

中国人口空间分布不均匀性分析及其可视化

韩嘉福^{1,2}, 张忠¹, 齐清文¹

(1 中国科学院地理科学与资源研究所, 北京 100101; 2 中国科学院研究生院, 北京 100049)

摘要: 中国人口空间分布的不均匀性是一种客观存在的现象, 自从 1935 年胡焕庸先生提出“爱辉—腾冲线”以来, 中国人口东西部的不均匀现象并没有发生显著的变化。这种不均匀不仅仅存在于“胡焕庸线”两侧, 而是广泛存在的。本文引用经济学家研究收入分配平均程度常用的 LORENZ 曲线方法, 结合相关图、表, 对中国人口空间分布的不均匀性进行研究。通过对曲线特征点的分析, 结合地图可视化可以形象地反映人口空间分布的特点, 同时对人口空间分布进行定量化描述。

关键词: LORENZ 曲线; 人口分布; 不均匀性; 可视化

1 引言

多元统计分析中的 Logistic 回归、因子分析、聚类分析等是人口学家进行人口研究常用的方法^[1]。从地理学角度进行人口空间分布格局分析的基础是将统计数据空间化, 形成 GIS 的人口地理信息数据库, 在此基础上通过一定的技术手段进行分析。

中国人口空间分布的不均匀特征早有研究, 统计分析、比较分析是常用的方法^[2~4], LORENZ 曲线法也是研究人口空间分布不均匀性研究的一种方法^[5~7]。中国人口空间分布的基本特征是, 地区差异大, 既有每平方公里 500 人以上的高度密集的地区, 也有每平方公里不足 10 人的人烟稀少的地区。人口分布的总体格局是东部密集, 西北稀疏^[3]。

人口密度是衡量人口空间分布的有效手段, 但由于掌握信息尺度的不同, 往往得到不同的结果, 宏观尺度往往会掩盖人口空间分布的这种不均匀性。研究人口空间分布最直接最可靠的数据是人口普查数据, 人口普查是被广泛采用的搜集人口资料的一种科学方法, 是提供全国基本人口数据的主要来源, 统计单元越小, 越能反映人口密度变化的细部特征, 统计单元越大则越有可能掩盖这种不均匀性。

本文利用汇总到乡、镇、街道办事处的人口普查数据, 通过绘制中国人口空间分布 LORENZ 曲线和中国人口分布高程 LORENZ 曲线, 从水平和垂直两个方向定量化研究人口空间分布的不均匀性, 并对不同特征的人口空间分布进行地图可视化, 以形象直观地表达这种不均匀性。

2 人口空间分布研究方法

(1) LORENZ 曲线是美国经济学家 LORENZ 在 20 世纪初提出的, 利用累积次数分配曲线描述一个国家或地区收入分配平均程度的一种图示方法。将各组数据按收入从少到多排序后, 以相对累积人口数作为横轴, 相对累积总收入作为纵轴, 将各组数据表示在笛卡儿坐标系中, 形成的曲线即为 LORENZ 曲线, 如图 1 所示曲线 I 和曲线 II 表示两个地区的 LORENZ 曲线。45°对角线为绝对平均线, 135°对角线为对称轴。曲线越接近 45°对角线表明收入越平均, 反之则出现两极分化的趋势。

意大利统计学家基尼 (Gini) 根据洛伦茨曲线的特性, 将它指数化并命名为基尼系数, 基尼系数是定量描述 LORENZ 曲线常用的方法。它是洛伦茨曲线图中绝对均匀线与 LORENZ 曲线之间的面积和绝对均匀线与绝对不均匀线之间的面积之

收稿日期: 2006-07-12 修回日期: 2007-09-25

作者简介: 韩嘉福 (1975-), 男, 曾多年从事地图制图及 GIS 数据获取、处理、分析相关技术及管理工作, 现为硕士研究生, 主要从事地图可视化及 GIS 应用研究。

