SHARE PROGRAM LIBRARY AGENCY

PROGRAM NUMBER

175003

University of Miami

1365 MEMORIAL DRIVE - CORAL GABLES, FLORIDA
(305) - 284-6257

SHARE PROGRAM LIBRARY SUBMITTAL FORM

SHARE PROGRAM LIBRARY AGENCY
Triangle Universities Computation Center
Post Office Box 12076
Research Triangle Park, North Carolina USA 27709

SPLA CONTROL NUMBER: CC240

This form should be completed and submitted with the program package to the SHARE Program Library Agency at the address shown above. Standards and instructions for submitting programs are in the SHARE Reference Manual, Section 6.

(1) Program Number (to be filled by SPLA) 360D-17.5.003

(2) Title of Program . . . . . TRANSECT

(3) System Type(s) (Machine) . . . . All major machine types, has been run on CDC, Burroughs and Dec

(4) Search Key(s) . . . .
Line transect sampling
Population density estimation

(5) Programming Systems/Languages . . . . ANSI FORTRAN IV

(6) Primary Subject Code . . . . 17-5 Biology

(7) Minimum System Requirements 150_8 K words-CDC6600 w/o overlay, input & 1 output file peripherals for 1

(8) New (N) or Revision (R) (if revision, show prior Program Number in Item 1) N

(9) Date of Submittal 4/21/80

(10) Documentation (number of original pages submitted) . . . . User's Manual - 20p., Sup. Document. - 7p.

(11) Author's Name and Address . . . . .

Jeffrey L. Laake	Kenneth P. Burnham
IATTC	USFWS, WELUT
c/o Scripps Inst. Ocean.	2625 Redwing Rd.
La Jolla, Ca 92037	Ft.Collins,Co 80521

(12) Direct Technical Inquiries to Name & Address (if different than Author)
Jeffrey L. Laake
Inter-American Trop. Tuna Com.
c/o Scripps Inst. of Oceanography
La Jolla, Ca 92037

(13) Submitter's Installation Membership Code . . . . .

(14) Abstract (should contain sufficient information for a reader to determine the value of the program). Listed on the reverse side of this form are subjects which may serve as a guide for a descriptive abstract.

SPLA 1001 Revised 11-76

SHARE PROGRAM LIBRARY SUBMITTAL FORM

Subject Guide:

a. Purpose
b. Programming Language used
c. Version and modification level or release number
d. Field of application
e. Type of routine (main program, subroutine, etc.)
f. Specific description of machine requirements

DISCLAIMER

Triangle Universities Computation Center (TUCC) serves solely as the distribution agent for contributed programs and does not test or maintain them. They are distributed essentially in the original form submitted by the author. Neither TUCC nor SHARE, INC., makes any warranty, expressed or implied, as to the documentation, function, or performance of the contributed programs.

Program TRANSECT(version 1.1) is written in ANSI FORTRAN IV and is composed of a main routine and 55 subroutines. TRANSECT provides an analysis of data from line transect sampling for the purpose of estimating the density of biological populations. The analysis consists of graphical representations of the data(e.g.,histograms,cdf plots and cross plots) and point and interval estimation of the model parameters and of the density. The specific estimation techniques used in TRANSECT are described thoroughly in 'Density Estimation of Biological Populations from Line Transect Sampling' by Kenneth P. Burnham, David R. Anderson, and Jeffrey L. Laake, Wildlife Monograph No. 72.

TRANSECT should run on any machine which has a FORTRAN IV compiler and at least the equivalent of 150K(octal) words(on the CDC 6600) of addressable main memory. This latter requirement can be reduced considerably if the program is overlaid and/or if a virtual array capability exists on the machine. An overlaid and slightly modified version has been constructed to run on a 16 bit word machine(PDP 11/34) limited to 32K words of addressable memory. As a minimum the I/O requires 2 files: 1)input-80 character record file and 2)output-132 character record file.

(Please attach additional pages if necessary) . . . . . Total pages attached ________

An "Acknowledgement of Assistance" statement must be attached to this Submittal Form.

Permission to Publish

"I hereby give the SHARE Program Library Agency permission to reprint, reproduce, and distribute this program"

(15) Signature of Submitter and Date Jeffrey L. Laake 4/21/80

(15) Signature of Installation Addressee ________

SPLA 1001 (R) Revised 11-76

TRANSECT

360D-17.5.003

TAPE KEY

File 1	Program Source (EBCDIC) 7357 card images Logical Record length 80 characters Block size 8000 characters (73 blocks of 8000 + final block of 4560)
File 2	Sample Data (EBCDIC) 1093 card images Logical Record length 80 characters Block size 8000 characters (10 blocks of 8000 + final block of 7440)
File 3	Sample Output (EBCDIC) 11,836 line images Logical Record length 132 characters Block size 1320 characters (1183 blocks of 1320 + final block of 792)
File 4	Supplementary Documentation (EBCDIC) 441 card images Logical Record length 80 characters Block size 8000 characters (4 blocks of 8000 + final block of 3280)

SUPPLEMENTAL DOCUMENTATION

This documentation will attempt to explain some points about TRANSECT which are not described in either the User's Manual nor the monograph('Estimation of Density from Line Transect Sampling of Biological Populations',by Kenneth P. Burnham,David R. Anderson and Jeffrey L. Laake, Wildlife Monograph(no.72)). The monograph describes thoroughly the purpose and methods used in TRANSECT and the User's Manual describes the control card and data structures. This documentation will address program limitations, environment requirements, program testing and the sample data and output.

PROGRAM LIMITATIONS

The major limitation of the program is in the sample size restrictions which can be changed without too much dificulty. the following limitations are imposed:

no. of observations with ungrouped data	1600
no. of replicates in the data(e.g.,lines)	100
no. of intervals for cut points and grouped data	20
no. of parameters in the estimator	15

The last two are fairly minor and will probably never restrict a user. The other two could possibly be a restriction so the necessary program changes have been indicated through comment cards in the source code.

The other limitation is in the accuracy of the program. Some parts of the program compute maximum likelihood estimates which are dependent on mumerical methods for the maximization of a function and for integration. The accuracy of the maximization can be improved by allowing the technique a larger number of iterations. It is presently limited to 40 iterations. The numerical integration is accomplished via a Newton-Cotes method and its accuracy can be increased by dividing the range into smaller intervals. All of the numerical methods are done with double precision variables. The program has successfully run on a PDP 11/34 with a 16 bit-word and a 32 bit double precision variable without any major numerical problems occurring.

ENVIRONMENT REQUIREMENTS

The environment requirements are quite minimal except that the program requires a large portion of main memory to execute unless it is overlaid. As a minimum the program requires two physical files for input and output. These can be from a card reader and printer or both from a disk or other medium. The size of the program will be the biggest factor limiting its use. On a CDC 6600 it requires approximately 150 K(octal) words without an overlay. This can be reduced considerably by reducing the sample sizes and/or overlaying the program. Two overlay diagrams are included to help initiate the design of an overlay structure. Figure 1 shows a structure with overlay areas in which manual loading of modules could be used. Figure 2 shows a tree structure which has been used on the PDP 11/34 with automatic loading of modules.

