基于多源遥感数据及 DEM 的人口统计数据空间化—以浙江省为例

杨续超[1]，高大伟[2]，丁明军[3]，刘林山[4]

（1.浙江省气象科学研究所，浙江 杭州 310008；2.浙江省气候中心，浙江 杭州 310017；3.江西师范大学鄱阳湖湿地与流域研究教育部重点实验室，江西 南昌 330022；4.中国科学院地理科学与资源研究所，北京 100101）

【摘 要】利用 DMSP/OLS 遥感夜间灯光数据进行人口等社会经济数据的空间化时，往往受到其较低的空间分辨率、像元过饱和以及像元溢出现象的影响。植被指数(如 NDVI)与不透水面呈负相关关系，与夜间灯光数据在反映人类活动、提取建成区方面可以互补，将这两种数据融合可以有效减少夜灯数据像元过饱和等因素引起的误差。通过进一步融合 DEM 数据对基于 DMSP/OLS 夜间灯光数据和 NDVI 构建的人居指数进行了海拔修正，基于修正后的人居指数与统计人口之间很强的线性相关建立人口空间化模型，获得 2010 年浙江省 1km×1km 分辨率下的人口密度空间分布。模拟结果显示，浙江省平均人口密度为 515 人/km2，模拟的平均相对误差为 18.3%，相比海拔订正前的模拟误差减少约 5%，表明利用多源遥感数据融合后的人居指数在省级尺度上模拟人口空间分布的精度较高。

【关键词】人口空间化；DMSP/OLS；NDVI；人居指数；浙江

【中图分类号】P966　　**【文献标识码】**A　　**【文章编号】**1004-8227(2013)06-0729-06

长期以来，人口数据以行政区域为基础统计单元进行数据的收集、存贮和分析，这种统计方法在实际应用时存在着许多问题，如空间分辨率低、与其他数据空间单元尺度不同等。这使得不同部门、行业以及学科之间的数据在空间定位上无法相互引用，不能充分揭示社会经济现象的空间差异性，妨碍了社会经济要素与自然要素的数据共享、融合以及综合分析。此外，人口等社会经济要素精确、详细的空间分布信息是灾害风险管理的 重要组成部分。其中，人口是最重要的受灾体，精细、准确的人口空间分布信息是防灾减灾工作的重要依据，开展有效的灾害预警预报和风险识别等需要获取精准的风险人群的空间分布特征。人口统计数据的空间化是解决上述问题的有效手段，人口密度网格化比人口密度行政单元化更为接近人口的实际分布，是实现人口数据与其他社会经济数据、自然资源、环境数据的有效融合和跨学科研究，提高人口、资源、环境综合管理能力的重要途径之一，对于跨学科的研究具有重要价值。

目前，人口空间化的方法主要包括人口密度-距离衰减模型[1]、格点内插法[2]、土地利用类型影响模型[3~7]、多源数据融合[8~10]、遥感估算法[11，12]以及移动通讯基站[13]等。随着遥感(RS)和地理信息系统(GIS)技术的迅速发展，采用遥感估算法对人口等社会经济数据进行空间化是目前最为重要的手段之一，一些高分辨率的遥感数据被用来进行人口分布的估算，但是由于价格昂贵

收稿日期:2012-06-19;修回日期:2012-09-26
基金项目:中国气象局气候变化专项(CCSF-201336);浙江省科技厅公益技术研究社会发展项目(2011C23051);国家自然科学基金项目(41001023)
作者简介:杨续超(1980~),男,河南省信阳人,高级工程师,博士,主要从事气象灾害风险评估以及气候变化方面的研究.E-mail:yangxuchao@gmail.com

