Selecting representative high resolution sample images for land cover studies. Part 1: Methodology

J. Cihlar[1], R. Latifovic[2], J. Chen[1], J. Beaubien[3], Z. Li[1]

1) Canada Centre for Remote Sensing, Ottawa, Ontario

2) Intermap Technologies, Inc.

3) Canadian Forest Service, Quebec City, Quebec

Corresponding author: Josef Cihlar, Canada Centre for Remote Sensing, 588 Booth Street, Ottawa, Ontario, K1A 0Y7; <josef.cihlar@ccrs.nrcan.gc.ca>

Abstract

This is the first of two papers which explore the combined use of coarse and fine resolution data in land cover studies. It describes the development and evaluation of an objective procedure to select representative sample of tiles of high resolution images that complements a coarse resolution coverage of an entire region of interest. The second paper explores the use of the procedure for an accurate estimation of cover type composition at the regional scale. The Purposive Selection Algorithm (PSA) assumes that a relationship exists between land cover compositions at the two spatial scales. It selects one tile at a time, seeking the sample which most closely resembles the composition of the coarse resolution map. Two selection criteria were used, fraction of cover types and contagion index. PSA was evaluated using two land cover maps for a 288kmx165km area in central Saskatchewan, Canada derived from Landsat Thematic Mapper images (30m pixels) and Advanced Very High Resolution Radiometer (AVHRR, 1000m pixels), each divided into 64 tiles. The performance of an intermediate sensor (480m pixels) was assessed by resampling the TM map. When using cover type composition alone, It was found that the procedure rapidly converges on a representative set of tiles with land cover composition very similar to the full coverage. The match between the domain and sample cover type fractions was very close, with errors less than 0.002% once about 1/5 to 1/3 of the tiles were selected and no discernible bias in the selected sample. Compared to the TM whole area coverage, samples selected with AVHRR classification were as representative as those obtained using the TM map. The performance of samples selected by a combination of cover composition and contagion index responded to the characteristics of individual tiles in terms of the selection criteria. A rigorous application of the algorithm with spatial heterogeneity measures such as the contagion index is computationally very demanding. It is concluded that PSA provides an efficient and effective tool to select a representative sample for land cover studies in which both large area coverage and local detail are desired.

Introduction

The interest in land cover analysis at regional to global scales has grown dramatically in the last decade, stimulated by global environmental change and by improved mapping tools. The International Geosphere - Biosphere Program (IGBP) identified a strong need for regional and global land cover information (Townshend et al., 1994) to serve a variety of IGBP projects. Earlier work in using NOAA AVHRR data for mapping land cover (e.g., Loveland et. al., 1991) led to the execution of a global 1 km data land cover mapping initiative formulated in response to the IGBP and other requirements (Eidenshink and Faundeen, 1994). Various regional studies have also been undertaken, as were methodological studies to improve land cover information extraction procedures over large areas (e.g., Defries and Townshend, 1994; Belward, 1996; Cihlar et al., 1996). As a result, quality land cover data sets over large terrestrial areas are emerging, and they will become reality with improved data sources such as provided by the MODIS instrument (Salomonson, 1988) and planned new activities, e.g., the Global Observation of Forest Cover project (Ahern et al., 1998).

Maps showing large areas as a virtual 'snapshot' in time are a fundamentally new type of earth science information, not available until the recent advent of the appropriate remote sensing technology and analytical know-how. Nevertheless, they do not provide all the land cover information needed for detailed analysis, primarily because of the limitation by the spatial resolution. Even with the planned sensors operating in the 200-300m range, the resolution will not be sufficient for analysis and process studies at the local (stand or patch) scale. Experience from the Boreal Ecosystem-Atmosphere Study (BOREAS; Sellers et al., 1995), the GAP Analysis Project (Jennings, 1995) and similar investigations makes it clear that a resolution of 10-30m (such as provided by the Landsat Thematic Mapper, TM and the SPOT High Resolution Visible, HRV) is optimum for studies of ecological and other landscape processes. However, at this stage it is not feasible to implement a sustained continental or global mapping program which would provide consistent, time-specific (e.g., within one year) land cover data sets over large areas and at such high resolution. The limitations are of a practical nature, particularly the absence of suitable automated information extraction technology and financial resources.

The premise of this study is that it should be feasible to employ high and low resolution data in an optimum fashion to characterize land cover at the large (regional or continental) scale through a judicious combination of coarse and fine resolution data. In this case, the coarse resolution data would cover the entire domain of interest, while only a sample would be provided by the fine resolution data. Such samples are useful for studies of land cover composition, in the design of terrestrial sampling networks, and for the planning of large-scale experiments. The principle of sampling for land cover analysis is well established (e.g., Belward, 1996; Walsh and Burk, 1993). A key question is the sample selection strategy. Random sampling is statistically appealing because of the applicability of the classical statistical procedures. However, it is not generally an efficient approach. For this reason, other sampling designs such as stratified random, systematic grid and others (Cochran, 1963) may be preferred for land cover analysis. However, the sampling problem is complicated by the nature of fine resolution satellite data. Since the data are acquired as orbits and later subdivided the sampling unit is an image with a fixed size (e.g., a 185x185 km for a Landsat scene), not a single pixel. For reasons of costs and efficiency of using the acquired data, it is much preferable to select scenes that will make the greatest contribution to the characterization of land cover over the entire domain and at fine resolution.

The purpose of this paper is to outline an objective method for selecting a sample of fine resolution images for land cover analysis using a coarse resolution coverage of the entire area of interest, and to test the performance of the method. A companion paper explores the use of this methodology to estimate land cover composition over a region in central Canada.

Selection methodology

The proposed method is intended to select subareas to be imaged at high resolution, using a coarse resolution map of an entire area. For example, the area may be mapped using AVHRR 1 km data, and the sample provided by Landsat images (full scenes or ¼ scenes).

For the global land surface or any part thereof (termed ‘domain’ hereafter) one can readily obtain (i) complete coverage of images with coarse resolution data (e.g., 1 km) and (ii) the data acquisition framework (called ’tiles’ below) describing the potential coverage with fine resolution data. For Landsat, the tiles are specified by the World Reference System (NASA,

1982) path and row lattice (or quarter scenes within the images which are the smallest Landsat data granules); other satellite fine resolution imaging sensors use similar reference systems. Given that land cover composition is available in map form at the domain level one can also compute land cover composition for each tile, using the coarse resolution data. It is then postulated that a representative sample of the domain is that ensemble of tiles that together provide the same domain-level information as when the entire domain is mapped. The specific meaning of 'information' depends on the goals of land cover analysis but could include total area of land cover by type, spatial distribution of individual cover types, relative spatial distribution of several cover types, and others. In developing the algorithm and its testing in the companion paper, we concentrated on composition by land cover type, and to a lesser extent on spatial distribution. However, in general the algorithm may need to be adjusted with respect to the objective of the sampling.

Two descriptors of land cover are used below to describe composition and distribution, respectively. For composition, we employed the Euclidean distance *ED* between the cover fractions at the domain and tile levels, respectively:

$$ED_{d,j} = \sqrt{\sum_{i=1}^{n} \left(f_{d,i} - f_{j,i}\right)^2}, \qquad [1]$$

where $f_{j,i}$ = fraction of the area covered by cover type i in tile j (dimensionless), d represents the domain, and n is the number of classes.

