INFORMATIONAL LEAFLET NO. 175

STOCK SEPARATION STUDIES OF ALASKAN SALMON BASED ON SCALE PATTERN ANALYSIS

By:

Paul V. Krasnowski

and

Michael L. Bethe

STATE OF ALASKA
Jay S. Hammond, Governor
DEPARTMENT OF FISH AND GAME
Ronald O. Skoog, Commissioner
Subport Building, Juneau 99801

May 1978

STOCK SEPARATION STUDIES OF ALASKAN SALMON

BASED ON SCALE PATTERN ANALYSIS

By:

Paul V. Krasnowski

and

Michael L. Bethe

Statewide Salmon Stock Separation Project
Division of Commercial Fisheries
Anchorage, Alaska

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF APPENDIX TABLES

Page

LIST OF APPENDIX FIGURES

ABSTRACT

The Statewide Stock Separation Project was initiated in July, 1976 to research and apply new techniques of stock identification for use in mixed-stock salmon fisheries. Scales collected from sockeye, chum, coho and chinook salmon have been examined using a projection microscope at 100x magnification. Numbers of circuli and distances from the focus to annuli and supplementary checks were the commonly measured characteristics. Actual stock identification was based on pattern recognition procedures using discriminant function analysis of scale characteristics. These techniques have successfully applied to sockeye salmon (Bristol Bay, Cook Inlet, Kodiak), chum salmon (Norton Sound, Kotzebue Sound, Yukon River), coho salmon (Cook Inlet) and chinook salmon (Cook Inlet). Application of stock identification techniques based on scale pattern recognition to mixed-stock fishery management is logistically and statistically feasible.

STOCK SEPARATION STUDIES OF ALASKAN SALMON BASED ON SCALE PATTERN ANALYSIS

By

Paul V. Krasnowski, Research Project Leader

and

Michael L. Bethe, Fishery Research Biologist

Statewide Salmon Stock Separation Project
Division of Commercial Fisheries
Anchorage, Alaska

INTRODUCTION

The Statewide Salmon Stock Separation Project of the Alaska Department of Fish and Game, Division of Commercial Fisheries, was first funded for Fiscal Year 1977 (beginning July 1, 1976). Its objectives are the research, development and application of new techniques of stock identification which will permit determination of stock composition for Pacific salmon (Oncorhynchus sp.) harvested in areas where fish from more than one system are present.

For purposes of this report, stock is defined as a somewhat discrete group of fish which originates from the same river system. A "stock" may include more than one spawning group or population but, although there may be more genetic similarity within a stock than between neighboring stocks, the term is mostly a matter of convenience and does not imply a strictly genetic basis for identifying or separating these groups (Larkin 1972). Unless otherwise indicated a specific stock will refer to all the component sub-populations of a particular river drainage including all lake basins and tributaries.

Generally, the management of Alaskan salmon, operating on the principles of optimum sustained yield, is done on the basis of discrete stocks, that is, by river system. In some areas, escapement enumeration projects and statistical catch allocation techniques have provided sufficient data for development of spawner-recruit models. More commonly, only escapement estimates and rough catch figures or catch per unit of effort (CPUE) are available for determination of escapement requirements. It is on the basis of this information that escapement goals must be set. Consequently, decisions to open or close a fishery are only as good as the catch and escapement data upon which they are based. Obviously, if the commercial harvest is operating on mixed stocks, the ability to harvest the surplus of the healthy stocks and protect the stocks that may fall below escapement requirements, must be based on some stock identification technique. It is, therefore, essential to determine the proportion of each stock in the commercial harvest.

The development of techniques to identify individual stocks of fish would enable the design of sampling programs to describe the movements of each stock through time in the areas of concern to the fishery managers. Management decisions can then be implemented which, by opening and closing various district and sub-districts, optimize the harvest based on stock composition.

Although investigations into the application of x-ray fluorescence spectroscopy and protein electrophoresis for identification of Alaskan salmon stocks have been underway for several years, the Division's stock separation project has been directed toward scale characteristic analysis. There are several advantages of scale characteristic methods. Scale sampling for age determination is already an integral part of the research and management programs in many areas. Scale collection is a quick, logistically simple, and inexpensive operation even when handling live fish. Scales do not require special preservation and preparation time and expense are minimal. Finally, where scale analysis techniques are applicable to existing collections of catch and escapement scales, it may be possible to allocate to the systems of origin the numbers of fish taken in past commercial harvests and provide the background data on total return which is necessary for the determination of spawner-recruit relationships.

Scale development begins when salmon fry reach approximately 25-40 mm in length, depending on the species. Magnified images of salmon scales appear as a series of concentric rings called circuli. The different species of Pacific salmon can be recognized by their different scale patterns (Koo 1962; Bilton et. al. 1964). In all salmon species, there is an overall correlation of growth of the fish with radial growth of the scale (Clutter and Whitesel 1956). Changes in growth rate due to environmental and/or physiological conditions, therefore, are reflected in changes in the spacing of the circuli (Major and Craddock 1962; Bilton 1972; Bilton and Robins 1971a,b,c). Salmon that spend a significant portion of their life cycle in fresh water have a central portion of the scale within which the spacing of the circuli is more compact than that in the outer portion. This is referred to as the freshwater growth zone (Figure 1). Outward from the freshwater zone, the circuli are typically widely spaced, probably reflecting the rapid

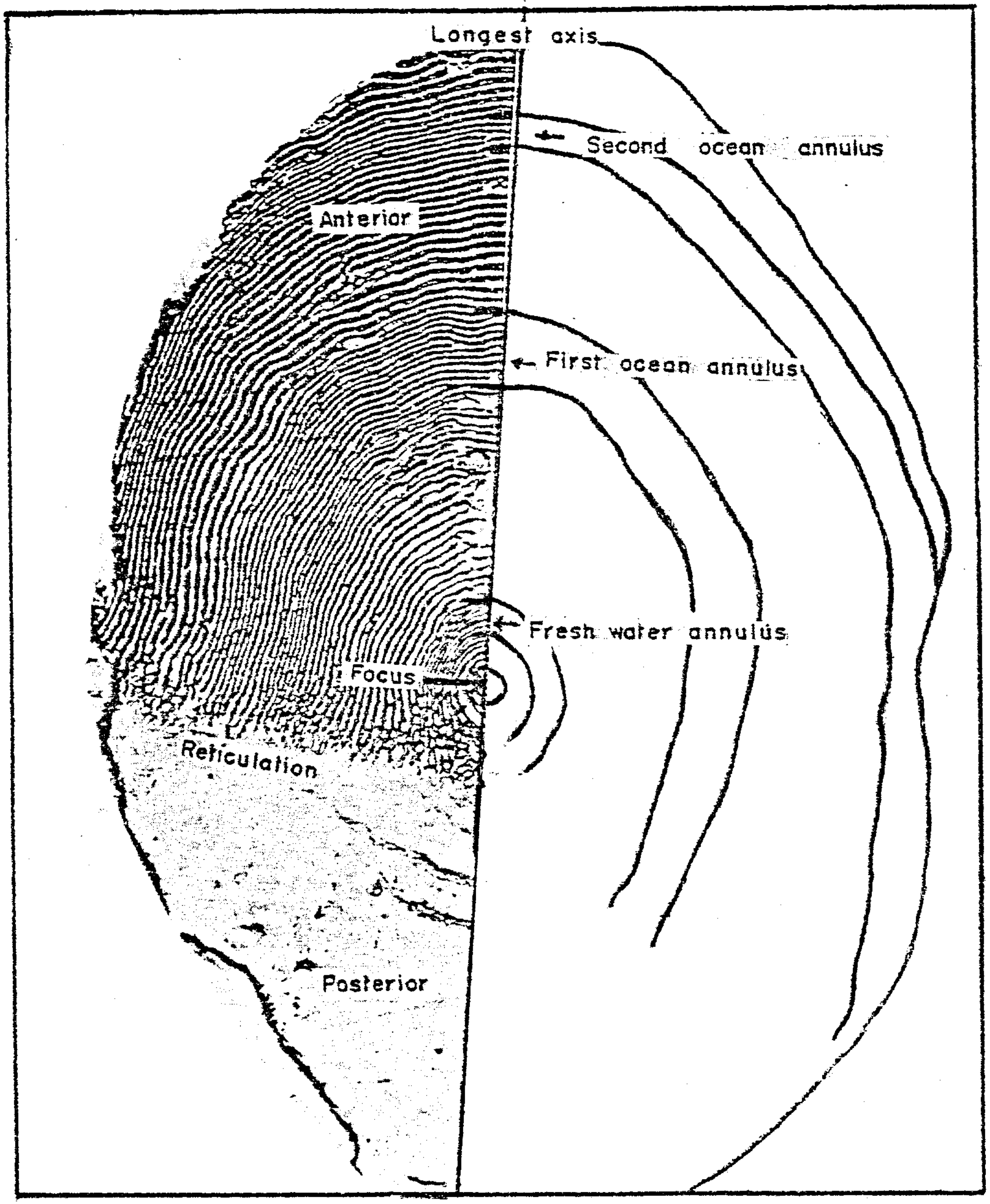


Figure 1. Age 4_2 sockeye salmon scale showing location of the various scale characteristics used in discriminant analysis (after Bilton, 1964).

growth which occurs early in the marine life stage. Ocean circuli are generally more broadly spaced than freshwater circuli since most of the rapid growth occurs in the marine environment. Additionally, bands of very closely spaced circuli, many of which are branched or broken, occur in both the freshwater and ocean zones of scales. The close spacing of circuli in these annular rings or annuli is the result of extreme environmental changes due to winter weather and their effect on growth (Bilton and Messinger 1975). Determination of fish age by reading annular rings is common practice in many areas.

Growth in general, and scale growth specifically, are genetically moderated, environmentally influenced and are recorded in the scale pattern. Differences in scale patterns between fish from different systems have been noted for various scale characteristics. Commonly, the number of circuli and the width of the scale for each year's growth have been examined. Much of the research based on scale pattern analysis has been aimed at allocation of the harvest of salmon by the Japanese high seas fishery to continent of origin (Pearson 1964; Mason 1967; Mosher 1963, 1972; Anas and Murai 1969). These researchers have examined scale patterns of pink (*O. gorbuscha*), chinook (*O. tshawytscha*), and sockeye salmon (*O. nerka*). Research conducted in Canada has described recognizable differences between Canadian and North American stocks of pink and sockeye salmon (Bilton 1970, 1971; Bilton and Messinger 1975). In addition, researchers have attempted to identify stocks harvested within inshore fisheries (Wright 1965) as well as attempting identification of sub-stocks occurring within complex river systems such as the Skeena (Bilton and Smith 1969) and the Fraser River (Henry 1961).

In general, most research has involved groupings of stocks from large geographical areas. Applications of these techniques to inshore fisheries in Alaska has been attempted in some areas (Wright 1965; Bergander 1977). However, in most cases the differences between populations are not sufficient to use with standard measuring and analytical techniques. In such cases, larger data bases must be constructed and more involved, multivariate statistical methods utilized to recognize possible scale pattern differences between stocks.

The development in Canada of a reasonably-priced, high resolution projector and semi-automated measuring and data encoding equipment designed for use on fish scales has provided the necessary means to generate the required data. Availability of high speed data processing through the University of Alaska Computer Network has provided the capability for rapid analysis of scale characters.

The Statewide Salmon Stock Separation Project presently consists of two permanent fishery biologists and from one to four temporary technicians.

From July 1976 through December 1977, more than 25,000 scales from sockeye, chum (O. keta), coho (O. kisutch), and chinook salmon have been processed. Geographical areas of concern have included Norton Sound, Kotzebue Sound and the Yukon River, Kodiak and the Alaska Peninsula, Cook Inlet, and Bristol Bay.

The following report covers the sampling, sample processing, measurement and analysis of the data and includes general summaries of the results by area and species.

MATERIALS AND METHODS

Scale Collection and Processing

An attempt was made to have all scales collected from a preferred area on the left side of the body below the insertion of the dorsal fin and two or three rows above the lateral line (INPFC 1961). Scales were mounted on gum cards and impressions of the scale surface were made on cellulose acetate cards using methods similar to those described by Clutter and Whitesel (1956). Initial examination and aging was accomplished by using a portable microfische reader. Ages are described using Gilbert-Rich notation.

