Landscape Ecol (2010) 25:561–572
DOI 10.1007/s10980-009-9433-x

RESEARCH PAPER

Derivation of a yearly transition probability matrix for land-use dynamics and its applications

Takenori Takada · Asako Miyamoto · Shigeaki F. Hasegawa

Received: 19 March 2008 / Accepted: 19 November 2009 / Published online: 10 December 2009

Abstract Transition matrices have often been used in landscape ecology and GIS studies of land-use to quantitatively estimate the rate of change. When transition matrices for different observation periods are compared, the observation intervals often differ because satellite images or photographs of the research site taken at constant time intervals may not be available. If the observation intervals differ, the transition probabilities cannot be compared without calculating a transition matrix with the normalized observation interval. For such calculation, several previous studies have utilized a linear algebra formula of the power root of matrices. However, three difficulties may arise when applying this formula to a practical dataset from photographs of a research site. We examined the first difficulty, namely that plural solutions could exist for a yearly transition matrix, which implies that there could be multiple scenarios for the same transition in land-use change. Using data for the Abukuma Mountains in Japan and the Selva el Ocote Biosphere Reserve in Mexico, we then looked at the second difficulty, in which we may obtain no positive Markovian matrix and only a matrix partially consisting of negative numbers. We propose a way to calibrate a matrix with some negative transition elements and to estimate the prediction error. Finally, we discuss the third difficulty that arises when a new land-use category appears at the end of the observation period and how to solve it. We developed a computer program to calculate and calibrate the yearly matrices and to estimate the prediction error.

Keywords Abukuma Mountains (Japan) · Computer program · Multiple scenarios · n-th power roots of matrices · Observation interval

Electronic supplementary material The online version of this article (doi:10.1007/s10980-009-9433-x) contains supplementary material, which is available to authorized users.

T. Takada (✉) · S. F. Hasegawa
Graduate School of Environmental Earth Science, Hokkaido University, N10W5, Kita-ku, Sapporo 060-0810, Japan
e-mail: takada@ees.hokudai.ac.jp

A. Miyamoto
Forestry and Forest Products Research Institute, Matsunosato 1, Tsukuba 305-8687, Japan

Introduction

In the late 1990s, a number of international research projects such as the Land Use and Cover Change project (Messerli 1997) began to examine the intensity of land-use change and the implications for global environmental change (Lambin and Geist

2006; Turner et al. 2007). These projects examined relatively large areas, including suburbs and cities, and focused on land-use change induced by human activities. The results indicated the necessity for intensive studies of land-use changes to determine the rate of changes and the associated driving forces. To quantitatively estimate the rate of land-use changes, satellite images, aerial photographs, and geographic information systems (GIS) have been widely used to identify and examine land-use and land cover change (Ehlers et al. 1990; Meyer and Turner 1991; Hathout 2002; Braimoh and Vlek 2004). The type, amount, and location of land-use changes can now be quantified, and some GIS software now provides a flexible environment for displaying, storing, and analyzing the digital data necessary to detect such changes. The software includes a procedure to classify the patterns of land use and land cover and to calculate transitions in areas of these classifications of land use. Area-based tables can be constructed using these procedures, allowing users to conveniently grasp the transition at a glance.

Probability-based transition tables, such as Markovian models or cross-tabulation matrices, are often obtained from area-based transition as a theoretical tool of landscape ecology. These tables provide a simple method for comparing the dynamics among research sites of different sizes and have been the focus of extensive theoretical studies (Usher 1981; Kachi et al. 1986; Gardner et al. 1987; Baker 1989a, b; Gustafson and Parker 1992; Lewis and Brown 2001; Pontius 2002; Pontius et al. 2004; Pontius and Cheuk 2006). Therefore, since the 1990s, many researchers have used Markovian models or cross-tabulation matrices (Meyer and Turner 1991; Mertens and Lambin 2000; Hathout 2002; Braimoh and Vlek 2004; Mundia and Aniya 2005; Braimoh 2006; Flamenco-Sandoval et al. 2007) to grasp dynamical characteristics of land use such as the diversity, driving forces, or scale dependence of land use (Turner et al. 1989; Turner 1990; Lo and Yang 2002).

Transition probability matrices are used to predict land-cover distributions and to generate land-cover projections as follows:

$$\mathrm{x}_{t+c} = \mathrm{x}_t \boldsymbol{A} \tag{1}$$

$\mathbf{x}_t$ is a 1-by-n row vector that gives the proportion of each category at the initial time t, where n is the number of categories in a land use classification. c is the number of years between the initial year t and the subsequent year of observation and $\mathbf{A}$ is a n-by-n matrix in which each element a_{ij} is the conditional probability that a pixel transition to category j by time $t + c$ given that it is category i at time t. Therefore, Eq. (1) means that the area vector of land-use categories after c years can be obtained by the product of that area vector in the current year and the transition matrix expressing the transition rule. Using this equation iteratively, the subsequent series of area vectors, i.e., $\mathrm{x}_{t+2c}, \mathrm{x}_{t+3c}, \mathrm{x}_{t+4c}, \ldots$, can be calculated to forecast and estimate future dynamics under the assumption that the transition rule is invariant. The probabilities of transitions from one land-use category to another usually differ among different observation periods. The differences are caused by historical, political, economic, or biological changes in the research sites, and comparisons among observation periods are the first step in understanding the background of dynamic changes. However, one of the problems that sometimes arise when comparing transition matrices is that observation intervals may differ among several observation periods because satellite images or photographs of the research site are not always prepared every year or at a constant time interval. If the observation intervals differ, the transition probabilities cannot be compared directly. For example, suppose that there are three aerial photographs, and the observation interval of the first two is 7 years and that of the last two is 14 years. Even if the observed transition probability is of the same magnitude, say 0.64, one cannot conclude that they are the same because 0.64 for the latter is equivalent to 0.8 for the former. Therefore, they should be adjusted on the basis of the same observation intervals and compared under the normalized observation interval.

