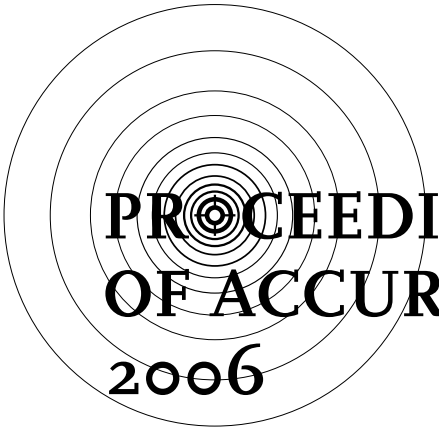

PROCEEDINGS OF ACCURACY 2006

PROCEEDINGS OF ACCURACY 2006

7th International Symposium
on Spatial Accuracy Assessment
in Natural Resources and Environmental Sciences

Editors
Mário Caetano
Instituto Geográfico Português
Marco Painho
Instituto Superior de Estatística e Gestão de Informação - UNL

Published by

Lisboa - Portugal - 2006

Plenary lectures

7th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences.
Edited by M. Caetano and M. Painho.

A methodology for translating positional error into measures of attribute error, and combining the two error sources

Yohay Carmel[1], Curtis Flather[2] and Denis Dean[3]

1 Faculty of Civil and Environmental Engineering
The Technion, Haifa 32000, Israel
Tel.: + 972 4 829 2609; Fax: + 972 822 8898
yohay@technion.ac.il

2 USDA, Forest Service
Rocky Mountain Research Station
2150 Centre Ave, Bldg A
Fort Collins, CO 80526-1891
Tel.: +001 970 295 5910; Fax: + 001 970 295 5959
cflather@fs.fed.us

3 Remote Sensing/GIS Program
Colorado State University
113 Forestry Building
Fort Collins, CO 804523-1472
Tel.: +001 970 491 2378; Fax: +001 970 491 6754
denis@cnr.colostate.edu

Abstract

This paper summarizes our efforts to investigate the nature, behavior, and implications of positional error and attribute error in spatiotemporal datasets. Estimating the combined influence of these errors on map analysis has been hindered by the fact that these two error types are traditionally expressed in different units (distance units, and categorical units, respectively. We devised a conceptual model that enables the translation of positional error into terms of thematic error, allowing a simultaneous assessment of the impact of positional error on thematic error – a property that is particularly useful in the case of change detection. Linear algebra-based error model combines these two error types into a single measure of overall thematic accuracy. This model was tested in a series of simulations using artificial maps, in which complex error patterns and interactions between the two error types, were introduced. The model accommodated most of these complexities, but interaction between the two error types was found to violate model assumptions, and reduced its performance. A systematic study of the spatiotemporal structure of error in actual datasets was thus conducted. Only weak and insignificant interactions were found between the two error types. Application of this error model to real-world time series data indicated that such data are much less accurate than is typically thought. These results highlight the importance of reducing positional error. The second part of our paper presents an analysis of how to reduce the impacts of positional error through aggregation (i.e., increasing the observation grain). Aggregation involves information loss, and thus, the choice of a proper cell size for aggregation is important. A model was developed to quantify the decay in impact of positional error, with increasing cell size. Applying the model to actual data sets, a major reduction in effective positional error was found for cell sizes ranging between 3-10 times the average positional error (RMSE). The model may help users in deciding on an optimal aggregation level given the tradeoff between information loss and accuracy gains.

Keywords: positional accuracy, attribute accuracy, thematic accuracy, post-classification change detection, Combined Location-Classification model

1 Introduction

Anthropogenic processes are modifying land cover at unprecedented rates. Change detection is a critical step in describing and analyzing these processes at regional and local scales. The accuracy of change detection estimates is a complex issue, involving both spatial and temporal dimensions (Cherrill and McClean 1995; Dai and Khorram 1998; Townshend et al. 1992). This paper investigates various aspects of thematic accuracy in spatiotemporal data (= post-classification change detection). Two typical questions in the context of land cover / land use change estimates are (1) what is the probability that a reported land cover / land use transition, inferred from a spatiotemporal dataset, indeed occurred, and (2) how can one increase the reliability of spatiotemporal datasets.

The first question was addressed by developing an error model capable of estimating the accuracy of specific transitions, as well as of the accuracy of the entire spatiotemporal dataset. This model will be presented in section 2. The second question can be answered by evaluating means to increase the reliability of spatiotemporal datasets. Section 3 focuses on one such technique, namely, cell aggregation. Aggregation reduces uncertainty in spatiotemporal datasets, but it results in information loss. Thus, gains in accuracy should be weighted against losses in information content, in order to decide on an optimal aggregation level for a particular application. We will present a simple tool for estimating accuracy gains for a given aggregation level, and exemplify the use of this tool in an actual case study. More details on all of the studies summarized here are provided in several publications, available at http://envgis.technion.ac.il/.

