potentially biasing PRI toward zero where populations were concentrated near rivers, county seats, or urban centers. Early-period PRI estimates should therefore be interpreted with caution, particularly at fine spatial scales. Nevertheless, aggregation at the national and basin scales may reduce the influence of local allocation errors, although it cannot eliminate this uncertainty.

Second, differences between national- and basin-scale driver contributions may partly reflect spatial aggregation and regional heterogeneity. National-scale aggregation can smooth or offset variations in climate, hydrology, population distribution, and socio-economic development among basins, thereby emphasizing drivers associated with broad national trends. In contrast, basin-scale models retain regional conditions and development pathways. In addition, national socio-economic variables were obtained directly from official national statistics, whereas basin-scale variables were estimated from provincial statistics using population-density-weighted averages. These differences in scale and data construction may affect the relative importance of individual predictors. The driver analysis should therefore be interpreted as scale-dependent explanatory evidence rather than direct causal attribution.

Third, our analysis focused on inland river networks and did not account for proximity to lakes and coastlines, which may also influence population distribution. Finally, although the Random Forest models were optimized, their black-box nature limits mechanistic interpretation, and estimates of predictor importance may remain sensitive to the selection, number, and characteristics of the included variables.

5. Conclusions

This study developed the Proximity to Rivers Index (PRI) to quantify changes in population-density gradients relative to major rivers in China. Four main conclusions emerged. First, although the median population–river distance remained close to 10 km, PRI showed a progressively negative trend from 1953 to 2020, indicating strengthening population concentration near rivers. Second, PRI trajectories varied markedly among basins: PRI generally became more negative in northern basins, whereas it became less negative or showed more complex changes in several southern and coastal basins; statistically supported breakpoints were identified in four basins. Third, the relative importance of drivers was scale and region dependent. Socio-economic development was more strongly associated with PRI changes in most northern basins, while natural variability showed greater importance in several southern and coastal basins. In the Southwest Basin, the observed PRI pattern may be consistent with the combined effects of economic development and terrain-related constraints on human settlement and development. Fourth, urban settlements were consistently closer to rivers than rural settlements, whereas relationships between per-capita GDP and river proximity varied among Eastern, Central, and Western China rather than following a uniform national pattern.

These findings demonstrate that human–river proximity is multidimensional and geographically heterogeneous. PRI provides a consistent framework for characterizing these changes and supports basin-specific coordination of urban development, river-corridor protection, water-resource allocation, and climate adaptation.

CRediT authorship contribution statement

Xuejing Zhang: Writing – review & editing, Writing – original draft, Visualization, Methodology, Formal analysis, Data curation, Conceptualization. **Wenfang Cao:** Writing – original draft, Formal analysis, Data curation, Conceptualization. **Dandan Zhao:** Writing – review & editing, Methodology. **Yanlong Guan:** Writing – review & editing, Methodology, Conceptualization. **Pengfei Wang:** Visualization. **Jeffrey J. McDonnell:** Writing – review & editing, Conceptualization. **Junguo Liu:** Supervision, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors thank the anonymous reviewers and editors for their valuable feedback that helped refine this paper. This work was supported by the Programme of Introducing Talents of Discipline to Universities (111 Project) (D25014), the National Foreign Experts Program (Category S) (S20240116), the Henan Provincial Key Laboratory of Hydrosphere and Watershed Water Security, the Henan Province Foreign Scientist Studio (GZS2024013), the Henan Provincial Talent Program (254000510004), and the National Natural Science Foundation of China (42561144298 and 42401379). This work was performed as part of the UNESCO Chair on Stepwise Ecological Restoration of Watersheds, and the IAHS HELPING Working Group on Stepwise Ecological Restoration of Watersheds and on Acceptance of Nature-based Solutions & Their Implementation in Guidelines.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work, the authors used DeepL and ChatGPT to review the spelling and grammar. The authors reviewed and edited the output as needed and take full responsibility for the content of the published article.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhydrol.2026.136710.

Data availability

Data will be made available on request.

References

Alifu, H., Hirabayashi, Y., Imada, Y., Shiogama, H., 2022. Enhancement of river flooding due to global warming. Sci. Rep. 12, 20687. https://doi.org/10.1038/s41598-022-25182-6.

