

文章编号: 1001-8166(2009)02-0192-12

基于多源数据融合方法的中国 1 km 土地覆盖分类制图

冉有华, 李 新, 卢 玲

(中国科学院寒区旱区环境与工程研究所, 甘肃 兰州 730000)

摘 要:精确的全球及区域土地覆盖数据是陆地表层过程研究的重要基础。在集成研究兴起和多种数据并存的背景下,利用多源信息融合技术进行大尺度土地覆盖制图具有重要的现实意义。证据理论清楚地表达了由于不确定和不完全信息所带来的对命题认识的“无知”,能够确定相应的假设在目前的认知与知识状态下的确定、不确定和“无知”程度,是多源数据决策融合的重要方法。基于证据理论,将 2000 年中国 1:10 万土地利用数据、中国植被图集(1:100 万)的植被型分类、中国 1:10 万冰川图、中国 1:100 万沼泽湿地图和 MODIS 2001 年土地覆盖产品(MOD12Q1)进行了融合,最终基于最大信任度原则进行决策,产生了新的、IGBP 分类系统的 2000 年中国土地覆盖数据。新的土地覆盖数据在保持了土地利用数据的总体精度的同时,补充了中国植被图中对植被类型及植被季相的信息,更新了中国湿地图,增加了中国冰川图最新信息,使分类系统更加通用。

关 键 词:土地覆盖;土地利用;遥感;证据理论

中图分类号: TP75 **文献标志码:** A

1 引 言

精确的全球及区域土地覆盖数据是陆地表层过程研究的重要基础。土地覆盖的变化将影响生态系统的结构与功能及其他一系列的地球表层过程,如能量交换、水循环、生物地球化学循环和植被生产力等^[1~6]。日益深入的全球地表过程研究需要确定各种地表覆盖类型及其特征参数^[4],许多参数通过土地覆盖类型给定,如 SB2、CoLM 模型的许多的时不变参数(如植被反照率、冠层粗糙度、冠层 0 平面位移等)是依赖土地覆盖类型给定的^[7,8]。

由于土地覆盖数据的重要意义,世界各国和许多国际组织相继运用不同的图像处理技术和数据,如 Landsat 7、AVHRR、SPOT 4 VEGETATION、EOS

计划的 Terra 及 Aqua 星上的中分辨率成像光谱仪(MODIS)数据,开展了以土地利用和土地覆盖为主体的区域、洲际和全球尺度的土地覆盖遥感制度研究。产生了一系列土地覆盖数据,如 IGBP DISCOVER^[9]、美国马里兰大学^[10]、GLC2000^[11]和 MODIS 每季度更新的全球土地覆盖数据产品^[12]。Ran^[13]对现存的几个全球土地覆盖数据中国陆域部分的精度进行了评价比较,认为都不能满足中国陆面过程模拟的需要。

在中国,大量与土地覆盖相关的工作开展了许多,取得了众多有价值的信息,如中国科学院在“八五”、“九五”期间组织开展了全国范围的以遥感手段为主的全国土地资源调查,建立了全国 1:10 万土地利用数据库^[14~18],还有中国 1:100 万植被图^[19],

* 收稿日期: 2008-09-16;修回日期: 2008-10-20

* 基金项目:中国科学院西部行动计划(二期)项目“黑河流域遥感—地面观测同步试验与综合模拟平台建设”(编号:KZCX2-XB2-09);国家自然科学基金面上项目“陆地碳循环遥感与模型模拟的融合方法研究”(编号:40871190);国家自然科学基金“中国西部环境和生态科学重大研究计划”项目“西部生态与环境科学数据中心”(编号:90502010)联合资助。

作者简介:冉有华(1980-),男,甘肃镇原县人,硕士研究生,主要从事遥感、地理信息系统与多源数据融合方法应用研究及科学数据共享与服务。Email: ranyh@lzb.ac.cn

中国 1 10 万冰川图^[20],中国 1 100 万沼泽湿地图^[21]等,与土地覆盖相关的数据资源非常丰富,但这些数据没有一个能完整地反映土地覆盖的全貌,目的各有侧重,如中国的土地利用制图主要面向国家资源调查等。一个新的问题是如何集成中国境内可用的土地覆盖相关数据,应用或发展适合于土地覆盖分类信息融合的多源数据融合方法,产生新的、精度更高的、分类系统通用的中国土地覆盖类型数据。

本文基于证据理论,收集了中国境内的多种土

地覆盖相关数据,经过推理决策,发展了中国全境精度更高的、IGBP 土地覆盖分类系统代表 2000 年土地覆盖状况的中国土地覆盖类型数据。

2 分类系统、输入数据

IGBP 分类系统已经大尺度土地覆盖分类中得到了广泛的应用,本文采用 IGBP 分类体系进行中国土地覆盖分类制图;IGBP 分类系统见表 1。

五类数据被作为证据输入决策系统进行融合,

表 1 IGBP 土地覆盖分类系统定义^[22]
Table 1 The definition of IGBP land cover classification system

编号	名 称	含 义
1	常绿针叶林	指覆盖度 > 60% 和高度超过 2 m, 且常年绿色, 针状叶片的乔木林地
2	常绿阔叶林	指覆盖度 > 60% 和高度超过 2 m, 且常年绿色, 具有较宽的叶片的乔木林地
3	落叶针叶林	指覆盖度 > 60% 和高度超过 2 m, 且有一定的落叶周期, 针状叶片的乔木林地
4	落叶阔叶林	指覆盖度 > 60% 和高度超过 2 m, 且有一定的落叶周期, 具有较宽叶片的乔木林地
5	混交林	指前四种森林类型的镶嵌体, 且每种类型的覆盖度不超过 60%
6	郁闭灌木林	指覆盖度 > 60%, 高度低于 2 m, 常绿或落叶的木本植被用地
7	稀疏灌木林	指覆盖度在 10% ~ 60% 之间, 高度低于 2 m, 常绿或落叶的木本植被用地
8	有林草地	森林覆盖度在 30% ~ 60% 之间, 高度超过 2 m, 和草本植被或其它林下植被系统组成的混合用地类型
9	稀树林地	森林覆盖度在 10% ~ 30% 之间, 高度超过 2 m, 和草本植被或其它林下植被系统组成的混合用地类型
10	草地	由草本植被类型覆盖, 森林和灌木覆盖度小于 10%
11	永久湿地	常年或经常覆盖着水 (淡水、半咸水或咸水) 与草木或木本植被的广阔区域, 是介于陆地和水体之间的过渡带
12	农田	指由农作物覆盖, 包括作物收割后的裸露土地。永久的木本农作物可归类于合适的林地或者灌木覆盖类型
13	城市与建筑用地	被建筑物覆盖的土地类型
14	农业与自然植被镶嵌体	指由农田、乔木、灌木和草地组成的混合用地类型, 且任何一种类型的覆盖度不超过 60%
15	冰川	指常年由积雪或者冰覆盖的土地类型
16	裸地与稀疏植被	指裸地、沙地、岩石, 植被覆盖度不超过 10%
17	水体	海洋、湖泊、水库和河流, 可以是淡水或咸水

