

引用格式:孙云华,郭涛,崔希民,等.基于行为聚类算法的土地利用聚类适宜性分析研究[J].地球信息科学学报,2016,18(3):396-405. [Sun Y H, Guo T, Cui X M, *et al.* 2016. Suitability analysis on behavior-based aggregation of land use classification in Yunnan Province. Journal of Geo-information Science, 18(3):396-405.] DOI:10.3724/SPJ.1047.2016.00396

基于行为聚类算法的土地利用聚类适宜性分析研究

孙云华^{1,2}, 郭涛^{3*}, 崔希民¹, 崔伟宏²

1. 中国矿业大学(北京)地球科学与测绘工程学院, 北京 100083; 2. 中国科学院遥感与数字地球研究所, 北京 100101;
3. 湖南科技大学计算机科学与工程学院, 湘潭 411201

Suitability Analysis on Behavior-based Aggregation of Land Use Classification in Yunnan Province

SUN Yunhua^{1,2}, GUO Tao^{3*}, CUI Ximin¹ and CUI Weihong²

1. College of Geoscience and Surveying Engineering, China University of Mining & Technology, Beijing 100083, China; 2. Institute of Remote Sensing and Digital Earth, CAS, Beijing 100101, China; 3. School of Computer Science and Technology, Hunan University of Science and Technology, Xiangtan 411201, China

Abstract: Comparison of the same study area between two time points on the same categorical variable can reveal changes occurred among categories over time, such as transitions among land categories. Therefore, it is an effective method to aggregate a large number of categories into a smaller number of broader categories in order to simplify interpretation and give insights concerning categorical change over time. We use an algorithm to aggregate categories in a sequence of steps based on the categories' behaviors in terms of gross losses and gross gains. The behavior-based algorithm aggregates net gaining categories with net gaining categories, and aggregates net losing categories with net losing categories, but doesn't aggregate a net gaining category with a net losing category. The algorithm's steps are summarized into three phases. The first phase is the Exclusive Zero phase, when our algorithm aggregates pairs of Exclusive Zero categories until all Exclusive Zero categories are aggregated into one comprehensive Exclusive Zeroes category, which remains as one of the three categories at the end of the algorithm. The second phase is the Exclusive Loser and Gainer phase, when the algorithm aggregates pairs of Exclusive Gainer categories and pairs of Exclusive Loser categories. The third phase is the swapping phase. In this final phase, our algorithm computes the transition sum for every possible pair of categories that have the same direction of net change. Moreover, the behavior-based algorithm at each step in the sequence keeps the values of net change and maximizes swap change. This article introduces three mathematical principles and defines six aggregation behaviors. In the last section of this article, we present a case study, in which the data are obtained in Yunnan Province between 1990 and 2010 for 20 land categories, in order to prove the advantage and feasibility of this algorithm in terms of area change and intensity change. The results show that the behavior-based algorithm produces a set of 9 categories that retains almost the original amount of change, while giving a total area change of 5.10%, which is only 0.06% lower than the original change. In contrast, the common used classification system of 6 categories produces a total area change of 4.7%, which is 0.4% lower than the results of our algorithm. It has been found that the behavior-based algorithm is an effective method to retain land use dynamic change information.

Key words: behavior-based aggregation; land use; classification; net change; swap change; Yunnan Province

***Corresponding author:** GUO Tao, E-mail: guotao0628@outlook.com

摘要: 为了在较少地类的基础上,深入研究土地利用变化过程,可把较多的地类合并成较少的具有重要变化特征的地类。本

收稿日期 2015-05-30;修回日期:2015-07-07.

基金项目:国家自然科学基金项目(51474217,71150001)。

作者简介:孙云华(1985-),女,山东临清人,博士生,主要从事土地利用变化研究。E-mail: yunhua07@163.com

*通讯作者:郭涛(1990-),男,甘肃平凉人,硕士生,主要从事土地利用/植被覆盖模拟以及土地利用系统开发集成研究。
E-mail: guotao0628@outlook.com

文运用基于行为聚类的方法,将净增加和净增加的地类合并或净减少和净减少的地类合并,但不能将净增加和净减少的地类合并。主要分为3个阶段实现:第1个阶段是完全不变阶段,聚类结果是形成综合的完全不变地类;第2个阶段是完全减少和完全增加阶段,该阶段将完全减少地类和完全增加地类分别进行合并;第3个阶段是转变阶段,该阶段计算了每一对有相同净变化方向地类的相互转变信息。基于行为聚类的算法在每一步的计算过程中都会保留净变化面积为常数,转变变化面积保留最大化。本文对3个聚类原则进行了数学证明,定义了6种聚类行为,以云南省土地利用分类体系为例,从面积变化和强度变化2个角度阐述了行为聚类方法的可行性和优势性。最后,与刘纪远等建立的土地利用遥感监测分类系统进行比较,结果表明:行为聚类算法聚类到9种类型时,类型总变化面积达到5.10%,比原始类型的总变化面积下降了0.06%;而基于遥感监测分类系统的6个一级分类将总变化减少至4.7%,与聚类算法比较,总变化面积减少了0.4%。实践证明,行为聚类的算法可更好地保留土地利用动态变化信息,证明了该聚类方法的有效性。

关键词 行为聚类;土地利用;分类系统;净变化;转变变化;云南省

1 引言

土地涵盖了地球陆地表层的一切物质,结构复杂,与周围的环境不同,土地在地域之间存在着差异性,导致其可利用性、生产能力和利用方式不同。将不同类型的土地按照一定的规律方式,建立从属关系,形成了土地利用分类系统。长期以来,为了满足不同的需要,不同部门形成了不同的分类系统,主要包括:(1)全国农业区划委员会制定的土地利用分类系统,包括8个一级类型和46个二级类型^[1];(2)中国科学院结合遥感特点制定的土地利用分类系统^[2],包括6个一级类型和25个二级类型,并对耕地进行细分,增加了8个三级类型;(3)国土资源部制定的土地利用现状分类国家推荐标准^[3],包括12个一级分类和57个二级分类。人们在利用这些分类系统时,均会根据工作需要或区域特点因地制宜地进行必要的调整和完善。

