Alberta Biodiversity
Monitoring Institute

www.abmi.ca

Manual for Species Modeling and Intactness

Version 2016-04-14

ALBERTA BIODIVERSITY MONITORING INSTITUTE

Disclaimer
These standards and protocols were developed and released by the ABMI. The material in this publication does not imply the expression of any opinion whatsoever on the part of any individual or organization other than the ABMI. Moreover, the methods described in this publication do not necessarily reflect the views or opinions of the individual scientists participating in methodological development or review. Errors, omissions, or inconsistencies in this publication are the sole responsibility of ABMI.

The ABMI assumes no liability in connection with the information products or services made available by the Institute. While every effort is made to ensure the information contained in these products and services is correct, the ABMI disclaims any liability in negligence or otherwise for any loss or damage which may occur as a result of reliance on any of this material. All information products and services are subject to change by the ABMI without notice.

Suggested Citation: Alberta Biodiversity Monitoring Institute. 2015. Manual for Species Modeling and Intactness (20029), Version 2016-04-14. Alberta Biodiversity Monitoring Institute, Alberta, Canada. Report available at: abmi.ca.

Contact Information
If you have questions or concerns about this publication, you can contact:
ABMI Information Centre
CW-405 Biological Sciences Centre
University of Alberta
Edmonton, Alberta, Canada, T6G 2E9
Phone: (780) 492-5531
E-mail: abmiinfo@ualberta.ca

Table of Contents

Summary

The Alberta Biodiversity Monitoring Institute (ABMI) monitors hundreds of species of vertebrates, invertebrates, plants, lichens and habitat structures at 1656 sites systematically located across the province of Alberta, and at additional targeted sites. Two of the main goals of the institute are to support natural resources management by providing: i) empirical models of the relationship of species and habitat elements to natural vegetation types, human footprint and climate and geographical gradients, and ii) credible and understandable indices of the status of biodiversity. This document details the statistical methods used to produce models for species, and to derive reference conditions and the indices ABMI has developed methods to assess the intactness (or deviation from reference condition) for species, groups of species and habitat features. The methods presented here are continuously in revision, and updated versions of this document will be released periodically.

1. Background on ABMI

The Alberta Biodiversity Monitoring Institute (ABMI) was initiated in 1997 through a broad partnership of industry, government and academia. ABMI is tasked with providing an effective way to track status and change to biodiversity at local, regional and provincial scales, and provide relevant and objective information to policy-makers, scientists and the general public.

The institute collects information on thousands of terrestrial and aquatic species (mammals, birds, mites, vascular plants, bryophytes, lichens, and aquatic macro-invertebrates) and habitat structures at 1656 sites spaced systematically on a 20-kilometre grid across the entire province. At full implementation, each of the 1656 sites will be sampled once every 5 years using a set of scientifically reviewed protocols. The same protocols are used at additional "off-grid" sites targeted to complement the main sites to address questions, such as the effects of gradients of human disturbance on a species abundance. This standardized data collection is designed to reduce duplication and increase cost efficiency for provincial and regional monitoring commitments, and provide a more complete understanding of cumulative impacts on the environment from multiple industries and human activities.

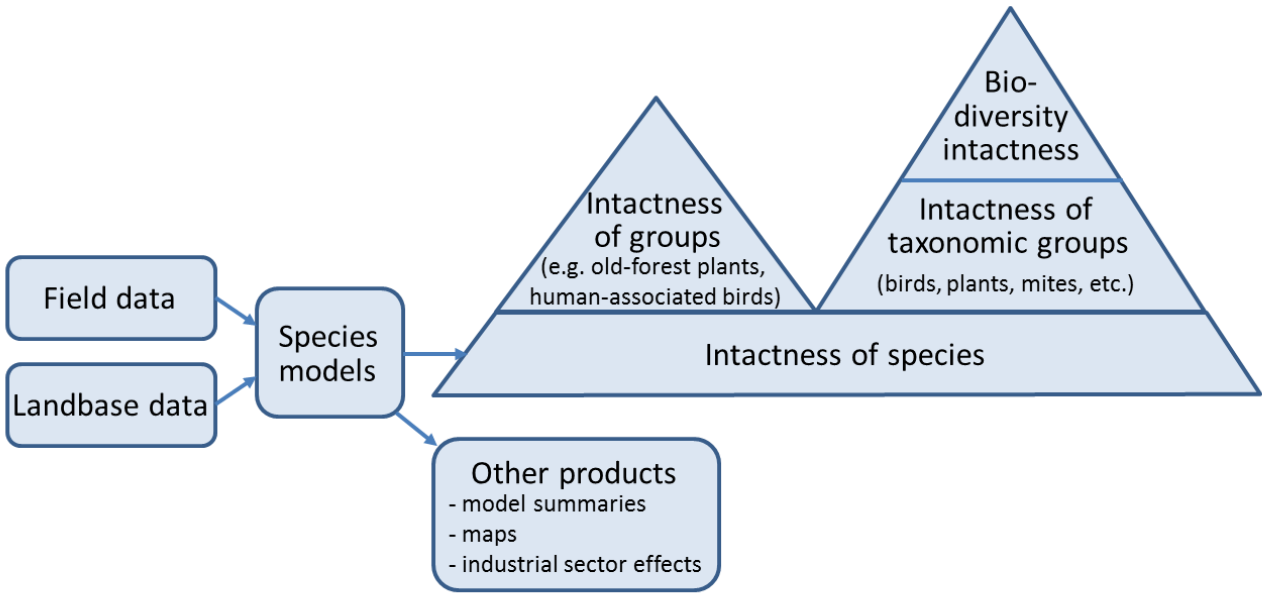
A main goal of the institute is to provide scientifically valid models of the relationships between species or habitat elements and vegetation types, human footprint and climate and geographical gradients (collectively referred to as "habitat models"). These habitat models have several purposes. One of these is to generate indices of the intactness of biodiversity for regions within the province and for the entire province of Alberta. These indices identify how the state of biodiversity has changed in comparison to an estimated reference condition.

