Supplementary Information to the paper:

'Spatial pattern analysis of water-related ecosystem service and the evaluation of grassland carrying capacity of the Heihe River Basin under land use change'

Part 1. InVEST models

The InVEST (Version.3.6.0) suite of tools has been developed to enable decision makers to assess trade-offs within and among ecosystem services and to compare the consequences of different future change scenarios, for example those related to land use or climate (Sharp et al., 2016). This study used the Water Yield model (for water provision service), the Sediment Delivery Ratio model (for sediment retention service), and the Nutrient Delivery Ratio model (for water purification service) to evaluate the corresponding ecosystem services in Heihe River Basin (HRB).

Part 1.1 Water yield (WY) model

The annual water yield for pixel *i* on LULC$_j$, Y_{ij} (mm/yr), is estimated based on average annual precipitation and the Budyko curve as follows:

$$Y_{ij} = (1 - \frac{AET_i}{P_i}) \cdot P_i$$

where AET_{ij} (mm/yr) is the actual annual evapotranspiration for pixel *i* on LULC$_j$, and P_i (mm/yr) is the annual precipitation for pixel *i*.

For vegetated LULC, the evapotranspiration portion of the water balance, $\frac{AET_i}{P_i}$, is based on an expression of the Budyko curve proposed by Fu et al. (1981) and Zhang et al. (2004):

$$\frac{AET_i}{P_i} = 1 + \frac{PET_i}{P_i} - [1 + (\frac{PET_i}{P_i})^{\omega}]^{\frac{1}{\omega}}$$

where PET_{ij} is the potential evapotranspiration and ω_i is a non-physical parameter that characterizes the natural climate-soil properties.

Potential evapotranspiration, PET_{ij}, is defined as:

$$PET_i = K_{c,j} \cdot ET_{0,i}$$

where $ET_{0,i}$ is the reference evapotranspiration from pixel *i* and k_{ij} is the vegetation evapotranspiration coefficient associated with the pixel *i* on LULC$_j$

$$\omega_i = Z \cdot \frac{AWC_i}{P_i} + 1.25$$

$$AWC_i = Min\,(Rest\_layer\_depth_i, Root\_depth_i) \cdot PAWC_i$$

where ω_i is a non-physical parameter that characterizes the natural climate-soil properties; *Z* is a dimensionless constant, ranging from 1 to 30, that captures the local precipitation pattern and hydrogeological characteristics; AWC_i (mm) is the volumetric plant-available water content; the 1.25 term is the minimum value of ω_i (Donohue et al., 2012); k_{ij} is the evapotranspiration coefficient for pixel *i* on LULC$_j$; ET_0 (mm/yr) is the reference evapotranspiration for pixel *i*; and *PAWC* (mm) is the plant-available water capacity.

For non-vegetated LULC (e.g., water, construction land), the actual annual evapotranspiration is computed directly from the reference evapotranspiration and has an upper limit defined by the precipitation:

$$AET_{ij} = Min\,(k_{ij} \cdot ET_{0_i}, P_i)$$

where P_i (mm/yr) is the annual precipitation for pixel *i*.

Part 1.2 Sediment Delivery Ratio (SDR) model

The InVEST sediment delivery model maps overland sediment generation and delivery to the stream. The sediment export from a pixel *i*, $Export_i$ (ton·ha^{-1}·yr^{-1}), and the total sediment export of the evaluated area, $Export_{tot}$(ton·ha^{-1}·yr^{-1}), are defined by the equations:

$$Export_i = usle_i \cdot SDR_i$$

$$Export_{tot} = \sum_i Export_i$$

The amount of annual soil loss on pixel *i*, $usle_i$, is determined by the revised universal soil loss equation:

$$usle_i = R_i \cdot K_i \cdot LS_i \cdot C_i \cdot P_i$$

where R_i (MJ·mm·(ha·hr)$^{-1}$) is the rainfall erosivity, K_i (ton·ha·hr·(MJ·ha·mm)$^{-1}$) is the soil erodibility, LS_i is the slope length-gradient factor, C_i is the crop-management factor, and P_i is the support practice factor.

The soil retention is computed by the model as follows:

$$Soilretention = R_i \cdot K_i \cdot LS_i \cdot (1 - C_i \cdot P_i) SDR_i$$

It represents the avoided soil loss by the current land use compared with bare soil.

