Septic Tank Setback Distances: A Way to Minimize Virus Contamination of Drinking Water\textsuperscript{a}

by Marylynn V. Yates\textsuperscript{b} and S. R. Yates\textsuperscript{c}

ABSTRACT

Septic tanks are the most frequently reported causes of contamination in ground-water disease outbreaks associated with the consumption of untreated ground water in the United States. The placement of septic tanks is generally controlled by county-wide or state-wide regulations, with little consideration given to the local hydrogeologic, climatic, and land-use conditions. Using the travel time necessary to achieve a seven-order-of-magnitude reduction in virus number as the criterion, a wide range of septic tank setback distances (from less than 15 m to greater than 300 m) were calculated for a part of the Tucson Basin. This study uses 3-D disjunctive kriging to calculate the conditional probabilities associated with the setback distance estimates. The results are presented in two different ways: first, given a setback distance (e.g., prescribed by law) the probabilities that the level of viruses will be within acceptable limits are calculated; and second, the desired probability level is specified (e.g., 90\%) and the setback distances required to achieve that level of confidence that the water will be free of virus contamination are calculated. The methods have potential for use by local government officials for land-use planning purposes.

INTRODUCTION

In 1980, it was estimated that there were 22 million septic tanks in the United States, serving approximately one-third of the population (U.S. EPA, 1986). Septic tanks contribute more than one trillion gallons of waste to the subsurface every year (OTA, 1984); this waste is the most frequently reported cause of ground-water contamination (U.S. EPA, 1977). In addition, the overflow or seepage of sewage, primarily from septic tanks and cesspools, was responsible for 38\% of the outbreaks and 58\% of the cases of illness caused by the use of contaminated, untreated well water from 1971 to 1980 (Craun, 1986a).

Several potentially harmful chemical substances may be present in domestic waste water, including heavy metals (from pigments in cosmetics), toxic organic chemicals (from cleaners), and nitrates. A review of the occurrence of harmful chemical substances in household waste water is provided by Viraraghavan and Hashem (1986), and the role of these compounds in disease production is reviewed by Craun (1986b).

In addition to chemical contaminants, septic tank effluent may contain potentially infectious microorganisms, including bacteria, parasites, and viruses. Pathogenic microorganisms which have been found in domestic waste water include \textit{Salmonella}, \textit{Shigella}, \textit{Entamoeba}, \textit{Giardia}, hepatitis A virus, rotavirus, and poliovirus. Yates (1985) reviewed the ground-water-borne disease outbreaks caused by microorganisms present in septic tank effluent.

In the past, most government agencies have regulated septic tank placement by requiring minimum setback distances between septic tanks and drinking-water wells (Perkins, 1984). Setback distances range from 15 to 91 meters, with typical values averaging 15 to 30 m (Plew, 1977). These setback distances are generally imposed over at least a county-wide area, with little consideration given to the local geology, hydrology, and meteorology. Numerous studies have shown that microorganisms can travel considerable distances in the subsurface; these have been reviewed by Yates and Yates (1988). Viruses, in particular, due to their small size (20 to 200 nm) and long survival times, can migrate very large distances in soil and ground water; as much as 1600 m have been reported for certain viruses in karst terrain (Gerba, 1984b) and up to 400 m in sandy soil (Keswick and Gerba, 1980).

Indirect evidence of microorganism movement in the subsurface has been obtained from water-borne disease outbreaks in which it was shown using dye tracers that a septic tank was the source of well contamination (Craun, 1979). There have, however, been studies that directly examined the movement of viruses from a septic tank to a ground-water well (Hain and O’Brien, 1979; Stramer, 1984; and Vaughn \textit{et al.}, 1983). These

\textsuperscript{a}The research described in this article was supported in part by the U.S. Environmental Protection Agency through an interagency cooperative research agreement, no. DW-129326310-01-0.
\textsuperscript{b}Department of Soil and Environmental Sciences University of California, Riverside, California 92521.
\textsuperscript{c}U.S. Salinity Laboratory, Riverside, California 92501.

