10095

One-aircraft tandem survey - analysis of pod sighting data from line transect aerial surveys using one aircraft to repeat selected track segments.

Summary

The ML point estimate for g(0) from porpoise sightings unstratified by sighting conditions was 0·68 and the effective search width 356 metres. The 95% confidence interval on the g(0) estimate was from 0·21 from 1·0. For sightings stratified by subjective sighting conditions the g(0) estimate was 0·86 under good and 0·24 under moderate conditions. The improvement in fit provided by the subjective assessment of sighting conditions was highly significant.

The point estimates for g(0) and for the mean rate of pod displacement are about twice those from SCANS data. Constraining pod displacement rate to that estimated from SCANS gives g(0) estimates identical to those from SCANS (0·25 for unstratified sightings). The observed differences are not statistically significant and are probably due to the limited number of trailing track segments completed sufficiently close to the leading track. Additional circling experiments using a slightly modified flight pattern should provide estimates of g(0) and effective search width with confidence intervals that are narrow enough to be useful. At the same time it would be worth considering whether local conditions could have led to increased rates of pod displacement and also whether the records of pod movement direction could help in discriminating duplicate sightings.

Final report to NMFS/NEFSC under order number 4UENNF700128 by Conservation Research Ltd.,110 Hinton Way, Great Shelford, Cambridge, CB2 5AL, England.

Introduction

The previous report under this contract described a number of simulation tests of the one-aircraft version of the tandem line-transect survey technique. Initial field trials of the data collection protocol were carried out in October 1997 and flights were then conducted through June 1998 to collect data suitable for estimation of effective strip width (*esw*) to porpoise pods. Data were collected using program VOR to log aircraft location and help in carrying out repeat flights over selected sightings. The location log and end-points of the repeat flight segments were transferred to a database following the flight. Changes in sighting conditions and details of sightings were then typed into the database and integrated with the location log to identify periods of single and duplicate sighting effort and assign the sightings to those periods. Effort periods were split by sighting conditions that were recorded separately for observers on the left and right of the aircraft.

Review of data collected.

Flights were conducted along zigzag tracks in sea areas "C" and "D" off the east coast of the United States between latitude 41N and 44N, and 35N and 41.5N respectively. The tracks were chosen for comparison with earlier surveys and were not designed as part of this project, which was therefore restricted to estimation of *esw* and did not include estimation of porpoise pod abundance.

Flights on tracks in area C were conducted on 31st October, and 18th and 19th November 1997. Flights on tracks in area D were conducted on 2nd and 3rd February, 5th March, 8th April, 28th May and 10th June 1998. 92 porpoise pods were sighted on the area C flights, 47 of which initiated a "racetrack" flight pattern to attempt a second pass along the track segment from which the pod had been sighted. The remaining 45 were seen during "dead-time" periods imposed by program VOR to avoid excessive circling by the aircraft. 20 porpoise pods were seen on the 47 second-pass track segments. Only 3 porpoise pods were sighted on the area D tracks, each of which initiated a second-pass flight. No pods were seen on those second-pass flights. 25¼ hours of effort were spent in area C, 4/5 hours of which were spent on the repeated track segments (designated as "leading" and "trailing" by analogy with the two-aircraft tandem technique used during SCANS surveys). 9¾ hours of effort were spent in area D, 5¼ minutes of which were spent on the repeated track segments

Program TANDEM was used to estimate *esw* from all the sighting data combined. TANDEM does not require a decision as to whether any sighting on a trailing segment is or is not the same as the sighting on the leading segment. The likelihood for the sightings data is calculated by considering all possible pairings of leading and trailing sightings (including the case where none of the leading and trailing sightings are the same). The likelihood is maximised with respect to parameters of the pod sighting and movement models to derive a ML estimate for *esw*. Data on perpendicular distance to sightings allow parameters defining reduction in detection probability with distance to be estimated. But it is the proximity of trailing to leading sighting positions that allows the proportion of pods detected near the track to be estimated, so the reliability of that estimate depends on the number of potential duplicate sightings available. Unfortunately, the number of potential duplicates among the 47 leading sightings in area C was reduced by lateral displacement of the trailing track segment in a number of cases. The leading and trailing tracks were more than 200m apart for about half (24 out of 47) the leading sightings. Re-sighting was then only possible if the track displacement was in the same direction as the original sighting, as sighting rate declined rapidly beyond 200m (figure 1). In 11 of the 47 sightings the tracks were more than 500m apart, making it impossible for the leading sighting to be re-sighted on the trailing segment.

Figure 3 compares the frequency distribution of mean displacement of trailing from leading tracks with that for SCANS survey flights. Figures 4a to 4e illustrate some of the flight tracks completed during the duplicate trials. The position from which the leading sighting was made is at the origin, and the track plot starts about a minute before the sighting and continues to the edge of the figure. Thus in figure 4a the aircraft enters from the left and the trailing flight track lies close to leading one throughout. In figure 4b the trailing track converges with the leading one only just before the origin and in the rest of the figures the convergence is too late, so that the leading sighting is probably out of range of the trailing segment. It is probable that the pilot, having completed the second turn, noted the displacement from the leading track (inevitable, given the existence of cross winds) and then converged gradually in order to maintain a smooth flight. It is therefore necessary to re-emphasise the need to converge with the leading track as soon as possible after the second turn, even at the expense of some uncomfortable corrections to the flight path.

For eleven of the leading sightings a trailing sighting occurred within 15s of the time the aircraft, on its second pass, came abeam of the leading sighting position (see figure 5). Any pod seen more than 15s before or after that position is unlikely to be the same as the leading pod. Aircraft speed is about 50ms^{-1} and if a pod were to move more than 750m during the time between the two passes (about 4½ minutes) its mean speed would need to exceed 10kph. It is impossible to determine which of the sighting pairs in the figure were genuine duplicates. For five of the pairs lateral displacement was over 500m making it unlikely that they were duplicates unless we accept a rate of displacement considerably higher than that estimated from the SCANS data. Unfortunately the small effective number of duplicate trials did not allow the TANDEM program to distinguish clearly between a low rate of displacement and low number of duplicates, and a high rate of displacement and high number of duplicates. The resulting point estimates for g(0) and displacement, given in the following section, are probably too high and the CV's very wide.

Before running program TANDEM the data were inspected to identify any variation in mean density of porpoise pods which might bias the estimate of *esw*. In the version of the analysis used for SCANS data the possibility that a proportion of duplicate track segments had been flown in areas of zero density was incorporated in the likelihood function. Although the *esw* estimator is reasonably robust to variations in density, the estimator is biased upwards when density is zero over a significant proportion of the survey area. This is because the mean density over the whole area is then significantly less than the mean density in those areas where duplicate trials are completed, so that the risk of leading and trailing sightings in close proximity being of different pods is underestimated. However, the modification of the likelihood used for SCANS data is not appropriate to the one-aircraft version, and this aspect was neglected when simulation trails of the one-aircraft version were conducted. Prior to analysis of the current data set a second program, STITCH was written to overcome this difficulty.

STITCH considers the spacing of sightings along the succession of leading and single track segments that make up the entire survey track, and generates a frequency distribution of numbers of sightings per sector when the track is divided in to a number of equal-length sectors. The sector length is chosen such that the expected proportion of empty sectors is 30% assuming random distribution of pods over the entire survey area. The observed number of empty segment will normally exceed the expected 30% and can be used to estimate what proportion of the survey area has zero pod density. This assumes that the non-zero regions of the survey area have constant pod density and STITCH plots the expected frequency distribution under this assumption for different values of the zero-density proportion. Visual comparison of the observed and expected frequency distributions suggests a value of around 70% for the zero-density proportion over the current survey area. However, rather than apply that estimate STITCH outputs the actual number of empty sectors and this value is read in when TANDEM is run and used to modify the likelihood. The zero-density proportion is introduced as an extra free parameter for maximising the modified likelihood.

This modification was subjected to simulation testing and gave satisfactory results. These are described in Hiby (1998), a copy of which is included as an Appendix to this report. The simulation trials included a number of runs to check the validity of likelihood ratio confidence limits derived from program TANDEM. The unconstrained maximum likelihood was compared to maximum likelihood obtained with the g(0) parameter constrained to its value in the simulation. In 5% of runs the constrained log likelihood was more than 1.92 lower than the unconstrained log likelihood so that in 5% of runs the 95% confidence limits on the estimate of g(0) would have failed to include the population value.

Results of the TANDEM analysis

First the files required as input to program STITCH (STITCH.EFF & STITCH.SIG) were generated by pressing hotkey F6 on the PORP97 menu. The files were copied to the directory containing the STITCH program and the program was then run, resulting in 59 empty sectors and 22 non-empty sectors. Figure 7a compares the frequency distribution for observed numbers of sightings against that expected for random distribution of pods at constant density. In figure 7b the expected frequency distribution is modified by assuming 70% of the survey area has zero pod density and gives a close match to the observed distribution.

Program TANDEM was first applied to all survey data without stratification by sighting conditions. That is, in the PROTOLEG table the "conditions" attribute was set to 11 for every porpoise sighting where the "subjective" attribute was in the set {gg,mg,gm,mm}. Hotkey F5 was then pressed to generate the PORPOISE.LEG and PORPOISE.SIG files, these were copied to the subdirectory containing the executable TANDEM program and the program run, entering the values of 59 empty and 22 non-empty sectors output from program STITCH. The resulting ML estimates were:

```
INITIAL POP DENSITY              =    0.140D-06
PROPN ZERO                       =    0.724D+00
GAMMA SHAPE PARAMETER            =    0.655D+01
GROWTH IN GAMMA SCALE (M/S)      =    0.450D+00
HAZARD RATE PARAMETER A          =    0.236D+00
HAZARD RATE PARAMETER B          =    0.678D+01
G(0)                             =    0.680D+00

EFFECTIVE STRIP WIDTH = 356M

MAX. LOG LIKELIHOOD = -239.63138
```

To calculate 95% confidence limits on the estimate of g(0) the program was run with g(0) fixed at values below and above the ML estimate of 0.68 to give log likelihood 1·92 less than the unconstrained maximum of −239·63. The lower 95% CL was 0·21 and the upper limit greater than 1·0.

Program TANDEM was then run with the data stratified by subjective sighting conditions. In the PROTOLEG table the conditions attribute was set to 11 where the subjective attribute was gg, to 12 where the subjective attribute was gm, to 21 where the subjective attribute was mg and to 22 where the subjective attribute was mm. The conditions can be set to 3 where the subjective assessment is that sighting conditions are unacceptable, in case they are classified as unacceptable to one side of the aircraft but still acceptable to the other side. For example, conditions code 23 would mean that sighting conditions were moderate to the left of the trackline and unacceptable to the right. That situation was not encountered in the data set, however. The resulting ML estimates were:

```
INITIAL POP DENSITY              =    0.186D-06
PROPN ZERO                       =    0.724D+00
GAMMA SHAPE PARAMETER            =    0.696D+01
GROWTH IN GAMMA SCALE (M/S)      =    0.423D+00
HAZARD RATE PARAMETER A          =    0.236D+03
HAZARD RATE PARAMETER B          =    0.617D+01
G(0)                             =    0.859D+00
G(0) REDUCN.                     =    0.278D+00
HAZA REDUCN.                     =    0.790D+00

EFFECTIVE STRIP WIDTHS = 454M AND 100M

MAX. LOG LIKELIHOOD = -211.52150
```

The two extra parameters are the decrease in g(0) and the decrease in the scale parameter for the sighting function when moving from good to moderate sighting conditions. The large increase in the maximum log likelihood indicates that the effect of the recorded subjective conditions on sighting rate is highly significant ($P \ll 1\%$). Figure 2 shows the frequency distribution for perpendicular distance to pod sightings for leading and trailing trackline segments. In figure 1 the distribution is split by sighting conditions and the fitted sighting functions are shown by the continuous curves.

Figure 8 compares the frequency distribution of inter-sighting intervals on leading and trailing tracks with that expected if pods are randomly distributed (at a spatial scale corresponding to the duration of a leading or trailing flight). It does not indicate any small-scale aggregation of pods that would lead to upward bias in the g(0) estimate.

Discussion

The point estimates of g(0) obtained from the current analyses are higher than would be expected given what is known of porpoise surfacing behaviour. The g(0) estimate from the SCANS data was around 0·25 compared to 0·68 from the current analysis. The mean pod size estimate from the current analysis was higher than for SCANS (1·87 as compared to 1·2 in block I' of SCANS, which contained the bulk of duplicate effort) so some increase in g(0) is expected but not as large as that observed. Furthermore, the point estimates for the parameters defining the Gamma distribution for pod displacement correspond to a much higher mean displacement rate. The mean is given by the product of the scale and shape parameters and is about $3 ms^{-1}$ for the current analysis as compared to $1{\cdot}5 ms^{-1}$ for SCANS. This difference is crucial because if the Gamma parameters are constrained to their SCANS values the resulting ML estimate for g(0) is almost identical to that from SCANS. The difference is also evident in comparison of figure 5 with the equivalent plot for the SCANS data in figure 6. This is partly due to the longer delay between leading and trailing aircraft using the circling technique (about 4½ minutes instead of 3 minutes) but even taking this difference into account the lateral movements in figure 5 exceed those in figure 6.

