A method for estimation of a land-cover change matrix from error-prone unit-level observations

Steen Magnussen

Abstract: Coregistration and classification errors can seriously compromise direct unit-level (pixel) estimation of land-cover change from remotely sensed data. A more robust alternative to a pixel-based estimation of change is warranted. In a proposed method, spatially adjacent pixels are grouped into 3×3 clusters, and the change matrix is obtained from cluster-specific and land cover specific pixel counts at two points in time. The diagonal of a change matrix is estimated by combining an estimate of the temporal correlation of cover type specific, cluster-level counts with an estimate of the odds ratio of no change. Off-diagonal elements are least-squares solutions to a set of linear constraints or obtained by iterative proportional fitting under a model of quasi-independence. In a study with data from five sites, the proposed method produced less biased estimates on three sites if the mean coregistration error was in excess of 0.3–0.7 pixels and on four sites if classification accuracy dropped below 0.9.

Résumé : Les erreurs de correspondance géométrique et de classification peuvent compromettre sérieusement l'estimation du changement des surfaces par une analyse des pixels de l'image obtenue par télédétection. Une alternative plus robuste à l'estimation du changement pixel-à-pixel est nécessaire. La méthode proposée dans cette étude utilise les pixels voisins rassemblés par groupes de 3×3 pixels pour produire une matrice de changement des histogrammes associés aux groupements de pixels et aux surfaces pour un point, à deux moments dans le temps. La diagonale de la matrice de changement est calculée en combinant la corrélation temporelle du dénombrement de pixels par type de surface avec une estimation des probabilités qu'il n'y ait pas de changement. Les valeurs hors-diagonale sont des solutions par la méthode des moindres carrés d'un ensemble de contraintes linéaires ou sont obtenues par un ajustement proportionnel itératif fait selon un modèle de quasi indépendance. Dans une étude avec les données de cinq sites, la méthode proposée a produit des estimations moins biaisées sur trois sites si l'erreur moyenne de correction géométrique était supérieure à 0,3–0,7 pixels et sur quatre sites si l'exactitude de la classification était inférieure à 0,9.

[Traduit par la Rédaction]

Introduction

Estimation of changes in land cover, land use, and land status is a primary objective in most natural large area natural resource monitoring activities (Rosenfield et al. 1982; Smith and Annoni 1999; Anderson 2002; Brown 2002; Coomes et al. 2002; Corona et al. 2002; Parr et al. 2002). For a closed population composed of N ultimate units labeled to one of K distinct categorical classes at two points in time, a $K \times K$ change matrix **n** with elements $n_{kk'}$, $k, k' = 1, 2, ..., K$ provides a succinct and sufficient summary of the changes that have taken place; $n_{kk'}$ is the number of units that changed from k to k' during the period in question. The composition of the population at the beginning of the time period (t_1) is given by the row sums of this matrix $(n_{k\cdot}, i = 1, 2, ..., K)$, where a centre dot in a subscript stands for a summation over the subscript it replaces. Conversely, the composition at the end of the period (t_2) is obtained by its column sums $(n_{\cdot k}, k = 1, 2, ..., K)$. With unit-by-unit observations at each point in time, a direct estimation of the change matrix would follow from a simple count-

ing of units in the K^2 change classes. In a sampling context, the direct estimation would be by a design-consistent method (Czaplewski and Catts 1992; Green et al. 1992; Reams and Van Deusen 1999; Magnussen and Köhl 2005).

However, land cover change estimation from remotely sensed data is riddled by a unique set of problems due to registration errors when two images are coregistered (Coppin and Bauer 1996; Townsend et al. 2000; Coppin et al. 2004), point-spread functions that "smear" information from one image unit (pixel) across neighbouring units (Collins and Woodcock 1999), image resampling of lost or missing units (Garcia-Gigorro and Saura 2005), different atmospheric conditions (Song et al. 1999; Schroeder et al. 2006), topography, differences in viewing and sun angles (Lunetta and Lyon 2004), and possible classification errors (Congalton 2001; Pontius and Lippitt 2006). Therefore, a registered change in a unit can misrepresent the actual change event. Robust procedures for change estimation when unit-level data is error prone are warranted. Although techniques for reducing the bias due to classification errors are readily available (Stehman and Czaplewski 1998; Czaplewski 2003; for example, Van Deusen 1994; Biging et al. 1999), their efficiency depends critically on an accurate estimate of a $K^2 \times K^2$ confusion matrix, which is rarely available in practical applications. A locally adaptive threshold for accepting an apparent change event as "real" has been proposed for mittigating coregistration errors, but the

Table 1. Change matrix for four land cover classes (1–4). $n_{kk'}$ number of units that changed from class k to class k' between time t_1 and time t_2.

Class	1	2	3	4	t_1
1	n_{11}	n_{12}	n_{13}	n_{14}	$n_{1\cdot}$
2	n_{21}	n_{22}	n_{23}	n_{24}	$n_{2\cdot}$
3	n_{31}	n_{32}	n_{33}	n_{34}	$n_{3\cdot}$
4	n_{41}	n_{42}	n_{43}	n_{44}	$n_{4\cdot}$
t_2	$n_{\cdot 1}$	$n_{\cdot 2}$	$n_{\cdot 3}$	$n_{\cdot 4}$	$n_{\cdot\cdot}$

Note: $n_{k\cdot}$, number of units in class k at t_1 and $n_{\cdot k}$, the number of units in class k at t_2; $n_{\cdot\cdot}$, total number of units ($= N$ in a census).

Table 2. Constraints on change matrix counts $n_{kk'} = \{1,2,3,4\}$ and $k \neq k'$.

Constraint	Equation
1	$n_{12} + n_{13} + n_{14} = n_{1\cdot} - \hat{n}_{11}$
2	$n_{21} + n_{23} + n_{24} = n_{2\cdot} - \hat{n}_{22}$
3	$n_{31} + n_{32} + n_{34} = n_{3\cdot} - \hat{n}_{33}$
4	$n_{41} + n_{42} + n_{43} = n_{4\cdot} - \hat{n}_{44}$
5	$n_{21} + n_{31} + n_{41} = n_{\cdot 1} - \hat{n}_{11}$
6	$n_{12} + n_{32} + n_{42} = n_{\cdot 2} - \hat{n}_{22}$
7	$n_{13} + n_{23} + n_{43} = n_{\cdot 3} - \hat{n}_{33}$
8	$n_{14} + n_{24} + n_{34} = n_{\cdot 4} - \hat{n}_{44}$
9	$n_{13} + n_{23} + n_{14} + n_{24} = n_{(1+2)\cdot} - \hat{n}_{(1+2)(1+2)}$
10	$n_{12} + n_{32} + n_{14} + n_{34} = n_{(1+3)\cdot} - \hat{n}_{(1+3)(1+3)}$
11	$n_{12} + n_{42} + n_{13} + n_{43} = n_{(1+4)\cdot} - \hat{n}_{(1+4)(1+4)}$
12	$n_{21} + n_{31} + n_{24} + n_{34} = n_{(2+3)\cdot} - \hat{n}_{(2+3)(2+3)}$
13	$n_{21} + n_{23} + n_{41} + n_{43} = n_{(2+4)\cdot} - \hat{n}_{(2+4)(2+4)}$
14	$n_{31} + n_{32} + n_{41} + n_{42} = n_{(3+4)\cdot} - \hat{n}_{(3+4)(3+4)}$

Note: Constraints 1–8 apply to the original 4×4 change matrix. Constraints 9–14 are obtained from the six possible 3×3 matrices that can be formed from a 4×4 change matrix by joining two of the four classes. Subscript $(k + k')$ indicate that class k and class k' have been joined.

method requires a good estimation of the coregistration error process to be efficient (Bruzzone and Cossu 2003).

This study proposes a new method for estimating a change matrix from clustered (pooled) data units as opposed to a direct unit-by-unit observation of change. It is assumed that change estimates derived from clusters of spatially adjoining units are less sensitive to the above-mentioned problems than estimates derived from direct unit-level observations of change. Only cluster-specific counts of the number of units in each cover type at t_1 and t_2 would be used during estimation of a change matrix. Of course, the loss of information on unit-level change increases dramatically with cluster size. Hence, the clusters must necessarily be small yet large enough to mitigate the negative impact of the aforementioned problems. In this study, a cluster size of 3×3 units was chosen as a compromise between intrinsically conflicting goals.

In the proposed method, estimates of the elements along the diagonal in a change matrix ($\hat{n}_{kk}$) are obtained via a combination of an estimate of the temporal correlation of cluster-level, cover type specific unit counts and an estimate of the odds ratio of no change (Magnussen 2004). Off-diagonal elements are either least-squares solutions to a derived set of

linear constraints on row and column sums or estimated by iterative proportional fitting (IPF) under a model of quasi-independence of row and columns (Agresti 1992; Booth et al. 2005).

The performance of the proposed method is assessed for a 4×4 change matrix with data from five sites, but it extends naturally to a $K \times K$ change matrix, albeit with a nontrivial increase in computational costs. Stochastic coregistration errors in t_2 data with means ranging from 0 to 1.1 unit (pixel) in steps of 0.1 are simulated and their impact on bias is assessed. The impact of multinomially distributed classification errors at both times of observation is also assessed in simulations with the classification accuracy dropping from 1.0 to 0.5 in steps of 0.1.

