Canadian Atlantic Fisheries
Scientific Advisory Committee

CAFSAC Research Document 86/75

Comité scientifique consultatif des
pêches canadiennes dans l'Atlantique

CSCPCA Document de recherche 86/75

Calculation of mid-year estimates in SPA

by

George White
Biological Sciences Branch
Marine Fish Division
Bedford Institute of Oceanography
Fisheries and Oceans
P.O. Box 1006
Dartmouth, N.S. B2Y 4A2

Abstract

Mid-year estimates from sequential population analysis are often used for calibration. Because the population numbers at mid-year depend on the exploitation pattern, no single formula is appropriate in all cases. It is suggested that assessment documents should present the actual formula that was used.

Résumé

Les estimations semestrielles provenant des analyses séquentielles de population servent souvent à l'étalonnage. Etant donné que les chiffres de population relevés en milieu d'année dépendent des caractéristiques de l'exploitation, aucune formule donnée ne convient à tous les cas. On propose que les documents relatifs aux estimations indiquent la formule utilisée.

Introduction

The basic SPA formulae provide estimates for population numbers at the endpoints of the intervals (normally the calendar year) for which catch is accumulated. Mid-year estimates from SPA calculations are often required in order to calibrate against independent abundance series such as those derived from commercial catch rates or research surveys. Since population numbers at mid-year depend on the exploitation patterns in each fishery, there is no single formula that is appropriate in all cases.

Several commonly used interpolation formulae will be presented in both mathematical and computer (APL) formulae. A series of figures is presented to illustrate the difference between the various formulae.

Interpolation Formulae

Table 1 provides a list of commonly used mathematical formulae together with their APL equivalents. Selection of an appropriate formula for calculation of mid-year estimates must be made in light of information concerning the timing of catches. In some cases it may be necessary to use different formulae for certain age groups. This would be the case if seasonal fisheries exploit different age compositions.

The differences between the formulae in Table 1 were examined by comparing the estimates for mid-year numbers with a simulated population generated using hypothetical monthly mortality rates. Two mortality patterns were used, one with catches concentrated over the summer months, and the second with the catches narrowly concentrated in February. The actual monthly mortality rates used in the simulations were based on the following patterns:

	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
a.	0.05	0.05	0.05	0.1	0.3	0.6	0.6	0.4	0.3	0.2	0.1	0.05
b.	0.15	1.0	0.15	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05

For each simulation, these values were scaled to produce an average annual instantaneous rate of fishing mortality, F, of 0.1, 0.3, or 0.5. Figures 1-9 present the comparisons for various values of M and F. In each case the simulated numbers are shown as solid curves and the values calculated using the various formulae are shown as small circles. The upper curve in these graphs corresponds to the mortality pattern in line a. of the table. Because the catches are taken later in the year, the annual mortality rate was the same for each pattern, the endpoints of the two curves were the same. The APL functions used to produce these figures are presented in the Appendix.

Use of an inappropriate formula can result in large errors. If the catches are concentrated near the middle of the year, then either average numbers or linear interpolation performs well. If the catches are concentrated near the start of the year, the interpolation formula must be chosen with care. Errors will be greater when F is large compared to M. It must also be noted that the rate of natural mortality is likely to vary seasonally. Thus significant errors may occur even if the interpolation formula is chosen with care. Although use of seasonal catch data does not help overcome the problem of seasonal variation in natural mortality rates, it would eliminate the need for interpolation, thus removing a potential source of error.

Conclusions

The terminology for the various interpolation formulae is not adequate and has been abused in the past. Perhaps because the average population biomass is sometimes used as an estimate of the mid-year numbers, many assessment programs have variables with names such as "POPBIOMASSΔMIDYR" (Rivard 1982, p. 143) to represent the "average biomass" calculated as the produce of mid-year weights and the average numbers obtained using the last formula in Table 1. Variable names in computer programs are often poor or incomplete descriptions of the formula used to derive the variable. While mnemonic variable names are desirable, it is seldom possible to capture all the information in a complex mathematical formula in a name.

Few assessment documents provide an adequate description of the interpolation formula. Terms such as "log-linear" interpolation refer to a general formula. In a given instance it is important to know the location (in time) of the interpolated point. Even if a standard terminology was adopted by CAFSAC, the possibility for confusion when dealing with existing documents or those prepared outside CAFSAC remains. Mathematical notation is unambiguous and its meaning does not vary over time. Thus the best way to document what has been done is to provide the acutal mathematical formula used for the calculation.

Literature Cited

Rivard, D. 1982. APL programs for stock assessment (revised). Can. Tech. Rep. Fish. Aquat. Sci. 1091: 146 p.

Table 1. Interpolation formulae used to obtain estimates of midyear population numbers from SPA numbers at the beginning of the year.

interpolation formula/ APL expression	key assumption
$N(a,t+\Delta t) = N(a,t)e^{-M\Delta t}$ `NΔt ← N×*-M×Δt`	catch taken **after** time t + Δt
$N(a,t+\Delta t) = N(a,t)e^{-M\Delta t-F(a,t)}$ `NΔt ← N×*-(M×Δt)+F`	catch taken **before** time t + Δt
$N(a,t+\Delta t) = N(a,t)e^{-(M+F(a,t))\Delta t}$ `NΔt ← N×*-Δt×F+M`	F(a,t) is constant over the year, resulting in more of the catch being taken early in the year (log-linear interpolation)
$N(a,t+\Delta t) = N(a,t)\times(1-\Delta t + \Delta t e^{-(F(a,t)+M)})$ `NΔt ← N×(1-Δt)+Δt×*-F+M`	linear interpolation

although not strictly an interpolation formula, the average population:

$$\bar{N}(a,t) = N(a,t)(1 - e^{-(F(a,t)+M)}) \div (F(a,t)+M)$$

`NA ← N×(1 - *-F+M)÷F+M`

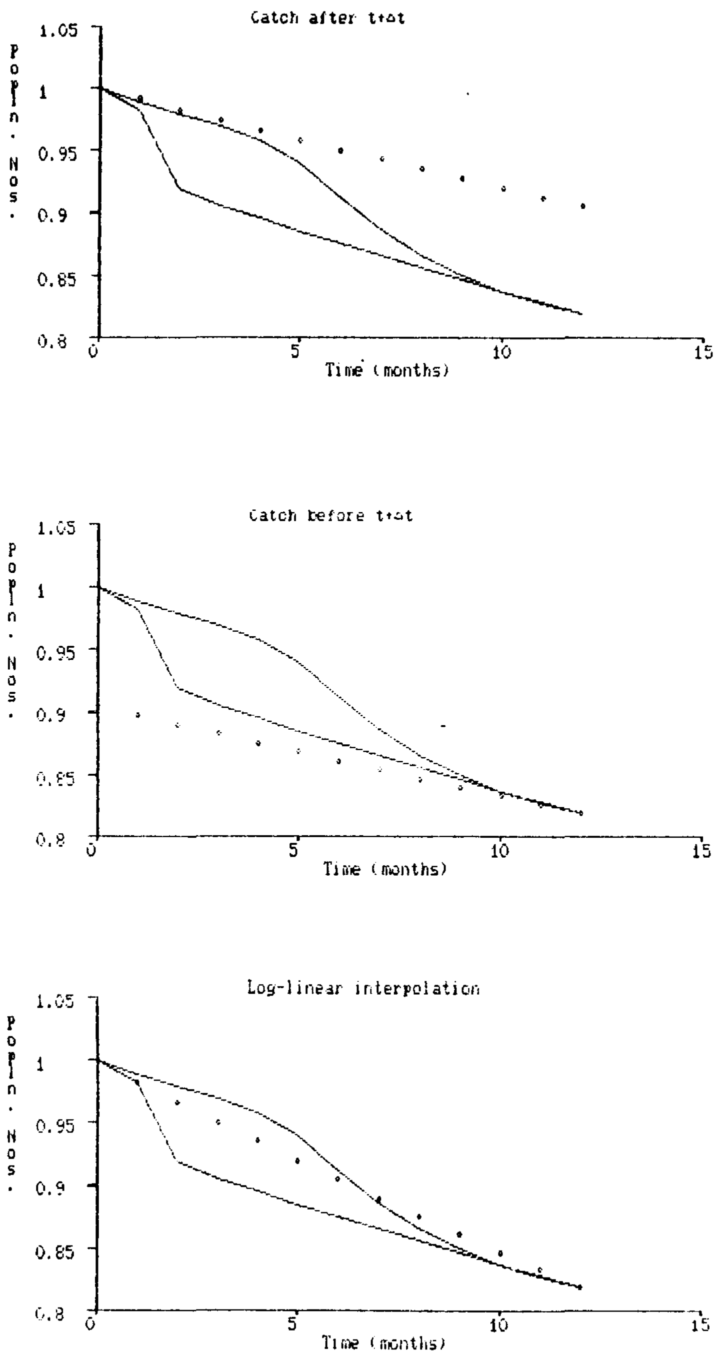
is sometimes used as an estimate for the mid-year numbers.

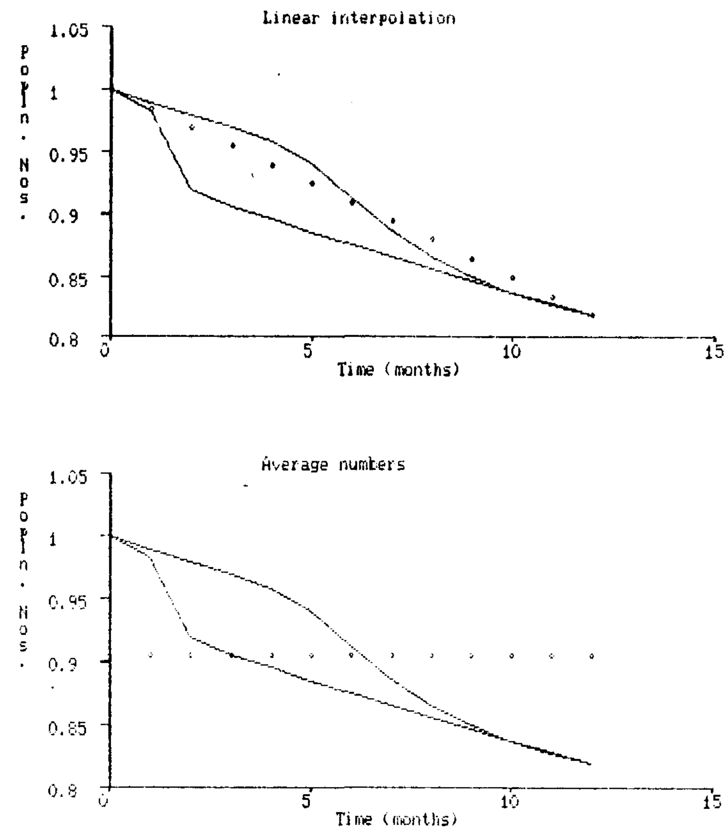
Notation:

- $N(a,t)$ = beginning of the year population numbers at age a, time t, where $a=a_0,\ldots,a_1$ and $t=t_0,\ldots,t_1$
- M = (assumed constant) instantaneous rate of natural mortality
- $F(a,t)$ = annual average instantaneous rate of natural mortality, defined as the constant rate which would produce the same change in numbers as was observed in the population
- Δt = the fraction of the year at which population numbers are to be estimated, $0 \le \Delta t \le 1$

APL terms:

- N = the array whose elements are the values N(a,t), for all a,t
- NΔt = the array whose elements are the values N(a,t+Δt), for all a,t
- F = the array whose elements are the values F(a,t) for all a,t
- NA = the array whose elements are the values $\bar{N}(a,t)$ for all a,t