This document describes the analysis methods used to generate the habitat models for species and habitat elements, and how these models are used to estimate intactness indices at the regional scale. Species level indices, along with important habitat structures, are ABMI's focus because they are the most clearly defined levels of biodiversity. However, information is also rolled up to provide indices for groups of species, and for biodiversity overall.

1.1 Overview

Figure 1 outlines the general procedure for producing species models and using them to calculate intactness for species, groups of species and overall biodiversity. Field surveys provide data on the relative abundances of species and habitat elements at ABMI sites. We use province-wide maps of vegetation and human footprint created by ABMI to describe the landbase at each site. Statistical analyses generate models of how each species' abundance relates to the vegetation and human footprint at the sites. These models are summarized and applied in different ways to generate several products, including estimates of relative abundance of each species in each vegetation or human footprint types, and maps of the predicted current distribution of the species across the province. The models are also applied to "reference" vegetation maps, in which human footprint has been replaced or "back-filled" with the original vegetation. Intactness of a species compares the predicted current abundance of the species to the abundance predicted under the reference condition with no human footprint. Intactness of a group of species, such as old-forest plants or human-associated birds, or of a larger taxonomic group, such as all mites, is calculated by averaging the intactness across all species in the group. Overall biodiversity intactness is an average of the intactness values for the larger taxonomic groups.

Field protocols are described in http://www.abmi.ca/abmi/reports/reports.jsp?categoryId=0. The landbase data are described in http://www.abmi.ca/abmi/rawdata/rawdataselection.jsp. Section 2 of this document details the statistical analysis that combines these into habitat models for each species. Section 3 outlines how the models are used to generate intactness values for species, and how these are combined into intactness for groups and overall biodiversity. Section 4 covers the analysis of data from ABMI's wetland monitoring sites. Section 5 describes how we produce other products from the habitat models: model summaries, maps and estimates of the effects of individual industrial sectors.



2. Habitat Models for Species

Modeling the relationship of species to habitat is at the heart of many ABMI analyses, including direct descriptions of those habitat relationships, mapping and species intactness. "Habitat" is used in a broad sense, including: 1) Human footprint, 2) Descriptions of the natural ecosystems – vegetation (stand types by age classes) in forested areas, or soils in prairie areas, 3) Climate and spatial variables (latitude, longitude), and 4) Footprint surrounding a site. (The effects of footprint surrounding a site are described below, but these are not currently used in our products, for reasons we discuss.) We use a flexible analysis for the habitat modeling, which can be adopted to address specific questions, such as the effects of particular industrial sectors.

2.1. Relative abundance by human footprint and ecosystem type

2.1.1. Modeling species

The first, most basic step of the analysis, is to estimate the relative abundance of a species in different human footprint types and ecosystem types. Because most ABMI sites cover several habitat types, we use a multiple regression approach to separate the effect of each type. The main footprint and ecosystem variables used in the models (section 2.1.2) and modeling procedure (section 2.1.3) are the same between taxa, but details of the modeling differ slightly due to different sampling methods for the taxa and habitat elements.

Vascular plants and mites: The data for both taxa are treated as occurrences in 4 quadrats (plants) or soil samples (mites) in the 1-ha central area of each site, using a logit-link binomial distribution with 4 trials. Vegetation and human footprint variables are summarized for the 1ha central area.

Bryophytes and lichens: The data since 2009 are similar to the vascular plant data, and are also analyzed as 0-4 quadrat occurrences per site. Prior to 2009, these taxa were surveyed with a more complex design that stratified by microhabitat across the central 1-ha plot. Those data are reduced to simple presence/absence at the site. The pre- and post-2009 data are analysed together, using a logit-link binomial distribution with 1 (pre-2009) or 4 (post-2009) trials, and an additional factor for old versus new protocol. The protocol factor adjusts for the different area of each occurrence (1ha versus 50mx50m) and different search protocol.

