University of Cape Town

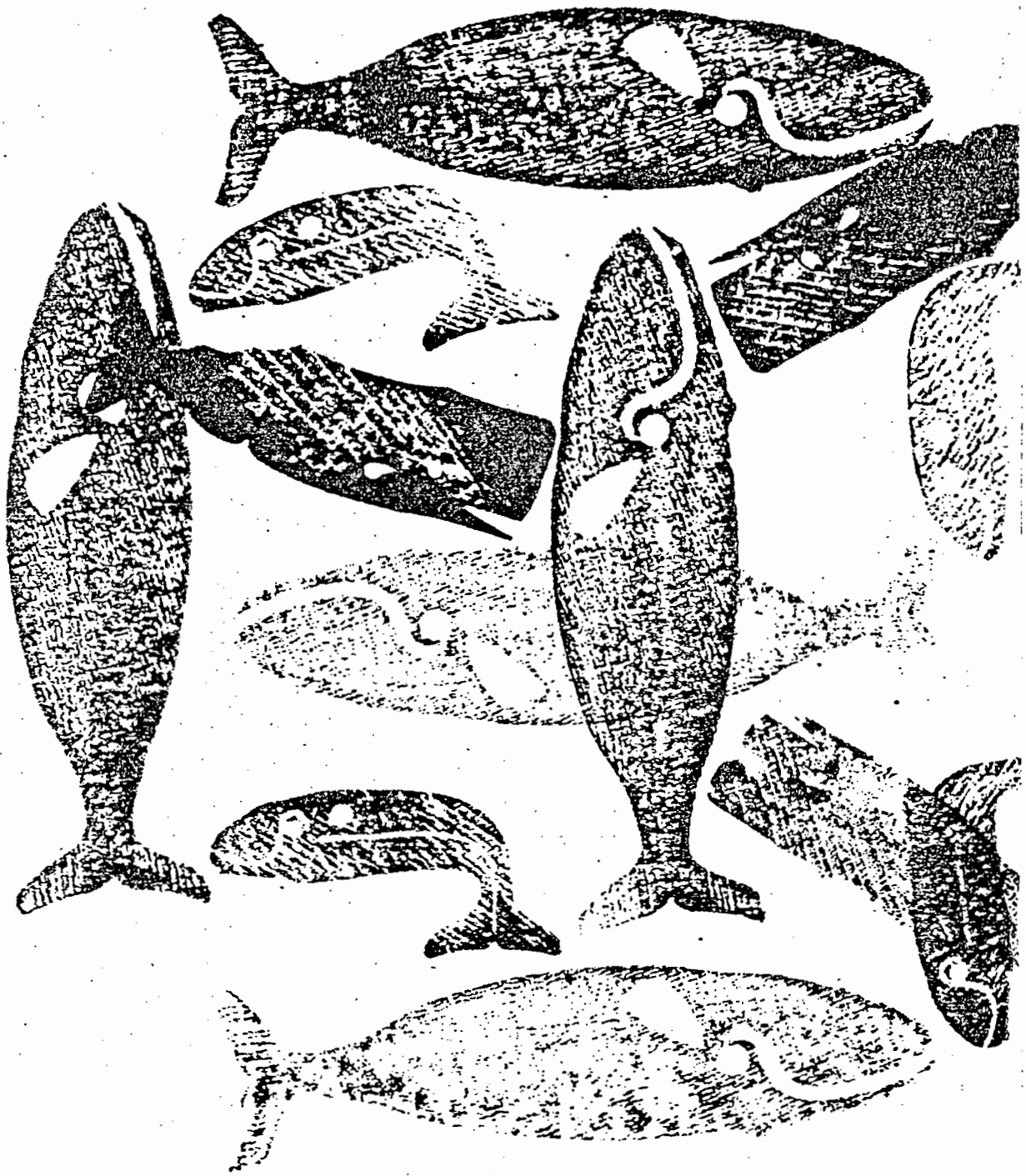
Department of Mathematical Statistics

A SIMULATION STUDY OF THE EFFECT OF TARGET MOTION ON SIGHTING ESTIMATES OF MINKE WHALE POPULATION DENSITY

by

M BASSON

A thesis prepared under the supervision of Dr D.S. Butterworth and Dr L.G. Underhill in partial fulfilment of the degree of Master of Science in Operations Research.



Page from a whaling ship's journal

From: Songs the Whalemen sang by Gale Huntington

Barre Publishers, Massachusetts, 1964

To my parents

and JP

ACKNOWLEDGEMENTS

I wish to express my sincere thanks to my supervisors Dr D.S. Butterworth and Dr L.G. Underhill: Dr Butterworth for his enthusiasm and for always finding time to help and encourage; and Dr Underhill for his help with the statistical analysis and especially for his support in the final stages. My thanks also to the CSIR for their financial support. Finally, I am indebted to UNIVAC 1100 without 'whom' this study would not have been possible.

ABSTRACT

Line transect methods used to estimate population density assume stationarity of targets. Violation of this assumption leads to overestimation of the true density. A simulation study based on a hazard-rate model is used to assess the resulting bias. The model is calibrated to generate sighting data resembling real data from minke whale sighting surveys. The procedure currently used to calculate a corrected negative exponential density estimate from sighting data is duplicated using simulated data. The resulting estimates are compared to the true population density determined by the simulation.

Results reveal that in the case considered, the method of calculating the g(0) factor (which corrects for the fact that all animals on the trackline are not sighted) leads to a greater degree of overestimation than the effect of target motion at 3 knots. Shortcomings of the model are pointed out and possible improvements suggested. It is also suggested that further research be focused initially on the calculation of the g(0) correction factor rather than on effects of target motion.

Table of contents

Table of contents

Table of contents

1. INTRODUCTION

In ecological studies it is often necessary to determine the abundance of a biological population. The most direct way to do this is to count all individuals in a given known area. An estimate of population density can then be obtained simply by dividing the number counted by the size of the area sampled. This approach can obviously be very time consuming and is often impractical if not impossible. An alternative approach, considered here, is that of line transect methods.

Assume that the density of some population has to be estimated in an area of known boundaries and size A. An observer moves along the transect, which is a straight line placed within the study area, and records all sightings of individuals of the population. Apart from recording the number of observed individuals, certain quantitative measurements must also be taken at the moment of detection. Usually one or more of the following three measurements are recorded (figure 3-9, section 3.6):

y = the perpendicular distance from transect line to object

r = the radial or direct distance from observer to object

θ = the sighting angle, i.e. the angle between the line of travel and the radial r

Initially, line transects were often thought of as very long narrow quadrats (Forbes and Gross 1921). The procedure of estimating animal abundance based on the recorded distances seems to go back only to the 1930s. Gates (1979) credited R.T. King as being the first to recognise that not all animals are seen in the

sampled strip. King's idea was to try and estimate the effective half width of the transect, in order to adjust for missed animals. Instead of the original density:

$$D = n/A$$

where n = the number of individuals counted

A = the total (known) area,

the estimated density would be:

$$ED = n/EA$$

where ED = the estimated density

EA = the estimated area effectively sampled.

Since the length, say L, of the transect line is known, one can write the effective area EA as:

$$EA = 2L.EW$$

where EW = the estimated half width or strip width.

(Note that E is used here to indicate estimators, instead of the conventional ^). King used the average sighting or radial distance as the estimate of W.

The effective strip width W can be interpreted or defined in many ways. Gates (1979) defines it as follows: W is that distance from the trackline such that the number of unseen individuals located closer to the line than W, equals the number of individuals seen at distances greater than W.

Another basic approach to the estimation of density from line transects was introduced by Kelker (1945). He suggested determining a strip width s about the transect line of length L,

within which all animals are seen. One then defines n' as the number of animals seen in the area 2Ls and an estimator of density is then:

$$ED = n'/2Ls$$

Hayne (1949) was probably the first to create a mathematical foundation for line transects. He also provided the first estimator that has a rigorous justification in statistical theory. Until 1968 almost no significant theoretical advances appeared in the literature although some evaluations of the line transect method were presented (for example Robinette et al. 1956) and the method was used frequently on a variety of species (for example dead deer- Robinette et al. 1954; seals- Eklund and Atwood 1962).

The first truly rigorous statistical development of a line transect estimator was presented by Gates et al. (1968). The estimator is applicable only to untruncated, ungrouped perpendicular distance data.

At this stage it is necessary to define the detection curve g(y) as the probability of an animal being seen, given that it is a perpendicular distance y from the line transect path. The function g(y) is such that g(0)=1, i.e. all animals on the trackline are seen. Gates et al. (1968) assumed g(y) to have a negative exponential form, $g(y)=\exp(-by)$, where b is an unknown parameter to be estimated. They developed the optimal estimator of b (based on the perpendicular distances) and constructed the estimator of population size.

A second important paper that appeared in 1968 was that of Eberhardt. He suggested the less restrictive and more appropriate approach of adopting a family of curves to model g(y). The two

families he suggested were a power series and a modified logistic family. However, no estimators based on the logistic model were developed.

Between 1968 and 1976 rigorous models and estimators were developed mainly for untruncated ungrouped perpendicular data and the approach was predominantly parametric. Anderson and Pospahala (1970) introduced some basic ideas that underlie a nonparametric approach to the analysis of line transect data. Seber (1973:pp 28-30) presented a general model structure based on perpendicular distances, but equations were left at a conceptual stage.

Burnham and Anderson (1976) attempted the general formulation of line transect sampling and showed a basis for the general construction of estimators. Their formulation is suitable for developing parametric or nonparametric estimators.

Work on parametric and nonparametric procedures has also been done by, amongst others, Pollock (1978), Crain et al. (1978), and Quinn and Galucci (1980).

An alternative model proposed independently by Schweder (1977) and Hayes and Buckland (in press) is the so-called hazard-rate model. If a function $k(r,y)$ is specified, the corresponding expression for $g(y)$ can be found. The meaning of the function $k(r,y)$ is the following: if an animal at right angle distance y and radial distance r has not yet been sighted, then $k(r,y)dx$ represents the incremental probability of sighting the animal as the observer advances a distance dx along the transect line. Burnham and Anderson (1976) recognised the main problem of analysis as the modeling of $f(y)$, the probability density function of the perpendicular distance data, and the subsequent estimation of

f(0). Note that f(y) is simply g(y) normalised. In the light of this, it is clear how estimators can be derived for the hazard rate model. Hayes and Buckland conclude that a nonparametric perpendicular distance model will generally provide more reliable estimation than any radial distance model.

This review is by no means exhaustive and various proposed models and estimators have not been mentioned. The most comprehensive reference on line transect sampling is probably the monograph by Burnham et al. (1980) and the reader is referred to this and the other references for further detail concerning line transect methods and its applications.

2. THE PROBLEM

The basic assumptions underlying line transect methods as listed by Seber (1973) are given below:

1. The animals are randomly and independently distributed over the population area.
2. The sighting of one animal is independent of the sighting of another.
3. No animal is counted more than once.
4. When animals are seen through being flushed into the open, each animal is seen at the exact position it occupied when startled.
5. The response behaviour of the population as a whole does not substantially change in the course of running the transect.
6. The individuals are homogeneous with regard to their response behaviour, regardless of sex, age etc.
7. The probability of an animal being seen, given that it is a right angle distance y from the line transect path (irrespective of which side of the path it is on), is a simple function g(y), say, such that g(0)=1 (i.e. all animals on the trackline are seen).

Any of these assumptions can be violated in a variety of ways in practical line transect sampling. Since the line transect methods were originally developed for application to terrestrial animals and plants, a number of special problems arise in the case of whale sighting surveys.

For the minke whale (Balaenoptera acutorostrata) population, for

example, assumptions 1 and 2 are believed not to hold true (Best and Butterworth 1980). This can be remedied by stratifying the searching effort and calculating density for schools rather than for individuals. The violation of assumption 4, however, causes more serious problems. Note that this assumption implies that animals are stationary, at least until detection.

Assumption 4 may be relaxed slightly to allow movement that is independent of and slow relative to the observer before detection (Burnham et al. 1980). An analysis by Skellam (1958) deals with circumstances where motion of the animals can be assumed to be relatively uninfluenced by the observer.

In the case of whale sighting surveys, the above is not applicable. Although more data on the average swimming speed of undisturbed whales are needed, indications are that the average swimming speed is less than 5 knots (Butterworth and Best 1982). Compared to the normal searching speed of 12 knots, it is clear that whale movement cannot be considered as being slow relative to the observer. The possibility of whale reaction to the vessel has been investigated (Butterworth and Best 1982). Although inconclusive, results show that there is no direct indication of ship-seeking, which could lead to abundance overestimation.

In his study of search theory, Koopman (1956) investigates the problem of moving targets. He considers an observer moving with constant velocity v and with a fixed detection radius R. Targets move with constant speed u, but random direction. An expression for the number of targets detected per unit time is derived. This shows that, under the said circumstances, an overestimate of population density results when targets are non-stationary. Unfortunately this model is very unrealistic for application to

whale sighting data, because of the particular sighting criterion used. The fixed detection range sighting criterion implies that all targets within a radius R of the observer are sighted with certainty. This is unrealistic for two main reasons. Firstly, sightability decreases with distance from the observer. Secondly, a whale is only visible when it surfaces or blows i.e. a whale can be within the detection area without being visible.

A preliminary study using a simulation model to estimate the degree of overestimation under more realistic conditions was attempted by Hiby (1982).

It is clear from the above that target movement influences the estimates of population density. The main aim of this study is therefore to attempt to assess the bias in abundance estimates resulting from random whale movement.

Three approaches to the problem of assessing this bias exist. The first approach is to develop a theoretical model and attempt to find an analytical expression for the bias. For a simple case, such as considered by Koopman, this approach is possible, but for more realistic (and therefore more complex) cases it is not always easy or possible to obtain an analytical expression, even if it is possible to develop a model. A second possibility is to make use of experimental methods. These methods are often (as in this case) expensive and impractical. The third approach is to develop a computer simulation model. This is considered to be the most suitable approach in this specific case.

There are a number of advantages in using a simulation model rather than a theoretical model. A simulation model is usually more tangible than a complicated mathematical model. Once the

relevant variables and elements of the system to be simulated have been identified, it is relatively easy to ensure that all necessary assumptions are satisfied. It is then easy to violate one assumption (e.g. in this case the stationarity assumption, 4) and analyse the effect it has on the system. It is also possible to investigate the system's sensitivity to changes in given parameters, something that might not be easy in a mathematical model.

A simulation model can hardly duplicate a physical system and one usually has to be satisfied with a simplification of such a system. The following additional assumptions (i.e. simplifications of reality) were made:

1. Whale paths are straight lines.
2. Whales are always potentially sightable.
3. Whales do not aggregate.

For a first estimate of bias due to random movement, straight line paths were considered to be acceptable. A sophistication of the model would be to consider random walk paths. However, since not much is known about the swimming patterns of minke whales, this might not necessarily mean a more accurate representation of reality.

In real sighting experiments, whales are sightable when they break or disturb the surface of the sea or blow. Clearly, a blow would be visible much further away from the vessel than would disturbance of the surface of the sea. Data on dive times and blow rates would be needed to develop a more realistic model.

If density is calculated for schools rather than for individuals, in the case of real sighting data, simplification 3 poses no

problems and an individual whale in the simulation model can effectively represent a pod.

The simulation model is discussed in more detail in Chapter 4.

3. MATHEMATICAL ANALYSIS

In this section the mathematical basis of the simulation model is developed and discussed.

It is important to note that there are two frames of reference when considering the sea, vessel or platform and target (in this case a whale). In the sea frame, absolute or true speeds and directions are considered. In the platform frame however, speeds and directions relative to the platform (i.e. as if the platform were at rest) are considered.

I shall subsequently refer to the sea-frame as the s-frame and the platform-frame as the p-frame.

3.1. The choice of axes.

In the s-frame a set of perpendicular x-y axes is chosen. The sighting platform moves in the minus x direction in this frame. A set of x-y axes is also chosen in the p-frame. The platform is fixed at the position with x-y coordinates (R,0.0) and this set of axes moves along with the platform. See figure 3-1. below.

The detection area (the area within which a whale can possibly be sighted, but outside of which no sighting can occur) also moves with the platform and is positioned in such a way that the p-frame y-axis is tangent to it. In the simplest case considered, the detection area is a circle with radius R and the platform situated at its centre. See figure 3-2. below.

Figure 3-1. Frames of reference

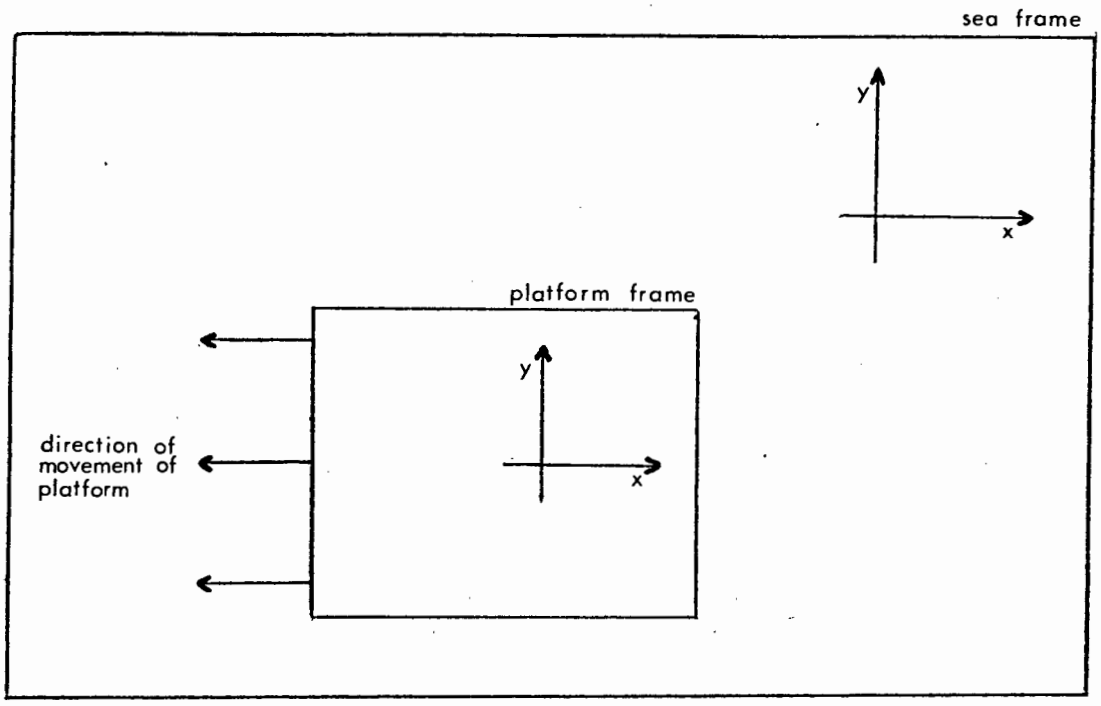
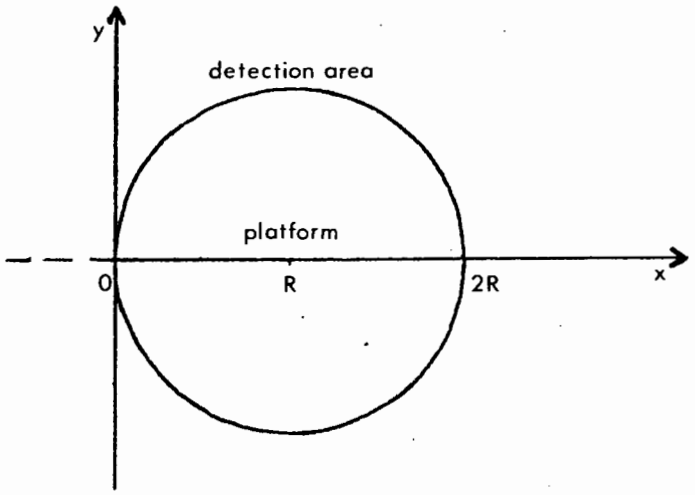


Figure 3-2. Detection area



3.2. The Limiting Angle.

Let the platform move (in the -x direction) at true speed v knots and a whale in any direction at true speed u knots. Note that u and v are positive values. The following possibilities exist:

(a) $u < v$ (of which u=0 is a special case)

(b) $u = v$

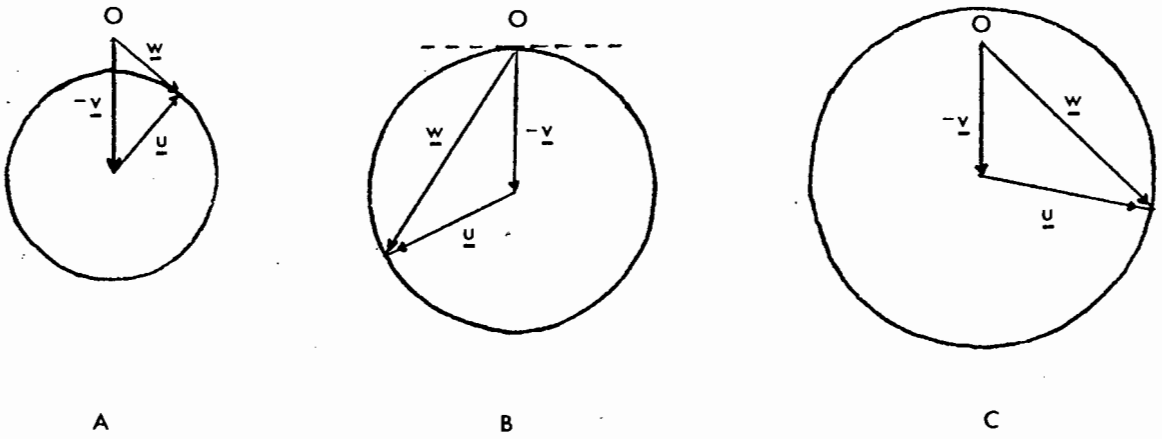
(c) $u > v$

Let $\underline{u}$ and $\underline{v}$ be the vectors of the respective whale and platform speeds and let $\underline{w}$ be the vector of whale speed relative to the platform. From the theory of relative motion, it is known that

$$\underline{w} = \underline{u} - \underline{v} \qquad (3\text{-}1)$$

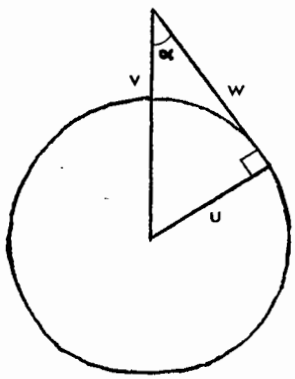
It is convenient to draw the circular diagrams A,B and C of figure 3-3. corresponding to the cases a,b,c above .

Figure 3-3. Limiting Angle



In each case the distance from O (the origin of $\underline{w}$) to the centre of the circle is v and the radius of the circle is u. The resultant vector $\underline{w}$ has its origin at O and extremity on the circle. In cases (a) and (b), the angle formed between the vectors $-\underline{v}$ and $\underline{w}$ reaches its maximum to either side when $\underline{w}$ is tangent to the circle. Let this maximum angle be α, the so-called limiting angle.