Several authors (Mertens and Lambin 2000; Petit et al. 2001; Flamenco-Sandoval et al. 2007) have tried to construct yearly transition matrices using mathematical formulae from stochastic process theory (Cinlar 1975; Lipschutz 1979). Mertens and Lambin (2000) used four satellite images of East Province in Cameroon that were taken in 1973, 1986, 1991, and 1996 (at one 13-year and two 5-year intervals). They constructed 2×2 transition matrices with forest and non-forest land-use categories, obtained the yearly transition matrices, and compared them to detect the annual rate of changes in land cover. Flamenco-Sandoval et al. (2007) also conducted a similar analysis with 7 × 7 transition matrices in 1986, 1995, and 2000. Obtaining yearly transition matrices from the original transition matrices is becoming increasingly popular in land-use analysis. The above two papers established the mathematical formulae for obtaining the yearly transition matrix. However, it is not well known that several practical difficulties arise in the general way of obtaining the yearly matrix, although Flamenco-Sandoval et al. (2007) experienced one of the difficulties.

In the present paper, we clarify the three practical difficulties and the reasons they occur. The first difficulty is that the yearly transition matrix basically has plural solutions, which implies that multiple scenarios may exist for the same transition in land-use change. The second is that the yearly transition matrix could have some negative elements. We show two examples of land-use change, one from the Abukuma Mountains of central Japan and the other from Flamenco-Sandoval et al.'s (2007) study. The third is that a new land-use category may appear at the end of the observation period. A new land-use category could appear when the land-use change is very large. In this case, the transition matrix is not a square matrix, and we cannot apply the established formula. Finally, we propose ways of solving these problems and construct an algorithm to obtain the yearly transition matrix, presented by Mathematica and C++ programs.

Methods

To transform a transition matrix that has an arbitrary observation interval into one that has a normalized interval, the normalized interval is usually set as 1 year because of the seasonality of climatic conditions and the fact that arbitrary observation intervals could include prime numbers. Assuming that the transition rule is invariant within one observation interval, i.e., c years, and setting a yearly transition matrix as $\mathbf{B}$, Eq. (1) can be written as:

$$\mathbf{x_{t+c}} = \mathbf{x_t A} = \mathbf{x_t} \underbrace{\mathbf{BBB \cdots BB}}_{\mathbf{c\,times}} = \mathbf{x_t B^c}. \quad (2)$$

Therefore, the n-by-n yearly matrix, $\mathbf{B}$, is the c-th power root of an original transition matrix, $\mathbf{A}$.

It is simple to calculate the power root matrix numerically, as long as the solution is unique, because methods of numerical calculation that are used to find a root of higher-order simultaneous equations are common and are sometimes prepared as a toolbox in programming languages. However, this method is not adequate when there are many solutions because the method is heuristic and it is very difficult to obtain all of the numerical solutions. Therefore, two formulae on the c-th power root of a matrix are employed. One is

$$\mathrm{B} = \mathrm{A}^{\frac{1}{c}} = \mathrm{expm}^{\frac{1}{c}\mathrm{logm}\mathbf{A}}, \quad (3)$$

where expm is the matrix exponential and logm is the matrix logarithm (Mertens and Lambin 2000). The other was given in the recent paper of Flamenco-Sandoval et al. (2007) as follows:

$$\mathbf{B} = \mathbf{A}^{\frac{1}{c}} = \mathbf{U} \begin{pmatrix} (\lambda_1)^{1/c} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & (\lambda_n)^{1/c} \end{pmatrix} \mathbf{U}^{-1}$$
$$\mathbf{U} = \begin{pmatrix} \mathbf{u_1} & \cdots & \mathbf{u_n} \end{pmatrix}, \quad (4)$$

where λ_i is the i-th eigenvalue of matrix $\mathbf{A}$ and $\mathbf{u}_i$ is its corresponding eigenvector (n-by-1 column vector). This formula can be derived from stochastic process theory (Cinlar 1975; Lipschutz 1979) and is conditional as follows: "if an n-by-n matrix has n distinct eigenvalues and all of them are not equal to zero" (see Appendix 1 for the proof). The above authors obtained yearly transition matrices using the above two formulae.

We developed computer programs to solve Eq. (4). Here, we describe the calculation of several examples and clarify the three practical difficulties in obtaining the yearly transition matrices in the "Result" section.

Result

The first practical difficulty

In calculating the yearly transition matrix, we may encounter the difficulty of obtaining more than one yearly matrix. The number of c-th power roots of the matrix is easily obtained from Eq. (4), in which $\lambda^{1/c}$ represents the c-th power root of the scalar λ. Because λ could be a complex number, we can set

$\lambda = \mathrm{r}e^{i\theta} = r(\cos\theta + i\sin\theta) \quad (r > 0 \text{ and } 0 \le \theta < 2\pi)$, using polar coordinates. Therefore, $\lambda^{1/c} = r^{1/c}\left(\cos\frac{\theta+2\pi k}{c} + i\sin\frac{\theta+2\pi k}{c}\right)$ for $k = 0, 1,\ldots, c{-}1$ generally has c solutions, including complex numbers, as long as λ is not equal to zero. Therefore, the number of combinations for n λ_is is c^n, and the number of whole solutions is c^n.

