

引用格式: Jiang Luyuan, Xiao Pengfeng, Feng Xuezhi, *et al.* Assessment of Large-scale Land Cover Datasets in Typical Areas of China based on Sub-fractional Error Matrix[J]. Remote Sensing Technology and Application, 2015, 30(2): 353-363. [蒋璐媛, 肖鹏峰, 冯学智, 等. 基于亚分数混淆矩阵的中国典型区大尺度土地覆盖数据集评价[J]. 遥感技术与应用, 2015, 30(2): 353-363.]
doi: 10. 11873/j. issn. 1004-0323. 2015. 2. 0353

基于亚分数混淆矩阵的中国典型区 大尺度土地覆盖数据集评价

蒋璐媛, 肖鹏峰, 冯学智, 李云, 朱榴骏

(南京大学地理与海洋科学学院地理信息科学系, 江苏 南京 210023)

摘要:在中国东北、华北、华中、华南、西北、青藏、内蒙古7个自然地区分别选择典型区A、B、C、D、E、F、G,以Landsat TM/ETM+影像分类结果为参考数据,采用亚分数混淆矩阵对5种大尺度土地覆盖数据集的精度进行定量评价,为数据集的使用提供科学依据。亚分数混淆矩阵可避免参考数据与待评价数据尺度转换时引入的误差,能反映不同优势类比重情况下数据集的总体精度和分类方法误差。结果表明:GLC2000在全部典型区的总体精度最高,为65.64%;UMD总体精度最低,为43.06%。GLC2000在主要土地覆盖类型为林地和耕地以及草地区域具有较高的分类精度;UMD在各区域的分类精度均最低或较低。5种土地覆盖数据集对于城镇、其他的分类精度在各典型区均较低;对于草地和水体的分类精度则是在西北干旱区和青藏高原区的典型区较高。

关键词:亚分数混淆矩阵;大尺度;土地覆盖数据集;精度评价

中图分类号: TP 79 文献标志码: A 文章编号: 1004-0323(2015)02-0353-11

1 引言

大尺度土地覆盖数据集是全球陆地表层过程研究及全球变化研究的重要基础,是对地表覆盖特征的模拟和概括,广泛应用在生态系统评估和环境模拟的研究中^[1-4]。目前,已有多种基于遥感影像生产的全球尺度土地覆盖数据集,常用的有马里兰大学生产的全球土地覆盖数据集(University of Maryland Global Landcover Datasets, UMD)^[5];美国地质调查局为国际地圈—生物圈计划建立的全球土地覆盖数据集(International Geosphere-Biosphere Programme Data and Information System Cover, IGBP DISCover)^[6];基于MODIS生产的全球土地覆盖数据集(MODIS Land Cover Type Product collection 4, MOD12Q1)^[7];欧盟联合研究中心生产的全球土地覆盖数据集(Global

Land Cover 2000, GLC2000)^[8];以及欧洲空间局生产的全球土地覆盖数据集(Global Land Cover Product-GlobCover)。UMD、IGBP DISCover、MOD12Q1 3种土地覆盖数据集使用的是IGBP分类体系,GLC2000、GlobCover两种土地覆盖数据集使用的是联合国粮农组织制定的Land Cover Classification System(LCCS)分类体系。由于生产各土地覆盖数据集使用的数据来源、分类系统以及分类方法各不相同,导致分类结果差异较大,用户选用合适的数据集较为困难^[9]。因此开展数据集精度评价研究,对提高数据集应用水平具有重要意义。

目前国内外学者针对土地覆盖产品的精度评价已开展的研究工作主要包括:①全球尺度的多种数据集比较分析(如面积一致性、空间一致性)和精度评价^[10-15];②区域尺度的数据集精度评价^[16-19]。在

收稿日期:2014-05-06;修订日期:2015-03-16

基金项目:国家973计划项目(2011CB952001)资助。

作者简介:蒋璐媛(1990—),女,陕西安人,硕士研究生,主要从事数字图像处理和资源环境遥感研究。E-mail:yuanyuanc042@126.com。

通讯作者:肖鹏峰(1979—),男,湖南宁乡人,副教授,主要从事遥感数字图像处理和资源环境遥感研究。Email:xiaopf@nju.edu.cn。

中国区域,则多根据 2000 年中国 1:10 万土地利用数据评价各数据集的总体精度^[20-22],为各数据集在中国的应用提供了支持及有效依据。

利用中分辨率参考数据评价低分辨率土地覆盖数据集的研究中,通常是将参考数据与土地覆盖数据集重采样至相同的分辨率,这不可避免地引入了新的误差,且无法反映数据集在亚像元尺度的精度^[23];同时研究均采用单一时间点的参考数据评价不同时间段的土地覆盖数据集,忽略了时间差异引起的土地覆盖变化,从而引入了新的误差。基于亚分数混淆矩阵的精度评价无需考虑中分辨率参考数据与低分辨率待评价数据之间的尺度差异,避免重采样过程所引入的误差。它通过直接比较待评价数据与参考数据的差异,评价亚像元尺度的数据集总体精度,反映不同优势类比重情况下数据集的总体精度,在精度评价研究中具有较大优势^[16]。Latifovic 等^[16]以加拿大的土地覆盖数据作为参考,首次采用亚分数混淆矩阵的精度评价方法,对 UMD、IGBP DISCover、GLC2000、MODIS 4 种土地覆盖数据集在加拿大的总体精度进行评价,并分别分析了参考数据空间分辨率、研究区土地覆盖总体均质性、参考数据类别分辨率与精度评价的关系。Ran 等^[24]在对 UMD、IGBP DISCover、GLC2000、MODIS 4 种土地覆盖数据集分类误差的分析中基于 1:10 万土地利用数据,在中国选取 21 个 200 km × 200 km 的样区,分别计算 21 个样区与 4 种土地覆盖数据集的亚分数混淆矩阵,基于 21 个样区的亚分数混淆矩阵对 4 种土地覆盖数据集进行误差分析。结果显示 GLC2000 与 MODIS 2001 土地覆盖数据集具有更高的整体分类精度。IGBP 数据集的整体分类精度次之,UMD 分类精度整体最低。误差产生的地方主要分布在云贵高原、重庆和东南丘陵区;分类误差在不同的土地覆盖类型之间明显不同,结果表明在 4 种数据集中湿地类型和建筑用地具有最大的标定误差。

本文在基于 Latifovic 和 Ran 的土地覆盖数据集精度评价工作的基础上,根据我国自然地理条件复杂的特点,采用自然地理分区的策略分别评价不同自然地区各土地覆盖数据集的精度,将土地覆盖数据集的精度评价工作首次与自然区划方案中的不同自然地区相结合,力图更全面地反映数据集精度。针对不同自然地区中的典型区,分别以 5 种全球土地覆盖数据集同时期的 Landsat TM/ETM+ 影像分类结果作为参考数据,减小由时间差异引起的土

