A Comparison of Fuzzy Models

in Similarity Assessment

of Misregistered Area Class Maps

by

Scott Brown

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Arts

Approved November 2010 by the
Graduate Supervisory Committee:

Elizabeth Wentz, Chair
Soe Myint
Sharolyn Anderson

ARIZONA STATE UNIVERSITY

December 2010

ABSTRACT

Spatial uncertainty refers to unknown error and vagueness in geographic data. It is relevant to land change and urban growth modelers, soil and biome scientists, geological surveyors and others, who must assess thematic maps for similarity, or categorical agreement. In this paper I build upon prior map comparison research, testing the effectiveness of similarity measures on misregistered data. Though several methods compare uncertain thematic maps, few methods have been tested on misregistration. My objective is to test five map comparison methods for sensitivity to misregistration, including sub-pixel errors in both position and rotation. Methods included four fuzzy categorical models: fuzzy kappa's model, fuzzy inference, cell aggregation, and the epsilon band. The fifth method used conventional crisp classification. I applied these methods to a case study map and simulated data in two sets: a test set with misregistration error, and a control set with equivalent uniform random error. For all five methods, I used raw accuracy or the kappa statistic to measure similarity. Rough-set epsilon bands report the most similarity increase in test maps relative to control data. Conversely, the fuzzy inference model reports a decrease in test map similarity.

ACKNOWLEDGMENTS

I gratefully acknowledge the support and encouragement I received from my advisor, Dr. Elizabeth Wentz, without whom none of this would be possible. I also give thanks for the helpful and positive feedback of my graduate committee.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

Chapter 1

INTRODUCTION

In a virtual workshop in 2004, Boots and Csillag (2006) identified four common reasons to compare area-class maps: to assess accuracy in a map relative to ground truth (Congalton and Green 2009), to detect change over time in a mapped area (Fisher et al. 2006), to assess predictive models of such temporal change (Pontius, Huffaker, and Denman 2004), and to compare landscape patterns across space and time (White 2006). In general, such comparisons propose to discover some meaningful change or difference between two maps. However, unknown error in either map can show false changes or conceal interesting differences. In this context, unknown imprecision and error inherent in geographic data is referred to as uncertainty (Goodchild 2008). The rising popularity of area-class maps classified using areal imagery coupled with a growing interest in land change analysis necessitates a critical understanding of how common sources of uncertainty affect map comparison. Research addressing spatial uncertainty falls loosely into six categories: qualifying sources of uncertainty, such as in geocoding (Karimi, Durcik, and Rasdorf 2004), rasterizing vector data (Bregt et al. 1991) or historical map digitization (Bolstad, Gessler, and Lillesand 1990; Leyk, Boesch, and Weibel 2005); modeling fuzzy (Schneider 2008), rough (Worboys and Clementini 2001) and probabilistic data objects (Cheng 2009); measuring accuracy in crisp (Congalton and Green 2009) and fuzzy data (Ricotta 2005); simulating map variation (Goodchild and Dubuc 1987; Parker et al. 2003; Grunwald 2009); assessing uncertainty propagation and sensitivity in derived products (Crosetto and Tarantola 2001), such as in slope stability assessment (Davis and Keller 1997); and communicating vagary to users (e.g. Ehlschlaeger and Goodchild 1994). My research on comparison methods' sensitivity to misregistration relates most to the study of uncertainty propagation. Map comparison methods commonly provide both local results in the form of difference maps and global results as measures of overall map similarity. My work emphasizes the latter, where similarity is a single metric of agreement between compared maps, measured in terms of class proportion and distribution. In this chapter I address two categories of uncertainty research; I qualify misregistration as a source of uncertainty in area-class map comparison, and discuss philosophies for modeling such uncertainty. Similarity measures and map comparison are discussed further in chapter two.

Misregistration is a significant source of uncertainty in area-class map comparison (Boots and Csillag 2006). Area-class maps divide space into bounded regions, where class membership describes attributes common to each region. Also called thematic maps and nominal fields, these categorical maps are unique in that geometry and attributes are interdependent. This dependency forms boundary areas between classes that are distinct from lines and other cartographic phenomena, where points may belong to several classes, to no class, or to a class not addressable at this scale (Mark and Csillag 1989; Rossiter 2001; Goodchild 2003). Zhang and Goodchild (2002) describe these as unique phenomena in spatial uncertainty. Rossiter (2001) and Foody (2002) review the complex interaction of area-class boundaries with several sources of uncertainty, including sensor imprecision, misregistration, and issues of scale, as well as classifier bias and taxonomy fitness. Misregistration produces two patterns of error in area-class maps. In the first, movement of class boundaries results in highly autocorrelated geometric error, where uncertainty in the class identity at one point increases class uncertainty in nearby points. This phenomenon may obscure true land cover boundary change (Foody 2002). The second error pattern relates to the Modifiable Areal Unit Problem (MAUP), where an offset sample area may change identity due to mixed-pixel class confusion. Independent studies show that misregistration by one-fifth of a pixel may cause a 10% error in measured similarity (Townshend et al. 1992; Dai and Khorram 1998). In general, Foody (2002) associates most error observed along thematic boundaries with these two patterns of uncertainty. The deleterious effect of misregistration on map comparison has motivated further research.

Although methods exist to minimize misregistration of area-class data (Bruzzone and Prieto 2000; Stow and Chen 2002; Bruzzone and Cossu 2003), some uncertainty will always remain. Many classical map comparison statistics do not address the spatial distribution of error within a map (Foody 2002). Goodchild (2003) concludes that such measures cannot address area-class uncertainty. Recognizing this challenge, researchers in map comparison adapt crisp statistics using formal uncertainty models (Hagen 2002; Leyk, Boesch, and Weibel 2005). Following the success of this adaptation, modeling uncertainty has become a fundamental component of map similarity assessment (Kuzera and Pontius 2008).

Fuzzy set theory may be the most prevalent foundation for uncertainty models, but it is not without competition from multi-valued logics and probability theory (Elkan et al. 1994). Uncertainty modeling began in the 1920's with the first trivalent logic published by Łukasiewicz (1970), admitting values for *true*, *false*, and *possible*. This work later developed into rough set theory, which defined a set as two collections of definite and possible members (Worboys and Clementini 2001). Łukasiewicz's work was precursor to the seminal introduction of fuzzy sets in 1965 (Zadeh). Zadeh (1978) generalized rough sets to include continuous degrees of set membership ranging from zero to one, formalizing a logical calculus he termed *possibility theory*. Where a rough set may represent a forest boundary as a constant band of *possible* membership, fuzzy sets allow a continuous transition between areas of definite inclusion and exclusion (Mark and Csillag 1989). A prototypical example from Zhu (1997) represents vague soil types using fuzzy class membership. Bennett (2001) raises serious questions on the utility of fuzzy set theory over multi-valued logics. Despite this, Kay (1994) observes fuzzy approaches in widespread use. For a full treatment of this ambiguity, see (Elkan et al. 1994). Probabilistic logic also developed in parallel with fuzzy set theory (Nilsson 1986). Probabilistic representations describe the likelihood of an event. In this logic, event outcomes are known, but conditions driving an event are not. While numerically similar, fuzzy sets and probability theory are nevertheless philosophically distinct in terms of *what* each represents. Probability theory represents uncertainty of a process, while fuzzy set theory represents uncertainty in a result (Gudder 2000). To illustrate, Zadeh (1968) shows propositions such as “in eight coin tosses there are *many* more tails than heads,” which describes probabilistic processes working on fuzzy concepts, suggesting that probability theory and fuzzy logic may be complimentary (Zadeh 1995).

Models termed *fuzzy* include many intellectual descendents of Zadeh’s (1965) seminal work that can represent registration uncertainty in area-class maps. Here the umbrella term may refer to any model that generalizes crisp membership to represent vague notions. In general, these models represent uncertainty in pixel location or category using a fuzzy pixel with partial membership in nearby or similar classes (Hagen 2003). Such models vary in their approach to fuzziness. For instance, areal aggregation techniques generalize class membership to a coarse geometry, describing regions by the *proportion* of underlying class membership (Pontius, Huffaker, and Denman 2004). This effectively smoothes away small errors in class geometry shown in registration error (Carmel 2004). Conversely, models related to information theory may describe an object using continuous values of *belief* in uncertain statements (Zadeh 1975; Shafer 1976). With proper inference rules, such an approach may also suit misregistration. Rough set theory may prove useful too, as egg-yolk models effectively generalize *possible* boundaries while maintaining crisp class membership elsewhere (Cohn and Gotts 1996). Any of these fuzzy models can adapt a crisp similarity statistic to cases of uncertain registration, but the question of their performance at this task remains open (White 2006).

Sundaresan, Varshney, and Arora (2007) assessed sensitivity of a fuzzy similarity measure under misregistration of remotely sensed land cover images. They compared Markov random field change detection (MRFCD) against image differencing (ID) with a neighborhood softening technique comparable to that used by Hagen (2002), showing that MRFCD is more robust to linear and rotational misregistration than softened ID. Although their study may have been the first to compare different similarity measures under misregistration (Sundaresan, Varshney, and Arora 2007), assessment of only two fuzzy measures limits their results. The unique qualities of misregistered thematic maps should inform the choice of fuzzy model used in map comparison (Stehman 1999; Cross and Sudkamp 2002). In my research I assess registration sensitivity in several fuzzy models used for map comparison. This assessment will address the question of which thematic map similarity measures can better differentiate between registration error and legitimate classification change.

To assess model sensitivity, I will adapt the approach used by Sundaresan, Varshney, and Arora (2007) for use on categorical maps. I examine four fuzzy models for registration sensitivity: the raster-based methods called fuzzy kappa (Hagen 2003) and cell aggregation (Pontius, Huffaker, and Denman 2004), and the vector-based methods known as fuzzy inference (Power, Simms, and White 2001) and epsilon bands (Mark and Csillag 1989). These models were chosen to provide a representative sample of different approaches that could dampen registration error. Test data included 22 simulated area-class maps and a case study region in northwestern Arizona. These were progressively misregistered in both liner and rotational patterns. For each misregistration, I created a control map with identical error magnitude and random error distribution. I then compared similarity metric results, using crisp similarity as a performance baseline.

As discussed in this chapter, fuzzy models are integral components to map comparison that may reduce registration sensitivity. Area-class map boundaries are unique phenomena significant to change detection research. Classification uncertainty due to misregistration increases along these boundaries, which can cause significant error in map similarity measures. Error along such boundaries should inform fuzzy model selection to reduce similarity measure error (Stehman 1999; Cross and Sudkamp 2002). In the following chapter, I address literature in map comparison by component, examining specific fuzzy models, per-pixel comparison techniques, and confusion matrix rules. Chapter three describes experimental test data and model parameters, whose results are presented in chapter four. I discuss these results critically in chapter five. Finally, chapter six supports conclusions of my research and future avenues for advancement.

