

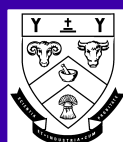
Using Flexi to detect a trend in count and binary longitudinal data

Jim Young

Research Report No: 96/12
October 1996

RESEARCH REPORT

ISSN 1173-8405



Centre for Computing and Biometrics

The Centre for Computing and Biometrics (CCB) has both an academic (teaching and research) role and a computer services role. The academic section teaches subjects leading to a Bachelor of Applied Computing degree and a computing major in the BCM degree. In addition it contributes computing, statistics and mathematics subjects to a wide range of other Lincoln University degrees. The CCB is also strongly involved in postgraduate teaching leading to honours, masters and PhD degrees. The department has active research interests in modelling and simulation, applied statistics and statistical consulting, end user computing, computer assisted learning, networking, geometric modelling, visualisation, databases and information sharing.

The Computer Services section provides and supports the computer facilities used throughout Lincoln University for teaching, research and administration. It is also responsible for the telecommunications services of the University.

Research Report Editors

Every paper appearing in this series has undergone editorial review within the Centre for Computing and Biometrics. Current members of the editorial panel are

Dr Alan McKinnon	Dr Keith Unsworth
Dr Bill Rosenberg	Dr Don Kulasiri
Dr Clare Churcher	Mr Kenneth Choo
Dr Jim Young	

The views expressed in this paper are not necessarily the same as those held by members of the editorial panel. The accuracy of the information presented in this paper is the sole responsibility of the authors.

Copyright

Copyright remains with the authors. Unless otherwise stated, permission to copy for research or teaching purposes is granted on the condition that the authors and the series are given due acknowledgement. Reproduction in any form for purposes other than research or teaching is forbidden unless prior written permission has been obtained from the authors.

Correspondence

This paper represents work to date and may not necessarily form the basis for the authors' final conclusions relating to this topic. It is likely, however, that the paper will appear in some form in a journal or in conference proceedings in the near future. The authors would be pleased to receive correspondence in connection with any of the issues raised in this paper. Please contact the authors either by email or by writing to the address below.

Any correspondence concerning the series should be sent to:

The Editor
Centre for Computing and Biometrics
PO Box 84
Lincoln University
Canterbury, NEW ZEALAND
Email: computing@lincoln.ac.nz

Using Flexi to detect a trend in count and binary longitudinal data

Jim Young

Centre for Computing and Biometrics

Lincoln University

Canterbury, New Zealand

young2@lincoln.ac.nz

1. Abstract

The Department of Conservation uses an aerial transect survey to monitor the number of Hector's dolphins around Banks Peninsula. Flights are made repeatedly along a set of 15 transects. Relative dolphin abundance can be measured by the number of dolphins counted in each transect, or by the presence or absence of dolphins in each transect. At times consecutive flights are only days apart and so consecutive measurements on the same transect are correlated. Flexi, Bayesian software for smoothing time series, can be used to detect a trend in count or binary longitudinal data. Flexi's estimate of the trend can approximate the estimate from a generalised estimating equation model, or show the effect of measurement error.

2. Introduction

Flexi is software for smoothing a time series (Wheeler and Upsdell 1994). It has a generalised linear model framework like that of McCullagh and Nelder (1989 p26-32). The user can choose from a variety of error distributions and link functions. Flexi can be used to detect a trend in longitudinal data. Measurements made at different times on the same subject do not need to be independent; measurements do not need to be normally distributed at each point in time. The generalised estimating equation (GEE) models of Liang and Zeger (1986) are also appropriate for this sort of data. In this paper, I compare results from GEE and Flexi models, using count and binary data from an aerial transect survey of Hector's dolphin.

3. The Hector's dolphin aerial transect survey

Hector's dolphin is a rare species found only in New Zealand waters. The Department of Conservation established a marine sanctuary around Banks Peninsula in November 1988, to reduce the number of Hector's dolphins being caught in set nets (Slooten and Lad 1991). The sanctuary extends from the coast out to four nautical miles offshore.

In 1990, the Department of Conservation began an aerial transect survey to monitor the number of Hector's dolphins in the sanctuary. Fifteen points were picked at random along the sanctuary's coastline. At each point, a transect extends perpendicular from the coast out to sea (Figure 1). Over three months each summer, a plane is used to count the number of Hector's dolphins seen in each of these 15 transects. Sea and cloud condition are also recorded, although flights are made only in light winds. Each flight follows a standard pattern: flights start at the same time each day relative to sunrise; the same flight path is followed at the same speed and altitude (Department of Conservation 1992).

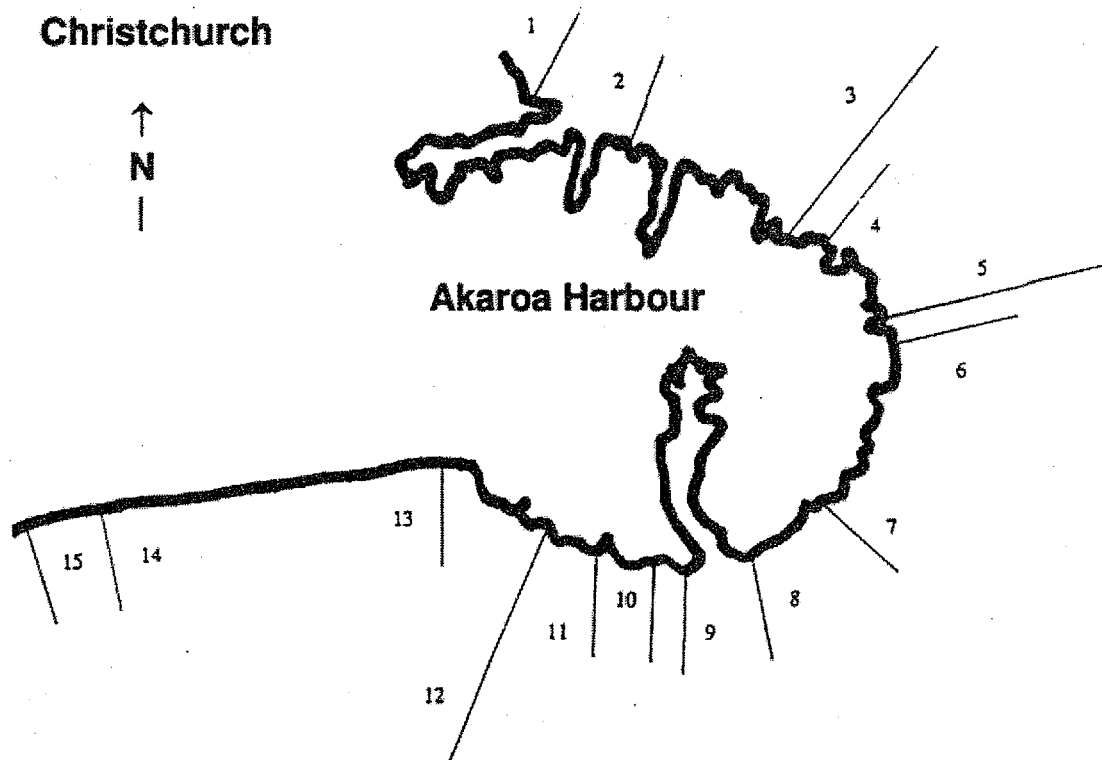


Figure 1. The 15 transects of the Hector's dolphin aerial transect survey.

