

# ABMI Species Website Manual

*Alberta Biodiversity Monitoring Institute*

*Version: 2017-03-23*

# Contents

<b>1 Introduction</b>	<b>3</b>
Version . . . . .	3
Suggested citation . . . . .	4
License . . . . .	4
Contact . . . . .	4
Rationale for presenting model results . . . . .	4
<b>2 Habitat associations</b>	<b>6</b>
Full habitat models for relatively common species, systematic (non-wetland) sites	6
Methods . . . . .	7
Limitations . . . . .	8
Downloadable results . . . . .	10
Full habitat models for relatively common species, wetland sites . . . . .	13
Methods . . . . .	13
Limitations . . . . .	14
Downloadable results . . . . .	14
Use-availability summaries for rarer species . . . . .	14
Methods . . . . .	15
Limitations . . . . .	15
Downloadable results . . . . .	15
<b>3 Predictive maps</b>	<b>16</b>
Methods . . . . .	16
Limitations . . . . .	17
Downloadable results . . . . .	18
<b>4 Sector effects</b>	<b>19</b>
Methods . . . . .	19
Limitations . . . . .	20
Downloadable results . . . . .	20
<b>5 Intactness</b>	<b>21</b>
Limitations . . . . .	22
<b>6 Richness</b>	<b>23</b>

Methods . . . . .	23
Limitations . . . . .	23
<b>7 Uniqueness</b>	<b>25</b>
Methods . . . . .	25
Limitations . . . . .	26
References . . . . .	27

# Chapter 1

## Introduction

This manual serves as background material for *individual species* results available from the Alberta Biodiversity Monitoring Institute (ABMI). The document describes the content, methods, and limitations/caveats associated with species results which are currently accessible in two places:

- Species Website (<http://species.abmi.ca>): The ABMI Species Website was developed, and is maintained by the ABMI Science Centre. This site was developed to publicly share species results, and present summaries of new analyses as results become available.
- Data & Analytics Portal (<http://www.abmi.ca/home/data-analytics>; launched December 2016): The ABMI Data & Analytics (DA) Portal was developed, and is maintained by the ABMI Information Centre. The DA Portal serves to present a variety of ABMI data and mapping products, including not only the results presented in the Species Website but also a number of other data and mapping products and functions. This site allows users to access and interact with ABMI data in a variety of ways.

Because the DA Portal has only been recently launched, these two websites will continue to function in parallel for the immediate future. This allows users to learn about the new website, it allows the Science Centre to responsively update data products and present new analyses, and it ensures seamless access to all materials provided at both sites.

## Version

This version 2017-03-23 (date of updating the manual) of the manual describes results for version 4.0 (2016-09-12) species results available from the ABMI Species Website (<http://species.abmi.ca>) and the ABMI Data & Analytics Portal (<http://abmi.ca/home/data-analytics>) (date of updating the website results).

## Suggested citation

Alberta Biodiversity Monitoring Institute. 2016. ABMI Species Website Manual, Version: 2017-03-23. Alberta Biodiversity Monitoring Institute, Alberta, Canada. URL: <http://species.abmi.ca>.

## License

© Alberta Biodiversity Monitoring Institute (2014–2016) under a [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/).

## Contact

In case of questions related to ABMI Species Website Manual, please contact [Peter Solymos](#).

## Rationale for presenting model results

ABMI collects data on thousands of species, and evaluates how their relative abundance varies among natural and human-created habitats. The raw data and model results are made available for public use on the ABMI website. However, some species are detected only a few times, other species are not well detected using the ABMI sampling method, and some species produce models that are not plausible.

## Sample size

ABMI conducts separate analysis for two regions: the Northern (Boreal, Canadian shield, Foothills and Rocky Mountain natural regions) which is vegetation-based and Southern (Grassland, Parkland and Dry Mixedwood regions) which is soil based. The vegetation-based models require on average 20 degrees of freedom (df) with habitat/age/climate terms in the models. The soil-based models require on average 10 df with soil/climate terms in the model. Accordingly, modelling is carried out for all species, except birds, that occur in at least 20 sites in an analysis region. Since a great deal more information is available for birds, a species must have at least 100 detections in the Northern analysis region, and 50 detections in the Southern analysis region to be modelled.

## Sampling design adequacy and detection issue

ABMI is designed to determine habitat association and track population trend for common species. The chosen sampling methods, however, do not capture all species effectively, including species that live in rare patchy habitats.

**Birds.** Point count methods are inadequate for monitoring the following groups of birds: waterfowl, shorebirds, raptors, owls, nightjars because:

- Their habitats are not well sampled (waterfowl, shorebirds),
- Their nocturnal activity periods are not well sampled (rails, owls),
- They vocalize irregularly or infrequently,
- Their home range is at much larger spatial scale than sampled by point counts (raptors, owls).

The taxonomic information for bird species as part of the downloadable material provides details about which bird species are adequately monitored. The classification follows family level taxonomy.

**Vascular plants, bryophytes and lichens.** A number of species belonging to these taxa are inadequately sampled or detected, because:

- Survey season is not optimized (too early/too late) to sample the species when their characteristic features (flowers, seeds) for field identification are available to distinguish them from similar species in the field. Problematic vascular plants are downgraded to genus or family level in the data, \* Survey protocol is not designed to sample them (e.g., crustose lichens),
- Species is part of taxonomically complex groups which makes it difficult to distinguish them. This applies particularly for lichens and bryophytes, and a few vascular plants. These species are analysed at genus level or species group level.

## Model output plausibility

Once a species meets the minimum detection requirements, visual inspection of the model results is done to evaluate the plausibility of the output. Visual inspection includes:

- Determining whether habitat associations are compatible with known biology of species,
- Determining whether confidence intervals in habitat associations are extreme (relative to mean abundance),
- Determining whether the predicted distribution of the species is compatible with the known spatial distribution of the species (range maps, detections).

