



**KOOGLER & ASSOCIATES**  
**ENVIRONMENTAL SERVICES**

4014 NW THIRTEENTH STREET  
GAINESVILLE, FLORIDA 32609  
352/377-5822 • FAX/377-7158

KA 123-97-01  
MEMORANDUM

RECEIVED

SEP 02 1998

BUREAU OF  
AIR REGULATION

TO: Syed Arif, FDEP

FROM: Pradeep Raval

DATE: August 27, 1998

SUBJECT: North MAP/DAP Plant  
Farmland Hydro, L.P.

This is a follow up to our telephone conversation yesterday regarding the suggested wording for measures to be implemented to improve scrubber performance during MAP production. Please consider the following:

"The permittee shall install improved spray nozzles in the HI-MOL scrubber system in order to reduce fluoride and particulate matter emissions during MAP production. Upon completion of performance testing, the Department shall review the performance test data and, if necessary, require additional improvements to the existing air pollution control equipment to achieve an allowable fluoride emission limit during MAP production which is closer to 0.0417 lb F/ton P205. The Department may also review the particulate matter emission limit, if warranted.

The performance testing during MAP production, not to be used for compliance purposes, shall consist of four quarterly tests over a 12-month period, using EPA Method 13A or 13B for fluorides; and, EPA Method 5 for particulate matter. Each test shall consist of three complete runs, pursuant to Rule 62-297, FAC. A report shall be submitted to FDEP's Bureau of Air Regulation to document the test results and data analysis to determine the appropriate allowable fluoride and particulate matter emission limits during MAP production. Compliance tests during MAP and DAP production shall be conducted subsequent to the performance testing. At this time the particulate limit during DAP production will also be reviewed. The report shall document the scrubber operating parameters during the tests, as required by this permit."

Please revise the permit expiration date to May of 2000 to accommodate the testing.

If you have any questions, please call me.

par.



KOOGLER & ASSOCIATES  
ENVIRONMENTAL SERVICES  
4014 NW THIRTEENTH STREET  
GAINESVILLE, FLORIDA 32609  
352/377-5822 • FAX/377-7158

KA 123-97-01

August 18, 1998

RECEIVED

AUG 19 1998

BUREAU OF  
AIR REGULATION

Mr. A. A. Linero  
Florida Department of  
Environmental Protection  
2600 Blair Stone Road  
Tallahassee, FL 32399-2400

Subject: U.S. Fish and Wildlife Service,  
Air Quality Branch  
Letter to Al Linero from Ellen Porter  
Dated July 28, 1998

Dear Al:

I have had a chance to review the letter to you from Ellen Porter dated July 28, 1998, regarding Farmland's draft MAP-DAP permit and have a couple comments I would like to make. I understand from Pradeep that matters addressed in this letter might have already been resolved and I hope this is the case.

The comments that I have relate both to the Farmland matter and to statistical analyses presented to the Department or prepared by the Department relating to any matter.

Porter's letter states:

"... the calculations of the stack test result confidence intervals are seriously flawed."

My comment is that the analysis presented in Porter's letter, while statistically correct, is not an estimator of the confidence interval for the variability (or spread) of individually measured emission rates. The estimator presented in Porter's letter establishes confidence limits for the mean (or average) of an entire set of emission data.

In other words, let's assume a company provided a set of data to the Department representing emissions from a particular operation and further assume the mean (or average) and the standard deviation of the data set had been calculated. The company then established the 95th percentile confidence limits for the data set; the interval into which 95 out of every 100 emission measurements would fall. Five out of every 100

Mr. A. A. Linero  
Florida Department of  
Environmental Protection

August 18, 1998  
Page 2

emission measurements would fall outside of this interval (2.5 percent above the interval and 2.5 percent below the interval).

The 95th percentile confidence interval is calculated as the mean (or average) of the data set plus or minus approximately two standard deviations. In other words, the lower bound of the confidence interval would be the mean minus approximately two standard deviations and the upper bound of the confidence interval would be the mean plus approximately two standard deviations. This is the approach that has been used by the Department to set emission limits for existing operations undergoing BACT analyses.

Another statistical test that can be conducted on the data set is to evaluate the reliability of the sample mean, or the reliability of the average of the data set. The data set presented by a company represents a small sampling of the emission rates that occur every hour that the plant is operating under normal conditions. If emission data were available for every hour the plant operated and if all of these emission rates were averaged, one would have the true mean (or true average) emission rate for the plant. The mean (or average) of the data set presented by the hypothetical company represents an estimate of the true average emission rate from the plant. Using statistical analyses, one can establish confidence limits for the range of the true sample mean based on the mean calculated from the limited data set. The analysis presented in Porter's letter establishes confidence limits for the mean (or average) emission rate and does not establish confidence limits for individual emission rate measurements making up the data set.

The attached example is presented to demonstrate the difference between the confidence interval for an entire data set and the confidence interval presented in Porter's letter representing variability in the sample mean.

These comments are provided to you to clarify a matter that hopefully is no longer an issue in the case of Farmland. It is an issue that I am sure will arise many times in the future, however. If you have any questions or comments regarding this information, please do not hesitate to contact me.

Very truly yours,

KOOGLER & ASSOCIATES

  
John B. Koogler, Ph.D., P.E.

JBK:wa  
Enc.

cc: S. Arif, BAK  
SWD  
polk Co.



**HYPOTHETICAL DATA SET REPRESENTING 17 HOURLY  
EMISSION RATE MEASUREMENTS ON A PLANT THAT  
OPERATES THOUSANDS OF HOURS (AND THEREFORE  
HAS THOUSANDS OF HOURLY EMISSION RATES)**

SAMPLE NUMBER	EMISSION RATE (LB/HR)
1	16
2	22
3	21
4	20
5	23
6	21
7	19
8	15
9	13
10	23
11	17
12	20
13	29
14	18
15	22
16	16
17	25
AVG	20

Number of samples (n)	= 17
Average emission rate ( $\bar{x}$ )	= 20 lb/hr
Standard deviation (s)	= 3.98 lb/hr
Number of standard deviations from mean(+ and -) that will include 95% of individual samples (t)	= 2.12*
Standard error ( $s_x$ ) = $s/\sqrt{n}$	= 0.97

\* "t" value for probability of 0.05 with 17-1 = 16 degrees of freedom

Based on the 17 emission rates that represent a very large data set (the emission rate for every hour a plant operates), the mean (average or  $\bar{x}$ ), standard deviation (s) and standard error ( $s_x$ ) have been calculated. The interval into which 95 out of every 100 measured emission rates will fall is given by the equation:

$$(\bar{x} - 2.12s) \leq \mu \leq (\bar{x} + 2.12s)$$

Where  $\mu$  = the true mean of the total data set (the average of the emission rates for every hour the plant operated). The true mean is estimated by  $\bar{x}$  from the sample population of 17 ( a subset of the total data set).

In our example, the 95th percentile confidence limits for individual hourly emission rates are:

$$\text{Lower C.L.} = 20 - 2.12 (3.98) = 11.56 \text{ lb/hr}$$

$$\text{Upper C.L.} = 20 + 2.12 (3.98) = 28.44 \text{ lb/hr}$$

In other words, for every 100 emission rates measured, 95 will be between 11.56 and 28.44 lb/hr, two or three will be less than 11.56 lb/hr and two or three will be greater than 28.44 lb/hr.

The next question is, how well does  $\bar{x}$  estimate  $\mu$ ? - how well does the average emission rate of the 17 samples in the limited data set ( $\bar{x} = 20$ ) represent the average ( $\mu$ ) of the total data set? The 95th percentile confidence interval for this can be calculated using the equation:

$$(\bar{x} - 2.12 s_x) \leq \mu \leq (\bar{x} + 2.12 s_x) \text{ or}$$

$$(\bar{x} - 2.12 s/\sqrt{n}) \leq \mu \leq (\bar{x} + 2.12 s/\sqrt{n})$$

(This is the equation proposed in Porter's letter for calculating the 95th percentile confidence interval for the individual samples in the data set).

The 95th percentile confidence interval for the sample mean is:

$$\text{Lower C.L.} = 20 - 2.12 (0.97) = 17.94 \text{ lb/hr}$$

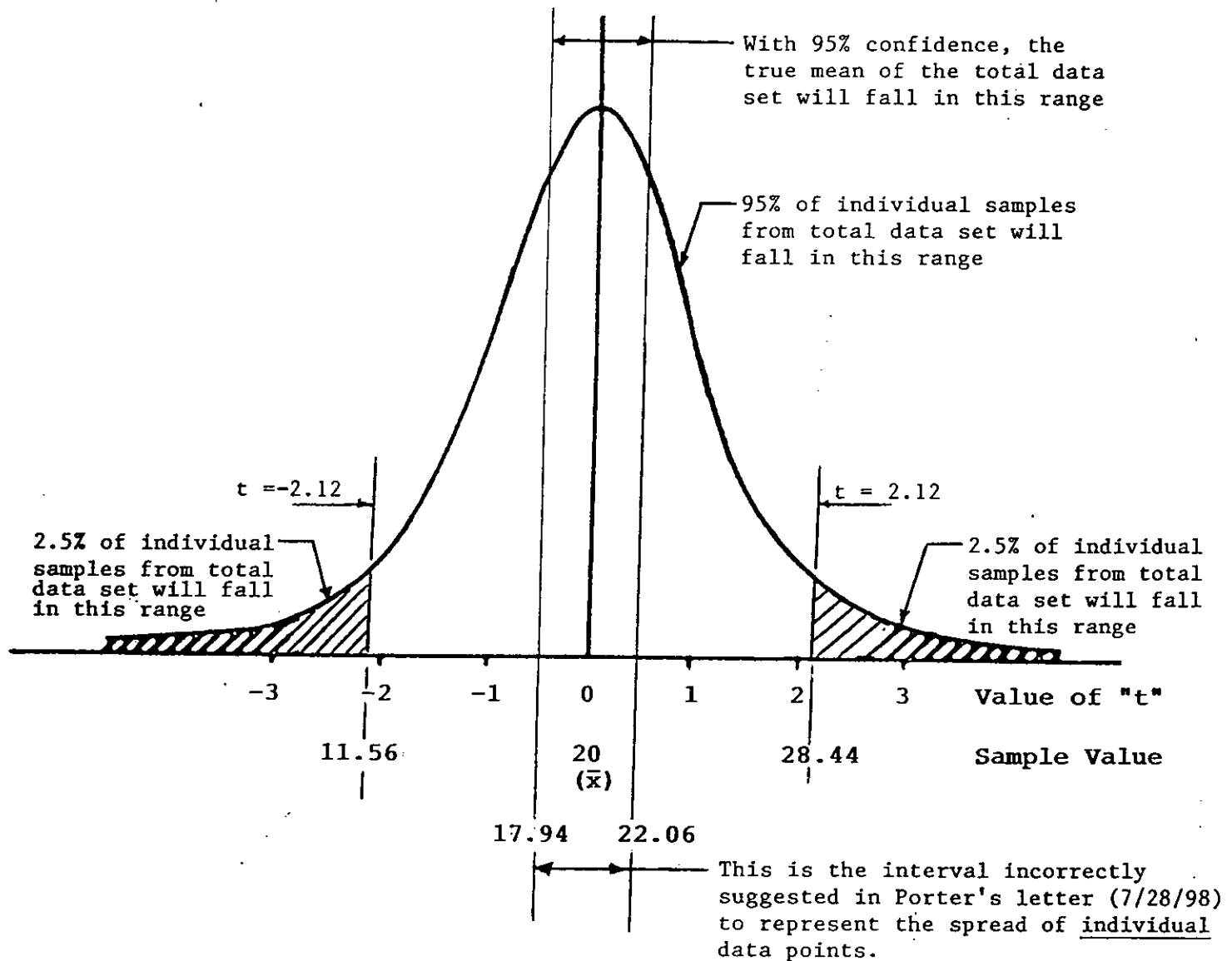
$$\text{Upper C.L.} = 20 + 2.12 (0.97) = 22.06 \text{ lb/hr}$$

In other words, the true mean ( $\mu$ ) of the entire data set (the true average emission rate for all hours the plant operates), with 95% confidence, is between 17.94 and 22.06 lb/hr.

These two estimates are presented graphically in the following figure and are explained in the attached text from a statistical reference.

# CHARACTERISTICS OF HYPOTHETICAL LIMITED DATA SET (17 SAMPLES)

$n = 17$  (16 degrees of freedom)  
 $\bar{x} = 20$  (mean of limited data set)  
 $s = 3.98$  (standard deviation of limited data set)  
 $s_x = 0.97$  (standard error of limited data set)  
 $t = 2.12$  (0.05 probability with 16 d.f.)



FIFTH EDITION

# STATISTICAL METHODS

APPLIED TO EXPERIMENTS IN AGRICULTURE AND BIOLOGY

by **GEORGE W. SNEDECOR**

Former Professor of Statistics  
and Director, Statistical Laboratory  
Iowa State University

*With Chapter 17 on Sampling by*

**WILLIAM G. COCHRAN**

Professor of Statistics  
Harvard University



The Iowa State University Press, *Ames*, Iowa, U.S.A.

must go back to the original fractions, add their numerators and denominators separately, then divide.

EXAMPLE 1.17.1—Three dairy herds of a certain community showed the following reaction to a test for tuberculosis:

Number cows in herd	40	100	10
Percentage reactors	5	2	60

Calculate the average, 6.7%. Do you think this is a better average than  $(5 + 2 + 60)/3 = 22.3\%$ ?

EXAMPLE 1.17.2—The percentages of noxious weed seeds in two samples of timothy are 0.01% and 0.05%. If each sample consisted of 10,000 seeds, what is the average percentage in the two? Ans. 0.03%.

EXAMPLE 1.17.3—If the samples in the foregoing example were 80,000 and 20,000 seeds, respectively, what would be the average? Ans. 0.018%, quite properly nearer the percentage of the larger sample.

EXAMPLE 1.17.4—Schott and Lambert reported that the numbers in table 1.17.1 are averages for 7 years so that the total number of males was 6,972 and of females, 7,126, the sex ratio being 97.84 males per 100 females. Test the hypothesis that the population sex ratio is 100. Ans.  $\chi^2 = 1.68$ . Note: If the averages were used, chi-square would be 0.24, only one-seventh the correct value.

## REFERENCES

1. HERBERT ARKIN and RAYMOND R. COLTON. *Graphs: How To Make and Use Them*. Harper and Brothers, New York. Revised Edition (1940).
2. R. A. FISHER. *Journal of the Royal Statistical Society*, 85:87 (1922).
3. MERLE T. JENKINS and MARTIN A. BELL. *Genetics*, 15:253 (1930).
4. E. W. LINDSTROM. *Cornell University Agricultural Experiment Station Memoir* 13 (1918).
5. JOHN W. MACARTHUR. *Transactions of the Royal Canadian Institute*, 18:1 (1931).
6. ALEXANDER MCFARLANE MOOD. *Introduction to the Theory of Statistics*. McGraw-Hill Book Company, Inc., New York (1950).
7. J. NEYMAN. *First Course in Probability and Statistics*. Henry Holt and Company, New York (1950).
8. KARL PEARSON. *Philosophical Magazine*, Series 5, 50:157 (1899).
9. R. G. SCHOTT and W. V. LAMBERT. *Iowa State Journal of Science*, 4:343 (1930).
10. J. G. SNEDECOR. *Journal of Experimental Zoology*, 110:205 (1949).
11. CATHERINE M. THOMPSON. *Biometrika*, 32:187 (1941).
12. RAY E. WAKELEY. *Iowa State University Research Bulletin* 249 (1938).
13. HELEN M. WALKER and N. DUROST. *Statistical Tables; Their Structure and Use*. Teachers College, Columbia University, New York (1936).
14. S. S. WILKS. *Mathematical Statistics*. Princeton University Press, Princeton, New Jersey (1946).
15. G. U. YULE. *Journal of the Royal Statistical Society*, 85:95 (1922).
16. These data were furnished by courtesy of Dr. T. A. Brindley, leader of the co-operative project.

## Sampling From a Normally Distributed Population

**2.1—Normally distributed population.** In the first chapter, sampling was mostly from a population with only two kinds of individuals; odd or even, alive or dead, infested or free. Random samples of  $n$  from such a population made up a *binomial distribution*. The variable, an enumeration of successes, was discrete. Now we turn to another kind of population whose individuals are measured for some characteristic such as height or yield or income. The variable flows without a break from one individual to the next—a continuous variable with no limit to the number of individuals with different measurements. Such variables are distributed in many ways, but we shall be occupied mainly with the *normal distribution*.

