

## THE 2008 AAFCO FEED COLLABORATIVE CHECK SAMPLE PROGRAM

### Background

The Association of American Feed Control Officials (AAFCO) Collaborative Check Sample Program is an analytical series of 12 feed samples available by yearly subscription. The series includes a wide variety of feeds, many or most containing minerals, vitamins and drugs or antibiotics currently in use. To allow sufficient time for analysis, a sample is sent two months before the results are due and one month before the analyses are to be done, e.g., the sample to be analyzed in July is shipped in June and the report is due in August.

Samples are commercially manufactured products carefully ground and blended to ensure uniformity of composition throughout all collaborators' packets by Able Laboratory, Pikeville, Tennessee ([www.ablelaboratory.com](http://www.ablelaboratory.com)). This firm, which also prepares samples for other check sample programs such as the AOCS Smalley Series and TFI Magruder Fertilizer Series, has earned a justifiably fine reputation for its care in preparing and handling such samples.

### Participation

Every participating laboratory may choose to do as few or as many determinations as it wishes. However, a laboratory which performs as many determinations as possible will receive the most benefit from the statistical evaluation of the results.

Each participant is asked

- (1) to perform **two and only two** analyses on different days for each constituent of interest and record the results on an Excel report available by e-mail (5-digit codes are used to specify the methods used according to a code sheet supplied to each participant upon annual subscription to the program and available for down-load from the AAFCO website);
- (2) to forward a single **legible** copy of the report sheet to the Check Sample Committee Chair, preferably by e-mail ([vsiegel@purdue.edu](mailto:vsiegel@purdue.edu)), no later than the 10th of the month following receipt of the sample.

### Reporting Results

Participants report their results using an Excel spreadsheet form. Forms and program information may be downloaded from the AAFCO website ([www.aaftco.org](http://www.aaftco.org), News and Information, Check Sample Program). The participant should indicate as precisely as possible the method used by selecting an appropriate five-digit code from the code sheet. All results not coded with a method extension are assigned "99" in the last two positions of the method code number (the extension), implying an "unidentified" or "miscellaneous" method was used for the pair of determinations. All results entered as "unidentified" or "miscellaneous" are not included in any statistical treatment or summary. All single results are summarily rejected.

Those using official AOAC methods should ensure the method code corresponding to the AOAC method is entered. Those using other methods should locate an appropriate listing. All participants need to ensure that reported units correspond to those on the method code sheet.

Please contact the Program Chair with suggestions for method code needs. The Chair will survey participants to confirm whether sufficient data would be reported using the suggested method code prior to updating the method code document.

Participants are asked to check over their reports before submitting to the Chair to ensure that duplicate results are reported, that the AAFCO confidential laboratory number is properly entered, that the results have two digits following the decimal in the method code entered, and that no gross errors are present.

## **Treatment of the Data**

The Chair forwards the received reports to AAFCO's statistical consultant for computerized statistical treatment. The computer program is identical to that employed in the Magruder Fertilizer Series.

The aggregate results for each method code are first screened for extreme values by an estimation procedure not affected by outliers. The screen accomplishes this by counting the number of results by a given method of analysis (code) and ranking results high to low. By counting in from each end of the ranking by sixteen percent of the count, the program isolates the 68% at the center. The mean is calculated from this 68% and the standard deviation estimated by taking one-half the difference between top and bottom values of the 68%. (The center 68% is the fraction of normally distributed results which lies between one-sigma boundaries on either side of the mean. Consequently, the values at the count boundaries are two standard deviations apart. One-half of the difference between them estimates the standard deviation of the set of analyses without any influence from outliers. The mean of the central 68% is also entirely free of outlier effect. However, this calculation cannot be used if there are less than six reported values.)

Because the count-boundary estimate, although not affected by outliers, is based on only part of the available information, the program uses the estimates determined from the above data for mean and standard deviation to identify outliers through a preliminary screening at 3.5-sigma control limits. This technique avoids over-censoring laboratory analyses. All of the data lying between  $\pm 3.5$  sigma are used to calculate a mean, standard deviation, and average range of duplicates for the method which in turn serve as the basis for determining the 3.0-sigma control limits which the program uses to identify outlier averages and outlier average ranges of duplicates not identified as such during the preliminary screening.

Outlier averages and outlier average range of duplicates found by this first pass through the full-scale statistical calculations are omitted from a second pass that produces the summary statistics – the grand averages, estimates of standard deviation and average range of duplicates – for each method code and for the entire method set.

## **Evaluating Results**

### *a. Summary Report*

The report lists the mean and standard deviation for each method reported by those laboratories submitting data for the current check sample. The precision within laboratories is

reported as the average range because it relates directly to a laboratory's difference between duplicates, the range being defined as the absolute value of the difference between results. An average range is larger by a factor of 1.128 than the standard deviation which can be calculated from the same data. The upper 99.7% control limit for ranges of duplicates is the average range multiplied by 3.267. (The lower control limit is zero).

*b. Performance Card*

Each participant in the program receives a confidential performance report (score card). All values reported are included in the calculation of a laboratory's performance unless specifically excluded by the reporting laboratory's placing an X in the appropriate column. Averages and ranges which are outliers are flagged in the report next to the related index value.

