United States
Environmental Protection
Agency

Risk Reduction
Engineering Laboratory
Cincinnati, OH 45268

Research and Development

EPA/600/S2-90/014 July 1990

EPA

Project Summary

Composite Sampling for Detection of Coliform Bacteria in Water Supply

Wesley O. Pipes and Harvey A. Minnigh

Coliform bacteria occurring at low densities in treated water can be introduced into distribution systems when they become established and grow in protected habitats such as the interior surfaces of the pipes. Water samples containing organisms from these protected habitats may confuse attempts to use coliform bacteria as indicators of sewage-contaminated drinking water. At the present time, the limit of detection of coliform bacteria in a single sample is about 1 per 100 mL, and methods for detecting them at lower densities are needed to develop information about possible sources of low density contamination.

There are two requirements for detecting coliform bacteria when they occur in water at very low densities: (1) at least one coliform must be captured in a water sample and (2) the sample must be tested in such a manner that the coliform reproduces in the test medium and produces the characteristic reaction by which coliforms are recognized. Some coliform bacteria in potable water samples will not grow in the medium used for testing, and some that grow in the test medium will not produce the characteristic reaction. This project was not concerned with the second aspect of detection, that of recognition of coliform bacteria in water samples, but only with the question of capturing coliform bacteria in a water sample.

The methods of increasing the probability of capturing coliform bacteria in water samples are examining larger volume samples, a larger number of samples, and a composite of a large number of small volume samples. Increasing either the number of samples or the volume examined per sample is merely a matter of increasing the amount of sampling. Clearly, increasing the amount of sampling will increase the probability of capturing any coliform bacteria that may be present, but the pertinent question is how to increase the probability for any predetermined sampling effort. The one possible method is collection and examination of composite samples.

This project was concerned with the use of composite samples to improve the probability of capturing coliform bacteria. Its specific aims were (1) to model mathematically the capture of coliform bacteria from very low densities in a continuous flow of water from a treatment plant, (2) to develop and test equipment and protocols for composite sampling of potable water, (3) to determine the optimum values of the parameters of composite sampling, and (4) to verify the mathematical models in laboratory and field studies.

This Project Summary was developed by EPA's Risk Reduction Engineering Laboratory, Cincinnati, OH, to announce key findings of the research project that is fully documented in a separate report of the same title (see Project Report ordering information at back).

Mathematical Models

When the membrane filter (MF) method is used to detect coliform bacteria, the limit of detection is one organism per 100 mL for an individual grab sample. The results from one sample do not, however, provide much information about the density of coliform bacteria in the body of water that was sampled. If several samples are collected from a relatively large amount of water containing only a few coliform bacteria, the limit of detection is approximate and depends on the number of samples examined and the probability of capturing coliform bacteria in any sample. For any body of water, there is a probability of capturing one or more coliform bacteria in any 100 mL sample, which can be represented at P{+/100 mL}. The probability of capturing coliforms with n samples can be calculated from

$$P\{\text{detection}\} = 1-[1-P\{+/100\ \text{mL}\}]^n.$$

This formula is independent of the frequency distribution of the coliform counts and can be generalized for other sample volumes by substituting P{+/V mL} as the probability of a positive result in a sample of V mL. The manner in which P{+/V mL} changes with a change in V can be calculated if the frequency distribution is known. This information can then be used to calculate the difference in the probability of detecting coliform bacteria by compositing a large number of small volume samples or by using a single large sample having the same total volume.

The frequency distributions used for the mathematical models were the Poisson as a representation of a random dispersion and the lognormal as a representative of aggregated dispersion. These frequency distributions have been used to describe coliform data on large numbers of samples from water distribution systems.

