

基于 DMSP/OLS 的人口分布网格精细化研究^{*}

郭山山, 龚俊, 尹晶飞

(浙江省地震局, 浙江 杭州 310013)

摘要: 采用浙江省已有的乡镇级行政区域属性数据、乡镇级的建筑物数据和灯光数据, 利用加权计算的方法, 得到了浙江省 250 m × 250 m 的网格人口分布数据。与传统公里网格人口分布数据进行对比, 所得到的网格数据不仅能有效体现出人口分布特征, 而且能适用华东地区的中小地震灾害评估。

关键词: 人口分布; 建筑物分布; 250 m × 250 m 网格; DMSP/OLS 夜间灯光数据

中图分类号: P315-39

文献标识码: A

文章编号: 1000-0666(2016)02-0321-06

0 引言

人口密度是表现人口分布的主要形式, 获取不同尺度上高精度的人口分布信息对于发生重大灾害时的应急救援工作意义重大。近几年浙江省时常发生具有一定社会影响的震群事件, 例如 2006 年 2 月 4 日在温州文成一泰顺交界地区发生了 $M_L 4.6$ 的震群活动, 给当地造成直接经济损失 9 000 多万元, 2014 年同一位置又发生 $M_L 4.8$ 地震震群。所以对突发性的地震事件我们要一直保持警惕, 积极作好应急准备工作。

长期以来地震工作者通过不同方式来防御和减少地震带来的损失, 目前主要有两种手段 (李东平等, 2011a, 2011b), 一是进一步加强地震预测工作、做好震前防御工作, 二是在突发地震事件后提高反应能力, 使受灾人员得到及时救助, 救灾工作高效有序, 迅速恢复和重建灾区 (范开红等, 2014)。一旦地震发生, 各有关部门急需震区灾情信息, 以指导应急和救灾工作 (苏幼坡, 刘瑞兴, 2000)。因此, 各相关部门在地震灾害应急过程中, 要做到行动迅速、准确、高效。如果不能及时了解灾情, 就无法有针对性地提供救援, 就会造成重大的损失。日本阪神地震后对灾情的错误评估 (危福泉等, 2005), 错过了抢救的最佳时机, 导致了更多的人员伤亡和财产损失。

地震灾害损失盲估, 是在震后尽量短的时间内, 根据平时建立的模型方法和有限的现场信息, 给出损失的快速评估结果, 以便政府和有关部门根据灾害规模做出应急对策, 部署救灾行动, 争取外援。其主要要求是速度快, 但不苛求十分精

确 (李智, 王晓青, 2010)。现今, 如果在经济发达、人口稠密的浙江省发生一次破坏性地震, 通常需要一周以上的时间才能得到评估结果, 其产生的灾害损失就将大大增加, 还可能会导致政府职能的发挥失常, 甚至政府的信任危机等各种弊端也会接踵而至。相反如果在地震发生后, 能在最短的时间内准确估算出地震灾害损失的分布、大小, 就可以为政府提供科学的应急决策依据, 积极有效地组织实施抢险救灾行动, 最大程度上减少经济损失并挽救成百上千人的生命 (李平, 罗云, 1994), 同时有效提升政府的公众形象。因此, 地震灾害损失快速评估是必要的。

为了提高地震灾害损失快速评估的水平, 笔者将对浙江省的人口网格数据进行精细化、实时化研究。

1 浙江省数据源及数据预处理

基础地理信息数据是进行空间分析、模型模拟以及专题制图的基础, 笔者利用现有的浙江省行政区划数据、人口分布数据、建筑物数据、夜间灯光数据等来模拟浙江省人口的空间分布。

1.1 行政区划数据

笔者使用的行政区划数据来源于 2013 年浙江省地震局地震应急基础数据库。按照该数据库结构, 行政区划数据分为地市代码库、区县代码库和乡镇代码库。这 3 类数据中的属性库部分, 包含浙江省 11 个地 (市)、90 个区 (县、市)、1 324 个乡、镇、街道行政区划, 数据表中包含有行政区划代码、行政区名称等。空间库为 shp 格式, 投影方式为经纬度。空间库与代码库一一对应, 并

^{*} 收稿日期: 2015-01-13.

