COMPARING STRATIFICATION SCHEMES FOR AERIAL MOOSE SURVEYS

John R. Fieberg[1] and Mark S. Lenarz[2]

[1]Minnesota Department of Natural Resources, Biometrics Unit, 5463-C West Broadway, Forest Lake, Minnesota 55025; [2]Minnesota Department of Natural Resources, Forest Wildlife Populations and Research Group, 1201 East Highway 2, Grand Rapids, Minnesota 55744, USA.

ABSTRACT: Stratification is generally used to improve the precision of aerial surveys. In Minnesota, moose (*Alces alces*) survey strata have been constructed using expert opinion informed by moose density from previous surveys (if available), recent disturbance, and cover-type information. Stratum-specific distributions of observed moose from plots surveyed during 2005-2010 overlapped, suggesting some improvement in precision might be accomplished by using a different stratification scheme. Therefore, we explored the feasibility of using remote-sensing data to define strata. Stratum boundaries were formed using a 2-step process: 1) we fit parametric and non-parametric regression models using land-cover data as predictors of observed moose numbers; 2) we formed strata by applying classical rules for determining stratum boundaries to the model-based predictions. Although land-cover data and moose numbers were correlated, we were unable to improve upon the current stratification scheme based on expert opinion.

Aerial surveys are used to estimate moose (*Alces alces*) numbers and population parameters throughout much of North America, and in northeastern Minnesota they have been conducted annually since 1959 (Karns 1982). The Minnesota survey is based on a stratified random plot design, in which aerial units having similar expected moose density are grouped into strata to minimize the sampling variance of the population estimate. Variance among stratum-specific abundance estimates does not contribute to the sampling variance of the total population estimate, and thus, even poor stratification will usually improve precision compared to simple random sampling (Cochran 1977, Gasaway et al. 1986).

Gasaway et al. (1986) suggested using counts from pre-survey flights to form strata, but logistical and financial considerations have made pre-survey flights infeasible in Minnesota. Instead, moose survey strata have been constructed using expert opinion informed by moose density from previous surveys (when available), recent disturbance (e.g., logging, insect damage, blowdown, fire), and cover-type information; plots are periodically re-stratified to incorporate real or perceived changes in these factors. Despite best efforts, we have occasionally observed few moose in high density stratum plots or many moose in low or medium density stratum plots. It has generally been assumed that more formal numerical methods could be employed that would integrate cover-type data to better stratify the survey plots (see Moen et al. 2011). Presumably, better stratification would increase the precision of population estimates and increase the power to detect population trends (Lenarz et al. 2010, Lenarz 2011).

Moose abundance in Minnesota is currently estimated using a sightability model estimator, which inflates raw counts by

plot-level sampling probabilities as well as estimates of the probability of detecting each observed moose group (Steinhorst and Samuel 1989, Giudice et al. 2012). Three random processes create uncertainty in the estimated abundance: 1) the random sampling of survey plots; 2) random detection (and failed detection) of independent groups within surveyed plots; 3) uncertainty associated with parameters used to model detection probabilities. The coefficient of variation from past survey estimates (years 2004-2010) has ranged from 14-23% (mean = 17%). On average, 45, 16, and 39% of the variance has been attributed to sampling, detection, and model estimation processes, respectively. Our primary objective is to report on our efforts to develop alternative stratification schemes using cover-type composition data which has the potential to reduce the sampling variance component.

METHODS
Study Area

The aerial sampling frame used for the survey encompasses 15,056 km^2 of northeastern Minnesota (47°40'N, 91º25'W; Fig. 1). The forests were transitional between Canadian boreal forests and northern hardwood forests further south and includes all of the Boundary Waters Canoe Area Wilderness (BWCAW, Pastor and Mladenoff 1992). Wetlands including small to medium sized lakes, bogs, marsh, and small streams are interspersed on a low plateau that rises abruptly from Lake Superior to a crest approximately 700 m above sea level (Heinselman 1996). A northeast-southwest continental divide runs down the middle of the plateau with water flowing southeast into Lake Superior or northwest into the Hudsonian watershed.