Figure 3-4. Limiting angle, case A



Since w is tangent to the circle, $\sin(\alpha)=u/v$ and therefore

$$\alpha = \text{asin}\ (u/v) \qquad (3\text{-}2)$$

Note that when u=0, $\underline{w}=-\underline{v}$ and the limiting angle is 0. Furthermore, when u=v (i.e. case b), asin(u/v)=asin(1) and α is 90 degrees to either side of $-\underline{v}$.

In case (c) where u>v, the origin of $\underline{w}$, namely O, lies within the circle and the limiting angle is clearly 180 deg. to either side of $-\underline{v}$. Note the implication in this case that the target can overtake the platform or enter the detection area from behind the platform.

3.3. Determining Wmax.

The sighting platform usually maintains a cruising speed of 12.0 knots and it has been estimated that minke whales swim at an average speed of less than 5.0 knots when undisturbed (Butterworth and Best 1982). It is therefore realistic to assume that $u<v$ and only this case, i.e. case (a), will be considered in this simulation model. Under this assumption whales cannot enter the detection area from behind the platform and it is sufficient to generate the whales along the y-axis perpendicular to the trackline, ahead of the platform (see figure 3-2.). This implies that all whale paths start at some point on the y-axis.

An expression for the length of y along which whales have to be generated so as not to miss any possible sightings, has to be found.

Let umax be the maximum whale speed that will be considered, i.e. $u < umax < v$. Let α_{max} be the corresponding limiting angle and therefore, from eq. (3-2):

$$\alpha_{max} = \text{asin}\ (umax/v)$$

Since asin is monotonically increasing for the relevant values of u and v, α will be less than or equal to α_{max} and α_{max} can be considered as the limiting case.

A whale starting at any point P (fig.3-5.) along the y-axis can have at its maximum, a direction of movement at an angle α_{max} to either side of the x direction. The whale path can only intersect the detection area if it starts at a point P', below point P, where P' is positioned such that P'Q' is tangent to the detection area and at an angle α_{max} to the x direction.

Figure 3-5. Generating line

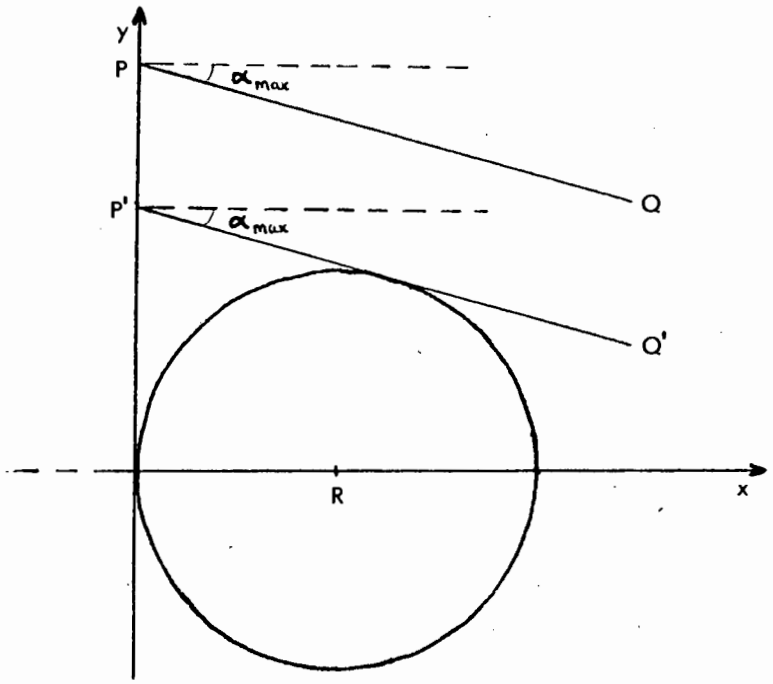
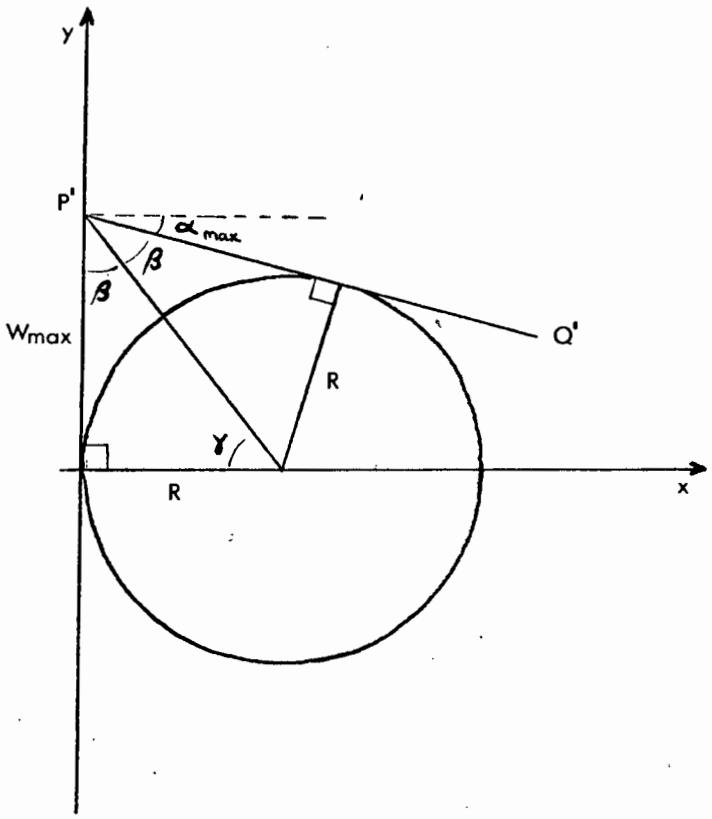


Figure 3-6. Determining Wmax



Let the distance from O to P' be Wmax. Note that because of the symmetry of the detection area, there will be an equal distance to the other side of the track.

From figure 3-6. it is clear that:

$$W_{max} / R = \tan \gamma$$

and

$$\gamma = (\pi/2 + \alpha_{max})/2$$

therefore

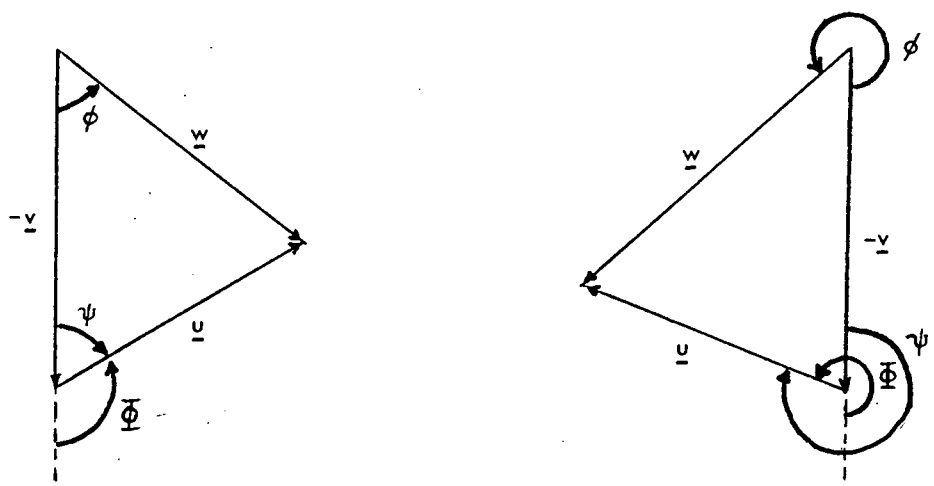
$$W_{max} = R.\tan\left(\frac{\pi/2 + \alpha_{max}}{2}\right) \qquad (3\text{-}3)$$

In the simulation itself umax=10.5 knots was usually used.

3.4. The distribution of the angle

Let the vectors $\underline{u}, \underline{v}$ and $\underline{w}$ be as before. Let ϕ be the angle between $-\underline{v}$ and $\underline{w}$ (figure 3-7.). Let ψ be the angle between $-\underline{v}$ and $\underline{u}$ and Φ the angle between the extended $-\underline{v}$ and $\underline{u}$. This angle indicates the direction of movement of a whale in the s-frame.

Figure 3-7. Notation



It is known that: $\underline{w} = \underline{u} - \underline{v}$

Splitting each vector into its x- and y-components, the following is obtained:

$$\underline{u} = u.\cos\Phi.\underline{i} + u.\sin\Phi.\underline{j}$$

$$\underline{v} = -v.\underline{i} \quad \text{and} \quad -\underline{v} = v.\underline{i}$$

now $$\underline{w} = (u.\cos\Phi + v)\underline{i} + u.\sin\Phi.\underline{j} \qquad (3\text{-}4)$$

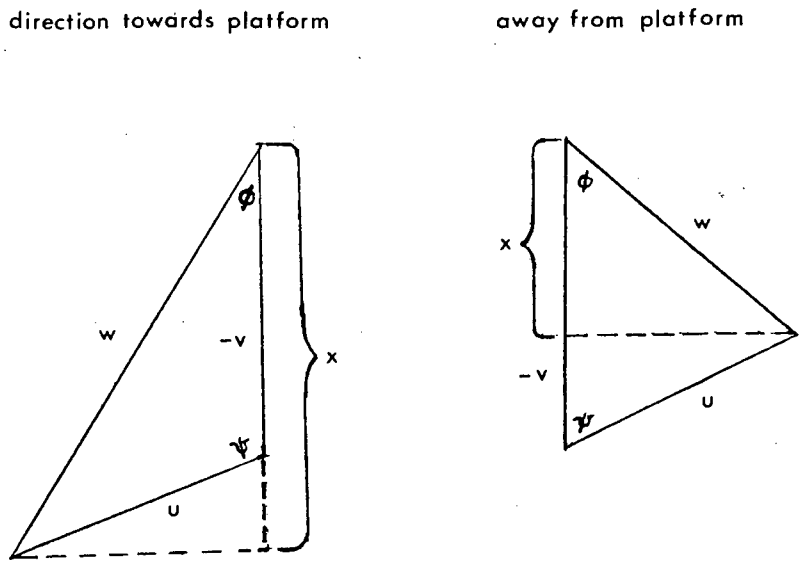
Clearly $\underline{w} = v.\underline{i} = -\underline{v}$ when $u = 0$.

Since Φ plays no role in the case where u=0, Φ can take on any value. In terms of the simulation it is therefore sensible to draw a value for Φ using a random number generator, irrespective of the value of u.

The case where u>0 is now considered. Assuming that the directions of movement of the whales are randomly distributed, one would expect a stationary sighting platform to encounter equal proportions of whales moving away from it and whales moving towards it. As soon as the platform starts moving in a fixed direction, it tends to encounter a larger proportion of whales moving in directions towards it than the proportion moving away from it. This shift in proportions is related to the platform and whale speeds and relative directions.

The relative shift in proportions moving towards and away from the platform can be given by the ratio of the x component of $\underline{w}$ to the x component of $-\underline{v}$, i.e. simply v, as the rate at which whales may be encountered is proportional to the component of the whale velocity relative to the vessel, in the platform- or vessel-frame.

Figure 3-8. Directions of movement



Let x (as in fig. 3-8) denote the x component of $\underline{w}$. The desired ratio is then:

$$x/v$$

Since $x/w=\cos\phi$ the ratio can also be written as:

$$(w/v)\cos\phi$$

The probability density function of Φ, $f(\Phi)$, now becomes:

$$f(\Phi) = \frac{1}{2\pi} \cdot \frac{w}{v} \cos\phi$$

or

$$f(\Phi) = \frac{1}{2\pi}\left[1 - \frac{u}{v}\cos\Phi\right] \qquad (3\text{-}5)$$

The $1/2\pi$ factor is simply a normalising factor so that $f(\Phi)$ is a probability density function.

In order to simulate this distribution of Φ, the following procedure has to be followed:

(a) find $F(\Phi)$, the distribution function of Φ

(b) set $F(\Phi) = q$ where q is a random number

(c) solve for $\Phi = F^{-1}(q)$ where F^{-1} is the inverse of F

(a)
$$F(\Phi) = \int_0 f(\Phi)\, d\Phi$$
$$= 1/2\pi \int_0 (1 - \tfrac{u}{v}\cos\Phi)\, d\Phi$$
$$F(\Phi) = 1/2\pi \left(\Phi - \tfrac{u}{v}\sin\Phi\right)$$

(b) A random number q between 0.0 and 1.0 is drawn. Note that max $F(\Phi) = F(2\pi) = 1.0$ and min $F(\Phi) = F(0) = 0.0$

(c) The solution i.e. the desired angle Φ is given by the inverse function of $F(\Phi)$, namely $F^{-1}(q)$. Since F^{-1} cannot be found analytically, a numerical method must be employed to solve for Φ in $q = 1/2\pi \left(\Phi - \tfrac{u}{v}\sin\Phi\right)$

The Newton Raphson method is employed in the following way:

Let Φ_i indicate the value of Φ at the i-th iteration.

Set $\Phi_0 = q/2\pi$ and solve recursively, using equation (3-6).

$$\Phi_{i+1} = \Phi_i - \frac{g(\Phi_i)}{g'(\Phi_i)} \qquad (3\text{-}6)$$

where $g(\Phi_i) = \Phi_i - \tfrac{u}{v}\sin\Phi_i - 2\pi q$

$$g'(\Phi_i) = 1 - \tfrac{u}{v}\cos\Phi_i$$

Iteration is terminated when a desired level of accuracy is obtained i.e. when $|\Phi_{i+1} - \Phi_i| < e$

where e is the desired level of accuracy, e=0.00001

3.5. Moving a whale

Let the vectors $\underline{u}$, $\underline{v}$ and $\underline{w}$ be as before. Equation (3-4) gives the x and y components of $\underline{w}$. Let wx and wy indicate the respective components, then

$$wx = u.\cos\Phi + v$$
$$wy = u.\sin\Phi$$

Assume the whale is at a point (x,y) in the p-frame. During a time dt it will move a distance wx.dt in the x direction and a distance wy.dt in the y direction, thus being at a new point (x+wx.dt,y+wy.dt). This can however be written as (x+dx,y+dy) where:

$$dx = (u.\cos\Phi + v)dt$$
$$dy = (u.\sin\Phi)dt \qquad (3\text{-}7)$$

The process described above is an attempt to approximate the differential

$$\frac{\partial B}{\partial x} = FB \qquad (3\text{-}8)$$

by the difference equation:

$$B(x,y) = B(x+\delta x, y)[1 - F\delta x] \qquad (3\text{-}9)$$

where B(x,y) is the probability that the whale has not been seen by the time it reaches the point (x,y) and $F\delta x$ is the probability of a whale (moving at a right-angle distance y from the vessel) being seen between the points (x+dx,y) and (x,y).

Let us assume F to be constant for the moment. Now the solution to (3-8) is:

$$B = B_0 e^{-Fx}$$

and by series expansion of the exponential:

$$B \approx B_o \left(1 - Fx + (Fx)^2/2! - \cdots + \cdots\right)$$

B can be approximated by $B_o(1 - Fx)$ only if $Fx \ll 1$ and terms of the order of $(Fx)^2$ can be neglected.

Similarly, for the case where F=F(x,y) where the solution to eq. (3-8) is:

$$B = B_o e^{-\int_x^\infty F(x,y)\,dx}$$

the approximation (eq.3-9) is reasonable only if $F\delta x \ll 1$ and terms of the order of $(F\delta x)^2$ can be neglected.

In the case of the negative exponential sighting rate (section 3.6.), the maximum value F(x,y) can take on can be obtained as follows:

$$F(x,y) = f(r).h(\theta).1/v$$
$$\max f(r) = f(0) = \mu$$
$$\max h(\theta) = h(0) = 1$$

i.e. $\max F(x,y) = \mu/v$

Therefore we require $(\mu/v)\delta x \ll 1$

and since $\delta x/v = \delta t$ the choice of δt must be such that $\mu\delta t \ll 1$

In the simulation δt was usually set to 0.01. Note that δt could preferably be smaller still, but CPU-time increased to such an extent that it was considered acceptable to proceed with δt =0.01.

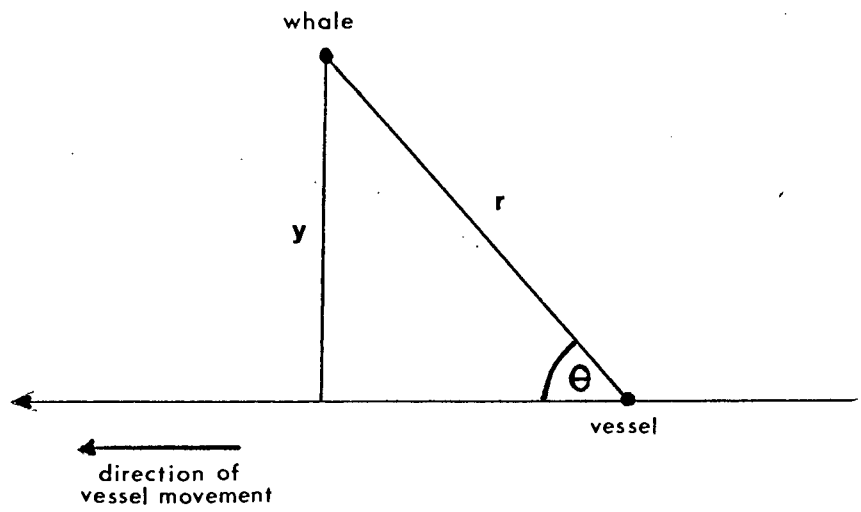
3.6. Sighting Criteria

Let (r, θ) indicate a whale's position, r being the radial distance between animal and observer and θ the angle between the trackline and the radial (figure 3-9). Assume that the probability per unit time, p, of sighting a whale at a position (r, θ) is given by:

$$p = f(r).h(\theta) \qquad (3\text{-}10)$$

The $h(\theta)$ factor describes the relative concentration of sighting effort at different directions from the track and the f(r) factor describes the relative sightability at different radial distances from the vessel.

Figure 3-9. Notation: sighting data



3.6.1. The fixed detection range sighting criterion.

The simplest case considered in this study is the fixed detection range sighting criterion (FDR). This criterion assumes that within the given detection area, all whales are detected. The detection area is in this case a circle with radius R (figure 3-2).

Clearly this sighting criterion is not very realistic. It implies equal sighting effort in all directions. It also implies that sightability remains constant with increasing distance from the vessel up to a distance R. At distances greater than R there is however no chance of any sighting. A more realistic f(r) would for example be of a negative exponential form.

The advantage of this sighting criterion is however, that it can be handled analytically because of its simplicity.

Koopman (1956) condsiders an observer moving with fixed velocity v and targets moving in random directions with fixed speed u. Assuming the number of targets per unit area to be N and conditions as described above i.e. the FDR sighting criterion and detection circle of radius R. He then shows (Appendix I) that the number of targets detected per unit time Nt is given by:

$$N_t = \frac{4RN}{\pi}(u+v)\,E(\sigma)$$

where

$$E(\sigma) = \int_0^{\pi/2} \sqrt{1 - \sin^2\sigma . \sin^2\psi}\; d\psi$$

$$\sin\sigma = 2\sqrt{uv}/(u+v)$$

When targets are stationary, i.e. u=0, the expression simplifies to:

$$N_t(u=0) = 2RNv$$

The ratio Nt/Nt(u=0), referred to as the Koopman ratio in this study, is thus given by:

$$KR = \frac{N_t}{N_t(u=0)} = \frac{2}{\pi} \cdot \frac{(u+v)}{v} E(\sigma) \qquad (3\text{-}11)$$

This ratio is an increasing function of target speed u and reflects the fact that target motion will increase the contact rate on random targets. Skellam (1958) also derives this result. The Koopman ratio can be used in the validation process of the simulation model as discussed in section 5.1.

Implementing the FDR sighting criterion simply means determining, at every position along a whale's path, whether r is less than or equal to R i.e. whether the whale is in the detection area. If so, the whale is sighted and the relevant information recorded. Note that this implies that a whale will be sighted at the first position it occupies within the detection area.

3.6.2. The negative exponential sighting criterion.

The second case to be considered has the advantage that a functional form of g(y), the probability distribution of the right angle distances, exists for the case where whale velocity is zero.

Two forms of the negative exponential are considered for the sighting rate f(r) (figure 3-10):

$$(i) \quad f(r) = \mu e^{-\lambda r} \qquad (3\text{-}12)$$

$$(ii) \quad f(r) = \frac{\mu}{r} e^{-\lambda r}$$

where μ and λ are positive parameters. The $\exp(-\lambda r)$ factor

would account for decreasing sightability with distance. At larger distances, the 1/r factor (in case (ii)) would allow for decreasing apparent size of the object. It is clear, however, that this factor diverges as r tends to zero and a realististic and less problematic approach might be to combine cases (i) and (ii). This can be done as follows for example:

$$f(r) = \begin{cases} \mu e^{-\lambda r} & 0 \le r \le R \;;\; R>0 \\ \frac{\mu}{r} e^{-\lambda r} & r > R \end{cases} \qquad (3\text{-}13)$$

For simplicity $h(\theta)$ is taken as:

$$h(\theta) = \begin{cases} 1 & 0 \le \theta \le \pi/2 \\ 0 & \pi/2 < \theta < \pi \end{cases} \qquad (3\text{-}14)$$

i.e. only the forward directions are searched.

Butterworth (1982) derives the corresponding functional form of g(y) for (i) and (ii), using the above form of $h(\theta)$.

For case (i) $$g(y) = 1 - e^{-\alpha y K_1(\lambda y)} \qquad (\alpha = \tfrac{\mu}{v}) \qquad (3\text{-}15)$$

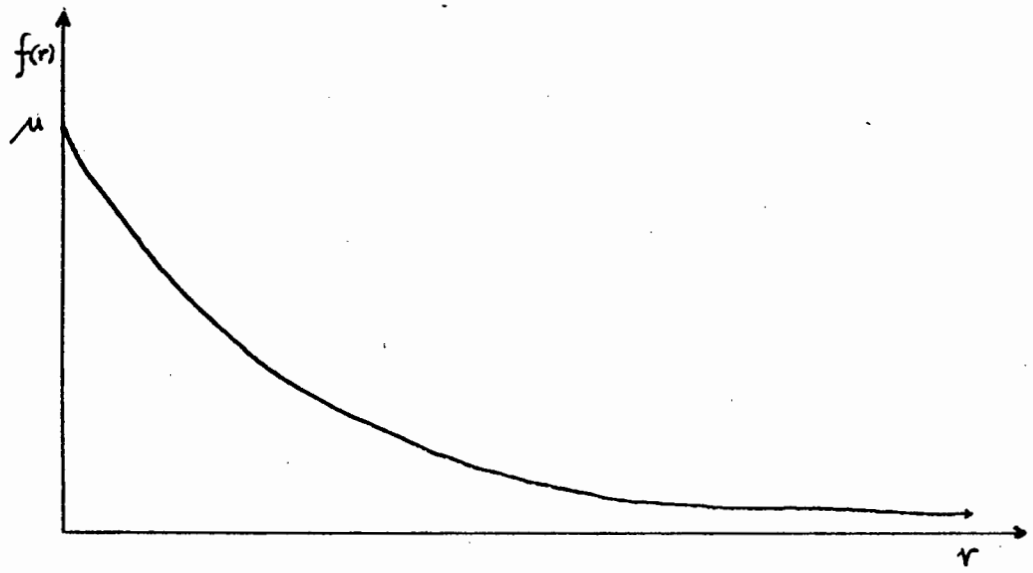
and case (ii) $$g(y) = 1 - e^{-\alpha K_0(\lambda y)} \qquad (\alpha = \tfrac{\mu}{v}) \qquad (3\text{-}16)$$

where K_0 is a zero order and K_1 a first order modified Bessel function.