基金项目: 浙江省地震局科技计划项目 (2016ZJJ01) 与地震行业科研专项 (201508026) 联合资助。

完成关联。

1.2 人口分布数据

浙江省人口分布数据来源于浙江省地震局地震应急基础数据库中的乡镇人口统计表(街道办事处),其中包含浙江省 1 324 个乡、镇、街道的人口统计数据, ID 与浙江省行政区划数据中的乡镇代码库的 ID 对应,并完成关联,数据保存为 mdb 格式。人口分布数据包括行政区名称、总人口、家庭户户数、大于 65 岁人口、0~14 岁年龄人口和居住本地、户口在本地人口数。

2013 年底,浙江省常住人口约为 5 494 万人,全国排名第 10,人口密度为 540 人/km²,居全国第 8 位。从分区域来看,浙江省人口比重最大的分别是温州市、杭州市和宁波市,3 个地(市)的人口总数占全省人口总数近一半,位于浙江省西南方的丽水市由于地处山区,人口分布相对稀疏。

图 1 为浙江省地震局应急数据库中 2013 年乡镇人口分布图。从图中可以看出,浙江省东、西部的人口数差距很大,东部一个乡镇的人口数大于 25 000 人,而部分西南山区乡镇的人口数却小于 15 000 人。

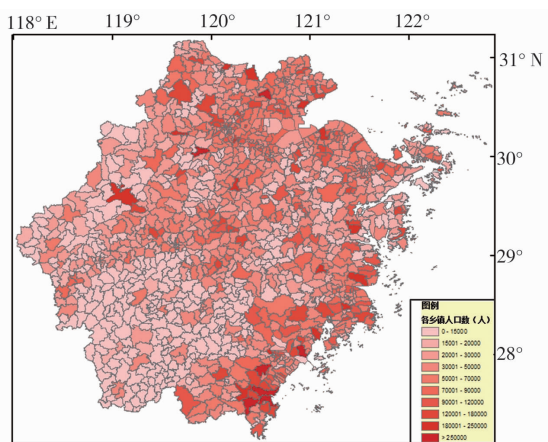


图 1 2013 年浙江省乡镇人口分布图

Fig. 1 Distribution of population for town in Zhejiang Province in 2013

1.3 建筑物分布情况

地震中建筑物的倒塌造成大量的人员伤亡,建筑物分布能够实时反映人口聚居以及分布情况。根据《2013 年浙江省国民经济和社会发展统计公报》(浙江省统计局,国家统计局浙江调查总队,2014),浙江省城镇居民人均住房建筑面积 38.82 m²,农村居民人均居住面积 60.82 m²。浙江省测绘与地理信息局提供 2013 年浙江省 1:1 万房屋外廓分布图,包含浙江省全省的建筑物,反映出浙江省建筑物的分布情况。

1.4 建筑物分布与人口分布对比

图 2 为浙江省各地(市)总人口与总建筑面积归一化分布图。从图中可以看出两条折线具有明显的相似性,因此在一定程度上,建筑物面积与当地人口总数是成正比的。

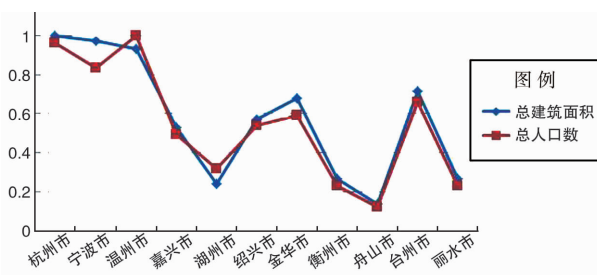


图 2 浙江省各地市总人口与总建筑面积归一化分布图

Fig. 2 Normalization Distribution of total population and total construction area for cities in Zhejiang Province

