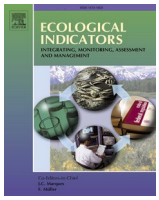
Ecological Indicators 148 (2023) 110046

Contents lists available at ScienceDirect

Ecological Indicators

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



Driving mechanisms of urbanization: Evidence from geographical, climatic, social-economic and nighttime light data

Siyi Huang [a,b], Lijun Yu [a,*], Danlu Cai [a,*], Jianfeng Zhu [a], Ze Liu [c,d], Zongke Zhang [a], Yueping Nie [a], Klaus Fraedrich [e]

[a] Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China
[b] University of Chinese Academy of Science, Beijing, China
[c] Research Center of Territorial & Spatial Planning, The Ministry of Natural Resources of the People's Republic of China, Beijing, China
[d] China Land Surveying and Planning Institute, Beijing, China
[e] Max Planck Institute for Meteorology, Hamburg, Germany

ARTICLE INFO

Keywords:
Urban land extraction
NPP/VIIRS
Geodetector
Attribution analysis
Contribution dynamics

ABSTRACT

Urbanization induced changes have attracted widespread attention. Key challenges arise from the inherent uncertainties in attribution models diagnosing the driving mechanisms and the interrelationships of the attributes given by the complexity of interactions within a city. Here, we investigate urbanization dynamics from nighttime light signals before analyzing their driving mechanisms from 2014 to 2020 on both provincial and regional scale and a flat versus mountainous urbanization comparison. Model uncertainties are discussed comparing the contribution results from Geodetector and the Gini importance from Random Forest analyses. The method is applied to Shaanxi Province, where flat urban land is located mainly in its center and mountainous urban land is situated in the North and South. The following results are noted: i) Employing the Geodetector based maximum contribution method for urban region extraction of night time light reveals a notable accuracy improvement in flat urban land compared with the closest area method. ii) Geographical factors attain high contribution for mountainous urban land of Shannan, while for flat urbanization land dynamics, economic factors and community factors prevail. iii) The most obvious driving mechanisms are economic factors which, associated with local urban development strategies, show highest contribution values in 2014 (2018) over the flat (mountainous) urban land of Guanzhong Plain (Northern Shaanxi Plateau or Shanbei region) linked with an early (late) development. iv) Population factors achieve high contribution values in the initially low populated urban land of the northern mountainous land initiating huge migration. v) The contributions resulting from Geodetector are in agreement with the Gini importance from Random Forest in agriculture, geographical and population factors ($R^2 > 0.5$) but not in economy, community and climatic factors ($R^2 < 0.5$). The dynamics of driving mechanisms for urbanization provides insights in connecting urban geographical expansion with multi-factors and thus to assist municipal governments and city stakeholders to design a city with geographical, climatic and social-economic changes and interactions in mind.

1. Introduction

Urbanization refers to the population shift from rural to urban areas (Parnell and Walawege, 2011). The corresponding increase in the proportion of people living in urban areas and increases in material demands of production, human consumption and urban waste discharge have recently emerged as a sustainability challenge related to urban dynamics and associated with environmental consequences during the urbanization processes (Avtar et al., 2019; Chen et al., 2018). Research on urban dynamics includes urban expansion (Jiao, 2015; Li et al., 2018; Zhang et al., 2018b), population growth (Bongaarts, 2016; Qizhi et al., 2016; Shang et al., 2018), economic growth (Ahmad et al., 2021; Bakirtas and Akpolat, 2018; Minh Ha and Nguyen, 2017), and electric power consumption (Bilgili et al., 2017; Sheng et al., 2017; Xie et al., 2020). The associated environmental consequences refer to the urbanization/anthropogenic induced changes, including increased atmospheric greenhouse gas concentration (Ahmed et al., 2019; Liu and Bae, 2018; Wang et al., 2019b), urban heat island effect (Carmona et al.,

* Corresponding authors.
E-mail addresses: yulj@aircas.ac.cn (L. Yu), caidl@radi.ac.cn (D. Cai).

https://doi.org/10.1016/j.ecolind.2023.110046
Received 29 July 2022; Received in revised form 15 January 2023; Accepted 16 February 2023
Available online 24 February 2023

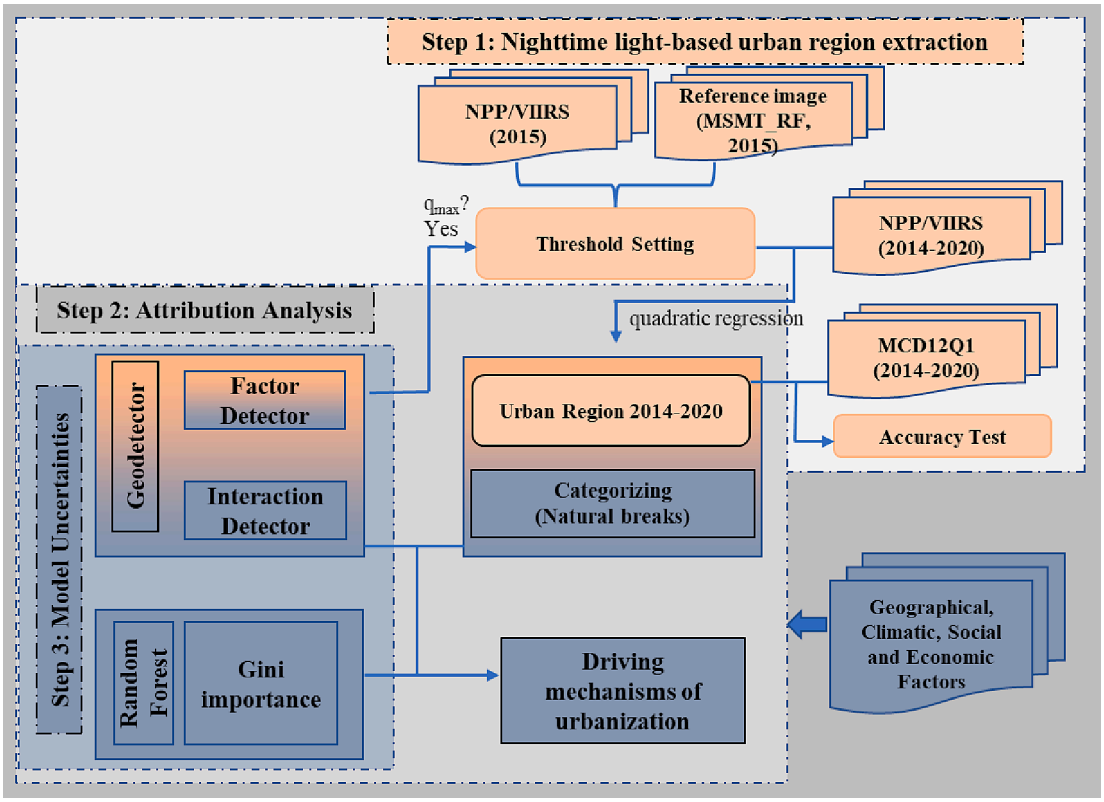


Fig. 1. Methodology flowchart.

2017; Fu and Weng, 2016; Zhou et al., 2019), extreme climate events (Lin et al., 2020; Wang et al., 2022; Zhang et al., 2018a), aerosol emissions (Jiang et al., 2017; Wang et al., 2021a; Xue et al., 2019), and land use and land cover change (Haregeweyn et al., 2012; Hou and Wen, 2020; Ul Din and Mak, 2021).

As a rapidly changing developing country with the largest population in the world, the urbanization rate in China has increased steadily over the last decades. Chinese urbanization commenced since mid-twentieth century (Liu et al., 2021), rose from 18% of the total population in China in 1970 s, to 36% in 2000, to around 51% in 2011, and up to approximately 60% till 2018 (Gaughan et al., 2016; Luo et al., 2018; Yu, 2021). The following direct changes in Chinese urbanization have been observed: i) Urban expansion has occurred over 95% of the Chinese cities from 1970 to 2010 (Wang et al., 2020). ii) Cropland reduction has increased from 47.29% in 2000 to 77.46% in 2015 (Chunyu et al., 2019). iii) Gross domestic product has increased with an annually averaged rate of 9.5% from 1995 to 2012 (Zhao and Tang, 2018). And iv) the annual growth rate of energy consumption has increased from 5.5% in 1990s to 9.6 % in 2000s (Zheng and Walsh, 2019) and reaching 12% in the 2020s (Elheddad et al., 2020). All those observed changes in Chinese urbanization show negative influence on changing the urban surface climate, including greenhouse gas emissions (Feng et al., 2017; Wang and Li, 2018; Zhang et al., 2019b), temperature increase (Wang et al., 2021b; Wang et al., 2019a; Wang et al., 2021c), PM2.5 increase (Du et al., 2019; Lin et al., 2018; Wu et al., 2018), precipitation rise in the city core (Liu and Niyogi, 2019), etc..

Accordingly, a great number of studies analyze the driving factors of urbanization and the corresponding changes to guide a sustainable urban development in China. Mathematical methods such as regression or spatial regression models and machine learning are popular. Traditional studies related to the driving mechanisms of urbanization focus on the regression relationship between social economic factors and an urbanization indicator (Cai and Fangyuan, 2020; Guo et al., 2021). For example, Yang et al. (2019) analyze the driving mechanisms of urbanization by associating the real estate investment, per capita fiscal expenditure, and the urban–rural income ratio with the ratio of urban to total population. Recent research has shown advantages in using artificial intelligence and machine learning. For example, Wu et al. (2021) and Zhang et al. (2019a) identify driving factors for land use change using the random forest method. Ye et al. (2021) apply Principal Component Analysis and a stepwise regression to obtain the contribution from driving factors.

However, regression models require the assumption that the probability distribution of the data and whether the function of independent variables is linear or nonlinear. While spatial regression models may suffer from multicollinearity problems (see (Gao et al., 2012)). Artificial intelligence (AI) and machine learning (ML) methods show advantages in performing complex tasks and making decisions based on deep data analysis.

Here, a simple Geodetector method (Wang et al., 2016) and the machine learning (ML) methods of Random Forest (Breiman, 2001) are used. The Geodetector provides a geostatistical diagnostic of the observed spatial stratified heterogeneity and attributes the spatial stratified heterogeneity to selected impact factors and thereby detects whether the impact factor is more similar within strata than between strata. One of the advantages of the diagnostics is that the calculated attributions are without any linear assumption (Ju et al., 2016; Shen et al., 2015; Wang et al., 2016; Wang et al., 2010; Yang et al., 2016). Together with the Geodetector diagnostic, the interrelationship between multi-factors and the annual attribution dynamics from 2014 to 2020 could be easily analyzed. Furthermore, Geodetector method based on driving mechanisms is compared with factor contributions from machine learning (ML) methods of Random Forest.

Instead of using present urban land use products (MCD12Q1, (Liu et al., 2018)), which define urban land depending on physical discrepancies by integrating the presence of human-made structures and materials. Remotely sensed nighttime light signals are chosen due to their potential to obtain more comprehensive characteristics relating to urban

Table 1
Illustration of urbanization related data, and geographical, climatic, social-economic factors. For details of the selected 40 social-economic factors see Appendix Table 1.