Mammals: Mammals were originally surveyed using snow track transects with a triangle of 3km per side. The protocol was changed to a 10km linear transect in 2005. In both methods, occurrences of mammal species were recorded separately for 1km segments. Mammals are analysed as presence/absence on each 1km segment, with vegetation and human footprint summarized in a 250m buffer around each segment. Adjustments are made to estimated standard errors to compensate for the dependence of segments along the same transect, and transects are resampled together in bootstrapping (section 3.3). A factor is included in the analysis to account for the change in protocol from triangles to transects. Additionally, tracking data is directly affected by days-since-snow when the survey was conducted. A quadratic relationship with days-since-snow is included in the models to account for this.

ABMI has recently (2015) switched from snow track transects to cameras for monitoring mammals. Different abundance metrics will be used when that data becomes available for analysis, but the overall analysis is expected to be similar to other taxa.

Birds: The ABMI bird data is modeled as the number of occurrences in the 9 point counts at a site, using a logit-link binomial distribution with 9 trials. The raw data for the bird surveys records the number of individuals of a species at a point count station, but we reduce this to presence/absence at each station, because most values are 0 or 1, and repeatability of the data for multiple individuals at a station is low. Vegetation and human footprint variables are summarized for the total area occupied in 150m-radius circles around each of the 9 count stations at a site.

The above description applies to analyses of only the bird data collected by ABMI. However, most results presented by ABMI, including on the website and in intactness analyses, use a larger dataset provided by the Boreal Avian Modeling (BAM) project. The BAM bird data includes ABMI data, Breeding Bird Survey data and datasets collected by Environment Canada and researchers throughout the boreal forest. These analyses are more detailed because the combined dataset for birds is richer than that for other taxa. The models produced, however, are the same as those produced for other taxa, and they are used in the same way to produce our various products.

The various data sources were standardized using statistical offsets (see Sólymos et al. (2014) for additional references and more detailed description of the bird modeling).

Habitat elements – trees, snags and logs: Counts of these structures, divided into classes by diameter and species groups, are analysed with log-link negative binomial count models, which can capture the highly aggregated nature of some structures (a few high counts, many zeroes).

Predicted counts from these models can be converted into densities (/ha) and then into basal areas of trees or snags (m^2/ha) or volumes of downed wood (m^3/ha), using quadratic mean diameters of the size classes and the known sampling areas or transect lengths.
Habitat elements – cover layers: Canopy, shrub, etc. cover layers are analysed as a logit-linked binomial variable, as this produced more stable models than alternatives such as beta distributions or other transformations.
Habitat elements – miscellaneous: pH is modeled as a log-normal distribution. Organic depth and soil carbon have a normal error distribution.

2.1.2. Human footprint and ecosystem variables

The human footprint types and way of describing ecosystems differ between the North (boreal, foothills, Canadian shield and Rocky Mountains) and the South (grasslands, parklands, dry mixedwood and some of the central mixedwood subregion of the boreal).

North: Human footprint types in the North are grouped as cultivation, hard linear (non-vegetated: roads, railways, etc.), soft linear (vegetated: pipelines, powerlines, road margins, seismic lines, etc.), urban/industrial (combining industrial sites like well-pads, compressor stations, factories, with urban and rural residential areas, which cannot currently be distinguished reliably in our footprint maps), human-made water bodies and forestry. Forestry sites are classified according to their pre-disturbance stand as deciduous or mixedwood, upland spruce+fir, or pine, and their age (0-10yr, 10-20yr, 20-40yr, 40-60yr, 60-80yr). Ecosystems are described using vegetation types, which are combinations of broad forest types (deciduous, mixedwood, upland spruce+fir, pine, black spruce) and age classes (0-10yr, 10-20yr, 20-40yr, 40-60yr…140+ yr), along with open vegetation types – larch fens, upland grass, upland shrub, wet grass, wet shrub, open water, rock and other barren areas. All areas/polygons in the North are classified as one of these human footprint or natural vegetation types.

South: Human footprint types in the South are grouped as cultivation, hard linear, soft linear, urban-industrial and human-made water body. There is no forestry in the South region. Ecosystems are currently described by two variables: Soil groups and pAspen. Soil groups are productive (loamy, silty types), clay, saline and rapid-draining (sandy and badlands). pAspen is the probability that aspen would be found under natural (no footprint) conditions in a plot at the site, based on a model of the species that uses climatic, topographic and lithology variables.

2.1.3. Estimating species' relative abundances in human footprint and ecosystem types

North: In the North analysis region, a two-stage procedure is used to estimate species' relative abundances in the human footprint and vegetation types. The first step estimates relative abundances for the human footprint types and the broad vegetation types (combining stand ages within the forested types). The second stage estimates the species' relationship with age of forest within the broad stand types.