Received March 1988, revised July 1988, accepted July 1988.
Discussion open until September 1, 1989.
studies showed that viruses could move as far as 65 m from the septic tank (Vaughn et al., 1983) and persist for up to 131 days in the ground water (Stramer, 1984). Based on the results of these studies, it becomes clear that a setback distance of 15 to 30 m may not be adequate to prevent viral contamination of ground water, and possibly waterborne disease outbreaks, under certain environmental conditions.

Previous studies (Yates et al., 1986; Yates and Yates, 1987) demonstrated the variability in septic tank setback distances over a city-wide area that resulted when hydrogeologic variables and virus inactivation rates at each well were used in the setback calculations. In the present study, in addition to calculating the septic tank setback distances, the probabilities that these setback distances are adequate to protect ground water from viral contamination were calculated using disjunctive kriging, a nonlinear estimation technique. Two situations were considered: (1) Given a setback distance (e.g., specified by regulation), what is the probability that this would be adequate to protect the ground water from viral contamination at different locations in the city? and (2) Given a desired probability level, what setback distance would be necessary to be that confident that the ground water would be protected from contamination by viruses?

METHODS
Calculation of Setback Distances

Setback distances between septic tanks and drinking-water wells were calculated using the simplest model available, a modified form of Darcy’s law (Freeze and Cherry, 1979):

\[ D = \frac{(tK)}{n_e} \]  

where \( D \) is the setback distance (m); \( t \) is the travel time (d); \( K \) is the hydraulic conductivity (m d\(^{-1}\)); \( i \) is the hydraulic gradient (m m\(^{-1}\)); and \( n_e \) is the effective porosity of the aquifer.

Virus inactivation rates were determined experimentally using the ground water obtained from 71 pumping municipal drinking-water-supply wells (Yates et al., 1986). Sample locations are shown in Figure 1. Travel times were calculated for each sample location using the virus inactivation rates and are based on the amount of time required to achieve a seven-order-of-magnitude reduction in virus numbers. The seven-order-of-magnitude reduction in virus numbers was chosen for the following reason: the World Health Organization (WHO) has recommended that there be no viruses detectable in 1000 liters of water (Gerba, 1984a).

If we assume that 10 viruses per ml (10\(^6\) per liter) of septic tank effluent travel through the soil and reach ground water, then a decrease of seven orders of magnitude in virus numbers would be required to approach the WHO’s recommendation of zero viruses per 1000 liters (there would actually be one virus per 1000 liters if a seven-order-of-magnitude reduction occurred).

Hydraulic conductivity values were calculated for each sample location based on transmissivity values at each location provided by the State of Arizona Department of Water Resources (the aquifer thickness was assumed to be constant). Hydraulic gradients at each sample site were calculated from a water-table elevation map obtained from the City of Tucson.

Geostatistical Analyses

The techniques used to estimate the setback distances and conditional probabilities are referred to as geostatistical methods. This field of statistics, unlike classical statistics, assumes that the samples are not independent of one another, and that the value of a variable at one sample location is related to its value at another location. Kriging is a geostatistical technique which allows one to estimate the value of a variable at an unsampled location using known values at nearby locations, based on a linear weighted averaging. A complete discussion of geostatistical techniques can be found in Journel and Huijbregts (1978). Disjunctive kriging, which is a nonlinear technique, also allows one to calculate a value at an unsampled location using surrounding known values. In addition, the conditional probability that the estimated value is greater than a prescribed cutoff level is calculated.

Fig. 1. Sample collection sites. Origin is Arizona quadrant D, township 14S, range 14E, section 19, CCC.
The Disjunctive Kriging Estimator

Detailed derivations of the disjunctive kriging estimator and the conditional probability can be found in Matheron (1976), Journal and Huijbergs (1978), Yates et al. (1986a), and Yates (1986); therefore, only a brief description and salient results will be given here. A FORTRAN computer program which calculates the disjunctive kriging estimator and conditional probability can be found in Yates et al. (1986b).

