Supplementary Information for

Regime transition identification for water utilization in river basin developments

Shuang Song, Shuai Wang, Bojie Fu, Xutong Wu (complete author list)

Shuai Wang.
E-mail: shuaiwang@bnu.edu.cn

This PDF file includes:

Supplementary text
Figs. S1 to S12
Tables S1 to S2
SI References

Supporting Information Text

This supplementary document consists of 5 *Methods* sections, 9 *SI Appendix* figures and 2 *SI Appendix* tables. Firstly, we introduced our study area, the Yellow River Basin, and its subdivisions in the section Methods S1. Then, we give a detailed description of used datasets and analysis their uncertainties in the section Methods S2. Further, we introduced how to harmonized different datasets with various spatial and temporal scales in Methods S3. Then, the indicators corresponding to the IWRU are introduced in the section Methods S4. Finally, we proposed a sensitivity analysis in the section Methods S5, to test the robustness of our results.

Methods S1. Definition of study area

Region divisions of Yellow River Basin. The study area is the Yellow River Basin (YRB, see Fig. S1-A), which had experienced the most intense water exploitation and the most dramatic shifts of management regime in China. According to the Yellow River Conservancy Commission (YRCC), an administrative government directly under the Ministry of Water Resources at the basin level, the upper, middle and lower reaches of the Yellow River can be distinguished by the characteristic of river. However, there is another scheme suggesting that the upstream only refers to river source areas with little human disturbance and high water retention capacity. Anyhow, since the socio-economic and natural conditions were both considered in this study, we integrated the two schemes above and divided the Yellow River Basin into four regions, which can be distinguished by three important hydrological control stations (see Fig. S1-B). Previous studies have also shown that such a division is valid when both social water use and the natural conditions of the basin are considered, as the regions exhibit strong heterogeneity among themselves (see Fig. S1-C) (1):

- **Source Region (SR):** Over 50% of natural runoff was produced in this region. The most ecology function here is water conservation, as sparsely populated and less economically developed.
- **Upper Region (UR):** With the highest per capita irrigated land area, there are numbers of large irrigation lands in this region. However, because of backward production methods, efficiency of irrigation are used to be very low.
- **Middle Region (MR):** Crossing Loess Plateau, famous rich-sand area, Yellow River loads most of its sediments here with the highest soil erosion risk. To reverse this situation, the grain for green project changed the water utilization here strikingly.
- **Lower Region (LR):** With dense population and the traditional agricultural trajectory, lower region used to be the largest water use region. However, as the industrial transformation going, proportion of agriculture keeps decreasing, but LR is still the largest water use region in each aspect.

In general, there are inter-regional differences in the economic layout, distribution of water resources, distribution of water consumption, and population distribution of the Yellow River Basin (see Fig. S2). On the basis of these fundamental differences, social development and watershed management continue to influence and reshape their changes, making the Yellow River Basin the world's most intimately connected and dramatically changing large river basin. Thus, as a case study for analysing the evolution of the human-water relationship, it possesses typicality.

Water resources make an irreplaceable contribution to the development of society in the YRB. Firstly, each aspect of economic development highly depended on water resources, since the most water use patterns are positively correlated with the region's key economic indicators (see Fig. S2-A, B and D).

Changes brought from human activities on the YRB. Humans are constantly modifying the water cycle processes in the watersheds as society develops.

- Firstly, the YRB has been subject to strong intervention by human activities since ancient times, while the last 60 years have seen the most dramatic changes. Under human influence, the Yellow River's surface runoff, sediment transport, and human water consumption patterns have all undergone multiple regime shifts in the last 60 years
- Secondly, landscapes in the YRB have significantly altered by human activities, which can both change the natural and the social water cycle. For an example, grain for green project has produced a significant change of landscapes in the middle region (MR) of YRB. With the addition forest in an erosion-prone area, water use patterns in the area and surface runoff patterns in the middle and lower reaches have been radically altered.
- Thirdly, a series of management practices have promoted in order to govern the YRB. Since the establishment of the Yellow River Water Conservancy Commission in XX, the agency has continued to reform and expand, eventually forming a basin management agency with unified coordination, scheduling and regulatory functions. Since the promulgation of the Water Law in xxx, the Ministry of Water Resources, the YRCC and other relevant agencies have issued a series of policies to carry out comprehensive river basin management under the guidance of national laws (see *SI Appendix* Table S2).

Historic and recent river basin management practices have strong impacts on water utilization.

Methods S2. Detailed information on datasets

Multiple types of dataset were used in this study (see Table S1).

Statistical datasets. For statistics, we used GDP data, water resources uses data extracted from the 2nd National Water Resources Assessment Program (2) and statistical yearbook data of YRB. GDP data from the China Macro data in the Wind database, which firstly aggregated from annual reports of the provinces. Water resources use dataset was published by Zhou et al. (2), which records water utilization in different sectors along with social-economic situations in perfects level. This dataset was mainly extracted by 2nd National Water Resources Assessment Program launched in 2002, led by the National Development and Reform Commission and the Ministry of Water Resources (see ref (1) and http://www.mwr.gov.cn/english/publs/ for more details). Since then, the statistics from the survey using the same criteria have been supplemented and harmonized to the 2013 administrative divisions. The data covers a total of subcategories of water use under four broad categories: agriculture, industry, urban and rural, but does not distinguish between surface water and groundwater. There is uncertainty at the county scale for each disaggregated water use sector, but because the data are corrected for statistical information using the water balance method, the data are adequate for the regional scale and the four broad water use categories used in this study. Finally, in order to make a simple distinction between the proportion of surface water and groundwater extraction where needed, we use basin-scale water resources annual reporting data. Yearbook data of YRB documenting surface water resources and groundwater resource extraction in each watershed and province.

Hydrological datasets. For hydrological datasets, reservoirs data and a measured runoff data were used in this study. The reservoir dataset were collected by Wang et al. (1), which introduced includes the major new reservoirs built in the Yellow River Basin since 1949. Of these, we consider the reservoirs marked as pivot projects by the YRCC to be more important, as they are directly involved in water resources management in the basin (http://www.yrcc.gov.cn/hhyl/sngc/). Annual runoff data derived from hydrological station measurements.