The observed differences are probably due to the limited sample size and are not statistically significant. However, it is worth considering whether there could be a real difference in displacement rate between the current area and the SCANS area because a high displacement rate, if genuine, would reduce the potential of the technique to provide useful estimates in the future. A displacement rate of $3 ms^{-1}$ corresponds to 10kph. This might not be unreasonable as a swimming speed but it would be surprising if most pods in a survey region continued at this rate in a consistent direction over the 4½ minutes between the leading and trailing aircraft. All porpoise sightings had a swimming direction recorded – does this imply that they were generally moving in a consistent direction? Is there a reason to expect a difference in average behaviour for porpoises in the currently surveyed area and those in the SCANS areas (primarily blocks I' and Y)? Each displacement is a combination of animals swimming, water currents and positioning errors. There is no reason why abeam times and distances would have been less accurately recorded than during the SCANS surveys and the performance of the GPS equipment should have been equivalent. Could current speeds in the current survey area exceed, on average, those in the SCANS aerial survey areas and, if so, would pods be expected to move with the water mass or maintain their position?

If displacement rates are really higher than anticipated it is worth considering whether the delay time between leading and trailing segment in the circling method could be shortened. It would probably be worth reducing the periods prior to circling back and rejoining the trackline used by program VOR to give duplicate segments of about 40s rather than 60s duration, provided there was still time to converge back to the leading trackline before the original sighting position was reached.

A related question is whether swimming direction could be used to aid in identification of duplicates. An additional term could be introduced to the likelihood for the sighting locations of a duplicate pod - the pdf of the difference in swimming direction recorded from the leading and trailing segments. The modified analysis could then be applied to the current data but some guidance as to the way this value is recorded would be required. For example, though most swimming direction estimates are rounded to ten degrees some observers record values to the nearest degree – what is the reason for this difference?

Fig. 1: Hazard rate sighting functions corresponding to ML estimates for scale and shape parameters and, under moderate sighting conditions, a reduction in the scale parameter.

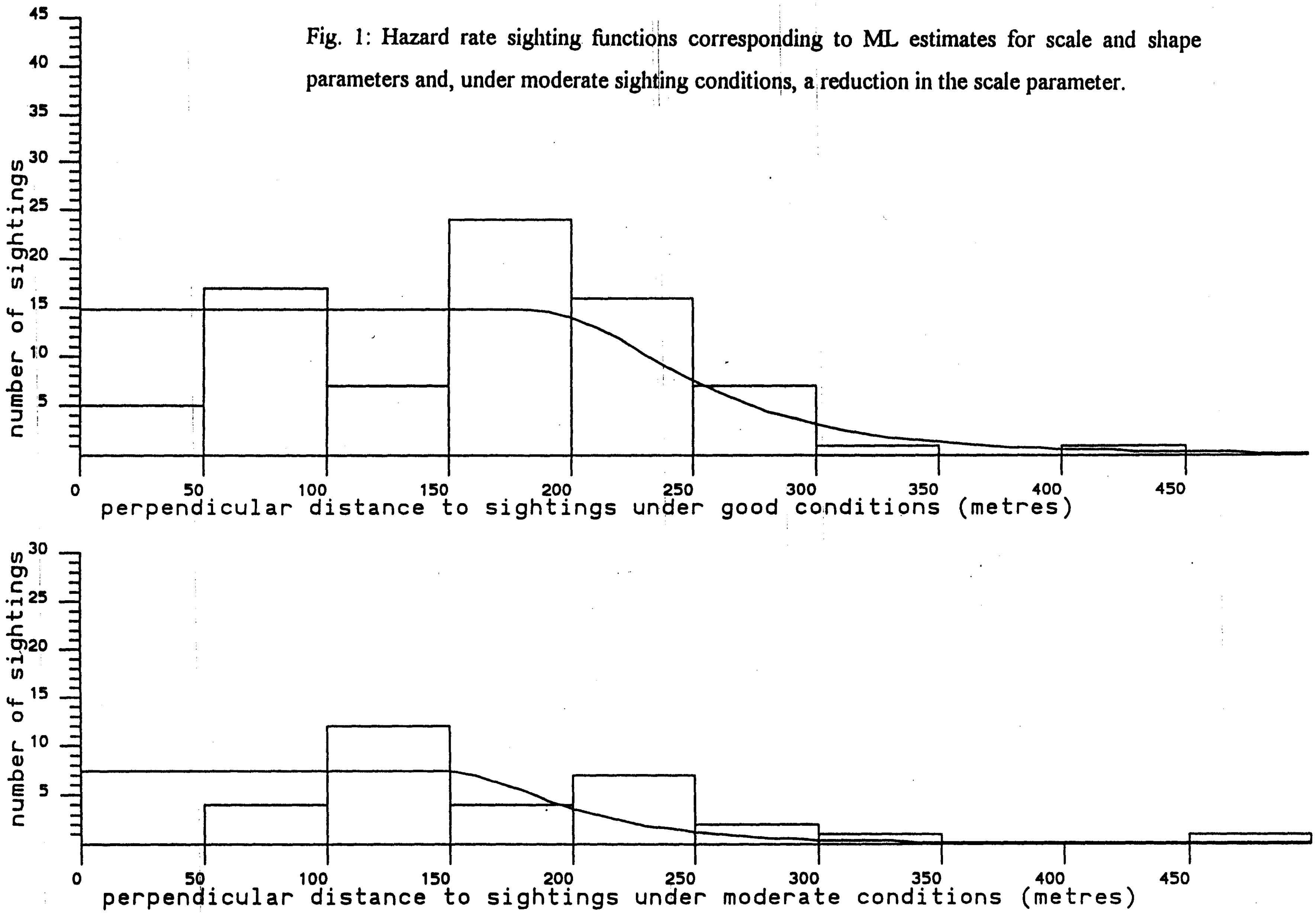
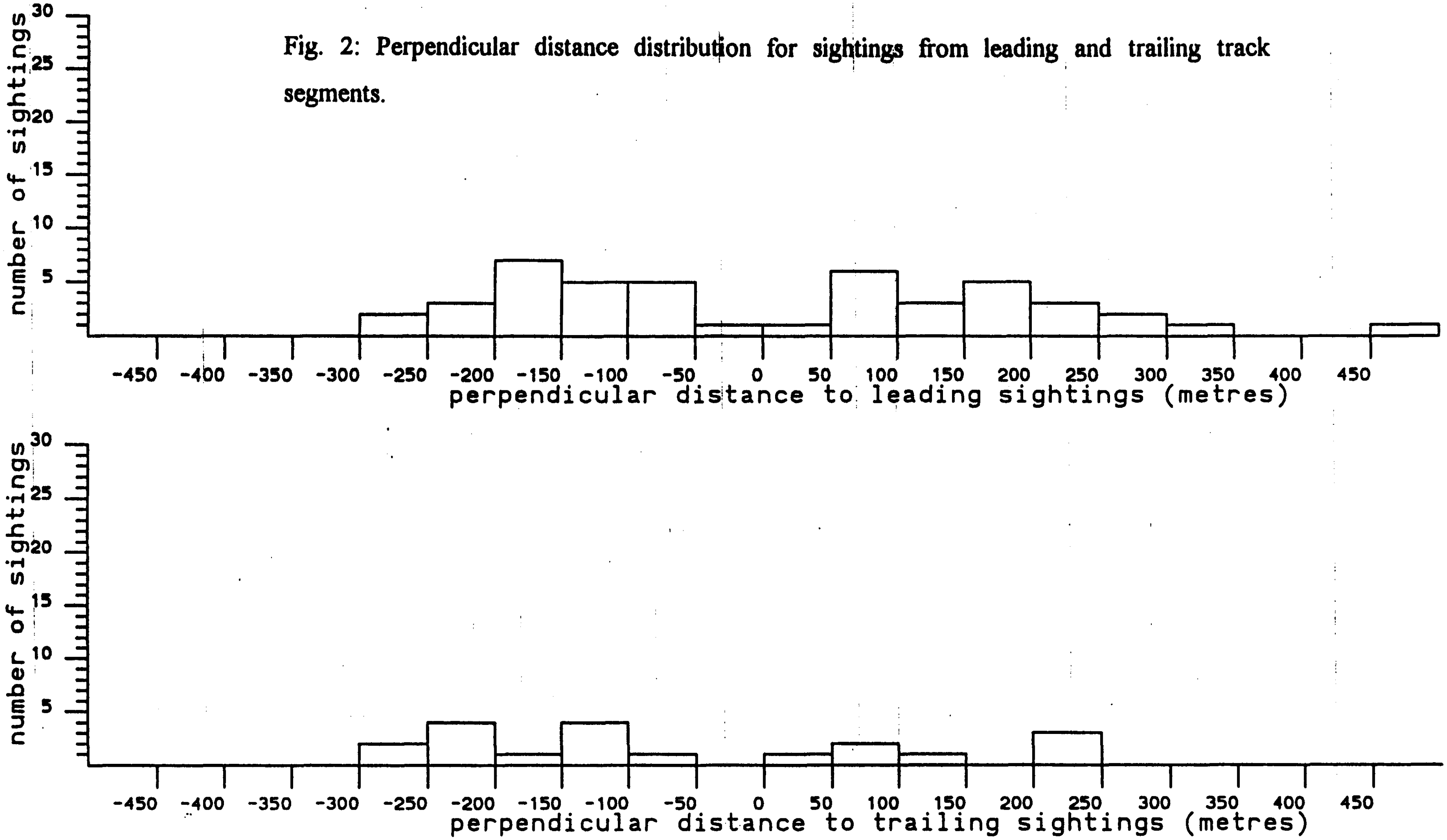


Fig. 2: Perpendicular distance distribution for sightings from leading and trailing track segments.



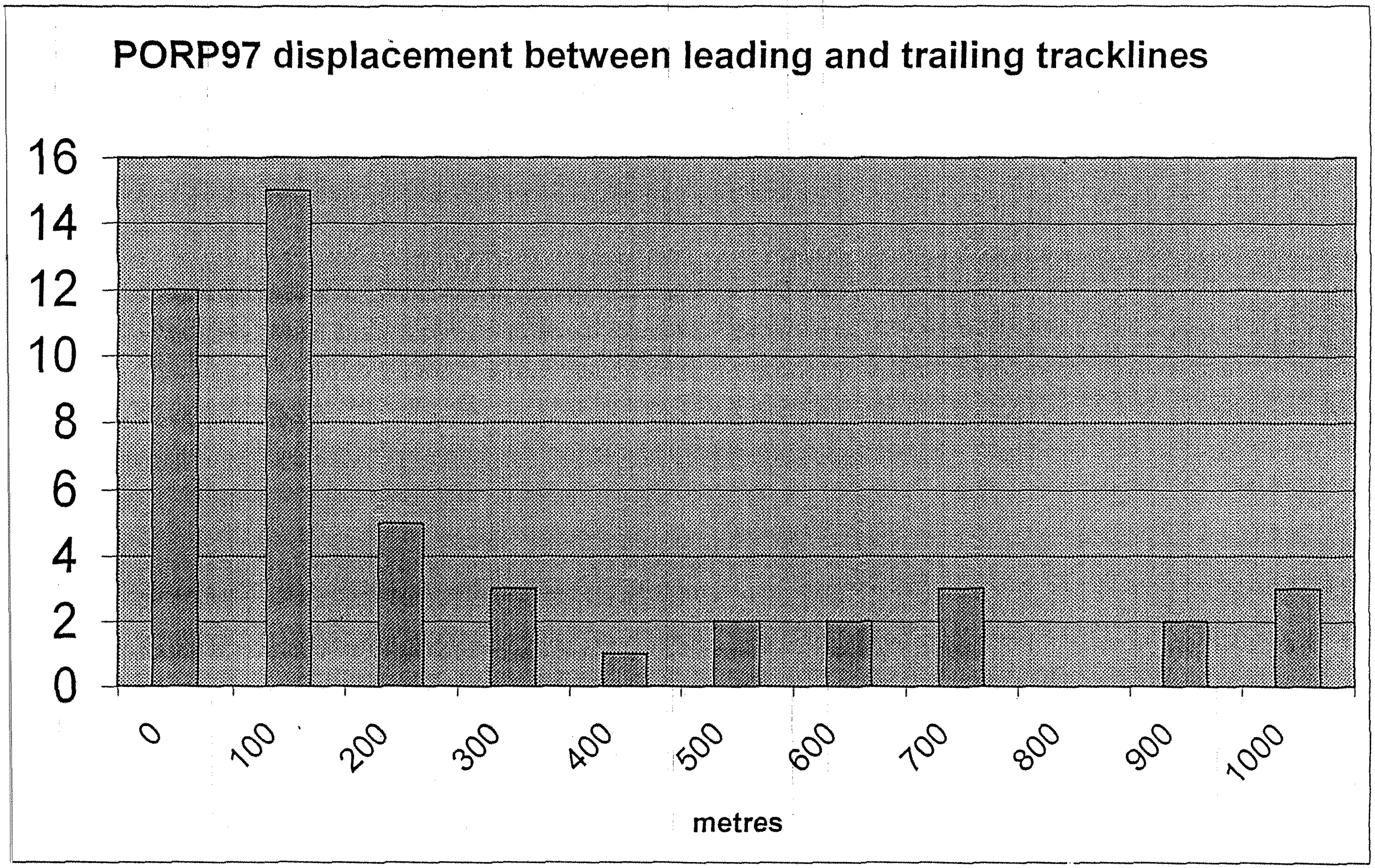


Fig. 3: Frequency distribution for displacement of trailing from leading track segment perpendicular to each leading sighting. 3a: Current survey, 3b: SCANS.

