

## Research Article

# Accuracy of discontinuous binary surfaces: a case study using boreal forest fires

TARMO K. REMMEL

Ontario Forest Research Institute, Ontario Ministry of Natural Resources, 1235 Queen St. East, Sault Ste. Marie, Ontario, P6A 2E5 Canada.  
e-mail: tarmoremmel@hotmail.com

and AJITH H. PERERA

Ontario Forest Research Institute, Ontario Ministry of Natural Resources, 1235 Queen St. East, Sault Ste. Marie, Ontario, P6A 2E5 Canada.  
e-mail: ajith.perera@mnr.gov.on.ca

*(Received 12 July 2000; accepted 7 May 2001)*

**Abstract.** Confidence in the conclusions of GIS and remote sensing analyses depends on our ability to specify their accuracy. The square error matrix, which is commonly used for accuracy assessment when the database contains continuous digital choropleth data, allows computation of user's, producer's, overall, and Kappa coefficient accuracy values. However, when the underlying assumptions of the square error matrix method are violated, these accuracy values may become unreliable. We demonstrate how traditional forms of accuracy assessment can fail for discontinuous binary surfaces and provide an improved spatial assessment method. We developed 15 new accuracy measures that allow analysts to assess scene- and feature-level accuracy from many new perspectives.

## 1. Introduction

Assessing the accuracy of continuous categorical maps, generated by classifying remotely sensed imagery, is common. Several well-documented accuracy assessment procedures using point-in-polygon techniques (Blakemore 1984), polygon overlays (Chrisman and Lester 1991), and error matrices (Congalton 1991) have been developed. Error matrices are widely used and simple to construct (Rosenfield 1986, Congalton 1991) and allow the computation of Kappa (Cohen 1960, Rosenfield and Fitzpatrick-Lins 1986), a measure of actual agreement minus the agreement expected by chance (Næsset 1996). However, these procedures can become unreliable when their underlying assumptions are violated (Congalton 1991).

The error matrix assumes that map categories exhibit spatial continuity and evenness thus allowing simple random sample points realistically to represent the surface. With binary representations of data, this continuity and evenness of categories can be compromised. Further, if the category of interest only represents a small proportion of the entire image, polygons (or cells) representing the category of

interest may be considered discrete features on a vast background. We refer to this categorical condition as a discontinuous binary surface. Sampling methods specific to the error matrix may yield biased results for discontinuous binary categorical maps (Congalton 1988).

The Kappa coefficient is a preferred measure of accuracy compared with the traditional overall accuracy value from the error matrix (Rosenfield and Fitzpatrick-Lins 1985; Fitzgerald and Lees 1994). The estimator of the Kappa coefficient, KHAT, has been widely used to test significant differences between and among various error matrices (Congalton and Mead 1983, Congalton *et al.* 1983). However, if simple random sampling is not conducted, the confidence of the Kappa coefficient may also be biased for discontinuous binary surfaces (Stehman 1992).

To reduce attribute errors due to positional inconsistencies between maps, all data must be properly co-registered and checked by comparing distances between point locations (Barrette *et al.* 2000) or by measuring the root-mean-square (RMS) error (Jensen 1996, Fernández *et al.* 1997). These assessments are especially important when datasets are combined for analysis, which can result in multiplication and propagation of errors leading to improper conclusions (Heuvelink *et al.* 1989, Congalton 1991, Lunetta *et al.* 1991).

This paper provides a new approach for computing accuracy for discontinuous binary maps, using fire mapping in a northern boreal forest as a case study. Annual fires in northern Ontario (discrete features) represent  $\approx 1\%$  of the 43 million ha study area (background), providing an ideal discontinuous binary surface for analysis. The methods we describe allow several variations of spatial accuracy to be quantified, explained and presented both numerically and graphically. In general, this paper demonstrates the superiority of a discrete-feature accuracy assessment over the traditional error matrix and Kappa approaches for discontinuous binary surfaces.

