An Ecological Framework for Evaluating Map Errors Using Fuzzy Sets

John H. Lowry, Jr., R. Douglas Ramsey, Lisa Langs Stoner, Jessica Kirby, and Keith Schulz

Abstract
The use of fuzzy sets to assess uncertainty in land-use/cover maps provides a robust conceptual framework for examining unique characteristics of map error. By recognizing the possibility of gradations of error, fuzzy sets can be used to assess errors due to class similarity, or the sensitivity of the map legend to class boundaries. Building on the theoretical work of Gopal and Woodcock (1994), we present a practical methodology for assessing map errors using fuzzy sets. A key component of our methodology focuses on improving the decision-making process map experts assume when conducting a fuzzy set assessment of map errors. Using an ecological context to define varying levels of land-cover class similarity, we demonstrate how a decision framework guides the map experts' decisions and provides a more meaningful assessment of map errors. Our methodology differs from traditional fuzzy set error assessment methods in that the map expert evaluates misclassifications within the error matrix (off-diagonal cells) rather than individual reference sites. Advantages to a matrix-based approach include a reduction in the time required by map experts to evaluate map errors, and a relatively simple means of conveying map error information to the map user. We conclude that establishing criteria for determining multiple set memberships in a fuzzy set error assessment is an important methodological procedure that is commonly overlooked. Our methodology, designed to explicitly identify land-cover class similarities based on ecological criteria, serves as a practical example of how to address this issue.

Introduction
Assessing land-cover map accuracy is a significant concern for remote sensing-based mapping projects. The error matrix and kappa statistic have emerged as the *de facto* standard for evaluating map accuracy and presenting this information to the map user (Foody, 2002). Appeal for the error matrix lies partly in its simplicity, but also in its utility; it provides information about errors of commission and omission for individual map classes, as well as a quantitative measure of overall map accuracy in a single table (Congalton, 1991; Congalton and Green, 1999).

Other approaches aimed at delivering additional information to map users have been explored, notably information gained through error analysis based on fuzzy set theory. Zadeh (1965 and 1973) introduced the concept of fuzzy sets to deal with imprecision in complex systems, and the concept was extended to error assessment in remote sensing-based mapping by Gopal and Woodcock (1994). The applicability of fuzzy sets to error assessment in land-cover mapping arises from recognition that natural variation in land-use/cover types do not fit unambiguously into discrete classes. Fuzzy set theory in a mapping environment assumes that varying levels of membership are possible for multiple map categories. In an error assessment of land-cover maps this means that gradations of error are possible. Recognizing the possibility of gradations of error provides a robust conceptual framework for dealing with, and examining, unique characteristics of map errors, particularly errors attributed to class similarity. Class similarity can be defined as similarities between land-use/cover classes in regards to their composition and/or function.

A fundamental component of fuzzy set error assessment is the construction of a "linguistic measurement scale" to assign degrees of correctness for classification errors. Gopal and Woodcock (1994) suggest five levels of linguistic values that map experts can use when evaluating a map product relative to reference samples, these are: *absolutely wrong, understandable but wrong, reasonable or acceptable, good*, and *absolutely right*. Determining the appropriate error level for any given reference site is subject to the judgment of the "map expert." This subjectivity can be problematic when more than one expert is involved in the fuzzy set assessment process (Woodcock and Gopal, 1992). Determining "good" versus "reasonable or acceptable" is not only subject to the judgment of each expert, but the intended application of the map.

The fuzzy set framework has applications beyond assessing the accuracy of a single map. Fritz and See (2005) used fuzzy sets to describe uncertainty associated with differences in the classification legends for the Global Land Cover 2000 (Fritz *et al.*, 2003) and the MODIS land-cover products (Friedl *et al.*, 2002). Their methods reconciled differences between land-cover classes from the two maps by quantifying similarity among classes through surveys administered to mapping experts. Experts were asked to express how easy or difficult it was for them to differentiate land-cover classes from both legends. Using five linguistic categories ranging from *very easy* to *very difficult* [to differentiate] the authors demonstrated an effective approach for comparing land-cover maps, taking into consideration inherent differences in both the number

John H. Lowry, Jr., R. Douglas Ramsey and Lisa Langs Stoner are with the RS/GIS Laboratory, College of Natural Resources, Utah State University, Logan, UT (jlowry@gis.usu.edu).

Jessica Kirby is currently with the School and Institutional Trust Lands Administration, State of Utah, Salt Lake City, UT, and formerly with the RS/GIS Laboratory.

Keith Schulz is with NatureServe, Boulder, CO.

Photogrammetric Engineering & Remote Sensing
Vol. 74, No. 12, December 2008, pp. 1509–1519.

0099-1112/08/7412–1509/$3.00/0

of classes, and the class definitions for distinct map products. Other approaches using fuzzy sets for reconciling differences between distinct map legends have been presented by Hagen (2003) and Hagen-Zanker *et al*. (2005).

Previous use of fuzzy set theory for assessing uncertainty in land-use/cover maps suggests there is a critical need for more objective methods of defining the gradations between membership sets. Using a fuzzy set approach to uncertainty analysis in land-use/cover maps requires a mapping expert to make inherently subjective decisions regarding the appropriateness of multiple set memberships, often implicitly defined by some measure of class similarity. A primary objective of this paper is to present a methodological approach aimed at improving the decision-making process map experts assume when assessing map errors using fuzzy set theory. In addition, the paper demonstrates how this approach can be extended to the examination and characterization of map errors.

Background

The methods presented in this paper were developed, in part, through our effort to validate a regional land-cover map developed for the Southwest Regional Gap Analysis Project (SWReGAP). SWReGAP was a multi-state collaborative effort involving five land-cover mapping teams and a team of vegetation ecologists. The five-state region covers approximately 1.4 million square kilometers and includes the states of Arizona, Colorado, Nevada, New Mexico, and Utah. The region was divided into 25 spectral-physiographic mapping zones, which facilitated image classification and assisted in project tracking and management (Lowry *et al*., 2007). The land-cover legend was based on the Terrestrial Ecological System Classification framework developed by NatureServe (Comer *et al*., 2003). For the five-state region, approximately 93,000 ground truth samples were available for training and map validation.