Once inspection is done, notes are added to the species lookup table and final decisions are made. This way the steps in the evaluation process are transparent and decisions can be re-evaluated if needed. These are incorporated into the species taxonomic information table to make the process and any caveats transparent.

# Chapter 2

## Habitat associations

### Full habitat models for relatively common species, systematic (non-wetland) sites

Models are used to show how species, relative abundance differed among vegetation, soil and, human footprint types in the northern, and southern regions of Alberta. Models were only created for species that had enough detections. The following figures are presented to describe habitat associations:

- **Species–habitat associations in northern Alberta.** Predicted species abundance in each vegetation and human footprint type is shown with bars. Vertical lines indicate 90% confidence intervals. Dots within the forested habitat types show predicted species abundance in forest harvest areas of various ages.
- **Species–habitat associations in southern Alberta.** Due to climate and soil conditions natural disturbances and vegetation succession, varying amounts of aspen and other trees may be present on each soil type in the south (**Non-treed** sites and **Treed** sites); the presence/absence of trees greatly affects the presence and abundance of many other biota. Separate figures are presented for south **Non-treed** sites and **Treed** sites. Predicted species abundance in each soil/human footprint type is shown with bars. Vertical lines indicate 90% confidence intervals.
- **Linear footprint relationships in northern Alberta.** For linear footprint in northern Alberta, the pairs of points show the change expected from the average habitat with no linear footprint when 10% linear footprint is added.
- **Linear footprint relationships in southern Alberta.** For linear footprint in southern Alberta, the pairs of points show the change expected from the average habitat with no linear footprint when 10% linear footprint is added.

## Methods

We use a set of statistical models to relate the relative abundance of each species measured at ABMI sites to two sets of variables: vegetation or human footprint type, and geographic location or climate variables. The analyses are done separately for two overlapping regions: North (boreal, foothills, Canadian shield, parkland), and South (grasslands, parkland, dry mixedwood subregion of the boreal).

Relative abundance for vascular plants, bryophytes, and lichens is the probability that the species will occur in a 50 m x 50 m ABMI quadrat, or a soil sample for mites. For birds, relative abundance is the number of birds detected at a point count, adjusted for the species' detection distance and standardized for other sampling variables. For mammals, we are using winter snow-track transect data, with relative abundance being the probability of finding the animal's tracks on a 1-km segment 5 days after a snowfall. These are the values on the y-axis of the habitat model figures, and in the coefficient tables. We also model the number of non-native vascular plant species, besides individual vascular plant species.

Vegetation types in the North region are based on several province-wide GIS layers of vegetation variables, and include main forest stand types (white spruce, pine, deciduous, mixedwood, black spruce) by broad age classes (0-9 years, 10-19 years, then 20-year classes to 140+ years) and several categories of open vegetation (larch fens, upland grass, upland shrub, wet grass, and wet shrub). In the South, 'vegetation types' are broad soil types, such as productive, rapid draining, etc. The South models also include a term for the probability that the site is or was naturally treed. Open wetlands are sampled and modeled separately.

Several categories of human footprint are distinguished in the models, including cultivation (typically forage or tame pasture in the North and crops in the South), urban/industrial, forestry, soft (vegetated) linear features and hard (unvegetated) linear features. Forestry footprint is differentiated by broad stand type and the same age classes as natural forest (although we have few samples in forestry areas >20 years old).

The scale of the vegetation or soils and human footprint information used in the models is a square 1-ha area for plants, lichens and mites, a 150m-radius circle for birds, and a 250m buffer around the 1-km snow-track transect segment for mammals.

Flexible relationships with latitude, longitude and a set of broad climate variables are also included in the models to represent the geographic and climatic distribution of the species. The bioclimatic variables used in the modeling were calculated at a 4-km resolution using monthly climate normals of temperature and precipitation averaged over 1961-1990.

The analysis is conducted in a 'model selection framework', which means that the models for each species are only as complex as the data for that species support. For example, the models for a rarer species with few records may not be able to separate different age classes or even broad stand types. All age classes or several stand types therefore have the same average value in the resulting habitat model. This does not mean that there are truly no differences between these types for that species. It just means that we do not yet have enough data to estimate abundances in all those individual vegetation types.

This habitat modeling requires moderate sample sizes, so we can only do the analyses for species found in at least 20 sites in an analysis region (25 sites for birds; 50 transect segments for mammals).

Full details of the habitat modeling can be found in the ABMI Species Modeling Protocol[1].

## Limitations

There are many limitations and caveats about the use of these habitat models:

**We use simple abundance measures.** For vascular plants, bryophytes, lichens and mites, we are only analyzing occurrence of the species at 0-4 of the 4 quadrats or soil samples per site. The number of individuals, percent cover, total volume or biomass, reproductive status, etc. of the species is not part of this measure — a single sapling of a tree species is an occurrence counted as much as complete cover of large trees. There is also no adjustment for the probability that the species was present but was missed by the field technicians. That probability is probably higher in complicated quadrats, species rich habitats, rare species, and for groups of species that are hard to distinguish in the field, all of which could affect the habitat models.

For birds, the measure is based on counts of birds heard in the field by trained observers in some studies, or heard on recordings made using various technologies in others (including ABMI). These are mainly, but not exclusively, singing male birds. Adjustments are made for different detection distances in the different studies, habitats and for different species, but these adjustments all have statistical uncertainty. The number of vocalizing birds may also not directly reflect the total number of birds or reproductive success in different habitat types.

For mammals, ABMI is switching to camera-based indices, but these are not available yet. The models are currently based on snow-track transects, so they reflect how far the species travels in different habitat or human footprint types during winter. This can be very different from the habitats that they use for foraging, or during summer. Additionally, ABMI transects were along linear features that could be travelled by snowmobile, so the results are confounded by how those features affect winter travel by the species.

