

Determination of the Requirement to Include Water Quality Standards-Based Limits in CDPS Permits Based on Reasonable Potential

Procedural Guidance

February 2003

Reasonable Potential Guidance

- I Introduction**
 - A. Background**
 - B. Basis of RP Determination Described by this Document**
- II. Quantitative Evaluation of Reasonable Potential**
 - A. Overview**
 - B. Method of Quantitative Evaluation**
- III. Qualitative Evaluation of Reasonable Potential**
- IV. Ongoing Evaluation of Reasonable Potential**
 - A. Pollutants with a Permit Limitation**
 - B. Pollutants with “Monitor Only” Requirements**

APPENDICES

- Appendix 1: Reasonable Potential Multiplying Factors Table**
- Appendix 2: Instructions for Calculating Coefficient of Variation for Data sets**
 - For which all values are above the detection limit**
- Appendix 3: Instructions for Calculating Coefficient of Variation for Data Sets with**
 - Values Below the Detection Limit**
- Appendix 4: Rationale**

GLOSSARY/ACRONYMS

Antidegradation Limit

An effluent limit that, when met, causes no significant degradation of the current water quality of a specified water body.

BPJ

Best professional judgment

Maximum Allowable Pollutant Concentration

This is the concentration of a specified pollutant in an effluent that is either included as a CDPS permit limit or calculated in the same manner as if it were to be incorporated as a CDPS permit limit. This concentration takes into account water quality standards, background concentrations, and available instream dilution. In some instances, the maximum allowable pollutant concentration may be derived from a Total Maximum Daily Load (TMDL) allocation or an antidegradation based effluent limit. The maximum allowable pollutant concentration is identified as the “calculated assimilative capacity” in the “Water Quality Assessment” which is sometimes included as an appendix to the permit rationale.

Colorado Discharge Permit System (CDPS)

Regulation No. 61. 5 CCR 1002-61.

Coefficient of Variation (CV)

A measure of relative dispersion around a mean that is applicable only when the mean is not equal to zero.

Maximum Estimated Pollutant Concentration

An estimate (using the statistical method described in this guidance) of the pollutant concentration in an effluent that exceeds the 99th percentile of the data set, at the 99% confidence level.

Multiplier

The statistically derived number taken from Appendix 1--Reasonable Potential Multiplying Factors Table (based upon *Technical Support Document for Water Quality-based Toxics Control* (USEPA 1991)) that is multiplied by the highest, non-outlier, value in the data set to arrive at the maximum estimated pollutant concentration.

Permittee

Any entity (individual, corporation, municipality, etc.) that holds a CDPS permit for the discharge of pollutants from a point source to state waters.

Practical Quantitation Limit (PQL)

The lowest concentration of a pollutant that can be measured reliably within specified limits of precision and accuracy under routine laboratory conditions.

Reasonable Potential (RP):

The likelihood that the concentration of a pollutant in a discharge would cause or contribute to an exceedance of water quality standards.

I. INTRODUCTION

A. BACKGROUND

The Colorado Discharge Permit System Regulations (Regulation No. 61, 5 CCR 1002-61) require that permit limitations be placed upon any discharged pollutant that causes or contributes to, or that has the reasonable potential to cause or contribute to, an exceedance of water quality standards (see Section 61.8(2)(b)(i)(A)).

The purpose of this document is to describe how the Water Quality Control Division (the “Division”) will determine whether pollutant concentrations in a discharge are such that the discharge has the “reasonable potential” to:

- 1) cause or contribute to an instream exceedance of a water quality standard or
- 2) in the case of reviewable or outstanding waters, cause or contribute to an exceedance of the significant concentration threshold or current water quality, respectively.

Using the criteria in this guidance, if the Division determines that a pollutant has “reasonable potential,” a water quality standards-based limitation for that pollutant will be included in the permit. If the Division determines that *no* reasonable potential exists, a limit will *not* be placed in the permit, although monitoring requirements may be placed in the permit under appropriate circumstances. This guidance does not apply to determination of RP for a discharge to cause or contribute to an exceedance of the narrative water quality standard for the purpose of deciding whether to impose a limit for whole effluent toxicity in a permit.

This document is written primarily for the use of those persons, both inside and outside of the Division, who are involved in the writing of permits issued under the Colorado Discharge Permit System (CDPS). It is intended to serve as guidance and, as such, the Division reserves the right to use best professional judgement in cases that differ from those anticipated by this policy. In such a case, if the permit writer chooses to deviate from the guidance, clear documentation for the deviation shall be included in the permit rationale.