Methods

An estimate of the change between two points in time (t_1 and t_2) in a population composed of N discrete units labelled to one of K distinct land-cover classes is desired. A $K \times K$ change matrix **n** provides this estimate. A change matrix for $K = 4$ is given in Table 1. Unit-level observation of change events is considered error prone due to either imperfect coregistration, classification errors, or both. For this reason, a direct estimation based on unit-level counts by change class may be seriously biased.

Estimation of the change matrix by the proposed method begins with a complete tesselation of the population into M size 9 clusters with units in each cluster arranged in a regular 3×3 array. The observed frequencies of unit-level class change between time t_1 and t_2 in the ith cluster ($n_{kk'}^i$) are conveniently arranged in a $K \times K$ matrix $\mathbf{n}_i$. Because unit-level observations are considered to be error prone, only cluster-level counts of the number of units in each class at t_1 and t_2 will be used for the estimation of the desired change matrix. Specifically, the t_1 data were $\mathbf{n}_1^{(i)} = \{n_1^{(i)}, \ldots, n_K^{(i)}\}$, $i = 1, \ldots, M$, where $n_{k\cdot}^{(i)} = \sum_{k'} n_{kk'}^{(i)}$, i.e., the kth row sum of $\mathbf{n}_i$, is the number of class k units in the ith cluster at time t_1. For the given cluster size, we have $\sum_{k=1}^{K} n_{k\cdot}^{(i)} = 9$. Time t_2 data are $\mathbf{n}_2^{(i)} = \{n_{\cdot 1}^{(i)}, n_{\cdot 2}^{(i)}, \ldots, n_{\cdot K}^{(i)}\}$, $i = 1, 2, \ldots, M$, where $n_{\cdot k}^{(i)} = \sum_{k'} n_{k'k}^{(i)}$, i.e., the kth column sum of $\mathbf{n}_i$. Sufficient statistics for the state of the population at t_1 and t_2 are the vectors of total counts by class $\mathbf{n}_1 = \{n_{1\cdot}, n_{2\cdot}, \ldots n_{K\cdot}\}$, and $\mathbf{n}_2 = \{n_{\cdot 1}, n_{\cdot 2}, \ldots n_{\cdot K}\}$. The total sum of observed counts at both t_1 and t_2 is $\sum_k n_{k\cdot} = \sum_k n_{\cdot k} = n_{\cdot\cdot}$. In a census, $n_{\cdot\cdot} = N$. In a sampling context, the data would arise from a sampling of m of the M clusters according to some sampling protocol.

Estimation of diagonal elements

Estimation of the elements along the diagonal in the change matrix (i.e., n_{kk}, $k = 1, 2, \ldots, K$) has been detailed by Magnussen (2004). Only a brief outline is provided here. First, Pearson's correlation coefficient (ρ_{kk}) of cluster-level cover type specific counts at t_1 and t_2 $\left(\{n_{k\cdot}^{(i)}, n_{\cdot k}^{(i)}\} \right.$, $i = 1, 2, \ldots, M$, $k = 1, 2, \ldots, K \right)$ is computed. From these coefficients, a preliminary estimate $\tilde{n}_{kk}$ is obtained by solving the following equation for the unknown n_{kk} (Murtaugh and Phillips 1998):

Table 3. Change matrix for Prince George, British Columbia (BC), population ($N = 121\,104$).

Class	1	2	3	4	All (t_1)
1	15 047 (12.4)	10 660 (8.8)	1 879 (1.6)	666 (0.5)	28 252 (23.3)
2	5 846 (4.8)	33 255 (27.5)	5 040 (4.2)	3216 (2.7)	47 357 (39.1)
3	3 257 (2.7)	5 468 (4.5)	7 961 (6.6)	1964 (1.6)	18 650 (15.4)
4	3 181 (2.6)	10 977 (9.1)	8 604 (7.1)	4083 (3.4)	26 845 (22.2)
All (t_2)	27 331 (22.6)	60 360 (49.8)	23 484 (19.4)	9929 (8.2)	121 104 (100.0)

Note: See Table 1 for a format specification. Percentages of total are given in parentheses. Percentages may not add to 100 due to rounding.

Table 4. Change matrix for Hinton, Alberta (HI), population ($N = 129\,600$).

Class	1	2	3	4	All (t_1)
1	69 850 (53.9)	7 117 (5.5)	1885 (1.5)	5 169 (4.0)	84 021 (64.8)
2	64 (<0.1)	1 923 (1.5)	1367 (1.1)	809 (0.6)	4 163 (3.2)
3	8 481 (6.5)	1 549 (1.2)	4373 (3.4)	5 136 (4.0)	19 539 (15.1)
4	8 313 (6.4)	1 010 (0.8)	2066 (1.6)	10 488 (8.1)	21 877 (16.9)
All (t_2)	86 708 (66.9)	11 599 (9.0)	9691 (7.5)	21 602 (16.7)	129 600 (100.0)

Note: See Table 1 for a format specification. Percentages of total are given in parentheses. Percentages may not add to 100 due to rounding.

$$[1] \qquad \hat{\rho}\left(n_{k\cdot}^{(i)}, n_{\cdot k}^{(i)}\right) \approx \frac{n_{kk}n_{\cdot\cdot} - n_{k\cdot}n_{\cdot k} \times n_{\cdot\cdot}^{-1}}{\sqrt{n_{k\cdot}(n_{\cdot\cdot} - n_{k\cdot})n_{\cdot k}(n_{\cdot\cdot} - n_{\cdot k})}},$$
$$k = 1,\ 2,\ ...,\ K$$

A second preliminary estimate $\breve{n}_{kk}$ is obtained by finding an odds ratio of no change for class k

$\left(\theta_k \equiv n_{kk}\sum_{k \neq k'}n_{k'k'} \times (n_{k''} - n_{k'k'})^{-1}(n_{\cdot k'} - n_{k'k'})^{-1}\right)$ (Fleiss 1981, Ch. 5.3) that maximizes the likelihood of the observed counts $\mathbf{n}_t^{(i)}$, $i = 1, 2, ..., M$, $t = 1, 2$. The likelihood of maximization is detailed elsewhere (Magnussen 2004, eq. 11, p. 1707). The second prelimary estimate is then

$$[2] \qquad \breve{n}_{kk} \approx \frac{n_{\cdot\cdot} - n_{k\cdot} - n_{\cdot k} + \hat{\theta}_k(n_{k\cdot} + n_{\cdot k}) \pm \hat{\Upsilon}_k}{2(\hat{\theta}_k - 1)} \quad \text{and} \quad \hat{\Upsilon}_k \approx \sqrt{\begin{array}{l} n_{\cdot\cdot}^2 + (1 - 2\hat{\theta}_k + \hat{\theta}_k^2)n_{k\cdot}^2 + 2(\hat{\theta}_k - 1)n_{\cdot k}n_{\cdot\cdot} + (1 - 2\hat{\theta}_k + \hat{\theta}_k^2)n_{\cdot k}^2 \\ +2\left(\hat{\theta}_k - 1 + (1 - \hat{\theta}_k^2)n_{\cdot k}n_{\cdot\cdot}^{-1}\right)n_{k\cdot}n_{\cdot\cdot} \end{array}}$$

where the solution is the one that satisfies $n_{k\cdot}n_{\cdot k} \leq \min(n_{k\cdot}, n_{\cdot k})$. Empirical evidence suggests taking $\hat{n}_{kk}$ as $\frac{2}{3}\tilde{n}_{kk} + \frac{1}{3}\breve{n}_{kk}$ (Magnussen 2004, p. 1707).

Estimation of off-diagonal elements

The estimation of the $K(K - 1)$ off-diagonal elements $n_{kk'|k \neq k'}$ $k, k' = \{1, \ldots, K\}$ proceeds in one of two ways: (*i*) a fast IPF of an initial proposal change matrix that has the above diagonal elements and positive real integers in the off-diagonal positions that are chosen in such a way that either the row sums of the proposed matrix matches those of $\mathbf{n}_1$ or the column sums matches those of $\mathbf{n}_2$ (Bishop et al. 1975; Magnussen and Köhl 2005), or (*ii*) a computationally expensive least-squares solution (CLS) to a set of derived constraints on the off-diagonal elements of the $K \times K$ change matrix.

IPF estimation of off-diagonal elements

The IPF approach yields the following estimates: $\breve{n}_{kk'} = \hat{c}_k \times \hat{c}_{k'} \times (n_{k\cdot} - \hat{n}_{kk}) \times (n_{\cdot k} - \hat{n}_{kk})$, $k \neq k'$ where $\hat{c}_k$ and $\hat{c}_{k'}$ are constants to be estimated by IPF. IPF estimates satisfy $\sum_{k \neq k'}\breve{n}_{kk'} \approx n_{k\cdot} - \hat{n}_{kk}$ and $\sum_{k \neq k'}\breve{n}_{kk'} \approx n_{\cdot k} - \hat{n}_{kk}$ with approximation errors (here) less than 0.3%. A change matrix estimated by IPF is denoted as $\hat{\mathbf{n}}_{\text{IPF}}$. Note that IPF

assumes conditional row and column independence for the off-diagonal element with conditioning on the observed row and column sums and the estimated diagonal elements. That is a model of quasi-independence (Agresti and Caffo 2000, Ch. 8; Booth et al. 2005).

