Identification of de facto protected areas in boreal Canada.
Biological Conservation, 146 (1). pp. 97-107.

http://researchrepository.murdoch.edu.au/23062/

Copyright: © 2011 Elsevier B.V.
It is posted here for your personal use. No further distribution is permitted.
Identification of *de facto* protected areas in boreal Canada

Margaret E. Andrew\(^1\^, Michael A. Wulder\(^1\), Nicholas C. Coops\(^2\)

Author Affiliations:

\(^1\)Canadian Forest Service (Pacific Forestry Centre), Natural Resources Canada, 506 West Burnside Road, Victoria, British Columbia, V8Z 1M5, Canada

\(^2\)Department of Forest Resource Management, 2424 Main Mall, University of British Columbia, Vancouver, British Columbia, V6T 1Z4, Canada

\(^\text{^1}\)Corresponding author:
Margar et Andrew
Canadian Forest Service, Pacific Forestry Centre
506 West Burnside Rd., Victoria, BC V8Z 1M5;
Phone: 250-298-2428; Fax: 250-363-0775; Email: margaret.andrew@nrcan-rmc an.gc.ca

**Pre-print of published version.**

**Reference:**

**DOI:**
[http://dx.doi.org/10.1016/j.biocon.2011.11.029](http://dx.doi.org/10.1016/j.biocon.2011.11.029)

**Disclaimer:**
The PDF document is a copy of the final version of this manuscript that was subsequently accepted by the journal for publication. The paper has been through peer review, but it has not been subject to any additional copy-editing or journal specific formatting (so will look different from the final version of record, which may be accessed following the DOI above depending on your access situation).

**Keywords:** forest fragmentation, frontier forests, Landsat, landscape configuration, maximum entropy (Maxent), MODIS, wilderness.
Abstract. Canada is dominated by remote wilderness areas that make important conservation contributions, but are currently only protected de facto by their inaccessibility. Mechanisms for the identification and formal protection of such areas can help ensure that they continue to function naturally and provide essential ecosystem services. However, a lack of spatially explicit, publicly available sources of data on anthropogenic disturbances and natural resource extraction challenges the development of detailed wilderness inventories. We suggest that landscape structure can be used to classify areas of natural landscapes, as trained by the landscape structure of protected areas, and demonstrate this approach by mapping de facto protected areas in Canada’s boreal forest. Overall, between 50%, based on landscape structure, and 80%, based on anthropogenic infrastructure alone, of Canada’s boreal zone exists in large, intact blocks. The true extent of boreal wilderness likely falls within this range, as existing infrastructure datasets may omit disturbance and the protected area network in far northern areas proved inadequate to train effective wilderness classifications. We anticipate that such efforts may be improved by refining the identification of training areas or by classifying along additional landscape metrics. Nevertheless, the areas identified are valuable candidates for protected area expansion, and can contribute to a reserve network that meets national and regional conservation targets and is representative of the range of vegetation productivities, which was used as a biodiversity surrogate. Our general approach need not be limited to the boreal forest, as it has the potential to successfully identify relatively undisturbed (or less disturbed) areas over a range of systems and across levels of human influence.

1. Introduction

The current coverage of Canada’s protected area system (encompassing 9.6% of its terrestrial area; CARTS, 2011) falls below the global average (12.7% of the world’s land area protected as of 2010; IUCN and UNEP, 2010). This finding is not remarkable for boreal forests, however, which make up a large portion of the Canadian landmass and are considered underprotected globally (Mittermeier et al., 2003; Schmitt et al., 2009). There are a number of considerations for protected area expansion to achieve adequate coverage of boreal forests and other underprotected systems. Most academic reserve network design exercises focus on the principle of representation, identifying a set of areas that will provide protection to all species, communities, ecosystems, or ecological regions that are present within the planning area (Margules and Pressey, 2000). However, such analyses might select areas that are extremely unrealistic for effective conservation. For example, area-based optimization algorithms may select sites that have purchase prices that are well beyond the entire budget for conservation efforts (Ando et al., 1998; Moilanen and Arponen, 2011). Likewise, they may select sites that are so transformed or that contain so much pre-existing infrastructure as to be impractical or impossible to restore and protect, or that are simply unattainable due to patterns of property ownership. In practice, conservation network design must balance representation goals with the realities of the landscape (e.g., Hoctor et al., 2000).

The current area of formal protection in the Canadian boreal is complemented by the vast expanses of wilderness that remain (Bryant et al., 1997; Lee et al., 2010; McCluskey and Spalding, 1989; Mittermeier et al., 2003; Nogueiral, 2002; Potapov et al., 2008; Sanderson et al., 2002), continuing to support a portfolio of natural processes and providing opportunities for
protected area expansion that may be relatively cost-effective to conserve (Mittermeier et al., 2003). These areas can be considered to enjoy de facto protection (Joppa et al., 2008), largely due to harsh climates and inaccessibility (Bryant et al., 1997), but are increasingly threatened by climate change and expanding natural resource exploitation (Schindler and Lee, 2010).

Developing large-area wilderness inventories, such as those cited above, presents many challenges, not the least of which is a general lack of data on local anthropogenic disturbances, such as secondary roads, resource extraction and related infrastructure (Canadian Boreal Initiative, 2005). Resulting wilderness maps are either very coarse (e.g., Noguerón, 2002; Sanderson et al, 2002) or rely upon manual image interpretation to identify disturbances (e.g., Lee et al., 2010; Potapov et al., 2008), a subjective process that is time- and labor-intensive (e.g., two years for the forested area of Canada, Lee et al., 2006). Relatively rapid, objective, automated wilderness classifications may now be possible, supported by unprecedented datasets of fine-scale landscape structure over large extents (e.g., Heilman et al., 2002; Wulder et al., 2008a).