包括 MODIS 全球土地覆盖数据 2001 年产品 (MOD12Q1)^[12, 23], 中国植被图集 (1 100 万) 的植被型^[17], 来源于中国 1 10 万土地利用数据库的 2000 年中国 1 km 土地利用分类图, 中国 1 10 万冰川图^[20]和中国 1 100 万沼泽湿地图^[21]。这五类数据各有特点, 信息相互补充。MODIS 全球土地覆盖数据产品提供了分类系统通用的、一致性好的全球土地覆盖分类信息; 中国 1 km 土地利用图来源于中国 1 10 万土地利用数据库, 对其进行合并、矢栅转换 (面积最大法) 而来。中国 1 10 万土地利用数据库基于 Landsat MSS、TM 和 ETM 遥感图像, 主要通过全国各地的相关专家根据对图像光谱、纹理、色调等的认识结合地形图目视解译而成。数据库经过了野外实地考察验证, 精度达到 95%^[15]。但是中国 1 10 万土地利用数据采用的是一个分层的土地覆盖

分类系统, 将全国分为 6 个一级类 (耕地、林地、草地、水域、城乡、工矿、居民用地和未利用土地), 31 个二级类, 分类系统中缺少对植被类型和季节特性的描述, 在陆面过程模拟中难以应用; 而中国 1 100 万植被型分类图正好可以补充土地利用数据缺少的信息, 它是根据半个世纪以来全国各地开展植被调查所积累的丰富的第一手资料, 并利用航空遥感和卫星影像等现代技术所获得的材料以及有关地质学、土壤学和气候学最新的研究成果编制而成, 它详细反映了我国 11 个植被类型组、54 个植被型的 796 个群系和亚群系植被单位的分布状况、水平地带性和垂直地带性分布规律; 中国 1 10 万冰川图和中国 1 100 万沼泽湿地图则提供了精确的冰川和湿地分布信息。

3 方法与数据处理

3.1 D-S证据理论

证据理论^[24, 25]是由 Dempster 和 Shafer 在 20 世纪 60 年代末、70 年代初对 Bayes 理论进行扩展的基础上提出的,已经在各种图像分类问题中得到了广泛的研究^[26~34]。证据理论是把概率论中的事件扩展成了命题,把事件的集合扩展成了命题的集合,并提出了基本概率赋值函数、信任函数和似然函数的概念,建立了命题和集合之间的一一对应关系,从而把命题的不确定性问题转化为集合的不确定性问题^[24, 25]。证据理论清楚地表达了由于不确定和不完全信息所带来的对命题认识的无知,能够确定相应的假设在目前的认知与知识状态下的确定、不确定和无知程度,是一种无参的方法,可以处理任意分布、任意数量数据源、任何尺度的数据^[29]。

证据理论中最基本的概念是辨识框架,记为 Ω 。辨识框架定义为一个互不相容事件的集合^[24, 25]。的选取依赖于我们的知识与认识水平及所知道的和想要知道的。鉴别框架 Ω 的一个划分,即 Ω 的子集。不像 Bayes 理论中是对 Ω 中的元素进行运算,证据理论是对 Ω 的幂集 2^Ω 中的元素进行运算,并且满足下列条件:

$$\begin{cases} m(\phi) = 0 \\ m(A) = 1 \\ A \in 2^\Omega \end{cases}$$

$m(A)$ 称为事件 A 的基本概率赋值,在本文中事件是指某一土地覆盖类型,它表示了对类型 A 的支持程度。

在遥感分类中,辨识框架对应于待分类地物类型的集合。本文中以 IGBP 的土地覆盖类型定义作为辨识框架,IGBP 分类系统将地球表层土地覆盖类型划分为 17 类进行,这 17 个类别组成的集合就是我们所说的辨识框架。

3.2 基本概率赋值即证据生成

D-S 融合规则是通过基本概率赋值来确定各证据对每个命题的支持程度的,确定基本概率赋值就是确定各证据源对所决策命题集的支持程度。常用的基本概率赋值的方法是基本概率赋值函数。本文是对不同来源的土地覆盖分类数据在决策层次上融合,由于各信息源是分类结果,在真实分类结果未知的情况下,几种信息源各类型对目标分类的支持程度不是一个直接的函数关系。因此,对应这种分类数据源与目标分类系统不匹配的融合问题,我们

分类型、分部分考虑,目标类型得到每个数据源某个类型或者多个类型的支持程度由两部分组成:

$$M_{ij} = E_{ij} \cdot C_{ij}$$

其中: i 对应于 5 种信息源, j 对应于 17 种土地覆盖类型; E 代表数据源本身的精度; C 代表 3 种信息源类型与 17 种目标土地覆盖类型在定义上的相关程度。

对于中国植被类型图,本研究通过植被型与目标分类系统各类型的关系定义并在充分考虑国内相关植被学专家所给的定义相关系数的基础上,确定了植被型与目标土地覆盖类型在定义上的相关关系,如表 2。假设植被图中植被型的精度为 95%,最终产生了中国植被图对目标土地覆盖各类型的证据。

对于中国土地利用数据,同样也是通过土地利用数据类型与目标类型的关系定义并在充分考虑国内相关土地利用专家所给的 2 个分类系统定义相关系数的基础上,确定了中国土地资源分类系统与 IGBP 土地覆盖类型在定义上的相关关系。中国土地利用数据的验证精度为 95%^[16],相关系数与精度相乘最终产生了中国土地利用类型对 IGBP 土地覆盖类型的证据,支持程度如表 3 所示。

MODIS 土地覆盖产品 (IGBP 分类) 的基本概率赋值主要来自文献 [13] 的结果,将 7 类的验证结果精度直接分配到 17 类上去,但是经典 D-S 理论的合成规则存在悖论^[35],即当几个信息源对某一类型的支持程度都很小的时候,反而得到这个类型的最终信任度最大,从而违背了人的常识推理结果。因此,我们在不改变各类型精度相对高低的情况下进行了修正,使基本概率赋值不存在特别微波的值。因为分类系统与目标分类系统是一样的,所以相关系数 C 为 1;最终的基本概率赋值如表 4。

冰川图只有一种类型,本文假设对 IGBP 类型 15 的支持程度为 0.95;沼泽湿地图也提供了一种类型的支持,本文假设对 IGBP 类型 11 的支持程度为 0.95。

需要说明的是,不同类型之间相关性的确定在量上可能不精确,但可以区分各数据源的相对重要性,这对最后输出的总信任度可能会有影响,但不影响决策结果。最终的各证据表现为对 IGBP 类型有不同支持程度的多个图层,每一层的属性值为基本概率值。

