Available online at www.sciencedirect.com
SCIENCE DIRECT®
Remote Sensing of Environment 87 (2003) 161–170
Remote Sensing of Environment
www.elsevier.com/locate/rse

Updating cover type maps using sequential indicator simulation

Steen Magnussen[a,*], Sytze de Bruin[b,1]

[a] Canadian Forest Service, Natural Resources Canada, 506 West Burnside Road, Victoria, BC Canada V8Z 1M5
[b] Centre Geo-Information, Wageningen University, P.O. Box 339, 6700 AH, Wageningen, The Netherlands

Received 28 January 2003; received in revised form 16 April 2003; accepted 21 May 2003

Abstract

Maximum posterior probability (MAP) maps of forest inventory (FI) cover type classes were produced from a maximum likelihood (ML) classified TM image and 5% (2%) systematic reference sampling of actual cover types for of nine 2 × 2 km study sites in New Brunswick, Canada. MAP cover type maps were obtained via sequential indicator simulation (SIS) using collocated indicator cokriging. A 5% reference sampling increased the coefficient of accuracy of MAP cover type maps by about 0.2 compared to the accuracy of the ML classified maps. MAP prediction errors were obtained for global and small area estimates of cover type extent. MAP-based cover type statistics of extent and precision were compatible with corresponding results for maximum likelihood bias-corrected estimates (MLE). Spatial autocorrelation of MAP prediction errors declined rapidly with distance and were near 0 for distances of more than 3–4 Landsat TM pixels. MAP cover type maps produced by SIS are attractive when both global and local estimates of precision of map-derived statistics are needed.

Keywords: Map accuracy; Maximum likelihood; Collocated indicator cokriging; Prediction error; Spatial autocorrelation; Small area estimation

1. Introduction

Forest cover type classification of a single remote sensing image by maximum likelihood (ML) or by clustering techniques (Andersen, 1998; Franklin, Gillespie, Titus, & Pike, 1994; Friedl & Broadley, 1997) is the common and most basic approach to transform a satellite image to a cover type map. Map accuracy is generally estimated from an independent reference sample (Stehman, 1999; Stehman & Czaplewski, 1998). Reference data are used for estimation of map accuracy and bias-corrected marginal cover type frequencies (Card, 1982; McRoberts, Wendt, Nelson, & Hansen, 2002; Tenenbein, 1972). However, the classified map, usually remains fixed and disconnected from the results of the reference sample. Ideally, the map should be updated by integrating the results of the reference sample (Moody & Woodcock, 1996). Spatial smoothing (Cressie, 1991), nearest neighbour smoothing (Franco-Lopez, Ek, & Bauer, 2001), and interpolation techniques (Soares, 1998, 2001) can achieve this integration. However, the accuracy of the updated map, while improved, remains unknown unless an estimate of the spatial distribution of prediction errors is obtained as part of the update process.

Global estimates of map accuracy are useful in their own right but do not apply directly to small areas of the map. Map accuracy is not constant across the map and map errors are generally spatially autocorrelated (Conese & Maselli, 1992; Foody, 1999, 2002; Masselli, Conese, & Petkov, 1994; Stehman & Czaplewski, 1998). Integration of the reference data into an updated cover type map and estimation of a spatial distribution of prediction errors would seem desirable from a map user's perspective.

The contiguous spatial extent of forest cover types suggests that the reference sample and the classified image can be integrated into a spatial prediction model of cover types (Atkinson & Lewis, 2000; Kyriakidis & Dungan, 2002; Rossi, Dungan, & Beck, 1994). Solow (1986) was among the first to suggest mapping a binary set of classes by simple indicator kriging from a set of points of known cover type. Solow further showed that the mapping was unbiased for an isotropic stationary random binary field. Goovaerts (1997) gives an overview of methods for spatial predictions of categorical attributes.

* Corresponding author. Tel.: +1-250-363-0712; fax: +1-250-363-0775.
E-mail addresses: smagnuss@nrcan.gc.ca (S. Magnussen), sytze.debruin@wur.nl (S. de Bruin).
[1] Fax: +31-317-41-90-0.

0034-4257/$ - see front matter © 2003 Published by Elsevier Inc.
doi:10.1016/S0034-4257(03)00138-X

Yet, the accuracy of a cover type statistic derived from a map that has been updated by integration of reference data remains unknown, unless an estimate of the spatial distribution of prediction errors is provided. Estimation of the spatial distribution of prediction errors is complicated by the conditional nature of the integration process and spatial interdependencies (Cressie, 1991). Sequential indicator simulation (SIS) in which indicators (here synonymous with cover types) are predicted for all map pixels in a random sequential order (Deutsch & Journel, 1998; Goovaerts, 1997; Soares, 1998) is well suited for this estimation problem. By repeating the SIS process, a large number of times one obtains a series of stochastic realizations of a cover type map suited for estimation of the spatial distribution of prediction errors.

The objective of our study is to produce a series of cover type maps compatible with maximum likelihood bias-corrected estimates (MLE) of relative cover type extent and to estimate, from this series of maps, global and local estimates of accuracy. The objective is pursued by extending the SIS approach of De Bruin (2001) to a multivariate setting of cover type mapping (Gaudard, Karson, Linder, & Sinha, 1999; Klein & Press, 1992). De Bruin demonstrated how SIS with prediction via collocated indicator cokriging could be used to generate a series of stochastic map realizations from which the marginal distribution of the extent of a particular olive grove could be obtained. Nine 2×2 km study sites within a single Landsat TM image scene from New Brunswick, Canada are used for our purpose.

2. Material and methods

Generation of a maximum posterior probability (MAP) cover type map involved four steps: (i) image classification; (ii) sampling of reference data; (iii) SIS with collocated indicator cokriging predictions; and (iv) MAP cover type labeling.

2.1. Study sites and image data

We worked with nine 2×2 km study sites (plots) located on a 20×20 km grid in the southeastern part of New Brunswick, Canada. The size and locations of the study sites were determined by the national forest inventory grid (Natural Resources Canada, 1999).

A Landsat 5 TM scene (Track 9, Frame 28) from July 31, 1995, covered all nine areas. The image was geometrically corrected with 37 control points obtained from forest inventory (FI) maps. The estimate of the overall root mean square positional error based on a first-order nearest neighbor sampling technique was 0.46 pixels (Glasby & Horgan, 1995). No radiometric correction was attempted since reliable data on local atmospheric conditions were not available. The image was cloud-free over the nine areas.

2.2. Image classification

All pixels in the study areas were classified into one of seven cover type classes (Table 1) using a Gaussian ML classifier (McLachlan, 1991). The classifier was trained on 7000 pixels with known cover type in a contiguous area separated from the nine study sites by a minimum distance of 20 km. The known cover types were obtained from a contemporary FI cover type map (New Brunswick Department of Natural Resources and Energy, Forest Management Branch). For computational convenience, the first four principal components (explaining 94.5% of the variation in the data) served as predictors of cover type class (Magnussen, Boudewyn, Wulder, & Seemann, 2000).

