

# **Boreal Avian Modelling Project**

## **Biophysical Database - Covariate Sampling Distribution**

Steve Cumming, Mélanie L. Le Blanc

June 25, 2010



© Samantha Song, Canadian Wildlife Service

## Contents

1	Introduction	4
2	Covariates	5

# 1 Introduction

Our ability to detect gradients or thresholds in avian to climatic or other covariates depends strongly on our sample having adequate contrast with respect to each covariate. This is a question of sample design in the technical sense. The power of a sample design to detect effects is related to the comparative distributions of the covariates in the sample and in the population. Our sample of locations comparative distributions of the covariates in the sample and in the population. Our sample of locations was not designed a priori, but it nevertheless represents a sample design of some sort, in this technical sense. The purpose of this part of the BAM Auto-Documentation System is to present an initial view of the sampling design of the covariates. As noted in Part 2, the current version of the database now contains 106 climatic covariates, and 25 remote-sensed covariates. Part 3 is presently restricted to the climatic covariates only. Remote sensed data have been excluded due to technical difficulties in processing these large datasets in R. The raw climatic data are interpolated to a regular 300 arc-second grid and to the exact locations of the BAM stations. The area of the grid cells is approximately 100km<sup>2</sup> in the southern latitudes. We define the population as the set of cell-level values that intersect the BAM boreal region. The sample is defined in two ways: at the station level and at the grid level. The station-level sample includes all the station level climate values interpolated to the exact station locations. The grid-level sample includes the unique cell-level values for all and only those 300 arc-second cells that include at least one station. These two definitions are different because stations that lie within the same cell are not independent samples of the climate surface. We conjecture that differences between values of stations within the same cell must largely reflect differences in elevation, due to the large distances between climate stations in most of the boreal. The sampling design for each climate covariate is presented as a figure of four panels. The first panel (upper left) is an image plot of the 300-arc second surface, with the station locations overlaid as points. The second panel is a histogram of the cell-level distribution in the population. The third and fourth panels (bottom left and right, respectively) show histograms of the station and cell-level samples. All three histograms are on the same scales, to facilitate comparison. The sample distribution panels also include the results of a Kolmogorov-Smirnov goodness-of-fit test between the sample and the population. The test statistic value and significance level are printed at the top of each histogram. It is evident that in all cases, the sampling and population distributions differ significantly. Further diagnostic tests could easily be included in the document generation procedure. We welcome specific suggestions from the BAM Technical Committee and other readers as to how this document could be made more informative.

## 2 Covariates















































































































