Closed Form Model

If coliform bacteria were ever randomly dispersed in the entire body of water (i.e., the colony counts in unit volumes of water fitting a Poisson distribution), the probability of obtaining a count of x coliform bacteria in a volume, V, would be given by

$$P\{x\} = e^{-m}(m^x/x!)$$

where **m** is the mean colony count in samples of volume **V**. In this hypothetical case, composite sampling will not change the probability of capturing coliforms because as the sample volume is changed, the mean density per sample would change in proportion to the change in the sample volume. The probability of capturing coliform bacteria in V mL is given by

$$P\{+/V\ \text{mL}\} = 1-P\{0\} = 1-e^{-m}$$

where e^{-m} is the Poisson probability of obtaining a zero count because both m^x and $x!$ are 1 when $x = 0$. If V is increased (or decreased) by some factor, then **m** is also increased (or decreased) by the same factor, e.g., an average of 1 organism per 10 L implies an average of 10 organisms per 100 L and so forth. If the total sample volume is fixed, it makes no difference how the sample is collected; one 100 L sample gives the same probability as one thousand 100 mL samples.

A characteristic property of the Poisson distribution is that the variance is equal to the mean. Studies of coliform bacteria in potable water systems have shown that the variance of the coliform counts is much greater than the mean; i.e., the dispersion of bacteria is aggregated rather than random. The lognormal distribution was selected to represent an aggregated dispersion because the computations are more efficient and because it is better known among water works personnel than other frequency distributions that might be used.

A lognormal distribution is described by two parameters, the geometric mean (GM) and the geometric standard deviation (GSD). In the report, x was used to represent coliform density and y, the natural logarithm of the coliform density ($y = 1n\ x$). The arithmetic mean coliform density of the water distribution system was m_x, the mean of the logarithms of the coliform densities was m_y, the variance of the coliform density was s_x^2, and the variance of the logarithms of the coliform densities was s_y^2. From these definitions, $m_y = 1n\ GM$ and $s_y = 1n$ GSD. The probability density function (PDF) for the lognormal is

$$\text{lognormal PDF} = (2\pi s_x^2 x)^{-1/2} \exp -[(1n\ x - m_x)/2(s_x^2)]^2$$

and the arithmetic mean density may be calculated from

$$m_x = \exp [m_y + S_y^2/2].$$

For the samples from any body of water, if the sample volume is increased, the GM increases proportionately but the GSD remains the same. If the values of m_y and s_y are known, the probability of one or more coliforms occurring in an individual sample can be obtained from a table of probabilities associated with th standardized normal (or **Z**) distributio Since the logarithm of 1 is 0, the **Z** valu for one or more bacteria per sample given by $Z = (0-m_x)/s_x$.

Calculations based on the lognorm probabilities for one or more colifor bacteria in a given volume of water sho that compositing a large number of sm volume samples gives a higher prob bility of capturing one or more coliform than does an equal volume of wat collected as a single large sample. Sinc coliform bacteria are not randomly di persed in water, multiple discre samples always provide a higher prob bility of finding them, if they are preser than does a single, large sample equivalent volume.

Simulation Model

To test the conclusions of the close form model, a series of Monte-Carlo typ simulations were carried out. A series normally distributed random numbe was needed with the mean and varianc selected. A pseudorandom number ge erator was used to create the series random numbers with the mean ar variance input for each simulate distribution. The numbers were interpre ed as the logarithms of the colifor densities so that the antilog of the mea was the GM and the antilog of th standard deviation was the GSD. The G was selected over a range betwee 0.000335 and 0.00000000206, and th GSD, between 10 and 120.

The antilogs of the numbers generate with the program were treated as colifor densities for the water passing th sampling point during the samplin period. The basic question for th simulations was: To what degree doe composite sampling increase the prob bility of coliform detection? Gra samples were simulated in the same run and for the same data sets as continuou portions to verify the operation of th simulation program, since these prob bilities can be estimated with absolut accuracy for known parameters for given distribution. The question of wh volume of composite sample to collect really a question of what level contamination is the maximum acce table for the sampled stream; i.e., th volume that provides, even with n coliforms present in the sample, 95 certainty that the water sampled ha fewer than that permissible maximum.