Chapter 2

BACKGROUND

The findings of this study contribute to literature in thematic map comparison. Boots and Csillag (2006) identify four pursuits in comparison research: accuracy assessment comparing maps to data with known error (Rossiter 2001; Foody 2002), change detection comparing maps over time (Mas 2005; Radke et al. 2005), validation of land use simulation against known data (Hagen 2003; Pontius, Huffaker, and Denman 2004), and landscape pattern analysis measured by dozens of assorted metrics (McGarigal and Marks 1995; Zhu 1997; Gustafson 1998; White 2006; Williams and Wentz 2008). These categories are not independent; they share several important comparison techniques. For example, Hagen (2003) designed a fuzzy kappa measure to measure simulation validity and to perform multi-temporal land change comparison, and the approach by Power, Simms, and White (2001) analyzes landscape patterns within simulated maps. Furthermore, when classifiers and class definitions are well known, change detection and accuracy assessment are strongly related, and they share use of the confusion matrix (Rees 2008).

Although the breadth of research in these four areas of study limits generalization of the field as a whole, I describe fuzzy comparison methods according to three common functional components: a fuzzy softening technique, a set of local comparison functions, and rules for deriving global statistics. The general process for map comparison using these three components is illustrated in Figure 1. Functions for local comparison of fuzzy pixels or polygons are the point

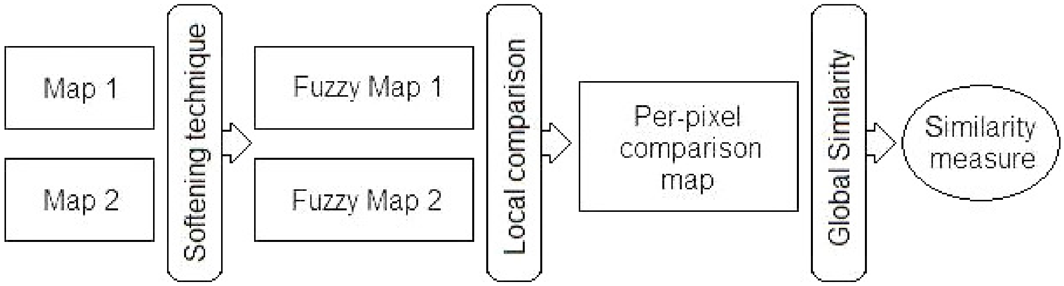

Figure 1. A flow-chart of the general map comparison process

of contact between softening techniques for fuzzy models and global similarity measures. In this chapter I address each component, beginning with techniques for softening crisp maps into fuzzy models. This leads into a discussion of local comparison functions. Following that, I discuss global similarity statistics calculated from local comparisons.

I. *Softening Techniques*

Published measures of area-class map similarity have implicit models of class membership. Traditional crisp measures define areas as polygons or raster cells, assigning one class per area and one area to every point in the study (Foody 2006; Rees 2008). Power, Sims and White (2001) claim that such classical approaches are extremely susceptible to errors in map registration, and are liable to propagate such errors. As discussed in chapter one, fuzzy models for class membership can reduce this risk. Fuzzy measures of map similarity take advantage of this with a softening technique to represent crisp map data in a fuzzy model. Formally, I define a crisp area-class map as a set of pixels or polygons $\boldsymbol{P}$ classified into a set of categories $\boldsymbol{N}$ by $\boldsymbol{U} \in \boldsymbol{N}^{\boldsymbol{P}} \equiv \{f : \boldsymbol{P} \rightarrow \boldsymbol{N}\}$. A softening technique μ then takes the form $\mu : \boldsymbol{N}^{\boldsymbol{P}} \rightarrow \boldsymbol{S}^{(\boldsymbol{P} \times \boldsymbol{N})}$, $\boldsymbol{S}^{(\boldsymbol{P} \times \boldsymbol{N})} \equiv \{f : \boldsymbol{P}, \boldsymbol{N} \rightarrow \boldsymbol{S}\}$, where $\boldsymbol{S}$ is the set of qualifiers in the underlying fuzzy model. For example, $\boldsymbol{S}$ is the realvalued interval (0,1) for fuzzy sets, or ***S***={*true*, *false*, *possible*} for Łukasiewicz's (1970) trivalent logic. Softening techniques that increase spatial autocorrelation in crisp maps have been shown to limit the effect of registration error (Power, Simms, and White 2001).

Power, Simms, and White (2001) soften classified data through a novel method based on Mamdani fuzzy inference. Built to model the vagary of human speech, Mamdani inference compiles fuzzy evidence with rules to measure strength of a theory (Jang 1993). This model is epistemologically distinct from probabilistic or possibilistic models, as it describes *belief* in a given class membership. To soften crisp classifications, unique polygons are formed and weighted based on their areal proportion of category agreement. Fuzzy linguistic labels such as *poor* and *very good* are then applied to describe match quality. This construction allows fuzzy inference to compare high-level patterns in thematic maps while ignoring small changes (White 2006). The approach differs slightly from the process illustrated in Figure 1, as polygon overlay unifies compared data into a single map prior to softening. Map comparison using polygon overlay has been widely critiqued due to the occurrence of spurious polygons (Chrisman 1989; Goodchild, Guoqing, and Shiren 1992). Power, Simms, and White partially address this by using polygon area as a weight. Discrete and continuous epsilon bands may further address the issue of spurious polygons in map comparison (Worboys and Clementini 2001; Mas 2005).

Epsilon bands represent one of the earliest and most pervasive models of area-class map uncertainty (Mark and Csillag 1989). The model softens nominal data using a distance decay function for class membership near boundaries, while areas outside this epsilon distance band retain crisp membership in their attendant class. Fuzzy membership then describes class *proximity*, in contrast to the class *belief* described in Mamdani inference. Originally presented in 1958 by Julian Perkal (1986) as a methodology for measurement and comparison of planar curves, the first of these models was a simple inclusion-exclusion model strongly related to egg-yolk models (Cohn and Gotts 1996) and rough set theory (Worboys and Clementini 2001). Mark and Csillag (1989) later applied a probabilistic interpretation, proposing several cumulative normal functions for class membership within epsilon bands, which have since been applied to land cover change detection (Mas 2005).

Goodchild and Dubac (1987) criticized the epsilon band approach as failing to represent autocorrelative effects of error, noting that neighboring points along a line are not independent samples. Elaborating, Goodchild (2003) stated that any general model of area-class map uncertainty must address class membership confusion at every point, observing that the epsilon band model does not address confusion beyond the distance band. However, these criticisms do not apply to the case of misregistration error. Misregistration is autocorrelated along class boundaries, and the epsilon band reflects this correlation. A model of error along class boundaries may have no need to address class confusion elsewhere. Mark and Csillag (1989) also criticize Perkal's (1986) discrete formulation, proposing that class membership should vary continuously within epsilon bands.

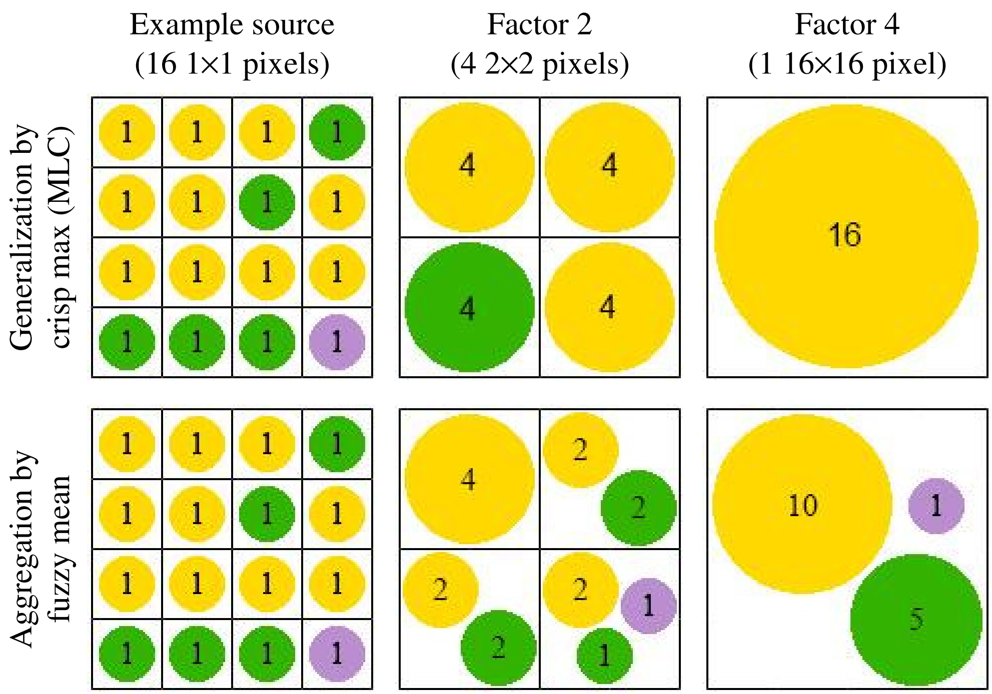

Figure 2. Crisp generalization performed by the maximum likelihood classifier, and fuzzy cell aggregation performed by fuzzy mean

Recent literature has disagreed with this position, questioning the utility of continuous models for uncertain objects (Bennett 2001).

The field of multiple resolution map comparison has identified numerous methods for reducing the number of cells or polygons within a map (Costanza 1989; Pontius 2002; Kuzera and Pontius 2008). Land use simulations are traditionally validated at an aggregate scale, obscuring individual pixel locations in favor of overall classification pattern (Boots and Csillag 2006). The most common aggregation approach calculates the *areal proportion* of class membership in a square neighborhood, replacing that neighborhood with a single fuzzy pixel (Pontius, Huffaker, and Denman 2004). Figure 2 illustrates fuzzy raster aggregation by factors of two and four. Note that location information is lost while global class membership remains constant. A crisp equivalent is also shown, using a maximum likelihood classifier instead of areal proportion. Because class information is lost, the crisp process is referred to as generalization rather than aggregation. Carmel (2004) found that aggregation is effective at reducing registration error when using a factor 3-10 times as large as the average pixel offset distance. However, these results were established against only a single source map.

II. *Local Comparison*

While comparison functions for crisp polygons or pixels are obvious and intuitive, extension of these principals to their softened counterparts is not. Crisp categorical similarity at the local scale is described by a binary function, which yields *true* (or one) when class membership is equal, and *false* (or zero) otherwise. Difference is traditionally defined as similarity's compliment; crisp difference then is the binary negation of this similarity measure. However, there is no single agreed upon definition for similarity of vague categorical maps. The *extension principal* from fuzzy set theory allows crisp measures of global similarity to be used on fuzzy data (McBratney and Odeh 1997; Woodcock and Gopal 2000). Formally, the extension principal allows that for similarity measures $f : \boldsymbol{N^P}, \boldsymbol{N^P} \rightarrow \Re$ relating crisp maps $\boldsymbol{U}, \boldsymbol{U'} \in \boldsymbol{N^P}$ to a real value using crisp set operations, any softening method of the form $\mu : \boldsymbol{N^P} \rightarrow (0,1)^{(\boldsymbol{P} \times \boldsymbol{N})}$ can be composed with f such that $f(\mu(\boldsymbol{U}), \mu(\boldsymbol{U'})) = v$, $v \in \Re$ (Cross and Sudkamp 2002). However, the extension principal does not provide a unique mapping, which means that a crisp similarity measure may have several fuzzy equivalents.

