

MODIS 土地覆盖产品研究进展

卫亚星, 王莉雯

辽宁师范大学, 海洋经济与可持续发展研究中心, 城市与环境学院, 辽宁 大连 116029

摘要 遥感技术在近几十年来发展迅速, 在生态和环境等领域得到了广泛的应用。MODIS 是一个比较新的数据源, 它的一系列产品中, 土地覆盖产品是一个比较重要的产品, 在全球和区域模型中经常应用。文章比较全面地介绍了 MODIS 土地覆盖产品的生成过程, 探讨了基于神经网络和基于决策树的土地覆盖分类方法的特点, 强调了土地覆盖分类中方向信息的重要性, 比较详细地介绍了变化矢量分析方法和基于人工神经网络的土地覆盖变化侦察方法, 分析了 IGBP 定义的全球 17 种土地覆盖类型, 并且与另外的 3 种分类系统进行了比较。

关键词 土地覆盖; 遥感; MODIS

中图分类号: TP75 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2010)07-1848-05

引言

1999 年 7 月发射 Terra 卫星以后, 可以提供每个季度的全球 1 km 分辨率的土地覆盖产品。1 km 的土地覆盖变化产品是近期开发的。1 km 的土地覆盖和土地覆盖变化产品依赖于 1 km 的栅格数据库, 该数据库是由 MODIS 2 级和 3 级产品生成的^[1]。生成 1 km 的土地覆盖产品需要输入: (1) EOS 的陆地和水掩模, 用于限制陆地区和浅水区的分类; (2) 利用 BRDF 调整过的天底反射系数, 由 MODIS 陆地波段 (1~7) 的 BRDF/ALBEDO 产品生成; (3) 空间纹理从第 1 波段 (红光, 250 m) 产生; (4) 以 16 天为一个周期的 1 km 的方向反射信息; (5) 以 16 天为一个周期的 1 km 的 MODIS 增强型植被指数 (EVI); (6) 8 天为一个周期的 500 m 雪盖产品; (7) 8 天为一个周期的 1 km 陆地表面温度; (8) 地形高程信息。以上这些数据构成了全球统一的多时相 1 km 栅格数据库, 输入以后优化了分类算法^[2]。

陆地覆盖参数根据国际地圈生物圈计划 (IGBP) 的分类系统分了 17 个土地覆盖类型。这个系统包括 11 类自然植被, 3 类开发过的土地, 3 类非植被土地类型。然后利用决策树和人工神经网络分类算法, 基于训练数据指定土地覆盖类型, 通过处理 32 天的数据库产生土地覆盖类型。为了减少计算费用, 增加灵活性, 分类是按照全球 7 大洲分别进行的^[3]。

土地覆盖产品在获得某区域 32 天的合成数据集 3 个月以后发布。全球的土地覆盖产品在获得全球的输入产品以后

可以获得。期间, 可以产生一系列的原型土地覆盖产品。获得土地覆盖变化参数, 需要 2 年的数据^[4]。大多数研究基于植被覆盖的结构特征, 遥感能够敏感地反映这个特征。通过研究 MODIS 数据的光谱、时相、空间特征、方向信息, 可以识别一系列的土地覆盖类型。土地覆盖变化产品的目标是发现和量化土地覆盖的变化, 研究自然和人类的影响过程, 完善全球和区域模型^[5]。

1 全球尺度的土地覆盖数据

当前土地覆盖数据集用于全球气候模型中的做法起源于以前已经存在的大量地图和地图集。这种方法受几点限制。第一种情况是, 参考源本身具有不同的日期、空间尺度、分类方案。从参考类单元到分类系统, 以及土地表面数据集中所用的比例尺的混乱, 可以引起最终产品中的错误。例如, 当基于植物和基于气候的分类不兼容时, 可能需要两者结合重新分类, 生成新的覆盖类型。第二种情况是, 如果结果数据集中包含统计数据, 则可以认为它已经包含了采集源中存在的误差。第三种情况是, 一些数据集反映的是植被最好时期的地图, 并且它是从气候变量 (如温度和降雨量) 推断的, 而不是真正的植被类型^[6]。

