International Council for
the Exploration of the Sea

ICES C.M. 1996/D:9 Ref. H
Statistics Committe

Back-calculated length at age in Norwegian spring-spawning herring (*Clupea harengus* L.): Estimating the common intercept or origin by tuning back-calculated length frequencies to observed ones

by

Jens Christian Holst
Institute of Marine Research
P. O. Box 1870, N-5024 Bergen, Norway

Abstract

In Norwegian spring-spawning herring the common intercept of the back-calculating formula, as estimated by linear regression of the fish-scale relationship, grosssly overestimated the back-calculated length frequencies distributions as compared to the corresponding observed ones. By tuning the back-calculated length frequency distributions to fit optimally with the corresponding observed ones, an "optimum fit" intercept was estimated. This intercept was observed to increase with the age back-calculated. Thus, the growth trajectory of single fish becomes non-linear, and the origin of the growth trajectory is probably situated close to the "biological intercept" as proposed by Campana (1990). The quality of the "optimum fit" intercept depend upon factors as size-selective mortality and representative sampling and should be evaluated further.

Introduction

In applying traditional methods to estimate the population based intercept for back-calculating lengths at age in Norwegian spring-spawning herring, an unacceptable discrepancy was noted between the back-calculated and the corresponding observed length frequencies. In the present paper an alternative method is proposed to estimate the common intercept or origin, in order to obtain back-calculated lengths in better accordance with the corresponding observed ones.

A population based intercept, or origin, is required for individual back-calculation of lengths at age from growth increments in otoliths or scales of fishes (Campana, 1990; Ricker, 1992). The intercept may either be estimated as the Y-intercept of the fish length (L) - scale size (S) regression or as a "biological intercept" based on the fish length at first scale formation. The former is currently the most commonly used method (Francis, 1990), although there may be advantages in applying the latter one (Campana, 1990).

When the statistical intercept referred above was applied for back-calculating lengths at age from scales of Norwegian spring-spawning herring, a discrepancy of unacceptable magnitude was observed between the back-calculated length frequency distributions and the corresponding observed ones. The difference was too large to be explained by size-selective mortality and was probably due to methodological factors. The biological intercept was also regarded unreliable due to problems related to the determination of its position. As an alternative way of determining the intercept, it was proposed to tune the back-calculated length frequencies to fit optimally with the corresponding observed length frequency by varying the intercept and adopt this intercept as an "optimum fit" intercept.

The main aim of this study is to develop a method to estimate back-calculated length frequencies which fit the corresponding observed length frequencies with the least possible error.

Material and methods

Samples of Norwegian spring-spawning herring collected by the Institute of Marine Research, Bergen, Norway, (IMR) from commercial and scientific catches taken by drift-net, purse-seine and trawl during the period 1935 to 1994 were available for this study. The observed length frequency distributions were evaluated to be of the best quality in the data collected after 1970 and as the body length vs. scale relationship for this period was considered representative also for the period from 1935 to 1970, only the yearclasses 1973 to 1989 were included in the analysis.

Biological data such as total fish length, weight, sex, maturation stage etc. were collected from individual fish together with up to four scales from the area just behind the operculum. Each scale was cleaned by hand and attached to glass plates coated with gelatine. Scales were only collected from individuals larger than approximately 15 cm. The slides were mounted under a binocular (13X or 33X) fitted with translucent light and a measuring scale. The growth zones were counted and measured along an axis usually not exceeding about ±20 degrees off the mid axis (Fig. 1). Only the first 6 annuli were used in this study, implying that only the lengths up to age 6 were back-calculated.

Back-calculating procedure

The back-calculating formula (BCF) applied was based on the Fraser-Lee method (Fraser, 1916; Lee, 1920):

$$L_i = a + \frac{(L_c - a)}{S_c} S_i \qquad [1]$$

where:

L_i= back-calculated length at age i
L_c= length of fish at capture
S_c= radius of scale at capture
S_i= scale measurement at annulus i
a= intercept in the linear body-scale regression:

$$L = a + bS \qquad [2]$$

Statistical intercept

The statistical intercept (a in [2]) was estimated by yearclass, using linear regression of total fish length (L) on scale size (S) of all body-scale pairs available. An unweighted aritmetric mean of the statistical intercept over all yearclasses was estimated.

Observed length frequencies

Observed length frequencies by age were compiled for all yearclasses involved. Only fish caught during the period corresponding to the formation of the winter ring of the scale (November - April) were included in the analysis. Fish caught during this period were treated as belonging to the same agegroup, and consequently one year was added to the age of those caught in November and December. All available samples were included to obtain the best possible geographic spreading and inclusion of all sizes in a yearclass at a given age.

Optimum fit intercept

By varying the intercept, "a" in [1], in steps of 0.1 from 0 to 11, and calculating the corresponding back-calculated length distribution, 111 back-calculated length distributions were compiled by yearclass and age. The square deviation between each of these back-calculated length frequencies by age and the corresponding observed length frequency by age was calculated as:

$$Square\ Deviation = \sum_{X=f} (X_O - X_{BC})^2 \qquad [3]$$

,where f is the frequency range, X_O is the observed length, and X_{BC} is the back-calculated length.

The square deviation was plotted as a function of the intercept (from 0 to 11 by 0.1) and the intercept value corresponding to the minimum of the curve was found by eye. This value corresponded to the least square deviation and was adopted as the "optimum fit" intercept for that yearclass and age.

This method is sensitive to the shape and range of the observed and back-calculated length frequency distributions and the estimated "optimum fit" intercepts may be biased when large deviations in either parameter occur. Therefore, the estimated "optimum fit" intercepts were rejected in yearclasses and ages where the shape or the range spanned by the two frequencies deviated substantially. The deviation in the range spanned by the two distributions was not allowed to exceed approximately 15 %. The acceptable deviation in the shape of the two distributions was not given a numeric notion, but was merely based on a critical evaluation of the shapes of the distributions.

The 1983 yearclass was used to illustrate the effect of the different intercepts on its backcalculated length frequency distribution.

Data handling and statistics

The statistical analysis and plotting was carried out using the SAS system (SAS, 1990).

Results

The linear statistically estimated intercept by yearclass varied from 4.2 to 14.4 with a mean of 9.4 (Table 1). In most yearclasses, both the statistical intercept by yearclass and the mean statistical intercept over all yearclasses produced back-calculated length frequencies at age which were characterized by a far higher mean length than the corresponding observed length frequency distribution (1983 yearclass, age 1-6: Fig. 2-7). The discrepancy decreased by age, but was present in most yearclasses and ages.

Several of the estimated "optimum fit" intercepts by yearclass and age were rejected (Table 2) due to unacceptably large deviations either in the range (Fig. 8) or in the shape (Fig. 9) of the observed and back-calculated length frequencies. Linear regression of the accepted optimum fit intercepts (Table 2) on age indicated a significant relation:

$$\text{Optimum fit intercept} = 0.16 + 1.58*\text{age}, \qquad [4]$$

($n=44$, $p<0.001$, $r^2=0.58$). Inserting this age dependent intercept model in the backcalculating formula [1], the back-calculated length at age in Norwegian spring-spawning herring can be calculated by the following modified Fraser-Lee equation:

$$L_i = (0.16+1.58*age) + \frac{(L_c-(0.16+1.58*age))}{S_c} S_i \qquad [5]$$

Although this BCF is linear by age, it produces a non-linear growth trajectory of the individual fish (Fig. 10).

The backcalculated length frequency distributions of the 1983 yearclass from age 1 to 6 using the linear regression intercept (10.8), the estimated optimum fit intercept and the proposed model intercept [5], together with the corresponding observed length frequency distribution is plotted in figures 2-7.

