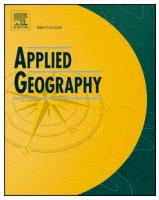
Applied Geography 158 (2023) 103045

Contents lists available at ScienceDirect

Applied Geography

journal homepage: www.elsevier.com/locate/apgeog



Effects of cluster plot design parameters on landscape fragmentation estimates: A case study using data from the Swedish national forest inventory

Habib Ramezani [a,*], Andrew Lister [b]

[a] Department of Forest Resource Management, Swedish University of Agricultural Sciences, SLU, SE-901 83 Umeå, Sweden
[b] USDA Forest Service, Forest Inventory and Analysis Unit, Northern Research Station, 3460 Industrial Highway, York, PA 17402, USA

ARTICLE INFO

Keywords:
Landscape pattern
Sample-based estimation
Forest monitoring
Landscape change
Forest fragmentation monitoring
Forest degradation monitoring

ABSTRACT

Forest fragmentation is commonly characterized using indices derived from analyses of classified land cover maps. An alternative is to use data obtained from sampling, such as those from a national forest inventory (NFI). The main objective of the current study is fill knowledge gaps on the performance of sample-based forest fragmentation metrics calculated with different cluster plot designs and under different forest conditions. A set of NFI cluster plot designs, each with different geometric properties, was created from Swedish NFI data. Each member of the set was used to calculate the fragmentation metrics mean patch size (MPS) and perimeter-area ratio (PA). Impacts of plot design parameters on metric estimates and their precision were assessed.

Important differences in metric values were observed both within and between regions under different plot design scenarios; within regions, ranges of PA and MPS values were large, and confidence intervals for the minimum and maximum metric values did not overlap. Weighted least squares regression significance testing results suggest that subplot separation distance was an impactful design factor whereas number of subplots and cluster shape were less important. However, cluster plots with more and widely-separated subplots yielded estimates that were more precise (lower relative sampling errors) than smaller, more compact clusters. We suggest that care should be taken when interpreting the physical meaning of the metrics under study.

1. Introduction

Forest ecosystems play an important role in protecting biodiversity (Corona et al., 2019). However, their integrity has been threatened by forest fragmentation due to human activities, including forest management, agriculture, industrialization, and urbanization (Grantham et al., 2020). Forest fragmentation is a dynamic process in which a contiguous tract of forest is broken into smaller and more-isolated patches (Tolentino & Anciães, 2020). Fragmentation has many negative impacts on vegetation, wildlife, biodiversity, and ecosystem services provided by forested landscapes (Lister et al., 2019; Shapiro et al., 2016). It is also recognized that climate change and fragmentation may have combined effects on habitat loss (Pyke, 2004). Therefore, it is important that fragmentation be accurately quantified for management and monitoring purposes, as well as for its potential to inform forest degradation analyses for countries interested in participating in deforestation and degradation reduction incentives programs (Lister et al., 2019).

The most common method of quantifying forest fragmentation is using Geographic Information System (GIS) software to analyze raster data, such as a satellite or aerial images, in which pixels have been assigned a land cover or use class (Hassett et al., 2011; Lister et al., 2019; McGarigal & Marks, 1995). Other methods involve delineating cover type patches manually or with Object-Based Image Analysis (OBIA) on high resolution imagery. These methods are not without problems. For instance, manual patch delineation and classification on aerial photos can be extremely labor-intensive, and OBIA requires fine-tuning of segmentation parameters and can require post-processing to manually adjust maps (Ye et al., 2018). By contrast, automatic techniques like supervised or unsupervised classification are more time-efficient than using manual approaches, but costs can increase due to the time required for correction (Ramezani et al., 2013). Furthermore, land cover maps based on medium-resolution satellite imagery often have low accuracy for certain rare classes (Fang et al., 2006), and it is costly to obtain information from high-resolution imagery such as Quickbird or

* Corresponding author.
E-mail addresses: Habib.Ramezani@slu.se, Ramezani.Habib@gmail.com (H. Ramezani), Andrew.lister@usda.gov (A. Lister).

https://doi.org/10.1016/j.apgeog.2023.103045
Received 21 September 2022; Received in revised form 7 June 2023; Accepted 3 July 2023
Available online 14 July 2023
0143-6228/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Fig. 1. Locations of Swedish national forest inventory regions, with regions 4 and 5 with bold outline. Imagery credit: Earthstar Geographics World Imagery service.

IKONOS (Hassett et al., 2011). In some cases, existing raster maps may not be available for the area or time period of interest, or, if available, they might use landscape class definitions that do not align with the desired ones. Finally, some maps are created with different methods, resolutions, or definitions for different time periods, making trend analyses very difficult (Nelson et al., 2009).

Many countries use cluster plot designs for their national forest inventories (NFIs) (Ramírez et al., 2022; Tomppo et al., 2010), and these can have different designs in different subpopulations within a country (Axelsson et al., 2009; Fridman et al., 2014). In addition, cluster plot protocols using ocular interpretation of high-resolution imagery, such as Collect Earth Online, have recently played important parts in global efforts to monitor deforestation (Saah et al., 2019). An interesting alternative to estimation of fragmentation metrics using raster maps is to use data obtained from sampling with cluster plots like these.

One important issue in sample-based estimation of the metrics is that many raster-based metrics cannot be estimated using sampling data like those obtained with NFI plots. This is due to the fact that many of the metrics were originally developed for wall-to-wall categorical maps such as land cover or use maps. NFI plots, on the other hand, were not originally designed to measure fragmentation. However, Kleinn (2000) developed an approach that allows for the calculation of mean patch size (MPS) and perimeter-area (PA) metrics from forest inventory cluster plots (which are commonly used in NFIs), and Nelson et al. (2009) used that method with NFI data from the United States. Ramezani and Ramezani (2015) proposed a point vector-based contagion metric (C) based on similar methods, where calculation of C relies on land cover class assignments made at subplot centers, and Lister et al. (2019) calculated several sample-based landscape metrics using NFI data from the US state of Maryland.

To our knowledge, there have been few or no previous studies undertaken to understand the impacts of cluster plot design factors on forest fragmentation metric values. There is therefore a knowledge gap in the forest monitoring community regarding the calculation and interpretation of fragmentation statistics using NFI and similar cluster plots with different designs under different landscape heterogeneity scenarios. The main objective of the present study is to fill this gap by determining how variation in NFI cluster plot designs can influence the magnitude and precision of estimates of sample-based fragmentation metrics mean patch size (MPS) and perimeter-area ratio (PA) in regions with different spatial patterns of forest. We hypothesized that there would be important variation in MPS and PA values under different cluster plot design scenarios and across regions with different forest patterns. The overall goal is to fill the knowledge gaps we identified and improve the capacity of practitioners to calculate and interpret sample-based fragmentation metrics.

2. Materials and methods

2.1. Study site characteristics

A dataset from a subset of one five-year cycle (2007–2011) of the Swedish NFI was used for the analyses. Sweden's forests, which occupy approximately 60% of the country, are comprised of large areas of pine, wetland, and spruce forest types with a smaller area comprised of a mixture of hardwood and mixed species forest types. Nonforest areas are mostly agricultural and a smaller area of developed land uses (Olsson & Ledwith, 2021). The country is divided into six forest inventory

Table 2
Information on the cluster plot designs used in the study for inventory regions 4 and 5. Original (**), permanent (P) and temporary (T) plot designs are labeled. Total # of subplots in each plot design = Total # of plots × number of subplots in design. The maximum distance (*d*) is calculated as the maximum distance between subplot centers in the design, and separation distance is between adjacent subplot centers.