Consider a second-order stationary random function $Z(x)$ which has been sampled over two dimensions at $N$ locations: $x_1$, $x_2$, ..., $x_N$. It is assumed that $Z(x)$ is spatially correlated and this correlation can be described by a semivariogram under a second-order stationarity hypothesis, which states that the spatial correlation function for $Z(x)$ exists and does not depend on position. The linear ordinary kriging estimator, $Z_{ok}(x_0)$, has the form

$$Z_{ok}(x_0) = \sum_{i=1}^{n} \lambda_i Z(x_i)$$

where $Z(x_i)$ is the measured value of $Z(x)$ at location $x_i$, and $\lambda_i$ is a kriging weight. Equation (2) can be considered a special case of a more general and nonlinear disjunctive estimator, $Z_{dk}(x_0)$. The disjunctive kriging estimator has the form

$$Z_{dk}(x_0) = \sum_{i=1}^{n} f_i [Z(x_i)]$$

where $f_i$ is an unknown function corresponding to the data value at location $x_i$. The linear and disjunctive kriging estimators are equivalent when the unknown functions, $f_i$, are linear functions of $Z(x_i)$.

To use the disjunctive kriging method, a transformed variable, $Y(x_i)$, must be determined. It is assumed that the transformed data values have a univariate and bivariate normal distribution. The transform relationship, $\phi[Y(x)] = Z(x)$, between $Z(x_i)$ and $Y(x_i)$ is written in terms of a Hermite polynomial with coefficients, $C_k$

$$\phi[Y(x)] = Z(x) = \sum_{k=0}^{\infty} C_k H_k[Y(x)]$$

The Hermite coefficients are found using numerical integration, the properties of orthogonality (see Yates et al., 1986a) and the data values. In an analogous manner to ordinary kriging, an unbiased estimator with minimum variance is sought, that is,

$$E[Z(x_0) - Z^*(x_0)] = 0$$

$$\text{Var}[Z(x_0) - Z^*(x_0)] = \text{min}$$

The solution for the unknown functions, $f_i$, in equation (3) requires several steps. First, the original data must be transformed, using equation (4), into a new variable, $Y(x)$, that has a standard normal distribution where pairs of values are bivariate normal. This step provides values for the $C_k$'s. Next, the linear kriging system is solved "K" times and provides the kriging weights, $b_{ik}$, in equation (6) which are used to calculate an estimate of the Hermite polynomial, $H_k^*(x_0)$, in equation (7) at the estimation site from values of the Hermite polynomials at the sample locations. Once the K values for $H_k^*(x_0)$ have been obtained, they are used in equation (8) along with the appropriate value of $C_k$ to give the disjunctive kriging estimate

$$\sum_{i=1}^{n} b_{ik} (\rho_{ij})^k = (\rho_{oi})^k; \quad j = 1, 2, 3, \ldots, n \quad (6)$$

$$H_k^*(x_0) = \sum_{i=1}^{n} b_{ik} H_k[Y(x_i)] \quad (7)$$

$$Z_{dk}(x_0) = \sum_{k=0}^{K} C_k H_k^*[Y(x_0)] \quad (8)$$

where $\rho_{ij}$ is the autocorrelation function, $\rho(x_i - x_j)$.

The disjunctive kriging estimator uses the autocorrelation function for determining the kriging weights in equation (6); therefore, second-order stationarity conditions are required so that the variance exists. In this case the autocorrelation function can be written in terms of the semivariogram

$$\rho(h) = 1 - \gamma(h)/\gamma(\infty) \quad (9)$$

where $\rho(h)$ is the autocorrelation function; $\gamma(h)$ is the semivariogram; $\gamma(\infty)$ is the sill value of the semivariogram; and $h$ is the distance vector $(x_i - x_j)$.

The estimation variance for disjunctive kriging, $\text{var}[Z(x) - Z_{dk}(x)]$, is

$$\sigma_{dk}^2 = \sum_{k=1}^{K} k! C_k^2 \left[1 - \sum_{i=1}^{n} b_{ik} (\rho_{oi})^k \right] \quad (10)$$