3.3 证据合成

证据理论提供了一个正交和 (8) 的证据结合规则,假如 $m_1(X)$ 表示信息源 1 的基本概率赋值, $m_2(Y)$ 表示信息源 2 的基本概率赋值,则综合概率

表 2 中国植被类型与 IGBP 土地覆盖类型定义的相关性

Table 2 The correlation between two definitions: Chinese vegetation class and IGBP land cover class																
植被型	1	2	3	4	5	6	7	8	9	10	11	12	14	16		
寒温带和温带山地针叶林	0.5		0.5													
温带针叶林	0.6		0.4													
亚热带针叶林	1.0															
热带针叶林	1.0															
亚热带和热带山地针叶林	1.0															
温带针叶、落叶阔叶混交林					1.0											
亚热带山地针叶、常绿阔叶、落叶阔叶混交林				1.0												
温带落叶阔叶林				1.0												
温带落叶小葉林				1.0												
亚热带落叶阔叶林				1.0												
亚热带常绿、落叶阔叶混交林					1.0											
亚热带常绿阔叶林		1.0														
亚热带硬叶常绿阔叶林		1.0														
热带季雨林		1.0														
热带雨林		1.0														
亚热带和热带竹林及竹丛		1.0														
温带落叶灌丛						0.4	0.6									
亚热带、热带常绿阔叶、落叶阔叶灌丛						1.0										
热带珊瑚岛肉质常绿阔叶灌丛和矮林		0.5				0.5										
亚热带、热带旱生常绿肉质多刺灌丛						1.0										
亚高山落叶阔叶灌丛						1.0										
亚高山革质常绿阔叶灌丛						1.0										
亚高山常绿针叶灌丛	0.2					0.8										
矮半乔木荒漠								0.5							0.5	
灌木荒漠							0.5								0.5	
草原化灌木荒漠							0.3	0.3							0.4	
半灌木、矮半灌木荒漠							0.5								0.5	
多汁盐生矮半灌木荒漠							0.5								0.5	
一年生草本荒漠										0.5					0.5	
垫状矮半灌木高寒荒漠										0.5					0.5	
温带禾草、杂类草草甸草原										1.0						
温带丛生禾草草原										1.0						
温带丛生矮禾草、矮半灌木荒漠草原							0.5			0.5						
禾草、薹草高寒草原										1.0						
温带草丛						0.2				0.8						
亚热带、热带草丛						0.3				0.7						
禾草、杂类草草甸										1.0						
禾草、薹草及杂类草沼泽化草甸										1.0						
禾草、杂类草盐生草甸										1.0						
嵩草、杂类草高寒草甸										1.0						
寒温带、温带沼泽											1.0					
亚热带、热带沼泽											1.0					
热带红树林						1.0										
高寒沼泽	1.0															
高山苔原							0.5			0.5						
高山垫状植被										1.0						
高山稀疏植被							0.5			0.5						
一年一熟短生育期耐寒作物田														1.0		
一年一熟粮食作物及耐寒经济作物田														1.0		
一年一熟粮食作物及耐寒经济作物田、落叶果树园				0.3									0.4	0.3		
两年三熟或一年两熟旱作田和落叶果树园				0.3									0.4	0.3		
一年两熟粮食作物田及常绿和落叶果树园与经济林		0.1		0.1									0.4	0.4		
一年两熟或三熟粮食作物田及常绿果树园、亚热带经济林		0.3											0.4	0.3		
一年三熟粮食作物田及常绿果树园和经济林		0.3											0.4	0.3		
无植被地段															1.0	

表 3 中国土地利用类型针对 IGBP 土地覆盖类型的基本概率值
Table 3 The basic probability value for Chinese land use map

土地利用类型	1	2	3	4	5	6	7	8	9	10	11	12	13	15	16	17
水田												0.95				
旱地												0.95				
有林地	0.57	0.57	0.57	0.57	0.57											
灌木林						0.6	0.4									
疏林地							0.5	0.5								
其他林地	0.5	0.5	0.5	0.5	0.5											
高覆盖度草地										0.95						
中覆盖度草地										0.95						
低覆盖度草地										0.8					0.2	
河渠																0.95
湖泊																0.95
水库坑塘																0.95
永久性冰川雪地													0.95			
滩涂																0.95
滩地																0.95
城镇用地													0.95			
农村居民点													0.95			
其他建设用地													0.95			
沙地																0.95
戈壁																0.95
盐碱地																0.95
沼泽地											0.95					
裸土地																0.95
裸岩石砾地																0.95
其他										0.37						0.6

表 4 MODIS 土地覆盖类型基本概率值^[13]

Table 4 The basic probability value for MODIS land cover product

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
0.5	0.5	0.5	0.5	0.5	0.45	0.45	0.5	0.45	0.45	0.1	0.54	0.1	0.4	0.6	0.6	0.6

赋值可用下式计算：

$$m_1 \otimes m_2(Z) = \frac{m_1(X)m_2(Y)}{1 - k}$$

其中，X 和 Y 分别代表了证据理论中定义的辨识框架的子集，在本文中指某一土地覆盖类型。k 是一个归一化常数：

$$k = \sum_{X \cap Y = \phi} m_1(X)m_2(Y)$$

本研究采用了 Idrisi 软件^[36] belief 模块进行了各证据的合成。

3.4 分类决策

为了确定每一个像元的土地覆盖类型，必须根据证据合成的结果进行决策。在证据理论中，有 3 个指标可做决策依据，包括信任度、似真度和信任

区间，信任函数 Bel(A) 表达了对命题的信任度，定义为：对于任意的 A ⊆：

$$Bel(A) = \sum_{B \subseteq A} m(B)$$

用似真函数表达不反对 X (或不反驳 X) 的程度。对于任意的 A ⊆，似真函数定义为：

$$Pl(A) = 1 - Bel(\bar{A})$$

信任函数和似真函数组成信任区间 [Bel(A), Pl(A)]，也叫不确定区间，它描述了 A 的不确定性，不确定性区间越大，决策结果的不确定性也大。

在 Idrisi 软件的 belief 模块中，对各证据合成后，可输出 IGBP 各类型的总的信任度、似真度和置信区间。本研究中，选择最大信任度作为决策规则，对输出的各类型的信任度进行比较，取信任度最大的类型，最终得到融合结果。

4 结果与不确定性分析

4.1 融合结果

根据最大信任原则最终综合得到了代表中国 2000 年土地覆盖信息的土地覆盖类型图,如图 1 所示。该图主要融合了《中国植被图集》(1:100 万)中对于植被的基本外貌(乔木、灌木和草本)、叶的属性(常绿和落叶)、叶型(阔叶和针叶)信息,中国土地利用图中水体、城镇和建筑用地、裸地等信息,

中国冰川图的冰川分布和中国沼泽湿地图的湿地分布信息。

4.1.1 相对于中国土地利用图的改进

为了说明土地覆盖融合结果与中国土地利用图在融合后分类精度方面的变化,我们将两种数据的分类系统统一综合到一个 7 类的体系下,通过面积和空间一致性两方面进行分析,新的分类综合方案如表 5 所示。

