CORRO-CONSULTA

8081 Diane Drive | Rudolf H. Hausler | Kaufman, TX 75142

Tel: 972 962 8287 (office) | *rudyhau@msn.com* | *Fax: 972 932 3947*

Tel: 972 824 5871 (mobile)

Memorandum

To: Richard Webster, Esq. June 23, 2006

From: Rudy Hausler

Subject: **Discussion of Corrosion Monitoring Methodologies At Oyster Creek Nuclear Plant Dry Well**

SUMMARY

The corrosion on the outside of the Oyster Creek drywell steel liner, particularly in the former sandbed region, is of great concern in regards to the structural integrity of the liner. Various structural integrity calculations had been performed by Amergen/Exelon in the past to arrive at various wall thickness criteria. Subsequently these criteria were compared to actual measurements of remaining wall thicknesses. Going forward, continuing corrosion rates have been discussed, and times at which possible minimum wall thickness, as defined by the criteria, have been derived by the operator.

This study critically reviews what is known about the corrosion in the sandbed area, the way the corrosion measurements had been evaluated, and the conclusions that had been drawn.

As it turns out, only a very small fraction of the total sandbed area had been examined, which poses the problem as to whether in fact the most severely corroded areas had been observed, and whether extrapolation of these observation to the entire surface are justified.

The measurements were performed with a 6inch by 6inch template and consisted of point measurements at one-inch spacings. As a consequence assessments of corrosion flaws could only be made in the z-direction (depth) while the x/y dimensions of the flaws remained unexplored. However, acceptance criteria are based on spatial dimensions, which consequently had to be guessed at.

It had been assumed that pitting corrosion rates in the sand bed area would be constant in time. This assumption is not justified based on an analysis of the corrosion mechanism. It had also been assumed that the pit distribution would be Gaussian, and that therefore the deepest measured pits which were beyond

the 2s limit could be dropped from consideration. This is considered an unprofessional approach for two reasons. First: no measurements should ever be excluded from consideration (on statistical considerations only) unless it can be demonstrated that such measurements are flawed technically. Second: Pit distributions are not Gausian, but exponential, hence the deepest pits are of vital importance.

Amergen/Exelon evaluated corrosion rates based on average remaining wall thickensses. However it is well known that structures do not fail by averages but rather by extremes, namely where due to corrosion the wall thickness had become thinnest.

Consequently, evaluation of the available data by extreme value statistics demonstrated that the most probable deepest pits (corrosion anomalies) were deeper than those assessed by the operator or Oyster Creek.

At this point in time, there are no valid assessments of possible corrosion (pitting) rates. The operator assumed that conditions might have been constant over time and would remain constant in the future. However, this assumption cannot be justified under any condition.

It is, therefore, considered of primary importance that a) the entire drywell surface be examined with UT technology capable of assessing corrosion anomalies spatially. It is furthermore essential that the coating, which is well past its useful lifetime be examined with methodology other than just visual, in order to completely assess whether it is still protective. Programmatic aging surveillance must include such measurements much more frequently than every 10 years, because deterioration of the coating is not linear in time either.

I. Background

It is well established that serious corrosion occurred over the years on the outside of the drywell containment of the nuclear reactor at Oyster Creek [1)]. While corrosion occurred in all areas on the outside of the drywell, which experienced temporary or permanent wetting due to water leaks, the most severe damage was observed in the sandbed region [2)]. In 1992 the sandbed was removed and the corroded areas were coated with an epoxy coating. The coating was specified to have “an estimated life of 8 to 10 years”. Subsequently three UT inspections were performed in 1992, 1994 and 1996. Based these inspection results, projections were made to the effect that no corrosion would occur over the next 10 years. There are many concerns within this paradigm, which need further examination and discussion. The most striking are:

[1)] see for instance e-mail correspondence from George Beck (Exelon Corp.) to Donnie Ashley (dja1@nrc.gov) 4/5/06)

[2)] see for instance GPU Nuclear Corporation letter to US NRC September 15, 1995

- The assertion of no further corrosion based on the ’92, ’94 and ‘96 UT measurements was erroneously based on the assumption that conditions would remain constant, i.e. the epoxy coating of the dry well liner and concrete floor, as well as the elastomer used to seal the crevice between the floor and the drywell liner, all part of the sandbed region, would not deteriorate with time.

- The analysis of the results erroneously assumed that averaging 49 individual UT measurements, which were conducted over a 6x6 inch grid at 1 inch spacings would adequately represent the corrosion damage occurring in each bay, and hence these averages could therefore be used to conduct the necessary structural integrity calculations.

Embedded in these major concerns are a number of issues dealing with basic assumptions made in the evaluation of the corrosion measurements. These are listed below and will require some discussion:

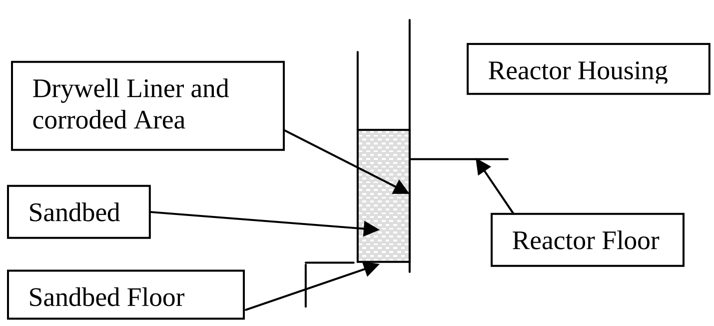
- Amergen/Exelon have assumed that the growth of the observed localized attach (pitting) would be linear with time, hence the corrosion rate) pit penetration rate) would be constant with time. The known pitting mechanisms will not support this assumption

- It has been furthermore assumed and so stated in many supporting documents that the pit size distribution would be normal (Gaussian). This assumption has then led to a number of conclusions and actions, which must be revised.

- It has also been assumed that averages of observed pit sizes would be representative of the corrosion processes, and that such averages from observations over time could be used to extract the “corrosion rate” (or more precisely, the pit penetration rate).

II. Some Comments Regarding the Corrosion Mechanism

A simple model as follows is being considered in order to delineate the major processes and parameters, which control them:



The fact is that the sandbed was essentially soaked with water, either periodically or permanently. This water was initially aerated which caused corrosion, even if the pH is above 7. As corrosion in the wet sandbed continues, the wet environment in the sand becomes depleted of oxygen. However, there is an almost inexhaustible reservoir of oxygen just above the sandbed – the air space. As a consequence, the steel surfaces embedded in the sand become anodic, while the cathodic reaction takes place on the areas which are richer in oxygen – the typical situation for crevice corrosion. The anodic reaction is not uniform, but pits will be forming. Initially, there will be a plethora of small shallow pits. Eventually some grow deeper than others, in fact at the cost of others. The frequency distribution of the pits is not normal, rather one can observe an exponential distribution – the frequency of pit depth decreases exponentially with pit depth. The fact that often a normal distribution is observed is an artifact, simply because the smaller pits are not normally measured, but are attributed to surface roughness and hence not included in the histogram. For this reason it is not proper to evaluate pit depth distribution on the basis of Gaussian statistics, and it is even less proper to discard deep pits outside the 95% confidence limits as atypical. Rather, deep pits, which have been measured, are a fact of life and must be included in any statistical evaluation, unless the measurement can be shown to be faulty for technical (not statistical) reasons. We will therefore show below the application of extreme value statistics to some data obtained from Oyster Creek.

