

Performance of a spatio-temporal error model for raster datasets under complex error patterns

Y. CARMEL

Faculty of Civil and Environmental Engineering, Technion—Israel institute of Technology, Haifa 32000, Israel; e-mail: yohay@tx.technion.ac.il

and D. J. DEAN

Department of Forest Sciences, Colorado State University, Fort Collins, CO 80523, USA

(Received 13 February 2003; in final form 10 November 2003)

Abstract. The CLC (Combined Location Classification) error model provides indices for overall data uncertainty in thematic spatio-temporal datasets. It accounts for the two major sources of error in such datasets, location error and classification error. The model assumes independence between error components, while recent studies revealed various degrees of correlation between error components in actual datasets. The goal of this study is to determine if the likely violation of model assumptions biases model predictions. A comprehensive algorithm was devised to simulate the entire process of error formation and propagation. Time series thematic maps were constructed, and modified maps were derived as realizations of underlying error patterns. Error rate and pattern (positive autocorrelation) were controlled for location error and for classification error. The magnitude of correlation between errors from different sources and correlation between error at different time steps was also controlled. A very good agreement between model predictions and simulation results was found in the absence of correlation in error between time steps and between error types, while the inclusion of such correlations was shown to affect model fit slightly. Given our current knowledge of spatio-temporal error patterns in real data, the CLC error model can be used reliably to assess the overall uncertainty in thematic change detection analyses.

1. Introduction

As spatial data produced via change detection analyses become more widespread and more widely used, there is a growing need to understand the error and uncertainties present in these datasets. These measures of uncertainty can be used in subsequent spatial analyses and decision making to estimate reliability of the analyses' products (Stoms *et al.* 1992, Mowrer 1994, 1997, 1999, Antrop 1998, Kyriakidis and Dungan 2001). Therefore, realistic estimates of the degree and nature of uncertainty are becoming essential elements to accompany change detection products (Chrisman 1991).

Error modelling has become a major means to provide spatially-explicit information on the uncertainty of estimated parameters. Such information is

increasingly required by data-users for evaluating the risk that a specific outcome of further analysis of the information will be incorrect (Fisher 1998), or for incorporating the variability of the parameters of interest into ecological and environmental models using stochastic simulations (Phillips and Marks 1996, Mowrer 1997, Kyriakidis and Dungan 2001).

Error in spatio-temporal data has several components, originating from different sources. For example, thematic datasets derived from remotely sensed data contain both location error (sometimes called positional error or misregistration) and classification errors, while radiometric errors play an indirect role only, affecting classification error (Carmel *et al.* 2001). In the case of multi-temporal datasets of the type used in change detection analysis co-registration errors are present as well (Townshend *et al.* 1992, Dai and Khorram 1998, Stow 1999), but in principle, this error is an extension of individual image location errors. The integration of all of these components into an overall measure of uncertainty is not trivial.

Until recently, each of these errors was measured separately. Reporting error in this disjointed fashion makes it difficult for an analyst to grasp the overall level of error present in the dataset. Recently, combined error models have been developed that integrate these various error components into a single measure of uncertainty. Two such models have been proposed for the two major types of change detection: Stow (1999) proposed a combined error model appropriate for change detection based on the image differencing approach, and Carmel *et al.* (2001) presented a model suitable for thematic change detection, based on post-classification analyses. Both models have a common feature, which is the integration of the impact of misregistration on overall uncertainty. The error model developed by Carmel *et al.* (2001) combines both location and classification errors, as well as the interaction between them, to produce an estimate of the overall error in the data. We will refer to this as the CLC (Combined Location and Classification) model.

1.1. The CLC model

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

1. For each classified image in a multi-temporal spatial dataset, a standard classification error matrix is constructed.

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 accuracy matrix, identical in form to a conventional accuracy matrix is then constructed to describe these errors.

3. The standard classification and 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 (this assumes that images are first rectified and then classified). In order to combine the two matrices, we traced the fate of pixels when location error and then classification error are introduced into a map. A^{LOC} is

Spatio-temporal error model performance

the accuracy matrix for misregistration only, A^{CLASS} for classification error only and A^{BOTH} for both error sources combined (see table 1 for A^{BOTH}). 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 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 A^{BOTH} is more complicated. Consider $n_{1,2}$ in A^{BOTH} —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) A class 2 pixel is assigned to class 1 due to misregistration ($A^{LOC}_{1,2}$), thereafter correctly classified to class 1 ($A^{CLASS}_{1,1}$). The number of pixels that would be subject to this process is therefore $A^{LOC}_{1,2}A^{CLASS}_{1,1}/A^{CLASS}_{n_{+1}}$.
- (b) Class 2 pixel is assigned correctly to class 2 in spite of misregistration ($A^{LOC}_{2,2}$), thereafter assigned to class 1 due to classification error ($A^{CLASS}_{1,2}$). The respective term for this event is $A^{LOC}_{2,2}A^{CLASS}_{1,2}/A^{CLASS}_{n_{+2}}$.
- (c) Class 2 pixel is assigned to class 3 due to location error ($A^{LOC}_{3,2}$), thereafter assigned to class 1 due to classification error ($A^{CLASS}_{1,3}$). This event is denoted by $A^{LOC}_{3,2}A^{CLASS}_{1,3}/A^{CLASS}_{n_{+3}}$.