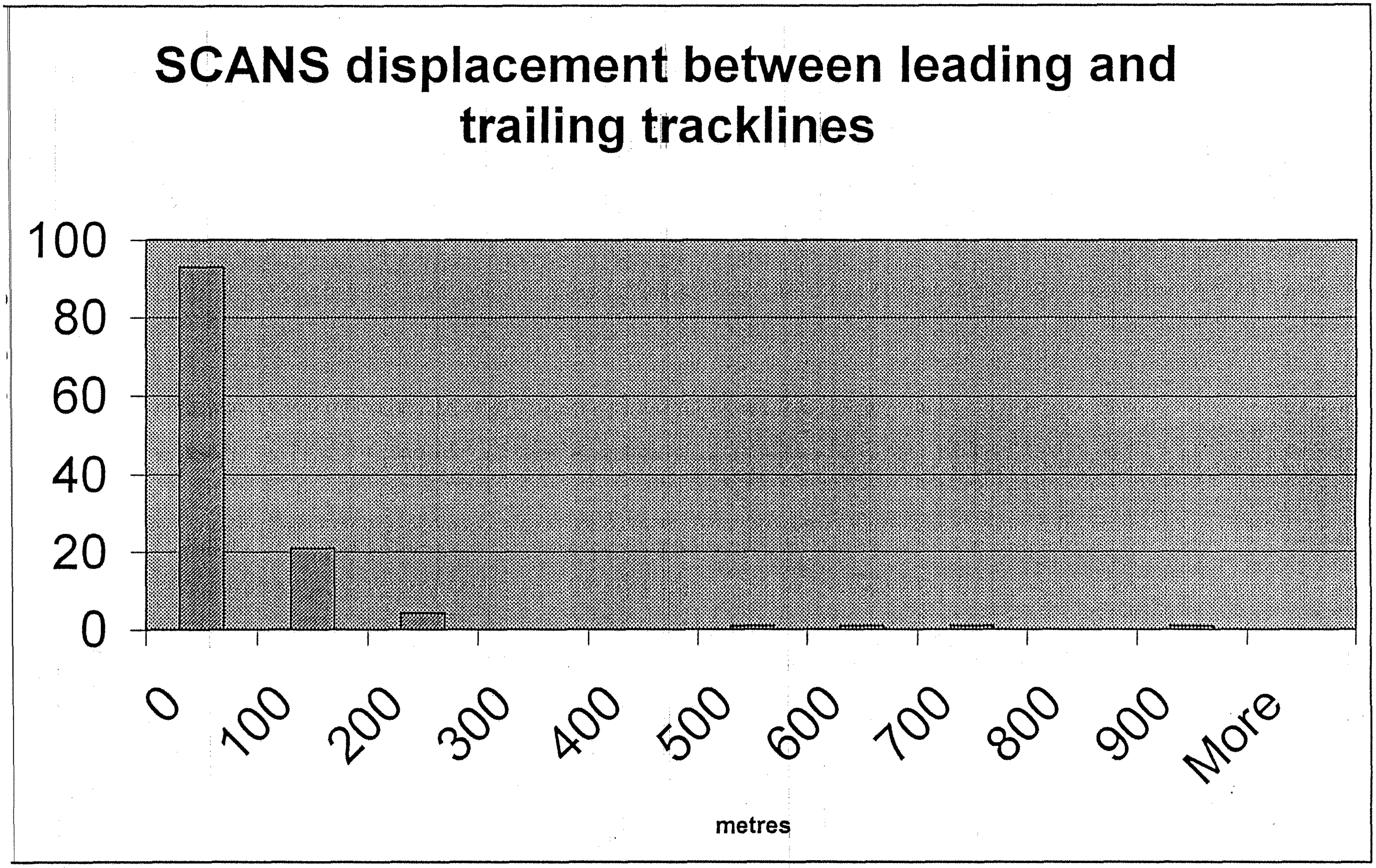

SCANS displacement between leading and trailing tracklines
100
80
60
40
20
0
0
100
200
300
400
500
600
700
800
900
More
metres

Fig. 3b

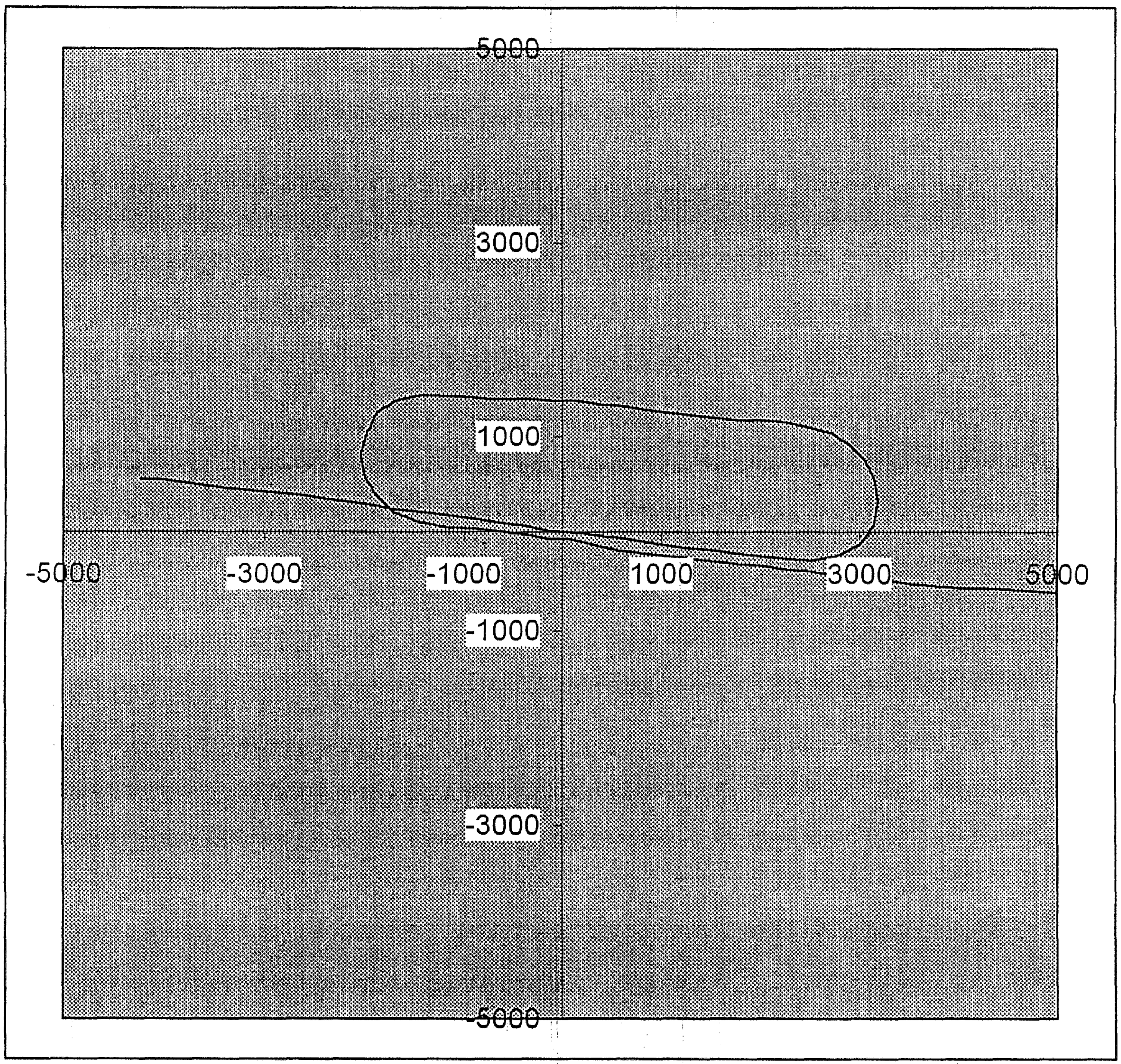


Fig. 4a: Flight path completed when circling back to attempt re-sighting from trailing track segment. Aircraft enters on left, exits on right.