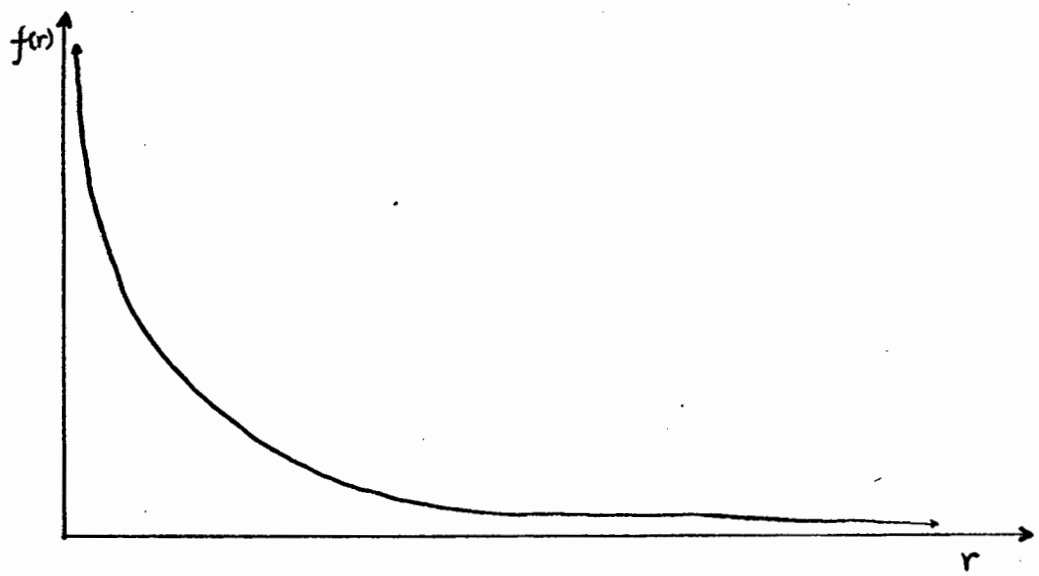
Implementing both cases (i) and (ii), implies the following in terms of the simulation model. Assuming that the whale is at point (r, θ), $f(r)h(\theta)dt$ is calculated, where dt is the time increment. A random number q between 0.0 and 1.0 is drawn. If $f(r)h(\theta)dt$ is less than q, the whale is NOT sighted. However, if $f(r)h(\theta)dt$ is greater than q, the animal is sighted and the relevant information recorded.

The problem that arises in case (ii) as r tends to zero, is overcome by setting r = e, where e=0.0000001, whenever r=0 in the simulation.

Figure 3-10. Two functional forms of f(r)



(i) $f(r) = \mu e^{-\lambda r}$



(ii) $f(r) = \frac{\mu}{r} e^{-\lambda r}$

4. THE SIMULATION MODEL

The mathematical basis for the simulation model has been discussed in chapter 3 (mathematical analysis). The following is a simple description of the model.

A sighting platform is placed at a fixed point relative to the chosen x-y axes. A detection area is placed around the platform. Note that the detection area can take on different shapes depending on the sighting criteria used. It need not be circular or symmetrical, but is assumed convex. See figure 3-2.

A true speed u and a randomly chosen direction determined in the sea-frame, i.e. irrespective of the platform, are associated with a whale in the system. The path of the whale starts outside the detection area, at a randomly determined position along the y-axis (see figure 3-5.). In order to simulate the platform moving at fixed speed in a given direction, each animal is given an additional velocity similar in magnitude to that of the platform, but in the opposite direction. The relative speed and direction (animal relative to platform) are calculated and used to determine the straight line path of the animal. An animal is moved through the system by updating its position every dt seconds, dt being a small time increment. In other words a whale is moved from a point (x,y) to a new position (x+dx,y+dy). Expressions for dx and dy are given in section 3.5.

As a whale is moved through the system, one of three mutually exclusive events occur:

(a) its path does not intersect the detection area, i.e there

is no chance of its being sighted

(b) its path intersects the detection area and it is sighted

(c) its path intersects the detection area, but it is not sighted.

Note that the above is the general case. When the fixed detection range sighting criterion is used for example, event (c) cannot occur.

It is important to note that since the animal paths are straight lines, an animal can only enter and leave the detection area once, given that the detection area is convex.

As a whale is entered into the system, it is determined whether event (a) will occur. If it does, the animal is immediately 'forgotten' (only recorded as having passed through the system, contributing one to the number of unseen animals) and a next animal is entered. If (a) will not occur, the animal is moved through the system. At every position (x+dx,y+dy), a probability of having been sighted in the interval from (x,y) to (x+dx,y+dy) is calculated and random number generation is used to simulate the event of the whale being sighted or not. When the animal is sighted, the following relevant values are recorded (Fig 3-9.):

(i) the y-coordinate of position (i.e perpendicular distance from trackline to animal)

(ii) the direct or radial distance r from platform to animal and

(iii) the angle θ between the trackline and the radial.

The animal is only forgotten when it has been sighted and the relevant information recorded (case (b)) or when it leaves the

detection area without having been sighted (case (c)).

It is important to bear in mind that this procedure excludes the possibility of multiple sightings i.e. counting the same animal more than once, thus satisfying one of the neccessary conditions for applying line transect theory.

The way in which whales are entered is based on the assumption that whale speed never exceeds platform speed (both relative to the sea-frame). This is an acceptable assumption, since the sighting platform usually maintains a cruising speed of 12 knots when in the sighting mode, i.e. when gathering sighting data. It has been estimated that minke whales swim at an average speed of less than 5 knots when undisturbed. (Butterworth and Best 1982). Under this assumption it is impossible for an animal to enter the detection area from behind the platform or to overtake the platform. Therefore new whales are entered along a line (in this case the y-axis) perpendicular to the trackline, ahead of the platform (fig 3-2.). A position on the generating line is chosen randomly. If the whales are assumed to be stationary, they have no direction of movement relative to the s-frame. However, if the animals are assumed to be non-stationary their direction of movement relative to the sea-frame is chosen from the appropriate angular distribution which is not uniform. The angular distribution is discussed in section 3.4 .

5. VALIDATION

5.1. Fixed Detection Range Sighting Criterion.

It is known that target (in this case whale) motion leads to density overestimation when line transect theory is applied to sighting data (chapter 2). The bias for this effect can be given by the ratio of the number of targets detected per unit time at target velocity u to the number of targets detected per unit time when targets are stationary.

For the special case where the fixed detection range sighting criterion is used, Koopman (1956) and Skellam (1958) derived an expression for this ratio. (See section 3.5, eq. 3-11). I shall subsequently refer to this special case ratio as the Koopman ratio or KR.

Table 5-1 shows theoretical values of the KR calculated using eq. (3-11) and three values of u. From the table it is clear that in the case of target velocity u=0, the theoretical value of the KR is unity, as one would expect. The table also shows the fact that the KR is an increasing function of target velocity.

Table 5-1. KOOPMAN RATIOS

Koopman ratios for fixed platform velocity v=12 knots	
whale velocity (in knots)	theoretical Koopman ratio
0.0	1.0000
5.0	1.0438
10.0	1.1831

In the first stage of validation the simulation results are tested against the Koopman ratios. The simulation model can be considered to behave favourably if the observed Koopman ratios are significantly close to the theoretically expected values.

Because the simulation model handles one whale at a time, the number encountered per unit time can not be used as a measure. Instead, the total number of animals sighted, corrected for area difference, is used. The measure is in other words:

$$kr = (n/N)x(A/a) \qquad (5\text{-}1)$$

where n = the number of whales sighted

N = the total no. of whales that passed through the system

a = the area over which sighting occurred

A = the total area covered

and a = tracklength x 2r, where r=radius of detection circle

A = tracklength x 2Wmax, where Wmax=half the length along

which whales are generated

This expression is applicable since the fixed detection range sighting criterion is used and an exact value for a exists.

The simulation model is first tested with whale velocity u=0 to confirm that the observed ratio is significantly close to 1.0. Results for u=0 are presented in section 5.1.1. Whale velocities of 5.0 and 10.0 knots are subsequently used and results are presented in sections 5.1.2 and 5.1.3 respectively. Throughout, the platform velocity is fixed at 12.0 knots. Also note that the fixed detection range sighting criterion is used in each of the three above mentioned cases.

5.1.1. Results : u=0 Knots

The results of the first stage of the validation process, for the case where whale velocity u=0 are presented in table 5-2 below.

Three sets of results are given. Each set was obtained by using a different value for the total number of whales N, that passed through the system, in each run. Note that this reflects the length of time that the simulation was run. The s-value in the table indicates the number of completed runs (each done with different random number generating seeds) for the specific case. s therefore also indicates the number of observations (observed Koopman ratios) used in obtaining the mean, standard deviation and confidence interval.

Table 5-2. VALIDATION, u=0.0

Expected Koopman ratio = 1.00 Whale velocity u=0.0 Platform velocity v=12.0 knots			
case	1	2	3
N =total number of whales	1 000	10 000	100 000
s =number of runs	26	12	11
Mean Koopman ratio	0.99588	1.00693	1.00050
Standard deviation	0.04487	0.01329	0.00531
t statistic	-0.4682	1.8063	0.3123
d.o.f	25	11	10
p value	0.6437	0.0983	0.7612
95% Confidence interval			
lower bound	0.97775	0.99849	0.99693
upper bound	1.01401	1.01537	1.00407
Minimum observed KR	0.89549	0.98769	0.99094
Maximum observed KR	1.07459	1.03081	1.00905

The values of the mean Koopman ratio are all close to the expected value 1.00. A t-test is used to test the null hypothesis (Ho) that the observed mean KR is equal to the theoretical value (1.00), against a two-sided alternative. Results show that in all three cases Ho is accepted. This is also reflected in the 95% confidence intervals given in table 5-2 above, since the theoretical value lies within the given intervals.

As one would expect, the standard deviation decreases as the total

number of whales (N) increases (see Appendix II). The standard deviation in case 3 is at an acceptable level as are the minimum and maximum observed values. Within this section of validation, most attention will therefore be attached to case 3 results (i.e N=100 000).

One may therefore conclude that in the case of u=0 the observed mean Koopman ratio is significantly close to the expected or theoretical value. The case where whale velocity is non-zero, can now be considered.

5.1.2. Results : u=5 knots

The results for whale velocity u=5.0 knots are presented in table 5-3 below. The notation is as in table 5-2.

Table 5-3. VALIDATION, u=5 KNOTS

Expected Koopman ratio = 1.0438			
Whale velocity u=5.0	Platform velocity v=12.0 knots		
Case	1	2	3
N =total number of whales	1 000	10 000	100 000
s =number of runs	26	10	10
Mean Koopman ratio	1.04155	1.04378	1.04352
Standard deviation	0.04947	0.01433	0.00362
t statistic	-0.2319	-0.0044	-0.2446
d.o.f	25	9	9
p value	0.8185	0.9966	0.8123
95% Confidence interval			
lower bound	1.02156	1.03353	1.04119
upper bound	1.06154	1.05403	1.04611
Minimum observed KR	0.92531	1.02451	1.03820
Maximum observed KR	1.14424	1.06132	1.04971

In this case the null hypothesis (Ho) tests whether the observed mean Koopman ratio is equal to the theoretical value, 1.0438. Ho is accepted in all three cases and once again this is also shown in the 95% confidence intervals, since the value 1.0438 lies

within the intervals.

As in the previous case, the standard deviation decreases as the total number of whales (N) increases as does the interval between the minimum and maximum observed Koopman ratios.

5.1.3. Results : u=10 Knots

The results for whale velocity u=10.0 knots are presented in table 5-4. below. The notation is as in tables 5-2 and 5-3.

Table 5-4. VALIDATION, u=10,0

Expected Koopman ratio = 1.1831 Whale velocity u=10.0 Platform velocity v=12.0 knots			
Case	1	2	3
N =total number of whales	1 000	10 000	100 000
s =number of runs	26	10	10
Mean Koopman ratio	1.17370	1.18596	1.18185
Standard deviation	0.04596	0.01567	0.00437
t statistic	-1.0429	0.5852	-0.9045
d.o.f	25	9	9
p value	0.3070	0.5728	0.3893
95% Confidence interval			
lower bound	1.15513	1.17475	1.17872
upper bound	1.19227	1.19717	1.18498
Minimum observed KR	1.07459	1.15419	1.17511
Maximum observed KR	1.28353	1.21057	1.18742

Similar to the previous case, the appropriate null hypothesis is accepted in all three cases. Here the theoretical value of the KR is 1.1831. Once again the 95% confidence intervals contain the value 1.1831.

It is also clear from table 5-4. that the standard deviation and interval between minimum and maximum observed KR's decrease as the total number of whales (N) increase.

5.1.4. Conclusion

Under the fixed detection range sighting criterion, for zero and non-zero whale velocities, the simulation model behaves favourably and as theoretically expected. The first step of validation has now been completed successfully.

5.2. Negative Exponential Sighting Rate

The second stage of the validation process consists of analysing results obtained by using two forms of the negative exponential sighting rate for the case where whale velocity is zero. Expressions for the two forms are given in section 3.6.2, eq. (3-12).

Butterworth (1982) derives the corresponding functional forms of g(y), the distribution of the right angle distance y at sighting. (Section 3.6.2, eqs. 3-15 and 3-16)

The object of this section is to compare the distribution (in histogram form) of the y-distances obtained from the simulation runs to the theoretically expected distribution i.e. g(y) given in equations (3-15) and (3-16).

The frequencies of observed y-distances are plotted in histogram

form with class intervals of 0.2 Nm and with a cut-off at R=15.0 Nm. Outside the circle of radius 15.0 Nm, the probability of being sighted is so small that it can be taken as zero for all practical purposes. The relevant g(y) is then fitted to the data and a Chi-square goodness of fit test completed for each run.

The procedure for fitting g(y) to the observed frequencies is as follows. Using a numerical integration method (NAGLIB computer package, routines D01AHE, S18ACE, S18ADE), integrals of g(y) are calculated over:

(a) each of the class intervals i, i=1,2,....,I

(b) the whole range, y=0 to y=R the cut-off

The expected number of observations within interval i, is then calculated as:

(Total no. of observations) x (integral over i) / (total integral)

The total number of observations refer to the number of sightings counted in the simulation and the total integral refers to the value obtained in (b) above.

5.2.1. Results: case (i), negative exponential

Results were obtained using the sighting rate $f(r) = \mu \exp(-\lambda r)$ and setting $\mu = 2.5$ and $\lambda = 1.0$. The constant μ is chosen so that a sufficient number of sightings is obtained when the simulation is run for a reasonable time, e.g. N=100 000. It is important to note that because the simulation approximates a continuous process by a discrete one, results will be valid as long as the factor $\mu \cdot dt$ is not too large (see section 3.5).

Results of the Chi-square goodness of fit tests on 5 runs are presented in table 5-5 below. N refers to the total number of whales that passed through the system. D-square refers to the calculated or observed Chi-square value ,d.o.f. to the degrees of freedom and the p value to the observed significance level.

Table 5-5. Goodness of fit tests for case i

Run no.	N	d.o.f	D-square	p value
1	100 000	25	23.13	0.5700
2	100 000	25	14.67	0.9488
3	100 000	25	19.39	0.7780
4	100 000	24	11.89	0.9811
5	100 000	24	22.21	0.5667

From table 5-5 it is clear that since the p value is in each case much greater than 0.05, the null hypothesis cannot be rejected. Here the null hypothesis is that the underlying distribution is that given in equation (3-15).

Figure 5-1 show the histograms of the observed y-values and the theoretical curves for runs 1 and 2 and clearly support the outcome of the chi-square tests.

5.2.2. Results: case (ii), negative exponential

In this case $f(r)=(\mu/r).\exp(-\lambda r)$ and μ was set to 50.0 and λ =0.0. As already mentioned, r was set to 0.000001 whenever r=0.0. Note that in this case the 1/r factor overrides the dt factor for small r, so that a somewhat larger μdt is acceptable. Results of the Chi-square goodness of fit tests on 4 runs are presented in table 5-6 below. The notation is as in table 5-5 above.

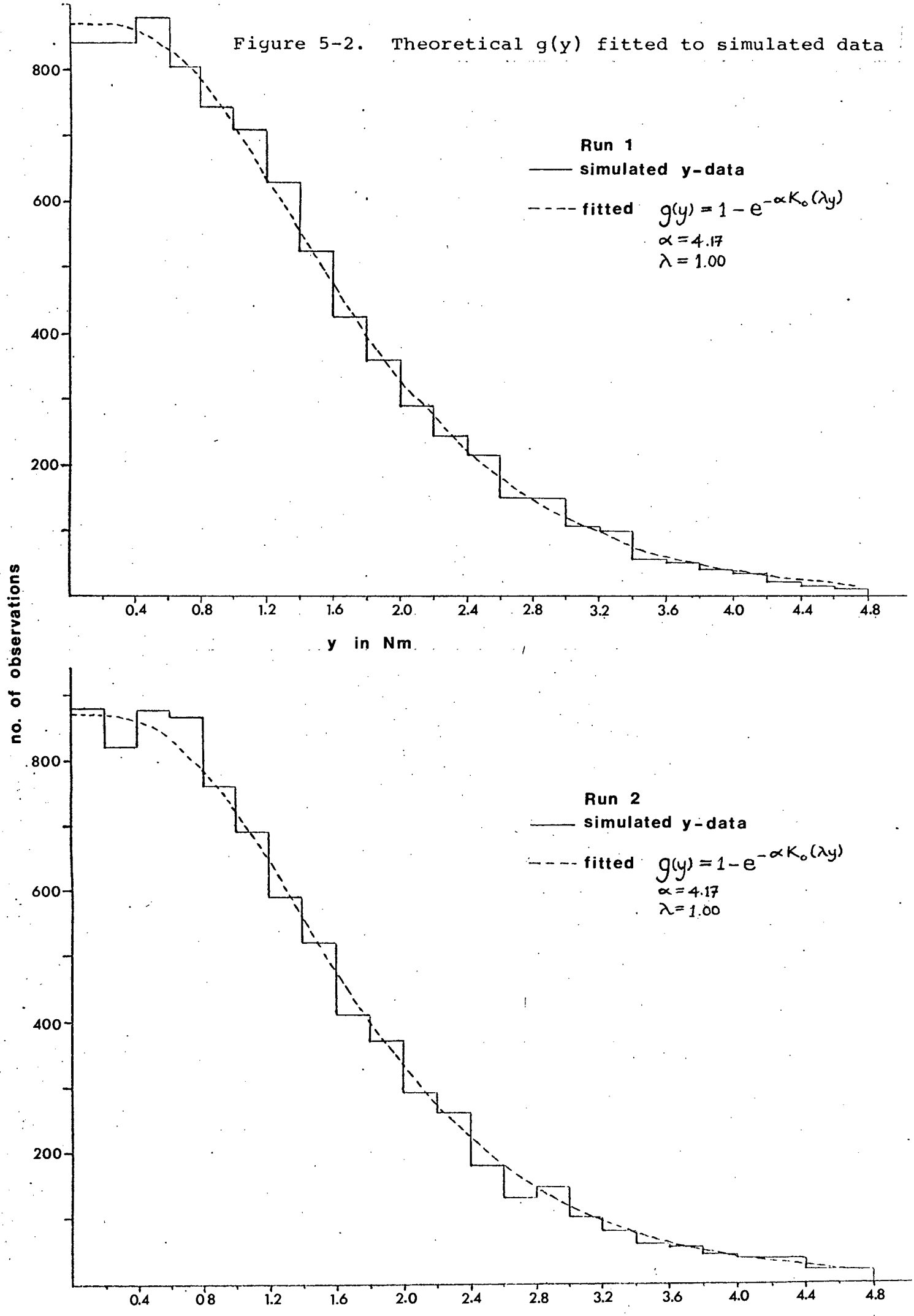
Table 5-6. Goodness of fit tests for case ii

Run no.	N	d.o.f	D-square	p value
1	100 000	30	31.82	0.3759
2	100 000	30	35.86	0.2127
3	100 000	29	27.62	0.5383
4	100 000	29	31.34	0.3496

Table 5-6 indicates that the null hypothesis (that the underlying y distribution is as that given in equation 3-16) can be accepted for each of the 4 runs.

Figure 5-2 showing the theoretical g(y) curves and the observed y-values of runs 1 and 2 in histogram form also reflect the good fit of the data.

Figure 5-2. Theoretical g(y) fitted to simulated data

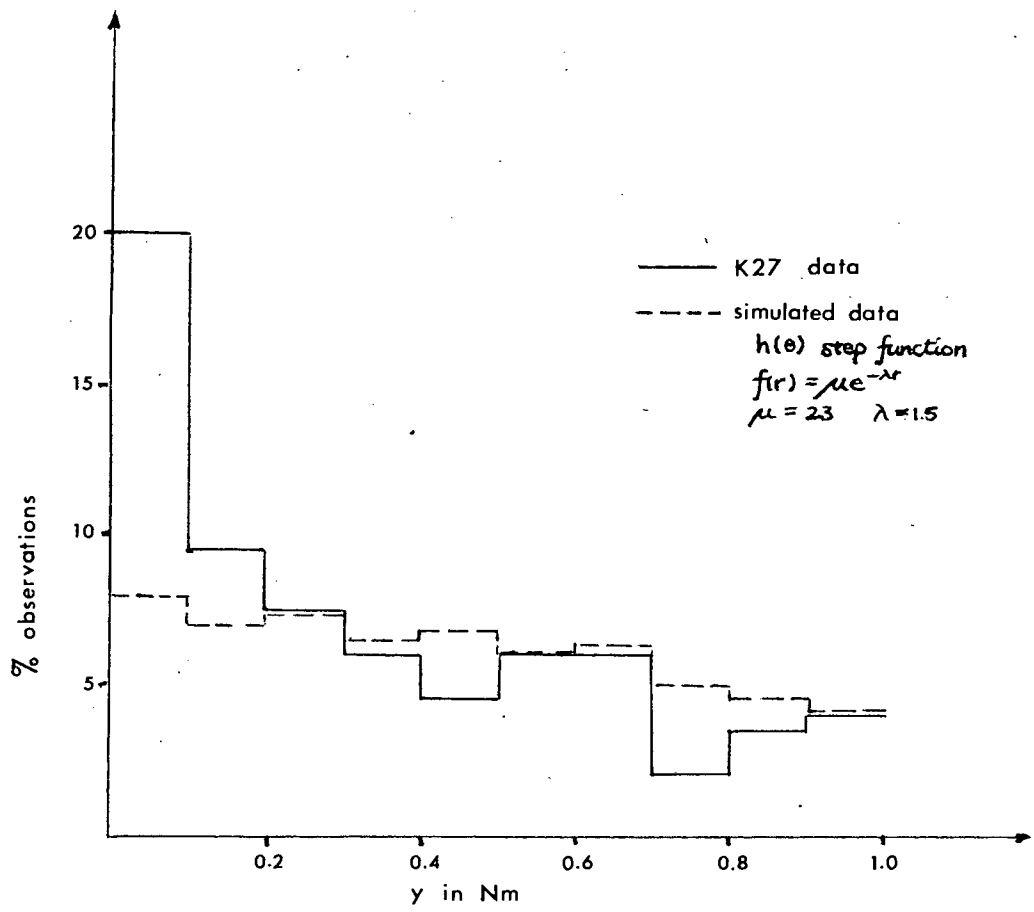


5.2.3. Conclusion

One can conclude from the above that the expected results are obtained for both forms of the negative exponential sighting rate under zero whale velocity. A more realistic parametrisation for $h(\theta)$ can now be developed.