Discussion

The choice of an appropriate body-scale relationship in back-calculating lengths at age in fishes is not obvious and may have consequences to the back-calculated lengths (Hile, 1970; Kang, 1979; Bartlett et al., 1984; Campana, 1990; Rubin and Perrin, 1990; Ricker, 1992). The most common method applied appears to be the linear approach, both in herring (e.g. Lea, 1910; Lee, 1920; Moores and Winters, 1982; Sjøstrand, 1992) and in other fishes (Francis, 1990). Considering the body-scale plot of the largest yearclass in this study, the 1983 (Fig. 11), both a linear or non-linear model may seem appropriate. The underside of the plot is slightly concave while the top is convex. However, the convex top is not a true convex surface but made up of two lines; the asymptotic length of the herring at about 38 cm (from scale size 5 mm and larger) and the minimum scale size at a given fish length. The downwards concave lower part of the plot reflects the maximum scale size of a given fish length and could be regarded as a biological constant at each length. It may be argued that, given the nature of the body - scale plot, a non-linear BCF would be more correct. However, the proposed back-calculating method is actually based on the principles of the linear approach, although it produces a non-linear growth trajectory of the individual fish (Fig. 10). Thus, it may be argued that the present method of back-calculation falls in between the linear and non-linear approach, which in this case makes the choice of a certain body-scale relationship less important.

The statistical intercepts overestimated the back-calculated length frequency distributions as compared to the corresponding observed ones. Important causes to these over-estimated intercepts are most probably found in the nature of the body-scale plots. The smallest herring in the samples were approximately 13 cm in total length, and the large range with no values, from 0 to 13 cm, caused relatively large confidence intervals of the estimated intercepts. The shape of the body-scale plot is furthermore particularly influenced by the asymptotic length of the individual yearclass and probably also by the fact that relatively small scales of small individuals are more easily rejected than small scales of large fishes. It was thus concluded that the present data was not suitable for statistical analysis in estimating the common intercept, neither by linear nor non-linear methods.

The precision of the "optimum fit" intercept depend critically upon certain assumptions: (1)Length dependent mortality should be negligible; (2) All lengths present in a yearclass at a given age should be proportionally represented in the observed length frequency; and (3) The fish used for back-calculation should correspond proportionally to the fish in the observed length frequency. A major criticism against the optimum fit intercept is obviously that the assumptions should be fulfilled all at a time, and it is of importance to consider the extent of deviation from the assumptions in the present data.

If size selective mortality have acted upon the yearclasses, this may have biased the estimated intercept either in positive or negative direction. The size selective mortality may either be caused by natural mortality factors or from different catchabilities of different sizes (Ricker, 1969). The yearclasses studied were distributed throughout a wide geographic range and experienced different temperature and predator regimes in their nurseries (Hamre, 1990; Røttingen, 1990). Whether this led to strong size selective mortality due to natural factors is uncertain. Barros (1995) investigated the mortality of the young herring in the Barents Sea and concluded that the most severe mortality occurred during the first year of life. After this the natural mortality decreased sharply and remained at a lower level. To demonstrate the effect of size selective mortality on the mean length of a ("non-growing") yearclass, Ricker (1969) applied a size selective mortality ranging by length from 0 to 1.5 (mean 0.75), increasing by 0.1 for each 5 mm increase in initial length. This resulted in a decrease in the mean length at 2 mm in a year. To demonstrate a decrease at 10 mm an average mortality at 3.75 was required. In comparison, the instantaneous natural mortality used for compiling the VPA in this stock is 0.13 (Anon., 1993) for all ages. Both the steep size selective mortality function used by Ricker (1969) and the high average mortality is definitely unrealistic for this stock, and if size selective mortality due to natural causes acts upon this stock it is certainly of much smaller magnitudes than those in Ricker's (1969) exercise. In conclusion, size selective mortality due to natural causes is not believed to have affected the estimated optimum fit intercepts appreciably.

The minimum landing size of the herring was 25 cm during most of the period studied, so size selective mortality due to fishing was probably of minor importance to the immature herring. Due to the nonsynchronous maturation of the two components of the stock (Dragesund et al., 1980; Holst and Slotte, 1996), some size selective mortality due to fishing may have occurred during the recruiting phase of the components. However, the instantaneous fishing mortality was at a mean level below 0.1 for these ages during the entire period studied (Anon., 1993) and also size selective mortality due to the fisheries appears negligible to the estimated intercepts.

In general, a representative stock sampling scheme will require a profound knowledge about the geographic distribution, growth patterns, components and migratory pattern of the yearclass studied. To assure that all size groups are sampled proportionally to their share of the entire yearclass, a large sampling effort will be required, which is perhaps not realistic in most cases. The observed length frequencies were collected in order to compile representative length frequencies for stock management purposes, which implied that the sampling scheme was emphasised. Despite this, groups were missing in some yearclasses and ages, which is evident when comparing the observed and back-calculated length frequencies (e.g. Fig. 8 and 9). However, given the widespread sampling strategy and large number of samples available in most of the yearclasses studied, one may be confident that in most yearclasses, all size groups are represented. It is more doubtful if all groups were sampled in proportion to their share of the yearclass.

The fish used for back-calculation were mainly caught at the spawning grounds and are therefore supposed to give a representative composition of the spawning stock. Holst and Slotte (1996) showed that the two components as defined in this stock, mix reasonably well at the spawning grounds, and there is little reason to believe that any part of the spawning stock were missed out. The main concern about the precision of the estimated intercepts was therefore that the fish in the observed and back-calculated length frequencies had been sampled proportionally to the size groups present in the individual yearclass by age. The only way to minimize this possible source of error was by selection of acceptable intercepts by age and yearclass after an objective evaluation of the shape and range of the observed and corresponding back-calculated length frequencies. In cases where the size range of the observed length frequency was evidently smaller than the back-calculated length frequency (e.g. Fig. 8), the corresponding intercept was easily rejected. Rejection was also evident when the shapes of the distributions deviated significantly (Fig. 9). It may be argued that this method for accepting or rejecting the individual intercept relies too heavily on a subjective evaluation. This was of concern during the development of this method, and preferentially better methods should be developed in order to try to reduce this source of bias and error.

Although the present method is based on a series of linear BCF's, the resulting formula falls within the family of non-linear back-calculation methods. The growth trajectory of an individual fish (Fig. 10) runs very much like a power function ($L=aS^b$), except with a negative intercept ($L=c+aS^b$, $c < 0$). However, the scale is formed when the herring larvae metamorphose at a total length of about 4 cm and the growth trajectory is not defined below this length. Thus, the growth trajectory proposed by the presented method probably has its origin close to the biological intercept as defined by Campana (1990). The degree of congruency between the biological intercept and the origin of the proposed growth trajectory is uncertain since the position of the biological intercept has not been empirically investigated in this stock. Provided such work is carried out the relative performance of the present method could be further evaluated.

Acknowledgements

I'm very much indebted to Harald Gjøsæter and Egil Ona for advice and fruitful scepticism, and to Sigmund Myklevoll and Jan Henrik Nilsen for drawing figure number 1 and 10. My work was funded by the Norwegian Research Council.

References

Anon. 1993. Report of the Atlanto-Scandian herring and Capelin working group. ICES CM /Assess:6.

Barros, P. 1995. Variable natural mortality of juvenile Norwegian spring-spawning herring (*Clupea harengus* L.) in the Barents Sea. Doctoral thesis, University of Bergen, 1995.

Bartlett, J. R., Randerson, P. F., Williams, R., and Ellis, D. M. 1984. The use of analysis of covariance in the back-calculation of growth in fishes. Journal of Fish Biology, 24: 201-213.

Campana, S. E. 1990. How reliable are growth back-calculations based on otoliths ? Canadian Journal of Fisheries and Aquatic Sciences, 47: 2219-2227

Dragesund, O., Hamre, J., and Ulltang, Ø. 1980. Biology and population dynamics of the Norwegian spring-spawning herring. *In* The Assessment and Management of Pelagic Fish Stocks, pp 43-71. Ed. by A. Saville.

Francis, R. I. C. C. 1990. Back-calculation of fish length: a critical review. Journal of Fish Biology, 36: 883-902.

Fraser, C. McL. 1916. Growth of the spring salmon. Transactions of the Pacific Fisheries Society, 1915: 29-39.

Hamre, J. 1990. Life history and exploitation of the Norwegian spring-spawning herring. *In* Biology and fisheries of the Norwegian spring spawning herring and blue whiting in the northeast Atlantic, pp. 5-39. Ed. by T. Monstad. Institute of Marine Research, Bergen, Norway.

Hile, R. 1970. Body scale relation and calculation of growth in fishes. Transactions of the American Fisheries Society, 99: 468-474.

