IDEQ Contract No. C771   Work Order IDEQ.002185.019

WASTEWATER LAGOON SEEPAGE TEST STATISTICAL REVIEW

STATISTICAL REVIEW OF SEEPAGE TESTS AND ASSOCIATED
CALCULATIONS IN SUPPORT OF COMPLIANCE WITH DEQ'S
*GUIDANCE FOR EVALUATING WASTEWATER LAGOON
SEEPAGE RATES*

OCTOBER 2011

Prepared for:

IDAHO DEPARTMENT OF ENVIRONMENTAL QUALITY
State Office
1410 N. Hilton Street
Boise, Idaho 83706

Prepared by:

720 Park Boulevard
Boise, Idaho 83712

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1.0   INTRODUCTION

The Idaho Department of Environmental Quality (DEQ) is evaluating the guidance and process by which wastewater lagoons are regulated. The Idaho DEQ requested URS to perform a statistical review of the seepage tests and associated calculations in support of compliance with DEQ's "Guidance for Evaluating Wastewater Lagoon Seepage Rates". The intent of this effort is to develop defensible methods to determine whether seepage test submittals meet the Idaho "Wastewater Rules", IDAPA 58.01.16.493.

2.0   DOCUMENT REVIEW

The Idaho DEQ provided URS with three primary documents for review and consideration during the statistical analysis. These are:

   a.   Idaho Department of Environmental Quality's "Guidance for Evaluating Wastewater Lagoon Seepage Rates", Revised April 2009;
   b.   Ham, J.M. Uncertainty Analysis of the Water Valance Techniques for Measuring Seepage from Animal Waste Lagoons. J. Environ. Q. (2002) 31:1370-1379;
   c.   DEQ's "Seepage Calculation Spreadsheet"

These documents provide a basis for existing guidance, techniques, and technical approaches used for regulating wastewater lagoons under Idaho law. URS has reviewed these documents and applied them during the statistical analysis.

3.0   PURPOSE AND SCOPE

DEQ tasked URS with the following technical assignments:

   1.   Evaluate DEQ's method for incorporating sampling error and equipment error in seepage tests for manual and electronic equipment and testing, based on the "5-day" and "15-day" tests documented in the DEQ spreadsheet http://www.deq.idaho.gov/assistance-education/for-engineers-developers/guidance.aspx.

   2.   Develop a statistically-based method for DEQ to evaluate whether the seepage rate determined during the test exceeds the regulatory threshold.

URS has completed the tasks and the findings are presented in this document.

4.0   REVIEW OF SPREADSHEET ERROR CALCULATIONS

4.1   MANUAL APPROACH

4.1.1 Error Types

Error calculations in the DEQ spreadsheet for the manual approach generally relate to three types of error:

(1) *Error associated with accuracy limitations of individual pieces of measurement equipment.* Data obtained from the equipment are used to calculate parameters in the seepage rate equation, which propagate errors associated with the inaccuracy. Examples of parameters subject to measurement error are the lagoon surface elevation change (ES), net lagoon evaporation (I_L), and evaporation pan error (Epan). The ES and Epan errors are calculated based on the manufacturer's stated accuracy for the device used, taking the square root of the summed squared errors associated with the starting and ending measurements for a time period. This follows the root-sum-error (RSS) approach recommended by Ham 2002. The *I_L error* is calculated by multiplying the P coefficient, the Epan error, and the precipitation factor (which is typically 1.0). The unit of the errors involved is +/- inches per day;

(2) *Total error associated with the various types of equipment used to measure water surfaces during the test and applied to a given day's seepage rate estimate.* This error is termed *Equipment Error* in the DEQ spreadsheet, a term which sums the ES and I_L errors. The error is calculated by taking the square root of the summed squares for ES and I_L and dividing this quantity by the number of days between measurements, as per Ham 2002. The unit of the error is +/- inches per day;

(3) *Total error associated with the measurement of water surfaces (lagoon and evaporative pan). This error is termed Sampling Error in the DEQ spreadsheet, a term which sums the standard error of the lagoon and evaporative pan measurements.* The error is calculated in the spreadsheet by taking the square root of the summed squares of the standard error of the mean for the daily replicate hook gauge measurements for the stilling well (lagoon) and the evaporative pan. The squared error for the evaporative pan term (I_L) is also multiplied by the P coefficient before summation with the squared ES error.

4.1.2 Assessment of Error Propagation

The spreadsheet is well organized and generally does a good job of producing the appropriate calculations and statistics needed for decision-making. Based on the analysis, URS recommends the following modifications:

a. The formula for ES error twice sums the square of the value in J9. This presumes that the equipment used to measure the water level in the lagoon is the same equipment (or has equal accuracy) for the equipment used to measure the water level in the evaporative pan. As a quality assurance feature, the spreadsheet would benefit if the values in J9 and J11 were summed in the (perhaps unusual) case where equipment of differing accuracies is used.

b. The formulas in cells D35 through D39 in the "Example 15-Day" tab contain a syntax error and should be changed. A "plus" (+) symbol is used instead of a comma in the SUMSQ function to calculate the sum of squares. This syntax error produces a higher value for the sum of squares than is appropriate. A simple example can illustrate the effect. If the squares of data values 2 and 3 are to be summed, the result should be $2^2 + 3^2$

= 13. Using the existing syntax, the result would be $(2 + 3)^2 = 5^2 = 25$. Thus, the current formula overstates the equipment error.

c. Two modifications to the error calculations that contribute to the "Best Case" and "Worst Case" intervals are recommended. At the present time, the average equipment error$_{SR1}$ (Eqt$_{Sr1}$) is being calculated but the average sampling error$_{Sr1}$ (Samp$_{Sr1}$) is not being calculated. Samp$_{Sr1}$ should be calculated and could be displayed in cell E41. In order to calculate the appropriate error interval around the average daily seepage rate, which is calculated in cells C41 and C42, the calculation should sum the squares of the errors associated with equipment (Eqt$_{Sr1}$) and sampling (Samp$_{Sr1}$), then take the square root. The result could be calculated in cell E42, replacing the existing calculation. This value should be subtracted from the value in cell D41 to obtain the lower value of the uncertainty interval ("Best Case"). Likewise, the result should be added to the value in cell D41 to obtain the upper value of the uncertainty interval ("Worst Case"). The application of this uncertainty interval is described in Section 5.0 under Category Determination (Section 5.2.3).