The Conditional Probability

One important advantage the disjunctive kriging method has over ordinary kriging is that an estimate of the conditional probability that the value at an estimation site is less (or greater) than an arbitrary critical value, $y_c$, can be calculated. This is shown diagrammatically in Figure 2 where the conditional probability is the shaded area under the curve and to the left of the critical value, $y_c$.
Prob(x), that the value of a property at $x_o$ is less than a specified cutoff level, $y_c$ [note that $y_c$ is the transformed cutoff level; see equation (4)], is

$$\text{Prob}(x_o) = G(y_c) - g(y_c) \sum_{k=1}^{K} H_{k-1}(y_c)H_k^*[Y(x_o)]/k!$$

where $G(y_c)$ and $g(y_c)$ are the cumulative and probability density functions, respectively, for a standard normal variable; and $H_k^*[Y(x_o)]$ is found using equation (7). The estimated conditional probability density function, Pdf(u), can be obtained from equation (11) by taking the derivative with respect to $y_c$

$$\text{Pdf}(x_o) = g(u) \{1 + \sum_{k=1}^{K} H_k(u)H_k^*[Y(x_o)]/k!\}$$

**Estimation of Setback Distances**

Since there is a unique relationship between the conditional probability and the critical setback distance (i.e., the cutoff value), the relationship can be inverted. Taking the inverted relationship allows the critical setback distance given a specified conditional probability level to be obtained. This method for presentation of the results has an advantage in that the actual minimum setback distance can be given to assure a desired probability level; because it is more likely that the probability level (of safety) will be known, this gives the minimum setback distance more directly. In other words, it is not necessary to calculate a series of cutoff levels to find the one that gives the desired probability.

**RESULTS AND DISCUSSION**

The experimental semivariogram resulting from the calculated septic tank setback distance was modeled using a spherical equation with a nugget of 24.0, a sill of 389.0, and a range of 3.0 km (Figure 3). A contour map of the septic tank setback distances estimated by disjunctive kriging of the calculated setback distances at each well is shown in Figure 4. The values range from less than 15 m to over 75 m, with the higher setbacks generally located in the north-central area of the map. The location of the higher setback distances corresponds to an area of high transmissivity, where the wells are adjacent to an intermittent stream. The virus inactivation rates in these wells are low due to the relatively cool temperature (19°C) of the ground water in this area. [Groundwater temperature has a significant positive correlation with virus inactivation rates (Yates et al., 1985).] The combination of low virus inactivation rates (which makes t, the time for seven orders of magnitude reduction in virus number, large) and high transmissivity (which, owing to the fairly
uniform thickness of the aquifer, makes the hydraulic conductivity large) results in the calculation of large distances to minimize the possibility of viruses being present in drinking water.

Case 1: Probabilities Associated with Specified Setback Distances

Probability maps were calculated for two setback distances for comparative purposes. The probabilities estimated for a specified setback of 15 m are shown in Figure 5A. Comparing the 15-m contours in Figure 4 with the corresponding contours in Figure 5A, one can see that these contours have a probability of 0.70. In other words, if a 15-m cutoff level is specified, there is a 70% probability that 15 m would be adequate to prevent virus contamination of ground water at the 15-m contours.

Figure 5B shows the probability contour map calculated using a 30-m cutoff level. Looking at the 15-m contours on Figure 4 once again, and comparing them with the corresponding contours on Figure 5B, the contours now have a 0.85 probability. This is because, at this location, we had estimated that 15 m would be adequate to protect the ground water from contamination. Now we have imposed a 30-m setback at this location. It follows that the probability of 30 m being adequate is higher (85%) as compared with the probability estimated for 15 m (70%).

Case 2: Setback Distances Associated with Specified Probabilities

In this case, rather than specifying a setback distance and calculating the associated probabilities, the desired probability level is specified and the associated setback distances are calculated. In the first example, a probability level of 0.9 was specified. In other words, what setback distance is necessary to be 90% certain that the actual setback distance is less than or equal to that distance? Comparing the 15-m contours of Figure 4 with those roughly corresponding on Figure 6A, it can be seen that a 40-m setback distance would be required to be 90% certain that the ground water would be adequately protected from virus contamination. If one wanted to be 99% certain that the setback was adequate to prevent viral contamination, an 80-m setback distance would be required in those locations where 15-m distances were calculated (Figure 6B).

These methods of calculating septic tank setback distances have potential for use as management decision-making aids in regulating septic tank placement in a community. Although a very simple model was used here for illustrative purposes, it is expected that if the proposed techniques were to be used by a municipality, a more comprehensive ground-water travel time model would be used to calculate setback distances. Vertical transport through the unsaturated zone and the presence of pumping wells are among the factors which would have to be considered.

