Numerical Simulation of Population Distribution in China

T. X. Yue
Y. A. Wang
S. P. Chen
J. Y. Liu
D. S. Qiu
X. Z. Deng
M. L. Liu
Y. Z. Tian
Chinese Academy of Sciences

A model for simulating population distribution (MSPD) of China is developed based on the grid generation method and the Control of MapObjects of geographical information system. Elevation, net primary productivity, land use and land cover, city sizes and their spatial distribution, and spatial distribution of transport infrastructures are taken into full account in the MSPD. The result from the MSPD shows that in 2000, 90.8% of the total population of China distributed on the southeastern side of the Heihe-Tengchong line. The ratio of population on the northwestern side to total population of China has been increasing since 1935. The yearly growth rate was 0.8% from 1935 to 1990 and 6.1% from 1990 to 2000. One important advantage of the MSPD is that when scenarios of land cover, spatial distributions of transport infrastructures and cities are available, scenarios of spatial population distribution can be developed on the basis of total population forecast.

KEY WORDS: population distribution; numerical simulation; grid generation; geographical information system.

Please address correspondence to T. X. Yue, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, 917 Building, Datun, Anwai, 100101 Beijing, P. R. China; e-mail: yue@www.lreis.ac.cn.

Population and Environment, Vol. 25, No. 2, November 2003 © 2003 Human Sciences Press, Inc.

INTRODUCTION

The analysis of human population distribution represents a particularly influential line of enquiry (Woods & Rees, 1986). It traces back to the statistical revolution that began in Europe and America in the 19th century and resulted from the combination of regular population censuses with vital registration data organized and manipulated within a framework of administrative units (Cullen, 1975; Cassedy, 1984). Approaches to estimating population distribution within administrative units can be divided into two categories that are areal interpolation and surface Modelling (Deichmann, 1996).

Areal interpolation is the transformation of data between different sets of areal units (Goodchild & Lam, 1980). The set of zones, for which data are available, is termed source zones. The second set of zones, for which estimates need to be derived, is termed target zones. The third set of zones, for which auxiliary information can be incorporated in the interpolation process, is termed control zones (Goolchild et al., 1993; Moxey & Allanson, 1994). The methods of areal interpolation based on alternative hypotheses include radially symmetric kernel functions (Parr, 1985; Bracken & Martin, 1989), maximally smooth estimation (Tobler, 1979), piecewise approximation (Flowerdew et al., 1991), uniform target-zone densities (Goodchild & Lam, 1980), and uniform control-zone densities (Flowerdew & Green, 1989).

Surface Modelling is aimed at formulating population in a regular grid system, in which each grid contains an estimate of total population that is representative for that particular location. Representing data in grid form has at least three advantages: (1) regular grid can be easily re-aggregated to any areal arrangement required; (2) producing population data in grid form is one way of ensuring compatibility between heterogeneous data sets; and (3) converting data into grid form can provide a way of avoiding some of the problems imposed by artificial political boundaries (Martin & Bracken, 1991; Deichmann, 1996). Compiling population data in grid form is by no means a new approach. For instance, Adams (1968) presented a computer generated grid map of population density in West Africa; Population Atlas of China presented grid population data for several regions in China (Institute of Geography of Chinese Academy of Sciences, 1987).

The surface Modelling consists of three basic steps: (1) a surface of weighting factors is created in a regular grid system for the study areas; (2) the basic weights derived in the first step are adjusted by using auxiliary data sources, and (3) total population in the study areas is distributed to the corresponding grids in proportion to the weights constructed in the previous steps. In this paper, a model for simulating population distribution is developed on the basis of an improvement on surface modelling by introducing grid generation method.

METHODS

The Model for Simulating Population Distribution (MSPD)

Sir Isaac Newton propounded his Law of Universal Gravitation in 1687, i.e., any two bodies attract each other in proportion to the product of their masses and in inverse proportion to the square of their distance. In analogy to physical gravity model, the concept of potential population distribution was developed, which is a measure of average accessibility of a given location with respect to the size and location of other features (Plane & Rogerson, 1994; Deichmann, 1996). The influence of a city upon a grid is assumed to be proportional to the city's size, weighted inversely by the distance of separation between the grid and the city. Within a given threshold distance, the potential population distribution is formulated as (1)

$$P_{ij} = \sum_{k=1}^{M} \frac{S_k}{(d_{ijk})^a} \tag{1}$$

where p_{ij} is the population at grid (i, j); S_k is the size of city k; d_{ijk} is the distance between grid (i, j) and city k; M is the total number of cities within the given threshold distance; and a is exponent to be simulated.

In fact, spatial distribution of population is greatly influenced by natural factors, especially net primary productivity (NPP) and elevation. The influence of each of these factors is not simple. It depends on economic factors such as spatial distribution of transport infrastructures and cities. For instance, when a railway traverses Qinghai-Xizang plateau of China, or when a new city appears, the distribution of population is immediately modified. The growth of means of transport is fundamental in distribution of population. In China, an increasing contrast becomes manifest between regions already reached by modern and those not so affected. The former develops rapidly, in which great cities expanded at an accelerating pace. Transport favors the growth of cities. Cities are one of the major elements in the distribution of population. In general it is possible to distinguish a central and more densely peopled nucleus and a peripheral residential area and extended star-like along the lines of transportation. The influence of cities on the distribution of population is exercised directly through the concentration of people and also indirectly through the swarms of suburbanites scattered over their peripheral countryside in dormitory suburbs or satellite towns.

Therefore, in addition to the size of city k and the distance between grid (i, j) and city k embodied in the formulation (1), land cover, net primary productivity, transport infrastructures and elevation are involved in the MSPD. The MSPD introduced in this paper is generally formulated as,

$$MSPD_{ij}\ (t) = G(n,t) \cdot \mathrm{W}_{\mathrm{ij}}(\mathrm{t}) \cdot f_1\ (Tran_{ij}(t)) \cdot f_2\ (NPP_{ij}(t)) \cdot f_3\ (DEM_{ij})) \cdot f_4(u_{ij}(t)) \tag{2}$$

where t is the time variable; $G(n,t)$ is a parameter determined by total population in administrative division n where grid (i, j) is located; $\mathrm{W}_{\mathrm{ij}}(\mathrm{t})$ is an indicative factor of water area; $f_1(Tran_{ij})$ is a function determined by the condition of transport infrastructures of grid (i, j); $f_2(NPP_{ij})$ is a function determined by the condition of net primary productivity of grid (i, j); $f_3(DEM_{ij})$ is a function determined by the elevation of grid (i, j); $f_4(u_{ij}(t))$ is a function determined by contribution of urban areas to population density at grid (i, j); $u_{ij}(t) = \sum_{\mathrm{k}=1}^{\mathrm{M(t)}} \frac{(S_{\mathrm{k}}\ (t))^{\mathrm{a}_1}}{(d_{\mathrm{ijk}}(t))^{\mathrm{a}_2}}$, $S_k(t)$ is size of the kth city, $M(t)$ is the total number of cities, d_{ijk} is the distance from grid (i, j) to the core grid of the kth city, a_1 and a_2 are exponents to be simulated.