Figure 5-3.

Comparing K27 and simulated y-data



6. FURTHER ANALYSIS

The object of the next stage is to produce simulation data that resemble the real data more closely than has been the case so far. This is done by introducing a more realistic form of $h(\theta)$ and by finding appropriate values for the parameters μ and λ in form (i) of the negative exponential sighting rate (eq. 3-12(i)). (See fig. 5-3)

6.1. Fitting $h(\theta)$

The $h(\theta)$ used till now implies that all forward directions are searched with equal intensity (sections 3.6.1 and 3.6.2). The new form of $h(\theta)$ implies that directions straight ahead are searched with greater intensity than directions to the sides of the vessel.

6.1.1. Analysing the real data

The new form of $h(\theta)$ is based on the real data presented in Doi et al. (1982). Frequencies of search effort by angle in the case of two topmen were recorded and the data are presented in figure 6-1 below. Although the histogram shows a slight assymmetry about the trackline, $h(\theta)$ is still assumed to be symmetrical. This assumption is made since the real data are merely used to suggest a general form for $h(\theta)$ and in the absence of more reliable data it is logical to assume symmetry.

The histogram of the data (after averaging over the frequencies of angles to the left and right of the trackline to render the distribution symmetrical, (figure 6-2) shows that the normal

distribution is probably a good candidate.

Figure 6-1. Frequencies of search effort by angle

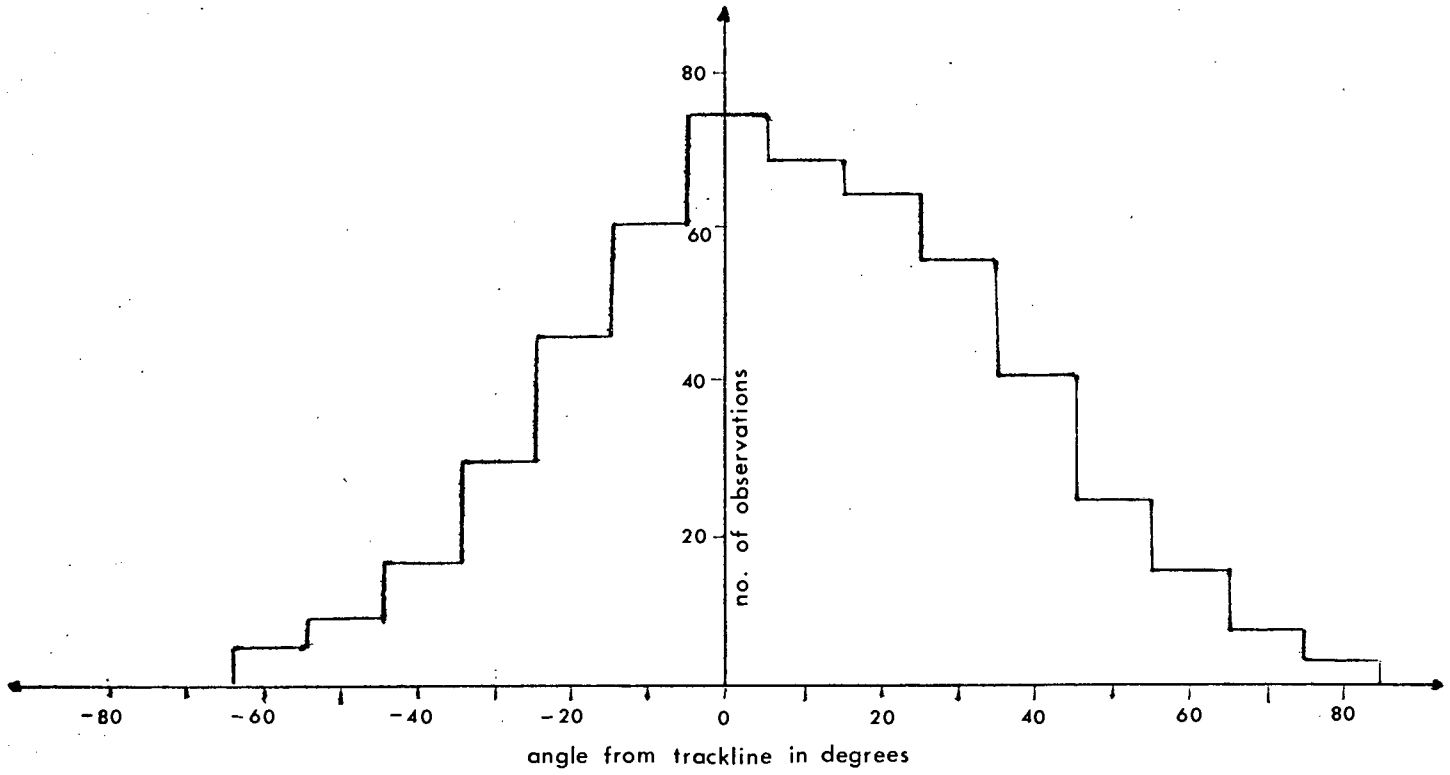
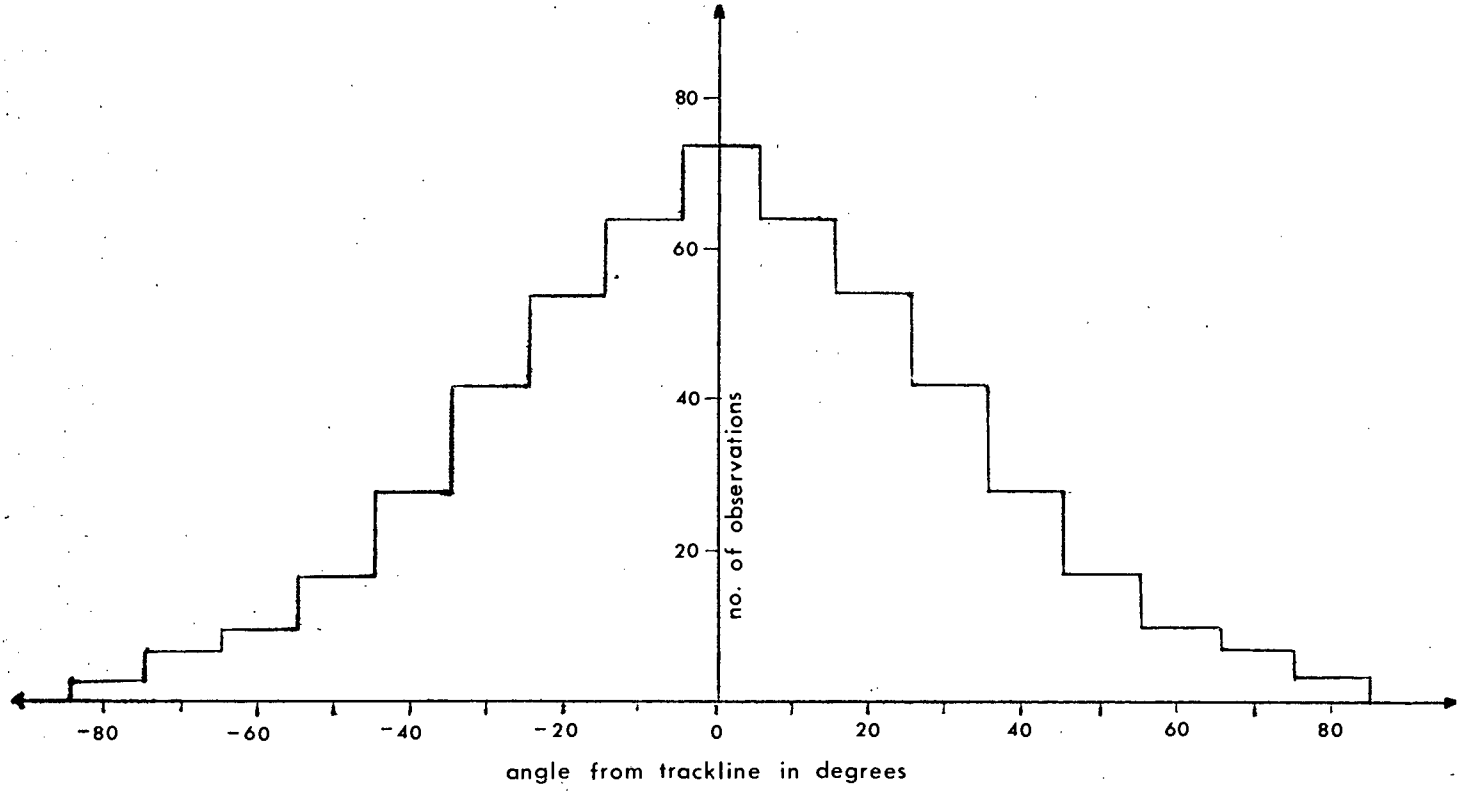


Figure 6-2. Symmetrical form of figure 6-1 data



6.1.2. Results

An estimate of the standard deviation to be used in the normal distribution cannot be calculated from the frequency data and an intuitive value of about 30 was chosen. Goodness of fit tests between the real data and the normal distribution with mean 0 for various values of σ were completed. Results are presented in table 6-1. below.

Table 6-1. Goodness of fit tests

σ	25	30	35	40
D^2	61.00	3.66	11.70	35.70
d.o.f	12	14	14	16
$\chi^2_{0.05}$	21.03	23.69	23.69	26.30
p	0.0000	0.9972	0.6304	0.0032

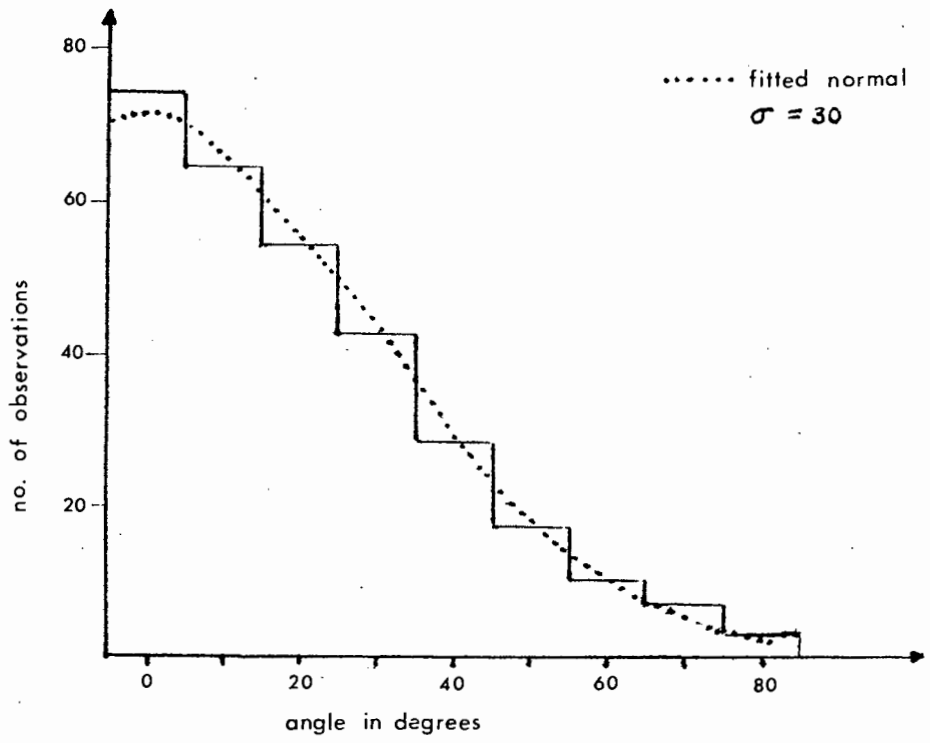
Table 6-1. clearly indicates that the normal distribution with mean 0 and standard deviation 30 gives the best fit (figure 6-3). A standard deviation of 35 is also acceptable, but since 30 offers a better fit, this was chosen. Note that it is possible to continue the search for a better σ, but at this stage a standard deviation to the nearest 5 is satisfactory. In all subsequent runs, $h(\theta)$ takes on the form of the normal distribution with mean 0 and standard deviation 30. However, since $p=f(r)h(\theta)$ the constant in $h(\theta)$ is incorporated with that of $f(r)$ and therefore:

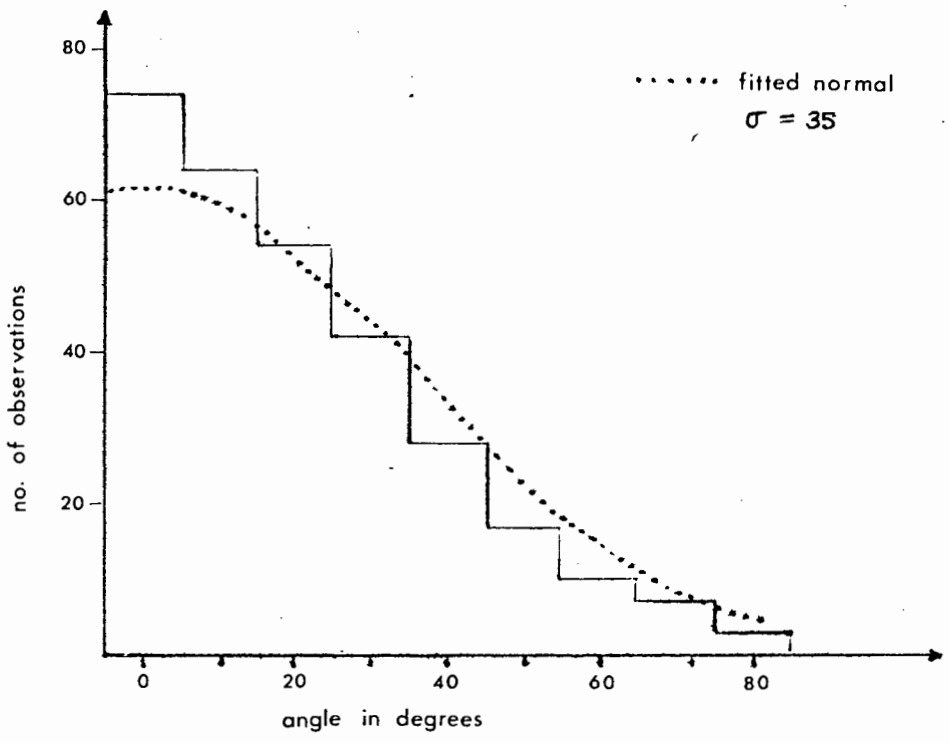
$$f(r) = \mu e^{-\lambda r}$$
$$h(\theta) = \begin{cases} e^{-1/2\,(\theta/\sigma)^2} & 0 \le \theta \le \pi/2 \\ 0 & \pi/2 < \theta < \pi \end{cases} \qquad (6\text{-}1)$$

Figure 6-3.

Normal distribution fitted to figure 6-2 data

(only presenting right half of graph)





6.2. Finding values for μ and λ

The next step is to find suitable values for μ and λ. Only one case of real sighting data will be considered, that obtained by the Kyo Maru 27 (K27) on the 1980/81 cruise. These data are typical of whale sighting data, whereas some of the data from the other vessels for this particular cruise are somewhat atypical.

Since the aim is to create simulation data that resemble the real data, various values of μ and λ are used in simulation runs and the r and θ distributions of these runs are compared to the distributions of the real data. (Note that r and θ are the radial distance and sighting angle respectively (see fig. 3-9).) For this purpose a simple analysis of the real sighting data is necessary.

6.2.1. Analysing the real sighting data

Only sightings by topmen while the vessel was in sighting mode are considered and the observed r and θ values are extracted from the sighting records.

About 99% of the θ values were recorded as multiples of 10 degrees. Figure 6-5 below, therefore presents the data in a bar graph rather than a histogram. It was also decided to round the simulated values to the nearest 10 degrees and rather compare the grouped rounded data to the real data. Almost 9% of the observed θ 's were recorded as 90 degrees. This caused a peak in the distribution that is incompatible with the normal distribution form of $h(\theta)$. As already mentioned, this form of $h(\theta)$ implies very little or no search effort in the 90 deg. direction and therefore

very little or no observations in that direction. The observations of $\theta > 85$ deg. are therefore ignored for the purpose of this study. The frequencies of the r values are presented in figure 6-4 and the percentage sightings within given intervals for r and for θ at at given direction are listed in Appendix IV.

Figure 6-4. K27 radial distance data

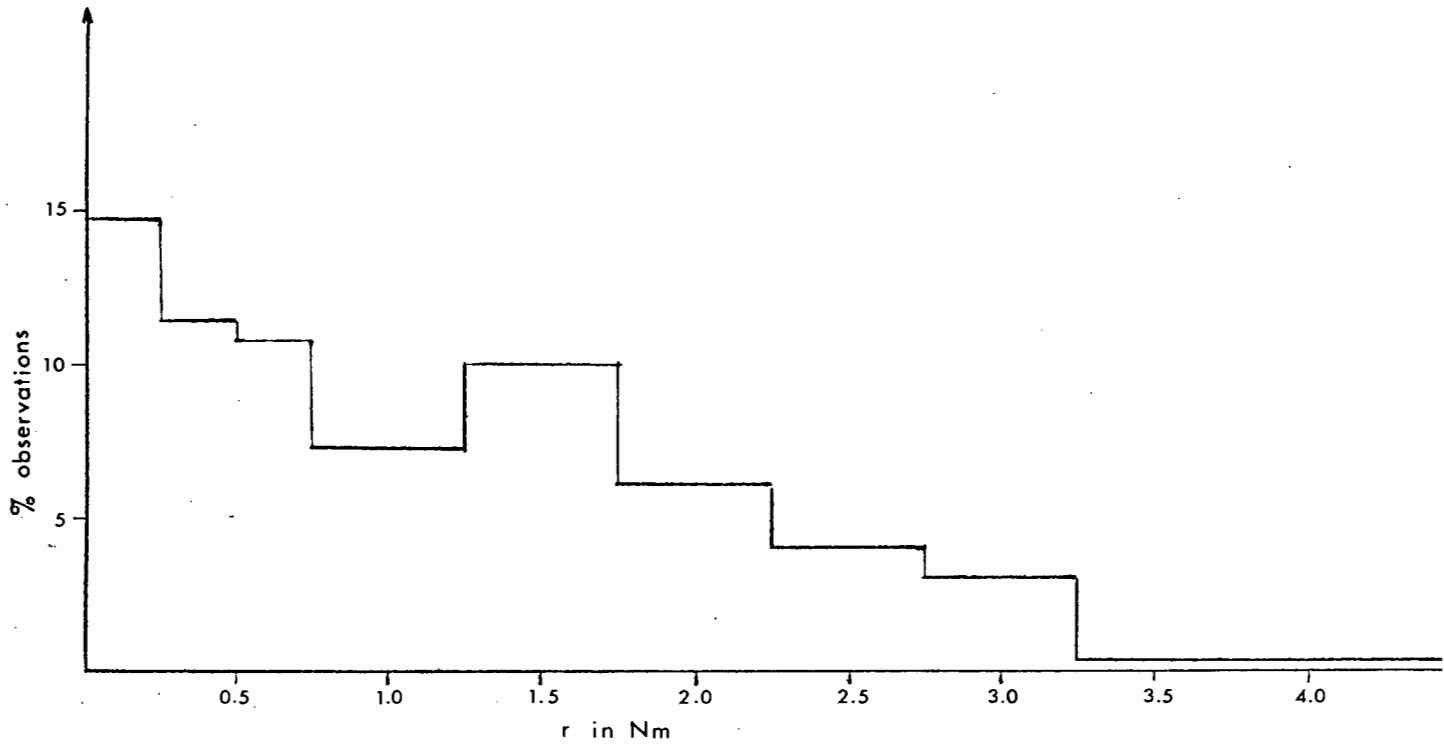
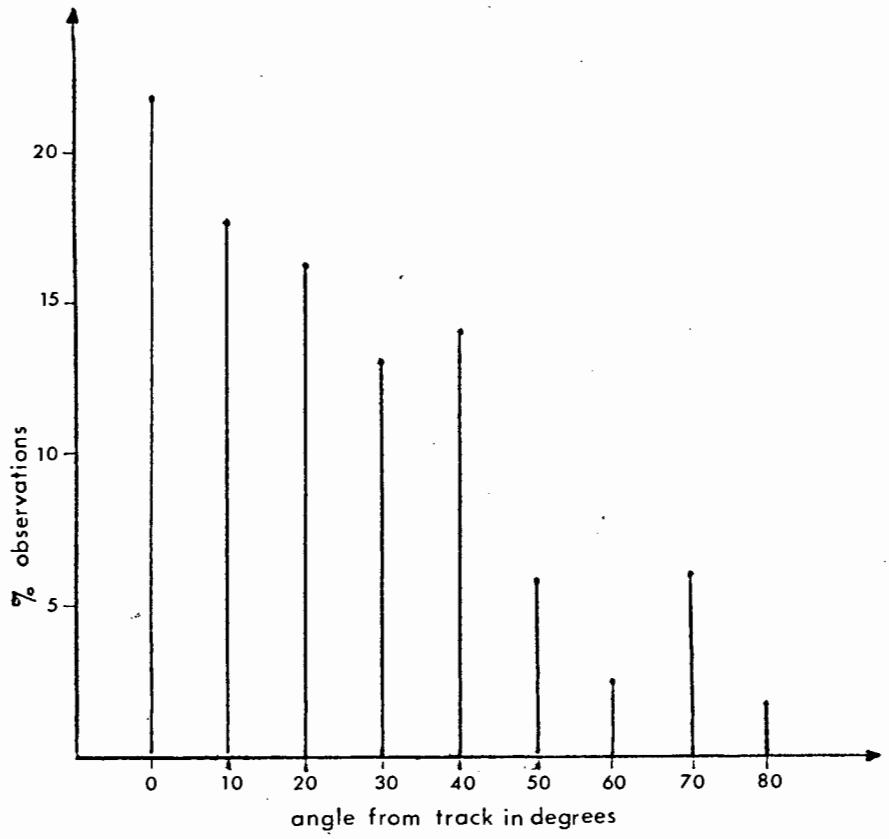


Figure 6-5. K27 sighting angle data



6.2.2. Results

For each chosen value of μ and λ an average of three runs, each with an N value (total number of whales) of 100 000 were done. It is important to note that whale velocity u=0 was used. The real data actually corresponds to the case where u is not zero. According to Butterworth and Best (1982) u can be estimated at about 3 knots. Strictly speaking one should therefore compare simulation data for u=3 knots to the real data. For the sake of simplicity (and economy) it was decided to use u=0 to obtain an initial fit. It is also reasonable to expect that the r and θ distributions and therefore the μ and λ values will not change drastically with the small change of u from 0 to 3 knots.