Inventory region	Cluster plot size ($m \times m$)	Maximum distance (m)	Number of subplots in a cluster (Total # of plots)	Subplot Separation distance (m)
Region 4				
1**P	800 × 800	1130	8 (1185)	400
2**T	400 × 800	890	6 (1064)	400
3 P	800 × 800	1130	4 (1185)	800
4 T	400 × 400	570	4 (1064)	400
5 P	400 × 800	890	4 (1185)	400,800
6 P	400 × 800	1130	5 (1185)	400
7 P	800 × 800	1130	3 (1185)	800
8 P	800 × 800	570	3 (1185)	400
9 P	800	800	3 (1185)	400
10 P	400	400	2 (1185)	400
Region 5				
11**T	600 × 300	670	6 (369)	300
12 T	600 × 300	670	4 (369)	300,600
13**P	300 × 300	420	4 (849)	300
14 P	300 × 300	420	3 (849)	300
15 T	600	600	3 (369)	300
16 T	300	300	2 (369)	300

Table 1
Description of the sample and plot designs (permanent and temporary) associated with data used in the study, including the total area of the region and percent of the total area that is forest land area. The original plot size shown is the distance between the centers of the subplots located in the corners of the square plot. Sample sizes are for a five-year cycle (with count of subplots per cluster in parenthesis). The NFI cycle used was 2007–2011 in each inventory region.

Inventory region	Total area (km^2) (Forest land %)	Original cluster plot side length (km)		Cluster plot sample size (number of subplots)	
		Permanent[a]	Temporary[b]	Permanent	Temporary
4	116,848.48 (60%)	0.8 × 0.8	0.8 × 0.4	1185 (8)	1064 (6)
5	34,476.76 (50%)	0.3 × 0.3	0.6 × 0.3	849 (4)	369 (6)

[a] Subplots have a radius of 10 m.
[b] Subplots have a radius of 7 m.

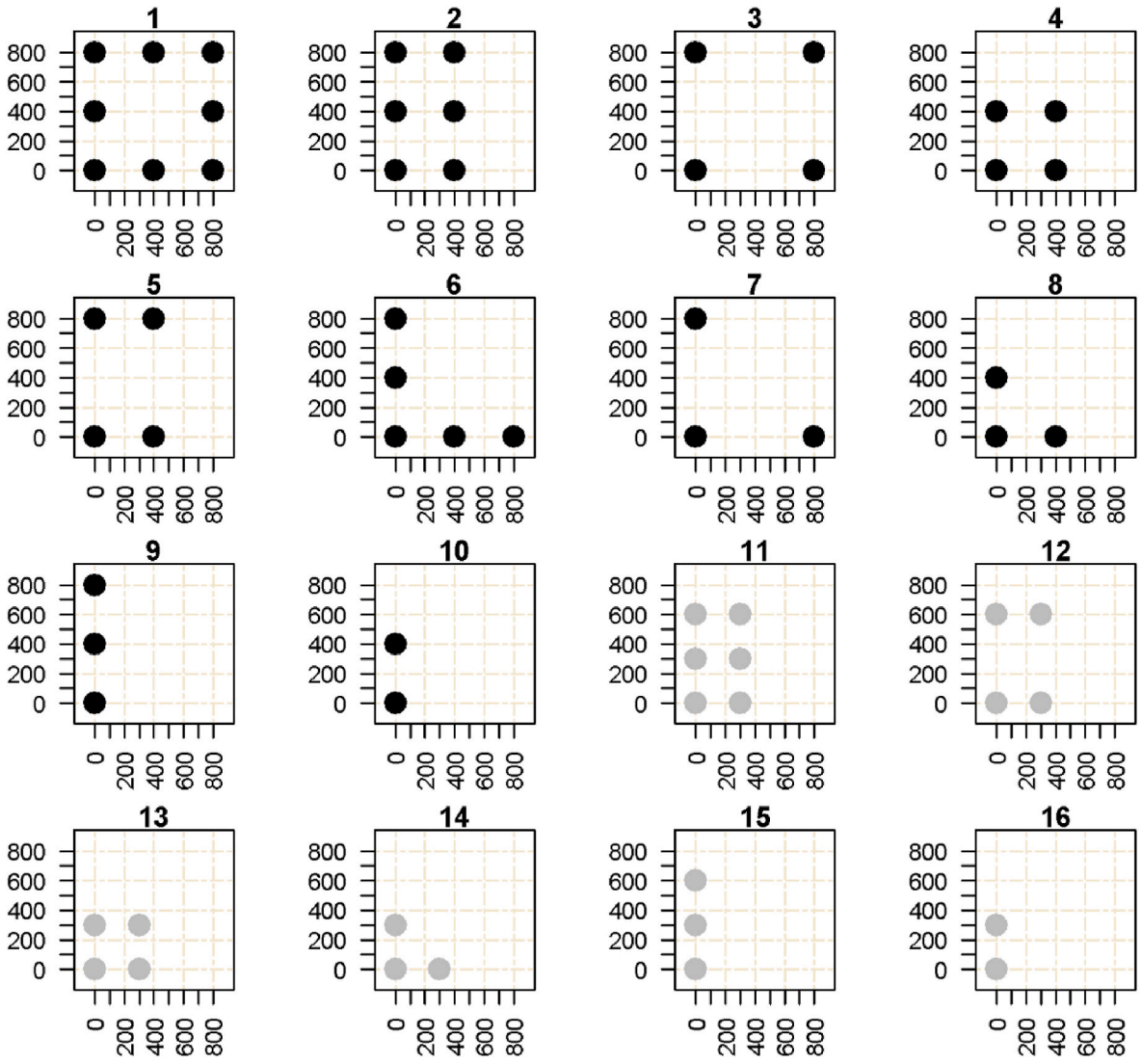


Fig. 2. Illustration of the sixteen cluster shapes and subplot separation distances (in meters) applied in the study. Region 4 figures have black points and region 5 figures have grey points. Cluster plots 1, 2, 11, and 13 are the original NFI cluster plot shapes, while other cluster plot designs were extracted from the originals. Detailed plot characteristics can be seen in Table 2.

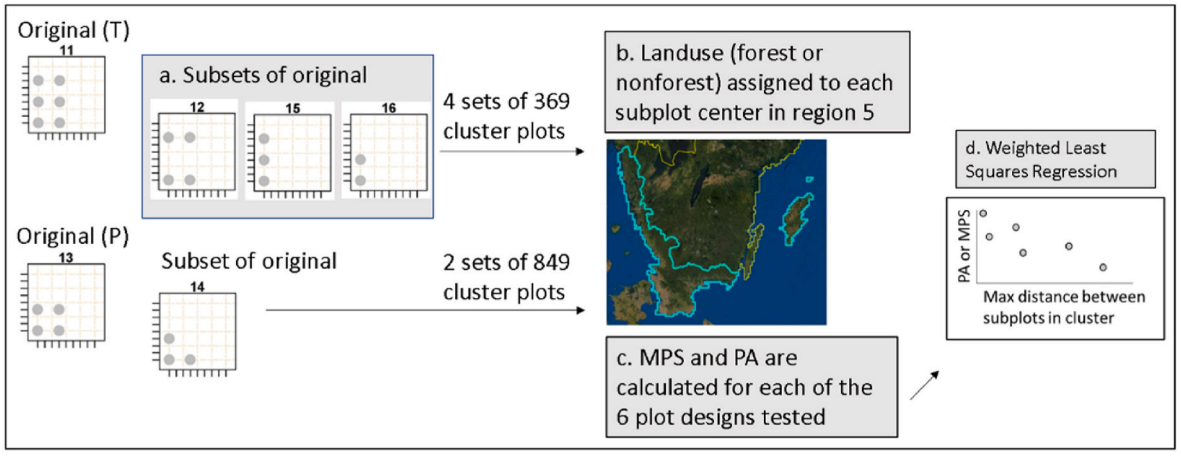


Fig. 3. The analysis workflow, depicted for region 5. Subsets of original Swedish NFI plots (T = temporary plots, P = permanent plots) are created, land cover is assigned to each subplot center in each of the six cluster plot designs, and fragmentation metrics (PA and MPS) are calculated as described in the text and relationships between these and design factors are assessed.