**There is often high uncertainty in individual values.** The confidence intervals on individual habitat values for a species are often wide, and should be integral to any use of the habitat models. There are two main reasons for wide uncertainty in our results: 1) We estimate values for many habitat and human footprint types, particularly in the North, with many combinations of broad stand types and age classes. Ideally we would have hundreds of observations of a species to estimate so many values. Instead, we use a fairly complicated modeling procedure to extract reasonable values for many species with lower sample sizes by simplifying the complexity of habitat types. For all but the most common species, the resulting habitat models are simplified (some habitat types are assumed to be the same) and estimated with large error. 2) ABMI sites are systematically located, and the Alberta landbase often contains a fine-scale mix of several habitat types and human

footprint. Individual sites, even 1-ha areas, therefore usually fall in more than one habitat and/or human footprint type. Our models are designed to separate out the effects of each individual habitat or human footprint type as best as possible, but the analysis is necessarily imperfect with limited sample sizes. A highly specialised species, for example, will almost inevitably show some predicted abundance in habitat types that it does not use, when these co-occur with preferred habitat types.

**The explanatory variables are confounded.** We have limited ability to separate effects of each vegetation and human footprint type, and spatial and climate variables. In addition to the challenges generating uncertainty in the individual values, ABMI is not an experimental study. We do not design and manipulate habitat and human footprint types, and simply use what is sampled in the province. As a result, habitat and human footprint types are partially confounded with geographic location and climate variables. Our modeling separate out these effects as best as possible, but without impossible province-wide experiments, we can never be sure how well we have estimated the separate habitat, human footprint and geographic/climate effects.

**We use the same model set for all species.** Our model selection approach weights different models in a model set based on the support provided by each species' data. However, we use the same set of models for all species. We currently do not have the resources to develop separate model sets for individual species to benefit from the differences in natural history.

**There are errors in species identification.** All taxa have potential for errors in species identification, including the mammal track surveys and the bird recording interpretation. Taxonomic issues are common with the plants, bryophytes, and lichens, where many species groups are difficult to distinguish and taxonomy changes frequently. This adds to uncertainty in models for some species.

**The vegetation and soil classification is simplified and contains errors.** ABMI's vegetation map[2] is compiled from various sources, not all of which cover the entire province. Some layers are less accurate than others. Also, even if individual polygons are accurate, their boundaries may not be precise, and many ABMI sites fall on the edges of vegetation polygons. Resulting error in the vegetation types at sites increases error in the habitat models. Some habitat types may have additional challenges. Grass and shrub sites in the North, for example, may be true permanent open habitats, but the same designation is often given to early seral stages of forested sites after fire. Grass and shrub sites in the prairies are also poorly mapped. Mixedwood stands are also poorly mapped, because they represent a gradient between deciduous and upland conifer stands, which can often change as a stand ages.

**The soil classification[3] currently uses only broad groupings of individual soil types.** Few species show strong responses to soil type in our models, suggesting that these groups may not be well-defined or relevant to many species. We have too few sites to treat lowland sites in the South separately.

**We have only broad classes of human footprint.** Human footprint[4] is generally mapped more precisely than vegetation, but we use only broad groups of footprint types.

Industrial sites and urban or rural residential areas, for example, cannot be reliably distinguished. Values are particularly difficult to estimate reliably for linear features, because these usually occupy only a small portion of a site. We also have few sites with older forestry cutblocks. Values for those habitat types combine our rough estimates with an assumed rate of convergence of aging forestry and natural stands.

**No other effects of human activities are included in our models.** Many sites in the South are grazed to varying degrees but this information is not available. Thus grazing effects cannot be included in our models, and if grazing differs among native habitat types the effects confound those difference. Any wide-scale human effects, like pollution away from visible human footprint or recent climate change, are also not included in our modeling. Additionally, we only include effects of footprint at the local scale — the 1-ha scale we sample at, or larger sampling areas for birds and mammals. Due to sample size constraints and confounding among variables, we do not include any additional effects that might be caused by human footprint in the larger surrounding area.

## Downloadable results

There are separate tables of coefficients for the North and South analysis regions. The North region is the Boreal, Canadian Shield, Foothills and Parkland natural regions. The South region is anywhere that soil information is available (see soil maps on website), including: Grassland and Parkland natural regions (including Peace River), most of the Dry Mixedwood subregion of the Boreal region and smaller areas of the Central Mixedwood subregion and southern Foothills regions. There is an extensive area of overlap between these analysis regions, where results from both regions can be used.

The coefficient tables give relative abundances for each analysed species in each habitat or soil type, and human footprint types. For all taxa except birds, these coefficients are the probabilities of detecting the species in a standard ABMI plot in that habitat or soil type. For birds, the value is an index of relative density, adjusted for the different detection distances of species in unlimited distance point counts, and the different survey methods used in some non-ABMI data sources that are included in the bird models.

The coefficients are for a simple linear additive model: Relative abundance of a species in a study area is defined as  $Coef_A \times Area_A + Coef_B \times Area_B + \dots$ , where A, B, etc. refer to the different habitat/soil and human footprint types. Because this is a simple additive equation, it can be applied on any scale individual 1ha units similar to ABMI plots, 1 km<sup>2</sup> raster units, entire regions, etc. If the Area values are expressed as a proportion of the total area of the study unit, then the resulting abundance value for all taxa except birds will be the probability of detecting the species in a standard ABMI plot located randomly within the study area. But any units of area can be used in the equation, as the results are just an index of relative abundance.

**Northern coefficients table.** For each species, the table gives the estimate of the species' relative abundance in each habitat type (natural or human footprint), along with the lower and upper 90% confidence intervals on the estimates (.LCI and .UCI columns). Forested

habitat types are a combination of broad stand types and an age classes. The broad stand types are:

- “White spruce”. This is really “upland non-pine conifer”. It is dominated by white spruce, but also includes other upland conifer species like Engelmann spruce or hybrids, balsam fir and Douglas-fir where those species occur.
- “Pine”. Any pine species.
- “Deciduous”. Any stand type dominated by deciduous trees, primarily upland but including some areas of lowland deciduous.
- “Mixedwood”. Any stand mapped as mixedwood.
- “BlackSpruce”. Lowland conifer stands, which may including species other than black spruce, but excluding stands where the dominant tree is larch.