Political datasets. For the policy data set, we select the laws and policies related to the Yellow River basin promulgated and implemented by departments at and above the river Basin level (such as Yellow River Conservancy Commission) in the Comprehensive Planning for the Yellow River Basin (such as national ministries and commissions) (3). In addition, we also obtain the organizational evolution of the Yellow River Water Conservancy Commission as a watershed management unit from the book Organizational Evolution of the Yellow River Water Conservancy Commission to reflect the implementation of watershed management measures (4).

Methods S3. Harmonization of datasets

Due to the wide sources of our data set and the different spatial scales, we need to reconcile the data on the regional scale (SR, UR, MR and DR), and have different processing methods according to the spatial scale of the data.

- 1. Watersheds: Due to the fact that the subdivision of a basin is determined by hydrological subdivision, hydrological annual data and measured runoff data can be directly divided according to their corresponding hydrological stations.

- 2. Perfects: At the scale of cities and counties, we respectively calculate the area occupied by cities and counties in different watershed districts, and select the largest region. If this area accounts for more than 95% of the area of the county, the city is considered to belong to this region, that is:

  $$S_{ij} = MAX(S_i j/S_i)$$

  where i refers a certain perfect and j refers a region within YRB, i.e. SR, UR, MR, or DR. S_i refers the area of perfect i, and S_{ij} refers intersecting area between perfect i and region j. We define perfect i belongs to region j if their intersecting area over 95% of S_i, i.e.:

  $$MAX(S_{ij}) > 0.95 * S_i$$

- 3. Province: Data at the provincial scale (i.e., GDP data of different industries) are more used for qualitative description, and do not carry out horizontal proportion with data at the county scale, so the accuracy of the data is no longer important. We have adopted a rough treatment. That is, according to the major provinces contained in different regions, determine which region the data of that province is merged into. Referring to the traditional division practice, we made the following division:

  - SR: Qinghai Gansu and Sichuan,
  - UR: Ningxia and Neimeng,
  - MR: Shanxi and Shaanxi,
  - DR: Shandong, Hebei and Henan.

Finally, when we process the location data (such as the location data of reservoir construction), we judge the province it belongs to according to its location, and then process it according to the regional classification method on the provincial scale. After the above processing, all the data can be well coordinated at the regional scale.

Methods S4. Indicators of IWRU

The Integrated Water Resources Utilization (IWRU) index of water resources utilization considers the three basic dimensions of human utilization of water resources at the watershed scale, corresponding to the three questions of “how much water”, “how to use” and “for whom, for where” respectively. In order for these dimensions to be quantified, we found or invented a total of three metrics corresponding to the trend indicating their change: SFV Index, Non-provisioning Share and Configuration Entropy Metric.

Stress: SFV-index. There is how much water can be used in a certain basin? How many stresses has societies put on the water resources of the basin? Answering these questions are crucial for water utilization. Various metrics, therefore, proposed for water stress (e.g. water scarcity, water stresses index, scarcity-flexibility-variability index), where the aspect of human impact are increasingly valued. Among of them, by taking changes of water flexibility and variability into account, the scarcity-flexibility-variability (SFV) index focus more on dynamic responses to water resources in developing perspective, which considered a valid metric of temporal changes in water stresses (5).

To apply this method, we need to combine three metrics following:

First for scarcity, $A_{i,j}$ is the total water consumption as a proportion of regional multi-year average runoff volume, in year j and region i:

$$A_{i,j} = \frac{WU_{i,j}}{R_{i,avg}}$$

Second for flexibility, $B_{i,j}$ is the inflexible water use $WU_{inflexible}$ (i.e. for thermal power plants or humans and livestock) as a proportion of average multi-year runoff, in year i and region j:

$$B_{i,j} = \frac{WU_{inflexible}}{R_{i,avg}}$$

Finally for variability, the capacity of the reservoir and the positive effects of storage on natural runoff fluctuations are also considered.

$$C_i = C1_i * (1 - C2_i)$$

$$C1_{i,j} = \frac{R_{i,std}}{R_{i,avg}}$$

$$C2_i = \frac{RC_i}{R_{i,avg}},\ ifRC < R_{i,avg}$$

$$C2_i = 1,\ ifRC >= R_{i,avg}$$

In all the equations above, $R_{i,avg}$ is the average runoff in region i, RC_i is the total storage capacities of reservoirs in the region i, $R_{i,std}$ is the standard deviation of runoff in the region i.

Finally, assuming three metrics (scarcity, flexibility and variability) have the same weights, we can calculate SFV index after normalizing them:

$$V = \frac{A_{normalize} + B_{normalize} + C_{normalize}}{3}$$

$$a = \frac{1}{V_{max} - V_{min}};$$

$$b = \frac{1}{V_{min} - V_{max}} * V_{min}$$

$$SFV = a * V + b$$

Priority: Non-provisioning share. How to use water resources? Today, the world’s non-provisioning form of water supply to human beings are increasing more than 20 times faster growth, including industry, services and other ways, the human society is transformed from “consumption of water” to “make use of water”. Here, we use Non-Provisioning Shares (NPS) of used water resources as a metric to indicate priority of water utility:

$$NPS_i = \frac{WU_{non-pro,i}}{WU_{pro,i} + WU_{non-pro,i}}$$

Where i refers a certain region, or the whole basin, i.e: $I_P = NPS_{basin}$ In this study, we consider water for livestock, water for rural or urban residents and water for agriculture as provisioning water. Correspondingly, non-provisioning water is often industrial water, mainly energy water. Extensive studies have shown that there is a wide connection between water, food and energy, which is a manifestation of water supply tendency, in which provisioning water is often the key to support the production of agricultural products, while non-provisioning water is more involved in industrial and other social processes. For China as a whole, if only agricultural production is considered, China is the recipient country of virtual water trade. But if industrial consumption is taken into account, China is exporting virtual water. However, as a river basin with high degree of water resource scarcity, the trade-off between the two has always been treated with greater value; thus, the adoption of this ratio as an indicator can simply, directly and effectively reflect the change of priority with water.

