ACCURACY ASSESSMENT METHOD FOR WETLAND OBJECT-BASED CLASSIFICATION

M. Grenier*, S. Labrecque, M. Benoit and M. Allard

Environment Canada, Canadian Wildlife Service, Québec region, 1141 route de l'Église, Québec, Québec, G1V 4H5

KEY WORDS: Accuracy assessment, object-based classification, error matrix, fuzzy logic, wetland mapping

ABSTRACT:

The object-based classification approach needs an adaptation of the traditional accuracy assessment methodology. This paper presents a methodology that considers the inherent variability within each wetland class for the sampling size calculation based on objects and a stratified random sampling to select samples for each category. The polygons selected for validation are overlaid on the data used for classification. The analyst uses all ancillary data available to build an interpretation key which defines the labels of the polygons. The error matrix representing the whole territory could then be calculated. Finally, considering that wetland classes are not completely discrete and that the delineation is rarely categorical, fuzzy logic is used to analyze the accuracy. Decision rules are built based on class description to accept a relative level of confusion between similar classes.

1. INTRODUCTION

1.1 Accuracy assessment considerations

The knowledge required to manage territory is increasingly based on information and maps created from remote sensing images (Campbell 2007; Jensen 2005). These maps generally provide an approximate and synthetic description of the geographic reality, and therefore the results obtained through image analysis contain information that is not completely accurate (Casemier et al., 1998). The validation of results from image processing has become an inherent component of mapping projects using remote sensing (Song et al., 2005; Mowrer, 2000; Congalton and Green 1999). Frank (1998) reported that the evaluation of the precision of results must inform the user on the quality of the product as well as the limits of its use. According to Congalton and Green (1999), the validation informs about the quality of the results obtained by using remote sensing images and comparing them to the reference data, assuming that these data are accurate. Validation must therefore provide information on the exactness, based on a degree of precision defined and confirmed by a test of hypothesis (USGS, 1994). The development of remote sensing validation methods has increased considerably, going from the qualitative confidence-building assessment to the quantitative evaluation of these results based on statistical methods (Congalton, 2004; Kyriakidis and Dungan, 2001; Mowrer, 2000). The error matrix has become a standard in the evaluation of remote-sensing results (Congalton, 2004, Congalton and Green, 1999; Hammond and Verbyla, 1996; Congalton, 1991; Story and Congalton, 1986). However, if the matrix is created without taking into account statistical requirements, the validation can be misleading (Mower, 2000; USGS, 1994). Hammond and Verbyla (1996) recommended inclusion of methodological details for the calculation of error matrix, namely the sample size, sample plan, data and references. This way, the user will be able to evaluate the risk associated with the decisions based on these results. Additional methods such as the standard error matrix (Congalton, 1991), the Monte Carlo Method (Goodchild et al., 1992), geostatistics (Kyriakidis and Dungan, 2001) and bootstrapping (Verbyla and Hammond, 1995) have been developed overtime to improve the required statistical conditions. Despite these developments, validation of results is still a concern for projects based on image analysis (Campbell 2007; Jensen 2005). In remote sensing, calculation of minimum sample size needed for population representativeness is based on a law of probability corresponding to a binomial and multinomial distribution (Congalton and Green, 1999; Congalton, 1991; Story and Congalton, 1986). Although this equation is independent from the reference unit, it seems to be better adapted to a traditional pixel-based image classification, since the number of samples is generally very high (Grenier et al., 2007; Wang and Howart, 1993). In the case of a classification using an object-based approach, the image is broken down into several smaller spatial objects that are relatively homogenous (segments). These objects are then classified following traditional supervised classification methods (nearest neighbor) or according to fuzzy logic using statistical membership functions. The thematic map resulting from the object-based classification must therefore be evaluated by taking the following characteristics into account: thematic accuracy must be based on the object; ground control data or reference data must reflect the object; and spatial precision must validate the fit between the segment and the object of interest (Tiede et al., 2006; Schöpfer and Lang, 2006, Zhan et al., 2005). In real world, land use classes are not completely distinct, and limits are rarely categorical. Furthermore, area covered by one pixel may encompass several categories thus resulting in a mixed pixel (Zhan et al., 2005; Campbell, 2007). Assigning these mixed pixels to a given class may result in overestimating the class. In order to take this uncertainty into account, fuzzy logic was introduced (Fritz and See, 2005; Falzarano and Thomas, 2004; Zhu, 2001; Woodcock and Gopal, 2000).