A large number of fragmentation metrics have been developed based on the size, shape, and pattern of habitat patches in a landscape (e.g., McGarigal and Marks, 1995), and these are capable of distinguishing between natural and anthropogenic landscape structures. For example, boreal forest landscapes are naturally patchy (Wade et al., 2003; Wulder et al. 2011), with mosaics of wetlands, open water, and forests of different types and age classes (Cumming et al., 1996; Esseen et al., 1997; Leroux et al., 2007; Sjöberg and Ericson, 1997). Disturbed forest landscapes containing cutblocks and linear elements such as roads or seismic lines are quite distinct in both pattern (Franklin and Formann, 1987; Hansson and Larsson, 1997; Mladenoff et al., 1993; Reed et al., 1996a; Ripple et al., 1991) and scale (Johnson et al., 1998). (However, note that in areas with extensive fire suppression, forest harvest may be the main driver of landscape variation in forest composition and age structure, contributing beneficially to the landscape mosaic (e.g., Bergeron et al., 2004).) Natural landscape structure is a criterion of wilderness (Aplet et al., 2000), but is rarely explicitly considered in wilderness classifications. However, fragmentation metrics have been used to identify different landscape types (Cardille and Lambois, 2010; Leimgruber et al., 2003), suggesting that they may also be useful for wilderness classification.

Existing protected areas offer the most logical set of landscapes with which to train an automated wilderness classification. Protected areas are often successful at reducing habitat conversion and resource exploitation (Asner et al., 2006; Bruner et al., 2001; Joppa et al., 2008; Nagendra, 2008; Oliviera et al., 2007), and thus forestal alterations to landscape structure. Protected areas also have fewer signs of conspicuous human presence or access (Cardille and Lambois, 2010; Lee and Cheng, 2011; Leu et al., 2008), and often correspond to large remnants of intact habitat (Goetz et al., 2009; Heilman et al., 2002; Alberta case study in Lee et al., 2006). Although there are important exceptions (Soverel et al., 2010; Timoney, 1996), in the aggregate, protected areas are relatively unimpaired and can function as valuable benchmarks (Arcese and Sinclair, 1997; Sinclair, 1998; Wiersma, 2005) against which to evaluate landscape structure.

The goal of this study is to describe a new approach to identify de facto protected areas - those landscapes that are relatively free of human disturbance - and apply it over the Canadian boreal forest using datasets documenting forest fragmentation and anthropogenic infrastructure. Further, we evaluate the ability of these areas to contribute to overall protected area system goals.
2. Methods

2.1. Study area

This study was conducted over the full extent of the Canadian boreal forest, as delineated by Brandt (2009). This product demarcates the boundaries of the North American boreal biome, including altitudinal tree lines, but includes naturally treeless areas within these bounds. The boreal zone covers over 5,500 km² (55%) of Canada. Boreal landscapes are complex mosaics of wetlands, open water, and (primarily coniferous) forests in a diverse patchwork of stand ages and sizes, maintained by fire and other disturbances (Cumming et al., 1996; Esseen et al., 1997; Leroux et al., 2007; Sjöberg and Ericson, 1997).

2.2. Datasets

Information on boreal forest landscape structure in Canada was supplied by the national forest fragmentation dataset (Wulder et al., 2008a), which includes a wide variety of forest fragmentation metrics calculated over several measurement extents from a 25 m resolution binary forest/non-forest classification derived from a land cover product of classified Landsat image data (Wulder et al., 2008b). However, many of these variables are strongly intercorrelated, and abstracted from ecological processes. Thus, the present study selected six fragmentation metrics that contain independent information content and exhibit known, intuitive responses to anthropogenic landscape disturbances, all estimated over 1 km extents: the proportion of forest land cover, both by pixel (farea) and by patch (fprop); the number of forest patches (fpatch); the mean (farea) and standard deviation (fsarea) of forest patch size; and the density of forest edges (fdense).

Human access was indicated by the distance to roads and settlements, as computed over a 1 km resolution grid by Wulder et al. (2011) from the national road network (Statistics Canada, 2008) and the Defense Meteorological Survey Program Nighttime Lights dataset (NOAA, 2000). The latter dataset is sensitive to any stable anthropogenic light source, including not only settlements, but also industrial, mining, and oil/gas developments (Elvidge et al., 1997).

Protected area boundaries within the boreal zone were taken from polygon features in the World Database on Protected Areas (Fig. 1; IUCN and UNEP, 2009) and rasterized to the 1 km grid of the fragmentation and access data. Overall, 6.8% of Canada’s boreal forest was protected in 878 reserves, representing 41% of Canada’s protected area system, by area. The Canadian ecozone framework (Fig. 1; Ecological Stratification Working Group, 1995) was used as a spatial stratification. Following Kull et al. (2006), the Boreal Shield and Taiga Shield ecozones were split into eastern and western units, to better represent patterns of climate and forest processes.

The distribution of protected areas and wilderness across the boreal forest was evaluated against elevation (Shuttle Radar Topography Mission; Rabus et al., 2003) and patterns of vegetation productivity, represented as the minimum annual productivity, integrated annual productivity, and seasonality (coefficient of variation of productivity), using the dynamic habitat index framework of Coops et al. (2008). The dynamic habitat index is calculated from monthly estimates of the fraction of absorbed photosynthetically active radiation (fPAR) at 1 km resolution from the satellite-borne MODerate-resolution Imaging Spectrometer (MODIS) sensor.