图 2 显示了融合结果与中国 2000 年土地利用

表 5 一个综合的 7 个类型的分类系统 (镶嵌体根据重分类后的参考数据进行了分配)
Table 5 The common classification system include seven aggregated classes
(The mosaic type is reclassified by reference data aggregated)

类型名称	融合结果	中国土地利用图
乔木林地	常绿针叶林、常绿阔叶林、落叶针叶林、落叶阔叶林、混交林、稀疏林地	有林地、疏林地、其他林地
灌木草地	郁闭灌木林、稀疏灌木林、草地、有林草地	灌木林、高覆盖度草地、中覆盖度草地、低覆盖度草地
农田	耕地	水田、旱地
裸地冰川	冰川、裸地与稀疏植被	永久性冰川雪地、沙地、戈壁、盐碱地、裸土地、裸岩石砾地、其他
建筑用地	城市和建筑用地	城镇用地、农村居民点、其他建设用地
湿地	永久湿地	沼泽地、滩涂
水体	水体	河渠、湖泊、水库坑塘、滩地

图在 7 类的水平上各类型总面积的差异情况。从图 2 可以看出,在 7 类水平上,融合了多种信息的融合结果与原始的土地利用图在各类型的总面积上差异不大,基本保持一致,特别是城市、湿地和水体类型,保持了与中国土地利用图在 7 类水平上类型面积的总体一致性,说明融合过程是合理的。对于空间一致,通过建立融合结果与原始的土地利用图的误差矩阵来评价,如表 6 所示。从表 6 可以看出,农田、城市、裸地冰川和水体类型的一致性都达到 95% 以上,其次是草地灌木和湿地类型,都达到 88% 以上,林地类型也接近 70%。总体一致性达到 88.84%, Kappa 系数为 0.85,因为中国土地利用数据的验证精度达到 95%^[16],说明融合结果在很大程度上保持了中国土地利用图 7 类水平上分类精度的同时,显著提高了农田、城市、裸地冰川和水体类型的精度达到 90% (95% × 95%),湿地和灌木草地的精度也达到 80% (88% × 95%) 以上。

总之,融合结果在保持了中国土地利用图总体分类精度的同时,在分类系统中增加了常绿针叶林、常绿阔叶林、落叶针叶林、落叶阔叶林等几种类型,

采用国际上较通用的 IGBP 土地覆盖分类系统。

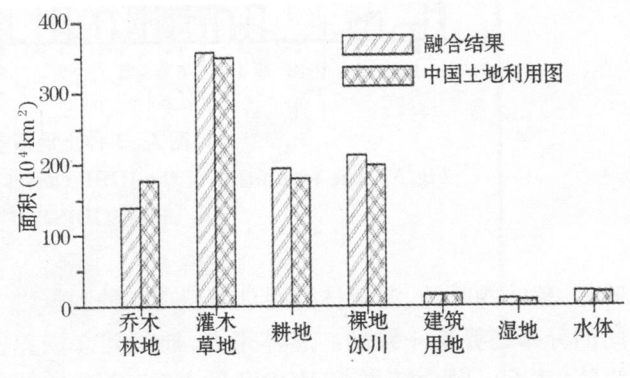


图 2 融合结果与中国土地利用图在 7 类水平上总面积的比较

Fig 2 The total areas of the seven aggregated classes in China from the result map and Chinese land use map respectively

4.1.2 相对于常用的全球土地覆盖数据的改进

IGBP DISCover^[9]与 MODIS 土地覆盖产品^[12,23]是两种常用的、分类系统采用 IGBP 分类方案的全球土地覆盖数据。本文的融合工作希望通过融合中

国本土多年积累的土地覆盖相关调查数据成果提高中国区域土地覆盖图的精度,为了验证此目标,我们从面积和空间分布两个方面与 IGBP DISCover 和 MODIS 的 2001 年土地覆盖产品进行比较。

对于面积,图 3 显示了融合结果、IGBP DISCover 和 MODIS 的 2001 年土地覆盖产品各类型总面积的

比较结果。3 种数据在面积分布上基本一致,除了单个类型的农田、城市、裸地、冰川和水体类型的精度前面已经证明有较高的精度外,综合类型灌木草地的总体精度虽然较高,但对于郁闭灌木林、稀疏灌木林、草地、有林草地四种单个类型的面积在 3 种数据之间差异较大。常绿针叶林、常绿阔叶林、落叶针

表 6 融合结果与中国土地利用图在 7 类水平上的误差矩阵

Table 6 The error matrix of result map with Chinese land use map in seven classes level

	乔木林地	灌木草地	农田	裸地冰川	建筑用地	湿地	水体
乔木林地	69.16	4.43	0.53	0.05	0.22	0.59	0.43
灌木草地	23.76	88.28	0	3.23	0	0	0
农田	5.58	1.3	99.25	0.07	2.88	3.62	1.32
裸地冰川	0.23	5.74	0.01	96.53	0.18	0.69	0.33
建筑用地	0	0	0	0	96.34	0	0
湿地	0.9	0.12	0.07	0.06	0.1	88.13	1.14
水体	0.37	0.13	0.14	0.05	0.27	6.97	96.78
总体精度 = 88.84% Kappa系数 = 0.8507							

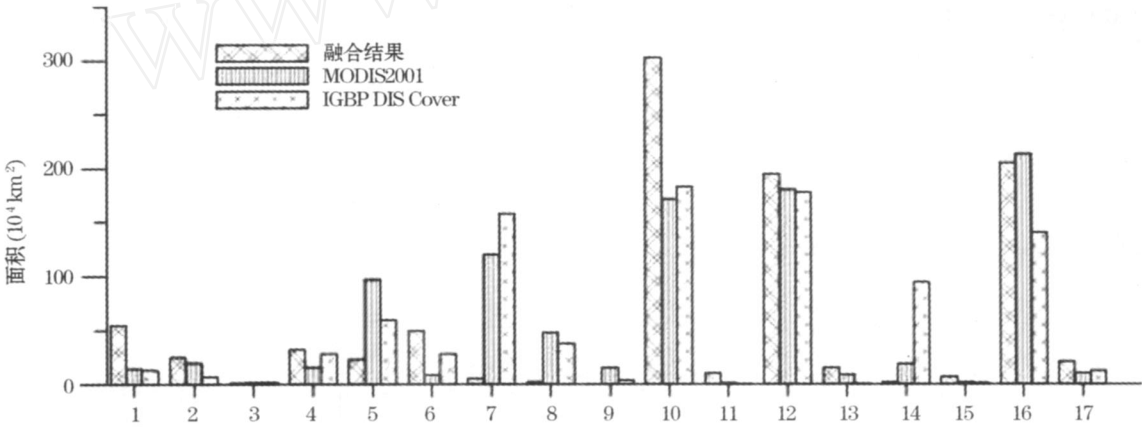


图 3 3 种土地覆盖数据各类型面积比较

Fig 3 The total areas of the IGBP classes in China from the result map, IGBP DISCover and MODIS2001 products respectively