During the first five years of the survey, ten flights were made each summer. To fit all ten flights in over three months, flying only when conditions were suitable, meant that at times consecutive flights were only days apart. With flights only days apart, counts on the same transect are likely to be correlated. Figure 2 shows data from the first four nautical miles offshore - some (but not all) transects extend out to ten nautical miles - for the first five years of the survey. For each year, the left graph shows the average number of dolphins seen per transect, and the right graph shows the proportion of transects in which dolphins were seen. Each point in Figure 2 is a summary of 150 observations (less three missing observations in 1990). What follows is a more detailed analysis of this data.

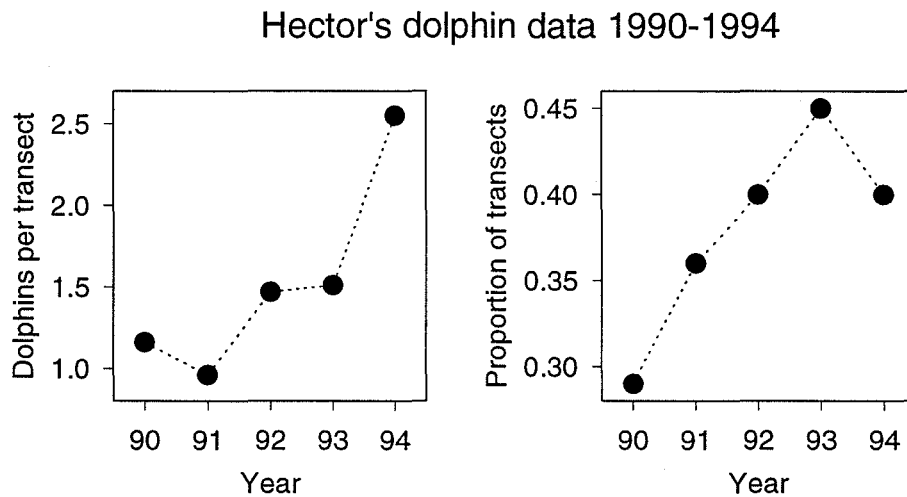


Figure 2. The average number of dolphins seen per transect and the proportion of transects in which dolphins were seen for the first five years of the survey.

The Department of Conservation is naturally interested to know if the number of Hector's dolphins in the sanctuary is increasing or decreasing. To answer this question, I considered a regression model with some measure of dolphin abundance as the response variable, and time in days since the survey began as the predictor variable. Assuming observers see a constant proportion of the dolphins in the water, evidence of a positive slope for this regression model is evidence of an increase in the number of dolphins in the sanctuary. As measures of dolphin abundance, I used the presence or absence of dolphins in a transect (logistic regression) and the number of dolphins counted in a transect (Poisson regression). I analysed both binary and count data because I was interested to see if it would make any difference. In the first five years, 74% of counts were either zero or one, and so I wanted to see how much power I'd lose if I converted every count to a binary presence or absence.

In early analyses, with only three and then four years' data, I assumed observations on the same transect were independent. I was cautiously optimistic: these basic regression models suggested an increase in number, but I was unsure what effect the likely correlation between observations would have. I was concerned by high variability in the data. For example, the highest number of dolphins counted in one flight was 57; yet only a single dolphin was seen in a subsequent flight ten days later. I thought that with such variable data, the evidence for an increase in numbers might depend on my assumption of independent observations. To get more definitive results, I developed logistic and Poisson regression models using generalised estimating equations.

4. GEE models

Liang and Zeger's (1986) generalised estimating equations are a way of analysing correlated longitudinal data within a generalised linear model framework. To avoid specifying a joint distribution for observations on the same subject, they proposed using a 'working correlation matrix' - a model for all the pairwise correlations between observations on the same subject. Their method has some nice properties. Provided that the relationship between the response

and predictor variables is modelled correctly, GEE estimates are consistent. And the closer the assumed correlation model is to the true correlation, the more efficient the estimate.

Liang and Zeger (1986) give five correlation models. Of these, the most appropriate for the Hector's dolphin data is to assume the correlation between observations y_{it} and $y_{it'}$ on the i^{th} transect is $\text{corr}(y_{it}, y_{it'}) = \alpha^{|t-t'|}$, where t is time in days. This 'first-order autoregressive' model says that the correlation between observations on the same transect decreases as the time between observations increases. The model has a single parameter α to be estimated from the data.

As far as I know, Genstat is the only statistical package with a GEE procedure (see Kenward and Smith 1995). A GEE macro written in SAS/IML by Rezaul Karim has been around for some time. SAS (according to their website) will add GEE capabilities to PROC GENMOD in a maintenance release to version 6.11. Xiangyang Liu has recently written 'Quator' - shareware for Windows. David Smith and Peter Diggle have written 'Oswald' - an add-on package for S-plus. Statlib has code for GEE in S-plus and XLISP.

I wrote my own programs in Genstat - Genstat's GEE procedure wasn't available then. Liu's software does not have an autoregressive option, and the autoregressive option in the SAS macro does not handle missing data. With an autoregressive model, Liang and Zeger (1986) suggest that since $E(\hat{r}_{it} \cdot \hat{r}_{it'}) \equiv \alpha^{|t-t'|}$, the slope of the regression of $\log(\hat{r}_{it} \cdot \hat{r}_{it'})$ on $\log(|t-t'|)$ is an estimate of α , where \hat{r}_{it} is the Pearson residual. How they arrive at this conclusion is not clear to me, and what one does when two Pearson residuals are opposite in sign is not clear either. I estimated α using Genstat's FITNONLINEAR directive.

5. *Flexi models*

Flexi was initially developed by Martin Upsdell (Wheeler and Upsdell 1994 p8). Flexi is Bayesian software: the user selects a covariance function and the degree of polynomial expected for the mean, given what is known from theory about the data. Other Bayesian smoothers prescribe specific mean and covariance functions as part of the method (Upsdell 1996). The user can specify non-normal error distributions, and can restrict the range of expected values with an appropriate link function. Flexi uses this information 'in a similar way to generalised linear models (McCullagh and Nelder 1989), by iteratively forming an adjusted dependent variable with associated weights' (Wheeler and Upsdell 1994 p182).