The connectivity index *IC* is defined as:

$$IC = \log_{10}(\frac{D_{up}}{D_{dn}})$$

D_{up} is the upslope component, defined as:

$$D_{up} = \overline{C}\,\overline{S}\sqrt{A}$$

where $\overline{C}$ is the average C factor of the upslope contributing area, $\overline{S}$ (m/m) is the average slope gradient of the upslope contributing area, and *A* (m^2) is the upslope contributing area. D_{dn} is the downslope component, defined as:

$$D_{dn} = \sum_i \frac{d_i}{C_i S_i}$$

where C_i and S_i are the *C* factor and the slope gradient on pixel *i* and d_i (m) is the length of the flow path along the pixel *i*.

The *SDR* for a pixel *i*, SDR_i, is derived from the connectivity index *IC* as follows:

$$SDR_i = \frac{SDR_{\max}}{1 + \exp(\frac{IC_0 - IC_i}{k})}$$

where SDR_{max} is the maximum theoretical *SDR* and IC_0 and *k* are calibration parameters that define the shape of the *SDR-IC* relationship.

Part 1.3 Nutrient Delivery Ratio (NDR) model

The InVEST Nutrient Delivery Ratio model maps nutrient sources from watersheds and nutrient transport to streams. Nutrient export from each pixel is calculated based on the product of the load and the *NDR* :

$$X_{\exp ort_i} = load_{surf,i} \times NDR_{surf,i} + load_{subs,i} \times NDR_{subs,i}$$

$$X_{\exp ort_{tot}} = \sum_i X_{\exp ort_i}$$

Each pixel's load is modified to account for the local runoff potential, which can be divided into surface and subsurface runoff. The ratio between these two types of nutrient sources is given by the parameter *proportion_subsurface*$_i$; therefore, the *load* (kg·ha^{-1}·yr^{-1}) for pixel *i* is defined as:

$$load_{surf,i} = (1 - proportion\_subsurface_i) \times \mathrm{mod}ified\_load_i$$

$$load_{subsurf,i} = proportion\_subsurface_i \times \mathrm{mod}ified\_load_i$$

$$\mathrm{mod}ified\_load_i = load_i \times RPI_i$$

$$RPI_i = \frac{RP_i}{RP_a}$$

where RPI_i is the runoff potential index for pixel$_i$, RP_i is the nutrient runoff proxy for runoff on pixel *i*, and RP_a is the average *RP* over the entire area.

The delivery ratios ($NDR_{surf,i}$ and $NDR_{subs,i}$) are computed based on the concept of the nutrient delivery ratio.

(1) Surface NDR

The surface NDR is the product of a delivery factor, representing the ability of downstream pixels to transport nutrients without retention, and a topographic index, representing the position on the landscape. For pixel *i*:

$$NDR_{surf,i} = NDR_{0,i}(1 + \exp(\frac{IC_i - IC_0}{k}))^{-1}$$

where IC_0 and *k* are calibration parameters, IC_i is a topographic index, and $NDR_{0,i}$ is the proportion of nutrient that is not retained by downstream pixels (irrespective of the position of the pixel on the landscape).

$$NDR_{0,i} = 1 - eff_i^{'}$$

$$eff_i^{'} = \begin{cases} eff_{LULC_j} \cdot (1 - s_i) & if\ down_i\ is\ a\ stream\ pixel \\ eff_{down_i}^{'} \cdot s_i + eff_{LULC_j} \cdot (1 - s_i) & if\ eff_{LULC_j} > eff_{down_i}^{'} \\ eff_{down_i}^{'} & otherwise \end{cases}$$

where $eff_i^{'}$ is retention efficiency for pixel i, eff_{LULC_j} is the maximum retention efficiency that $LULC_j$ can reach, $eff_{down_i}^{'}$ is the effective downstream retention on the pixel directly downstream from pixel *i*, and S_i is the step factor defined as:

$$s_i = \exp(\frac{1 - 5l_{i_{down}}}{l_{LULC_i}})$$

where $l_{i_{down}}$ is the length of the flow path from pixel *i* to its downstream neighbor, l_{LULC_i} is the LULC retention length of the land cover type on pixel *i*.

IC is the index of connectivity:

$$IC = \log_{10}(\frac{D_{up}}{D_{dn}})$$

$$D_{up} = \overline{S}\sqrt{A}, \quad D_{dn} = \sum_i \frac{d_i}{S_i}$$

where $\overline{S}$ (m/m) is the average slope gradient of the upslope contributing area, A (m^2) is the upslope contributing area, and d_i (m) is the length of the flow path along the pixel *i*.