The program is written in ANSI FORTRAN IV so naturally the environment will require a FORTRAN compiler. There is only one known nonstandard type of statement in the program. This is the use of END=sn in the READ statement where sn is a statement number. These will not compile on CDC computers which use an EOF(fn) where fn is a file number. All of these statements have been flagged in the source code by comment cards which the code for the necessary changes.

PROGRAM TESTING

TRANSECT has been tested thoroughly by running it on data sets for which a correct result was known and also by working through examples by hand. The descriptive functions of TRANSECT (means, standard deviations, crossplots, histograms and cdf plots) can easily be and have been checked by hand. The estimation functions of TRANSECT have been checked by simulating data via the computer and physical simulation from models with known parameters(e.g., a specified pdf for perpendicular distances or a specific type of flushing curve for sighting distances and angles) and comparing the estimated parameters to their known counterparts.

SAMPLE DATA AND OUTPUT

Several sets of sample data and output have been included with the program to allow the program to be checked at a particular installation and to further illustrate the control and data structures. These data sets have been used as examples throughout the monograph and they are explained thoroughly in it. A minor difference in the construction of the histograms for ungrouped data has caused some slight differences between the output in the monograph and the sample output. Presently, TRANSECT constructs histograms by including the right interval end-point in each interval and for the first interval it also includes zero(the left interval end-point). The program previously included the left interval end-points and the last interval included the final right end point. This would not cause any differences at all if the data were truly continuous without any rounding. However, there is a considerable amount of rounding with sample data sets and slight differences in the output result.

OVERLAY AREA 0	MAIN ROUTINE
OVERALY AREA 1	ROUTINES 1-6
OVERLAY AREA 2	ROUTINES 7-17
OVERLAY AREA 3	ROUTINES 18-29
OVERLAY AREA 4	ROUTINES 30-33

Figure 1. The overlay structure for manual loading is shown above. The object module or group of concatenated modules associated with each number are listed in table 1. The routines within an overlay area are independent so they can overlay one another. The routines in different areas may reference one another.

```
                  MAIN
                   1
 -------------------------------------
 1           1           1           1
 1           2           5           6
             1                       1
          ---------              ---------
           1     1                1     1
           3     4                7     8
                                  1
                           ---------------
                            1      1     1
                            9      10    11


                NULL ROOT
                    1
 ---------------------------------------
 1      1      1      1      1      1    1
 12     13     14     15     16     17   18


                NULL ROOT
                    1
            ---------------
            1      1      1
            19     20     21
```

Figure 2. The tree structure for overlaying TRANSECT is composed of three co-trees. The object module or group of concatenated object modules associated with each number are listed in table 2. Routines in different co-trees may reference one another as well as routines which lie on the same path(branch). Routines on different branches within the same tree are independent and overlay one another.

Table 1. List of object modules associated with numbers in figure 1. The dash infers that the modules are concatenated

NO.	ROUTINE NAMES
1	CNTRL-KEYWRD
2	GRPD
3	UNGRPD
4	SUMARY
5	PEREST
6	SIGEST-ELLIPS
7	MLE-MARQUA
8	CRSPLT
9	PARAM
10	MEAN
11	PRINT-LINE
12	READ
13	UNIVAR-PHIST
14	FSEST
15	PLTMOD
16	TEST
17	EXPT-UTEST
18	HISTGM
19	FREE-SKIP
20	SORT-PERMUT-PART
21	COR
22	CHI
23	TVALUE
24	CELLS-CHECK
25	UMLE
26	INVERT-MINV
27	PLT-[illegible]-SETUP-CLEAR-DVSION
28	PCDF-PLOTF
29	CALC-STATS-START-UPDATE-XLOG-DOMAIN
30	ONE
31	NCNI
32	NAME
33	HEADER

Table 2. List of object modules associated with the numbers in figure 2. The dash indicates the modules are concatenated.

NO.	ROUTINE NAMES
1	CNTRL-KEYWRD-GRPD
2	UNGRPD
3	CRSPLT-MEAN-PRINT-LINE-READ
4	UNIVAR-PHIST
5	SUMARY
6	UTEST
7	PEREST
8	SIGEST-ELLIPS
9	FSEST-PLTMOD
10	TEST-EXPT
11	MLE-MARQUA
12	CALC-STATS-START-UPDATE-XLOG-DOMAIN-HISTGM-FREE-HIST-SKIP-SORT-PERMUT-PART
13	COR-CHI-TVALUE-PARAM-NAME
14	CELLS-CHECK
15	UMLE
16	INVERT-MINV
17	PCDF-PLOTF
18	PLT-PRNT-SETUP-CLEAR-DVSION
19	ONE
20	NCNI
21	HEADER

USER'S MANUAL FOR PROGRAM *TRANSECT*

by

Jeffrey L. Laake[1,3]

Kenneth P. Burnham[2]

David R. Anderson[3]

[1]Inter-American Tropical Tuna Commission, c/o Scripps Institution of Oceanography, La Jolla, Calif. 92037

[2]U. S. Fish and Wildlife Service, Western Energy and Land Use Team, Ft. Collins, Colo. 80526

[3]Utah Cooperative Wildlife Research Unit, Logan, Ut. 84321

Cooperators:
1. U. S. Fish and Wildlife Service
2. Utah State University
3. Utah Division of Wildlife Resources
4. Wildlife Management Institute

CONTENTS

ACKNOWLEDGMENTS

We thank G. White for his help in developing and testing the program. We also express our appreciation to R. G. Streeter, U. S. Fish and Wildlife Service, who provided encouragement during this work, and to B. Johnson for the typing of this manuscript.

This work was performed under contract 14-16-0008-1223 of the Coal Program, Office of Biological Services, U. S. Fish and Wildlife Service, to the Utah Cooperative Wildlife Research Unit. Funds for this work were made available to the Fish and Wildlife Service as part of the Federal Interagency Energy/Environment Research and Development Program, Office of Research and Development, U. S. Environmental Protection Agency (IAG-EPA-D5-E385). Publication costs were paid by the Office of Cooperative Units, U. S. Fish and Wildlife Service.

INTRODUCTION

TRANSECT is a comprehensive computer program which provides an analysis of line transect data for estimation of density. TRANSECT was developed at Utah State University under a contract with the U.S. Fish and Wildlife Service to further develop the theory and applications of line transect sampling. This work also resulted in a publication by Burnham et al. (in prep.) which provides a thorough reference on many aspects of line transect sampling. It should be consulted in reference to the techniques used in TRANSECT. Many of the calculations required to compute the estimates of density are quite tedious, which makes a computer program like TRANSECT quite useful. In addition, the program gives various plots and graphs which can be useful aids in an analysis.

