

**AERIAL LINE TRANSECT SURVEY PROTOCOLS  
AND DATA ANALYSIS METHODS TO MONITOR  
MOOSE (*Alces alces*) ABUNDANCE AS APPLIED ON  
THE INNOKO NATIONAL WILDLIFE REFUGE,  
ALASKA**

**Prepared By:**

**Ryan M. Nielson  
and Lyman L. McDonald  
Western EcoSystems Technology, Inc.  
2003 Central Ave  
Cheyenne, WY 82001**

**and**

**Steven D. Kovach  
Innoko National Wildlife Refuge  
US Fish and Wildlife Service  
PO Box 69  
McGrath, AK 99627**

**January 2006**



**AERIAL LINE TRANSECT SURVEY PROTOCOLS AND DATA  
ANALYSIS METHODS TO MONITOR MOOSE (*Alces alces*)  
ABUNDANCE AS APPLIED ON THE INNOKO NATIONAL  
WILDLIFE REFUGE, ALASKA**

By:  
Ryan M. Nielson<sup>1</sup>, Lyman L. McDonald<sup>1</sup>, and Steven D. Kovach<sup>2</sup>  
January 2006

**REPORT CITATION:**

Nielson, R. M., L. L. McDonald, and S.D. Kovach. 2006. Aerial line transect survey protocols and data analysis methods to monitor moose (*Alces alces*) abundance as applied on the Innoko National Wildlife Refuge, Alaska. Technical report prepared for US Fish and Wildlife Service, McGrath, Alaska, USA.

Mention of specific products or trade names does not constitute endorsement by the US Fish and Wildlife Service, Department of Interior, or the US Government.

<sup>1</sup> Western EcoSystems Technology, Inc., 2003 Central Ave., Cheyenne, WY 82001

<sup>2</sup> Innoko National Wildlife Refuge, US Fish and Wildlife Service, PO Box 69, McGrath, AK 99627

## Table of Contents

<b>1.0 BACKGROUND.....</b>	<b>5</b>
1.1 Introduction.....	5
1.2 History of the Project.....	5
<b>2.0 Line Transect Sampling .....</b>	<b>7</b>
2.1 Purposes and Objectives .....	7
2.2 Description .....	7
2.3 Theory .....	8
2.4 Critical Assumptions and Weaknesses .....	9
2.5 Other Uses of Line Transect Sampling .....	11
2.6 Other Potential Methods.....	11
2.7 Comparison of Line Transect Sampling to Plot Based Procedures .....	13
<b>3.0 Methods and Protocols .....</b>	<b>14</b>
3.1 Sampling .....	14
3.1.1 Stratification .....	14
3.1.2 Transect Selection .....	16
3.1.3 Aircraft .....	17
3.2 Map and Waypoint Preparation.....	17
3.3 Field Procedures .....	18
3.3.1 Checklist .....	19
3.3.2 Safety.....	20
3.3.3 Required Snow Conditions .....	21
3.3.4 Survey Airspeed and Altitude .....	21
3.3.5 Crew Duties .....	21
3.3.6 Recording Survey Conditions .....	22
3.3.7 Measuring Distance to Moose .....	22
3.3.8 Moose Observations.....	24
3.3.9 Double Sampling .....	25
3.3.10 Data Management .....	26
3.4 Data Analysis .....	27
3.4.1 Data Preparation.....	27
3.4.2 Estimating Moose Abundance.....	27
3.4.3 Assessing Survey Accuracy .....	32
3.4.4 Group Size Alternatives .....	32
3.4.5 Bootstrapping.....	33
3.4.6 Data Pooling .....	33
3.4.7 Assumptions and Availability Bias .....	34
3.4.8 Census Counts .....	34
3.5 Estimation of Change and Trend.....	34
3.6 Analysis examples.....	35
<b>4.0 Literature Cited.....</b>	<b>36</b>
<b>Appendix 1: Example Field Maps .....</b>	<b>41</b>
<b>Appendix 2: Field Data Form .....</b>	<b>44</b>
<b>Appendix 3: Voice/GPS Survey Program Manual .....</b>	<b>45</b>
INSTALLATION.....	45
Setup .....	45
Moving Map Option.....	47
TRANSCRIBE Program .....	49
Creating a File of Header Names .....	49

Creating a file of Species Names .....	49
Selecting the Voice File .....	50
Transcribing Voice Recordings .....	50
Deleting, Inserting and Editing in the Observation List .....	50
Changing the Font Size in the Species List or the Observation List .....	50
Saving the Transcribed Observations .....	51
File Management Considerations .....	51
File Examples .....	52
<b>Appendix 4: Data Files .....</b>	<b>54</b>
A.4.1 Data Files .....	54
Transect_File .....	54
Survey_Observations .....	54
Transects_Flown .....	56
Transects_Flown_Minus_Islands .....	58
Bootstrap_Distance_Output_Combined .....	59
Generated Files .....	59
A.4.2 Presenting Results .....	60
A.4.3. Transects Flown AML .....	61
<b>Appendix 5: SAS and R Code .....</b>	<b>63</b>
A.5.1 SAS Code for Data Analysis .....	63
A.5.2 R Code for Histograms .....	67
A.5.3 SAS Code for Logistic Function .....	67
A.5.4 SAS Code for Bootstrap .....	68
A.5.5 SAS Code for Merging Bootstrap Information and Calculating Standard Errors, Bias, and Confidence Intervals .....	77
<b>Appendix 6: Using the Program DISTANCE .....</b>	<b>80</b>
A.6.1 Using DISTANCE to Fit a Detection Function .....	80
A.6.2 Using DISTANCE When Bootstrapping .....	103

# 1.0 BACKGROUND

## 1.1 Introduction

This document constitutes the protocol for planning, conducting, and analyzing aerial line transect surveys for moose (*Alces alces*) abundance as applied on Innoko National Wildlife Refuge (NWR), Alaska. Recommendations in this protocol on survey design, data collection, and statistical analysis are based on the line transect sampling literature and experiences gained while conducting moose surveys on Innoko NWR and similar areas. This document is accompanied by a CD which contains an example datasheet, versions 3.1 and 6.2 of the Voice/GPS Survey program files and manual (John Hodges, Migratory Bird Management, US Fish and Wildlife Service, Juneau, AK), example data files, and code for data analysis in both SAS (SAS Institute 2000) and R (R Development Core Team 2005). Copies of all figures used in this document are also contained on the CD, as are Microsoft Word and Adobe Acrobat PDF copies of this document.

This protocol begins with a brief history of moose monitoring on Innoko NWR, including an evolution of the aerial survey method and the involvement of Western EcoSystems Technology, Inc. (WEST, Inc.) on the project. The purposes and objectives of using line transect sampling to monitor moose populations on Innoko NWR are described. Next, a brief description of line transect sampling theory, assumptions, and inference is given. This section on theory and inference is brief because there exists extensive documentation in the literature. For examples see Buckland et al. (1993), Buckland et al. (2001), Borchers et al. (2002), and Buckland et al. (2004). A brief comparison of line transect sampling to other potential methods for estimating moose abundance is provided. Procedures, recommendations, and cautions for line transect survey design, data collection, and statistical analyses are presented. The statistical analysis section addresses potential choices that must be made when analyzing data from a moose survey, and provides guidance on how to make well-informed decisions along the way. Appendices 1 – 6 contain example survey maps, data sheet, and data files to help the researcher with survey preparation and data management, a manual for operating the Voice/GPS Survey program used during survey operation, SAS and R code for analysis of survey data, and detailed instructions for using the program DISTANCE (Thomas et al. 2004) to model probability of detection of moose during the survey, a key component of line transect data analysis.

## 1.2 History of the Project

This history of moose surveys on Innoko NWR was reconstructed from files, unpublished data, and annual narratives on file at Innoko NWR, McGrath, Alaska, as well as Skinner et al. (1997). Innoko NWR began flying moose surveys in 1982. Three survey plots, approximately 52 km<sup>2</sup> (20 mi<sup>2</sup>) in size, were established with the help of the Alaska Department of Fish and Game (ADFG). Two of the plots were near the Iditarod River; the third plot was along the Innoko River near the junction with the Iditarod River. All 3 plots were flown in December 1982, December 1983, and March 1984; the Innoko River plot was also flown in January 1983, March 1983, and December 1984. Starting with the December 1984 survey, 2 additional plots were located along the Innoko River that reached up to Cripple Landing. ADFG repeated the upper Innoko River (Cripple Landing-Camp Lake) plot in November 1985.

Innoko NWR conducted its first survey to estimate moose density in fall 1987 following procedures outlined by Gasaway et al. (1986). This survey was conducted in the southwest portion of Innoko NWR where a radio telemetry study was being conducted at the time. This survey was only partially completed due to adverse weather conditions. The next survey conducted in fall 1989 was another Gasaway et al. (1986) survey. This survey was focused on the northern half of Innoko NWR; however, adverse weather limited the survey to the eastern portion of the survey area. This was also the first time that Innoko NWR was divided into north and south halves for the purposes of surveying moose. No moose surveys were conducted in 1981, 1986 and 1988.

In the fall of 1990, Innoko NWR attempted to estimate moose density using techniques other than those used by ADFG (i.e., the ‘Gasaway’ method mentioned above). This survey, again in the northeast portion of Innoko NWR, utilized strips flown with a helicopter rather than fixed-wing aircraft. The survey area was divided into 3 strata – river, muskeg, and hills. One of the objectives of the survey was to obtain a density estimate for moose with a 90% confidence interval (CI) of  $\pm 25\%$  or less of the point estimate. Adverse weather prohibited completion of the survey. April 1991 saw a second effort to carry out a helicopter survey flying strips in the northeast portion of Innoko NWR. This time the weather cooperated and the survey was completed with a 95% CI of  $\pm 19.8\%$  of the point estimate. The success of this effort convinced Innoko NWR biologists that line strips or transects flown with a helicopter may in fact be a viable alternative to the Gasaway et al. (1986) technique currently used by ADFG.

Innoko NWR contacted Dr. Lyman McDonald of WEST, Inc. later in 1991 to assist with the development of a moose survey protocol that employed line transect methods. Dr. McDonald brought his experience from designing line transect surveys for caribou (*Rangifer tarandus*) (Drummer and McDonald 1987), polar bears (*Ursus maritimus*) (Manly et al. 1996; McDonald et al. 1999), and walrus (*Odobenus rosmarus divergens*) (Gilbert et al. 1990; Gilbert 1999) to the project. An initial protocol was developed and peer reviewed by Drs. David Anderson and Jeff Laake in early 1992 (Innoko NWR files). March 1992 saw the first attempt at using line transect methods flown with a helicopter to estimate moose abundance on Innoko NWR.

No surveys were conducted in 1993, as data from the 1992 survey was analyzed further and modifications were being made to the 1992 protocol. At this time a plan was developed to test the revised protocol not only with a helicopter, as was used since 1990, but to compare the results against those obtained from Piper Super Cub (PA-18) and Cessna (C-185) fixed-wing aircraft during the 1994 spring survey. In 1993, Mr. Wallace Erickson of WEST, Inc. became the lead statistician on this project.

Results of the 1994 surveys comparing helicopter data against fixed-wing data found that not only was the data from the helicopter better (e.g., more moose observed), but surveys from fixed-wing aircraft violated some of the visibility assumptions of line transect methods (Skinner et al. 1997). By 1996, most of the protocol for estimating moose abundance on Innoko NWR via the line transect method had been developed (Skinner et al. 1997). However, new techniques and equipment have continued to be tested, and recommendations have been made to improve the safety, efficiency, reliability, and precision of the surveys. For example, in 2000 double-observer procedures were used to assess violation of the assumption of 100% detection of moose

on or near the transect line, a primary assumption of line transect sampling. Less than 100% detection was found during the 2000 survey, thus a double-observer modification was integrated into the standard protocol for moose surveys on Innoko NWR. In 2004, laser range finders were used to accurately measure perpendicular distances to observed moose, an improvement on the original technique of using distance bins based on window markings. In addition to the double-observer method and use of laser range finders, adjustments in stratification have made the surveys and analysis more efficient. In 2002 Mr. Ryan Nielson of WEST, Inc. joined the project, assisting Mr. Erickson. Mr. Nielson was then assigned as the lead research statistician for this project in 2004.

## **2.0 Line Transect Sampling**

### **2.1 Purposes and Objectives**

Line transect aerial surveys for moose on Innoko NWR were developed for two purposes: 1) to obtain unbiased estimates of moose abundance with desired levels of precision; and 2) to obtain unbiased locations of moose to use in developing habitat selection models (Skinner et al. 1997). An original objective of evaluating the line transect technique at Innoko NWR was “to develop safe, reliable, efficient, and cost-effective procedures that obtain unbiased estimates of both moose populations and trends in large areas...containing in part, dense coniferous forest, with a 90% confidence interval (CI) of  $\pm 25\%$  or less of the estimate with a crew of 4 to 6 people in few total flight hours (40 or less for 3,000 square miles),” (Skinner et al. 1997:3). Experience has shown that coefficients of variation ( $100\% \times [SE/Estimate]$ ) range from 12.9% to 19.1%. This level of precision should allow managers to detect a 38% change in abundance using a 90% CI (equivalent to a significance level of  $\alpha = 0.10$ ).

Unbiased estimates of moose abundance can be used to set regulations for moose harvest. Trends in moose abundance can also be estimated to help management regulate moose harvest or potentially identify negative impacts from habitat loss or degradation. Obtaining unbiased estimates of moose locations along with moose abundance is one of the principal advantages of line transect sampling, although other survey procedures could be modified so they to yield unbiased estimates of moose locations. Estimates of moose locations can be related to habitat characteristics to describe landscapes selected or avoided by moose (Manly et al. 2002).

### **2.2 Description**

Basic concepts and background information for line transect (distance) sampling at Innoko NWR closely follow Buckland et al. (1993), Buckland et al. (2001), and literature cited therein. These authors describe line transect sampling as a plotless technique for sampling biological populations that can be practical, efficient, and relatively inexpensive. A basic description of line transect sampling is that a transect line is randomly located in some manner and an observer(s) travels the line. Individual animals or groups of individuals observed are recorded, along with group size, and either the perpendicular distance of the individual or group from the transect line (Fig. 1) or the sighting distance and sighting angle. Not all individuals that are potentially detectable from the line will be seen; generally, the further the individual or group is from the transect line, the greater the probability that the individual will be missed. The sighting

distance/sighting angle method, while a viable option for some surveys, is not recommended for aerial moose surveys and is not considered further in this protocol.

Line transect methods were largely developed for surveying sparsely distributed populations (Buckland et al. 2001), thus are well suited for estimating moose abundance, especially in a large region (Brochers et al. 2002). The method is regarded as “simple, economical, and relatively precise” (Cassey and McArdle 1999). Line transect methods allow for calculating moose density with a measure of precision. This enables managers to compare density estimates between surveys from other areas or points in time.

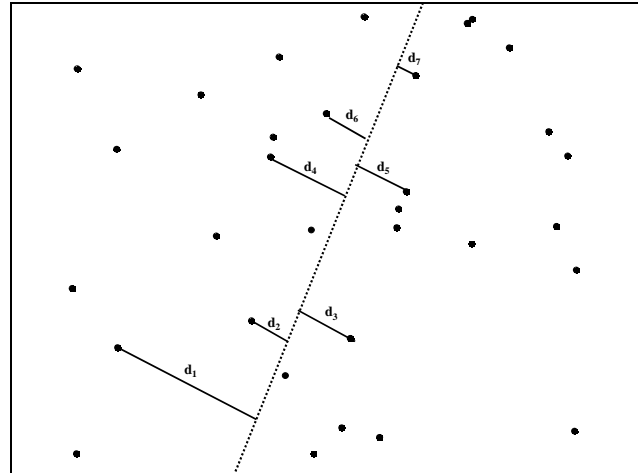


Fig. 1. Example of line transect sampling with a single, randomly placed line. Seven objects were detected at distances  $d_1, d_2, \dots, d_7$ . Modified from Buckland et al. (2001:4).

### 2.3 Theory

Data are recorded on groups of moose, including groups of size 1 (i.e., one individual). If placement of transects are random relative to the locations of groups, and the probability of detection of a group on the transect line (or on the inside edge of the survey strip) is 1.0 (Fig. 2A), or can be estimated (Fig. 2B), then the density of moose within the survey area can be estimated. Random placement of transects relative to group locations implies that if all groups

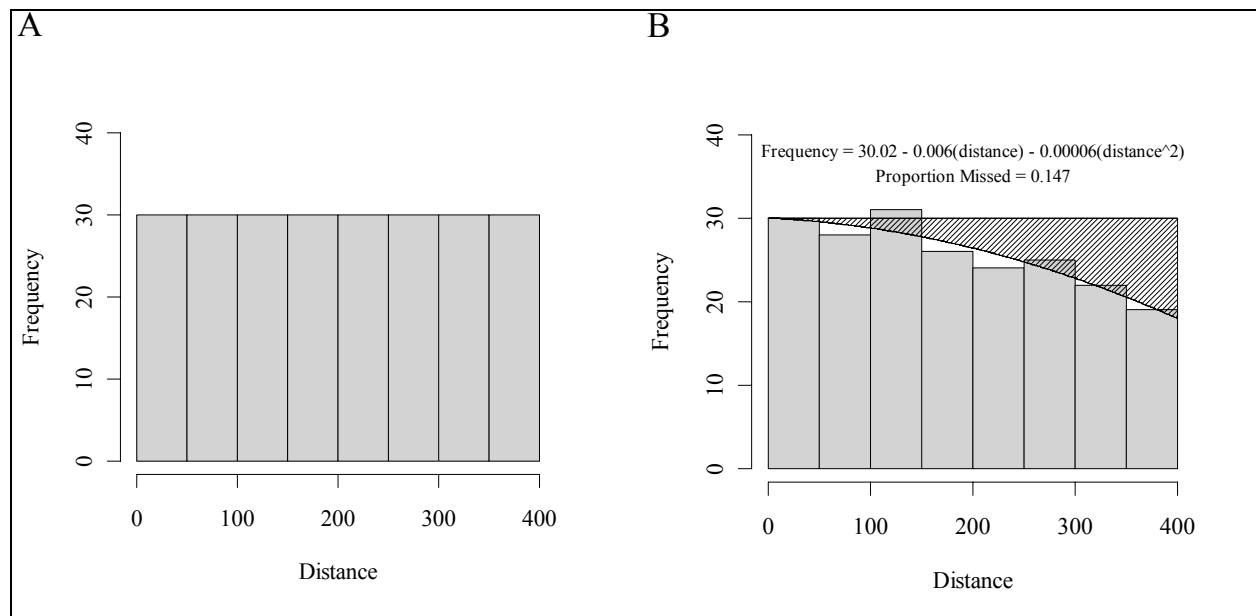


Fig. 2. Conceptual basis for line transect sampling: (A) the expected number of groups within eight distance classes with 100% detectability; (B) realistic survey data indicating a tendency to miss groups at greater distances from the transect line, with the proportion of groups missed (diagonally shaded area) estimated and used to correct for the number of groups in the search area. Modified from Buckland et al. (2001:21).



are observed during the survey then the probability distribution function of perpendicular sighting distances is approximately uniform (Fig. 2A) for relatively large populations. However, if more groups are missed further away from the transect line, and the probability of detection on the line is estimated, then the proportion of groups missed during the survey can still be estimated (Fig. 2B).

Correcting for missed groups is done by: fitting a detection function to the perpendicular distances to observed groups; and adjusting the number of observed groups by an estimated probability of detection. To illustrate, if 100 groups are observed and the average probability of detection is estimated to be  $\frac{1}{2}$ , then the estimated number of groups is  $100/(1/2) = 200$ . An adjustment must be made for average number of individuals in a group and a more complex analysis is required if there is less than 100% probability of detection on the transect line. For Fig. 2B, the number of observed groups is divided by the estimated average probability of detection; i.e.,  $(\text{count})/(1 - 0.147) = \text{count}/0.853$  in this example.

## 2.4 Critical Assumptions and Weaknesses

There are six critical assumptions for successful application of line transect sampling:

1. Groups on the transect line (in this case at the minimum available sighting distance) are always detected, OR an unbiased estimate of detection at the minimum available sighting distance can be obtained.;
2. Locations of groups are recorded at their initial sighting position;
3. Groups do not move before being detected and none are counted more than once. Alternatively, movement is “random” relative to survey lines;
4. Perpendicular distances from the transect line are measured without bias and with a relatively high degree of precision;
5. Sightings are independent events;
6. Transect lines are located “randomly” with respect to the location of individuals or groups.

Assumption 1 is difficult to guarantee when terrain or vegetative cover keep some animals from being detected on the transect line. To avoid negative bias in estimating abundance, a double-observer modification to standard line transect sampling can be used to estimate detection on the transect line (Manly et al. 1996, Quang and Becker 1997, Buckland et al. 2001).

An example of violation of assumption 2 is when animals are detected after being flushed by the aircraft. Researchers should search for violations of assumption 2 and address its potential effect. There is evidence of movement prior to detection if substantially more individuals were sighted some distance away from the transect line. Histograms of observed distances can be inspected for movement prior to detection, which is denoted by a spike in the frequency of

observations at some distance away from the transect line (in this case minimum available sighting distance).

On the surface it might appear that animal movement from one transect to another would be problematic (assumption 3), but random movement within the study area relative to the progression of the survey is allowed. However, movement of a large segment of the population, towards or away from the general progression of the survey, is not allowed.

Measurement of perpendicular distances can be obtained with a high degree of precision (assumption 4) using equipment such as laser range finders, locator binoculars, or a global positioning system (GPS). Distance bands drawn on windows of the aircraft can also be used if precise measurement is not possible. If distances can be measured without bias but precision is low, binning can also be used during data analysis (Buckland et al. 2001).

Care should be taken so that detection of a moose group is not dependent on detection of the prior moose group. For example, if the helicopter slows down due to detection of one group, and then another group that might otherwise have been missed is detected. Maximum likelihood methods for estimating a detection function assume independence among detections (assumption 5), but dependence should not influence estimates of the detection function or density (Buckland et al. 2001). Lack of independence will influence estimates of precision, but the bootstrapping methods described in this protocol should not be affected. In fact, Buckland et al. (2001:36) state, “failure of this assumption [independence] has little effect on the point estimators, and the robust variance estimators [bootstrapping of transect lines] we recommend”.

Assumption 6 can be met by randomly placing transect lines relative to potential moose habitat within strata.

Gasaway and DuBois (1987) outlined the major weaknesses of line transect sampling as related to moose surveys. Of those discussed, only three remain relevant. The first weakness identified was that “the method is restricted to relatively flat terrain” (Gasaway and DuBois 1987:607). The landscape of Innoko NWR is relatively flat, and in addition, line transect methodology can be used successfully over moderately uneven terrain if a constant survey altitude can be maintained by the pilot, and one can measure the perpendicular distances or distance bins with a high degree of accuracy (Buckland et al. 2001). Use of a radar altimeter increases a pilot’s ability to maintain a consistent survey altitude. Use of laser range finders that measure horizontal distances from the transect line can be used to reduce bias and increase precision of measured distances over slight to moderate uneven terrain because they adjust for survey altitude. Measurements of perpendicular distances to moose groups will not be perfect, especially in uneven terrain, however if there is no systematic bias (i.e., always overestimating or underestimating distance) then minor effects are expected on estimation of the distribution of observed distances (Fig. 2B) and the estimated probability of detection.

The second weakness is that “sightability of moose tends to be low, causing the SCF and its sampling variance to be large”. If average probability of detection (the SCF) is very small (e.g.,  $\leq 0.30$ ) then obviously, division by a small value can introduce unstable estimates. However, the estimated average probability of detection in past surveys on the Innoko NWR has varied from

0.44 to 0.63; values that should not introduce problems with unstable estimates. Features of the survey design can be adjusted to increase precision, such as: using double-observers to estimate the proportion of groups missed at the minimum observed distance; increasing the number of transects surveyed; using stratification; pooling data from previous surveys for estimating key parameters; and including covariates in the detection function (Buckland et al. 2004).

The last relevant criticism was that “bias and precision of the SCF estimators for moose surveys have not been adequately evaluated”. Research over the intervening 18 years suggests that little or no bias is expected under proper line transect survey design, data collection, and statistical analysis (Cassey and McArdle 1999; Buckland et al. 2001).

## **2.5 Other Uses of Line Transect Sampling**

Line transect sampling is believed to dominate estimation of abundance of cetaceans and large terrestrial mammals (Borchers et al. 2002). In addition to sampling moose in some regions of Alaska, line transect distance sampling has been used to survey brown bears (*Ursus arctos middendorffi*) (Quang and Becker 1996), caribou (Drummer and McDonald 1987), elk (*Cervus elaphus*) (Samuel et al. 1987), wild horses (*Equus caballus*) (Walter and Hone 2003), mule deer (*Odocoileus hemionus*) (White et al. 1989), harbor porpoise (*Phocoena phocoena*) (Laake et al. 1997), polar bears (Manly et al. 1996, McDonald et al. 1999, Evans et al. 2003), pronghorn (*Antilocapra americana*) (Johnson et al. 1991, Whittaker et al. 2003), seals (crabeater [*Lobodon carcinophagus*], leopard [*Hydrurga leptonyx*], Ross [*Ommatophoca rossi*], and Weddell [*Leptonychotes weddelli*]) (Southwell et al. 2002), hard shelled (*Cheloniidae*) sea turtles (Beavers and Ramsey 1998), whales (fin [*Balaenoptera physalus*], minke [*Balaenoptera acutorostrata*], and humpback [*Megaptera novaeangliae*]) (Hay 1982, Buckland et al. 2001, Hedley and Buckland 2004), macropodids in Australia (Southwell 1989), as well as other species (Buckland et al. 2001). Line transect methods are also popular for surveying bird populations (e.g., Good et al. 2004)

## **2.6 Other Potential Methods**

Non-controversial alternative methods for estimating moose density include: catch-effort and change-in-ratio methods (e.g., Borchers et al. 2002); mark-recapture (e.g., Amstrup et al. 2005); and plot or strip sampling (e.g., Gasaway et al. 1986, Ver Hoef 2001, Borchers et al. 2002).

Catch-effort methods require correct identification of individuals, random removal of individuals from the study area, and tend to have poor precision and the potential for large bias (e.g., Borchers et al. 2002). Usually, a large fraction of the population needs to be “caught” for acceptable performance. While moose are harvested on Innoko NWR, hunter access is physically limited and the harvest is limited to males only; non-Alaskan residents are further limited to large males only. The change-in-ratio method requires that the number of animals harvested be known, a high percentage of the targeted component of the population be removed, and the ratio of sexes or ages can be estimated with low bias both before and after the harvest (e.g., Borchers et al. 2002). This method is impractical as well for Innoko NWR as unbiased sex and age ratios of the moose population cannot be obtained, nor is a substantial proportion of the population harvested each year.

Mark-recapture methods would require substantially greater field costs to obtain similar levels of precision as line transect sampling, and estimates can be more sensitive to failures of assumptions (Borchers et al. 2002). In addition, a high degree of heterogeneity in capture probabilities across animals can substantially bias results and be difficult to account for in the analysis. If the population is extremely small or individuals are hard to detect during line transect surveys, mark-recapture may be preferable to line transect sampling if random samples of populations can easily be marked. One benefit of the mark-recapture method is that it allows for obtaining detailed information on each captured animal (e.g., survival, movement, etc.). Following a comparison of three aerial survey techniques for wild horses, Walter and Hone (2003) recommended line transect sampling over mark-recapture and strip counts because the assumptions of line-transect sampling were easier to satisfy and the method easily accounted for lack of perfect detectability.

The Gasaway technique (Gasaway et al. 1986), and other plot-based methods such as strip counts, have been used for estimating density of moose and other large mammals in many areas around the world, including Innoko NWR. These methods appear to be more appropriate than line transect sampling for dense populations in smaller study areas (Buckland et al. 2001). However, these methods are known to be expensive and time consuming when the population is sparsely distributed over a large area (Ward et al. 2000, Borchers et al. 2002). Key to obtaining unbiased estimates of animal abundance using plot sampling involves the assumption that all animals within the plot or strip are detected, or that a sightability correction factor (SCF) can be estimated without bias. Gasaway et al. (1986) incorporated a SCF into their protocol by intensively sampling a small portion of the plot just sampled, and using the ratio of counts from the intensive effort to the standard effort as the SCF. Search effort in plot-based surveys largely determines sightability; Gasaway et al. (1986) recommended increasing search effort if the SCF was  $>1.15$ , meaning the percentage of moose observed is  $<85\%$  using the standard effort. The Gasaway et al. (1986) technique relies on stratifying the study area immediately prior to the survey. Therefore, any delay in either initiating or completing the survey (e.g., multiple day weather delay during the survey) resulted in invalidating the stratification. This would result in researchers either abandoning the effort or starting over with a new stratification and restarting the survey.

A recent variation on the plot-based method incorporates the rapidly emerging field of geospatial statistics. Ver Hoef (2001, 2002) introduced this new method in Alaska in the late 1990's. While the spatial method (Ver Hoef 2001, 2002) also requires stratifying the survey area, the stratification is more broad than that from the Gasaway et al. (1986) method and is applicable from one year to the next, provided that the survey is conducted during the same time of year and significant changes to the habitat have not occurred (Ver Hoef 2000). ADFG moved away from the Gasaway et al. (1986) method and has adopted the spatial method (Ver Hoef 2000, 2001, 2002) for estimating moose densities within Alaska, as have some Federal land managers (Burch et al. 2004). Like the Gasaway et al. (1986) method, the spatial method (Ver Hoef 2000, 2001, 2002) requires multiple Piper Super Cub type fixed-wing aircraft to carry out the survey in a timely manner. Advantages that the spatial method (Ver Hoef 2000, 2001, 2002) has over the Gasaway et al. (1986) method are: only a single stratification effort is needed unless habitat conditions change; the survey can be conducted over a longer time period compared to the Gasaway et al. (1986) method; precision is higher; and it can be applied to smaller areas if

needed. Additionally, the method can be applied by sampling the same project area at a lower intensity more frequently to get a better identify on population trend (Ver Hoef, Alaska Department of Fish and Game, Fairbanks, Alaska, personal communication to S. Kovach, 1999).

## **2.7 Comparison of Line Transect Sampling to Plot Based Procedures**

As discussed above, ADFG has moved away from the Gasaway et al. (1986) method to the spatial technique (Ver Hoef 2000, 2001, 2002) for surveys on state lands. Differences from the Gasaway et al. (1986) technique include: smaller plots (generally half the size); decreased precision in defined strata due to more broadly applied stratification; allowance for a portion (10%) of plots surveyed to be selected non-randomly; and analysis based on Kriging (Ver Hoef 2000, Burch et al. 2004). The spatial analysis method used (i.e., Kriging), allows for a non-random sample of survey units, but comes with added assumptions about the model used for analysis (Ver Hoef 2002).

The line transect sampling approach to estimating moose density is based on a design based analysis approach compared to the spatial approach which relies on a model-based analysis. Model-based analyses tend to be more flexible, but require additional assumptions. The spatial technique, like the Gasaway technique, requires multiple fixed-wing aircraft conducting intensive surveys (~3 min/km<sup>2</sup>) for each plot; however, there is no estimation of a sightability correction factor for the spatial technique. The assumption that 100% of moose are being counted in all habitat types implicitly implies that the counts are an index on population density with an unknown proportion of individuals being missed. A main advantage of the line transect method is that correction for sightability bias is inherent in the method and does not require repeated surveys over some area or assumptions about detection rates.

The protocols for line transect sampling calls for recording moose locations with a high degree of accuracy, which can be used for habitat analyses and future survey stratification. This requires very little additional time or effort when using line transect methodology. To obtain unbiased locations of moose when using the Gasaway or similar technique, the aircraft would have to alter course to fly over the moose group to record its location using a GPS unit, and then return to the last position on the survey route; an activity not performed by many workers for every observation (Kovach, unpublished data). In addition to providing unbiased estimates of moose abundance, the line transect method, with its randomly selected transect lines, gives a better representation of areas of use by moose across geographic space and within habitats.

Pooling of observations from multiple observers in line transect sampling should result in minimum bias in estimation of the sightability correction factor (probability of detection) if the proportion of groups missed on the line is approximately the same for all observers. That is, line transect estimates of the average probability of detection are robust to pooling of data from multiple observers (Buckland et al. 2001). Similarly, if multiple observers are used on a plot type survey, each observer should estimate a unique SCF, or contribute to a pooled estimate, as there can be large differences in observer skill and experience. LeResche and Rausch (1974) found that even under the most ideal flying conditions, experienced observers saw only 68% of moose present during a plot-based survey, and inexperienced observers saw only 43%.

Both the Gasaway and spatial techniques appear to be just as, if not more, expensive and time consuming than the line transect methods for moose density estimation when total effort and costs are evaluated (Kovach, unpublished data). There are alternatives to line transect sampling, but this relatively efficient and cost effective method requires few assumptions and may be the best available technique for estimating moose abundance on Innoko NWR.

## **3.0 Methods and Protocols**

### **3.1 Sampling**

#### **3.1.1 Stratification**

Stratification can be used to improve precision and reduce bias of estimates, or for administrative convenience. Stratification can occur prior to sampling based on the known distribution of moose or moose habitat, and/or after sampling (post-stratification) to estimate density in domains of interest based on habitat types, environmental conditions, time, etc. Stratification requires knowing the proportion of total area that is in each stratum.

If moose are known to occur in high densities in some of the survey area and in low densities in other parts, then stratification of the survey area prior to sampling can reduce survey time and costs while increasing precision of estimates for the entire study area (Cochran 1977). However, if moose are distributed uniformly across the study area or their distribution is unknown, stratification prior to sampling may not be beneficial.

Domain estimation (post-stratification) procedures are used for the same reason as pre-sample stratification. Strata should be relatively large. Stratification procedures should generally be based on habitat information and other physical boundaries that are not overly influenced by current observed data. Workers are cautioned that unintended negative biases are likely introduced if current observed densities in a study area are used to define relatively small domains of interest.

Different stratifications may be used for estimating detection functions and population size within strata. If strata are defined on habitat types, then many survey parameters may not be constant across the strata, such as probability of detection on the line and probability of detection at greater distances from the aircraft. Pre-survey stratification is generally preferred over post-survey stratification. Reliable estimation of detection function parameters require large sample sizes, so it is often not advisable to post-stratify unless the number of transects and moose group sightings in each stratum are large (e.g., >20 transects and 60-80 observations of moose groups) (Buckland et al. 2001). Fortunately, detection functions are known to be pooling robust (Buckland et al. 2001), minimizing the need to post-stratify for this portion of the analysis. The concept of pooling robustness is that “data can be pooled over many factors that affect detection probability and still yield a reliable estimate of density,” (Buckland et al. 2001:41). Pooling robustness minimizes the need to post-stratify for this portion of the analysis. Alternatively, other parameters such as encounter rate or group size can be estimated without bias from relatively small samples. For these reasons, different stratifications are generally used for estimation of detection functions, and estimation of population size within strata. For example,

average group size might be estimated for each stratum separately, but an estimate of the detection function obtained by pooling data across strata.

Initially, 4 strata were identified within both the northern and southern halves of Innoko NWR (Table 1). Modifications were made to the strata through time, with some splitting and some lumping of individual strata. No documentation exists on the rationale behind the lumping of the northern strata, or adding the Yetna River strata in the south; however, the Western strata was split in 2000 in an effort to increase precision within that large area.

Table 1. Strata used in previous moose density estimation efforts on Innoko NWR. Buffer refers to the distance (in meters) from the river centerline to the edge of the strata.

Year	North Half		South Half	
	Strata	Buffer	Strata	Buffer
1994	Innoko River	4,000		
	Yukon River	4,000		
	Tributaries <sup>a</sup>	2,000		
	Other	N/A		
1996			Western <sup>b</sup>	N/A
			Innoko River	3,200
			Iditarod River	2,400
			Other	N/A
1998	Innoko River	4,000		
	Yukon River	4,000		
	Tributaries <sup>a</sup>	2,000		
	Other	N/A		
2000			Yukon River	4,000 <sup>c</sup>
			Holikachuk/Shageluk Sloughs	2,800
			Innoko River	3,200
			Iditarod River	2,400
			Yetna River	2,400
			Other	N/A
2002	Innoko River	4,000		
	Yukon River	4,000		
	Other <sup>d</sup>	N/A		
2004			Western <sup>b</sup>	N/A
			Innoko River	3,200
			Iditarod River <sup>e</sup>	2,400
			Yetna River <sup>e</sup>	2,400
			Other <sup>e</sup>	N/A

<sup>a</sup> Includes Dishna River, Mud River, and Papa Willie Slough.