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

$$A^{BOTH}_{1,2} = A^{LOC}_{1,2}A^{CLASS}_{1,1}/A^{CLASS}_{n_{+1}} + A^{LOC}_{2,2}A^{CLASS}_{1,2}/A^{CLASS}_{n_{+2}} + A^{LOC}_{3,2}A^{CLASS}_{1,3}/A^{CLASS}_{n_{+3}}.$$

The expected counts for the remainder of the cells in the combined matrix A^{BOTH} are calculated similarly (table 1 (c)).

4. 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 dataset. For any individual time period t , the probability that a pixel assigned to category C actually belongs to that category is denoted by $p(C_t)$. Congalton and Green (1999) term this probability user accuracy. These user accuracies can be computed easily via column summaries of the CLC error matrices.

5. Finally, errors are accumulated across time periods by computing transition probabilities. A transition probability is simply the probability that a specific sequence of errors (or non-errors) occurred over the T time periods involved in a multi-temporal dataset. Thus, the transition probability that a cell that remained in informational category C throughout all T time periods was correctly classified is given by:

$$p(C_1 C_2 \dots C_T) = p(C_1) \cap p(C_2) \cap \dots \cap p(C_T) \quad (1)$$

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

$$p(C_1 C_2 \dots C_T) = p(C_1) \bullet p(C_2) \bullet \dots \bullet p(C_T) \quad (2)$$

Equation (2) indicates that transition probabilities can be calculated from user accuracies. Given that the CLC model can calculate user accuracies (using steps (1) through (4) above), it is clear that 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.* 2001).

Table 1. A hypothetical matrix for a single map with three map categories. A^{LOC} is the accuracy matrix for misregistration only, A^{CLASS} for classification only and A^{BOTH} for both error sources combined. Matrix cells contain pixel counts, where i indicates the row (representing the observed image) and j indicates the column (representing the actual image). The count n in cell i,j of A^{LOC} represents the number of class j pixels that were assigned to class i under misregistration. Similarly, n_{ij} in A^{CLASS} is the count of class j pixels that were assigned to class i under misclassification. The total number of actual class j pixels is denoted by n_{+j} . After Carmel *et al.* (2001).

Observed (i)	Actual (j)		
	1	2	3
1	$A^{LOC}_{1,1}(A^{CLASS}_{1,1}/n_{+1})$ $+ A^{LOC}_{2,1}(A^{CLASS}_{1,2}/n_{+2})$ $+ A^{LOC}_{3,1}(A^{CLASS}_{1,3}/n_{+3})$	$A^{LOC}_{1,2}(A^{CLASS}_{1,1}/n_{+1})$ $+ A^{LOC}_{2,2}(A^{CLASS}_{1,2}/n_{+2})$ $+ A^{LOC}_{3,2}(A^{CLASS}_{1,3}/n_{+3})$	$A^{LOC}_{1,3}(A^{CLASS}_{1,1}/n_{+1})$ $+ A^{LOC}_{2,3}(A^{CLASS}_{1,2}/n_{+2})$ $+ A^{LOC}_{3,3}(A^{CLASS}_{1,3}/n_{+3})$
2	$A^{LOC}_{1,1}(A^{CLASS}_{2,1}/n_{+1})$ $+ A^{LOC}_{2,1}(A^{CLASS}_{2,2}/n_{+2})$ $+ A^{LOC}_{3,1}(A^{CLASS}_{2,3}/n_{+3})$	$A^{LOC}_{1,2}(A^{CLASS}_{2,1}/n_{+1})$ $+ A^{LOC}_{2,2}(A^{CLASS}_{2,2}/n_{+2})$ $+ A^{LOC}_{3,2}(A^{CLASS}_{2,3}/n_{+3})$	$A^{LOC}_{1,3}(A^{CLASS}_{2,1}/n_{+1})$ $+ A^{LOC}_{2,3}(A^{CLASS}_{2,2}/n_{+2})$ $+ A^{LOC}_{3,3}(A^{CLASS}_{2,3}/n_{+3})$
3	$A^{LOC}_{1,1}(A^{CLASS}_{3,1}/n_{+1})$ $+ A^{LOC}_{2,1}(A^{CLASS}_{3,2}/n_{+2})$ $+ A^{LOC}_{3,1}(A^{CLASS}_{3,3}/n_{+3})$	$A^{LOC}_{1,2}(A^{CLASS}_{3,1}/n_{+1})$ $+ A^{LOC}_{2,2}(A^{CLASS}_{3,2}/n_{+2})$ $+ A^{LOC}_{3,2}(A^{CLASS}_{3,3}/n_{+3})$	$A^{LOC}_{1,3}(A^{CLASS}_{3,1}/n_{+1})$ $+ A^{LOC}_{2,3}(A^{CLASS}_{3,2}/n_{+2})$ $+ A^{LOC}_{3,3}(A^{CLASS}_{3,3}/n_{+3})$