2.3. Reference data

Reference data for accuracy assessment and bias correction were obtained by a systematic grid sampling with a random start location of pairs of coregistered pixels in the classified image and in the FI cover type map. FI data serve as a benchmark reference. Two sampling intensities of 5% and 2% were employed, yielding 217 and 72 pixels located on a regular grid within each 2×2 km study area. Reference sample data were used in the SIS procedure (see below), for bias correction of ML classifications, and for a direct estimation of the extent of a cover type. The latter applies to both study areas (S) and to small areas of interest (polygons) where only samples inside a polygon of interest are used (SIP).

The statistical inference applied to reference sample data use estimators applicable to simple random sampling (Köhl, Innes, & Kaufmann, 1994) since systematic and simple random sampling designs have identical variance estimators when the sample attribute is a binary (indicator variable) (Collett, 1991; Särndal, Swensson, & Wretman, 1992).

2.4. SIS

The main objective of SIS is to produce an estimate of the spatial uncertainty given the available data (Goovaerts,

Table 1
Cover-type classes

Class	Abbreviation	Examples
Exposed land	el	roads, railroad, airstrips, cities, towns, military impact zone, winter road
Shrubs	sh	barren land, wetland, cultivated blueberry, burn, scattered outcrops, newly planted after clearcutting, alder thickets
Hardwoods	tb	hardwood species >80% by volume
Mixedwoods	tm	hardwoods and softwoods are 21–79% by volume
Softwoods	tc	softwoods >80% by volume
Water	wa	lakes, ponds, rivers, stream, ocean
Crops	cp	cultivated farm land, pastures, fields

1997 pp. 376–379). For each pixel in a random path of all nonreference pixels in a study area, SIS first computes a conditional probability p_i* of membership in a given cover type class i ($i=1,\ldots,7$) given the 'soft' and 'hard' data for the eight closest pixels for which 'hard' data in the form of a reference FI cover type class or a prior SIS cover type exist. Soft data are vectors of the probability of cover type class memberships given the pixel's spectral signature as generated during the ML classification process. A stochastic cover type for the pixel is then simulated by choosing a class i randomly with probability proportional to p_i*. Stationarity of the simulated pixel labels is thus conserved. Further details are below.

The probability of membership in a cover type class (p_i*) was estimated from (Goovaerts, 1997, p. 313, Eq. 7.46):

$$p_i^* = [\boldsymbol{\delta}_i \quad p_i' \,] \times \begin{bmatrix} \boldsymbol{\lambda}_{i1} \\ \\ \lambda_{i2} \end{bmatrix} \tag{1}$$

where $\boldsymbol{\delta}_i$ is an 8×1 column vector of nearby cover type indicator variables $\delta_i(j)$, $j=1,\ldots,8$ with $\delta_i(j)=1$ if the predicted or actual cover type class of the jth nearest pixel is i and 0 otherwise, p_i' is the pixel-specific adjusted probability of membership in class i as per the ML classification (adjusted via a multinomial logistic regression to prevent bias), $\boldsymbol{\lambda}_{i1}$ is an 8×1 vector of class indicator weights, and λ_{i2} is the weight given to the collocated 'soft datum' (p_i'). The weights combine hard and soft data in a way that minimizes the root mean square error of predictions. A sum-to-one constraint is imposed on the weights to prevent bias. Any p_i* value outside the interval [0;1] was replaced by the closest value inside the allowed interval. A sum-to-one condition was assured by dividing each p_i* by $\sum_i p_i$*. Eq. (1) is identical to the prediction equation used in collocated indicator cokriging (Goovaerts, 1997). Note that p_i* depends on the random path through the pixels. We assume that the conditions necessary for kriging have been met, i.e., second-order stationarity and ergodic, and isotropic spatial covariance processes. Binary indicator variables are less sensitive to violations of these conditions than continuous variables (Solow, 1986). Simple indicator kriging with a local prior (the classified image) was tried as a natural alternative to the collocated indicator cokriging but was found to be less performing, probably due to the nonlinear and variable relationship between the soft and hard data. Increasing the number of neighbouring pixels from 8 to 10 or even 12 had virtually no effect on predictions. Further details are in Appendix A.

The above simulations generated, as expected, nonsensical results along edges of water bodies due to smoothing of local contrast. In an attempt to maintain a high ML classification accuracy for water, our search protocol for the eight nearest neighbours of any pixel located on or within one pixel from a water's edge in the classified image was therefore restricted to either the land or the water side of that edge. Edges were identified with a Roberts filter (Gonzales & Woods, 1992; Pratt, 1991). A similar adjustment was considered for string-like features in the classified image but the rules needed to prevent potential kriging artefacts were deemed too restrictive for the present purpose.

2.5. MAP cover type map

A MAP cover type map was obtained for each area by assigning to each pixel the cover type class predicted most often for that pixel during 200 repeat SIS runs. Preliminary results indicated that 200 SIS realizations sufficed to keep the chance that a different set of 200 realizations would change the MAP cover type of pixel below 0.01.

2.6. MAP prediction errors

Cover type assignments during a single SIS realization depend on the specific constellation of (simulated) hard and soft data used for the collocated indicator cokriging prediction for a single pixel. Errors obtained after a single SIS run will be biased due to this dependency. Instead, we consider the 200 SIS cover type assignments for a particular pixel as an estimate of the distribution of predictions over all possible constellations of hard and soft data for that pixel. The mean and variance of this distribution has lessened the above dependency (through averaging) to a level assumed insignificant for the intended use. Spatial covariance of predictions are obtained by extending this principle to pairs of pixels. Specifically, a SIS label for a pixel is either equal to or different from the final MAP prediction. Hence, the 200 SIS realizations can be transformed into 200 binary maps taking the value 1 if the realized cover type equals the MAP prediction and 0 otherwise. We compute the variance and covariances of MAP predictions from these binary maps. Specifically, we first obtained a 7×200 matrix (**SI**) of indicator variables for each pixel in an area. The elements of **SI** are δ_{ij} ($i=1,\ldots,7$; $j=1,\ldots,200$), where $\delta_{ij}=1$ if the jth SIS cover type is equal to the final MAP prediction and 0 otherwise. Variances and covariances of MAP predictions were then computed by standard procedures applied to these **SI** matrices (Searle, 1982). For two pixels, say i and j, their 7×7 prediction covariance (or variance if $i=j$) matrix is $\hat{\boldsymbol{\Omega}}_{ij} = 1/(200)\mathbf{SI}_i \times \mathbf{SI}_j' - 1/(200^2)(\mathbf{SI}_i \times \mathbf{J}')(\mathbf{SI}_j \times \mathbf{J}')$ where $\mathbf{SI}'$ is the transpose of **SI** and **J** is a 7×200 matrix of ones. The MAP prediction variance associated with a specific MAP cover type map area (A) is simply $\sum_{i \in A} \sum_{j \in A} \hat{\boldsymbol{\Omega}}_{ij}$ (Matern, 1980). Note, the above variances (covariances) are invariant under a permutation of the binary labeling convention (Collett, 1991). We provide estimates of MAP prediction errors of cover type extent for each study area and for local application to small areas (polygons) of interest within each study area.