This documentation provides a reference for the use of TRANSECT. It explains all of the necessary control cards and the data structure needed to execute the program. Also, it gives detailed examples of the input and an overview of the output as well. TRANSECT is a fairly general and flexible program with many different options. However, the input is very simple if the default options are chosen. Many of the options which are discussed may not be relevant for all users. We have tried to use many examples to help illustrate the program input. The example input has been set off from the text in different type to prevent any confusion.

TRANSECT has been oriented to the analysis of perpendicular distances for density estimation because there has been much more research done on estimators for perpendicular distance and because the basic model is simpler and requires fewer assumptions than sighting distance-angle models. TRANSECT does analyze sighting distance-angle data but the majority of the options are slanted towards perpendicular distances. We strongly suggest analyzing the data for most situations on the basis of perpendicular distances.

PROGRAM INPUT

Overview

The input for TRANSECT is divided into two sections: (1) the control cards and (2) the data. The control cards are used to describe the data format and to give commands for the types of analyses to be performed. The purpose of the data section is self-evident.

The control cards and data can be input into the program from the same or different computer files. The control cards are always input from File 5 which by default is usually a card reader on most computer systems. The data is also input from File 5 by default; however, it can be input from another file by specifying the file number in the control cards (discussed later). File 6 is the only file which cannot be specified for input because it is reserved for program output (which by default is usually a printer). It is the responsibility of the user to establish the links between the physical and logical files if they are not done so by default or if they are changed from default (this will vary between computer systems). Some possible combinations of input and file structure are:

(1) read both control cards and data from File 5 (a card reader)

(2) read both control cards and data from File 5 (where File 5 is defined as a disk file)

or (3) read control cards from File 5 (a card reader) and the data from File n (where n specifies a disk file or other medium).

This latter combination allows the data to be stored while the control cards are manipulated. One of these combinations ought to be suitable for most situations. The only limitation is that the file for the control cards should be specified such that there are 80 characters per record (i.e., a card image).

TRANSECT allows any number of sets of control cards and data to be analyzed in one program execution. If multiple sets are used, they should be separated by a flag card with the characters END. in the first four card columns. Figure 1a illustrates the input structure for multiple sets of control cards and data when they are both on the same file. If the control cards and the data are read from separate files and multiple sets are given, then it is necessary to use the flag card END. between the sets for each file. This is necessary to allow a control card set and a data set to be skipped if an error is encountered. Figure 1b illustrates the input structure for multiple sets on separate files.

Control Cards

An analysis of line transect data can vary depending on the type of data being analyzed and the types of analyses to be performed. The variations in the basic data from line transect studies include the following:

1) the kind of data measured (i.e., perpendicular distance, sighting distance, and/or sighting angle),

2) how the data were measured (i.e., measured exactly (ungrouped) or taken in intervals (grouped)),

and 3) whether the data were truncated or not (i.e., a specified transect width).

a)

File 5

Control card set 1
Data set 1

END.

Control card set 2
Data set 2

END.

.
.
.

Control card set r
Data set r

[end of file]

b)

File 5

Control card set 1

END.

Control card set 2

END.

.
.
.

Control card set r

[end of file]

File n ($n \neq 5$ or 6)

Data set 1

END.

Data set 2

END.

.
.
.

Data set r

[end of file]

Figure 1. Example illustration of multiple sets of control cards and data on a single file (a) and separate files (b). The [end of file] will vary depending on the computer system.

The types of analyses can also vary. Density estimates can be based on perpendicular distances or sighting distances and angles. Also there are a variety of estimators within each of these categories. And finally there are various ways to describe the basic data graphically.

The control cards allow the user to specify the various attributes described above for each particular analysis. The control cards are a series of card images which follow a specified logical order and which give either a label, keywords for various options and/or numbers for various attributes. The format for the control cards is very unrestrictive and with only a few constraints. In general, all data items are separated by commas or any number of blanks and character data must be enclosed by asterisks. If a control card requires more than one card image, then a dollar sign ($) should be put in column 80 to indicate that the card will be continued. At least one blank is necessary between any data and the dollar sign if it is used. Numbers only need a decimal point if they are not integer values (e.g., the numbers 9, 9., 9.0, or 9.5 are all valid). The character data on the control cards is either for a user specified label or for keywords. This character data must never exceed 80 characters for any one control card, even though a control card may be continued. The keywords used throughout the control cards are 4 letter word mnemonics. These keywords can be separated by any number of blanks or other characters or they can be run together. The only requirement is that they are spelled correctly. They have been separated by commas throughout the documentation for convenience only.

In the following, each of the control cards will be explained in the order in which they appear in the input. The format for the control card will be given followed by its purpose, some examples and any limitations. A synopsis of the order of the control cards, their content and when they are used is presented in Table 1 as a quick reference.

General Label

Format:

    Any informative label n

where n is the data input file number which is 5 by default.

Purpose:

The general label is used for identification purposes in the output. The label is printed out near the top of each page to identify the output for the analyses from the particular set of control cards.

This control card is also used to specify the file number for the data if it is different from File 5. Any file number may be used (except for 6) but the user must establish the links for the

Table 1. Synopsis of the control cards for TRANSECT. The optional parts of the control cards are in brackets.

Control Card	Content and Keywords	Used When?
General Label	Any informative label, [file no.]	Always
Conversion Factors (three cards)	Measurement keyword, [unit label] , [Explanatory material] Keywords = DIST, LENG, AREA	Always
Line Lengths	$\ell_1, \ell_2, \ell_3, \ldots, \ell_r$	Always
Sample Sizes	$n_1, n_2, n_3, \ldots, n_r$	Always
Options	Keywords, [transect width] Keywords = HELP, GRPD, DEFT, DESC PEST, SEST, CUTP, NPOL	Always
Data Status	Keywords = PERP, SIGH, ANGL	Ungrouped data
Descriptive Output	Keywords = LIST, HIST, CPLT, CDFP	Ungrouped data when DESC specified
Perpendicular Distance Estimators (NPD + 1 cards)	NPD, [keyword = STRT or MSET] *{Estimator keyword, [NP or S_i]	when PEST is specified
Sighting Distance-Angle Estimators	Keywords = GHYN, HAYN, MHYN	Ungrouped data when SEST is specified
Cut Points for Perpendicular Distances (Ncut + 1 cards)	Ncut {$C_1, C_2, C_3, \ldots, C_{kcut}$	Ungrouped data when CUTP is specified

*This bracket indicates the card is repeated.

file number to the physical file. As an example, if a disk file is used for the data and the file number is 10, then a 10 should be put on the control card.

Example:

1) An example showing the use of an alternate file.

**ANALYSIS OF EXAMPLES FOR MONOGRAPH* 10*

2) An example using the default file (unit 5).

ANALYSIS OF PORPOISE DATA

Limitations:

The general label can consist of any legal characters or numbers except an asterisk. Also it cannot exceed 80 characters.

Conversion Factors

The conversion factors require a series of three control cards each with the same format. These cards do not have to be in any specific order among themselves.

Format:

* Explanatory material with keyword and .units label.* c

where c is the conversion factor. The keywords are the following:

1) DIST[ance]

2) LENG[th]

3) AREA

The keyword is the first four characters and the rest of the characters in brackets are optional.