CLS estimation of off-diagonal elements

The fact that diagonal elements of any change matrix can be estimated quite well is exploited in the computationally expensive CLS approach to derive an expanded set of linear constraints on the off-diagonal elements in the change matrix and then finding a least-squares solution to these constraints. Given the marginal counts $(n_{k\cdot}, n_{\cdot k})$ and the above estimates $\hat{n}_{kk}$, $k = 1, ..., K$, we can formulate a trivial set of $2K$ constraints on the row and column sums of the elements in the change matrix. For $K = 4$, the constraints (1–8) are listed in Table 2. However, they are not very useful for the estimation of the $K^2 - K$ off-diagonal elements, because their rank is only $2K - 1$. For a reasonable estimation, we would need to increase the number of nonredundant constraints. It turns out that we can construct an expanded set of constraints with a rank higher than $2K - 1$. How this is done is described next and detailed for $K = 4$.

If we were to estimate a 3×3 change matrix, we could

Table 5. Change matrix for Latium, Italy (IT), population ($N = 208\,675$).

Class	1	2	3	4	All (t_1)
1	30 771 (14.7)	89 (<0.1)	5117 (2.5)	4 262 (2.0)	40 239 (19.3)
2	2 (<0.1)	1455 (0.7)	80 (<0.1)	12 (<0.1)	1 549 (0.7)
3	3 582 (1.7)	3321 (1.6)	64 977 (31.1)	1 876 (0.9)	73 756 (35.3)
4	630 (0.3)	30 (<0.1)	364 (0.2)	92 107 (44.1)	93 131 (44.6)
All (t_2)	34 985 (16.8)	4895 (2.4)	70 538 (33.8)	98 257 (47.1)	208 675 (100.0)

Note: See Table 1 for a format specification. Percentages of total are given in parentheses. Percentages may not add to 100 due to rounding.

Table 6. Change matrix for Sussex, New Brunswick (NB), population ($N = 181\,068$).

Class	1	2	3	4	All (t_1)
1	25 528 (14.1)	765 (0.4)	78 (<0.1)	2 387 (1.3)	28 758 (15.9)
2	668 (0.4)	56 414 (31.2)	4 931 (2.7)	7 343 (4.1)	69 356 (38.3)
3	120 (0.1)	497 (0.3)	21 696 (12.0)	1 636 (0.9)	23 949 (13.2)
4	336 (0.2)	376 (0.2)	251 (0.1)	58 042 (32.1)	59 005 (32.6)
All (t_2)	26 652 (14.7)	58 052 (32.1)	26 956 (14.9)	69 408 (38.7)	181 068 (100.0)

Note: See Table 1 for a format specification. Percentages of total are given in parentheses. Percentages may not add to 100 due to rounding.

obtain a rank five set of constraints for the six unknown off-diagonal elements after having obtained the above estimates for the diagonal. The ratio of unknowns (six) to the rank of the linear constraints (five) is the maximum possible for the estimation problem at hand. Hence, if we could break the estimation problem down to a size 3×3 problem, our least-squares solution would be the best possible. Because any $K \times K$ change matrix can be transformed into a 3×3 change matrix by keeping the identity and the observed counts for two of the original K classes and joining the counts for the remaining $K - 2$ classes into one aggregate third class, we could obtain the best possible least-squares solution for this 3×3 change matrix. Only the two off-diagonal estimates for the two original classes in this matrix would be of interest, of course. A repeat of this procedure for all $\binom{K}{2}$ possible 3×3 change matrices would provide the desired estimates for the off-diagonal elements. The six possible 3×3 change matrices that can be formed this way from a single 4×4 change matrix are as follows:

$$\begin{bmatrix} n_{33} & n_{34} & n_{3(1+2)} \\ n_{43} & n_{44} & n_{4(1+2)} \\ n_{(1+2)3} & n_{(1+2)4} & n_{(1+2)(1+2)} \end{bmatrix}$$

$$\begin{bmatrix} n_{22} & n_{24} & n_{2(1+3)} \\ n_{42} & n_{44} & n_{4(1+3)} \\ n_{(1+3)2} & n_{(1+3)4} & n_{(1+3)(1+3)} \end{bmatrix}$$

$$\begin{bmatrix} n_{22} & n_{23} & n_{2(1+4)} \\ n_{32} & n_{33} & n_{3(1+4)} \\ n_{(1+4)2} & n_{(1+4)3} & n_{(1+4)(1+4)} \end{bmatrix}$$

$$\begin{bmatrix} n_{11} & n_{14} & n_{1(2+3)} \\ n_{41} & n_{44} & n_{4(2+3)} \\ n_{(2+3)1} & n_{(2+3)4} & n_{(2+3)(2+3)} \end{bmatrix}$$

$$\begin{bmatrix} n_{11} & n_{13} & n_{1(2+4)} \\ n_{31} & n_{33} & n_{3(2+4)} \\ n_{(2+4)1} & n_{(2+4)3} & n_{(2+4)(2+4)} \end{bmatrix}$$

$$\begin{bmatrix} n_{11} & n_{12} & n_{1(3+4)} \\ n_{21} & n_{22} & n_{2(3+4)} \\ n_{(3+4)1} & n_{(3+4)2} & n_{(3+4)(3+4)} \end{bmatrix}$$

Counts for joined classes have subscripts $(k + k')$ to indicate that classes k and k' have been joined ($k, k' = 1, 2, 3, 4$). Joined classes been placed in the third column and third row of each matrix.

However, estimates obtained in this sequential fashion would not be simultaneously the best possible in a least-squares sense. The rank of all constraints considered at once remains below the number of variables, so we must find a smaller set of constraints with a maximum rank (Searle 1982, p 190). Table 2 lists (for $K = 4$) a set of 14 linear constraints on the off-diagonal elements with the maximum possible rank of nine. Many other equivalent rank-nine sets of constraints can easily be found by simple matrix operations on the design matrix of the entire set of constraints. However, least-squares solutions would be identical. A change matrix estimated by CLS is denoted as $\hat{\mathbf{n}}_{\text{CLS}}$. All CLS estimates reported here also satisfy some additional contraints:

$$\hat{n}_{kk'}^{\text{CLS}} \geq 0 , \quad \sum_{k,k'} \hat{n}_{kk'}^{\text{CLS}} = n_{..}, \text{ and } n_{kk'}$$
$$\supset \text{Integer} , \ k, k' = \{1, 2, \ \ldots, K = 4\}$$

The latter was achieved by rounding of estimates for the di-

Table 7. Change matrix for Selangor, Indonesia (SE), population ($N = 166\,464$).

Class	1	2	3	4	All (t_1)
1	15 922 (9.6)	3 742 (2.2)	12 813 (7.7)	6 311 (3.8)	38 788 (23.3)
2	3 472 (2.1)	9 626 (5.8)	10 323 (6.2)	1 0398 (6.2)	33 819 (20.3)
3	2 770 (1.7)	5 298 (3.2)	38 709 (23.3)	22 423 (13.5)	69 200 (41.6)
4	1 095 (0.7)	1 566 (0.9)	6 881 (4.1)	15 115 (13.5)	24 657 (14.8)
All (t_2)	23 259 (14.0)	20 232 (12.2)	68 726 (41.3)	54 247 (32.6)	166 464 (100.0)

Note: See Table 1 for a format specification. Percentages of total are given in parentheses. Percentages may not add to 100 due to rounding.

Table 8. Absolute bias of estimated change matrices.

Site	Absolute bias ($\hat{\mathbf{n}}_{\mathrm{CLS}}$)	Absolute bias ($\hat{\mathbf{n}}_{\mathrm{IPF}}$)
BC	$\begin{bmatrix} 1.5 & 0.8 & 0.5 & -0.2 \\ 1.4 & 3.2 & 1.5 & 0.3 \\ 0.6 & 0.3 & 2.0 & 1.0 \\ 0.4 & 2.1 & 0.0 & 1.6 \end{bmatrix}$	$\begin{bmatrix} \cdots & 3.2 & 1.2 & 0.5 \\ 1.6 & \cdots & 0.5 & 1.2 \\ 0.9 & 0.2 & \cdots & 0.9 \\ 1.1 & 0.1 & 2.7 & \cdots \end{bmatrix}$
HI	$\begin{bmatrix} 0.7 & 0.2 & 0.4 & 0.5 \\ 0.8 & 0.2 & 0.5 & 0.1 \\ 0.2 & 0.4 & 1.1 & 1.3 \\ 1.3 & 0.5 & 0.1 & 2.0 \end{bmatrix}$	$\begin{bmatrix} \cdots & 1.5 & 0.6 & 0.1 \\ 1.2 & \cdots & 0.8 & 0.2 \\ 0.2 & 0.9 & \cdots & 1.8 \\ 1.9 & 0.7 & 0.8 & \cdots \end{bmatrix}$
IT	$\begin{bmatrix} 1.2 & 0.3 & 1.5 & 0.7 \\ 0.0 & 0.1 & 0.0 & 0.2 \\ 1.6 & 1.1 & 1.2 & 0.6 \\ 0.4 & 0.9 & 0.3 & 0.2 \end{bmatrix}$	$\begin{bmatrix} \cdots & 1.5 & 0.7 & 1.1 \\ 0.1 & \cdots & 0.0 & 0.0 \\ 1.1 & 1.4 & \cdots & 1.4 \\ 0.1 & 0.1 & 0.2 & \cdots \end{bmatrix}$
NB	$\begin{bmatrix} 2.1 & 1.6 & 0.4 & 0.1 \\ 0.8 & 1.9 & 1.4 & 0.3 \\ 0.9 & 0.4 & 1.8 & 0.5 \\ 0.4 & 0.2 & 0.1 & 0.3 \end{bmatrix}$	$\begin{bmatrix} \cdots & 0.9 & 1.1 & 0.2 \\ 1.6 & \cdots & 0.6 & 0.3 \\ 0.5 & 0.9 & \cdots & 0.5 \\ 0.0 & 0.1 & 0.2 & \cdots \end{bmatrix}$
SE	$\begin{bmatrix} 0.3 & 0.9 & 1.0 & 0.5 \\ 1.0 & 1.1 & 0.3 & 0.2 \\ 1.2 & 1.5 & 1.2 & 0.9 \\ 0.1 & 0.5 & 0.1 & 0.2 \end{bmatrix}$	$\begin{bmatrix} \cdots & 0.7 & 1.1 & 2.2 \\ 0.8 & \cdots & 0.1 & 0.4 \\ 1.0 & 0.3 & \cdots & 2.0 \\ 0.1 & 0.0 & 0.3 & \cdots \end{bmatrix}$

