Available online at www.sciencedirect.com
ScienceDirect
Remote Sensing of Environment 112 (2008) 1226–1241
Remote Sensing of Environment
www.elsevier.com/locate/rse

Extending post-classification change detection using semantic similarity metrics to overcome class heterogeneity: A study of 1992 and 2001 U.S. National Land Cover Database changes

Ola Ahlqvist *

The Ohio State University, Department of Geography, 1049 Derby Hall, 154 N Oval Mall, Columbus, OH 43210, USA

Received 1 March 2007; received in revised form 9 June 2007; accepted 5 August 2007

Abstract

The use of post-classification change methods for the analysis of land cover change provides intuitive and potentially reliable results. A recurring problem is the difference in land cover nomenclature that can occur over time or across space when multiple data sources are required. Building on work that uses category semantics as a foundation for reasoning with land cover classes, this paper uses a fuzzy sets based approach to develop attribute based prototype definitions of land cover classes. These formalized category descriptions are used to look at land cover changes as a semantic change evaluated through semantic similarity metrics. The methodology is illustrated on the U.S. National Land Cover Data from 1992 to 2001 over Chester County, PA. The results demonstrate that the proposed method is more versatile than the standard post-classification method in that it can both provide an overall, spatially explicit evaluation of land cover change, as well as nuanced assessments of graded changes for heterogeneous land cover types.

Keywords: Land cover change; Land use change; Category semantics; Classification; Fuzzy; Uncertainty

1. Introduction

In 1993 a group of federal agencies formed the Multi-Resolution Land Characteristics (MRLC) Consortium to meet the needs of Federal agencies for consistent land-cover data covering the entire United States, and two data sets called the National Land Cover Dataset (NLCD 1992) (USGS, 2006b) and the National Land Cover Database (NLCD 2001) (USGS, 2006a) have been produced. With the recent completion of the 2001 Land Cover Dataset many users are likely to look for changes in the landscape since the mapping from 1992. Although none of these data sets were primarily intended for per-pixel based analysis at a local level (Vogelman et al., 2001) a reasonable assumption would be that these and similar data could be used to indicate regional as well as nation wide land use and land cover changes over time. However, slight to moderate differences in class definitions between the land cover classes in the 1992 and 2001 data makes a straightforward post-classification change analysis practically impossible without an alternative approach. The purpose of this paper is to suggest how recent work on information semantics can be applied to remote sensing data classifications for landscape change analysis in situations like these.

This introductory section provides a summary of post-classification landscape change analysis methods and how fuzzy and semantically informed approaches recently have been suggested as an alternative to tackle some of the problems associated with category uncertainty. A Data and methods section then briefly describes the National Land Cover Database for the United States and identifies class heterogeneities as a specific challenge to use it for a straightforward change assessment. The Methods section also describes the rationale behind semantic similarity metrics and how they can be applied to a post-classification land cover change assessment. Section 3 presents a case study of the suggested approach where some of the semantic change analysis possibilities are demonstrated. Section 4 provides some concluding comments. Thus, this paper does not seek to evaluate the actual land cover changes in the study area, but rather illustrate how the proposed methodology can overcome the major hurdle of incompatible classification systems and still provide a

* Corresponding author. Tel.: +1 614 247 7997.
E-mail address: ahlqvist.1@osu.edu.

0034-4257/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2007.08.012

quantitative and spatially explicit picture of overall and detailed landscape change.

1.1. Representation and analysis of category semantics

There are a large number of methods to analyze landscape change (Lu et al., 2004; Mas, 1999; Singh, 1989). Mas (1999) separate between three general methods for change detection using satellite images; image enhancement including numerical combination of image bands using e.g. PCA, one-step classification of multi temporal data, and post-classification methods where two independent land cover classifications are compared. One of the most commonly used techniques is the post-classification approach (Foody & Boyd, 1999; Lu et al., 2004). Given that it is possible to overcome issues of time consumption and required expert knowledge to produce reliable land cover classifications, the main advantages of this method is the detailed information that can be gained from the produced change matrix and the limited impact that image calibration, atmospheric and environmental differences will have on the multi-temporal image comparison (Lu et al., 2004). A further advantage of the post-classification approach is its intuitive interpretation as opposed to numerically based image analysis methods that need careful interpretation to asses what the identified changes mean. This advantage is mainly due to the rich semantics of land cover class labels, but the semantics is also noted by several authors as problematic due to the usually limited descriptions of exactly how land cover labels should be understood (Comber et al., 2004b). In many situations certain types of land use or land cover change is more or less important so that either original classes need to be reclassified into more relevant categories for the analysis or the change analysis on original categories need to be further analyzed and weighted depending on the importance put on certain types of change. Additionally we often find that data on land use and land cover from different times are classified using different classification systems (Comber et al., 2004a). In these situations a standard post-classification change assessment can be very complicated if not impossible.

Many scholars have acknowledged a need to negotiate and compare information from different origins, such as data that use different classification systems. The work on semantic uncertainty (Salge, 1995) and semantic interoperability (Bishr, 1998) of geographic information reflect this concern. Work in computer science, artificial intelligence and information science have also tackled the issue of translations between heterogeneous information sources and many see a potential for using formalized descriptions, or ontologies, that can describe category semantics, to address this (Fonseca et al., 2002; Guarino, 1995; Sowa, 2000). The two main approaches to formally represent category semantics that have emerged out of that work use either formal *axioms*, such as structured taxonomies of classes and sub-classes, or *prototypes* where objects are described using a set of characteristic properties. For many natural resource classification systems both of these representational approaches are feasible. A land cover system often includes several levels that are organized in a tree-like hierarchy of classes and subclasses (c.f. Anderson, 1976) that can readily be represented formally using the axiomatic approach. At the same time there are often more or less detailed descriptions of the classes in terms of characteristic attributes such as “>70% vegetative cover”, “ trees generally make up more than 20% of the vegetation cover”. Such statements can be turned into quantitative properties in a formal prototype definition. Once a classification scheme has been transformed into a formalized categorization a translation can be achieved by matching the concepts in one system with concepts in another, either directly or through an intermediate classification. Conceptually, these computational approaches largely follow suggestions from the cognitive sciences on how categories are mentally constructed and this has informed some recent examples of negotiating different nomenclatures including some that specifically target incompatible land use and land cover taxonomies (Ahlqvist, 2005; Feng & Flewelling, 2004; Kavouras & Kokla, 2002).

1.2. Graded notions of classification

Traditional post-classification change analysis follows a straightforward methodology of image overlay where pixels have been unambiguously categorized into one of many alternative classes. The resulting change image together with a cross-tabulated change matrix typically provides quantitative summaries of differences between the two images. The logic applied in this analysis is usually based on traditional set theory, and spatial visualization of change is often restricted to single category maps, e.g. change from non-urban to urban development (Yuan et al., 2005), or summarizing a fairly limited number of categories e.g. Anderson level 1 (Loveland et al., 2002). Because post-classification change detection is essentially a binary analysis, evaluated as either change or no change, it does not differentiate between subtle and more dramatic changes. For example a subtle change from “row crops” to “pasture” is treated equally as a drastic change from “row crops” to “strip mine”. Many scholars have pointed out that this simplistic treatment of information in classified pixels is both problematic and inaccurate, and several alternative ways of understanding the content of a pixel have emerged (Pontius & Cheuk, 2006).

One common conceptualization is founded on the understanding of a pixel class as a statement of the probability that any given crisp class is the true class. Consequently, this approach uses conventional probabilistic reasoning instead of the standard set theory. Note that the probabilistic reasoning is still treating the classes as crisp and unambiguous, but the reasoning allows for uncertainty in terms of “what is the chance that class X is present here?” A different mode of reasoning is when the notion of vagueness in a classification is considered. In practical remote sensing applications we often encounter borderline cases; ecotones are neither of the bordering categories, and mixes of trees can make neither “deciduous” nor “evergreen” a correct label. In these cases probabilistic reasoning does not help, instead the theory of fuzzy sets (Zadeh, 1965) has been proposed as a way to handle vagueness (Fisher & Pathirana, 1990). Each object or pixel is allowed to have a degree of membership to any one class so that one place can be assigned to several classes with partial membership. Yet another understanding of pixels, potentially also including probabilistic and/or fuzzy reasoning is that of mixed, sub-pixel or multi-resolution pixels (Foody & Cox, 1994). Under this interpretation a pixel is recognized to be a mixture of different things and the proportions of those things determine the class assignment. Obviously both the sub-pixel entities as well as the final pixel class could be stated crisply, as a probability, or as a fuzzy membership.

So far most work on these 'soft' remote sensing classifications have looked at analyses confined within the same classification where the classes have been defined as mutually exclusive, albeit with a possibility of confusion or error. Thus the uncertainty handling has so far mostly addressed the relationship between an observation and a target category. In contrast, the notion of category semantics and metrics of category similarity, introduced in the previous section, is concerned with the vagueness inherent in category definitions and semantic relations between categories. These relations are specifically relevant in situations where available data use heterogeneous classification systems.

1.3. Existing approaches to semantic analysis of land cover change

Some recent work have proposed semantically grounded approaches that acknowledge the graded nature of land use land cover categories and their use in change analysis. For example, Comber et al. (2004a) used the notion of *inconsistency* applied to a land cover change analysis based on crisp classes but where some ancillary data contained underlying probabilistic information on alternative classes. Their main concern however was to reconcile different classification systems in the two data sets they used. They employed a knowledge engineering approach based on semantic lookup tables generated from expert input that evaluated whether the class found in the most recent data set could be "expected" based on what it was classified as at an earlier point in time given that no change had occurred. If an "unexpected" relationship was found this could be caused either by one or both classifications having an error, or there had been a change in the landscape between the two recordings. Thus, their end result seeks to identify a crisp "change" or "no change" situation for every location.

In a similar vein Fritz and See (2005) also used a lookup table approach to compare the Global Land Cover 2000 data with the MODIS global land cover product. They had experts evaluate the "degree of difficulty in distinguishing between" any two land cover classes. It is not entirely clear from their reporting whether this relates to the difficulty to distinguish classes spectrally (a probabilistic uncertainty) or as an interpreter in the field (a fuzzy uncertainty). They generally followed the methodology developed by Gopal and Woodcock (1994) and Woodcock and Gopal (2000) but they also generated a spatially explicit picture of the agreement between the two data sets on a graded scale from complete agreement to severe disagreement. A similar approach is also used in Jung et al. (2006) to combine three incompatible global land cover data sets into one "best estimate" map.

The following section seeks to further extend this research direction by interpreting land cover categories as continuous phenomena with classes being more or less similar and as a consequence argue that land cover change also can be interpreted in a continuous fashion. Building on previous research by Ahlqvist (2004) the method proposed in this paper provides a cognitive foundation for a detailed description of land cover category semantics. A semantic analysis of change images is then used to identify land cover change on a continuous scale, from subtle to drastic, based on calculated semantic relationships between land cover classes.

2. Data and method description

The proposed method is demonstrated through an applied example that uses publicly available U.S. land cover data sets from 1992 to 2001 described below. This section also provides a brief description of the Uncertain Conceptual Spaces method (Ahlqvist, 2004) used here to formalize land cover category definitions, and then introduces the use of semantic similarity metrics as a measure of land cover change.

2.1. Data and study area

The demonstration example uses data from Chester County, located in southeastern Pennsylvania bordering with the states of Maryland to the south and Delaware to the southeast. The landscape is largely agricultural in the western parts of the county and more urban in the east. According to Arthur et al. (2000) there is continued suburban expansion, influenced by the Philadelphia metropolitan region and bedroom communities in Delaware. This process has increased the proportion of urban areas from about 11% in 1987 to around 19% in 1996 with accompanying decreases in forest and agricultural areas. Hence, an assessment of land cover changes using data from 1992 to 2001 should display significant changes in this region, particularly related to urban expansion. The NLCD 1992 and 2001 land cover data sets were downloaded from the USGS seamless server (http://seamless.usgs.gov/) and reprojected to an Albers Equal Area conic projection, Fig. 1. Associated metadata was also acquired to guide the development of formalized semantic descriptions of each land cover class in the two data sets as detailed in the next section.

Additional USGS Digital Orthophoto (quarter–quadrangle) imagery was downloaded from the Pennsylvania Spatial Data Access portal (www.pasda.psu.edu) to support interpretation and illustrations in the Results section.