以及数据处理的难度，这种方法很难用于区域尺度的人口密度估算。因此，一些中等分辨率的遥感数据常被用来进行区域人口分布的估算。其中，美国军事气象卫星 Defense Meteorological Satellite Program(DMSP)搭载的 Operational Lines can System(OLS)传感器能够探测到城市、居民地、火光等发出的低强度灯光，在监测人类夜间活动方面有其独特的优越性，是目前实现人口及其他社会经济数据空间化较理想的数据源之一。国内外很多研究基于 DMSP/OLS 夜间灯光数据对不同区域尺度的人口空间分布进行了模拟[14, 15]，如美国能源部橡树岭国家实验室(ORNL)开发的 Land Scan 全球人口动态统计分析数据库中使用夜间灯光、土地利用数据、道路以及海拔等信息获得全球高分辨率(30″)人口分布数据[16, 17]。

近年来，DMSP/OLS 夜间灯光数据也被用来模拟中国的人口空间分布。其中，Lo[18]使用夜灯数据估算了中国省、市、县 3 级行政单元上的平均人口密度；卓莉等[11]利用夜灯数据进行了网格单元上中国人口密度的模拟估算；曹丽琴等[19]以湖北省为例进行了基于夜灯数据的城市人口估算。但是夜灯数据由于其传感器自身的原因也存在一些缺陷，国外学者的研究发现，夜灯数据的应用受到像元过饱和(Saturation)、像元溢出(Overglow)的影响，在实际应用中需要对其进行适当的处理和修正[20~22]。然而，以上针对中国的研究中都没有对夜灯数据的像元溢出和过饱和现象进行订正。最近，Zeng 等[23]将土地利用数据与夜灯数据相结合，对中国人口进行了空间化，两种数据的结合提高了人口模拟的精度。值得注意的是，海拔对我国人口的空间分布有着重要影响[9]，但是以往利用夜灯数据对人口的模拟研究并未考虑海拔的影响。

植被指数(如 NDVI)与不透水面呈负相关关系，也被用来提取建成区[24, 25]。与 DMSP/OLS 夜间灯光在反映人类活动、提取建成区方面可以互补，将两种数据融合可以有效减少夜灯数据像元过饱和等因素引起的误差[26]。Lu 等[27]利用 DMSP/OLS 夜间灯光和 NDVI 提出了一种新的指数--人居指数，这两种遥感数据的融合有效减少了夜灯数据像元饱和效应，大大提高了对人居状况空间分布的提取精度。本研究以浙江省为例，在人居指数的基础上进一步考虑了海拔对人口分布的影响，将多源遥感数据(包括 DM-SP/OLS 夜间灯光、NDVI 和 DEM)进行融合，在 GIS 技术的支持下建立人口空间化模型，进行浙江省人口密度的模拟。以期为提高区域尺度上人口、资源和环境的综合管理能力，以及灾害风险评估中基于人口的承灾体精细化分析提供数据支持。

1 研究区域、数据及预处理

1.1 研究区域

浙江省地处我国东南沿海长江三角洲南翼，东西和南北的直线距离均为 450km 左右，陆域面积 10.18 万 km^2，为全国的 1.06%，是我国面积最小的省份之一。浙江地形复杂，山地和丘陵占 70.4%，平原和盆地占 23.2%，河流和湖泊占 6.4%，耕地面积仅 208.17 万 hm^2，故有“七山一水两分田”之说。全省下辖 11 个地级市，市辖区 32 个，22 个县级市，县 36 个(其中一个为自治县)。2010 年底，全省总人口 4 747.95 万，其中非农人口 1 468.90 万。

1.2 数据及其预处理

(1)DMSP/OLS 遥感夜间灯光数据:来源于美国国家地球物理数据中心(NGDC)提供的 2010 年稳定夜间灯光产品，空间分辨率约为 1km，将其投影由 Geographic(LAt/LON)转换为 Albers 等面积投影，再对其进行标准化处理；(2)2010 年 SPOT-VEGETATION 逐旬 NDVI 数据:来自于比利时佛莱芒技术研究所(Flemish Institute for Technological Research，VITO)，空间分辨率为 1km，将其投影转换为 Albers 等面积投影；(3)DEM 数据:源于美国地质调查局 EROS 数据中心的 GTOP30 数据集，经投影转换及重采样，生成 1km 分辨率的栅格数据；(4)2010 年人口数据:来自于 2011 年《浙江省统计年鉴》中各市县的人口统计数据，根据各市县面积计算平均人口密度；(5)浙江省 1∶5 万地理信息基础数据。