Scale Examination

Scales were projected onto a table surface utilizing equipment similar to that described by Bilton (1970) and later modified by Peter Ryan of the Canadian Fisheries and Marine Service (Ryan and Christie 1975). Photographs of the scale projection and measuring equipment are presented in Appendix Figures 1-8.

The basic projector is a Leitz Micro-Promar projection microscope equipped with a wide-field eyepiece and plano objective lenses (Appendix Figures 2-3). A high contrast image is achieved by use of a Prado Universal 250 watt quartz iodine lamp. The microprojector and lens system, used in conjunction with an overhead adjustable mirror and frame assembly, is designed to deliver a flat, undistorted image to the table surface. The table surface is constructed of flat white formica to enhance the contrast of the projected image. High and low magnifications can be achieved by selecting different combinations of occular and objective lenses and adjusting the height of the overhead mirror.

All adult salmon scales were projected at a magnification of 100x. To ensure continuity of measurements, magnification at the table surface was calibrated at frequent intervals using a gridded millimeter microscope slide.

After a scale was selected for measurement, the image was projected onto a large sheet of white bond paper which is pre-printed with nine axis lines (Appendix Figure 4). For sockeye, an axis line was oriented such that it intersected the center of the nuclear area and lay along a radius which was 20° ventral to the anterior - posterior axis of the scale (Clutter and Whitesel 1956). For other species, the longest scale axis was selected. Selection of these axes afforded examination of the longest portion of the scale that had the fewest broken or branched circuli. Where each circulus crossed the axis, a tracing approximately one inch long was made on the paper. Only those circuli which continued more or less intact within a distance of 0.1 mm on both sides of the 20° ventral axis (i.e., within a distance of 1.0 cm on the projected image) were counted. Selection criteria were adapted from Bilton (1971) and are further detailed in Appendix Table 1.

Digitizer and Linear Encoder

To enable rapid generation of digital output of scale measurements in computer readable format, a digitizer and linear encoder were employed. This equipment is described in Ryan and Christie (1975) and shown in Appendix Figures 5-8.

The digitizer unit consists of electronics which format and output scale data in computer readable, fixed format (ASCII). Sample identification data is input via a twelve digit code representing location of sampling, scale card number, length of fish, sex, year sampled, and age. The linear encoder is connected to the digitizer and facilitates entry of measurements via a remote control. The encoder is oriented parallel to axis lines and the sliding Rouchi rule is zeroed at the nuclear portion of the scale drawing. Depression of the "record" button on the remote control panel initiates the automatic recording and formatting of all data on to a typed sheet and to punched paper tape on an ASR33 teletype. The index of the rule is then moved across the scale to the end of the first annulus or to the first characteristic of concern while manually depressing the count button for each circulus passed. When the index is aligned, the record button is depressed causing the interval measurement and the circuli count to be recorded on tape as described above. The digitizer is capable of recording seven paired measurements (numbers of circuli and distance from focus) per record.

At present, data are either keypunched or paper tapes are processed by the University of Alaska, Geophysical Institute in Fairbanks for conversion to NOVA magnetic tape and subsequent data entry. On-line terminal - computer data entry has been necessary during field operations. Acquisition of a paper tape reader will allow direct entry of raw data to computer file when interface of the equipment is complete this year. Format of data output from the digitizer is shown in Appendix Figures 9-11.

As described above, circuli counts and distance measurements to each point of interest are paired. The data output are cumulative, representing total circuli and distance from the scale origin (focus). An editing program has been developed which eliminates records with obvious errors in the fixed data (e.g., sex code, district code, age) and provides general editing criteria to be applied to the scale data. Additionally, the output of this program converts the measurements from cumulative to incremental and compacts the data. Examples are provided in the Appendix. These edited data are in FORTRAN format (3F2.0, F3.0, F1.0, F2.0, 12F4.0) where the column assignments are:

Variable	Column Assignment
District Code	1 - 2
Card (AWL#)	3 - 4
Year	5 - 6
Length (mm)	7 - 9
Sex Code	10
Age	11 - 12
NC1	13 - 16
ID1	17 - 20
NC2	21 - 24
ID2	25 - 28
NC3	29 - 32
ID3	33 - 36
NC4	37 - 40
ID4	41 - 44
NC5	45 - 48
ID5	49 - 52
NC6	53 - 56
ID6	57 - 60

where NC_i = the number of circuli along the selected axis for the ith characteristic

ID_i = the interval distance along the selected axis for the ith characteristic

Most of the research to date has dealt with scales from adult sockeye salmon. In most cases, measurements of sockeye scales have been from the focus to the outside of the first annulus (A_1) and thence from the outside edge of each annulus to the outside of the succeeding annulus. However, there are many different characteristics that can be used. The measurements as used in the sockeye and described above, are referred to as standard measurements. In other cases (e.g., sockeye smolt and adults of other species) other characteristics have been used. These have been further described in Appendix Table 1.

Statistical Techniques

Linear discriminant function analysis was developed because of a need to distinguish statistically two or more groups, and is based upon work done by R.A. Fisher (1936). Applications of the technique to biological data have developed rapidly in recent years due to the advent of digital computers. The utility of discriminant function analysis as applied to stock separation based on scale characteristics is based on the concept that two groups (stocks of fish in our case) may differ slightly in the mean and distribution of values for some measureable characteristics (e.g., circuli counts and radii). Characteristics taken singly and measured from samples of two populations may be useless for identifying group membership since, despite slight differences in means, the degree of overlap of values between the two groups is generally so great as to render the individual characteristic useless for discriminating between groups.

Because no single characteristic will allow identification of group membership, discriminant analysis techniques attempt to do this by using a multivariate approach which combines variables to yield "discriminant functions" which serve to better identify membership. These functions may be linear or nonlinear. Since scale measurements have been found to generally satisfy the required assumptions for linear discriminant analysis (Anas and Murai 1969; Cook, personal communication) and since software for the University of Alaska Honeywell computer provides linear discriminant function analysis routines, the work described in this report utilizes the linear methods.

The analysis first requires measurements from samples of known group membership. These samples, also called standards or learning samples, provide the data required to formulate the discriminant function, essentially the discriminating model. The program selects the discriminating variables in a stepwise fashion. The order of selection of variables for inclusion into the analysis reflects the relative between-group variability of each characteristics, i.e., their relative discriminating ability. The discriminant functions are of the form:

$$D_i = d_1 z_{i1} + d_2 z_{i2} + \dots + d_p z_{ip},$$

where D_i is the discriminant score for the ith scale, $d_1, d_2, \dots, d_p$ are weighting coefficients and $z_{i1}, z_{i2}, \dots, z_{ip}$ are standardized values of the measurements from the ith scale. In other words,

$$z_{ij} = \frac{x_{ij} - m_j}{s_j}$$

where x_{ij} is the value of the jth measurement from the ith scale, m_j is the mean of the jth measurement for all scales, and s_j is the standard deviation of the jth measurement, again for all scales.

The discriminant functions define p-dimensional hyperplanes which cut across the intermixed clusters of points so that as many as possible of the members of one group have high values of D_i and most of the other group members have low values of D_i. The weighting coefficients $d_1, d_2, \dots, d_p$ are calculated so that the discriminant scores D_i are standard normal variables, and the mean discriminant scores for all scales is zero, with a standard deviation of one.

For each group taken singly, the mean of the discriminant scores for all its members is called the centroid and describes the most probable location of that group in discriminant function space. The distance between the group centroids is an indication of the distance by which two groups are separated (again, along these dimensions in hyperspace). The midpoint between two centroids ($D_{0.5}$) serves as a decision point and unknowns can be classified as to probable group membership based on which side of $D_{0.5}$ they fall. In practice, the SPSS (Nie et. al. 1975) and BMD (Dixon 1965) programs used in the scale analysis project also output classification functions, one equation for each group, which are more convenient for classifying unknowns.

In cases where classification of more than two groups is required, the problem of visualizing the discriminant functions becomes more difficult. The number of discriminant functions generated is equal to the number of original discriminating variables (p) or to one less than the number of groups (g-1), whichever is less. Generally, the (g-1) limitation has been used in these analyses. Each resulting discriminant function is orthogonal (at right angles) to the previous functions and the resulting discriminant scores are taken to be (g-1) dimensional descriptions of the locations of the g groups in discriminant function space.

To test the efficacy of the analysis, a trial classification is made using the above mentioned classification function and the standards (learning samples). The classification equations are of the form

$$C_i = c_o + c_1 x_{i1} + c_2 x_{i2} + \ldots + c_p x_{ip}$$

where the C_i are the classification scores which are the sum of the c_j's (classification coefficients) times the raw variable values (x_{ij}'s) plus the classification constant c_o. Since there is one classification function for each group, there will be g classification scores for each case classified. An unknown case is classified as a member of the group in which it has the highest score. Since the actual group membership for each case is known, the results of this trial classification can be summarized and tabulated as below:

Actual Group Membership		Classified Group Membership: Group A	Group B	
Group A	number	Aa	Ab	N_a
	proportion	P_{aa}	P_{ab}	1.0
Group B	number	Ba	Bb	N_b
	proportion	P_{ba}	P_{bb}	1.0
		K_a	K_b	

where Aa and Bb are the numbers of their respective groups that were correctly classified and Ab and Ba are the numbers incorrectly identified. Furthermore, the proportions correctly and incorrectly classified are taken as estimates of the probabilities of classification accuracy and classification error, i.e., P_{aa} is the estimated probability of correctly classifying an unknown sample that is actually a member of group A, whereas P_{ab} is the estimated probability of misclassifying an unknown sample that is actually a member of group A as a member of group B. Given equal sample sizes from each group and normal distributions of discriminant scores differing only in mean values, the apparent numbers in groups A and B (i.e., the sum of the correctly classified members plus the misclassified members of the other group) should be approximately equal. If not, both SPSS and BMD have options to make a priori adjustments which affect the probabilities of group membership and can serve to equalize the misclassification errors. However, a thorough understanding of the affects of manipulating a priori probabilities is needed before attempting this.

When classification of an actual mixed sample of unknown composition (group membership proportions) is completed, the results represent the apparent or observed numbers and frequencies of each group (similar to the a and b above) and represent both the correctly classified members of each group plus the misclassified numbers of the other group. Since the trial classification of the learning samples provides the estimates of the probabilities associated with correct and erroneous classification, the observed frequencies can be adjusted with these probabilities to estimate the actual proportions present in the sample (Worlund and Fredin 1962). This adjustment procedure is accomplished through the solution of a set of simultaneous equations. Since the observed number of each group in the mixed sample is the sum of the correct decisions for that group plus the incorrect decisions for the other groups, and since in each case these decisions are the product of classification probabilities (P_{jk}) and the actual number of fish from that group in the sample, a series of equations can be constructed. For a three-group situation these equations are:

$$K_a = P_{aa}N_a + P_{ab}N_b + P_{ac}N_c,$$

$$K_b = P_{ba}N_a + P_{bb}N_b + P_{bc}N_c, \text{ and}$$

$$K_c = P_{ca}N_a + P_{cb}N_b + P_{cc}N_c,$$

where

K_a, K_b and K_c are the numbers of fish classified to each system;

N_a, N_b and N_c are the estimated numbers of fish from each system in the mixed sample (unknown); and

P_{jk} are the proportions of fish from system k classified as from system j (known -- estimated from the training sets).

Confidence interval estimates for the two group classification model are given by Worlund and Fredin (1962). However, these estimates assume that the P_{jk} are known without error and therefore the intervals are too narrow. A method for calculation of confidence intervals for g groups which takes into account the variability of P_{jk} is in the final stages of development and will be used with future classifications based on scale characters.

Required sample sizes for both the analysis and the classification aspects are not yet well worked out. As with most sample statistics the variance of estimates decreases with increased sample size. Some investigators have recommended sample sizes of 50 minimum for each group standard. Rod Cook (Fisheries Research Institute, personal communication) has used as few as 25. In general, the results reported here are based on a minimum sample size of 50 scales.

For classification of a mixed sample we recommend a minimum of 100 scales. The variance of the estimated frequencies will be inversely proportional to the size of the sample classified.