2.2. Semantic representation and similarity analysis

Detailed formal descriptions of the NLCD 1992 and 2001 land use and land cover categories were developed based on an evaluation of class definitions provided in the metadata accompanying each data set. A compilation of these class definitions is provided in Appendix A. The formalization was done using the Uncertain Conceptual Spaces approach which follows the idea of a prototype representation where each land cover class is defined by a set of characteristic attributes, Table 1.

The choice of descriptive attributes was informed by the classifiers used in the Food and Agriculture Organization of the United Nations Land Cover Classification System, LCCS (Di

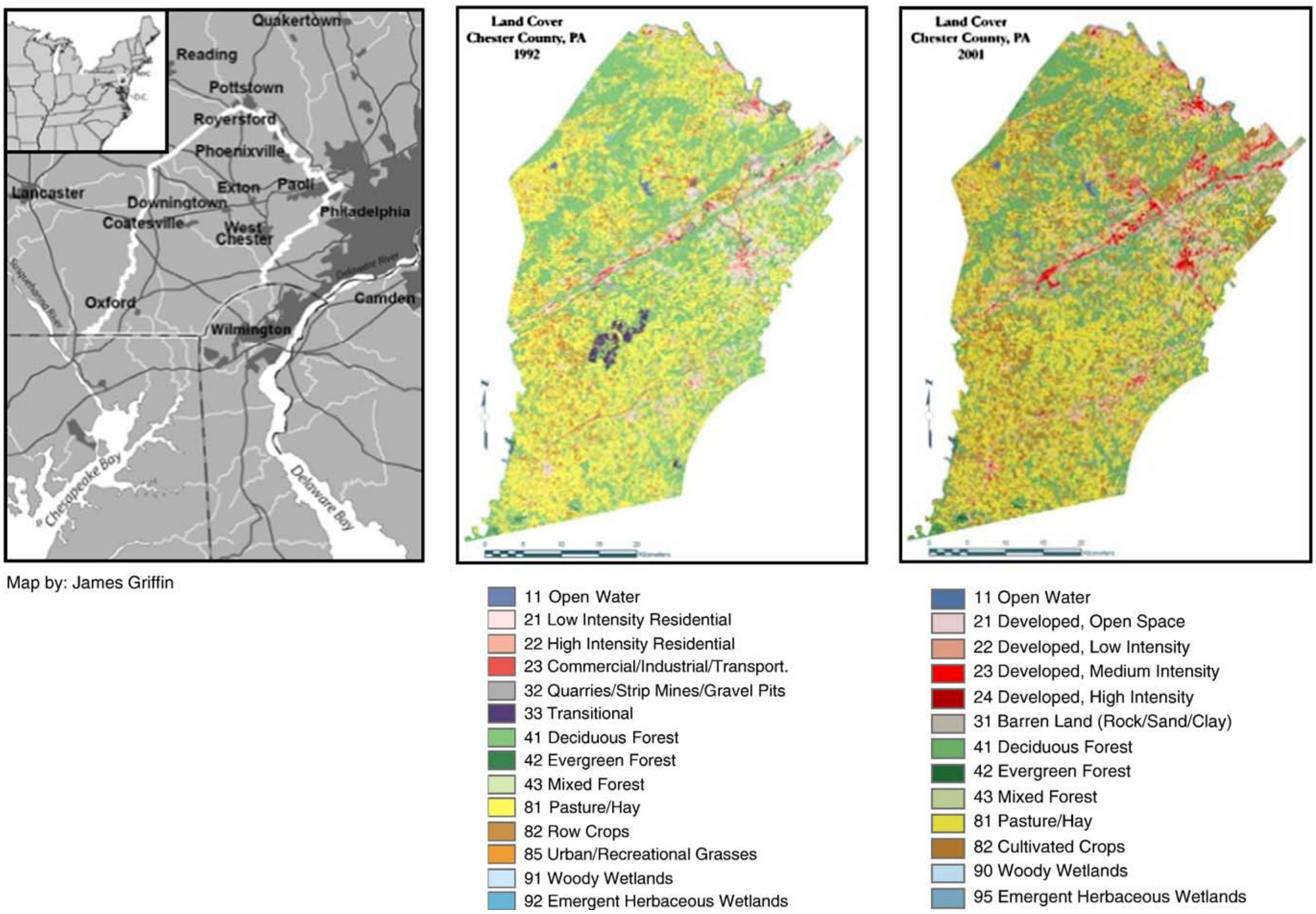

Fig. 1. Study area overview and the two land cover data sets, NLCD 1992 (left) and NLCD 2001 (right), used in the change analysis.

Gregorio & Jansen, 2000) but with modifications in the formal implementation to represent attribute values as fuzzy numbers (Kaufmann & Gupta, 1985) that enable an explicit representation of vagueness at the attribute level (Ahlqvist, 2005). This representation makes it possible both to define attributes as intervals and to attach membership values that describe uncertainty to all or portions of an attribute value. For example, say that a forest category is defined as a natural land cover with trees covering >70–60% of the surface. The attribute Woody tenure is set to 1, full membership, for the "natural" value, and 0, no membership, for the "cultivated" value. The Tree cover attribute values are set so that the interval from 60 to 100% cover has a membership value of 1, and the interval from 0 to 60% cover is set to 0. There is also a possibility to define a fuzzy transition e.g. from 60 to 70% to acknowledge the vague character of that threshold. Graphically this definition can be illustrated as the first row of Fig. 2 where the two first attributes % tree cover and tenure are assigned membership functions according to the mentioned class description. Additional attributes are possible to add in order to further specify each class.

As other classes get specified it is possible to compare them, attribute by attribute, in a systematic fashion. Hence, each category description can be thought of in terms of a multi-dimensional attribute space made up of a collection of defining attribute domains, such as vegetation cover, tenure, water cover etc. in which a concept definition occupies a certain region. The main difference between this and a standard multi-dimensional data space is the fuzzy nature of the uncertain conceptual space occupied by a category. Moreover, for any concept definition, each attribute can be assigned a certain importance, or salience, in relation to other properties of the concept. This enables us to

Table 1
List of descriptive attributes to define the land cover classes in NLCD 1992 and 2001 data

Characteristic attribute	Values
Water cover %	[0,100] %
Water phase	{ice, water}
Water persistence	{permanent, periodically, waterlogged}
Impervious cover	[0,100] %
Vegetation cover	[0,100] %
Development type	{residential, commercial, mining}
Surface Type	{earth, constructed}
Tree cover	[0,100] %
Tree height	[0,20) m
Deciduous cover	[0,100] %
Evergreen cover	[0,100] %
Shrub cover	[0,100] %
Grass/herb cover	[0,100] %
Woody tenure	{natural, cultivated}
Grass/herb tenure	{natural, cultivated}
Crop type	{Row crops, Small grains, Fallow, Hay, Grass}

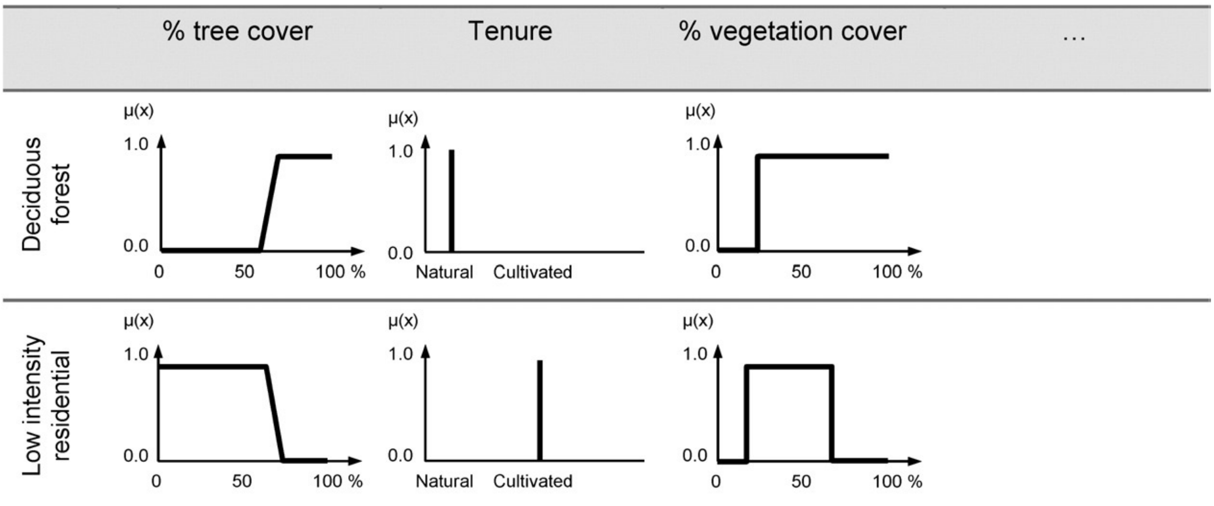

Fig. 2. Schematic illustration of two land cover classes with their attribute definitions specified as fuzzy membership functions.

declare some properties as more important than others for defining a concept. In some respect the suggested technique is similar to the algebraic techniques of image differencing and change vector analysis where a numerical difference between images is used as an indicator of change (Lambin & Strahler, 1994). In fact, the change vector analysis also uses several indicators to calculate the overall difference between two sets (vectors) of indicators. A major difference though is that the semantic similarity assessment is done on fuzzy definitions of classes, not empirical spectral values.

Because of the close structural resemblance between a standard multivariate data space and a multivariate conceptual space we find many analytical parallels. For example, commonly used multi-variate distance metrics such as the Euclidean and city-block metric have been suggested as semantic metrics that can be used to evaluate category similarity (Gärdenfors, 2000; Nosofsky, 1986; Shepard, 1987). A characteristic trait of the cognitive literature is the awareness of context and asymmetry effects in evaluations of category similarity such that people would respond differently to the statement “a forest is similar to an estuarine woody wetland” than they would to “an estuarine woody wetland is similar to a forest”. Many metrics of semantic similarity tries to take this into account and make evaluations in both “directions” for any two categories.

I will use two metrics that address different aspects of semantic similarity between the land cover classes; an overlap and a distance metric. The overlap metric is calculated using the following equations:

$$o(a_A, a_B) = \int \min(f_{P_A}(x), f_{P_B}(x))\mathrm{d}x \Big/ \int f_{P_B}(x)\mathrm{d}x. \quad (1)$$

$$O(C_A, C_B) = \sqrt{\sum_i w_{B_i} * o(a_{A_i}, a_{B_i})^2}. \quad (2)$$

In Eq. (1), the attribute value overlap, o, is measured as the overlap of two fuzzy functions $f_{PA}(x)$ and $f_{PB}(x)$, each defining attribute values a_A, a_B for land cover class A and B respectively.

The class *overlap* metric, O, in Eq. (2), is then a weighted average of the attribute specific overlaps from Eq. (1). The class overlap metric applies the perspective of land cover class B by using $f_{PB}(x)$ as the denominator in Eq. (1) and by applying weights, W_B, in Eq. (2), used to assign different importance to the defining attributes. The weights are scaled to sum to one, $\sum_i W_{Bi} = 1$, so that classes that overlap a lot will have values close or equal to 1 and non-overlapping classes will have a value of 0.

The class *distance* metric follows a similar approach, but here we employ a Euclidean distance metric formalized as:

$$d(C_A, C_B) = \sqrt{\sum_i w_{B_i}(a_{A_i} - a_{B_i})^2}, \quad (3)$$

where we calculate the difference $a_{Ai} - a_{Bi}$ using the fuzzy dissemblance index (Kaufmann & Gupta, 1985) that calculates the distance between two membership functions. Again the weights are adjusted to sum to 1, $\sum_i W_{Bi} = 1$, so that identical land cover classes will have a distance of 0 and very different classes will have values approaching 1. Fig. 3 provides a graphic illustration of how the two metrics are applied to two fuzzy membership functions defined on an example attribute domain.

The class distance and overlap metrics are obviously an indication of how similar two classes are in their formalized definitions. In addition, the overlap metric can identify degrees of class–subclass relations. If for example land cover class A fully

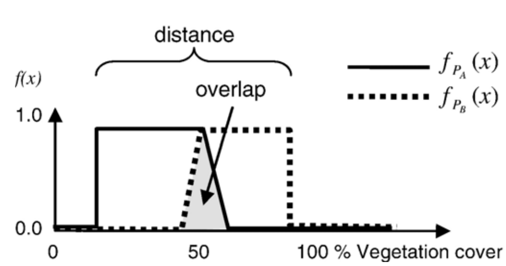

Fig. 3. Graphic illustration of the distance (bracket) and overlap (shaded area) metrics measured on one descriptive attribute, vegetation cover %.