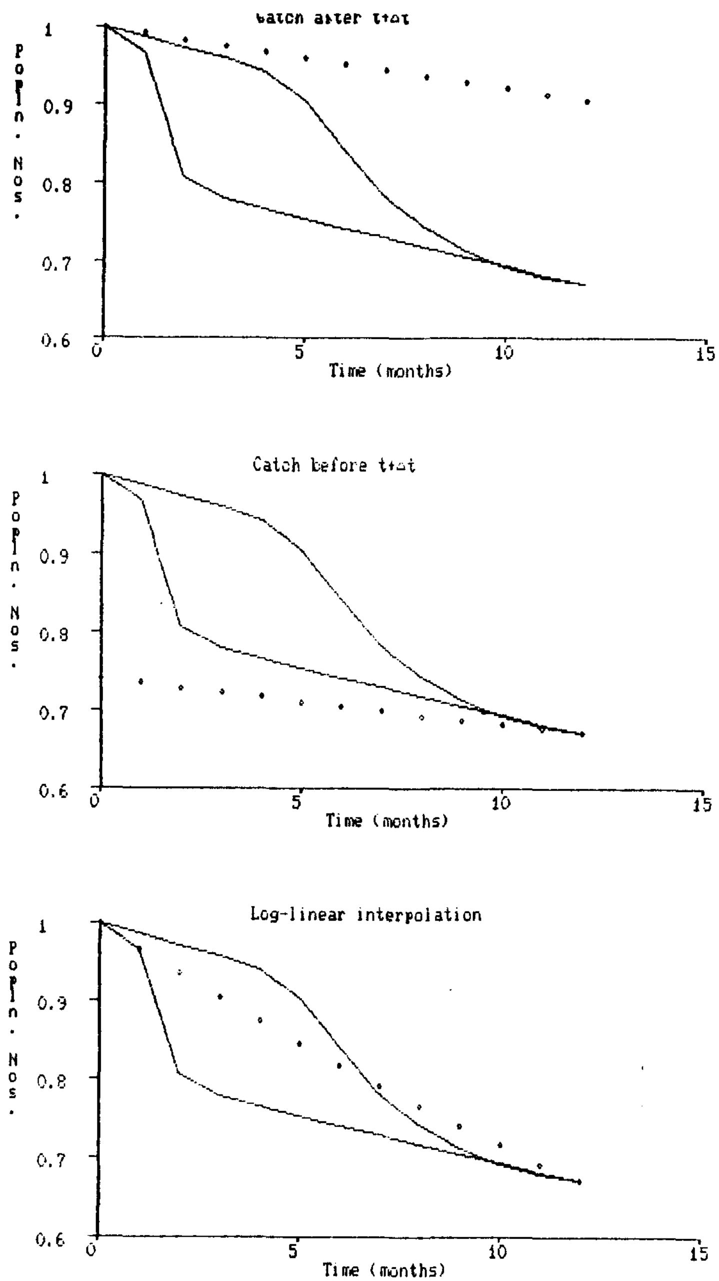
Compare

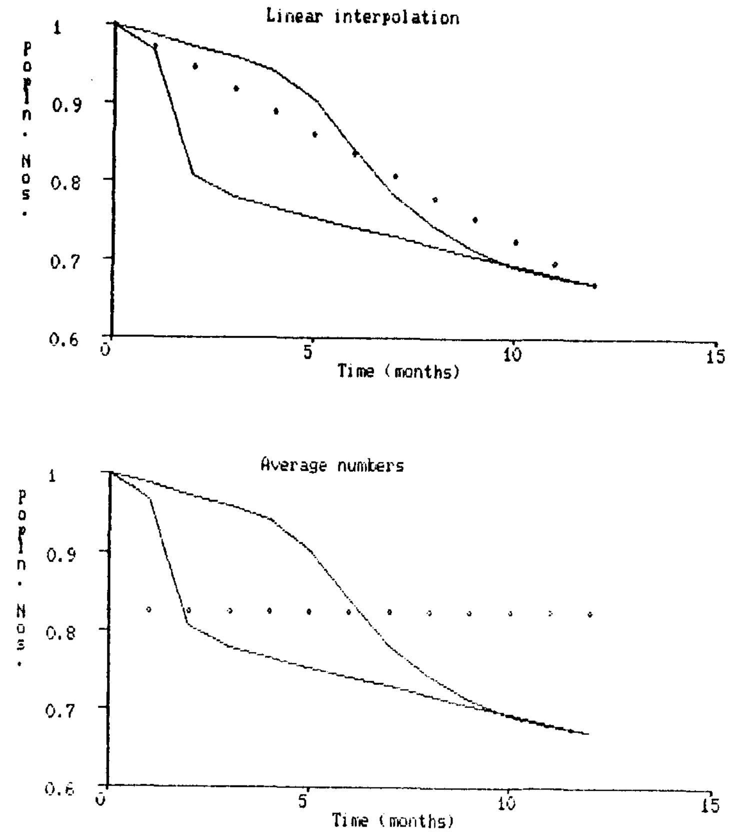
Natural Mortality: 0.1
Average F: 0.1 0.1

Figure 1. Interpolation curves for the formulae in Table 1.

$M = 0.1$
$F = 0.1$

Solid curves - hypothetical populations
Dotted lines - interpolated values





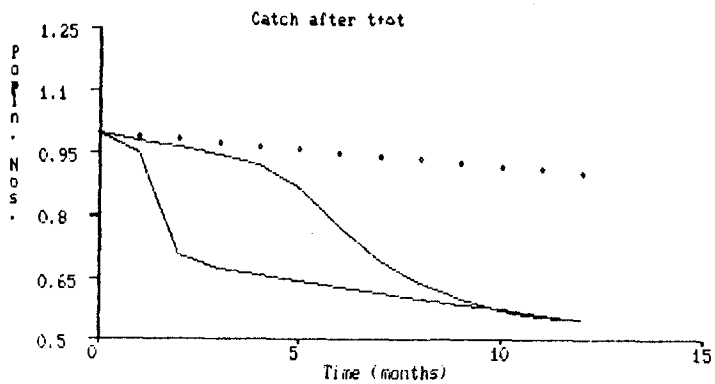
Compare

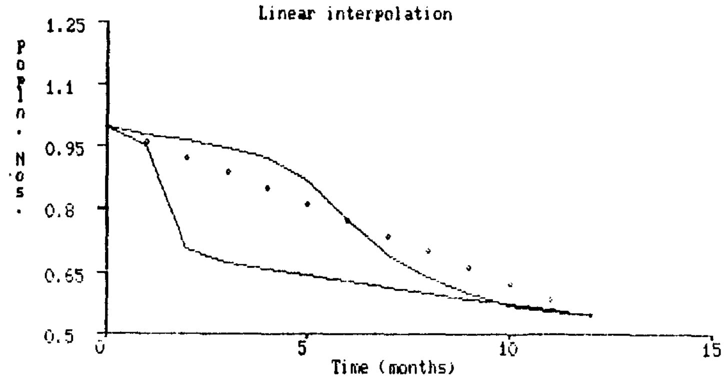
Natural Mortality: 0.1
Average F: 0.3 0.3

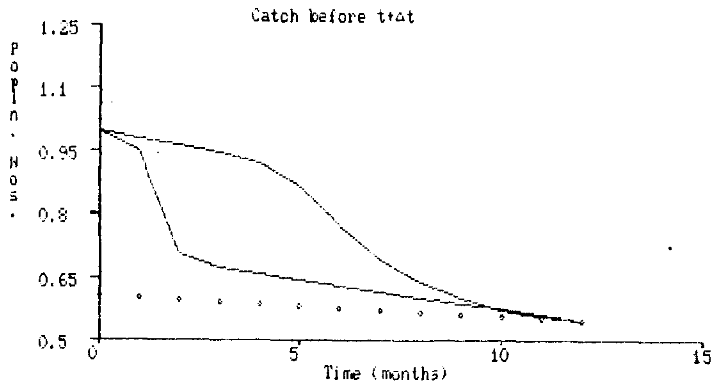
Figure 2. Interpolation curves for the formulae in Table 1.