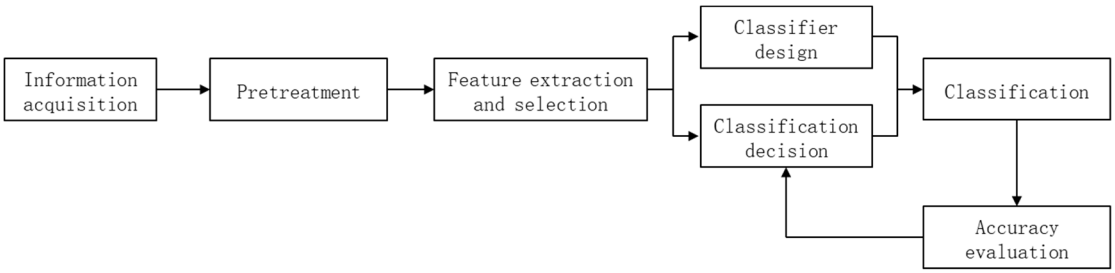
(2) Subsurface NDR

$$NDR_{subs,i} = 1 - eff_{subs}(1 - e^{\frac{-5l_i}{l_{subs}}})$$

where eff_{subs} is the maximum nutrient retention efficiency that can be reached through subsurface flow, l_i is the distance from the pixel to the stream, and l_{subs} is the subsurface flow retention length (i.e. the distance after which it can be assumed that soil retains nutrient at its maximum capacity).

Part 2. Classification algorithm used for land use/land cover

Remote sensing mapping of land use classification by computer generally involves the following five steps[3] and the flow chart is as follows:



1) Data collection and preprocessing include geometric and radiometric correction, noise reduction filtering, signal enhancement, feature extraction and selection, data compression, etc.

2) Selection of training sample area for unsupervised classification, sample area needs to be selected to assist classification; For supervised classification, the training sample area is used to extract various characteristic parameters to simulate various types.

3) The object element is analyzed, and the classification algorithm is used to classify any pixel into the most appropriate class according to the eigenvalue of the pixel. Pixel features can be a spectral reflection, texture features of adjacent pixels and geometric features of their location (such as height, slope, aspect, etc.).

4) The post-processing of classification results includes various filtering, re-classification of analysis results, the geometric transformation of classification results according to the requirements of map projection, decoration of classification map, etc. Generally, it is difficult to achieve a good classification effect only by the original remote sensing image. In order to achieve the desired effect, different image processing technologies need to be adopted for different areas, different terrain and different soil vegetation types, supplemented by some non-remote sensing information.

5) Evaluate the classification accuracy. Through random sampling and ground investigation, the classification results are compared with the known accurate types to obtain the objective resolution of the classification map.

Table S1. LULC class definitions from the system of remote sensing investigation and assessment of the 18-year change in ecological environments in China (2000-2018) used in the maps for HRB (2000 and 2018).

Code	Class/ Value	Descriptions
1	Forest	All areas characterized by tree cover (natural or semi-natural woody vegetation, generally greater than 6 m tall); tree canopy accounts for 25% to 100% of the cover.
2	Grass-land	Areas dominated by graminoid or herbaceous vegetation, generally greater than 80% of total vegetation. These areas are not subject to intensive management such as tilling, but can be utilized for grazing. Areas of grasses, legumes, or grass-legume mixtures planted for livestock grazing or the production of seed or hay crops, typically on a perennial cycle. Pasture/hay vegetation accounts for greater than 20% of total vegetation.
3	Cultivated	Includes cultivated crops – Cultivated crops are described as areas used for the production of annual crops, such as corn, soybeans, vegetables, tobacco, and cotton, and also perennial woody crops such as orchards and vineyards. This class also includes all actively tilled land.
4	Developed	Includes developed open spaces with a mixture of some constructed materials, but mostly vegetation in the form of lawn grasses such as large-lot single-family housing units, parks, golf courses, and vegetation planted in developed settings for recreation, erosion control, or aesthetic purposes. Also included are lands of low, medium, and high intensity development with a mixture of constructed materials and vegetation, such as single-family housing units, multifamily housing units, and areas of retail, commercial, and industrial uses.
5	Wetland	All areas of open water, generally with less than 25% vegetation or soil cover. Includes woody wetlands and herbaceous wetlands – Areas where forest or shrub land vegetation accounts for greater than 20% of vegetative cover and the soil or substrate is periodically saturated with or covered with water. This class also includes areas where perennial herbaceous vegetation accounts for greater than 80 percent of vegetative cover and the soil or substrate is periodically saturated with or covered with water.
6	Bare	Areas of bedrock, pavement, scarps, talus, slides, glacial debris, strip mines, gravel pits, and other accumulations of earthen material. Generally, vegetation accounts for less than 15% of total cover.