Mapping responsibilities were divided among the five state teams by mapping zones, each state team having responsibility for three to six mapping zones. Each mapping zone functioned as a separate working unit and was independently assessed for map errors. Twenty percent of the available ground-based training samples from each land-cover class for each mapping zone were randomly selected and used for the validation exercise. Reference samples were intersected with the land-cover map and the output formatted as an error matrix with a corresponding kappa statistic (KHAT) (Congalton 1991; Congalton and Green 1999). The number of land-cover classes varied with the size and complexity of each mapping zone and ranged from 13 classes in the least complex mapping zone to 53 classes in the largest and most complex mapping zone.

Methods

A common approach to fuzzy set error assessment requires a "map expert" to assign multiple land-cover labels to each reference site as acceptable alternatives to the most acceptable, or "true" call (Green and Congalton, 2004). The mapping analyst is unaware of the actual mapped class for that specific reference site, thus assuring independence between the alternate labels and the land-cover map. Reference sites and the mapped land-cover classes are compared, determining the level of correspondence, or "correctness," between the reference data and the land-cover map.

The methodology we present differs from this common approach. Instead of evaluating individual reference sites, the mapping analyst evaluates misclassification errors (i.e., off-diagonals) within the error matrix using a decision framework of criteria established *a priori* that explicitly recognizes similarities among mapped land-cover classes. An important objective of developing these *a priori* criteria is to minimize the subjectivity inherent in the map expert's determination of acceptable alternatives to the most correct, or "true" label.

Establishing a Decision Framework

We begin by recognizing two requirements for establishing the decision framework for fuzzy set map error assessment. First, we identify the *context* within which we consider similarities among land-cover classes, and second we define *explicit criteria* for error evaluation given those recognized similarities. We suggest the context for recognized similarities be determined by the land-cover classification legend. For example, mapping projects dealing primarily with land-use may choose to identify similarities among land *uses* as the context for establishing evaluation criteria. Because the land-cover classification legend for SWReGAP focused on natural and semi-natural land-cover classes utilizing NatureServe's Ecological Systems as the primary mapping unit, we identify an *ecological* context for our evaluation criteria.

We recognize four basic *types* of ecological similarity among mapped land-cover classes (Table 1). *Ecological similarity types* are defined by specific ecological conditions shared by two distinct land-cover classes. For example, two land-cover classes may have the same physiognomic structure (Type A), determined by how the Ecological Systems nest within the 2001 National Land-cover Dataset (NLCD) classification legend (Homer *et al*., 2003). Some land-cover classes share dominant or diagnostic species (Type B) as identified by the Ecological System class descriptions published by NatureServe. Other classes may be commonly juxtaposed on the landscape, or may form a mosaic where patch or linear land-cover classes commonly occur within a matrix of another land-cover class (Type C). Finally, some land-cover classes may share similar special substrates (Type D).

Next, we recognize that any two land-cover classes may share multiple combinations of these basic ecological similarity types. Constructing a systematic ranking of all possible ecological similarity types provides a framework for assigning *ecological similarity categories* (Table 2). The idea of ecological similarity categories builds on Gopal and Woodcock's (1994) concept of "scaled linguistic values" to describe levels, or categories, of fuzzy set membership. An important difference with our approach, is that it explicitly defines an ecological context for determining membership using the combination of ecological similarity types to form criteria, or rules, for set membership (from Table 1). Another important difference is that our approach is designed to evaluate pairs of land-cover classes for recognized ecological similarities. For example, using this framework, an Inter-Mountain Basins Big Sagebrush Steppe land-cover class is considered "moderately similar" to an Inter-Mountain Basins Big Sagebrush Shrubland land-cover class by virtue of two ecological similarity types: Type B (shared dominant or diagnostic species) and Type C (common juxtaposition). This differs from more common fuzzy set assessment methods that focus on assigning individual reference sites to membership categories. Figure 1 presents additional examples of shared ecological similarity types and categories.

Computing Fuzzy Set Matrices: A Case Study

The ecological context and explicit ecological criteria established in Tables 1 and 2 form the foundation upon which we performed fuzzy set error assessments for each of the 25 mapping zones in SWReGAP (see **http://earth.gis.usu.edu/swgap/maperror.html**). Using these ecological criteria,

TABLE 1. DESCRIPTIONS OF FOUR BASIC *ECOLOGICAL SIMILARITY TYPES* WITH CORRESPONDING *ECOLOGICAL SIMILARITY CODES* (A, B, C, and D)

Ecological Similarity Code	Ecological Similarity Type	Ecological Similarity Description
A	Physiognomic Structure	Reference and mapped classes share the same National Land-cover Dataset (NLCD) class group: NLCD 30 Barren (Includes all Barren Lands) NLCD 40 Forest (Includes all Deciduous Forest, Evergreen Forest and Mixed Forest types) NLCD 50 Shrubland (Includes all Shrub, Dwarf Shrub and Shrub/Scrub types) NLCD 70 Grassland/Herbaceous (Includes all Grassland, Herbaceous, Savanna and Shrub Steppe types) NLCD 90 Wetlands (Includes all Wetland, Riparian, Emergent Wetlands, Wet Meadows, and Greasewood Flats)
B	Dominant Species Composition	Reference and mapped classes share dominant/diagnostic species as specified in concept of Ecological Systems. For example, if systems share *dominant* or *co-dominant* species, then species composition is similar. If systems share species that are only *present*, then species composition is not similar. Would also apply between systems where the dominant/co-dominant species is common, but has been identified to a different subspecies (e.g. *Artemisia tridentata* spp.).
C	Juxtaposition	Reference and mapped classes commonly form a mosaic, such as patch or linear systems occurring within matrix systems, or where broad ecotonal boundaries between the classes occur with regularity. This often relates to minimum mapping unit (scale) issues with mosaics of similar land-cover types. Refrain from using this code when the possibility of juxtaposition is only a rare occurrence.
D	Special Substrates	Reference and mapped classes share substrates with special properties that ecologically define each Ecological System. Apply with the following substrates only: - Eolian (sandsheets and dunes) - Bedrock (exposed weathering parent material); sparse vegetation (Barren) classes - High Salinity (exposed marine shales, saline overflow/playas)