4b to 4e: selected paths showing delayed convergence to trackline

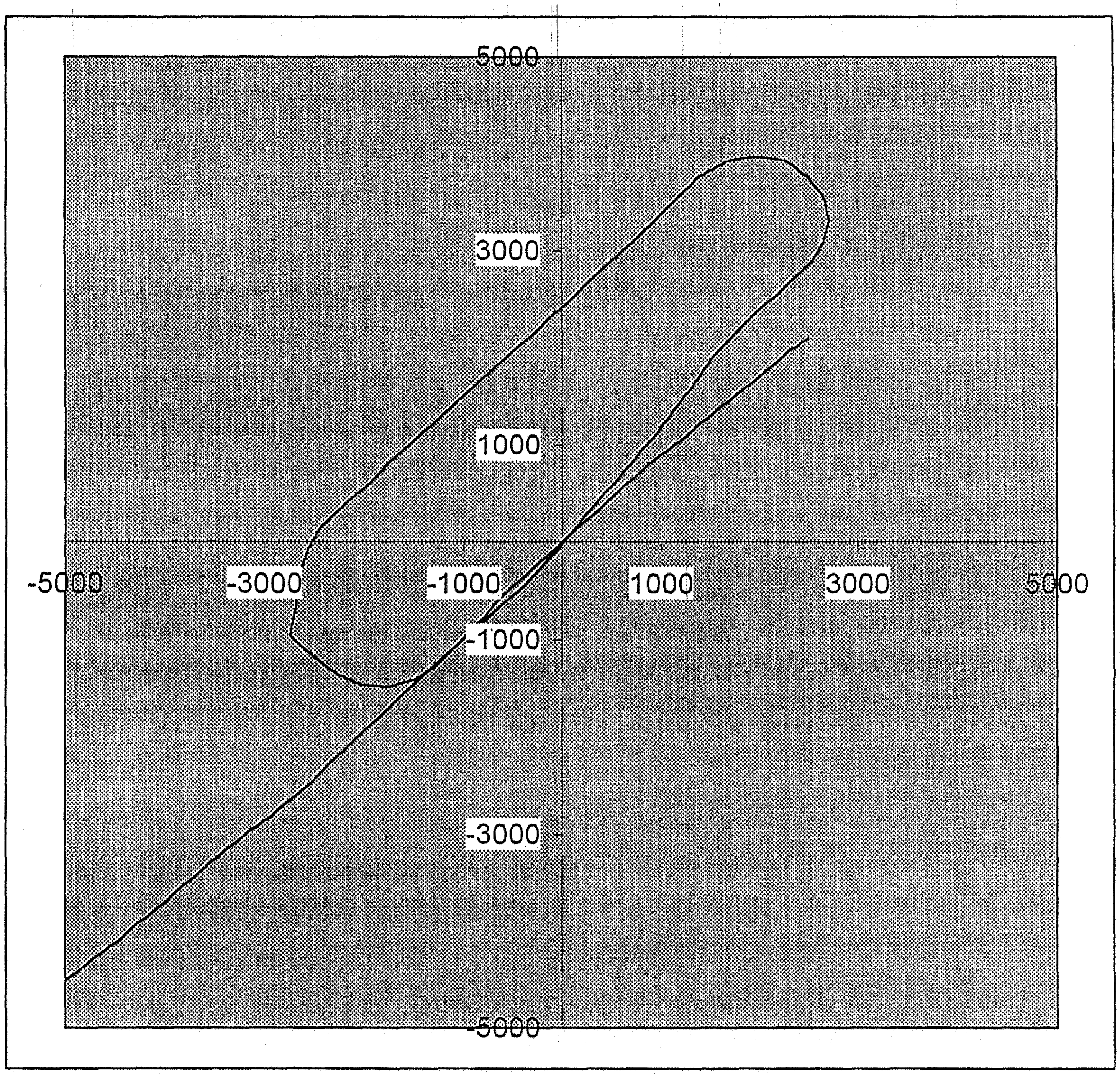


Fig. 4b

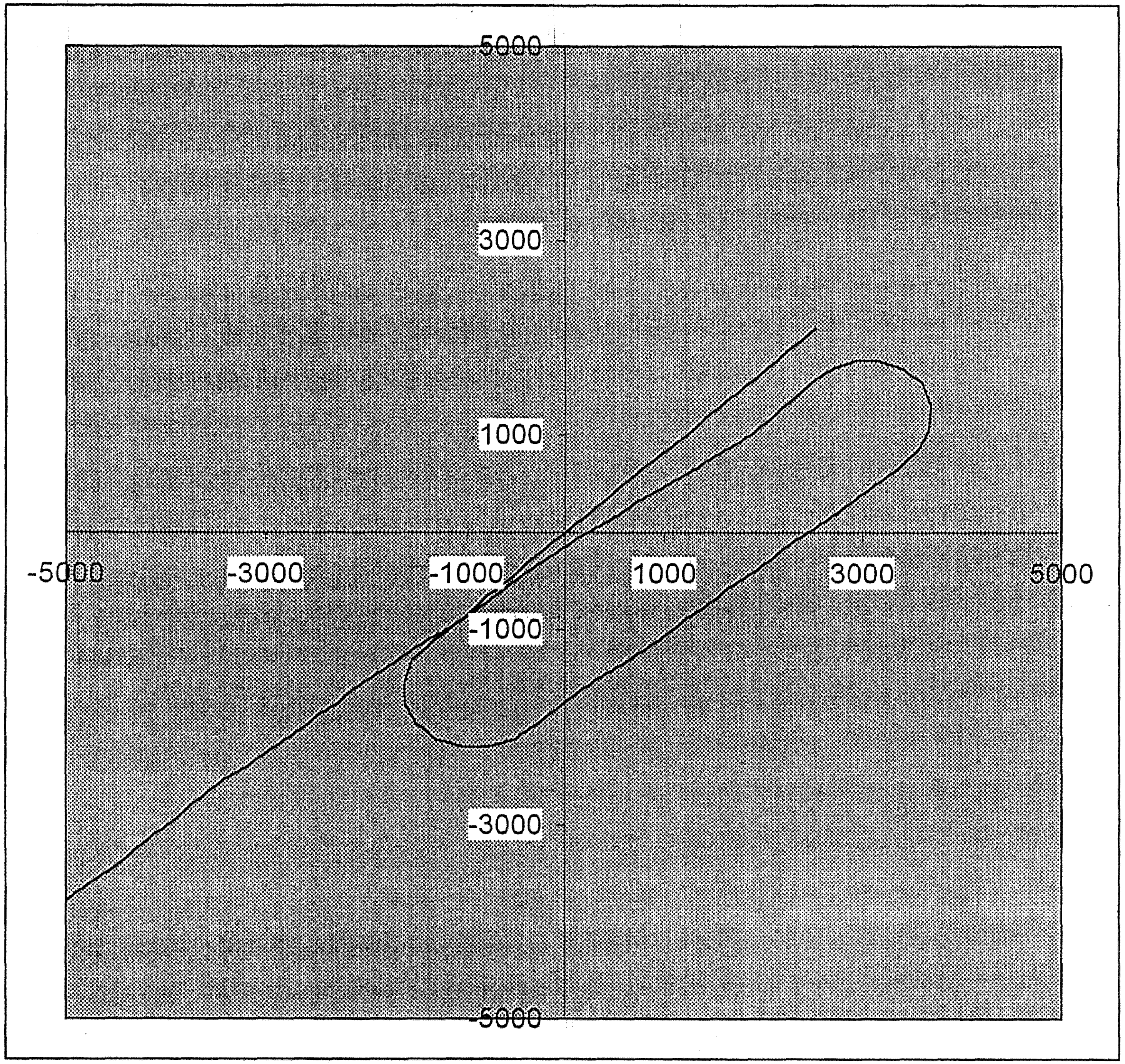


Fig. 4c

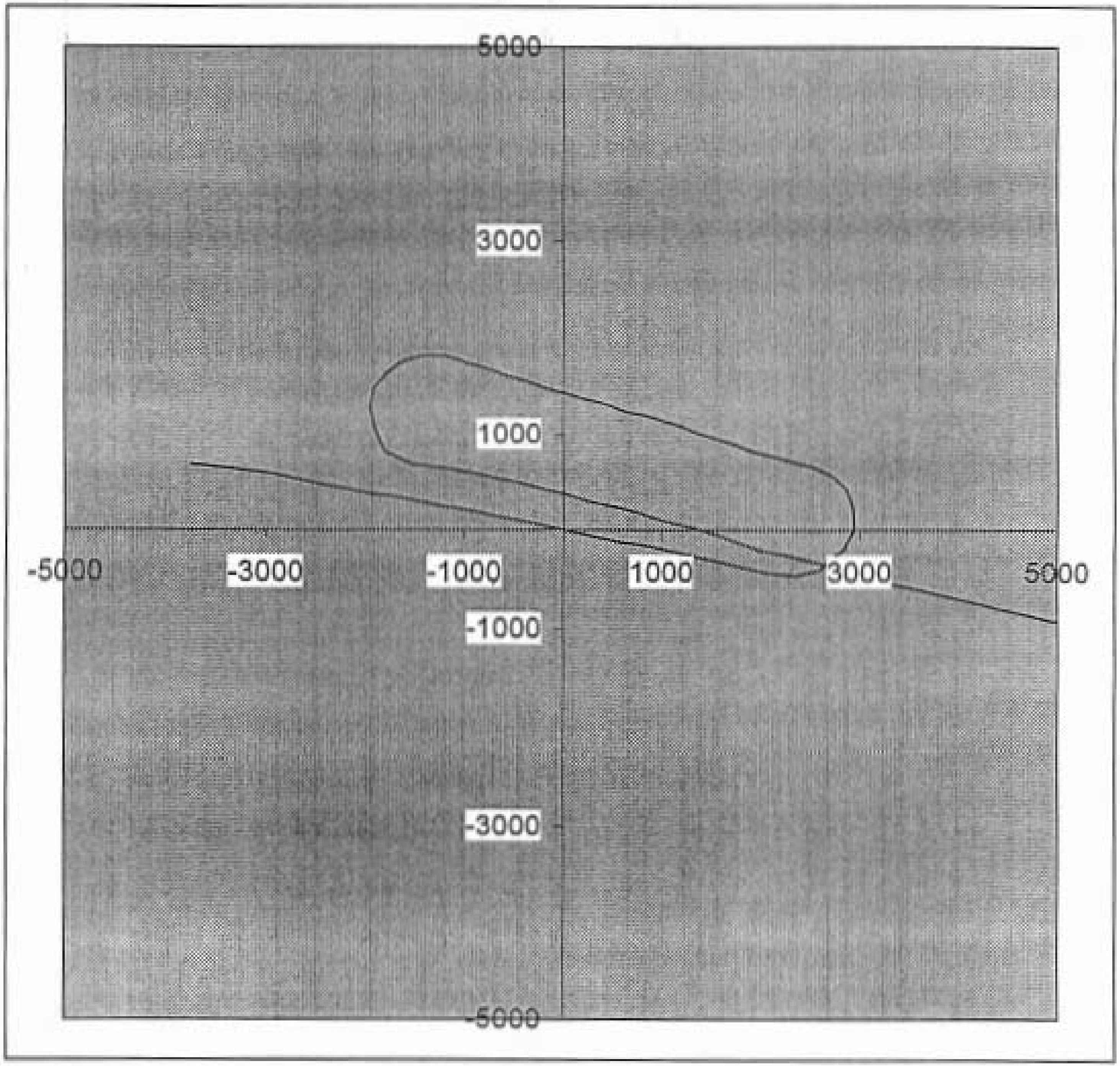


Fig. 4d

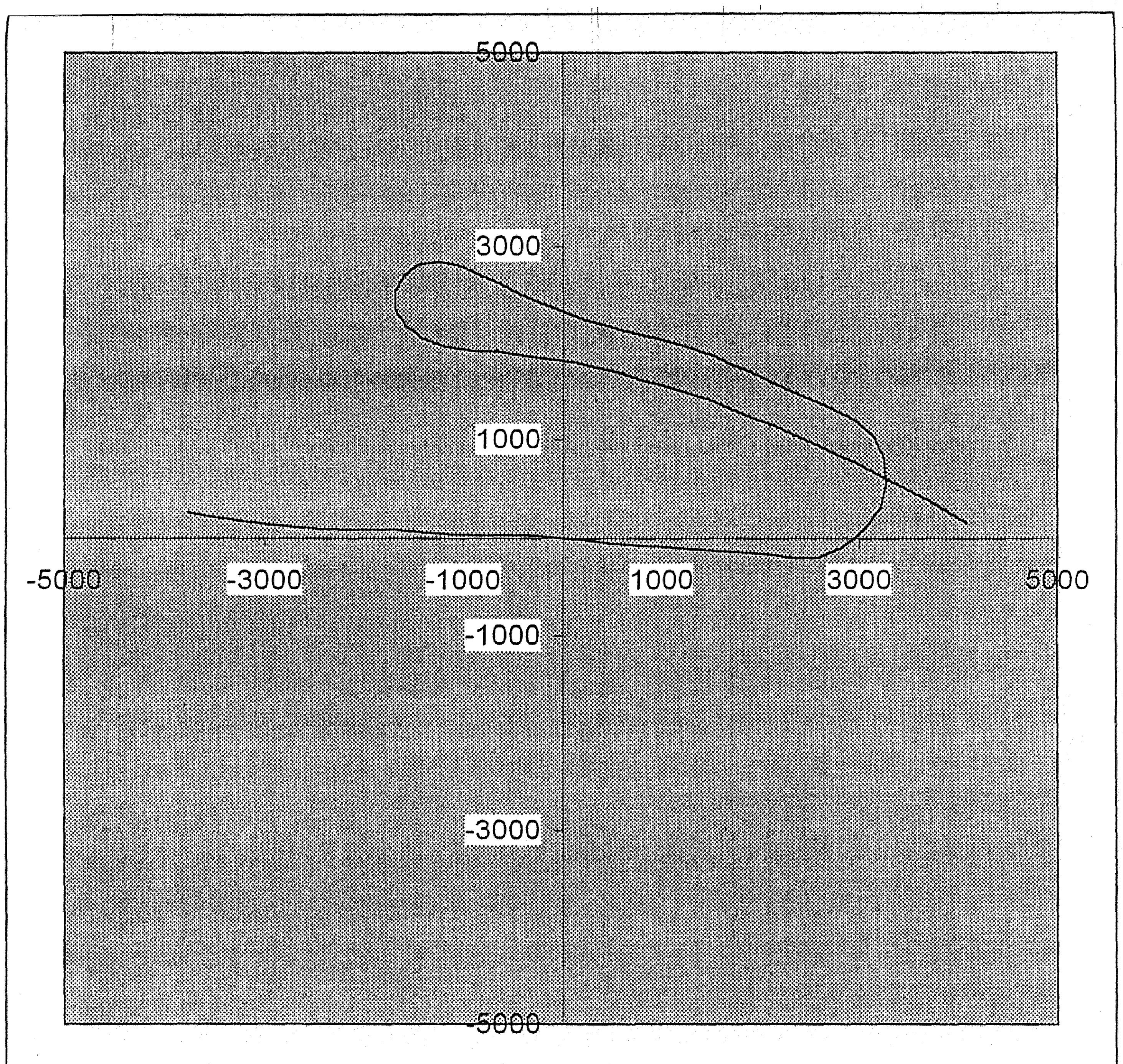


Fig. 4e

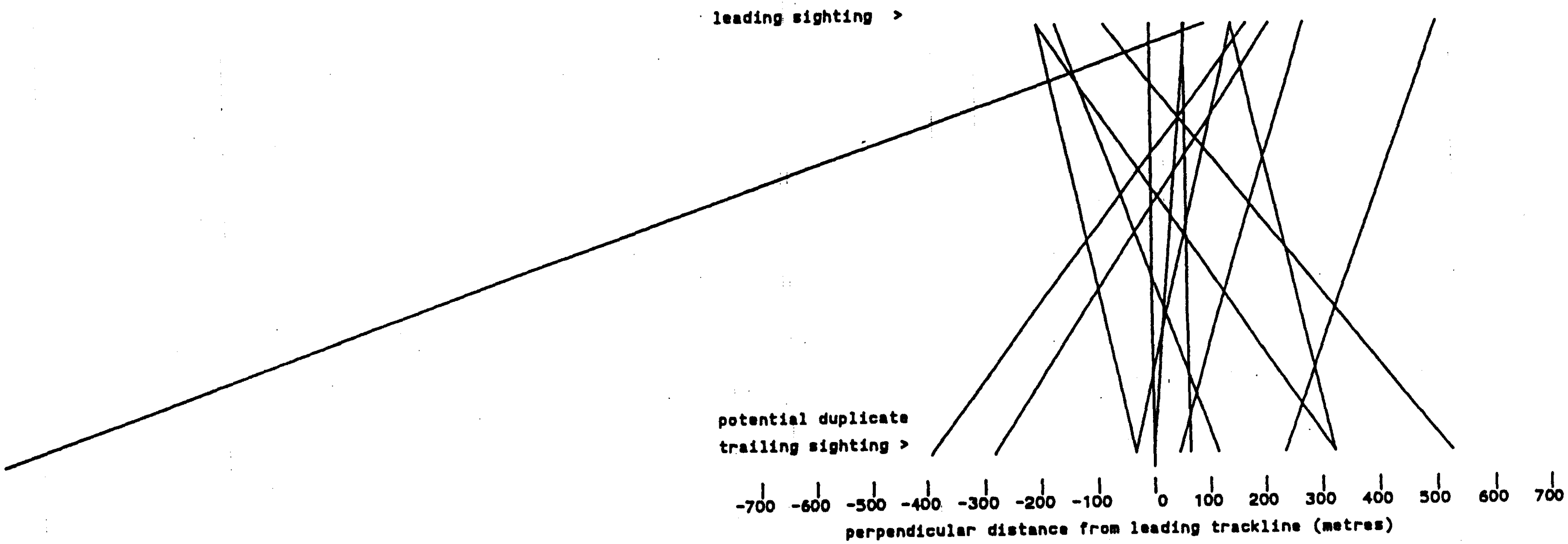


Fig. 5: Lateral movement relative to leading track for all leading/trailing sighting pairs where the trailing sighting is within 15s of the time expected for a duplicate sighting. Current survey.