Table 3
Summary of inventory design characteristics and estimates of $\widehat{MPS}$ (km^2) and $\widehat{PA}$ (km/km^2) (and their components) for each of 16 forest inventory plot designs in two inventory regions in Sweden. SE % in parenthesis of two last columns is standard error of the estimate (Eqs. A7 and A8) as a percent of the estimate.

Inventory Region	Design ID (Fig. 1)	Total subplots in forestland	Number of Subplots in Design	Percent Forest ($\hat{p}$) (%)	P_{int} (%)	P_{incl}	Maximum separation distance (d) (km)	$\widehat{MPS}$ (SE %)	$\widehat{PA}$ (SE %)
4	1**P	4606	8	49	69	0.90	1.1	0.52 (3.8)	1.38 (2.6)
4	2**T	3321	6	52	66	0.86	0.9	0.37 (2.0)	1.65 (2.4)
4	3 P	2447	4	52	46	0.90	1.1	1.33 (6.4)	0.87 (6.2)
4	4 T	2198	4	52	55	0.90	0.6	0.23 (4.7)	2.08 (4.5)
4	5 P	2430	4	51	59	0.86	0.9	0.44 (3.4)	1.51 (3.5)
4	6 P	3041	5	51	58	0.77	1.1	0.58 (2.0)	1.31 (1.8)
4	7 P	1804	3	51	52	0.77	1.1	0.73 (3.1)	1.17 (2.0)
4	8 P	1826	3	51	40	0.77	0.6	0.31 (3.6)	1.8 (2.7)
4	9 P	1815	3	51	42	0.64	0.8	0.38 (4.0)	1.62 (3.5)
4	10 P	1204	2	51	27	0.64	0.4	0.23 (7.1)	2.08 (7.0)
5	11**T	801	6	36	51	0.86	0.7	0.17 (6.0)	2.46 (5.0)
5	12 T	546	4	37	43	0.86	0.7	0.25 (6.5)	2.02 (6.0)
5	13**P	532	4	36	39	0.90	0.4	0.12 (6.5)	2.84 (6.5)
5	14 P	918	3	36	33	0.77	0.4	0.13 (5.5)	2.81 (4.8)
5	15 T	403	3	36	34	0.64	0.6	0.16 (7.3)	2.47 (7.0)
5	16 T	264	2	36	20	0.64	0.3	0.12 (8.0)	2.91 (10.0)

administrative regions that are distributed from north to south, with lower sample intensities toward the north of the country.

Data from inventory regions 4 and 5, which are located in the south of the country (Fig. 1), were used in the study. Region 4, which is approximately 117,000 km^2, is approximately 60% forest that consists mostly of pine and spruce forest types as well as lesser amounts of mixed and deciduous forests. Nonforest cover types include predominantly agricultural areas, water, sparsely-vegetated natural areas, and human-impacted areas (Olsson & Ledwith, 2021). Region 5, which is approximately 35,200 square kilometers and 50% forest, has a similar forest composition, but a higher proportion of agricultural and nonforest vegetated land than region 4 (Olsson & Ledwith, 2021). Both regions have a mix of both conifer and deciduous tree species, including pine, spruce, ash, elm, and oak, as well as many species of shrubs, grasses, and herbs that can be found in the forest and many wetland areas. Both regions have relatively flat topography and climate is milder than in other parts of the country.

2.2. National forest inventory (NFI) data

The Swedish NFI consists of a systematically distributed set of a combination of temporary and permanent cluster plots, with intensity varying by region. Beginning in the 1950s, a square cluster plot design (tract) was introduced into the NFI. Modifications to the design occurred over the years, but since the 1980s, when permanent square cluster plots with circular subplots were established, the plot and sample designs have essentially remained the same. The temporary cluster plots have different spacing than the permanent ones (Table 1), and their subplots have smaller radii (7 vs 10 m); despite these differences, plot types were intermingled for all analyses and thus distinctions between them are not considered. In inventory regions 4 and 5, the number of circular subplots per plot varies from 4 to 8 (Table 1).

2.3. Analysis strategy

2.3.1. Analysis goals

In this study, we applied a sample-based forest fragmentation measurement method (Kleinn, 2000) to a set of different plot designs derived from subsets of the original NFI plot designs and assessed the relationship between the metric values and plot design parameters. The forest patch characteristics with which the metrics are associated (mean patch size and patch perimeter:area ratio) are commonly linked to ecological processes such as species richness (Saura & Carballal, 2004). Our first goal was to test our hypothesis that cluster plot design has important influences on metric values. A second goal was to clarify the mechanisms behind metric performance, and how these relate to landscape pattern and plot design.

Enhancing the clarity of the information will aid in making better-informed decisions regarding the interpretation and use of metrics.

2.3.2. Plot design subset creation

Due to the unique structure of the Swedish NFI cluster plots, it is possible to extract subsets of subplots from each cluster with specific shapes. For example, an L-shape cluster can be made with a subset of the original square shape cluster. The full set of subplot designs used in the study is described in Table 2 and can be seen in Fig. 2. Fig. 3 gives a stylized example of the analysis framework for two original NFI plot designs in region 5. In the example, cluster plot design 11, which is the original temporary plot design from the NFI in region 5 (Table 2 and Fig. 2), is composed of six subplots, each of which has the land cover class (forest or nonforest) assigned. Three subset designs can be extracted from the original design to create a set of 4 designs in total. For each design, $\widehat{PA}$ and $\widehat{MPS}$ (and their sampling errors) were calculated using the land cover assignments and plot design parameters in the equations given in Appendix A. An identical process is shown in Fig. 3