1.2 Wetland considerations

Wetlands are complex and can vary depending on water conditions. They often have various vegetation structures which, in addition to the presence of water, results in an heterogeneous spectral response (Founier et al., 2007; Ozesmi and Bauer, 2002). Object-based classification improves wetlands identification. The segmentation process takes into account neighboring pixels and provides a better ecological

* Corresponding author (Marcelle.Grenier@ec.gc.ca)

definition of these environments (Grenier et al., 2007; Hurd et al., 2006; Blaschke and Strobl 2001). Grenier et al., (2007), developed a method for mapping wetlands with object-based classification using Landsat-ETM and RADARSAT-1 images. This top-down approach aims to map the five wetlands classes, as defined by The Canadian Wetland Classification System (CWCS) (NWWG, 1987). These five classes (shallow water, marsh, wetland, fen and bog) are mapped with a minimum geographic unit of one (1) hectare. Since the characterization of wetlands is based on medium-resolution satellite images, the description of classes is a generalization of the characteristics of each wetland category. For example, in the case of bogs and fens, objects are classified in one of the peatland classes according to ombrotrophic and minerotrophic dominance of the object, regardless of ecological criteria such as formation and structure. An error in identifying a bog versus a fen is therefore less important than the distinction between wetland and upland (Fournier et al., 2007). Remote sensing may introduce sources of error during processing steps (Jensen 2005). As part of this study, we assumed that the errors generated during acquisition and pre-processing of images were overall uniform, and therefore had no specific effect on determining wetland types. Statistical rigor is essential to accurately quantify results obtained from image analysis (Congalton, 2004). The validation approach suggested in this paper is adapted from the method developed by Morisette et al., 2004, as part of the Global Land Cover project (Global Land Cover 2000) along with the fuzzy accuracy approach of Green and Congalton 2004 and Falzarano et al. 2004 in the NIMA project. The main objective of this article is to propose a practical validation method for object-based classification of wetlands in order to inform on the accuracy as well as respect the statistical requirements for measuring precision.

2. DATA COLLECTION

2.1 Site description

As part of the methodological development phase of Canadian Wetland Inventory, several sites in Quebec, Canada have been tested (Grenier et al., 2007). Among these sites, the Montérégie region was analyzed more closely because of the greater availability of ancillary data. The region of Montérégie covers an area of 11,883 km² and is part of the Mixedwood Plains ecozone (Ecological Stratification Working Group, 1995). This ecozone is characterized by a high agricultural density with some of the most productive agricultural soils in Canada. On the other hand, forest cover is relatively small (less than 10%). Approximately half of Canadian population lives in this ecozone.

2.2 Field data

Orthophotos at 1: 40,000 acquired in 1998-1999-2000 by the *Agence de géomatique Géomont* (www.geomont.qc.ca) were made available for this study. Ground control data were collected using random stratified sampling. Sample size was predefined to 50 per class based on Congalton and Green (1999) recommendations. Only the five wetland classes were sampled with a total of 250 samples overall. Of that number, 90 were validated in the field or were flown over by plane in the fall of 2007. The remaining samples were validated by visual interpretation on aerial photos. Note that the sample size calculation based on variability, as suggested in this study, was carried out after the field campaign.