The aerial sampling frame was a mosaic of conifer communities classified as the Northern Superior Upland section (Minnesota Department of Natural Resources [MNDNR] 2007). The landscape was characterized by northern white cedar (*Thuja occidentalis*), black spruce (*Picea mariana*), and tamarack (*Larix laricina*) on the lowlands, and balsam fir (*Abies balsamea*), white spruce (*P. glauca*), and jack (*Pinus banksiana*), white (*P. strobus*), and red pine (*P. resinosa*) on the uplands. Deciduous species, primarily quaking aspen (*Populus tremuloides*) and white birch (*Betula papyrifera*), occurred on the uplands in hardwood stands or were intermixed with conifers.

Although logging is not allowed in the BWCAW portion of the aerial sampling frame (Fig. 1), extensive disturbance occurred in 1999. A severe windstorm broke off or uprooted millions of trees in an area of roughly 1,900 km^2 (USDA Forest Service 2002). Large wildfires disturbed an additional 430 km^2 in 2006 and 2007 (Fites et al. 2007). Outside the BWCAW, extensive commercial logging has occurred on state, federal, and private lands.

Survey Data and Stratification Variables

We limited our analysis to surveys after 2005 when we made 3 major changes to the protocol. We switched from fixed-wing aircraft to helicopters, we shifted from using irregularly shaped survey plots to a grid of rectangular plots (4.3 x 8.0 km), and we changed from using double sampling to adjust for sightability bias (Gasaway et al. 1986) to the development and use of a sightability model (Anderson and Lindsey 1996, Quayle et al. 2001, Giudice et al. 2012). Prior to the 2005 and 2010 surveys, we re-stratified all survey plots. Low density plots were expected to contain ≤7 moose (0.2 moose/km^2), medium density plots 8-20 moose (0.2-0.6 moose/km^2), and high density plots ≥21 moose (0.6 moose/km^2).

As described in Lenarz et al. (2011), we modified a land use and land cover (LULC) raster layer provided by MNDNR (1998) for potential use in determining strata. The base source layer was derived from LANDSAT 30-meter thematic satellite imagery dated summers 1991-1996 and then divided into 16 classes based on imagery dated 1995 and
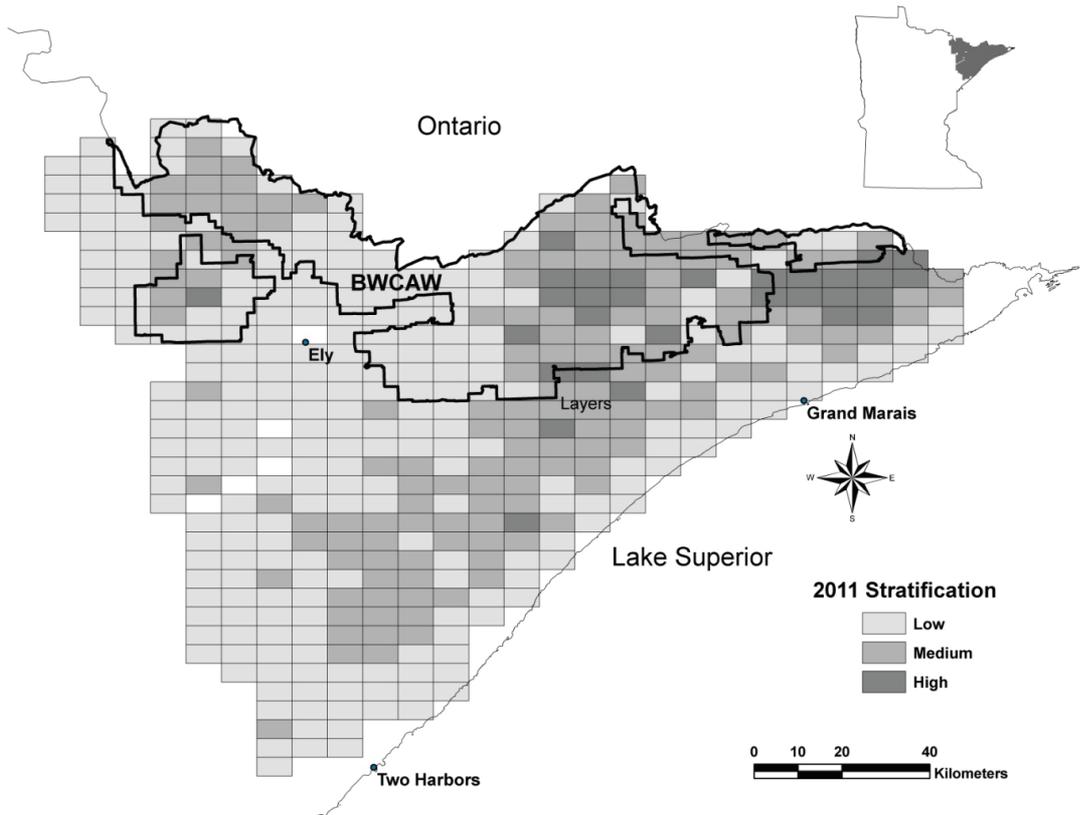
Fig.1. Sampling frame used for the aerial moose survey in northeastern Minnesota, USA, 2005-2011. Stratification indicated by shades of gray. White grid blocks (urban, iron mines) were expected to contain 0 moose and were not included in the survey.