Table S2. Data requirements for the InVEST model (Water yield model=WY; Nutrient delivery ratio model=NDR; Sediment delivery ratio model=SDR).

Data	Type	Data Source	Note	Related Model
Digital Elevation Model (DEM)	Raster	Geospatial Data Cloud, http://www.gscloud.cn	Resolution is 30m×30m	NDR, SDR
Annual average precipitation	Raster	China Meteorological Data Center, http://data.cma.cn/	Interpolated based on annual data, resolution is 30m×30m	WY, NDR, SDR
Reference evapotranspiration	Raster	MODIS Global Evapotranspiration Project (MOD16) http://www.ntsg.umt.edu/project/mod16	Resolution is 30m×30m	WY
Plant-available water content	Raster	Environmental and Ecological Science Data Center for West China, http://westdc.westgis.ac.cn/	Calculated based on the soil data (Harmonized World Soil Database) according to the model proposed by Zhou (2005), resolution is 30m×30m	WY
Land use / land cover	Raster	Resource and environment data cloud platform, Chinese Academy of Sciences, http://www.resdc.cn/	LULC of year 2000, and 2018, including forest, grassland, developed land, cultivated land, wetland, and bare land, resolution is 30m×30m	WY, NDR, SDR
Depth to root restricting layer	Raster	Environmental and Ecological Science Data Center for West China, http://westdc.westgis.ac.cn/	Derived from the soil data (Harmonized World Soil Database), resolution is 1km×1km	WY
Watersheds	Shapefile	Geospatial Data Cloud, http://www.gscloud.cn	A shapefile determined by DEM raster using ArcGIS tool	WY, NDR, SDR
Rainfall erosivity index	Raster	China Meteorological Data Center, http://data.cma.cn/	Calculated based on precipitation according to the model proposed by Zhang and Fu (2003), resolution is 30m×30m	SDR
Soil erodibility	Raster	China Meteorological Data Center, http://data.cma.cn/	Calculated based on precipitation according to the model proposed by Cao et al. (2015), resolution is 30m×30m	SDR
Biophysical data	.CSV file	Literature ([7]; [8]), and the InVEST user's guide ([9])	Including attributes of each LULC, K_c (the plant evapotranspiration coefficient), load of nutrients, efficiency of nutrient retention, etc.	WY, NDR, SDR

Table S3. Key parameters used in the present study

Parameters	Description	Computation
k_{ij}	Evapotranspiration coefficient for each pixel	Defined according to the literature and InVEST user's guide
$load\_n$ $load\_p$	Load of nitrogen and phosphorus for each LULC	Defined according to the literature data[10,11]
eff_n eff_p	The maximum retention efficiency of nitrogen and phosphorus for each LULC, varying between 0 and 1.	Defined according to the literature data [10,11]
$SDR\_\max$	The maximum theoretical SDR	Defined as 0.8 according to the InVEST user's guide

Table S4. Critical parameter settings in the biophysical table

Biophysical table (Wang et al. 2016, Sharp et al. 2016, China Soil Map Based on Harmonized World Soil Database 2015). Kc: the plant evapotranspiration coefficient for each LULC class; root_depth: the maximum root depth for vegetated land use classes in mm ; usle_c and usle_p: two parameters in the SDR model (equation 19); sedret_eff: the maximum soil retention efficiency for each LULC ; load_n: the original nitrogen loading for each LULC in kg/ha•yr; load_p: the original phosphorus loading for each LULC in kg/ha•yr; eff_n: the maximum nitrogen retention efficiency for each LULC; eff_p: the maximum phosphorus retention efficiency for each LULC; LULC_vegetation: a flag to distinguish bare land from vegetated land, 1 for vegetated land and 0 for others.