An important result of using th lognormal distribution is that som portions will have very large colifor densities, and these portions are cruci

for estimating the sampled distribution. For instance, a lognormal distribution with GM = 0.00000112 and GSD = 100 has an arithmetic mean density lower than another lognormal distribution with GM = 0.00112 and GSD = 20; however, it can produce samples with densities of one or more coliforms per sample precisely because, with its greater variance, it is more likely to produce samples with a very high density. As an example, for the GM - GSD pairs cited above, the arithmetic mean and variance and the probability of a positive sample are shown in Table 1.

GSD values reported for water distribution systems ranged from about 10 to 100. When the GSD is low, composite sampling gives little advantage for finding coliform bacteria. For example, a system with a GM of 0.1 and a GSD of 10 has an arithmetic mean density of 1.42/100 mL, and composite sampling will provide little advantage when compared with grab sampling since the probability of a positive 100 mL sample is 0.16. The more challenging case is the system that has the same arithmetic mean but much higher variability. A system with a GM of 0.000025 and a GSD of 100 also has an arithmetic mean density of about 1 per 100 mL, but very few 100 mL aliquots (about 1 in 100) will have any coliforms at all. In this case, composite sampling would be the method of choice.

Summary on Mathematical Models

It was shown by both closed form models and computer simulations that composite sampling from a lognormal distribution is superior to grab sampling, if the variance of the coliform count is large. The larger the variance, the greater the advantage of composite sampling for capturing coliform bacteria. When, however, the value of the variance approaches the mean density, composite sampling does not provide any great advantage for detecting coliform bacteria. The agreement of the simulated results with the predictions of the closed form model provides some assurance of the reliability of the conclusions.

Laboratory Study

The laboratory study included controlled experiments for testing hypotheses about the effectiveness of composite sampling for capturing bacteria entering or in a water distribution system. Evidence of the efficacy and efficiency of composite sampling as compared with grab sampling was developed by means of these experiments. The laboratory results turned out to be the primary verification of the results from the mathematical models, since few coliform bacteria were found in either grab or composite samples during the field study.

Methods

The criteria that governed selection of a sampler were that it should (1) be easy to clean, (2) be constructed so that the parts coming in contact with the water could be sterilized or sanitized, (3) require only standard fittings and fixtures to install, and (4) be portable and easy to set up. The intent was to select a stock or modified commercial sampler; however, none of the suppliers queried had or knew of such a stock sampler. The Sigma Model 6301* (American Sigma PO Box 300, Middleport, NY 14105-0300) was selected largely because of its ability to index through a series of sample containers, which would provide sequential composite samples. Additional samplers were also constructed consisting of variable rate peristaltic pumps and polypropylene sample containers (Nalgene 2319 series).

To produce large variances of the coliform densities, a syringe pump driven by a very accurate timer was selected to deliver the inoculum. The composite samples were collected by means of a peristaltic pump operated at a rate selected to provide the sample size desired. Tubing and pump heads were selected to provide residence times in the experimental system consistent with residence times expected or experienced in field applications.

The laboratory composite sampling systems were modified from time to time as experimental requirements made additional controls necessary. Composite samples were collected with variable rate peristaltic pumps set to collect the desired volume. Volumes collected for the composite samples ranged between 0.1 and 4.0 L/hr.

All composite samples were analyzed in their entirety. Intermittent inoculation was used to simulate the occurrence of coliform bacteria in a real distribution system; i.e., relatively high densities were injected over relatively short periods of time and relatively small portions of the flow contained most of the bacteria. Both *Escherichia coli* (ATCC 8739) and *Enterobacter cloacae* (ATCC 13047) were used in the laboratory experiments, and there were no differences in the results obtained using the two different organisms. The inoculum density was changed by using different dilutions of the initial culture; using different inoculation rates; varying the flow rate; and, to a much lesser extent, varying the periodicity of inoculation.

Results

The laboratory experiments were designed to show that (1) the recovery of coliform bacteria by composite samples from flows of water with known dispersions, in which very little of the water under test contained all the coliforms, were not significantly different from the predicted frequency of occurrence, (2) the threshold for coliform detection was generally in the range predicted, and (3) the mean coliform densities of the samples were not significantly different from those calculated from the inoculum density. The control of the dispersions of bacteria in the flow of water to mimic the occurrence of coliform bacteria in actual water systems was a central theme of these experiments.