Configuration: Configuration Entropy Metric. Where to use water resources? The configuration of water resources has always been a difficult problem, requiring trade-offs between upstream and downstream and even different water sectors. However, there is still a lack of comprehensive indicators to quantify. Here we designed a "Configuration Entropy Metric (CEM)", which can simply and directly express the configuration of water allocation (which tends to be more evenly distributed or more focused on a certain region and sector). We imitate the idea of information entropy in physics –When the distribution of information (here is water resource) is more equal, its entropy increases. The CEM is a metrics for different regions (SR, UR, MR, DR) and different sectors (agriculture, industry, domestic and services) in configurations of Water Use (WU), calculated by:

$$CEM = \sum_{i=1}^{n} -log(p_i) * p_i$$

Where i represents the unit (a region or a sector) in configuration of water utilization, p_i represents the proportion of water used by the units. For different regions, when water resources are completely evenly used, (i.e. for any region i, $p_i = 1/n$), the maximum value of CEM is:

$$CEM_{max} = \sum_{i=1}^{n} -log(p_i) * p_i = -log(\frac{1}{n})$$

I_C, an indicator to measure the configuration dimensions of water resources utilization, is composed of three parts:

$$I_C = \frac{CEM_r * CEM_s}{CEM_{rs}}$$

- Configuration between regions CEM_r:

$$CEM_r = \sum_{i=1}^{4} -log(p_i) * p_i$$

  where p_i is the proportion of region i in water resource usage (WU_i), in the total water resource consumption in all regions ($i = 1$ to 4 represents SR, UR, MR, or, DR respectively), i.e.:

$$p_i = \frac{WU_i}{\sum_{i=1}^{4} WU_i}$$

- Configuration between sectors CEM_s:

$$CEM_s = \sum_{j=1}^{3} -log(p_j) * p_j$$

  where p_j is the proportion of sector j in water resources usage (WU_j), in the total water consumption in all sectors ($j = 1$ to 3 represents agriculture, industry and domestic & services respectively), i.e.:

$$p_j = \frac{WU_j}{\sum_{j=1}^{3} WU_j}$$

- Division of sectors' function between regions CEM_{rs}

$$CEM_{rs} = \frac{1}{3} * \sum_{j=1}^{3} \sum_{i=1}^{4} -log(p_{ij} * p_{ij})$$

  where p_{ij} refers water use proportion of region i in the sector j, in the total water consumptions of all regions in this sector, i.e.:

$$p_{ij} = \frac{WU_ij}{\sum_{i=1}^{4} WU_ij}$$

The indicator I_C (Configuration of water utilization), therefore, not only take into account the distribution between different regions and industries, but also considers the complicated professional division with the development of society.

Sensitive analysis

A common criticism of breakpoint detection methods is that the results can be manipulated by adjusting the use of the methods, so we conducted a sensitivity analysis of the breakpoint recognition process to demonstrate that our results are robust.

Firstly, we performed three different normalization methods in construction of IWRU for breakpoint detection:

- Min-max normalization:
$$normalized(X) = \frac{X - X_{min}}{X_{max} - X_{min}}$$
- Mean-max normalization:
$$normalized(X) = \frac{X - X_{mean}}{X_{max} - X_{min}}$$
- Z-score normalization:
$$normalized(X) = \frac{X - X_{mean}}{X_{std}}$$

Where X is a time series and X_{mean}, X_{std}, X_{min}, X_{max} are mean, standard deviation, minimum, maximum of it. No matter which of the above standardized methods we use in combining the indicators of the three dimensions, and our detection results for IWRU breakpoints are the same.

Another place where there is room for manipulation is the choice of significance threshold (i.e. threshold of the p-value) in the Pettitt Change points detection method. Here, we test 1000 possible p-values between 10^{-6} and 0.01, and most of the results give the same detection in breakpoints. It has not given another breakpoint before the threshold is near to 10^{-2} (see *SI Appendix* Fig. S12: A). Not only that, when three breakpoints are detected, the previous two breakpoints still exist (1978 and 1994), and the trend difference of the new breakpoints is not obvious (see *SI Appendix* Fig. S12: B and C).

Taken together, the results of three periods in this study with the two breakpoints detected (1978 and 1994) are robust.