地覆盖变化而引入的误差。采用基于亚分数混淆矩阵的方法分别对各土地覆盖数据集的精度、各典型区内的精度以及各土地覆盖类型的精度进行评价,并对空间匹配误差对精度评价的影响进行研究。

2 研究区概况

土地覆盖主要受气候、土壤和地貌等自然环境条件的影响,相同自然地理条件下土地覆盖的组成、结构相似性较大,同一综合自然地理区内自然地理条件也较为相似^[25]。因此,本文从不同的自然地区中选择典型区进行精度评价,以反映各数据集在不同自然地理条件下的精度情况。

已有学者根据不同标准提出多种中国自然地理分区方案,主要有:①罗开富^[26]以景观作为划分对象,以植物与土壤作为景观的标志,提出的“中国自然地理区划草案”;②黄秉维等^[27]根据地表热量的地域差异提出的“中国综合自然区划图(草案)”;③任美镔^[28]依据自然差异的主要矛盾,以及利用和改造自然的不同方向提出的“中国综合自然区划”;④赵松乔等^[29]根据温度、水分条件的组合及其在土壤、植被等方面的反映提出的“综合自然分区”;⑤赵济^[30]根据自然区划原则、中国自然地理特点及地域分异规律提出的“中国自然区划”。

本文采用赵济提出的“中国自然区划”方案分析不同自然地区条件下各土地覆盖数据集的精度。依据这一区划方案,在 7 个二级自然地区中分别选取 120 km × 120 km 的典型区作为研究区,对常用的 5 种大尺度土地覆盖数据集进行评价。在进行典型区的选取时,尽量使其自然地理特征和土地覆盖特点与所在的二级自然地区相一致。图 1 为 7 个典型区在中国的位置示意图,表 1 为选取的 7 个典型区的自然地理特征。

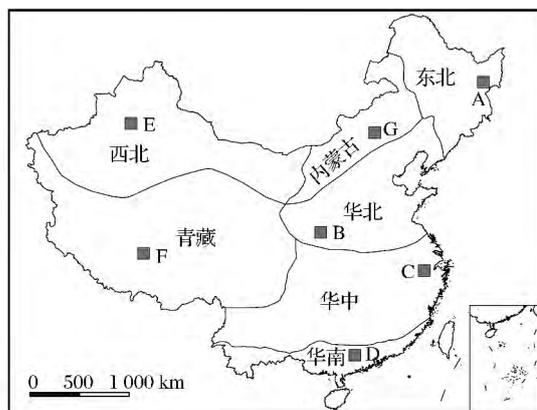


图 1 7 个典型区位置示意图(修改自赵济^[30],1995)

Fig. 1 Position schematic diagram of the seven typical areas (Modified from Zhao Ji^[30], 1995)

表 1 7 个典型区的自然地理特征

Table 1 Physical geography characteristics of the seven typical areas

典型区	中心 经纬度	典型区自然地理特征及 主要土地覆盖类型	所属 自然地区	所属自然地区 主要土地覆盖类型
A	45°56'N 130°52'E	位于三江平原南部,属中温带湿润气候区,主要土地覆盖为林地、耕地	东北 地区	丰富的温带冷湿性森林资源,松嫩平原和三江平原的耕地面积占全区的 80% 以上
B	34°02'N 108°50'E	位于渭河流域中部关中盆地,属暖温带半湿润的季风气候区,主要土地覆盖为耕地	华北 地区	耕地面积占绝对优势,是东部季风区内耕地比例最大而林地比例最小的区域
C	30°18'N 119°56'E	位于浙江省北部,属亚热带季风气候区,主要土地覆盖为林地、耕地	华中 地区	广泛分布湿热亚热带常绿林,北部为肥沃的冲积平原,利于农作物的生长
D	23°30'N 111°40'E	位于广西省东部,属亚热带湿润季风气候区,主要土地覆盖为林地	华南 地区	广泛分布着热带季雨林、热带季风常绿阔叶林、热带灌丛等
E	44°40'N 86°00'E	位于新疆准格尔盆地南缘,受水汽来源、地形和纬度等因素影响,流域垂直分带特征明显,属温带干旱气候区,主要土地覆盖为草地、永久积雪	西北 地区	广泛分布着荒漠、荒漠草原及干草原自然景观,同时区内冰川地貌广阔分布
F	31°49'N 89°33'E	位于西藏北部,青藏高原腹地,区内内陆湖泊星罗棋布,河流纵横其间,属亚寒带气候区,主要土地覆盖为草地、水体	青藏 地区	为中国著名的湖区,冰川、冻土广布,高原寒漠、草甸、牧草地占全区面积过半
G	43°11'N 115°57'E	位于内蒙古自治区中部,重要的草原牧区。属温带半干旱气候区,主要土地覆盖为草地、其他	内蒙 古地区	天然草场面积辽阔,是重要的草原牧区,但同时荒漠也广泛分布

3 数据与方法

3.1 待评价数据

本文评价的 5 种大尺度土地覆盖数据集包括:

① UMD 数据集; ② IGBP DISCover 数据集; ③ MOD12Q1-2001 数据集; ④ GLC2000 数据集; ⑤ GlobCover2009 数据集。表 2 为 5 种土地覆盖数据集的特点。

表 2 5 种土地覆盖数据集的特点^[15,32]Table 2 Characteristics of the five land cover datasets^[15,32]

各土地覆盖数据集	UMD	IGBP DISCover	GLC2000	MOD12Q1-2001	GlobCover2009
数据时相	1992/04~1993/03	1992/04~1993/03	2000/01~2000/12	2001/01~2002/01	2009/01~2009/12
分类方法	K-Means	决策树	神经网络、决策树	ISODATA	分区专家分类
传感器	NOAA/AVHRR	NOAA/AVHRR	SPOT/Vegetation	Terra/MODIS	MERIS
分辨率	1 km	1 km	1 km	1 km	300 m
分类系统	IGBP 简化(14 类)	IGBP(17 类)	LCCS(23 类)	IGBP(17 类)	LCCS(23 类)

由于 5 种土地覆盖数据集所采用的分类系统各不相同,对于相同描述类别也无法完全对应,不能直接用来比较,因此需对分类系统进行转换。考虑到 Landsat TM/ETM+ 的空间分辨率及光谱分辨率,新的分类系统包括林地、耕地、草地、水体、城镇及其他大类。转换方式参考 Latifovic^[16]、Herold 等^[9]、朱榴骏等^[31],如表 3 所示。

3.2 参考数据

本文采用中分辨率的 Landsat 影像分类结果作为参考数据。由于 5 种土地覆盖数据集的时间不同,且获取与 5 种土地覆盖数据集相同时间的高质量 Landsat 影像较为困难,故分别选取 3 个时间段

