Environ Ecol Stat (2008) 15:447–467
DOI 10.1007/s10651-007-0061-9

Spatial structure effects on the detection of patches boundaries using local operators

Mathieu D. Philibert · Marie-Josée Fortin · Ferenc Csillag

Received: 1 January 2006 / Revised: 1 June 2006 / Published online: 13 December 2007

Abstract Landscapes exhibit various degrees of spatial heterogeneity according to the differential intensity and interactions among processes and disturbances that they are subjected to. The management of these spatially dynamical landscapes requires that we can accurately map them and monitor the evolution of their spatial arrangement through time. Such a mapping requires first the delineation of various spatial features present in the landscape such as patches and their boundaries. However, there are several environmental (spatial variability) as well as technical (spatial resolution) factors that impair our ability to accurately delineate patches and their boundaries as polygons. Here, we investigate how the spatial structure and spatial resolution of the data affect the accuracy of detecting patches and their boundaries over simulated landscapes and real data. Simulated landscapes consisted of two patches with parameterized spatial properties (patches' level of spatial autocorrelation, mean value and variance) separated by a boundary of known location. Real data allowed the investigation of a more complex landscape where there is a known transition between two forest domains with unknown spatial properties. Boundary locations are defined using the lattice-wombling edge detector at various aggregation levels and the degree of

F. Csillag—Deceased.

M. D. Philibert
Department of Geography, Simon Fraser University, Burnaby, BC, Canada V5A 1S6

M. D. Philibert (✉)
National Public Health Institute of Québec, 190 Crémazie East blvd, Montréal, QC, Canada H2P 1E2
e-mail: mathieu.philibert@inspq.qc.ca

M.-J. Fortin
Department of Zoology, University of Toronto, 24 Harbord St., Toronto, ON, Canada M5S 3G5
e-mail: mjfortin@zoo.utoronto.ca

F. Csillag
Department of Geography, University of Toronto at Mississauga, Mississauga, ON, Canada L5L 1C6

patch homogeneity is determined using Getis-Ord's G^*. Results show that boundary detection using a local edge detector is greatly affected by the spatial conditions of the data, namely variance, abruptness of the spatial gradient between two patches and patches' level of spatial autocorrelation. They also suggest that data aggregation is not a panacea for bringing out the ecological process creating the patches and that indicators derived from local measures of spatial association can be complementary tools for analysing spatial structures affecting boundary delineation.

Keywords Edge detection · Local statistics · Spatial autocorrelation · Stationarity

1 Introduction

Landscapes exhibit various degrees of spatial heterogeneity according to the differential intensity and interactions among natural and anthropogenic processes and disturbances that they are subjected to. The management of these spatially dynamical landscapes requires that we can accurately map them and monitor the evolution of their spatial arrangement through time. Maps can be produced by photo-interpreting aerial photographs, by classifying remotely sensed images or using field data. However, such mapping requires first the delineation of various spatial features forming the landscape such as patches and their boundaries. Ideally, natural patches should be represented using polygons (i.e., vector format) to facilitate their comparison with anthropogenic landscape features (e.g., cities, counties). In natural landscapes, however, there are several ecological (e.g., dispersal, disturbance, spatial competition) as well as technical factors (e.g., spatial resolution of the measurements, edge detection techniques) that impair our ability to accurately delineate patches and their boundaries as polygons (Fortin and Edwards 2001).

Ecological patches (e.g., forest stand) are usually determined assuming that for at least one variable (qualitative or quantitative) its values within the patch are homogeneous. The natural landscape features with sharp boundaries are mostly of anthropogenic origins (e.g., forest logging, especially traditional clear cutting and road construction) while natural boundaries vary usually over space creating less abrupt (i.e., gradual, fuzzy) boundaries (e.g., at a regional scale, from a deciduous to a coniferous forest). Furthermore, most of the ecological patches are spatially structured (Fig. 1) due to spatial dependence and spatial autocorrelation processes (Fortin and Dale 2005). Spatial dependence arises from interactions with other environmental processes (e.g., drainage, topography, climate) while spatial autocorrelation is an intrinsic characteristic of an ecological process (e.g., seeds dispersal, competition). In addition, patches can be highly variable (i.e., high variance, Fig. 1) and their strength can depend on the data themselves and their spatial resolution (Gosz 1993), making the detection of patches (i.e., difference in patches average; Fig. 1) difficult (Fortin and Edwards 2001).

In this study, we investigate the effect of spatial structure and spatial resolution on the accuracy of detecting patches and their boundaries using both simulated and real data. Specifically, we examine how "within-patch" spatial structures (i.e., spatial autocorrelation) and variability (i.e., mean and variance) affect the ability of a local edge detector, here the lattice-wombling (Fortin and Drapeau 1995), to accurately find

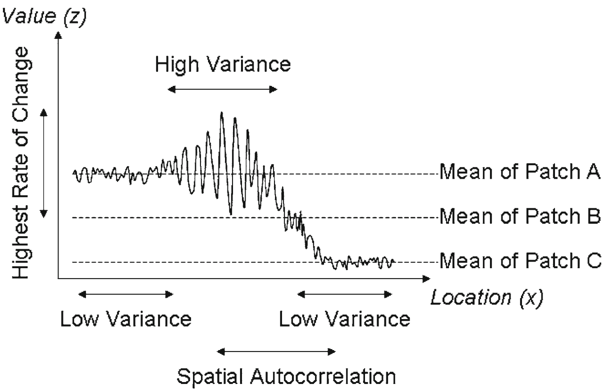

Fig. 1 Illustration of the effects of high values of variance and spatial autocorrelation in the ability to detect boundaries among three patches

the location and the width of ecological boundary. This edge detector was favoured among others (Pitas 2000) as it can detect boundary region/area, not just boundary line (Fortin and Drapeau 1995; Fortin and Dale 2005). The use of a local measure of spatial associations (local Getis statistics; Getis and Ord 1992) as an exploratory analysis for identifying patches is also evaluated.

2 Methods

2.1 Patch and boundary

Operationally, boundaries are defined as connected locations at which large changes occur between two patches based on a given attribute (i.e., species abundance, species type) where each patch shows a relative level of homogeneity for that given variable (Jacquez et al. 2000). In order to evaluate how spatial structures and variability, as well as scale of analysis, affect the accuracy of local operators, both simulated and real data were used: simulated landscapes allowed to control for (*parameterize*) spatial and statistical properties in two patches; real data allowed to investigate how these local operators are affective on a more complex landscape where there is a known transition between two forest domains, black spruce and balsam fir but where the spatial properties of each domain are unknown (*unparameterized* values). First we explain how the simulated data were simulated and described the real data. Then we review the local edge detector algorithm as well as the local spatial association statistic.

2.2 Simulated data

To investigate the effects of known degree of spatial autocorrelation and statistical values (mean and variance), simulations were produced using a conditional autoregressive model (CAR; Csillag et al. 2001; Fotheringham et al. 2000; Cressie 1993; Haining 1990). The conditional probability distribution of a value at a location is given by:

$$P[z(s_i)] :\sim \text{Gauss}(\rho^* \text{ave}(zN_i), \tau_i),$$

where $\text{ave}(zN_i)$ is the average of the neighbours of location s_i, ρ the spatial autocorrelation value attributed to a patch and τ the conditional variance assigned to the entire landscape.