## 2. Methods

### 2.1. Fire mapping

Decreases in the Normalized Difference Vegetation Index (NDVI), calculated from the Advanced Very High Resolution Radiometer (AVHRR), have been used to map vegetation senescence and forest fire disturbances in various biomes (Razafimpanilo *et al.* 1995). By subtracting pixel values between two spatially complementary images and applying a threshold (Lunetta and Elvidge 1998), it has been shown in boreal Alaska that all change values exceeding the threshold value can be coded as fires (Kasischke *et al.* 1993, Kasischke and French 1995). Three thresholds developed in Alaska (table 1) were applied to imagery (Eidenshink and Faundeen 1994) obtained for the area north of  $51^\circ$  latitude in Ontario, Canada, for the 1992,

Table 1. Descriptions and references for AVHRR/NDVI threshold values used for fire mapping in this study.

Mapping method	Threshold value*	Reference
1	23	Kasischke <i>et al.</i> (1993)
2	24	Kasischke <i>et al.</i> (1993)
3	23 and 18 combined	Kasischke and French (1995)

\*Threshold values represent actual AVHRR/NDVI values ( $-10$  to  $+10$ ) scaled to a range of 0 to 200 on  $1063\text{ m} \times 1063\text{ m}$  spatial resolution, 10-day composite imagery.

1993, and 1995 fire seasons to map large fires. Unmanaged boreal forests that undergo little or no fire suppression dominate this geographical area, allowing fires to reach 1000s to 10 000s of ha. More detailed descriptions of the fire-mapping methods are provided in Rimmel (2000) and Rimmel and Perera (2000, 2001).

The resulting binary fire maps (images) were resampled from 1063 m to 531.5 m spatial resolution (doubling the original resolution) and geocorrected to match the projection of the GIS reference data. The increased resolution ensured a smoother fire boundary after the second-order geocorrection with nearest neighbour resampling (ERDAS Inc. 1997). Twenty-one ground control points were selected along distinctive hydrological features (e.g. shorelines, islands, rivers, and coastlines) as per the approach of Martín and Chuvieco (1995). The ground control and corresponding image reference points were used to calculate the RMS error. A RMS error of approximately 400 m was considered acceptable (Martín and Chuvieco 1995, Lunetta and Elvidge 1998). Each contiguous cluster of fire pixels was assigned a unique *code* for discriminating individual mapped fires. Finally, fire maps were converted to vector format.

## 2.2. Reference data

Reference data for assessing the accuracy of fire mapping consisted of a fire-pattern GIS database describing the dates and locations of fires in Ontario between 1921 and 1995 (Perera *et al.* 1998). We only used reference fires for the years of study (1992, 1993, and 1995). The database was compiled from hardcopy maps derived from field data, aerial photographs, and interpretation of Landsat images. The fire perimeters were originally digitized at varying scales, representative of their source data. These represent a historical account of fire in the province.

Fire polygons within the reference database often abutted or were adjacent to other fire polygons, forming clusters (i.e. not always dispersed). Given the variability in delineation scale and source, these nearby fire polygons could represent different areas of the same fire *event*, or each fire *event* could consist of several independently digitized polygons. Thus, all fire polygons within 1063 m (one AVHRR pixel dimension) of each other and belonging to the same year, were coded to indicate that they are a part of a common reference fire *event*. To remove effects of varying map scale among reference fire *events*, the reference database was rasterized to the same resolution as the mapped fires (531.5 m) and converted back to vector format, allowing direct area comparisons. Since 531.5 m resolution was larger than the fire boundary variability within the original reference data, this coarser resolution served to standardize the scale.

## 2.3. Preparation for accuracy assessment

For each treatment (NDVI threshold value), the geometric intersection (union) between mapped fires and the reference database were computed using ArcInfo (ESRI Inc. 1997). Overlapping polygons were split and given attributes from both original polygons (figure 1). The original reference and mapped fire polygons remained available by querying with original polygon identifiers (i.e. *event* and *code*). These part-polygons in figure 1 were sufficient to calculate proportions of areas mapped within reference boundaries but not for calculating the total contiguous mapped area associated with a reference fire *event* (discrete feature).