2 A conceptual model for combining location and classification errors

2.1 Overview

A major impediment to estimating accuracy of change detection in thematic datasets is the existence of two very different sources of error in such data. The first source is positional error (also termed location error or misregistration). In the context of spatiotemporal datasets, positional error is known to have a strong impact on change estimates, since a shift of objects in space due to positional error may be interpreted as a change over time. The second source of error is classification uncertainty (termed also thematic or attribute uncertainty). In the case of temporal change, this thematic uncertainty is a combination of possible error in each one of the time steps. These two error types are fundamentally different—being measured by different methods and quantified with different units. Positional error is typically quantified in terms of root-mean-square error (RMSE), which expresses the average Euclidean distance between a point on the map and its actual location in the field. Its units are distance. Classification error is measured by enumerating cases where actual and observed land cover classes differ, summarized in an accuracy matrix.

Recent models describe the impact of these errors on spatiotemporal data accuracy (Liu and Zhou 2004; Stow 1999; Stow and Chen 2002; Van Oort 2005; Wang and Ellis 2005). However, these models consider either one error source or the other, but not both. Reporting these errors separately makes it difficult for the user to grasp the overall level of error present in a data set. Combining these two error metrics into a single measure of overall uncertainty is challenging, especially given the different units of measure used for these error types. The model presented combines both location and classification errors, as well as the interaction

between them, to produce an overall estimate of the error in a database. We will refer to this as the Combined Location and Classification (CLC) error model.

2.2 Description of the model

The CLC model is implemented via a five-step process:

1. **A standard classification error matrix is constructed** for each classified image in a spatiotemporal data set, and the location error of each image is estimated (using a standard evaluation of control points that can be identified on each image).
2. **The impact of location error on each individual classified image is then estimated.** This is accomplished by shifting each raster cell by *x* pixels horizontally and *y* pixels vertically, where *x* and *y* represent the horizontal and vertical components of the RMS of location error in the individual image. The resulting map is compared to the original on a pixel by pixel basis. Pixels whose original classifications differ from their predicted classifications are flagged as suffering from locational-origin attribute errors. A location error matrix, identical in form to a conventional error matrix is then constructed to describe these errors.
3. **The standard classification and the location error matrices described above are combined to yield an overall error matrix (hereafter, the CLC error matrix)**. Location error is assumed to precede classification error, thus the outcome of the latter is conditioned on the former (assuming that images are first rectified and then classified).

These matrix combination procedures trace the fate of pixels when location error and then classification error are introduced into a map. Assume $\mathbf{A}^{LOC}$, $\mathbf{A}^{CLASS}$, and $\mathbf{A}^{CLC}$ are the accuracy matrices for misregistration only, classification error only, and both error sources combined, respectively. Matrix cells contain pixel counts, where *i* indicates the row (representing the observed image) and *j* indicates the column (representing the actual image). For example, the count *n* in cell *i,j* of $\mathbf{A}^{LOC}$, represents the number of class *j* pixels that were assigned to class *i* under misregistration. The total number of actual class *j* pixels is denoted by n^{+j}.

An interpretation of $n_{i,j}$ in $\mathbf{A}^{CLC}$ is more complicated. Consider $n_{1,2}$ in $\mathbf{A}^{CLC}$ – it represents the number of class 2 pixels that were assigned to class 1 due to the combined effect of both errors. This may result from any one of three scenarios:

- A class 2 pixel is assigned to class 1 due to misregistration ($\mathbf{A}_{1,2}^{LOC}$), thereafter correctly classified to class 1 ($\mathbf{A}_{1,1}^{CLASS}$). The number of pixels that would be subject to this process is therefore $\mathbf{A}_{1,2}^{LOC} \times \mathbf{A}_{1,1}^{CLASS} \Big/ \mathbf{A}_{n^{+1}}^{CLASS}$.
- Class 2 pixel is assigned correctly to class 2 in spite of misregistration ($\mathbf{A}_{2,2}^{LOC}$), thereafter assigned to class 1 due to classification error ($\mathbf{A}_{1,2}^{CLASS}$). The respective term for this event is $\mathbf{A}_{2,2}^{LOC} \times \mathbf{A}_{1,2}^{CLASS} \Big/ \mathbf{A}_{n^{+2}}^{CLASS}$.

- Class 2 pixel is assigned to land cover class L due to location error ($A^{LOC}_{L,2}$), thereafter assigned to class 1 due to classification error ($A^{CLASS}_{1,L}$). This event is denoted by $A^{LOC}_{L,2} \times A^{CLASS}_{1,L} \,/\, A^{CLASS}_{n+L}$.

Thus, The value for ($A^{CLC}_{1,2}$) is denoted by:

$$A^{CLC}_{1,2} = \frac{A^{LOC}_{1,2} \times A^{CLASS}_{1,1}}{A^{CLASS}_{n+1}} + \frac{A^{LOC}_{2,2} \times A^{CLASS}_{1,2}}{A^{CLASS}_{n+2}} + \ldots + \frac{A^{LOC}_{L,2} \times A^{CLASS}_{1,L}}{A^{CLASS}_{n+L}} \quad (1)$$

Values in the remainder of the cells in the combined A^{CLC} matrix are calculated similarly (equations for calculating cells in the three matrices are given in Table 1 in Carmel et al 2001a).