Next to the binomial, the normal distribution was the earliest to be developed. De Moivre published its equation in 1733, 20 years after Bernoulli had given a comprehensive account of the binomial. That the two are not unrelated is clear from figure 2.1.1. On the left is the graph of a symmetrical binomial distribution similar to that in figure 1.6.1. In this new figure the sample size is 48 and the population sampled has equal numbers of the two kinds of individuals. An indefinitely great

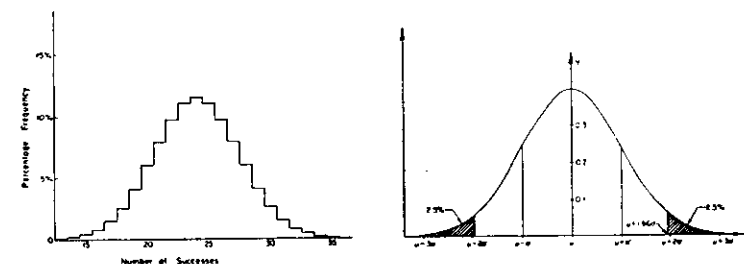


FIG. 2.1.1—At left: Binomial Distribution of successes in samples of 48 from 1:1 population. At right: Normal Distribution with mean  $\mu$  and standard deviation  $\sigma$ . (see table 8.7.1); the shaded areas comprise 5% of the total.



number of samples were drawn so that the frequencies are expressed as percentages of the total. Successes less than 13 and more than 35 do occur, but their frequencies are so small that they cannot be shown on the graph.

Imagine now that the size of the sample is increased without limit, the width of the intervals on the horizontal axis being decreased correspondingly. The steps of the histogram would soon become so small as to look like the continuous curve at the right. Indeed, the normal distribution is related to the binomial in some such manner as that described. The discrete variable has become *continuous* and the frequencies have merged into each other without a break.

This normal distribution is completely determined by two constants or *parameters*. First, there is the *mean*,  $\mu$ , which locates the center of the distribution. Second, the *standard deviation*,  $\sigma$ , measures the spread or variation of the individual measurements; in fact,  $\sigma$  is the *scale* (unit of measurement) of the variable which is normally distributed.

From the figure you see that during one sigma on either side of  $\mu$  the frequency is decreasing ever more rapidly but beyond that point it decreases at a continuously lesser rate. By the time the variable,  $X$ , has reached  $\pm 3\sigma$  the percentage frequencies are negligibly small. Theoretically, the frequency of occurrence never vanishes entirely, but it approaches zero as  $X$  increases indefinitely. The concentration of the measurements close to  $\mu$  is emphasized by the facts that over  $\frac{3}{4}$  of the observations lie in the interval  $\mu \pm \sigma$  while some 95% of them are in the interval  $\mu \pm 2\sigma$ . Beyond  $\pm 3\sigma$  lies only 0.26% of the total frequency.

You are doubtless wondering why such a model is being presented since it obviously cannot describe any real population. It is astonishing that this normal distribution has dominated statistical practice as well as theory. Some of the reasons will be noted at suitable places (sections 3.4, 5.6, and 5.7), but three of them can be indicated here. First, there are many biological variables whose distributions are approximately normal, such as heights of men, for example, or lengths of ears of corn, or dressing percentages of swine. Second, it has been learned from both theory and experience that the inferences we shall make from sampling experimental populations are little affected by ordinary deviations from normality. Third, the mathematical treatment of the equation of the normal distribution is surprisingly easy and has been productive of a large body of theory with practical applications. Further discussion of the normal distribution will be found in chapter 8.

**2.2—Estimators of  $\mu$  and  $\sigma$ .** While  $\mu$  and  $\sigma$  are seldom known, they may be estimated from random samples. To illustrate the estimation of the parameters, we turn to the data reported in table 2.2.1. In 1936 the Council on Foods of the American Medical Association sampled the vitamin C content of commercially canned tomato juice by analyzing a specimen from each of the brands that displayed the seal of the Council (3). The data are shown in the second column of the table.

TABLE 2.2.1  
VITAMIN C CONCENTRATION OF 17 SPECIMENS OF COMMERCIAL CANNED  
TOMATO JUICE, 1936\*

Observation Number	Vitamin C Concentration Mg. Per 100 g.	Deviation From Mean	Deviation Squared
$n$	$X$	$x = X - \bar{x}$	$x^2$
1	16	-4	16
2	22	+2	4
3	21	+1	1
4	20	0	0
5	23	+3	9
6	21	+1	1
7	19	-1	1
8	15	-5	25
9	13	-7	49
10	23	+3	9
11	17	-3	9
12	20	0	0
13	29	+9	81
14	18	-2	4
15	22	+2	4
16	16	-4	16
17	25	+5	25
Totals	340	-26	+26

$$\bar{x} = 340/17 = 20 \text{ mg. per 100 grams}$$

$$s^2 = \Sigma x^2 / (n - 1) = 254/16 = 15.88 \quad s = 3.98 \text{ mg./100 g.}$$

$$s_1^2 = s^2/n = 15.88/17 = 0.934 \quad s_1 = s/\sqrt{17} = 0.965 \text{ mg./100 g.}$$

\* Slightly modified, as is our custom, to make calculation easy. The conclusions are unaltered. For the original data see example 2.12.1.

Assuming random sampling from a normal population,  $\mu$  is estimated by an average called the *mean of the sample* or, more briefly, the *sample mean*. This is calculated by the familiar process of dividing the sum of the observations,  $X$ , by their number. Representing the sample mean by  $\bar{x}$ ,

$$\bar{x} = 340/17 = 20 \text{ mg. per 100 grams of juice}$$

The symbol,  $\bar{x}$ , is often called "bar- $x$ " or " $x$ -bar." We say that this sample mean is an estimator of  $\mu$  or that  $\mu$  is estimated by it.

As for the standard deviation, the simplest estimator of it is based on the *range* of the sample observations, that is, the difference between the largest and smallest measurements. For the vitamin C data,

$$\text{range} = 29 - 13 = 16 \text{ mg./100 g.}$$

From the range, sigma is estimated by means of a fraction which depends on the sample size; see table 2.2.2 (13, 18). For  $n = 17$ , halfway between 16 and 18, the fraction is 0.279, so that

$$\sigma \text{ is estimated by } (0.279)(16) = 4.46 \text{ mg./100 g.}$$

TABLE 2.2.2  
RATIO OF  $\sigma$  TO RANGE IN SAMPLES OF  $n$  FROM THE NORMAL DISTRIBUTION.  
EFFICIENCY OF RANGE AS ESTIMATOR OF  $\sigma$ . NUMBER OF OBSERVATIONS WITH  
RANGE TO EQUAL 100 WITH  $\sigma$

$n$	$\frac{\sigma}{\text{Range}}$	Relative Efficiency	Number per 100	$n$	$\frac{\sigma}{\text{Range}}$	Relative Efficiency	Number per 100
2	0.886	1.000	100	12	0.307	0.815	123
3	.591	0.992	101	14	.294	.783	128
4	.486	.975	103	16	.283	.753	133
5	.430	.955	105	18	.275	.726	138
6	.395	.933	107	20	.268	.700	143
7	.370	.912	110	30	.245	.604	166
8	.351	.890	112	40	.231	.536	186
9	.337	.869	115	50	.222	.49	204
10	.325	.850	118				

Quite easily, then, we have made a *point estimate* of each parameter of a normal population; these estimators constitute a summary of the information contained in the sample. The sample mean cannot be improved upon as an estimator of  $\mu$ , but we shall learn to estimate  $\sigma$  more efficiently. Also we shall learn about interval estimates and tests of hypotheses. Before doing so, it is worth while to examine our sample in greater detail.

The first point to be clarified is this: What population was represented by the sample of 17 determinations of vitamin C? I have raised this question tardily; it is the first one to be considered in designing any sampling. The report makes it clear that not all brands were sampled, only the seventeen that were allowed to display the seal of the Council. The dates of the packs were mostly August and September of 1936, about a year before the analyses were made. The Council report states that the vitamin concentration "... may be expected to vary according to the variety of the fruit, the conditions under which the crop has been grown, the degree of ripeness and other factors." About all that can be said, then, is that the sampled population consisted of those year-old containers still available to the 17 selected packers.

Other details are discussed in the following sections.

**2.3—The array and its graphical representation.** Some of the more intimate features of a sample are shown by arranging the observations in order of size, from low to high, in an *array*. The array of vitamin contents is like this:

13, 15, 16, 16, 17, 18, 19, 20, 20, 21, 21, 22, 22, 23, 23, 25, 29

For a small sample the array serves some of the same purposes as the frequency distribution of a large one.

The range, from 13 to 29, is now obvious. Also, attention is attracted to the concentration of the measures near the center of the array and to

their thinning out at the extremes. In this way the sample may reflect the distribution of the normal population from which it was drawn. But the smaller the sample, the more erratic its reflection may be.

In looking through the vitamin C contents of the several brands, one is struck by their variation. What are the causes of this variation? Different processes of manufacture, perhaps, and different sources of the fruit. Doubtless, also, the specimens examined, being themselves samples of their brands, differed from the brand means. Finally, the laboratory technique of evaluation is never perfectly accurate. Variation is the very essence of statistical data.

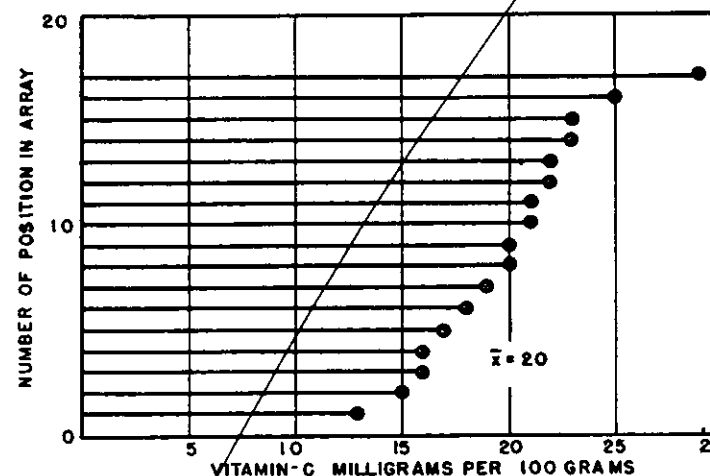


FIG. 2.3.1—Graphical representation of an array. Vitamin C data.

Figure 2.3.1 is a graphical representation of the foregoing array of 17 vitamin determinations. A dot represents each item. The distance of the dot from the vertical line at the left, proportional to the concentration of ascorbic acid in a brand specimen, is read in milligrams per 100 grams on the horizontal scale.

The diagram brings out vividly not only the variation and the concentration in the sample, but also two other characteristics: (i) the rather symmetrical occurrence of the values above and below the mean, and (ii) the scarcity of both extremely small and extremely large vitamin C contents, the bulk of the items being near the middle of the set. These features recur with notable persistence in samples from normal distributions. For many variables associated with living organisms there are averages and ranges peculiar to each, reflecting the manner in which

each seems to express itself most successfully. These norms persist despite the fact that individuals enjoy a considerable freedom in development. A large part of our thinking is built around ideas corresponding to such statistics. The words, pig, daisy, man, each raises an image which is quantitatively described by summary numbers. It is difficult to conceive of any progress in thought until memories of individuals are collected into concepts like averages and ranges of distributions.

**2.4—Symbolical representation.** The items in any set may be represented by

$$X_1, X_2, X_3, \dots, X_n,$$

where the subscripts 1, 2,  $\dots$ ,  $n$ , may specify position in the set of  $n$  items (not necessarily an array). The three dots accompanying these symbols are read "and so on." Matching the symbols with the values in column 2 of table 2.2.1,

$$X_1 = 16, X_2 = 22, \dots, X_{17} = 25 \text{ mg./100 g.}$$

The sample mean is represented by  $\bar{x}$ , so that

$$\bar{x} = (X_1 + X_2 + \dots + X_n)/n$$

This is condensed into the form,

$$\bar{x} = (\Sigma X)/n,$$

where  $X$  stands for every item successively. The symbol,  $\Sigma X$ , is read, "summation  $X$ " or "sum of the  $X$ ." Applying this formula to the values of  $X$  in table 2.2.1,

$$\Sigma X = 340, \text{ and } \bar{x} = 340/17 = 20 \text{ mg./100 g.}$$

**2.5—Deviations from sample mean.** The individual variations of the items in a set of data may be well expressed by the *deviations* of these items from some centrally located number such as the sample mean. For example, the deviation-from-mean of the first  $X$ -value in table 2.2.1 is  $16 - 20 = -4$  mg. per 100 g.; that is, this specimen falls short of  $\bar{x}$  by 4 mg./100 g. Of special interest is the whole set of deviations calculated from the array in section 2.3:

$$-7, -5, -4, -4, -3, -2, -1, 0, 0, 1, 1, 2, 2, 3, 3, 5, 9$$

These deviations are represented graphically in figure 2.3.1 by the distances of the dots from the vertical line drawn through the sample mean.

Deviations are almost as fundamental in our thinking as are averages themselves. "What a whale of a pig" is a metaphor expressing astonishment at the deviation of an individual's size from the speaker's concept of the normal. Gossip and news are concerned chiefly with deviations from accepted standards of behavior. Curiously, interest is wont to center in departures from norm, rather than in that background of averages

against which the departures achieve prominence. Statistically, freaks are freaks only because of their large deviations.

Deviations are represented symbolically by lower case letters. That is:

$$x_1 = X_1 - \bar{x}$$

$$x_2 = X_2 - \bar{x}$$

$$\vdots \quad \vdots \quad \vdots$$

$$x_n = X_n - \bar{x}$$

Just as  $X$  may represent any of the items in a set, or all of them in succession, so  $x$  represents deviations from sample mean. In general,

$$x = X - \bar{x}$$

It is easy to verify the theorem that the sum of a set of deviations-from-mean is zero; that is  $\Sigma x = 0$ . The set listed in table 2.2.1 adds to zero, the sum of the positive deviations being equal to the sum of the negatives. This theorem about deviations-from-mean is useful for verifying the calculation of a set of deviations—be sure that the sum is zero. As a consequence of the theorem, it follows that the *mean* of the deviations is zero. This is a theorem about which you will be reminded later.

**EXAMPLE 2.5.1**—The weights of 12 staminate hemp plants in early April at College Station, Texas, (16), were approximately:

13, 11, 16, 5, 3, 18, 9, 9, 8, 6, 27, and 7 grams

Array the weights and represent them graphically. Calculate the sample mean, 11 grams, and the deviations therefrom. Verify the fact that  $\Sigma x = 0$ . Show that  $\sigma$  is estimated by 7.4 grams.

**EXAMPLE 2.5.2**—The heights of 11 men are 64, 70, 65, 69, 68, 67, 68, 67, 66, 72, and 61 inches. Compute the sample mean and verify it by summing the deviations. Are the numbers of positive and negative deviations equal, or only their sums?

**EXAMPLE 2.5.3**—The yields of alfalfa from 10 plots were 0.8, 1.3, 1.5, 1.7, 1.7, 1.8, 2.0, 2.0, 2.0, and 2.2 tons per acre. How many deviations are positive and how many negative? Is their sum zero? Estimate  $\sigma$ . Ans. 0.46 ton per acre.

**EXAMPLE 2.5.4**—The weights of 11 forty-year-old men were 148, 154, 158, 160, 161, 162, 166, 170, 182, 195, and 236 pounds. Contrast the graphical representation of this array with that of the preceding example. Notice the fact that only three of the weights exceed the sample mean. Would you expect weights of men to be normally distributed? A test of symmetry will be given in section 8.5.

**EXAMPLE 2.5.5**—The following were the yields of two varieties of oats in five successive years (bushels per acre):

Variety	Year				
	1	2	3	4	5
A	34	30	41	25	45
B	30	17	33	25	25

Calculate the 5 differences,  $A - B$ . Might these differences be a sample from a normal population of differences? Assuming that they are, estimate  $\mu$  and  $\sigma$ . Ans. 9.0 and 8.6 bushels per acre.

**EXAMPLE 2.5.6**—The following data are adapted from Reddy's (14) investigations of the differences in yield attributable to the disinfection of *Diplodia* infected seeds of maize. The figures represent bushels per acre.