B. BASIS OF RP DETERMINATION AS DESCRIBED IN THIS GUIDANCE

The reasonable potential determination (“the RP Determination”) may be done through both quantitative and qualitative analysis of pertinent data. The Division’s method for quantitative analysis of reasonable potential data is based upon a modification of the EPA’s guidance for establishing reasonable potential as found in Chapter 3 and Appendix E-1 of the *Technical Support Document for Water Quality-based Toxics Control* (USEPA 1991), hereafter referred to as the “TSD” method. The Division’s method for qualitative analysis of reasonable potential is unique to the Division, but based upon some of the general ideas presented in the TSD.

II. QUANTITATIVE EVALUATION OF REASONABLE POTENTIAL

A. OVERVIEW

The quantitative method for determining reasonable potential is a two step process that includes: 1) determination of the maximum value in the data set and 2) a statistical analysis of facility-specific effluent data to estimate the highest expected concentration of a pollutant in that effluent that, as appropriate, would also include identification of any outliers. The value of the estimated highest pollutant concentration varies with the number of samples in the data set and the variability of the data set. A highly variable data set with few data points is more unpredictable and thus will result in higher estimates of the highest expected pollutant concentration in the effluent. Conversely, a less variable data set with many data points is more predictable and will result in a lower estimate of the highest expected effluent concentration.

The variability is expressed as the “coefficient of variation” or “CV”. The coefficient of variation is calculated for each individual discharger’s data set as described in section B below.

Given the sample size and coefficient of variation (see section B), the Reasonable Potential Multiplying Factors Table (Appendix 1) is used to choose a “multiplier.” This multiplier is then applied to the highest pollutant concentration in the data set. The resultant product is the maximum estimated pollutant concentration—the pollutant concentration that exceeds the 99th percentile of the distribution.

Once derived using the method described above, the maximum estimated pollutant concentration in the effluent and any outlier value(s) are compared with the maximum allowable pollutant concentration calculated during the permitting process. The maximum allowable effluent concentration, the maximum concentration of a specified pollutant that can be discharged by an individual discharger without causing an exceedance of water quality standards, accounts for dilution, background pollutant concentration, and pollutant loading(s) contributed by other discharges. The maximum allowable pollutant concentration is listed as the “calculated assimilative capacity” in the “Water Quality Assessment” included as an appendix to the permit rationale.

If the maximum estimated pollutant concentration in the effluent or any outlier data value is greater than the maximum allowable pollutant concentration, then there is reasonable potential for an exceedance of water quality standards and a limit is placed in the permit. If the maximum estimated pollutant concentration in the effluent is less than the maximum allowable pollutant concentration, then the RP determination process continues with the qualitative portion of the analysis.

B. METHOD OF QUANTITATIVE ANALYSIS FOR RP

Step One-Determine the Pollutants of Concern

The permit writer, with the cooperation of the permittee, must use best professional judgement to determine the pollutants of concern (“POC’s”). POC’s are pollutants that might be expected in the effluent. POC’s may be:

- pollutants that have been detected in the effluent (through compliance monitoring, priority pollutant monitoring, optional monitoring, or other monitoring) in the last 5 years;
- pollutants with known sources;
- pollutants that are known to commonly occur in similar effluents;
- pollutants that are present in the influent or at other sampling points in the treatment or collection systems;
- pollutants that are present in the biosolids or other treatment residuals;
- other pollutants which, in the permit writer’s best professional judgement, may be found in the effluent.

Step Two- Determine Whether the Effluent Data Meet the Minimum Requirements

The effluent data should be assembled and checked to assure that the data meets the minimum requirements (see box below).

Data points for like statistically determined values should be grouped together. For example: all daily maximum data for a given parameter should be grouped together; all monthly average data for a given parameter should be grouped together, and all seasonal data for a given parameter should be grouped with like data in the same season. If a permittee does not have the required data, a compliance schedule may be placed in the permit to require the collection of the data. A final RP determination may then be postponed until the required amount of data is received.

If a permittee has more than the minimum data available, all recent data (5 years old or less) must be used for the analysis unless a shorter period of record is appropriate as determined by the Division.