LULC_desc	lucode	Kc	root_depth	usle_c	usle_p	sedret_eff	load_n	eff_n	load_p	eff_p	LULC_veg	crit_len_p	crit_len_n	load_subsurface_n	load_subsurface_p	proportion_subsurface_n
Forest	1	1	7000	0.003	0.2	0.6	1.8	0.8	0.011	0.8	1	150	150	0.18	0.0011	0
Grassland	2	0.65	1700	0.008	0.2	0.4	11	0.4	1.5	0.4	1	150	150	1.1	0.15	0
Cultivated	3	0.6	1500	0.5	0.4	0.25	11	0.25	3	0.25	1	150	150	1.1	0.3	0.3
Developed	4	0.3	500	0.001	0.001	0.05	8	0.05	3.4	0.05	0	150	150	0.8	0.34	0
Wetland	5	1	1000	0.001	0.001	0.8	0.001	0.05	0.001	0.05	0	150	150	0.0001	0.0001	0
Bare	6	0.2	10	0.25	0.01	0.2	4	0.05	0.001	0.05	0	150	150	0.4	0.0001	0

Table S5. The area changes of different reaches in the Heihe River Basin in 2000 and 2018

reaches	lulc	2000 area(km^2)	2018 area(km^2)	changes area(km^2)
the upper reaches	Forest	1037.07	1400.16	363.09
	Grassland	6341.87	5919.52	-422.36
	Cultivated	43.42	18.77	-24.65
	Developed	17.21	11.27	-5.93
	Wetland	909.23	230.75	-678.48
	Bare	2163.32	2932.11	768.80
the middle reaches	Forest	2925.02	3533.19	608.17
	Grassland	20862.34	13084.32	-7778.03
	Cultivated	5394.54	5567.08	172.54
	Developed	610.87	676.71	65.84
	Wetland	560.26	971.58	411.32
	Bare	9829.24	16350.79	6521.55
the lower reaches	Forest	1367.96	924.42	-443.54
	Grassland	2552.18	9632.52	7080.34
	Cultivated	628.50	772.68	144.17
	Developed	128.59	179.36	50.77
	Wetland	110.80	386.56	275.76
	Bare	72808.19	65718.64	-7089.54