2 研究方法

2.1 确定适宜居住区与不适宜居住区

首先根据 2010 年 SPOT 10d 最大合成的 ND-VI 数据计算年最大 NDVI 值，NDVI 值小于 0.1 的区域基本属于裸岩及水体，NDVI 大于 0.9 的区域基本属于郁闭度较高的森林，这些区域不适宜人类居住[11]，需要先行剔除。浙江水域众多，再利用浙江省 1∶5 万的地理信息基础数据制作水体掩膜，与年最大 NDVI 叠加后剔除水体。

2.2 利用遥感夜间灯光指数和 NDVI 构建人居指数

图 1 给出了 2010 年浙 江省 DMSP/OLS 夜间灯光和年最大 NDVI 值的空间分布，两者有很好的空间对应关系，夜灯值高的区域 NDVI 值低。相比夜灯数据，NDVI 的像元值更为精细，其原因在于夜灯数据的原始分辨率为 2.7km，而且在城市中心区还存在像元过饱和现象，此外，像元溢出则导致钱塘江等城市区的水域也有很高的夜灯值。因此，基于这两种遥感数据的互补性，引入 NDVI 可以有效的缓解夜灯数据的过饱和现象，也能在一定程度上减少由于水面反射等引起的像元溢出现象。参考 Lu 等[27]的研究，利用以下公式构建人居指数(Human Settlement Index，HSI)：

$$HSI=\frac{(1-NDVI_{\max})+OLS_{\mathrm{nor}}}{(1-OLS_{\mathrm{nor}})+NDVI_{\max}+OLS_{\mathrm{nor}}\times NDVI_{\max}}$$