Treatment	Pairs of Plots in 1933					
	1	2	3	4	5	6
Treated	68.1	74.6	64.4	69.2	61.8	57.9
Untreated	64.7	62.5	66.8	69.2	53.9	58.5
Differences	3.4	12.1	-2.4	0.0	7.9	-0.6

Pairs of Plots in 1934										
1	2	3	4	5	6	7	8	9	10	
18.0	24.0	18.8	17.8	18.5	27.2	23.6	23.9	20.3	11.9	
10.9	24.4	15.1	16.8	13.2	21.6	13.7	17.5	16.3	15.5	
7.1	-0.4	3.7	1.0	5.3	5.6	9.9	6.4	4.0	-3.6	

Clearly the yields in the two years are not samples from the same population, but the differences may be. Represent graphically the array of differences. Estimate  $\mu$  and  $\sigma$  in the population of differences. Ans. 3.7 and 4.4 bushels per acre.

**EXAMPLE 2.5.7**—If you sum the deviations from 3.7 bushels per acre in the foregoing example you will not get zero. Why? If you compute the sample mean of the deviations and add it to 3.7, will you get the exact sample mean of the differences?

**EXAMPLE 2.5.8**—If you should calculate the sample mean of the yields of the 16 untreated plots in example 2.5.6, would it estimate the parameter of any population that you can describe?

**EXAMPLE 2.5.9**—Suppose you wish to estimate the yield of a field of 300 rows of corn. You actually harvest 10 rows, chosen at random, and determine the sample mean yield, 5 bushels per row. Would you hesitate to fix the field yield at 1,500 bushels? You would be assuming that the field mean is the same as that of the 10 harvested rows, and would be using the theorem,  $\Sigma X = n\bar{x}$ .

**EXAMPLE 2.5.10**—If you have some skill in algebra, prove the theorem that  $\Sigma x = 0$ . Starting with the relation,  $x = X - \bar{x}$ , sum both members, then substitute  $\Sigma X = n\bar{x}$ .

**EXAMPLE 2.5.11**—If you have two sets of data which are paired as in example 2.5.5, and if you have calculated the resulting set of differences, prove that the sample mean of the differences is equal to the difference between the sample means of the two sets. Verify this theorem by use of the data in example 2.5.5.

**2.6—Another estimator of  $\sigma$ ; the sample standard deviation.** The range, dependent as it is on only the two extremes in a sample, has a more variable sampling distribution than an estimator based on the whole set of deviations-from-mean in a sample, not just the largest and smallest. Such a set, with 17 deviations, was shown in table 2.2.1 and again as an array in section 2.5. What kind of average is appropriate to summarize these deviations, and to estimate  $\sigma$  with the least sampling variation?

Clearly, the sample mean of the deviations is useless as an estimator because it is always zero. But a natural suggestion is to ignore the signs,

calculating the sample mean of the absolute values of the deviations. The resulting measure of variation, the *mean absolute deviation*, had a considerable vogue in times past. Now, however, we have other estimators, more efficient and more easily calculated.

One of the more efficient estimators is the *sample standard deviation* which we shall denote by  $s$ . Its calculation is set out in the right-hand part of table 2.2.1. First, each deviation is squared. Next, the *sum of squares*,  $\Sigma x^2$ , is divided by the number of *degrees of freedom*, one less than the sample size. The result is the *mean square*,  $s^2$ . Finally, the extraction of the square root recovers the original unit of measurement (in this example, mg. per 100 g.). Before further discussion of this average, its calculation should be fixed in mind by the working of a few examples.

**EXAMPLE 2.6.1**—The five differences,  $A - B$ , in example 2.5.5 were 4, 13, 8, 0, and 20 bushels per acre. Calculate  $s$ . Ans. 7.8 bu./acre. Compare this with the estimate based on the range.

**EXAMPLE 2.6.2**—In example 2.5.1, calculate the sample standard deviation. Ans. 6.7 grams. Compare this with your first estimate of  $\sigma$ .

**EXAMPLE 2.6.3**—Calculate  $s$  for the alfalfa yields of example 2.5.3. Ans. 0.41 ton per acre.

It may be a little surprising to have the divisor,  $n - 1$ , proposed for computing an average; you have always calculated the sample mean by using the divisor,  $n$ . In computing  $s^2$ , it is necessary to divide by the degrees of freedom if you wish to avoid a bias in estimating  $\sigma^2$ . Division by  $n$  does produce an estimate of  $\sigma^2$  but it is a *biased estimate*. In the problems we shall consider there is no occasion to use any but the *unbiased estimate*.

You now have two estimators of  $\sigma$ , one of them easier to calculate than the other, but *less efficient*. You need to know what is meant by "less efficient" and what governs the choice of which to use. Both pieces of information are supplied by the fourth columns of table 2.2.2. As an example, if  $n = 10$ , the estimate from the range is only 85% as efficient as that from  $s$ ; meaning that, for the same precision, a sample of 10 with  $s$  as estimator is equivalent to a sample of  $10/0.85 = 12$  using range. The argument will become clearer as you proceed, but the table indicates right away that, other things being equal, you have to weigh the cost of calculating  $s$  against the cost of more observations; in this instance, 12 instead of 10. Now there are some operations where observations are taken for other purposes and are then available to the statistician at no extra cost. For estimating  $\sigma$  he could have a few extra just by copying them. His cue would be to use the range. But consider the alfalfa experiment of example 2.6.3. How much would it cost the investigator to provide 2 extra plots? There is the cost of land and equipment to be considered, together with their availability for use, there are salaries and wages, and finally there is the sale price of the alfalfa. The net cost of the 2 extra plots is to be balanced against the few cents or the few minutes it would take to calculate  $s$ . I suggest that, if you value the information to be obtained

from the experiment, you should proceed with the calculation of  $s$ . This will give you the maximum information to be obtained from the data. The advantage of the range is that it provides a quick preliminary estimate. Also, since computers sometimes make mistakes, the estimate from the range is an easy approximate check on the calculation of  $s$ . For this purpose, it is well to fix in mind a few of the fractions,  $\sigma/\text{range}$ . Remember:

If $n$ Is Near This Number	Then $\sigma$ Is Roughly Estimated From Dividing Range by
5	2
10	3
25	4
100	5

In statistical practice, this rough-and-ready estimator of  $\sigma$  proves itself a most useful device. But for the usual run of experimental data it pays to calculate  $s$ .

If you study mathematical statistics, you may hear a good deal about the Principle of Least Squares. The sample mean and deviations therefrom are related to that principle in this manner: if deviations are measured from the sample mean, then the sum of their squares is a minimum. In particular (and reversing the statement) if deviations in table 2.2.1 are measured from some number different from the sample mean, 20, the sum of their squares will be greater than 254. Verify this by trying deviations from, say, 19 then 22.

It seems rather characteristic that large things vary much and small things little. For this reason it is often convenient to express the sample standard deviation as a fraction of the sample mean, the resulting statistic being called *relative standard deviation* or *coefficient of variation*,  $C$ . As an example, it is reported (1) that the average stature of one-year and eighteen-year girls are 74.4 and 161.0 cm. respectively, with sample standard deviations 2.64 and 6.12 cm. The two coefficients of variation are

$$C_1 = 2.64/74.4 = 0.036$$

$$C_{18} = 6.12/161.0 = 0.038,$$

almost the same. Usually  $C$  is expressed as a percentage,  $C_1$  for example, being 3.6%. Discussion of this characteristic of the sample standard deviation is resumed in section 2.16.

**EXAMPLE 2.6.4**—The birth weights of 20 guinea pigs, borne in litters of two, were: 30, 30, 26, 32, 30, 23, 29, 31, 36, 30, 25, 34, 32, 24, 28, 27, 38, 31, 34, 30 grams. Estimate  $\sigma$  in 3 ways: (i) by the rough approximation, one-fourth of the range (Ans. 3.8 g.); (ii) by use of the fraction, 0.268, in table 2.2.2 (Ans. 4.0 g.); (iii) by calculating  $s$  (Ans. 3.85 g.). N.B.: Observe the time required to calculate  $s$ .

**EXAMPLE 2.6.5**—In the preceding example, how many birth weights would be required to yield the same precision if the range were used instead of  $s$ ? Ans. 29 weights.

**EXAMPLE 2.6.6**—If it takes 5 minutes to weigh a guinea pig (removing and returning to cage, weighing and recording) and 2 minutes to estimate  $\sigma$  using the range, would you have saved or lost time by calculating  $s$ ?

**EXAMPLE 2.6.7**—Suppose you lined up according to height the 16 men in 2 squads of 18-year-old freshmen, then measured the height of the shortest, 64 inches, and of the tallest, 72 inches. Would you accept the midpoint of the range,  $(64 + 72)/2 = 68$  inches as a rough estimate of  $\mu$ , and  $8/3 = 2.7$  inches as a quick-and-easy estimate of  $\sigma$ ?

**EXAMPLE 2.6.8**—The mean yield of hay from 15 plots of alfalfa was 2.2 tons per acre, with  $s = 0.35$  ton per acre. Using table 2.2.2, approximate the range on the assumption that  $\sigma = 0.35$  ton per acre. Ans. 1.2 tons per acre. Would this suggest that the highest yielding plot bore about 2.8 tons per acre?

**2.7—"Student's"  $t$ -distribution.** We now have adequate point estimators for  $\mu$  and  $\sigma$ . Next to be considered are *interval estimates* and *tests of hypotheses*.

First we require a sampling distribution analogous to that of chi-square. Known as "*Student's*"  $t$ -distribution, it was discovered by W. S. Gosset in 1908 (15) and perfected by R. A. Fisher in 1924 (6). This distribution has revolutionized the statistics of small samples. In the next chapter you will be asked to verify the distribution by the same kind of sampling process as you used for chi-square; indeed it was by such sampling that Gosset first learned about it.

The quantity  $t$  is given by the equation,

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

That is,  $t$  is the deviation of the estimated mean from that of the population, measured in terms of  $s/\sqrt{n}$  as the unit. Both  $\bar{x}$  and  $s$  are calculated from a sample of  $n$  observations, assumed to be a random sample from a normal population. We do not know  $\mu$  though we may have some hypothesis about it. Without  $\mu$ ,  $t$  cannot be calculated; but its sampling distribution has been worked out.

The denominator,  $s/\sqrt{n}$ , is a useful quantity estimating  $\sigma/\sqrt{n}$ , the *standard error*. We shall call  $s/\sqrt{n}$  the *sample standard error* and symbolize it by  $s_x$ . Further explanation will be given in chapter 3. For vitamin C, table 2.2.1,  $s_x = 3.98/\sqrt{17} = 0.965$  mg./100 g.

The distribution of  $t$  is laid out in table 2.7.1. In large samples it is practically normal with  $\mu = 0$  and  $\sigma = 1$ . It is only for samples of less than 30 that the distinction becomes obvious.

Like the normal, the  $t$ -distribution is symmetrical about the mean. This allows the probability in the table to be stated as that of a larger absolute value, sign ignored. As an example, look at the value,  $t = 1.96$ , for infinite ( $\infty$ ) degrees of freedom, the normal distribution. The probability indicated is 0.05. This means that among samples of great size, drawn at random from a normal population, 5% of them are expected to have either  $t > 1.96$  or  $t < -1.96$ . Figure 2.7.1 shows such values of  $t$

TABLE 2.7.1  
THE DISTRIBUTION OF  $t^*$ 

Degrees of Freedom	Probability of a Larger Value, Sign Ignored								
	0.500	0.400	0.200	0.100	0.050	0.025	0.010	0.005	0.001
1	1.000	1.376	3.078	6.314	12.706	25.452	63.657		
2	.816	1.061	1.886	2.920	4.303	6.205	9.925	14.089	31.598
3	.765	.978	1.638	2.353	3.182	4.176	5.841	7.453	12.941
4	.741	.941	1.533	2.132	2.776	3.495	4.604	5.598	8.610
5	.727	.920	1.476	2.015	2.571	3.163	4.032	4.773	6.859
6	.718	.906	1.440	1.943	2.447	2.969	3.707	4.317	5.959
7	.711	.896	1.415	1.895	2.365	2.841	3.499	4.029	5.405
8	.706	.889	1.397	1.860	2.306	2.752	3.355	3.832	5.041
9	.703	.883	1.383	1.833	2.262	2.685	3.250	3.690	4.781
10	.700	.879	1.372	1.812	2.228	2.634	3.169	3.581	4.587
11	.697	.876	1.363	1.796	2.201	2.593	3.106	3.497	4.437
12	.695	.873	1.356	1.782	2.179	2.560	3.055	3.428	4.318
13	.694	.870	1.350	1.771	2.160	2.533	3.012	3.372	4.221
14	.692	.868	1.345	1.761	2.145	2.510	2.977	3.326	4.140
15	.691	.866	1.341	1.753	2.131	2.490	2.947	3.286	4.073
16	.690	.865	1.337	1.746	2.120	2.473	2.921	3.252	4.015
17	.689	.863	1.333	1.740	2.110	2.458	2.898	3.222	3.965
18	.688	.862	1.330	1.734	2.101	2.445	2.878	3.197	3.922
19	.688	.861	1.328	1.729	2.093	2.433	2.861	3.174	3.883
20	.687	.860	1.325	1.725	2.086	2.423	2.845	3.153	3.850
21	.686	.859	1.323	1.721	2.080	2.414	2.831	3.135	3.819
22	.686	.858	1.321	1.717	2.074	2.406	2.819	3.119	3.792
23	.685	.858	1.319	1.714	2.069	2.398	2.807	3.104	3.767
24	.685	.857	1.318	1.711	2.064	2.391	2.797	3.090	3.745
25	.684	.856	1.316	1.708	2.060	2.385	2.787	3.078	3.725
26	.684	.856	1.315	1.706	2.056	2.379	2.779	3.067	3.707
27	.684	.855	1.314	1.703	2.052	2.373	2.771	3.056	3.690
28	.683	.855	1.313	1.701	2.048	2.368	2.763	3.047	3.674
29	.683	.854	1.311	1.699	2.045	2.364	2.756	3.038	3.659
30	.683	.854	1.310	1.697	2.042	2.360	2.750	3.030	3.646
35	.682	.852	1.306	1.690	2.030	2.342	2.724	2.996	3.591
40	.681	.851	1.303	1.684	2.021	2.329	2.704	2.971	3.551
45	.680	.850	1.301	1.680	2.014	2.319	2.690	2.952	3.520
50	.680	.849	1.299	1.676	2.008	2.310	2.678	2.937	3.496
55	.679	.849	1.297	1.673	2.004	2.304	2.669	2.925	3.476
60	.679	.848	1.296	1.671	2.000	2.299	2.660	2.915	3.460
70	.678	.847	1.294	1.667	1.994	2.290	2.648	2.899	3.435
80	.678	.847	1.293	1.665	1.989	2.284	2.638	2.887	3.416
90	.678	.846	1.291	1.662	1.986	2.279	2.631	2.878	3.402
100	.677	.846	1.290	1.661	1.982	2.276	2.625	2.871	3.390
120	.677	.845	1.289	1.658	1.980	2.270	2.617	2.860	3.373
$\infty$	.6745	.8416	1.2816	1.6448	1.9600	2.2414	2.5758	2.8070	3.2905

\* Parts of this table are reprinted by permission from R. A. Fisher's *Statistical Methods for Research Workers*, published by Oliver and Boyd, Edinburgh (1925-1950); from Maxine Merrington's "Table of Percentage Points of the  $t$ -Distribution," *Biometrika*, 32:300 (1942); and from Bernard Osile's *Statistics in Research*, Iowa State College Press (1954).

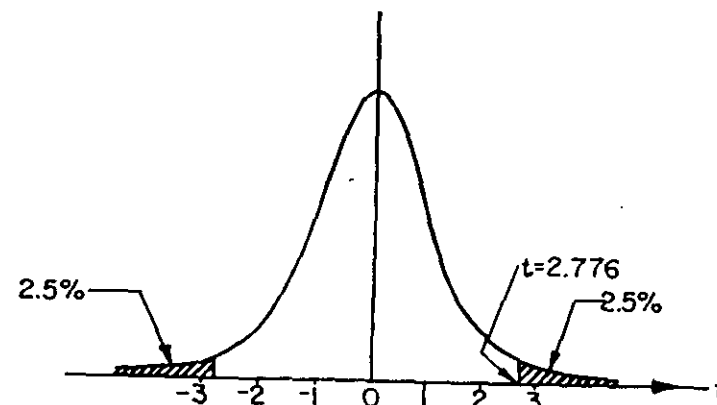


FIG. 2.7.1—Distribution of  $t$  with 4 degrees of freedom. The shaded areas comprise 5% of the total area. The distribution is more peaked in the center and has higher tails than the normal.

in the shaded areas; 2.5% of them are in one tail and 2.5% in the other. Effectively, the table shows the two halves of the figure superimposed, giving the sum of the shaded areas (probabilities) in both. So, the probability in the two tails of the  $t$ -distribution corresponds to that in one tail of chi-square, table 1.14.1. The reason for making the table this way will shortly appear.