2.3. Wilderness classifications

2.3.1. Low-access areas

As a preliminary analysis and point of comparison, wilderness areas were first mapped by simply excluding all areas within 1 km of a road or settlement. Roads are often cited as a widely available indicator of anthropogenic disturbance, and have wide-ranging ecological effects
Together with settlements, roads increase human access, providing the potential for increased resource extraction in a given area (Asner et al., 2006; Laurance et al., 2002; 2006; Wilkie et al., 2000). Roads have been found to be the most important component of the human footprint to habitat fragmentation (Hawbaker et al., 2006; Reed et al., 1996b) and some wildlife responses (Blake et al., 2008).

Low-access areas were further grouped on the basis of unit size. Two minimum size thresholds were chosen: 540 km² and 3,000 km². The first is the mean size of existing protected areas in Canada’s boreal forest; the second is the minimum area recommended to maintain wide-ranging mammal populations and landscape dynamics (Wiersma et al., 2005).

This rule-based approach is akin to many existing wilderness classifications (McCloskey and Spalding, 1989; Noguerón, 2002; Sanderson et al, 2002), and is likely to overestimate the area of true wilderness as not all anthropogenic disturbances are considered (e.g., forest harvest, oil exploration), although they might be reasonably proxied by the road network. However, it is also important to consider the converse: that roads provide the potential for disturbance, rather than actual disturbance, and do not necessarily preclude conservation. Much of our understanding of roads as landscape drivers comes from the tropics, and may be less relevant in the boreal. Further, current patterns are not necessarily permanent and roads may be decommissioned.

2.3.2. Maximum entropy classification of de facto protected areas

To refine the above wilderness classification based solely on readily available infrastructure datasets, we conducted an automated maximum entropy (Maxent) classification of forest fragmentation to identify areas of “natural” forest landscape structure. Maxent is a presence-only machine learning algorithm that has recently been applied to species distribution modeling (Phillips et al., 2006), typically out-performing alternative algorithms (Elith et al., 2006). Maxent models the geographic distribution of a target (typically species, but in this study de facto protected areas) such that the statistical distribution of environmental predictor variables from the modeled area is as close as possible to that of the study region overall, but constrained by the variable means of the training set (Elith et al., 2011). Maxent considers not only the raw environmental variables, but also higher order and composite products, in order to represent nonlinearities and interactions (Phillips et al., 2006; Phillips and Dudík, 2008).

Maxent offers a number of advantages to wilderness classifications over standard expert- and rule-based approaches: (1) It is automated, objective, and repeatable; (2) Maxent’s measures of variable importance provide information about the driving variables; (3) the quantitative entropy output indicates degrees of wilderness character, rather than simply a binary wilderness/not-wilderness classification result; and (4) similarly, it facilitates the mapping of different types of wilderness, for example, corresponding to the different types of protected areas as classified by the International Union for Conservation of Nature (IUCN; Dudley, 2008).

We performed Maxent classifications on the six forest fragmentation metrics, trained with the existing protected areas. All protected pixels within 1 km of a road, settlement, or the protected area boundary were excluded, to minimize anthropogenic disturbances within the training set. To reduce confounding variation due to regional differences in forest structure (Kneeshaw and Gauthier, 2003) or different types of landscapes under different types of protection (Dudley, 2008), separate classifications were performed for each ecozone and IUCN protected area category. Note that not all combinations of ecozones and IUCN categories exist. Due to computational limitations, both the training and background sets were subsampled for
each classification. Protected areas were split into two sets, and, within a chosen set, stratified by site, with 1, 10, and 100 pixels randomly sampled from reserves < 100 km², < 1000 km², and ≥ 1000 km² in area, respectively. Background samples were a random sample of 5% (up to 25,000) of the unprotected pixels in an ecozone. To overcome variation resulting from the random sampling, 20 Maxent models were run for each ecozone and averaged. The two protected area splits were sampled to train 10 iterations each.

The aggregated Maxent models for each IUCN category in each ecozone were evaluated with the area under the receiver operating characteristic curve (AUC), computed across all pixels in the ecozone. Models with AUC < 0.6 were discarded. Continuous Maxent outputs were converted to a binary classification with the threshold that maximized the sum of sensitivity and specificity, a recommended approach to threshold selection (Liu et al., 2005). Variable importance was evaluated by the percentage that each variable contributed to individual models.

Classification outputs were aggregated by IUCN category into strictly protected (categories Ia-IV) and all protected sets and ecozone-level results were stitched together into a national product. A risk of pixel-level classifications such as this one is that the outputs may be patchy and therefore not ideal for prioritizing new protected areas, which are not selected or managed at the level of individual 1 km pixels. Classification outputs might be refined by a tool such as PatchMorph (Girvetz and Greco, 2007), which smooths pixel outputs into patches that are more ecologically relevant, by applying density filters, filling internal holes, and excluding habitat spurs. However, as our results were relatively continuous, it was decided not to further refine the raw classification maps. As with the low-access analyses, de facto protected areas were grouped by size into areas larger than 540 km² and 3,000 km².

2.4. Representation and bias of de facto protected areas

Previously, we have suggested that remotely sensed vegetation productivity might serve as a valuable biodiversity surrogate (Andrew et al., 2011a) and be used to evaluate protected area system coverage (Andrew et al., 2011b). Consequently, we estimated the bias and representation of de facto protected areas along each of the three productivity axes, and elevation, to determine their ability to contribute to a representative reserve network. Representation was calculated for 100 discrete, equal-width intervals of each productivity and elevation axis, and is the proportion under protection by area of each level. Consequently, representation is extremely information rich (e.g., in the present study it includes 4 variables * 100 levels = 400 values), but, as a result, it may be difficult to quickly interpret or to distill into a general assessment of the distribution of protection. Bias, on the other hand, is a single value that rapidly conveys if there is systematic unevenness in the distribution of protected areas along a quantitative environmental gradient. For example, protected areas worldwide are reputed to be biased to high elevations and low productivities (i.e., there is relatively greater protection at high than low elevations, and the reverse is true for productivities; Joppa and Pfaff, 2009). Bias was calculated as the difference between the median of an environmental variable in the de facto protected area network and its median throughout the study extent, standardized as a proportion of the variable’s range in the study extent (Pressey et al., 2000).

