A Simulation Study on the Urban Population of China Based on Nighttime Light Data Acquired from DMSP/OLS

Qingxu Huang [1], Yang Yang [2],*, Yajing Li [2] and Bin Gao [1,3]

[1] Center for Human-Environment System Sustainability, State Key Laboratory of Earth Surface Processes and Resource Ecology, Beijing Normal University, 19 Xinjiekouwai Street, Beijing 100875, China; qxhuang@bnu.edu.cn (Q.H.); Bing0402@126.com (B.G.)

[2] Teaching and Research Section of Land Resources Management, Department of Public Administration, Law & Politics School, Ocean University of China, 238 Songling Road, Qingdao 266100, China; lijing6520@126.com

[3] College of Resources Science & Technology, Beijing Normal University, 19 Xinjiekouwai Street, Beijing 100875, China

* Correspondence: yangyang@ouc.edu.cn; Tel.: +86-523-6678-1123

Abstract: The urban population (UP) measure is one of the most direct indicators that reflect the urbanization process and the impacts of human activities. The dynamics of UP is of great importance to studying urban economic, social development, and resource utilization. Currently, China lacks long time series UP data with consistent standards and comparability over time. The nighttime light images from the Defense Meteorological Satellite Program's (DMSP) Operational Linescan System (OLS) allow the acquisition of continuous and highly comparable long time series UP information. However, existing studies mainly focus on simulating the total population or population density level based on the nighttime light data. Few studies have focused on simulating the UP in China. Based on three regression models (*i.e.*, linear, power function, and exponential), the present study discusses the relationship between DMSP/OLS nighttime light data and the UP and establishes optimal regression models for simulating the UPs of 339 major cities in China from 1990 to 2010. In addition, the present study evaluated the accuracy of UP and non-agricultural population (NAP) simulations conducted using the same method. The simulation results show that, at the national level, the power function model is the optimal regression model between DMSP/OLS nighttime light data and UP data for 1990–2010. At the provincial scale, the optimal regression model varies among different provinces. The linear regression model is the optimal regression model for more than 60% of the provinces. In addition, the comparison results show that at the national, provincial, and city levels, the fitting results of the UP based on DMSP/OLS nighttime light data are better than those of the NAP. Therefore, DMSP/OLS nighttime light data can be used to effectively retrieve the UP of a large-scale region. In the context of frequent population flows between urban and rural areas in China and difficulty in obtaining accurate UP data, this study provides a timely and effective method for solving this problem.

Keywords: urban population; DMSP/OLS; China; non-agricultural population; regression analysis

1. Introduction

Currently, countries worldwide, and developing countries in particular, are undergoing an unprecedented urbanization process [1]. As one of the most direct indicators that reflects the urbanization process and the impacts of human activities, the urban population (UP) measure is

of great importance to urban economic and social development and resource utilization [2]. Since the onset of economic reforms, the proportion of the UP in China has increased from 17.9% in 1978 to 52.6% in 2012. If the current trend continues, the UP in China will reach one billion over the next 20 years [3,4]. Meanwhile, China is implementing a new urbanization strategy, called the "New-Type Urbanization Plan". Therefore, the accuracy of the UP indicator is directly related to judgments on urbanization processes in China. Acquiring information on the UP of China in a timely and accurate manner is of great significance for formulating scientific urban system plans while reasonably optimizing the UP distribution.

Currently, two common methods are used to acquire information on the UP in China. The first method involves acquiring information on the UP through a census. Since the founding of the People's Republic of China, China has conducted six successive national censuses (in 1953, 1964, 1982, 1990, 2000, and 2010), whereby the population was studied and measured on a general, door-to-door and person-by-person basis. These censuses have facilitated the acquisition of relatively accurate statistical data on the UP [5–8]. However, acquiring information on the UP through a census still presents some disadvantages. First, each census was completed at the cost of large amounts of manpower and material resources. Second, the interval between two consecutive censuses was approximately 10 years. Therefore, the data lacks timeliness. As a result, acquiring statistical data on the UP through censuses presents certain limitations regarding research relevance and practical applications. More importantly, of the six censuses conducted, those conducted in 1953, 1964, and 1982 employed a relatively consistent statistical definition for UP that was, however, different from the one used for the censuses conducted in 1990, 2000, and 2010. This difference in statistical definitions is undoubtedly unfavorable to the continuity and comparability of the data [7,9,10].

The second common method involves acquiring information on the UP by substituting the non-agricultural population (NAP) for the UP. Due to their relatively high levels of continuity, data on the NAP have been extensively used in statistical publications and in urban studies in China. However, substituting the NAP for the UP also presents many problems. Most importantly, there is a significant difference between the UP and NAP. The UP cannot be equated precisely with the NAP. By definition, the UP refers to the population that lives within the boundaries of cities and is mainly statistically determined based on the division between urban and rural areas. In comparison, the NAP refers to the population that is engaged in non-agricultural production activities and to its supported population, and it is statistically determined based on the agricultural or non-agricultural status recorded in the household registration system [11]. In 2010, the UP accounted for 49.68% of the total population of China, whereas the NAP accounted for only 34.17% of the total population. The difference between the UP and NAP was approximately 200 million. This difference is mainly attributable to the fact that the numerous migrant workers in cities that originated from rural areas are not recorded in a given city's NAP, and thus, substituting the NAP for the UP often underestimates the real UP. This problem is particularly prominent in large cities with high labor demands but strict control over people with household registrations [12]. Therefore, China currently lacks a long time series set of UP data that are accurate, consistent and comparable over time [6].

The nighttime light image acquired by the Operational Linescan System (OLS) sensor of the US Defense Meteorological Satellite Program (DMSP) provides an alternative method for the acquisition of continuous, highly-comparable long time series information on the UP [13–17]. On the one hand, nighttime light data acquired through the DMSP/OLS reflect surface nighttime light information across the globe [18], exhibit visual and realistic accounts of human activities, and have been successfully used in a broad range of research, including retrieving urban land use and human settlement [19–26], measuring urbanization process [27,28], simulating population density [13–17] and energy consumption [29–32], estimating economic activity [33–35], evaluating urbanization impact [36,37], and as a proxy measure of human well-being [38]. On the other hand, the time series dataset of global DMSP/OLS nighttime lights published by the National Geophysical Data Center (NCDC) of the US National Oceanic and Atmospheric Administration (NOAA) includes annual

data from 1992 to 2013 and, thus, presents outstanding advantages in terms of providing continuous and highly comparable urban information [25,27,28,39–42]. However, existing studies still mainly focus on simulations of the total population or population density based on nighttime light data [13–17]. Few studies focus on simulating the UP.