Data		Original Spatial resolution	Period	Sources
Urbanization related data	Nighttime light data	500 m	2014–2020	NPP/VIIRS
	Reference image	30 m	2015	MSMT_RF (Zhang et al., 2020)
	Accuracy test image	500 m	2014–2020	MCD12Q1.006
Geographical factors	Digital elevation model	30 m	/	SRTM
	Slope	30 m	/	See section 2.3
	multiresolution index of the ridge top flatness (MrRTF)	30 m	/	See section 2.3
	multiresolution index of valley bottom flatness (MrVBF)	30 m	/	See section 2.3
Climatic factors	Temperature tendency	1000 m	2000–2020	MOD11A1.006
	Precipitation tendency	0.05°	1981–2020	CHIRPS(Funk et al., 2015)
	Drought index tendency	2.5 arc min	1958–2020	TerraClimate(Abatzoglou et al., 2018)
Social-economic factors	Community indicators Population indicators Economy indicators Agriculture indicators	Prefectures level	2014–2020	Shaanxi Provincial Bureau of Statistics, https://tjj.shaanxi.gov.cn/ (Accessed 12 May 2022)

dynamics and the associated environmental consequences during the urbanization processes (Ma et al., 2012). Present methods extracting urban land indices include i) area optimized thresholding by the closest area method (Liu et al., 2012; Small et al., 2005; Yu et al., 2021), ii) vegetation adjusted urban index (Li et al., 2016), and iii) supervised classification (Dou et al., 2017). The closest area method extracts the urban land by determining an optimal threshold when the urban land area from the nighttime light data is closest to the area of reference image. Vegetation adjusted urban index extracts the urban land by integrating the nighttime light and vegetation information to reduce the nighttime light saturation effect in urban cores and over-glowing effect. Supervised classification is based on the idea that a user can select sample pixels that are representative of urban land, and then direct the image processing software to categorize the nighttime light signals into urban land and non-urban land (see (Liu et al., 2019)).

To understand the driving mechanisms of urbanization and the uncertainties of the analysis, it is the first time that (i) Geodetector analysis is employed to obtain the urban region, and (ii) this in terms of nighttime lighted area selected by an intensity threshold which is given by the maximum contribution of the nighttime light data relative to the reference impervious (urban) surface distribution. And subsequently, both Geodetector (2014–2020) and Random Forest (2015) are used to attribute urban dynamics (2014–2020) to geographical, climatic and social-economic factors. Thus, a three-step spatial analysis (section 2) is designed: Step-1, temporally, nighttime light data is employed to quantify the long-term annual urban region, whose accuracy is tested with present land cover data. Step-2, spatially, Geodetector is applied to diagnose the non-linear statistical attribution dynamics of geographical, climatic and social-economic impact factors for all single years' urban region. Step-3, the contribution result from Geodetector is compared with the Gini importance from Random Forest for better understanding the uncertainty and limitation of these two contribution models. This is followed by an application (section 3) to Shaanxi Province, China, and a subsequent discussion of the results (section 4), and finally concludes the analysis (section 5).

2. Materials and methods

Data and preprocessing: Nighttime light signals (2014–2020) derived from the Suomi National Polar-orbiting Partnership Visible Infrared Imaging Radiometer Suite (NPP/VIIRS) are selected for obtaining the urban region. The NPP/VIIRS nighttime light monthly composite from the NOAA/NGDC is available since 2012 (https://ngdc.noaa.gov/eog/index.html, accessed on 3 July 2020) with a spatial resolution of 15 arc-seconds (0.5 km*0.5 km). Multisource, multitemporal random forest (MSMT_RF) map 2015 (Zhang et al., 2020) with an overall accuracy of 95.1% and a spatial resolution of 30 m*30 m is used for the urban region threshold setting. MODIS land cover data MCD12Q1 (2014–2020) is used for the accuracy test of nighttime light-based urban region extraction. Note that, to be comparable, all remote sensing images are resampled as the spatial resolution of 0.5 km*0.5 km and projected to the Asia North Albers Equal Area Conic projection. The methodology consists of urban region extraction and attribution analysis (flowchart see Fig. 1). In the attribution analysis, geographical, climatic and social-economic factors (Table 1 and Appendix Table 1) are used.

2.1. Geodetector

Geodetector is designed to assess the relationship between resultant outcome (the dependent variable *Y*) and driving factors (the independent variable *X*) through the spatial heterogeneity (see (Wang et al., 2016)). The Geodetector diagnostic consists of four detectors, including factor detector, interaction detector, ecological detector and risk detector. Both numerical and qualitative data can be listed as independent variables after categorizing (Chen et al., 2020; Wang et al., 2016). The following two detectors are used.

Factor detector: The factor detector q tests whether one particular factor X is the cause for a certain resultant outcome Y by comparing the total Y-variance in a subregion (or category, σ^2_{sub}) with the Y-variance in the entire region σ^2. N represents the number of pixels in the entire region, and N_{sub} represents the number of pixels in a subregion (or category):

$$q = Factordetector(X,Y) = 1 - \frac{\sum_{h=1}^{L} N_{sub}\sigma^2_{sub}}{N\sigma^2} \quad (1)$$

The factor detector q ranges from *0* to *1*. When q approaches *0*, categories of the independent variable X cannot explain the distribution of the dependent variable Y. Contrarily, when q approaches *1*, categories of the independent variable X explain the distribution of dependent variable Y and can be listed as one of the driving factors. The higher the q value, the higher is the contribution of the driving factor.

Interaction detector: The interaction detector calculates the inter-contribution of two factors (X_1, X_2) (inter-q value) on the dependent variable Y (more detailed information see (Ran et al., 2019)):

Enhance and nonlinear: $q(X_1 \cap X_2) \rangle q(X_1) + q(X_2)$
Independent: $q(X_1 \cap X_2) = q(X_1) + q(X_2)$
Enhance and bivariate: $q(X_1 \cap X_2) \rangle Max(q(X_1), q(X_2))$
Weaken and univariate: $Max(q(X_1), q(X_2)) \rangle q(X_1 \cap X_2) \rangle Min(q(X_1), q(X_2))$
Weaken and nonlinear: $q(X_1 \cap X_2) \langle Min(q(X_1), q(X_2))$

Note that datasets of driving factors (the independent variable X) should be categorized before entering the diagnostics (more details see

Table 2
Accuracy test. Note that OA represents overall accuracy.

region	method	index	2014	2015	2016	2017	2018	2019	2020
Shaanxi	Closest area	OA(%)	98.21	98.27	98.28	98.12	98.03	98.00	98.00
		Kappa	0.44	0.45	0.46	0.45	0.45	0.45	0.45
	Maximum contribution	OA(%)	98.48	98.43	98.41	98.35	98.30	98.24	98.20
		Kappa	0.47	0.47	0.47	0.47	0.48	0.47	0.47
Shanbei	Closest area	OA(%)	99.40	99.45	99.43	99.40	99.39	99.41	99.35
		Kappa	0.23	0.25	0.24	0.23	0.21	0.24	0.19
	Maximum contribution	OA(%)	99.42	99.46	99.46	99.44	99.38	99.43	99.37
		Kappa	0.22	0.25	0.25	0.24	0.21	0.24	0.20
Guanzhong	Closest area	OA(%)	95.35	95.48	95.49	94.93	94.58	94.27	94.13
		Kappa	0.51	0.50	0.50	0.50	0.48	0.48	0.47
	Maximum contribution	OA(%)	96.50	96.26	96.18	96.03	95.84	95.55	95.25
		Kappa	0.55	0.53	0.53	0.54	0.53	0.52	0.50
Shannan	Closest area	OA(%)	99.52	99.52	99.48	99.40	99.32	99.38	99.40
		Kappa	0.49	0.48	0.50	0.47	0.45	0.48	0.49
	Maximum contribution	OA(%)	99.50	99.45	99.42	99.36	99.30	99.34	99.32
		Kappa	0.4	0.47	0.48	0.46	0.44	0.47	0.48

section 2.3).

2.2. Nighttime light based urban region extraction

Unlike previous urban region extraction methods mentioned in the introduction, factor detector (see section 2.1) is used here. Only the nighttime light data and the reference impervious surface distribution are required as driving factor (the independent variable X) and resultant outcome (the dependent variable Y), respectively using the data of 2015. Further detailed processes are noted:

Threshold setting: Two methods are employed and compared for a better threshold setting.

i) Closest area method: It determines the optimal threshold, when area from nighttime light data is closest to the area from the reference impervious surface (for a more detailed description see also (Yu et al., 2021)).

ii) Maximum contribution method: It is novel that the Geodetector diagnostic is to be used for the first time to extract the optimal threshold of the nighttime light signal-based urban region. This is not performed by comparing urban area differences, but by integrating the spatial distribution heterogeneity between the nighttime light signal and the reference impervious surface distribution (for detailed advantages see introduction).

That is, the optimal digital number (or threshold), DN_{opt}, of the nighttime light signal for defining the urban region as regions with nighttime light exceeding the DN_{opt} threshold, is when the maximum contribution q_{max} of the nighttime light data on the reference impervious surface distribution is obtained (see Eq. (2)):

$$DN_{opt} = f^{-1}\{q_{max}\} = f^{-1}[Factordetector_{max}(\text{nightlight}_{\geq DN_x}, \text{reference image})]$$

$$\text{where } DN_x \in \{0.1, \cdots, 60, step = 0.1\} \tag{2}$$

Thresholds adjustment: Based on the above-obtained optimal threshold with the reference impervious surface distribution of 2015, annual urban region from 2014 to 2020 is calculated with nighttime light exceeding the $DN_{opt2015}$. To reduce the effects of over-glow in a certain year, usually caused by anthropogenic activity in undeveloped areas that were not taken into account in statistical data of urban development in the subsequent year, a further quadratic regression has been applied to adjust the urban region from 2014 to 2020 with a F-test (p-value < 0.05 is used unless mentioned otherwise); the yearly thresholds are adjusted accordingly.

Accuracy test: Followed by the threshold adjustment, the extracted urban region from 2014 to 2020 is tested with MODIS Land Cover Product (MCD12Q1, 500 m spatial resolution) by an input of a random sample selection with a third of the total pixel number into the 'compute confusion matrix' tool in ArcGIS software 10.7 (application result see section 3 and Table 2).

2.3. Attribution: geographical, climatic and social-economic factors

Geographical, climatic, social-economic factors are compared and attributed to the urbanization obtained from the nighttime light data. Geographical and climatic factors require preprocessing and all driving factors require a categorizing process:

Geographical indicators: Slope, multi-resolution index of the ridge top flatness (MrRTF, (Gallant and Dowling, 2003)) and of valley bottom flatness (MrVBF, (Gallant and Dowling, 2003)) are calculated from digital elevation model (DEM) using SAGA GIS software 8.2. As a widely used geographical factor, slope describes the change in elevation. MrRTF is a topographic index designed to identify high flat areas, while the MrVBF index is designed to identify flat valley bottoms.

Climatic tendency: The linear regression of long term yearly averaged climate parameter Cli_{yr}, including temperature, precipitation and drought index, is calculated as the climatic independent variable input (data descriptions see Table1).

$$tendency(x) = \frac{n^*\sum_{yr=1}^{n} yr^* Cli_{yr} - \left(\sum_{yr=1}^{n} yr\right)\left(\sum_{yr=1}^{n} Cli_{yr}\right)}{n^*\sum_{yr=1}^{n} yr^2 - \left(\sum_{yr=1}^{n} yr\right)^2} \tag{3}$$

Note that n represents the number of years, $yr \in \{1, 2, \cdots, n\}$.

Categorizing: As mentioned before, Geodetector requires the independent factors to be classified into categories. Here Jenks Natural Breaks Classification (Natural Breaks) in ArcGIS software 10.7 is used due to its advantage in minimizing the variance within the classes and maximizing the variance between classes (Jenks, 1967).