3. Results
3.1. Low-access areas

A large proportion of Canada’s boreal forest is distant from human infrastructure (Fig. 2, Table 1). Nearly all of this low-access area occurs in blocks larger than 3,000 km². In addition, most of the boreal ecozones are largely remote, with the exception of boreal portions of the
Atlantic Maritime ecozone, none of which occurs in large blocks; and the Boreal Plains, half of which is fragmented by roads and settlements (Table 1).

3.2. De facto protected areas

The low-access areas were substantially reduced when constrained to locations of natural landscape structure, as defined by the forest fragmentation characteristics of existing protected areas (Fig. 3, Table 1). Overall, about half of the Canadian boreal forest occurs in large blocks with landscape structure consistent with that of protected areas. This drops to 37% when basing the landscape structure classification only on strictly protected areas. The distribution of de facto protected areas is also highly variable between ecozones. The landscape structure of most protected area types in most ecozones could be modeled successfully by Maxent classifications; only 4 out of 40 models failed to meet the 0.6 AUC criterion (Table 2).

The proportion of a pixel that was forested was generally the most important variable for differentiating the landscape structure of protected areas (Table 3). Mean forest patch size tended to be more important in mountainous ecozones, and forest edge density was of greater importance in the Taiga Plains and Taiga Shields ecozones (Table 3). Patterns were similar when looking at variable importance across protected area categories (not shown): mean forest patch size was most important for classes III and V, both of which are relatively rare in the Canadian boreal; forest edge density was most important for protected areas with unknown designations; and total forest cover was most important for all other categories.

3.3. Representation and bias of de facto protected areas

The statistical distributions of productivity and elevation variables within de facto protected areas generally paralleled those of boreal Canada overall (Fig. 4a), indicating little bias in protection along these metrics. However, biases were greater for individual ecozones (Table 4). In general, de facto protected areas were less environmentally biased than existing protected areas, especially when combining all protected area categories. Low-access areas were unbiased regardless of spatial extent (Table 4).

There was also relatively high and consistent representation of productivities and elevations in all of the wilderness classifications (Fig. 4b), which agrees with the overall area classified (Table 1) and low biases (Table 4). Representation levels were lowest for rare conditions such as high minimum annual productivities and elevations; the highest integrated annual productivities; and both extremes of seasonality. There was conspicuously low representation of moderately high seasonalties (seasonality ≈ 1.3; Fig. 4b), corresponding to regions along the northern and altitudinal margins of the boreal zone with few areas classified as de facto protected.

4. Discussion

Our analyses have identified candidate areas within which protection might be practical (Figs. 2, 3). Although not considered here, additional insight might be gleaned from the continuous Maxent results. Wilderness is not a binary state (Lesslie and Taylor, 1985), but existing approaches to delineate wilderness do not have the capability to indicate degrees of wilderness. Further, we have demonstrated that these areas can make important contributions to reserve networks at both regional and national scales, increasing representation of biodiversity surrogates and improving the distribution (reducing the bias) of protection across environmental gradients (Fig. 4; Table 4). If this information is used strategically, the bias of a future protected area system may be reduced below that of the de facto protected areas in which candidate protected areas are delineated. This is because the bias of an actual protected area network will
only correspond to that of the de facto set if environmental conditions are protected proportionate to their availability in pristine areas. In contrast, if conditions are protected proportionate to their availability in the planning region as a whole (but still selecting pristine areas to protect), bias can be eliminated. This will be possible if the area of de facto protection for each level of the biodiversity surrogates meets or exceeds the representation target for the planning region. Areal conservation targets promoted for boreal Canada include 12% (Environment Canada, 2006), 20% (Canadian Boreal Initiative, 2003), and as high as 50% (Canadian Boreal Initiative, 2005). Additionally, the Convention on Biological Diversity’s protected area target has recently been raised to 17% of the world’s terrestrial area by 2020 (UNEP, 2010). Many of these goals can be met by the identified de facto protected areas, both overall and for most individual biodiversity surrogates assessed (ecozones, Table 1; and productivity levels, Fig. 4b). In fact, for many of the surrogates the de facto protected area is much greater than the target, indicating considerable flexibility to meet conservation goals in boreal Canada, an encouraging situation that is likely relatively unique globally, given patterns of wilderness loss (Sanderson et al., 2002). Similarly, evaluations of the United States Forest Service’s roadless holdings have determined that, should these areas be granted strict protection, they would make substantial contributions to the conservation of underprotected elevation ranges, habitat types, ecoregions, and threatened species, as well as increase protected area network connectivity (Crist et al., 2005; DeVelice and Martin, 2001; Loucks et al., 2003; Strittholt and DellaSala, 2001).

From the opposite perspective, the identification of de facto protected areas can indicate where it will not be possible to meet conservation targets with relatively unimpacted lands, given patterns of habitat conversion and anthropogenic disturbances. For example, we found that even relatively modest areal targets cannot be met in the most productive, least seasonal, and highest elevation areas of the boreal zone (Fig. 4b), although this accounts for only a small portion of Canada’s boreal region. In our case, some of these shortfalls may actually be due to how the study extent was defined. The exclusion of alpine areas above the treeline makes sense in the context of defining natural conditions on the basis of forest landscape structure. However, it also artificially fragments the study extent, and thus the classification results, in mountainous areas, causing the elimination of otherwise pristine areas that no longer meet the minimum size requirement. Canada’s existing protected area network, when evaluated nationally without a forested-land constraint, substantially over-represents high elevations (Andrew et al., 2011b), so it is likely that the present failure to represent high elevations is an artifact. Likewise the chosen study extent may also be responsible for the noted low representation of seasonalities around 1.3 (Fig. 4b). These seasonalities occur largely in the small portion of the Southern Arctic ecozone which is included in the boreal zone but does not currently contain any protected areas (the area north of ecozone 9 in Fig. 1) and, consequently, did not support any classifications. In contrast, the low representation of high productivities in de facto protected areas is likely real, and driven by preferential human use of high productivity ecosystems (Andrew et al., 2011b; Joppa and Pfaff, 2009). As productive areas are likely to be of great value to biodiversity, it is of concern that these also have reduced capacity to support protected area expansion in relatively undisturbed habitats.

