Available online at www.sciencedirect.com

ScienceDirect

Remote Sensing of Environment 112 (2008) 2272–2283

Remote Sensing of Environment

www.elsevier.com/locate/rse

Mapping land-cover modifications over large areas: A comparison of machine learning algorithms

John Rogan [b,*], Janet Franklin [a], Doug Stow [b], Jennifer Miller [c], Curtis Woodcock [d], Dar Roberts [e]

[a] *Department of Biology, San Diego State University, San Diego CA 92182, United States*
[b] *Department of Geography, San Diego State University, San Diego CA 92182, United States*
[c] *Department of Geology and Geography and Geology, West Virginia University, Morgantown, WV 26506, United States*
[d] *Department of Geography, Boston University, Boston, MA 02215, United States*
[e] *Department of Geography, University of California Santa Barbara, Santa Barbara, CA 93106, United States*

Received 13 January 2006; received in revised form 7 October 2007; accepted 13 October 2007

Abstract

Large area land-cover monitoring scenarios, involving large volumes of data, are becoming more prevalent in remote sensing applications. Thus, there is a pressing need for increased automation in the change mapping process. The objective of this research is to compare the performance of three machine learning algorithms (MLAs); two classification tree software routines (S-plus and C4.5) and an artificial neural network (ARTMAP), in the context of mapping land-cover modifications in northern and southern California study sites between 1990/91 and 1996. Comparisons were based on several criteria: overall accuracy, sensitivity to data set size and variation, and noise. ARTMAP produced the most accurate maps overall (~84%), for two study areas — in southern and northern California, and was most resistant to training data deficiencies. The change map generated using ARTMAP has similar accuracies to a human-interpreted map produced by the U.S. Forest Service in the southern study area. ARTMAP appears to be robust and accurate for automated, large area change monitoring as it performed equally well across the diverse study areas with minimal human intervention in the classification process.

Keywords: Land-cover change; Machine learning; Large area monitoring

1. Introduction

"The Holy Grail of (digital) change detection is still total automation and high accuracy." (Loveland et al., 2002, p. 1098). Over the coming decades, the global effects of land-cover/use change may be as significant, or more so, than those associated with potential climate change (IPCC, 2000). In spite of this there is a lack of comprehensive information on the types and rates of land-cover/use change, and even less evidence of natural and anthropogenic causes and consequences of such change (Turner et al., 1999). As a result, several large area land-cover monitoring programs have been established over the past five years to comprehensively address this issue (Wulder et al., 2004). Monitoring programs, unlike most research-oriented studies, employ change mapping methods that require processing and interpretation of large volumes of *in situ*, remotely sensed and ancillary data (Cilhar 2000; Franklin & Wulder, 2002). Very large data volumes and time-consuming data processing, integration and interpretation make automated and accurate methods of change mapping highly desirable (Aspinall, 2002; Dobson & Bright, 1994; Hansen et al., 2002; Rogan & Chen, 2004).

Complex land change processes are of particular interest to researchers involved in large area monitoring (Roberts et al., 2002), where many different types of land-cover changes can occur and must be characterized (e.g., forest pest infestation, logging, wildfire, and suburbanization) (Rogan & Miller 2006). Thus, increased automation can ensure that the classification process is objective and repeatable in processing large volumes of data over complex and phenologically diverse landscapes (DeFries & Chan 2000; Gong & Xu 2003). Consequently, classification algorithm selection and performance have become

* Corresponding author. Current address: Department of Geography, Clark University, 950 Main Street, Worcester, MA 01610, United States. Tel.: +1 508 793 7562.
E-mail address: jrogan@clarku.edu (J. Rogan).

0034-4257/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2007.10.004

particularly important, because large area change monitoring can only realistically be achieved (i.e., low cost and generalizable results) through techniques that minimize time-consuming human interpretation and maximize automated procedures for data analysis (Woodcock et al., 2001).

A wide variety of classification algorithms have been used to map land-cover/use and changes to land-cover/use, from remotely sensed data. Unsupervised classification and cluster labeling is the dominant method for large area land-cover mapping and monitoring (see Wulder et al., 2004). Recently, machine learning algorithms (MLAs) have emerged as more accurate and efficient alternatives to conventional parametric algorithms (e.g., maximum likelihood), when faced with large data volumes and complex measurement spaces (Foody, 1995; Muchoney & Williamson 2001; Kasischke et al., 2004). The effectiveness of MLAs has been demonstrated primarily in single-date land-cover mapping studies (Friedl et al., 1999; Pal & Mather 2003). However, their effectiveness in change mapping has also recently been addressed (Gopal & Woodcock 1996; Liu & Lathrop 2002; Chan & Chan 2002).