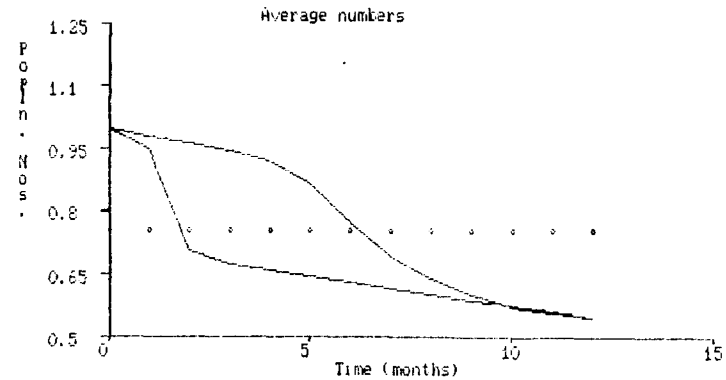
$M = 0.1$
$F = 0.3$

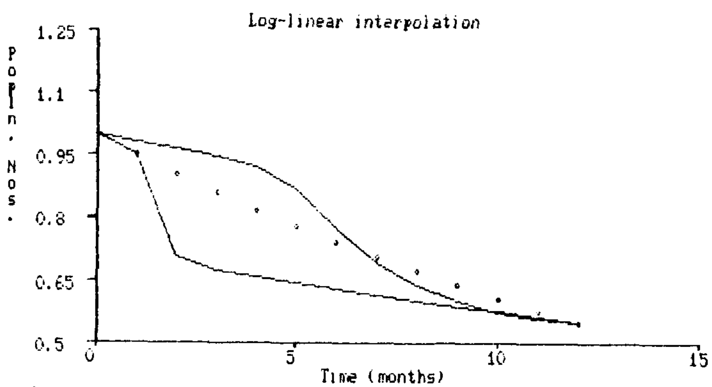
Solid curves - hypothetical populations
Dotted lines - interpolated values











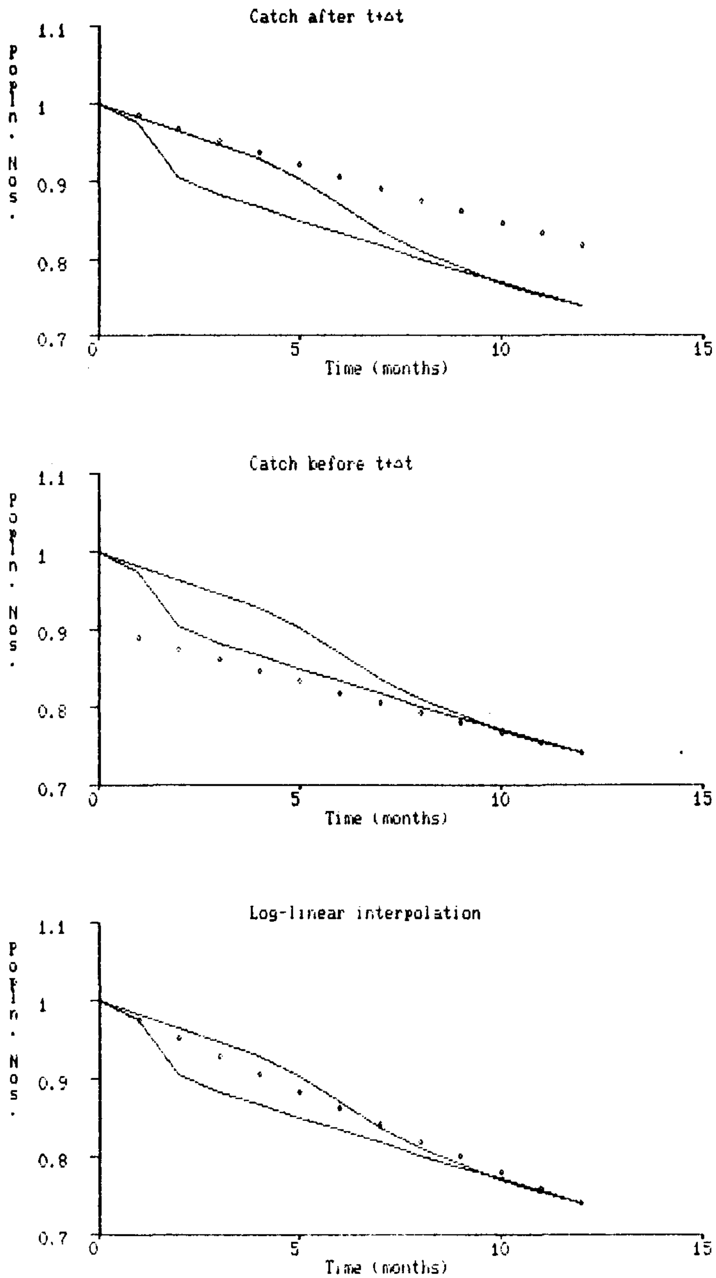
Compare

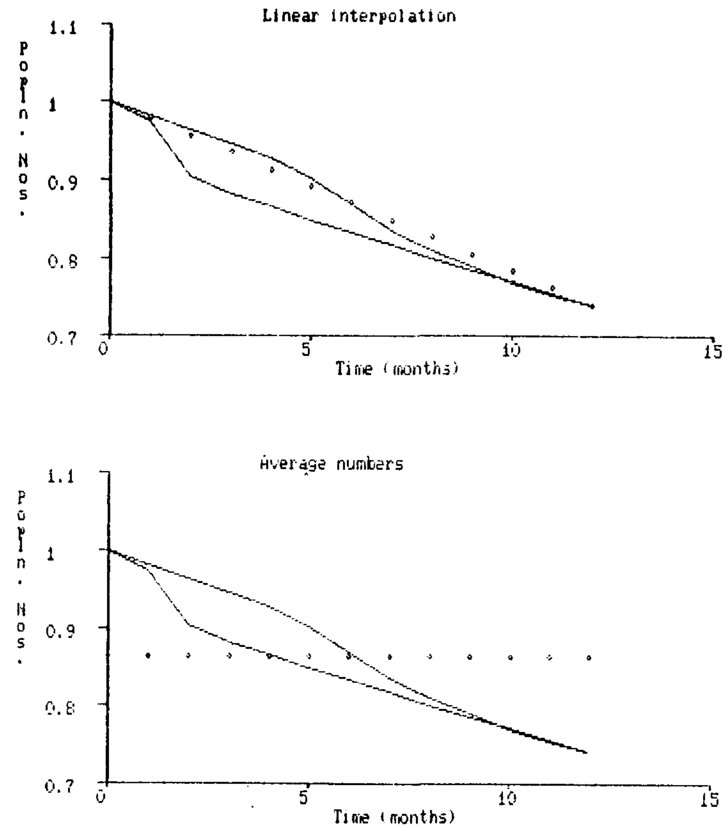
Natural Mortality: 0.1
Average F: 0.5 0.5

Figure 3. Interpolation curves for the formulae in Table 1.