1. **User accuracy for each CLC error matrix is calculated.** The error combination process described in step (3) results in T CLC error matrices, one for each of the T time periods in the multi-temporal data set. For any individual time period t, the probability that a pixel assigned to land cover class L actually belongs to that class is denoted by p(Lt). (Congalton and Green 1999) term this probability user accuracy. These user accuracies can be computed easily via column summaries of the CLC error matrices.
2. **Finally, errors are accumulated across time periods by computing transition probabilities, and the probability that a specific sequence of reported transitions indeed occurred is calculated.** A transition probability is the probability that an observed transition (or sequence of transitions) over the T time periods involved in a multi-temporal data set indeed occurred.

Thus, the transition probability that a cell that remained in land cover category L throughout all T time periods was correctly classified, is given by:

$$p(L_1 L_2 \ldots L_T) = p(L_1) \cap p(L_2) \cap \ldots \cap p(L_T) \quad (2)$$

Assuming independence of error between time steps, equation (1) becomes:

$$p(L_1 L_2 \ldots L_T) = p(L_1) \bullet p(L_2) \bullet \ldots \bullet p(L_T) \quad (3)$$

Equation 3 indicates that transition probabilities can be calculated from user accuracies. User accuracies may be calculated using steps (1) through (4) above, and thus the CLC model can be used to calculate all possible transition probabilities. Once these transition probabilities are available, multi-temporal cumulative error indices may be derived (Carmel et al. 2001a).

2.3 Model application: California woodland dynamics

In order to illustrate the use of the model, we applied it to a case study in which the dynamics of oak woodlands are studied over a period of 56 years. Full details of the image analyses, and on the ecological analyses, are provided in Carmel et al. (2001a), and Carmel and Flather (2004), respectively. Four aerial photos of the Hastings Natural History Reservation and surroundings (Monterey County, California), taken in 1939, 1956, 1971, and 1995, were orthorectified. Location error for each of the four orthophotos was assessed using 40 points located across the scene (each of which could be identified in all photos). For a given point in a

specific image, location error was defined as the Euclidean distance between point location on the image and the average of locations of that point in all four images. These individual point location error estimates were summarized by calculating RMSE for each orthophoto.

The classification scheme applied to the images included three distinct vegetation classes: oak woodland, chaparral, and grassland. Ground truthing was used to assess classification accuracy for the 1995 image, while a stereoscope-aided manual photo interpretation was used for the older images. In order to construct location accuracy matrices, a shifted image was constructed for each image in the database, in which the shift along the *x* and *y* axes corresponded to the *x* and *y* RMSE values. The shifted images were compared to their precursors on a pixel by pixel basis, thus deriving the location accuracy matrices. The combined location-classification accuracy matrix was then calculated for each time period.

Thematic accuracy for a multi-temporal database is usually described for two time steps, and its extension to *n* time steps deserves a brief discussion. Merging the various time-specific matrices using cross-tabulation was used for describing two images with two classes (Congalton and Macleod 1994). This method has the advantage of presenting the entire accuracy information for the entire dataset. Yet, when more than two time steps are involved, the combined matrix becomes large. For example, the Hastings dataset accuracy assessment resulted in four matrices with 3 classes. Cross tabulation of these matrices would result in an 81 X 81 matrix. The interpretation of such matrices becomes difficult. Alternatively, the time-specific accuracy matrices may serve to construct other, more concise indices. We suggest three such indices (ordered here in increasing degree of compaction:

The transition-specific probability, $p(L_{1939}L_{1956}L_{1971}L_{1995})$ is the probability that an observed transition is correct for all time steps. For example, the transition CGCC represents a pixel that was classified as chaparral (C) in 1939, changed to grassland (G) in 1956 (probably due to a fire event), changed to chaparral again in 1971, and remained classified as chaparral in 1995. p(CGCC) is the probability that this set of transitions indeed occurred. These probabilities were calculated as the products of the respective 'user-accuracy' values for the relevant class in each time step.

The class-specific probability is the probability of any transition involving a specific map class i to be correct. It is calculated as the average of all the probabilities associated with transitions involving that specific class in at least one time step.

The spatiotemporal proportion classified correctly (PCC) is the probability that the assigned transition of any given pixel in the scene is correct. It can be calculated as the product of all PCC's derived from each time-step accuracy matrix.

The classification accuracy for all images was relatively high; PCCs for the classification accuracy matrices were in the range of 0.90 – 0.94 (Table 1). Location accuracy was also relatively high (RMSE was 3.53 m for 1939, 1.97 m for 1956, 2.42 m for 1971 and 2.51 m for 1995). However, location error had a large effect on thematic accuracy (presumably, due to the small pixel size, 0.6 m, relative to the RMSE, and due to the heterogeneous nature of the classified images). PCCs for the location accuracy matrices were consistently lower than those for the classification error matrices, in the range of 0.62 – 0.80 (Table 1). As a result, the combined accuracy PCCs were in the range of 0.58 – 0.74, being (as expected) lower than both their constituents (Table 1).

Table 1 PCC values for Hastings vegetation maps.