Note that the modifications recommended in (b) and (c) will result in a smaller interval than is currently calculated. Using the approach above, the uncertainty interval around the average seepage rate of 0.1288 inches per day is +/- 0.0076. This produces an interval between 0.1212 and 0.1364 inches per day, which is smaller than the "Best Case" and "Worst Case" values in the spreadsheet.

The testing performed by Ham 2002 accumulates data on different time scales (daily point-in-time vs. hourly averages, etc.). Because of this , the duration of the test has an significant effect on the uncertainty surrounding the seepage rate. Ham 2002 recommends dividing the seepage errors by the length of the test. URS understands the manual test is to be composed exclusively of point-in-time measurements separated by two days of inactivity. URS recommends DEQ consider adjusting the calculation of Eqt$_{Sr1}$ as the error is incurred only at the initial and final measurement in the example. In contrast, the Epan error occurs five times because refilling introduces a discrete error event, which is also a point-in-time measurement.

4.2   ELECTRONIC APPROACH

The spreadsheet calculations for the electronic approach follow many of the procedures, yet incorporate adjustments recommended by Ham 2002 for measurements taken on different time scales.

1. Both the sampling error and the equipment error for the electronic testing are larger (almost two orders of magnitude) than the equivalent errors in the manual testing. Because the electronic testing is assumed to be more accurate, this increase is unexpected. URS recommends the equipment error formula (D65 through D66) be divided by the square root of the number of seepage tests (6) as the error is being over-propagated. This will produce a value of 0.0147, which is more in line with the manual testing error.

---

2. The sampling error is being calculated based on the variability of the seepage estimates during the course of the day. For the manual test, the sampling error is based on measurement variation of measurements of water levels in the lagoon and evaporation pan. Thus, two different calculation systems are being used. The sampling error is also approximately as large as the seepage rate limit for lagoons built after April 15, 2007. With sampling errors this large, permittees have little chance of being compliant even if seepage is almost zero. The sampling error used in the manual calculations was based on hook gauge measurements, which are not taken in the electronic approach. Even though the sampling error is not included in the interval calculation, URS recommends renaming Sampling Error to be *Daily Seepage Rate Variation*.

5.0   STATISTICAL APPROACH FOR SEEPAGE COMPLIANCE DETERMINATION

5.1   CURRENT APPROACH

The current DEQ guidance states that measurements must be taken and seepage rates must be calculated until a *consistent pattern* is evident. A consistent pattern is defined in the DEQ guidance as:

> *The calculated seepage for each test period shall be within 20% of the calculated seepage for the four previous test periods.*

If an inconsistency is noted, testing is to continue until a consistent pattern (as defined) is achieved.

The current approach offers a clear-cut decision rule for compliance determination by applying the 20% rule. However, estimating lagoon seepage rates is a complex endeavor, involving multiple measurement types as well as uncontrolled dynamics that are often not well understood or quantifiable. Existing guidance offers a single, arbitrary metric as a firm basis for decision-making. Whereas the consistency metric is necessary, it is not sufficient to account for the complexity of the seepage rate determination. It is also subject to question due to the choice of a subjective interval range.

Despite their vulnerability to challenge, arbitrary standards, professional judgment, and other methods that do not provide an objective "bright line" for decision-making or compliance determination should not be excluded from the compliance determination process. Other approaches that offer more definitive metrics should also be included to add defensibility in DEQ decision-making. Thus, the existing approach contains certain deficiencies, but offers an opportunity to augment and amplify the process to include a multi-perspective approach that includes objective statistical metrics. In this way, the inherent complexity of seepage evaluation can include appropriate metrics that balance and include professional judgment, arbitrary standards, and defensible statistical methods.

5.2   PROPOSED APPROACH

A multi-metric approach to compliance determination is proposed. The components recommended create a decision rubric to assist both DEQ and seepage testers. With a better understanding of how DEQ will judge a seepage rate submittal, seepage testers are able to submit more complete packages, allowing DEQ to make faster and more accurate determinations. The metrics are (1) Compliance; (2) Completeness/Validity, (3) Category; (4) Consistency; (5) Comparability; (6) Data Sufficiency; and, (7) Data Quality. Table 1 summarizes these components and their functions. Detailed descriptions of the rubric components are provided below.

Lagoon Seepage Assessment Rubric		
Parameter	**Type**	**Description and Objectives**
Compliance	Objective	Assess whether the anticipated average seepage rate exceeds the regulatory limit (0.125 or 0.25 inches per day) and screen to a first tier compliance category status (compliant/non-compliant)
Completeness	Objective	Determine if a sufficient number of calculated seepage rates exist to make a valid decision for compliance determination
Category	Objective	Evaluate the impact of the errors/uncertainties on Category assignment; make a determination whether the lagoon status is Category 1, 2, 3, or 4 (second tier)
Consistency	Quantitative with Arbitrary Limits	Determine if the most recent seepage rate conforms to the +/- 20% standard
Comparability	Professional Judgment/ Subjective	Evaluate the degree to which levels in the lagoon and evaporative pan compare. Conclusion is based on the time series plot in the DEQ spreadsheet
Data Sufficiency	Professional Judgment/ Subjective	Determine if the minimum number of data/measurements obtained during the course of the testing were obtained; and,
		Determine if sufficient hook gauge measurements were obtained
Data Quality	Professional Judgment/ Subjective	Determine whether data were collected at the proper times/intervals and whether appropriate procedures and methods were employed.