As prior information for my analyses, I used TYPE:=AUTOREGRESS, MORDER:=2 and ORDER:=0. These parameters represent my expectation that firstly, repeated observations on the same transect will be correlated; and secondly, the mean response will be higher at one end of the series (ie. the mean function polynomial should have two terms). If MORDER is greater than ORDER, the model will have a deterministic component. With MORDER:=2 and ORDER:=0, Flexi will estimate constant and slope parameters and their standard errors.

Flexi has two variances and is essentially fitting a generalised linear mixed model (a 'random effects' model) using REML equations. Of the two independent variances, one is the variance of the curve about the mean (the 'random effects' variance), and the other is the variance in measuring the response (the 'error' variance). It's a Bayesian version of Genstat's GLMM procedure (Welham 1993), except that GLMM uses a diagonal covariance matrix for its random effects, while Flexi uses a 'structured' covariance which typically contains off-diagonal elements.

Parameter estimates from a random effects model are not the same as 'population-averaged' estimates from GEE or basic regression models ('marginal models'). With this survey, each of the 15 transects could perhaps have its own intercept and slope, and a random effects model would describe how each transect's intercept and slope varies about average values. These average values are given as estimates of the deterministic component of the model, and they are not the same thing as estimates of a response curve for the population (see Diggle, Liang and Zeger 1995 p137-142). Looking at different responses in different transect is of some interest, but what the Department of Conservation really wants to know is what these 15 transects have to say about the population's response over time.

Random effects and marginal models are compared by Zeger, Liang and Albert (1988) and Neuhaus, Kalbfleisch and Hauck (1991). Some simple relationships exist if subjects (here transects) have different random intercepts but a common slope, as long as the distribution of random intercepts is Gaussian. With a logit link, the absolute value of a random effects parameter will always be greater than the absolute value of the equivalent marginal parameter, but its standard error will be proportionately greater too, so that a random effects model gives approximately the same inference about whether a parameter is zero (Zeger *et al* 1988). With a log link, the two models will have different intercepts, but all other parameters and their standard errors will be the same (Zeger *et al* 1988).

But assuming random intercepts and a common slope implies an equal correlation between any two measurements on the same subject (Diggle *et al* p56). This model for correlation does not account for any serial correlation between measurements on the same subject. And perhaps transects have different random intercepts and different random slopes, and then with a logit link, 'simple general statements regarding the relationship between [random effects and marginal parameters] do not seem to be available' (Neuhaus *et al* 1991). Gromping (1996) shows that with a log link, random effects models can be made to give correct 'population-averaged' parameters by including additional predictor variables in the deterministic component of a random effects model. In all cases, the distribution of random effects must be correctly specified; otherwise estimates from a random effects model will not be consistent (Zeger *et al* 1988).

In theory then, Flexi and GEE models with the same deterministic component will give different estimates for slope unless there is an equal correlation between any two measurements on the same subject. When some sort of serial correlation is expected, Flexi estimates for slope will only approximate GEE estimates. As the next section shows, the approximation turns out to be quite good for the survey data but there seems no way to generalise this result to other situations.

6. Detecting a trend

Approximate 95% confidence intervals for the slope parameter are shown in Figure 3, for logistic and Poisson regression models. Each interval is an estimate of slope plus or minus two standard errors. Intervals are given for a basic model, where one assumes observations on the same transect are all independent, and for Flexi and GEE models. Each model has an interval for three (1990-92), four (1990-93) and five (1990-94) years' data.

Given four and five years' data, there is good evidence of a positive slope - that is, an increase in numbers. Confidence intervals for the basic model are too narrow and with three years' data, results from the basic model could be misleading. Flexi does a good job of reproducing GEE results, but its intervals tend to be too wide under Poisson regression.

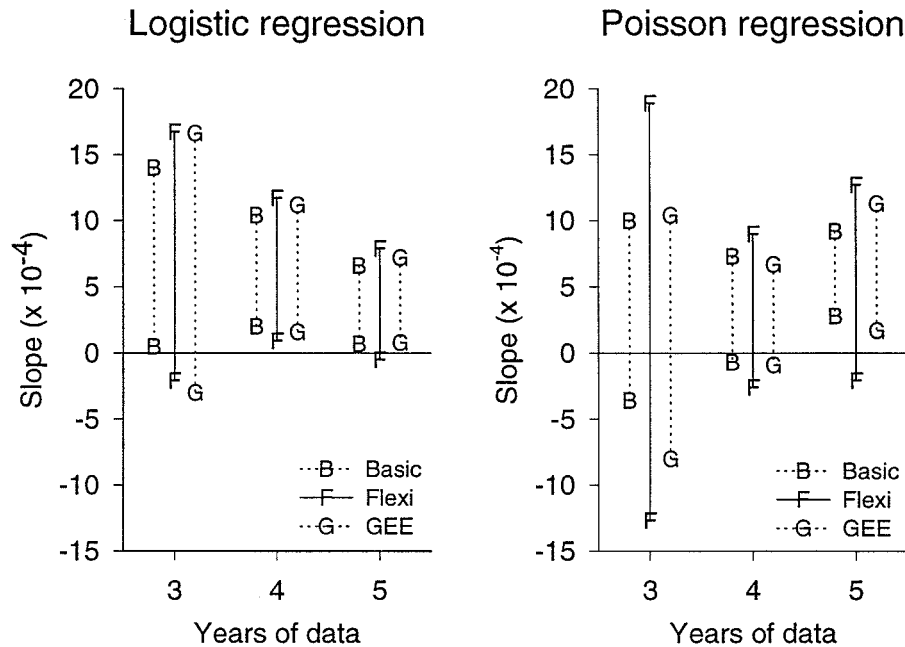


Figure 3. Approximate 95% confidence intervals for slope using logistic and Poisson regression with basic, Flexi and GEE models

Estimates of α from GEE models can be of some practical use. With logistic regression, values of α are 0.48, 0.34 and 0.45 for three, four and five years' data respectively. Taking 0.50 as an upper limit for α implies that presence or absence observations made on the same transect but a week apart are essentially independent (with a correlation of less than 0.01). With Poisson regression, values of α are higher (0.73, 0.79, 0.94) implying two weeks to a month must pass before counts on the same transect can be considered independent. The more independent consecutive observations are, the more information is gained from each consecutive flight. From 1995 on, the Department of Conservation plans to fly only five times each summer. If these flights are at least a week apart, the data collected will carry more information with more precise estimates of slope as a result. Consecutive binary observations will be essentially independent, so that basic logistic regression models can be used to detect trend. Figure 3 suggests that with this data set, binary data are as informative as the count data from which the binary data were derived.