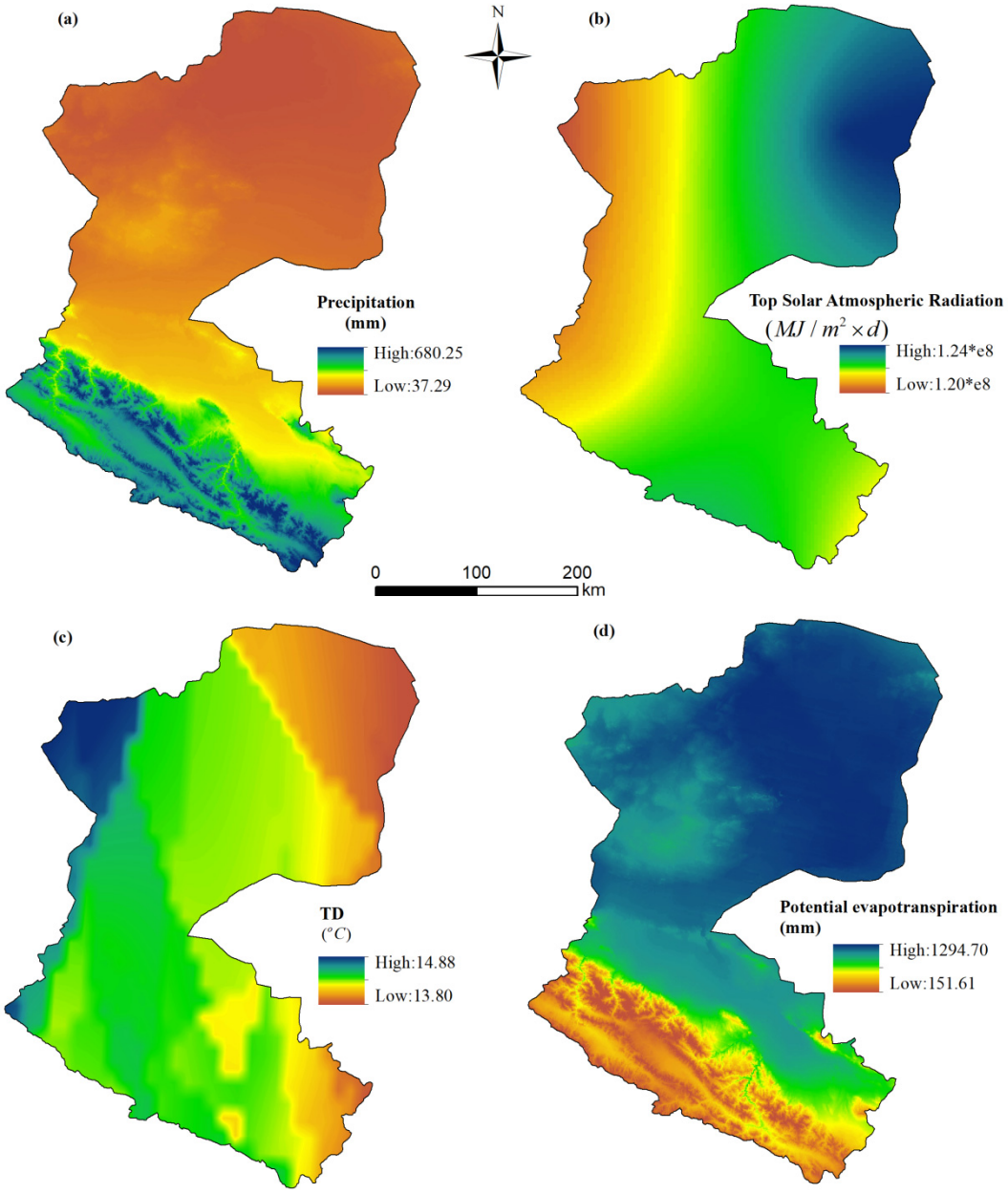


Figure S1. Spatial distribution in the Heihe River Basin of (**a**) precipitation; (**b**) top solar atmospheric radiation; (c) temperature difference (TD) between the mean maximum temperature and the mean minimum temperature; (d) potential evapotranspiration.

Table S6. Soil erosion transition matrix of the Heihe River Basin from 2000 to 2018. (unit:km^2)

		2018						
		I	II	III	IV	V	VI	Total
2000	I	114398.76	3430.44	366.26	128.81	92.44	90.91	118507.62
	II	3401.08	2993.40	424.03	126.49	73.08	43.17	7061.25
	III	383.69	411.08	175.02	73.45	49.77	30.89	1123.89
	IV	149.68	123.20	71.40	41.67	34.02	24.89	444.85
	V	119.90	70.57	48.33	33.19	36.41	33.40	341.80
	VI	138.18	42.07	30.32	23.77	32.69	66.83	333.85
	Total	118591.27	7070.77	1115.36	427.37	318.41	290.08	127813.27