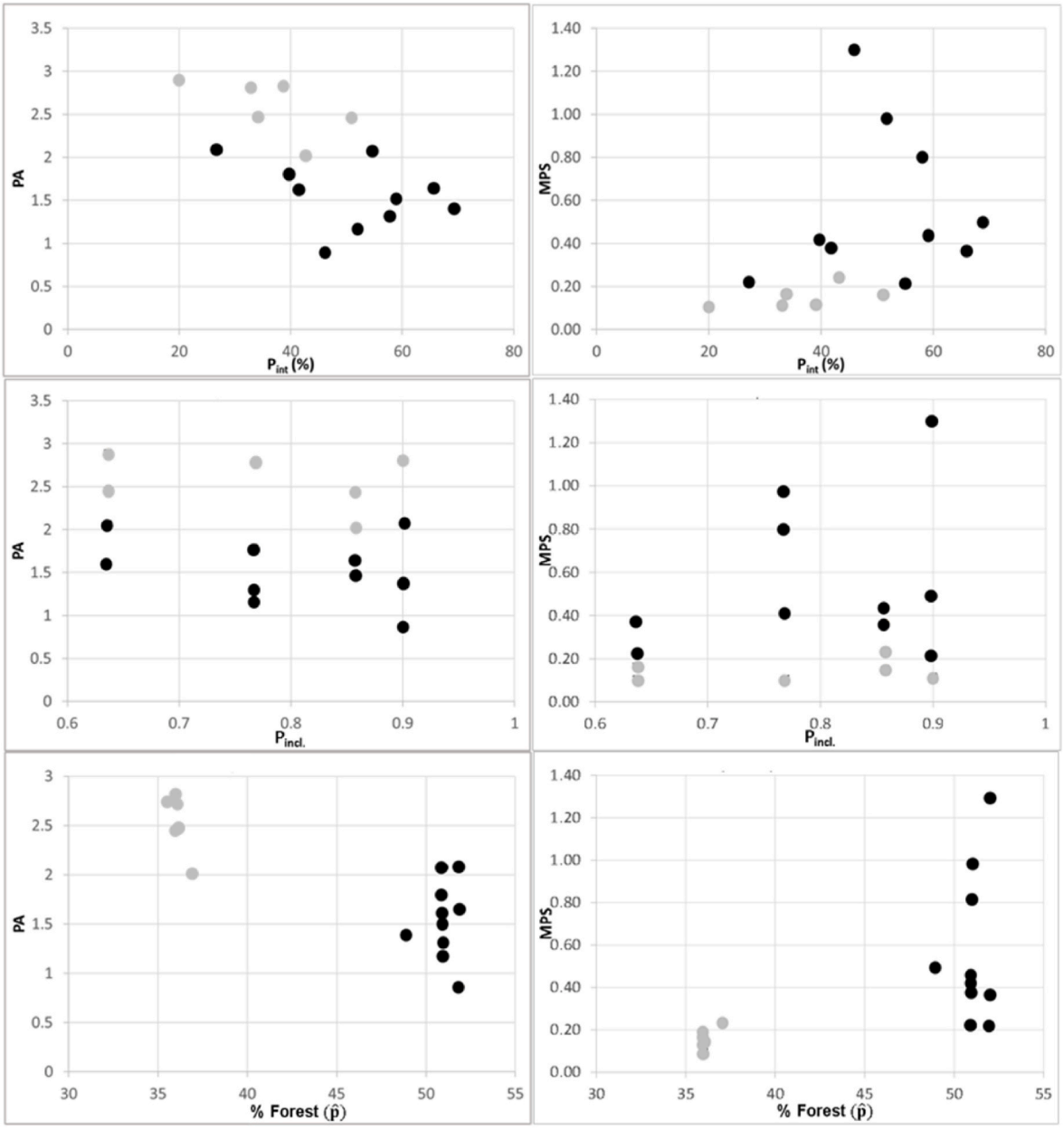


Fig. 4. Relationships between the landscape metrics Perimeter-area ratio (PA, km/km^2) and Mean Patch Size (MPS, km^2) and their components $\hat{p}_{int.}$, $p_{incl.}$, and $\hat{p}$. Grey points are for inventory region 5 and black points are for region 4.

for cluster plot design 13, and replicated for all other plot designs shown in Fig. 2 and Table 2).

2.3.3. Analyses conducted

Descriptive statistics and scatter plots were used to summarize results and identify relationships between metric values and design parameters, and a weighted least squares (WLS) regression (Draper & Smith, 1981) was used with the base R *lm* function (R Core Team, 2020) to test our null hypothesis of no relationship between the metrics (dependent variables) and their constituents and associated plot design parameters. WLS regression was used because the metric values had different levels of precision for each design, and weights were assigned as the inverse of the variances of the estimates. Confidence intervals were generated for $\widehat{PA}$ and $\widehat{MPS}$, and the overlap of these confidence intervals between all pair combinations of plot designs was assessed to provide evidence for the significance of differences between metric values generated with different designs.

3. Results

3.1. General results for $\widehat{PA}$, $\widehat{MPS}$, $\hat{p}$, $p_{incl.}$, and p_{int}

Summary statistical results ($\widehat{PA}$, $\widehat{MPS}$ and their sampling errors), as well as mean values for metric components are shown in Table 3. There are important differences in metric values both between and within regions. For example, the percent differences between the highest and lowest $\widehat{PA}$ in regions 4 and 5 are 139% and 44%, respectively, and those for $\widehat{MPS}$ are even larger (491% and 118%, for regions 4 and 5 respectively). In addition, confidence intervals for the highest and lowest values in each region do not overlap (Tables C1 and C2). This supports our conclusion that plot design choices are important when designing and analyzing results of sample-based fragmentation analyses using these metrics.

When averaged across all plot designs by region, the proportion

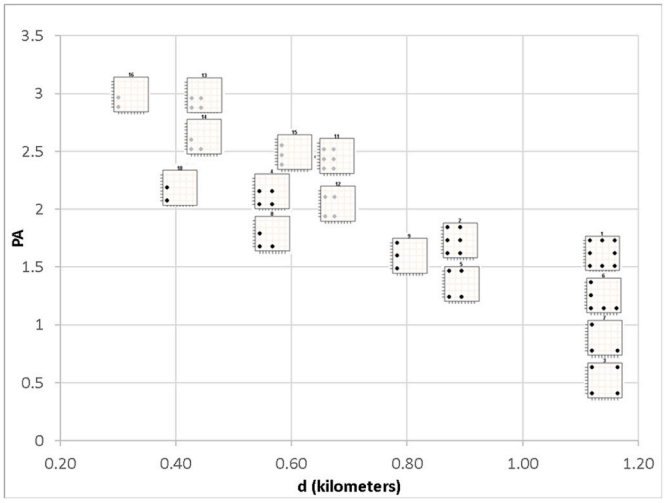


Fig. 5. Relationship between landscape metric Perimeter-area ratio (PA, km/km^2) and maximum subplot center separation distance (d) of the corresponding cluster design. Grey points are for inventory region 5 and black points are for region 4.

Table 4
Weighted least squares regression models, F statistics, p values, and R^2. Regressions represent the impacts of metric components and related plot design factors (Dist = maximum distance between subplots; Numsp = number of subplots in design; Pctfor = proportion forest ($\hat{p}$), Pint = the proportion of plots crossing forest/nonforest boundaries; and Pincl is the parameter linked to plot shape (Table A1)) on the dependent variables, $\widehat{PA}$ and $\widehat{MPS}$. Lines in bold are considered significant.

Region	Model	F	P value	R^2
4	**PA = -1.34(Dist) + 2.73**	**24.464**	**0.001**	**0.75**
4	PA = 0.02(Numsp) + 1.28	0.16	0.70	0.02
4	PA = 0.03(Pctfor) - 0.21	0.08	0.78	0.01
4	PA = 0.0007(Pint) + 1.33	0.005	0.95	<0.01
4	PA = 0.027(Pincl) + 1.35	0.0004	0.98	<0.01
5	**PA = -0.53(Pctfor) + 21.61**	**14.425**	**0.02**	**0.78**
5	PA = −1.81(Dist) + 3.52	5.586	0.08	0.58
5	PA = −1.47(Pincl) + 3.55	1.203	0.33	0.23
5	PA = −0.01(Pint) + 2.92	0.728	0.44	<0.01
5	PA = −0.03(Numsp) + 2.47	0.055	0.83	0.01
4	**MPS = 0.54(Dist) - 0.06**	**21.071**	**0.002**	**0.72**
4	MPS = −0.08(Pctfor) + 4.29	1.605	0.24	0.17
4	MPS = 0.003(Pint) + 0.20	0.762	0.41	0.09
4	MPS = 0.024(Numsp) + 0.27	0.675	0.44	0.08
4	MPS = −0.15(Pincl) + 0.50	0.074	0.79	0.01
5	**MPS = 0.11(Pctfor) - 3.69**	**9.053**	**0.04**	**0.69**
5	**MPS = 0.18(Dist) + 0.06**	**7.63**	**0.05**	**0.66**
5	MPS = 0.002(Pint) + 0.08	2.528	0.19	0.39
5	MPS = 0.012(Numsp) + 0.11	1.27	0.32	0.24
5	MPS = 0.1655(Pincl) + 0.03	1.143	0.34	0.22

forest ($\hat{p}$) and proportion of plots that cross nonforest boundaries ($p_{int.}$) for region 5 were 29% lower than those for region 4 (Table 3, Fig. 4). However, the coefficient of variation (CV) of the $\hat{p}$ values for both regions (1.7% and 1.1% for regions 4 and 5, respectively) were much lower than those of $p_{int.}$ (24.9% and 28.5%), suggesting that forest area estimates are much less impacted by plot design than are measures like $p_{int.}$, which are tied to the geometric properties of forest patch configuration.

