

A STATISTICAL ANALYSIS OF KINGBIRD NEST PREDATION

By

B.M. Sarath Gamini Banneheka

B.Sc.(Mathematics Special), University of Sri Jayewardenepura, Sri Lanka, 1986

M.Sc.(Mathematics), Kings College London, University of London, UK, 1988

THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
in the Department
of
Mathematics and Statistics

© B.M. Sarath Gamini Banneheka 1991

Simon Fraser University

September 1991

All rights reserved. This work may not be
reproduced in whole or in part, by photocopy
or other means, without permission of the author.

APPROVAL

Name: B.M.S.G. Banneheka

Degree: Master of Science

Title of thesis: A Statistical Analysis of Kingbird Nest Predation

Examining committee:

Chairman: Dr. A. H. Lachlan

Dr. T. Swartz
Senior Supervisor

Dr. D. Eaves

Dr. C. Dean

Dr. M. A. Stephens
External Examiner

Date Approved: October 10, 1991

PARTIAL COPYRIGHT LICENSE

I hereby grant to Simon Fraser University the right to lend my thesis, project or extended essay (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

Title of Thesis/Project/Extended Essay

A STATISTICAL ANALYSIS OF KINGBIRD NEST
PREDATION

Author: _____
(signature)

B.M.S.G. BANNEHEKA
(name)

28/10/91
(date)

ABSTRACT

The logistic regression model and the Cox proportional hazard model are employed to identify the factors influential on kingbird nest predation. The high overhead visibility of the nest and the high percentage of water and trees around the nest are found to decrease the probability of predation.

ACKNOWLEDGEMENTS

I am extremely grateful to my senior supervisor, Dr. Tim Swartz for his stimulating ideas, constant encouragement, invaluable guidance and above all patience, throughout the successful completion of this work.

I would also like to extend my sincere appreciation and gratitude to my committee members Dr. D. Eaves, Dr. C. Dean and Dr. M. A. Stephens and all others of the Department of Mathematics and Statistics who have undoubtedly been extremely helpful to me in numerous ways throughout my course of study. Also I would like to thankfully acknowledge Ms. Joanne Siderius, Department of Biological Sciences, for providing me with these data sets and Mr. Francois Bellavance, Director, Statistical Consulting Service of the Department of Mathematics and Statistics for helping initiate this work.

At last but not the least, I would like to extend my thanks to all my colleagues.

DEDICATION

To my parents

TABLE OF CONTENTS

Approval	ii
Abstract	iii
Acknowledgements	iv
Dedication	v
List of Tables	vii
List of Figures	ix
1. INTRODUCTION	1
2. THE DATA SET	5
2.1. ENVIRONMENT	4
2.2. BEHAVIOUR	10
3. THE ANALYSIS	13
3.1. Preliminary analysis of ENVIRONMENT	13
3.2. Logistic analysis of ENVIRONMENT	20
3.3. Summary of analysis of ENVIRONMENT	32
3.4. Preliminary analysis of BEHAVIOUR	33
3.5. Logistic analysis of BEHAVIOUR	36
3.6. Survival analysis of BEHAVIOUR	39
3.7. Summary of analysis of BEHAVIOUR	42
4. CONCLUDING REMARKS	44
APPENDIX 1	46
APPENDIX 2	56
REFERENCES	60

LIST OF TABLES

Table	Page
3.1.1 Total number of observations and number of predations in 1988, 1989 and 1990.	13
3.1.2 F-values and p-values from one-way analysis of variance corresponding to 15 variables in reduced ENVIRONMENT.	20
3.2.1 Deviance decreases and associated p-values in logistic regression with explanatory variables TREE, WAT and OVER.	24
3.2.2 Estimated parameters and estimated standard deviations from logistic model with explanatory variables TREE, WAT and OVER	25
3.2.3 Summary of 27 logistic regression fittings based on 92 observations.	29
3.2.4 Scaled deviances and outliers associated with models 26 and 27.	29
3.2.5 Estimated parameters and corresponding p-values in models 26 and 27.	30
3.2.6 Estimated parameters and corresponding p-values in refit models 26 and 27.	31
3.2.7 Fitted probabilities of predation in 8 covariate classes.	31
3.5.1 Days to predation and newly defined PRED.	37
3.5.2 Deviance decreases and corresponding p-values associated with logistic regression on BEHAVIOUR.	38
3.6.1 Survival times and censoring indicators used in survival analysis on BEHAVIOUR.	41
3.6.2 Estimated coefficients and corresponding p-values associated with variables in survival analysis of BEHAVIOUR.	42
5.1.1 ENVIRONMENT.	46
5.1.2 Created variables for ENVIRONMENT.	52

5.1.3	BEHAVIOUR.	54
5.1.4	Dates of visits for 9 preyed upon nests in BEHAVIOUR.	55
5.2.1	Correlation matrix of 12 visibility variables associated with north, south, east and west.	56
5.2.2	Correlation matrix of OVER, VN, VS, VE and VW.	56
5.2.3	Correlation matrix of 15 variables in reduced ENVIRONMENT.	57
5.2.4	Description of 15 variables in reduced ENVIRONMENT.	58
5.2.5	Correlation matrix of variables in BEHAVIOUR.	59

LIST OF FIGURES

Figure	Page
3.1.1 Histogram of OVER for all nests	15
3.1.2 Relative frequency of OVER	16
3.1.3 Relative frequencies of VN, VS, VE and VW	17
3.1.4 Histogram of WAT for all nests	18
3.1.5 Histogram of WAT for preyed upon nests	18
3.1.6 Histogram of WAT for fledged nests	19
3.1.7 Histogram of TREE for all nests.	19
3.2.1 Plot of OVER vs NEST	25
3.2.2 Plot of WAT vs NEST	26
3.2.3 Plot of TREE vs NEST	26
3.2.4 Estimated probability of predation vs OVER for 8 covariate classes	32
3.4.1 VIS vs INTE for preyed upon nests	34
3.4.2 VIS vs INTE for fledged nests	34
3.4.3 BOTH vs INTE for preyed upon nests	35
3.4.4 BOTH vs INTE for fledged nests	35

Chapter 1

INTRODUCTION

The eastern kingbird (*Tyrannus Tyrannus*) is a defiant and fearless small bird and is black and white in colour. It can be found in North and South America. It starts nesting in May, lays eggs in mid June through early July and the young fledge by August. The eggs take 14 days to hatch. After hatching, the mother bird provides heat to the young by sitting in the nest for approximately six days. It takes 8-12 more days before the young can fledge.

During the 32 days from the laying of eggs,

- (a) the eggs or the young can be preyed upon.
- (b) the young can die due to lack of food, diseases etc.

Predation can easily be distinguished from other causes of death because afterwards one can see the scattered nest, egg shells or wings of the young on the ground around the nest. The most common nest predators are crows, ravens, magpies, kestrels and hawks. In addition, there are ground predators like squirrels, racoons and snakes. We will refer to the nests in which the young were not preyed upon as the 'fledged nests'.

Ms. Joanne Siderius is a graduate student in Biological Sciences, Simon Fraser University (SFU) and she is interested in studying the factors which may be influential on the predation of kingbirds. She has collected some data on kingbirds in the Creston Valley Wildlife Management Area, Creston-Courtenay mountains in British Columbia. These data were collected during 1989-1990. In the first two years she collected data on the visibility and the environment surrounding the nests. In addition, in 1990, Ms. Siderius collected some information on the behaviour of the parent birds during their nest vigilance. We name the data set containing environmental factors 'ENVIRONMENT' and the data set containing behavioural factors 'BEHAVIOUR'.

Ms. Siderius came to the Statistical Consulting Service in the Department of Mathematics and Statistics, SFU with these data sets, when I was taking Statistical Consulting II as a course requirement for my M.Sc. degree. She wanted us to help her

recognize the factors which are influential on kingbird nest predation by means of more sophisticated statistical methods than those which she had previously attempted. Subsequently, Dr. Tim Swartz, Department of Mathematics and Statistics, who is my graduate supervisor, Mr. Francois Bellavance, the director of the Statistical Consulting Service and I had a discussion with Ms. Siderius concerning her data set. This discussion focused on her data collection methods and the problems of interest.

According to Ms. Siderius, the kingbird is quite different from other birds in its nesting pattern and behaviour. She says that most birds hide their nests. But as we shall see later, a considerable portion of the nests of kingbirds are not hidden. In contrast to most birds, whose concern is restricted to the immediate vicinity of their nest, the kingbird's attention reaches far out. Its perch always commands a good view of the surrounding area; the kingbird is always on the watch for an enemy. It seems to consider any big bird its enemy; it does not wait for one to come near but, assuming the offensive, dashes out at crows and big hawks; size seems to make no difference to the kingbird and it practically always wins. The kingbird is often noisier than other birds. It has a high sibilant, jerky voice. Sometimes it is noisy close to the nest while sometimes it is noisy far from the nest.

Ms. Siderius raised the following questions :

- (a) How important is nest visibility in attracting predators? Are birds in visible nests more liable to lose their nests to predators than those who live in less visible nests?
- (b) Do noisy birds attract predators to their nests? Do birds which call close to the nest lose their nests to predators more often than birds which call less often or call further from the nest?
- (c) Is nest vigilance effective in preventing predation? Do birds that spend more time at the nest lose their offspring to predators less often than birds that spend less time at the nest?
- (d) In general, what are the environmental and behavioural factors, if any, upon which predation depends?

Chapter 2 provides a description of the data and the data collection methods together with a running commentary on the experimental design. We include the data set in table 5.1.1 of appendix 1. The measurements of a few variables were transformed into other scales and a few were rounded off considering the nature of the measurements. A description of these changes is also presented in chapter 2.

We present our analysis in chapter 3. In any statistical analysis, a preliminary exploration of the data set using graphical methods and other statistical tools plays an important role. In our attempt to address the problems proposed by Ms. Siderius, we do a preliminary analysis of the 2 data sets. The preliminary analyses of ENVIRONMENT and BEHAVIOUR are presented in sections 3.1 and 3.4 of chapter 3.

In an attempt to provide a more detailed analysis of kingbird nest predation we note that the fledge or the occurrence of predation can be regarded as the outcome of a binary variable. Therefore, from a statistical point of view we are interested in the relationship between this binary response variable and certain explanatory variables. Generalized linear regression methods like logistic or probit regression analysis are widely used in this type of situation as a way of modelling relationships. We report our findings using logistic regression analysis on ENVIRONMENT and BEHAVIOUR in sections 3.2 and 3.5 of chapter 3.

In a further attempt to provide a more detailed analysis of kingbird nest predation we are reminded that binary variables provide information only about whether or not an event has occurred. However, in some situations the time taken for an event to occur may carry additional information about the event. If the event is a death or an expiration of an object, the time is generally referred to as a 'survival time'. Survival analysis techniques are employed to extract information concerning covariates using the survival time. In a subsequent discussion with Ms. Siderius we came to realize that she had recorded the dates of her visits from which we could calculate an approximate length of time prior to predation. Even though our problem does not fit exactly into the usual framework of survival analysis, it seems that we can still use these techniques to extract information, if any, from our data set. We include our attempt to answer Ms. Siderius's questions by means of survival analysis in section 3.6 of chapter 3.

Sections 3.3 and 3.7 of chapter 3 provide summaries of the analyses on ENVIRONMENT and BEHAVIOUR respectively. Finally, chapter 4 offers some concluding remarks.

Chapter 2

THE DATA SET

Ms. Siderius started collecting data in 1988 at the Creston Valley Wildlife Management Area in the Creston-Courtenay mountains in British Columbia. She started visiting the site (by this we mean the area where the data collection took place) when the birds started building their nests. As soon as she noticed that a nest was being built, if the nest was within climbing height, that is less than about 10 meters, a number, for identification purposes, was assigned to the nest. She visited the site every three days. In this way she could tell approximately the date on which the first egg was laid and also the approximate date that the nest was preyed upon if that was the case. During the breeding seasons of kingbirds, in 1988 and 1989, she collected data on the visibility and the environment surrounding the nests. These variables are time independent. In 1990, on a few more nests, in addition to the variables mentioned above, she gathered some information about the behaviour of the parent birds during their parenting period. These variables are time dependent.

Therefore we have some nests on which only the time independent variables were observed and a few nests on which both the time independent variables and the time dependent behavioural variables were observed. We therefore have two sets of data. The first, which will be called 'ENVIRONMENT' contains the time independent visibility and environmental measurements and the second, which will be called 'BEHAVIOUR' contains the time dependent behavioural measurements. ENVIRONMENT contains 95 observations on 95 nests while BEHAVIOUR contains 48 observations on 23 nests.

We now describe the variables contained in the two data files.

2.1 ENVIRONMENT

1. NEST Each nest was given a number for identification purposes. In this data set we have nest numbers up to 103. Even though Ms. Siderius gave a number as soon as she found a nest, 8 nests were later found unused

for the laying of eggs. Therefore we have only 95 nests with at least one explanatory variable.

2. PRED This is a bernoulli response variable.
1 - if the nest was preyed upon
0 - if the young had fledged.

All the measurements of variables 3-21 were taken after the young had fledged or the nest had been preyed upon. The reason for this was that Ms. Siderius had to be very close to the nest to obtain these measurements and this might have caused unwanted problems to Ms. Siderius as well as to the birds.

3. OVER This is supposed to be an estimate of the visibility of the nest from above the nest. Ms. Siderius looked up from directly below the nest and estimated the percentage of sky that could be seen through a one meter diameter circle, centred at the nest. It is assumed that if the nest is covered from above, then the percentage of the sky that one can see from below the nest is equivalent to the percentage of visibility of the nest for a predator flying over the nest.

4. - 7.

PN, PS, PE, PW

These variables are the estimated percentage visibility of the nest at 15 meters distance from the nest in four directions north, south, east and west respectively. Ms. Siderius walked towards the nest from a distance and observed that around 15 meters she could see the nest. Therefore she selected 15 meters as her fixed distance to estimate the visibility.

Clearly variables 3-7 are prone to measurement error as there was no measurement device used. We mention that there is no reason to suspect that any of the variables 4-7 are more accurate than any other. This observation is used in chapter 3.

After the young had fledged or had been preyed upon, a box with 3 black stripes and 4 white stripes on each of four sides was placed at the nest location. Ms. Siderius then estimated the visible number of black stripes and the visible number of white stripes. These measurements were also taken at 15 meters distance from the nest.

As kingbirds are black and white in colour, the use of above technique was intended to mimic the situation when there were birds in the nest. For instance, if she could see 2 black stripes and 1 1/2 white stripes, she recorded them as 20 black stripes and 15 white stripes. In her data set she had values like 3, 8 and 28 which meant that she could see 0.3, 0.8 and 2.8 stripes. These measurements are again not very reliable. As these are also visibility measurements like variables 4-7, we transformed these values into percentages to form the following variables.

8. - 11.

BN, BS, BE, BW

These are the estimated percentage visibilities of black stripes from the north, south, east and west respectively.

12 - 15.

WN, WS, WE, WW

These are the estimated percentage visibilities of white stripes from the north, south, east and west respectively.

Ms. Siderius noticed that kingbirds start chasing predators or any other birds when they are as close as 100 meters to the nest and they chase them as far as 300 meters from the nest. She thinks that the distance a kingbird can see may be an important predatory variable since by attacking a predator at a great distance the predator may not notice the nest. Therefore she is interested in seeing if it makes a difference if the parent birds can see far from the nest. She therefore looked far to each of four sides at the nest level and estimated the distance that she could see. Ms. Siderius says that when she could see very far, then she would record it as 1000 meters. Unfortunately, we have missing values for approximately half of the nests as she did not collect data on these variables in 1989.

16-19.

DN, DS, DE, DW

These are the distances from the nest in meters that Ms. Siderius could see, in north, south, east and west directions respectively.

Sometimes Ms. Siderius could not reach the nest because there was a bee hive close to the nest or the tree was thorny or the nest was on a tiny branch over the water. In some cases such as these we have missing values for variables 8-19.

Ms. Siderius thinks that kingbirds can see farther in higher nests. Also if ground predators were involved, higher nests would be more secure. She therefore measured the nest height when she could reach the nest level. She also measured the nest tree height because she was interested in knowing if kingbirds preferred nesting in particular portions of the tree like the upper 1/4th or the middle 2/3rd of the tree. We have many missing values for tree height corresponding to trees which were too tall for her to climb. Some missing values may also be due to the other reasons mentioned above.