Grid Generation

In 1960s, grid generation techniques began to be developed (Morrison, 1962; Sidorov, 1966; Ahuja & Coons, 1968). The successful development of numerical grid generation has already formed a separate mathematical discipline. In the 1990s, grid techniques reached a new stage. The unique aspect of grid generation on general domain is that grid generation is not obliged to have any specified formulation and any foundation may be suitable for the purpose if the grid generated is acceptable. The most important step is to find an appropriate transformation between computational domain and physical domain for purposes (Liseikin, 1999).

A number of grid generation methods have been developed and every one of them has its strengths and weaknesses. This paper chooses the most efficient method, which is based on the gravity model, for the specific issue of population distribution of China, taking net primary productivity, elevation, city sizes and their spatial distribution, spatial distribution of transport infrastructures and grid location into full account.

MATERIALS AND DATA ACQUISITION

Spatial Distribution of Cities in China

Urbanization is a process of the concentration of population in cities. From 1978 to 2000, the mean growth rate of urban population in China is 7.5% annually, which occupied front place in the world. However, proportion of urban population to the total population in China is still 15% lower than the average one of the world. Urbanization in China will speed up in near future (Li, 2000). Spatial distribution of cities and proximity to cities are essential factors for population distribution of China.

The spatial distribution of cities in China has had the feature that city density is much higher in eastern China than in western China in modern history. Cities in China spatially concentrates in coastal area, especially Yangtze River Delta, Peal River Delta, and Beijing-Tianjin-Tangshan area. According to statistical data (Urban Society and Economy Survey Team of National Bureau of Statistics of People's Republic of China, 2001), 42.1% of the 667 major cities of China distributes in eastern China where area accounts for 9.5% of the whole area of China; 34% distributes in middle China where area accounts for 17.4%; 23.8% distributes in western China where area accounts for 70.4% (as seen in Figure 1 and Table 1). The distribution densities of cities in eastern China and in middle China are respectively 13.1 times and 5.8 times the one in western China. The biggest urban agglomerations are distributed in eastern China, such as the urban agglomeration surrounding Shanghai-Najing-Hangzhou, the one surrounding Beijing-Tianjin-Tangshan, and the one surrounding Guangzhou-Hong Kong-Macao.

Spatial Distribution of Transport Infrastructure in China

Transport infrastructure is a primary indicator of population distribution (Dobson et al., 2000). Roads and railways are especially indicative because of their vital role in human well-being. The transport trunk network (i.e., railways and roads) of China is composed of 6 east-west channels and 7 south-north channels (Chen & Zhang, 2000). The 6 east-west channels are Suifenhe-Harbin-Manzhouli channel, Dandong-Shengyang-Beijing-Baotou-Lanzhou channel, Qingdao-Jinan-Shijiazhuang- Taiyuan-Xian channel, Lianyungang-Zhengzhou-Lanzhou-Ürümqi-Alatam pass channel, Shanghai-Wuhan-Chongqing-Chengdu channel, and Shanghai-Hangzhou- Nanchang-Zhuzhou-Guiyang-Kunming channel. The 7 south-north channels are Harbin-Shenyang-Dalian–Shanghai-Guangzhou channel, Tianjin-Jinan-

TABLE 1

Spatial Distribution of the Major Cities not Including Taiwan, Hong Kong and Macao Temporarily in 2000

Region	Area (10^3 km^2)	City Number	Classification According to Non-Agricultural Population				
			More 2 Million	1 to 2 Million	0.5 to 1 Million	0.2 to 0.5 Million	Less 0.2 Million
Western China	*6754.6*	*159*	*3*	*5*	*4*	*46*	*101*
Inner Mongolia	1143.3	20	0	1	1	7	11
Guangxi	236.4	19	0	0	2	4	13
Chongqing	82.4	5	1	0	0	3	1
Sichuan	485	32	1	0	1	12	18
Guizhou	176.2	13	0	1	0	3	9
Yunnan	383.6	15	0	1	0	2	12
Tibet	1204.2	2	0	0	0	0	2
Shaanxi	205.8	13	1	0	0	5	7
Gansu	405.6	14	0	1	0	2	11
Qinghai	717.2	3	0	0	1	0	2
Ningxia	51.8	5	0	0	0	2	3
Xinjiang	1663.1	19	0	1	0	7	11
Middle China	*1670.8*	*227*	*3*	*8*	*25*	*71*	*120*
Shanxi	156.5	22	0	1	1	4	16
Anhui	140.2	22	0	1	4	10	7
Jiangxi	167	21	0	1	0	8	12

Henan	165.7	38	0	2	7	8	21
Hubei	186	36	1	0	4	12	19
Hunan	211.8	29	0	1	3	8	17
Jilin	189.2	28	1	1	0	11	15
Helongjiang	454.4	31	1	1	6	10	13
Eastern China	*912.71*	*276*	*7*	*14*	*23*	*100*	*132*
Beijing	16.4	1	1	0	0	0	0
Tianjin	11.8	1	1	0	0	0	0
Hebei	188.3	34	0	3	3	5	23
Liaoning	146.8	31	2	2	6	7	14
Shanghai	7.6	1	1	0	0	0	0
Jiangsu	105	41	1	3	3	23	11
Zhejiang	105	35	0	1	2	7	25
Fujian	122.3	23	0	1	1	4	17
Shandong	157.8	48	0	3	6	23	16
Guangdong	17.79	52	1	1	2	29	19
Hainan	33.92	9	0	0	0	2	7

Source: National Bureau of Statistics, 2001.