Note that these estimates of α may not be very accurate (see discussion in Kenward and Smith 1995). They recommend a second method of fitting GEE models ('GEE2') when α is of interest. This second method requires both the correct regression model and the correct correlation model before parameter estimates are consistent. Liang and Zeger's GEE ('GEE1') gives consistent estimates of regression parameters even if the correlation model is wrong. Fitzmaurice, Laird and Rotnitzky (1993 - see also the discussion following their paper) recommend GEE1 if regression parameters are of interest and α considered a nuisance parameter, as in this example.

With Poisson regression, scale parameter estimates indicate overdispersion. The variability in counts is five to seven times what would be expected if counts were distributed Poisson. As Hector's dolphins are usually seen in small groups, a compound Poisson model seems

appropriate here. Counts of animal groups are distributed Poisson; group size is an independent and identically distributed variable. The resulting distribution of counts is compound Poisson. This model is consistent with overdispersed count data (McCullagh and Nelder 1989 p198) and is often appropriate in ecology (Feller 1968 p289). I am not aware of any use of the compound Poisson model in a regression context.

7. Using covariates

Sea and cloud conditions recorded during each flight are potential covariates. In similar surveys, both covariates have had significant effects. Calm seas and clear skies are expected to lead to higher counts (Barlow, Oliver, Jackson and Taylor 1988). Figure 4 shows (as open circles) the average sea and cloud conditions in the first five years of the survey. Sea condition (the left graph) is recorded as calm (zero) or rough (one); while cloud cover (the right graph) is recorded in eighths, from a clear sky (zero) to complete cloud cover (eight).

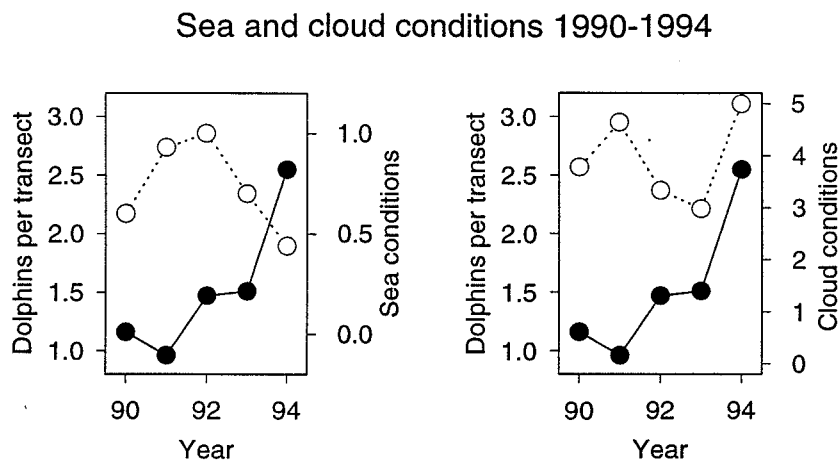


Figure 4. Average sea and cloud conditions (open circles) plotted against the average number of dolphins seen per transect (closed circles) for the first five years of the survey

Of the two, sea condition is the more promising covariate. With a basic model and five years' count data, the estimated coefficient is -0.36 (standard error 0.17). With a GEE model and the same data, the estimate is -0.31 (standard error 0.25). So counts tend to decrease when made under rough conditions, but the evidence for this relationship is not conclusive. However, a sizeable correlation (-0.34) between time and sea coefficients is evidence of multi-collinearity: one cannot untangle the separate effects of time and sea condition on the number of dolphins counted. The left graph of Figure 4 shows that over time the average number of dolphins counted increased as average sea conditions improved.

Covariates are easily added to basic and GEE models - these are now multiple regression models instead of simple regression models. With Flexi, one can use only a single predictor variable. To include information on sea conditions, I adjusted measures of dolphin

abundance to what would be expected if observations were always made in calm conditions. Once adjusted, counts recorded in rough conditions were the original integer values plus a constant; presence or absence data for rough conditions were the original one or zero plus a constant. These constants were calculated from the parameters of Poisson (or logistic) regression models where the response was dolphins counted (or presence or absence of dolphins), and the predictor was sea condition. In effect "It is as though each Y were moved parallel to the sample regression line until above X, and then measured as a new or adjusted Y" (Steel and Torrie 1980 p251).

Flexi has some useful graphs for exploring whether this approach is valid. Figures 5 and 6 show Poisson models in Flexi for data collected in calm (closed circles, solid lines) and rough sea conditions (open circles, dotted lines). Thick lines denote the mean response, while the thinner lines give an 83% confidence interval for the mean. If sea condition is to be a useful covariate, Poisson models under different conditions should have different intercepts but the same slope. In Figure 5, the mean response under calm conditions appears greater than the mean response under rough conditions, although the evidence is not conclusive. Figure 6 shows the slope (and an 83% confidence interval on that slope) for calm and rough sea conditions. There doesn't seem to be any difference in slope. Note both graphs have wider confidence intervals for calm conditions because only 200 out of the 750 observations were made in truly calm conditions.