20. NH This measurement is the nest height in inches. Ms. Siderius chose only the nests which were within about 10 meters from the ground for her study.

21. TH This measurement is the nest tree height in inches. Ms. Siderius did not collect this information in 1989 and hence we have 42 missing values for this variable.

22. NTRE This is the distance in feet to the nearest tree exceeding 5 cm in diameter and exceeding 15 meters in height. This measurement gives some idea of the distance between the nest tree and other trees.

Most of the time there was a number of thin trees surrounding the nest tree. Ms. Siderius calls these stems. They may provide a cover to the nest. It is also a measure of the vegetation around the nest.

23. STEM This is the number of stems within 5 meters of the nest which are greater than 1 cm in diameter and taller than the nest height.

24. NPER This is the distance to the nearest perch measured in meters. A perch should be taller than the nest so that the predatory birds can see the nest when they perch there. Ms. Siderius says that the perch may or may not be a tree. It could even be a telephone pole.

Ms. Siderius says that kingbirds almost always feed within 100 meters from the nest and that most of the time they remain nearby. These birds feed on small insects in the long grass and larger insects like dragon-flies on the water. Therefore the availability of food depends on the environment surrounding the nest. If there is more water close to the nest, they can feed on larger insects and therefore they may not have to leave the nest as often. Furthermore, the type of predators around the nest may depend on the habitat. Collecting data on the following variables, getting a feeling for the environment surrounding the nest and thereby investigating their effects on nest predation was intended.

25 - 29.

WAT, FIELD, SHRUB, MARSH, TREE.

These are the estimated percentages of water, field, shrub, marsh and trees within 100 meters of the nest. When these do not add up to 100, Ms. Siderius explains that it may be due to roads, rocks, etc.

As mentioned before, if there is water close to the nest it may make it easier for kingbirds to find their food. If the nest is directly over the water the nest may be protected against ground predators. Furthermore, the water tends to moderate the influence of temperature and therefore the female bird may not require as much effort in incubation to keep a constant temperature in the nest. Ms. Siderius provided measurements of distance to the water from the nest.

30. DWAT This is the distance in meters to the water from the nest.

The variable WAT is the percentage of water within 100 meters of the nest. Therefore $WAT=0$ means there was no water within 100 meters. But we have values like 0, 15, 25, 30 and 70 for DWAT corresponding to nests which have $WAT=0$. There is clearly something wrong with this data. We also have $DWAT=800$ corresponding to $WAT=25$, which is also contradictory. We have 22 missing values of the variable DWAT. Considering the unreliability and the missing values, we decided to ignore this variable in the analysis.

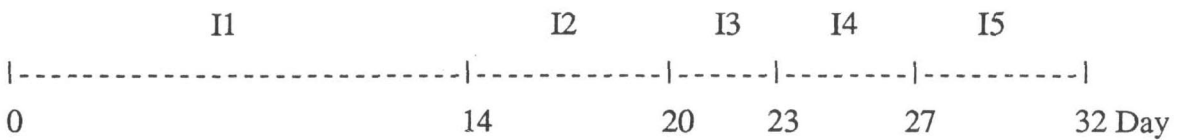
31. YEAR This is the year of data collection. 1-1988, 2-1989, 3-1990.

The data set ENVIRONMENT appears in table 5.1.1 of appendix 1.

2.2 BEHAVIOUR

In 1990, in addition to visiting the site every three days to see if the young had been preyed upon Ms. Siderius monitored each nest several times during the period after hatching, and collected information on the behaviour of the parent birds and some other factors.

Ms. Siderius divided the period of 32 days between the laying of eggs and fledging into 5 intervals as follows.



She believes that we can assume that the behaviour of the parent birds is more or less constant within these intervals. She monitored each nest for 90 minutes during each or some of these intervals in order to estimate the behavioural measures. These observational periods were chosen to be either between 5 a.m - 11 a.m or between 6 p.m - 9 p.m. Ms. Siderius claims that these are the most active times of the day for kingbirds. However, she had never witnessed a predation. This would suggest that she did not monitor the nests at the proper times or no predation during active times.

1. NEST Nest number.
2. INTE This is the interval in which the data was collected (I1, I2, I3, I4, and I5 will be denoted by 1, 2, 3, 4 and 5 respectively).
3. PRED This is a bernoulli response variable.
 1 - if the nest was preyed upon
 0 - otherwise.

This corresponds to the final status of the nest. Later in the analysis, we change this definition according to the status of the nest within the interval.

Kingbirds typically lay 2-5 eggs at a time. Three or four is the most common number and five is rare. Usually there are the same number of young in the beginning

and at the time of fledging or predation. Ms. Siderius says that if the nest is preyed upon, all the young will be killed. We have for nest 96, 3 young in I2 and 2 young in I3. She says that one young may have died due to some other reason in which we are not interested. In the study this is the only case of a death due to causes other than predation. She thinks that when there are more young in the nest, the parent birds may visit the nest more often than when there are fewer young in the nest and thereby the nests which have more young may get better protection. Furthermore, when there are more young they may have more strength to defend themselves against predators. On the other hand, she adds, that more young may result in a noisier nest and this may attract predators.

4. YNG Number of young in the nest.

Each observation amongst the behavioural variables 5-13 was obtained during a 90-minute observational period. For a nest, there is at most one observation of variables 5-13 in any of the intervals. Ms. Siderius stayed at a place from where she could observe the nest. She was careful to choose locations where her presence would not disturb the kingbirds' behaviour. The closest such location was about 15 meters from the nest.

Kingbirds visit their nests very frequently. When the young get older, the number of visits by the parent birds can be expected to increase because they have to be fed very frequently. This may also depend on the number of young in the nest. In some of their visits they are very noisy within two meters of the nest. Ms. Siderius calls these visits noisy visits. The reason for noisy visits is not clear. It may be a way of communicating an impending visit to the young or to the other parent bird. However, this noise may attract predators to the nest. Sometimes the parent birds were noisy at a distance more than two meters from the nest. These were not associated with the visits to the nest. She calls these noisy perches. The reason for noisy perches is also not understood. However, to see whether these kinds of behaviour have an effect on the nest predation, she collected observations on the following variables.

5. VIS This is the number of visits to the nest by the parent birds.

6. NVIS This is the number of noisy visits.

- 7. NOIP This is the number of noisy perches.
- 8. DNOI This is the mean distance in meters to the noisy perches.
- 9. SDNOI This is the standard deviation of distances in meters to the noisy perches.

As we mentioned earlier kingbirds chase any threatening bird that they see flying around the nest. The number of chasings might be a measure of how aggressive the kingbirds are and it might provide some information on the effect of aggressive behaviour in protecting the nest. Therefore she counted the number of crows and ravens that she could see and the number of aggressive encounters made by the kingbirds.

- 10. AG This is the number of aggressive encounters.
- 11. SEE This is the number of crows and other predatory birds that Ms. Siderius saw flying around the nest.

Ms. Siderius wonders if a nest is less likely to be preyed upon if the parent birds maintain close contact with the nest. By 'close' she means within a distance that the parent birds could see the nest. Usually this means within 100 meters of the nest. She therefore observed the following variables.

- 12. BOTH Duration in minutes where both parent birds were close to the nest.
- 13. ONE Duration in minutes where there was only one parent bird close to the nest.

The data set BEHAVIOUR appears in table 5.1.3 of appendix 1.

Chapter 3

THE ANALYSIS

3.1 PRELIMINARY ANALYSIS OF ENVIRONMENT

Kingbirds are averse to having another pair of kingbirds nest near them. Ms. Siderius told us that the closest two nests were more than 100 meters apart. This helps justify the assumption of independence between nests. In this analysis we assume that the observations are independent of each other.

In the data set ENVIRONMENT we have 95 observations on 95 nests of which 50 are preyed upon. Missing values of the variables is a major concern. Unfortunately, in this problem we could not find a reasonable way to simultaneously estimate all missing values. We therefore considered using only the complete cases. However if we take into account all the variables simultaneously, we are left with only 38 complete cases. This is certainly not a desirable number of observations for an analysis with 30 explanatory variables. Therefore our first concern was to condense the available information into a reasonable number of variables. Note that the variable NEST takes no role in the analysis except that it is useful for identification purposes. Furthermore it was decided to eliminate DWAT due to its unreliability. Therefore the number of explanatory variables has been reduced to 28.

The data were collected over three years. We would like to test if there is any significant difference of overall probability of predation in these three years. Let P_1 , P_2 and P_3 be the overall probabilities of predation in 1988, 1989 and 1990 respectively. The following contingency table (table 3.1.1) shows the total number of observations, number of preyed upon nests and the number of nests where the young fledged. It seems that there is an upward trend for predations over three years.

	1988	1989	1990	Total
Fledgings	18	17	10	45
Total	31	36	28	95

Table 3.1.1

We test the following hypothesis:

$$H_0: P_1=P_2=P_3$$

$$\text{vs } H_A: \text{not the case that } P_1=P_2=P_3.$$

The value of the χ^2 statistic based on this contingency table is 2.97 with 2 degrees of freedom. The corresponding p-value is 0.23. Therefore H_0 is not rejected. We conclude then that the probability of predation in the three years does not differ significantly. Therefore YEAR does not provide any information about the probability of predation and the number of predictor variables reduces from 28 to 27.

The variables 4-15 are essentially the measures of visibility of the nest from the four directions, obtained from three different techniques. Let us examine the correlation between these 12 variables. The correlation matrix of the variables is given in table 5.2.1 in appendix 2. The high correlations between the variables, extracted from this table are shown below.

	PN	BN		PS	BS
BN	0.673		BS	0.632	
WN	0.756	0.908	WS	0.646	0.942
	PE	BE		PW	BW
BE	0.570		BW	0.587	
WE	0.554	0.908	WW	0.657	0.857

As one would expect, for each direction, the three visibility variables are fairly correlated. We also note that we have no reason to believe that any of the three techniques is more accurate than other. Therefore, we form the following 4 visibility variables for north, south, east and west directions respectively. This reduces the number of explanatory variables from 27 to 19.

- VN rounded average of PN, BN and WN
- VS rounded average of PS, BS and WS
- VE rounded average of PE, BE and WE
- VW rounded average of PW, BW and WW

The correlation matrix of VN, VS, VE and VW appears in table 5.2.2 of appendix 2. It shows that these variables are not highly correlated.

Let us now examine the following figures (figures 3.1.1-3.1.2).

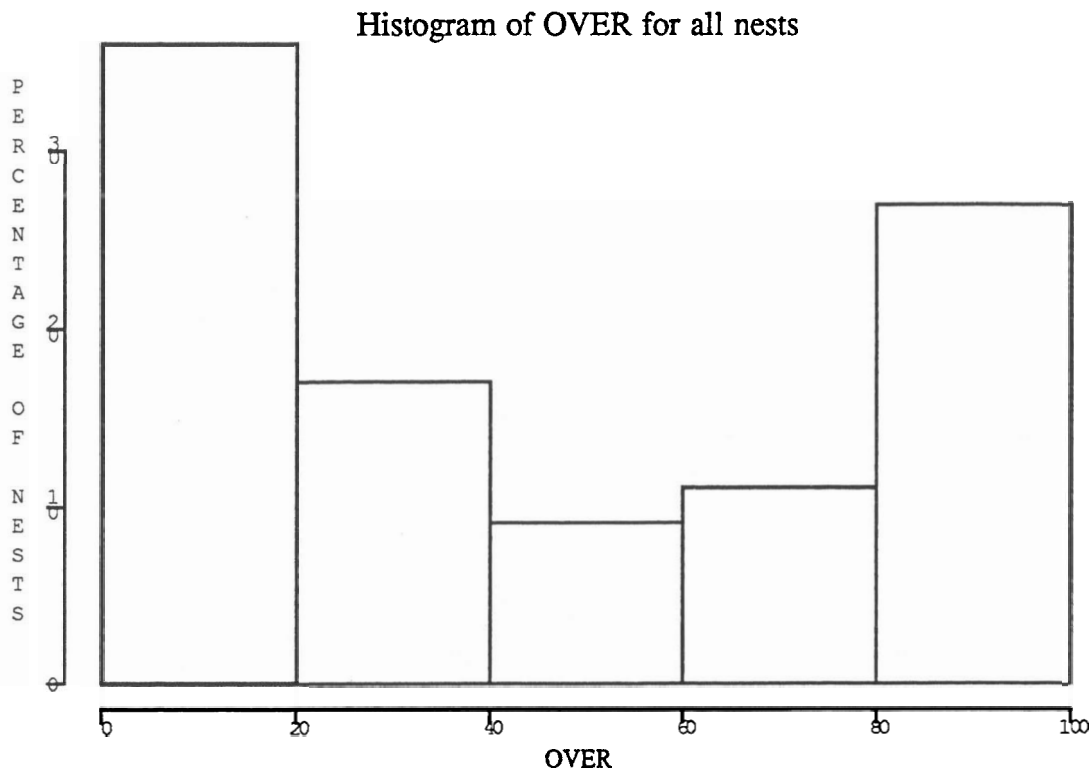


Figure 3.1.1

According to Ms. Siderius, most other types of birds hide their nests. However the histogram in figure 3.1.1 clearly shows that a considerable portion of kingbird nests are visible from above. Noticeably there is a high frequency of nests with overhead visibility greater than 80 percent.

Figure 3.1.2 contains two plots of the relative frequency of the variable OVER. The solid line and the dotted line correspond to the preyed upon nests and the fledged nests respectively. The range of OVER is divided into 5 intervals of equal length and the relative frequency in each interval is plotted against the middle point of the interval. Surprisingly, the highest relative frequency of the fledged nests has overhead visibility greater than 80 percent. The relative frequency for the fledged nests is less than that for preyed upon nests when OVER is below 80 and the situation is reversed

when OVER exceeds 80. This result may be explained by noting that kingbirds with high visibility may be alert to nearby predators.

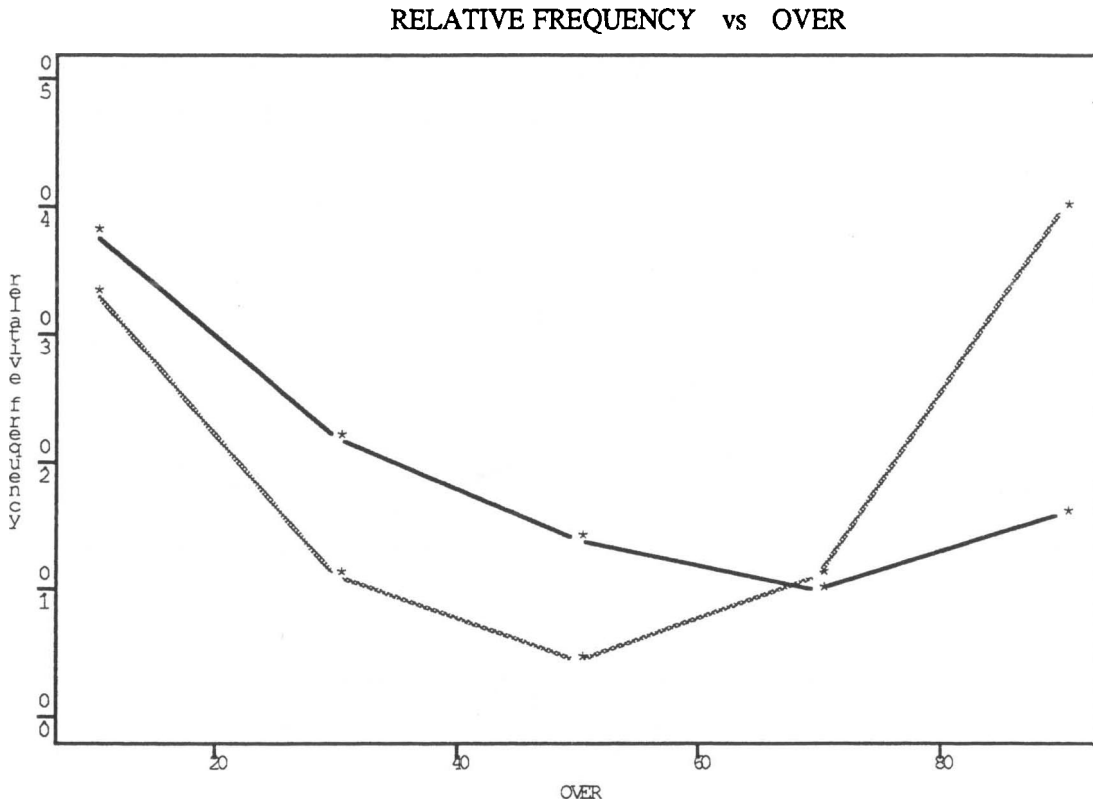


Figure 3.1.2

Figure 3.1.3 gives the plots of relative frequencies of VN, VS, VE and VW respectively. The graphs are produced as described above. Again, the solid line and the dotted line correspond to the preyed upon nests and the fledged nests respectively. In contrast to OVER, the variables VN, VS, VE and VW have the highest frequency at value zero for both the preyed upon and the fledged nests. It is apparent that kingbird nests are more covered from the sides than from above. It is also clear that a moderate proportion of nests have more than 50 percent visibility from the four directions. There is no strong evidence that these variables have different distributions for the two populations of fledged and preyed upon groups. These variables are the measures of nest visibility taken from the ground. If there were ground predators involved, we may expect these variables to be influential. However, as far as the ground predators are concerned, a more reasonable variable to look at may be either the average of the four variables or the maximum of the four

variables. We therefore chose to consider the following variables and thereby reduce the number of explanatory variables from 19 to 17. Note that we are left with only OVER and 2 of the original 12 visibility variables.