Minimum Effluent Data Requirements

- The RP evaluation requires a minimum of 10 data points collected over a period of at least one year. Generally, each calendar quarter (Jan-March; April- June; July – September; October – December) must be represented by at least one data point.
- If applicable water quality standards are seasonal, samples must coordinate with the seasonal limits. (For example, when evaluating RP to exceed a June limitation, sampling should be done in June.)
- Data that was collected prior to significant changes in the service area, contributing sources, or plant operations; or other modifications that resulted in a change in effluent quality should not be included in the analysis. Significant changes in service area may include situations like the addition of a new type of industrial user. Significant plant changes may include expansions or changes in treatment process.

Handling Newly Discovered False Positives

Data submitted and certified on a discharge monitoring report is presumed to be valid data. However, it is possible for lab, sampling, data entry or other errors to produce false positive results that come to light only after results are reported. In such a case, the permittee may submit an amended discharge monitoring report with an explanation of the evidence. If the permittee can demonstrate, to the Division's satisfaction, that the result is truly a false positive, it may be removed from the permittee's record, and the RP analysis may continue without further consideration of the known false positive result.

It is important to note that the calculation of the maximum estimated pollutant concentration is only necessary in situations where the maximum pollutant concentration in the effluent data set has not exceeded the maximum allowable pollutant concentration. If the maximum actual pollutant concentration in a discharge has exceeded the maximum allowable pollutant concentration, then it is presumed that the discharge has the reasonable potential to cause or contribute to an exceedance of water quality standards for the specified pollutant. In such instances a permit limitation will be placed in the permit for the specified parameter and there is no need to proceed further with the analysis described in this guidance.

Step Three: Calculate the Coefficient of Variation:

Enter like data for the same parameter into the "RP Spreadsheet," following the instructions included. In most cases, the RP Spreadsheet is used to calculate the coefficient of variation. The RP Spreadsheet includes input instructions and can be found on the Division's common drive or obtained from the permit writer.

Some special situations where the permit writer may choose to modify the RP Spreadsheet method of calculating the coefficient of variation are outlined in the boxes below:

Calculating the CV for Data Sets that Follow a Normal Distribution

If there is evidence that the data follow a normal distribution instead of a lognormal distribution (which is the base assumption when doing this analysis) then, at the permittee's request and given valid documentation that the data set follows a normal distribution, the permit writer may calculate the coefficient of variation assuming a normal distribution according to the following formula:

$$CV = (\text{Standard deviation})/(\text{mean})$$

This calculation can be done using Microsoft Excel or similar software.

Calculating the CV for Data Sets with Values that are Below the PQL

For data sets that include data reported as "below the practical quantitation limit" ("PQL"), the coefficient of variation should be determined using the "Robust Log-Probability Regression" method in the "MDLWIN" software that can be found at www.practicalstats.com or on the Division's common drive. (The RP Spreadsheet should not be used for data sets with data points below the PQL.)

The data must meet **BOTH** of the following criteria in order to use the MDL Program:

- At least 3 data points must be at or above the detection limit.
- At least 30% of the data must be at or above the detection limit.

Again, at least 10 data points are required. Please see Appendix 3 for details on the use of this method.

If MDLWIN cannot be used because the data set contains too few data points above the PQL, the permit writer should use BPJ when determining whether limitations or monitoring requirements should be placed in the permit. If all of the data are less than the PQL then a "default" finding of no reasonable potential will be made unless the permit writer, considering the following qualitative criteria, determines that there is a basis for a finding of reasonable potential.

- the proximity of the maximum observed concentration to the maximum allowable pollutant concentration;
- the proximity of the PQL to the maximum allowable pollutant concentration;
- the items listed in Section II.B above, "Determining Pollutants of Concern"
- the items listed in Section III below, "Quantitative Evaluation of Reasonable Potential";
- any other relevant items

Where there are data both above and below the PQL, the permit writer shall evaluate the maximum value against the maximum allowable value as described below.

Managing Statistical Outliers

If the permittee believes that the data set used in the RP analysis contains values that are inconsistent with the remainder of the data (outliers) then, at the permittee's request, the permit writer may exclude the outlier from the calculation of the CV and the subsequent calculation of the maximum estimated pollutant concentration provided that:

- the permittee can provide valid statistical analysis that the value is a statistical outlier.