Table 1: Decision Rubric Components

5.2.1 Compliance

The ultimate standard for seepage rate compliance is specified in DEQ guidance in the form of an average allowable seepage rate per day from a lagoon. According to Idaho Wastewater Rules, this is either 0.25 inches per day or 0.125 inches per day, depending upon whether the facility was constructed and approved before April 15, 2007 or after April 15, 2007.

When the lagoon seepage rate testing has been completed, the seepage rates for each time period are averaged in the DEQ spreadsheet to obtain an expected daily seepage rate. The average must then be compared to the appropriate regulatory limit. If the average seepage rate is less than or equal to the applicable regulatory limit, the lagoon is deemed conditionally compliant. If the average seepage rate is greater than the regulatory limit, the lagoon is considered conditionally non-compliant. Note that individual seepage rates for a time period may exceed the regulatory limit; however, as long as the average seepage rate does not exceed the regulatory limit, the conditional status of compliant may be retained.

Whereas the compliance status is conditional until other metrics are brought into play, the compliant/non-compliant determination is an objective measure. Additional objective measures must be applied to confirm or refute the initial compliance status.

5.2.2   Completeness/Validity of Seepage Rate Estimates

5.2.2.1 Completeness Issues

The existing guidance and accompanying spreadsheet enable the permittee to calculate seepage rates at regularly-spaced time intervals. Ham (2002) indicates a minimum of five seepage rate calculations is necessary to ensure that stabilization of measurements has occurred and that errors in the estimated rates are unlikely to cause decision errors. DEQ guidance currently reflects the need for a minimum of five calculated seepage rates.

Despite a thorough, sophisticated, and scientifically-based approach to lagoon seepage monitoring, the Ham rule-of-thumb of five days is based ultimately on professional judgment, which could be called into question in certain situations. To mitigate this potential issue, appropriate statistical approaches used in other regulatory environments may be applied to seepage rate assessments. These approaches answer very basic and fundamental questions such as, "Is there a minimum number of data points that provides a valid test?"; and, "How can DEQ determine if a sufficient number of data points are available to assess if a statistically-valid has been submitted?"

5.2.2.2 Completeness Metrics

The United States Environmental Protection Agency's (USEPA) *Data Quality Objectives* (DQO) guidance recommends using a *decision performance goal diagram*, also known as a *discomfort curve*, to determine data completeness and validity of a calculated average value for a parameter of interest. In general, the discomfort curve is useful when an average value of a parameter of interest (such as a seepage rate) is calculated and compared to a fixed regulatory standard (e.g. 0.25 or 0.125 inches per day).

The discomfort curve is a type of statistical power curve (EPA 2006) used to determine sample sufficiency in the DQO process. It provides the confidence associated with the decision as well. The curve manages simultaneously two types of uncertainty and decision error. These are *Type I error*, or deciding a regulatory limit has been exceeded when it has not, and *Type II error*, deciding the regulatory limit has not been exceeded when it has (Myers 1997). For the Idaho DEQ, a Type I error would be declaring a lagoon non-compliant when it is compliant; a Type II error would be declaring a lagoon -compliant when it is non-compliant. Statistical confidence is associated with the Type I error rate and statistical power is associated with the Type II error rate. These error rates are controlled in the discomfort curve.

Three principle dynamics underlie and influence the discomfort curve results. The first relates to the difference in value between the average and the regulatory limit. Intuitively, it makes sense that fewer samples are necessary to prove a mean (average) value is less than a regulatory limit when the mean is far below the limit. Conversely, when the mean is close to the limit, more samples are needed to prove compliance. Second, the discomfort curve factors in the variability (standard deviation) of the data. If the variability is high, more uncertainty exists and more samples will be required than for less variable data sets. Third, as higher levels of statistical confidence are desired, more samples are required.