叶林、落叶阔叶林、混交林类型总面积在 3 种数据之间的分布趋势是一致的。总体来看,通过建立误差矩阵(表 6),融合结果与 MODIS 的 2001 年土地覆盖产品的一致性达到 47.59%,Kappa 系数为 0.38;融合结果与 IGBP DISCover 的一致性达到 40.2%,Kappa 系数为 0.31。

为了直接表现 3 种数据空间特征的差异,我们在全中国选择 3 个区域,分别位于中国第二大内陆河流域黑河流域中游、齐齐哈尔地区和安徽省与湖北省交界地区,在视觉上对融合结果、IGBP DISCover 和 MODIS2001 进行比较,如图 4 所示。图 4 显示,3 种土地覆盖数据都反映出了局部土地覆盖类型的分布特征,但各反映的局部特征细节有较大差异,其

中 WESTIDCover 反映的细节特征最为详细。如对于区域 A,3 种数据都反映出了黑河流域的植被分布的形态,区域中间的金塔绿洲轮廓明显,但在 IGBP DISCover 数据中表现为大部分为草地,有少量农田,MODIS 产品类似,而融合结果则清楚的反映了黑河流域几个绿洲的分布类型,如金塔绿洲的农田、黑河河道等;区域 B 中,IGBP DISCover 数据反映了大片林地分布,而根据中国林业调查数据(<http://www.cfsdc.org>),该区并没有大片林地分布,融合结果明显反映了嫩江河道和湿地分布信息;区域 C 类似。由此进一步说明,融合了多源信息的中国土地覆盖数据在保持了中国土地利用图的总体精度,增加了中国植被图叶属性和叶型信息的同时,所反映的局