We obtained all of the solutions of the following example. Suppose

$$\mathbf{B} = \begin{array}{c} \\ \mathrm{W} \\ \mathrm{G} \\ \mathrm{C} \\ \mathrm{B} \end{array}\begin{array}{c} \begin{array}{cccc} \mathrm{W} & \mathrm{G} & \mathrm{C} & \mathrm{B} \end{array} \\ \begin{pmatrix} 0.3 & 0.4 & 0.2 & 0.1 \\ 0.1 & 0.2 & 0.4 & 0.3 \\ 0.4 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.1 & 0.5 & 0.2 \end{pmatrix} \end{array} \tag{5}$$

and

$$\mathbf{B}^3 = \begin{pmatrix} 0.26 & 0.225 & 0.313 & 0.202 \\ 0.274 & 0.233 & 0.301 & 0.192 \\ 0.26 & 0.234 & 0.306 & 0.2 \\ 0.265 & 0.243 & 0.301 & 0.191 \end{pmatrix} = \mathbf{A}, \tag{6}$$

where W, G, C, and B represent woodland, grassland, cropland, and built-up area, respectively. If we have only two aerial photographs with an interval of 3 years, and **A** is a Markovian matrix of land-use classification obtained from the photographs, we can calculate the yearly transition matrix using Eq. (4), and one of the solutions must be the same as **B**. Using Eq. (4) and our computer program, we obtained $3^4 = 81$ solutions, including matrices that have imaginary numbers or negative values as a part of elements in the matrix. The elements of the yearly matrix should be real and range from 0 to 1 because

$$\mathbf{B}_1 = \begin{pmatrix} 0.3 & 0.4 & 0.2 & 0.1 \\ 0.1 & 0.2 & 0.4 & 0.3 \\ 0.4 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.1 & 0.5 & 0.2 \end{pmatrix},$$

$$\mathbf{B}_2 = \begin{pmatrix} 0.366 & 0.106 & 0.347 & 0.180 \\ 0.316 & 0.404 & 0.172 & 0.109 \\ 0.261 & 0.125 & 0.313 & 0.301 \\ 0.071 & 0.370 & 0.397 & 0.162 \end{pmatrix}$$

$$\mathbf{B}_3 = \begin{pmatrix} 0.098 & 0.211 & 0.413 & 0.279 \\ 0.388 & 0.089 & 0.329 & 0.194 \\ 0.260 & 0.296 & 0.209 & 0.235 \\ 0.347 & 0.337 & 0.285 & 0.032 \end{pmatrix}$$

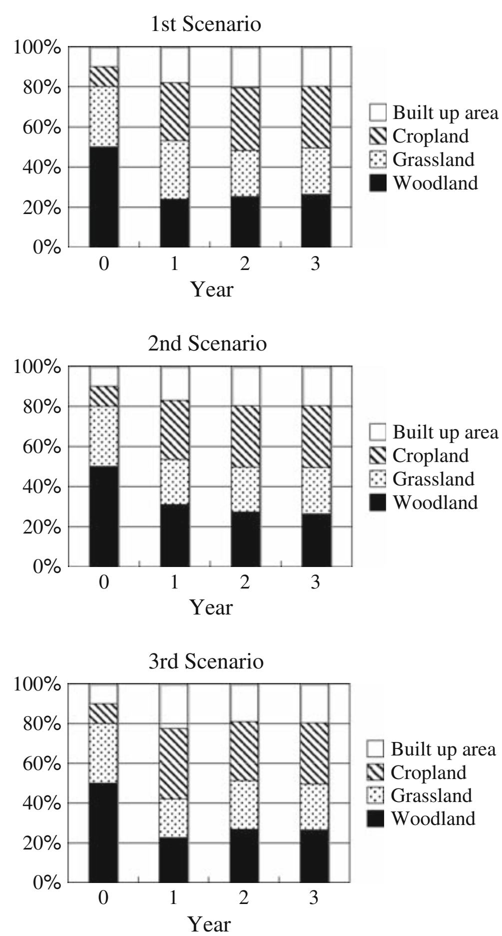

Fig. 1 The time course of change in area proportions in 3 years for three scenarios. Beginning from the initial area distribution (0.5, 0.3, 0.1, 0.1), the dynamics of the area distributions are shown using each yearly transition matrix. The top is $\mathbf{B_1}$, the middle is $\mathbf{B_2}$, and the bottom is $\mathbf{B_3}$. The third scenario shows the different changes in area distributions from the first scenario in transition years; all of the scenarios lead to the same area distribution in the third year

they are transition probabilities. Only three solutions of the 81 satisfy the criteria; these are:

$$A^{\frac{1}{3}} = \begin{pmatrix} 0.3 & 0.4 & 0.2 & 0.1 \\ 0.1 & 0.2 & 0.4 & 0.3 \\ 0.4 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.1 & 0.5 & 0.2 \end{pmatrix}, \begin{pmatrix} 0.366 & 0.106 & 0.347 & 0.180 \\ 0.316 & 0.404 & 0.172 & 0.109 \\ 0.261 & 0.125 & 0.313 & 0.301 \\ 0.071 & 0.370 & 0.397 & 0.162 \end{pmatrix}, \begin{pmatrix} 0.098 & 0.211 & 0.413 & 0.279 \\ 0.388 & 0.089 & 0.329 & 0.194 \\ 0.260 & 0.296 & 0.209 & 0.235 \\ 0.347 & 0.337 & 0.285 & 0.032 \end{pmatrix}, \quad (7)$$

which are referred to as the first, second, and third scenarios ($\mathbf{B_1}$, $\mathbf{B_2}$, and $\mathbf{B_3}$). $\mathbf{B_1}$ actually agrees with $\mathbf{B}$, and we obtained two additional solutions for the transition matrices. This means that there are three scenarios that can lead to the same transition. Actually, starting from an initial area frequency distribution of $\mathbf{x}_0 = (0.5, 0.3, 0.1, 0.1)$, the dynamics of the area distributions $\mathbf{x}_0, \mathbf{x}_0\mathbf{B}_i, \mathbf{x}_0\mathbf{B}_i^2, \mathbf{x}_0\mathbf{B}_i^3$, which are 1-by-$n$ row vectors, can be calculated using each yearly transition matrix (Fig. 1). Each scenario shows a different change in area distribution in transition years and, of course, all of the scenarios lead to the same area distribution in the third year.