This simulation model is an iterative process from which a given iteration is the result of the modification of the values at locations sampled from the previous iteration. Locations for which the values to be modified are chosen from a random sub-lattice and their values are assigned following given conditions. The conditions are to follow a Gaussian distribution with a patch-specific mean and a global variance (i.e., for the whole landscape) and to keep a given level of spatial autocorrelation within a patch.

From an original image of 50×50 cells showing two perfectly homogeneous patches separated by a crisp boundary of known location, location values inside each patch were modified following the three parameters defined above as follow: (a) 1, 2 and 10 for the conditional variance; (b) a difference of 0, 1, 2 and 10 between patches' average; and (c) 0.0, 0.95 and 0.9999 for the spatial autocorrelation (1 being the strongest level of spatial autocorrelation). Combining all the possible values of the three parameters leads to 72 sets of conditions. For each set of conditions, 20 replicates were generated as preliminary analyses on sub-sets of simulations indicated that 40 and 60 replicates yielded to similar results (not shown). For all these sets, the step of the random sub-lattice used for aggregating was set to three (see Sect. 2.4) and a new iteration was launched after aggregating a total of 100 locations. The simulation was stopped after 150 iterations after a visual assessment indicated stability in the spatial structure of the landscapes. Figure 2 illustrates the landscape spatial differences while using different values for the three parameters.

2.3 Real data

The real data originate from a digitised forest inventory (SIFORT digitised database) made in the boreal forest of northern Québec, Canada (Fig. 3). A portion of the Abitibi region, where the black spruce (*Picea mariana*) and the balsam fir (*Abies balsamea*) forests meet, was selected from the SIFORT database. This sector will therefore serve to test the ability of the methodology described below to find the boundary between the black spruce domain (north) and the balsam fir domain (south). Numerical values of the data represent black spruce percentage coverage and the resolution of the pixels is 14 ha (350 pixels in longitude and 425 pixels in latitude).

2.4 Local edge detector

Most commonly used edge detectors are local operators that are detecting boundaries among adjacent locations in terms of strength of gradient (i.e., quantitative difference; Pitas 2000). These detectors operate as moving kernels, usually defined by square

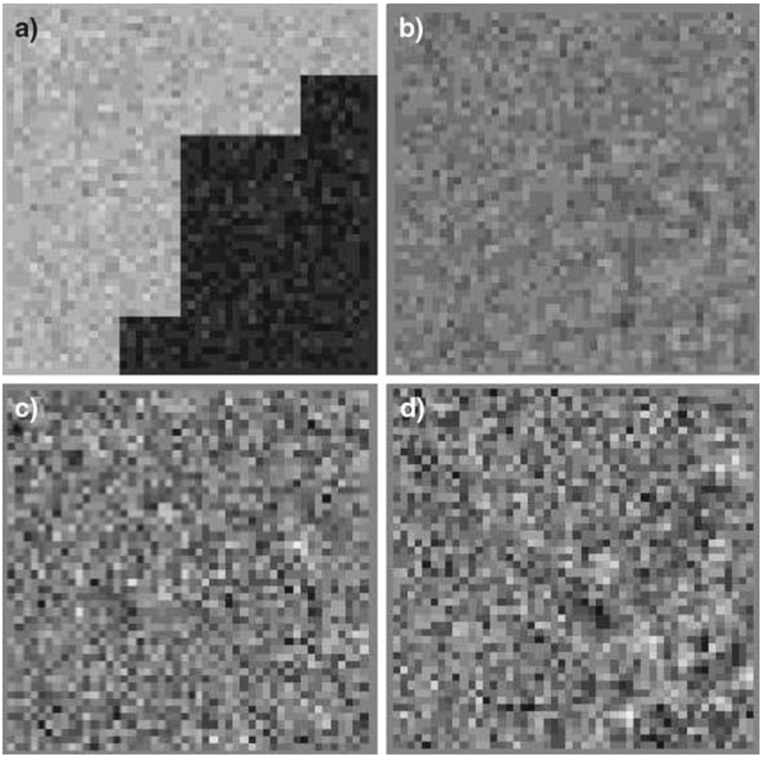

Fig. 2 Examples of simulated landscapes following various combinations of spatial and statistical conditions. (**a**) Is a landscape where there is no spatial autocorrelation within each patch, the global variance is 1 and the difference between patches' average is 10; (**b**) shows no spatial autocorrelation in the upper-left patch but a high level (0.9999 on a 0–1 scale) of spatial autocorrelation in the lower-right patch, a conditional variance of 1 and no difference between patches' average; (**c**) is a landscape where spatial autocorrelation in each patch and the difference between patches' average were set to 0 but the amount of global variance was at its highest (10); (**d**) the same conditions as (**c**) except for the amount of spatial autocorrelation in the lower-right patch, which was set to 0.9999

grids of $n \times n$ spatial units (e.g., 3×3 or 5×5, etc.), centered at x, y coordinates or at a given pixel. Such local operators are therefore sensitive to local heterogeneity found in the spatial data which will result in local edges that are not corresponding to boundary (Fig. 1).

Given that the purpose of this study is to investigate the effect of within-patch properties on the detection of boundaries, most of the existent local edge detectors could have been used. Here we used the lattice-wombling edge detector that computes first-order derivatives, m, over immediately adjacent four spatial units forming a square (Barbujani et al. 1989). The choice of the lattice-wombling is justified by the fact that it is an edge detector known to ecologists and that tests were developed to assess the significance of the identified boundaries (Fortin et al. 1996; Oden et al. 1993).

This magnitude of rate of change (hereafter simply referred as edges), m, is computed for the centroid of every 2×2 overlapping window and is position at the centroid

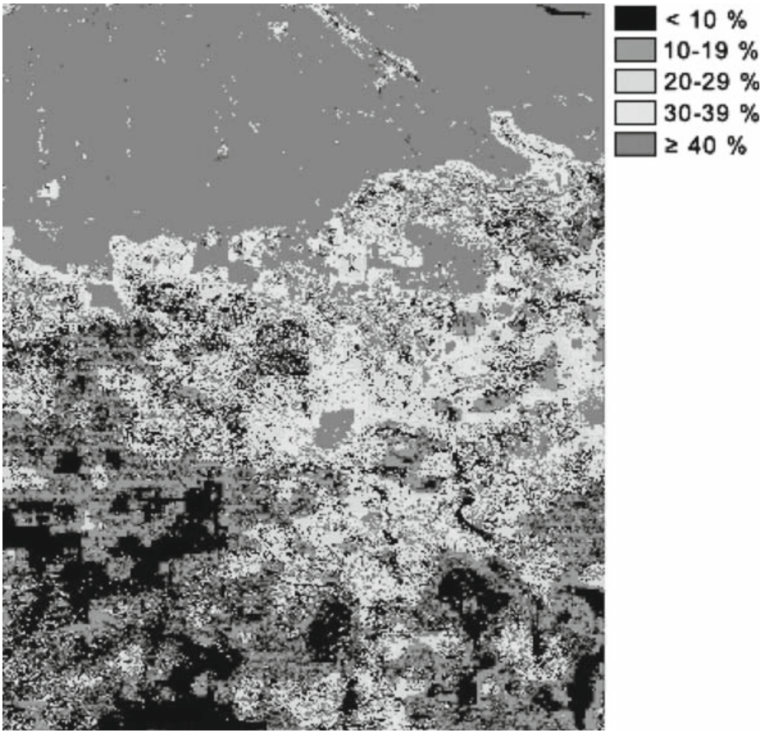

Fig. 3 Percentage coverage of black spruce (*Picea mariana*) in the boreal forest of northern Québec, Canada

of these four adjacent units creating the search window:

$$m = \sqrt{[\partial f(x, y)/\partial x]^2 + [\partial f(x, y)/\partial y]^2}$$

where $f(x, y)$ is a bilinear function of the values at the x and y coordinates of the units. The maximum number of rates of change computed is the number of rows minus one times the number of columns minus one.

