Journal of Hydrology, 28 (1976) 317–329

PROBABILITY MODELS OF WASTEWATER TREATMENT PLANT OPERATION

Jonathan W. Bulkley

Associate Professor
University of Michigan
Ann Arbor, Michigan

1. INTRODUCTION

This paper represents a bried report upon the question of failure in a wastewater treatment plant. Typically the engineer designs a wastewater treatment plant to remove a fixed percentage of constituents of the wastewater. For example, the design may be to remove 90% (each) of the volatile suspended solids (VS), Biochemical Oxygen Demand (BOD) and Total Phosphorous as P. In evaluating proposed wastewater treatment plants, the implicit assumption is that plants will operate at their design level. This study originated as an effort to examine actual operating data from a wastewater treatment facility for the purpose of comparing performance data with anticipated design standards. In particular, we set out to utilize probability models as representatives of plant performance. This may prove to be valuable for several purposes. First, one may be able to associate specified probability models with specific types of treatment processes. Secondly, explicit means would be available for comparing alternative treatment techniques in terms of alternative probability models. Third, the utilization of probability models for treatment plant operations may provide a more systematic basis for the use of Monte-Carlo simulations of alternative treatment plants planned for a specific location and provide more complete information upon anticipated environmental impacts of the wastewater effluent upon the receiving waters.

2. DATA ACQUIRED

In order to accomplish the research objectives, three years of daily operating data were obtained from the wastewater treatment plant at Ann Arbor, Michigan. The parameters studied were Voltaire Suspended Solids (VS), Biochemical Oxygen Demand (BOD), and Total Phosphorous (P). On a daily basis, the data for each of these parameters was coded for four different points in the treatment system, i.e., incoming to the plant, primary influent, primary effluent, and final effluent. The Ann Arbor wastewater treatment plant is a secondary treatment plant which is designed to remove 90% of the voltile solids and BOD. During the period of investigation, the wastewater treatment plant implemented chemical removal of phosphorous by addition of ferric chloride prior to primary treatment. Also, during the entire three year period, the Ann Arbor treatment plant has been operating near its design capacity of 15 mgd. The data analyzed is from October 1, 1969 through September 30, 1972, and comprises a data set of 1096 cases - each case represents a single day; for each day the four data points for each of the three determinants are included.

3. STATISTICAL ANALYSIS

Previous investigation has demonstrated the importance of time series analysis as opposed to conventional statistical analysis in the performance of wastewater treatment plant However, this investigation has been directed toward more conventional statistical analysis through utilization of the concept of removal efficiency as the key indicator for examination. Accordingly the following conversion was performed on a daily basis for each of the three parameters being examined:

Volatile Solid Removal Efficiency =
(VSREMEFF)

$$\frac{\text{Volatile Solids}_{\text{incoming}} - \text{Volatile Solids}_{\text{Final Effluent}}}{\text{Volatile Solids}_{\text{incoming}}}$$

BOD Removal Efficiency =
(BODREMEFF)

$$\frac{BOD_{incoming} - BOD_{Final\ Effluent}}{BOD_{incoming}}$$

Phosphorous Removal Efficiency =
(PHREMEFF)

$$\frac{P_{incoming} - P_{Final\ Effluent}}{P_{incoming}}$$

For example, if on a particular day $BOD_{incoming}$ = 200 mg/1

$BOD_{Final\ Effluent}$ = 20 mg/1

Then the BODREMEFF for that day is as follows:

$$BODREMEFF = \frac{200-20}{200} = \frac{180}{200} = .90$$

or a removal efficiency of 90%.
These conversions were performed upon the data and the results plotted utilizing the MIDAS (Michigan Interactive Data Analysis System) at the Computing Center at the University of Michigan. The examination of the plotted results for the entire data set of 1096 days indicated a sharp break in the removal efficiency of both volatile solids and Biochemical Oxygen Demand once the phosphorous removal activity became stable. The following partitioning of the entire data base reflects the situation of stabilizing phosphorous removal:

DATA1 885 Days
October 1, 1969 - March 3, 1972

DATA2 211 Days
March 4, 1972 - September 30, 1972

DATAT (Entire Data Set) 1096 Days
October 1, 1969 - September 30, 1972

The plotted results indicated that the technique of phosphorous removal became stable after March 4, 1972. Table 1 specifies the summary statistics for volatile solids removal efficiency, biochemical oxygen demand removal efficiency and phosphorous removal efficiency for DATA1 (October 1, 1969 - March 3, 1972) and for DATA2 (March 4, 1972 - September 30, 1972). From Table 1, it is clear that the operating performance of the wastewater treatment plant is much improved once the phosphorous removal system had been perfected. For example, the mean removal efficiency for phosphorous removal increased from .274 to .891; the coefficient of variation for phosphorous removal decreased from 1.11 to .063. Similar improvements in removal efficiency are also observed for both volatile solids and biochemical oxygen demand. Given this dramatic improvement in plant performance, it was decided to analyze DATA1 and DATA2 separately for the purpose of determining whether or not probability functions could be utilized to describe treatment plant performance in terms of removal efficiency.

For removal efficiency the performance data would normally be bounded by the (0,1) range. Table 1 indicates that in the DATA1 set negative removal efficiencies were observed for both volatile solids and phosphorous - i.e., in the extreme case more volatile solids and phosphorous left the treatment plant in the final effluent than came into the plant. This situation reflected operational difficulties which have been corrected. With the $0 \rightarrow 1.0$ range for performance data, a Beta distribution appeared to be appropriate for representation of the efficiency removal. Accordingly, six separate Beta distributions were estimated for the following cases:

DATA1	DATA2
(1) VSREMEFF	(4) VSREMEFF
(2) BODREMEFF	(5) BODREMEFF
(3) PHREMEFF	(6) PHREMEFF

The Beta probability density function defined over the interval (0,1) is

$$f(x;\gamma,\eta) = \frac{\Gamma(\gamma+\eta)}{\Gamma(\gamma)\Gamma(\eta)} x^{\gamma-1} (1-x)^{\eta-1} \qquad \begin{array}{l} 0 \leq x \leq 1 \\ 0 < \gamma \\ 0 < \eta \end{array}$$

TABLE 1: Removal Efficiencies

Volatile Solids Remocal Efficiency

DATA SET	N	MEAN	STANDARD DEVIATION	MIN	MAX	COEFFICIENT OF VARIATION
DATA1	883	.717	.233	-2.733	.982	.326
DATA2	209	.869	.052	.60	.966	.059

Biochemical Oxygen Demand Removal Efficiency

DATA SET	N	MEAN	STANDARD DEVIATION	MIN	MAX	COEFFICIENT OF VARIATION
DATA1	881	.842	.116	.181	.987	.137
DATA2	210	.922	.025	.818	.977	.028

Phosphorous Removal Efficiency

DATA SET	N	MEAN	STANDARD DEVIATION	MIN	MAX	COEFFICIENT OF VARIATION
DATA1	868	.274	.305	-4.3	1.0	1.11
DATA2	187	.891	.063	.449	1.0	.063

The exact shape of the Beta Distribution is a function of the two parameters, γ, η. In the removal efficiencies case, the random variable x represents removal efficiency - i.e., from 0 - 1.0. The Beta Distribution parameters, γ, η are estimated by the following equations:

$$\hat{\eta} = \frac{(1-\bar{X})}{S^2}[\bar{X}(1-\bar{X}) - S^2]$$

$$\hat{\gamma} = \frac{\bar{X}}{1-\bar{X}} \hat{\eta}$$

where $\bar{X}$ = mean of each removal efficiency

S = standard deviation of each removal efficiency

Table 2 summarizes the calculated estimates of the Beta Distri·

TABLE 2: Beta Distribution Parameters.