1.5 夜间灯光数据

美国 DMSP (Defense Meteorological Satellite Program) 卫星的 OLS (Operational Linescan System) 搭载在 DMSP 卫星上,以 14 轨/d 的速度覆盖地球一周,能获取全球范围的黎明、白天、黄昏和夜晚 4 个时段的影像数据。OLS 传感器有可见光和热红外 2 个通道,其中可见光波段范围为 0.4~1 m,光谱分辨率为 6 bit,灰度值范围 0~63。OLS 具有很强的光电放大能力,能够探测到地表的城市灯光、辉光、极光、闪电和火光等,OLS 传感器每天都能获取覆盖全球的黑夜和白天的图像,在夜间工作时对可见近红外光高度敏感,其量级是普通传感器的 4 倍,高增益性能使 OLS 传感器探测城镇灯光等发出的电磁波,因此在国外已被应用于城市监测,火灾监测等研究(何洋等,2014)。

目前,DMSP/OLS 夜间灯光影像主要用于城镇扩展研究、社会经济因子估算以及其他环境、灾害、渔业、能源等领域(米晓楠等,2013)。自 1992 年,OLS 传感器获取的数据在美国国家地理数据中心由美国空军和 NOAA 处理成为数字格式的文档,并通过相应程序研发生成夜间灯光影像数据,其数据每年 1 幅大小约为 2.7 GB,分辨率为 0.008 33°(约 0.93 km)。

由于 DMSP/OLS 数据在低灰度值(<3)的数据受噪声干扰比较严重,因此笔者对 3~63 级灰度值进行阈值处理。笔者下载 1992~2010 年的数据进行分析,图 3 为杭州市部分地区夜间灯光影像及

建筑物叠加分布情况。

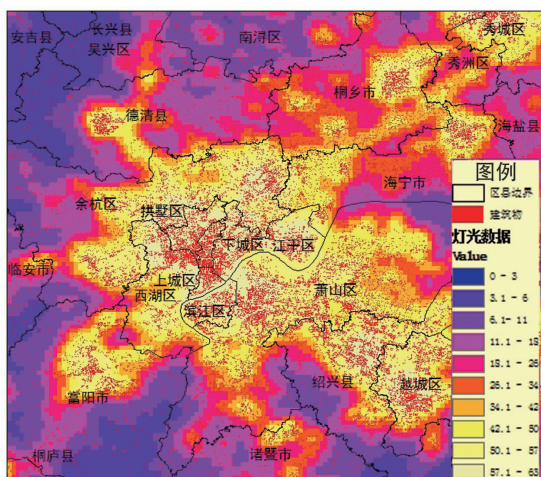


图3 杭州市部分地区灯光影像及建筑物分布情况

Fig.3 Distribution of night's light images and buildings in part region of Hangzhou City

2 浙江省细网格人口分布研究

Sutton (1997)、卓莉等 (2005) 研究发现, 人口密度与夜间灯光数据及建筑物具有较大的关系。在一定程度上, 夜间灯光数据越高代表本地人口密度越大, 建筑物越多代表本地人口密度越大, 反之亦然, 但单一的元素不能够很好地反映较精细化的人口密度。

为了使各地人口更加接近实际情况, 笔者选取 $250\text{ m} \times 250\text{ m}$ 网格的人口密度情况进行研究, 具体实施步骤如下:

(1) 建立数据库。将调查搜集到的浙江省各乡镇人口数据、全省建筑物分布情况数据、全省夜间灯光数据等, 建立属性数据库。

(2) 人口数据的处理。根据《区域级抗震救灾指挥部地震应急基础数据库格式规范 (修订稿)》^① 的规定, 数据的编码系统为 14 位, 编码的前 9 位与乡镇代码库的 ID 对应。在处理浙江省 2013 年乡镇街道人口数据前还需要检查空间和人口数据的年份、数量、编码是否一致, 检查无问题后, 使用 ArcGIS Toolbox 中的链接功能, 使人口和空间数据进行匹配, 匹配成功后就可以根据含有人口属性表的空间数据计算出浙江省的人口密度, 进而统计分析人口分布影响因子。