Therefore, the present study aims to develop a means of acquiring long time series information on the UP in China based on DMSP/OLS nighttime light data. The method can account for current deficiencies in statistical UP data for China. Specifically, the main objectives of the present study are to (1) compare and discuss the relationship between DMSP/OLS nighttime light data and the UP based on multiple regression models and to (2) simulate the UP of 339 major cities in China for 1990 to 2010 based on established optimal regression models and to evaluate the accuracy of the simulation results.

2. Study Area and Data

Three main types of data were used in the present study: DMSP/OLS stable nighttime light data, statistical data, and geographical information system (GIS) auxiliary data (Figure 1). The first type of data includes DMSP/OLS stable nighttime light data, which are provided by the NGDC of the NOAA. DMSP/OLS stable nighttime light data are data on stable lights emitted from urban areas, rural areas, and other locations, excluding occasional noise such as flames, and have a spatial resolution of 30" (curvature). The value of DMSP/OLS stable nighttime light data represents the mean light intensity (range: 1–63). DMSP/OLS stable nighttime light data are created through the strict screening of all usable data archived by the DMSP/OLS each year and through the de-clouding of selected data. As DMSP/OLS time series data are collected by several sensors from different satellites, and are not calibrated against radiation, the difference between different sensors, the difference in transit time between different satellites and the deterioration of sensors have a certain effect on the comparability of the nighttime light data. Therefore, a relatively large difference can be found between data acquired by different satellites and data acquired by the same satellite over different years. In other words, the data are not comparable, thereby limiting the practical application of nighttime light data [39–41]. Following the method proposed by Liu *et al.* [25], the data were calibrated from 1992 to 2010 of mainland China. Specifically, stable nighttime light data were first subjected to a series of calibration treatments, including relative radiation calibration, intra-annual calibration and inter-annual series calibration, to enhance the continuity and comparability of the data. The grid resolution was resampled to 1 km, and projections were converted to Albers equal-area projections.
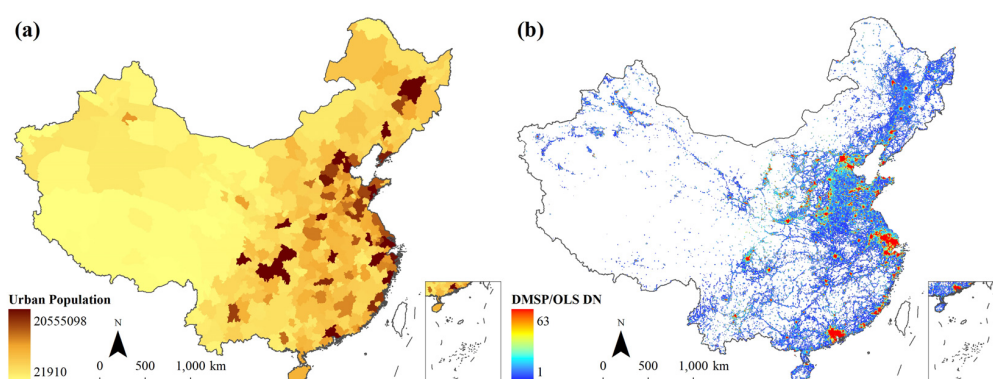
Figure 1. Urban population (UP) and the digital number (DN) of US Defense Meteorological Satellite Program's Operational Linescan System (DMSP/OLS) nighttime light data in mainland China for 2010.

The second type of data used is statistical data on the UP and NAP in China for 1990, 2000, and 2010. Considering the lack of statistical data on the populations of Hong Kong, Macao, and Taiwan, we examined 339 major cities in 31 provinces and regions in China as our study area. In the present study, census data were used to collect UP data; and the *Yearbooks of Statistics* of Chinese cities

were used for NAP data [43]. The third type of data is GIS auxiliary data, which are 1:4,000,000 data of the administrative boundaries of provinces and prefecture-level cities in China published on the *National Geomatics Center of China* website [44].

3. Methods

3.1. Regression Method

Regression analysis is a common method used to establish the relationship between the total digital number (TDN) and population data based on DMSP/OLS nighttime light data. Accordingly, urban populations in 1990, 2000, and 2010 were simulated with regression models. As the DMSP/OLS dataset recorded the nighttime light from 1992, we used the nighttime light data in 1992 and the census data in 1990 to develop the simulation model for urban population in 1990. In accordance with the literature [45], we discussed the relationship between the DMSP/OLS nighttime light and UP using three regression models—the linear regression model, the exponential regression model, and the power function regression model. The equations used were as follows:

$$UP_i = a + b \cdot TDN_i \tag{1}$$

$$UP_i = m \cdot e^{n \cdot TDN}{}_i \tag{2}$$

$$UP_i = p \cdot TDN_i^q \tag{3}$$

where UP_i represents the statistical value of the UP of city i; TDN_i represents the total light brightness of city i; and a, b, m, n, p, and q represent corresponding parameters of the regression equations.