In figure 1 all parts of the large reference polygon (comprising of both R and  $D \cap R$ ) contain a common *event* number. However, the adjoining part-polygon labelled D, has

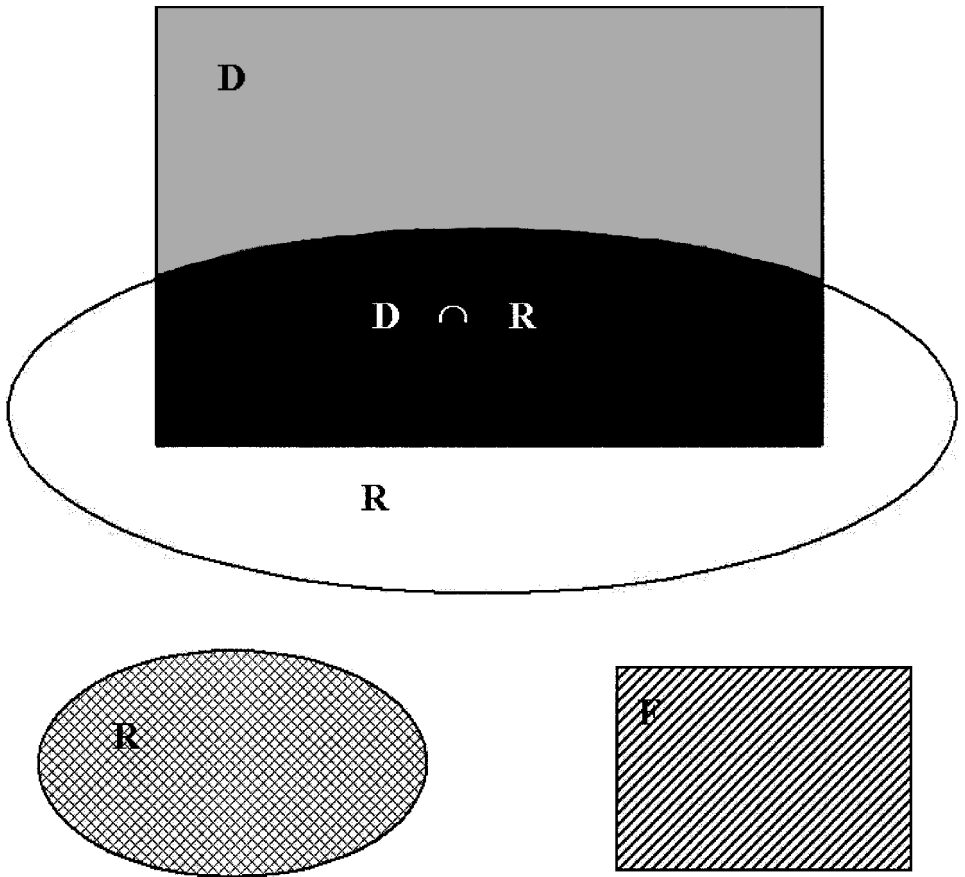


Figure 1. Hypothetical example of possible components formed by the geometric intersection of reference and mapped fire polygons where D = detected, R = reference, and F = falsely detected.

no *event* associated with it. To compute various measures of total mapped areas associated with reference *events*, unique *event* numbers were added to the external, contiguous part-polygon records using an algorithm written by the authors.

To transfer the *event* number to D in all unique cases, the devised algorithm located all part-polygons adjoining reference polygons and associated them with the proper number. These linkages were possible because of the GIS database structure (table 2). The devised algorithm selected a unique *event* number (e.g. 4) from the union database. It would store each unique *poly-id* value (reference fire, part-polygon identifier) associated with that fire *event* (i.e. 29). Then, for each of those *poly-id* values, the algorithm stored all unique *code* values (i.e. 0, 95, 105). These *codes* were then parsed in sequence; e.g. if *code* > 0, then the *event* number was updated. Each unique fire *event* was automatically processed until all updates were completed.

#### 2.4. Traditional error matrix approach

For each treatment, an error matrix was generated. Two thousand reference points were selected for the study area using a simple random sampling design. These reference data were compared with the fire map data in each treatment to

Table 2. Union database structure with sample data.

Record	Poly-id	Detected	Reference	Event	Code
1	29	D	R	4	95
2	29	D		4	0
3	1		R	0	95
4	29	D	R	4	105
5	1		R	0	105

*Notes:* The detected and reference columns are populated to indicate whether polygons represent detected fires or reference fires. Since a reference fire may consist of more than one contiguous detected area, combinations of *code* and *poly-id* are used to update the discrete event number encompassing all components belonging to a single fire *event*.

compute producer's, user's, overall, and Kappa coefficient measures of accuracy (Congalton 1991).