TABLE 2. EXAMPLES AND EXPLANATIONS OF *ECOLOGICAL SIMILARITY CATEGORIES* BASED ON ECOLOGICAL SIMILARITY CODES (FROM TABLE 1), WITH CORRESPONDING RELATIVE SIMILARITY SCORES (RSS)

Ecological Similarity Code	Relative Similarity Category	Example	Explanation	Relative Similarity Score (RSS)
No Similarity (blank)	ABSOLUTELY INCORRECT	*Inter-Mountain Basins Cliff and Canyon (S009)* CONFUSED WITH *Great Basin Xeric Mixed Sagebrush Shrubland (S055)*	No *Ecological Similarity Types* are shared between these two Ecological Systems.	**1**
A C D	SOMEWHAT SIMILAR	*Rocky Mountain Aspen Forest and Woodland (S023)* CONFUSED WITH *Great Basin Pinyon-Juniper Woodland (S040)*	These two Ecological Systems are nested within the same NLCD Class (Forest—NLCD 40) and therefore share physiognomy (Similarity Code A). No other *Ecological Similarity Types* are shared. They are considered SOMEWHAT SIMILAR.	**2**
B AB AC AD BC BD CD	MODERATELY SIMILAR	*Inter-Mountain Basins Big Sagebrush Shrubland (S054)* CONFUSED WITH *Inter-Mountain Basins Montane Sagebrush Steppe (S055)*	These two Ecological Systems share dominant/diagnostic species (Similarity Code B) and are commonly juxtaposed on the landscape (Similarity Code C). They are considered MODERATELY SIMILAR.	**3**
ABC ABD ACD BCD ABCD	VERY SIMILAR	*Inter-Mountain Basins Big Sagebrush Shrubland (S054)* CONFUSED WITH *Great Basin Xeric Mixed Sagebrush Shrubland (S055)*	These two Ecological Systems share a common physiognomy (Similarity Code A), share dominant/diagnostic species (Similarity Code B) and are commonly juxtaposed on the landscape (Similarity Code C). They are considered VERY SIMILAR.	**4**
Diagonal Cell	ABSOLUTELY CORRECT	*Great Basin Pinyon-Juniper Woodland (S040) MAPPED AS Great Basin Pinyon-Juniper Woodland (S040)*	The reference and mapped classes are identical.	**5**

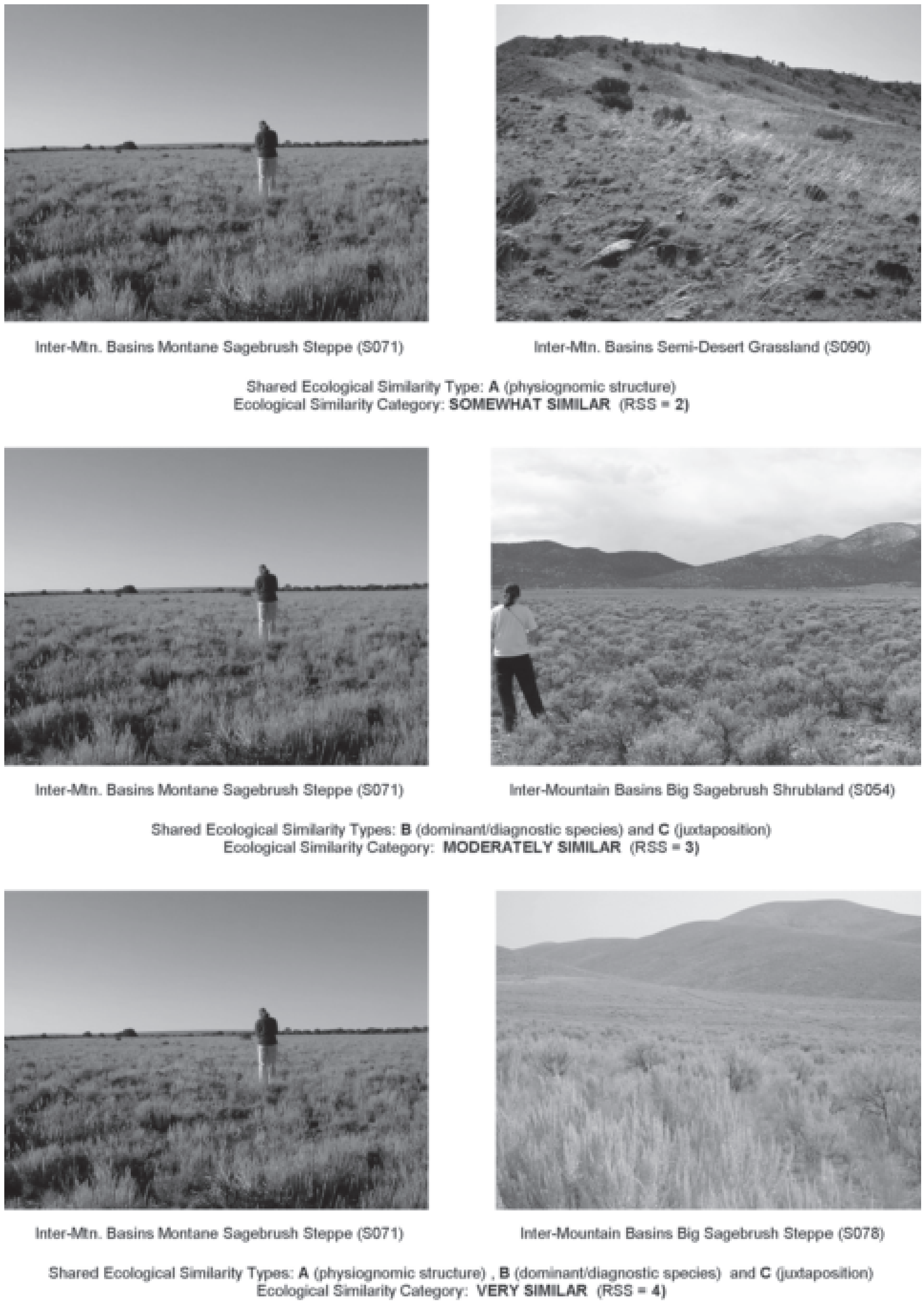

Figure 1. Examples of land-cover classes that are SOMEWHAT, MODERATELY, and VERY SIMILAR to the Inter-Mountain Basins Montane Sagebrush Steppe (S071) land-cover class, based on shared ecological similarity types (A, B, C, or D from Table 1). Ordinal *relative similarity scores* (RSS) corresponding to nominal *ecological similarity categories* are in parentheses.

mapping analysts (i.e., experts) constructed a series of *supporting* matrices (described below) used in conjunction with the error matrix to compute revised fuzzy set error matrices. Thus, the same ecological criteria from the decision framework were applied in the fuzzy set error assessment for all 25 mapping zones. We will use a small mapping zone in northeastern Nevada as a case study to demonstrate the ecological decision framework approach for map error assessment.