3.2. Estimation of $\widehat{PA}$

Results for $\widehat{PA}$ indicate important relationships between metric values, some of their components, and associated plot design factors. The relationships between the components of Eq. (A1) and $\widehat{PA}$ are shown in Figs. 4 and 5. Within each region, Fig. 4 suggests no or weak, negative relationships between $\widehat{PA}$ and $\hat{p}_{int.}$, $p_{incl.}$, and $\hat{p}$, and WLS regression results support this. The only significant regression ($p \leq 0.02$, $R^2 = 0.78$) is for $\hat{p}$, and it occurs in region 5 (Table 4). However, Fig. 5 suggests that $\widehat{PA}$ showed a strong decreasing trend with increasing maximum subplot separation distance (d). This conclusion aligns with the WLS results; there is a significant ($p \leq 0.02$) regression that contains d for region 4, and a nearly significant ($p < 0.08$) regression for region 5, with high R^2 values (0.58 and 0.75, respectively). This suggests that subplot separation distance is an impactful plot design parameter.

With respect to number of subplots in the design, inspection of Fig. 5 suggests that the number of subplots in the design is weakly associated with the magnitude of $\widehat{PA}$. However, inspection of Table C1 reveals that there often exists overlap between confidence intervals (CIs) that are in the same or similar separation distance classes, suggesting a lack of significant differences in $\widehat{PA}$ values. For example, looking at plot designs with the largest maximum subplot separation distance of approximately 1.1 km (Fig. 5), CIs for design 1 (8 subplots) and design 6 (5 subplots) overlap, whereas designs 3 and 7 do not overlap (Table C1). This ambiguous pattern repeats for other separation distance classes, and suggests that number of subplots is a less impactful plot design variable.

3.3. Estimation of $\widehat{MPS}$

As with $\widehat{PA}$, Fig. 4 suggests no or weak relationships between $\widehat{MPS}$ and its components $\hat{p}_{int.}$, $p_{incl.}$, and $\hat{p}$, and this again aligns with WLS regression results (Table 4). The only significant regression model made with these $\widehat{MPS}$ components contained $\hat{p}$ ($p \leq 0.04$, $R^2 = 0.72$) (Table 4). However, $\widehat{MPS}$ generally showed a strong increasing trend with d in both inventory regions (Fig. 6), and WLS results indicate significant regressions when using d as a predictor in both regions ($p \leq 0.05$, $R^2 > 0.69$). Again, as with $\widehat{PA}$, there existed an ambiguous relationship between $\widehat{MPS}$ and number of subplots (Fig. 6, Table C2).

3.4. Sampling errors

Relative standard errors (sampling errors) were higher in region 5 than in region 4, and were generally lower for designs with more subplots and greater separation distance (Fig. 7). In region 4, sampling errors are smaller than in region 5 for a given number of subplots. For a separation distance of 1.1 km, a square cluster shape with four subplots results in the highest sampling errors for both metrics.

4. Discussion

4.1. Summary statistical results for $\widehat{PA}$, $\widehat{MPS}$, $\hat{p}$, $p_{incl.}$, and p_{int}

To begin, results suggest that region 5 has less forest and more forest fragmentation than region 4 (Table 4, Fig. 4). It is known that inventory region 5 has less, more fragmented forest and more agricultural land, and region 4 has a more homogeneous, mostly-forested landscape (Olsson & Ledwith, 2021). Previous studies that assessed the fragmentation status of inventory regions 4 and 5 using the fragmentation metrics contagion (C) and aggregation index (AI) support this conclusion as well (Ramezani & Ramezani, 2015, 2021).

Lack of overlapping confidence intervals support the hypothesis that plot design has an important impact on the outcome of sample-based fragmentation assessments using certain metrics. This aligns with the conclusion that plot design is a crucial factor to consider when interpreting results and that metric values should not necessarily be interpreted as having a direct physical meaning Kleinn (2000).

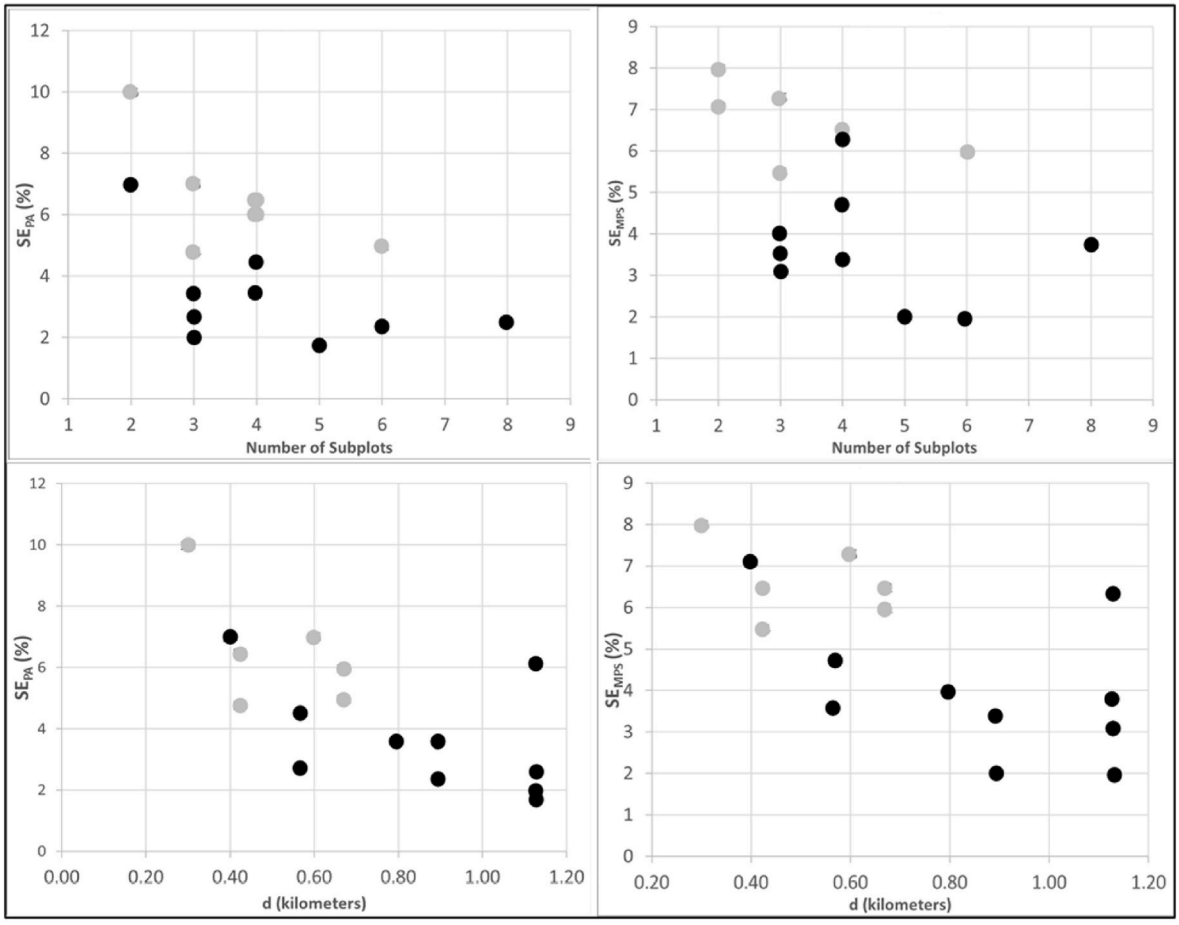


Fig. 6. Relative standard errors (% sampling error) of landscape metrics Perimeter-area ratio (PA, km/km^2) and Mean Patch Size (MPS, km^2) for designs with differing numbers of subplots and maximum separation distance (*d*) scenarios. Grey points are for inventory region 5 and black points are for region 4.