2.3 Images

Wetlands in the Montérégie sector were mapped with four Landsat-ETM+ images acquired in 2001 and 2002, and three RADARSAT-1 images acquired in the spring of 2004. The object-based classification method developed to extract the five wetland classes is described in Grenier et al. (2007). Spatial resolution of summer images was resample from 25 m to 12,5 m using the *Pansharp* module from PCI Geomatica. Classification results of Landsat-ETM and RADARSAT-1 images for wetland mapping are presented in Table 1.

Table 1: Distribution of classes based on the area and number of polygons

Classes	Area (ha)	% area	nb polygons	% of polygons
Bog	21803	2,25	236	9,45
Fen	1404	0,15	23	0,92
Swamp	15367	1,59	960	38,43
Marsh	9846	1,02	623	24,94
Shallow water	569	0,06	53	2,12
Other	852907	88,13	301	12,05
Open water	65933	6,81	302	12,09
Total	967829		2498	

3. METHOD

3.1 Reference Unit

In the case of object-based classification, the attributes used for classification are calculated based on the object; consequently, the sampling unit for validation is the object or polygon (Tiede et al., 2006; E. Schöpfer and S. Lang, 2006; Zhan et al., 2005). Nevertheless, in the case of multi-level classification, the final result is a map made up of contiguous polygons belonging to the same class. Since the objective of validation is to inform about the result and not about the classification process, polygons from different levels were merged to produce a single polygon. This final polygon, resulting from the merger of objects, is used for validation.

3.2 Sample Size

In practice, the number of samples is limited by the operational constraints of a study, and often represents a compromise between the wish to obtain a precise measurement and the desire to remain efficient or able to process all samples properly. According to Congalton and Green (1999), the evaluation of the number of points required to validate the results of an image is based on several criteria, including the number of classes (or sampling strata), and their proportion. From a statistical perspective, the number of samples to be validated must be adequate for measuring the variability associated with the variable tested. The sample size needed to validate the thematic map with several classes depends on a multinomial distribution (Equation 1) (Congalton and Green, 1999).

$$N = \frac{B \prod_i (1 - \prod_i)}{b^2_i} \quad (1)$$

Where:

$\prod_i$ is the proportion of a population (number of polygons) in the *i*th class out of k classes that the proportion closest to 50% (in our case, swamp is 38%), b is the desired precision (5%), is the upper of the chi square (X^2) distribution with 1 degree of

freedom (1- as X^2 (1, 0.99285) = 7.04 and k is the number of classes. N = **7.**04 (0.38) (1-0,38)/0,05², i.e., 666 samples/7 classes = 95 samples per class. This number is very high and requires far too much effort compared to the classification process. Congalton (1991) and Congalton and Green (1999) suggest to use a rule of thumb and collect a minimum of 50 samples for each class in the error matrix (*N*=350, 7 classes X 50 samples). In theory, the sample size should be biased against the variability of a class (Congalton, 2004; Congalton and Green 1999; Congalton, 1991). A wetland category is considered variable when it is difficult to classify. This high variability can be explained at the spectral level, when the wetland is highly heterogeneous, and at the spatial level, when it is very abundant. It is then deemed to cover a large spectral signature. Note that a class that is barely present on a territory could also be difficult to classify, but this is mainly due to missing information required to establish a valid spectral signature. To calculate the variability of a class, averages and standard deviations are used for the main attributes that serve to classify the objects. These statistics were calculated for the entire Montérégie territory including all polygons for each wetland class as well as the *open water* and *other* classes. We can then calculate the variance. The variance is used to combine all values within a dataset to obtain the measure of dispersion. Variance of a discrete variable made up of *n* observations is defined as follows:

$$s^2 = \frac{\sum (x - \bar{x})^2}{n}$$

The pooled standard deviation for all sites is then calculated using Equation 3 (Thalheimer, W. and Cook, S., 2002)

$$s_{pooled} = \sqrt{\frac{(n_t - 1)s_t{}^2 + (n_c - 1)s_c{}^2}{n_t + n_c}} \qquad (3)$$