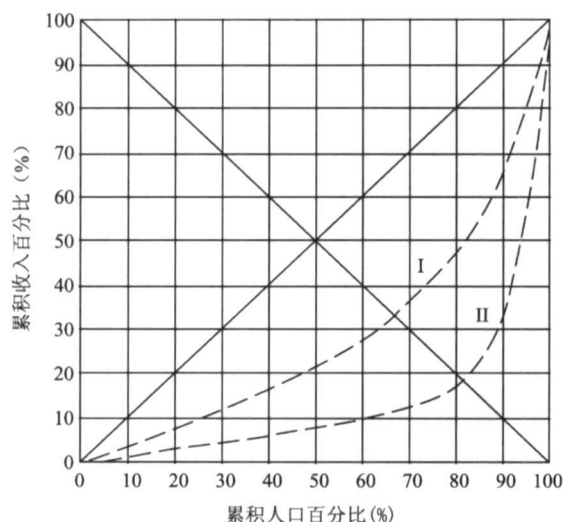


图 1 LORENZ 曲线

Fig. 1 LORENZ curves

比^[6]，在数值上等于曲线与绝对平均线构成图形的面积的 2 倍。其值域为 $[0, 1]$ ，如果所描述的现象均匀分布，则 $G = 0$ 。如果绝对不均匀则 $G = 1$ 。基尼系数的大小定量的反映了所描述现象的均衡程度。

在人口空间分布中我们可以借鉴 LORENZ 曲线方法：将所有乡、镇、街道办事处按密度从小到大排序，从人口密度最低的乡镇算起，进行人口数和面积的累积，分别求得该数值占全国总人口和总面积的百分比，每个乡、镇、街道办事处对应一组数值。以累积人口百分比为横轴，以累积面积百分比为纵轴，将该组数值绘制到平面直角坐标系中，连接相邻点得到中国人口分布 LORENZ 曲线；将全国 1:25 万 DEM 按 50m 一个高程带，将高程在 50 m 以下的分为第一带，50~100 m 为第 2 带，100~150 m 为第三带，依次类推将全国 DEM 重新分类。通过 GIS 的空间叠加分析方法，将乡、镇、街道办事处空间点位数据与 DEM 进行空间叠加，获取所有乡、镇、街道办事处驻地的高程。将全国乡、镇街道办事处按照高程从小到大的顺序排序，得到每个高程带的面积百分比和人口百分比，分别按人口百分比和面积百分比进行累加，并绘制在二维平面直角坐标系中，连接相邻点形成人口-高程 LORENZ 曲线。上述曲线对应的基尼系数可以量化的描述中国人口空间分布的不均匀性。

(2) 地图如同语言文字一样是人类社会信息

交流的基础，在地理学研究中，传统地图学主要是以视觉传输的形式实现其功能的，地图学家通过对地学现象进行深刻理解综合分析以及抽象等一系列复杂科学思维和创造，利用符号来反映地物、现象和地理过程。地图作为信息传递的工具其最大的优点是具有直观性，是地理学研究的一种重要的工具。利用地图可以分析制图对象的分布规律、各种现象的相互联系、现象的动态变化等。在人口空间分布规律的研究中地图可视化是一个有效的手段，通过对人口空间分布数据进行适当的处理，选择合适的表现形式可以形象直观的反映人口空间分布的规律。

(3) 普查人口数据是一种典型的统计数据，不具备明确的空间特征，需要对统计的人口数据进行地理编码，实现人口统计数据与空间位置信息的统一。选用 1:25 万基础数据库，提取其中的镇级居民点，与第五次人口普查汇总到乡镇的人口数据进行连接，并根据现势资料，对发生变化的乡镇进行空间定位，可以得到全国所有乡、镇、街道办事处的人口空间分布数据。

从国家 1:25 万基础数据库中提取相关基础数据如：数字高程模型 (DEM)、各级居民地数据 (省会、地级市、县、乡镇及街道办事处、村庄) 等，作为人口空间分布研究的基础数据。这些数据是实现人口统计数据空间化的基础。根据居民地代码提取乡镇级居民地，形成 shp 文件，根据乡、镇、街道办事处的行政区划代码可以实现空间数据与人口数据的连接，输入乡、镇、街道办事处的面积数据即可得到相应的平均人口密度。