The first stage uses a set of 7 models. The first model includes each human footprint and vegetation type (broad stand types, no age classes). The other models combine similar human footprint or vegetation types that are difficult to estimate into coarser groups. The estimates for each habitat type are combined with AIC-based model-averaging, which weights each model based on a metric of how well the data for the species supports each model. These models provide estimates of a species' relative abundance in each human footprint type (including the age classes and stand types of forestry cutblocks, although classes or stand types are combined in some models and hence have the same estimate), in the non-forest vegetation types, and in the broad stand types.

The second stage estimates how the species' relative abundance within each of the broad stand types changes as a function of stand age. The age analysis is done separately for each stand type. Only sites with >10% of their area in the target stand type are included. Sites included in the analysis are weighted by the proportion of their area in that stand type (for example, a site with 20% pine would have a weight of 0.2, compared to a weight of 1 for a site that was 100% pine). Some recently burned upland forests are classified as grass or shrubs in our vegetation data. Grass and shrub sites are therefore included as age 0-10 or age 10-20 stands, respectively. They are weighted as 0.5 X the proportion of the target upland stand type in the surrounding $1km^2$ area. The 0.5 weighting is to reflect the fact that some of these grass and shrub sites are edaphic (permanent) openings, not early-seral stages of the stand types. With this dataset for each stand type, a curve is fit between the relative abundance of the species and age. A predicted abundance of the species at each site is first generated from the relative abundance coefficients in step 1, and this value is used as an offset in the age analysis. The analysis is therefore looking at the residual (incremental) effect of age on the species' abundance in the stand type. These age curves are used to predict the relative abundance of the species in each age class of each stand type.

After estimates of relative abundance have been generated for each human footprint and vegetation type using the above models, two adjustments are made: 1) Estimates for hard and soft linear features are often highly uncertain, because these features never occupy a high proportion of ABMI sites, making it difficult to separate their effect from the effects of other vegetation types. We therefore modify the estimate for hard linear features by combining it with the estimate for the similar urban/industrial footprint type, using an inverse-variance weighted estimated (the weight given to the original hard linear or urban/industrial estimate is proportional to 1/variance of the estimate). For soft linear features, we use an inverse-variance weighted average with recent forestry cutblocks (for upland sites) and the youngest age class of black spruce (for lowland sites). With inverse-variance weighting, the poorer the estimate for hard or soft linear footprint, the more strongly the final estimate is pulled towards urban/industrial or forestry/young lowland estimates. 2) Estimates for cutblocks >20yr are poor, because there are few sites with old forestry. We adjust the estimates for cutblocks >20yr so that they converge on natural stands by age 80yr. Based on expert opinion of the timing of convergence of harvested and natural stands, the estimate for spruce and pine cutblocks 20-40yr is adjusted 50% of the way towards the estimate for natural 20-40yr stands, 40-60yr cutblocks are adjusted 85% of the way and 60-80yr cutblocks 96% of the way. Convergence rates are somewhat higher for deciduous stands: 70%, 91% and 97% at 20-40yr, 40-60yr and 60-80yr.

For bird data analysis we used Poisson generalized linear models with a log link. The response variable was the total number of individuals counted per survey. We used the statistical offsets to account for differences in sampling protocol and nuisance parameters affecting detectability (time of day, time of year, tree cover and composition). We applied forward stepwise variable selection to minimize bias in predictions. The best supported model was determine based on the

Akaike's information criterion (AIC). Local habitat conditions within a 150 m radius circular buffer assessed based on current land cover composition. Dominant habitat type was assigned to each survey station based on a simple majority rule. Habitat age effects and interactions between forest types and age were considered. Age was calculated as the difference between year of the bird survey and the year of origin. Year of origin was estimated based on area-weighted age of forest polygons within the 150 m radius circular buffer. Age was modelled as a linear or quadratic relationship using age and square root of age. We have used expert-based trajectories to differentiate between stands of natural (fire) vs. cutblock origin. We assumed that the growth trajectory converges to natural disturbance trajectory at year 80 post disturbance. When the dominant habitat type was forestry footprint, the backfilled habitat class was used. This was done so that forestry activities did not form a separate habitat class but instead was treated as young forest with a modified trajectory. Bird models also incorporated the presence of roads and the proportion of soft linear features within the 150 m buffers. see Sólymos et al. (2014) for additional references and more detailed description of the bird modeling.

South: There are no stand ages in the South analysis, so relative abundances in each human footprint and soil type are estimated in a single step. Eight models are used in an AIC-based model-averaging framework. The first four models combine difficult-to-estimate soil types or human footprint types in different ways. The other four models add a term for pAspen to the first four models. These models are used to estimate species' relative abundances in each human footprint and soil type. (For presenting results, separate estimates are generated for sites with pAspen=0 and pAspen=1 – i.e., sites predicted to be completely treeless or completely forested prior to human disturbance). As in the North, estimates for hard linear footprint are modified using an inverse-variance weighted average with urban/industrial footprint. There are no estimates for forestry in the South, so estimates for soft linear features are not adjusted.