The age class codes for these stands are (all these stands are of natural origin; harvested areas are designated separately, see below): R = 0–9 years, 1 = 10–19 years, 2 = 20–39 years, 3 = 40–59 years, 7 = 120–139 years, 8 =  $\geq$  140 years.

Additional habitat types are:

- “LarchFen”. Lowland stands with larch present. All ages classes are combined in this rare stand type, except for the birds, where more data allows separate estimates by age class.
- “Grass”. Upland edaphic grass areas, or herbaceous areas. Mapping of this type is imprecise, and some early seral post-fire stands may also be included here.
- “Shrub”. Upland edaphic shrub areas. Mapping of this type is imprecise, and some early seral post-fire stands may also be included here.
- “WetlandGrass”. Lowland grass- or herb-dominated areas. Mapping of wetland grass and wetland shrub is not precise, and there may be considerable overlap.
- “WetlandShrub”. Lowland shrub-dominated areas. Mapping of wetland grass and wetland shrub is not precise, and there may be considerable overlap.
- “Swamp”. These are lowland tree-dominated areas that where leading species is not black spruce or larch (i.e. not treed bogs or treed fens). This category is distinguished for birds only.

Coefficients are also given for a set of human footprint types:

- “CC” preceding a stand-age designations indicates a cutblock in that stand type. The age classes are the same as for natural stands, and refer to time since harvest. Mixedwood cutblocks are combined with deciduous for all taxa except birds.
- “Cult”. Cultivated land, primarily tame (plowed) forage and pasture in the north and crops in the south.
- “UrbInd”. Urban areas, industrial areas (including wellsites, compressor stations, etc.) and rural residential areas.
- “SoftLin”. Soft linear footprint, which are vegetated linear features potentially capable of succeeding to natural vegetation eventually. These include seismic lines, pipelines, powerlines and road margins. Note that these linear features are treated as areas here, like all other habitat types. If they are mapped as linear vectors, they have a buffer width assigned along their length to give them an area. Values in the tables reflect

estimates under 100% soft linear disturbance which does not naturally occur in our samples. The column “SoftLin10” and figures on the website represent 10% linear footprint, i.e. composed as a mix of two strata: 90% natural (AverageCoef) + 10% footprint (no confidence intervals provided for this average).

- “HardLin”. Hard linear footprint, which is permanently alienated, primarily roads, but also rail lines. These linear features are treated as areas. If they are mapped as linear vectors, they have a buffer width assigned along their length to give them an area. Values in the tables reflect estimates under 100% hard linear disturbance which does not naturally occur in our samples. The column “HardLin10” and figures on the website represent 10% linear footprint, i.e. composed as a mix of two strata: 90% natural (AverageCoef) + 10% footprint (no confidence intervals provided for this average). Note that estimation of hard linear footprint effects is different for birds where instead of the amount of footprint in the buffer only the presence of a hard linear feature is noted rendering the count a ‘roadside survey’. As a result, the estimate reflects the contrast between the average abundance for an off-road survey vs. a roadside survey.
- “AverageCoef” refers to the mean of coefficients across all natural habitats and human footprint types not counting footprint classes. This is used as a reference value in the website figures for linear features (no confidence intervals provided for this average).

Note that some features are assumed to have coefficients of 0 for all terrestrial species, these are open water (natural or human created) and active mine sites.

**Southern coefficients table.** The southern coefficients are based on broad soil types derived from the GSC soil layers (see soil map on website[3]). For areas mapped with multiple soil types, the dominant type is used. As in the north, the mean estimate is presented, along with the lower and upper 90% confidence intervals on the estimates (.LCI and .UCI columns). Soil types modeled are:

- “Productive” soils include loamy, silty, limy and subirrigated soils.
- “Clay”.
- “Saline” soils include blowouts, overland flow and saline lowlands.
- “RapidDrain” soils include gravel, sand, badland, and several other rapidly draining soils

ABMI has too few samples in wetland types in the south to provide coefficients for those rare types. A separate sampling protocol and data summary provides information for open wetlands through the southern region.

In addition, coefficients are given for the human footprint types (detailed descriptions are presented under Northern coefficients):

- “Cult”
- “UrbInd”
- “SoftLin”
- “HardLin”

There are separate tables listing coefficient values in *treed* and *non-treed* locations. This reflects the difference based on probability that the site is or was naturally treed. The average

coefficient provided for the soft and hard linear feature coefficients and the linear feature plot on the website in the south refers to the average of natural habitat coefficients under non-treed conditions.

Water and active mine sites are assumed to have coefficients of 0 for all terrestrial species.

## Full habitat models for relatively common species, wetland sites

### Methods

In addition to the systematic sites, ABMI surveys the small permanent wetland with open water that is nearest to each of the systematic sites. Surveys use transects to monitor vascular plants and net sweeps to sample aquatic invertebrates. For plants, the transects extend into the wet margin area surrounding the water.

Habitat models for species found in wetlands include covariates describing the wetland and human footprint in the surrounding area. Three sets of covariates are used: 1) Physical and chemistry variables measured in the wetland, including water depth, temperature, pH, salinity, dissolved oxygen, dissolved organic carbon, total nitrogen and total phosphorus. 2) Broad vegetation types or soil groups in the surrounding area. Vegetation types are used in the north and include stand types (e.g., deciduous, pine, see description under systematic sites) as well as open habitat types. Soil groups are used in the south, with broad groupings (e.g., productive, rapidly draining soils, see description under systematic sites). 3) Latitude, longitude and climate variables (e.g., mean annual temperature, precipitation, see description under systematic sites). Within each set of covariates, multiple models are examined in a model-selection framework. Another model-selection step is then used to weight the importance of each set of covariates to the species being modeled. This two-stage model selection approach is intended to find the best set of covariates for each species, even for species that have low sample sizes (rarer species with few records).