It is likely that our results underestimate wilderness areas, particularly in the northern boreal ecozones, and therefore provide a conservative estimate of potential contributions to protected area expansion. This is a shortcoming of using existing protected areas to train the wilderness classification. When protected areas are biased in their distribution, as is frequently the case (Joppa and Pfaff, 2009), classification algorithms will detect (Kamei and Nakagoshi,
and propagate those biases. This has clearly occurred in our wilderness classifications of the Taiga Shield and Hudson Plains ecozones (Fig. 3), which were determined to have implausibly low amounts of de facto protected areas (Table 1). All three of these ecozones have few protected areas with biased geographic distributions (Fig. 1), which is reflected in the classification results (Fig. 3). Existing protected areas in the Hudson Plains and Taiga Shield East are restricted to the northern portions of these ecozones which are only sparsely forested; so too are the classification results concentrated in the north. Further, protected areas of the Taiga Shield East, none of which are under strict protection, are primarily protected salmon rivers and lakes. By having fewer types of protected areas containing fewer types of landscapes, less area in this ecozone is deemed as de facto protected by the landscape structure classification. Protected areas of the Taiga Shield West predominantly fall in the eastern portion of this ecozone (Fig. 1), and evidently fail to represent the forest landscape structures of the western half, which were left unclassified (Fig. 3). In such regions where neither anthropogenic use nor protected areas are well developed, the simpler low-access analysis (Fig. 2) is likely a more accurate portrayal of wilderness conditions.

The landscape structure classification might also be improved by incorporating additional variables. The six forest fragmentation metrics used in this study might not be the most sensitive indexes of boreal landscape change. For example, 1 km cells entirely covered by either old growth forest or by heavily managed production forest will yield identical fragmentation metrics based on a binary forest/non-forest classification, as will cells that are entirely non-forested either naturally (supporting, for example, wetland or alpine land cover) or due to forest harvest or conversion. Landscape composition and the configuration of specific forest types or age classes may contain meaningful information on boreal landscape condition (Schmiegelow and Mönkkönen, 2002) and wilderness classification. Another consideration is the spatial extent within which fragmentation metrics are calculated. O’Neill et al. (1998) recommend that measurement extents should be 2-5 times the size of patches. Forest patches in the boreal zone can be much larger than our 1 km² calculation landscapes. However, Cumming et al. (1996) observed relatively fine spatial scales of landscape variation in a Canadian boreal landscape, suggesting that 1 km² might adequately represent this landscape pattern.

Because our approach does not intrinsically rely on specific input variables, it is expected to be quite generalizable and portable to other regions and systems. Although we focused on gross forest pattern, the de facto protected area classifications can be pursued using any measures of landscape composition and configuration. Indeed, even within the boreal forest, we found the relative importance of landscape pattern metrics to be somewhat variable by region (Table 3) and protected area category. With the careful selection of relevant input variables, a classification of natural landscapes should be possible in any system. For example, the landscape classification of Cardille and Lambois (2010) considered 1921 landscape metrics computed from a 21-class land cover product, successfully characterizing landscapes in all biomes of the United States and providing some indication of naturalness, or lack thereof. Further, we believe that the approach of classifying de facto protected areas should be robust across levels of human impact, potentially much more so than a simple inventory of low-access areas. Many regions no longer possess large, intact natural areas (Sanderson et al., 2002), however, all are likely to exhibit a gradient of naturalness. Provided that a de facto protected area classification is trained with areas within a planning region experiencing less anthropogenic disturbance, it should be able to identify additional relatively natural areas that remain. Because existing protected areas tend to successfully reduce human impacts (DeFries et al., 2005; Joppa et al., 2008; Nagendra, 2008),
we expect they will be appropriate training regions regardless of the level of anthropogenic pressure in the study extent, although other criteria, such as expert judgment, may be necessary to choose training areas in regions with biased or poorly developed protected area systems. In regions with greater human domination, minimum size thresholds and distance buffers of human infrastructure will need to be reduced, if not eliminated altogether, and greater reliance will be placed on the classification based on landscape structure. In sum, although individual parameters (selection of landscape metrics, size and isolation thresholds, and potentially the protocol for defining training areas) will require fine tuning to ensure relevance to the conditions and pressures present, the general framework of a de facto protected area classification along landscape structure holds great potential regardless of regional context.