(1989~1995、1999~2001、2006~2009 年)的 Landsat TM/ETM+ 影像,依次进行辐射校正及最大似然法分类,并以分类结果为参考数据评价 5 种土地覆盖数据集。1989~1995 年的 Landsat TM 影像分类结果用于评价 UMD 和 IGBP DISCover, 1999~2001 年的 Landsat TM 影像分类结果用于评价 GLC2000 和 MOD12Q1-2001, 2006~2009 年的 Landsat TM/ETM+ 影像分类结果用于评价 GlobCover2009。此外,为保证参考数据的质量,采用 Google Earth 高分影像与目视判读结合的方法选取 256 个随机验证样本点分别评价每景影像的分类结果。精度评价的结果表明参考数据的总体精度

表3 5种土地覆盖数据集的分类系统转换^[9,16,31]Table 3 Transformation in classification system of the five land cover datasets^[9,16,31]

统一分类和编码	IGBP/MOD12Q1-2001	UMD	GLC2000	GlobCover2009
1 林地	1 常绿针叶林, 2 常绿阔叶林, 3 落叶针叶林, 4 落叶阔叶林, 5 混交林, 6 郁闭灌木林, 7 稀疏灌木林, 8 树林	1 常绿针叶林, 2 常绿阔叶林, 3 落叶针叶林, 4 落叶阔叶林, 5 混交林, 6 稀疏林地, 8 郁闭灌木林, 9 稀疏灌木林	1 落叶针叶林, 2 常绿针叶林, 3 常绿阔叶林, 4 落叶阔叶林, 5 灌木, 6 稀疏林地, 24 森林镶嵌体	30 自然植被与耕地镶嵌体, 40 常绿阔叶林, 50 常绿针叶林, 60 落叶阔叶林, 70 落叶子阔叶林, 90 落叶针叶林, 100 混交林, 110 林地/灌木与草原镶嵌体, 120 草原与林地/灌木镶嵌体, 130 郁闭灌木林, 160 定期淹没阔叶林, 170 永久淹没阔叶或灌木
2 耕地	12 耕地	11 耕地	21 耕地, 23 耕地镶嵌体	11 水田, 14 旱地, 20 耕地与自然植被镶嵌体
3 草地	9 稀树草原, 10 草原, 14 耕地与自然植被镶嵌体草原	7 稀疏草原, 10 草原	8 定期淹没草地, 9 山坡草地, 10 平原草地, 11 沙漠草地, 12 草地, 22 高山亚高山草地	140 草原, 180 定期淹没草地
4 水体	17 水体	0 水体	14 河流, 15 湖泊	210 水体
5 城镇	13 城镇	13 城镇	13 城镇	190 城镇
6 其他	11 永久湿地, 16 裸地, 15 雪与冰	12 裸地	7 海滨湿地, 15 雪与冰, 16 湿地, 18 裸地, 19 砾石	150 稀疏植被, 200 裸地, 220 雪和冰覆盖区, 20 荒漠

均大于 85%, Kappa 系数均大于 0.8, 符合对参考数据的精度要求。图 2 为 7 个典型区不同时相 Landsat TM/ETM+ 影像分类结果与相应的 5 种大尺度土地覆盖数据集的比较。

3.3 亚分数混淆矩阵

亚分数混淆矩阵是指不考虑参考数据与待评价数据的尺度差异, 直接进行相同空间位置上参考数据与待评价数据的类型比较所得到的混淆矩阵^[16]。具体建立方法如下:

(1) 建立待评价数据与参考数据的空间对应关系。

(2) 针对待评价数据的每一个像元, 统计其对应空间位置上参考数据各类型的比重, 并与待评价像元的类型进行对比, 从而得到待评价像元被正确分类的比重及被错分为其他类型的比重。图 3 为待评价像元与参考数据对比示例。图 3(a) 为待评价像元, 分辨率为 1 km, 类型为林地。图 3(b) 为对应空间位置上的参考数据, 分辨率为 30 m, 统计其各类型的比重分别为林地 0.5、耕地 0.25 和水体 0.25。对比参考数据类型与待评价像元类型, 便得到待评价像元的分类精度和误差: 在待评价像元中, 0.5 的林地被正确分类, 0.25 的耕地被错分成林地, 0.25 的水体被错分成林地。

(3) 统计待评价数据所有像元的对比结果, 从而得到待评价数据完整的亚分数混淆矩阵。

基于亚分数混淆矩阵的精度评价方法具有以下优点: ①对待评价数据与参考数据进行全面对比, 更加客

观地描述数据集与参考数据的一致性; ②包含亚像元尺度的精度信息, 可以反映混合像元的精度特点; ③无需尺度转换, 可避免尺度转换引入的不确定性。

3.4 评价指标

根据亚分数混淆矩阵, 可以得到总体精度、制图精度和用户精度等精度评价指标, 以及参考数据与待评价数据总体一致性的度量指标 Kappa 系数。制图精度和用户精度可以反映不同类别的精度, 总体精度和 Kappa 系数可以反映数据集整体的精度, 采用这种指标组合能够全面地反映数据集的精度^[33]。

4 结果与分析

4.1 全部典型区的精度

由全部典型区的参考数据与各待评价数据集分别建立亚分数混淆矩阵, 得到各土地覆盖数据集在全部典型区的总体精度与 Kappa 系数(表 4)。由于篇幅限制, 未将各亚分数混淆矩阵一一列举。其中 GLC2000 的总体精度与 Kappa 系数最高, 分别为 65.64% 和 0.53, UMD 的总体精度与 Kappa 系数最低, 分别为 43.06% 和 0.20, 其余土地覆盖数据集的总体精度及 Kappa 系数介于这两者之间。

由于不同的土地覆盖数据集使用的数据源、分类体系、分类方法和辅助数据各不相同, 造成不同土地覆盖数据集的精度差异较大^[12,14]。其次生产各种土地覆盖数据集时所使用的辅助数据也是造成不