5.2.2.3 Metric Implementation

Figure 1 is a discomfort curve for a seepage test with an average seepage rate of 0.095 inches per day, a standard deviation on the seepage rate of 0.0133, and a regulatory limit of 0.125 inches per day. Tolerable Type I (labeled *alpha* on the figure) and Type II (labeled *beta* on the figure) error rates are set at five percent each. The graph indicates that with these input parameters, four samples are necessary to provide a valid average seepage rate. In other words, four samples provide a 95 percent confidence that the seepage rate will not exceed the regulatory limit. Assuming the Ham/DEQ minimum of five seepage rates was obtained, sufficient seepage rates have been provided as evidence of compliance and to satisfy decision error constraints. Note that if the variability had been higher, more than four samples would be necessary. Similarly, if decision errors need to be limited to one percent (99 percent confidence), additional samples would be required. Type I and Type II error rates may be relaxed or tightened based on DEQ objectives.
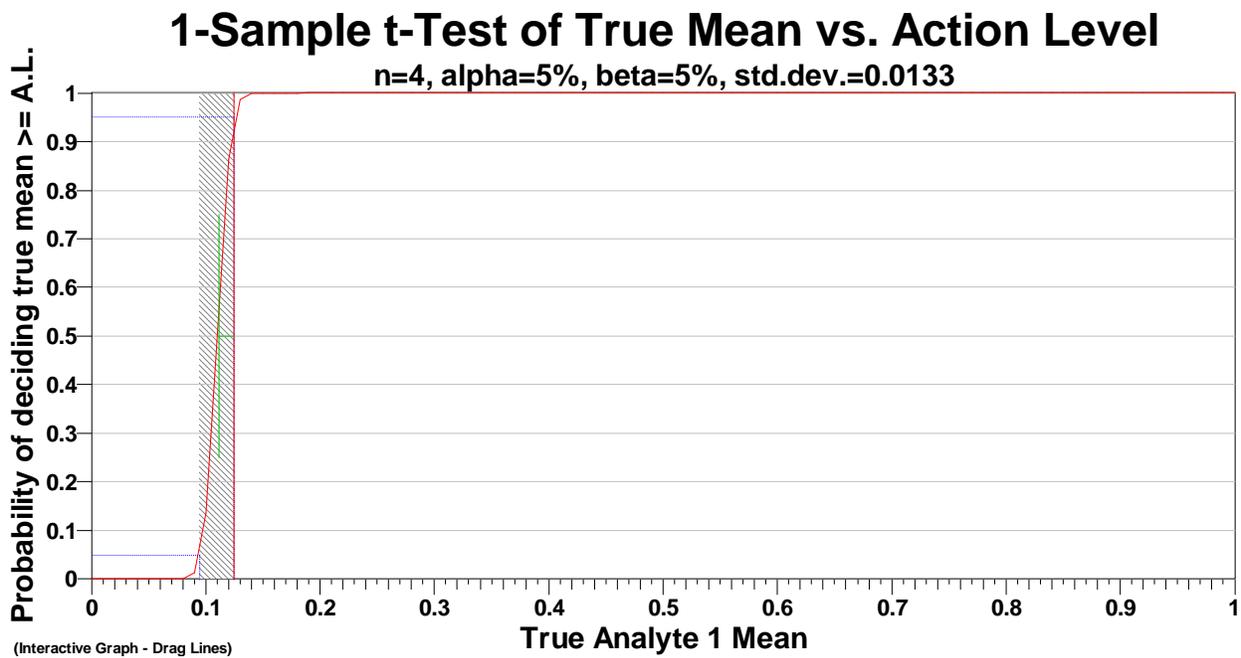

Figure 1: Discomfort Curve for a Seepage Test with an Average Seepage Rate of 0.095 Inches per Day

Figure 2 is a discomfort curve with an average seepage rate of 1.05 inches per day. In this case, the average seepage rate is closer to the regulatory limit than in Figure 1. This means it will be more difficult to distinguish with confidence whether the true seepage rate is below 0.125. Consequently, 21 seepage rates are required to obtain a valid and complete average. With only five seepage rate values, insufficient data are available to pass the test, even though on face value the average seepage rate is below the regulatory limit. As in Figure 1, if the standard deviation of the seepage rates were higher or the decision error limits tightened, more than 21 samples would be needed.
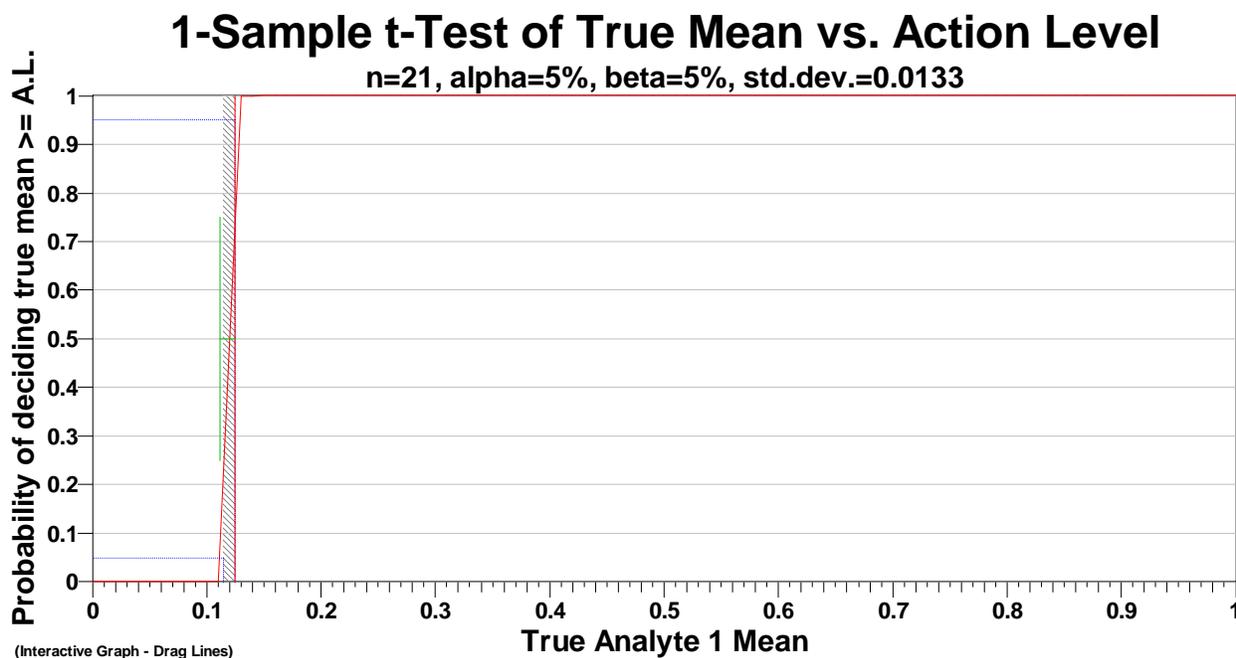
1-Sample t-Test of True Mean vs. Action Level
n=21, alpha=5%, beta=5%, std.dev.=0.0133

(Interactive Graph - Drag Lines)

Figure 2: Discomfort Curve for a Seepage Test with an Average Seepage Rate of 1.05 Inches per Day

Figures 1 and 2 were created using the Visual Sample Plan (VSP) software developed by the Pacific Northwest National Laboratory (PNNL) under the auspices of the United States Department of Energy (DOE). VSP is public domain software and may be downloaded without charge at http://vsp.pnnl.gov/. URS recommends that DEQ obtain the VSP software and use it in the Completeness/Validity analysis for lagoon seepage rate submissions. Permittees could also benefit by using VSP so they have a better indication of their potential compliance status before submission.