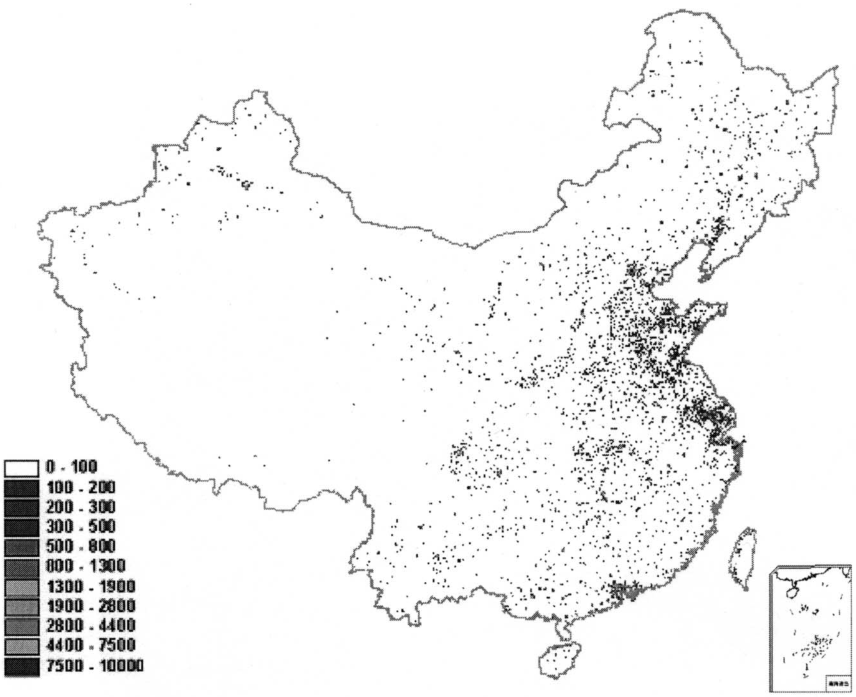

FIGURE 1. Spatial distribution of cities in 2000 in China (unit: thousand persons grouped by non-agricultural population in urban district).

Source: Resources and Environment Data Center, Institute of Geographical Science and Natural Resources Research, Chinese Academy of Sciences.

Xuzhou- Nanjing-Shanghai channel, Beijing-Wuhan-Guangzhou channel, Beijing-Jiujiang- Shenzhen-Jiulong channel, Datong-Taiyuan-Jiaozuo-Zhijiang-Liuzhou-Zhanjiang channel, Ankang-Chongqing-Guiyang-Liuzhou-Nanning-Fangcheng channel, and Zhongwei-Baoji-Chengdu-Kunming Channel. The transport infrastrucutre densities of eastern China and middle China are respectively 6 times and 3 times the one of western China. The provinces where the transport network densities are less than 100km per thousand square kilometer include Tibet, Xinjiang, Qinghai, Inner Mongolia and Gansu. They are all located in the western China. The provinces where the transport network densities are greater than 400 km per thousand square kilometer include Guangdong, Beijing, Tianjin, Shanghai, Hainan, Shandong and Fujian (National Bureau of Statistics of People's Republic of China, 2001). They are all distributed in the eastern China. The average transport situation in China is that from northwest to southeast the transport network density becomes greater and greater (as seen in Table 2, Figure 2 and Figure 3).

Land Cover and Spatial Distribution of NPP in China

Land cover is a good indicator of spatial population distribution. In most regions, population would range from extremely low density in desert, water, wetlands, ice, or tundra land cover to high density in developed land cover associated urban land cover, between which arid grasslands, forests, and cultivated lands would range (Dobson et al., 2000). The land cover database of China in 2000 is derived from Landsat Thematic Mapper (TM) imagery at 30-m resolution (Figure 4)

Net primary productivity (NPP) is the difference between accumulative photosynthesis and accumulative autotrophic respiration by green plants per unit time and space (Lieth and Whittaker, 1975). The Boreal Ecosystems Productivity Simulator (Liu et al., 1997) is employed for analyzing spatial distribution of NPP in China. It integrates different data types that include NOAA/AVHRR data at 1-km resolution in lambert conformal conic projection, daily meteorological data in Gaussian grided systems, and soil data grouped in polygons.

The general situation in China is that from southeast to northwest NPP becomes smaller and smaller gradually. Most of the NPP is distributed in the East of the rainfall line where the annual precipitation is 410 mm (as seen in Figure 5 and Figure 6), excepting that there is higher NPP in the southern slopes of Tianshan mountains and Altai mountains in Xinjiang. The maximum NPP appears in Xiaoxinganling mountain and Changbai mountain in the northeast China, Yunnan-Guizhou plateau, Guangxi, Hainan, Chongqin and provinces along middle and lower reaches of Yangtze river.

In terms of land cover types, on the average (as seen in Table 3), NPP of shrub and open forest is 1071 $gC \cdot m^{-2} \cdot yr^{-1}$, evergreen broad-leaved forest 975 $gC \cdot m^{-2} \cdot yr^{-1}$, deciduous broad-leaved forest 928 $gC \cdot m^{-2} \cdot yr^{-1}$, coniferous and broad-leaved mixed forest 870 $gC \cdot m^{-2} \cdot yr^{-1}$, farmland system 752 $gC \cdot m^{-2} \cdot yr^{-1}$, evergreen coniferous forest 587 $gC \cdot m^{-2} \cdot yr^{-1}$, deciduous coniferous forest 585 $gC \cdot m^{-2} \cdot yr^{-1}$, and grassland 271 $gC \cdot m^{-2} \cdot yr^{-1}$ (Liu, 2001).

TABLE 2

Spatial Distribution of Rail and Road not Including Taiwan, Hong Kong and Macao Temporarily in 2000

Region	Railway + Road (km/10^3 km^2)	Railway (km/10^3 km^2)	Road (km/10^3 km^2)
Western China	*86*	*4*	*82*
Inner Mongolia	64	5	59
Guangxi	234	10	224
Chongqing	362	7	355
Sichuan	193	5	187
Guizhou	206	10	197
Yunnan	291	5	286
Tibet	19	0	19
Shaanxi	227	13	214
Gansu	105	8	97
Qinghai	28	2	26
Ningxia	212	16	196
Xinjiang	23	2	21
Middle China	*262*	*19*	*243*
Shanxi	379	25	354
Anhui	338	21	317
Jiangxi	243	21	222
Henan	417	28	389
Hubei	328	17	311
Hunan	304	17	287
Jilin	206	20	186
Helongjiang	125	15	111
Eastern China	*511*	*26*	*486*
Beijing	940	111	829
Tianjin	840	82	758
Hebei	346	32	314
Liaoning	345	34	310
Shanghai	621	52	569
Jiangsu	282	14	269
Zhejiang	408	12	396
Fujian	425	7	418
Shandong	473	25	448
Guangdong	5845	78	5768
Hainan	520	7	513

Source: National Bureau of Statistics, 2001.