There is a bias associated with the estimates of classification accuracy and error, the P_{jk}, which is due to the fact that the classification functions are not tested with additional random samples of each group but are tested with the same cases that were used to generate functions. Therefore, these data will fit the models slightly better than might be expected. Some initial research has indicated that this self-classification bias might be as much as +4-6%, with the sample sizes we have been using. The bias can be avoided by collecting additional samples of knowns which are used to test the classification function but which are not used to generate the functions. There are other methods to avoid the bias, but we have not yet incorporated them. If separate samples are used, we recommend equal sample sizes of 25 or more.

RESULTS AND DISCUSSION

To date more than 25,000 scales have been measured, digitized, edited and analyzed. The data have provided several hundred discriminant analyses which cover four of the salmon species (sockeye, chum, coho, chinook) and five geographical areas of Alaska. Approximately six thousand of these scales were collected in conjunction with "in-season" stock separation projects conducted in Cook Inlet and Bristol Bay during 1977. Detailed analyses which will provide information on timing, distribution, and catch allocation of component stocks are underway and will be described in later reports. The results summarized below were derived from scales collected as routine samples from on-going management and research projects prior to the development of the Stock Separation Project. Scale measurements of chum, coho and chinook were non-standard and are detailed in the Appendix.

Bristol Bay - Sockeye Salmon

Prior to initiating in-depth discriminant analyses, a series of analyses of variances were performed on scale characteristic measurements from Bristol Bay sockeye salmon collected for routine escapement samples in 1970 through 1975. In general, for individual river systems, fish of the same freshwater age and brood year but of different ocean ages have very significant differences ($p < 0.01$) for all variables. Similarly within stocks, there are very significant differences between fish of differing freshwater and total age for all variables. In addition, there frequently are significant differences between sexes within years, age classes, and river systems. The indications, then are that new standards (learning samples) should be developed for each system, year and age class examined using discriminant analysis. Although identification of group membership would probably be further enhanced by examining each sex independently, the increased sample sizes required would make the small gain in accuracy very costly.

For the Naknek and Kvichak Rivers (Figure 2) scales from age 4_2, 5_3, and 6_3 fish sampled in 1970 through 1975 were examined. For most years, overall classification accuracies were above the mid-80% level. Discriminant analysis of data from these same years from Naknek, Kvichak, and Ugashik age 5_3 fish (3-way analyses) produced classification accuracies in the low 70% range. Egegik and Ugashik, for all age classes, appear separable with overall accuracies varying between 80% and 85%. A three-way analysis of Naknek, Kvichak and Egegik, produced overall accuracies that were quite variable, ranging from about 60% to 80% accuracy. Naknek, Kvichak, Egegik and Ugashik (four-way discriminant analyses) produced overall accuracies in the low 70% range with Naknek being the least separable (largest misclassification error), and Kvichak and Ugashik being the most distinctive. In general, it appears that in most years there are enough distinct differences between all systems on the east side of Bristol Bay and for all the major age classes (particularly age 5_3 and 6_3 fish) to provide an effective tool for stock identification.

In general, the systems on the west side of Bristol Bay produce fish that spend one year in freshwater (4_2 and 5_2). For Wood River and Nuyakuk River, there were sufficient scales to examine one year of data from each age class. The data from age 4_2 and 5_2 (1972) fish yielded overall accuracies in excess of 90%. Small samples from age 5_3 and 6_3 (1971) fish produced overall accuracies in the low 70% range. Several years data from Wood, Nuyakuk and Igushik Rivers were compared in a three-way discriminant analysis (age 5_2). The results were highly variable ranging from a low of 41% to highs of approximately 90% overall classification accuracy. Nuyakuk and Igushik scale characters are frequently quite similar.

Despite the size of the data base, there are only a few age classes and years in which there are sufficient data to compare more than four systems at a time. Comparison of Naknek, Kvichak, Ugashik, Wood and Nuyakuk (age 4_2) produced 68% overall accuracy. However, the Ugashik and Wood fish were classified accurately only in 48% of the cases, whereas Nuyakuk showed no misclassification error. Samples from only one year provided sufficient data to compare age 5_2 and 6_3 fish from Naknek, Kvichak, Igushik, Wood and Nuyakuk Rivers. The data yielded 51% overall accuracy

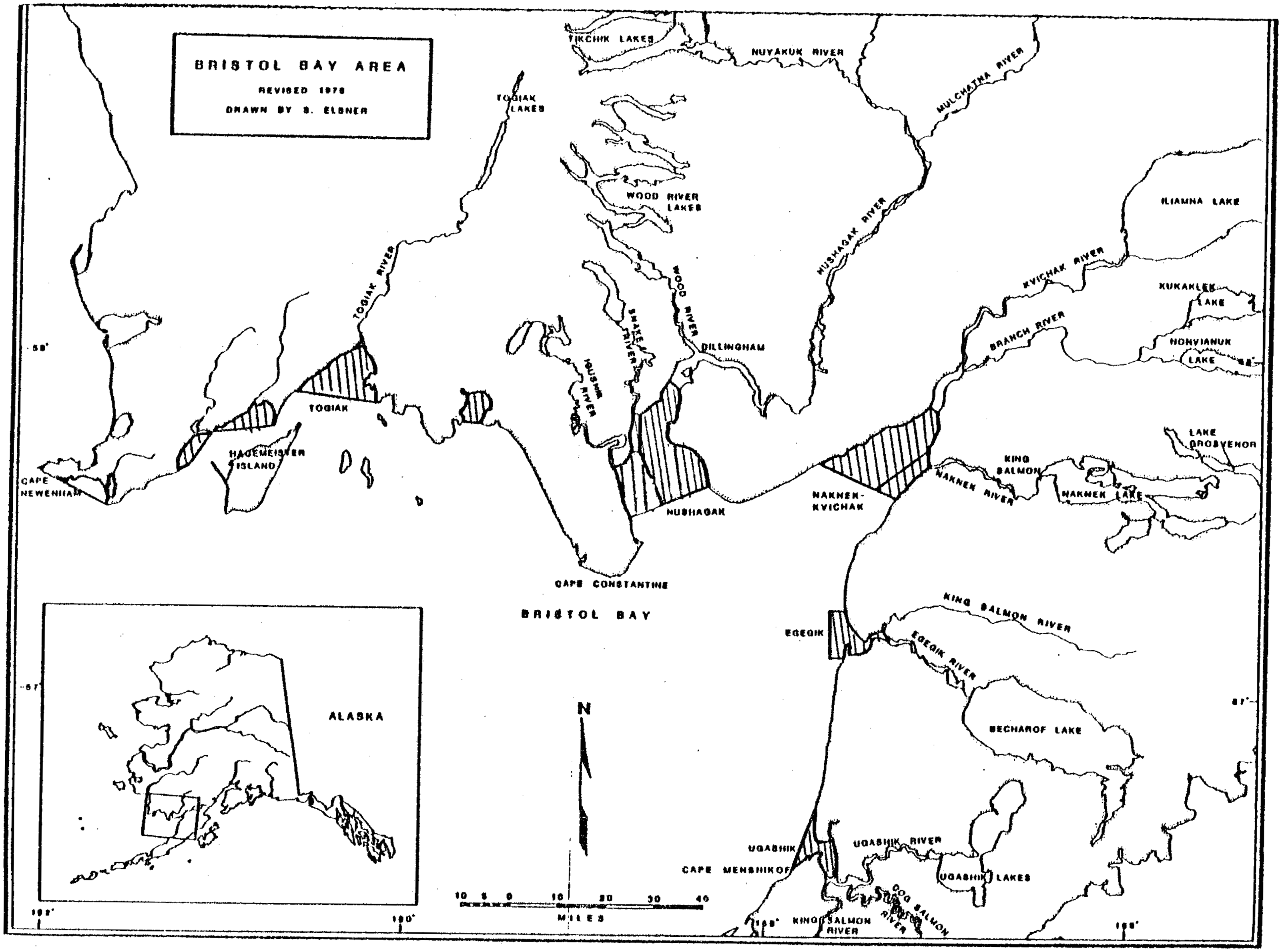


Figure 2. The Bristol Bay area. Shaded areas delineate fishing districts.

for the age 5_2 and 49% overall accuracy for the age 6_3 fish. For age 5_3 fish from 1975, all seven of the major systems were analyzed and resulted in overall classification accuracies in the high 50% range. Generally, overall classification accuracy decreases with increasing numbers of discriminating groups.

Several years' data from one and two check smolt were analyzed for the Kvichak and Naknek Rivers. The results from scale analysis of smolt which had spent two full years in freshwater ranged from the low to mid-70% range. Data from one check smolt provided generally poor overall accuracies. Data derived from smolt scales would be of limited value in identification of future adult returns since subsequent measurements of the freshwater portions of the adult scales from these same brood years indicates a strong selective pressure for the larger smolt during the marine life stage. In most instances, the means for the same characteristics measured on the returning adults were significantly larger than the corresponding measurements of smolt scales. In most cases, the measurement of adults were larger by at least one standard deviation.

In the final series of tests, from the east-side systems (Kvichak, Naknek, Egegik and Ugashik) and the west-side systems (Wood, Nuyakuk and Igushik) were pooled by age group. An offshore test fishing program near Port Moller on the north side of the Alaska Peninsula provides abundance estimates for the entrance of Bristol Bay sockeye based on CPUE, approximately 6-8 days prior to the arrival of the fish in the districts. Since scale samples are routinely collected, application of an east-side systems versus west-side systems pooled function would allow refinement of these estimates to provide additional timing and abundance information. For age 4_2 fish, the overall accuracy range (east vs. west) was in the high 70's; for 5_2 and 6_3 samples accuracies were in the mid 80% range and for age 6_3 fish, performance was in the low 80's.

Data from all years can be pooled without significantly reducing classification accuracy. However, there are significant between-year variabilities within the groups which prevents these years' pooled functions from being used as an "in-season" tool prior to obtaining new standards from the escapements.

In general for the east-side systems, scale measurements of the freshwater zone provided the best discriminating variables in most years. However, scale measurements from the first marine years were in many instances excellent discriminating characteristics. For fish collected in the rivers in Nushagak Bay (west side), marine characteristics were frequently chosen in the stepwise procedure as the variables showing the largest between group variances.

Cook Inlet - Sockeye Salmon

The major sockeye salmon producing systems in the Cook Inlet area of southcentral Alaska (Figure 3) are the Susitna, Kenai and Kasilof Rivers. Analysis of scales collected from fish wheel samples in 1975 provided overall classification accuracies of about 80%. Kenai fish appear more similar to Susitna than to the neighboring Kasilof system. Freshwater variables were the most effective, followed by measurements from the first marine year. Analysis of data from Cook Inlet sockeye collected in 1976 for all age classes examined (4_2, 5_2 and 5_3), proved quite successful. The age 5_3 fish separated with overall accuracies in the high 90% range. In a two-way analysis, Kenai and Kasilof, age 5_3, sockeye are separable with nearly 100% accuracy in that year. Analysis of the age 4_2 and 5_2 fish from that year yielded overall accuracies in the low 70% range with a considerable number of Kenai fish being misclassified as Susitna fish. However, the number of samples available from the Susitna was quite limited and this generally causes a loss of accuracy. For 1976 samples, the marine characteristics were the most effective variables.

Analysis of data collected in 1977 is not yet complete. Preliminary results indicate overall accuracies in the mid-70 percent range. Samples from the multiple basin Kenai system again show the greatest error of misclassification.

Cook Inlet - Coho Salmon

Although there are many systems in the Cook Inlet area that produce coho salmon, as much as 80 to 90% of the production may be attributed to the Kenai and Susitna Rivers. Approximately 80% overall accuracy has been achieved in separating Susitna from Kenai coho salmon based upon scale characteristics (data from 1975 and 1977 age 4_3). Measurements from the marine portions of the scales were the most successful in separating the two stocks. Recent investigations by the staff of the Sport Fish Division, ADF&G (unpublished data) indicate that mean fish weight may be substantially different between the two stocks. Future research will incorporate fish length and weight with scale measurement data to determine their value as discriminant characteristics.