Attribution (Geodetector): Factor detector and interaction detector (see section 2.1) are used to define whether one particular factor is the reason for the nighttime light-based urbanization and to explore the interrelationships between the driving factors of the urbanization process. Note that all the results related to contribution shown in this paper have passed the significance test unless mentioned otherwise.

Attribution (Random Forest): The Gini importance from Random Forest model (Menze et al., 2009) is frequently used for measuring the independent variable impurity with respect to a reference image (Jin et al., 2018) and thus to obtain the importance of different driving forces (Huang et al., 2020; Ma et al., 2020). With the Gini importance (the package randomForest in R could be used for calculation), the larger the reduction in Gini importance caused by independent variable changes is, the more important the variable (Foody and Arora, 1997).

3. Application and results: Shaanxi Province

The three-step spatial temporal attribution analysis is applied to

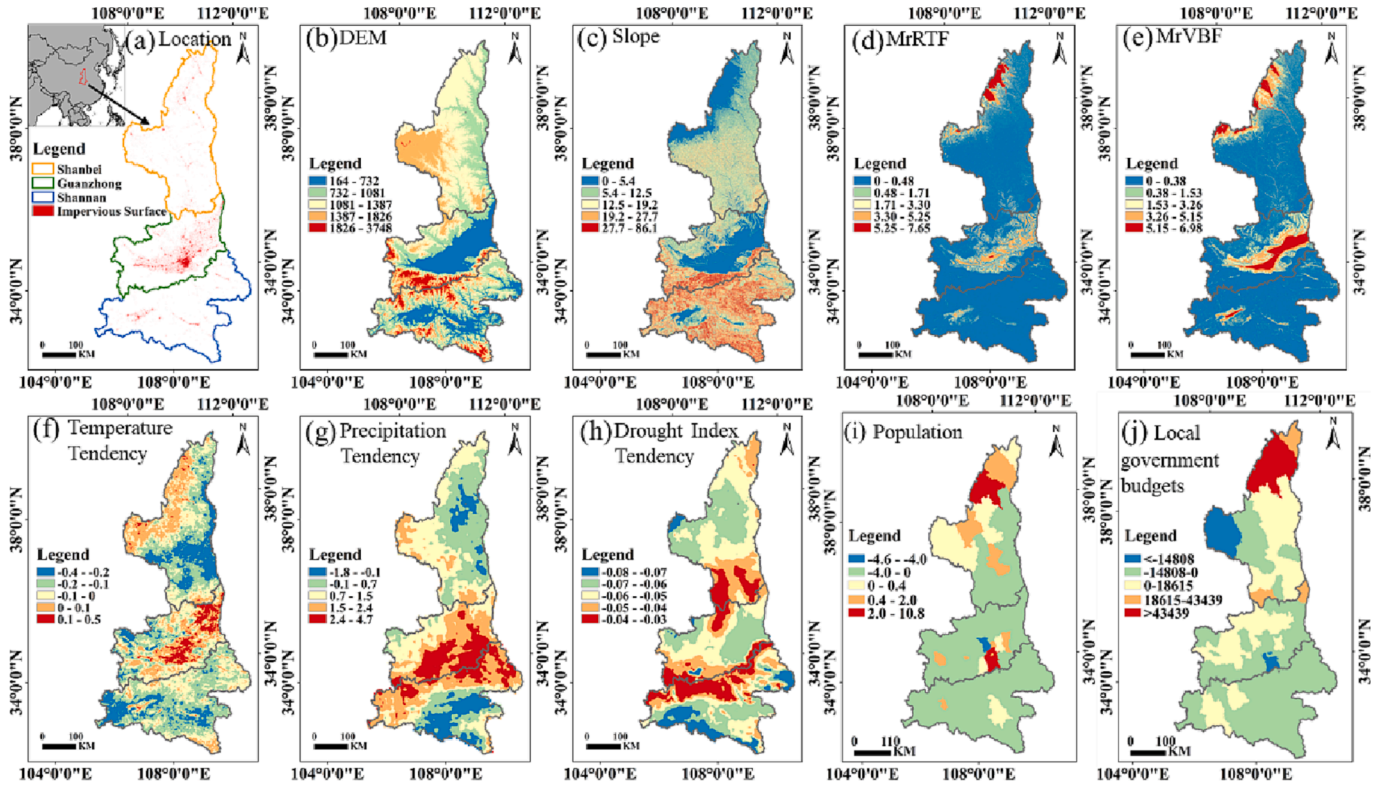


Fig. 2. Geographical setting of geographical, climatic and social-economic factors in Shaanxi Province: (a) the location map of the study area; (b) digital elevation model (DEM); (c) slope; (d) MrRTF; (e) MrVBF; (f) temperature tendency; (g) precipitation tendency; (h) drought index tendency; (i) population (prefectures); and (j) Local government budgets (prefectures). Note that the three sub-regions (Northern Shaanxi Plateau or Shanbei, Guanzhong Plain, and Shannan Mountain region) and the reference built-up area (MSMT_RF, 2015 (Zhang et al., 2020)) are highlighted in (a).

Shaanxi Province, which comprises topographically, social-economically and climatically complex areas due to the Qinling Mountains and local different urbanization stages. Three sub-regions are included (Fig. 2): 1) Northern Shaanxi Plateau or Shanbei region, with an elevation from 900 m to 1900 m, which is relatively cold, dry and known for local coal resources. 2) Guanzhong Plain lies in the middle of Shaanxi Province and is covered by urban agglomerations attracting most of the population. 3) Shannan Mountain region, south of the Qinling-Huaihe Line, is relatively humid and, in contrast to Guanzhong Plain, contribution to emigration. Here, threshold setting of the night-time light-based urbanization is obtained on both provincial and sub-region scale using both Geodetector based maximum contribution and closest area methods. Employing an accuracy test, attribution analysis relates geographical, climatic and social-economic factors with the urbanization process.

3.1. Geographical, climatic and social-economic factors

Geographical, climatic and social-economic factors in Shaanxi Province (Fig. 2) are characterized in the following:

i) Geographically (Fig. 2 b and c), *90%* urban land is located in regions where elevation and slope are less than $\sim$ *1* km and $\sim$ *3°*, respectively, which fits Zipf's power-law implying that small occurrences are extremely common, whereas large instances are extremely rare (Adamic et al., 2000). That is, linking urban planning with the preference for certain geographical features shows that, instead of mountainous regions, flat land appears to be preferable because of lower building costs and better urban expansion possibilities (see (Zhao et al., 2014)), which is in agreement with previous research (Ma and Xu, 2010; Mundia and Aniya, 2005; Pan et al., 2021).

ii) Climatically (Fig. 2 f-h), the temperature (precipitation) has been significantly increasing (decreasing) in Guanzhong Plain in the middle of Shaanxi Province, where most of the urban agglomerations are located. However, for most regions over the Northern Shaanxi Plateau or Shanbei and Shannan Mountain region, the temperature (precipitation) has been significantly decreasing (increasing). Although the above-mentioned differences of temperature and precipitation occur over three sub-regions, the whole province is getting drier (see Fig. 2h) according to the Palmer drought index computed using precipitation, temperature and soil characteristics (more details see (van der Schrier et al., 2011)).

iii) The social-economical aspect shows that after a rapid decreasing from 2016 to 2017, local government budgets of Shaanxi Province and for the three regions of Guanzhong Plain, Northern Shaanxi Plateau or Shanbei region and Shannan Mountain region, indicate a stagnation. Guanzhong Plain accounts for about 57% local government budgets of the total while Northern Shaanxi Plateau or Shanbei region and the Shannan Mountain region occupy about 34% and 9% of the total government budgets. This is inconsistent with the proportion of permanent residents. The number of permanent residents of Shaanxi Province is stable from 2014 to 2020 with a slope of 0.84 (not significant), with about 63%, 15% and 22% residing in the Guanzhong Plain, Northern Shaanxi Plateau or Shanbei region and Shannan Mountain region, respectively. That is, the local government budgets per person is highest (lowest) over Northern Shaanxi Plateau or Shanbei region (Shannan Mountain region) region.

3.2. Geodetector analysis of urbanization

Based on the heterogeneity of geographical, climatic and social-economic distribution in Shaanxi Province (section 3.1), threshold setting, adjustment, accuracy test and attribution analysis are carried out depending on both provincial and subregion scale.

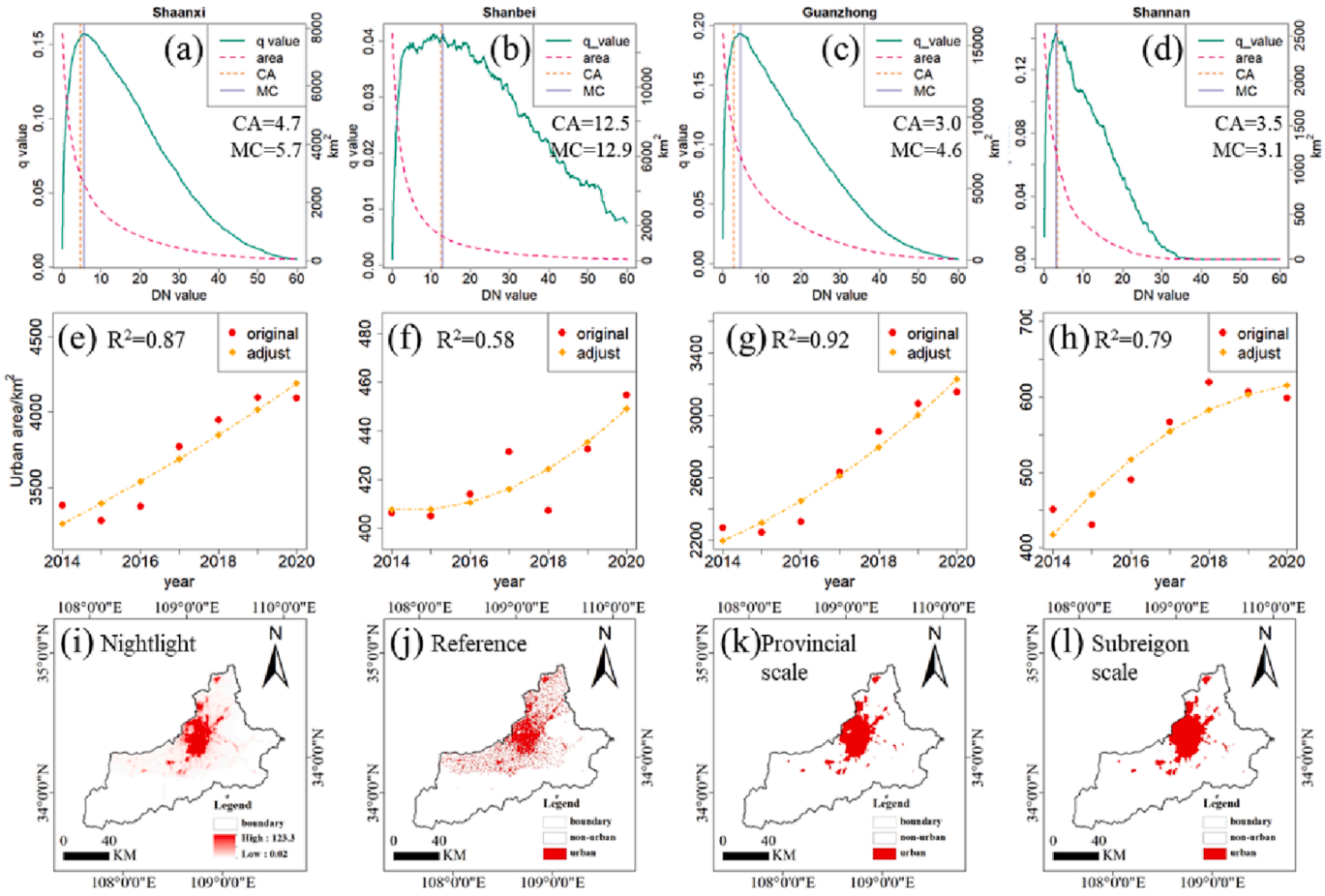


Fig. 3. Illustration of urban extraction: threshold setting (1st row), urban area adjustment (2nd row), distribution of urban extraction (3rd row). Comparison of Geodetector based maximum contribution result versus the closest area-based result is shown on both provincial scale (a and e) and sub-regions scale (b and f, c and g, d and h). Extraction results of urban region in Xi'an (2015) are shown for example: (i) nighttime light data, (j) reference data (MSMT_RF, 2015, see (Zhang et al., 2020)), and Geodetector based maximum contribution from (k) provincial scale threshold and (l) sub-regions scale threshold. Note that CA represents closest area method and MC represents maximum contribution method.