式中:NDVImax 为 SPOT NDVI 在 2010 年的最大值；OLSnor 为 2010 年经过标准化的夜间灯光指数(0～1)。

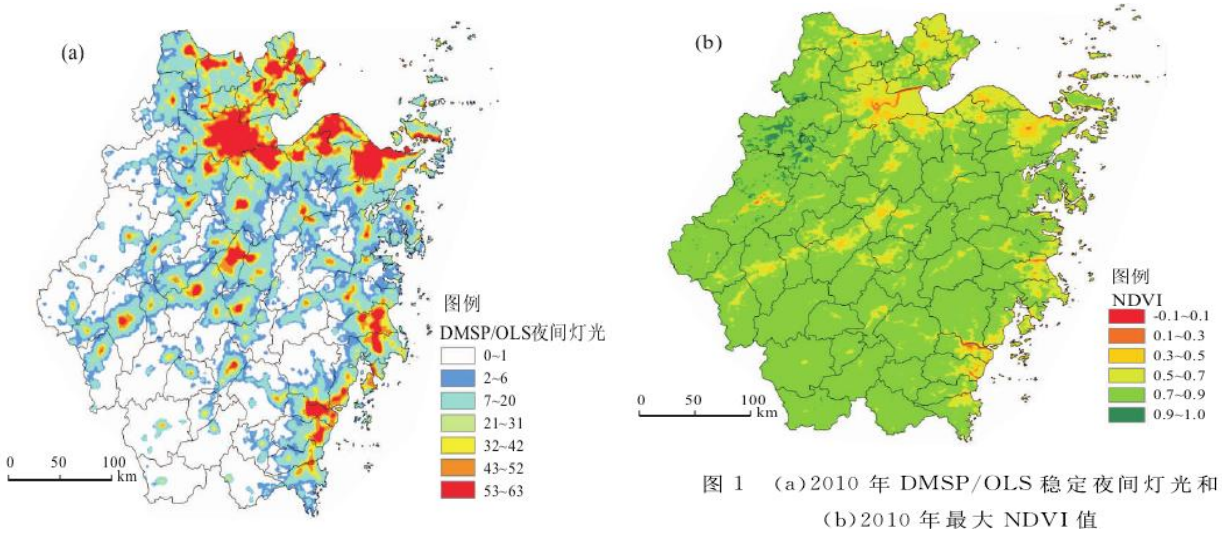

图 1 (a)2010 年 DMSP/OLS 稳定夜间灯光和 (b)2010 年最大 NDVI 值

2.3 人口模拟的误差检验

人口模拟误差检验的计算公式如下：

$$MPE=\frac{\sum_{i=1}^{n}|(RE)_i|}{n}$$

其中：

$$RE=\frac{POP_m-POP_a}{POP_a}\times 100\%$$

式中:MPE(Mean Percentage Error)为平均相对误差；RE(Relative Error)为相对误差；POPm 为市县模拟人口数；POPa 为市县统计人口数；n 为浙江省市县个数。

3 结果分析

3.1 利用人居指数对浙江省市县人口的初步模拟

图 2 给出浙江省各市县人居指数累计值与人口统计数据的关系。结果显示，将 DMSP/OLS 夜间灯光数据和 SPOT NDVI 进行融合后构建的人居指数与人口有着很强的线性相关(R^2=0.89)。利用人居指数与统计人口数据之间的线性关系模拟各市县的总人口数，平均相对误差为 23.48%。对各县市模拟误差的分析发现，浙江西南部高海拔市县(平均海拔都在 600 m 以上)的模拟人口显著偏大，误差都在 50%以上(图 3)，原因在于人居指数并没有考虑海拔对人口分布影响，导致高海拔地区的人口模拟结果显著偏大，因此需要根据海拔对人居指数进行修正。