The pooled standard deviation (S *pooled*) obtained for each class can then be used as a weighting factor in Equation 4.

$$n_{var} = S_{class} / S_{total} * N \qquad (4)$$

S_{class} is the standard deviation for the class over total standard error for all S_{total} classes. *N* (350) is the sample size (7 classes X 50 samples). Thus, a high standard deviation means that the class has a greater spectral range and is therefore difficult to classify. Since the objective is to weight samples based on the variability of classes, the sample size should be reduced for classes that are easier to identify, or increased for those that are not. However, less frequent classes are generally the most difficult ones to identify. In this case, all available objects are evaluated.

3.3 Sampling design

Statistically speaking, sampling must be representative of the whole territory and selected without bias to ensure the validity of the statistical analyses (Congalton, 2004). In this study we used a systematic unaligned sampling, which combines the advantages of randomness and stratification with the benefits of systematic sampling, without falling into the trap of its periodicity (Congalton and Green, 1999; Jensen, 2005). Initially, the image to be validated is divided into cells systematically distributed to give *N* equal cells. The number of cells is equivalent to the sample size. Within these cells, the points are placed randomly in order to select the polygons to be validated. This approach ensures that the samples are selected randomly in each strata of wetland, but are distributed uniformly or systematically throughout the image. We then assume that all polygons in a cell have an equal probability to be chosen within the stratification. Within these cells, a point is randomly placed. It is then possible to select polygons of different classes that are closest to the point. Figure 1 gives an example for the selection of a bog polygon. These polygons will then be selected for validation of wetland classes.
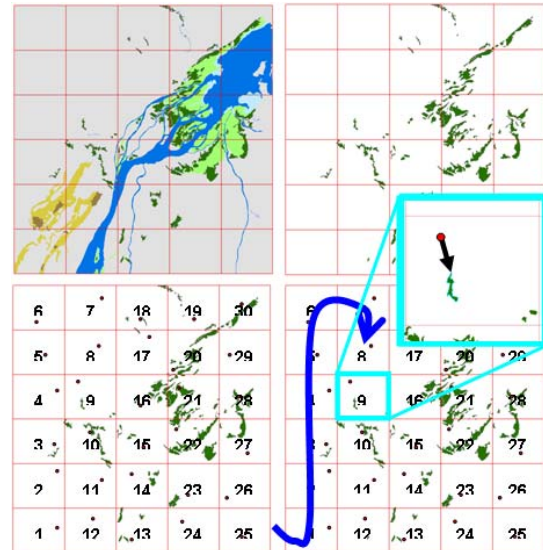

Figure 1: Example of the selection of a "bog" polygon based on stratified unaligned sampling.

3.4 Visual evaluation of results

Interpreting results. To insure the independence between data, the polygons to be validated are replaced without labels on the Landsat/ RADARSAT color composite used for classification. The analyst who carries out the validation at no time participates in classifying the image. All ancillary data that can inform on the type of wetland are used as references to built interpretation keys. Very often, ground control points and existing maps provide context that is different from the one observed on the images (e.g.: Emerging plant size in the marsh, water level in the wetlands). It is preferable to use ancillary data as a support to the interpreter's decision. Figure 2 shows a wetland polygon selected for validation superimposed over a Landsat color composite image. From this image, the analyst can assign one of the wetland classes to the polygon.