In summary, the discomfort curve uses an EPA-approved statistical power curve to incorporate a fixed regulatory limit, a calculated average seepage rate, and the variability around the seepage rate to create a defensible and objective "bright line" number of test days needed for completeness of seepage rate testing.

5.2.3 Category Determination

The initial compliance determination offers only a conditional status (first tier). In order to finalize and categorize the compliance status, the equipment and sampling uncertainties need to be applied to the decision-making process, to establish a more substantive and informative category determination (second tier).

Figure 3 shows a conceptual model of the category determination approach. The small, square boxes on the diagram represent valid average seepage rates (i.e., averages that conform to the completeness requirement that have been validated by the discomfort curve analysis). The boxes contain equal-length vertical error bars that represent the amount of error or uncertainty contributed by the equipment and sampling errors calculated in the DEQ spreadsheet.
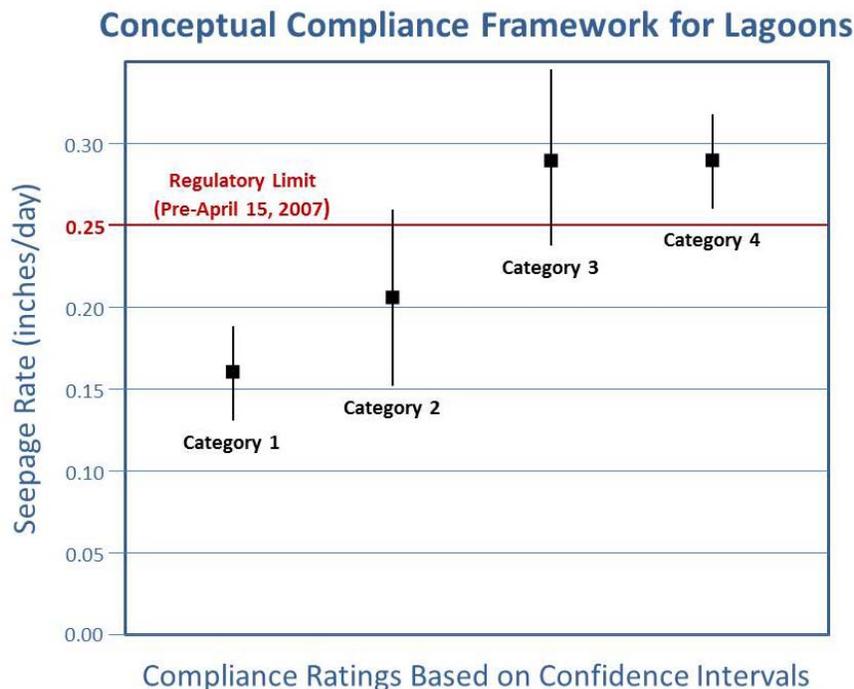
Conceptual Compliance Framework for Lagoons

Compliance Ratings Based on Confidence Intervals

Figure 3: Conceptual Model for Categorization in Compliance Assessment for Lagoons Constructed Prior to April 15, 2007

To be classified as Category 1, the average must be below the regulatory limit of 0.25 inches per day and the highest point on the upper error bar must fall below the regulatory limit. This is the most stringent but also the most desirable classification category. Category 2 also has a valid average that falls below the regulatory limit, but the error bar extends above the regulatory limit. Thus, the uncertainty associated with equipment and sampling invokes a caveat for decision-making (i.e., there is a possibility the lagoon is non-compliant). This aspect of the categorization process (i.e., the possibility of shifting from a Category 1 status to a Category 2 status because less accurate equipment was used) should encourage permittees to employ accurate equipment in order to avoid slipping from a Category 1 to a Category 2 classification.

A similar approach applies to lagoons constructed and permitted after April 15, 2007 (Figure 4). The process is the same; the difference is a lower regulatory limit.