$M = 0.1$
$F = 0.5$

Solid curves - hypothetical populations
Dotted lines - interpolated values





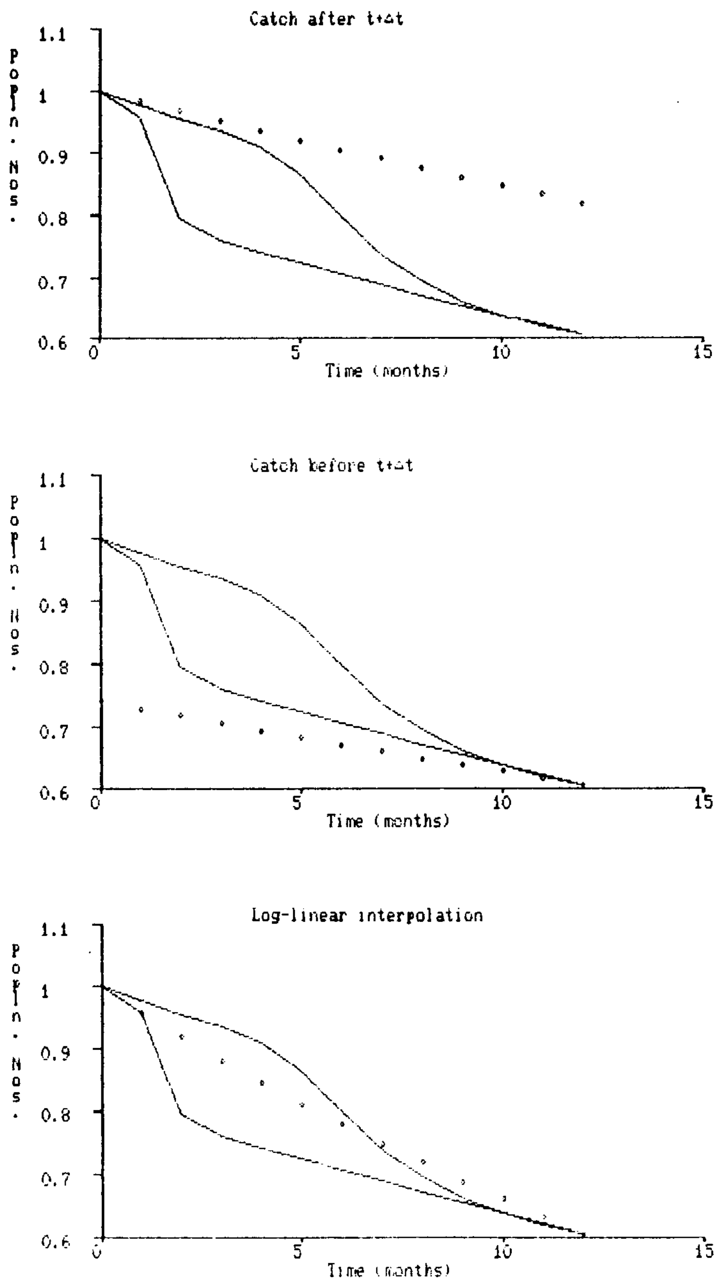
Compare

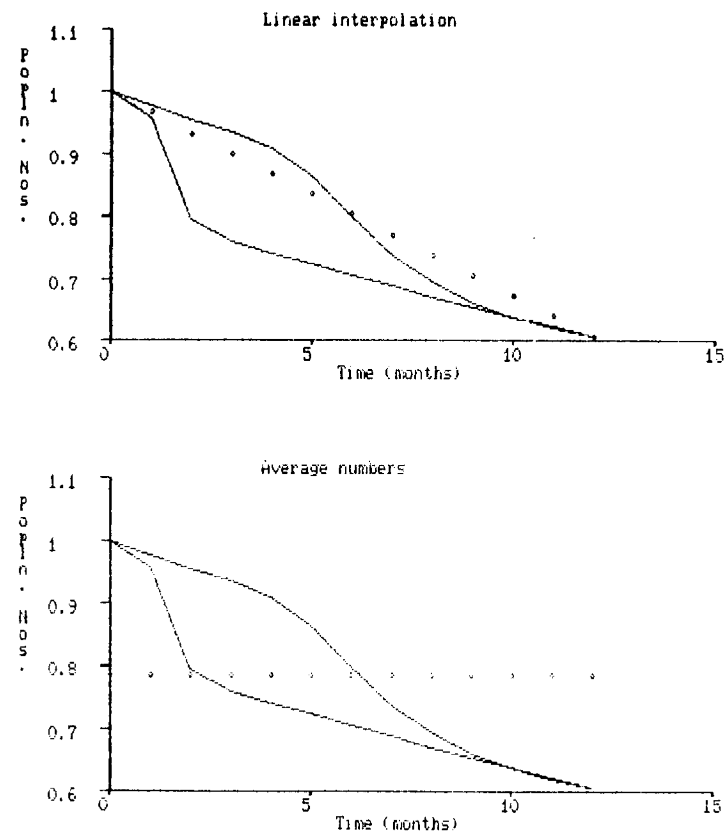
Natural Mortality: 0.2
Average F: 0.1 0.1

Figure 4. Interpolation curves for the formulae in Table 1.

$M = 0.2$
$F = 0.1$

Solid curves - hypothetical populations
Dotted lines - interpolated values





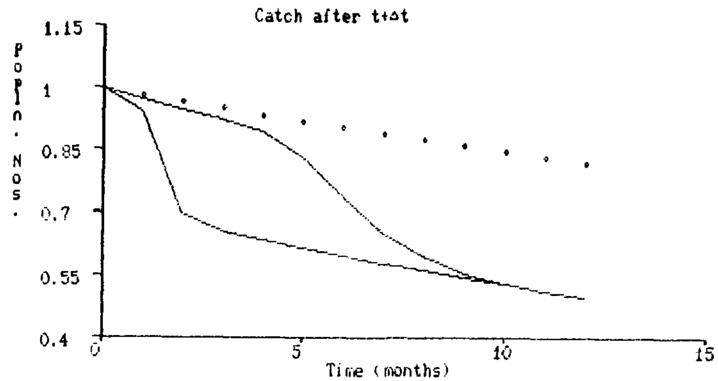
Compare

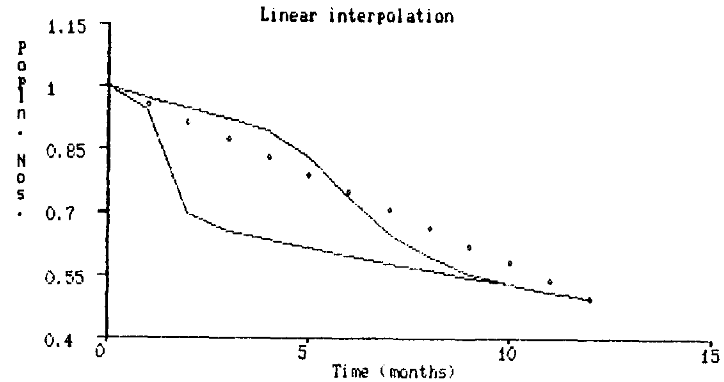
Natural Mortality: 0.2
Average F: 0.3 0.3

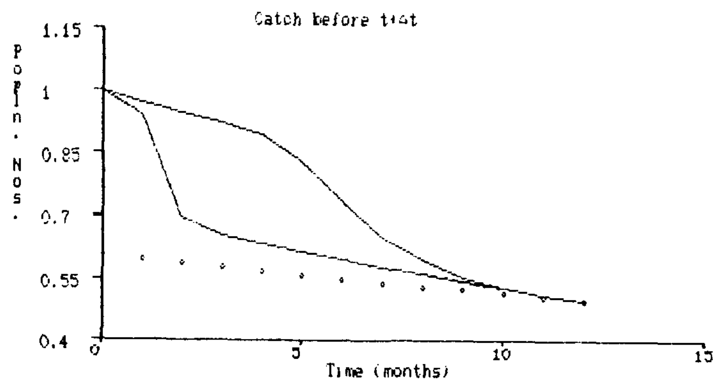
Figure 5. Interpolation curves for the formulae in Table 1.

