NPAFC
Doc. No. 567
Rev. No. _____

Developing and deploying a high-resolution imaging approach for scale analysis

Peter T. Hagen

Dion S. Oxman

Beverly A. Agler

Alaska Department of Fish and Game
P.O. Box 25526
Juneau, Alaska 99821-5526

Submitted to the

NORTH PACIFIC ANADROMOUS FISH COMMISSION

By

The United States Party

October 2001

This paper may be cited in the following manner:
Hagen, P.T., D.S. Oxman, and B.A. Agler 2001. Developing and deploying a high resolution imaging approach for scale analysis. (NPAFC Doc. 567). 11 p. Mark, Tag, and Age Lab, Alaska Department of Fish and Game, P.O. Box 25526, Juneau, Alaska.

ABSTRACT

Collecting and analyzing salmon scales to obtain age information is a fundamental component of fishery management and monitoring programs. In Alaska, tens of thousands of salmon scales are routinely collected and examined each year and then archived. These collections can be extensive. For example, there has been over 45 years of uninterrupted sampling for some stocks. In addition to yielding age information, scale patterns can also be measured and used to discriminate among stocks, and to relate growth patterns to production and environmental trends. However, the methods used to extract measurements from scales are labor-intensive, rely on subjective determinations, and utilize technology that in many cases is obsolete and no longer available. As a result, measurements have been extracted from a relatively small number of scales, and the data obtained are generally not accessible for additional analysis and cannot be combined with other datasets for comparison.

Recently, we embarked on a project to develop a comprehensive approach for deploying high-resolution image analysis to scale patterns as a means to address these concerns. The approach includes three components: 1) a data management system to efficiently combine individual sampling information with scale images and pattern data; 2) the application of high-resolution imaging technology for capturing, storing and sharing images of scales; and 3) the application of image analysis routines to automate the extraction of pattern data from images using reader supervision. This document provides the rationale behind this approach, gives detail on the hardware and software components we are using, and discusses areas for further development and the opportunities for collaboration among other researchers.

INTRODUCTION

The formation of salmonid scales is strongly affected by the environmental conditions encountered during development of the individual fish. Consequently, circuli patterns observed on scales can be used to interpret and assess a variety of life history traits, including age, growth, age at migration to sea, and age at spawning (Schluchter and Lichatowich 1977, Ihssen et al. 1981, Fukuwaka and Kaeriyama 1997). This age information, combined with sex and size data, is critical for monitoring salmon stocks and is the focus of sampling efforts by hundreds of projects within Alaska each year. Many of these are long-term projects, and over time protocols for collecting scales, interpreting their patterns and storing them as acetate impressions have been established. Collections of these scales and associated data are maintained in regional Alaska Department of Fish and Game (ADFG) offices as well as many smaller field offices. The number of archived samples is estimated to be in the millions, and some collections represent 45 years of sampling effort. Unfortunately, there is no central archive or inventory of these collections, and much of the associated data are only available in the original format.

In addition to providing age information, scales also have been used to discriminate among stocks (Anas and Murai 1968, Tanaka et al. 1968, Cook 1982) and to relate growth patterns to production and environmental trends (Ruggerone and Rogers 1998, Bumgarner 1993, Isakov 1998). These applications require extracting measurements from the scales and are more limited in scope involving relatively fewer samples. The technologies currently used to obtain growth data include microprojectors, digitizing tablets, and low-resolution imaging systems. Unfortunately, this equipment is outdated and replacement parts are difficult to find. In addition, the methods employed are labor intensive, rely on subjective criteria, and the data obtained are not available for additional analysis and cannot be combined with other datasets for comparisons.

Recently, we became involved in collaborative projects with the University of Alaska, Fairbanks (UAF), the National Marine Fisheries Service (NMFS) and the U.S. Geological Survey – Biological Research Division (USGS-BRD) to create a time series of sockeye and chum salmon scale growth in Alaska. The purpose was to explore the relationship of salmon growth to climate change and other factors that could affect growth history. We used this as an opportunity to develop a different approach to analyzing scales. These projects required the creation of a digital archive of scale images to share among researchers and the collection of measurements and counts of circuli and annuli from the focus of the scale to its outer margin on approximately 20,000 scales. To achieve these goals, we developed a data structure by which we could efficiently access the sampling data associated with scales and store the associated measurements of scale patterns. The project also required the application of high-resolution digital imaging techniques to produce the scale images and extract the measurements. Our goal was to develop a system that could be adapted for use in a wide variety of scale analyses. The primary advantages of using an imaging approach was that measurement criteria could be standardized and automated, decision rules could be documented, production rates increased, and the images and data made available for other analyses. Such a database could be easily exchanged among agencies and researchers worldwide and, given the number of scales collected each year, be used to improve population forecasts and stock identification.