Cook Inlet - Chinook Salmon

The major systems which produce significant runs of chinook salmon in Cook Inlet are the Susitna, Kenai, Ninilchik, and Anchor Rivers. Scales collected from age 5_2 fish from these systems (1977) were measured for discriminant analysis.

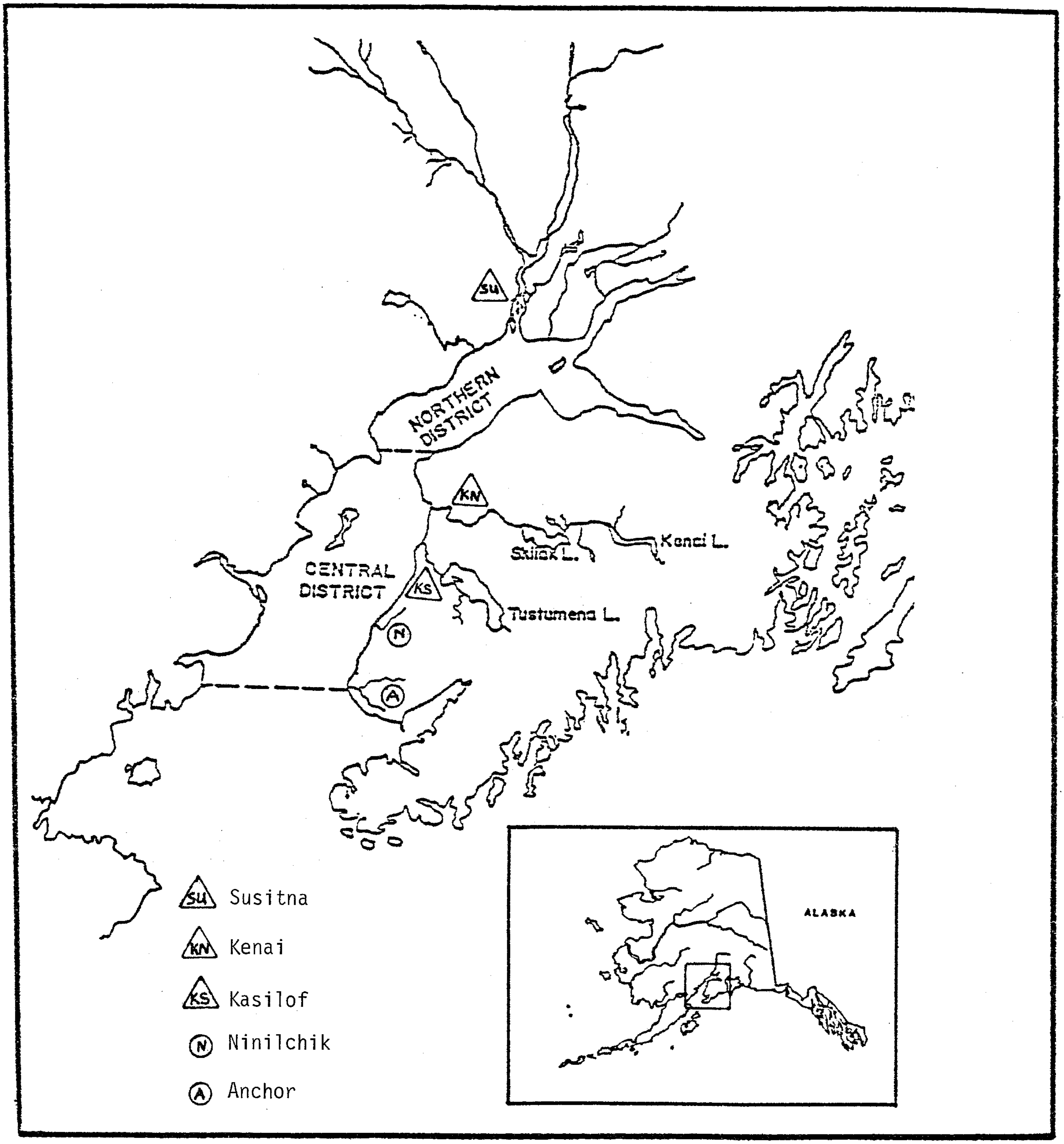


Figure 3. The upper Cook Inlet area showing locations where scale samples were collected for stock separation studies.

Pairwise (2-way) analyses of the four major systems provided overall accuracies ranging from 71% to 83%. A four-way analysis yielded 51% overall classification accuracy. As in the sockeye analyses, the Kenai system showed the greatest variability and therefore the highest misclassification error. The scale measurements reflecting the first and second marine summers growth were consistently the best discriminating variables.

Kodiak and South Peninsula - Sockeye Salmon

Data from age 5_3 sockeye salmon scales collected in 1976 from the Karluk and Frazer Rivers on Kodiak Island and from Chignik on the south side of the Alaska Peninsula provided 80% overall classification accuracy despite small sample sizes (see Figure 4). Ninety-six percent of the Chignik samples were correctly identified, whereas, there were 66% correct Karluk decisions (26% of the Karluk fish were misclassified as Frazer) and 74% correct decisions of Frazer fish (all errors for Frazer were misclassification as Karluk). Fish length and freshwater scale measurements were consistently selected as the best discriminating variables. Discriminant analysis using only freshwater variables, provided overall accuracies in the high 60% range. Karluk and Frazer examined in a two-way analysis yielded an overall accuracy of approximately 70% based primarily on freshwater characteristics.

Yukon River - Chum Salmon

Chum salmon scales were measured in a non-standard manner. The data include the measurement from the scale focus to a false check (transition or migration check) which occurs before the first winter check.

Age 3_1 and 4_1 scale samples from the Sheenjek River, a large tributary of the Porcupine system in Northeast Alaska and the Toklat and Delta Rivers, both Tanana River tributaries in the Central Interior were available for 1976 (Figure 5). Overall accuracy was 76% with individual group accuracies of 89% for Sheenjek, 67% for Toklat and 72% for Delta River. The two Tanana tributaries were most similar in scale measurements and only a small proportion of the errors were misclassification as Sheenjek.

Some additional samples of age 4_1 chum salmon collected in 1974 from the Sheenjek and Toklat Rivers were also examined but produced classification accuracies in the low 60% range.

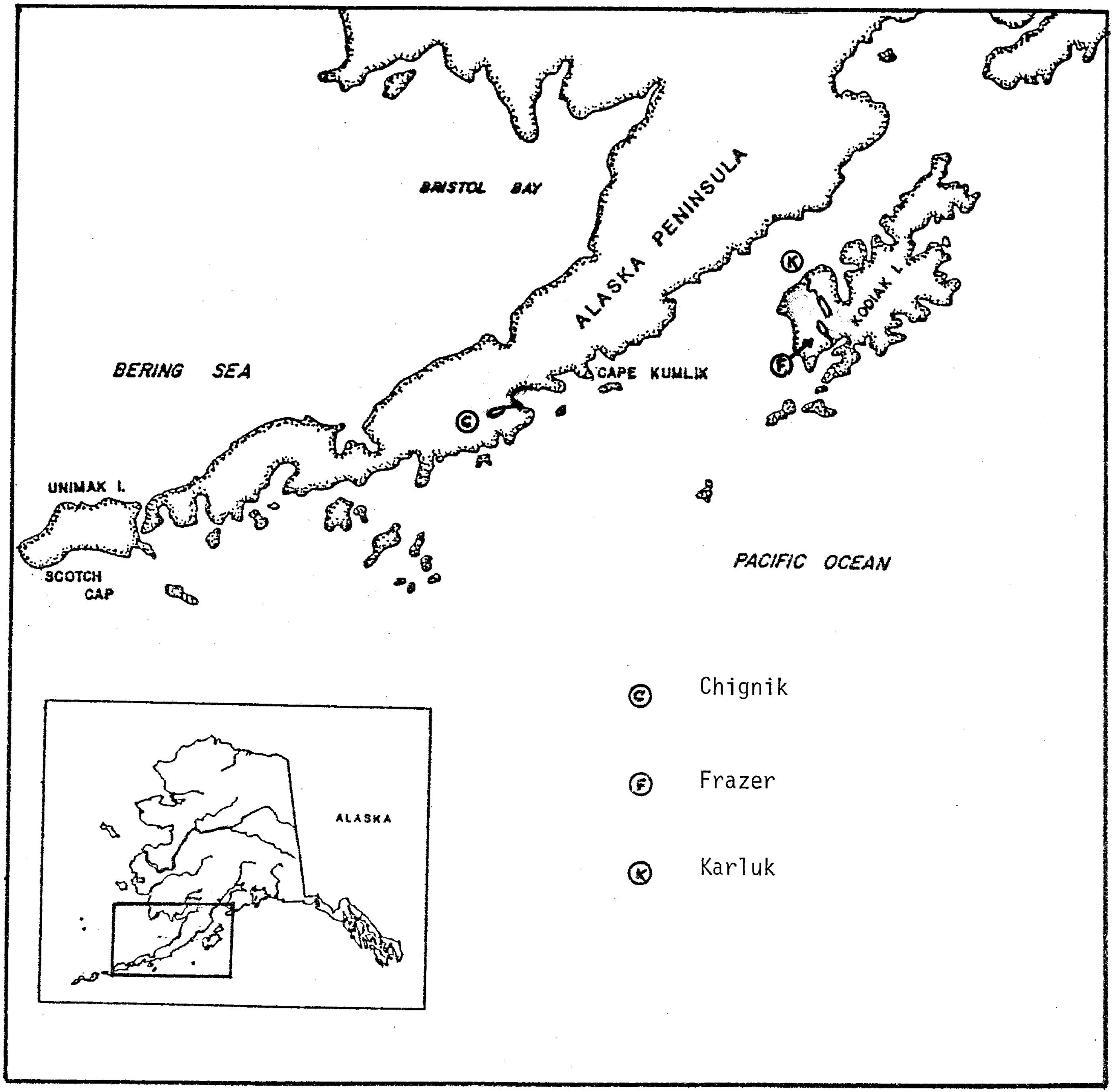


Figure 4. The Kodiak - South Peninsula area showing locations where sockeye salmon scales were collected for stock separation studies.

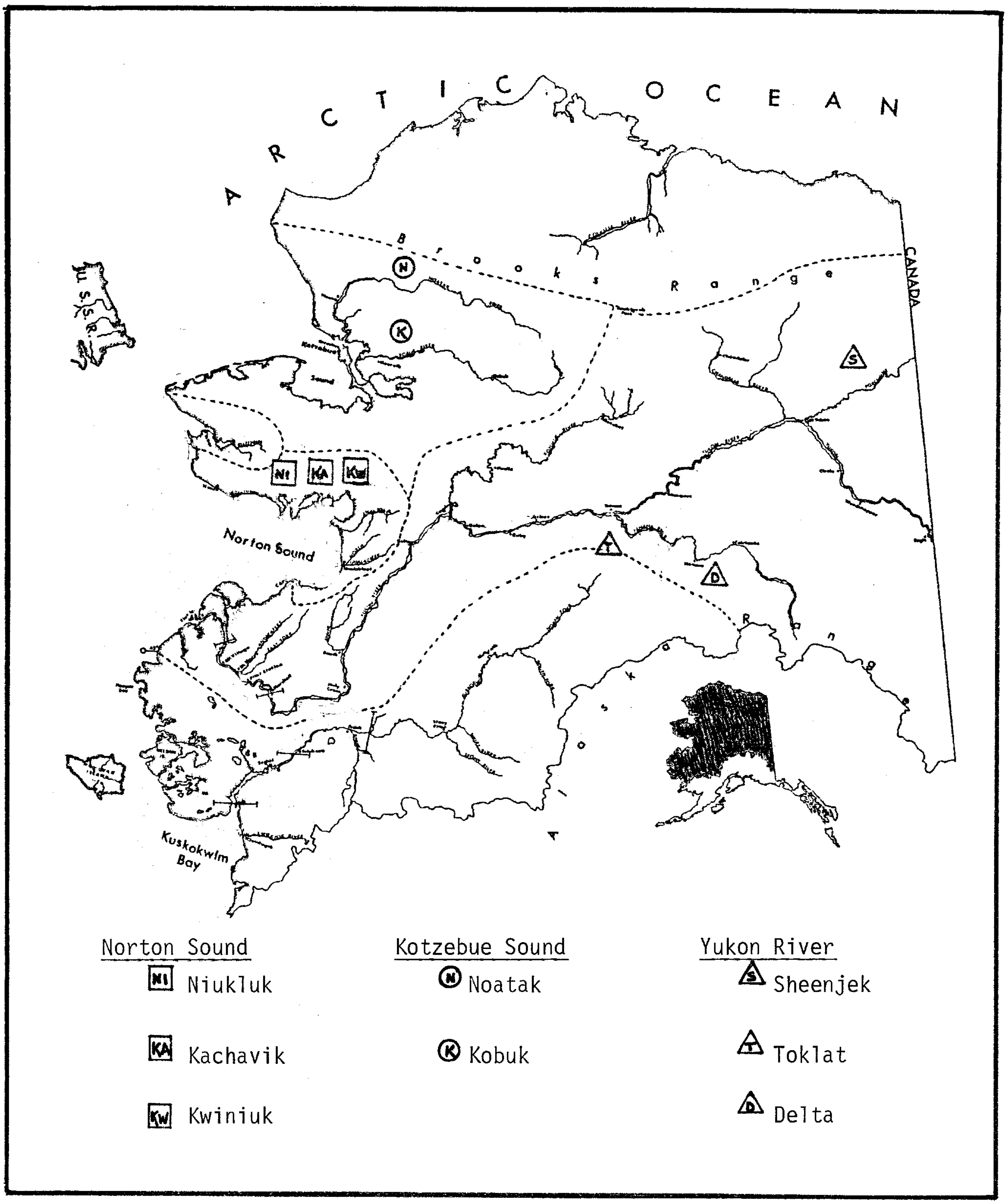


Figure 5. The Arctic-Yukon-Kuskokwim region showing areas where chum salmon scales were collected for stock separation studies.