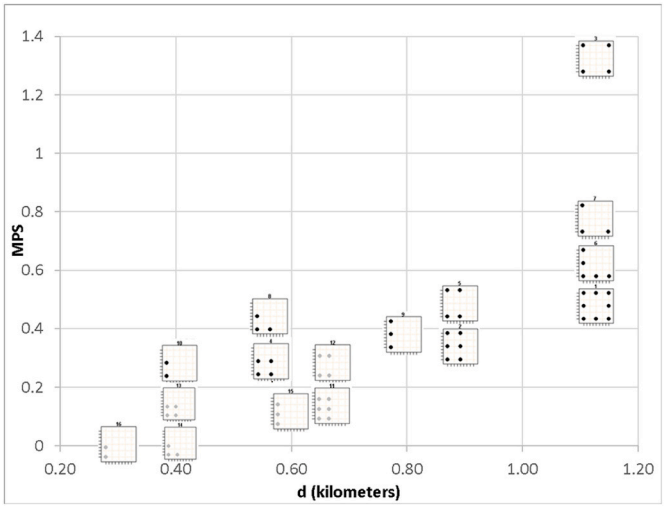


Fig. 7. Relationship between landscape metric Mean Patch Size (MPS, km^2) and maximum subplot center separation distance (*d*) of the corresponding cluster design. Grey points are for inventory region 5 and black points are for region 4.

Our results suggest that the maximum distance between subplots in a design (*d*) is the most impactful plot design factor, particularly in region 4, while the other design parameters (number of subplots and $p_{incl.}$, which is related to shape) are generally less important. Plot design

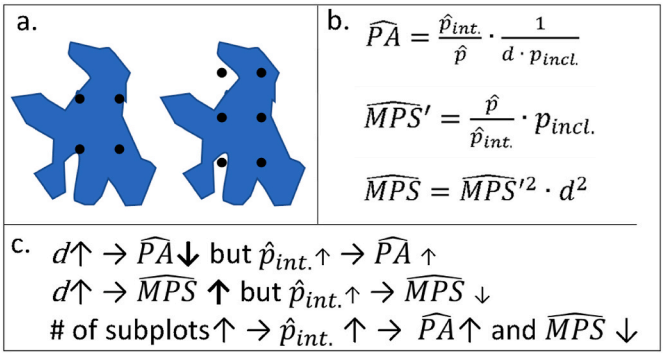


Fig. 8. Conceptual model of the process that affects changes in landscape metrics PA and MPS due to plot design. a) A hypothetical patch, with a 4- and 6-subplot design superimposed. b) Equations 1 through 3. c) Arrow diagram representing the effects of the relationship between *d* and $\hat{p}_{int.}$ on $\widehat{PA}$ and $\widehat{MPS}$. Size and boldness of vertical arrows represent strength of impact or relationship.

factors and metric components can impact metric estimates (Fig. 8). Increasing *d* and the number of subplots in a design should increase the likelihood that the plot will cross a forest/nonforest boundary. In region 5, where forests are sparser and more fragmented, $\hat{p}$ had a significant impact on the metrics, more so than the maximum distance between subplots.

The factors that affect the magnitudes of estimators also affect their variances (Fig. 7). Larger plots with more dispersed subplots capture

more information and have values closer to the sample mean, leading to lower variance. This aligns with previous findings on the impact of plot design on the precision of estimates (Kleinn, 2000; Lister & Leites, 2021a).

4.2. Alignment with results of other studies and future opportunities

Kleinn's (2000) method was also applied by Nelson et al. (2009) to estimate $\widehat{PA}$ and $\widehat{MPS}$ using cluster plots from the United States national forest inventory. As with our study, the authors found that sample-based estimation of metrics has potential to detect fragmentation regions with different spatial patterns of forest. They also found that sample-based results were comparable to those using raster-based metrics, as did Lister et al. (2019).

The expansion of ground-based forest inventories as well as monitoring systems relying on ocular interpretation of high-resolution satellite imagery (e.g., Saah et al., 2019) have created many opportunities to use sample-based tools like those we present to monitor forest fragmentation or degradation, which may be a key contributor to climate change. For example, remeasurement of permanent inventory plots would allow for monitoring in changes of $\widehat{PA}$ and $\widehat{MPS}$, and corresponding inferences about fragmentation dynamics. Traditional approaches to measuring forest fragmentation, such as raster analytics, have their advantages but do not provide estimates with uncertainty metrics that are interpretable through the lens of sampling theory, nor are they necessarily compatible with existing forest inventory plot networks on which other ground data are collected. Our study provides tools and greater understanding of how sample-based methods can improve the science of forest fragmentation and degradation monitoring.

Our study provides insights that can guide decisions on design of sample-based studies and interpretation of results from such studies. For example, the variability of the metrics' values under different plot design scenarios suggests these methods are more suitable for regional comparisons (as we showed with comparisons of region 4 and 5), and potentially for monitoring relative change in fragmentation status on remeasured plots. It also highlights opportunities for how sample-based forest monitoring networks in general might be used to characterize landscape dynamics, and suggests that more research, such as simulation studies, is needed to help determine optimal plot design guidelines under different landscape fragmentation levels and types.

4.3. Caveats

There are several caveats to consider when either designing a sample-based fragmentation monitoring system or adapting existing data for this purpose. First, $\widehat{PA}$ and $\widehat{MPS}$ consist of a ratio between two random variables multiplied by a scalar. When using ratio estimators in a sampling context, there is a small bias introduced, of order 1/n (Cochran, 1977, p. 155); the bias is therefore small with reasonably large sample sizes. The results of Ramezani et al. (2010) agree with this assessment.

Second, acceptable allowable error of estimates (e.g., 10% error at the 95% confidence level), is often a forest inventory design requirement. Although our results suggest that larger plots with more widely-separated subplots will lead to better precision estimates, one must factor in the impacts of cost when designing an inventory. There is a point of diminishing returns from separating subplots and enlarging plots beyond which precision improves little but costs increase greatly, and thus subsampling experiments like ours or factorial simulation experiments can shed light on this important consideration and help planners make better decisions (Lister & Leites, 2021b).

Third, Kleinn (2000) points out that $\widehat{MPS}$ in Eq. (A3) is a metric related to the mean size of patches, not to be interpreted as a direct estimate of mean patch area, and it and $\widehat{PA}$ assume that curvature of patch boundaries does not occur at a scale finer than d. In other words, spatial information that occurs at distances less than d are not observable by this method, and can lead to smaller $\widehat{PA}$ and larger $\widehat{MPS}$ values than expected. In addition, if adjacent patches are less than distance d from each other, and if patch boundaries are not straight lines over distance d, the assumptions underpinning the relevance of $p_{incl.}$ become less valid.

5. Conclusions

In this study, we have revealed how plot design has important effects on estimates of two landscape metrics that are commonly applied to raster data but have been adapted to use with cluster plots. Through a case study that uses data from the Swedish NFI, we found that subplot separation distance has important impacts on metric values. While other factors, such as percent forest on the landscape, also impact metric values, plot shape and number of subplots in the design have a much smaller impact.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author statement

Habib Ramezani: conceptualization, formal analysis, methodology, and writing the manuscript.

Andrew Lister: conceptualization, writing and editing the manuscript.

Acknowledgments

The data from the Swedish NFI were kindly provided by Dr. Cornelia Roberge, the Department of Forest Resources Management (SLU). We are grateful to the thoughtful comments of anonymous reviewers.

Appendix A

Calculation of PA and MPS and corresponding variance estimators

Landscape metrics

The $\widehat{PA}$ and $\widehat{MPS}$ estimation equations (Eq. 1 and 2) rely in part on land cover class assignments made at subplot centers by field crews. These assignments allow for the calculation of two measures needed to calculate $\widehat{PA}$ and $\widehat{MPS}$: $\hat{p}$, the proportion of subplot centers in the population that are labeled as forest, and $\hat{p}_{int.}$, the proportion of cluster plots that cross forest boundaries. This method is based on a virtual buffer with a fixed width d ($d/2$ on either side of the boundary), which is assumed on both sides of the forest boundary. Buffer width d corresponds to the diagonal length of the cluster plot (the maximum distance between two subplot centers for a given plot design). See Appendix B for an example of the calculation procedure for $\hat{p}_{int.}$.