3.2.1. Threshold setting and adjustment: Maximum contribution versus closest area

To obtain the 2015 urban region threshold for suitably defining the urbanization, a novel Geodetector based maximum contribution method is introduced and compared with the traditional closest area method over both provincial and sub-regions scale (Fig. 3). The following results are noted:

i) The maximum contribution method indicates that as the nighttime light threshold increases, the Geodetector based contribution q from the area of the region exceeding the threshold (independent variable X) to the reference impervious surface (dependent variable Y), increases up to a peak value and then decreases (see green in Fig. 3 a-d,). Meanwhile, the closest area method shows area of region greater than threshold approaches to the area of the reference impervious surface then apart (see pink in Fig. 3 a-d).

ii) Slight differences of the final thresholds from the two methodologies occurred in the subregions of Northern Shaanxi Plateau or Shanbei region and Shannan Mountain region (Fig. 3b and 3d), while more obvious differences of the final thresholds from the two methodologies were detected for the subregion of Guanzhong Plain scale (threshold = *4.6* versus *3.0*, Fig. 3c). The Geodetector based maximum contribution method has a notable improvement in the accuracy of extracting nighttime light based urban region in the subregion of Guanzhong Plain which is mainly covered by flat urban land (see Fig. 3).

iii) The final thresholds from the two methodologies show high values (threshold greater than 12) in the subregion scale of Northern Shaanxi Plateau or Shanbei region with low Kappa value of $\sim$ *0.2* compared with thresholds (threshold $<$ *6*) from both the whole province scale and the other two subregions (see Table 2). This may be related to Northern Shaanxi Plateau or Shanbei region, which is one of the main coal-producing regions in China (Liu et al., 2015) where non-urban area light is observed like in the districts of the oil industry causing nighttime light misclassification near Williston, in North Dakota (US) (Cai et al., 2017).

3.2.2. Accuracy test with MODIS data

Obtaining the 2015 urban region thresholds from the novel Geodetector based maximum contribution method and the traditional closest area method, urban areas from 2014 to 2020 are calculated. A further urban area adjustment (Fig. 3 e-h) is carried out based on the regression of the urban area from 2014 to 2020 to reduce the effects of over-glow in a certain year. Results from the two methods over both provincial and sub-regional scale are tested with MODIS data and show that the Geodetector based maximum contribution method is superior in both overall accuracy and in Kappa values compared with the traditional closest area method, with the exception of the Shannan Mountain region. Thus, the closest area method appears to be more suitable in the

Table 3
The Geodetector based contribution statistics of geographical, climatic, social and economic factors. Note that, *40* social and economic factors are integrated in the Geodetector, and the contribution values (q-value) greater than *0.5* are listed.

type	Contributor	q value
Geographical factors	DEM	0.44
	MrVBF	0.41
	MrRTF	0.38
	Slope	0.28
Climatic factors	Drought index tendency (1958–2015)	0.18
	Temperature tendency (2000–2015)	0.15
	Precipitation tendency (1981–2015)	0.08
Social-Economic factors	Total retail sales of consumer goods	0.90
	Per capita annual disposable income of rural households	0.85
	Per capita annual disposable income of urban households	0.71
	Number of neighborhood committees	0.69
	Local general budgetary revenue	0.65
	Number of street communities	0.61
	Number of legal entities	0.58
	Number of towns	0.55

mountainous regions, while the Geodetector based maximum contribution method has a notable accuracy advantage in flat urban land. This may be related to a higher concentration of urban distribution over flat terrain, which leads to a relatively smaller variance associated with a higher contribution value q (see Eq. (1)).

3.3. Attribution analysis

Since the urban region is located mainly over the flat regions, the novel Geodetector based maximum contribution method is used in the following attribution analysis due to its superiority in the overall accuracy test (see Table 2). For the year 2015, contributions from geographical, climatic and the top 9 social-economic factors are listed in Table 3. The highest contribution for urban region arises from social-economic factors and (in decreasing order) from total retail sales of consumer goods > per capita annual disposable income of rural households > per capita annual disposable income of urban households > number of neighborhood committees > local general budgetary revenue > number of street communities > number of legal entities > number of towns (only contribution rates greater than *0.5* are listed). For geographical factors, the contributions of elevation (*0.44*), is greater than MrVBF and MrRTF (0.41 and 0.38, respectively), while slope is with less importance of *0.28*. The climatic factors appear to be less relevant in the urban region of the year 2015 with contributions ranking from tendencies of drought index (*0.18*) > temperature (*0.15*) > precipitation (*0.08*).

In general, for a net contribution, social-economic factors indicate a relatively higher value than geographical factors, and both of which exceed the influence of climatic factors. This is because social-economic factors directly improve the urban infrastructure system and working opportunities which attracting population immigration, supports a further urban development. In addition, geographical factors influence shape and boundaries of urban expansion.

To better understand the driving mechanisms of urbanization, contribution dynamics and factor interactions are calculated. Results from 2014 to 2019 (see Figs. 4 to 7) are presented (without the agriculture data missing in 2020). Along the diagonal line (*slope* = *1*), the contribution (q value) of a single factor is displayed with values ranging from *0* to *1* (color from blue to red), while inter-contributions of each two factors are shown below the diagonal-line with the same color ranges. Likewise for the factor interactions: Interaction of the two categories is shown above the diagonal line. Here, interactive types of bivariate and nonlinear enhance are detected. The following results are noted:

i) On the whole provincial scale (Fig. 4), the following single factor

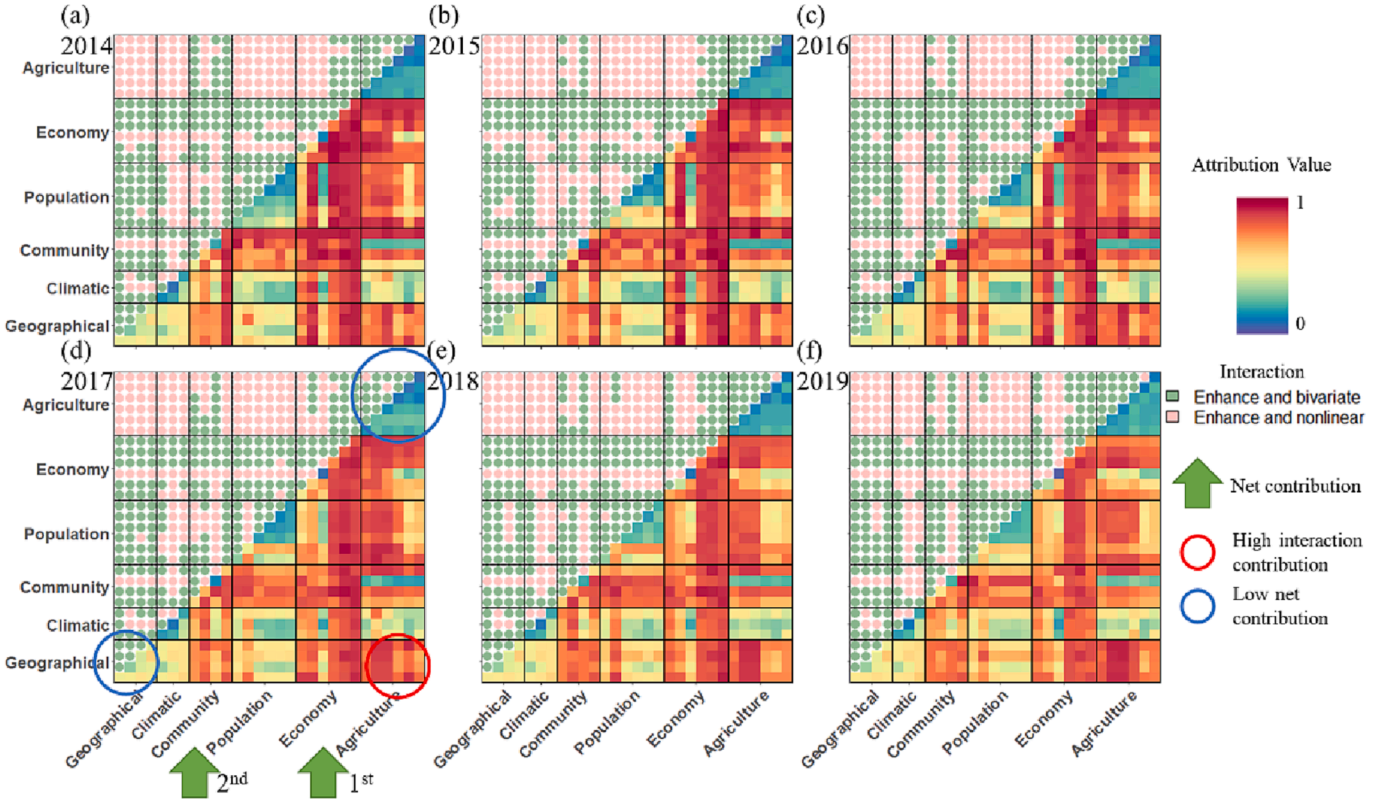


Fig. 4. Attribution of geographical, climatic and social-economic factors to urbanization and related factor interactions over Shaanxi Province from 2014 to 2019.