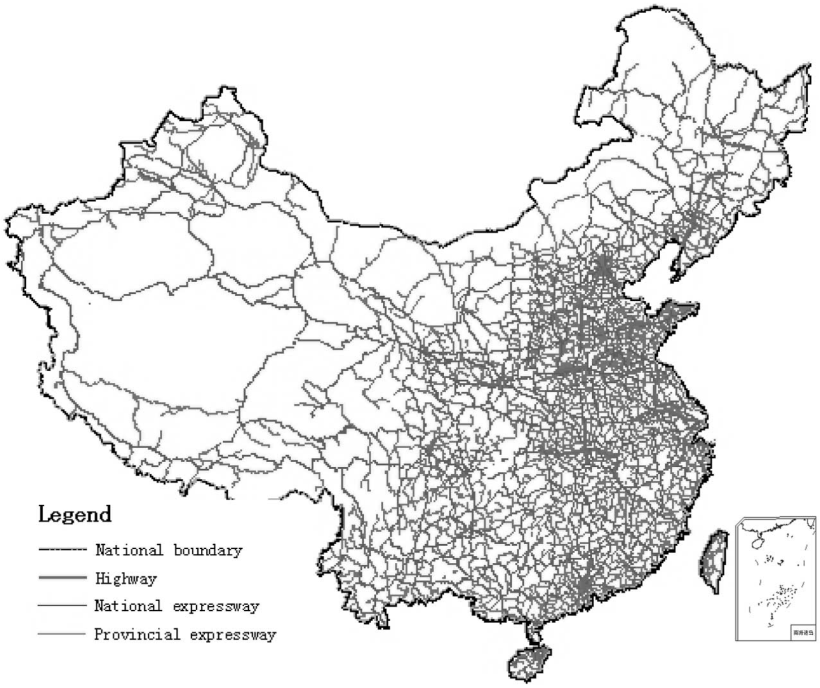

FIGURE 2. Spatial distribution of roads in 2000 in China.

Source: Resources and Environment Data Center, Institute of Geographical Science and Natural Resources Research, Chinese Academy of Sciences.

Elevation

Elevation is an important variable in population distribution modelling of China because most human settlements occur on lower elevation in China. The terrestrial parts of China are broadly divided into three steps as seen in Figure 7 from Qinghai-Xizang Plateau eastward (Zhao, 1986). The lofty and extensive Qinghai-Xizang Plateau is the first great topographic step. Its eastern and northern borders roughly coincide with the 3000m contour line. It generally has an elevation of 4000m to 5000m and hence is called the roof of the world.

From the eastern margin of the Qinghai-Xizang Plateau eastward up to the Da Hinggan-Taihang-Wushan mountains lies the second great topographic step. It is mainly composed of plateaus and basins with elevations of 1000 to 2000m, such as the Nei Mongol, Ordos, Loess and Yunnan-Guizhou plateaus and the Tarim, Junggar, and Sichuan basins.

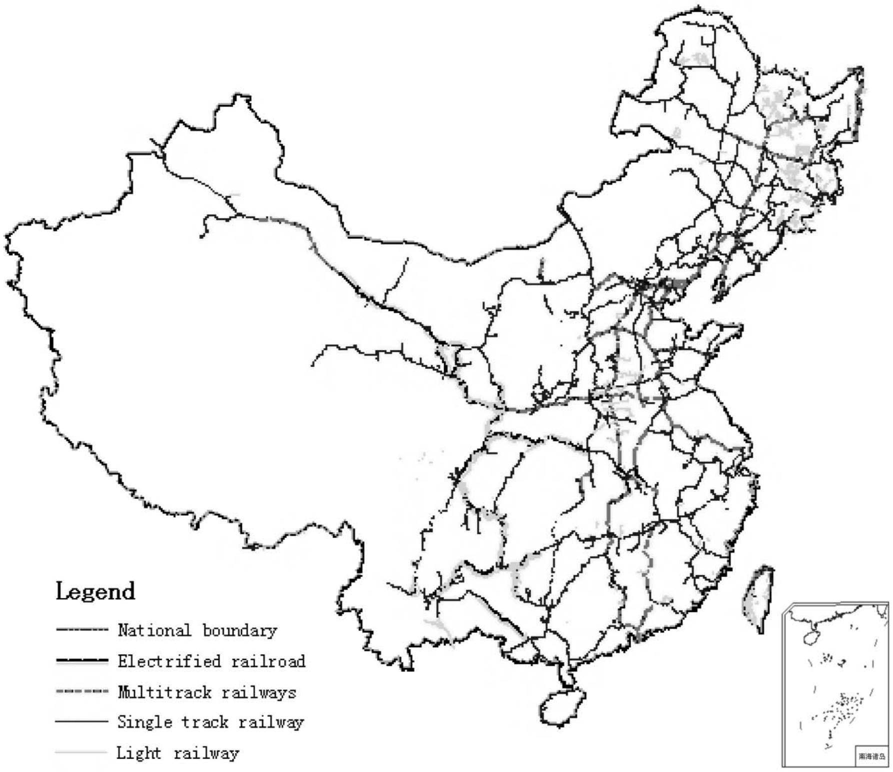

FIGURE 3. Spatial distribution of rails in 2000 in China.

Source: Resources and Environment Data Center, Institute of Geographical Science and Natural Resources Research, Chinese Academy of Sciences.

From the eastern margin of the second step eastward up to the coast is the third great topographic step. The largest plains of China, the Northeast China Plain, the North China Plain and the middle and lower Changjiang Plain are distributed in this step, which generally lie at elevation of below 200m.

SIMULATION PROCESS AND RESULTS

The Simulation Process

The major auxiliary tools of grid generation include the Control of MapObjects of geographical information system and Delphi computer lan-

TABLE 3

NPP in Terms of Land Cover Types on the Average

Land Cover type	NPP ($gC \cdot m^{-2} \cdot yr^{-1}$)
Shrub and open forest	1071
Evergreen broad-leaved forest	975
Deciduous broad-leaved forest	928
Coniferous and broad-leaved mixed forest	870
Farmland system	752
Evergreen coniferous forest	587
Deciduous coniferous forest	585
Grassland	271

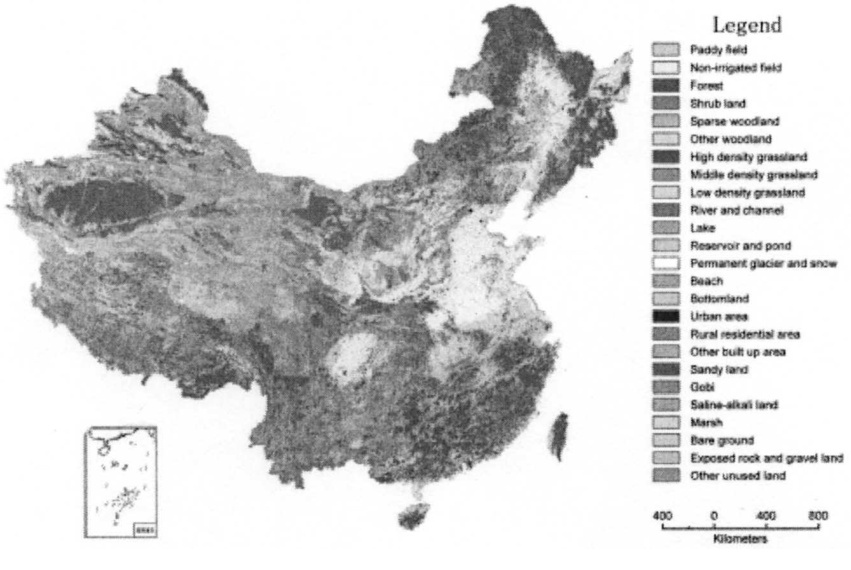

FIGURE 4. Land cover of China in 2000.