1996. Forest disturbance data (P. T. Wolter, University of Wisconsin, Green Bay, unpublished data) were added to reflect areas of regenerating deciduous and coniferous forest following logging (1975-2000), as well as areas recovering from the 1999 windstorm event. We combined blowdown, hardwood regeneration, and conifer-regeneration classifications from Wolter's data with the cover type listed as "cutover" from the LULC layer into a single cover type to reflect disturbed habitats. Cutover represented areas where commercial timber was removed between 1980 and 1995. We also used 6 other cover types from the LULC layer: mixed, conifer, bog, deciduous, marsh, and water. In Arc-GIS 9.2 (Environmental Systems Research Institute, Redlands, CA, USA) we clipped the vegetative cover layer for each survey plot and computed the area (m^2) of each cover type within each plot. Finally, we calculated the proportion of the 7 cover types that made up each survey plot.

Alternative Stratification Schemes

Several classical rules have been developed for forming stratum boundaries using a single stratification variable, x, assumed to be correlated with the survey response variable, y. In particular, the cumulative square root frequency of sampling units ("$cum\sqrt{f}$") rule developed by Dalenius and Hodges (1959) is widely applied in practice (see Cochran 1977: 127-131 for a description and exemplification of the general approach). We illustrate the rule using x = the proportion of each sample plot that is classified as "disturbed" habitat. To apply this rule, one first forms a frequency

table containing counts of the number of observations falling into (usually, equally sized) class intervals; in our example, we considered 30 intervals (Table 1). Next, one constructs a column holding the square root of the frequency count, $\sqrt{f}$ (column 3 of Table 1), and a column holding the cumulative sum of $\sqrt{f}$, $cum\sqrt{f}$ (column 4 of Table 1). Stratum boundaries are chosen to give roughly equally sized intervals on the $cum\sqrt{f}$ scale. For example, to form 3 strata we would divide the total $cum\sqrt{f} = 75.15$ by $3 \approx 25$, resulting in the following 3 strata definitions: $cum\sqrt{f} \leq 25$ (stratum 1), $25 < cum\sqrt{f} \leq 50$ (stratum 2), and $cum\sqrt{f} > 50$ (stratum 3). Values of disturb = 0.1 and 0.201 create stratum boundaries that best approximate this goal.

Fabrizi and Trivisano (2007) suggested a method for determining stratum boundaries when multiple stratification variables are available along with prior-survey response data. To apply their approach, stratum boundaries are formed using a 2-step process: 1) a model is fit to prior survey data which allows prediction of survey responses from the group of available stratification variables; 2) stratum boundaries are formed by applying the $cum\sqrt{f}$ rule to these model-based predictions (rather than a single predictor, x). Following this general approach, we evaluated alternative stratification schemes that incorporated land cover-type composition data. We explored 2 different predictive modeling approaches using past survey counts as the response data: 1) we fit a Poisson regression model with linear and additive effects of land-cover variables (on the log scale); 2) we formed predictions using boosted regression trees (hereafter 'boosted trees'), a non-parametric method that tends to perform well in settings where interactions are prevalent or the relationship between response and predictor variables is highly non-linear (De'ath 2007, Elith et al. 2008). Boosted trees were among the best-performing methods for forming strata across a range of simulation scenarios considered by Fabrizi and Trivsano

Table 1. Application of the $cum\sqrt{f}$ rule with "disturb" (percent of the sample plot classified as disturbed habitat) as the stratification variable.