To test where edges are "high enough" to identify the position of boundaries (hereafter referred as BLs for boundary locations) several approaches can be used: statistical, geographical and arbitrary. In the statistical approach, the significance of each edge is assessed (Oden et al. 1993). It has been shown however that this method is not effective because the randomization procedures can have higher values than those observed so that no edge would be found significant (Fig. 4; Fagan et al. 2003). In the geographical approach, edges are mapped in a descending order (i.e., from the highest value to the lowest one) up to the level at which the study area is split into two patches (Fortin and Drapeau 1995). This method can result in having several isolated edges (called singletons; Jacquez et al. 2000) over the entire study area making comparisons among data sets more difficult. In the arbitrary one, edges above a numerical threshold or proportion are considered as boundary locations (BLs) and the significance test

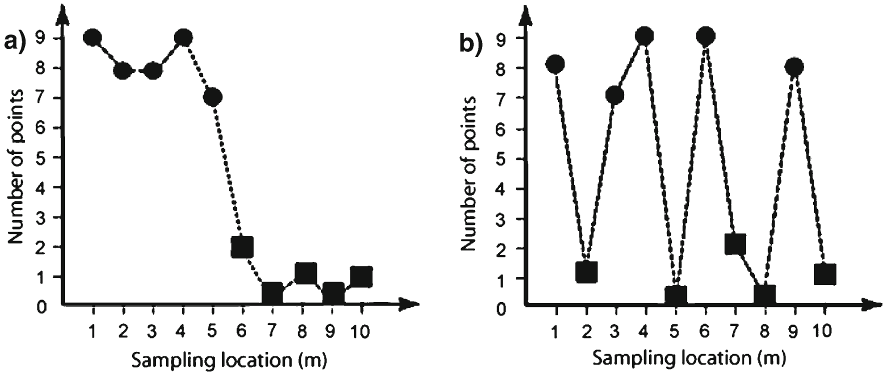

Fig. 4 Example of observed quantitative values belonging to two patches (circles and squares) along a transect (**a**). When the observed values are randomized (**b**) highest differences are found than in the observed data which could be interpreted as if there were not distinct patches in the observed data

is based on these boundary units as entities using boundary statistics rather than the individual edges (Oden et al. 1993). This last approach, although subjective, allows the comparison among different simulated maps and spatial resolutions. Hence, locations where the rate of change value was above a given threshold were identified as BLs and those below the threshold were considered as being part of a homogeneous patch.

For the simulated data (Fig. 2), there are 50×50 cells for a total of 2,401 edges ($n = (n\text{col} - 1) \times (n\text{row} - 1) = (50 - 1) \times (50 - 1)$). From these edges, BLs were selected using two thresholds: 3% and 10%. The first 3%-threshold corresponds to the actual number of BLs forming the known boundary separating two homogeneous patches, where there are 72 BLs ($72/2401 \approx 3\%$). A second threshold arbitrarily set at 10% was also used as it has been shown to be a good estimator in boundary detection of ecological and simulated data sets (Fortin 1994). Using that 10%-threshold implies the definition of 240 BLs (10% of 2,401 cells). Because there is no known boundary in the real data set, only the 10%-threshold was used for detecting the boundary between black spruce and balsam fir domains.

2.5 Data aggregation

The lattice-wombling being a local edge detector, it is sensitive to local variation (noise; Fig. 1). Data aggregation (using the average value) into a coarser resolution with a 2×2 window is expected to reduce the effect of local variation and to favour the detection of meaningful ecological boundaries rather than high rates of change caused by local noise. To avoid any confusion with other terms related to the concept of scale, *aggregation level* will be used hereafter. The simulated data were aggregated twice: a first aggregation (aggregation level 2), where the value of each pixel corresponds to the average value a of 2×2 window from the original resolution (aggregation level 1), and a second aggregation (aggregation level 3), where each pixel covers a 4×4 cell size from the original resolution (aggregation level 1). Real data from boreal forest were aggregated at six aggregation levels, producing various resolutions equivalent to the following cell sizes at original resolution: 2×2, 4×4, 8×8, 16×16, 32×32 and 64×64.

For the simulated landscapes, the true boundary counts 36 BLs at aggregation level 2 and 27 BLs at aggregation level 3. The proportion of BLs found over the original image at these two aggregation levels is not equal to 3% as it was the case for the analysis of the first level the original resolution; at level of aggregation 2, the 36 BLs found to represent approximately 6%, whereas the BLs from the level of aggregation 3 correspond to 20% of the total spatial units. The three latter thresholds (3, 6 and 20%) for the three aggregation levels are hereafter termed *true boundary thresholds* at the respective aggregation levels.

2.6 Sensitivity analysis

In order to assess the sensitivity in spatial location of the boundary detected by the lattice-wombling according to spatial and statistical conditions, a binomial test was used to evaluate the significance of the number of BLs from a given simulated landscape found at the same location as the BLs from the original image (Barbujani and Sokal 1991). Indeed, overlapping BLs from simulated landscapes with those from the original image gives an indication of the sensitivity of the edge detection algorithm by looking for direct overlaps between the two (Fortin et al. 1996). It is the number of these direct overlaps that is tested under a binomial distribution (overlaps versus non-overlaps), where the number of trials is the total number of BLs forming the true boundary at each aggregation level.

Multiscale representations of the results from boundary detection over simulated and real data were obtained by mapping the amount of spatial occurrences of BLs across all levels of aggregation (cell sizes from 2×2 to 64×64), for a given location.

2.7 Local spatial association

While edge detectors allow the identification of heterogeneous areas in a landscape, indices of local spatial association provide information on local levels of homogeneity, which can also contribute to patches' identification (Fortin and Dale 2005). Here, we used the Getis-Ord's G^* (Getis and Ord 1992; Boots 2002) which is a local measure of spatial association. This statistic defines the level of association resulting from a concentration of similar values in an area (defined by weighted locations) in relation to all the values of the study area. G^* is expressed by:

$$G_i^* = \frac{\sum_j W_{ij}(d)x_j}{\sum_j x_j}$$

where $i = j$, W_{ij} is the connectivity matrix, d is the distance, x a quantitative variable at a location and i and j are a given location and neighbours respectively. The null hypothesis for G_i^* states that the sum of the values of all sites j within a radius d (i.e., neighbourhood search distance) around site i is not different than the one obtained under random conditions. If there is spatial association in the data set, one will observe a spatial clustering (i.e., patch) of high or low values (relatively to the average of the distribution). Spatial clusters of high values will be indicated by a positive result of G_i^* while a negative result indicates clusters of low values.

One of the limits of G_i^* is that it identifies clusters, areas, that are spatially homogeneous, because they share similar values that are either high ("hot spots") or low ("cold spots") ones, but the determination of these clusters are all relative to the average value to start with. Also, significance testing are associated with the Getis-Ord statistics, as well as with other local indicators of spatial association, is problematic as the presence of global autocorrelation over the study area violates the assumption of randomness of the Getis' statistics, that is the absence of spatial structure at the landscape level (see Ord and Getis 1995 for a detailed discussion). Because of the problems of significance testing, G_i^* is not used here on the basis of its statistical significance but rather as an indicator of local spatial variation for an exploratory analysis.