Discrete multi-valued logics such as rough sets allow multiple comparison functions. If comparison results are constrained to the output set {*true*, *false*, *possible*}, Worboys and Clementini (2001) note that functions must bias either toward definite or indefinite response. To illustrate, consider a pair of cells $a,b \in \boldsymbol{P}$ in the softened map $\mu(\boldsymbol{U})$ defined over three classes, *a* with known membership (*true*, *false, false*) and *b* with vague membership (*possible*, *false, possible*). A similarity result of *possible* would bias toward indefinite response. A result of *true* or *false* shows definite bias. Importantly, similarity and difference can now have independent meaning (Cross and Sudkamp 2002). For example, dissimilarity of cells may be given definite positive (*true*-leaning) bias while similarity is assigned indefinite bias. For rough set models, it may be inappropriate to measure continuous similarity (Bennett 2001).

Fuzzy set theory supports numerous varieties of continuous local comparison. While crisp sets use the set *difference* operation to measure similarity, Cross and Sudkamp (2002) report that fuzzy set operations *union*, *intersection*, *compliment*, and *difference* suffer ongoing contention among researchers, leaving no formulation generally accepted. Fisher et. al. (2006) favor the Bounded Difference over traditional fuzzy intersection for local similarity. Kuzera and Pontius (2008) compare a local similarity measure based on multiplication with one based on the arithmetic minimum. However, they only examine fuzzy membership values that sum to unity; some softening techniques do not share this constraint (Cross and Sudkamp 2002). In addition, some fuzzy measures (e.g. Power, Simms, and White 2001) define similarity as noncommutative, meaning that the order of maps matters during comparison, or more formally that $\exists \boldsymbol{U},\boldsymbol{U}'\in \boldsymbol{N}^{\boldsymbol{P}} \mid f(\mu(\boldsymbol{U}), \mu(\boldsymbol{U}')) \neq f(\mu(\boldsymbol{U}'), \mu(\boldsymbol{U}))$.

III. *Global Similarity*

Similarity statistics for area-class maps seek improved numeric estimates for the vague human notion of compatibility (Cross and Sudkamp 2002). The earliest and most direct of these, a ratio of common to total map area, has been called *classification accuracy*, the *coefficient of areal agreement*, *fraction correct*, *area of agreement*, *crisp difference*, and *raw similarity* among other terms (Foody 2002; Foody 2006). Accuracy statistics are often used to measure map similarity; this is justified in fuzzy maps when class definitions do not change (Rees 2008). Raw accuracy is widely agreed to overestimate similarity (Power, Simms, and White 2001), prompting further methods research. Accuracy provides an excellent example of global similarity measures' dependence on local comparison. In rough sets, accuracy measures with definite bias will clearly overestimate crisp classification accuracy, while indefinite bias will underestimate the crisp case (Cohn and Gotts 1996).

The confusion matrix is perhaps the most common and well established method for quantifying accuracy for nominal maps. It is alternately known as the error matrix, contingency matrix, and the cross-tabulation matrix (Pontius, Shusas, and McEachern 2004; Rees 2008). The matrix describes omission and commission errors in nominal data, giving a per-class fractional accuracy for a map along with an overall fraction correct. There are countless methods to compile confusion matrices with desirable properties by varying local comparison techniques (Kuzera and Pontius 2008).

The Kappa statistic was adapted from the confusion matrix by Cohen (1960) to account for chance agreement due to category frequency. Later studies showed that this statistic consistently underestimates map similarity (Congalton 1991). It follows that all results of map similarity assessment should fall between raw accuracy and kappa measures (Power, Simms, and White 2001). Despite this noted drawback, Cohen's kappa has seen widespread use and popularity in accuracy and similarity assessment (Congalton 1991; Monserud and Leemans 1992; Foody 2002; Foody 2006; Hagen-Zanker 2009).

Spatial scientists have used the extension principal to apply other formerly crisp geostatistics to fuzzy nominal maps (McBratney and Odeh 1997). Davis and Keller combine fuzzy data layers by computing the minimum class membership among a collocated set of pixels (Davis and Keller 1997); this is identical to a fuzzy set intersection operation (Cross and Sudkamp 2002). One noteworthy example is the fuzzy kappa similarity metric, which softens map data using a per-pixel equivalent to the epsilon band model (Hagen 2003). Hagen-Zanker (2009) improved upon this statistic to account for spatial autocorrelation. A recent review compiles further examples (Kuhnert, Voinov, and Seppelt 2005).

In this chapter I addressed methods and literature in area-class map comparison. I discussed softening techniques used in change detection, simulation validation, and landscape pattern analysis. Softening techniques are found to effect local map comparison of (Cross and Sudkamp 2002); for instance, a representation of fuzzy class *belief* may use a different measure of similarity than a representation of class *proximity* or *areal proportion*. I also examined traditional measures of similarity that have been adapted for use in fuzzy map comparison. These statistics are also tightly related to local comparison methods (Kuzera and Pontius 2008). In chapter three, I propose an experiment to assess these concepts of fuzzy similarity for sensitivity to misregistration.

Chapter 3

EXPERIMENT

I designed an experiment to compare five methods of similarity assessment under various conditions of misregistration. This experiment design relates to that used by Sundaresan, Varshney, and Arora (2007) on areal imagery. My experiment formed 328 test cases: 64 misregistrations of a case study area and 12 misregistrations each of 22 simulated maps. Cases separated equally into two test series, *L* and *R*, respectively built using linear and rotational misregistration. Misregistration distances were examined in 0.25 pixel increments up to a total offset of 8 pixels. Additionally, the experiment includes a control dataset with randomly distributed error for each test case in *L* and *R*, denoted by the series *C*(*L*) and *C*(*R*). My experiment examines five comparison methods for area-class maps: fuzzy kappa, epsilon band, cell aggregation, fuzzy inference, and a null method for crisp similarity. Each comparison method yields a similarity measure based on raw accuracy; the fuzzy kappa, epsilon band, and null comparisons further weight this for expected accuracy to derive a form of the kappa statistic. Including parameter variation, this experiment measures 18 similarity statistics for each test case and its control, producing 11,808 statistical results. In this chapter I discuss the case study area and procedures used to misregister it. I go on to describe simulation of random test maps. Finally, I describe the five comparison methods.

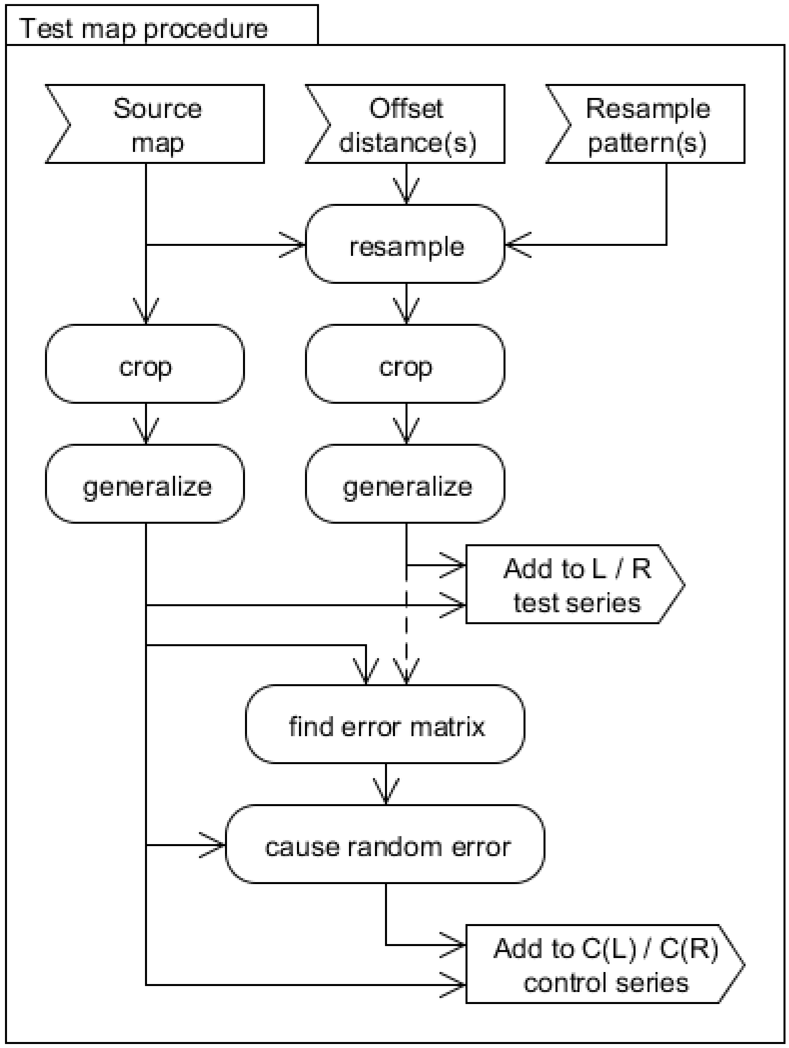

Figure 3. The procedure used to generate test maps for case study and simulated source data.

I. *Case study*

For this case study I chose a map from the US Geological Survey's (USGS) National Land Cover Database (NLCD) published in 2003. The chosen map categorizes 80,000 km^2 of remotely sensed imagery at a 30 meter resolution; it describes parts of Arizona, Utah, Nevada and California with fifteen distinct land cover types defined by the NLCD 2001 land cover class descriptions (Homer et al. 2004). From this area, I selected a square test region 2^{13} pixels across, removing highly localized map features in the process. The resulting square map lacked atypical features, and was deemed representative of those nominal maps most commonly compared using the methods under evaluation. The USGS reports cross-validation accuracy of 0.946 for the case study map. My experiment uses this map to simulate ground truth, so accuracy metadata was discarded before further analysis.

This experiment emulates linear and rotational misregistration with a resampling technique described here. The overall procedure for test map creation is shown in Figure 3. In the first step, misregistration was accomplished by resampling, or assigning class membership for each pixel (x,y) in a new map from a source map pixel located at (x',y'). Linear misregistration is the most common form of registration error used in research (Townshend et al. 1992; Dai and Khorram 1998; Carmel 2004; Sundaresan, Varshney, and Arora 2007). Here I induced linear misregistration by resampling with $x=x'$ and $y=y'+k$, where k is a constant whole number of pixels (i.e. the offset distance times four). Rotational misregistration is a less common alternative to linear shifts in uncertainty studies. The resampling formula for rotational maps defines source pixels by an angle of rotation about map center. For my experiment, I derived an angle of rotation to approximate the root mean squared error (RMSE) in a one-pixel linear offset. A reasonable first guess was the rotation angle required to shift a cell in the middle of any map edge by two pixels. This guess was trimmed experimentally, resulting in an angle of 0.027976456° in the case study map.

It is important to realize that resampling was performed on source maps whose dimensions were each eight times that of test data found in sets *L* and *R*. There are two causes for this change in resolution. First, misregistration leaves undefined pixels near map edges. Following a resample, the test map procedure crops undefined pixels, halving map dimensions. The second cause of resolution change relates to sub-pixel misregistration. To replicate sub-pixel misalignment using the whole-pixel resampling method described above, this procedure generalizes cropped maps. This generalization was modeled on the maximum likelihood classifier; it assigns each square region of 16 cells the most common class found within (see Figure 2). Ties are resolved at random, as suggested in (Saura and Martínez-Millán 2000). The test map procedure reduced case study map dimensions to 2^{10} (1024) pixels on a side, with a pixel resolution of 120 meters.