We then use the best set of covariates in models to estimate how human footprint in the area surrounding the wetland affects the species. We use broad human footprint categories, such as agriculture or all other footprint types in the south, and successional or alienating footprint in the north. We use a flexible regression to estimate how the amount of surrounding footprint affects the species, accounting for the effect of the covariates at each wetland. These models can be used to predict what effect human footprint has had on each wetland species in a region, which is the basis of ABMI's reporting on species intactness.

Full details of the habitat modeling can be found in the ABMI Species Modeling Protocol[1].

## Limitations

Limitations and caveats of the modeling described for systematic sites also apply to the wetland models. Several additional caveats include:

- The wetland models use vegetation or soils and human footprint in the surrounding 250m-wide buffer area. This is a fairly arbitrary area used as on an interim bases until the actual hydrological catchment of the wetland is determined.
- The physical and chemical variables used in the model are sampled on a single day and time for each wetland. Some of these variables are highly dynamic, so the single sample introduces considerable uncertainty in the relationships.
- As with the systematic models, the analysis does not account for human effects beyond visible footprint. Grazing in wetlands, agricultural runoff, other pollution and hydrological disruptions (e.g., from roads) are all potentially important effects on wetland species that we do not incorporate.
- Our wetland monitoring uses stratified sampling to account for the different distinct wetland zones. We do not have information on how the area of the zones may have been altered in the past, or information on how many wetlands have disappeared altogether (and hence are not available for sampling).
- Confounding of the explanatory variables is a particular challenge with the wetland models. The physical and chemistry variables, surrounding vegetation or soil, and geographic location and climate are all correlated with each other. With limited samples and no opportunity to manipulate those variables experimentally, it is difficult to reliably assign effects to each individual variable.
- Wetland models are produced for any species found at 20 or more wetlands. Because our sampling transects for plants extends into the adjacent upland, some of these species also have habitat models from the systematic sites. The two sets of models are quite different. The systematic models directly show the relative abundance of the species in each vegetation or soil and human footprint type. In contrast, the wetland models show how those vegetation or soil and human footprint types in the surrounding land area modify the abundance of the species in or around wetlands. The two types of models therefore need to be applied in different ways.

## Downloadable results

This is yet to be determined.

## Use-availability summaries for rarer species

If it was not possible to create complex habitat association models for species, we present a coarse index of habitat use the proportion of detections in each native vegetation and human footprint type in comparison to the proportional availability of the habitat types.

## Methods

ABMI has information on many species that are detected too few times for full habitat models. For species detected at 3 or more sites but less than 20 sites (less than 50 or 100 for birds; less than 50 transect segments for mammals), we summarize habitat and human footprint use with a simple use-availability index. The index is based on the proportions of each broad habitat or human footprint type in the sites where the species is present versus the proportions of habitat and human footprint type across all sites that we have sampled. A value of 0 for a habitat type means that the habitat type is as abundant at the sites where the species occurs as it is across all sampled sites; increasingly positive values show that the habitat type is used disproportionately more than its overall availability; a negative value means that it is relatively underused by the species.

## Limitations

In addition to the limitations listed for the full habitat models, uncertainty is extreme in the use-availability figures due to the small sample sizes for these poorly detected species. We also do not have enough data to separate out confounded effects of geographic or climatic distributions from habitat preferences. For example, a species with a southern range will have higher use-availability indices for grass, shrub and deciduous forest, which are more common in the south, even if it does not really prefer these types to conifer forest. The figures are simply intended to summarize habitat and human footprint use information in our data for these rarer species. They should not be interpreted as general habitat models for the species.

## Downloadable results

The use-availability indices are available for each species in the following land cover types in **northern** study region: Deciduous, Mixedwood, White Spruce, Pine, Black Spruce, Larch, Open, Wet, Forestry, Cultivated, Urban/Industrial, Hard Linear, Soft Linear. In the **southern** study region, we use the following classes: Productive, Clay, Saline, Rapid Drain, Cultivated, Urban/Industrial. Interpretation of the land cover types is similar to as explained above for habitat modeling, with some of the classes were merged. The values range between -1: avoidance, 0: neutral, +1: preference, given availability.

# Chapter 3

## Predictive maps

Habitat association models, plus models describing how species varied spatially and with climate gradients were used to predict species abundance in 1 km<sup>2</sup> spatial units under reference and current conditions.

We present the following figures:

- **Predicted relative abundance throughout Alberta under reference conditions.** Habitat association models, plus models describing how species varied spatially and with climate gradients were used to predict species abundance in 1 km<sup>2</sup> spatial units under reference conditions. Predictions of relative abundance of species in each 1 km<sup>2</sup> unit were made after all human footprint in the 1 km<sup>2</sup> unit had been ‘backfilled’[2] based on native vegetation in the surrounding area.
- **Predicted relative abundance throughout Alberta under current conditions.** Habitat association models, plus models describing how species varied spatially and with climate gradients were used to predict species abundance in 1 km<sup>2</sup> spatial units under current conditions. Predictions of relative species abundance in each 1 km<sup>2</sup> unit were made based on the vegetation and human footprint present in the 1 km<sup>2</sup> unit under current conditions (in year 2012).
- **Difference in relative abundance between current and reference conditions.** For each 1 km<sup>2</sup> unit the difference between predicted current and reference conditions was determined.
- **Uncertainty in predicted current conditions.** To highlight the degree of uncertainty in the models, we estimated the standard error of prediction for each 10 km x 10 km units.