By any reckoning, Canada has large amounts of wilderness. These large blocks of intact landscapes with limited human access are prime candidates for protected area expansion. Conservation prioritization schemes differ largely in their treatment of vulnerability (Brooks et al., 2006). However, proactive preservation of large, intact wilderness areas is an important component of effective conservation (Watson et al., 2009). Although the remaining wilderness areas such as boreal forests are not especially species rich, they contain regionally important species, such as caribou. Further, recent assessments have shown that boreal ecosystem services can be substantially more valuable than natural resource extraction in these areas (Anielski and Wilson, 2009). Boreal wilderness thus merits expanded and ensured protection. This is especially so as vulnerability and intactness can change rapidly, particularly in the areas that are currently least impacted. The earliest roads in a road network have the greatest effect on fragmentation (Forman and Alexander, 1998; Hawbaker et al., 2006), and initial anthropogenic disturbance leads contagiously to continued habitat modification (Boakes et al., 2010). De facto protection plays valuable conservation roles (e.g., Carroll and Miquelle, 2006), but is not a viable option for long-term conservation, as it can be quickly overcome by changing market pressures and technological and infrastructure developments (Huggard et al., 2006). Such processes are already evident in the boreal forest: forest harvest and oil and gas development are expanding (Schindler and Lee, 2010; Smith and Lee, 2000), although rates and geographic patterns again reflect economic conditions (Masek et al., 2011). Lastly, large wilderness areas may actually be quite vulnerable, albeit to threats other than those typically considered in conservation prioritization (e.g., habitat destruction and overexploitation). High latitudes have and will continue to experience the greatest warming due to climate change (Solomon et al., 2007), altering the distribution and levels of protection of northern land covers and biomes (Lee and Jetz, 2008; Scott et al., 2002) and causing observable declines of northern species populations that do not correspond with their current relatively high levels of protection (Kujala et al., 2011).

5. Conclusions

Very few areas of the globe possess large amounts of wilderness impacted by minimal anthropogenic disturbances. Canada does, awarding it the enviable potential to establish proactive conservation measures and forward-thinking land-use planning for those regions. Our analyses, demonstrating a rationale for automated wilderness classifications, reveal that 80% of Canada’s boreal forest is remote from anthropogenic infrastructure, and that 50% has natural landscape structure, based on evaluations against protected area benchmarks, and is considered to be wilderness, or de facto protected areas. In addition to their vast areal extent, these areas span a wide range of environmental conditions, notably ecozones, vegetation productivities, and elevations. We do not suggest that all of these areas should be protected, or expect that they will.
Rather, we present tools that can be used as one component of integrated conservation planning. Quantitative classifications of *de facto* protected areas should generate valuable information regardless of system or region, and will likely provide more subtle indicators of naturalness than heuristics based solely on anthropogenic infrastructure. The *de facto* protected areas that are identified provide flexible options for the expansion of existing protected area systems into large, representative, connected protected area networks that maintain biodiversity and ecosystem functions in the context of sustainable resource management and global climate change.

**Acknowledgements**
This research was enabled through funding of “BioSpace: Biodiversity monitoring with Earth Observation data” through the Government Related Initiatives Program (GRIP) of the Canadian Space Agency.
References


### Table 1. Proportion of each ecozone classified as *de facto* protected by the low-access and Maxent analyses.

<table>
<thead>
<tr>
<th>Ecozone</th>
<th>low-access</th>
<th>Maxent, all protected</th>
<th>Maxent, strictly protected</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>all</td>
<td>&gt;540km²</td>
<td>&gt;3000km²</td>
</tr>
<tr>
<td>Atlantic Maritime</td>
<td>0.604</td>
<td>0.194</td>
<td>0.000</td>
</tr>
<tr>
<td>Boreal Cordillera</td>
<td>0.956</td>
<td>0.943</td>
<td>0.937</td>
</tr>
<tr>
<td>Boreal Plains</td>
<td>0.777</td>
<td>0.650</td>
<td>0.539</td>
</tr>
<tr>
<td>Boreal Shield East</td>
<td>0.893</td>
<td>0.814</td>
<td>0.754</td>
</tr>
<tr>
<td>Boreal Shield West</td>
<td>0.958</td>
<td>0.940</td>
<td>0.923</td>
</tr>
<tr>
<td>Hudson Plains</td>
<td>0.996</td>
<td>0.994</td>
<td>0.994</td>
</tr>
<tr>
<td>Montane Cordillera</td>
<td>0.937</td>
<td>0.909</td>
<td>0.845</td>
</tr>
<tr>
<td>Taiga Cordillera</td>
<td>0.959</td>
<td>0.947</td>
<td>0.947</td>
</tr>
<tr>
<td>Taiga Plains</td>
<td>0.967</td>
<td>0.961</td>
<td>0.946</td>
</tr>
<tr>
<td>Taiga Shield East</td>
<td>0.995</td>
<td>0.992</td>
<td>0.991</td>
</tr>
<tr>
<td>Taiga Shield West</td>
<td>0.997</td>
<td>0.997</td>
<td>0.997</td>
</tr>
<tr>
<td><strong>Overall</strong></td>
<td><strong>0.850</strong></td>
<td><strong>0.821</strong></td>
<td><strong>0.791</strong></td>
</tr>
</tbody>
</table>

\(^\dagger\) The Taiga Shield East does not currently contain any protected areas in the strictly protected categories. Consequently, a classification of *de facto* protected areas corresponding to the forest landscape structure of strictly protected areas could not be performed for this ecozone.

### Table 2. AUC scores and Maxent thresholds (in parentheses) for *de facto* protected area classifications, by ecozone and IUCN protected area designations. Dashes indicate classifications that could not be performed due to the absence of a particular protected area category in a given ecozone.