In addition to an overall land cover composition, the patchiness of land cover at various spatial scales may also be important (Johnson et al., 1999). To describe the spatial distribution we selected the Contagion Index *CI* proposed by O'Neill et. al. (1988), as modified by Li and Reynolds (1993):

$$CI_j = 2n\ln(n) + \sum_{k=1}^{n}\sum_{m=1}^{n} P_{k,m}\ln(P_{k,m}), \qquad [2]$$

where $P_{k,m}$ is the probability that a pixel of land cover type k is found adjacent to a pixel of type m, and n is the number of land cover types in tile j. CI thus quantifies the likelihood that two

adjacent randomly selected pixels in the map belong to cover types *k* and *m*, based on the fraction of these types within the map and their spatial distribution. The reason for selecting *CI* was to characterize the degree of local intermixing of cover types. This index has been widely used in other studies (Turner, 1990; Graham et al., 1991; Gustafson and Parker, 1994).

Given the two descriptors, a selection algorithm can be defined for application to a domain map. It consists of the following steps (see Figure 1):

A. Preparation

1. Divide the domain land cover map into the desired tiles, and identify the extent of each tile on the domain map.

2. Determine *f(d,i)* and *f(j,i)* for the domain *d* and all tiles *j*. Also compute *CI* for each tile and for the domain using Eq. [2]. Put the results into 'source list' which contains the candidate tiles.

3. Compute the Euclidean distance (Eq.[1]) between the composition of the domain and each tile.

B. First tile

4. Select the first tile as that with minimum Euclidean distance *ED(d,j)*.

C. Second and subsequent tiles

5. For each tile *j* not yet selected, compute $ED_{d,s}$ (i.e., the distance between the domain and the sample tiles selected so far) and $ED_{d,s+j}$ (which would result if *j* were added to the sample). The change in *ED* is then determined as:

$$C_j = | ED_{d,s} - ED_{d,s+j} |, \qquad [3]$$

where C_j is the change that would result from adding tile *j,* and *s+j* is a hypothetical sample that includes tiles already selected and tile *j*. The absolute value is used because the difference could become temporarily negative.

6. Compute the relative change for each not-yet-selected tile as:

$$RC_j = \frac{C_{\max} - C_j}{C_{\max}}, \qquad [4]$$

where C_{max} is the maximum C value among all the remaining tiles (including tile j).

7. Identify as 'candidate tiles' those for which $RC(j) \leq Thr$ where threshold $Thr \geq 0$ is a user-defined value to provide a window of opportunity for the contagion index in the selection. Note that if $Thr=0$, the selection is based on ED only.

8. Among the candidate tiles that meet the Thr criterion select the tile that has the closest CI to the domain CI.

9. Return to Step 5 unless all tiles are selected.

Thus, the algorithm seeks to select the minimum set of tiles which most effectively represent the domain; for brevity, it is referred to below as PSA (purposive selection algorithm).

The PSA algorithm produces a plot of selection step vs. ED based on which the sample can be selected. This is described in the following sections.

Data and analysis procedure

The general approach to evaluating PSA was to prepare a domain coverage with coarse and fine resolution data; to test PSA with coarse resolution data, and to evaluate the results with fine resolution data considered as 'the truth'. For this reason, three land cover maps of the domain were prepared: coarse resolution (AVHRR-derived); fine resolution (TM-derived); and medium resolution, obtained by generalizing the fine resolution map to simulate future satellite data types, specifically MODIS.

Input Data

A land cover classification of a part of the BOREAS Region (Sellers et al., 1995) was used to test the PSA methodology. The area includes various boreal forest cover types in the northern part, cropland and grassland cover in the south. It is contained within two Landsat Thematic

Mapper (TM) scenes (Table 1). The TM scenes were classified using the Enhancement-Classification Method (ECM; Beaubien et al., 1999). The essence of ECM is to optimally enhance the input data, compress these without losing significant land cover information (as judged by a knowledgeable interpreter), and label the clusters after nearest neighbour classification using ancillary information.

Prior to the classification, the two scenes were radiometrically normalized using the overlapping area and time-invariant targets as determined through visual interpretation. The resulting clusters were assigned to various classes of the classification legend (Table 2). A qualitative evaluation of the accuracy of the classification was made through a comparison with colour infrared, stereoscopic aerial photographs obtained in the summer of 1994 along transects over parts of the BOREAS Region.

The AVHRR data to be classified (Table 1) were processed for the entire 1995 growing season (refer to Cihlar et al. (1997a) for details of the AVHRR processing). The classification was performed for all of Canada (Cihlar and Beaubien, 1998). ECM was also employed in this case, using the same basic steps but the resulting clusters were labeled using Landsat transparencies or prints from various parts of Canada. A qualitative assessment of the accuracy of the classification was carried out by a comparison of the classified AVHRR image with approximately 100 Landsat scenes, the latter being interpreted visually in the process. As a cross-check on the AVHRR – derived map it may be noted that in comparison to the independently obtained TM map, the average absolute difference (*DAB*, Eq. [6]) for the domain was 0.03% and the relative difference (*DRE*, Eq. [7]) 1.6%.

Computation of f and CI

The TM-derived map (9600x5504, 30 m pixels) was divided into 64 tiles (Figure 7), each consisting of 1200x688pixels. Secondly, the TM classification was transformed into an equivalent coarser classification by assigning the most frequent cover type within a 16x16 pixels window to all 30m pixels in that window. From this map, another set of 64 tiles was created and is referred to as P-MODIS. While the 30 m pixel size was retained in P-MODIS for analytical purposes, each tile was in fact equivalent to 75x43, 480mx480m pixels. The third domain coverage was provided by AVHRR. The AVHRR-derived map (1000m pixels) was registered to

the TM map, and a 30m pixel AVHRR coverage was created by nearest-neighbour resampling. The tiles in all three data sets were co-registered.

For each tile and domain, values of *f*, *ED*, and *CI* were computed. For *CI* the FRAGSTATS implementation of Li and Reynolds (1993) formula was used (McGarigal and Marks, 1993). A weighted absolute difference (*WAD*) between the domain and the sample was computed for various samples *s* as:

$$WAD = \sum_{i=1}^{n} f_{d,i} \mid f_{s,i} - f_{d,i} \mid, \qquad [5]$$

$$f_{d,i} = \frac{NP_{d,i}}{NP_d},$$

$$f_{s,i} = \frac{NP_{s,i}}{NP_s},$$

where *NP* is the number of pixels, *n* is the number of classes, and *i* refers to individual classes.

In addition to various combinations of data sets and *Thr* values, an additional test was made, a Random selection in which the tiles were chosen at random (without replacement), using *Thr*=0.

The differences between the domain and selected sample were quantified using mean values for the absolute (*DAB*) and relative (*DRE*) difference between the domain and the sample:

$$DAB = \frac{1}{n} \sum_{i=1}^{n} \mid f_{d,i} - f_{s,i} \mid, \qquad [6]$$

$$DRE = \frac{1}{n} \sum_{i=1}^{n} \frac{\mid f_{d,i} - f_{s,i} \mid}{f_{d,i}}. \qquad [7]$$

Results and discussion

Figure 2 shows the effect of tile selection for three data sets (TM, P-MODIS, AVHRR) using *Thr*=0. Two measures express the difference between whole area and the selected tiles, *ED* (Eq. [1]) and *WAD* (Eq.[5]). In all cases, the difference diminishes rapidly at first and then very gradually until most tiles are selected. The difference became quite small after about 1/6 (for *WAD*) to 1/3 (*ED*) of the tiles were selected, depending on data set and measure. It was smallest for the TM set and largest for the AVHRR. This is likely because the small scale variability was retained in case of TM, thus improving the representation of class proportions at the tile level. For AVHRR, the local variability was reduced and the tile consisted of fewer pixels, thus requiring additional tiles to obtain a broader, more representative sample. The P-MODIS result was intermediate between the two. The rate of convergence of the sample to the domain would thus depend on the relationship between the landscape heterogeneity and pixel size, and secondly on the size of the tile relative to that of the domain. Figure 2 also shows that the trends of *ED* and *WAD* were similar, although *ED* had a wider range. This is because a large difference in a few classes will affect *ED* more than *WAD*. Thus, *ED* is a more appropriate criterion for the selection of tiles because it leads to a faster convergence of the sample to the domain. The advantage of *ED* is obvious in some cases, e.g. the third selection step (Figure 2) for which *ED* decreased strongly (signifying fast convergence to domain values) but *WAD* increased somewhat.