2.2. Additional variation due to climate, geographic location and surrounding human footprint

The analyses in section 2.1 provide estimates of species' relative abundances in human footprint and ecosystem types that are the same throughout the analysis zone. In addition to varying due to these factors, we know that many species also vary geographically and/or along climate gradients. Species may also be affected by the amount of footprint in an area around a site, not just the footprint at the site. For example, a species that uses native vegetation may be less common at a site with native vegetation if it is surrounded by human footprint than if it is part of a larger area of native vegetation.

The estimates of abundance by human footprint and ecosystem type are first used to predict a species' abundance at each site. These are then used as offsets in models to estimate the residual effects of geographic location, climate and surrounding human footprint. We use 14 models with different individual or combined climate variables, such as mean annual temperature, potential evapotranspiration, etc. An additional 14 models add latitude and longitude, and a further 14 models add latitude, longitude and latitude*longitude. The best of these 42 models of residual effects of climate + geographic location is selected using BIC.

The climate and/or spatial variables from this best model are then used as covariates in a set of 7 models estimating the residual effects of human footprint in a 1km^2 unit surrounding each site.

The 7 models use different ways of grouping the human footprint types. These 7 models of surrounding-footprint effects, including the best climate and spatial covariates, are combined using BIC-weighted model-averaging. Currently, however, we are not using the surrounding-footprint effects in the predictions used to generate intactness estimates, maps and industrial sector effects. Confounding of surrounding footprint with footprint at the site, climate variables and geographic location make it difficult to reliably estimate the incremental effects of surrounding footprint for all species.

2.3 Using sites with more than one visit

Some sites have been surveyed more than once. These repeat samples are included in intactness models but weighted by the number of repeats (1/n; in other words, each site has the same overall weighting even if some sites include data from more than one survey). During bootstrap analysis (section 3.3), the site is treated as the unit of resampling, such that data from all surveys at a site are resampled together.

Once data from the planned re-surveys of sites become available in the future, additional terms will be added to the models to allow the species abundance to differ between initial and subsequent sampling periods. The predicted abundance for the second (or later) measurement cycle could therefore be greater or less than the abundance predicted from habitat and footprint alone. That difference would reflect trends in the species' abundance beyond those due to footprint changes. Intactness calculations after a second set of samples is begun will include both the footprint effect and any additional changes in the species. This will allow us to attribute changes in species abundances to both the effects of changes in human footprint and other changes unrelated to footprint.

2.4 "Targeted" sites

In addition to the 1656 sites on the systematic grid across the province, ABMI samples many "targeted" sites. These sites are chosen to complement the systematic sites for the purpose of addressing specific short-term questions, and are sampled with the same protocols as the systematic sites. A main objective for targeted sites is to improve sampling coverage along the gradient of human footprint levels, and thus to improve estimation of species-footprint relationships. Targeted sampling is therefore focused on the end of the footprint gradient that is less common in a region: high footprint sites in the boreal and foothills region, and low footprint sites in the grasslands and parklands. Effort is made to find sites satisfying these conditions while also being widely distributed and representative of the ecosystem types in the particular region (although this is not always possible). Some targeted sites are chosen in underrepresented footprint types, such as large industrial sites, or address specific management questions, such as the rate of species recovery in older cutblocks. Because we calculate reference and current abundances of species using complete maps of the region of interest rather than just our sampled sites (section 3), our intactness results are not affected by the fact that the targeted sites are not a representative sample of the footprint in the region.

3. Intactness

3.1 Species Intactness

ABMI's intactness index compares the predicted current abundance of each species to its predicted reference abundance. Our reference condition is the current abundance with the effects of human footprint statistically removed using our habitat models (Nielsen et al. 2007). Alternative definitions of reference conditions have been used elsewhere, but these are problematic: 1) Abundances in protected areas cannot be used, because protected areas are rarely representative of all ecosystems. In Alberta, many protected areas are in remote or high elevation regions with low productivity. 2) Time-zero approaches set the reference condition as abundance in a certain year. However, human footprint has affected parts of Alberta for over a century, and many areas had extensive development when systematic monitoring began in 2007. 3) Desired or target abundances of species are sometimes used as references, but these are social values that differ for different people, and it is infeasible to set targets for thousands of species. Using a reference condition based on statistically removing the effects of human footprint overcomes these problems. However, this "de-footprinted" reference condition cannot account for any past changes in the species' abundance that are not due to local or regional footprints, including historical exploitation, effects of diseases or introduced species that are not associated with footprint, climate change or past effects outside the province for migratory species.

The species intactness index compares the predicted relative abundance of each species across the reporting region to the predicted abundance for that species under zero human footprint in the same region. This measure of intactness is scaled between 0 and 100, with 100 representing current abundance equal to that expected under reference conditions, and 0 representing species abundance as far from reference condition as possible. Both over- and under-abundances are viewed as deviations from intact conditions. The index is estimated as:

Current / Reference × 100%, when Current<Reference, or

Reference / Current × 100%, when Reference<Current

A value of 50%, for example, means that the species is either half as abundant as the reference condition, or twice as abundant. Because the intactness index for individual species decreases from 100% with either downwards or upwards differences from reference conditions, an "increaser" species does not cancel out a "decreaser" species. Instead, both contribute to lowering the average intactness.