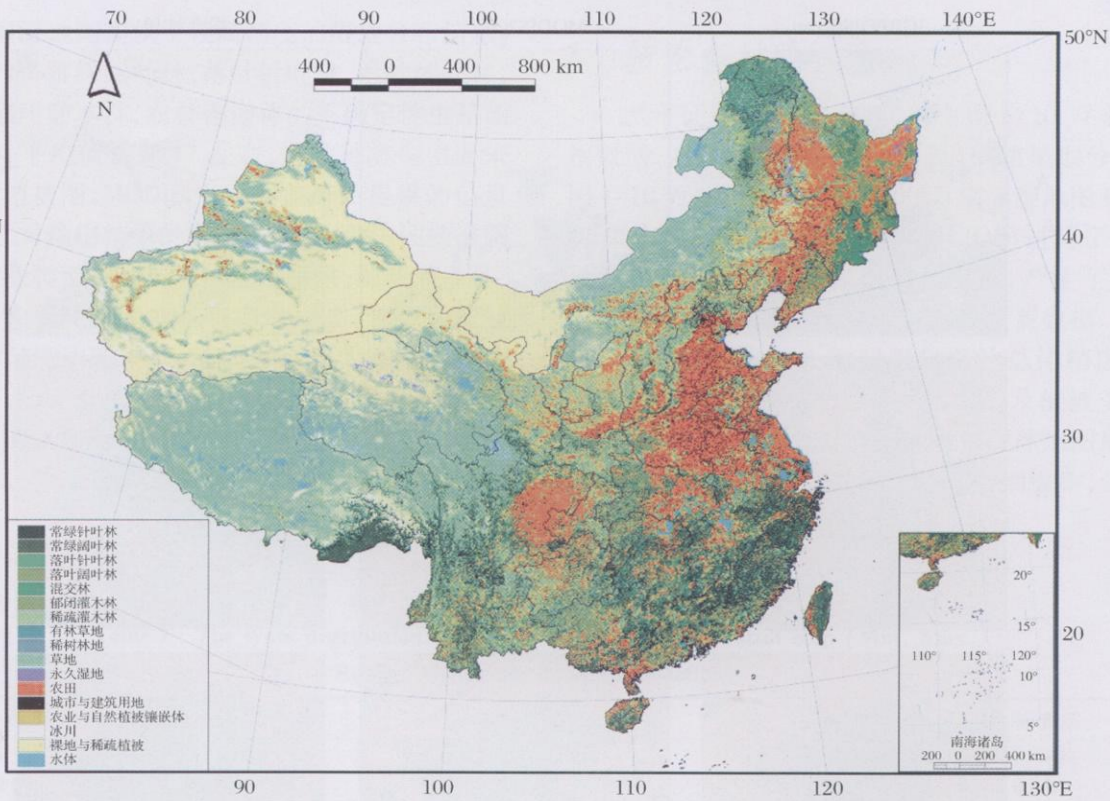


图 1 基于证据理论的中国土地覆盖分类融合结果

Fig. 1 The Chinese land cover classification map fused the multi-source data

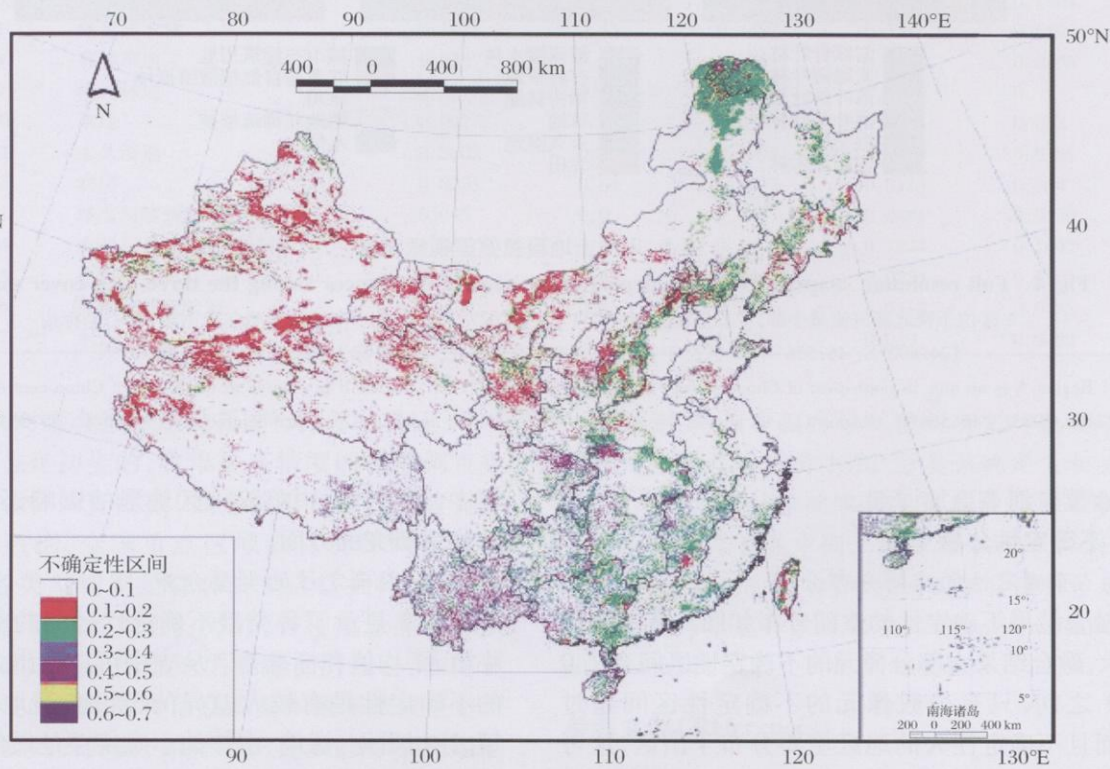


图 5 融合结果不确定性区间的空间分布

Fig. 5 The spatial distribution of interval of uncertainty of result map

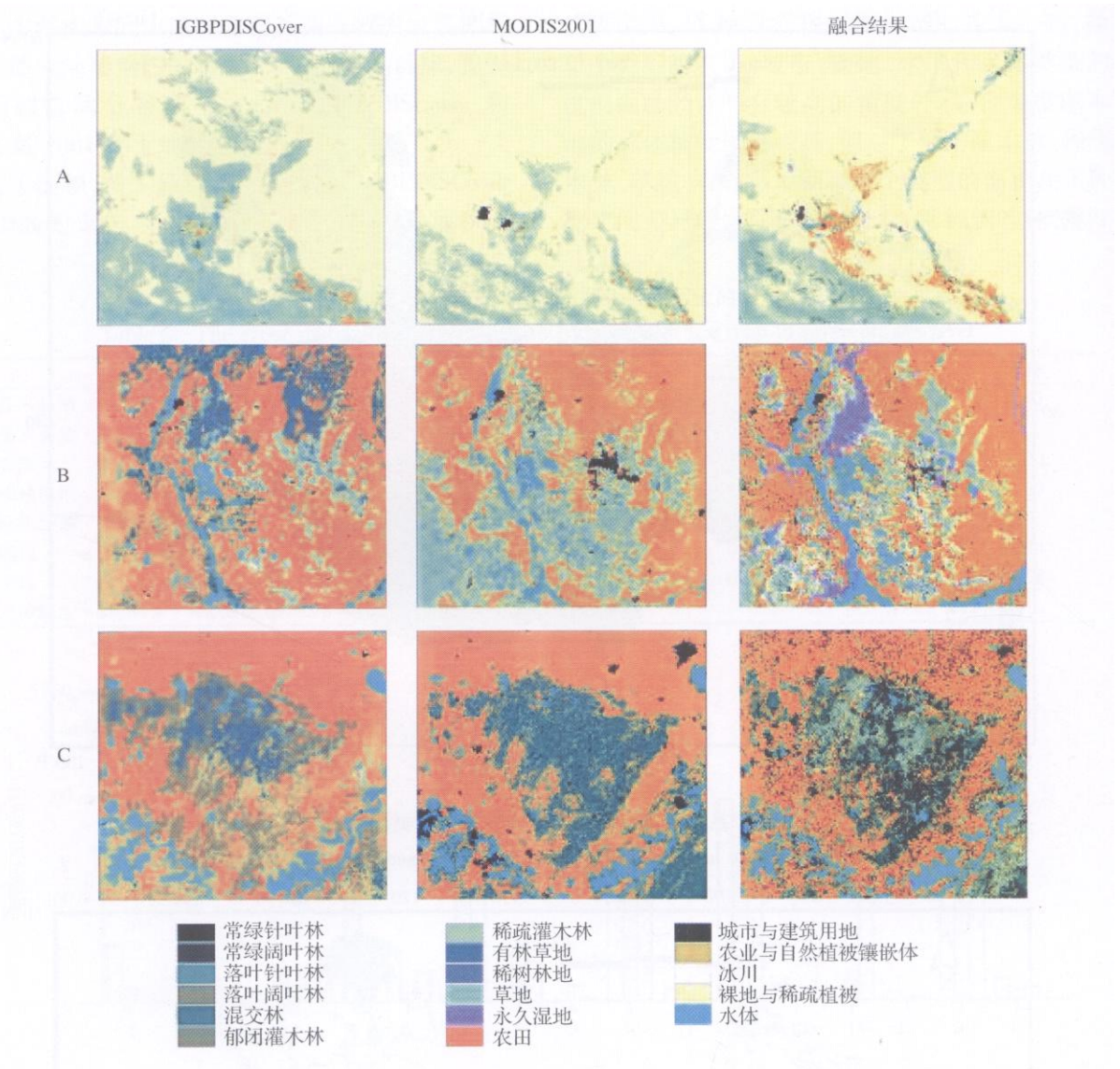


图 4 3种土地覆盖数据视觉对比

Fig 4 Full resolution snapshots of areas depicting the general differences among the three land cover maps

A 区位于西北黑河流域中部,中心坐标为 99. 1234 °E, 40. 077 °N; B 区位于东北齐齐哈尔市附近,中心坐标为 124. 602 °E, 46. 556 °N; C 区位于安徽省与湖北省交界处,中心坐标为 115. 916 °E, 30. 9643 °N
Region A is an area in north-west of China centred at 99. 1234 °E, 40. 077 °N. Region B is an area in north-east of China centred at 124. 602 °E, 46. 556 °N. Region C is an area in the junction between Anhui and Hubei province centred at 115. 916 °E, 30. 9643 °N

部土地覆盖细节更加详细。

4 2 不确定性分析

4. 2. 1 不确定性的空间分布分析

融合结果不确定性的空间分布如图 5 所示。图 5 显示,融合结果大部分像元的不确定性区间介于 0 ~0. 1 之间,只有少数像元的不确定性区间超过 0. 4,而且不确定性大的地区主要分布于山区,这可能说明了山区复杂的气候环境所造成复杂的土地覆盖类型,同时也反映了基于遥感的土地覆盖分类过

程中,遥感数据中消除山区地形的影响还是一个有待加强研究的方面。

4. 2. 2 不确定性的类型分布

表 7 显示了各类型不确定性区间的最大值、最小值、平均值和标准差。从表中可以看出,类型之间的不确定性是有较大差异的,草地、湿地、耕地、城镇、冰川雪地、裸地、水体的不确定性是最低的,都低于 0. 05,这可能是因为中国土地利用图提供了较高精度的草地、耕地、城镇、冰川雪地、裸地和水体分布