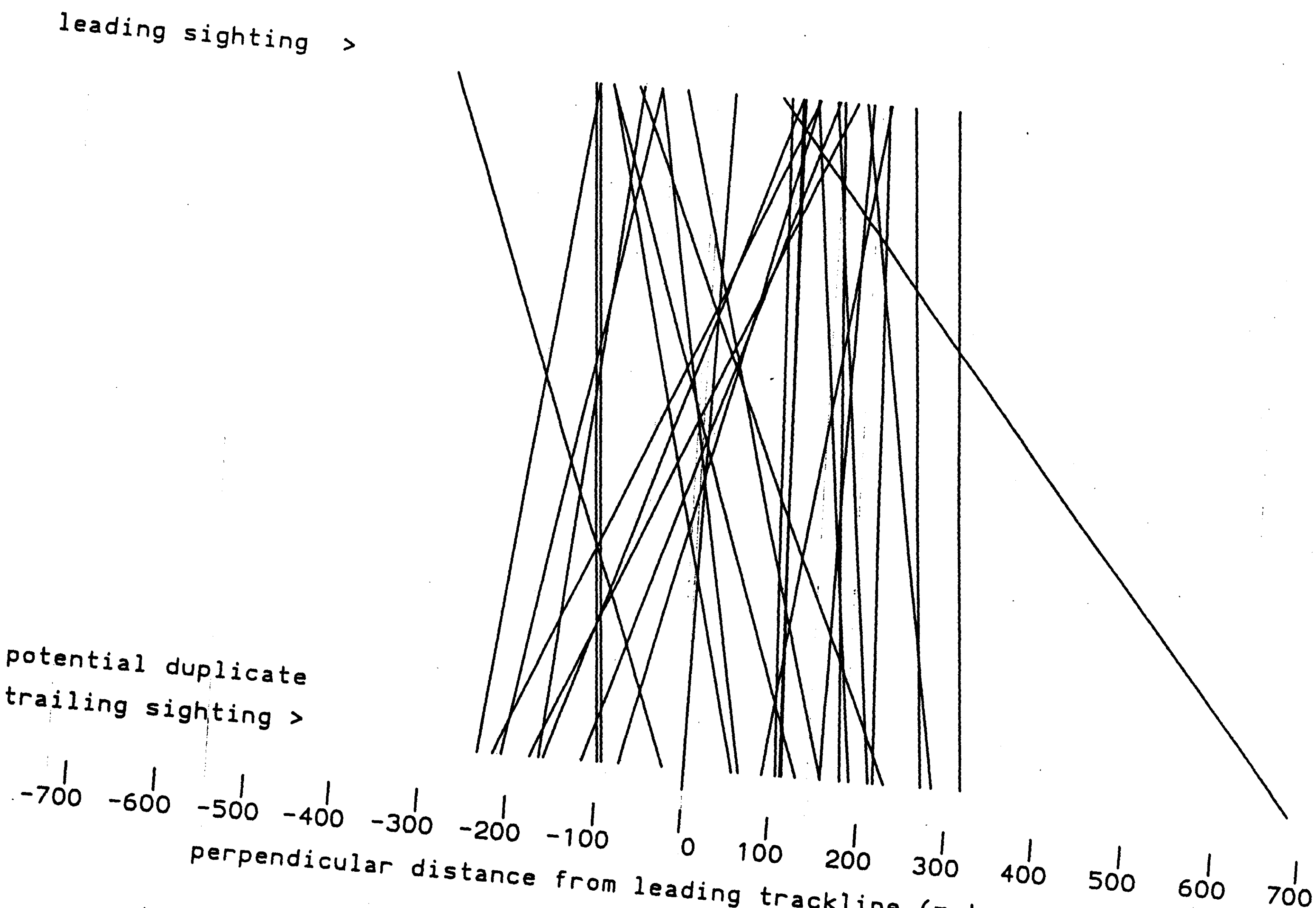


Fig. 6: Lateral movement relative to leading track for all leading/trailing sighting pairs where the trailing sighting is within 10s of the time expected for a duplicate sighting. SCANS.

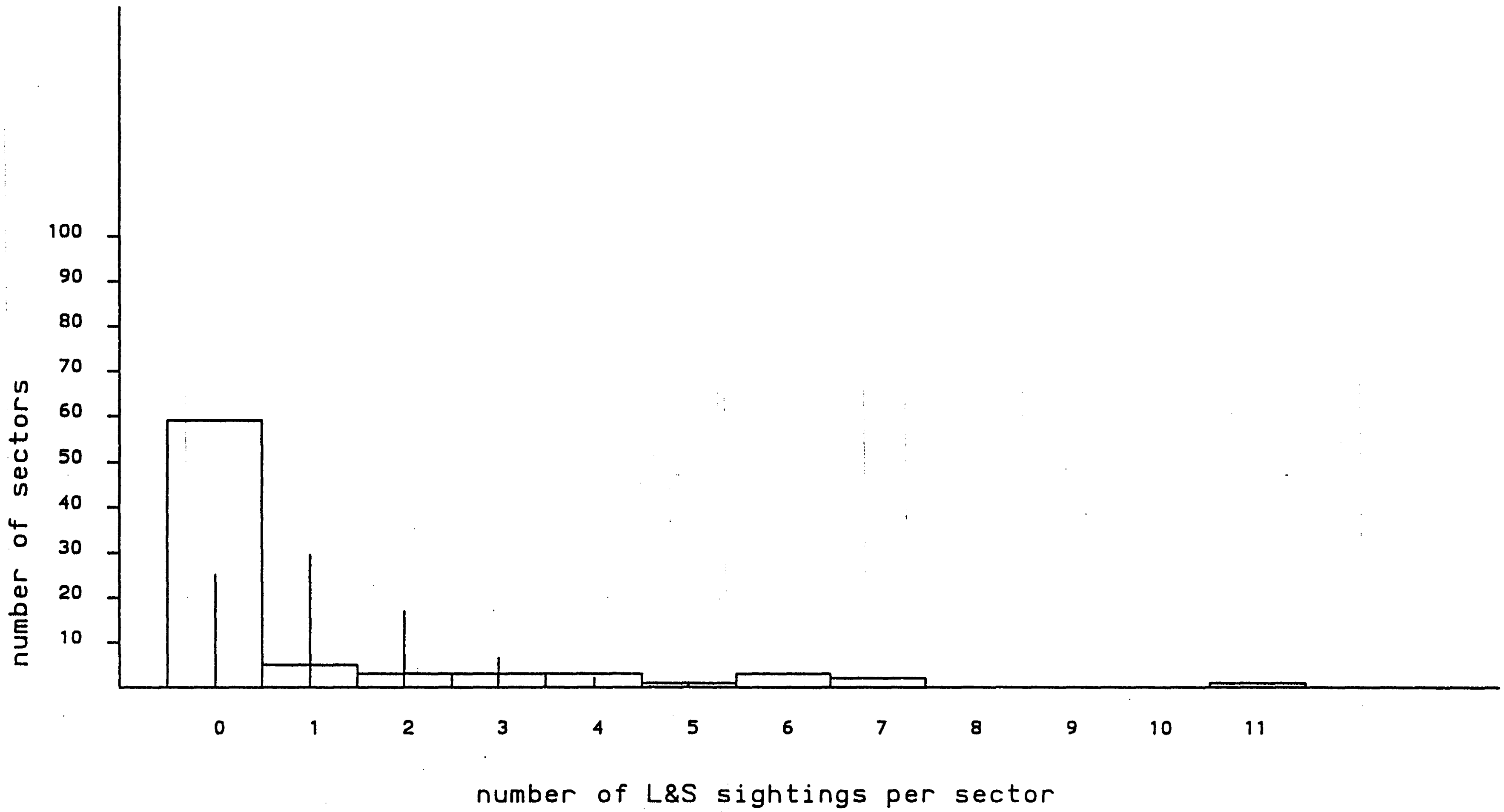


Fig. 7: Frequency distribution for number of sightings per sector when total survey divided into 81 successive surveys of equal length. Vertical lines give expected distribution. 7a: random distribution of pods over whole survey area. 7b: zero pod density over 70% of survey area.

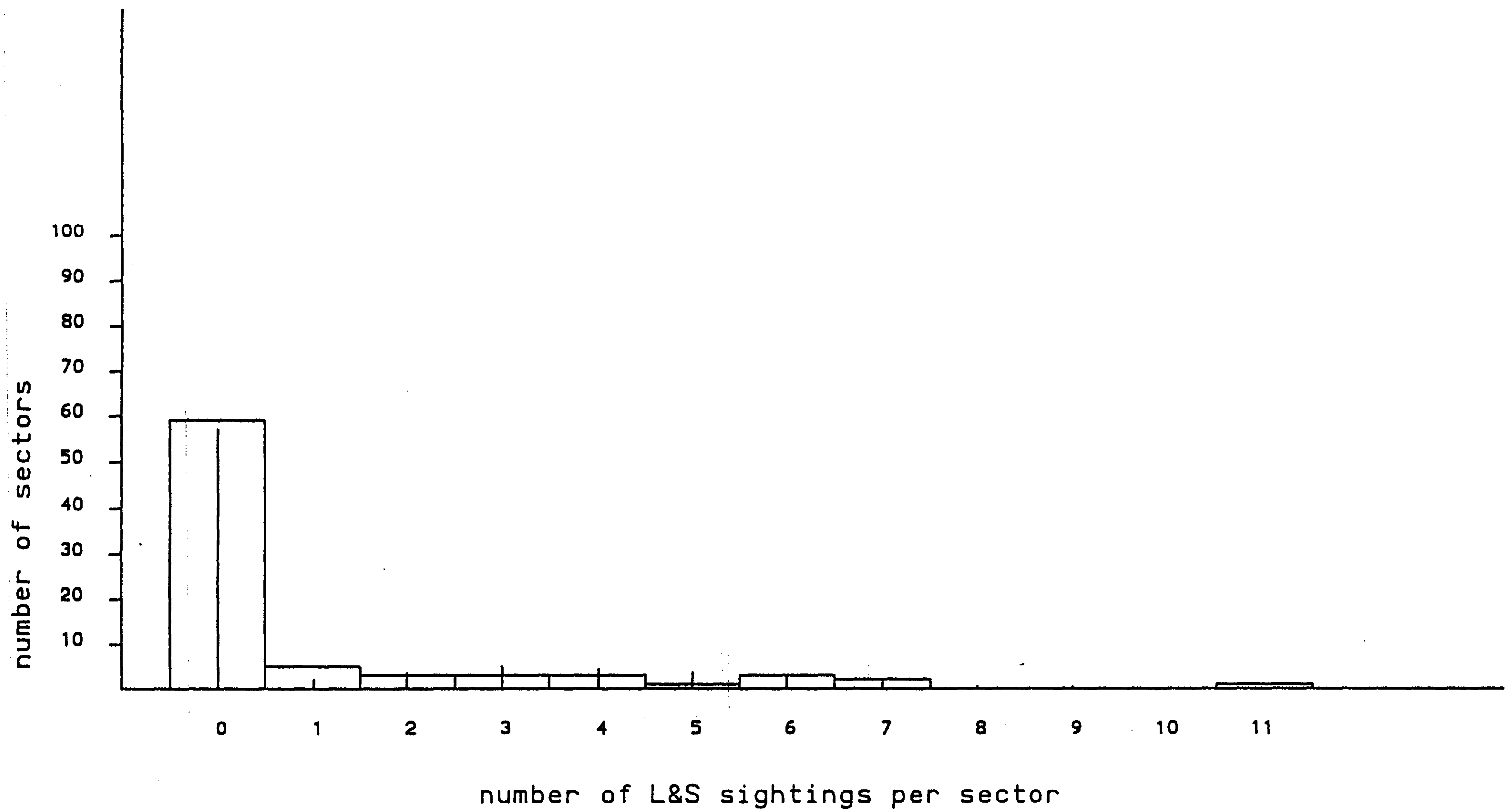

number of sectors
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
5
6
7
8
9
10
11
number of L&S sightings per sector