1.2. *Model assumptions*

The CLC error model is based on two key assumptions: (1) location and classification errors are independent of one another, and (2) errors in each time step are independent of errors in other time steps. Unfortunately, one could imagine scenarios where these assumptions would be violated. For example, steep terrain may affect both classification and location accuracy, and thus some correlation between location and classification error may occur. Furthermore, when a single DEM is used to georectify all of the remotely sensed images in a multi-temporal database, location errors might be correlated across time steps. While a recent study failed to find evidence of the first type of correlation (between location and classification errors within a single time period), it did find strong evidence of the second type of correlation (between errors in different time periods, Carmel 2004).

The goal of the present study is to evaluate the CLC error model, and in particular its adequacy under complex error patterns, in order to assess its suitability for modelling accuracy of spatio-temporal thematic datasets.

2. **Methods**

2.1. *Overview*

The goal of this study requires full knowledge of the spatial pattern of both error sources, and a large number of datasets. Field data are not amenable for this aim, since error amount 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, Haining and Arbia 1993, Mower 1994, Veregin 1995, Griffith *et al.* 1999). The basic approach adopted for this study is based on the evaluation of numerous spatio-temporal datasets, where the amount of error and its spatio-temporal structure are controlled. Towards this end, the entire process of error formation and propagation is simulated.

2.2. *Creating simulated multi-temporal datasets with known error characteristics*

The process of building an artificial multi-temporal dataset with a controlled, non-random error pattern was not trivial. This section describes the principles of this process, while further details on the structure and controls of the simulation can be found in the appendix. We used ArcInfo's Grid module and Arc Macro Language (AML, ESRI 2001) to construct a 512×512 thematic raster map, with each cell randomly assigned to one of three possible classes, according to pre-determined class proportions. An iterative smoothing procedure was then applied to the map in order to form a patchy pattern with a controlled degree of clustering (we used a focal majority function in the Grid module of ArcInfo, and its radius determined the degree of clustering, figure 1 (a) and (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. 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 degree of autocorrelation in these probability maps determined the magnitude of error clustering (or error patchiness) in the final maps produced by this process.