3.2. Optimal Regression Model

Due to the significant regional heterogeneities of mainland China, we selected the optimal regression models at two scales: the national and provincial scales (Figure 2). At the national scale, regressions were conducted between the UP and the TDN in 339 major cities of mainland China. At the provincial scale, the regression analysis used the UP and the TDN in each province. Since directly controlled municipalities (*i.e.*, Beijing, Tianjin, Shanghai, and Chongqing) in China are also provincial-level administrative units, data on the UPs and TDNs in Beijing and Tianjin were combined with the counterparts in Hebei Province; data in Shanghai were combined with the counterparts in Jiangsu Province; and data in Chongqing were combined with the counterparts in Sichuan Province.
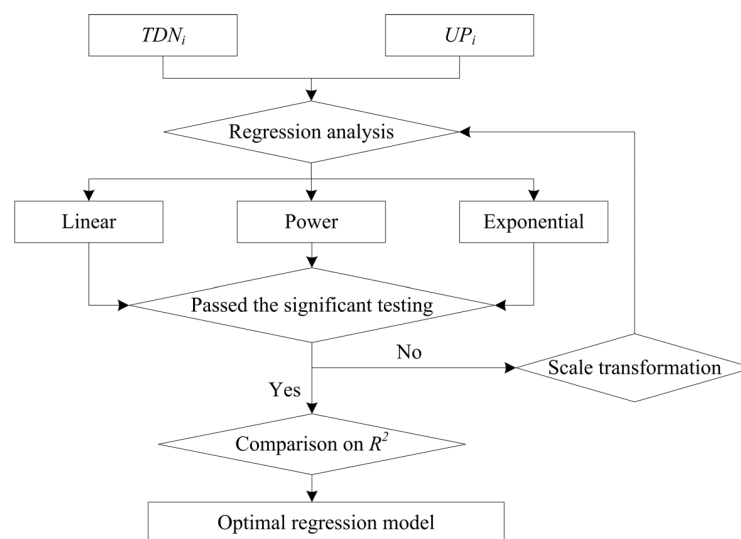

Figure 2. Flow chart for choosing an optimal regression model.

The optimal regression models at the two scales were determined by the coefficient of determination (R^2) and significance test of the regression equations. At the provincial scale, when none of the three regression models passed the significance test, the national optimal regression model was used as the optimal regression model for a given province/region. When different forms of optimal regression were obtained for a given province among different years, the regression form that occurred most frequently was used as the optimal regression for a given province over time.

4. Results

4.1. National Scale

Nighttime light data can be used to retrieve a city's urban population in China. All three regression models passed the significance test at the 0.001 level. However, the power function regression model was the optimal regression model between the TDN of nighttime light data and UP data for 1990, 2000, and 2010, with a goodness of fit (R^2) greater than 0.674 (Table 1). Specifically, the R^2 values of the linear regression model for the TDN and UP data for 1990, 2000, and 2010 were 0.562, 0.616, and 0.563, respectively. The R^2 values of the exponential regression model for 1990, 2000, and 2010 were 0.353, 0.430, and 0.508, whereas the R^2 values of the power function regression model for the three years were 0.696, 0.711, and 0.674, respectively. Clearly, the R^2 of the power function regression model was higher than that of the other two regression models for each of the aforementioned three years.

Table 1. The regression results of the DMSP/OLS nighttime light brightness data and statistical data for the UP at the national scale.

Equation/Year	Linear			Power			Exponential		
	1990	**2000**	**2010**	**1990**	**2000**	**2010**	**1990**	**2000**	**2010**
n	339	339	339	339	339	339	339	339	339
R^2	0.562	0.616	0.563	0.696	0.711	0.674	0.353	0.430	0.508
Sig.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

4.2. Provincial Scale

At the provincial scale, the optimal regression model varied across provinces (Table 2), while the linear regression model was the optimal model for most of provinces. Of the three regression models for 1990–2010, the linear regression model was the optimal regression model for 64% of the provinces. For example, the mean R^2 value of the linear regression model for Hubei Province is as high as 0.941. The power function regression model was the optimal regression model for 20% of the provinces. For instance, the mean R^2 value of the power function regression model for Qinghai Province was 0.810. The exponential regression model was the optimal regression model for 16% of the provinces. The mean R^2 value of the exponential regression model for Beijing-Tianjin-Hebei was 0.894. In addition, the R^2 value of the optimal regression model for each province/region shows an increasing and then decreasing trend from 1990 to 2010. The mean R^2 values for 1990, 2000, and 2010 were 0.721, 0.798, and 0.746, respectively. None of the fitting results of the three regression models for the Ningxia Hui Autonomous Region and Hainan Province passed the significance test at the 0.05 level, therefore, the regression model at the national scale was adopted in these two provinces.

Table 2. The regression results of the total digital number from the DMSP/OLS nighttime light data and statistical data for the UP.

Region	Optimal Regression Model	n	R^2		
			1990	2000	2010
Fujian	Linear	9	0.618 *	0.827 ***	0.700 **
Gansu	Linear	14	0.941 ***	0.950 ***	0.786 ***
Guangdong	Linear	21	0.516 ***	0.694 ***	0.514 ***
Guangxi	Linear	14	0.778 ***	0.684 ***	0.570 **
Guizhou	Linear	9	0.761 **	0.714 **	0.569 *
Henan	Linear	17	0.481 **	0.682 ***	0.850 ***
Hubei	Linear	14	0.927 ***	0.961 ***	0.935 ***
Hunan	Linear	14	0.506 **	0.745 ***	0.842 ***
Jilin	Linear	9	0.884 ***	0.949 ***	0.920 ***
Jiangxi	Linear	11	0.840 ***	0.895 ***	0.881 ***
Liaoning	Linear	14	0.778 ***	0.810 ***	0.827 ***
Neimenggu	Linear	12	0.638 **	0.734 ***	0.522 *
Shandong	Linear	17	0.386 *	0.690 ***	0.744 ***
Tibet	Linear	7	0.870 **	0.936 ***	0.726 *
Yunnan	Linear	16	0.920 ***	0.934 ***	0.822 ***
Chongqing & Sichuan	Linear	22	0.701 ***	0.891 ***	0.953 ***
Anhui	Power	16	0.768 ***	0.799 ***	0.781 ***
Heilongjiang	Power	13	0.728 ***	0.789 ***	0.833 ***
Qinghai	Power	8	0.789 **	0.878 ***	0.762 **
Shaanxi	Power	10	0.769 ***	0.627 *	0.517 *
Xinjiang	Power	15	0.453 *	0.589 ***	0.411 *
Shanxi	Exponential	11	0.737 ***	0.574 *	0.682 **
Zhejiang	Exponential	11	0.639 **	0.816 ***	0.812 ***
Beijing, Tianjin &Hebei	Exponential	13	0.865 ***	0.899 ***	0.919 ***
Shanghai & Jiangsu	Exponential	14	0.720 ***	0.880 ***	0.772 ***
Ningxia [1]	N/A	4	N/A	N/A	N/A
Hainan [1]	N/A	3	N/A	N/A	N/A

* represent the 0.05 statistical significance levels. ** represent the 0.005 statistical significance levels. *** represent the 0.001 statistical significance levels. [1] As none of the regression models for the Ningxia Hui Autonomous Region and Hainan Province pass the significance test at the 0.05 level, the optimal regression model for the national scale is used as the optimal regression model. The relative errors between census data and simulated results in Ningxia in 1990, 2000, and 2010 were 18.47%, 42.36%, and 7.11%, respectively. The relative errors in Hainan for the three years were −47.25%, −41.72%, and −4.19%, respectively.