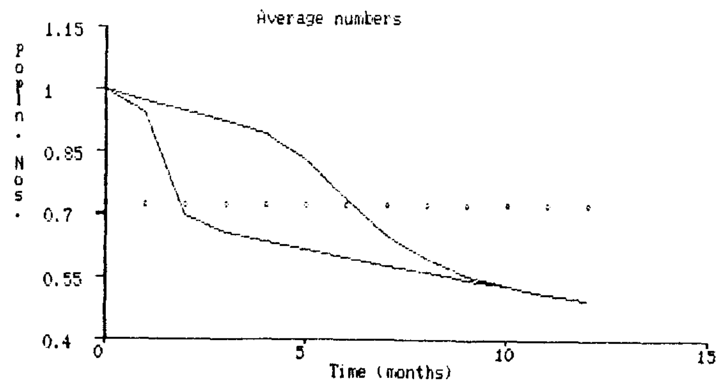
M = 0.2
F = 0.3

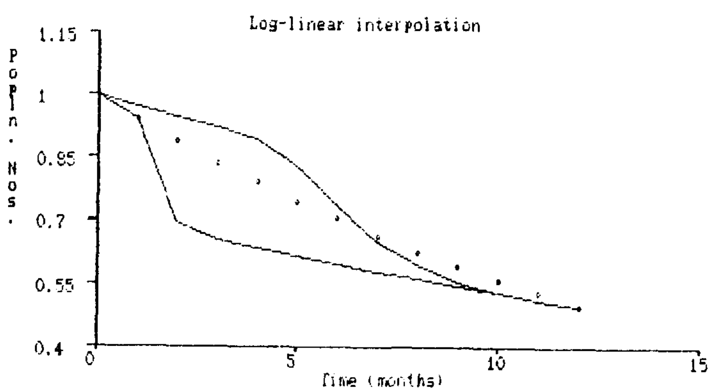
Solid curves - hypothetical populations
Dotted lines - interpolated values











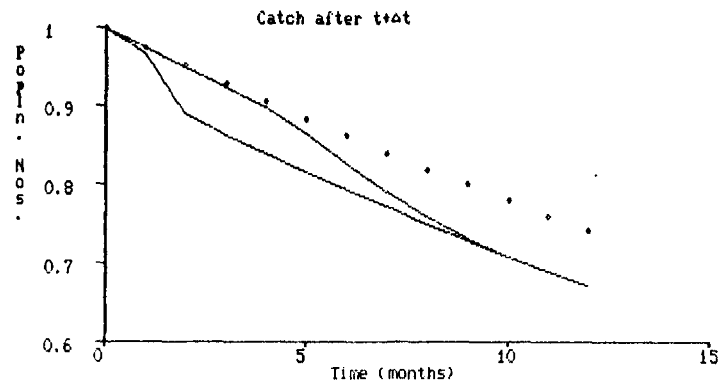
Compare

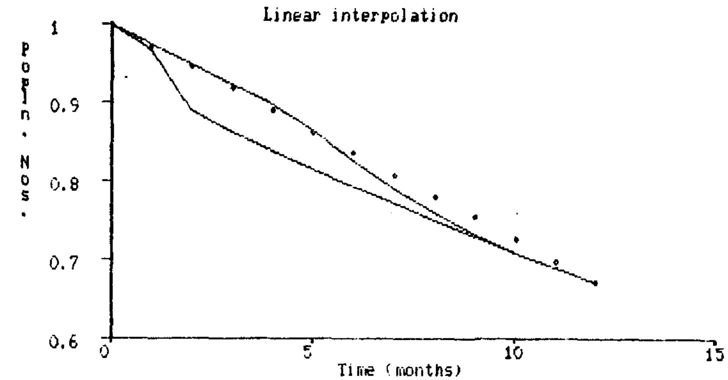
Natural Mortality: 0.2
Average F: 0.5 0.5

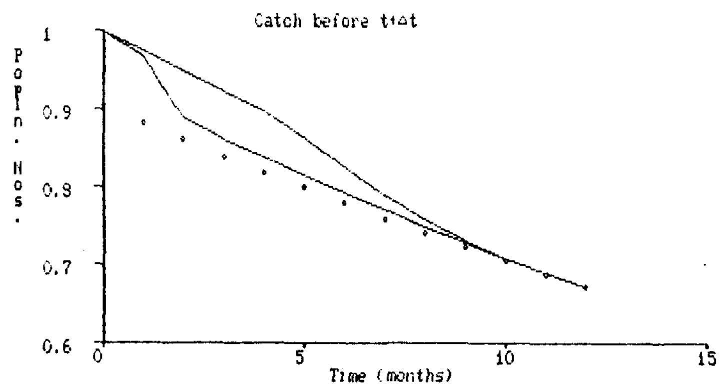
Figure 6. Interpolation curves for the formulae in Table 1.

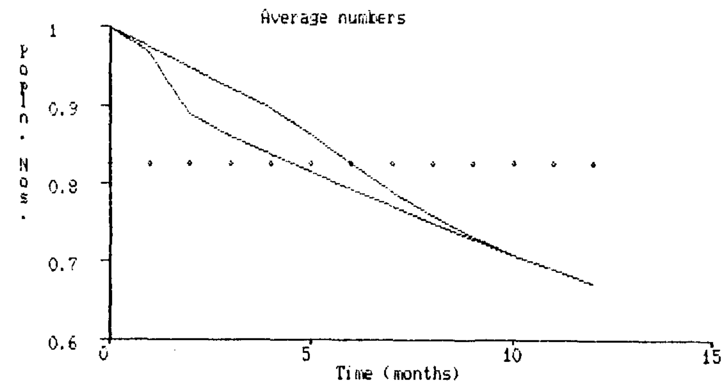
M = 0.2
F = 0.5

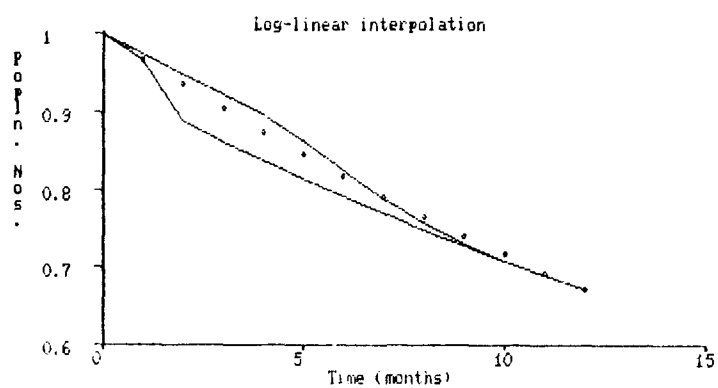
Solid curves - hypothetical populations
Dotted lines - interpolated values