<table>
<thead>
<tr>
<th>Ecozone</th>
<th>Ia(^\dagger)</th>
<th>Ib(^\dagger)</th>
<th>II(^\dagger)</th>
<th>III(^\dagger)</th>
<th>IV(^\dagger)</th>
<th>V(^\dagger)</th>
<th>VI(^\dagger)</th>
<th>unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atlantic Maritime</td>
<td>-</td>
<td>0.621 (0.378)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Boreal Cordillera</td>
<td>0.609 (0.477)</td>
<td>0.747 (0.301)</td>
<td>0.673 (0.302)</td>
<td>-</td>
<td>0.711 (0.451)</td>
<td>-</td>
<td>0.826 (0.258)</td>
<td>-</td>
</tr>
<tr>
<td>Boreal Plains</td>
<td>-</td>
<td>0.776 (0.408)</td>
<td>0.557 (-)</td>
<td>-</td>
<td>-</td>
<td>0.781 (0.275)</td>
<td>0.821 (0.341)</td>
<td>0.691 (0.304)</td>
</tr>
<tr>
<td>Boreal Shield East</td>
<td>0.808 (0.372)</td>
<td>0.890 (0.177)</td>
<td>0.864 (0.322)</td>
<td>-</td>
<td>0.719 (0.412)</td>
<td>-</td>
<td>0.423 (-)</td>
<td>0.566 (-)</td>
</tr>
<tr>
<td>Boreal Shield West</td>
<td>0.821 (0.365)</td>
<td>0.843 (0.225)</td>
<td>0.662 (0.314)</td>
<td>-</td>
<td>0.803 (0.359)</td>
<td>0.620 (0.452)</td>
<td>0.807 (0.299)</td>
<td>0.772 (0.372)</td>
</tr>
<tr>
<td>Hudson Plains</td>
<td>-</td>
<td>-</td>
<td>0.647 (0.315)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.658 (0.450)</td>
</tr>
<tr>
<td>Montane Cordillera</td>
<td>-</td>
<td>0.649 (0.434)</td>
<td>0.710 (0.386)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.550 (-)</td>
</tr>
<tr>
<td>Taiga Cordillera</td>
<td>-</td>
<td>-</td>
<td>0.849 (0.298)</td>
<td>0.631 (0.454)</td>
<td>0.828 (0.298)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Taiga Plains</td>
<td>0.724 (0.363)</td>
<td>-</td>
<td>0.624 (0.388)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.776 (0.359)</td>
<td>-</td>
</tr>
<tr>
<td>Taiga Shield East</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.827 (0.345)</td>
<td>0.634 (0.271)</td>
<td>-</td>
</tr>
<tr>
<td>Taiga Shield West</td>
<td>-</td>
<td>0.817 (0.333)</td>
<td>0.840 (0.281)</td>
<td>-</td>
<td>-</td>
<td>0.826 (0.281)</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

\(^\dagger\) IUCN protected area categories are defined as: Ia - strict nature reserves, Ib - wilderness areas, II - national parks, III - natural monuments, IV - habitat/species management areas, V - protected cultural landscapes, and VI - protected sustainable use areas.
Table 3. Variable importance (% contribution) of the forest fragmentation metrics (defined in section 2.2 of the text) to de facto protected area classifications, by ecozone.

<table>
<thead>
<tr>
<th>Ecozone</th>
<th>fdense</th>
<th>farea</th>
<th>fpch</th>
<th>fprop</th>
<th>frarea</th>
<th>Fsarea</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atlantic Maritime</td>
<td>9.63</td>
<td>2.31</td>
<td>10.30</td>
<td>0.63</td>
<td>55.34</td>
<td>21.78</td>
</tr>
<tr>
<td>Boreal Cordillera</td>
<td>10.85</td>
<td>28.85</td>
<td>12.53</td>
<td>12.89</td>
<td>24.67</td>
<td>10.20</td>
</tr>
<tr>
<td>Boreal Shield East</td>
<td>22.34</td>
<td>12.93</td>
<td>7.43</td>
<td>19.41</td>
<td>27.33</td>
<td>10.56</td>
</tr>
<tr>
<td>Boreal Shield West</td>
<td>14.81</td>
<td>20.88</td>
<td>7.82</td>
<td>19.11</td>
<td>29.47</td>
<td>7.90</td>
</tr>
<tr>
<td>Hudson Plains</td>
<td>10.03</td>
<td>19.89</td>
<td>11.59</td>
<td>11.96</td>
<td>40.84</td>
<td>5.69</td>
</tr>
<tr>
<td>Montane Cordillera</td>
<td>10.40</td>
<td>36.06</td>
<td>12.73</td>
<td>15.11</td>
<td>11.57</td>
<td>14.15</td>
</tr>
<tr>
<td>Taiga Cordillera</td>
<td>15.66</td>
<td>26.92</td>
<td>11.12</td>
<td>2.23</td>
<td>37.71</td>
<td>6.36</td>
</tr>
<tr>
<td>Taiga Shield East</td>
<td>32.31</td>
<td>12.60</td>
<td>17.02</td>
<td>4.65</td>
<td>24.23</td>
<td>9.19</td>
</tr>
<tr>
<td>Taiga Shield West</td>
<td>28.69</td>
<td>16.40</td>
<td>12.73</td>
<td>7.29</td>
<td>30.39</td>
<td>4.49</td>
</tr>
</tbody>
</table>

Table 4. Productivity biases of existing and de facto protected areas, as identified from low-access and Maxent analyses. Only bias results from de facto protected areas larger than 3000 km² are shown. Biases by greater than 5% are shown in bold.