AV rounded average of VN, VS, VE and VW

MV Maximum of VN, VS, VE and VW

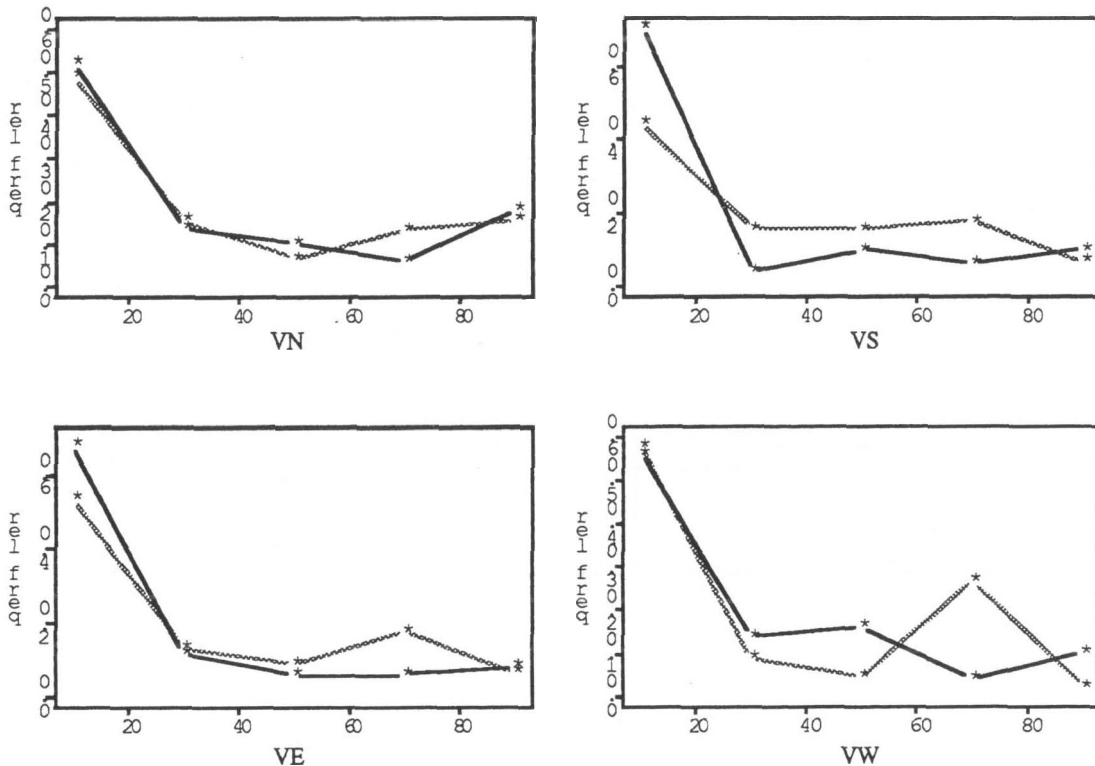


Figure 3.1.3

Recall that variables 16-19 provided information on whether the birds could see as far as 100 meters or 300 meters. Hence, it would be appropriate to look at the number of sides that they can see as far as 100 meters and the number of sides they could see as far as 300 meters. We create the following two new variables from variables 16-19. This reduces the number of explanatory variables in ENVIRONMENT from 17 to 15.

NS1 The number of sides that Ms. Siderius could see more than 80 meters.

NS2 The number of sides that Ms. Siderius could see more than 250 meters.

The correlation matrix of these 15 variables appears in table 5.2.3 in appendix 2. It shows that these 15 variables are almost uncorrelated except for the high correlation between NS1 and NS2 and the high correlation between AV and MV. The correlation between NH and NS1 is 0.292 and the correlation between NH and NS2 is 0.154. Therefore the speculation that the kingbirds can see farther in higher nests is not supported.

We now examine the histograms in figures 3.1.4-3.1.7. The histograms of WAT (figures 3.1.4-3.1.6) show that about 3/4 of the nests have at least a small amount of water close to the nest. Noting that there are 95 nests a simple calculation shows that about 3/4 of the nests which do not have water within 100 meters are preyed upon.

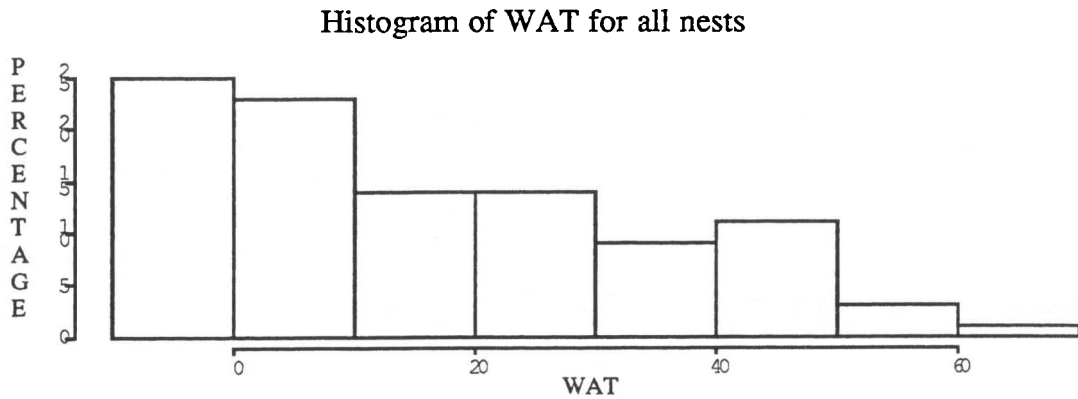


Figure 3.1.4

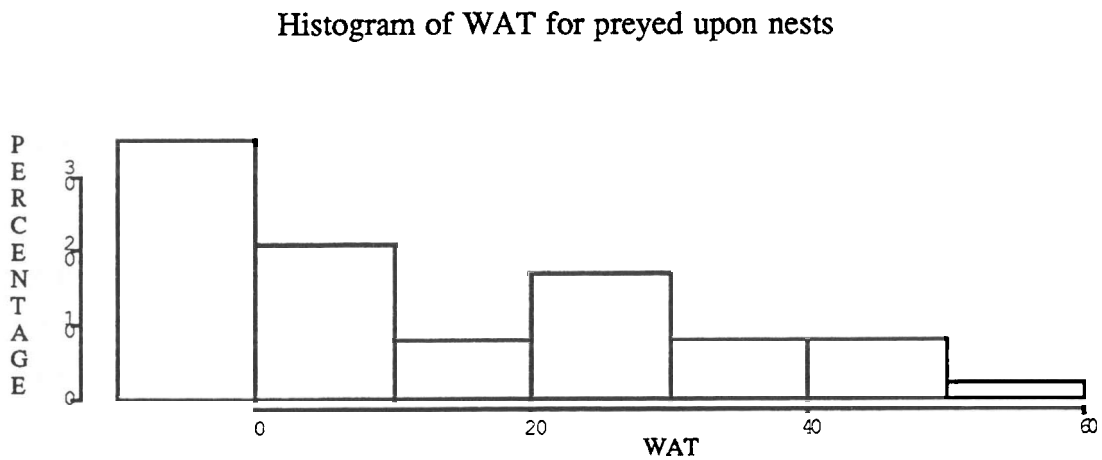


Figure 3.1.5

Histogram of WAT for fledged nests

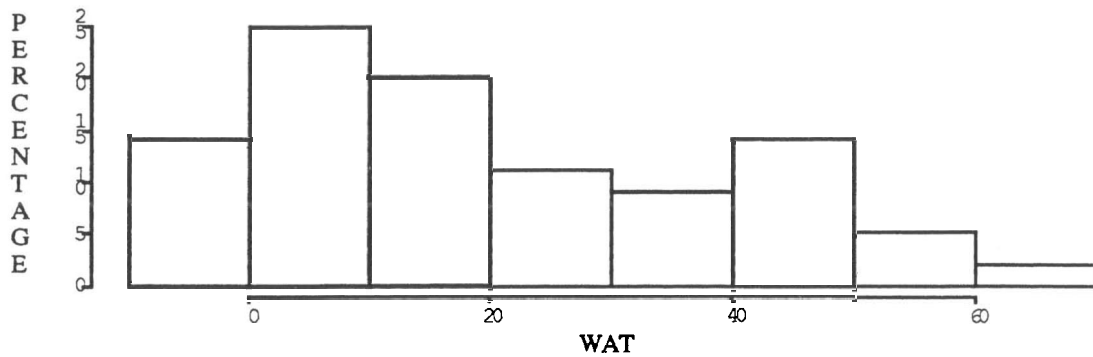


Figure 3.1.6

Histogram of TREE for all nests



Figure 3.1.7

The histogram of TREE (figure 3.1.7) shows that about 1/2 of the nests do not have other trees within 100 meters of the nest and approximately 90% of the nests are in areas where there are less than 20 percent trees within 100 meters of the nest. Therefore kingbirds seem fond of nesting in clear areas close to water.

The population of kingbird nests can be considered to consist of two groups; fledged nests and preyed upon nests. If a variable is associated with this grouping, it may be reasonable for us to expect different characteristics of the variable in the two groups. Table 5.2.4 in appendix 2 contains a description of some simple distributional properties of each of the variables by the two groups. According to these descriptions clearly none of the variables is closely normally distributed. As a very conservative way of trying to choose the variables which are significant predictors between fledged

and preyed upon groups, we did a one-way analysis of variance for each of these 15 variables. Table 3.1.2 shows the p-values associated with the variables.

Variable	F-value	p-value
OVER	3.88	0.05
AV	1.07	0.30
MV	0.27	0.60
NS1	0.00	1.00
NS2	2.79	0.10
NH	0.13	0.72
TH	0.36	0.55
NTRE	0.84	0.36
STEM	0.90	0.35
NPER	0.14	0.71
WAT	3.39	0.07
FIELD	2.46	0.12
SHRUB	0.48	0.49
MARSH	0.09	0.76
TREE	7.16	0.01

Table 3.1.2

In order to allow for the non-normality of the variables, the variables with corresponding p-values less than 0.25 were selected for further consideration. With this approach we have reduced the data set to 5 explanatory variables; namely, OVER, NS2, WAT, FIELD and TREE. We have 52 complete observations with these variables.

3.2 LOGISTIC REGRESSION ANALYSIS OF ENVIRONMENT

PRED is a bernoulli response variable which takes the value 1 if the nest is preyed upon and which takes the value 0 otherwise. For simplicity, hereafter it will be denoted by Y.

Then $Y_i = 1$ if NEST i is preyed upon
 $= 0$ if the young in NEST i eventually fledge.

Let $P_i = \Pr(Y_i=1)$.

Our aim is to find a relationship, if any, between P_i and the set of variables OVER, NS2, WAT, FIELD and TREE. Let X be the design matrix comprised of the columns of these explanatory variables together with a constant term.

Let

$$\theta = H^{-1}(P) = g(P) = X \beta \quad (3.2.1)$$

be our statistical model where $H(\theta)$ is a strictly increasing function of θ with $-\infty < \theta < \infty$, $H(-\infty)=0$ and $H(\infty)=1$. Note that a function H satisfies these requirements if it is the cumulative distribution function of any continuous random variable on $(-\infty, \infty)$. H^{-1} or g is called the link function. A widely used link function is the logit link $g(P)=\log(P/(1-P))$ which corresponds to the distribution function of the logistic distribution. Our basic interest is to find the components of β which are significantly different from zero. The maximum likelihood estimate b of β can be found by solving the system of non-linear equations given by $\partial l / \partial \beta = 0$ where l is the log-likelihood function.

Suppose $L(b_m ; y)$ is the maximum achievable likelihood corresponding to some maximal model and $L(b ; y)$ is the maximum likelihood for a sub-model of interest. If the model of interest describes the data well then $L(b ; y)$ should be approximately equal to $L(b_m ; y)$. If the model is poor then $L(b ; y)$ will be much smaller than $L(b_m ; y)$. This suggests the use of the generalized likelihood ratio statistic $\lambda = L(b_m ; y) / L(b ; y)$ as a goodness of fit statistic. Equivalently, $\log(\lambda) = \{l(b_m ; y) - l(b ; y)\}$ where l is the log-likelihood, can be used. It has been shown that $D = 2 \log(\lambda)$ has an asymptotic χ_r^2 distribution, if the fitted model describes the data well. Here r is the difference in the number of parameters between the maximal model and the sub-model. D is called the deviance of the fitted model from the maximal model.

There are two types of binary data:

- (a) binary data without common covariates (no covariate classes)
- (b) binary data with common covariates (with covariate classes).

McCullagh and Nelder[8] (pp. 121) point out that for binary data without any common covariates, the large sample theory does not apply to the distribution of the deviance D . Aitkin et al[1] says that in this case, failure of the model to fit at individual observations has to be assessed by residual examinations.

Let the estimated value of P_i be denoted by p_i . For binary data, the Pearson residuals are defined as $e_i = (y_i - p_i) / \sqrt{p_i(1-p_i)}$ where y_i is 0 or 1. These are approximately standardized variables with mean 0 and variance approximately 1. These residuals are not normally distributed as e_i can take only two possible values a_i and $-1/a_i$ for each i , where $a_i = \sqrt{p_i/(1-p_i)}$. Large values of Pearson residuals indicate failure of the model to fit at the corresponding points.

According to McCullagh and Nelder[8], when we have covariate classes, the large sample theory can be applied to the residual deviances given that the following assumptions are satisfied.

1. The observations are distributed independently according to the binomial distribution.
2. Letting n be the number of covariate classes, m_i the number of observations in the i -th covariate class and P_i the corresponding probability of predation, the approximation is based on a limiting operation in which n is fixed, $m_i \rightarrow \infty$ and $m_i P_i(1-P_i) \rightarrow \infty$.

If the above assumptions are satisfied then D is asymptotically distributed as a χ^2 random variable with $n-q$ degrees of freedom, where q is the number of fitted parameters. The sum of squares of Pearson residuals is the Pearson goodness of fit statistic and it is distributed as χ^2 with $n-q$ degrees of freedom for large samples. If the fitted model is adequate, we may expect D and χ^2 to be close to $n-q$. McCullagh and Nelder[8] point out that the χ^2 assumption is usually quite accurate for differences of deviances even though it is inaccurate for the deviances themselves. The Statistical package GLIM allows us to fit the model (3.2.1) with several link

functions. In GLIM output, the values of D and the Pearson χ^2 statistic are given as the scaled deviance and %X2.

We now report on the procedure to fit a logistic regression model $\log(P/1-P)=X\beta$ using variables OVER, NS2, WAT, FIELD and TREE as explanatory variables. We fit the model using a stepwise regression procedure. In each step we make use of the maximum number of complete observations for the combinations of variables concerned.