Purpose:

These control cards specify the labels for the units of the distance, length and area measurements which can vary depending on the particular study. Also, because it is often convenient to convert the measurements recorded in the field at the analysis stage, these control cards can be used to specify conversion factors (multiplicative) for the measurement units. The conversion factors and unit labels are identified for each particular measurement by the keywords listed above. Explanatory material can also be included on the card about the units or conversion. The material on the cards is printed in the section INTERPRETATION OF CONTROL CARDS as a reminder. The label for the measurement units is delimited from the rest of the material by periods. If a conversion factor is not specified, then it is assumed to be a one (i.e., no conversion). If no labels are given for the units, then no output is given for the labels. If both conversion factors and labels are not wanted, it is still necessary to input the control cards with the keywords on them.

The conversion factors for length and distance measurements are used to directly transform the data. The area conversion factor is different in that it specifies the conversion after the other two measurements have been converted. As an example, if distance is changed from feet to meters and the line length from miles to kilometers, then the area has to be transformed from meters X kilometers to some other unit. If it was desired to express the density in numbers per hectare, then the conversion factor in this case would be 10.

Examples:

1) The example described above follows. There are .3048 meters in a foot, 1.604 kilometers in a mile and 10 meter X kilometers (or 10 X 1000 meters squared) in a hectare which gives the following example. The labels are meters, kilometers and hectares.

```
*DISTANCE MEASURED IN FEET CONVERTED TO .METERS.* .3048
*LINE LENGTH MEASURED IN MILES CONVERTED TO .KILOMETERS.* 1.604
*AREA CONVERTED TO .HECTARES.* 10.0
```

2) As a minimum, the above example could have been specified by

```
*DIST* .3048
*LENG* 1.604
*AREA* 10.0
```

3) An example with no conversion factors needed but with labels

```
*DISTANCE MEASURED IN .NAUTICAL MILES.*
*LINE LENGTH MEASURED IN .NAUTICAL MILES.*
*AREA EXPRESSED IN .SQUARE NAUTICAL MILES.*
```

Limitations:

The explanatory material including the keyword and units label must not exceed 80 characters. The unit label must not in itself exceed 25 characters. Also, no other periods should be in the explanatory material or unit label other than the two which delimit the label.

Line Lengths

Format:

$\ell_1, \ell_2, \ell_3, \ldots, \ell_r$

Purpose:

A line transect survey usually consists of replicate lines rather than just one line. This control card provides the lengths of the r replicate lines. These lengths are used directly (after conversion if any) in the estimate of density. They are also used to compute an estimate of the weighted variance of the number of observations, n.

Example:

1) The line lengths for a survey with 4 replicate lines

20.56,19.98,2.50,12.28

2) An example with equal line lengths.

4,4,4,4,4,4,4

Limitations:

The number of line lengths (i.e., replications) must not exceed 100.

<u>Sample Sizes</u>

Format:

$n_1,n_2,n_3,\ldots,n_r$

Purpose:

This control card gives the numbers of objects observed on the r replicate lines corresponding to the line lengths on the previous control card. These sample sizes are not used for reading in the data. They are used exclusively for calculating a sample variance for the total sample size n ($n = n_1 + n_2 + \ldots + n_r$). As explained in Burnham et al. (in prep.), the variance for the density estimate is partly a function of the variance of the sample size. There are two ways of calculating a variance for n: (1) assume some distribution for the sample size which specifies the variance or (2) calculate a variance empirically from replications. The former approach is often used with the assumption that the objects are spatially distributed randomly so the statistical distribution for n is Poisson. The estimate for the sampling variance in this case is the sample size n. This is the default assumption if only one replicate is given. Thus, it can be implemented by adding all of the ℓ_i and the n_i and treating them as one replicate. If more than one replicate is given, then an empirical estimate of the variance is made.

Examples:

1) An example of sample sizes for 4 replicate lines corresponding to the example in the line length section.

16,14,5,11

2) An example with 8 replicate lines showing that lines with no observations (i.e., $n_i = 0$) must be included.

8,4,10,0,1,16,4,0

Limitations:

The number of sample sizes (i.e., replications) must not exceed 100 and it should equal the number of line lengths.

Options

Format:

Keyword 1, keyword 2,..., keyword n transect width

The keywords and their meanings are:

1) HELP - provide informative material
2) GRPD - grouped data
3) DEFT - default options
4) DESC - descriptive output
5) PEST - perpendicular distance estimation
6) SEST - sighting distance - angle estimation
7) CUTP - cut points for histograms of perpendicular distances
8) NPOL - a separate density estimate for each replicate (no pooling)

Purpose:

All of the control cards explained so far provide basic data and are input for any analysis. This control card is also input for any analysis and it is designed to explain certain features of the data and to specify the types of analyses to be performed. On the basis of this control card, TRANSECT determines the other control cards it will need to interpret. Each of the keywords will be explained in the order they are presented above.

The keyword HELP indicates to the program that explanatory material about the program should be output. This material includes a narrative at the beginning of the output which discusses the program in general and defines various symbols. It also includes narratives about the descriptive output, about density estimation in general, and about estimation based on perpendicular distances and sighting distances and angles. These latter narratives are given throughout the output and only if the specific analyses are requested. This material is valuable for the first initial runs by a user or for anyone who needs to interpret the output without the aid of this documentation.

The keyword GRPD signifies that the data have been recorded in terms of frequencies within intervals (grouped) rather than exact measurements (ungrouped). If this keyword is not given, then it is assumed that the data are ungrouped. This keyword is only relevant for perpendicular distances because TRANSECT will not presently analyze grouped sighting distances and angles. The input and output for grouped and ungrouped data are very different and many of the options described in the rest of the documentation will not be relevant for grouped data.

The keyword DEFT specifies that the user will accept all of the default options for the analyses of the data. In the default mode, descriptive output and an estimate of density from perpendicular distances by the Fourier series estimator are given. The types of descriptive output for ungrouped data given in the default mode are listed in the discussion of the descriptive output control card. An estimate of density based on a Fourier series was chosen as the default for the following reasons: (1) an analysis based on perpendicular distances requires fewer assumptions than that of sighting distance and angle models and (2) the Fourier series estimator is a robust estimator based upon the criteria in Burnham et al. (in prep.) (e.g., model robust, pooling robust, shape criterion and statistically efficient). If the data are ungrouped and sighting distance and angle are measured, then perpendicular distance will be computed by the program. The only time TRANSECT will not default to estimation based on perpendicular distances is if the sighting distance alone is recorded. In this case, the Hayne estimator is used for the density estimate. However, the assumption of a uniform sighting angle sine cannot be tested so there is no way of determining the validity of the model. We advise against recording only sighting distance. If only one measurement is recorded, then perpendicular distance should be recorded.