Local spatial associations were estimated at different scales by using various window sizes within which G_i^* was calculated: 3×3, 5×5, 7×7, 9×9 and 11×11. For each location, two measures were used to summarise the local spatial association (Wulder and Boots 1998): the maximum absolute value of G^* (Gmax) found throughout all G^*s computed at various scales and the maximum G^* distance (MGD), which corresponds to the scale at which Gmax was found. Gmax thus indicates the highest level of local spatial association found across scales while MGD identifies the range (or local extent, neighbourhood size) at which this maximum level of local stationarity (i.e., Gmax) was detected. These two values will be used as exploratory tools to determine the presence of spatial non-stationarity at local scale and its spatial extent.

3 Results

3.1 Local edge detector

The lattice-wombling results, over simulated landscapes after a thresholding procedure using an a priori knowledge of the real boundary location (true boundary threshold, i.e., ∼3% at aggregation level 1), are shown in Table 1. It can be seen that when the difference between the patches' average and spatial autocorrelation are fixed, the accuracy of the boundary detection tends to decrease as the level of variance increases. Conversely, increasing the difference between the patches' average augments the accuracy of the boundary delineation when other parameters are fixed. When only spatial autocorrelation is kept constant, the accuracy of the boundaries' location follows the ratio of variance to difference between the patches' average; accuracy is high when the ratio is small (low variance, high difference) and it is low when the ratio is high (high variance, low difference). For a fixed difference between the patches' average and a fixed variance, the algorithm appears to be sensitive to the global level of spatial autocorrelation: as the global level of spatial autocorrelation increases, the lattice-wombling shows a diminishing accuracy in finding the boundary at the real location. Fewer replicates show a significant number of BLs detected at the real location when one or both patches have no spatial autocorrelation (Table 1). The low accuracy of

Table 1 Number of landscapes (replicates) in which a significant number of direct overlaps was found at each aggregation level using a true boundary threshold for a total of 20 replicates, where the aggregation level 1 refers to the original resolution, level 2 is equivalent to a 4 × 4 kernel at the original resolution and level 3 to a 8 × 8 kernel at the original resolution

Variance	Regions' spatial autocorrelation	Difference between regions' mean											
		0			1			2			10		
		Aggregation levels											
		1	2	3	1	2	3	1	2	3	1	2	3
1	(0.0, 0.0)	0	0	0	12	20	20	20	20	20	20	20	20
	(0.0, 0.95)	2	1	0	5	13	7	17	20	20	20	20	20
	(0.0, 0.9999)	2	0	0	3	8	3	14	20	20	20	20	20
	(0.95, 0.95)	1	0	0	1	7	2	5	16	20	20	20	20
	(0.95, 0.9999)	0	0	0	0	4	3	7	16	19	20	20	20
	(0.9999, 0.9999)	0	0	0	2	2	1	2	14	20	20	20	20
2	(0.0, 0.0)	1	0	1	2	6	1	12	20	19	20	20	20
	(0.0, 0.95)	2	1	0	1	0	0	5	13	8	20	20	20
	(0.0, 0.9999)	0	2	0	3	2	0	5	11	7	20	20	20
	(0.95, 0.95)	0	0	0	0	2	0	0	4	1	20	20	20
	(0.95, 0.9999)	0	0	0	2	2	0	2	4	2	20	20	20
	(0.9999, 0.9999)	0	0	0	1	0	0	0	2	0	20	20	20
10	(0.0, 0.0)	2	1	0	3	1	0	1	0	0	10	20	19
	(0.0, 0.95)	4	1	0	1	3	0	0	2	1	6	12	6
	(0.0, 0.9999)	7	0	0	6	1	0	2	0	0	9	11	5
	(0.95, 0.95)	2	0	0	0	0	0	2	0	0	2	3	4
	(0.95, 0.9999)	1	0	0	2	0	0	2	1	0	2	1	0
	(0.9999, 0.9999)	1	1	0	2	0	0	2	1	0	7	3	0

boundary delineation due to a high ratio of variance to patches' difference is further decreased by a high level of spatial autocorrelation. Results from edge detection using a 10%-threshold (Table 2) suggest that it is an appropriate arbitrary estimator as they are comparable to those obtained using the true boundary threshold.

3.2 Aggregation

The aggregation level has an asymmetric effect on the lattice-wombling algorithm depending on the spatial structures of the simulated data. When the variance is low and the difference between patches' average is high (i.e., small ratio of variance to patches' difference), results are enhanced in comparison to those obtained at the original resolution (Table 1). When the variance is high, results from level of aggregation 2 are less different than those from the lattice-wombling, even less accurate. There is also a slight decrease in sensitivity to spatial autocorrelation as when one or both patches have no spatial autocorrelation, the algorithm shows a more accurate delineation of the true boundary. At aggregation level 3 the algorithm performs slightly better than at both first and second levels of aggregation when the variance is low (Table 1). However, when the ratio of variance to patches' difference is high, there is no improvement in the algorithm's accuracy. When compared to the second level of

Table 2 Number of landscapes (replicates) in which a significant number of direct overlaps was found at each aggregation level using an arbitrary 10%-threshold for a total of 20 replicates, where the aggregation level 1 refers to the original resolution, level 2 is equivalent to a 4 × 4 kernel at the original resolution and level 3 to a 8 × 8 kernel at the original resolution

Variance	Regions' spatial autocorrelation	Difference between regions' mean											
		0			1			2			10		
		Aggregation levels											
		1	2	3	1	2	3	1	2	3	1	2	3
1	(0.0, 0.0)	2	0	0	11	19	20	14	20	20	20	20	20
	(0.0, 0.95)	1	1	0	8	15	5	18	20	20	20	20	20
	(0.0, 0.9999)	2	0	0	3	10	5	15	20	20	20	20	20
	(0.95, 0.95)	0	0	0	1	8	5	3	15	19	20	20	20
	(0.95, 0.9999)	0	0	0	0	2	2	9	14	19	20	20	20
	(0.9999, 0.9999)	1	0	0	3	2	4	2	14	19	20	20	20
2	(0.0, 0.0)	0	0	0	2	8	5	14	20	19	20	20	20
	(0.0, 0.95)	2	1	0	0	1	0	5	11	8	20	20	20
	(0.0, 0.9999)	0	2	0	1	2	1	4	13	8	20	20	20
	(0.95, 0.95)	0	0	1	0	1	0	2	5	3	20	20	20
	(0.95, 0.9999)	0	0	0	1	4	0	1	5	4	20	20	20
	(0.9999, 0.9999)	0	0	0	0	0	0	0	2	7	20	20	20
10	(0.0, 0.0)	1	0	0	2	1	0	0	3	0	12	20	20
	(0.0, 0.95)	4	2	0	2	2	1	2	1	0	3	12	8
	(0.0, 0.9999)	2	1	0	2	0	0	1	2	0	7	11	7
	(0.95, 0.95)	0	0	0	0	0	0	1	1	0	0	3	5
	(0.95, 0.9999)	0	0	0	0	0	0	2	1	1	1	1	1
	(0.9999, 0.9999)	2	0	0	0	0	0	1	0	0	4	2	3

aggregation, there is a decrease in the algorithm's ability to detect the true boundary that follows an increase in variance and a diminution in the difference between patches' average. With regards to spatial autocorrelation, the algorithm is less sensitive at the second level of aggregation than at the first. Results from the third level of aggregation suggest an enhanced accuracy in comparison to the original resolution but not to the second level of aggregation.