The initial supporting matrix is the *ecological similarity code* matrix. This matrix has the same structure as the original error matrix, with an equal number of columns and rows for the number of land-cover classes mapped. The analyst evaluates each paired combination of land-cover classes and assigns ecological similarity code(s) to each cell in the matrix (Table 3). Similarity code assignment is determined by examining the published description for each Ecological System from the classification legend and applying the criteria specified in Table 1. Multiple combinations of ecological similarity codes representing different types of ecological similarity are possible. Empty cells indicate no ecological similarity between paired land-cover classes, and self-similar land-cover classes (i.e., diagonal cells) are given a code of "X". The analyst then uses the criteria from Table 2 to translate the nominal ecological similarity codes (and combinations thereof) in Table 3 to identify similarity categories with ordinal *relative similarity scores* (RSS) (Table 4).

The final step is a cell-by-cell evaluation of the original error matrix (Table 5) in relation to the relative similarity score matrix (Table 4) which results in a computed "fuzzy set" matrix. During computation, misclassification errors in the off-diagonal cells from the original error matrix are added to the diagonal in accordance with their respective relative similarity score from Table 4. For example, in the "very similar" fuzzy set computation, cells with misclassification errors in the original error matrix corresponding to a RSS of 4 (very similar) are adjusted by adding the misclassification value (i.e., number of misclassified sites) to the diagonal cell in the computed fuzzy set matrix. Errors of omission and commission for the new matrix are re-calculated along with the kappa statistic.

Fuzzy set matrices may be computed two ways. Classification errors can be moved *vertically* to the diagonal "revising" errors of omission, or classification errors can be moved *horizontally* to the diagonal "revising" errors of commission. In this paper, we demonstrate the approach by moving off-diagonal errors vertically, thus revising errors of omission.

Results

Revised "fuzzy set" error matrices (with revised KHAT) are produced for each of three relative similarity categories: very similar (level 4), moderately similar (level 3), and somewhat similar (level 2). Fuzzy set error matrices for levels 1 and 5 are not produced because they either provide the same information as the original error matrix (level 5) or make all classes 100 percent accurate (level 1) which is not appropriate. Each new matrix is a revision of the original matrix based on increasingly liberal (i.e., RSS $\geq$ 4, RSS $\geq$ 3, and RSS $\geq$ 2) thresholds of ecological similarity. Table 6 presents, as an example, the computed error matrix for the *very similar* (RSS $\geq$ 4) relative similarity category.

Gopal and Woodcock (1994) demonstrate several approaches for extracting additional information during error analysis using fuzzy sets. Specifically, they show how information about the frequency, magnitude, source, and nature of errors can be explored and presented. Using our matrix-based approach, it is possible to explore and present information on the frequency and magnitude of error in a similar manner (see Lowry *et al.*, unpublished white paper at **http://earth.gis.usu.edu/swgap/**). Since the results of our methods are similar to Gopal and Woodcock's (1994) for *frequency* and *magnitude* of error, we refer readers to their work. Our matrix-based approach differs with respect to our evaluation and presentation of the *sources* and *severity* of confusion between classes within the land-cover map as a function of ecological similarities.

To assess the frequency of multiple set memberships at given thresholds of ecological similarity, we return to an evaluation of the original error matrix (Table 5) in relation to the RSS matrix (Table 4). The evaluation is facilitated by constructing a new matrix (Table 7a) that combines information

TABLE 3. *ECOLOGICAL SIMILARITY CODE* MATRIX DERIVED USING TABLE 1

LAND-COVER CLASS	Map Code	S009	S023	S028	S040	S050	S054	S055	S065	S071	S078	S090	S096	S118
Inter-Mountain Basins Cliff and Canyon	S009	X		C	C	C				C		C		
Rocky Mountain Aspen Forest and Woodland	S023		X	AC	A					C				C
Rocky Mountain Subalpine Dry-Mesic Spruce-Fir Forest and Woodland	S028	C	AC	X	AC					C				
Great Basin Pinyon-Juniper Woodland	S040	C	A	AC	X	C	C	C		C	C	C		C
Inter-Mountain Basins Mountain Mahogany Woodland and Shrubland	S050	C			C	X	AC	A	A	C	C			
Inter-Mountain Basins Big Sagebrush Shrubland	S054				C	AC	X	ABC	AC	BC	BC	C	C	C
Great Basin Xeric Mixed Sagebrush Shrubland	S055				C	A	ABC	X	AC	C	BC	C	C	
Inter-Mountain Basins Mixed Salt Desert Scrub	S065					A	AC	AC	X	C	C	BCD		
Inter-Mountain Basins Montane Sagebrush Steppe	S071	C	C	C	C	C	BC	C		X	ABC	A		C
Inter-Mountain Basins Big Sagebrush Steppe	S078				C	C	BC	BC	C	ABC	X	AC	C	C
Inter-Mountain Basins Semi-Desert Grassland	S090	C			C		C	C	C	A	AC	X	C	
Inter-Mountain Basins Greasewood Flat	S096							C	BCD		C	C	X	AC
Great Basin Foothill and Lower Montane Riparian Woodland and Shrubland	S118		C	C	C		C			C	C		AC	X