Described in this paper are the methods we are currently applying to the analysis of scales from the ADFG sockeye and chum salmon collections. It should be noted that this approach is still being developed. Other hardware and software components may work equally well or better, and technological improvements will no doubt be available in the future. We purposely tried to avoid linking this approach to one hardware or software solution but instead use established protocols for data management and electronic storage as a means to compile and share data on salmon scale growth.

METHODS

Figure 1 presents a flow chart of the steps we used to create a database of scale measurements for archived scales. The chart emphasizes the relationship between image capture, data management and data extraction. The processes of obtaining images and extracting data from the images are not directly joined. Thus, technological advances can be incorporated into the system without the need to upgrade all components.

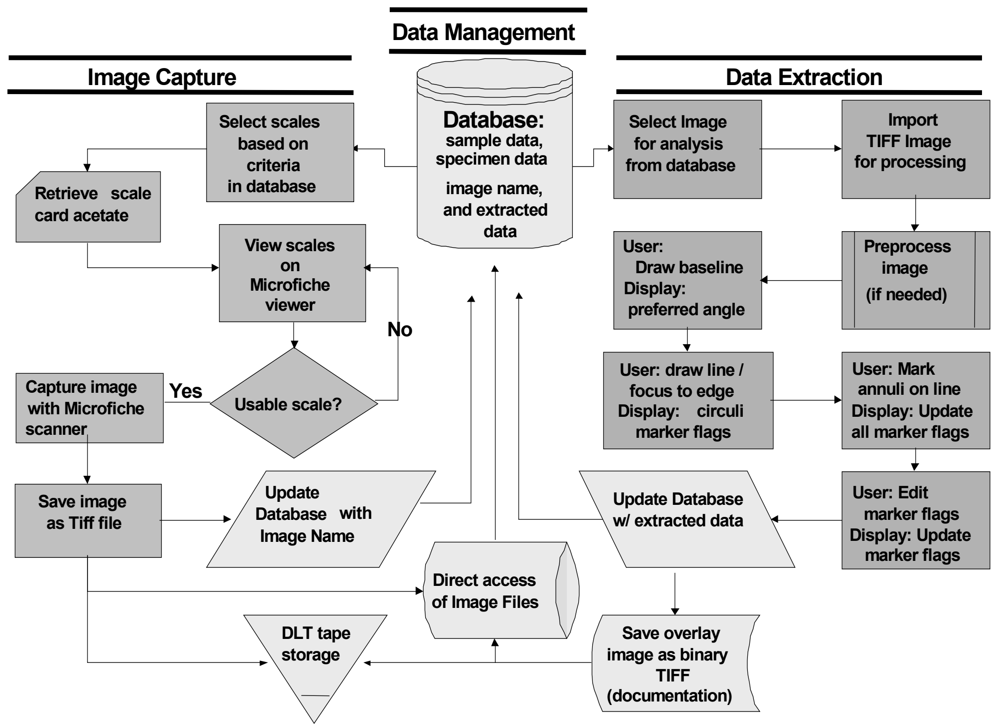

Figure 1. Flow chart showing the steps used to track the selection of scales in a database through the acquisition and data extraction steps.

The components of the system we use include readily available computer hardware and software, as well as customized applications. We provide here a brief description of the components and the procedures used to capture scale images and extract the data.

A – Database Management and Image Capture:

Microsoft Access® Database: Microsoft Access is used as our relational database to store, sort, and organize sampling data associated with each scale collection and the individual measurements associated with each fish sampled. The database is maintained on a network server and supports simultaneous use. For the projects currently underway, each system is contained in its own database to facilitate portability. However, if a common statewide approach is developed, it will be necessary to use a more robust software approach with a larger database serving as a central repository.

Custom Interface Programs: Custom Visual Basic programs allow the user to search the database for scales that meet the project criteria or to input sample data to the database when only paper copies are available. The programs generate file names based on the selection criteria. The file names are stored in a data field and serve as a link between the scale's measurement data and other tables in the database. During selection, the file name is passed to the image capture software and used as the image file name for storage.