The second practical difficulty

The next difficulty is the extreme opposite of the first. We may obtain no positive Markovian matrix and only a matrix with partly negative or complex numbers, even if the original matrix obtained from the GIS data was positive. We show two examples below.

(1) The Abukuma Mountains

The study area (10,000 ha; 10 × 10 km) is located in the southern part of the mountainous Abukuma region in central Japan, located at approximately 36°53′–36°59′N and 140°32′–140°39′E. Data on past forest landscapes were obtained from aerial photographs at four points in time, in 1947, 1962, 1975, and 1997. The land-use patterns were classified into five categories: coniferous planted forest, secondary forest, old-growth forest, grassland, and other land use. All land-use maps were prepared as vector maps using the GIS software TNTmips Ver. 6.8 (MicroImages Inc.). We then constructed three area-based transition tables (Table 1) and probability-based transition tables, i.e., transition matrices (Table 2).

We obtained the yearly transition matrices in the Abukuma Mountains using our computer program. The matrix between 1947 and 1962 has 15^5 solutions, elements of which could include negative and complex numbers, as explained previously. We omitted solutions with negative or complex numbers after obtaining the whole solutions and could not obtain any positive solutions. The computer program was then modified to detect solutions with real elements ≥ -0.1, taking into account approximately positive solutions.

At the first stage (1947–1962), there was only one valid solution ($b_{ij} \geq -0.1$, where b_{ij} is the transition probability from the i-th land-use category to the j-th in yearly matrix $\mathbf{B}$) among the 15^5 solutions (Table 3a); all other solutions contained elements ≤ -0.1 and/or complex numbers. At both the second and third stages (Table 3b and c, respectively), we similarly detected a single appropriate solution among the 13^5 and 22^5 solutions, respectively. Most diagonal elements of the yearly transition matrices are >90%, which means that the land-use changes in all of the observation periods are very slow on a yearly basis.

We presumed that the obtained matrices with negative elements close to zero would be appropriate solutions. To confirm that these negative elements might actually be considered as zero or approximately small values, we constructed the n-by-n calibrated yearly transition matrices ($\mathbf{B_{calibrated}}$) such that negative elements are zero and all column sums are equal to 1:

$$\begin{cases} \text{if } b_{ij} < 0, \text{ then } b_{ij,\text{calibrated}} = 0 \\ \text{if } b_{ij} > 0, \text{ then } b_{ij,\text{calibrated}} = b_{ij} \Big/ \sum_{\Theta_i} b_{ij}, \end{cases} \quad (8)$$

where $\Theta_j = \{i | b_{ij} > 0\}$ and $b_{ij,\text{calibrated}}$ is the transition probability from the i-th land-use category to the j-th in $\mathbf{B_{calibrated}}$ (Table 3).

According to Eq. (1), the area vector at time $t + c$ (the final year of the observation period) should be equal to the product of the c-th power of a calibrated yearly matrix and the area vector at time t (the initial year of the observation period). Therefore, we

Table 1 Area-based transition tables among land-use categories in the Abukuma Mountains

(a) From 1947 to 1962	1962					Total in 1947
1947	Coniferous planted forest	Secondary forest	Old forest	Grassland	Other	
Coniferous planted forest	1,642.4	143.8	0.0	19.6	1.3	1,807.1
Secondary forest	852.6	4,940.0	0.0	87.0	67.5	5,947.0
Old forest	211.7	164.7	437.4	0.6	2.4	816.7
Grassland	96.1	483.4	0.0	85.6	26.1	691.2
Other	1.1	7.3	0.0	0.9	741.5	750.9
Total	2,803.9	5,739.1	437.4	193.8	838.7	10,012.9
(b) From 1962 to 1975	1975					Total in 1962
1962	Coniferous planted forest	Secondary forest	Old forest	Grassland	Other	
Coniferous planted forest	2,656.5	77.6	0.0	55.2	14.6	2,803.9
Secondary forest	1,805.1	3,433.0	2.5	391.3	107.3	5,739.1
Old forest	49.3	65.5	228.8	93.8	0.0	437.4
Grassland	92.0	24.3	0.0	77.4	0.0	193.8
Other	33.0	5.5	0.0	11.2	789.1	838.7
Total	4,635.9	3,606.0	231.3	626.8	911.0	10,012.9
(c) From 1975 to 1997	1997					Total in 1975
1975	Coniferous planted forest	Secondary forest	Old forest	Grassland	Other	
Coniferous planted forest	3,249.8	1,265.7	18.2	29.7	70.3	4,633.7
Secondary forest	1,110.8	2,158.6	31.2	190.3	115.4	3,606.4
Old forest	45.9	39.7	145.3	0.0	0.9	231.9
Grassland	308.6	165.3	3.8	124.6	26.1	628.5
Other	127.6	159.3	3.3	76.5	545.6	912.4
Total	4,842.8	3,788.7	201.9	421.2	758.4	10,012.9

The numbers in each cell represent the area of transition from one category to another (ha). The total area is approximately 100 km^2

compared the observed area vector at time $t + c$ with this product and calculated the percentage errors in the three observation periods. The first and second rows in each table in Table 4 are the observed area vectors in the initial and final years in Table 1, i.e., $\mathbf{x}_t$ and $\mathbf{x}_{t+c}$, respectively. The estimated $\mathbf{x}_{t+c}$ in Eq. (1) is calculated in the third row using the observed area vector in the initial year and the calibrated yearly transition matrix $\left(\mathbf{x}_t \mathbf{B}^c_{\text{calibrated}}\right)$. The difference between the observed and estimated $\mathbf{x}_{t+c}$ is calculated in the fourth row. The calibrated matrices were reasonably good estimators of the area vectors in the final year, and all of the errors were <1%.

