PREPARATION OF A DRY PRODUCT FROM CONDENSED MENHADEN SOLUBLES

Statistical Supplement

SPECIAL SCIENTIFIC REPORT–FISHERIES No. 194

UNITED STATES DEPARTMENT OF THE INTERIOR
FISH AND WILDLIFE SERVICE

United States Department of the Interior, Fred A. Seaton, Secretary
U. S. Fish and Wildlife Service

PREPARATION OF A DRY PRODUCT FROM CONDENSED MENHADEN SOLUBLES: STATISTICAL ANALYSIS OF THE DATA

(Supplement to the Fish and Wildlife Service Research Report 45)

by

C. F. Lee and R. J. Monroe

Special Scientific Report - Fisheries No. 194

Washington, D. C.

December 1956

CONTENTS

TABLES

ABSTRACT

A statistical analysis has been made, as a supplement to U. S. Fish and Wildlife Service Research Report 45, Preparation of a Dry Product from Condensed Menhaden Solubles.

Simple correlation studies show that 4 of 10 chemical and physical characteristics are important in preparing suitable dry condensed solubles. These are ammonia, corrected protein content, total ash and water insoluble matter, and desirable limits are suggested for each.

Multiple correlation studies of other selected data show that the specific gravity of condensed solubles is largely determined by its total ash and fat content. Dry solids content and the refractive index or specific gravity are not correlated either separately or in combination. The study indicates that the variability in composition of condensed solubles does not account for either the extreme variation in viscosity of the condensed solubles or for differences in moisture-absorptive characteristics of the related dry solubles.

By means of an analysis of variance it was shown that the solubles produced by plants in the South Atlantic and Gulf of Mexico areas differed in composition from the product of plants in the Central Atlantic area. Otherwise, there were not sufficient data to permit determination of the effect of type of plant or of the month of production upon the properties of the condensed solubles.

PREPARATION OF A DRY PRODUCT FROM CONDENSED MENHADEN SOLUBLES: STATISTICAL ANALYSIS OF THE DATA

(SUPPLEMENT TO THE U. S. FISH AND WILDLIFE SERVICE RESEARCH REPORT 45)

by C. F. Lee[1] and R. J. Monroe[2]

INTRODUCTION

In the course of an investigation of the practicability of preparing a dry product from condensed fish solubles (50 percent solids), a number of commercially produced samples of condensed menhaden solubles were collected. The 20 plants producing these samples were located along the Atlantic Coast from New Jersey to Florida and along the coast of the Gulf of Mexico from West Florida to Texas. Most of the samples were drawn directly from storage tanks, although some were taken from the transfer pipes as the product was being pumped from the evaporators. The samples were collected from various areas during October 1952 and from July to November 1953.

Thirty-two samples of condensed solubles were used to prepare pure dry solubles on a laboratory drum dryer. This process, the description of the dry solubles produced, and the composition and physical properties of the condensed solubles used as starting material for the dry solubles are described in detail in the U. S. Fish and Wildlife Service Research Report 45, Preparation of a Dry Product from Condensed Menhaden Solubles.

The present paper is a supplement to Research Report 45 and consists primarily of a statistical analysis and an interpretation of some of the data presented in table 1, page 9 of that report. Information on the source of samples and on the 12 physical and chemical tests used to compare these samples is contained on pages 13-16 of the Research Report. On page 12 of the Report, the procedure used to derive a "Suitability Factor" for each sample of dry solubles is described. This value is derived from various characteristics observed during the drying operation and from certain properties of the dry product and is intended to give an objective comparison of the various dry solubles. The Suitability Factor permits an estimation of the suitability of the corresponding sample of condensed solubles to make a satisfactory dry product.

1/ Chemical Engineer, Fishery Technological Laboratory, U. S. Fish and Wildlife Service, College Park, Maryland

2/ Acting Head, Department of Experimental Statistics, North Carolina State College, Raleigh, North Carolina

There was a considerable range in each set of data for all characteristics of the condensed solubles that were studied and for the Suitability Factors. To illustrate the magnitude of this variability and the correlation, or more properly, the lack of a marked degree of correlation of these values to the Suitability Factor, scattergrams were drawn in which the values for each characteristic were plotted against the values of the corresponding Suitability Factor. These are reproduced in the Research Report as figures 3 to 12, inclusive.

