Journal of the Economic Geographical Society of Korea
Vol.16, No.3, 2013(458~471)

Assessing the Metric to Measuring Land-Use Change Suitability

Oh Seok Kim*

토지 이용 변화 예측 모형의 정확도 검정을 위한 통계량 연구

김오석*

Abstract : This paper addresses the limitation of a map comparison metric entitled Figure of Merit through employing a simple land change model. The metric was originally designed to overcome limitations of other existing statistics, such as Kappa, when assessing predictive accuracy of land change models. A series of comparisons between null and predicted outcomes at multiple resolutions as well as a multi-resolution Figure of Merit analysis techniques of validation are compared for spatially segregated calibration and validation datasets. The Figure of Merit at the null resolution in this paper was 57%, although future research must be done to determine if this was simply a coincidence. A Figure of Merit greater than 50% would seem to represent a "Resolution of Merit" in that the Figure of Merit at that resolution becomes greater than the error. Thus, these two metrics should be used in tandem to assess predictive accuracy of a land change model.

Key Words : land change modeling, predictive accuracy, Figure of Merit, Resolution of Merit

요약 : 본 논문은 토지 이용 변화를 예측하는 계량 모형의 정확도 평가에 필수적인 통계량인 성능 지수를 심도 있게 이해하는 것을 목적으로 한다. 이 통계량은 기존의 토지 이용 변화 연구에서 소개된 예측 모형의 정확도를 평가하는 다른 통계량들 (예: 카파 통계량)의 단점을 보완하여 만들어진 것이나, 이 또한 계량 모형의 예측력을 명확하게 평가하고 해석하기에는 제한적이다. 본 논문에서는 성능 지수의 보다 명확한 해석을 위해서 결과물의 공간해상도를 고정해야 함을 밝히고, 그 특정 공간해상도를 "성능 해상도"라 정의한다. 성능 해상도는 예측오류가 현격하게 줄어들면서 계량 모형의 예측력이 증가하는 시점의 공간해상도를 일컫는다. 따라서 토지 이용 변화 예측 모형의 예측력을 정확하게 평가하기 위해 두 통계량, 즉 성능 지수와 성능 해상도를 함께 이용할 것을 제안한다.

주요어 : 토지 이용 변화 예측모형, 예측력, 성능 지수, 성능 해상도

1. Introduction

Urban land change models are geared towards explaining and simulating urban sprawl in a spatially explicit fashion (Dietzel and Clarke, 2007; Pontius Jr. and Malanson, 2005; Pontius Jr. and

* Research Associate, Division of Environmental Science and Ecological Engineering, College of Life Sciences, Korea University
(고려대학교 생명과학대학 환경생태공학부 선임연구원, oskim@korea.ac.kr)

Spencer, 2005; Silva and Clarke, 2002; Tan et al., 2005), and if a prediction turns out to be accurate then this could potentially support any important decision making in terms of urban policy and development. In this context, it is indispensable to systematically assess predictive accuracy of land change models whether to compare diverse modeling approaches or to evaluate underlying theories/hypotheses embedded in the prediction.

The dilemma, however, is such that we have limited toolsets to verify or validate such modeling outcomes. That is, although there are numerous theories to employ either in Economic or Urban Geography, such as the von Thünen (Angelsen, 2007), Forest Transition (Mather, 1992), Ricardian (Polsky, 2004), or Spatial-Interaction Model (Fotheringham, 1983), and have highly sophisticated modeling approaches, such as agent-based modeling (Brown et al., 2004; Evans et al., 2011), we must acknowledge the fact that without systematically verifying the outcomes of such modeling approaches, it is unlikely that research could have broader impact to the society. Pontius Jr. et al. (2008) present a novel work by arguing a systematic validation in land change modeling is crucial and pointing out there should be more measurements to account for modeling comparison. Although predicting future land change is not the only purpose of land change modeling, validating prediction outcomes are indeed useful in the discipline (Castella et al., 2007). With this in mind, this paper attempts to understand how a new map comparison technique behaves when different spatial resolution are employed.

Figure of Merit is relatively a new metric to measure predictive accuracy of different land change models. On the one hand, its logic is quite similar to the Kappa family (Pontius Jr., 2000; Pontius Jr. et al., 2008); on the other hand, the Producer's and User's Accuracies (Congalton, 1991) in remote sensing could be considered variants of the Figure of Merit (Kim, 2010). This implies that Figure of Merit could be potentially replace the tradition Kappa statistics that have been frequently used in the literature of land change modeling (Hall et al., 1995). However, it is important to note that there is little to no point evaluating a land change model's predictive accuracy based on a fixed range of Figure of Merit. That is, even though a range of Figure of Merit was generated based scientific experiments conducted on several different regions the Figure of Merit itself, by design, does not provide the full picture of a land change model's accuracy. This metric should always be accompanied by the net change of land-cover and/or land-use; otherwise, the measurement is incomplete, hence can be misleading. That is, if one is to specify and offer the range of Figure of Merit, then he or she must specify the range of net change of land-cover or land-use accordingly to accurately measure a model's predictive accuracy (Sloan and Pelletier, 2012; Pontius Jr. et al., 2008). Pontius Jr. et al. (2008) employ a Cartesian coordinate system to compare different land change models' predictive accuracy where the y-axis denotes Figure of Merits of numerous land change models, e.g., Conversion of Land Use and its Effects in Small regions ([CLUE-S], Verburg et al., 2002), and the x-axis denotes observed net changes of land-cover and land-use for each case study. Then it becomes vivid that a larger net change tends to guarantee a higher Figure of Merit, so the net areal change must be controlled. That being said, although the Figure of Merit is considered an improved measurement further experiments

should be conducted systematically to understand its performance.