Note: The bias is $(\hat{\mathbf{n}}_M - \hat{\mathbf{n}})/(m \times 9) \times 100\%$ where $\hat{\mathbf{n}}$ and $\hat{\mathbf{n}}_M$ are the means in 200 replications of $m = 200$ samples of 3×3 clusters, $M = \{\mathrm{CLS}, \mathrm{IPF}\}$. Bias of diagonal elements are identical in the two models and only shown for $\hat{\mathbf{n}}_{\mathrm{CLS}}$. Sites are as follows: BC, Prince George, British Columbia; HI, Hinton, Alberta; IT, Latium, Italy; NB, Sussex, New Brunswick; SE, Selangor, Indonesia.

agonal and by integer least squares solutions for off-diagonal elements.

CLS-fitted off-diagonal elements in rows viz. columns with large values of $n_k^{(t)} - \hat{n}_{kk}$, $t = \{1, 2\}$, $k = \{1, 2, ..., K\}$ will have the smallest relative error. Hence, the relative error of an element in a row viz. a column with a small value

of $n_i^{(t)} - \hat{n}_{ii}$ could be large. A weighted least-squares solution with weights inversely proportional to, for example, $n_i^{(t)} - \hat{n}_{ii}$ would address this error-allocation problem. Only the unweighted option is pursued here.

Performance assessment

In five case studies with unit-level land cover change data in four land-cover classes ($K = 4$), the best (least biased) of the proposed estimators $\hat{\mathbf{n}}_{\mathrm{IPF}}$ and $\hat{\mathbf{n}}_{\mathrm{CLS}}$ are compared with an estimate $\hat{\mathbf{n}}$ obtained by a direct counting of unit-level change events $\left(n_{kk'}^{(i)}, k, k' = 1, 2, 3, 4; i = 1, 2, ..., m\right)$. The hypothesis that $\hat{\mathbf{n}}_{\mathrm{CLS}}$ provides a better fit to $\hat{\mathbf{n}}$ than $\hat{\mathbf{n}}_{\mathrm{IPF}}$ was tested with partial χ^2 test statistics $\hat{\chi}^2(\hat{\mathbf{n}}_{\mathrm{IPF}} - \hat{\mathbf{n}}) - \chi^2(\hat{\mathbf{n}}_{\mathrm{CLS}} - \hat{\mathbf{n}})$ with 2 degrees of freedom and adjusted for the effect of clustering (Brier 1980; Magnussen and Köhl 2006). The hypotheses that both $\hat{\mathbf{n}}_{\mathrm{IPF}}$ and $\hat{\mathbf{n}}_{\mathrm{CLS}}$ provide a better fit to $\hat{\mathbf{n}}$ than an estimate $\hat{\mathbf{n}}_0$ obtained under the null model of row and column independence (i.e., $\hat{\mathbf{n}}_0 = \hat{\mathbf{n}}_1 \otimes \hat{\mathbf{n}}_2$) were also tested with partial χ^2 tests.

To illustrate the performance of the proposed estimator(s) and to facilitate Monte Carlo simulations of unit-level errors of coregistration and classification (see next section), the assessment is carried out with data from 200 replications of simple random sampling of $m = 200$ 3×3 clusters of unit-level change data from five sets of two coregistered cover-type maps. Linear regressions of $\hat{n}_{kk'}$ on $n_{kk'}^{\mathrm{CLS}}$ (or $n_{kk'}^{\mathrm{IPF}}$) are used to illustrate bias and lack of fit in individual replications.

Error-prone unit-level data motivated the proposed estimator(s). Because $\hat{\mathbf{n}}_{\mathrm{IPF}}$ and, in particular, $\hat{\mathbf{n}}_{\mathrm{CLS}}$ adds nontrivial computational costs, they must be justified by expected benefits. With this objective, the performance of the better of the two estimators is demonstrated in applications when unit-level data are corrupted by simulated coregistration errors viz. classification errors. The data used are (*i*) the original data, (*ii*) the original data subject to simulated coregistration errors at time two, and (*iii*) the original data subject to independently simulated classification errors at time t_1 and t_2. For the second and third data sets, various levels of mean coregistration errors and classification errors are simulated.

Monte Carlo simulation of coregistration and classification errors

Coregistration errors in t_2 data arise when the t_2 image of the remotely sensed population units is coregistered to the t_1 image. Eleven levels of the mean coregistration error for t_2 data were simulated as random events at the cluster level. With a probability of P_0, the "true" t_2 data in a 3×3 cluster had no coregistration error ($P_0 = 0, 0.1, ..., 0.9, 1.0$). With a probability of $P_1 = \frac{2}{3}(1 - P_0)$, the t_2 location of a

Fig. 1. Scatterplot of $\hat{n}_{kk'}^{\mathrm{CLS}}$ versus $\hat{n}_{kk'}$ in 200 replicated samples of size $m = 200$ in Prince George, British Columbia (BC).
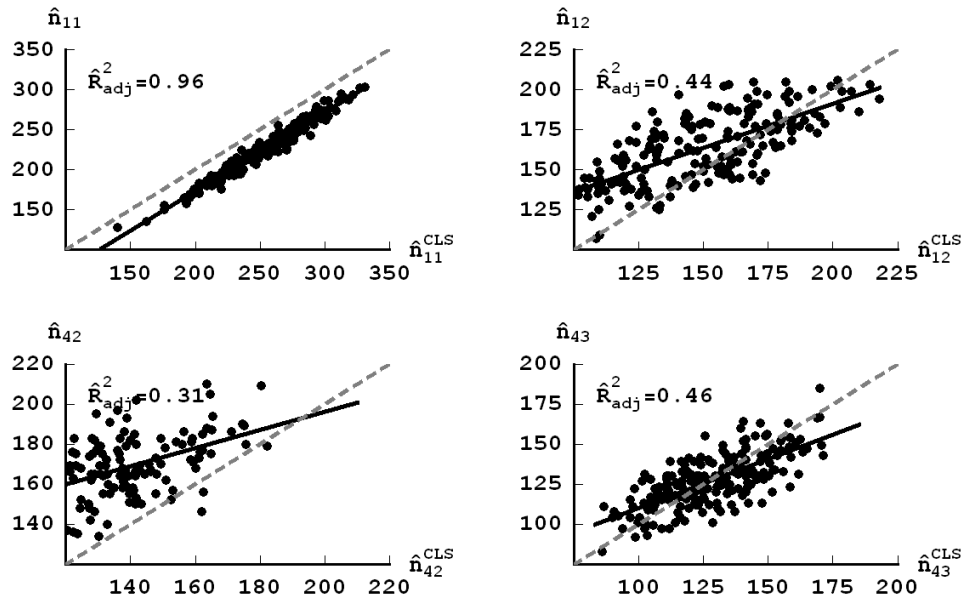
Fig. 2. Scatterplot of $\hat{n}_{kk'}^{\mathrm{CLS}}$ versus $\hat{n}_{kk'}$ in 200 replicated samples of size $m = 200$ in Hinton, Alberta (HI).
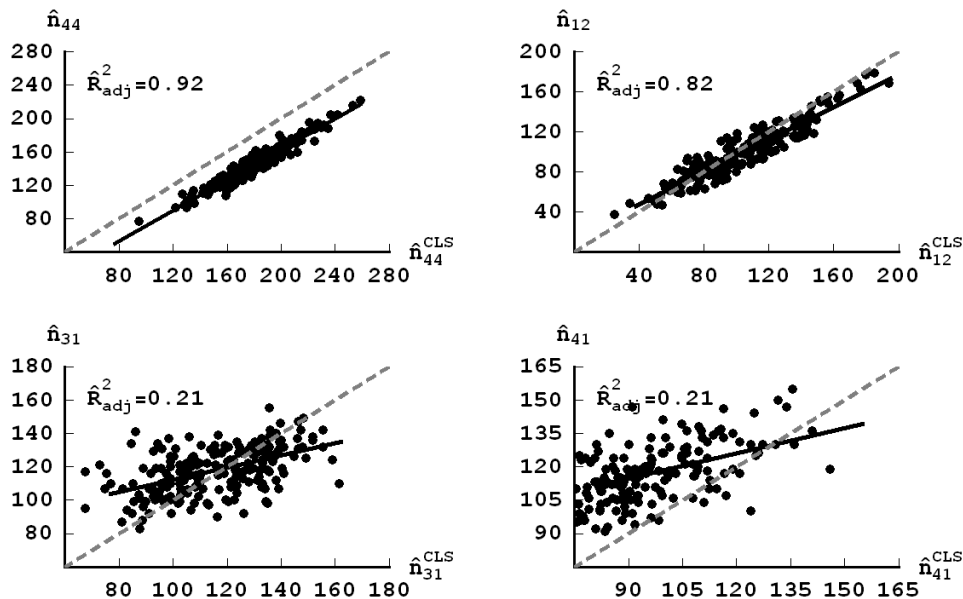

3×3 cluster was shifted one column to the left (right) or one row up (down) from its true position. The four directions of the shift were equally probable. With a probability of $P_2 = \frac{1}{3}(1 - P_0)$, the location of a cluster was shifted one column to the left (right) and one row up (down) relative to its true location. Again, the four possible shifts were equally probable. Accordingly, the mean coregistration error was varied from 0 to approximately 1.14 units in steps of ≈ 0.11. Note that, for a 3×3 cluster, the proportion of units that change cluster membership because of a registration error ranges from 0 ($P_0 = 1.0$) to $\approx 0.41(P_0 = 0)$. The assumed robustness of the proposed estimator(s) rests on these lower rates of units in error. For each level of P_0

and each of 200 replicates, the estimate based on a direct unit-level counts of change events ($\hat{\mathbf{n}}$) and the estimate from the proposed estimator $\hat{\mathbf{n}}_{\mathrm{CLS}}(\hat{\mathbf{n}}_{\mathrm{IPF}})$ were compared with the true "error-free" estimate in terms of mean absolute bias (MAD). Registration errors in each of the 200 replications of a random sample of 200 clusters were distributed at random across clusters.