Figure 3 illustrates the effect of adding the contagion index as a selection criterion. The overall tendency is to approximate the Random selection result (top curve, Figure 3a). For both the TM (Figure 3a) and P-MODIS (3b) data sets, the difference between the selections with *Thr*=0 (i.e., no *CI* used) and *Thr*=0.3 was small and not systematic. At higher *Thr* values, *ED* initially followed the *Thr*=0 curve but then moves towards the Random selection case. The point at which it starts to deviate depends on *Thr*. At high *Thr* (1.0), the selection considers a broad range of tiles at every step. Thus, after the tiles which together can strongly contribute to the domain are selected, the selection becomes so broad as to be effectively random. The selection is then more strongly influenced by *CI*. For lower *Thr* (0.5), the selection is restricted to a narrower range of tiles, and therefore the point at which the influence of *CI* begins to dominate arrives later. This trend is the same for TM and for P-MODIS, and also holds for *WAD*. For the AVHRR (Figure 3c), *ED* values for intermediate threshold values (0.3, 0.5) were also between the *Thr*=0 and Random cases throughout most of the selection process. The trend shown in Figure 3a and 3b was present as well (initial *ED* decrease followed by an increase) but it never reached the

Random case. On the other hand, *Thr*=1.0 produced the same *ED* as the Random case. It suggests that if the spatial distribution of cover types is also important (as described by *CI*), a larger number of scenes will be required to represent a domain.

A further insight into the effect of *CI* on the selection can be obtained from Figure 4 which shows *CI* values for individual selected tiles. The *CI* values for the domain were 21.5 (TM) and 57.1 (AVHRR). There is a general trend to increasing *CI* values starting from the domain value and the selection based on *ED* only (*Thr*=0). In other words, the tiles with the highest diversity of cover types (and thereby lowest *CI*) were selected first, and subsequent scenes tended to be more homogenous. Second, for low *Thr* the *CI* values of adjacent tiles fluctuated substantially but this fluctuation was dampened as the *Thr* value increased. For higher *Thr*, the *CI* values of selected tiles increased almost monotonically once the heterogeneous tiles were used up. Thus, the selection was guided by *CI* once the initial heterogeneous scenes were exhausted. In effect, the *RI* values (Eq.[4]) would differ less between the remaining tiles, thus permitting a larger number of tiles to become candidates for selection (Step 7). The same trend was observed for TM (Figure 4a) and AVHRR (4b). The main difference was the earlier start of the monotonic increase for AVHRR, throughout the *Thr*=0.3 as opposed to a second half of *Thr*=0.5 selection (TM). The TM curves also show that the monotonic increase was stronger for higher *Thr* values. Figure 4 thus implies that the *ED* trends in Figure 3 result from the combined effect of cover type heterogeneity (in the candidate tiles) and the *Thr* value. When the former is low and the latter high, the *ED* curve for the selected tiles will approximate Random case and *CI* will increase monotonically for adjacent tiles. This trend is due in part to the way *CI* is used in the algorithm (Step 8), and is discussed further below.

Figure 5 shows the effect of increasing the sample size on the difference between the domain and selected tiles within the same data set, both on individual classes and the combined effect. A value of *Thr*=0 was used. For one scene, the relative difference between the actual and estimated area can fluctuate widely. In case of TM (Figure 5a), it varied between 0% (class 3) and 97% (class 17: true fraction 2.7%, estimated 0.1%). This was reflected in the average errors for all classes, both absolute (*DAB*, 1.5%, Eq. 6) and relative (*DRE*, 41.2%, Eq. 7).

Increasing the number of tiles dampened the fluctuation, leading to average *DAB* (*DRE*) errors of 0.3% (10%) for 10 tiles and 0.2% (10%) for 20 tiles (i.e., 31% of the area). Although further

reductions were obtained they were relatively small. For example, by increasing the number of tiles by 50% (to 30), the average absolute error decreased by only 0.1% and the relative error by 4.1%. It should be noted that the relative error was strongly influenced by two small classes, representing 0.7% and 0.3% of the area respectively. Without these classes, *DRE* was 7.6% (10 tiles), 5.9% (20), and 2.6% (30).

A similar trend was observed for the P-MODIS data (Figure 5b). From the high average values after one tile (*DAB*= 2.0%, *DRE*= 62.7%), the magnitude decreased to 0.2% absolute and 12.3% (9.2% without two small classes) relative after 20 tiles. The addition of the next 10 tiles reduced the errors by 0% and 5.1% (2.0% without the two classes), respectively. The actual AVHRR data behaved in a similar way, although the fluctuations were larger. After one tile, the *DAB* (*DRE*) was 3.1% (155.4%). These were reduced to 0.5% (22.9%) after 10 scenes and to 0.3% (15.2%) after 20. As in the case of P-MODIS, the addition of further scenes decreased appreciably only *DRE* (to 8.8%) while absolute error changed by 0.1%. This is because of three small classes (0.6 to 0.9% of the area); without these, the DRE values were 10.2% (with 10 tiles), 8.0% (20), and 4.8% (30). The trends observed in Figure 5 suggest that the tile selection scheme is less efficient if small classes are present and must also be well represented, unless they are spatially associated with larger, more ubiquitous classes.

Since the TM map completely covered the area of interest it is possible to accurately evaluate the extent to which a sample of the AVHRR tiles would represent the area if it were imaged at high resolution, i.e. the feasibility of using coarse resolution coverage to select a high resolution sample. Figure 6 shows *DRE* values for the three data types and different reference data (*Thr*=0): domain coverage by the same data type (Figure 6a); domain coverage by TM (Figure 6b); and domain coverage by TM but ignoring the smallest class (representing 0.3% of the area) in *DRE* computation. Several observations can be made. First, the relative error decreased first rapidly and then more gradually, as also noted in Figure 5. Only in one case (AVHRR, 10 tiles) did a partial increase (Figure 6b) occur. Second, the *DRE* values for coarser resolution tiles were smaller when compared to 30m pixels (Figure 6b) than to the same resolution (Figure 6a), by about 30% for 20 tiles and both AVHRR and P-MODIS. At the higher resolution, the individual classes appear to be represented more accurately because the small classes were not averaged out. Third, a comparison of Figure 6b and 6c shows that the small classes magnified the overall relative error. For P-MODIS, the relative error was reduced by 30% when leaving out the

smallest class; for AVHRR, the reduction was 20%. Fourth, when TM was used as a reference, the difference in the *DRE* values between tiles selected using AVHRR or TM was small: 14.5% (for 10 tiles), 0.9% (20), and 0.1% (30). Similarly, the difference in relative errors was also small for P-MODIS (0.5% for 10 tiles, 1.5% for 20, 0.1% for 30). This shows that full coverage by coarse or medium resolution data can be successfully used to select representative areas to derive statistics that are nearly as accurate as if the sample were selected from full coverage of high resolution data. Although there is an implied requirement for accurate coarse resolution map, it is more important that the cover classes be internally consistent, i.e. each coarse resolution class should be comprised of a reasonably stable fractions of individual classes (that are resolved at high resolution).