<table>
<thead>
<tr>
<th>existing protected areas</th>
<th>Atlantic Maritime</th>
<th>Boreal Cordillera</th>
<th>Boreal Plains</th>
<th>Boreal Shield East</th>
<th>Boreal Shield West</th>
<th>Hudson Plains</th>
<th>Montane Cordillera</th>
<th>Taiga Cordillera</th>
<th>Taiga Plains</th>
<th>Taiga Shield East</th>
<th>Taiga Shield West</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>minimum fPAR</td>
<td>-0.019</td>
<td>0.000</td>
<td>-0.027</td>
<td>-0.003</td>
<td>0.021</td>
<td>0.000</td>
<td>-0.056</td>
<td>0.000</td>
<td>0.008</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>integrated fPAR</td>
<td>-0.007 -0.050</td>
<td>-0.039</td>
<td>0.000</td>
<td>0.018</td>
<td>-0.163 -0.082</td>
<td>0.001</td>
<td>0.071</td>
<td>-0.045</td>
<td>0.019</td>
<td>-0.004</td>
<td>-0.000</td>
<td>-0.004</td>
</tr>
<tr>
<td>seasonality</td>
<td>0.005</td>
<td>0.035</td>
<td>0.032</td>
<td>0.001</td>
<td>-0.056 0.117</td>
<td>0.049</td>
<td>0.085</td>
<td>-0.095</td>
<td>0.042</td>
<td>0.040</td>
<td>0.003</td>
<td>0.003</td>
</tr>
<tr>
<td>elevation</td>
<td>0.034</td>
<td>0.032</td>
<td>-0.130</td>
<td>-0.026</td>
<td>0.016</td>
<td>-0.170</td>
<td>0.095</td>
<td>-0.201</td>
<td>-0.019</td>
<td>-0.053</td>
<td>-0.149</td>
<td>-0.020</td>
</tr>
<tr>
<td>low-access</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>minimum fPAR</td>
<td>-</td>
<td>0.000</td>
<td>0.014</td>
<td>-0.009</td>
<td>-0.007</td>
<td>0.000</td>
<td>-0.014</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>integrated fPAR</td>
<td>-0.001 -0.009</td>
<td>-0.040</td>
<td>-0.010</td>
<td>0.001</td>
<td>-0.019 -0.001</td>
<td>0.001</td>
<td>-0.004</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.004</td>
<td>-0.000</td>
<td>-0.023</td>
</tr>
<tr>
<td>seasonality</td>
<td>-0.003 -0.009</td>
<td>0.031</td>
<td>0.012</td>
<td>0.000</td>
<td>0.024 -0.002</td>
<td>0.005</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.027</td>
<td></td>
</tr>
<tr>
<td>elevation</td>
<td>-0.006 -0.034</td>
<td>0.017</td>
<td>-0.009</td>
<td>0.001</td>
<td>-0.004 0.003</td>
<td>0.002</td>
<td>0.000</td>
<td>0.000</td>
<td>0.001</td>
<td>0.001</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>Maxent, all protected</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>minimum fPAR</td>
<td>-</td>
<td>0.000</td>
<td>0.034</td>
<td>-0.009</td>
<td>-0.007</td>
<td>0.000</td>
<td>-0.056</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>integrated fPAR</td>
<td>-0.007 -0.009</td>
<td>-0.039</td>
<td>-0.011</td>
<td>-0.113 -0.073</td>
<td>0.009 -0.005</td>
<td>-0.014</td>
<td>-0.007</td>
<td>-0.015</td>
<td>0.015</td>
<td>-0.034</td>
<td>-0.015</td>
<td>-0.015</td>
</tr>
<tr>
<td>seasonality</td>
<td>-0.008 -0.030</td>
<td>0.024</td>
<td>0.013</td>
<td>0.066 -0.105</td>
<td>0.015 0.006</td>
<td>0.015</td>
<td>0.034</td>
<td>0.014</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>elevation</td>
<td>-0.005 -0.057</td>
<td>0.018</td>
<td>-0.007</td>
<td>-0.044</td>
<td>-0.038</td>
<td>-0.069</td>
<td>0.001</td>
<td>-0.015</td>
<td>-0.018</td>
<td>-0.003</td>
<td>-0.002</td>
<td></td>
</tr>
<tr>
<td>Maxent, strictly protected</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>minimum fPAR</td>
<td>-</td>
<td>0.000 0.068</td>
<td>-0.009</td>
<td>-0.007</td>
<td>0.000</td>
<td>-0.056</td>
<td>0.000</td>
<td>0.008</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>integrated fPAR</td>
<td>-0.003 0.047</td>
<td>-0.039</td>
<td>-0.015</td>
<td>-0.136 -0.075</td>
<td>0.010 0.042</td>
<td>0.134</td>
<td>-0.015</td>
<td>0.014</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>seasonality</td>
<td>-0.005 -0.044</td>
<td>0.023</td>
<td>0.013</td>
<td>0.085 0.105</td>
<td>0.015 -0.050</td>
<td>-0.150</td>
<td>0.050</td>
<td>-0.013</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>elevation</td>
<td>-0.009 -0.042</td>
<td>0.019</td>
<td>0.005</td>
<td>-0.071 -0.038</td>
<td>-0.070</td>
<td>0.007</td>
<td>0.007</td>
<td>-0.020</td>
<td>-0.002</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 2. Low-access areas of the boreal zone, distinguishing areas larger than either 3,000 km$^2$ or 540 km$^2$. 
Figure 3. Classification of de facto protected areas based on landscape structure. Areas are distinguished by the degree of protection to which their landscape structure corresponds (strictly protected: PA I-IV, or all others: PA V-VI, including unknown) as well as the size of the de facto protected area (larger than either 3,000 km² or 540 km²). The inset shows the increased level of detail as well as the reduced area of the de facto protected area classification relative to the low-access areas.
Figure 4. Bias (a) and representation (b), relative to productivity and elevation gradients, of low-access and de facto protected areas (>3,000 km²), corresponding to all or strictly protected categories, in boreal Canada. Panels in (a) plot histograms (frequency of 1 km² pixels) of productivities and elevations in boreal Canada overall and in the identified de facto protected areas. Those in (b) plot the areal proportion of each productivity/elevation range that occurs within the de facto protected area sets. These proportions were calculated from the areas in (a), i.e., by dividing the area for each de facto protected area set by that of boreal Canada at a given productivity. Dotted horizontal lines highlight proportional representation targets of 12% and 50%. In both (a) and (b), rows, from top to bottom, correspond to minimum annual productivity, integrated annual productivity, seasonality of productivity, and elevation.