To demonstrate the effect of adding pumping wells to the regional ground-water flow in the travel time calculations, a simple one-well case was used. The well chosen is pumped at a rate of $9.46 \times 10^3$ m$^3$ sec$^{-1}$ (150 gpm). In the disjunctive kriging calculation, in which only regional ground-water flow was used in the setback distance calculation, this well was located on a 60-m contour (Figure 4). When the $9.46 \times 10^3$ m$^3$ sec$^{-1}$ pumping rate is added to the travel time calculation, a setback distance of 156 m is required to achieve a
seven-order-of-magnitude reduction in virus number (Figure 7). If only four orders of magnitude of virus inactivation are required, the setback distance would be 93 m, which is 1.5 times greater than that calculated without adding the effects of pumping. The actual calculations would be more complicated than described here, as the effects of all of the wells pumping would have to be included to get an accurate picture of the flow field in the Basin. This simple example does show, however, that pumping has a large impact on the travel time, and thus setback distance calculations, and must be considered if the method is to be used for municipal planning purposes.

With the appropriate modifications to model the specific situation of interest, the methods could be used for community planning purposes. The first case described, namely calculating the conditional probabilities given a specified cutoff level, would be useful in a situation where the minimum setback distance was specified by regulation. For example, a certain community has a regulation stating that 30 m is the minimum separation between a well and a septic tank. Disjunctive kriging could be used to generate a conditional probability contour map. A decision to allow a septic tank to be placed in a certain location could then be based on the calculated probabilities. For example, it might be decided that if the probability was 75% or greater, a septic tank would be permitted on any lot, provided that soil percolation test requirements were met. If the probability was between 50% and 75%, soil percolation test requirements could be made more stringent or the minimum lot size could be increased in order for a septic tank permit to be issued. If the probability was less than 50%, it might be decided that septic tanks would not be allowed at all.

The approach described in the second case could also be used for community planning purposes, in that a desired probability level could be specified (e.g., in a regulation), and the setback distances necessary to achieve that level would be calculated. One advantage of using this method is that the implicit assumption that the hydrogeologic characteristics of the area are constant would be avoided. The regulations would only have to specify a probability level to be met in order to allow a septic tank permit.

Models of this type will become more important in the future, especially in light of the proposed maximum contaminant level goal (MCLG)
for viruses in ground water. In November 1985, the U.S. Environmental Protection Agency proposed an MCLG of 0 viruses in drinking water. Rather than require monitoring all drinking water for the presence of viruses (as is done for coliform bacteria), it was proposed that all ground waters must be disinfected prior to distribution. It was anticipated that variances from the mandatory disinfection requirement might be granted if it could be shown that it is unlikely that viruses could contaminate the drinking water. It has been suggested that if it could be shown, using a model, that the travel time of domestic waste from the source to the well will result in an eight- to ten-order-of-magnitude reduction in virus numbers, a variance could be granted (Gerba, 1984a). In this model, the setback distance was calculated based on a seven-order-of-magnitude reduction in virus numbers, although this could easily be modified for any amount of virus inactivation.

REFERENCES


Gerba, C. P. 1984a. Strategies for the Control of Viruses in Drinking Water. American Association for the Advancement of Science, Washington, DC.


Marylyn V. Yates received a Ph.D. in Microbiology from the University of Arizona in Tucson. She also has an M.S. in Chemistry from New Mexico Institute of Mining and Technology and a B.S. in Nursing from the University of Wisconsin-Madison. Currently, she is a Hydrologist/Toxicologist in the Department of Soil and Environmental Sciences at the University of California in Riverside.

Scott R. Yates received his B.S. in Geology from the University of Wisconsin in 1980, M.S. in Hydrology from New Mexico Institute of Mining and Technology in 1982, and Ph.D. in Soil and Water Science from the University of Arizona in 1985. He is currently a Soil Scientist for the USDA/ARS at the U.S. Salinity Laboratory in Riverside, California. His research interests include soil physics, modeling the fate and transport of contaminants in variably saturated media, and geostatistics.

208

NATIONAL WATER WELL ASSN.
LIBRARY