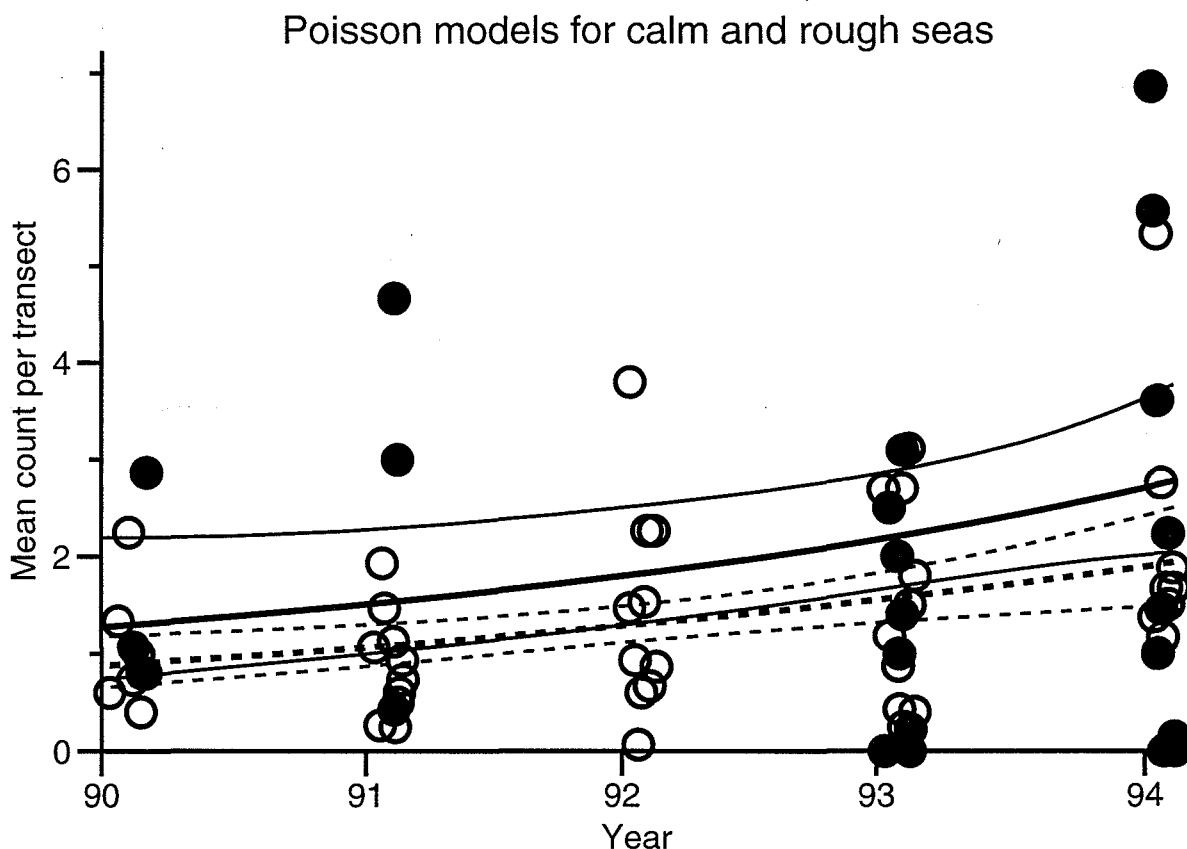


Figure 5. Poisson models for calm (closed circles, solid lines) and rough (open circles, dotted lines) sea conditions.

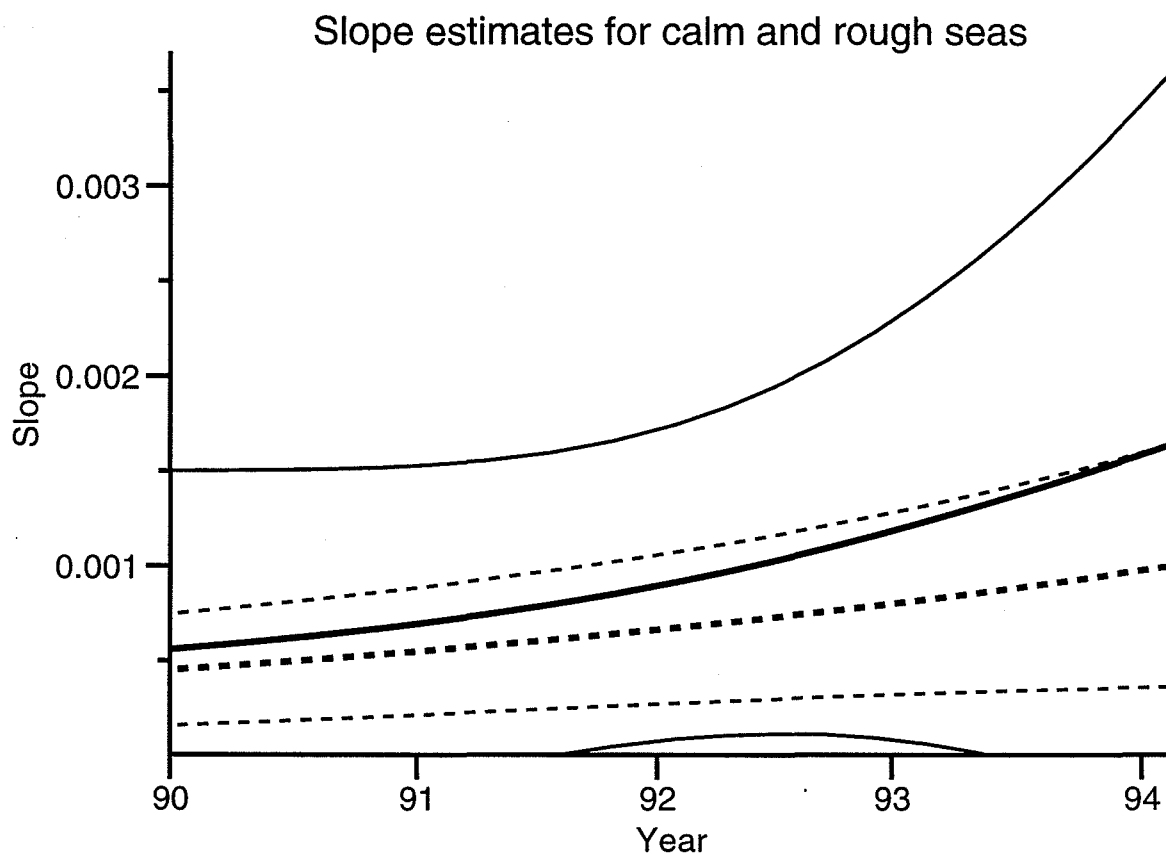


Figure 6. Slope estimates from Poisson models for calm (solid lines) and rough (dotted lines) sea conditions.

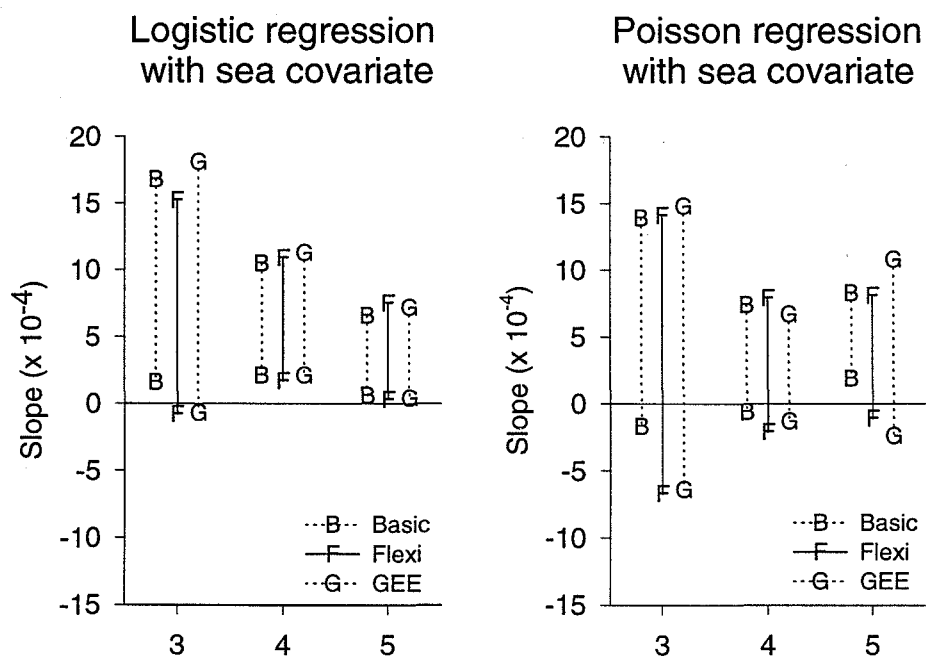


Figure 7. Approximate 95% confidence intervals for slope using logistic and Poisson regression models with sea condition as a covariate.