Class interval[1]	Frequency (f)[2]	$\sqrt{f}$	$cum\sqrt{f}$
[0,0.0164]	40	6.32	6.32
(0.0164,0.0332]	20	4.47	10.80
(0.0332,0.05]	17	4.12	14.92
(0.05,0.0668]	14	3.74	18.66
(0.0668,0.0836]	16	4.00	22.66
(0.0836,0.1]	18	4.24	26.90
(0.1,0.117]	26	5.10	32.00
(0.117,0.134]	13	3.61	35.61
(0.134,0.151]	12	3.46	39.07
(0.151,0.168]	12	3.46	42.54
(0.168,0.184]	21	4.58	47.12
(0.184,0.201]	14	3.74	50.86
(0.201,0.218]	6	2.45	53.31
(0.218,0.235]	8	2.83	56.14
(0.235,0.252]	7	2.65	58.78
(0.252,0.268]	5	2.24	61.02
(0.268,0.285]	6	2.45	63.47
(0.285,0.302]	5	2.24	65.71
(0.302,0.319]	6	2.45	68.16
(0.319,0.336]	1	1.00	69.16
(0.336,0.352]	1	1.00	70.16
(0.352,0.369]	1	1.00	71.16
(0.369,0.386]	1	1.00	72.16
(0.386,0.403]	1	1.00	73.16
(0.403,0.42]	1	1.00	74.16
(0.42,0.436]	0	0.00	74.16
(0.436,0.453]	0	0.00	74.16
(0.453,0.47]	0	0.00	74.16
(0.47,0.487]	0	0.00	74.16
(0.487,0.504]	1	1.00	75.16

[1] Class interval for the "disturb" land-cover variable (describing the amount of disturbed habitat in a sample plot). The (x1, x2] format indicates the interval includes all values > x1 and ≤ x2.

[2] Count of the number of observations falling within the class interval.

(2007). Boosted trees require specification of several tuning parameters; see Appendix I for model fitting details. In addition, we recom-

mend De'ath (2007) and Elith et al. (2008) for an introduction to boosted trees.

We evaluated the alternative stratification schemes by constructing side-by-side box plots illustrating stratum-specific distributions of observed moose. In addition, we compared the expected precision of a stratified random sample of size 40, with (stratum-specific) sampling rates determined using optimal allocation formulas (Cochran 1977). Without any adjustment, these comparisons would result in optimistic performance measures for the model-based stratification schemes since the same data would be used for model fitting and evaluation. By contrast, strata based on expert opinion were formed prior to data collection (although past survey data were occasionally used to change stratum assignments, these temporal changes were updated in our survey data only for future, and not prior years). We used a 10-fold cross-validation procedure to allow for a more fair comparison of the stratification schemes. Specifically, we split the data into 10 groups and fit both models (Poisson regression, boosted tree) 10 times, leaving out 1/10 of the data on each occasion. We then used the fitted models to predict moose numbers for the observations not used in the model-fitting process. These "out-of-sample" predictions were combined into a new dataset and the $cum\sqrt{f}$ rule was applied to form strata. The out-of-sample predictions were also used to allocate sampling effort among strata, but the expected precision was calculated using the (true) observed moose data.

We used functions in Program R (R Development Core Team 2010) for model fitting and for testing stratification schemes. We used the 'glm' function to fit the Poisson regression models. We used functions in the dismo and gbm packages (Ridgeway 2010, Hijmans et al. 2011) to fit the boosted tree model. Lastly, we used functions in the stratification package (Baillargeon and Rivest 2011) to apply the $cum\sqrt{f}$ rule to model-based predictions, specifying that we desired 3 strata.

RESULTS

Between 2005 and 2011, 237 plots were surveyed as part of the annual moose survey in northeastern Minnesota. Cover types were ubiquitous across the survey plots examined; 86% of the surveyed plots contained all 7 cover types and the remaining 14% had 6 cover types. Disturbed, conifer, marsh, mixed, and water were found in 100% of the survey plots examined; bog was found in 99% and deciduous 88% of the plots.