信息,中国湿地图提供了较高精度的湿地分布信息;常绿针叶林,常绿阔叶林,落叶针叶林,落叶阔叶林,混交林,郁闭灌木林,农林镶嵌体的不确定性也是相对较低的,平均值在 0.17 左右,这几种类型主要来源于中国植被图,MODIS 土地覆盖产品提供的信息权重较低,对输出结果的影响不大,因此中国植被图提供了这些类型较高精度的输出结果;稀疏灌木林、有林草地和稀疏林地的不确定性较高,其中稀疏林地的不确定性达到 0.55,这可能是因为在融合系统的输入中,没有支持这几种类型的直接证据,导致支持该类型的证据有较大的不确定性,从而导致结果的不确定。

5 结论与讨论

基于证据理论,将 2000 年中国 1:10 万土地利用数据、中国植被图集(1:100 万)的植被型分类、中国 1:10 万冰川图、中国 1:100 万沼泽湿地图和 MODIS2001 年土地覆盖产品(MOD12Q1)进行了融合,最终基于最大信任度原则进行决策,产生了新的、IGBP 分类系统的 2000 年中国土地覆盖数据。融合结果在保持了中国土地利用数据的总体精度的同时,补充了中国植被图中对植被类型及植被季相的信息,更新了中国湿地图,增加了中国冰川图最新信息,使得中国土地覆盖数据的精度得到提高,分类系统采用 IGBP 分类系统而更加通用。

表 7 不确定性区间在各类型的分布
Table 7 The types distribution of interval of uncertainty of result map (Min, Max, Range, Mean and Standard deviation)

编号	土地覆盖类型	最小	最大	变程	平均值	标准差
1	常绿针叶林	0.0215	0.55	0.5285	0.0955	0.0828
2	常绿阔叶林	0.0215	0.55	0.5285	0.1338	0.124
3	落叶针叶林	0.0989	0.5	0.4011	0.1495	0.0619
4	落叶阔叶林	0.0215	0.5	0.4785	0.0705	0.0702
5	混交林	0.0215	0.5	0.4785	0.1762	0.1298
6	郁闭灌木林	0.0236	0.55	0.5264	0.155	0.1304
7	稀疏灌木林	0.1018	0.55	0.4482	0.3386	0.1195
8	有林草地	0.025	0.55	0.525	0.2512	0.0296
9	稀树林地	0.55	0.55	0	0.55	0
10	草地	0.0027	0.64	0.6373	0.0433	0.068
11	永久湿地	0.0002	0.9	0.8998	0.0353	0.0381
12	农田	0.0023	0.64	0.6377	0.0413	0.064
13	城市与建筑用地	0.045	0.9	0.855	0.0494	0.0139
14	农业与自然植被镶嵌体	0.06	0.6	0.54	0.1844	0.1492
15	冰川	0.001	0.4	0.399	0.0253	0.0303
16	裸地与稀疏植被	0.002	0.4	0.398	0.0313	0.0485
17	水体	0.002	0.4	0.398	0.039	0.0482

但是,对融合结果还没有进行完整的验证,对结果的分析也是初步的,收集更高精度的植被调查数据特别是森林调查数据(如全国第六次连续清查森林分布图数据,全国重点区域森林资源调查数据等)针对各类型特别是森林类型进行抽样验证,是需要进一步研究的内容。另外,各种信息源基本概率赋值的确定是证据理论决策融合的核心,如何更好的更科学的确定各种信息源基本概率赋值也是需要进一步研究的方面,如可组织更多的专家利用层次分析法(AHP)进行各种信息源基本概率赋值的确定,可能降低证据生成的不确定性。

致谢:感谢中国科学院植物研究所倪健研究员对本文中国植被图与 IGBP 分类系统定义相关性的指导,感谢中国科学院寒区旱区环境与工程研究所王建华副研究员在中国土地资源分类与 IGBP 分类定义相关性的指导。同时感谢国家自然科学基金委员会环境与生态科学数据中心(<http://westdcwestgis.ac.cn>)提供本文所使用的数据。

参考文献 (References):

[1] Turner B L II, Skole D, Sanderson S, et al Land-use and land-cover change, in Science/Research Plan[R]. HDP Report 7/IGBP Report 35, Stockholm and Geneva, 1995.