Object of the Statistical Analysis

The object of this statistical analysis and interpretation was to determine whether significant relationships existed between the variable factors determined for the samples of condensed solubles and their suitability for preparing dry solubles, as represented by the Suitability Factor. A similar study of simple and multiple correlations was made for various other interrelated data. In some instances, these were for data where correlations would be expected, for example refractive index and specific gravity to total dry solids. In most cases inspection of the scattergrams or experience had lead to assumption of certain correlations and the statistical analysis was used to verify these assumptions and test their significance. Some discussion of these assumed correlations may be found on pages 18 and 24 of Research Report 45.

A secondary object of the analysis of data was to explore the possibility that certain of the variables studied might be selected as indicators of lots of condensed solubles that would be generally suited for making dry solubles. It was evident from the scattergrams that no single variable was highly correlated, but certain selected combinations might give increased accuracy of prediction.

Introduction to the Statistical Analysis

The large volume of data made it impractical to carry out the desired statistical work at the College Park Laboratory. It was done through a cooperative agreement with the Department of Experimental Statistics, North Carolina State College, Raleigh, North Carolina. Dr. R. J. Monroe supplied the data presented in tables 1 to 9, and has also included the introductory paragraphs immediately following to facilitate interpretation of the tables. The senior author had appended a discussion of a less technical nature of the data in the tables to emphasize practical aspects and implications.

Simple correlation, multiple correlation and analysis of variance techniques were used to provide evidence of association between physical and chemical characteristics and to measure the effects of certain external factors on these characteristics.

Simple Correlations

The simple correlation coefficients were determined from the sample n = 32. Each simple correlation coefficient was tested for significance using the standard tables at 30 degrees of freedom. The results are shown in table 1. One asterisk indicates that the sample value exceeds the table value at the 5-percent level of significance; 2 asterisks, at the 1-percent level; and 3 asterisks, at the 0.1-percent level.

Multiple Correlations

The multiple correlations vary in degrees of freedom, depending on the number of variables included. In general, degrees of freedom = n - p - 1, where p is the number of independent variables included. The data are tabulated in tables 2 to 5 inclusive.

Discussion of Tests for Significance

The testing of either simple or multiple correlation coefficients using the standard tables and levels of significance is always a test of the null hypothesis that the true correlation in the population is exactly zero. Hence it is frequently observed that "significant" correlations do not always indicate a strong association. Another measure of the strength of the association is the square of the correlation coefficient, which rather roughly expresses the percent variability in the dependent variable associated with the independent variable (or variables). For example, the simple correlation between water insoluble matter and Suitability Factor is -0.4086 significantly different from zero at the 5-percent level. Yet, $r^2 \times 100$ = 16.7 percent of the variability in Suitability Factor is associated with water insoluble matter leaving 83.3 percent still to be accounted for. The value of r^2 or (R^2) therefore should always be kept in mind when interpreting the significance of correlation coefficients.

Analysis of Variance

The analysis of variance was applied to the information on history of samples, to determine what affect these factors had on sample variability. The technique used was that of a single classification. The data were not extensive enough to allow a cross-classification of areas, processes, and dates. No attempt was made to evaluate the factor plant because so many plants were represented by only one sample. These data are tabulated in tables 6 to 9.

DISCUSSION

Simple Correlations Related to the Suitability Factor

The analyses show that there was less correlation than had been supposed originally between the Suitability Factor and the 10 characteristics listed in table 1. Rather unexpectedly, the total ash had the highest simple correlation coefficient, namely r = .60*, significant at the 0.1-percent level. Ammonia was the only other characteristic which was positively correlated, namely r = .40, significant at the 5-percent level. A set of values not charted showed the greatest negative correlation; namely, corrected protein with a value of r = -.56, indicating that, when the corrected protein (crude protein corrected for ammonia nitrogen) is high, the solubles usually gave a good dry product. Only one other characteristic, water insoluble matter, had a negative correlation, significant at the 5-percent level.