Intactness in large regions

The reference and current abundances of a species in a reporting region are calculated by applying the habitat models and models of residual climate/location/surrounding footprint to landbase data from every 1km^2 unit in the region. The landbase data comes from province-wide maps of vegetation, soils and human footprint produced by ABMI. Compiling and updating the human footprint information for the whole province is time-consuming, so regional intactness results usually represent human footprint levels 2 or 3 years prior to reporting.

When we calculate intactness for large diverse regions, such as those used by the Land Use Framework in Alberta, we may have different numbers of species modeled for different parts of the region. Specifically, we have more species modeled in the northern regions of the province, where we have more samples. Simply averaging all species found in a region would bias results towards the northern parts of the region. Instead, we first calculate intactness for each natural sub-region within the region of interest, and then calculate an area-weighted average intactness of the sub-regional values. This procedure treats each area of the region equally, regardless of how many species we have modeled in the area.

Within each subregion, we also exclude species that are predicted to have very low abundances in that subregion. The specific criterion is to exclude any species whose predicted current and reference abundances in the region are both <1% of the species' predicted maximum abundance in the province. This criterion excludes modeled species from sub-regions where they are unlikely to occur in reality, while not favouring common species over species that are uncommon everywhere in the province.

Predicted current abundances rather than observed

The intactness index uses the predicted current abundance for each 1km^2 unit when the footprint variables are set to their current levels, rather than directly using the counts at surveyed sites, for 3 reasons:

1) This approach allows us to do the calculations using 1km^2 units that cover the entire area of the region, rather than just the small subsample of the area where we have ABMI sites with the direct counts. That is particularly important for reporting on small subregions.
2) Measurement error is large for single counts, and substantial even across multiple sites in a subregion. We do not want to compound true changes in species caused by footprint and the effects of measurement error.
3) The reference condition is calculated using regression models with a link function (log or logit). The current condition is also calculated with this link function, to avoid differences due simply to different types of scaling (parallel to the difference between arithmetic and geometric means).

Predicted intactness over time

When we report on changes in intactness over time in a region, we use a subset of 3x7km areas overlaying ABMI sites (i.e., on a 20km x 20km grid) in which human footprint has been mapped from satellite images for most years from 1999 to present. These 3x7 areas are a 5.25% sample of the all regions. Because they are a sample of a region, rather than a complete inventory, they add sampling error to the intactness estimates. Our analyses show that this sampling error adds 2-3% additional error to the intactness estimates for species, which is small compared to the uncertainty in the models generating the current and reference predictions.

The importance of "increaser" species

In areas with low footprint, such as much of northern Alberta, species that increase with human footprint tend to have lower intactness than species that decrease. With 5% footprint in a region, for example, it is rare to have a species that decreases more than 5%. To do so, the species would have to have a strong negative response to footprint and footprint would have to be concentrated in preferred habitat for that species. In contrast, it is more common for some species to increase several-fold as soon as any footprint occurs in a region, because several species associated with footprint are expected to be extremely rare under no-footprint reference conditions – the proportional increase when they show up in footprint is therefore large. Although species that increase with footprint can be an ecological concern because of negative effects on other species, species that decrease with footprint are often of a greater priority to managers. For this reason, we often report intactness of decreaser and increaser species separately, or highlight decreaser species or groups of species.

3.2 Intactness for groups, taxa and overall biodiversity

Intactness for groups of species within a taxon, such as old-forest birds or berry-producing shrubs, is calculated as the average intactness of species in the group. Similarly, intactness for a taxon like bryophytes or mites is the average intactness of all analysed species in that taxon.

Overall biodiversity intactness is the average of the intactness for each of the taxa surveyed by ABMI (birds, mammals, vascular plants, bryophytes, mites, lichens). Each taxon is weighted equally, regardless of how many species it contains.

We considered many other ways of combining species when calculating intactness for species groups, but concluded that a simple mean is the most appropriate. Geometric, harmonic or other types of averages have undesirable properties, including giving excessive weight to individual species with extreme values, which were often the most poorly estimated results (for rare or highly aggregated species). Alternatives that weighted species by the precision of their intactness estimate are biased toward common species. More complicated methods that statistically correct for the expected relationship between abundance and precision could be used to remove this bias with precision-weighted averages. However, the loss of transparency and ease of understanding the results outweigh any possible statistical benefits of such a procedure.

