ELSEVIER

Agriculture, Ecosystems and Environment 85 (2001) 239–248

Agriculture Ecosystems & Environment

www.elsevier.com/locate/agee

# Land-cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA

R. Gil Pontius Jr.*, Laura C. Schneider

*Graduate School of Geography, Clark University, 950 Main Street, Worcester, MA 01610-1477, USA*

**Abstract**

Scientists need a better and larger set of tools to validate land-use change models, because it is essential to know a model's prediction accuracy. This paper describes how to use the relative operating characteristic (ROC) as a quantitative measurement to validate a land-cover change model. Typically, a crucial component of a spatially explicit simulation model of land-cover change is a map of suitability for land-cover change, for example a map of probability of deforestation. The model usually selects locations for new land-cover change at locations that have relatively high suitability. The ROC can compare a map of actual change to maps of modeled suitability for land-cover change. ROC is a summary statistic derived from several two-by-two contingency tables, where each contingency table corresponds to a different simulated scenario of future land-cover change. The categories in each contingency table are actual change and actual non-change versus simulated change and simulated non-change. This paper applies the theoretical concepts to a model of deforestation in the Ipswich watershed, USA. © 2001 Elsevier Science B.V. All rights reserved.

*Keywords:* ROC; LUCC; Land cover; Suitability map; Simulation model; Validation

## 1. Introduction

### 1.1. Validation in land-cover change models

The international scientific community has called for research into land-cover change, specifically for research into models that predict spatial patterns of future change (Turner et al., 1995; Lambin et al., 1999). Modelers are satisfying this need with a variety of approaches (Wilkie and Finn, 1988; Baker, 1989; Lambin, 1994, 1997; Pontius, 1994; Hall et al., 1995; Veldkamp and Fresco, 1996; Geoghegan et al., 1997; Mertens and Lambin, 1997; Liverman et al., 1998; Wu and Webster, 1998). In most cases, the models are connected to a raster-based geographic information system (GIS). The models predict which grid cells are likely to experience future land-cover change. Scientists must develop statistical methods to validate such models, because it is essential to know a model's prediction accuracy. This paper offers a method of validation that uses a quantitative measurement called the relative operating characteristic (ROC). The ROC technique applies to any model that predicts a homogenous category in each grid cell.

Fig. 1 shows the general flow of logic of a typical method to calibrate, run and validate a land-cover change model. The calibration phase uses maps of factors that guide the location of future change, such as protection status, slope, and proximity to roads. These factor maps can be combined with one or more maps of historical land cover to complete calibration. The model usually uses the information in the

* Corresponding author. Tel.: +1-508-793-7761;
fax: +1-508-793-8881.
*E-mail address:* rpontius@clarku.edu (R.G. Pontius Jr.).

0167-8809/01/$ – see front matter 
PII: S0167-8809(01)00187-6

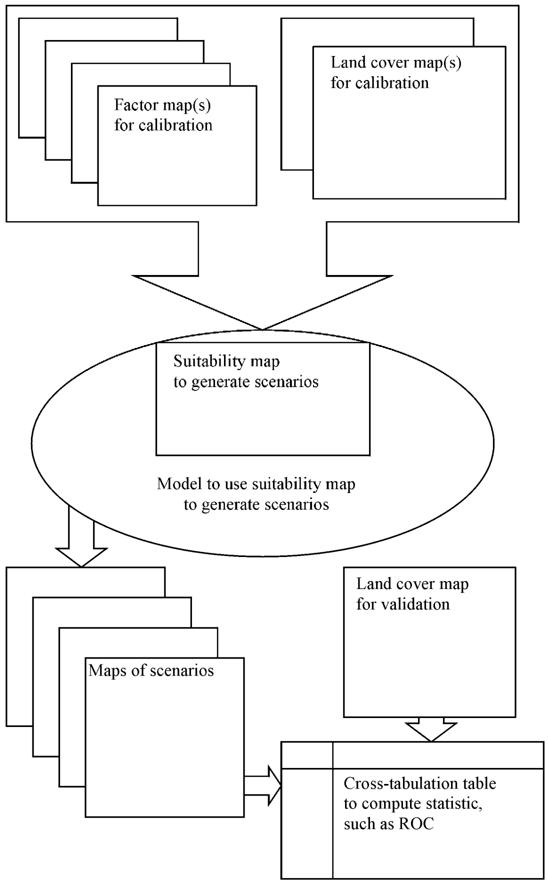


Fig. 1. Flow of logic among components necessary to validate a land-cover change model.

calibration maps to construct some type of map of suitability or probability of land-cover change, such as Fig. 2. Suitability maps can be created through logistic regression, or through multi-criteria methods (Schneider and Pontius, 2001; Hall et al., 1995). Each model run uses the map of suitability to generate a map of simulated future change, placing simulated change in cells that have the largest suitability values. If the suitability map were perfect, the order of the suitability values would match the order in which humans change the landscape, with the largest suitability values being changed first. For validation, a map of simulated future change is compared to a map of recent real land-cover change, such as Fig. 3. For appropriate validation, the map of reality used for validation should not be used in calibration. An index of agreement measures similarity between the simulated change and real change. Several model runs can generate a series of simulated maps, each with a different quantity of change, hence each with a different level of agreement with the map of reality.