Since pits are anodic areas where iron ions are being generated it stands to reason that anions must migrate into the pit, generally through a corrosion product layer, which fills the pit, such as iron hydroxide (two valent), or iron oxyhydroxide (three valent). The anions, which are present in the water at the highest concentration, are most likely to accumulate at the bottom of the pit where iron ions are being generated. The water in the sand bed is said to have contained as much as 500 ppm of chloride ions. This is more than the concentration the hydroxide ion at a pH of 7 or 1.5×10^{-5} mol/l Cl^- vs. 10^{-7} mol/L for OH^-. Chloride therefore will accumulate at the bottom of the pit. This will cause the pH in the pit to decrease to perhaps as little as 1 or 2. (This chemistry is well known and has been described in the literature many times). Lowering of the pH in the pit will accelerate pit growth, provided that the mass transport of water into the pit can sustain a higher corrosion rate at the bottom of the pit.

- It is therefore no forgone conclusion that the pit growth rate is constant with time. In fact, depending on the nature of the corrosion product in the pit, the mass transport into the pit can either be shut down, or sustain an accelerated corrosion rate due to the lower pH.

- Organic coatings will greatly reduce the transfer of both water and oxygen to the pit area. However, as the coating ages, such mass transfer is again accelerated. The unverified assumption that the coating will shut down pit growth for all eternity is totally unjustified. Furthermore, the unverified assumption that visual observation of the coated areas is sufficient to assert that no corrosion occurs is also unjustified. The assumption that if the coating held for 10 or 15 year it will hold for another 20 years is also unjustified and contradictory to general observations. (Coating life has been specified for 8 to 10 years).

- More disturbing, however, the fact that pit depth of 600 mils can easily be demonstrated statistically. This corresponds locally to a remaining wall thickness of about 550 mils or close to the 490 mil criterion for small areas. (This criterion, as we understand it is based not on buckling considerations, but on pressure calculations.) If isolated pits of that size exist, and extreme value statistics predict such pits with a high probability (see below), then the specter of chloride induced fatigue cracking is raised. Again, the danger is based on the fact that chloride is present in the base of the pit (and has actually been found there), that the pH in the pit is low, and that the stress at the pit tip is approaching a limiting value. All this contributes to stress corrosion and/or fatigue cracking. **It will therefore be necessary to examine the corroded areas, and in fact all areas susceptible to corrosion, for the possible existence of cracks in the dry well liner wall.**

III. Monitoring Frequency Is Too Long And Monitoring Periods Must Adapt To Safety Margins

Because the sand bed region is severely corroded, margins of safety are now much thinner than when the plant was first built. For example, in parts, over 0.5 inches of metal have corroded away from the steel drywell shell over areas larger than just single pits, leaving a metal thickness of just over 0.6 inches. According to AmerGen, no part of the drywell shell in the sand bed region should be thinner than 0.49 inches. Thus, to maintain safety, the monitoring regime must be able to predict how fast the metal could corrode to safety-critical levels, and must ensure that testing of areas that are closest to the margins occurs before there is any possibility that the metal has corroded too much.

The monitoring regime proposed does not achieve this goal because the monitoring frequency is too low and is not adaptive to how close the shell thickness is to the acceptance criteria, degraded areas of the shell would not be systematically identified and tested, the quality assurance for the measurements is inadequate, and the statistical techniques used in data analysis are flawed. This Memorandum discusses these issues in detail.

1. Overview

AmerGen has stated that it derived the proposed one in ten year testing frequency from the standard in service interval. Ex. NC 4 at 63. This is totally inadequate.

To insure that margins of safety are maintained, AmerGen must predict the worst case corrosion rate that could occur before the next scheduled round of monitoring. The monitoring regime should show that in the worst case the acceptance criteria will continue to be met. Interestingly, in the past the reactor operator has recognized this need to some extent. For example, in 1992 a calculation estimated that with 95% confidence, the mean thickness of area 13A would not go below 0.736 inches before June 1995. Ex. NC 7 at 9. The operator also predicted the minimum mean thickness at the 95% confidence level at the date of the next scheduled monitoring to verify that it was less than the acceptance criterion. Id. at 10.

However, more recently AmerGen has not estimated the corrosion rate at the sand bed because it has assumed that it is zero, which, far from being the worst case, is actually the best possible case. See NC 1 at 19 to 30. Furthermore, although the reactor operator used to provide 95%ile confidence limits for its predictions, AmerGen has ceased to do this for the sand bed region, id., while continuing to do this for the upper drywell. Ex. NC 6 at 8. AmerGen attempts to justify this on the basis that visual inspection of the sand bed is sufficient. Ex. NC 1 at 32. However, the coating could deteriorate between inspections, because it is already well past its 8 to 10 year expected life. Ex. NC 8 at 56. In addition, corrosion behind the coating could occur and not be noted visually. Furthermore the committed visual inspection period is once every ten years, the same as the UT testing period. Therefore, visual inspections will not provide any information on changes in conditions between UT tests.

In addition, because past analyses relied on prediction of the mean thickness, they failed to apply a corrosion rate to the measurements at individual points to ensure that even in the worst case they will remain thicker than the 0.490" acceptance criterion before the next scheduled monitoring. Furthermore, they failed to predict the rate of growth of the areas below 0.736 inches in each bay to ensure that they will also remain less than one square foot before the next scheduled monitoring.

At present, AmerGen has insufficient data to predict the worst case corrosion rate without sand. As discussed in more detail below, one reasonable approach to resolve this problem would be to use results taken before the sand was removed, derive a statistically valid worst case corrosion rate, and see how soon acceptance criteria could be violated using that rate. For example, AmerGen has stated that the thinnest individual result that has been measured is 0.603 inches. Ex. NC 1 at 7. The acceptance criterion for individual points is 0.490 inches. The uncertainty in each measurement is around 5% or 0.03 inches meaning that the thinnest real condition consistent with the measurement is around 0.573 inches.[3] This yields a current margin of safety of approximately 0.083 inches. The second highest long term corrosion rate estimated was 0.017 inches per year. Ex. NC 1 at 20. Thus, assuming that the next round of monitoring shows no further deterioration, and that

---

[3] As discussed below, AmerGen should make a more rigorous estimate of this parameter using appropriate statistical measures.

the worst case corrosion rate could be around 0.020 inches per year, further testing would be needed in approximately four years.

Turning to the area below 0.736 inches, bay 13 was closest to the safety margin when measurements were taken from the outside in 1992. The results showed that nine areas below 0.736" were widely scattered over a large area in this bay. Ex. NC 3 at Sheet 26-29. The outside of the shell was found have indentations from a thickness of around 0.800 inches that were "about 12 to 18" in diameter . . . at about 12 inches apart." Id. at Sheet 24. Measurements of nine one to two inch diameter areas at the thinnest parts of these indentations showed thicknesses ranging from 0.618 inches to 0.728 inches. Id. at Sheets 26, 28. The areas below 0.736 inches were "not more than 1 to 2 inches in diameter," except for location 7 which could have been 6 inches square with an average thickness of 0.677 inches. Id. at Sheet 26.

Applying the one square foot below 0.736 inches acceptance criterion to these measurements, the total area measured below 1 square foot was around 0.3 square feet. However, this area is very sensitive to additional corrosion because in a length of around 5 inches, the thickness changed from around 0.736 inches to 0.800 inches. Assuming that the edge of the hole is a straight line, this means that a change of 0.064 inches in depth occurs over about 5 inches in length. Thus, for the radius of the thin area to change by two inches, the depth would have to change by only 0.026 inches. If this occurred the total area below 0.736 inches would be approximately 1.6 square feet, well beyond the current acceptance criterion. Assuming a worst case corrosion rate of 0.020 inches per year shows that the area acceptance criterion could be violated in around a year, even if the thin areas have not grown bigger since they were last measured in 1992.