The error probability maps were then used to construct error precursor maps,

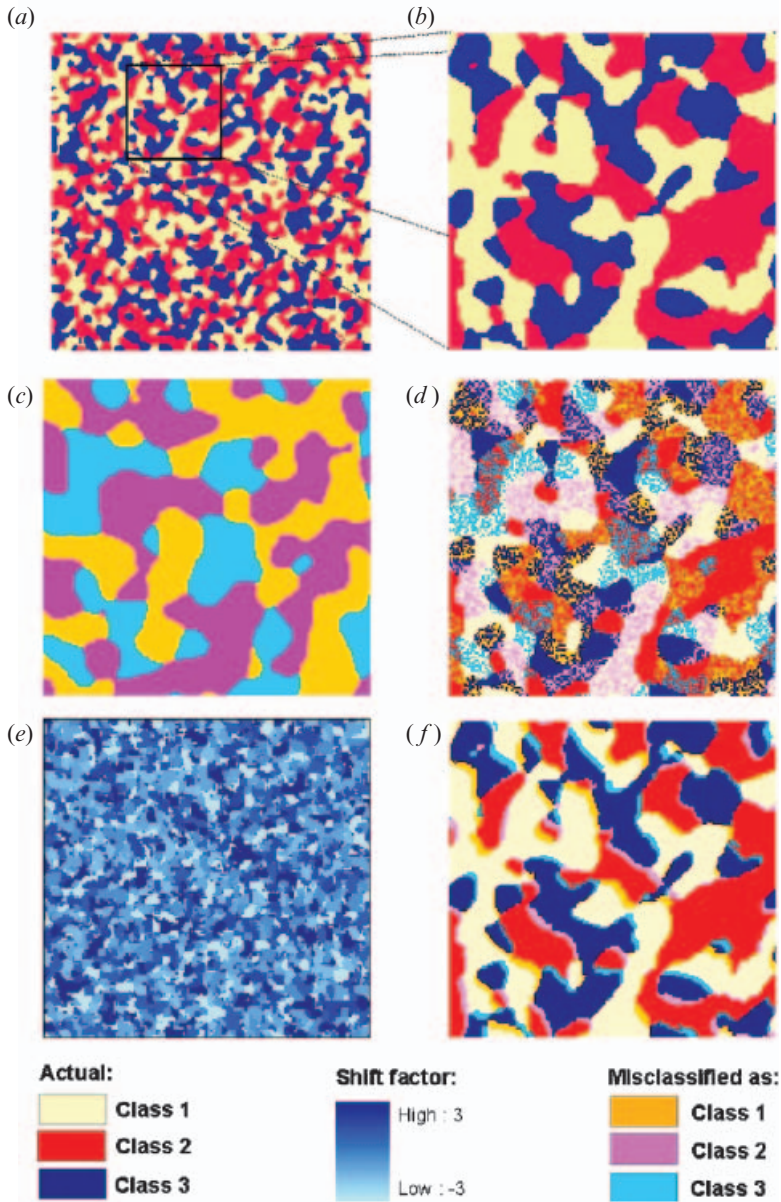


Figure 1. Examples of the various steps of the simulations. (a) The 'actual' map for time a . (b) A small part of the map of time a , enlarged. The rest of the images in this figure correspond to the same area represented in figure 1(b). (c) A classification error precursor map, used to construct the 'classified map', which contains a controlled degree of classification error. (d) Classification error map, resulting from overlaying a user-determined proportion of the error-map (depicted in (c)) over the 'actual' map (depicted in (b)). Pixels classified correctly are shown in their original colours (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', which 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 (c)) over the 'actual' map (depicted in (b)). Pixel colours are the same as in (d).

which in turn were used (along with the true map described earlier) to construct final maps with known error characteristics. In the case of classification error, this process was relatively simple. The classification error precursor map was a binary realization (e.g. 1=an error was present, 0=no error was present) of the classification error probability map. Combining this classification error precursor map with the true map to create a final map containing a known amount of classification error was a two part process:

1. For cells identified in the classification error precursor map as containing no error, cell values were simply copied from the true map to the final classification error map.
2. For cells identified in the classification error precursor map as containing an error, cells values were copied from another random, patchy map. The patchiness of this latter map controlled the spatial pattern of classification error, regardless of class values (figure 1 (c)). In some cases, cells were assigned back to their original values; yet this phase preceded the assessment of error structure and amount, and thus did not interfere with the evaluation of the model. Figure 1 (d) portrays the final classification error map.

In the case of location errors, the process of going from the error probability to the error precursor to the final map was somewhat more complex. Based on the probabilities in the location error precursor map, two error precursor maps were constructed independently, showing the amount of location error present in each raster cell in the x and y directions, respectively. These location errors were integer values ranging from -3 to 3 cells, and higher probabilities in the error precursor map translated into larger location errors (figure 1(e)). These location errors were applied to the true map by shifting cells within the true map by the cell-specific location error amounts (figure 1(f)). In order to preclude edge effects, we did not include image edges in the evaluation stage; 500×500 subset rasters were used for the evaluation stage of the process.

A combined error map was then constructed by first shifting the original image, and then introducing the classification error. The program can induce a controlled amount of correlation between location error and classification error patterns. This was achieved by constructing another random grid, and combining it with both the location error probability and classification error probability maps. The products of this process were two probability grids with a controlled degree of correlation between them.

This algorithm is applied twice, to represent multi-temporal dataset 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 (using a change precursor grid, similar to the error precursor grids). The program may induce correlation between error in the two time steps, and controls for its magnitude, in the same way as explained above for correlation between the two error types. In each simulation run, all the relevant correlations were assessed. Moran I was used to assess autocorrelation in the various 'true' and error maps. Pearson correlation coefficient was used to estimate correlation in error between points in time and between error types.

2.3. *Experimental design and analysis*

As a preliminary test, the robustness of the CLC error model was tested for correlation-free error structure. The following parameters were varied in a set of 50

simulation runs: number of map classes, proportion of each class, and rate of error of each type. Model performance in each simulation was recorded (for details on model assessment, see below).

Monte Carlo methods are typically used for stochastic simulations. In our simulations, where all the relevant parameters are pre-determined, there is little stochasticity. When simulations are run several times with all parameters kept constant, results differ from each other by $<10^{-3}$. Using Monte Carlo would yield little gain in information on the magnitude of deviance between model predictions and simulation results, and we chose not to use it.