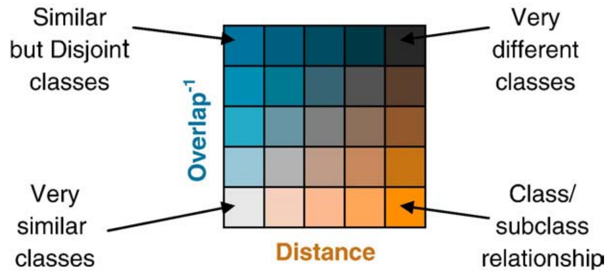

Fig. 4. Bi-variate color scheme used to symbolize the semantic distance and overlap metrics. Increasing class distance render more intense orange, decreasing overlap render increasing blue, and combinations thereof blend to increasingly darker gray. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

overlaps class B, and B partially overlaps A, it would indicate that B can be held a sub-class of A. Since the two metrics provide different aspects of semantic relations they should be used together to enable a better characterization of the semantic relationship between categories. In this paper I will use a bi-variate color scheme illustrated in Fig. 4 to simultaneously symbolize both metrics in subsequent maps and tables. It mixes shades of orange and blue so that four general types of semantic relations, revealed by the two similarity metrics, can be distinguished.

Generally this color scale means that we symbolize larger difference as increasingly saturated blue or orange hues or gray mixes. Thus, in the following illustrations, significant land cover changes (large semantic difference) will show as dark gray. Intermediate changes will show as more or less saturated blue, and the orange tones can be either an actual change or a change of the classification so that the current land cover is a subclass of the previously recorded land cover. No-change situations will show as very pale colors or no color at all.

Another useful visual summary of the similarity relationships will be provided by a spatialization of the semantic distance matrix. Spatialization is often used as a technique to reduce the dimensionality of a multivariate data space (Skupin & Fabrikant, 2003). The multi-dimensional scaling program PROXCAL in the SPSS statistical software (SPSS Inc., 2005) was therefore used to summarize the multi-dimensional conceptual space by using the matrix of pair wise distance values as input. A reduction to two dimensions was generated such that classes can be represented as points in a standard, two-dimensional scatter plot, and relative distances between class points correspond roughly to the relative semantic difference between classes.

2.3. Fuzzy change estimations

Although post-classification change analysis can present estimates of the area of change and the rate of change, there is a debate about how to best translate fuzzy statements about class membership into for example an area estimate. Consider one 30×30 m pixel classified as “Low intensity residential” that changes into “Developed, medium intensity”. While the standard, crisp change logic will count this pixel as changed, a fuzzy interpretation of change will somehow view this as only a slight change based on class similarity. Both Fisher et al. (2006) and Pontius and Cheuk (2006) point out that the interpretation of a change matrix under the assumption of fuzzy categories will differ from the standard one where diagonal elements hold instances of no change and off diagonal elements hold instances of category gains and losses. The question is then what metric and what type of logic an area calculation based on similarity metrics should follow. Woodcock and Gopal (2000) as well as Comber et al. (2004a) used different fuzzy and belief values thresholds respectively to calculate total areas above a certain level. In the study by Fritz and See (2005), fuzzy metrics of semantic similarity was used to produce an overall agreement measure between two maps. As long as the basis and logics for calculated area estimates are given in detail these differences need not be a big problem and it should be possible to provide robust metrics of land cover change using fuzzy values.

However, change assessment based on heterogeneous classification systems impose yet one important restriction that makes standard use of a change matrix invalid. The main diagonal

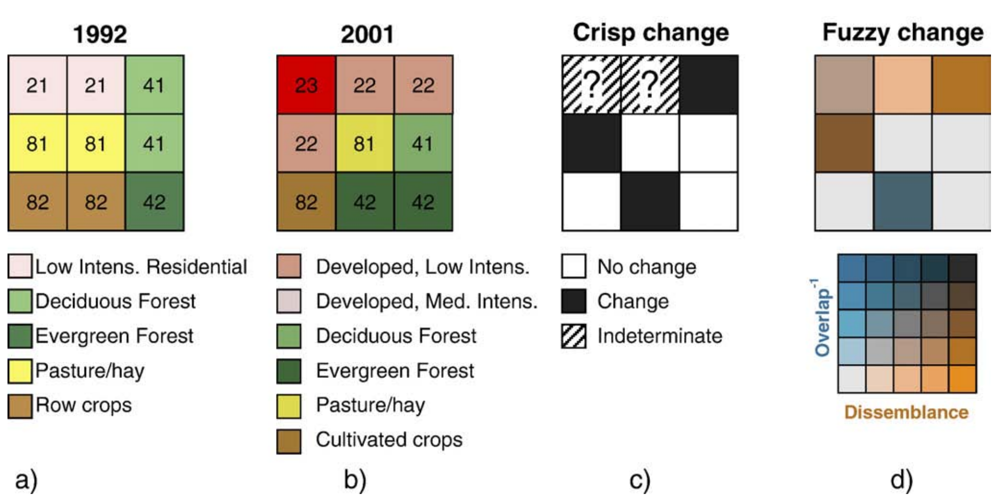

Fig. 5. Hypothetical 3×3 pixel land cover change scenario from a) 1992 to b) 2001, with evaluation of change using c) crisp change detection based on matching classes and d) fuzzy change detection based on bi-variate semantic metrics.

Table 2
Cross tabulated semantic distance values for land cover classes in NLCD 1992 and 2001 data

	1992 Classes																				
1992 Classes	11	12	21	22	23	31	32	33	41	42	43	51	61	71	81	82	83	84	85	91	92
11	0	0.5	0.58	0.6	0.6	0.55	0.55	0.55	0.62	0.62	0.62	0.58	0.6	0.58	0.58	0.58	0.58	0.55	0.58	0.3	0.19
12	0.5	0	0.58	0.6	0.6	0.55	0.55	0.55	0.62	0.62	0.62	0.58	0.6	0.58	0.58	0.58	0.58	0.55	0.58	0.58	0.54
21	0.68	0.68	0	0.19	0.48	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.66	0.66
22	0.7	0.7	0.19	0	0.45	0.67	0.67	0.67	0.72	0.72	0.72	0.69	0.71	0.69	0.69	0.69	0.69	0.67	0.69	0.73	0.7
23	0.7	0.7	0.48	0.45	0	0.67	0.67	0.67	0.72	0.72	0.72	0.69	0.71	0.69	0.69	0.69	0.69	0.67	0.69	0.73	0.7
31	0.55	0.55	0.53	0.56	0.56	0	0	0	0.57	0.57	0.57	0.53	0.55	0.53	0.52	0.53	0.53	0.03	0.53	0.58	0.54
32	0.66	0.66	0.65	0.67	0.67	0.45	0	0.45	0.68	0.68	0.68	0.65	0.67	0.65	0.65	0.65	0.65	0.45	0.65	0.69	0.66
33	0.26	0.26	0.21	0.29	0.29	0	0	0	0.32	0.32	0.32	0.18	0.27	0.18	0.18	0.18	0.18	0.04	0.18	0.34	0.22
41	0.51	0.51	0.45	0.52	0.52	0.47	0.47	0.47	0	0.38	0.19	0.32	0.43	0.27	0.46	0.46	0.46	0.47	0.46	0.22	0.28
42	0.51	0.51	0.45	0.52	0.52	0.47	0.47	0.47	0.38	0	0.19	0.32	0.43	0.27	0.46	0.46	0.46	0.47	0.46	0.22	0.28
43	0.48	0.48	0.42	0.5	0.5	0.45	0.45	0.45	0.19	0.19	0	0.28	0.4	0.22	0.43	0.44	0.44	0.45	0.44	0.16	0.24
51	0.52	0.52	0.47	0.54	0.54	0.49	0.49	0.49	0.27	0.27	0.27	0	0.46	0.24	0.47	0.47	0.47	0.48	0.47	0.26	0.26
61	0.6	0.6	0.52	0.62	0.62	0.55	0.55	0.55	0.5	0.5	0.5	0.53	0	0.53	0.53	0.53	0.53	0.55	0.53	0.52	0.55
71	0.55	0.55	0.48	0.57	0.57	0.57	0.51	0.51	0.51	0.28	0.28	0.24	0.48	0	0.41	0.41	0.41	0.49	0.41	0.32	0.1
81	0.63	0.63	0.58	0.65	0.65	0.61	0.61	0.61	0.59	0.59	0.59	0.58	0.58	0.54	0	0.38	0.38	0.45	0.38	0.61	0.54
82	0.63	0.63	0.58	0.65	0.65	0.61	0.61	0.61	0.6	0.6	0.6	0.58	0.59	0.54	0.38	0	0.38	0.45	0.38	0.61	0.54
83	0.63	0.63	0.58	0.65	0.65	0.61	0.61	0.61	0.6	0.6	0.6	0.58	0.59	0.54	0.38	0.38	0	0.45	0.38	0.61	0.54
84	0.63	0.63	0.63	0.64	0.64	0.5	0.5	0.5	0.66	0.66	0.66	0.65	0.65	0.66	0.55	0.55	0.55	0	0.55	0.67	0.66
85	0.69	0.69	0.65	0.7	0.7	0.67	0.67	0.67	0.66	0.66	0.66	0.64	0.64	0.61	0.5	0.5	0.5	0.55	0		
91	0.6	0.71	0.68	0.75	0.75	0.71	0.71	0.71	0.54	0.54	0.54	0.57	0.66	0.6	0.71	0.71	0.71	0.71	0.71	0	0.25
92	0.57	0.67	0.65	0.7	0.7	0.67	0.67	0.67	0.56	0.56	0.56	0.55	0.66	0.51	0.62	0.62	0.62	0.66	0.62	0.27	0

	2001 Classes											
1992 Classes	11	12	21	22	23	24	31	41	42	43	52	71
11	0	0.5	0.62	0.59	0.58	0.6	0.55	0.62	0.62	0.62	0.58	0.58
12	0.5	0	0.62	0.59	0.58	0.6	0.55	0.62	0.62	0.62	0.58	0.58
21	0.68	0.68	0.49	0.46	0.45	0.48	0.66	0.65	0.65	0.65	0.65	0.66
22	0.7	0.7	0.58	0.52	0.46	0.45	0.67	0.72	0.72	0.72	0.69	0.69
23	0.7	0.7	0.58	0.52	0.46	0.45	0.67	0.72	0.72	0.72	0.69	0.69
31	0.55	0.55	0.57	0.54	0.53	0.56	0.05	0.56	0.56	0.56	0.52	0.52
32	0.66	0.66	0.68	0.66	0.65	0.67	0.04	0.67	0.67	0.67	0.65	0.65
33	0.26	0.26	0.33	0.25	0.2	0.29	0.05	0.3	0.3	0.3	0.18	0.18
41	0.51	0.51	0.44	0.44	0.47	0.52	0.49	0.02	0.38	0.19	0.32	0.27
42	0.51	0.51	0.44	0.44	0.47	0.52	0.49	0.38	0.02	0.19	0.32	0.27
43	0.48	0.48	0.41	0.41	0.44	0.5	0.47	0.19	0.19	0.02	0.29	0.23
51	0.52	0.52	0.46	0.46	0.48	0.54	0.51	0.26	0.26	0.26	0.02	0.25
61	0.6	0.6	0.11	0.11	0.22	0.36	0.57	0.5	0.5	0.5	0.53	0.54
71	0.55	0.55	0.48	0.48	0.5	0.57	0.54	0.28	0.28	0.28	0.23	0.03
81	0.63	0.63	0.44	0.44	0.47	0.52	0.63	0.59	0.59	0.59	0.58	0.54
82	0.63	0.63	0.44	0.45	0.47	0.52	0.62	0.59	0.59	0.59	0.58	0.54
83	0.63	0.63	0.44	0.45	0.47	0.52	0.62	0.59	0.59	0.59	0.58	0.54
84	0.63	0.63	0.56	0.53	0.52	0.53	0.51	0.66	0.66	0.66	0.65	0.66
85	0.69	0.69	0.21	0.41	0.44	0.49	0.68	0.66	0.66	0.66	0.64	0.61
91	0.6	0.71	0.67	0.67	0.7	0.75	0.73	0.55	0.55	0.55	0.57	0.61
92	0.57	0.67	0.65	0.66	0.67	0.7	0.69	0.56	0.56	0.56	0.55	0.51