<sup>b</sup> Includes the Yukon River, Holikachuk/Shageluk Sloughs, lowest portion of Innoko River, and some areas listed as Other in 2000 survey.

<sup>c</sup> Draft survey report does not identify actual buffer size; graphic appears to indicate a 4,000 m buffer width.

<sup>d</sup> Includes all tributaries as well as Other from 1994 and 1998 surveys.

<sup>e</sup> Scheduled but not completed due to weather.

### 3.1.2 Transect Selection

Systematic placement of parallel transects with a random start point has been used for line transect sampling in each stratum on the Innoko NWR. This placement ensures good coverage with a minimum number of transect lines. Transects that cross strata should be placed perpendicular to the stratification gradient, if one exists (e.g., east to west gradient of low, medium, and high density moose habitat).

Transects that cross major habitat gradients should also run perpendicular to the habitat gradient. Transects placed parallel to major habitat gradients can result in violation of assumption 6—individuals are distributed uniformly within the search area relative to distance from the transect line. Narrow strata, such as along river riparian corridors (Fig 3), make it difficult to efficiently run transects perpendicular to the gradient, however, narrow river riparian corridors typically have a fairly uniform distribution of moose. Thus, we can establish transects running parallel to the river, in a saw-tooth pattern (Fig. 3) for winding river corridors, without violating assumption 6.

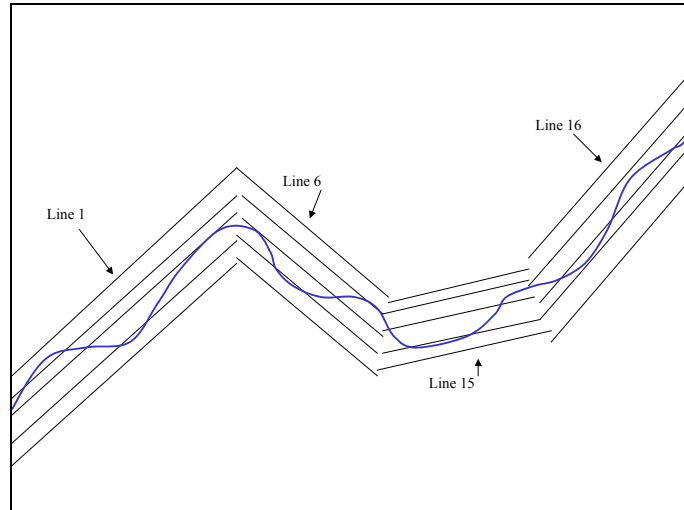


Fig. 3. Hypothetical layout of transects within a narrow river riparian corridor.

Spacing of transects should guarantee that stationary or slow traveling moose are not likely to be counted more than once. Given a survey altitude of 122 m (400 ft) above ground level (AGL), the suggested transect spacing is  $\geq 1$  km (0.6 mi) to prevent search areas from overlapping. To optimize survey effort it is advisable to increase sampling effort and reduce spacing of transects in strata with higher moose densities, and sample fewer transects per unit area in lower density strata. On Innoko NWR, 1.6 to 4.8 km (1 to 3 mile) spacing between transects has been used in moderate to high moose density habitat, and 4.8 to 8 km (3 to 5 mile) spacing for lower density habitat. On Innoko NWR, this transect density has resulted in 18% of the high density habitats along the Yukon River, 25% of the moderate density habitats along the Innoko River, and 10% of the low density, non-river habitats to be surveyed (assuming an 800 m [0.5 mile] survey width). A parallel saw-tooth pattern for transects within river corridors has been used in past surveys on Innoko NWR (Fig. 3). This pattern simplifies sampling within long, narrow river corridors and increases the amount of actual survey time relative to total flight time.

Once strata (or survey areas if using post-stratification) have been identified, and the spacing of transects has been determined, the actual transect lines need to be identified. In order to randomly select transect lines for sampling, a number of transect lines need to be available for selection. For example, Innoko NWR laid out parallel transect lines at 50 m (165 ft) intervals within each strata identified. The outer-most transects were located 400 m (1,312 ft) inside the outer boundary that is parallel to the transect line. Transects to be included in the survey are randomly selected from the total available set, maintaining the spacing requirements for the



strata, until the total number of transects for the area have been selected. This process should be repeated for each stratum and each leg within a river corridor.

Transect end points are located at stratum edges if the topography allows for safe flying (i.e., if a transect end point is located at the base of a cliff it is moved down the transect line so that the end point can be safely accessed by the helicopter). Transect lengths are not required to be the same throughout the study area, but should be long enough to ensure some moose observations on most transects and short enough to give the observers a quick break from time to time during the day. Once transects have been selected, they can be used for subsequent surveys (i.e., there is no need to select a new set of transects the next time the survey is conducted).

### **3.1.3 Aircraft**

Skinner et al. (1997) found that a helicopter was the best tool for line transect surveys for moose for the types of habitats and topography found on Innoko NWR. Sightability close to the aircraft is better than when using most fixed-wing aircraft. In addition, a helicopter has more maneuverability, and the ability to hover, if necessary, while verifying moose group size and composition. Perhaps of main importance is the ability of a helicopter to fly relatively slowly (e.g. <100 km/h [62 mph]) over dense cover for higher detection rates. Although the cost of an individual fixed-wing aircraft is much lower than a helicopter, Skinner et al. (1997) tested 2 common fixed-wing aircraft used on Alaskan refuges (i.e., PA-18 and C-185) and determined them to be unsuitable due to a wide obstructed view under the plane and faster airspeeds. Skinner et al. (1997) found that the cost of contracting a single helicopter was equal to the cost of 4-5 Government owned PA-18 Super Cub aircraft or 2-3 vendor owned PA-18 Super Cub aircraft for conducting a moose survey in 1996. In 2005 dollars, the cost of a single vendor Hughes 500D helicopter for a 40 hour survey (excluding fuel and ferry costs) is approximately the same as 5-6 Government owned PA-18 Super Cubs or 3 vendor owned PA-18 Super Cubs.

Skinner et al. (1997) recommended that the survey helicopter be equipped with bubble windows for the rear seat observers. All of their work was with a Bell 206B III model helicopter. Nielson et al. (2004) found that a Hughes 500D helicopter equipped with wedge windows afforded the same visibility as the Bell 206B III equipped with bubble windows. Regardless of which platform is used to conduct a moose survey, the helicopter should be equipped with a radar altimeter to assist the pilot at maintaining the survey altitude of 122 m (400 ft) AGL.

Past surveys on Innoko NWR have flown 1,000-1,200 km (621-746 mi) of survey lines with 30-36 hours of flight time (actual time surveying plus ferry time between base and transect lines). Therefore, a single helicopter can efficiently complete a survey in a timely manner.

## **3.2 Map and Waypoint Preparation**

Prior to starting the survey, electronic and hard copies of all waypoint coordinates need to be made. Both the electronic and hard copy of transect waypoints need to be given to the helicopter pilot for entry into the onboard GPS unit. In addition, a hard copy of the waypoints should be brought on board the aircraft during the survey should waypoints need to be re-entered into GPS units.

Selection of spheroid and projection of transect waypoints is not critical. What is critical is that the datum, spheroid, and projection remain consistent between GPS units, field maps, and the Voice/GPS Survey program (Appendix 3) for the entire survey. It is strongly recommended that GPS units be allowed to stay in their native datum (usually WGS84) as different GPS units vary in their ability to translate between datums. However, because the Voice/GPS Survey program requires a NAD27 datum, the GPS datum, as well as all the waypoints and map datum needs to be set to NAD27. Additionally, a transect line file must be made for the Voice/GPS Survey program (see Appendices 3 and 4).

Prior to the actual start of the survey, the project leader should ensure that all GPS units (aircraft and handheld) do in fact contain all transect waypoints and are set to the same datum.

Field maps need to be prepared prior to starting the survey as well. For large survey areas (e.g., >5,000 km<sup>2</sup>), field maps based on USGS 1:250,000 scale topographic maps have proved useful. One large (e.g., 56x86 cm [22x34 in]) wall map is useful for planning and tracking of the survey. Smaller, 28x43 cm (11x17 in) flightline maps should be brought on board the aircraft for in-flight navigation; laminating these maps increases their durability. Reduced scale example maps are located in Appendix 1.

### **3.3 Field Procedures**

When developing a field plan prior to a survey, develop several alternatives for completing the survey. Weather and unforeseen events can require frequent changes to the initial survey plan. The line transect method is not affected by the sequence in which transects are completed both within a single survey and between subsequent surveys. Location and amounts of fuel caches need to be discussed prior to the survey with the helicopter pilot, land management personnel, and the fixed-wing pilot who will be delivering the fuel. Caches may be placed in advance, provided a secure location is available.

If Government owned aircraft (helicopter for the survey and fixed-wing for transporting fuel caches and crew) are not available, then a contract must be prepared well in advance of the survey (about 3-4 months). The contract should state that actual survey dates will be determined by weather and that the survey must begin within relatively short notice (24-48 hours). The survey helicopter must have certain equipment (see 3.3.1. Checklist/Equipment below). The project leader must discuss survey procedures, expected number of flight hours, and logistics (fuel, housing, and food) with the helicopter pilot prior to the start of the survey.

For those unfamiliar in applying the line transect technique for moose density estimates, or if new observers are on the crew, a practice flight(s) for training purposes must be taken before the actual survey begins. This will allow all observers, as well as the pilot and data recorder, to become familiar with all survey protocols, procedures, and equipment. If possible, practice in an area of moderate to high moose density outside the actual survey area to fully prepare the crew. Project leaders need to be sure to include any needed training flight when planning their survey, and budget accordingly.

### 3.3.1 Checklist

#### Crew

- In-flight crew consists of a pilot, data recorder (located in front passenger seat), back left observer, and back right observer.
- Ground support (e.g., fuel handling, fuel hauling, flight following) may be required depending on the location and specifics for each survey.

#### Equipment

- 4-seat helicopter equipped with GPS navigation, radar altimeter, backseat bubble (Bell 206B III) or wedge (Hughes 500D) windows, 12-volt power supply (cigarette lighter receptacle type) for a laptop computer, and all way intercom.
- Nomex flight suit and gloves, helicopter approved helmet, and personal survival kit/vest for each crewmember.
- Data GPS unit with external antennae to attach to helicopter windscreen and appropriate cable(s) for interfacing with laptop. The data GPS unit is used in combination with a laptop computer to record moose locations and flight-path. The Voice/GPS Survey program requires the GPS unit to have NMEA 0183 output with a minimum baud rate of 4800. Carry enough batteries for the GPS unit to maintain battery strength at  $\geq 50\%$  throughout the day.
- Laptop computer with  $\geq 2$  batteries, a power cord, and appropriate jacks for the microphone and data GPS unit. The laptop computer should be plugged into the helicopter's power supply during the survey, but battery power can be used as backup. The laptop computer should have a minimum of 2.00 GHz processor speed, 500 Mb of RAM, and 10 Mb of hard disk storage after all software and data files have been installed. A hand-held external mouse can make recording observations easier, but is not necessary for a successful survey.
- If the laptop computer does not have a power cord for a 12-volt power source, then a power inverter will be needed to supply power to laptop computer (and possibly the data GPS unit) while in flight.
- External heat source(s) for laptop computer and data GPS if the environment inside the helicopter may be below the minimum operating specifications
- Software
  - GPSTRACK (Anthony and Stehn 1994) was used for in flight tracking during the 1994 and 1996 surveys. Since 2000, the in-flight tracking program Voice/GPS Survey (executable files: MOVEMAP, RECORD, and TRANSCRIBE; Appendix 3) has been used with good success.
  - Software for uploading waypoints into GPS Units (e.g., MapSource [Garmin International, Inc., Olathe, Kansas, USA] for Garmin GPS units).
  - Spreadsheet or database manager for editing and assembling daily data files.
- Pre-survey data

- Transect line endpoints for Voice/GPS Survey (Appendix 3, Table A.4.2)
  - Transect end points for data GPS units
  - Graphic files depicting survey areas and transect lines
- Backup disks (e.g., floppy, CD, or DVD) with all pre-survey data.
  - Disks for daily backup of collected data.
  - Wall map showing all intended flight lines, based on USGS 1:250,000 scale topographic maps for flight planning.
  - Flight line maps, 28x43 cm (11x17 in), at 1:100,000 to 1:150,000 scale for in-flight navigation.
  - Printouts with latitude/longitude locations of all flight line waypoints and turning locations; at least 1 copy to be on the survey helicopter.
  - Data forms (Appendix 2; minimum of 4 per transect), clipboards (3), pencils (1 box), ruler (1), calculator (1), and paper.
  - Laser Range Finders and manuals for both back seat observers. These range finders must be capable of measuring perpendicular distances between the transect line and the observed moose group. In 2004, Impulse XL laser range finders (Laser Technology, Inc., Englewood, Colorado, USA) were used to measure perpendicular distances. These range finders have a built-in clinometer to adjust for the height of the aircraft above the ground.
  - Clinometer for measuring angles that delineate the distance bands marked on the side windows of the helicopter.
  - Grease pencils (black) for marking windows of aircraft at distance band angles.

It is very important to have extra equipment of all types in case items are broken, lost, or malfunction - survey success depends upon backup equipment. Examples of backup equipment include: data GPS and external antenna; laptop computer, cables, and software; power inverter; and clinometer. While not specifically equipment, having an extra person available to serve as an observer can save time and money by keeping a survey going should someone become ill. Again, any observer not familiar with any of the survey equipment must be trained on the ground and then practice in-flight distance measurements and data recording procedures prior to the start of a survey.

### **3.3.2 Safety**

For Department of Interior (DOI) agencies, surveys will be conducted with DOI Aviation Management approved pilot and aircraft. The pilot must have a low level endorsement. All aircraft associated with the survey must be equipped with a winter survival kit, sleeping bags for each crew member, an emergency locator transmitter, and satellite telephone. For the survey helicopter, pilot and crewmembers must wear appropriate winter helicopter safety gear (i.e.,

helmet, Nomex winter suit, and Nomex or leather gloves). Each crewmember in the aircraft must have a personal survival kit either in a survival vest or in pockets of their clothing, and at least one crewmember should have a hand held radio in his/her flight suit. The weather service should be called daily for current weather conditions and future forecast. All surveys, or other flights, must have at least 4.8 km (3 mi) visibility and winds <37 km/h (23 mph or 20 knots). Each aircraft must file a flight plan daily with the Federal Aviation Administration (FAA) or a person on the ground. The flight plan must state the number of people in the aircraft, the flight route, and the latest time the flight will end. All aircraft associated with the survey must be on the same radio frequency. Both the survey helicopter and support fixed-wing aircraft when in flight but not surveying should fly at least 305 m (1,000 ft) AGL.

### **3.3.3 Required Snow Conditions**

Snow conditions have a major influence on sightability of moose. Therefore, satisfactory snow conditions must exist throughout the survey area throughout the survey period. Gasaway et al. (1986) recommended moderate to good snow conditions as being required for a successful survey. They defined moderate snow conditions as: some low vegetation showing if fresh (i.e., <7 days old) or moderately aged (i.e., 7-13 days old) snow existed; or complete snow cover with snow aged  $\geq 14$  days old. They defined good snow conditions as complete snow cover with fresh or moderately aged snow. Complete snow cover includes adequate depth to cover low vegetation; horizontal logs must have sufficient snow on them to help distinguish them as such and not bedded moose with snow on their backs. Root wads from upturned trees are problematic for all surveys; sufficient snow draped over root wads helps observers to distinguish root wads from moose when searching at greater distances from the helicopter.

### **3.3.4 Survey Airspeed and Altitude**

Moose are harder to see in dense conifer cover than treeless cover. Airspeed of the helicopter should vary accordingly, and stay consistent within habitat types. Surveys have been conducted at about 48-64 km/h (30-40 mph or 26-35 knots) in dense conifer cover and 145-161 km/h (90-100 mph or 79-88 knots) in open habitat types. Line transects are to be flown at 122 m (400 ft) AGL. Skinner et al. (1997) found that aircraft were causing movement of animals when flying at 76 m (250 ft) AGL during shallow snow winters. In addition, they found that the ground appeared to be moving by very rapidly when flying at 76 m (250 ft) AGL, which could cause moose to be missed near the aircraft. Aircraft safety is improved with a higher survey altitude, but Skinner et al. (1997) found that at 152 m (500 ft) AGL too many moose are missed in dense cover. The survey altitude of 122 m (400 ft) AGL has been found to be a good compromise to improve survey safety and minimize violations of line transect assumptions.

### **3.3.5 Crew Duties**

The data recorder is responsible for recording all sightings from all participants on the computer (Appendix 3). The data recorder will announce the ID number of each observation and the appropriate back seat observer will record this ID number on their data sheet. The data recorder is also responsible for searching for moose in all zones. Additionally, the data recorder is responsible for recording all relevant survey conditions (see 3.3.6).

Each back seat observer is responsible for searching for moose as soon as a transect has started (i.e., both sides of the transect are searched for moose), classifying moose observed, determining the distance to the moose group observed, and reporting and recording their observations.

The pilot is primarily responsible for flying transect lines, maintaining prescribed survey altitude and airspeed, and determining when weather conditions are no longer within the survey parameters (i.e., winds >37 km/h [23 mph or 20 knots], visibility <4.8 km [3 mi]). Secondly, the pilot searches for moose within Zone 0 (see 3.3.7) and beyond. The pilot is to inform the data recorder of moose observations within Zone 0. To avoid confusion, the pilot is not to call out sightings of moose in Zones 1-6 independently except those that are missed by the back seat observer on the pilot side of the helicopter.

### **3.3.6 Recording Survey Conditions**

At the beginning and end of each survey flight when the helicopter is on the ground or in transit to the survey area, the data recorder is responsible for documenting relevant survey information such as crew names and their positions within the helicopter, weather conditions, flight line numbers to be flown, etc. Back seat observers are responsible for recording their position within the helicopter and who is in the other positions, date, strata to be flown, and page number on the field data form (Appendix 2).

The data recorder is also responsible for categorizing weather and sighting conditions throughout each survey flight. Weather conditions include cloud cover (e.g., 10ths), ceiling height, precipitation, approximate wind direction and velocity, and lighting conditions (e.g., bright intense light, flat low light, etc.). Sighting conditions are categorized into good, fair, and poor based upon degree of overcast, precipitation, and lighting conditions. These observations are recorded on the transect log and coded to the corresponding time given by the computer (see Appendix 3 for more information).

Location of the helicopter, time, and date are recorded automatically by the GPS/computer at 10-second intervals (Appendix 3). This permits plotting and comparison of the actual flight path versus the planned transects. The data recorder is responsible for recording the start and end points into the computer using the Voice/GPS Survey program. These locations can be recorded using the microphone (Appendix 3) to describe whether the location is a start or end point.

The radar altimeter is used to help the pilot keep the aircraft at the intended survey altitude. If laser range finders are not used to measure distance to moose and distance bins are used, then one needs to measure the discrepancy between the actual versus intended survey altitude. The recorder should view the radar altimeter at fixed time intervals (e.g., 5 minute intervals) and record the data. Airspeed can also be recorded in a similar fashion; however, airspeed is used for descriptive purposes only and is not required for statistical analysis of moose abundance.

### **3.3.7 Measuring Distance to Moose**

Perpendicular distance from the transect line (i.e., location of helicopter when flying on the planned transect line) to the center of the observed moose group is to be measured. When used properly, laser range finders can measure the perpendicular distance to the center of the moose group with a high degree of accuracy ( $\leq 3$  meters). To measure the distance, observers are to aim

the indicator at the feet of moose in order to avoid measuring past the moose (thereby resulting in a greater distance measurement). The Impulse XL laser range finders used in the past have failed at distances  $\geq 600\text{m}$  (1,969 ft); failures were indicated by an error message in place of a distance reading. If laser range finders (or alternative equipment) are not available or malfunction, seven distance zones (Table 2, Fig. 4) need to be marked on the back windows of the helicopter prior to the start of the survey.

Table 2. Angles corresponding to distance zones to the nearest 1/2 degree, for above ground level flight of 122 m. Distance zone widths and midpoints are in meters.

Zone	Distance Zone Width	Midpoint	Angle Ranges
0	0 - 45	22.5	70.0 – 90.0
1	46 - 100	73	50.5 – 70.0
2	101 - 165	133	36.5 – 50.5
3	166 - 240	203	27.0 – 36.5
4	241 - 320	280.5	20.0 – 27.0
5	321 - 400	360.5	17.0 – 20.0
6	> 400	na	< 17.0

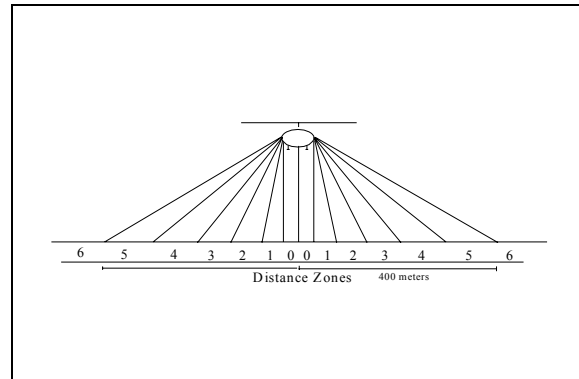


Fig. 4. Illustration of distance zones.

Each crewmember marks his/her window with the aid of a clinometer and a grease pencil at the angles representing each of the distance zones. The observer's head and body should be in a relaxed and comfortable but fixed position relative to the marks. This is best accomplished with an observer sitting in their seat and the helicopter door closed. The observer sights through the clinometer and uses their finger to indicate the proper location. Another worker standing outside the helicopter then marks the window at the location indicated by the observer. Several marks are needed across the length of the window in order to properly mark the zone boundary. Putting the grease pencil marks on the outside of the window also prevents them from being rubbed off by observers during flights. Obviously, observers must be wearing all their flight gear in order to properly mark the windows. Distance angle between Zones 0 and 1 can be different for each helicopter because of the different angles of view of the ground immediately adjacent to the aircraft. However, if multiple helicopters are used on one survey, identical distance zones must be used. Also, people of different heights view the ground at different angles. The shortest person participating in the survey should determine the cut-point between Zones 0 and 1 for the survey helicopter so that all observers see the same area. Distance zones are numbered 0 through 6, with Zone 6 having no distant boundary (Table 1). Total width of Zones 0-5 is approximately 400 m on each side of the helicopter. As zones become further from the transect, their width increases (Buckland et al. 2001:262).

The area directly under the helicopter that cannot be viewed from the back seat is Zone 0 (Fig. 4). Zone 0 needs to extend to the point where both eyes of the rear seat observers have an unobstructed view of the ground. That narrow strip where only 1 eye can see the ground is to be included into Zone 0 as it is difficult to see animals in this narrow visible strip. The break point between Zones 0 and 1 will vary slightly between helicopter type and observer height. This

angle must always be measured and recorded prior to the start of each survey. It is important that the shortest observer also mark the junction between Zones 0 and 1 on the data recorder's window; this aids the data recorder in determining moose detections near this critical line (see 3.3.9 below). Zone 0 and most of Zone 6 observations are not used for the statistical calculations pertaining to population size.

### **3.3.8 Moose Observations**

A moose sighting consists of a group of moose, including groups of single individuals, sighted near each other while flying along the transect line at the assigned survey altitude and an appropriate airspeed. The definition of what constitutes a moose "group" must be identified prior to the start of each survey (e.g., individuals within 15-20 m of each other), and must be consistent from survey to survey. The location of a sighting is defined to be the latitude-longitude of the aircraft on a particular transect at the point of intersection of a perpendicular projection from the center of the moose group (Fig. 1). It is important not to record moose seen while flying off-transect (e.g., flying between transects) as though they were observed while on-transect, as this will bias results. As the number of moose in each group is used for estimating density, the group size and composition is recorded.

Classification of moose observations in Alaska differs between fall and late winter surveys (Gasaway et al. 1986). Late winter survey moose classifications are: combined number of yearling and adult moose; and number of calves. Fall survey moose classifications are: number of large bulls with antlers  $\geq 127$  cm (50 in); number of medium-sized bulls with antlers  $< 127$  cm (50 in), but larger than yearling bulls; number of yearling bulls with spike or forked antlers; number of cows without calves; number of cows with 1 calf; number of cows with 2 calves; total number of lone calves; and number of sex-age unknown. These sex and age groupings can be altered to suit individual survey needs (Gasaway et al. 1986). Workers should note, however, that altering these groups from one survey to the next can negatively effect the ability to assess trends across time in a given group, or shifts in composition (relative abundance of groups) across time.

For each moose group observed by a rear seat observer, the observer will announce a moose or group of moose has been sighted just prior to when the moose group is positioned perpendicular to the helicopter to enable the data recorder to "mark" the location. The observer will state the position of the moose as either left or right of the transect line, the distance to the center of the moose group (or distance band), the number of moose in each sex/age category, and the estimated percent cover in an approximate 10 m (33 ft) and 50 m (165 ft) radius buffer around the center of the moose group. For example, if during a late winter survey the back left observer sees a cow and calf 250 m (825 ft) from the transect line in cover estimated to be 50% within 10 m of the group and 30% within 50 m of the group, the observer will announce "left side, 250 m, 1 adult, 1 calf, 50%, 30%". In order to reduce the number of moose groups missed by observers, whenever an observer detects and announces a moose group, the pilot should reduce air speed, by half or more if safe to do so, while the observer measures the distance and records the data. If all data were not collected, the helicopter must return to the marked location to allow the observer to collect the remaining data.



Each backseat observer is responsible for recording all of the sightings on their side of the aircraft on the observer data sheets (Appendix 2). This provides quality control of the data entry by the data recorder, and can be used as a backup if the laptop computer fails. Observer data sheets are crosschecked with the computer data files at the end of each survey day.

Observers must concentrate on detecting moose within Zone 1 (i.e., within 100 m [328 ft] of the helicopter). While the double observer observations (see 3.3.9 below) provides a measure of the moose missed in Zone 1, it is very important to detect all moose possible in this area to minimize violations of assumption 1 and the variance of the estimated detection probability at this distance. After detecting all moose in Zone 1 (the first area indicated by the grease pencil lines), Zone 2 (the second area indicated by the grease pencil lines) is then scanned until one is certain all moose possible are detected. Zones 3-5 are scanned in exactly the same manner. In dense cover, the outer zones will not receive the same search effort as the inner zones due to visual obstruction caused by the vegetation. Line transect methods are very robust to data pooling (see sections 2.7 and 3.4.2) for the detection function, enabling workers to pool data collected across cover types.

Using a computer mouse and microphone, the recorder records the location on the transect line perpendicular to the observed moose group by left-clicking the mouse when using the Voice/GPS Survey program, and while holding down the mouse button use the microphone to voice record observation data such as which side of the aircraft observed the group, the group size, and perpendicular distance (see Appendix 3 for more information).

The pilot should refrain from calling out a moose sighting in Zone 1 or beyond until the moose group has been passed and it becomes clear that the observer seated behind the pilot did not detect the moose group. Information regarding the pilot's observations will be recorded into the field data forms by the rear seat observer, but will not be used in the analysis to estimate detection functions and overall density. To conserve flight time and fuel, observations made only by the pilot will not be circled for verification or distance measurements.

### **3.3.9 Double Sampling**

Perfect detection of individuals on or near the transect line is a primary assumption of standard distance sampling and analysis procedures (Buckland et al. 2001), yet in many cases, detection of animals on or near the line is <100% (Borchers et al. 2002). During the surveys, the double-observer (Manly et al. 1996, Quang and Becker 1997) approach is to be used on the passenger side of the helicopter. Under this approach, the data recorder is charged with detecting moose at the inside edge of the survey strip (i.e., the junction between Zones 0 and 1). An experienced data recorder should search for moose groups within all distance zones. More double sampling observations provide more information (e.g., larger sample size) on detection probabilities. Past experience has shown that having the data recorder also search for moose does not unduly detract from his/her primary duty of data recording.

Double-count data collection procedures on the double-observer side of the aircraft allow estimation of detection rates on or near the transect line by the rear seat observer, which can then be used in the analysis to correct for those individuals missed during the surveys. The double-observer sampling approach requires that the data recorder and observer on the passenger side of

the helicopter operate independently of each other, and that each correctly records: moose detected by both observers; moose detected by the front seat observer and not the rear seat observer; and moose detected by the rear seat observer and not the front seat observer. To help in estimating an unbiased probability of detection at the minimum available sighting distance, rear seat observers should change seats at least daily to ensure a representative sample of double-count observations.

Under the double-sampling protocol, the data recorder simply records the GPS location of the moose group when it is perpendicular to the helicopter but does not announce the observation. Only once the group has passed and out of sight of the back seat observer (depending on cover and airspeed, an average of 5 to 10 seconds after initial sighting), and no other moose are in sight on that side, does the data recorder announce their sighting. If necessary, the pilot will leave the transect line in a fashion so that the group under question can be verified. To save time and fuel, moose groups observed on the double-observer side of the aircraft should be circled only when it is unclear whether both observers saw the same moose group, group size was not determined, or the perpendicular distance of the group from the transect line was not obtained. It is important to reconcile differences in moose observations by the double-observers before proceeding along the transect. It is also important that once off-line, observers pay attention only to the location of the group in question so that the pilot can efficiently circle and locate the group. A determination will be made whether both the recorder and observer saw the group in question (this can be as simple as the recorder and observer discussing the observation to as complex as various indicator systems). The pilot will then navigate to the location on the transect line where they came off-line and continue with the survey. To save time and fuel, moose groups observed on the pilot side of the aircraft should be circled only when necessary information is missing. The observer will announce moose observations as described above (section 3.3.8). The data recorder will inform the observer if they also observed the same moose group.

There are alternatives to estimating the probability of detection at the minimum available sighting distance, which include using the pilot as another double-observer to obtain a larger sample size for double-count trials. In addition, the pilot and front seat observer could be used as regular observers for the area close to the transect line, which is fully visible from the front seats of a helicopter and not observable by the rear seat observers. Potentially, observations from the pilot and front seat observer could be used to help estimate the proportion of groups being missed by rear seat observers at the minimum observed distance (i.e., junction of Zone 0 and 1). Although pilots often have the best visibility and are skilled at detecting animals during flight, safety should be the primary concern during surveys and the pilot should not be responsible for looking for moose unless weather is relatively calm and terrain is relatively flat. In general, data from the pilot and front seat observer are of limited value for density estimation calculations, but can be important for resource selection analysis.

### **3.3.10 Data Management**

Daily data management for the surveys includes cross checking computer data files with field data forms, adding each day's new moose observations to a data file, and backing up all computer data files. The initial data file for moose observations can be created using a spreadsheet program such as Excel (Microsoft Corp. Redmond, Washington, USA). This data file must contain one row for each moose group observation with the moose group size,

perpendicular distance from the transect, percent cover in the 10 m and 50 m radius buffers around the group, transect line or segment, and who observed the moose group (see Appendix 4 for an example).

Upon returning from the survey, proper data management includes integration of the initial moose location data files into refuge databases for inclusion into the appropriate geodatabase(s). This includes using the GIS to re-project the geographic information (i.e., latitude/longitude values) from the NAD27 datum used with the Voice/GPS Survey software to match current regional standards (NAD83 and Alaska Albers Equal Area Conic projection as of this writing). Data sheets are to be scanned into a single pdf document. Metadata associated with geodatabases must be updated as well. All updated electronic files are to be backed up following current applicable office protocols.

### **3.4 Data Analysis**

#### **3.4.1 Data Preparation**

During a survey, the program Voice/GPS Survey (Appendix 3) is used to record the actual transect flight lines flown, locations along transects where moose groups were observed, and relevant information for each moose observation. At the end of each survey day, the associated program TRANSCRIBE (Appendix 3) is used to obtain locations on each transect where each individual moose group was observed, along with the actual start and endpoints of each transect. These data are then transferred to a master data file; each day's data are appended to the end of the master data file (Appendix 4.1).

At the completion of the survey, 2 files (3 if areas with empirical counts were conducted) must be created to begin the data analysis process, and another file must be created during the data analysis process (Appendix 4.1). Creating the file of survey observations is a straightforward process as it is created from the master observation file. Creating the file containing the lengths of each transect flown an easy, but cumbersome process involving the use of a GIS. The last file created from multiple runs of program DISTANCE is another easy file to create. Data file formats and production processes are detailed in Appendix 4.1.

#### **3.4.2 Estimating Moose Abundance**

##### **3.4.2.1 Search width**

Due to the pooling robustness of the detection function, a separate detection function is not needed for each stratum. However, if it is believed the detection function is substantially different between strata, a separate detection function could be created for each strata provided there are many, say  $> 60$ , observations in each stratum. Usually, small sample sizes within some stratum will preclude estimating separate detection functions.

Prior to density estimation, the search strip width for the aerial line transects must be defined. This strip defines the minimum and maximum available sighting distances,  $W_1$  and  $W_2$ , respectively, to be used in the analysis, and allows for calculation of the total area searched for moose within each stratum. Survey protocol requires the helicopter fly at 122 m (400 ft) AGL on all transects. Flying at 122 m AGL results in a swath of approximately 45 m (148 ft) under the aircraft on either side that cannot be observed so  $W_1$  can usually be set to 45 m. Recall that

this minimum available sighting distance has been measured using laser range finders, and is consistent with calculations based on clinometer measurements of the maximum downward sighting angle from the horizon from the rear seat of a Hughes 500D with wedge windows and a Bell 206B III with bubble windows. Workers are strongly encouraged to verify  $W_1$  at the start of a survey. If a different survey altitude, helicopter, or window style is used for the survey, the minimum available sighting distance must be measured using both the clinometer and a laser range finder (if available). Offsetting the transect line due to inadequate view under the platform or to alleviate other problems with the data on or near the transect line is called left-truncation (Buckland et al. 2001).

During the survey observers are to record all moose groups seen, regardless of a group's distance from the transect line. Inspection of histograms of observed distances can aid in identification of outliers at large distances, provide evidence that "heaping" of distances occurred (see Buckland et al. 2001 for an example of heaping distances), or that animals moved prior to detection. Several different histograms using alternative bin sizes should be used for investigation of the assumptions that locations of groups are recorded at their initial sighting position, and there was no movement prior to detection. If data are not truncated in the field (e.g., only recording observations within 400 m), Buckland et al. (2001) recommend removing around 5% of the observations with the largest perpendicular distances prior to analysis. (We do not recommend truncating observations made during the survey as observers tend to "wish" moose into the survey strip.) This right truncation facilitates modeling of the data and usually results in little loss of precision (Buckland et al. 2001). In past surveys on Innoko NWR, all observations within 400 m (1,312 ft) of either side of the helicopter have been used, and  $W_2$  was equal to the maximum distance a moose group was sighted during that survey, provided it was  $\leq 400$  m.

If perpendicular distances of observed moose groups from the transect line are measured using relatively precise methods (e.g., laser range finder), raw distances can be used in the analysis. If distances to groups are very imprecise, based on bin markings on the window, or heaping occurs, then binning should be used in the analysis (Buckland et al. 2001). If window distance band markings are used to estimate perpendicular distances, the bin midpoints should be used as the observed distances during the analysis.

#### **3.4.2.2 Estimating detection on the transect line**

Standard line transect analysis assumes 100% detection of groups at or near the minimum available sighting distance,  $W_1$ , or  $\hat{g}(W_1) = 1.00$  (Buckland et al. 2001, Borchers et al. 2002). Following determination of the search width ( $W_2 - W_1$ ) for transects surveyed, probability of detection of moose groups at the minimum available sighting distance by the rear seat observers is to be estimated using the double-count observations. Visibility bias, or probability of detection, at the minimum available sighting distance can be estimated using logistic regression and all observations within the interval  $[W_1, W_2]$  by just the front seat observer (data recorder) or both the front and rear seat observers on the double-observer side of the aircraft. Essentially, the researcher models the probability of detection by the rear-seat observer given the front seat observer saw the moose group using logistic regression (McCullagh and Nelder 1989). For this analysis, the response variable = 1 if both the front and rear seat observers detected the moose group, and 0 if the front seat observer saw the moose group but the rear-seat observer did not (Appendix 5.1). Observations of moose groups by the back seat observer not detected by the

front seat observer are not used for estimating detection at the minimum available sighting distance.