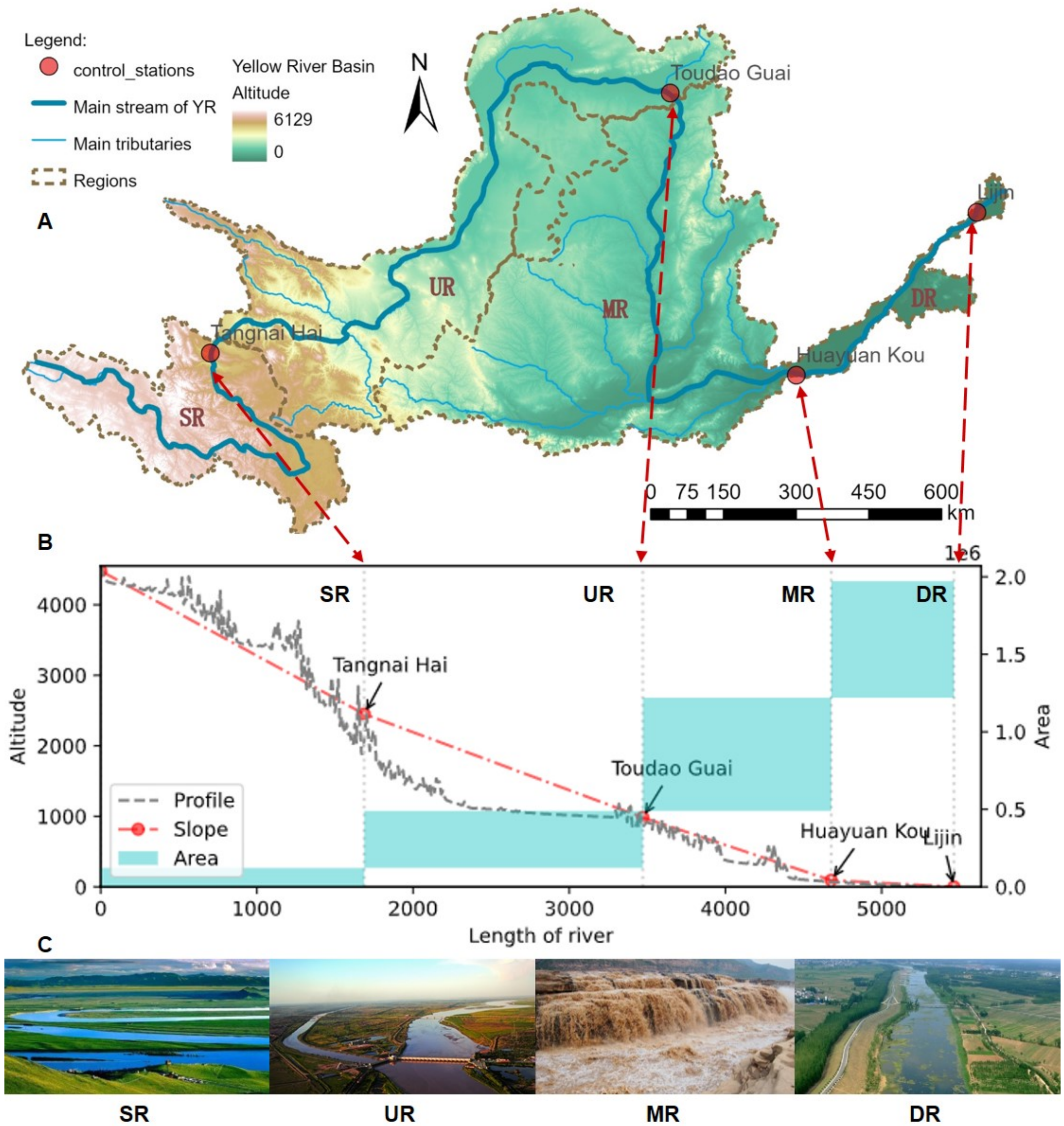

Fig. S1. The study area. **A.** Diagram of the YRB and the subdivision of the basin (SR: Source Region, UR: Upper Region, MR: Middle Region, DR: Downstream region). **B.** Profile of main channel of the Yellow River. **C.** Typical landscapes in different regions in the YRB.

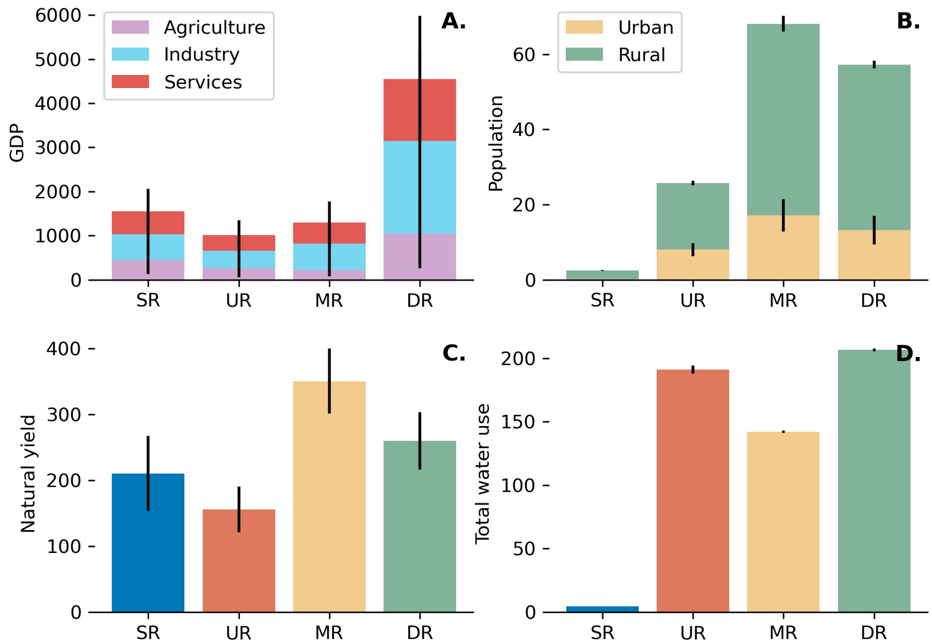

Fig. S2. Basic socio-economic status and water resource conditions of different regions in the YRB, the average from 1980 to 2000 shown. **A.** Proportion of GDP in different industries (primary, secondary or tertiary industry) in different regions. **B.** Population of urban or rural area in different regions. **C.** Natural water yield in different regions. **D.** Total water use in different regions.