在同一研究区域内,对2期或多期时相的相同地类进行比较,可揭示地类随时间的动态变化。例如,不同地类之间的相互转变,如果地类过多,则难以从整体上把握变化趋势。对地类进行聚类的目的就是2种或更多种地类进行合并得到较少但更具概括性的地类。聚类可简化解译过程,也可影响土地利用变化过程的分析研究。一些基于先验聚类可简化过程,但会忽略某些重要的动态变化。Pontius和Malizia^[4]对土地利用的聚类提出了5个土地利用变化行为原则,从数学角度进行了描述。此外,许多学者认为聚类可影响转移矩阵的结果^[5-7]。Aldwaik^[8]在前人聚类原则的基础上,提出了6种行为聚类方式,但是并不一定适合国内土地利用系统的分类。因此,本文借鉴Aldwaik的基本理论,将其首次应用于国内常用土地利用分类系统的聚类过程,以探索该方法的可用性和普适性。

2 行为聚类算法

2.1 行为聚类的基本公式解释

式(1)~(6)分别定义了 L_k 、 G_k 、 T_{ij} 、 D 、 N 、 S 。式(1)表示在时间间隔内地类 k 的总减少变化面积, L_k 是地类 j 在 t 时刻的总变化面积(括号内的公式)与主对角线保持不变地类面积 C_{kk} 的差值, J 表示地类总数。式(2)表示在时间间隔内地类 k 的总增加变化面积,是地类 i 在 $t+1$ 时刻的总变化面积(括号内的公式)与主对角线保持不变地类面积 C_{kk} 的差值。式(3)的 T_{ij} 表示关联表中除了对角线保持不变的地类面积,即地类 i 和地类 j 之间相互转变的总和; C_{ij} 表示某一时间间隔内由初始时刻的地类 i 转变为末期时刻地类 j 的面积; C_{ji} 表示某一时间间隔内由初始时刻的地类 j 转变为末期时刻地类 i 的面积。式(4)计算了在时间间隔 $[Y_t, Y_{t+1}]$ 内,所有地类的总变化面积 D ,等于 t 时刻所有减少地类 k 的总面积,也等于在 $t+1$ 时刻所有增加地类 k 的总面积。式(5)定义了净变化 N ,来源于地类 k 的总增加面积 G_k 与总减少面积 L_k 之差的绝对值,除以2的原因是在 t 时刻地类 k 的总减少面积 L_k 和 $t+1$ 时刻总增加面积 G_k 同时存在。式(6)表示转变变化面积 S 等于总变化面积 D 与净变化面积 N 的差值。

$$L_k = \left(\sum_{j=1}^J C_{kj} \right) - C_{kk} \quad (1)$$

$$G_k = \left(\sum_{i=1}^J C_{ik} \right) - C_{kk} \quad (2)$$

$$T_{ij} = C_{ij} + C_{ji} \quad (i > j) \quad (3)$$

$$D = \sum_{k=1}^J L_k = \sum_{k=1}^J G_k \quad (4)$$

$$N = \frac{\sum_{k=1}^J |L_k - G_k|}{2} \quad (5)$$

$$S = D - N \quad (6)$$

2.2 行为聚类的3个原则

Pontius 和 Malizia^[4]根据地类随时间的变化,提出了5个聚类原则,分析了按照一定聚类方式对总变化的净变化面积 N 和转变变化面积 S 产生的影响,这2部分变化之和即为整个时间间隔内的总变化面积 D 。总变化的净变化部分 N 是每种地类 k 的面积总数在2个时间点的差值,增加面积 G_k 大于减少面积 L_k ,则地类 k 为净增加地类,反之则是净减少地类;总变化的转变变化部分面积 S 是地类 k 在空间位置重新分配的过程,即同一时间间隔内,某一地类在一个空间位置发生了像元变化,在另外一个时间点同一空间位置的地类发生了像元重新分配。

对于聚类前的遥感解译影像,利用任意2个地类之间聚类可能性,得到聚类后的结果。本文从数学角度对下述的原则1、2、3进行了证明,聚类后的总变化面积 D 、净变化部分面积 N 和转变变化部分面积 S 取决于聚类前遥感解译影像中的地类 A 与地类 B 之间是否存在相互转换,以及是否存在时间间隔 $[Y_i, Y_{i+1}]$ 内地类 A 与地类 B 转变时表现为净减少或净增加。3个原则都是基于整个时间间隔内的总变化 D 等于净变化面积 N 和转变变化面积 S 之和(式(6))。因此,由聚类产生的整个时间间隔内总变化面积差值 $\Delta Total_M$ 等于净变化面积差值 ΔNet_M 和转变变化面积差值 $\Delta Swap_M$ 之和,如式(7)所示。

$$\Delta Total_M = \Delta Net_M + \Delta Swap_M \quad (7)$$

原则1可应用于所有情况,也是提出其他原则的基础,原则2和原则3应用于减少地类和减少地类合并或增加地类和增加地类合并的情况。原则2描述了对总变化中净变化部分面积聚类产生的影响,原则3描述了总变化中转变变化部分面积在聚类过程中产生的影响。原则1和原则2证明了原则3。

原则1:假设对于任意相互发生转变的地类都可进行聚类,则聚类会使整个时间间隔内总变化面积减少,差值用 $\Delta Total_M$ 表示。式(8)从数学角度表达了任意地类 i 和地类 j 相互发生转变,对地类 i 和地类 j 进行聚类后,总变化面积差值为 $\Delta Total_M$ 。

$$\Delta Total_M = -(C_{ij} + C_{ji}) \quad (8)$$

任意2种地类合并后,二者之间的转变由非主对角线元素转为对角线元素。地类 i 和 j 的合并会减少总变化面积,其值变为总变化面积 D 减去 $C_{ij} + C_{ji}$,

原因是 $C_{ij} + C_{ji}$ 转变为保持不变的面积,即转变为对角线元素。地类合并导致遥感解译地图上总变化面积减少。聚类后总变化唯一不会减少的特例是没有发生任何转变的地类进行合并,即 $C_{ij} + C_{ji} = 0$ 。

