THE SOCIETY OF NAVAL ARCHITECTS AND MARINE ENGINEERS
601 Pavonia Avenue, Suite 400, Jersey City, New Jersey 07306 USA
Paper presented at the Marine Structural Inspection, Maintenance, and Monitoring Symposium
Sheraton National Hotel, Arlington, Virginia, March 18-19, 1991

A Probabilistic Approach for Determining the Effect of Corrosion on the Life Expectancy of Marine Structures

G.J. White, U.S. Naval Academy, Annapolis, Maryland
B.M. Ayyub, University of Maryland, College Park, Maryland

ABSTRACT

This paper looks at a means of evaluating the effect of corrosion on the estimated service-life of marine structures. Both the mean value and standard deviation of the corrosion rate are treated as random variables. Estimates for these values are developed using semivariogram analysis with kriging estimation. These values are then included in an extreme-value analysis by using Monte Carlo simulation to determine the likelihood of a particular level of wastage in a given time period. As the structure ages and undergoes in-service inspections, the estimators for the corrosion rate are updated and revised estimates of service-life can be provided. An example procedure for including the effects of corrosion on the service-life of a structure for a specific mode of failure is examined and discussed.

INTRODUCTION

The problem of the loss of structural material to corrosion has been a thorn in the side of marine designers since metals first went to sea. In recent years tremendous amounts of money and effort have gone into the development of coatings and protection systems to mitigate this loss of structural material. Even with these systems there is still some problems both with the pitting and general wastage of the structural material.

When performing a service life analysis of marine structures both pitting and general wastage need to be included in the limit states. They can affect a number of potential failure modes through (1) the loss of structural strength through the general wastage of structural material under corrosive attack, and (2) the potential hazards associated with loss of structural integrity due to a through thickness penetration from localized pitting of the structure. The determination of the rate of corrosion and the rate of pitting has been a major difficulty in designing cost-effective and reliable structures. Part of the difficulty is associated with the errors in the measuring devices, but a larger part is due to attempting to characterize a field value (the mean thickness, thus the thickness loss) from a series of point estimations. This is especially difficult since corrosion is a random stochastic process.

The objective of this paper is to look at one possible way of including the effects of corrosion in a service-life estimation analysis using probabilistic methods. This includes both the determination of the corrosion rates and how to incorporate them into the analysis procedure. In order to determination the rates, the basic concepts of semivariogram analysis and kriging estimation are discussed. The procedure to include the information on rates is based on updating existing knowledge through regression analysis and the use of Monte Carlo simulation with Variance Reduction Techniques (VRT).

SEMIVARIOGRAM ANALYSIS AND KRIGING ESTIMATION

Steel plating which has been exposed to a corrosive environment on one side will exhibit a characteristically rough surface. One would expect that the thickness of the plating would vary from point-to-point on the surface of the plating. One would also expect that this variation would be local in nature; that areas which have experienced relatively large amounts of wastage would be interspersed with areas which have experienced relatively little loss of thickness. Over a large area, such as the bottom of a ship, one would expect little correlation in thickness loss between nonadjacent areas. However, there would be some correlation in local areas, ones which are relatively close to one another. The semivariogram function is a mathematical description of this relationship. If the mean thickness of an area, such as a panel of plating, is required, sample thickness measurements within the area can be used. The weights attached to each sample point will be based on their spatial relationship to one another and to the area being investigated. The semivariogram function is the means for determining the weights with kriging providing the best unbiased linear estimates of the weights. The following discussion is largely based on the works of McCuen and Synder [1], and McCuen, et al [2], and Ayyub and McCuen [3].

Semivariogram Analysis

The thickness of a material at any location X over the surface of the material can be denoted as Z(X). The thickness at any other location, a distance h away from the initial point at location X, can be written as $Z(X + h)$. If the distance h is relatively small then it is likely that there is some correlation between Z(X) and $Z(X + h)$. For large distances h this correlation is likely to be 0; that is Z(x) and $Z(X + h)$ are independent. The separation distance at which the autocorrelation between Z(X) and $Z(X + h)$ becomes zero is denoted as r, the range of influence. The range of influence can be evaluated from test data as will be described later.