There are numerous statistical methods to validate a land change model's predictive accuracy, other than the Figure of Merits or Kappa variants (Pontius Jr. and Millones, 2011; Robin et al., 2011; Peterson et al., 2008; Pontius Jr. et al., 2008; Pontius Jr. and Spencer, 2005). Some are geared towards validating accuracies of rank maps that indicate, for example, distribution of species or transition potential of land-use/cover change (Eastman et al., 2005; Peterson et al., 2008), while others are to evaluate decreasing or increasing accuracies in predicting with respect to different spatial and temporal resolutions (Pontius Jr. et al., 2008; Pontius Jr. and Spencer, 2005). As such, changing resolutions of a prediction and measuring the series of associated accuracies is one of the main approaches in accuracy assessment. Given that this paper aims to demonstrate the varying measures of Figure of Merit when different spatial resolutions are employed, the objective of this demonstration is to better understand the Figure of Merit so that land change modelers can further utilize the measurement in a precise manner. To do so, very simple land change modeling is conducted to provide test data for the statistic.

As there is a growing concern about how residential areas are affecting the sustainable use of freshwater (Runfola et al. 2013), the Greater Boston region was chosen as study area (Figure 1). Land-use/cover data of 1991 and 1999 available from the Massachusetts Office of Geographic Information (MassGIS) and United States Census data can only be matched for the years of 1990 and 2000, meaning that study of change must be restricted to that single time interval. Here, it is assumed that land-uss/cover does not change significantly within a

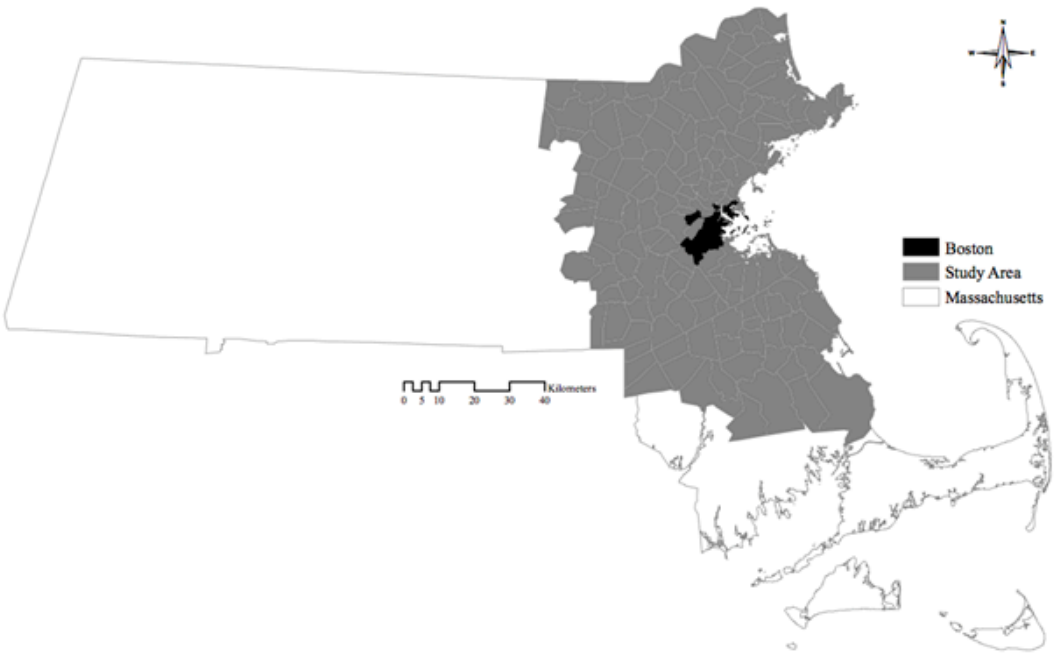

Figure 1. Study Area: Greater Boston

year, thus the difference of 1990 and 1991 (and 1999 and 2000) is negligible. Residential development comprises approximately 2/3 of total development in the study area, meaning that an effective understanding of processes of residential development is by far the most important component of an understanding of broader patterns of urban development in this region.

2. Data

The variables of median neighborhood commute time, median neighborhood income, terrain slope, and land-use/cover category are examined with respect to their effect on the probability of residential development. The data are explained as follows in more detail:

1) Median neighborhood commute time

1990 Census data were obtained at the tract level for disaggregated commute time. An excel spreadsheet algorithm was created to calculate a median commute time for each tract from this disaggregated commute time data, using a linear interpolation method along with a cumulative distribution function to interpolate estimated median times within category boundaries. Non-residential portions of census tracts were removed, and the triangulated irregular network interpolation is used to create a potential commute time field around existing residential areas in 1991. Figure 2 shows the percent built of residential areas (i.e., the ratio of residential and non-residential) classified by median commute time, and it appears there is little to no (linear) relationship between the two.