If the default mode is not chosen, then TRANSECT requires keywords for each of the analyses to be performed. These keywords are DESC, PEST, SEST, and CUTP. All of these keywords indicate to the program that other control cards will follow to further explain the option. The keyword, DESC, specifies that descriptive output should be given for the data. This keyword only applies if the data are ungrouped. If the data are grouped, then the following is always output: (1) a listing of the cut points and frequencies and (2) a histogram of the data. This descriptive output is given for each analysis so there is no need for an option nor this keyword. The keyword, PEST, is used to invoke estimation of density from perpendicular distances and likewise SEST is used for sighting distances and angles. Again, because TRANSECT will analyze grouped perpendicular distance but not grouped sighting distance-angle data, the keyword, SEST, should not be used if the keyword, GRPD, has been used. The keyword, CUTP, is specifically for ungrouped perpendicular distances. It signals to TRANSECT that the user will define the cut points for histograms of perpendicular distances. These cut points are used for histograms in the descriptive output as well as for plots of the fit of the estimated probability density function f(x) to the shape of the histogram of the perpendicular distance data. If the keyword, CUTP, is not used, then TRANSECT will automatically use three different sets of cut points which it calculates on the basis of the sample size.

The keyword NPOL indicates that there is more than one replicate line and that an analysis should be performed separately for each of the replicates. This option was designed to provide an alternative method of deriving a variance estimate for density. An estimate of density is made for each replicate and an average density estimate and its variance are calculated from these replicates weighted by line length (if the sample size for a replicate is zero (i.e., $n_i = 0$), then the density estimate is appropriately estimated as zero for that replication). This approach avoids making any assumptions about the variance of the sample size, and it does not depend on a theoretical variance estimate of f(0) (see Burnham et al. (in prep.)). However, it does require a substantial sample size to be able to make density estimates for individual replicates. If NPOL is not specified, then the default is to make just one density estimate from the entire data set. This option will be explained further in the data section.

The final piece of information which can be specified on this control card is the transect width. If the data were collected in such a manner that measurements were recorded for any object seen, then the data are considered untruncated and the transect does not have a specified width (i.e., unbounded or infinite width). Whereas, if a specific transect width was chosen such that objects beyond that distance were ignored, then the data are truncated. In the latter case, the width at which the data were truncated should be put on this control card (in the original measurement units). If no width is given, then the data are assumed untruncated.

The transect width can also be used to truncate the data during the analysis. However, this truncation only applies to perpendicular distances. Often the fit of an estimator for perpendicular distances is adversely affected by a few observations which were recorded at extremely long distances. If it is desirable to delete these observations, then a transect width can be specified which is less than the extreme distances. If the data were untruncated originally, then they will be considered truncated in the analysis. There are three points which must be accounted for in truncating the data: (1) only perpendicular distances are deleted so any analysis based on sighting distances and angles will include the whole sample, (2) when truncating the data if the variance for the sample size is being computed empirically, then the sample sizes given on the control card should reflect the number of observations after truncation, and (3) the descriptive output (except for histograms) will also include the whole sample and will not be affected by truncation. We suggest initially running the program with all of the data to get descriptive output and density estimation with the whole sample and then making additional runs, if desired, for density estimation based on perpendicular distances with truncation of the sample and with adjustments made on the sample size control card.

Examples:

1) An example with the default options and untruncated, grouped data.

   DEFT,GRPD

2) An example with ungrouped, truncated data and the default options.

   **DEFT* 16.0*

3) In this example, the user has chosen to specify the descriptive and perpendicular distance estimation options for ungrouped, untruncated data.

   DESC,PEST

Limitations:

None

<u>Data status</u>

Format:

keyword 1,...,keyword n

The keywords and their meanings are:

1) PERP - <u>perp</u>endicular distance
2) SIGH - <u>sigh</u>ting distance
3) ANGL - sighting <u>angl</u>e

Purpose:

There are three different measurements which can be recorded in line transect studies. This control card indicates to the program which of the three measurements it can expect to find in the data section. This control card is only read if the data are ungrouped. As stated previously, the program analyzes grouped perpendicular distances but not grouped sighting distances and angles so there is no need to state the type of data collected.

Based on this control card, the program determines a status code for the data which reflects the types of analyses which can be performed on the data. The determination of the status code takes into account that some of the data can be computed from the recorded measurements. For example, the sighting angle is computed when perpendicular distances and sighting distances are recorded. Table 2 gives the possible status codes and the measurements which are recorded and computed.

Examples:

1) All three measurements recorded.

PERP,SIGH,ANGL

2) Sighting distance and angle recorded.

SIGH,ANGL

Limitations:

None

Descriptive Output

Format:

keyword 1,...,keyword n

The keywords and their meanings are:

1) LIST - data listing and summary statistics
2) CPLT - crossplots of the data
3) HIST - histograms of the data
4) CDFP - cumulative distribution functions plots of the data

Purpose:

This control card is read if the descriptive option is indicated on the options control card and if the data are ungrouped. It specifies the user's choice of types of descriptive output. The LIST, HIST and CPLT options are set if the default model (DEFT) is chosen on the options control card; however, the type and amount of output depends on the data status.

The keyword LIST gives a complete listing of the data which was recorded and computed (after conversion, if any). It also gives summary statistics which includes means, standard deviations, standard errors and correlations.

The keyword CPLT gives crossplots of the measured sighting angle vs. the computed sighting angle and/or the sighting distance vs. the sine of the sighting angle. These plots are only given if the appropriate data has been collected. The latter plot is useful in examining the correlation between the two measurements because in sighting distance-angle models they are usually assumed to

Table 2. The possible data status codes and the data which are recorded (R) and computed (C).

Status code	Perpendicular distance	Sighting distance	Sighting angle
1	R	R	C
2	-	R	-
3	R	-	-
4	R	R	R,C
5	C	R	R
6	R	-*	R
7	-	-	R

*The sighting distance is not given because it is not computable when both perpendicular distance and sighting angle are zero.

be uncorrelated. The former plot is valuable in finding any measurement errors and in determining the reliability of the measurements.

The keyword HIST plots histograms of the data. The histogram of a variable provides information about the underlying distribution of the variable. However, the shape of a histogram is a function of the number of intervals used in its construction. This is especially true with small samples. Therefore, for each variable, three histograms with different numbers of intervals are given to avoid any misleading conclusions about the distribution. For perpendicular distances, the user has the additional option of specifying the cut points (intervals) for the histogram with the CUTP keyword on the options control card.

The keyword CDFP gives plots of the empirical cumulative distribution function (cdf) of the data. Cdf plots do not depend on intervals and thus they provide a much truer picture of the distribution of the variable. However, they are not as easy to interpret as histograms so they have not been included as part of the default options. When either the CDFP or HIST keywords are specified, then the order statistics (measurements sorted in ascending order) for the data are listed. This allows easy construction of other histograms or reconstruction of the cdf plot.

Examples:

1) Cdf plots and crossplots

   CDFP,CPLT

2) All descriptive output.

   LIST,HIST,CPLT,CDFP

Limitations:

None

Perpendicular Distance Estimators

A series of control cards are used to specify the estimators for perpendicular distances. The first control card specifies the number of estimators which will be used and whether the number of parameters will be set or whether starting values for the parameters will be given.