Both the Poisson regression model and the boosted tree model suggested the 7 land-cover predictors were correlated with observed moose numbers. Regression coefficients in the Poisson model were all positive and statistically significant at the $P = 0.01$ level, suggesting moose were more likely to be found in plots that contain these 7 cover types, relative to those that have more "shrub" or "other" (the only 2 land-use categories not included in the model).

Functional relationships are more difficult to infer with boosted tree models, although various tools and metrics have been developed to help interpret model fit and predictor importance (again, refer to De'ath 2007 and Elith et al. 2008 for useful reviews). In particular, partial dependence plots can be used to graphically explore how predictions vary with each covariate (Friedman 2001). The partial dependence plot for covariate X_i is formed by averaging predictions across all observations while holding X_i constant. These plots indicate fairly linear responses (up to or down to a threshold) in relation to the relative abundance of bog, conifer, mixed, and disturbed cover types (Fig. 2); predicted moose numbers decreased with increases in % bog and increased with % conifer, % mixed, and % disturbed. Responses to the other covariates were non-linear, with modal responses observed for % marsh and % water. Indices of predictor importance (given in the x-axis labels in Fig. 2) were large for all predictors, but suggested that % bog was the most infor-
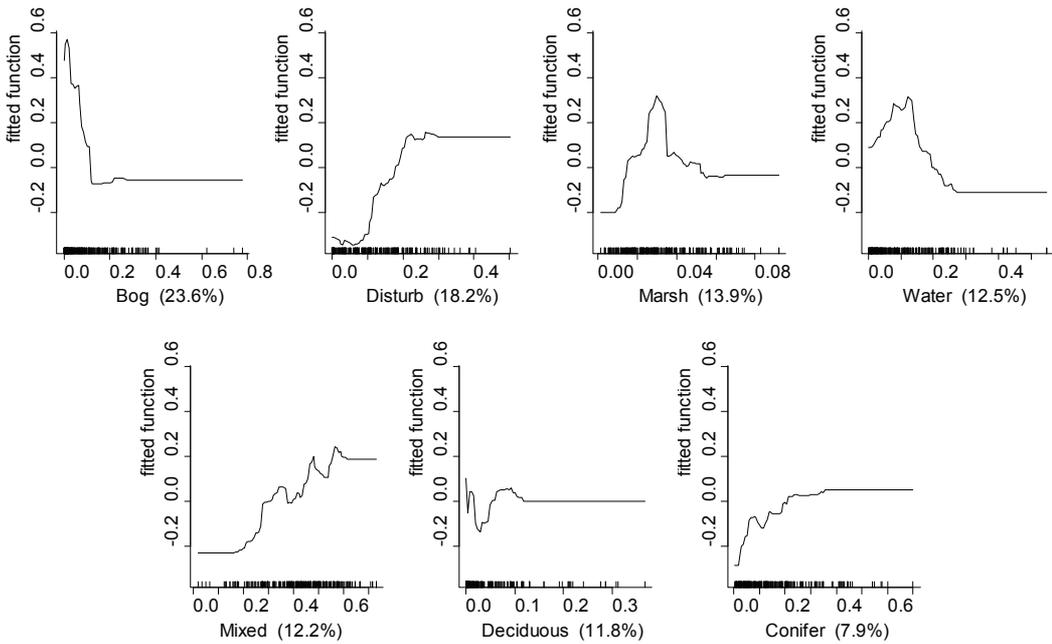
Fig. 2. Partial dependency plots depicting marginal predictor-response relationships from the boosted tree model. The partial dependence plot for covariate X_i is formed by averaging predictions across all observations while holding X_i constant. Tick marks along the x-axis depict the distribution of observed data. Percentages incorporated in the x-axis labels give predictor importance measures.

mative predictor.