	1992 Classes																				
2001 Classes	11	12	21	22	23	31	32	33	41	42	43	51	61	71	81	82	83	84	85	91	92
11	0	0.5	0.58	0.6	0.6	0.55	0.55	0.55	0.62	0.62	0.62	0.58	0.6	0.58	0.58	0.58	0.58	0.55	0.58	0.3	0.19
12	0.5	0	0.58	0.6	0.6	0.55	0.55	0.55	0.62	0.62	0.62	0.58	0.6	0.58	0.58	0.58	0.58	0.55	0.58	0.58	0.54
21	0.71	0.71	0.49	0.58	.58	0.68	0.68	0.68	0.65	0.65	0.65	0.65	0.64	0.65	0.65	0.65	0.65	0.68	0.47	0.65	0.66
22	0.69	0.69	0.46	0.52	0.52	0.66	0.66	0.66	0.64	0.64	0.64	0.65	0.64	0.65	0.65	0.65	0.65	0.66	0.47	0.65	0.66
23	0.68	0.68	0.45	0.46	0.46	0.65	0.65	0.65	0.67	0.67	0.67	0.66	0.66	0.66	0.66	0.66	0.66	0.65	0.48	0.68	0.66
24	0.7	0.7	0.48	0.45	0.45	0.67	0.67	0.67	0.72	0.72	0.72	0.69	0.71	0.69	0.69	0.69	0.69	0.67	0.52	0.73	0.7
31	0.66	0.66	0.66	0.67	0.67	0.45	0.04	0.45	0.69	0.69	0.69	0.65	0.68	0.65	0.66	0.65	0.65	0.45	0.65	0.7	0.66
41	0.5	0.5	0.45	0.52	0.52	0.47	0.47	0.47	0.02	0.38	0.19	0.31	0.43	0.26	0.46	0.46	0.46	0.47	0.46	0.22	0.28
42	0.5	0.5	0.45	0.52	0.52	0.47	0.47	0.47	0.38	0.02	0.19	0.31	0.43	0.26	0.46	0.46	0.46	0.47	0.46	0.22	0.28
43	0.48	0.48	0.42	0.5	0.5	0.44	0.44	0.44	0.19	0.19	0.02	0.27	0.4	0.21	0.43	0.43	0.43	0.44	0.43	0.16	0.23
52	0.52	0.52	0.47	0.54	0.54	0.48	0.48	0.48	0.28	0.28	0.28	0.02	0.46	0.24	0.47	0.47	0.47	0.48	0.47	0.27	0.26
71	0.56	0.56	0.49	0.57	0.57	0.52	0.52	0.52	0.3	0.3	0.3	0.25	0.49	0.03	0.41	0.41	0.41	0.49	0.41	0.34	0.11
81	0.69	0.69	0.63	0.7	0.7	0.66	0.66	0.66	0.63	0.63	0.63	0.6	0.62	0.58	0.03	0.41	0.41	0.49	0.41	0.64	0.59
82	0.7	0.7	0.65	0.71	0.71	0.67	0.67	0.67	0.63	0.63	0.63	0.65	0.63	0.65	0.47	0.48	0.48	0.49	0.48	0.65	0.67
90	0.71	0.71	0.68	0.74	0.74	0.71	0.71	0.71	0.54	0.54	0.54	0.57	0.66	0.59	0.7	0.7	0.7	0.71	0.7	0.02	0.24
95	0.68	0.68	0.66	0.71	0.71	0.68	0.68	0.68	0.57	0.57	0.57	0.55	0.66	0.51	0.62	0.62	0.62	0.66	0.62	0.28	0.02

	2001 Classes											
2001 Classes	11	12	21	22	23	24	31	41	42	43	52	71
11	0	0.5	0.62	0.59	0.58	0.6	0.55	0.62	0.62	0.62	0.58	0.58
12	0.5	0	0.62	0.59	0.58	0.6	0.55	0.62	0.62	0.62	0.58	0.58
21	0.71	0.71	0	0.1	0.24	0.36	0.7	0.64	0.64	0.64	0.65	0.65
22	0.69	0.69	0.1	0	0.15	0.27	0.67	0.65	0.65	0.65	0.65	0.65
23	0.68	0.68	0.24	0.15	0	0.12	0.66	0.67	0.67	0.67	0.66	0.66
24	0.7	0.7	0.36	0.27	0.12	0	0.67	0.72	0.72	0.72	0.69	0.69
31	0.66	0.66	0.7	0.67	0.66	0.67	0	0.69	0.69	0.69	0.65	0.65
41	0.5	0.5	0.44	0.44	0.47	0.52	0.48	0	0.38	0.19	0.31	0.26
42	0.5	0.5	0.44	0.44	0.47	0.52	0.48	0.38	0	0.19	0.31	0.26
43	0.48	0.48	0.41	0.41	0.44	0.5	0.46	0.19	0.19	0	0.27	0.22
52	0.52	0.52	0.46	0.47	0.49	0.54	0.5	0.27	0.27	0.27	0	0.25
71	0.56	0.56	0.49	0.49	0.51	0.57	0.54	0.29	0.29	0.29	0.25	0
81	0.69	0.69	0.46	0.48	0.51	0.57	0.68	0.63	0.63	0.63	0.6	0.58
82	0.7	0.7	0.46	0.46	0.49	0.56	0.68	0.63	0.63	0.63	0.65	0.66
90	0.6	0.71	0.67	0.67	0.7	0.74	0.72	0.55	0.55	0.55	0.57	0.6
95	0.58	0.68	0.66	0.66	0.67	.71	0.69	0.57	0.57	0.57	0.55	0.51

Table 3
Cross tabulated semantic overlap values for land cover classes in NLCD 1992 and 2001 data

	1992 Classes																				
1992 Classes	11	12	21	22	23	31	32	33	41	42	43	51	61	71	81	82	83	84	85	91	92
11	1	0.87	0.72	0.65	0.65	0.87	0.87	0.87	0.62	0.62	0.62	0.74	0.79	0.74	0.79	0.74	0.74	0.82	0.74	0.71	0.81
12	0.87	1	0.72	0.65	0.65	0.87	0.87	0.87	0.62	0.62	0.62	0.74	0.79	0.74	0.79	0.74	0.74	0.82	0.74	0.5	0.64
21	0.22	0.22	1	0.79	0.65	0.49	0.49	0.49	0.55	0.55	0.55	0.5	0.65	0.5	0.51	0.5	0.5	0.51	0.67	0.6	0.55
22	0.55	0.55	0.89	1	0.89	0.71	0.71	0.71	0.37	0.37	0.37	0.58	0.58	0.58	0.58	0.58	0.58	0.71	0.73	0.45	0.63
23	0.55	0.55	0.78	0.89	1	0.71	0.71	0.71	0.37	0.37	0.37	0.58	0.58	0.58	0.58	0.58	0.58	0.71	0.73	0.45	0.63
31	0.71	0.71	0.72	0.65	0.65	1	1	1	0.56	0.56	0.56	0.69	0.74	0.69	0.74	0.69	0.69	0.96	0.69	0.62	0.74
32	0.64	0.64	0.64	0.58	0.58	0.89	1	0.89	0.5	0.5	0.5	0.62	0.67	0.62	0.67	0.62	0.62	0.86	0.62	0.56	0.67
33	0.82	0.82	0.83	0.75	0.75	1	1	1	0.64	0.64	0.64	0.8	0.86	0.8	0.86	0.8	0.8	0.95	0.8	0.72	0.86
41	0.8	0.8	0.81	0.76	0.76	0.89	0.89	0.89	1	0.85	0.85	0.85	0.93	0.93	0.93	0.93	0.93	0.89	0.93	1	0.93
42	0.8	0.8	0.81	0.76	0.76	0.89	0.89	0.89	0.85	1	0.85	0.85	0.93	0.93	0.93	0.93	0.93	0.89	0.93	1	0.93
43	0.8	0.8	0.81	0.76	0.76	0.89	0.89	0.89	0.85	0.85	1	0.85	0.93	0.93	0.93	0.93	0.93	0.89	0.93	1	0.93
51	0.76	0.76	0.77	0.66	0.66	0.87	0.87	0.87	0.82	0.82	0.82	1	0.89	0.91	0.91	0.91	0.91	0.87	0.91	0.91	0.91
61	0.63	0.63	0.65	0.51	0.51	0.81	0.81	0.81	0.8	0.8	0.8	0.72	1	0.72	0.88	0.88	0.88	0.81	0.88	0.8	0.72
71	0.76	0.76	0.73	0.62	0.62	0.87	0.87	0.87	0.91	0.91	0.91	0.91	0.91	1	0.91	0.91	0.91	0.77	0.91	0.82	1
81	0.62	0.62	0.61	0.53	0.53	0.73	0.73	0.73	0.76	0.76	0.76	0.67	0.77	0.86	1	0.86	0.86	0.72	0.86	0.66	0.79
82	0.71	0.71	0.67	0.57	0.57	0.8	0.8	0.8	0.76	0.76	0.76	0.76	0.85	0.85	0.93	1	0.93	0.81	0.93	0.66	0.85
83	0.71	0.71	0.67	0.57	0.57	0.8	0.8	0.8	0.76	0.76	0.76	0.76	0.85	0.85	0.93	0.93	1	0.81	0.93	0.66	0.85
84	0.72	0.72	0.66	0.56	0.56	0.87	0.87	0.87	0.64	0.64	0.64	0.64	0.73	0.73	0.73	0.73	0.73	1	0.73	0.57	0.67
85	0.66	0.66	0.74	0.65	0.65	0.75	0.75	0.75	0.71	0.71	0.71	0.71	0.79	0.84	0.84	0.84	0.84	0.73	1	0.62	0.84
91	0.65	0.53	0.54	0.41	0.41	0.62	0.62	0.62	0.85	0.85	0.85	0.77	0.77	0.67	0.67	0.67	0.67	0.62	0.67	1	0.85
92	0.79	0.7	0.65	0.55	0.55	0.76	0.76	0.76	0.8	0.8	0.8	0.8	0.8	0.87	0.8	0.8	0.8	0.68	0.8	0.87	1

	2001 Classes											
1992 Classes	11	12	21	22	23	24	31	41	42	43	52	71
11	1	0.87	0.75	0.72	0.72	0.65	0.69	0.59	0.59	0.59	0.7	0.66
12	0.87	1	0.75	0.72	0.72	0.65	0.69	0.59	0.59	0.59	0.7	0.66
21	0.22	0.22	0.89	0.91	0.88	0.79	0.46	0.54	0.54	0.54	0.46	0.45
22	0.55	0.55	0.89	0.89	0.89	1	0.61	0.3	0.3	0.3	0.54	0.54
23	0.55	0.55	0.89	0.89	0.89	1	0.61	0.3	0.3	0.3	0.54	0.54
31	0.71	0.71	0.75	0.72	0.72	0.65	0.85	0.46	0.46	0.46	0.6	0.54
32	0.64	0.64	0.67	0.64	0.64	0.58	0.88	0.41	0.41	0.41	0.53	0.48
33	0.82	0.82	0.86	0.83	0.83	0.75	0.79	0.53	0.53	0.53	0.69	0.62
41	0.8	0.8	0.83	0.82	0.77	0.76	0.89	0.95	0.79	0.79	0.79	0.88
42	0.8	0.8	0.83	0.82	0.77	0.76	0.89	0.79	0.95	0.79	0.79	0.88
43	0.8	0.8	0.83	0.82	0.77	0.76	0.89	0.79	0.79	0.95	0.79	0.88
51	0.76	0.76	0.84	0.78	0.73	0.66	0.8	0.76	0.76	0.76	0.92	0.82
61	0.63	0.63	0.85	0.83	0.76	0.71	0.79	0.71	0.71	0.71	0.6	0.59
71	0.76	0.76	0.84	0.73	0.71	0.62	0.73	0.86	0.86	0.86	0.86	0.85
81	0.62	0.62	0.82	0.72	0.71	0.65	0.64	0.7	0.7	0.7	0.6	0.73
82	0.71	0.71	0.87	0.77	0.76	0.69	0.68	0.7	0.7	0.7	0.7	0.69
83	0.71	0.71	0.87	0.77	0.76	0.69	0.68	0.7	0.7	0.7	0.7	0.69
84	0.72	0.72	0.81	0.8	0.78	0.67	0.74	0.6	0.6	0.6	0.6	0.62
85	0.66	0.66	0.95	0.88	0.87	0.8	0.62	0.65	0.65	0.65	0.66	0.72
91	0.65	0.53	.65	0.58	0.45	0.41	0.62	0.83	0.83	0.83	0.74	0.63
92	0.79	0.7	0.74	0.65	0.63	0.55	0.65	0.78	0.78	0.78	0.78	0.77