An agreement was obtained between the theoretical frequency of occurrence of various densities and the actual densities obtained for all overnight (18 to 24 hr) composite sampling runs. These results indicated that the composite sample densities were substantially in agreement with the densities that were introduced into the experimental system by the syringe pump. The probability of capturing a coliform was directly related to the highest density in the flow of water sampled. The 95% probability of capturing a single coliform in the composite was the 95% probability of a single coliform occurring in the sample portion given the sampled stream density.

The data showed that composite sampling provided equal or superior performance for times of both high and low probability of capturing coliform bacteria and hence, for both levels coliform occurrence. All experiments accepted for further analysis were from systems tested for and found free of coliform contamination. Bacterial densities in composite samples were consistent with stream densities calculated from inoculum densities and from the volume of flow for all experiments. Experiments designed to test the effect of variability on the frequency of positive sample results followed predictions of the closed form

*Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Table 1. *Effect of Variability of Coliform Counts on Arithmetic Parameters*

GS	GSD	Arithmetic		Probability of	
		Mean	Variance	Any Positive Result	Result >1
0.00112	10	0.0159	0.0503	0.0016	0.0015
	50	2.36	2.46×10^{7}	0.0412	0.0381
	100	45.12	3.305×10^{12}	0.0700	0.063
0.00000112	10	0.00002	5.03×10^{-8}	1×10^{-9}	1×10^{-9}
	50	0.00236	24.6182	0.0002	0.0002
	100	0.04512	3.305×10^{6}	0.0015	0.0014

and simulation models; i.e., that, for any given mean density, higher variability will provide a greater frequency of positive portions. Further, the frequency of positive portions was directly related to the maximum stream density for any given size of aliquot. Finally, experiments directly comparing composite and grab sampling from the same stream conclusively demonstrated the superiority of composite sampling for capturing coliform bacteria.

Field Studies

Field sampling results were obtained from waters from two different distribution systems, West Chester and Downingtown. Both of these systems are in Chester County, Pennsylvania, about 30 miles west of Philadelphia. Neither have a record of violations of the microbiological maximum contaminant level. The objectives of the field sampling were to test the use of the composite sampler in realistic situations, to attempt to obtain further verification of the mathematical model and to compare grab samples with the composite samples.

Methods

The composite sampling setups for the field studies were identical to those used in the laboratory. The West Chester sites were both at system pressure, approximately 160 to 180 psi at the plant and 140 to 160 psi at the tank. The Downingtown sampler was not used on a pressurized line but drew from the top of the filters and from the filter effluent. Composite and grab samples were collected over 18- to 24-hr periods and protected from sunlight during transport in insulated containers with artificial coolant packs. All microbiological procedures were completed within 6 hr of collection.

Each sample, composite and grab, was tested for total coliform bacteria and heterotrophic plate count. Large aliquots could be filtered by the MF procedure for all the finished water samples and both 300 and 500 mL portions were routinely filtered with no evident reduction in filtrate flow due to occlusion or matting of filter surfaces. In addition to the microbiological analyses, each sample source was tested each day for free and total chlorine, temperature, and pH. The turbidity of all except West Chester tank samples was also determined each sampling day.

West Chester Results

The East Branch of the Brandywine Creek was the source of raw water during the period of sampling. The raw water had very high densities of coliform bacteria from agricultural land runoff and wastewater discharges. Treatment consisted of prechlorination, flocculation, settling, rapid sand filtration, and post chlorination. Composite samples were concurrently taken of the treated water (the high pressure side of one of the high service pumps) and the flow to the storage tank. Overall, 36 sets of 2 composite samples and 5 individual composite samples (77 total composite samples) were obtained from this system from June to November 1987. In addition, grab samples of the raw, finished, and distribution system water were collected. The distribution system samples were obtained from both private residences and fire hydrants.