2) Median neighborhood income

1990 Census tract-level data on median household income were obtained from the U.S. Census Bureau; the triangulated irregular network interpolation was used to create a potential household income field around existing residential areas in 1991 in the same manner as for commute time. Figure 3 indicates the distribution of median in-

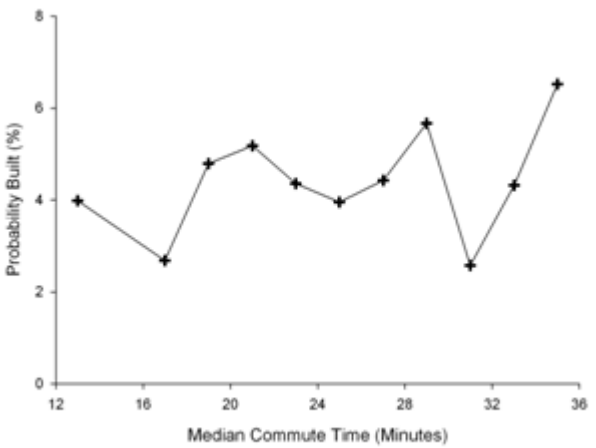

Figure 2. Development frequency by median commute time

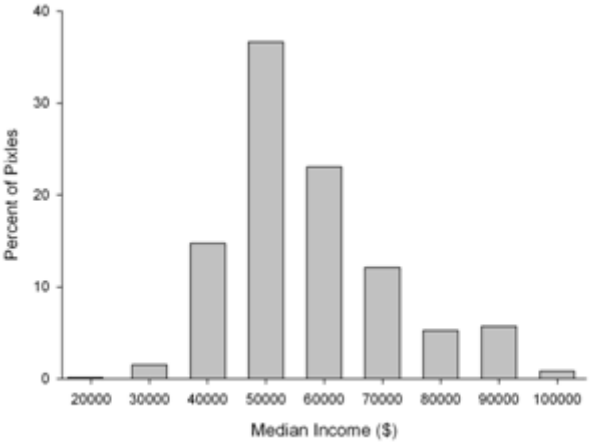

Figure 3. Study area distribution by median income

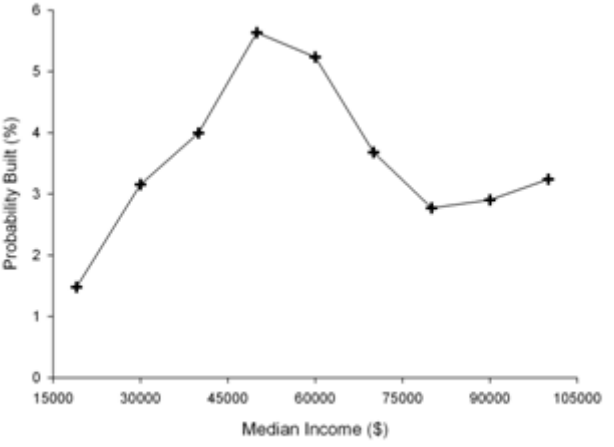

Figure 4. Development frequency by median income

come in the study area, and the 50,000 US Dollars band has the highest frequency. Figure 4 shows the relationship between the percent built and median income, and they seem to have a positive linear relationship if median income ranges from 20,000 US Dollars and 50,000 US Dollars. However, the trend is inverted at the range of 50,000 US Dollars and 80,000 US Dollars indicating a negative linear relationship. Finally, the trend is inverted again if median income is above 80,000 US Dollars.

3) Terrain slope

Data on slope were produced based on the digital elevation model obtained directly from the MassGIS database. According to Figure 5, the overall terrain of the study area seems suitable for developing built environments as there are not many slopy areas. Figure 6 shows the development frequency of residential areas classified by slope. The frequency and degree of slope appear having a positive linear