Although comparable results after 20 or so tiles can be obtained this does not mean that exactly the same tiles will be selected. Figure 7 shows the tiles selected from the three data sets. Using the TM data set as a reference, only 7 tiles (35%) were selected from the P-MODIS data set and 11 (55%) from the AVHRR data. On the other hand, the first 5 tiles were chosen in the same order for TM and P-MODIS while the entire selection sequence was different between the AVHRR and the TM. A full correspondence is not to be expected because small differences in *ED* may lead to different selection paths. Once a different tile is selected it affects the subsequent sequence because those tiles are selected to balance the ones already chosen. The similarity of the relative errors after a number of tiles were selected (Figure 6) means that various combinations of tiles can provide similar results.

How closely do the statistics for the selected tile approximate the domain, and are they unbiased? Figure 8 shows the mean difference per class between the domain (TM map) and the sample with sign considered, computed as in the Eq. [6] without the absolute values. For *Thr*=0, the curves rapidly converged to 0%, especially for TM and P-MODIS. The convergence was more gradual for AVHRR but even there the mean difference was only 0.0017% after 20 tiles were selected. The difference from 0 (and thus the bias) is therefore negligible, even without taking into consideration other sources of error (such as classification accuracy) in the domain data sets. For *Thr*=0.5, the curve behaved more erratically. This is because the importance of the contagion index increases at some stage of the selection process (refer also to Figure 3). Neither convergence nor zero bias can therefore be assured in this case.

Comments

The above results show the practical feasibility of selecting a set of tiles that are 'representative' of the entire area. In this paper, representativeness was considered in terms of the cover type composition, as expressed by the fractions occupied by each class and by the spatial distribution measured by the contagion index. Using cover fractions only, the tiles selected with 480m or 1000m resolution maps provided virtually the same statistics as tiles selected with 30m data once approximately 1/4 to 1/3 of the area was selected. The number of tiles included in the coverage would depend somewhat on the importance of small classes. If the representation of such classes is not critical, a smaller number of tiles can be used. After a certain point, more tiles makes only a small contribution to the cover type composition, as evidenced by the increase from 20 to 30 tiles (31 to 47% of the area) which decreased the residual errors only marginally. Figure 2 indicates that the difference between the sample and domain statistics was appreciably reduced further only when almost all tiles were selected.

When using *CI* as a selection criterion, the process becomes more complicated because of the combinations of land cover distribution within individual tiles. The algorithm performed as expected, i.e. it selected, as far as possible, scenes with *CI* similar to the domain. The problem is that *CI*s for most individual tiles will be higher than those for the domain because of the reduced complexity of land cover distribution over a smaller area. The complexity increases as the number of selected tiles grows. Thus the way the *CI* was applied here is not optimal because only the *CI* of individual tiles was considered (Step 8). The preferred approach would be to compute a combined *CI* for all tiles already selected, and then compute the change in the sample *CI* if each individual tile were added – similarly as for *ED* in Step 5. Unfortunately, this presents a formidable computational demand, especially for areas of appreciable size. It would also require some simplification, i.e. the selected tiles would have to be assumed to constitute a contiguous map, thus incurring inaccuracies at the seams between tiles.

For the study area, the sample selected using the algorithm closely represents the region of interest and the residual errors are very small. In general, the difference between the domain and the sample is likely to depend on the heterogeneity of the land cover distribution in relation to the pixel size and on the size of the tiles. Considering typical data types, the tiles can vary between 60 km (SPOT images) and 185 km (TM full scenes), larger than those used here. Thus the convergence between domain and sample statistics is likely to be faster because individual tiles

will have a more balanced representation of cover types; this is especially true if full TM scenes are being selected. Nevertheless, from a statistical viewpoint the exact cover type composition of the domain should not be based on a simple average of the selected tiles. Simple averaging assumes random sampling so that the sample units can be considered independent. In our case, the selection is not random as the goal is to select the minimum representative subset. Three observations can be made. First, the procedure is designed to produce an unbiased sample (as compared to the domain of the same data type). Second, the sample will also be unbiased if a close relationship exists between the domain maps at coarse and fine resolutions. This is likely to be the case for larger areas, simple landscapes, large parcels of individual cover types, or a combination of these. Third, the 'worst case' situation will be infrequent cover types occurring in small patches. They should be represented well if their occurrence is correlated with the presence of more frequent cover types, e.g. cutovers co-occurring with contiguous dense forest stands. When small, infrequent cover types occur independently of others, their representation will depend on the spatial distribution within the tiles. In this study, the worst case (smallest class) occupied 0.3% of the area in the TM domain coverage. The *DRE* values for this class with 20 (30) tiles were 25.2% (15.1%) for TM, 21.7% (36.7%) for P-MODIS, and 0% (15.1%) for AVHRR. This shows that coarse resolution does not necessarily cause underrepresentation of small classes. It is also important to note that using random selection does not ensure representativeness unless independent information about the domain is available (which implies some form of coarse resolution mapping in the broad sense). Even when such information is available there is no assurance that a randomly or systematically selected sample contains the small classes. On the other hand, if the classes are mapped at both resolutions PSA ensures that the sample is selected in the most efficient way to represent the entire domain for a given sample size.

When analyzing land cover characteristics it is desirable to obtain measures of statistical confidence. In principle, these can be obtained if the probability of each tile being selected is known and statistical methods based on unequal probability sampling are applied (Cochran, 1963). Stuart (1976) demonstrated that in unequal probability sampling, the selection should be made as nearly proportional as possible to the values of the variable in the population. Various methods have been used to combine coarse and fine resolution data for land cover analysis (e.g., Walsh and Burk, 1993; Moody and Woodcock, 1996; Mayaux and Lambin, 1995, 1997; Cihlar et al., 1997b; Moody, 1998). These methods should yield more accurate results if the selected high

resolution sample closely represents the domain, and an objective and reproducible method of selecting such sample is strongly preferable to a subjective procedure. This topic is explored in the companion paper.

As described, PSA does not provide an objective cutoff for the sample size. This is a decision to be made by the analyst depending on the requirements of a particular study and the financial and other resources available. The changes in *ED* (Figure 2), *CI* (Figure 4), *DRE* (Figure 6), and mean difference (Figure 8) or *DAB* with increasing number of tiles provide the foundation for the tradeoff decisions.

Summary and conclusions

The increasing availability of coarse and fine resolution land cover maps presents a requirement for methods to optimally combine these different sources of information for studies of landscape characteristics at various spatial scales. In practice, it is readily feasible to produce land cover maps for large areas at coarse resolution and for smaller areas at high resolution. The question then arises, where should the high resolution sample be taken?

In this paper, we have developed and tested an algorithm for an objective sample selection, concentrating on the parts of the domain which provide the most information content at the high resolution. PSA seeks to find, at each iteration, the sample unit (tile) which would make the greatest contribution to closing the gap between the characteristics of the whole domain and those of the sample. We used two descriptors of land cover composition for the selection, fractional distribution of various cover types (measured by *ED*) and interspersion (quantified by Contagion Index, *CI*). When applied to an area mapped from both AVHRR (1000 m) and TM (30 m) images for a 47,500km^2 area in central Saskatchewan, Canada we found that:

1. Using *ED* alone, the difference between the domain and the sample diminished rapidly at first and later only slowly with an increasing sample size. Depending on the data set, the difference was small to negligible after 1/6 to 1/3 of the domain was selected.
2. Selection using data with 1000 m pixels (AVHRR) was nearly as efficient as with 30 m (TM) or 480 m (resampled TM). The average difference in cover type proportions between the domain (mapped at 30 m) and the sample of 20 tiles was 0.0002% (TM), -0.0017%

(AVHRR), and 0.0007% (P-MODIS). The differences were also small for lower numbers of selected tiles.