EXAMPLE 2.7.1—In the vitamin C sampling of table 2.2.1, set up the hypothesis that  $\mu = 17.954$  mg./100 g. Calculate  $t$ . Ans. 2.12.

EXAMPLE 2.7.2—For the vitamin C sample, degrees of freedom =  $17 - 1 = 16$ , the denominator of the fraction giving  $t$ . From table 2.7.1, find the probability of a value of  $t$  larger in absolute value than 2.12. Ans. 0.05. This means that, among random samples of  $n = 17$  from normal populations, 5% of them are expected to have  $t$ -values below  $-2.12$  or above  $2.12$ .

EXAMPLE 2.7.3—If samples of  $n = 17$  are randomly drawn from a normal population and have  $t$  calculated for each, what is the probability that  $t$  will fall between  $-2.12$  and  $+2.12$ ? Ans. 0.95.

EXAMPLE 2.7.4—If random samples of  $n = 17$  are drawn from a normal population, what is the probability of  $t$  greater than 2.12? Ans. 0.025.

EXAMPLE 2.7.5—What size of sample would have  $|t| > |2|$  in 5% of all random samples from normal populations? Ans. 61. (Note the symbol for "absolute value," that is, ignoring signs.)

EXAMPLE 2.7.6—Among very large samples ( $d.f. = \infty$ ), what value of  $t$  would be exceeded in 2.5% of them? Ans. 1.96.

2.8—The interval estimate of  $\mu$ ; the confidence interval. The argument about the interval estimate for  $\mu$  is a bit complicated so I am going to tell you how to set the interval before explaining why. As illustration, recall the vitamin C determinations in table 2.2.1;  $n = 17$ ,  $\bar{x} = 20$  and

$s_s = 0.965$  mg./100 g. To get the 95% confidence interval (interval estimate):

1. Enter the table with  $df = 17 - 1 = 16$  and in the column headed .05 take the entry,  $t_{.05} = 2.12$ .

2. Calculate the quantity,

$$t_{.05}s_s = (2.12)(0.965) = 2.05 \text{ mg./100 g.}$$

3. The confidence interval is from

$$20 - 2.05 = 17.95 \text{ to } 20 + 2.05 = 22.05 \text{ mg./100 g.}$$

If you say that  $\mu$  is covered by the interval from 17.95 to 22.05 mg./100 g., you will be right unless a 1-in-20 chance has occurred in the sampling.

The point and 95% interval estimate of  $\mu$  may be summarized this way:  $20 \pm 2.05$  mg./100 g. (The formula is  $\bar{x} \pm t_{.05}s_s$ .)

The explanation of these rules rests on the selection of a particular value for  $t$  in the table. If for  $df = n - 1$  the value  $t_{.05}$  is chosen, it may be said that this value of  $t$  is expected to be exceeded in absolute value in 5% of all samples drawn at random from normal populations. Or the statement may be changed to this:  $t$  will lie between  $-t_{.05}$  and  $+t_{.05}$  in 95% of such samples. That is, the probability is 0.95 that  $-t_{.05} \leq t \leq t_{.05}$ . Substituting the expression for  $t$ , the probability is 0.95 that

$$-t_{.05} \leq \frac{\bar{x} - \mu}{s_s} \leq t_{.05}$$

Multiplying both sides of each inequality by  $s_s$ , the probability is 0.95 that

$$-t_{.05}s_s \leq \bar{x} - \mu \leq t_{.05}s_s$$

Transposing  $\bar{x}$ , changing signs, and reversing the terms, the probability is 0.95 that

$$\bar{x} - t_{.05}s_s \leq \mu \leq \bar{x} + t_{.05}s_s$$

The interpretation is: Before the sample is drawn, the probability is 0.95 that the interval indicated will include  $\mu$ . After the sample is drawn and the values of  $\bar{x}$ ,  $s$ , and  $n$  are substituted, it may be said with confidence that the interval includes  $\mu$  and the statement will be correct unless a 1-in-20 chance has occurred in the sampling.

An interval estimate for  $\sigma$  will be presented in section 2.14.

EXAMPLE 2.8.1—For the yields of alfalfa in examples 2.5.3 and 2.6.3,  $n = 10$ ,  $\bar{x} = 1.70$  and  $s = 0.41$  ton/acre. Set 95% confidence limits on the mean of the population from which this is a random sample. Ans. 1.41 and 1.99 tons/acre.

EXAMPLE 2.8.2—In examples 2.5.5 and 2.6.1, the 5 differences had  $\bar{d} = 9.0$  and  $s_D = 7.8$  bu./acre. Set the 99% interval estimate on  $\mu$ . Ans. From  $-7.0$  to  $25.0$  bu./acre. Might the population difference be zero?

EXAMPLE 2.8.3—In an investigation of growth in school children of 8 private schools (7), the sample mean height of 265 boys of age 13.5 to 14.5 years was 63.84 inches with standard deviation, 3.08 inches. What is the 95% confidence interval for  $\mu$ ? Ans. 63.5 to 64.2 inches. Calculate  $C = 4.8\%$

2.9—Estimates and tests of differences. Experiments are most often designed to discover and evaluate differences between effects rather than the effects themselves. It is differences between yields produced by fertilizers or differences between gains produced by feeds that are wanted. One of the simplest of such experiments is designed to contrast the effects of two treatments. Pairs of similar individuals are selected, one of the treatments being applied to each. The individuals in the pairs may be field plots or pigs or colonies of bees. If there were only a single pair it would be impossible to say whether the difference in behavior is to be attributed to the two treatments imposed or to the natural variability of the individuals or partly to each. Hence, there must be two or more pairs, or replications, one member of every pair being chosen at random to receive the first treatment, the other member the second. The differences between the measurements of the two pairs constitute the sample data upon which inferences are to be based. If there were no individual variation the differences would presumably be all alike. Always there is variation.

In many experiments it may be assumed that the differences make up a random sample from a normal population. Commonly the objective is to learn the size of the mean of this population and particularly if it is different from zero. Let us examine such an experiment.

Youden and Beale (19) wished to find out if two preparations of the mosaic virus would produce different effects on tobacco leaves. The method employed was to rub half a leaf of a tobacco plant with a piece of cheesecloth soaked in one preparation of the virus extract, then to rub the other half similarly with the second preparation. The measurement of potency was the number of local lesions appearing on the half leaf; the measurement is assumed to be a continuous variable. The data reported in table 2.9.1 are taken from leaf number 2 on each of 8 plants. The differences,  $D = X_1 - X_2$  in the fourth column make up the random sample in which the experimenter is interested. The question posed by the experiment is this: Does the preparation of the virus affect the number of lesions? In statistical terms, is the mean of the sampled population equal to zero or is it different from zero?

One way to answer the question is to set confidence limits on the population mean difference,  $\mu_D$ . With  $\bar{d} = 4$  lesions,  $s_d = 1.52$  lesions and  $t_{.05} = 2.365$ , it may be said that, unless a 1-in-20 chance has occurred,

$$4 - (2.365)(1.52) \leq \mu_D \leq 4 + (2.365)(1.52)$$

That is,  $\mu_D$  is expected to be greater than  $4 - (2.365)(1.52) = 0.4$  lesions and less than 7.6. Since zero is not included in the confidence interval,

# Farmland Hydro, L.P.

Charles W. Jenkins  
Manager of Environmental and Safety Services

Green Bay Plant  
County Road 640  
Post Office Box 960  
Bartow, Florida 33831  
Tele: 941 533-1141  
Fax: 941 533-8793

July 29, 1998

**RECEIVED**

JUL 30 1998

**BUREAU OF  
AIR REGULATION**

Bureau of Air Regulation  
Florida Department of Environmental Protection  
2600 Blair Stone Road, Mail Station #5505  
Tallahassee, FL 32399-2400

**RE: AFFIDAVIT OF PUBLICATION ON  
AIR PERMIT NO. 1050053-020-AC (PSD-FL-246)  
CONSTRUCTION PERMIT FOR NORTH MAP/DAP PLANT AND SHIPPING**

Enclosed please find the signed and notarized Affidavit of Publication for the above referenced Air Construction Permit. Please note the date of publication was Monday, July 27, 1998.

Should you have any questions or concerns regarding this matter, call me at (941) 533-1141, extension 334.

Sincerely,



Charles W. Jenkins  
Manager of Environmental and Safety Services

CWJ:jp\129-98  
enc.

cc: Merle Farris, V.P. - Operations  
Farmland Hydro, L.P.

cc: S. Arief, BAR  
SWP  
pelk co.



A Delaware Limited Partnership





# AFFIDAVIT OF PUBLICATION

## THE LEDGER

### Lakeland, Polk County, Florida

Case No .....

STATE OF FLORIDA)  
COUNTY OF POLK )

Before the undersigned authority personally appeared Nelson Kirkland, who on oath says that he is Classified Advertising Manager of The Ledger, a daily newspaper published at Lakeland in Polk County, Florida; that the attached copy of advertisement, being a

#### Public Notice Of Intent

in the matter of .....

DEP File No. 1050053-020-AC (PSD-Fl-246)

in the .....

Court, was published in said newspaper in the issues of .....

July 27;

1998

Affiant further says that said The Ledger is a newspaper published at Lakeland, in said Polk County, Florida, and that the said newspaper has heretofore been continuously published in said Polk County, Florida, daily, and has been entered as second class matter at the post office in Lakeland, in said Polk County, Florida, for a period of one year next preceding the first publication of the attached copy of advertisement; and affiant further says that he has neither paid nor promised any person, firm or corporation any discount, rebate, commission or refund for the purpose of securing this advertisement for publication in the said newspaper.

Signed   
Nelson Kirkland  
Classified Advertising Manager

By Nelson Kirkland who is  
personally known to me

Sworn to and subscribed before me this 29TH

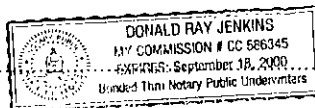
day of July A.D. 1998

(Seal)

Notary Public

My Commission Expires .....

Order#706854  
Farmland Hydro



#### Attach Notice Here

#### PUBLIC NOTICE OF INTENT TO ISSUE AIR CONSTRUCTION PERMIT

STATE OF FLORIDA  
DEPARTMENT OF ENVIRONMENTAL PROTECTION  
DEP File No. 1050053-020-AC (PSD-Fl-246)  
North Monocammium/Diammonium Phosphate (MAP/DAP) Plant  
Farmland Hydro, L.P., Green Bay Facility  
Polk County

The Department of Environmental Protection (Department) gives notice of its intent to issue an air construction permit to Farmland Hydro, L.P. to increase the production rates as well as storage and shipping rates of the North monochlorophosphate (MAP) and diammonium phosphate (DAP) plant at its Green Bay facility. The plant is located at 4393 County Road 640 West, Bartow, Polk County, A First Available Control Technology (BACT) determination was required for fluorides and particulate matter, pursuant to Rule 62.212-400, F.A.C. and 40 CFR 52.71, Prevention of Significant Deterioration (PSD). The applicant's name and address are: Farmland Hydro, L.P., P.O. Box 980, Bartow, Florida 33831.

The MAP production rate will be increased from 120 to 200 tons per hour and the DAP production rate will be increased from 100 to 150 tons per hour. The shipping and storage process rate will be increased to 120 tons of P.D. per hour. Controls for fluoride emissions consist of scrubbers using process pond water. Particulate emissions are also controlled by scrubbers.

An air quality impact analysis was conducted. The project is predicted to have no significant impact in the PSD Class II area in the vicinity of the facility or on the Chassawhatchee National Wilderness Area PSD Class I area located approximately 100 kilometers northwest of the plant.

The Department will issue the final permit with the attached conditions unless a response received in accordance with the following procedures results in a different decision or significant change of terms or conditions.

The Department will accept written comments concerning the proposed permit issuance action for a period of 30 (thirty) days from the date of publication of "Public Notice of Intent to Issue Air Construction Permit." Written comments should be provided to the Department's Bureau of Air Regulation at 2600 Blair Stone Road, Mail Station #5505, Tallahassee, FL 32399-2400. Any written comments filed shall be made available for public inspection. If written comments received result in a significant change in the proposed agency action, the Department shall revise the proposed permit and require, if applicable, another Public Notice.

The Department will issue the permit with the attached conditions unless a timely petition for an administrative hearing is filed pursuant to sections 120.569 and 120.57 of the Florida Statutes. The procedures for petitioning for a hearing are set forth below.

A person whose substantial interests are affected by the proposed permitting decision may petition for an administrative proceeding (hearing) under sections 120.569 and 120.57 of the Florida Statutes. The petition must contain the information set forth below and must be filed (received) in the Office of General Counsel of the Department at 3900 Commonwealth Boulevard, Tallahassee, Florida 32399-3000. Petitions filed by the permit applicant or any of the parties listed below must be filed within fourteen days of receipt of this notice of intent. Petitions filed by any persons other than those entitled to written notice under section 120.60(3) of the Florida Statutes must be filed within fourteen days of publication of the public notice or within fourteen days of receipt of this notice of intent, whichever occurs first. Under section 120.60(3), however, any person who asked the Department for notice of agency action may file a petition within fourteen days of receipt of that notice regardless of the date of publication. A petitioner shall mail a copy of the petition to the applicant at the address indicated above at the time of filing. The failure of any person to file a petition within the appropriate time period shall constitute a waiver of that person's right to request an administrative determination (hearing) under sections 120.569 and 120.57 F.S., or to intervene in this proceeding and participate as a party to it. Any subsequent intervention will be only at the approval of the presiding officer upon the filing of a motion in compliance with rule 28-106.205 of the Florida Administrative Code.

A petition that disputes the material facts on which the Department's action is based must contain the following information: (a) The name and address of each agency affected and each agency's file or identification number, if known; (b) The name, address, and telephone number of the petitioner, the name, address, and telephone number of the petitioner's representative, if any, which shall be the address for service purposes during the course of the proceeding; and an explanation of how the petitioner's substantial interests will be affected by the agency determination; (c) A statement of how and when petitioner received notice of the agency action or proposed action; (d) A statement of all disputed issues of material fact. If there are none, the petition must so indicate; (e) A concise statement of the ultimate facts alleged, as well as the rules and statutes which entitle the petitioner to relief; and (f) A demand for relief.

A petition that does not dispute the material facts upon which the Department's action is based shall state that no such facts are in dispute and otherwise shall contain the same information as set forth above as required by rule 28-106.301.

Because the administrative hearing process is designed to formulate final agency action, the filing of a petition means that the Department's final action may be different from the position taken by it in this notice. Persons whose substantial interests will be affected by any such final decision of the Department on the application have the right to petition to become a party to the proceeding in accordance with the requirements set forth above.

Mediation is not available in this proceeding.

A complete project file is available for public inspection during normal business hours: 8:00 a.m. to 5:00 p.m., Monday through Friday, except legal holidays, at:

Polk County Public Works Department - Air Division 4189 Ben Durrance Road Bartow, Florida 33830 Telephone: 941/534-7377 Fax: 941/534-7374	Dept. of Environmental Protection Bureau of Air Regulation 111 S. Magnolia Drive, Suite 4 Tallahassee, Florida 32301 Telephone: 850/488-0114 Fax: 850/922-6979	Dept. of Environmental Protection Southwest District 3804 Coconut Palm Drive Tampa, Florida 33619-8218 Telephone: 813/744-0100 Fax: 813/744-6084
--	---	---

The complete project file includes the application, technical evaluations, Draft Permit, and the information submitted by the responsible official, exclusive of confidential records under Section 403.111, F.S. Interested persons may contact the Administrator, New Resource Review Section at 111 South Magnolia Drive, Suite 4, Tallahassee, Florida 32301, or call 850/488-0114 for additional information 8-742-7-27, 1998.