These results show that the currently proposed monitoring interval of ten years is far too long. If the worst case corrosion rate is around 0.020 inches, the total area under 0.736 inches could increase beyond the safety margin in about a year. Thus, monitoring would be needed at least once per year. Finally, if the next round of measurements shows that the margin of safety is less than it was in when the last valid round of testing occurred (in 1992 or 1994), the testing intervals must be increased accordingly.

2. Proposed Area To Be Measured Is Too Small

Large variations in remaining wall thickness have been observed. Minimum wall thicknesses of as little as 0.603 inches have been reported within the 6x6 inch grids. In addition, many other thin areas, with thickness measurements as low as 0.618, have been observed from the outside of the drywell. It is therefore entirely unreasonable to assume that the small 6x6 inch areas on top of the sandbed are representative of the over 3 foot thickness, Ex. NC 8 at 40, of the entire sandbed area, simply because around two thirds of the sand bed shell below the 6x6 inch grid was not accessible from the inside. See Ex. NC 10.

Furthermore, the spatial scope of the monitoring must be sufficient to allow meaningful comparison with the acceptance criteria that are to be applied to the results. In various submissions AmerGen has laid out how the monitoring was done in the sand bed region in 1992, 1994, and 1996. Initial investigations, carried out before the sand was removed, measured the thickness of the drywell shell in the sand bed region from the inside "at the lowest accessible locations." Ex. NC 5 at 11. However, because the interior concrete floor and curb is over two feet higher than the exterior floor this meant around two thirds of the sand bed area was not tested. To see if the corrosion extended to these areas the reactor operator dug a trench into the floor in bays 17 and 5 and found that the thinning below the floor level in bay 17 was similar to that observed above the floor, but eventually became less severe. Id. This confirmed that much of the area below the interior floor was corroding, showing that this area should not have been omitted from the monitoring regime.

In bays where initial investigations found significant wall thinning, 49 readings were taken within a 6 inch by 6 inch square centered at elevation 11'3". Ex. NC 2 at 5. In other bays, 7 readings were taken along a 7 inch horizontal line at the same elevation. Id. Thus, the initial selection of the points to be monitored periodically was fundamentally flawed because it omitted to establish monitoring of known thin areas below the interior floor level, and failed to even attempt to identify thin areas below the floor level in eight of the ten bays.

Measurements conducted from the outside of the drywell shell in 1992 highlighted these deficiencies in the initial investigations. The 1992 measurements demonstrated that there are extensive areas in bays 1 and 13 that are not proposed to be tested, but are already well below 0.736 inches thick. Ex. NC 3. For example in bay 13, nine areas below 0.736" were widely scattered over a large area. Id. at Sheet 26-29. Measurements of nine one to two inch diameter areas showed thicknesses ranging from 0.618 inches to 0.728 inches. Id. at Sheets 26, 28. Figure 13 on Sheet 29, shows the locations. To give an idea of scale, the distance between locations 5 and 7 was "about 30 inches apart." Id. at Sheet 26. For point 7 alone, the area below 0.736 inches was conservatively estimated to be 6 by 6 inches with a thickness of 0.677 inches on average. Id. at Sheet 26. Similarly, the measurements in bay 1 showed eight areas below 0.736 inches, whose thickness ranged from 0.700 to 0.726 inches. Id. at Sheet 11. The thinnest area was at location 7, which was located well below the "bathtub ring" and so cannot be easily monitored from the inside of the drywell. Id. at Sheet 12.

These results show that the spatial scope of the proposed monitoring is wholly inadequate to assess whether the drywell shell is meeting the acceptance criteria. Many areas that are thinner than 0.736 inches limit are not proposed to be monitored at all. Even those that have been monitored once are not fully characterized. To fully address all the areas that are below 0.736 inches, AmerGen must devise a systematic approach to identify and measure all such areas. Thereafter, each area must be measured and tracked to enable AmerGen to estimate the worst case corrosion rate and the worst case rate at which the thin areas could expand.

Because AmerGen is now proposing to measure at the same locations that it measured in 1992, 1994, and 1996, the scope of the monitoring will remain inadequate, even though the exterior of the sand bed is now accessible, so that the cause of the initial inadequacy no longer exists. Unless AmerGen can devise a way to monitor through the concrete in the interior of the drywell, it appears likely that future monitoring will need to be conducted from the outside of the shell.

A second, less difficult problem is that the square grid pattern employed in the previous testing may miss extended areas of thinness that are not square. For instance, if a 5 inch by 30 inch horizontal trough were present in the shell and intersected the measured area, its area would only be estimated as 5 inches by 6 inches because of the area limitation of the measurements. Thus, its area would be estimated as 0.2 square feet, whereas the actual area would be over 1 square foot, in violation of an acceptance criterion. This means that if the testing finds points below 0.736 inches on the outside of the grid, it will underestimate the continuous area that is below 0.736 inches. To avoid this error AmerGen should expand the search area where or when readings at the edge of the grids show readings of less than 0.736 inches. It has failed to propose such a change.

3. The Quality Assurance For The Measurements Is Inadequate

Recently, the NRC concluded that the quality of the calibration for the UT measurements taken after 1992 is in question. Transcript of Meeting on June 1, 2006, attached as Citizens' Exhibit NC 4, at 28. Further, NRC said that the 1996 results are anomalous because they show that the drywell shell got dramatically thicker between 1994 and 1996. Id. at 28, 31. AmerGen responded that they had spent a lot of time trying to find the source of the problem, but were unable to explain why the results were so high. Id. at 29. AmerGen also acknowledged that it could not explain the increase between 1994 and 1996, Id. at 31, but would do additional calibration to see if the coatings on the inside and outside of the drywell affected the results. Id. at 29.

The systematically higher wall thicknesses observed in 1996 cannot be explained purely by the presence of the epoxy coating, because the coating was present when the previous two measurements were taken from the inside in 1992 and 1994. One potential explanation for the anomalous 1996 results is the start of coating deterioration. It is known that certain poly-epoxides tend to swell in the presence of humidity and at elevated temperature. It is proposed, as a working hypothesis, that the higher measurements in 1996 may well be due to such swelling, which could not have been calibrated out of the measurements. As a consequence, the actual thickness of the drywell shell in 1996 might well have been lower than the 1994 measurements due to ongoing corrosion, albeit a slower pace than pre-1992. What is clear is that the 1996 results cannot be used to predict future corrosion rates, and that even in the 1992 and 1994 post-coating results are in question.

Had AmerGen had an effective quality assurance program in place when the results were taken in 1996, it would have identified any problems with the data close to the time that they were taken. As illustrated by my memo of May 3, 2006, the anomaly in the 1996 results was not difficult to find, provided systematic rather than random error was the focus. Thus, Amergen obviously did not have an adequate quality assurance program in place. AmerGen has recently stated that the same methodology will be used to analyze the 1992, 1994, and 1996, will be used for the new UT results. Ex. NC 2 at 2. This means that AmerGen will continue to fail to identify questionable data in a timely manner, unless it changes its approach to the identification of systematic error.

Furthermore, although AmerGen realized at some point that there were questions about the reliability of the thickness data taken after 1992, especially the 1996 results, it has continued to use these data to predict the thickness of the drywell shell during any license renewal period. See e.g. Citizens' Exhibit NC 1 at 19-30. This is wholly unjustifiable. Unless questions about calibration of the results taken after the coating can be answered, the post-coating thickness data provide little knowledge about the actual thickness of the drywell shell, let alone the corrosion rate.