3. When *ED* and *CI* are used in combination, the selected sample represents a combination of the two attributes, and may not converge to the domain uniformly.
4. The final number of tiles to be included in the sample is a compromise decision which involves the residual differences between the domain and the sample at each selection step as well as practical (resource) considerations.

It is concluded that the PSA algorithm provides an efficient way to identify a sample of high resolution data for multi-resolution studies.

Acknowledgements
We wish to acknowledge the helpful comments of two anonymous reviewers.

References

Ahern, F. J., A. C. Janetos, and E. Langham, 1998. Global Observation of Forest Cover: a CEOS' Integrated Observing Strategy, Proceedings of 27th International Symposium on Remote Sensing of Environment, Tromsø, Norway, June 8-12: 103-105.

Beaubien, J., J. Cihlar, G. Simard, and R. Latifovic. 1999. Land cover from multiple Thematic Mapper scenes using a new enhancement-classification methodology. Journal of Geophysical Research (accepted).

Belward, A. (Ed.). 1996. The IGBP-DIS global 1 km land cover data set: proposal and implementation plan. IGBP-DIS Working Paper #13, The International Geosphere-Biosphere Programme, IGBP Data and Information System Office. 61p.

Cihlar, J., H. Ly, Z. Li, J. Chen, H. Pokrant, and F. Huang. 1997a. Multitemporal, multichannel AVHRR data sets for land biosphere studies: artifacts and corrections. Remote Sensing of Environment 60: 35-57.

Cihlar, J., and J. Beaubien. 1998. Land Cover of Canada 1995 Version 1.1. Digital data set documentation, Natural Resources Canada, Ottawa, Ontario.

Cihlar, J., H. Ly, and Q. Xiao. 1996. Land cover classification with AVHRR multichannel composites in northern environments. Remote Sensing of Environment 58: 36-51.

Cihlar, J., J. Beaubien, Q. Xiao, J. Chen, and Z. Li. 1997b. Land cover of the BOREAS Region from AVHRR and Landsat data. Canadian Journal of Remote Sensing 23: 163-175.

Cochran, W.G. 1963. Sampling techniques. John Wiley&Sons, Inc., New York. 413p.

DeFries, R.S., and J.R.G. Townshend. 1994. NDVI-derived land classifications at a global scale. International Journal of Remote Sensing 15: 3567-3586.

Eidenshink, J.C., and J.L. Faundeen. 1994. The 1 km AVHRR global land data set: first stages in implementation. International Journal of Remote Sensing 15 (17): 3443-3462.

Graham, R.L., C.T. Hunsaker, R.V. O'Neill, and B. Jackson. 1991. Ecological risk assessment at the regional scale. Ecological Applications 1: 196-206.

Gustafson,. E.J., and G. R Parker. 1994. Relationships between landcover proportion and indices of landscape spatial pattern. Landscape Ecology 7: 101-110.

Jennings, M.D. 1995. Gap analysis today: A confluence of biology, ecology, and geography for management of biological resources. Wildlife Society Bulletin 23:658-662.

Johnson, G.D., W.L. Myers, G.P. Patil, and C. Taillie. 1999. Multiresolution fragmentation profiles for assessing hierarchically structured landscape patterns. Ecological Modelling 116: 293-301.

Li, H., and J. F. Reynolds. 1993. A new contagion index to quantify spatial patterns of landscapes. Landscape Ecology 8: 155-162.

Loveland, T.R., J.W. Merchant, D.O. Ohlen, and J.F. Brown. 1991. Development of a land-cover characteristics database for the conterminous U.S. Photogrammetric Engineering and Remote Sensing 57: 1453-1463.

Mayaux, P., and E. Lambin. 1995. Estimation of tropical forest area from coarse spatial resolution data: a two-step correction function for proportional errors due to spatial aggregation. Remote Sensing of Environment 53: 1-16.

Mayaux, P., and E. Lambin. 1997. Tropical forest area measured from global land-cover classifications: inverse calibration models based on spatial textures. Remote Sensing of Environment 59: 29-43.

McGarigal, K., and B.J. Marks. 1994. FRAGSTATS. Spatial Analysis Program for Quantifying Landscape Structure. Version 2.0, Forest Science Department, Oregon State University, Corvallis, OR. 67p.

Moody, A. 1998. Using landscape spatial relationships to improve estimates of land-cover area from coarse resolution remote sensing. Remote Sensing of Environment 64: 202-220.

Moody, A., and C.E. Woodcock. Calibration-based models for correction of area estimates derived from coarse resolution land-cover data. Remote Sensing of Environment 58: 225-241.

NASA. 1982. LANDSAT-4 World Reference System (WRS) User Guide. National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, MD.

O'Neill, R.V., J.R. Krummel, R.H. Gardner, G. Sugihara, B. Jackson, D.L. DeAngelis, B.T. Milne, M.G. Turner, B. Zygmunt, S.W. Christensen, V.H. Dale, and R.L. Graham. 1988. Indices of landscape pattern. Landscape Ecology 1: 153-162.

Salomonson, V.V. 1988. The moderate resolution imaging spectrometer (MODIS). IEEE Geoscience and Remote Sensing Newsletter, August 1988: 11-15.

Sellers, P., F. Hall, H. Margolis, R. Kelly, D. Baldocchi, J. den Hartog, J. Cihlar, M. Ryan, B. Goodison, P. Crill, J. Ranson, D. Lettenmeier, and D. Wickland. 1995. The Boreal Ecosystem-Atmosphere Study (BOREAS): an overview and early results from the 1994 field year. Bulletin of the American Meteorological Society 76: 1549-1577.

Stuart, A. 1976. Basic ideas of scientific sampling. Griffin's Statistical Monographs and Courses #4, Hafner Press, New York. 106p.

Townshend, J.R.G., C.O. Justice, D. Skole, J.-P. Malingreau, J. Cihlar, P. Teillet, F. Sadowski and S. Rutenberg. 1994. The 1 km resolution global data set: needs of the International Geosphere Biosphere Programme. International Journal of Remote Sensing 15: 3417-3441.

Turner, M.G. 1990. Spatial and temporal analysis of landscape patterns. Landscape Ecology 4: 21-30.

Walsh, T.A., and T.E. Burk. 1993. Calibration of satellite classifications of land area. Remote Sensing of Environment 46: 281-290.

Table 1. Satellite data employed

Sensor	Location	Date	Bands
Landsat TM	36/22-23* 37/22-23*	1996/07/30 1991/08/09	3,4,5 3,4,5
NOAA AVHRR	Canada	1995 growing season	1,2,Nm**

- Path/Row; ** Mean value of the Normalized Difference Vegetation Index.