Classification errors were assumed to be equal at both times of observation and to be independent both temporally and spatially. Given the true class k ($k = 1, 2, 3, 4$) of a unit, the probability that the classifier would classify the unit to class $k'(k' = 1, 2, 3, 4)$ is $P(\text{Classif.} = k'|\text{True} = k) = p_{k'|k}$. The following symmetric 4×4 confusion matrix $\mathbf{P}$ was used:

Fig. 3. Scatterplot of $\hat{n}_{kk'}^{\text{CLS}}$ versus $\hat{n}_{kk'}$ in 200 replicated samples of size $m = 200$ in Lambia, Italy (IT).
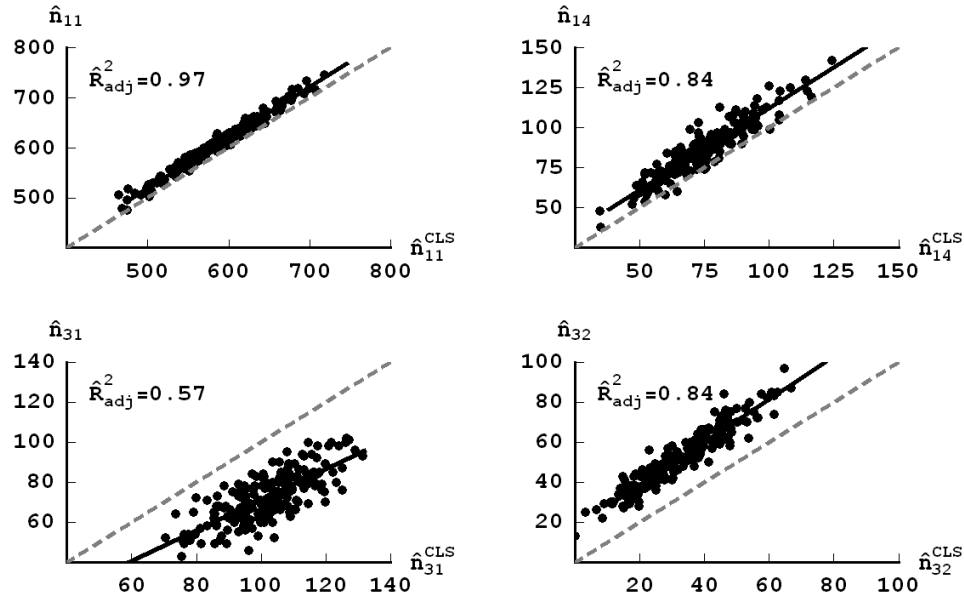

Fig. 4. Scatterplot of $\hat{n}_{kk'}^{\text{CLS}}$ versus $\hat{n}_{kk'}$ in 200 replicated samples of size $m = 200$ in Sussex, New Brunswick (NB).
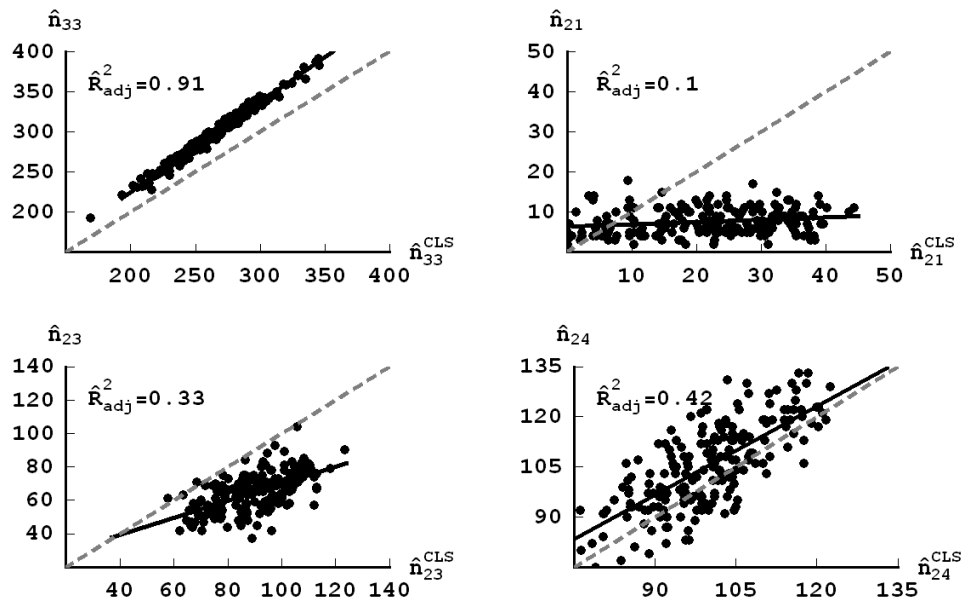
$$[3] \quad P = \begin{bmatrix} & \text{True} = 1 & \text{True} = 2 & \text{True} = 3 & \text{True} = 4 \\ \text{Classif.} = 1 & p_{k|k} & \frac{17}{24}\left(1 - p_{k|k}\right) & \frac{5}{24}\left(1 - p_{k|k}\right) & \frac{2}{24}\left(1 - p_{k|k}\right) \\ \text{Classif.} = 2 & \frac{17}{24}\left(1 - p_{k|k}\right) & p_{k|k} & \frac{2}{24}\left(1 - p_{k|k}\right) & \frac{5}{24}\left(1 - p_{k|k}\right) \\ \text{Classif.} = 3 & \frac{5}{24}\left(1 - p_{k|k}\right) & \frac{2}{24}\left(1 - p_{k|k}\right) & p_{k|k} & \frac{17}{24}\left(1 - p_{k|k}\right) \\ \text{Classif.} = 4 & \frac{2}{24}\left(1 - p_{k|k}\right) & \frac{5}{24}\left(1 - p_{k|k}\right) & \frac{17}{24}\left(1 - p_{k|k}\right) & p_{k|k} \end{bmatrix}$$

Fig. 5. Scatterplot of $\hat{n}_{kk'}^{CLS}$ versus $\hat{n}_{kk'}$ in 200 replicated samples of size $m = 200$ in Selangor, Indonesia (SE).
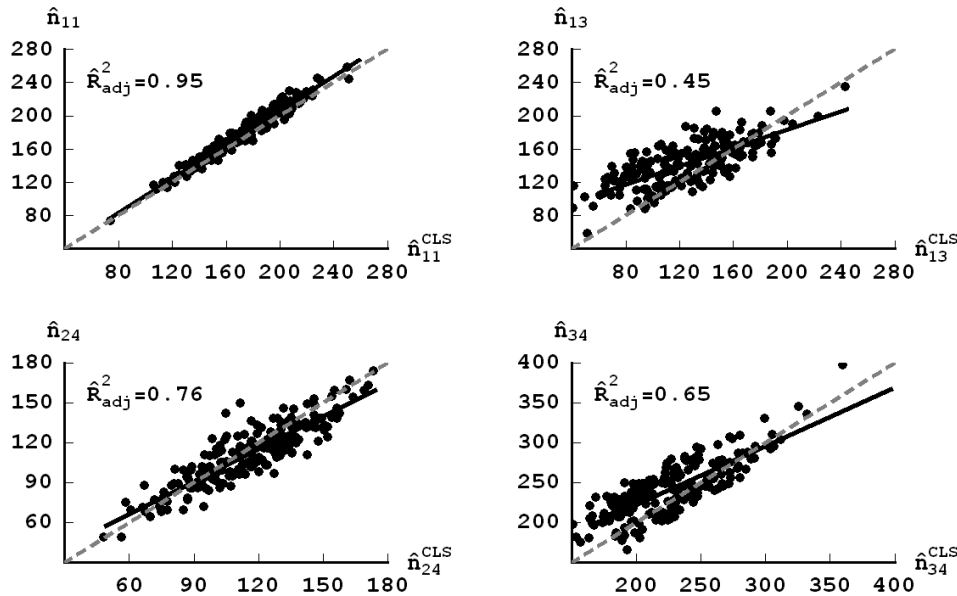
The value of $p_{k|k}$ is the classification accuracy (Congalton 2001), and it was varied from 0.5, 0.6,..., 0.9 with the value fixed during one simulation of 200 replications of 200 randomly sampled clusters. For each unit in a cluster, the classified class given the true class was determined by a random draw from a multinomial distribution with the conditional probabilities as in eq. 3 for a given value of $p_{k|k}$. Compensating errors (errors in opposite direction) are the main factors contributing towards the presumed robustness of the proposed estimator.