Holst, J. C. and Slotte, A. 1996. Long term trends in the geographic spawning recruitment pattern of the Barents Sea component and the coastal component of the Norwegian spring spawning herring (*Clupea harengus* L.). Submitted to ICES Journal of Marine Sciences.

Kang, Y. J. 1979. A new method for the calculation of the growth of chum salmon, *Oncorhynchus keta*, from scale measurements. Bulletin of the National Fisheries University of Busan, 19(2): 43-54.

Lea, E. 1910. On the methods used in the Herring-

Investigations. Publications de Circonstance, No. 53: 7-25.

Lee, R. M. 1920. A review of the methods of age and growth determination in fishes by means of scales. Fishery Investigations, London, Ser. 2 4(2): 32 p.

Moores, J. A. and Winters, G. H. 1982. Growth patterns in a Newfoundland Atlantic herring (*Clupea harengus harengus*) stock. Canadian Journal of Fisheries and Aquatic Sciences, 39: 454-461.

Ricker, W. E. 1969. Effects of size-selective mortality and sampling bias on estimates on estimates of growth, mortality, production, and yield. Journal of the Fisheries Research Board of Canada, 26: 479-541.

Ricker, W. E. 1992. Back-calculation of fish lengths based on proportionality between scale and length increments. Canadian Journal of Fisheries and Aquatic Sciences, 49: 1018-1026.

Rubin, J. F., and Perrin, N. 1990. How does the body-scale model affect back-calculated growth: The example of Arctic charr, *Salvelinus alpinus* (L.), of Lake Geneva (Switzerland). Aquatic Science, 2/3: 287-295.

Røttingen, I. 1990a. The 1983 year class of Norwegian spring spawning herring as juveniles and recruit spawners. *In* Biology and fisheries of the Norwegian spring spawning herring and blue whiting in the northeast Atlantic, pp. 165-204. Ed. by T. Monstad. Institute of Marine Research, Bergen, Norway.

Sjøstrand, B. 1992. Changes in length at age in Baltic Herring, studied by back-calculation from scales. ICES C.M /J:21.

SAS Institute Inc. 1990. SAS/STAT® User's Guide, Version 6, 4.ed, Vol. 2, Cary, NC: SAS Institute Inc., 1990. 846 pp.

Table 1. Statistical intercepts, estimated by linear regression of fish length on scale size of the yearclasses 1973 to 1989.

Yearclass	N	Intercept
1973	10873	8.26
1974	6349	9.77
1975	1660	12.72
1976	4768	14.36
1977	3021	8.48
1978	3548	9.43
1979	6468	9.31
1980	946	4.18
1981	704	11.64
1982	1670	7.56
1983	28673	10.80
1984	1002	8.17
1985	2213	8.27
1986	355	5.50
1987	1241	11.82
1988	3716	11.41
1989	4238	8.67
Sum	81445	Mean 9.43
		St.dev. 2.55

Table 2. Estimated optimum fit intercepts for back-calculating age 1 to 6 of the yearclasses 1973 - 1989, with estimates of mean by age over accepted yearclasses and their standard deviation. r and s indicate rejection due to an unacceptably large deviation in the range (r) or shape (s) or both (rs) of the corresponding observed and back-calculated length frequency. m indicate missing data.

	YEARCLASS								
AGE	1973	1974	1975	1976	1977	1978	1979	1980	1981
1	m	r	r	0.0	s	rs	s	2.0	s
2	8.0	4.3	s	s	s	s	rs	rs	rs
3	s	5.3	s	s	5.2	6.0	5.1	s	rs
4	s	9.5	8.9	6.6	8.3	2.8	3.7	s	s
5	6.5	12.3	13.1	11.6	-	6.5	12.1	rs	s
6	13.0	-	-	10.0	10.0	6.3	-	-	rs

	YEARCLASS								Mean 73-89	St. dev.
AGE	1982	1983	1984	1985	1986	1987	1988	1989		
1	1.2	1.3	1.7	0.9	s	1.0	1.7	s	1.23	0.62
2	s	4.3	2.0	s	s	s	3.8	s	4.48	2.18
3	s	s	2.0	s	s	s	s	7.2	5.13	1.72
4	rs	5.5	3.7	3.1	s	s	s	s	5.78	2.63
5	s	5.0	5.7	5.9	s	s	s	s	8.74	3.40
6	9.9	6.1	9.5	s	s	s	s	m	9.26	2.39

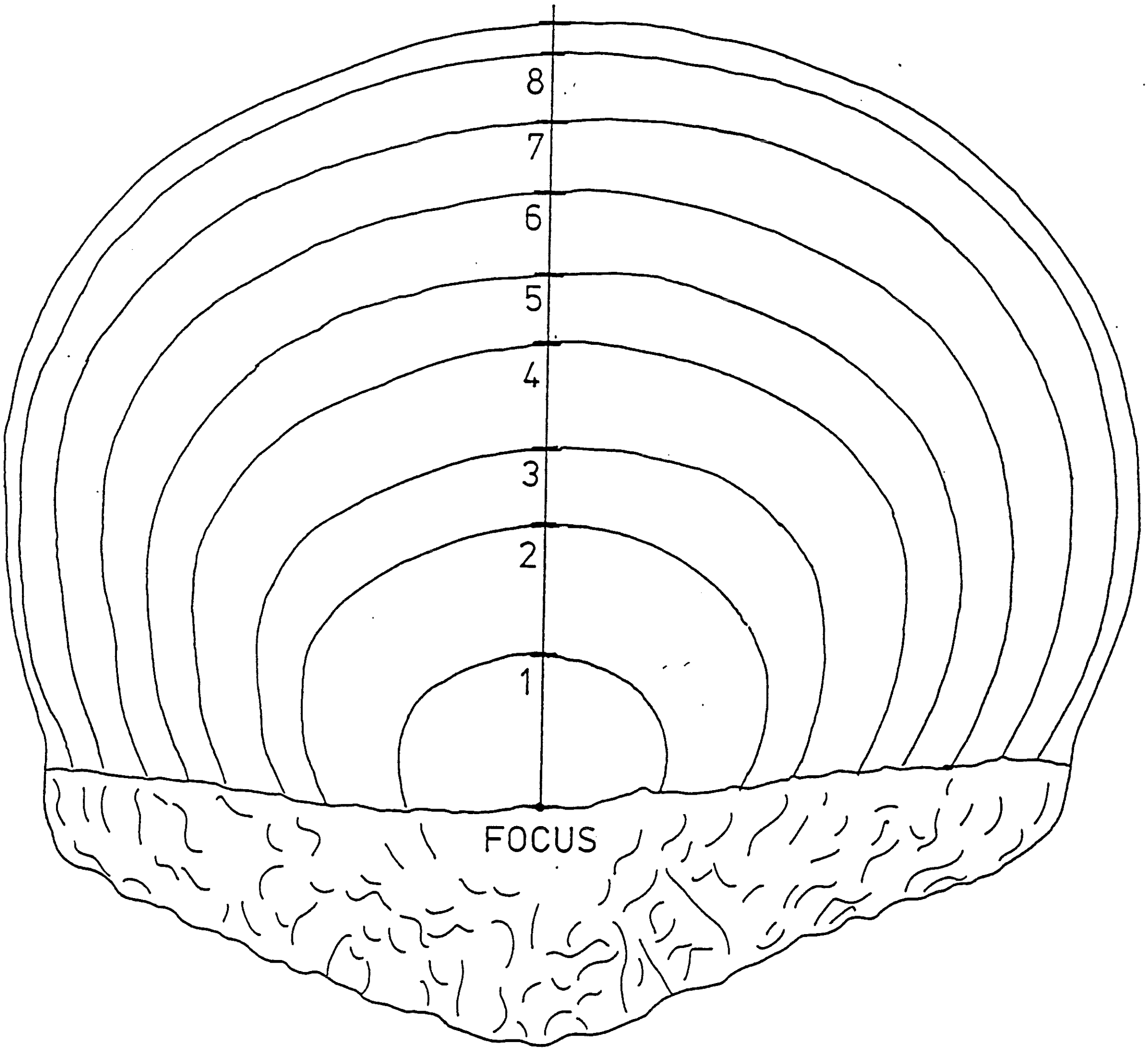

Figure 1. Schematic drawing of a herring scale from an 8-year-old individual caught in summer. Approximate line used for measuring growth increments with age of corresponding annuli indicated.