将 DEM 重新采样为 $1\text{km} \times 1\text{km}$ 。利用 ArcGIS 的空间分析功能，获取每个乡镇点的高程。

下文提到的数据应遵循以下约定：

①为叙述方便，下文凡涉及到乡、镇、街道办事处级居民地时均用乡镇代替。

②人口数据来源于 2000 年第五次人口普查汇总到乡镇的数据，该数据中不包括 2000 年现役军人。

③面积数据来源于《中国政区大典》的公布面积，该面积中不包括沙漠、戈壁、原始森林、大面积水域。

④参与分析的乡镇点中不包括部分 (约 7%) 没有正式审批的虚拟乡、虚拟镇、虚拟街道办事处，以及其他没有正式审批的乡镇级政府。因此

参与分析的总人口约 11.8 亿, 总面积为所有乡镇的公布面积之和约 804 万 km^2 。

3 人口空间分布的不均衡性分析

3.1 人口空间分布在平面上的不均匀性

根据基尼系数的计算方法, 可以计算中国乡镇人口分布 LORENZ 曲线的基尼系数为 0.79, 这种不均匀性主要是由于人口密度较小的乡镇占有了相对较多的面积, 而人口数相对较少造成的。观察图 2 中曲线, 可以发现: 距离坐标原点较近端, 人口密度较小的乡镇占据着较大的面积, 而居住在该面积中的人口则只占全国人口很小的比重; 而从距离原点较远的一端可以发现, 人口密度较大的乡镇拥有较多的人口, 而面积则相对较小。

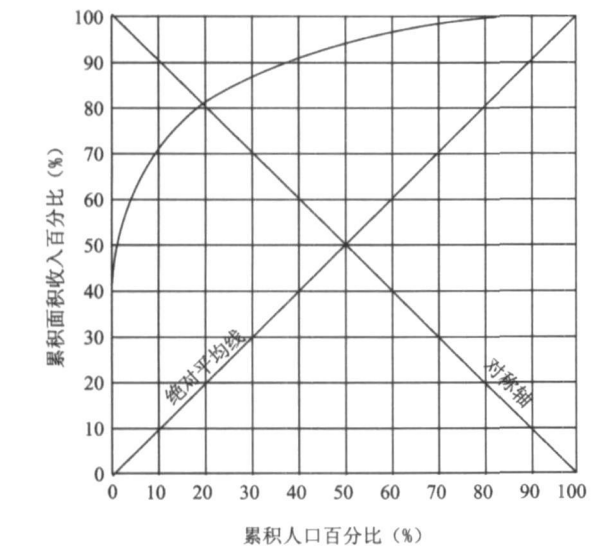


图 2 中国乡镇人口分布 LORENZ 曲线
Fig. 2 China's township population distribution LORENZ curve

根据 LORENZ 曲线的生成方法, 我们可以知道任何一个人口百分比数值和面积百分比数值所对应的镇平均密度, LORENZ 曲线的缺点是只能定量描述人口空间分布的不均匀性, 却无法确定这种不均匀性在空间上的存在状态。根据这个密度对全国乡镇点进行地图可视化可以弥补这方面的不足。将 LORENZ 曲线反映的事实表现在地图上, 得到比较形象、直观的印象。结合表 1 和图 2 我们可以发现:

(1) 人口密度小于 19 人 / km^2 的地区面积约占总面积的 49%, 人口则只有 1% 左右, 密度大

于 8 518 的乡镇面积仅占约不足 0.1% 的地区聚集着 10% 的人口。以人口密度 19 人 / km^2 作为分界点, 建立乡镇人口密度分布图, 见图版 I 彩图 1。

表 1 中国乡镇人口分布 LORENZ 曲线坐标表
Tab 1 Coordinate of China's township population distribution LORENZ curves

单位: 人 / km^2				
截止密度	面积百分比 (%)	人口百分比 (%)	平均密度	乡镇个数 (个)
0	49.335	1.022	3.0	
19	22.249	9.050	59.8	2713
112	9.719	9.998	151.1	9677
200	5.900	10.009	249.2	6920
310	4.016	9.957	364.2	5390
425	3.037	9.996	483.6	4538
551	2.367	9.992	620.3	3974
702	1.805	9.997	813.7	3545
987	1.088	9.984	1348.4	3121
2097	0.397	9.999	3700.2	2384