Five examples

Remotely sensed data from five large forested areas representing different regional landscapes with contrasting cover-type composition and rates of presumed change are used for demonstrating the performance of the proposed alternative estimator(s) of a change matrix when data are potentially error prone. The sites are near Prince George, British Columbia (BC); Hinton, Alberta (HI); Latium, Italy (IT); Sussex, New Brunswick (NB); and Selangor, Indonesia (SE). The data domains vary in size from 109 km^2 (BC) to 188 km^2 (IT). Cover-type classified Landsat-7 ETM+ images at two time-points 5–10 years apart (BC, HI, and SE) or a combination of a classified Landsat-7 ETM+ and a contemporary tessellated forest inventory cover-type map (IT and NB) furnished the change data. Although the changes in the latter two cases are dominated by disagreement between two classification processes, they were nevertheless deemed to mimic land-cover changes over an unspecified period of time. Further details on the data can be found in Magnussen (2004) or Magnussen et al. (2000). The original cover-type classes were reduced to a $K = 4$ class system by merging the most similar classes. Population sizes in ultimate units (a Landsat pixel of approximately 30 m × 30 m) were 121 104 (BC), 129 600 (HI), 208 675 (IT), 181 068 (NB), and 166 464 (SE). Population change matrices **n** as derived by counting of unit-level change events are given in Tables 3–7. The proportion of units in a change class varied

from a low of <0.1% to a high of 54% (HI class 1 × 1). Overall, between 14% (NB) and 52.3% (SE) of the population units changed class between the two (implied) points in time.

Results

Comparing bias of $\hat{n}_{CLS}$ and $\hat{n}_{IPF}$

Estimation of off-diagonal elements by IPF under the model of quasi-independence is orders of magnitude faster than with CLS. However, in two of the five examples (BC and HI), the mean absolute bias of off-diagonal elements in $\hat{n}_{IPF}$ was about 1.5 times larger than in $\hat{n}_{CLS}$ (Table 8). In the three remaining cases (IT, NB, and SE), the average performance was about equal (estimates were within 5% of each other). On two of these three sites (IT and NB), both estimators performed poorly (relative bias > 20%). Partial χ^2 test of the difference in goodness of fit confirmed that the IPF fit was significantly poorer than the CLS fit in BC and HI (P values under the hypothesis of equal goodness of fit were 0.002 and <0.001, respectively). No significant difference emerged on sites IT ($\hat{P} = 0.26$), NB ($\hat{P} = 0.55$), and SE ($\hat{P} = 0.18$). Unless prior knowledge exists to justify a quasi-independence model, it appears that the cumbersome CLS model provides a better fit than IPF when the off-diagonal elements are subject to nontrivial row by column interactions. Both $\hat{n}_{IPF}$ and $\hat{n}_{CLS}$ were significantly closer to the actual change matrix than an estimate based on a model of indendent rows and columns (no interaction). P values of partial χ^2 statistics were consistently 0.002 under the hypothesis of equal goodness-of-fit (Lloyd 1999).

Detailing the performance of $\hat{n}_{CLS}$

The bias of $\hat{n}_{CLS}$ estimates was in most cases nontrivial (Table 8). Diagonal elements were generally estimated with less bias than off-diagonal elements, a trend that is especially clear on a relative scale where the bias of diagonal elements, with the exception of BC, was 2–10 times lower

Fig. 6. Mean absolute bias (MAD) of elements of $\hat{\mathbf{n}}_{CLS}$ and $\hat{\mathbf{n}}$ (direct) plotted against mean coregistration error (unit: pixel of 30 m × 30 m). See Table 8 for site abbreviations.
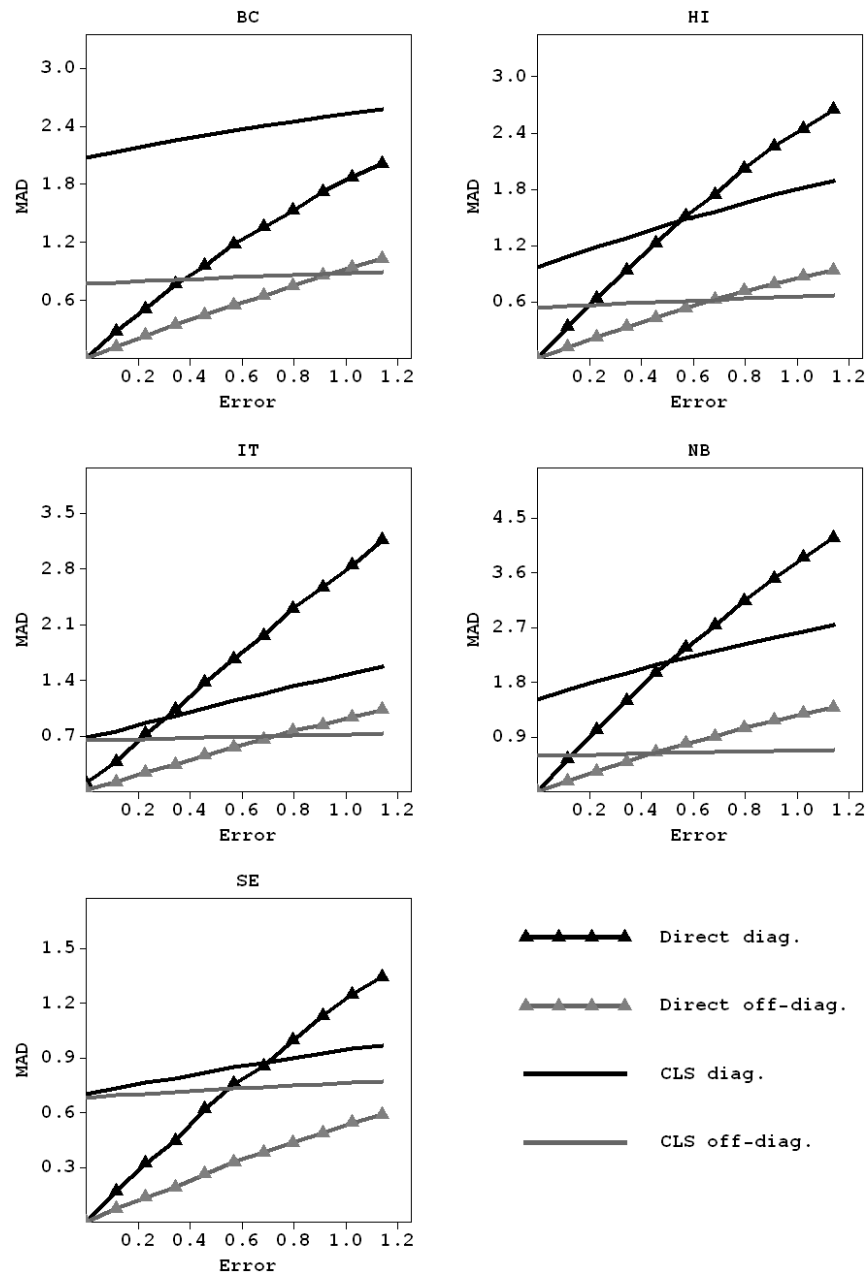
than for the off-diagonal change events. Relative bias of estimated diagonal elements was in the range of 5% to 10%.

Scatterplots in Figs. 1–5 of paired elements of $\hat{\mathbf{n}}_{CLS}$ and $\hat{\mathbf{n}}$ across 200 replications of the 200 randomly sampled clusters give additional insight to the performance of CLS. There are four plots per site: one for the diagonal element closest to the mean of the four diagonal elements and three for the off-diagonal elements with the highest true counts ($n_{kk'}$). In general, the ability to predict an off-diagonal element improves as $n_{kk'}$ increases. For the diagonal elements, the relationship was typically linear with $\hat{R}^2_{adj} = 0.91–0.97$ yet with a persistent bias. For the off-diagonal elements, the scatterplots suggest at linear relationship with a correlation

that varied from poor ($\hat{R}^2_{adj} < 0.3$) to fairly strong ($\hat{R}^2_{adj} = 0.7–0.9$) and, in most cases, a persistent bias.