### 2.5. Discrete-feature approach

The discrete-feature approach allowed 15 accuracy measurements to be computed from two levels of observation: (1) scene and (2) feature. The levels of observation were further dissected by considering three different perspectives (i.e. *map*, *event*, and *reference*). Scene-level variables returned one non-site-specific accuracy value per treatment, while the individual reference *event*-level variables returned accuracy values for each detected fire within a treatment. The benefit of the feature level of observation is that accuracy values for each *event* can be linked back to the original fire polygons as attributes, allowing accuracy to be mapped spatially for each *event*.
































Figure 2 describes the possible components resulting from the intersection of reference fire polygons with mapped fire polygons. Each component is given a unique hatch pattern that is used to describe various area combinations used for calculating the 15 new accuracy variables. The scene-level variables consider undetected and falsely detected fire areas even when they are not associated with features (*events*) because accuracy values are generalized for the entire scene (non-site-specific). The perspective defines the denominator for each variable, regardless of the observation level. At the feature level of observation, there are no map perspective variables because falsely detected fires would never intersect with reference fire polygons; therefore, those components would never enter computations.

The variables defined in figure 2 are aligned to indicate similarities between observation levels. The scene-level *correct*, *incorrect*, and *omission* variables correspond to the feature-level *interior*, *exterior*, and *omission* variables. The differences stem from the inclusion of undetected fire and falsely detected fire areas in scene-level variables. The mapped-to-reference ratio variable measures the ratio of mapped to reference area for each fire *event*, regardless of their proportion of intersection. The map perspective calculates accuracy relative to the total area mapped by the fire-detection algorithm. The event perspective computes accuracy relative to the total area of mapped events, disregarding all falsely detected fires. Finally, the reference perspective computes accuracy relative to the total area of reference fires.

## 3. Results/discussion

### 3.1. Traditional accuracy measures

The simple random selection of reference points across the background-dominated study area resulted in the majority of reference locations representing the

Scene-level				Feature-level			
Measure	Perspective			Measure	Perspective		
	Map	Event	Reference		Map	Event	Reference
Correct				Interior (of reference)	N/A		
							
Incorrect				Exterior (to reference)	N/A		
							
Omission				Omission (within reference)	N/A		N/A
							
Mapped-to- reference ratio	N/A	N/A	N/A	Mapped-to- reference ratio	N/A	N/A	
							






	Correctly mapped portion of fire
	Extraneously mapped portion of fire
	Unmapped portion of fire
	Undetected fire
	Falsely detected fire

Figure 2. Measures of accuracy for the discrete-feature method are computed for three perspectives (map, event, and reference) at the scene and feature levels. Unique hatching patterns describe all possible components resulting from the geometric intersection of reference and mapped fire polygons. Hatching patterns are used to describe the denominator and numerator combinations used for each of the 15 new accuracy variables.

background class. This led to almost perfect producer's and user's accuracy values for background. However, the few points residing in the fire class returned more varying results (table 3). Given the low count of reference points within the fire class, the accuracy values do not necessarily represent true ground conditions.

The misrepresentation of accuracy is viewed as a drawback of the error matrix approach. Since a simple random sample is required to meet the assumptions of Kappa coefficient and overall accuracy calculation (Congalton 1988), the discontinuous binary surface accuracy is automatically biased owing to a disproportionate distribution of class areas.

The overall accuracy reported for each treatment was >99%. However, when the numbers of correctly and incorrectly detected fires are considered (table 4), this level of accuracy appears much too high. Kappa coefficient values (table 3) appear to be better estimates of overall accuracy, but given the drastic differences in reference sample numbers between the two classes, the Kappa coefficient and overall accuracy calculations can be easily inflated (Fitzgerald and Lees 1994).

### 3.2. Feature-oriented, scene-level accuracy measures

Three measures of accuracy (i.e. *correct*, *incorrect*, and *omission*) were computed from each of three perspectives (i.e. *map*, *event*, and *reference*) for the scene level of each treatment (table 5). Since these values were based on all mapped and reference areas, the entire population, rather than just a simple random sample was assessed for accuracy. The scene-level feature-oriented assessments of accuracy differ drastically from the values of overall accuracy calculated from the error matrix approach. Since the feature-oriented values were more realistic, they confirm the need for these alternative assessment approaches. Furthermore, nine measures per treatment are

Table 3. Error matrix summary with accuracy and Kappa coefficient values for each treatment.