Format:

NPD *keyword 1*

where NPD is the number of perpendicular distance estimators. The keywords and their meanings are:

1) MSET - number of parameters set
2) STRT - starting values given

Only one of the keywords is used because the keyword STRT implies the number of parameters from the number of starting values given. The rest of the control cards in the series give a keyword for the estimator and either the number of parameters or starting values, if any are needed for that particular estimator.

Format:

Estimator keyword NP or $S_1, S_2, \ldots, S_{NP}$

where NP is the number of parameters and S_i is the starting value for the i^{th} parameter. The keywords for the estimators are:

1) FSER - Fourier series (default)
2) EXPS - exponential power series
3) EXPL - exponential polynomial (quadratic)
4) NEXP - negative exponential
5) HNOR - half-normal

Purpose:

This series of control cards is read when the keyword PEST is used on the options control card. It was designed for the following three reasons: (1) to specify estimators other than the default estimator (Fourier series), (2) to set the number of parameters in estimators with variable numbers of parameters and (3) to give starting values to the iterative procedures which calculate the parameter estimates.

There are five estimators based on perpendicular distances which have been included into TRANSECT. Many more estimators could have been included; however, we chose to only represent certain classes of estimators which were developed with a firm theoretical foundation. These classes of estimators are: (1) nonparametric linear estimators (Fourier series), (2) generalized exponential estimators (exponential power series and exponential polynomial) and (3) simple parametric estimators (negative exponential and half-normal). The advantages and disadvantages of each of these estimators are discussed in Burnham et al. (in prep.). Each of these estimators are included in TRANSECT and can be used with grouped or ungrouped and truncated or untruncated perpendicular distance data with the exception of the exponential power series and Fourier series which are not implemented for untruncated data. Instead of using infinity as the width, they use the greatest observed distance. The Fourier series must always have a bounded range and using a bounded range for the exponential power series avoids the difficulty of numerically integrating to infinity (or some large value). This is not a real disadvantage because the fit provided by using the greatest observed distance will not be much different than if infinity had been used.

The Fourier series was chosen as the default estimator because of its robustness properties. It should be a suitable estimator for most situations. However, the other estimators have been provided if a user is interested in other estimators.

The Fourier series is the only estimator of the five provided in TRANSECT which has a variable number of parameters. For ungrouped data, the Fourier series determines the number of parameters to use on the basis of a stopping rule which attempts to provide the best fit to the data. The maximum number of parameters is limited to 6 in TRANSECT as suggested in Burnham et al. (in prep.). If TRANSECT stops at 6 parameters (which it will indicate) and the user is interested in more parameters, then the number of parameters can be set using the MSET keyword and specifying 7 or greater parameters on the estimator control card. For grouped data, the number of parameters is determined by fitting the data with a successively greater number of parameters and calculating a likelihood ratio test statistic for the successive fits. In this case a user may also be interested in examining the fit for a specified number of parameters which can be done with the MSET keyword.

Of the five estimators provided, there are only three which in certain situations will provide a closed form solution. These include: (1) the Fourier series with ungrouped data, (2) the negative exponential with ungrouped, untruncated data and (3) the half-normal with ungrouped, untruncated data. For the other estimators and other situations (e.g. truncated and grouped), the parameter estimates are obtained through an iterative procedure which maximizes a function in one or more dimensions. This iterative procedure which is described in Burnham et al. (in prep.) can be sensitive to the starting values (i.e. initial parameter estimates) that it is given. The starting values are determined by the program from moment estimators which assume the data are untruncated and ungrouped. The iterative procedure is not guaranteed to converge to the correct answer, although it will most of the time. Messages about the iteration procedure are given in the output which help in determining if the procedure is converging correctly. If the convergence is questionable or if the procedure does not converge, then the user can examine the fit to the data and try different starting values to determine if the best fit was obtained.

Examples:

1) If a user was interested in other estimators, then the following might be used. The first card specifies NPD as 5, so 5 estimator keyword control cards follow.

5
FSER
EXPS
EXPL
NEXP
HNOR

2) If a user had convergence problems with the Fourier series estimator with grouped data, then the following might be used to examine the fit for a series of parameters and the likelihood ratios could be calculated by hand.

```
3. *STRT*
*FSER* 0.01
*FSER* 0.01,0.005
*FSER* 0.0,0.0,0.0
```

Limitations:

The number of estimators (NPD) cannot exceed 10 and no more than 15 parameters can be specified.

<u>Sighting Distance-Angle Estimators</u>

Format:

keyword 1,...,keyword n

The keywords and their meanings are:

1) HAYN - <u>Hayn</u>e estimator

2) GHYN - <u>g</u>eneralized <u>Hay</u>ne estimator

3) MHYN - <u>m</u>odified <u>Hay</u>ne estimator

Purpose:

This card is read whenever the keyword, SEST, is specified on the options control card. It is used to indicate which of the sighting distance-angle estimators should be used in analyzing the data. These estimators are straightforward and do not require any of the necessary complexities with perpendicular distance estimators.

Examples:

1) Generalized Hayne and modified Hayne,

```
*GHYN,MHYN*
```

2) All three estimators,

```
*MHYN,HAYN,GHYN*
```

Limitations:

None

Cut Points for Perpendicular Distance

This is a series of control cards which gives from 1 to 5 sets of cut points for histograms of perpendicular distances. The first control card simply gives the number of cut point sets which will be input.

Format:

NCUT

The rest of the control cards each give a set of cut points.

Format:

$c_1, c_2, c_3, \ldots, c_{kcut}$

The above card should be repeated NCUT times.

Purpose:

This series of control cards allows the user to define the cut points (interval end points) by which the histograms of perpendicular distances are constructed for ungrouped data only. These cut points are used for the histograms in the descriptive output and also in representing the fit of the estimator to the data in plot form. TRANSECT calculates three sets of cut points which usually provide a fairly representative picture of the distribution if the user does not define the cut points. However, there could be various reasons why a user may want to define the cut points. For instance, in illustrating the fit of the estimator to the data, often an extreme outlier will make the rest of the plot unbalanced or unreadable. In these cases the final cut point C_{kcut} can be defined to be less than the largest measurement on the specified width. Thus the user can examine all of the histogram or only the first part with any interval sizes.

Examples:

1) An example illustrating that the intervals need not be equal.

```
2
5,10,20,35,50,75,100
10,40,60,80,100
```

2) An example illustrating that the final cut points need not be the same.

```
3
5,10,20,45
5,10,20,45,60,80
5,10,20,45,60,90,100
```

Limitations:

The number of cut point sets (NCUT) cannot exceed 5. The number of cut points within a set (kcut) cannot exceed 20. Additional cut point sets could be specified in another program execution.

Data

The format for the data section is quite different for grouped and ungrouped data so they will be described separately. Each section will discuss how the basic data is entered and how replicate lines are handled.