	1992 Classes																				
2001 Classes	11	12	21	22	23	31	32	33	41	42	43	51	61	71	81	82	83	84	85	91	92
11	1	0.87	0.72	0.65	0.65	0.87	0.87	0.87	0.62	0.62	0.62	0.74	0.79	0.74	0.79	0.74	0.74	0.82	0.74	0.71	0.81
12	0.87	1	0.72	0.65	0.65	0.87	0.87	0.87	0.62	0.62	0.62	0.74	0.79	0.74	0.79	0.74	0.74	0.82	0.74	0.5	0.64
21	0.37	0.37	0.78	0.68	0.68	0.57	0.57	0.57	0.66	0.66	0.66	0.6	0.73	0.6	0.61	0.6	0.6	0.57	0.75	0.7	0.65
22	0.21	0.21	0.83	0.68	0.68	0.49	0.49	0.49	0.59	0.59	0.59	0.52	0.67	0.52	0.53	0.52	0.52	0.48	0.69	0.64	0.58
23	0.3	0.3	0.92	0.69	0.69	0.53	0.53	0.53	0.51	0.51	0.51	0.51	0.64	0.51	0.53	0.51	0.51	0.55	0.68	0.57	0.57
24	0.55	0.55	0.81	0.92	0.92	0.71	0.71	0.71	0.37	0.37	0.37	0.58	0.58	0.58	0.58	0.58	0.58	0.71	0.73	0.45	0.63
31	0.64	0.64	0.63	0.71	0.71	0.89	1	0.89	0.49	0.49	0.49	0.66	0.66	0.66	0.66	0.66	0.66	0.89	0.66	0.55	0.71
41	0.8	0.8	0.81	0.76	0.76	0.89	0.89	0.89	0.98	0.82	0.82	0.84	0.92	0.92	0.92	0.92	0.92	0.89	0.92	0.99	0.92
42	0.8	0.8	0.81	0.76	0.76	0.89	0.89	0.89	0.82	0.98	0.82	0.84	0.92	0.92	0.92	0.92	0.92	0.89	0.92	0.99	0.92
43	0.8	0.8	0.81	0.76	0.76	0.89	0.89	0.89	0.82	0.82	0.98	0.84	0.92	0.92	0.92	0.92	0.92	0.89	0.92	0.99	0.92
52	0.77	0.77	0.77	0.66	0.66	0.87	0.87	0.87	0.81	0.81	0.81	0.99	0.87	0.91	0.91	0.91	0.91	0.87	0.91	0.9	0.91
71	0.77	0.77	0.73	0.71	0.71	0.87	0.87	0.87	0.91	0.91	0.91	0.91	0.91	1	0.91	0.91	0.91	0.77	0.91	0.81	1
81	0.65	0.65	0.64	0.61	0.61	0.77	0.77	0.77	0.7	0.7	0.7	0.81	0.81	0.91	1	0.91	0.91	0.77	0.93	0.7	0.84
82	0.57	0.57	0.58	0.46	0.46	0.72	0.72	0.72	0.71	0.71	0.71	0.63	0.77	0.65	0.78	0.79	0.79	0.74	0.78	0.71	0.63
90	0.65	0.53	0.55	0.41	0.41	0.62	0.62	0.62	0.84	0.84	0.84	0.75	0.76	0.66	0.66	0.66	0.66	0.62	0.66	0.98	0.84
95	0.79	0.7	0.65	0.63	0.63	0.76	0.76	0.76	0.8	0.8	0.8	0.8	0.8	0.87	0.8	0.8	0.8	0.68	0.8	0.86	1

	2001 Classes											
2001 Classes	11	12	21	22	23	24	31	41	42	43	52	71
11	1	0.87	0.75	0.72	0.72	0.65	0.69	0.59	0.59	0.59	0.7	0.66
12	0.87	1	0.75	0.72	0.72	0.65	0.69	0.59	0.59	0.59	0.7	0.66
21	0.37	0.37	1	0.89	0.82	0.78	0.55	0.63	0.63	0.63	0.56	0.55
22	0.21	0.21	0.89	1	0.82	0.78	0.47	0.55	0.55	0.55	0.46	0.45
23	0.3	0.3	0.89	0.89	1	0.8	0.47	0.51	0.51	0.51	0.49	0.47
24	0.55	0.55	0.89	0.89	0.89	1	0.61	0.3	0.3	0.3	0.54	0.54
31	0.64	0.64	0.67	0.64	0.63	0.71	1	0.37	0.37	0.37	0.58	0.58
41	0.8	0.8	0.85	0.81	0.78	0.76	0.89	1	0.85	0.85	0.85	0.93
42	0.8	0.8	0.85	0.81	0.78	0.76	0.89	0.85	1	0.85	0.85	0.93
43	0.8	0.8	0.85	0.81	0.78	0.76	0.89	0.85	0.85	1	0.85	0.93
52	0.77	0.77	0.87	0.78	0.73	0.66	0.8	0.82	0.82	0.82	1	0.88
71	0.77	0.77	0.87	0.72	0.72	0.71	0.78	0.91	0.91	0.91	0.91	1
81	0.65	0.65	0.89	0.74	0.74	0.74	0.71	0.71	0.71	0.71	0.82	0.91
82	0.57	0.57	0.8	0.74	0.68	0.64	0.71	0.73	0.73	0.73	0.64	0.66
90	0.65	0.53	0.68	0.58	0.47	0.41	0.62	0.85	0.85	0.85	0.76	0.66
95	0.79	0.7	0.77	0.64	0.64	0.63	0.69	0.8	0.8	0.8	0.8	0.87

cannot be treated as no-change instances since there is no guarantee that two categories map to each other unambiguously. However, using a lookup table with estimates of land cover class correspondence we can use that to guide our interpretation of a change matrix. In this work we will use the semantic similarity matrices to get an indication of where the cells of smallest and largest change are located. The combination of semantic relations and areas affected will be provided as a change matrix color coded to represent the semantic relationship between the classes represented by each cell in the matrix. This will serve as a semi-quantitative tool for evaluating overall magnitudes and trajectories of land cover change.

A standard post-classification assessment can also present a spatial picture of identified changes by recoding the change image into a binary image of change and no change pixels. If we are interested in particular types of change we can similarly recode and produce a number of change maps for each type of change. This may turn out problematic though since the number of possible types of changes can be very large ($n \times m$, with n classes at time 1 and m classes at time 2). The conditions of this study is particularly challenging since many categories in the 1992 classification do not have a one-to-one match in the 2001 classification (see Appendix A). The simplified change scenario in Fig. 5a and b illustrates a 3×3 pixel area that undergoes some urban expansion and intensification as well as some forest expansion on previous agricultural fields. In this scenario it is not always clear what a crisp change image can tell us since the user needs to figure out what some of the changes, e.g. from “Low intensity residential” to “Developed medium intensity”, really means, see Fig. 5c.

The use of semantic metrics to identify change can help to distinguish not only between different degrees of change, such as from low to medium density urban land, but also when heterogeneous classes are used. The semantic change matrices can be used to replace the change identity, e.g. a pixel changing from “21 Low intensity residential” to “23 Developed medium intensity”, with a metric for their semantic relationship, e.g. a distance value of 0.46 and an overlap value of 0.89. These values can then be visualized with the bi-variate color legend described previously to display a map of semantic change, Fig. 5d. In the scenario above, semantic metrics would indicate a large similarity between identical or almost identical classes, a slight difference between grades of developed land, and larger differences when land change from forest or cultivated to developed land. Class–subclass and disjoint class relations also creates a more nuanced picture of the land cover changes.

3. Demonstration and results

3.1. Semantic analysis

In Tables 2 and 3 we see the result of the semantic similarity analysis of the land cover class definitions for the two NLCD data sets. These were produced by calculating overlap and distance values for all class pair relationships and then organize these metrics into two contingency matrices, one for overlap and one for distance values, so that each cell got the value from comparing one class with another.

Table 2 shows the distance metric. The upper left quarter section holds values for comparisons within the 1992 classes and

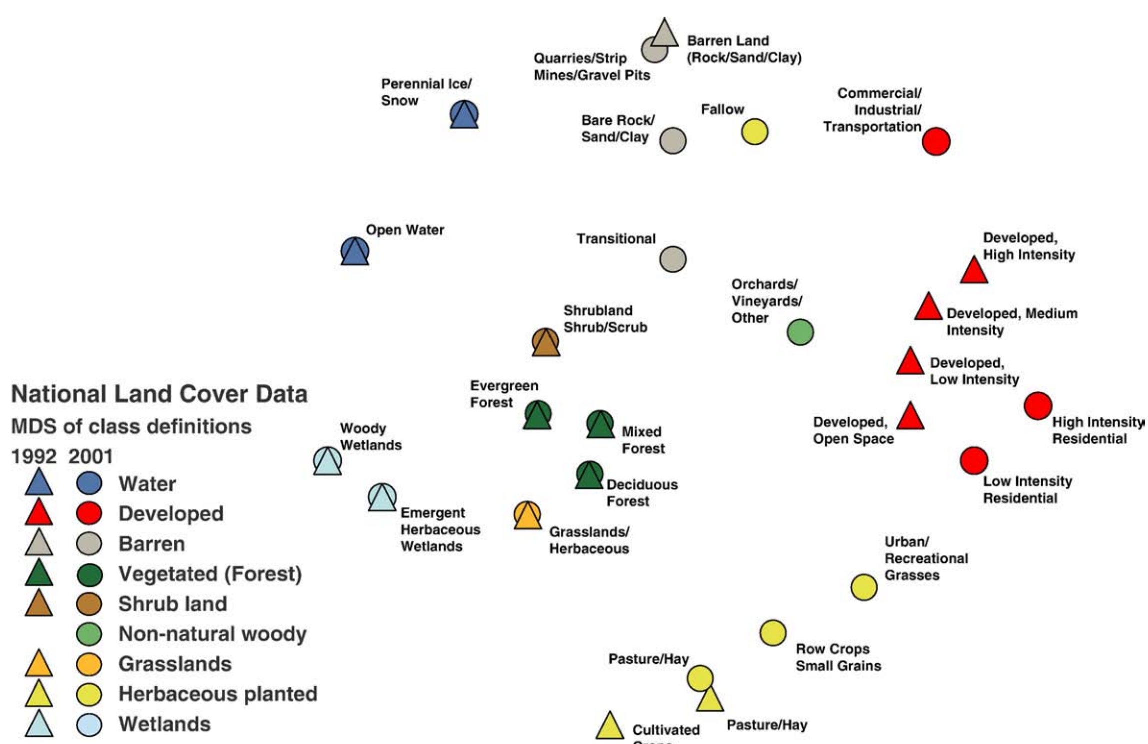

Fig. 6. A two-dimensional scatter plot generated from a multi-dimensional scaling (PROXCAL) of the 16 attribute dimensions of the NLCD 1992 and 2001 land cover classes.