DATA1	DATA2
(1) VSREMEFF	(1) VSREMEFF
$\hat{\eta}$ = .77	$\hat{\eta}$ = 5.41
$\hat{\gamma}$ = 1.95	$\hat{\gamma}$ = 35.75
(2) BODREMEFF	(2) BODREMEFF
$\hat{\eta}$ = 1.4	$\hat{\eta}$ = 8.6
$\hat{\gamma}$ = 7.5	$\hat{\gamma}$ = 101.6
(3) PHRE,EFF	(3) PHREMEFF
$\hat{\eta}$ = 828	$\hat{\eta}$ = 2.56
$\hat{\gamma}$ = .313	$\hat{\gamma}$ = 21.

DATA1		
VSREMEFF	$\hat{\eta} < 1, \hat{\gamma} > 1$	Distribution J-Shaped
BODREMEFF	$\hat{\eta} > 1, \hat{\gamma} > 1$	Single Peak* (X=.94)
PHREMEFF	$\hat{\eta} < 1, \hat{\gamma} < 1$	U-Shaped
DATA2		
VSREMEFF	$\hat{\eta} > 1, \hat{\gamma} > 1$	Single Peak* (X=.887)
BODREMEFF	$\hat{\eta} > 1, \hat{\gamma} > 1$	Single Peak* (X=.93)
PHREMEFF	$\hat{\eta} > 1, \hat{\gamma} > 1$	Single Peak* (X=.927)

* Peak located at $X = \frac{\hat{\gamma}-1}{\hat{\gamma}+\hat{\eta}-2}$

bution parameters, $\hat{\eta}$ and $\hat{\gamma}$ for the six cases investigated. As previously indicated the exact shape for the Beta Distribution is determined by $\hat{\gamma}$ and $\hat{\eta}$. Accordingly, the six cases examined lead to the following shapes for the Beta Distribution:

Given $\hat{\eta}$, $\hat{\gamma}$ one may utilize tabulated references to obtain the comulative Beta Distribution for the six cases under investigation.

It should be noted that for these cases where the value of the Beta Distribution parameter, $\hat{\gamma}$, exceeds 20, the Normal Distribution may be utilized in place of the Beta Distribution. Accordingly, for the three parameters of removal efficiency in DATA2, VSREMEFF, BODREMEFF, and PHREMEFF, Normal Distributions were utilized in addition to Beta Distribution to obtain commulative distribution functions for these three removal efficiencies.

Incoming wastewater

Table 3 summarizes the data regarding the strength characteristics of the wastewater incoming to the treatment plant for the three parameters of Volatile Solids, Biochemical Oxygen Demand, and Phosphorous (as P).

Goodness of fit to cummulative distributions of removal efficiencies

The identification of specific probability models which may be utilized to represent the removal efficiency in subsequent analysis is a first step. However, one must test the goodness of fit between the observed distribution function and the assumed distribution function. In this study of six removal efficiency cases, six of the assumed distributions are Beta Distributions with parameters $\hat{\gamma}$, $\hat{\eta}$ as shown in Table 3. Three additional assumed distributions are Normal with means and standard deviations as specified in Table 1. The Normal Distributions are associated with removal efficiencies of volatile solids, biochemical oxygen demand, and phosphorous in data set, DATA2.

The Kolmogorov-Smirnov statistic was utilized in all nine cases to test the goodness of fit [4]. As indicated in the references, the Kolmogorov-Smirnov statistic is the maximum difference between the observed cumulative distribution function

TABLE 3: INCOMING TO TREATMENT PLANT

(parts per million)

	DATA SET	N	MEAN	STANDARD DEVIATION	MIN	MAX	COEFFICIENT OF VARIATION
Volatile Solids	DATA1	884	145.89	48.513	64	680	.332
	DATA2	210	129	40.45	68	376	.313
	DATAT	1094	142.7	47.507	64	680	.333
Biochemical Oxygen Demand	DATA1	883	173.67	39.694	94	472	.229
	DATA2	211	139.76	24.891	85	265	.178
	DATAT	1094	167.13	39.619	85	472	.237
Phosphorous (P)	DATA1	868	9.79	2.20	3.0	20.6	.225
	DATA2	191	7.34	1.55	0.0	14.7	.212
	DATAT	1059	9.34	2.30	0.0	20.6	.246

and the assumed cumulative distribution function. If this difference becomes too large, one rejects the null hypothesis - namely that the assumed distribution function does not differ in a significant fashion from the observed data. The procedure followed in all nine cases was the same. The significance level for all cases was taken at five percent.