TABLE 4. *RELATIVE SIMILARITY SCORE* (RSS) MATRIX DERIVED FROM TABLE 2 AND TABLE 3

LAND-COVER CLASS	Code	S009	S023	S028	S040	S050	S054	S055	S065	S071	S078	S090	S096	S118
Inter-Mountain Basins Cliff and Canyon	S009	5	1	2	2	2	1	1	1	2	1	2	1	1
Rocky Mountain Aspen Forest and Woodland	S023	1	5	3	2	1	1	1	1	2	1	1	1	2
Rocky Mountain Subalpine Dry-Mesic Spruce-Fir Forest and Woodland	S028	2	3	5	3	1	1	1	1	2	1	1	1	1
Great Basin Pinyon-Juniper Woodland	S040	2	2	3	5	2	2	2	1	2	2	2	1	2
Inter-Mountain Basins Mountain Mahogany Woodland and Shrubland	S050	2	1	1	2	5	3	2	2	2	2	1	1	1
Inter-Mountain Basins Big Sagebrush Shrubland	S054	1	1	1	2	3	5	4	3	3	3	2	2	2
Great Basin Xeric Mixed Sagebrush Shrubland	S055	1	1	1	2	2	4	5	3	2	3	2	2	1
Inter-Mountain Basins Mixed Salt Desert Scrub	S065	1	1	1	1	2	3	3	5	1	2	2	4	1
Inter-Mountain Basins Montane Sagebrush Steppe	S071	2	2	2	2	2	3	2	1	5	4	2	1	2
Inter-Mountain Basins Big Sagebrush Steppe	S078	1	1	1	2	2	3	3	2	4	5	3	2	2
Inter-Mountain Basins Semi-Desert Grassland	S090	2	1	1	2	1	2	2	2	2	3	5	2	1
Inter-Mountain Basins Greasewood Flat	S096	1	1	1	1	1	2	2	4	1	2	2	5	3
Great Basin Foothill and Lower Montane Riparian Woodland and Shrubland	S118	1	2	2	2	1	2	1	1	2	2	1	3	5

from Tables 4 and 5. Here, the first numeral represents the number of reference sites in the cell (from Table 5) and the second numeral (in bold) indicates the RSS for that cell (from Table 4). Table 7b summarizes information in the combined matrix presenting the frequency of multiple set memberships from a producer's perspective. The frequency of multiple set memberships for each RSS threshold provides information about the relative heterogeneity (or generality) of the mapped land-cover classes based on gradations of recognized ecological similarity (i.e., RSS thresholds).

To provide a single metric of set membership errors we compute the *weighted mean* RSS for errors, which is the mean relative similarity score for each land-cover class. It is calculated by multiplying the frequency of errors for each RSS category by the RSS, summing the products for all RSS categories, and dividing by the total number of sites found off the diagonal (i.e., total number of errors). Formally it is expressed:

$$R_j = \frac{\sum_{i=1}^{n} (f_r * r)}{e_j}$$

where R_j is the *weighted mean* RSS of errors for land-cover class j, f is the frequency of errors for membership category r, e is the total number of errors for class j, and n is the total number membership categories that constitute errors (e.g., with five similarity categories there are four that represent some level of error, so $n = 4$). The weighted mean RSS value therefore, is an index of set membership based on the observed errors (Table 5) and recognized similarities (Table 4). It can also be thought of as a measure of the average severity of the errors for a given class due to class similarity, assuming that errors with less similarity are more severe. The index ranges from 1.00 to 4.00. A value of 1.00 indicates that, on average, the errors for a given land-cover class have no ecological similarity with the "true" or intended land-cover class, indicating high misclassification severity. A value of 4.00, on the other hand, indicates that on average, errors for a given class are "very similar" to the true land-cover class. If there are no errors, then the index is not applicable, and a value of NA is given. It should be noted that confidence in this metric will be determined by the number of error sites, or sample size, per land-cover class (Lohr, 1999).

The intent of set membership analysis is to identify possible sources of error by revealing degrees of similarity among misclassified land-cover classes. This is useful when interpreting the quality of the map product, but may have greater utility during early phases of the mapping process to determine whether land-cover class definitions need refinement. For example, from Table 7a we see that classes S054 and S055 are highly confused with each other. From a producer's perspective (i.e., errors of omission, or errors along the vertical axis) we note that S054 is confused with four other classes: S055, S071, S078, and S090 (for class names, see Table 7a). Including the relative similarity score (in bold) for each cell highlights the ecological similarity of the errors for each mapped class. For example, of the five reference sites omitted from the correct classification of S054, two of them were mapped as S055, which is considered ecologically *very similar* to S054.

In summary, by combining the RSS matrix with the original error matrix, map users can readily identify the severity of confused classes and map producers receive additional insight into the sources of map errors. In addition to facilitating error interpretation, the weighted mean RSS for errors provides a single metric of ecological similarity of the

Table 5. Original Error Matrix

	LAND-COVER CLASS	Code	REFERENCE													Total	User (%)
			S009	S023	S028	S040	S050	S054	S055	S065	S071	S078	S090	S096	S118		
MAPPED	Inter-Mountain Basins Cliff and Canyon	S009	5						1							6	83%
	Rocky Mountain Aspen Forest and Woodland	S023		4												4	100%
	Rocky Mountain Subalpine Dry-Mesic Spruce-Fir Forest and Woodland	S028			5											5	100%
	Great Basin Pinyon-Juniper Woodland	S040				17						1				18	94%
	Inter-Mountain Basins Mountain Mahogany Woodland and Shrubland	S050					1									1	100%
	Inter-Mountain Basins Big Sagebrush Shrubland	S054				1		54	12	2	2	6	3	1		81	67%
	Great Basin Xeric Mixed Sagebrush Shrubland	S055						2	8	1	2	1				14	57%
	Inter-Mountain Basins Mixed Salt Desert Scrub	S065							1	3						3	67%
	Inter-Mountain Basins Montane Sagebrush Steppe	S071	1	2			1	1	3		18	2	1	1		30	60%
	Inter-Mountain Basins Big Sagebrush Steppe	S078						1				0		1		2	0%
	Inter-Mountain Basins Semi-Desert Grassland	S090						1					3			4	75%
	Inter-Mountain Basins Greasewood Flat	S096							1					1		2	50%
	Great Basin Foothill and Lower Montane Riparian Woodland and Shrubland	S118													6	6	100%
		Total	6	6	5	18	2	59	25	6	22	9	8	4	6	176	
		Producers (%)	83%	67%	100%	94%	50%	92%	32%	33%	82%	0%	38%	25%	100%		70%
																	KHAT = 0.63

errors for each land-cover class, based on the observed errors and the criteria established by the ecological similarity decision framework.