Barrington-Leigh, C., Millard-Ball, A., 2015. A century of sprawl in the United States. Proc. Natl. Acad. Sci. 112, 8244–8249. https://doi.org/10.1073/pnas.1504033112.
Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B (Methodol.) 57, 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.
Bosker, M., Buringh, E., 2017. City seeds: geography and the origins of the European city system. J. Urban Econ. 98, 139–157. https://doi.org/10.1016/j.jue.2015.09.003.
Cao, W., Liu, J., Ceola, S., Mao, G., Macklin, M.G., Montanari, A., Ciais, P., Yao, Y., Tarolli, P., 2023. Landform-driven human reliance on rivers in imperial China. J. Hydrol. 620, 129353. https://doi.org/10.1016/j.jhydrol.2023.129353.
Ceola, S., Laio, F., Montanari, A., 2015. Human-impacted waters: new perspectives from global high-resolution monitoring. Water Resour. Res. 51, 7064–7079. https://doi.org/10.1002/2015WR017482.
Chen, W., Liu, J., 2022. Migration transition in China, 1950-2015. Chinese Political Science Review 7, 181–196. https://doi.org/10.1007/s41111-021-00193-2.
Chen, X., Yu, L., Li, Y., Liu, T., Liu, J., Peng, D., Zhang, X., Fang, C., Gong, P., 2024. China's ongoing rural to urban transformation benefits the population but is not evenly spread. Commun. Earth Environ. 5, 416. https://doi.org/10.1038/s43247-024-01580-8.
Chi, H., Wu, Y., Zheng, H., Zhang, B., Sun, Z., Yan, J., Ren, Y., Guo, L., 2023. Spatial patterns of climate change and associated climate hazards in Northwest China. Sci. Rep. 13, 10418. https://doi.org/10.1038/s41598-023-37349-w.
Davies, R.B., 1987. Hypothesis testing when a nuisance parameter is present only under the alternative. Biometrika 74, 33–43. https://doi.org/10.1093/biomet/74.1.33.
Deng, X., Huang, J., Rozelle, S., Zhang, J., Li, Z., 2015. Impact of urbanization on cultivated land changes in China. Land Use Policy 45, 1–7. https://doi.org/10.1016/j.landusepol.2015.01.007.
Deng, Y., Jiang, W., Ling, Z., Liu, L., Sun, S., 2025. Revealing the driving factors of urban wetland park cooling effects using Random Forest regression and SHAP algorithm. Sustain. Cities Soc. 120, 106151. https://doi.org/10.1016/j.scs.2025.106151.
Di Baldassarre, G., Viglione, A., Carr, G., Kuil, L., Salinas, J.L., Blöschl, G., 2013. Socio-hydrology: conceptualising human-flood interactions. Hydrol. Earth Syst. Sci. 17, 3295–3303. https://doi.org/10.5194/hess-17-3295-2013.
Fang, Y., Ceola, S., Paik, K., McGrath, G., Rao, P.S.C., Montanari, A., Jawitz, J.W., 2018. Globally universal fractal pattern of human settlements in river networks. Earth's Future 6, 1134–1145. https://doi.org/10.1029/2017EF000746.
Fang, Y., Jawitz, J.W., 2019. The evolution of human population distance to water in the USA from 1790 to 2010. Nat. Commun. 10, 430. https://doi.org/10.1038/s41467-019-08366-z.
Gong, J., Jian, Y., Chen, W., Liu, Y., Hu, Y., 2022. Transitions in rural settlements and implications for rural revitalization in Guangdong Province. J. Rural Stud. 93, 359–366. https://doi.org/10.1016/j.jrurstud.2019.10.037.
Gou, J., Miao, C., Ni, J., Sorooshian, S., Duan, Q., Yan, D., Slater, L., Xu, Z., Borthwick, A. G.L., Su, L., Zhang, Q., Wada, Y., 2025. Warming climate and water withdrawals threaten river flow connectivity in China. Proc. Natl. Acad. Sci. 122, e2421046122. https://doi.org/10.1073/pnas.2421046122.
Han, H., Guo, X., Yu, H., 2016. Variable selection using mean decrease accuracy and mean decrease gini based on random forest. In: 2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS). IEEE, pp. 219–224. https://doi.org/10.1109/ICSESS.2016.7883053.