(2) The Selva el Ocote Biosphere Reserve in Mexico studied by Flamenco-Sandoval et al. (2007)

They examined land-use shift in the Selva el Ocote Biosphere Reserve in Mexico with 7 × 7 transition matrices in 1986, 1995, and 2000. Their seven categories of land use were agriculture and pasture (A/P), temperate forest (TemF), tropical forest (TroF), shrub and savanna (S/S), second-growth temperate forest (SGTemF), second-growth tropical forest (SGTtroF), and second-growth forest with slash and burn agriculture (SGF + SBA). They constructed two transition matrices, from 1986 to 1995 and from 1995 to 2000, and obtained the yearly matrices to discern whether they were significantly different by a log-linear statistical test.

We also obtained the yearly matrices using our computer program. At the first and second periods

Table 2 Transition matrices among land-use categories in the Abukuma Mountains

(a) From 1947 to 1962	1962				
1947	Coniferous planted forest	Secondary forest	Old forest	Grassland	Other
Coniferous planted forest	0.909	0.080	0	0.011	0.001
Secondary forest	0.143	0.831	0	0.015	0.011
Old forest	0.259	0.202	0.536	0.001	0.003
Grassland	0.139	0.699	0	0.124	0.038
Other	0.002	0.010	0	0.001	0.988
(b) From 1962 to 1975	1975				
1962	Coniferous planted forest	Secondary forest	Old forest	Grassland	Other
Coniferous planted forest	0.947	0.028	0	0.020	0.005
Secondary forest	0.315	0.598	0	0.068	0.019
Old forest	0.113	0.150	0.523	0.214	0
Grassland	*0.475*	0.125	0	0.400	0
Other	0.039	0.007	0	0.013	0.941
(c) From 1975 to 1997	1997				
1975	Coniferous planted forest	Secondary forest	Old forest	Grassland	Other
Coniferous planted forest	0.701	0.273	0.004	0.006	0.015
Secondary forest	0.308	0.599	0.0009	0.053	0.032
Old forest	0.198	0.171	0.627	0	0.004
Grassland	*0.491*	0.263	0.006	0.198	0.042
Other	0.140	0.175	0.004	0.084	0.598

The numbers in each cell represent transition probabilities from one land-use category to another

(1986–1995 and 1995–2000), no positive solutions and only one appropriate solution ($b_{ij} \geq -0.1$) were found among the 9^7 and 5^7 solutions (Table 5); all other solutions contained elements ≤ -0.1 and/or complex numbers. At the first period, the negative elements were very small, on the order of 10^{-6}, and do not appear in Table 5a explicitly, whereas five negative elements appeared at the second period (italicized cells in Table 5b). Most diagonal elements of the yearly transition matrices are >90%, which means that the land-use changes in all of the observation periods are very slow on a yearly basis.

Table 5c shows the calibrated matrix of the second period, including the error estimation (that of the first period is the same in the range of three decimal digits because of extremely small negative elements; see Table 5a). The result of error estimation is very low, at most 1.1%, implying that the calibrated matrix is a reasonably good estimator of the area vectors in the final year.

The third practical difficulty

The other problem is that we cannot obtain a yearly transition matrix if the original matrix is not a square matrix, e.g.,

$$\mathbf{C} = \begin{array}{c} \\ \mathrm{W} \\ \mathrm{G} \\ \mathrm{C} \end{array} \begin{array}{c} \begin{array}{cccc} \mathrm{W} & \mathrm{G} & \mathrm{C} & \mathrm{B} \end{array} \\ \begin{pmatrix} 0.31 & 0.40 & 0.27 & 0.02 \\ 0.43 & 0.23 & 0.33 & 0.01 \\ 0.32 & 0.45 & 0.22 & 0.01 \end{pmatrix} \end{array} \quad (9)$$

where W, G, C, and B represent woodland, grassland, cropland, and built-up area, respectively. Equation (2) cannot be applied to the matrix because it is not a square matrix. The matrix means, in practice, that the "built-up" area in the land-use classification appears for the first time at the end of the observation period. The appearance of a new land-use category could occur in cases in which human activity is strong and

Table 3 Yearly and calibrated transition matrices in the Abukuma Mountains

(a) From 1947 to 1962	1962				
1947	Coniferous planted forest	Secondary forest	Old forest	Grassland	Other
Coniferous planted forest	0.9931	0.0052	0	0.0016	0.0000
Secondary forest	0.0109	0.9860	0	0.0023	0.0008
Old forest	0.0225 (0.0224)	0.0189 (0.0189)	0.9592 (0.9585)	−0.0007 (0.0000)	0.0002 (0.0002)
Grassland	0.0105	0.1193	0	0.8652	0.0002
Other	0.0000	0.0008	0	0.0002	0.9992
(b) From 1962 to 1975	1975				
1962	Coniferous planted forest	Secondary Forest	Old forest	Grassland	Other
Coniferous planted forest	0.9949	0.0026	0	0.0022	0.0004
Secondary forest	0.0287	0.9597	0	0.097	0.0018
Old forest	−0.0005 (0.0000)	0.0163 (0.0163)	0.9514 (0.9507)	0.0330 (0.0329)	−0.0002 (0.0000)
Grassland	0.0530 (0.0530)	0.0178 (0.0178)	0.0 (0.0)	0.9296 (0.9293)	−0.0004 (0.0000)
Other	0.0027	0.0005	0	0.0015	0.9953
(c) From 1975 to 1997	1997				
1975	Coniferous planted forest	Secondary forest	Old forest	Grassland	Other
Coniferous planted forest	0.9800 (0.9790)	0.0202 (0.0202)	0.0002 (0.0002)	−0.0010 (0.0000)	0.0006 (0.0006)
Secondary forest	0.0199	0.9708	0.0006	0.0066	0.0021
Old forest	0.0116 (0.0116)	0.0104 (0.0104)	0.9789 (0.9780)	−0.0008 (0.0000)	0.0000 (0.0000)
Grassland	0.0504	0.0190	0.0004	0.9263	0.0039
Other	0.0036	0.0106	0.0001	0.0094	0.9763