(3) 建立基础剖分网。网格是进行人口数据

网格化研究的目标, 是各类参数载体以及各种空间分析的主要对象。本研究基于 ArcGIS 与 Matlab, 构建了 $250\text{ m} \times 250\text{ m}$ 的规则网格以均匀覆盖浙江省各个乡镇。

(4) 单元网格点数据处理方法。对单元网格进行空间分析, 考虑单元网格内所有影响因子的累积影响, 将单个单元格的数据拓扑到所在单元网格中心点上, 通过加权叠加等空间分析方法获取各个单元格对应的人口密度数据。应用空间分析模块对单元网格中心点上人口密度数据进行分析, 获取该单元格对应的人口密度。

①在每个 $250\text{ m} \times 250\text{ m}$ 的网格中, 将该单元网格中心点对应的夜间灯光数据表示为 L_i 、建筑物面积数据表示为 B_i 、乡镇的人口数据 P_i 为该网格所在乡镇的人口数; ②一般情况下可以认为灯光数据 L_i 越大代表本处高层建筑物越多, 假设单个 $250\text{ m} \times 250\text{ m}$ 网格的 W_i 等于该网格中建筑物面积 B_i 与夜间灯光数据 L_i 的乘积 (式 (1)); 分别将各乡镇 W_i 相加获取 $\sum W_i$, 而每个网格的实际权重 W_i' 等于 W_i 除以 $\sum W_i$ (式 (2)); ③任一 $250\text{ m} \times 250\text{ m}$ 的网格计算得到的人口数 P_i' 等于该乡镇的实际人口数 P_i 乘以该网格的实际权重 W_i' (式 (3))。

$$W_i = B_i * L_i, \quad (1)$$

$$W_i' = W_i / \sum W_i, \quad (2)$$

$$P_i' = P_i * W_i'. \quad (3)$$

依据上述评价方法, 获取浙江省 $250\text{ m} \times 250\text{ m}$ 网格人口密度分布, 见图 4。

2013 年浙江省的统计人口总数 5 494 万人, 由上述方法经过网格化、数据分析加权等过程得到的浙江省总人口数为 5 472 万人, 分布在约 $101\,800\text{ km}^2$ 的区域内, 约有 $67\,260\text{ km}^2$ 区域无人居住。

3 成果对比

中国科学院地理科学与资源研究所利用第 6 次普查数据以及 GIS 空间分析技术生成了浙江省 $1\text{ km} \times 1\text{ km}$ 网格人口数据 (陈振拓, 2012), 如图 5 所示。

在实际情况中, 公里网格人口密度存在一定程度的不足, 没有考虑乡村的存在及对人口分布的影响。其主要原因是无法科学地确定如村级居民地的平均人口密度及其影响半径。在插值时如果能够合理地考虑上述因素对人口密度的影响, 最终结果可能会更接近真实。

①中国地震局. 2006. 区域级抗震救灾指挥部地震应急基础数据库格式规范 (修订稿).

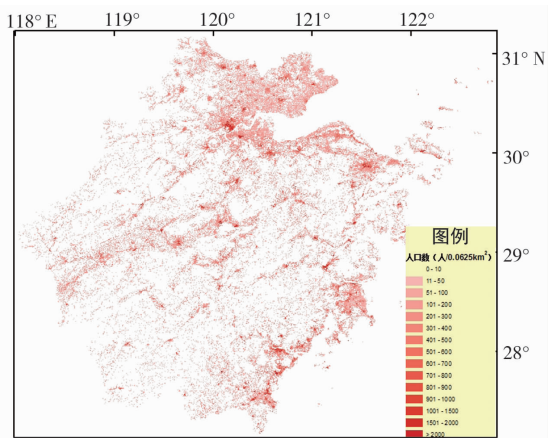


图4 2013年浙江省250 m×250 m 网格人口分布情况
Fig. 4 Distribution of population for grid of 250 m×250 m in Zhejiang Province in 2013