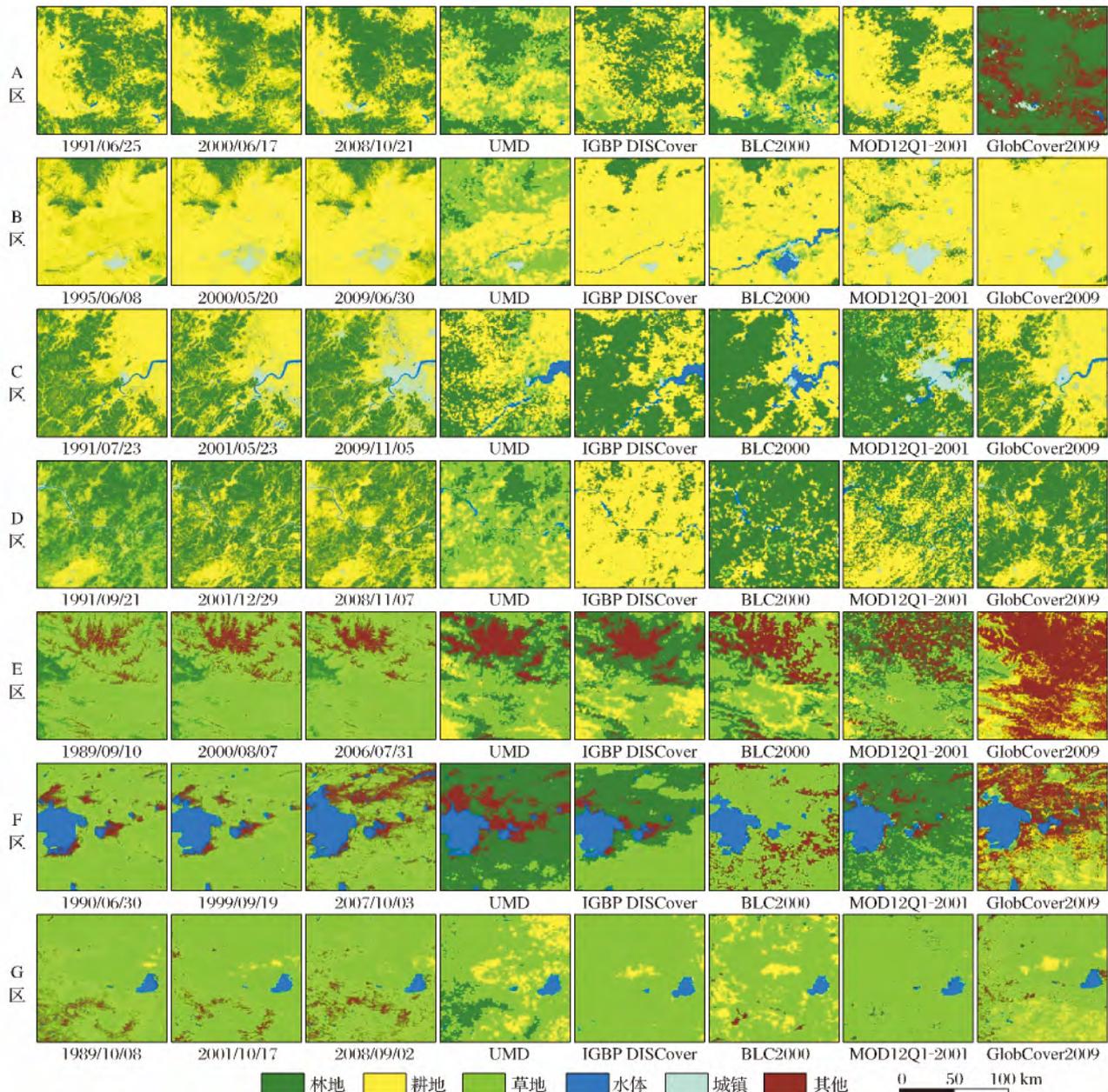


图 2 7 个典型区的 Landsat TM/ETM+ 影像分类结果与 5 种土地覆盖数据集的比较

Fig. 2 The classified results for Landsat TM/ETM+ senses of 7 typical areas compared to the five land cover data sets

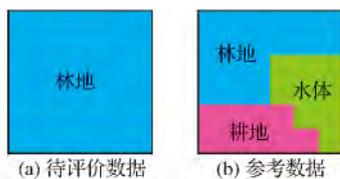


图 3 待评价数据与参考数据空间关系示意图

Fig. 3 Spatial relationship among classification data and reference data

表 4 5 种土地覆盖数据集总体精度和 Kappa 系数

Table 4 Overall accuracy and Kappa coefficient of the five land cover datasets

土地覆盖数据集	UMD	IGBP DISCover	GLC 2000	MOD12Q1-2001	GlobCover 2009
总体精度/%	43.06	56.51	65.64	57.17	53.56
Kappa 系数	0.20	0.41	0.53	0.42	0.40

同土地覆盖数据集精度差异的重要原因。

4.2 各典型区的精度

表 5 是 5 种大尺度土地覆盖数据集在各典型区

的总体精度与 Kappa 系数。在东北、华北、华中、华南 4 个自然地区内的 A、B、C、D 4 个典型区中,东北区的典型区 A 土地覆盖平整,整齐分布着大面积的

林地与耕地。分析各数据集在 A 区的分类精度可知, GLC2000 与 MOD12Q1-2001 的总体精度较高, UMD 与 IGBP DISCover 在林地与耕地的交接处将林地错分为草地, 部分耕地也被错分为草地, GlobCover 2009 将大面积的耕地错分为林地与其他类型, 导致分类总体精度最低。华中区的典型区 C 和华南区的典型区 D 的组成模式较为接近, 土地覆盖

较为破碎, 天然植被均以森林为主, 由于人类活动范围的不断扩展, 导致林地与耕地镶嵌分布。分析各土地覆盖数据集在 C、D 两区的精度可得, GLC2000 与 GlobCover2009 的总体精度较高, 但对城镇面积的估计却存在严重的低估。UMD 总体精度较低, 林地错分为草地和耕地的现象较为严重。对于华北地区的典型区 B, 人类活动对于土地覆盖组成的

表 5 5 种土地覆盖数据集在各典型区的精度

Table 5 Accuracy in each typical area of the five land cover datasets

土地覆盖数据集典型区	UMD		IGBP DISCover		GLC2000		MOD12Q1-2001		GlobCover2009	
	总体精度/%	Kappa	总体精度/%	Kappa	总体精度/%	Kappa	总体精度/%	Kappa	总体精度/%	Kappa
A	49.69	0.24	56.62	0.20	58.78	0.36	61.20	0.31	48.25	0.13
B	34.34	0.04	74.16	0.23	68.03	0.20	67.87	0.17	76.56	0.20
C	55.66	0.21	62.14	0.27	68.48	0.44	56.27	0.27	64.08	0.41
D	35.48	0.13	29.53	0.05	66.67	0.17	55.62	0.16	67.77	0.33
E	46.83	0.17	40.12	0.12	52.19	0.25	45.30	0.17	29.03	0.07
F	18.68	0.13	52.40	0.27	69.94	0.30	25.70	0.15	43.34	0.21
G	60.75	0.10	80.61	0.09	75.38	0.11	88.24	0.20	67.62	0.10