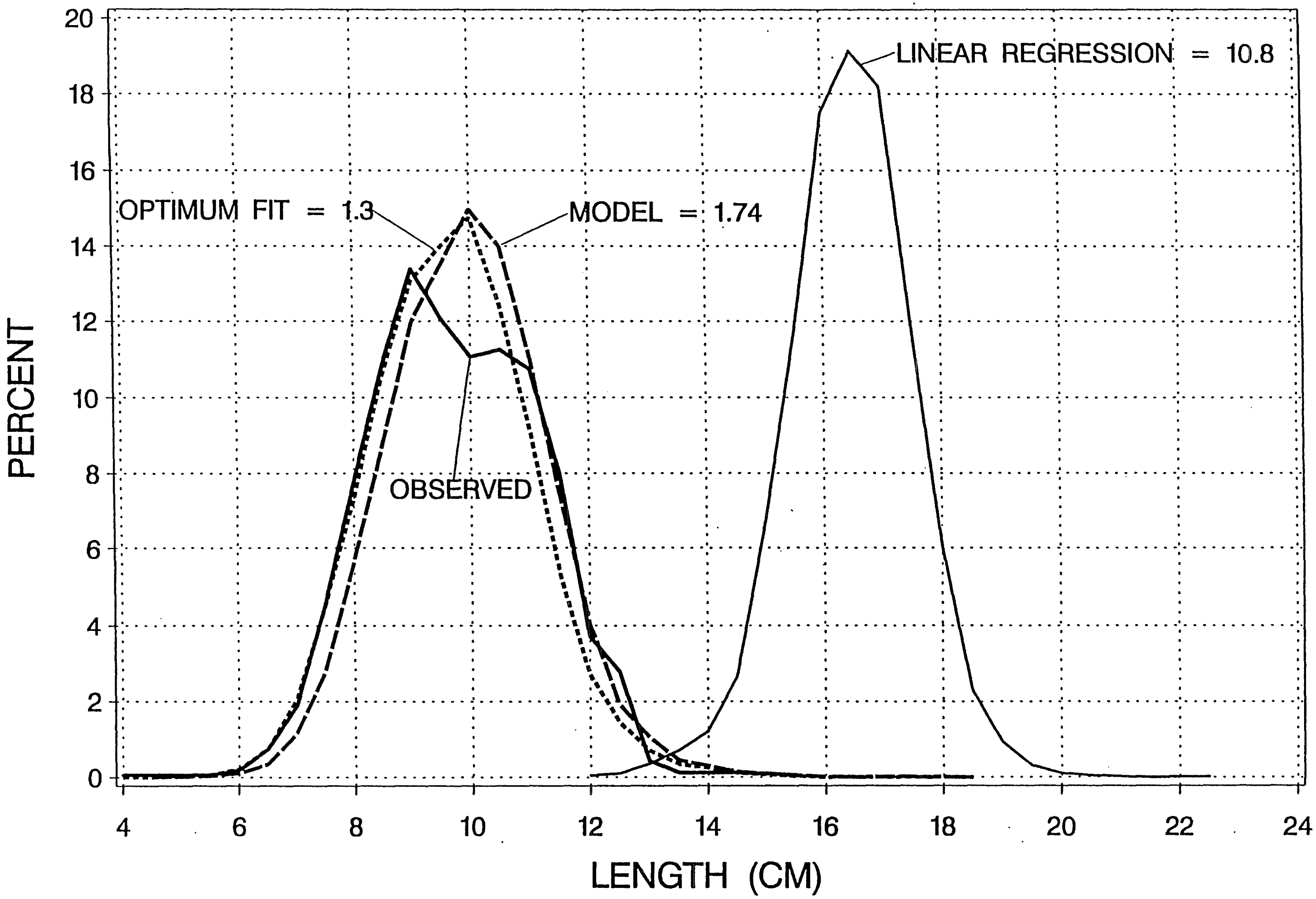

Figure 2. Observed length frequency distribution and back calculated length frequency distributions using the linear regression intercept, the optimum fit intercept and the proposed model intercept of the 1983 yearclass at age 1.

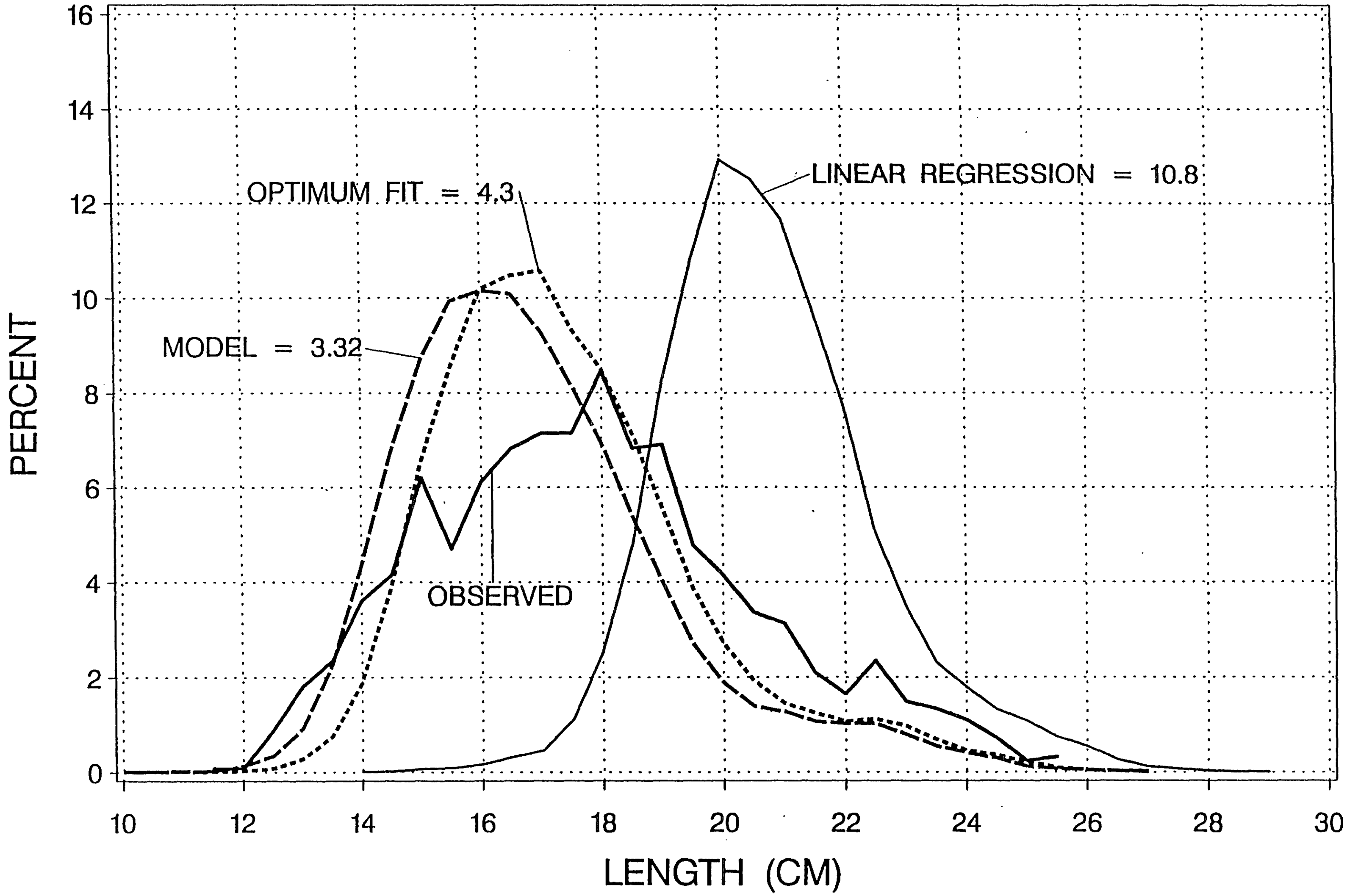

Figure 3. Observed length frequency distribution and back calculated length frequency distributions using the estimated linear regression intercept, the estimated optimum fit intercept and the proposed model intercept of the 1983 yearclass at age 2.

