20.0
30.0

1185.

JOB PROGRESS REPORT

As required by

FEDERAL AID IN FISHERIES RESTORATION ACT

TEXAS

Federal Aid Project No. F-6-R-19

FISHERIES INVESTIGATIONS - REGION 5-B

Job No. VII K Factor Correlation

Project Leader: John C. Barron

James U. Cross
Executive Director
Texas Parks and Wildlife Department
Austin, Texas

Fred G. Lowman
Branch Head, Inland Fisheries

Roy T. Huffman
Director, Current Operations

January 27, 1972

SUMMARY

The K factor index is a method of partitioning the ecological variability of K factors of fishes and forming a "common denominator" between species. Field collections from three river basins in southern Texas have been collected and stored on computer cards. Length-weight data from these collections will form a data bank for comparing against K factors of samples.

Samples will be compared with the data bank by use of a polynomial regression model of at least third degree. Differences between the sample model and the data bank model will be measured statistically. The probability of the differences will be summed and its average will be the K factor index (50 = no difference).

JOB PROGRESS REPORT

State Texas

Project No. F-6-R-19 Project Title: Fisheries Investigations - Region 5-B

Job No. VII Job Title: K Factor Correlation

Period Covered: January 1, 1971 to December 31, 1971

Background:

Derivation of K Factor Index: If one divides the weight of an animal by its length, the resulting ratio is indicative of the animal's body condition per unit of length. In order to determine the volumetric equivalent of this ratio, the length is cubed before the division is made producing the cubic function:

$$K = W/L^3.$$

A large constant, 10^6, is usually applied to K to attain a value close to unity. This length-weight ratio is the well-known condition factor, and it is used extensively in fisheries management to estimate robustness.

In addition to body condition, numerous other variables including age, sexual dimorphism, and gamete development contribute to the variability of K within a species, and comparison between species is impossible due to different body shapes. Therein lies the major problems in working with K factors. Due to the influence of these variables, one can seldom partition the K factor values into their various components and determine how much of the magnitude is due to body condition and how much is due to ecological constraints. The K factor index is proposed as a method of partitioning ecological variability and developing a "common denominator" between species.

To make use of the K factor index, volumes of length-weight data must be collected from waters of adjacent areas (for instance within a river basin or bay system). By computer methods, parameters of the data will be stored on magnetic tape to form the basic data bank. Sample data taken from the study area may then be compared with the data bank and the differences measured statistically. These differences converted to their expected probabilities will be summed to form the K factor index.

Field Collections: Beginning in December 1963, monthly length-weight data collections using variable mesh gill nets were begun at Lake Medina. This lake is in the upper reaches of the San Antonio River basin. Draught conditions prevailed until September 1964, when the lake filled. K factors

Background: (Con.)

for most species were quite low since this lake is rather low in nutritional value. Gizzard shad (Dorosoma cepedianum), white bass (Roccus chrysops), flathead catfish (Pylodictis olivaris), and carp (Cyprinus carpio) were some of the predominant species taken.

During the 1965 calendar year, monthly collections were made at Falcon Lake in extreme southern Texas in the Rio Grande River basin. A plankton bloom had engulfed this lake due to the rising water level covering for the first time in several years, fertile flatlands around the shores. K factors were high and collections were very large. Along with gizzard shad, white crappie (Pomoxis annularis) dominated the collections with large numbers of channel catfish (Ictalurus punctatus).

A third southern Texas river system, the Nueces, was sampled in 1966 when collections were made at Lake Corpus Christi. K factors from this lake were generally between the ones from Medina and Falcon, and the samples were usually large. Blue catfish (Ictalurus furcatus) were caught in large numbers, along with gizzard shad.

Procedures:

Data Processing: Computation of K factor parameters and comparison techniques were attempted by the use of manual and unit record methods. These techniques were not feasible, however, due to the large volume of data and the great number of statistical computations necessary. It was decided to delay further work on this job until systems analysis and computer time were available.

1. The analytical technique to correlate K factors with regional values will consist of computing the statistical difference between the sample mean and the geographical area mean as determined from a regression model. The probability of the difference will be used as a random variable and its mean will be defined as the K factor index.

   The difficulty is in the construction of the regression model. A polynomial equation of at least third degree must be fitted as in the following:

$$Y = a + b_1X + b_2X^2 + b_3X^3 + \cdots + b_nX,^n$$

   where Y = the K factor and X = day of the year number. The computer programs to calculate the polynomial regression models have been obtained. To construct a cubic regression model, the following nine constants must be computed from the input data:

Procedures: (Con.)

$X_1=X$	$X_2=X_1^2$	$X_3=X_1^3$
$\sum X_1$	$\sum X_2$	$\sum X_3$
$\sum X_1^2$	$\sum X_2^2$	$\sum X_3^3$
$\sum X_1X_3$	$\sum X_2X_3$	
	$\sum X_1X_3$	

2. By knowing the above parameters plus the number of observations which went into their make-up, one can compute a third degree polynomial equation. It is these 10 numbers that will form the basis of K factor data bank. There will be a set of these parameters computed and stored for male and female of each species for each geographical division.

3. Three computer programs are actually needed for this system.

Edit Program: This program will be used to test whether sample data have been coded and punched correctly. It will check for length-weight data lying outside sensible boundaries, check for illegal code numbers, and make various other tests on the input cards. The program will write valid data on to magnetic tape. Attempts will be made to resolve any errors that are detected in this edit.

Regression Program: After all records in error have either been resolved or voided and a "clean tape" will have been built, and the tape will be sorted by sex within species within area. The previously mentioned regression parameters will be computed for each set in the sample. Magnetic tape records with these parameters will then be written to be used as input into the next program.

Update Program: Two input tapes will be used in this program: the tape with the regression parameters from the sample and the data bank tape with regression parameters from the area. Each sex-species set from the sample will be compared with its matching set from the area. The statistical difference as measured by student's t and the probability of t will be computed, printed, and stored in the machine memory. After processing each set, the sample and area parameters will be added and these sums written on an output tape forming a new data bank of updated information.

Findings:

1. During this segment, computer runs were made which showed that the basic logic for fitting K factor on time in a polynomial model was feasible. Significant fits were obtained and useful graphical printouts developed.

2. The design of tape layouts and input-output documents was begun. Other research work forced a delay on completing this portion of the objectives, but the ideas have been established and only need to be documented.

3. The completion of the computer program designs and writing the programs is the primary portion which needs to be finished. Although the most difficult part (the matrix inversion to obtain the regression parameters) has been done, tying the three programs together to form a workable system will take considerable time. This portion will be completed during the next segment.

Prepared by John C. Barron
Systems Analyst II

Approved by Fred H. Lowman

Date January 27, 1972

Elgin M. C. Dietz
Assistant for Inland Fisheries