对比中国科学院地理科学与资源研究所制作的2010年1 km×1 km 网格人口分布以及本文的250 m×250 m 网格人口密度分布。对图4和图5中杭州市部分城区及人口分布情况作放大处理,如图6所示。

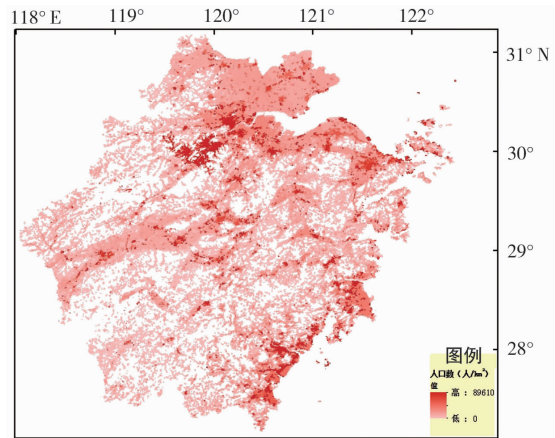


图5 2010年浙江省1 km×1 km 网格人口分布
(付晶莹等, 2014)

Fig. 5 Distribution of population for grid of 1 km×1 km in Zhejiang Province in 2010 (based on Fu et al., 2014)

图6a为杭州市部分地区250 m×250 m 网格人口分布情况,图6b为杭州市部分地区1 km×1 km 网格人口分布情况。

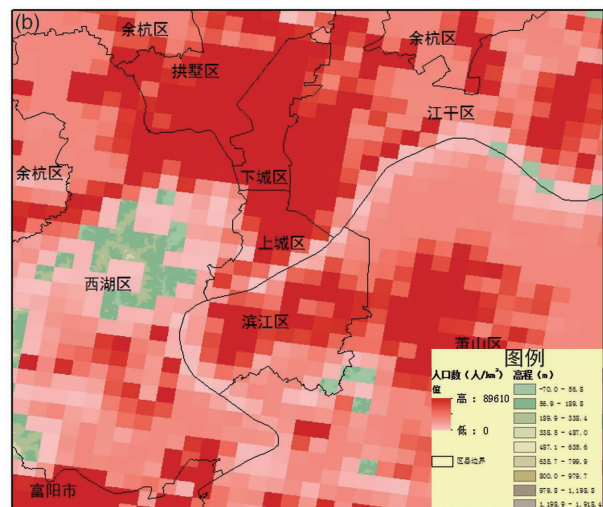
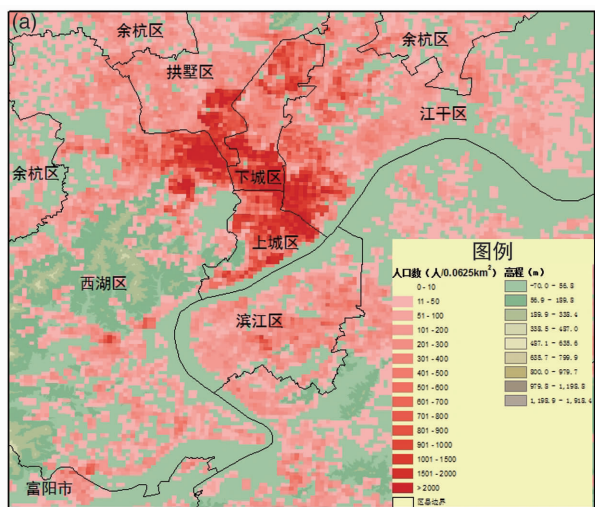


图6 杭州市部分地区250 m×250 m (2013) (a) 和1 km×1 km (2010) (b) 网格人口分布情况

Fig. 6 Distribution of population for grid of 250 m×250 m (2013) (a) and 1 km×1 km (2010) (b) in part region of Hangzhou City