2.7. Coefficient of map accuracy (Kappa)

The accuracy of MAP and ML cover type maps was gauged by the Kappa coefficient of map accuracy ($0 \le \kappa \le 1$) (Hudson, 1987; Lunetta et al., 1991). A perfect map has $\kappa = 1.0$ while a map where classification is no better than chance has $\kappa = 0$. Approximate standard errors of Kappa were computed as outlined by, for example, Czaplewski (1994). The accuracy is computed only for pixels *not* in the reference sample. Kappa coefficients of MAP and ML were compared statistically by means of a one-sided test of the hypothesis that the accuracy of a MAP-based cover type map is higher than that of a ML classified map.

2.8. MLE of cover type extent

Estimates of cover type extent obtained by pixel counts from the classified maps are biased due to imperfection of the ML classifier (Cochran, 1977). MLEs of cover type extent were obtained by premultiplying the vector of relative ML classified cover type frequencies with the matrix of reference-sample-based estimates of the probabilities that the FI class is i given that the classified class is j ($i,j = 1,\ldots, 7$) (Card, 1982; Tenenbein, 1972). Precision of MLEs was estimated by standard procedures (Card, 1982; Tenenbein, 1972). Localized MLEs for small areas (polygons) of interest were obtained in a similar way by limiting the computations to the pixels representing the polygon(s) in question.

2.9. MAP, MLE, and FI statistics

To infer whether the MAP cover type maps are compatible with associated MLEs of cover type extent, a comparison was done under the null hypothesis of no difference. The comparisons were done for global map estimates and estimates applicable to small areas of interest. Our comparison only serves to benchmark the MAP results.

Paired t-tests (Snedecor & Cochran, 1971) gauged the significance of the average absolute difference (MAD) between a MAP estimate and a MLE across the nine study areas. Average squared differences were tested for significance with a t-test of log-transformed estimates (Lin, Hedayat, Sinha, & Yang, 2002). Confidence intervals (95%) of MAD absolute difference were determined as outlined by Lin et al. (2002). Estimated precision (root mean square errors) of estimates of cover type extent were compared and tested for statistical significant differences with a randomization test (Good, 1993). Wilcoxon's signed rank test statistic (z_w) for paired observations (Conover, 1980) was computed after each round of randomization of the precision estimates. The probability of obtaining the observed z_w value under the null hypothesis was estimated as the relative frequency by which the test statistic under randomization exceeds the observed value.

MLE and MAP estimates of cover type extent were also compared to their corresponding FI benchmarks in order to gauge accuracy, and also against the reference sample estimates (S) in order to gauge the overall gain in precision attributable to either the ML classifications or to the MAP predictions. Statistical inference for these comparisons was the same as outlined above. The agreement between FI and MAP (viz. MLE) estimates was quantified by both the concordance correlation coefficient ($0 \le \rho_c \le 1$) and the coefficient of accuracy ($0 \le \rho_a \le 1$) measures of reproducibility and accuracy of a method (Lin et al., 2002).

3. Results

3.1. Intensive reference sampling

Examples of the MAP cover type maps obtained after 200 SIS with collocated indicator cokriging predictions of cover type are provided in Fig. 1. Kappas of MAP predictions for pixels not included in the reference sample were, as expected, greater than Kappas of the ML classification. The Kappa of each of the nine MAP cover type maps was higher (0.23–0.45) than the corresponding Kappa of a ML classified map ($P < 0.01$). The mean of the former was 0.60 ± 0.04, while the mean of latter was 0.42 ± 0.05. The improvement in the per-pixel accuracy achieved by the use of the reference sample in SIS is manifested. Related to the increased accuracy of the MAP

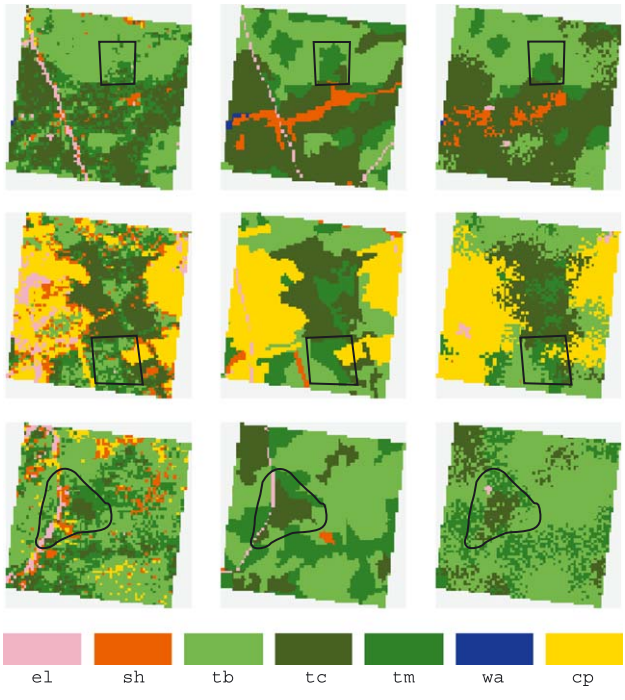

Fig. 1. Cover type maps for three of the nine study areas. Left column: ML classified TM cover type maps. Center column: photo-interpreted forest inventory (FI) cover type maps (benchmark). Right column: MAP cover type maps. Pixels included in the 5% reference sample have the same FI cover type across columns. Small areas of interest (polygons) are delineated by a black line. See Table 1 for the cover type legend.

Table 2
Average relative frequencies of left-to-right joined pixels with identical cover type (5% reference sampling)

	el	sh	tb	tc	tm	wa	cp	$\hat{X}^{2a}$
FI	0.02	0.05	0.25	0.23	0.17	0.16	0.12	na
MAP	0.02	0.05	0.26	0.23	0.14	0.16	0.14	88
ML	0.03	0.04	0.31	0.24	0.08	0.19	0.11	359

[a] χ^2 statistic of squared deviations from the true distribution.

cover type maps is an improvement in the predicted frequency distribution of joined (left-to-right) pixels with the same cover type (Table 2).

Fig. 2 shows examples of the spatial distribution of the relative frequency of prediction of a MAP cover type. Although the fine detail has been suppressed, it is clear that the distribution is not uniform across a study area. A patchy mosaic of higher and lower relative frequencies is the norm. This mosaic only partly reflects the joint effects of the spatial distribution of cover types and the cover-type-specific ML classification accuracies. Across all pixels the relative frequency of a MAP prediction was about 0.08 higher than that of the ML classification. As expected, a MAP prediction for a pixel next to a reference sample pixel was, in general, more accurate (2–30%) than a prediction made for pixels further away from a reference sample pixel. MAP prediction errors tend to increase along FI boundaries between two cover types due to the smoothing of local variability.