Perimeter-area ratio ($\widehat{PA}$)

Because the NFI uses subplots separated in space, it is impossible to estimate $\widehat{PA}$ for individual forest patches. As an alternative, the ratio can be estimated at the landscape level with information from all sampled cluster plots. According to Kleinn (2000), a landscape level estimator of $\widehat{PA}$ is defined as

$$\widehat{PA} = \frac{\widehat{p}_{int.}}{\widehat{p}} \bullet \frac{1}{d \bullet p_{incl.}}, \tag{A1}$$

where $\widehat{p}_{int.}$ is the proportion of cluster plots that the cross forest patch boundaries (see Appendix B), $\widehat{p}$ is the proportion of subplot centers that fall in forestland (i.e., the estimated proportion forest), d is the virtual buffer width (km), and $p_{incl.}$ is the conditional inclusion probability of an edge intersecting a buffer of a specified shape.

The inclusion probability is a function of cluster shape; in other words, regardless of cluster size, as long as the shape remains in the same proportions, $p_{incl.}$ remains the same. However, buffer width (d) is a function of both cluster size and shape because it corresponds to the diameter of the circumcircle of the cluster plots (the longest distance between two points on a certain geometric shape). For instance, for a given rectangular shape (Fig. A1), d can be calculated by $(2 \bullet (a+b))/(\pi \bullet \sqrt{a^2+b^2})$ where a and b are the sides of a rectangular cluster. Table A1 gives values for $p_{incl.}$ for several common shapes, including those used in this study.

Mean patch size ($\widehat{MPS}$)

$\widehat{MPS}$ is based on a relation between estimates of forest patch area and perimeter. One advantage of this metric is that its calculation is not affected by patch shape. The estimator of $\widehat{MPS}$ is defined by Kleinn (2000) as

$$\widehat{MPS}' = \frac{\widehat{p}}{\widehat{p}_{int.}} \bullet p_{incl.}. \tag{A2}$$

Since $\widehat{MPS}'$ depends on buffer width (cluster size) through the $p_{incl.}$ term, it is impossible to compare inventory regions with different cluster configurations. To overcome this problem, Kleinn (2000) suggested a standardized $\widehat{MPS}$ metric that is independent of the size of the clusters. A standardized $\widehat{MPS}$ metric can be estimated by incorporating information on the buffer width d:

$$\widehat{MPS} = \widehat{MPS}'^{2} \bullet d^2. \tag{A3}$$

Variance estimation

In the present study, a conventional variance estimator was used to estimate variance of $\widehat{PA}$ and $\widehat{MPS}$. In Eqs. A1 and A2 there are two random variables, $\widehat{p}$ (the estimated proportion of forest land) an $\widehat{p}_{int.}$ (the proportion of clusters that partially intersect forest patch boundaries) in the form of a ratio estimator. According to Thompson (2002, p. 70) the estimator of variance of an estimate of a ratio is

$$\widehat{v}(\widehat{R}) = \frac{1}{n \bullet \underline{x}^2} \bullet \frac{\sum_{i=1}^{n} (y_i - \widehat{R}x_i)^2}{(n-1)}. \tag{A4}$$

Let $\widehat{p}_{int.} = \frac{1}{n}\sum_{i=1}^{n} y_i$, where $= 1$ if cluster plot i crosses a forest border and $= 0$ otherwise. Let $\widehat{p} = \frac{1}{n}\sum_{i=1}^{n} x_i$, where $= 1$ if subplot i is in forest land and $= 0$ otherwise. Then $\widehat{R} = \frac{\sum_{i=1}^{n} y_i}{\sum_{1=1}^{n} x_i}$.

Thus, the variance estimator of $\widehat{PA}$ is

$$\widehat{v}(\widehat{PA}) = \widehat{v}(\widehat{R}) \left(\frac{1}{d^2 \bullet p_{incl.}^2} \right). \tag{A5}$$

The variance of $\widehat{MPS}$, where $\widehat{R}_{MPS} = \frac{\widehat{p}}{\widehat{p}_{int.}}$, is estimated as

$$\widehat{v}(\widehat{MPS}) = \widehat{v}(\widehat{R}_{MPS} \bullet p_{incl.}) = p_{incl.}^2 \widehat{v}(\widehat{R}_{MPS}). \tag{A6}$$

Relative sampling error (SE %) for $\widehat{PA}$ and $\widehat{MPS}$ is calculated as

$$\frac{\sqrt{\widehat{v}(\widehat{PA})}}{\widehat{PA}} \bullet 100 \tag{A7}$$

and

$$\frac{\sqrt{\widehat{v}(\widehat{MPS})}}{\widehat{MPS}} \bullet 100, \tag{A8}$$

respectively.

Appendix B

Calculation of $\widehat{p}_{int.}$

For example, assume that the study area contains 50 square cluster plots with eight subplots each, like for cluster plot design number 1 in Fig. 2. There are three possible positions of the cluster plot relative to a forest patch boundary that are found in this dataset: 1) seven cluster plots have all eight subplot centers in forest, 2) thirteen cluster plots have all eight subplot centers in non-forest, and 3) thirty cluster plots have some of their eight subplots in forest and some in nonforest. That means that thirty cluster plots cross the forest boundary, as determined by a virtual line drawn that connects all subplots. Thus, $\widehat{p}_{int.}$ can easily be calculated as

$$\widehat{p}_{int.} = \frac{\#\ of\ cluster\ plots\ that\ cross\ the\ forest\ boundary}{total\ number\ of\ cluster\ plots} = \frac{30}{50} = 0.6$$

Appendix C

Table A1
The inclusion probability of intersection ($p_{incl.}$) for different geometric shapes, based on Kleinn (2000).

Geometrics shapes	Inclusion probability
Circle	1
Square	0.9003
Rectangle	0.8545
Triangle	0.8269
L-shape	0.7685
Line	0.6366

Table C1
Pairwise comparison of the $\widehat{PA}$ 95% confidence interval (CI) overlap of subplot designs 1–16. Sampling errors were converted to 95% CIs (SE% × $\widehat{PA}$ × 1.96/100). O indicates overlap, NO indicates no overlap. For example, the CI for design 1overlaps with itself and that of designs 5 and 6 (designs are described in Fig. 2 and Table 2). Bolded region (designs 11–16) are region 5, non-bold design (1–10) are from region 4.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	O	NO	NO	NO	O	O	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
2		O	NO	NO	O	NO	NO	O	O	NO	NO	NO	NO	NO	NO	NO
3			O	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
4				O	NO	NO	NO	NO	NO	O	O	O	NO	NO	O	NO
5					O	NO	NO	NO	O	NO	NO	NO	NO	NO	NO	NO
6						O	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
7							O	NO	NO	NO	NO	NO	NO	NO	NO	NO
8								O	O	O	NO	O	NO	NO	NO	NO
9									O	NO	NO	NO	NO	NO	NO	NO
10										O	O	O	NO	NO	O	O
11											**O**	**O**	**O**	**O**	**O**	**O**
12												**O**	**NO**	**NO**	**O**	**NO**
13													**O**	**O**	**O**	**O**
14														**O**	**O**	**O**
15															**O**	**O**
16																**O**