Year	Method*	Class name	Reference totals	Classified totals	Number correct	Accuracy (%)			
						Producer's	User's	Overall	Kappa
1992	1	Background	1992	1987	1980	99.40	99.65	99.05	0.091
		Fire	8	13	1	12.50	7.69		
	2	Background	1992	1990	1983	99.55	99.65	99.20	0.107
		Fire	8	10	1	12.50	10.00		
	3	Background	1992	1981	1977	99.25	99.80	99.05	0.292
		Fire	8	19	4	50.00	21.05		
1993	1	Background	1994	1997	1994	100.00	99.85	99.85	0.666
		Fire	6	3	3	50.00	100.00		
	2	Background	1994	1999	1994	100.00	99.75	99.75	0.285
		Fire	6	1	1	16.67	100.00		
	3	Background	1994	1994	1993	99.95	99.95	99.90	0.833
		Fire	6	6	5	83.33	83.33		
1995	1	Background	1988	1990	1984	99.80	99.70	99.50	0.543
		Fire	12	10	6	50.00	60.00		
	2	Background	1988	1991	1985	99.85	99.70	99.55	0.569
		Fire	12	9	6	50.00	66.67		
	3	Background	1988	1988	1984	99.80	99.80	99.60	0.665
		Fire	12	12	8	66.67	66.67		

\*Fire mapping methods are defined in table 1.

Table 4. Number of reference fires, detected fires, and incorrectly detected fires for each treatment.

Year	Mapping method*	Number of fires		
		Reference	Detected	Incorrectly detected
1992	1	12	8	101
	2	12	6	88
	3	12	8	117
1993	1	10	7	40
	2	10	6	18
	3	10	7	45
1995	1	34	13	13
	2	34	13	10
	3	34	13	14

\*Fire mapping methods are defined in table 1.

Notes: No specification of accuracy is provided by these numbers; they indicate whether or not contiguous clusters of detected pixels exhibited any overlap with reference fire polygons.

available for assessment rather than a single measure as in the square error matrix approach.

The *map* perspective answers the question: how well (or poorly) were fires mapped with respect to the total mapped area? The *event* perspective indicates how well (or poorly) fires were mapped with respect to all mapped areas associated with reference fire *events*. Finally, the *reference* perspective allows accuracy to be assessed with respect to all reference fire areas.

Table 5. Three measures of accuracy from three perspectives at the scene level.

Year	Method*	Sensor correct (%)	Perspective incorrect (%)	Omission (%)	Event correct (%)	Perspective incorrect (%)	Omission (%)	Reference correct (%)	Perspective incorrect (%)	Omission (%)
1992	1	7.5	92.5	41.4	60.5	746.0	334.0	15.3	189.1	84.7
1992	2	8.5	91.5	55.5	63.1	678.4	411.6	13.3	142.9	86.7
1992	3	8.9	91.1	23.0	59.9	612.6	154.7	27.9	285.5	72.1
1993	1	18.5	81.5	54.2	63.2	277.9	184.9	25.5	112.0	74.5
1993	2	31.4	68.6	143.2	62.0	135.6	282.9	18.0	39.3	82.0
1993	3	19.1	80.9	20.9	57.5	244.4	63.1	47.7	202.5	52.3
1995	1	69.1	30.9	90.5	78.4	35.1	102.7	43.3	19.4	56.7
1995	2	72.3	27.7	100.0	79.9	30.7	110.6	41.9	16.1	58.1
1995	3	56.8	43.2	52.9	73.6	56.0	68.5	51.8	39.4	48.2

\* Fire mapping methods are defined in table 1.

The three measures within each perspective provided a means for further discriminating areas correctly mapped from those not mapped or incorrectly mapped. These measures indicate the relative proportion of the indicated condition with respect to a reference area as governed independently by each of the three perspectives (figure 2).

The accuracy values reasonably reflect the expected differences among mapping methods and years. For example, mapping method 3, representing the lowest threshold of NDVI change, correspondingly exhibits lowest *omission*. It does not always exhibit highest *correct* values because the lenient nature of the threshold increases *incorrect* detection simultaneously; increasing the denominator acts to decrease observed accuracy. Mapping method 2 uses the strictest (highest) threshold value; consequently, *omission* errors are highest owing to the limited areas mapped. Conversely, values for *incorrect* fire mapping are lowest. Values for *correctly* mapped fires vary depending on the amount of mapped/non-mapped area influencing the denominator.