影响表现得更加明显, 城市逐步扩张, 林地面积不断减少, 耕地面积不断增加且大面积连续分布, 主要土地覆盖类型为耕地。分析各土地覆盖数据集在 B 区的精度可得, IGBP DISCover 的总体精度最高, 对于耕地面积的估计与参考数据最为接近。UMD 的总体精度最低, 林地面积被严重低估, 且与耕地、草地的混分现象严重。整体而言, 对于连续分布的大面积林地与耕地的区域, GLC2000 与 MOD12Q1-2001 的总体精度较高, UMD 与 GlobCover2009 的总体精度较低; 对于土地覆盖呈破碎状, 主要为林地与耕地的区域, GLC2000 与 GlobCover2009 的总体精度较高, UMD 的总体精度较低; 对于连续分布的大面积耕地区域, IGBP DISCover 与 GlobCover2009 的总体精度较高, 但 GlobCover2009 对于水体的估计存在较大错误。5 种数据集中 GLC2000 在 4 个典型区内均具有较高的分类精度, 对各主要土地覆盖类型面积的估计与参考数据相比也最为接近, 但在对城镇的分类上存在较大错误, 城镇多被错分为水体。

在西北干旱区内的西北、内蒙两个自然地区的 E、G 典型区中, E 区的天然植被主要为连续分布的大面积草地, 裸露的岩石也是区内土地覆盖类型的一种。分析 E 区内各土地覆盖数据集的精度可得, GLC2000 具有最高的分类精度, 对于草地面积的估计与参考数据最为接近。UMD、IGBP DISCover 中将草地错分为林地的现象较为严重, MOD12Q1 中草地与林地的混分现象较为严重, GlobCover2009

中则将大面积的草地错分为其他类型, 导致分类精度在 5 种数据集中最低。G 区的土地覆盖主要为破碎分布的荒漠草原与荒漠, 同时也破碎分布小面积的稀疏灌木林。5 种土地覆盖数据集在 G 区的总体精度均较高, MOD12Q1 的总体精度最高。但是 5 种数据集对荒漠面积的估计远低于参考数据, 荒漠被错分为荒漠草原, 荒漠草原被错分为耕地。对于青藏自然地区的 F 典型区, 其土地覆盖类型以连续分布的高山草甸, 破碎分布的水体、盐碱地为主。GLC2000 在 F 区的分类精度最高, 对草地面积、水体面积的估计与参考数据也最为接近。UMD、IGBP DISCover、MOD12Q1-2001 均将大面积的草地错分为林地, GlobCover2009 则将大面积的草地错分为其他类型。整体而言, 对于连续分布着大面积草地的区域, GLC2000 的总体精度较高; 对于荒漠草原与荒漠呈破碎分布的区域, MOD12Q1-2001 的总体精度较高。5 种数据集中 GLC2000 在 E、F、G 3 个典型区的精度均较高, UMD 与 GlobCover 的总体精度较低。

表 5 中部分区域存在总体精度较高而 Kappa 系数极低的现象, 如 B、G 区。对于这一现象, 已有研究指出, Kappa 系数在评价一致性时存在两个悖论: ①如果正确分类的数值在各类别间存在显著差异(即亚分数混淆矩阵中对角线上数值差异很大), 这样即使总体精度较高, Kappa 系数仍会极低; ②行列合计值对称时的 Kappa 系数小于不对称时的 Kappa 系数^[34-36]。以 IGBP DISCover 在 G 区的亚

分数混淆矩阵为例, G 区被正确分类的草地为 11 428 km², 耕地为 7 km², 水体为 173 km², 林地、城镇及其他被正确分类的面积均为 0, 由于正确分类的数值在各类别间存在较大差异, 即使部分区域

总体精度很高, 仍会出现 Kappa 系数极低的现象。

4.3 各土地覆盖类型的精度

图 4 为各土地覆盖类型在不同数据集和不同典型区的制图精度及用户精度, 结果表明:

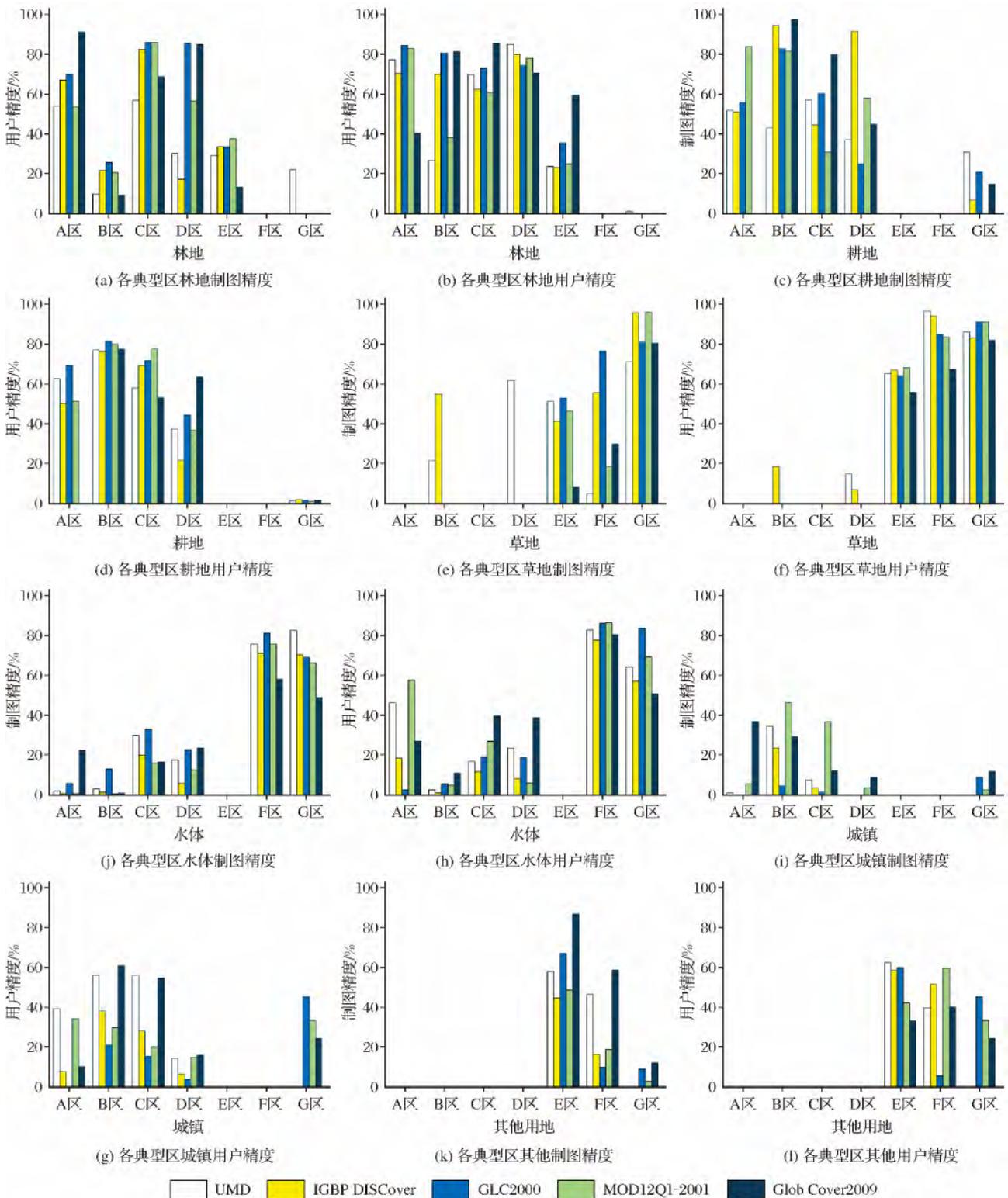


图 4 各土地覆盖类型在各典型区的制图精度与用户精度

Fig. 4 Producer accuracy and user accuracy of each land cover type in the different data sets and typical areas