从图版 I 彩图 1 可以发现, 这部分地区是中国人口分布最稀疏的地区, 该类地区主要集中在胡焕庸线以西及其附近地区, 以及东北的大、小兴安岭及长白山地区, 小部分乡镇零散分布于云贵高原、大巴山地区、南岭地区、浙闽丘陵地带。这类地区的特点是地理环境恶劣, 多为戈壁、沙漠、原始森林等。该类地区面积广阔但不适宜人居住, 相信这种现象即使在将来相对长时间内也不会有太大的改变。

(2) 密度大于 8 518 人 / km^2 的乡镇街道办事处面积仅占中国不足 0.1%, 该地区聚集着约 10% 的人口, 以 8 518 人 / km^2 作为分界点, 建立中国乡镇级居民地人口密度分布图参见图版 I 彩图 2。通过与全国县及县以上居民地的叠加分析, 我们可以发现, 该类乡镇级居民地主要分布在城市中心点 10km 范围内, 说明这部分乡镇主要是由各类街道办事处构成的, 主要为城镇居民。

(3) 从图 2 曲线中我们可以发现, 中国人口的空间分布状态也符合二八定律: 约 80% 的人口分布在约 20% 的地区, 另外约 20% 的人口分布在约 80% 的地区。符合二八定律的人口密度分界点在 190 / km^2 人左右, 也就是说人口密度大于 190 / km^2 的乡镇集中了 80% 的人口。

我们以人口密度为 190 人 / km^2 作为分界点建立中国乡镇人口密度分布图, 详见图版 I 彩图 3。

可以发现,人口密度大于 190人 /km² 的乡镇街道办事处主要分布在“胡焕庸线”以东。该分布具有集中分布于零星分布相结合的特征,集中分布地区主要包括黄淮海平原、长江中下游平原、四川盆地、东南沿海地区及其连接的珠江三角洲和广西盆地、关中盆地、东北地区以哈尔滨 – 长春 – 沈阳连线的附近地区、云贵高原部分地区等;零散分布则主要是在城市周围地区。

同样道理可以根据实际情况,以任意数值作为阈值建立中国乡镇人口空间分布密度图。结合上述 LORENZ 曲线可以对该分布进行定量描述。

从上分析发现,中国人口空间分布的不均匀性是广泛存在的,在水平方向上这种不均匀性尤为明显,主要表现在两个方面:适宜人类居住地的虽然面积狭小,但聚集了较多的人口,条件比较恶劣的地区的人口则比较稀疏,这部分地区并不仅仅分布于西部,在东部的人口密集区中同样具有局部稀疏地区;城市是典型的人口密集区,该类地区以相对较小的面积,聚集了较多的人口,人口密度明显高于偏远地区。

3.2 人口空间分布在垂直方向上的不均匀性分析

中国人口高程分布 LORENZ 曲线对应的基尼系数为 0.58 从总体上来说,人口在垂直方向上是不均匀的,人口空间分布随着高程的升高而减少,但这种不均匀性没有水平方向的不均匀性明显。结合表 2 及图 3 曲线我们可以发现,中国人口在垂直方向的分布十分广泛,从吐鲁番地区的高程小于 0 到青藏高原 5600 m 左右均有乡镇级居民地分布。高程较低的各带人口分布较密,低于 100 m 地带分布有 51.6% 的人口,面积只占 13.5%,平均密度约 561.2 人 /km²; 2300 m 以上面积占 23.9%,人口不足 1%,平均密度只有 5.7 人 /km²; 同时,在局部各高程带的人口分布是相对均匀的,如高程在 100 ~ 1000 m 范围内的各高程带人口占全国人口的百分比和面积百分比均为 38% 左右,接近于均匀分布,平均密度约为 145.9 人 /km²。正是这个高程带的人口空间分布的均匀性减弱了中国人口空间分布的总体不均匀性。

图版 I 彩图 4 则以形象直观的方式定量地表示了上述不均匀性的实际分布情况。高程 100 m 以下的地区是中国人口分布最稠密的地区,从图版 I 彩图 4 可以发现该类地区主要集中在沿海、