In the past, two logistic regression models have been fitted for the probability of detection at a given distance by the rear seat observer. Models used distance from the aircraft and group size, but other covariates such as percent cover or observer could also be included in the models (see Buckland et al. 2004 for examples), given sufficient sample sizes. If a variable for group size or habitat type is in the final logistic function, an estimate of  $g(W_1)$  is needed for each level of each covariate, and an average estimate of  $g(W_1)$  should be used. Another approach would be to stratify estimation of the detection function (see 3.4.2.3 below), or include the same covariates in both the logistic function and the detection function. Our experience suggests that it is probably sufficient to only use distance from the transect line in the logistic regression function.

Using the determined minimum and maximum available sighting distances of  $W_1$  and  $W_2$ , respectively, the logistic regression analysis assumes the probability of a group of moose being sighted at a perpendicular distance  $x_1$  from the rear seat observer is well approximated by the following logistic function:

$$g(x) = \exp(\beta_0 + \beta_1 x) / \{1 + \exp(\beta_0 + \beta_1 x)\}, \quad (1)$$

where the  $\beta_i$  are unknown coefficients, and  $W_1 \leq x \leq W_2$ .

If more than one logistic regression model is fit to the double-count data, we recommend using an information-theoretic approach for model selection (Burnham and Anderson 2002). Generally, sample sizes for fitting the logistic regression models will be small enough to warrant use of the second-order variant of Akaike's Information Criterion (AICc) to select the final model (lower is better). AICc for each model can be calculated using

$$AICc = -2 \ln(Likelihood) + 2pn / (n - p - 1), \quad (2)$$

where  $p$  is the number of parameters in the model (including intercept term),  $n$  is the number of observations in the sample, *Likelihood* is the value of the logistic likelihood evaluated at the maximum likelihood estimates, and 'ln' is the natural logarithm. A basic check of the final logistic function should indicate decreasing probability of detection for the rear seat observer as distance increases. A plot of the estimated detection function over the observed distances is used to further evaluate model fit (Appendix 5.3). A drop-in-deviance test (McCullagh and Nelder 1989) could be used to provide a more formal evaluation of model fit, although this is rarely necessary.

Using the final logistic regression model or a weighted average approach (Burnham and Anderson 2002), the probability of detection at  $W_1$  can be calculated along with a 95% CI for this prediction. To calculate the probability of detection at the minimum available sighting distance, or  $\hat{g}(W_1)$ , simply plug  $x_1 = W_1$  into the estimated logistic regression model (equation 1). SAS Proc Genmod (SAS Institute 2000) is used to estimate the logistic regression model and obtain a 95% CI for the estimation of  $g(W_1)$  (Appendix 5.3).

### 3.4.2.3 Estimating a detection function

Standard line transect distance analysis involves fitting a detection function to the distance data, integrating this function over the search width ( $W_1$  to  $W_2$ ), and then dividing by the search width to obtain an estimate of the average probability of detection,  $\hat{P}$ , provided  $g(W_1)$  is known or can be estimated. This estimate of the average detection probability is then used to expand the number of groups detected to account for groups missed within the survey strip. Based on this protocol, density can be estimated using

$$\hat{D} = \frac{n\hat{E}(s)}{\hat{g}(W_1)\hat{P}A}, \quad (3)$$

where  $n$  is the number of moose groups detected,  $\hat{E}(s)$  is the expected group size,  $A$  is the total area searched ( $A = \text{total length of transects} \times 2[W_2 - W_1]$ ),  $\hat{g}(W_1)$  is the estimated probability of detection at the minimum available sighting distance by all rear seat observers combined, and  $\hat{P}$  is the estimated average probability of detection of a moose group within the search area, given detection at  $W_1$  was 100%. In equation (3),  $\hat{g}(W_1) \times \hat{P}$  is the estimated average probability of detecting a moose group between  $W_1$  and  $W_2$  units of the transect line.

The freeware program DISTANCE (Thomas et al. 2004) is used to estimate  $\hat{P}$  based on observations by the rear seat observers (Appendix 6). Often, detection of individuals or groups are greatly influenced by observer experience or skill, habitat type, and survey conditions. When detection of animals at  $g(W_1)$  is known or can be estimated without bias, and the fitted detection function is flexible enough, estimators of density (equation 3) for the study area are unbiased even if all factors other than distance are ignored during estimation (Buckland et al. 2004). However, if detection probabilities vary greatly across habitat types, more accurate and precise estimates of density can be obtained for the specific habitat types by incorporating covariates into the detection function. For more information on these procedures, see chapters 2 and 3 in Buckland et al. (2004). Alternatively, if sample sizes are large enough to provide adequately fitting models with high precision, individual estimates of  $P$  can be obtained for each combination of relevant factors. In the past, sample sizes obtained during Innoko NWR moose surveys have been inadequate to calculate individual estimates of  $P$ .

Using the program DISTANCE, the following four detection models should be fit to the survey data: uniform key functions with a cosine expansion; a uniform key function with simple polynomial expansion; a half-normal key function with a hermite polynomial expansion; and a hazard-rate key function with a cosine expansion (Buckland et al. 2001). The number of expansion terms in each model should be determined using the program DISTANCE and a stepwise model building process that calculates model AICc values to determine the most parsimonious model (Burnham and Anderson 2002).

These four semi-parametric models are considered sufficiently flexible to provide good fit to most distance data, can yield model robust estimation, and can satisfy the shape criterion described by Buckland et al. (2001). The model with the lowest AICc can be chosen as the final

model for estimating  $\hat{P}$ , or modeling averaging can be used if there is sufficient model selection uncertainty and estimates of  $P$  vary greatly from model to model (Burnham and Anderson 2002). We are not in favor of using a chi-square test to select or evaluate the fit of the final model for estimating  $\hat{P}$  due to the sensitivity of the test to the choice of binning used. We have found that the bias and variance estimates generated from the bootstrap exercise (see below) provide enough information on quality of each model.

It is important to recognize that when left-truncation of the distances occur due to not being able to see moose directly under the aircraft, the measured perpendicular distances need to be shifted by subtracting  $W_1$  prior to importing the data into DISTANCE. In addition, a right truncation distance of  $W_2 - W_1$  must be specified in DISTANCE to obtain reliable estimates of  $P$  for the search interval  $[W_1, W_2]$  (Appendices 5 and 6).

Equation (3) requires that the estimated group size,  $\hat{E}(s)$ , be calculated. Average moose group size for groups observed within a short to moderate distance from the transect line (e.g. 200 m [656 ft]), by any of the three observers, should be used to estimate expected group size. Groups observed further away are subject to detection bias (higher probability of observing larger groups), hence leading to overestimates of mean group size. Moose groups seen within 200 m (656 ft) of the helicopter have been used in previous Innoko NWR surveys to estimate average group. If moose group size is expected to change, or does change substantially across strata, then expected group size should also be estimated for each strata.

Bootstrapping (Manly 1997) is used to estimate the variance and bias of estimated densities within each stratum and their combined area. This involves taking a simple random sample with replacement of transects flown in each stratum, and re-running the entire analysis, possibly including the detection model selection if a fair amount of model selection uncertainty exists, to obtain a large set of re-sampled estimates of moose abundance. The bootstrap process should include an estimation of  $\hat{g}(W_1)$ , model selection for the detection function in the program DISTANCE, and model selection of the logistic regression function estimating probability of detection based on double-observer data.

The number of transects sampled with replacement for a bootstrap should equal the number of transects flown in each stratum. At least 1,000 bootstrap samples should be obtained for reliable estimates of variance and bias. An alpha level of  $\alpha = 0.1$  (i.e., 90% CI) has been used for most moose surveys conducted in Alaska. An approximate 90% bootstrap CI can be calculated using the standard bootstrap CI,

$$\text{Estimate} \pm 1.64(\text{Bootstrap Standard Deviation}). \quad (4)$$

This method assumes the sample of bootstrap estimates is approximately normally distributed (Neter et al. 1996), but even moderate violations of this assumption should result in approximately correct CI's. If a histogram of bootstrap estimates for the parameter of interest (e.g., moose density) is highly skewed then the percentile method should be used; select the 5<sup>th</sup> and 95<sup>th</sup> percentiles of the bootstrap distribution as the lower and upper endpoints for the 90% CI.

Bias of a density estimate (or any other parameter) can be calculated as

$$\text{Bias}(\text{Estimate}) \approx \text{Mean}(\text{Estimate}_b) - \text{Estimate}, \quad (5)$$

where  $\text{Mean}(\text{Estimate}_b)$  is the mean of the bootstrap estimates from the 1,000 samples. A bias  $\geq 5\%$  is cause for concern.

### 3.4.3 Assessing Survey Accuracy

It may be possible to evaluate the accuracy of the estimated total number of moose by attempting to conduct a full census of moose in one or several small areas, provided survey effort and visibility warrant assuming that all moose can be detected. For example, in 2004 all portions of transects over islands on the southern portion of the Yukon River were flown under the standard protocol, and then an attempt was made to count all moose on the islands. All observations obtained under the usual distance sampling protocol with a double-observer were used to estimate probability of detection on the transect line and the overall detection function. However, the length of transect and area searched on the islands were removed from the analysis, because the total number of moose counted on the islands was added to the total number of moose estimated for the remaining study area. Zero variance was assigned to the moose island counts, because it was judged that no island moose were missed and movement to and from the islands by moose was random for the duration of the study. This process reduced the coefficient of variation and percent CI half-width for estimates of the total number of moose in the study area by 7% and 10%, respectively (Nielson et al. 2004).

A finite population correction factor (fpc) should be used if: the population of moose can be considered fixed during the course of the study (i.e., the survey window is short enough such that movement in or out of the study area does not occur); individual moose are not likely to be counted  $>1$  time during one survey (e.g., search areas do not overlap and movement from one transect to another by an individual is unlikely); and no moose are entering or leaving the study area (Cochran 1977). If these assumptions are met, ignoring the fpc will result in an overestimate of the standard error of moose abundance. To adjust the standard errors for moose density or total number of moose for having surveyed a large proportion of the study area, use

$$SE_{fpc} = SE \times (1 - A / A_t), \quad (6)$$

where  $SE$  is the standard error estimated by bootstrapping (i.e., standard deviation of bootstrap estimates),  $A_t$  is the total area in the sampled stratum, and  $A$  is the total area searched in the stratum. If the sampling fraction does not exceed 10%, the effect of ignoring the fpc will be negligible.

### 3.4.4 Group Size Alternatives

Buckland et al. (2001:13) define a group as “a relatively tight aggregation of objects”. It is recommended that the researcher use all groups seen at closer distances to the transect line, where the detection probability remains relatively high (Buckland et al. 2001). Visual inspections of several histograms with alternative bin widths can aide in this decision. While not



used in the past on Innoko NWR, other methods for estimating group size include using a weighted average of cluster sizes and various regression estimators (Buckland et al. 2001). Alternatively, data analysis could be stratified by cluster size or cluster size could be used as a covariate when fitting the detection function.

### **3.4.5 Bootstrapping**

As indicated above, bootstrapping for estimation of variances should include model selection. If model averaging is used, bootstrapping should include that component of the analysis as well. Model averaging can substantially increase the amount of time and computing power needed for bootstrapping, although this is not sufficient justification for not considering model averaging when model selection uncertainty is high. If detection on the line is known to be 100%, bootstrapping can be done entirely within the program DISTANCE (see program manual and/or help files). If detection on the line is estimated to be <100% based on the survey data, the bootstrap process involves inputting each bootstrap sample into DISTANCE and merging the results (see Appendices 5.4, 5.5, and 6.2). The component of the bootstrap process completed in program DISTANCE can be troublesome due to limitations with the program, including unintelligible warning or error messages. The researcher must decide if error or warning messages are the result of improper use of the program, program limitations, or can be linked to a specific bootstrap sample. Usually, one would consider discarding information from bootstrap replicates associated with error messages to avoid estimating an artificial bias or unrealistic variance and CI. However, this response should only apply to situations where a small percentage, say < 5%, of the bootstrap replicates result in error messages. More frequent errors are likely indicative of a larger problem with the data or model assumptions. The analyst should carefully consider the answers provided by any analysis. For example, if one of the models fit in the program DISTANCE suggests that 100% of the animals were seen within 400 m (1,312 ft) of the transect line, it is wise to investigate this result to determine if this result is a program error or supported by the sample data.

### **3.4.6 Data Pooling**

If detection probability (SCF) is quite different between strata, and data are pooled to estimate a detection function, stratum specific estimates may be biased, while estimates of abundance for the larger study area are believed to be pooling robust (Buckland et al. 2001). It is possible to avoid this pooling bias by fitting a separate detection function for each stratum, but small sample sizes may prohibit this approach. An alternative method is to include covariates into the detection function (Buckland et al. 2004). This method and the potential biases associated with pooling data needs to be investigated further.

Another approach to increasing precision involves pooling data across years to better estimate probabilities of detection. Pooling should only be used if survey protocol is identical, survey crews are equally skilled at observing moose, and survey conditions are similar; here, judgment of the project leader is needed when comparing different surveys to determine if pooling is appropriate. Careful visual inspection of histograms of observed distances, coupled with expert opinion, can help determine if the relative frequency of observations at different distances are similar across surveys, indicating that pooling may be appropriate. If pooling is used, separate estimates based on original data should be compared to pooled estimates. Generally, pooling is used to increase precision, so if estimates of probability of detection change substantially when

data are pooled, we recommend against using pooled data. Pooling across survey years was done on Innoko NWR in 1998 (Erickson and Skinner 1998). That analysis showed that pooling only slightly decreased the CV and size of the CIs.

### **3.4.7 Assumptions and Availability Bias**

Every assumption should be carefully evaluated to insure that the survey protocol and survey conditions can be expected to result in quality data for statistical analysis. It is also important to note that methods described in this protocol do not adjust for individual or groups of moose near the minimum available sighting distance but unavailable to be seen by an observer due to obstructions or snow cover. The only method we are aware of for adjusting for this availability bias involves using known moose locations to check observer detections. For example, if several moose were collared with radiotelemetry or GPS collars, pre-survey flights could verify their location and then actual survey flights could be used to determine sightability. Moose not available to be seen near the minimum available sighting distance results in estimates of abundance that are biased low. However, if availability bias is relatively constant over time, the biased estimates of moose abundance from line transect sampling can still be used to obtain reliable estimates of trend.

### **3.4.8 Census Counts**

As mentioned earlier, empirical counts over some areas can help increase the precision of abundance estimates, if it is reasonable to assume that all individuals are actually detected during the census. Islands on the Yukon River within Innoko NWR offer a unique opportunity to intensively survey a very small area containing relatively large numbers of moose to obtain a reliable moose count, based on the assumption of 100% detection of all individuals. These islands are relatively small and do not have thick, dense cover, allowing for the possibility of an accurate count of the total number of moose using all three observers during an exhaustive search. In 2004, comparison of moose abundance estimated from line transect sampling using the standard survey protocol on the islands (87 moose) to the empirical "complete count" (88 moose) indicated that few if any moose were missed during a census of the islands. Conversely, one could also argue that the standard line transect survey was very representative and this exercise provides a test of the accuracy of the line transect survey in similar habitat on the remainder of the Refuge.

It is recommend that periodically, surveys include standard line transect surveys in small areas with high moose density where census counts might be used to compare the methods and improve precision of the overall survey. If the assumptions of both methods are satisfied, the estimates of moose abundance should be approximately the same on these small "test" areas and confidence in both methods would be gained. Precision of estimated totals for the survey area would also be improved. Project leaders should evaluate the costs of the two techniques on such areas and determine if the comparisons and reduction in variance is justified. Consideration of using the two techniques in other habitat types for further evaluation of relative costs, precision, and accuracy should also be given.

## **3.5 Estimation of Change and Trend**

Two of the most difficult challenges in wildlife and environmental research are modeling change and testing for trend in data (Edwards 1998). To further complicate issues of designing and

analyzing surveys over time, the researcher has the choice of estimating net change between two points in time (e.g., 2004 abundance vs. 2000 abundance) or estimating the average net change over time (e.g., average trend from 1992 through 2004). We believe that estimation of a net change between two points in time (i.e., the difference between moose population sizes in 2000 and 2020), and estimation of the average net change (i.e., the average trend in moose population sizes from 2000 to 2020), are secondary objectives of the line transect surveys.

If interest is focused on net change in abundance between just two surveys, we can calculate a 90% confidence interval for a difference in two population totals (Zar 1999). If the 90% confidence interval contains 0, one can declare there was no evidence of a significant difference between the two totals at the  $\alpha = 0.1$  level.

If interest is focused on net change in abundance across three or more surveys, it is often easier to detect (i.e., higher power) a significant average net change in population size using linear regression compared to a test for net change between the two extreme surveys in time. This requires that the assumptions of a linear regression analysis are not violated and the survey protocols and efforts were consistent.

To estimate average net change in population size, the slope statistic from a linear regression analysis can be estimated with time as the independent variable. A real increase or decrease in population size might be exponential; hence a logarithmic transformation of the response might be used prior to regression analysis. A formal test for average net change using linear regression would involve calculating a 90% CI for the trend, or slope, of the regression line (Neter et al. 1996). If a 90% CI for the slope contains 0 one can declare there was no evidence of a significant average net change in population size at the  $\alpha = 0.1$  level. If estimated variances for population totals are quite different across years, say  $> 20\%$ , weighted-least squares regression might be more appropriate.

It can not be emphasized enough that in order to estimate net change, or average net change, is facilitated by consistent stratification, field data collection protocol, and data analysis procedures. Statistical power to detect change is increased for repeated surveys over the same transect lines. Trends can be estimated when stratification changes, but trends in individual strata will not be available.

### **3.6 Analysis examples**

Appendix 4 provides guidance and contains examples of all the data files needed for analyzing data collected from a line transect survey. The examples are from previous Innoko NWR moose surveys; Appendix 5 contains SAS and R code for calculating estimated moose abundance, precision, and bias discussed in the Data Analysis section (i.e., section 3.4). Details for using program DISTANCE to develop detection functions are contained in Appendix 6. These example data files and analysis code are not intended to create a black box for workers to blindly input their data; as has been identified earlier, many decisions need to be made along the analysis path. Rather, the examples in Appendix 4 and the code in Appendix 5 are intended serve as a start point for analyses of moose abundance based on line transect sampling. The code in Appendix 5 contains notes and descriptions to help the reader identify what each section of code is doing. The code (Appendix 5) and most of the example data sets (Appendix 4) came from the

2004 survey conducted within two strata in the southern portion of Innoko NWR. In 2004, empirical island moose counts supplemented estimates of the total number of moose in the remaining habitat off the islands, and the SAS code reflects that additional sampling effort. The SAS and R code, along with example data files, can be used to re-create the survey results reported in “2004 Aerial Survey for Moose (*Alces alces*) on the Innoko NWR” (Nielson et al. 2004).

## 4.0 Literature Cited

- Amstrup, S. C., T. I. McDonald, and B. F. J. Manly. 2005. Handbook of capture-recapture analysis. Princeton University Press, Princeton, New Jersey, USA.
- Anthony, R. M., and R. A. Stehn. 1994. Navigating aerial transects with a laptop computer map. *Wildlife Society Bulletin* 22:674-676.
- Beavers, S. C., and F. L. Ramsey. 1998. Detectability and analysis in transect surveys. *Journal of Wildlife Management* 62:948-957.
- Beyer, H. 2005. Hawth's analysis tools for ArcGIS, version 3.21. <http://www.spatial ecology.com/htools/index.php>.
- Borchers, D. L., S. T. Buckland, and W. Zucchini. 2002. Estimating animal abundance. Springer, London, England.
- Buckland, S. T., D. R. Anderson, K. P. Burnham, and J. L. Laake. 1993. Distance sampling: Estimating abundance of biological populations. Chapman and Hall, London, England.
- \_\_\_\_\_, \_\_\_\_\_, \_\_\_\_\_, \_\_\_\_\_, D. L. Borchers, and L. Thomas. 2001. Introduction to distance sampling. Oxford University Press, Oxford, England.
- \_\_\_\_\_, \_\_\_\_\_, \_\_\_\_\_, \_\_\_\_\_, \_\_\_\_\_, and \_\_\_\_\_, editors. 2004. Advanced distance sampling. Oxford University Press, Oxford, England.
- Burch, J., T. Meier, M. Reid, and J. Lawler. 2004. Moose monitoring protocol for the central Alaska network. Unpublished report prepared for the National Park Service.
- Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Second Edition. Springer, New York, New York, USA.
- Cassey, P., and B. H. McArdle. 1999. An assessment of distance sampling techniques for estimating animal abundance. *Environmetrics* 10:261-278.
- Cochran, W. G. 1977. Sampling techniques. John Wiley and Sons, New York, New York, USA.

- Drummer, T. D., and McDonald, L. L. 1987. Size bias in line transect sampling. *Biometrics* 43:13-21.
- Edwards, D. 1998. Issues and themes for natural resources trend and change detection. *Ecological Applications* 8:323-325.
- Erickson, W. 2002. Results of 2002 aerial survey of moose on the Innoko NWR. Draft. Unpublished report for US Fish and Wildlife Service. Western EcoSystems Technology, Inc., Cheyenne, Wyoming, USA
- \_\_\_\_\_, and R. Skinner. 1998. Results of the 1998 aerial survey of moose on the Innoko NWR. Draft. Unpublished report for US Fish and Wildlife Service. Western EcoSystems Technology, Inc., Cheyenne, Wyoming, USA
- Evans, T. J., A. Fischbach, S. Schliebe, B. Manly, S. Kalxdorff, and G. York. 2003. Polar bear aerial survey in the eastern Chukchi Sea: a pilot study. *Arctic* 56:359-366.
- 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* 22:1-108.
- \_\_\_\_\_, and \_\_\_\_\_. 1987. Estimating moose population parameters. *Swedish Wildlife Research Supplement* 1. pp. 603-617.
- Gilbert, J. R. 1999. Review of previous Pacific walrus surveys to develop improved survey designs. Pages 75-86 in Garner, G. W., S. C. Amstrup, J. L. Laake, B. F. J. Manly, L. L. McDonald, and D. G. Robertson, editors. *Marine mammal survey and assessment methods*. Balkema, Rotterdam, The Netherlands.
- \_\_\_\_\_, L. P., D. Douglas, and L.L. McDonald. 1990. Protocol for estimating the numbers of pacific walruses. Unpublished report for Alaska Fish and Wildlife Research Center, U. S. Fish and Wildlife Service, Anchorage, Alaska, USA.
- Good, R. E., R. M. Nielson, H. H. Sawyer, and L. L. McDonald. 2004. A population level survey of golden eagles (*Aquila chrysaetos*) in the western United States. Prepared for U.S. Fish and Wildlife Service, Denver, Colorado, USA. URL [http://mountain-prairie.fws.gov/species/birds/golden\\_eagle/](http://mountain-prairie.fws.gov/species/birds/golden_eagle/).
- Hay, K. 1982. Aerial line-transect estimates of abundance of humpback, fin and long-finned pilot whales in the Newfoundland-Labrador area. Report to the International Whaling Commission 32:475-486.
- Hedley, S. L., and S. T. Buckland. 2004. Spatial models for line transect sampling. *Journal of Agricultural, Biological, and Environmental Statistics* 9:181-199.
- Johnson, B. K., F. G. Lindzey, and R. J. Guenzel. 1991. Use of aerial line transect surveys to estimate pronghorn populations in Wyoming. *Wildlife Society Bulletin* 19:315-321.

- Laake, J., J. Calambokidis, S. D. Osmeck, and D. J. Rugh. 1997. Probability of detecting harbor porpoise from aerial surveys: estimating  $g(0)$ . *Journal of Wildlife Management* 61:63-75.
- LeResche, R. E., and R. A. Rausch. 1974. Accuracy and precision of aerial moose censusing. *Journal of Wildlife Management* 38:175-182.
- Manly, B. J. F. 1997. Randomization, bootstrap and monte carlo methods in biology. Second Edition. Chapman and Hall, London, England.
- \_\_\_\_\_, L. L. McDonald, and G. W. Garner. 1996. Maximum likelihood estimation of the double-count method with independent observers. *Journal of Agricultural, Biological, and Environmental Statistics* 1:170-189.
- \_\_\_\_\_, \_\_\_\_\_, D. L. Thomas, T. L. McDonald, and W. P. Erickson. 2002. Resource selection by animals: statistical design and analysis for field studies. Chapman and Hall, London.
- McCullagh, P. and J. A. Nelder. 1989. Generalized linear models. Second Edition. Chapman and Hall, Boca Raton, Florida, USA.
- McDonald, L. L., G. W. Garner, and D. G. Robertson, D. G. 1999. Comparison of aerial survey procedures for estimating polar bear density: results of pilot studies in northern Alaska. Pages 37-52 in Garner, G. W., S. C. Amstrup, J. L. Laake, B. F. J. Manly, L. L. McDonald, and D. G. Robertson, editors. *Marine Mammal Survey and Assessment Methods*. Balkema, Rotterdam, The Netherlands.
- Nielson, R. M., W. Erickson, and L. L. McDonald. 2004. 2004 aerial survey for moose (*Alces alces*) on the Innoko NWR. Unpublished report for US Fish and Wildlife Service. Western EcoSystems Technology, Inc., Cheyenne, Wyoming, USA
- Neter, J., M. H. Kutner, C. J. Nachtsheim, and W. Wasserman. 1996. Applied linear statistical models. Fourth Edition. McGraw-Hill, Boston, Massachusetts, USA.
- Overton, W. S., and Stehman, S. V. 1995. Design implication of anticipated data uses for comprehensive environmental monitoring programmes. *Environmental and Ecological Statistics* 2: 287-303.
- \_\_\_\_\_, and \_\_\_\_\_. 1996. Desirable design characteristics for long-term monitoring of ecological variables. *Environmental and Ecological Statistics* 3(4):349-361.
- Quang, P. X., and E. F. Becker. 1996. Line transect sampling under varying conditions with application to aerial surveys. *Ecology* 77:1297-1302.
- \_\_\_\_\_, and \_\_\_\_\_. 1997. Combining line transect and double count sampling techniques for aerial surveys. *Journal of Agricultural, Biological, and Environmental Statistics* 2:229-242.

- R Development Core Team. 2005. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org>.
- SAS Institute, Inc. 2000. SAS/STAT users guide. SAS Institute, Cary, North Carolina, USA.
- Samuel, M. D., E. O. Garton, M. W. Schlegel, and R. G. Carson. 1987. Visibility bias during aerial surveys of elk in northcentral Idaho. *Journal of Wildlife Management* 51:622-630.
- Skinner, R., W. Erickson, L. L. McDonald, and G. Minick. 1997. Estimating moose populations and trends using line transect sampling. Unpublished report. US Fish and Wildlife Service, McGrath, Alaska, USA.
- Southwell, C. 1989. Techniques for monitoring the abundance of kangaroo and wallaby populations. Pages 659-693 in G. Grigg, P. Jarman, and I. Hume, editors. *Kangaroos, wallabies and rat-kangaroos*. Surrey Beatty and Sons, Sydney, Australia.
- \_\_\_\_\_, B. de la Mare, M. Underwood, F. Quartararo, and K. Cope. 2002. An automated system to log and process distance sight-resight aerial survey data. *Wildlife Society Bulletin* 30:394-404.
- Thomas, L., J. L. Laake, S. Strindberg, F. F. C. Marques, S. T. Buckland, D. L. Borchers, D. R. Anderson, K. P. Burnham, S. L. Hedley, J. H. Pollard, and J. R. B. Bishop. 2004. Distance 4.1. Release 2. Research Unit for Wildlife Population Assessment, University of St. Andrews, UK. URL <http://www.ruwpa.st-and.ac.uk/distance/>.
- Ver Hoef, J. 2000. A spatial moose survey estimation method. Unpublished report. Alaska Department of Fish and Game, Fairbanks, Alaska, USA.
- \_\_\_\_\_. 2001. Predicting finite populations from spatially correlated data. 2000 Proceedings of the Section on Statistics and the Environment of the American Statistical Association, pp 93-98.
- \_\_\_\_\_. 2002. Sampling and geostatistics for spatial data. *Ecoscience* 9:152-161.
- Walter, M. J., and J. Hone. 2003. A comparison of 3 aerial survey techniques to estimate wild horse abundance in the Australian Alps. *Wildlife Society Bulletin* 31:1138-1149.
- Ward, R. M. P., W. C. Gasaway, and M. M. Dehn. 2000. Precision of moose density estimates derived from stratification survey data. *Alces* 36:197-203.
- White, G. C., R. M. Bartmann, L. H. Carpenter, and R. A. Garrott. 1989. Evaluation of aerial line transects for estimating mule deer densities. *Journal of Wildlife Management* 53:625-635.

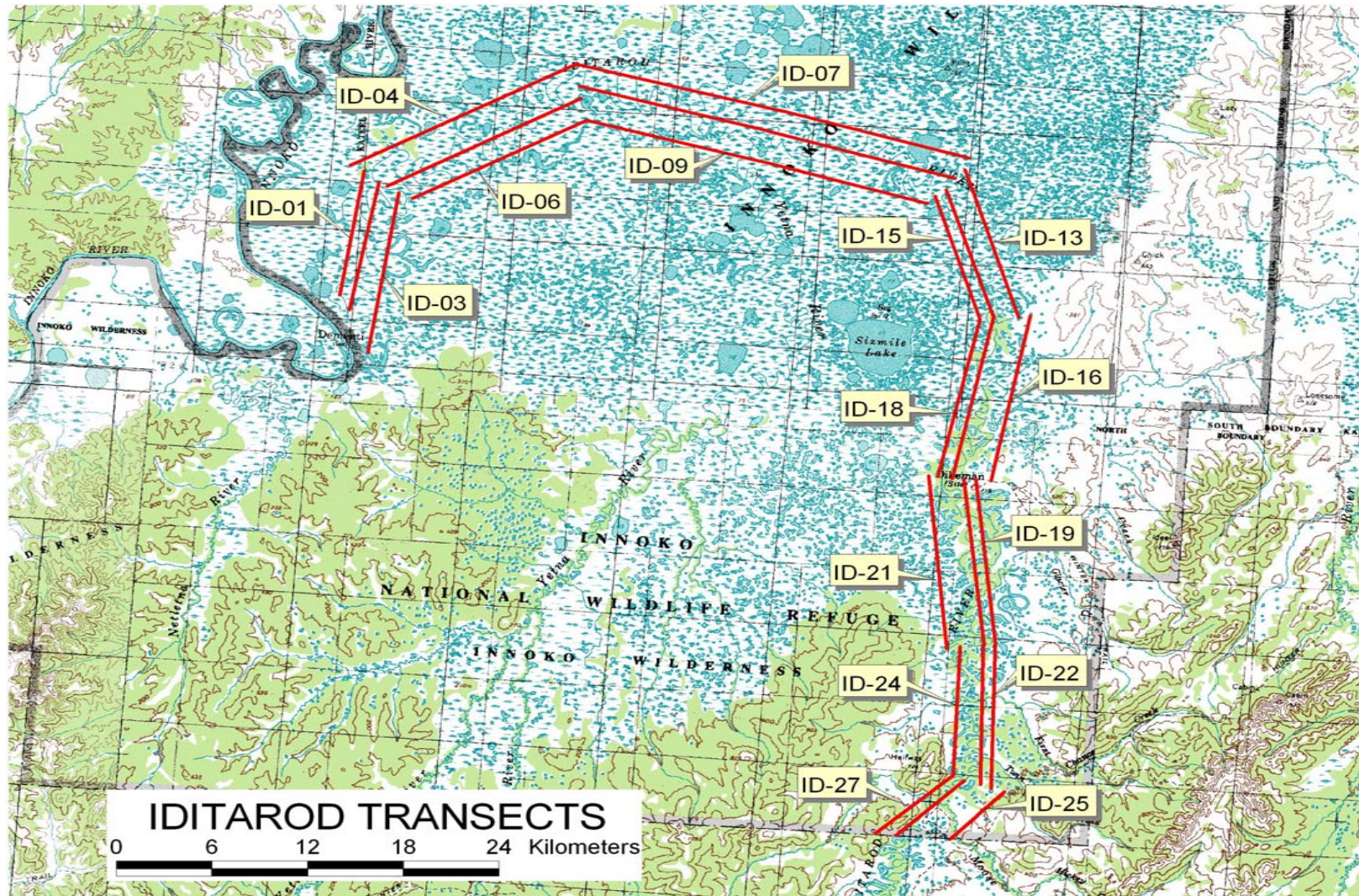
Whittaker, D. G., W. A. Van Dyke, and S. L. Love. 2003. Evaluation of aerial line transect for estimating pronghorn antelope abundance in low-density populations. *Journal of Wildlife Management* 31:443-543.

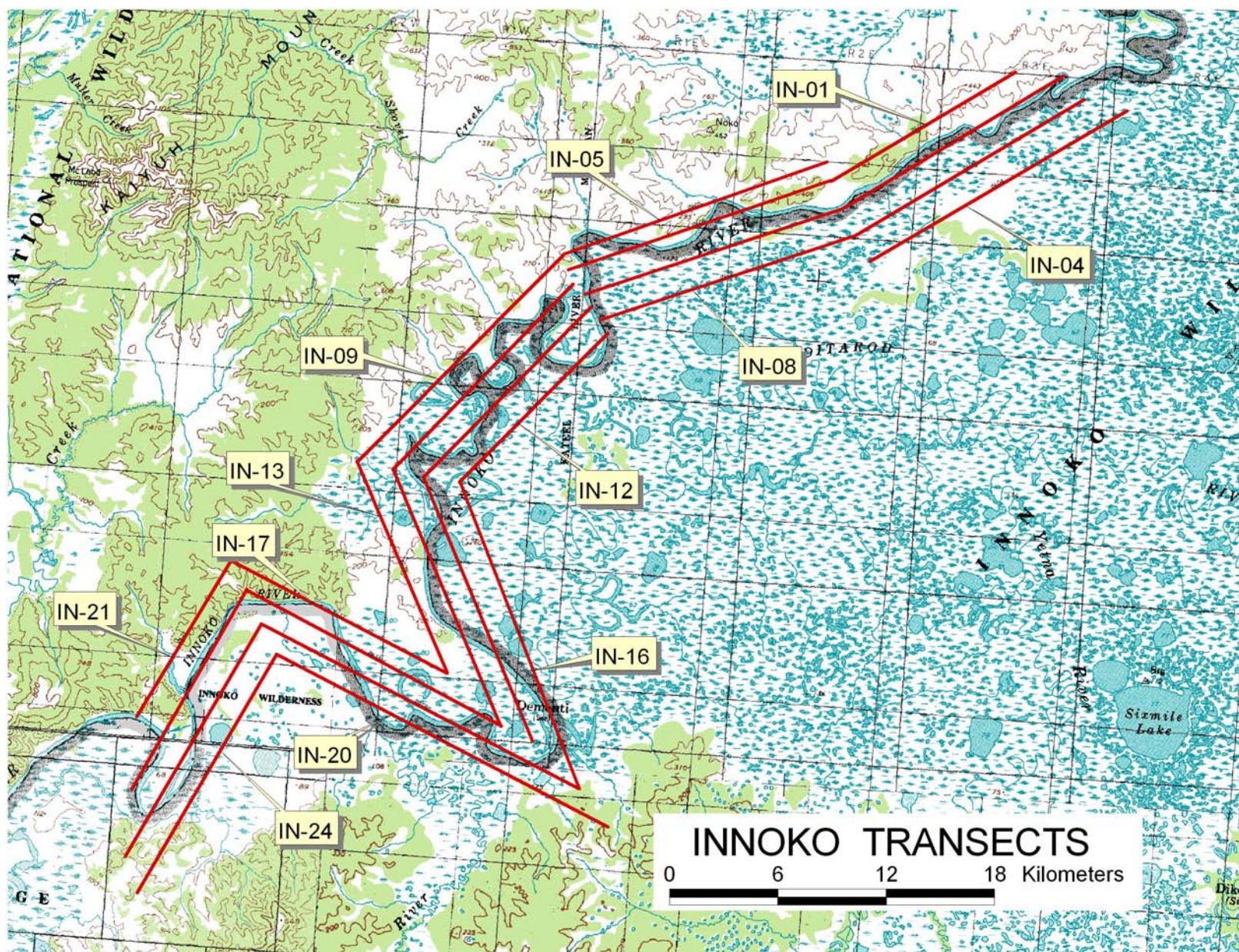
Zar, J. H. 1999. *Biostatistical analysis*. Fourth Edition. Prentice Hall, New Jersey, New Jersey, USA.