Norton Sound and Kotzebue Sound - Chum Salmon

Age 4_1 chum salmon scales collected in 1977 from the Kobuk and Noatak Rivers in Kotzebue Sound (Figure 5) were measured using the same scale characteristics as the Yukon River chums. Overall classification accuracy was 69.3% with 20.0% classification error for Noatak and 41.2% classification error for Kobuk. The bimodal distribution of discriminant scores from the Kobuk samples may indicate the presence of two discrete stocks in that system, perhaps from the Squirrel River, a major tributary and mainstream stocks of the Kobuk. Additional sampling in each tributary may improve the accuracies. In Norton Sound (Figure 5) samples were available from only three of the contributing systems: Kwiniuk, Nuikluk and Kachavik. The three-way analysis yielded 63% overall accuracy for age 4_1 chum salmon. Whereas the separation of the Kotzebue Sound stocks was based primarily on characteristics from the second marine year, the Norton Sound samples had greater between group variability based on the first year characteristics.

Samples Grouped by Geographical Area

Throughout the various areas and species studied there appears to be an overall tendency for variation in scale patterns to be partially a function of geographic distance between systems. For example, the Kotzebue Sound and Norton Sound stocks, although showing considerable scale pattern differences within the areas, show substantial differences between these areas if each is considered as a discriminant group. Discriminant analysis of this pooled-area sample yields classification accuracy in the 80% range. Also, the means from measurements of the Yukon River chums are considerably different from Norton or Kotzebue Sound fish. In the Kodiak data, Karluk and Frazer fish have somewhat similar scale characteristics. These lakes are both on Kodiak Island. The scales from Chignik, located on the Alaska Peninsula west of Kodiak Island, are easily distinguishable from the Kodiak Island stocks. As stated above, there is a general similarity among the systems in Bristol Bay when pooled into groups of east and west-side systems.

Within some smaller areas, some analyses have combined systems in order to reduce the number of discriminant groups. Since, in general, overall accuracy increases with fewer groups, this may be a valid technique where multiple group analyses yield poor accuracies, and separation into individual river systems is not essential. For example, in Bristol Bay, comparison of Wood, Nuyakuk and Igushik River data provide low accuracies in some years. Pooling data from Wood and Nuyakuk and comparing this pooled sample with Igushik does, in some cases, substantially improve the accuracy. However, preliminary research into the application of resulting classification equations to a mixed sample indicates that this may have a secondary effect of increasing the variance of the group frequency estimates (T.L. Robertson, personal communication).

Since scale measurements are reflections of fish growth through various life stages, similarity in scale patterns of fish from adjacent systems might reflect the similar environmental influences affecting those stocks. However, in many instances, the best discriminating characteristics were measurements from growth in the second and third marine years. This would seem to suggest that either there were different environmental influences affecting these stocks on the high seas, or that genetic factors affecting fish growth and subsequently scale development are to some measurable extent responsible for these population differences. Taken a step further, if it is assumed that genetic similarity would be greatest among neighboring populations (systems), it would explain why scale pattern differences are more pronounced with increasing geographical distance.

LITERATURE CITED

Anas, R.E. and S. Murai. 1969. Use of scale characters and a discriminant function for classifying sockeye salmon (Onchorhynchus nerka) by continent of origin. Bull. Int. North Pac, Fish. Comm., No. 26: 157-192.

Bergander, F.E. 1977. Sockeye stock contribution and identification, S.E. Alaska. Alaska Dept. of Fish and Game Tech. Rpt., Project No. AFC-51-2. 22p.

Bilton, H.T. 1970. Data and computations used in racial analyses of age 1.2 and 1.3 North American sockeye on the basis of their scale characters. Fish. Res. Bd. Can., Tech. Rpt. No. 223, 57p.

_____. 1971. Data and computations used in racial analysis of even- and odd-year North American pink salmon on the basis of their scale characters. Fish. Res. Bd. Can., Tech. Rpt. No. 234, 10p.

_____. 1972. A comparison of body and scale growth of young sockeye salmon (Onchorhynchus nerka) reared under two light periods and in total darkness. Fish. Res. Bd. Can., Tech. Rpt. No. 330, 12p.

Bilton, H.T., D.W. Jenkinson and M.P. Shepard. 1964. A key to five species of Pacific salmon (Genus Oncorhynchus) based on scale characters. J. Fish. Res. Bd. Can., 21(5): 1267-1268.

Bilton, H. T. and H. B. Messinger. 1975. Identification of major British Columbia and Alaska runs of 1.2 and 1.3 sockeye from their scale characters. In: Bulletin, International North Pacific Fisheries Commission 32:109-128.

Bilton, H.T. and G.L. Robins. 1971a. Effects of feeding level on circulus formation on scales of young sockeye salmon (Oncorhynchus nerka). J. Fish. Res. Bd. Can., 28: 861:868.

_____. 1971 b. Effects of starvation, feeding and light period on circulus formation on scales of young sockeye salmon (Onchorhynchus nerka) Ibid. 28: 1749-1755.

_____. 1971 c. Response of young sockeye salmon (Oncorhynchus nerka) to prolonged periods of starvation. Ibid. 28: 1757-1761.

Bilton, H.T. and D.H. Smith. 1969. Scale characteristics of sockeye salmon (Oncorhynchus nerka) originating from small nursery areas of the Skeena River system. Fish Res. Bd. Can. Tech. Rpt. No. 133, 33p.

Clutter, R.I. and L.E. Whitesel. 1956. Collection and interpretation of sockeye salmon scales. Bull. Int. Pac. Salmon Fish. Comm., No. 9, 159p.

Dixon, W.J. (ed.). 1976. BMD Biomedical Computer Programs. Univ. Calif. Press, Los Angeles, 773p.

Fisher, R.A. 1936. The use of multiple measurements in taxonomic problems. Ann. Eugenics 7: 179-188.

Henry, K.A. 1961. Racial identification of Fraser River sockeye salmon by means of scales and its application to salmon management. Bull. Int. Pac. Salmon Fish. Comm. No. 12, 97p.

International North Pacific Fisheries Commission. 1963. Annual Report 1961: 167p.

Klecha, W.R. 1975. Discriminant Analysis. In: SPSS Statistical Package for the Social Sciences. (ed. N.H. Nie, C.H. Hull, J.G. Jenkins, K. Steinbrenner and D.H. Brent). McGraw Hill, N.Y. pp 434-467.

Koo, T.S.Y. 1962. Differential scale characters among species of Pacific salmon. In: Studies of Alaskan red salmon (Ed. T.S.Y. Koo) Univ. Wash. Press, Seattle. pp. 127-135.

Larkin, P.A. 1972. The stock concept and management of Pacific salmon. In The stock concept in Pacific salmon (ed. R.C. Simson and P. A. Larkin) Univ. Brit. Columbia, pp 11-15.

Major, R.L. and D.R. Craddock. 1962. Marking sockeye salmon scales by short periods of starvation. U.S. Fish and Wildl. Serv., Spec. Sci. Rpt, Fish., No. 416, 12 p.

Mason, J.E. 1967. Scale studies of sockeye salmon. Int. North Pac. Fish. Comm., Ann. Rpt. 1966; 99-111.

Mosher, K.H. 1963. Racial analysis of red salmon by means of scales. Bull. Int. North Pac. Fish. Comm., No. 11, p 31-55.

_____. 1972. Scale features of sockeye salmon from Asian and North American coastal regions. Fishery Bull. 70(1): 141-183.

Nie, N.H., C.H. Hull, J.G. Jenkins, K. Steinbrenner and D.H. Bent. 1975. SPSS Statistical Package for the Social Sciences. McGraw-Hill, New York, 675p.

Pearson, R.E. 1964. Pink salmon scale studies. In: Report on the investigations by the United States for the International North Pacific Fisheries Commission - 1963. Int. North Pac. Fish. Comm., Annual Rpt, pp 162-165.

Ryan, P. and M. Christie. 1976. Scale reading equipment. Fisheries and Marine Service, Canada, Technical Report No. PAC/T. -75-8. 38p.

Worlund, D.D. and R.A. Fredin. 1962. Differentiation of stocks. In: Symposium on pink salmon. H.R. MacMillan Lectures in Fisheries, Univ. Brit. Columbia, Vancouver, pp 143-153.

Wright, A.T. 1965. The use of scale circuli spacings as a means of separating races of Prince William Sound pink salmon. Alaska Dept. of Fish and Game, Informational Leaflet No. 66, 27p.

APPENDIX

Appendix Table 1. Scale characteristics measured, by species. Data were recorded as paired measurements of NC_i and ID_i where:

NC_i = Number of circuli in ith characteristic

ID_i = Interval distance of ith characteristic

A_j = Annulus formed during jth winter growth period

fw = freshwater

oc = ocean

Unless otherwise noted, measurements to or from an annulus (Aj) include the circuli and distance to or from the last closely spaced circulus forming a part of that annulus.