We are interested in assessing the variability between the two measurements taken a distance h apart. The variogram, which is given as $2\gamma(h)$, provides the characterization of the variability of the property z between two points [2];

$$2\gamma(h) = \frac{1}{n}\sum_{i=1}^{n} [z(X_i) - z(X_i + h)]^2 \tag{1}$$

in which n is the number of measurement made at a separation distance h, and X_i is the location of a point with respect to some set of coordinates. It should be apparent that Equation (1) has the form of the expected value of the variable $[z(X_i) - z(X_i + h)]^2$:

$$2\gamma(h) = E\{[z(X_i) - z(X_i + h)]^2\} \tag{2}$$

The application of Equations (1) and (2) assumes that the value of the variogram only depends on the distance that the points are separated, h, and not on the location of the sample points, X, within the area being investigated. Another way of looking at this assumption is that the field of measurements represents a statistically stationary field. This is not to say that the differences for each pair of points a distance h apart must be equal, only that they must be from the same statistical population. Points that are some other separation distance apart may be from a different statistical population [1].

Dividing the variogram values from Equations (1) and (2) by 2 yields the semivariogram $\gamma(h)$. Typically, values of $\gamma(h)$ are found for discrete distances h and the results presented in the form of data points on a plot of semivariogram vs. separation distance. It is then very useful to fit a mathematical model of a semivariogram function to the data in order to obtain an estimate of the variation between data points. The most frequently used semivariogram model is called the spherical model and takes the following form [1]:

$$\gamma(h) = \begin{cases} \gamma_r & \text{for } h > r \\ \dfrac{\gamma_r}{2}\left[\dfrac{3h}{r} - \left[\dfrac{h}{r}\right]^3\right] & \text{for } h \le r \end{cases} \tag{3}$$

where γ_r is the semivariogram model parameter called the sill. The *sill* represents an upper bound on the value of the semivariogram. It is the characteristic of the model which says the variation of the samples at a distance r apart is equal to the variance of all of the data. There is no special relationship between the sample points at this separation distance. The spherical model is just one of many models available. It is the most widely used model because it has the shape and scale properties that characterize many real data sets.

Another form of semivariogram model which has been found useful in fitting data from corroded plates is a combination model. Because there is often a very rapid increase in variance, even at short distances, on the surface of a corroded plate, the variance data never approaches zero. The spherical model can be modified by including a constant, γ_n , sometimes called the *nugget effect*, to account for this phenomena. The combination form of Equation (3) would be [1]

$$\gamma(h) = \begin{cases} \gamma_r & \text{for } h > r \\ \gamma_n + (\gamma_r - \gamma_n)\left[\dfrac{3h}{2r} - \left[\dfrac{h}{2r}\right]^3\right] & \text{for } h \le r \end{cases} \tag{4}$$

The purpose of the semivariogram analysis is to provide a means of estimating the mean value of thickness of the plating over some specified area. However, the best estimate of that value is not the only thing we are after, we also want a measure of the accuracy of the estimated mean. This accuracy measure can be provided by the standard error of the estimate, or the error variance.

Error Variance of the Mean

If we collect only one thickness measurement z(X) in a field that has dimensions of L by W, then that value represents our best estimate for the mean value of the thickness of the plating. The accuracy of that estimate can be characterized by the error variance $2\gamma(h)$. The standard error of the estimate S_e would then be the square root of the error variance.

Since our sample consists of only one point in a field, the error variance is made up of two components. The first is the average variation between the sample point, S, and every other point within the field. From this we

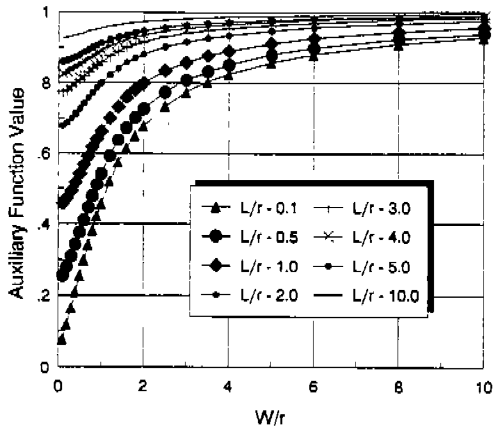


Figure 1. Auxiliary Function

must subtract the variation within all of the points in the field. The second component is subtracted because we are not interested in finding the standard error of the estimate for all points in the field, but rather the variation of the average value within the field. The error variance can be written as [3]:

$$S_e^2 = 2\Gamma(S,z) - \Gamma(z,z) \tag{5}$$

The variance component $\Gamma(S,z)$ is a function of the length and width of the field, the location of the sample point within the field, and the underlying semivariogram model. The computation of $\Gamma(S,z)$ involves integration and can be very tedious. Fortunately, a standardized auxiliary function $g(L,W)$ is available for the case of a sample taken in the corner of a field of length L and width W, assuming a spherical semivariogram model. Figure (1) provides some example values for this function based on the field dimensions being normalized by the range of influence r. The auxiliary function also assumes that both the sill and the range of influence are equal to 1.0. For a field where the sample point is not at one corner, Figure (1) may still be used by dividing the field into four rectangles with dimensions L_i and W_i and finding the weighted average of the values of $g(L_i,W_i)$ by [3]

$$\Gamma(S,z) = \gamma_r \left[\frac{1}{LW}\sum_{i=1}^{4} L_i W_i\, g(L_i,W_i)\right] \tag{6}$$

The term within the brackets is the auxiliary function for the case where the sill value is equal to one. To find the auxiliary function for the case of interest the bracketed term is multiplied by γ_r, the sill value for the current case.

The variance component $\Gamma(z,z)$ in Equation (5) is the field auxiliary function. This is the semivariogram between each point in the field and every other point in the field. The computation of this component involves carrying out the following integration [2]:

$$\Gamma(z,z) = \frac{1}{(LW)^2}\int_0^L\int_0^L\int_0^W\int_0^W \gamma(P_1 - P_2)\, dP_1 dP_2 dP_1 dP_2 \tag{7}$$

Because the integration in Equation (7) must be carried out numerically for all but the simplest models, a standardized set of solutions were developed. Examples of these solutions are presented in Figure (2) for the standardized case where $r = 1.0$, and the length and width dimensions are normalized by **r**. Reference [1] provides tables for other cases.

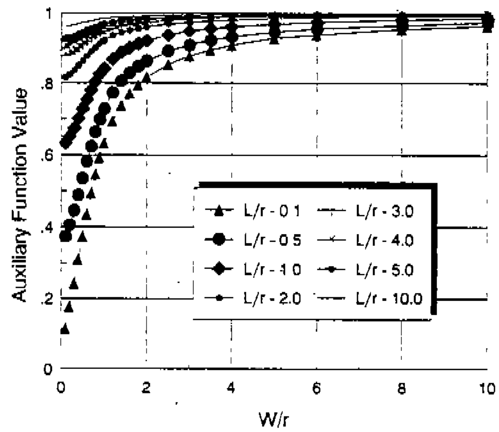


Figure 2. Field Auxiliary Function

If we have collected thickness measurements from n points in the field, then a reasonable estimate of the mean value could be given by:

$$\bar{z} = \sum_{i=0}^{n} w_i\, z(X_i) \tag{8}$$

where w_i is the weighting factor associated with data point $z(X_i)$. The weighting factor is an indication of the importance of measurement $z(X_i)$ in determining the mean value of thickness. Including more points in the sample also changes the form of the error variance. Equation (5) now becomes [3]

$$S_e^2 = 2\Gamma(S,z) - \Gamma(z,z) - \Gamma(S,S) \tag{9}$$

The third term in Equation (9), $\Gamma(S,S)$, represents the variation within the sample points. This variation must be subtracted from the first term because it represents variation that is imbedded in the first term. The within sample variance component is the average semivariogram value for every pair of points in the sample and can be computed for each pair using the spherical model of Equations (3) or (4).