A Chi-square measure of error is calculated for each run as well as for the data of the three or more runs combined. The Chi-square measure of error is calculated as follows:

$$D^2 = \sum_i \frac{(o_i - e_i)^2}{e_i}$$

where o_i - % observations in i'th cell: simulation data

e_i - % observations in i'th cell: real data

The above measure of error is calculated for grouped r data and grouped θ rounded to the nearest 10 degrees. Note that the e_i values are in each case the values given in the second column of tables in Appendix IV. Also note that the measure of error cannot be considered a true Chi-square test value, since in that case the last cells of the expected distribution would have to be combined to give an e_i value greater than 5.

Results of the Chi-square measure of error for the combined data (from three or more runs) of r and θ are given in table 6-2

below.

Table 6-2. Measures of error

D^2 measures of error for r and θ						
μ \ λ	0.50	1.00	1.25	1.50	1.75	2.00
2.5	12.86	12.80	14.84	11.86	12.18	18.55
	546.60	92.13	27.04	17.10	24.21	37.76
5.0			13.15	12.85	12.62	11.40
			29.40	16.25	17.87	33.46
20.0			11.81	16.30	9.80	
			39.38	17.81	20.79	
23.0	EACH CELL :	D^2_θ		13.77		
		D^2_r		22.07		
30.0			14.51	13.69	15.00	
			58.70	23.36	20.24	
40.0			12.04	11.62	11.96	
			65.26	24.33	26.33	

Initially μ was fixed at 2.5 and various values of λ from 0.25 to 2.00 were used. The measure of error for θ is relatively insensitive to change in λ . However, great fluctuation was observed in this measure. For the same values of μ and λ for example, under identical conditions but only varying the random number seeds, D^2 values for θ of from 9.0 to 21.0 were observed.

Much longer runs or variance reduction techniques would be required to obtain a better fit and therefore more accurate values of μ and λ.

The D^2 values for r, however, are clearly a minimum in the area of λ =1.5. Subsequently runs using values of λ from 1.25 to 1.75 were considered sufficient for a given μ. For most values of μ the minimum D_r^2 value is situated at λ =1.5. Only in the case where μ =30.0 is it at 1.75. The values of D_r^2 (for r) for λ =1.5 and 1.75 are so close however (16.25 and 17.87) that λ probably lies in the region of 1.50 and 1.75. As μ is increased, the minimum D^2 for r also increases and the fit deteriorates.

The combination μ =5.0, λ =1.5 was considered to give a reasonable fit and was therefore chosen for further analysis.

Let g(y) be the detection curve as defined in Chapter 1, then g(0) indicates the probability of seeing a whale on the trackline in the case where u=0. Although one of the basic assumptions of line transect theory is that g(0)=1, this is not the case when the function $f(r)=\mu e^{-\lambda r}$ (Butterworth 1982). In real sighting surveys g(0) is also not unity, partly because whales are not sightable (visible) at all times and also because they are not stationary and therefore not on the trackline all the time. Considering the analytical expression for g(0),

$$g(0) = 1 - e^{-\mu/\lambda v} \qquad (6\text{-}3)$$

(see appendix V) and setting μ =5.0 and λ =1.5 with vessel speed v the normal 12 knots, a value of g(0)=0.24 is obtained.

Butterworth et al. (1982) develop a method of determining g(0) experimentally. Using this method, an experimental value for g(0)

can be calculated from the real data. The g(0) value obtained from the 1980/81 cruise is g(0)=0.72. This seems realistic and clearly g(0)=0.24 is very unrealistic in comparison. A corresponding value of μ=23.0 (λ=1.5,v=12) has to be used in equation 6-3 to obtain g(0)=0.72. It was therefore decided to analyse this case as well, even though the fit of the r and θ distributions are not as good as in the case where μ=5.0.

The graphs of the simulated and real r-data (figure 6-6) show that the sighting function seems incapable of producing 'enough' observations at small r (0.00 - 0.50 Nm.). The reason why μ=5.0 gives a better D_r^2 value, indicating better fit, is probably because of the fact that it has a shorter tail on the lower (left hand) side than is the case for μ=23.0. The problem most probably lies with the form of f(r) used.

A number of approaches can be followed to obtain greater similarity between the simulated and the real data. As already mentioned, variance reduction techniques can be introduced and a greater number of much longer runs can be done. A smaller value for dt can be used and a different form of f(r) in the hazard-rate model can be investigated. An example of an alternative f(r) is given in section 3.6.2 eq. 3-13.

At this stage we shall, however, proceed with the forms of f(r) and h(θ) and the values of μ and λ as determined in this section.

Figure 6-6. Simulated versus real r-data

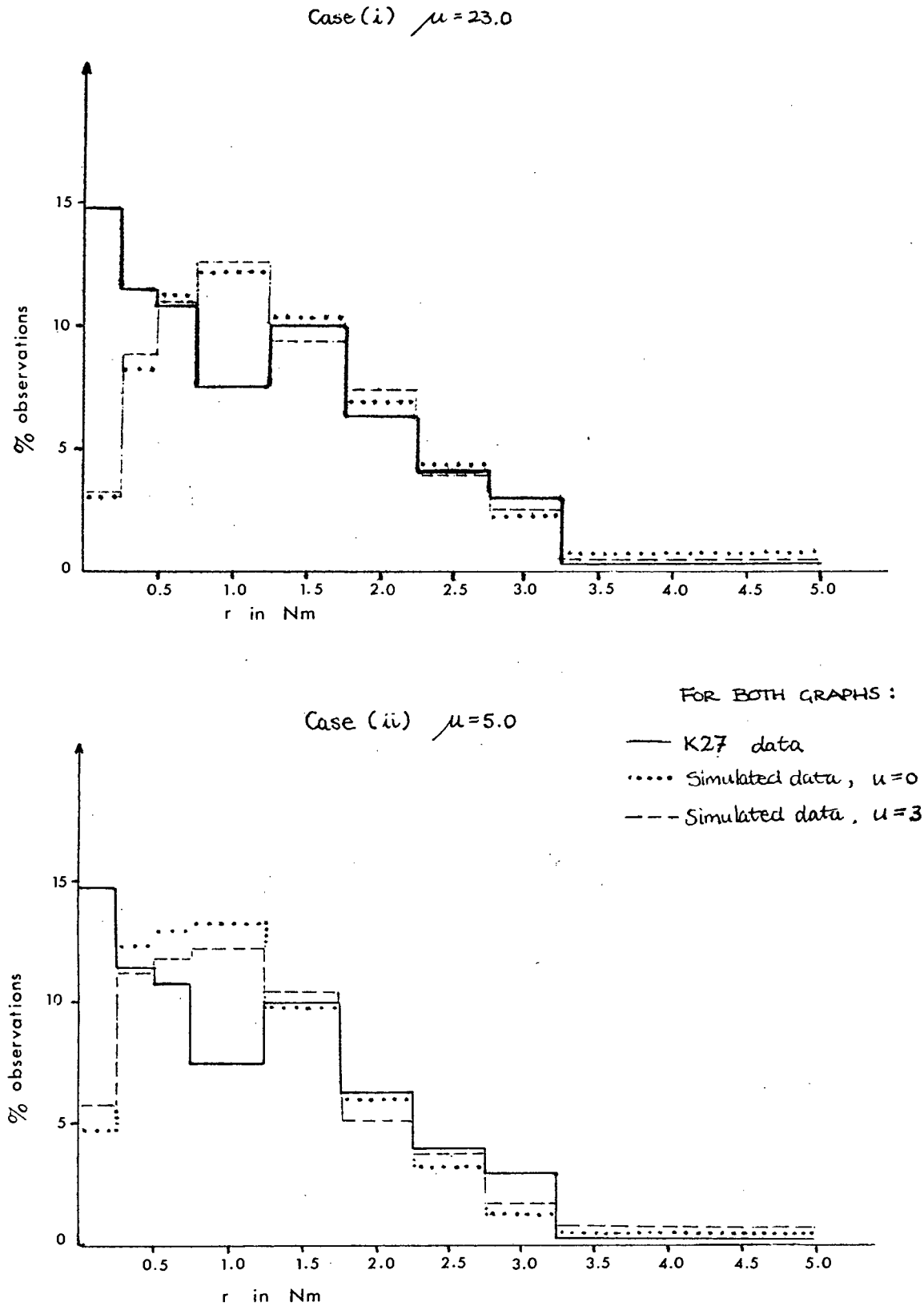
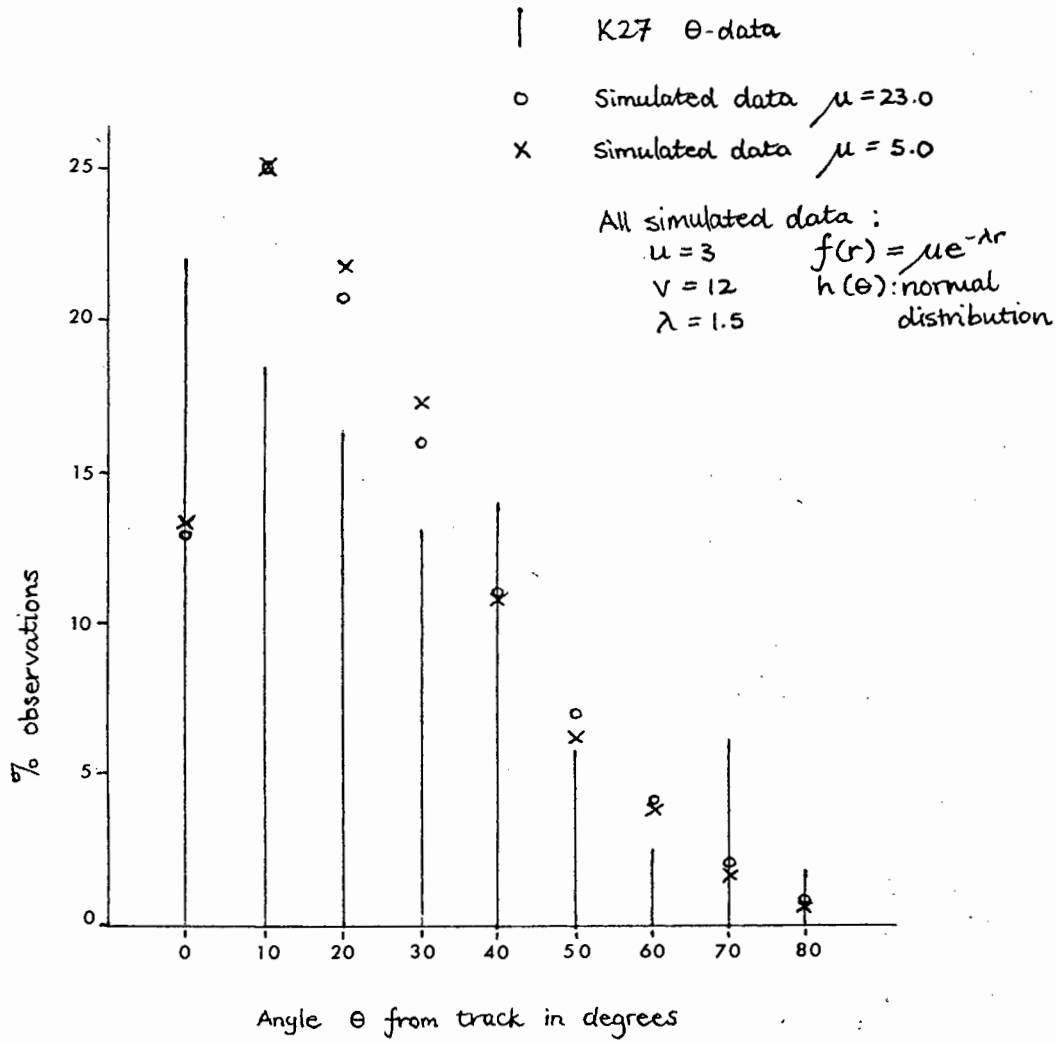


Figure 6-7.

Simulated vs. K27 Θ-data



7. FINAL RESULTS

It is now perhaps useful to recall that the main aim of this study is to attempt to assess the bias in density estimates resulting from random whale movement. In this section the relevant density estimates will therefore be obtained and compared.

The sighting function as given in form 1.2 (appendix III) for the two cases: (i) μ =23.0, λ =1.5 and (ii) μ=5.0, λ =1.5 will be considered. This sighting function will be used in the simulation and for each case a negative exponential density estimate (NE estimate) for zero whale velocity and an experimental value for g(0) will be calculated. The density estimate corrected for the fact that g(0) is not unity will then be compared to the true density which is known from the simulation model. Furthermore, the uncorrected density estimate for whale velocity u=3 knots will be obtained and compared to the true density and the estimate for u=0.

The use of the NE density estimate and the procedure for calculating the g(0) correction factor duplicates the way in which the density estimate is obtained from the real data (Butterworth and Best 1982). This is also the main reason for using the NE estimator. It is easy to calculate and has been used traditionally in minke whale sighting surveys. It is not, however, necessarily the most appropriate estimator.

Although the validation process showed that the simulation model behaves favorably, the case where $h(\theta)$ is the step function as defined in form 1.1 (appendix III) will be used to verify that the correct experimental value for g(0) is obtained. Only case (i)

above will be considered here.

7.1. Notation

Let $D_V(u)$ (sometimes v and u are omitted) indicate the NE density estimate determined under vessel speed v and whale speed u. See appendix VIII for the expression for the NE estimate. Let $g_V(0)$ indicate the experimentally determined value of g(0) under vessel speed v. Any reference to a theoretical value of g(0) implies the value obtained using eq. (6-3) above. True D refers to the true density as known from the simulation model. Since the simulation model is based on a fixed number of whales being generated over a length Wmax and whales moving through the system one at a time, it is not possible to determine the tracklength L as such. The true density given here is therefore actually the density multiplied by the tracklength L (see appendix VII). Similarly, $D_V(u)$ is actually the NE estimate multiplied by L (appendix VIII).

7.2. Calculating g(0) for h(θ) the step function.

For the case where the sighting function is of the form no.1.1 (appendix III), Butterworth (1982) has calculated the correction factors to be used when a negative exponential form is fitted to data with underlying distribution $g(y)=1-e^{-\alpha y K_1(\lambda y)}$. In other words, if D is the NE estimate, D multiplied by the correction factor would be a better estimate of the true density. The correction factors C, given in table 7-1, consist of a g(0) factor and a so-called shape factor, say s. The g(0) factor compensates for the fact that g(0) is not unity and the shape factor for the fact that the negative exponential is used instead of the true

distribution g(y) given above. Since an expression for g(0) exists (appendix V), it is easy to calculate the g(0) factors for given $\mu/\lambda v$. The correction factor can then be split into its two factors. The separate values for s and g(0) are given in table 7-1.

Table 7-1. Correction factors

$$g(y) = 1 - e^{-\frac{\mu}{v} y K_1(\lambda y)}$$

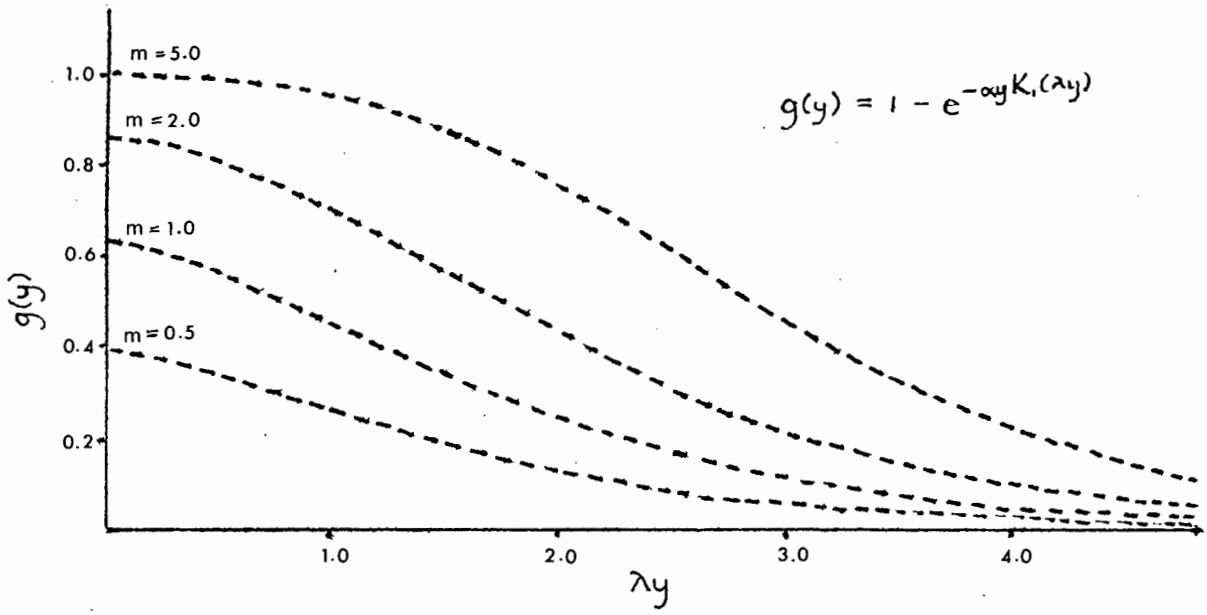
$$g(0) = 1 - e^{-\mu/\lambda v}$$

$\mu/\lambda v$	C	g(0)	s
0.1	8.45	0.10	0.85
0.2	4.40	0.18	0.79
0.4	2.38	0.33	0.79
0.6	1.71	0.45	0.77
0.8	1.38	0.55	0.76
1.0	1.19	0.63	0.75
1.2	1.06	0.70	0.74
1.4	0.97	0.75	0.73
1.6	0.90	0.80	0.72
1.8	0.85	0.83	0.71
2.0	0.81	0.86	0.70
2.5	0.74	0.92	0.68
3.0	0.70	0.95	0.67
4.0	0.65	0.98	0.64
5.0	0.62	0.99	0.61

The values for C were taken from Butterworth (1982) table 1.

The plots of g(y) (after Butterworth 1982) illustrate the fact that g(0) increases and s decreases as $\mu/\lambda v$ increases (figure 7-1). Note that the m value in the graph refers to $\mu/\lambda v$ (i.e. $m=\mu/\lambda v$).

Figure 7-1. Plots of g(y) as given in eq.(3-15)



The ratio of the true density to the NE density estimate gives an estimate of the correction factor to be used. Note, however, that True $D/D_v(0)$ includes both the g(0) and s factors. Results for vessel speed 7 and 12 and parameters μ=23.0 and λ=1.5 are presented in table 7-2 below.

Table 7-2. Estimated correction factors

V	$\mu/\lambda v$	$\hat{C}$
7	2.19	0.7534
12	1.28	0.9890

It can be seen from table 7-1 that for $\mu/\lambda v$ between 2.0 and 2.5, C lies between 0.81 and 0.74 and for $\mu/\lambda v$ between 1.2 and 1.4, C lies between 1.06 and 0.97. It is clear that the estimates of C as given in table 7-2 lie within the expected intervals.

Experimental values of g(0) are calculated using (a) the NE density estimates for v=7 and v=12 knots and (b) the NE estimates corrected for shape (i.e. Ds). Results are presented in table 7-3 below.

Table 7-3. Calculated g(0) values

	(a) D uncorrected	(b) D corrected for shape
D_7/D_{12}	1.3127	1.2359
$g_{12}(0)$	0.61	0.72
$g_7(0)$	0.80	0.88

It is clear that using the density estimate corrected for shape in calculation of g(0) leads to the correct values of g(0) whereas using the uncorrected values of D gives an incorrect g(0) value.

This shows that when the negative exponential function does not fit the right angle distance data well, calculation of g(0) by the experimental procedure can lead to an incorrect value, subsequently leading to an over- or underestimate of the true density. In this particular case a too small g(0) results and therefore an overestimate of True D is obtained.

Plots showing simulated y-data and the true underlying distribution as well as a negative exponential function fitted to the data are presented in figures 7-2, 7-3 below. The relevant parameters common to both figures are: u=0, μ =23.0, λ =1.5. Note how the histogram flattens out for small values of y, especially when v=7. This illustrates how the NE estimate, which is based on the intercept, overestimates the true density.

Figure 7-2.