He, C., Wei, Y.D., Xie, X., 2008. Globalization, institutional change, and industrial location: economic transition and industrial concentration in China. Reg. Stud. 42, 923–945. https://doi.org/10.1080/00343400701543272.
He, Q., Bertness, M.D., Bruno, J.F., Li, B., Chen, G., Coverdale, T.C., Altieri, A.H., Bai, J., Sun, T., Pennings, S.C., Liu, J., Ehrlich, P.R., Cui, B., 2014. Economic development and coastal ecosystem change in China. Sci. Rep. 4, 5995. https://doi.org/10.1038/srep05995.
Kuang, X., Liu, J., Scanlon, B.R., Jiao, J.J., Jasechko, S., Lancia, M., Biskaborn, B.K., Wada, Y., Li, H., Zeng, Z., Guo, Z., Yao, Y., Gleeson, T., Nicot, J.-P., Luo, X., Zou, Y., Zheng, C., 2024. The changing nature of groundwater in the global water cycle. Science 383, eadf0630. https://doi.org/10.1126/science.adf0630.
Kummu, M., de Moel, H., Ward, P.J., Varis, O., 2011. How close do we live to water? A global analysis of population distance to freshwater bodies. PLoS One 6, e20578. https://doi.org/10.1371/journal.pone.0020578.
Li, H., Zhang, J., Shan, Y., Wang, G., Tian, Q., Wang, J., Ma, H., 2024a. Study on spatial distribution dispersion evaluation and driving forces of rural settlements in the Yellow River Basin. Land 13, 1181. https://doi.org/10.3390/land13081181.
Li, L., Knapp, J.L.A., Lintern, A., Ng, G.-H.-C., Perdrial, J., Sullivan, P.L., Zhi, W., 2024b. River water quality shaped by land–river connectivity in a changing climate. Nat. Clim. Chang. 14, 225–237. https://doi.org/10.1038/s41558-023-01923-x.
Long, H., Zou, J., Liu, Y., 2009. Differentiation of rural development driven by industrialization and urbanization in eastern coastal China. Habitat Int. 33, 454–462. https://doi.org/10.1016/j.habitatint.2009.03.003.
Macklin, M.G., Lewin, J., 2015. The rivers of civilization. Quat. Sci. Rev. 114, 228–244. https://doi.org/10.1016/j.quascirev.2015.02.004.
Mård, J., Di Baldassarre, G., Mazzoleni, M., 2018. Nighttime light data reveal how flood protection shapes human proximity to rivers. Sci. Adv. 4, eaar5779.
Masrur, A., Dewan, A., Botje, D., Kiselev, G., Murshed, M.M., 2022. Dynamics of human presence and flood-exposure risk in close proximity to Bangladesh's river network: an evaluation with multitemporal satellite imagery. Geocarto Int. 37, 14946–14962. https://doi.org/10.1080/10106049.2022.2093410.
Mishra, R.R., Saxena, S., 2024. Cities and rivers: a symbiotic relationship, in: Managing Urban Rivers. Elsevier, pp. 3–24. https://doi.org/10.1016/B978-0-323-85703-1.00010-9.
Muggeo, V.M.R., 2003. Estimating regression models with unknown break-points. Stat. Med. 22, 3055–3071. https://doi.org/10.1002/sim.1545.
Pizaña, A.D., Alejandro, B.A., 2021. Proximity to the river a mutual role: a phenomenological study. GIIRJ 9, 151–159. https://doi.org/10.17605/OSF.IO/6UMDS.
Ridolfi, E., Mondino, E., Di Baldassarre, G., 2020. Hydrological risk: modeling flood memory and human proximity to rivers. Hydrol. Res. 52, 241–252. https://doi.org/10.2166/nh.2020.195.
Salman, H.A., Kalakech, A., Steiti, A., 2024. Random forest algorithm overview. Babylonian J. Mach. Learn. 2024, 69–79. https://doi.org/10.58496/BJML/2024/007.
Sivapalan, M., Savenije, H.H.G., Blöschl, G., 2012. Socio-hydrology: a new science of people and water. Hydrol. Process. 26, 1270–1276. https://doi.org/10.1002/hyp.8426.