Table 2. Land cover types for the Landsat TM and AVHRR classifications

2.1 Thematic Mapper legend

Forest

Coniferous

1 High crown density (>60%)
2 High crown density younger (>60%)
3 Low crown density (25-40%)
4 Low crown density with lichens (25-40%)
5 Very low crown density often treed wetland (10-25%)
6 Very low crown density with lichens (10-25%)

Deciduous

7 High crown density (>50%)
8 Low crown density, mostly regeneration (25-50%)

Mixed

9 Coniferous >50%
10 Deciduous >50% (occasionally very open forest)

Open land (tree crown density open forest <10%)

11 Burns
12 Burn with more vegetation (also very open lichen conifers)
13 Wetland
14 Wetland or cropland
15 String bogs

Cropland

16 High vegetation cover
17 Medium vegetation cover
18 Low vegetation cover
19 Very low or without vegetation

Others

20 Water bodies
21 Clouds

2.2 Advanced Very High Resolution Radiometer (AVHRR) map legend

Forest land

Evergreen needleleaf forest

1 *High density*

Medium density

2 Southern forest

3 Northern forest

Low density

4 Southern forest

5 Northern forest

6 **Deciduous broadleaf forest**

Mixed forest

7 Mixed needleleaf forest

Mixed intermediate forest

8 Mixed intermediate uniform forest

9 Mixed intermediate heterogeneous forest

10 Mixed broadleaf forest

Burns

11 Low green vegetation cover

12 Green vegetation cover

Open land

13 Transition treed shrubland

14 Wetland/shrubland (medium density)

15 Grassland

Developed land

Cropland

16 High biomass

17 Medium biomass

Mosaic land

18 Cropland-Woodland

19 Cropland-Other

Other

20 Water

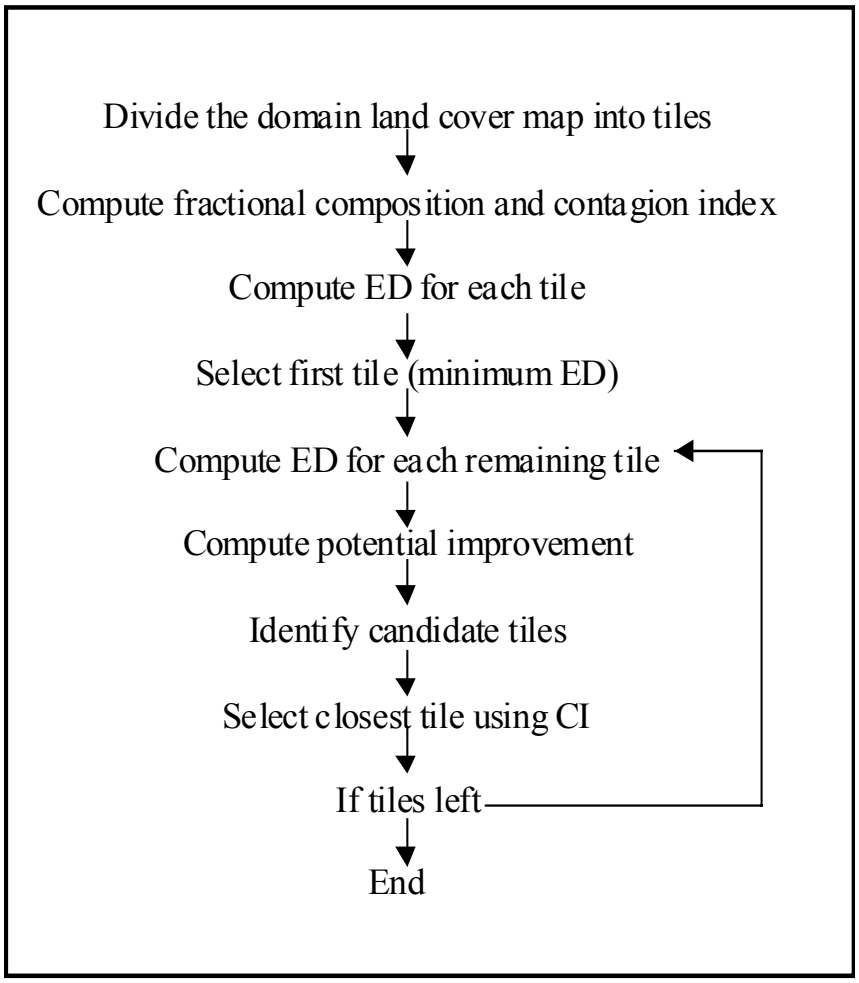

Figure 1. Flowchart of the tile selection algorithm

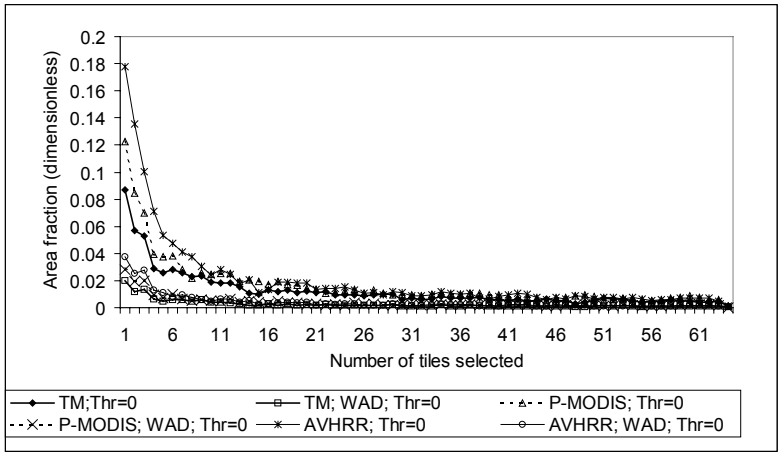

Figure 2. The mean Euclidean distance (*ED*) and the mean weighted absolute difference (*WAD*) between the domain and the tiles selected using threshold *Thr*=0.

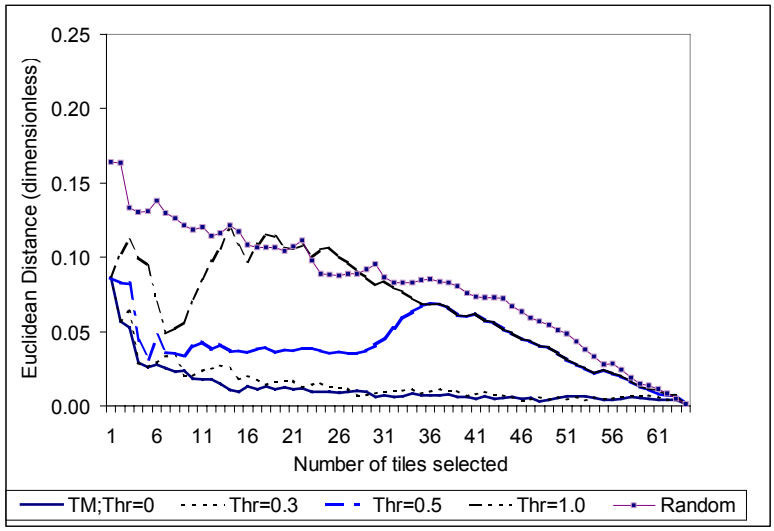

Figure 3a

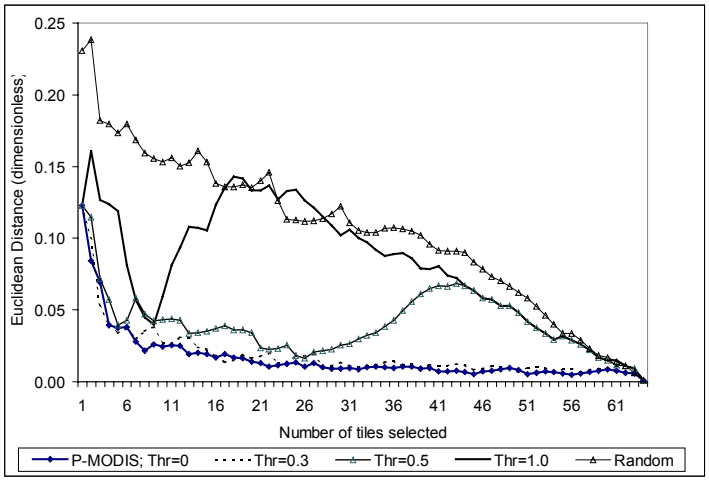

Figure 3b

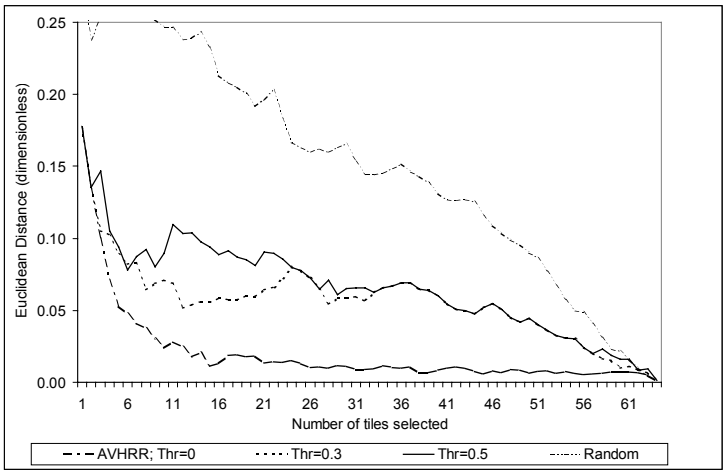

Figure 3c

Figure 3. The mean Euclidean distance (*ED*) between the domain and the tiles selected using various threshold values, and for a random tile selection: TM map (Figure 3a), P-MODIS map (3b), AVHRR map (3c).