Despite a few outliers, side-by-side box plots of the distribution of observed moose in each stratum indicated that the current expert opinion process has been effective at discriminating among sample plots (Fig. 3a). The boosted tree model resulted in a better stratification scheme than the Poisson regression model (Fig. 3d, e versus Fig. 3b, c). Further, when the same data were used to fit and evaluate the model, the boosted tree approach resulted in a smaller expected SE than the current stratification scheme (264 versus 289; Fig. 3a, d). This evaluation, however, assumes that residual error (i.e., differences between current survey data and predictions from models fit to these same data) will provide an adequate approximation to prediction error in future surveys. This assumption might be reasonable if no new plots will be sampled in future years and moose counts are relatively constant over time (within a plot). We examined moose counts of 52 sample units that were surveyed multiple times, and although counts

for some plots were relatively constant, others exhibited a high degree of variability (Fig. 4). Thus, the cross-validation approach, which approximates future survey prediction error by comparing model predictions to survey data not used in the model fitting process, should provide a more accurate evaluation of future performance. The use of out-of-sample predictions (from cross-validation) degraded the performance of both model-based stratification schemes (compare Fig. 3c, e to Fig. 3b, d), and the expected SE for the boosted regression tree approach (350) was no longer smaller than the corresponding SE for the current expert opinion based system (289). Thus, we conclude that the boosted tree approach may be competitive with (but not a clear improvement over) the current stratification scheme.

DISCUSSION

The precision of population estimates obtained from a stratified random survey will be influenced by several survey design choices, including the number of strata, choice of strati-

fication variables and stratum boundaries, and sample allocation among strata. We focused here on the choice of stratification variables and stratum boundaries, assuming 3 strata were desired and that optimal sampling formulas would be used to allocate sampling effort among strata. We followed Fabrizi and Trivisano's (2007) general approach to strata formation, but were unable to demonstrate a clear improvement over the current stratification scheme. This (negative) result is likely due, in part, to the effectiveness of the current stratification scheme based on expert opinion (Fig. 3a), but may also reflect the difficulty of predicting moose numbers from coarse-level land-cover layers developed from remote sensing data that were acquired several years prior to the survey. Nonetheless, the 2-step stratification approach of Fabrizi and Trivisano (2007) may prove useful in other applications. The approach is transparent, repeatable, and easy to apply in practice.

Although the current stratification scheme for Minnesota's aerial moose survey is determined by expert opinion, land cover data are considered during this process (albeit more qualitatively). Past survey data and other qualitative information also play a role, and these additional information sources can be difficult to incorporate more formally into a quantitative approach (e.g., using past data as a predictor would result in much missing
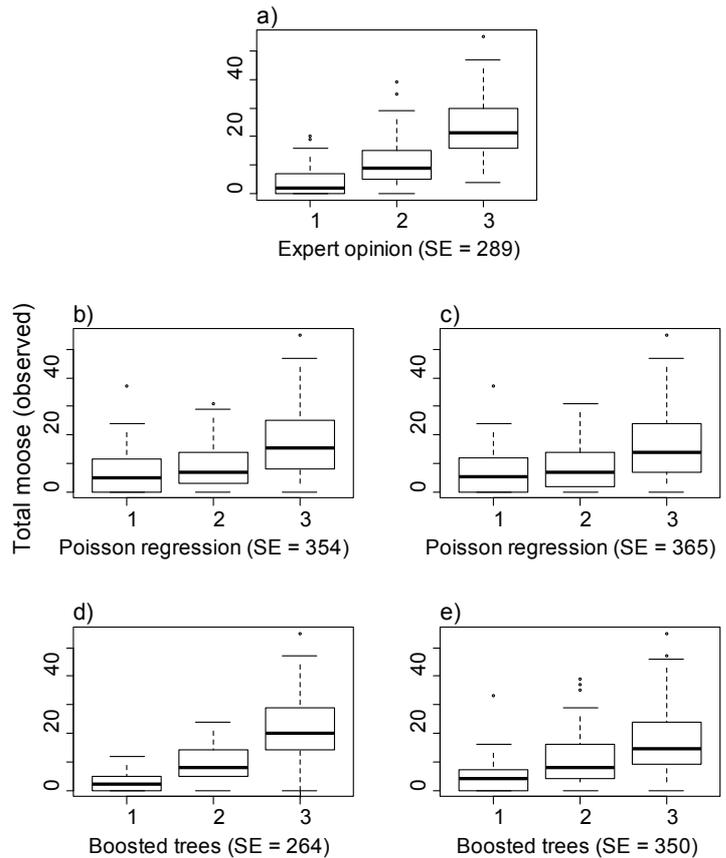
Fig. 3. Side-by-side box plots of total moose observed in survey plots by stratum in northeastern Minnesota, USA, 2005-2011 for: a) the current expert opinion based stratification scheme; b) the Poisson regression model based stratification scheme (using the same data to fit models and determine strata); c) the Poisson regression model based stratification scheme using out-of-sample predictions (from cross validation); d) boosted tree stratification scheme (using the same data to fit models and determine strata); e) boosted tree stratification scheme using out-of-sample predictions (from cross validation). The x-axis labels give the expected standard error associated with an estimate of the total number of observed moose from a sample of 40 plots allocated using optimal allocation formula.