B742



U.S. FISH & WILDLIFE SERVICE  
AIR QUALITY BRANCH  
P.O. BOX 25287, Denver, CO 80225-0287

---

FACSIMILE COVER SHEET

---

**Date:** July 28, 1998

**Telephone:** (303) 969-2617

**Fax:** (303) 969-2822

**To:** Al Linero

**From:** Ellen Porter

**Subject:** PSD-FL-246: Farmland Hydro, L.P.—MAP/DAP Draft Permit

FDEP is proposing to issue a permit to Farmland Hydro (FH) for expansion of its MAP/DAP facility. In our initial comments, we recommended that fluoride (F) and particulate matter (PM) emission limits should not exceed the 0.0417 lb F/T and 0.19 lb PM/T limits required by other permits issued by FDEP. FDEP is proposing limits of 0.0417 lb F/T and 0.3 lb PM/T for DAP and 0.06 lb F/T and 0.3 lb PM/T for MAP.

It appears that the limits proposed by FDEP are based upon statistical analyses of stack test results supplied by FH. However, the calculations of the stack test result confidence intervals are seriously flawed. The correct formula for calculating the confidence interval for a normally distributed population is:

$$X - Z\sigma/\sqrt{n} \leq \mu \leq X + Z\sigma/\sqrt{n}$$

Where X is the average of the samples, Z is the number of standard deviations from the mean for a given level of confidence,  $\sigma$  is the population standard deviation, n is the number of samples, and  $\mu$  is the population mean. FH neglected to divide the standard deviation by the square root of the number of samples—the result of that error is a much-inflated confidence interval.

We re-calculated the statistical analyses for F and PM for MAP, and the PM emissions for DAP. In evaluating the F emissions from MAP, we looked at two scenarios:

1. Table MAP-F-1 included the February 1998 test results and estimated production rates during the runs by dividing the measured hourly F emission rates by the corresponding calculated emission rates in lb F/ton  $P_2O_5$ . Combined emissions/ton versus test date are depicted in Figure MAP-F-1a, while combined emissions/ton versus production rate are depicted in Figure MAP-F-1b, and R/G stack emissions/ton versus production rate are depicted in Figure MAP-F-1c.
2. Because of our uncertainty about the very low production rate during the February 1998 run #1, and the exceptionally high R/G stack F emission rates during the March 1997 and February 1998 tests, all of those values were excluded from Table MAP-F-2 and the remaining data are depicted in Figure MAP-F-2.

Figure MAP-F-1a indicates that F emissions are increasing with the age of the MAP plant—all of the tests run after 1996 are higher than any test run before then. Figure MAP-F-1b indicates that the increase in emissions could also be due to pushing production rate beyond some threshold (60 ton/hr?) at which the scrubber loses effectiveness, but that correlation is very weak compared to the age correlation. Figure MAP-F-1c focuses on the R/G stack as the source of the increase in emissions. Even if the suspect results from March 1997 and February 1998 are retained, the upper ends of the confidence intervals in Table MAP-F-1 range from 0.0431 to 0.0464 lb F/ton  $P_2O_5$ . Table MAP-F-2 shows that, if the suspect test results are excluded, the upper ends of the confidence intervals range from 0.0253 to 0.0270 lb F/ton  $P_2O_5$ . In either case, the corrected limits are significantly lower than the 0.06 lb F/ton limit proposed in the draft permit.

In evaluating the PM emissions from MAP, we looked at three scenarios:

1. Table MAP-PM-1 included the February 1998 test results and assumed that production rates during the runs were the same as for the Fluoride tests. Emissions/ton versus test date are depicted in Figure MAP-PM-1a, while emissions/ton versus production rate are depicted in Figure MAP-PM-1b.
2. Table MAP-PM-2 excluded the February 1998 test results from run #1 because of its unusually low production rate, and assumed that production rates during the other runs were the same as for the Fluoride tests. Emissions/ton versus production rate are depicted in Figure MAP-PM-2.
3. Table MAP-PM-3 excluded the February 1998 test because of the uncertainty over our calculation of production rates. Emissions/ton versus production rate are depicted in Figure MAP-PM-3.

None of the scenarios showed any meaningful correlation between emissions and time or production rate. The high ends of the confidence intervals ranged from 0.18 to 0.20 lb PM/ton of  $P_2O_5$ , and provide little justification for FDEP relaxing its previous BACT limit of 0.19 lb PM/ton.

In evaluating PM emissions from DAP, we simply corrected FH's error in calculating the confidence interval by dividing the standard deviation by the square root of 12, the number of samples. This yielded an upper bound to the 99% confidence interval of 0.15lb PM/ton, half of the value calculated by FH and proposed in the draft permit, and below the limit previously proposed by FWS.

### **Conclusions & Recommendations**

The validity of the results from MAP test run #1 on February 18, 1998 should be checked because the production rate appears abnormally low. Production rate data should be supplied for all of the February 1998 tests.

FH incorrectly calculated the proposed limits for its modified MAP and DAP operations due to statistical errors. When those errors are corrected, typical historic emissions are found to be much lower.

Fluoride emissions from MAP production have increased dramatically over the last two years and should not be allowed to justify a permit limit higher than applied to other fertilizer plants. Instead, FH should be encouraged to investigate and remedy the cause of the increased emissions.

Particulate emissions from MAP production are relatively consistent and do not justify a permit limit higher than applied to other fertilizer plants.

When FH's statistical error is corrected, particulate emissions from DAP do not justify a permit limit higher than applied to other fertilizer plants.

Emission limits should not exceed the 0.0417-lb F/T and 0.19 lb PM/T limits required by other permits issued by FDEP.

Contact: Don Shepherd (303) 969-2075

*Number of Pages: 17  
(Including this cover sheet)*

---

*Office Location: 7333 West Jefferson Ave, Suite 450, Lakewood, CO 80235*

**MAP February 1998 Fluoride Results**

Total F							
Measured Fluoride		Calculated	Dryer Stack Fluoride		R/G Stack Fluoride		
(lb/hr)	(lb/ton)	P2O5 ton/hr	(lb/hr)	(lb/ton)	(lb/hr)	(lb/ton)	
2.296	0.048	47.8	0.687	0.014	1.609	0.034	
3.880	0.062	62.6	0.505	0.008	3.375	0.054	
5.207	0.084	62.0	0.510	0.008	4.697	0.076	

Table MAP-F-1. All MAP Fluoride Emissions Data

Test Date	DRYER					R/G					Combined	
	TPH-P205	F-LB/HR	F-LB/TON			TPH-P205	F-LB/HR	F-LB/TON			F-LB/TON	
Feb-98	47.8	0.687	0.0144			47.8	1.609	0.0336			0.0480	
Feb-94	56.4	0.700	0.0124			56.5	0.083	0.0015			0.0139	
Feb-94	56.4	0.645	0.0114			56.5	0.074	0.0013			0.0127	
Feb-94	56.4	0.537	0.0095			56.5	0.915	0.0162			0.0257	
Feb-95	56.5	0.710	0.0126			56.1	1.059	0.0189			0.0314	
Feb-95	56.5	0.787	0.0139			56.1	1.123	0.0200			0.0339	
Feb-95	56.5	0.753	0.0133			56.1	0.29	0.0052			0.0185	
May-96	58.5	0.616	0.0105			58.6	0.317	0.0054			0.0159	
May-96	58.5	0.781	0.0134			58.6	0.061	0.0010			0.0144	
May-96	58.5	0.655	0.0112			58.6	0.041	0.0007			0.0119	
Feb-98	62.0	0.510	0.0082			62.0	4.697	0.0758			0.0840	
Feb-98	62.6	0.505	0.0081			62.6	3.375	0.0539			0.0620	
Mar-97	62.2	0.474	0.0076			63.0	1.99	0.0316			0.0392	
Mar-97	62.2	0.378	0.0061			63.0	1.912	0.0303			0.0364	
Mar-97	62.2	0.346	0.0056			63.0	2.318	0.0368			0.0424	
COUNT			15					15			15	
AVERAGE			0.0105					0.0222			0.0327	
MEDIAN			0.0112					0.0189			0.0314	
ST. DEV			0.0029					0.0220			0.0206	
95% CI			0.0016	0.0091	0.0120			0.0111	0.0110	0.0333	0.0104	0.0223
99% CI			0.0019	0.0086	0.0125			0.0146	0.0075	0.0368	0.0137	0.0190