	Location	Classification	Combined
1939 PCC	0.8	0.9	0.74
1956 PCC	0.62	0.88	0.58
1971 PCC	0.76	0.94	0.72
1995 PCC	0.77	0.91	0.72

In order to describe the overall uncertainty in the dataset, we present three example indices (Table 2). First, we look at pixels that apparently remained as grassland throughout the 65-year period (43% of 1939 grassland area). For any single pixel, the probability that this transition sequence is correct is only 0.21. Similarly, the combined probability of a transition sequence involving forest to be correct is 0.2. The spatiotemporal PCC, which is the probability of any given pixel in this dataset being assigned to the correct class in all four time periods is 0.22. Note that in all cases, the contribution of location error to the overall uncertainty is much larger than the contribution of classification error. This indicates that for the Hastings dataset, reducing the impact of positional error is a prerequisite, before any meaningful analysis can take place. One way to increase positional accuracy is by reducing resolution. Quantification of the gain in accuracy for a given decrease in resolution is the subject of part 3 of this paper.

Table 2 Indices for the accuracy of the Hastings multi-temporal dataset. a. Transition specific accuracy, the probability of a specific set of transitions to be correct. Three examples of possible transition sets are presented. b. Class specific accuracy, the probability for a transition involving a particular class to be correct, averaged for all relevant transitions. c. Spatiotemporal PCC, the probability for any given pixel, that its assigned set of transitions is correct, averaged for all pixels in the scene.

	Transition type	Nature of transition	Proportion[a]	Probability of being correct, given:		
				location error	Classification error	Combined error
a	GGGG	Grassland does not change	0.43	0.28	0.69	0.21
	CGGG	Chaparral burnt in 1955 fire	0.06	0.29	0.60	0.22
	GTTT or GGTT or GGGT	Grassland turns into forest at some point	0.37	0.31	0.69	0.25
b	All 69 transitions involving forest		0.66	0.26	0.59	0.20
c	Averaged across the entire study area		1	0.29	0.67	0.22

a Proportion of 1939 class undergoing transition.

2.4 Model validation

2.4.1 Overview

The CLC error model is based on two key assumptions: (1) errors in each time step are independent of errors in other time steps, and (2) location and classification errors are independent of one another. Unfortunately, one could imagine scenarios where these assumptions would be violated. For example, when a single DEM is used to rectify all of the remotely sensed images in a multi-temporal data base, location errors might be correlated across time steps. Furthermore, steep terrain may affect both classification and location accuracy, and thus some correlation between location and classification error may occur. A systematic process of model validation is therefore necessary in order to evaluate the CLC error model, and in particular its adequacy under complex error patterns.

Validation of an error model requires full knowledge of the spatial pattern of both error sources, and a large number of datasets. Typically, field data do not support this kind of analysis since error magnitude and pattern can not be controlled. Artificially generated datasets were successfully used in previous analyses of error propagation in spatial data (Goodchild et al. 1992; Griffith et al. 1999; Haining and Arbia 1993; Mowrer 1994; Veregin 1995). The basic approach adopted for this study is based on the evaluation of numerous artificial spatiotemporal datasets, where the magnitude of error, as well as its spatiotemporal structure are controlled.

2.4.2 Simulating multi-temporal datasets with controlled error characteristic

The process of building an artificial multi-temporal data set with controlled, non-random error pattern is not trivial. A brief description of our algorithm follows (full details are available in Carmel and Dean 2004). We used ArcInfo's Grid module and Arc Macro Language (AML, ESRI 2001) to construct a 512 X 512 thematic raster map, with each cell randomly assigned to one of 3 possible classes, according to pre-determined class proportions (Figure1a-b). This map was viewed as the "true" thematic map to which the remainder of the process would assign known amounts of error. Next, two maps were generated to represent the cell-specific probability of suffering from a classification error and from a location error, respectively (Figure 1c,e). The probability distribution used to populate these maps, and the degree of spatial autocorrelation among the probabilities within each map were controlled by the user. The error maps (Figure 1d,f) were derived from these precursors using a probability function with user-controlled parameters. The algorithm was applied twice, to represent multi-temporal data set layers from two points in time. The true map of time b is derived from the true map of time a, controlling for both the rate of change and its pattern. The program induced correlation between error in the two time steps, and controlled for its magnitude.