### Methods

Province-wide maps are generated by applying the models of a species’ relationship with vegetation or soil types, human footprint, geographic location and climate to a set of 1-km<sup>2</sup> pixels covering the province. Within each pixel, the proportion of each vegetation or soil and

human footprint type in the 1-km<sup>2</sup> area is determined, along with the latitude, longitude and climate values at the centre of the pixel. The ‘current’ maps summarize information from the current landbase, including current human footprint (2012 human footprint[4] is being used presently). The ‘reference’ maps summarize information from a ‘back-filled’ map[2], in which footprint has been removed and replaced with the vegetation type that was expected prior to the footprint being created. Back-filling is based on adjacent vegetation types and rules about what vegetation types different footprint types can occur in (e.g., forestry occurs only in harvestable-age upland stands). The ‘difference’ maps show the areas where the species is predicted to be higher or lower in the current map compared to the reference map.

Separate predictions are made using the North and South models. In the overlap area where both of these predictions are available, a weighted average prediction is used. The weighting is based on the mapped probability that the site is or was treed, with the weight for the North prediction being that probability and the weight for the South prediction being 1 minus that probability. In other words, the South predictions apply to the untreed proportion of the raster, while the North applies to the treed part.

Province-wide maps are only produced for species that we can analyze from the systematic sites. The models generated for species in the wetland sites cannot be applied province-wide for two reasons: 1) We do not have GIS layers describing where all the wetlands of the type we sample occur in the province, 2) The wetland models include physical and chemistry variables that need to be sampled in the field for each wetland, and therefore are not available for unsampled wetlands.

## Limitations

There are several limitations to ABMI species maps:

- The maps are based on predicted habitat quality. They show the average expected abundance of a species given the habitat types, footprint, geographic location and climate of the 1 km<sup>2</sup> area. We do not know the actual abundances of each species in each of the 1 km<sup>2</sup> area in the province. There are many variables beyond habitat quality that determine why a species may be absent or rare at a particular place where we predict it to be common, and vice versa.
- There is uncertainty in the habitat models for each species. Uncertainty of habitat models is high for rarer species and for rarer habitat types. Uncertainty is also higher in the corners of the province, particularly the northwest and southwest, where we have limited sampling and are extrapolating from our models. Accompanying uncertainty maps show the absolute prediction uncertainty across the province.
- The map predictions are most accurate at broad regional scales. At smaller scales, the true abundance of the species can vary from the prediction, because many factors beyond habitat quality affect a species. At small scales — individual pixels through townships — the maps show the habitat quality for the species, but should not be interpreted as showing the actual abundance of the species in that small area.
- The underlying vegetation maps used to make the predictions are imperfect. There

are classification errors in some places with greater errors for some particular habitat types (e.g., grass and shrub, mixedwood), and areas missing some of the input layers particularly outside of commercial forestry areas. We make assumptions about stand ages in about 1/3 of the forest area that is missing that information. Human footprint is more accurate, but may be 2 or 3 years out of date.

- Reference maps only show the effects of statistically removing the effects of visible human footprint. They do not show how the species would have been distributed under ‘pristine’, ‘natural’ or ‘pre-historic’ conditions.
- Current maps do not show effects of other human disturbances, like pollution, cattle grazing, climate change or changes on wintering grounds for migratory species.

## Downloadable results

We provide comma separated text (CSV) files with predicted relative abundance (reference and current) for each species in 1 km<sup>2</sup> prediction grid in Alberta, Canada. The following fields are in the CSV files:

- “Row\_Col”: ID number defining raster row and column IDs (see Projection)
- “Ref”: normalized relative abundance under reference landscape conditions
- “Curr”: normalized relative abundance under current landscape conditions

Normalized relative abundance ( $Ref$ ,  $Cur$ ) for each pixel was calculated from original values ( $Ref_0$ ,  $Curr_0$ ) as:  $Ref = round(Ref_0/max(Ref_0, Curr_0))$ ,  $Cur = round(Curr_0/max(Ref_0, Curr_0))$ .

The Row\_Col field links the raster cells to these products:

- Geodatabase: [http://ftp.public.abmi.ca/species.abmi.ca/gis/Grid1km\\_working.gdb.zip](http://ftp.public.abmi.ca/species.abmi.ca/gis/Grid1km_working.gdb.zip)
- CSV with latitude/longitude (NAD\_1983\_10TM\_AEP\_Forest projection): [http://ftp.public.abmi.ca/species.abmi.ca/gis/Grid1km\\_working.csv.zip](http://ftp.public.abmi.ca/species.abmi.ca/gis/Grid1km_working.csv.zip)

# Chapter 4

## Sector effects

The effect of a sector on a species is the product of the area of the sector's footprint (width of the bars), and the average "per unit area" effect of that sector's footprint on the species (height of the bars). The area of footprint for a sector is the sum of the footprint belonging to it. The average effect is the difference between the predicted current and reference abundances in that sector's footprint type based on the species' model.

### Methods

We calculate the effects of different industrial sectors by predicting the abundance of each species in the current landbase versus a landbase in which the footprint associated with a particular industrial sector has been back-filled. 'Back-filling'[2] means replacing that sector's footprint with the vegetation or soil type that would have been present before the footprint was made. The difference between the total predicted population in the current and the back-filled landbase is the predicted effect of that sector on the abundance of a species in the region.

For the sector effect figures, we express the effect as the percent population change per unit area of the sector's footprint (height of the bar). The width of the bar for each sector shows the total area of the region occupied by that sector's footprint. The height multiplied by the width of the bar (i.e., the area of the bar = the effect per unit area of footprint multiplied by the area of footprint) is the total effect of the sector on the population of the species in the region. A bar with a large area means that that sector has had a large effect on the population — increasing the population if the bar is above 0, decreasing it if the bar is below 0.