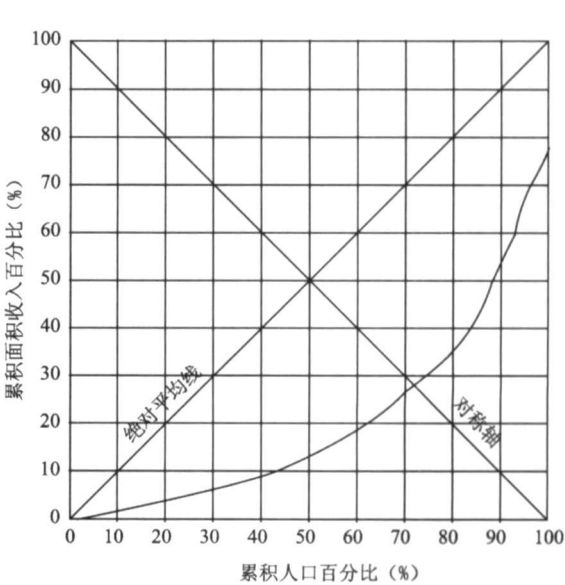


图 3 中国人口高程分布 LORENZ 曲线

Fig. 3 China's population elevation distribution LORENZ curves

沿江、沿湖地带,包括黄淮海平原和长江中下游平原地区,这些地区由于地势低平,气候适宜特别适合人类居住,同时我国近年来的经济政策主要偏向沿海地区,该地区城市发展迅速,城镇化程度较高,所以具有较高的人口密度,2 300 m 以上人口密度较低的高程带,该类地区主要集中在青藏高原地带和新疆的部分地区,较高的海拔和恶劣的环境是该地区人口稀疏的主要原因。这种现状一般不会随着时间的推移而发生显著的变化。一般来说人口自然迁移主要趋向于海拔较低的地区,因此这种稀疏状态可能会加剧。但是由于人口基数较小,受人口自然增长率的影响可能会较大,高人口自然增长率会对现状略有改变,同时低人口自然增长率则会加剧这种稀疏状态。

表 2 中国人口高程分布 LORENZ 曲线特征点表

Tab 1 Coordinate of China's population elevation distribution LORENZ curve				
单位: 人 /km ²				
高程 (m)	面积百分 比 (%)	人口百分 比 (%)	平均密度	乡镇个数 (个)
< 0	13.499	51.557	561.2	16432
100	37.141	36.891	145.9	20782
1000	25.430	10.622	61.4	7870
2300	23.930	0.930	5.7	1935
5650				

可以发现, 在垂直方向上人口空间分布的不均匀性同样存在, 总体趋势是人口密度随高程降低, 但与水平方向的不均匀性具有不同的特点, 高程较低的地区人口密度明显高于平均水平, 高程较高的地区人口密度明显低于平均水平, 而高程处于中间的地带如 100~1000 m 范围内的人口空间分布则是接近均匀的, 从图版 I 彩图 4 我们可以发现, 该类地区主要以平原丘陵为主, 包括东北平原、四川盆地等地区。

4 结语

人口空间分布 LORENZ 曲线可以依据统计数据建立, 反映的是数据的统计特征。而地图可视化反映的是人口实际的空间分布状态。将人口空间分布 LORENZ 曲线与地图可视化方法结合起来, 一方面可以定量描述人口空间分布的统计特征和现实状态, 同时可以根据 LORENZ 曲线的特征, 选择不同特征点指导地图分级, 科学合理的分级方法是进行地图分析的基本前提, 不合理的分级方式有可能会掩盖这种特征。

通过绘制不同形式的人口分布 LORENZ 曲线, 结合相关图表可以定量分析这种人口空间分布的不均匀性。与地图可视化的结合可以将这种量化的结果形象地表现在地图上, 限于篇幅本文仅就人口-乡镇 LORENZ 曲线及人口高程 LORENZ 曲线中的重要特征点进行了地图可视化, 并对其反映的现象进行了分析与描述。通过对相关 LORENZ 曲线的深入分析, 并结合地图可视化, 绘制人口空间分布系列地图, 建立人口空间分布