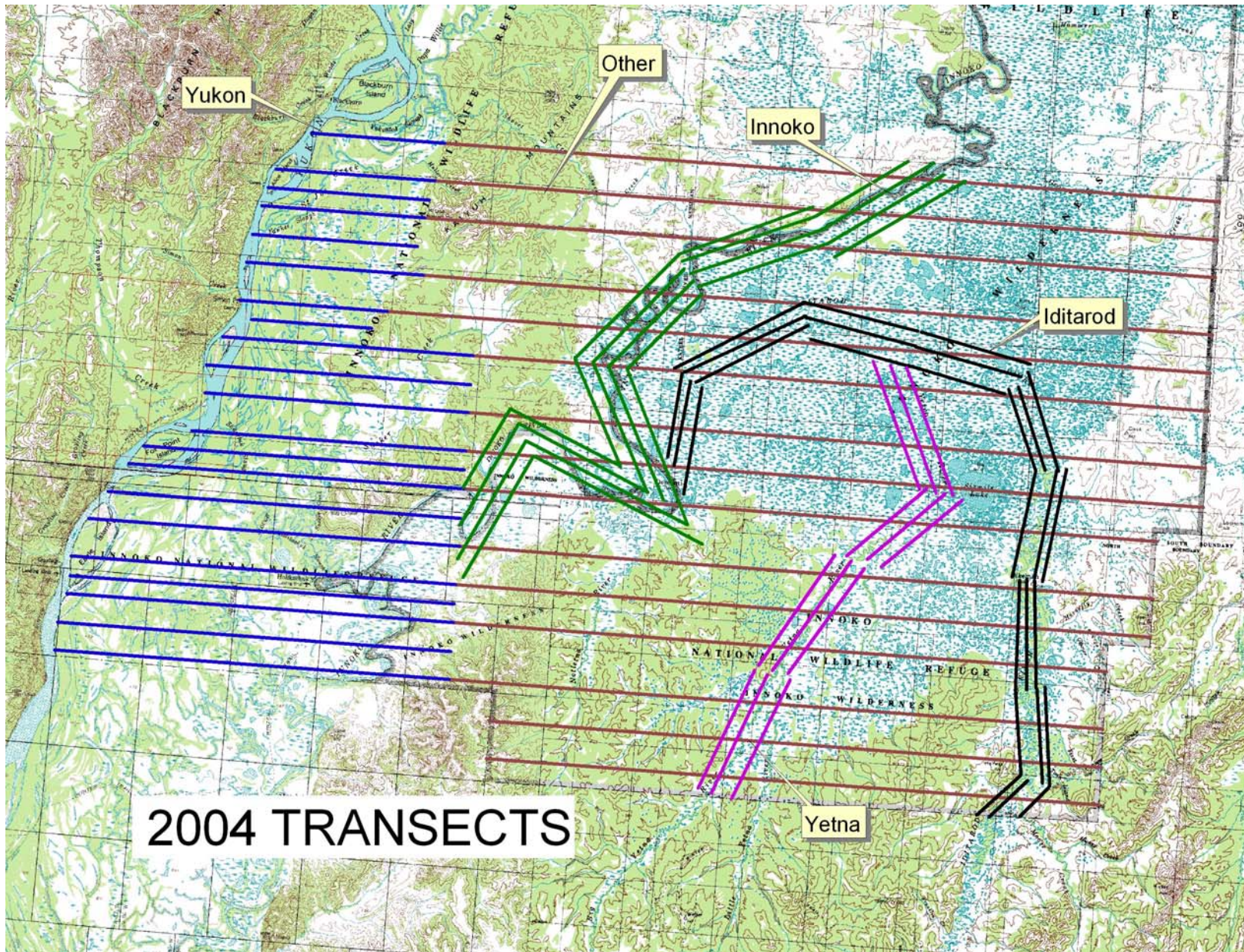


## Appendix 1: Example Field Maps

These example field maps are reduced from their original size, but are sufficient to provide the reader with examples.







## Appendix 2: Field Data Form

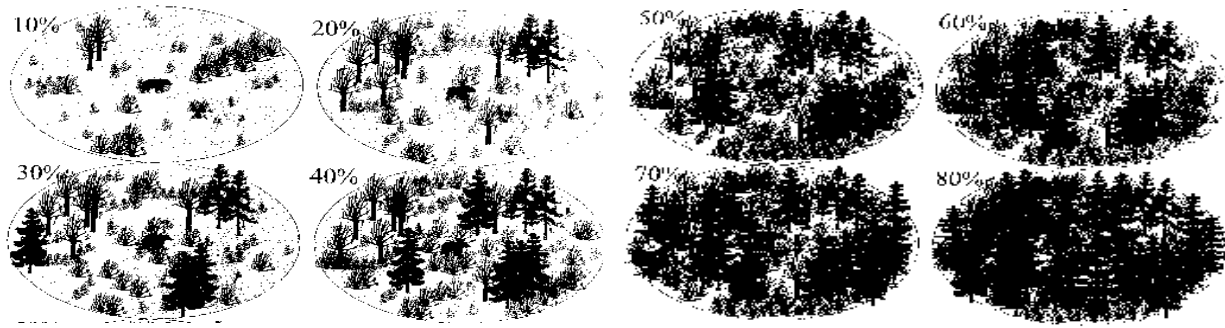
Date \_\_\_\_\_ Name \_\_\_\_\_ Page # \_\_\_\_\_

Data Recorder \_\_\_\_\_ Back Left \_\_\_\_\_ Back Right \_\_\_\_\_

Strata \_\_\_\_\_

Notes \_\_\_\_\_

Line ID	Rec ID	Seen By (FR, BR, BL)	Distance (m)	Total Adults	Total Calves	%Cover		Notes
						10 m	50 m	



## Appendix 3: Voice/GPS Survey Program Manual

This computer program (versions 3.1 and 6.2) allows an observer to talk through a microphone into a computer and attach GPS positions to observations. An option is available for simultaneously displaying a moving map at 1:250,000 or 1:1,000,000 scales. The full suite of files needed for version 3.1 is included on the CD that accompanies this document. In addition, there are a few files associated with version 6.2, but we have not been able to obtain all the necessary files to make this version usable at this time. Hopefully, in the near future we can obtain the necessary files for version 6.2. Much of the following was copied directly from the manuals (versions 3.1 and 6.2) included with the program (Voice/GPS Survey Program Manual, USFWS, Migratory Bird Management, Juneau, Alaska, USA, August 2003), with modifications for application to moose surveys. The program files, as well as the original manual, are included on the CD that accompanies this document.

### INSTALLATION

Place the CD into the CD drive. Copy all files from the folder ‘\VoiceGPS version 3.2’ from the CD to your hard-drive. Double-click on *setup.exe*.

This will create a program group called *Voice/GPS Survey*. Within this group are two programs, Record and Transcribe. *RECORD program*

### Setup

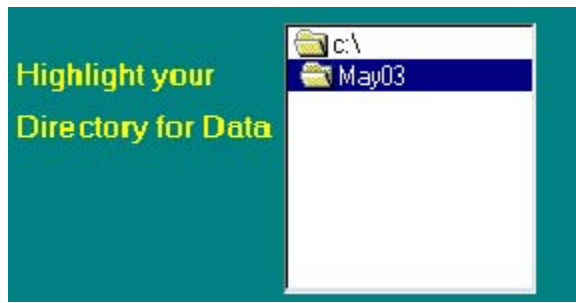
Before you run the Record program for the first time, create a directory on your hard disk where the voice files and other files related to your survey will be stored (e.g., C:\survey).

The first time you run the Record program you must specify a survey name. It may have, at most, 5 characters. It will serve as the root file name for all voice files, position files, header file, species file, sequential log file, and stored history file.

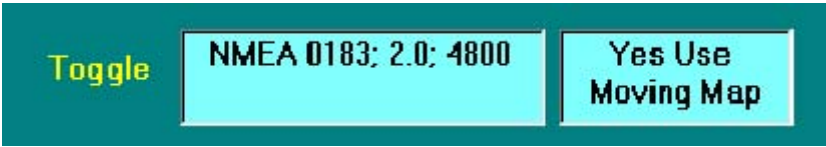


An example of a survey name would be IN04 (for Innoko NWR 2004). Voice files for this survey would look like IN04001.WAV (for the 1<sup>st</sup> wav file from the Innoko survey in 2004). Position files would look like IN04001.POS (for the 1<sup>st</sup> pos file from the Innoko survey in 2004). The header name file would be IN04HEA.DER. The species list file would be IN04SPE.CIE. The sequential log file would be IN04SEQ.LOG. The stored history file would be IN04STO.RED. See the File Description section for a more detailed description of each file type.

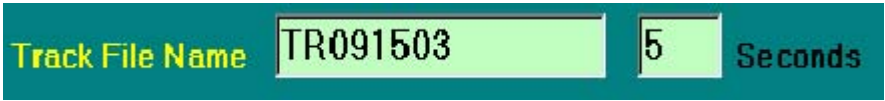
Next highlight the directory where your survey files will be stored.



If using GPS, toggle through the GPS output types to select the type that matches the output from your GPS. If you are using the moving map then also click on the box to the right to toggle it to read "Yes Use Moving Map".



The Track File Name will show the default file name. It has the letters TR followed by the current date (example: TR022904 for February 29, 2004). Click on the box and type a new name if desired. The default time between stored track positions is 5 seconds, which needs to be changed to 10 seconds for moose surveys.



If the time is not correct, use the spin arrows to set it. When the GPS receives its first position fix, the computer clock will reset the minutes and the seconds to match the GPS satellites. The hour will remain unchanged.



You may test your microphone by placing the cursor on [Test Mic] and holding the right or left mouse button down while you talk into the microphone. When you release the mouse button, it will play back the test.



The next screen will have a white background unless the moving map option is in effect. The root file name with the sequential file number (example: IN04001) will be displayed. This file

name increments one sequential file number each time you decide to store a block of data by hitting the left mouse button. The latest file name used is recorded in the -----SEQ.LOG file. Do not ever move or delete the -----SEQ.LOG file from your survey directory.

Note: The GPS Fix box will remain red until a position is received using the satellites, at which time it will turn green. If at any time the latest satellite fix is more than 15 seconds old, the GPS Fix box will turn red and an audio warning will sound. This indicates that the GPS receiver is currently not receiving enough quality signals for a fix.

To record a moose observation, press and hold the right mouse button down. Release the button to stop recording. The status box will read [RECORDING] while recording, and [STOPPED] when not recording. The Location Number will increment upwards with each moose group observation. The mouse button may be held down for a maximum of 60 seconds at which time recording of that observation will automatically stop. This is to protect against enormously large audio files occurring because the mouse button is unwittingly held down. Therefore, limit talking to less than 60 seconds per push of the mouse button.

Note: Pushing the right mouse button down marks the time of the observation, which will occur between two known GPS locations. If those known locations are less than 15 seconds apart, the program interpolates the observation position on a straight line between the two known points.

To store a block of observations on to the hard drive, press the left mouse button. After storing, the status box returns to [OFF] and the computer is ready to start a new block of recordings. Observations should be stored fairly often to limit the size of the voice files, usually 50 to 100 observations per file. The maximum number of observations per file is 500.

To quit this screen, the last recordings must have been saved and the left mouse button must be used on the "QUIT" button.

## **Moving Map Option**

If you have toggled to "Yes Use Moving Map", there are additional options and considerations.

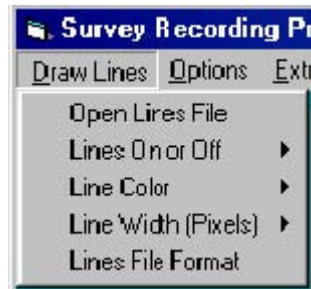
If you will be entering point location data using a keyboard, you must have clicked on the Point File Name in the opening window, and selected a file name where these data will be saved. This will hold the manually entered data as opposed to the voice-recorded data.



User lines may be displayed on the moving map by selecting the menu item Draw Lines. Selecting the submenu item Open Lines File will allow you to choose the file where the line coordinates are stored. The lines file is an ASCII file of Latitude and Longitude pairs describing the end points of each desired line. An example of this type of file would be:

63.13653, -158.75990, 63.05544, -158.77750  
63.13020, -158.74110, 63.04658, -158.76420  
63.12392, -158.71500, 63.01925, -158.73700

Previous track files may also be displayed with the Draw Lines option. The color of the lines and the width of the lines may be selected using the Draw Lines menu items Line Color and Line Width. The recording window will show the track of the helicopter overlaid onto a digital map. The map will be the USGS 1:250,000 topographic map if it is found in the hard drive or the CD Drive. If it is found on the CD Drive it will be copied onto the C: drive for future use. To observe the 1:1,000,000 aeronautical charts you must zoom out twice. A special map provided by the user may also be overlaid if desired.



The compass direction of the helicopter path across the ground is displayed in degrees (magnetic) inside the white box inside the green box. If your GPS unit does not provide magnetic variation information, the reading will be degrees true. The ground speed is also shown below this. The ground speed is averaged over the last 30 seconds. The circle and needle can be used to maintain a heading. Click on the circle and the needle will reset to point to the current direction of the aircraft. The needle will continue to point towards that heading until the green box is clicked again.

There are buttons for zooming in or out. The buttons around the [Center] button allow you to place yourself off center of the map in any of eight quadrants.

Location data may be tied to the current position of the aircraft or a position selected off of the helicopter track. Data can be typed observations or voice recordings. Remember that if data is to be typed, a point file name must have been specified on the opening screen

The right and left mouse buttons have distinctly different uses. No matter where the arrow is on the screen, using the right mouse button will grab the current location of the aircraft. If a point file name was specified and if the right mouse button is just clicked (not held down), it will open the attribute box for typing in the observation. If the right mouse button is held down it will begin recording of a voice observation through the microphone. The location will correspond to the aircraft position when the right mouse button was pushed. If a point file name was specified, then the left mouse button may be used to grab the position of the arrow if it is located on the map. This allows the user to select a location separate from the location of the aircraft. Clicking the left mouse button will open the attribute box for typing the observation. The location will correspond to the position of the arrow, not the position of the aircraft. If the user wants to record the observation into the microphone rather than typing it, he/she has 5 seconds to begin



the recording with the right mouse button after the left mouse button click has been used to grab the position of the arrow. Whenever an observation is stored, its position is marked on the map with a circle. The most current 200 observations will always be redrawn when the screen is redrawn.

To save data with the left mouse button, you must click on a white portion of the screen (not on the map or any of the other buttons that have other purposes).

The digital map tiles must be in the C:, D:, E: or F: drives. The 1:1,000,000 scale maps must be in a subdirectory named \MOVEMAPS\MILLION\ and they will have a name such as MN60W155.bmp. The name indicates the corner of the map tile which has the smallest absolute values for latitude and longitude, in this case Latitude=60N and Longitude=155W. Each map is 5 degrees of latitude by 5 degrees of longitude. The 1:250,000 scale maps are in subdirectories by degree of latitude. An example would be \MOVEMAPS\LAT58\. Each tile is .5 degrees of latitude by 1 degree of longitude. A tile file example would be N585W135.bmp which indicates the southeast corner is 58.5N and 135W. If this were in the southern hemisphere, the directory name would be Lat-58.

A special map provided by the user may be overlaid on the screen. The program expects this map to be rectangular (Decimal Degrees Projection) and in bitmap form. If the map file name is MYMAP.bmp then a corresponding reference file is necessary called MYMAP.ref. This file has on the first line the latitude and longitude of the northwest corner in decimal degrees, and on the second line the latitude and longitude of the southeast corner.

While in the simulation mode, the simulation speed may be selected from the Simulation Speed menu.

### **TRANSCRIBE Program**

Provide the survey name and the directory where the data files reside as detailed under the RECORD Program portion of this manual. Then press [Continue].

If a header file cannot be found, a message box will advise you. Similarly, if a species list file is not found for this survey, a message will advise you.

### **Creating a File of Header Names**

To create a file of header names, use a text editor such as NotePad or WordPad. The file is simply a list of header names. The maximum number of header names is 30. Save the header list with the name -----HEA.DER. For our example it would be IN04HEA.DER.

### **Creating a file of Species Names**

To create a file of species names, use a text editor such as NotePad or WordPad . The file is simply a list of three types of locations stored: (1) 'MOOSE'; (2) 'START'; and (3) 'STOP'. Species name files may be copied from one survey to another such as IN04SPE.CIE copied to IN06SPE.CIE .

## **Selecting the Voice File**

During the recording process, files were assigned sequential numbers (i.e., IN04001.WAV, IN04002.WAV, IN04003.WAV, etc.). Choose the File Sequence number for the desired file by inputting it directly or by using the spin up and spin down arrows.

## **Transcribing Voice Recordings**

To hear the first observation, the mouse must be used to click the [Forward] button. Use the [Pause], [Repeat], [Back], and [Rewind] buttons as needed to hear the desired recordings. Input header values by clicking to the right of the header name, then typing a new value. Continue this process until all desired header changes have been made. Use the scroll bar at the right if necessary. During subsequent transcribing, a shortcut key may be used in place of the [Forward] button. The shortcut key is the upper left key with the characters ~ and ` on it. You may wish to label this key as the shortcut forward key. The [Hint] button reminds the user of this shortcut.

An observation consists of a location number, the type of location (i.e., either a transect starting point, stopping point, or a moose group observation) and a number seen (for moose group locations). Click on the 'species' name from the species list. The location number and the species name will appear in the boxes at the top of the list. If hot keys are used, they will substitute for clicking on a species name unless the focus has been changed to another part of the form such as the Header Box.

Next, input the number seen. If Enter is pressed at this time, the observation will be transferred to the observation list at the left part of the screen. If not, the observation will not be transferred until a species name is selected again, or the Forward key is hit.

To add additional or auxiliary information to an observation, you may place a comma after entering the number of animals observed and then typing the additional characters.

## **Deleting, Inserting and Editing in the Observation List**

To delete an observation from the left hand observation list, first highlight the observation by clicking on it, then click the [Delete] button.

To insert an observation into the left hand observation list, first highlight the observation which will end up below the inserted observation, then click on the [Insert] button. This observation and those below it will move down one notch and leave a blank space. Then create the new observation as usual and hit the Enter key. It will be inserted into the vacant slot on the observation list.

To edit an observation in the observation list, click on the observation. It will appear highlighted at the top of the list box. Edit the observation as desired and then press Enter. The edited observation will return to its proper position in the list.

## **Changing the Font Size in the Species List or the Observation List**

The font size for the species names or the observations may be changed using the pull down menu Fonts.

## **Saving the Transcribed Observations**

Before a new file of voice recordings is selected for transcription or the header values are changed, the transcribed observations must be saved by clicking the [SAVE] button. A dialog box will appear asking for a file name to append the data to. This is the final product. For each observation there are header values, the sequential file name, the latitude and longitude position, the number of seconds after midnight, the number of seconds lag time between the last GPS fix and the beginning of the voice recording, the species name, and the number seen. There are no restrictions on what this file name can be.

## **File Management Considerations**

There are nine file types. Six of them may be edited by the user. These are the Header File (-----HEA.DER), the Species File (-----SPE.CIE), the Header Check File (-----HEA.CHK), the Final Data File (Example\_Transcribe\_Output), the Track File (TR022904) and the Stored History File (-----STO.RED). The other three files are handled by the program only. They should never be altered by the user. The first is the Log File (-----SEQ.LOG) which stores the last sequential file number used and other information about user preferences. The second are the Position Files (-----001.POS) which store for each observation the start and stop positions in the .WAV file, the latitude and longitude, the number of seconds since midnight and the time delay from the last GPS fix. The third are the Sound Files (-----001.WAV), which store the digital sound recordings. The Position Files and the Sound Files are numbered sequentially as voice recordings are periodically saved during the survey.

The Stored History File is not a critical file. It merely provides a history of the sequential file numbers that have been transcribed and stored to help the user keep track.

## File Examples

IN04HEA.DER  
Innoko2004

IN04SPE.CIE  
START  
MOOSE  
STOP

IN04SEQ.LOG  
Last Record File = 12  
GPS Yes  
Trimble 2000; R1, 9600

IN04001.POS  
21

0	3000	62.84807	-159.56810	48187.5	0.0
3000	7000	62.84884	-159.29874	48779.87	0.0
7000	10000	62.84945	-159.21285	48985.57	0.0
10000	12000	62.82192	-159.22526	49509.4	0.0
12000	13000	62.82193	-159.22814	49514.45	0.0
13000	16000	62.82174	-159.28807	49630.62	0.0
16000	21000	62.82111	-159.30344	49839.11	0.0
21000	25000	62.82091	-159.31716	49867.18	0.0
25000	27000	62.82088	-159.35710	49954.35	0.0
27000	30000	62.82067	-159.61008	50503.55	0.0
30000	31000	62.82062	-159.62388	50543.86	0.0
31000	34000	62.82072	-159.63120	50566.0	0.0
34000	38000	62.82074	-159.63296	50571.0	0.0
38000	42000	62.82053	-159.65846	50774.22	0.0
42000	47000	62.82057	-159.66863	50800.69	0.0
47000	48000	62.82065	-159.67370	50814.1	0.0
48000	54000	62.82032	-159.72164	50921.8	0.0
54000	58000	62.81996	-159.73900	50962.5	0.0
58000	62000	62.81836	-159.95899	51460.79	0.0
62000	66000	62.81784	-160.05559	51679.01	0.0
66000	68000	62.82072	-160.05697	51713.12	0.0

IN04001.WAV      These files are standard audio wave files.

Example\_Transcribe\_Output.txt

4042004,TR4042004,63.1351,-159.7615, 44275.83,0.0,START,1  
4042004,TR4042004,63.1391,-159.2149, 44375.37,0.0,STOP,1  
4042004,TR4042004,63.1601,-159.7013,44477.22,0.0,START,1  
4042004,TR4042004,63.1602,-159.6731,44560.19,0.0,MOOSE,1  
4042004,TR4042004,63.1574,-159.4372,44598.73,0.0,STOP,1  
4042004,TR4042004,63.1748,-159.405,45026.87,0.0,START,1  
4042004,TR4042004,63.1744,-159.7007, 45105.72,0.0,STOP,1  
4052004,TR4042004,63.1929,-159.7224, 45229.74,0.0,START,1  
4052004,TR4042004,63.1917,-159.693, 45242.24,0.0,MOOSE,1  
4052004,TR4042004,63.1917,-159.6778,45277.53,0.0,MOOSE,1  
4052004,TR4042004,63.1925,-159.5891, 45368.79,0.0,MOOSE,1  
4052004,TR4042004,63.1943,-159.3743, 45400.86,0.0,STOP,1

TR022904

63.10775,-159.40262, 35973.43  
63.10772,-159.40527, 35977.60  
63.10770,-159.40792, 35981.83  
63.10767,-159.41182, 35987.43

IN04STO.RED

1 2 3 4

## Appendix 4: Data Files

### A.4.1 Data Files

The Voice/GPS Survey program (Appendix 3), analysis code for SAS (Appendix 5.1), bootstrap code for SAS (Appendix 5.4 and 5.5), and program DISTANCE (Appendix 6) require various user created input data files (Table A.4.1). To aid the researcher in data management and analysis, this appendix describes the various user created data files needed for analysis of line transect moose data. Tables display portions of data sets from previous Innoko NWR moose surveys to further illustrate the data file descriptions. Copies of all example data files are included on the accompanying CD.

Table A.4.1. User created data input files for line transect moose survey data collection and analysis.

Data File	Used By
Transect_File.csv	Voice/GPS Survey
Survey_Observations.csv	Analysis.SAS and Bootstrap.SAS
Transects_Flown.csv	Analysis.SAS and Bootstrap.SAS
Transects_Flown_Minus_Islands.csv	Analysis.SAS and Bootstrap.SAS
Bootstrap_Distance_Output_Combined.csv	Bootstrap_Distance.SAS

#### Transect\_File

In order for transect lines to be displayed on the computer during a survey, a file containing the transect end points is needed. Transect waypoints for the Voice/GPS Survey program (Appendix 3) are contained in a comma delimited ASCII file (Table A.4.2). Each row contains the start and end points for one transect line in decimal degrees, NAD27 datum. RECORD will not accept input waypoint files that are not comma delimited or have column or row identifiers. Data are organized as latitude 1, longitude 1, latitude 2, and longitude 2. Latitude 1 and longitude 1 identify the start of transect whereas latitude 2 and longitude 2 identify the end of the transect. As most of Alaska is in the western hemisphere, longitude values are negative values; latitude values are always positive due to Alaska being located in the northern hemisphere.

Table A.4.2. First 6 rows of transect waypoint locations from the file 'Example Transect File.csv' for the Voice/GPS Survey program. Input sequence is Latitude1, Longitude1, Latitude2, Longitude2.

63.13653, -158.75990, 63.05544, -158.77750
63.13020, -158.74110, 63.04658, -158.76420
63.12392, -158.71500, 63.01925, -158.73700
63.21410, -158.50820, 63.14027, -158.77790
63.19258, -158.49820, 63.12812, -158.73070
63.17817, -158.48900, 63.12117, -158.69720

#### Survey\_Observations

At the end of each survey day, the program TRANSCRIBE (Appendix 3) is used to both obtain the start and end points of each transect flown, as well as integrate recorded moose locations with sound files containing data associated with each moose observation. The format of the data are: header information (e.g., IN04 for Innoko NWR, 2004), sequential position file name, latitude and longitude position in decimal degrees, the number of seconds after midnight, the

number of seconds lag time between the last GPS fix and the beginning of the voice recording (if the time between known points before and after the observation exceeds 4 seconds), the species name or start and end point identifier, and the number seen (Table A.4.3).

Table A.4.3. First 7 records from the file Example\_Transcribe\_Output.txt, a raw data file produced by the Voice/GPS Survey program TRANSCRIBE.

4042004,TR4042004,63.1351,-159.7615,44275.83,0.0,START,1
4042004,TR4042004,63.1391,-159.2149,44375.37,0.0,STOP,1
4042004,TR4042004,63.1601,-159.7013,44477.22,0.0,START,1
4042004,TR4042004,63.1602,-159.6731,44560.19,0.0,MOOSE,1
4042004,TR4042004,63.1574,-159.4372,44598.73,0.0,STOP,1
4042004,TR4042004,63.1748,-159.405,45026.87,0.0,START,1
4042004,TR4042004,63.1744,-159.7007,45105.72,0.0,STOP,1

The electronic data, as well as the voice recordings themselves must be cross checked against the field data forms completed by the data recorder and both observers. Discrepancies among any of the field data forms or between the field data forms and the voice recordings the must be reconciled as quickly as possible to ensure data quality.

Once completed, the Transcribe\_Output.txt file needs to be imported into either a database manager or a spreadsheet program. Past workers have found that using a spreadsheet program, such as Excel, has proved very useful while a survey was in progress. A master data file which contains all the data for each observation, as well as all observer names, and survey condition information is then created (Table A.4.4).

Table A.4.4. Field names and descriptions for master data file Survey\_Observations.

Field Name	Description
RecNo	Record number, a sequential numbering of observation records
Date	Survey date in mm/dd/yyyy format
Transect	Transect identifier (e.g., IN-03, YU-12, RW-01)
Recorder	Full name of the data recorder/navigator at the time the observation was made
BackLeft	Full name of the back left seat observer at the time the observation was made
BackRight	Full name of the back right seat observer at the time the observation was made
Obs	Who observed the moose group, values limited to BL, BR, FR, and FR and BR (this assumes the right side of the helicopter is the double observer side)
Wav	Name of the WAV file generated by the Voice/GPS Survey program with the data for the observation
ID	The ID number assigned by the Voice/GPS Survey program
Dist	Distance to the moose group in meters
Total	Total number of moose observed; calculated by summing Adults and Calves
Adults	Total number of adult ( $\geq 1$ -year-old) moose observed in the group
Calves	Total number of calf moose observed in the group
Cov_10	Percentage of vegetative cover with 10 m of the moose group observed
Cov_50	Percentage of vegetative cover with 50 m of the moose group observed
Cloud	Amount of cloud cover at the time the transect was started (e.g., 20%, scattered)
Vis	Visibility conditions at the time the transect was started (e.g., good, moderate, poor)
Wind	Wind conditions at the time the transect was started (e.g., calm, light, moderate)
Light	Lighting conditions at the time the transect was started (e.g., bright, flat)

Table A.4.4 *continued*. Field names and descriptions for master data file Survey\_Observations.

Field Name	Description
Island	Was the observation made on an island (values are Y or N)
Notes	Miscellaneous notes about the observation

This master data file has been named Survey\_Observations.xls in the past as Excel was used to create the file. At the conclusion of each survey day, the Transcribe\_Output.txt data file is appended to Survey\_Observations file. At the conclusion of the survey, a subset file is generated and

Table A.4.5. First 6 records from the spreadsheet Example\_Survey\_Observations.xls prior to creating the file Example\_Survey\_Observations.csv. See Table A.4.4 for field names and meanings.

RecNo	Transect	Obs	Dist	Total	Cov_10	Cov_50	Island
1	YU-10	BL	131	1	40	40	N
2	YU-12	BL	363	1	20	30	Y
3	YU-13	BL	279	1	20	10	N
4	YU-08	FR	224	1	30	40	N
5	YU-08	BL	125	1	30	40	N
6	YU-08	BL	137	1	30	20	N

saved as a csv file named Survey\_Observations.csv (Table A.4.5). Survey\_Observations.csv is used as an input file for both Analysis.SAS (Appendix 5.1) and Bootstrap.SAS (Appendix 5.4).

### Transects\_Flown

At the conclusion of the survey, the actual start and end points of each transect flown are extracted from the Transcribe\_Output.txt files. This is most easily accomplished by importing the Transcribe\_Output.txt files into a text editor (e.g., Notepad). Arrange the endpoint data as start longitude, start latitude and end longitude, end latitude (Table A.4.6). It is important that there is no space after the comma. Also, note that the value 'end' delimits the end of each line definition and that the entire file is also completed with an additional 'end' value. When finished editing the file, save as Transects\_Flown.txt. Note that each transect is identified by a numeric parameter and not an alphanumeric. This requires the user to create a list that relates the transect name (e.g., YU-08) with the numeric identifier in the Transects\_Flown.txt file. This will be critical in a later step.

Table A.4.6. Transect endpoint data for 2 transect lines arranged for importing into ArcInfo (second 2 transects from the file Transects\_Flown.txt).

3	-159.405,63.1748
	-159.7007,63.1744
end	
4	-159.7224,63.1929
	-159.3743,63.1943
end	
end	

Depending on the user's familiarity with ArcInfo, there are 2 ways to import the Transects\_Flown.txt file and obtain the length of each transect. One is to use the menu system within ArcTools, the other is to use (modified as needed) the Arc Macro Language file t-fln.aml (Appendix 4.3).

To import the Transects\_Flown.txt file using the ArcTools menus, follow this procedure. First, start an ArcInfo session, at the arc prompt (Arc: ) enter the command "arctools edit" (without the quotes) followed by <Enter>. Next, set the workspace (i.e., working directory) from the



ArcTools/Workspace menu. Next, open the Command Tools menu bar by selecting Command Tools... from the Tools option; acknowledge the warning message by clicking on OK.

To import the Transects\_Flown.txt file select Conversion from the Command Tools menu bar; select To Arc then Generate. This will open the Generate dialog box. Enter the coverage name t\_fln\_dd in the Output Coverage box, select single precision, select file as the input source, and select lines from the Generate Options. At this point the Select A File dialog box opens, navigate to the directory with the Transects\_Flown.txt file and select that file. Click on OK to close the Select a File dialog box; click OK to generate the coverage. Next select Edit from the Command Tools menu bar; select Coordinates and Define Projection. This opens the ProjectDefine dialog box. Right click in the Cover box to bring up the Select a Coverage dialog box; again, navigate to the directory containing the t\_fln\_dd coverage and select it. Next, click on the Define button to bring up the Input Projection dialogs. First select World as the geographic extent, then scroll down to Geographic and select it. In the Geographic Parameters dialog box select DD for units; right click in the Datum box and select NAD27 from the scroll list. Click on OK to close the Geographic Parameters dialog box, click on OK to close the Input Projection dialog box, and then click OK to execute the action.

The t\_fln\_dd coverage must be re-projected into an Albers projection. This is accomplished by selecting Edit from the Command Tools menu bar; select Coordinates and Project Coverage. This brings up the Project Cover dialog box. Under Input Projection, right click in the Input Cover box to bring up the Select a Coverage dialog box; again, navigate to the directory containing the t\_fln\_dd coverage and select it. Next, left click in the Output Cover box under Output Projection and enter the cover name t\_fln\_alb. Next, click on the Define button to bring up the Output Projection dialog box. Select Continent as the Geographic Extent and select Albers from the scroll list. This brings up the Albers Projection Parameters dialog box. Fill in this dialog box with the values shows in Fig. A.4.1. The Datum is entered by right clicking in the Datum box and selecting NAD83 from the scroll list. Click on OK to close the Albers Projection Parameters dialog box; click on Apply in the Project Cover dialog box; then click Cancel to close the project Cover dialog box.

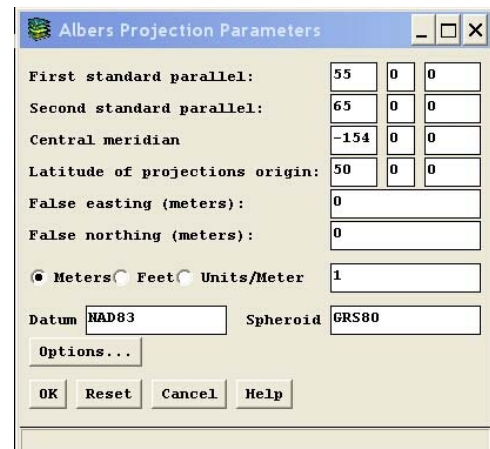


Fig. A.4.1. Albers projection parameters for Alaska.

Next, the t\_fln\_alb coverage must have topology built. This is done from the Command Tools menu bar. Select Edit/Topology/Build features from the menu to bring up the Coverage Topology dialog box. Again, right click in the Coverage box to bring up the Select a Coverage dialog box; select the t\_fln\_alb coverage and click OK. Select Line under the Feature Class scroll list. Ensure that the radio button for Build is selected, then click the Apply button to execute the command. Click the Cancel button to close the Coverage Topology dialog box. The lengths of each transect line has now been calculated in meters.

Table A.4.7. Trans\_Length.txt example file with 4 transect lines created from data derived from Transect\_Output.txt (Table A.4.3).

01/17/06		PAGE	1
T_FLN_ALB-ID	LENGTH		
1	27,502.800		
2	13,280.200		
3	14,857.620		
4	17,479.950		

Table A.4.8. First 6 records from the file Transects\_Flown.csv, a data file containing the lengths of the transects flown during a survey.

Transect,Length (miles)
YU-25,26.32
YU-24,26.31
YU-23,26.25
YU-22,26.26
YU-21,26.18

Next, the transect identifiers and their lengths need to be exported in a format that Excel (or any other spreadsheet) can use. From the Command Tools menu bar select Analysis/Tabular/Report writer to bring up the Create a Report dialog box. In the File to create box type Trans\_Length.txt. Select the radio button for INFO (right of the File to analyze box); then right click in the File to analyze box and select T\_FLN\_ALB.AAT, click OK to close the dialog box. Under the Available Items scroll list select T\_FLN\_ALB-ID and then LENGTH; click OK to create the report (Table A.4.7). Click the Quit button in the popup window displaying a copy of the report.