Performance of $\hat{\mathbf{n}}_{CLS}$ when data is error prone

Coregistration errors in data from t_2 make the direct estimate ($\hat{\mathbf{n}}$) biased and also increased the bias of $\hat{\mathbf{n}}_{CLS}$. The bias of direct estimates increased with the magnitude of the mean coregistration error at a rate that was three to four times higher than seen in $\hat{\mathbf{n}}_{CLS}$ (Fig. 6). In other words, $\hat{\mathbf{n}}_{CLS}$ is less sensitive to coregistration errors, as expected. As a result, coregistration errors could increase to a point where the mean (absolute) bias in $\hat{\mathbf{n}}$ would be larger than in $\hat{\mathbf{n}}_{CLS}$. In BC, this happended for the off-diagonal elements at

Fig. 7. Mean absolute bias (MAD) of elements of $\hat{\mathbf{n}}_{CLS}$ and $\hat{\mathbf{n}}$ (direct estimate) plotted against classification accuracy (%). See Table 8 for site abbreviations.
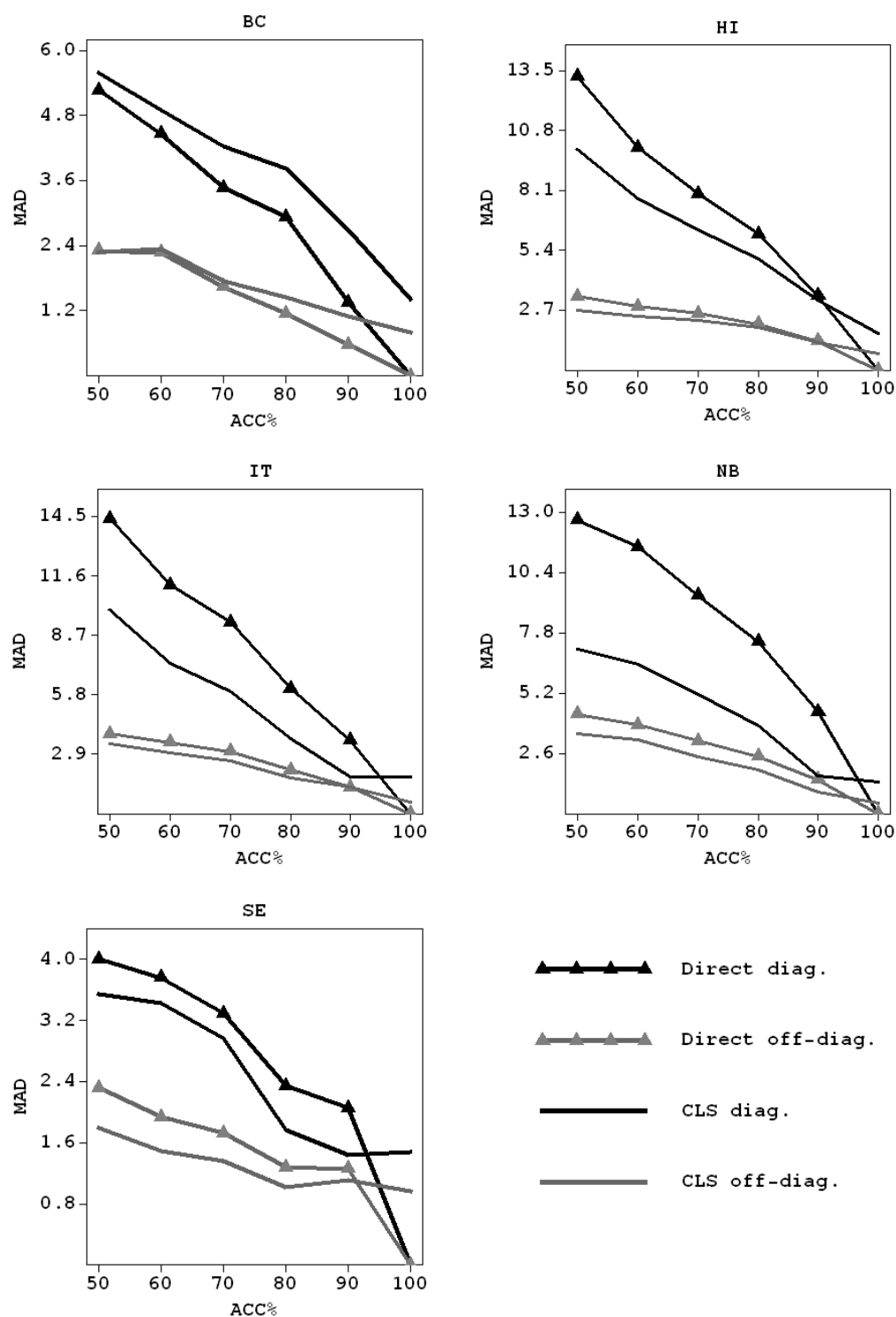
a mean coregistration error of 0.94 units (pixels), whereas the trend for the diagonal elements suggests that the crossover point might be at an unusual high coregistration error of 1.6 units (estimate obtained by linear extrapolation). For HI, the break-even point was at 0.5 for the diagonal elements and at 0.7 for the off-diagonal elements. Corresponding points for IT, NB, and SE were estimated at 0.3 and 0.4,

0.5 and 0.4, and 0.7 and 2.99 (by linear extrapolation), respectively. Coregistration errors beyond 1.1 should be rare in practice.

A less than perfect classification accuracy generates a serious bias in both the direct estimates ($\hat{\mathbf{n}}$) and in $\hat{\mathbf{n}}_{CLS}$ (Fig. 7). Bias of the diagonal elements increased at a rate that was three to six times higher than in off-diagonal ele-

ments. This was expected due to the high proportion of units (pixels) with no change on all five sites (from 0.5 in BC and SE to 0.8 and 0.9 in IT and NB, respectively). A classification error in these units generates a change event where none occurred. The rate at which bias increased as classification accuracy decreased was about 1.5 times higher in $\hat{\mathbf{n}}$ than in $\hat{\mathbf{n}}_{CLS}$. A critical accuracy level around 0.9, below which $\hat{\mathbf{n}}_{CLS}$ would be less biased than $\hat{\mathbf{n}}$, emerged in four of the sites (HI, IT, NB, and SE) but not in BC. On three sites (HI, IT, and NB), the differential in favor of $\hat{\mathbf{n}}_{CLS}$ appears important at accuracies <0.80.

Discussion and conclusions

A model-based estimation of a change matrix from cluster-level counts of units per class at two points in time is, of course, less accurate and less efficient than a direct estimation from observed unit-level change events. It is clear that pooling of unit-level information of change comes at a considerable price in terms of bias in estimates and a large increase in computationally costs of estimation. The proposed "robust" method of estimation is clearly only an option when unit-level change is seriously compromised by errors. To minimize the loss of change information, the formed data clusters should be small. The choice of a 3×3 cluster was a compromise between conflicting goals. A larger cluster would be more robust against errors of the kind studied here but also exponentially increases the number of possible change matrices with row and column sums identical to those observed for a single cluster, an increase that adversely affects the estimates. A smaller (e.g., 2×2) cluster would be less robust against registration and classification errors without offering any significant computational advantages.

Although the IPF approach to estimation of off-diagonal elements under the quasi-independence model for the change matrix (Booth et al. 2005) reduces the computational burden significantly — without seriously compromising the results — the computation remains an issue. When diagonal elements can be assumed to account for the majority of the row × column interactions in a change matrix (Von Eye and Spiel 1996), the IPF procedure is preferable. The proposed estimation procedure performed poorly for rare change events, as expected. Few events to "model" invariably produce poor estimates (Venette et al. 2002; Christman 2000; Thompson 1992). One might consider collapsing cells with low counts, as is often done in analysis of contingency tables, and then recover the estimates for these cells by some form of postcalibration (Kateri and Iliopoulos 2003; Agresti 1992).

The Monte Carlo simulations confirmed the sensitivity of unit-level change estimates derived directly from remotely sensed data to errors of coregistration and classification (Lunetta and Elvidge 1999; Bruzzone and Cossu 2003). Mean coregistration errors in the range of 0.3–1.1 units (pixels) are not uncommon even after concerted efforts have been invested in the coregistration process (Coulter et al. 2003; Bruzzone and Cossu 2003; Toutin 2004; Armston et al. 2002). Classification accuracies of 0.7–0.9 are commonly reported for Landsat ETM+ derived forest cover type maps (Foody 2002; Holmgren and Thureson 1998). Errors of this

magnitude generate a serious bias in change estimates, and every effort must be made to mittigate their negative effects. This may include bias reduction by calibration with a known confusion matrix or adopting advanced image-analytical procedures to reduce the mean coregistration error (Coulter et al. 2003; Bruzzone and Cossu 2003). In practice the two types of errors would co-occur and interact in a complex way. The examples illustrate that the impact is site specific, depending on the distribution of land cover patch sizes and the distribution of change events.

The proposed "robust" procedure was, as expected, less sensitive to coregistration and classification errors. It produced less biased results on three (four) of five sites when the mean coregistration error exceeds 0.3–0.5 or the classification accuracy dropped below 0.9. Because the performance improves as the error level increases, a point may be reached where, despite the computational costs, the proposed method does become an option to be considered. Change estimation from two independently produced vector-formatted land-cover maps (Stehman 2005; Kleinn et al. 2002; Biging et al. 1999) is probably the application domain that would qualify the most.

Improvements of the proposed estimation procedure are needed to reduce the bias in the estimates. A reliable estimator of the variance of an estimate is also needed. Although a bootstrap estimate of error can be obtained by fairly conventional means (Shao 1996), the computational burden is an issue

The amount of computational work needed to increase, by only 2, the rank of the linear constraints on the elements of the change matrix seems excessive. An empirical Bayes' approach with a Dirichlet prior (Bishop et al. 1975; Congalton 2001) on the change matrix and pseudodata in the form of all possible cluster-specific change matrices consistent with the observed counts $\mathbf{n}_1^{(i)}$ and $\mathbf{n}_2^{(i)}$, $i = 1, 2, ..., m$ deserves an investigation. Adopting a Bayesian framework would also make an estimate of the variance of estimates straightforward.

References

Agresti, A. 1992. A survey of exact inference for contingency tables. Stat. Sci. **7**: 131–153.

Agresti, A., and Caffo, B. 2000. Simple and effective confidence intervals for proportions and differences of proportions from adding two successes and two failures. Am. Stat. **4**: 280–200.

Anderson, A.B. 2002. Detecting changes in natural resources using land condition trend analysis data. Environ. Manage. **29**: 428–436. doi:10.1007/s00267-001-0017-z. PMID:11830771.