In order to assess model performance under complex error pattern, ~ 200 simulated datasets were constructed, as unique combinations of the following parameters:

1. amount of location error (from -3 to $+3$ pixels in both the x and y directions).
2. amount of classification error (percent classified correctly (PCC) from 0.5 to 0.99).
3. degree of spatial autocorrelation within each type of error (Moran I from 0.05 to 0.95)
4. degree of spatial correlation between location and classification error (Pearson coefficient between 0 and 0.6).
5. degree of correlation in error between time steps (Pearson coefficient between 0 and 0.8).

Each simulated multi-temporal dataset was evaluated against an analytical solution counterpart. Location, classification, and combined error matrices were constructed using the equations shown in table 1. Values in these error matrices and the CLC model were used to compute the model-prediction of the probability of each transition type $p(C_1 C_2 \dots C_T)$ 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:

$$D = \left| p(C_1 C_2 \dots C_T)^{\text{observed}} - p(C_1 C_2 \dots C_T)^{\text{predicted}} \right| \quad (3)$$

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.

3. Results

3.1. Reference

As a form of reference, we simulated situations where classification and location error were uncorrelated, as were errors between time periods, conforming to the assumptions of the CLC model. A very good agreement between model predictions

and simulation results was found in the absence of correlation in error between time steps or between error types (D_{avg} did not exceed 0.002, and D_{max} did not exceed 0.01). These results were consistent for different numbers of classes (2–4), for various proportions of each class (between 0.01 and 0.99), and for both low and high levels of error of each source (PCC values ranged between 0.5 and 0.99).

3.2. Correlation between location and classification errors in a single time step

Introducing 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 (figure 2). 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 \sim 0.4) coincided with strong correlation between these sources (\sim 0.6). In those cases D_{avg} reached 0.015, while D_{max} reached 0.056 (figure 2).

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 $R^2=0.91$, $p<0.001$).

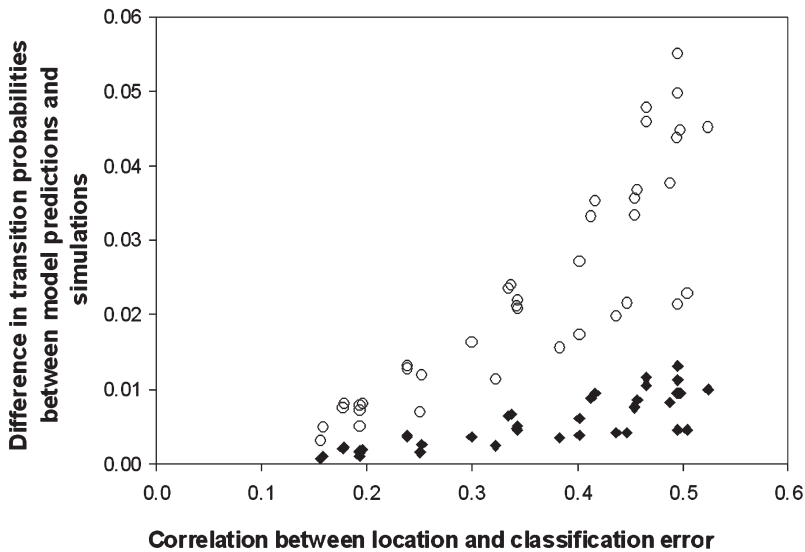


Figure 2. The \blacklozenge average and \circ maximum difference in transition probabilities between simulation results and the error model, as a function of the correlation between location and classification error patterns. Pearson coefficient is used to estimate correlation between error types.

3.3. Correlation between error patterns in time a and time b

Effects of correlation between error pattern in time a and time b were very small, but significant (figure 3(a) and (b), for location and classification error, respectively). The overall agreement between model predictions and simulations was 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.

