International Journal of Applied Earth Observation and Geoinformation 35 (2015) 229–238

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag

Behavior-based aggregation of land categories for temporal change analysis

Safaa Zakaria Aldwaik[a], Jeffrey A. Onsted[b], Robert Gilmore Pontius Jr[c,*]

[a] Ramallah Municipality, 4 Issa Ziadeh Street, Ramallah 600, Palestine
[b] Department of Earth & Environment, Florida International University, 11200 SW 8th Street, Miami, FL 33199, USA
[c] School of Geography, Clark University, 950 Main Street, Worcester, MA 01610-1477, USA

ARTICLE INFO

Article history:
Received 29 May 2014
Accepted 12 September 2014
Available online 31 October 2014

Keywords:
Aggregation
Category
Change
Land
Map
Transition

ABSTRACT

Comparison between two time points of the same categorical variable for the same study extent can reveal changes among categories over time, such as transitions among land categories. If many categories exist, then analysis can be difficult to interpret. Category aggregation is the procedure that combines two or more categories to create a single broader category. Aggregation can simplify interpretation, and can also influence the sizes and types of changes. Some classifications have an a priori hierarchy to facilitate aggregation, but an a priori aggregation might make researchers blind to important category dynamics. We created 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 never aggregates a net gaining category with a net losing category. The behavior-based algorithm at each step in the sequence maintains net change and maximizes swap change. We present a case study where data from 2001 and 2006 for 64 land categories indicate change on 17% of the study extent. The behavior-based algorithm produces a set of 10 categories that maintains nearly the original amount of change. In contrast, an a priori aggregation produces 10 categories while reducing the change to 9%. We offer a free computer program to perform the behavior-based aggregation.

© 2014 Elsevier B.V. All rights reserved.

Introduction

Purpose

Categorical scale concerns the level of detail of a set of categories. Categorical scale is also known as thematic scale, which describes the type of information that a map shows. The choice of categorical scale is one of the central challenges for geographers in mapping land cover, land use, vegetation type, soil type and other categorical variables. Comparison of maps that show a categorical variable at two time points for the same study extent can reveal change during the time interval. However, if maps have a large number of categories, then analysis can become difficult to interpret. Category aggregation is the process of merging detailed categories to create a smaller number of broader categories. Category aggregation can simplify interpretation, but can also reduce the amount of apparent temporal change, depending on which categories are aggregated. If category aggregation is performed strategically, then aggregation can play an important role in data mining, because strategic category aggregation can help to reveal information that might otherwise be lost by other types of aggregations. This article presents a new algorithm to perform a sequence of categorical aggregations in a manner that gives insights concerning categorical change over time. Our algorithm is based mainly on the temporal behavior of each category in terms of its gross gain and gross loss.

Literature

Briassoulis (2000) describes various classification systems that are popular for land change science. Many of these systems have a hierarchical structure, whereby detailed categories are grouped under a smaller number of conceptually broader categories. Anderson et al. (1976) established a popular system where the detailed categories can be aggregated to a coarser level in a hierarchy based on similarity of land use. For example, if we

* Corresponding author. Tel.: +1 508 793 7761; fax: +1 508 793 8881.
E-mail addresses: saldwaik@gmail.com (S.Z. Aldwaik), jonsted@fiu.edu (J.A. Onsted), rpontius@clarku.edu (R.G. Pontius Jr).

http://dx.doi.org/10.1016/j.jag.2014.09.007
0303-2434/© 2014 Elsevier B.V. All rights reserved.

Table 1
Glossary.

Term	Meaning
Contingency table	Square table where the number of rows and the number of columns equals the number of categories. The entry in row *i* and column *j* of the table gives the size of the study extent that is category *i* at the initial time and category *j* at the final time.
Transition	A particular off-diagonal entry in the contingency table.
Total change	Sum of all off-diagonal entries in the contingency table. Total change can be expressed as the sum of two components called net and swap.
Net change	Component of change that is attributable to differences in the quantity of each category between the initial time and the final time.
Swap change	Component of change that is attributable to gross gain of a category in some locations and gross loss of the same category in other locations during the time interval.
Exclusive loser	Category that has positive gross loss and zero gross gain.
Exclusive zero	Category that has zero gross loss and zero gross gain.
Exclusive gainer	Category that has zero gross loss and positive gross gain.
Swapping loser	Category for which its positive gross loss is greater than its positive gross gain.
Swapping zero	Category for which its positive gross loss equals its positive gross gain.
Swapping gainer	Category for which its positive gross loss is less than its positive gross gain.

have various Agricultural subcategories classified according to the detailed Anderson level II system, we can move to the coarser Anderson level I system by aggregating all Agricultural subcategories into one broader Agricultural category. Aggregation can have important implications for subsequent analyses. For example, if the subcategories transition with each other, then aggregation will eliminate those transitions. Also, if one subcategory is responsible for change while other subcategories persist, then aggregation will lose information concerning which subcategory is responsible for the change. Conway (2009) demonstrated that various category aggregations can influence the calibration and validation of a land change simulation model. Aggregation can also influence the results of pattern metrics (Ahlqvist and Shortridge, 2010; Buyantuyev and Wu, 2007; Buyantuyev et al., 2010).