Sockeye (age 4_2, 5_2)

i = 1 = focus to A_1 to A_2 (1st fw year)
2 = A_1 to A_2 (1st oc year)
3 = A_2 to A_3 (2nd oc year)
4 = A_3 to A_4 (3rd oc year - age 5_2 only)

Sockeye (age 5_3, 6_3)

i = 1 = focus to A_1 (1st fw year)
2 = A_1 to A_2 (2nd fw year)
3 = A_2 to A_3 (1st oc year)
4 = A_3 to A_4 (2nd oc year)
5 = A_4 to A_5 (3rd oc year - age 6_3 only)

Sockeye smolt (age I)

i = 1 = focus to first circulus of A_1 (1st fw summer)
2 = within A_1 (1st fw winter)
3 = A_1 to scale margin (fw plus growth)

Sockeye smolt (age II)

i = 1 = focus to first circulus of A_1 (1st fw summer)
2 = within A_1 (1st fw winter)
3 = A_1 to first circulus A_2 (2nd fw summer)
4 = within A_2 (2nd fw winter)
5 = A_2 to scale margin (fw plus growth)

Appendix Table 1. (cont.)

chum (age 3_1, 4_1)

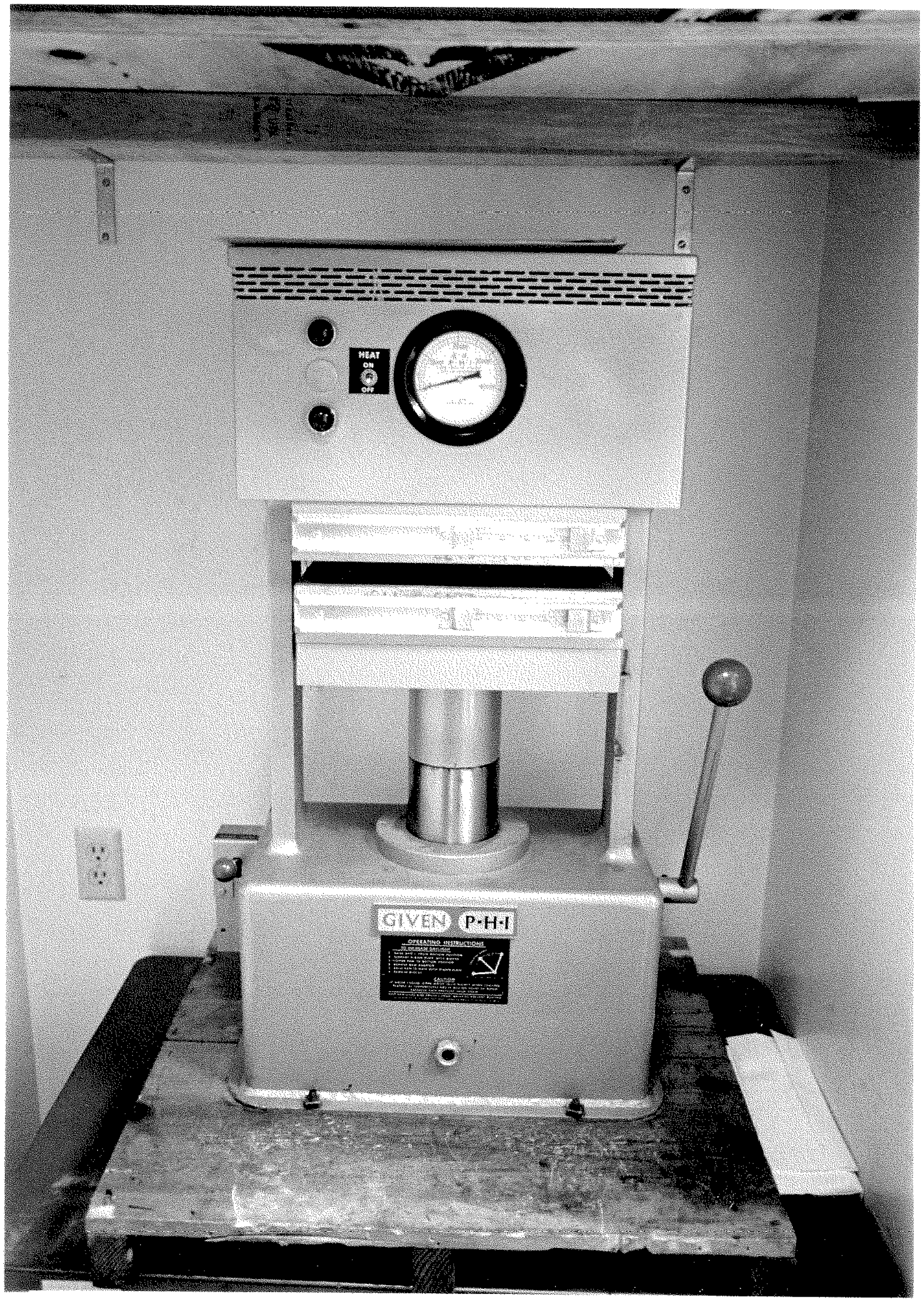
i = 1 = focus to last circulus supplementary (false) check
2 = supplementary check to first circulus A_1 (1st oc summer)
3 = within A_1 (1st oc winter)
4 = A_1 to first ciculus A_2 (2nd oc summer)
5 = within A_2 (2nd oc winter)

Coho (age 4_3)

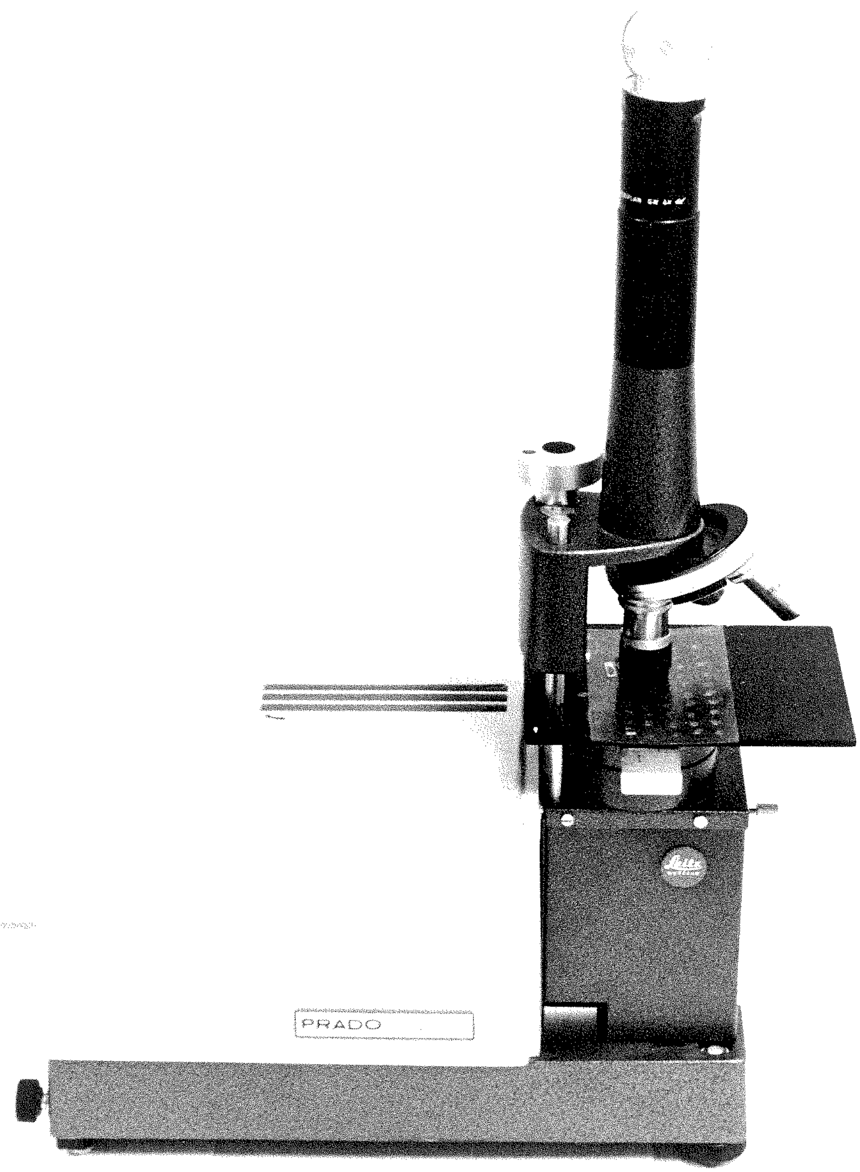
i = 1 = focus to A_1 (1st fw year)
2 = A_1 to A_2 (2nd fw year)
3 = A_2 to A_3 (1st oc year)

Chinook (age 5_2)

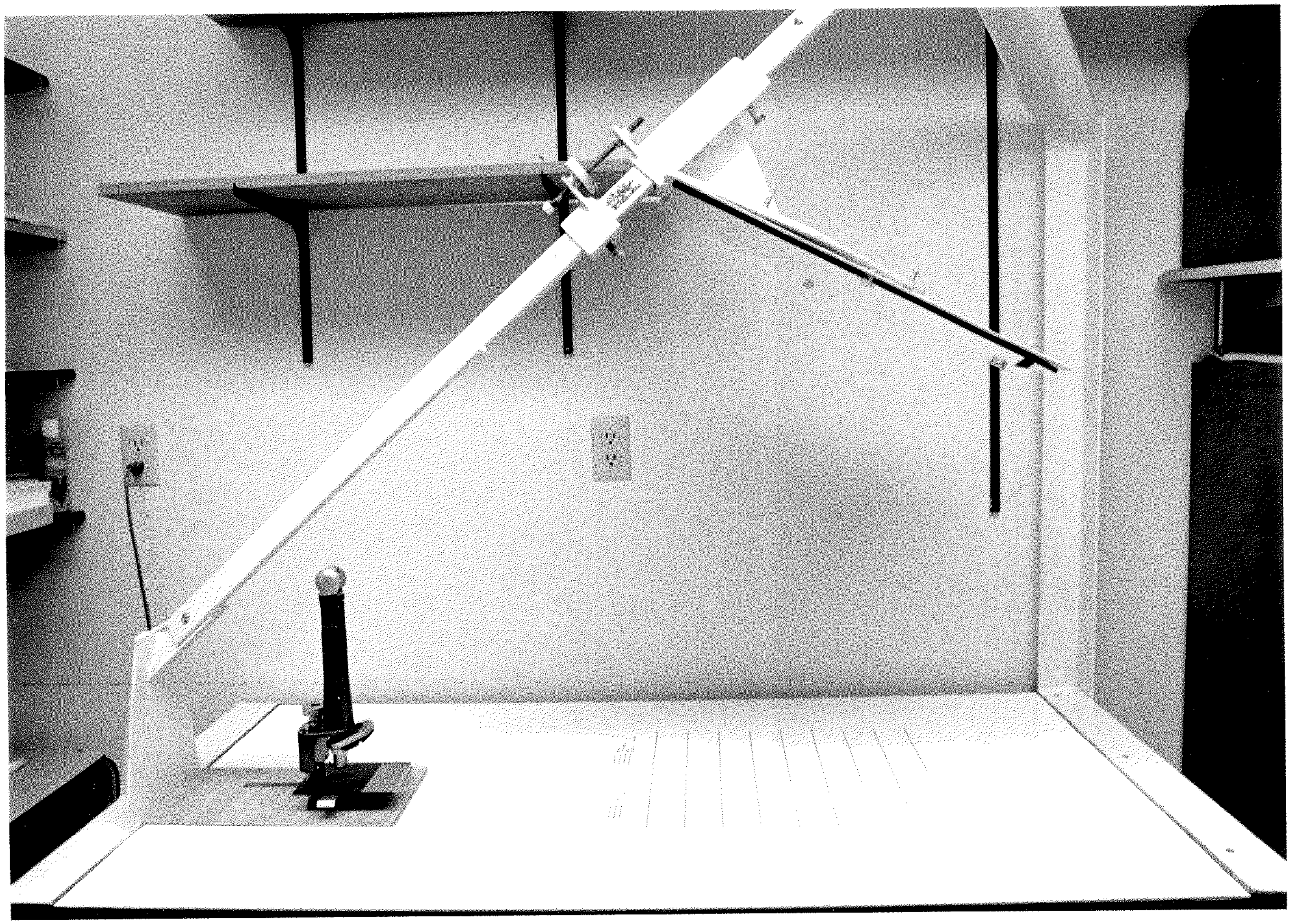
i = 1 = focus to first circulus A_1 (fw summer)
2 = within A_1 (fw winter)
3 = A_1 to first circulus A_2 (1st oc summer)
4 = within A_2 (1st oc winter)
5 = A_2 to first circulus A_3)2nd oc summer)
6 = within A_3 (2nd oc winter)