Discussion

This paper demonstrates how a decision framework, based on pre-defined ecological similarity types, can be used to derive an ordinal relative similarity score matrix from which new, revised "fuzzy matrices" are computed for each RSS threshold. New user and producers accuracies are calculated, as is a new KHAT statistic.

Congalton and Green (1999) suggest errors within the error matrix can be attributed to four possible sources: (a) errors in the reference data, (b) sensitivity of the classification legend, (c) inappropriateness of remote-sensing technology for mapping a specific land-cover class, and (d) true "random" mapping error. The primary objective of error analysis is to learn more about *why* reference sites do not match their intended land-cover classes. The use of fuzzy sets for error analysis provides an ideal conceptual and objective framework for a closer examination of map errors by providing additional information beyond the binary right-or-wrong response. In general, the use of fuzzy sets in error assessment provides insight into the four potential sources of error outlined by Congalton and Green (1999). The focus of our methodology is to glean information from errors resulting from the complexity or generality of the classification legend which are indicative of the sensitivity of the classification legend. By default, if our analysis cannot attribute the source of error to the sensitivity of the classification legend, errors must be attributed to one or more of the other three sources.

As the field of remote sensing-based mapping matures, land-use/cover classification legends are becoming increasingly more complex (Green and Congalton, 2004). In general, this means that attempts are made to map a greater number of land-use/cover classes with subtler distinctions, sometimes referred to as ambiguous or fuzzy class boundaries (Gopal and Woodcock, 1994; Congalton and Green, 1999). To address the question of ambiguous class boundaries, this paper has focused on similarities between classes defined by their ecological composition and function. With categorical classifications based on a continuous field such as tree size or canopy closure, error variance between classes (e.g., 1 to 10 percent versus 11 to 30 percent canopy closure classes) can be addressed from a fuzzy set perspective by simply accepting as "correct," errors that fall plus or minus one class from the intended tree size or canopy closure class (Congalton and Green, 1999). Quantifying class variance for land-use/cover maps that are not based on a continuous

Table 6. Revised "Fuzzy Set" Matrix Computed from the Original Matrix (Table 5) and the *Relative Similarity Score* (rss) Matrix (Table 4) Using a rss Threshold of ≥4 (VERY SIMILAR). User and Producer Accuracies and khat are Recalculated. For Illustrative Purposes, Values that have Changed are in Bold Italics

	LAND-COVER CLASS	Code	REFERENCE													Total	User (%)
			S009	S023	S028	S040	S050	S054	S055	S065	S071	S078	S090	S096	S118		
MAPPED	Inter-Mountain Basins Cliff and Canyon	S009	5						1							6	83%
	Rocky Mountain Aspen Forest and Woodland	S023		4												4	100%
	Rocky Mountain Subalpine Dry-Mesic Spruce-Fir Forest and Woodland	S028			5											5	100%
	Great Basin Pinyon-Juniper Woodland	S040				17							1			18	94%
	Inter-Mountain Basins Mountain Mahogany Woodland and Shrubland	S050					1									1	100%
	Inter-Mountain Basins Big Sagebrush Shrubland	S054				1		***56***		2	2	6	3	1		***71***	***79%***
	Great Basin Xeric Mixed Sagebrush Shrubland	S055						2	***20***	1	2	1				***24***	***83%***
	Inter-Mountain Basins Mixed Salt Desert Scrub	S065							1	***3***						***4***	***75%***
	Inter-Mountain Basins Montane Sagebrush Steppe	S071	1	2			1	1	3		18		1	1		28	60%
	Inter-Mountain Basins Big Sagebrush Steppe	S078						1				***2***		1		***4***	***50%***
	Inter-Mountain Basins Semi-Desert Grassland	S090						1					3			4	75%
	Inter-Mountain Basins Greasewood Flat	S096												1		2	50%
	Great Basin Foothill and Lower Montane Riparian Woodland and Shrubland	S118													6	6	100%
		Total	6	6	5	18	2	***59***	***25***	***6***	22	***9***	8	4	6	176	
		Producers (%)	83%	67%	100%	94%	50%	***95%***	***80%***	***50%***	82%	***22%***	38%	25%	100%		***80%***
																KHAT = ***0.75***	

field is much more problematic, and underlies the goals of a context-based decision framework such as the one we describe in this paper.

A legitimate question regarding the use of fuzzy sets for map error assessment is whether or not we are actually dealing with the concept of "hierarchical classification." That is, when we conduct a fuzzy set error assessment, are we not revising our perspective of the errors by simply using a coarser thematic scale? We suggest that in a peripheral sense we are, but this is not the principle function or outcome of a fuzzy set error analysis. Instead, the intent of a fuzzy set error analysis is to provide a perspective that includes the recognition of acceptable levels of variance between and among land-cover classes. If the principle outcome of a fuzzy set assessment were simply a revised perspective from a coarser thematic scale, the result would be a summary of map errors with fewer *functional* classes. This distinction is important because the notion that a fuzzy set assessment is simply the "lumping up" of classes belies the utility gained through a better understanding of the sources of map error.

A key point of this paper, and the methodology we present, is that a more meaningful assessment of errors due to class similarity can be made if the context and criteria of class similarities are explicitly established. Our design for the decision framework for the SWReGAP maps focused on ecological similarities among land-cover classes because map units in the classification legend were defined by groups of plant communities with similar ecological processes, substrates and/or ecological gradients. This is important, as it makes interpretation of the fuzzy set analysis more meaningful and potentially useful to map users. As an example, the similarity between the Inter-Mountain Basins Big Sagebrush Shrubland (S054) and Great Basin Xeric Mixed Sagebrush Shrubland (S055) is considered "very similar" due to common physiognomic structure, dominant species, and juxtaposition on the landscape. This information is available to the map user using the ecological similarity matrix, and aptly describes why these two classes are considered very similar. Map users can use this information to decide whether the accuracy of the map is acceptable, given the ecological similarities between these classes. Providing revised "fuzzy set" matrices at multiple levels of similarity allows the map user to choose the level of similarity they wish to accept. For example, map users may find the revised error matrix based on the "very similar" threshold of similarity to be reasonable, but the revised error matrix based on the "moderately similar" threshold too liberal.