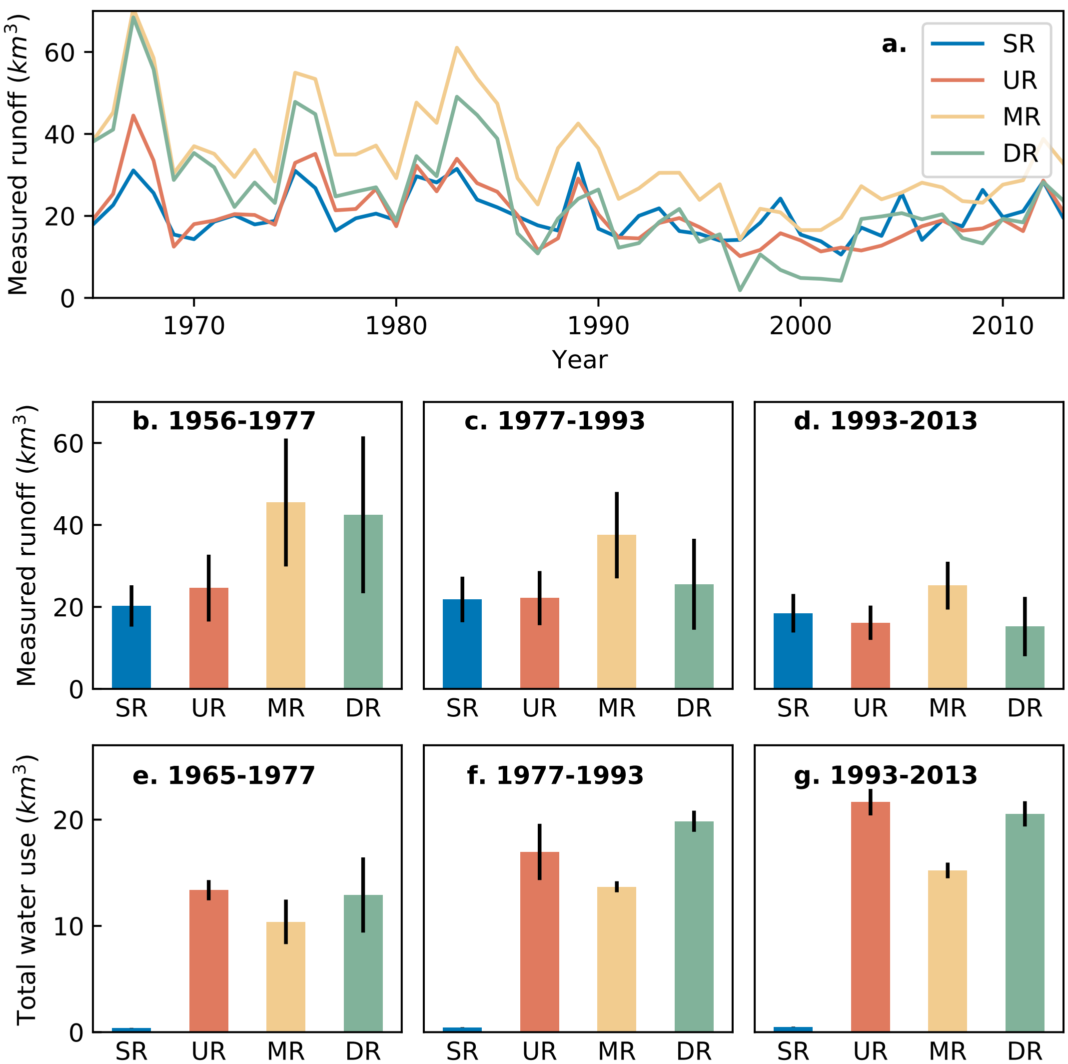

Fig. S3. Water resources in different regions. **A,** changing trend of measured runoff, **B, C and D** average measured runoff within different periods. **E, F and G** average total water consumptions within different periods.

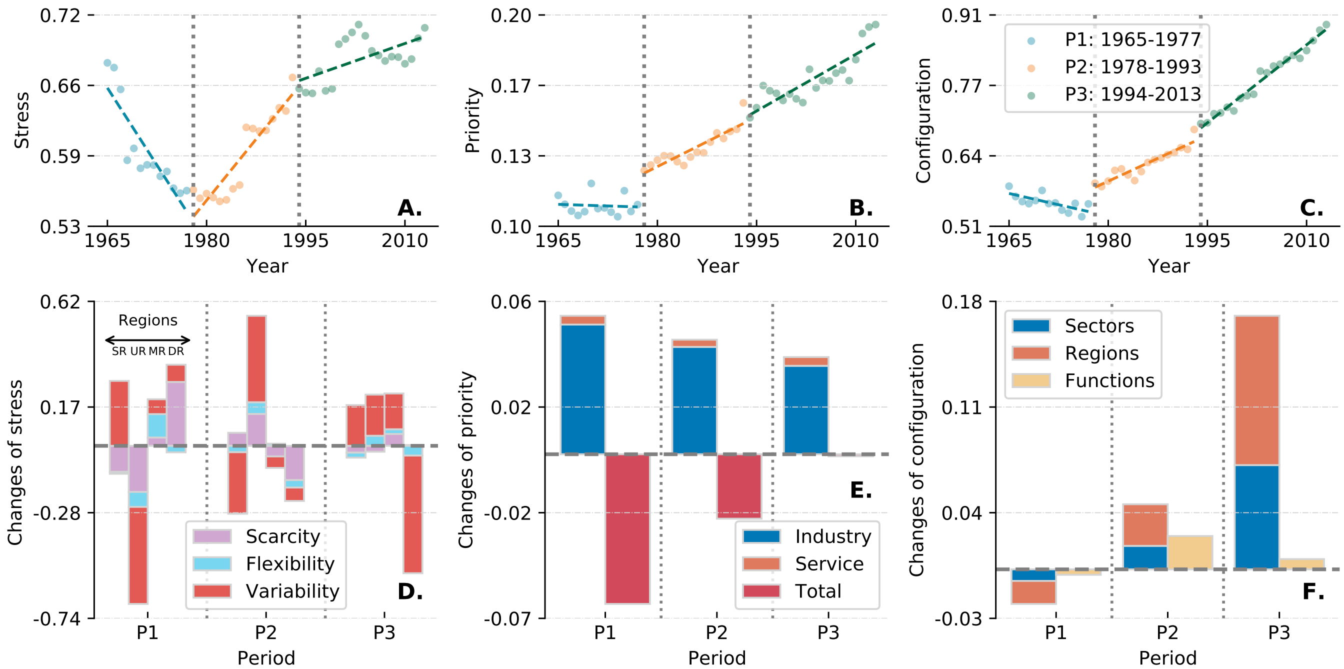

Fig. S4. Changes in different dimensions of water resources utilization regimes and their main contributors. **A,** changes of water utilization stress, indicated by unstandardized scarcity-flexibility-variability water stresses index (SFV-index, see *Methods* and *SI Appendix* Method S4). **B,** changes of water utilization priority, indicated by non-provisioning water shares (see *SI Appendix* Methods S4). **C,** changes of water utilization configuration, indicated by unstandardized distribution information entropy index (*Methods* and *SI Appendix* Method S4). **D,** Main impact factors to water utilization stresses in each period or region, and their change contributions to unstandardized SFV-index. **E,** Main impact water uses to water utilization priority, and their change contributions to non-provisioning water shares. **F,** Main impact factors to water utilization configurations, and their change contributions of related unstandardized distribution information entropy index (*Methods* and *SI Appendix* Method S4).