Song, H., Thisse, J.-F., Zhu, X., 2012. Urbanization and/or rural industrialization in China. Reg. Sci. Urban Econ. 42, 126–134. https://doi.org/10.1016/j.regsciurbeco.2011.08.003.
Tang, Z., Wang, P., Li, Y., Sheng, Y., Wang, B., Popovych, N., Hu, T., 2024. Contributions of climate change and urbanization to urban flood hazard changes in China's 293 major cities since 1980. J. Environ. Manage. 353, 120113. https://doi.org/10.1016/j.jenvman.2024.120113.
Varis, O., Taka, M., Kummu, M., 2026. Mapping world's coastal population facing water-related risks. Geogr. Sustain. 7, 100411. https://doi.org/10.1016/j.geosus.2026.100411.
Wang, E., Li, Q., Hu, H., Peng, F., Zhang, P., Li, J., 2019. Spatial characteristics and influencing factors of river pollution in China. Water Environ. Res. 91, 351–363. https://doi.org/10.1002/wer.1044.
Wang, F., Wang, Y., Zhang, K., Hu, M., Weng, Q., Zhang, H., 2021. Spatial heterogeneity modeling of water quality based on random forest regression and model interpretation. Environ. Res. 202, 111660. https://doi.org/10.1016/j.envres.2021.111660.
Wang, N., Sun, F., Koutsoyiannis, D., Iliopoulou, T., Wang, T., Wang, H., Liu, W., Sargentis, G.-F., Dimitriadis, P., 2023. How can changes in the human-flood distance mitigate flood fatalities and displacements? Geophys. Res. Lett. 50, e2023GL105064. https://doi.org/10.1029/2023GL105064.
Wang, Y., Borthwick, A.G.L., Ni, J., 2022. Human affinity for rivers. River 1, 4–14. https://doi.org/10.1002/rvr2.12.
Wang, Y., Deng, Q., Yang, H., Liu, H., Yang, F., Zhao, Y., 2024. Spatial-temporal differentiation and influencing factors of rural settlements in mountainous areas: an example of Liangshan Yi Autonomous Prefecture, Southwestern China. J. Mt. Sci. 21, 218–235. https://doi.org/10.1007/s11629-023-8191-6.
Wang, Z., Xie, J., 2025. Water constraint mitigation and agricultural productivity: Evidence from the China's South-to-North water diversion project. Agric. Water Manag. 314, 109511. https://doi.org/10.1016/j.agwat.2025.109511.
Wu, H., Gao, L., Song, D., Yang, Y., Xu, C., Yang, X., 2022. A dataset of district/county-level population distribution of China's six national censuses. Chin. Sci. Data. 7, 1–7. https://doi.org/10.11922/csdata.2021.0034.zh.
Xu, H., Chen, H., Wang, H., 2022. Future changes in precipitation extremes across China based on CMIP6 models. Int. J. Climatol. 42, 635–651. https://doi.org/10.1002/joc.7264.
Zhang, L., Njepu, A., Xia, X., 2021. Minimum cost solution to residential energy-water nexus through rainwater harvesting and greywater recycling. J. Clean. Prod. 298, 126742. https://doi.org/10.1016/j.jclepro.2021.126742.
Zhang, X., Huang, P., Sun, L., Wang, Z., 2013. Spatial evolution and locational determinants of high-tech industries in Beijing. Chin. Geogr. Sci. 23, 249–260. https://doi.org/10.1007/s11769-012-0581-6.
Zheng, J., Wang, H., Liu, B., 2022. Impact of the long-term precipitation and land use changes on runoff variations in a humid subtropical river basin of China. J. Hydrol. Reg. Stud. 42, 101136. https://doi.org/10.1016/j.ejrh.2022.101136.
Zheng, X., Ma, C., Lv, H., Wan, F., Wang, H., 2025. Evolution and drivers of water use in Northern China. J. Hydrol. Reg. Stud. 59, 102419. https://doi.org/10.1016/j.ejrh.2025.102419.
Zong, S., Hu, Y., Zhang, Y., Wang, W., 2021. Identification of land use conflicts in China's coastal zones: from the perspective of ecological security. Ocean Coast. Manag. 213, 105841. https://doi.org/10.1016/j.ocecoaman.2021.105841.