Table 7a. Combined Information from Original Matrix (First Numeral) and Relative Similarity Score Matrix (Second Numeral in Bold)

	LAND-COVER CLASS	Code	REFERENCE S009	S023	S028	S040	S050	S054	S055	S065	S071	S078	S090	S096	S118	Total	User (%)
MAPPED	Inter-Mountain Basins Cliff and Canyon	S009	5:**5**						1:**1**							6	83%
	Rocky Mountain Aspen Forest and Woodland	S023		4:**5**												4	100%
	Rocky Mountain Subalpine Dry-Mesic Spruce-Fir Forest and Woodland	S028			5:**5**											5	100%
	Great Basin Pinyon-Juniper Woodland	S040				17:**5**							1:**2**			18	94%
	Inter-Mountain Basins Mountain Mahogany Woodland and Shrubland	S050					1:**5**									1	100%
	Inter-Mountain Basins Big Sagebrush Shrubland	S054				1:**2**		54:**5**	12:**4**	2:**3**	2:**3**	6:**3**	3:**2**	1:**2**		81	67%
	Great Basin Xeric Mixed Sagebrush Shrubland	S055						2:**4**	8:**5**	1:**3**	2:**2**	1:**3**				14	57%
	Inter-Mountain Basins Mixed Salt Desert Scrub	S065							1:**3**	2:**5**						3	67%
	Inter-Mountain Basins Montane Sagebrush Steppe	S071	1:**2**	2:**2**			1:**2**	1:**3**	3:**2**		18:**5**	2:**4**	1:**2**	1:**1**		30	60%
	Inter-Mountain Basins Big Sagebrush Steppe	S078						1:**3**				0:**5**		1:**2**		2	0%
	Inter-Mountain Basins Semi-Desert Grassland	S090						1:**2**					3:**5**			4	75%
	Inter-Mountain Basins Greasewood Flat	S096								1:**4**				1:**5**		2	50%
	Great Basin Foothill and Lower Montane Riparian Woodland and Shrubland	S118													6:**5**	6	100%
		Total	6	6	5	18	2	59	25	6	22	9	8	4	6	176	
		Producers (%)	83%	67%	100%	94%	50%	92%	32%	33%	82%	0%	38%	25%	100%		70%

Table 7b. Summary of Combined Matrix (Table 7a) from a Producer's Perspective Showing Frequency of Errors for Each Set Membership Group (i.e., rss Threshold)

SET MEMBERSHIPS as frequency	S009	S023	S028	S040	S050	S054	S055	S065	S071	S078	S090	S096	S118
No membership (RSS = 1)							1					1	
Multiple set membership (RSS ⩾ 2)	1	2		1	1	1	3		2		5	2	
Multiple set membership (RSS ⩾ 3)						2	1	3	2	6			
Multiple set membership (RSS ⩾ 4)						2	12	1		2			
Single set membership (RSS = 5)	5	4	5	17	1	54	8	2	18		3	1	6
Total Sites	6	6	5	18	2	59	25	6	22	8	8	4	6
Total Errors	1	2	0	1	1	5	17	4	4	8	5	3	0
Weighted Mean RSS for Errors	2.00	2.00	NA	2.00	2.00	3.20	3.41	3.25	2.50	3.25	2.00	1.67	NA

An important advantage to a matrix-based approach to fuzzy set error analysis is the ability to calculate the kappa statistic (KHAT) at each similarity level (i.e., KHAT can be calculated for each revised "fuzzy set" matrix). KHAT measures the degree to which the agreement between the map and the reference data could occur by chance, and is determined by the relationship between the correctly mapped reference sites (diagonal cells) and "chance agreement" from the row and column totals (Congalton, 1991; Jensen, 2005). As the number of reference sites in the off-diagonals increase, KHAT decreases, indicating errors are occurring more randomly across all classes. When all the reference sites lie on the diagonal, KHAT is 1.0. Thus, it is often reported that a higher KHAT indicates there is a strong agreement between the reference data and the map, and a lower KHAT indicates poor agreement, or in other words, the results of the error matrix are more likely an occurrence of chance (Landis and Koch, 1977). A measure of variance around KHAT can also be calculated indicating the significance of the statistic at given confidence interval (e.g., 95 percent confidence) (Congalton and Green, 1999).

Using our approach, calculating KHAT for each of the revised "fuzzy matrices" based on RSS thresholds is carried out in the same manner as the original error matrix. As reference sites are taken from the off-diagonal cells and added to the diagonal cells, KHAT increases. (As demonstrated in this case study, KHAT increased from 0.63 in the original matrix to 0.75 in the revised fuzzy matrix computed with an RSS $\geq$ 4). Mathematically this is to be expected. Conceptually, this is understandable because the chance for agreement between the map and the reference data is being assessed under more liberal terms with each RSS threshold. In other words, the likelihood of chance agreement between the map and the reference data decreases as the requirements for agreement between the map and reference data become more liberal.

Conclusions

Our approach is unique among previous efforts to assess map uncertainty using fuzzy sets since the methodology involves evaluating errors within the error matrix rather than errors at individual reference sites. One of the advantages to this approach is that it reduces the amount of time it takes to evaluate map errors using fuzzy sets. Rather than evaluating individual reference sites, the mapping expert evaluates the off-diagonal cells in the error matrix, and makes decisions regarding set membership based on criteria established by the decision framework. An additional benefit of the error matrix format is that it offers a relatively simple means of conveying map error information to map users in a manner that is commonly understood (Foody, 2002). As demonstrated in Table 7a it is possible to extend the utility of the conventional error matrix by including additional information provided by relative similarity scores. It should be noted, however, that because our method focuses on an analysis of the error matrix rather than individual reference sites, it cannot detect variability among individual sample locations. In other words, the methodology assumes that all reference sites are labeled correctly.

Previous efforts (Woodcock and Gopal 1992; Laba *et al.*, 2002; Fritz and See, 2005) have used a fuzzy set approach for determining multiple set membership. In these efforts, experts used their mapping knowledge and experience to guide their decisions regarding acceptable alternative classes for multiple set memberships. Others (Hansen *et al.*, 2004) used decision rules to guide the fuzzy set assessment process, but these have generally been applied to relatively few land-cover classes within a single legend. One of the requirements of our methodology was the need to apply decisions to an undefined number of classes. (In SWReGAP there were 125 classes for the entire region, however not all 125 classes occurred in each mapping zone, and many mapping zones shared land-cover classes). By establishing a decision framework, rather than individual rules for each land-cover class, our methodology meets this requirement. We note that while the decision framework presented in this paper focuses on ecological similarities, it is the concept of a decision framework that can be applied to other land-cover mapping efforts, not necessarily the specific criteria presented in this paper. For example, a more liberal or conservative decision framework could be constructed. The merit of a decision framework is that it provides a standardized decision process multiple mapping experts can use to determine "acceptable" misclassifications at different levels of class similarity for an undefined number of land-cover classes. This does not completely remove subjectivity from the decision-making process, but instead channels subjectivity to pre-established bounds defined within an accepted context.