These four characteristics, that is, a sample having low values for total ash and ammonia, and high values for corrected protein and water-insoluble matter, when taken together, indicate that it should dry on a drum dryer to give a satisfactory dry product. It might be noted that the very high simple correlation coefficient of total ash to specific gravity (r = .75), which is developed in the first of the multiple correlation studies, results in the rather unusual relationship of specific gravity to the Suitability Factor (r = .32). This value just misses being significant at the 5-percent level. There was almost no relationship whatever of dry solids, fat, pH, refractive index, or viscosity to predict whether a sample of condensed solubles will make a satisfactory dry product. ($R^2 \times 100$ = 0 to 4 percent).

Minor Simple Correlations

Other simple correlations showed, as expected, a highly significant relationship of fat to separation of oil during drying (r = .82) and an almost as positive relationship of ammonia to the production of objectionable odors noted during the drying operation (r = .62). As was pointed out on page 24 of the Research Report, high values for ammonia resulted from more or less deliberate spoilage of the stickwater before evaporation, which is sometimes used as a means of reducing viscosity of the finished condensed solubles. Ammonia content of the condensed solubles had a barely significant positive correlation to the rather annoying hot plasticity of the dry solubles on the drums, often observed during the drying process.

---

* Note all values of r have been rounded to the nearest hundredth.

Table 1.--Simple correlation values

Factor correlated	r	r^2 x 100
		Percent
Suitability Factor to:		
Dry solids	-0.1779	3
Ammonia	0.4008*	16
Corrected protein	-0.5595***	31
Fat	0.0922	1
Total ash	0.5975***	36
Water insoluble matter	-0.4086*	17
pH	0.0156	0
Refractive index	0.1683	3
Specific gravity	0.3208	10
Viscosity	-0.2089	4
Fat to:		
Oil separation	0.8199***	67
Ammonia to:		
Hot plasticity	0.3714*	14
Vapor odor	0.6177***	38

* significant at the 5-percent level; **, at the 1-percent level; and ***, at the 0.1-percent level.

Multiple Correlations

Several very interesting multiple correlation relationships were found.

Factors Determining Specific Gravity

The data in table 2 confirm the conclusions stated in the section on Specific Gravity on page 27 of the Research Report, that specific gravity is affected by fat content and total ash. Total ash was positively correlated and fat negatively correlated to the specific gravity to an extent that the combined value for $R = .93$, or R^2 accounted for 86 percent of the variability. This value is high enough that formulas could be derived predicting the third value based on any other two, if any practical purpose was to be gained by such calculations.

Factors Affecting Total Dry Solids

The conclusion reached from inspection of figure 13 and stated on page 24 of the Research Report, namely that there was little correlation between the dry solids content and either the specific gravity or refractive index, was also confirmed by the statistical analysis, table 3. There was no significant simple correlation of either one to the dry-solids content, and the combination gave no appreciable improvement. The R^2 value was only 14 percent.

Factors Affecting Viscosity

The next study was suggested in the hope that differences in composition might explain the very great range in viscosity of the condensed solubles that had been observed in the samples tested. The conclusion reached from the data in table 4 was that none of the variables tested, that is, dry solids, ammonia, fat, total ash, or water insoluble matter, either individually or collectively could be significantly associated with variability in viscosity.

Factors Affecting Moisture Absorption

In a similar manner, four variables were selected for study to determine their effect on the moisture absorption of the dry solubles. The data in table 5 show that individually, none of them, ammonia, corrected protein, fat, or total ash had a significant simple correlation coefficient, but it was suggested that the improvement considering all variables is quite good. The multiple correlation just lacked significance at the 5-percent level, and consequently, it is not of great practical value in predicting the probability that a given sample of condensed solubles would yield dry solubles having low moisture absorption.

Table 2.--Multiple correlation with specific gravity

Factor correlated	Fat	Total ash	Specific gravity
Dry solids	0.6890***	-0.2618	-0.2790
Fat		-0.1349	-0.5461**
Total ash			0.7535***

Note: R = 0.9282*** with 28 d.f. and R^2 x 100 = 86 percent. Total ash alone will account for about 57 percent, but the addition of dry solids and fat increases this to 86 percent.