Next, open the Trans\_Length.txt file in a spreadsheet. Delete the 5 header rows from the file. Next, replace the transect id number with the transect name (use the listing of transect numbers and names created at the beginning of this section). Transect line lengths are listed in meters and need to be converted to miles. Calculate the transect lengths in miles by multiplying the meters length by 0.0006213; limit the miles distance to 2 decimal places. Create the Transects\_Flown.csv file (Table A.4.8) by using the File/Save As function of the spreadsheet program. Note that the first row does contain header information as shown in Table A.4.8. Transects\_Flown.csv is used in both Analysis.SAS (Appendix 5.1) and Bootstrap.SAS (Appendix 5.4).

### Transects\_Flown\_Minus\_Islands

In 2004, Innoko NWR attempted to reduce variance within its highest density strata by conducting empirical counts of moose found on the islands in the Yukon River after completing the normal transects which intersected each island. These islands tend to have large numbers of moose on them and typically have low vegetative cover, thereby making visibility high; these islands tend to be small in size. The probability of obtaining accurate empirical counts was considered to be very high. In order to evaluate density estimates for all the transect lines (including the islands) versus the lines without the islands, the data recorder included a data point to indicate when the transect line reached the edge of the Yukon River. This added data point served as a marker along the transect. This added marker was then used to truncate the transect line from its full length. The Transects\_Flown\_Minus\_Islands.csv file is created following the same steps outlined for the Transects\_Flown.csv file above.

## Bootstrap\_Distance\_Output\_Combined

When program DISTANCE is run following the instructions in Appendix 6.2 to analyze the bootstrap estimates generated by Bootstrap.SAS (Appendix 5.4), program DISTANCE produces 4 output files – one from each run. (Note, users with large datasets will have a larger number of runs due to more files generated by Bootstrap.SAS.) Each of

these program DISTANCE output files (Bootstrap Distance Output 1.txt, Bootstrap Distance Output 2.txt, Bootstrap Distance Output 3.txt, Bootstrap Distance Output 4.txt), all of which are ASCII files, needs to be opened in a spreadsheet program and cleaned so the only data in the file is the replicate number, the model name, and the P value (Table A.4.9). Note that the first row of Example\_Bootstrap\_Distance\_Output\_Combined.csv is a header record. As there are 1,000 bootstrap samples, Bootstrap\_Distance\_Output\_Combined.csv has a total of 1,001 records. This file is used as the input file for Bootstrap\_Distance.SAS (Appendix 5.5) for the final analysis.

Table A.4.9. First 6 records from Example\_Bootstrap\_Distance\_Output\_Combined.csv for use in Bootstrap\_Distance.SAS (Appendix 5.5).

---

Replicate	Model	P
1	Uniform/Cosine	0.59691
2	Uniform/Cosine	0.59868
3	Uniform/Cosine	0.63741
4	Half-normal/Hermite	0.72505
5	Uniform/Cosine	0.65241

---

## Generated Files

Analysis.sas (Appendix 5.1), Histograms.R (Appendix 5.1), Logistic Function.sas (Appendix 5.3), Bootstrap.sas (Appendix 5.4), Merge Bootstrap Results.sas (Appendix 5.5), and DISTANCE (Appendix 6) produce a number of files. Of the files created, most are interim steps in the analytical process. However, users should be aware of 6 files in particular.

- Distance\_Input.csv is a file created by Analysis.sas specifically for use in program DISTANCE. It is important for workers to understand that for distance observations to be properly evaluated by program DISTANCE when estimating a detection function, measured distances need to be shifted to the left by the minimum available sighting distance,  $W_1$  (i.e., measured distance – Zone 0 width). This shift can be found in the SAS code in Appendix 5.1 (or Analysis.sas).
- BS\_moose\_1.csv through BS\_moose\_4.csv are the results from the initial bootstrapping process conducted by Bootstrap.sas. These 4 files serve as input files for program DISTANCE for generating detection functions for the 1,000 bootstrap replicates. Workers with large numbers of moose observations will likely have more than the 4 files identified here (see Bootstrap.SAS for details).
- Bootstrap Distance Output 1.txt through Bootstrap Distance Output 4.txt are the output files created by program DISTANCE during the bootstrap.

## A.4.2 Presenting Results

Once all the data analysis has been completed, results need to be displayed in an orderly fashion.

For past surveys at Innoko NWR, a spreadsheet program has been used to present the data and

make the final calculations for reporting purposes. Beginning with the 2004 survey, the final data presentation table was broken into 2 tables for easier readability (Tables A.4.10 and A.4.11).

Table A.4.10. Example presentation of parameter estimates used in density calculations for each stratum. FPC is the finite population correction factor. From the file Density\_Estimator.xls (Nielson et al. 2004).

Strata	P	g(45)	n <sup>a</sup>	E(s)	Area Searched		Total Area		FPC
					Km <sup>2</sup>	Mi <sup>2</sup>	Km <sup>2</sup>	Mi <sup>2</sup>	
Western	0.645	0.819	77	1.453	490	189	1,634	631	0.70
Innoko	0.645	0.819	22	1.143	251	97	583	225	0.57
Total			99		741	286	2,217	856	0.67

<sup>a</sup> number of moose groups observed

<sup>b</sup> average size of moose groups observed

Table A.4.11. Example presentation of final data analysis results. FPC is the finite population correction factor. From the file Density\_Estimator.xls (Nielson et al. 2004).

Estimates	Strata		Total
	Western	Innoko	
Values without FPC			
Density (#/mi <sup>2</sup> )	1.121	0.491	
SE (Density)	0.246	0.154	
CV (Density)	22%	31%	
Total Moose	707	110	818
SE (Total Moose)	155.0	34.7	160.0
CV (Total Moose)	22%	31%	20%
Values with FPC			
Density (#/mi <sup>2</sup> )	1.121	0.491	
SE (Density)	0.206	0.116	
CV (Density)	18%	24%	
Total Moose	707	110	817
SE (Total Moose)	129.7	26.2	131.0
CV (Total Moose)	18%	24%	16%
Bias			
Bootstrap Mean (Density)	1.143	0.488	
Density (#/mi <sup>2</sup> )	1.121	0.491	
Bias	0.022	-0.003	
90% Confidence Intervals Adjusted with FPC			
Density Lower Limit	0.782	0.300	
Density Upper Limit	1.459	0.682	
Total Moose Lower Limit	494	67	602
Total Moose Upper Limit	920	153	1,032
% Half Width	30%	39%	26%

### A.4.3. Transects Flown AML

Following is the Arc Macro Language (AML) script for importing Transects\_Flown.txt, and creating the t\_fl\_n\_dd and t\_fl\_n\_alb coverages. Before starting ArcInfo, be sure that Transects\_Flown.txt, t-fln.aml, and ddtoalb154densify.prj are all located in the same directory. After starting ArcInfo, but before invoking the t-fln.aml, be sure that the workspace is properly set.

Table A.4.12. Contents of the t-fln.aml.

---

```

/* name - t-fln.aml
/*
/*This aml imports a text file (transects_flown.txt) to create an arc
/*coverage, defines the original projection (Geographic), reprojects the
/*coverage into Albers, then builds line topology. User must initiate an
/*Info session to export line id's and lengths (in meters) to an ascii file,
/*or extract the data using ArcMap.
/*
/*The workspace MUST be set to the directory containing the
/*transects_flown.txt and ddtoalb154densify.prj files, as well as this aml.
/*
/* created by S.D. Kovach, Innoko National Wildlife Refuge, Jan 1999
/*
generate t_fl_n_dd
input transects_flown.txt
lines
q
projectdefine cover t_fl_n_dd
projection geographic
units dd
datum nad27
parameters
project cover t_fl_n_dd t_fl_n_alb ddtoalb154densify.prj
build t_fl_n_alb line
/*End of aml.

```

---

See Table A.4.13 for the contents of the ddtoalb154densify.prj file called by the t-fln.aml.

Following completion of the t-fln.aml you will need to export the transect lengths. This is accomplished by following the instructions in the next to last paragraph in the Transects\_Flown section in Appendix 4.1 above, or by following Script A.4.1 below. Script A.4.1 takes you through an Info session to export line id numbers and their lengths (in meters) to an ASCII file named trans\_lengths.txt. Users will need to modify the path name to fit their situations. Start an info session by entering the command INFO at the arc prompt (Arc: ). User input in the script is in BOLD; all user input must be in upper case. Info prompts without visible user input are created by depressing the Enter key on the keyboard.

Table A.4.13. Contents of the ddtoalb154densify.prj file.

---

```

INPUT
PROJECTION GEOGRAPHIC
UNITS DD
DATUM NAD27 NADCON
DENSIFY .01
PARAMETERS
OUTPUT
PROJECTION ALBER
UNITS METERS
DATUM NAD83 NADCON
PARAMETERS
55 0 0
65 0 0
-154 0 0
50 0 0
0
0
END

```

---

Script A.4.1. Info Session Script for exporting flown transect lengths to an ASCII file.

---

Arc: **INFO**

INFO EXCHANGE CALL

17/01/2006 16:08:13

INFO 9.42 11/11/86 52.74.63\*

Copyright (C) 1994 Doric Computer Systems International Ltd.

All rights reserved.

Proprietary to Doric Computer Systems International Ltd.

US Govt Agencies see usage restrictions in Help files (Help Restrictions)

ENTER USER NAME>**ARC**

ENTER COMMAND >**SEL T\_FLN\_ALB.AAT**

RECORD(S) SELECTED

ENTER COMMAND >**OUTPUT**

**D:\PROJECTS\TEMP\TRANS\_LENGTHS.TXT**

ENTER COMMAND >**REPORT**

2 ENTER COLUMN CONTENT>**T\_FLN\_ALB-ID**

ENTER REPORT OPTIONS>**16**

ENTER COLUMN HEADINGS>**ID**

ENTER COLUMN HEADINGS>

18 ENTER COLUMN CONTENT>**LENGTH**

ENTER REPORT OPTIONS>**16**

ENTER COLUMN HEADINGS>**LENGTH**

ENTER COLUMN HEADINGS>

35 ENTER COLUMN CONTENT>

ENTER REPORT TITLE>'**TRANSECT LENGTHS (M)**'

DO YOU WISH TO EXECUTE THIS FORM ( Y OR N ) >**Y**

OUTPUT TO PRINTER(Y OR N)?>**Y**

LINES PER PAGE?>**9999**

ENTER REPORT OPTIONS>

ENTER COMMAND >**Q STOP**

---

## Appendix 5: SAS and R Code

This appendix contains SAS and R code used for analysis and graphics. Notes and comments are embedded in the code to describe what is happening at each step. The latest version of the R software (R Development Core Team 2005) can be obtained free of charge at <http://www.r-project.org/>. Once at this web page, locate the closest CRAN Mirror to download the latest precompiled binary distribution. This code is not a black box; users must modify this code to fit their specific needs. Notes and comments are embedded in the code to describe what is happening at each step. The Excel (Microsoft Corp., Redmond, Washington, USA) file 'Density Estimator.xls' or a similar spreadsheet program can be used to obtain final estimates of moose abundance based on equation (3) in section 3.4.2.

### A.5.1 SAS Code for Data Analysis

```
/*  
#####  
# Analysis.sas#  
#####
```

SAS Code for analysis of line transect moose survey data. Example data are from the 2004 winter Innoko NWR moose survey (Nielson et al. 2004).

Comments included.

In this example, 2 strata were surveyed, and a separate density estimate is desired for each strata. In addition, a census of moose on Yukon River Islands was taken following surveying the islands using the line transect protocol. This allowed us to reduce variance of total population estimates without reducing our sample size for estimating the detection function (Nielson et al. 2004).

This code was written as a general guide for estimating moose density based on aerial line transect surveys and a double-observer on one side of the aircraft.

See 'Bootstrap.sas' (Appendix 5.4) and 'Merge Bootstrap Results.sas' (Appendix 5.5) for code and notes on how to use bootstrapping to estimate variances.

Ryan Nielson (WEST, Inc.) 2006

```
*/  
  
/*Location of data files*/  
%let datadirectory=C:\Ryan\Projects\innoko_nwr\SOP;  
libname moose "&datadirectory";  
  
/*Read in survey data (see "2004_Survey.csv") in comma delimited format*/  
data observations;  
    length date $9 notes $20;  
    infile "&datadirectory.\Survey_Observations.csv" delimiter=',';
```

```

recl=32767 firstobs=2;
    input Record Transect $ Observer $ Distance Total_moose Cover_10
    Cover_50 Island $;
strata = scan(transect, 1, '-');
group_size = total_moose;
keep Record Transect Observer Distance group_size Cover_10 Cover_50 strata Island;
proc print data=observations; title "Observations"; run; title;

/*Export data as .csv file for plotting histogram of all observations in R. See "Histograms.R"*/
data short;
    set observations (keep=distance);
    if distance ^= .;

proc export data=short outfile="&datadirectory.\example_observed_distances.csv" replace; run;

/*Logistic regression to estimate detection at minimum available sighting distance. In this
example, double observers were set up on the right side of the aircraft.*/
data double_count_observations;
    set observations;
    if distance ^= .;
    if observer = "FR" or observer = "FR & BR";
    if observer = "FR & BR" then BR = 1;
    else BR = 0;
    if distance <= 396; /*396 was determined to be the maximum search distance to
use in this analysis, (should be re-examined for every survey).*/
proc print data=double_count_observations; var BR distance group_size;
title "Double Count Observations"; run; title;

proc logistic data=double_count_observations descending;
    model BR = distance; run;

proc logistic data=double_count_observations descending;
    model BR = distance group_size; run;

/*Calculate total search area*/
data transects;
    infile "&datadirectory.\Transects_Flown.csv"
    delimiter=',' lrecl=32767 firstobs=2;
    input transect $ length;
    strata = scan(transect, 1, '-');
    area = length*2*0.2181013; /*sq miles: 0.2181023 is
396m-45m transformed into miles*/

proc sort data=transects;
    by strata;

```



```

proc means data=transects n sum;
    by strata;
    var area; title "Number of Transects and Total Area Searched (sq-miles)";

/*Caclulate total search area after removing islands were we obtained empirical counts*/
data transects_no_islands;
    infile "&datadirectory.\Transects_Flown_Minus_Islands.csv"
        delimiter=',' lrecl=32767 firstobs=2;
    input transect $ length;
    strata = scan(transect, 1, '-');
    area = length*2*0.2181013; *sq miles;

proc sort data=transects_no_islands;
    by strata;

proc means data=transects_no_islands n sum;
    by strata;
    var area;
    title "Number of Transects and Total Area Searched (minus islands censused)"; run;

/*Caclulate number of groups observed in each strata by back seat observers*/
data back_seat;
    set observations;
    if distance ^= .;
    if observer = "FR & BR" or observer = "BR" or observer = "BL";
    if distance le 396;
    strata = scan(transect, 1, '-');

proc sort data=back_seat;
    by strata;

proc means data=back_seat n;
    by strata;
    var group_size;
    title "Number of Groups Observed by Rear Seat Observers"; run; title;

/*Caclulate number of groups observed in each strata by back seat observers with island
observations removed*/
data back_seat;
    set observations;
    if distance ^= .;
    if island = "N";
    if observer = "FR & BR" or observer = "BR" or observer = "BL";
    if distance le 396;
    if distance ^= .;
    if group_size ^= .;

```

```

        strata = scan(transect, 1, '-');

proc sort data=back_seat;
    by strata;

proc means data=back_seat n;
    by strata;
    var group_size;
    title "Number of Groups Observed by Rear Seat Observers with Islands Removed";
run; title;

/*Exporting data for program DISTANCE. Data contains observations from the rear seats only*/
data short;
    set observations;
        if group_size ^= .;
        if distance ^= .;
        if observer ^= "FR";
        if observer ^= "NA";
        if distance le 396;
        /*Must subtract minimum available sighting distance to obtain correct
        estimates of probability of detection from DISTANCE program*/
        distance = distance - 45;
    keep Record distance transect;

proc freq data=short;
    table distance; run;

proc export data=short outfile="&datadirectory\Distance_Input.csv" replace; run;

/*Estimating average group size for using observations within 200m*/
data short;
    set observations;
        if observer = "NA" then delete;
        if distance <= 200;

proc means data=short mean n stderr;
    var group_size; title "Average Group Size"; run;

/*Estimating average group size for each stratum*/
proc sort data=short;
    by strata;

proc means data=short mean n stderr;
    by strata;
    var group_size; title "Average Group Size by Strata";

```

/\*The remainder of the analysis is done in DISTANCE, with the exception of bootstrapping for variance and bias. See Bootstrap.SAS and Bootstrap\_Distance.SAS\*/

## A.5.2 R Code for Histograms

```
#####  
#Histograms.R#  
#####  
#  
#Histogram function available in R.  
#  
#Use 'help(hist)' for help files.  
#  
#Realize that R is case sensitive.  
#  
#Ryan Nielson (WEST, Inc.) 2006  
  
#Read in data containing distances of moose observations.  
distances<- read.csv("Observed_Distances.csv", header=T)  
  
#Call function to create histogram of moose observations.  
#Example 1  
break.points = seq(45, max(distances$Distance)+75, by=75)  
hist(distances$Distance, xlim=c(min(break.points), max(break.points)),  
      breaks=break.points, xlab="Distance (meters)",  
      main=" ", xaxt='n', col="steelblue")  
axis(1, at=break.points, labels=T, tick=T)  
  
#Example 2  
break.points = seq(45, max(distances$Distance)+65, by=65)  
hist(distances$Distance, xlim=c(min(break.points), max(break.points)),  
      breaks=break.points, xlab="Distance (meters)",  
      main=" ", xaxt='n', col="steelblue")  
axis(1, at=break.points, labels=T, tick=T)  
  
#Use the mouse to move the pointer select the Graphics Device window and then  
#select 'File->Save As' to save the image as a jpeg.
```

## A.5.3 SAS Code for Logistic Function

```
/*  
#####  
# Logistic Function.sas#  
#####
```

Plot probability of detection by the rear-seat observer, given the moose group was observed by the front-seat observer, and calculate 95% confidence interval for the probability of detection at

the minimum available sighting distance. This is the logistic regression function estimated from the double-observer data. Visual inspection of this function should provide some insight into the relationship between probability of detection and distance from transect line.

```
*/
```

```
/*In this example, the "best" model had 'distance' as the only predictor variable. Look at probability of detection by back seat observer, given moose group was seen by front seat observer*/
```

```
data detection_at_zero;  
    * X to Y by 1: X and Y are minimum available sighting distances for the survey;  
    do distance = 45 to 396 by 1;  
        *parameters below estimated from logistic model fit—modify as necessary;  
        probability = exp(1.557 - 0.001*distance)/(1+exp(1.557 - 0.001*distance));  
        output;  
    end;
```

```
proc gplot data=detection_at_zero;  
    plot probability*distance; run; quit;
```

```
/*Use the mouse to move the pointer and select 'File->Save As' to save the image as a jpeg.
```

```
/*Prediction and 95% CI at minimum available sighting distance*/
```

```
data pred;  
    distance = 45; *change, if necessary;
```

```
data pred;  
    set double_count_observations pred;  
    keep BR distance;
```

```
proc logistic data=pred descending;  
    model BR = distance;  
    output out=preds predicted=pred lower=lower_CI upper=upper_CI;  
run;
```

```
data preds;  
    set preds;  
    if BR = . and distance = 45; *change, if necessary;
```

```
proc print data=preds; title "Predicted Probability of Detection at 'G(0)"; run; title;
```

#### **A.5.4 SAS Code for Bootstrap**

```
/*
```

```
#####
```

```
#Bootstrap.sas#
```

```
#####
```

SAS Code for bootstrapping line transect moose survey data. Example data are from the 2004 winter Innoko NWR moose survey (Nielson et al. 2004).

Comments are included for explanation.

This code bootstraps individual transects within each stratum, estimates number of groups observed and average group size for each stratum, and then outputs each bootstrap sample for analysis in the program DISTANCE. In addition, the probability of detection at the minimum available sighting distance is estimated based on each bootstrap sample.

Using observations on islands for estimation of  $P$ ,  $G(0)$ , and  $E(S)$ , and creating one estimate of density using all transect survey data for calculation of number of groups seen and total area searched, and other estimate based on a combination of line transect sampling and empirical island counts to increase precision.

Ryan Nielson (WEST, Inc.) 2006

```
*/
```

```
/*Location of data files*/
```

```
%let datadirectory=C:\Ryan\Projects\innoko_nwr\SOP;  
libname moose "&datadirectory";
```

```
/*Maximum sighting distance used in analysis*/
```

```
%let max_distance = 396;
```

```
/*Minimum available sighting distance used in analysis*/
```

```
%let min_distance = 45;
```

```
/*Number of bootstrap samples*/
```

```
%let reps = 1000;
```

```
/*Macro for generating bootstrap samples and exporting sample data for analysis in  
program DISTANCE*/
```

```
options mprint symbolgen;
```

```
%macro bootstrap();
```

```
    /*Import list and lengths of transects flown*/
```

```
    data yukon_transects innoko_transects;
```

```
        infile "&datadirectory\Transects_Flown.csv" delimiter=',';
```

```
        lrecl=32767 firstobs=2;
```

```
        input transect $ length;
```

```
        strata = scan(transect, 1, '-');
```

```
        if strata = "YU" then output yukon_transects;
```

```
        if strata = "IN" then output innoko_transects;
```

```

/*Take a random sample of transects, with replacement, from each stratum.
Use the same sample size as the original sample*/
proc surveyselect data=yukon_transects method=urs sampsize=25 rep=&reps
    out=yukon_samples;

proc surveyselect data=innoko_transects method=urs sampsize=24 rep=&reps
    out=innoko_samples;

/*Save bootstrap samples as permanent SAS datasets*/
data moose.yukon_samples;
    set yukon_samples; run;

data moose.innoko_samples;
    set innoko_samples;

/*Combine samples from the 2 strata*/
data samples;
    set yukon_samples innoko_samples;

/*Calculate total area searched in each stratum for each sample*/
data transects_w_islands;
    infile "&datadirectory\Transects_Flown.csv" delimiter=','
        lrecl=32767 firstobs=2;
    input transect $ length;
    strata = scan(transect, 1, '-');
    keep strata transect length;

data transects_no_islands;
    infile "&datadirectory\Transects_Flown_Minus_Islands.csv" delimiter=','
        lrecl=32767 firstobs=2;
    input transect $ length;
    strata = scan(transect, 1, '-');
    length_no_islands = length;
    keep strata transect length_no_islands;

proc sort data=transects_w_islands;
    by strata transect;

proc sort data=transects_no_islands;
    by strata transect;

proc sort data=samples;
    by strata transect;

data transects;
    merge samples (in=a) transects_w_islands (in=b) transects_no_islands (in=c);

```

```

        by strata transect;
        if a and b and c;
        do i = 1 to NumberHits;
        output; end;

proc sort data=transects;
    by replicate strata;

proc means data=transects noprint;
    by replicate strata;
    var length length_no_islands;
    output out=area sum=total_length total_length_no_islands;

data area;
    set area;
        searched = (total_length)*(&max_distance-
            &min_distance)*0.0006213712*2;
        *converting to square-miles;
        searched_no_islands = (total_length_no_islands)*(&max_distance-
            &min_distance)*0.0006213712*2;
        *islands removed;

proc transpose data=area out=area_searched;
    by replicate;
    var searched;
    id strata;

proc transpose data=area out=area_searched_no_islands;
    by replicate;
    var searched_no_islands;
    id strata;

data area_searched;
    set area_searched;
        yu_area = YU;
        in_area = IN;
        keep yu_area in_area replicate;

data area_searched_no_islands;
    set area_searched_no_islands;
        yu_area_no_islands = YU;
        keep yu_area_no_islands replicate;

proc sort data=area_searched;
    by replicate;

```

```

proc sort data=area_searched_no_islands;
    by replicate;

data area_searched;
    merge area_searched area_searched_no_islands;
    by replicate;

/*Assigning observations to the sampled transects*/
data obs;
    length date $9 notes $20;
    infile "&datadirectory.\Survey_Observations.csv" delimiter=','
        lrecl=32767 firstobs=2;
    input Record Transect $ Observer $
        Distance Total_moose Cover_10 Cover_50 Island $;
    strata = scan(transect, 1, '-');
    group_size = total_moose;
    keep transect observer distance group_size cover_10 cover_50 strata island;

proc sort data=obs;
    by transect;

%do i=1 %to &reps;

    data rep_&i;
        set samples;
        if replicate=&i;

/*For each transect, get the observations*/
proc sort data=rep_&i;
    by transect;

    data rep_&i;
        merge obs (in=a) rep_&i (in=b);
        by transect;
        if a and b;
        do i=1 to NumberHits;
            output;
        end;

proc append base=moose.bootstrap_observations data=rep_&i force; run;

%end;

/*Calculate average group sizes for each bootstrap sample*/
data group_sizes;
    set moose.bootstrap_observations;

```



```

        if observer ^= "NA";
        if distance ^= .;
        if group_size ^= .;
        if distance le 200;
        strata = scan(transect, 1, '-');

proc sort data=group_sizes;
    by replicate strata;

proc means data=group_sizes noprint;
    by replicate strata;
    var group_size;
    output out=group_sizes mean=group_s;

proc transpose data=group_sizes out=group_sizes;
    by replicate;
    var group_s;
    id strata;

data group_sizes;
    set group_sizes;
        Yukon_group_size = YU;
        Innoko_group_size = IN;
    keep Yukon_group_size Innoko_group_size replicate;

/*Calculate number of groups observed by back seat observers for each bootstrap
sample*/
data numbers_with_islands;
    set moose.bootstrap_observations;
        if observer ^= "NA";
        if observer ^= "FR";
        if distance ^= .;
        if group_size ^= .;
        if distance le &max_distance;
        strata = scan(transect, 1, '-');

proc sort data=numbers_with_islands;
    by replicate strata;

proc means data=numbers_with_islands noprint;
    by replicate strata;
    var group_size;
    output out=numbers n=n_groups;

proc transpose data=numbers out=numbers;
    by replicate;

```

```

var n_groups;
id strata;

data numbers_with_islands;
set numbers;
    Yukon_number_groups = YU;
    Innoko_number_groups = IN;
keep Yukon_number_groups Innoko_number_groups replicate;

data numbers_without_islands;
set moose.bootstrap_observations;
    if island = "N";
    if observer ^= "NA";
    if observer ^= "FR";
    if distance ^= .;
    if group_size ^= .;
    if distance le &max_distance;
    strata = scan(transect, 1, '-');
    if strata = "YU";

proc sort data=numbers_without_islands;
    by replicate strata;

proc means data=numbers_without_islands noprint;
    by replicate strata;
    var group_size;
    output out=numbers n=n_groups;

proc transpose data=numbers out=numbers;
    by replicate;
    var n_groups;
    id strata;

data numbers_without_islands;
set numbers;
    Yukon_number_groups_wo_islands = YU;
keep Yukon_number_groups_wo_islands replicate;

/*Using logistic regression and the double-observer data to estimate detection at the
minimum available sighting distance for each bootstrap replicate*/
data dbl;
set moose.bootstrap_observations;
BR = 0;
if distance le &max_distance;
if observer = "FR & BR" then BR = 1;
if observer ^= "BL";

```

```

        if observer ^= "BR";
        if observer ^= "NA";

proc sort data=dbl;
    by replicate;

/*Proportion of moose seen by back-right observer, this will be used if proc logistic
does not converge*/
proc means data=dbl noprint;
    by replicate;
    var BR;
    output out=br_moose mean=p_seen;

/*Estimate Detection function for back-right observer*/
proc logistic data=dbl descending;
    ods output ParameterEstimates=parms ConvergenceStatus=Convergence;
    by replicate;
    model BR = distance;

proc sort data=parms;
    by replicate;

proc sort data=convergence;
    by replicate;

data ok_parms;
    merge parms convergence;
    by replicate;
    if Reason="Complete separation of data points detected." then delete;
    if Reason="Quasicomplete separation of data points detected." then delete;

proc transpose data=ok_parms out=logistic;
    by replicate;
    id variable;
    var estimate;

data logistic;
    set logistic;
        g_0 = exp(intercept + distance*&min_distance)/(1+exp(intercept
            + distance*&min_distance));
        keep replicate g_0; run;

data on_the_line;
    merge logistic br_moose;
    by replicate;
    if g_0=. then g_0=p_seen;

```

```

        if g_0=. then delete;

/*Merging estimates of group size, number of groups, area searched, and g(0)*/
data moose.bs_parms;
    merge on_the_line numbers_with_islands numbers_without_islands group_sizes
        area_searched;
    by replicate; proc print; run;

/*Export data for program DISTANCE*/
/*Must export data and run in batches of 250, otherwise DISTANCE will stall or*/
/*crash. For extremely large datasets, may need to reduce batch size to 100 bootstrap*/
/*replicates.*/
data short;
    set moose.bootstrap_observations;
    if distance le &max_distance;
    distance = distance - &min_distance;
    if observer ^= "NA";
    if observer ^= "FR";
    if distance ^= .;
    if group_size ^= .;
    keep transect replicate distance;

proc sort data=short;
    by replicate transect;

data temp;
    set short;
    if replicate le 250;
proc export data=temp outfile="&datadirectory.\BS_moose_1.csv" replace;
run;

data temp;
    set short;
    if 250 < replicate <= 500;
proc export data=temp outfile="&datadirectory.\BS_moose_2.csv" replace;
run;

data temp;
    set short;
    if 500 < replicate <= 750;
proc export data=temp outfile="&datadirectory.\BS_moose_3.csv" replace;
run;

data temp;
    set short;
    if 750 < replicate <= 1000;

```

```

proc export data=temp outfile="&datadirectory.\BS_moose_4.csv" replace;
    run;

proc datasets; delete temp; run; quit;

run;
%mend;
/*Macro call*/
%bootstrap;

/*Check means of bootstrap estimated parameters*/
proc means data=moose.bs_parms mean;
    var Yukon_number_groups Yukon_number_groups_wo_islands Innoko_number_groups
        Yukon_group_size Innoko_group_size yu_area yu_area_no_islands in_area;
run;

```

### **A.5.5 SAS Code for Merging Bootstrap Information and Calculating Standard Errors, Bias, and Confidence Intervals**

```

/*
#####
#Merge Bootstrap Results.sas#
#####
SAS Code for merging bootstrap information from DISTANCE and SAS. Example data taken
from 2004 winter Innoko NWR moose survey data.

```

Comments are included for explanation.

This code imports bootstrap output from DISTANCE and merges this information with the permanent SAS data files created by ‘SAS Code for Bootstrap.SAS’ in Appendix 5.3. Estimates of variance (standard errors) and bias are calculated, along with 90% confidence intervals for parameters of interest.

Ryan Nielson (WEST, Inc.) 2006

```

*/
/*Location of data files*/
%let datadirectory=C:\Ryan\Projects\innoko_nwr\SOP;
libname moose "&datadirectory";

/*Import Results from program DISTANCE. Output results were cleaned up and merged into
one .csv file (see 'Distance_Bootstrap_Output_Combined.csv')*/
data distance;
    length model $20;
    infile "&datadirectory\Bootstrap_Distance_Output_Combined.csv"
        delimiter=' ' lrecl=32767 firstobs=2;
    input replicate model $ p;

```

```

        if p = 1 then delete; /*Had error messages associated with 2 bootstrap samples*/
proc print; run;

/*How often each model was selected as "BEST"*/
proc freq data=distance;
    table model; run;

/*Merging bootstrap information*/
proc sort data=moose.bs_parms;
    by replicate;

proc sort data=distance;
    by replicate;

data parms;
    merge moose.bs_parms (in=a) distance (in=b);
    by replicate;
    if a and b then output parms; run;

/*Calculating mean, variance and bias*/
data parms;
    set parms;
    Yukon_density = (Yukon_number_groups*Yukon_group_size)/(g_0*p*yu_area);
    Yukon_density_no_islands =
        (Yukon_number_groups_wo_islands*Yukon_group_size)
        /(g_0*p*yu_area_no_islands);
    Yukon_total = Yukon_density*631.277; /*Area in Yukon Stratum including
    Islands*/
    Yukon_total_no_islands = Yukon_density_no_islands*608;
        /*Area in Yukon Stratum not including islands*/
    Yukon_total_no_islands = Yukon_total_no_islands + 88; /*88 moose seen on
    Islands*/
    Log_Yu_BS = log10(Yukon_total);
    Log_Yu_BS_no_islands = log10(Yukon_total_no_islands);
    Innoko_density = (Innoko_number_groups*Innoko_group_size)/(g_0*p*in_area);
    Innoko_total = Innoko_density*225; /*Area in Innoko R. Stratum;
    Log_In_BS = log10(Innoko_total);
    total = Yukon_total + Innoko_total;
    total_no_islands = Yukon_total_no_islands + Innoko_total;

proc means data=parms noprint;
    var Yukon_density Yukon_density_no_islands Innoko_density Log_Yu_BS
        Log_Yu_BS_no_islands Log_In_BS Yukon_total Yukon_total_no_islands
        Innoko_total total total_no_islands;
    output out=est mean=BS_Mean_Yu_density BS_Mean_Yu_density_no_islands
        BS_Mean_In_density BS_Mean_logN_Yu S_Mean_logN_Yu_no_islands

```

```

        BS_Mean_logN_In BS_Mean_Yu_total BS_Mean_Yu_total_no_islands
        BS_Mean_In_total BS_mean_total BS_mean_total_no_islands
stddev=SE_Yu_density SE_Yu_density_no_islands SE_In_density SE_logN_Yu
        SE_logN_Yu_no_islands SE_logN_In SE_Yu_total
        SE_Yu_total_no_islands SE_In_total SE_total SE_total_no_islands;
proc print data=est; title "Bootstrap Results"; run; title;

```

```

/*Standard 90% Confidence Intervals*/

```

```

data est;

```

```

    set est;

```

```

        Log_Yu_Total = log10(707); /*707 is estimated total number of moose for
        Yukon*/

```

```

        Log_Yu_Total_no_islands = log10(601); /*601 is estimated total number of
        moose for Yukon minus islands*/

```

```

        Log_In_Total = log10(110); /*110 is estimated total number of moose for
        Innoko*/

```

```

        LL_Yu_Total = 10**(Log_Yu_Total - 1.64*SE_logN_Yu);

```

```

        UL_Yu_Total = 10**(Log_Yu_Total + 1.64*SE_logN_Yu);

```

```

        LL_Yu_Total_no_islands = 10**(Log_Yu_Total_no_islands -
        1.64*SE_logN_Yu_no_islands);

```

```

        UL_Yu_Total_no_islands = 10**(Log_Yu_Total_no_islands +
        1.64*SE_logN_Yu_no_islands);

```

```

        LL_In_Total = 10**(Log_In_Total - 1.64*SE_logN_In);

```

```

        UL_In_Total = 10**(Log_In_Total + 1.64*SE_logN_In);

```

```

proc print data=est; title "Standard 90% CIs"; run; title;

```

```

/*Bootstrap percentile confidence intervals (90% CI)*/

```

```

proc univariate data=parms;

```

```

    var Yukon_total; /*Innoko_total Yukon_total Yukon_total_no_islands */

```

```

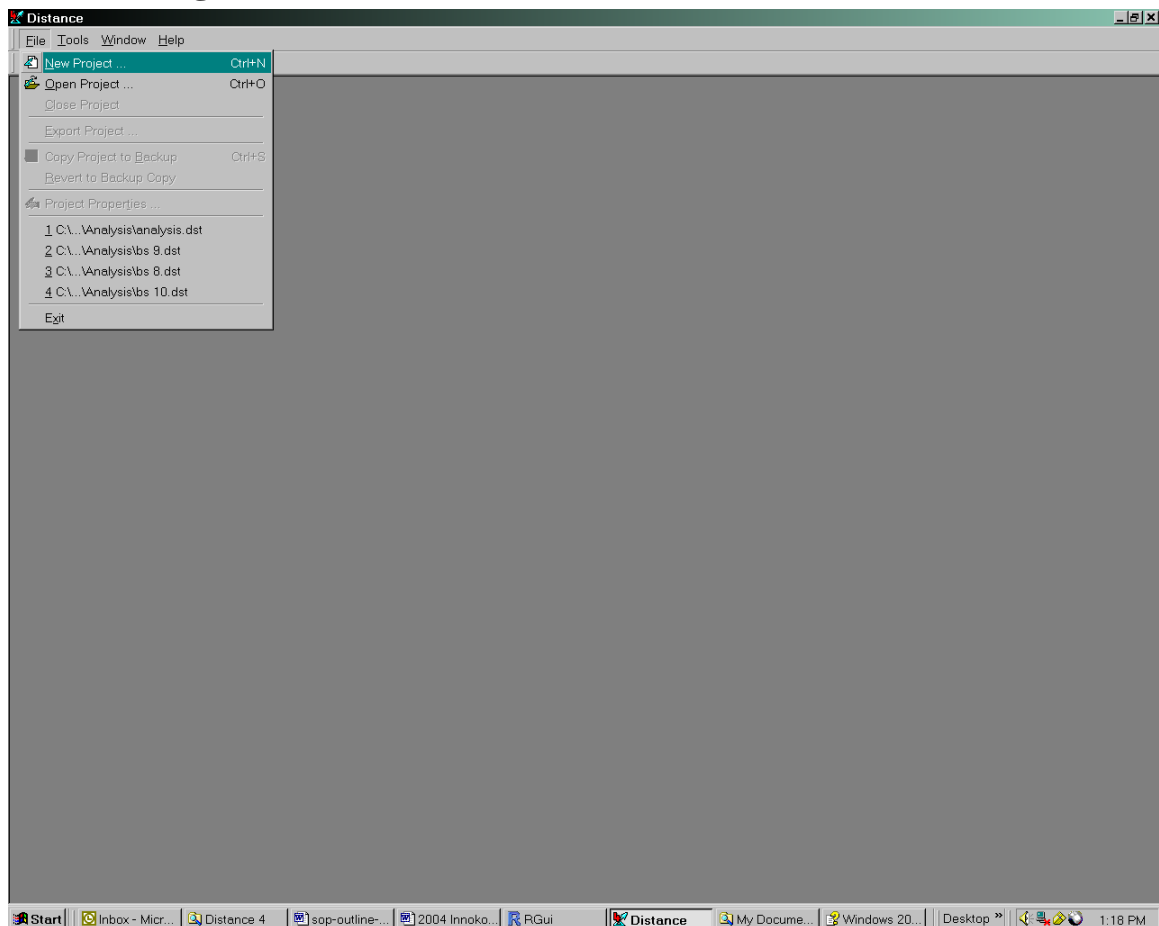
    output pctlpre=P pctlpts=5 95; proc print; title "Percentile CIs"; run

```

## Appendix 6: Using the Program DISTANCE

This appendix contains step-by-step instructions for running the program DISTANCE (Thomas et al. 2004), which can be obtained free of charge at <http://www.ruwpa.st-and.ac.uk/distance/>, after registering as a new user (at the same web address). More detailed instructions and explanation of different software components and uses can be found in the Distance User's Guide On-line manual, which is available under "Help" when you open the DISTANCE program.