Figure MAP-F-1a. Emissions vs. Time

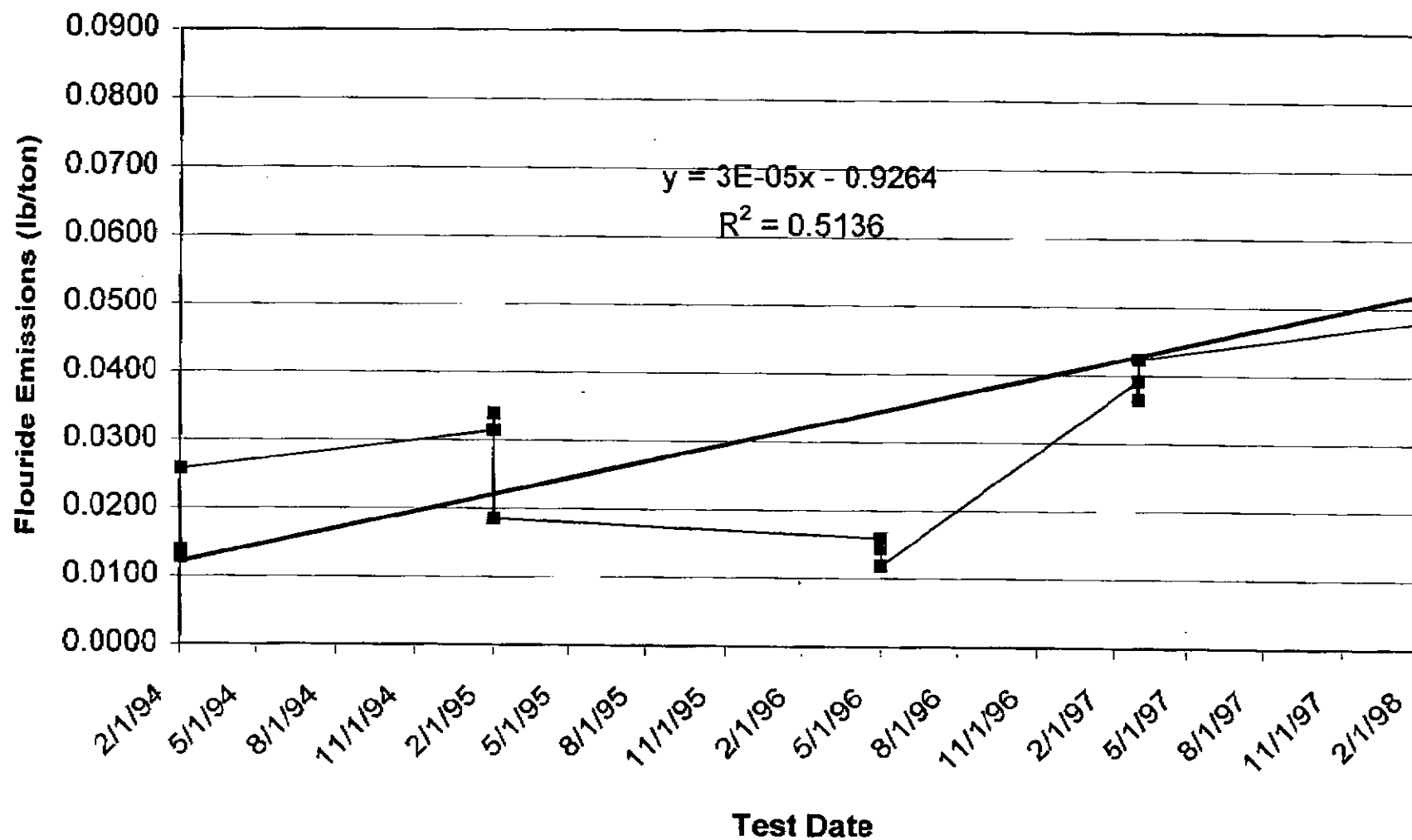


Figure MAP-F-1b. MAP Product vs. Emissions

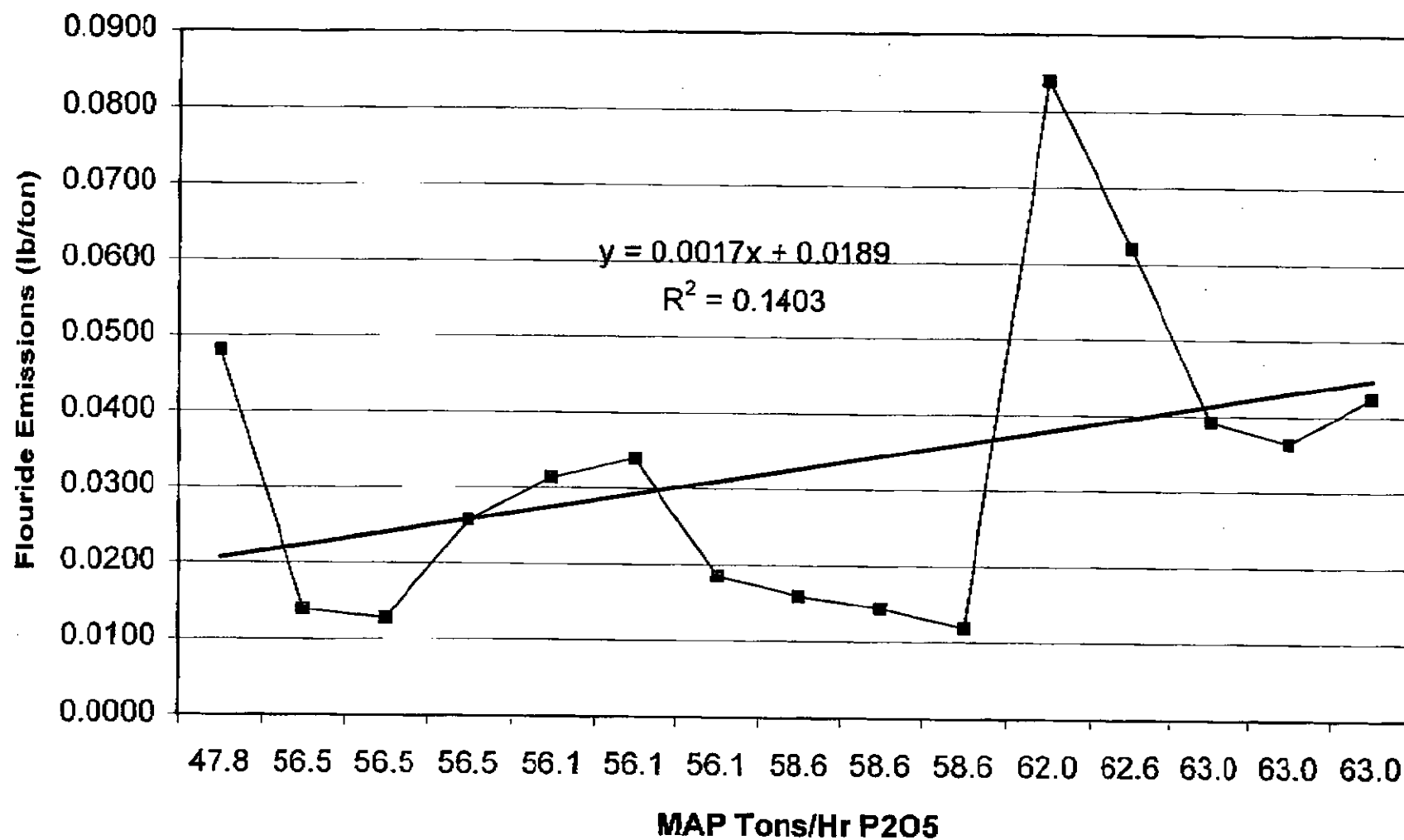




Figure MAP-F-1c. MAP Product vs. R/G Stack Emissions

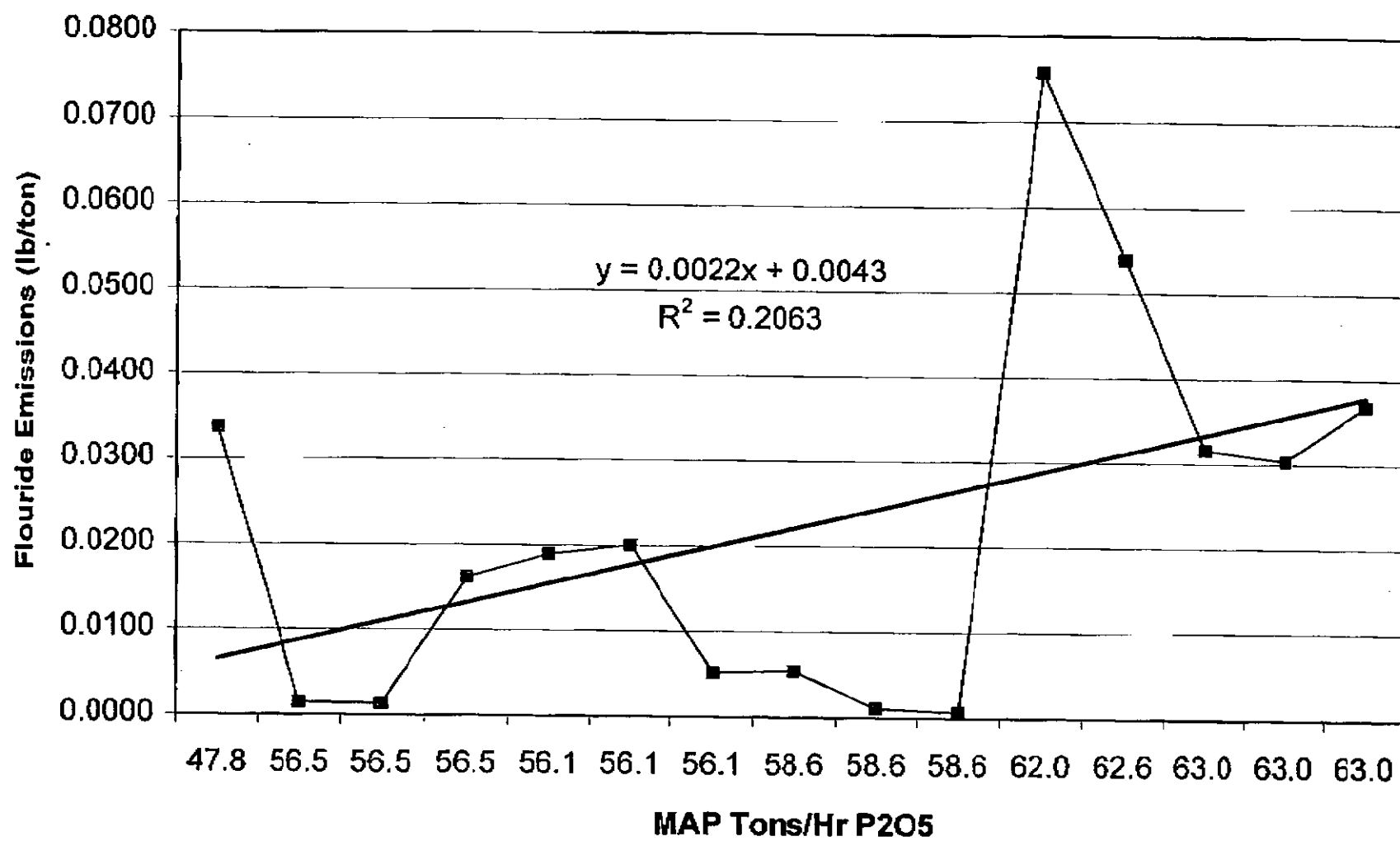


Table MAP-F-2. Edited Fluoride Emissions Data

Test Date	DRYER					R/G					Combined		
	TPH-P2O5	F-LB/HR	F-LB/TON			TPH-P2O5	F-LB/HR	F-LB/TON			F-LB/TON		
Feb-98	47.8	0.687	0.0144			47.8	1.609	0.0336			0.0480	excluded	
Feb-94	56.4	0.700	0.0124			56.5	0.083	0.0015			0.0139		
Feb-94	56.4	0.645	0.0114			56.5	0.074	0.0013			0.0127		
Feb-94	56.4	0.537	0.0095			56.5	0.915	0.0162			0.0257		
Feb-95	56.5	0.710	0.0126			56.1	1.059	0.0189			0.0314		
Feb-95	56.5	0.787	0.0139			56.1	1.123	0.0200			0.0339		
Feb-95	56.5	0.753	0.0133			56.1	0.29	0.0052			0.0185		
May-96	58.5	0.616	0.0105			58.6	0.317	0.0054			0.0159		
May-96	58.5	0.781	0.0134			58.6	0.061	0.0010			0.0144		
May-96	58.5	0.655	0.0112			58.6	0.041	0.0007			0.0119		
Feb-98	62.0	0.510	0.0082			62.0	4.697	0.0758			0.0840	excluded	
Feb-98	62.6	0.505	0.0081			62.6	3.375	0.0539			0.0620	excluded	
Mar-97	62.2	0.474	0.0076			63.0	1.99	0.0316			0.0392	excluded	
Mar-97	62.2	0.378	0.0061			63.0	1.912	0.0303			0.0364	excluded	
Mar-97	62.2	0.346	0.0056			63.0	2.318	0.0368			0.0424	excluded	
COUNT			9					9			9		
AVERAGE			0.0120					0.0078			0.0198		
MEDIAN			0.0124					0.0052			0.0159		
ST. DEV			0.0015					0.0082			0.0084		
95% CI			0.0010	0.0111	0.0130			0.0053	0.0025	0.0131	0.0055	0.0143	0.0253
99% CI			0.0013	0.0108	0.0133			0.0070	0.0008	0.0148	0.0072	0.0126	0.0270

Figure MAP-F-2. MAP Product vs. Edited Emissions

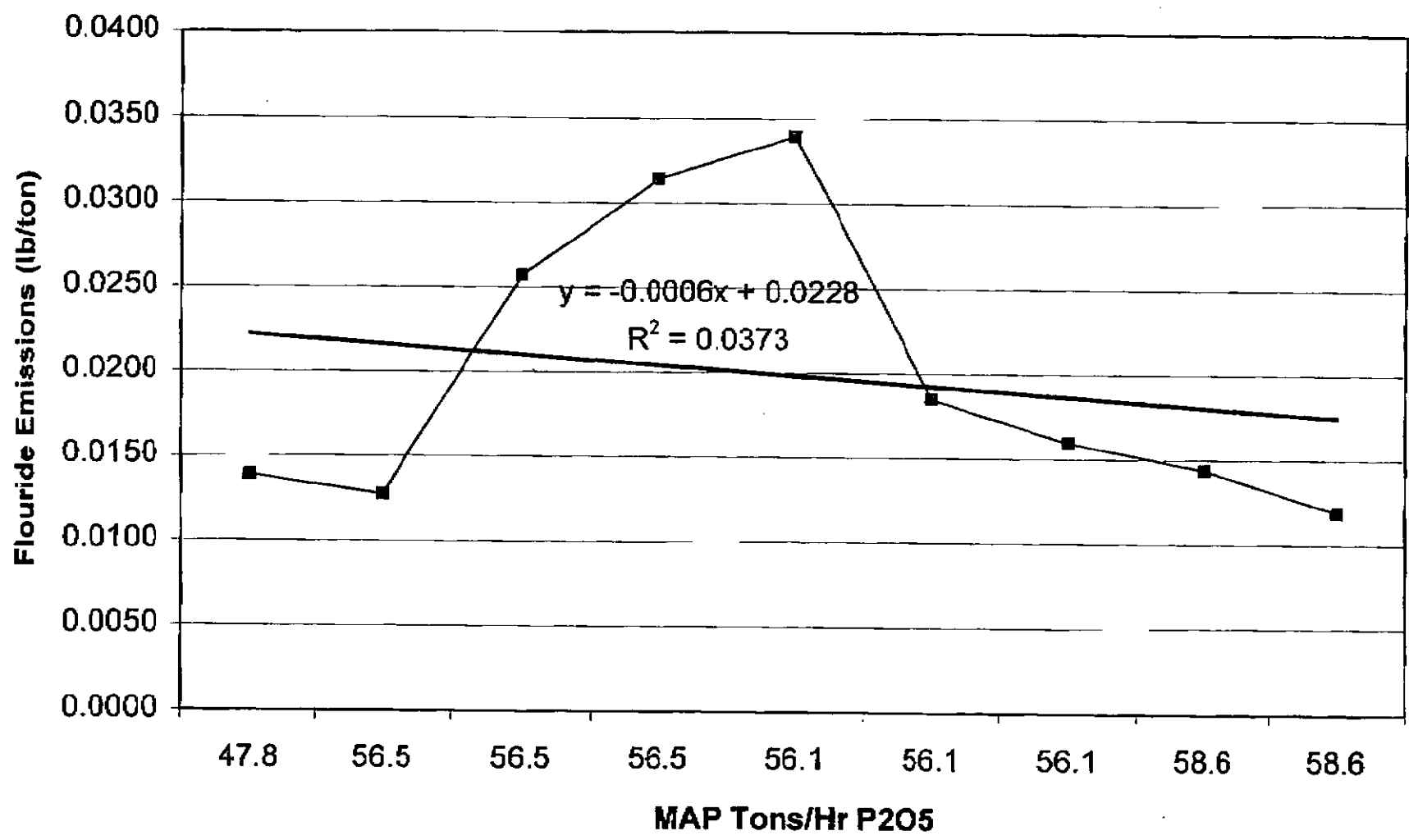


Table MAP-PM-1. All MAP PM Emissions Data

Test Date	DRYER					R/G					Combined	
	TPH-P2O5	PM-LB/HR	PM-LB/TON			TPH-P2O5	PM-LB/HR	PM-LB/TON			PM-LB/TON	
Feb-98	47.8	4.810	0.1006			47.8	1.599	0.0335			0.1341	
Feb-94	56.4	5.413	0.0960			56.5	0.940	0.0166			0.1126	
Feb-94	56.4	7.192	0.1275			56.5	0.857	0.0152			0.1427	
Feb-94	56.4	5.024	0.0891			56.5	1.570	0.0278			0.1169	
Feb-95	56.5	13.194	0.2335			56.1	2.268	0.0404			0.2739	
Feb-95	56.5	6.780	0.1200			56.1	2.610	0.0465			0.1665	
Feb-95	56.5	11.493	0.2034			56.1	1.846	0.0329			0.2363	
May-96	58.5	6.240	0.1067			58.6	2.588	0.0442			0.1508	
May-96	58.5	8.691	0.1486			58.6	3.292	0.0562			0.2047	
May-96	58.5	8.596	0.1469			58.6	2.121	0.0362			0.1831	
Feb-98	62.0	6.856	0.1106			62.0	3.897	0.0629			0.1734	
Feb-98	62.6	2.096	0.0335			62.6	1.490	0.0238			0.0573	
Mar-97	62.2	2.741	0.0441			63.0	1.489	0.0236			0.0677	
Mar-97	62.2	2.967	0.0477			63.0	1.538	0.0244			0.0721	
Mar-97	62.2	4.306	0.0692			63.0	3.790	0.0602			0.1294	
COUNT			15					15			15	
AVERAGE			0.1118					0.0363			0.1481	
MEDIAN			0.1067					0.0335			0.1427	
ST. DEV			0.0559	Min	Max			0.0152	Min	Max	0.0611	Min
95% CI			0.0283	0.0836	0.1401			0.0077	0.0286	0.0440	0.0309	0.1172
99% CI			0.0371	0.0747	0.1490			0.0101	0.0262	0.0464	0.0407	0.1074