Table 3.--Multiple correlation with dry solids

Factor correlated	Specific gravity	Dry solids
Refractive index	0.0347	0.2414
Specific gravity		-0.2790

Note: $R = 0.3755$ with 29 d.f. and $R^2 \times 100 = 14$ percent. Including both refractive index and specific gravity results in no appreciable improvement.

Table 4.--Multiple correlation with viscosity

Factor correlated	Ammonia	Fat	Total ash	Water insoluble matter	Viscosity
Dry solids	-0.6324***	0.6890***	-0.2618	0.3495*	0.1352
Ammonia		-0.4563**	0.6329***	-0.2520	-0.2247
Fat			-0.1349	0.4834**	0.0103
Total ash				-0.3882*	-0.1121
Water insoluble matter					0.1953

Note: R = 0.3703 with 26 d.f. and R^2 x 100 = 14 percent. Viscosity cannot be associated with any of the other variables singly or as a group.

Table 5.--Multiplo correlation with moisture absorption

Factor correlated	Corrected protein	Fat	Total ash	Moisture absorption
Ammonia	-0.3484	-0.4563	0.6329	-0.1235
Corrected protein		0.2301	-0.7049	-0.0296
Fat			-0.1349	-0.0423
Total ash				0.2460

Note: R = 0.5079 with 27 d.f. and R^2 x 100 = 26 percent. Here the improvement considering all variables is quite good. The r = 0.5079 just misses the tabular value of 0.536 for the 5-percent level at n - p - 1 = 27 and p = 4. The R^2 x 100 = 26 percent is not particularly striking.

Study of Source Factors of Solubles by Analysis of Variance

Area Effects on Composition of Solubles

The remainder of the statistical analysis was devoted to a study of the effect of areas, processes, and month of sample collection on various characteristics of the condensed solubles. The statistical analysis indicates that the number of samples from any single area of plant location is too small to justify any definite conclusions as to the effect of area. This conclusion is particularly true for areas 5 and 6, in the Gulf of Mexico. In the east Gulf (area 5) for example, 3 of the 4 samples came from different storage tanks at the same plant on the same date, and all of the west Gulf samples were collected at the same time, 4 of them from 2 plants. This may result in what should be considered as a processing factor being listed as an area effect.

As an example of other interrelations of sample history and solubles properties, the plants in the South Atlantic and Gulf areas, have difficulty "cooking" their solubles to the desired 50-percent solids concentration, table 6. This may be related to the high air and water temperatures in these areas, since these high temperatures result in "soft" fish as a normal condition, and some of the fish are in pretty poor condition before they are pumped from the boat hold. Even when the fish are in fairly good condition when landed, the press water is often held in temporary storage or in settling tanks until partial protein breakdown occurs. Some of the plants in these areas use a gravity separation process and this makes it much easier to separate the oil. It also reduces the viscosity of the finished solubles so that the stickwater can be handled in evaporators of the air-stream type, all of which were located in the Southern areas. Another effect is to increase the ammonia content as a result of this protein breakdown (table 6). On the other hand, there is one atypical plant in each of the Gulf of Mexico areas. These two plants process stickwater almost directly from the centrifuges, and it will be noted in the table on page 9 of the Research Report that these plants, code nos. EG 2 and WG 2, produced condensed solubles with a lower ammonia content, higher fat content and higher viscosity as compared with other samples from these areas. There is no explanation that can even be suggested to account for the apparent area differences in the total ash content.

Effect of Process

The section headed "Process Means" needs an explanation of the code numbers, since only two processes, the multistage vacuum evaporators and air-stream evaporators, are involved. The vacuum-evaporated samples are further divided into (1) samples drawn from storage tanks and (2) samples taken directly from the process line from evaporator to storage. Only four samples were obtained from air-stream evaporators, and all were from storage tanks.