The effect of an industrial sector on a species is determined by three factors: 1) How much area is occupied by the footprint of that sector, 2) How strongly — positively or negatively, or a mix of both — the species responds to each of the sector's footprint types, 3) How much of the sector's footprint is in higher- versus lower-quality habitat for the species. For example, a species that lives in old upland forest may be more affected by the forestry sector than the

energy sector, because forestry focuses on those older merchantable stands.

## Limitations

The sector effect results are based on each species' habitat models applied to the maps of vegetation or soils and footprint in the region, so all the limitations of the modeling and mapping apply to these results. There are additional caveats:

- Where footprint types overlap, such as a well site in a cultivated field, or a road through a forestry cutblock, we assign the sector footprint type as being 'on top', and its sector is credited with the effect of footprint in that overlap area.
- We cannot tell which roads are associated with particular sectors, such as roads to well sites or forestry cutblocks, so all roads are called transportation sector.
- Two sectors sometimes operate together to reduce cumulative effects, such as forestry cutblocks being placed where energy developments are planned. A forestry cutblock on a future wellsite, for example, would be assigned to forestry until the wellpad was built, when that area of footprint would change to the energy sector.
- We have difficulty separating some types of footprint, such as urban and industrial areas, so those types are assigned to only one of the two sectors involved.
- Sector effects only include direct effects of footprint, not indirect effects (e.g., pollution, noise, access effects) or possible cumulative effects where two or more sectors interact (e.g., roads allowing weeds into an area, where they can then colonize harvest areas).

## Downloadable results

Tables (separately for the northern and southern study regions) represent the total population effect (area of rectangles) and the "unit-effect" (y-axis value) for the different sectors (columns) per species (rows). The percent of human footprint (x-axis value) can be calculated as the ratio of total population effect and unit effect, or from this list:

- Agriculture sector: 2.13% in the north, 50.75% in the south.
- Forestry sector: 6.03% in the north, 0.71% in the south.
- Energy sector: 1.99% in the north, 1.99% in the south.
- Rural/Urban sector: 0.20% in the north, 1.88% in the south.
- Transportation sector: 0.52% in the north, 2.55% in the south.

# Chapter 5

## Intactness

ABMI's intactness is a measure of how much human footprint has affected species' abundances. We first use habitat models for each species to estimate its relative abundance in the current landbase, and in a reference landbase in which human footprint has been removed and back-filled with the vegetation that was there prior to the footprint. If the species is predicted to be less abundant in the current landbase than in the reference, intactness is defined as current abundance / reference abundance  $\times$  100%. An intactness of 80%, for example, would mean that the species is currently at 80% of the abundance expected in the reference landbase with no footprint. If the species is predicted to be more abundant in the current landbase, intactness is calculated as reference abundance / current abundance  $\times$  100%. In that case, an intactness of 80% would mean that the species is currently 1.25 times as abundant as expected in the reference landbase ( $1/1.25 \times 100\% = 80\%$ ). Intactness therefore declines as the current abundance differs from the reference abundance, regardless of whether the species is predicted to be less or more abundant currently.

Intactness of a taxon, like birds, is simply the average of the intactness of each species we can analyse in that taxon. Biodiversity intactness is the simple average of intactness for the 6 taxa: birds, mammals, mites, vascular plants, bryophytes and lichens. Because any difference between current and reference, positive or negative, lowers intactness of a species, species that increase with human footprint do not cancel out species that decrease with footprint.

For intactness maps, we calculate the intactness in each 1-km<sup>2</sup> pixel covering the province. For tables that summarize intactness for particular regions, we calculate each species' current and reference abundance by summing the 1-km<sup>2</sup> raster values in the region, then calculate intactness for the species from those totals. Species intactness values are then averaged to give the intactness of the taxon for the region, and the taxa are averaged for biodiversity intactness. For the wetland analysis, we can only calculate intactness using sampled wetlands. Wetland intactness maps therefore only show points at sampled wetlands across the province.

Intactness uses the total current and reference population in a reporting region. One human footprint type can have a positive effect on a species in a region and another type have a negative effect. These offsetting effects would keep intactness relatively high for that species, even though each footprint type has an effect.

## Limitations

Because intactness is based on the models for each species, it incorporates the limitations of those models . However, intactness is an average across many species, making error in individual species' models less important. A main limitation is that intactness is largely a function of human footprint. Maps of intactness are therefore highly correlated with maps of human footprint, regardless of the taxon. Intactness maps differ from human footprint maps only because some footprint types have greater or lesser effects on species in a taxon. Intactness does not show how much species have changed over time, because many factors besides habitat change due to human footprint can change species' populations.

# Chapter 6

## Richness

The richness index across all taxa was derived by taking a simple average of the standardized richness index (see below for explanation) for each 1 km<sup>2</sup> grid cell across the six taxa: mammals, birds, vascular plants, bryophytes, lichens and mites.

### Methods

Species richness for each taxon was modelled by stacking predictions from individual species habitat association models. Individual species models were built by relating the species' occurrence data to three sets of environmental variables: vegetation types, human footprint, and geographic location. Using these relationships, the mean occurrence probability of each species was projected for each 1 km<sup>2</sup> grid cell in the province. The species richness index for each taxon was produced by simply summing these predicted probabilities and scaling the values to range between 0 and 1 by dividing each 1 km<sup>2</sup> grid cell value by maximum value over all grid cells. For birds, relative abundance was rescaled to probability as  $1 - \exp(-abundance)$ , the probability of >0 abundance in a 1 ha area.

### Limitations

**The index of richness does not adjust for multiple habitat types** occurring within each 1 km<sup>2</sup> grid cell. For each species probability of occurrence was modeled for each habitat type, and an area-weighted average across habitat types determined for each 1 km<sup>2</sup> grid cell. By summing across species an index of richness was determined. This estimate was really just the average of point richness throughout each 1 km<sup>2</sup> grid cell, and did not incorporate the fact that most 1 km<sup>2</sup> grid cells span multiple habitat types and thus have higher richness than any single type alone, or having higher richness due to edge effects. However, a 1 km<sup>2</sup> grid cell dominated by habitats having relatively high richness will have higher index than a 1 km<sup>2</sup> grid cell dominated by habitats with low richness.