Townshend 等比较了一个地区的大规模的土地覆盖数据集的统计值和空间分布。他们不赞同这些产品中基本土地覆盖类型所占的相对百分比。更多人的意见承认特定植被类型的相对覆盖面积, 但认为空间分布不对。这样的数据库具

收稿日期: 2009-09-02, 修订日期: 2009-12-06

基金项目: 教育部人文社会科学重点研究基地项目 (08JJD790142) 和辽宁省教育厅人文社会科学研究项目 (2007T095) 资助

作者简介: 卫亚星, 1969 年生, 辽宁师范大学城市与环境学院副教授 e-mail: wyx9585@sina.com

有明显的局限性。

很多研究人员试图用 NOAA 系列的气象卫星上所搭载的粗空间分辨率的 AVHRR 传感器的高时间频率来建立区域尺度的土地覆盖数据库。几乎毫无例外,这些工作都是将 AVHRR 的第 1 波段和第 2 波段转变成归一化植被指数 (NDVI) 值。配准的一个时间序列的 NDVI 图象构成了一个数据集,将这个时间序列的最大的 NDVI 值作为每一个像元位置的输出值。构成过程中的选择措施强烈地依赖于大气和气溶胶的分布,它们在红光和近红外波段的分布减少了 NDVI 值。云在红光和近红外波段的反射都很强烈,经过云污染处理以后的 NDVI 值更低。NDVI 值的构成与变化的视角和光照几何学相联系,靠近上散方向的测量法趋向于选稍微高的 NDVI 值。合成时间的长短取决于植被变化的频率和产生无云图象所必需的最小时间^[7]。

虽然 NDVI 不直接提供土地覆盖类型,但是一般可以量化陆地表面的生物物理活动。无论如何,一个时间序列的 NDVI 值能够根据物候或季节信号将不同的土地覆盖类型分开。Reed 等人已经利用多时相的物候规律从 AVHRR 数据中得出了土地覆盖分类系统。Lambin 等发现利用一个时间序列的地表温度与 NDVI 的比值比只用 NDVI 提供的分类更稳定。

Townshend 等根据合成的全球 NDVI 数据完成了南美的监督分类。虽然他们没有用实地数据验证他们的成果,但他们发现通过增加该时间序列中的图象数量可以充分地改善训练样地的精度。Koomanoff 利用 NOAA 的全球植被指数产品的一年集成的 NDVI 值生成了全球植被图,这项工作的结论是有 9 种植被类型并不依赖于 NDVI 的季节性。Lloyd 使用了一个二进制分类器,它是基于某一个时间序列的 NDVI 数据的综合指数,植物物候历变量包括最大光合活动的日期、生长季的时间长度、每天 NDVI 值的平均数,将变量通过已经分层的像元输入到二进制决策树分类器中,首先考虑的是最大 NDVI 的日期,然后考虑的是生长季的时间长度,最后考虑的是每日 NDVI 值的平均数。

作为国际卫星陆地表面气候项目的一个活动,粗分辨率的全球陆地表面参数数据库被拷贝在 5 个压缩盘中发布。数据库包括土地覆盖类型、光合作用有效辐射的吸收部分 (FPAR)、叶面积指数 (LAI)、粗糙度、冠层的绿度值、以及全球的气象、土壤、和水文数据,数据库的空间尺度是 1 度乘 1 度。如 FPAR、LAI、冠层的绿度值是从 AVHRR NDVI 8 km 的合成数据中得到的^[8]。土地覆盖类型是基于 8 km 数据到 1° 的聚合,采用了 NDVI 的监督分类。

Friedl 等已经归档了 1 度分辨率, 11 个土地覆盖类型的全球土地覆盖数据。全球土地覆盖图 and 数据库已经被用于估计和推断地球表面参数。1° 的 AVHRR 分析是基于对前人所做地图的认可,来定义训练和测试数据的。Loveland 等用非监督分类法得出全球的土地覆盖类型,目前正在被验证之中,他们所用的训练和测试点数据是基于 Landsat 的 MSS 和 TM 数据上的多边形来确定的 11 个土地覆盖类型。