- [2] Crutzen P J, Andreae M O. Biomass burning in the tropics: impact on atmospheric chemistry and biogeochemical cycles[J]. *Science*, 1990, 250(4 988): 1 669-1 678.
- [3] Keller M, Jacob D J, Wofsy S C, *et al*. Effects of tropical deforestation on global and regional atmospheric chemistry[J]. *Climatic Change*, 1991, 19: 139-158.
- [4] Hederson-Sellers A, Wilson M F. Surface albedo data for climatic modeling[J]. *Review of Geophysics and Space Physics*, 1983, 21: 1 743-1 778.
- [5] Li Xiaobing. International research on environmental consequence of land use/cover change[J]. *Advances in Earth Science*, 1999, 14(4): 395-400. [李晓兵. 国际土地利用—土地覆盖变化的环境影响研究[J]. *地球科学进展*, 1999, 14(4): 395-400.]
- [6] Li Xiaobing, Chen Yunhao, Yu Feng. Global and regional cover mapping from remote sensing data: Status, strategies and trends[J]. *Advances in Earth Science*, 2004, 19(1): 71-80. [李晓兵, 陈云浩, 喻锋. 基于遥感数据的全球及区域土地覆盖制图——现状、战略和趋势[J]. *地球科学进展*, 2004, 19(1): 71-80.]
- [7] Yongjiu Dai, Xubin Zeng, Robert E, *et al*. Common Land Model (CLM) (Technical Documentation and User's Guide) [Z]. 2001.
- [8] Sellers P J, Randall D A, Collatz G J, *et al*. A revised land surface parameterization (SB2) for atmospheric GCMs Part I: Model formulation[J]. *Journal of Climate*, 1996, 9(4): 676-705.
- [9] 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): 1 303-1 330.
- [10] Hansen M C, Defries R S, Townshend J R G, *et al*. Global land cover classification at 1 km spatial resolution using a classification tree approach[J]. *International Journal of Remote Sensing*, 2000, 21(6/7): 1 331-1 364.
- [11] 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): 1 959-1 977.
- [12] 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): 287-302.
- [13] Ran Y, Li X, Lu L. Evaluation of four remote sensing based land cover products over China[J]. *International Journal of Remote Sensing*, 2008 (in press).
- [14] Liu J, Liu M, Deng X, *et al*. The land use and land cover change database and its relative studies in China[J]. *Journal of Geographical Sciences*, 2002, 12(3): 275-282.
- [15] Liu J, Liu M L, Tian H Q, *et al*. Spatial and temporal patterns of China's cropland during 1990-2000: An analysis based on Landsat TM data[J]. *Remote Sensing of Environment*, 2005, 98(4): 442-456.
- [16] Liu Jiyuan. Macro-Scale Survey and Dynamic Study of Natural Resources and Environment of Chinese by Remote Sensing[M]. Beijing: Chinese Science and Technology Press, 1996. [刘纪远. 中国资源环境遥感宏观调查与动态研究[M]. 北京: 中国科学技术出版社, 1996.]
- [17] Liu Jiyuan. Study on national resources and environment survey and dynamic monitoring using remote sensing[J]. *Journal of remote sensing*, 1997, 1(3): 225-230. [刘纪远. 国家资源环境遥感宏观调查与动态监测研究[J]. *遥感学报*, 1997, 1(3): 225-230.]
- [18] Liu Jiyuan, Zhang Zengxiang, Zhuang Dafang, *et al*. A study on the spatial-temporal dynamic changes of land-use and driving forces analyses of China in the 1990s[J]. *Geographical Research*, 2003, 22(1): 1-12. [刘纪远, 张增祥, 庄大方, 等. 20世纪90年代中国土地利用变化时空特征及其成因分析[J]. *地理研究*, 2003, 22(1): 1-12.]
- [19] Hou Xueyu, ed. Vegetation Map (1:1 000 000) in China[M]. Beijing: Science Press, 2001. [侯学煜主编, 1:1 000 000中国植被图集[M]. 北京: 科学出版社, 2001.]
- [20] Wu Lizong, Li Xin. China Glacier Information System [M/CD]. Beijing: Ocean Press, 2004. [吴立宗, 李新. 中国冰川信息系统(配光盘)[M/CD]. 北京: 海洋出版社, 2004.]
- [21] Zhang Shuqing. An introduction of wetland science database in China[J]. *Scientia Geographica Sinica*, 2002, 22(2): 188-189. [张树清. 中国湿地科学数据库简介[J]. *地理科学*, 2002, 22(2): 188-189.]
- [22] Hansen M C, Reed B. A comparison of the IGBP DISCover and University of Maryland 1 km global land covers products[J]. *International Journal of Remote Sensing*, 2000, 21(6&7): 1 365-1 373.
- [23] Hodges J C F, Friedl M A, Strahler A H. The MODIS global land cover product: New data sets for global land surface parameterization[C]. Proceedings of the Global Change Open Science Conference, Amsterdam, 2001.
- [24] Dempster A P. Upper and lower probabilities induced by multi-valued mappings[J]. *Annals of Mathematical Statistics*, 1967, 38: 325-329.
- [25] Shafer G. A Mathematical Theory of Evidence[M]. Princeton, NJ: Princeton University Press, 1976.
- [26] Peddle D R. MERCURY: An evidential reasoning image classifier[J]. *Computers and Geosciences*, 1995, 21(10): 1 163-1 176.
- [27] Soh L, Tsatsoulis C, Gineris D. ARKTOS: An intelligent system for SAR sea ice image classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2004, 42(1): 229-247.
- [28] Comber A J, Law A N R, Lishman J R. A comparison of Bayes, Dempster-Shafer and Endorsement theories for managing knowledge uncertainty in the context of land cover monitoring[J]. *Computers, Environment and Urban Systems*, 2004, 28(4): 311-327.
- [29] Sun W, Liang S, Xu G, *et al*. Mapping plant functional types from MODIS data using multisource evidential reasoning[J]. *Remote Sensing of Environment*, 2008, 112(3): 1 010-1 024.
- [30] Liu Chunping. Studies on the Methods and Application of Multi-Sources Information Fusion in Remote Sensing[D]. Nanjing: Nanjing University of Science and Technology, 2002. [刘纯平.

- 多源遥感信息融合方法及其应用研究 [D]. 南京: 南京理工大学, 2002.]
- [31] Zhu Xiaokun. Application of Fusion Model of Dempster-Shafer Evidence Reasoning in Remote Sensing[D]. Wuhan: Wuhan University, 2005. [祝晓坤. Dempster-Shafer 证据推理融合模型在遥感分类中的应用 [D]. 武汉: 武汉大学, 2005.]
- [32] Peng Tianqiang, Li Bicheng. Remote sensing image classification method based on evidence theory and neural networks[J]. *Journal of Data Acquisition & Processing*, 2003, 18(2): 170-174. [彭天强, 李弼程. 一种基于证据理论与神经网络的遥感影像分类方法 [J]. 数据采集与处理, 2003, 18(2): 170-174.]
- [33] Wang Xuhong, Zhou Mingquan, Geng Guohua. Reseach of the application of the theory of dempster_Shafer in intelligent remote sense classification [J]. *Computer Applications and Software*, 2004, 21(9): 28-29. [王旭红, 周明全, 耿国华. "Dempster-Shafer 证据理论在智能化遥感分类中应用研究 [J]. 计算机应用与软件, 2004, 21(9): 28-29.]
- [34] Wang Yang, Zheng Qinbo, Zhang Junping. Target classification of the data fusion in multichannel using Dempster-Shafer method [J]. *Journal of Infrared and Millimeter Waves*, 2002, 21(3): 229-232 [汪洋, 郑亲波, 张钧屏. 用证据理论方法进行多波段数据融合的目标分类 [J]. 红外与毫米波学报, 2002, 21(3): 229-232.]
- [35] Zadeh L A. Fuzzy sets as a basis for a theory of possibility[J]. *Fuzzy Sets and Systems*, 1978, 1(1): 3-28.
- [36] Ronald Eastman J. *DRISI Kilimanjaro Guide to GIS and Image Processing*[Z]. Clark University, 2003.

China Land Cover Classification at 1 km Spatial Resolution Based on a Multi-source Data Fusion Approach

RAN Youhua, LI Xin, LU Ling

(Cold and Arid Regions Environmental and Engineering Research Institute,
Chinese Academy of Sciences, Lanzhou 730000, China)

Abstract: Land cover plays a significant role in the earth system science, which reflects the influence of human activities and environmental changes. In China, many land use/cover maps can be used in recent years. How to combine these data effectively to better produce land cover map is a key question. This paper presents a method to map land cover classification using Dempster-Shafer evidential reasoning (DS). DS theory is a non-parametric classifier and can handle data from any number of sources at any scale of measurement and has an explicit mechanism for handling information uncertainty through the use of the concept of ignorance. The 1km land use data of China in 2000 from China landuse database, the 1:1000000 vegetation map, the 1km glacier map from 1:1000000 glacier map, the 1:1000000 wetland map and the MODIS land cover classification product (MOD12Q1) as multi-source of evidence to support each IGBP land cover class. These evidences are then combined using Dempster's Rule of combination. Finally, a new land cover map with higher accuracy is generated by a decision rule based on maximum belief function and could be representative of the Chinese land-cover status in 2000. A simple compare validation is taken and result shows that the accuracy of each land cover class of new land cover map has been strongly promoted relative to IGBP DISCover and MOD12Q1 land cover map especially for the accuracy of the crop lands, urban, glacier, wetland and water bodies class. The evidence theory can be used for fusing multi-classification information. The fusion result is sensitive for the basic probability assignment of input data. The results will be validated further and the basic probability for multi-source input data will be assigned more scientifically.

Key words: Land cover; Land use; Dempster-Shafer theory; Remote sensing