**Rare species were not included** when determining the index of richness. ABMI only models species habitat associations for species that have sufficient data to create robust models. Since the index of richness was created from the species models, the index only includes species with more than the minimum number of detections. It is possible that there are some habitat types may have a disproportionate number of rare species (e.g., possibly some wetland types) and our index of richness that does not incorporate rare species does not capture this.

**Spatial irregularities in sampling intensity** influence the index of richness. In regions with higher intensity of ABMI surveys, more species will have reached the minimum number of detections required for modeling, and thus more species will have been included in the index of richness. ABMI sampling intensity has been relatively low in the Rocky Mountain natural region, and in northwestern Alberta. Thus, species that are most commonly found in these regions will be less likely to have been included in the index of richness and 1 km<sup>2</sup> grid cell in these regions may have artificially low indices of richness.

# Chapter 7

## Uniqueness

Our measure of uniqueness was based on the degree to which species composition in each pixel was distinct compared to all other pixels within the Natural Region. Uniqueness was evaluated separately for each Natural Region to highlight unique areas with each region. If uniqueness had been calculated across natural regions, the boundaries between natural regions would have had the highest uniqueness simply because they contained species from both natural regions.

### Methods

The relative uniqueness measure shows the degree to which species composition in each pixel was distinct compared to all other pixels within the Natural Region. The uniqueness analysis was conducted separately for each Natural Region because the regional species pool and their predicted relative abundances differ between these. Relative abundance of each species in each pixel was predicted using habitat association models that relate species relative abundance to native habitat types. The uniqueness index was calculated based on two patterns of distribution in relative abundance of species: (1) the co-occurrence pattern among species pairs (species dissimilarity matrix) in the region and (2) the degree to which the relative abundance of species was clumped in the region. The species uniqueness for a given pixel in a region was calculated as:

$$SU_k = \sum_{i=1}^{s-1} \sum_{j=i-1}^{s-1} d_{ij} \sqrt{p_{ik} p_{jk}}$$

where  $SU_k$  is the uniqueness value for pixel  $k$ ,  $d_{ij}$  is the dissimilarity between species  $i$  and  $j$  in the natural region, and  $p_{ik}$  and  $p_{jk}$  the proportion of species  $i$  and  $j$  in site  $k$ . The species pairwise dissimilarity matrix was calculated using the relative abundance of species in the region and based on Bray-Curtis dissimilarity index. The relative abundance was standardized to species maximum so that rare and common species were equally weighted.

In general, highly restricted species (i.e., species with a patchy distribution) were less likely to occur with other species than were common or widely distributed species. This results in restricted species having relatively larger distance in the species pairwise association matrix. Moreover, by being restricted to a few places, most of the species regional population for a clumped species was concentrated in a few places resulting in the regional abundance of these species being relatively higher in the localities where they occur. The relative uniqueness index for a given locality was calculated by multiplying the distance between all species pairs and their respective regional proportion in the locality, and summing the products for all species pairs in the locality.

The analysis was conducted separately for each Natural Region, with the values in the region rescaled to between 0 and 1 by dividing each 1km<sup>2</sup> grid cell value by maximum value over all cells in the region.

## Limitations

**The uniqueness of species composition is a relative measure and values estimated for specific sites are partly dependent on the scale of study.** For example, a given area might be regarded less unique in a relatively homogeneous subset area where it is found but that particular subset area could have high uniqueness compared to a larger scale of study.

**The index of relative uniqueness is not comparable across natural regions.** For each taxon, relative uniqueness was estimated for each 1km<sup>2</sup> grid cell in each natural region, and values were rescaled to 0-1 range by dividing the value by the maximum of all values in the region. Since the maximum values differed among regions, the index was scaled differently among regions and thus is not comparable among different regions. However, the maximums were similar among forested regions (Boreal, Foothills and Shield natural regions), and for the non-forested regions (Parkland and Grassland natural regions). Thus, results have been presented separately for forest and non-forest regions.

**The data for uniqueness is based on predictions made using native habitat only.** The species habitat association analysis was conducted using native habitats only. This means that the role of other environmental (eg., climate) and spatial variables (eg., latitude/longitude) were not incorporated in the index.

**Relative uniqueness is not a measure of species composition.** The index of uniqueness identifies the degree to which species composition in a pixel differs from the regional centroid. Pixels with high uniqueness had species assemblages that differed greatly from the average, but this does not mean that they had similar species composition to each other.

**Relative uniqueness is sensitive to the dissimilarity measure and data transformation.** The index of uniqueness was strongly influenced by species that had clumped distributions. The species dissimilarity matrix was computed using a dissimilarity index that emphasized co-occurrences of species independent of species' rarity and commonness giving all species equal weight.

**Other measures of similarity/dissimilarity are possible** and different patterns of uniqueness may emerge using these.

## References

[1]: 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](http://abmi.ca).

[2]: Alberta Biodiversity Monitoring Institute. 2015. Alberta Backfilled Wall-to-Wall Vegetation Layer (Version 5) Metadata. Alberta Biodiversity Monitoring Institute, Alberta, Canada. Report available at: [abmi.ca](http://abmi.ca).

[3]: Alberta Biodiversity Monitoring Institute. 2015. Soil Layers Metadata. Alberta Biodiversity Monitoring Institute, Alberta, Canada. Report available at: [abmi.ca](http://abmi.ca).

[4]: Alberta Biodiversity Monitoring Institute. 2015. ABMI Human Footprint Inventory for 2012 conditions (Version 3). Alberta Biodiversity Monitoring Institute, Alberta, Canada. Report available at: [abmi.ca](http://abmi.ca)