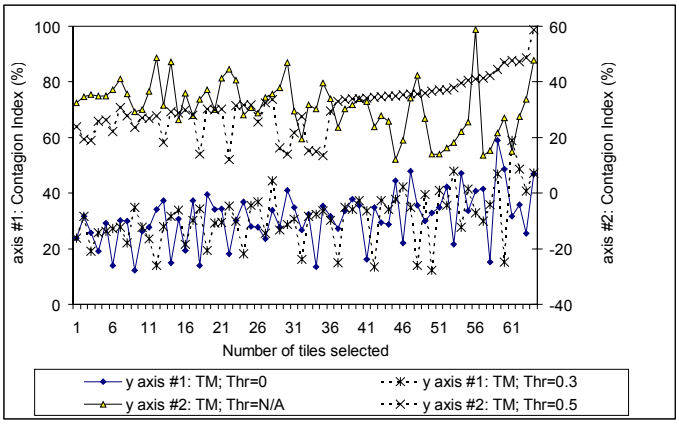

Figure 4a

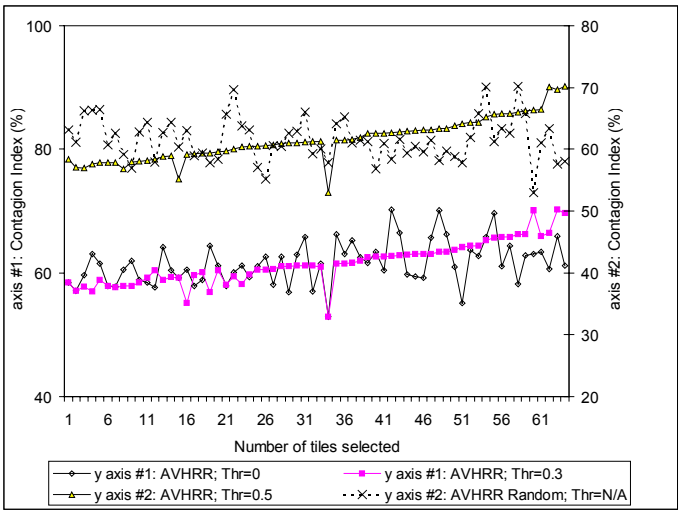

Figure 4b

Figure 4. The contagion index (CI) between the domain and the tile selected at each step for various threshold values: TM map (Figure 4a), AVHRR map (4b).

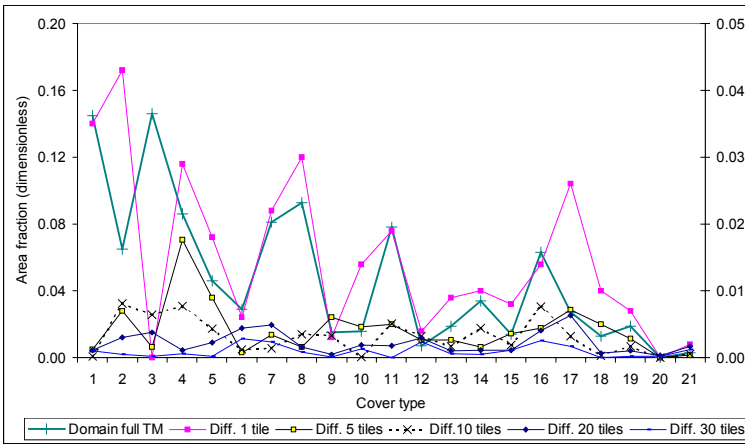

Figure 5a

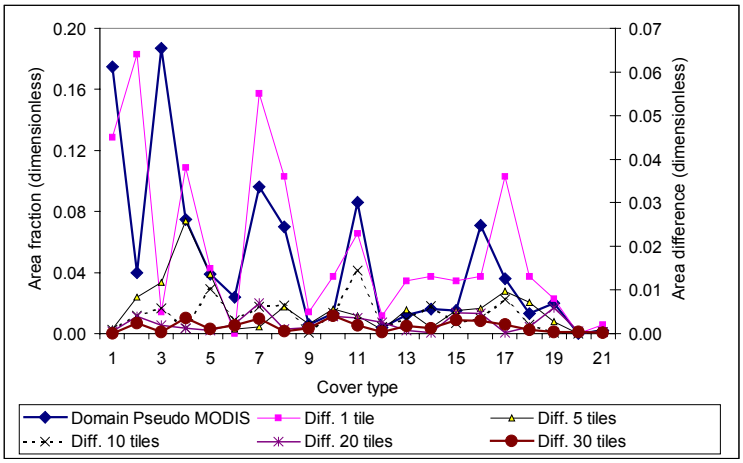

Figure 5b

Figure 5. The cover type area fraction and the absolute difference between the fraction of land cover type in the domain map and the sample (right-hand y axis), for individual cover types: TM (Figure 5a), P-MODIS (5b), AVHRR (5c). The difference is computed between the domain values for that date type and the sample.