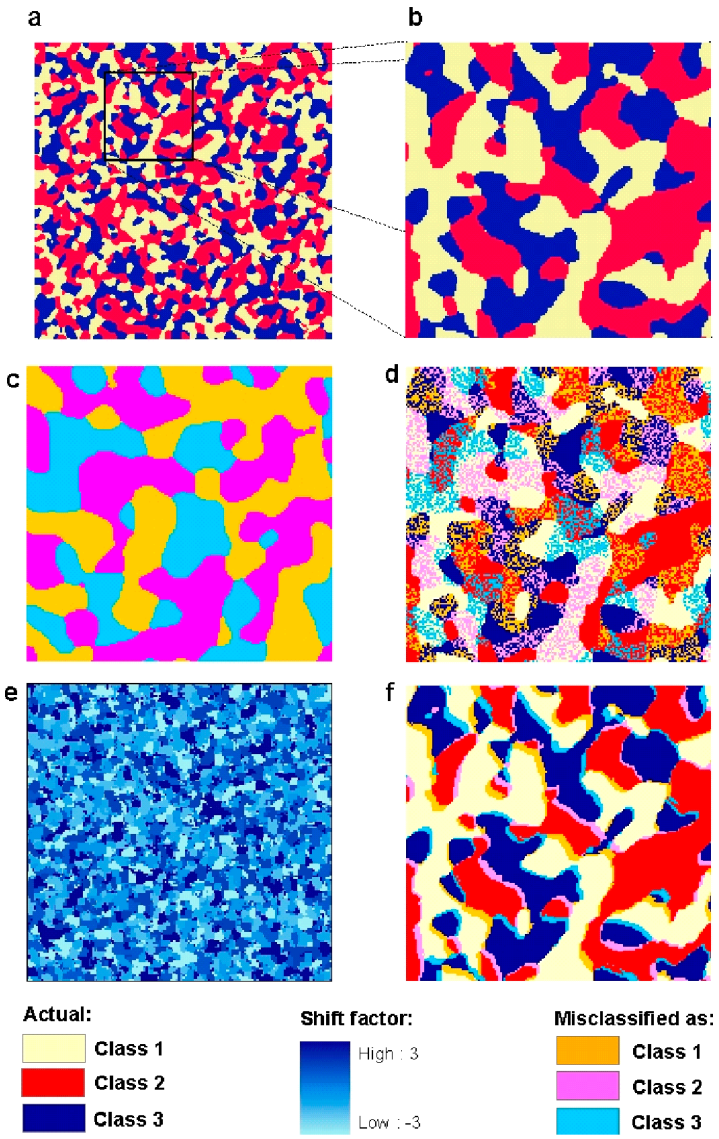

Figure 1 Examples of the steps of the simulations. (a) The 'actual' map for time a. (b) A small part of the map of time a, enlarged. (c) A classification error precursor map, used to construct the 'classified map', that contains a controlled degree of classification error. (d) Classification error map, resulting from overlaying a user-determined proportion of the error-map (depicted in 2c) over the 'actual' map (depicted in 2b). Pixels classified correctly are shown in their original colors (white, blue and red, for classes 1, 2, and 3, respectively), while pixels classified incorrectly are shown in orange, pink, and cyan, respectively. (e) location error precursor map, used to construct the 'shifted map', that contains a controlled degree of location error. (f) Location error map, resulting from overlaying a user-determined proportion of the location error pattern (depicted in 2c) over the 'actual' map (depicted in 2b).

2.4.3 Validation scheme and statistical analysis

The robustness of the CLC error model was tested in a set of 250 simulation runs, where the following parameters were varied:

- Proportion of each class (0.01-0.99).
- Number of map classes (2-4 classes).
- Location error magnitude (from -3 to +3 pixels on both the x and y directions).
- Classification error magnitude (PCC from 0.5 to 0.99).
- Spatial autocorrelation within each type of error (Moran *I* between 0.05 and 0.95).
- Correlation in error between time steps (Pearson coefficient between 0 and 0.8).
- Spatial correlation between the two error types (Pearson coefficient from 0 to 0.6).

Each simulated multi-temporal data set was evaluated against an analytical solution counterpart. Location, classification, and combined error matrices were constructed using model equations. Values in these error matrices and the CLC model were used to compute the model-prediction of the probability of each transition type *p(C1C2…CT)* to be correct. As a reference, the respective 'observed' (in simulation) probabilities were computed by comparing the final error-laden maps to the 'true' maps. We defined D as an index of deviance between the observed (in simulation) and predicted (by the model) transition probabilities:

$$D = \left| p(C_1 C_2 ... C_T)^{OBSERVED} - p(C_1 C_2 ... C_T)^{PREDICTED} \right| \quad (4)$$

For each simulation run we calculated the average deviance (D_{avg}) and maximum deviance (D_{max}) across all transition types. In the present study, three classes and two time steps yielded nine transition types for each simulation, thus D_{avg} was calculated as the average of nine values and D_{max} was the maximum of nine values. Multiple regression analyses were used to assess the impact of correlation between error sources and of error rate on model performance. D_{avg} was the dependent variable, and the correlation between the two error sources in a single time step, correlation between error in different time steps, and error rate were the independent (predictor) variables.

2.4.4 Simulation results

Good agreement between model predictions and simulation results was found **in the absence of any correlation in error patterns** (D_{avg} did not exceed 0.002, and D_{max} did not exceed 0.01). These results were consistent for different number of classes, for various proportions of each class, and for both low and high levels of error of each source. Under these conditions, the model predicted correctly the combined effects of location and classification errors. Differences between model predictions and observed results reflected stochasticity in the simulation process.

Introducing **correlation between error pattern in time *a* and time *b*** had a small effect on the agreement between model predictions and simulations, which was still very high. Even when the correlation coefficient between attribute errors in the two time steps was as high as 0.8, D_{avg} never exceeded 0.003, and D_{max} did not exceed 0.012. In contrast, **correlation between the two sources of error** affected the fit between model predictions and simulation results significantly. This deviation increased with the rate of correlation between error sources. It was also affected by both location error rate and classification error rate. For equal values of correlation between error types, larger error rates in the simulation resulted in larger differences between model predictions and simulation results (note that in the absence of such

a correlation, model fit was not affected by error rate). The largest reduction in model fit occurred when the largest errors from both sources (PCC~0.4) coincided with strong correlation between these sources (~0.6). In those cases D_{avg} reached 0.015, while D_{max} reached 0.056. The regression analyses revealed a significant impact of correlation between error types on model fit (with D_{avg} as the dependent variable). In a univariate analysis, with correlation between location error and classification error as the sole predictor, a strong effect on the dependent variable was obvious ($R^2 = 0.77$, $p < 0.001$). When error rates of both sources for both time steps were added to the regression model in a stepwise forward procedure, all variables had a significant effect on the dependent variable (Adjusted $R^2 = 0.91$, $p < 0.001$).