(1)在对林地的分类中,UMD、IGBP DISCover、MOD12Q1-2001 3种数据集均使用IGBP分类体系,3种数据集的相似性较高,3种数据集在西北干旱区和青藏高原区的E、F、G3个典型区均存在将草地错分为林地的情况。5种数据集在东部季风区各典型区的用户精度较高。

(2)在对耕地的分类中,5种数据集在西北干旱区的E、G两个典型区均存在将草地错分为耕地情况,在东部季风区各典型区的制图精度与用户精度较高,并且5种数据集在东部季风区的典型区B均具有最高的用户精度。IGBP DISCover在D区将大面积的林地错分为耕地,导致耕地面积高估近50%。GlobCover2009在A区将耕地完全错分成林地,在F区将30%的草地错分为耕地。

(3)在对草地的分类中,5种数据集在东部季风区的各典型区均存在将耕地、林地错分为草地的现象。在西北干旱区和青藏高原区各典型区的用户精度较高IGBP DISCover在G区将沙地和盐碱地全部错分为草地。MOD12Q1-2001在F区将78%的草地错分为林地。GlobCover2009在E、F区将大量的草地错分为其他。

(4)在对水体的分类中,5种数据集在东部季风区各典型区的用户精度均小于西北干旱区和青藏高原区各典型区,分析各亚分数混淆矩阵可知,东部季风区各典型区中水体与耕地、林地的混分现象比较严重,而在西北干旱区和青藏高原区各典型区,由于水体面积相对较大且与周围地物的光谱特征存在较大差异,分类精度普遍高于东部季风区。5种数据集在青藏高原区的典型区F均具有最高的用户精度。

(5)在对城镇的分类中,除MOD12Q1-2001数据集在B、C区存在高估外,5种数据集在各典型区都存在不同程度的低估,且分类精度较低,分析各亚分数混淆矩阵可得,7个典型区中大部分的城镇都与草地、耕地、水体、林地存在严重的混分。这是因为在中低分辨率影像中城镇本身均为混合像元,极易被错分为其他类型。

(6)在对其他类型的分类中,除GlobCover2009在东部季风区的A区中将大量的耕地错分为其他外,5种数据集在东部季风区的其他各典型区均具有较高的精度。在西北干旱区与青藏高原区各典型区中,5种数据集在E、F、G3个典型区的用户精度均较低。

在对以上各土地覆盖类型的精度评价中发现,不同的土地覆盖数据集对不同土地覆盖类型具有不

同的分类精度。造成这一现象的主要原因有:①5种土地覆盖数据集所使用的分类框架具有较大的差异^[12,14];②由于不同区域的土地覆盖组成模式存在较大差异,不同的分类框架对不同的土地覆盖组成模式的响应也存在差异,也会造成同地类的分类精度在不同产品中各不相同。现有研究认为这一现象与局部区域的土地覆盖组成相关^[16]。

4.4 空间配准误差对精度评价的影响

参考数据与待评价数据集的分辨率相差较大,空间配准上的误差不可避免,这类误差将会影响精度评价的准确性。现有研究表明空间配准会造成数据集精度的低估^[37];对于复杂的土地覆盖组成模式,空间配准误差对精度评价的影响会高于均质的地表覆盖情况^[6]。为评估空间配准误差对精度评价的影响,对参考数据沿经线和纬线方向以0.3 km为步长进行偏移,分别计算每次偏移后的参考数据和土地覆盖数据所构成的亚分数混淆矩阵,得到每次偏移后的土地覆盖数据的总体精度,与参考数据无偏移时土地覆盖数据的总体精度进行比较。

图5为5种土地覆盖数据集与参考数据在不同偏移距离下的评价误差,评价误差=(无偏移处精度-偏移处精度)/无偏移处精度。当空间偏移在3 km×3 km之内时,5种土地覆盖数据集空间匹配误差对精度评价结果的影响均小于6%。以总体精度最高的GLC2000为例,当数据集与参考数据的偏移距离为0 km时,总体精度为65.64%;当参考数据在纬线方向偏移3 km时,利用亚分数混淆矩阵评价的数据集总体精度下降5.94%,为61.75%。因此以Landsat TM/ETM+分类数据评价5种土地覆盖数据集的精度时,参考数据与待评价数据在空间配准上的误差对精度评价的结果影响较小。

5 结 语

本文以5种大尺度土地覆盖数据集同时期的中国7个典型区Landsat TM/ETM+影像分类结果为参考数据,采用亚分数混淆矩阵对5种大尺度土地覆盖数据集在各典型区的精度、各土地覆盖类型精度进行评价,并对空间匹配误差对精度评价的影响进行了研究。主要结论如下:

(1)对于各数据集在全部典型区的总体精度,GLC2000的总体精度和Kappa系数最高,分别为65.64%、0.53;UMD的总体精度和Kappa系数最低,分别为43.06%、0.20;其余3种数据集介于这两者之间。