Fig. 7b

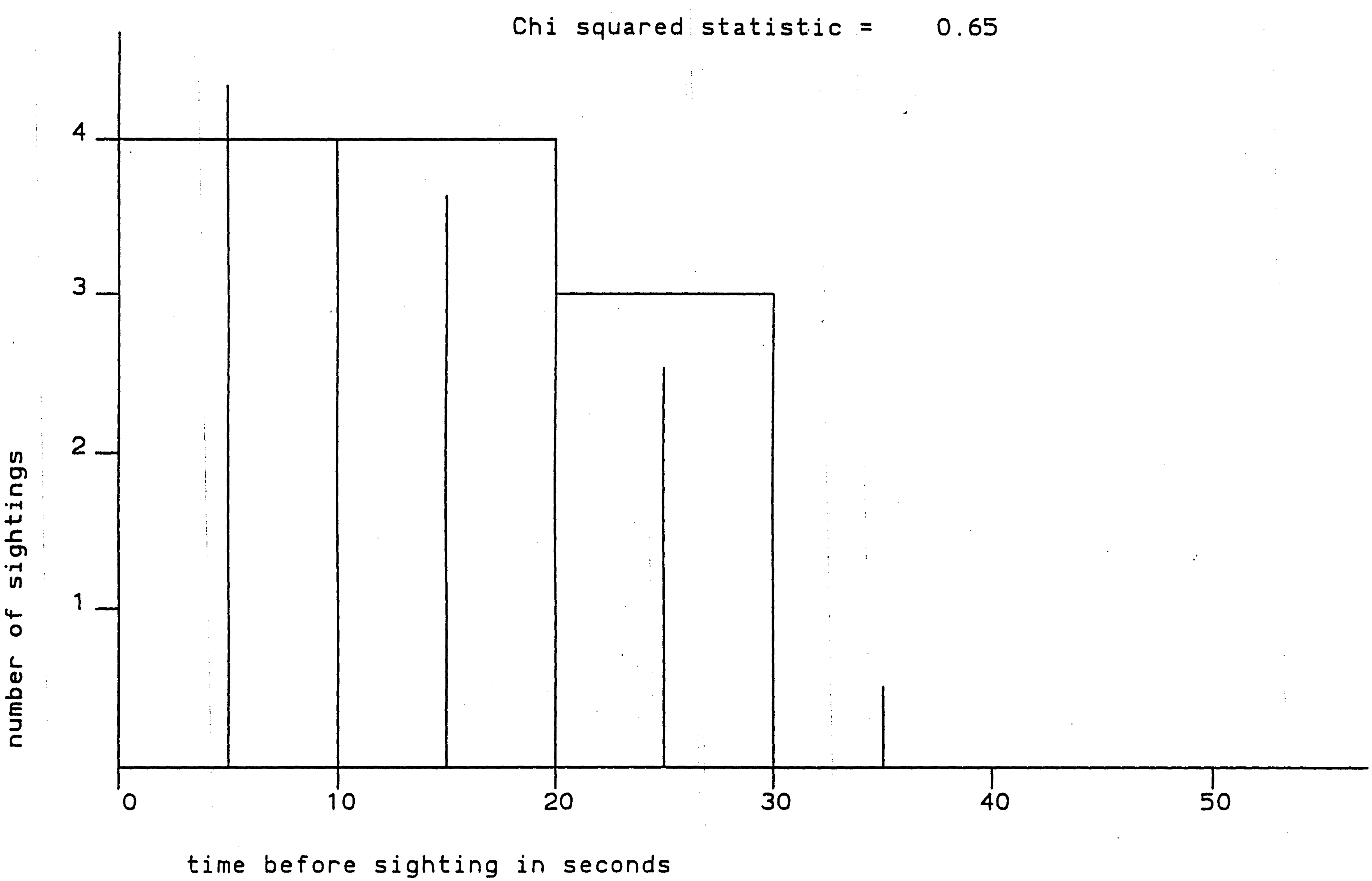


Fig. 8: Frequency distribution for inter-sighting intervals on leading and trailing track segments. Vertical lines give expected distribution given random local distribution of pods.

Appendix: (submitted for publication in proceedings of Seattle meeting on monitoring marine mammals)

The objective identification of duplicate sightings in aerial survey for porpoise.

Lex Hiby
Conservation Research Ltd., Cambridge, England

ABSTRACT: This paper describes simulation trials of an estimator for effective strip width, based on duplicate sightings from a single aircraft. The aircraft circles back after each sighting to overfly the location of the sighting a second time. Pods seen during these repeat flights are called "trailing" sightings and those seen before the aircraft circled back "leading" sightings, because the data collected is similar to that collected from two aircraft in tandem formation. Because it is not known which (if any) of the trailing sightings are re-sightings of the leading ones, the likelihood of the sighting locations is summed over all possible pairings. The resulting estimates of effective strip width are robust to failure of many of the assumptions used in derivation of the likelihood

1 INTRODUCTION

Aircraft are often used for line transect surveys of cetaceans. They offer a number of advantages over shipboard survey. For example, aerial survey allows waters between islands or near convoluted coastlines to be effectively covered and access to coastal airports allows surveys to be conducted only under optimum conditions. Distance estimation is more accurate, any error in declination angle representing a smaller proportion of the angle being measured, and most distances can be measured abeam, eliminating the need to measure horizontal angles to sightings.

However, for cetaceans normally existing as solitary animals or in small pods, there is bound to be a serious risk of missing diving animals from a fixed wing aircraft. One option is to accept the existence of an unknown level of negative bias and use the survey to generate relative abundance estimates only (e.g. Heide-Jorgensen, Mosbech, Teilmann, Benke & Schultz 1992, Barlow, Oliver, Jackson & Taylor 1988). However, because the bias may be large there is a risk that it will differ significantly between different surveys and hence invalidate comparison of the relative estimates. If an estimate of absolute abundance is required there are two ways of quantifying the risk of missing animals. One is to use behavioural measurements to estimate the probability an animal will be "available" as the aircraft passes overhead, and independent observers on the aircraft to estimate what proportion of available animals are actually seen. For some species, however, availability may depend on factors such as water depth and turbidity, making it difficult to apply the behavioural measurements to the survey context. The second option is to use independent observers on separate platforms to estimate both availability and the proportion of available pods seen, as a single composite parameter. That method was used to estimate the effective strip width (*esw*) during the SCANS surveys for porpoise in the Baltic and North Sea (Hiby & Lovell 1995). Independent observer teams were located on aircraft flying in tandem formation at a separation of about three minutes along the trackline. The technique was subjected to simulation tests to check for robustness to any failure of the model assumptions (Hiby & Lovell 1998).

Duplicate sightings were also collected during the shipboard components of the SCANS surveys (Borchers, Buckland, Clarke & Cumberworth 1995). Although both teams of independent observers were on the same vessel, the regions observed by the two teams were separated by having one team use binoculars to search an area far ahead of that searched by the other team. Whereas duplicate sightings made during the shipboard survey could be identified by using a third team to track sightings, there was no way to identify duplicates between the aircraft in real time. The likelihood for the aerial survey data was therefore calculated by summing the likelihood for observed sighting locations over all possible duplicate pairings between the leading and trailing aircraft.

This paper describes how the same technique could be applied to data collected from one aircraft circling back to overfly selected sightings a second time. Insufficient survey data have been collected using this technique up to now so the results reported below are restricted to analysis of simulated data.

2 METHODS

2.1 *Data collection*

The data collection protocol for the aerial survey components of the SCANS surveys is described in Hiby & Lovell (1995). Briefly, observers were located on two aircraft flown in tandem formation about nine kilometres apart. At this separation the trailing aircraft encounters a pod about three minutes later than the leading aircraft. This is well in excess of the average duration of the dive cycle (Westgate, Read, Berggren, Koopman & Gaskin 1995), hence the probability that the pod is in the near-surface phase of its dive cycle when passed by the trailing aircraft is approximately independent of its status when passed by the leading aircraft. The output from the GPS navigator in each aircraft was logged continuously and following each flight these records were combined with the sighting times (recorded to the nearest second) and declination angles to generate a database of pod sighting positions for the leading and trailing aircraft.

Simulation trials suggest that the estimates of *esw* from the SCANS aerial survey data were reasonably robust, in terms of bias, to failure of the assumptions used to derive the likelihood for sighting locations. However, the coefficient of variation (CV) on the *esw* estimate was about 0.25, based on a likelihood ratio calculation, or 0.2 based on the simulation results. Either way, this represents a significant contribution to the error on the eventual estimate of pod abundance. The costs of mounting a tandem survey are also very high and the logistics complex. These considerations motivated design of a revised technique which could exploit the advantages of the tandem estimator but use data collected from a single aircraft. A single aircraft technique would halve the cost of the survey, simplify the logistics and, by establishing a protocol which could be carried out routinely, allow sufficient data to accrue from a given survey team to reduce the CV to a low level.

If duplicate sightings are to be obtained using a single aircraft certain sections of the trackline have to be repeated. The track design for SCANS used a “zigzag” flight path between waypoints located on the boundaries of the survey area. The first option considered with respect to placement of repeat flights was to fly back along the last few miles of each transect “zig” or “zag” before moving to the next waypoint. The flight over the last few miles up to the waypoint would then be regarded as that by the leading aircraft and the return flight away from the waypoint as that by the trailing aircraft. One disadvantage of that approach is that the derivation for the probability of detection by the leading aircraft only, as given by equations (2) and (3) of Hiby & Lovell (1998), is no longer valid because the separation between leading and trailing aircraft is no longer constant, even approximately. A second disadvantage is that unless pod density is very high, many flight segments adjacent to the waypoints may provide no sightings at all and hence no information on the proportion of pods missed along the trackline.

Both these disadvantages can be avoided by restricting repeat flights to segments which have already yielded at least one pod sighting and conducting the repeat flight in the same direction as the original flight. A possible flight pattern is illustrated below. The aircraft conducts survey flights along transects lying between successive waypoints, as usual. A short "dead-time" is imposed following departure from each waypoint. Any pod sighting following the dead-time then initiates a break-off from the trackline after an interval of about 30 seconds, or 1.5 km, to allow the aircraft to fly back and rejoin the trackline at a point about 1.5 km before the point from which the sighting occurred. The aircraft then resumes survey along the original trackline, passing the location of the original sighting, and continues to the next waypoint. This results in a “racetrack” flight pattern that is familiar to most pilots as a holding pattern for aircraft waiting to land at a busy airport (figure 1).

$$r(t) \sim \frac{e^{-r/(t\beta)}(t\beta)^{-\alpha}\, r^{\alpha-1}}{\Gamma(\alpha)} \tag{2}$$

The scale parameter was proportional to t, corresponding to movement in constant direction at constant speed is. α and β were introduced as the third and fourth free parameters.

4. The hazard rate function (Buckland 1985) was used for the probability of detecting a pod at perpendicular distance y from the aircraft given it is near the surface:

$$g(y) = 1.0 - \exp\left\{-(a/y)^b\right\} \tag{3}$$

Detection of pods is impossible during their diving phase so that unconditionally the probability of detection of a pod passing abeam at distance y is S/C times $g(y)$. a and b were introduced as the fifth and sixth free parameters.

The models were used to define the probability, P, of a pod's detection by the leading aircraft and the probability, p_{01}, of its detection by the trailing aircraft only. For example,

$$P = \frac{S}{C}\int_{-.5}^{.5} g(y)dy. \tag{4}$$

P and p_{01} were then used to define the likelihood for each possible arrangement of leading and trailing sightings on a survey leg into duplicate pairs (including the case where none of the sightings were considered to be duplicates). Those likelihoods were then summed over the possible arrangements to give the likelihood for the sighting locations observed on that leg and the logs of those sums summed over all legs to give the log likelihood for the whole survey. For example, a given arrangement on a given leg will generate n_{10}, n_{01} and n_{11} sightings by the leading, trailing and both aircraft on that leg. The joint density for the intervals preceding the $n_{10} + n_{01} + n_{11}$ sightings is then:

$$[VD(P + p_{01})]^{(n_{10}+n_{01}+n_{11})} \cdot \exp[-VD(P + p_{01})T] \tag{5}$$

where T is the duration of the leg. This is the first of three terms required for the likelihood of that arrangement. The second and third terms (equations (5) and (6) in Hiby & Lovell 1998) deal with the probability density for the locations of the pods seen by either the leading or trailing aircraft only, and the locations of the pods seen by both.