When observing the effect of aggregation on BLs location from their spatial occurrences in simulated landscapes (Fig. 5), one notices that when there is no spatial autocorrelation in either patch, a high ratio of variance to patches' difference makes it impossible to identify a cohesive boundary. When the variance and the patches' difference are of the same value (ratio = 1), the true boundary is visible but at a coarse resolution as it was already visible after the edge detection at the third level of aggregation. A low ratio leads to a clear picture of the true boundary, mostly resulting from the BLs found at the second and third levels of aggregation. In the presence of spatial autocorrelation in one of the two patches, using a multiscale representation provides more or less the same results as when spatial autocorrelation is null for both patches in cases where the ratio of variance to patches' difference is low. It is when the ratio is close to or greater than one that such a representation brings new perspectives to identifying the landscape's patches. Indeed, even though the boundary does not appear

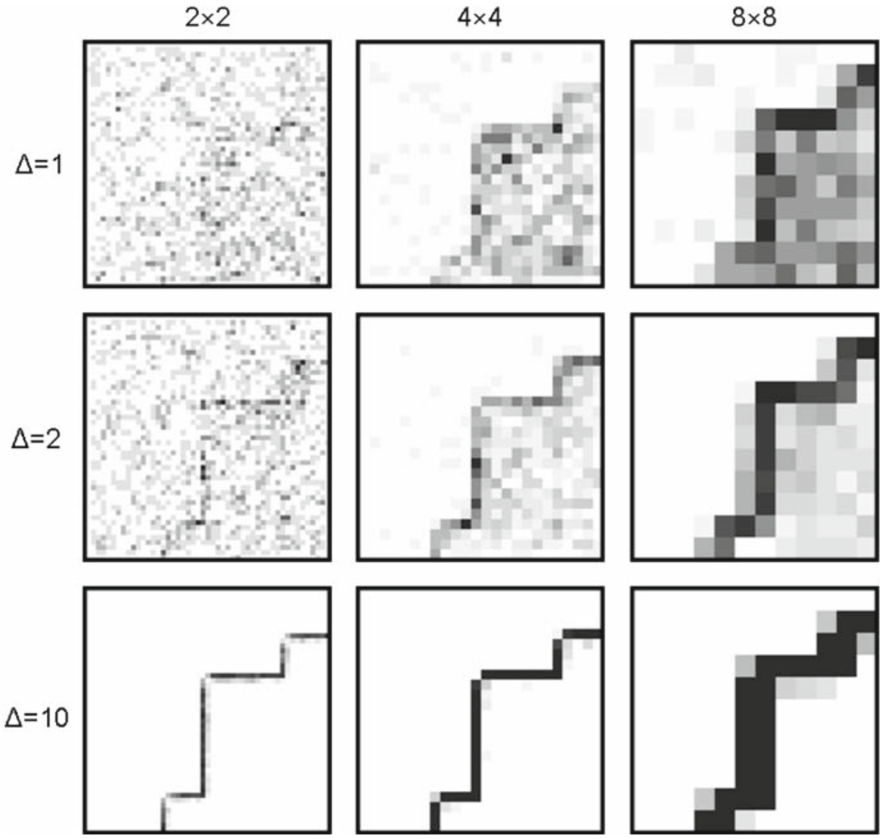

Fig. 5 Sum of spatial occurrences of BLs over the 20 replicates (white = 0 and black = 20) at the three aggregated levels (2 × 2, 4 × 4 and 8 × 8) where a conditional variance of 1 and there is no spatial autocorrelation in the upper-left patch but a high level (0.9999) in the lower-right patch, while the difference between patches' average is: 1 in the upper row, 2 in the middle row and 10 in the lower row

clearly under these conditions, this type of representation allows us to appreciate the differences in spatial structures between the two patches, especially those detected at the lower aggregation levels.

The results on simulated landscapes shown that a multiscale representation of the boundaries can be informative on the behaviour of spatial structures across aggregation levels. In Fig. 6, each pixel shows the number of times this location was defined as a boundary across all aggregation levels over the real data. High spatial occurrences of BLs across aggregation levels suggest stability in the spatial structure of ecotones (i.e., gradual boundary area rather than a sharp boundary). Moreover, mapping spatial frequencies of boundaries can provide information on the level of fuzziness of the boundaries across scales by indicating less stable transition zones. One can see that darker areas in Fig. 6 correspond to sharp transitions, from 20% to 40% black spruce cover or from 10% to 30% (Fig. 3), and that the boundaries' fuzziness (i.e., areas of lower BLs occurrences) can be associated to zones of high local heterogeneity in cover

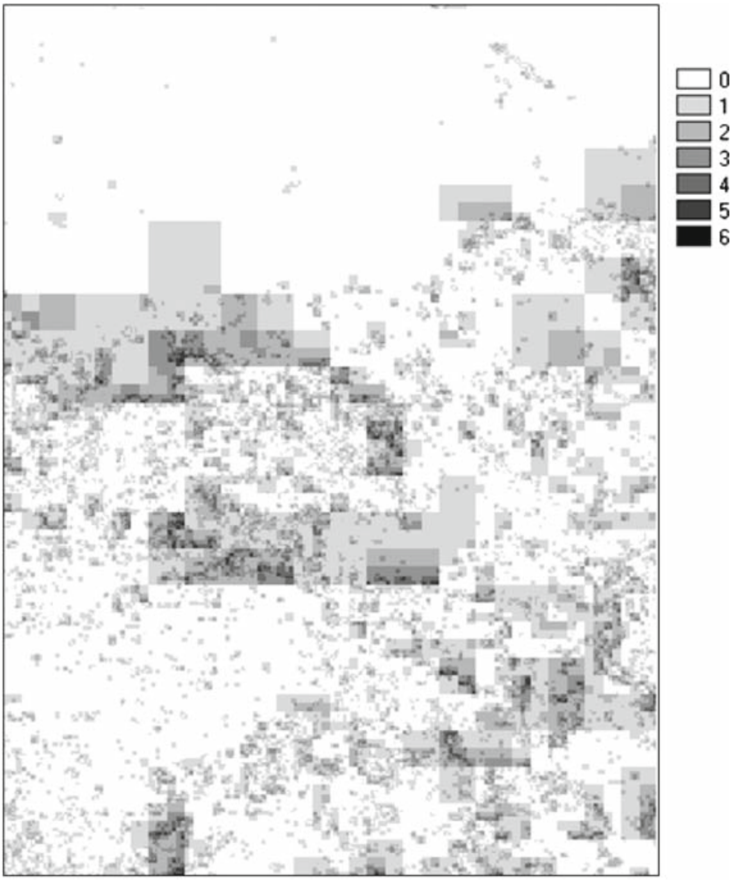

Fig. 6 Spatial occurrences of BLs across aggregation levels over the boreal forest data. Each location is given the value corresponding to the number of times (from 0 to 6) it has been detected as a BL across all analysed aggregation levels

percentage. Table 3 provides similar information from the proportion of areas defined as BLs at each aggregation level. Let us note that due to the various resolutions used after aggregating the data, the number of pixels can be misleading; it is preferable to consider them as areas defined as transition zones across all analysed aggregation levels (1 pixel = 14 ha at original resolution). These results also illustrate that the spatial structure of the transition between the black spruce domain and the balsam fir domain is relatively stable across aggregation levels (around 10% of whole landscape).