Estimation of the Mean Thickness by Kriging

If weights are assigned to each sample point when determining the mean value, as in Equation (8), then the error variance needs to include the sample weights. Therefore, the estimate of the error variance can be given by [1]:

$$S_e^2 = 2\sum_{i=1}^{n} w_i\,\gamma(S_i,z) - \sum_{j=1}^{n} w_i w_j\,\gamma(S_i,S_j) - \Gamma(z,z) \qquad (10)$$

If we impose the constrain that the sum of the weights, w_i, must be equal to one, then we are able to solve for the values of the weight factors by minimizing the following function [1]:

$$\text{Minimize} \quad S_e^2 - \lambda \left[\left[\sum_{i=1}^{n} w_i \right] - 1 \right] \qquad (11)$$

Here λ is an unknown quantity and Equation (11) represents an example of a Langrangian optimization, with Equation (10) as the objective function. The λ term would then be the Lagrange multiplier and the solution would come from a set of $n+1$ equations for the $n+1$ unknowns, here λ and the n values of w_i. The solution vector of the weights would provide the minimum possible error variance for the given sample measurements, semivariogram model, and criterion function (mean value).

Sample Analysis for Steel Plating

In order to evaluate the effectiveness of the foregoing techniques, four ten-inch square pieces of heavily rusted mild steel plating were carefully measured. The nominal original thickness of the plating was 12 ga. (.128 in.) and they had been in active use in a marine environment for 12 years. The plates were weighed to determine the mean value of remaining thickness. Thickness measurements of the plates were mechanically taken at one-half inch intervals over the entire area of the plate. Using Equation (1), the semivariograms at separation distances, h, of from 0.5 to 7.0 inches were computed for each plate, The values of the semivariograms were plotted against separation distance, as shown for Plate A in Figure (3). Various values for the sill, γ_r, the nugget effect, γ_n, and the range of influence, r, were tried in equations (3) and (4) in order to find the curve with the best fit to the data. The scatter in the data at separation distances above 5 inches is a result of the size of the plates being used. There are fewer pairs of points that can be used to evaluate the semivariogram as the separation distance increases, thus decreasing the confidence in the resulting semivariogram estimates. The sill value, γ_r, is based on the variance of all of the data points on all of the plates and is considered to be an accurate measure. The results from the four plates suggested using γ_r = 110 mils^2, γ_n = 50 mils^2, and r = 5 inches.

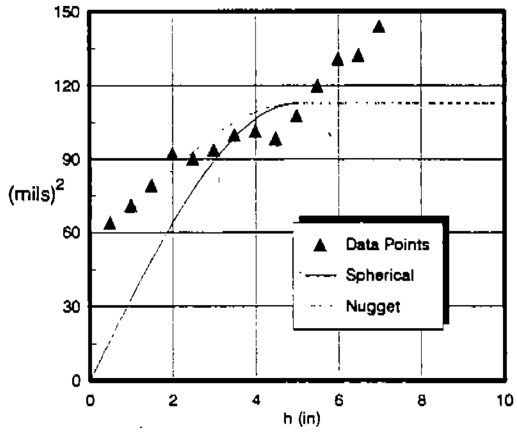


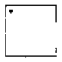
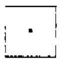
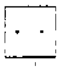
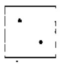
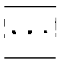
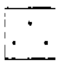
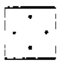
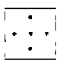
Figure 3. Fitted Semivariogram Models for Plate A

Estimation of Mean Thickness

In order to demonstrate how one might use semivariogram analysis and to see the effect of spatial location of the sample points, eight different sets of sample numbers and locations were investigated using one of the sample ten-inch square corrode plates. The plate, Plate A, had a mean thickness, found by weighing, of .097 inches. The results are given in Table (1). Using one sample point, the effect of moving in from the corner to the center of the plate reduces the estimated error. The change in the mean value estimate is just chance. Increasing the number of points to two dramatically reduces the standard error. The location of the two points on the plate also has a big effect on the error. If the points are moved too close together the area within their combined range of influence is reduced, thus causing an increase in the error term. If they are moved too far apart, there is no interaction between points and again the error is increased. There is also a point of diminishing return on increasing the number of sample points. The extra sample point taken when going from four points to five points in Table 1 only decreased the error by about 3%.