### *1.2. Validation complication*

When the modeler overlays the map of a simulated landscape on the map of reality, a contingency table can summarize the results, as in Table 1, for cases where each grid cell is a homogenous land type. The rows of Table 1 show categories of the map of the model's output and the columns show the categories of the map of reality. The entries are the number (or proportion) of cells that fall into each category combination. Therefore the proportion correct is $(A + D)/(A + B + C + D)$.

The proportion correct is sensitive to the model's ability to simulate quantity, which is the marginal distribution in the row totals in Table 1. Fig. 4 shows how the model's quantity of change restricts the percent correct. The horizontal axis shows the percent of grid cells that the model specifies as change. The vertical axis shows the percent of grid cells classified correctly by the model. Fig. 4 shows an example where the quantity of change in the map of reality is 10%, while the quantity of change in a modeled scenario can range from 0 to 100%.

For Fig. 4, let us consider the upper limit on percent correct. In order for the model to attain 100% correct classification, it must specify the quantity of change correctly as 10%, and it must specify the location correctly via the map of suitability. Notice that many different suitability maps can lead to 100% correct classification. If the top 10% of suitability values are distributed in any way among the locations of actual change, then there will be 100% correct

Table 1
Two-by-two contingency table showing the proportion (or number) of grid cells in a map of reality versus a map of a modeled scenario

| *Model* | Reality | | |
|---|---|---|---|
| | Change | Non-change | Total |
| Change | A | B | A + B |
| Non-change | C | D | C + D |
| Total | A + C | B + D | A + B + C + D |

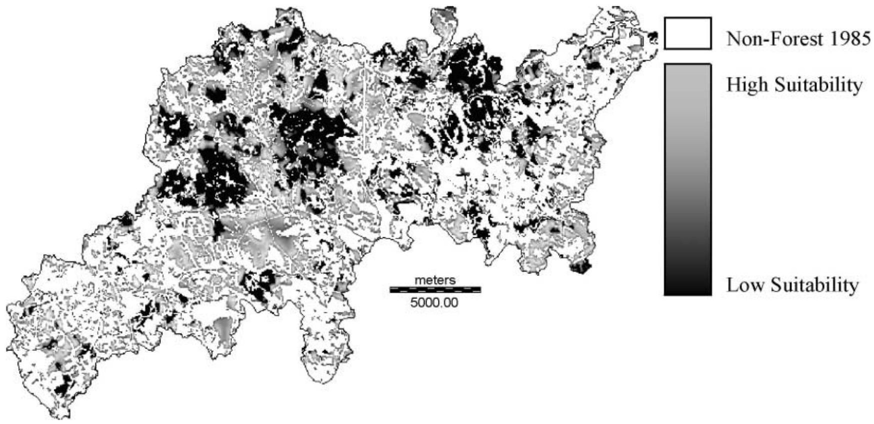


Fig. 2. Map of suitability for deforestation after 1985 in the Ipswich watershed.

classification, regardless of how the bottom 90% of suitability values are distributed. However, the maximum percent success is constrained by the quantity of simulated change in the modeled scenario. The percent of success cannot occur in the solid white upper left and upper right triangles of Fig. 4. If the model specifies some quantity of change other than quantity of change found in the map of reality, then the simulation's percent correct must be less than 100%, even when the map of suitability is perfect.

Now consider the lower limit on percent correct. The solid black lower left and lower right triangles of Fig. 4 show levels of percent correct that are not attainable due to the model's specification of quantity