The situation for the one-aircraft technique is the same except for the existence of the "single-effort" sections between those chosen for re-survey. Each duplicate section now lasts only about one minute, with single sections running from the end of one duplicate section to the start of the next. The single sections have to be incorporated because the duplicate sections will have a higher than average pod density. If not balanced by the intervening sections, this would distort the probability that a "trailing" sighting (i.e. one seen during a re-survey section) is new rather than a re-sighting of a "leading" one. The likelihood for each single-effort section was obtained simply by eliminating the elements relating to the trailing aircraft. So, for example, the joint density for the intervals preceding the n_{10} sightings on a single section is:

$$[VDP]^{n_{10}} \cdot \exp[-VDPT]. \tag{6}$$

One further modification was made so that the assumption that pods are randomly distributed at constant density could be relaxed. If a survey is split into a number of successive sections of constant length, K, the number of empty sectors will normally be found to exceed the number expected by Poisson distribution of pod sightings. The survey area may include regions that do not represent suitable porpoise habitat and have zero density. The mean density over the entire region will thus be less than the density in those areas where potential duplicate sightings are recorded. The value of D required in the above equations will thus be underestimated, resulting in an underestimate of the risk that leading and trailing sighting in close proximity may be different pods. To overcome this difficulty the survey was split into sections of length K and the number of empty sections, s_0 and non-empty sectors, s_1 noted. As the number of leading sightings on a survey leg, $n_{10} + n_{11}$ is independent of the chosen pairing arrangement, the probability density for the intervals preceding leading sightings do not need to be included in the recursive code used to deal with the trailing sightings. They could thus be combined with the single sightings over all legs to give

$$[VDP]^{(n_L+n_S)} \cdot \exp[-VDP\Psi] \tag{7}$$

where n_L and n_S are the number of leading and single sightings over the whole survey and Ψ is the duration of the whole survey. Let z be the probability that a section of length K falls within a zero-density region. The probability of a section having zero sightings is then z plus $(1-z)$ times the probability that a sector in the non-zero region will have no sightings. Given s_0 empty sectors and s_1 non-empty sectors the density for intervals preceding all leading and single sightings is thus modified to

$$[VDP]^{(n_L+n_S)} \cdot \left((1-z)\exp[-VDPK]\right)^{s_1}\left(z+(1-z)\exp[-VDPK]\right)^{s_0}. \tag{8}$$

z was included as the seventh and last free parameter used to maximise the likelihood for the sighting locations.

2.3 *Simulation model*

The performance of the ML estimators was checked against simulated data. Initially, the stochastic models used in the simulation model corresponded exactly to those used to derive the likelihood, then certain aspects were changed to check the robustness of the estimates to departures from the assumptions.

The simulation designed to test the one-aircraft technique was based on that designed to test the two-aircraft technique. To simulate sightings, encounter times were generated by incrementing the previous encounter time by a number drawn at random from an exponential distribution (with expectation determined by the ground speed of the aircraft and the assumed pod density). Each "zig" or "zag" of a survey flight was assumed to last for 30 minutes so if an encounter time exceeded 1800 seconds a new 30-minute section was initiated. The number of sections per run of the simulation was varied according to the assumed mean pod density to give the same expected number of pod detections on each simulation run (about 300). The pod was either detected or missed by the "leading" aircraft, its probability of detection depending on the probability it was near the surface at that time and its distance from the trackline, which was uniformly distributed from 0 to ½ km. The hazard rate function was used for g(y), the detection probability for a pod near the surface at perpendicular distance y from the aircraft.

The probability a pod was near the surface when the aircraft passed overhead was equated to the fraction of the dive cycle, C, occupied by the near-surface phase, S. C and S were set at 60 and 30 seconds respectively for most simulation runs. By this construction detection is certain for a pod near the surface and close to the trackline, and any risk of missing such a pod in a real survey would be subsumed into an underestimate of the near-surface fraction of the dive cycle. The probability it was still, or again, at the surface by the time it came abeam of the trailing aircraft was then calculated by function $U(t)$. The separation time of three minutes was, however, substantially longer than the mean dive cycle duration so that the probability it was near the surface when it came abeam of the trailing aircraft was always near S/C. The perpendicular distance of the pod from the trailing aircraft was calculated by picking the distance over which it moved while waiting for the trailing aircraft, $r(t)$, from the Gamma distribution. The direction of movement was chosen from a uniform distribution between 0 and 360 degrees.

The probability the pod was detected by the trailing aircraft was again calculated using the hazard rate function. On detection by either the leading or trailing aircraft the time and declination angle were recorded and subjected to normally distributed and independent recording errors.

Any of the parameters involved in the simulation could be chosen interactively and the starting values for the estimation equated to those used in the simulation or perturbed from those values.

In order to generate data for the revised, one-aircraft, method, those trailing sightings that did not occur from 2.5 to 3.5 minutes after any leading sighting were eliminated. In the revised procedure the aircraft re-joins the trackline 2.5 minutes after the sighting that initiated the break-off occurred and reaches the break-off point 3.5 minutes after the sighting occurred. Sightings on the repeat flight between these points would be classed as "trailing" sightings by the data management program used to update the database of sighting locations. Furthermore, the leading sightings used to define these intervals were restricted by imposing dead-times, as described above. Thus some sightings were generated on "single" legs and were recorded as single sightings for the estimation code and the remainder were recorded as leading or trailing, according to whether they were seen on the first or the repeat flight over the 2.5 to 3.5 minute interval. Note that by this method of construction there must be a minimum of one leading sighting on each duplicate-effort section and any further leading sightings must occur during the second half of that section, but there are no restrictions on the numbers of trailing or single sightings.

3 RESULTS

A number of simulation trials reported in Hiby & Lovell (1998) investigated the effect of allowing the simulation models to diverge from those used to derive the likelihood function. These were repeated using data simulated for the single-aircraft technique. They included the following departures from the model assumptions:

1. The alternating Poisson process for near-surface and deep-dive phases of the dive cycle was replaced by normally distributed near-surface and deep-dive durations with means and standard deviations as suggested by data from time-depth recorders (Westgate et al. 1995).
2. The Gamma distribution for pod displacement was replaced by movement at the same speed by each pod, so that the locus for the position of any pod at the moment the trailing aircraft passed overhead was a circle centred on its original position. Secondly a diffusion model, corresponding to porpoises "milling" rather than moving in straight lines, was used.
3. The uniform distribution for the direction of pod displacement was replaced by movement away from the trackline only (selecting at random from the semicircle of divergent angles), to simulate avoidance reaction to the leading aircraft

The results were the same as for the tandem technique, that is, (1) and (2) produced no more than 2% bias in estimated effective strip width. The simulated avoidance behaviour in (3) did produce serious downward bias in estimated *esw* but also generated a clear indication of avoidance behaviour in the sighting distances to a sample of potential duplicates (i.e. sightings by the trailing aircraft within ten seconds of the expected time for a duplicate sighting). The SCANS aerial survey data did not show a similar level of avoidance behaviour and the bias was eliminated when the simulated avoidance was restricted to pods passed within 0.1km by the leading aircraft.

A number of additional simulation runs were carried out using data simulated for the single-aircraft technique. These were targeted towards the effect of varying the amount of "dead-time" imposed to prevent excessive circling, and the effect of increasing and varying pod density.

Figure 2 first illustrates the performance of the single-aircraft technique using a uniform pod density of 0·1 pods per km^2 and a dead-time of five minutes, the minimum practical value. The histograms to the left of the figure show the distribution of the maximum likelihood estimate for each free model parameter expressed as a percentage of the value used in the simulation (the values used in the simulation are given in the figure legend). The histogram to the right of the figure gives the distribution for the corresponding effective strip width estimates. The results were based on 1000 replicate simulation runs. The *esw* estimate is unbiased with a CV of about 0·11. The bias evident in the parameter estimates for the displacement model results from inclusion in the simulated data of measurement error for time and declination angle to sightings. As measurement error is not included explicitly in derivation of the likelihood its effect is subsumed in the estimates for the displacement parameters.

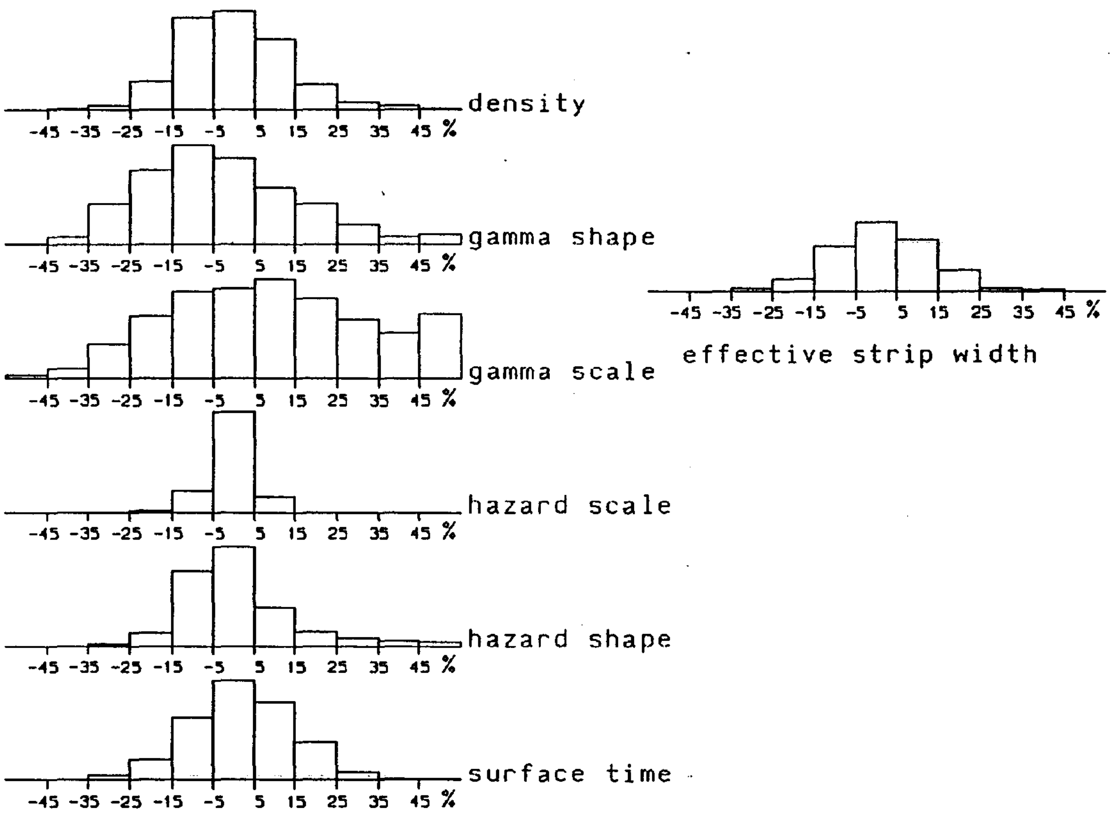


Figure 2. Results from 1000 runs of a simulation of leading, trailing and single sightings that would be obtained using the data collection protocol illustrated in figure 1. Maximum likelihood estimates of the model parameters used in the simulations were calculated from the data generated by each run. The frequency distribution for each parameter estimate, expressed as a percentage of the value used in the simulation, is given to the left of the figure. To the right is the distribution of estimated effective strip width as a percentage of the value corresponding to the parameter values used in the simulation. These were as follows: mean pod density, 0·1 pods per km^2; shape and scale parameters for the gamma distribution of pod displacement, 4 and 0·4 $m.s^{-1}$; shape and scale parameters for the hazard rate sighting function, 5 and 0·2 km; duration of near-surface phase of the dive-cycle, 30 s (dive-cycle duration was 60 s). The CV for the *esw* estimate was 0·11.

The tandem and revised, one-aircraft, techniques use the proximity of trailing sightings to the locations expected for duplicates to provide information of the number of duplicates which actually occurred on a survey. It might therefore be expected that under higher pod densities, when many non-duplicate trailing sightings will occur close to the locations expected for duplicates, this information will be degraded and the CV on the *esw* estimate will therefore increase. This was confirmed by increasing the mean pod density for the next simulation to 0·3 pods per km^2 (and reducing the trackline length by a factor of three to produce the same expected sighting frequency as the previous simulation) which resulted in an increase of 26% in the CV of the *esw* estimate. Thus, for a given number of pod sightings, estimates of *esw* are better when collected at lower pod densities. This suggests that attempting to estimate *esw* for a given survey team quickly by surveying very high-density areas is not an optimal approach. It would be better to adopt the method as a standard protocol over a longer period, which should be possible given that the equipment required for this technique does not represent a significant increase over that normally needed.

The CV on the *esw* estimate was also increased, by 10%, when the dead-time used in the simulations was increased from five to eight minutes, which is as expected given the reduction in the potential number of duplicates.