Source: Resources and Environment Data Center, Institute of Geographical Science and Natural Resources Research, Chinese Academy of Sciences.

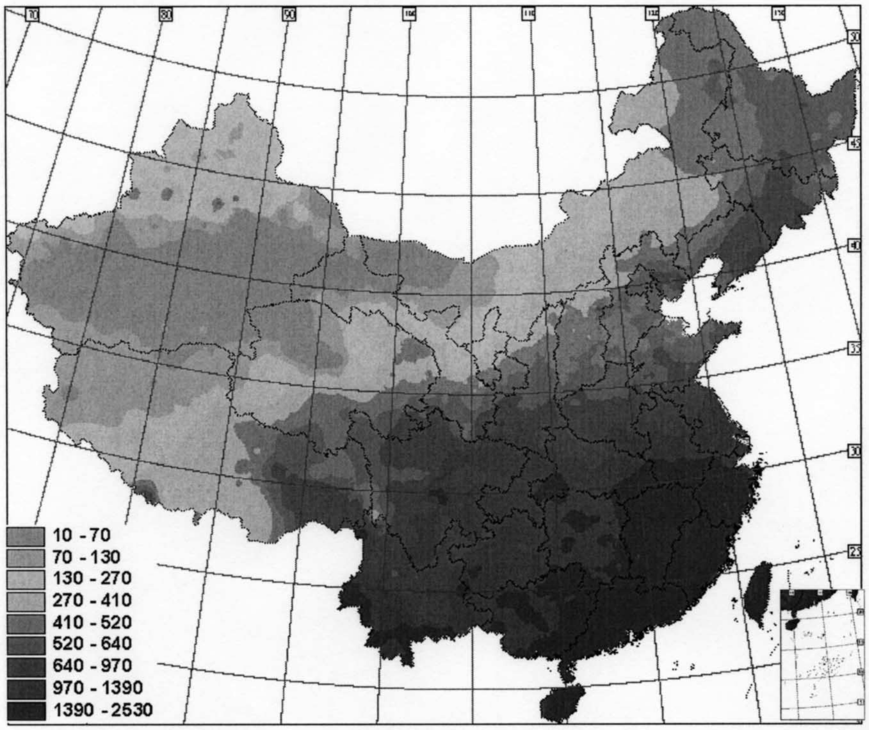

FIGURE 5. Annual precipitation of China in 2000 (unit: mm).

Source: Resources and Environment Data Center, Institute of Geographical Science and Natural Resources Research, Chinese Academy of Sciences.

guage. Eight data layers are involved, which are NPP (net primary productivity), DEM (digital elevation model), WA (water area), GridRail (railway network), GridRoad (road network), Chbnd (administrative boundary), Chzh(urban area) and Cityshp (geographical coordinate of city). The data are first pre-processed as follows: (1) converting NPP into vector data, (2) overlaying Chbnd with GridRoad and GridRail by Intersect and creating a data layer, ChBndNew, (3) adding fields, CityFlag for urban code and rural code and CityArea for areas of urban districts, in Chzh, (4) overlaying Chzh with ChBndNew by Intersect and creating a data layer, ChCity (5) overlaying NPP with ChCity by Intersect and creating a data layer, NppNew; (6) overlaying DEM with NppNew by Intersect and creating a data layer, DNpp; (7) overlaying WA with DNpp by Intersect and creating a data layer, WDNpp.

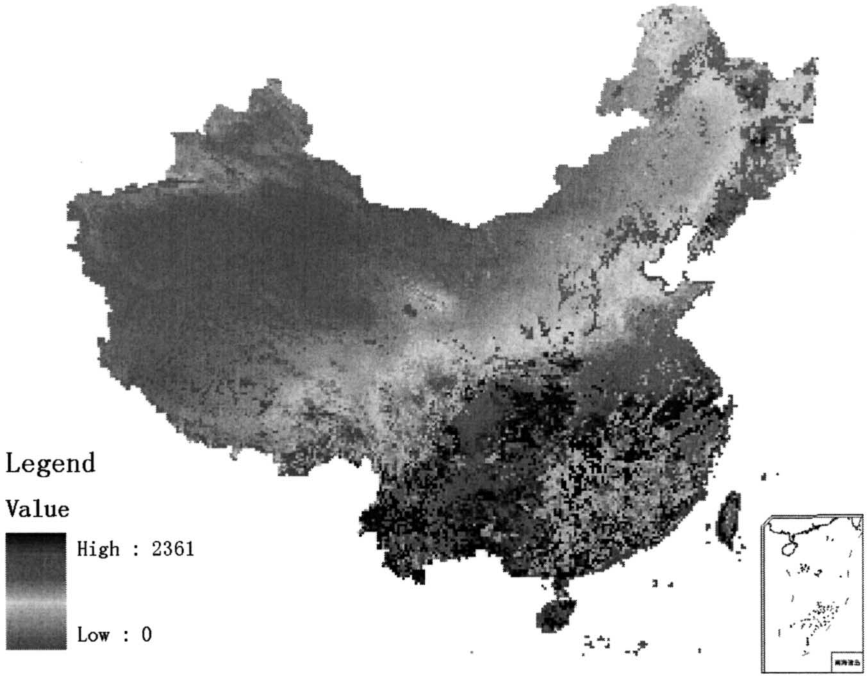

FIGURE 6. Spatial distribution of the mean NPP in 1990s in China. (unit: $gC \cdot m^{-2} \cdot yr^{-1}$).
Source: After Liu, 2001.

The MSPD grids are generated on the basis of WDNpp and CityShp, which includes 6 steps: (1) to read the attribute values of natural and socio-economic indicators at every MSPD grid, (2) to simulate the contribution of NPP and elevation to MSPD, (3) to define search radius of a MSPD grid and to search cities and transport infrastructures that have considerable effects on the MSPD grid, (4) to simulate the contribution of the searched cities and transport infrastructures to the MSPD grid, (5) to calculate the MSPD, and (6) Text file of the calculated result is converted into point vector data and grid data is created from the point vector data.