Negative exponential and true distribution fitted to simulated data for v=7

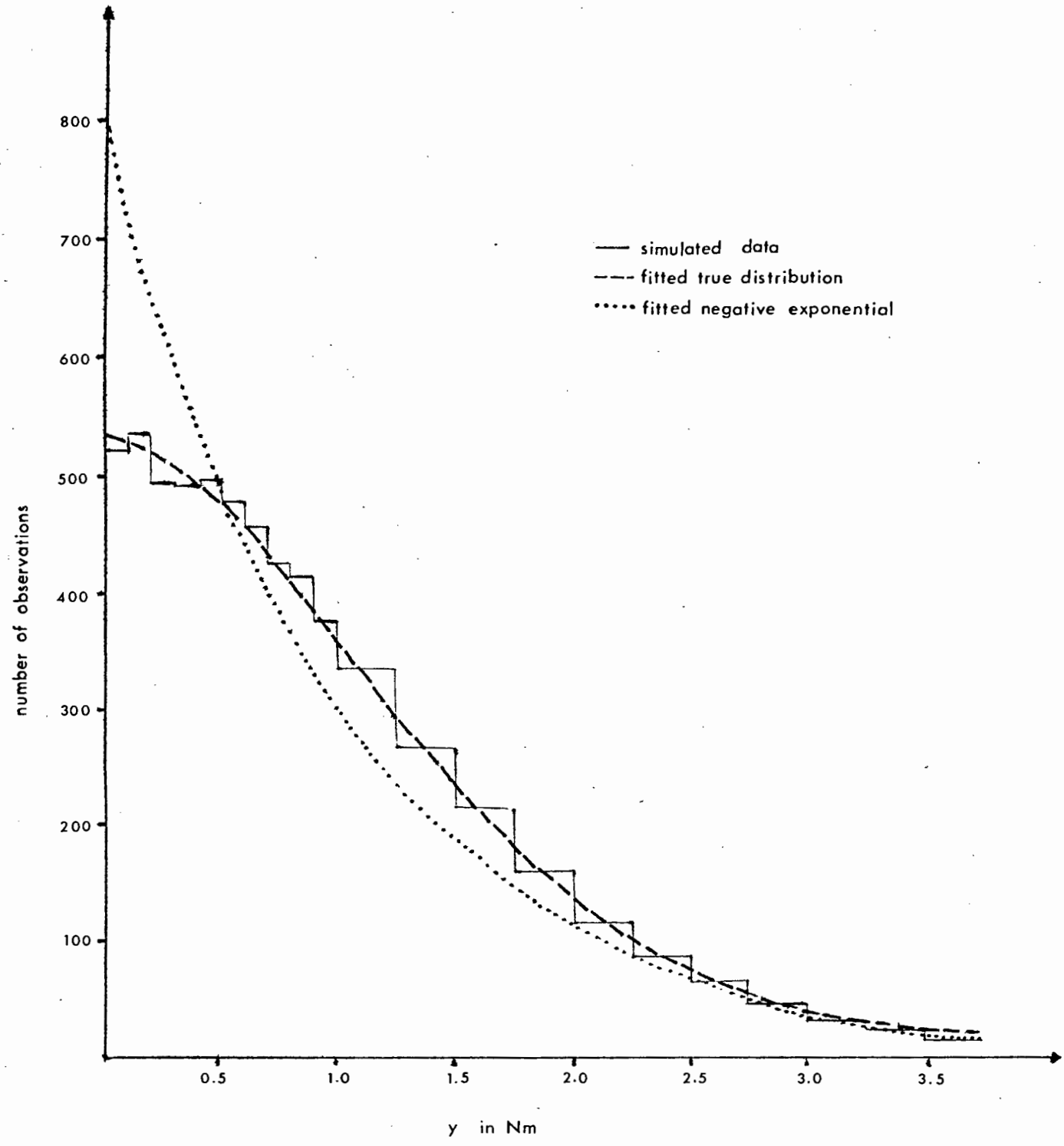
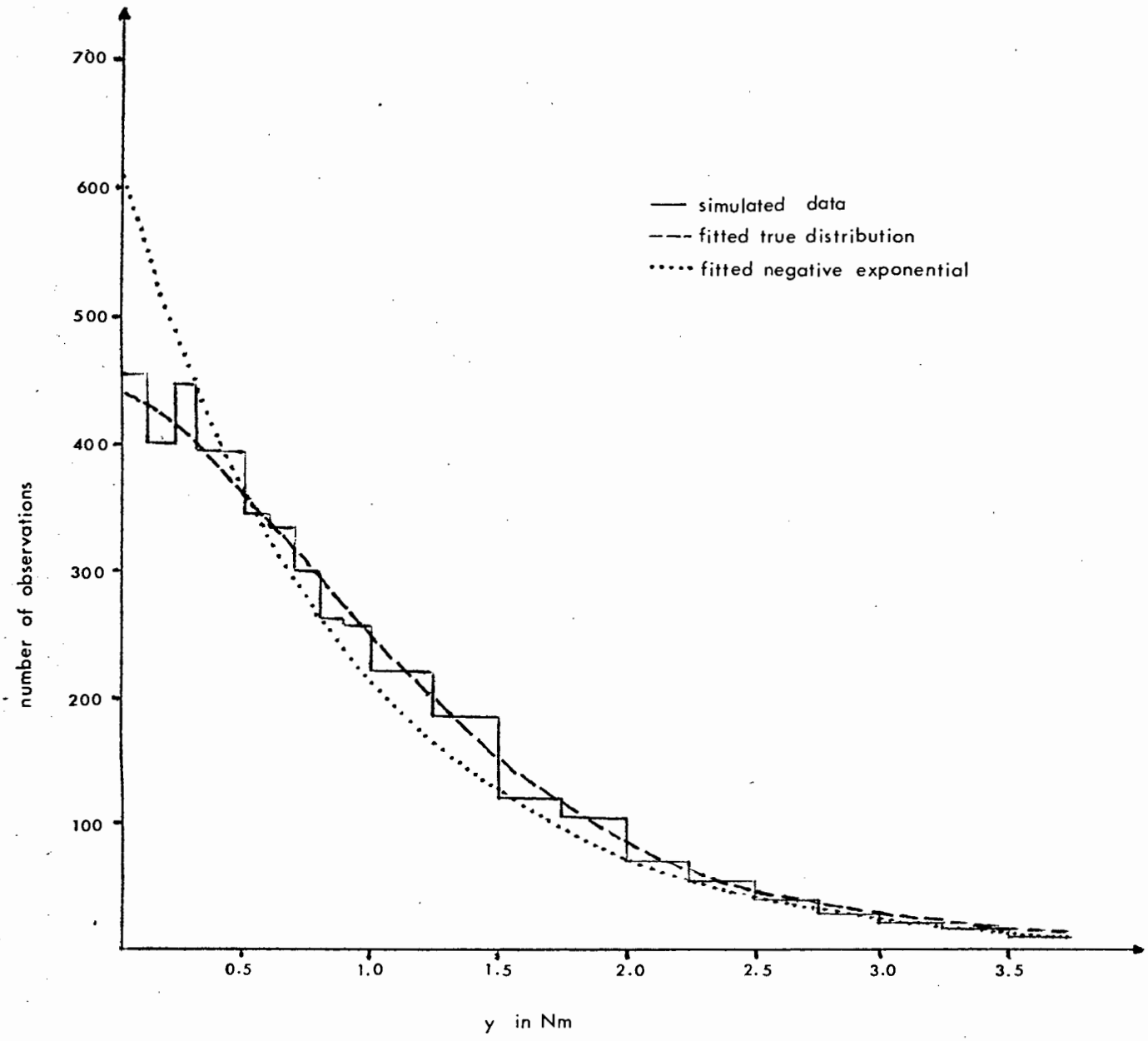


Figure 7-3.

Negative exponential and true distribution fitted to simulated data for v=12



7.3. Calculating g(0) for h(θ) normally distributed

In the following section the more realistic form for h(θ), namely the normal distribution with standard deviation 30 degrees is used.

Once again a g(0) value is calculated experimentally. Recall that the theoretical g(0) value is still given by eq.(V-1), appendix V. Results for cases (i) and (ii), each for u=0 and u=3 knots are presented in table 7-4 below. The negative exponential density estimates are obtained by calculating the mean density estimates for 5 runs done under identical conditions, only changing the random number seeds. Note that no theoretical value of g(0) is given in the case where u=3. The reason for this is that g(0) requires careful definition when whales are non-stationary and the expression of g(0) for u=0 is not necessarily applicable. In table 7-4, CD indicates the NE estimate corrected using the calculated g(0) and TCD the same NE estimate corrected using the theoretical g(0).

Table 7-4. Results : experimental g(0)

Case (i) $\mu = 23.0$ $\lambda = 1.5$		
Whale speed	u=0	u=3
D_7 / D_{12}	1.38	1.36
Calculated $g_{12}(0)$	0.52	0.54
Theoretical g(0)	0.72	-
$CD_{12} = D_{12} / g_{12}(0)$	4351.88	4327.07
$TCD = D_{12} / 0.72$	3245.30	-
True D	3015.11	3015.11
CD_{12} /True D	1.44	1.44
TCD /True D	1.08	-

Case (ii) $\mu = 5.0$ $\lambda = 1.5$		
Whale speed	u=0	u=3
D_7 / D_{12}	1.53	1.61
Calculated $g_{12}(0)$	0.30	0.17
Theoretical g(0)	0.24	-
$CD_{12} = D_{12} / g_{12}(0)$	2343.83	3969.00
$TCD = D_{12} / 0.24$	2929.79	-
True D	3015.11	3015.11
CD_{12} /True D	0.78	1.32
TCD /True D	0.97	-

In case (i) the experimental g(0) (table 7-4 above) is too small,

leading to the corrected estimate CD_{12} being an overestimate of the true density. This happens because D_7 slightly overestimates and D_{12} greatly underestimates the true density in both the cases u=0 and u=3 (see table 7-5). Once again this is mainly due to the fact that the negative exponential function does not fit the y-data very well, especially for small values of y (see figure 7-4). As in the case where $h(\theta)$ is the step function, the histograms flatten out for small values of y, though not to the same degree. This is apparent when the intervals in this area (y<1.0) are about 0.1 Nm or smaller.

In case (ii) where u=0, the experimental g(0) is too big, leading to an underestimate of the true density. However, where u=3, g(0) once again leads to an overestimate of true density. Here (for u=0 and u=3) both D_7 and D_{12} greatly underestimate the true density. It is not quite clear exactly why $g_{12}(0)$ for u=0 is too large, whereas in all the other cases it is too small. Further comment will be made in section 7.5.

When comparing the experimentally obtained g(0) values to the theoretical values, it is important to remember that one would not necessarily expect them to be very close, since the shape factor has not yet been extracted from the experimental values. Recall the test case (section 7.2) where it was possible to extract the shape factor and where two different sets of g(0) values were obtained, depending on whether the density estimates had been corrected for shape or not. In this case ($h(\theta)$ the normal distribution) a shape factor has not been calculated since a relatively simple analytical expression for g(y), necessary for computing a shape factor, has not been determined.

7.4. Density estimates

Table 7-4 shows the corrected density estimates (CD_{12}) obtained by using the experimental $g_{12}(0)$ values given in the same table. The case (i) corrected estimates for u=0 and u=3 both greatly overestimate the true density. The ratio of the corrected estimate to the true density shows the degree of overestimation (approximately 44%).

The case (ii) corrected estimates, however, show an underestimate for u=0 and an overestimate for u=3. The underestimate obviously results from the very large $g_{12}(0)$ value that was used. Once again the ratio defined above reveals the degree of over/underestimation (approximately 20 to 30%). It is interesting to note that the NE estimates corrected using the theoretical g(0) (i.e. only for u=0) are reasonable estimates of the true density.

7.5. The effect of whale movement

The ratio $D_V(3)/D_V(0)$ gives an idea of the effect of whale movement on the NE estimates. The uncorrected estimates are used in order to keep the effect of the incorrect g(0) value separate from the movement effect. The ratios for both cases (i) and (ii) are presented in the last column of table 7-5. For each case the ratios for vessel speed v=7 and v=12 are given. The ratios D_V/True D simply indicate the degree of over- or underestimation of the uncorrected NE estimates.

Table 7-5. Relevant ratios

	u=0	u=3	D (3)/D (0)
Case (i) $\mu = 23.0$ $\lambda = 1.5$			
D_7	3122.17	3177.77	1.02
D_7/True D	1.04	1.05	
D_{12}	2262.98	2336.62	1.03
D_{12}/True D	0.75	0.78	
Case (ii) $\mu = 5.0$ $\lambda = 1.5$			
D_7	1073.58	1083.52	1.01
D_7/True D	0.36	0.36	
D_{12}	703.15	674.73	0.96
D_{12}/True D	0.23	0.22	

In case (i) both ratios show that the $D_v(3)$ value is slightly greater than the $D_v(0)$ value. In case (ii), the first ratio (i.e. v=7) indicates the same effect. However, the second ratio (v=12) shows that $D_{12}(3)$ is smaller than $D_{12}(0)$. This is contradictory to what one expects, since it implies that random whale movement leads to an underestimate of the density. The standard deviation s , of the density estimate is large and it is probably simply the effect of random fluctuation that caused $D_{12}(0)$ to be greater than $D_{12}(3)$ in case (ii).

It is also important to remember that the D values are the means of very small sets of observations (only 5 observations were used

to calculate each D value and its standard deviation). Using larger data sets might easily rid us of this problem. Cases where u=10 for example, can be investigated since one expects the movement effect to be greater for increased u. However, since the swimming speed of minke whales is estimated at about 3 knots and since this is the case we are interested in, it is important that the variance problem be solved.

Although it is not possible to attach a reliable figure to the percentage overestimation resulting from random whale movement at 3 knots, table 7-5 shows very clearly that compared to the effect of using an incorrect g(0) value (see table 7-4), the effect of whale movement is very small.

Table 7-6. Measures of variation

Case (i) $\mu=23.0$ $\lambda=1.5$		
Whale speed	u=0	u=3
D_7	3122.17	3177.77
std.dev. s	84.11	60.63
CV	37.60	27.11
SEM	0.03	0.02
D_{12}	2262.98	2336.62
std.dev. s	12.51	89.38
CV	5.59	39.97
SEM	0.01	0.04

Case (ii) $\mu=5.0$ $\lambda=1.5$		
Whale speed	u=0	u=3
D_7	1073.58	1083.52
std.dev. s	47.86	28.95
CV	23.93	14.48
SEM	0.04	0.03
D_{12}	703.15	674.73
std.dev. s	38.15	18.03
CV	17.06	8.06
SEM	0.05	0.03

std.dev. — standard deviation
CV — coefficient of variance
SEM — standard error of the mean

Figure 7-4.

Negative exponential fitted to simulated y-data

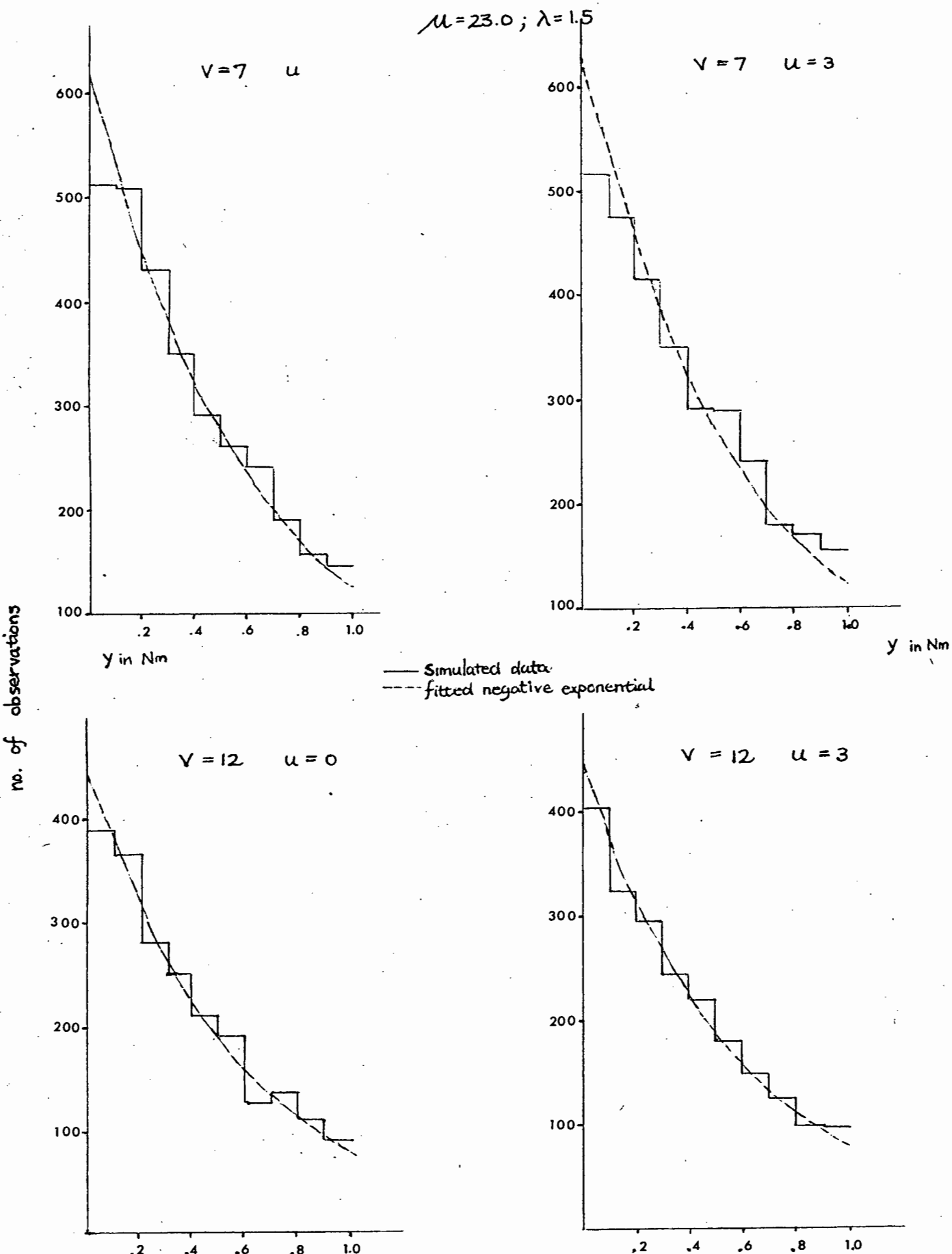
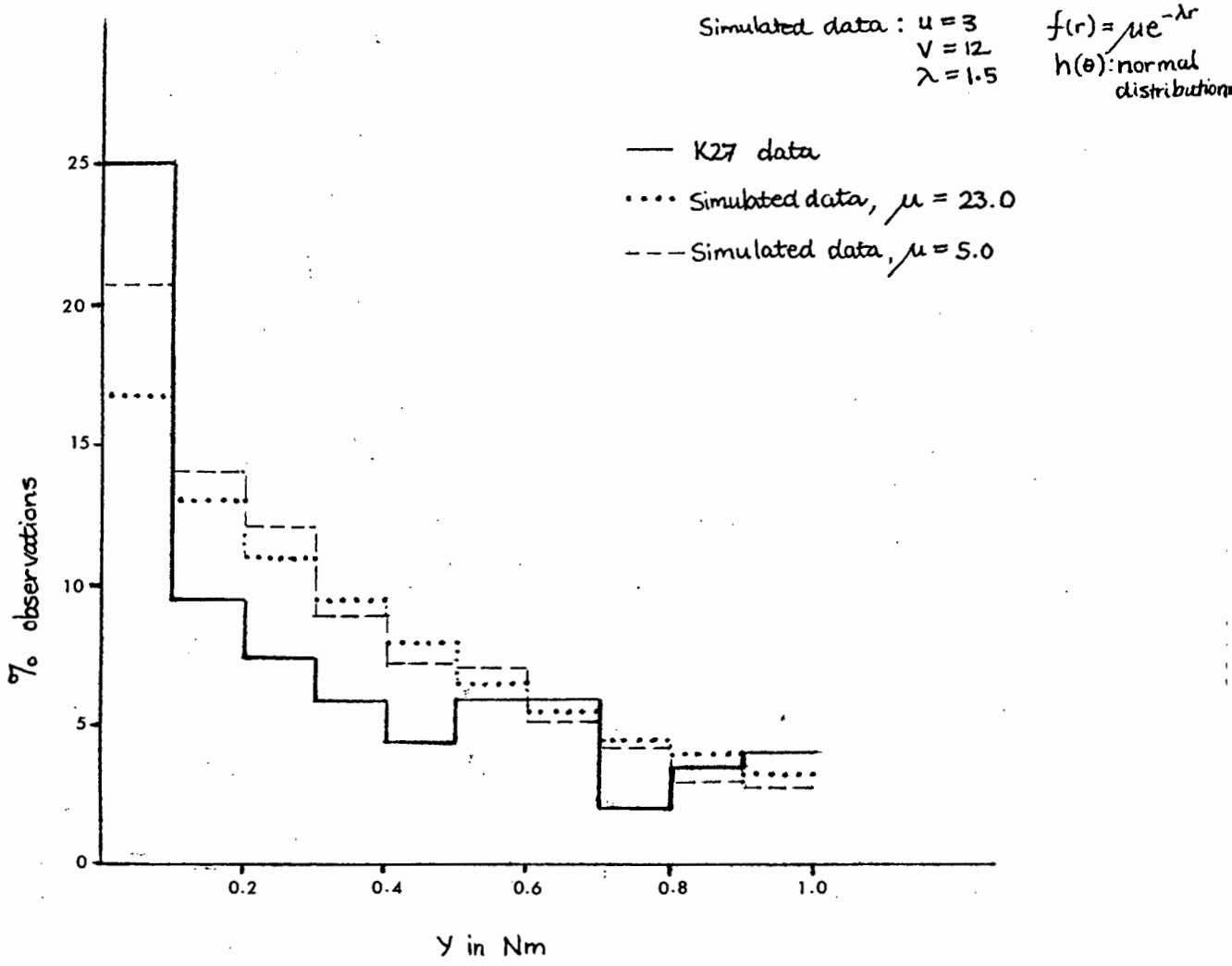


Figure 7-5.

Comparing K27 and simulated y-data



8. DISCUSSION AND CONCLUSION

The aim of this study, as set out in chapter 2, was to attempt to assess the bias in density estimates due to random whale movement. Simulated data were used to duplicate the estimation procedure followed for real data. Density estimates obtained from the simulated data were compared to the true density that is known from the simulation.

8.1. Discussion

As mentioned in the previous chapter, results from five runs, for each relevant case were used to calculate a mean negative exponential density estimate and standard deviation. The standard deviations, coefficients of variance and standard error of the means (table 7-6) show the degree of variation observed in the estimates. Clearly, if the effect of whale movement creates a bias that is smaller than or of the same order as the observed standard error of the mean, the ratio $D_v(3)/D_v(0)$ will not be a reliable estimate of the bias and it might even be impossible to detect the movement effect. This seems to be the case here. From tables 7-5 and 7-6 it is clear that both the standard error of the mean and the bias effect range from about 1 to 5 percent. Note that in case (ii) for v=12 (table 7-5) the observed bias is negative instead of positive as one would expect. This is discussed in section 7.5.

At this stage the only conclusion that can be drawn is that for the particular form of the sighting function used in this study, the negative exponential density estimates do not show a

significantly greater bias than the 1.04 (i.e. 4%) estimate currently used as correction factor in calculating real estimates (Butterworth and Best 1982). Note that this 4% correction factor is based on the fixed detection radius model with vessel speed 12 knots and whale speed 5 knots.

It is, however, disturbing to note how badly the corrected estimates, CD_{12}, overestimate the true density (table 7-4). The main reason for this was shown to be the effect of using an incorrect experimental g(0) value which resulted from the negative exponential function fitting the right-angle distance data badly. Most important is the fact that the g(0) effect is an order of magnitude greater than the whale movement effect.

A first step towards more reliable estimates of the bias caused by g(0) and whale motion would be the use of much larger data sets in calculating the mean density estimates. However, it would probably be wiser and more useful first to obtain greater resemblance between simulated and real data. This can be done by finding a better functional form for the sighting function and/or more accurate values for the parameters μ and λ.