The test map procedure described in Figure 3 also creates a control map. Such maps are used to compare model performance on two patterns of error: misregistration and random distribution. Control maps are formed in two stages. First, a standard crisp error matrix calculates per-class omission and commission errors caused by resampling. The procedure then reproduces the error matrix in another copy of the generalized source map, locating error with a uniform random distribution. Example control maps for 8-pixel offset members of *R* are shown in Figure 4.

I ran the test creation procedure on the case study map 32 times each for linear and rotational misregistration. Offset distances ranged from 0.25 to 8 pixels

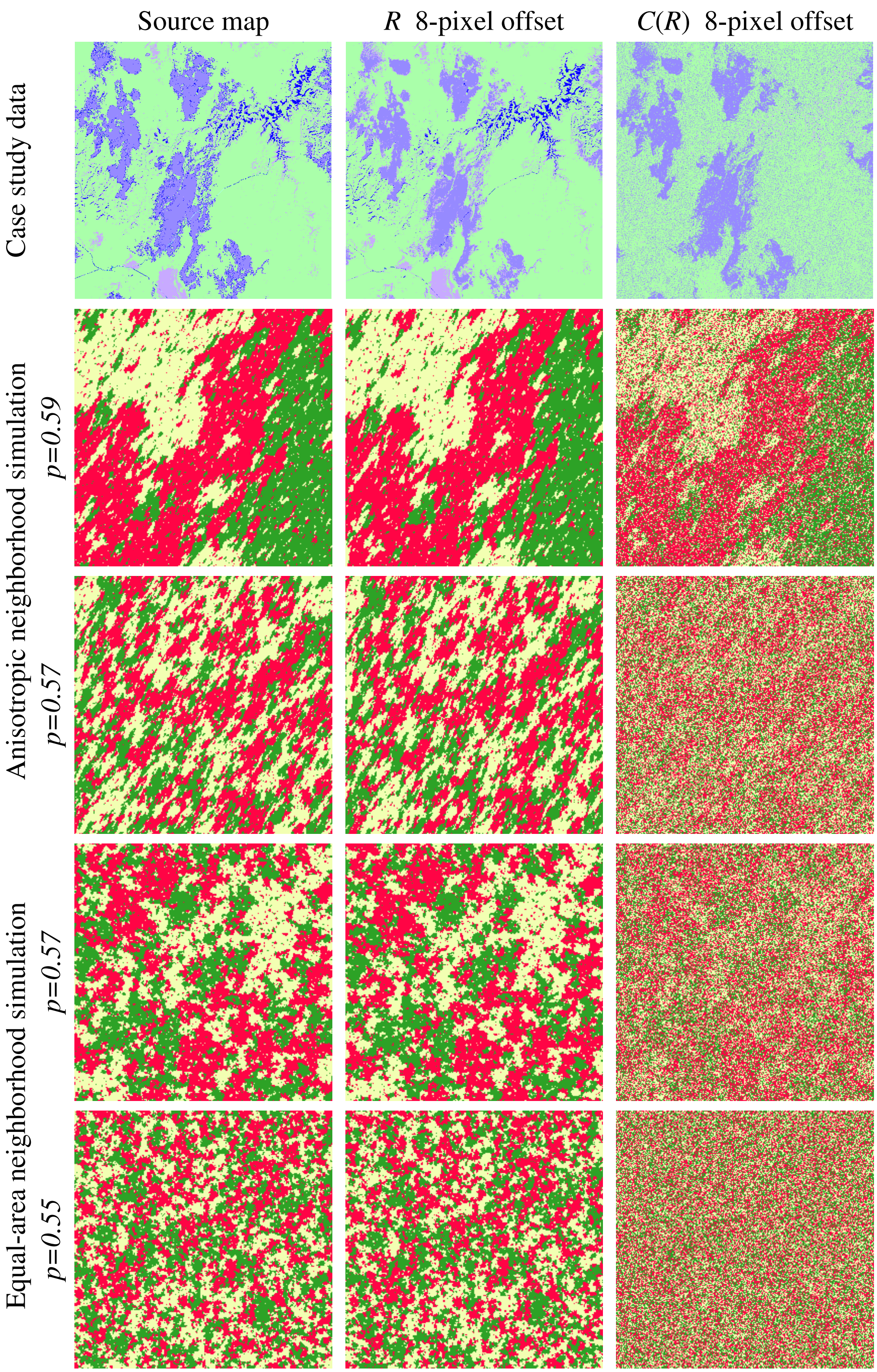

Figure 4. Sample test maps, pixel color indicating crisp class identity.

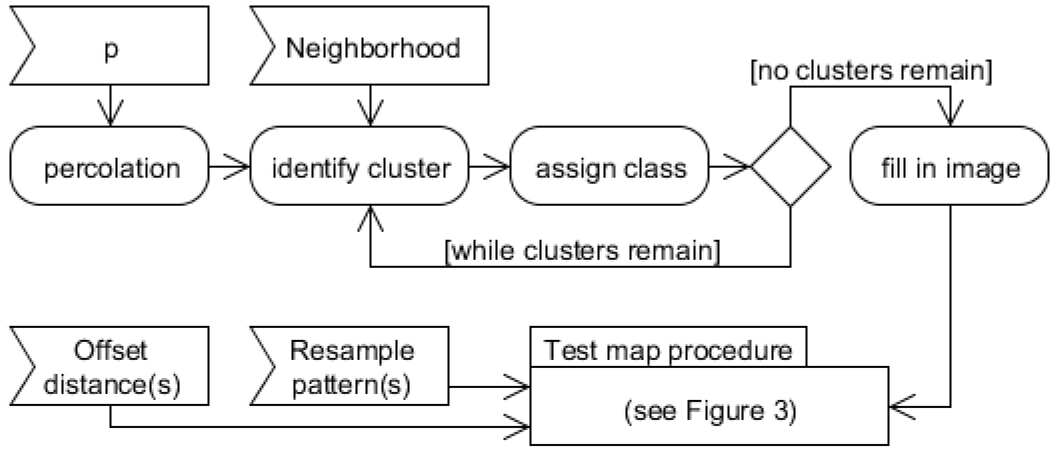

Figure 5. Map simulation procedure

with a 0.25 pixel linear increment. Later analysis refers to linear and rotational misregistration sets from case study data as L_{CS} and R_{CS}, respectively.

II. *Simulated maps*

To generalize beyond my single case study, I require a number of categorical maps with properties representative of area-class phenomena. For this problem, Saura and Martínez-Millán (2000) propose an efficient landscape simulation algorithm. This approach creates artificial maps with the appearance of authentic landscape classes (Hagen-Zanker 2009). The model works in four stages, as illustrated in Figure 5. The first step builds a binary percolation map, flagging each map cell with probability *p*. This probability controls average cluster size of a simulated map. The second simulation step identifies clusters of neighboring flagged cells. Saura and Martínez-Millán (2000) suggest several possible neighborhoods for this stage, including a default equal-area neighborhood and one that induces anisotropy or directional effects in data. In the third step, each cluster is assigned a class with equal probability. Finally, cells not marked in the percolation stage are assigned the class most frequent among their neighboring eight cells.

I generated 22 random maps of three classes: 11 each for isotropic and anisotropic neighborhoods. For these maps, p ranged from 0.50 to 0.60 in 0.01 increments. For each random source map, test maps were created using the same procedure as applied to case study data. However, note that offset distances for simulated source maps are exponential, i.e. {0.25,0.5,1,2,4,8}, reducing 32 test maps for the case study area to 6 for each simulated map. Figure 4 illustrates the effects of both changes to p and the neighborhood parameter following the crop and generalization steps. Dimensions of simulated maps following the crop and generalize operations are 256 pixels square. In chapter 4 I discuss linear and rotational misregistration sets from simulated data, referring to them as L_{Sim} and R_{Sim}, respectively.

III. *Compared measures*

My experiment assesses registration sensitivity in five area-class map comparison methods: fuzzy kappa, epsilon band, cell aggregation, fuzzy inference, and null (a traditional crisp comparison). Recall the three components of map comparison discussed in the previous chapter: softening techniques, local comparison schemes, and measures of global similarity. I chose these five comparison methods to represent the breadth of research in these three component areas, including four different softening techniques for uncertain thematic maps, four methods of local comparison, and fuzzy approximations of raw accuracy and the kappa statistic: the two most popular measures for accuracy assessment

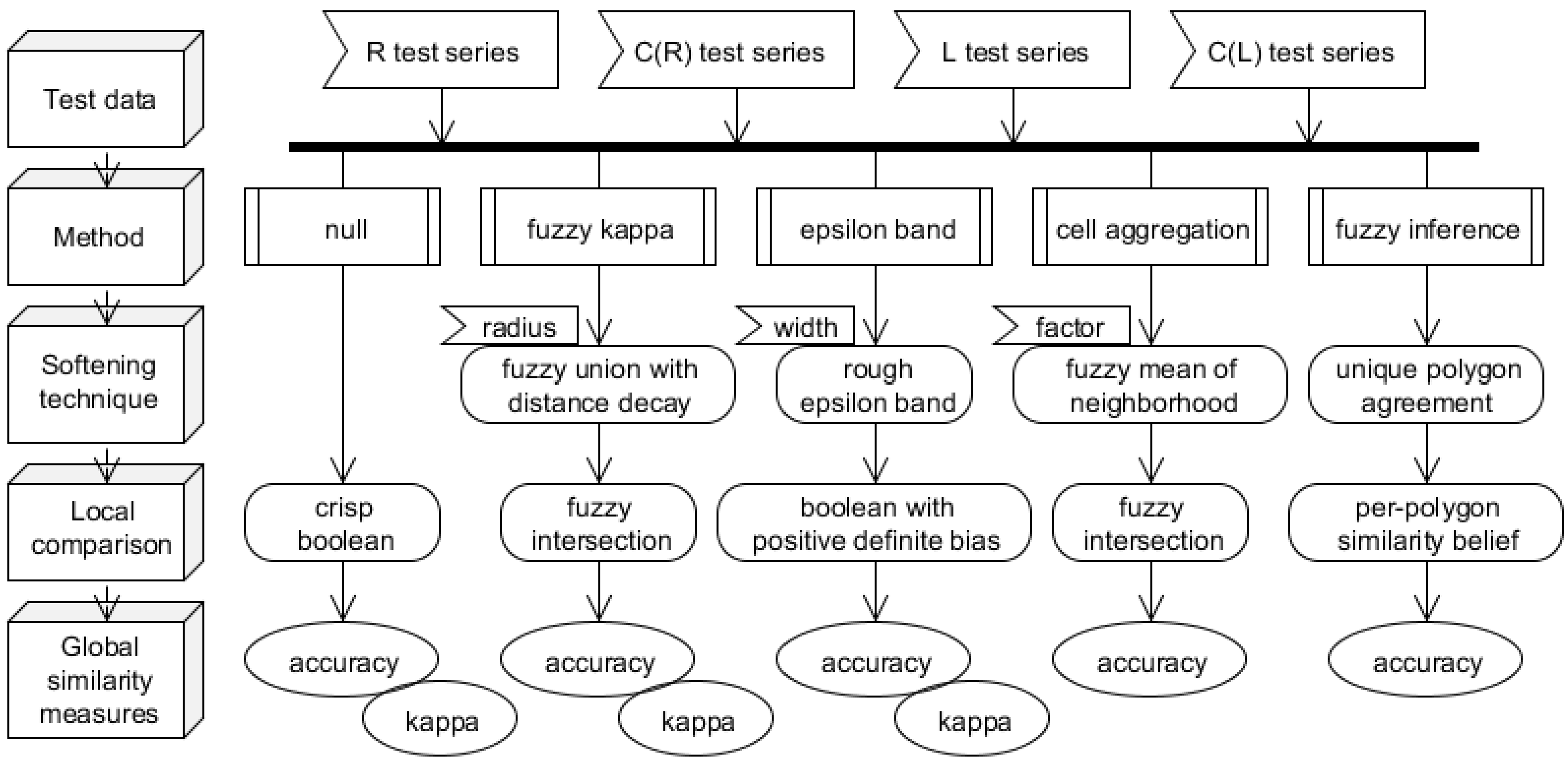

Figure 6. Description of comparison method components

(Foody 2006). Figure 6 describes these comparison method components. Below I describe each comparison method, emphasizing their fuzzy softening technique.