Loveland 等用 IGBP 的分类系统和季节土地覆盖区域分类系统制作了北美的土地覆盖地图。这些地图是基于一年的

按月合成的 AVHRR-LAC 数据来生成美国非监督分类的土地覆盖类型图。同时基于辅助的环境数据,例如高程和生态区域,结果类型被进一步分层。类型的指定基于时间曲线和大量的辅助源。由于被合成的 NDVI 数据的质量限制和辅助源精度的限制,该数据集最多代表 1 km 空间分辨率的 AVHRR 数据的大面积分类。

在 IGBP-DIS 的赞助下, Loveland 的努力得到了扩展,基于 1 km 的 AVHRR 全球数据库的接受时间是 1992 年 4 月到 1993 年 9 月,这些数据已经被集中到一起,在 EROS 的数据中心按 10 天合成了。10 天合成的 AVHRR 数据用最大的 NDVI 值合成了每个月的 NDVI 值,达到去云、消除地形影响和极端偏离星下点像元、以及辐射率对扫描角的依赖。使用按月合成的 AVHRR 数据可能有一点问题。一些学者的分析认为合成偏离了选择离星下点近的像元,尤其在北半球冬季月份的扫描视角。对任何大面积投影,他们发现有效的地图单元在地理位置上是可变的,在这种情况下,采用了 Goode 的等面积投影系统和重采样方法,由于没有进行传感器校正,所以多时相 NDVI 信号的时间轨道是混乱的^[9]。

全球 NDVI 数据为土地覆盖分类提供了一个多时相的数据库,可以采用非监督分类法分类。IGBP 的全球产品已经用全球网络的 400 个样点进行了验证,这些样点是利用较高分辨率的 Landsat TM 和 SPOT-XS 数据通过专家解译得到的。

2 利用神经网络进行土地覆盖分类

在过去的 15 年中,神经网络已经被证明是改善遥感信息提取最有效的方法。利用神经网络进行遥感数据的分类,在遥感文献中大约是 15 年前开始出现的。从那以后,例子和应用案例逐渐多了起来。利用神经网络分类器处理的遥感数据集包括: Landsat MSS、Landsat TM、合成孔径雷达、SPOT HRV、AVHRR、航空扫描数据。一些研究中也使用了辅助数据,例如,地形和纹理。虽然很多人经常用监督分类法,但是也采用自组织神经网络的非监督分类法。几乎在所有的情况下,已经证明神经网络分类器优于传统分类器,整体精度提高 10%~20%。由于成功应用神经网络分类案例数量的增多,清楚地表明在遥感中基于神经网络的分类可以比传统方法产生更高精度的结果。原因包括: (1)神经网络分类器没有对数据的分布做先验假设,能够学会类型分布的非线性和不连续模式; (2)神经网络可以容纳间接数据,例如:构造信息、坡度、坡向、高程; (3)神经网络是非常灵活的,可以提高解决特殊问题的能力。

遥感中的大多数神经网络分类,用后向电磁辐射传播算法训练过的多层前馈网络,是基于递归学习过程伴随逐步下降搜索。无论怎样,这种训练过程对选择初始网络参数是比较敏感的。神经网络用适应共振理论 (ART) 克服这些问题。根据 ART 原理组织的网络在学习过程中是稳定的,同时保持了足够的可塑性来学习新模式。MODIS 土地覆盖分类和变化探测用了模糊 ARTMAP 的神经网络方法。模糊 ARTMAP 是一种监督学习系统,综合了模糊逻辑和适应性共振

理论模型,最近已经被用于森林植被制图。

最近的研究突出了模糊 ARTMAP 体系结构及它在土地覆盖分类中的应用。在一项研究中,西非一年的按月合成的 1.1 km 空间分辨率的 AVHRR LAC NDVI 数据,用模糊 ARTMAP 分类器分成了 6 个主要的生活型种类。每个类的精度范围在 66%~98% 之间,平均精度是 83%。另一项研究用奥地利维也纳城北部的一景 TM 评价多层感知器、光线基本函数、模糊 ARTMAP 网络的性能^[10]。更全面的结果(精度和时间)是 MLP(不考虑权重)伴随的模糊 ARTMAP 获得的。试验数据的分类精度用模糊 ARTMAP 法是 98%,而最好的 MLP 体系结构只能到 90%。