原则2:假设净减少地类和净减少地类合并或净增加地类和净增加地类合并,聚类后的遥感影像的净变化面积 N 保持不变。用数学语言来讲,如果 $(C_{+i} - C_{i+} \leq 0 \text{ 且 } C_{+j} - C_{j+} \leq 0)$ 或 $(C_{+i} - C_{i+} \geq 0 \text{ 且 } C_{+j} - C_{j+} \geq 0)$ 。其中, C_{+i} 代表末期类型 i 的总量; C_{+i} 代表的是初期类型 i 的总量; C_{j+} 代表末期类型 j 的总量; C_{j+} 代表初期类型 j 的总量(式(9))。

$$\Delta Net_M = 0 \quad (9)$$

原则3:假设净减少地类和净减少地类合并或净增加地类和净增加地类合并,聚类后的遥感解译影像上的转变变化面积减少。从数学角度来看,如果 $C_{+i} - C_{i+} \leq 0 \text{ 且 } C_{+j} - C_{j+} \leq 0$,或 $C_{+i} - C_{i+} \geq 0 \text{ 且 } C_{+j} - C_{j+} \geq 0$,则表示为式(10):

$$\Delta Swap_M = -(C_{ij} + C_{ji}) \quad (10)$$

当合并的地类显示为净变化方向时,地类 i 和 j 的合并会使转变变化部分面积减少了 $C_{ij} + C_{ji}$ 。式(10)是式(8)、(9)取代式(7)的结果。因此,原则3是原则1、2产生的直接结果,因为所有涉及地类聚类的原则都有相同的净变化方向。原则3适用于原则2所适用的所有条件。

2.3 行为聚类的6种行为方式

根据式(1)、(2)定义了地类 k 在初始分类时的6种变化行为:

(1)当总减少 $L_k > 0$ 且总增加 $G_k = 0$,聚类后遥感影像地类 k 在某一空间位置出现总减少,把这种聚类行为定义为完全减少地类(Exclusive Loser);

(2)当 $L_k = G_k = 0$,聚类后的遥感影像地类 k 既不出现增加也不出现减少的现象,把这种聚类的行为方式定义为完全不变地类(Exclusive Zero);

(3)当 $L_k = 0$ 且 $G_k > 0$,聚类后的遥感影像地类 k 在空间某一位置仅出现了总增加,把这种行为聚类方式定义为完全增加地类(Exclusive Gainer);

(4)当 $L_k > G_k > 0$,聚类后的遥感影像地类 k 在空间某一位置增加,在空间其他位置减少,且地类 k 的总减少大于总增加,地类 k 在空间位置发生了重新分配现象,把这种行为聚类方式定义为转变减少地类(Swapping Loser);

(5) 当 $L_k = G_k > 0$, 聚类后的遥感影像地类 k 在空间上发生了地类增加和地类减少, 且减少地类和增加地类面积相同, 则把这种地类的行为聚类方式定义为转变为零地类 (Swapping Zero);

(6) 当 $G_k > L_k > 0$, 聚类后的遥感影像地类 k 增加和减少在空间位置发生分配, 并且增加地类的面积大于减少地类的面积, 把这种地类行为聚类方式定义为转变增加地类 (Swapping Gainer)。

2.4 行为聚类算法流程

基于行为聚类的算法可在保持净变化不变的前提下合并地类, 该算法不会将净增加地类和净减少地类合并。首先, 将 T_{ij} 最小的 2 种地类进行合并, 每一步都会先将 2 种地类沿着同一个净变化方向进行合并, 以最大可能保证转变变化部分的面积。

图 1 为行为聚类算法的流程, 分为 3 个阶段。

第 1 阶段是完全不变阶段 (Exclusive Zero Phase), 用该算法将所有完全不变 (Exclusive Zero) 的地类进行聚类, 直到所有完全不变地类 (Exclusive Zero) 合并为一种综合的完全不变地类, 在算法结束时仅保留 3 种地类中的一种。完全不变阶段通过选择具有最小变化信息的地类, 决定了地类合并的先后次序。

第 2 阶段是完全减少和完全增加阶段, 该阶段将完全增加地类和完全减少地类分别进行合并。该算法对完全增加地类的界定是在所有完全增加地类中具有最小的 T_{ij} ; 同理, 对完全减少地类的限

定是在所有完全减少地类中具有最小的 T_{ij} 。先比较 2 种已经界定的地类, 然后将具有最小 T_{ij} 的地类合并。如果 2 种或更多的完全增加或减少地类有相同的 T_{ij} , 该算法就会将包括保持不变类型的具有最小面积总和的地类合并。随后, 进入下一步地类合并的过程。第 2 阶段产生 2 种地类, 一种是完全减少地类, 是所有完全减少地类的综合聚类结果; 另一种是完全增加地类, 对应的是所有完全增加地类的综合聚类结果。这 2 种地类是该算法下一个阶段进行聚类的可选地类。在该算法的前 2 个阶段, 净变化和转变变化均为常数。

第 3 阶段是转变阶段, 用该算法计算出每一对有相同净变化方向地类的 T_{ij} 。地类对的界定是在所有减少地类对和所有增加地类对中具有最小的 T_{ij} 。该算法会将已界定的转变减少地类或转变增加地类进行合并, 同样具有最小的 T_{ij} 。如果 2 种或更多种转变地类有相同的 T_{ij} , 该算法就会将具有最小总变化的地类对合并。如果超过一对的地类有最小的 T_{ij} 且有相同的最小总变化, 该算法就会将具有最小总面积的地类对合并。新形成的地类又会参与到随后的聚类过程中, 使用同样的逻辑原理, 转变为零的地类可和增加地类、减少地类或另一个转变为零的地类形成一对进行聚类。可能的转变地类对是: 增加地类和增加地类, 增加地类和为零地类, 减少地类和减少地类, 减少地类和为零地类, 为零地类和为零地类。随着聚类过程的推进, 地类总数越来越少, 净变化始终保持常数, 而转变变化面