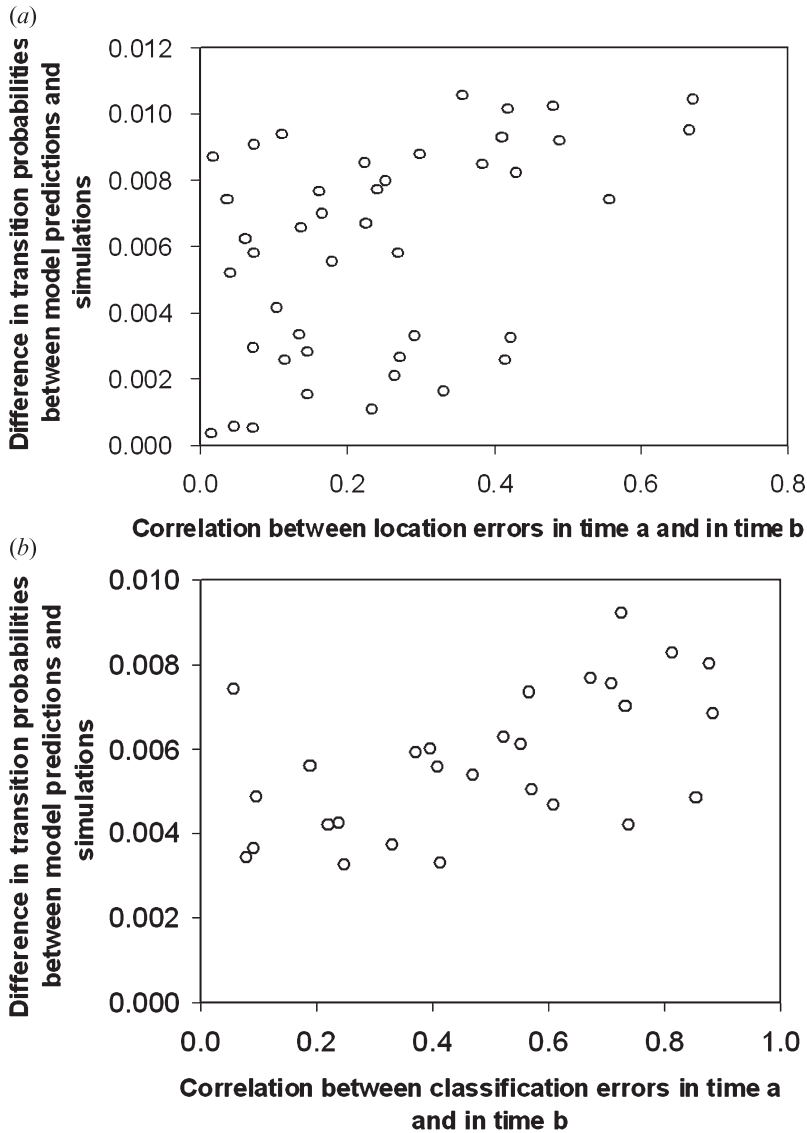


Figure 3. Maximum difference in transition probabilities between simulation results and the error model, as a function of correlation between error pattern in time a and time b , for (a) location error ($R^2=0.27$, $p<0.05$), and (b) classification error ($R^2=0.56$, $p<0.01$). Pearson coefficient is used to estimate correlation between error in time a and in time b .

3.4. *Effects of the spatial pattern of a single error source.*

As expected, the model was found to be insensitive to the rate of autocorrelation (clustering) in error pattern for both error types. D_{avg} was below 0.002 in all simulations, regardless of error pattern (Moran I values for error pattern ranged from 0.25 through 0.95).

4. Discussion

The CLC error model was designed to produce indices of overall data uncertainty in multi-temporal datasets used in thematic change detection analyses. The CLC model assumes independence between location and classification errors within each data layer, and independence of errors between data layers. This study assessed the CLC model reactions to departures from these assumptions. The CLC error model was found to be a precise descriptor of overall error when both the location error and classification error in the multi-temporal data are essentially randomly distributed. It was also found that the CLC model was insensitive to spatial autocorrelation of either location error, classification error, or both, within any single data layer. This finding, albeit expected (since autocorrelation of this type does not violate any of the CLC model's assumptions), is important since strong positive autocorrelation in error pattern was reported previously for both location error (Pugh and Congalton 2001) and classification error (Cherrill and McClean 1995).

The CLC model did react to correlation between errors in different time steps. The CLC model assumes that the error probability for a specific location (i,j) in a specific time step is independent of error probabilities at the same location in other time steps (see equation (2)). This study shows that the CLC model was sensitive to violations of this assumption. Thus, correlated errors of this type may, in principle, induce some uncertainty on the accuracy estimates produced by the model. However, it was also found that the magnitude of the problems caused by this type of correlation was minor, and for practical purposes, it is negligible.

Finally, this study evaluated how the CLC model reacts to correlations between location and classification error within an individual data layer. Correlations of this sort clearly violate the CLC assumptions, so it was expected that the CLC model would react to these types of correlation. The results of this study showed that for correlation values <0.2 , model performance was practically unaffected. However, for larger degrees of correlation, model performance was reduced more notably. For datasets with strong correlation between error types, the accuracy estimates produced by the model may be uncertain. For an entire dataset, the actual error estimate would differ by $\pm 1.5\%$ of the true value, while for specific transition types, this estimate may go up to $\pm 5\%$ of the actual value, in the case of the strongest possible correlation between error types.

A recent study (Carmel 2004) found that correlations between location and classification errors were less than 0.17, and insignificant, in all the five real-world datasets studied. This implies that in most actual situations, the CLC model is unlikely to suffer badly from correlations of this sort.

