Available online at www.sciencedirect.com
ScienceDirect
Geoderma 136 (2006) 774–787
GEODERMA
www.elsevier.com/locate/geoderma

Fuzzy soil mapping based on prototype category theory

Feng Qi [a,*], A-Xing Zhu [b,c], Mark Harrower [c], James E. Burt [c]

[a] Department of Political Science and Geography, University of Texas at San Antonio, 6900 NW Loop 1604, San Antonio, TX 78249, USA
[b] State Key Laboratory of Resources and Environmental Information System, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Building 917, Datun Road, An Wai, Beijing 100101, China
[c] Department of Geography, University of Wisconsin–Madison, 550 North Park Street, Madison, WI 53706, USA

Received 11 January 2005; received in revised form 26 May 2006; accepted 6 June 2006
Available online 1 August 2006

Abstract

An essential component of soil mapping is classification, a process of assigning spatial soil entities to predefined categories (classes). However, by their nature soils exist as a continuum both in the spatial and attribute domains and often cannot be fitted into discrete categories without introducing errors or at least over-simplification. One approach to mitigate this problem in digital soil mapping is the combination of fuzzy logic-based class assignment with a raster GIS representation model which allows the continuous spatial variation of soils to be expressed at much greater detail than has been achieved in conventional (analog) soil survey. However, applications of fuzzy soil mapping face two significant challenges: defining the central concept of a soil category and determining the degree of membership to the central concept. Prototype category theory is presented here as a potential solution to these difficulties. Emerging from ideas of family resemblance, centrality and membership gradience, and fuzzy boundaries (fuzzy set theory), prototype category theory stresses the fact that category membership is not homogenous and that some members are better representatives of a category than others. A prototype can be viewed as a representation of the category, that 1) reflects the central tendency of the instances' properties or patterns; 2) consequently is more similar to some category members than others; and 3) is itself realizable but is not necessarily an instance. Based on this notion, we developed a prototype-based approach to acquire and represent knowledge on soil–landscape relationships and apply the knowledge in digital soil mapping under fuzzy logic. The prototype-based approach was applied in a case study to map soils in central Wisconsin, USA. Our approach created maps that were more accurate in terms of both soil series prediction and soil texture estimation than either the traditional soil survey or a case-based reasoning approach.

Keywords: Soil map; Fuzzy logic; Cognitive theory; Prototype category theory; GIS; SoLIM

1. Introduction

In traditional soil mapping it is long-standing convention to classify soils and depict soil classes as discrete polygons on an 'area-class' map (Mark and Csillag, 1989). However, soils are known to exist more or less as a continuum in both geographic space and attribute space (Burrough, 1996; Zhu, 1997a) as one soil type blends into another. Fitting this continuous spatial character of soils into discrete soil categories with full memberships over-generalizes the inherent complexity of soil variation, which in turn degrades the accuracy of soil spatial information products. In addition to the over-generalization of

* Corresponding author. Tel.: +1 210 458 5611; fax: +1 210 458 5430.
E-mail addresses: feng.qi@utsa.edu (F. Qi), azhu@wisc.edu (A.-X. Zhu), maharrower@wisc.edu (M. Harrower), jburt@geography.wisc.edu (J.E. Burt).

0016-7061/$ - see front matter © 2006 Published by Elsevier B.V.
doi:10.1016/j.geoderma.2006.06.001

soil variations, manual production in traditional soil mapping also faces other limitations, primarily low speed and high cost of production (Zhu et al., 2001).