Table 1 is a glossary of terms that the literature and the remainder of our manuscript uses. We define these terms in a broad manner so that they have applications beyond land change science.

Pontius and Malizia (2004) investigated the influence of category aggregation on measurement of change over time. They derived five mathematical principles that dictate the effects of aggregation on the net change and the swap change, which are two components that sum to the total change. Net change is the component of change that derives from a difference between two time points in the number of pixels of each category. Swap change is the component of change where a category appears to reallocate. Reallocation occurs when a category experiences both gross gain in some pixels and gross loss in other pixels during the time interval. The five principles rely on characterizing each category according to the category's net change, thus categories are labeled as net gainers or net losers. A net gainer is a category for which its gross gain is greater than its gross loss. A net loser is a category for which its gross loss is greater than its gross gain. Principle 1 dictates that aggregation cannot increase the total change. If the aggregation is either a net loser with another net loser or a net gainer with another net gainer, then Principle 2 dictates that the net change is maintained and Principle 3 dictates that the swap change decreases by the sum of the transitions between the aggregated categories. If a net loser is combined with a net gainer, then Principle 4 dictates that the net change decreases and Principle 5 dictates that the swap change can decrease, increase, or be maintained. These principles derive from the fundamentals of set theory.

This article introduces a new algorithm to perform a sequence of category aggregations that are based exclusively on the behavior of the categories. Our goal is to produce a new type of aggregation based on information that existing aggregation systems ignore. The mathematical foundation of our approach is the framework concerning net change and swap change, which is also known respectively as quantity difference and allocation difference (Pontius and Millones, 2011). Specifically, our algorithm uses the second and third principles of Pontius and Malizia (2004) to maintain net change and to maximize swap change at each step in the aggregation sequence. Thus if individual detailed categories are involved in large transitions, then those categories tend to be maintained as unaggregated until latter steps in the aggregation sequence.

Case study in Redland, Florida, USA

We illustrate the characteristics and advantages of our algorithm with a case study from Redland, Florida, USA. Redland is part of the Florida Coastal Everglades (FCE) site of the United States National Science Foundation's Long Term Ecological Research network. Redland is 20 miles southwest of Miami and between two national parks. FCE researchers are interested in the rapid suburbanization in some locations and conservation in other locations. Various government agencies are involved in conservation through zoning and acquisition of farmland via eminent domain. Researchers at FCE are examining the impact that zoning has on the calibration of a cellular automaton urban growth model (Onsted and Roy Chowdhury, 2014). Researchers have also been interviewing Redland residents regarding land use with respect to pesticides, fertilizers and water (Harris et al., 2012). We chose Redland as a study area because its data contain 64 categories, thus it illustrates well the need for aggregation. The Redland land categories are organized in an a priori hierarchy based on similarity of use, therefore this hierarchy can serve as an initial framework for use-based aggregation. However, the use-based aggregation can make scientists blind to important signals of land change that might otherwise be apparent with a different combination of aggregated categories. The remainder of this paper compares the use-based aggregation versus behavior-based aggregations that derive from our algorithm.

Methods

Data

We use raster maps from two time points to illustrate our algorithm. Each raster has 270 columns and 233 rows of pixels at the 200-m resolution, where each pixel is classified as exactly one category. Fig. 1a shows maps of 64 categories at year 2001 and year 2006. The rows of Fig. 2 show how the data have an a priori hierarchical structure that groups the categories conceptually into ten broader categories based on use. Fig. 1b shows these ten use-based categories. We compare the use-based aggregation to the behavior-based aggregations that our algorithm produces. The columns of Fig. 2 show the 10-category aggregation that our behavior-based algorithm produces. We present Fig. 2 now so that the reader can understand our goal. Fig. 1c shows these ten categories that our algorithm produces.

Download English Version:

https://daneshyari.com/en/article/4464717

Download Persian Version:

https://daneshyari.com/article/4464717

Daneshyari.com