Table 4
Summary of land cover change in Chester County, PA, from 1992 to 2001 as detected by the NLCD data

1992 major classes	1992 Area, km^2 (%)	Change, km^2 (%)	2001 Area, km^2 (%)	2001 major classes
Water	11.7 (0.3)	−4.2 (+0.1)	7.5 (0.4)	Water
Developed	167.7 (8.5)	91.7 (+4.7)	259.4 (13.2)	Developed
Barren	23.8 (1.2)	−3.4 (−0.2)	20.4 (1.0)	Barren
Forest	821.5 (41.8)	−266.8 (−13.6)	554.7 (28.2)	Forest
Shrubland	0 (0)	0	0 (0)	Shrubland
Non-natural woody	0 (0)	N/A	N/A	
Grasslands	0 (0)	0	0 (0)	Grasslands
Planted/cultivated	931.8 (47.4)	151.1 (+7.6)	1082.9 (55.0)	Planted/cultivated
Wetlands	11.3 (0.6)	31.5 (+1.6)	42.8 (2.2)	Wetlands

the lower right quarter section holds values from comparing the 2001 classes with each other. In the other two quarter sections of the table we find distance values for semantic comparison between the two systems. In the upper right section we find values of semantic distance when 1992 classes are compared with 2001 classes as a reference and the lower left section show the inverse comparison, 2001 classes evaluated with 1992 classes as reference. In this way asymmetries in the semantic relations can be read from the table. The main diagonal in this full comparison table holds values for a class compared with itself and therefore have distance values of zero, i.e. identical classes. The overlap matrix in Table 3 follows the same layout but this metric has almost an inverse interpretation in that zero values means no overlap and semantically different values, whereas full overlap has a value of one indicating a class–subclass relation. Again the main diagonal, holding comparisons of a class with itself will exhibit full overlap and have values of one. The large number of cells in these tables makes a comprehensive interpretation very complex and time consuming but it can be consulted for detailed information on specific land cover types. A more informative summary view of the two classifications is provided by the multi-dimensional scaling plot. Fig. 6 shows the result from a multi-dimensional scaling of the semantic distances between land cover classes. 1992 classes are symbolized as colored triangles and

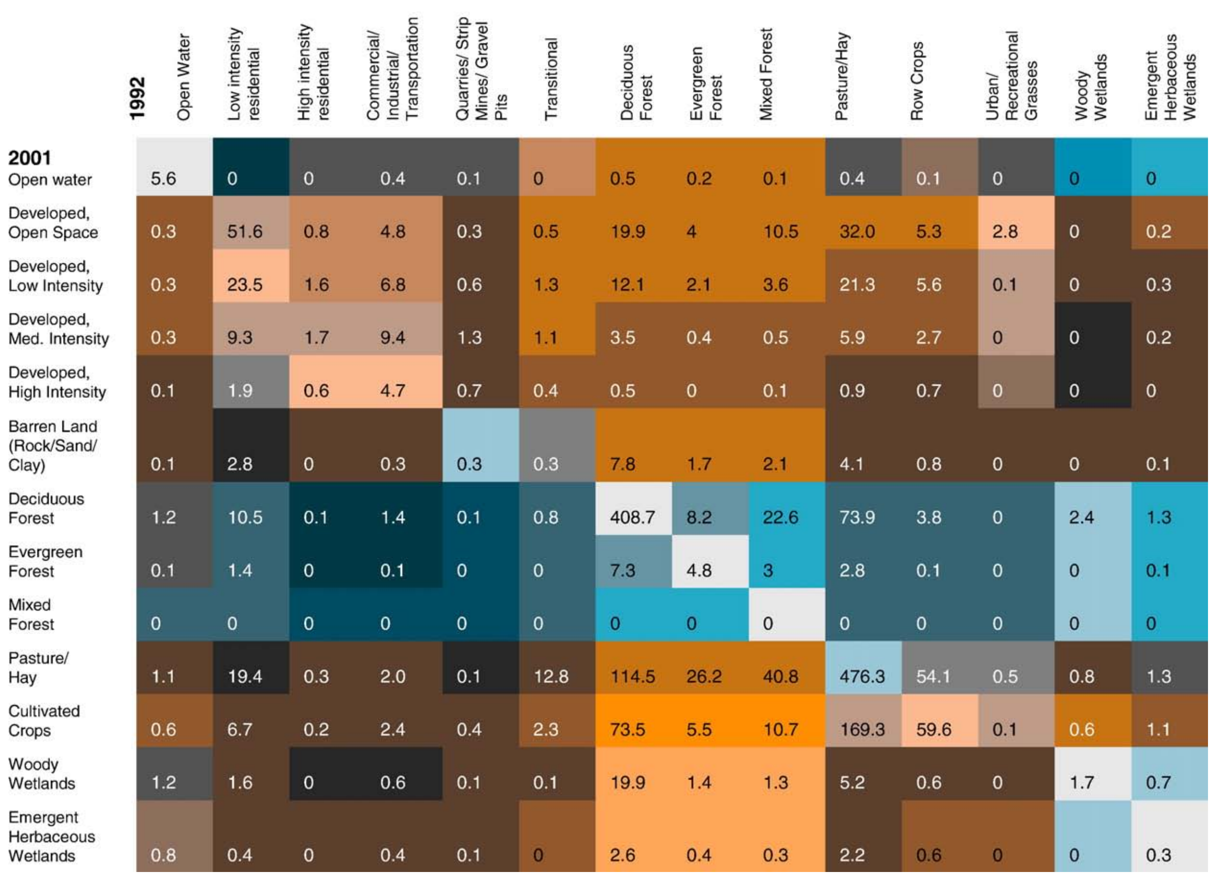

Fig. 7. Full semantic change matrix for the Chester County study area. Values are square kilometers and color coding follows the bi-variate color scheme introduced in Fig. 4. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2001 classes are symbolized as colored circles all with class labels. The colors closely follow those used by USGS in maps that depict these land cover data sets.

From this summary graph we can see that many categories are identical or largely the same in the 1992 and 2001 classification systems with a few significant departures for the Developed, Transitional, and Agriculture classes. This is also noted by Homer et al. (2004) in their description of the NLCD 2001 development. The clustering and separation between major groups of land cover types (shown in different colors) are mostly distinct. Also the close proximity between similar major groups reflect the semantic similarity between land cover classes, for example the location of woody wetlands towards the left where the water classes are, and the 1992 class urban/recreational grasses located near the residential classes. Obviously, a close scrutiny of this plot as well as the separate distance and overlap matrices can reveal much other information. It may for instance reveal the thematic context of a classification system, such as if it has a certain focus with many classes crowding a specific portion of the graph.

3.2. Fuzzy change assessment in Chester County

A standard cross tabulation of the change images is summarized in Table 4. Here some classes have been aggregated to major groups for clarity. We see that the major changes seem to occur in the Developed, Forest, and Planted/cultivated categories. Most of the loss is recorded for forest that loose about 270 km^2 to other land cover types from 1992 to 2001. The two biggest gains are for Developed land (~ 90 km^2) and Planted/cultivated land cover (~ 150 km^2). The magnitude and direction of gain in developed land largely follows the trend noted by (Arthur et al., 2000). However, while their data show a large decrease of

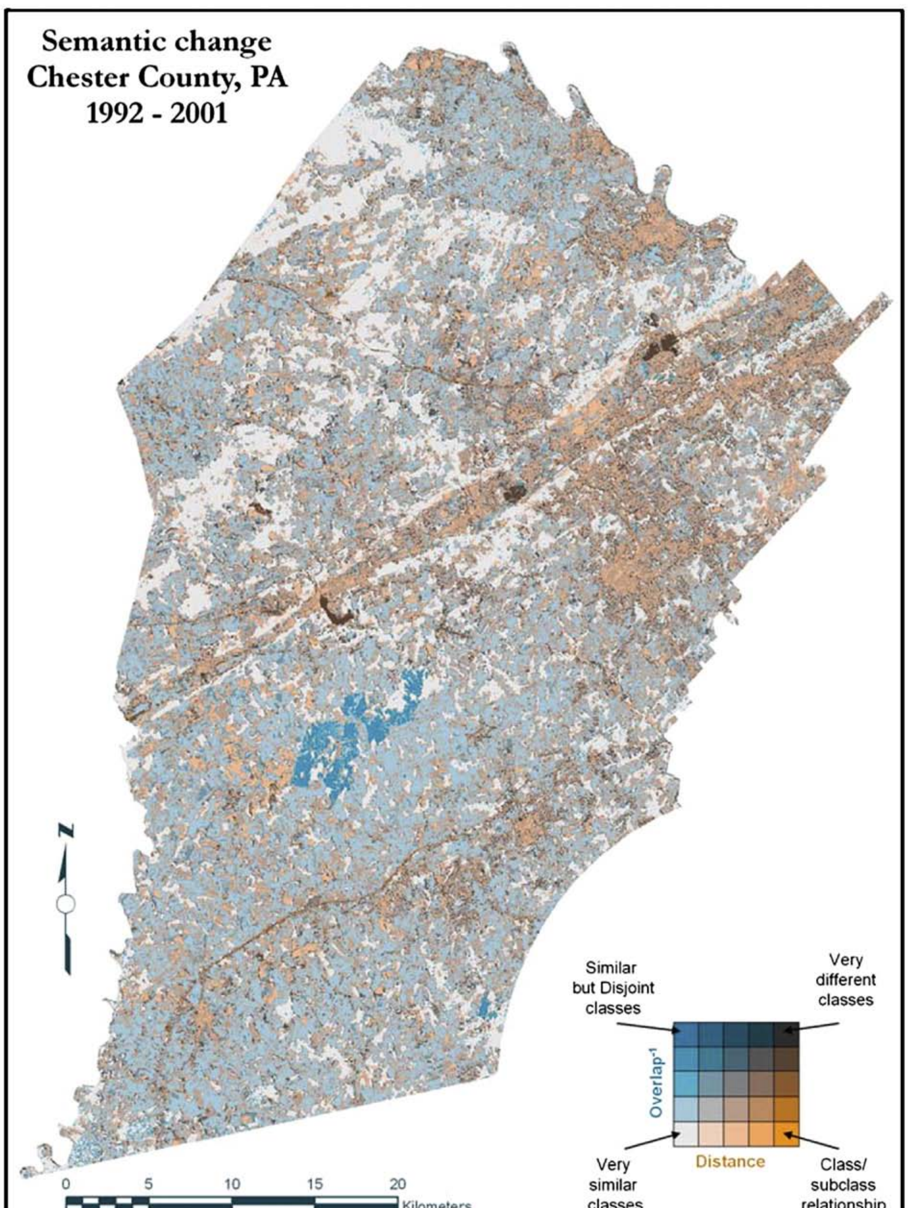

Fig. 8. Semantic change map for Chester County, PA according to NLCD 1992 and 2001 data and the semantic similarity metrics given in Tables 2 and 3.

agricultural and small decrease of forested area, the NLCD data suggest a very large decrease of forest areas but the opposite trend for agricultural areas.

The full change matrix obviously has a lot more detail on the specific changes, but it also provides the challenge of comparing the heterogeneous classes and not having a major diagonal of no-change to guide the interpretation. Fig. 7 shows the change matrix for the entire study area with values in square kilometers. With the help of the semantic similarity values from the distance and overlap matrices the change matrix is color coded to symbolize semantic distance and overlap values following the bi-variate color scheme introduced earlier.

In the colored change matrix the light, pale colors that symbolize small semantic differences between the classes show very similar classes in some instances along the major diagonal, specifically for the water, forest, and wetland categories. But we can also see some large similarities in the off diagonal elements. For example the 2001 *Developed*, *open Space* class shows signs of being a sub-class of the 1992 *Urban*/*Recreational Grasses* class. The 2001 wetland categories also indicate a close and largely overlapping, sub-class relationship with the 1992 forest classes. For the forest categories the half saturated blue shades around a bright diagonal illustrate matching classes along the diagonal and largely similar classes off the diagonal but where the blue tone indicates a disjoint semantic relationship (lower overlap). Darker blue shades (nearly no overlap), such as for changes from urban/transitional to forest, indicate more prominent land cover changes than the fairly subtle change from e.g. mixed forest to deciduous forest.