The popular method by Hagen (2003) represents fuzziness of location for his well-known fuzzy kappa statistic. This approach defines degrees of class membership for a cell as the fuzzy union of all cell classes within a given radius, subject to specified distance decay. Assuming crisp class definitions, membership for a class in a given cell is equal to the distance decay function evaluated for the nearest member of that class, including the cell itself. This approach reflects methods common in fuzzy map comparison literature; it views fuzzy membership as *continuous possibility* through location. My experiment tested three models based on linear distance decay with a radius of 2, 4, and 8 pixels, examining raw accuracy and the kappa statistic for each.

The epsilon band model used in my study defines boundaries between two classes as *possibly* belonging to each neighboring class within a constant distance epsilon (Cohn and Gotts 1996). Cells beyond this distance have crisp set membership in a single class. This possibilistic model takes no position on the likelihood or degree of membership inside a boundary. It was included in this study to represent egg-yolk models and rough set theory. My experiment examined three discrete bands of 2, 4, and 8 pixel epsilon distances, and compared them using raw accuracy and the kappa statistic.

This experiment evaluates a cell aggregation method developed by Pontius et. al. (2004) for softening hard classified maps. Their approach reduces the number of pixels in a crisp map, aggregating pixel values using the fuzzy mean operation. The factor parameter determines the length of a square aggregation neighborhood in pixels. Each neighborhood will become a single larger pixel whose fuzzy class is the *areal proportion* of crisp membership within overlapped pixels. Figure 2 illustrates this. Categorical map aggregation methods stem from research in generalization and multiple-resolution map comparison. I examine cell aggregation using factors of 2, 4, and 8 pixels

The fuzzy inference system tested here was developed by Power, Simms, and White (2001) and implemented by Visser and de Nijs (2006). This model distinguishes one map as the *template*, measuring similarity for each template map polygon. For this experiment, the ground truth dataset was always chosen as the template map. Importantly, Power, Simms, and White's (2001) method does not measure per-class fuzzy membership for template polygons. Their method instead measures fuzzy *belief* in two tiers of linguistic concepts. The first tier approximates the crisp proportion of areal agreement in a template polygon using five fuzzy concepts: *very low*, *low*, *medium*, *high*, and *very high*. Disagreement is judged similarly, and polygon size is approximated with fuzzy concepts for *small* and *large*. The second tier applies fuzzy logic to these twelve variables, estimating belief in five terms for overall local similarity, from *very poor* to *perfect*. Finally, these beliefs are "defuzzified" into a single similarity term for each polygon in (0,1). The fuzzy inference system as formulated in 2001 produces one global similarity metric, computed in the same way as raw accuracy. Power, Simms, and White (2001) report this value as comparable with both accuracy and the kappa statistic.

My experiment employs a null model reflecting the traditional measures of similarity in thematic data. In this model, class membership is crisp and unequivocal. Though the approach has no parameters, it supports the now familiar raw accuracy and kappa coefficient statistics. These results provide a benchmark for minimal performance of the other four models.

This experiment is designed to reveal the relationship between misregistration and map similarity, and to qualify changes in this relationship due to misregistration pattern, fuzzy softening technique, offset distance, map cluster size, and anisotropy. Control data and the null model provide a baseline for comparing these effects. In the following chapter, I discuss experimental results.

Chapter 4

RESULTS

Classification error introduced in simulated test data is described in Table 1. As intended, a strong positive correlation is evident between misregistration distance and error introduced in the L_{Sim} and R_{Sim} datasets. The additional error due to misregistration of anisotropic maps was also expected. It is interesting to note that isotropic error in the L_{Sim} series shows consistently lower mean accuracy than anisotropic error in R_{Sim}. It seems that the neighborhood used to simulate source maps will not produce sufficient anisotropy to supersede error caused by the rotational resampling approach. The precise relationship between error induced by linear and rotational misregistration cannot be identified here, but data supports the general claim that rotational misregistration may cause more error in area-class maps than linear offset with equivalent RMSE.

Table 1. Mean and standard deviation for raw accuracy of simulated data by pixel offset distance

Mean of raw accuracy							
		0.25	**0.5**	**1**	**2**	**4**	**8**
L_{Sim}							
	anisotropic	0.914	0.831	0.698	0.563	0.473	0.425
	isotropic	0.926	0.858	0.744	0.613	0.506	0.440
R_{Sim}							
	anisotropic	0.935	0.883	0.790	0.669	0.558	0.477
	isotropic	0.939	0.889	0.800	0.680	0.565	0.481
Standard deviation of raw accuracy							
L_{Sim}							
	anisotropic	0.029	0.058	0.104	0.144	0.153	0.146
	isotropic	0.026	0.049	0.089	0.130	0.151	0.150
R_{Sim}							
	anisotropic	0.022	0.041	0.072	0.112	0.139	0.145
	isotropic	0.022	0.039	0.069	0.109	0.137	0.145

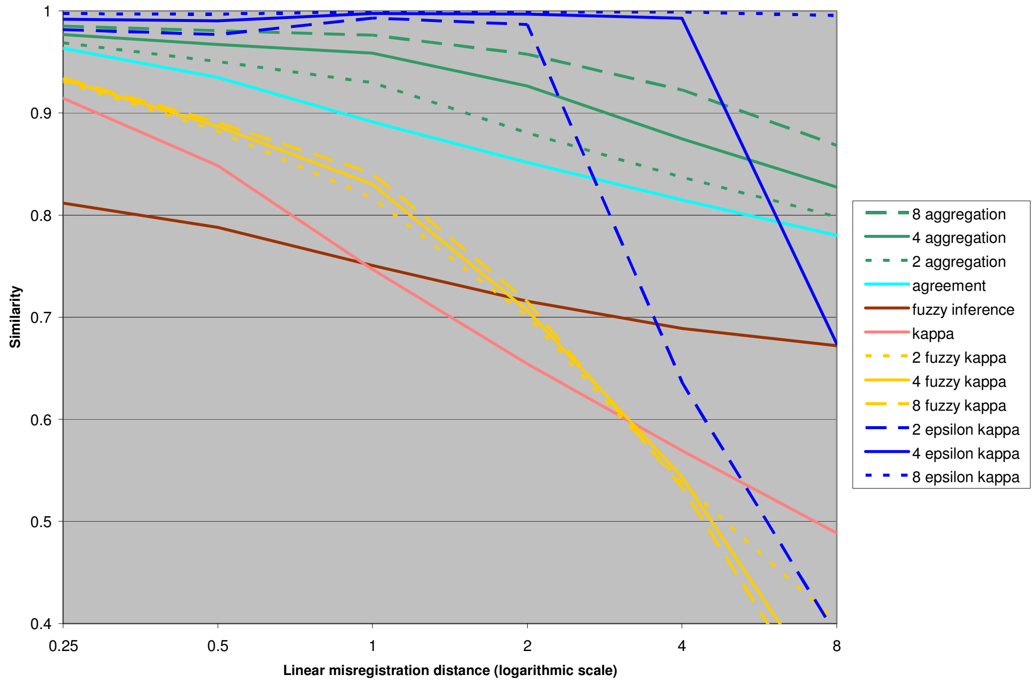

Figure 7. Method similarity results with varying radii by logarithmic misregistration distance, averaged over linear and rotational misregistration.

This study assessed sensitivity of individual metrics to variation in model parameters. Figure 7 shows similarity results from aggregation, fuzzy kappa, and epsilon kappa measures plotted using three radii values. Increase to the fuzzy kappa statistic radius seems to have limited effect on its outcome. The aggregation method shows a slight lag due to changes in radius, but aggregation similarity never drops below crisp agreement, as expected. Transitions in this graph merit investigation. On tests over two pixels in offset distance, fuzzy kappa reports lower similarity estimates than the traditional kappa statistic. Furthermore, fuzzy inference underestimates both kappa results for sub-pixel misregistration. The most unusual of the three expected error measures on display here is the epsilon kappa. This measure shows notable decline once misregistration distance extends beyond its radius. It seems that this decline does not result in lower similarity estimates than those of fuzzy kappa.

Misregistration distance was shown to have a logarithmic relationship with accuracy-based measures for all metrics except discrete epsilon bands. The effect is illustrated in Figure 8. This graph shows results only from case study data, plotting similarity method results for each quarter pixel offset. Neighborhood-based softening methods were expected to behave similarly under increasing misregistration distance. The cell aggregation and fuzzy kappa methods show this correlation. It was surprising that the epsilon band metric did not share this logarithmic relation at low pixel offset distances. Such a relationship seems to exist for offsets above 2 pixels; this is an artifact of averaging results for the epsilon band widths described in Figure 7.

To understand the sensitivity of comparison methods to linear and rotational error, I compared average accuracy statistics for test and control data. Figure 8 shows average results for $C(L_{CS})$ and $C(R_{CS})$ as dashed lines. Green lines depict the null model. As expected, test and control accuracy for the null model overlap, while fuzzy models report greater similarity in test data. Strangely, fuzzy inference reports greater similarity on control data than it does on misregistered test maps. This behavior is repeated in L_{Sim} and R_{Sim}. A summary of test and control results for simulated data is illustrated in Figure 9. Misregistration pattern and simulated anisotropy do not greatly affect differences between test and control results.