3.3 Local spatial association

Figure 7 shows a map of the maximum values of spatial association found across all the aggregation levels analysed (*Gmax*): dark green and dark blue colours indicate the highest levels of spatial association where positive spatial association (i.e., homo-

Table 3 Total number of BLs and proportion of landscape defined as transition areas at each aggregation scale from the boreal forest data

Aggregation scale	Number of BLs	% Landscape
2×2	13,472	0.1218
4×4	10,896	0.0985
8×8	10,368	0.0938
16×16	10,432	0.0943
32×32	11,520	0.1042
64×64	12,288	0.1111

geneity) is presented in green and negative spatial association (i.e., heterogeneity) in blue. The northern part of the area seems relatively homogeneous with a vast region of the same colour, that is, dark green indicating clusters of high values. The southern part appears more heterogeneous as many locations indicate low spatial association (yellow tints). However, low G^* values have to be interpreted with caution as they could also indicate clusters of medium values (Wulder and Boots 1998). Figure 8 illustrates the aggregation levels at which *Gmax* values were obtained (MGD values). The homogeneous patch in the northern area appears clearly as it reached its highest level of spatial association at the highest aggregation level used (11×11 window

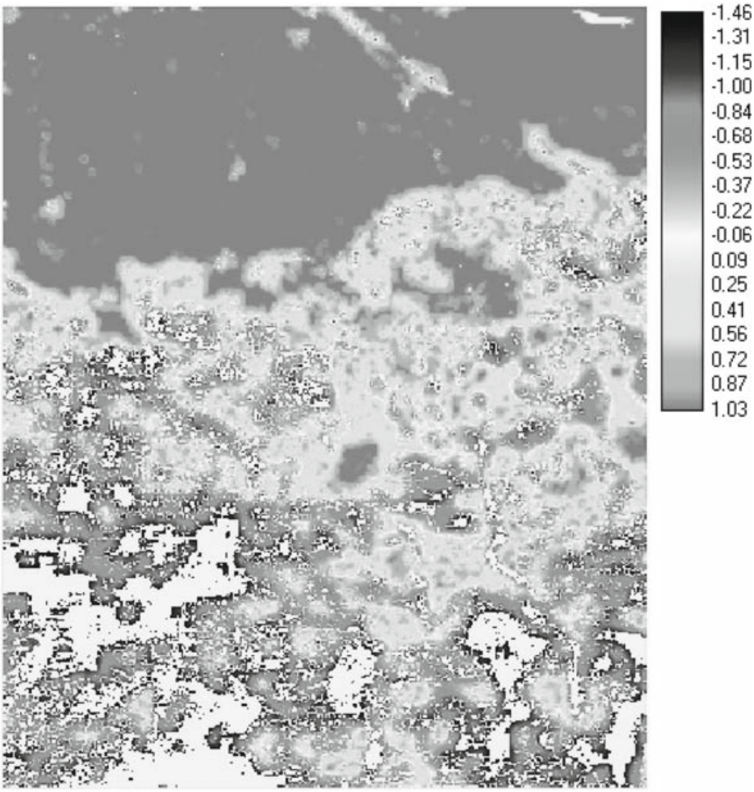

Fig. 7 Variations in Gmax values over the boreal forest data

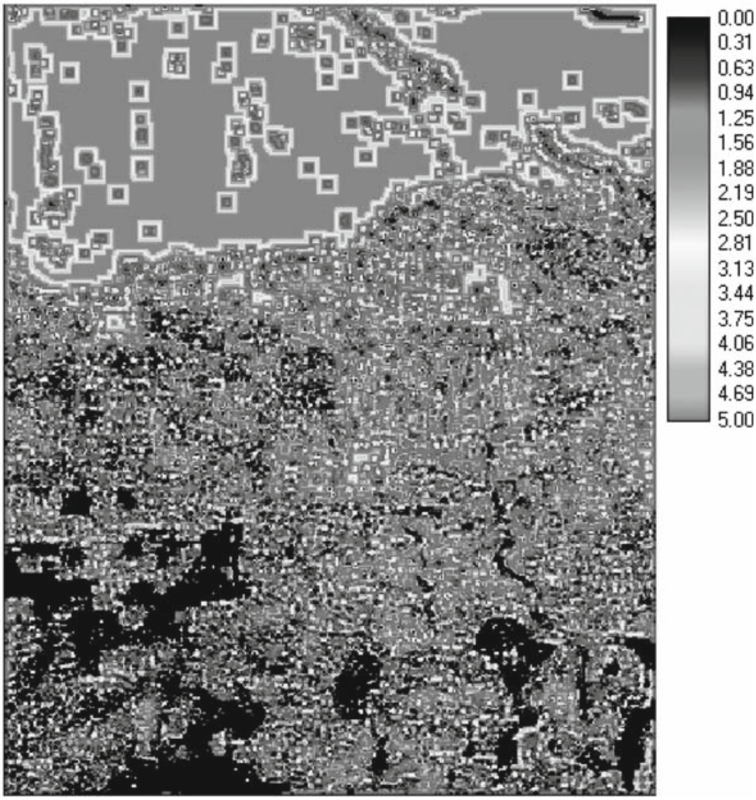

Fig. 8 Variations in MGD values over the boreal forest data

size). However, MGD does not allow for the identification of other clearly defined homogeneous patches as the southern portion of the landscape is characterised by a highest level of spatial homogeneity obtained at the lowest aggregation level (3×3 window size), suggesting a rather heterogeneous area.

4 Discussion

This study aimed at evaluating the effect of within-patch (local) spatial structures on the accuracy of local operators for locating patches and their boundaries. It was found that spatial autocorrelation greatly contributes to variations in the accuracy of boundary delineation by the lattice-wombling algorithm at different aggregation scales. Spatial autocorrelation affects the delineation of the true boundary by creating "inner-patches", that is, small clusters inside each patch. These inner-patches (or sub-patches) create "inner-boundaries" that are detected by the lattice-wombling as BLs. Local spatial heterogeneity (measured heterogeneity) thus impacts on the algorithm's performance in accurately delineating boundaries.

Variance plays a determining role on the accuracy of boundary delineation as well. The true boundary is found less often as the variance of the simulated landscapes increases. For boundary delineation, having a high level of variance in a landscape implies that some cells will have values that are not representative of the ecological dynamic creating the true boundary (Bradshaw and Fortin 2000). Increasing the difference between patches' average leads to an increase in the accuracy of the delineation of the true boundary by the lattice-wombling. Given the fact that the lattice-wombling is in essence computing differences among adjacent cells, this result was expected. A high difference between the patches' average creates a sharp rupture in the landscape that is easily detectable by the lattice-wombling; in such cases, adjacent cells located on each side of the boundary will present higher rates of change than if they are part of two patches sharing similar values. Hence, the ratio of variance to difference between patches' average affects a local edge detector's precision in finding the true boundary by varying the sharpness of the spatial gradient between two patches in the measured landscape. Therefore, when delineating boundaries, one has to consider, if not control for, both the combined effect of patches' average and variance and the presence of spatial autocorrelation in the data. The accuracy of a boundary delineation using a local edge detector will decrease with the presence of spatial autocorrelation in the data as well as with an increase in the ratio of variance to difference between the patches' average.