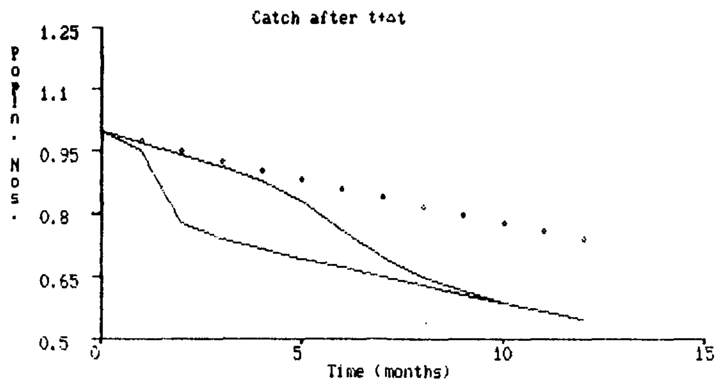
Compare

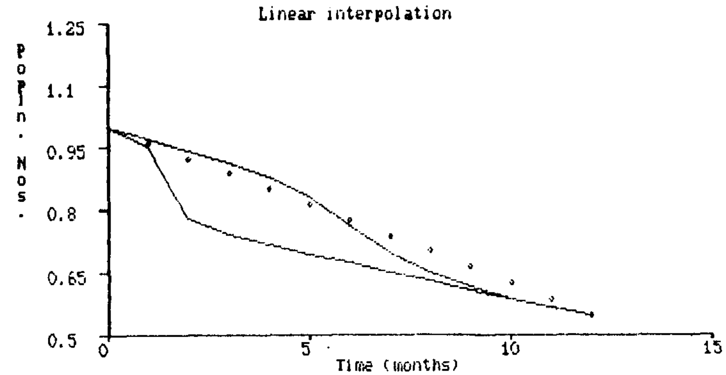
Natural Mortality: 0.3
Average F: 0.1 0.1

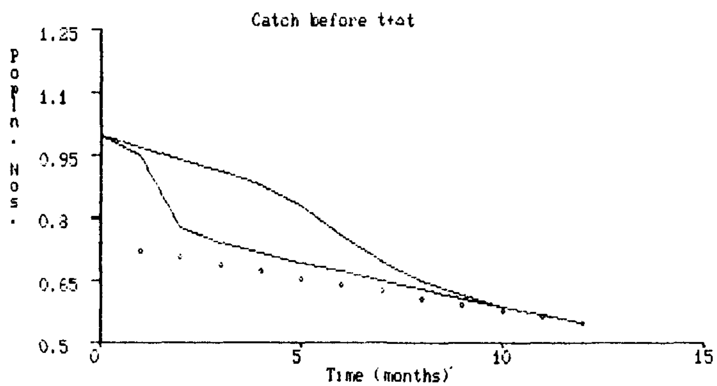
Figure 7. Interpolation curves for the formulae in Table 1.

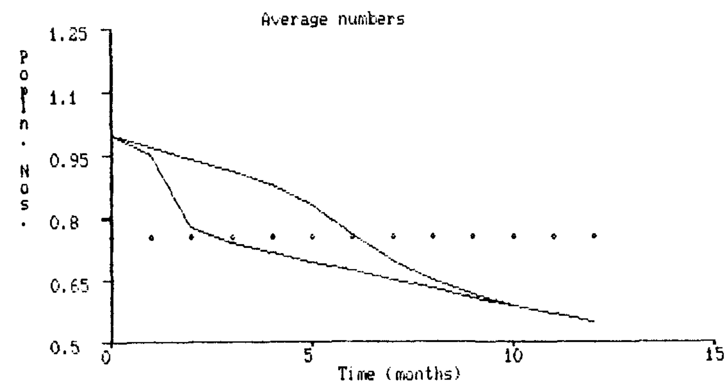
M = 0.3
F = 0.1

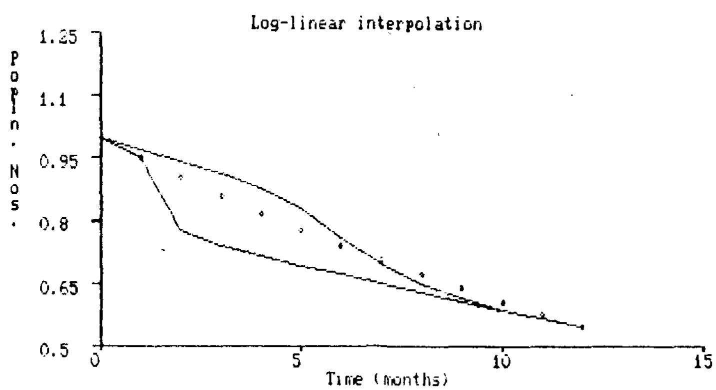
Solid curves - hypothetical populations
Dotted lines - interpolated values











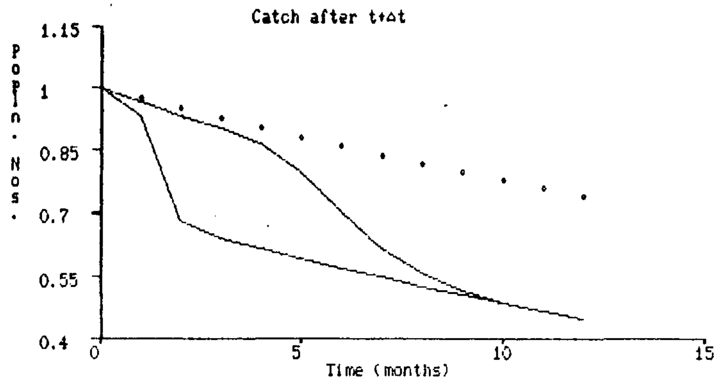
Compare

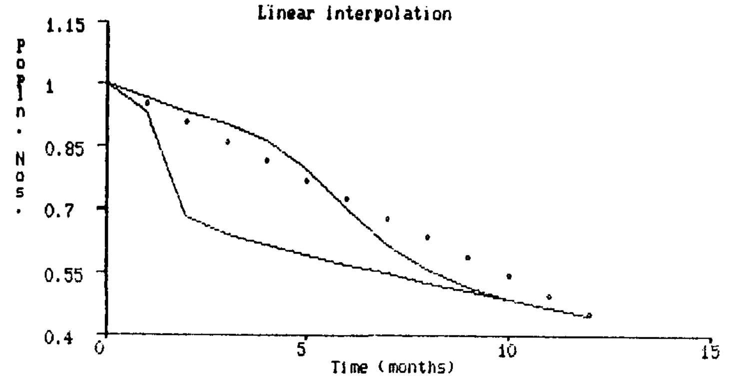
Natural Mortality: 0.3
Average F: 0.3 0.3