Screenscan® Microfiche Scanner: This high-resolution line camera attaches to an Indus 4601 microfiche reader and a Windows PC computer. It digitizes the scale image directly from the microfiche screen then stores the digital image as a Group IV TIFF ("Tagged Image File Format") file at 8-bit depth and 3352 x 4425 pixels. This results in an uncompressed file size of 14.5 Mb. This resolution allows the entire scale to be viewed for aging and provides enough pixels between narrow circuli to insure accurate measurements of circuli spacing. The program supplied by the manufacturer uses ISIS protocol to communicate with a PC computer, requiring about 6 seconds for image capture and transfer to the PC. The price for scanner and microfiche is approximately $8,500.

Image Storage and Transfer: During the image capture and processing steps images are stored on network file server. For long-term and archival storage they are stored in batches on Digital Linear Tapes (DLT). Each tape stores 40 gigabytes of uncompressed information and 80 gigabytes of compressed data. Recordable digital media, such as DVD-RAM and CD-R, are also available for storage and distribution of images. Because of the large file size and the need to preserve detail on circuli spacing, transfer through the Internet will likely incorporate wavelet-based file compression, such as that found in the new JPEG 2000 format and available from companies such as LuraTech®.

PROCEDURES

Sockeye and chum salmon scales, as well as the associated sex, age, length, and weight (AWL) data were provided by regional offices of ADFG. Sampling protocols in the field called for each scale to be removed from the "preferred area" of the fish, located two rows above the lateral line where it intersects with the row of scales descending from the posterior region of the base of the dorsal fin. Scales were subsequently placed on gum cards from which acetate impressions were made for data replication, long-term storage, and analysis.

The sampling goal for digitizing was 30 scales per age/sex/year. Scales were viewed on the microfiche viewer and the accompanying data was examined through a custom program tied to the database or from paper copies.

A scale was selected for image capture only if: (1) the circuli and annuli were clearly defined; (2) the scale reader agreed with the age as previously reported by a trained

ADFG reader; and (3) the scale's shape indicated that it was removed from the preferred area of the fish.

Scales were discarded if: (1) the accompanying data regarding a fish's length, gender, collection date, or location of capture were missing or incomplete; (2) the scale impression was poorly defined or the acetate card too warped to be placed in the microfiche reader; (3) the scale pattern was too confusing to measure accurately because of numerous false checks and/or branching; (4) the scale was damaged, regenerated, or possessed numerous irregular or wavy annuli; or (5) the scale showed significant resorption along the measurement axis, making age determination uncertain.

Once a scale had been selected for analysis, the Screenscan microfiche scanner was used to digitize the scale impression directly from the microfiche reader to the networked file server. The file name was stored as a field in the database computer for subsequent selection and importing into the image processing program.

B –Data Extraction - Image Processing:

Optimas® 6.5 Image Analysis Program: Saved images are loaded into Optimas 6.5 image processing software to collect measurement data using a customized program written in the Optimas programming language (ALI). Optimas is installed on a PC computer equipped with dual monitors. One standard monitor is used to display Windows software, while a digital LCD flat panel tablet is used to display the scale image. A more limited version of Optimas, called Optimate® can also run the customized program developed under Optimas and is currently used on one of our workstations. Optimas has a list price of $4000 while Optimate has a list price of about $1000. However, it is likely that both Optimas and Optimate will no longer be available in the near future. Another imaging company purchased the Optimas software product line, and its availability is uncertain. We are investigating other software packages, but do not have an alternative to recommend at this time.

Wacom® Digital Flat Panel Monitor w/ Touchscreen Interface: The scale images are displayed on a Wacom LCD tablet. These 13-inch or 15-inch digital LCD graphics tablets enable the user to interact directly with the scale image on the screen. The user can zoom and pan across the image, draw transect lines, place annuli markers, and edit the placement of circuli markers directly on the screen using a programmable pen. Use of these tablets may require compatible video cards in the PC. We use a Matrox G400 dual head graphics card with Digital Video Interface (DVI) adapter. Rapid changes are taking place with these products and prices for the 15-inch tablets are starting at about $1800.