Following aggregation, some inner-patches and noise are smoothed out after being averaged with their surrounding at aggregation level 2. Although the algorithm's accuracy is generally enhanced after aggregating the data, the third level of aggregation does not provide a better delineation of the true boundary for all conditions. It is possible that significant differences between cells (i.e., rates of change indicating the true boundary as opposed to those created by inner-patches) are also smoothed out during the aggregation procedure since they are not high enough in comparison to those created by a high variance (Fortin 1999; Palmer and Dixon 1990). Indeed, data aggregation by averaging their values tends to decrease the amount of variance that makes the scalable-wombling detect BLs where no real boundaries exist. When the variance was relatively low, there has been an increase in the accuracy of the detection of the true boundary following the aggregation of the data as the scale of the analysis was gradually shifted towards the overall tendency of the ecological process (i.e., the underlying process at the origin of the two patches) and account less for the local variability. However, when the variance is high, aggregating the data leads to a diminution of accuracy that is caused by the fact that the variability of the data is too high for the edge detection algorithm to capture well the underlying process at the origin of the true boundary.

The combination of variance and difference between patches' average leads to two tendencies with respect to scaling: aggregating the data decreases the accuracy of the boundary delineation in cases where a high variance is combined with a low difference between the two patches' average (high ratio of variance to patches' difference), whereas the performance of the algorithm is increased when a low variance is coupled with a high difference in mean values (low ratio of variance to patches' difference). It suggests that aggregating smoothes out small differences between the cells located at or near the boundary by averaging together cells that are part of different patches.

This leads to a smoothing out of true location of the boundary. Such smoothing of the true boundary can create spatial changes in the cells' values that do not represent the underlying ecological dynamic but rather that are resulting from the noise in the data. Thus, aggregating improves the accuracy of the boundary detection when the ratio of variance to difference between patches' average is low, but not is cases where it is high, and it does not show to make much difference within the spectrum of spatial autocorrelation conditions.

Even though these results suggest that the aggregation procedure leads to a reduction in the accuracy of the delineation of the true boundary under certain conditions, one has to keep in mind that these results were obtained from simulated landscapes, of which the extent was kept constant (i.e., the landscape always represented the same surface). Indeed, the scale at which a landscape is analysed is defined by both its grain and its extent (Turner et al. 2001; Csillag et al. 2000), where the grain refers to the size of the spatial units (i.e., resolution) and the extent is the surface area covered by a landscape. This study shows the limit of aggregating a data set by modifying only the grain (resolution) of the landscape. Thus, when the ratio of variance to patches' difference is high, it is sometimes necessary to expend the extent of a landscape in order to allow the overall spatial trend to be more explicit. By modifying the extent, it can be expected to find more homogeneity and therefore, to increase the accuracy of an edge detection algorithm.

One has nonetheless, as proven by these results, to remain cautious about the level of spatial autocorrelation found inside patches, regardless of the extent, as it affects greatly the accuracy of the edge detection. The accuracy of boundary delineation follows statistical properties based on the level of the ratio of variance to difference between patches' average but the presence of spatial autocorrelation confounds this relation: a small ratio coupled with a low level of spatial autocorrelation creates conditions where the true boundary can be found with reasonable accuracy whereas, at the opposite, an accurate delineation is more uncertain when both the ratio and the level of spatial autocorrelation are high. Even though each condition does not affect the accuracy of the boundary detection the same way, these results illustrate that the importance of these conditions have to be known or controlled for when delineating boundaries using local edge detection algorithms. Unfortunately, it is obviously impossible to determine such properties of the patches to be detected before conducting the analysis (or only by expecting the presence of spatial autocorrelation in the data based on the nature of the ecological process) and one has therefore to rely on exploratory analyses prior to the boundary detection procedure. Indices derived from the G_i^* statistic (Gmax and MDG) have been used on real data as examples of exploratory tools for getting information on local variations in the spatial structure of a landscape prior to a boundary analysis.

The multiscale representation obtained from mapping the occurrence of BLs at each location for all analysed aggregation levels enables to understand how spatial structures behave through scales. It seems that such a representation reproduces information obtained from boundary detection at individual aggregation levels except when there is a difference in the patches' level of spatial autocorrelation. It minimizes the uncertainty arising from the asymmetrical effect of aggregating the data that varies depending on the spatial structure of the landscape, which can be a useful precaution when spatial autocorrelation is suspected in the ecological process under investigation.

From the real data from the boreal forest one can see a resolution-variant pattern (Fig. 6), even an embedded structure, suggesting that a field representation using a single resolution might not be the most appropriate geographical model from which boundaries can be delineated. Indeed, because of the non-stationary state of the ecological dynamic from which originates the landscape, areas of transition (zones of local heterogeneity) are not at the same locations across scales and so, this ecological dynamic might be better represented using a scale-variant geographical model.

Identification of BLs across scales (Table 3) allows the observation of the behaviour of the spatial structure across scales. Having a proportion of BLs to the whole landscape that is stable around 10% indicates a certain stability in that the total areas defined as transition zones remains the same, even if these ecotones are not always delineated at the same locations depending on the scale of analysis. A variation in this proportion of BLs would indicate changes in the importance of transition areas in relation to the entire study area, possibly originating from shifts in the detected *ecological signal* across scales. It is worth mentioning that this proportion is not an artefact of the 10%-threshold used to define boundaries. In effect, this arbitrary threshold implies that BLs are defined following the value of the 10th decile of the ranked distribution of the rates of change. However, at this rank the actual value of the rate of change can be shared by many locations (cells) and therefore the cut-off between BLs and non-BLs is also re-adjusted to include all other locations sharing the same value of rate of change than the one found at the 10%-threshold. Therefore, the actual number of BLs can represent a different proportion of locations than 10%.

As for the simulated data, aggregating the data from the Abitibi study region only by modifying the resolution might not be enough for reaching other spatial structures. It is presumed that by modifying the extent of the area as well, it would have been possible to aggregate using coarser resolutions and probably observe abrupt changes in the proportion of transition areas across scales corresponding to shifts from one *functional scale* to another (Cao and Lam 1997). Cao and Lam (1997) describe the functional scale as the range of *observational scales* (spatial resolutions or grain sizes) in which the spatial structure remains relatively stable for there are no major changes in the ecological signal captured. Changes in functional scale occur when the most explicit ecological process changes, for example, from individual trees to a forest, that is from one ecological manifestation to another (Cao and Lam 1997). In the case of the Abitibi data, we could expect from such a shift the emergence of two distinct, and relatively homogeneous, forest domains (patches) in comparison to the somewhat heterogeneous landscape at original resolution (Fig. 3).

The northern region of the Abitibi data appears very smooth with a high level of spatial autocorrelation (i.e., a very homogeneous patch). A transition is also visible between this patch and the other one, in the south of the study area, which shows more local heterogeneity. *MGD* indicates local ranges of homogeneity, that is the vicinity within which the highest level of homogeneity is found and can therefore be complementary for detecting shifts in the functional scale. One can see that it allows the visual delimitation of the northern patch, delineating its boundary at the north of the transition towards the southern patch. The latter, however, does not appear as clearly but it can be presumed that it would have appeared more homogeneous if *MGD* had been computed using larger neighbourhoods (window sizes). Nonetheless, one can appreciate the usefulness of *MGD* for distinguishing patches with different spatial structures. Indeed, even without computing *MGD* over larger neighbourhoods, which often necessitate landscapes of a greater extent, it is possible to identify different patches, provided that their spatial structures' characteristics are used as attributes to define them. As shown by these results, *Gmax* can help identifying the presence of spatial non-stationarity and its variation throughout the landscape while *MGD* could provide useful insights for defining the resolution at which each area is better represented for boundary detection (Philibert 2002).