5. Discussion

5.1. Simulation Results of the UP Based on Nighttime Light Data Are Better than those of the NAP

To comparatively verify the reliability of the simulation results for the UP, the NAP of mainland China for 1990–2010 was also simulated based on statistical data for the NAP using the same approach shown in Figure 2. In addition, the accuracy of our simulation of the UP and NAP for 1990–2010 based on DMSP/OLS stable nighttime light data was also evaluated and compared at the national and provincial scales, respectively (Figures 3 and 4).
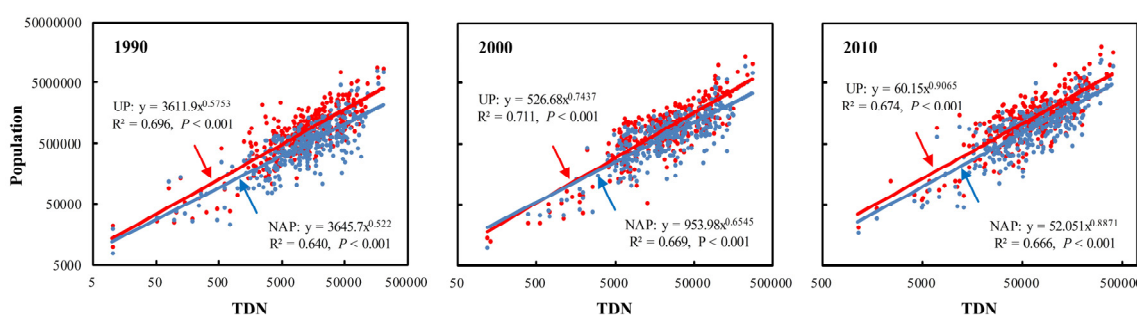
Figure 3. Optimal regression results of the total digital number (TDN) from the DMSP/OLS nighttime light data and UP/NAP at the national scale.
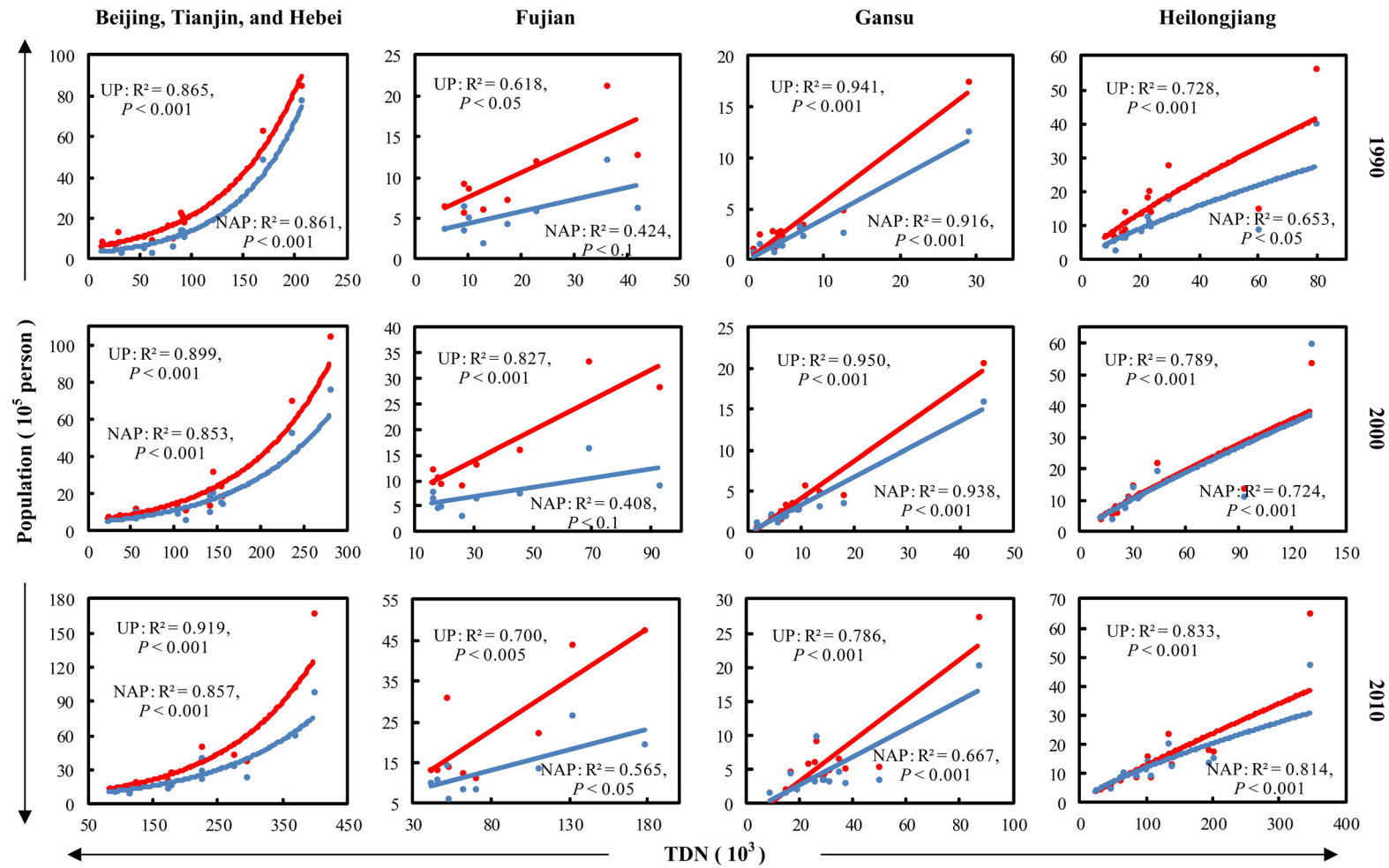
Figure 4. Optimal regression results of the TDN from the DMSP/OLS nighttime light data and UP/NAP at the province level.