Data Import Program: This Visual Basic program imports scale measurement data from text files created by Optimas into the Microsoft Access database. The program loads the data into several tables with the image name as the link to the sample table, allowing flexibility in extracting and exporting data to statistical packages for further analysis.

Procedures

The program we wrote using the Optimas language allows the user to conduct semi-supervised extraction of circuli and annuli counts and measurements between circuli. To do this, the reader first loads a scale image into the program and then establishes a reference line along the longest axis of the scale's reticulated region using the Wacom tablet and pen. The program then generates a line indicating the preferred angle for extracting data and the reader uses that line as a guide to draw the transect line from the center of the focus to just beyond the scale's edge. It is along this line that growth increments (circuli and annuli) are counted and measured. Annuli markers are placed manually by the reader, while the circuli are flagged automatically using edge detection algorithms that are applied to the scale image's luminescence profile (Figure 2). The reader then evaluates and edits the computer's marker placement using the tablet and the program's pan and zoom functions to ensure adequate placement. The edge detection algorithms used by Optimas place flags with an accuracy ranging from 60-100%, with an average of about 90%. Accuracy is determined by whether the reader overrides the original placement made by the computer or has to add or delete flags. It is dependent on the quality of the scale impression, the resolution of the scale image, and the age of the scale card (older cards tend to be warped and scratched).

The transect line, along with its associated annuli and circuli markers, are stored as separate binary images that can be recombined with the original scale image. This provides documentation on the reader's placement of annuli, thus it can be compared with other readers of the same image. The Optimas program counts circuli and annuli, and measures distances between circuli. Once count and measurement data are extracted, they are imported into a text file along with comments on scale quality and notes regarding unique features, such as amount of resorption along the scale's edge and amount of fresh or salt water plus growth. The data are saved as a comma and tab delimited text file but may also be saved as Microsoft Excel files. A Visual Basic program is used to import the data into the Microsoft Access® database, where measurement information is combined with data regarding that individual's stock (river of origin), year of return, brood year, age, month and day of collection, gender, fish length (mm), scale reader, acetate card number, scale location on acetate card, resorption code, image file name, and overlay image file name. Once a scale's growth increments have been measured, their images are removed from the computer and placed on DLT tapes for storage and backup.

DISCUSSION

Current production rates depend on the species of salmon, the quality of the scale impressions, and the experience of the reader. Scales are viewed and selected for image capture on the microfiche scanner at a rate of approximately 120-150 scales per day. Much of this time is spent handling and selecting scales that had been previously aged to see if they meet the project criteria, but capturing images could just as easily take place when scales are first examined, either as part of routine monitoring program or for specific projects that may make use of digital image analysis.

Extracting measurement data from scales takes more time. Our application requires collecting measurements from all circuli between the scale focus and the edge. Depending on the species and the amount of editing that takes place, the rate varies from 80-120 scales per day. Improving the algorithm used to mark circuli could increase these rates. The algorithm we are currently using, marks the trailing edge of the circuli based on a change in luminescence value of a specific shape and magnitude along the transect line. Typically readers edit about 10% of these marker flag placements. However in some cases the number of edits is much higher, which slows production considerably. Incorporating better edge detection algorithms will take additional research and may require the use of different imaging software than Optimas.

The readers currently mark the annuli, but algorithms could also be developed to allow the computer to identify annuli. In fact, one of the early PC-based imaging programs, OPRS (BioSonics Inc., Seattle WA.), had this capability though it is no longer manufactured. Including annuli detection would not necessarily increase production rates since users would still have to verify and perhaps override the computers decision. But it does represent another area for research, and automating annuli selection might help remove another source of variability among readers.

We have conducted preliminary studies on the consistency of the measurements between two readers using our system on Bristol Bay sockeye salmon scales, and the results appear encouraging. We examined 18 variables (average circuli counts, average circuli widths, and annuli widths for each of the two freshwater and three saltwater, and plus growth zones) using paired t-tests and found no statistical difference between two readers examining the same scales (n = 298) when the data were grouped by sex, age, and year (~25 scales in each group). This is still being reviewed, but it suggests that reader variability may not be a significant concern when conducting studies on the growth histories contained in scales.