- Step 1 Add each variable separately to the null model based on the available observations for the variable concerned. TREE gives the lowest p-value of 0.007 corresponding to a deviance decrease of 7.29. This is the difference between the deviances corresponding to the null model and the constant+TREE model based on 92 observations.
- Step 2 Add the rest of the variables separately to the constant+TREE model. WAT gives the lowest p-value of 0.034 corresponding to a deviance decrease of 4.47. This is the difference between the deviances corresponding to the constant+TREE model and the constant+TREE+WAT model based on 92 observations.
- Step 3 Drop TREE from the constant+TREE+WAT model. The deviance increases by 8.38. The p-value corresponding to this increase is 0.004. Therefore do not drop TREE at this step.
- Step 4 Add the rest of the variables separately to the constant+TREE+WAT model. OVER gives the lowest p-value of 0.01 corresponding to a deviance decrease of 6.66. This is the difference between deviances corresponding to the constant+TREE+WAT model and the constant+TREE+WAT+OVER model based on 92 observations.
- Step 5 Drop TREE from the constant+TREE+WAT+OVER model. The deviance increases by 7.26. The corresponding p-value is 0.007. Therefore retain TREE in the model. Drop WAT from the constant+TREE+WAT+OVER model. The deviance increases by 8.16. The corresponding p-value is 0.004. Therefore keep WAT in the model.

Step 6 Add the rest of variables separately to the constant+TREE+WAT+OVER model. NS2 gives the lowest p-value of 0.11 corresponding to a deviance drop of 2.56. This is the difference between deviances corresponding to the constant+TREE+WAT+OVER model and the constant+TREE+WAT+OVER+NS2 model based on 52 observations. Considering the large number of missing values and the large p-value we do not add NS2 to the model nor any other variable.

We have now recognized WAT, OVER and TREE as potential influential variables. Table 3.2.1 summarizes the deviance drops and the corresponding p-values of variables in the stepwise regression procedure.

Variable	Deviance Decrease	p-value
TREE	7.29	.007
WAT	4.47	.034
OVER	6.66	.010

Table 3.2.1

It seems that OVER, WAT and TREE are significant variables. The scaled deviance for the constant+TREE+WAT+OVER model based on 92 observations is 108.95 with 88 degrees of freedom. This model leaves one big residual of size 4.00 at NEST 34. This nest is a preyed upon nest with OVER, WAT and TREE taking values 60, 30 and 60 respectively. In general, at this step we should check for correctness of this observation with the experimenter. If it is correct the model should be modified, or else we accept the poor fit at this point as random variation. We mentioned earlier that some of the values of the variables in our data set are not very reliable. Ms. Siderius was unable to comment on the reliability of this outlier mentioned above. Therefore we consider the model based on 92 observations as well as the same model with the outlier removed. The scaled deviance corresponding to the model based on 91 observations after removing NEST 34 is 102.17 with 87 degrees of freedom. This model fits the data without leaving any residual greater than 3 in absolute value. Table 3.2.2 shows estimated parameters, estimated standard deviations of estimated parameters and the corresponding p-values associated with the variables in the refit model.

Variable	Estimated Parameter	Estimated Std Dev	t-value	p-value
constant	2.41	0.66	3.65	0.0002
	-0.07	0.02	-3.50	0.0004
	-0.04	0.01	-4.00	0.0001
OVER	-0.02	0.01	-2.00	0.02

Table 3.2.2

However this model may not seem physically plausible. One would hesitate to believe that log-odds of predation depends linearly on TREE, WAT and OVER. Therefore, before we proceed modify the model let us look at the variables TREE, WAT and OVER more closely. Figures 3.2.1, 3.2.2 and 3.2.3 are the plots of OVER, WAT and TREE against NEST respectively. The circled points correspond to the preyed upon nests.

Plot of OVER vs NEST (circled points are preyed upon nests)

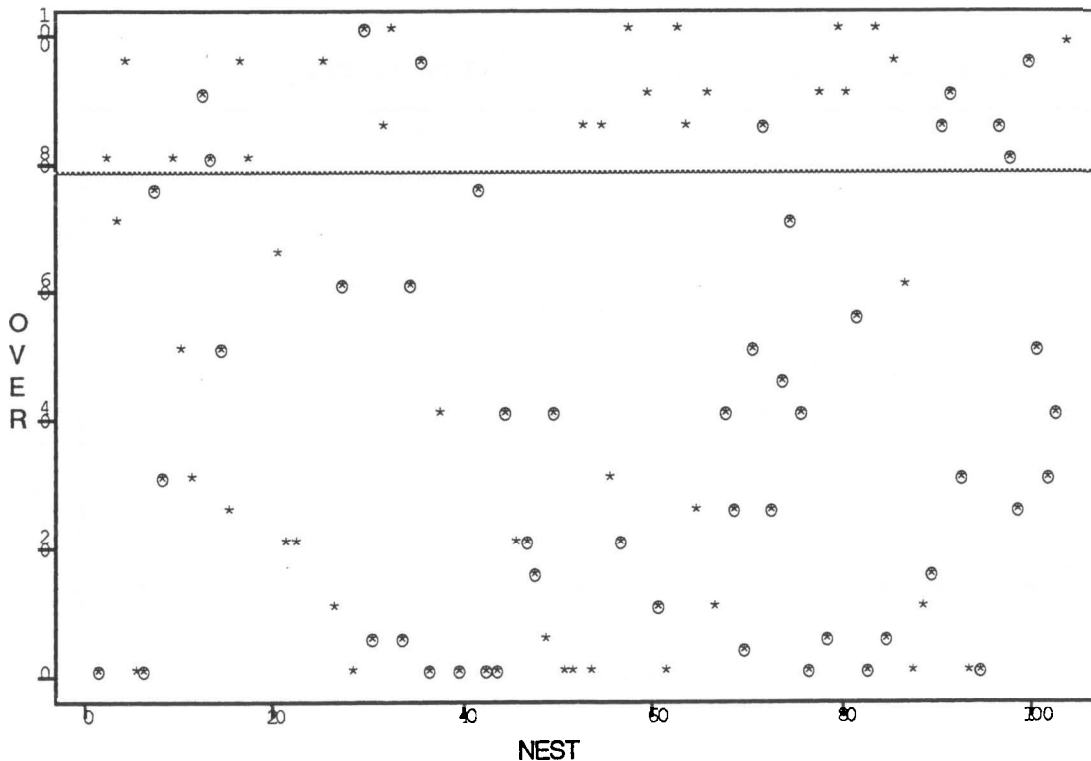
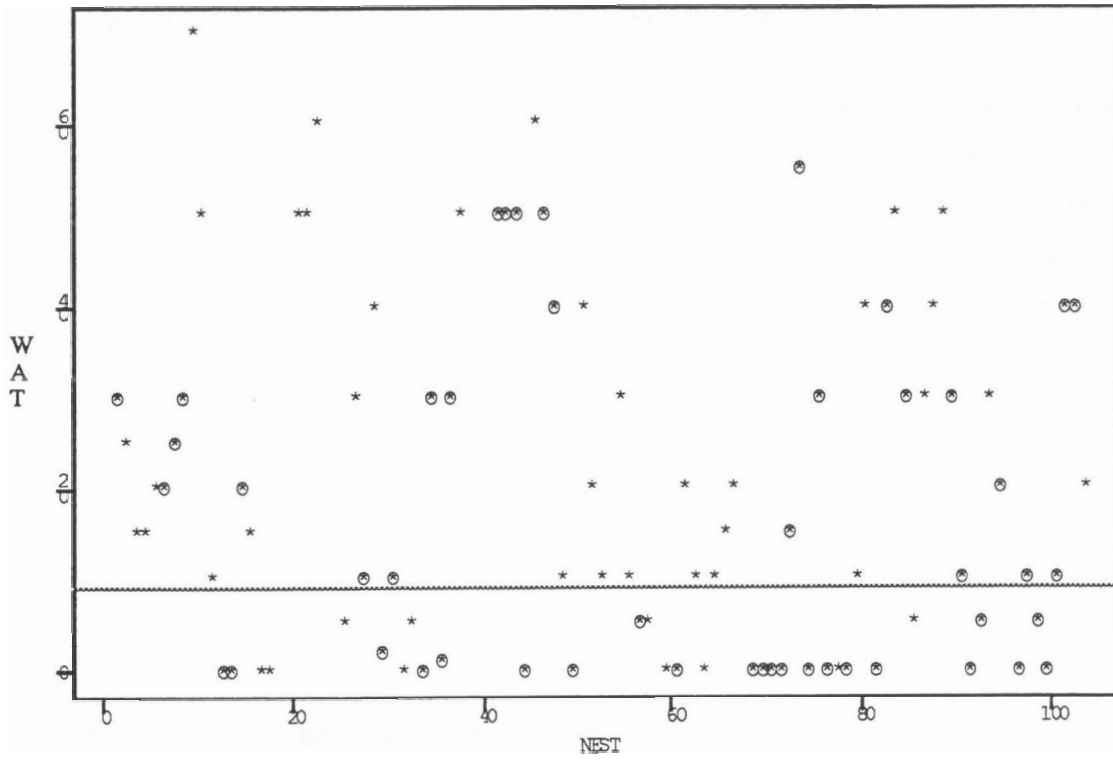


Figure 3.2.1

Plot of WAT vs NEST (circled points are preyed upon nests)



Plot of TREE vs NEST (circled points are preyed upon nests)

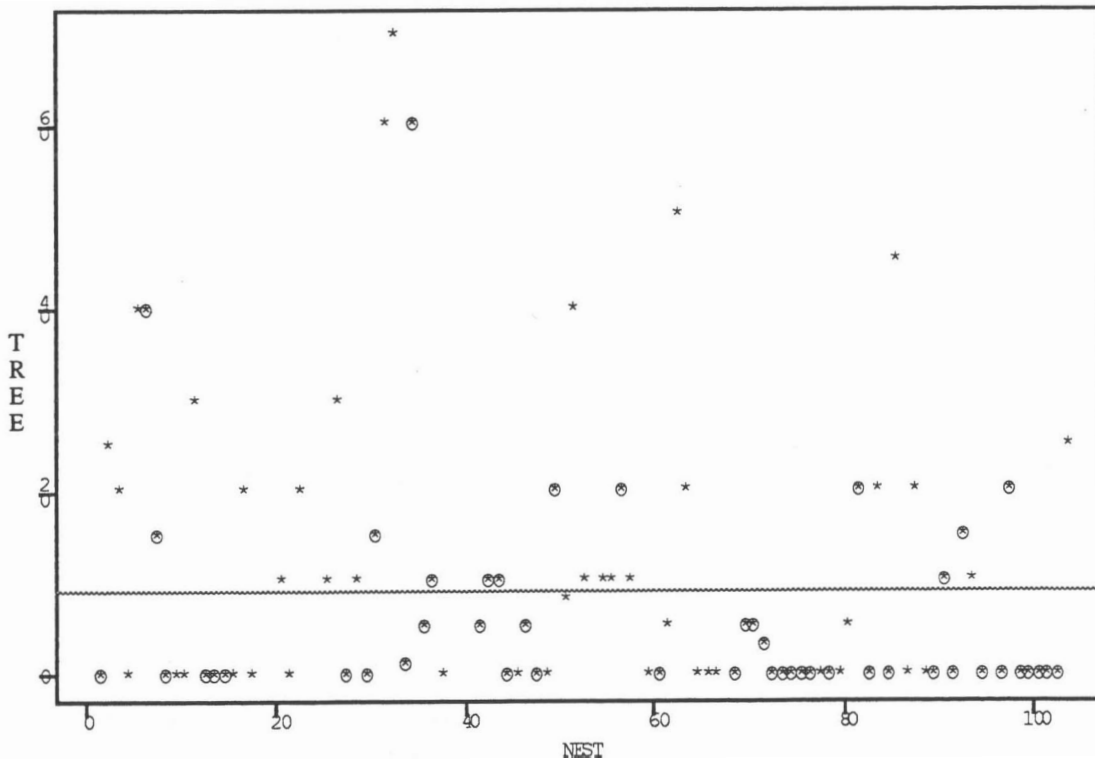


Figure 3.2.2 and Figure 3.2.3

In figure 3.2.1, the circled points are more dense when OVER is below 80. The percentage of preyed upon nests is 62.3 for $OVER < 80$ and 35.3 for $OVER \geq 80$. Figure 3.2.2 shows that the circled points are more dense when $WAT < 10$. The percentage of circled points in this region is 68.6 while it is 44.4 outside the region. Figure 3.2.3 also shows this kind of accumulation of preyed upon nests in a region. We find that 63.8 percent nests for which $TREE < 10$ are preyed upon and 37.5 percent of the nests for which $TREE \geq 10$ are preyed upon. Some possible biological explanations for these exploratory findings are given in section 3.3. At this point we would like to mention that although we are carrying out thoughtful exploratory data analysis, we should be careful to determine real trends and not simply artifacts of this particular data set.

Although figures 3.2.1-3.2.3 give us some clues regarding potentially influential variables, they do not specify the form of these variables. We therefore construct 3 new variable candidates from each of the old variables TREE, WAT and OVER by partitioning each old variable into 2 intervals as shown below. We then fit 27 models with different combinations of the three variables.

1. TR1 = TREE if TREE < 10
 = 0 otherwise.

- TR2 = TREE if TREE \geq 10
 = 0 otherwise.

- TR3 = 0 if TREE < 10
 = 1 otherwise.

2. WA1 = WAT if WAT < 10
 = 0 otherwise.

- WA2 = WAT if WAT \geq 10
 = 0 otherwise.

- WA3 = 0 if WAT < 10
 = 1 otherwise.

3. $OV1 = \begin{cases} OVER & \text{if } OVER < 80 \\ 0 & \text{otherwise} \end{cases}$
- $OV2 = \begin{cases} OVER & \text{if } OVER \geq 80 \\ 0 & \text{otherwise.} \end{cases}$
- $OV3 = \begin{cases} 0 & \text{if } OVER < 80 \\ 1 & \text{otherwise.} \end{cases}$

We now present a summary of fitting 27 logistic models in table 3.2.3. Each of these fits involves 92 observations. As we mentioned before, the measurements of the variables in our problem are not very reliable. Therefore when we find large residuals, (here we consider residuals which are greater than 3 in absolute value) we remove the associated observations and see whether the model fits the rest of the data.

Model	Variables Involved	Deviance	Outliers (NESTs)	Deviance (Refit Model)
1	TR1 WA1 OV1	122.00	--	--
2	TR1 WA1 OV2	118.97	--	--
3	TR1 WA1 OV3	119.51	--	--
4	TR1 WA2 OV1	116.70	--	--
5	TR1 WA2 OV2	108.53	--	--
6	TR1 WA2 OV3	109.37	--	--
7	TR1 WA3 OV1	114.29	--	--
8	TR1 WA3 OV2	102.96	90, 97	89.69
9	TR1 WA3 OV3	104.07	90, 97	89.68
10	TR2 WA1 OV1	115.53	--	--
11	TR2 WA1 OV2	112.98	--	--
12	TR2 WA1 OV3	113.42	--	--
13	TR2 WA2 OV1	109.54	--	--
14	TR2 WA2 OV2	102.20	--	--
15	TR2 WA2 OV3	102.85	--	--
16	TR2 WA3 OV1	107.96	--	--

17	TR2 WA3 OV2	98.33	90, 97	84.74
18	TR2 WA3 OV3	99.16	90, 97	84.25
19	TR3 WA1 OV1	115.65	--	--
20	TR3 WA1 OV2	112.43	--	--
21	TR3 WA1 OV3	112.98	--	--
22	TR3 WA2 OV1	110.50	--	--
23	TR3 WA2 OV2	102.76	--	--
24	TR3 WA2 OV3	103.52	--	--
25	TR3 WA3 OV1	109.10	--	--
26	TR3 WA3 OV2	98.74	90, 97	83.10
27	TR3 WA3 OV3	99.59	90, 97	82.33

Table 3.2.3

It appears that models 26 and 27 are the most promising models and we now investigate them more closely.