Figure MAP-PM-1a. MAP PM Emissions vs. Time

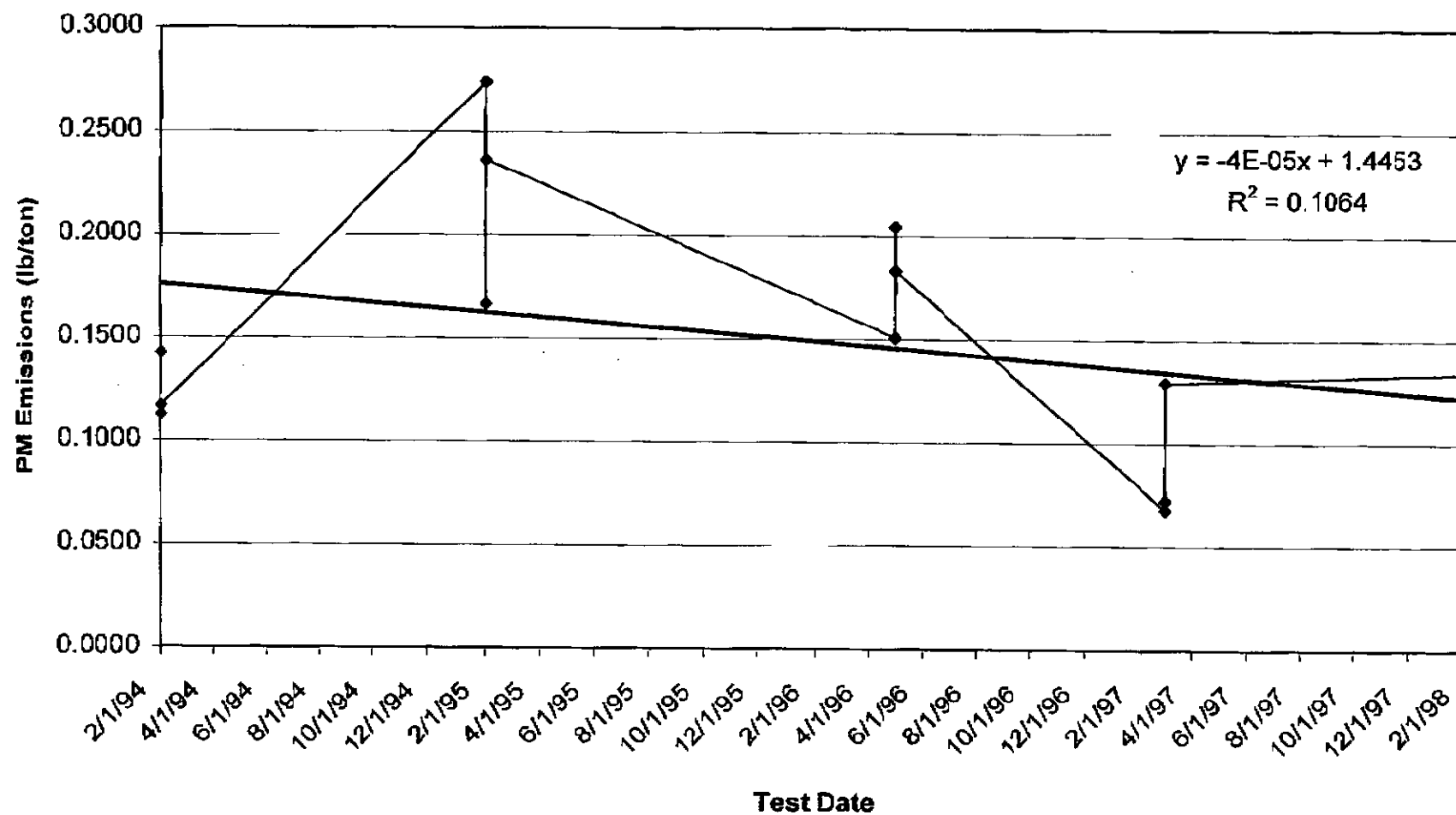


Figure MAP-PM-1b. MAP Product vs. PM Emissions

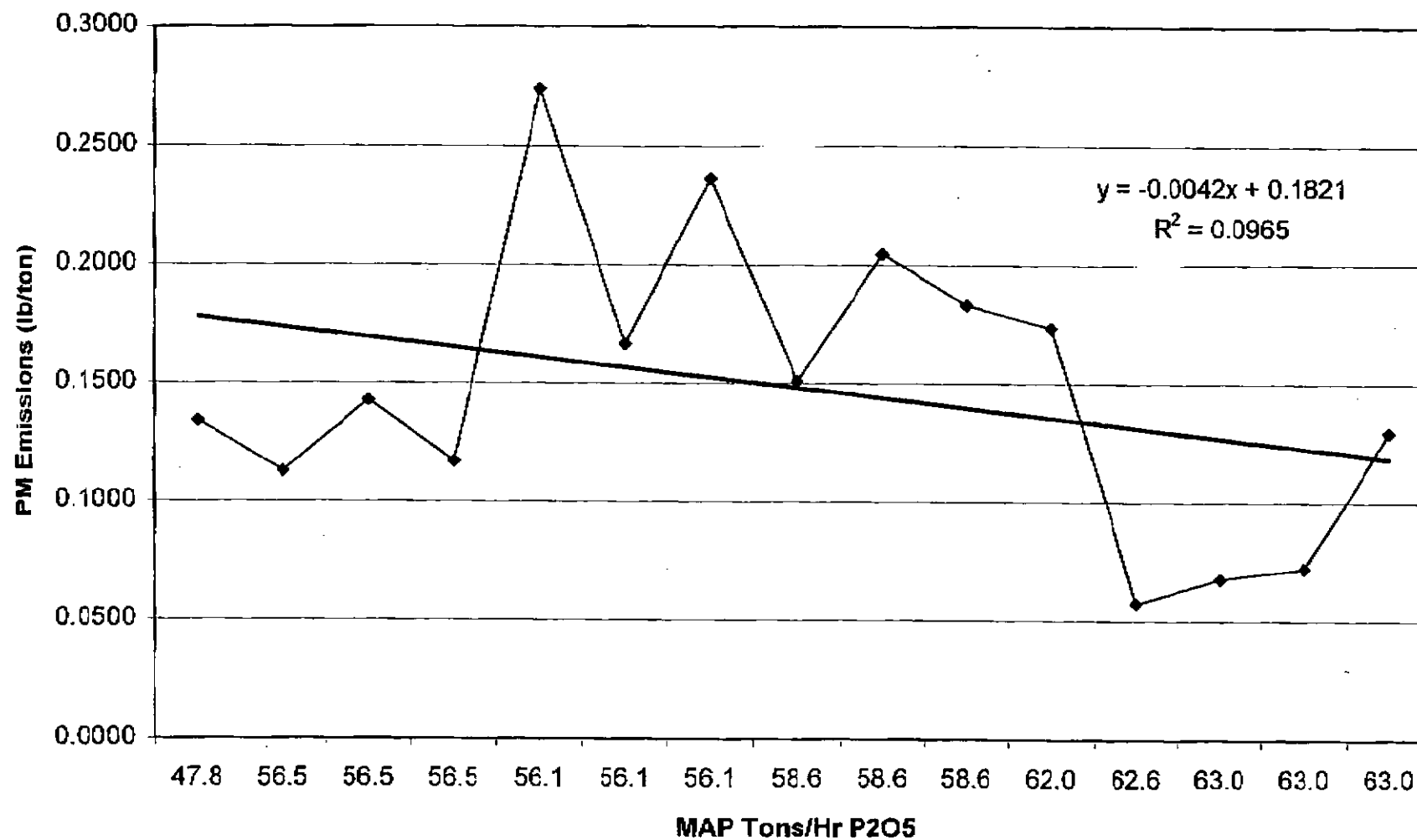


Table MAP-PM-2. Edited MAP PM Emissions Data

Test Date	DRYER					R/G					Combined	
	TPH-P2O5	PM-LB/HR	PM-LB/TON			TPH-P2O5	PM-LB/HR	PM-LB/TON			PM-LB/TON	
Feb-98	47.8	4.810	0.1006	excluded		47.8	1.599	0.0335	excluded		0.1341	excluded
Feb-94	58.4	5.413	0.0960			58.5	0.940	0.0166			0.1126	
Feb-94	58.4	7.192	0.1275			56.5	0.857	0.0152			0.1427	
Feb-94	56.4	5.024	0.0891			56.5	1.570	0.0278			0.1169	
Feb-95	56.5	13.194	0.2335			56.1	2.266	0.0404			0.2739	
Feb-95	56.5	6.780	0.1200			58.1	2.610	0.0465			0.1665	
Feb-95	56.5	11.493	0.2034			56.1	1.846	0.0329			0.2363	
May-96	58.5	8.240	0.1067			58.6	2.588	0.0442			0.1508	
May-96	58.5	8.691	0.1486			58.6	3.292	0.0562			0.2047	
May-96	58.5	8.596	0.1469			58.6	2.121	0.0362			0.1831	
Feb-98	62.0	6.856	0.1106			62.0	3.897	0.0629			0.1734	
Feb-98	62.6	2.096	0.0335			62.6	1.490	0.0238			0.0573	
Mar-97	62.2	2.741	0.0441			63.0	1.489	0.0236			0.0677	
Mar-97	62.2	2.967	0.0477			63.0	1.538	0.0244			0.0721	
Mar-97	62.2	4.306	0.0692			63.0	3.790	0.0602			0.1294	
COUNT			14					14			14	
AVERAGE			0.1126					0.0366			0.1491	
MEDIAN			0.1086					0.0346			0.1468	
ST. DEV			0.0579	Min	Max			0.0158	Min	Max	0.0633	Min
95% CI			0.0303	0.0823	0.1429			0.0083	0.0282	0.0447	0.0332	0.1159
99% CI			0.0398	0.0728	0.1525			0.0108	0.0256	0.0473	0.0436	0.1055

Table MAP-PM-2. MAP Product vs. PM Emissions

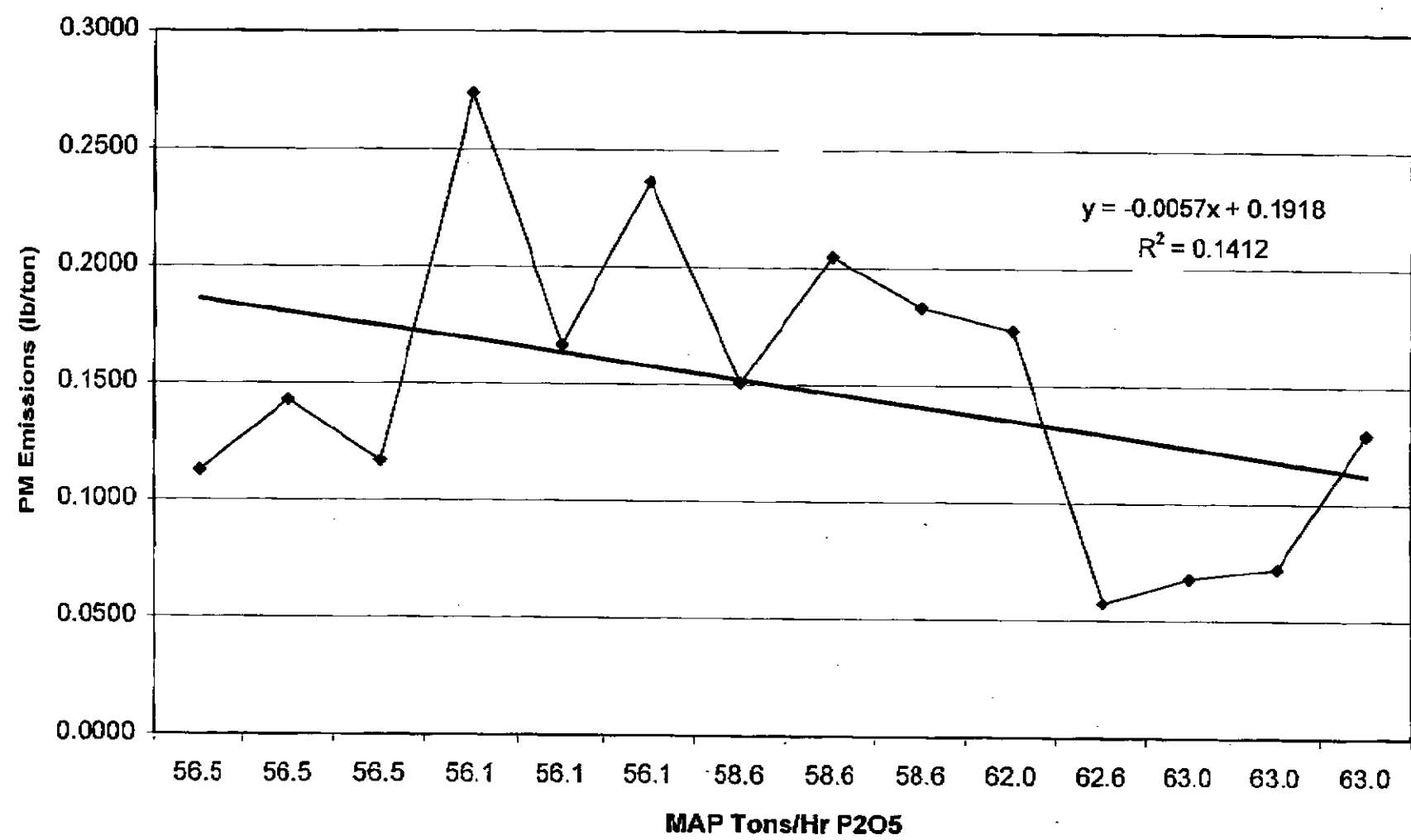
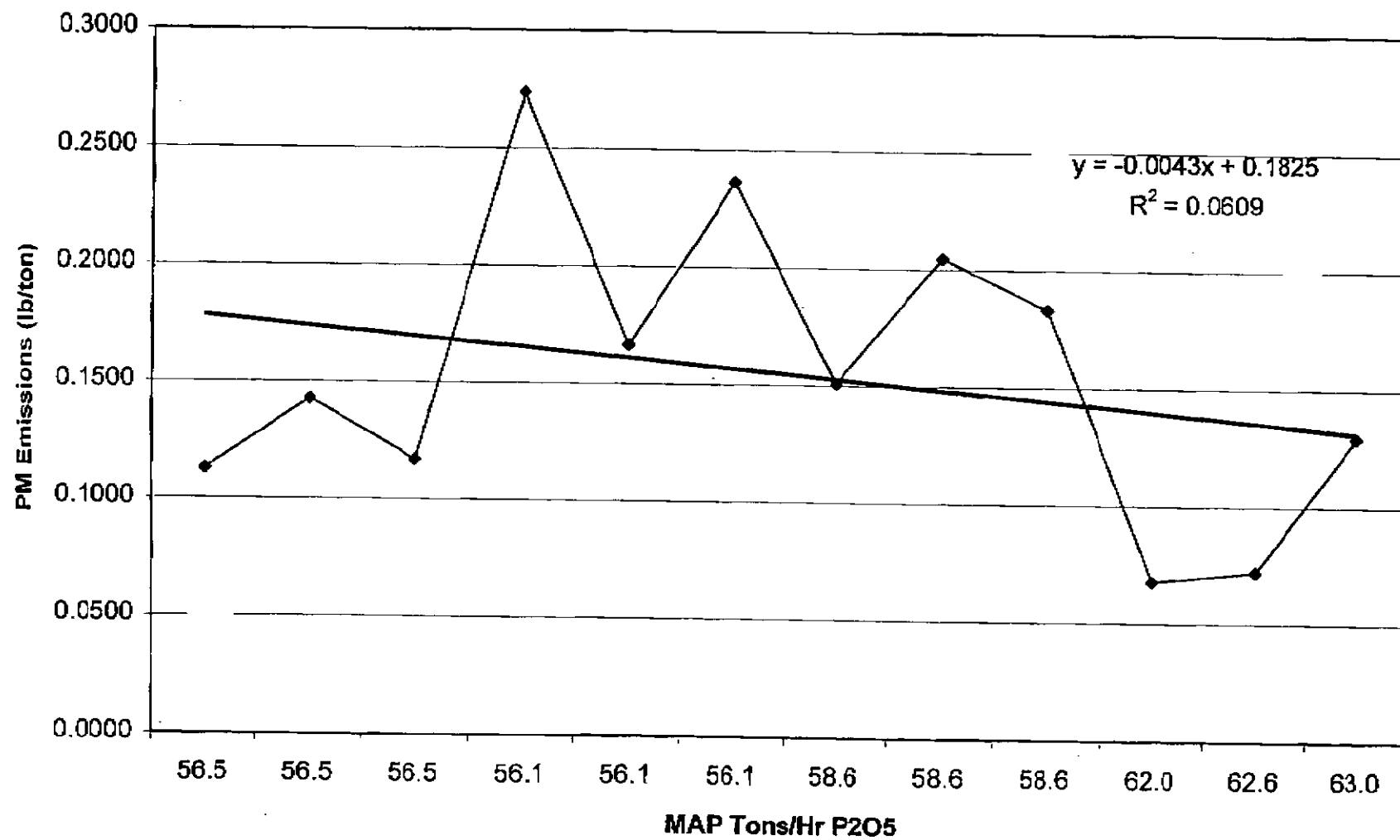




Table MAP-PM-3. MAP PM Emissions Data prior to Feb 98

Test Date	DRYER					R/G					Combined	
	TPH-P2O5	PM-LB/HR	PM-LB/TON			TPH-P2O5	PM-LB/HR	PM-LB/TON			PM-LB/TON	
Feb-98	47.8	4.810	0.1006	excluded		47.8	1.599	0.0335	excluded		0.1341	excluded
Feb-94	56.4	5.413	0.0960			56.5	0.940	0.0166			0.1126	
Feb-94	56.4	7.192	0.1275			56.5	0.857	0.0152			0.1427	
Feb-94	56.4	5.024	0.0891			56.5	1.570	0.0278			0.1189	
Feb-95	56.5	13.194	0.2335			56.1	2.288	0.0404			0.2739	
Feb-95	56.5	6.780	0.1200			56.1	2.610	0.0465			0.1665	
Feb-95	56.5	11.493	0.2034			56.1	1.846	0.0329			0.2363	
May-96	58.5	6.240	0.1067			58.6	2.588	0.0442			0.1508	
May-96	58.5	8.691	0.1486			58.6	3.292	0.0562			0.2047	
May-96	58.5	8.596	0.1469			58.6	2.121	0.0362			0.1831	
Mar-97	62.2	2.741	0.0441			63.0	1.489	0.0236			0.0677	
Mar-97	62.2	2.967	0.0477			63.0	1.538	0.0244			0.0721	
Mar-97	62.2	4.306	0.0692			63.0	3.790	0.0602			0.1294	
COUNT			12					12			12	
AVERAGE			0.1194					0.0353			0.1547	
MEDIAN			0.1133					0.0346			0.1468	
ST. DEV			0.0578					0.0146			0.0623	
95% CI			0.0327	0.0867	0.1521			0.0083	0.0271	0.0436	0.0353	0.1195
99% CI			0.0430	0.0764	0.1624			0.0109	0.0245	0.0462	0.0463	0.1084

Figure MAP-PM-3. MAP Product vs. PM Emissions





KOOGLER & ASSOCIATES  
ENVIRONMENTAL SERVICES

4014 NW THIRTEENTH STREET  
GAINESVILLE, FLORIDA 32609  
352/377-5822 ■ FAX/377-7158

RECEIVED

JUL 22 1998

BUREAU OF  
AIR REGULATION

KA 123-97-01

July 21, 1998

Mr. Perry Odom, Esq.  
Office of General Counsel  
Florida Department of  
Environmental Protection  
Twin Towers Office Building  
2600 Blair Stone Road  
Tallahassee, FL 32399-2400

Subject: Motion for Extension of Time to  
File a Petition

*Farmland*  
*PSD-FI-246*

Dear Mr. Odom:

Attached is a request for an extension of time to file for a hearing in accordance with Rule 28-106, FAC.

If you have any questions concerning this request, please do not hesitate to contact me.

Very truly yours,

KOOGLER & ASSOCIATES

John B. Koogler, Ph.D., P.E.

JBK:par  
Enc.

c: Mr. Syed Arif, DEP  
Mr. Charles Jenkins, Farmland

STATE OF FLORIDA  
DEPARTMENT OF ENVIRONMENTAL PROTECTION

In the Matter of an Application  
for Air Permit by

Farmland Hydro, L.P.  
P.O. Box 960  
Bartow, FL 33831

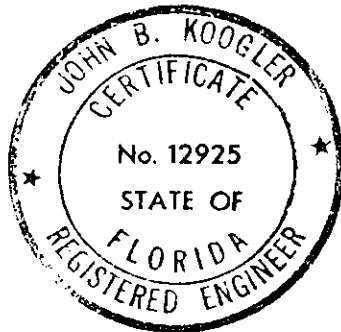
DEP File No. 1050053-020-AC and  
PSD-FL-246  
Polk County - AP

MOTION FOR EXTENSION OF TIME


The Applicant, Farmland Hydro, L.P. (Farmland), by and through its undersigned Engineer of Record and pursuant to Rule 28-106, FAC, requests the Secretary of DEP to grant a 60-day extension of time in which to file a petition. The additional time will allow Farmland to submit additional information to DEP on the North MAP/DAP Plant permit application review.