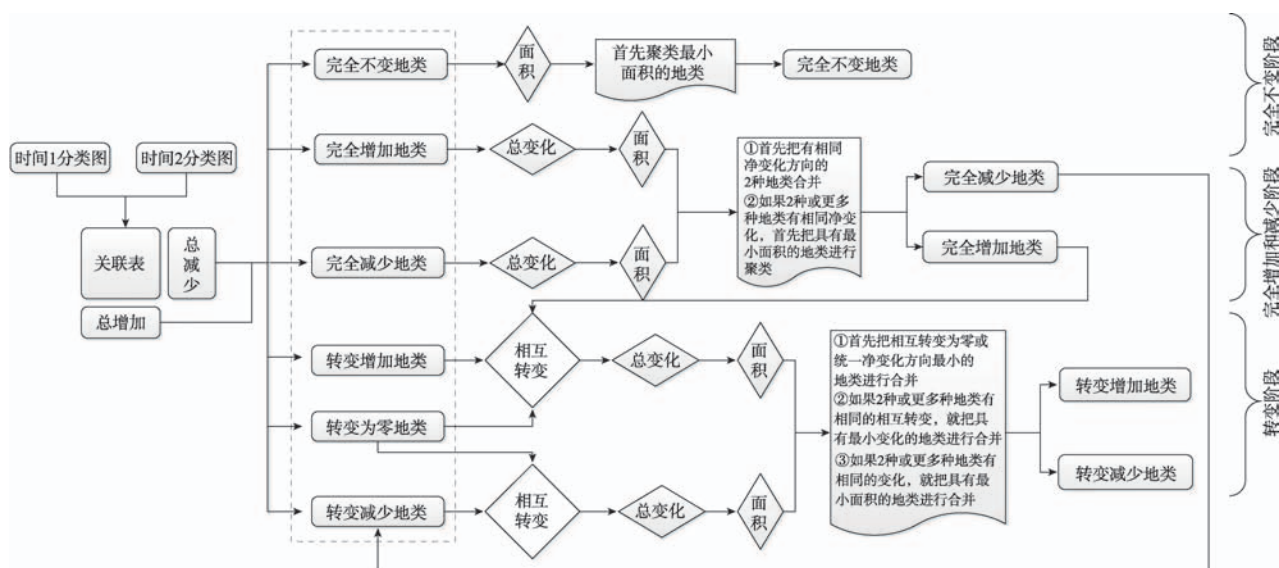


图 1 基于行为聚类的算法流程图

Fig. 1 Methodological flow of the algorithm

积在随后聚类的每一步均会用 T_i 来逐渐缩小。

2.5 地类的强度变化分析

强度分析理论是 Pontius 研究团队提出的理论框架,是一种自上而下的分层等级的解释性的框架结构,用于分析地类之间的变化差异,由间隔层次、地类层次和转变层次3个层次构成^[9-10]。本研究只涉及到强度分析的地类层次,关于强度分析的更多理论可访问: <https://sites.google.com/site/intensity-analysis/>获取。

式(11)给出了研究区域整个时间范围 $[Y_i, Y_T]$ 的统一变化强度,分子中的 C_{ij} 表示 t 时刻从地类 i 转变到地类 j 的面积, C_{ii} 表示 t 时刻转移矩阵中保持不变地类的面积。对于特定时间间隔 $[Y_i, Y_{i+1}]$, 式(12)、(13)分别给出了地类的年均观察增加变化强度和年均观察减少变化强度,也在每个特定时间间隔内为每种地类产生了一个输出结果。通过比较观察变化强度大小和统一变化强度,如果某地类的强度变化大于统一变化强度,就称该地类是活跃的;如果某地类的变化强度小于统一变化强度,就称该地类是沉寂的。

$$U = \frac{\sum_{t=1}^{T-1} \left\{ \sum_{j=1}^J \left[\left(\sum_{i=1}^J C_{ij} \right) - C_{ii} \right] \right\} / \left[\sum_{j=1}^J \left(\sum_{i=1}^J C_{ij} \right) \right]}{Y_T - Y_1} \times 100\% \quad (11)$$

$$G_{ij} = \frac{\left[\left(\sum_{i=1}^J C_{ij} \right) - C_{ii} \right] / (Y_{i+1} - Y_i)}{\sum_{i=1}^J C_{ij}} \times 100\% \quad (12)$$

$$L_{ii} = \frac{\left[\left(\sum_{j=1}^J C_{ij} \right) - C_{ii} \right] / (Y_{i+1} - Y_i)}{\sum_{j=1}^J C_{ij}} \times 100\% \quad (13)$$

3 算法实验结果与分析

利用中国科学院资源环境科学数据中心(<http://www.resdc.cn/Default.aspx>)提供的云南省1990年和2010年1:10万比例尺土地利用现状遥感监测数据为原始数据,采用刘纪远^[11]、张增祥^[12]等建立的分类体系,包括6个一级分类和20个二级分类。为保证获取数据的解译质量和一致性,他们对每期数据集都进行统一的质量检查和数据集成^[12]。通过对每期数据开展野外考察,实地调查(除台湾省以外)

各省土地利用情况,获取大量的外业调查记录、照片,对野外调查资料与外业实地记录,按照县数10%的比例随机抽样开展精度验证。土地利用一级类型综合评价精度达到94.3%以上,二级类型分类综合精度达91.2%以上,满足1:10万比例尺用户制图精度^[11-17]。本实验采用其二级分类体系,利用行为聚类算法,得到9种土地利用类型,最后与刘纪远^[11]等建立的一级分类体系结果进行比较。

本文从面积变化和强度变化2个方面,对6个一级类型和行为聚类结果的9个土地类型进行对比。在变化分析中,对比了2个分类体系的变化面积(包括总增加(G_k)、总减少(L_k)和保持不变3部分)以及变化强度(增加强度和减少强度2部分)。从变化面积来看,基于行为聚类方式保留了最大化的转变变化部分面积和保持净变化面积不变,而6类的分类体系出现了净变化和转变变化,从关联表的非对角线元素转移到对角线上,导致整个时间间隔内的总变化面积减少,使得研究区域的真实变化没有反映出来。从增加强度和减少强度来看,6类分类体系的林地表现为增加强度和减少强度均是活跃的,而9类分类体系可明显地看出森林地的减少强度和园地的增加强度是活跃的,结合转移矩阵可知,园地的增加来自森林地的减少,可更清晰地看出将林地细致分类的优点。

3.1 基于行为聚类结果

基于行为聚类的算法是按照3个原则和6种行为方式进行两两合并,只能将净减少地类和净减少地类合并,净增加地类和净增加地类合并,净增加地类与净减少地类不能合并。表1是对原始的分类体系利用行为聚类算法的合并次序。由表1可知,保持地类净变化不变、地类转变变化保留最大时,所合并的地类数为9类。此时,地类总变化由5.16%减少到5.10%,减少了0.06%,地类转变变化由3.60%减少到3.53%,减少了0.07%,地类净变化面积保持不变,这样保持地类总变化面积为最大。

图2是聚类结果,第1阶段是完全不变阶段,原始分类体系由20类进行两两聚类,聚类到17类时,地类总变化与原始分类结果地类总变化一样,没有发生变化。第2阶段是完全减少和增加阶段,这个阶段是将完全增加地类和完全减少地类进行界定,分别找到完全增加地类和完全减少地类界定的最小值进行合并,聚类到地类14;地类16是有林地和聚类18进行合并,地类15是聚类16与裸岩石砾地