data). Lastly, we note that current stratum assignments exhibit a high degree of spatial correlation (Fig. 1). In theory, one could incorporate spatial predictors or use models that allow for spatial correlation in step 1 of Fabrizi and Trivisano's (2007) approach to forming strata, but doing so would complicate the analysis considerably.
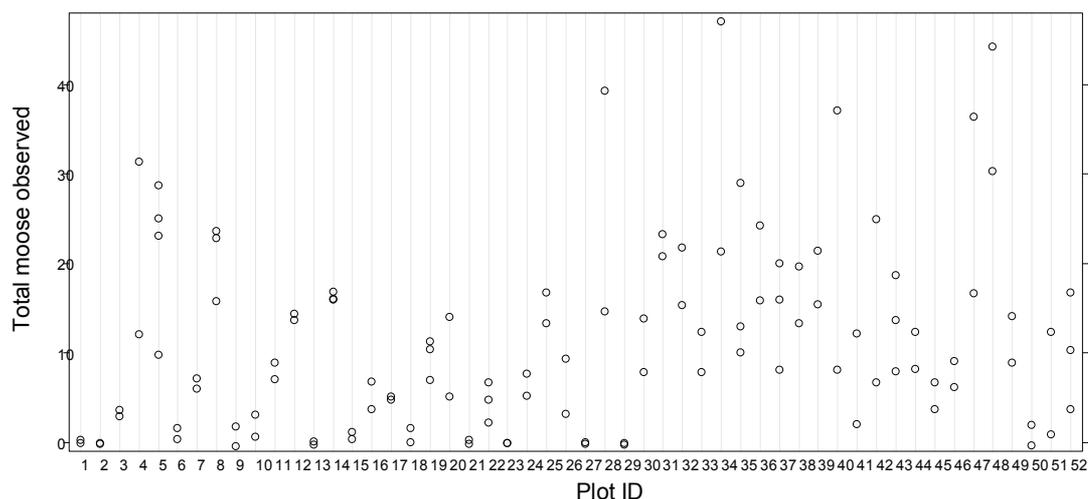
Fig. 4. Total moose observed for plots that were sampled more than once between 2005 and 2010. Each vertical line corresponds to a different sample plot. Observations have been jittered slightly to allow viewing of multiple observations with the same value (e.g., two observations of 0 moose).

REFERENCES

ANDERSON, C. R., and F. G. LINDZEY. 1996. Moose sightability model developed for helicopter surveys. Wildlife Society Bulletin 24: 247-259.

BAILLARGEON, S., and L-P. RIVEST. 2011. Stratification: univariate stratification of survey populations. R Package version 2.0-3. <http://CRAN.R-project.org/package=stratification> (accessed August 2011).

COCHRAN, W. G. 1977. Sampling Techniques. John Wiley and Sons, New York, New York, USA.

DALENIUS, T., and J. L. HODGES., JR. 1959. Minimum variance stratification. Journal of the American Statistical Association 54: 88-101.

DE'ATH, G. 2007. Boosted trees for ecological modeling and prediction. Ecology 88: 243-251.

ELITH, J., J. R. LEATHWICK, and T. HASTIE. 2008. A working guide to boosted regression trees. Journal of Animal Ecology 77: 802-813.

FABRIZI, E., and C. TRIVISANO. 2007. Efficient stratification based on nonparametric regression methods. Journal of Official Statistics 23: 35-50.