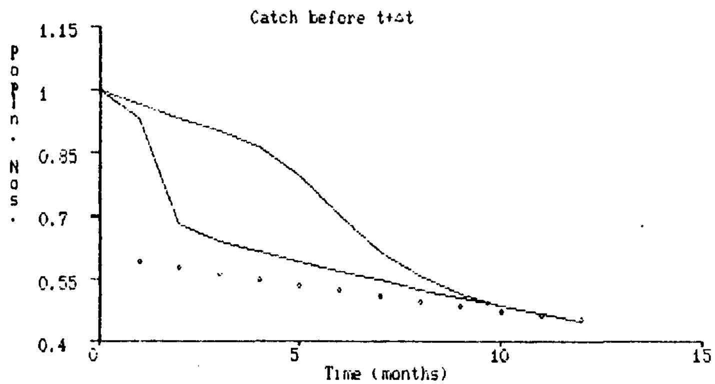
Figure 8. Interpolation curves for the formulae in Table 1.

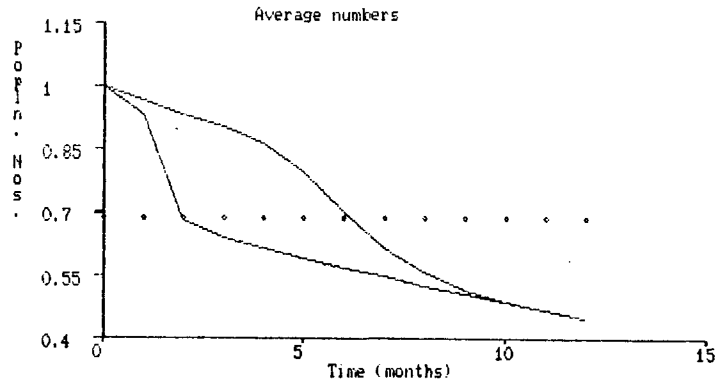
M = 0.3
F = 0.3

Solid curves - hypothetical populations
Dotted lines - interpolated values









Compare

Natural Mortality: 0.3
Average F: 0.5 0.5

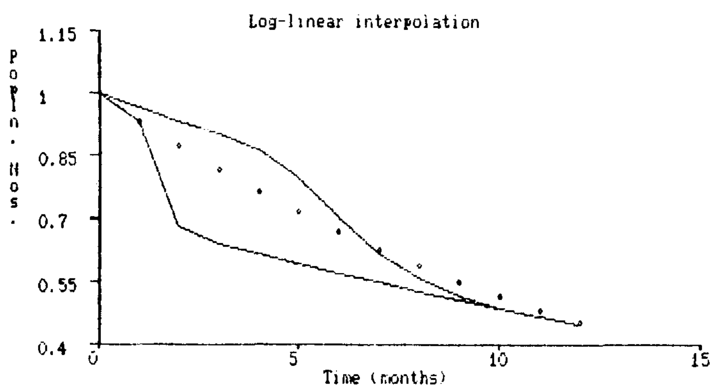


Figure 9. Interpolation curves for the formulae in Table 1.

M = 0.3
F = 0.5

Solid curves - hypothetical populations
Dotted lines - interpolated values

APPENDIX. APL functions used to generate Figures 1-9.

```
      )FNS
 Compare Fbar    N0      N1      N2      N3      N4     Nbar


    ∇ Z←F Compare M;dx;dy;tit;n;nΔt;Δt;f;M;f1;f2;f1bar;f2bar
[1]  ⍝ compare various mid-year formulae
[2]   n0←1
[3]   Δt←0,(⍳12)÷12
[4]   f1← 0.15 1 0.15 ,9⍴0.05
[5]   f2← 0.05 0.05 0.05 0.1 0.3 0.6 0.6 0.4 0.3 0.2 0.1 0.05
[6]   f1bar←(M+f1)Fbar n0
[7]   f1←F×f1÷f1bar
[8]   f2bar←(M+f2)Fbar n0
[9]   f2←F×f2÷f2bar
[10]  Z←'Compare',⎕TCNL
[11]  Z←Z,⎕←⎕TCNL,'Natural Mortality: ',(⍕M),⎕TCNL
[12]  Z←Z,⎕←'Average F: ',(⍕((M+f1)Fbar n0),(M+f2)Fbar n0),⎕TCNL
[13]  n1←(f1+M)N0 n0
[14]  n2←(f2+M)N0 n0
[15]  dx←'Time (months)' ⋄ dy←'Popln. Nos.' ⋄ tit←'Catch after t+Δt'
[16]  (x←0,⍳12)ΔPLOTXY 3 13 ⍴n1,n2,F N1 n0
[17]  tit←'Catch before t+Δt'
[18]  (x←0,⍳12)ΔPLOTXY 3 13 ⍴n1,n2,F N2 n0
[19]  tit←'Log-linear interpolation'
[20]  (x←0,⍳12)ΔPLOTXY 3 13 ⍴n1,n2,F N3 n0
[21]  tit←'Linear interpolation'
[22]  (x←0,⍳12)ΔPLOTXY 3 13 ⍴n1,n2,F N4 n0
[23]  tit←'Average numbers'
[24]  (x←0,⍳12)ΔPLOTXY 3 13 ⍴n1,n2,13⍴F Nbar n0
[25]
    ∇

    ∇ F←Z Fbar N0;N
[1]  ⍝ average F from monthly Z
[2]   N←¯1↑N0×1,*-+\Z×(1↓Δt)-¯1↓Δt
[3]   F←(⍟N0÷N)-M
    ∇

    ∇ N←Z N0 n0
[1]  ⍝ numbers from monthly Z
[2]   N←n0×1,*-+\Z×(1↓Δt)-¯1↓Δt
    ∇

    ∇ NΔt←F N1 N
[1]  ⍝ catch taken after time t+Δt
[2]   NΔt←N×*-M×Δt
    ∇

    ∇ NΔt←F N2 N
[1]  ⍝ catch taken before time t+Δt
[2]   NΔt←N×*-(M×Δt)+F
    ∇

    ∇ NΔt←F N3 N
[1]  ⍝ log-linear interpolation
[2]   NΔt←N×*-Δt×F+M
    ∇

    ∇ NΔt←F N4 N
[1]  ⍝ linear interpolation
[2]   NΔt←N×(1-Δt)+Δt×*-F+M
    ∇

    ∇ NA←F Nbar N
[1]  ⍝ average numbers
[2]   NA←N×(1-*-F+M)÷F+M
    ∇
```