表1 基于行为聚类方法的次序
Tab. 1 Behavior-based aggregation sequence

地类数量	地类总变化/(%)	转变变化/(%)	净变化/(%)	聚类次序	聚类类型
20	5.16	3.60	1.56	湖泊、沼泽地	G/净增加、G/净增加
19	5.16	3.60	1.56	河渠、水库坑塘	G/净增加、G/净增加
18	5.16	3.60	1.56	冰川与永久积雪、裸地	L/净减少、L/净减少
17	5.16	3.60	1.56	低覆盖度草地、滩地	G/净增加、G/净增加
16	5.16	3.60	1.56	有林地、聚类18	L/净减少、L/净减少
15	5.16	3.60	1.56	聚类16、裸岩石砾地	L/净减少、L/净减少
14	5.16	3.60	1.56	旱地、聚类20	G/净增加、G/N净增加
13	5.15	3.59	1.56	聚类14、聚类17	G/净增加、G/净增加
12	5.15	3.58	1.56	水田、聚类15	L/净减少、L/净减少
11	5.14	3.57	1.56	聚类13、聚类19	G/净增加、G/净增加
10	5.12	3.56	1.56	聚类11、工交建设用地	G/净增加、G/净增加
9	5.10	3.53	1.56	聚类10、城镇	G/净增加、G/净增加
8	5.05	3.49	1.56	疏林地、中覆盖度草地	L/净减少、L/净减少
7	5.01	3.45	1.56	聚类12、灌木林地	L/净减少、L/净减少
6	4.64	3.08	1.56	其他林地、农村居民点	G/净增加、G/净增加
5	4.64	3.08	1.56	聚类7、聚类8	L/净减少、L/净减少
4	4.09	2.52	1.56	聚类5、高覆盖度草地	L/净减少、L/净减少
3	3.06	1.50	1.56	聚类9、聚类6	G/净增加、G/净增加
2	2.95	1.38	1.56		

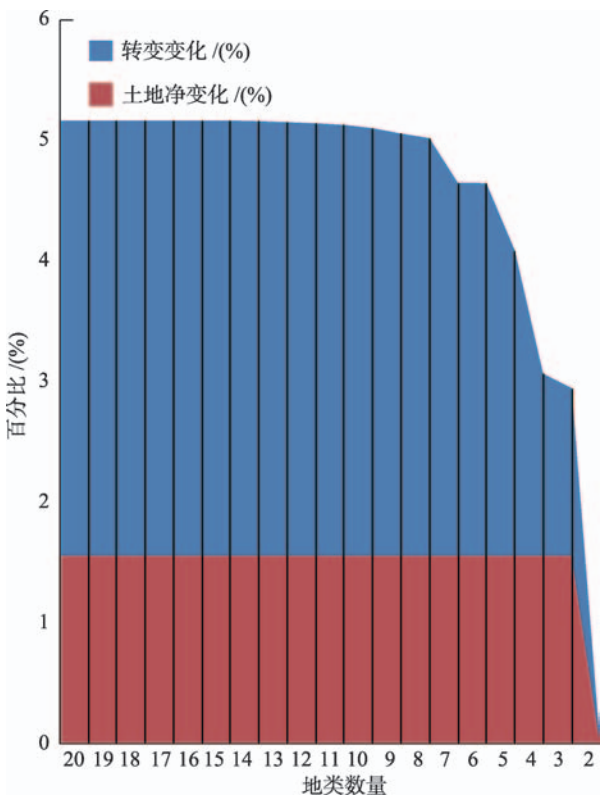


图2 1990–2010年行为聚类次序地类转换和地类净变化面积百分比

Fig. 2 Swap change and net change during 1990–2010 in terms of percentage on the study extent for the behavior-based aggregation sequence

进行合并。第3阶段是转变阶段,这部分是地类转换面积变化最为剧烈阶段,聚类到9类时转变变化开始出现下降,因此,本文选择9类作为最终结果。地类13是聚类14和聚类17进行合并,地类12是水田和聚类15进行合并,地类11是聚类13和聚类19进行合并,地类10是聚类11和交通建设用地进行合并,地类9是聚类10与城镇地类进行合并。随着聚类过程的推进,地类总数越来越少,净变化始终保持为常数,而转变变化在随后聚类的每一步都会逐渐减少。结合表1和图2可知,聚类为2类时,净变化仍保持常数,地类转变变化从3.60%减少到1.38%,转换变化部分达到最小。

根据行为聚类的方法,并结合合理的逻辑顺序,将20种地类重分类成9类,即:(1)耕地(11:水田,12:旱地);(2)森林地(21:有林地,22:灌木林地);(3)疏林地(23:疏林地);(4)园地(24:其他林地);(5)天然牧草地(31:高覆盖牧草地,32:中覆盖草地);(6)人工草地(33:低覆盖草地);(7)水域(41:河渠,42:湖泊,43:水库坑塘,44:冰川永久积雪,46:滩地,64:沼泽地);(8)建设用地(51:城镇,52:农村居民点用地,53:工交建设用地);(9)裸地(65:裸地,66:裸岩石砾地)。

3.2 1990–2010 年 6 个和 9 个土地利用类型变化面积和变化强度

图 3(a)是 1990–2010 年 6 个土地利用类型分类结果变化面积,建设用地、林地和水域的增加面积大于减少面积,并且都大于零,属于转变增加地类 (Swapping Gainer);耕地、草地和裸地的增加面积小于减少面积,但都大于零,属于转变减少地类 (Swapping Loser)。结合图 3(a)和表 2 的转移矩阵可看出,建设用地和林地增加面积主要来自于耕地和草地,也存在耕地、草地和林地之间的相互转换。

图 3(b)是 1990–2010 年行为聚类土地利用分类结果的变化面积,建设用地、园地、人工草地和水域的增加面积大于减少面积,并且都大于零,属于转变增加地类 (Swapping Gainer);森林地、疏林地、耕地、天然牧草地和裸地的减少面积大于增加面积,并且都大于零,属于转变减少地类 (Swapping Loser)。从表 3 的转移矩阵可看出,建设用地的增加主要源于耕地的减少,园地的增加源于森林地和天然牧草地的减少。这也可看出,原始森林转为橡

胶园茶园的面积逐年增加^[18],并且随着桉树等速生经济林的广泛种植^[19],虽然云南省林地总面积增加,但生物多样性降低、水源涵养性变弱、水土流失严重,干旱等自然灾害加重^[20]。