从图3可以看出,杭州市上城区、下城区建筑物较为密集,与此同时灯光数据也相对来说较大。这反映出,杭州人口分布较密集的地区集中在传统意义上的老城区:上城区、下城区以及萧山区,而钱塘江周边由于尚未完全建设成熟以及西湖景区限制导致人口分布较为稀疏,这种分布与图6一致。

但是,相较于250 m×250 m 网格人口分布,1 km×1 km 网格人口分布图在精度上存在一定程度上的不足。以图6中部分地区为例进行分析:(1)钱塘江及其沿岸区域,钱塘江作为浙江省最大的河流流经杭州市,其所在区域必没有人口居住,这种情况在图6a中反映很准确,图6b中则反映不是很明显;(2)西湖景区,西湖作为杭州著

名的旅游胜地，由西湖及其周边群山组成，在图 6a 中我们几乎可以准确地判断出西湖及其周边群山的位置，但是在图 6b 中却较难反映；（3）富阳市部分区域，为检验富阳市人口数据网格化结果的精度，本文对富阳市人口数据与网格化的人口数据结果进行对比验证。实地调查依据公里网格所确定的 $250\text{ m} \times 250\text{ m}$ 范围进行，对结果精度进行检验。结果显示，富阳市人口较之萧山及上城区、下城区少，而图 6b 显示富阳市人口与萧山区人口相差无几。

为验证本文提出的 $250\text{ m} \times 250\text{ m}$ 网格人口准确性，特在全省范围内随机选取 8 个乡镇街道进行对比实验。将乡镇街道实际人口数与网格化后人口数进行对比，实验结果如表 1 所示。

表 1 部分乡镇实际人口数与网格化后人口数对比

Tab. 1 Comparison between real statistical population data and population data after grid in some towns

乡镇街道	实际人口 /人	网格化后人口 /人	网格化后人口/ 实际人口 (%)
杭州清波街道	31 349	31 056	99.07
宁波文教街道	30 516	31 248	102.4
苍南县石坪乡	10 803	10 321	95.54
平湖市乍浦镇	54 809	55 286	100.87
上虞市岭南乡	12 265	12 895	105.14
柯城区华墅乡	19 210	19 983	104.02
青田县舒桥乡	14 319	13 996	97.74
龙泉市锦溪镇	9 435	9 621	101.97

由上述实验结果可知，本文 $250\text{ m} \times 250\text{ m}$ 网格人口分布在一定程度上较为准确。网格化后的数据与实际人口对比存在一定程度上的差异，可能是由于网格化的数据在乡镇边界部分隶属于两个乡镇。

综上所述，本文研究的 $250\text{ m} \times 250\text{ m}$ 网格人口分布较 $1\text{ km} \times 1\text{ km}$ 网格人口分布更加精确、细致，且本文的研究成果与实际情况较为一致。

4 结论

笔者将原本以乡镇级行政区为统计单位的人口以及建筑物分布数据转化为规则网格点上的人口密度及单位面积人口。然后，依据建筑物及夜间灯光数据分析该网格化获取数据的正确性。通

过研究结果的统计分析，得到以下结论：

（1）本文对浙江省人口分布情况做了具体分析，从海拔看，浙江西南山区人口分布较为稀少，水系穿过的区域也均无人分布。浙北平原地区以及东南沿海平原地区人口密集；从城市来看，杭州、宁波以及温州部分地区人口分布相对密集，而丽水等地相对人口较为稀疏。

（2）本文得到的浙江省 $250\text{ m} \times 250\text{ m}$ 网格人口密度分布较为合理地反映了浙江省单元行政区内部的人口分布情况。可以看出：在浙江省各个行政单元内部，城镇街区建筑物分布密集、乡村建筑物呈现片状零散分布，而偏远农村的建筑物分布稀疏，这 3 类建筑物的分布与本文所生成的高值、中间值和低值人口密度区有较好的空间耦合关系。具体来说，城镇地区的人口空间密度形态为片状，并依地形分布，城镇的边缘地区沿交通干道线状辐射；乡镇和农村地区人口分布以沿水系、道路分布为特征。同时本文将部分乡镇实际人口与网络化人口进行对比分析，生成的人口密度数据质量具有较高的精度，达到了预期效果。