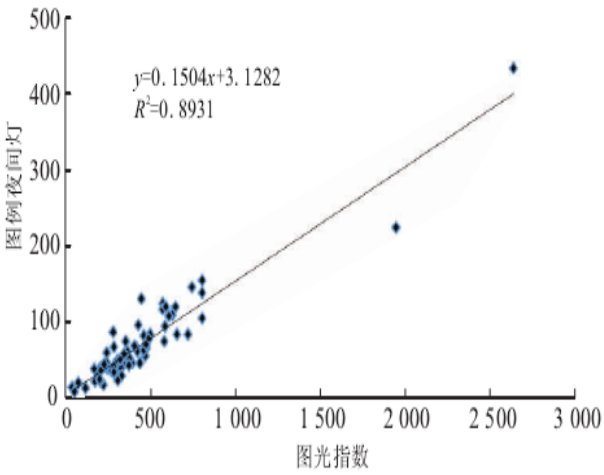

图 2 浙江省各市县人居指数累计值与统计人口的回归分析

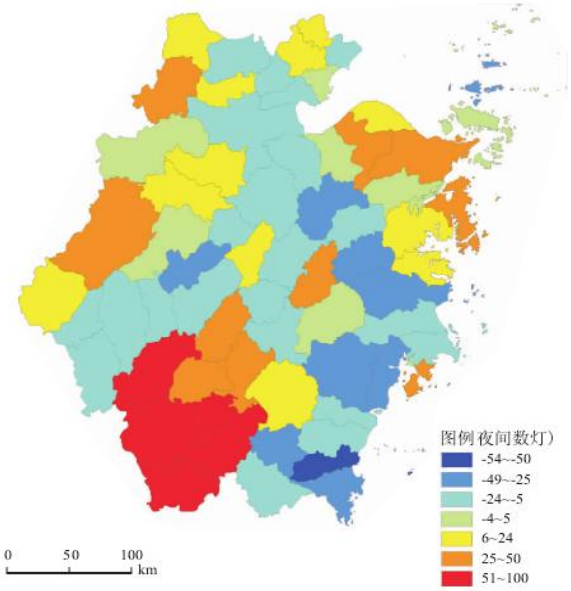

图 3 浙江省各市县人口模拟误差分布

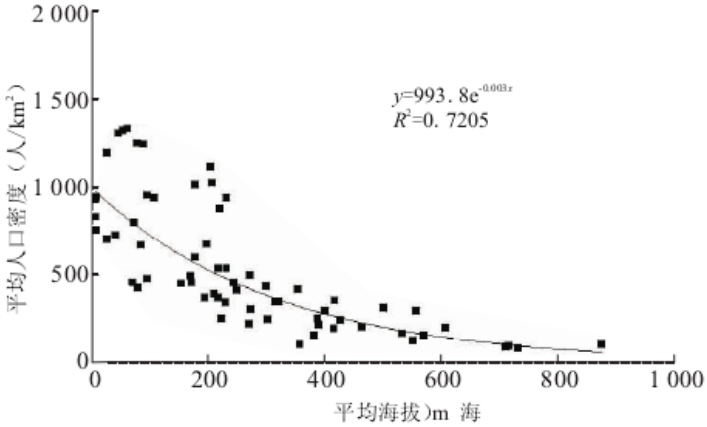

图 4 浙江省各市县平均海拔与平均人口密度的关系

图 4 给出了浙江省各市县平均人口密度与平均海拔高度的关系，采用指数函数回归分析之后显示回归方程的决定系数

R2=0.72，表明海拔高度与人口密度之间具有较强的负相关性，可以作为人口密度模拟的影响因子之一。利用人口密度与海拔的指数关系对海拔高于 600 m 的像元进行人居指数的海拔订正，订正后的人居指数计算公式如下:

$$HSI_{new}=\frac{(1-NDVI_{max})+OLS_{nor}}{(1-OLS_{nor})+NDVI_{max}+OLS_{nor}\times NDVI_{max}}\times$$

$$e^{-0.003DEM}$$

3.2 浙江省人口密度图的生成

修正后的人居指数与人口统计数据之间依然有很强的线性相关(图略)，决定系数有所提高(R^2=0.9)。基于海拔订正后的人居指数与人口的线性关系模型，将人口统计数据展布到空间，生成 1km×1km 分辨率的人口密度图。为了保证在市县级行政单元上模拟总人口与统计总人口相等，利用统计总人口与模拟总人口的比值对人口密度图进行修正。模拟结果显示(图 5)，浙江省人口密度平均值为 515 人/km^2，最大值为 10 675 人/km^2。人口密集区主要集中于以杭州、宁波为中心的浙北平原地区，以及台州、温州为中心的东南沿海地区。海拔 500m 以下地区分布的人口为 4 207 万人，占全省总人口的 88.61 %，而海拔在 1 000m 以上的人口只有 3.8 万人，占全省总人口的 0.08%，说明研究区人口分布受海拔高度的影响明显。

利用海拔修正后的人居指数对浙江省 68 个市县人口模拟的平均相对误差为 18.3%，相对于利用未经海拔订正的模拟结果，平均相对误差减少超过 5%。从各市县人口模拟结果的误差来看，有 64.7%的市县模拟总人口的相对误差绝对值小于 20%，其中 10 个市县的相对误差绝对值小于 5%。误差最小的是磐安县，与统计人口仅相差 0.4%。模拟误差较大的多为海岛县(如洞头)以及沿海市县(如平阳、苍南等)，其中，平阳模拟误差最大，偏小 52.3%。沿海市县模拟人口普遍偏小的原因在于沿海地区人口较为密集，浙江省海岸线漫长而曲折，周围有许多岛屿，对模拟结果进行栅格分区统计时边界细节取舍过程中容易造成像元的舍弃，从而造成总人口的缺失。