从图 3(a)和图 3(b)的结果可看出,2 种分类体系中的耕地、裸地、建设用地和水域变化面积基本一致。9 种分类结果中的天然牧草地和人工草地相对于 6 种分类结果仅一种草地而言,可更清晰地表明,发生动态变化较大的是天然牧草地。6 种分类结果中的林地对应于 9 种分类结果中的园地、森林地和疏林地,可看出园地增加更明显,疏林地和森林地减少较多,但 6 种分类结果中的林地仅表现出增加略大于减少。对比结果表明,基于行为聚类的方法可更清晰地揭示地类变化过程中的细微变化,这对于深入理解地类变化、挖掘变化来源,进而揭示土地利用变化所造成的宏观生态环境变化及其他环境效应具有重要意义。

图 4(a)是 6 类分类体系的强度分析结果,林地的增加强度和减少强度均位于统一强度线右侧,属

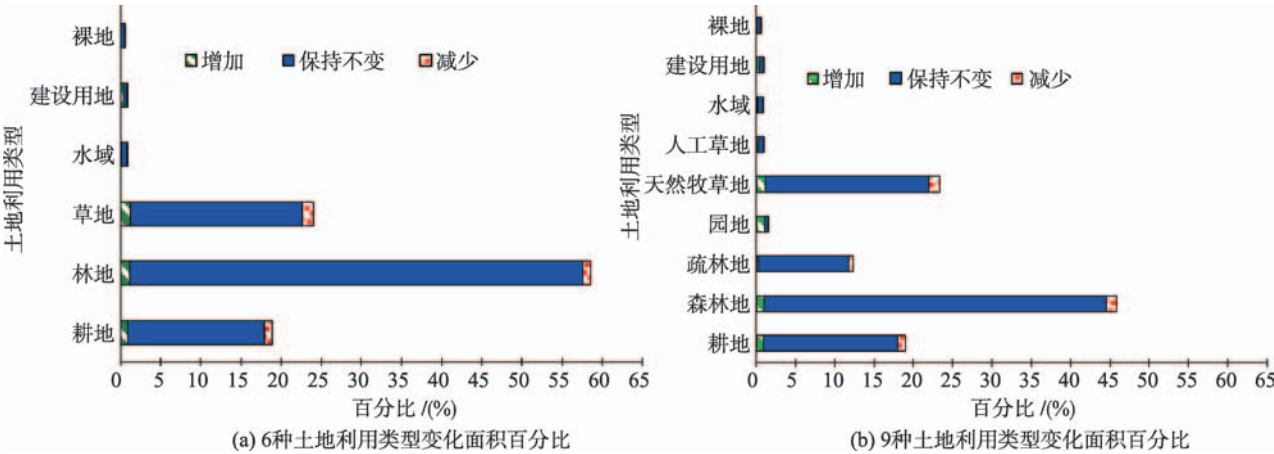


图 3 1990–2010 年土地利用类型变化面积百分比
Fig. 3 Percentage of area change of different land categories during 1990-2010

表 2 1990–2010 年云南省 6 个一级类型的转移矩阵(hm²)
Tab. 2 Cross-tab of six land categories during 1990-2010 in Yunnan Province (hm²)

1990 年	2010 年						总减少
	耕地	林地	草地	水域	建设用地	裸地	
耕地	6 530 509	120 729	146 418	6122	99 727	1407	374 403
林地	104 013	21 606 562	239 751	5735	16 805	766	367 072
草地	211 137	299 079	8 216 116	5935	21 057	1644	538 852
水域	2156	1413	8898	270 013	784	158	13 409
建设用地	5597	1023	1362	645	170 057	473	9100
裸地	2234	1413	57 804	275	96	151 779	61 822
总增加	325 137	423 658	454 233	18 712	138 469	4448	

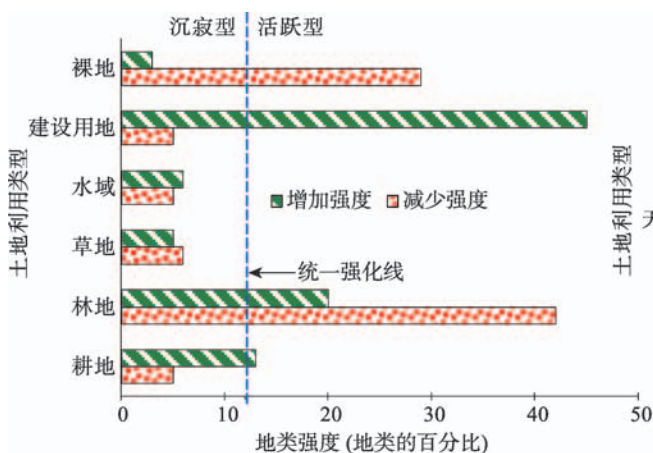
表3 1990-2010年云南省行为聚类结果的转移矩阵(hm²)Tab. 3 Cross-tab of nine land categories during 1990-2010 in Yunnan Province (hm²)

1990年	2010年									
	耕地	森林地	疏林地	园地	天然牧草地	人工草地	水域	建设用地	裸地	总减少
耕地	6 529 142	81 532	19 346	19 823	144 086	2244	6088	99 711	1407	374 237
森林地	60 799	16 675 113	17 398	226 482	179 115	31	4218	11076	321	499 443
疏林地	38 126	68 825	4 406 585	44 312	57 316	63	1446	4352	403	214 843
园地	5079	1289	199	157 009	3139	32	43	1373	42	11 196
天然牧草地	210 692	197 730	22 972	76 378	7 948 931	2085	5888	20 270	1621	537 636
人工草地	419	1652	240	23	6920	254 358	26	785	23	10 088
水域	2147	980	295	138	5563	3335	268 223	781	158	13 397
建设用地	5597	479	356	188	1362	0	645	170 052	473	9100
裸地	2234	1413	0	0	12 124	45 540	275	96	151 324	61 682
总增加	325 093	353 899	60 806	367 344	409 625	53 330	18 629	138 444	4448	374 237