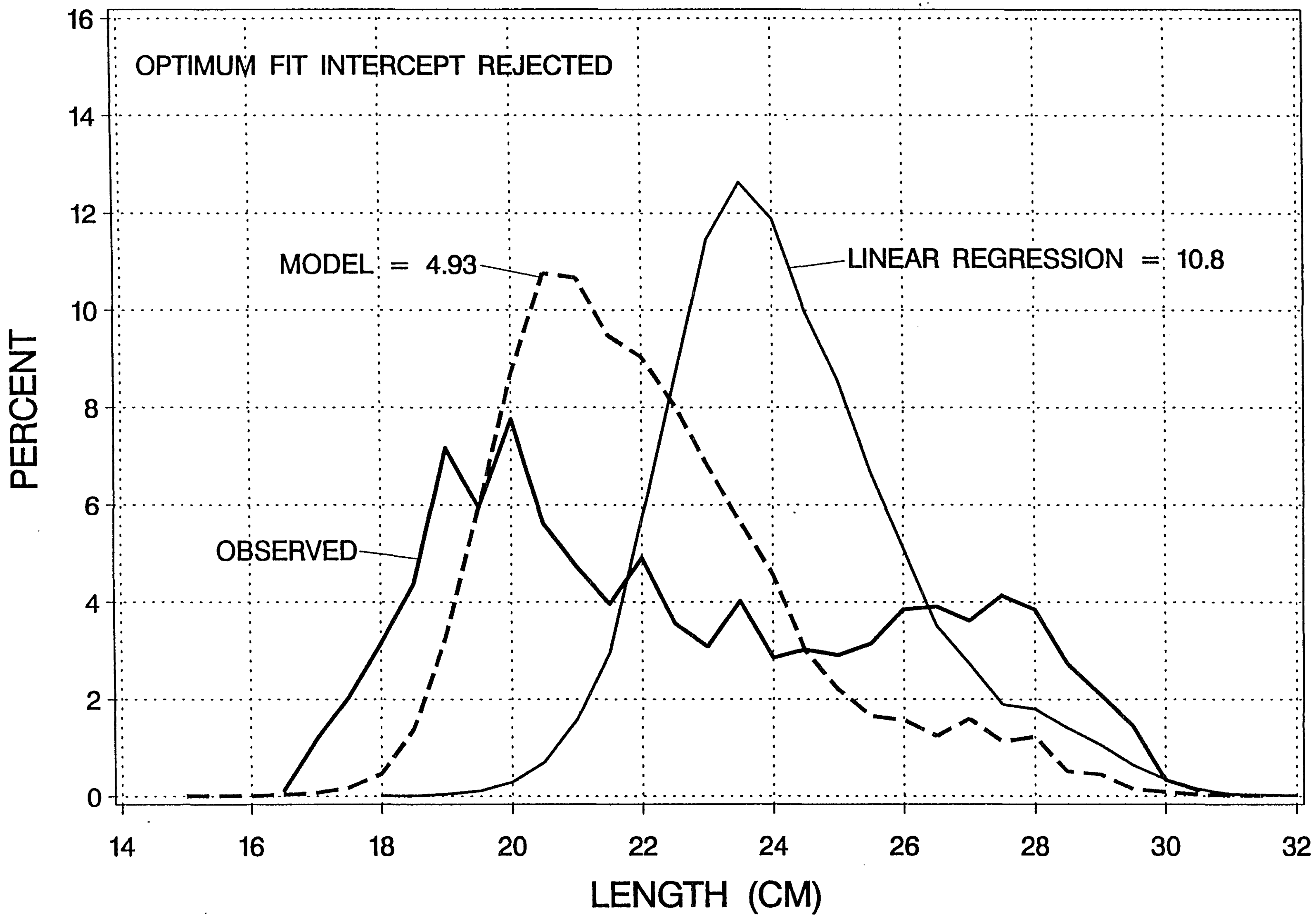

Figure 4. Observed length frequency distribution and back calculated length frequency distributions using the estimated linear regression intercept and the proposed model intercept of the 1983 yearclass at age 3. Due to rejection, the back calculated length frequency distribution based on the optimum fit intercept is not indicated.

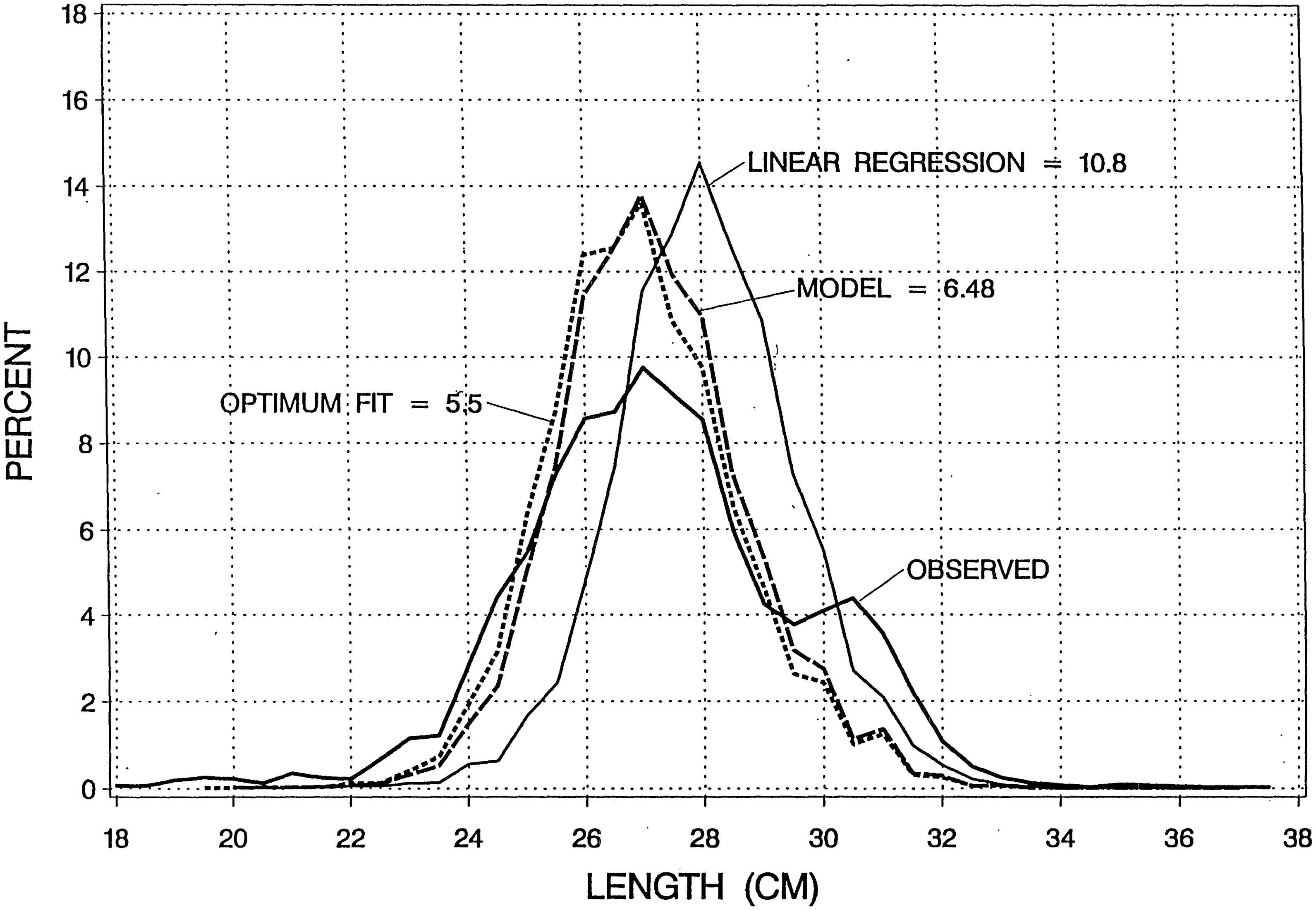

Figure 5. Observed length frequency distribution and back calculated length frequency distributions using the linear regression intercept, the optimum fit intercept and the proposed model intercept of the 1983 yearclass at age 4.