Acknowledgments

The authors thank Chris Garrard for programming the script used to compute fuzzy set matrices. The script is written as an ArcView® 3.X extension and is available free of charge from **http://earth.gis.usu.edu/swgap/mapquality.html.** We thank Bill Adair and two anonymous reviewers for the helpful criticism of earlier drafts of this paper. Finally, we also thank individuals involved in the SWReGAP mapping effort who helped implement and refine some of the methods described herein. These individuals are: Patrick Comer, Keith Pohs, Todd Sajwaj, Kathryn Thomas, Don Schrupp, Scott Schrader, Eric Waller, and Bret Wolk.

References

Comer, P., D. Faber-Langendoen, R. Evans, S. Gawler, C. Josse, G. Kittel, S. Menard, M. Pyne, M. Reid, K. Schulz, K. Snow, and J. Teague, 2003. *Ecological Systems of the United States: A Working Classification of U.S. Terrestrial systems*, Arlington, Virginia: NatureServe.

Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data, *Remote Sensing of Environment*, 37:35–46.

Congalton, R.G., and K. Green, 1999. *Assessing the Accuracy of Remote Sensed Data: Principles and Practices*, Lewis Publishers, Boca Raton, Florida, 137 p.

Friedl, M.A., D.K. McIver, J.C.F. Hodges, X.Y. Zang, D. Muchoney, A.H. Strahler, C.E. Woodcock, S. Gopal, A. Schneider, A. Cooper, A. Baccini, F. Gao, and C. Schaaf, 2002. Global land-cover mapping from MODIS: Algorithms and early results, *Remote Sensing of Environment*, 83:287–302.

Fritz, S., E. Bartholomé, A. Belward, A. Hartley, H.J. Stibig, H. Eva, P. Mayaux, S. Bartalev, R. Latifovic, S. Kolmert, P. Roy, S. Agrawal, W. Binfang, X. Wenting, M. Ledwith, F.J. Pekel, C. Giri, S. Mücher, and E. de Badts, 2003. *Harmonisation, Mosaicing and Production of the Global Land Cover 2000 Database (Beta Version)*, Luxembourg: Office for Official Publications of the European Communities.

Fritz, S., and L. See, 2005. Comparison of land-cover maps using fuzzy agreement, *International Journal of Geographical Information Science*, 19(17):787–807.

Foody, G., 2002. Status of land-cover classification accuracy assessment, *Remote Sensing of Environment*, 80:185–201.

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

Green, K., and R. Congalton, 2004. An error matrix approach to fuzzy accuracy assessment: The NIMA Geocover Project, *Remote Sensing and GIS Accuracy Assessment* (R.S. Lunetta and J.G. Lyon, editors), CRC Press, Boca Raton, Florida, pp. 163–172.

Hansen, M., J. Coles, K.A. Thomas, D. Cogan, M. Reid, J. Von Loh, and K. Schulz, 2004. *USGS-NPS National Vegetation Mapping Program: Wupatki National Monument, Arizona, Vegetation Classification and Distribution, Final Report*, U. S. Geological Survey, Southwest Biological Science Center, Flagstaff, Arizona.

Hagen, A., 2003. Fuzzy set approach to assessing similarity of categorical maps, *International Journal of Geographical Information Science*, 17(3):235–249.

Hagen-Zanker, A., B. Straatman, and I. Uljee, 2005. Further developments of a fuzzy set map comparison approach, *International Journal of Geographical Information Science*, 19(7):769–785.

Homer, C., C. Huang, L. Yang, B. Wylie, and M. Coan, 2003. Development of a 2001 National Land Cover Database for the United States, *Photogrammetric Engineering & Remote Sensing*, 69(7):829–840.

Jensen, J., 2005. *Introductory Digital Image Processing*, Pearson Education, Upper Saddle River, New Jersey, 526 p.

Laba, M., S.K. Gregory, J. Braden, D. Oguracak, E. Hill, E. Fegraus, J. Fiore, and S.D. DeGloria, 2002. Conventional and fuzzy accuracy assessment of the New York Gap Analysis Project land cover map, *Remote Sensing of Environment*, 81:443–455.

Landis, J., and G. Koch, 1977. The measurement of observer agreement for categorical data, *Biometrics*, 33:159–174.

Lohr, S.L., 1999. *Sampling: Design and Analysis*, Brooks/Cole Publishing Co., Pacific Grove, California, 494 p.

Lowry, J., R.D. Ramsey, K. Thomas, D. Schrupp, T. Sajwaj, J. Kirby, E. Waller, S. Schrader, S. Falzarano, L. Langs, G. Manis, C. Wallace, K. Schulz, P. Comer, K. Pohs, W. Reith, C. Velasquez, B. Wolk, W. Kepner, K. Boykin, L. O'Brien, D. Bradford, B. Thompson, and J. Prior-Magee, 2007. Mapping moderate-scale land-cover over very large geographic areas within a collaborative framework: A case study of the Southwest Regional Gap Analysis Project (SWReGAP), *Remote Sensing of Environment*, 108:59–73.

Woodcock, C., and S. Gopal., 1992. Accuracy assessment of the Stanislaus Forest vegetation map using fuzzy sets, *Proceedings of the 4th Biennial Remote Sensing Applications Conference*, Orlando, Florida, 17 April, pp. 378–194.

Zadeh, L., 1965. Fuzzy sets, *Information and Control*, 8:338–353.

Zadeh, L., 1975. Outline of a new approach to the analysis of complex or imprecise concepts, *IEEE Transactions: Systems, Man and Cybernetics*, SMC-3:28–44.

(Received 23 February 2007; accepted 23 April 2007; revised 28 June 2007)