<u>Non-native plants</u>: For non-native species, it does not make sense to calculate an average intactness, because the more rare non-native plant species that are included, the higher the intactness. For example a non-native species that is observed once in a region with 100 sites sampled has an intactness of 99%, and another non-native that is observed at 50 sites has an intactness of 50%. If these intactness values were simply averaged, the first record of a newly invading non-native species would raise the mean intactness, which is not a sensible result. For that matter, potential non-native species that have not yet been detected would have an intactness of 100%, and would raise the average intactness for the overall group even higher. Instead, a different approach is used for non-native plants. We assume that the reference condition is 0 for these species, then calculate a species' intactness as: SI_s = 100% - percent occurrence of the species. Percent occurrence is calculated at the quadrat (50m x 50m) level. For example, a non-native species that occurred in 20% of quadrats would have an intactness value of 80%. We calculate the overall intactness of ALL non-native plants as 100% - percent occurrence of any non-native species on a quadrat. This value is simply the percent of quadrats on which no non-native species was detected in a region.

3.3 Bootstrapping and confidence intervals

The sampling distribution of species intactness indices are estimated using bootstrapping, in which the original data are resampled with replacement and the entire analysis repeated 100 times. The bootstrap replicates are used to calculate the median reference condition and confidence intervals (based on percentiles of the 100 resampled values). Bootstrapping is required, rather than an analytical formula, because the current abundances and reference conditions are not independent, and the intactness calculation is complicated for the multi-stage, multi-model approach with different weighting of revisited sites. A blocked bootstrap is used, in which the resampling is done within pre-defined spatial blocks to preserve the spatial structure of the sample design. For mammal snow-tracking, the transect or triangle is used as the unit of resampling. When we are using the 3x7km areas to estimate intactness changes over time, the

3x7 areas in the reporting region are also resampled during the bootstrapping to include the additional error due to these areas being a sample of the region, rather than a complete inventory.

4 Wetland habitat models and intactness

In addition to the systematic sites, ABMI conducts additional surveys targeted at small wetlands with open water. The systematic sites can fall in upland or lowland habitat, but the field surveys cannot be conducted in open water. Open-water wetlands and their adjacent emergent, margin meadow and margin wooded habitats are important distinct ecosystem types that are most efficiently monitored with targeted sites and specific field methods. ABMI surveys the nearest open-water wetland to each of the systematic sites, with transect-based surveys of plants and aquatic invertebrates (www.abmi.ca/home/ publications/1-50/45.html).

4.1 Wetland habitat models

Analysis of footprint effects on wetland species takes a different approach than analysis with the systematic sites, because the relevant footprint is not just in the wetland itself, but also in the surrounding habitats. Additionally, habitat classification and mapping of wetland types is less developed than for terrestrial habitats, but there are other important ecosystem variables, such as wetland depth or chemistry.

Habitat modeling for wetland species therefore first analyzes three sets of ecosystem covariates: climate and spatial variables, wetland physical and chemical variables (e.g., wetland depth, pH, total nitrogen, etc), and broad surrounding vegetation (North) or soil (South) types. The best sets of these covariates are chosen with a model selection procedure. Using those best covariates, the analysis then examines what effect footprint in the surrounding area has on the species' abundance. We are currently using a 250m-wide buffer around the wetland's water edge to define the surrounding area. Eventually we will change that to a topographically defined catchment for the wetland, which is more relevant biologically. The analysis results in relationships of each species to surrounding footprint types, having factored out the effects of relevant covariates.

We summarize the wetland models with figures showing how the predicted relative abundance of the species changes with increasing levels of each footprint type in the surrounding area, at average values of the relevant covariates, and we tabulate the coefficients of the model for use by others. We cannot currently map wetland species or wetland abundance across the province, because we do not have accurate maps of where the types of wetlands we sample occur, and also because the models include wetland physical and chemical variables that have to be measured in the field for each wetland.

4.2 Wetland intactness

Because we cannot map wetland species' current and reference abundances across the province, the procedure for calculating regional intactness is simpler for wetland species than for species in the systematic sites. We predict the current and reference abundance of each wetland species at each sampled wetland – the only places we have the necessary field sampling of the physical and chemical covariates. The current abundance is predicted with footprint in the surrounding 250m buffer set to the levels currently observed; the reference abundance is predicted with those levels set to 0. As with the systematic sites, we use the predicted current abundance rather than the observed values for each species so that we are not including the effect of measurement error as part of the difference between current and reference values, and to allow an "apples-to-apples" comparison of the same logit-link values. For regional intactness, we simply sum the predicted current and reference abundances of a species for all wetlands sampled in the region, and calculate intactness with the same formula used in the systematic plots.

5 Additional products from the habitat models

In addition to regional intactness estimates, we use the human footprint + ecosystem models (section 2.1) and the models of residual effects of climate and geographic location (section 2.2) to produce several other products. Many of these products are on the ABMI species website (www.species.abmi.ca)

5.1 Upland and lowland intactness

Intactness can be calculated separately for upland and lowland habitat types (or for any other mapped subset of the landbase). The same approach of applying the habitat models to the current and reference landbases apply. A complexity is that, unlike overall intactness, we need to know what habitat types each type of footprint is in, within each 1km^2 unit. For example, if a 1km^2 unit contains both upland and lowland habitat, and forestry and energy footprint, we need to know how much of each footprint type is in upland versus lowland habitat, to know what effects footprint is having in those habitat types. Having to track all combinations of habitat types X footprint types in each 1km^2 unit creates much larger datafiles in the analysis, but otherwise the approach of predicting current and reference abundances, and calculating resulting intactness, is the same as for overall regional intactness.