MAP cover type prediction errors were, as expected, spatially correlated. The extent of the spatial covariance of prediction errors is illustrated in Fig. 3 by means of average standardized covariances (Pearson correlation coefficients) plotted against the distance (in pixels) between the pairs of pixels for which the predictions are made. Predictions appear to be independent once the separating distance is more than three to five pixels. A correlation as high as 0.20–0.30 was only observed for pixels separated by less than two pixels (approximately 60 m). This coincides with an observed median run lengths of about four pixels with the same land

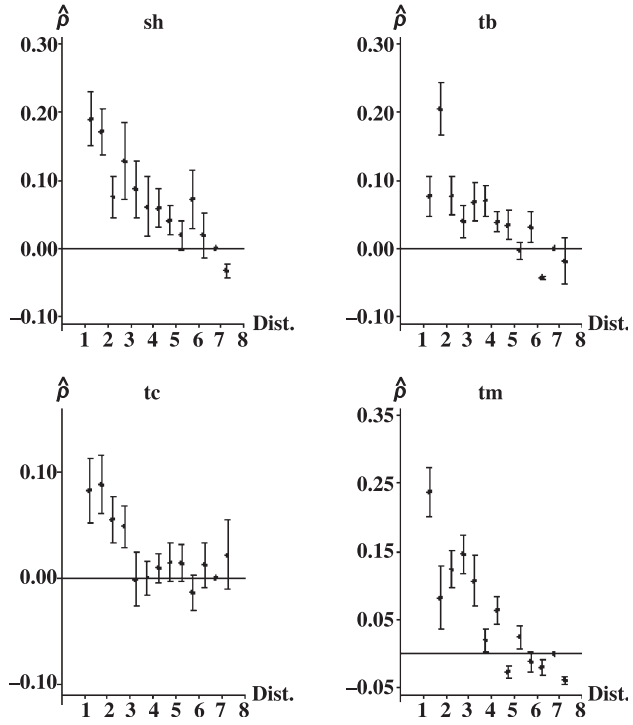

Fig. 3. Mean spatial autocorrelation of MAP cover type class prediction errors between pairs of pixels (Pearson correlation coefficients) plotted against the distance (Dist.) separating them (Dist. is measured in pixels with a nominal side length of 30 m). Standard errors of the means are indicated by the vertical error bars. Results are based on a simple random sample of 1500 pairs of pixels located no more than 8 pixels apart.

cover type (range for individual study areas was 2 (exposed land) to 16 (crops)). The trend in spatial autocorrelation of prediction errors appears to depend on the cover type.

All MAP estimates of relative cover type extent were within 0.02 of their corresponding FI benchmarks, overall, the average bias was 0.01 (Table 3). MLEs were, as a rule, slightly closer to their FI benchmarks (max absolute deviation ≤ 0.01). Yet only the MAP estimate of exposed land deviated significantly ($P=0.04$) from the FI bench-

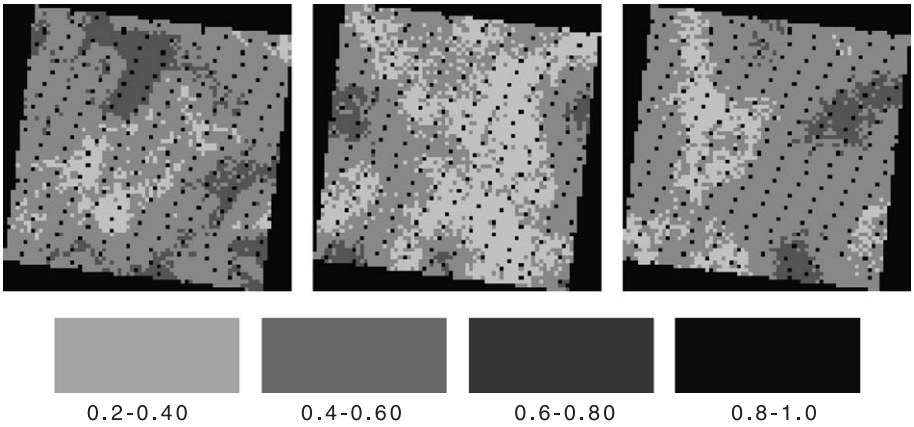

Fig. 2. Relative frequency of MAP cover type class predictions in three study areas (same as in Fig. 1). Pixels included in the reference sample are identified by a relative frequency of 1.0.

Table 3
Summary statistics of relative cover-type extent (π) for the 5% reference sampling

	el	sh	tb	tc	tm	wa	cp
π_{FI}	0.03 ± 0.03	0.05 ± 0.07	0.25 ± 0.16	0.23 ± 0.17	0.18 ± 0.09	0.14 ± 0.19	0.12 ± 0.14
$\hat{\pi}_S$	0.03 ± 0.03	0.05 ± 0.07	0.26 ± 0.16	0.23 ± 0.17	0.18 ± 0.10	0.14 ± 0.19	0.13 ± 0.15
$\hat{\pi}_{MLE}$	0.03 ± 0.03	0.05 ± 0.07	0.24 ± 0.16	0.23 ± 0.17	0.18 ± 0.10	0.13 ± 0.19	0.13 ± 0.16
$\hat{\pi}_{MAP}$	0.03 ± 0.03	0.05 ± 0.09	0.26 ± 0.19	0.24 ± 0.19	0.16 ± 0.10	0.13 ± 0.18	0.13 ± 0.16
$\hat{t}_{\lvert\hat{\pi}_S-\pi_{FI}\rvert}$ [a]	1.53 (0.16)	0.52 (0.61)	0.22 (0.83)	0.87 (0.21)	0.77 (0.23)	0.37 (0.72)	1.31 (0.23)
$\hat{t}_{\lvert\hat{\pi}_{MLE}-\pi_{FI}\rvert}$ [b]	1.72 (0.12)	0.65 (0.54)	0.67 (0.52)	0.31 (0.76)	0.35 (0.73)	2.21 (0.06)	1.98 (0.08)
$\hat{t}_{\lvert\hat{\pi}_{MAP}-\pi_{FI}\rvert}$ [c]	2.61 (0.04)	0.13 (0.90)	0.73 (0.49)	0.79 (0.45)	1.05 (0.33)	1.84 (0.10)	2.21 (0.06)
$\hat{t}_{\lvert\hat{\pi}_{MAP}-\hat{\pi}_{MLE}\rvert}$ [d]	− 1.48 (0.18)	0.46 (0.65)	1.33 (0.11)	0.87 (0.41)	1.33 (0.22)	1.31 (0.23)	1.13 (0.15)
$\hat{t}_{(\hat{\pi}_S-\pi_{FI})^2}$ [e]	0.11 (0.91)	0.08 (0.94)	0.26 (0.80)	0.19 (0.85)	0.15 (0.88)	0.17 (0.87)	0.16 (0.87)
$\hat{t}_{(\hat{\pi}_{MLE}-\pi_{FI})^2}$ [f]	0.10 (0.93)	0.15 (0.88)	0.30 (0.77)	0.16 (0.88)	0.15 (0.88)	0.09 (0.93)	0.19 (0.86)
$\hat{t}_{(\hat{\pi}_{MAP}-\pi_{FI})^2}$ [g]	0.18 (0.87)	0.24 (0.82)	0.46 (0.66)	0.33 (0.75)	0.47 (0.65)	0.20 (0.85)	0.22 (0.83)
$\hat{t}_{(\hat{\pi}_{MAP}-\hat{\pi}_{MLE})^2}$ [h]	0.29 (0.78)	0.15 (0.88)	0.53 (0.61)	0.37 (0.72)	0.48 (0.64)	0.25 (0.81)	0.33 (0.75)

Number following a ' ± ' is the standard deviation of individual study area results.