Model	Variables Involved	Scaled Deviance	Outliers (NESTs)
26	TR3 WA3 OV2	98.74	90, 97
27	TR3 WA3 OV3	99.59	90,97

Table 3.2.4

The deviance difference between these 2 models is negligible. Both models fail to fit at NESTs 90 and 97. Note that these are the only preyed upon nests which have OVER, TREE and WAT greater than or equal to 80, 10 and 10 respectively. TR3 and WA3 are related to the environment surrounding the nest. In models 26 and 27 these 2 variables appear in similar forms which may make the comparison easier. If we consider the outliers as chance variation and include them in the models, then with regard to the scaled deviance, both models will be equally preferable. Model 27 is slightly simpler and easier to interpret than the model 26. The estimated parameters and the corresponding estimated standard errors associated with the variables in the two models are presented in table 3.2.5.

Model	Parameter	Estimated Parameter	Estimated Std Dev	t-value	p-value
26	constant	3.01	0.83	3.63	0.0002
	TR3(2)	-1.02	0.50	-2.04	0.02
	WA3(2)	-2.48	0.81	-3.06	0.001
	OV2	-0.03	0.01	-3.00	0.002
27	constant	2.93	0.81	3.62	0.0003
	TR3(2)	-1.05	0.50	-2.10	0.02
	WA3(2)	-2.40	0.79	-3.04	0.002
	OV3(2)	-2.52	0.79	-3.19	0.001

Table 3.2.5

Note: If Z is a categorical variable with 2 levels, Z(2) will denote the regression coefficient associated with the second level of the variable, in GLIM output.

Next let us examine the models where the outliers have been removed. If we refit model 26 without NESTs 90 and 97, we get a fit with no outliers greater than 3 in absolute value. Removing outliers slightly increases the p-values corresponding to the constant, WA3(2) and OV2 terms and decreases the p-value corresponding to TR3(2). The deviance for this model is 83.10 with 86 d.f. The parameter estimates and the corresponding estimated standard errors associated with the variables in refit model 26 are presented in table 3.2.6. The refit model 27 omitting NESTs 90 and 97 does not leave any residual greater than 3 in absolute value although the standard deviations of the estimates of the constant, WA3(2) and OV3(2) terms increase considerably. Aitkin[1](pp. 175) has pointed out that the use of parameter standard errors obtained from the expected information matrix could be seriously misleading in small samples when the likelihood function has substantial skew. We should heed this warning and place little faith in the individual estimated values of the parameters. We make this statement due to the extreme differences between models with and without their outliers. In order to see the combined importance of the constant, WA3 and OV3 terms, we omitted them from the model one at a time and found that the corresponding deviance increases are very significant. The deviance for the refit model 27 is 82.33 with 86 d.f. The parameter estimates and the corresponding estimated standard errors associated with the variables in refit model 27 are also presented in table 3.2.6.

Model	Parameter	Estimated Parameter	Estimated Std Dev	t-value	p-value
26	constant	6.04	2.85	2.12	0.02
	TR3(2)	-1.39	0.55	-2.53	0.01
	WA3(2)	-5.48	2.85	-1.92	0.03
	OV2	-0.06	0.03	-2.00	0.02
27	constant	9.94	13.37	0.77	0.22
	TR3(2)	-1.46	0.55	-2.65	0.004
	WA3(2)	-9.35	13.37	-0.70	0.24
	OV3(2)	-9.64	13.37	-0.72	0.24

Table 3.2.6

We prefer model 27 since it is simpler and easier to interpret than model 26. Furthermore, the deviance for model 27 is slightly less than the deviance corresponding to model 26.

Therefore our final fitted model is:

$$\log_e(P/1-P) = 9.94 - 9.46 \text{ OV3} - 1.46 \text{ TR3} - 9.35 \text{ WA3}.$$

Table 3.2.7 presents the estimated probabilities of predation in each covariate class.

Covariate Class	OV3	TR3	WA3	N	n	Observed Probability	Fitted Probability
1	0	0	0	10	10	1.00	0.99
2	0	0	1	29	17	0.59	0.64
3	0	1	0	4	4	1.00	0.99
4	0	1	1	18	7	0.39	0.26
5	1	0	0	11	8	0.73	0.59
6	1	0	1	5	0	0.00	0.00
7	1	1	0	7	0	0.00	0.22
8	1	1	1	6	0	0.00	0.00

Table 3.2.7

The following graph (figure 3.2.4) shows the estimated probability of predation versus the percentage of overhead visibility for the 8 covariate classes. The numbers 1,2,...,8 refer to the covariate classes in table 3.2.7.

Estimated probability of predation vs OVER for 8 covariate classes

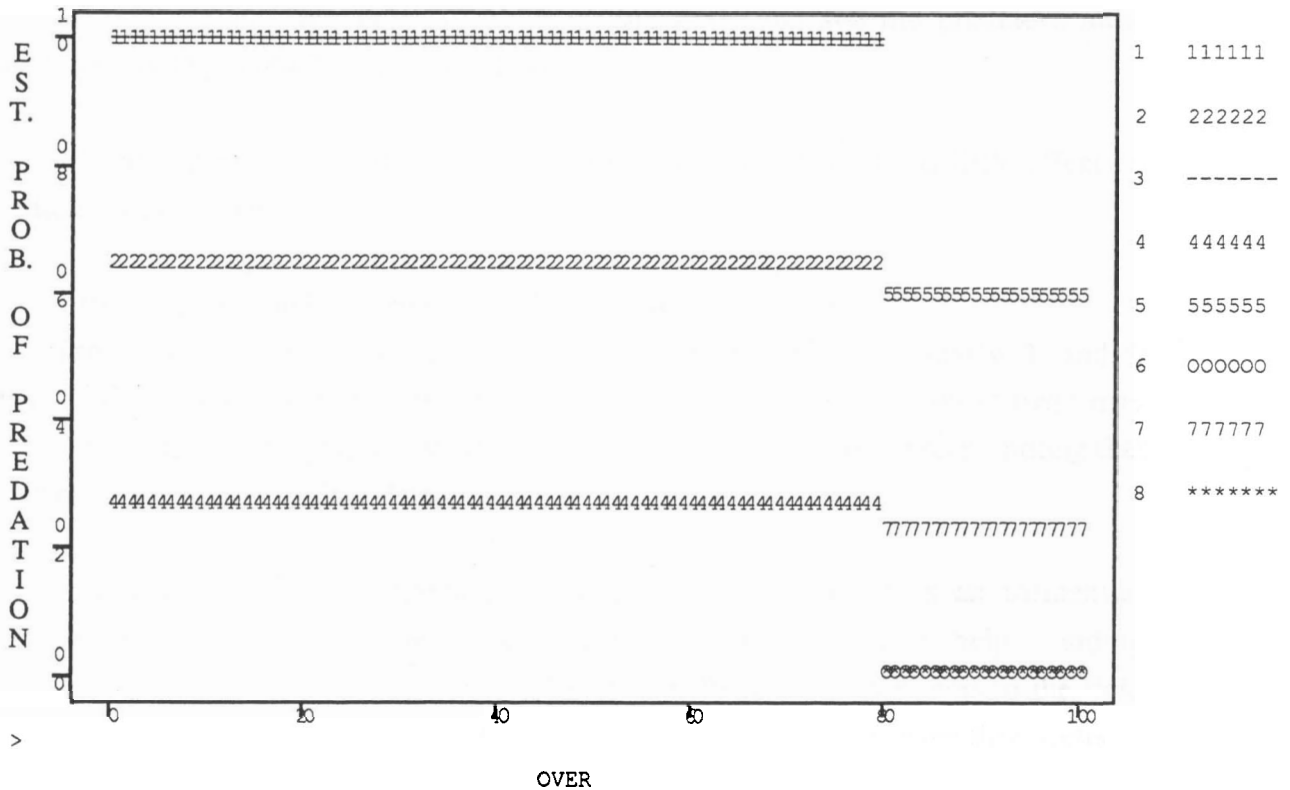


Figure 3.2.4

3.3 SUMMARY OF THE ANALYSIS OF ENVIRONMENT

After initially reducing the data set we found that the variables OVER, WAT and TREE are the most influential environmental variables affecting kingbird predation. Most of the other variables were eliminated by a conservative application of one-way analysis of variance. The remaining variables were eliminated using stepwise regression.

Among the visibility variables, only overhead visibility turned out to be important. The apparent effect of the overhead visibility is quite interesting. One of Ms. Siderius's initial questions was whether infant kingbirds in visible nests are more liable to be preyed upon. The actual effect is in fact the reverse. Figure 3.2.4 suggests that the visible nests more often lead to a fledge. The high overhead visibility ($\geq 80\%$) seems strangely to be an advantage to the kingbirds rather than to the predators. It may be possible that the birds in more visible nests can see the predators more easily, before they come close to the nest.

From figure 3.2.4 it also seems that the existence of trees has little effect on predation when either

(a) there is poor overhead visibility and little water

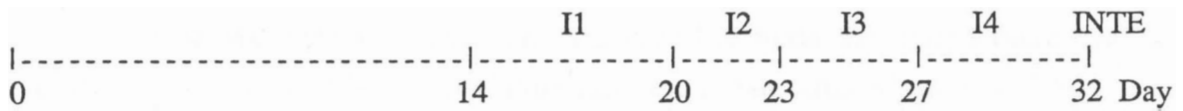
or (b) there is good overhead visibility and substantial water.

In these two cases the estimated probability of predation is nearly 1 and 0 respectively. In other scenarios it seems that the existence of some trees ($\geq 10\%$) may have a mild effect in helping the young to fledge. This may be explained by noting that trees offer protection against detection.

Figure 3.2.4 clearly shows that the percentage of water is an influential variable. It seems that the existence of some water ($\geq 10\%$) is a great help in aiding young kingbirds to fledge. As mentioned before kingbirds feed on insects in the field and on the water. With available water, kingbirds may not have to leave their nests as often since insects may be more plentiful and perhaps larger. This would aid kingbirds in their nest vigilance.

3.4 PRELIMINARY ANALYSIS OF BEHAVIOUR

The data set BEHAVIOUR contains 48 observations on 23 nests of which 9 were preyed upon. The interval of 32 days after the laying of eggs was divided into 5 intervals as described in section 2.2. One observation per nest in each or some of the intervals was collected. Although we have 5 intervals, we do not have any information in interval I1 because Ms. Siderius did not collect data in this interval. Therefore, for simplicity we define new INTE to be old INTE-1.



The number of kingbird visits, VIS is expected to increase with time. We therefore expect a set of increasing, parallel time series corresponding to different nests. If VIS is influential on predation we would also expect a difference between the plots for preyed upon nests and the plots for fledged nests. Figure 3.4.1 gives the plots of VIS versus INTE for the 6 preyed upon nests on which we have at least 2 observations in consecutive intervals. Figure 3.4.2 is the similar plot for the 7 fledged nests. These plots suggest that VIS varies with time. However we can not see any parallelism between lines for different nests, in either group, as expected. Figures 3.4.1 and 3.4.2 do not provide any evidence for the dependence of VIS between different intervals as no pattern is recognizable. (In comparing the plots, it could be noted that there is a scale difference in two plots.)

VIS vs INTE for preyed upon nests

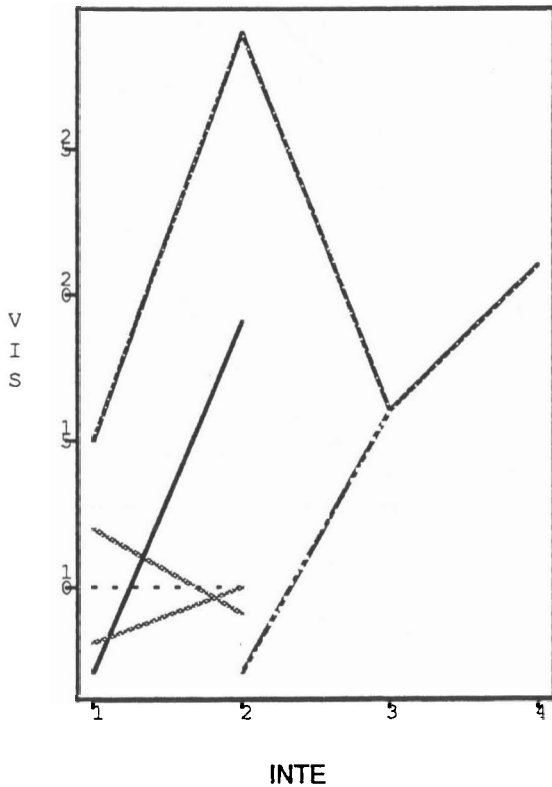


Figure 3.4.1

VIS vs INTE for fledged nests

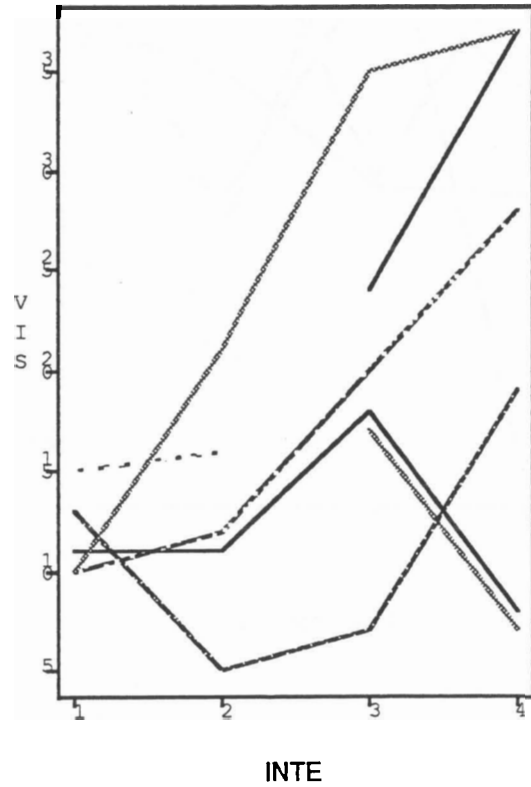


Figure 3.4.2

One of Ms. Siderius's questions was whether birds that spend more time at the nest lose their offspring to predators less often than birds which spend less time at the nest. Figures 3.4.3 and 3.4.4 are the plots of BOTH vs INTE for the 6 preyed upon nests and the 7 fledged nests respectively. These plots do not provide enough evidence to support the above speculation. Furthermore, these plots also do not provide enough evidence of dependence between intervals for a given nest. We also examined graphically the variables NVIS, NOIP and DNOI and obtained similar results to VIS and BOTH. We can neither identify a similarity of any variable within groups nor a difference of any variable between groups. (Again notice the scale difference in two plots)

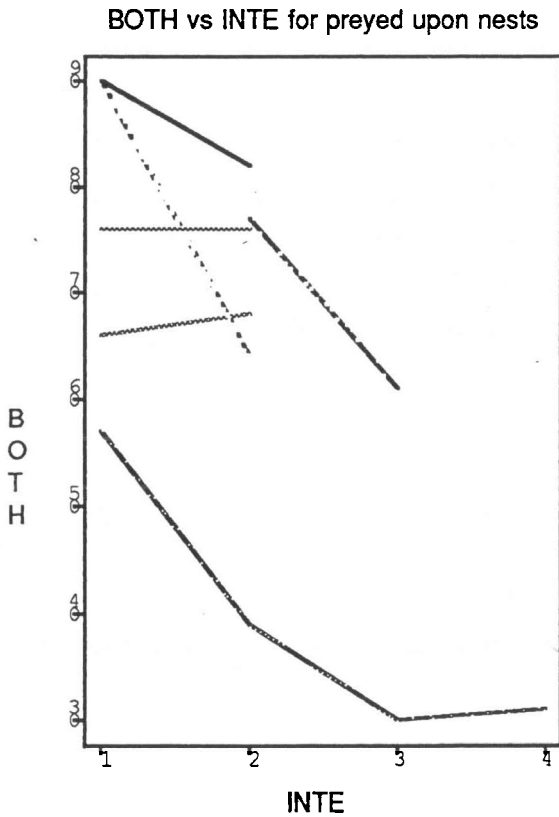


Figure 3.4.3

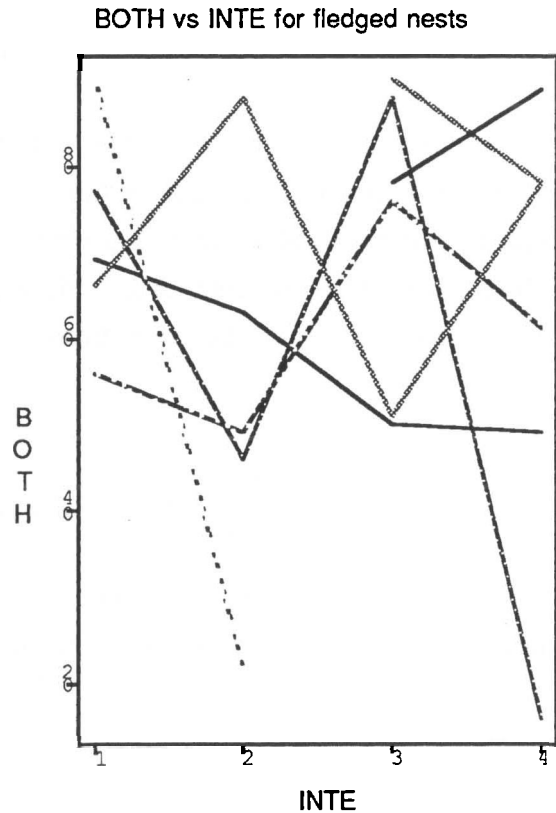


Figure 3.4.4

As mentioned in section 3.1, we can reasonably assume independence between nests. We also noticed that the observations for a given nest in different intervals do not show any dependence. Therefore it may be reasonable to assume that the 48 observations are independent of each other.