In order to validate the objectivity of the interpretation made by the analyst, the results of the validation derived from the interpretation were compared 1) with the ground survey and 2) with the interpretation of aerial photos made by a second analyst.
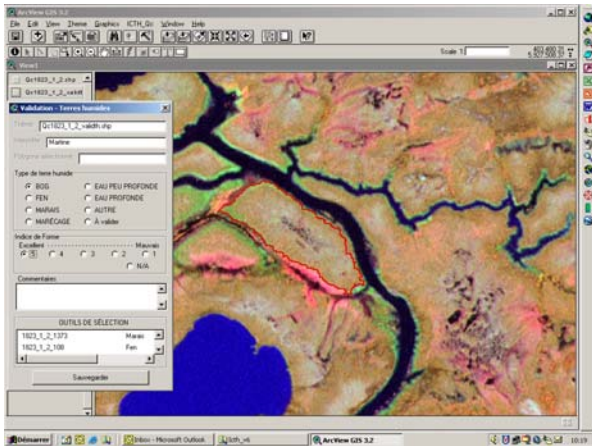
Figure 2: An example of the validation module using the ArcGis software

3.4.1 Fuzzy logic error matrix: Given that wetland classes' definitions from CWCS are based on biological criteria observed on the ground, use of remote sensing to map wetland could introduce overlapping between classes. However, overlapping may be explained when looking at the spectral response. Based on the classes' description, it is possible to create decision-making rules that completely or partially accept overlapping. A four-level agreement system was developed, and a new precision estimate of the wetland classes was calculated.

4. RESULTS

4.1 Reference Unit

The four hierarchical levels for segmentation/classification were merged to create polygons representative of wetland patches. These polygons were merged using ArcGis software from ESRI. Therefore, the population is the number of merged polygons, and the proportion of a class corresponds to the number of polygons associated with this class divided by the total number of merged polygons (table 1).

4.2 Sample Size

For the Montérégie site, the sample size calculated with the multinomial formula (1) gave 666 samples. To increase efficiency in the validation process, the number of samples should be reduced and adjusted to the inherent variability of each class with respect to statistical requirements. The sample size was based on a rule of thumb of 50 per class (7) giving 350 samples. It is therefore possible to weight the sample size according to the effort needed to classify a category.

Table 2: Sample size modified according to variability

Class	S pooled	n (var)
Shallow water	0,1350	3
Fen	1,2738	25
Open water	1,5310	31
Swamp	1,9440	39
Marsh	2,7315	55
Other	3,2330	65
Bog	6,6678	133
Total	17,5160	350

4.3 Sampling design

The territory was divided according to a systematic grid, to provide 350 equal cells, i.e., a grid of 14 by 25. Within these cells, points were placed randomly in order to select the polygons to be validated. For each class, the number of polygons selected is equivalent to the sample size.

4.4 Error matrix

For each polygon, the class resulting from image classification is compared against the class defined by the interpreter, which corresponds to the reference. Table 3 shows the error matrix with omission and commission errors.

Table 3: Error matrix for the Montérégie classification (B=bog, F=fen, SW= Shallow water, M=Marsh, S=Swamp, OW= open water and O=other, % of O=% of omission, % of C=% of Commission).

Class	Reference (image interpretation)							Tot
	B	F	SW	M	S	OW	O	
B	74	4	3	15	22	1	14	133
F	4	15	0	5	1	0	0	25
SW	0	0	3	0	0	0	0	3
M	1	0	1	45	7	0	1	55
S	2	0	0	1	36	0	0	39
OW	0	0	3	0	0	28	0	31
O	0	0	0	2	2	0	61	65
Total	81	19	10	68	68	29	76	351
% of O	9	21	70	34	47	3	20	-
% of C	44	40	0	18	8	10	6	-
Global accuracy								75
Kappa								69

Validation derived from the interpretation was compared with 90 wetland polygons in the field. Table 4 shows details for each class. Bog and fen classes were grouped under peatland class because there were no distinctions made in the field. 81% of peatland visited were correctly interpreted but only 55% of the peatland interpreted were real. Note that peatland is the most variable wetland class.

Table 4: Comparison of reference from ground truth and aerial survey with interpretation on the images (O=Other, P=Peatland, SW= Shallow water, M=Marsh, S=Swamp, OW= Open water, and EP= exploited peatland).