We believe that thin clouds present in the imagery in 1992 and 1993 acted to depress NDVI values resulting in poor accuracy (e.g. Kasischke and French 1997). The considerably better values in 1995 were partly due to the occurrence of several large, easily detected fires (Kasischke *et al.* 1993), but also to cloud-free imagery.

### 3.3. *Feature-oriented, feature-level accuracy measures*

Aside from calculating scene-level accuracy measures, the discrete-feature approach can partition accuracy into components providing a means for calculating accuracy measures for sub-scenes and individual features. We calculated accuracy measures (figure 2) for all detected fires individually, which allowed us to determine accuracy among and within individual fires on the background.

Feature-level assessment provides site-specific accuracy values. The *event* perspective's *interior* and *exterior* values indicate the percentage of the mapped area (without regard for absolute reference fire size) that lies inside/outside the reference fire perimeter with respect to the total area mapped. Thus, it measures the positional accuracy of the mapped area (overlap of mapped and reference area). The *omission* value indicates the reference fire area remaining unmapped in relation to the area of the mapped portion for that object.

The reference perspective's *interior* and *exterior* values indicate the percentage of reference area correctly and incorrectly mapped. The *mapped-to-reference* measure of accuracy provides insight to the areal correspondence between reference and mapped area without regard for absolute location. Thus, the *mapped-to-reference* ratio reduces the concern of spatial misregistration, an issue of confusion as indicated by Congalton and Green (1993).

The benefit to calculating percentages is that they can be used to compare among treatments and discrete features. Therefore, accuracy can be compared for individual features. Feature-oriented accuracy assessment does not assume homogeneous error across the entire image, allowing this potential variability to be tested from two perspectives, potentially providing a distribution of varying accuracy levels across the fire map.

### 3.4. *Spatial distribution of accuracy values*

The computed feature-level accuracy measures were linked back to the original GIS database as attributes for each reference fire *event*. This allowed the construction of choropleth maps showing accuracy distributions for each of the accuracy measures

independently. As an example, figure 3 displays the 1992, 1993, and 1995 *interior* accuracy results for the fire map generated by mapping method 3. The figure categorizes each fire into one of three levels of accuracy described in the caption.

Visualization of spatial accuracy could help indicate regions of the image where accuracy was especially poor or good. Significant differences in accuracy may be related to misregistration, ecological phenomena, or user error. By combining quantitative accuracy assessment results with the ability to spatially recognize patterns, overall evaluation will become more rigorous.

#### 4. Conclusions

Accuracy assessment techniques relying on the square error matrix method may return incorrect values for discontinuous binary surfaces. In this study, overall accuracy and Kappa coefficient values were unreasonable for the square error matrix, owing to the disproportionate distribution of reference data stemming from the simple random sampling design. Since simple random sampling is required to meet the assumptions for calculating these values, the error matrix approach was concluded to be insufficient for assessing mapping accuracy of discontinuous binary surfaces.

The proposed object-oriented method of accuracy assessment for discontinuous binary surfaces using an existing GIS database showed many benefits over the error matrix approach. Most importantly, an image could be partitioned, allowing accuracy assessment of either the entire scene or various subsets thereof. Furthermore, since the reference data were in polygon GIS format, all data were used for accuracy assessment, eliminating the need for a sampling procedure to select reference locations.