The MSPD of China can be formulated as a transformation from computational domain (i, j) to physical domain ($i, j, MSPD_{ij}$).

$$MSPD_{ij} = Pop_{n2000} \cdot \frac{P_{ij}}{\Sigma P_{ij}} \tag{3}$$

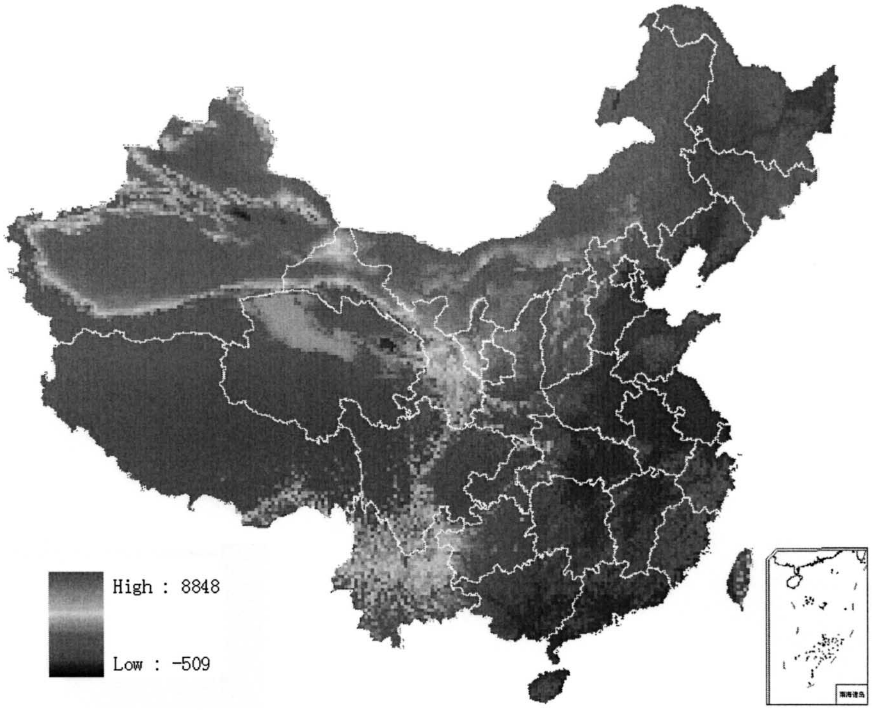

FIGURE 7. Digital elevation model of China (unit: m).

Source: Resources and Environment Data Center, Institute of Geographical Science and Natural Resources Research, Chinese Academy of Sciences.

$$P_{ij} = W_{ij}(Tran_{ij})^{1.3} \cdot (NPP_{ij})^{0.0001} \cdot (DEM_{ij})^{0.7} \cdot (u_{ij})^{1.2} \tag{4}$$

$$Tran_{ij} = \frac{ra_{ij} + ro_{ij}}{\underset{i,j}{Max}\{ra_{ij} + ro_{ij}\}} \tag{5}$$

$$NPP_{ij} = \exp\left\{ -\frac{(MNPP_{ij} - 800)^2}{10^6} \right\} \tag{6}$$

$$DEM_{ij}(t) = \begin{cases} \dfrac{500}{(dem_{ij}(\mathrm{t}))^2} & dem_{ij}(t) \geq 3700m \\ \dfrac{500}{dem_{ij}(t)} & 500m \langle dem_{ij}(t) \langle 3700m \\ 1 & dem_{ij}(t) \leq 500\mathrm{m} \end{cases} \tag{7}$$

$$u_{ij} = \sum_{k=1}^{M} \frac{S_k}{d_{ijk}} \tag{8}$$

where Pop_{n2000} is total population in administrative division n where grid (i, j) is located; W_{ij} is the indicative factor of water area, when grid cell (i, j) is located in water area $W_{ij} = 0$, or else $W_{ij} = 1$; ra_{ij} and ro_{ij} represent respectively rail density and road density at grid (i, j); S_k is size of the kth city; M is the total number of cities; d_{ijk} is the distance from grid (i, j) to the core grid of the kth city; $MNPP_{ij}$ is the mean net primary productivity annually in 1990s at grid (i, j); and dem_{ij} is elevation at grid (i, j).

Results

In 1935, Hu published his research results on the distribution of population in China and introduced a critical line, of which two end points Heihe city in Heilongjiang province and Tengchong city in Yunnan province. This Heihe-Tengchong line is located in the ecologically fragile zone where southeastern monsoon meets with westerlies (Chen, 2002). The elevation of the area on southeastern side of the Heihe-Tengchong line is 378m on an average. Hu found that 96% of population in China lived in the southeastern side of the Heihe-Tengchong line, where the area is 4.117 million square kilometers accounting for 42.9% of the whole of China. In 1990, the fourth census of China showed that on the southeastern side of the Heihe-Tengchong line there was 94.3% of the total population of China and population density on the southeastern side was 22 times the one on the northwestern side of the Heihe-Tengchong line (Zhang, 1997). The calculated result of MSPD shows that the southeastern side of the Heihe-Tengchong line had 90.8% of the total population of China in 2000 (as seen in Figure 8). In other words, the ratio of population on the northwestern side of the Heihe-Tengchong line to total population of China has been increasing since 1935. The growth rate was yearly 0.8% from 1935 to 1990 and yearly 6.1% from 1990 to 2000.

In terms of the provincially mean values, Shanghai, Tianjin and Beijing have the highest densities that are respectively 2089, 861 and 843 persons per square kilometer in 2000. Jiangsu, Shandong and Henan have higher densities that are respectively 719, 578 and 559 persons per square kilometer. The mean densities of Hunan, Hubei, Hebei, Chongqing, Anhui, Zhejiang and Guangdong range between 304 and 481 persons per square kilometer. The lowest densities appear in Tibet, Qinghai, Xinjiang and Inner Mongolia (as seen in Table 4 and Figure 9). In general, the average MSPD

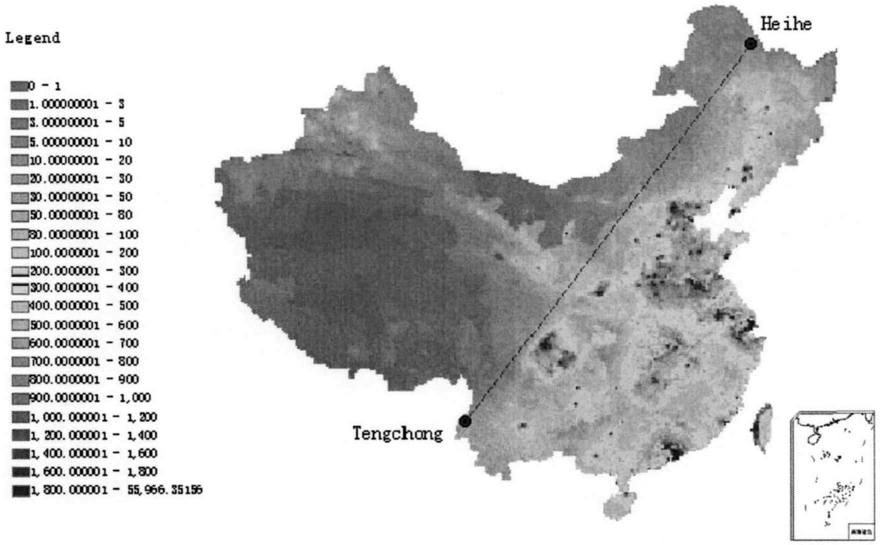

FIGURE 8. The population distribution of China in 2000 (unit: persons per square kilometer).

of eastern China and middle China are respectively 7.4 times and 5.7 times the one of western China in 2000.