2.5 Conclusions regarding the CLC approach

Location error between various layers in a GIS dataset may seriously affect the quality of any spatial analysis performed using these layers (Foody 2003; Stow and Chen 2002; Wang and Ellis 2005). This is particularly true in the case of change detection, where a time-series dataset is analyzed. Our California woodlands data illustrates the vulnerability of such data to location errors. The CLC approach described here quantifies location accuracy in terms of thematic accuracy. The CLC error model combines the location and classification accuracy matrices into a single matrix that represents the overall thematic accuracy for a single layer. The interactions between both error sources are explicitly quantified in our model.

The validation process revealed that model estimates were found to be highly accurate under a wide range of conditions. Differences between model predictions and observed results seemed to stem only from the stochastic nature of the simulation process. Two exceptions are the case of correlation between error in different time steps and between the two sources of error. When strong correlation between errors in different times was induced, there was some decrease in the accuracy of model predictions. Yet, even for unrealistically high correlation values, the magnitude of the problems caused by this type of correlation was insignificant. For weak correlations between location and classification error within an individual data layer (values < 0.2), model performance was practically unaffected. In the case of the strongest possible correlation between error types, the actual error estimate would differ by ±1.5% of the true value over the entire dataset, while for specific transition types, this difference may go up to ±5% of the actual value. A recent study (Carmel 2004a) found that correlations between location and classification errors in actual data were low (<0.17) and insignificant in all five real-world data sets studied. This implies that in most actual situations, the CLC model is unlikely to suffer badly from correlations of this sort.

The CLC error model produces time-specific matrices that can then serve to derive useful indices for estimating the overall uncertainty in a multi-temporal dataset. We suggest three such indices: the transition-specific probability, the class-specific probability, and the spatiotemporal PCC. Indices that account for the chance agreement between maps (e.g. kappa, Congalton 1991) may be derived from the time-specific combined matrices. Pontius (2000) suggests a specific derivation of kappa for maps with known location error. The extension of such indices for multi-temporal datasets is yet to be explored.

Application of the CLC model to real world datasets revealed that the reliability of observed transition in any specific location may be much lower than is usually assumed. Heterogeneous landscapes, and positional error that is large compare to pixel size, would act to worsen the problem, and reported change may be nearly meaningless. Yet, in the context of larger spatial units, the same report may be trustworthy. This seemingly contradiction may be explained by

the fact that positional error attenuates quickly when the grain of the basic spatial units increases. A quantitative estimation of this process is the topic of the next part of this paper.

3 Aggregation and accuracy

Several studies have suggested that reduction of spatial resolution enhances thematic map accuracy significantly (Dai and Khorram 1998; Townshend et al. 2000; Townshend et al. 1992). This is mainly due to a decrease in the impact of positional error, as explained above. On the other hand, the decrease in spatial resolution involves a loss of information that may be valuable for particular applications (Carmel et al. 2001b). A plot of thematic map accuracy as a function of spatial resolution would allow users to choose the specific spatial resolution and its associated uncertainty level that best fits the needs of their specific application. The goal of this part is to explore the relationship between spatial resolution and thematic map accuracy, and to develop a model that quantifies this relationship for thematic ('classified') maps.

3.1 Model description

For a single pixel, misregistration is translated into thematic error if its 'true' location is occupied by a pixel belonging to a different class. Let us define *p(loc)*, the probability that a pixel is assigned an incorrect class due to misregistration, as:

$$p(loc)_{rc} = p(i_{rc} \neq i_{r+e(x),c+e(y)}) \tag{5}$$

where *r* and *c* are pixel coordinates, i_{rc} is the class assigned to the pixel, *e(x)* and *e(y)* are the *x* and *y* components of location error, respectively (measured in pixels). Thus, $p(loc)_{rc}$ depends on the magnitude of location error, and on landscape fragmentation (map heterogeneity). *p(loc)* can be estimated empirically for a given map, based on map pattern and the magnitude of location error. Location error may not be uniform across an image (Fisher 1998). Thus, the general model scheme presented here does not assume a constant location error, but allows error to vary across the image. Location error is assumed, however, to be constant within each grid cell (the unit which the pixels are aggregated into).

Considering a larger cell size *A*, let us define a similar probability, $p^A(loc)$, which is the probability that a pixel within the framework of a larger cell was misclassified due to misregistration. For cell sizes larger than location error, this probability would be lower than the original probability *p(loc)*. This is due to the fact that misregistration would shift a certain proportion of the pixels only within the grid cell, and for those pixels, thematic error is cancelled at the grid cell level (Figure 2). Here, a conceptual model is presented, in which this probability is denoted by:

$$p^A(loc) = \alpha \cdot p(loc) \tag{6}$$

where α is the proportion of a cell of size *A* in which pixels are misplaced into neighboring cells, and may thus result in thematic error (Figure 2).