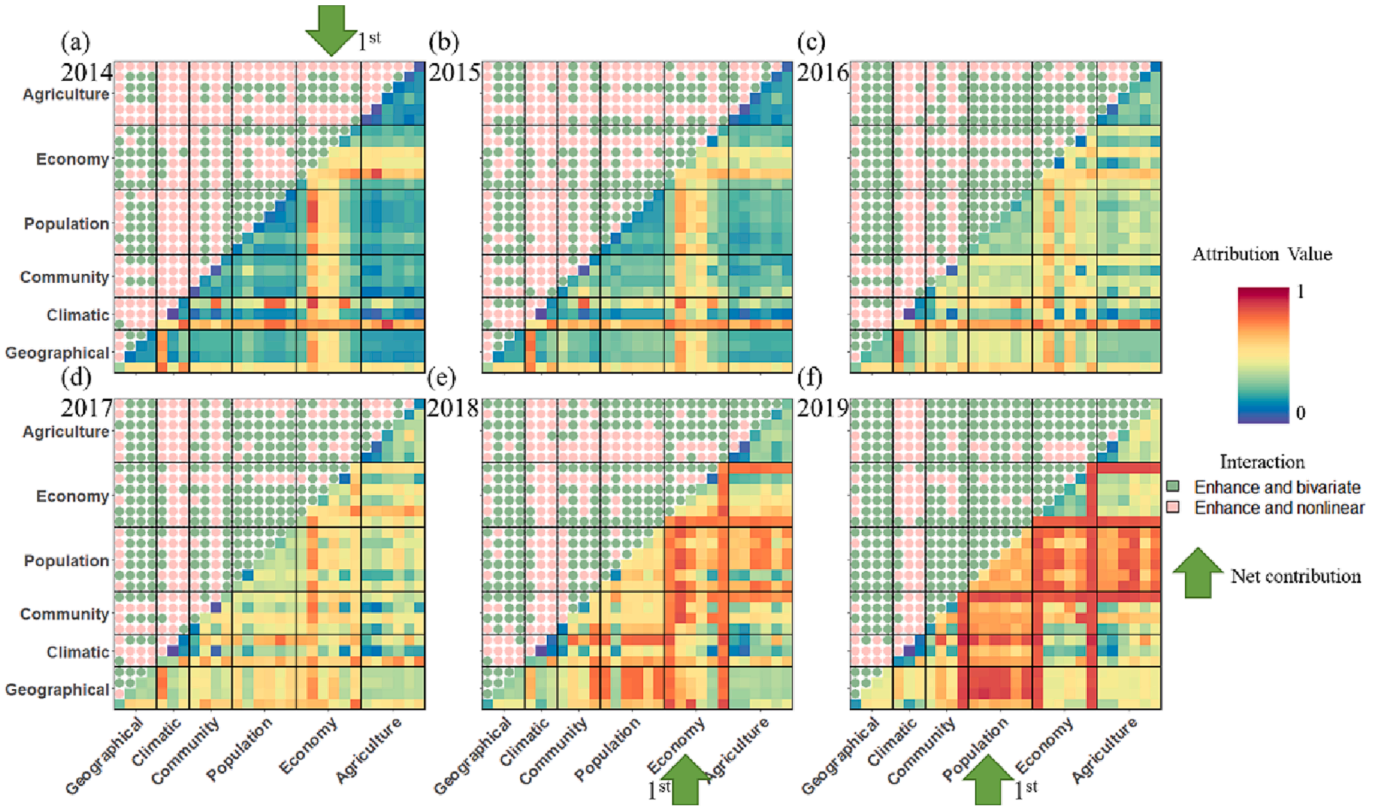


Fig. 5. Attribution of geographical, climatic and social-economic factors to urbanization and related factor interactions over Northern Shaanxi Plateau or Shanbei region from 2014 to 2019.

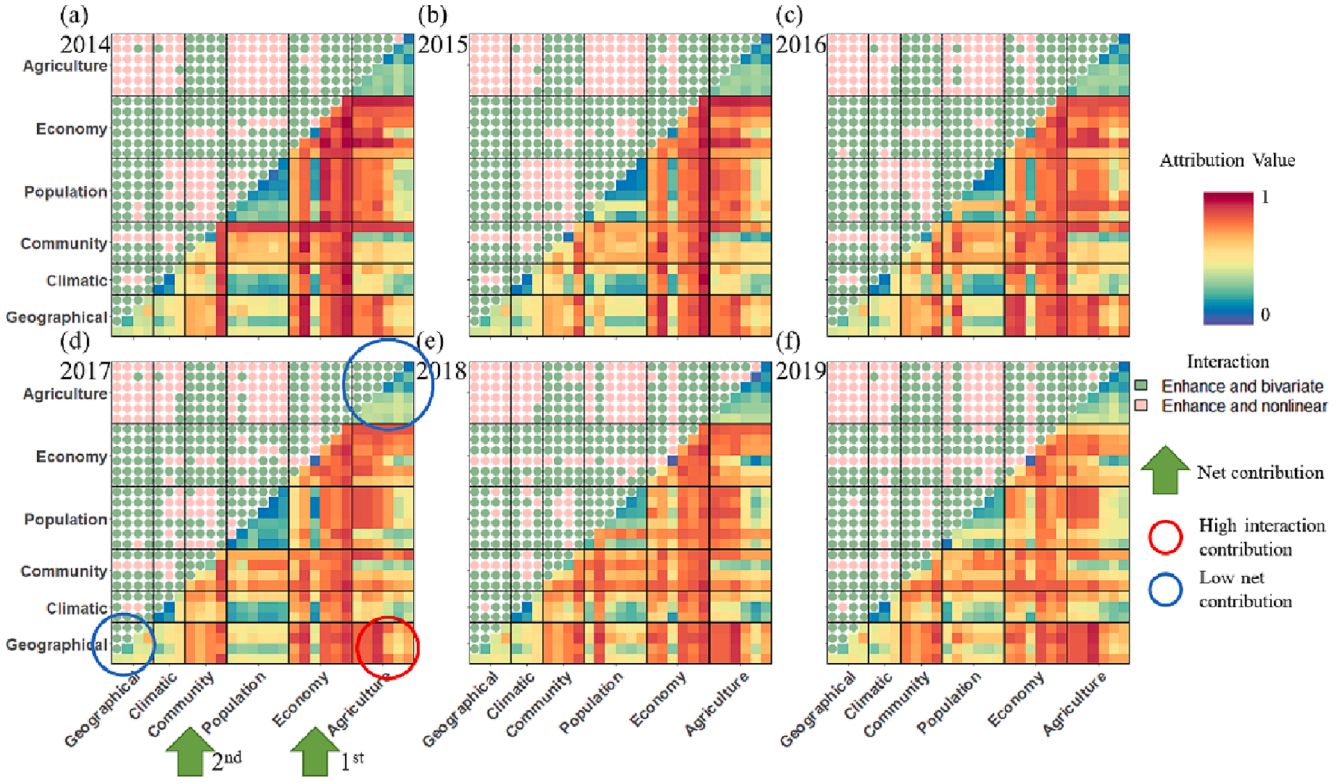


Fig. 6. Attribution of geographical, climatic and social-economic factors to urbanization and related factor interactions over Guanzhong Plain from 2014 to 2019.

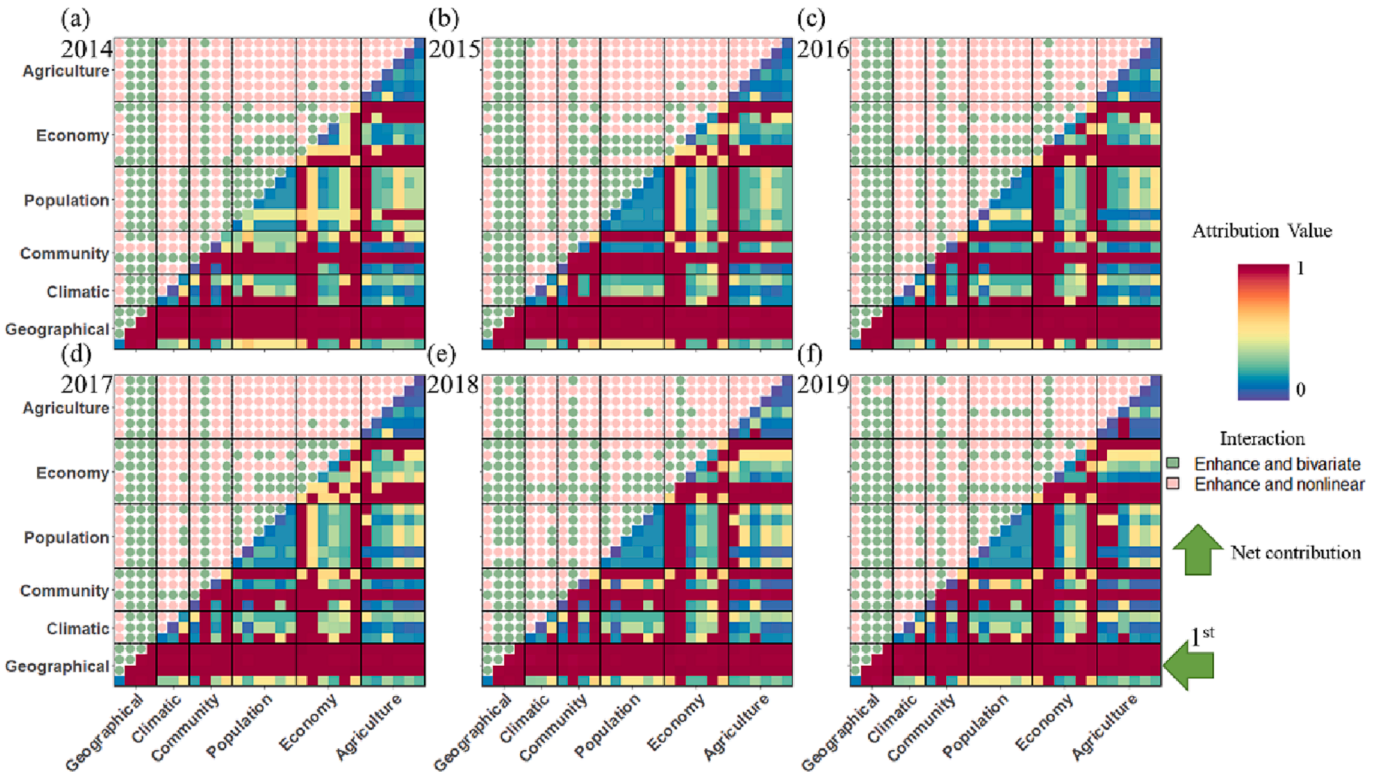


Fig. 7. Attribution of geographical, climatic and social-economic factors to urbanization and related factor interactions over Shannan Mountain region from 2014 to 2019.

contributions to urbanization show the following ranking: economy > community > population > geographical > climatic > agriculture. ii) For inter-contributions from two factors to urbanization, high contribution values are diagnosed as combinations with economy and/or community (dark red in the left, approaching to the maxima *1*). iii) In addition, although most of the agriculture related factors are linked with low contribution values, they show nonlinearity as enhanced interaction when being associated with other factors (see pink in the top). iv) As time increases (2014 to 2019), the contributions from economy and community factors to urbanization decrease (see dark red getting lighter or even turning yellow in the right of Fig. 4).

High contribution factors vary on the subregion scale (Figs. 5-7, representing contribution dynamics and factor interactions in Northern Shaanxi Plateau or Shanbei region, Guanzhong Plain and Shannan Mountain region):

i) For Northern Shaanxi Plateau or Shanbei region (Fig. 5), the high contribution values (*~0.5*) are mainly due to economy factors in 2014 (Fig. 5a). But as time continues, all factors, that is population, economy, community, geographical, agriculture, climatic, contribute almost equally in 2016 (Fig. 5c). Economy factors dominate the contributions to urbanization again in 2018 (Fig. 5e), but they are surpassed by the population factors in 2019 (Fig. 5f). In addition, enhanced nonlinear interaction between two factors is widely distributed in 2014 (see pink in the left), but it keeps decreasing until 2019 with only climatic and some of the agricultural factors remaining.

ii) For Guanzhong Plain (Fig. 6), high contribution values show similarity with the provincial scale which are contributed by economy and community factors (see dark red in the row of community and economy, approaching the maxima of *1*). With time progressing (from 2014 to 2019), the contributions from economy and community factors to urbanization decrease (dark red getting lighter or even turning yellow), while the contributions from other factors increase (blue getting lighter or even yellow). And for Shannan Mountain region (Fig. 7), geographical factors contribute too obvious to notice the other factors, $q = \sim 1$.