4. Statistical Analysis Of Results Is Flawed

a. Background

As the NRC has recognized, uncertainty is the key issue when analyzing the UT results. Ex. NC 4 at 63-64. In fact, there are a number of uncertainties, all of which need to be taken into account in the design of the monitoring regime. The first is that the UT results themselves are subject to uncertainty. This uncertainty means that the thickness at the time the measurement is taken is uncertain and it also means that the rate of corrosion is uncertain. Adding to the uncertainty in the corrosion rate is that conditions may change over time. For example, coatings may deteriorate, or the volume and composition of the water reaching the corroded area may change.

Since the actual original UT measurements were not available for a detailed statistical analysis, a hypothetical 6x6 inch grid was constructed to illustrate a point to be made here. Figure 1 shows hypothetical UT measurements in a 6x6 grid with 1 inch spacings. The average wall thickness over the grid is 0.81 inches. However, as is often the case in real life, a corrosion trough is depicted parallel to the y-axis with an average depth of 0.68 inches and a maximum depth of 0.55 to 0.60 inches. While this example is not a real life observation, it nevertheless illustrates how averages can be misleading. In this particular case, the corrosion trough exceeds the grid, and one could not tell whether corrosion would become more severe or less severe beyond the boundaries of the grid. Similarly, when Amergen talks about "isolated minimum thickness measurements", one does not know where these were recorded and whether there were others, which exceeded the average wall loss, but may have been above the quoted minimum wall thickness. When the same data shown in Figure 1 are plotted from a different perspective (Figure 2), conclusions may be different, but again it appears that there may be an extensive corrosion phenomenon on one side of the grid.

In the treatment of the current thickness, AmerGen has set various acceptance criteria: one for small areas of around 2.5 inches in diameter (0.490 inches), one for areas of less than 12 by 12 inches (0.535 inches), one for the total area where the wall thickness is less than 0.735 inches (one square foot), and one for the mean thickness of the vessel (0.753 inches). In comparing the measured data to the acceptance criteria, the reactor operator actually evaluated the UT results by comparing the means of the 6 by 6 inch grids to 0.535 inches, and comparing each measurement to the small area criterion. Ex. NC 2 at 11.

b. Modeling

It is generally accepted that failures do not occur as a result of average corrosion, but are generally occasioned by the weakest spot in the system. As a consequence, one cannot interpret the data by calculating averages and standard deviations. Figure 3 for instance shows a histogram of the 49 hypothetical UT measurements. It can clearly be seen that in this example a bimodal distribution exists. The first mode, covering small pit depths, is perhaps Gaussian, as is often observed for pit initiation, because the smallest pits, too difficult to count, are rarely included in the analysis. The second mode is represented by a skewed distribution, perhaps a Weibull distribution with very high extreme values. Again this type of distribution is often observed after pitting has progressed for some time. It would clearly be irrational to try and present data of this kind by a Gaussian distribution and disregard the values that are outside "confidence limits". Rather, data of this kind should be analyzed by Extreme Value statistic. It turned out, as shown in Figure 4 that a reasonably straight line is obtained when the pit depths are plotted as a function of the "reduced variate". Only 49 points were available for the correlation. Extrapolation to the virtual 100th point results in a pit depth of about 0.77 inch, a remaining wall thickness of about 0.4 inches, or in this hypothetical case, a remaining wall thickness of less than minimum allowable. Because a worst case analysis is necessary for a safety-critical condition, the data must be analyzed by a methodology similar to the one demonstrated in the above procedure. Unfortunately, at present AmerGen appears to take no account of the chance that the true value of the remaining wall thickness at each point could actually be substantially less than indicated. See Ex. NC 3 at Sheet 6.

Turning to the corrosion rate, AmerGen attempted to predict corrosion rates based on the '92, '94, and '96 UT measurements. They used the averages for each grid measured in each by over the time period indicated. (This procedure is based on the notion that all pits grow at the same rate, which is quite erroneous since the deepest pits usually grow faster than the smaller ones.) In most instances it turned out that the 92 averages were higher than the 94 averages, while the 96 averages were again higher then the previous two. This is shown in Figure 5. However, a statistical Analysis of Variance (ANOVA), Figure 6, shows that there is no significance to these variations from date to date if the data are amalgamated. However, the differences from location to location are indeed very significant. On the basis of these data Amergen concluded that the corrosion was arrested following the application of the epoxy coating. It is probably correct that *on average* the corrosion was significantly slowed or even arrested during the four years covered by the measurements. Whether the extreme corrosion rates were also similarly affected remains an open question. Nevertheless, it would be logical to expect that corrosion slowed down following the application of the coating, at least for a period of time. It is, however, stretching credulity to assume that such protection would last in excess of the stated lifetime of the coating, which was specified as 8 to 10 years.

Turning to the details of the analysis, the way in which AmerGen calculated corrosion rates was flawed in at least four ways. The calculation of estimated corrosion rates erroneously assumed that the rate would be constant over time, the means of the 49 point grid were used for curve fitting, the most extreme values were often omitted from the calculation of the means, and a ninety five percentile statistic is used as the appropriate level of uncertainty for future predictions.

Taking each of these flaws in turn AmerGen first made the erroneous assumption that "if corrosion is continuing, the mean thickness will decrease linearly with time." Citizens' Exhibit NC2 at 6. In fact, if the coating starts to fail, the corrosion rate could increase rapidly in a non-linear fashion. The projected coating life is around eight to ten years, and that life has now been exceeded by around four to six years. Ex. NC 8 at 54. In addition, other conditions could change. Thus, the assumption that the corrosion rate will be constant with time is simply invalid.

Second, using the means of the 49 points rather than the individual points to produce the curve fit that is used for future predictions only serves to mask the inherent uncertainty in the data, because the means are less variable than the individual points. See e.g. Ex. NC 1 at 21. The fit statistics from the curve fit program therefore do not fully represent the uncertainty in the fit because the errors are artificially lowered by only feeding in the means, rather than individual measurements. A more appropriate procedure would be to plot all the individual measurements and then do a curve fit and find the predicted errors on the curve fit at an appropriate level of uncertainty.

Third, AmerGen does not include the thinnest points in the means it reports, because it treats pits separately in the analysis when the data are not normally distributed. E.g. Ex. NC 5 at 25. A more recent analysis of upper region results by AmerGen best illustrates the problem. The analysis candidly states "points that were considered pits are . . . excluded from the mean." Ex. NC 6 at 15. Such a procedure obviously leads to an underestimation of the mean value and the corrosion rate. Thus, in some cases the mean values that have been plotted and fitted are actually thicker than the mean thicknesses of the areas that were measured. This is obviously a major problem with the analysis because one acceptance criterion is based on the mean values of the grids.

Fourth, only the ninety five percentile of extreme values are used for the prediction of the corrosion rate. E.g. Ex. NC 5 at 1. This means that even if the prediction is correct and the 95%ile confidence limit is taken as the worst case corrosion rate, there is a one in twenty chance that the actual corrosion rate will be higher than that calculated. For a safety-critical evaluation, this level of uncertainty is far too high. The statistical procedure must be redesigned to insure that safety margins are met with a substantial degree of certainty.