[a] Student's *t*-test statistics for hypothesis $|\hat{\pi}_S-\pi_{FI}|=0$ probability of *t*-test statistics is in brackets.

[b] As footnote a for the hypothesis $|\hat{\pi}_{MLE}-\pi_{FI}|=0$.

[c] As footnote a for the hypothesis $|\hat{\pi}_{MAP}-\pi_{FI}|=0$.

[d] As footnote a for the hypothesis $|\hat{\pi}_{MAP}-\hat{\pi}_{MLE}|=0$.

[e] Student's t-test statistics for hypothesis $|\hat{\pi}_s-\pi_{FI}|^2=0$ probability of t-test statistics is in brackets.

[f] As footnote e for the hypothesis $|\hat{\pi}_{MLE}-\pi_{FI}|^2=0$.

[g] As footnote e for the hypothesis $|\hat{\pi}_{MAP}-\pi_{FI}|^2=0$.

[h] As footnote e for the hypothesis $|\hat{\pi}_{MAP}-\hat{\pi}_{MLE}|^2=0$.

mark while bias in MAP estimates of relative crop extent approached the 0.05 level of significance due to a single large deviation of 0.04 in study area number 5 (row 2 in Fig. 1). The significant MAP bias for exposed land was no surprise. String-like structures (e.g. roads) with a distinct direction of maximum spatial continuity are generally unsuited for the type of isotropic spatial prediction embodied in our SIS model.

MAP estimates of relative cover type extent were, across the nine study areas, strongly and positively correlated with their corresponding MLEs and FI benchmarks ($\rho \gtrsim 0.92$). The reproducibility of MAP estimates, as measured by the concordance correlation coefficient, was only slightly lower (avg. 0.02), indicating a consistent and high coefficient of accuracy of MAP predictions (Lin et al., 2002). The length of the 95% confidence interval for the absolute mean difference between a MAP estimate and a MLE was always below 0.02. MAP estimates were, overall, slightly (approximately 15%) more variable than corresponding MLEs.

Area-specific MAP prediction errors of relative cover type class extent were, on average, slightly (5%) below their bias-corrected MLE counterparts. Although the two sets of predictions errors were positively correlated (Fig. 4) the correlation (0.68) was weak. Statistical tests indicated that the relationship could be adequately described by a straight line through origo and a slope of 1 without significantly ($P=0.89$) increasing the sum of squared residuals of a linear regression fit with an intercept of 0.0004 ($P=0.91$) and a slope of 0.936. Adding a quadratic term to the linear regression would not significantly improve the fit ($P=0.77$).

A MAP estimate of the extent of a cover type class within a polygon of interest was superior to a localized MLE in terms of absolute deviations from the FI benchmarks (Table 4). Mean absolute deviation of MAP estimates from the FI values was 20% versus 33% for localized MLEs ($P=0.03$) and 25% for estimates derived directly from reference sample points inside the polygon (SIP). Relative root mean square errors followed the trends for absolute deviations and were, on average, 18% for MAP, 31% for localized MLEs, and 22% for SIP. Although the MAP root mean square error of prediction was the lowest in six of nine polygons, a signed rank test comparison failed to identify any significant method-dependent differences ($P>0.07$). Average length of a 95% confidence interval of a MAP estimate was only about one-third as long as that of a localized MLE or a SIP estimate.

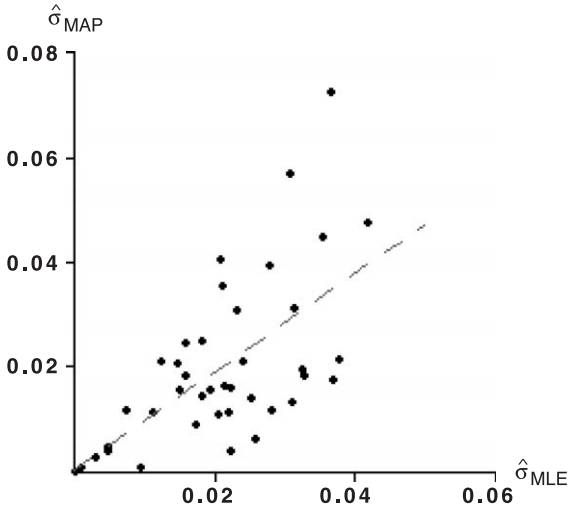

Fig. 4. Scatterplot of MAP (vertical axis) and MLE (horizontal axis) study area specific errors of predicted relative extent of land cover type classes. An ordinary least-squares regression line with a slope of 0.94 is indicated by the dashed gray line.

Table 4
Estimates of cover-type extent (Π) in polygons of interest (5% reference sampling)

Polygon area[a] (ha)	Cover type	Π_{FI} (ha)	Π_{SIP} (ha)	Π_{MLE} (ha)	Π_{MAP} (ha)
19	cp	6.7	5.3 ± 2.0	2.7 ± 4.0	4.7 ± 1.9
16	tm	4.5	6.5 ± 2.5	1.6 ± 2.9	2.5 ± 2.0
29	tm	9.7	9.0 ± 2.0	10.4 ± 0.7	6.4 ± 3.3
35	tc	11.4	10.1 ± 2.4	13.3 ± 1.8	11.2 ± 0.2
42	tm	13.1	12.6 ± 2.2	9.3 ± 3.8	9.8 ± 3.3
105	tb	36.2	35.5 ± 3.6	43.2 ± 7.0	32.5 ± 3.7
20	sh	3.6	0.0 ± 3.6	2.8 ± 0.9	3.6 ± 0.2
71	tc	18.5	16.1 ± 3.5	10.5 ± 8.0	15.6 ± 2.9
23	el	6.2	8.3 ± 2.7	1.5 ± 4.7	4.1 ± 2.1

Estimated root mean square prediction error follows the ' ± ' sign.
[a] One polygon per study area.