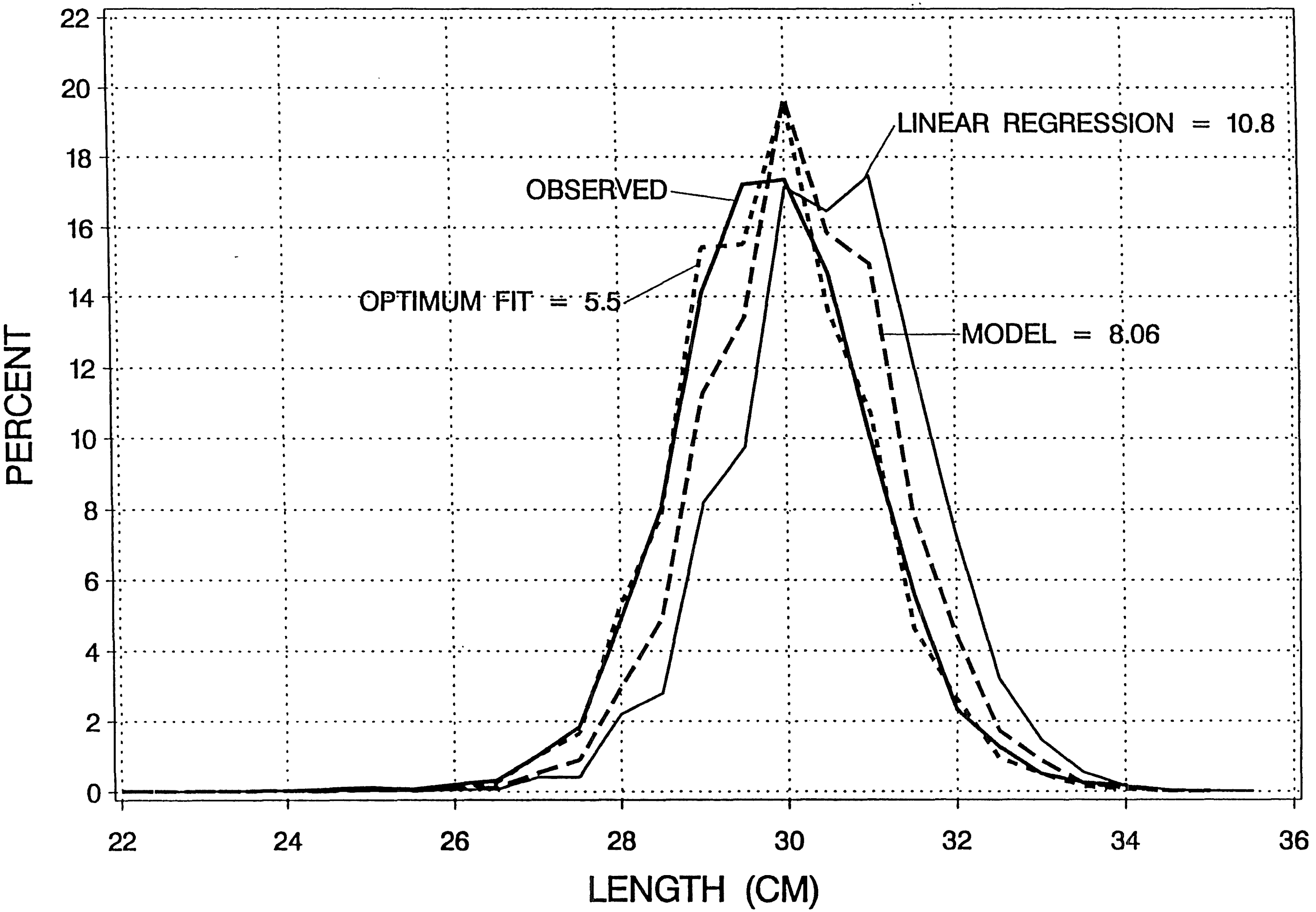

Figure 6. Observed length frequency distribution and back calculated length frequency distributions using the linear regression intercept, the optimum fit intercept and the proposed model intercept of the 1983 yearclass at age 5.

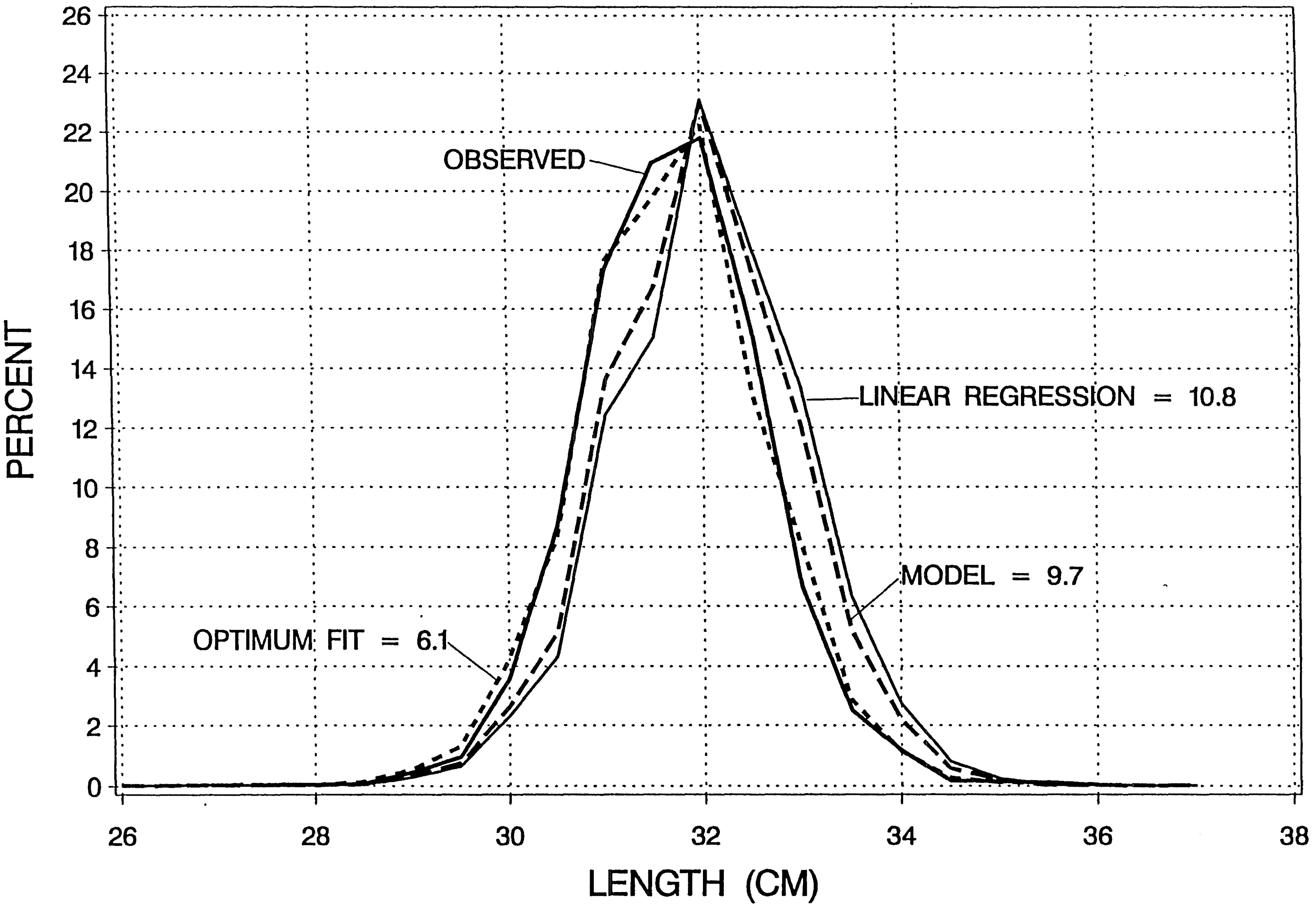

Figure 7. Observed length frequency distribution and back calculated length frequency distributions using the linear regression intercept, the optimum fit intercept and the proposed model intercept of the 1983 yearclass at age 6.