Figure 7 shows approximate 95% confidence intervals for the slope using logistic and Poisson regression models with sea condition as a covariate. There's not much difference with and without the covariate (Figure 7 versus Figure 3). Ideally, including a covariate should reduce bias and increase precision in estimates of the slope over time. Confidence intervals for basic models haven't changed much; Flexi intervals are a little narrower (particularly under Poisson regression); GEE intervals are little wider (particularly under Poisson regression). With three years' data, all confidence intervals have shifted slightly upwards. Again Flexi does a good job of reproducing GEE results, but this time two of its intervals are too narrow.

8. Sample size

Flexi will accept raw data or data summarised for each value of the predictor variable. I found Flexi slow when given large amounts of raw data, taking several hours with 750 individual values, and the algorithm often failed to convergence. Flexi still gives parameter estimates but their standard errors may be unreliable (Wheeler and Upsdell 1994 p146).

The alternative with count data is to summarise data as means, and to weight each mean (using $ERROR_STD=1/\sqrt{n}$, where n is a variable giving the number of observations in each mean). This way I had no convergence problems. On the other hand, some information is lost in representing counts simply by their means and relative standard errors. Estimates from raw data may have smaller standard errors where Flexi can converge.

With binary data, the user can give the number of 'successes' as the response variable, and the number of binomial trials as a special NBINOMIAL variable. I had no convergence problems with this method, nor is there any loss of information compared with a data set consisting of zeros and ones. The number of 'successes' can take non-integer values if the response needs to be adjusted for a covariate.

Flexi has been designed particularly for use with small samples. With small samples, prior information will have more influence and results may well differ from the results of a non-Bayesian method. I have used large samples here. My interest is in using Flexi, a random effects model, to give quick approximate estimates of 'population-averaged' parameters. As a rule of thumb for this survey's data, GEE confidence intervals will be wider than those of a basic model, but narrower than those of a Flexi model.

9. Long term trend

Using Flexi to approximate GEE results has meant fitting Flexi models that are less than ideal. With $MORDER:=2$ and $ORDER:=0$, Flexi estimates two deterministic parameters, but $ORDER:=0$ implies that the data have a stationary covariance (Wheeler and Upsdell 1994 p169). This would be true if there were no serial correlation between measurements on the same subject, and each subject's response over time had the same slope but a different intercept (Diggle *et al* 1995 p88-89). But an equal correlation between any two measurements on the same transect isn't likely with this survey's data.

Figure 8 shows response profiles for the counts in each transect over time. Transects obviously differ in slope. In Figure 8, the appearance of each line indicates which transect the line represents. The long dashes far apart represent transect 1; the short dashes close together represent transect 15; and the other transects have intermediate patterns with more dashes as the number of the transect increases (see Figure 1). Figure 8 clearly shows that transects at the edge of the sanctuary show little (or perhaps even negative) growth from 1990 to 1994.

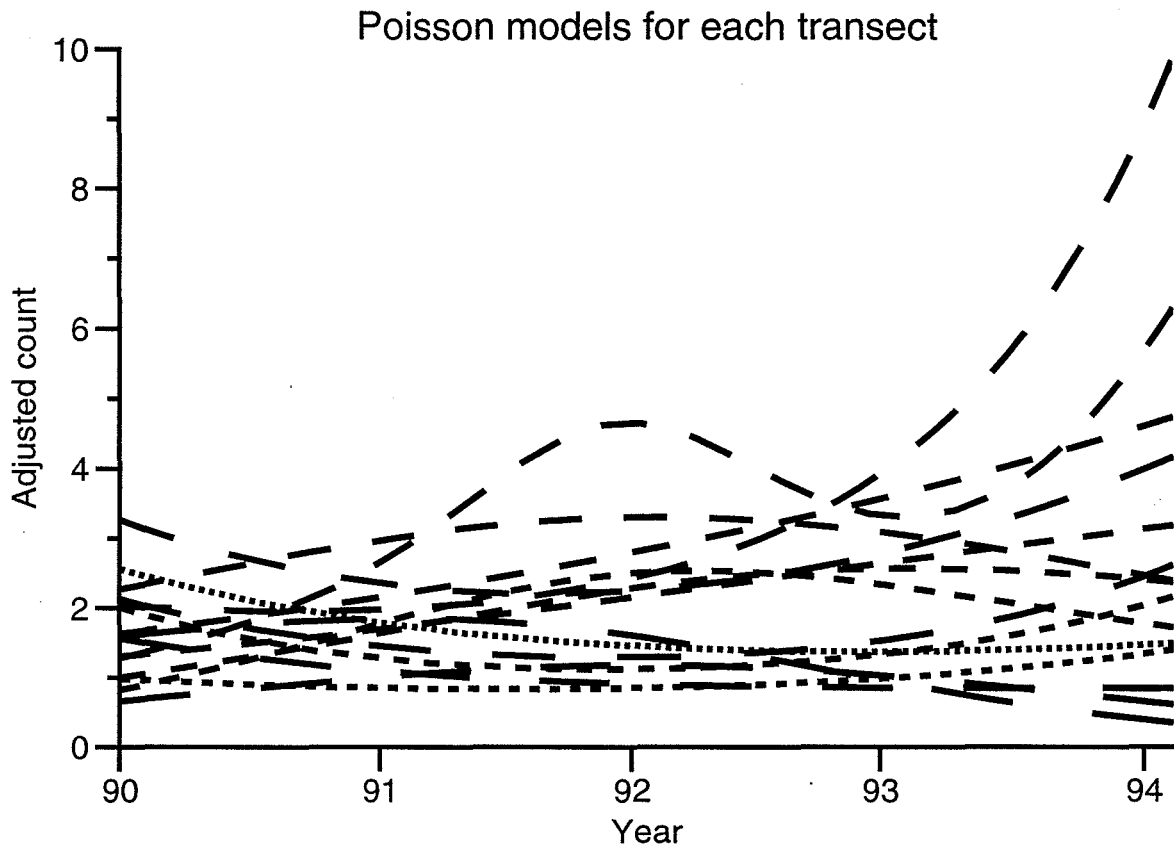


Figure 8. Poisson models for each of the 15 transects, using count data adjusted for the sea covariate.