Further flaws have crept into the analysis over time. In 1992 the reactor operator recognized that to estimate a 95%ile of the corrosion rate, at least four data sets are needed. Ex. NC 5 at 1. It further recognized that where only two points were available, the uncertainty in the individual points should be used to plot a straight line. Id. However, more recently AmerGen concluded the corrosion rate was zero based on only three points, one of which it has now recognized as unreliable. AmerGen now intends to confirm this conclusion by taking one more set of measurements before the start of any license extension period. Because at least four reliable sets of measurements are needed AmerGen would continue to have insufficient data to predict the corrosion rate reliably, even if conditions over time had remained constant.

In fact, it is highly likely that conditions have changed since 1994, therefore realistic wall thickness measurements must be made as soon as possible to establish the current baseline and margins of safety. In addition, a worst case corrosion rate needs to be established for the current time period. This is obviously impossible based on just one point. I therefore suggest a pragmatic solution. AmerGen should use the corrosion data it gathered previously to estimate a statistically valid worst case corrosion rate based on previous conditions, which were with water and sand present and without a coating. This approach should have some inherent conservatism because the removal of the sand and the coating appeared to slow the corrosion for the period from 1992 to 1994. Thus, even if the coating has now become ineffective, the previous conditions should continue to provide a worst case scenario, provided a statistically valid approach is used.

c. Modeling and Statistical Analysis with Actual Data

The calculation sheet (EX NC 3 (DRF 143071)) contains sufficient original data to analyze GPU Nuclear's evaluation of the UT and micrometer pit measurements.

Figure 7 is a schematic of what I understand was done to arrive at a representative remaining wall thickness in the former sandbed region in order to subsequently perform GE type vessel integrity calculations. First: UT measurements were taken from the inside as described earlier. Second: an imprint (or cast) was taken from the outside in order to characterize the roughness of the corroded surface in addition to the UT measurements. The roughness is also characterized as "dimples". The depth of the dimples was measured from the imprint by means of a micrometer. Thus, Figure 7 shows the UT measurement (1) from the inside, which characterizes the remaining wall thickness. Second, the dimple depths were measured (repeatedly) and averaged (2). This average was added to each UT measurement. Third, a characteristic average dimple depth was determined and used as a global average to be used in all areas were imprints were not available, or where such were performed in a reduced fashion. The reason for this procedure is not entirely clear, other than hopefully arriving at a representative average, which could be the basis for the integrity calculations.

As can be seen from Figure 7, the first location: if the average dimple depth is added to the UT remaining wall thickness, and then a global average dimple depth is being subtracted from the result, the actual pit depth may be reduced. In the second location the actual average pit depth may be increased by this procedure.

However, a more detailed analysis of the some of the available data shows that this procedure performed by GPU Nuclear may show milder corrosion than what actually prevails.

Detailed Data Analysis

Appendix A of above reference document lists the measurements of impressions taken from Bay # 13, presumably the Bay where corrosion was the roughest. The average of all "dimple" measurements is 0.13 in with a standard deviation of 0.07 in. GPU Nuclear used the same average plus one standard deviation to arrive at the value of 0.2 in for the characterization of the average roughness of the corroded surface. It is not clear why only on standard deviation was added to the average when in fact 2 standard deviations represent a confidence limit of 95%. Hence it is my opinion that 0.27 in should have been used to represent worst case, or 270 mils. If this had been done, for instance, for the UT measurements summarized in Table 1-b (page 11) of referenced document, five of the 8 locations cited would have been below the acceptable criterion of 736 mils, while GPU Nuclear found all eight locations acceptable.

It is therefore concluded that the procedure employed by GPU Nuclear is highly arbitrary, since the one vs. two standard deviations has not been explained.

Extreme value Statistics

Figure 8 shows the extreme value statistical evaluation of the UT Measurements in Bays 1 and 13. It can be seen that worst case penetrations can be predicted to be of the order of 550 to 600 mils, or dangerously close to the criteria for the remaining local wall thickness of 490 mils. Hence, predictions of this nature, which in the case of Bay 13 are reasonably accurate, (R^2 = 0.95), are considerably less optimistic than those of GPU Nuclear.

Figure 9 shows a comparison of the measurements in Bay 13 of the UT remaining wall thickness and the dimple depths. The correlations are reasonably good. The prediction for the most severe dimple depth is about 300 mils, or 50% larger than the average used by GPU Nuclear, and more in line with the use of 2 standard deviations.

Interestingly, the difference between the UT measured pit depth from the outside and the pit depth arrived at by micrometer measurements using the cast imprint turns out to be 200 mils at the lower pit depths and 300 mils at the higher pit depts. The 300 mil figure results in an average remaining wall thickness over the measured area of 1154 mils minus – 300 mils equals 854 mils, a number which has also been used in integrity calculations aimed at the buckling question. However, as pointed out earlier, this is clearly an average and one does not know how large the area is, which was further reduced by localized pitting.

And herein lies the difficulty of what has been done in the past and what Amergen/Exelon proposes to do in the future. 99% of the sandbed region has not been monitored overtime and even the small areas that have been monitored are incompletely characterized. The overall area of the sandbed region is of the order of 300 ft^2. AmerGen are proposing more measurements at 12 6 inch by 6 inch areas, or a total of 3 ft^2. Thus only 1% of the total area is proposed to be monitored. In those small areas, point UT measurements, as have been done in the past, using a template and positioning the sensor always at the same location give information about the remaining wall thickness at this location (z- direction), but contain no information about the extent of the reduction in wall thickness around the point measurement (x-, y-directions). Hence around 93% of the 0.25 ft^2 area of each template remains unexplored. **For these reasons it is urged that Amergen/Exelon consider using more modern UT methods which are capable of scanning large areas and can generate data in all three directions, x, y, and z.**

Conclusions

This brief analysis of the original data presented in 1993 (measured in 1992) depicts a more severe corrosion situation than was extracted by GPU Nuclear on the basis of averages. Hence we think that a much more detailed analysis of the integrity of the remaining wall thickness is warranted and required, the repeated assertions that the coating has arrested any further corrosion not withstanding.