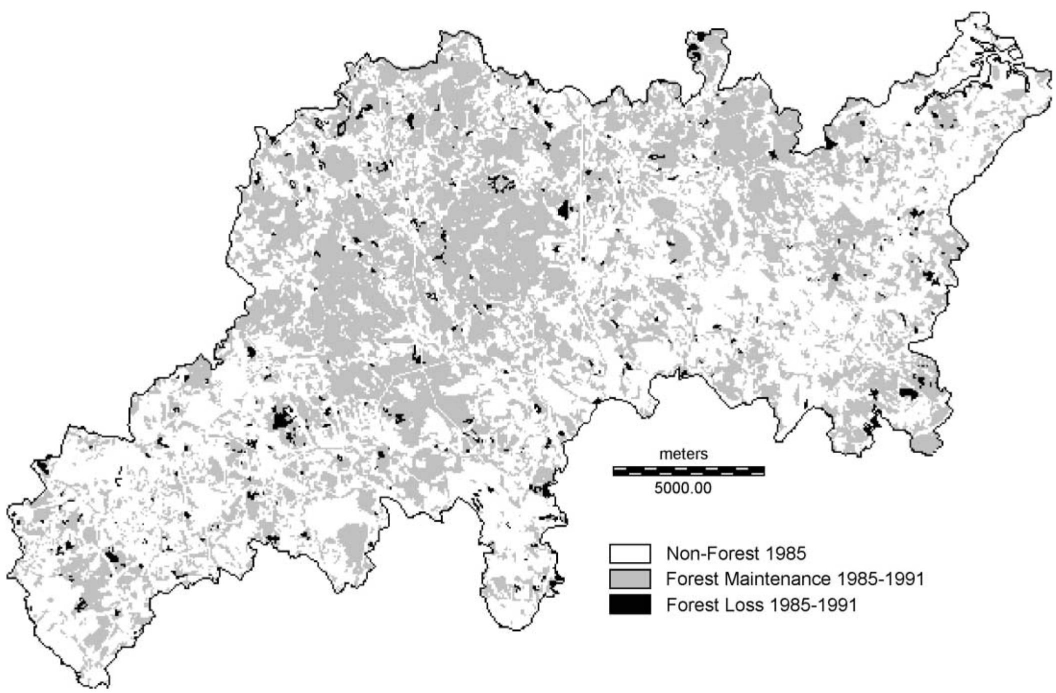


Fig. 3. Map of real deforestation from 1985 to 1991 in the Ipswich watershed.

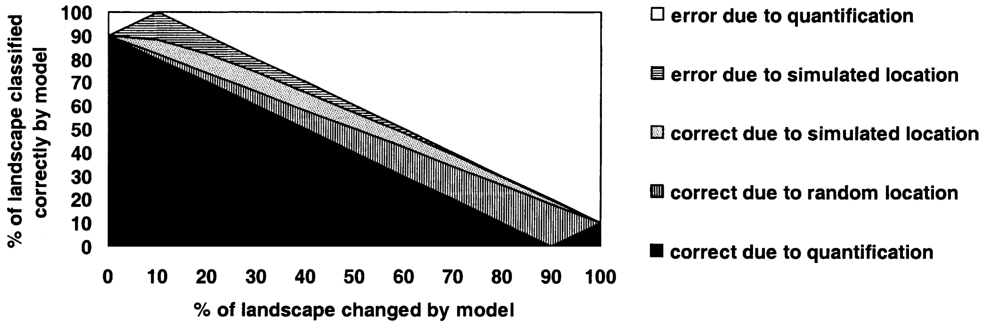


Fig. 4. Range for percent of landscape classified correctly by a model as a function of percent of landscape changed in a modeled scenario, where 10% of landscape is changed in reality.

of change. That is, the model is guaranteed to attain a percent of success greater than or equal to the upper boundary of the lower two triangles of Fig. 4. When the model specifies a percent of change near 10%, then the model is guaranteed to attain a high percent success, even when the suitability map's specification of location is poor. Even if none of the cells classified as "change" by the model are "change" in reality, the percent correct would be high.

Fig. 4 shows that the percent correct for any model must lie within the parallelogram defined by the points $(1 - R, 0)$, $(1, R)$, $(R, 1)$ and $(0, 1 - R)$, where $R$ is the proportion change in the map of reality. When $R$ is near 0 or 100%, then the model's percent correct is influenced dramatically by the model's ability to specify quantity. Notice in Fig. 4 that for a given specification of quantity in the model, the maximum and minimum percent correct are very close. Within this parallelogram, the model's actual percent correct is determined by the quality of the map of suitability. If the suitability map were to assign locations at random, then the percent success falls on the line that bisects the parallelogram. So if the map of suitability were better than random, then the percent success would lie with an extremely narrow range. Clearly, it is not reasonable to judge the quality of the map of suitability on its ability to attain a large percent correct for a single simulation run, because the percent correct fails to separate error due to quantification from errors due to location (Pontius, 2000). The result of any single simulation run depends on a single use of the map of suitability and a single specification of quantity. If the quantity of change were different, it is not clear how the model would perform. Furthermore, two substantially different suitability maps can give an identical percent correct.