4. Discussion

4.1. Model assumption

Area changes of continuous nighttime light represent the spatial extent of intensively used anthropogenic settlements, a further expansion of which depends on factors from geographical, climatic, social-economic reasons. The contribution results from those independent factors to the nighttime lighted urban are obtained from both Geodetector method and Random Forest with the aim to be of possible assistance to municipal governments and city stakeholders to city-design with geographical, climatic and social-economic changes and interactions in mind. Thus, an underlying assumption is the distribution of urbanization and/or its expansion exhibits a spatial similarity with the distribution of those factors. The greater the influence of the factor is, the higher the contribution value. This spatial 'factor-urbanization' consistency of geographical, climatic or social-economic factors has been indicated by previous research (Christensen and McCord, 2016; Henderson et al., 2017; Ma et al., 2021). For example, social economy, education, secondary industry and living environment are found to be the four dominating factors on the development of the Yangtze River Delta from 2007 to 2016 (Ma et al., 2021).

4.2. Comparison: Urban extraction and attribution

Nighttime lighted data are used to defining urban area, which bypasses the question that what land cover types should be listed into urban. With the same or modified DN-threshold, interannual urban development is thus comparable. And this methodology for defining urban with nighttime light data is frequently used (see (Liu et al., 2012; Yu et al., 2021; Zou et al., 2017)).

The closest area method determines urban regions with area

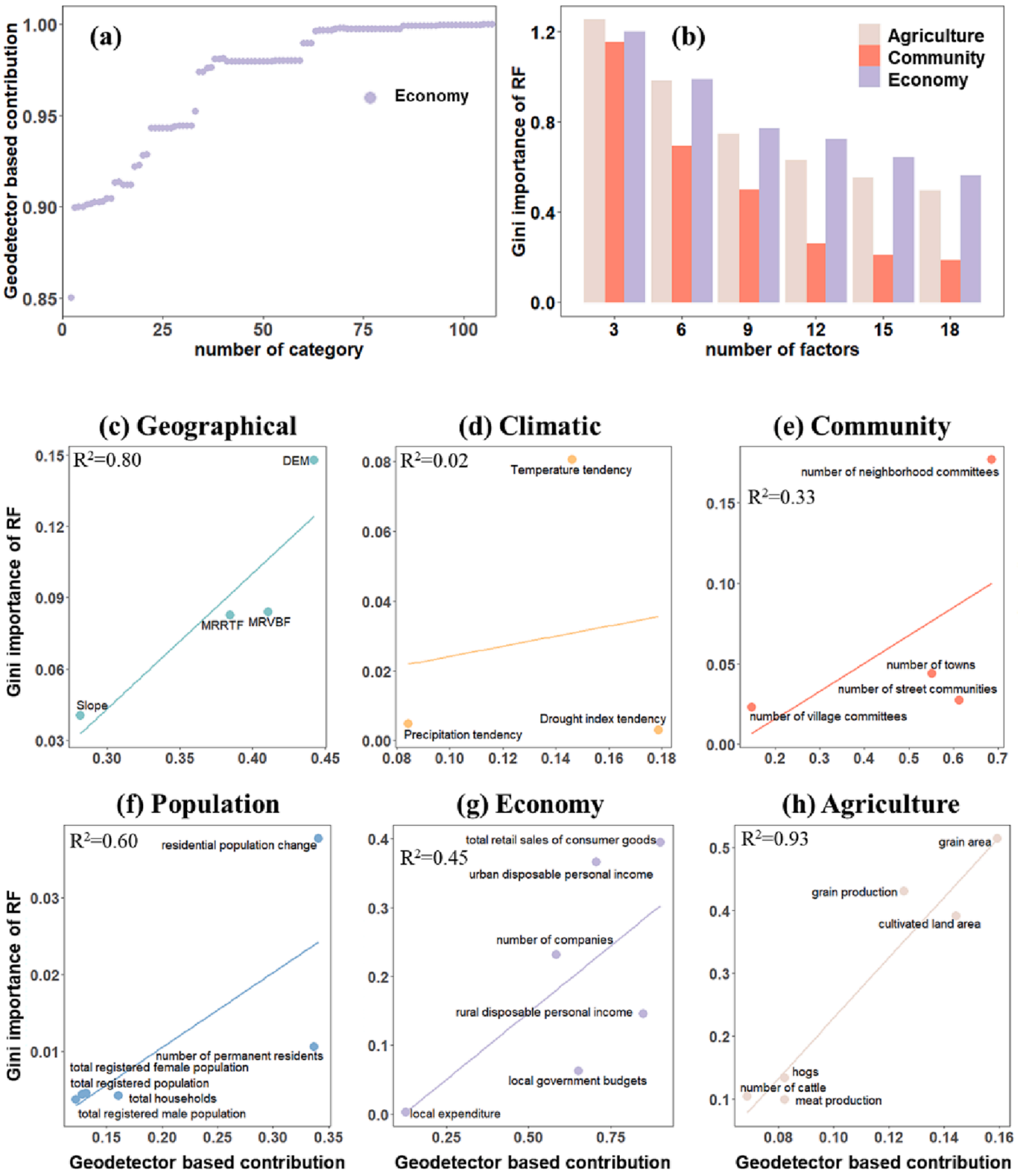


Fig 8. Comparison is carried out between Geodetector based contribution values and Random Forest (RF) based Gini importance: (a) Geodetector based contribution values changing with increasing number of categories, (b) the Gini importance of RF decreases as the number of independent variables increases, and diagrams of Random Forest based Gini importance versus Geodetector based contribution values for all categories (c-h, linear regression coefficients R^2, selected factors see Table 3).

similarity between nighttime light data and reference image. The maximum contribution method assumes the distribution of urbanization and/or its expansion exhibit a similar spatial distribution of nighttime lighted area. Here, nighttime lighted data in each pixel is defined by an intensity (referred to *DN* values). A suitable data discretization is necessary to define whether a pixel is or is not an urban area; that is, if it is above or below a particular *DN*-threshold. Thus, all the *DN*-threshold values (*600* trials) are tested by stepwise increasing *DN* from *0* to *60* in intervals of *0.1*. The greater the similarity of the nighttime lighted area is, compared with the fine urban land use distribution, the higher the contribution value. Pixels above the *DN*-threshold of the highest contribution value are defined as urban area.

4.3. Net and interactive attribution

More topographic indicators are tested (not shown), including digital elevation model (DEM), slope, aspect, convergence index, multi-resolution index of the ridge top flatness (MrRTF), multi-resolution index of valley bottom flatness (MrVBF), Terrain Ruggedness Index, Topographic Position Index and Topographical Wetness Index. However, Convergence Index and Topographic Position Index did not pass the significant test of the Geodetector contribution analyses. Furthermore, none of those factors show a higher contribution than the original local DEM, and only MrRTF and MrVBF indicate similar contribution values (*0.44*, *0.41* and *0.38*). That is, except for DEM, the results provide new insights that high flat areas and flat valley bottoms provide a high contribution to the urbanization process, particularly in mountainous region (e.g., Shannan Mountain region, *0.97*, *0.98*, respectively).

Although the net attribution from geographical and agricultural factors (blue circle in Fig. 4) in Shaanxi Province is low (*~0.37* and *~ 0.11)*, interactive attribution from both geographical and agricultural factors (red circle in Fig. 4) lead to an obvious improvement with the interactive contribution value rising up to *0.78*. Similar enhanced interactive contribution value could also be diagnosed in the flat urban region, Guanzhong Plain, but can not be diagnosed in Northern Shaanxi Plateau and Southern Shannan Mountain region. That is, the flat urban regions are significantly affected by an enhanced contribution of the combination of geographical with agricultural factors. And this is probably due to the competitive relationship between urban and agricultural land use.

4.4. Model priority and uncertainties

Geodetector is, for the first time, applied to define the nighttime lighted urban area by selecting the optimal DN-threshold, which provides the largest contribution to the reference landcover data. Geodetector requires all the independent variables to be categorized. This is suitable for defining the nighttime lighted urban which should also be categorized, namely above and below the DN-threshold representing the urban or non-urban domain. Furthermore, the selected nighttime lighted urban area is tested for its accuracy with present land cover data and compared with frequently used methods for diagnosing nighttime lighted urban boundary. The results indicate a better performance of the Geodetector over flat regions.

The novel comparison of the Geodetector based analysis with the Random Forest scheme reveals the following results:

i) Although the Geodetector is a widely used method, it should be used with some caution because all the independent variables need to be categorized. With increasing number of categories, the contribution values show ‘stair-case like’ increases (Fig. 8a). Thus, it is hard to define the most suitable number for categorizing. That is, valuable information may be inevitably lost during the categorization procedure, and thus it may be misleading the final contribution results.
ii) Random forest (RF) is a widely used machine learning method, and an increasing Gini importance of RF provides a measure of factor relevance of reference data without any categorizing (Jin et al., 2018). However, the model is less stable as the number of factors increases. That is, the Gini importance of RF decreases as the number of independent variables increases (Fig. 8b).
iii) Comparison is carried out between Geodetector based contribution values and Random Forest based Gini importance (Fig. 8 c-h). For Geodetector, all the independent variables are categorized into five classes using Natural break. The linear regression coefficients R2 between Geodetector based contribution value and the RF based Gini importance show: Agriculture (0.93) > Geographical (0.8) > Population (0.6) > Economy (0.45) > Community (0.33) > Climatic (0.02).

4.5. Limitation and future work

Geographical, climatic or social-economic factors and urbanization are mutually influencing each other in the development of a city and disparities suggest different opportunities for the urbanization process, meanwhile urbanization indicates influences on land surface parameter changes (e.g., geographical, ecological, and climatical parameters), as well as changes on local investment. However, in initializing a city, geographical, climatic and social-economic factors are of importance. For example, Christensen and McCord (2016) confirm the role of three factors on shaping and constraining urbanization distribution, including biophysical land suitability for agriculture, distance to major ports and terrain slope, with which they explain ~ 50% of the urbanization distribution variations. Thus, this research is carried out in one direction according to the definition of urbanization as a process whereby populations move from rural to urban areas, enabling cities and towns to grow, and this process typically brings the need for more housing and jobs associating with a need for land-use change. The influence from the urbanization to geographical, climatic and social-economic factors will be considered in the future work (Christensen and McCord, 2016).

5. Conclusion

Unlike trying to understand urbanization induced changes, this work is to understand and examine the driving mechanisms of urbanization, as well as their model uncertainties. This is important because it is the first step towards a holistic approach, embedding economic development, social change, ecological environment, soil and water resources and other fields, which provides guidance for planning a sustainable urban development. Beyond a statistical diagnostic between resultant outcome (the dependent variable *Y*) and driving factors (the independent variable *X*), Geodetector and Random Forest are employed to assess – through spatial heterogeneity – the relationship between resultant outcome and driving factors. Moreover, urbanization is an evolving process neither does it remain constant nor do the contributions from the driving factors of the urbanization. To capture the contribution dynamics from driving factors and from factor interactions to the changing urbanization, a novel three-step spatial–temporal attribution analysis has been designed and applied to Shaanxi Province, which includes nighttime light data based on the long-term urbanization, the change attribution of geographical, climatic and social-economic factors to this long-term urbanization process; and it includes model uncertainties by comparing attribution to urbanization between Geodetector and Random Forest. The main conclusions are noted::

i) Nighttime light data based long-term urbanization has stably increased in Shaanxi Province from 2014 to 2020. Selected methods for extracting the long-term urbanization measures include the closest area method, which is more suitable to be used in the mountainous region, while the Geodetector based maximum contribution method shows improved accuracy

Table A1
Detailed information for social-economic indicators.