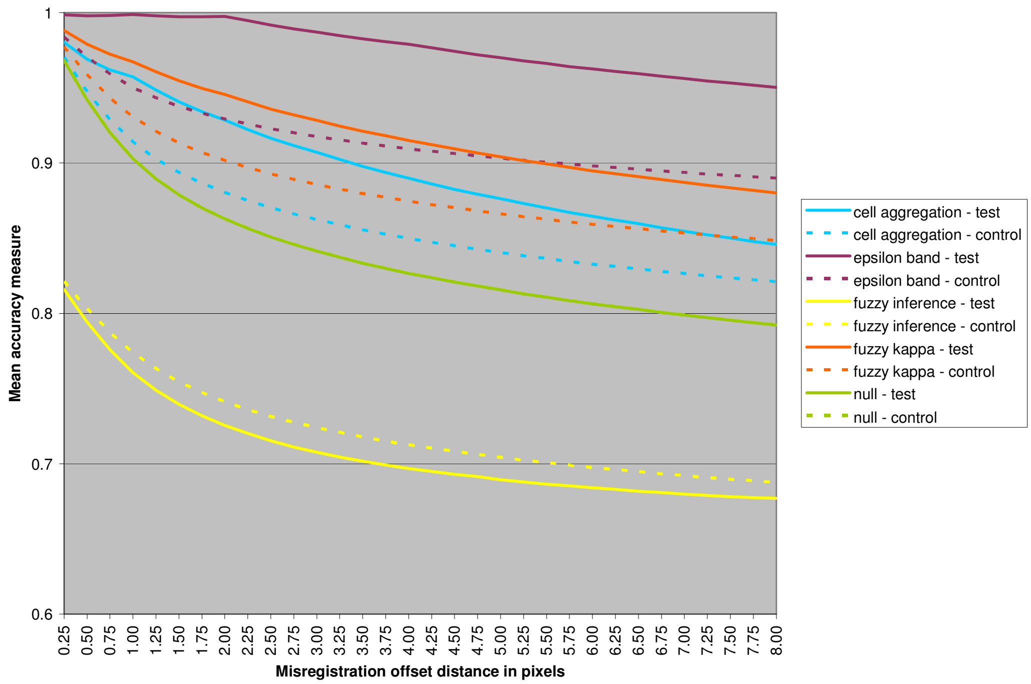

Figure 8. Method similarity results by equivalent misregistration distance, averaged over linear and rotational misregistration and method parameters

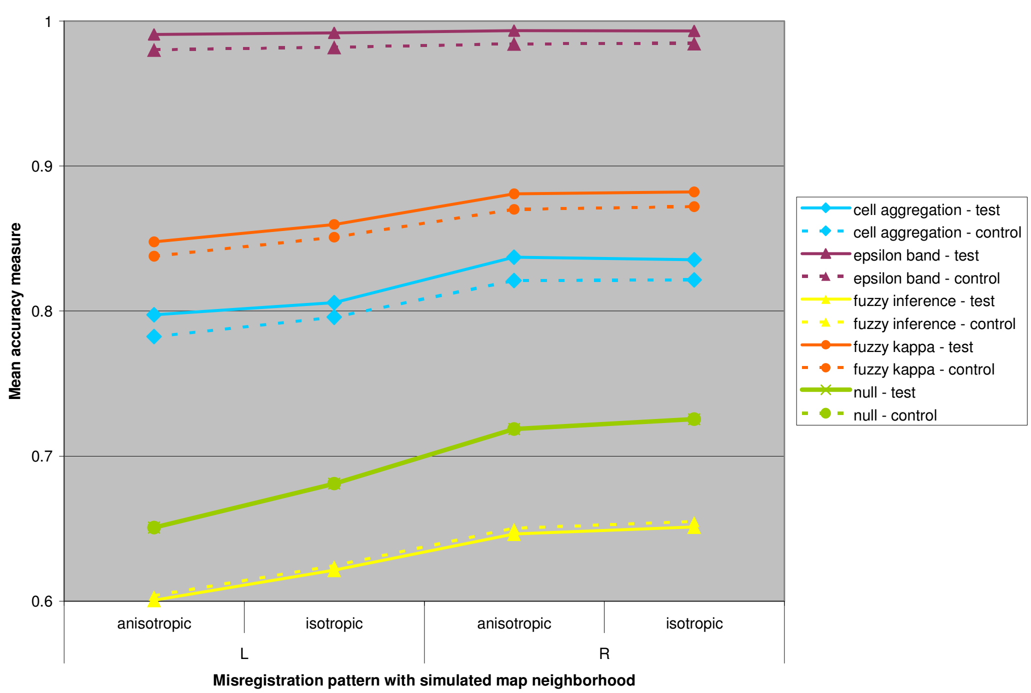

Figure 9. Method similarity results by rotational and linear misregistration, averaged over misregistration distance and method parameters

Table 2. Mean and standard deviation for accuracy statistics of simulated data by pixel offset distance

	0.25	0.5	1	2	4	8
Mean of accuracy						
cell aggregation	0.953	0.928	0.895	0.818	0.704	0.616
epsilon band	0.999	0.999	0.999	0.999	0.996	0.961
fuzzy inference	0.778	0.737	0.681	0.594	0.520	0.469
fuzzy kappa	0.976	0.956	0.925	0.856	0.777	0.715
null	0.928	0.865	0.758	0.631	0.525	0.456
Standard deviation of accuracy						
cell aggregation	0.014	0.020	0.027	0.049	0.076	0.085
epsilon band	0.001	0.001	0.001	0.001	0.004	0.032
fuzzy inference	0.024	0.033	0.048	0.083	0.097	0.098
fuzzy kappa	0.008	0.015	0.026	0.049	0.060	0.056
null	0.026	0.051	0.091	0.129	0.145	0.144

The distribution of raw similarity estimates for each of the five models is of interest in relating model performance on misregistration error. Mean and standard deviation of accuracy statistics from each similarity measure served as preliminary indication of sensitivity. Results of this investigation are listed in Table 2. Increasing offset distance shows positive correlation with standard deviation. For further analysis, Table 2 was expanded to describe changes per model. Results were plotted in Figure 10 to visually distinguish model performance under progressive error. Models are listed from left to right by increasing variance in reported similarity. Three significant observations come from this data. The epsilon method shows the lowest variance, reporting near-total similarity on offsets of two pixels or less. Despite offering the lowest absolute measures of similarity, the fuzzy inference metric is less sensitive to increasing misregistration than either fuzzy kappa or cell aggregation. Finally, fuzzy kappa and cell aggregation metrics show nearly identical variation due to misregistration distance.

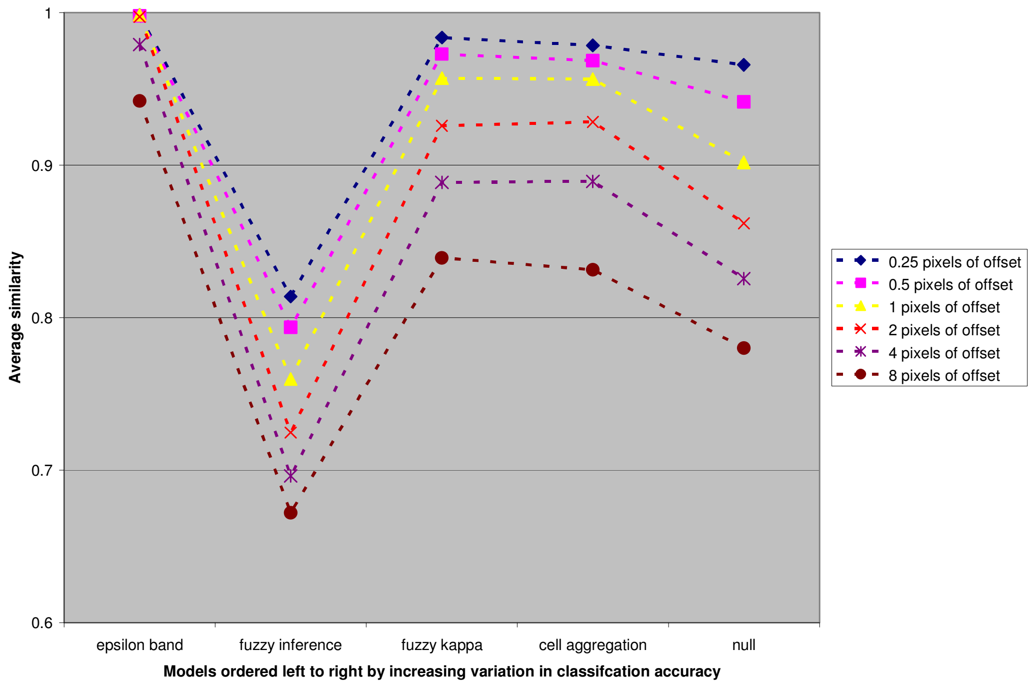

Figure 10. Misregistration distance similarity results by similarity measure, averaged over linear and rotational misregistration and method parameters

In this chapter I analyzed numerical results from the experiment described in Figures 4-6. This experiment produced valid test data with expected error properties. I have highlighted unique characteristics of accuracy and kappa results obtained from the five map comparison methods examined, including strong performance of the epsilon band method and questionable behavior of fuzzy inference. In the following chapter I discuss specific findings of my research.

Chapter 5

DISCUSSION

My study has demonstrated that rotational misregistration may lead to increased variance in similarity assessments when compared with linear misregistration of equivalent RMS error. As results were controlled for relative number of classification errors, this variance relates to the distribution of error by misregistration. While this distribution is known, its representation within uncertainty models is of principal concern. By sampling from autocorrelated error, per-pixel comparisons of misregistered maps also show spatially dependent error (Foody 2002). In light of this, per-pixel comparison maps for models of 0.25 and 4 pixel offsets are shown in Figures 11 and 12, respectively. Model variation due to error distribution are noticeable for most measures; these changes propagate through similarity statistics (Power, Simms, and White 2001), resulting in increased variation observed in Table 2. Uncertainty models of misregistration have generally favored linear offsets for experimental assessment. It is possible that such studies underestimate measure sensitivity due to registration error.

The epsilon band model faired remarkably well during analysis. It reported the highest raw similarity on all test datasets, and the lowest raw similarity variance of all five models. Its kappa statistic estimated near-total similarity for misregistration distances up to the radius parameter, with kappa scores consistently higher than fuzzy kappa beyond this distance. To explain this behavior, recall that aggregation represents class membership as a normalized proportion of area, and that fuzzy kappa membership describes nearest neighbor distance. While these philosophies credit partial similarity in misregistered classes, rough set theory ascribes total similarity to any class appearing within an epsilon band boundary for that class. Figure 11 shows epsilon model behavior on linear and rotational misregistration. The egg-yolk approach has intriguing properties for similarity comparison of misregistered maps.

While I expected a decrease in all neighborhood similarity estimates once misregistration distance exceeded method radius, this was not seen in aggregation or fuzzy kappa metrics. These measures show smooth transitions for parameter analysis in Figure 7. Furthermore, model results in Figure 10 show high variance due to misregistration, close to that measured in the null model. Aggregation and fuzzy kappa representations in Figures 6 and 7 agree with this assessment, showing close resemblance to the null model. Prior analysis also noted that these models produce nearly identical similarity estimates. Although these models showed few compelling traits for misregistration, demonstrating similarities between aggregation and fuzzy kappa models is an interesting result of this study.