5.2.4 Consistency
As currently stated in the DEQ guidance, the permittee must demonstrate consistency of seepage rate measurements. As defined in the current DEQ guidance, consistency compares the last calculated seepage to an interval around the average seepage rate of the preceding four calculations. At the discretion of DEQ, more than four seepage rates may be used. The interval is equal to plus-or-minus (+/-) 20 percent of the average value. If the last calculated seepage rate for a time period falls within the +/- 20 percent interval, consistency is considered to be proved.
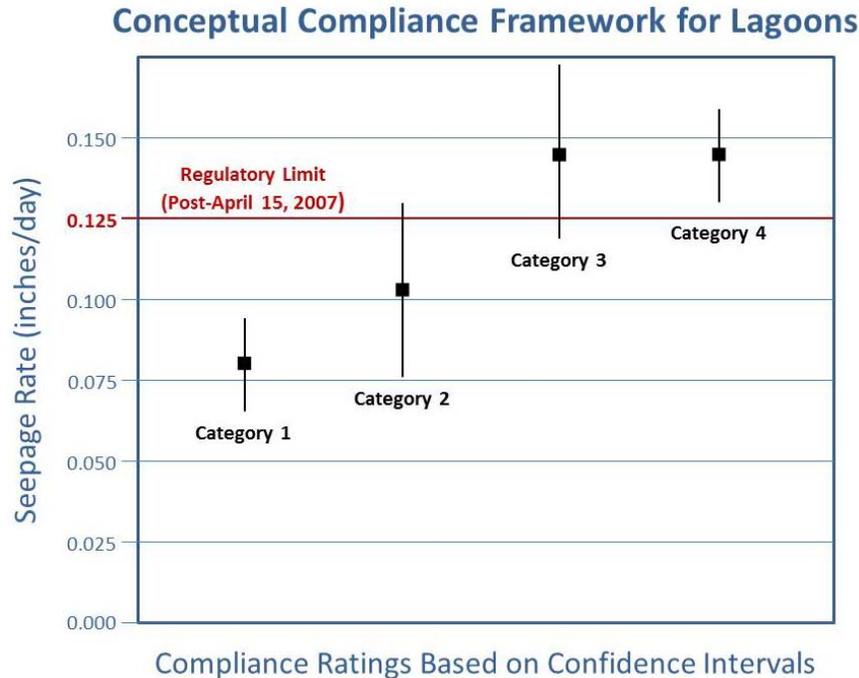
Conceptual Compliance Framework for Lagoons

Figure 4: Conceptual Model for Categorization in Compliance Assessment for Lagoons Constructed after April 15, 2007

No firm basis exists for the selection of a 20 percent interval versus, say, a 15 percent, 25 percent or other size percentile interval. In an attempt to employ more objective and statistical metrics, the statistical evaluation tested the applicability of using a 95 and 99 percent upper and lower confidence limit (LCL/UCL) interval. A range of expected and potential variability values (i.e. standard deviations) was applied to determine the sensitivity of the approach. The results indicate a high probability that a 95 or 99 percent UCL would accept seepage test data that would be rejected by other metrics such as compliance, completeness, and the 20 percent interval. Additionally, LCLs frequently were less than zero. Therefore, the LCL and UCL approach is not recommended for implementation. Despite the arbitrary nature of the interval size (20 percent), it appears to be a reasonable and robust metric. Moreover, being an arbitrary metric, it allows DEQ some flexibility in decision-making for situations where the final seepage measurement falls outside the 20 percent interval, but not excessively. Such an exception might be made for a small lagoon that is far above or sequestered from aquifers of concern.

5.2.5 Comparability

The DEQ spreadsheet plots a time series graph of the amount of evaporation in the lagoon overlain on the amount of evaporation in the evaporation pan. Because these two water bodies are exposed to the same weather conditions, theory predicts these two graphs should be synchronized over time, with rises and decreases in evaporation rates coinciding. In practice, however, variations occur between the two rates and, thus, the lines are not always synchronized.

Experience has shown that physical, chemical, biological, or other factors may come into play during the course of the seepage testing. These factors may contribute to unsynchronized graphs. DEQ should exercise professional judgment to assess first the degree of comparability between

the two evaporation rates to determine if it is acceptable. Deviations should be evaluated to determine if influences were affecting results and, if so, whether their impact can be considered tolerable. The evaluation may also determine, in some cases, that the correlation between the two evaporation rates is not up to DEQ standards. This metric may also be linked to consistency. For example, if the comparability is poor in the initial portion of the testing period but synchronizes well as the test progresses, a more favorable assessment might be given in contrast to a test where the initial comparability is good but degrades as the test continues.

5.2.6 Data Sufficiency

Multiple types of data are required to support the calculation of the seepage rates for each time period as well as for determining the equipment and sampling errors. The DEQ spreadsheet submitted by the permittee contains the summary of data collected and analyzed. DEQ currently evaluates the data to determine if, in their judgment, sufficient data exist. For example, if insufficient hook gauge replicates measurements were taken in either the lagoon or evaporation pan, or if other data are incomplete, professional judgment should be exercised by DEQ to determine the impacts on decision-making. This determination is semi-quantitative but is important to the ultimate determination.

As an aid to permittees, DEQ may consider developing a checklist for data sufficiency. Examples of this concept for manual and electronic testing appear in Tables 2 and 3 respectively.

Data Sufficiency Checklist				
Manual Test	**Status**	**Device**	**Minimum Measurement Frequency**	**Comments**
Temperature	Required	Temperature Recorder	Hourly	
Evaporation Pan	Required	Hook Gauge	Daily	
Lagoon Surface	Required	Hook Gauge	Daily	
Wind Speed	Recommended	Anemometer	Discretionary	
Precipitation	Required	Rain Gauge	Per Event	
Effluent/Influent Flow	Required	Flow Meter	Per Event	

Table 2: Potential Data Sufficiency Checklist for Manual Testing

Data Sufficiency Checklist				
Electronic Test	**Status**	**Device**	**Minimum Measurement Frequency**	**Comments**
Temperature	Required	Temperature Recorder	Hourly	
Evaporation Pan	Required	Transducer	Every four (4) hours for five (5) days	
Lagoon Surface	Required	Pressure Transducer	Every four (4) hours for five (5) days	
		Ultra Sonic		
		Float-Operated		
Wind Speed	Recommended	Anemometer	Discretionary	
Precipitation	Required	Rain Gauge	Per Event	
Effluent/Influent Flow	Required	Flow Meter	Per Event	

Table 3: Potential Data Sufficiency Checklist for Electronic Testing

5.2.7 Data Quality

Data of sufficient quality as well as quantity are required for compliance determination. DEQ performs a data quality review as part of its current process. The DEQ evaluates whether data were collected at the proper times or intervals and whether appropriate procedures were used. This includes, but is not limited to, a review of the quality assurance plan, equipment calibration, calculation performed outside the DEQ spreadsheet, equipment setup, and other items, many of which appear in the DEQ guidance. As with data sufficiency, this determination is semi-quantitative but is important to the ultimate determination.