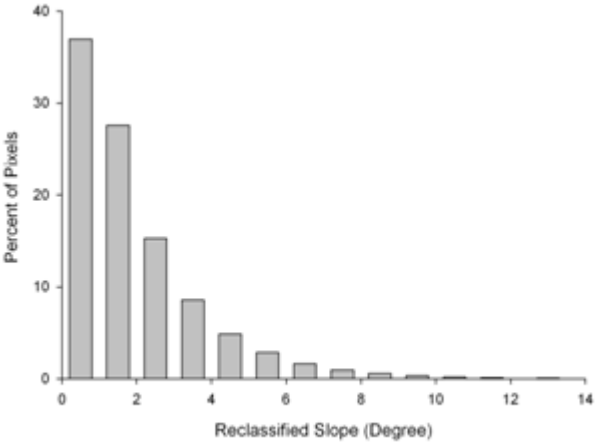

Figure 5. Study area distribution by slope

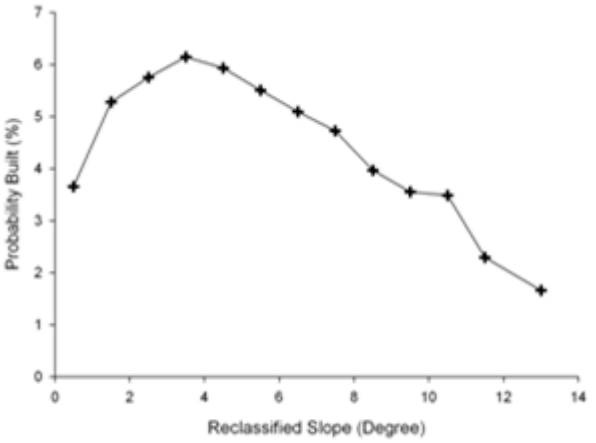

Figure 6. Development frequency by slope

relationship if slope ranges from 0 to 4 degrees, but the trend is inverted if slope is above 4 degrees indicating a negative linear relationship.

4) Developable and non-developable land categories

The existing land categories are aggregated into the following two categories: (1) developable and (2) non-developable. Forest, cropland, pasture, open land, and woody perennial categories are considered to be developable. Fresh and saltwater wetlands were found to have a development probability of less than .1% for any given pixel and as such are considered to be non-developable, along with non-residential built areas. This binary structure is necessary to run a logistic regression model, which will be explained in the method section.

3. Methods

Logistic regression was performed to determine linear correlation factors between commute time, income, and slope variables and the probability of new residential development within each dummy variable category. The development probability function developed through logistic regression was applied to the various calibration raster data layers using the raster calculator function in ArcMap to produce a simulated development favorability map for the calibration area. A favorability threshold was set for new development so that there was no disagreement in quantity of land-use/cover change (Pontius Jr., 2000) in between predicted and actual residential development in the calibration zone.

1) Logistic regression

Logistic regression detects a statistical relationship between three environmental variables and a binary event such as developable versus non-developable, where "1" indicates the developable and "0" indicates the non-developable. The basic assumption is that the probability of a binary event that takes the value of "1" follows a logistic curve, and it is expressed as follows:

$$P(y=1|X)=\frac{exp\sum BX}{1+exp\sum BX}$$

where y is a binary event, P is the probability of the binary event given column vector X, X is a column vector that give the values of the environmental variables, and B is a row vector that gives the estimated coefficients (Wooldridge 2006).

2) Calibration and validation

The study area was separated into calibration and validation zones, with the calibration set being performed using Middlesex and Norfolk counties, and the validation set being performed in Essex, Northern Bristol, and Plymouth counties. The validation process is essential in any predictive modeling, and usually a future observed value is employed to validate the associated predicted outcome (Pontius Jr. and Millones 2011, Pontius Jr. and Spencer 2005, Kim 2010). In this paper, the validation data are determined spatially. The logistic regression function determining development favorability was applied to the relevant raster data layers in the validation zone. Undeveloped pixels in 1991 with a development favorability above the construction threshold for the calibration zone are predicted as developed over the time interval. Figure of Merit and Multiple Resolution Analysis are applied to the validation area (Pontius et al., 2008). Disagreement between maps at pixel resolutions ranging from 1 to 1024 was calculated as sum of the absolute values of all pixels produced by the subtraction of one raster map from another, with the value of each pixel in these raster maps being the sum of the values of the pixels from which it was aggregated (at the highest resolution all prediction maps consisted of binary values, with "0" corresponding to non-residential and "1" corresponding to total, persistent, or new development).

3) Figure of Merit

Figure of Merit basically employs the concept of Venn diagram. That is to say, the statistic shows the predictive accuracy of a prediction map of t_1

by overlaying the map with the observed maps of t_0 and t_1. That is, unlike Kappa, Figure of Merit employs three maps because with two maps there is no way to measure a land change model's performance. Figure of Merit is numerically expressed as follows:

$$\text{Figure of Merit} = B/(A+B+C)$$

where A is a number of pixels for "error due to observed change predicted as persistence" (or misses), B is a number of pixels for "correct due to observed change predicted as change" (or hits), and C is a number of pixels for "error due to observed persistence predicted as change" (or false alarms). The value ranges from 0 to 100 percent, where 100 percent indicates perfect prediction (Pontius Jr. et al., 2008).

4) Multiple Resolution Analysis

Multiple Resolution Analysis (Pontius et al., 2008) assesses how closely in space the simulated change is to the observed change. It compares each simulation map to its reference map by aggregating pixels from finer to coarser resolutions. The raw pixels are aggregated into coarser pixels by computing the proportion of the three land-use/cover categories of the raw pixels within each coarser pixel (Pontius Jr., 2002), and such partial membership of the pixels is calculated by the Validate module in Idrisi (Eastman, 2012). Each finer resolution is completely nested within each coarser resolution where the side of each coarser pixel is a multiple of the side of a raw pixel in a geometric sequence such as 1, 2, 4, 8, ... , 1024. A more precise simulation has the finer null resolution than a less accurate simulation (Pontius Jr. and Malanson, 2005).