For example, consider the case of BODREMEFF for DATA2. The MIDAS data analysis computer program was utilized to obtain the descriptive statistics reported in Table 1. MIDAS also prepared a histogram for the (0,1.0) interval in increments of 1/100. The information displayed in this histogram - Figure 1 (for the case of BODREMEFF for DATA2) was utilized to calcu-

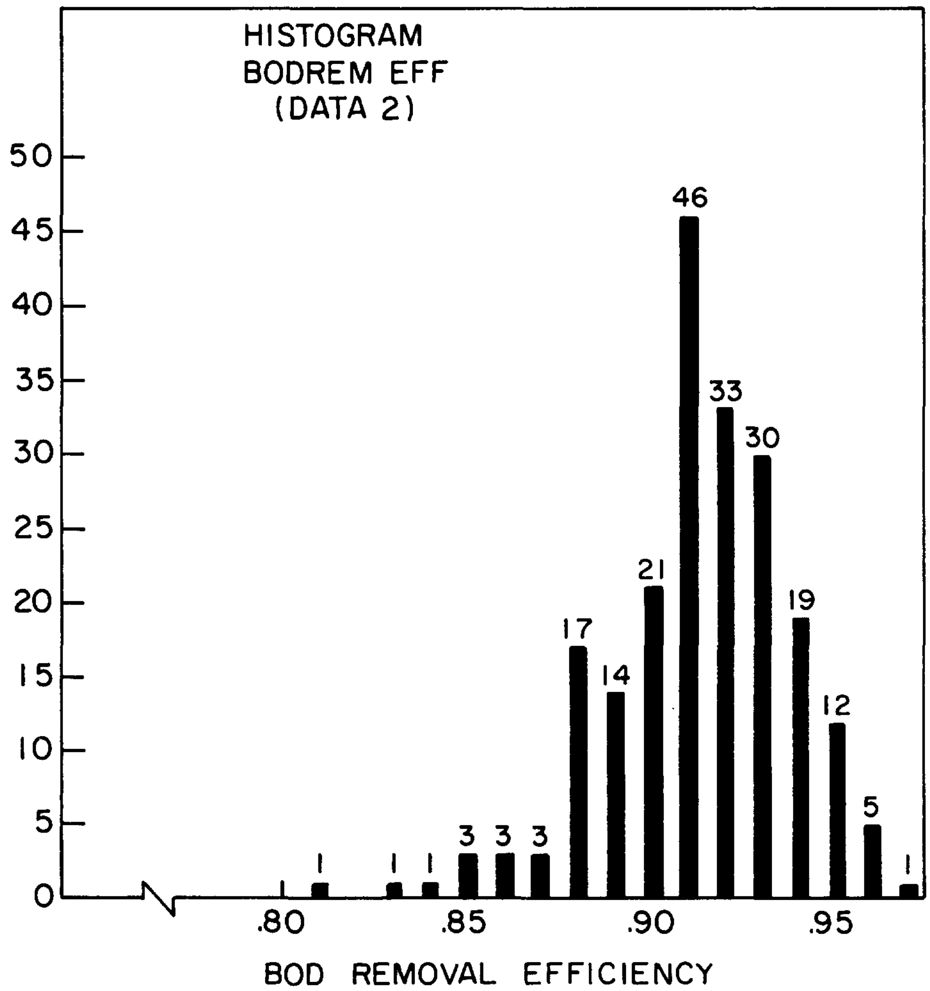

Figure 1

late the observed cumulative frequency distribution. Next, the observed cumulative frequency distribution and the assumed cumulative frequency distribution were plotted in order to determine the maximum (absolute) difference between the two distributions. Figure 2 is a plot of the cumulative frequency distributions for BODREMEFF (DATA2). The maximum observed difference was compared with calculated acceptance limits as a function of significance level (5%) and sample size.