The notable features are the high ammonia content and low fat content of the hot air-evaporated samples, and the high viscosity of the samples taken directly from the vacuum evaporators. Interesting, even if non-significant, is the lower ammonia content of the process samples compared to the storage samples and the somewhat higher total ash of the hot-air-evaporated samples. The averages for the ash content of two groups of vacuum-evaporated samples are nearly equal, while the average for the four hot-air-evaporated samples is about 25 percent higher. There is a temptation to ascribe this to the scrubbing action of the stickwater spray on the hot-air stream, which would undoubtedly remove dust, ash, etc., from the drying air. This may, indeed, have an effect, but the range in total ash values from 5.5 to 13.0 for condensed solubles produced in vacuum evaporators makes it plain that other unknown sources of variability in total ash content are involved.

Effect of Month of Collection

The month during which the samples were collected had no effect, and none would be expected. All samples from the Southern areas were collected during one trip, and only from the Chesapeake Bay area was a sufficient number of samples obtained so that the possible effect of season might have been developed. This effect, if any, was lost in averaging all samples for a given month in the statistical analysis.

Table 6.--Analysis of effects of areas, processes, and dates of collection

Variable	Area means*						Test
	1	2	3	4	5	6	
Number of samples	1	4	14	4	4	5	
Dry solids	52.5	49.8	49.2	43.8	45.3	44.9	**
Ammonia	0.39	0.57	0.93	1.34	1.12	2.72	**
Fat	8.3	7.2	8.4	2.8	3.0	5.5	**
Total ash	7.7	9.2	9.3	10.6	6.1	10.6	**
Water insoluble matter	3.8	5.3	4.4	3.7	4.5	5.0	N.S.
Viscosity	775	829	70	137	70	167	N.S.
(Omitting 3000° value)	775	106	70	137	70	167	**

* The areas represented by the numbers are as follows: 1) Long Island, Northern New Jersey, 2) Southern New Jersey, Delaware, 3) Chesapeake Bay, 4) North Carolina, Florida, 5) Gulf of Mexico east of Mississippi delta, 6) Gulf of Mexico west of the delta.

Note: Analysis of variance ignoring processes and adjusting for the unequal numbers in each area indicates that areas may have some effect in all cases except water insoluble matter. Note that including the one observation of viscosity equal to 3000° so increases the variability that even the large differences cannot be regarded as significant. Omission of this observation changes the significance of values for area 2, but not for the other areas.

Specific comments on data in table 6:

Dry solids — The values for areas 4, 5, and 6 appear to differ from those of areas 1, 2, and 3.

Ammonia — The values for area 6 appear to be a good bit higher. The other areas are not different among themselves.

Fat — The values for areas 4 and 5 are significantly lower than for the others.

Total ash — The values for areas 1 and 5 are significantly lower than for the others.

Viscosity — After omitting the value of 3000° in area 2 sample the analysis indicates that the values in area 1 may be regarded as higher than the others which do not differ among themselves.

Table 7.--Process means

Variable	Mean process number			Test
	1	2	3	
Number of samples	23	5	4	
Dry solids	48.2	47.4	44.08	N.S.
Ammonia	1.10	0.65	2.61	**
Fat	7.1	6.5	2.7	*
Total ash	8.9	9.1	11.3	N.S.
Water insoluble matter	4.7	3.9	4.2	N.S.
Viscosity	87	849	123	N.S.
(Omitting 3000° value)	87	312	123	*

Note: The test results are for processes after ignoring areas and dates.

Specific comments on data in table 7:

Ammonia — The values for process 3 are definitely higher than for 1 and 2; and processes 1 and 2 do not differ significantly.

Fat — The values for process 3 are significantly lower than for 1 and 2.

Viscosity — The values for process 2 still appears to be the highest when the 3000° value is omitted.

Table 8.--Means by month of sample collection

Variable	July	Aug.	Sept.	Oct.	Nov.	Test
Number of samples	4	13	7	7	1	
Dry solids	49.5	46.8	46.6	49.4	43.7	N.S.
Ammonia	1.26	1.57	1.16	0.72	0.42	N.S.
Fat	8.5	5.5	6.2	7.9	1.3	N.S.
Total ash	10.1	8.7	9.7	9.1	9.6	N.S.
Water insoluble matter	4.7	4.8	4.2	4.7	1.0	N.S.
Viscosity	88	170	98	490	58	N.S.
(Omitting 3000° value)	88	170	98	72	58	N.S.