（3）本文获取的 $250\text{ m} \times 250\text{ m}$ 网格数据较之前的 $1\text{ km} \times 1\text{ km}$ 网格人口及人口数据精确且细致。相较于传统公里网格人口数据不够严密的特点，本文获取的人口数据更加真实，且更加能反映实际人口分布情况。

参考文献：

- 陈振拓. 2012. 服务于地震应急的人口数据格网化方法研究——以云南省为例[D]. 北京：中国地震局地质研究所.
- 范开红, 林洋, 申源. 2014. 地震应急指挥技术系统产出信息面向对象分类与应用[J]. 地震研究, 37(2): 317–322.
- 付晶莹, 江东, 黄耀欢. 2014. 中国公里网格人口分布数据集[EB/OL]. (2014–06) [2015–01–13]. <http://www.geodoi.ac.cn/webcn/doi.aspx?Id=131>
- 何洋, 程辉, 唐亮. 2014. 基于 DMSP/OLS 数据的我国省级经济发展水平研究[J]. 地理空间信息, 12(2): 79–82.
- 李东平, 龚俊, 陈军徽等. 2011a. 浙江省震害盲估属地化研究[J]. 地震研究, 34(2): 222–226.
- 李东平. 2011b. GIS 的发展趋势与数字地震应急救援的实现技术[J]. 计算机技术与发展, 21(1): 214–217.
- 李平, 罗云. 1994. 地震灾害经济损失的宏观估算方法[J]. 中国地质灾害与防治学报, 5(4): 86–89.
- 李智, 王晓青. 2010. 地震震害微观与宏观方法快速盲估综述[J]. 地震, 30(2): 134–142.
- 米晓楠, 白林燕, 谭雪航等. 2013. 基于 DMSP/OLS 数据的城市中心城区提取新方法[J]. 地理信息科学学报, 15(2): 255–261.
- 苏幼坡, 刘瑞兴. 2000. 城市地震灾害紧急救助的时序特性分析

- [J]. 灾害学, 15(2): 33–37.
- 危福泉, 刘高焕, 姚新等. 2005. 地震灾害预测和应急模拟系统的设计与应用——以永安市城市应急系统为例[J]. 地理研究, 24(5): 749–756.
- 卓莉, 陈晋, 史培军等. 2005. 基于夜间灯光数据的中国人口密度模拟[J]. 地理学报, 60(2): 266–277.
- 浙江省统计局, 国家统计局浙江调查总队. 2014. 2013 年浙江省国民经济和社会发展统计公报[J]. 统计科学与实践, 3(1): 4–10.
- Sutton P. 1997. Modeling population density with nighttime satellite imagery and GIS[J]. Computers, Environment and Urban System, 21(3–4): 227–244.

Study on Grid Refinement for Population Distribution Based on DMSP/OLS

GUO Shanshan, Gong Jun, YIN Jingfei

(Earthquake Administration of Zhejiang Province, Hangzhou 310013, Zhejiang, China)

Abstract

We applied the attribute data of township-level administrative areas, the data of township-level buildings and DMSP/OLS light in Zhejiang Province to obtain the grid data of $250\text{ m} \times 250\text{ m}$ by the weighted calculation method. Contrast with the $1\text{ km} \times 1\text{ km}$ traditional grid data of population distribution which was made by Chinese Academy of Sciences, it showed that the grid data obtained in this paper can not only reflect the characteristics of population distribution effectively in Zhejiang Province but also be applied to the disaster assessment of moderate and small earthquake in East China.

Key words: distribution of population; distribution of buildings; grid of $250\text{ m} \times 250\text{ m}$; data of DMSP/OLS for night's light