The semantic change map in Fig. 8 shows the overall spatial pattern of semantic change in Chester County. The main features of this map are the orange tones associated with developed areas and the light blue tones associated with the rural areas. These mostly reflect the changes in definition of the agriculture and developed land cover classes from the 1992 to the 2001 mapping mentioned previously. Some large, dark gray areas show locations of significant change. Most of these areas represent change from the 1992 class Quarries/Strip Mines/Gravel Pits to the 2001 Developed classes. These class descriptions generally have opposite or very different values in their characteristic attributes which results in large distance and small overlap values. Another prominent feature is the large, irregular, blue area just left and down from the map center. This area was classified as Transitional in the 1992 data but turns out to be mostly Pasture/Hay in the 2001 classification. Although it seems that the 1992 classification is wrong this type of change is recognized as a

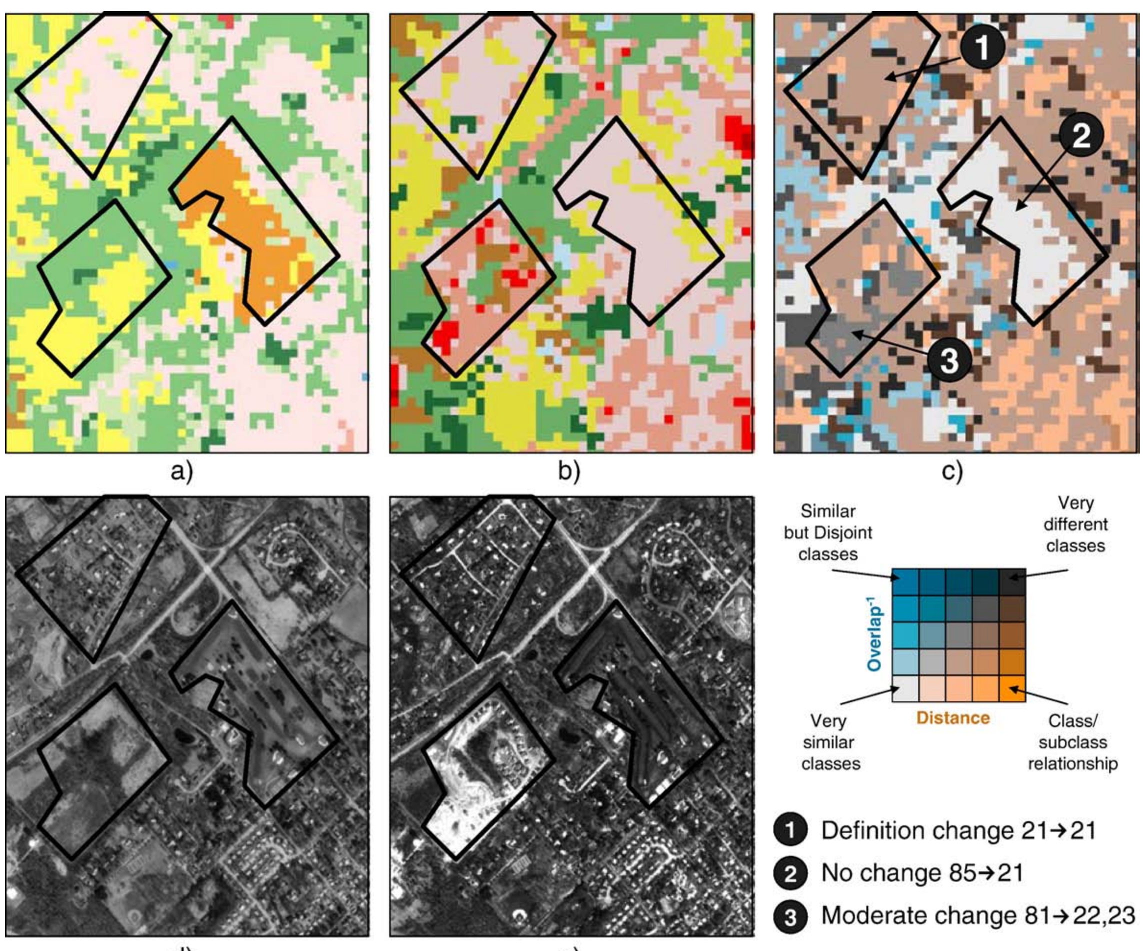

Fig. 9. Examples of land cover change details from a 1.5 km^2 sub-section of Chester County, PA. a) NLCD1992 data, b) NLCD 2001 data, c) semantic change image with three numbered change cases, d) orthophoto from 1992, e) orthophoto from 1999.

change to a very similar (small distance), but separate (small overlap) class from the original Transitional category.

To further exemplify some of the details revealed by the semantic change map Fig. 9 shows five images of the same enlarged sub-section of the study area. The first two land cover images represent a) the 1992 classification and b) the 2001 classification. Fig. 9c shows the semantic change image, color coded according to the bi-variate color scheme introduced in Fig. 4. Three case areas with different types of land cover change are numbered and will be discussed below. For further reference two orthophotos are provided that roughly correspond to the same dates as the land cover data above them. The image in Fig. 9d was taken 1992 and the image in Fig. 9e was taken in 1999.

For change case 1 in Fig. 9 the area with single-house development is mainly classified as Low Intensity Residential in 1992 and as Developed Open Space in 2001. The semantic change image indicates a slight to moderate change for this area based on the semantic mismatch between these two land cover categories. As mentioned previously, the land cover classes related to developed areas changed significantly in the 2001 classification. The 2001 Developed Open Space category is most closely corresponding to the Urban/Recreational Grasses class in the 1992 classification system. The other three 2001 developed classes, low, medium, and high density, have substantial overlaps with the two residential classes in the 1992 classification system. A further complicating factor is the inclusion of Commercial/Industrial/Transportation areas in the 2001 Developed categories. These types of areas had a separate class in the 1992 data.

The golf course area in change case number 2 is interesting in that it illustrates a true, no-change situation that was classified as Urban/Recreational Grasses in 1992 and recognized as Developed, Open Space in 2001. Again, these two categories could possibly be interpreted as a direct match, but there are slight differences in their descriptions (see Appendix A for details). Nevertheless, the descriptive attributes are very similar with a large overlap and therefore shows up as a bright, no-change area in the semantic change image.

Finally case 3 illustrates a grass/brush/tree area, classified as Deciduous Forest and Pasture/Hay in 1992, that undergoes development into a single house residential area classified mostly as Low-Medium Intensity Developed in the 2001 data. This change is also recognized by the semantic change image as a slight to moderate change with intermediate overlap and distance values. It is interesting to note that large portions of this true change has similar semantic similarity values as the areas in case 1 above, and only the areas that get classified as changing into Medium Intensity Developed are darker and less similar in the semantic change image. Here, the definition changes between the 1992 and 2001 mappings creates a problem of falsely indicating changes in situations such as case 1, thus creating a noise that masks moderate changes in the land cover such as situations illustrated by case 3, Fig. 9.

4. Concluding comments

The semantic similarity change methodology is still in its infancy. Much research is needed on the validity of formal semantic descriptions, to develop infrastructure for collaborative development of different classification systems, and further development of methods to estimate area changes as well as accuracy assessments of the change results. The development of formalized axioms or prototypes can use automated methods such as Natural language Processing of category definitions (Jensen & Binot, 1987) or be done manually by eliciting knowledge from domain experts (Feng & Flewelling, 2004). Evaluating the validity of derived specifications probably entails a collaborative process where many users can evaluate and negotiate proposed class specifications according to their own understanding of data (Gahegan & Pike, 2006). In terms of assessing accuracy for the change analysis most work has been concerned with the validation of single image accuracy, and only very little research has been directed towards accuracy assessment of change detection. Nevertheless, the error matrix approach used in many standard procedures is still a useful method to evaluate the accuracy of change analyses (Khorram et al., 1999). Specifically the methods developed by Gopal and Woodcock (1994) and Congalton and Green (1999) seem to suggest viable approaches for fuzzy land cover data in general. The exact implementation of a fuzzy interpretation of land cover change remains to be investigated. Nevertheless, initial work on the joint effect of error and vagueness in land cover data suggests that accuracy assessments need to include complementary analyses of both aspects (Ahlqvist and Gahegan, 2005).

In some respect the proposed methodology suffers from similar problems as other methods based on algebraic evaluation of image characteristics. Although much of the problems associated with radiometric and atmospheric correction of the images are eliminated by the post-classification approach, it is hard to establish at what level the image difference values separate actual change from apparent change. In the semantic case this has a specific translation in that a category may have changed a lot in its definition although it is still a sub-category of the original class.

The use of medium resolution satellite data products in e.g. regional planning applications rests on the ability to not only provide global summaries of changes, but also to enable users to compare spatial patterns of those changes. Most post-classification techniques for spatially explicit evaluation of land cover change rely on the use of identical classification systems over time. Furthermore they generate binary change images with no recognition of major and minor landscape changes or provide only a class-by-class based evaluation. The suggested methodology is more versatile in that it can both provide an overall, spatially explicit evaluation of land cover change throughout the study area, as well as nuanced assessments based on changes of heterogeneous land cover types or aspects of land cover such as vegetative cover.

Acknowledgements

Jamison Conley, Penn State University helped provide the bi-variate color scheme. Thanks to James Griffin for producing the study area overview map in Fig. 1 as part of class work in Cynthia Brewer's Advanced Cartography class, Penn State University. I also wish to thank the reviewers for their helpful and encouraging comments. This work was partially supported by the National Science Foundation under Grant No. BCS-9978052.

Appendix A

Summary of land cover class definitions in the NLCD 1992 and 2001 data with approximately matching classes lined up next to each other

1992	2001
11. Open water — All areas of open water, generally with less than 25% cover of vegetation/land cover.	11. Open water — All areas of open water, generally with less than 25% cover of vegetation or soil.
12. Perennial ice/snow — All areas characterized by year-long cover of ice and/or snow.	12. Perennial ice/snow — All areas characterized by a perennial cover of ice and/or snow, generally greater than 25% of total cover.
Developed — Areas characterized by high percentage (approximately 30% or greater) of constructed materials (e.g. asphalt, concrete, buildings, etc.).	
See 85	21. Developed, open space — Includes areas with a mixture of some constructed materials, but mostly vegetation in the form of lawn grasses. Impervious surfaces account for less than 20% of total cover. These areas most commonly include large-lot single-family housing units, parks, golf courses, and vegetation planted in developed settings for recreation, erosion control, or aesthetic purposes.
21. Low intensity residential — Includes areas with a mixture of constructed materials and vegetation. Constructed materials account for 30–80% of the cover. Vegetation may account for 20 to 70% of the cover. These areas most commonly include single-family housing units. Population densities will be lower than in high intensity residential areas.	22. Developed, low intensity — Includes areas with a mixture of constructed materials and vegetation. Impervious surfaces account for 20–49% of total cover. These areas most commonly include single-family housing units.
22. High intensity residential — Includes heavily built up urban centers where people reside in high numbers. Examples include apartment complexes and row houses. Vegetation accounts for less than 20% of the cover. Constructed materials account for 80–100% of the cover.	23. Developed, medium intensity — Includes areas with a mixture of constructed materials and vegetation. Impervious surfaces account for 50–79% of the total cover. These areas most commonly include single-family housing units.
23. Commercial/industrial/ transportation — Includes infrastructure (e.g. roads, railroads, etc.) and all highways and all developed areas not classified as High Intensity Residential.	24. Developed, high intensity — Includes highly developed areas where people reside or work in high numbers. Examples include apartment complexes, row houses and commercial/industrial. Impervious surfaces account for 80 to100% of the total cover.

Appendix A (*continued*)

1992	2001
Barren — Areas characterized by bare rock, gravel, sad, silt, clay, or other earthen material, with little or no "green" vegetation present regardless of its inherent ability to support life. Vegetation, if present, is more widely spaced and scrubby than that in the "green" vegetated categories; lichen cover may be extensive.	
31. Bare rock/sand/clay — Perennially barren areas of bedrock, desert, pavement, scarps, talus, slides, volcanic material, glacial debris, and other accumulations of earthen material.	31. Barren land (rock/sand/clay) — Barren areas of bedrock, desert, pavement, scarps, talus, slides, volcanic material, glacial debris, sand dunes, strip mines, gravel pits, and other accumulations of earthen material. Generally, vegetation accounts for less than 15% of total cover.
32. Quarries/strip mines/gravel pits — Areas of extractive mining activities with significant surface expression.	
33. Transitional — Areas of sparse vegetative cover (less than 25% that are dynamically changing from one land cover to another, often because of land use activities. Examples include forest clearcuts, a transition phase between forest and agricultural land, the temporary clearing of vegetation, and changes due to natural causes (e.g. fire, flood, etc.).	
Forested upland — Areas characterized by tree cover (natural or semi-natural woody vegetation, generally greater than 6 m tall); Tree canopy accounts for 25–100% of the cover.	
41. Deciduous forest — Areas dominated by trees where 75% or more of the tree species shed foliage simultaneously in response to seasonal change.	41. Deciduous forest — Areas dominated by trees generally greater than 5 m tall, and greater than 20% of total vegetation cover. More than 75% of the tree species shed foliage simultaneously in response to seasonal change.
42. Evergreen forest — Areas characterized by trees where 75% or more of the tree species maintain their leaves all year. Canopy is never without green foliage.	42. Evergreen forest — Areas dominated by trees generally greater than 5 m tall, and greater than 20% of total vegetation cover. More than 75% of the tree species maintain their leaves all year. Canopy is never without green foliage.
43. Mixed forest — Areas dominated by trees where neither deciduous nor evergreen species represent more than 75% of the cover present.	43. Mixed forest — Areas dominated by trees generally greater than 5 m tall, and greater than 20% of total vegetation cover. Neither deciduous nor evergreen species are greater than 75% of total tree cover.