3 利用决策树进行土地覆盖分类

决策树分类技术已经被成功地用于大范围的分类问题,但是只到最近才在遥感中做了详细试验。这项技术对于遥感分类问题的解决具有实质的优点,这在于它的灵活、简单、计算效率高。尤其在洲际和全球尺度,用决策树分类算法解决土地覆盖分类问题的方法正获得越来越高的认可。对于已经很好理解了的数据集的分类问题,分类树可以只根据专家的意见定义^[11]。这种方法是由 Running 等提出的,他们根据 AVHRR 的 NDVI 的阈值定义了全球植被的分类策略。在这个框架中,基于生态和遥感知知识给阈值的实际值定义了一个先验值。然而,在实践中,由于 NDVI 阈值的精确值随时空变动,这个过程是难以实现的,所以只用用户的知识指定比较困难。

更普遍的做法是,利用决策树定义的分类结构是用统计方法从训练数据中估计的。最近一系列的工作已经证明,决策树提供了解决土地覆盖分类问题的精确有效的方法。例如,DeFries 等用 8 km 的 AVHRR 数据集,采用决策树成功地绘制了土地覆盖图。相似地,Friedl 等证明决策树为洲际到全球尺度的土地覆盖制图问题提供了有力的分类方法。决策树的优点中,对遥感问题尤其有用的是它们处理噪音和缺损数据的能力。同时,它们不需要关于输入数据分布情况的假设,提供了一种直接的分类结构。

4 土地覆盖分类的方向信息

虽然在文献中关于遥感数据的分类使用光谱、时相、空间信息的报道已经很长时间了,但是一直到 1997 年,才认识到方向信息的重要性^[12]。Barnsley 等利用一个农业试验点的有方向性的航空扫描图象,注意到通过用主成分转换和给两波段的光谱数据加三个视角,分类精度大约提高了 20%。Abuelgasim 从 NASA 的 ASAS 航空仪器获得方向反射比图象,利用非监督和监督混合的神经网络分类器区分美国明尼苏达州泰加林中的 5 种主要的土地覆盖类型。他们利用在一个单一的近红外波段中获得的 7 个方向图象,用混合分类器的精度达到了 89%,比较表明,传统的前馈后向传播神经网络

的精度为 85%,最大似然法分类器的精度是 61%。

5 土地覆盖单元

在区域和全球模型研究中,土地覆盖参量的主要目的是加强从土地覆盖获得的生物物理信息的干涉。因而,土地覆盖所需的特殊分类单位不仅用高精度的遥感和辅助数据可以分辨,而且需要与地表的物理特性和主要的植被直接建立关系。为了这个目的,专门联合 IGBP 的核心计划,IGBP-DIS 开发了一套含 17 个全球土地覆盖类型的集合。它们被用于 EROS 数据中心的全球 1 km AVHRR LAC NDVI 数据集的分类。因为 IGBP 的这套系统是为 1 km 分辨率的全球土地覆盖产品开发的,目的相同(模型的生物物理参量化),IGBP 的分类也被用于 MODIS 的土地覆盖产品。

表 1 提供了 IGBP 土地覆盖单位的名单,并给予了描述。名单包括 11 类自然植被类型,3 类开发过的镶嵌地,3 类非植被土地。自然植被单元区分了常绿和落叶、阔叶和针叶林,混合林,密闭灌丛和稀疏灌丛,萨瓦那和木本萨瓦那,草原,大片的永久湿地。3 类开发过的镶嵌地包括农田,城市和建筑用地,农田和自然植被镶嵌体。非植被土地覆盖单元包括冰雪,荒地,水体^[13]。