5. Conclusions

The CLC error model describes the overall error in spatio-temporal data, combining effects of image misregistration and classification error. It was found accurate and robust, insensitive to the magnitude and pattern of images, and to the magnitude of pattern of error, for uncorrelated error sources. The model assumes

independence between error types, and independence between error at different times. Violations of these assumptions are shown to affect the model. However, the small magnitude of these effects renders them negligible for practical purposes. Given our current knowledge on spatio-temporal error patterns in real data, the CLC error model can be applied reliably to spatio-temporal datasets.

Appendix. The simulation

This appendix contains details and explanations regarding the simulations. The general structure of the simulation is described first, followed by descriptions of the controls that the user has on the various parameters of the simulation.

A1. The simulation structure

1. Define a grid to serve as the template for the simulation (in this case, a 512×512 pixels image).
2. Create a simulated grid (hereafter ‘time_a’) that serves as the ‘true’ map for the first time-point in the simulation.
3. Modify ‘time_a’ to simulate the change in the scene with time. Construct ‘time_b’, as the ‘true’ map for the second point in time.
4. Modify ‘time_a’ to simulate location error that remains in the maps following georectification (‘loc_err_a’).
5. Modify ‘time_a’ to simulate classification error present in the map (‘class_err_a’).
6. Modify ‘time_a’ to simulate the observed map that combines both error types (hereafter, ‘observed_a’). This is done by inducing location error (as in step [4]), and then inducing classification error into the map (as in step [5]).
7. Repeat steps [4]–[6] for ‘time_b’, to construct ‘loc_err_b’, ‘class_err_b’ and ‘observed_b’, respectively.
8. Calculate the simulation CLC error matrices and transition probabilities, and compare with the respective values calculated using model equations.

A2. The simulation controls

1. *Controlling the number of classes and the proportion of each class in the original grid (‘a’).* A random function constructs a grid of 100 classes, with equal proportions of each class. Then, a conditioning function, based on threshold values, lumps these 100 percentiles into the desired number of classes, with the desired proportion for each class.
2. *Controlling the pattern of the original grid.* A focal majority function operates on a random grid, and constructs a ‘patchy’, clustered grid. The magnitude of clustering depends on window size. This process controls the autocorrelation in the image (more clustered image implies stronger autocorrelation). Moran I is used to quantify the autocorrelation in each image.
3. *Controlling the magnitude of change between time steps.* The simple and common way to induce change in spatial simulations is to randomly choose pixels for modification, and randomly choose a different class value to assign to the modified pixels. However, this random change is unrealistic, and results in a map with a more fragmented pattern than the original. The challenge is to modify class values of a specified proportion of the pixels in the original grid, while conserving the spatial pattern of the map. In our simulation, this goal is achieved using an alternative grid ‘alt_a’, constructed

similar to the original, with the same class proportions and pattern, but based on a different random grid. Each pixel in the ‘time b grid’ is assigned the value of either ‘a’ or ‘alt_a’, based on a threshold function coupled to a precursor grid. The precursor itself may be a clustered grid (similar to ‘time_a’), resulting in a non-random spatial pattern of change.

4. *Controlling location error magnitude and pattern.* The simple way to induce location error is to shift the entire grid by a specific distance on each axis. However, in order to create a more realistic pattern for location error, the extent of the shift needs to vary across the image. Towards this end, two precursor grids are used, ‘x_loc_err_precursor’, and ‘y_loc_err_precursor’, containing the pixel-specific shift in the *x* and *y* directions, respectively. These two precursor grids have their own spatial pattern that determines the spatial pattern of location error. They are constructed in the same way as the original grid (see items #1 and #2), only with a different range of values. The location error grid is produced by shifting each pixel in the original grid by the extent specified by the precursor grid.
5. *Controlling classification error magnitude and pattern.* Here, the challenge is to modify class values of a specified proportion of the pixels in the original grid, while controlling the spatial pattern of the modified pixels. This is achieved in a similar way to the construction of the ‘time_b’ grid (see item #3). An alternative grid ‘alt_a’, is constructed similar to the original ‘time_a’, with the same class proportions and pattern, but based on a different random grid. Each pixel in the classification error grid is assigned the value of either ‘a’ or ‘alt_a’, based on a threshold function coupled to a random grid of error probability, ‘class_err_prob’. The error probability map itself has a clustered pattern (constructed similar to ‘time_a’), resulting in a non-random spatial pattern of classification error. The autocorrelation in the classification error pattern is also controlled by the user, via the extent of clustering in the error probability map.
6. *Controlling the correlation between error from different types.* This sort of correlation is achieved by constructing the ‘class_err_precursor’ grid as a modified product of the ‘loc_err_precursor’ grid. The degree of modification determines the degree of correlation between the different error types.
7. *Controlling the correlation between error in different time steps.* Correlation in error between the two time steps is induced by deriving the error grids in both time steps from a common precursor. The degree of this correlation is controlled by the degree to which the common precursor is modified in the process of constructing the error maps for each point in time.