Table 6
Estimates of cover-type areas (Π) in polygons of interest (2% reference sampling)

Polygon area (ha)	Class	Π_{FI} (ha)	Π_{SIP} (ha)	Π_{MLE} (ha)	Π_{MAP} (ha)
19	cp	6.7	0.0 ± 6.7	2.7 ± 3.9	4.1 ± 2.5
16	tm	4.5	0.0 ± 4.5	1.7 ± 2.8	1.8 ± 2.7
29	tm	9.7	15.7 ± 6.3	12.9 ± 3.3	9.4 ± 0.3
35	tc	11.4	11.4 ± 2.0	11.4 ± 0.2	10.5 ± 1.0
42	tm	13.1	14.6 ± 2.7	10.1 ± 3.1	9.8 ± 3.4
105	tb	36.2	21.3 ± 15.2	52.2 ± 16.0	31.2 ± 4.9
20	sh	3.6	0.0 ± 3.6	2.6 ± 1.0	2.5 ± 1.1
71	tc	18.5	10.4 ± 8.4	11.7 ± 6.8	18.5 ± 0.2
23	el	6.2	8.1 ± 2.5	1.8 ± 6.6	4.0 ± 3.2

3.2. Less intensive reference sampling

With a 2% reference sampling, the Kappa coefficient of accuracy of the MAP cover decreased (mean = 0.50 ± 0.04) and was only slightly better (mean difference = 0.08) than that of a ML classified map. Trends reported for the 5% reference sampling were, otherwise, largely confirmed, albeit with a general increase in mean absolute deviations and root mean square errors. MAP estimates of relative cover type extent of exposed land and shrubs deviated significantly ($P \leq 0.05$) from corresponding MLEs (Table 5). Although test statistics of squared differences were higher than at 5% sampling, none reached the 0.05 level of significance (Table 5). The various correlation coefficients between estimates derived from MAP, MLE, and FI were, apart from a systematic drop of about 0.02, almost the same as reported for the 5% sampling.

MLE and SIP estimation of the extent of a cover type within a polygon of interest appears problematic with the 2% reference sampling (Table 6). Their root mean square error of prediction was, on average, in excess of 40%. Three of the nine polygons did not contain any reference sample points. MAP predictions, in contrast, were locally more accurate, as expected a priori of kriging-based predictions.

4. Discussion and conclusions

Integration of cover type reference sample data into a classified cover type map is needed to achieve compatibility between the map and MLEs of cover type extent (Card, 1982; Czaplewski, 1999; McRoberts et al., 2002; Tenenbein, 1972). The integration was accomplished in a MAP cover type map by the combined use of 'hard' and 'soft' data in collocated indicator cokriging predictions of cover type (Goovaerts, 1997) which also ensures unbiasedness and minimum variance of predictions. The employed SIS procedure merely organized the production of pixel-specific predictions into a series of replicate stochastic realizations of the cover type map. MAP cover type estimates of extent and precision obtained for the nine study areas were generally comparable to corresponding MLEs, a result we interpret as an achievement of the first part of the stated study objective. Apart from producing estimates comparable to MLEs the MAP cover type map also enjoys an increase in accuracy vis a vis the accuracy of a cover type map classified prior to the sampling of reference data. For the types of forested landscapes studied these benefits of a MAP-based approach to

Table 5
Summary statistics of relative cover-type extent π (2% reference sampling)

	el	sh	tb	tc	tm	wa	cp
π_{FI}	0.03 ± 0.03	0.05 ± 0.07	0.25 ± 0.16	0.23 ± 0.17	0.18 ± 0.09	0.14 ± 0.19	0.12 ± 0.14
$\hat{\pi}_S$	0.03 ± 0.03	0.05 ± 0.07	0.24 ± 0.17	0.23 ± 0.18	0.19 ± 0.10	0.14 ± 0.19	0.12 ± 0.14
$\hat{\pi}_{MLE}$	0.03 ± 0.03	0.05 ± 0.06	0.23 ± 0.17	0.23 ± 0.17	0.18 ± 0.10	0.13 ± 0.18	0.14 ± 0.17
$\hat{\pi}_{MAP}$	0.02 ± 0.02	0.04 ± 0.05	0.24 ± 0.15	0.24 ± 0.19	0.19 ± 0.11	0.13 ± 0.18	0.14 ± 0.15
$\hat{t}_{\lvert\hat{\pi}_S - \pi_{FI}\rvert}$	1.20 (0.28)	0.04 (0.98)	0.54 (0.61)	0.17 (0.87)	1.18 (0.28)	0.14 (0.89)	0.00 (1.00)
$\hat{t}_{\lvert\hat{\pi}_{MLE} - \pi_{FI}\rvert}$	0.29 (0.78)	0.03 (0.98)	0.83 (0.43)	0.13 (0.90)	0.27 (0.80)	2.23 (0.06)	1.34 (0.22)
$\hat{t}_{\lvert\hat{\pi}_{MAP} - \pi_{FI}\rvert}$	3.35 (0.01)	1.66 (0.14)	0.29 (0.78)	1.18 (0.27)	0.70 (0.51)	1.83 (0.11)	1.20 (0.27)
$\hat{t}_{\lvert\hat{\pi}_{MAP} - \hat{\pi}_{MLE}\rvert}$	2.71 (0.03)	2.44 (0.04)	0.33 (0.75)	0.96 (0.37)	0.46 (0.66)	1.06 (0.32)	0.18 (0.86)
$\hat{t}_{(\hat{\pi}_S - \pi_{FI})^2}$	0.17 (0.87)	0.15 (0.89)	0.53 (0.61)	0.31 (0.77)	0.35 (0.73)	0.18 (0.86)	0.28 (0.78)
$\hat{t}_{(\hat{\pi}_{MLE} - \pi_{FI})^2}$	0.18 (0.86)	0.19 (0.85)	0.55 (0.60)	0.29 (0.78)	0.39 (0.71)	0.12 (0.91)	0.39 (0.71)
$\hat{t}_{(\hat{\pi}_{MAP} - \pi_{FI})^2}$	0.25 (0.81)	0.23 (0.82)	0.60 (0.56)	0.36 (0.73)	0.47 (0.65)	0.21 (0.84)	0.54 (0.60)
$\hat{t}_{(\hat{\pi}_{MAP} - \hat{\pi}_{MLE})^2}$	0.22 (0.83)	0.64 (0.54)	0.64 (0.54)	0.42 (0.69)	0.48 (0.64)	0.26 (0.80)	0.55 (0.60)

See Table 3 for symbol definitions.

cover type mapping appear to require a fairly intensive (>3%) sampling of reference data before they emerge as practically important. The expected large number of pixels that are surrounded by two or more different FI cover types (approximately 14%) poses an upper limit to gain in accuracy that can be achieved by collocated indicator cokriging.

The second part of the study objective, the provision of global and local estimates of precision was achieved through the SIS generation of a binary matrix of prediction errors for every pixel. Only standard matrix operations are needed to convert these matrices into estimates of precision of global and local MAP predictions. In small area applications, MAP estimates appear preferable to more erratic localized MLEs and to estimates obtained directly from a small number of reference sample points in the area (SIP). MAP estimates are, accordingly, less sensitive to the number of reference points in small areas (Flores & Martínez, 2000) than either localized MLEs or SIP estimates.

Adaptation of the SIS procedure to local circumstances will be necessary in most cases. We modified the prediction of the water class to avoid a priori known problems along edges of water bodies. Related problems inherent in the kriging paradigm (Goovaerts, 1997) are likely to appear in practical applications. Variations of the workaround used here should be easy to implement and can be detailed around various edge detection methods (Wallerman et al., 2002; Gonzales & Woods, 1992). Also, instead of using an omnidirectional variogram as a default for all cover type classes, local circumstances may dictate the adoption of anisotropic variogram models.