Armston, J.D., Danaher, T.J., Goulevitch, B.M., and Byrne, M.I. 2002. Geometric correction of Landsat MSS, TM, and ETM+ imagery for mapping of woody vegetation cover and change detection in Queensland. *In* Proceedings of the 11th Australasian Remote Sensing and Photogrammetry Conference, Sept. 2002, Brisbane, Australia. pp. 1–23.

Biging, G.S., Colby, D.R., and Congalton, R.G. 1999. Sampling

systems for change detection and accuracy assessment. *In* Remote sensing change detection. *Edited by* R.S. Lunetta and C.D. Elvidge. Taylor & Francis, London. pp. 281–308.

Bishop, Y.M.M., Fienberg, S.E., and Holland, P.W. 1975. Discrete multivariate analysis: theory and practice. MIT Press, Cambridge, Mass.

Booth, J.G., Capanu, M., and Heigenhauser, L. 2005. Exact conditional *P* value calculation for the quasi-symmetry model. J. Comput. Graph. Stat. **14**: 716–725.

Brier, S.S. 1980. Analysis of contingency tables under cluster sampling. Biometrika, **67**: 591–596.

Brown, S. 2002. Measuring, monitoring, and verification of carbon benefits for forest-based projects. Philos. Trans. R. Soc. London Ser. A Math. Phys. Eng. Sci. **360**: 1669–1683.

Bruzzone, L., and Cossu, R. 2003. An adaptive approach to reducing noise effects in unsupervised change detection. IEEE Trans. Geosci. Remote Sens. **41**: 2455–2465. doi:10.1109/TGRS.2003. 817268.

Christman, M.C. 2000. A review of quadrat-based sampling of rare, geographically clustered populations. J. Agric. Biol. Environ. Stat. **5**: 168–201.

Collins, J.B., and Woodcock, C.E. 1999. Geostatistical estimation of resolution-dependent variance in remotely sensed images. Photgr. Eng. Rem. Sens. **65**: 41–50.

Congalton, R.G. 2001. Accuracy assessment and validation of remotely sensed and other spatial information. Int. J. Wildland Fire, **10**: 321–328. doi:10.1071/WF01031.

Coomes, D.A., Allen, R.B., Scott, N.A., Goulding, C., and Beets, P. 2002. Designing systems to monitor carbon stocks in forests and shrub-lands. For. Ecol. Manage. **164**: 89–108. doi:10.1016/ S0378-1127(01)00592-8.

Coppin, P.R., and Bauer, M.E. 1996. Digital change detection in forest ecosystems with remote sensing imagery. Remote Sens. Environ. **13**: 207–234.

Coppin, P., Jonckheere, I., Nackaerts, K., Muys, B., and Lambin, E. 2004. Digital change detection methods in ecosystem monitoring: a review. Int. J. Remote Sens. **25**: 1565–1596. doi:10. 1080/0143116031000101675.

Corona, P., Chirici, G., and Marchetti, M. 2002. Forest ecosystem inventory and monitoring as a framework for terrestrial natural renewable resource survey programmes. Plant Biosyst. **136**: 69–82.

Coulter, L.L., Stow, D.A., and Baer, S. 2003. A frame center matching technique for precise registration of multitemporal airborne frame imagery. IEEE Trans. Geosci. Remote Sens. **41**: 2436–2444. doi:10.1109/TGRS.2003.819191.

Czaplewski, R.L. 2003. Accuracy assessment of maps of forest conditions: statistical designs and methodological considerations. *In* Remote sensing of forest environments: concepts and case studies. *Edited by* M.A. Wulder and S.E. Franklin. Kluwer Academic Publishers, Boston. pp. 115–140.

Czaplewski, R.L., and Catts, C.P. 1992. Calibration of remotely sensed proportion or area estimates for misclassification error. Remote Sens. Environ. **39**: 29–43. doi:10.1016/0034-4257(92) 90138-A.

Fleiss, J.L. 1981. Statistical methods for rates and proportions. Wiley, New York.

Foody, G.M. 2002. Status of land cover classification accuracy assessment. Remote Sens. Environ. **80**: 185–201. doi:10.1016/ S0034-4257(01)00295-4.

Garcia-Gigorro, S., and Saura, S. 2005. Forest fragmentation estimated from remotely sensed data: comparison across scales possible? For. Sci. **51**: 51–63.

Green, E.J., Köhl, M., and Strawderman, W.E. 1992. Improved es-

timates for cell values in a two way table. Biometr. Inf. Med. Biol. **23**: 24–30.

Holmgren, P., and Thureson, T. 1998. Satellite remote sensing for forestry planning — a review. Scand. J. For. Res. **13**: 90–110.

Kateri, M., and Iliopoulos, G. 2003. On collapsing categories in two-way contingency tables. Statistics, **37**: 443–455.

Kleinn, C., Corrales, L., and Morales, D. 2002. Forest area in Costa Rica: a comparative study of tropical forest cover estimates over time. Env. Monit. Assess. **73**: 17–40.

Lloyd, C.J. 1999. Analysis of categorical variables. Wiley, New York.

Lunetta, R.S., and Elvidge, C.D. 1999. Remote sensing change detection. Environmental monitoring methods and applications. Taylor & Francis, London.

Lunetta, R.S., and Lyon, J.G. 2004. Remote sensing and GIS accuracy assessment. CRC Press, Boca Raton, Fla.

Magnussen, S. 2004. Prediction of 2 × 2 tables of change from repeat cluster sampling of marginal counts. Can. J. For. Res. **34**: 1703–1713. doi:10.1139/x04-041.

Magnussen, S., and Köhl, M. 2005. A variance estimator for constrained estimates of change in relative categorical frequencies. Environ. Monit. Assess. **105**: 391–410. doi:10.1007/s10661-005-4437-8. PMID:15952530.

Magnussen, S., Boudewyn, P., Wulder, M.A., and Seemann, D. 2000. Predictions of forest inventory cover type proportions using Landsat™. Silva Fenn. **34**: 351–370.

Magnussen, S., and Köhl, M. 2006. A better alternative to Wald's goodness-of-fit test. For. Ecol. Manage. **121**: 123–132.

Murtaugh, P.A., and Phillips, D.L. 1998. Temporal correlation of classifications in remote sensing. J. Agric. Biol. Environ. Stat. **3**: 99–110.

Parr, T.W., Ferretti, M., Simpson, I.C., Forsius, M., and Kovacs-Lang, E. 2002. Towards a long-term integrated monitoring programme in Europe: network design in theory and practice. Environ. Monit. Assess. **78**: 253–290. doi:10.1023/ A:1019934919140. PMID:12374392.

Pontius, R.G., Jr., and Lippitt, C.D. 2006. Can error explain map differences over time? Cartogr. Geogr. Inf. Sci. **33**: 159–171.

Reams, G.A., and Van Deusen, P.C. 1999. The southern annual forest inventory system. J. Agric. Biol. Environ. Stat. **4**: 346–360.

Rosenfield, G.H., Fitzpatrick-Lins, K., and Ling, H.S. 1982. Sampling for thematic map accuracy testing. Photogramm. Eng. Remote Sens. **48**: 131–137.

Schroeder, T.A., Cohen, W.B., Song, C., Canty, M., and Yang, Z. 2006. Radiometric correction of multi-temporal Landsat data for characterization of early successional forest patterns in western Oregon. Remote Sens. Environ. **103**: 16–26. doi:10.1016/j.rse. 2006.03.008.

Searle, S.R. 1982. Matrix algebra useful for statistics. Wiley, New York.

Shao, J. 1996. Bootstrap model selection. J. Am. Stat. Assoc. **91**: 655–665. doi:10.2307/2291661.

Smith, P., and Annoni, A. 1999. Spatial analysis of land-use changes as knowledge tools in support of European spatial policies and ecosystem health. Ecosyst. Health, **5**: 275–284.

Song, C., Woodcock, C.E., Seto, K.C., Lenney, M.P., and Macomber, S.A. 1999. Classification and change detection using Landsat TM data: when and how to correct atmospheric effects? Remote Sens. Environ. **75**: 230–244.

Stehman, S.V. 2005. Comparing estimators of gross change derived from complete coverage mapping versus statistical sampling of remotely sensed data. Rem. Sens. Env. **96**: 466–474.

Stehman, S.V., and Czaplewski, R.L. 1998. Design and analysis for thematic map accuracy assessment: Fundamental principles. Re-

mote Sens. Environ. **64**: 331–344. doi:10.1016/S0034-4257(98)00010-8.

Thompson, S.K. 1992. Sampling. Wiley, New York.

Toutin, T. 2004. Geometric processing of remote sensing images: models, algorithms and methods. Int. J. Remote Sens. **25**: 1893–1924.

Townsend, J.R.G., Huang, C., Kalluri, S.N.V., Defries, R.S., Liang, S., and Yang, K. 2000. Beware of per-pixel characterization of land cover. Int. J. Remote Sens. **21**: 839–843.

Van Deusen, P.C. 1994. Correcting bias in change estimation from

thematic maps. Remote Sens. Environ. **50**: 67–73. doi:10.1016/0034-4257(94)90096-5.

Venette, R.C., Moon, R.D., and Hutchison, W.D. 2002. Strategies and statistics of sampling for rare individuals. Annu. Rev. Entomol. **47**: 143–174. doi:10.1146/annurev.ento.47.091201.145147. PMID:11729072.

Von Eye, A., and Spiel, C. 1996. Standard and nonstandard log-linear symmetry models for measuring change in categorical variables. Am. Stat. **50**: 300–305. doi:10.2307/2684924.