The numbers in parentheses represent the calibrated transition probabilities

the cases are sufficiently probable in areas that suffer a dramatic change in land use such as suburbs or exploited forest. Therefore, the next question is how to obtain a yearly transition matrix in the case where a new land-use category appears.

If we can assume that the newly appeared "built-up" area remains a "built-up" area during the observation period, we could obtain the yearly transition matrix by setting the fourth row in Eq. (9) as (0, 0, 0, 1). For example, when the observation period is 3 years, the cubic root of Eq. (9), the yearly Markovian matrix is calculated as follows:

$$\mathbf{C}^{1/3} = \begin{array}{c c} & \begin{array}{cccc} \mathrm{W} & \mathrm{G} & \mathrm{C} & \mathrm{B} \end{array} \\ \begin{array}{c} \mathrm{W} \\ \mathrm{G} \\ \mathrm{C} \\ \mathrm{B} \end{array} & \begin{pmatrix} 0.114 & 0.529 & 0.344 & 0.013 \\ 0.575 & -0.014 & 0.440 & 0.000 \\ 0.398 & 0.607 & -0.006 & 0.000 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}.$$

Here, we are confronted with the second difficulty again and could obtain the calibrated yearly Markovian matrix using Eq. (8). If an obtained matrix includes many and large negative elements, the result of error estimation would become large. Then, we should not adopt the solution because the assumption of setting the fourth row in Eq. (9) to (0, 0, 0, 1) does not hold.

Discussion

A problem in the examination of land-use changes using satellite images or aerial photography is that photographs are sometimes lacking such that transition matrices with constant observation intervals cannot be obtained. Thus, we developed a computer program to calculate a yearly transition matrix from an original transition matrix that has an arbitrary observation period. In comparing differences among

Table 4 Error estimation between the observed area and estimated area using a calibrated transition matrix

Area vector (ha)	Coniferous planted forest	Secondary forest	Old forest	Grassland	Other
(a) From 1947 to 1962 in the Abukuma Mountains; c = 15					
At time t	1,807.1	5,947.0	816.7	691.2	750.9
At time $t + 15$	2,803.9	5,739.1	437.4	193.8	838.7
Estimated area vector	2,802.8	5,741.6	432.6	196.5	839.5
Error (%)	−0.01	0.02	−0.05	0.03	0.01
(b) From 1962 to 1975 in the Abukuma Mountains; c = 13					
At time t	2,803.9	5,739.1	437.4	193.8	838.7
At time $t + 13$	4,635.9	3,606.0	231.3	628.8	911.0
Estimated area vector	4,637.3	3,605.6	229.4	627.0	913.6
Error (%)	0.01	0.00	−0.02	−0.02	0.03
(c) From 1975 to 1997 in the Abukuma Mountains; c = 22					
At time t	4,633.7	3,606.4	231.9	628.5	912.4
At time $t + 22$	4,842.8	3,788.7	201.9	421.2	758.4
Estimated area vector	4,792.3	3,786.9	198.8	474.5	760.4
Error (%)	−0.50	−0.02	−0.03	0.53	0.02

transition matrices that have different observation intervals, Eq. (4) provides a useful tool and avoids misunderstandings of the processes of land-use change. An example is the 10-year matrices (10th power matrices of yearly matrices) for each period in the Abukuma Mountains, obtained using the calibrated yearly transition matrices (Table 6). These make it easier to understand the rate of land-use change intuitively because the yearly rate of change in the forest area is usually very slow. The transition probabilities from "grassland" to "coniferous planted forest" are almost the same, both at the second and third stages in the original matrices (italicized cells in Table 2b, c), whereas those in the 10-year matrices differ (italicized cells in Table 6b, c) and their order is reversed. Therefore, the normalization of transition matrices with different observation intervals is necessary to accurately estimate land-use changes and to understand the cultural and historical causes of the changes, i.e., motive forces. Without such calculations, the transition probabilities of the original matrices might be misread.

The present paper has described three practical difficulties in obtaining the yearly transition matrix. One is that the number of appropriate solutions for yearly matrices could be >1, as in Eq. (7). This implies that there could be multiple scenarios that lead to the same aerial photographs in the final year of the observation period, which could be caused by different driving forces among the scenarios, i.e., political reasons (e.g., adoption of new ordinances), economic reasons (e.g., price reductions in the timber market), or environmental reasons (e.g., soil erosion). Mathematically, there would be no way to identify which matrix (or which driving force) is correct. To determine the correct transition matrix among plural solutions, extra aerial photographs are required for a middle year (m) during the observation period (c years; $c > m$). Using the observed area distribution in the initial year, $\mathbf{x_0}$, the discrepancy between the observed area distribution from the extra photograph ($\mathbf{x}_m$) and the expected area distribution $\mathbf{x}_0\mathbf{B}_i^m$ can be calculated for each scenario ($\mathbf{B}_i$). The most appropriate scenario can then be chosen such that the norm of $\left\|\mathbf{x}_0\mathbf{B}_i^m - \mathbf{x}_m\right\|$ is minimized.