MAP via SIS is computationally intensive. Exploiting the Markov–Bayes theorem to decompose the complex stochastic process behind the prediction of a cover type map into a sequence of independent simpler predictions is straightforward in the details but does require powerful computers in practical applications. Potential alternative methods such as *p*-field probability simulation and block-kriging (Goovaerts, 1997; Matern, 1980) would, in our view, be as computationally demanding as the approach taken here. Efficient search algorithms and fast routines for obtaining the indicator predictions are essential to reduce computing time. A considerable reduction in the time required to estimate the spatial prediction error of a polygon or a map can be achieved by subsampling. In our experience, simple random subsampling of 20–30% of the prediction covariances will provide approximations of the prediction variances and covariances that are within 5% of the estimate obtained via exhaustive computations.

MAP cover type mapping via SIS is best suited to quantify risk and uncertainty associated with data derived from a single or several overlaid maps (De Bruin, 2001; Hess & Bay, 1997; Kyriakidis & Dungan, 2002; Stehman & Czaplewski, 1998). The value of a forest will depend, among other attributes, on its cover type composition. An estimate of the spatial distribution of cover type prediction errors enables a decision-maker to obtain a probability density distribution of the estimated value. Areas less than 100 km^2 in size and characterized by a mosaic of spatially contiguous cover type or land use polygons, each 10–1000 ha in size, are probably best suited for the demonstrated MAP cover type prediction approach.

Acknowledgements

Valuable comments and critque of an earlier version of this manuscript were kindly provided by Drs. F. Csillag, G.B.M. Heuvelink, and R. McRoberts.

Appendix A

Class and pixel specific estimates of $\boldsymbol{\lambda}_1$ and λ_2 in Eq. (1) were solutions to the following system of equations (Goovaerts, 1997, p. 313, Eq. 7.47):

$$\begin{bmatrix} \boldsymbol{\Gamma}_1 & \boldsymbol{\Gamma}_2' & 1 \\ \boldsymbol{\Gamma}_2 & \boldsymbol{\Gamma}_3 & 1 \\ \mathbf{1}' & 1 & 0 \end{bmatrix} \times \begin{bmatrix} \boldsymbol{\lambda}_1 \\ \lambda_2 \\ \nu \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Gamma}_4 \\ \boldsymbol{\Gamma}_5 \\ 1 \end{bmatrix} \tag{2}$$

where $\boldsymbol{\Gamma}_1$ is a 8×8 matrix of distance-dependent class *i* indicator covariances between the eight nearest neighbors of the pixel for which a prediction is to be made, $\boldsymbol{\Gamma}_2$ is an 8×1 column vector of class *i* distance-dependent estimates of the covariance between 'hard' and 'soft' data, $\boldsymbol{\Gamma}_3$ is the binomial variance of p_i, $\mathbf{1}$ is a 8×1 column vector of ones, ν is a Lagrangian multiplier for the sum to one constraints on the weights $\boldsymbol{\lambda}_1$ and λ_2. Finally, $\boldsymbol{\Gamma}_4$ is a 8×1 column vector of class-specific distance-dependent indicator covariances between the pixel and its eight nearest neighbors, and $\boldsymbol{\Gamma}_5$ is the class-specific indicator covariance between the true and the conditional class membership probability from the initial ML classification. For each of the seven cover type classes, Gaussian elimination techniques were used to solve the 10 equations with 10 unknowns in Eq. (2). Valid solutions were found for all pixels.

Covariance estimates needed for solving Eq. (2) were obtained from semivariogram models of indicator covariances as a function of distance and global estimates of the indicator covariance between the ML classifier and the FI cover type class. The 7000 training pixels forming a contiguous area were used to obtain nonlinear least square estimates of the semivariogram model parameters and the global covariance estimates.

The cover-type-specific indicator covariance depends on the spatial distribution and frequency of the cover type pixels in the spatial domain of interest (Cressie, 1991). At distance 0, the class indicator covariance is the binomial variance of the class frequency. The covariance between cover type class (*i*) indicators separated by a distance (Eu-

clidian) h was estimated from $\hat{cov}(h|i) = \hat{cov}(0|i) - \hat{\gamma}(h|i)$, where $\hat{cov}(0|i)$ is the estimated covariance at zero distance, and $\hat{\gamma}(h|i)$ is the estimated semivariogram for distance h. The variogram $2\hat{\gamma}(h|i)$ for a second-order stationary process is an estimate of $2(\hat{cov}(0|i) - \hat{cov}(h|i))$ (Cressie, 1991, p. 83, Eq. 2.5.2).

For an ergodic and isotropic process, $2\hat{\gamma}(h|i)$ is obtained as the mean squared difference between pairs of cover type i indicator variables separated by a distance h. The mean is taken over all possible pairs of cover type i training pixels separated by a distance h. For indicator variables $2\hat{\gamma}(h|i)$ is a measure of the likelihood that two locations separated by a distance h are of the same cover type (Goovaerts, 1997). Omnidirectional spherical variogram models (Cressie, 1991, p. 61, Eq. 2.3.8) for 30 m $< h <$ 2000 m were fitted to values of $2\hat{\gamma}(h|i)$ derived from the data used for training the ML classifier. The omnidirectional spherical variogram models provided an overall satisfactory fit (they explained >80% of the variation in the data). The three parameters of the spherical models ($c_0(i)$, $c_s(i)$, $a_s(i)$ in Eq. (3)) were estimated by nonlinear least squares techniques (Gallant, 1987).

$$\gamma(h \mid i) = \begin{cases} 0 \text{ if } h = 0 \\ g_0(i) + g_s(i) \times \left(\dfrac{3h}{2a_s(i)} - \dfrac{h^3}{2a_s(i)^3} \right), \text{if } 0 < h < a_s(i) \\ g_0(i) + g_s(i), \text{ if } h > a_s(i) \end{cases} \quad (3)$$

Indicator cross-variograms between the true cover type class and the class predicted by the ML classifier were derived in a similar fashion, except that paired indicator variables represent an observed and a predicted class, respectively.

Although visual inspection of the training area and the nine studied plots may suggest local anisotropy (particularly for linear elements such as roads), there was no consistent direction of maximum continuity. Hence, all variograms use Euclidian distances as the most consistent predictor of covariance.

Three alternative variogram models (Cressie, 1991) were compared before choosing the above variogram model but no practically important difference emerged.

References

Andersen, G. L. (1998). *Classification and estimation of forest and vegetation variables in optical high resolution satellites: A review of methodologies.* IR-98-085, Int. Inst. Appl. Syst. Analysis, Laxenburg, AU.

Atkinson, P. M., & Lewis, P. (2000). Geostatistical classification for remote sensing: An introduction. *Computers & Geosciences*, *26*, 361–371.

Card, D. H. (1982). Using known map category marginal frequencies to improve estimates of thematic map accuracy. *Photogrammetry and Remote Sensing*, *48*, 431–439.

Cochran, W. G. (1977). *Sampling techniques*. New York: Wiley.

Collett, D. (1991). *Modelling binary data*. London: Chapman & Hall.

Conese, C., & Maselli, F. (1992). Use of error matrices to improve area estimates with maximum likelihood classification procedures. *Remote Sensing of Environment*, *40*, 113–124.