Let us next examine the correlations between the variables in this data set. The correlation matrix of the variables appears in table 5.2.5 of appendix 2. The correlation between BOTH and ONE is very high as the sum of these is 90 minutes or very close to 90 minutes. One of Ms. Siderius's questions was whether noisy birds attract predators. The variables NVIS and NOIP can be considered as measures of the noisiness. SEE is the number of predatory birds that fly around the nest. The correlation between NVIS and SEE is 0.103 while the correlation between NOIP and SEE is 0.118. Therefore there is no evidence to support the idea that noisy birds attract predators. SDNOI, the standard deviation of the distances from the nest to where the parent birds were noisy is greater than DNOI, the mean of those distances. Therefore, due to its variability DNOI may not be a good summary statistic for the distances. The second highest correlation is 0.549 between NVIS and NOIP. This implies that there are no other significant linear relationships between variables.

We would like to emphasize strongly that with an extremely limited number of observations it is very difficult to recognize variables that are influential on kingbird nest predation. Even when there is a real difference of a variable in the two groups, 7 observations is hardly enough to visualize such a difference. In sections 3.5 and 3.6 we try to identify significant variables, if any, using more advanced methods.

3.5 LOGISTIC REGRESSION ANALYSIS OF BEHAVIOUR

In Ms. Siderius's data set, PRED was a binary variable taking value 1 if the nest was preyed upon within 32 days and 0 otherwise. We change this definition as follows.

PRED = 1 if the nest was preyed upon within the interval concerned.
= 0 otherwise.

This definition does not change the values of PRED in Ms. Siderius's data set for fledged nests. It does however change the values of PRED for preyed upon nests. Ms. Siderius provided us a record of dates from which we could estimate the approximate length of time prior to the predation. This list of dates is presented in table 5.1.4 in appendix 1. Table 3.5.1 shows the estimated length of time prior to the predation and

the values of the variable PRED according to the new definition. For example, Ms. Siderius found that the first egg was laid approximately on the 28th of June, 1990 in NEST 94. Her last visit of the nest prior to the predation was on the 22nd of July and the first visit after the predation was on the 27th of July. From this information we estimate the length of time prior to the predation to be 27 days. Therefore it has been preyed upon in the interval I3. However Ms. Siderius has observed behavioural variables for this nest only in I1 and I2. Within these two intervals the nest was not preyed upon. Therefore PRED takes the value 0 corresponding to these two intervals. In Ms. Siderius's original data set, both values of PRED were set equal to 1 because the nest was preyed upon within 32 days. NEST 96 is similar. The time prior to the predation for NEST 39 is not known. However we have an observation on this nest in I1. The reason that we do not have an observation in I2 may be that the nest had already been preyed upon when Ms. Siderius visited the nest for the second time. Therefore we assume that the nest was preyed upon at the end of the interval I1.

NEST	INTE	Days to Predation	PRED
39	I1	20	1 +
75	I1		0
75	I3	26	1
76	I1		0
76	I2	23	1
81	I1		0
81	I2		0
81	I3		0
81	I4	30	1
82	I1	20	1
91	I1		0
91	I2	23	1
92	I1		0
92	I2	23	1
94	I1		0
94	I2	27	0 ++
96	I1		0
96	I2	27	0 ++

Table 3.5.1

- + Assume that the nest was preyed upon at the end of the interval I1.
- ++ Nest was preyed upon in I3. But we have observations in I1 and I2 only.

As we discussed in the previous section, in this analysis we make the strong assumption that the 48 observations in the data set BEHAVIOUR are independent of each other. We also assume that the explanatory variables within an interval are solely responsible for the outcome within that interval. A logistic regression model was tried, considering newly defined PRED as the binary response variable. We have 41 observations corresponding to fledgings and only 7 observations corresponding to predations in this analysis. Considering the possible time effect, we included INTE as an explanatory variable. Since BOTH and ONE are highly correlated, it is satisfactory to include only one of them in the model. We have chosen BOTH as it seems biologically preferable to ONE. The deviance decreases and the corresponding p-values associated with the variables are given in table 3.5.2.

Variable	Deviance decrease	p-value
INTE	0.9719	0.32
VIS	0.102	0.75
NVIS	0.252	0.62
NOIP	2.1287	0.14
DNOI	1.8499	0.17
AG	0.213	0.64
BOTH	0.314	0.58

Table 3.5.2

These p-values imply that none of the variables in BEHAVIOUR are significantly related to the probability of predation. If anything there seems to be some mild evidence that noisier kingbirds may encourage predation. The binary variable PRED may not be reflecting the variation contained in the explanatory variables well. In a further attempt to explore the problem, we are reminded that the time prior to the predation may contain additional information. In the next section, we carry out an analysis using the time prior to predation as our response variable.

3.6 SURVIVAL ANALYSIS OF BEHAVIOUR

Survival analysis is a loosely defined statistical term that encompasses a variety of statistical techniques for analysing positive valued random variables. Typically, the value of the random variable is the time to the death of a biological unit or the time to the failure of a physical component. Time in such a case is generally referred to as 'survival time' and we denote it by T . In general, survival time carries more information than the simple observation of whether or not a death or a failure has occurred. Survival analysis techniques are employed to extract information concerning covariates using the survival time.

Let $T \geq 0$ have the density $f(t)$ and the distribution function $F(t)$. The survival function $S(t)$ (i.e. probability of survival up to time t) is,

$$S(t) = 1 - F(t) = \Pr(T > t)$$

and the hazard rate or the hazard function $h(t)$ is

$$h(t) = \frac{f(t)}{1 - F(t)}$$

The hazard rate can be interpreted as the instantaneous failure rate at time t given survival up to time t . That is;

$$h(t)dt = P\{t < T < t + dt \mid T > t\}.$$

It can then be easily shown that

$$S(t) = \exp\left(-\int_0^t h(u)du\right).$$

A characteristic feature of survival data is the occurrence of censored observations, i.e. observations on which the complete survival time is not observed. Censoring is usually on the right; that is the observed censored time is less than the actual survival time. Sometimes the distribution of T is known and then parametric models are used to model the relationship between the hazard and the covariates. When the distribution of T is unknown, non-parametric methods have to be used. Cox's proportional hazards model is a widely used semi-parametric model used to

describe the relationship between the hazard function and the covariates when the distribution of T is unknown.

Let x be the vector of covariates. Under the Cox proportional hazards model, the hazard function is,

$$h(t) = \lambda(t)\exp(x'\beta)$$

where $\lambda(t)$ is an arbitrary unspecified function of t ; the base-line hazard when the covariate $x = 0$. The unknown parameter vector β has components which are the coefficients of the covariates of interest.

Let us now consider our problem. The predation in our problem is equivalent to the death or the failure in the usual framework of survival analysis. Let T be the time prior to the predation and it is equivalent to the survival time. We have four intervals in which observations were collected. We assume that the outcome in an interval depends solely on the corresponding covariates in the interval. We noticed in section 3.4 that there is no evidence of dependence between intervals. In this analysis we assume that the base-line hazards in intervals I_1, I_2, I_3 and I_4 are $\lambda_1, \lambda_2, \lambda_3$ and λ_4 respectively and that they are independent.

In any of the first three intervals the outcome is either a predation or a non-predation. When it is a predation, the corresponding survival time is the time prior to the predation within the interval. If it is a non-predation, we consider it as a censored observation with censored time equal to the length of the interval. In the fourth interval, the outcome is either a predation or a fledge. The predation can be considered as before. The case of fledge can be considered as non-predation in other intervals and the time for the fourth interval is censored at 5 days.

Let h_{ij} be the hazard function of the i -th nest in the j -th interval. Our model is given by

$$h_{ij}(t) = \lambda_j \exp(x_{ij}'\beta) \tag{3.6.1}$$

where x_{ij} is the vector of covariates for the i -th nest in the j -th interval and β is the vector of coefficients of the covariates. Again, making the strong assumption that the 48 observations can be considered independent as in section 3.5, the marginal likelihood of β can be written as

$$L(\beta) = \prod_{j=1}^4 L_j(\beta),$$

where $L_j(\beta)$ is the marginal likelihood arising from the j -th interval. Kalbfleisch and Prentice[7] have extensively discussed the estimation of β maximizing the marginal likelihood function. The program that we used to estimate β is a modification of several sub-routines provided in Kalbfleisch and Prentice[7] and Forsythe, Malcolm and Moler[6]. It is available upon request.

Let IND be the censoring indicator.

IND =1 if t is uncensored
 =0 if t is censored.

INTE											
I1			I2			I3			I4		
NEST	T	IND	NEST	T	IND	NEST	T	IND	NEST	T	IND
39	6	1	76	3	1	75	3	1	79	5	0
75	6	0	79	3	0	79	4	0	80	5	0
76	6	0	81	3	0	80	4	0	81	3	1
77	6	0	83	3	0	81	4	0	83	5	0
79	6	0	85	3	0	83	4	0	86	5	0
81	6	0	86	3	0	86	4	0	93	5	0
82	6	1	88	3	0	93	4	0	100	5	0
83	6	0	91	3	1	103	4	0	103	5	0
85	6	0	93	3	1						
86	6	0	94	3	0						
87	6	0	96	3	0						
91	6	0	99	3	0						
92	6	0	103	3	0						
93	6	0									
94	6	0									
96	6	0									
97	6	0									
98	6	0									
103	6	0									

Table 3.6.1

Table 3.6.1 shows the survival times and the corresponding censoring indicator for each observation as used in the analysis. We fitted model 3.6.1 with 7 explanatory variables, one at a time. Estimated values of β were found by maximizing the marginal likelihood function. Table 3.6.2 contains the estimated coefficients and the corresponding approximate p-values associated with the explanatory variables.

Variable	Estimated Parameter	Estimated Std Dev	Approximate p-value
YNG	-0.41	0.37	0.13
VIS	0.11	0.05	0.02
NVIS	0.09	0.09	0.16
NOIP	-0.005	0.03	0.41
DNOI	0.001	0.02	0.46
AG	0.06	0.25	0.42
BOTH	-0.01	0.01	0.19

Table 3.6.2

The Cox proportional hazards model indicates that the variable VIS, the number of visits by parent birds may be an influential behavioural variable. The positive coefficient implies that the hazard and hence the probability of predation increases with VIS.

3.7 SUMMARY OF THE ANALYSIS OF BEHAVIOUR

Both the logistic regression analysis and the survival analysis assume the independence of the 48 observations. Observations on different nests were assumed to be independent because the nests were reasonably far apart. Observations on the same nest in different intervals were assumed independent merely because there was not enough evidence of dependence from the graphs based on the 13 nests. The logistic regression model was unable to identify any significant variables. VIS, the number of visits by parent birds was found to increase the probability of predation, by the survival analysis approach. For this result it is difficult to find a biological explanation. Further more, the figures 3.4.1 and 3.4.2 do not show any difference of the

variable VIS in preyed upon and fledged groups. It seems to us that nest vigilance and protection has little to do with the actual number of visits by kingbirds and more to do with the percentage of time that the nest was protected. We would have expected BOTH to be a significant variable. We emphasize again that the validity of the result depends on the independence of the observations on the same nests and the integrity of the data.

In addition to the odd result involving VIS, the reader may notice that the p-value for VIS using survival methods (0.02) differ markedly from the p-value for VIS using logistic regression (0.75). The most likely explanation for this phenomena is that the model (3.6.1) is incorrect. Although (3.6.1) is semi-parametric it imposes a great deal of structure. Besides the strong independence assumption (3.6.1) requires that the proportionality assumption in the hazard behaves the same way across intervals even though the intervals are of different lengths. Also testing is based on asymptotic normality. With so few observations it is doubtful that normality has “kicked in”. For all of these reasons we will conclude that the data gives no evidence linking behavioural variables with predation.

Chapter 4

CONCLUDING REMARKS

1. As far as the visibility of the nest is concerned, only the overhead visibility turned out to be significant. The interesting feature is that in contrast to what one would expect, overhead visibility decreases the possibility of predation.
2. Among the environmental factors considered, only the availability of water and trees around the nest are found to have significant effects on nest predation.
 - (a) Regardless of overhead visibility, availability of water reduces the possibility of predation.
 - (b) When water is present, tree density makes a difference in predation only when the overhead visibility is less than 80 percent. On the other hand, when there is little water, tree density makes a difference only when the overhead visibility is greater than 80 percent. In both cases the effect of more trees is to decrease the possibility of predation.
3. The analysis on the behavioural factors reveals that out of 7 variables considered, none are influential in predicting predation.

We now offer some final remarks:

- (1) Ms. Siderius had never witnessed a predation during her monitoring period. This may suggest that she had not monitored the nests at the ideal times. Furthermore, 90 minutes may be a too small fraction of time for monitoring the behaviour of parent kingbirds. Therefore the variables available in BEHAVIOUR may not reflect the true effects.
- (2) The analysis on both data sets was based on the assumption of independence between observations. In the case of ENVIRONMENT, the assumption of independence may be justified by the fact that nests are considerably far apart

and kingbirds do not live in groups. However, in the case of BEHAVIOUR the assumption of independence is doubtful because we have several observations on the same nest.

- (3) As mentioned in section 3.6 the survival model (3.6.1) is highly questionable. Inferences obtained from this model should be regarded cautiously. Despite this, we believe that the modelling exercise was worthwhile since the weakness of the data did not yield any conclusions using less demanding models. As a general rule, the more that one assumes in a model, the more that one can infer.
- (4) Most importantly, as discussed in the running commentary in chapter 2, the data is at times unreliable and sparse. With more complete data more convincing results may have been obtained.