Combinations of three perspectives, seven accuracy measures, and two levels of

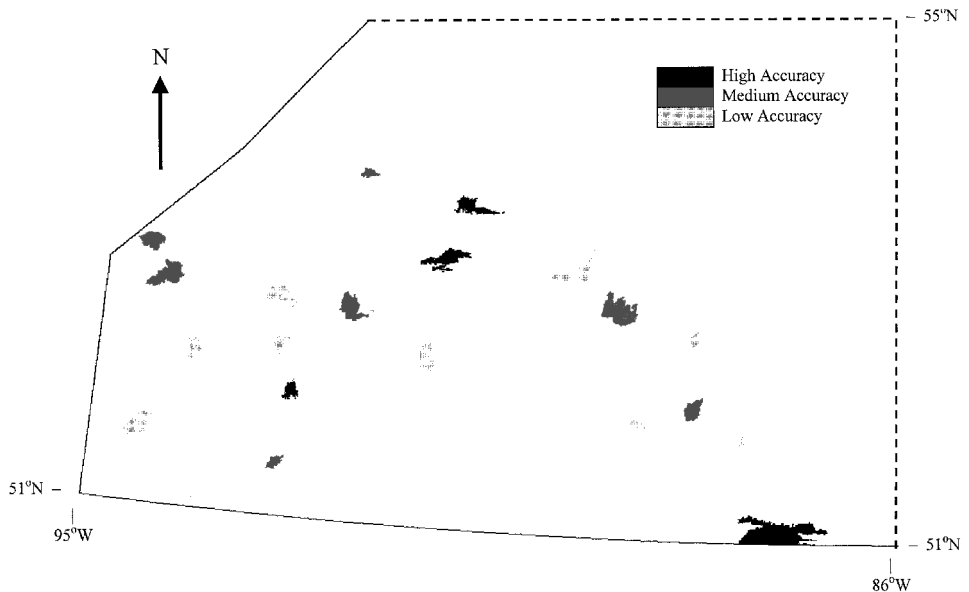


Figure 3. Spatially mapped *interior* accuracy at the feature level of observation for fires mapped with method 3 in 1992, 1993, and 1995; high accuracy >80%, medium accuracy 70 to 80%, and low accuracy 50 to 70%.

observation allow various aspects of accuracy to be assessed. This reduces the reliance on an all-encompassing accuracy value, and allows discrimination of accuracy depending on specific user requirements. Underlying this is the ability to focus on areal and/or positional accuracy. By disregarding precise positional accuracy, areas of individual mapped features can still be compared to reference areas. Thus, if calculating fire areas were the major concern, small spatial misregistration problems could be ignored.

Generation of choropleth maps from accuracy measures linked to the original GIS database as attributes provides a powerful method for visualizing the distribution of accuracy. These distribution maps can be used to isolate regions of high or low accuracy leading to further hypotheses, analyses, and explanations of data quality.

## References

- BARRETTE, J., AUGUST, P., and GOLET, F., 2000, Accuracy assessment of wetland boundary delineation using aerial photography and digital orthophotography. *Photogrammetric Engineering and Remote Sensing*, **66**, 409–416.
- BLAKEMORE, M., 1984, Generalisation and error in spatial data bases. *Cartographica*, **21**, 131–139.
- CHRISMAN, N., and LESTER, M., 1991, A diagnostic test for error in categorical maps. *Technical papers ACSM-ASPRS Annual Convention, Baltimore*, **6**, 330–348.
- COHEN, J., 1960, A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, **20**, 37–46.
- CONGALTON, R. G., 1988, A comparison of sampling schemes used in generating error matrices for assessing the accuracy of maps generated from remotely sensed data. *Photogrammetric Engineering and Remote Sensing*, **54**, 593–600.
- 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 GREEN, K., 1993, A practical look at the sources of confusion in error matrix generation. *Photogrammetric Engineering and Remote Sensing*, **59**, 641–644.
- CONGALTON, R. G., and MEAD, R. A., 1983, A quantitative method to test for consistency and correctness in photointerpretation. *Photogrammetric Engineering and Remote Sensing*, **49**, 69–74.
- CONGALTON, R. G., ODERWALD, R. G., and MEAD, R. A., 1983, Assessing Landsat classification accuracy using discrete multivariate analysis statistical techniques. *Photogrammetric Engineering and Remote Sensing*, **49**, 1671–1678.
- EIDENSHINK, J. C., and FAUNDEEN, J. L., 1994, The 1-km AVHRR global land data set: first stages in implementation. *International Journal of Remote Sensing*, **15**, 3443–3462.
- ERDAS, Inc., 1997, ERDAS Imagine version 8.3.1, Build 59 (Windows NT), ERDAS Inc.
- ESRI, Inc., 1997, Arc/Info version 7.1.1 (UNIX), Environmental Systems Research Institute, Inc. Redlands, California, USA.
- FERNANDEZ, A., ILLERA, P., and CASANOVA, J. L., 1997, Automatic mapping of surfaces affected by forest fires in Spain using AVHRR NDVI composite image data. *Remote Sensing of Environment*, **60**, 153–162.
- FITZGERALD, R. W., and LEES, B. G., 1994, Assessing the classification accuracy of multisource remote sensing data. *Remote Sensing of Environment*, **47**, 362–368.
- HUVELINK, G. B. M., BURROUGH, P. A., and STEIN, A., 1989, Propagation of errors in spatial modelling with GIS. *International Journal of Geographical Information Systems*, **3**, 303–322.
- JENSEN, J. R., 1996, *Introductory digital image processing, A remote sensing perspective, Second Edition* (New Jersey: Prentice Hall).
- KASISCHKE, E. S., and FRENCH, N. H. F., 1995, Locating and estimating the areal extent of wildfires in Alaskan boreal forests using multiple-season AVHRR NDVI composite data. *Remote Sensing of Environment*, **51**, 263–275.
- KASISCHKE, E. S., and FRENCH, N. H. F., 1997, Constraints on using AVHRR composite imagery to study patterns of vegetation cover in boreal forests. *International Journal of Remote Sensing*, **18**, 2403–2426.