Analysis revealed uncommon behavior in the fuzzy inference metric. As mentioned during a review of literature, any metric for classification accuracy should produce values between those of raw accuracy and the kappa statistic (Power, Simms, and White 2001). While Power, Simms, and White (2001) report fuzzy inference scores falling within appropriate bounds, this did not hold for sub-pixel misregistration. Figure 7 depicts fuzzy inference values dropping below fuzzy- and crisp kappa estimates on test maps offset by one pixel or less. The unusual behavior of fuzzy inference comparison can be partly explained by its

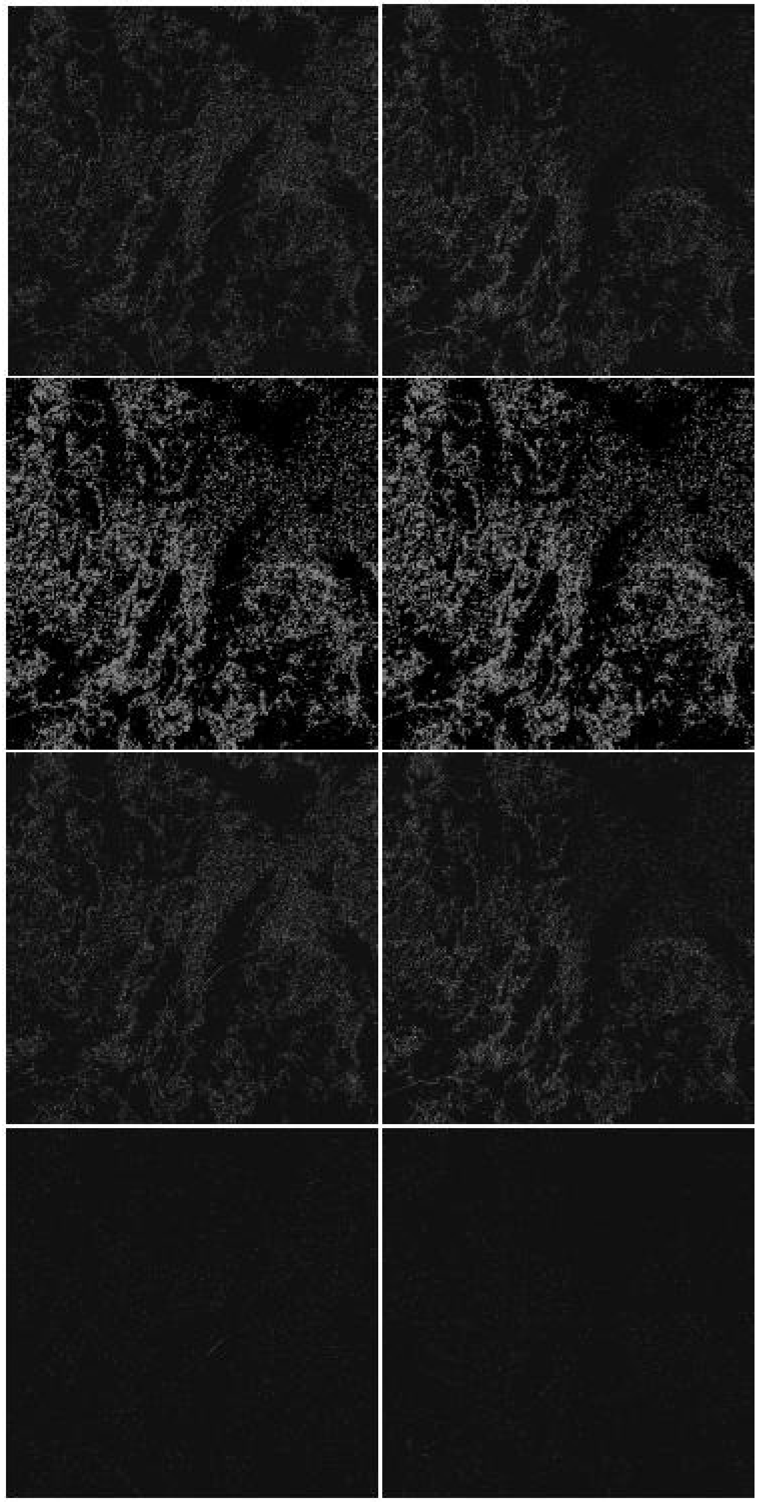

Figure 11: Difference intensity maps of the 0.25-pixel misregistration of *L* and *R* datasets (from left). From top: null model, aggregation model, fuzzy kappa model, and egg-yolk model. Brighter intensity indicates a higher degree of difference

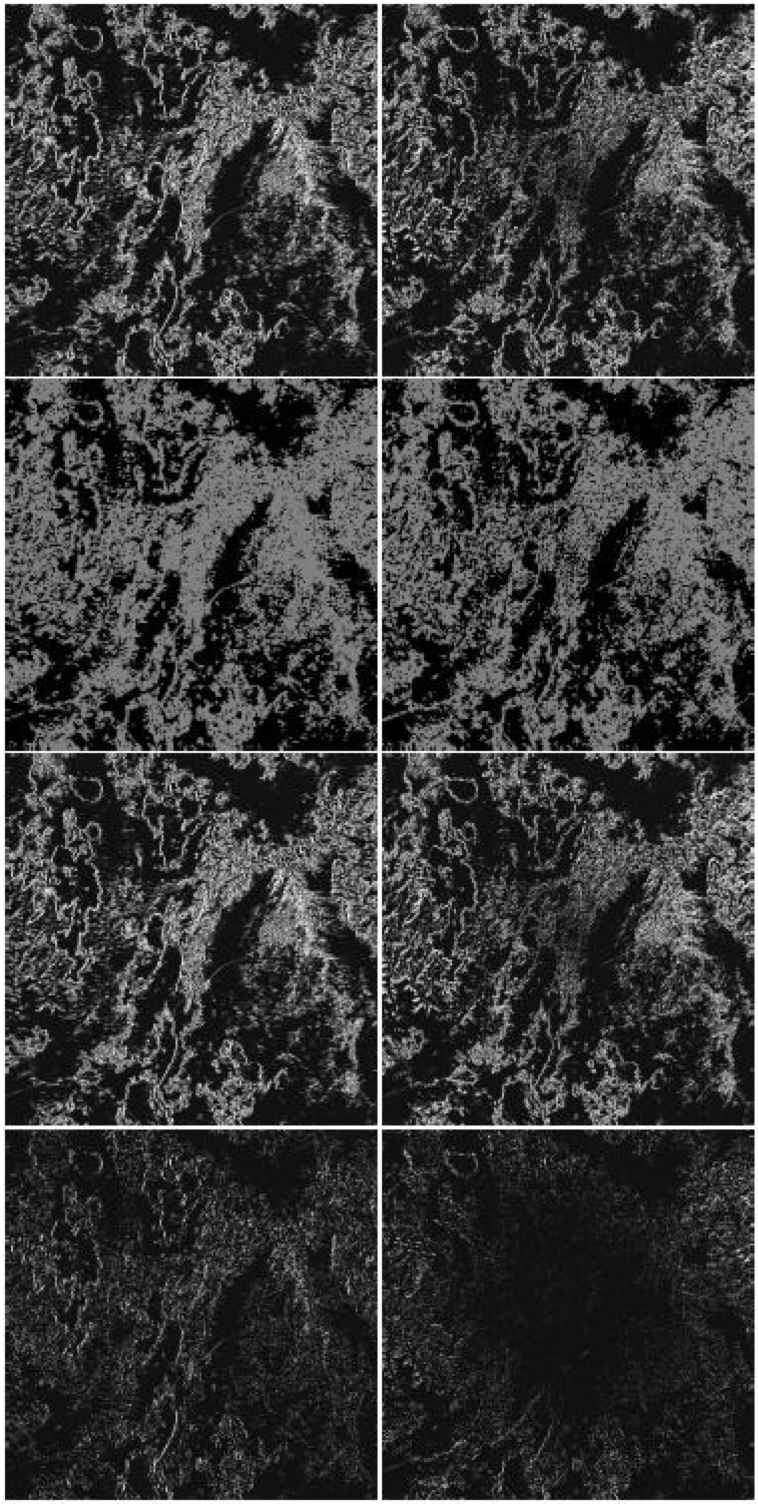

Figure 12: Difference intensity maps of the 4-pixel misregistration of *L* and *R* datasets (from left). From top: null model, aggregation model, fuzzy kappa model, and egg-yolk model. Brighter intensity indicates a higher degree of difference

unique philosophic basis. As discussed, the fuzzy inference model quantifies belief in map similarity, rather than fraction of agreement or a kappa equivalent. This may prohibit unqualified comparison between fuzzy inference and other models.

My analysis is limited by my choice of control and test data. Simulated and case study data were chosen to represent common area-class maps encountered during map comparison. However, I did no conduct a detailed analysis of patch size, edge length, or autocorrelation in mapped classes. Furthermore, not all interesting map changes are approximated by a uniform random distribution.

In this chapter I discussed the principal findings of my research. I discussed why the epsilon band model may show uncommon effectiveness in modeling misregistration uncertainty. Neither fuzzy kappa nor cell aggregation faired as well in this experiment, despite being neighborhood-oriented comparison methods. I also questioned behavior observed in the fuzzy inference measure. In the following chapter I discuss the significance of my work.

Chapter 6

CONCLUSION

The results of my study demonstrate that a discrete epsilon band method shows remarkably low sensitivity to small registration error. By using a possibilistic view of class membership, local movement of class boundaries are effectively ignored. This allows the discrete epsilon band to achieve high accuracy assessments on misregistered test data with low variability due to increasing misregistration. These traits encourage further research on epsilon bands in area-class map similarity assessment, and reassert the utility of Łukasiewicz' trivalent logic.

In contrast, my experiment demonstrated poor performance of the fuzzy inference approach for modeling misregistration. As discussed in the prior section, this method gave invalid similarity results when comparing maps with sub-pixel registration errors. Furthermore, fuzzy inference consistently reported greater similarity on control data than in the test dataset. It is therefore inappropriate to use fuzzy inference to model misregistration error.

This research addresses a significant outstanding problem in area-class map comparison. Although numerous comparison methods exist, many have unknown performance under common error conditions. Investigations of land use change often encounter data with questionable registration, such as historical land use maps. Correcting error in such cases is often impossible. My research suggests that a discrete epsilon band may address this best of the representative sample of models tested. This knowledge can improve the accuracy of land change analysts or other experts working with uncertain area-class maps.

Propagation of uncertainty is an ongoing issue in similarity assessment. Along with improved approaches for map registration, future research should pursue comparison methods resistant to registration error. Sensitivity assessment of more comparison methods such as those used in pattern analysis will contribute to this agenda. Also significant is the representation of underlying map data; an experiment adapting raster-based comparison methods to vector maps may provide different results than this approach, which evaluated vector-based comparison methods with raster data. Finally, many applications of map comparison seek to identify nonsystematic map change with patterns other than a uniform random distribution. To further address this, future assessments can misregister maps containing such changes.

REFERENCES

Bennett, B. 2001. What is a forest? on the vagueness of certain geographic concepts. *Topoi* 20 (2):189-201.

Bolstad, P. V., P. Gessler, and T. M. Lillesand. 1990. Positional uncertainty in manually digitized map data. *International Journal of Geographical Information Science* 4 (4):399 - 412.

Boots, B., and F. Csillag. 2006. Categorical maps, comparisons, and confidence. *Journal of Geographical Systems* 8 (2):109-118.

Bregt, A. K., J. Denneboom, H. J. Gesink, and Y. Van Randen. 1991. Determination of rasterizing error a case study with the soil map of The Netherlands. *International Journal of Geographical Information Science* 5 (3):361 - 367.

Bruzzone, L., and R. Cossu. 2003. An adaptive approach to reducing registration noise effects in unsupervised change detection. *Geoscience and Remote Sensing, IEEE Transactions on* 41 (11):2455-2465.

Bruzzone, L., and D. F. Prieto. 2000. Automatic analysis of the difference image for unsupervised change detection. *Geoscience and Remote Sensing, IEEE Transactions on* 38 (3):1171-1182.

Carmel, Y. 2004. Controlling data uncertainty via aggregation in remotely sensed data. *Geoscience and Remote Sensing Letters, IEEE* 1 (2):39-41.

Cheng, R. 2009. Querying and cleaning uncertain data. In *Quality of Context*, 41-52. Berlin / Heidelberg: Springer.