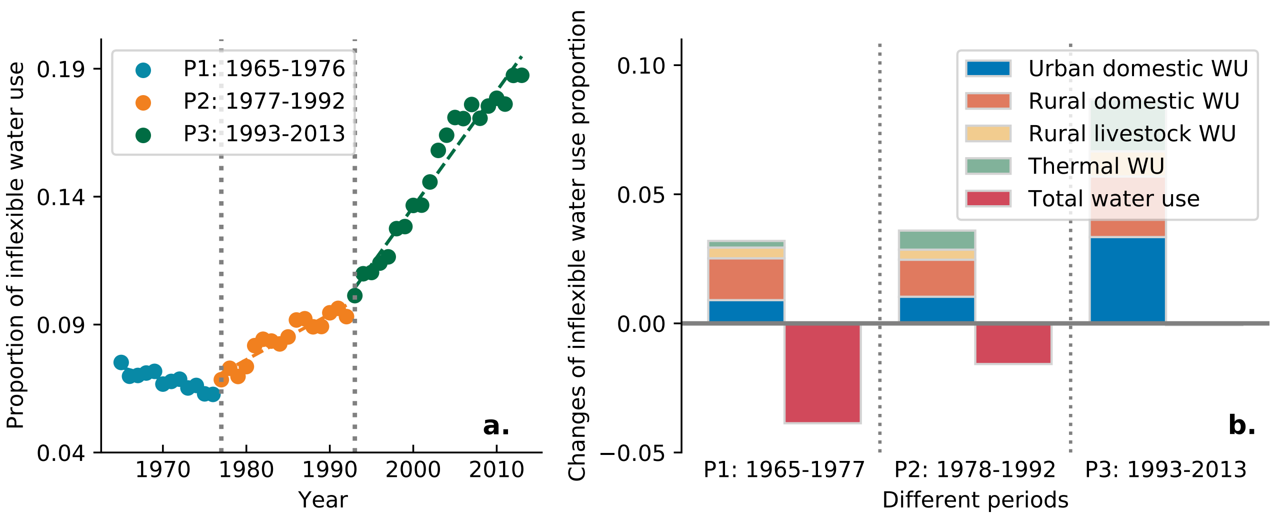

Fig. S5. Flexibility and its change contributions of water utilization. **A,** Proportion of inflexible water uses. **B,** Change contributions of inflexible water uses.

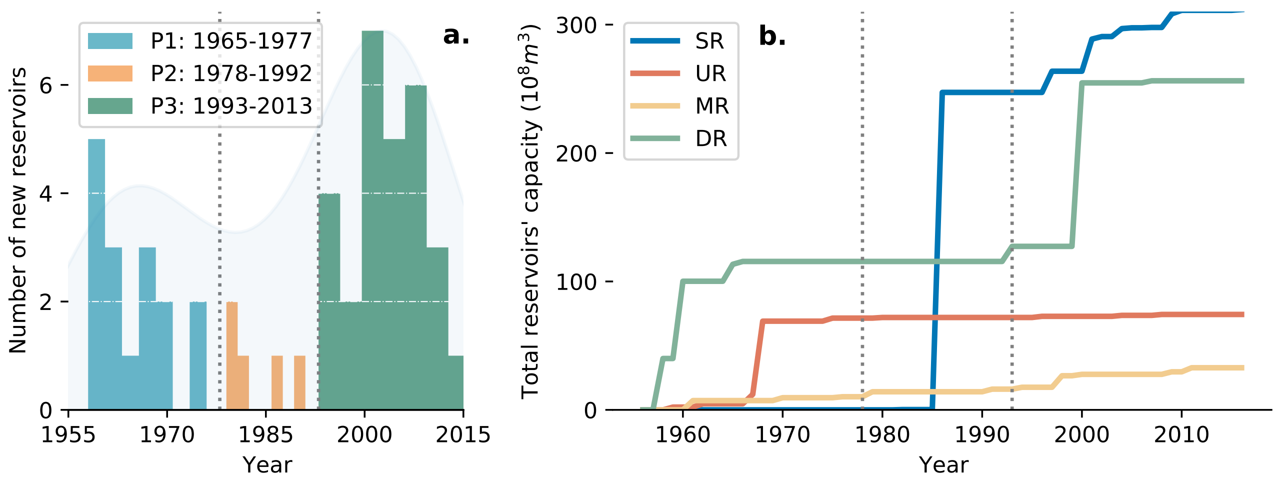

Fig. S6. Reservoirs and their storage. **A,** number of new reservoirs in different periods. **B,** total storage capacities in different regions.

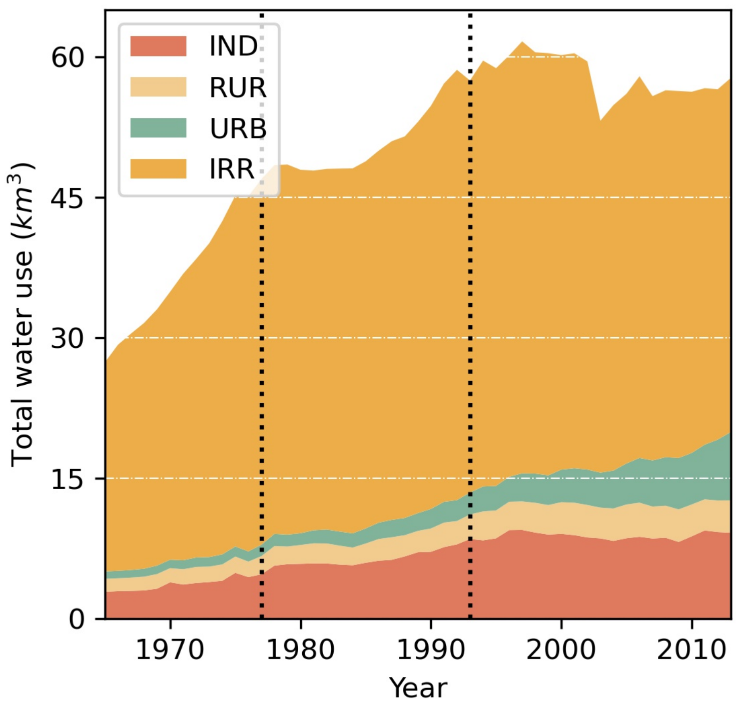

Fig. S7. Proportions of total water use between the different sectors

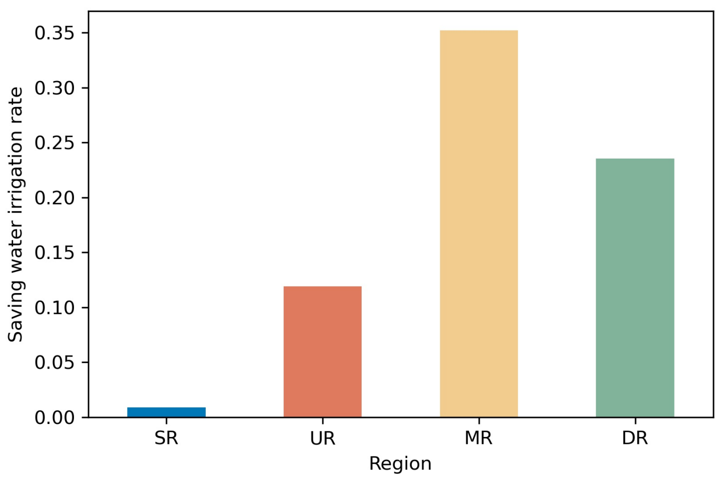

Fig. S8. Saving water irrigated area ratio in 2013.