- KASISCHKE, E. S., FRENCH, N. H. F., HARRELL, P., CHRISTENSEN, N. L. JR., USTIN, S. L., and BARRY, D., 1993, Monitoring of wildfires in boreal forests using large area AVHRR NDVI composite image data. *Remote Sensing of Environment*, **45**, 61–71.
- LUNETTA, R. S., CONGALTON, R. G., FENSTERMAKER, L. K., JENSEN, J. R., MCGWIRE, K. C., and TINNEY, L. R., 1991, Remote sensing and geographic information system data integration: Error sources and research issues. *Photogrammetric Engineering and Remote Sensing*, **57**, 677–687.
- LUNETTA, R. S., and ELVIDGE, C. D., 1998, *Remote sensing change detection, environmental monitoring methods and applications* (Michigan: Ann Arbor Press).
- MARTÍN, M. P., and CHUVIECO, E., 1995, Mapping and evaluation of burned land from multitemporal analysis of AVHRR NDVI images. *EARSeL Advances in Remote Sensing*, **4**, 7–13.
- NÆSSET, E., 1996, Use of the weighted Kappa coefficient in classification error assessment of thematic maps. *International Journal of Geographical Information Systems*, **10**, 591–604.
- PERERA, A. H., BALDWIN, D. J. B., SCHNEKENBURGER, F., OSBORNE, J. E., and BAE, R. E., 1998, Forest fires in Ontario: A spatio-temporal perspective. Forest Research Report No. 147, Ontario Ministry of Natural Resources, Ontario Forest Research Institute, Sault Ste. Marie, Ontario, Canada.
- RAZAFIMPANILO, H., FROUIN, R., IACOBELLIS, S. F., and SOMERVILLE, C. J., 1995, Methodology for estimating burned area from AVHRR reflectance data. *Remote Sensing of Environment*, **54**, 273–289.
- REMMEL, T. K., 2000, Fire disturbance mapping in a northern boreal forest using AVHRR/NDVI imagery: comparing techniques of change detection and substrate correction. M.Sc.F. thesis, Lakehead University, Thunder Bay, Ontario.
- REMMEL, T. K., and PERERA, A. H., 2000, Accuracy of remotely sensed fire mapping assessed with a GIS database. In *GIS 2000 Conference Proceedings*, CD-ROM.
- REMMEL, T. K., and PERERA, A. H., 2001, Fire mapping in a northern boreal forest: assessing AVHRR/NDVI methods of change detection. *Forest Ecology and Management*, **162**, 119–129.
- ROSENFELD, G. H., 1986, Analysis of thematic map classification error matrices. *Photogrammetric Engineering and Remote Sensing*, **52**, 681–686.
- ROSENFELD, G. H., and FITZPATRICK-LINS, K., 1986, A coefficient of agreement as a measure of thematic classification accuracy. *Photogrammetric Engineering and Remote Sensing*, **52**, 223–227.
- STEHMAN, S. V., 1992, Comparison of systematic and random sampling for estimating the accuracy of maps generated from remotely sensed data. *Photogrammetric Engineering and Remote Sensing*, **58**, 1343–1350.