Nb poly.	Reference (ground truth and aerial survey)								
	O	P	SW	OW	M	S	EP	Tot	% O.
O				1				1	0
P		11			4	4	1	20	55
SW		1	13	1	3			18	72
OP			2					2	0
M	3		1		19	1		24	79
S		5			4	14		23	61
EP							2	2	100
Total	3	17	16	2	30	19	3	90	
% of C.	0	81	50	63	74	65	67		65

A second validation for the objectivity of the interpretation made by the analyst was compared with the interpretation of aerial photos made by an independent analyst.

Table 5: Comparison of interpretation of orthophotos and interpretation on the images

Nb poly.	Reference (interpretation of orthophotos)								
	O	P	SW	OW	M	S	E.P	Tot	% O
O	2	2		1			1	6	33
P	2	15	1		2	9		29	52

SW			6	5	11			24	25
OW				7	16			23	69
M	7	8	2		8	4	2	31	26
S	7	9			8	17	4	41	41
EP							4	4	100
Tot	18	34	16	22	29	30	11	158	
% of C.	1	44	37	73	27	57	36		43

4.5 Fuzzy logic matrix

Based on the four-level agreement decision rules, a new precision estimate for wetland classes was calculated. The calculation used the maximum potential agreement to evaluate the % of correspondence for each class (table 6).

Table 6: Error matrix using fuzzy logic for the object-based classification, image interpretation as reference

	Reference (image interpretation)								
Class	B	F	SW	M	S	OW	O	Tot	Max
B	296	8	3	15	22	0	0	344	532
F	8	60	0	5	2	0	0	75	100
SW	0	0	12	0	0	0	0	12	12
M	1	0	2	180	7	0	0	190	220
S	2	0	0	1	144	0	0	147	156
OW	0	0	3	0	0	112	0	115	124
O	0	0	0	0	0	0	244	244	260
Tot	307	68	20	201	175	112	244	1127	1404
Global Accuracy									80,3

5. RESULTS ANALYSIS AND DISCUSSION

The calculation of the sample size must be adjusted for use with object-based classification. If not, the number of objects to validate requires a disproportionate effort compared to the classification. Calculation of variability for wetland classes allows optimization of validation time by focusing on those classes that are difficult to identify. It should be noted that the variability measured according to standard deviation of attributes clearly reflects the effort of the analyst during wetland classification. Agreement level of 65% between image interpretation and ground surveys shows that the validation by interpretation approach is reliable. Main confusions happen 1) in classes that are often contiguous in space and for which transition is sometime fuzzy (marsh vs. swamp; marsh vs. shallow water) 2) in situations where the geographical context is not considered (peatland vs. swamp). A weaker agreement level of 43% between interpretation of orthophotos and interpretation on the images shows that it is wiser to ensure that the interpretation keys used for reference are calibrated against the ground control data. In the same way as for the sample size calculation, class variability could define the proportion of ground data needed to better calibrate the interpretation approach. The calculation of the error matrix based on fuzzy logic is better adapted to the definition of classes obtained via image analysis. Global accuracy increases from 75% to 80%, which represents a stronger validation estimate when compared to other land cover classes (e.g. forest, agriculture). However, decision-making rules must also respect bio-geographical and climatic characteristics valid for a pre-defined territory. Ecological limits should be preferred (eco-regions and eco-districts).

6. CONCLUSIONS AND RECOMMENDATIONS

Validation of remote sensing results remains an issue and needs additional work particularly for adaptation of classification methods to object-based approaches. Validation methods should also take into account externals factors that impact its efficiency such as budget limitations and the need for national or global coverage. A similar issue arises in the case of very high resolution images where large amount of information also requires higher validation levels. Spectral and spatial variability can weight the sample size, and thus optimize the validation efforts. Therefore, a minimum of ground surveys is required to calibrate interpretation keys. Interpretation can fill information gap and enables the validation of the entire territory while achieving statistical requirements and reducing the time and costs associated with this step. This approach using interpretation can also serve as a quality control process and could also guide field survey operations. In addition, it can be applied to other themes than wetlands. Information derived from remote sensing images is increasingly used in management decisions. Therefore, validation of image classification must provide quantitative information on its exactness as well as the inherent limitations in using these products.