Grouped Data

The data input for the analysis of grouped data consists simply of three cards. These cards provide a label for the particular data set, the cut points (end points) for the intervals of perpendicular distance and the frequencies (number observed) within the intervals. The first card, the label card, is simply a label which identifies the data. This label is distinct from the general label because it can be used to describe subsets of the data (e.g., replicate lines). The first four characters of this label cannot be END. or it will be confused with the flag card. The label is put in the first 30 columns of the card image. The second card, for cut points, is input in the same format as the control cards.

Format:

$c_1, c_2, c_3, \ldots, c_{kcut}$

where kcut is less than or equal to 20. The first cut point is the right end point of the first interval. It is assumed that the first interval starts at the origin. The last cut point, c_{kcut}, should be the transect width if the data are truncated. If the data are untruncated, then c_{kcut} is actually infinity. It is treated as such in the analysis by all of the estimators except for the Fourier series and the exponential power series which need a value for c_{kcut}. Also, because it is not possible to represent infinity in descriptive output, a value for c_{kcut} should be given within which all the observations were made. The third control card, for frequencies, also takes the format of the control cards.

Format:

$f_1, f_2, f_3, \ldots, f_{kcut}$

where kcut is less than or equal to 20. These are simply the numbers observed in each of the intervals. It is important that the number of frequencies and cut points agree.

The above description of data input applies if the data from the replicate lines in the survey are pooled for one estimate of density. If the NPOL keyword is set on the options control card, then the data should be presented for each replicate. If there are r replicate lines, then a series of three cards should be input for each replicate which has a nonzero number of observations.

Examples:

1) An example with unequal intervals.

RABBIT DATA
1,2,5,10,20
28,16,17,14,17

2) An example in which the NPOL keyword was used and estimates of density are made for three replicates.

RABBIT DATA, LINE 1
1,2,5,10,20
8,6,7,4,3
RABBIT DATA, LINE 2
1,2,5,10,20
7,5,3,5,2
RABBIT DATA, LINE 3
1,2,5,10,20
2,4,3,7,2

Ungrouped Data

An individual card image is used for each observation with ungrouped data. The format for these cards is more restrictive than the control cards. The information on this card includes an identifying label in the first 30 columns, the perpendicular distance in columns 31 to 35, the sighting distance in columns 36 to 40, and the sighting angle in columns 41 to 45. A decimal point should be included with the distances and angles. If any of the measurements are not recorded, then the field should be left blank. For example, if only perpendicular distance was collected, then columns 36-45 should be left blank.

The label on the card identifies the observation as part of a subset of the data (the first four characters must not be END.). The only subsets which are of concern to the program are those which identify them as being from different replicate lines. If the keyword, NPOL, is not used (the normal situation), then the program does not use the labels except to use the first label it encounters as a label in the output. If the keyword, NPOL, is used, then TRANSECT expects to find r data sets for the r replicate lines except for those replicates in which no observations were made. TRANSECT identifies these replicates on the basis of the label. Whenever it comes to a card with a different label, it will create a new data set. Therefore, all of the observations within a replicate must be contiguous within the file.

It may be valuable to have other ancillary data in the label as well. For example, this might be the sex of the animal or whether it is a juvenile or adult. Then estimates of density can be made for juveniles and adults separately or together. TRANSECT does not have any mechanism for sorting subsets of data so they must be sorted manually. If ancillary information is used in the label, then the program can be made to ignore this information with a card which is put at the beginning of the data that acts like an eraser. For every column in the label to be ignored, a '1' should be put in the corresponding column on the eraser card. These card columns will be ignored when TRANSECT compares the labels to determine the label for the output or to identify different replicates when NPOL is used. For example, if the label was given as RABBIT DATA in the first 11 columns and the line number in columns 13 and 14 and the age (either juvenile or adult) in columns 15 to 25, then an eraser card with ones in columns 15 to 25 would allow the data to be analyzed separately for each line.

Examples:

1) In this example only perpendicular distance was collected and only one estimate of density is being calculated (i.e., NPOL not used). The eraser card has been used so the output will not show the line number and age. The sample size is very small for illustrative purposes.

```
             1111111111111111111
RABBIT DATA, LINE 1, JUVENILE 10.0
RABBIT DATA, LINE 3, ADULT     5.2
RABBIT DATA, LINE 5, ADULT     0.5
```

2) In this example both sighting distance and sighting angle were recorded and estimates of density are being calculated for each replicate line (i.e., NPOL used). Again the sample sizes would be normally much larger.

```
                     1111111111
RABBIT DATA, LINE 2, JUVENILE     12.0  45.
RABBIT DATA, LINE 2, JUVENILE      5.0  20.
RABBIT DATA, LINE 3, ADULT         2.0  72.
RABBIT DATA, LINE 3, JUVENILE     25.0  18.
```

DETAILED EXAMPLES

This section will attempt to tie together what has been presented by giving some examples using all of the control cards. Four examples will be given showing the typical commands for an analysis of grouped and ungrouped data.

The first example is a situation in which a survey is being conducted on marine mammals from the air and exact measurements of distance cannot be made so the data are grouped. The measurements of distance and line length are in nautical miles and the area is in square nautical miles. The survey consists of 4 replicate lines and a total of 40 observations. The data are truncated at one nautical mile. Only one estimate of density will be made from the 40 observations, but the observations were identified by their line number so a variance estimate of the sample size can be calculated. Using the default estimator (Fourier series), the following set of instructions and data might be used:

```
*MARINE MAMMAL SURVEY FOR 1979*
*DISTANCE IS IN .NAUTICAL MILES.*
*LINE LENGTH IS IN .NAUTICAL MILES.*
*AREA IS EXPRESSED IN .SQUARE NAUTICAL MILES.*
1000,562,989,1100
12,2,17,9
*DEFT,GRPD* 1.0
MARINE MAMMALS, ALL LINES
.1,.2,.3,.4,.5,.6,.7,.8,.9,1.0
10,9,5,5,4,3,0,2,1,1
```