4. RESULTS

Removal Efficiencies

After testing each of the removal efficiency distributions using the Kolmogorov-Smirnov statistic, the null hypothesis is accepted for the following cases:

VSREMEFF (DATA2)

Beta Distribution ($\hat{\gamma}$ = 36; $\hat{\eta}$ = 5.5)

Normal Distribution ($\bar{X}$ = .869; SD = .052)

BODREMEFF (DATA2)

Normal Distribution ($\bar{X}$ = .922; SD = .025)

The null hypothesis was rejected for the remaining case in DATA 2 as well as for all three cases in DATA1.

5. CONCLUSIONS

This research investigation has demonstrated the possibility of developing probability models to represent performance of a secondary activated sludge wastewater treatment plant with regard to removal of volatile solids and biochemical oxygen demand. The results demonstrated that once the chemical process for removing phosphorous had been stabilized the removal efficiencies of both volatile solids and biochemical oxygen demand improved significantly and both could be represented by Normal Distributions. From the analysis performed upon the 1096 cases, the primary factor which appears to be dominant in allowing mathematical representation of removal efficiencies is the stabilization of treatment achieved through phosphorous removal. The precipitation of the phosphate through addition of ferric chloride resulted in significant improvement in sub-

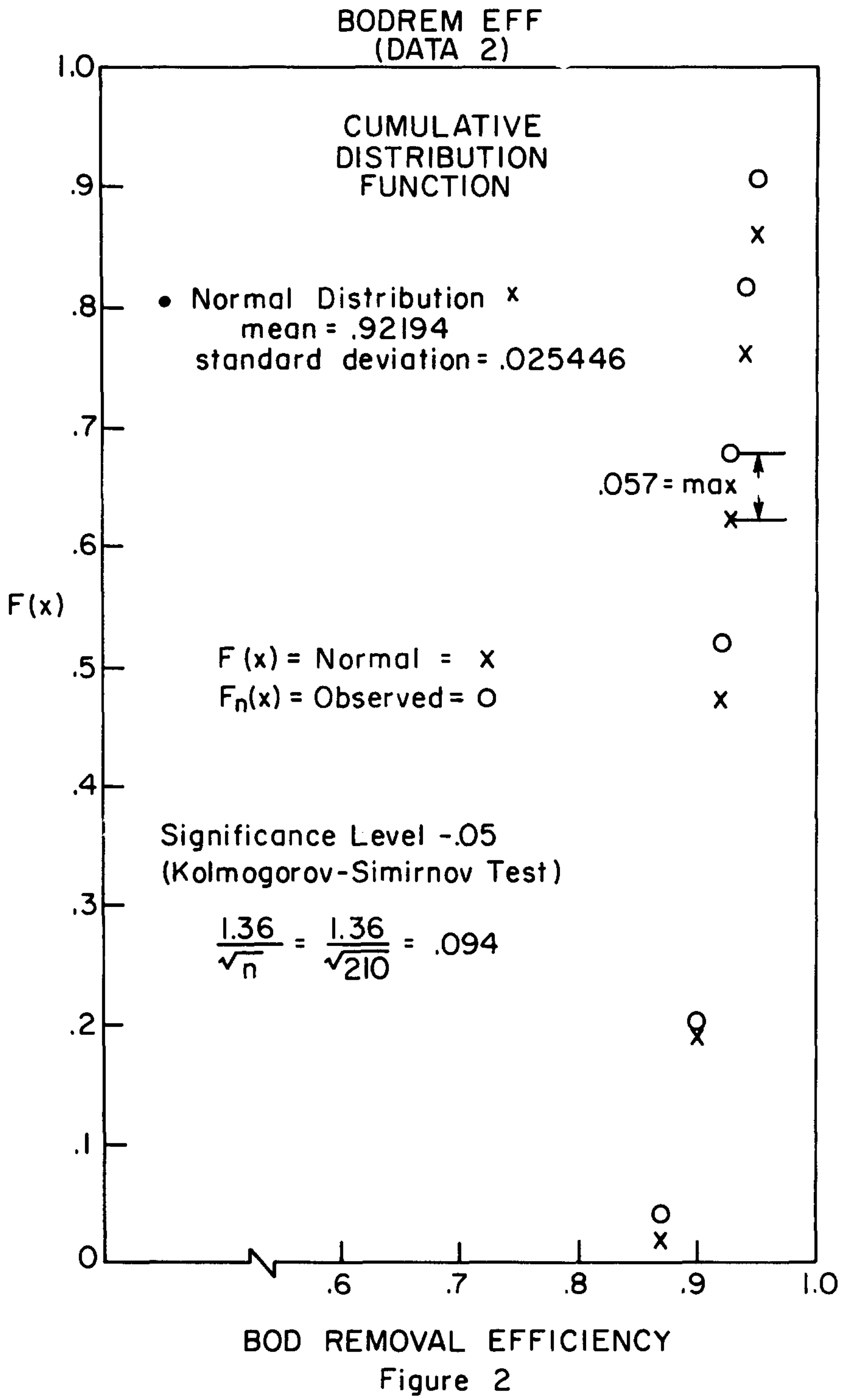
BODREM EFF
(DATA 2)
CUMULATIVE DISTRIBUTION FUNCTION
• Normal Distribution x
mean = .92194
standard deviation = .025446
.057 = max
F(x) = Normal = X
Fn(x) = Observed = O
Significance Level -.05
(Kolmogorov-Simirnov Test)
1.36/√n = 1.36/√210 = .094
F(x)
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
0
.6
.7
.8
.9
1.0
BOD REMOVAL EFFICIENCY