Apart from implementing the second negative exponential form for f(r) (form 2.2, Appendix III), a combination of the two forms as given in form 3.1, Appendix III, can be used. It is known that a peculiarity of minke whale sighting data is the prominent peak at small y values in the y distribution. Even with the use of the normally distributed $h(\theta)$, the sighting function (form 1.2, Appendix III) cannot produce such a peak. One might be forced to use somewhat less realistic forms of f(r) and especially $h(\theta)$ so as to force such a peak in the simulated data. (Use of an $h(\theta)$ with a smaller standard deviation for example, even if this is not

supported by the data of Doi et al. (1982).) The peak at small y values is of great importance since the intercept g(0), of the functional form fitted to the observed y-data is crucial in the calculation of the density estimate. (See fig. 7-5)

The main problem that arises from attempts to find suitable values of the parameters μ and λ is the great variance observed in the measures of error. A great number of very long runs (with total no. of whales N=2 500 000 and corresponding CPU-time of about 500 minutes!) might solve this problem, but the use of variance reduction techniques may prove more useful and economical. Although use of a smaller value for the time increment dt, greatly increases the CPU-time, this might also prove useful since it implies a better approximation of the differential equation, eq.(3-8) (see section 3.5).

It has already been pointed out that the overestimation resulting from use of too small a g(0) value is due to the bad fit of the negative exponential function to the simulated data. The negative exponential function might fit the real data slightly better than it fits the simulated data, because of the greater peak in the y distribution. If this is the case, the g(0) effect might not be as substantial as results seem to indicate here. This will obviously come to light when simulated data that resemble the real data more closely, are generated. It is, however, a good idea also to investigate the behaviour of other estimators. Both the g(0) problem and the whale motion problem could benefit from such an investigation.

A popular approach that can be followed here, is fitting a Fourier series to the y-data and using the corresponding Fourier series estimator. This nonparametric approach leads to results that are

said to be quite good when the underlying g(y) is concave, nearly linear or mildly convex (Burnham et al. 1980, p66.). Since the real y-data seem to indicate an underlying g(y) that is perhaps more than 'mildly' convex, care will have to be taken when this procedure is applied. Fortunately, the simulation model can easily be used to determine how good the Fourier series estimator is under given circumstances.

As mentioned in chapter 2, the simulation model leaves room for improvement, notably the introduction of sighting cues or a given blow rate (see Hiby 1982) as opposed to potential sightability at all times. The effects of dive times that are not Poisson distributed can be investigated and random walk paths instead of straight line paths can be considered. However, I believe that further research should first be conducted on this simple model before an attempt is made to study a more realistic model.

8.2. Conclusion

Although results from this study cannot be considered conclusive, there is a very definite warning that the procedure of determining g(0) experimentally is under question. The amount of overestimation resulting from use of an incorrect g(0) factor is an order of magnitude greater than that resulting from whale movement at 3 knots. At this stage it is therefore most important to concentrate further research on methods of calculating g(0) rather than on effects of target motion.

1. APPENDIX I

Derivation of the Koopman ratio.

Consider an observer moving with constant velocity v and with a fixed detection radius R, among a uniform random distribution of targets. The targets move with constant speed u, but random direction. Assume that there are on average N targets per square mile. Because the track angles ϕ (angle between track and true direction of target) are uniformly distributed and independent of position, the average number of targets with track angle between ϕ and $\phi+d\phi$ will be $N.d\phi/2\pi$. A target with particular track angle ϕ, moving at velocity $\underline{w}$ relative to the observer has to be in the shaded region of figure 1-1 (Appendix I) if it is to enter the circle in a unit of time (one hour). Since the area of the shaded region is 2Rw, the number of targets of track angle between ϕ and $\phi + d\phi$ that enter the detection circle per unit time is, to quantities of first order in the differential $2RW.Nd\phi/2\pi$

The total number N_t is now given by integration over ϕ

$$N_t = \frac{RN}{\pi}\int_0^{2\pi} w.d\phi$$

since $w = (u^2 + v^2 - 2uv.\cos\phi)^{1/2}$

$$N_t = \frac{RN}{\pi}\int_0^{2\pi} (u^2+v^2-2uv\cos\phi)^{1/2}\, d\phi$$

$$= \frac{RN}{\pi}\int_0^{2\pi} \left[(u+v)^2 - 2uv(1+\cos\phi)\right]^{1/2} d\phi$$

$$= \frac{RN}{\pi}\int_0^{2\pi} (u+v)\left[1 - \frac{2uv}{(u+v)^2}(1+\cos\phi)\right]^{1/2} d\phi$$

Let $\psi = \frac{\pi - \phi}{2}$ i.e. $\phi = \pi - 2\psi$

and $1 + \cos\phi = 1 + \cos(\pi - 2\psi)$

$= 2\sin^2\psi$

Now

$$N_t = \frac{RN}{\pi}(u+v)\int_{-\pi/2}^{\pi/2}\left(1 - \frac{4uv}{(u+v)^2}\sin^2\psi\right)^{1/2}.2.d\psi$$

Let $\frac{4uv}{(u+v)^2} = \sin^2\sigma$ i.e. $\frac{2\sqrt{uv}}{(u+v)} = \sin\sigma$

then

$$N_t = \frac{4RN}{\pi}(u+v)\int_0^{\pi/2}(1 - \sin^2\sigma.\sin^2\psi)^{1/2}\,d\psi$$

$$N_t = \frac{4RN}{\pi}(u+v)\,E(\sigma)$$

where $E(\sigma)$ is the complete elliptic integral of the second kind.

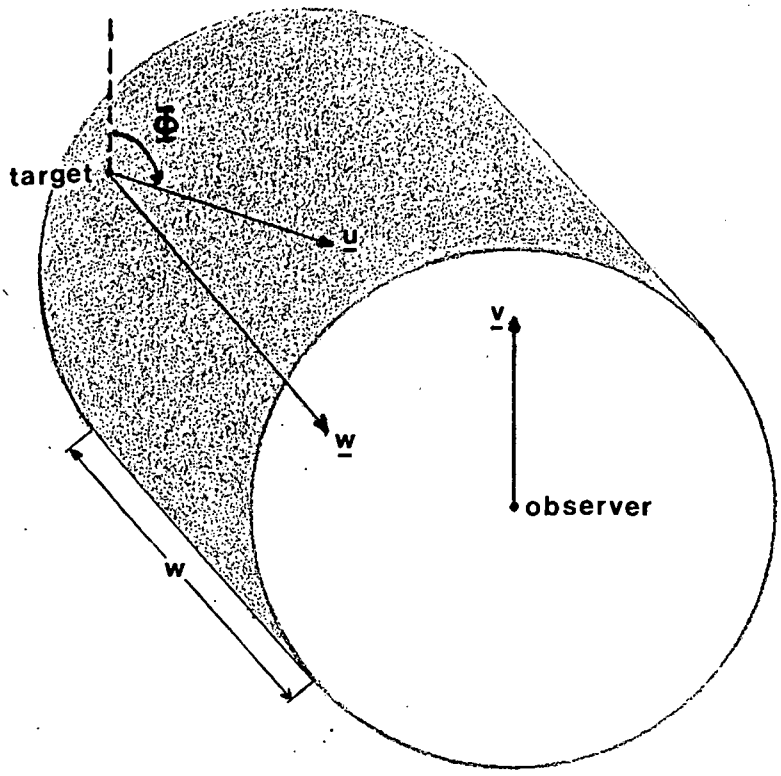
When u=0, N_t simplifies to:

$$N_t = 2RNv$$

and now the Koopman ratio is given by:

$$\frac{N_t}{N_t(u=0)} = \frac{2}{\pi}.\frac{u+v}{v}E(\sigma) \qquad \text{(I-1)}$$

Figure 1-1. (Appendix I)



2. APPENDIX II

Expected value and variance of the Koopman ratio, for u=0

The measure used for the KR is:

$$KR = \frac{nA}{Na}$$

(See section 5.1) In terms of the simulation N,A and a are fixed values and n is the only variable. Therefore, determining the expected value of the KR simplifies to determining the expected value of n.

$$E(KR) = E\left(\frac{nA}{Na}\right)$$

$$= \frac{A}{Na} E(n)$$

Similarly

$$Var\,(KR) = Var\left(\frac{nA}{Na}\right)$$

$$= \left(\frac{A}{Na}\right)^2 Var\,(n)$$

Note that only the fixed detection range sighting criterion for the case where targets are stationary is considered here. In this case whales move in straight lines parallell to the track line at velocity -v relative to the vessel and therefore a whale will be sighted only if generated within a distance R either side of the track. R is the radius of the detection circle.

Since the starting points are generated according to the uniform distribution on the interval (-Wmax,Wmax), the probability of a whale being sighted, which is equal to the probability of a whale

having starting position in the section (-R,R), is given by:

$$q = R/Wmax$$

Let X_i be the random variable associated with the i^{th} whale, such that:

$$x_i = \begin{cases} 1 \text{ if the } i^{th} \text{ whale is sighted} \\ 0 \text{ if the } i^{th} \text{ whale is not sighted} \end{cases}$$

Now $P(X_i=1) = q$

$P(X_i=0) = 1-q$

If the total number of whales that pass through the system is N, it is clear that the number of whales that are sighted, n, is given by:

$$n = \sum_{i=1}^{N} X_i$$

Since the sighting of one whale is independent of the sighting of any other whale, the X_i variants are independent. The variants are also identically distributed.

Now E(n) and Var(n) are given by:

$$E(n) = N E(X_i)$$

$$Var(n) = N Var(X_i)$$

It is easy to show that:

$$E(X_i) = q$$

$$Var(X_i) = q - q^2$$

Therefore $E(n) = Nq$

$Var(n) = N(q-q^2)$

and since q=R/Wmax=a/A

$$E(KR) = \frac{A}{Na} \cdot Nq = 1$$

Note that this is what one would expect, since the theoretical value of the KR for this specific case is unity.

$$Var(KR) = \left(\frac{A}{Na}\right)^2 N(q-q^2)$$

$$Var(KR) = \frac{1-q}{Nq} \qquad \text{(II-1)}$$

It is clear from eq. (II-1) that as N increases, the variance of the KR decreases. The variance was calculated for a few values of N, using eq. (II-1) and q=0.302 as used in the simulation. These values together with the observed standard deviations squared are presented in table 2-1 below.

Table 2-1.

N	1 000	10 000	100 000
Var(KR) calculated	0.00232	0.00023	0.00002
Observed s^2	0.00201	0.00018	0.00003
D^2 calculated	22.52	9.39	16.50
χ^2 critical	40.65	21.92	20.48
d.o.f	25	11	10

The calculated D^2-values are for testing the null hypothesis Ho: σ^2=var(KR) (calculated), against a two-sided alternative. At the 5% significance level Ho is accepted in all three cases.

3. APPENDIX III

Table of sighting functions.

FORM NO.	SIGHTING FUNCTION
1.1	$f(r) = \mu e^{-\lambda r}$ $h(\theta) = \begin{cases} 1 & 0 \le \theta \le \pi/2 \\ 0 & \pi/2 < \theta < \pi \end{cases}$
1.2	$f(r) = \frac{\mu}{r} e^{-\lambda r}$ $h(\theta) = \begin{cases} 1 & 0 \le \theta \le \pi/2 \\ 0 & \pi/2 < \theta < \pi \end{cases}$
2.1	$f(r) = \mu e^{-\lambda r}$ $h(\theta) = \begin{cases} e^{-1/2(\theta/\sigma)^2} & 0 \le \theta \le \pi/2 \\ 0 & \pi/2 < \theta < \pi \end{cases}$
2.2	$f(r) = \frac{\mu}{r} e^{-\lambda r}$ $h(\theta) = \begin{cases} e^{-1/2(\theta/\sigma)^2} & 0 \le \theta \le \pi/2 \\ 0 & \pi/2 < \theta < \pi \end{cases}$
3.1	$f(r) = \begin{cases} \mu e^{-\lambda r} & 0 \le r \le R \ (R>0) \\ \frac{\mu}{r} e^{-\lambda r} & r > R \end{cases}$

4. APPENDIX IV

Real data used in section 6.2.

The summary of data given below is that of the Kyo Maru 27 (K27) from the 1980/81 southern hemisphere minke whale assessment cruise.

Data of the radial distances r in Nautical miles are presented below. Instead of the exact number of observations, the percentage of observations within a given interval is given. Note that all the intervals are not of equal length.

Interval in Nm.	% Observations
0.00 - <0.25	14.68
0.25 - <0.50	11.51
0.50 - <0.75	10.71
0.75 - <1.25	14.69
1.25 - <1.75	19.84
1.75 - <2.25	12.30
2.25 - <2.75	7.94
2.75 - <3.25	5.95
3.25 - <5.00	1.99

Frequencies of observed θ (sighting angle) are given below as percentages of all sightings.

Sighting angle	% Observations
0	21.93
10	18.42
20	16.23
30	13.16
40	14.04
50	5.70
60	2.63
70	6.14
80	1.75

5. APPENDIX V

An expression for g(0)

From Butterworth (1982) we have the following expression for g(y):

$$g(y) = 1 - e^{-\int_0^\infty F(x,y)dx}$$

where $F(x,y) = f(r).h(\theta).\frac{1}{V}$ and $f(r)$ and $h(\theta)$ can take on various forms. In particular for $y=0$

$$g(o) = 1 - e^{-\int_0^\infty F(x,o)dx}$$

$$g(o) = 1 - e^{-\int_0^\infty f(r).h(o)\frac{dx}{V}}$$

In the case where:

$$f(r) = \mu e^{-\lambda r}$$

$$h(\theta) = \begin{cases} e^{-\frac{1}{2}(\theta/\sigma)^2} & 0 \le \theta \le \pi/2 \\ 0 & \pi/2 < \theta < \pi \end{cases}$$

it is clear that $h(o)=1$ and therefore

$$g(o) = 1 - e^{-\frac{1}{V}\int_0^\infty f(r)dx}$$

since $\frac{1}{V}\int_0^\infty f(r)dx = \frac{\mu}{V}\int_0^\infty e^{-\lambda r}dr$ $\quad (y=0 \rightarrow x=r \;,\; \rightarrow dx=dr)$

$$= \frac{\mu}{\lambda V}$$

now

$$g(o) = 1 - e^{-\mu/\lambda V} \qquad \text{(V-1)}$$

Note that the above expression for g(0) is also valid when h(θ) is the step function as defined in Appendix III, form 1.1.

6. APPENDIX VI

Determining g(0) experimentally.

From Butterworth et al. (1982) the following expression is obtained:

$$\frac{D_v}{D_{12}} = \frac{g_v(0)}{g_{12}(0)} \qquad \text{(VI-1)}$$

where D_v and D_{12} indicate the line-transect estimates of density at vessel speed v and 12 respectively and $g_v(0)$ and $g_{12}(0)$ the value of the detection function on the trackline, i.e. for y=0.

Using the hazard-rate model of Butterworth (1982) with the functions f(r) and h(θ) as defined in form 1.1 (appendix III) and the corresponding expression for g(0) as determined in Appendix II, eq. (VI-1) can be rewritten as:

$$\frac{D_v}{D_{12}} = \frac{g_v(0)}{g_{12}(0)} = \frac{1 - e^{-\alpha/v}}{1 - e^{-\alpha/12}} \qquad \text{(VI-2)}$$

An α value can be obtained from the estimate of relative density through eq. VI-2. Note, however, that α is assumed to be independent of vessel speed v. The estimate of g(0) at the normal searching speed of 12 knots, is calculated as:

$$g_{12}(0) = 1 - e^{-\alpha/12} \qquad \text{(VI-3)}$$

(See Butterworth et al. 1982 for a more detailed derivation).

7. APPENDIX VII

An expression for the true density.

A population density measure is given by the number of animals in a given area, divided by the area. In terms of the simulation model, the true density will therefore be given by:

$$D=N/2WmaxL$$

where N - the total no. of whales passing through the system
2Wmax - the length over which whales are generated
L - the tracklength

(i.e. 2WmaxL is the total area 'containing' the N whales). It is, however, not possible to determine L as such since the simulation model handles one whale at a time and runs till all N whales have passed through the system.

Instead of considering D, we therefore consider DxL. Note that the tracklength used in obtaining the estimated density is the same L as the one above. Therefore, let the true density be given by:

$$\text{True } D = DL$$

$$\text{True } D = N/2Wmax \qquad \text{(VII-1)}$$

8. APPENDIX VIII

The negative exponential density estimate.

Let the detection function g(y) be of the form:

$$g(y) = e^{-by}$$

where b is a positive parameter to be determined from the sightings data. It can be shown (Gates et al. 1968) that the maximum likelihood estimate of b is given by:

$$\hat{b} = \frac{n-1}{\sum_{i=1}^{n} y_i} \qquad \text{(VIII-1)}$$

where n is the number of animals sighted and the y_i are the observed right angle distances.

The corresponding density estimate, d say, is given by:

$$d = \frac{n\hat{b}}{2L}$$

where L is the length of the trackline. Substituting eq. (VIII-1) we obtain:

$$d = \frac{n(n-1)}{2L \Sigma y_i} = \frac{(n-1)}{2L\bar{y}}$$

Since L is unknown in the simulation, dL will be considered instead of d and $D_v(u)$ (as defined in Chapter 7) is given by:

$$D_v(u) = \frac{(n-1)}{2\bar{y}} \qquad \text{(VIII-2)}$$

where the v and u simply indicate the respective vessel and whale

speeds under which the estimate is obtained.

9. APPENDIX IX

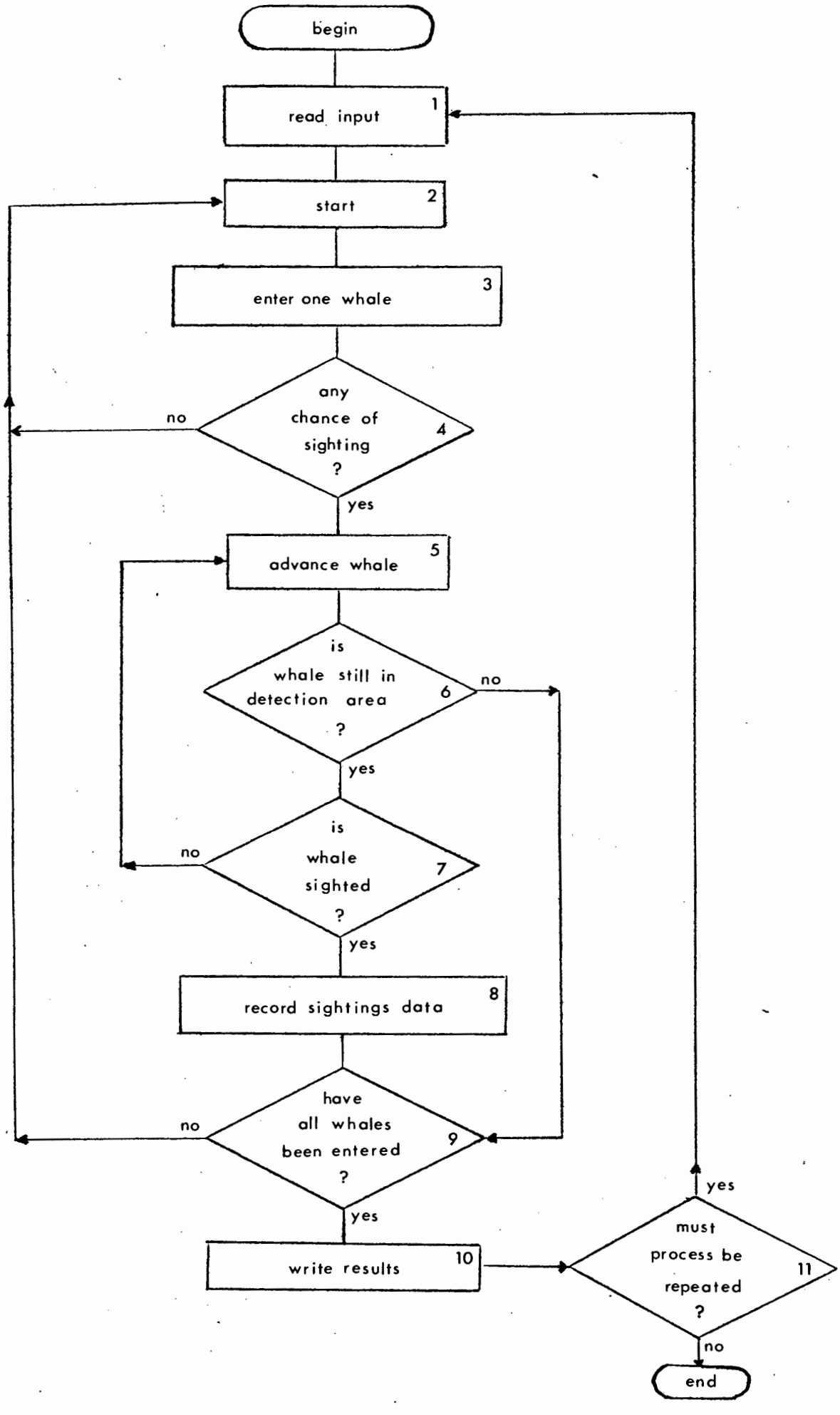
Program documentation

The FORTRAN names of the most relevant variables in the simulation model are listed below.

Variable	FORTRAN equivalent	
N	NTOT	total no. of whales entering system within 1 sampling
n	COUNT	no. of whales counted i.e. sighted
u	VW	absolute whale speed
v	VP	absolute vessel or platform speed
Wx	WX	relative whale speed, x component
Wy	WY	relative whale speed, y component
dt	DELT	time increment
λ	LAMBDA	parameter in sighting rate function
μ	MEW	parameter in sighting rate function
R	RR	radius of detection circle
r	RR	radial distance from whale to vessel, at sighting
θ	TH	sighting angle
y	Y	perpendicular distance from track
Wmax	WMAX	half the length of the generating line
α_{max}	AMAX	the maximum limiting angle

9.1. Flowchart of the program.

A flowchart illustrating the structure of the program is presented below.


begin
read input
1
start
2
enter one whale
3
any chance of sighting ?
4
no
yes
advance whale
5
is whale still in detection area ?
6
no
yes
is whale sighted ?
7
no
yes
record sightings data
8
have all whales been entered ?
9
no
yes
write results
10
must process be repeated ?
11
yes
no
end

9.2. Description of flowchart blocks

1. Read input.

Some variables are initialised in the program itself because they are kept fixed through most of the runs. The variables that are read in are:

Variable	FORTRAN input format
NTOT	I8
DELT	F7.5
VW	F4.2
N	I6
NN	I6
NSAMPL	I3
LAMBDA	F5.2
MEW	F5.2
NF	I1

2. Start.

Set the relevant data vriables to zero, e.g. COUNT, YSUM and the frequency counters of the arrays called through the subroutine HISTO: RBIN, YBIN and THBIN.

3. Enter a whale.

In this section a starting position and direction of movement are determined. The subroutine THETAA is called to determine the direction if whale speed is non-zero. Subsequently the relative whale velocity is calculated as well as the increments in the x and y directions (i.e. dx and dy). These increments are used to calculate the new coordinates of position after every time increment dt.

whole process is repeated from block 3.