Annual data	Impact Factors
Four Community indicators	number of towns; number of street communities; number of village committees; number of neighborhood committees;
Six Population indicators	number of permanent residents; residential population change; total households; total registered population; total registered male population; total registered female population
Nine Economy indicators	number of companies; gross domestic product (GDP); GDP growth rate; local government budgets; local expenditure; urban disposable personal income; rural disposable personal income; total retail sales of consumer goods; growth in total retail sales of consumer goods
Twenty-one Agriculture indicators	gross output value of Farming, Forestry, Animal Husbandry and Fishery (FFAF); gross output value change in FFAF; cultivated land area; agricultural energy consumption; chemical fertilizers consumption, agricultural sheeting consumption; grain area; grain production; oil production; cotton production; vegetable production; fruit production; apples production; meat production; egg production; milk production; number of cattle, milk cow, hogs, sheep and poultry.

notably in flat urban land. The peak in the maximum contribution method represents the strong spatial similarity between the reference and the nighttime light.

ii) The most obvious contribution dynamics depends on local urban development strategies associated with government investment (see economy factors). For example, urban regions over Guanzhong Plain (mountainous Northern Shaanxi Plateau or Shanbei region) obtain an early (late) development with high contributing values from economy factors since 2014 (2018).

iii) Controlling factors and their contributions vary in flat (Guanzhong Plain) and mountainous urban regions (Northern Shaanxi Plateau or Shanbei region and Shannan Mountain region) from 2014 to 2020. For example, high flat areas and flat valley bottoms provide great contribution to the urbanization process, particularly in mountainous region (e.g., Shannan Mountain region, 0.97, 0.98, respectively).

iv) Besides government investment, population factors attain high contribution values in mountainous regions of low population. Migration into the Northern Shaanxi Plateau or Shanbei mountainous urban regions, which is due to mining of local coal resources, contributes to the local urbanization with the highest value, occurring in 2019.

v) Geodetector based contribution values increase 'staircase-like' with increasing numbers of categorized classes of independent variables. For Random Forest, the Gini importance decreases as the numbers of independent variables increase. In this sense, we provide a novel uncertainty analysis scheme for the Geodetector, which is based on the suitable choice of the categorizing number and the number of independent factors.

Generally, geographical and climatic factors contribute highly to the urbanization process in its very beginning (of the establishment of a city), but they are contributing less during the present-day urbanization process when compared with economic and even population factors in the mountainous regions.

CRediT authorship contribution statement

Siyi Huang: Conceptualization, Software, Validation, Formal analysis, Investigation, Writing – original draft, Visualization. **Lijun Yu:** Writing – review & editing. **Danlu Cai:** Conceptualization, Resources, Writing – original draft, Writing – review & editing, Supervision. **Jianfeng Zhu:** Writing – review & editing. **Ze Liu:** Writing – review & editing. **Zongke Zhang:** Writing – review & editing. **Yueping Nie:** Writing – review & editing. **Klaus Fraedrich:** Writing – review & editing.

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.

Data availability

Data will be made available on request.

Acknowledgements

Supported by the National Key Research and Development Program of China (No. 2019YFD1100705).

Appendix A

References

Abatzoglou, J.T., Dobrowski, S.Z., Parks, S.A., Hegewisch, K.C., 2018. TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. Sci. Data 5, 170191.
Adamic, L., Huberman, B., Barabasi, A.-L., Albert, R., Jeong, H., Bianconi, G., 2000. Power-Law Distribution of the World Wide Web. Science 287, 2115.
Ahmad, M., Jiang, P., Murshed, M., Shehzad, K., Akram, R., Cui, L., et al., 2021. Modelling the dynamic linkages between eco-innovation, urbanization, economic growth and ecological footprints for G7 countries: Does financial globalization matter? Sustain. Cities Soc. 70, 102881.
Ahmed, Z., Wang, Z., Ali, S., 2019. Investigating the non-linear relationship between urbanization and CO2 emissions: An empirical analysis. Air Qual. Atmos. Health 12, 945–953.
Avtar, R., Tripathi, S., Aggarwal, A.K., Kumar, P., 2019. Population–Urbanization–Energy Nexus: A Review. Resources 8, 136.
Bakirtas, T., Akpolat, A.G., 2018. The relationship between energy consumption, urbanization, and economic growth in new emerging-market countries. Energy 147, 110–121.
Bilgili, F., Koçak, E., Bulut, Ü., Kuloğlu, A., 2017. The impact of urbanization on energy intensity: Panel data evidence considering cross-sectional dependence and heterogeneity. Energy 133, 242–256.
Bongaarts J. Development: Slow down population growth. Nature 2016; 530: 409-412.
Breiman, L., 2001. Random Forests. Mach. Learn. 45, 5–32.
Cai, W., Fangyuan, T., 2020. Spatiotemporal characteristics and driving forces of construction land expansion in Yangtze River economic belt. China. PLOS ONE 15, e0227299.
Cai, D., Fraedrich, K., Guan, Y., Guo, S., Zhang, C., 2017. Urbanization and the thermal environment of Chinese and US-American cities. Sci. Total Environ. 589, 200–211.
Carmona, P., Tran, D., Pla, F., Myint, S., Caetano, M., Kieu, H., 2017. Characterizing the relationship between land use land cover change and land surface temperature. ISPRS J. Photogramm. Remote Sens. 124, 119–132.
Chen J, Wang D, Li G, Sun Z, Wang X, Zhang X, et al. Spatial and Temporal Heterogeneity Analysis of Water Conservation in Beijing-Tianjin-Hebei Urban Agglomeration Based on the Geodetector and Spatial Elastic Coefficient Trajectory Models. GeoHealth 2020; 4: e2020GH000248.
Chen, J., Guo, F., Wang, H., Wang, Z., Wu, Y., 2018. Urban Land Revenue and Sustainable Urbanization in China: Issues and Challenges. Sustainability 10, 2111.
Christensen, P., McCord, G.C., 2016. Geographic determinants of China's urbanization. Reg. Sci. Urban Econ. 59, 90–102.
Chunyu, W., Xiaofang, S., Meng, W., Junbang, W., Qingfu, D., 2019. Chinese Cropland Quality and Its Temporal and Spatial Changes due to Urbanization in 2000–2015. Journal of Resources and Ecology 10 (174–183), 10.
Dou, Y., Liu, Z., He, C., Yue, H., 2017. Urban Land Extraction Using VIIRS Nighttime Light Data: An Evaluation of Three Popular Methods. Remote Sens. (Basel) 9, 175.
Du, Y., Wan, Q., Liu, H., Liu, H., Kapsar, K., Peng, J., 2019. How does urbanization influence PM2.5 concentrations? Perspective of spillover effect of multi-dimensional urbanization impact. J. Clean. Prod. 220, 974–983.
Elheddad, M., Djellouli, N., Tiwari, A.K., Hammoudeh, S., 2020. The relationship between energy consumption and fiscal decentralization and the importance of urbanization: Evidence from Chinese provinces. J. Environ. Manage. 264, 110474.
Feng, T.-t., Yang, Y.-s., Xie, S.-y., Dong, J., Ding, L., 2017. Economic drivers of greenhouse gas emissions in China. Renew. Sustain. Energy Rev. 78, 996–1006.