Acknowledgment

This study was supported by The International Arid Lands Consortium (IALC) project 02R-17. Avi Bar Massada helped with the simulation runs.

References

- ANTROP, M., 1998, Uncertainty, errors and doubts using GIS and remote sensing in environmental studies. *Natuurwet. Tijdschr.*, **78**, 1–12.
- CARMEL, Y., 2004, Complex error patterns in spatio-temporal datasets: the case of time-series vegetation maps derived from aerial-photos. *IEEE Transactions on Geoscience and Remote Sensing*, in press.
- CARMEL, Y., DEAN, D. J., and FLATHER, C. H., 2001, Combining location and classification

Spatio-temporal error model performance

- error sources for estimating multi-temporal database accuracy. *Photogrammetric Engineering and Remote Sensing*, **67**, 865–872.
- CHERRILL, A., and MCCLEAN, C., 1995, An investigation of uncertainty in field habitat mapping and the implications for detecting land cover change. *Landscape Ecology*, **10**, 5–21.
- CHRISMAN, N. R., 1991, The error component in spatial data. In *Geographical Information Systems: Principles and applications*, edited by D. J. Maguire, M. F. Goodchild, and D. W. Rhind (New York: Longman), pp.165–174.
- CONGALTON, R. G., and GREEN, K., 1999, *Assessing the Accuracy of Remotely Sensed Data: Principles and practices* (New York: Lewis Publishers).
- DAI, X. L., and KHORRAM, S., 1998, The effects of image misregistration on the accuracy of remotely sensed change detection. *IEEE Transactions on Geoscience and Remote Sensing*, **36**, 1566–1577.
- ESRI, 2001, *ArcGIS Users Guide* (Redlands, CA: Environmental Systems Research Institute).
- FISHER, P., 1998, Improved modeling of elevation error with geostatistics. *GeoInformatica*, **2**, 215–233.
- GOODCHILD, M. F., GUOQING, S., and 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., and ARBIA, G., 1999, Uncertainty and error propagation in map analyses involving arithmetic and overlay operation: inventory and prospects. In *Spatial Accuracy Assessment: Land information uncertainty in natural resources*, edited by K. Lowell and A. Jaton (Chelsea, MI: Ann Arbor), pp.11–25.
- HAINING, R. P., and ARBIA, G., 1993, Error propagation through map operations. *Technometrics*, **35**, 293–305.
- KYRIAKIDIS, P., and DUNGAN, J., 2001, A geostatistical approach for mapping thematic classification accuracy and evaluating the impact of inaccurate spatial data on ecological model predictions. *Environmental and Ecological Statistics*, **8**, 311–330.
- MOWRER, T. H., 1994, Monte Carlo techniques for propagating uncertainty through simulation models and raster-based GIS. *Proceedings of the International Symposium on Spatial Accuracy of Natural Resource Data Bases: Unlocking the puzzle*, edited by R. G. Congalton (Bethesda, MD: American Society for Photogrammetry and Remote Sensing), pp.179–198.
- MOWRER, T. H., 1997, Propagating uncertainty through spatial estimation processes for old-growth subalpine forests using sequential Gaussian simulation in GIS. *Ecological Modelling*, **98**, 73–86.
- MOWRER, T. H., 1999, Accuracy (re)assurance: selling uncertainty assessment to the uncertain. *Spatial Accuracy Assessment: Land information uncertainty in natural resources*, edited by K. Lowell and A. Jaton (Chelsea, MI: Ann Arbor), pp.3–10.
- PHILLIPS, D. L., and MARKS, D. G., 1996, Spatial uncertainty analysis: propagation of interpolation errors in spatially distributed models. *Ecological Modelling*, **91**, 213–229.
- PUGH, S. A., and CONGALTON, R. G., 2001, Applying spatial autocorrelation analysis to evaluate error in New England forest-cover-type maps derived from Landsat Thematic Mapper data. *Photogrammetric Engineering and Remote Sensing*, **67**, 613–620.
- STOMS, D. M., DAVIS, F. W., and COGAN, C. B., 1992, Sensitivity of wildlife habitat models to uncertainties in GIS data. *Photogrammetric Engineering & Remote Sensing*, **58**, 843–850.
- STOW, D. A., 1999, Reducing the effects of misregistration on pixel-level change detection. *International Journal of Remote Sensing*, **20**, 2477–2483.
- TOWNSHEND, J. R. G., JUSTICE, C. O., GURNEY, C., and MCMANUS, J., 1992, The impact of misregistration on change detection. *IEEE Transactions on Geoscience and Remote Sensing*, **30**, 1054–1060.
- VEREGIN, H., 1995, Developing and testing of an error propagation model for GIS overlay operations. *International Journal of Geographical Information Systems*, **9**, 595–619.