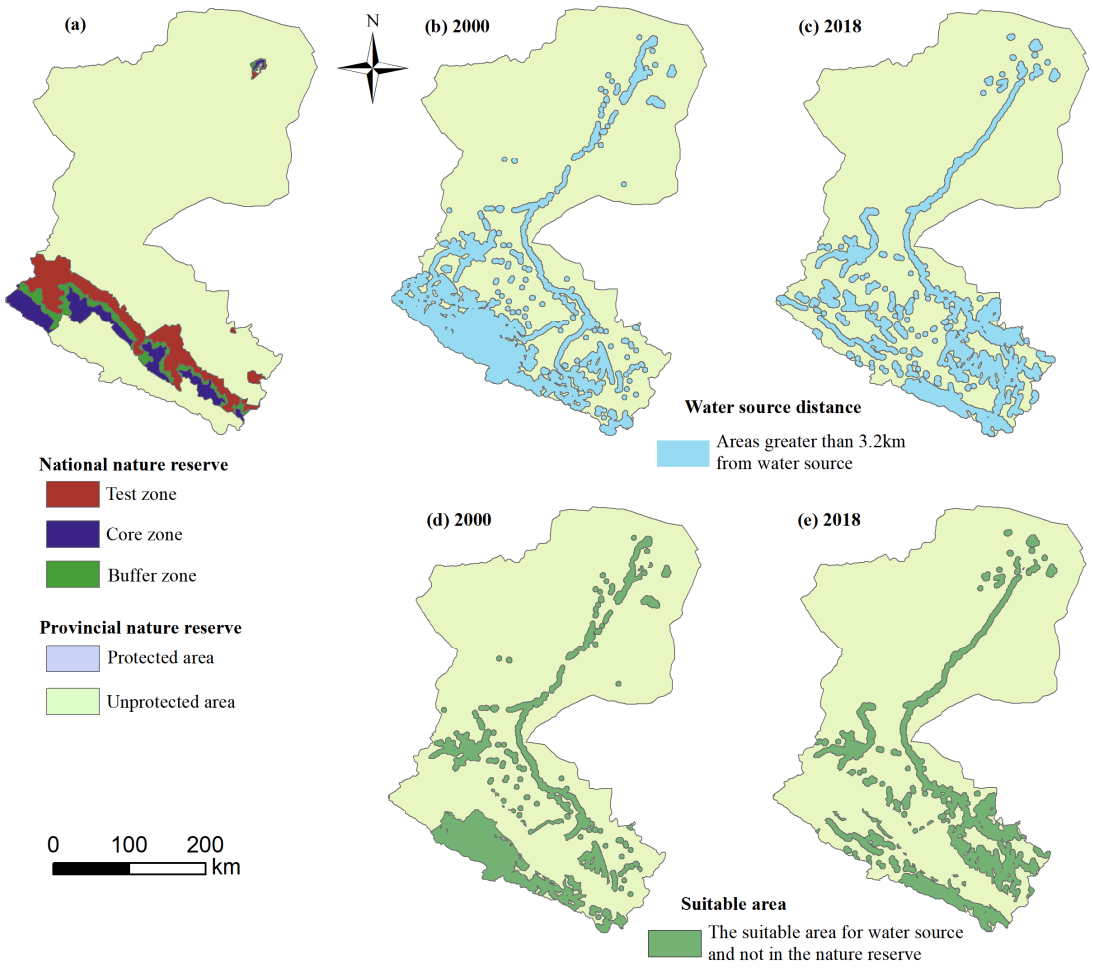


Figure S2. Spatial distribution in the Heihe River Basin (a) nature reserve; (b) water source distance in 2000; (c) water source distance in 2018; (d) the suitable area for water source and not in the nature reserve in 2000; (e) the suitable area for water source and not in the nature reserve in 2018.