The fanning out of response profiles over time is common with growth data. Figure 8 looks very similar to a simulation in Diggle *et al* (1995 p90) where subjects have both random intercepts and random slopes over time. In this case, the covariance is non-stationary, with a quadratic increase over time. So a more sensible model here for the covariance is a stationary (autoregressive) covariance, after a second differencing of the covariance function. In Flexi, this is ORDER:=2. Note that using inappropriate values of ORDER and TYPE may cause convergence problems (Wheeler and Upsdell 1994 p146).

But if MORDER:=2 and ORDER:=2, there is no deterministic component to the model. In a sense, this is as it should be:

‘It is clear that for many applications, the assumption of a stochastic trend is often more realistic than the assumption of a deterministic trend. This is of special importance in forecasting a time series, since a stochastic trend does not necessitate the series to follow the identical pattern that it has developed in the past.’ (Box, Jenkins and Reinsel 1994 p97).

Setting ORDER:=1 allows an estimate of a possible deterministic linear trend in the data, while still providing a better description of the likely covariance function. Figure 9 compares response curves (and 83% confidence intervals for the mean response) for ORDER:=0 and ORDER:=1, using count data adjusted for the sea covariate. The mean response curve for ORDER:=1 (the dotted lines) is much more sensible than the mean response curve for ORDER:=0 (the smooth lines).

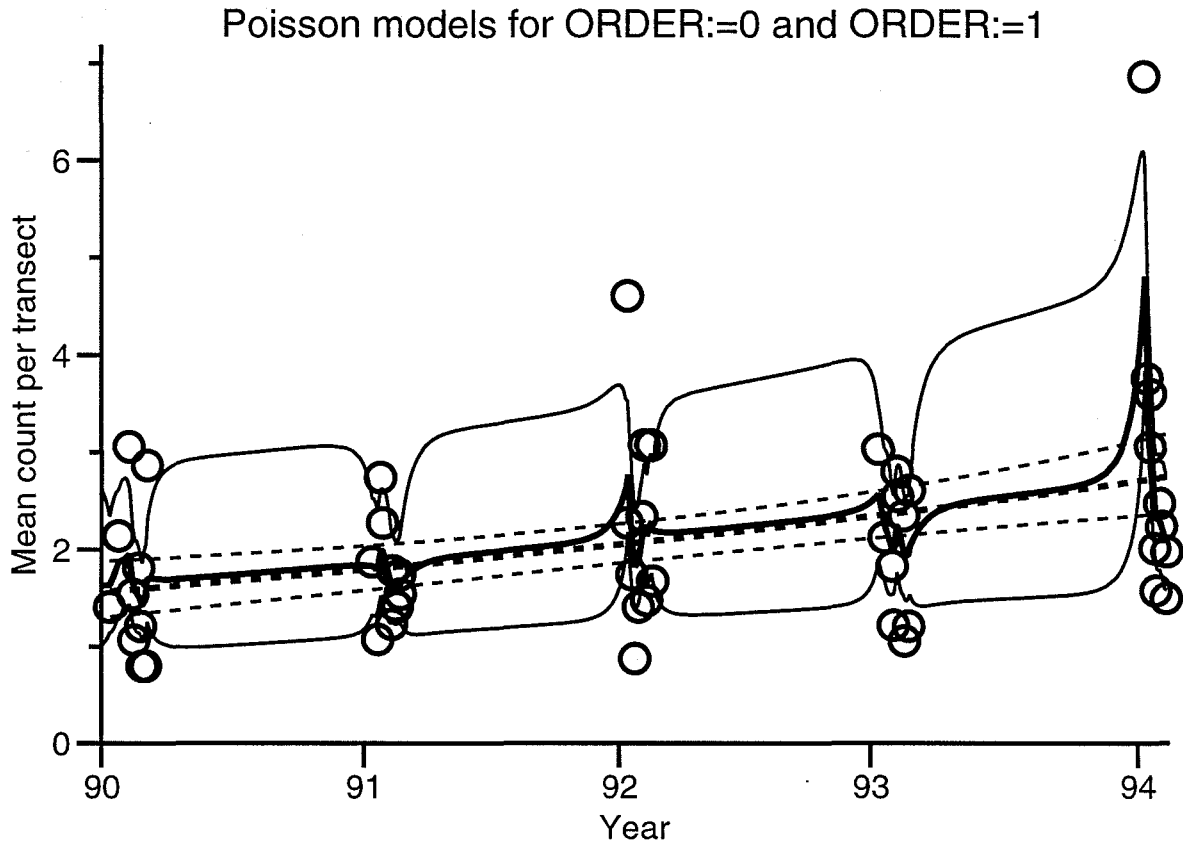


Figure 9. Poisson models for ORDER:=0 (smooth lines) and ORDER:=1 (dotted lines).

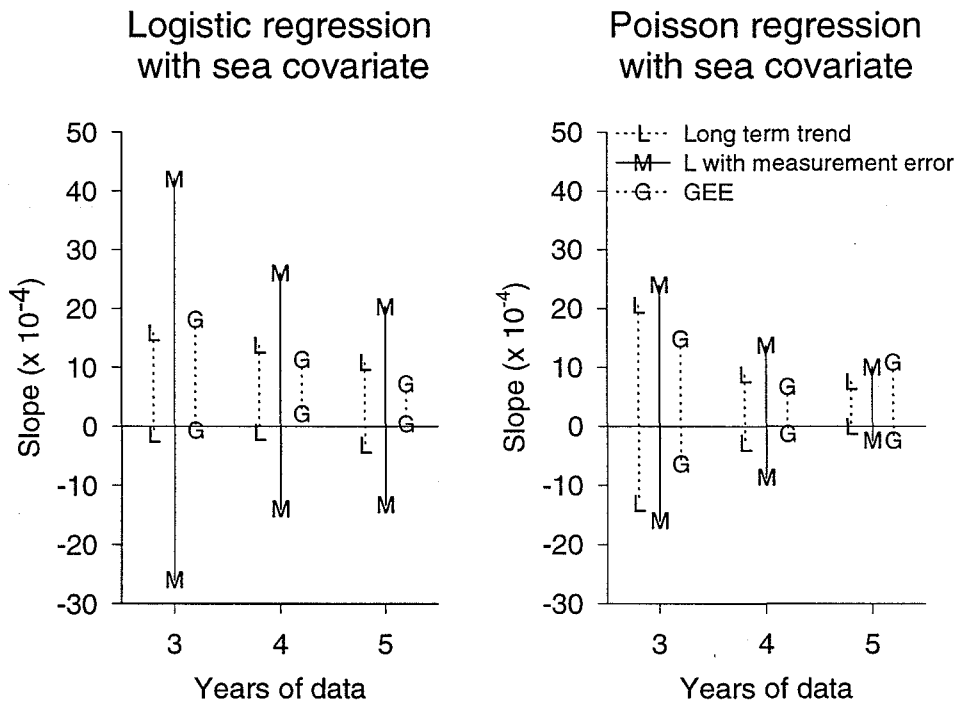


Figure 10. Approximate 95% confidence intervals for a 'long term' linear trend with and without measurement error.