6.0   APPLICATION OF THE PROPOSED STATISTICAL APPROACH

Figure 5 is a decision logic diagram based on the decision rubric. It provides a visualization of the proposed decision-making process. Using Figure 5, URS applied Example 15-Day test data in the DEQ spreadsheet to the process. The results are shown in Table 4.

The initial compliance determination is that the lagoon is non-compliant. This is because the average seepage rate of 0.1288 inches per day exceeds the regulatory limit of 0.125 inches per day. Completeness is not applicable because the discomfort curve is not designed to handle averages above the regulatory limit. Note that this condition does not preclude the permittee from continuing the seepage testing, where new data may demonstrate the average seepage rate is below 0.125 inches per day. Continued testing, however, does not ensure compliance; in fact, it may serve to prove more definitively that the lagoon is non-compliant. Thus, the permittee would continue seepage testing at risk.

Based on DEQ spreadsheet calculations, the lower error bar terminates below the regulatory limit. This suggests the possibility that true seepage rate is below the limit and that the permitee experienced "bad luck" during the test, hence the potential desire for a permittee to continue seepage testing. However, using inferior (less accurate) equipment may increase the length of the error bar. Permittees could see this as an opportunity to make a non-compliant lagoon look potentially compliant by increasing uncertainty.

According to the consistency metric, the lagoon seepage data fall within the prescribed 20 percent interval. A specious argument could be made claiming overall compliance based on consistency. In fact the reverse is true. Because the data have stabilized, this adds to the certainty that the average seepage rate exceeds the regulatory limit. Finally, the comparability, data sufficiency, and data quality metrics have all been deemed acceptable.
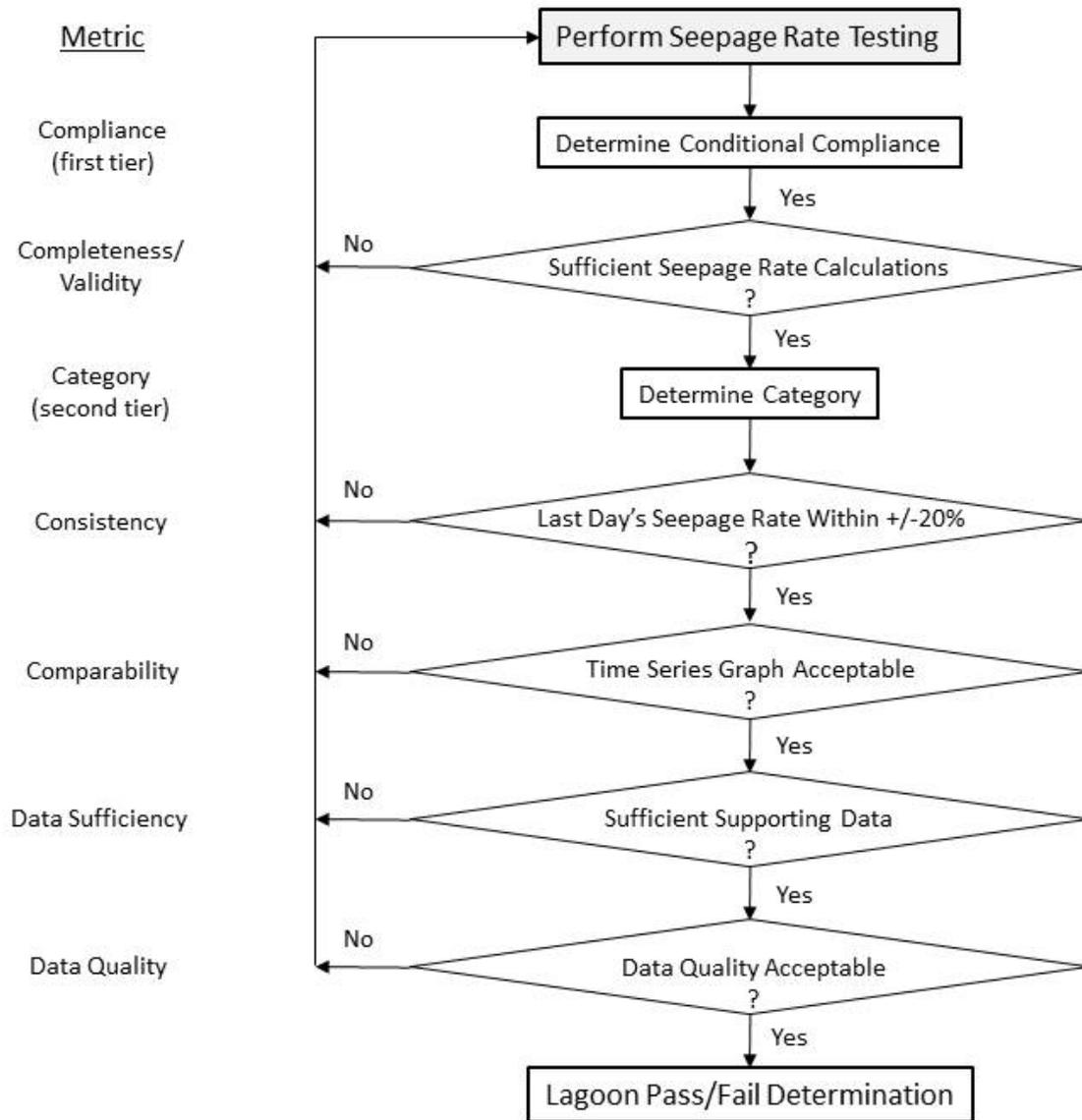
Figure 5: Example Logic Flow Diagram for the Decision Rubric