This proportion, α, is termed here the effective location error. It is a function of cell size A and of location error magnitude. The effective location error α is the proportional area of the dark gray region in the cell (Figure 2) and is denoted by:

$$\alpha = \frac{A \cdot e(x) + A \cdot e(y) - e(x) \cdot e(y)}{A^2} \quad (7)$$

where location error components e(x) and e(y) are assumed constant for all pixels within a single cell. Using Equation 6, the reduction in effective location error when cell size increases may be illustrated easily (here, for the special case where e(x)=e(y), Figure 3). Effective location error α declines rapidly from 1 for cell sizes ≤ the magnitude of location error, to 0.36 for cell sizes five times the magnitude of location error (Figure 3).

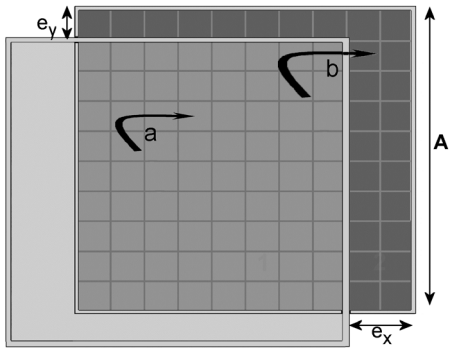

Figure 2 Quantification of the effect of location error on thematic accuracy. An image is 'shifted' against itself. The results are shown here for a single grid cell within the image. Here, cell size A = 10 x 10 pixels. The original image is shown in light gray and the shifted image is shown in dark gray. e_x and e_y denote the x and y components of location error, respectively (in this figure e_x = 2 pixels and e_y = 1 pixel). Pixels in the area of overlap between the two images would remain within the cell and thematic accuracy at the cell level would not be affected (see arrow a). Pixels in the dark gray region are shifted into neighbouring cells, and may result in thematic error (see arrow b). The effective location error α is the proportional area of the dark gray region within the cell (Equation 3).

In order to determine p(loc), α and p^A(loc), one may employ the practice of 'shifting' a map against itself, as discussed above. This is a simulation approach that enables one to estimate the thematic consequences of misregistration between two maps. The process results in two identical maps, where one is spatially 'shifted' off the other by a lag equal to the estimated location error (here, the map is shifted against itself by RMSE(x) and RMSE(y) on the x and y axes, respectively). A typical thematic error analysis can then be applied, to form an 'error matrix'. This matrix contains essentially two types of pixels, (1) those that were not affected by the shift between maps, and (2) those for which the shift resulted in 'thematic error'. p(loc) is derived as the proportion of type (2) pixels in the map. Finally, α and p^A(loc) are calculated for the entire image, using Equations 2 and 3.

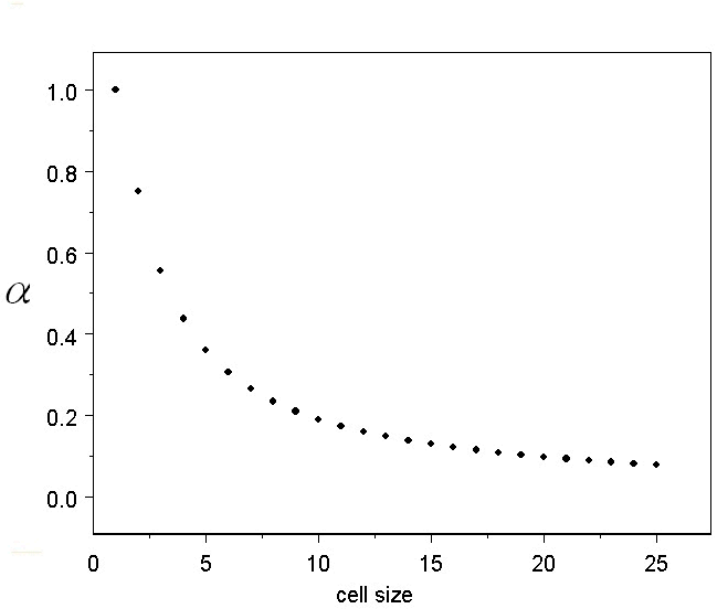

Figure 3 Effective location error α as a function of cell size. Cell size is shown here in multiplies of pixels, where average location error is set to 1 pixel in both x and y directions. Cell size varies between 1 and 25 times the magnitude of location error.

In this procedure, location error was assumed homogeneous. In fact, location error may vary largely across an image, since its major sources, topography and quality of ground-control points, imply inherent spatial pattern). Thus, α, p(loc), and p^A(loc), should ideally be estimated for each grid cell in the map. An alternative approach, that interpolates location error into location error surface (Fisher 1998), is described in (Carmel 2004b). A case study of actual data, in which p^A(loc) is estimated, and plotted against cell size, is provided in Carmel (2005).