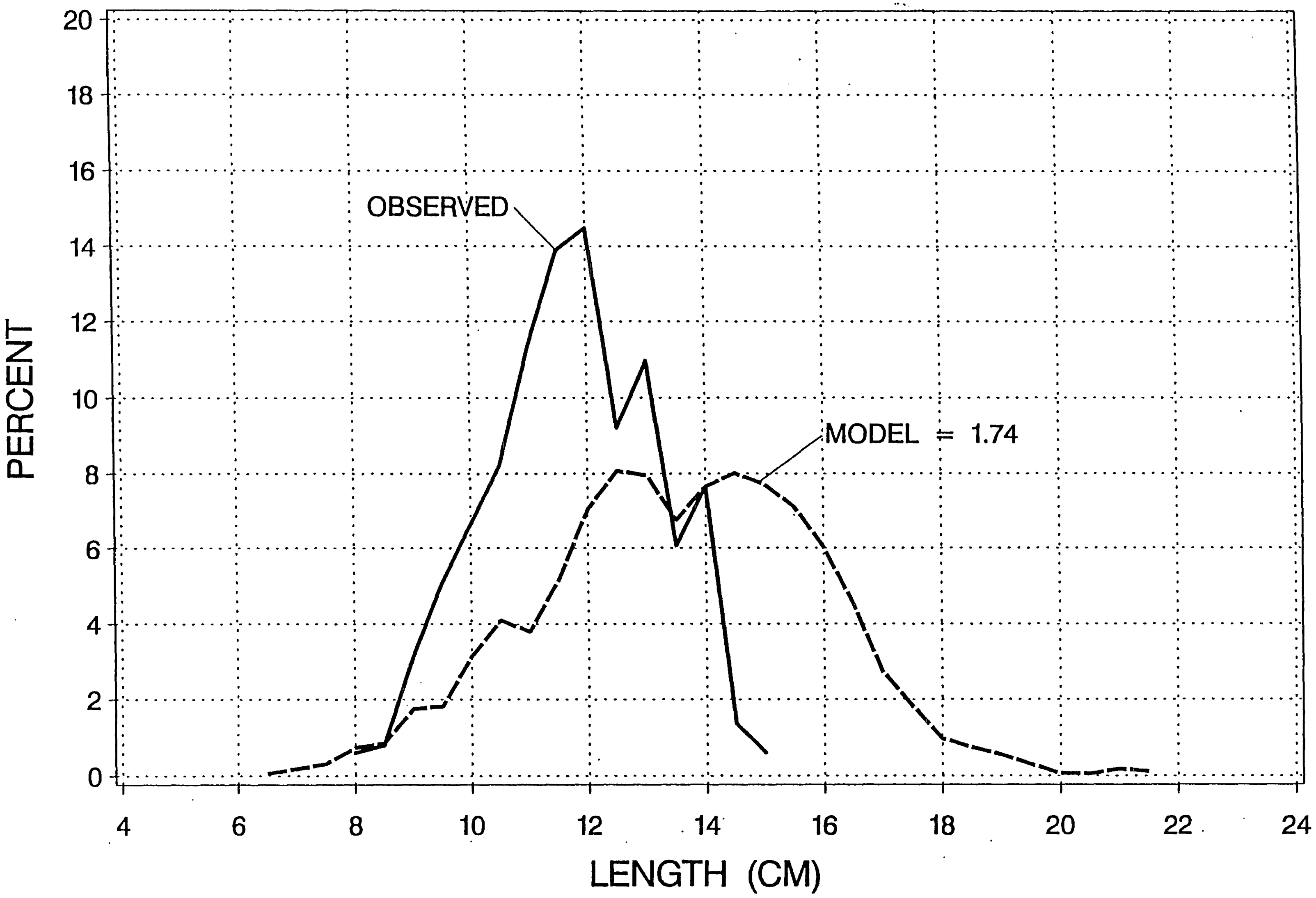

Figure 8. 1975 yearclass at age 1. Unrepresentative sampling of the observed length frequency distribution as compared to the backcalculated one. Rejection due to deviating range (code=r in table 2).

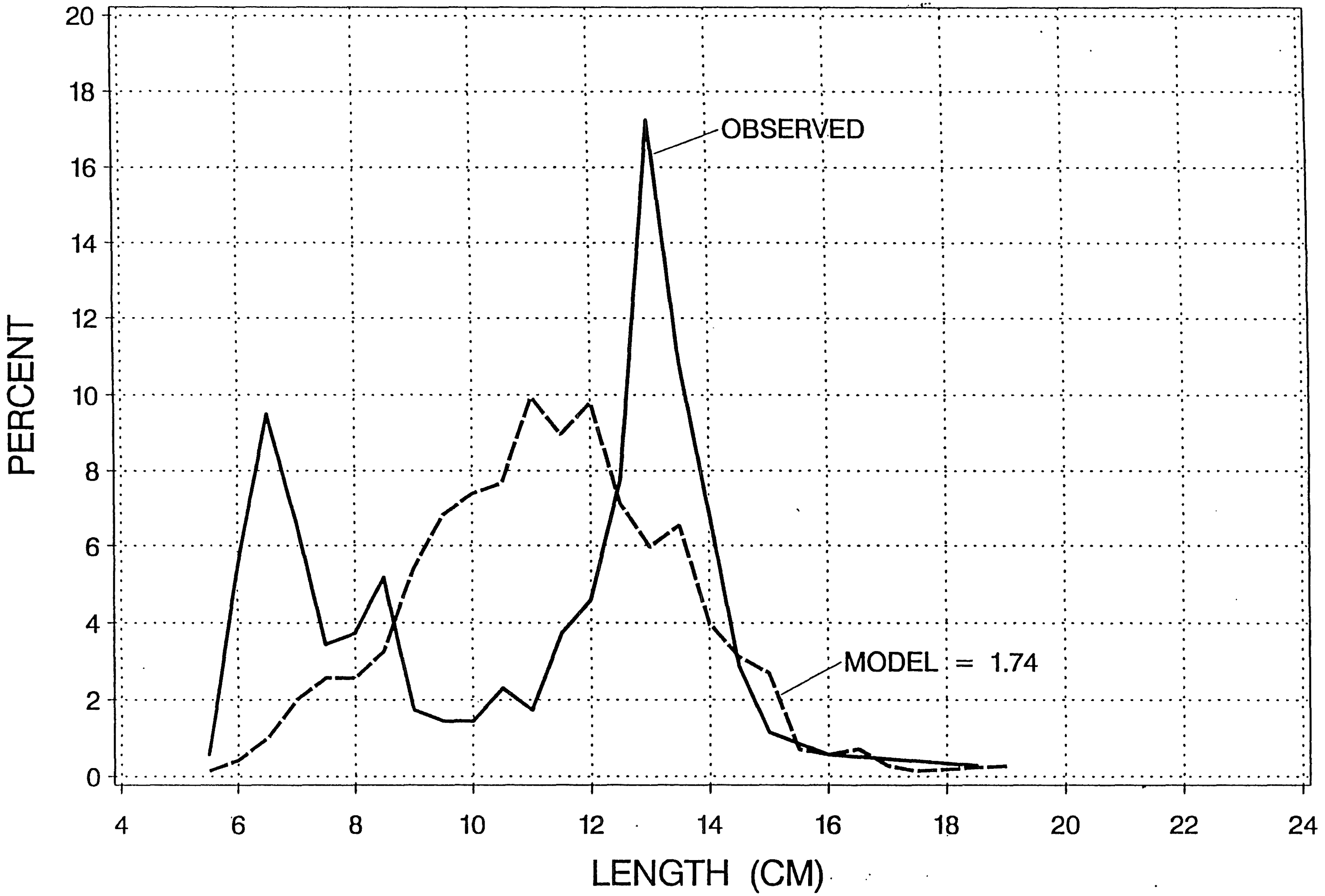

Figure 9. 1981 yearclass at age 1. Unrepresentative sampling of the observed length frequency distribution as compared to the backcalculated one. Rejection due to deviating shape (code=s in table 2).