The second difficulty is that only a yearly transition matrix with negative elements close to zero may be obtained, rather than transition matrices with positive elements. Previous studies did not refer explicitly to these two points. Mertens and Lambin (2000) calculated the yearly transition matrices of 2×2 matrices and obtained a positive matrix. Using

Table 5 Yearly transition matrices recalculated from Flamenco-Sandoval et al. (2007)

(a) From 1986 to 1995	1995						
1986	A/P	TemF	TroF	S/S	SGTemF	SGTtroF	SGF + SBA
A/P	0.995	0.000	0.000	0.000	0.000	0.005	0.000
TemF	0.001	0.997	0.000	0.000	0.002	0.000	0.000
TroF	0.001	0.000	0.988	0.000	0.000	0.009	0.002
S/S	0.006	0.000	0.000	0.994	0.000	0.000	0.000
SGTemF	0.000	0.000	0.000	0.000	1.000	0.000	0.000
SGTtroF	0.006	0.000	0.000	0.000	0.000	0.994	0.000
SGF + SBA	0.001	0.000	0.000	0.000	0.000	0.006	0.993

(b) From 1995 to 2000	2000						
1995	A/P	TemF	TroF	S/S	SGTemF	SGTtroF	SGF + SBA
A/P	0.959	0.000	0.000	0.003	0.001	0.036	0.000
TemF	0.010	0.971	0.000	−0.000	0.020	*−0.001*	0.000
TroF	0.013	−0.000	0.913	−0.000	*−0.001*	0.057	0.019
S/S	0.008	−0.000	−0.000	0.993	−0.000	*−0.001*	−0.000
SGTemF	0.037	−0.000	−0.000	−0.000	0.966	*−0.003*	−0.000
SGTtroF	0.069	0.000	−0.000	0.002	*−0.000*	0.929	−0.000
SGF + SBA	0.059	−0.000	0.000	−0.001	0.028	0.127	0.787

(c) Calibrated matrix from 1995 to 2000	2000						
1995	A/P	TemF	TroF	S/S	SGTemF	SGTtroF	SGF + SBA
A/P	0.959	0.000	0.000	0.003	0.001	0.036	0.000
TemF	0.010	*0.970*	0.000	0.000	0.020	0.000	0.000
TroF	0.013	0.000	*0.911*	0.000	0.000	0.057	0.019
S/S	0.008	0.000	0.000	*0.992*	0.000	0.000	0.000
SGTemF	0.037	0.000	0.000	0.000	*0.963*	0.000	0.000
SGTtroF	0.069	0.000	0.000	0.002	0.000	0.929	0.000
SGF + SBA	0.059	0.000	0.000	0.000	0.028	0.127	*0.786*
Error (%)	0.00	−0.40	−0.62	1.11	−0.07	0.18	−0.52

our computer program, we confirmed only one positive solution existed in their case. Similarly, Flamenco-Sandoval et al. (2007) obtained two 7 × 7 yearly matrices, one of which included a few small negative elements, as shown in the present paper. This could occur because of the non-stationarity of the Markov process in land-use change or errors such as mistaken image analysis in land-use classifications. If the transition among land-use categories is not stationary during the observation period, the possibility of not obtaining positive yearly transition matrices would increase. The percentage error between the observed and estimated area vectors in Table 4 would express the index of the non-stationarity of the Markovian process. Furthermore, an improbable transition can be picked up from photographs because of the precision of GIS software. To avoid mistakes in classification, a technique to compute the transition probabilities for soft-classified pixels would be useful, as proposed by Pontius and Cheuk (2006). From our experience, small negative elements in yearly transition matrices are likely to occur when many zero or small elements are included in the original matrix. For example, in the second period in the Abukuma Mountains (Table 2b), there are six zero elements among the 5 × 5 elements

Table 6 Ten-year transition matrices in the Abukuma Mountains

(a) From 1947 to 1962	1962				
1947	Coniferous planted forest	Secondary forest	Old forest	Grassland	Other
Coniferous planted forest	0.936	0.054	0	0.009	0.000
Secondary forest	0.100	0.879	0	0.013	0.008
Old forest	0.189	0.152	0.655	0.002	0.002
Grassland	0.097	0.631	0	0.241	0.001
Other	0.001	0.006	0	0.001	0.992
(b) From 1962 to 1975	1975				
1962	Coniferous planted forest	Secondary forest	Old forest	Grassland	Other
Coniferous planted forest	0.957	0.022	0	0.016	0.004
Secondary forest	0.253	0.671	0.000	0.061	0.015
Old forest	0.075	0.126	0.603	0.195	0.001
Grassland	*0.398*	0.112	0.000	0.489	0.002
Other	0.030	0.005	0	0.011	0.954
(c) From 1975 to 1997	1997				
1975	Coniferous planted forest	Secondary forest	Old forest	Grassland	Other
Coniferous planted forest	0.825	0.162	0.002	0.005	0.007
Secondary forest	0.171	0.764	0.005	0.043	0.018
Old forest	0.104	0.091	0.801	0.002	0.001
Grassland	*0.344*	0.154	0.003	0.470	0.028
Other	0.054	0.093	0.002	0.063	0.789

because transitions among categories are usually slow in forest ecosystems.

There is a kind of trade-off between the first and second problems. Matrix **A** (Eq. (6), original matrix) includes sufficiently large positive elements, indicating that there are large transitions among the land-use categories and increasing the possibility of obtaining plural solutions. This is likely to occur for land-use changes in cities or agricultural areas, where the effect of agricultural innovation or human activity is large. In contrast, in natural forests, transitions are usually slow, and the original matrices include many small elements. In this case, a yearly transition matrix with negative elements close to zero is sometimes obtained.