DISCUSSION

In addition to MSPD, the recent studies related to population distribution of China include transforming population data from census to grid (GPW) (Tobler et al., 1997), apportioning census counts to each grid cell based on probability coefficients (LandScan) (Dobson et al., 2000), and modeling the population of China using nighttime light data (MNLD) (Lo, 2001).

Population data are routinely collected by censuses and compiled for administrative units. Although this approach is essential for certain types of analyses, it has a limit to cross-disciplinary studies. The population data need to be referenced to a uniform coordinate system. This led to the initial version of GPW, which was released in 1995 (Tobler et al., 1995). It simply allocated total population on an administrative unit over grid cells proportionally under the assumption that population is distributed evenly over the

TABLE 4

The Provincially Average MSPD in 2000 not Including Taiwan, Hong Kong and Macao Temporarily in 2000

Region	Population Size (million persons)	Area (km^2)	Population Density (person/km^2)
Western China	*355.31*	*6725746*	*51*
Inner Mongolia	23.76	1143327	21
Guangxi	44.89	236544	190
Chongqing	30.90	82390	375
Sichuan	83.29	483759	172
Guizhou	35.25	176109	200
Yunnan	42.88	383101	112
Tibet	2.62	1201653	2
Shaanxi	36.05	205732	175
Gansu	25.62	404622	63
Qinghai	5.18	716677	7
Ningxia	5.62	51785	109
Xinjiang	19.25	1640111	12
Middle China	*415.64*	*1670726*	*293*
Shanxi	32.97	156563	211
Anhui	59.86	140165	427
Jiangxi	41.40	166960	248
Henan	92.56	165619	559
Hubei	60.28	185950	324
Hunan	64.40	211815	304
Jilin	27.28	191093	143
Helongjiang	36.89	452561	82
Eastern China	*491.33*	*1203528*	*379*
Beijing	13.82	16386	843
Tianjin	10.01	11620	861
Hebei	67.44	188111	359
Liaoning	42.38	146316	290
Shanghai	16.74	8013	2089
Jiangsu	74.38	103405	719
Zhejiang	46.77	103196	453
Fujian	34.71	122468	283
Shandong	90.79	157119	578
Guangdong	86.42	179776	481
Hainan	7.87	40070	196

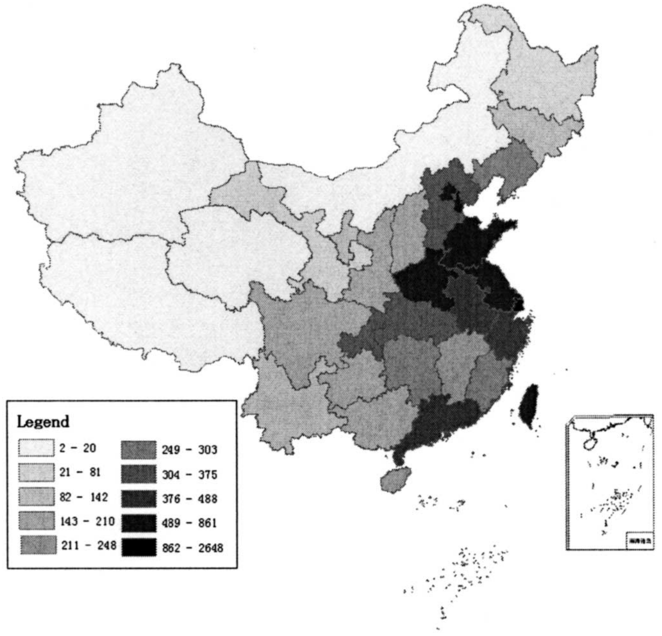

FIGURE 9. The provincially average MSPD in 2000 (unit: persons per square kilometer).

administrative unit. The second version, which released in 2001 (http://www.ciesin.org/datasets/gpw/globldem.doc.html), placed its emphasis on improving the resolution of the input data layers of administrative boundaries and on producing better population estimates for each unit.

The LandScan 2000 Global Population Database was developed by Oak Ridge National Laboratory for the United States Department of Defense(http://www.ornl.gov/gist/landscan/LandScan_2000_Release.htm). Best available census counts were distributed to grids in terms of probability coefficients that based on publicly available databases offering worldwide coverage of roads, slope, land cover and nighttime lights. The Landscan 2000 processed and transformed all data into a 30-by-30-second latitude/longitude grid system by means of a smart interpolation procedure (Sutton et al., 1997).

Lo (2001) estimated population distribution of China in 1997 at three different spatial scales of province, county, and city, using Radiance-calibrated nighttime light data of China from Operational Linescan System (OLS)

TABLE 5

Comparisons of MSPD with GPW, LandScan and MNLD

Summary	MSPD	LandScan	MNLD	GPW
Resolution (grid size)	1km	1km	1km	5km
Input variables of model	Census data of population; Land cover; transport infrastructures; DEM; spatial distribution of cities; net primary productivity based on NOAA/AVHRR data, daily meteorological data, and soil data.	Census data of population; land cover; roads; DEM; nighttime lights.	Night-time lights.	Census data of population.

of Defense Meteorological Satellite Program (DMSP) acquired between March 1996 and January–February 1997. Allometric growth models and linear regression models were developed to estimate population distribution of China in terms of light area, light volume, pixel mean, and percent light area.

All these models compared their results with statistical population data. They found that their results overall closely approximated the statistics and even some of them showed no difference with the statistical population. However, the accuracy of the statistical data is doubtable in many areas. The verification and validation studies should be conducted on the basis of spatially sampling, in which global positioning system must be used so that the most essential data on geographical coordinates of every sample could be acquired.

By comparing MSPD with Landscan, MNLD and GPW (as seen in Table 5), we can find that MSPD is based on much more essential data than other 3 ones methodologically. Especially, when scenarios of land cover, spatial distributions of transport infrastructures and cities are available for the MSPD, scenarios of spatial population distribution can be developed on the basis of total population forecast.

ACKNOWLEDGMENT

This work is supported by Projects of National Natural Science Foundation of China (40371094 & 90202002), and by National Basic Research Priorities Program (2002CB4125) of Ministry of Science and Technology of the People's Republic of China.

REFERENCES

Adams, J., (1968). A population map of West Africa, Graduate School of Geography discussion paper No. 26. London: London School of Economics.

Ahuja, D.V., & Coons, S.A., (1968). geometry for construction and display. *IBM Systems Journal, 7,* 188–205.

Bracken, I., & Martin, D. (1989). The generation of spatial population distribution from census centroid data. *Environment and Planning A 21*, 537–543.