Note: The means do not differ significantly compared to the "within the month" variation. Through November has several means much lower than the others, the fact that this is based on only one sample means that it carries very little weight in the overall analysis.

Table 9.--Number of observations in the area-process subclasses

Process	Area number*						Total
	1	2	3	4	5	6	
1	0	2	14	1	4	2	23
2	1	2	0	1	0	1	5
3	0	0	0	2	0	2	4
Total	1	4	14	4	4	5	32

* The areas represented by the numbers are as follows: 1) Long Island, Northern New Jersey, 2) Southern New Jersey, Delaware, 3) Chesapeake Bay, 4) North Carolina, Florida, 5) Gulf of Mexico east of Mississippi delta, 6) Gulf of Mexico west of the delta.

Note: It is clear from this table that means for process 3 included in only areas 4 and 6. Hence the conclusions regarding the effects of process 3 are not independent of similar conclusions regarding areas 4 and 6. Again, areas 3 and 5 contain only process 1, hence a comparison of areas 3 and 5 depends on a single process and cannot be generalized to include other processes. Some caution therefore should be exercised in claiming great generality for the results of these analyses.

SUMMARY

1. The following significant simple correlations were found:

   a. Total ash and ammonia were positively correlated (low value for desirable effect), and corrected protein and water insoluble matter were negatively correlated (high value for desirable effect) to the Suitability Factor.

   b. Fat content of condensed solubles was positively correlated to oil separation during drying.

   c. Ammonia content was positively correlated to the development of objectionable vapor odors during drying.

   d. Ammonia content was positively correlated (but at a barely significant level) to hot plasticity of the dry solubles on the dryer drums.

2. The multiple correlation studies indicated:

   a. Total ash was positively correlated and fat content negatively correlated to the specific gravity of the condensed solubles. The combination of these two factors accounted for almost all the variability in specific gravity.

   b. Although four selected characteristic components, namely, ammonia, corrected protein, fat, and total ash, were individually not significantly correlated to moisture absorption of the dried solubles; collectively, they almost had a significant correlation at the 5-percent level.

3. An analysis of variance indicated:

   a. The location of plant had an effect on dry solids; lower values being observed in the three Southern areas.

   b. Ammonia was significantly higher, and fat lower in solubles produced in hot-air evaporators as compared with those produced in multistage vacuum evaporators.

   c. Viscosity was significantly higher in condensed solubles sampled directly from the evaporators.

4. The statistical analysis showed that there were no significant correlations in the following data:

   a. Neither the specific gravity nor the refractive index was correlated to the total solids content of the condensed solubles.

b. None of five selected characteristic components of the condensed solubles, namely total solids, ammonia, fat, total ash, or water-insoluble matter, neither individually or collectively accounted for the large range in viscosity of the condensed solubles.

c. In general, the number of observations in each subclass for area and process were too few, with several areas not having any of the process groupings, to permit any generalizations as to the significance of these data except as indicated in 3a, b, and c.

d. The data were insufficient to permit determining any seasonal effect on composition of condensed solubles.

CONCLUSION

The original data published in Research Report 45 show that the 32 randomly collected samples of condensed solubles differed considerably in chemical and physical characteristics, and produced dry solubles of widely varying properties. No single characteristic of the condensed solubles showed a degree of correlation to the desired properties of a dry product that would justify its selection as a criterion of choice of condensed solubles for use in producing dry solubles. The statistical analyses suggest, however, that condensed solubles with low ammonia and high corrected protein content, and low ash and high water insoluble matter content would dry easily and produce a satisfactory dry product. The first two relate to the freshness of the stickwater from which the condensed solubles are produced. Specifically, an ammonia content less that 1.25 percent and total ash content less than 8.5 percent; and corrected protein content more than 29 percent and water insoluble matter more than 4.5 percent would be the suggested minima and maxima. It has not been possible to test the practical aspects of this conclusion since none of the producing plants is equipped with drum drying equipment at the present time.

INT.-DUP. SEC., WASH., D.C. 9393