The DEP Permitting Engineer, Mr. Syed Arif, has indicated that he has no objection to such an extension.

Dated the 21st day of July, 1998 in Gainesville, Alachua County, Florida.

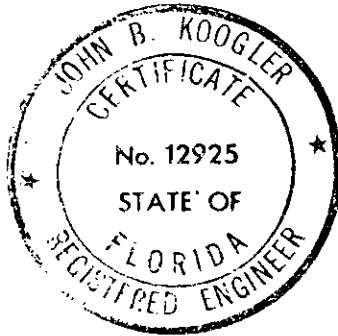


Koogler & Associates  
Environmental Services

  
\_\_\_\_\_  
John B. Koogler, Ph.D., P.E.  
Florida Registration No. 12925  
4014 N.W. 13th Street  
Gainesville, FL 32609  
(352) 377-5822  
Engineer of Record for  
Farmland Hydro, L.P.

CERTIFICATE OF SERVICE

I hereby certify that a copy of the foregoing has been furnished to Mr. Perry Odom (OGC) and Mr. Syed Arif (BAR), DEP, 2600 Blair Stone Road, Tallahassee, Florida 32399-2400 and Mr. Charles Jenkins, Manager Environmental & Safety Services, Farmland Hydro, L.P., P.O. Box 960, Bartow, FL 33831, by FAX and by U.S. Mail, this 21st day of July 1998.



  
\_\_\_\_\_  
John B. Koogler, Ph.D., P.E.



KOOGLER & ASSOCIATES  
ENVIRONMENTAL SERVICES  
4014 NW THIRTEENTH STREET  
GAINESVILLE, FLORIDA 32609  
352/377-5822 • FAX/377-7158

KA 123-97-01

July 21, 1998

RECEIVED

JUL 22 1998

BUREAU OF  
AIR REGULATION

Mr. Syed Arif  
Florida Department of  
Environmental Protection  
Twin Towers Office Building  
2600 Blair Stone Road  
Tallahassee, FL 32399-2400

Subject: Comments on Draft Permit  
North MAP/DAP Plant  
Farmland Hydro, L.P.  
DEP File No. 1050053-020-AC, PSD-FL-246

Dear Mr. Arif:

We have received and reviewed FDEP's draft permit, dated July 2, 1998, for the North MAP/DAP Plant. As we have some comments on the draft permit which may require additional time to address, an extension of time to file a petition for an administrative hearing has been submitted to OGC (copy attached).

The following comments are submitted for your consideration:

1. Specific Condition No. 5

It is possible that the rule citation should reflect Rule 62-212.400 (instead of 62-212.410), FAC.

2. Specific Condition No. 9

The sulfur dioxide emission limits should be removed as there is no applicable standard. It should be noted that the SO<sub>2</sub> limits in another company's recent DAP project were applicable based on a BACT analysis. This is not the case in the proposed project. The emissions estimates provided in Farmland's permit application were to demonstrate that the expected emissions would not trigger any regulatory requirements for SO<sub>2</sub>. Farmland does not object to the portion of the condition which states:

"During periods of firing No. 2 fuel oil with a maximum sulfur content of 0.05 percent by weight, the firing rate shall not exceed 50 MMBtu/hr and 3.1 million gallons per year. The permittee shall maintain records of the fuel oil supplier's sulfur content analysis."

3. Specific Condition No. 10

This condition, containing nitrogen oxides emission limits, should be deleted as there is no applicable standard. It should be noted that the NOx limits in another company's recent DAP project were applicable based on a BACT analysis. This is not the case in the proposed project. The emissions estimates provided in Farmland's permit application were to demonstrate that the expected emissions would not trigger any regulatory requirements for NOx.

4. Specific Condition No. 11

It is our understanding that the continuous pressure drop monitoring requirement applies to the scrubbers used as pollution control equipment and not to those which serve as process design equipment. Accordingly, the pressure drop monitoring requirement should not be required for the HI-MOL scrubber. Furthermore, it is requested that the condition allow for measurement of fan amps, in place of pressure drop, as allowed under draft Title V permit conditions.

Also, it is possible that the rule citation should reflect Rule 62-204.800 (instead of 62-296.800), FAC.

5. Specific Condition No. 14

In view of comment 2, EPA Method 7E should be removed from this condition as it is not applicable.

6. Specific Condition No. 16

It is possible that the rule citation should reflect Rule 62-204.800 (instead of 62-296.800), FAC.

7. Specific Condition No. 21

The product storage and shipping rate provided to FDEP was incorrect. We apologize for the oversight. The 120 tons per hour (tph) P205 rate conveyed to the Department corresponds to the rate associated with the North MAP/DAP Plant of 106.1 tph P205 storage and 120 tph P205 loadout, or a maximum rate of 120 tph P205. However, the South DAP Plant also contributes to the storage and shipping building with a permitted rate of 46 tph P205. Consequently, the storage and shipping building would handle a combined total of 152.1 tph P205 storage and 180 tph P205 loadout, or a maximum rate of 180 tph P205.

As stated in the permit application, the PM emission rate is not expected to change as a result of the proposed project as no changes are proposed

Mr. Syed Arif  
Florida Department of  
Environmental Protection

July 21, 1998  
Page 3

to the exhaust flow rate or the existing scrubber's operating parameters. With the fan operating at the same rate for 8760 hours both before and after the proposed project, and at the same exhaust particulate loading, no change in the mass emission rate is expected. Therefore, in our opinion, the storage and shipping building emissions would not be subject to PSD/BACT review. However, Farmland is willing to accept more stringent emission limits (discussed below) to expedite the permitting process.

Per your request, the historical particulate matter compliance test information is summarized in Attachment 1. Based on the compliance test results, Farmland should be able to comply with a particulate matter emission limit of 4.1 pounds per hour and 18 tpy. Information requested by you on the existing scrubber, is presented in Attachment 2.

Please note that the fluoride emission limit in the current operating permit is no longer applicable as GTSP is no longer manufactured and stored at the facility. A federally enforceable condition in the 1992 North MAP/DAP Plant PSD permit required that the GTSP production capability be removed.

If you have any questions, please call Pradeep Raval or me.

Very truly yours,

KOOGLER & ASSOCIATES

  
John B. Koogler, Ph.D., P.E.

JBK:par

c: Charles Jenkins, Farmland Hydro, L.P.

(CC: file)  
SWD  
Pork Co  
EPA  
NPS  
C. Holladay, BAR



ATTACHMENT 1

PM COMPLIANCE TEST RESULTS  
NORTH MAP/DAP PLANT

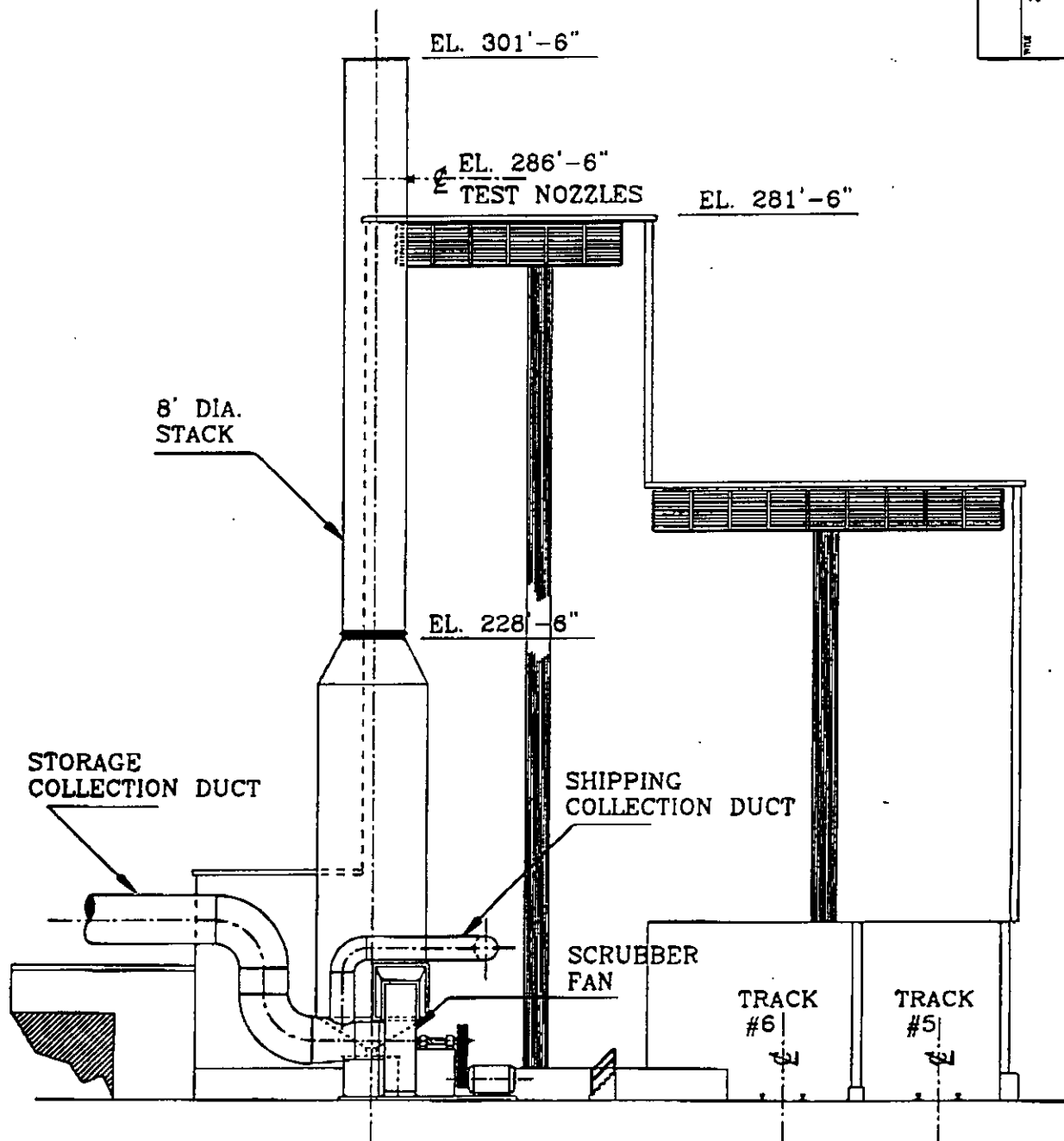
Test Date		PM Emission Rate (lbs/hr)
1992	Run 1	1.308
	Run 2	0.989
	Run 3	1.166
1993	Construction	
1994	Run 1	0.954
	Run 2	0.939
	Run 3	0.949
1995	Run 1	1.072
	Run 2	0.769
	Run 3	1.072
1996	Run 1	0.363
	Run 2	0.364
	Run 3	0.641
1997	Run 1	0.502
	Run 2	0.501
	Run 3	0.449
1998	Run 1	1.089
	Run 2	6.993
	Run 3	0.924
Average		1.169
S. Dev.		1.482
Avg + 2* SD		4.133

ATTACHMENT 2

EXISTING STORAGE/SHIPPING SCRUBBER INFORMATION  
NORTH MAP/DAP PLANT

Manufacturer:	ARCO (Detroit)
Model:	S.O. # 8928
Installation Date:	1965
Scrubber Dimensions:	14 feet diameter, 52 feet tall
Scrubbing Medium:	Pond water
Liquid Flow Rate:	1000 gpm
Gas Flow Rate:	80,000 cfm
Pressure Drop:	4 inches H2O

FARMLAND HYDRO L.P. BARTOW, FLORIDA		DATE 1975	BY NONE	9/23/85
STORAGE AND SHIPPING SCRUBBER		DATE 1975	BY NONE	9/23/85
SOURCE NO. 020		DATE 1975	BY NONE	9/23/85



# SHIPPING/STORAGE SCRUBBER

VIEW LOOKING WEST

KOOGLER & ASSOCIATES  
ENVIRONMENTAL SERVICES4014 NW THIRTEENTH STREET  
GAINESVILLE, FLORIDA 32609  
352/377-5822 • FAX/377-7158PROJECT 123-97-01

## FAX TRANSMITTAL FORM

TO:

Mr. Perry Olson  
Mr. Scott O'Neil  
Mr. Charles Jenkins

FAX NO. \_\_\_\_\_

FROM:

John Koozler

DATE:

7/21/98

SENT BY:

Wendy

The text being transmitted consists of 3 page(s) PLUS this one. If you do not receive all of the pages or if there are difficulties with this transmission, please call (352) 377-5822.

REMARKS: \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

This message is intended for use only by the individual to whom it has been addressed and may contain confidential or privileged information. If you are not the intended recipient, please note that the use, copying or distribution of this information is not permitted. If you have received this FAX in error, please destroy the original and notify the sender immediately at (352) 377-5822 so that we may prevent any recurrence. Thank you.



KOOGLER & ASSOCIATES  
ENVIRONMENTAL SERVICES

4014 NW THIRTIETH STREET  
GAINESVILLE, FLORIDA 32609  
352/377-0022 • FAX/377-7158

KA 123-97-01

July 21, 1998

Mr. Perry Odom, Esq.  
Office of General Counsel  
Florida Department of  
Environmental Protection  
Twin Towers Office Building  
2600 Blair Stone Road  
Tallahassee, FL 32399-2400

Subject: Motion for Extension of Time to  
File a Petition

Dear Mr. Odom:

Attached is a request for an extension of time to file for a hearing in  
accordance with Rule 28-106, FAC.

If you have any questions concerning this request, please do not hesitate  
to contact me.

Very truly yours,

KOOGLER & ASSOCIATES

  
John B. Koogler, Ph.D., P.E.

JBK:par  
Enc.

c: Mr. Syed Arif, DEP  
Mr. Charles Jenkins, Farmland

STATE OF FLORIDA  
DEPARTMENT OF ENVIRONMENTAL PROTECTION

In the Matter of an Application  
for Air Permit by

Farmland Hydro, L.P.  
P.O. Box 960  
Bartow, FL 33831

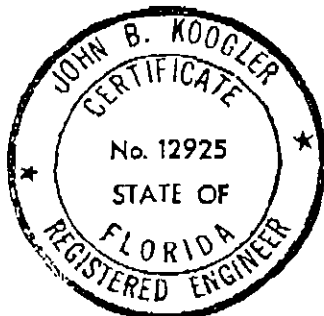
DEP File No. 1050053-020-AC and  
PSD-FL-246  
Polk County - AP

MOTION FOR EXTENSION OF TIME


The Applicant, Farmland Hydro, L.P. (Farmland), by and through its undersigned Engineer of Record and pursuant to Rule 28-106, FAC, requests the Secretary of DEP to grant a 60-day extension of time in which to file a petition. The additional time will allow Farmland to submit additional information to DEP on the North MAP/DAP Plant permit application review.

The DEP Permitting Engineer, Mr. Syed Arif, has indicated that he has no objection to such an extension.

Dated the 21st day of July, 1998 in Gainesville, Alachua County, Florida.

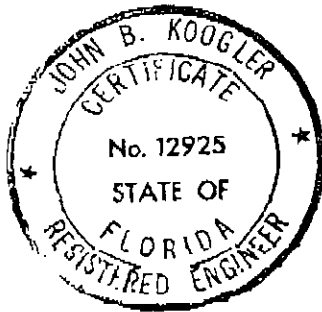


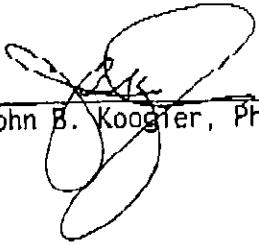
Koogler & Associates  
Environmental Services

  
John B. Koogler, Ph.D., P.E.  
Florida Registration No. 12925  
4014 N.W. 13th Street  
Gainesville, FL 32609  
(352) 377-5822  
Engineer of Record for  
Farmland Hydro, L.P.

CERTIFICATE OF SERVICE

I hereby certify that a copy of the foregoing has been furnished to Mr. Perry Odom (OGC) and Mr. Syed Arif (BAR), DEP, 2600 Blair Stone Road, Tallahassee, Florida 32399-2400 and Mr. Charles Jenkins, Manager Environmental & Safety Services, Farmland Hydro, L.P., P.O. Box 960, Bartow, FL 33831, by FAX and by U.S. Mail, this 21st day of July 1998.



  
John B. Koogler, Ph.D., P.E.