Signed

Rudolf H. Hausler

Figure 1

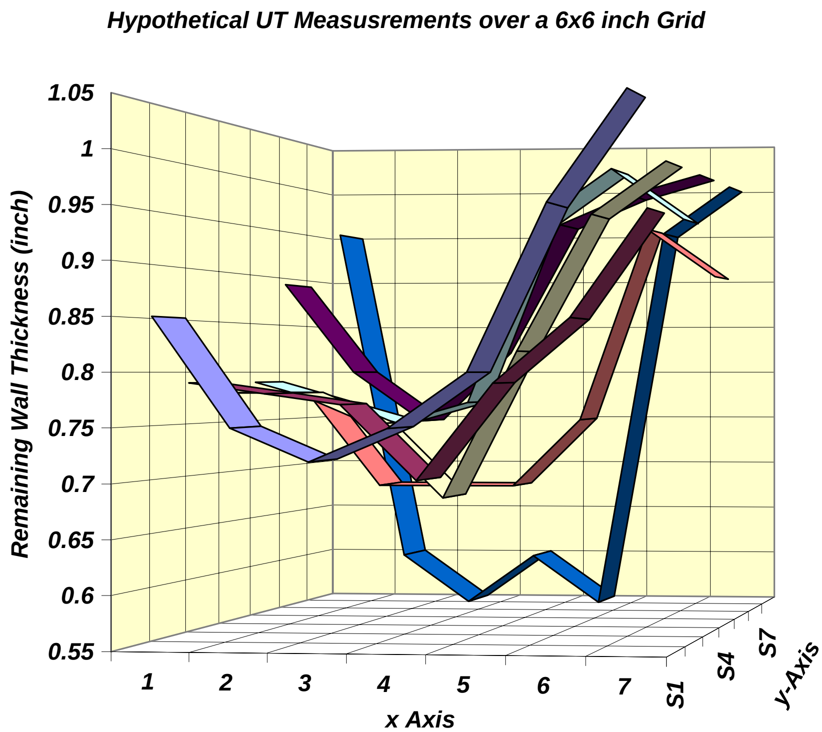


Figure 2

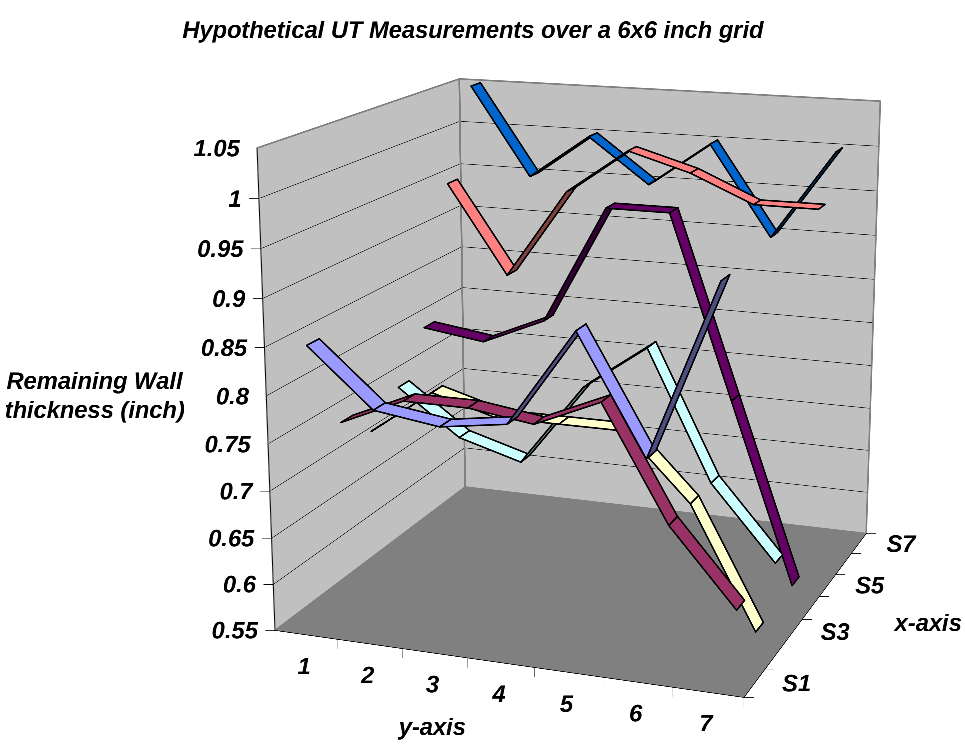
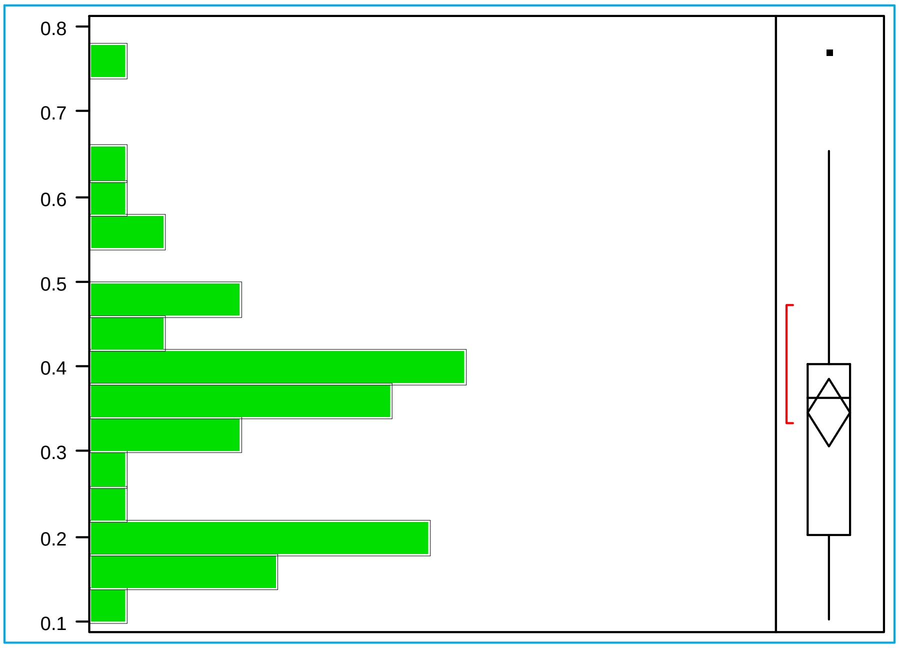


Figure 3

pit depth



Quantiles

Moments

Mean	0.34655
Std Dev	0.14139
Std Error Mean	0.02000
Upper 95% Mean	0.38674
Lower 95% Mean	0.30637
N	50.00000
Sum Weights	50.00000

Test for Normality

Shapiro-Wilk W Test

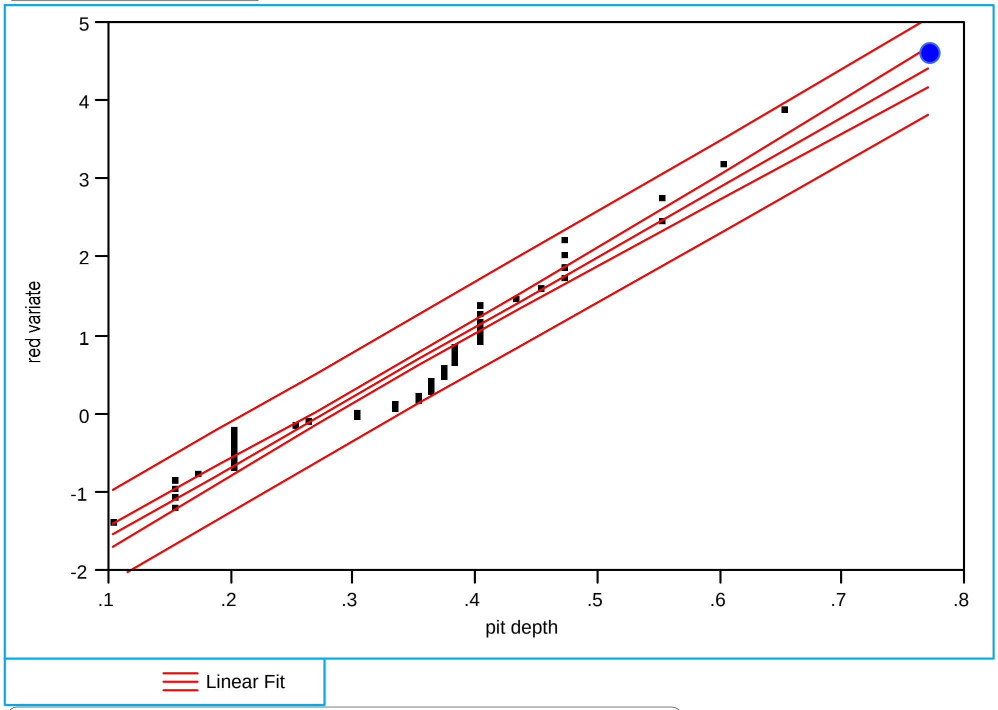
W	Prob<W
0.943173	0.0284

Test Mean=value

Histogram of 49 Hypothetical UT Measurements over a 6x6 inch grid with 1 inch spacing.