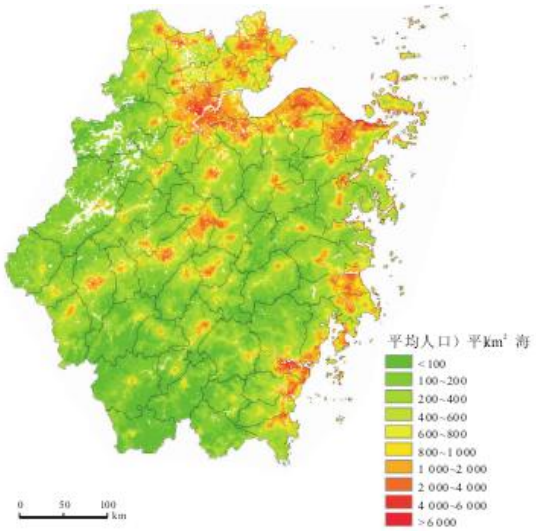

图 5 2010 年浙江省 1 km×1 km 分辨率的人口密度图

4 结论与讨论

以行政区界限为标准统计人口密度的传统方法，无论在理论概念和实际应用上都有很大的局限性，社会的可持续发展以及精细化的灾害风险评估迫切要求更高精度、更加灵活有效的统计和表示人口密度的新方法。将传统的基于行政区划的人口密度计算转变为空间离散的格网计算，使人口密度的空间分布更趋于精确且更具有实际意义。利用 DMSP/OLS 夜间灯光产品模拟人口

密度是人口空间化的重要研究方法之一，但是由于其数据本身像元过饱和和像元溢出等的影响，使其应用受到一定的局限。以往研究多使用夜灯数据来进行城市人口的空间化，如卓莉等[11]利用夜灯进行中国人口空间化时，需要针对灯光区外单独建立模型，曹丽琴等[19]仅利用夜灯数据对湖北省的城镇人口进行了估算。本文在前人研究的基础上，基于遥感数据融合理论，将夜灯数据与 NDVI 融合后构建人居指数，对其进行海拔订正后可以用于区域尺度上总人口数的空间化模拟。以浙江省为例的模拟结果显示，浙江平均人口密度为 515 人/km^2，最大人口密度达到 10675 人/km^2，模拟的平均相对误差为 18.3%，各市县模拟总人口与统计总人口之间线性回归分析的决定系数为 0.9，表明模拟值与统计数据的吻合较好，模拟结果精度较高。

相对于 DMSP/OLS 夜间灯光数据，SPOT NDVI 数据有着更高的空间分辨率，两种数据的融合有效的减少了夜灯像元过饱和，也在一定程度上缓解了由于水面反光等引起的像元溢出问题，而且 NDVI 本身与定居点分布有较好的相关关系，在没有夜灯的区域可以利用这种关系来进行人口模拟，从而不需要针对非夜灯区单独建立模型进行网格化。值得注意的是，以往利用夜灯数据进行人口空间化时没有考虑海拔高度的影响，本文的研究结果表明，对夜灯数据及人居指数进行海拔订正是非常必要的。由于夜间灯光数据本身就包含了城市、居民点、道路等与人口分布密切相关的信息，因此在模拟人口密度时无需再考虑这些因素，相对于利用其他数据模拟人口密度而言，所需数据量较少，易于实现。该方法还有一个重要优点在于所使用的 DMSP/OLS 遥感夜灯数据和 NDVI 数据都在不断更新，且都是免费公开获取的，因此可以实现区域人口密度的动态更新。

需要注意的是，受到时间和经费的限制，本文仅采用传统的误差分析方法对模拟结果进行了检验，误差检验更科学的方法应该包括 1km×1km 范围内实地抽样调查数据，希望在以后的研究中能够将实地抽样调查数据应用于人口空间化模拟及其误差分析中来。此外，由于浙江省海岸线漫长而曲折且海岛众多，导致分区统计时像元缺失明显，使得沿海和海岛市县的模拟人口显著偏小，未来可以尝试将夜间灯光影像与更高分辨率的遥感数据进行融合，开展更加精细的人口空间化模拟。

参考文献:

[1] 冯健.杭州市人口密度空间分布及其演化的模型研究[J].地理研究，2002，21(5):635-646.

[2] 刘纪远，岳天祥，王英安，等.中国人口密度数字模拟[J].地理学报，2003，58(1):17-24.

[3] 江东，杨小唤，王乃斌，等.基于 RS、GIS 的人口空间分布研究[J].地球科学进展，2002，17(5):734-738.

[4] 杨小唤，江东，王乃斌，等.人口数据空间化的处理方法[J].地理学报，2002，57(增刊):70-75.

[5] 田永中，陈述彭，岳天祥，等.基于土地利用的中国人口密度模拟[J].地理学报，2004，59(2):283-292.

[6] 廖顺宝，李泽辉.四川省人口分布与土地利用的关系及人口数据空间化试验[J].长江流域资源与环境，2004，13(6):557-561.