3.2 Conclusions

The results of this study support previous indications that the impact of misregistration on map accuracy can be large (Dai and Khorram 1998; Stow 1999; Townshend et al. 1992), and reveal that aggregation is a very effective means of reducing this impact. This study develops a conceptual model to quantify the effect of aggregation on map accuracy. Given the tradeoff between resolution and accuracy, potential utility of this model would be to help the user choose the proper cell size for a specific application. A first approximation of the gain in accuracy with increased aggregation level can be visualized using a simple procedure: solve Equation 3 for a range of relevant cell values, and portray α (the effective location error) as a function of cell size (Figure 3). This procedure is particularly easy to apply if RMSE is taken to represent e. Further information can be gained by estimating the actual probability of error, p^A(loc), for various aggregation levels. This stage requires spatially-explicit simulations that manipulate the actual map.

In conclusion, this methodology provides an effective tool for assessing the impact of aggregation on thematic map accuracy, and evaluating it against information loss, in order to decide on a proper level of map aggregation. The most effective reduction in error was achieved when cell-size was in the range of 3-10 times the size of average location error. Map-specific error rates may somewhat alter this conclusion.

References

Carmel, Y., 2004a. Characterizing Location and Classification Error Patterns in Time-Series Thematic Maps. *Ieee Geoscience and Remote Sensing Letters* 1, 11-14.

Carmel, Y., 2004b. Controlling Data Uncertainty via Aggregation in Remotely Sensed Data. *Ieee Geoscience and Remote Sensing Letters* 1, 39-41.

Carmel, Y., 2005. Aggregation as a Means of Increasing Thematic Map Accuracy, In *Geodynamics*. eds P. Atkinson, D. Martin, G. Foody, M. CRC Press, London.

Carmel, Y., Dean, D.J., 2004. Performance of a spatio-temporal error model for raster datasets under complex error patterns. *International Journal of Remote Sensing* 25, 5283-5296.

Carmel, Y., Dean, D.J., Flather, C.H., 2001a. Combining location and classification error sources for estimating multi-temporal database accuracy. *Photogrammetric Engineering & Remote Sensing* 67, 865-872.

Carmel, Y., Flather, C.H., 2004. Comparing landscape scale vegetation dynamics following recent disturbance in climatically similar sites in California and the Mediterranean basin. *Landscape Ecology* 19, 573-590.

Carmel, Y., Kadmon, R., Nirel, R., 2001b. Spatiotemporal predictive models of Mediterranean vegetation dynamics. *Ecological Applications* 11, 268-280.

Cherrill, A., McClean, C., 1995. An investigation of uncertainty in field habitat mapping and the implications for detecting land cover change. *Landscape Ecology* 10, 5-21.

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., Green, K., 1999. *Assessing the accuracy of remotely sensed data: principles and practices*. Lewis Publishers, New York.

Congalton, R.G., Macleod, R., 1994. Change detection accuracy assessment on the NOAA Chesapeake Bay pilot study, In *Proceedings of the International Symposium on Spatial Accuracy of Natural Resource Data Bases: unlocking the puzzle*. ed. R.G. Congalton, pp. 78-87. American Society for Photogrammetry and Remote Sensing, Bethesda, Maryland.

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

ESRI, 2001. *ArcGIS Users Guide*. Environmental Systems Research Institute, Redlands, California.

Fisher, P., 1998. Improved modeling of elevation error with geostatistics. *GeoInformatica* 2, 215-233.

Foody, G.M., 2003. Uncertainty, knowledge discovery and data mining in GIS. *Progress in Physical Geography* 27, 113-121.

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

Griffith, D.A., Haining, R.P., Arbia, G., 1999. Uncertainty and error propagation in map analyses involving arithmentic and overlay operation: inverntory and prospects, In *Spatial accuracy assessment: land information uncertainty in natural resources*. eds K. Lowell, A. Jaton, pp. 11-25. Ann Arbor Press, Chelsea, Michigan.

Haining, R.P., Arbia, G., 1993. Error propagation through map operations. *Technometrics* 35, 293-305.

Liu, H., Zhou, Q., 2004. Accuracy analysis of remote sensing change detection by rule-based rationality evaluation with post-classification comparison. *International Journal of Remote Sensing* 25, 1037-1050.

Mowrer, T.H., 1994. Monte Carlo techniques for propagating uncertainty through simulation models and raster-based GIS, In *Proceedings of the International Symposium on Spatial Accuracy of Natural Resource*

Data Bases: unlocking the puzzle. ed. R.G. Congalton, pp. 179-198. American Society for Photogrammetry and Remote Sensing, Bethesda, Maryland.

Pontius, R.G., 2000. Quantification error versus location error in comparison of categorical maps. *Photogrammetric Engineering & Remote Sensing* In Press.

Stow, D.A., 1999. Reducing the effects of misregistration on pixel-level change detection. *International Journal of Remote Sensing* 20, 2477-2483.

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

Townshend, J.R.G., Huang, C., Kalluri, S., Defreis, R., Liang, S., 2000. Beware of per-pixel characterization of land cover. *International Journal of Remote Sensing* 21, 839-843.

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

Van Oort, P.A.J., 2005. Improving land cover change estimates by accounting for classification errors. *International Journal of Remote Sensing* 26, 3009-3024.

Veregin, H., 1995. Developing and testing of an error propagation model for GIS overlay operations. *International Journal of Geographical Information Systems* 9, 595-619.

Wang, H.Q., Ellis, E.C., 2005. Image misregistration error in change measurements. *Photogrammetric Engineering and Remote Sensing* 71, 1037-1044.