Table C2
Pairwise comparison of the $\widehat{MPS}$ confidence interval (CI) overlap of subplot designs 1–16. Sampling errors were converted to 95% CIs (SE% × $\widehat{MPS}$ × 1.96/100). O indicates overlap, NO indicates no overlap. For example, the CI for design 1 overlaps with itself and that of design 6 (designs are described in Fig. 2 and Table 2). Bolded region (designs 11–16) are region 5, non-bold design (1–10) are from region 4.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	O	NO	NO	NO	NO	O	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
2		O	NO	NO	NO	NO	NO	NO	O	NO	NO	NO	NO	NO	NO	NO
3			O	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
4				O	NO	NO	NO	NO	NO	O	NO	O	NO	NO	NO	NO
5					O	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
6						O	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
7							O	NO	NO	NO	NO	NO	NO	NO	NO	NO
8								O	NO	NO	NO	NO	NO	NO	NO	NO
9									O	NO	NO	NO	NO	NO	NO	NO
10										O	NO	O	NO	NO	NO	NO

(*continued on next page*)

Table C2 (*continued*)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
11											**O**	**NO**	**NO**	**NO**	**O**	**NO**
12												**O**	**NO**	**NO**	**NO**	**NO**
13													**O**	**O**	**NO**	**O**
14														**O**	**O**	**O**
15															**O**	**O**
16																**O**

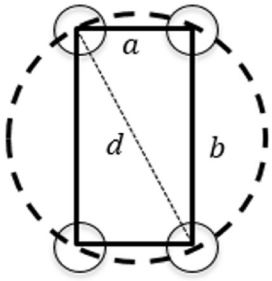


Fig. A1. An example of rectangular cluster plot, where *d* (the maximum distance between two subplot centers) is the diameter of the circumcircle of a rectangle.

References

Axelsson, A.-L., Ståhl, G., Södererg, U., Peterson, H., & Fridman, J. (2009). Development of Sweden's national forest inventory. In E. Tomppo, & T. Gschwantner (Eds.), *National forest inventories: Pathways for common reporting* (pp. 541–553). Heidelberg: Springer.

Cochran, W. (1977). *Sampling techniques*. New York: John Wiley & Sons.

Corona, P., Fattorini, L., & Franceschi, S. (2019). Estimating tree diversity in forest ecosystems by two-phase inventories. *Environmetrics, 30*(2), Article e2502. https://doi.org/10.1002/env.2502

Draper, N., & Smith, H. (1981). *Applied regression analysis* (2nd ed.). New York: John Wiley & Sons.

Fang, S.-F., Gertner, G., Wang, G. X., & Anderson, A. (2006). The impact of misclassification in land use maps in the prediction of landscape dynamics. *Landscape Ecology, 21*, 233–242.

Fridman, J., Holm, S., Nilsson, M., Nilsson, P., Ringvall, A.-H., & Ståhl, G. (2014). Adapting national forest inventories to changing requirements – the case of the Swedish National Forest Inventory at the turn of the 20th century. *Silva Fennica, 48*, 1–29.

Grantham, H.-S., Duncan, A., & Evans, T.-D. (2020). Anthropogenic modification of forests means only 40% of remaining forests have high ecosystem integrity. *Nature Communications, 11*(1), 5978. https://doi.org/10.1038/s41467-020-19493-3

Hassett, E.-M., Stehman, S.-V., & Wickham, J.-D. (2011). Estimating landscape pattern metrics from a sample of land cover. *Landscape Ecology, 27*, 133–149.

Kleinn, C. (2000). Estimating metrics of forest spatial pattern from large area forest inventory cluster samples. *Forest Science, 46*, 548–557.

Lister, A.-J., & Leites, L.-P. (2021a). Designing plots for precise estimation of forest attributes in landscapes and forests of varying heterogeneity. *Canadian Journal of Forest Research, 51*(10), 1569–1578. https://doi.org/10.1139/cjfr-2020-0508

Lister, A.-J., & Leites, L.-P. (2021b). Cost implications of cluster plot design choices for precise estimation of forest attributes in landscapes and forests of varying heterogeneity. *Canadian Journal of Forest Research, 52*(2). https://doi.org/10.1139/cjfr-2020-0509, 188-00.

Lister, A., Lister, T., & Weber, T. (2019). Semi-automated sample-based forest degradation monitoring with photointerpretation of high-resolution imagery. *Forests, 10*, 896.

McGarigal, K., & Marks, B.-K. (1995). Fragstats: Spatial pattern analysis program for quantifying landscape structure. In *Gen. Tech. Rep. PNW-GTR-351*. Pacific Northwest Research Station, Portland, OR: U.S. Department of Agriculture, Forest Service.

Nelson, M.-D., Lister, A.-J., & Hansen, M.-H. (2009). Estimating number and size of forest patches from FIA plot data. In R. E. McRoberts, G. A. Reams, P. C. Van Deusen, & W. H. McWilliams (Eds.), *Proceedings of the eighth annual forest inventory and analysis symposium; 2006 October 16-19; Monterey, CA. Gen. Tech. Rep. WO-79* (pp. 159–164). Washington, DC: U.S. Department of Agriculture, Forest Service.

Olsson, B., & Ledwith, M. (2021). *National land cover database, version1.0, 2020-08-26*. Stockholm, Sweden: The Swedish Environmental Protection Agency. https://gpt.vicmetria.nu/data/land/NMD/NMD2018_ProductDescription_ENG.pdf. (Accessed 1 January 2022).

Pyke, C. (2004). Habitat loss confounds climate change impacts. *Frontiers in Ecology and the Environment, 2*(4), 178–182.

R Core Team. (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available from: version 4.0.5. https://www.R-project.org/. (Accessed 1 January 2022).

Ramezani, H., Holm, S., Allard, A., & Ståhl, G. (2010). Monitoring landscape metrics by point sampling: Accuracy in estimating Shannon's diversity and edge density. *Environmental Monitoring and Assessment, 164*(1), 403–421.

Ramezani, H., Holm, S., Allard, A., & Ståhl, G. (2013). A review of sampling-based approaches for estimating landscape metrics. *Norwegian Journal of Geography, 67*(2), 61–71.

Ramezani, H., & Ramezani, F. (2015). Potential for the wider application of national forest inventories to estimate the contagion metric for landscapes. *Environmental Monitoring and Assessment, 187*(3), 187.

Ramezani, H., & Ramezani, A. (2021). Forest fragmentation assessment using field-based sampling data from forest inventories. *Scandinavian Journal of Forest Research, 36*(4), 289–296. https://doi.org/10.1080/02827581.2021.1908592

Ramírez, C., Alberdi, I., Bahamondez, C., & Freitas, J. (Eds.). (2022). *National forest inventories of Latin America and the Caribbean – towards the harmonization of forest information*. Rome: FAO. https://doi.org/10.4060/cb7791en.

Saah, D., Johnson, G., & Ashmall, B. (2019). Collect Earth: An online tool for systematic reference data collection in land cover and use applications. *Environmental Modelling & Software, 118*, 166–171. https://doi.org/10.1016/j.envsoft.2019.05.004

Saura, S., & Carballal, P. (2004). Discrimination of native and exotic forest patterns through shape irregularity indices: An analysis in the landscapes of Galicia, Spain. *Landscape Ecology, 19*(6), 647–662.

Shapiro, A.-C., Aguilar-Amuchastegui, N., Hostert, P., & Bastin, J.-F. (2016). Using fragmentation to assess degradation of forest design Democratic Republic of Congo. *Carbon Balance and Management, 11*(1), 1–15.

Thompson, S.-K. (2002). *Sampling*. New York: John Wiley & Sons.

Tolentino, M., & Anciães, M. (2020). Effects of forest fragmentation on the lekking behavior of White-throated Manakins in Central Amazonia. *Journal of Field Ornithology, 91*(1), 31–43.

Tomppo, E., Gschwantner, T., Lawrence, M., & McRoberts, R.-E. (Eds.). (2010). *National forest inventories: Pathways for common reporting*. New York, NY: Springer.

Ye, S., Pontius, R.-G., & Rakshit, R. (2018). A review of accuracy assessment for object-based image analysis: From per-pixel to per-polygon approaches. *International Society for Photogrammetry and Remote Sensing, 141*, 137–147.