APPENDIX 1

Table 5.1.1-ENVIRONMENT

ROW	NEST	PRED	OVER	PN	PS	PE	PW	BN	BS	BE	BW	WN
1	1	1	0	0	0	0	0	0	0	0	0	0
2	2	0	80	0	0	5	0	10	0	33	0	38
3	3	0	70	0	0	0	0	100	0	100	50	100
4	4	0	95	0	50	0	0	0	67	0	0	0
5	5	0	0	5	0	0	0	33	43	0	0	25
6	6	1	0	0	0	0	0	0	0	0	0	0
7	7	1	75	*	*	*	*	100	0	33	33	100
8	8	1	30	0	30	0	0	0	0	0	7	0
9	9	0	80	80	80	20	10	100	100	33	67	100
10	10	0	50	0	0	0	40	67	0	0	17	38
11	11	0	30	0	0	0	90	*	*	*	*	*
12	12	1	90	0	90	80	5	0	67	83	50	13
13	13	1	80	0	0	0	0	*	67	17	*	*
14	14	1	50	0	0	0	40	33	0	0	17	50
15	15	0	25	5	30	0	10	0	50	0	10	13
16	16	0	95	20	80	40	60	50	83	100	50	38
17	17	0	80	0	10	30	0	0	67	67	0	0
18	20	0	65	80	0	30	0	100	100	100	0	100
19	21	0	20	0	0	0	0	0	67	33	0	0
20	22	0	20	0	0	0	0	3	0	100	0	3
21	25	0	95	70	0	0	30	10	0	0	67	45
22	26	0	10	5	0	0	50	0	0	0	0	8
23	27	1	60	0	0	0	0	27	93	33	43	38
24	28	0	0	0	0	0	0	0	67	0	0	0
25	29	1	100	0	0	0	0	0	0	0	0	0
26	30	1	5	10	0	0	0	93	0	0	0	0
27	31	0	85	100	0	0	40	100	100	100	100	100
28	32	0	100	100	100	80	30	100	100	100	100	100
29	33	1	5	0	2	5	2	0	0	0	60	0
30	34	1	60	75	15	5	10	77	67	10	0	75
31	35	1	95	0	98	15	0	0	100	3	0	0
32	36	1	0	0	80	0	0	0	100	0	0	0
33	37	0	40	0	0	0	0	0	0	0	0	0
34	39	1	0	0	0	0	0	67	0	17	0	25
35	41	1	75	0	80	80	10	100	100	0	100	100
36	42	1	0	0	0	0	0	*	*	*	*	*
37	43	1	0	0	10	0	0	0	0	0	0	0
38	44	1	40	100	100	75	100	100	100	83	90	100
39	45	0	20	0	0	0	0	0	50	0	0	0
40	46	1	20	80	0	0	0	67	0	33	0	50
41	47	1	15	0	0	0	75	0	0	0	27	0
42	48	0	5	5	0	0	0	0	67	0	0	0
43	49	1	40	50	0	90	50	67	0	17	33	38
44	50	0	0	0	0	0	0	0	0	0	0	0
45	51	0	0	0	0	0	0	33	0	17	0	38
46	52	0	85	90	20	0	75	33	50	0	10	75
47	53	0	0	100	50	0	0	100	17	0	0	100
48	54	0	85	25	75	20	75	33	67	33	50	50
49	55	0	30	0	0	0	0	0	0	0	0	0
50	56	1	20	0	0	0	0	*	*	*	*	*
51	57	0	100	10	0	0	0	100	100	100	100	100

	NEST	PRED	OVER	PN	PS	PE	PW	BN	BS	BE	BW	WN
52	59	0	90	0	0	0	0	33	0	0	17	38
53	60	1	10	0	10	0	10	0	0	33	33	13
54	61	0	0	0	0	0	0	*	*	*	*	*
55	62	0	100	0	25	0	40	0	50	0	83	0
56	63	0	85	0	20	30	30	0	33	67	67	25
57	64	0	25	0	0	5	0	0	0	43	0	0
58	65	0	90	0	5	0	0	0	23	0	0	13
59	66	0	10	100	0	50	0	33	0	50	0	75
60	67	1	40	0	0	0	0	0	0	0	0	0
61	68	1	25	*	*	*	*	100	0	100	67	100
62	69	1	3	0	0	0	0	0	0	0	0	0
63	70	1	50	15	0	0	0	17	0	0	33	25
64	71	1	85	90	70	0	80	83	17	10	67	75
65	72	1	25	0	0	0	0	33	0	0	0	33
66	73	1	45	*	*	*	*	43	0	17	0	63
67	74	1	70	30	0	5	0	*	*	*	*	*
68	75	1	40	0	100	0	100	0	100	0	100	25
69	76	1	0	0	0	0	0	0	0	0	0	0
70	77	0	90	0	0	35	10	83	50	83	0	45
71	78	1	5	0	0	0	0	0	0	0	0	0
72	79	0	100	0	0	0	0	100	100	100	100	100
73	80	0	90	30	35	0	0	33	67	0	0	50
74	81	1	55	90	0	0	30	100	0	0	83	100
75	82	1	0	60	10	50	0	100	0	67	0	88
76	83	0	100	80	0	55	40	100	10	100	100	75
77	84	1	5	0	0	0	0	0	0	0	0	0
78	85	0	95	0	30	0	40	0	83	17	50	25
79	86	0	60	60	0	5	0	33	0	0	0	50
80	87	0	0	90	0	30	0	100	0	67	33	100
81	88	0	10	0	0	0	0	0	0	0	0	0
82	89	1	15	100	0	0	0	100	0	0	0	100
83	90	1	85	0	0	0	50	*	*	*	*	*
84	91	1	90	80	100	0	100	43	100	0	77	50
85	92	1	30	0	0	0	0	0	33	0	0	0
86	93	0	0	0	0	0	0	0	0	0	0	0
87	94	1	0	0	0	0	0	0	0	0	0	0
88	96	1	85	*	*	*	*	0	33	33	0	0
89	97	1	80	90	20	5	60	83	33	17	10	88
90	98	1	25	10	0	50	0	67	0	17	0	75
91	99	1	95	10	5	0	15	83	83	0	100	63
92	100	1	50	0	0	10	10	0	0	100	17	25
93	101	1	30	100	0	100	100	100	0	100	100	100
94	102	1	40	0	0	60	80	33	0	100	100	33
95	103	0	98	50	60	90	70	67	83	100	67	75

ROW	PRED	WS	WE	WW	DN	DS	DE	DW		NTRE	STEM	
1	1	0	0	0	40	40	0	6	36	114	300	0
2	0	0	33	0	100	15	200	7	228	282	20	5
3	0	0	100	50	500	0	200	100	169	180	40	6
4	0	100	0	0	1	50	6	0	59	83	150	110
5	0	33	0	0	100	0	0	50	106	239	400	3
6	1	0	0	13	500	20	30	0	88	236	400	3
7	1	0	38	38	150	500	500	500	213	413	50	0
8	1	0	0	70	500	6	2	300	212	290	50	112
9	0	100	63	50	0	200	30	500	134	*	30	25
10	0	0	0	25	500	10	8	18	181	278	30	7
11	0	*	*	*	*	*	*	*	4	5	10	15
12	1	75	88	75	0	30	10	1	103	*	500	0
13	1	83	88	*	500	500	500	5	35	74	500	3
14	1	0	0	50	500	4	15	200	78	122	60	24
15	0	50	0	25	500	0	0	100	69	106	300	0
16	0	88	100	75	0	700	700	15	65	236	20	0
17	0	75	50	0	1	200	300	1	56	104	200	38
18	0	88	95	0	*	*	*	*	67	157	9	4
19	0	70	43	0	*	*	*	*	47	118	8	46
20	0	0	100	0	2	100	150	3	176	394	35	11
21	0	0	0	88	80	15	5	500	202	256	15	97
22	0	0	20	0	60	6	18	40	33	315	50	0
23	1	83	50	58	8	300	500	7	37	67	150	0
24	0	75	0	0	3	35	1	12	42	197	40	0
25	1	0	0	0	8	8	7	3	24	472	12	12
26	1	0	0	0	5	0	40	100	54	256	25	112
27	0	100	100	100	*	*	*	*	95	135	25	42
28	0	100	100	100	*	*	*	*	100	159	6	27
29	1	0	63	13	*	*	*	*	36	75	50	4
30	1	75	8	0	500	300	4	500	111	210	40	1
31	1	100	8	0	*	*	*	*	92	147	16	7
32	1	50	0	0	*	*	*	*	82	*	30	*
33	0	0	0	0	*	*	*	*	111	*	160	*
34	1	0	13	0	*	*	*	*	*	*	*	*
35	1	100	0	50	*	*	*	*	138	*	120	*
36	1	*	*	*	*	*	*	*	141	*	10	*
37	1	0	0	0	*	*	*	*	113	*	*	*
38	1	100	75	93	*	*	*	*	43	*	20	*
39	0	50	13	0	*	*	*	*	76	*	100	*
40	1	0	25	0	*	*	*	*	41	*	4	*
41	1	13	0	38	*	*	*	*	29	*	100	*
42	0	38	0	0	*	*	*	*	41	*	150	*
43	1	0	25	50	*	*	*	*	73	*	25	*
44	0	0	0	0	*	*	*	*	42	*	*	*
45	0	0	13	0	*	*	*	*	108	*	*	*
46	0	63	75	100	*	*	*	*	107	*	30	*
47	0	25	0	0	*	*	*	*	*	*	*	*
48	0	75	50	63	*	*	*	*	59	*	150	*
49	0	0	0	0	*	*	*	*	*	*	*	*
50	1	*	*	*	*	*	*	*	*	*	*	*

ROW	PRED	WS	WE	WW	DN	DS	DE	DW	NH	TH	NTRY	STEM
51	0	100	100	100	*	*	*	*	94	*	50	*
52	0	38	0	38	*	*	*	*	42	*	175	*
53	1	0	0	33	*	*	*	*	53	*	60	*
54	0	*	*	*	*	*	*	*	7	*	*	*
55	0	50	0	63	*	*	*	*	233	*	4	*
56	0	50	63	100	*	*	*	*	105	*	1000	*
57	0	0	63	0	*	*	*	*	94	*	*	*
58	0	38	0	0	*	*	*	*	55	*	60	*
59	0	0	63	63	*	*	*	*	20	*	250	*
60	1	0	0	0	*	*	*	*	39	*	*	*
61	1	13	88	50	*	*	*	*	48	*	*	*
62	1	0	0	0	*	*	*	*	53	*	*	*
63	1	0	0	50	*	*	*	*	50	*	*	*
64	1	100	25	50	*	*	*	*	117	*	100	*
65	1	0	0	0	*	*	*	*	117	*	15	*
66	1	0	63	0	*	*	*	*	*	*	*	*
67	1	*	*	*	*	*	*	*	142	*	*	*
68	1	100	13	100	200	300	1000	200	96	156	100	18
69	1	0	0	0	1	5	360	2	59	93	110	33
70	0	63	75	0	300	100	20	1	122	171	15	40
71	1	0	0	0	1000	1000	2	20	156	221	*	0
72	0	100	100	100	300	1000	200	1000	98	102	45	0
73	0	75	0	0	70	70	20	6	95	*	90	29
74	1	0	25	63	1000	2	1	6	43	130	130	60
75	1	0	88	0	400	100	1000	7	53	102	200	69
76	0	13	100	100	50	30	30	50	33	89	30	7
77	1	0	0	0	1000	1	1	1	83	138	150	98
78	0	88	25	63	100	0	100	1000	146	229	15	4
79	0	0	13	0	1000	200	300	6	136	201	9	2
80	0	0	63	50	70	5	1000	1000	96	192	70	10
81	0	0	0	0	5	4	1000	4	75	142	12	31
82	1	0	0	0	1000	100	200	400	109	129	70	58
83	1	*	*	*	10	10	1000	1000	158	173	*	33
84	1	100	0	75	50	10	2	20	107	172	110	0
85	1	50	0	0	1000	1	7	700	253	*	7	1
86	0	0	0	0	100	1	1000	1000	*	*	10	0
87	1	0	0	0	20	300	1	7	127	190	200	78
88	1	50	50	25	8	700	8	700	58	129	700	20
89	1	25	25	88	50	50	60	1000	108	120	50	5
90	1	0	58	0	30	30	100	1	85	154	25	56
91	1	75	0	88	1000	800	20	700	71	93	700	0
92	1	0	75	8	1000	1000	1000	1	64	88	88	*
93	1	0	100	100	1000	200	15	0	*	*	50	3
94	1	0	100	88	1	2	50	50	68	104	20	39
95	0	63	100	75	100	600	100	1000	88	175	100	10

ROW	PRED	NPER	WAT	FIELD	SHRUB	MARSH	TREE	DWAT	YEAR
1	1	10	30	65	5	0	0	0	1
2	0	7	25	40	10	0	25	25	1
3	0	15	15	50	10	0	20	25	1
4	0	25	15	0	35	50	0	0	1
5	0	200	20	40	0	0	40	50	1
6	1	250	20	40	0	0	40	20	1
7	1	50	25	60	0	0	15	800	1
8	1	6	30	20	30	20	0	0	1
9	0	*	70	0	30	0	0	100	1
10	0	7	50	0	50	0	0	0	1
11	0	1	10	30	30	0	30	300	1
12	1	10	0	90	10	0	0	0	1
13	1	100	0	0	5	40	0	100	1
14	1	3	20	0	30	50	0	200	1
15	0	30	15	75	10	0	0	20	1
16	0	20	0	50	30	0	20	25	1
17	0	100	0	50	50	0	0	200	1
18	0	5	50	0	30	10	10	0	1
19	0	*	50	0	50	0	0	0	1
20	0	3	60	10	20	0	20	0	1
21	0	7	5	0	80	0	10	30	1
22	0	6	30	30	10	0	30	0	1
23	1	20	10	90	0	0	0	70	1
24	0	50	40	20	30	0	10	0	1
25	1	4	2	70	28	0	0	0	1
26	1	3	10	60	15	0	15	10	1
27	0	4	0	30	10	0	60	30	1
28	0	1	5	10	15	0	70	40	1
29	1	5	0	80	19	0	1	25	1
30	1	3	30	10	0	0	60	50	1
31	1	3	1	50	45	0	5	40	1
32	1	*	30	0	40	20	10	0	2
33	0	*	50	40	10	0	0	0	2
34	1	*	*	*	*	*	*	*	2
35	1	*	50	20	20	0	5	0	2
36	1	*	50	0	35	0	10	0	2
37	1	*	50	0	35	0	10	0	2
38	1	*	0	40	5	50	0	*	2
39	0	*	60	10	25	0	0	0	2
40	1	*	50	0	40	0	5	0	2
41	1	*	40	20	40	0	0	0	2
42	0	*	10	0	40	30	0	0	2
43	1	*	0	30	50	0	20	0	2
44	0	*	40	35	7	0	8	0	2
45	0	*	20	40	0	0	40	*	2
46	0	*	10	70	10	0	10	*	2
47	0	*	*	*	*	*	*	*	2
48	0	*	30	30	30	0	10	*	2
49	0	*	10	30	50	0	10	*	2
50	1	*	5	55	20	0	20	*	2

ROW	PRED	NPER	WAT	FIELD	SHRUB	MARSH	TREE	DWAT	YEAR
51	0	*	5	70	15	0	10	*	2
52	0	*	0	75	25	0	0	*	2
53	1	*	0	50	50	0	0	*	2
54	0	*	20	70	5	0	5	*	2
55	0	*	10	40	0	0	50	*	2
56	0	*	0	80	0	0	20	100	2
57	0	*	10	0	50	40	0	*	2
58	0	*	15	0	5	80	0	*	2
59	0	*	20	0	5	75	0	0	2
60	1	*	*	*	*	*	*	*	2
61	1	*	0	80	20	0	0	*	2
62	1	*	0	80	15	0	5	*	2
63	1	*	0	80	15	0	5	*	2
64	1	*	0	80	17	0	3	*	2
65	1	*	15	80	0	0	0	*	2
66	1	*	55	5	40	0	0	0	2
67	1	*	0	0	40	60	0	0	2
68	1	2	30	65	5	0	0	1	3
69	1	2	0	80	20	0	0	120	3
70	0	10	0	60	20	0	0	15	3
71	1	15	0	60	40	0	0	400	3
72	0	6	10	75	15	0	0	40	3
73	0	4	40	10	40	5	5	0	3
74	1	3	0	0	40	40	20	0	3
75	1	2	40	0	40	20	0	0	3
76	0	3	50	0	30	0	20	0	3
77	1	*	30	0	20	50	0	5	3
78	0	15	5	50	0	0	45	60	3
79	0	30	30	60	10	0	0	0	3
80	0	5	40	15	25	0	20	0	3
81	0	2	50	20	30	0	0	2	3
82	1	3	30	20	50	0	0	8	3
83	1	7	10	0	70	10	10	25	3
84	1	7	0	50	50	0	0	120	3
85	1	5	5	80	0	0	15	90	3
86	0	100	30	50	10	0	10	10	3
87	1	1	20	10	60	10	0	8	3
88	1	2	0	85	15	0	0	70	3
89	1	6	10	60	10	0	20	80	3
90	1	25	5	75	15	0	0	100	3
91	1	*	0	90	10	0	0	100	3
92	1	1	10	30	10	0	0	*	3
93	1	50	40	30	30	0	0	0	3
94	1	2	40	40	15	0	0	3	3
95	0	3	20	65	0	0	25	60	3