To solve the above difficulties, an adequate procedure is to find all the solutions and then select the appropriate solutions, including those with a few small negative elements. If a solution is unique, that is the solution we want to obtain. If a solution includes negative elements, we should check whether it can be calibrated and the extent to which the calibrated matrix causes the percentage error between the observed and estimated area vectors (Table 4). We developed a computer program using Mathematica (Wolfrum Research, Inc.) and C++, which can be accessed on the website http://hosho.ees.hokudai.ac.jp/~takada/enews.html. The procedure used in our computer program could be easily incorporated into popular GIS software as a standard subprogram. The subprogram can be used for the comparison of yearly transition matrices among different observation periods when temporal change of exogenous driving factors occurs, and for the analysis of spatial heterogeneity in yearly transition matrices.

Acknowledgments We express our sincere thanks to Masahiro Ichikawa, Takashi Kohyama, Toru Nakashizuka, and Ken-Ichi Akao for their helpful suggestions. Prof. Ichikawa encouraged us to continue this study. Profs. Kohyama and Nakashizuka provided the opportunity to solve the mechanism of the dynamics of land use. Prof. Akao provided mathematical advice at an early stage of our study. This research was funded in part by Grants-in-Aid from the

Japanese Society for the Promotion of Science (JSPS) for Scientific Research (nos. A-21247006, B-20370006 and B-21310152) and project 2–2 "Sustainability and Biodiversity Assessment on Forest Utilization Options" and D-04 of the Research Institute for Humanity and Nature.

References

Baker WL (1989a) Landscape ecology and nature reserve design in the Boundary Waters Canoe Area, Minnesota. Ecology 70:23–35

Baker WL (1989b) A review of models of landscape change. Landscape Ecol 2:111–133

Braimoh AK (2006) Random and systematic land-cover transitions in Northern Ghana. Agric Ecosyst Environ 113 (1–4):254–263

Braimoh AK, Vlek PLG (2004) Land-cover dynamics in an urban area of Ghana. Earth Interact 8:1–15

Cinlar E (1975) Introduction to stochastic processes. Prentice-Hall, Englewood Cliffs

Ehlers M, Jadkowski MA, Howard RR et al (1990) Application of SPOT data for regional growth analysis and local planning. Photogramm Eng Remote Sensing 56:175–180

Flamenco-Sandoval A, Ramos MM, Masera OR (2007) Assessing implications of land-use and land-cover change dynamics for conservation of a highly diverse tropical rain forest. Biol Conserv 138:131–145

Gardner RH, Milne BT, Turner MG, O'Neill RV (1987) Neutral models for the analysis of broad-scale landscape pattern. Landscape Ecol 1:19–28

Gustafson EJ, Parker GR (1992) Relationships between land-cover proportion and indices of landscape spatial pattern. Landscape Ecol 7(2):101–110

Hathout S (2002) The use of GIS for monitoring and predicting urban growth in East and West St. Paul, Winnipeg, Manitoba. Can J Environ Manag 66:229–238

Kachi N, Yasuoka Y, Totsuka T et al (1986) A stochastic model for describing re-vegetation following forest cutting—an application of remote sensing. Ecol Model 32:105–117

Lambin EF, Geist HJ (2006) Land-use and land-cover change. Springer Verlag, Berlin

Lewis HG, Brown M (2001) A generalized confusion matrix for assessing area estimates from remotely sensed data. Int J Remote Sensing 22(16):3223–3235

Lipschutz S (1979) Probabilities, course et problems. Série Schaum, McGraw-Hill, Paris

Lo CP, Yang X (2002) Drivers of land-use/land-cover changes and dynamic modeling for the Atlanta, Georgia metropolitan area. Photogramm Eng Remote Sensing 68(10):1073–1082

Mertens B, Lambin E (2000) Land-cover-change trajectories in southern Cameroon. Ann Ass Am Geogr 90(3):467–494

Messerli B (1997) Geography in a rapidly changing world. Int Geogr Bull 47:65–75

Meyer WB, Turner BL II (1991) Changes in land use and land cover: a global perspective. Cambridge University Press, Cambridge

Mundia CN, Aniya M (2005) Analysis of land use/cover changes and urban expansion of Nairobi City using remote sensing and GIS. Int J Remote Sensing 26:2831–2849

Petit C, Scudder T, Lambin E (2001) Quantifying processes of land-cover change by remote sensing: resettlement and rapid land-cover changes in south-eastern Zambia. Int J Remote Sensing 22:3435–3456

Pontius RG Jr (2002) Statistical methods to partition effects of quantity and location during comparison of categorical maps at multiple resolutions. Photogramm Eng Remote Sensing 68(10):1041–1049

Pontius RG Jr, Cheuk ML (2006) A generalized cross-tabulation matrix to compare soft-classified maps at multiple resolutions. Int J Geogr Info Sci 20(1):1–30

Pontius RG Jr, Shusas E, McEachern M (2004) Detecting important categorical land changes while accounting for persistence. Agric Ecosyst Environ 101(2–3):251–268

Turner MG (1990) Spatial and temporal analysis of landscape patterns. Landscape Ecol 4(1):21–30

Turner MG, Dale VH, Gardner RH (1989) Predicting across scales: theory development and testing. Landscape Ecol 3(3/4):245–252

Turner BL II, Lambin EF, Reenberg A (2007) The emergence of land change science for global environmental change and sustainability. PNAS 104:2066–2071

Usher MB (1981) Modeling ecological succession, with particular reference to Markovian models. Vegetatio 46–7: 11–18