[7] 唐华秀，马劲松，战金艳，等.江苏省人口数据空间离散化研究及其精度分析[J].长江流域资源与环境，2008，17(4):506-510.

[8] 廖顺宝，孙九林.基于 GIS 的青藏高原人口统计数据空间化[J].地理学报，2003，58(1):25-33.

[9] YUE T X，WANG Y A，CHEN S P，et al.Numerical simulation of population distribution in China[J].Population & En-vironment，2003，25(2):141-163.

[10] 王春菊，汤小华.基于 GIS 的福建省人口统计数据空间化[J].地理与地理信息科学，2004，20(4):71-74.

[11] 卓 莉，陈 晋，史培军，等.基于夜间灯光数据的中国人口密度模拟 [J].地理学报，2005，60(2):266-276.

[12] LU D，WENG Q，LI G.Residential population estimation u-sing a remote sensing derived impervious surface approach[J].International Journal of Remote Sensing，2006，27(16):3553-3570.

[13] 毛 夏，徐蓉蓉，李新硕，等.深圳市人口分布的细网格动态特征 [J].地理学报，2010，65(4):443-453.

[14] AMARAL S，MONTEIRO A，CAMARA G，et al.DMSP/OLS night-time light imagery for urban population estimatesin the Brazilian Amazon[J].International Journal of RemoteSensing，2006，27(5):855-870.

[15] CHOWDHURY P K R，MAITHANI S，DADHWAL V K.Estimation of urban population in Indo Gangetic Plains using night time OLS data [J].International Journal of RemoteSensing，2012，33(8):2498-2515.

[16] DOBSON J E，BRIGHT E A，COLEMEN P R，et al.Land Scan:Aglobal population database for estimating populationsat risk[J].Photogrammetric Engineering and Remote Sens-ing，2000，66(7):849-857.

[17] BHADURI B,BRIGHT E,COLEMAN P,et al.LandScan:Locating people is what matters[J].Geoinformatics,2002,5(2):34-37.

[18] LO C.Modeling the population of China using DMSP opera-tional linescan system nighttime data [J].Photogrammetric Engineering and Remote Sensing，2001，67(9):1037-1047.

[19] 曹丽琴，李平湘，张良培.基于 DMSP/OLS 夜间灯光数据的城市人口估算———以湖北省各县市为例[J].遥感信息，2009，1:83-87.

[20] DOLL C.CIESIN thematic guide to night-time light remotesensing and its applications [R].Center for International Earth Science Information Network(CIESIN)，2008.

[21] LETU H，HARA M，YAGI H，et al.Estimating energy consumption from night-time DMPS/OLS imagery after correcting for saturation effects [J].International Journal of Re-mote Sensing，2010，31(16):4443-4458.

[22] TOWNSEND A，BRUCE D. The use of night-time lights satellite imagery as a measure of Australia′s regional electricity consumption and population distribution [J].International Journal of Remote Sensing，2010，31(16):4459-4480.

[23] ZENG C，ZHOU Y，WANG S，et al.Population spatialization in China based on night-time imagery and land use data[J].Inter-national Journal of Remote Sensing，2011，32(24):9599-9620.

[24] BAUER M E，HEINERT N J，DOYLE J K，et al.Impervious surface mapping and change monitoring using landsat remote sensing[C].ASPRS Annual Conference Proceedings 2004，Denver，Colorado.

[25] WENG Q，LU D，SCHUBRING J.Estimation of land surface temperature-vegetation abundance relationship for urban heatisl and studies[J].Remote Sensing of Environment，2004，89(4):467-483.

[26] ROY CHOWDHURY P, MAITHANI S. Monitoring growth of built-up areas in indo gangetic plain using multi-sensor re-mote sensing data [J]. Journal of the Indian Society of Re-mote Sensing, 2010, 38(2):291-300.

[27] LU D, TIAN H, ZHOU G, et al. Regional mapping of human settlements in southeastern China with multisensor remotely sensed data[J]. Remote Sensing of Environment, 2008, 112(9):3668-3679.