We have not yet evaluated if reader variability could be a factor in stock separation studies. In those applications, not all of the data are pooled as in growth studies, but individual fish are classified based on their own measurement set. It should be noted; however, that one of the advantages of using digital image analysis is that additional data, which do not require human judgment, can be easily collected. For instance, the luminescence profile along the transect can be saved as a vector of values. Reducing that data to its frequency components though Fourier analysis could provide another set of variables for use in stock analysis. In fact a single transect line may not even be necessary because the entire scale can be expressed by its frequency components. The first step in exploring these alternative approaches is to start applying digital methods for capturing scale images.

The system we have described here continues to be improved. Currently, the system costs approximately $18,000 for the hardware and software components, but this price continues to decline. The ability to capture high-resolution images through a microfiche scanner is a much better approach than using a microscope and a camera, which has been traditionally applied to scale image analysis. One problem we faced, however, is the amount of storage space necessary to archive the images. Given the rapid advances in digital technologies this becomes less of a problem all the time. There is also concern about the availability of imaging software for new imaging stations if Optimas is no longer available; however, there are a number of products on the market that might work as well or better. What we required from the software is not highly advanced in terms of image analysis, but the ability to create custom programs to allow rapid processing is important. We are continuing to monitor changes in the field of imaging.

We believe the digital image approach we are using could have potential benefits when applied to other projects involving growth and stock identification studies using scales. Particularly when sharing information with researchers in other countries, the benefits of using a method in which measurement criteria and decision rules can be documented and images and data exchanged electronically, are apparent. By pooling resources, large data sets of stock-specific growth patterns combined with near-shore and marine sampling programs could provide managers and researchers through out the North Pacific a greater understanding of the factors influencing salmon production. However it also important to emphasize that these benefits are likely only to be obtained if they involve the use of a compatible data structure and the willingness to collect and share the data for that purpose.

ACKNOWLEDGMENTS

Kevin Trick, Benjamin Gregg and Bill Johnson provided the computer support for this project. Clyde Andrews, Wendy Whelan, Joe Cashen, and Jaym'e Schricker assisted greatly by digitizing the scales.

LITERATURE CITED

Anas, R.E. and S. Murai. 1968. Use of scale characters and a discriminant function for classifying sockeye salmon (*Oncorhynchus nerka*) by continent of origin. Int. North Pac. Fish. Comm. Bull. 26:157-192.

Bumgarner, J.D. 1993. Long-tern trends in the growth of sockeye salmon (*Oncorhynchus nerka)* from the Chignik Lakes, Alaska. M. S. Thesis. University of Washington, Seattle. 86 pp.

Cook, R.C. 1982. Stock identification of sockeye salmon (*Oncorhynchus nerka*) with scale pattern recognition. Can. J. Fish. Aquatic Sci. 39:611-617.

Dahlberg, M.L., and D.E. Phinney. 1968. A microprojector for use in scale studies. Prog. Fish-Cult. 30:118-119.

Fukuwaka, M. and M. Kaeriyama. 1997. Scale analysis to estimate somatic growth in sockeye salmon, *Oncorhynchus nerka*. Can. J. Fish. Aquatic Sci. 54:631-636.

Ihssen, P.E., H.E. Booke, J.M. Casselman, J.M. McGlade, N.R. Payne, and F.M. Utter. 1981. Stock identification: Materials and Methods. Can. J. Fish. Aquatic Sci. 38:1838-1855.

Isakov, A. G. Ocean growth of sockeye salmon from the Kvichak river, Bristol Bay based on scale analysis. M. S. Thesis, University of Alaska, Fairbanks 126 pp.

Groot, G. and L. Margolis (eds.). 1991. Pacific salmon: life histories. Univ. of British Columbia Press, Vancouver, BC. 564 pp.

Ruggerone, G.T. and D.E. Rogers. 1998. Historical analysis of sockeye salmon growth among populations affected by the Exxon Valdez oil spill and large spawning escapements. Exxon Valdez Oil Spill Restoration Final Rept. 96048-BAA.

Schluchter, M.D. and J.A. Lichatowich. 1977. Juvenile life histories of Rogue River spring chinook salmon *Oncorhynchus tshawytscha*, as determined by scale analysis. OR Dept. Fish and Game, DACW57-77-C-0027.

Tanaka, S., M.P. Shepard, and H.T. Bilton. 1968. Origin of chum salmon (*Oncorhynchus keta*) in offshore waters of the north Pacific in 1956-1958 as determined from scale studies. Int. North Pac. Fish. Comm. Bull. 26:57-120.