Therefore, in order to assess the quality of a model, it is advantageous: (1) to use an indicator other than percent success, (2) to measure its performance over a variety of scenarios of quantity of change, and (3) to illustrate the validation with a figure that shows clearly how a high agreement differs from a low agreement. The ROC satisfies all these goals. Although the ROC does not separate explicitly error due to quantity from error due to location, the ROC aggregates into a single index of agreement the success of several model scenarios of various quantities of change. The ROC is presented visually in a success space that is re-scaled, so that it is much more illustrative than Fig. 4. Scientists in Engineering, Medicine, Meteorology, Psychology, and several other fields have used the ROC to measure the relationship between a signal and reality (Ogilvie and Creelman, 1968; Egan, 1975; Metz, 1978; Swets, 1986a,b, 1988). This paper brings the ROC to the field of land-cover change modeling to measure relationship between simulated change and real change. Many authors have shown clever ways to glean useful information from contingency tables (Carstensen, 1987; Congalton, 1991; Congalton and Green, 1999). The ROC should be added to those methods.

## 2. Methods

### 2.1. Map preparation

To compute the ROC, the methodology described in this paper assumes the modeler has both: (1) a suitability map that shows the relative likelihood that a cell undergoes land-cover change, and (2) a map of reality used for validation, where each cell is categorized as either change or non-change. The suitability map shows the sequence in which the model selects grid cells for change (Fig. 2).

The ROC is explained in terms of a suitability map that shows the priority in which grid cells are selected for change because it is easiest to understand the ROC in terms of a suitability map. Many modelers will want to use a similar technique to evaluate their own maps of suitability or probability. However, it is not necessary that the model use a "suitability map" per se. For example, cellular automata models can select locations for predicted deforestation according to a variety of decision rules (Wagner, 1997). ROC can validate any model that can generate a series of simulation maps where each grid cell is classified as a homogenous category. It is best if each simulated scenario has a different quantity of change.

The ROC works for exactly two land types. If the modeler has more than two land-cover types, then the modeler can create an ROC for each land type. This can be accomplished by reclassifying the maps into the category of interest versus other, thus each category can have its own ROC.

The first step in creating the ROC is to reclassify the map of suitability by slicing it into a map of several suitability percentile groups, for example 10 deciles. A possible reclassification rule assigns 10% of the cells that have the largest suitability values to group 1. The rule assigns another 10% of the cells that have the next largest suitability values to group 2, and so on till the rule assigns 10% of the cells that have the smallest suitability values to group 10. Therefore, each of the 10 groups contains 10% of the cells in the study area. It is not absolutely necessary for each group to have the exact same percent of grid cells, however, it is desirable to have the slices between groups spread among various percentiles of suitability values. It is possible theoretically to have many more than 10 groups. The maximum number of groups is the number of unique suitability values. An increase in number of groups increases accuracy of the estimated ROC, but increases also the complexity of calculations.

Next, the modeler overlays the map of suitability groups with the map of reality. A tabular result shows the proportion of grid cells of each suitability group that are categorized as change versus non-change in reality, as in the first two columns of Table 2. Think of Table 2 as an efficient way to show two-by-two contingency tables for 11 scenarios of land-cover change. Each of the 11 scenarios uses the same map of reality and the same map of suitability, however, the quantity of change in each scenario is different. The scenario in the top row of the table is a special case in which the quantity of change is 0, therefore none of the 10 suitability groups are classified as change in the scenario. The next scenario simulates change in only suitability group 1, the subsequent scenario simulates change in suitability groups 1 and 2, and so on till the last scenario simulates change in all 10 groups. For each scenario, the two cumulative columns of Table 2 give the percent of grid cells that each scenario clas-

Table 2
Cross tabulation of grid cells in a map of reality versus a map of suitability for 11 scenarios of deforestation. All numbers are percents

| *Scenario's suitability group* | Reality | | Reality (cumulative) | | Scenario statistics | | | |
|---|---|---|---|---|---|---|---|---|
| | Change | Non-change | Change | Non-change | Change | Correct | False positive | True positive |
| – | – | – | – | – | 0.0 | 95.4 | 0.0 | 0.0 |
| 1 | 0.4 | 3.5 | 0.4 | 3.5 | 3.9 | 92.3 | 3.7 | 8.7 |
| 2 | 0.1 | 1.0 | 0.5 | 4.5 | 5.0 | 91.4 | 4.7 | 10.7 |
| 3 | 0.6 | 5.0 | 1.1 | 9.5 | 10.6 | 87.0 | 10.0 | 23.1 |
| 4 | 0.7 | 8.2 | 1.8 | 17.7 | 19.5 | 79.5 | 18.6 | 39.4 |
| 5 | 0.7 | 9.9 | 2.5 | 27.6 | 30.1 | 70.3 | 28.9 | 55.0 |
| 6 | 0.5 | 8.6 | 3.0 | 36.2 | 39.3 | 62.2 | 38.0 | 66.2 |
| 7 | 0.6 | 9.6 | 3.6 | 45.9 | 49.5 | 53.2 | 48.1 | 78.8 |
| 8 | 0.5 | 10.1 | 4.1 | 56.0 | 60.1 | 43.6 | 59.7 | 90.5 |
| 9 | 0.2 | 5.7 | 4.3 | 61.1 | 66.0 | 38.1 | 64.6 | 94.6 |
| 10 | 0.2 | 33.8 | 4.6 | 95.4 | 100.0 | 4.6 | 100.0 | 100.0 |