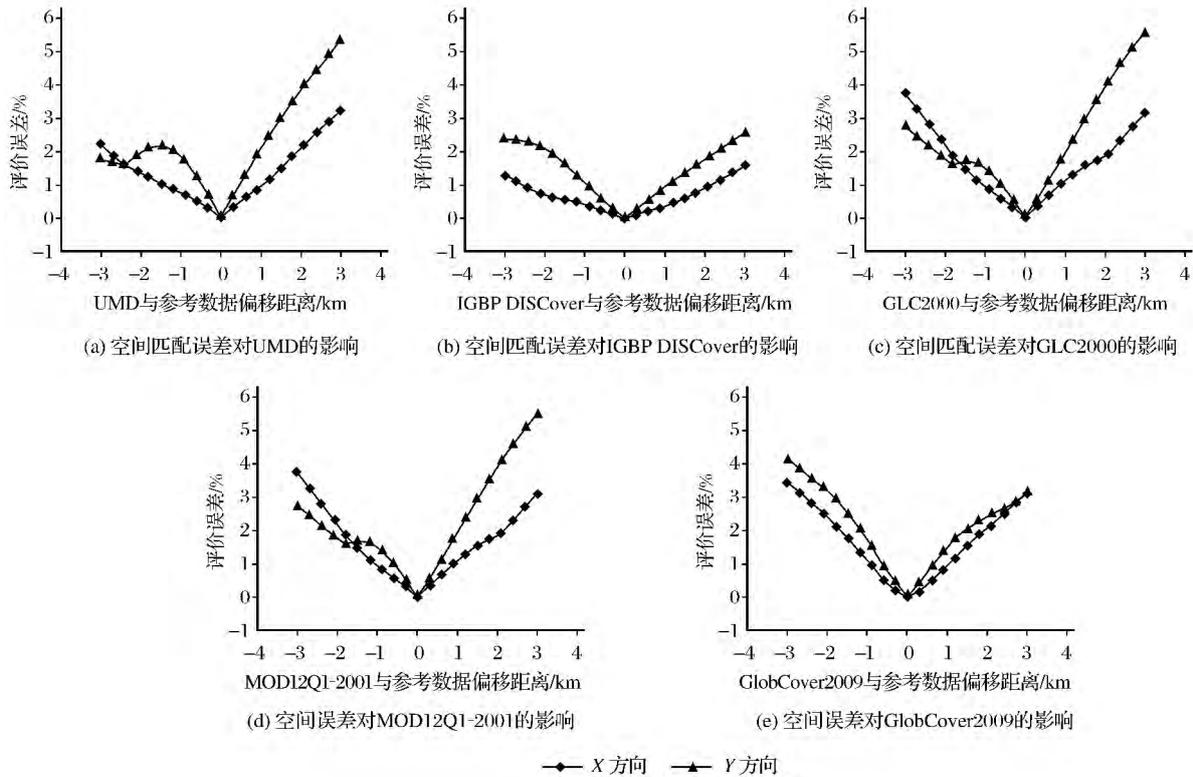


图 5 空间配准误差对精度评价的影响

Fig. 5 Influence of misregistration error on accuracy assessment

(2) 对于连续分布的大面积林地与耕地区域, GLC2000 与 MOD12Q1-2001 的总体精度较高; 对于土地覆盖呈破碎状, 主要为林地与耕地的区域, GLC2000 与 GlobCover2009 的总体精度较高; 对于仅连续分布的大面积耕地区域, IGBP DISCover 与 GlobCover2009 的总体精度较高; 对于连续分布着大面积草地的区域, GLC2000 的总体精度较高; 对于荒漠草原与荒漠呈破碎分布的区域, MOD12Q1-2001 的总体精度较高。

(3) 关于各土地覆盖类型的精度, 对林地和耕地的分类, 5 种数据集在东部季风区各典型区的用户精度较高; 对草地和水体的分类, 5 种数据集在西部干旱区和青藏高原区各典型区的用户精度较高; 对城镇的分类, 5 种数据集除 MOD12Q1-2001 外在各典型区均存在不同程度的低估; 对于其他的分类, 5 种数据集在西北干旱区和青藏高原区各典型区的用户精度均较低。

(4) 通过实验模拟空间配准误差对精度评价结果的影响, 结果表明当以 Landsat TM/ETM+ 为参考数据评价 5 种土地覆盖数据集的精度时, 参考数据与待评价数据在空间配准上的误差对精度评价的结果影响较小, 均小于 6%。

参考文献 (References):

- [1] Gallo K P, Easterling D R, Peterson T C. The Influence of Land Use Land Cover on Climatological Values of the Diurnal Temperature Range [J]. *Journal of Climate*, 1996, 9 (11): 2941-2944.
- [2] Sellers P J, Dickinson R E, Randall D A, *et al.* Modeling the Exchanges of Energy, Water, and Carbon between Continents and the Atmosphere [J]. *Science*, 1997, 275(5299): 502-509.
- [3] Brown J F, Loveland T R, Ohlen D O, *et al.* The Global Land-cover Characteristics Database: The Users' Perspective [J]. *Photogrammetric Engineering and Remote Sensing*, 1999, 65 (9): 1069-1074.
- [4] Cihlar J. Land Cover Mapping of Large Areas from Satellites: Status and Research Priorities [J]. *International Journal of Remote Sensing*, 2000, 21(6-7): 1093-1114.
- [5] Hansen M C, Reed B. A Comparison of the IGBP DISCover and University of Maryland 1km Global Land Cover Products [J]. *International Journal of Remote Sensing*, 2000, 21(6-7): 1365-1373.
- [6] Loveland T R, Reed B C, Brown J F, *et al.* Development of a Global Land Cover Characteristics Database and IGBP DISCover from 1 km AVHRR Data [J]. *International Journal of Remote Sensing*, 2000, 21(6-7): 1303-1330.
- [7] Friedl M A, McIver D K, Hodges J C F, *et al.* Global Land Cover Mapping from MODIS: Algorithms and Early Results