7. ACKNOWLEDGEMENTS

8. REFERENCES

Blaschke, T. and Strobl, J. 2001. What's Wrong with Pixels ? Some recent developments interfacing remote sensing and GIS. GIS - *Zeitschrift für Geoinformationssysteme*, 14(6), pp. 12-17.

Campbell, J.B. 2007. *Introduction to Remote Sensing*, Fourth Edition. The Builford press, New York, London, 625 p.

Cazemier, D., Lagacherie, P. and Martin-Clouaire, R. 1998. A fuzzy constraint approach for handling uncertainty in regional soil maps. In: *Data Quality in Geographic Information from Error to Uncertainty*. Edited by Michael Goodchild and Robert Jeansoulin. Hermers, pp.121-126.

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

Congalton, R.G. 2004. Putting the map back in map accuracy assessment. In: *Geospatial Data Accuracy Assessment*. Edited by Lunetta, R.L. and Lyons, J.G. U.S. Environmental Protection Agency. Las Vegas. Report No. EPA/600/R-03/064, pp. 1-11.

Congalton, R. G. and Green, K. 1999. *Assessing the accuracy of remotely sensed data: Principles and Practices*. Lewis Publisher, Boca Raton.

Ecological Stratification Working Group. 1995. *A National Ecological Framework for Canada*. Research Branch, Agriculture and Agri-Food Canada and Environment Canada, Ottawa, Ont.

Falzarano, S.R. and Thomas, K. A. 2004. Fuzzy set and spatial analysis techniques for evaluation of thematic accuracy of a land-cover map. In,: *Geospatial Data Accuracy Assessment*. Edited by Lunetta, R.L. and J. G. Lyons. U.S. Environmental Protection Agency. Las Vegas. Report No. EPA/600/R-03/064, p. 189- 207.

Fournier, R.A., Grenier, G., Lavoie, A., and Hélie, R. 2007. Towards a strategy to implement the Canadian Wetland

Inventory using satellite remote sensing. *Canadian Journal of Remote Sensing*, 33, Supp. 1, pp. S1-S16.

Frank, A. 1998. Metamodels for data quality description. In: *Data Quality in Geographic Information from Error to Uncertainty.* Edited by Michael Goodchild and Robert Jeansoulin. Hermers, pp. 15-30.

Fritz, S., and See, L. 2005. Comparison of land cover maps using fuzzy agreement. International Journal of Geographical Information Science, 19(7), pp. 797-807.

Goodchild, M. F., G. Q. Sun, and S. Yang. 1992. Development and test of an error model for categorical data. *International Journal of Geographical Information Systems*, 6(2), pp. 87–104.

Green, K. and Congalton, R.G. 2004. An error matrix approach to fuzzy accuracy assessment: The NIMA Geocover Project. In: *Geospatial Data Accuracy Assessment*. Edited by Lunetta, R.L. and Lyons, J.G. U.S. Environmental Protection Agency. Las Vegas. Report No. EPA/600/R-03/064, pp. 163-172.

Grenier, M., Demers, A.-M., Labrecque, S., Benoit, M., Fournier, R.A., and Drolet, B. 2007. An object-based method to map wetland using RADARSAT-1 and Landsat-ETM images: test case on two sites in Quebec, Canada. *Canadian Journal of Remote Sensing,* 33, Supp. 1, pp. S28-S45.

Hammond T.O. and Verbyla, D.L. 1996. Optimistic bias in classification accuracy assessment. *International Journal of Remote Sensing,* 17(6), pp. 1261-1266.

Hurd, J.D., Civco, D.L., Gilmore, M.S., Prisloe, S., and Wilson, E.H. 2006. Tidal wetland classification from Landsat imagery using an integrated pixel-based and object-based classification approach. In: American Society for Photogrammetry and Remote Sensing (ASPRS) 2006 Annual Conference. Reno, 1-5 May 2006, 11 p.