The following brief description defines (a) the terms used and (b) the basis for the numbers and letters that are on your monthly report card.

(1) The z Value

Each laboratory's results for a given method are judged on how well they agree with the grand average of all results for that method. This is done by taking the difference between **each** (not the average of the two) individual result and the grand average, i.e.,  $x_n - \bar{x}$ , and dividing by the standard deviation. The quotient is called a normalized value, an index value, or a z value. The z value tells how well a result agrees with the grand average, i.e., a low z value indicates that a result lies close to the mean while the sign indicates the direction of bias.

(2) Precision

Precision describes the repeatability of results compared to the grand average standard deviation.

There is often confusion concerning the "precision" ranking on a lab report card. It is **not true** that the closer the two replicate results are for a particular method, the better the precision should be on the report card. The precision on the report card measures how close a laboratory's two results lie to the grand average – **not to each other**. The precision within a lab is reported as the average range because it relates directly to a laboratory's difference between duplicates, the range being defined as the absolute value of the difference.\*

(3) Accuracy

Accuracy as used on the report card is a combination of bias and precision. The size of both the bias and the precision determines accuracy. The relationship is:

$$accuracy\ index = \sqrt{bias^2 + precision^2}$$

---

\* See Appendix

#### (4) Letter Grades

Letters which accompany evaluations of bias, precision and accuracy help to indicate performance. The number of laboratories whose evaluations are within 99.7% confidence limits is divided into thirds. The best one-third is marked A; the next best one-third is marked B; and the remaining one-third is marked C. Evaluations which are outside the 99.7% confidence limits are marked D. Plus/minus signs indicate the direction of the bias; thus, a B+ is not better than a B, it means the reported average is higher than the grand average. Therefore, performance is graded in comparison to the results of other reporting labs and not to any pre-determined level of performance.

#### **Information**

If a sample is lost or damaged in shipping, or if you wish to receive additional material, send an email to the Program Chair, including your lab code number, the AAFCO number of the damaged or missing sample, your name and Company name and current mailing address.

Dr. Victoria S. Siegel  
Chair, AAFCO Feed Check Sample Program  
Office of Indiana State Chemist  
Purdue University  
175 S. University St.  
West Lafayette, IN 47907-2063

Voice: 765-494-1565  
Fax: 765-494-8722  
e-mail: vsiegel@purdue.edu

Many questions can be answered using information available on the AAFCO website. Forms and program information may be downloaded from the site ([www.aaftco.org](http://www.aaftco.org)). The Check Sample Program page is available on the public access site by selecting "News and Information, Check Sample Program" from the drop-down list. You do not need to register or log-in to access this information.

## APPENDIX: ILLUSTRATIONS OF CALCULATIONS

### A. Bias, Precision and Accuracy for a Laboratory's Results Compared to Other Results of the Same Method Code

$x_1$  and  $x_2$  represent data from a check sample analyzed in duplicate by one laboratory. The accuracy, bias and precision are calculated as follows:

$$\begin{array}{cccccc} \bar{x}^* & x_1 & x_2 & z_1 & z_2 & \\ \hline 16.1082 & 16.07 & 15.85 & -0.142103 & -0.960494 & \end{array}$$

\*For 150 labs with a standard deviation of 0.26882

$$accuracy\ index = \sqrt{\frac{z_1^2 + z_2^2}{2}} = \sqrt{\frac{0.0201931 + 0.922549}{2}} = 0.68656$$

$$bias = \frac{z_1 + z_2}{2} = \frac{0.142103 + 0.960494}{2} = -0.55130$$

$$precision = \sqrt{\left(\frac{z_1^2 + z_2^2}{2}\right) - \left(\frac{z_1 + z_2}{2}\right)^2} = \sqrt{0.1674} = 0.4092$$

### B. The Effect of Within Laboratory Duplication (Repeatability) on Overall Performance

Laboratory	Analyte	$x_1$	$x_2$	$z_1$	$z_2$	$\bar{x}^*$	$s^*$
A	Protein	15.48	15.46	-0.57427	-0.65799	15.6172	0.23891
	Fat	27.2	27.22	2.469149	2.542822	26.5297	0.27147
	Fiber	49.46	49.47	2.032298	2.064531	48.8295	0.31024
B	Protein	15.48	15.46	-0.57427	-0.65799	15.6172	0.23891
	Fat	27.2	<b>26.53</b>	2.469149	0.001105	26.5297	0.27147
	Fiber	49.46	<b>48.83</b>	2.032298	0.001612	48.8295	0.31024

\*For 200 labs

	Normalized Values	
	Laboratory A	Laboratory B
Bias	1.3128	0.5453
Accuracy Index	1.9025	1.3534
Precision	1.3771	1.2386

Lab A's results show very good repeatability yet the overall "precision" as described in this report is poorer than Lab B. Lab B's repeatability for fat and fiber are much further apart, yet the overall precision improves considerably.

A second result obtained under conditions other than those prescribed by the check sample protocol may be the same as or very close to the first, but if it does not include all sources of variation, the agreement may actually hurt the performance.