Figure 10 gives approximate 95% confidence intervals for this 'long term' trend. That is, if there is a deterministic trend in this time series, what do the data say about its value? The resulting confidence intervals are similar to 'population-averaged' confidence intervals from the equivalent GEE model.

10. Measurement error

Information about measurement error can be included in Flexi. The ROUNDOFF parameter can be set to the standard deviation of the data points. From a set of 30 duplicate counts from two independent observers, I estimated the variance in counts (or binary data) as half the variance between the duplicates (Diggle *et al* 1995 p87). I converted the variance to a standard error of the mean, as the data were summarised in Flexi as a mean count (or proportion) at each point in time. This estimate of measurement error is not going to be very accurate, based on so few duplicates and on the assumption that both sets of observations are equally variable. Nevertheless, Figure 10 shows how important it is to assess measurement error. Confidence intervals for slope are often much wider when information about measurement error is included in the model.

11. Conclusion

Diggle *et al* (1995 p79) list three sources of random variation in longitudinal data: random effects, serial correlation and measurement error. They speculate (p88) that perhaps:

'Whilst serial correlation would appear to be a natural feature of any longitudinal data model, in specific applications its effects may be dominated by the combination of random effects and measurement error.'

The GEE approach taken here models variation as serial correlation, rather than as random effects or measurement error. Diggle *et al* (1995 p86) note that with serial correlation models of this sort, 'there is no straightforward extension to accommodate measurement error or random effects.'

In a way, Flexi addresses all three components of random variation in longitudinal data. Its variance for random effects can incorporate a serial component; the 'error' variance has measurement error as its lower bound. A choice of seven covariance functions, which can then be integrated to non-stationary forms, means that Flexi is indeed flexible at modelling the covariance structure of the data. These are the strengths of Flexi, along with its power with small data sets because of its Bayesian approach. Flexi's estimates are not 'population-averaged, but I would use Flexi to estimate the 'long term' trend, because Flexi models variation principally as random effects and measurement error and these may well be the dominant sources of variation. Comparing Flexi and GEE estimates will show whether GEE's 'population-averaged' estimates are to be trusted. I also found Flexi's quick graphical displays very informative: for example, the response profiles for each transect, and for different values of a covariate.

12. References

Barlow, J; Oliver, C W; Jackson, T D; Taylor, B L. (1988) Harbor-porpoise, *Phocoena Phocoena*, abundance estimation for California, Oregon, and Washington: II. Aerial surveys. *Fishery Bulletin* 86(3), 433-444.

- Box, G E P; Jenkins, G M; Reinsel, G C. (1994) *Time series analysis: forecasting and control* (third edition). Prentice Hall, London, 598p.
- Department of Conservation (1992) *Banks Peninsula Marine Mammal Sanctuary technical report*. Department of Conservation, Christchurch, 110p.
- Diggle, P J; Liang, K-Y; Zeger, S L. (1995) *Analysis of longitudinal data*. Clarendon Press, Oxford, 253p.
- Feller, W. (1968) *An introduction to probability theory and its applications* (volume one, third edition). John Wiley, New York, 509p.
- Fitzmaurice, G M; Laird, N M; Rotnitzky, A G. (1993) Regression models for discrete longitudinal responses (with discussion). *Statistical Science* 8(3), 284-309.
- Gromping, U. (1996) A note on fitting a marginal model to mixed effects log-linear regression data via GEE. *Biometrics* 52, 280-285.
- Kenward, M G; Smith D M. (1995) Computing the generalized estimating equations with quadratic covariance estimation for repeated measurements. *Genstat Newsletter* 32, 50-62.
- Liang; K-Y; Zeger, S L. (1986) Longitudinal data analysis using generalized linear models. *Biometrika* 73(1), 13-22.
- McCullagh, P; Nelder J A. (1989) *Generalized linear models* (second edition). Chapman and Hall, London, 511p.
- Neuhaus, J M; Kalbfleisch, J D; Hauck, W W. (1991) A comparison of cluster-specific and population averaged approaches for analyzing correlated binary data. *International Statistics Review* 59(1), 25-35.
- Steel, R G D; Torrie, J H. (1980) *Principles and procedures of statistics: a biometrical approach* (second edition). McGraw-Hill, London, 633p.
- Slooten, E; Lad, F. (1991) Population biology and conservation of Hector's dolphin. *Canadian Journal of Zoology* 69, 1701-1707.
- Upsdell, M P. (1996) Choosing an appropriate covariance function in Bayesian smoothing. In Bernardo, J M; Berger, J O; Dawid, A P; Smith, A F M (editors). *Bayesian Statistics 5*. Clarendon Press, Oxford, 747-756.
- Welham, S J (1993) Procedure GLMM. In Payne, R W; Arnold, G M; Morgan, G W (editors). *Genstat 5 Procedure Library Manual, release 3[1]*. Lawes Agricultural Trust, Rothamsted, 187-192.
- Wheeler, D M; Upsdell, M P. (1994) *Flexi 2.2: Bayesian Smoother, reference manual* (second edition). New Zealand Pastoral Agricultural Research Institute, Hamilton, 330p.
- Zeger, S L; Liang, K-Y; Albert, P S. (1988) Models for longitudinal data: a generalized estimating equation approach. *Biometrics* 44, 1049-1060.

13. Acknowledgments

I am grateful for financial support from the Department of Conservation. Credit for many long hours dolphin-spotting goes to Martin Rutledge, Chris Woolmore and Andy Grant. Richard Sedcole, here at Lincoln, helped with the Genstat programming. Clare Salmond, of the Wellington School of Medicine, suggested using Flexi for this sort of data and I very much appreciate Martin Upsdell's prompt advice. Copies of Flexi are available from Martin Upsdell, Ruakura Agricultural Research Centre, Private Bag 3123, Hamilton, New Zealand.