Only one of the composite samples, which was obtained from the tank site, had coliform bacteria. Four 500-mL portions were positive, three with one organism each and one with two. This gave a sample density of 0.00098 per mL for the approximately 2-gal composite, thereby demonstrating the ability of composite sampling to detect coliform bacteria at low densities. Three of 104 grab samples of the finished water and 2 of 96 grab samples from private residences had coliform bacteria. More coliform bacteria were detected in the grab samples than in the composite samples. Overall coliform densities were much lower than 1.0 per 100 mL, a these results were probably just happe stance.

Dowingtown Results

Two water sources, Copeland Run a Beaver Creek, supplied about 30% a 70% of the water treated, respectivel There were approximately 8 miles mains, most 6 or 8 in., and 2,100 servi connections. The system had three st age sites, a **4-MG** open storage reserv at the plant and two 2-**MG** steel grou storage tanks on elevations north a south of the Borough proper.

A history of poor turbidity removal w one reason for selecting this plant f sampling. There was no provision f continuous sludge removal. Filter influe turbidities were sometimes very high, a it was hoped that some coliform positi composite samples would result from th sampling series despite the 2 mg/L fr chlorine residuals maintained through t process.

The composite samples of the influe were analyzed in full. These data show that composite sampling can provide da at least equivalent to even the stricte grab sampling regimen. In these expe ments, grab samples were taken alte nately from the top of the filters (ha dipped) and the effluent sampling port a rate of between 8 and 20 per hr f each sampling location. Composite sa plers were operated at a rate around 3 mL/hr, which provided a 1.5 gal 24- composite sample.

The results of the field sampling d not achieve the objectives of this part the project. The occurrence of coliform in the plant effluent in these systems w not high enough to provide the da needed. Unfortunately, the samples wi coliform bacteria present came mo often from the grab samples than fro the composite sampler. This does n prove that the theoretical model or t laboratory results are incorrect, only th the systems selected for the fie sampling did not provide the conditio

needed for the composite sampler performance to be demonstrated.

Conclusions

1. If coliform bacteria were ever randomly dispersed in a body of water (i.e., the variance of the count per unit volume is equal to the mean density per unit volume), the probability of capturing coliform bacteria would be proportional to the total volume of water collected and would not depend on the total number of samples or how the samples were collected. A random dispersion of bacteria in a body of water can be achieved by complete mixing of the water; this, of course, is very unlikely to occur in a water distribution system.
2. For water supply systems in which the coliform bacteria are not randomly distributed but show an aggregated arrangement (i.e., one in which the variance of the counts per unit volume is greater than the mean density), composite sampling is more effective than grab sampling in capturing coliform bacteria in any volume of water tested. This conclusion is based on a closed form mathematical model, computer simulation of sampling from a lognormal distribution, and a laboratory study of composite sampling.
3. The results of sampling municipal water systems during this project did not demonstrate the superiority of composite sampling for collection of coliforms under field conditions.
4. The commercially available composite samplers could not be adapted for truly continuous, aseptic sampling of a stream treated water under pressure. The best that could be achieved was very frequent intermittent collection of very small volumes of water.
5. For intermittent composite sampling of a stream of water from a treatment plant, the probability of capturing coliform bacteria in any given sample volume is a function of the volume of sample tested, the size of the stream of water sampled, and the frequency of collecting the component aliquots.

The full report was submitted in fulfillment of Cooperative Agreement No. CR 813337 by Drexel University under the sponsorship of the U.S. Environmental Protection Agency.

Wesley O. Pipes and Harvey A. Minnigh are with Drexel University, Philadelphia, PA 19104.
***Donald J. Reasoner** is the EPA Project Officer (see below).*
The complete report, entitled "Composite Sampling for Detection of Coliform Bacteria in Water Supply," (Order No. PB 90 192-758/AS; Cost: $17.00, subject to change) will be available only from:
National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
Telephone: 703-487-4650
The EPA Project Officer can be contacted at:
Risk Reduction Engineering Laboratory
U.S. Environmental Protection Agency
Cincinnati, OH 45268

United States
Environmental Protection
Agency

Center for Environmental Research
Information
Cincinnati OH 45268

Official Business
Penalty for Private Use $300

EPA/600/S2-90/014