The scatter of the estimated thickness found using the semivariogram approach for the different sampling strategies shows two interesting points. First, the estimated thickness is highly dependant on the values at the sampling points. If there is any bias to the sampling strategy, the methodology proposed here would not be able to account for it. The analysis depends on the difference between sample points at a distance h apart being of the same statistical population. Table 1 further shows that as the error decreases, the scatter about the measured mean thickness also decreases. It is

Table 1. Results of Analysis of Plate A

Location	Number of Points	Estimated Mean Value (mils)	Standard Error, S_e (mils)
	1	110	7.91
	1	102	7.10
	2	93.5	5.08
	2	94.5	4.51
	3	96.5	3.98
	3	91.2	3.28
	4	100.4	2.60
	5	100.5	2.52

this relationship which allows the concept of semivariogram analysis to be used for developing sampling strategies.

Estimation of Corrosion Rates

One of the reasons for determining the remaining thickness of a corroding material is to use that information to determine a corrosion rate. Knowledge of the corrosion rate allows one to plan a measurement and inspection strategy which will increase reliability in a cost effective manner. However, the corrosion rate is really a stochastic random variable. It has some mean value, standard deviation, and distribution type. In previous work [4], the authors have found the lognormal distribution with a Coefficient of Variation (COV) of 0.25 to be a reasonable model for the statistics of an initial estimate for the corrosion rate [5].

Prior to taking samples one usually has an estimated value for the mean corrosion rate in a particular environment. This information can come from published reports or handbooks which are based on testing [5]. The mean value of thickness found using the analysis methods, proposed here, can be used to determine a new estimator of the mean corrosion rate. By subtracting the remaining thickness from the original thickness and dividing by the exposure time, an estimated mean corrosion rate is found. This rate can be assumed to be the outcome of an experiment and can be used to update the prior information using Bayesian updating [6]. Because the prior distribution for corrosion rate is lognormal, the conjugate distribution is also lognormal. That is, the posterior distribution for the mean corrosion rate is lognormal. The evaluation of the new mean value of corrosion rate is a fairly straight forward process given in most textbooks dealing with Bayesian methods [6]. It is interesting to note that the new mean value will be the weighted average of the prior mean and the mean determined from the semivariogram analysis. The weights are inversely proportional to the standard deviations of the estimated means.

The Wastage Allowance Model

As useful as the corrosion rate information is however, it will merely provide a point estimation at a specific instant in time. Because the corrosion rate changes with time the authors have chosen to treat the results of the semivariogram analysis as a new data point for estimating the mean wastage vs. time curve. When the structure is new, the mean wastage vs. time curve is a linear function of the estimated initial corrosion rate, because that is the only information available. The equation of the line takes the form

$$\text{mean value, } W_t = t\, R_m \qquad (12)$$

where W_t is the mean value of total wastage at time t, and R_m is the mean value of the initial estimated corrosion rate. But as information is obtained regarding wastage at later times, they can be used to estimate the mean wastage curve using regression analysis. The shape of the wastage curve will likely be adequately modeled as a power or exponential function.

The mean value isn't the only information available. An estimate of the shape of the distribution of wastage at time *t* can be provided by the sill value found from the semivariogram analysis. Typically, the coefficient of variation (COV) of the data is used as a measure of the distribution of values about the estimated mean. The COV of the wastage can be estimated from the data used in the semivariogram analysis. The COV of the wastage is also a function of time. The initial estimate comes from tabulated values, with later updates provided by the semivariogram analysis. As was proposed for the mean value of wastage, a regression analysis on the values for the COV of wastage could be performed. This would provide a relationship between COV and time similar to that available for the mean value. For this analysis it was assumed that initially there would be a linear increase in COV with time, given as

$$COV(W_t) = t\,COV(R) \tag{13}$$

where COV(R) is the initial estimate for the coefficient of variation of wastage and t is time .

EXTREME VALUE ANALYSIS

In order to make the wastage model compatible with the extreme value modeling of the load, it needs further development. According to the extreme value modeling, an extreme load is evaluated in a time period T, where T can be any value from the current time to the design life of the structure. The extreme load can occur at any point in time *t*, within the time period T. On the other hand, the plate wastage is a non-stationary stochastic process within the same time period T. This stochastic process can be simulated using Monte Carlo methods and converted into a random variable. The process can be summarized as follows:

1. Determine an initial value for time period T. This initial period must be set to an initial value larger than zero. The wastage at zero time period is zero.