The second example is a situation in which an aerial survey of African game animals is being made. Again, accurate measurements could not be made, so the perpendicular distances have been grouped into intervals. The measurements of distance and line length are in miles and it is desired to express the density estimate in hectares. The survey consists of 6 replicate lines and a total of 150 observations. The data were collected in the field at intervals of a quarter of a mile until 1 mile and then at half mile intervals out to 5 miles. All of the observations greater than 5 miles were also collected in one interval, so the data are untruncated. The field personnel felt that all the observations had been made within 7 miles. Therefore, 7 was used as the final cut point for the descriptive output and for the Fourier series which cannot extend to infinity. It was decided to examine both the Fourier series and exponential polynomial estimators. Also, it was decided to make a run which made one density estimate (pooled) and a run which made density estimates for each replicate (not pooled). On the basis of these decisions, the following two sets of control cards and data were used:

```
*AFRICAN GAME SURVEY FOR 1979 - POOLED*
*DISTANCE IS IN MILES CONVERTED TO .KILOMETERS.* 1.604
*LINE LENGTH IS IN MILES CONVERTED TO .KILOMETERS.* 1.604
*AREA IS CONVERTED FROM SQUARE KILOMETERS TO .HECTARES.* 0.01
100,20,195,235,150,175
15,40,0,35,42,18
*PEST,GRPD*
2.
*FSER*
*EXPL*
AFRICAN GAME ANIMALS, ALL LINES
```

```
.25,.5,.75,1,1.5,2,2.5,3,3.5,4,4.5,5,7,
18,15,5,9,23,20,18,12,9,6,5,3,7
END.
*AFRICAN GAME SURVEY FOR 1979 - NOT POOLED*
*DISTANCE IS IN MILES CONVERTED TO .KILOMETERS.* 1.604
*LINE LENGTH IS IN MILES CONVERTED TO .KILOMETERS.* 1.604
*AREA IS CONVERTED FROM SQUARE KILOMETERS TO .HECTARES.* 0.01
100,20,195,235,150,175
15,40,0,35,42,18
*PEST,GRPD*
2.
*FSER*
*EXPL*
AFRICAN GAME ANIMALS, LINE 1
.25,.5,.75,1.1.5,2,2.5,3,3.5,4,4.5,5,7
3,1,2,0,4,2,0,1,0,1,1,0,0
AFRICAN GAME ANIMALS, LINE 2
.25,.5,.75,1,1.5,2,2.5,3,3.5,4,4.5,5,7
5,6,0,3,8,4,3,3,2,1,1,1,3
AFRICAN GAME ANIMALS, LINE 4
.25,.5,.75,1,1.5,2,2.5,3,3.5,4,4.5,5,7
3,4,1,2,6,6,7,4,1,0,0,0,1
AFRICAN GAME ANIMALS, LINE 5
.25,.5,.75,1,1.5,2,2.5,3,3.5,4,4.5,5,7
5,3,0,4,3,6,6,3,3,2,3,2,2
AFRICAN GAME ANIMALS, LINE 6
.25,.5,.75,1,1.5,2,2.5,3,3.5,4,4.5,5,7
2,1,2,0,2,2,2,1,3,2,0,0,1
```

The third example is a situation in which a survey is being made for rabbits and exact measurements can be recorded. All three measurements of perpendicular distance, sighting distance and sighting angle are recorded to allow for error checking. The distance measurements are recorded in feet and the line length in miles but metric conversions are needed. The survey consists of 7 replicate lines each of which are 2.5 miles long. The data include 52 observations and no width was specified for the transect (untruncated). Again using the default estimator, the following series of control cards could be used:

```
*ANALYSIS OF RABBIT DATA*
*DISTANCE IS IN FEET CONVERTED TO .METERS.* .3048
*LINE LENGTH IS IN MILES CONVERTED TO .KILOMETERS.* 1.604
*AREA IS EXPRESSED IN .HECTARES.* 10.0
2.5,2.5,2.5,2.5,2.5,2.5,2.5
15,2,0,3,12,10,10
*DEFT*
*PERP,SIGH,ANGL*
                 1111111111111111
RABBIT DATA, LINE 1, JUVENILE 10.0 25.0 42.0
```

(continued for all 52 observations)

The fourth example will demonstrate a control card setup for an analysis of the third example for density estimation based on perpendicular distances in which the number of terms in the Fourier series is set at values greater than 6 and other estimators are examined.

```
*ANALYSIS OF RABBIT DATA*
*DISTANCE IS  IN FEET CONVERTED TO .METERS.* .3048
*LINE LENGTH IS IN MILES CONVERTED TO .KILOMETERS.* 1.604
*AREA IS EXPRESSED IN .HECTARES.* 10.0
2.5,2.5,2.5,2.5,2.5,2.5,2.5
15,2,0,3,12,10,10
*PEST*
*PERP,SIGH,ANGL*
5 *MSET*
*FSER* 7
*FSER* 8
*FSER* 9
*EXPS*
*EXPL*
          1111111111111111111
RABBIT DATA, LINE 1, JUVENILE  1.0 25.0 42.0
```

(continued for all 52 observations)

OVERVIEW OF PROGRAM OUTPUT

There are many examples of the output from TRANSECT given in Burnham et al. (in prep.) in the ILLUSTRATED EXAMPLES section. It is our purpose here to give a brief overview of the organization of the output. The amount of output for any execution of the program will vary depending on the options chosen. In this discussion, it will be assumed that all of the output will be given.

There are three major sections to the output. The first section is very short and is entitled INTERPRETATION OF CONTROL CARDS. This section reiterates how TRANSECT interpreted the control cards. If there were any errors in the control cards, they are output on this page. The next section is for descriptive output. For grouped data, this amounts only to a listing of the cut points and frequencies and a histogram of the data. For ungrouped data, the descriptive output begins with a data listing and summary statistics and crossplots. Then for each variable recorded and computed, there is a section for illustration of the distribution of the variables. This section gives the sorted measurements, histograms and cdf plots. The final section in the output is for density estimation. For each perpendicular distance estimator, the following is output: (1) information about the procedure which lead to the final parameter estimates (e.g., convergence of the iteration procedure, stopping rule and likelihood ratio tests for Fourier series), (2) a single page of output giving estimates of the model parameters, f(0), and density and their standard errors, coefficients of variation, interval estimates and information on confidence interval construction and (3) goodness of fit testing including chi-square tests with a plot illustrating the fit of the estimator to the data. For each sighting distance-angle estimator, a page of output is given which is nearly equivalent to 2 above which is then followed by goodness of fit tests if they are possible. If more than one estimator is computed for perpendicular distances and/or sighting distances-angles, then they are summarized in a page of output entitled SUMMARY OF ESTIMATION RESULTS.

If the NPOL keyword is used, then the second and third sections of output are repeated for each replicate line. At the end of these replicates, a page of output entitled AVERAGE DENSITY ESTIMATION FROM REPLICATIONS is given. This section calculates an average density estimate (weighted by line length) and its standard error, coefficients of variation and confidence intervals for each estimator chosen.

PROGRAM DETAILS

TRANSECT consists of a main routine and 57 subroutines which comprise approximately 7000 lines of code. It requires around 150K to run on a CDC 6600 without any overlay structure. This could be reduced considerably with the use of overlays if necessary. A typical run consisting of descriptive output and density estimation for perpendicular distances and sighting distances and angles and several estimators for ungrouped data costs in $1-$5 range on the Burroughs at Utah State University.

The program is written in ANSI FORTRAN IV with only a few known exceptions. These exceptions are documented with comment statements. The program has been complied on an ANSI compiler and it has been compiled and executed on Burroughs, CDC and DEC machines. It should run on most major computer systems with only very minor changes to accommodate input and output structures.

A magnetic tape with the source deck and the set of examples used in Burnham et al. (in prep.) is available from

SHARE Program Library Agency
P. O. Box 12076
Research Triangle Park, NC 27709

at a cost of approximately $40. Specifications for the tape (e.g., 7 or 9 track, 800 or 1600 bpi, etc.) should be given.

LITERATURE CITED

Burnham, K. P., D. R. Anderson and J. L. Laake. (In prep.) Estimation of density for line transect sampling of biological populations.