4. Results

No statistically significant correlation was found between median neighborhood income and residential development probability for the calibration zone. A weak positive correlation was found between commute time and development probability, and a somewhat stronger positive one found between terrain slope and development probability for most terrain categories. Based on the findings, a prediction map of residential areas is produced (Figure 7). The blurry areas of Figure 7 indicate the calibration zones that were used to fit the logistic regression, whereas the color-coded pixels portray the suitability of development. The reds are considered relatively unsuitable compared to the other colors; therefore, only the non-red pixels are predicted as change (Figure 7).

Figure 8 illustrates the error map of the prediction made by logistic regression. The grey indicates the land-use/cover category of persistence, meaning there had been no change between the two time periods. The amount of grey pixels is considerably larger than the other categories. The blue indicates errors due to either "observed change predicted as persistence" or "observed persistence predicted as change." Finally, the red indicates the "correct due to observed change predicted as change" (Pontius Jr. et al., 2008) and the Figure of Merit is about 2%.

The null resolution was linearly interpolated as approximately 71 pixels wide, corresponding to a pixel size of ~6,400 meters, and a Figure of Merit of ~57% (See Figures 9 and 10). Overall the model's

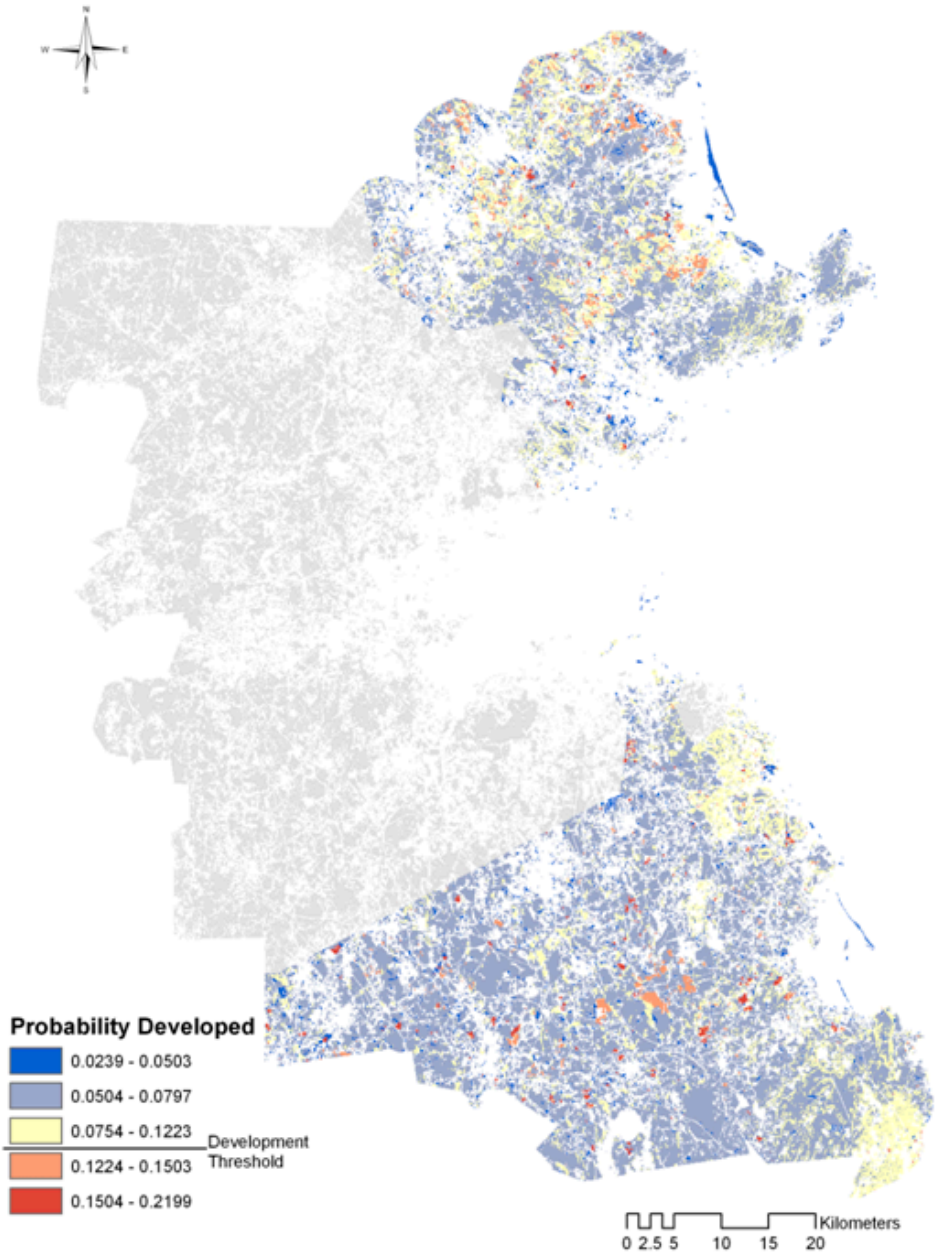

Figure 7. Predicted development favorability in validation areas based on logistic regression

predictive power was determined to be approximately 14% better than what would be expected from a random model preserving the same quantity of net change.