a. Figure 4

red variate By pit depth



Linear Fit

red variate = -2.4687 + 8.93965 pit depth

Summary of Fit

RSquare	0.954781
RSquare Adj	0.953839
Root Mean Square Error	0.277923
Mean of Response	0.629364
Observations (or Sum Wgts)	50

Analysis of Variance

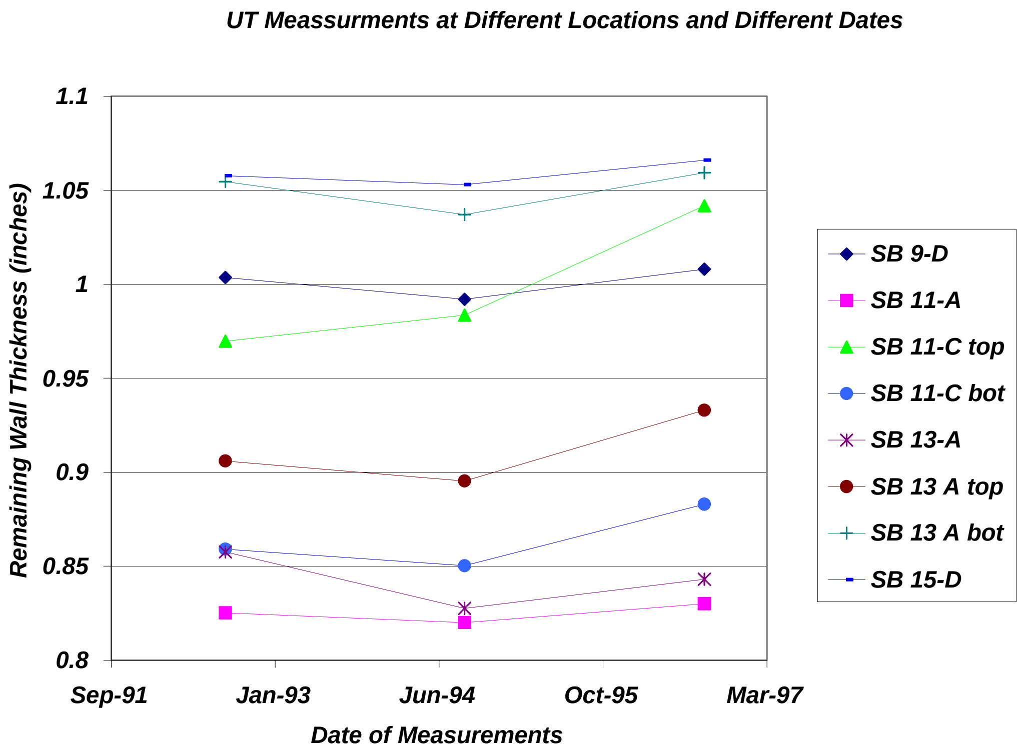
Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	78.284650	78.2846	1013.507
Error	48	3.707584	0.0772	Prob>F
C Total	49	81.992234		<.0001

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-2.468701	0.104952	-23.52	<.0001
pit depth	8.9396488	0.280806	31.84	<.0001

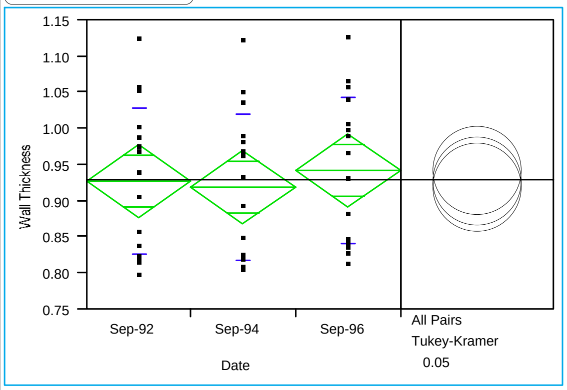
Extreme Value Statistical Plot of 49 Hypothetical UT Measurements over a 6x6 inch grid. The last point at 0.77 inch pit depth is the most probable pit depth obtained by extrapolation if 100 data point had been measured. It is within the statistical 95% boundaries for the fit.

b. Figure 5

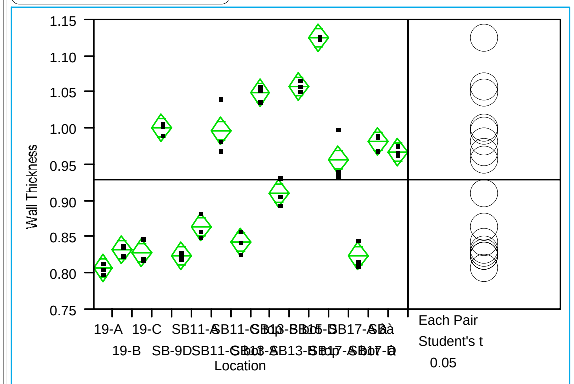

UT Meassurments at Different Locations and Different Dates
Remaining Wall Thickness (inches)
1.1
1.05
1
0.95
0.9
0.85
0.8
Sep-91
Jan-93
Jun-94
Oct-95
Mar-97
Date of Measurements
SB 9-D
SB 11-A
SB 11-C top
SB 11-C bot
SB 13-A
SB 13 A top
SB 13 A bot
SB 15-D

c. *Figure 6*

Wall Thickness By Date



Wall Thickness By Location



Wall Thickness By Date

Oneway Anova

Summary of Fit

RSquare	0.010207
RSquare Adj	-0.03378
Root Mean Square Error	0.101758
Mean of Response	0.930158
Observations (or Sum Wgts)	48

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	2	0.00480510	0.002403	0.2320
Error	45	0.46596232	0.010355	Prob>F
C Total	47	0.47076742	0.010016	0.7939

Means for Oneway Anova

Level	Number	Mean	Std Error
Sep-92	16	0.927588	0.02544
Sep-94	16	0.919394	0.02544
Sep-96	16	0.943494	0.02544

Std Error uses a pooled estimate of error variance

Means Comparisons

Dif=Mean[i]-Mean[j]	Sep-96	Sep-92	Sep-94
Sep-96	0.000000	0.015906	0.024100
Sep-92	-0.01591	0.000000	0.008194
Sep-94	-0.0241	-0.00819	0.000000

Alpha= 0.05

Comparisons for all pairs using Tukey-Kramer HSD

q*
2.42362

Abs(Dif)-LSD	Sep-96	Sep-92	Sep-94
Sep-96	-0.08719	-0.07129	-0.06309
Sep-92	-0.07129	-0.08719	-0.079
Sep-94	-0.06309	-0.079	-0.08719

Positive values show pairs of means that are significantly different.