为了与模型研究者当前和将来的系统相兼容,IGBP 的分类可以被重新标记。IGBP 的类可以被转变为其他 3 个方案中的土地覆盖类型。SiB2 是一个用于 GCM 的表面大气交互模型;Running 和 Nemani 分类的主要意图是全球碳和营养循环模型的参量化,Myneni 的分类被用于 MODIS 和 MIS LAI/FPAR 产品的辐射转换模型。这些方案中几乎所有的类,存在一个到多个 IGBP 类与它们等价物的直接映射。当一些类没有等价物时就会产生一个问题。例如,湿地和城区不出现在这 3 个方案中,可能是由于在模型中接收土地覆盖类型输入时不能认可。无论如何,这些类型可能追踪气体发射模型是很重要的。IGBP 的策略是认可那些对 IGBP 的所有模型准则最有用的类,要求模型专家考虑地球陆地表面有意义的面积范围内的所有相关类。

6 生物物理参量化

自然植被和开发过的土地单元可以被用于区分对于生态过程模型所必需的基本覆盖类型的差别。一年生与多年生习性的比较,通过植被是否保持一年的或多年的地上生物量来区分。本属性可以将多年生的植被(森林和木本灌木)与一年生的区别开来,一年生的农作物与牧草都经过不生长的季节,只有种子或地下结构。一年生与多年生的区别容许推论几个植物的关键生理属性。例如,在植物气体交换率的全球合成中,Korner 发现平均的一年生植物比多年生植物保持着高于 50% 的叶面光合能力。生物量与植物高度有关,是气候模型、能量和动力转换公式、表面粗糙度等参量中主要的植被决定因素。

Table 1 Global land cover types defined by IGBP

大类	覆盖类型	特征
自然植被	常绿针叶林	木本植被占优势的土地, 盖度>60%, 高度超过 2 m。几乎所有的树终年保持绿色。冠层从不会出现没有绿色树叶的情况
	常绿阔叶林	木本植被占优势的土地, 盖度>60%, 高度超过 2 m。几乎所有的树和灌木终年保持绿色。冠层从不会出现没有绿色树叶的情况
	落叶针叶林	木本植被占优势的土地, 盖度>60%, 高度超过 2 m。由季节针叶林群落组成, 每年的循环含有叶和无叶期
	落叶阔叶林	木本植被占优势的土地, 盖度>60%, 高度超过 2 m。由季节阔叶林群落组成, 每年的循环含有叶和无叶期
	混合林	树占优势的土地, 盖度>60%, 高度超过 2 m。由树群落组成, 点缀着其他 4 种森林类型的混合或镶嵌。没有那种森林类型超过景观的 60%
	密闭灌丛	低于 2 m 的木本植被和灌木冠层盖度>60%的土地。灌木树叶不是常绿的就是落叶的
	稀疏灌丛	低于 2 m 的木本植被和灌木冠层盖度在 10%~60%之间的土地。灌木树叶不是常绿的就是落叶的
	木本萨瓦那	草本和其他的林下叶层系统的土地, 并且森林冠层盖度在 30%~60%之间。森林高度超过 2 m
	萨瓦那	草本和其他林下叶层系统的土地, 并且森林冠层盖度在 10%~30%之间。森林高度超过 2 m
	草原	草本类型覆盖的土地。树和灌木盖度低于 10%
开发过的镶嵌地	永久湿地	水和草本或木本植被永久混合的土地。植被表现不是盐生的, 就是水生的。
	农田	被农作物覆盖的土地, 有收获和土壤裸露的时期(例如, 单季和多季耕作系统)。多年生的木本作物被分为适当的森林或灌木土地覆盖类型
	城市和建筑用地	被建筑物和其他人造结构覆盖的土地
非植被土地	农田和自然植被镶嵌体	农田, 森林, 灌丛, 和草原镶嵌的土地, 没有那一个成分超过整个景观的 60%
	冰雪	全年冰雪覆盖的土地
	荒地	暴露的土壤、沙子、岩石、或雪的土地, 一年当中的任何时期从不会有超过 10%的植被覆盖
	水体	海洋、湖泊、水库、河流, 不是淡水就是咸水

另一个基本的属性是叶子的寿命, 区分常绿和落叶植物的叶子。这个属性是植被动态碳循环中的关键变量, 对陆地表面季节反照率和能量转化很重要。树叶的寿命定义为是否植物必须每年完全地再生它的整个冠层或只是一部分, 可用于生态系统碳划分、落叶动态和土壤碳池的直接结果。Reich 认为冠导率、最大光合率与叶寿命是部分相反的。因此, 冠层气体交换量的全球属性可能是基于树叶的寿命标准推断的。