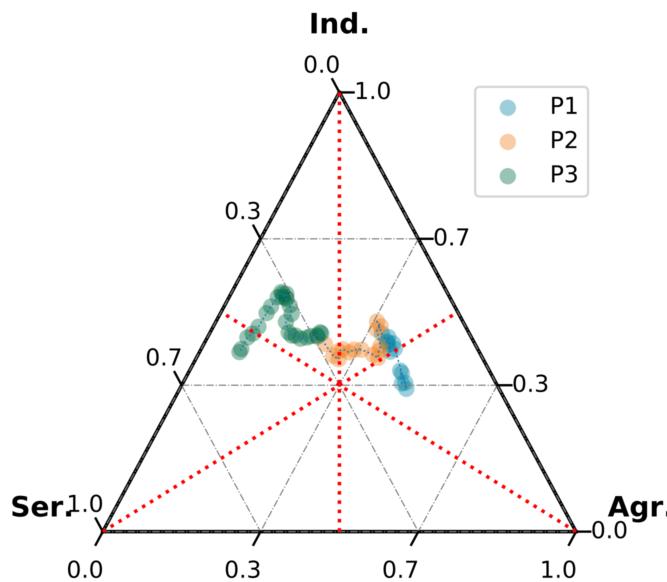

Fig. S9. Net contribution of agriculture, industry and services to GDP in different years.

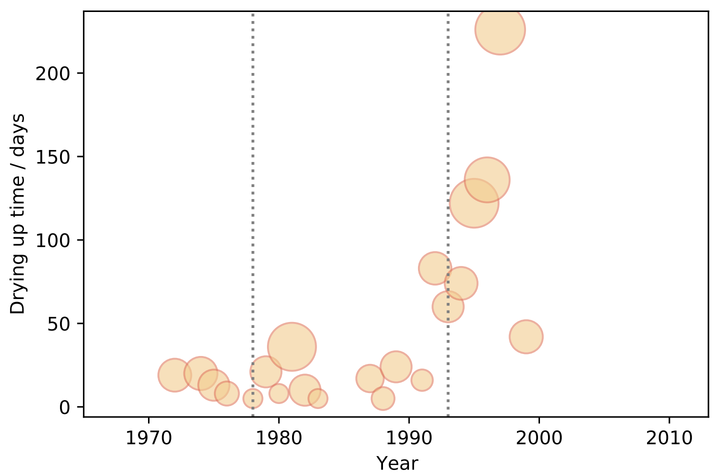

Fig. S10. Runoff outages records, size of points indicates length of drying up river segments.

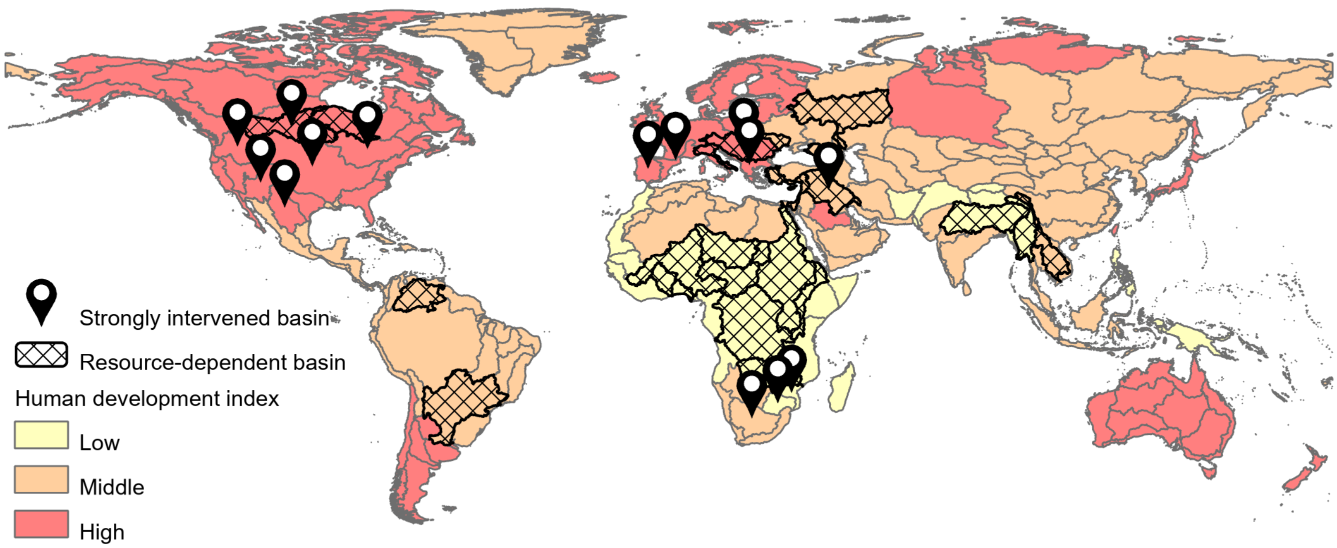

Fig. S11. The Human Development Index of the world's great river basins, and the major problems facing some basins. Overdependence on water resources for economic development in some basins, and high levels of human modification in others that have disrupted ecosystem structure. Data resources: Human Development Index in each basin refers from (**?**), strongly intervened basins and resource-dependent basins are come from Transboundary River Basins report by United Nations Environment Programme http://geftwap.org/publications/river-basins-technical-report (**?**).