Chrisman, N. R. 1989. Modeling error in overlaid categorical maps. In *Accuracy of Spatial Databases*, eds. M. F. Goodchild and S. Gopal, 21-34. London: Taylor and Francis.

Cohen, J. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 20 (1):37-46.

Cohn, A. G., and N. M. Gotts. 1996. The 'egg-yolk' representation of regions with indeterminate boundaries. In *Geographic Objects with Indeterminate Boundaries*, eds. P. A. Burrough and A. Frank, 171-187: Taylor & Francis.

Congalton, R. G. 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment* 37 (1):35-46.

Congalton, R. G., and K. Green. 2009. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*: CRC Press.

Costanza, R. 1989. Model goodness of fit: A multiple resolution procedure. *Ecological Modeling* 47 (3-4):199-215.

Crosetto, M., and S. Tarantola. 2001. Uncertainty and sensitivity analysis: tools for GIS-based model implementation. *International Journal of Geographical Information Science* 15 (5):415 - 437.

Cross, V. V., and T. A. Sudkamp. 2002. *Similarity and Compatibility in Fuzzy Set Theory*. Heidelberg; New York: Physica-Verlag.

Dai, X., and S. Khorram. 1998. The effects of image misregistration on the accuracy of remotely sensed change detection. *IEEE Transactions on Geoscience and Remote Sensing* 36 (5):1566-1577.

Davis, T. J., and C. P. Keller. 1997. Modeling uncertainty in natural resource analysis using fuzzy sets and Monte Carlo simulation: slope stability prediction. *International Journal of Geographical Information Science* 11 (5):409 - 434.

Ehlschlaeger, C. R., and M. F. Goodchild. 1994. Dealing with uncertainty In categorical coverage maps: defining, visualizing, and managing errors. In *Proceedings of the Workshop on Geographical Information Systems at the Conference on Information and Knowledge Management*, 86-91. Gaithersburg, Maryland.

Elkan, C., H. R. Berenji, B. Chandrasekaran, C. J. S. de Silva, Y. Attikiouzel, D. Dubois, H. Prade, P. Smets, C. Freksa, O. N. Garcia, G. J. Klir, Y. Bo, E. H. Mamdani, F. J. Pelletier, E. H. Ruspini, B. Turksen, N. Vadiee, M. Jamshidi, W. Pei-Zhuang, T. Sie-Keng, T. Shaohua, R. R. Yager, and L. A. Zadeh. 1994. The paradoxical success of fuzzy logic. *IEEE Expert* 9 (4):3-49.

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

Foody, G. 2006. What is the difference between two maps? a remote senser's view. *Journal of Geographical Systems* 8 (2):119-130.

Foody, G. M. 2002. Status of land cover classification accuracy assessment. *Remote Sensing of Environment* 80 (1):185-201.

Goodchild, M. F. 2003. Models for uncertainty in area-class maps. In *Proceedings of the Second International Symposium on Spatial Data Quality*, eds. W.

Shi, M. F. Goodchild and P. F. Fisher, 1-9. Hong Kong: Hong Kong Polytechnic University.

———. 2008. Imprecision and spatial uncertainty. In *Encyclopedia of GIS*, eds. S. Shekhar and H. Xiong. New York: Springer.

Goodchild, M. F., and O. Dubuc. 1987. A model of error for choropleth maps with applications to geographic information systems. In *Auto Carto 8*, 165-174. Falls Church, VA.

Goodchild, M. F., S. Guoqing, and Y. Shiren. 1992. Development and test of an error model for categorical data. *International Journal of Geographical Information Science* 6 (2):87-104.

Grunwald, S. 2009. Multi-criteria characterization of recent digital soil mapping and modeling approaches. *Geoderma* 152 (3-4):195-207.

Gudder, S. 2000. What Is fuzzy probability theory? *Foundations of Physics* 30 (10):1663-1678.

Gustafson, E. J. 1998. Quantifying landscape spatial pattern: what is the state of the art? *Ecosystems* 1 (2):143-156.

Hagen-Zanker, A. 2009. An improved Fuzzy Kappa statistic that accounts for spatial autocorrelation. *International Journal of Geographical Information Science* 23 (1):61 - 73.

Hagen, A. 2002. Multi-method assessment of map similarity. In *Proceedings of the 5th AGILE Conference on Geographic Information Science*, 171-182. Palma (Mallorca, Spain).

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

Homer, C., C. Huang, L. Yang, B. Wylie, and M. Coan. 2004. Development of a 2001 national land cover database for the United States. *Photogrammetric Engineering and Remote Sensing* 70 (7):829-840.

Jang, J. S. R. 1993. ANFIS: adaptive-network-based fuzzy inference system. *Systems, Man and Cybernetics, IEEE Transactions on* 23 (3):665-685.

Karimi, H. A., M. Durcik, and W. Rasdorf. 2004. Evaluation of uncertainties associated with geocoding techniques. *Computer-Aided Civil and Infrastructure Engineering* 19 (3):170-185.

Kay, I. W. 1994. A survey of fuzzy logic and applications, 55. Alexandria, VA: Institute for Defense Analyses.

Kuhnert, M., A. Voinov, and R. Seppelt. 2005. Comparing raster map comparison algorithms for spatial modeling and analysis. *Photogrammetric Engineering and Remote Sensing* 71 (8):975–984.

Kuzera, K., and R. Pontius. 2008. Importance of matrix construction for multiple-resolution categorical map comparison. *GIScience & Remote Sensing* 45 (3):249-274.

Leyk, S., R. Boesch, and R. Weibel. 2005. A conceptual framework for uncertainty investigation in map-based land cover change modelling. *Transactions in GIS* 9 (3):291-322.

Łukasiewicz, J. 1970. On three-valued logic. In *Jan Łukasiewicz: Selected Works*, ed. L. Borkowski, 87-88. North Holland, Amsterdam: North-Holland Publishing Company.

Mark, D. M., and F. Csillag. 1989. The nature of boundaries on 'area-class' maps. *Cartographica* 26 (1):65-78.

Mas, J. F. 2005. Change estimates by map comparison: a method to reduce erroneous changes due to positional error. *Transactions in GIS* 9 (4):619-629.

McBratney, A. B., and I. O. A. Odeh. 1997. Application of fuzzy sets in soil science: fuzzy logic, fuzzy measurements and fuzzy decisions. *Geoderma* 77 (2-4):85-113.

McGarigal, K., and B. J. Marks. 1995. FRAGSTATS: spatial pattern analysis program for quantifying landscape structure. In *General Technical Report PNW-GTR-351*, 122. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station.

Monserud, R. A., and R. Leemans. 1992. Comparing global vegetation maps with the Kappa statistic. *Ecological Modeling* 62 (4):275-293.

Nilsson, N. J. 1986. Probabilistic logic. *Artificial Intelligence* 28 (1):71-87.

Parker, D. C., S. M. Manson, M. A. Janssen, M. J. Hoffmann, and P. Deadman. 2003. Multi-agent systems for the simulation of land-use and land-cover change: a review. *Annals of the Association of American Geographers* 93 (2):314-337.

Perkal, J. 1986. On the length of empirical curves. In *Michigan Interuniversity Community of Mathematical Geographers. Papers. Reprint.*, ed. J. D.

Nystuen, Trans. R. Jackowski, 320-354. Ann Arbor, MI: Institute of Mathematical Geography. First published in *Zastosowania Matematyki* III, 3-4 (1958): 257–286.

Pontius, R. G. 2002. Statistical methods to partition effects of quantity and location during comparison of categorical maps at multiple resolutions. *Photogrammetric Engineering and Remote Sensing* 68 (10):1041-1050.

Pontius, R. G., D. Huffaker, and K. Denman. 2004. Useful techniques of validation for spatially explicit land-change models. *Ecological Modeling* 179 (4):445-461.

Pontius, R. G., E. Shusas, and M. McEachern. 2004. Detecting important categorical land changes while accounting for persistence. *Agriculture, Ecosystems & Environment* 101 (2-3):251-268.

Power, C., A. Simms, and R. White. 2001. Hierarchical fuzzy pattern matching for the regional comparison of land use maps. *International Journal of Geographical Information Science* 15:77-100.

Radke, R. J., S. Andra, O. Al-kofahi, and B. Roysam. 2005. Image change detection algorithms: a systematic survey. *IEEE Transactions on Image Processing* 14 (3):294-307.

Rees, W. G. 2008. Comparing the spatial content of thematic maps. *International Journal of Remote Sensing* 29 (13):3833 - 3844.

Ricotta, C. 2005. On possible measures for evaluating the degree of uncertainty of fuzzy thematic maps. *International Journal of Remote Sensing* 26 (24):5573 - 5583.

Rossiter, D. G. 2001. Assessing the thematic accuracy of area-class soil maps, 46. Enschede, the Netherlands: Social Science Division, ITC.

Saura, S., and J. Martínez-Millán. 2000. Landscape patterns simulation with a modified random clusters method. *Landscape Ecology* 15:661–678.

Schneider, M. 2008. Fuzzy spatial data types for spatial uncertainty management in databases. In *Handbook of Research on Fuzzy Information Processing in Databases*, ed. J. Galindo, 490-515.

Shafer, G. 1976. *A Mathematical Theory of Evidence*. Princeton, N.J.: Princeton University Press.

Stehman, S. V. 1999. Comparing thematic maps based on map value. *International Journal of Remote Sensing* 20 (12):2347-2366.

Stow, D. A., and D. M. Chen. 2002. Sensitivity of multitemporal NOAA AVHRR data of an urbanizing region to land-use/land-cover changes and misregistration. *Remote Sensing of Environment* 80 (2):297-307.

Sundaresan, A., P. K. Varshney, and M. K. Arora. 2007. Robustness of change detection algorithms in the presence of registration errors. *Photogrammetric Engineering and Remote Sensing* 73 (4):375–383.

Townshend, J. R. G., C. O. Justice, C. Gurney, and J. McManus. 1992. The impact of misregistration on change detection. *IEEE Transactions on Geoscience and Remote Sensing* 30 (5):1054-1060.

Visser, H., and T. de Nijs. 2006. The Map Comparison Kit. *Environmental Modelling & Software* 21 (3):346-358.

White, R. 2006. Pattern based map comparisons. *Journal of Geographical Systems* 8 (2):145-164.

Williams, E. A., and E. A. Wentz. 2008. Pattern analysis based on type, orientation, size, and shape. *Geographical Analysis* 40 (2):97-122.

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

Worboys, M. F., and E. Clementini. 2001. Integration of imperfect spatial information. *Journal of Visual Languages & Computing* 12 (1):61-80.

Zadeh, L. A. 1965. Fuzzy sets. *Information and Control* 8 (3):338-353.

———. 1968. Probability measures of fuzzy events. *Journal of Mathematical Analysis and Applications* 23 (2):421-427.

———. 1975. Fuzzy logic and approximate reasoning. *Synthese* 30 (3):407-428.

———. 1978. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems* 1 (1):3-28.

———. 1995. Discussion: Probability theory and fuzzy logic are complementary rather than competitive. *Technometrics* 37 (3):271-276.

Zhang, J., and M. F. Goodchild. 2002. *Uncertainty in Geographical Information*. New York: Taylor and Francis.

Zhu, A. X. 1997. A similarity model for representing soil spatial information. *Geoderma* 77 (2-4):217-242.