At the national scale, the fitting results of the UP based on DMSP/OLS nighttime light data were better than those of the NAP (Figure 3). The R^2 values of the optimal regression model for the TDN and the UP for 1990, 2000 and 2010 were 0.696, 0.711 and 0.674, respectively. In comparison, the R^2 values of the optimal regression model for the TDN and the NAP for 1990, 2000, and 2010 are 0.640, 0.669, and 0.666, respectively. All of the optimal regression models constructed based on TDN and UP/NAP data for 1990–2010 passed the significance test at the 0.001 level.

At the provincial scale, the fitting results of the UP based on DMSP/OLS nighttime light data were also better than those of the NAP. The mean R^2 value of the optimal regression models for the TDN and the UP for all provinces from 1990 to 2010 was 0.755, whereas the mean R^2 value of the optimal regression models for the TDN and the NAP was 0.713. A comparison between the UP and NAP fitting results for four representative regions that are different in location, development level, and optimal fitting models showed that, the R^2 value of the UP was higher than that of the NAP (Figure 4). For the Beijing-Tianjin-Hebei region, the optimal fitting model was the exponential regression model. The mean R^2 values of the UP and NAP based on the TDN was 0.894 and 0.857, respectively. For Gansu Province, the linear regression model was the optimal fitting model. The mean R^2 values of the UP and NAP was 0.892 and 0.840, respectively. For Fujian Province, the linear regression model was the optimal fitting model, while the mean R^2 values of the UP and NAP were 0.715 and 0.466, respectively. For Heilongjiang Province, the power function regression model was the optimal fitting model, while the mean R^2 values of the UP and NAP was 0.783 and 0.730, respectively.

In summary, the accuracy of the evaluation results showed that the fitting results of the UP based on DMSP/OLS nighttime light data were better than those of the NAP. It confirmed there was an undeniable difference between NAP data and the actual UP of China. Substituting the NAP for the UP would undoubtedly produce data errors. The difference between the NAP and UP was attributable to major population flows between urban and rural areas that have occurred with recent processes of rapid urbanization in China. However, due to unique features of Chinese household registration system, the large population that moved from rural to urban areas has not truly obtained urban resident status [46,47]. Data for the sixth national census showed that the total migrant population of China reached 260,100,000 in 2010, an 81.0% increase from the total migrant population in 2000. Migrant workers originating from rural areas accounted for much of the aforementioned migrant population. Therefore, simulating long time series information on the UP of China based on DMSP/OLS nighttime light data can compensate for the deficiencies in statistical datasets on the UP in China.

5.2. Simulation Accuracy Levels Varied Across Cities of Different Sizes

The size of a city is a key factor that affects the accuracy of simulations. We divided all the cities into four groups (*i.e.*, mega-, large, medium, and small cities) according to their urban population in 2010 [48]. The relative errors for the UPs in large cities and megacities were higher than the results in medium and small cities (Table 3). For cities with a statistical UP of more than 10,000,000 in 2010 (e.g., Shanghai, Beijing, and Chongqing), the relative errors of the simulation results of the UP for 2000 had a mean value of approximately 16.76%, while the mean relative error for a city with a statistical UP of 3,000,000–10,000,000 (e.g., Chengdu, Harbin, and Hangzhou) was 14.08%. For cities with a statistical UP of less than 3,000,000 for 2010 (e.g., Luoyang, Lanzhou, and Xiaogan), the mean relative error of the simulation results for was the smallest (11.29%).

Table 3. Comparisons between the accuracy of the simulation results for the UP and NAP across cities (the cities are ordered by their sizes in a descending manner).