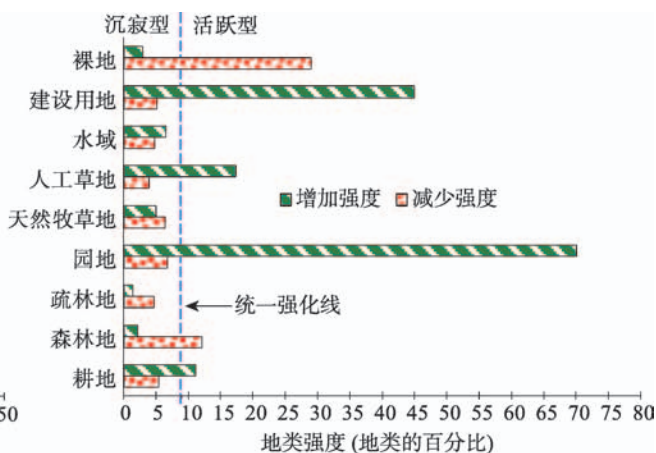
于活跃型地类。建设用地和耕地的增加强度位于统一强度线右侧,属于增加活跃型地类。裸地的减少强度位于统一强度线右侧,属于减少活跃型地类。水域和草地的增加强度和减少强度位于统一强度线左侧,属于沉寂型地类。由图3(a)和表2可知,在6种土地利用类型中林地是转变面积最大的地类,也体现为活跃型。这说明林地转换为其他地类具有最大的变化面积和最强的变化强度,这和1990年林地面积占整个研究区域的比重较大有关。此外,自1989年以来,中国为恢复和扩大森林植被、遏制水土流失、综合治理江河,首次实施了大规模的林业生态工程^[20]。20世纪末,云南开始实施退耕还林工程,尤其经济林的建设,这也是林地变化强度表现为活跃型的原因。由于建设用地具有不可逆性,表现为其他土地利用类型向建设用地转

入,而建设用地转变为其他土地利用类型的面积极少,因此,建设用地表现出了很强的增加强度,转变速度快,属于典型的增加活跃型地类。草地和水域距离统一强度线较近且具有相同变化方向的增加强度和减少强度。裸地表现为弱的增加强度和强的减少强度,原因是人们加大对裸地的开发程度,导致裸地转化为其他地类,所以,裸地减少强度大于增加强度。

图4(b)是利用行为聚类算法得到的9种土地利用类型的强度变化结果。裸地、建设用地、水域和耕地的变化强度方向和6个一级类型的变化强度方向一致。人工草地的增加强度位于统一强度线右侧,属于活跃型。园地的增加强度远远大于其他地类的增加强度,而森林地的减少强度也是活跃的,结合表3的转移矩阵可看出,园地的增加来自森林



(a) 6类土地利用类型增加强度和减少强度变化



(b) 9类土地利用类型增加强度和减少强度变化

图4 1990-2010年土地利用类型增加强度和减少强度变化

Fig. 4 Gross gain intensity and gross loss intensity of different land categories during 1990-2010

地的减少。查阅资料发现近20年来茶叶和天然橡胶涨势显著,特别是2006年以来橡胶价格暴涨,巨大的经济利益刺激了茶园和橡胶园等园地的扩张^[19]。

图4(a)和图4(b)相比,行为聚类的结果详细地体现出了土地利用类型中草地和林地的变化强度,森林地、疏林地和园地比6个一级类型的林地更加详细说明地类之间的变化,人工草地的增加强度较强,主要来自裸地的减少(表3)。

3.3 分类结果对比分析

图5分别对基于行为聚类、6类分类体系和原始分类结果进行比较,这三者都具有相同大小的地类净变化,地类转变变化减少的百分比由大到小依次是:原始分类结果>行为聚类结果>6类分类体系,其中行为聚类方式可保持地类转变部分信息最大,比6类的分类系统可更好反映土地利用变化动态过程。

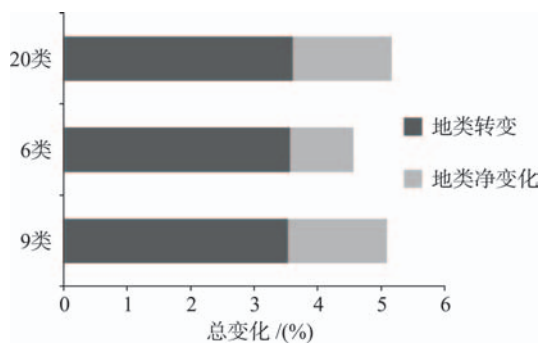


图5 基于行为聚类、常用分类系统和原始分类系统结果比较
Fig. 5 Comparison between the 20 original categories, the 6 common used categories and the 9 behavior-based categories

4 结论与讨论

通过建立行为聚类算法,对地类间合并的行为方式和合并方向进行界定。以1990年和2010年云南省20个二级类型的分类结果为数据源进行聚类,从变化面积和变化强度2方面进行了对比分析。

(1)与其他降维方法相比,基于行为聚类的算法是否合适取决于特定的研究问题。如果详细地类存在大量的行为方式(如上述6种行为都存在或存在多数),则行为聚类的算法合适。例如,科学家需理解林地的具体变化过程,则行为聚类的算法将林地分为3类是有益的。但是,如果详细地类存在一致的行为方式(如全部为转变增加地类或全部为转变减少地类),则行为聚类算法不合适。例如,6类中的耕地和9类中的耕地变化类似,则没有必要

再细分。

(2)云南省地处西南山区,林地占全省总面积的53%^[22],6个一级类型中的林地的增加和减少强度均是活跃的,从9个类型来看,具体体现为森林地的减少强度和园地的增加强度是活跃的。结合转移矩阵来看,园地的增加主要来自森林地的减少,这对云南省的生物多样性保护是不利的,需尽快采取措施,保护原始森林。2种分类体系中的建设用地均表现出强的增加强度,主要来自耕地的减少,这与野外调查结果一致。

(3)从数学角度证明了聚类算法的3个原则,定义了6种行为聚类方式。按照某一种方式两两地类之间合并,始终保持净变化不变。实验结果表明,在特定的样本环境下,该方法得到的分类结果能保持地类转变信息,并有效缩减分类的维度。但是,该算法没有充分考虑专家知识,可能在最后会出现把概念上属于2个地类的种类进行合并。下一步工作将会加入其他原则,假设增加地类和减少地类进行合并,使聚类结果加入专家知识后更合理,但可能导致净变化出现减少,转变变化出现减少、增加、保持不变的现象。

(4)基于行为的行为聚类算法具有广泛应用前景,可对植被覆盖、土壤侵蚀数据和其他数据进行聚类。本文使用的聚类算法只考虑了一个时间间隔的数据,下一步会利用大于一个时间间隔的多期数据进行聚类,从整个系统论的角度进行行为聚类,使其更符合实际需求。