2. Randomly generate time *t* according to a uniform probability distribution of the time period T.

3. At the generated time *t*, evaluated the statistics of wastage according to mean value and COV of wastage found from the regression analysis. The distribution type is Lognormal.

4. Randomly generate M wastage values according to the distribution as defined in step 3.

5. Repeat steps 2 to 4 N times.

6. Determine the mean value, COV and distribution type for the resulting M times N values of the wastage allowance W_a in the time period T as set in step 1.

7. Go to step 1, and increase the value of T, and repeat steps 2 to 6 for the new T value.

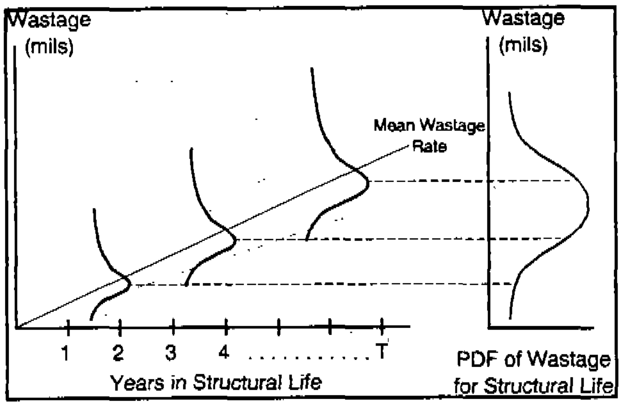


Figure 4. Wastage as a Function of Time

The above process is illustrated in Figure 4. For each time period T, a probability density function (PDF) of wastage for that period is generated. Each of these PDF's is then plotted and a curve constructed through the mean values, as shown in Figure 5. The resulting wastage allowance random variable W_a is a function of the time period T.

The simulation processes as described in the above steps was performed for assumed linear wastage rates R_m of 1, 2 and 3 mpy, COV(R) of 0.1, 0.25 and 0.4 and time periods T of 1, 5, 10, 15, 20, 25 and 30 years. Based on this parametric analysis, it can be concluded that the mean value of wastage allowance is dependent on the wastage rate and the period T, and is not dependent on the COV(R); while COV of wastage allowance is dependent on COV(R) and the time period, and is not dependent on the wastage rate R. These results were plotted in Figures 6,7, and 8.