sifies as non-change and change. For each scenario, a number of the "Reality (cumulative) Change" column of Table 2 shows a percent of grid cells that is analogous to a value of A in Table 1, and a number of the "Reality (cumulative) Non-change" column of Table 2 shows a percent of grid cells that is analogous to a value of B in Table 1. The bottom row of the "Reality (cumulative) Change" column of Table 2 shows the percent of grid cells that is analogous to a value of A + C in Table 1, and the bottom row of the "Reality (cumulative) Non-change" column of Table 2 shows a percent of grid cells that is analogous to a value of B + D in Table 1. The "Scenario Statistics Change" column is the sum of the two "Reality (cumulative)" columns, hence it gives the percent of cells modelled as change in the scenarios. The "Scenario Statistics Correct" column gives the percent classified correctly by the each scenario. The final two columns give the percent of true-positives and false-positives, which are explained next.

If a grid cell is simulated as change in a scenario, it is a 'positive'. Therefore, a 'true-positive' is a cell that is categorized as change in both reality and the modeled scenario. A 'false-positive' is a cell that is categorized as non-change in reality and as change in the modeled scenario. Therefore, in a two-by-two contingency table such as Table 1, the rate of true-positives is A/(A + C) and the rate of false positives is B/(B + D). For each of the 11 scenarios, Table 2 shows the rate of true-positives and false-positives.

The rates of true-positives and false-positives contain the same information as the rates of true-negatives and false-negatives, so it would be redundant to perform analysis on the rates of negatives. If one were to switch the order of the rows in the contingency table (Table 1), then one could analyze the situation in terms of negatives. Analysis of the positives gives the same ROC as analysis of the negatives.

### 2.2. Relative operating characteristic

To define the ROC, Fig. 5 plots the rate of true-positives on the vertical axis versus the rate of false-positives on the horizontal axis for each of the 11 scenarios. Each scenario corresponds to a point in the plotted space. The ROC statistic is the area under the curve that connects the plotted points. Eq. (1) uses integral calculus' trapezoidal rule to compute the area, where $x_i$ is the rate of false positives for scenario $i$, $y_i$ the rate of true positives for scenario $i$, and $n$ the number of suitability groups:

$$\text{Area under curve} = \sum_{i=1}^{n}[x_{i+1} - x_i][y_i + y_{i+1} - y_i/2] \quad (1)$$

If the sequence of the suitability values matches perfectly the sequence in which real land-cover change has occurred, then ROC equal to 1, because as the quantity of change in the scenarios increases from 0 to 100%, the ROC curve begins at the origin, goes up the horizontal axis to the point (0, 100%), then passes to the right to the point (100, 100%). More generally, ROC equal to 1 for any suitability map for which the locations that experience change have larger suitability values than the locations that do not experience change. Note several different suitability maps could have ROC of 1.

If the order of suitability values were assigned at random locations across the landscape, then the expected value of the ROC would be 0.5, as shown by the dashed line in Fig. 5. Under random allocation of suitability values, the expected number of true-positives is (A + B)(A + C)/(A + B + C + D) and the expected number of false-positives is

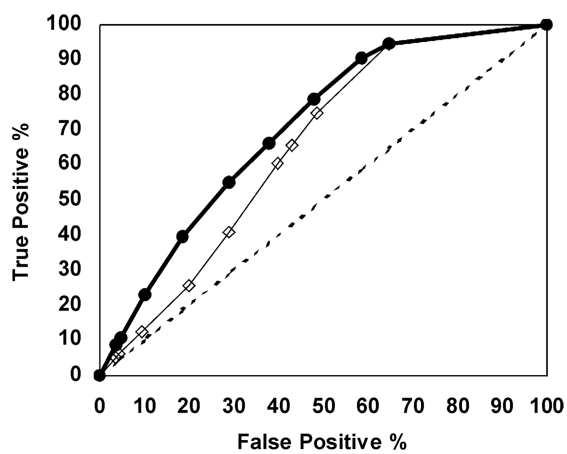


Fig. 5. ROC curves to validate models of deforestation in the Ipswich watershed between 1985 and 1991 using three suitability maps based on: random location (bottom ROC = 50%), logistic regression with protected areas (middle ROC = 65%), and MCE with a spatial filter and protected areas (upper ROC = 70%).