(continued on next page)

Appendix A (*continued*)

1992	2001
Shrubland — Areas characterized by natural or semi-natural woody vegetation with aerial stems, generally less than 6 m tall with individuals or clumps not touching to interlocking. Both evergreen and deciduous species of true shrubs, young trees, and trees or shrubs that are small or stunted because of environmental conditions are included.	
51. Shrubland — Areas dominated by shrubs; shrub canopy accounts for 25–100% of the cover. Shrub cover is generally greater than 25% when tree cover is less than 25%. Shrub cover may be less than 25% in cases when the cover of other life forms (e.g. herbaceous or tree) is less than 25% and shrubs cover exceeds the cover of the other life forms.	52. Shrub/scrub — Areas dominated by shrubs; less than 5 m tall with shrub canopy typically greater than 20% of total vegetation. This class includes true shrubs, young trees in an early successional stage or trees stunted from environmental conditions.
Non-natural woody — Areas dominated by non-natural woody vegetation; non-natural woody vegetative canopy accounts for 25–100% of the cover. The non-natural woody classification is subject to the availability of sufficient ancillary data to differentiate non-natural woody vegetation from natural woody vegetation.	
61. Orchards/vineyards/other — Orchards, vineyards, and other areas planted or maintained for the production of fruits, nuts, berries, or ornamentals.	See 82
Herbaceous upland — Upland areas characterized by natural or semi-natural herbaceous vegetation; herbaceous vegetation accounts for 75–100% of the cover.	
71. Grasslands/herbaceous — Areas dominated by upland grasses and forbs. In rare cases, herbaceous cover is less than 25%, but exceeds the combined cover of the woody species present. These areas are not subject to intensive management, but they are often utilized for grazing.	71. Grassland/herbaceous — Areas dominated by graminoid or herbaceous vegetation, generally greater than 80% of total vegetation. These areas are not subject to intensive management such as tilling, but can be utilized for grazing.
Planted/cultivated — Areas characterized by herbaceous vegetation That has been planted or is intensively managed for the production of food, feed, or fiber; or is maintained in developed settings for specific purposes. Herbaceous vegetation accounts for 75–100% of the cover.	
81. Pasture/hay — Areas of grasses, legumes, or grass–legume mixtures planted for livestock grazing or the production of seed or hay crops.	81. Pasture/hay — Areas of grasses, legumes, or grass–legume mixtures planted for livestock grazing or the production of seed or hay crops, typically on a perennial cycle. Pasture/hay vegetation accounts for greater than 20% of total vegetation.
82. Row crops — Areas used for the production of crops, such as corn, soybeans, vegetables, tobacco, and cotton.	82. Cultivated crops — Areas used for the production of annual crops, such as corn, soybeans, vegetables, tobacco, and cotton, and also perennial woody crops such as orchards and vineyards. Crop vegetation accounts for greater than 20% of total vegetation. This class also includes all land being actively tilled.
83. Small grains — Areas used for the production of graminoid crops such as wheat, barley, oats, and rice.	
84. Fallow — Areas used for the production of crops that are temporarily barren or with sparse vegetative cover as a result of being tilled in a management practice that incorporates prescribed alternation between cropping and tillage.	
85. Urban/recreational grasses — Vegetation (primarily grasses) planted in developed settings for recreation, erosion control, or aesthetic purposes. Examples include parks, lawns, golf courses, airport grasses, and industrial site grasses.	See 21
Wetlands — Areas where the soil or substrate is periodically saturated with or covered with water as defined by Cowardin et al.	
91. Woody wetlands — Areas where forest or shrubland vegetation accounts for 25–100% of the cover and the soil or substrate is periodically saturated with or covered with water.	90. Woody wetlands — Areas where forest or shrubland vegetation accounts for greater than 20% of vegetative cover and the soil or substrate is periodically saturated with or covered with water.
92. Emergent herbaceous wetlands — Areas where perennial herbaceous vegetation accounts for 75–100% of the cover and the soil or substrate is periodically saturated with or covered with water.	95. Emergent herbaceous wetlands — Areas where perennial herbaceous vegetation accounts for greater than 80% of vegetative cover and the soil or substrate is periodically saturated with or covered with water.

References

Ahlqvist, O. (2004). A parameterized representation of uncertain conceptual spaces. *Transactions in GIS*, *8*(4), 493–514.

Ahlqvist, O., & Gahegan, M. (2005). Probing the relationship between classification error and class similarity. *Photogrammetric Engineering and Remote Sensing*, *71*(12), 1365−1373.

Ahlqvist, O. (2005). Using uncertain conceptual spaces to translate between land cover categories. *International Journal of Geographical Information Science*, *19*(7), 831−857.

Anderson, J. R. (1976). *A land use and land cover classification system for use with remote sensor data.* US Govt. Print. Office.

Arthur, S. T., Carlson, T. N., & Ripley, D. A. J. (2000). Land use dynamics of Chester County, Pennsylvania, from a satellite remote sensing perspective. *Geocarto International*, *15*(1), 25−35.

Bishr, Y. (1998). Overcoming the semantic and other barriers to GIS interoperability. *International Journal of Geographical Information Science*, *12*(4), 299.

Comber, A., Fisher, P., & Wadsworth, R. (2004a). Integrating land-cover data with different ontologies: Identifying change from inconsistency. *International Journal of Geographical Information Science*, *18*(7), 691−708.

Comber, A., Fisher, P. F., & Wadsworth, R. (2004b). Assessment of a semantic statistical approach to detecting land cover change using inconsistent data sets. *Photogrammetric Engineering and Remote Sensing*, *70*(8), 931−938.

Congalton, R. G., & Green, K. (1999). *Assessing the accuracy of remotely sensed data: Principles and practices.* Lewis Publications.

Di Gregorio, A., & Jansen, L. J. M. (2000). *Land cover classification system: LCCS: Classification concepts and user manual.* Rome: Food and Agriculture Organization of the United Nations.

Feng, C. C., & Flewelling, D. M. (2004). Assessment of semantic similarity between land use/land cover classification systems. *Computers, Environment and Urban Systems*, *28*(3), 229−246.

Fisher, P., Arnot, C., Wadsworth, R., & Wellens, J. (2006). Detecting change in vague interpretations of landscapes. *Ecological Informatics*, *1*(2), 163−178.

Fisher, P. F., & Pathirana, S. (1990). The evaluation of fuzzy membership of land cover classes in the suburban zone. *Remote Sensing of Environment*, *34*(2), 121−132.

Fonseca, F. T., Egenhofer, M. J., Agouris, P., & Camara, G. (2002). Using ontologies for integrated geographic information systems. *Transactions in GIS*, *6*(3), 231−257.

Foody, G. M., & Boyd, D. S. (1999). Detection of partial land cover change associated with the migration of inter-class transitional zones. *International Journal of Remote Sensing*, *20*(14), 2723−2740.

Foody, G. M., & Cox, D. P. (1994). Sub-pixel land cover composition estimation using a linear mixture model and fuzzy membership functions. *International Journal of Remote Sensing (Print)*, *15*(3), 619−631.

Fritz, S., & See, L. (2005). Comparison of land cover maps using fuzzy agreement. *International Journal of Geographical Information Science*, *19*(7), 787−807.

Gahegan, M., & Pike, W. (2006). A situated knowledge representation of geographical information. *Transactions in GIS*, *10*(5), 727−749.

Gärdenfors, P. (2000). *Conceptual spaces: The geometry of thought.* Cambrige, MA: MIT Press.

Gopal, S., & Woodcock, C. (1994). Theory and methods for accuracy assessment of thematic maps using fuzzy sets. *Photogrammetric Engineering and Remote Sensing*, *60*(2), 181−188.

Guarino, N. (1995). Formal ontology, conceptual analysis and knowledge representation. *International Journal of Human Computer Studies*, *43*(5–6), 625−640.

Homer, C., Huang, C., Yang, L., Wylie, B., & Coan, M. (2004). Development of a 2001 National Landcover Database for the United States. *Photogrammetric Engineering and Remote Sensing*, *70*(4), 829−840.

Jensen, K., & Binot, J. L. (1987). Disambiguating prepositional phrase attachments by using on-line dictionary definitions. *Computational Linguistics*, *13*(3–4), 251−260.

Jung, M., Henkel, K., Herold, M., & Churkina, G. (2006). Exploiting synergies of global land cover products for carbon cycle modeling. *Remote Sensing of Environment*, *101*(4), 534−553.

Kaufmann, A., & Gupta, M. M. (1985). *Introduction to fuzzy arithmetic.* New York, NY: Van Nostrand Reinhold Co.

Kavouras, M., & Kokla, M. (2002). A method for the formalization and integration of geographical categorizations. *International Journal of Geographical Information Science*, *16*(5), 439.

Khorram, S., Biging, G. S., Chrisman, N. R., Colby, D. R., Congalton, R. G., Dobson, J. E., et al. (1999). In S. Khorram (Ed.), *Accuracy assessment of remote sensing derived change detection.* Bethesda, MD: American Society for Photogrammetry and Remote Sensing.

Lambin, E. F., & Strahler, A. H. (1994). Change-vector analysis in multitemporal space: A tool to detect and categorize land-cover change processes using high temporal-resolution satellite data. *Remote Sensing of Environment*, *48*(2), 231−244.

Loveland, T. R., Sohl, T. L., Stehman, S. V., Gallant, A. L., Sayler, K. L., & Napton, D. E. (2002). A strategy for estimating the rates of recent united states land-cover changes. *Photogrammetric Engineering and Remote Sensing*, *68* (10), 1091−1099.

Lu, D., Mausel, P., Brondízio, E., & Moran, E. (2004). Change detection techniques. *International Journal of Remote Sensing*, *25*(12), 2365−2407.

Mas, J. F. (1999). Monitoring land-cover changes: A comparison of change detection techniques. *International Journal of Remote Sensing*, *20*(1), 139−152.

Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of experimental psychology. General*, *115*(1), 39−61.

Pontius, R. G., & Cheuk, M. L. (2006). A generalized cross-tabulation matrix to compare soft-classified maps at multiple resolutions. *International Journal of Geographical Information Science*, *20*(1), 1−30.

Salge, F. (1995). Semantic accuracy. In S. C. Guptill & J. L. Morisson (Eds.), *Elements of spatial data quality* (pp. 139-151), 1st Ed. Oxford: Elsevier Science Ltd.

Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science*, *237*(4820), 1317−1323.

Singh, A. (1989). Digital change detection techniques using remotely-sensed data. *International Journal of Remote Sensing*, *10*, 989−1003.

Skupin, A., & Fabrikant, S. I. (2003). Spatialization methods: A cartographic research agenda for non-geographic information visualization. *Cartography and Geographic Information Science*, *30*(2), 99−120.

Sowa, J. F. (2000). *Knowledge representation: Logical, philosophical, and computational foundations.* MIT Press.

SPSS Inc. (2005). *SPSS 14.0 for windows.* Chicago, IL: SPSS Inc.

USGS (2006a). *National Landcover Dataset 1992.* Retrieved 08/31, 2006, from http://landcover.usgs.gov/natllandcover.php

USGS (2006b). *National Landcover Dataset 2001.* Retrieved 02/20, 2007, from http://www.mrlc.gov/mrlc2k_nlcd.asp

Vogelman, J. E., Howard, S. M., Yang, L., Larson, C. R., Wylie, B. K., & Van Driel, N. (2001). Completion of the 1990 s national land cover data set for the conterminous united states for Landsat Thematic Mapper data and ancillary data sources. *Photogrammetric Engineering and Remote Sensing*, *67*(6), 650−655.

Woodcock, C. E., & Gopal, S. (2000). Fuzzy set theory and thematic maps: Accuracy assessment and area estimation. *International Journal of Geographical Information Science*, *14*(2), 153−172.

Yuan, F., Sawaya, K. E., Loeffelholz, B. C., & Bauer, M. E. (2005). Land cover classification and change analysis of the twin cities (Minnesota) metropolitan area by multitemporal Landsat Remote Sensing. *Remote Sensing of Environment*, *98*, 317−328.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, *8*, 338−353.