Fu, P., Weng, Q., 2016. A time series analysis of urbanization induced land use and land cover change and its impact on land surface temperature with Landsat imagery. Remote Sens. Environ. 175, 205–214.
Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., et al., 2015. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. Sci. Data 2, 150066.
Gallant, J.C., Dowling, T.I., 2003. A multiresolution index of valley bottom flatness for mapping depositional areas. Water Resour. Res. 39.
Gao, J., Li, S., Zhao, Z., Cai, Y., 2012. Investigating spatial variation in the relationships between NDVI and environmental factors at multi-scales: a case study of Guizhou Karst Plateau. China. International Journal of Remote Sensing 33, 2112–2129.
Gaughan, A.E., Stevens, F.R., Huang, Z., Nieves, J.J., Sorichetta, A., Lai, S., et al., 2016. Spatiotemporal patterns of population in mainland China, 1990 to 2010. Sci. Data 3, 160005.
Guo J, Yu Z, Ma Z, Xu D, Cao S. What factors have driven urbanization in China? Environment, Development and Sustainability 2021.
Haregeweyn, N., Fikadu, G., Tsunekawa, A., Tsubo, M., Meshesha, D.T., 2012. The dynamics of urban expansion and its impacts on land use/land cover change and small-scale farmers living near the urban fringe: A case study of Bahir Dar, Ethiopia. Landscape and Urban Planning 106, 149–157.
Henderson, J.V., Storeygard, A., Deichmann, U., 2017. Has climate change driven urbanization in Africa? J. Dev. Econ. 124, 60–82.
Hou, K., Wen, J., 2020. Quantitative analysis of the relationship between land use and urbanization development in typical arid areas. Environ. Sci. Pollut. Res. 27, 38758–38768.
Huang, L., Shahtahmassebi, A., Gan, M., Deng, J., Wang, J., Wang, K., 2020. Characterizing spatial patterns and driving forces of expansion and regeneration of industrial regions in the Hangzhou megacity, China. Journal of Cleaner Production 253, 119959.
Jenks GF. The Data Model Concept in Statistical Mapping, 1967.
Jiang, Z., Huo, F., Ma, H., Song, J., Dai, A., 2017. Impact of Chinese Urbanization and Aerosol Emissions on the East Asian Summer Monsoon. J. Clim. 30, 1019–1039.
Jiao, L., 2015. Urban land density function: A new method to characterize urban expansion. Landsc. Urban Plan. 139, 26–39.
Jin, Y., Liu, X., Chen, Y., Liang, X., 2018. Land-cover mapping using Random Forest classification and incorporating NDVI time-series and texture: a case study of central Shandong. Int. J. Remote Sens. 39, 8703–8723.
Ju, H., Zhang, Z., Zuo, L., Wang, J., Zhang, S., Wang, X., et al., 2016. Driving forces and their interactions of built-up land expansion based on the geographical detector – a case study of Beijing, China. Int. J. Geogr. Inf. Sci. 1–20.
Li, Q., Lu, L., Weng, Q., Xie, Y., Guo, H., 2016. Monitoring Urban Dynamics in the Southeast U.S.A. Using Time-Series DMSP/OLS Nightlight Imagery. Remote Sens. (Basel) 8, 578.
Li, G., Sun, S., Fang, C., 2018. The varying driving forces of urban expansion in China: Insights from a spatial-temporal analysis. Landsc. Urban Plan. 174, 63–77.
Lin, L., Gao, T., Luo, M., Ge, E., Yang, Y., Liu, Z., et al., 2020. Contribution of urbanization to the changes in extreme climate events in urban agglomerations across China. Sci. Total Environ. 744, 140264.
Lin, Y., Zou, J., Yang, W., Li, C.-Q., 2018. A Review of Recent Advances in Research on PM2.5 in China. Int. J. Environ. Res. Public Health 15, 438.
Liu, X., Bae, J., 2018. Urbanization and industrialization impact of CO2 emissions in China. J. Clean. Prod. 172, 178–186.
Liu, Z., He, C., Zhang, Q., Huang, Q., Yang, Y., 2012. Extracting the dynamics of urban expansion in China using DMSP-OLS nighttime light data from 1992 to 2008. Landsc. Urban Plan. 106, 62–72.
Liu, X., Hu, G., Chen, Y., Li, X., Xu, X., Li, S., et al., 2018. High-resolution multi-temporal mapping of global urban land using Landsat images based on the Google Earth Engine Platform. Remote Sens. Environ. 209, 227–239.
Liu, Q., Jin, Z., Meng, Q., Wu, X., Jia, H., 2015. Genetic types of natural gas and filling patterns in Daniudi gas field, Ordos Basin. China. Journal of Asian Earth Sciences 107, 1–11.
Liu, S., Liao, Q., Liang, Y., Li, Z., Huang, C., 2021b. Spatio-Temporal Heterogeneity of Urban Expansion and Population Growth in China. Int. J. Environ. Res. Public Health 18, 13031.
Liu, J., Niyogi, D., 2019. Meta-analysis of urbanization impact on rainfall modification. Sci. Rep. 9, 7301.
Liu, J., Schlünzen, K.H., Frisius, T., Tian, Z., 2021a. Effects of urbanization on precipitation in Beijing. Physics and Chemistry of the Earth, Parts A/B/C 122, 103005.
Liu, C., Yang, K., Bennett, M.M., Guo, Z., Cheng, L., Li, M., 2019. Automated Extraction of Built-Up Areas by Fusing VIIRS Nighttime Lights and Landsat-8 Data. Remote Sens. (Basel) 11, 1571.
Luo, J., Zhang, X., Wu, Y., Shen, J., Shen, L., Xing, X., 2018. Urban land expansion and the floating population in China: For production or for living? Cities 74, 219–228.
Ma, J., Cheng, J.C.P., Jiang, F., Chen, W., Zhang, J., 2020. Analyzing driving factors of land values in urban scale based on big data and non-linear machine learning techniques. Land Use Policy 94, 104537.
Ma, W., Tian, W., Zhou, Q., Miao, Q., 2021. Analysis on the Temporal and Spatial Heterogeneity of Factors Affecting Urbanization Development Based on the GTWR Model: Evidence from the Yangtze River Economic Belt. Complexity 2021, 7557346.
Ma, Y., Xu, R., 2010. Remote sensing monitoring and driving force analysis of urban expansion in Guangzhou City. China. Habitat International 34, 228–235.
Ma, T., Zhou, C., Pei, T., Haynie, S., Fan, J., 2012. Quantitative estimation of urbanization dynamics using time series of DMSP/OLS nighttime light data: A comparative case study from China's cities. Remote Sens. Environ. 124, 99–107.
Menze, B.H., Kelm, B.M., Masuch, R., Himmelreich, U., Bachert, P., Petrich, W., et al., 2009. A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data. BMC Bioinf. 10, 213.
Minh Ha N, Nguyen L. The relationship between urbanization and economic growth: An empirical study on ASEAN countries. International Journal of Social Economics 2017; 45: 00-00.
Mundia, C., Aniya, M., 2005. Analysis of land use/cover changes and urban expansion of Nairobi city using remote sensing and GIS. Int. J. Remote Sens. 26, 2831–2849.
Pan, X., Wang, Y., Liu, Z., He, C., Liu, H., Chen, Z., 2021. Understanding Urban Expansion on the Tibetan Plateau over the Past Half Century Based on Remote Sensing: The Case of Xining City, China. Remote Sensing 13, 46.
Parnell, S., Walawege, R., 2011. Sub-Saharan African urbanisation and global environmental change. Glob. Environ. Chang. 21, S12–S20.
Qizhi, M., Long, Y., Wu, K., 2016. Spatio-Temporal Changes of Population Density and Urbanization Pattern in China (2000–2010). China City. Plan. Rev.
Ran, Q., Hao, Y., Xia, A., Liu, W., Hu, R., Cui, X., et al., 2019. Quantitative Assessment of the Impact of Physical and Anthropogenic Factors on Vegetation Spatial-Temporal Variation in Northern Tibet. Remote Sens. (Basel) 11, 1183.
Shang, J., Li, P., Li, L., Chen, Y., 2018. The relationship between population growth and capital allocation in urbanization. Technol. Forecast. Soc. Chang. 135, 249–256.
Shen, J., Zhang, N., Gexigeduren, H.B., Liu, C.-Y., Li, Y., et al., 2015. Construction of a GeogDetector-based model system to indicate the potential occurrence of grasshoppers in Inner Mongolia steppe habitats. Bull. Entomol. Res.
Sheng, P., He, Y., Guo, X., 2017. The impact of urbanization on energy consumption and efficiency. Energy Environ. 28, 673–686.
Small, C., Pozzi, F., Elvidge, C.D., 2005. Spatial analysis of global urban extent from DMSP-OLS night lights. Remote Sens. Environ. 96, 277–291.
Ul Din, S., Mak, H.W.L., 2021. Retrieval of Land-Use/Land Cover Change (LUCC) Maps and Urban Expansion Dynamics of Hyderabad, Pakistan via Landsat Datasets and Support Vector Machine Framework. Remote Sens. (Basel) 13, 3337.
van der Schrier, G., Jones, P.D., Briffa, K.R., 2011. The sensitivity of the PDSI to the Thornthwaite and Penman-Monteith parameterizations for potential evapotranspiration. J. Geophys. Res. Atmos. 116.
Wang, Y., Chen, L., Song, Z., Huang, Z., Ge, E., Lin, L., et al., 2019a. Human-perceived temperature changes over South China: Long-term trends and urbanization effects. Atmos. Res. 215, 116–127.
Wang, J., Chen, Y., Liao, W., He, G., Tett, S.F.B., Yan, Z., et al., 2021a. Anthropogenic emissions and urbanization increase risk of compound hot extremes in cities. Nat. Clim. Chang. 11, 1084–1089.
Wang, Y., Huang, C., Feng, Y., Zhao, M., Gu, J., 2020. Using Earth Observation for Monitoring SDG 11.3.1-Ratio of Land Consumption Rate to Population Growth Rate in Mainland China. Remote Sens. (Basel) 12, 357.
Wang, J.F., Li, X.H., Christakos, G., Liao, Y.L., Zhang, T., Gu, X., et al., 2010. Geographical Detectors-Based Health Risk Assessment and its Application in the Neural Tube Defects Study of the Heshun Region, China. Int. J. Geogr. Inf. Sci. 24, 107–127.
Wang, S., Li, C., 2018. The impact of urbanization on CO2 emissions in China: an empirical study using 1980–2014 provincial data. Environ. Sci. Pollut. Res. 25, 2457–2465.
Wang, Y., Li, X., Kang, Y., Chen, W., Zhao, M., Li, W., 2019b. Analyzing the impact of urbanization quality on CO2 emissions: What can geographically weighted regression tell us? Renew. Sustain. Energy Rev. 104, 127–136.
Wang, P., Luo, M., Liao, W., Xu, Y., Wu, S., Tong, X., et al., 2021b. Urbanization contribution to human perceived temperature changes in major urban agglomerations of China. Urban Clim. 38, 100910.
Wang, Y., Xiang, Y., Song, L., Liang, X.-Z., 2022. Quantifying the Contribution of Urbanization to Summer Extreme High-Temperature Events in the Beijing–Tianjin–Hebei Urban Agglomeration. J. Appl. Meteorol. Climatol.
Wang, Z., Xiao, Z., Tam, C.-Y., Pan, W., Chen, J., Hu, C., et al., 2021c. The projected effects of urbanization and climate change on summer thermal environment in Guangdong-Hong Kong-Macao Greater Bay Area of China. Urban Clim. 37, 100866.
Wang, J.-F., Zhang, T.-L., Fu, B.-J., 2016. A measure of spatial stratified heterogeneity. Ecol. Ind. 67, 250–256.
Wu, H., Lin, A., Xing, X., Song, D., Li, Y., 2021. Identifying core driving factors of urban land use change from global land cover products and POI data using the random forest method. Int. J. Appl. Earth Obs. Geoinf. 103, 102475.
Wu, J., Zheng, H., Zhe, F., Xie, W., Song, J., 2018. Study on the relationship between urbanization and fine particulate matter (PM2.5) concentration and its implication in China. J. Clean. Prod. 182, 872–882.
Xie, W., Wu, W.-Z., Liu, C., Zhao, J., 2020. Forecasting annual electricity consumption in China by employing a conformable fractional grey model in opposite direction. Energy 202, 117682.
Xue, R., Ai, B., Lin, Y., Pang, B., Shang, H., 2019. Spatial and Temporal Distribution of Aerosol Optical Depth and Its Relationship with Urbanization in Shandong Province. Atmos. 10, 110.
Yang, Z., Lei, J., Li, J.-G., 2019. Identifying the Determinants of Urbanization in Prefecture-Level Cities in China: A Quantitative Analysis Based on Spatial Production Theory. Sustainability 11, 1204.
Yang, R., Xu, Q., Long, H., 2016. Spatial distribution characteristics and optimized reconstruction analysis of China's rural settlements during the process of rapid urbanization. J. Rural. Stud. 47, 413–424.
Ye, J., Hu, Y., Zhen, L., Wang, H., Zhang, Y., 2021. Analysis on Land-Use Change and Its Driving Mechanism in Xilingol, China, during 2000–2020 Using the Google Earth Engine. Remote Sens. (Basel) 13.

Yu, B., 2021. Ecological effects of new-type urbanization in China. Renew. Sustain. Energy Rev. 135, 110239.
Yu, M., Guo, S., Guan, Y., Cai, D., Zhang, C., Fraedrich, K., et al., 2021. Spatiotemporal Heterogeneity Analysis of Yangtze River Delta Urban Agglomeration: Evidence from Nighttime Light Data (2001–2019). Remote Sens. (Basel) 13, 1235.
Zhang, Z., Liu, F., Zhao, X., Wang, X., Shi, L., Xu, J., et al., 2018b. Urban Expansion in China Based on Remote Sensing Technology: A Review. Chin. Geogr. Sci. 28, 727–743.
Zhang, D., Liu, X., Wu, X., Yao, Y., Wu, X., Chen, Y., 2019a. Multiple intra-urban land use simulations and driving factors analysis: a case study in Huicheng. China. GIScience & Remote Sensing 56, 282–308.
Zhang, Q., Liu, S., Wang, T., Dai, X., Baninla, Y., Nakatani, J., et al., 2019b. Urbanization impacts on greenhouse gas (GHG) emissions of the water infrastructure in China: Trade-offs among sustainable development goals (SDGs). J. Clean. Prod. 232, 474–486.
Zhang, X., Liu, L., Wu, C., Chen, X., Gao, Y., Xie, S., et al., 2020. Development of a global 30 m impervious surface map using multisource and multitemporal remote sensing datasets with the Google Earth Engine platform. Earth Syst. Sci. Data 12, 1625–1648.
Zhang, W., Villarini, G., Vecchi, G.A., Smith, J.A., 2018a. Urbanization exacerbated the rainfall and flooding caused by hurricane Harvey in Houston. Nature 563, 384–388.
Zhao, J., Tang, J., 2018. Industrial structure change and economic growth: A China-Russia comparison. China Econ. Rev. 47, 219–233.
Zhao, Y., Tomita, M., Hara, K., Fujihara, M., Yang, Y., Da, L., 2014. Effects of topography on status and changes in land-cover patterns, Chongqing City, China. Landscape and Ecological Engineering 10, 125–135.
Zheng, W., Walsh pp, 2019. Economic growth, urbanization and energy consumption — A provincial level analysis of China. Energy Econ. 80, 153–162.
Zhou, D., Xiao, J., Bonafoni, S., Berger, C., Deilami, K., Zhou, Y., et al., 2019. Satellite Remote Sensing of Surface Urban Heat Islands: Progress, Challenges, and Perspectives. Remote Sens. (Basel) 11, 48.
Zou, Y., Peng, H., Liu, G., Yang, K., Xie, Y., Weng, Q., 2017. Monitoring Urban Clusters Expansion in the Middle Reaches of the Yangtze River, China, Using Time-Series Nighttime Light Images. Remote Sens. (Basel) 9.