Cassedy, J. H. (1984). *American Medicine and Statistical Thinking, 1800–1860*. Cambridge, MA: Harvard University Press.

Chen, H., & Zhang, W.C. (2000). *Transport Geography of China*. Beijing: Science Press (in Chinese).

Chen, S. P., (2002). The spatial-temporal analysis of census. *China Population, Resources and Environment 12(4),* 3–7.

Contanza, R., d'Arge, R., de Groot, R., Farber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S., O'Neill, R.V., Paruelo, J., Raskin, R.G., Sutton, P., & van den Belt, M. (1997). The value of the world's ecosystem services and natural capitals. *Nature,* 387, 253–260.

Cullen, J. (1975). *The Statistical Movement in Early Victorian Britain: the foundation of empirical social science*. Brighton: Hassocks.

Deichmann, U. (1996). *A Review of Spatial Population Database Design and Modeling*. Technical Report 96-3, National Center for Geographic Information and Analysis, USA.

Dobson, J.E., Bright, E.A., Coleman, P.R., Durfee, R.C., & Worley, B.A. (2000). LandScan: a global population database for estimating populations at risk. *Photogrammetric Engineering & Remote Sensing,* 66(7), 849–857.

Flowerdew, R., & Green, M. (1989). Statistical methods for inference between incompatible zonal systems. In M.F. Goodchild & S. Gopal (Eds.), *Accuracy of Spatial Database* (pp. 239–248). London: Taylor and Francis.

Flowerdew, R., Green, M., & Kehris, E. (1991). Using areal interpolation methods in geographic information systems. *Papers in Regional Science,* 70, 303–315.

Goodchild, M.F., Anselin, L., & Deichmann, U. (1993). A framework for the areal interpolation of socioeconomic data. *Environment and Planning,* A 25, 383–397.

Goodchild, M.F., & Lam, N.S.N. (1980). Areal interpolation: a variant of the traditional spatial problem. *Geo-Preocessing,* 1, 297–312.

Hu, H.Y. (1935). The distribution of population in ChinaåLBOLE_LINKä. *Acta Geographica Sinica,* 2(2), 32–74 (in Chinese).

Hu, H.Y. (1983). *Discussions on Distribution of Population in China*. Shanghai: Eastern China Normal University Press (in Chinese).

Institute of Geography of Chinese Academy of Sciences. (1987). *The Population Atlas of China*. Hongkong: Oxford University Press.

Li, J.W. (2000). Opportunities and challenges facing the urban development of China in 21st century. In Liu, G.G. (Ed.), *The Urban Development of China in the 21st Century* (pp. 121–148). Beijing: Huixi Press (in Chinese).

Lieth, H., & Whittaker, R.H. (1975). *Primary Productivity of the Biosphere*. New York: Springer-Verlag.

Liseikin, V.D. (1999). *Grid Generation Methods*. Berlin: Springer–Verlag.

Liu, J., Chen, J.M., Cihlar, J., & Park, W.M. (1997). A process-based boreal ecosystem productivity simulator using remote sensing inputs. *Remote Sensing of Environment,* 62, 158–175.

Liu, J.Y., Liu, M.L., Zhuang, D.F., Zhang, Z.X., & Deng, X.Z. (2003). Study on spatial pattern of land-use change in China during 1995–2000. *Science in China (Series D)*, 46, 373–384.

Liu, M.L. (2001). Land-Use/Land-Cover Change and Vegetation Carbon Pool and Productivity of Terrestrial Ecosystems in China. *Doctoral Thesis of Institute of Geographical Sciences and Natural Resources Research of CAS*.

Lo, C.P. (2001). Modeling the population of China using DMSP Operational Linescan System nighttime data. *Photogrammetric Engineering & Remote Sensing,* 67(9), 1037–1047.

Martin, D., & Bracken, I. (1991). Techniques for modeling population-related raster databases. *Environment and Planning,* A 23, 1069–1075.

Mennis, J. (2003). Generating surface models of population using dasymetric mapping. *The Professional Geographer,* 55(1), 31–42.

Mirrison, D. (1962). Optimal mesh size in the numerical integration of an ordinary differential equation. *J. Assoc. Comput. Machinery,* 9, 98–103.

Moxey, A., & Allanson, P. (1994). Areal interpolation of spatially extensive variables: a comparison of alternative techniques. *International Journal of Geographical Information Systems,* 8(5), 479–487.

National Bureau of Statistics of People's Republic of China. (2001). *China Statistical Yearbook*. Beijing: China Statistics Press.

Parr, J.B. (1985). The form of the regional density function. *Regional Studies,* 19, 535–546.

Plane, D.A., & Rogerson, P.A. (1994). *The Geographical Analysis of Population: with Applications to Planning and Business*. New York: J. Wiley.

Sidorov, A.F. (1966). An algorithm for generating optimal numerical grids. *Trudy MIAN USSR,* 24, 147–151 (in Russian).

Sutton, P. (1997). Modelling population density with night-time satellite imagery and GIS. *Comput., Environ. and Urban Systems,* 21(3/4), 227–244.

Sutton, P., Elvidge, C., & Obremski, T. (2003). Building and evaluating models to estimate ambient population density. *Photogrammetric Engineering & Remote Sensing,* 69(5), 545–553.

Sutton, P., Roberts, D., Elvidge, C., & Baugh, K. (2001). Census from Heaven: an estimate of the global human population using night-time satellite imagery. *International Journal of Remote Sesnsing,* 22(16), 3061–3076.

Sutton, P., Roberts, D., Elvidge, C., & Meij, H. (1997). A comparison of nighttime satellite imagery and population density for the continental United States. *Photogrammetric Engineering & Remote Sensing,* 63(11), 1303–1313.

Tobler, W.R. (1979). Smooth pycnophylactic interpolation for geographical regions. *Journal of the American Statistical Association,* 74, 519–530.

Tobler, W., Deichmann, U., Gottsegen, J., & Maloy, K. (1995). *The Global Demography Project, Technical Report TR-95–6*. National Center for Geographic Information and Analysis, Santa Barbara, USA.

Tobler, W., Deichmann, U., Gottsegen, J., & Maloy, K. (1997). World population in a grid of spherical quadrilaterals. *International Journal of Population Geography, 3*, 203–225.

Urban Society and Economy survey Team of National Bureau of Statisitics of People's Republic of China. (2001). *Urban Statistical Yearbook of China*. Beijing: China Statistics Press (in Chinese).

Woods, R., & Rees, P. (1986). Spatial demography: themes, issues and progress. In Woods, R. & Rees, P. (Eds.), *Population Structures and Models: Developments in Spatial Demography* (pp. 1–3). London: Allen & Unwin.

Zhang, S.Y. (1997). *Population Geography of China*. Beijing: Business Press House (in Chinese).

Zhao, S.Q. (1986). *Physical Geography of China*. New York: John Wiley & Sons.