$(A + B)(B + D)/(A + B + C + D)$. Therefore, both the expected rate of true-positives and the expected rate of false-positives are $(A + B)/(A + B + C + D)$. However, for any particular assignment of random locations, there will be some variation from these expected rates. Therefore, Monte Carlo analysis is performed to compute the ROC for 10 000 runs, where the locations are selected at random for each run. This Monte Carlo analysis tests whether the model assigns locations that are significantly different than random, by comparing the ROC of the model to the range of the ROC values among the Monte Carlo runs.

### *2.3. Application to the Ipswich watershed, Massachusetts*

To illustrate the theoretical concepts, the ROC statistical methodology is applied to a land-cover change model that uses suitability maps such as Fig. 2 to predict the location of new deforestation from 1985 to 1991 in the watershed of Ipswich Massachusetts, USA. The suitability maps are calibrated with maps of socio-physical characteristics and forest areas in 1971 and 1985. The model assigns relatively high suitability values to those locations that have a combination of socio-physical characteristics similar to land that has experienced deforestation from 1971 to 1985. The State of Massachusetts supplies the maps for calibration and validation (MassGIS, 1999).

The ROC methodology compares two suitability maps that are generated by two different techniques. The first suitability map is generated by logistic regression where $Y = 1$ if a cell undergoes deforestation from 1971 to 1985, and 0 otherwise. The independent variables are elevation, slope and distance to residential areas. Then, the resulting map is overlaid with a map of protected areas, hence any protected areas have the minimum suitability value. The second suitability map is generated by multi-criteria evaluation (MCE) (Pereira and Duckstein, 1993; Eastman et al., 1995; Eastman and Jiang, 1996). Based on empirical analysis, the MCE assigns large suitability values to locations near existing residential areas and have topography desirable for residential development (Schneider and Pontius, 2001). Also, any protected areas are forced to have the minimum suitability value in the same manner as in the first suitability map. In addition, a spatial filter is applied so that the suitability value at any one cell is an average of the suitability values of neighboring cells.

The model is validated with a map of actual deforestation from 1985 to 1991, shown in Fig. 3. Forested locations are the only candidates for deforestation, therefore the study area for simulation and validation is restricted to cells in the watershed that were forest in 1985, which are black or gray in Figs. 2 and 3. Each suitability map generates one ROC.

The ROC methodology is applied with 10 suitability categories, however, it helpful to define some of the suitability groups for scenarios of special interest. For one scenario, the quantity of deforestation from 1985 to 1991 is based on linear extrapolation of the historic quantity of change. Therefore, the most suitable group (group 1) has a quantity of simulated deforestation equal to 4% of the forest area of 1985, which is the "best guess" at the real quantity of deforestation. The second most suitable group (group 2) is set such that the sum of the quantities in groups 1 and 2 is the quantity of deforestation that actually occurred between 1985 and 1991, which is 5% of the forested area. Most of the other groups have approximately 10% of the 1985 forest area. However, the last suitability group (group 10) consists of forested areas that have some form of legal protection, which constitutes 34% of the 1985 forest area.

## 3. Results

Logistic regression produces a suitability map with an ROC of 65%, and MCE produces a suitability map with an ROC of 70%. Table 2 gives the points on the curve of the MCE-based suitability map. Fig. 5 plots the ROC curves for the suitability maps.

The point at the origin in each ROC curve is the scenario that has zero simulated deforestation; therefore both false positives and true positives are zero. The next point on each curve has simulated deforestation in suitability group 1 only. This scenario shows a quantity of deforestation equal to the best guess at the quantity of deforestation. For the MCE-based suitability map, 92% is the percent success for the "best guess" scenario, however, most of those successes are correct simulation of non-change. The correct simulation of change is 9% of the total real change.

Each consecutive point on the ROC curve represents a scenario that has a larger quantity of simulated deforestation. The final point on the curve at (100%, 100%) shows a scenario for which the entire study area is simulated as deforested, hence both false positives and true positives are 100%. The second to last point shows the scenario where all unprotected forest of 1985 is simulated as deforested, and all protected land is simulated as forest maintenance. Notice that 5% of the real deforestation between 1985 and 1991 occurs on protected land.

The dashed line in Fig. 5 shows the expected ROC for a model that selects grid cells at random. This ROC expected due to chance is 0.5, however, for any particular random sequence of selected cells, there is some variation around 0.5. Over 10 000 Monte Carlo runs in which the sequence of grid cells were selected at random, the ROC ranged from a minimum of 48% to a maximum of 52%. Therefore, the ROCs of 65 and 70% are significantly better than random.