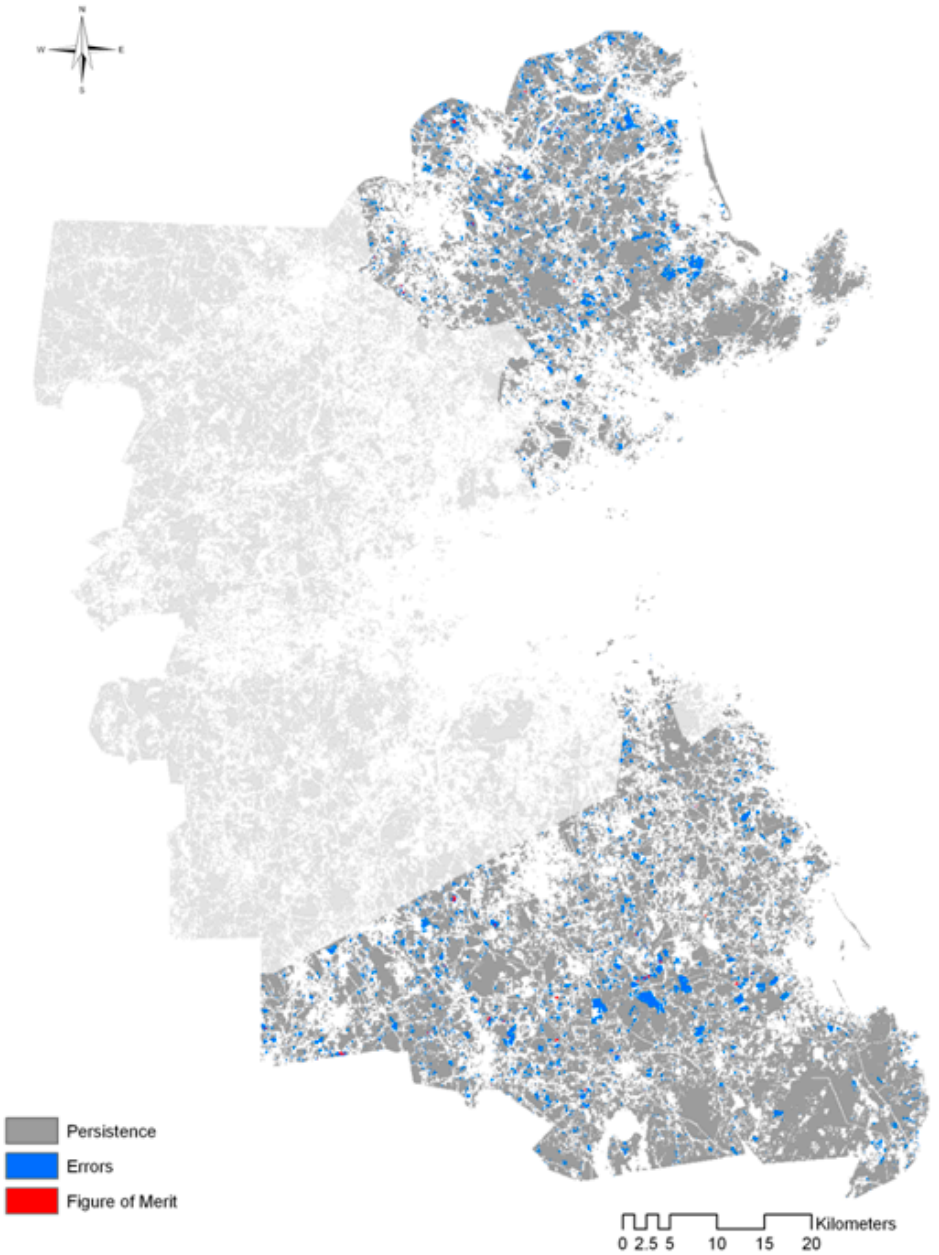

Figure 8. Error map in the validation zones

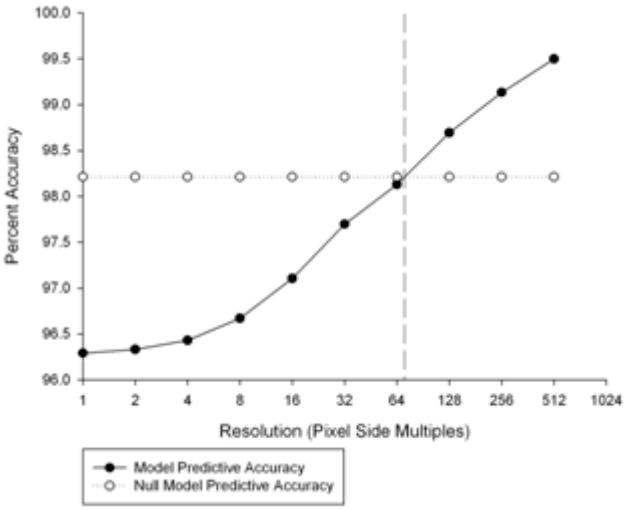

Figure 9. Null model and predictive accuracy of logistic regression (Null Resolution: 71 pixels or 6,400 meters)