10. Output results.

An example of output is given below. Apart from the relevant variables and parameters e.g. VW,VP DELT etc. that identifies the run, grouped y and r values as well as grouped values are given. It is also necessary to give grouped values rounded to 5 and 10 degrees respectively. The total number of whales that were sighted i.e. COUNT as well as NTOT, YSUM and YBAR are given.

11. Should process be repeated?

The variable NSAMPL indicates how many times the whole procedure should be repeated. Once again a counter is increased by 1 every time the procedure is started. If the counter is equal to NSAMPL, the run is completed. Note that a run consists of NSAMPL samples based on the same input values.

Some notes.

For very long runs (e.g. NTOT=1 000 000) the program was changed slightly so that cumulative data was output NSAMPL times. In this case one run consisted of 1 sample only and block 2 is not executed. A somewhat simplified model was used for the validation process. The output also varied as necessary.

9.3. Program listing.

A listing of the FORTRAN program as used to obtain the final results is presented below.

```
1234567890123456789012345678901234567890123456789012345678901234567890123
MINKE WHALE SIMULATION STUDY LISTING

C **** MINKE WHALE SIMULATION STUDY
C **********************************
C
C **** SIGHTING FUNCTION:
C
C       NEGATIVE EXPONENTIAL F(R) * NORMAL H(TH)
C       FORM 1: F(R)=MEW*EXP(-LAMBDA*R)
C       FORM 2: F(R)=(MEW/R)*EXP(-LAMBDA*R)
C
C **** MOST RELEVANT VARIABLES:
C
C       VP      =PLATFORM VELOCITY IN KNOTS
C       VW      =WHALE VELOCITY IN KNOTS
C       WX      =RELATIVE WHALE VEL. X COMPONENT
C       WY      =RELATIVE WHALE VEL. Y COMPONENT
C       X       =X COORDINATE OF POSITION
C       Y       =Y COORDINATE OF POSITION
C       DELT    =TIME STEP (INCREMENT) AS FRACTION OF AN HOUR
C       NTOT    =TOTAL NO. OF WHALES ENTERING SYSTEM WITHIN RUN
C       COUNT   =NO. OF WHALES SIGHTED (RECORDED)
C       THETA   =DIRECTION OF WHALE MOVEMENT  (SEA FRAME)
C       PHI     =RELATIVE DIRECTION OF WHALE MOVEMENT (PLATFORM FRAME)
C       R       =RADIUS OF DETECTION CIRCLE (CUT-OFF)
C       NF      =NO. OF THE FORM F(R) SHOULD TAKE ON
C
C **** DEFINING SOME VARIABLES
C
        INTEGER COUNT,YBIN(27),RBIN(27),THBIN(9)
        INTEGER TSBIN(19),TIOBIN(10),NF
        REAL VW,VP,PI,DELT,X,Y,WX,WY,DX,DY,THETA,WMAX,PHI,ALPHA
        REAL LAMBDA,MEW,YBAR
        COMMON PI,VW,VP,NN,YL
        DATA VP/12.0/,PI/3.141592654/,R/5.0/,DT/0.00001/
        DATA LAMBDA/0.5/,MEW/5.0/,YL/0.25/,NS/102030/
        DATA YBIN,RBIN/27*0,27*0/
C
C **** READING IN DATA
C
        READ(5,125)NTOT,DELT,VW,N,NN,NSAMPL,LAMBDA,MEW,NF
125     FORMAT(I8,F7.5,F4.1,I6,I6,I3,F5.2,F5.2,I1)
C
C **** INITIALIZING SOME VARIABLES
C
        IF(NF.EQ.0)GO TO 999
        COUNT=0
        PI2=PI/2
        AMAX=ASIN(10.0/12.0)
        WMAX=(R/COS(AMAX))*(1+SIN(AMAX))
        ALPHA=ASIN(VW/VP)
        YMAX=R*TAN((PI2+ALPHA)/2.0)
C
C **** WRITING HEADING AND IDENTIFYING PARAMETERS
C
        WRITE(15,150)
150     FORMAT('MINKE SIMULATION STUDY'//'************************'//)
        WRITE(15,250)R
250     FORMAT('SIGHTING : NEG EXPONENTIAL SIGHTING RATE, R=',F6.3)
        IF(NF.EQ.1)THEN
          WRITE(15,255)
 255      FORMAT('FUNCTION 1 : MEW*EXP(-LAMBDA*RR)'/)
        ELSE
          WRITE(15,256)
```

```
650   256    FORMAT('FUNCTION 2 : (MEW/RR)*EXP(-LAMBDA*RR)'/)
660          ENDIF
661          WRITE(15,75)LAMBDA,MEA
662   75     FORMAT(' LAMBDA=',F5.2,4X,'MEW=',F5.2/)
670          WRITE(15,350)VW,VP,DELT
680   350    FORMAT(' VW=',F6.3,' VP=',F6.3,' DELT=',F7.5/)
690          WRITE(15,450)NTOT,N,NN
700   450    FORMAT('MAX=',I8,'  N=',I6,'   NN=',I6/)
710          WRITE(15,550)XMAX,YMAX
720   550    FORMAT('XMAX=',F8.5,'   YMAX=',F8.5)
730   C
740   C  **** SET RELEVANT COUNTERS TO ZERO
750   C
760          YSUM=0.0
790          RSUM=0.0
800          DO 100 I=1,21
810            YBIN(I)=0
820            RBIN(I)=0
830            IF(I.LE.9)THBIN(I)=0
840   100    CONTINUE
850   C
851   C  **** SAMPLING DO LOOP
852   C
853   C
855          NII=NTOT/NSAMPL
856          DO 2000 NSAM=1,NSAMPL
860   C
890   C
900   C  **** MAIN DO LOOP (ENTERING WHALES)
910   C
920          DO 1000 ITERAT=1,NII
930            X=0.0
940            Y=(RANDOM(N)-0.5)*2.0*YMAX
944   C
945   C  **** DETERMINE DIRECTION OF MOVEMENT
946   C
950            IF(VW.EQ.0.0)THEN
960             THETA=RANDOM(NN)*2.0*PI
970            ELSE
980             CALL THETAA(TTHETA,DT)
990             THETA=PI-TTHETA
1000              IF(THETA.LT.0.0)THETA=2.0*PI+THETA
1010            ENDIF
1080  C
1081  C  **** CALCULATE DX AND DY
1082  C
1090            WX=VW*COS(THETA)+VP
1100            WY=VW*SIN(THETA)
1110            DX=WX*DELT
1120            DY=WY*DELT
1130            IF(WY.EQ.0.0)THEN
1140             PHI=0.0
1150            ELSE
1160             PHI=ATAN(WY/WX)
1170            ENDIF
1180            ANGL2=PI2-2.0*ATAN(R/ABS(Y))
1190            ANGL1=-ANGL2
1191  C
1192  C  **** ANY CHANCE OF SIGHTING WHALE?
1193  C
1200            IF(VW.EQ.0.0.AND.ABS(Y).GT.R)GO TO 1000
1210            IF(Y.GT.R.AND.PHI.GT.0.0)GO TO 1000
```

```
1220           IF(Y.LT.-R.AND.PHI.LT.0.0)GO TO 1000
1230           IF(PHI.LT.0.0)PHI=2.0*PI+PHI
1240           PHID=(PHI/(2.0*PI))*360.0
1241  C
1242  C ****  ADVANCE WHALE
1243  C
1250  500      X=X+DX
1260           Y=Y+DY
1270           XX=X-R
1280  C
1290  C ****   IS WHALE SIGHTED OR NOT?
1300  C
1310  C        (APPLICABLE ONLY FOR H(TH)=0 FOR ABS(TH).GT.90 DEG)
1311  C ****  HAS WHALE LEFT DETECTION AREA?
1312  C
1320           IF(XX.GT.0.0)GO TO 1000
1330           SIG=30.0
1340  C
1341  C ****  DETERMINE RADIAL DISTANCE R
1342  C
1350           RADSQ=(XX**2)+(Y**2)
1360           RR=SQRT(RADSQ)
1370           IF(RR.GT.R)GO TO 560
1380           RSQ=R**2
1381  C
1382  C ****  WHICH FORM OF F(R) SHOULD BE USED?
1383  C
1390           IF(NF.EQ.1)THEN
1400             FR=MEW*EXP(-LAMBDA*RR)*DELT
1410           ELSE
1420             IF(RR.EQ.0.0)RR=0.000001
1430             FR=(MEW/RR)*EXP(-LAMBDA*RR)*DELT
1440           ENDIF
1450           TH=ASIN(Y/RR)
1460           TH=(TH/(2.*PI))*360.0
1470           HTH=EXP(-.5*((TH/SIG)**2))
1475  CC       HTH=1
1476  C
1477  C ****  DETERMINE PROBABILITY OF SIGHTING
1478  C
1480           FR=FR*HTH
1490           RANS=RANDOM(NS)
1500           IF(FR.GT.RANS)THEN
1501  C
1502  C ****  WHALE IS SIGHTED!
1503  C
1510             COUNT=COUNT+1
1520             YSUM=YSUM+ABS(Y)
1530             RSUM=RSUM+RR
1540             CALL HISTO(Y,RR,TH)
1580             GO TO 1000
1590           ENDIF
1600  560      CH=2.0*R
1610           IF(X.LT.CH)GO TO 500
1620  1000   CONTINUE
1630  C
1640  C ****  END MAIN DO LOOP
1650  C
1690         THETAD=(THETA/(2.0*PI))*360.0
1700  C
1710  C ****  WRITING RESULTS
1810  C
```

```
1811 C       GROUPED Y AND R VALUES
1812 C
1820        WRITE(15,640)
1830 640    FORMAT(//'  GROUPED :       Y  AND     R VALUES')
1840        AA=0.0
1850        BB=0.1
1860        DO 1650 I=1,27
1870          WRITE(15,645)AA,BB,YBIN(I),RBIN(I)
1880 645      FORMAT(F5.2,'-',F5.2,3X,I8,5X,I8)
1890          AA=BB
1900          IF(I.LT.10)THEN
1901          BB=BB+0.1
1902          ELSE
1903          BB=BB+0.25
1904          ENDIF
1910          IF(I.EQ.26)BB=99999.9
1920 1650   CONTINUE
1930 C
1940 C       ANGLE THETA ROUNDED, GROUPED
1950 C
2040        AA=-5.0
2050        WRITE(15,648)
2060  648   FORMAT(/'OBS THETA ROUNDED TO 5 AND 10 DEG.',4X,'PERCENT'/)
2070        DO 1670 I=1,19
2080          PER=(100.*T5BIN(I))/COUNT
2090          AA=AA+5.0
2100          NT=I-(I/2)*2
2110          J=(I+1)/2
2120          PER2=(100.*T10BIN(J))/COUNT
2130          IF(NT.GT.0)THEN
2140          WRITE(15,649)AA,T5BIN(I),T10BIN(J),PER,PER2
2150  649     FORMAT(F5.1,4X,I8,6X,I8,10X,F6.2,4X,F6.2)
2160          ELSE
2170          WRITE(15,651)AA,T5BIN(I),PER
2180  651     FORMAT(F5.1,4X,I8,24X,F6.2)
2190          ENDIF
2200 1670   CONTINUE
2201 C
2202 C       CUMULATIVE COUNTS
2203 C
2205        NB=NSAM*NII
2210        WRITE(15,650)COUNT,NB,YSUM
2220 650    FORMAT(/' COUNT=',I8,'  N=',I8,6X,'YSUM=',F15.4)
2221  2000 CONTINUE
2222 C
2223 C  **** END SAMPLING DO LOOP
2224 C
2225 C
2226 C       WRITE THETA GROUPED (NOT ROUNDED)
2227 C
2230        WRITE(15,646)
2240 646    FORMAT(/'GROUPED OBSERVED THETA VAL. IN DEG.',4X,'PERCENT'/)
2250        AA=0.
2260        BB=10.0
2270        DO 1660 I=1,9
2280          PER=(100.*THBIN(I))/COUNT
2290          WRITE(15,647)AA,BB,THBIN(I),PER
2300 647      FORMAT(F5.1,' TO<',F5.1,3X,I8,13X,F6.2)
2310          AA=AA+10.
2320          BB=BB+10.
2330 1660   CONTINUE
2331 C
```

MINKE WHALE SIMULATION STUDY LISTING

```
2332 C       OTHER RELEVANT OUTPUT
2333 C
2400         WRITE(15,750)R,WMAX
2410 750     FORMAT(5X,'R=',F6.3,'  WMAX=',F6.3)
2420 C
2430         YBAR=YSUM/COUNT
2440         D=(COUNT-1)/(2*YBAR)
2450         TRUD=NTOT/(2*WMAX)
2460         WRITE(15,860)LAMBDA,MEW,YBAR
2470 860     FORMAT(19X,'LAMBDA=',F5.2,2X,'MEW=',F8.4,4X,'YBAR=',F8.4)
2480         WRITE(15,870)TRUD,D
2490 870     FORMAT(/19X,'TRUE D=',F15.2,4X,'D=(COUNT-1)/2*YBAR=',F15.2)
2500 C
2510 999     STOP
2520 C
2521 C
2522 C
2530 C  **** FUNCTIONS AND SUBROUTINES ****
2531 C
2532 C
2540 C
2550 C  **** RANDOM NUMBER GENERATOR
2560 C
2570         FUNCTION RANDOM(N)
2580         N=N*3125
2590         RANDOM=FLOAT(N)/34359738337.D0
2600         IF (RANDOM.LT.0.0)RANDOM=-RANDOM
2610         RETURN
2620 C
2621 C
2622 C  **** SUBROUTINE DETERMINING DIRECTION OF MOVEMENT
2630 C
2640         SUBROUTINE THETAA(TTHETA,DT)
2650 C
2660         INTEGER I
2670         REAL C1,C2,TJ,TJ1,E
2680         RAN=RANDOM(NN)
2690         C1=VA/VP
2700         C2=2.0*PI*RAN
2710         TJ=C2
2720         DO 1000 I=1,1000
2730           TJ1=TJ-(TJ-C1*SIN(TJ)-C2)/(1-C1*COS(TJ))
2740           E=ABS(TJ1-TJ)
2750           IF(E.LE.DT)GO TO 1500
2760           TJ=TJ1
2770 1000    CONTINUE
2780 1500    TTHETA=TJ1
2790         RETURN
2800 C
2801 C
2802 C
2810 C  **** SUBROUTINE GROUPING R AND Y VALUES
2811 C
2820         SUBROUTINE HISTO(Y,RR,TH)
2830 C
2835         YY=ABS(Y)
2836         IF(YY.LT.1.0)THEN
2837           YL=0.1
2838           K=1
2839         ELSE
2840           YL=0.25
2841           K=7
```

```
2842          ENDIF
2843          IF(RR.LT.1.0)THEN
2844            RL=0.1
2845            K2=1
2846          ELSE
2847            RL=0.25
2848            K2=7
2849          ENDIF
2860          NY=INT(YY/YL)+K
2870          IF(NY.GT.26)NY=27
2880          YBIN(NY)=YBIN(NY)+1
2890          NR=INT(RR/RL)+K2
2900          IF(NR.GT.26)NR=27
2910          RBIN(NR)=RBIN(NR)+1
2920          TT=ABS(TH)
2930          I=INT(TT/10.)+1
2940          IF(I.EQ.10)I=9
2950          THBIN(I)=THBIN(I)+1
2960          J=INT((TT+2.5)/5.0)+1
2970          T5BIN(J)=T5BIN(J)+1
2980          J=INT((TT+5.0)/10.0)+1
2990          T10BIN(J)=T10BIN(J)+1
3000  C
3010          RETURN
3020          END
```

9.4. Output.

An example of the output from the program as given above is presented below.

```
 1  MINKE SIMULATION STUDY
 2
 3  **********************
 4
 5
 6  SIGHTING : NEG EXPONENTIAL SIGHTING RATE, R= 5.000
 7  FUNCTION 1 : MEW*EXP(-LAMBDA*RR)
 8
 9   LAMBDA= 1.50     MEW= 5.00
10
11    VW=   .000 VP=12.000 DELT= .01000
12
13  MAX=  100000  N=467632   NN=801293
14
15  WMAX=16.58312    YMAX= 5.00000
16
17
18     GROUPED :        Y  AND     R VALUES
19      .00-   .10          135           11
20      .10-   .20          115            7
21      .20-   .30           75           26
22      .30-   .40           73           40
23      .40-   .50           48           33
24      .50-   .60           41           47
25      .60-   .70           32           38
26      .70-   .80           30           27
27      .80-   .90           35           41
28      .90-  1.00           24           41
29     1.00-  1.25           33           75
30     1.25-  1.50           22           70
31     1.50-  1.75            9           55
32     1.75-  2.00            4           43
33     2.00-  2.25            6           36
34     2.25-  2.50            5           20
35     2.50-  2.75            1           24
36     2.75-  3.00            0           16
37     3.00-  3.25            0           15
38     3.25-  3.50            1            4
39     3.50-  3.75            0           10
40     3.75-  4.00            0            3
41     4.00-  4.25            0            4
42     4.25-  4.50            0            1
43     4.50-  4.75            0            1
44     4.75-  5.00            0            1
45     5.00-*****              0            0
46
47  OBS THETA ROUNDED TO 5 AND 10 DEG.     PERCENT
48
49       .0          41            86          5.95       12.48
50      5.0          94                       13.64
51     10.0          88           175          12.77       25.40
52     15.0          82                       11.90
53     20.0          78           156          11.32       22.64
54     25.0          61                        8.85
55     30.0          64           107          9.29        15.53
56     35.0          44                        6.39
57     40.0          38            80          5.52        11.61
58     45.0          34                        4.93
59     50.0          22            52          3.19         7.55
60     55.0          18                        2.61
61     60.0           9            18          1.31         2.61
62     65.0           4                         .58
```

70.0	8	11	1.16	1.60
75.0	1		.15	
80.0	2	3	.29	.44
85.0	0		.00	
90.0	1	1	.15	.15

COUNT= 689 N= 100000 YSUM= 329.0193

GROUPED OBSERVED THETA VAL. IN DEG.		PERCENT
.0 TO< 10.0	175	25.40
10.0 TO< 20.0	175	25.40
20.0 TO< 30.0	125	18.14
30.0 TO< 40.0	96	13.93
40.0 TO< 50.0	65	9.43
50.0 TO< 60.0	31	4.50
60.0 TO< 70.0	14	2.03
70.0 TO< 80.0	6	.87
80.0 TO< 90.0	2	.29

R= 5.000 WMAX=16.583

LAMBDA= 1.50 MEW= 5.0000 YBAR= .4775

TRUE D= 3015.11 D=(COUNT-1)/2*YBAR= 720.3

ANDERSON,D.R. and POSPAHLA,R.S. 1970. Correction of bias in belt transects of immotile objects. J. Wildl. Manage. 34(1):141-146

BEST,P.B. and BUTTERWORTH,D.S. 1980. Report of the Southern Hemisphere Minke Whale Assessment Cruise, 1978/79. Rep. int. Whal. Commn. 30:257-283

BURNHAM,K.P. and ANDERSON,D.R. 1976. Mathematical models for non-parametric inferences from line transect data. Biometrics 32(2).325-336

BURNHAM,K.P., ANDERSON,D.R. and LAAKE,J.L. 1980. Estimation of density from line transect sampling of biological populations. Wildlife Monographs 72, 202pp.

BUTTERWORTH, D.S. 1982. A possible basis for choosing a functional form for the distribution of sightings with right angle distance : some preliminary ideas. Rep. int. Whal. Commn. 32: 555-558

BUTTERWORTH,D.S. and BEST,P.B. 1982. Report of the Southern Hemisphere Minke Whale Assessment Cruise, 1980/81. Rep. int. Whal. Commn. 32:835-873

BUTTERWORTH,D.S., BEST,P.B. and BASSON,M. 1982. Results of analysis of sighting experiments carried out during the 1980/81 Southern Hemisphere Minke Whale Assessment Cruise. Rep. int. Whal. Commn. 32:819-834

CRAIN,B.R., BURNHAM,K.P., ANDERSON,D.R. and LAAKE,J.L. 1978. A Fourier series estimator of population density for line transect sampling. Utah St. Univ. Press, Logan, Utah. 25pp.

DOI,T., KASAMATSU,F. and NAKANO,T. 1982. A Simulation study on sighting surveys of minke whales in the Antarctic. Rep. int. Whal. Commn. 32:919-928

EBERHARDT,L.L. 1968. A Preliminary appraisal of line transects. J. Wildl. Manage. 32(1):82-88

EKLUND,C.R. and ATWOOD,E.L. 1962. A population study of antarctic seals. J. Mammal. 43(2):229-238

FORBES,S.A. and GROSS,A.D. 1921. The orchard birds of an Illinois summer. Ill. Nat. Hist. Surv. Bull. 14:1-8

GATES,C.E. 1979. Line transects and related issues. Pp 71-154. In R.M. Cormack, G.P. Patil and D.S. Robson (Eds.) Sampling biological populations. Internatl. Co-op. Publ. House, Fairland, Md.

GATES,C.E., MARSHALL,W.H. and OLSON,D.P. 1968. Line transect method of estimating grouse population densities. Biometrics 24(1):135-145

HAYES,R.J. and BUCKLAND,S.T. Radial distance models for the line transect method. In press.

HAYNE,D.W. 1949. An examination of the strip census method for estimating animal populations. J. Wildl. Manage. 13(2):145-157

HIBY,A.R. 1982. The effect of random whale movement on density estimates obtained from whale sighting surveys. Rep. int. Whal. Commn. 32:791-793

KELKER,G.H. 1945. Measurement and interpretation of forces that determine populations of managed deer. Ph.D. thesis Univ. of Michigan, Ann Arbor. 422pp.

KOOPMAN,B.O. 1956. The theory of search I. Kinematic bases. Ops. Res. 4:324-346. Also in:

KOOPMAN,B.O. 1980. Search and screening. Pergamom Press Inc.

POLLOCK,K.H. 1978. A family of density estimators for line transect sampling. Biometrics 34(3).475-478

QUINN,T.J. and GALUCCI,V.F. 1980. Parametric models for line transect estimators of abundance. Ecology 61(2):293-302

ROBINETTE,W.L., JONES,J.S., GASHWILER,J.S. and ALDOUS,C.M. 1954. methods for censusing winter-lost deer. Trans. N. Amer. Wildl. Conf. 19:511-525

ROBINETTE,W.L., JONES,J.S., GASHWILER,J.S. and ALDOUS,C.M. 1956. Further analysis of methods for censusing winter-lost deer. J. Wildl. Manage. 20(1):75-78

SCHWEDER,T. 1977. Point process models for line transect experiments. Pp. 221-242. In J.R. Barba, F. Brodeau, G. Romier and B. Van Cutsem. (Eds.). Recent developments in statistics. North-Holland Publ. Co., New York, N.Y.

SEBER,G.A.F. 1973. The estimation of animal abundance. Hafner Publ. Co., Inc.,New York, N.Y. 506pp.

SKELLAM,J.G. 1958. The mathematical foundations underlying the use of line transects in animal ecology. Biometrics 14(3):385-400