信息图谱, 研究中国人口空间分布现状, 结合前几次人口普查数据进行人口时空特征分析是研究人口空间分布规律的有效方法。

参考文献

- [1] 刘金塘, 伍小兰. 多元统计分析技术在人口研究中的应用. 人口研究, 2002 (3): 1~6
- [2] 韩惠, 刘勇, 刘瑞雯. 中国人口分布的空间格局及其成因探讨. 兰州大学学报 (社会科学版), 2000, 28 (4): 16~21.
- [3] 宋健. 中国人口总体状况分析. 人口研究, 2005, 9: 92~96
- [4] 段成荣, 梁宏. 中国人口地区分布现状及其历史变迁. 市场与人口分析, 2002 (5): 67~73
- [5] 张果, 厉红兰. 成都市人口地域分布的研究. 四川师范大学学报 (自然科学版), 2004 (7): 414~418
- [6] 祝俊明. LORENZ 曲线在人口分布研究中的应用. 人口战线, 1994 (2): 11~16
- [7] 刘德钦, 刘宇, 薛新玉. 中国人口分布及空间相关分析. 测绘科学, 2004 (12): 76~79
- [8] Duangkamon Chotkapanich, William E Griffiths Averaging Lorenz curves. Journal of Economic Inequality, 2005 (3): 1~19.
- [9] Claudio Zoli. Intersecting generalized Lorenz curves and the Gini index. Soc. Choice Welfare, 1999 (16): 183~196
- [10] 刘岳, 沈益民, 奚国金. 中国人口分析与区域特征. 北京: 海洋出版社, 1991.
- [11] 廖克. 现代地图学. 北京: 科学出版社, 2003
- [12] 王英杰, 袁勘省, 余卓渊. 多维动态地学信息可视化. 北京: 科学出版社, 2003

Analysis and Visualization of the Uneven Distribution of Population in China

HAN Jiafu^{1,2}, ZHANG Zhong¹, QI Qingwen¹

(1 Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China;

2 Graduate School of the Chinese Academy of Sciences, Beijing 100039, China)

Abstract China's population distribution is quite inequable. Such a kind of inequalities exists widely and changes

slightly. Use for reference of economics we use the Lorenz curves to research these inequalities. Combining with the map visualization and corresponding tabular data we can make some quantitative analysis. We analyzed from two aspects, the horizontal and vertical direction. For convenient research we collect the 1:250 000 databases including residential area distribution and digital elevation modal data. Based on the research of China's township population distribution LORENZ curves we found some changes in special densities and generated corresponding maps in order to show these inequalities vividly, and by research of the China's population-elevation distribution LORENZ curves we can compile different elevation population distribution maps to show these inequalities. Compared with the Gini coefficient we found that the inequalities in horizontal and vertical directions have different characteristics.

Key words LORENZ curve; population distribution; inequality; visualization

祝贺中国科学院对地观测与数字地球科学中心成立

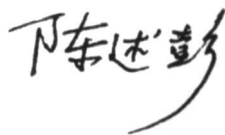
改革开放以来,从第六个五年规划开始,全国协作,自力更生,连续三个五年计划中,中国科学院主持承担起研究航空、航天传感器的任务,包括微波、近红外、可见光五类 17 种(当时国际商品大约 25 种),组织 14 次遥感中间试验和应用示范工程(包括净月潭、哈密、腾冲、二滩等)。完成三北防护林,长江、黄河洪涝灾情预警系统等重大攻关项目。突破了当时的国际技术封锁,为今天空间科学技术的发展,打下了初步的基础,争取了社会的认可,获得了政府的支持。

与此同时,在邓小平同志的亲切关怀下,组建了“遥感卫星地面接收站”,利用国际、国内两种空间信息资源,加以消化吸收,及时为国家资源、环境、能源、国土、海洋普查、城市建设,提供信息服务,促进了国际合作与交流。

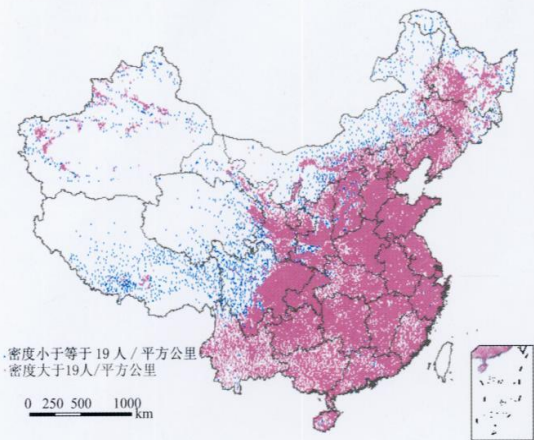
五十年来,我国空间科学技术快速成长、壮大,取得了非常辉煌的成就。在国防科工委的领导下,气象、海洋、资源和环境系列卫星上天,构成了业务运行系统,遥感、导航、地理信息系统也形成了产业链,进入国际市场。拥有的空间信息资源相当丰富,科技精英大批涌现。大专院校现已设置上百个专科,70 多个硕士、博士授予单位。群雄并起,百花齐放,欣欣向荣,一派盛世繁荣景象!