Lagoon Seepage Assessment			
Parameter	**Result**	**Comments**	**Basis**
Compliance	Non-Compliant	Possibility Exists that the True Seepage Rate is Compliant	Regulatory Limit
Completeness	Not Applicable	Mean Exceeds Regulatory Limit	Discomfort Curve
Category	3	Lower Error Bar Extends Below Regulatory Limit	Spreadsheet Error Calculations
Consistency	Pass	Last Day's Sample Within Range	+/- 20% Interval
Comparability	Acceptable	Minor Deviations	Time Series Graph
Data Sufficiency	Acceptable	No Comment	Minimum Data Requirements
Data Quality	Acceptable	No Comment	Collection Schedule/Procedeures

Table 4: Results of the Decision Rubric as Applied to the DEQ Example 15-Day Test

7.0 OBSERVATIONS AND RECOMMENDATIONS

Based on the statistical review, URS makes the following observations and recommendations:

1. The monikers "5-Day" and "15-Day" tests seem to give a false impression as to the ultimate duration of the testing. URS recommends calling the 5-Day test the "Electronic Test" and the 15-Day test the "Manual Test". The completeness and consistency requirements can easily increase the number of days required for testing; thus, titling a test based on length has a downside.

2. There does not appear to be a compelling statistical reason to perform manual testing with two-day hiatuses rather than taking measurements on five consecutive days. The consecutive day approach may have other limitations or risks that may make it more advantageous to space out the testing. A consecutive day approach would also reduce the likelihood of a precipitation event. Also, considering that electronic testing may be performed on consecutive days, requiring hiatuses for the manual approach appears to be inconsistent and may generate complaints.

3. In the current approach, precipitation of a duration greater than four hours triggers consequences. URS recommends a reevaluation to determine if using precipitation amount may be more appropriate or if both factors need to be represented. Moreover, seepage tests without precipitation events are highly preferred over those with precipitation events. This is an additional argument for consecutive-day testing.

4. For electronic testing, averaged measurements should represent a relatively short time interval. For example, water level measurements taken over a four-hour time frame should not be averaged. If the measurement device is able to provide one reading per second, a better approach would be to average one-second readings over, say, a one to five-minute (or other appropriate) time period at the start of each time interval. This approach reduces potential biases and errors that could be caused by evaporation over the four-hour time period.

    One must assume that seepage and evaporation are ongoing processes, which are measurable over one- to four-hour periods. "Snapshot" time measurements, averaged over say one to five minutes, minimize the potential biases introduced by averaging over a time period where measurable changes are known to occur, but occur to an unknown degree.

5. Daily Seepage Rate Variation could be included in the interval estimate for category determination. Summing the squares of cells D66 and E66 and then taking the square root of this quantity would give an alternative interval. The existing interval is +/= 0.0928; incorporating the Daily Seepage Rate Variation would produce an interval of +/- 0.0673. The existing Best/Worst Case summations on line 66 should be replaced or eliminated.

6. If a seepage test exceeds five days in length, DEQ should retain the use of the previous four days seepage rates for consistency calculations as the default option. However, DEQ

is not restricted from using more than four days of previous seepage data to determine consistency, which can be invoked at DEQ's discretion.

7. The spreadsheet should calculate and post the 20 percent consistency interval along with a cell that indicates whether the test statistic for consistency is a "Pass" or a "Fail".

8. The spreadsheet should have a location to post the number of seepage tests needed, based on completeness analysis from VSP.

9. The spreadsheet should calculate the standard deviation for the seepage rate data. This value can then be used directly in VSP to calculate completeness.

10. In the 15-Day calculations, columns 35 through 39, the value of three (3) is valid for the existing test. If seepage rates are calculated based on data measurements taken over the course of one (1) day, the denominator of the equation should be changed to one (1).

11. On the 5-Day spreadsheet, consider plotting (in overlay fashion) the daily seepage values on the Seepage S_{R1} Chart. This will provide a visual measure of variability/consistency.

12. On the 5-Day spreadsheet, Seepage S_{R1} Chart, the y-axis is labeled "inches"; should be "inches per day".

13. On the 5-Day spreadsheet, it might be useful to calculate the average standard deviation of the six seepage rates for the day (Column Z). A consistency test could also be run as a "soft" metric, similar to the consistency check on the average. A similar exercise could be performed for the coefficient of variation. In general, since the seepage rate is being calculated multiple times per day, these intermediate calculations might be examined and explored to see if they reveal other information.

14. To the extent DEQ accepts URS' recommendations, URS recommends further that they be incorporated in the DEQ guidance for lagoon seepage testing. One opportunity that appears to exist is the augmentation of Section 6, "Required Data". This topic could contain extensive descriptions, but is currently very brief. Some of the proposed items are consistent with the theme of Section 6 and could be included there.

8.0   REFERENCES

Ham, J.M. 2002. *Uncertainty Analysis of the Water Balance Technique for Measuring Seepage from Animal Waste Lagoons*. Journal of EnvironmentalQuality. 31:1370-1379.

Myers, J.C. 1997. *Geostatistical Error Management: Quantifying Uncertainty for Environmental Sampling and Mapping*. New York: John Wiley & Sons.

U.S. EPA. 2006. *Guidance on Systematic Planning Using the Data Quality Objectives Process*. EPA/240/B-06/001. Office of Environmental Information.