- [J]. *Remote Sensing of Environment*, 2002, 83(1-2): 287-302.
- [8] Bartholome E, Belward A S. GLC2000: A New Approach to Global Land Cover Mapping from Earth Observation Data [J]. *International Journal of Remote Sensing*, 2005, 26(9): 1959-1977.
- [9] Herold M, Mayaux P, Woodcock C E, *et al.* Some Challenges in Global Land Cover Mapping: An Assessment of Agreement and Accuracy in Existing 1 km Datasets [J]. *Remote Sensing of Environment*, 2008, 112(5): 2538-2556.
- [10] Scepan J, Estes J E. Thematic Validation of Global Land Cover Datasets-Procedures and Interpretation Methods [C]// *IEEE International Geoscience and Remote Sensing Symposium Sydney*, 2001: 1119-1121.
- [11] Friedl M A, McIver D K, Hodges J C F, *et al.* Global Land Cover Mapping from MODIS: Algorithms and Early Results [J]. *Remote Sensing of Environment*, 2002, 83(1-2): 287-302.
- [12] Giri C, Zhu Z L, Reed B. A Comparative Analysis of the Global Land Cover 2000 and MODIS Land Cover Datasets [J]. *Remote Sensing of Environment*, 2005, 94(1): 123-132.
- [13] Mayaux P, Eva H, Gallego J, *et al.* Validation of the Global Land Cover 2000 Map [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2006, 44(7): 1728-1739.
- [14] McCallum I, Obersteiner M, Nilsson S, *et al.* A Spatial Comparison of Four Satellite Derived 1 km Global Land Cover Datasets [J]. *International Journal of Applied Earth Observation and Geoinformation*, 2006, 8(4): 246-255.
- [15] Kaptué Tchuenté A T, Roujean J-L. Comparison and Relative Quality Assessment of the GLC2000, GLOBCOVER, MODIS and Ecoclimap Land Cover Data Sets at the African Continental Scale [J]. *International Journal of Applied Earth Observation and Geoinformation*, 2011, 13(2): 207-219.
- [16] Latifovic R, Olthof I. Accuracy Assessment Using Sub-pixel Fractional Error Matrices of Global Land Cover Products Derived from Satellite Data [J]. *Remote Sensing of Environment*, 2004, 90(2): 153-165.
- [17] Waser L T, Schwarz M. Comparison of Large-area Land Cover Products with National Forest Inventories and Corine Land Cover in the European Alps [J]. *International Journal of Applied Earth Observation and Geoinformation*, 2006, 8(3): 196-207.
- [18] Neumann K, Herold M, Hartley A, *et al.* Comparative Assessment of CORINE2000 and GLC2000: Spatial Analysis of Land Cover Data for Europe [J]. *International Journal of Applied Earth Observation and Geoinformation*, 2007, 9(4): 425-437.
- [19] Stehman S V, Wickham J D, Wade T G, *et al.* Designing a Multi-objective, Multi-support Accuracy Assessment of the 2001 National Land Cover Data (NLCD 2001) of the Conterminous United States [J]. *Photogrammetric Engineering and Remote Sensing*, 2008, 74(12): 1561-1571.
- [20] Wu W, Shibasaki R, Yang P, *et al.* Validation and Comparison of 1 km Global Land Cover Products in China [J]. *International Journal of Remote Sensing*, 2008, 29(13): 3769-3785.
- [21] Ran Youhua, Li Xin, Lu Ling. Evaluation of Four Remote Sensing based Land Cover Products over China [J]. *Journal of Glaciology and Geocryology*, 2009, 31(3): 490-500. [冉有华, 李新, 卢玲. 四种常用的全球 1km 土地覆盖数据集中国区域的精度评价 [J]. *冰川冻土*, 2009, 31(3): 490-500.]
- [22] Gao H, Jia G S. Spatial and Quantitative Comparison of Satellite-Derived Land Cover Products over China [J]. *Atmospheric and Oceanic Science Letter*, 2012, 5(5): 426-434.
- [23] Herold M, Woodcock C E, Gregorio A D, *et al.* A Joint Initiative for Harmonization and Validation of Land Cover Datasets [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2006, 44(7): 1719-1727.
- [24] Ran Y H, Li X, Lu L. Evaluation of Four Remote Sensing based Land Cover Products over China [J]. *International Journal of Remote Sensing*, 2010, 31(2): 391-401.
- [25] Turner B L, Moss R H, Skole D L. Relating Land Use and Global Land-cover Change: A Proposal for an IGBP-HDP Core Project [R]. *International Geosphere-Biosphere Programme*, Stockholm, Sweden, 1993: 24.
- [26] Luo Kaifu. The Draft Natural Divisions of China [M]. Beijing: Science Press, 1956. [罗开富. 中国自然区划草案 [M]. 北京: 科学出版社, 1956.]
- [27] Huang Bingwei. Comprehensive Physical Regionalization of China [J]. *Journal of Geographical Sciences*, 1958, 24(4): 348-363. [黄秉维. 中国综合自然区划的初步草案 [J]. *地理学报*, 1958, 24(4): 348-363.]
- [28] Ren Mei'e. Chinese Natural Geography Program [M]. Beijing: The Commercial Press, 1979. [任美铎. 中国自然地理纲要 [M]. 北京: 商务印书馆, 1979.]
- [29] Zhao Songqiao. A New Draft about Comprehensive Physical Regionalization of China [J]. *Journal of Geographical Sciences*, 1983, 38(1): 1-10. [赵松乔. 中国综合自然地理区划的一个新方案 [J]. *地理学报*, 1983, 38(1): 1-10.]
- [30] Zhao Ji. Physical Geography of China [M]. Beijing: Higher Education Press, 1995. [赵济. 中国自然地理 [M]. 北京: 高等教育出版社, 1995.]
- [31] Zhu Liujun, Xiao Pengfeng, Feng Xuezhi, *et al.* Multi-scale Accuracy Assessment of Land Cover Datasets based on Histograms [J]. *Journal of Remote Sensing*, 2013, 17(6): 1459-1474. [朱榴骏, 肖鹏峰, 冯学智, 等. 基于直方变差图的土地覆盖数据集多尺度精度评价 [J]. *遥感学报*, 2013, 17(6): 1459-1474.]
- [32] McCallum I, Obersteiner M, Nilsson S, *et al.* A Spatial Comparison of Four Satellite Derived 1 km Global Land Cover Datasets [J]. *International Journal of Applied Earth Observation and Geoinformation*, 2006, 8(4): 246-255.
- [33] Liu C, Frazier P, Kumar L, *et al.* Comparative Assessment of

- the Measures of Thematic Classification Accuracy[J]. Remote Sensing of Environment, 2007, 107(4): 606-616.
- [34] Feinstein A R, Cicchetti D V. High Agreement but Low Kappa; I. The Problems of Two Paradoxes[J]. Journal of Clinical Epidemiology, 1990, 43: 543-549.
- [35] Byrt T, Bishop J, Carlin J B. Bias, Prevalence and Kappa[J]. Journal of Clinical Epidemiology, 1990, 46: 423-429.
- [36] Lantz C H, Nebenzahl E. Behavior and Interpretation of the Kstatistic: Resolution of the Two Paradoxes[J]. Journal of Clinical Epidemiology, 1996, 49: 431-434.
- [37] Foody G M. Status of Land Cover Classification Accuracy Assessment[J]. Remote Sensing of Environment, 2002, 80: 185-201.

Assessment of Large-scale Land Cover Datasets in Typical Areas of China based on Sub-fractional Error Matrix

Jiang Luyuan, Xiao Pengfeng, Feng Xuezhi, Li Yun, Zhu Liujun

(Department of Geographic Information Science, Geographic and Oceanographic Sciences, Nanjing University, Nanjing 210023, China)

Abstract: This paper evaluated the accuracy of five large-scale land cover datasets based on sub-fractional error matrix, by taking Landsat TM/ETM+ image classification results of seven typical areas in China as the reference data, which provides the scientific basis for the use of datasets. The sub-fractional error matrix can avoid errors caused by the scale difference between reference data and datasets, and evaluate the accuracy on sub-pixel scale and reflect the classification accuracy and classification method error with different dominant fraction. The results show that; the overall accuracy of GLC2000 is the highest in all typical areas, is at 65.64%; and UMD is the lowest in all typical areas. GLC2000 has a higher classification accuracy in the areas covered by forest, cropland and grass; the classification accuracy of UMD is the lowest or the lower one in each typical area. The five land cover datasets have a lower classification accuracy in urban and other; while with a higher classification accuracy of grass and water in each typical area of arid region of northwest China and Tibet Plateau region.

Key words: Sub-fractional error matrix; Large-scale; Land cover datasets; Accuracy assessment