今天,中国科学院在第三期创新工程中,着重关注空间科学技术领域,落实科学发展观,加强对这一领域的领导,统筹协调、全面发展和战略部署,组建“对地观测与数字地球科学中心”,我们非常振奋!我们体会到院领导重振旗鼓,再创辉煌的决心。科学中心的建立,将有助于促进科学数据的集成与信息共享;有助于学科交叉,优势互补;有助于充分发挥人才资源储备和凝聚力量,有助于团结协作。有助于天空地一体化的落实。

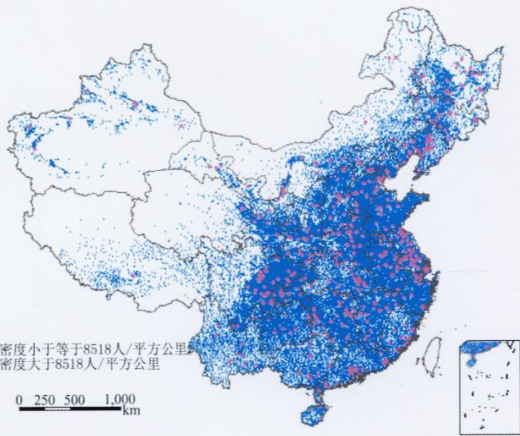
我们衷心祝愿科学中心为我国空间科学技术和地球系统科学的全面协调发展,作出开拓性的贡献;在应对全球变化,缓解资源需求与防治自然灾害的国际事务中,体现中国作为一个负责任的大国,能够引领“数字地球”潮流的发展,能够为世界和平和社会经济可持续发展,作出更多的新贡献!



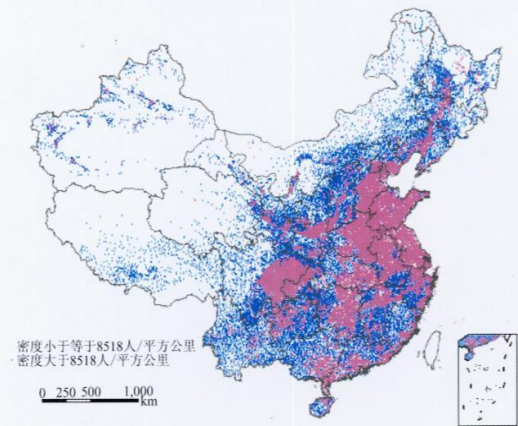
2007 年 11 月 21 日



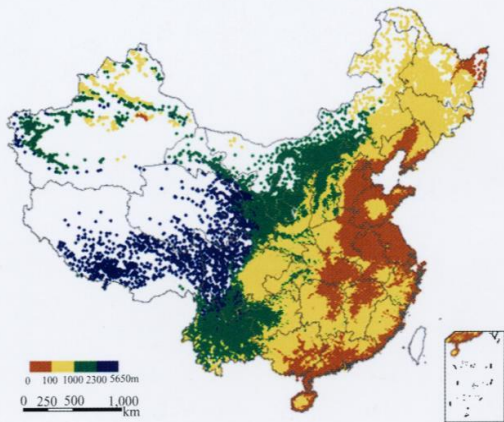
彩图 1 极低密度乡镇分布图
Fig.1 Township low density population
distribution map



彩图 2 高密度乡镇分布图
Fig.2 Township high density populaton
distribution map



彩图 3 符合二八定律的乡镇分布图
Fig.3 Township distribution map coincidence
with 2-8 rule



彩图 4 各高程带乡镇分布图
Fig.4 Map of township distribution based on
altitude