In order to overcome these limitations, many researchers started to explore the use of knowledge-based techniques and fuzzy logic concepts to improve soil mapping process and its products (see McBratney et al., 2003 for a review). Among these endeavors, the SoLIM approach (Zhu and Band, 1994; Zhu et al., 1996, 1997; Zhu, 1997a, b; Zhu et al., 2001) is a knowledge-based soil mapping system developed for mass production of soil survey. SoLIM is an automated soil inference system that combines fuzzy logic-based class assignment with a raster GIS representation model, which allows the continuous spatial variation of soils to be expressed at much greater detail so that the class transitions and within-class variations can be represented. SoLIM uses a *n*-dimensional similarity vector to depict soil properties at each pixel location (i, j) (Zhu, 1997a): $S_{ij}=(S_{ij}^1, S_{ij}^2,\ldots, S_{ij}^k, \ldots,S_{ij}^n)$, where S_{ij}^k represents the similarity value or fuzzy membership of the soil at location (i,j) to the prescribed soil class k, and n is the total number of these prescribed classes. The soil similarity vectors (S) are inferred under fuzzy logic based on the same concept $S=f(E)$ as in traditional soil survey, where the essential formative environment data E can be derived using GIS techniques (Zhu et al., 1996; McSweeney et al., 1994), and the relationship between soil and relevant environmental variables (the soil-landscape model) f is obtained through knowledge acquisition. f should reflect both the central concepts of soil classes and the transitions between central concepts.

Knowledge-based fuzzy digital soil mapping approaches currently face two major challenges in defining f: the representation of the central concept of a soil category and how membership to this central concept is modeled. These difficulties are largely due to two factors. First, knowledge acquisition from human experts has long been noted to be the bottleneck for the development of knowledge-based systems (Molokova, 1993; Weibel et al., 1995), especially in the case of knowledge-based soil mapping, where knowledge of the soil–landscape model largely exists as "tacit knowledge" (Hudson, 1992). Second, it is desirable that the extracted knowledge be represented in a form that is computable as well as readily communicable with soil scientists so that other soil scientists can validate and update this knowledge base. This makes it necessary for the knowledge representation to approximate the mental representation of soil scientists' understanding of the soil classes and the knowledge acquisition to be based on the cognitive process suited soil scientists' understanding of soil variation.

Early implementations of SoLIM, however, did not take into consideration the cognitive aspects of knowledge formulation and representation for soil classification. In previous SoLIM applications, soil scientists were required to provide either the exact forms of fuzzy membership functions (Zhu, 1999; Zhu et al., 2001) or a large set of typical cases for known soil types in the study area (Shi et al., 2004). Soil inference was then conducted through fuzzy inference or case-based reasoning. Lacking a cognitive basis, these previous approaches place unreasonable demands on soil scientists during knowledge acquisition by requiring that they reformulate their mental knowledge into the desired form. Difficulty may also arise when the acquired knowledge needs to be interpreted or updated. This paper presents an approach that addresses this issue by employing the principles of cognitive theory in both the knowledge representation and the associated knowledge acquisition process in an effort to provide guidance to practitioners of digital soil mapping using approaches similar to SoLIM.

It has been contended that better understanding of how human beings acquire, organize, and process domain knowledge eases the difficulties in acquiring knowledge from domain experts in the development of knowledge-based systems (McCracken and Cate, 1986; Ford et al., 1991; Zhu, 1999; Özesmi and Özesmi, 2004). Especially for soil classification, it has long been suggested that it would be possible and desirable to connect the design of knowledge-based classification systems with cognitive theories (McCracken and Cate, 1986). This paper presents an approach to obtaining and representing knowledge on soil–landscape relationships based on cognitive theories on human categorization, specifically, the prototype category theory (Rosch, 1973, 1978; Smith and Medin, 1981; Lakoff, 1987; Minda and Smith, 2001). We represent and acquire the knowledge of soil–landscape model from soil scientists in terms of *prototypes* and *membership gradations* from prototypes. Variations of soil properties can then be modeled through prototype-based reasoning. The next section of this paper gives a brief introduction to prototype theory and how it differs from "classical" category theory which underpins traditional approaches to soil mapping. A novel prototype-based fuzzy soil mapping method is then presented. The advantages of this new method are illustrated and evaluated through a case study.

2. Prototype theory

Developments in cognitive psychology over the last 30 years have radically changed our understanding of how humans comprehend and describe the world through the use of categories. The act of classification—the ability to