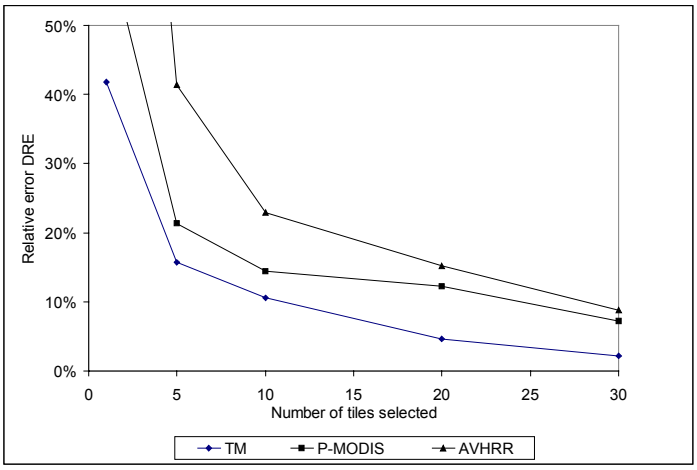

Figure 6a

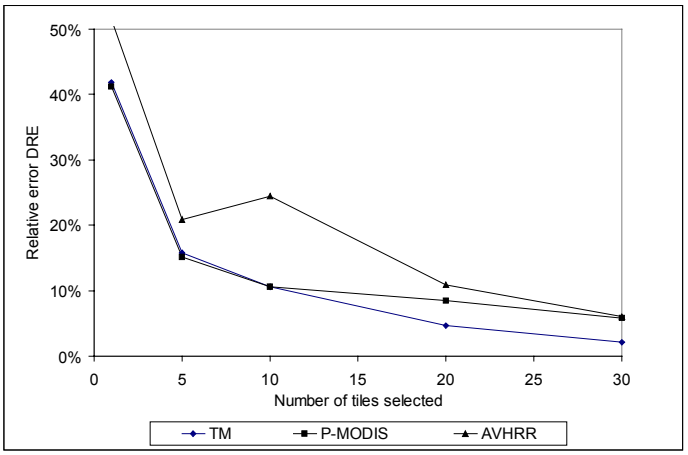

Figure 6b

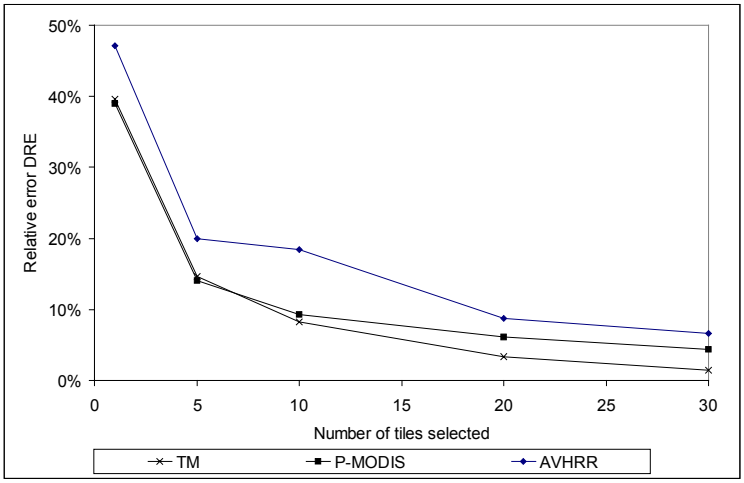

Figure 6c

Figure 6. The mean relative error (*DRE*) between the domain and the selected tiles with different reference data. Figure 6a: compared to the domain map of the same data type; Figure 6b: compared to the TM domain map; Figure 6c: compared to the TM domain map and without the smallest class (occupying 0.3% of the domain).

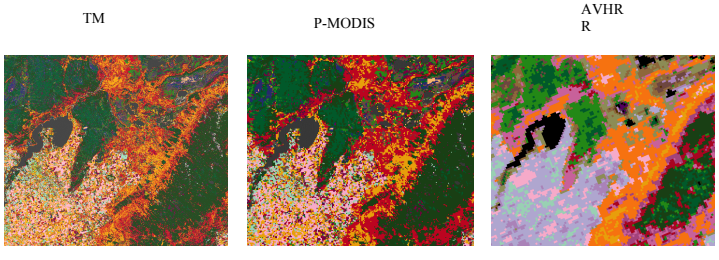

Figure 7

Thematic Mapper

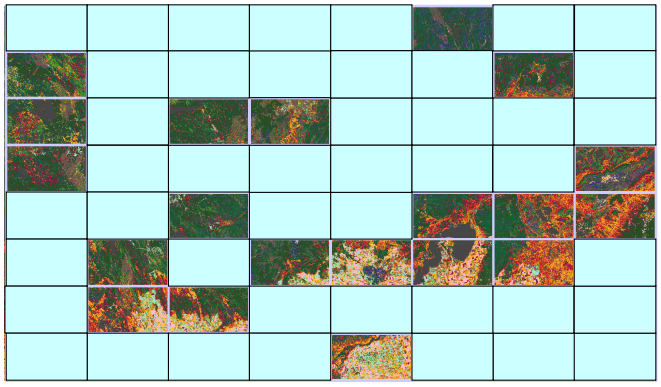

Figure 7a

Pseudo - MODIS

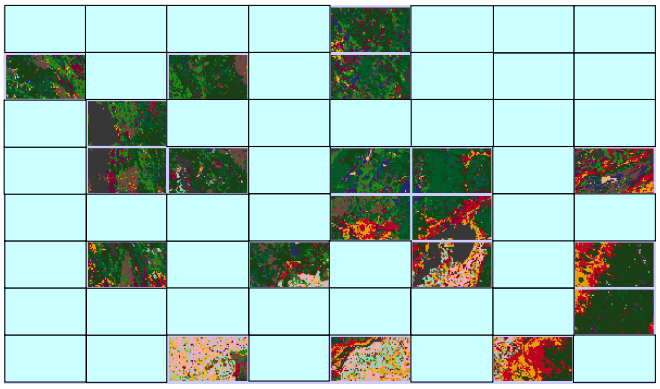

Figure 7b

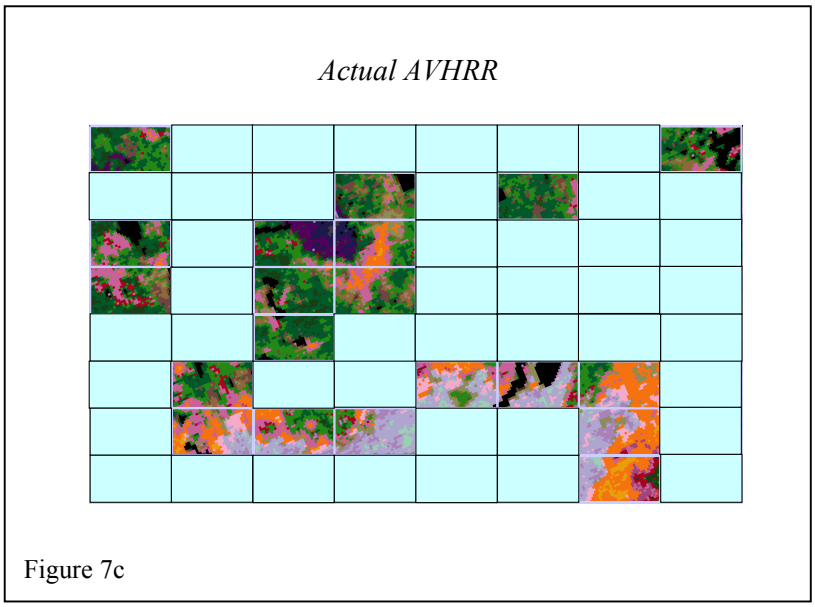

Figure 7. Tiles selected after 20 steps using TM map (Figure 7a), P-MODIS (7b), and AVHRR (7c).

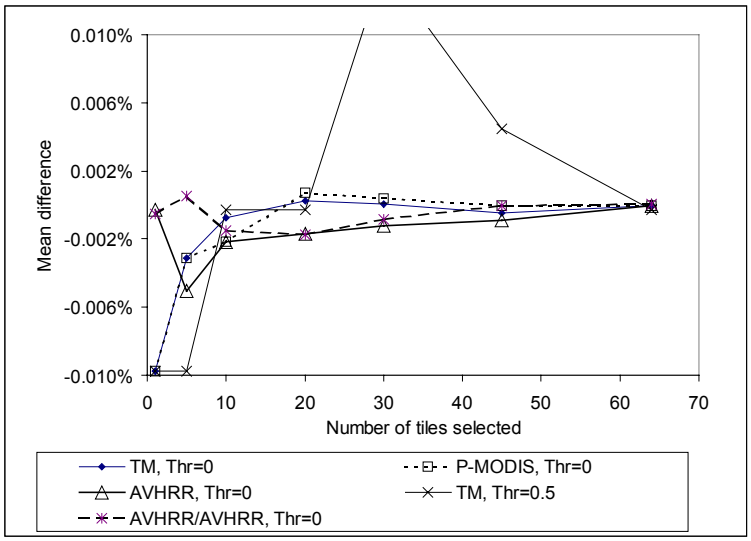

Figure 8. Mean difference (sign considered) between the fraction of the area occupied by a class in the domain and the sample. The reference used is the TM domain map, except for the AVHRR/AVHRR curve where the reference is the AVHRR domain map.