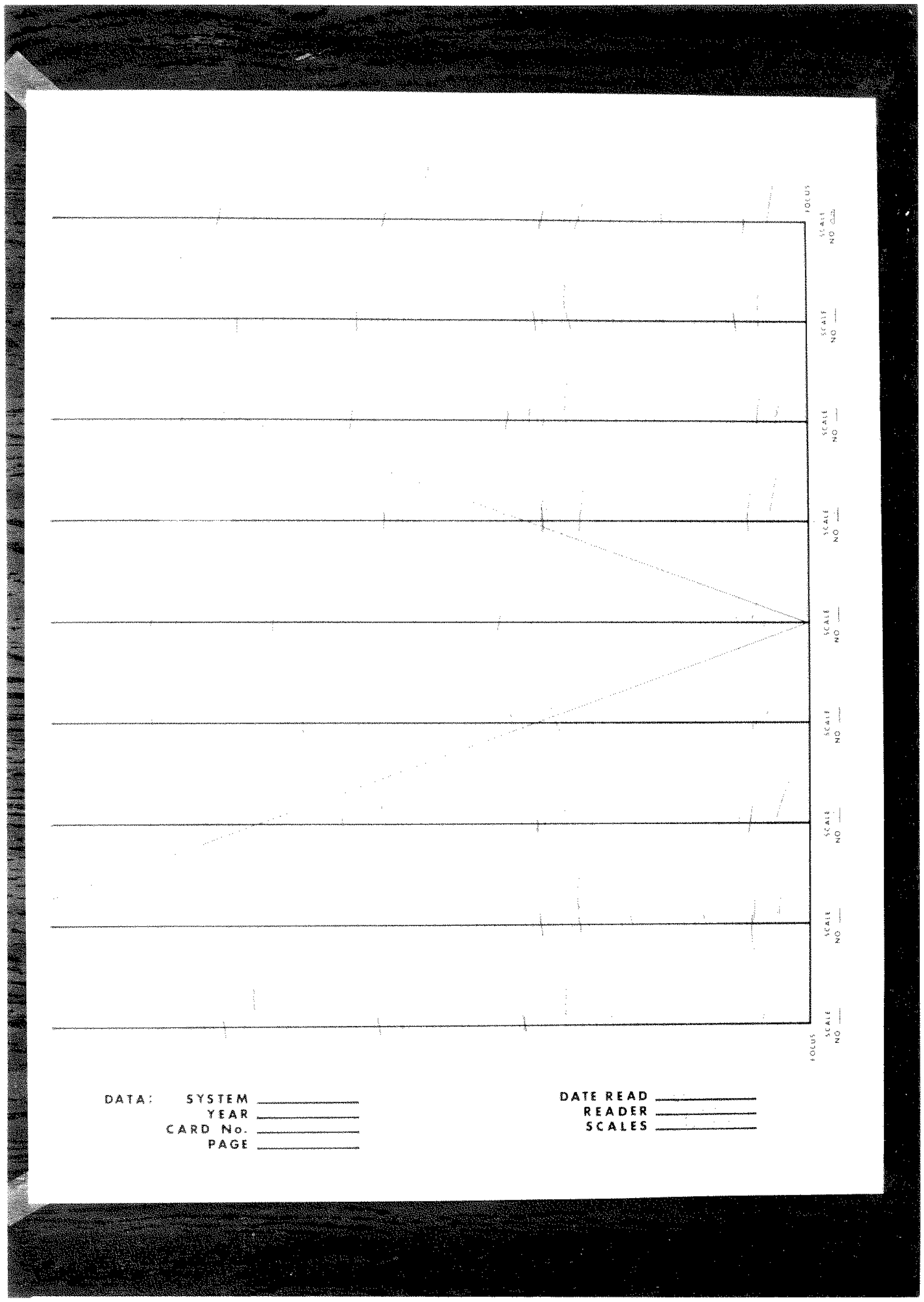
Appendix Figure 1. Hydraulic scale press used to make impressions of salmon scales on acetate plastic cards.



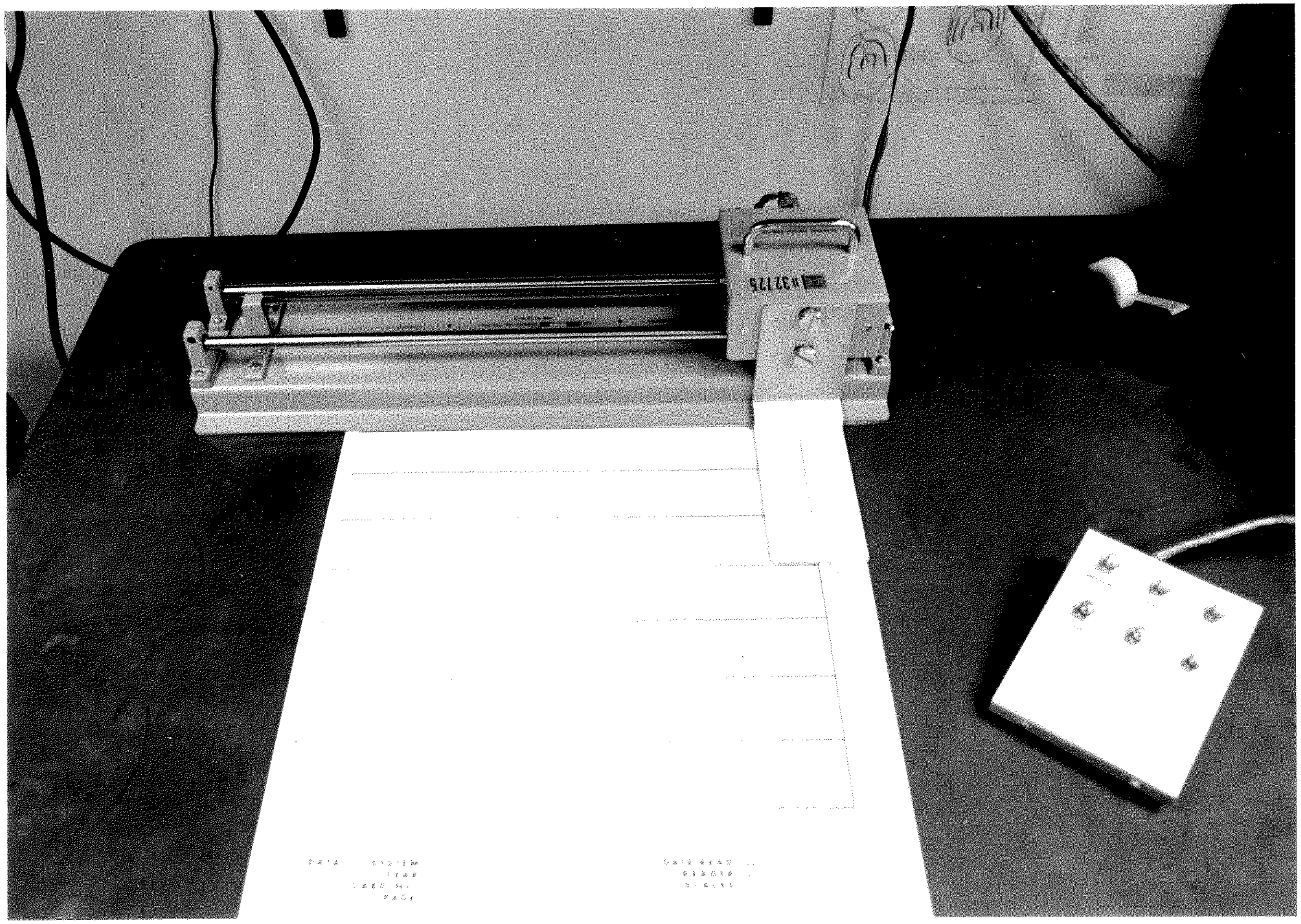
Appendix Figure 2. Leitz micro-projector and scale impression card.



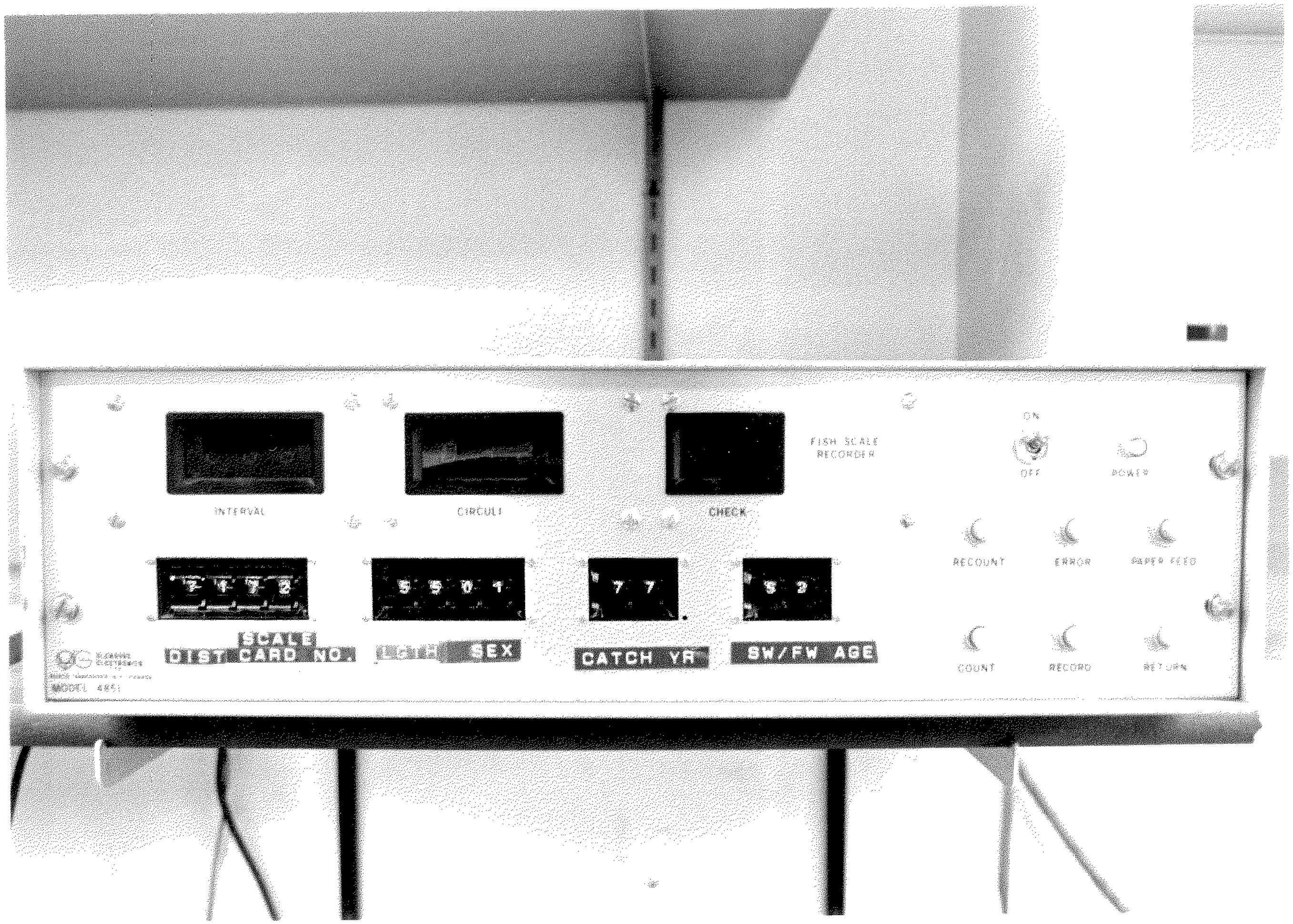
Appendix Figure 3. Micro-projector, mirror assembly and table surface for projection of scale images and marking of scale characteristics.