The remaining simulations addressed the effect of variation in pod density. Varying pod density over a three-fold range over the different legs of the survey had no effect on the estimate of *esw*. Similarly, allowing density to increase linearly from zero to 0·6 pods per km^2 along each 30-minute survey leg had no effect. Then a four-fold variation in density between successive five-minute sections of trackline was introduced. Such variation might be encountered by flying over high-densities associated with, say, a shelf area or tidal flow. The steps in the mean density parameter might be expected to generate some apparent coincidence between non-duplicate leading and trailing sightings and hence upward bias in the *esw* estimate. However, the bias was only around 4% and to generate serious bias in the *esw* estimate the pod locations had to be closely linked. For example, when each randomly encountered pod was followed by two further pods, each within five seconds (i.e. about 250 metres) of the previous one, the *esw* estimate was biased upward by 22%. However, such extreme clumping should be evident in the distribution of inter-sighting intervals for the leading and single sightings (see figure 3). It might result if the animals frequently form large aggregations that tend to split up into pods separated by a few hundred metres. It is therefore necessary to consider whether such behaviour is ever observed in the target species and if it is, to look out for such effects in the data by inspecting the distribution of inter-sighting intervals.

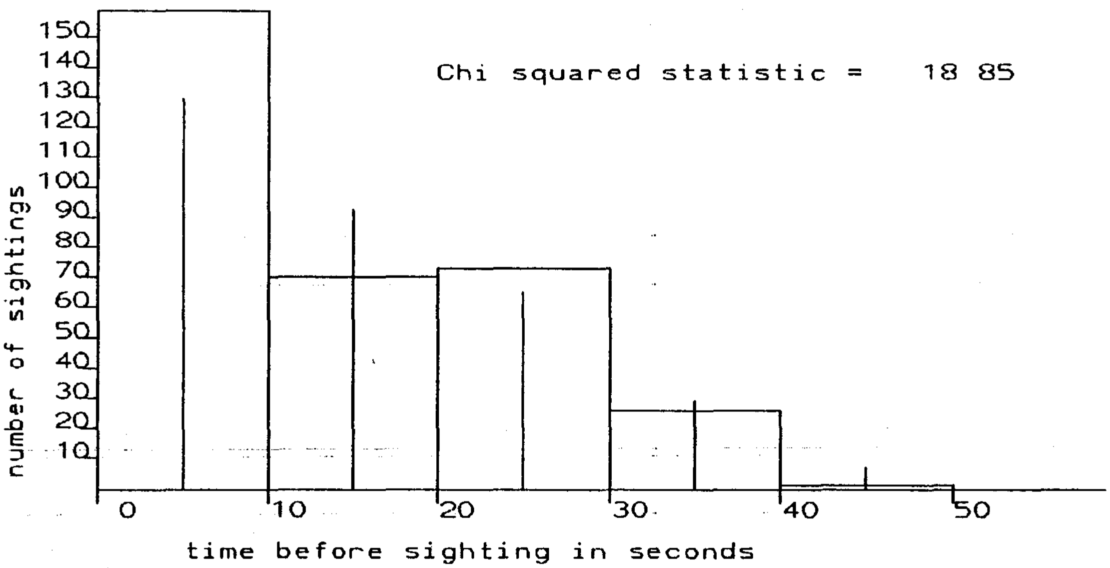


Figure 3. The frequency distribution of inter-sighting intervals from the simulation including linked sightings. The interval preceding the first sighting on each duplicate section is not included. Assuming pod sightings occur as a Poisson process, the interval from the first to any additional sighting on the section is uniformly distributed over the range zero to the remaining duration of the section. Given the number of additional sightings on each section the expected distribution of preceding intervals can therefore be calculated and is given as the vertical line at the centre of each histogram bin.

Finally, to check how effective the modified likelihood in equation (8) is in allowing for areas of zero density, the pod density was set to zero on 70% of the simulated survey legs. The total survey length was then split into equal sectors such that, if pods were randomly distributed throughout, 30% of sectors would be expected to be empty. s_0 and s_1 were then enumerated from the simulated sighting positions. Using the unmodified expression for the density of intervals preceding leading and single sightings (equation (7)) the *esw* estimate was biased upward by 10%. The bias was reduced to 3% when the expression in equation (8) was used instead.

4 SUMMARY AND CONCLUSIONS

Porpoise pods are particularly difficult targets for aerial survey. Surfacings are of short duration and a large proportion of animals are seen only beneath the surface as the aircraft passes overhead. The depth to which animals can be detected will vary with the turbidity of the water and the proportion of time spent near the surface may also depend on factors that can not be assessed from the air. It is also difficult to define a cue that can be counted from the air and used to provide a correction factor based on behavioural measurements. This leaves some type of independent sighting effort as the only way to correct for the proportion of pods undetected near the trackline. Clearly a second attempt to detect a sighting must be made following a sufficiently long interval for the probabilities of the two detections to be independent. But the fact that porpoises carry no distinctive markings that can be used for individual identification means that it is impossible to be certain which sightings, following that interval, are re-sightings and which are new pods. By summing the joint probability for sighting locations over all possible pairings of leading and trailing sightings it is possible to overcome this difficulty and make full use of the only available information relevant to duplicate identification, i.e. the proximity of trailing sighting locations to leading ones. The simulation trials suggest that this technique has the potential to provide unbiased estimates of effective strip width which are reasonably robust to variations in pod density and the rate and mode of pod displacement. They further suggest that when the technique is modified to use data from one aircraft circling back to overfly certain sections of the trackline the estimator is still satisfactory with respect to bias.

However, lack of precision in the estimates of effective strip width is more likely to cause problems than bias. Even if the estimator is unbiased, a CV of, say, 0·25 means that it is quite possible for a given survey to produce an estimate of effective strip width which is just half of the true value and hence overestimate pod abundance by 100%. Depending on the number of sightings and the actual duplicate proportion, the error on the estimate may well be this large.

Increasing sample size is one way of reducing the CV of an estimate but unfortunately, in this case, gaining more sightings by concentrating effort on high-density areas generates little improvement. At higher densities the proximity between leading and trailing sightings provides less evidence that those sightings are duplicates. Thus, in the simulation which produced the results in figure 2 the mean density was low, at 0·1 pods per km^2, around 100 duplicate sightings were generated by each run and the CV of the *esw* estimates was about 0·11. This is similar to the CV for a Petersen estimator of population size based on the same number of recaptures and suggests that not knowing, for sure, which of the sighting pairs are really duplicates has caused little difficulty. However, when the mean density was increased to 0·3 pods per km^2 the CV increased to 1·39, despite the fact that same number of duplicates was generated by the simulation.

How much effort needs to be expended to obtain sufficient duplicate sightings at low densities depends on the actual effective strip width and duplicate proportion. To generate the results illustrated in figure 2 each simulation used 150 30-minute flights, 75 hours of survey effort in all. The effective search width was 0.23 km and about one third of leading sightings were detected on the trailing flight. In that simulation pods were near the surface and hence "available" for sighting for half their dive-cycle. Had the near-surface phase been reduced to a third both the *esw* and duplicate proportion would have been reduced proportionally and the number of duplicates reduced in the ratio nine to four. Thus more than double the number of survey hours would have been required to generate to same number of duplicate sightings and achieve the same level of precision in the estimate.

It is difficult to think of ways to substantially reduce the effort required to achieve a low level of error on the *esw* estimate. If the mean density is low it should be possible to reduce the duration of the imposed dead-time to a minimum without spending an excessive amount of time circling back. Ensuring that the trailing sections are flown accurately along the original trackline is important and the use of differential GPS navigation would be worthwhile if available (cross-track errors of up to 50 metres were used in the simulations). If there were no cross-track error and no pod movement the duplicate proportion would be maximised, for a given *esw*, by a wide-shouldered detection function. Given that cross-track errors and pod movements are bound to occur it is probably not worth trying to direct searching effort to produce a desired shape of detection function, apart from pointing out that occasional sightings far from the trackline contribute little to the *esw* estimate. Reducing the interval between leading and trailing aircraft would increase the duplicate proportion, however the estimator would then not be robust to the model chosen for the alternation of dive and near-surface phases. There is, in any case, a limit to the rate at which the aircraft can be turned to fly over a selected section of trackline a second time.

It is conceivable that, in addition to location, other measures recorded for each sighting might be included in the likelihood function to help distinguish the duplicates, especially in high-density areas. Two possibilities are pod size and swimming direction. This would require modelling the distribution of pod size or swimming direction records from successive sightings of the same pod. If such models could be derived in the light of data available from, for example, TDR records, the extended likelihood could again be summed over all possible pairings between leading and trailing aircraft.

One possible source of bias for the one-aircraft technique, which was not addressed by the simulation trials, is an increased level of search effort during the re-sighting periods. The maximum potential bias, in the estimated risk of missing pods near the trackline, equals the proportion of that risk which is due to lack of concentration by the observers. The only way to assess whether that proportion is significant and, if it is, how much the level of concentration is increased during the re-sighting effort, is to use independent observers on the same aircraft. Given that effort periods during aerial survey are relatively short and observers can rest at the waypoints and during periods of unacceptable sighting conditions, the size of any bias from this source is unlikely to be large. Variations in sighting conditions along the trackline could also lead to dependence in sighting probabilities for the leading and trailing aircraft. However, as the calculated probability for each sighting is conditional on the sighting conditions recorded for that part of the survey (including any difference between left and right of the trackline) that source of bias has been eliminated.

It is worth considering whether, given the likely error on the available *esw* estimate, it is still worth using the method to provide an estimate of absolute abundance rather than accepting a relative abundance estimate. The answer depends on the main objective of the survey and also the level of pod density over the survey area. If consistency in observer effort can be ensured an index of relative abundance will have a lower CV, and thus be better able to monitor changes in abundance, than an estimate of absolute abundance based on this method. But it may be better, in that case, to simply use sighting rate to provide the index instead of incorporating distance data at all. Fitting a sighting function to the histogram of perpendicular sighting distances might compensate for variation in sighting conditions but such variations are just as likely to effect the risk of missing pods near the trackline. Sighting conditions could simply be incorporated as a qualitative factor in statistical comparison of sighting rates over time or between areas. Furthermore, any variation in the search pattern used by observers might have as much effect on the risk of missing pods near the trackline as on the perpendicular distribution. It may therefore be better to assume that their effective strip width remains constant rather than assuming that the risk of missing pods near the trackline remains constant.

Unfortunately the potential for sighting surveys to provide useful indices is usually low, especially in low density areas, and the main objective of the survey may then be to provide a lower or upper bound on abundance, for example to assess the potential effect of incidental catches. In that case an unbiased estimate of absolute abundance, even one with a high CV, may be sufficient to allow a management decision to be reached. When a lower bound is sought it may still be preferable to abandon an *esw* estimate based on duplicate sightings. For example, behavioural data might be used to place an upper bound on the proportion of time that pods are available and hence an upper confidence limit on effective strip width which is lower than that derived from the pattern of leading and trailing sightings. It is unlikely the same logic would apply to an upper bound on abundance. It is difficult to see how a lower bound on the risk of missing pods could be set, given that that risk depends not only on the behaviour of the animals but also on that of the observers and the effect of sighting conditions. In that case the pattern of leading and trailing sightings provides the only way to calculate a lower confidence limit on effective strip width, and hence an upper confidence limit on abundance.

REFERENCES

Barlow, J., C.W.Oliver, T.D.Jackson, & B.L.Taylor 1988. Harbour porpoise (*Phocoena phocoena)* abundance estimation for California, Oregon and Washington: II aerial surveys. *Fish. Bull. U.S.* 86: 433-444.

Borchers, D.L., S.T.Buckland, E.D.Clarke & S.L.Cumberworth 1995. Estimation of cetacean abundance from the SCANS shipboard surveys in summer 1994. In: P.S.Hammond (ed.) Distribution and abundance of the harbour porpoise and other small cetaceans in the North Sea and adjacent waters. *Final Report to the European Commission under project* LIFE92-2/UK/027

Cox, D.R. 1962. *Renewal Theory* London: Methuen.

Heide-Jorgensen, M.P., A.Mosbech, J.Teilmann, H.Benke & W.Schultz 1992. Harbour porpoise (*Phocoena phocoena*) densities obtained from aerial surveys north of Fyn and in the Bay of Kiel. *Ophelia* 35(2): 133-146.

Hiby A.R. & P.Lovell 1995. Estimation of harbour porpoise abundance from the SCANS aerial survey. In: Hammond P.S. (ed.) Distribution and abundance of the harbour porpoise and other small cetaceans in the North Sea and adjacent waters. *Final Report to the European Commission under project* LIFE92-2/UK/027

Hiby, A.R. & P.Lovell 1998. Using aircraft in tandem formation to estimate abundance of harbour porpoise. *Biometrics* (in press).

Westgate, A.J., A.J.Read, P.Berggren, H.N.Koopman & D.Gaskin 1995. Diving Behaviour of Harbour Porpoises, (*Phocoena phocoena*). *Canadian Journal of Fisheries and Aquatic Sciences* 52 (5): 1064-1073.