Conover, W. J. (1980). *Practical nonparametric statistics*. New York: Wiley.

Cressie, N. A. C. (1991). *Statistics for spatial data*. New York: Wiley.

Czaplewski, R. L. (1994). *Variance approximations for assessment of classification accuracy.* Research Paper RM-316, USDA Forest Service.

Czaplewski, R. L. (1999). Accuracy assessments and areal estimates using two-phase stratified random sampling, cluster plots, and the multivariate composite estimator. In H. T. Mowrer, & R. G. Congalton (Eds.), *Quantifying spatial uncertainty in natural resources: Theory and applications for GIS and remote sensing* (pp. 79–100). Chelsea, MI: Ann Arbor Press.

De Bruin, S. (2001). Predicting the areal extent of land-cover types using classified imagery and geostatistics. *Remote Sensing of Environment*, *74*, 387–396.

Deutsch, C. V., & Journel, A. G. (1998). *Geostatistical software library and user's guide* (2nd ed.). New York: Oxford University Press.

Flores, L. A., & Martínez, L. I. (2000). Land cover estimation in small areas using ground survey and remote sensing. *Remote Sensing of Environment*, *74*, 240–248.

Foody, G. M. (1999). The continuum of classification fuzziness in thematic mapping. *Photogrammetry and Remote Sensing*, *65*, 443–451.

Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, *80*, 185–201.

Franco-Lopez, H., Ek, A. R., & Bauer, M. E. (2001). Estimation and mapping of forest stand density, volume, and cover type using the k-nearest neighbors method. *Remote Sensing of Environment*, *77*, 251–274.

Franklin, S. E., Gillespie, R. T., Titus, B. D., & Pike, D. B. (1994). Aerial and satellite sensor detection of *Kalmia angustifolia* at forest regeneration sites in central Newfoundland. *International Journal of Remote Sensing*, *15*, 2553–2557.

Friedl, M. A., & Broadley, C. E. (1997). Decision tree classification of land cover from remotely sensed data. *Remote Sensing of Environment*, *61*, 399–409.

Gallant, A. R. (1987). *Nonlinear statistical methods*. New York: Wiley.

Gaudard, M., Karson, M., Linder, E., & Sinha, D. (1999). Bayesian spatial prediction. *Environmental and Ecological Statistics*, *6*, 147–171.

Glasby, C. A., & Horgan, G. W. (1995). *Image analysis for the biological sciences*. Chichester: Wiley.

Gonzales, R. C., & Woods, R. E. (1992). *Digital image processing*. Reading, MS: Addison-Wesley.

Good, P. (1993). *Permutation tests*. New York: Springer.

Goovaerts, P. (1997). *Geostatistics for natural resources evaluation*. New York: Oxford University Press.

Hess, G. R., & Bay, J. M. (1997). Generating confidence intervals for composition-based landscape indexes. *Landscape Ecology*, *12*, 309–320.

Hudson, W. D. (1987). Correct formulation of the Kappa coefficient of agreement. *Photogrammetry and Remote Sensing*, *53*, 421–422.

Klein, R., & Press, S. J. (1992). Adaptive Bayesian classification of spatial data. *Journal of the American Statistical Association*, *87*, 844–851.

Köhl, M., Innes, J. L., & Kaufmann, E. (1994). Reliability of differing densities of sample grids used for the monitoring of forest condition in Europe. *Environmental Monitoring and Assessment*, *29*, 201–220.

Kyriakidis, P. C., & Dungan, J. L. (2002). A geostatistical approach for mapping thematic classification accuracy and evaluating the impact of inaccurate spatial data on ecological predictions. *Environmental and Ecological Statistics*, *8*, 311–330.

Lin, L., Hedayat, A. S., Sinha, B., & Yang, M. (2002). Statistical methods in assessing agreement: Models, issues and tools. *Journal of the American Statistical Association*, *97*, 257–270.

Lunetta, R. S., Congalton, R. G., Fenstermaker, L. K., Jensen, J. R., McGwire, K. C., & Tinney, L. R. (1991). Remote sensing and geographic information system data integration: Error sources and research issues. *Photogrammetry and Remote Sensing*, *57*, 677–687.

Magnussen, S., Boudewyn, P., Wulder, M., & Seemann, D. (2000). Pre-

dictions of forest inventory cover type proportions using Landsat TM. *Silva Fennica*, *34*, 351–370.

Masselli, F., Conese, C., & Petkov, L. (1994). Use of probability entropy for the estimation and graphical representation of the accuracy of maximum likelihood classifications. *Journal of Photogrammetry and Remote Sensing*, *49*, 13–20.

Matern, B. (1980). *Spatial variation*. Berlin: Springer.

McLachlan, G. J. (1991). *Discriminant analysis and statistical pattern analysis*. New York: Wiley.

McRoberts, R. E., Wendt, D. G., Nelson, M. D., & Hansen, M. H. (2002). Using a land cover classification based on satellite imagery to improve the precision of forest inventory area estimates. *Remote Sensing of Environment*, *81*, 36–44.

Moody, A., & Woodcock, C. E. (1996). Calibration-based models for correction of area estimates derived from coarse resolution land-cover data. *Remote Sensing of Environment*, *58*, 225–241.

Natural Resources Canada (1999). *A plot-based national forest inventory design for Canada*. Victoria, BC: Canadian Forest Service, Pacific Forestry Center.

Pratt, W. K. (1991). *Digital image processing*. New York: Wiley.

Rossi, R. E., Dungan, J. L., & Beck, L. R. (1994). Kriging in the shadows: Geostatistical interpolation for remote sensing. *Remote Sensing of Environment*, *49*, 32–40.

Särndal, C. -E., Swensson, B., & Wretman, J. (1992). *Model assisted survey sampling*. New York: Springer.

Searle, S. R. (1982). *Matrix algebra useful for statistics*. New York: Wiley.

Snedecor, G. W., & Cochran, W. G. (1971). *Statistical methods* (6th ed.). Iowa: Iowa St. University Press.

Soares, A. (1998). Sequential indicator simulation with correction for local probabilities. *Mathematical Geology*, *30*, 761–765.

Soares, A. (2001). Direct sequential simulation and cosimulation. *Mathematical Geology*, *33*, 911–926.

Solow, A. R. (1986). Mapping by simple indicator kriging. *Mathematical Geology*, *18*, 335–352.

Stehman, S. V. (1999). Basic probability sampling designs for thematic map accuracy assessment. *International Journal of Remote Sensing*, *20*, 2423–2441.

Stehman, S. V., & Czaplewski, R. L. (1998). Design and analysis for thematic map accuracy assessment: Fundamental principles. *Remote Sensing of Environment*, *64*, 331–344.

Tenenbein, A. (1972). A double sampling scheme for estimating from misclassified multinomial data with applications to sampling inspections. *Technometrics*, *14*, 187–202.

Wallerman, J., Joyce, S., Vencatasawmy, C. P., & Olsson, H. (2002). Prediction of forest stem volume using kriging adapted to detected edges. *Canadian Journal of Forest Research*, *32*, 509–518.