Wall Thickness By Location

Oneway Anova

Summary of Fit

RSquare	0.979337
RSquare Adj	0.969651
Root Mean Square Error	0.017435
Mean of Response	0.930158
Observations (or Sum Wgts)	48

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	15	0.46103987	0.030736	101.1099
Error	32	0.00972755	0.000304	Prob>F
C Total	47	0.47076742	0.010016	<.0001

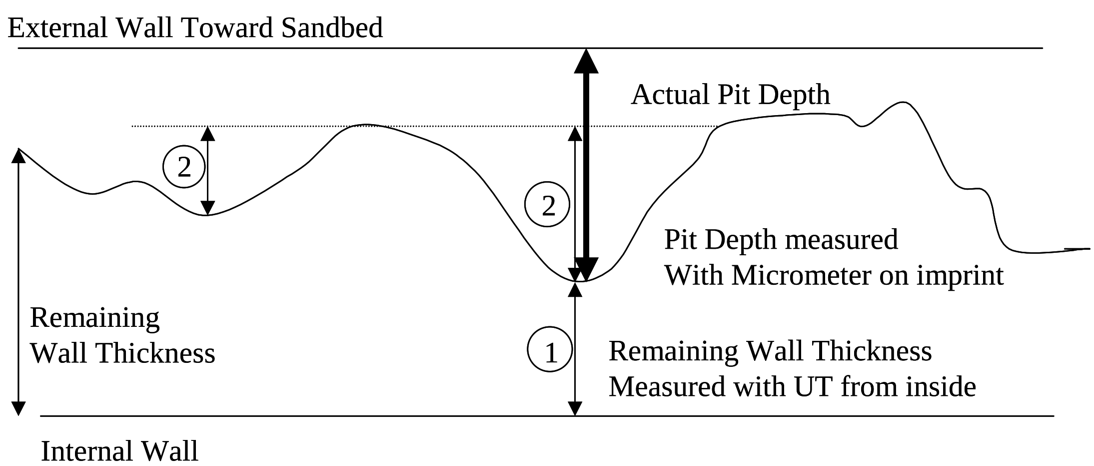
Means for Oneway Anova

Means Comparisons

Statistical Evaluations of all available UT Measurements performed in 1992, 1994 and 1996 on the drywell liner in the sandbed area

Figure 7

Schematic of Evaluation of Pit Depth Measurements and Averaging Procedure



Wall thickness used for integrity evaluations:
(1) + average of (2) – 200 = T (evaluation)

Figure 8

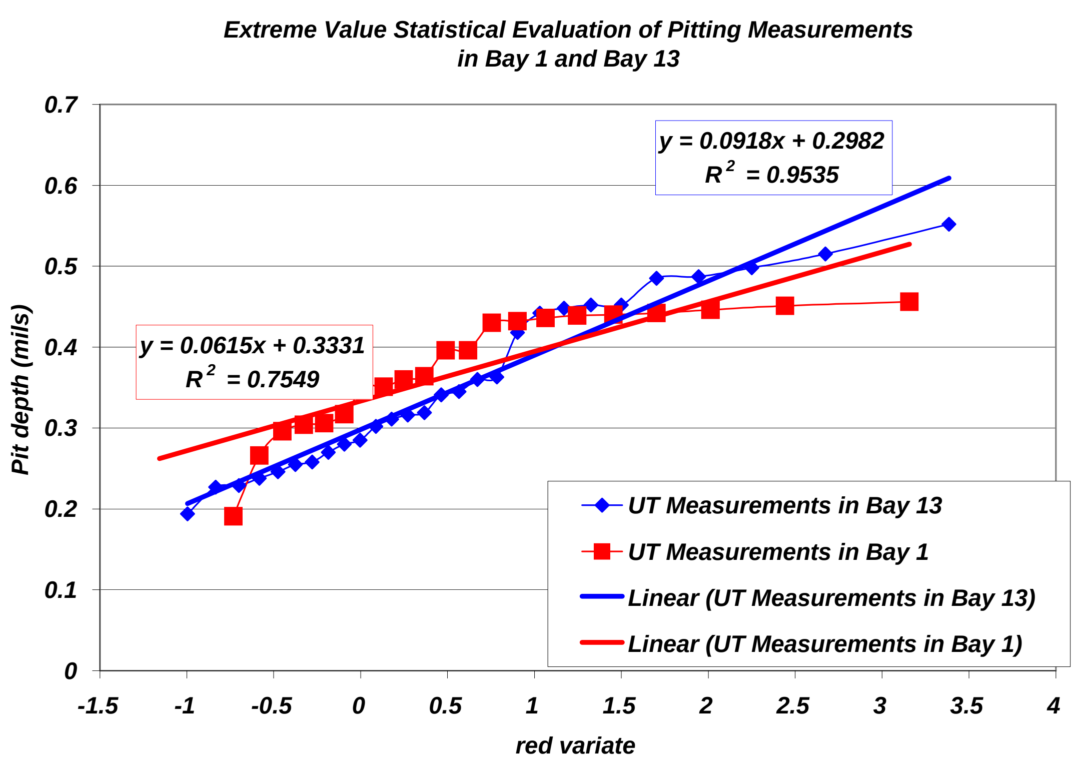
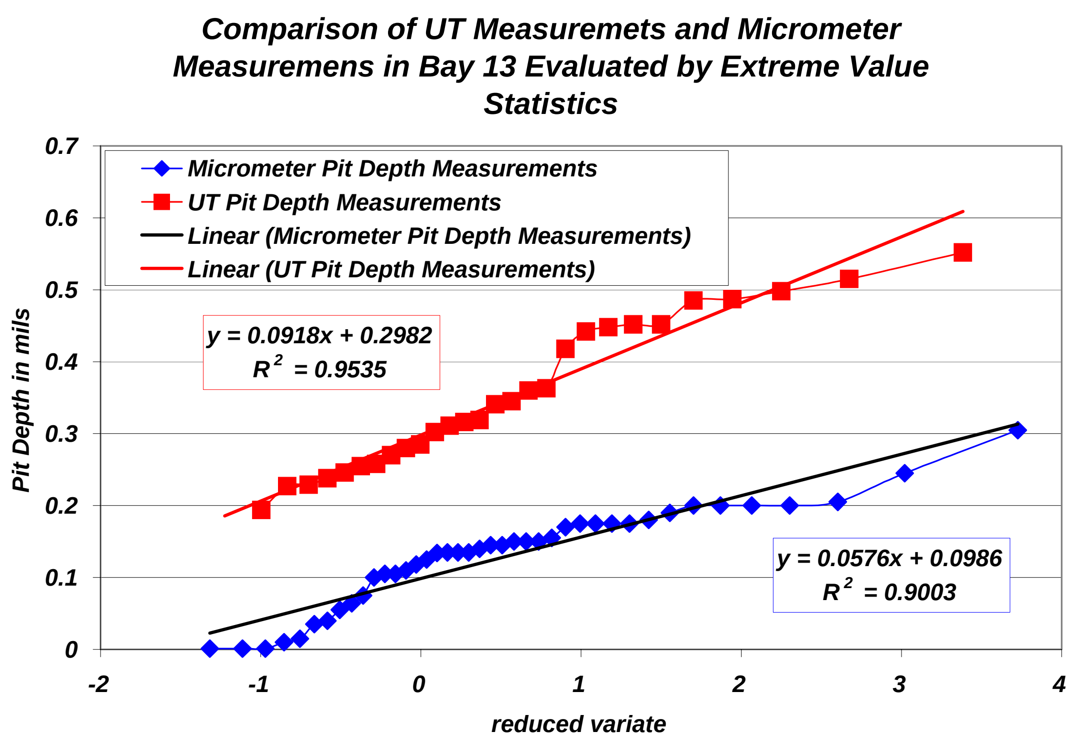


Figure 9


Comparison of UT Measuremets and Micrometer Measuremens in Bay 13 Evaluated by Extreme Value Statistics
Micrometer Pit Depth Measurements
UT Pit Depth Measurements
Linear (Micrometer Pit Depth Measurements)
Linear (UT Pit Depth Measurements)
y = 0.0918x + 0.2982
R^2 = 0.9535
y = 0.0576x + 0.0986
R^2 = 0.9003
Pit Depth in mils
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-2
-1
0
1
2
3
4
reduced variate