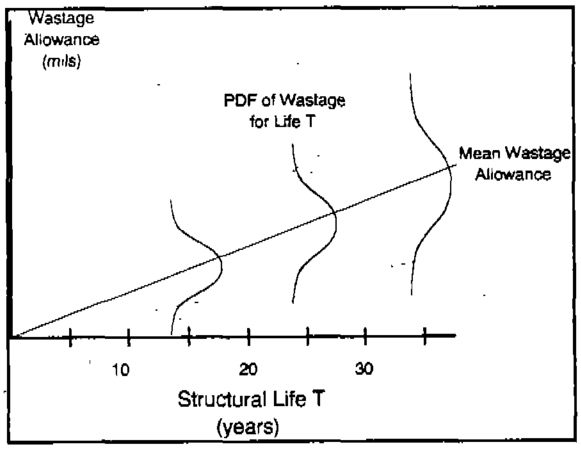


Figure 5. Wastage Allowance as a Function of Time

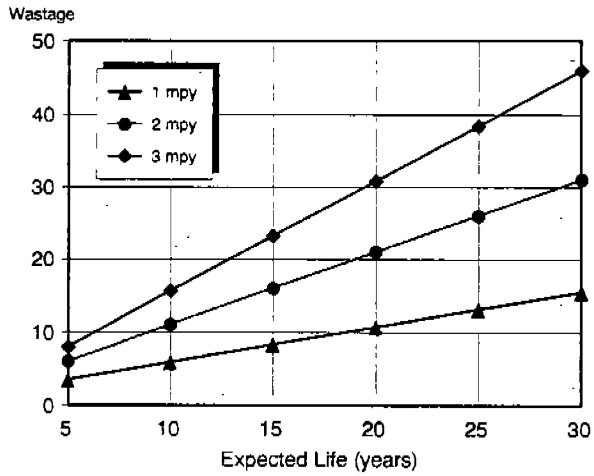


Figure 6. Mean Value of W_a as a Function of T and R_m

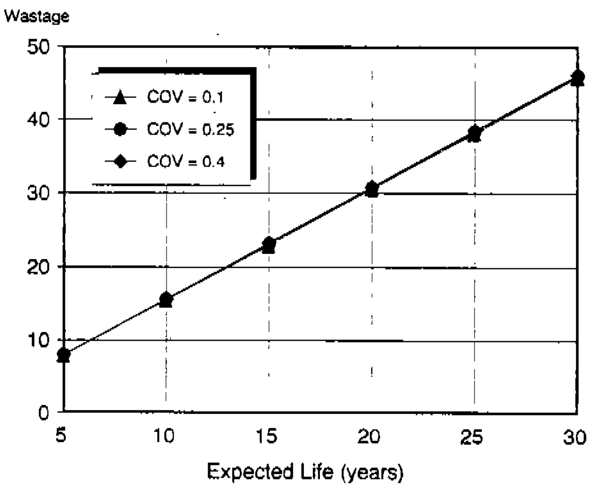


Figure 7. Mean Value of W_a as a Function of T and COV(R)

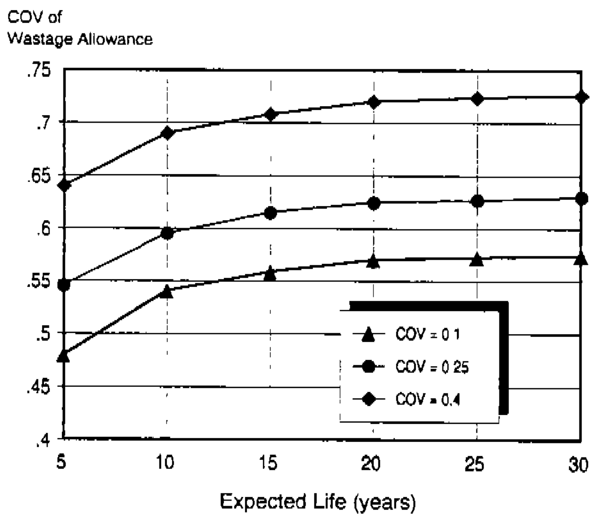


Figure 8. COV of W_a as a Function of T and COV(R)

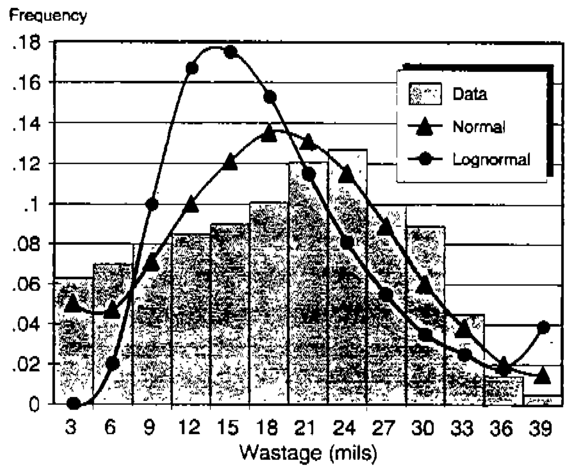


Figure 9. W_a PDF for an Expected Life of 10 Years

In order to examine the distribution type for the wastage allowance random variable, two cases were considered. In the first case, the mean value and COV of the wastage rate were taken to be 3 mpy and 0.1, respectively, and the time period equal to 10 years. The wastage allowance was simulated 2000 times, and a distribution goodness-of-fit was performed. The results are shown in Figure 9. In the second case, the mean value and COV of the wastage rate were taken to be 3 mpy and 0.25, respectively, and the time period equal to 25 years. Again the wastage allowance was simulated