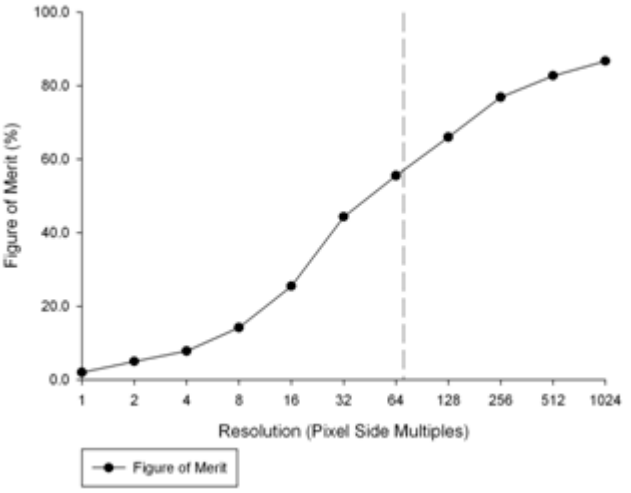

Figure 10. Figure of Merit by resolution (Null Resolution Figure of Merit: 57%)

5. Discussion

One interesting result can be salvaged from the validation procedure. There is currently no clear understanding of the relationship between Figure of Merit at multiple resolutions and the null-resolution. However, a Figure of Merit greater than 50% would seem to represent a "Resolution of Merit" in that the Figure of Merit at that resolution becomes greater than the error. A cursory intuitive assessment might also seem to indicate that null resolution would tend to correspond to a Figure of

Merit of exactly or approximately 50%, given that both null-resolution and Resolution of Merit seem to represent a threshold, at which the performance of the model somehow "breaks even." The Figure of Merit at the null-resolution in this instance was in fact 57%, although further research must be done to determine if this was simply a coincidence. Difference between null-resolution and Resolution of Merit was approximately 30%. At the moment it is not clear if this should be interpreted as a relatively large or small difference.

The methods in general do not yet have a mechanism to test a statistical difference of two (or more) distinct values generated by one accuracy assessing measurement. For instance, there is no way of comparing two Figure of Merits to test whether or not the numerical difference of the two values is statistically significant. The situation is identical even for the state-of-art map comparison measurement as the probability distribution of the measurement has not yet been found; hence it is unknown (Pontius Jr. and Millones, 2011). That is, fully evaluating a land change model's predictive accuracy in a statistical manner is unlikely at the moment, and more research has to be done to found the statistical distribution of the map comparison measurement as it is often done in spatial statistics (Kim and O'Kelly, 2008).

While certain aspects of this project could be seen as a case study in vaguely conceived, even questionably useful scholarship, it is hoped that some useful results can be found in the calculated relationship between Figure of Merit and null resolution, as well as observed empirical relationships between residential development probability and the variables of median income and terrain slope. The last two relationships in particular would seem to provide a relatively solid starting point for future model development and research. Determining whether this early understanding is correct, and to what degree it could improve our predictive capabilities with respect to urban growth, must await future research.

Acknowledgements

Clark Labs facilitated this work by creating the Geographic Information Systems and Remote Sensing (GIS/RS) software Idrisi.

References

Angelsen, A., 2007, "Forest Cover Change in Space and Time: Combining the von Thünen and Forest Transition Theories (WPS4117)," *World Bank Policy Research Working Paper*, p.1-43.

Brown, D.G., Walker, R.T., Manson, S., and Seto, K., 2004, Modeling Land-use and Land-cover Change, in Gutman, G., Janetos, A.C., Justice, C.O., Moran, E.F., Mustard, J.F., Rindfuss, R.R., Skole, D.L., Turner Jr., B.L., and Cochrane, M.A. (eds.), Land Change Science: Observing, Monitoring and Understanding Trajectories of Change on the Earth's Surface, Dordrecht: Springer Netherlands, pp.395-409.

Castella, J.-C., Kam, S.P., Quang, D.D., Verburg, P.H., and Hoanh, C.T., 2007, "Combining Top-down and Bottom-up Modelling Approaches of Land Use/cover Change to Support Public Policies: Application to Sustainable Management of Natural Resources in Northern Vietnam," *Land Use Policy* 24(3), pp.531-45.

Congalton, R.G., 1991, "A Review of Assessing the Accu-

racy of Classifications of Remotely Sensed Data," *Remote Sensing of Environment* 37(1), pp.35-46.

Dietzel C.K. and Clarke K.C., 2007, "Spatial differences in multi-resolution urban automata modeling," *Transactions in GIS* 11(1), pp.29-45.

Eastman, J.R., 2012, *IDRISI Selva Manual*, Worcester: Clark Labs, Clark University.

Eastman, J.R., Van Fossen, M.E., and Solórzano., L.A., 2005, Transition Potential Modeling for Land Cover Change, in Maguire, D., Batty, M., and Goodchild, M.F. (eds.), GIS, Spatial Analysis and Modeling, Redlands: ESRI Press, pp.339-68..

Evans, T.P., Phanvilay, K., Fox, J.M., and Vogler, J.B., 2011, "An Agent-based Model of Agricultural Innovation, Land-cover Change and Household Inequality: The Transition from Swidden Cultivation to Rubber Plantations in Laos PDR," *Journal of Land Use Science* 6(2-3), pp.151-73.

Fotheringham, A.S., 1983, "A new set of spatial-interaction models: the theory of competing destinations," *Environment and Planning A* 15(1), pp.15-36.