IGBP 单位中的第三个植被属性是优势植被覆盖的树叶类型或形状。各种类别中区分了 3 种叶子形状: 针叶、阔叶、禾草状。该属性也与生物地理化学模型中的关键生态参数密切相关。Running 和 Hunt 给针叶树、阔叶树、草分别定义了

最大叶面积指数为 1、6 和 3, 最大冠导率为 1.6、2.5 和 5.0 $\text{mm} \cdot \text{s}^{-1}$ 。

其他固有的生物物理参数可以从这些单位推断。例如, Dorman 和 Sellers 给一些植被类型分配了一系列的光学特性、生理特性、物理特性, 基于 Matthews 和 Kuchler 的研究成果, 建立了简单的陆地表面—大气交互作用的生物圈模型。在一些不同的应用中, Sellers 设计了从 1°分辨率的按月合成的 NDVI 中得到全球 FPAR、LAI 和冠层绿度的算法。虽然这个方法是基于 NDVI 的, 但他们提出的算法通过植被覆盖类型区分了 NDVI-FPAR-LAI 之间的关系, 利用了与 IGBP 分类相似的主要结构类型。在 MODIS 的 MODIS 15 产品中, 用单独的生物地球物理算法产生 FPAR 和 LAI^[14]。

参 考 文 献

- [1] Martin Jung, Kathrin Henkel, Martin Herold, et al. Remote Sensing of Environment, 2006, 101: 534.
- [2] Cihlar J. Int. J. Remote Sensing, 2000, 21: 1093.
- [3] Yang X, Zhang K, Jia B, et al. Journal of Arid Environments, 2005, 63(2): 517.
- [4] Hansen M C, DeFries R S, Townshend J R G, et al. Earth Interactions, 2003, 7(10): 15.
- [5] Boles S, Xiao X, Liu J, et al. Remote Sensing of Environment, 2004 90: 477.
- [6] Groten S M E, Ocatre R. International Journal of Remote Sensing, 2002, 23: 2797.
- [7] Schmidt H, Karnieli A. Journal of Arid Environments, 2000, 45: 43.
- [8] Peters A J, Eve M D. Environmental Monitoring and Assessment, 1995, 37: 273.
- [9] Zha Y, Gao J. Journal of Arid Environments, 1997, 37(3): 419.

- [10] Okin G S, Roberts D A. *Manual of Remote Sensing*, vol. 4, third ed. New York: Wiley, 2004. 111.
- [11] Weiss J L, Gutzler D S, Coonrod J E A, et al. *Journal of Arid Environments*, 2004, 58: 248.
- [12] Huang S, Siegert F. *Journal of Arid Environments*, 2006, 67: 308.
- [13] YANG Jian-ping, DING Yong-jian, CHEN Ren-sheng(杨建平, 丁永建, 陈仁升). *Acta Geographica Sinica(地理学报)*, 2005, 60(3): 467.
- [14] WANG Li-wen, WEI Ya-xing, NIU Zheng(王莉雯, 卫亚星, 牛 铮). *Spectroscopy and Spectral Analysis(光谱学与光谱分析)*, 2008, 28(12): 2956.

Progress in Research on Land Cover Products of MODIS

WEI Ya-xing, WANG Li-wen

Center for Marine Economic and Sustainable Development, College of Urban and Environment Science, Liaoning Normal University, Dalian 116029, China

Abstract Remote sensing technology has been rapidly developed in recent decades, and has been widely used in ecology and environment field. MODIS is a new data source, and in its many products, land cover product is an important product, and it has often been used in global and regional models. In the present review, the procedure of producing land cover product is clearly discussed, and the feature of the classification based network or decision tree is introduced. The paper emphasized the importance of direction information in classification, detailed introduction of change vector analysis methods and land cover change detection based artificial nerve network, analyzed global 17 land cover types defined by IGBP, and compared with other 3 classification systems.

Keywords Land cover; Remote sensing; MODIS

(Received Sep. 2, 2009; accepted Dec. 6, 2009)