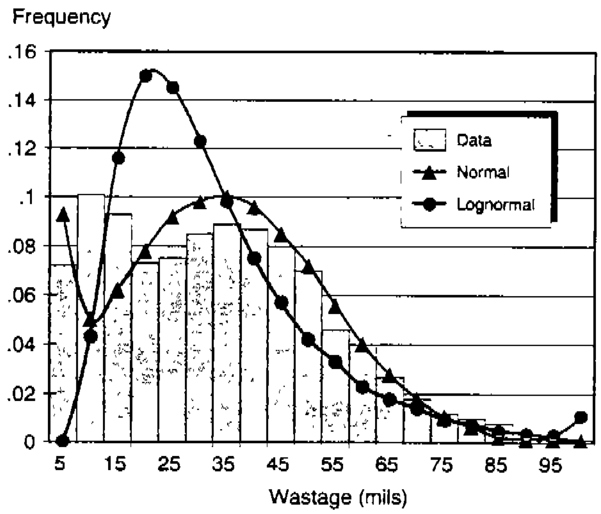


Figure 10. W_a PDF for an Expected Life of 25 Years

2000 times, and a distribution goodness-of-fit was performed. The results are shown in Figure 10. Based on these two cases, it can be concluded that the wastage allowance can be considered to follow a normal probability distribution with non-negative values, i.e., any simulated negative value for wastage allowance is replaced by a zero wastage allowance. This results in a heavier lower tail consistent with the actual distribution, resulting in reducing the statistical error in the assumed normal probability distribution model.

Conclusions and Recommendations

The information provided in this report gives a brief description of one means of including corrosion in a life expectancy analysis. The semivariogram analysis provides insight on the effects of measurement locations and the number required to get a desired level of confidence in the results. It has been shown that the location of the sample measurements can dramatically affect the level of error in the estimated mean. Of equal value is the knowledge that there is a decreasing benefit in taking more samples beyond a certain number. That certain number is tied to the area being investigated and the size of the range of influence. Determination of appropriate values for the sill and the range of influence is an important part of the analysis, and one which requires considerable effort. Calculations of these parameters for a variety of cases could be performed and tabulated in tables or charts. Confidence limits on the values for specific situations could also be provided and methods for including that uncertainty in the analysis developed.

The wastage allowance model provides a means of including the effect of corrosion in a life expectancy assessment of a marine structure. The procedure is more computationally efficient than it first seems. The development of the wastage allowance curve need only be done after new information is available, e.g. inspection results. Once the curve is developed, the equation for that mean value of wastage allowance for time period T can be used in the life assessment procedure.

There is still much work left to be done in order to make this technique a useful tool in engineering practice. A means for including the effects of time on both the sill value and the range of influence in semivariogram analysis needs to be developed. Errors associated with the measuring device need to be removed from the sample data so that the semivariogram is looking at the differences in thickness, not measurement errors. The authors are currently working on a means of investigating pitting using the approached described here.

Despite the remaining shortcomings, the semivariogram analysis with kriging estimation is still a very useful tool for making sense out of existing thickness measurements and shows promise of being able to do much more.

References

1. McCuen, R.H. and Synder, W.M., *Hydrologic Modeling: Statistical Methods and Applications*. Prentice-Hall, Englewood Cliffs, N.J., 1986.
2. McCuen, R.H., Aggour, M.S., and Ayyub, B.M., "Spacing for Accuracy in Ultrasonic Testing of Bridge Timber Piles," *Journal of Structural Engineering*, ASCE, Vol. 114, No. 12, December 1988.
3. Ayyub, B.M. and McCuen, R.C., "Optimum Sampling for Structural Strength Evaluation," *Journal of Structural Engineering*, ASCE, Vol. 116, No. 2, February 1990.
4. Ayyub, B.M., White, G.J., Bell-Wright, T.F., and Purcell, E.S., "Reliability-Based Comparative Life Expectancy Assessment of Patrol Boat Hull Structures," Report No. CG-D-02-90, U.S. Department of Transportation, Washington, D.C., January 1990.
5. Schumacher, M., *Seawater Corrosion Handbook*, Noyes Data Corporation, Park Ridge, N.J., 1979.
6. Ang, A.H-S. and Tang, W.H., *Probability Concepts in Engineering Planning and Design: Vol. I - Basic Principles*, John Wiley and Sons, New York, NY, 1975.