Hall, C.A.S., Tian, H., Qi, Y., Pontius Jr., R.G., and Cornell, J., 1995, "Modelling Spatial and Temporal Patterns of Tropical Land Use Change," *Journal of Biogeography* 22(4/5), pp.753-57.

Kim, O.S., 2010, "An Assessment of Deforestation Models for Reducing Emissions from Deforestation and Forest Degradation (REDD)," *Transactions in GIS* 14(5), pp. 631-54.

Kim, Y., and O'Kelly, M., 2008, "A Bootstrap Based Space-time Surveillance Model with an Application to Crime Occurrences," *Journal of Geographical Systems* 10(2), pp. 141-65.

Mather, A.S., 1992, "The Forest Transition," *Area* 24(4), pp.367-79.

Peterson, A,T., Papeş, M., and Soberón, J., 2008, "Rethinking Receiver Operating Characteristic Analysis Applications in Ecological Niche Modeling," *Ecological Modelling* 213(1), pp.63-72.

Polsky, C., 2004, "Putting Space and Time in Ricardian Climate Change Impact Studies: Agriculture in the U.S. Great Plains, 1969-1992." *Annals of the Association of American Geographers* 94(3), pp.549-64.

Pontius Jr., R.G., 2000, "Quantification Error Versus Location Error in Comparison of Categroical Maps," *Photogrammetric Engineering and Remote Sensing* 66(8), pp.1011-6.

Pontius Jr., R.G., 2002, "Statistical Methods to Partition Effects of Quantity and Location During Comparison of Categorical Maps at Multiple Resolutions," *Photogrammetric Engineering and Remote Sensing* 68(10), pp. 1041-9.

Pontius Jr., R.G., Boersma, W., Castella, J.-C., Clarke, K., Nijs, T., Dietzel, C., Duan, Z., Fotsing, E., Goldstein, N., Kok, K., Koomen, E., Lippitt, C.D., McConnell, W., Mohd Sood, A., Pijanowski, B., Pithadia, S., Sweeney, S., Trung, T.N., Veldkamp, A.T., Verburg, P.H., 2008, "Comparing the Input, Output, and Validation Maps for Several Models of Land Change," *Annals of Regional Science* 42, pp.11-37.

Pontius Jr., R.G. and Malanson, J., 2005, "Comparison of the Structure and Accuracy of Two Land Change Models," *International Journal of Geographical Information Science* 19(2), pp.243-65.

Pontius Jr., R.G. and Millones, M., 2011, "Death to Kappa: Birth of Quantity Disagreement and Allocation Disagreement for Accuracy Assessment," *International Journal of Remote Sensing* 32(15), pp.4407-29.

Pontius Jr., R.G. and Spencer, J., 2005, "Uncertainty in Extrapolations of Predictive Land-change Models," *Environment and Planning B: Planning and Design* 32(2), pp.211-30.

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., and Müller, M., 2011, "pROC: An Open-source Package for R and S+ to Analyze and Compare ROC Curves," *BMC Bioinformatics* 12(1), p.77.

Runfola, D.M., Polsky, C., Nicolson, C. Giner, N.M., Pontius Jr., R.G., Krahe, J., and Decatur A., 2013, "A Growing Concern? Examining the Influence of Lawn Size on Residential Water Use in Suburban Boston, MA, USA," *Landscape and Urban Planning* 119, pp.113-23.

Silva E.A. and Clarke K.C., 2002, "Calibration of the SLEUTH urban growth model for Lisbon and Porto, Portugal," *Computers, Environment and Urban Systems* 26, pp.525-552

Sloan, S. and Pelletier J., 2012, "How Accurately May We Project Tropical Forest-cover Change? A Validation of a Forward-looking Forest Baseline for Reduced Emissions from Deforestation and Degradation," *Global Environmental Change* 22(2), pp.440-53.

Tan, M., Li, X., Xie, H., and Lu, C., 2005, "Urban land expansion and arable land loss in China—a case study of Beijing-Tianjin-Hebei region," *Land Use Policy* 22(3), pp.187-96.

Verburg, P.H., Soepboer, W., Veldkamp, A., Limpiada, R., Espaldon, V., and Mastura, S.S.A., 2002, "Modeling the Spatial Dynamics of Regional Land Use: The CLUE-S Model," *Environmental Management* 30(3), pp. 391-405.

Wooldridge, J.M., 2006, *Introductory Econometrics: A Modern Approach*, 3rd edition, Cincinnati: South-Western College Publications.

교신: 김오석, 136-713, 서울시특별시 성북구 안암로 145 고려대학교 생명과학대학 환경생태공학부, 전화: 02-3290-3470, 이메일: oskim@korea.ac.kr

Correspondence: Oh Seok Kim, Division of Environmental Science and Ecological Engineering, College of Life Sciences, Korea University, 145 Anam-ro, Seungbuk-gu, Seoul, 136-713, Korea , Tel: 82-2-3290-3470, E-mail: oskim@korea.ac.kr

최초투고일 2013년 7월 13일
수 정 일 2013년 7월 30일
최종접수일 2013년 8월 3일