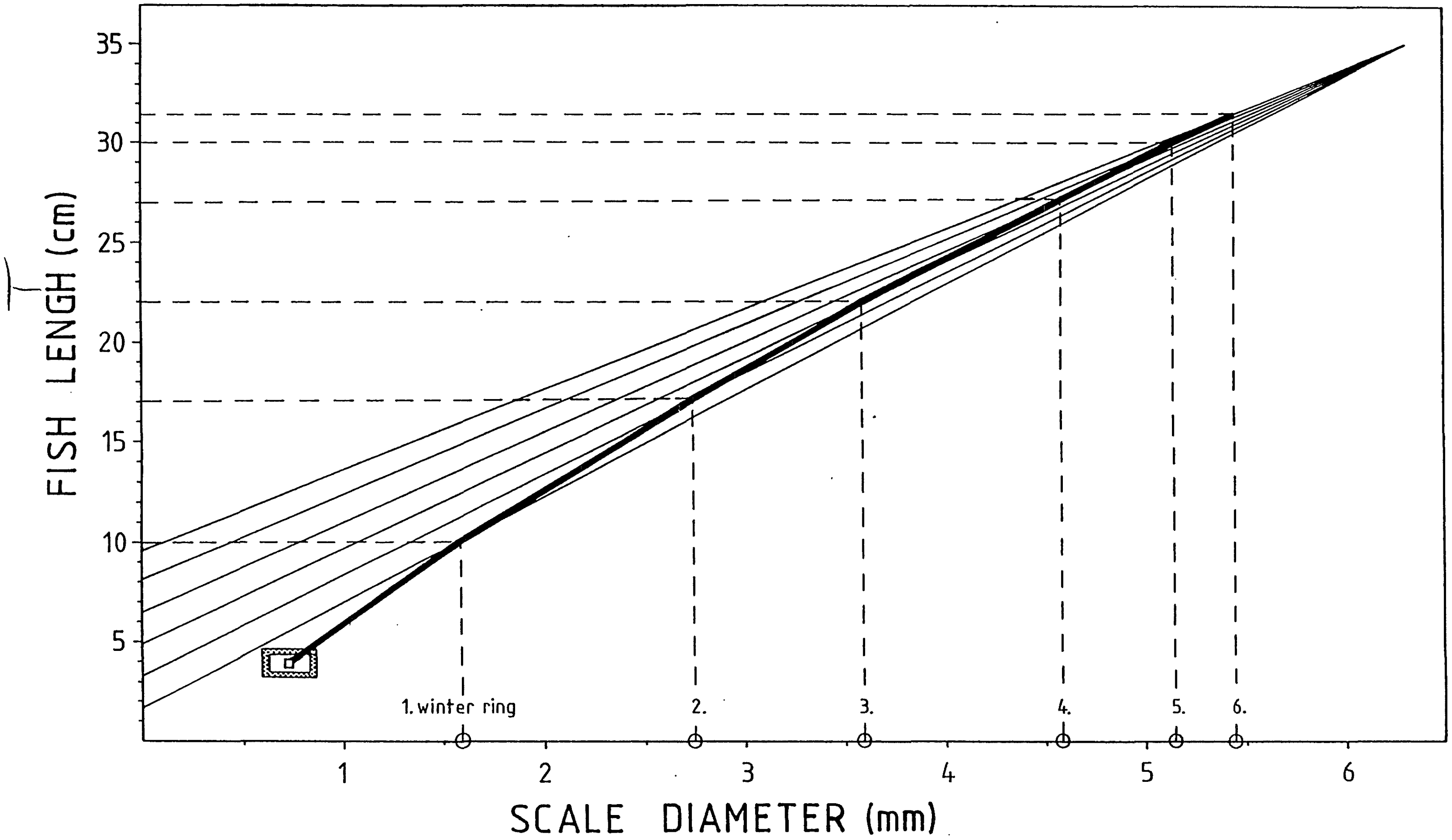

Figure 10. Growth trajectory up to age 6 of a 35 cm herring as suggested by the proposed age dependent model [5]. Back-calculated lengths at age indicated. □ = Biological intercept (or origin) estimated by interpolating the indicated growth trajectory backwards to metamorphosis. Box = Area within which the biological intercept is most probably situated, according to observed length at metamorphosis and scale diameter at formation.

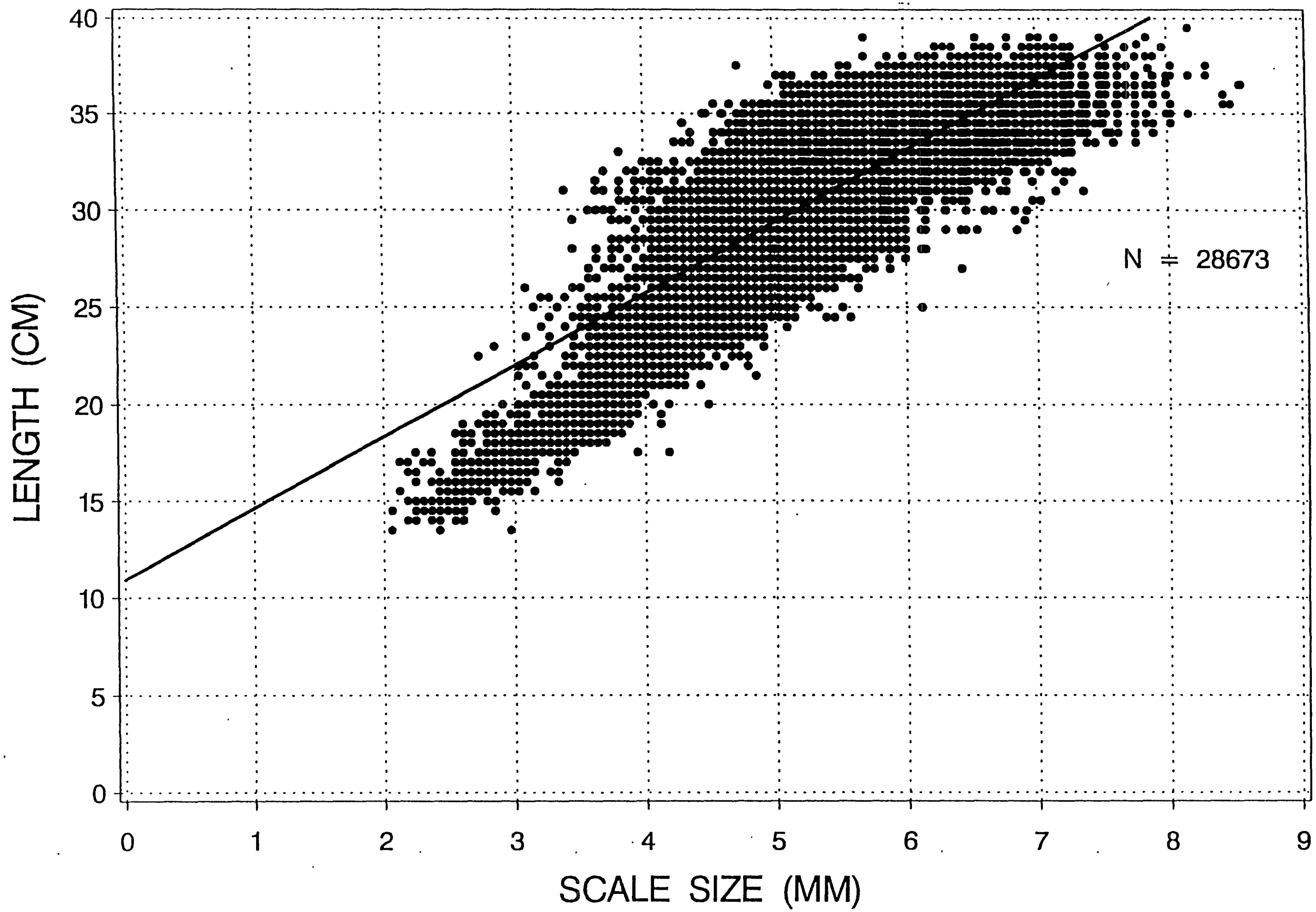

Figure 11. Body - scale plot of the 1983 yearclass. Estimated linear regression line indicated.