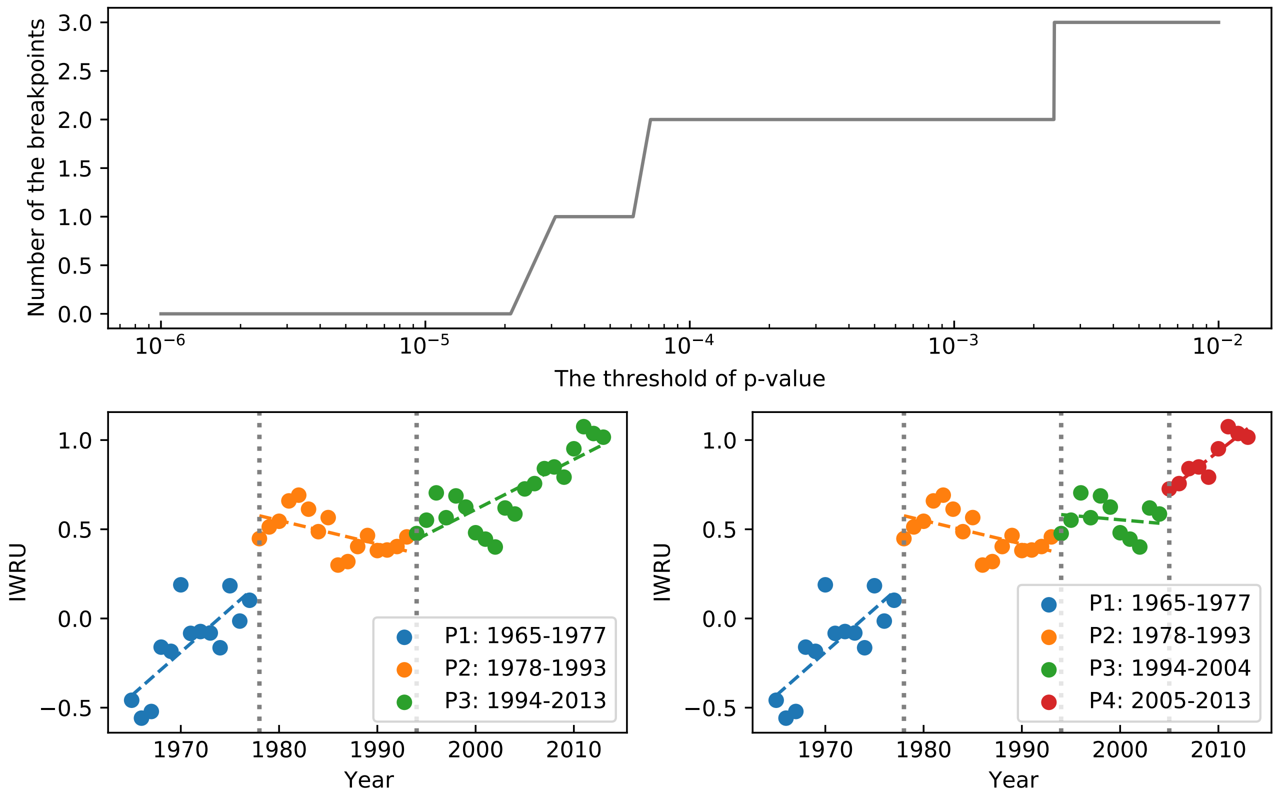

Fig. S12. Sensitivity analysis of the threshold of p-values. **A.** number of breakpoints in different p-values. **B.** A typical result of two breakpoints. **C.** A typical result of three breakpoints.

Table S1. Used datasets and their sources.

Dataset	Type	Spatial scale	Time limit	Source
1. Water use	Statistical	Perfects	1965-2013	2nd National Water Resources Assessment Program
2. GDP	Statistical	Province	1949-2019	Wind database
3. Groundwater and surface water use	Statistical	Watershed	2003-2019	Yearbooks of YRB by the YRCC.
4. Reservoirs	Hydrological	Location	1949-2015	Wang et al. (1)
5. Measured runoff	Hydrological	Location	1949-2019	Measured data from YRCC
6. Laws	Political	Documents	1949-2013	YRCC (3)
7. History of YRCC	Political	Documents	1949-2002	YRCC (4)

Table S2. Policies and regulations above YRB level which affected the whole basin in water utilization [a]

Name	Year	Agency
1. Water Law of PRC	1988	National People's Congress of the PRC
2. Water Law of PRC -revised 1	2009	National People's Congress of the PRC
3. Water Law of PRC -revised 2	2016	National People's Congress of the PRC
4. Regulations on the Administration of Water Drawing Licences and The Collection of water resource fees	2006	State Council of the PRC
5. Regulations on the Administration of Water Drawing Licences and The Collection of water resource fees -revised 1	2017	State Council of the PRC
6. Regulations on the Allocation of Water in the Yellow River	2006	State Council of the PRC
7. Yellow River water supply distribution scheme	1987	State Council of the PRC
8. Measures for the Administration of Water Drawing Permits	2008	Ministry of Water Resources of the PRC
9. Measures for the Administration of Water Drawing Permits -revised 1	2015	Ministry of Water Resources of the PRC
10. Measures for the Administration of Water Drawing Permits -revised 2	2017	Ministry of Water Resources of the PRC
11. Regulations on the Allocation of Water in the Yellow River	2006	State Council of the PRC
12. Annual distribution of available water supply of the Yellow River and main stream water dispatching scheme	1998	Ministry of Water Resources of the PRC
13. The Yellow River water dispatching management measures	1998	Ministry of Water Resources
14. Measures for the Implementation of the Yellow River Water Rights Conversion Management	2004	Ministry of Water Resources
15. Regulations on the Administration of Water Drawing Licences and The Collection of water resource fees	2006	State Council of the PRC
16. Measures for the implementation of the water drawing Permit system	1993	
State Council of the PRC		
17. Measures for the demonstration and management of water resources in construction projects	2002	Ministry of Water Resources of the PRC
18. Implementation Opinions on the Reform of Water Conservancy Project Management System	2006	State Council of the PRC

[a] If a policy was proposed by multiple legacies, we only choose the highest one to show.

References

1. Y Wang, W Zhao, S Wang, X Feng, Y Liu, Yellow River water rebalanced by human regulation. *Sci. Reports* **9**, 9707 (2019).
2. F Zhou, et al., Deceleration of China's human water use and its key drivers. *Proc. Natl. Acad. Sci.*, 201909902 (2020).
3. YRC Commission, *Yellow River Basin Comprehensive Plan (2012-2030).* (Yellow River Water Conservancy Press, Zhengzhou, China), First edition, (2013).
4. YR Archives, *Organizational History of the Yellow River Conservancy Commission.* (Yellow River Water Conservancy Press), (2004).
5. Y Qin, et al., Flexibility and intensity of global water use. *Nat. Sustain.* **2**, 515–523 (2019).