In the South, we have few systematic plots in lowland habitats, which are rare. Lowland intactness there relies exclusively on the wetland plots. In the North, many systematic sites fall in lowland, which includes black spruce forest and larch fens. We intend to combine intactness estimates from the lowland systematic plots and the wetland plots in that region, but have not yet worked out the details.

5.2 Summaries of habitat models

To allow others to evaluate and use our habitat models, we summarize them by presenting the relative abundance of each species in each footprint and habitat type, along with the coefficients for the residual effects of climate, geographic location and surrounding footprint (when the latter is being used). For each species, we make predictions for sites that are 100% of each human footprint or natural ecosystem type (vegetation, including age class for forest stands, in the North; soil types with pAspen at 0 or 1 in the South). These predictions are the expected relative abundance of the species in that footprint or ecosystem type. "Relative abundance" for all taxa except birds has the specific meaning of the expected probability of detecting that species in a standard ABMI survey. For example, a relative abundance of 0.2 for a plant species means that we would expect to find it in 20% of quadrats surveyed by ABMI in that habitat type. For birds, relative abundance is the expected number detected in a standardized point count. Confidence intervals are estimated directly from the models. These values are presented as figures for each species and tables with all species in a taxon.

For soft and hard linear footprint, there is never 100% of that one type at an ABMI site. For these linear features, we instead use the models to predict the effect of converting 10% of an average site to these linear features. Figures of those calculations show whether a species is predicted to have a positive or negative response to the linear features compared to an average undisturbed site.

For the climate and geographic variables, we present the coefficients of the residual models directly in tables. These relationships are also used in the province-wide maps (section 4.3).

For the models of the effects of surrounding footprint, we predict the change in abundance of a species at a site that has an intermediate level of the species as surrounding human footprint varies from 0 to 100%. This is graphed to illustrate the additional effects of surrounding human footprint, and the model coefficients are also available in tables. However, we are currently not using or presenting the surrounding footprint effects, because these results are not yet reliable for all species.

5.3 Maps of current and reference abundance

We combine information on the human footprint and habitat composition of 1km^2 units covering the province with the climate and location information for each unit, to map the predicted abundance of each species based on the habitat models and residual effects of location and climate. This is the predicted current abundance of the species.

To map reference abundance of each species, we repeat the above procedure, but using a map in which all human footprint has been "back-filled" – replaced by the vegetation type that was most likely to have occurred prior to footprint. Back-filling in the North is based on the vegetation types immediately surrounding the human footprint, as well as rules about which footprint types are restricted to certain vegetation types (e.g., forestry is only in mature upland forest, agriculture is only on upland sites). The resulting map is the predicted reference abundance of the species, the abundance it would have had if there was no footprint in the area. In the South, the ecosystem variables soil type and pAspen are mapped for all areas, including in human footprint, and therefore we do not need to back-fill human footprint to get reference conditions in this region. We simply ignore the effects of human footprint and predict all areas based only on the soil types and pAspen.

5.4 Effects of industrial sectors

We can also use the habitat models to estimate the effects of individual industrial sectors in a region. As for upland and lowland intactness (section 5.1), we start with a summary of all combinations of habitat types X footprint types in each 1km^2 unit, for both the current and reference (no footprint) conditions. We can then "turn on" only the footprint types associated with one industrial sector and compare the predicted abundances of each species with only that footprint type in the current landbase to the reference landbase. This shows how much the footprint from just that one industrial sector is predicted to have changed the abundance of that species. We do this calculation for the major industrial sectors of agriculture, forestry, energy, rural/urban and transportation. The analysis gives both the predicted total effect of the industrial sector on each species' abundance, and the effect per unit area of the industry's footprint (i.e., how "intensive" the effect of the industry's footprint is).

Literature Cited

Alberta Biodiversity Monitoring Program. 2014. Terrestrial field data collection protocols (abridged version), version 2014-3-21. Alberta Biodiversity Monitoring Program, Alberta, Canada. Report available at: http://www.abmi.ca/abmi/reports/reports.jsp?categoryId=0 [December 2014].

Nielsen, S.E., Bayne, E.M., Schieck, J., Herbers, J., and Boutin, S.A. 2007. A new method to estimate species and biodiversity intactness using empirically derived reference conditions. Biological Conservation 137: 403-414.

Sólymos, P., Mahon, C. L., and Bayne, E. M. 2014. Development of predictive models for migratory landbirds and estimation of cumulative effects of human development in the oil sands areas of Alberta. Technical report, Joint Oil Sands Monitoring: Cause-Effects Assessment of Oil Sands Activity on Migratory Landbirds, Edmonton, AB. pp. 829. http://www.borealbirds.ca/files/JOSM_Report_2014_AllSpecies.pdf [November 2015]