City	Statistical Data				Simulated Result				Relative Error (%)			
	UP		NAP		UP		NAP		UP		NAP	
	2000	2010	2000	2010	2000	2010	2000	2010	2000	2010	2000	2010
Shanghai	14,489,919	20,555,098	9,672,867	12,549,457	12,999,474	10,477,475	8,333,140	6,993,519	−10.29	−49.03	−13.85	−44.27
Beijing	10,522,464	16,858,692	7,628,441	9,931,140	9,023,602	12,554,614	6,267,903	7,594,441	−14.24	−25.53	−17.84	−23.53
Chongqing	10,095,512	15,295,803	6,350,940	11,069,952	8,065,198	14,306,374	4,921,309	10,254,142	−20.11	−6.47	−22.51	−7.37
Guangzhou	8,090,976	10,641,408	4,361,055	7,240,465	6,448,506	8,819,402	1,799,823	4,636,348	−20.30	−17.12	−58.73	−35.97
Tianjin	7,089,812	10,277,893	5,357,359	6,047,143	5,753,945	10,151,540	4,090,165	6,297,934	−18.84	−1.23	−23.65	4.15
Chengdu	5,967,819	9,237,015	3,458,970	6,509,118	7,459,475	9,814,294	4,561,005	7,042,435	24.99	6.25	31.86	8.19
Harbin	5,370,174	6,501,848	5,984,813	4,757,729	3,836,400	3,838,318	3,701,889	3,058,296	−28.56	−40.97	−38.15	−35.72
Hangzhou	4,033,397	6,372,650	2,269,946	3,544,784	4,348,305	6,489,276	1,864,825	2,700,392	7.81	1.83	−17.85	−23.82
Shenyang	5,066,072	6,247,700	4,333,209	4,692,388	4,867,256	6,146,719	3,992,731	4,563,064	−3.92	−1.62	−7.86	−2.76
Nanjing	4,355,280	6,238,186	3,095,203	5,421,980	3,139,684	4,256,668	2,267,173	3,271,116	−27.91	−31.76	−26.75	−39.67
Ningbo	3,323,736	5,195,162	1,420,327	2,020,381	3,863,717	6,068,642	1,673,728	2,535,452	16.25	16.81	17.84	25.49
Changsha	2,743,826	4,765,918	1,864,206	2,377,815	2,692,749	5,296,560	1,880,339	2,795,650	−1.86	11.13	0.87	17.57
Xuzhou	2,981,759	4,561,500	2,311,123	4,453,759	2,662,311	3,815,663	1,949,219	2,982,812	−10.71	−16.35	−15.66	−33.03
Fuzhou	3,359,051	4,408,076	1,651,957	2,663,508	2,536,278	3,570,545	1,039,726	1,827,731	−24.49	−19.00	−37.06	−31.38
Jinan	3,334,482	4,392,922	2,331,027	4,309,970	2,745,031	3,240,411	1,856,618	2,673,081	−17.68	−26.24	−20.35	−37.98
Kunming	3,176,380	4,337,798	1,891,491	2,251,570	2,929,754	3,557,160	1,755,077	1,704,434	−7.76	−18.00	−7.21	−24.30
Tangshan	2,265,605	3,850,975	1,907,110	2,397,964	2,222,904	5,993,640	1,659,212	3,958,881	−1.88	55.64	−13.00	65.09
Nanning	2,104,617	3,578,333	1,554,706	1,919,790	1,842,185	2,649,556	1,323,022	1,287,782	−12.47	−25.96	−14.90	−32.92
Taiyuan	2,755,726	3,467,987	2,039,238	2,630,159	2,119,664	2,149,781	1,418,734	1,486,624	−23.08	−38.01	−30.43	−43.48
Nanchang	2,115,437	3,313,235	1,758,950	2,340,239	2,155,403	2,924,979	1,865,327	1,932,236	1.89	−11.72	6.05	−17.43
Luoyang	1,870,406	2,888,355	1,468,523	1,912,413	1,699,607	2,927,984	1,250,806	1,828,237	−9.13	1.37	−14.83	−4.40
Lanzhou	2,070,949	2,758,558	1,597,481	2,029,221	1,968,707	2,306,056	1,496,757	1,657,460	−4.94	−16.40	−6.31	−18.32
Xiaogan	1,548,589	2,214,781	844,567	1,416,124	1,741,585	2,459,204	1,199,517	1,939,139	12.46	11.04	42.03	36.93
Shangqiu	1,023,886	2,171,557	1,024,097	1,731,242	1,021,530	2,354,475	877,274	1,513,271	−0.23	8.42	−14.34	−12.59
HulunBuir	1,750,292	1,722,795	1,630,192	1,796,552	1,587,301	1,940,984	1,316,459	1,452,134	−9.31	12.66	−19.25	−19.17
Huaihua	1,064,136	1,711,423	896,045	948,312	899,311	1,891,743	711,682	1,065,941	−15.49	10.54	−20.58	12.40
Yanbian	1,485,818	1,597,452	1,363,727	1,457,818	1,158,304	1,180,314	1,014,665	1,028,003	−22.04	−26.11	−25.60	−29.48
Wuhu	904,872	1,475,022	715,901	1,334,234	933,744	1,448,718	749,762	950,295	3.19	−1.78	4.73	−28.78
Tongren	531,433	803,990	333,122	486,045	483,633	513,925	250,665	213,179	−8.99	−36.08	−24.75	−56.14
Haidong	177,673	318,915	151,986	202,340	225,774	424,500	203,144	297,853	27.07	33.11	33.66	47.20

The relative errors of the simulation results of the NAP in large cities and megacities were also larger than those in medium and small cities. For large and megacities (UP > 10 million in 2010), the relative errors of the simulation results for the NAP for 2000 had a mean value of 27.32%. Meanwhile, the relative errors of the simulation results for the NAP in 2000 were 19.06% for medium cities (3 million < UP < 10 million in 2010) and 20.61% in small cities (UP < 3 million in 2010). Some plausible causes might account for the difference. First, the policies in China's large and megacities for urban population registration are stricter than the policies in medium and small cities. The flow of the urban population in medium and small cities is more flexible. Second, the frequent adjustment of administrative boundaries in large cities and megacities might lead to an incomparable urban population over time. During the last two decades, the administrative boundaries in a number of megacities and large cities have been adjusted dramatically, such as Chongqing, Beijing, Shanghai, Guangzhou, and Tianjin [49]. Therefore, the simulation accuracy in medium and small cities in China was higher than in large and megacities.

5.3. Limitations and Avenues for Future Research

The use of DMSP/OLS nighttime light data for UP simulations presented some shortcomings. First, the low spatial resolution of DMSP/OLS nighttime light data and issues related to saturated pixels in urban cores could limit the accuracy of UP simulations. In the future, corrected nighttime light data may be used to overcome the saturation issues [50]. Second, consumption and living habits vary across different areas in China, resulting in a difference in nighttime light data-based socioeconomic models for different regions. Therefore, integrating remote sensing products with higher spatial, temporal and spectral resolutions from multiple sources may improve the accuracy of simulation. For example, using newly published nighttime light data from the National Polar-Orbiting Partnership-Visible Infrared Imaging Radiometer Suite can improve the spatial and spectral resolutions of data while overcome data saturation issues in city centers, and consequently led to a better simulation of the UP [51].

6. Conclusions

DMSP/OLS nighttime light data can be used to retrieve long time series data on the UP of China. Due to regional variations in China, the retrieval model varied at different scales. At the national scale, all three regression models passed the significance test at the 0.001 level. However, the power function model produced the best results, with a R^2 greater than 0.67. At the provincial level, the linear regression model was the best regression model for 64% of the provinces. The power function model and the exponential model were the best regression model for 20% and 16% of the provinces, respectively.

The fitting results of the UP based on DMSP/OLS nighttime light data were better than those of the NAP. At the national scale, the mean R^2 values of the optimal regression models for the UP and NAP data from 1990 to 2010 were 0.694 and 0.658, respectively. At the provincial scale, the mean R^2 values of the optimal regression models between the UP and NAP data from 1990 to 2010 were 0.755 and 0.713, respectively. In addition, a city's size also affected the simulation results. Relative errors between the simulation results and statistical data on the UP for cities with an UP of fewer than 3,000,000 were smaller than those cities with an UP of more than 3,000,000. Therefore, we believe that fitting the UP based on nighttime light data is particularly suitable for China, a country with a large migratory population, which is undergoing rapid processes of urbanization.