FITES, J. A., A. REINER, M. CAMPBELL, and Z. TAYLOR. 2007. Fire behavior and effects, suppression, and fuel treatments on the Ham Lake and Cavity Lake fires. USDA Forest Service, Superior National Forest, Eastern Region, Milwaukee, Wisconsin, USA. <http://www.fs.fed.us/adaptivemanagement/projects/FBAT/docs/HamLake07_22_08.pdf> (accessed July 2011).

FRIEDMAN, J. H. 2001. Greedy function approximation: a gradient boosting machine. Annals of Statistics 29: 1189-1232.

GASAWAY, W. C., S. D. DuBOIS, D. J. REED, and S. J. HARBO. 1986. Estimating moose population parameters from aerial surveys. Biological Papers of the University of Alaska, No. 22. Fairbanks, Alaska, USA.

GIUDICE, J. H., J. R. FIEBERG, and M. S. LENARZ. 2012. Spending degrees of freedom in a

poor economy: a case study of building a sightability model for moose in north-eastern Minnesota. Journal of Wildlife Management 76: 75-87.

HEINSELMAN, M. 1996. The Boundary Waters Wilderness Ecosystem. University of Minnesota Press, Minneapolis, Minnesota, USA.

HIJMANS, R. J., S. PHILLIPS, J. LEATHWICK, and J. ELITH. 2011. dismo: Species distribution modeling. R package version 0.7-2. <http://CRAN.R-project.org/package=dismo> (accessed August 2011).

KARNS, P. D. 1982. Twenty-plus years of aerial moose census in Minnesota. Alces 18: 186-196.

LENARZ, M. S. 2011. 2011 Aerial Moose Survey. Minnesota Department of Natural Resources, St. Paul, Minnesota, USA. <http://files.dnr.state.mn.us/outdoor_activities/hunting/moose/moose_survey_2011.pdf> (accessed July 2011).

_____, J. FIEBERG, M. W. SCHRAGE, and A. J. EDWARDS. 2010. Living on the edge: viability of moose in northeastern Minnesota. Journal of Wildlife Management 74: 1013-1023.

_____, R. WRIGHT, M. W. SCHRAGE, and A. J. EDWARDS. 2011. Compositional analysis of moose habitat in northeastern Minnesota. Alces 47: 135-149.

(MNDNR) MINNESOTA DEPARTMENT OF NATURAL RESOURCES. 1998. Landsat-Based Land Use-Land Cover (raster). Minnesota Department of Natural Resources, St. Paul, Minnesota, USA. <http://deli.dnr.state.mn.us/metadata.html?id=L250000120604> (accessed April 2011).

_____. 2007. Ecological Classification System. Minnesota Department of Natural Resources, St. Paul, Minnesota, USA. <http://www.dnr.state.mn.us/ecs/index.html> (accessed April 2011).

MOEN, R., M. E. NELSON, and A. J. EDWARDS. 2011. Radiotelemetry locations, home ranges, and aerial surveys in Minnesota. Alces 47: 101-112.

PASTOR, J., and D. J. MLADENOFF. 1992. The southern boreal northern hardwood forest border. Pages 216-240 in H.H. Shugart, R. Leemans, and G. B. Bonan, editors. A Systems Analysis of the Global Boreal Forest. Cambridge University Press, Cambridge, England.

QUAYLE, J.F., A. G. MACHUTCHON, and D. N. JURY. 2001. Modeling moose sightability in south-central British Columbia. Alces 37: 43-54.

R DEVELOPMENT CORE TEAM. 2010. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. <http://www.R-project.org/> (accessed August 2011).

RIDGEWAY, G. 2010. gbm: Generalized Boosted Regression Models. R package version 1.6-3.1. <http://CRAN.R-project.org/package=gbm> (accessed August 2011).

STEINHORST, R. K., and M. D. SAMUEL. 1989. Sightability adjustment methods for aerial surveys of wildlife populations. Biometrics 45: 415-425.

USDA FOREST SERVICE. 2002. After the Storm: A Progress Report: July 2002. USDA Forest Service, Superior National Forest, Eastern Region, Milwaukee, Wisconsin, USA. <http://www.fs.fed.us/r9/forests/superior/storm_recovery/afterstorm/> (accessed July 2011).