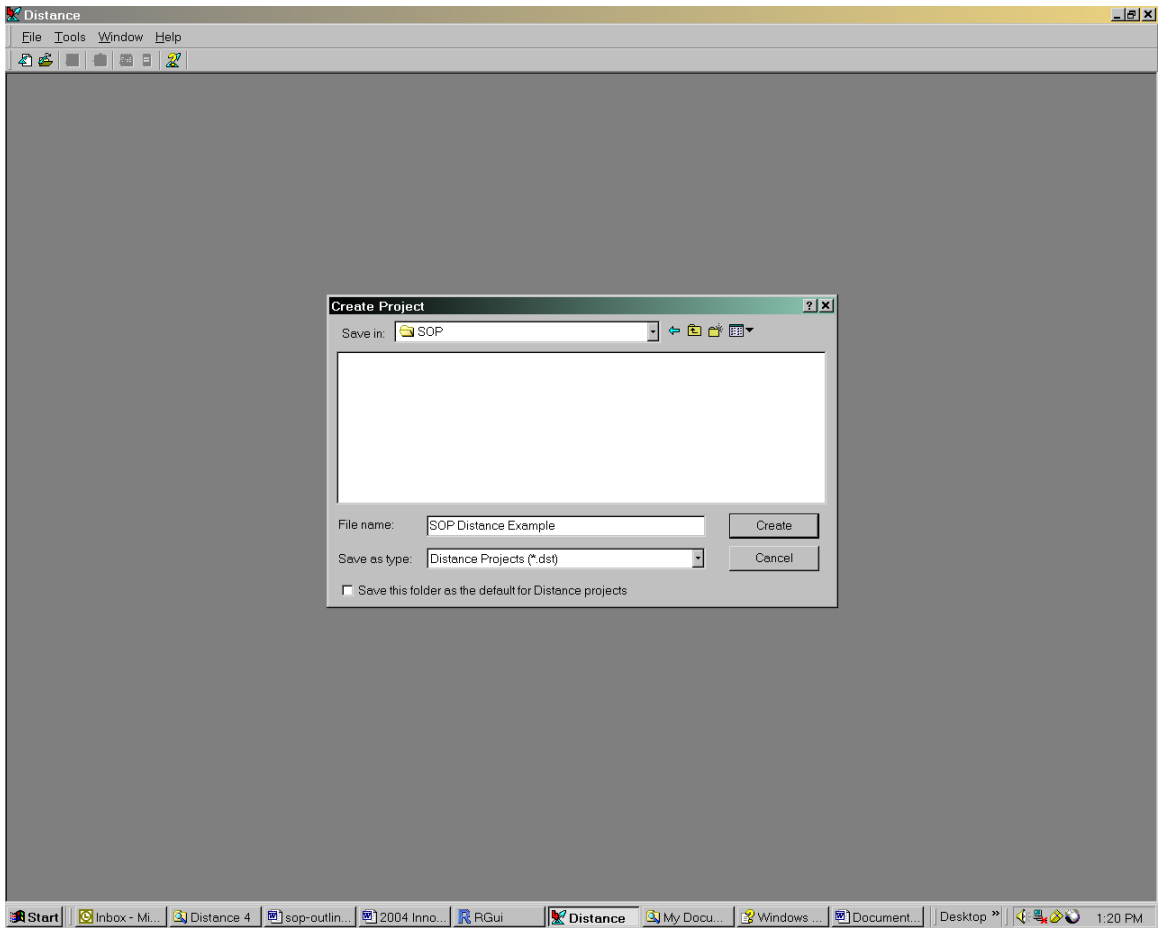
### A.6.1 Using DISTANCE to Fit a Detection Function



#### Step 1.

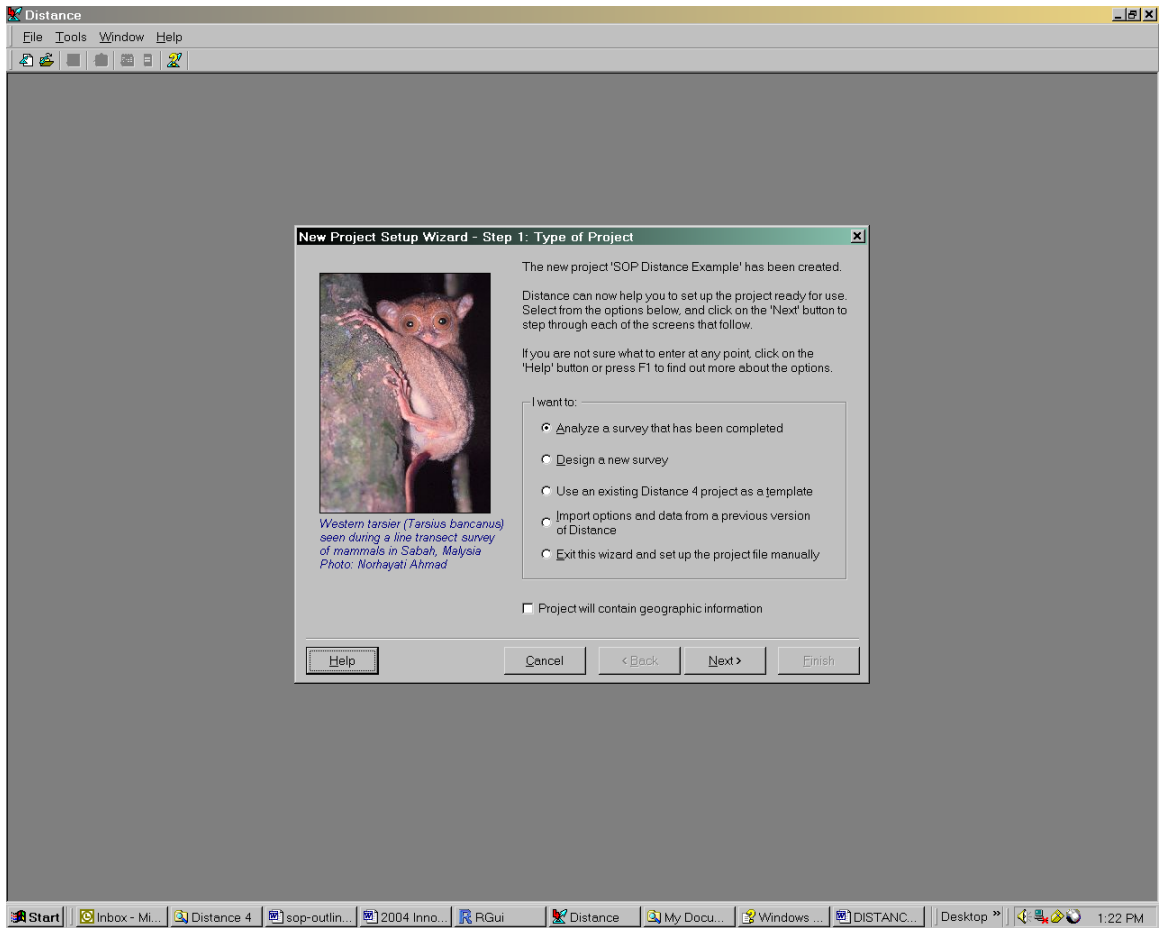
Start DISTANCE and open a new project.





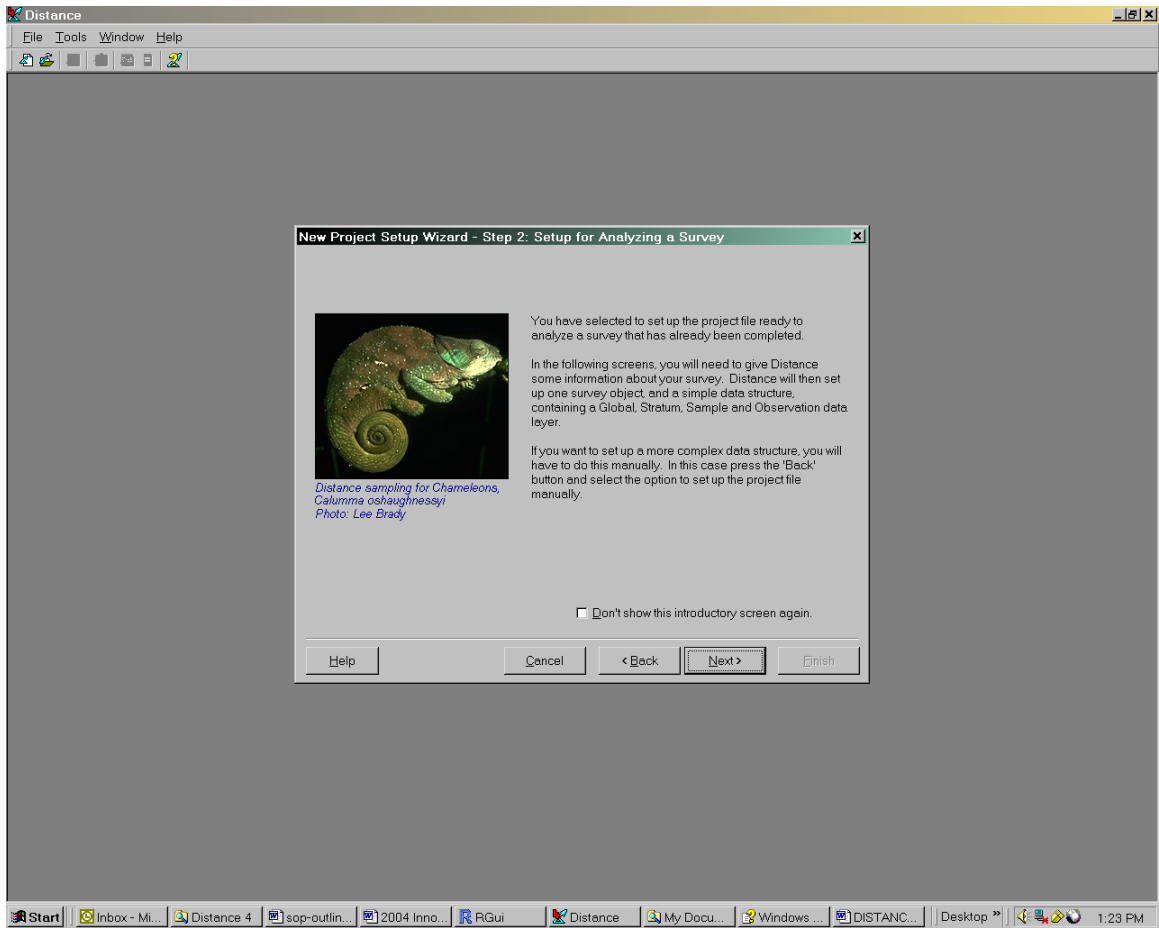
## Step 2.

You will then be asked to direct DISTANCE to the folder location where you want to save the new project. You will also be prompted to name the project. In the above example you can see the new project named “SOP Distance Example” is being saved in the folder titled “SOP”.



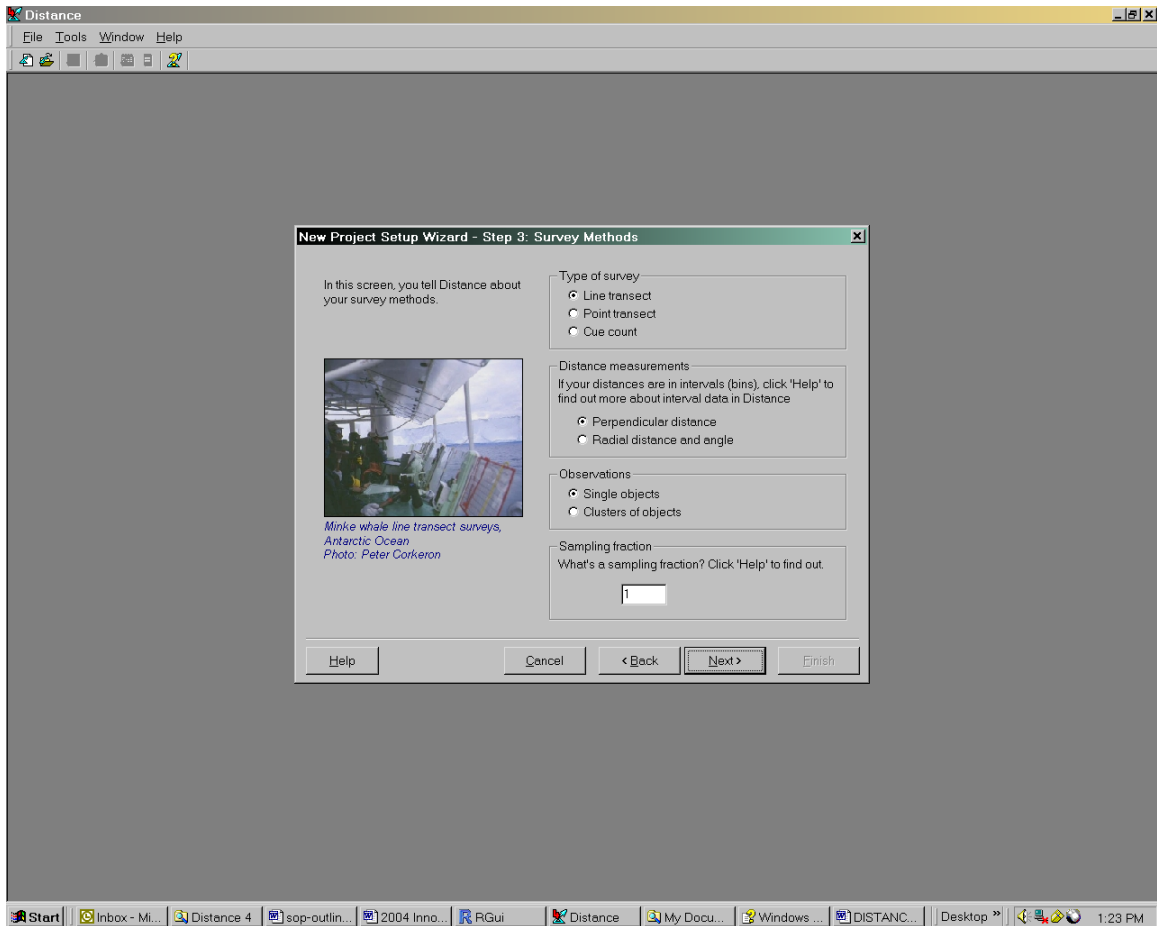
### Step 3.

In the Project Setup Wizzard select “Analyze a survey that has already been completed” and click the NEXT button.



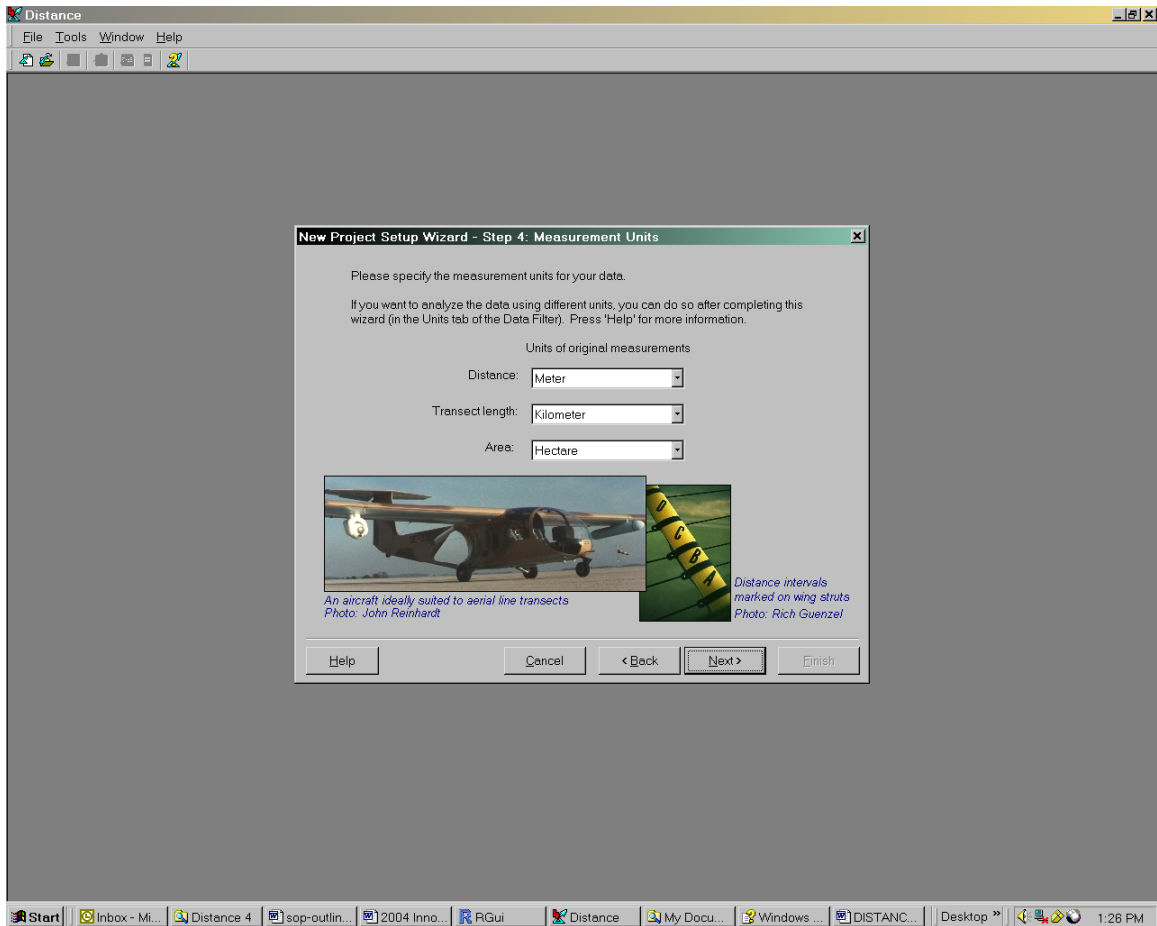
#### Step 4.

Click the NEXT button when this screen appears.



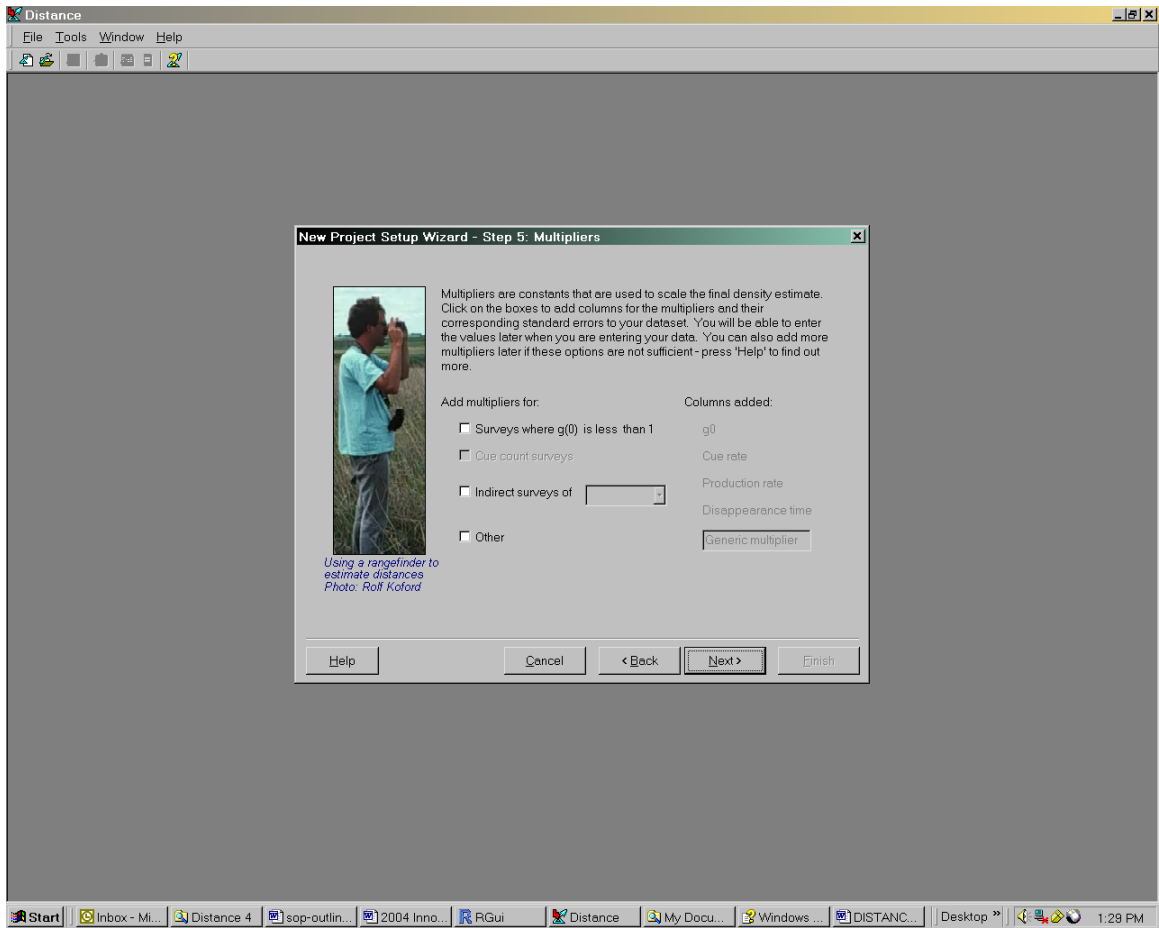
### Step 5.

Select “Line transect” as the type of survey with perpendicular distance measurements. Since we estimate group size and density of animals outside of DISTANCE (in SAS and Excel), just specify that the observations are “Single objects”. Click on the NEXT button after you make your selections.



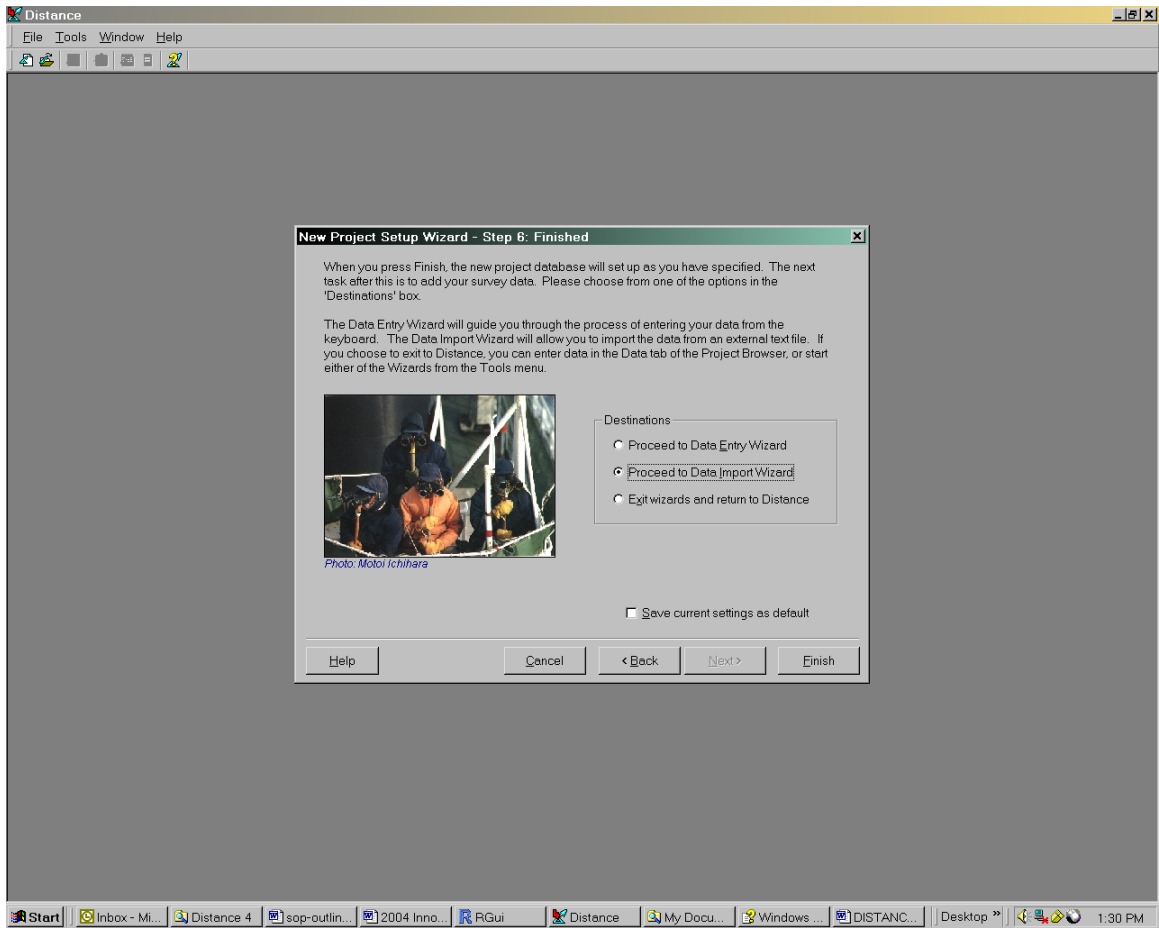
### Step 6.

Specify the proper distance units and click on the NEXT button. In this example we are using meters for measuring perpendicular distances, but because we are estimating density outside of DISTANCE, the “Transect length” and “Area” units are irrelevant.



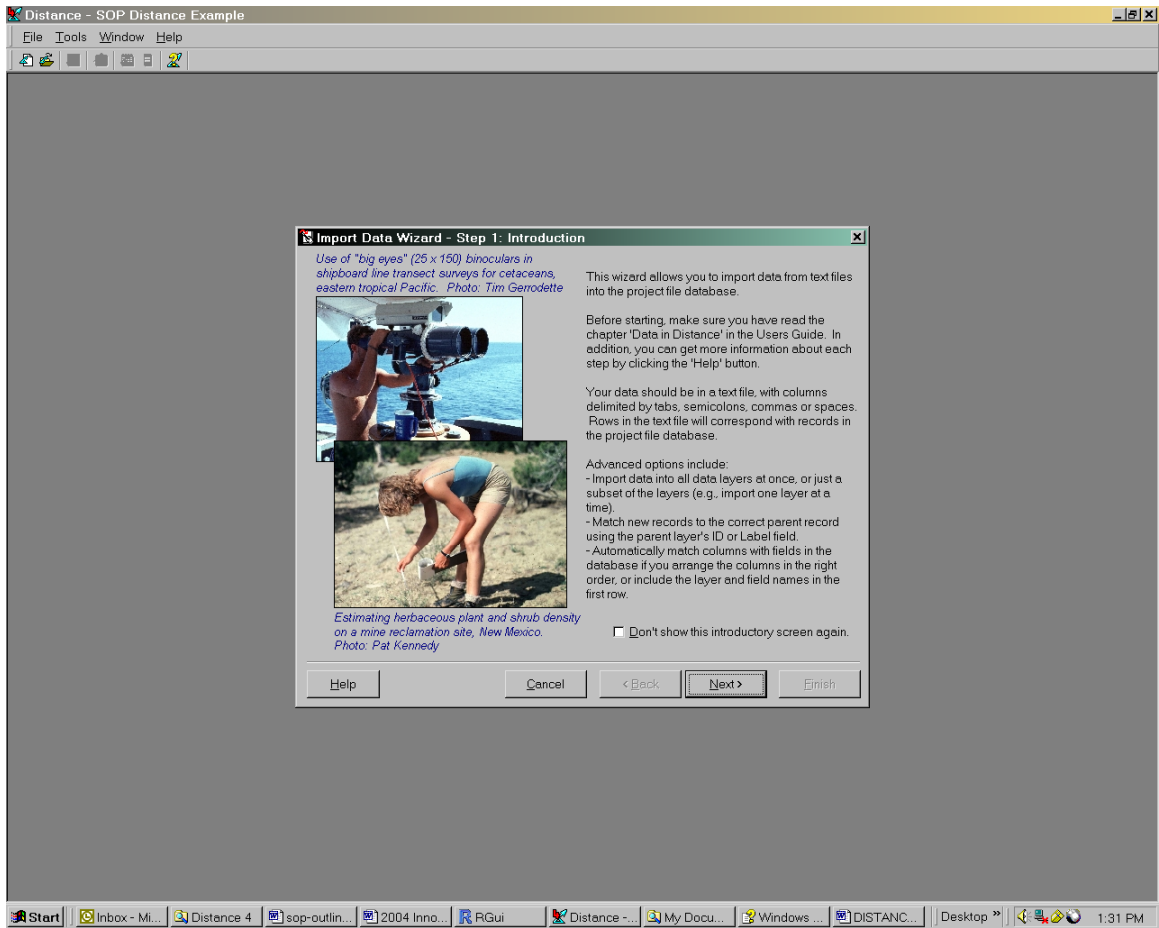
### Step 7.

We estimate density outside of DISTANCE, so we ignore this screen and click on the NEXT button.



### Step 8.

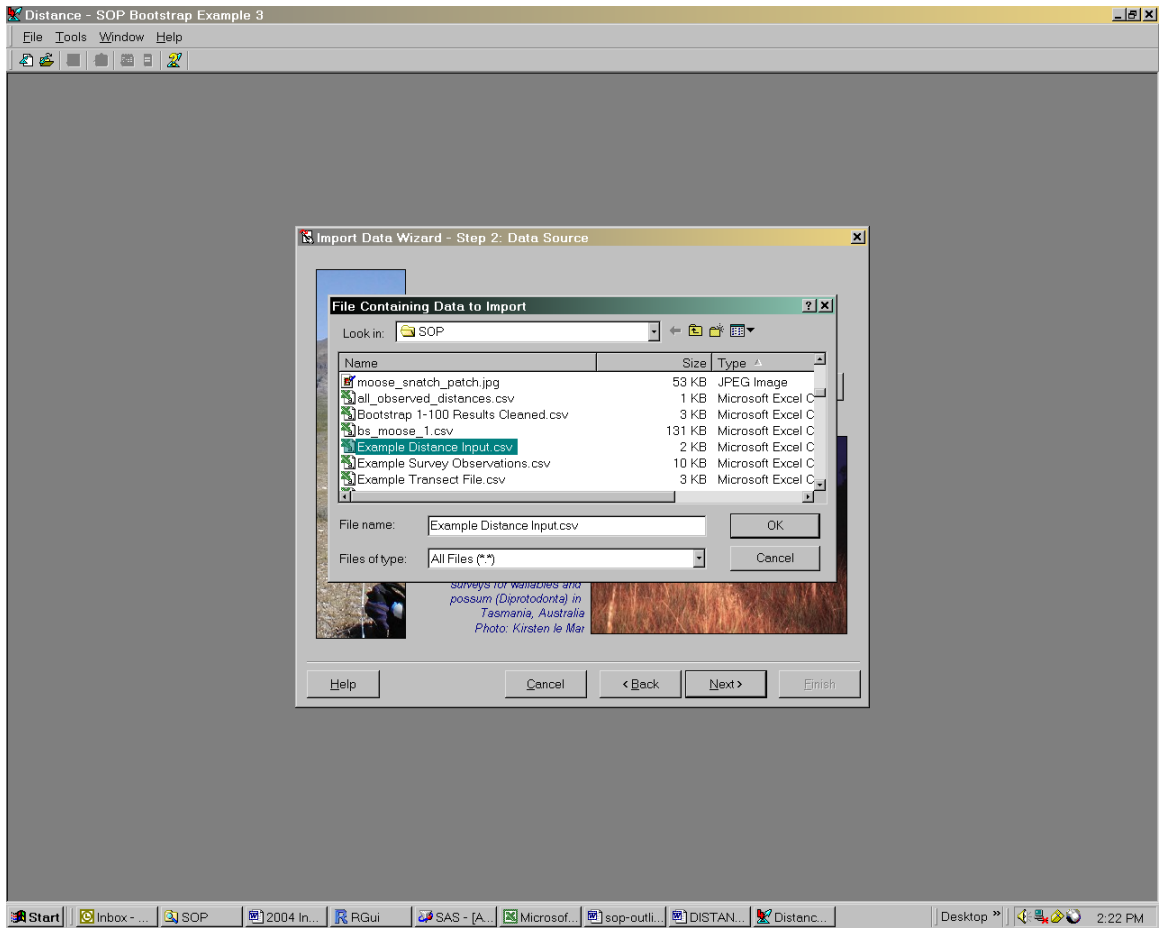
Select to “Proceed to Data Import Wizard” and click on the FINISH button.