The objective of this research was to compare the performance of three MLAs; two classification tree software routines (S-Plus and C4.5) and an artificial neural network (ARTMAP), in the context of mapping land-cover modifications in southern and northern California between 1990/91 and 1996. These algorithms were specifically chosen because they are increasingly used in land-cover mapping and monitoring using medium-coarse spatial resolution remotely sensed data (Rogan & Chen 2004), yet have not been compared with each other in a large area, large data volume context. This study is based on the US Forest Service and California Department of Forestry and Fire Protection statewide land-cover mapping and monitoring program (LCMMP) in California. LCMMP personnel required an evaluation of available machine learning algorithms to determine the feasibility of automating their land-cover change detection procedures. The current LCMMP is based on time-consuming unsupervised classification and cluster labeling (Levien et al., 1999).

Algorithm comparison followed two steps. Step one assessed how algorithm performance was affected by modifications in the model data. Performance was assessed using classification accuracy measures and the model data set was modified to test the following: (1) effect of partitioning data sets; (2) effect of training data set size; and (3) the effect of training data errors (i.e., noise). Step two compared a map of land-cover change, generated using the algorithm deemed best overall from step one, to a map produced by the LCMMP that was created through unsupervised classification and cluster labeling by human interpreters. The results of this work are intended to inform the growing number of resource managers engaged in other operational land-cover monitoring projects.

2. Background

2.1. Machine learning algorithms in land-cover change mapping

Machine learning refers to induction algorithms that analyze information, recognize patterns, and improve prediction accuracy through automated, repeated learning from training data (Malerba et al., 2001). There is now a large body of research that demonstrates the abilities of machine learning techniques, particularly classification trees and artificial neural networks, to deal effectively with tasks involving high dimensional data (Gahegan 2003). The increased interest in MLAs can be attributed to several factors:

- their non-parametric nature deals well with multi-modal, noisy and missing data (Hastie et al. 2001), but see Simard et al. (2000) in the case of classification trees;
- there is a significant reduction in computational demands when data measurement spaces are large and complex (Foody 2003);
- they readily accommodate both categorical and continuous ancillary data (Lawrence & Wright 2001);
- users can investigate the relative importance of input variables in terms of contribution to classification accuracy (Hansen et al., 1996; Foody & Arora 1997);
- they are flexible and can be adapted to improve performance for particular problems (Lees & Ritman 1991)
- and multiple subcategories per response variable can be accommodated (Gopal et al., 1999).

However, while promising, the above assertions have not been tested robustly in the context of land-cover change mapping. As a result, MLAs have yet to be fully incorporated in large area studies of land-cover change monitoring (Rogan & Miller 2006).

2.1.1. Classification trees and neural networks

Classification trees are a type of MLA used to predict membership of cases of a categorical dependent variable from their measurements on one or more predictor variables (De'ath & Fabricius 2000). Classification trees are developed using different measures that recursively split data sets into increasingly homogeneous subsets representing class membership. All classification tree approaches employ hierarchical, recursive partitioning of the data, resulting in decision rules that relate values or thresholds in the predictor variables with pixel classes (Friedl & Brodley 1997). An important advantage of classification trees is that they are structurally explicit, allowing for clear interpretation of the links between the dependent variable of class membership and the independent variables of remote sensing and/or ancillary data (Lawrence & Wright 2001).

Generally, a neural network learns a pattern by iteratively considering each training observation and then multiplying the explanatory variables by a set of weights, applying a set of transfer functions to their weighted sum, and finally predicting membership for each desired map class (Franklin et al., 2003). There are many different types of neural network algorithms available including multi-layer perceptrons, learning vector quantization, Hopfield, and Kohonen Self Organizing Maps. However, a large degree of uncertainty remains as to which network algorithm works best for a specific application (Borak & Strahler 1999).

MLA applications and comparisons in change mapping are relatively uncommon. Past studies range in complexity from