5 Conclusion

Within the scope of our simulations, we demonstrated that boundary detection using a local edge detector is greatly affected by the spatial conditions of the data, namely variance, abruptness of the spatial gradient between two patches and patches' level of spatial autocorrelation. Inner-patches and ratio of variance to patches' affect the accuracy of the delineated boundaries using local edge detectors, depending on their magnitude and mutual influence. One has therefore to be cautious when analysing spatial patterns of boundaries of non-stationary processes using such local edge detectors. Ideally, spatial and statistical conditions would be controlled for in order to allow for a "truer" representation of the ecological dynamic but unfortunately, "translating" raster information (i.e., continuous fields) into vector objects is a complex task as it necessarily implies uncertainty and loss of information.

It has also been shown that data aggregation is not a panacea for bringing out more clearly the ecological process creating the patches. Depending on the landscapes' conditions, aggregation can mask meaningful ecological variability in the vicinity of the boundary. However, the question of scale remains crucial as the resolution of a geographical model influences the accuracy of boundaries delineation, either by inducing local variability when the resolution is too fine, or by masking information that is meaningful for the analysis. Ideally, one would be able to study a given area using various extents, for example by sampling larger landscapes than needed. In practice, however, data collection is often limited by technical and financial resources, constraints under which the largest possible amount of information is usually gathered.

The multiscale representation obtained from the occurrences of BLs helps to visualise the presence of local variability of spatial structures for conditions under which the use of a local edge detector implies greater uncertainty (i.e., in the presence of spatial autocorrelation and variance). However, the pattern of the spatial frequencies suggests a multiscale spatial structure that is hardly detectable in whole without any uncertainty if a raster model is used. Developments in geographical models and analytical methods that allow a resolution-variant representation of landscapes would benefit from boundary detection in complex spatial structures.

Before going through a boundary detection process, one might have to adjust the aggregation level of the data (observational scale). Indicators derived from local measures of spatial association, such as *Gmax* and *MGD*, become complementary tools for analysing spatial structures affecting the accuracy of the patch/boundary delineation. This study has shown that they cannot only identify the local ranges of spatial autocorrelation, and from there the functional scales, but also directly contribute to the delineation of spatial objects as exploratory tools for informing on the local variability of the spatial structure.

Acknowledgements We are thankful to the National Center for Excellence Geomatics for Informed Decisions (GEOIDE) network of Canada for financial and operational support during this research (project ENV#4). Dr. M. Wulder from the Canadian Forest Services, Dr. T. Remmel from the University of Toronto and Mr. J. Stoodley from Simon Fraser University have been of considerable and much appreciated help by providing us computing resources. We also thank the anonymous reviewers for their insightful questions and constructive recommendations.

References

Barbujani G, Oden NN, Sokal RR (1989) Detecting regions of abrupt change in maps of biological variables. Syst Zool 38:376–389

Barbujani G, Sokal RR (1991) Genetic population structure of Italy. 1. Geographic patterns of gene frequencies. Hum Biol 63:253–272

Boots B (2002) Local measures of spatial association. Ecoscience 9:168–176

Bradshaw GA, Fortin M-J (2000) Landscape heterogeneity effects on scaling and monitoring large areas using remote sensing data. Geogr Inform Sci 6:61–68

Cao C, Lam N (1997) Understanding the scale and resolution effects in remote sensing and GIS. In: Quattrochi DA, Goodchild MF (eds) Scale in remote sensing and GIS. CRC Press, Boca Raton

Cressie NAC (1993) Statistics for spatial data. Wiley, New York

Csillag F, Boots B, Fortin M-J, Lowell K, Potvin F (2001) Multiscale characterization of boundaries and landscape ecological patterns. Geomatica 55:291–307

Csillag F, Fortin M-J, Dungan J (2000) On the limits and extensions of the definition of scale. Bull ESA 81:230–232

Fagan WF, Fortin MJ, Soykan C (2003) Integrating edge detection and dynamic modeling in quantitative analyses of ecological boundaries. Bioscience 53:730–738

Fortin M-J (1999) The effects of quadrat size and data measurement on the detection of boundaries. J Veg Sci 10:43–50

Fortin M-J (1994) Edge detection algorithms for two-dimensional ecological data. Ecology 75:956–965

Fortin M-J, Dale MRT (2005) Spatial analysis: a guide for ecologists. Cambridge University Press, Cambridge

Fortin M-J, Edwards G (2001) Accuracy issues related to the delineation of vegetation boundaries. In: Goodchild M, Friedl M, Case T, Hunsaker C (eds) Uncertainty data in ecology. Elsevier, New York

Fortin M-J, Drapeau P (1995) Delineation of ecological boundaries: comparison of approaches and significance tests. Oikos 72:323–332

Fortin M-J, Drapeau P, Jacquez GM (1996) Statistics to assess spatial relationships between ecological boundaries. Oikos 77:51–60

Fotheringham AS, Brunsdon C, Charlton M (2000) Quantitative geography: perspectives on spatial data analysis. SAGE, London

Getis A, Ord JK (1992) The analysis of spatial association by use of distance statistics. Geogr Anal 24:189–206

Gosz JR (1993) Ecotone hierarchies. Ecol Appl 3:369–376

Haining R (1990) Spatial data analysis in the social and environmental sciences. Cambridge University Press, Cambridge

Jacquez GM, Maruca S, Fortin M-J (2000) From fields to objects: a review of geographic boundary analysis. J Geogr Sys 2:221–241

Oden NL, Sokal RR, Fortin M-J, Goebl H (1993) Categorical wombling: detecting regions of significant change in spatially located categorical variables. Geogr Anal 25:315–336

Ord JK, Getis A (1995) Testing for local spatial autocorrelation in the presence of global autocorrelation. J Regional Sci 41:411–432

Palmer MW, Dixon PM (1990) Small-scale environmental heterogeneity and the analysis of species distributions along gradients. J Veg Sci 1:57–65
Philibert M (2002) Effect of spatial and statistical conditions on boundary detection across scales. M.Sc. Thesis, Simon Fraser University
Pitas I (2000) Digital image processing algorithms and applications. Wiley-Interscience, New-York
Turner MG, Gardner RH, O'Neill RV (2001) Landscape ecology in theory and practice: pattern and process. Springer, New York
Wulder M, Boots B (1998) Local spatial autocorrelation characteristics of remotely sensed imagery assessed with the Getis statistic. Int J Remote Sens 19:2223–2231

Author Biographies

Mathieu D. Philibert after completing his M.Sc. in geography working on spatial statistics and boundary detection, he joined the National Public Health Institute of Québec as a research professional. In his current work in spatial epidemiology, he applies spatial statistics into research on socio-environmental influences on population health.

Marie-Josée Fortin is a Full Professor of spatial ecology at the Department of Ecology and Evolutionary Biology, University of Toronto. Her principal research interests are in forest ecology, landscape ecology, conservation biology and spatial statistics. She investigates the interactions among the spatial dynamic processes responsible for creating and maintaining landscape spatial heterogeneity which affect biodiversity and species conservation.

Ferenc Csillag (late) was a Professor (and Chair) at the Department of Geography, University of Toronto at Mississauga. His research interests were focused on understanding uncertainty in geographical information analysis. His GUESS Lab has been involved in projects with the GEOIDE Network of Centres of Excellence, Parks Canada, the Canadian Forest Service, the Ontario Ministry of Natural Resources and others in combining GIS, remote sensing, spatial statistics and bio-geophysical modeling.