### Step 9.

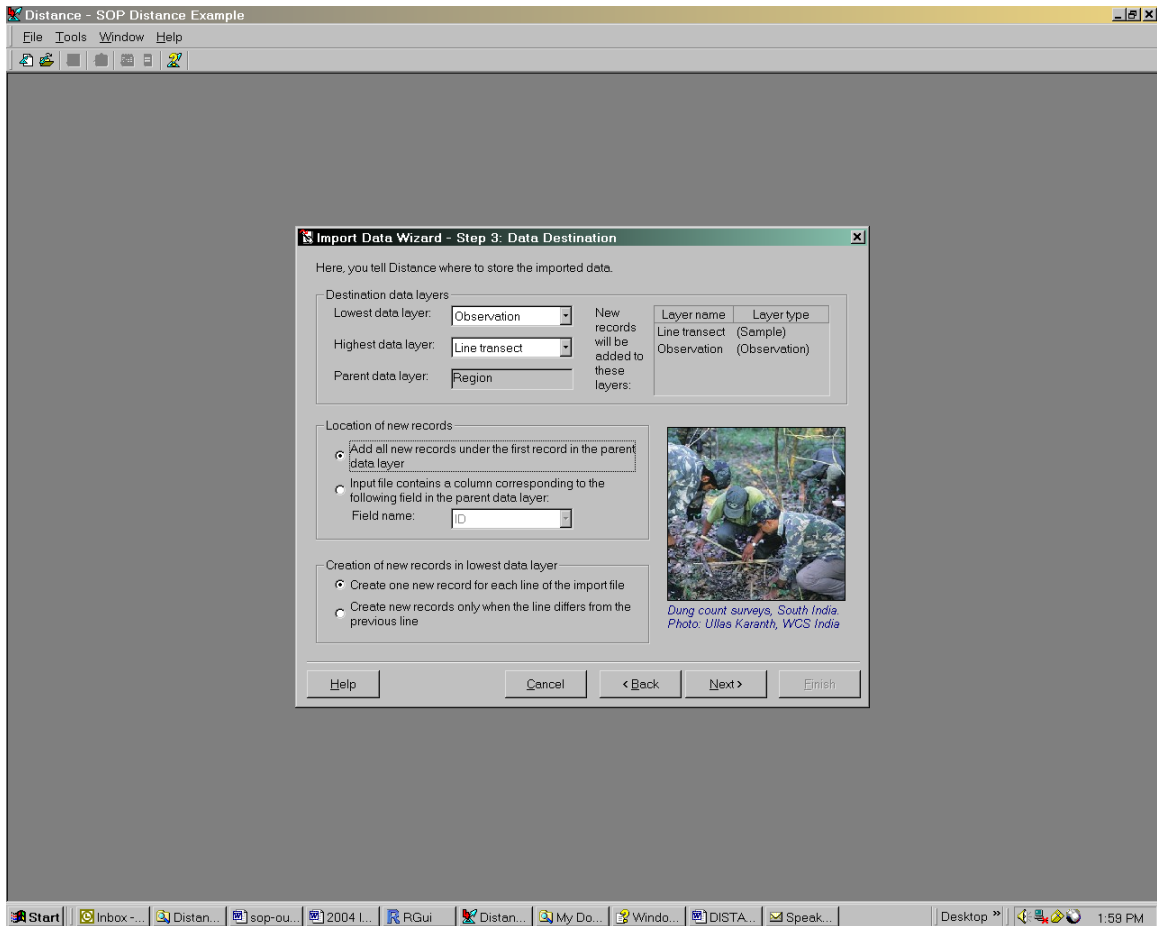
This screen offers some general advice on importing data. Click on the NEXT button to proceed with the Data Import Wizard.





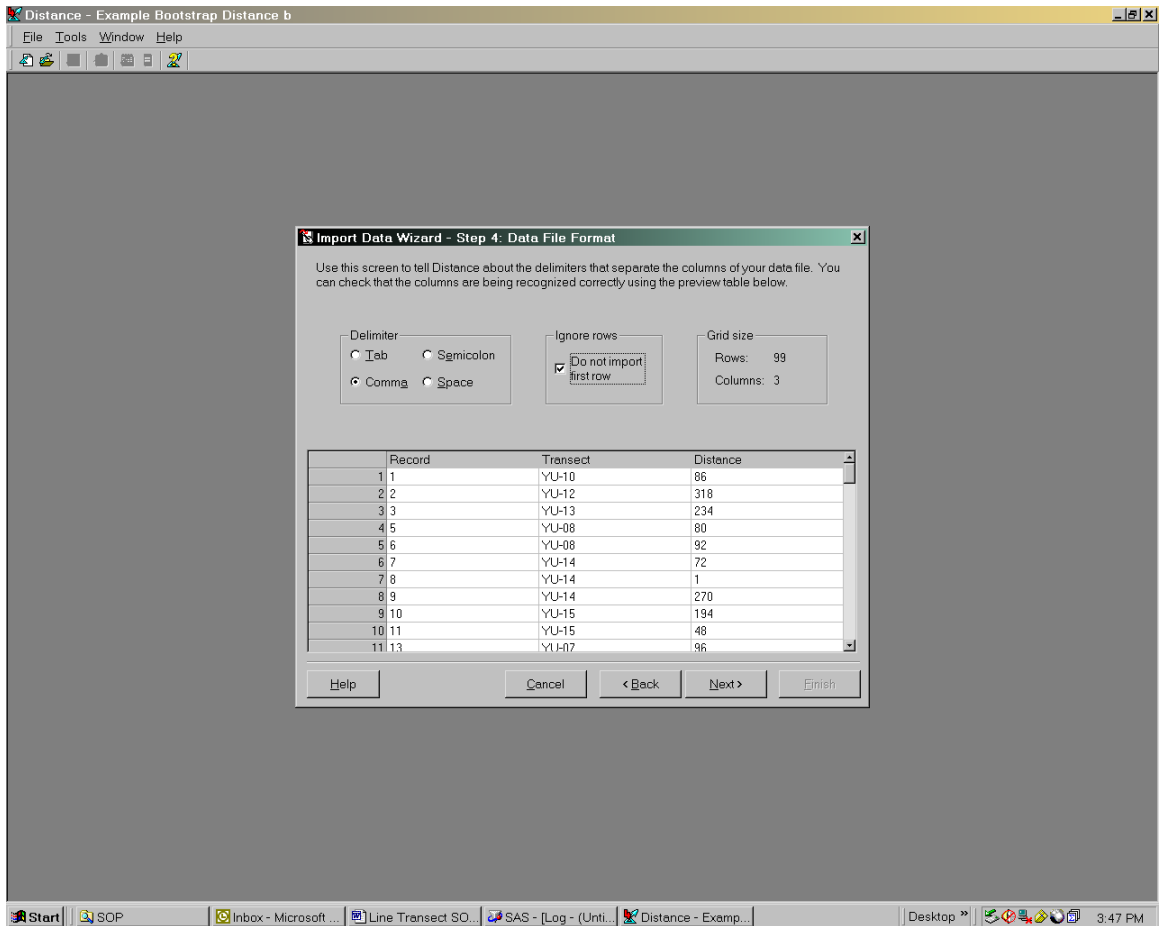
### Step 10.

Locate the file containing all valid observations from the rear seat observers within the search area and click on the OK button.



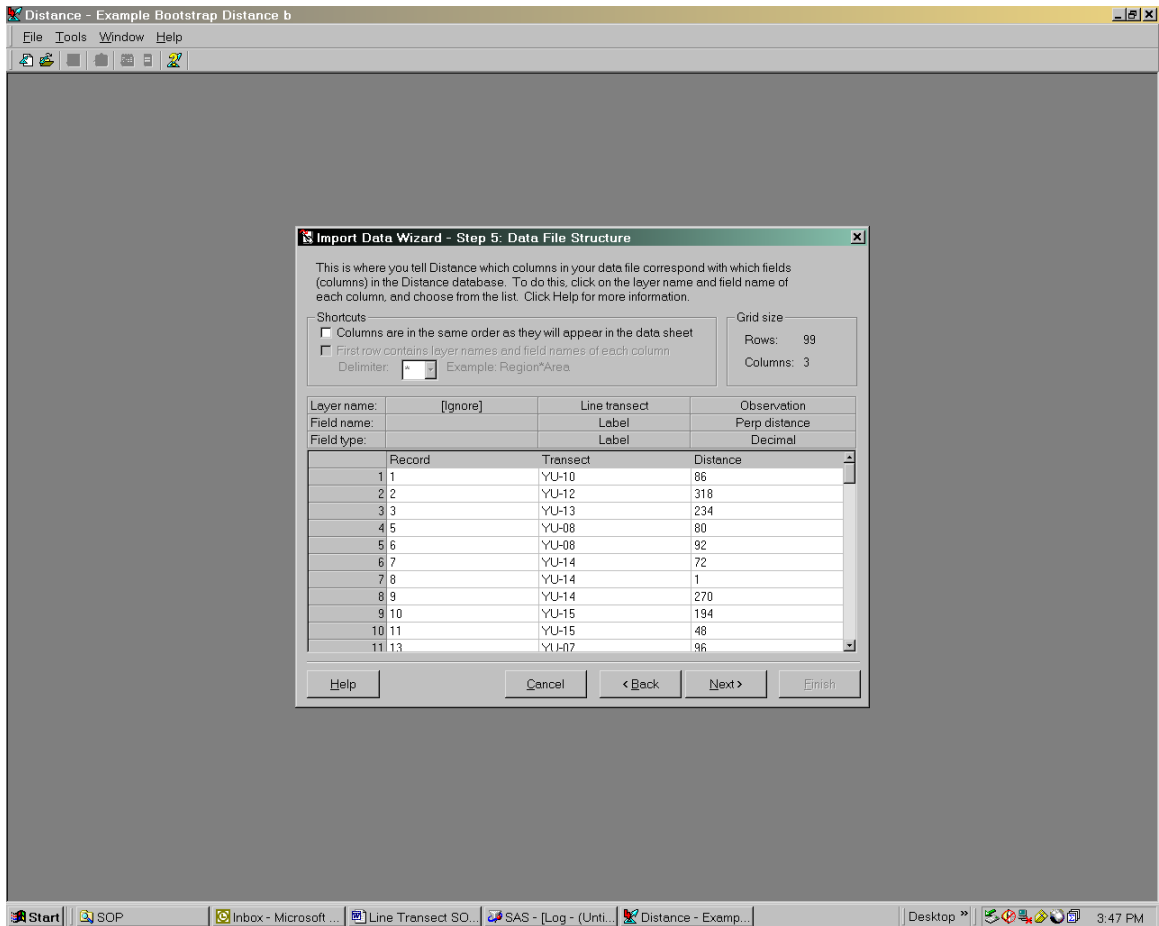
### Step 11.

Select “Observation” as the lowest data layer, and “Line transect” as the highest data layer. The remaining default settings for this screen are OK. Click on the NEXT button to proceed.



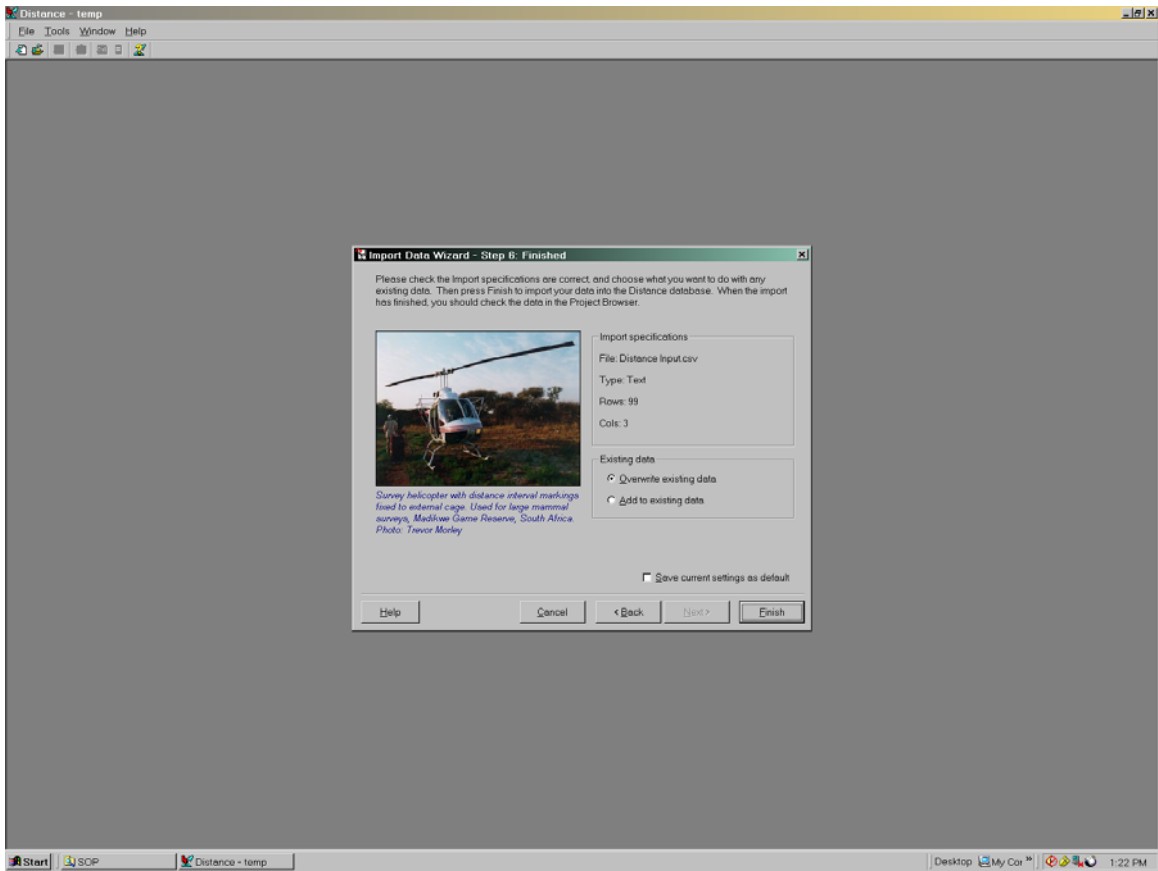
### Step 12.

Select the correct delimiter for the import data file. In this example, a CSV file is being imported and the first row in the file contains column headings, which are ignored by selecting “Do not import first row”. Click on the NEXT button to continue.



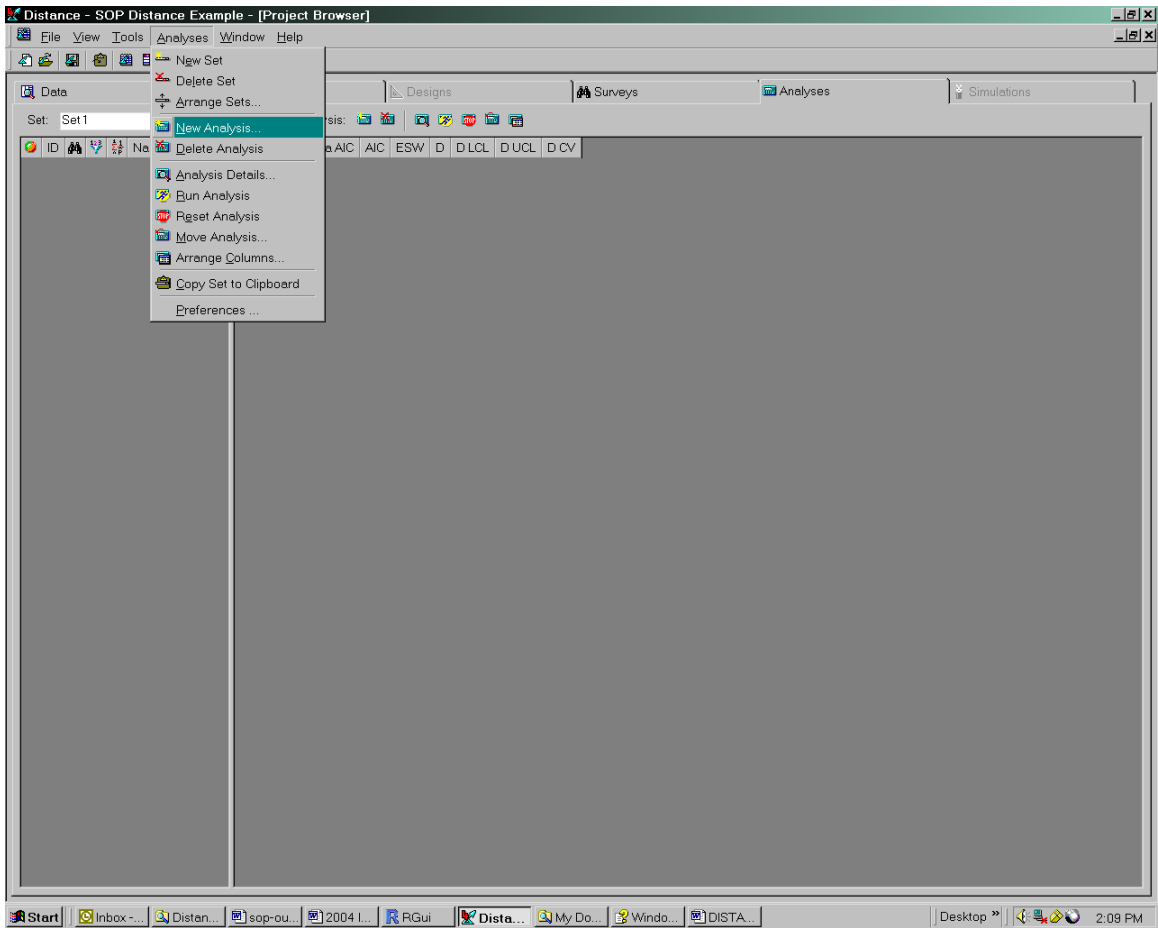
### Step 13.

Specify the data file structure. Click on [Ignore] above the second column and select “Line transect”. Then click on [Ignore] above the third column and select “Observation” and then “Perp distance”. Click on the NEXT button to proceed.



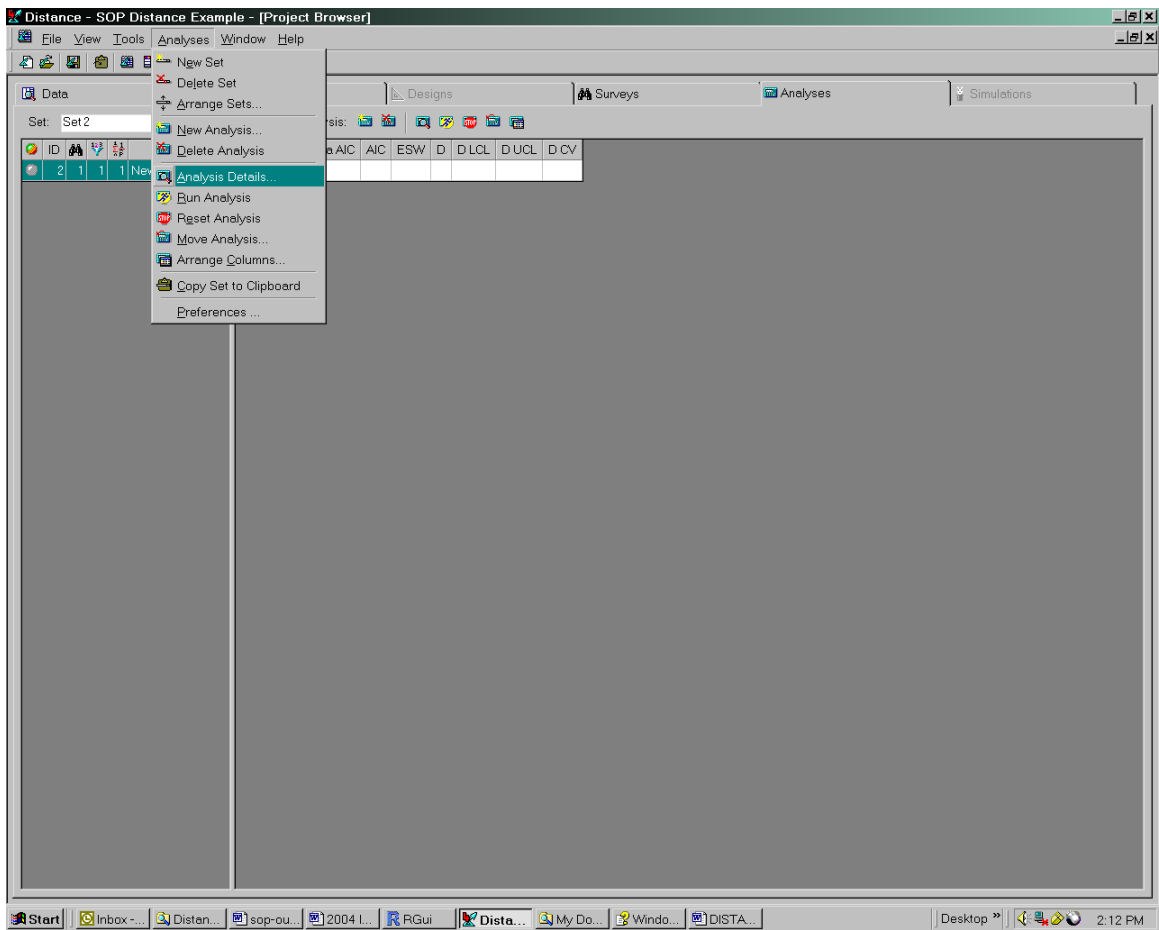
#### Step 14.

To be on the safe side, check the box specifying to “Overwrite existing data” and then click on the FINISH button to complete the Import Wizard.



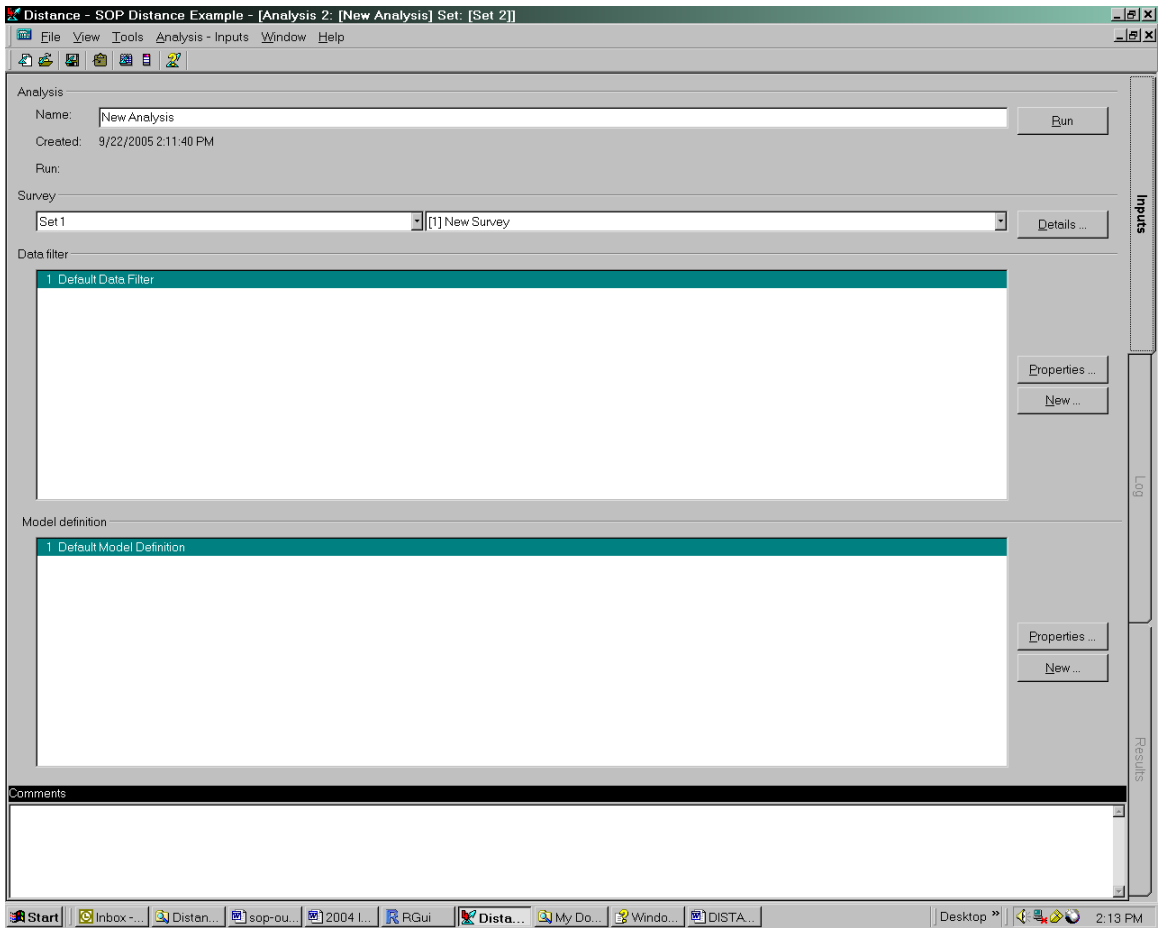
**Step 15.**

Select the Analysis tab, and then select “New Analysis”.



### Step 16.

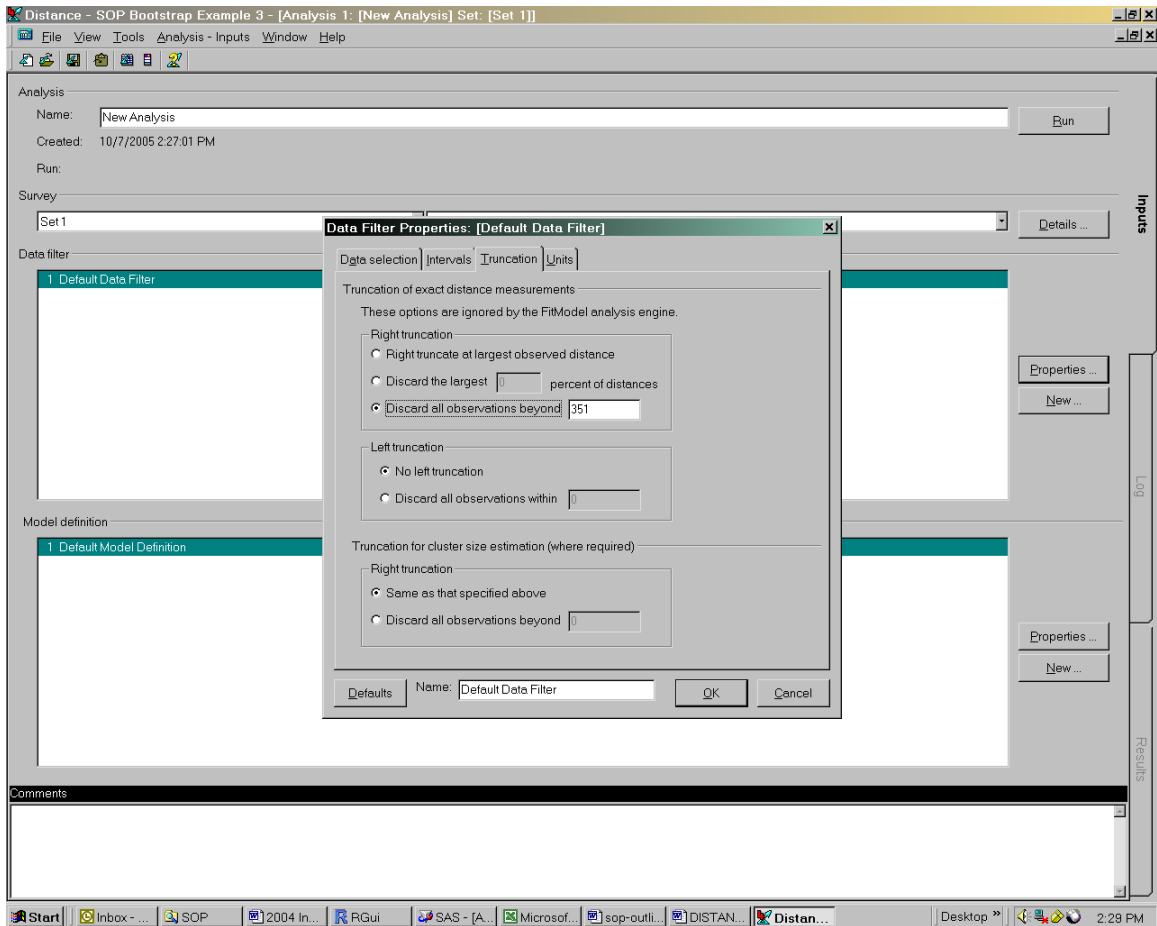
Select the Analyses tab and then “Analysis Details”.



**Step 17.**

When this window appears, click on the PROPERTIES button for the Data Filter.

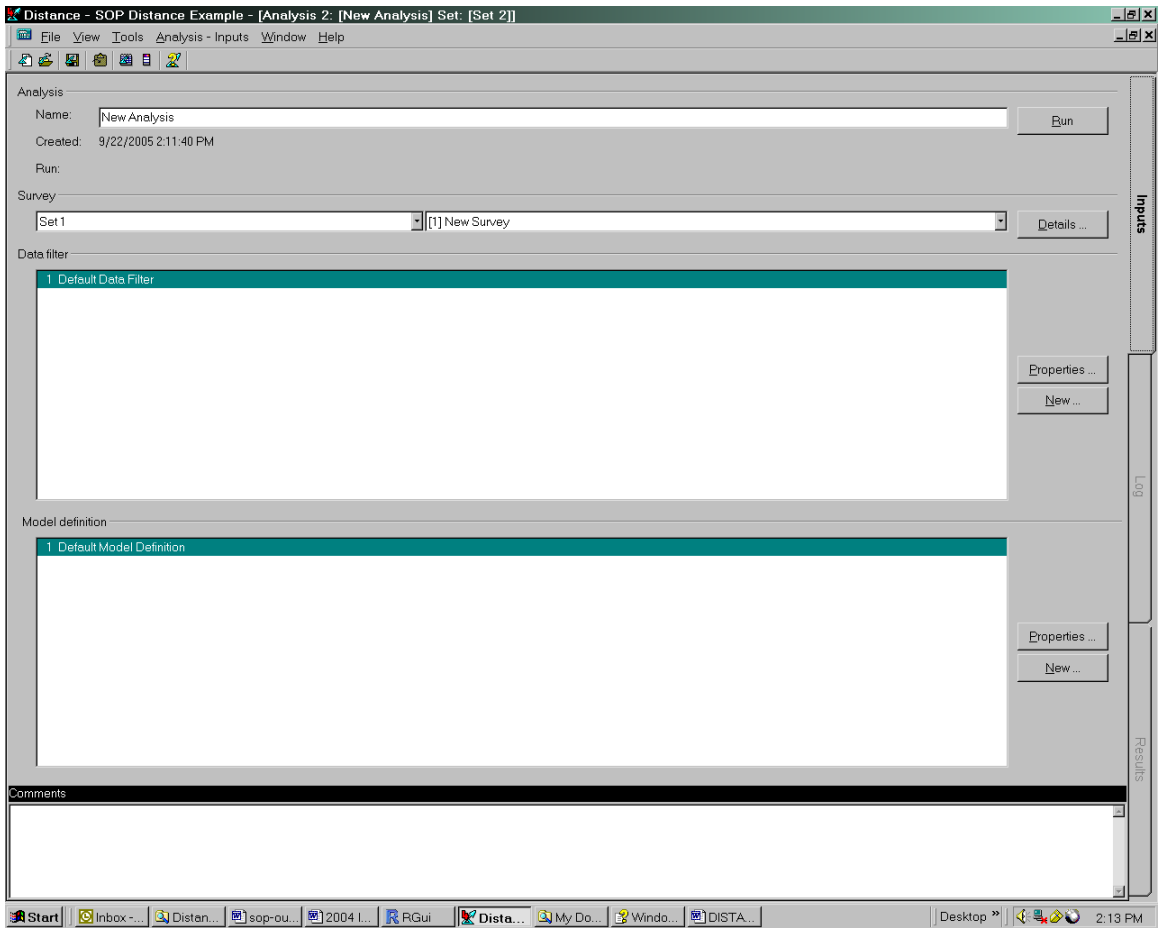




### Step 18.

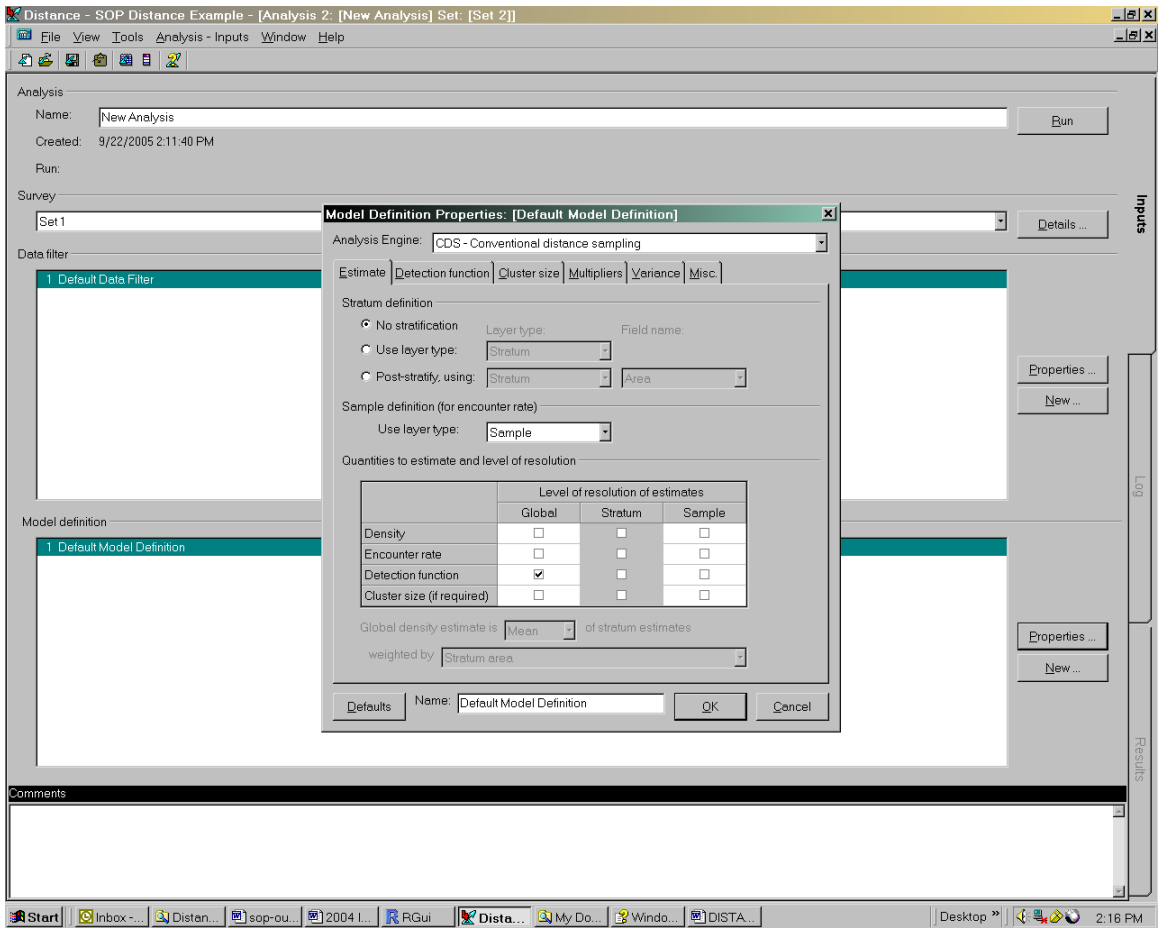
Selection the “Truncation” tab and select “Discard all observations beyond ...”.

Enter in the search width from one side of the aircraft. For example, if the maximum distance for a moose group used in the analysis is 396 m, and the minimum available sighting distance was 45 m, input 351 m ( $396 - 45 = 351$ ). Here you are telling DISTANCE to restrict its estimation of the average probability of detection to the interval  $[0, 351]$ . Prior to data export from SAS and importing into DISTANCE, all observed distance measurements were shifted to the left by subtracting the minimum available sighting distance (*see* Appendix 4 and Appendix 5.1). Click OK to proceed.