## 4. Discussion

What constitutes a "good" ROC? Any ROC above 50% is better than random. The results from the Monte Carlo analysis show that for the Ipswich example, any ROC above 52% is statistically better than random. But "better than random", does not necessarily mean "good". In other fields, ROCs have ranged from: 71 to 89% for weather forecasting, 75 to 97% for library information retrieval, 81 to 93% for medical imaging diagnosis, 68 to 93% for material strength testing, and 55 to 98% for polygraph lie detection. So the results of the Ipswich land-use change example are comparable with other fields.

Most importantly, the ROC helps to guide decisions concerning modeling strategy. The ROC has guided the decisions whether to use MCE or logistic regression, whether or not to force protected areas to have the minimum suitability values, whether or not to use a spatial filter, etc. In short, the ROC tells which modeling approach generates the best maps of suitability.

Results are sensitive to how the modeler defines the study area. Usually there are two approaches to define the study area. The first definition states that the study area consists of any cell that is in the region of interest, for example any cell in the watershed. The second definition is that the study area consists of the subset cells in the region of interest that are candidates for change, for example any forested cell in the watershed. The first definition usually will result in a higher percent success and a higher ROC because the model gets credit for predicting future non-forest at locations that did not have forest at the beginning of the simulation. In situations where forest re-growth is not of interest, the second definition should be used. The Ipswich watershed example used the second definition.

The ROC validates the model at quantities that can be far from the estimated best guess quantity. Even though the model is fairly accurate at predicting the quantity of change between 1985 and 1991, it is important to validate the model at several different quantities because it is necessary to know whether high (medium or low) suitability values have a relatively high (medium or low) amount of real change. This is important when the suitability map extrapolates the pattern of land change over several decades, hence into years for which there are large percents of land change. Even if the model does not extrapolate far into the future, it is desirable to confirm that small suitability values have little or no change. For example, the high rate of true positives on second to last point on the ROC curves shows that the suitability map uses the information of protected status in an intelligent manner. If the validation method had concentrated on only the best guess quantity, then the contribution of the protected status to the validation would have been less clear.

A weakness of the ROC is that it is does not account for the spatial arrangement of the model's successes and errors. All of the information for the ROC comes from contingency tables that show only cell-by-cell analyses of association between a map of real change and maps of simulated change, where each grid cell is a homogenous category. This weakness is shared by any index of agreement that is based on a contingency table, because a contingency table is a summarization that contains no information of location. Other indices that share this weakness are *percent correct*, any chi-square based statistic such as *phi*, and any *kappa* based statistic such as Kstandard, Kno, Klocation, or Kquantity (Pontius, 2000). Therefore the ROC should be supplemented with visual comparison and additional measures of association that account for spatial pattern (Costanza, 1989; Turner et al., 1989).

The next stage of development of the ROC and other non-spatial indices of agreement is to derive multiple resolution versions of them. A multiple resolution method would measure the agreement at various resolutions, by aggregating neighboring cells into an increasingly coarse grid, or window. At every grid resolution, the aggregation process will create heterogeneous cells that could have any proportion of the possible categories. Therefore, the method to measure agreement between a simulated cell and a real cell must be able to measure fractional agreement, i.e. agreement between 0 and 1. If a statistical method can address fractional agreement, then it will not be necessary for the cells at the finest resolution to be homogenous; hence the cells in the maps of simulation and reality can be fuzzy categories. The general method to derive multiple resolution statistics is the topic of an upcoming paper.

## 5. Conclusions

The ROC enables land-cover change modelers to validate a model's ability to specify location, while maintaining the freedom from committing to a specific quantity of change. ROC is used in Engineering, Medicine, Meteorology, Psychology, and several other scientific disciplines. Land-use change modelers should also use the ROC.

## Acknowledgements

The National Science Foundation (NSF) funded this research through the Water and Watersheds program grant DEB-9726862. Additional contributors include the Jesse B. Cox Charitable Trust, the Sweetwater Trust, and NSFs Long Term Ecological Research program OCE-9726921. Also, we thank our colleagues of the Integrated Modeling Group at Clark University and the Center for Integrated Studies of Global Environmental Change at Carnegie Mellon University, with which this work is increasingly tied intellectually and programmatically.

## References

Baker, W.L., 1989. A review of models in landscape change. Landscape Ecol. 2 (2), 111–133.

Carstensen Jr., L.W., 1987. A measure of similarity for cellular maps. Am. Cartographer 14 (4), 345–358.

Congalton, R., 1991. A review of assessing the accuracy of classification of remotely sensed data. Remote Sensing Environ. 37, 35–46.