致谢:感谢美国克拉克大学地理学院 Robert Gilmore Pontius, Jr 教授给予的指导。感谢中国科学院遥感与数字地球研究所张增祥研究员提出的宝贵意见。

参考文献(References):

- [1] 全国农业区划委员会.土地利用现状调查技术规程[M].北京:测绘出版社,1984. [National agricultural zoning committee. Technology procedures of land use status quo survey[M]. Beijing: Surveying and Mapping Press, 1984.]
- [2] 刘纪远.中国资源环境遥感宏观调查与动态研究[M].北京:科学出版社,1996. [Liu J Y. Macro-scale survey and dynamic study of natural resources and environment of china by remote sensing [M]. Beijing: China Science and Technology Press, 1996.]
- [3] 中华人民共和国国家质量监督检验检疫总局国家标准化管理委员会.中华人民共和国国家标准——土地利用现状分类(GB/T21010-2007)[S].北京:中国标准出

- 版社,2007. [Standardization Administration of China. The standard of the people's Republic of China: classification of land use status(GB/T21010- 2007) [S]. Beijing: Standards Press of China, 2007.]
- [4] Pontius Jr R G, Malizia N R. Effect of category aggregation on map comparison[J]. *Geographic Information Science*, 2005,3234:251-268.
- [5] Ahlqvist O, Shortridge A. Spatial and semantic dimensions of landscape heterogeneity[J]. *Landscape Ecology*, 2010,25(4):573-590.
- [6] Buyantuyev A, Wu J. Effects of thematic resolution on landscape pattern analysis[J]. *Landscape Ecology*, 2007, 22(1):7-13.
- [7] Buyantuyev A, Wu J, Gries C. Multiscale analysis of the urbanization pattern of the Phoenix metropolitan landscape of USA: Time, space and thematic resolution[J]. *Landscape and Urban Planning*, 2010,94:206-217.
- [8] Aldwaik S Z, Onsted J A, Pontius Jr R G. Behavior-based aggregation of land categories for temporal change analysis[J]. *International Journal of Applied Earth Observation and Geo-information*, 2015,35,Part B(1):229-238.
- [9] Aldwaik S Z, Pontius Jr R G. Intensity analysis to unify measurements of size and stationarity of land changes by interval, category, and transition[J]. *Landscape and Urban Planning*, 2012,106(1):103-114.
- [10] Pontius Jr R G, Gao Y, Giner N M, *et al.* Design and interpretation of intensity analysis illustrated by land change in Central Kalimantan, Indonesia[J]. *Land*, 2013,2(3):351-369.
- [11] 刘纪远,张增祥,庄大方,等.二十世纪九十年代中国土地利用变化的遥感时空信息研究[M].北京:科学出版社, 2005. [Liu J Y, Zhang Z X, Zhuang D F, *et al.* Remote sensing information study of land use change in China in 1990s [M]. Beijing: Sciences Press, 2005.]
- [12] 张增祥,赵晓丽,汪潇,等.中国土地利用遥感监测[M].北京:星球地图出版社,2012. [Zhang Z X, Zhao X L, Wang X, *et al.* Remote sensing monitoring of land use in China [M]. Beijing: StarMap Press, 2012.]
- [13] 刘纪远,匡文慧,张增祥,等.20世纪80年代末以来中国土地利用变化的基本特征与空间格局[J].*地理学报*,2014, 69(1):3-14. [Liu J Y, Kuang W H, Zhang Z X, *et al.* Spatio-temporal characteristics, patterns and causes of land-use changes in China since the late 1980s[J]. *Journal of Geographical Sciences*, 2014,24(2):195-210.]
- [14] 刘纪远,刘明亮,庄大方,等.中国近期土地利用变化的空间格局分析[J].*中国科学D辑*,2002,32(12):1031-1040. [Liu J Y, Liu M L, Zhuang D F, *et al.* Spatial pattern analysis of land use recent change in China[J]. *Science in China(Series D)*, 2002,32(12):1031-1040.]
- [15] 张增祥,汪潇,王长耀,等.基于框架数据控制的全国土地覆盖遥感制图研究[J].*地球信息科学学报*,2009,11(2): 216-224. [Zhang Z X, Wang X, Wang C Y, *et al.* National land cover mapping by remote sensing under the control of interpreted data[J]. *Journal of geo-information science*, 2009,11(2):216-224.]
- [16] Liu J Y, Liu M L, Zhuang D F, *et al.* Study on spatial pattern of land-use change in China during 1995- 2000[J]. *Science in China*, 2003,46(4):373-384.
- [17] Liu J Y, Zhang Z X, Xu X L, *et al.* Spatial patterns and driving forces of land use change in China during the early 21st century[J]. *Journal of Geographical Sciences*, 2010,20(4):483-494.
- [18] 刘晓娜,封志明,姜鲁光,等.西双版纳土地利用/土地覆被变化时空格局分析[J].*资源科学*,2014,36(2):233-244. [Liu X N, Feng Z M, Jiang L G, *et al.* Spatial-temporal pattern analysis of land use and land cover change in Xishuangbanna[J]. *Resources Science*, 2014,36(2):233-244.]
- [19] 刘晓娜.中老缅交界地区土地利用/土地覆被变化研究[D].北京:中国科学院地理科学与资源研究所,2013. [Liu X N. Study on land use and land cover change in the borerder region of China, Laos, and Myanmar[D]. Beijing: Institute of Geography Science and Natural Resources Research, Chinese Academy of Sciences, 2013.]
- [20] 赵元茂,杞明辉.全球变化中的我国云南气候与生态响应[J].*辽宁气象*,2005,1(1):20-22. [Zhao Y M, Qi M H. Climate and ecological response to global change in Yunnan Province[J]. *Liaoning qixiang*, 2005,1(1):20-22.]
- [21] 王瑞红,董慧秀.1950-2010年云南高原水资源利用与生态环境变迁[J].*黑龙江史志*,2014,1(6):65-70. [Wang R H, Dong H X. The use of water resources and the ecological environment change in yunnan plateau during 1950-2010[J]. *Heilongjiang local Chronicles*, 2014,1(6): 65-70.]
- [22] 云南省人民政府门户网站.云南省自然地理[EB/OL].2014. <http://www.yn.gov.cn/>. [People's government of Yunnan Province. Natural geography in Yunnan Province[EB/OL].2014. <http://www.yn.gov.cn/>.]