Table 5.1.2- Created Variables for ENVIRONMENT

ROW	PRED	VN	VS	VE	VW	AV	MV	NS1	NS2
1	1	0	0	0	0	0	0	0	0
2	0	16	0	24	0	10	24	2	0
3	0	67	0	67	33	42	67	3	1
4	0	0	72	0	0	18	72	0	0
5	0	21	25	0	0	12	25	1	0
6	1	0	0	0	4	1	4	1	1
7	1	100	0	36	36	43	100	4	3
8	1	0	10	0	26	9	26	2	2
9	0	93	93	39	42	67	93	2	1
10	0	35	0	0	27	16	35	1	1
11	0	0	0	0	90	23	90	*	*
12	1	4	77	84	43	52	84	0	0
13	1	0	50	35	0	21	50	3	3
14	1	28	0	0	36	16	36	2	1
15	0	6	43	0	15	16	43	2	1
16	0	36	84	80	62	66	84	2	2
17	0	0	51	49	0	25	51	2	1
18	0	93	63	75	0	58	93	*	*
19	0	0	46	25	0	18	46	*	*
20	0	2	0	67	0	17	67	2	0
21	0	42	0	0	62	26	62	1	1
22	0	4	0	7	17	7	17	0	0
23	1	22	59	28	34	36	59	2	2
24	0	0	47	0	0	12	47	0	0
25	1	0	0	0	0	0	0	0	0
26	1	34	0	0	0	9	34	1	0
27	0	100	67	67	80	79	100	*	*
28	0	100	100	93	77	93	100	*	*
29	1	0	1	23	25	12	25	*	*
30	1	76	52	8	3	35	76	3	3
31	1	0	99	9	0	27	99	*	*
32	1	0	77	0	0	19	77	*	*
33	0	0	0	0	0	0	0	*	*
34	1	31	0	10	0	10	31	*	*
35	1	67	93	27	53	60	93	*	*
36	1	0	0	0	0	0	0	*	*
37	1	0	3	0	0	1	3	*	*
38	1	100	100	78	94	93	100	*	*
39	0	0	33	4	0	9	33	*	*
40	1	66	0	19	0	21	66	*	*
41	1	0	4	0	47	13	47	*	*
42	0	2	35	0	0	9	35	*	*
43	1	52	0	44	44	35	52	*	*
44	0	0	0	0	0	0	0	*	*
45	0	24	0	10	0	9	24	*	*
46	0	66	44	25	62	49	66	*	*
47	0	100	31	0	0	33	100	*	*
48	0	36	72	34	63	51	72	*	*
49	0	0	0	0	0	0	0	*	*
50	1	0	0	0	0	0	0	*	*

ROW	PRED	VS	VE	VW	AV	MV	NS1	NS2
51	0	70	67	67	67	68	70	*
52	0	24	13	0	18	14	24	*
53	1	4	3	11	25	11	25	*
54	0	0	0	0	0	0	0	*
55	0	0	42	0	62	26	62	*
56	0	8	34	53	66	40	66	*
57	0	0	0	37	0	9	37	*
58	0	4	22	0	0	7	22	*
59	0	69	0	54	21	36	69	*
60	1	0	0	0	0	0	0	*
61	1	100	7	94	59	65	100	*
62	1	0	0	0	0	0	0	*
63	1	19	0	0	28	12	28	*
64	1	83	62	12	66	56	83	*
65	1	22	0	0	0	6	22	*
66	1	53	0	40	0	23	53	*
67	1	30	0	5	0	9	30	*
68	1	8	100	4	100	53	100	4
69	1	0	0	0	0	0	0	1
70	0	43	38	64	3	37	64	2
71	1	0	0	0	0	0	0	2
72	0	67	67	67	67	67	67	4
73	0	38	59	0	0	24	59	0
74	1	97	0	8	59	41	97	1
75	1	83	3	68	0	39	83	3
76	0	85	8	85	80	65	85	0
77	1	0	0	0	0	0	0	1
78	0	8	67	14	51	35	67	3
79	0	48	0	6	0	14	48	3
80	0	97	0	53	28	45	97	2
81	0	0	0	0	0	0	0	1
82	1	100	0	0	0	25	100	4
83	1	0	0	0	50	13	50	2
84	1	58	100	0	84	61	100	0
85	1	0	28	0	0	7	28	2
86	0	0	0	0	0	0	0	3
87	1	0	0	0	0	0	0	1
88	1	0	42	42	13	24	42	2
89	1	87	26	16	53	46	87	1
90	1	51	0	42	0	23	51	1
91	1	52	54	0	68	44	68	3
92	1	8	0	62	12	21	62	3
93	1	100	0	100	100	75	100	2
94	1	22	0	87	89	50	89	0
95	0	64	69	97	71	75	97	4

Table 5.1.3-BEHAVIOUR

ROW	NEST	INTE	PRED	YNG	VIS	NVIS	NOIP	DNOI	SDNOI	AG	SEE	BOTH	ONE
1	39	2	1	3	16	6	12	3	5	3	1	69	9
2	75	2	1	4	10	1	4	31	39	1	4	65	23
3	75	4	1	4	25	3	7	12	11	2	6	60	27
4	76	2	1	2	12	1	11	12	3	1	3	66	24
5	76	3	1	2	9	1	12	22	25	1	1	68	22
6	77	2	0	4	10	5	17	13	26	0	2	90	0
7	79	2	0	4	11	7	19	12	17	0	0	69	21
8	79	3	0	4	11	4	17	25	25	0	1	63	27
9	79	4	0	4	18	2	25	12	12	1	2	50	40
10	79	5	0	4	8	0	9	44	25	1	3	49	41
11	80	4	0	2	17	16	62	12	13	2	3	90	0
12	80	5	0	2	7	3	33	23	37	0	1	78	12
13	81	2	1	4	15	5	19	12	16	2	1	57	33
14	81	3	1	4	29	5	10	7	7	1	1	39	50
15	81	4	1	4	16	3	12	6	8	1	1	30	60
16	81	5	1	4	21	9	16	29	56	2	2	31	59
17	82	2	1	4	9	1	14	21	38	0	0	74	16
18	83	2	0	4	13	7	27	12	42	0	1	77	13
19	83	3	0	4	5	1	14	25	27	1	3	46	44
20	83	4	0	3	7	6	12	18	28	1	3	88	2
21	83	5	0	3	19	13	20	25	60	0	3	16	73
22	85	2	0	4	15	1	54	23	15	0	3	90	0
23	85	3	0	4	16	1	17	70	37	1	0	22	64
24	86	2	0	4	10	1	9	39	34	0	0	56	34
25	86	3	0	4	12	0	16	19	29	0	0	49	41
26	86	4	0	4	20	0	4	7	12	3	1	76	14
27	86	5	0	4	28	2	13	32	51	4	3	61	29
28	87	2	0	3	11	1	9	52	64	3	2	28	55
29	88	3	0	4	17	0	2	35	21	1	0	19	70
30	91	2	1	3	10	1	14	14	6	1	1	90	0
31	91	3	1	3	10	0	11	33	24	0	2	64	22
32	92	3	1	3	7	0	9	48	12	0	1	77	13
33	92	4	1	3	16	2	9	60	13	0	1	61	29
34	93	2	0	4	10	1	4	1	7	3	2	62	28
35	93	4	0	4	24	0	4	20	7	3	0	78	12
36	93	5	0	4	37	6	6	1	2	0	0	89	2
37	94	2	1	3	7	3	18	8	16	0	0	90	0
38	94	3	1	3	19	0	7	32	57	1	0	82	7
39	96	2	1	3	8	1	13	8	13	0	1	76	14
40	96	3	1	2	10	6	30	22	23	0	0	76	14
41	97	2	0	4	9	0	1	9	*	1	0	51	39
42	98	2	0	3	4	3	27	20	37	0	1	86	4
43	99	3	0	4	10	6	9	16	26	0	0	81	9
44	100	5	0	3	22	12	28	8	19	0	0	20	65
45	103	2	0	3	10	3	27	6	7	1	1	66	24
46	103	3	0	3	21	1	13	24	27	0	0	88	2
47	103	4	0	3	35	2	10	12	9	1	0	51	34
48	103	5	0	3	37	7	38	17	20	3	1	78	12

Table 5.1.4-Dates of visits for 9 preyed upon nests in BEHAVIOUR

NEST	date of first egg	last visit before predation	first visit after predation
39	*	*	*
75	June 15	July 10	July 12
76	June 21	July 12	July 15
81	June 19	July 18	July 20
82	June 20	July 8	July 11
91	June 20	July 12	July 14
92	It is known that the young were preyed upon at about 10 days		
94	June 28	July 22	July 27
96	June 17	July 11	July 16

APPENDIX 2

Table 5.2.1-Correlation matrix of 12 visibility variables associated with north, south, east and west.

	PN	PS	PE	PW	BN	BS	BE	BW
PS	0.242							
PE	0.359	0.356						
PW	0.355	0.412	0.343					
BN	0.673	0.131	0.408	0.226				
BS	0.127	0.632	0.185	0.228	0.186			
BE	0.269	0.098	0.570	0.213	0.464	0.240		
BW	0.306	0.387	0.422	0.587	0.466	0.423	0.468	
WN	0.756	0.201	0.416	0.302	0.908	0.238	0.486	0.520
WS	0.153	0.646	0.157	0.279	0.197	0.942	0.223	0.439
WE	0.320	0.107	0.554	0.255	0.476	0.257	0.908	0.473
WW	0.400	0.357	0.408	0.657	0.396	0.381	0.456	0.857

	WN	WS	WE
WS	0.251		
WE	0.536	0.256	
WW	0.516	0.398	0.491

Table 5.2.2-Correlation matrix of OVER, VN, VS, VE and VW

	OVER	VN	VS	VE
VN	0.271			
VS	0.541	0.221		
VE	0.329	0.523	0.268	
VW	0.482	0.451	0.412	0.466

Table 5.2.3-Correlation matrix of variables in reduced ENVIRONMENT

	OVER	AV	MV	NS1	NS2	NH	TH	NTRE
AV	0.548							
MV	0.468	0.858						
NS1	0.113	0.273	0.306					
NS2	0.117	0.179	0.184	0.809				
NH	0.225	0.009	0.053	0.292	0.154			
TH	0.017	-0.199	-0.208	-0.071	-0.203	0.532		
NTRE	0.084	-0.047	-0.082	-0.012	0.161	-0.174	-0.237	
STEM	-0.136	-0.189	-0.086	-0.179	-0.163	0.030	-0.052	-0.120
NPER	-0.243	-0.221	-0.276	-0.009	-0.020	-0.051	0.093	0.430
WAT	-0.310	-0.090	-0.020	-0.051	-0.184	0.080	0.184	-0.256
FIELD	0.142	0.078	-0.068	0.152	0.103	-0.036	-0.024	0.283
SHRUB	-0.043	-0.131	0.009	-0.265	-0.159	0.012	-0.023	-0.244
MARSH	-0.003	-0.065	-0.015	-0.086	0.018	-0.168	-0.210	0.063
TREE	0.147	0.225	0.179	0.072	-0.027	0.228	0.194	-0.094

	STEM	NPER	WAT	FIELD	SHRUB	MARSH
NPER	-0.196					
WAT	-0.000	-0.023				
FIELD	-0.381	0.023	-0.549			
SHRUB	0.436	-0.248	0.108	-0.508		
MARSH	0.473	-0.000	-0.116	-0.461	0.049	
TREE	-0.176	0.183	-0.090	-0.082	-0.320	-0.222

Table 5.2.4-Description of 15 variables in reduced ENVIRONMENT

	PRED	N	N*	MEAN	MEDIAN	STDEV	MIN	MAX	Q1	Q3
OVER	0	45	0	53.51	65.00	39.53	0.00	100.00	10.00	90.00
	1	50	0	38.86	35.00	32.94	0.00	100.00	5.00	71.25
AV	0	45	0	29.49	23.00	25.24	0.00	93.00	9.00	47.00
	1	50	0	24.34	20.00	23.17	0.00	93.00	4.75	41.50
MV	0	45	0	52.89	62.00	31.51	0.00	100.00	24.50	72.00
	1	50	0	49.20	50.00	36.86	0.00	100.00	17.50	84.75
NS1	0	23	22	1.739	2.000	1.251	0.000	4.000	1.000	3.000
	1	29	21	1.759	2.000	1.244	0.000	4.000	1.000	3.000
NS2	0	23	22	0.957	1.000	0.878	0.000	3.000	0.000	2.000
	1	29	21	1.414	1.000	1.053	0.000	3.000	0.500	2.000
NH	0	42	3	93.24	94.00	55.59	4.00	233.00	53.00	113.75
	1	46	4	89.04	80.00	51.67	24.00	253.00	49.50	114.00
TH	0	25	20	181.8	175.0	85.1	5.0	394.0	112.0	237.5
	1	28	22	166.7	134.0	96.6	67.0	472.0	102.5	205.0
NTRE	0	39	6	100.3	40.0	173.1	4.0	1000.0	15.0	150.0
	1	39	11	137.1	60.0	181.1	4.0	700.0	25.0	150.0
STEM	0	28	17	20.32	8.50	27.74	0.00	110.00	2.25	30.50
	1	30	20	28.30	9.50	35.66	0.00	112.00	0.75	56.50
NPER	0	26	19	25.35	7.00	44.50	1.00	200.00	3.75	26.25
	1	29	21	20.69	5.00	48.86	1.00	250.00	2.50	12.50
WAT	0	44	1	23.75	20.00	19.65	0.00	70.00	10.00	40.00
	1	48	2	16.52	10.00	18.02	0.00	55.00	0.00	30.00
FIELD	0	44	1	32.50	30.00	26.67	0.00	80.00	2.50	50.00
	1	48	2	42.29	45.00	32.65	0.00	90.00	6.25	78.75
SHRUB	0	44	1	21.75	17.50	18.10	0.00	80.00	10.00	30.00
	1	48	2	24.35	20.00	18.04	0.00	70.00	10.00	40.00
MARSH	0	44	1	6.59	0.00	18.83	0.00	80.00	0.00	0.00
	1	48	2	7.71	0.00	16.53	0.00	60.00	0.00	0.00
TREE	0	44	1	14.39	10.00	17.71	0.00	70.00	0.00	20.00
	1	48	2	6.12	0.00	11.49	0.00	60.00	0.00	10.00

N*-number of missing values

Q1-first quartile

Q3-third quartile

Table 5.2.5-Correlation matrix of variables in BEHAVIOUR

	YNG	VIS	NVIS	NOIP	DNOI	AG	SEE	BOTH
VIS	0.145							
NVIS	-0.219	0.248						
NOIP	-0.354	0.004	0.549					
DNOI	-0.002	-0.170	-0.299	-0.129				
AG	0.109	0.376	-0.034	-0.129	-0.072			
SEE	0.020	-0.045	0.103	0.118	0.034	0.295		
BOTH	-0.256	-0.144	-0.064	0.242	-0.302	-0.126	-0.037	
ONE	0.278	0.133	0.059	-0.236	0.301	0.098	0.034	-0.994

REFERENCES

1. Aitkin, M., Anderson, D., Francis, B. and Hinde, J. (1989). *Statistical Modelling in GLIM*. Clarendon Press, Oxford
2. Bent, A.C. (1942). *Life Histories of North American Flycatchers, Larks, Swallows and their allies*. United States Govt. Print. Off. Washington.
4. Cox, D.R. (1972). Regression models and life tables (with discussion). *J. Roy. Statist. Soc. B*, 34, 187-220.
5. Cox, D.R. and Oakes, D. (1984). *Analysis of Survival Data*. Chapman and Hall.
5. Dobsen, A.J. (1990). *An Introduction to Generalized Linear Models*. Chapman and Hall.
6. Forsythe, G.E., Malcolm, M.A. and Moler, C.B. (1977). *Computer Methods for Mathematical Computations*. Prentice-Hall.
7. Kalbfleisch, J.D. and Prentice, R.L. (1980). *The Statistical Analysis of Failure Time Data*. John Wiley and Sons.
8. McCullagh, P. and Nelder, J.A. (1983). *Generalized Linear Models*. Chapman and Hall.
9. Miller, R.G. (1981). *Survival Analysis*. John Wiley and Sons.