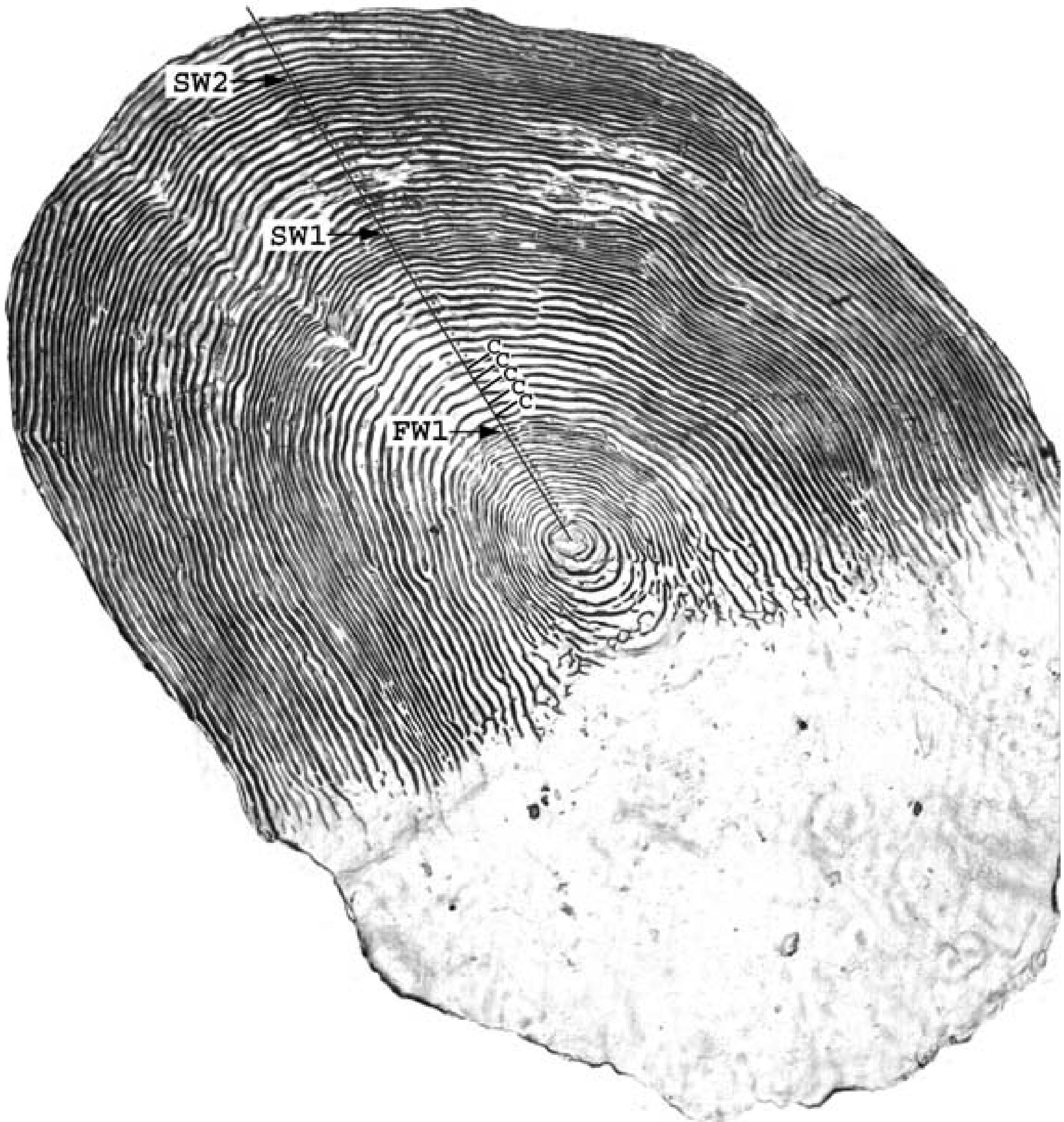

Figure 2. Scale removed from the preferred region of a female sockeye salmon collected from the Kvichak River during June 1995. FW1, SW1, and SW2 indicate the 1st year fresh water, 1st year marine, and 2nd year marine annuli, respectively. The "C" markers provide an example were the image analysis software flags and measures each circulus along the transect line. An image of the transect line and marker flags is stored in a separate binary file after the data is extracted. The line can be recombined with the image, providing documentation on the decisions used in placing annuli and circuli.