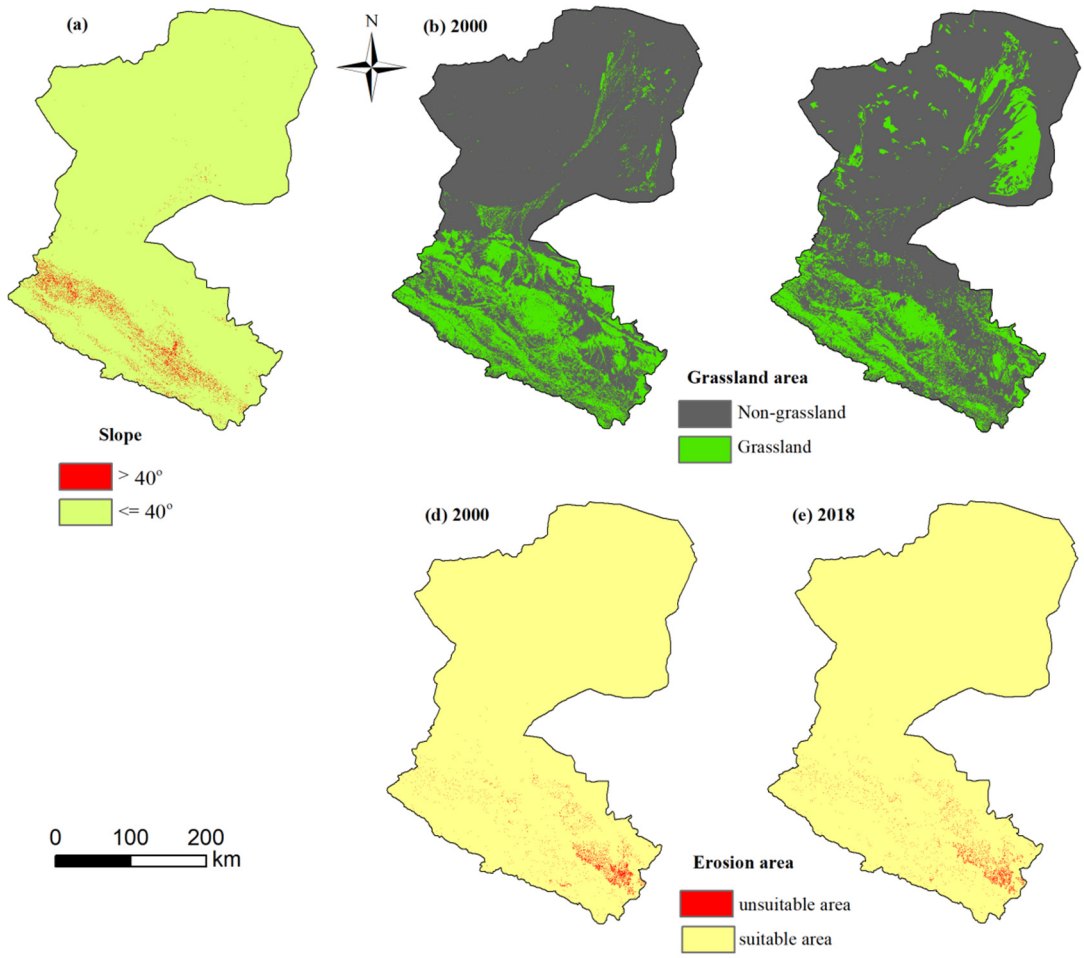


Figure S3. Spatial distribution in the Heihe River Basin (a) slope; (b) grassland area in 2000; (c) grassland area in 2018; (d) erosion area in 2000; (e) erosion area in 2018.

Table S7. Suitable graze area and graze capacity in 2000 and 2018.

	suitable graze area(km^2)	percent(%)
2000	8439.51	6.59
2018	5941.35	4.64

References:

1. Berg, C.E., Mineau, M.M., Rogers, S.H., 2016. Examining the ecosystem service of nutrient removal in a coastal watershed. Ecosystem Services, 20: 104-112.
2. Donohue, R., Roderick, M., Mcvicar. T., 2012. Roots, storms and soil pores: Incorporating key ecohydrological processes into Budyko's hydrological model. Journal of Hydrology, 436-437, 35-50.
3. Fu, B., 1981. On the calculation of the evaporation from land surface. Sci Atmos Sin, 5: 23-31 (in Chinese).
4. Hamel, P., Kramer, R.C., Simb, S., Mueller, C., 2015. A new approach to modeling the sediment retention service (InVEST 3.0): Case study of the Cape Fear catchment, North Carolina, USA. Science of the Total Environment, 524–525:166-177.
5. Hydraulic Design Manual. Texas Department of Transportation. 2016.1.
6. Katherine, M.K. 2011. Water yield in the Southern Appalachian mountains. Dissertation. University of Minnesota.
7. Kovacs, K., Polasky, S., Nelson, E., et al., 2013. Evaluating the Return in Ecosystem Services from Investment in Public Land Acquisitions. Plos One, 8(6): e62202.
8. Kochenderfer, J,N., Edwards, P.J., Wood, F., 1997. Hydrologic Impacts of Logging an Appalachian Watershed Using West Virginia's Best Management Practices, Northern Journal of Applied Forestry, 14(4): 207-218.
9. Line, D.E., Shaffer, M.B., Blackwell, J.D., 2011. Sediment export from a highway construction site in central north Carolina. Transactions of the ASABE. 54(1): 105-111.

10. O'Hop, J., and Wallace, J.B., 1983. Invertebrate drift, discharge, and sediment relations in a southern Appalachian headwater stream. Hydrobiologia, 98(1): 71-84.
11. PRISM Climate Group, Oregon State University, http://prism.oregonstate.edu, created 4 Feb 2018.
12. Robertson, D.M., Schwarz, G.E., Saad, D.A., Alexander, R.B., 2009. Incorporating uncertainty info the ranking of sparrow model nutrient yields from Mississippi/Atchafalaya river basin watersheds. Journal of the American water resources association, 45(2): 534-549.
13. Sharp, R., Tallis, H.T., Ricketts, T., Guerry, A.D., Wood, S.A., Chaplin-Kramer, R., et al., 2016. InVEST +VERSION+ User's Guide. The Natural Capital Project, Stanford University, University of Minnesota, The Nature Conservancy, and World Wildlife Fund.
14. Villines, J.A., Agouridis, C.T., Warner, R.C., Barton, C.D., 2015. Using GIS to delineate headwater stream origins in the Appalachia coalfields of Kentucky. Journal of the American water resources association. 51(6): 1667-1687.
15. Zhang, L., Hickel, K., Dawes, W.R., et al., 2004. A rational function approach for estimating mean annual evapotranspiration. Water Resources Research, 40(2): 89-97.
16. Zomer, R.J., Trabucco, A., Bossio, D.A., van Straaten, O., Verchot, L.V., 2008. Climate Change Mitigation: A Spatial Analysis of Global Land Suitability for Clean Development Mechanism Afforestation and Reforestation. Agric. Ecosystems and Envir. 126: 67-80.