Author Contributions: Qingxu Huang and Yang Yang conceived and designed the study. Yajing Li processed the data and performed analysis. Yang Yang and Qingxu Huang contributed to the interpretation of the results. Yang Yang and Yajing Li wrote the paper. All authors reviewed and edited the draft, approved the submitted manuscript, and agreed to be listed and accepted the version for publication. All authors have read and approved of the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1.　Montgomery, M.R. The urban transformation of the developing world. *Science* **2008**, *319*, 761–764. [CrossRef] [PubMed]
2.　Han, B.L.; Wang, R.S.; Tao, Y.; Gao, H. Urban population agglomeration in view of complex ecological niche: A case study on Chinese prefecture cities. *Ecol. Indic.* **2014**, *47*, 128–136. [CrossRef]
3.　Zhang, K.H.L.; Song, S.F. Rural-urban migration and urbanization in China: Evidence from time-series and cross-section analyses. *China Econ. Rev.* **2003**, *14*, 386–400. [CrossRef]
4.　Bai, X.M.; Shi, P.J.; Liu, Y.S. Realizing China's urban dream. *Nature* **2014**, *509*, 158–160. [CrossRef] [PubMed]
5.　Zhou, Y.X.; Ma, L.J.C. China's urbanization levels: Reconstructing a baseline from the fifth population census. *The China Q.* **2003**, *173*, 176–196.
6.　Zhou, Y.; Ma, L.J.C. China's Urban population statistics: A critical evaluation. *Eurasian Geogr. Econ.* **2005**, *46*, 272–289. [CrossRef]
7.　Qin, B.; Zhang, Y. Note on urbanization in China: Urban definitions and census data. *China Econ. Rev.* **2014**, *30*, 495–502. [CrossRef]
8.　He, C.F.; Chen, T.M.; Mao, X.Y.; Zhou, Y. Economic transition, urbanization and population redistribution in China. *Habitat Int.* **2016**, *51*, 39–47. [CrossRef]
9.　Ma, L.J.C.; Cui, G.H. Administrative changes and urban population in China. *Ann. Assoc. Am. Geogr.* **1987**, *77*, 373–395.
10.　Shen, J. Counting urban population in Chinese censuses 1953–2000: Changing definitions, problems and solutions. *Popul. Sp. Place* **2005**, *11*, 381–400. [CrossRef]
11.　Tan, M.H.; Li, X.B.; Lv, C.H.; Luo, W.; Kong, X.B.; Ma, S.H. Urban population densities and their policy implications in China. *Habitat Int.* **2008**, *32*, 471–484. [CrossRef]
12.　Tao, L.; Hui, E.C.M.; Wong, F.K.W.; Chen, T.T. Housing choices of migrant workers in China: Beyond the Hukou perspective. *Habitat Int.* **2015**, *49*, 474–483. [CrossRef]
13.　Sutton, P.; Roberts, D.; Elvidge, C.; Meij, H. A comparison of nighttime satellite imagery and population density for the continental united states. *Photogramm. Eng. Remote Sens.* **1997**, *63*, 1303–1313.
14.　Dobson, J.E.; Brlght, E.A.; Coleman, P.R.; Durfee, R.C.; Worley, B.A. Landscan: A global population database for estimating populations at risk. *Photogramm. Eng. Remote Sens.* **2000**, *66*, 849–857.
15.　Sutton, P.C.; Elvidge, C.D.; Obremski, T. Building and evaluating models to estimate ambient population density. *Photogramm. Eng. Remote Sens.* **2003**, *69*, 545–553. [CrossRef]
16.　Amaral, S.; Câmara, G.; Monteiro, A.M.V.; Quintanilha, J.A.; Elvidge, C.D. Estimating population and energy consumption in Brazilian Amazonia using DMSP night-time satellite data. *Comput. Environ. Urban Syst.* **2005**, *29*, 179–195. [CrossRef]
17.　Zhuo, L.; Ichinose, T.; Zheng, J.; Chen, J.; Shi, P.J.; Li, X. Modeling population density of China at pixel level based on DMSP/OLS non-radiance calibrated nighttime light image. *Int. J. Remote Sens.* **2009**, *30*, 1003–1018. [CrossRef]
18.　Sutton, P.C.; Anderson, S.J.; Elvidge, C.D.; Tuttle, B.T.; Ghosh, T. Paving the planet: Impervious surface as proxy measure of the human ecological footprint. *Prog. Phys. Geog.* **2009**, *33*, 510–527. [CrossRef]
19.　Imhoff, M.L.; Lawrence, W.T.; Stutzer, D.C.; Elvidge, C.D. A technique for using composite DMSP/OLS "city lights" satellite data to map urban area. *Remote Sens. Environ.* **1997**, *61*, 361–370. [CrossRef]
20.　Elvidge, C.; Baugh, K.; Hobson, V.; Kihn, E.; Kroehl, H.; Davis, E.; Cocero, D. Satellite inventory of human settlements using nocturnal radiation emissions: A contribution for the global tool chest. *Glob. Chang. Biol.* **1997**, *3*, 387–395. [CrossRef]
21.　Henderson, M.; Yeh, E.T.; Gong, P.; Elvidge, C.; Baugh, K. Validation of urban boundaries derived from global night-time satellite imagery. *Int. J. Remote Sens.* **2003**, *24*, 595–609. [CrossRef]
22.　He, C.Y.; Shi, P.J.; Li, J.G.; Chen, J.; Pan, Y.Z.; Li, J.; Zhuo, L.; Ichinose, T. Restoring urbanization process in China in the 1990s by using non-radiance-calibrated DMSP/OLS nighttime light imagery and statistical data. *Chin. Sci. Bull.* **2006**, *51*, 1614–1620. [CrossRef]
23.　Lu, D.; Tian, H.; Zhou, G.; Ge, H. Regional mapping of human settlements in southeastern China with multisensor remotely sensed data. *Remote Sens. Environ.* **2008**, *112*, 3668–3679. [CrossRef]

24.   Cao, X.; Chen, J.; Imura, H.; Higashi, O. A SVM-based method to extract urban areas from DMSP-OLS and spot vgt data. *Remote Sens. Environ.* **2009**, *113*, 2205–2209. [CrossRef]

25.   Liu, Z.F.; He, C.Y.; Zhang, Q.F.; Huang, Q.X.; Yang, Y. Extracting the dynamics of urban expansion in China using DMSP-OLS nighttime light data from 1992 to 2008. *Landsc. Urban Plan.* **2012**, *106*, 62–72. [CrossRef]

26.   Yang, Y.; He, C.; Zhang, Q.; Han, L.; Du, S. Timely and accurate national-scale mapping of urban land in China using defense meteorological satellite program's operational linescan system nighttime stable light data. *J. Appl. Remote Sens.* **2013**, *7*, 073535. [CrossRef]