Jensen, J. R. 2005. *Introductory digital image processing: a remote sensing perspective*. John R. Jensen. 3rd ed., Pearson Education Inc., Upper Saddle River.

Kyriakidis, P.C., and Dungan, J.L. 2001. A geostatistical approach for mapping thematic classification accuracy and evaluating the impact of inaccurate spatial data on ecological model prediction. *Environmental and Ecological Statistics*, 8(4), pp. 311-330

Morisette, J. T., Privette, J.L., Strahler, Q., Mayaux, P. and Justice, C.P. 2004. Validation of global land-cover products by the Committee on Earth Observing Satellites. In: *Geospatial Data Accuracy Assessment*. Edited by Lunetta, R.L. and Lyons, J.G. U.S. Environmental Protection Agency. Las Vegas. Report No. EPA/600/R-03/064, pp. 31-40.

Mowrer, H.T. 2000. Quantifying spatial uncertainty in natural resources: theory and applications for GIS and remote sensing. Edited by H. T. Mowrer and R. G. Congalton. In: *Second International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*. Sleeping Bear Press, Ann Arbor Press, USA.

National Wetlands Working Group. 1987. *The Canadian Wetland Classification System*, Provisional Edition. Ecological Land Classification Series, No. 21. Canadian Wildlife Service, Environment Canada, Ottawa.

Ozesmi, S. L. and Bauer, M.E. 2002. Satellite remote sensing of wetlands. *Wetlands Ecology and Management*, 10(5), pp. 381-402.

Schöpfer, E. and Lang, S. 2006. Object fate analysis – A virtual overlay method for the categorisation of object transition and object-based accuracy assessment. In: *Proceedings of the 1st International Conference on Object Based Image Analysis* (OBIA 2006), Bridging Remote Sensing and GIS, Salsburg, Austria.

Song, M., Civco, D.L., and Hurd, D. 2005. A competitive pixel-object approach for land cover classification. *International Journal of Remote Sensing.* 26(22), pp. 4981 – 4997.

Story, M. and Congalton, R.G. 1986. Accuracy assessment: A user's perspective. *Protogrammetric Engineering and Remote Sensing*, 52(3), pp. 397-399.

Thalheimer, W., & Cook, S. (2002, august). "How to calculate effect sizes from published research articles : A simplified methodology". http://work-learning.com/effect_sizes.htm. (accessed December 4th, 2007)

Tiede, D., Lang, S., and Hoffmann, C. 2006. Supervised and forest type-specific multi-scale segmentation for a one-level-representation of single trees. In: *Proceedings of the 1st International Conference on Object Based Image Analysis* (OBIA 2006), Bridging Remote Sensing and GIS, Salsburg, Austria.

USGS. 1994. Accuracy assessment procedures NBS/ NPS Vegetation Mapping Program. USGS Center for Biological Informatics, P.O. Box 25046 Denver, CO 80225, http://biology.usgs.gov/npsveg/index.html

Verbyla, D.L. and Hammond, T.O. 1995. Conservative bias in classification accuracy assessment due to pixel-by-pixel comparison of classified images with reference grids. *International Journal of Remote Sensing*, 16(3), pp. 581-587.

Wang, M. and Howart, P.J. 1993. Modeling errors in remote sensing image classification. *Remote Sensing of Environment*, 45(3), pp. 261-271.

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

Zhan, Q., Molenaar, M., Tempfli, K. and Shi, W. 2005. Quality assessment for geo-spatial objects derived from remotely sensed data. *International Journal of Remote Sensing,* 26(14), pp. 2953-2974.

Zhu, A.X. 1997 Measuring uncertainty in class assignment for natural resource map under fuzzy logic. *Photogrammetric Engineering and Remote Sensing,* 63(10), pp. 1195-1202.