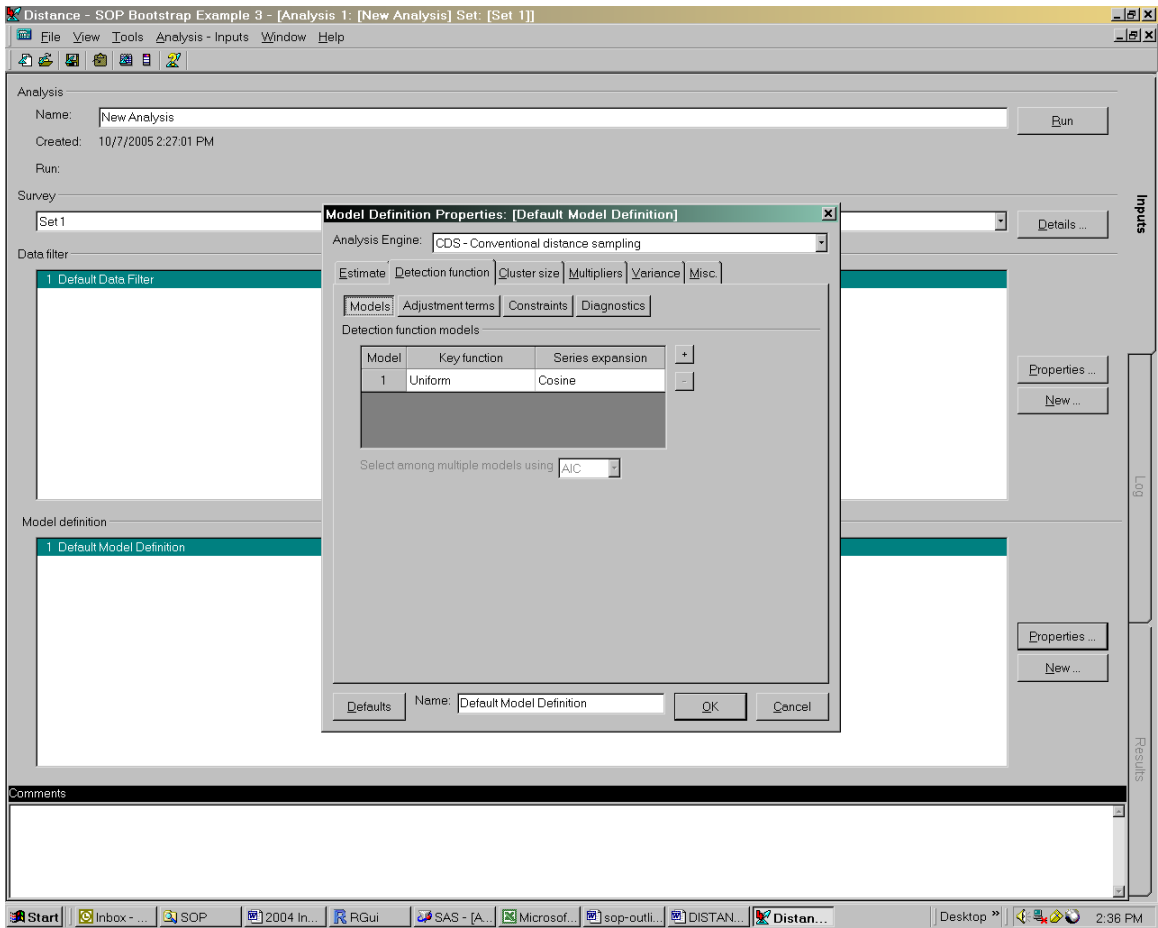
**Step 19.**

When this window appears again, click on the PROPERTIES button for the Model definition.



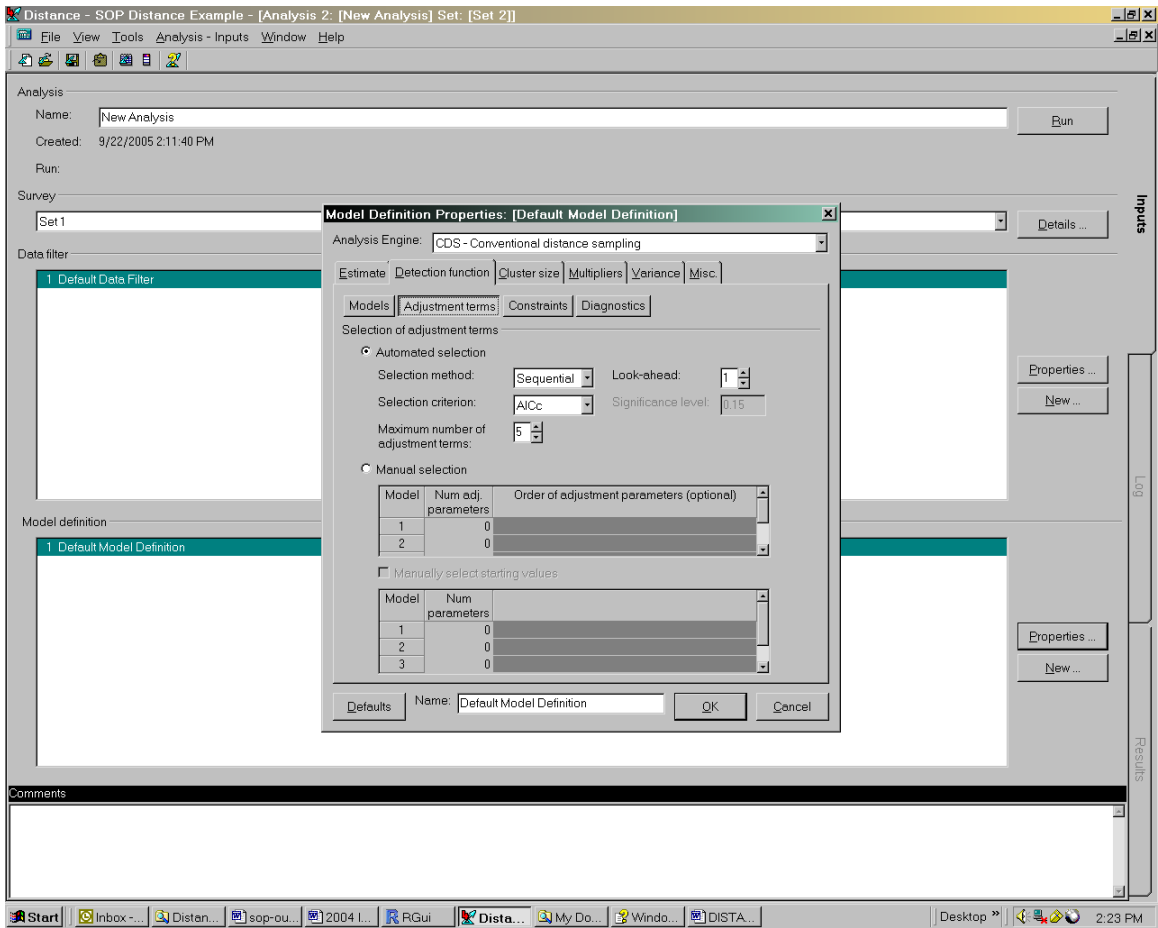
## Step 20.

Select the “Estimate” tab and specify that no stratification is needed and that only estimates of the detection functions are needed at the Global level.



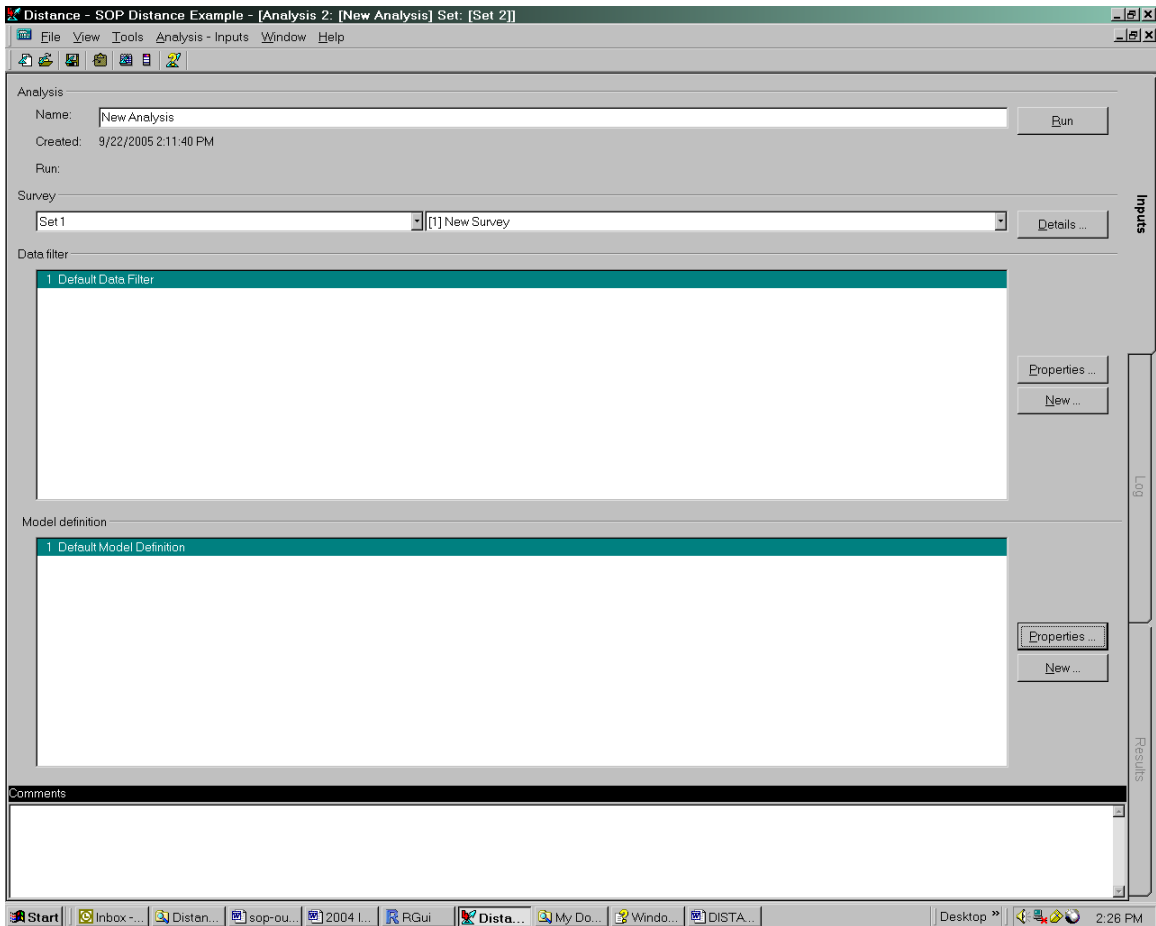
### Step 21.

Select the “Detection Function” tab and specify which detection function model you want to consider.



## Step 22.

Click on the ADJUSTMENT TERMS button within the “Detection Function” tab and specify “Automated Selection” using a sequential method and AICc up to a maximum of five adjustment terms. Click on OK at the bottom of the Model Definitions Properties box to continue.



### Step 23.

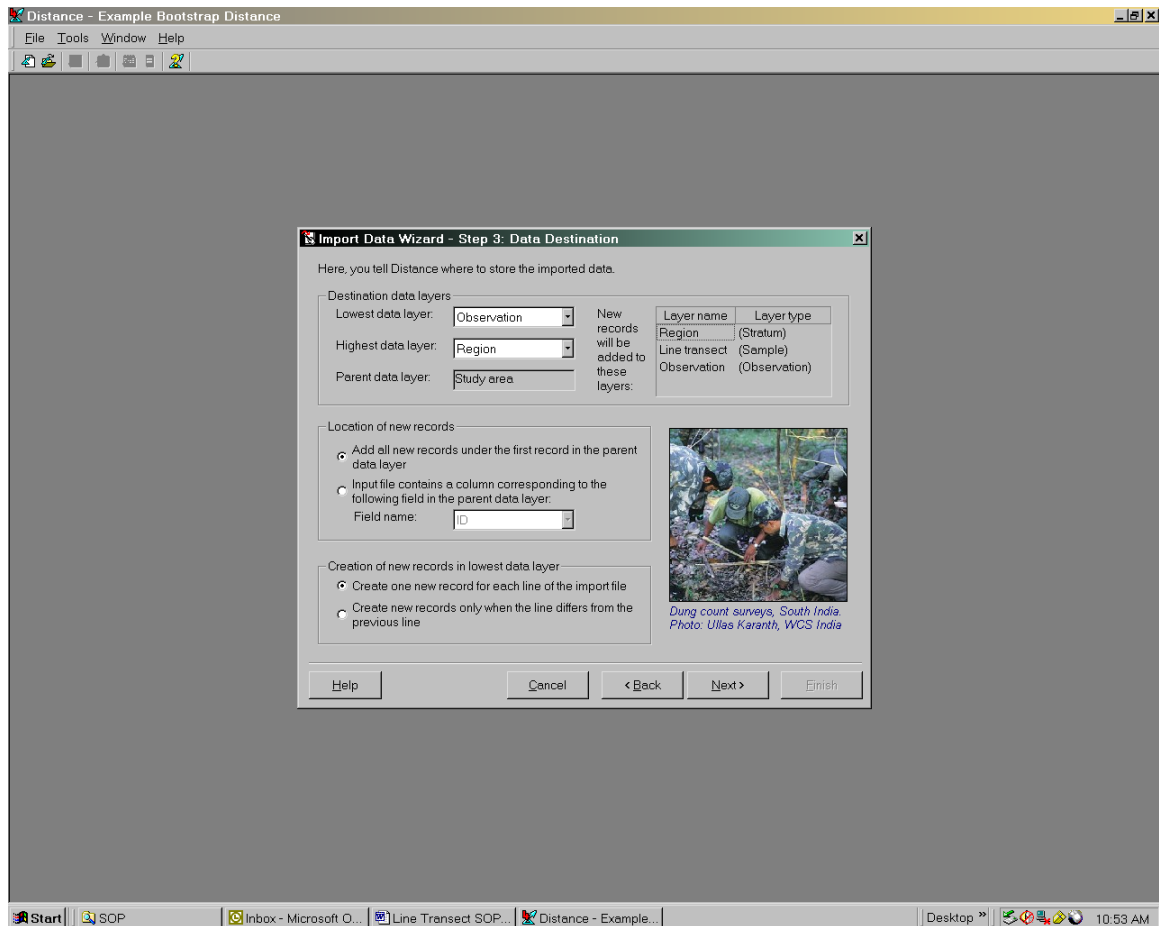
Click RUN to begin the analysis. Details of the analysis, including AICc values and the estimate of  $P$ , the average probability of detection within the search area given  $g(0) = 1$ , can be found by scrolling through the pages in the Results tab, along with the final estimated detection function plotted over several histograms of the observed distances.

Repeat steps (19) – (23) for each detection function model under consideration (uniform key functions with cosine expansion, uniform key function with simple polynomial expansion, half-normal key function with hermite polynomial expansion, and hazard-rate key function with cosine expansion all should fit the survey data). These steps must be repeated separately for each model to obtain estimates of  $P$ , along with model AIC or AICc values for each model.

## A.6.2 Using DISTANCE When Bootstrapping

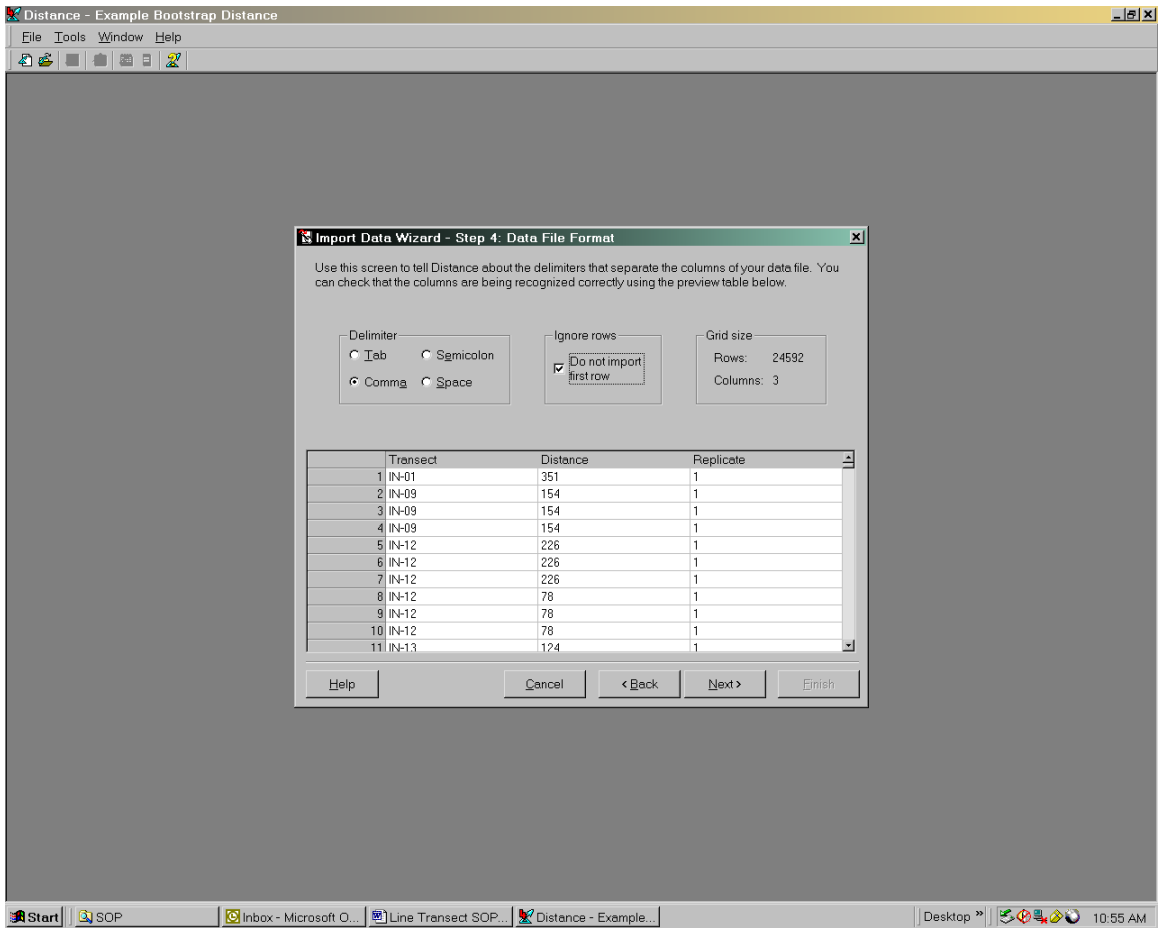
### Step 1.

Follow steps 1 – 10 in Appendix 6.1 to import a data file containing observations for multiple bootstrap samples (BS\_moose\_1.csv through BS\_moose\_4.csv created by SAS in Appendix 5.4).



### Step 11.

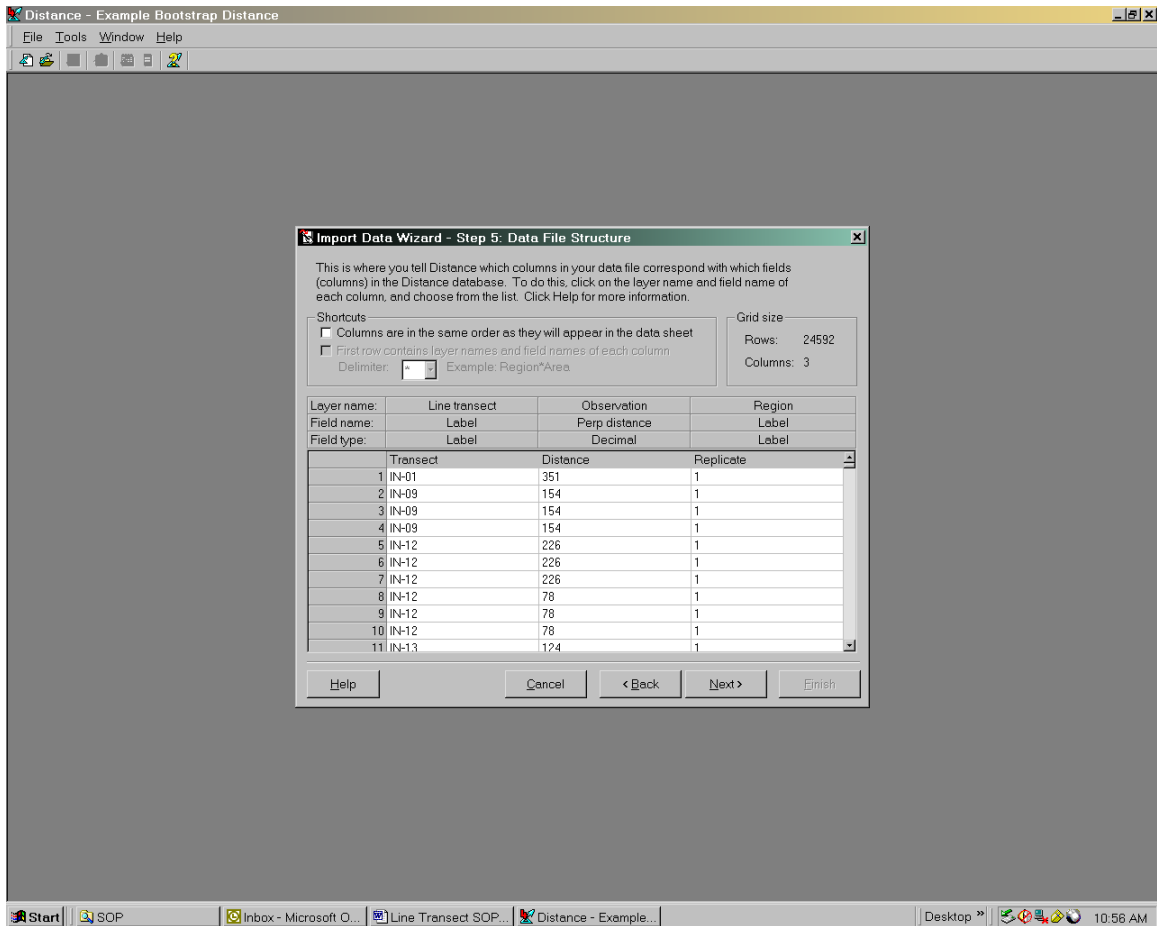
Select “Region” as the highest data layer and click on NEXT to continue.



### Step 12.

Select the correct delimiter for the import data file. In this example, a CSV file is being imported and the first row in the file contains column headings, which are ignored by selecting “Do not import first row”. Click on the NEXT button to continue.

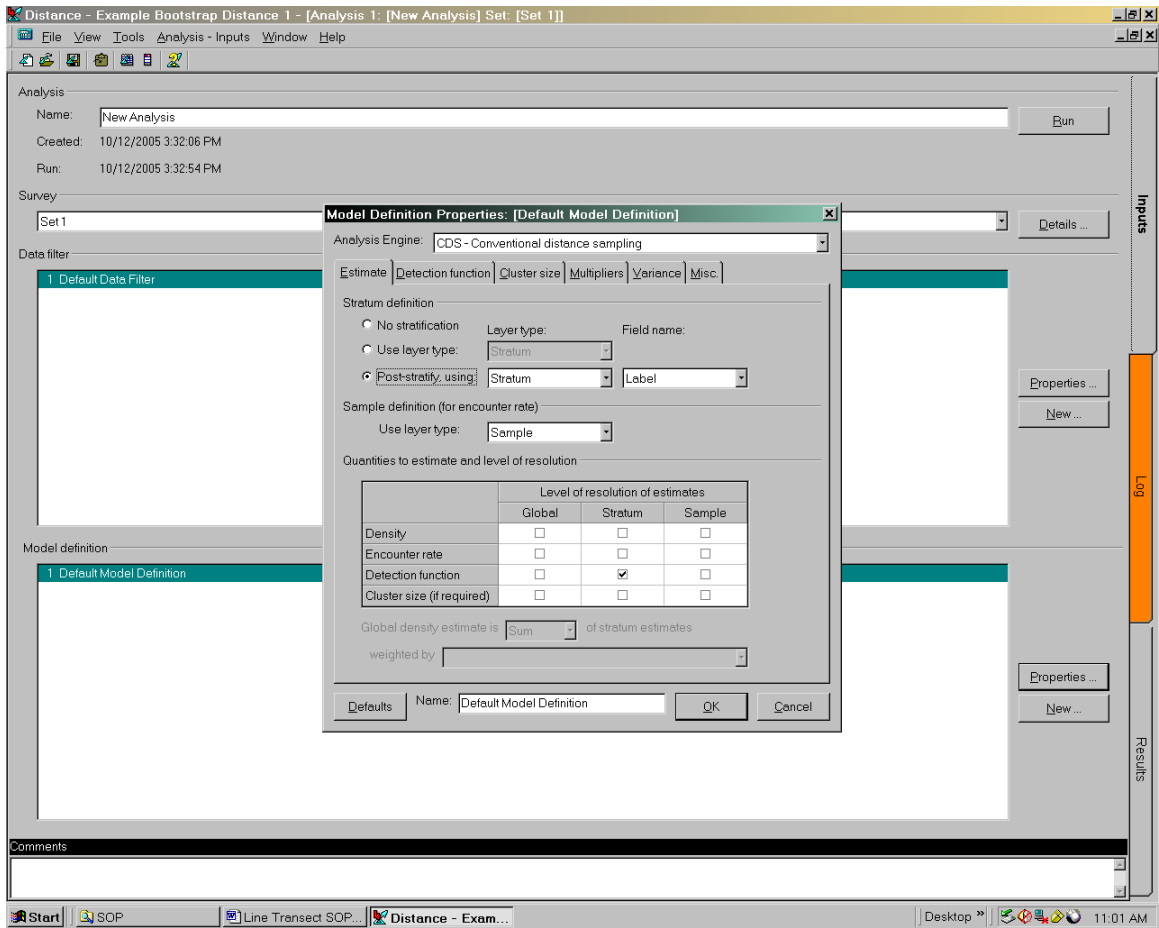




**Step 13.**

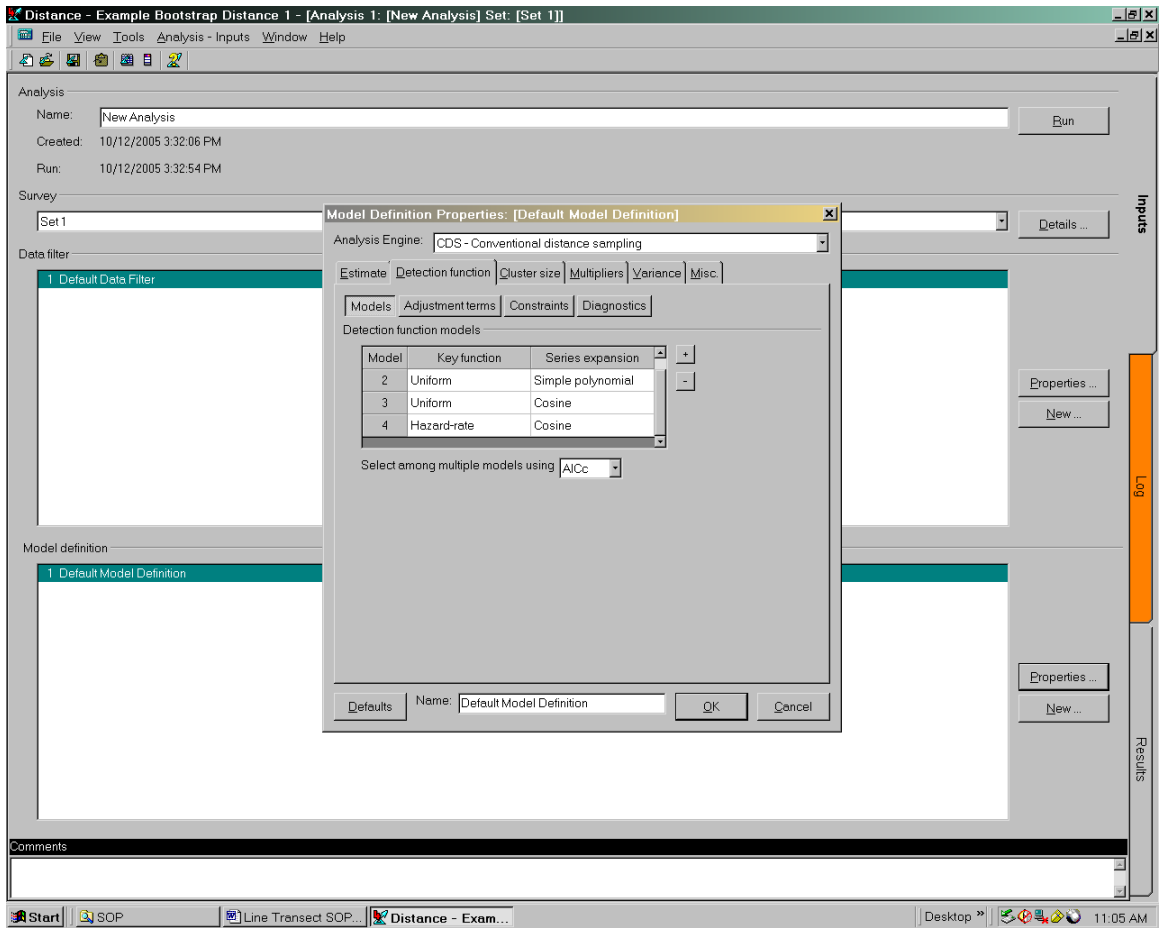
Specify the data file structure. Click on [Ignore] above the first column and select “Line transect”. Then click on [Ignore] above the second column and select “Observation” and then “Perp distance”. Finally, click on [Ignore] above the third column and select “Region” and click on the NEXT button to proceed.

Repeat steps 14 – 19 described in Appendix 6.1 to analyze the data with the 4 different detection functions.



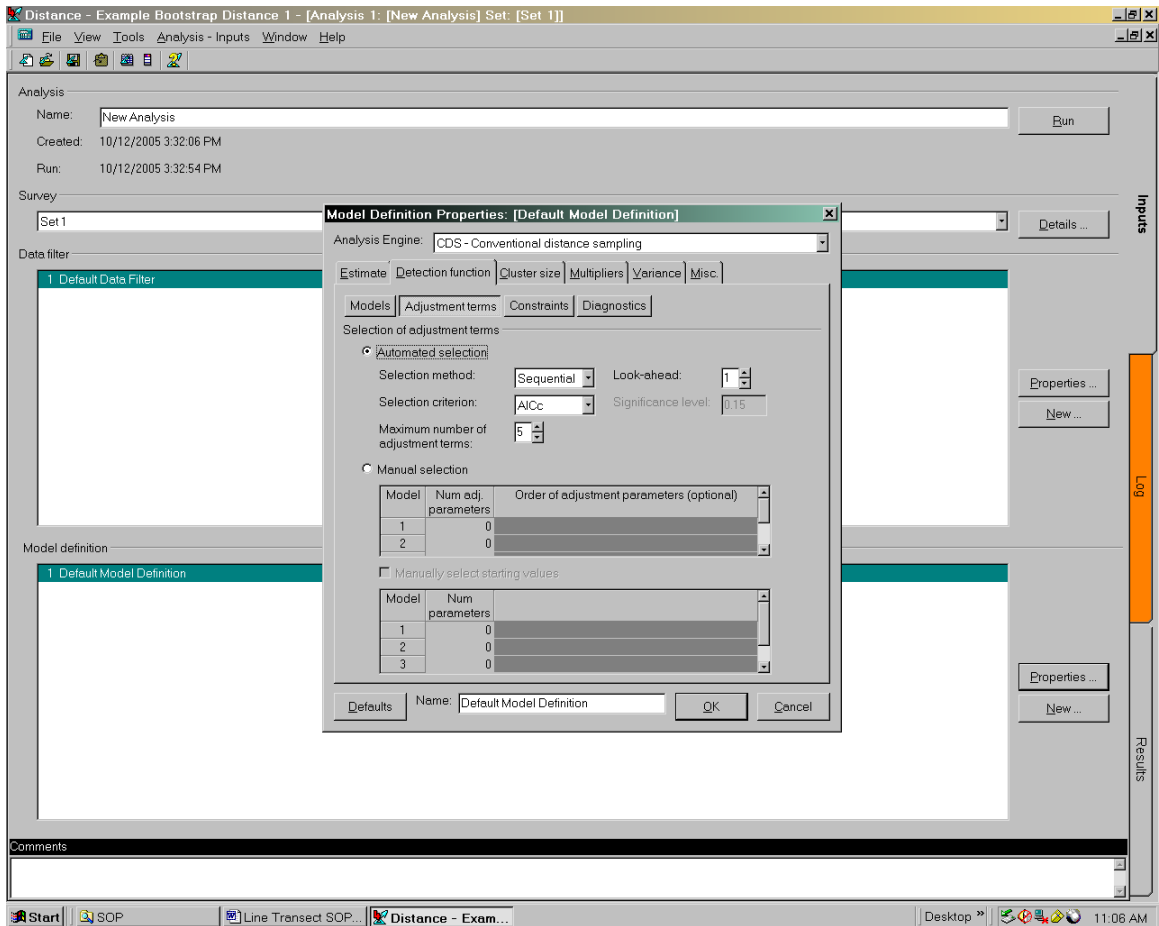
## Step 20.

Select the “Estimate” tab and specify post-stratification using “Stratum” as the Layer Type and “Label” as the Field Name. Then, select that the detection function is to be estimated at the “Stratum” level of resolution and click on OK to continue.



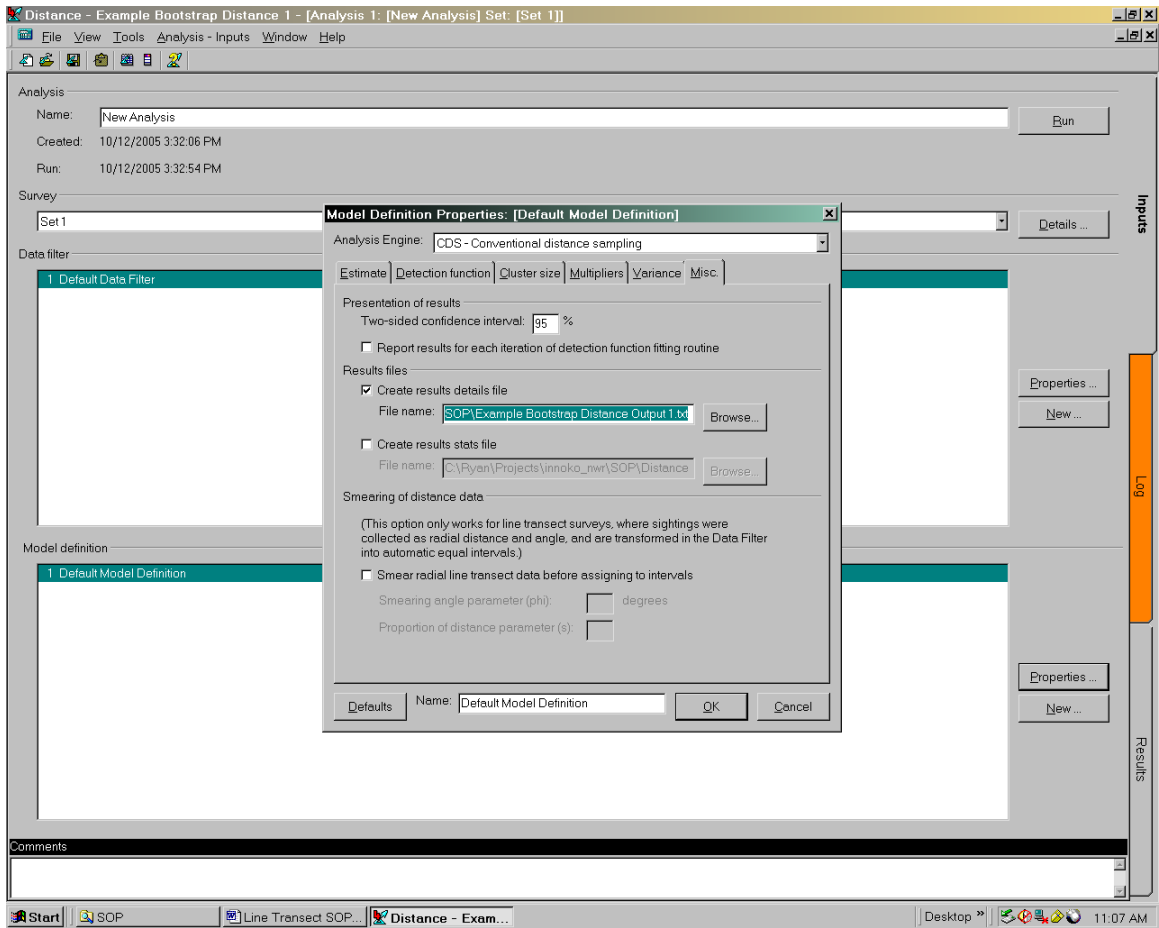
### Step 21.

Select the “Detection Function” tab and specify which detection function models you want to consider. Specify that AICc should be used to select among the multiple models.



## Step 22.

Click on the ADJUSTMENT TERMS button within the “Detection Function” tab and specify “Automated Selection” using a sequential method and AICc up to a maximum of five adjustment terms. Click on OK at the bottom of the Model Definitions Properties box to continue.



### Step 23.

Select the “Misc” tab and specify that a Results Details file should be created from DISTANCE. Finally, click on the OK tab and select RUN.