27.   Zhang, Q.L.; Seto, K.C. Mapping urbanization dynamics at regional and global scales using multi-temporal DMSP/OLS nighttime light data. *Remote Sens. Environ.* **2011**, *115*, 2320–2329. [CrossRef]

28.   Gao, B.; Huang, Q.X.; He, C.Y.; Ma, Q. Dynamics of urbanization levels in China from 1992 to 2012: Perspective from DMSP/OLS nighttime light data. *Remote Sens.* **2015**, *7*, 1721–1735. [CrossRef]

29.   Elvidge, C.D.; Ziskin, D.; Baugh, K.E.; Tuttle, B.T.; Ghosh, T.; Pack, D.W.; Erwin, E.H.; Zhizhin, M. A fifteen year record of global natural gas flaring derived from satellite data. *Energies* **2009**, *2*, 595–622. [CrossRef]

30.   Chand, K.T.R.; Badarinath, K.V.S.; Elvidge, C.D.; Tuttle, B.T. Spatial characterization of electrical power consumption patterns over India using temporal DMSP-OLS night-time satellite data. *Int. J. Remote Sens.* **2009**, *30*, 647–661. [CrossRef]

31.   Townsend, A.C.; Bruce, D.A. The use of night-time lights satellite imagery as a measure of Australia's regional electricity consumption and population distribution. *Int. J. Remote Sens.* **2010**, *31*, 4459–4480. [CrossRef]

32.   Coscieme, L.; Pulselli, F.M.; Bastianoni, S.; Elvidge, C.D.; Anderson, S.; Sutton, P.C. A thermodynamic geography: Night-time satellite imagery as a proxy measure of emergy. *Ambio* **2013**, *43*, 969–979. [CrossRef] [PubMed]

33.   Doll, C.N.H.; Muller, J.P.; Morley, J.G. Mapping regional economic activity from night-timelight satellite imagery. *Ecol. Econ.* **2006**, *57*, 75–92. [CrossRef]

34.   Lo, C.P. Urban indicators of China from radiance-calibrated digital DMSP-OLS nighttime images. *Ann. Assoc. Am. Geogr.* **2002**, *92*, 225–240. [CrossRef]

35.   Elvidge, C.D.; Sutton, P.C.; Ghosh, T.; Tuttle, B.T.; Baugh, K.E.; Bhaduri, B.; Bright, E. A global poverty map derived from satellite data. *Comput. Geosci.* **2009**, *35*, 1652–1660. [CrossRef]

36.   Milesi, C.; Elvidge, C.D.; Nemani, R.R.; Running, S.W. Assessing the impact of urban land development on net primary productivity in the southeastern United States. *Remote Sens. Environ.* **2003**, *86*, 401–410. [CrossRef]

37.   Imhoff, M.L.; Bounoua, L.; DeFries, R.; Lawrence, W.T.; Stutzer, D.; Tucker, C.J.; Ricketts, T. The consequences of urban land transformation on net primary productivity in the United States. *Remote Sens. Environ.* **2004**, *89*, 434–443. [CrossRef]

38.   Ghosh, T.; Anderson, S.; Elvidge, C.D.; Sutton, P.C. Using nighttime satellite imagery as a proxy measure of human well-being. *Sustainability* **2013**, *5*, 4988–5019. [CrossRef]

39.   Small, C.; Elvidge, C.D. Night on earth: Mapping decadal changes of anthropogenic night light in Asia. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *22*, 40–52. [CrossRef]

40.   Xu, T.; Ma, T.; Zhou, C.H.; Zhou, Y.K. Characterizing spatio-temporal dynamics of urbanization in China using time series of DMSP/OLS night light data. *Remote Sens.* **2014**, *6*, 7708–7731. [CrossRef]

41.   Xiao, P.F.; Wang, X.H.; Feng, X.Z.; Yang, Y.K. Detecting China's urban expansion over the past three decades using nighttime light data. *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.* **2014**, *7*, 4095–4106. [CrossRef]

42.   Huang, Q.X.; Yang, X.; Gao, B.; Yang, Y.; Zhao, Y.Y. Application of DMSP/OLS nighttime light images: A meta-analysis and a systematic literature review. *Remote Sens.* **2014**, *6*, 6844–6866. [CrossRef]

43.   *China City Statistical Yearbook 1992–2011*; National Bureau of Statistics of China, China Statistics Press: Beijing, China, 1993–2012.

44.   National Geomatics Center of China. Available online: http://ngcc.sbsm.gov.cn/article/khly/lyzx/ (accessed on 1 June 2015).

45.   Ma, T.; Zhou, C.; Pei, T.; Haynie, S.; Fan, J. 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.* **2012**, *124*, 99–107. [CrossRef]

46.   Chan, K.W.; Zhang, L. The Hukou system and rural-urban migration in China: Processes and changes. *China Q.* **1999**, *160*, 818–855. [CrossRef] [PubMed]

47. Chan, K.W. The Chinese Hukou system at 50. *Eurasian Geogr. Econ.* **2009**, *50*, 197–221. [CrossRef]
48. Gao, B.; Huang, Q.; He, C.; Sun, Z.; Zhang, D. How does sprawl differ across cities in China? A multi-scale investigation using nighttime light and census data. *Landsc. Urban Plan.* **2016**, *148*, 89–98. [CrossRef]
49. Zhang, L.; Lei, J.; Li, X.; Gao, C.; Zeng, W. The features and influencing factors of urban expansion in China during 1997–2007. *Prog. Geog.* **2011**, *30*, 607–614. (In Chinese)
50. Letu, H.; Hara, M.; Tana, G.; Nishio, F. A Saturated Light Correction Method for DMSP/OLS Nighttime Satellite Imagery. *IEEE Trans. Geosci. Remot. Sen.* **2012**, *50*, 389–396. [CrossRef]
51. Shi, K.; Yu, B.; Huang, Y.; Hu, Y.; Yin, B.; Chen, Z.; Chen, L.; Wu, J. Evaluating the Ability of NPP-VIIRS Nighttime Light Data to Estimate the Gross Domestic Product and the Electric Power Consumption of China at Multiple Scales: A Comparison with DMSP-OLS Data. *Remote Sens.* **2014**, *6*, 1705–1724. [CrossRef]