Congalton, R., Green, K., 1999. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices. Lewis Publishers, New York, 137 pp.

Costanza, R., 1989. Model goodness of fit: a multiple resolution procedure. Ecol. Model. 47, 199–215.

Eastman, R., Jiang, H., 1996. Fuzzy measures in multi-criteria evaluation. In: Proceedings of the Second International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, Fort Collins, CO.

Eastman, J.R., Jin, W., Kyem, P.A.K., Toledano, J., 1995. Raster procedures for multi-criteria/multi-objective decisions. Photogrammetric Eng. Remote Sensing 61 (6), 539–547.

Egan, J.P., 1975. Signal Detection Theory and ROC Analysis. Academic Press, New York.

Geoghegan, J., Wainger, L.A., Bockstael, N.E., 1997. Spatial landscape indices in a hedonic framework: an ecological economics analysis using GIS. Ecol. Econ. 23, 251–264.

Hall, C.A.S., Tian, H., Qi, Y., Pontius, G., Cornell, J., 1995. Modelling spatial and temporal patterns of tropical land use change. J. Biogeogr. 22 (4/5), 753–757.

Lambin, E.F., 1994. Modelling deforestation processes: a review. European Commission, Luxembourg.

Lambin, E.F., 1997. Modelling and monitoring land-cover change processes in tropical regions. Prog. Phys. Geogr. 21 (3), 375–393.

Lambin, E.F., Baulies, X., Bockstael, N., Fischer, G., Krug, T., Leemans, R., Moran, E.F., Rindfuss, R.R., Sato, Y., Skole, D., Turner II, B.L., Vogel, C., 1999. Land-use and land-cover change implementation strategy. Royal Swedish Academy of Sciences, Stockholm, Sweden.

Liverman, D., Moran, E., Rindfuss, R., Stern, P. (Eds.), 1998. People and Pixels. National Academy Press, Washington, DC.

MassGIS, 1999. Land use map. Executive Office of Environmental Affairs, Boston, MA, 2000.

Mertens, B., Lambin, E., 1997. Spatial modelling of deforestation in southern Cameroon. Appl. Geogr. 17 (2), 143–162.

Metz, C.E., 1978. Basic principles of ROC analysis. Seminars Nucl. Med. 8, 283–298.

Ogilvie, J.C., Creelman, C.D., 1968. Maximum likelihood estimation of ROC curve parameters. J. Math. Psychol. 5, 377–391.

Pereira, J.M., Duckstein, L., 1993. A multiple criteria decision-making approach to GIS-based land suitability evaluation. Int. J. Geogr. Inform. Syst. 7 (5), 407–424.

Pontius Jr., R.G., 1994. Modeling tropical land use change and assessing policies to reduce carbon dioxide release from Africa. Graduate Program in Environmental Science, SUNY-ESF, Syracuse, p. 177.

Pontius Jr., R.G., 2000. Quantification error versus location error in comparison of categorical maps. Photogrammetric Eng. Remote Sensing 66 (8), 1011–1016.

Schneider, L., Pontius Jr, R.G., 2001. Modeling land-use change: the case of the Ipswich watershed, Massachusetts, USA.. Agric. Ecosyst. Environ. 85, 83–94.

Swets, J.A., 1986a. Form of empirical ROCs in discrimination and diagnostic tasks: implications for theory and measurement of performance. Psychol. Bull. 99 (2), 181–198.

Swets, J.A., 1986b. Indices of discrimination for diagnostic accuracy: their ROCs and implied models. Psychol. Bull. 99 (1), 100–117.

Swets, J.A., 1988. Measuring the accuracy of diagnostic systems. Science 240, 1285–1293.

Turner, M.G., Constanza, R., Sklar, F.H., 1989. Methods to evaluate the performance of spatial simulation models. Ecol. Model. 48, 1–18.

Turner II, B.L., Skole, D., Sanderson, G., Fischer, L., Fresco, L., Leemans, R., 1995. Land-use and Land-cover Change: Science/Research Plan. IGBP/HDP, Stockholm and Geneva.

Veldkamp, A., Fresco, L.O., 1996. CLUE-CR: an integrated multi-scale model to simulate land use change scenarios in Costa Rica. Ecol. Model. 91 (1/3), 231.

Wagner, D.F., 1997. Cellular automata and geographic information systems. Environ. Plan. B 24, 219–234.

Wilkie, D.S., Finn, J.T., 1988. A spatial model of land use and forest regeneration in the Ituri forest of Northeastern Zaire. Ecol. Model. 41, 307–323.

Wu, F., Webster, C.J., 1998. Simulation of land development through the integration of cellular automata and multicriteria evaluation. Environ. Plan. B 25, 103–126.