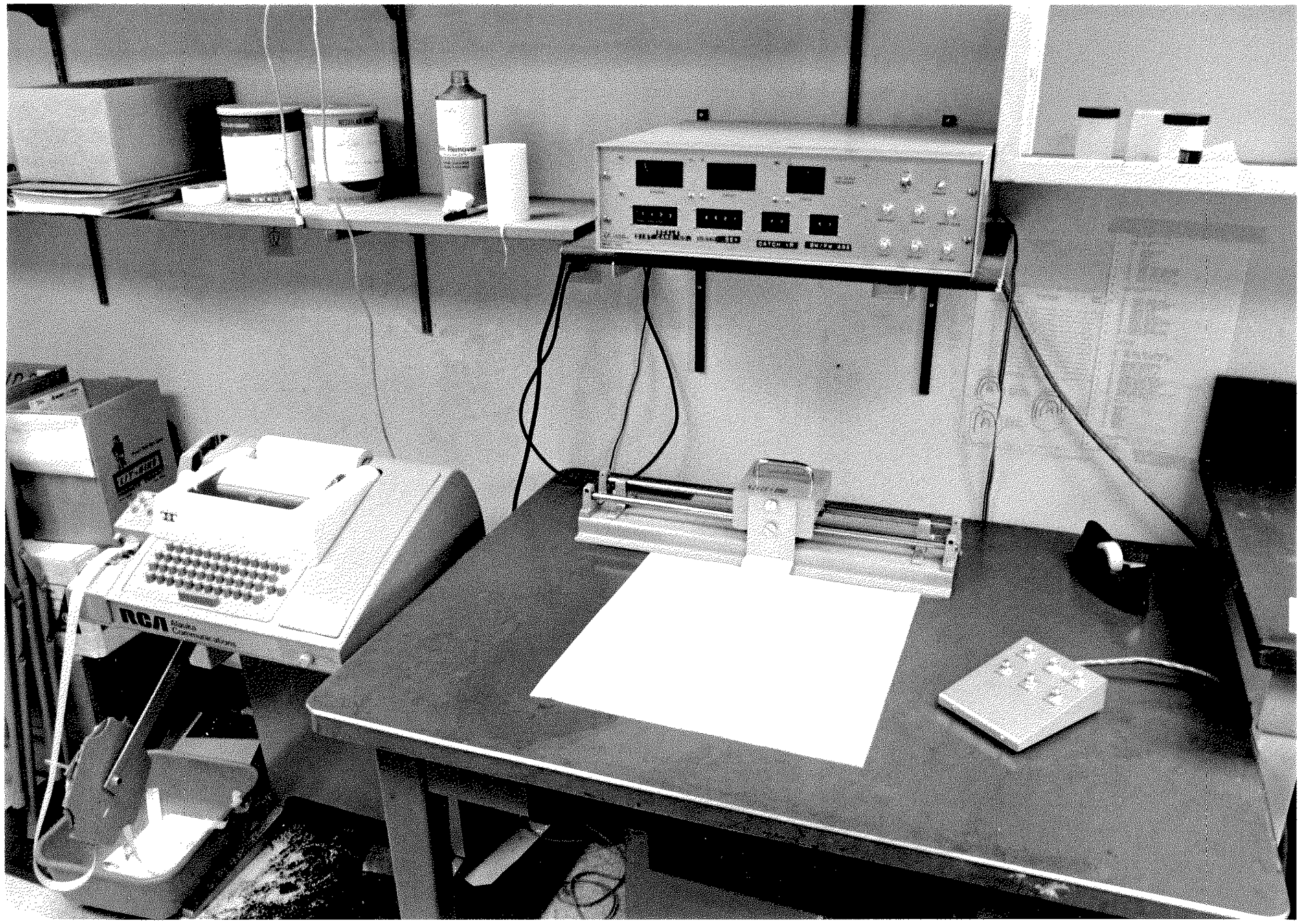
Appendix Figure 4. Scale drawing sheet with pre-printed axis lines. Short lines intersecting axes indicate positions of circuli, longer lines represent positions of annuli. (Actual paper size = 22 in. by 17 in.)



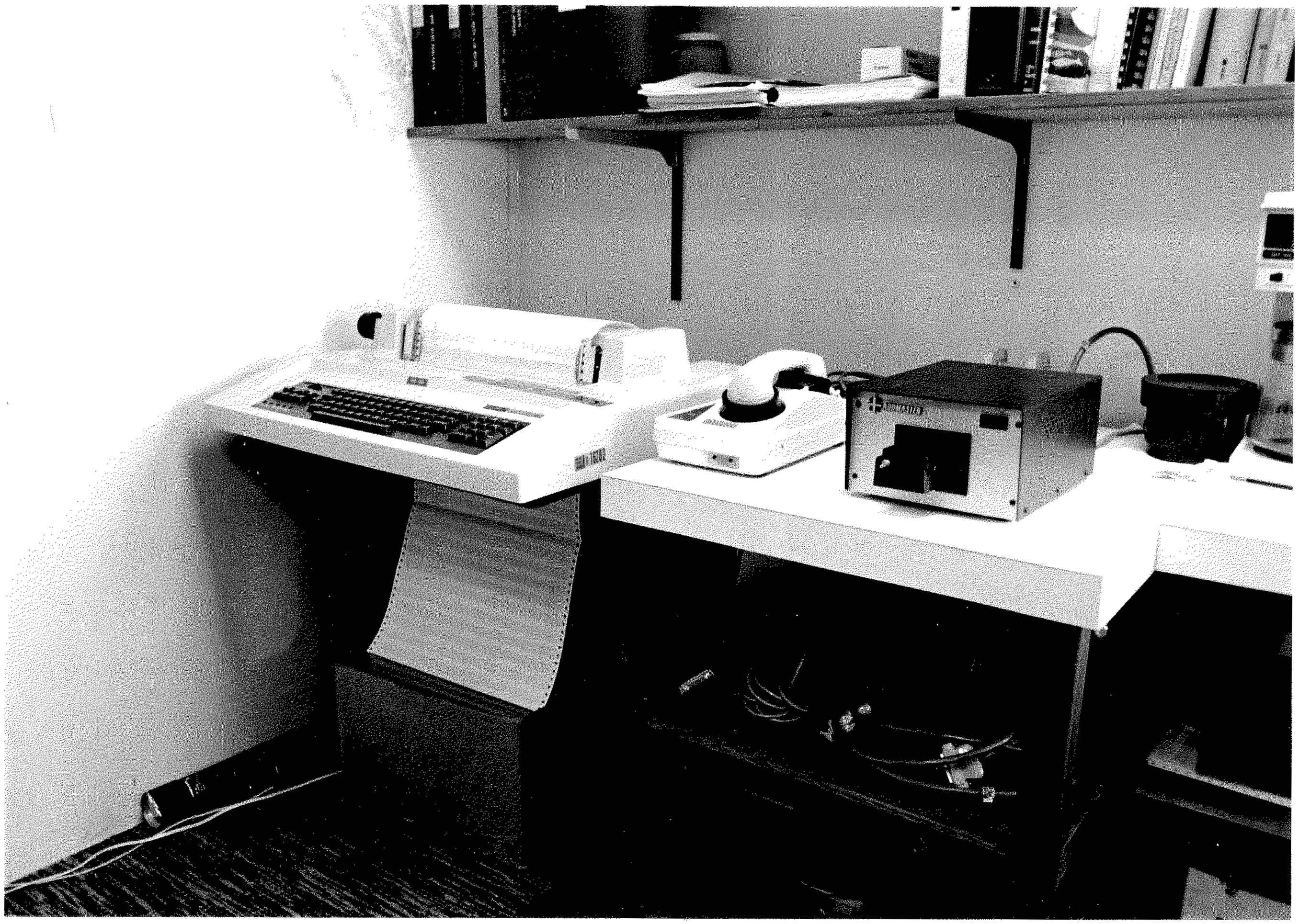
Appendix Figure 5. Linear digitizing rule (Rouchi rule) and remote digitizer controls.



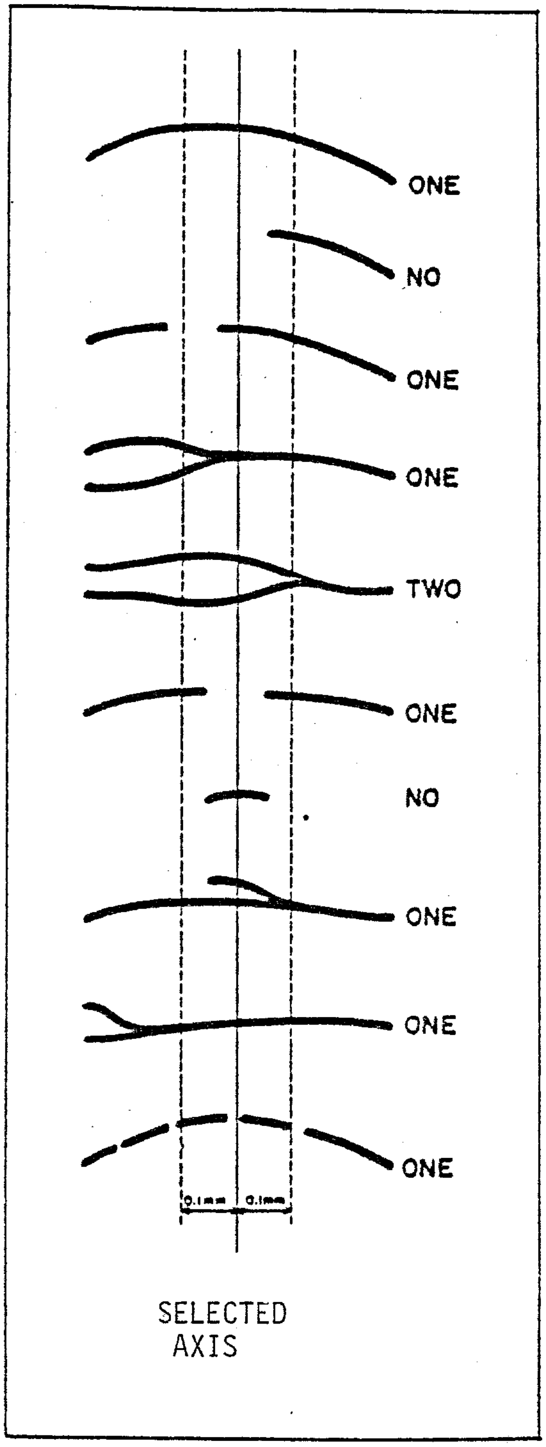
Appendix Figure 6. Digitizer electronics. Numbers visible in lower windows are fixed sample information and are controlled by resettable thumbwheels.



Appendix Figure 7. Configuration of Rouchi rule, digitizer and ASR-33.



Appendix Figure 8. Configuration of paper tape reader, computer terminal and acoustic coupler.



Appendix Figure 9. Diagrammatic representation of a section of a scale within 0.1 mm (10.0 mm magnified image) of the selected axis, showing "breakage" and "branching" of circuli and indicating the criteria for inclusion of circuli in counts and measurements. "One" and "two" indicate the number of circuli that would be included in counts and for which positions would be marked for measurement. "No" indicates that circulus would not be included in counts and that the position of its image would not be marked for measurement. (after Bilton, 1971)

```
8108 77 1111 52  000 000 012 085 041 381 062 672 079 797 085 862
8108 77 1111 52  000 000 006 051 039 364 067 617 085 778 089 822
8108 77 1111 52  000 000 009 078 038 366 062 625 080 771 085 827
8108 77 1111 52  000 000 011 082 039 354 070 609 089 775 093 817
8108 77 1111 52  000 000 009 062 041 332 065 522 082 674 087 728
8108 77 1111 52  000 000 008 063 039 329 065 586 080 731 085 790
8108 77 1111 52  000 000 013 100 045 390 071 624 087 778 089 870
8109 77 1111 52  000 000 009 066 046 357 071 546 087 685 093 761
8109 77 1111 52  000 000 007 063 042 397 066 628 081 775 083 800
8109 77 1111 52  000 000 009 073 041 388 065 601 084 789 088 840
8110 77 1111 52  000 000 009 073 038 347 067 616 086 816 090 857
8102 77 1111 52  000 000 008 065 044 413 067 618 088 822 091 851
8102 77 1111 52  000 000 011 095 042 383 072 626 090 790 093 829
8102 77 1111 52  000 000 011 072 042 375 073 666 093 863 098 928
8102 77 1111 52  000 000 010 082 043 412 061 583 080 782 083 824
8103 77 1111 52  000 000 012 083 040 371 072 632 086 786 094 887
8103 77 1111 52  000 000 007 073 039 402 067 687 084 868 089 930
```

Appendix Figure 10. Sample hardcopy output of digitized scale measurement data from ASR-33.

```
8102771111520010008200330330001801710019019900030042
8102771111520011007200310303000310291002001970005006 5
8102771111520011009500310288002802430020016400030039
8102771111520008006500360348002302050021020400030029
8103771111520007007300320329002802850017018100050062
8103771111520010008300300288003002620016015300080101
8108771111520012008500290296002102210017019500060065
8108771111520013010000320290002602340016015400020022
8108771111520008006300310266002602570015014500050059
8108771111520009006200320270002401900017015200050054
8108771111520011008200280272003102550019016600040042
8108771111520009007800290288002402390018016600050056
8108771111520006005100330313002802530018016100040044
8109771111520007006300350334002402310015014700020025
8109771111520009006600370291002501890016013900060076
8109771111520009007300320315002402130019018800040051
8110771111520009007300290274002902690019020000040041
```

Appendix Figure 11. Data shown in Figure 10 after editing, sorting, compacting and conversion from cummulative to incremental measurement.