Figure 2

sequent BOD and VS removal in the secondary treatment stage of the plant operation. To date, no parameter variation has been performed in order to find the range differences which are insignificant for the assumed distributions. This is a topic for future research investigation. The findings of this research effort will be of particular value in terms of Monte Carlo simulation of this particular plant into the future. The findings offer an opportunity for further work both in terms of time series analysis and comparative analysis of other wastewater plants to determine whether or not similar results will be observed with other secondary activated sludge wastewater treatment facilities which have incorporated phosphor removal into the treatment process.

Acknowledgement

The author wishes to acknowledge the assistance of Professor Gary M. Fowler, School of Natural Resources, the University of Michigan who provided excellent insight and assistance in the techniques of statistical analysis. In addition I wish to thank Professor Steven Pollock, Director of the Program in Engineering of Public Systems for providing extremely useful criticisms of this paper. Finally, I wish to thank my students at the University of Michigan in the Civil Engineering Systems Course who coded the data for computer analysis. The author is very appreciative of this professional assistance from colleagues and students. Deficiencies in the paper are the sole responsibility of the author.

REFERENCES

1. Thomann, Robert V., "Variability of Waste Treatment Plant Performance", Journal of the Sanitary Engineering Division, Proceedings of the American Society of Civil Engineers, June, 1970 (pp. 819-837).

2. Hahn, G.J. and Shapiro, S.S., Statistical Models in Engineering, John Wiley and Sons, New York, 1967.

3. Pearson, K., Tables of the Incomplete Beta-Function, University Press, Cambridge (England), 1968.

4. a) Lindgren, B.W., Statistical Theory, MacMillan Co., New York, 1963.

   b) Benjamin, J.R. and Cornell, C.A., Probability, Statistics and Decisions for Civil Engineers, McGraw-Hill Book Company, 1970.