It is important to note that the outlier is only excluded from the statistical portion of the RP analysis-which is a tool to help predict future pollutant concentrations. Outliers are still included in the comparisons to the maximum allowable pollutant concentrations described in Situations A-C in Step Six below. If the exclusion of the outlier results in a "no RP" determination, the permit writer may require monitoring in lieu of imposing a limit in the permit. With a "no RP" determination, the Division expects all actual pollutant concentrations to remain below the maximum allowable pollutant concentration. Therefore, in order to assure that the actual pollutant concentration stays below the maximum allowable pollutant concentration, the permit writer, if appropriate, will include, in the permit, a requirement for the permittee to notify the Division if monitoring results exceed the maximum allowable pollutant concentration. Such notification will be in writing, within 30 days of the permittee's receipt of laboratory results. Upon such notification, the permit writer may reopen the permit and impose a limit for the pollutant.

Step Four: Determine the Appropriate Multiplier

Given the sample size and the coefficient of variation, use the Reasonable Potential Multiplying Factors Table (Appendix 1) to determine the multiplier.

For data sets with greater than 100 data points, use the multiplier given for 100 data points.

Step Five: Determine if the Maximum Estimated Pollutant Concentration Exceeds the Maximum Allowable Pollutant Concentration

For evaluating RP to exceed acute standards:

For each applicable parameter, calculate the maximum estimated pollutant concentration by multiplying the highest concentration in the daily maximum data set with the multiplier from the Multiplier Table. This is the acute maximum estimated pollutant concentration. Compare this value with the calculated maximum allowable effluent concentration that is based on acute standards.

For evaluating RP to exceed chronic standards:

For each applicable parameter, calculate the maximum estimated pollutant concentration by multiplying the highest concentration in the "monthly average" data set with the

reasonable potential multiplying factors. Compare this value with the calculated maximum allowable pollutant concentration that is based on chronic standards.

For more recently issued permits, the maximum allowable pollutant concentration is listed as the “calculated assimilative capacity” in the “Water Quality Assessment” included in the appendix to the permit rationale. It is important to note that in the case of discharges to streams with TMDL’s, the maximum allowable pollutant concentration is the TMDL allocation (if applicable for the parameter in question). In the case of discharges to receiving waters that are considered “reviewable waters” under antidegradation regulations, the maximum allowable pollutant concentration is the antidegradation based effluent limit (or other limitation set under antidegradation guidelines.)

Step Six: Use the RP determination to determine the permit outcome

Situation A

If:

The maximum value in the data set, including outliers, or the maximum estimated pollutant concentration is *greater* than the maximum allowable pollutant concentration

Then:

A limitation will be placed in the permit.

Situation B

If:

The maximum value in the data set, including outliers, or the maximum estimated pollutant concentration is greater than 50% of the maximum allowable pollutant concentration

Then:

No limitation should be placed in the permit unless a need is indicated by the criteria in the qualitative analysis portion of this document. Routine monitoring requirements should be placed in the permit at a frequency commensurate with the size of the discharge.

Situation C

If:

The maximum value in the data set, including outliers, and the maximum estimated pollutant concentration are *less* than 50% of the maximum allowable pollutant concentration

Then:

No limitation or routine monitoring requirements should be placed in the permit unless a need for monitoring is indicated by the criteria in the qualitative analysis portion of this document.

III. QUALITATIVE EVALUATION OF REASONABLE POTENTIAL

The statistical analysis of effluent data described above is a valid method to analyze effluent data for reasonable potential purposes. However, it is only valid as long as the conditions at the wastewater treatment plant are relatively stable and accurately represented by the effluent data used in the analysis. Therefore, even if the outcome of the statistical analysis fits into Situation B, above, the permit writer may choose to impose a limit and/or additional monitoring requirements in situations that present potential for significant uncertainties in pollutant concentrations at the wastewater treatment plant. Similarly, qualitative analysis may indicate that, where the outcome of the statistical analysis fits into Situation C, above, continued monitoring is appropriate. This additional information will help the Division verify whether any anticipated changes or additional uncertainties reveal data that more accurately predicts actual effluent concentrations. Additional data for this purpose can include the results obtained using appropriate water quality modeling, as described in section 61.8(2)(b)(i)(B) of Regulation 61.

Situations that may warrant additional monitoring include but are not limited to the following:

- 1) There are intermittent changes in pollutant concentrations in amounts that could affect the outcome of the RP determination but that, due to the timing of RP sampling, are not reflected in the RP data.
- 2) Before the next permit renewal, planned growth, planned additions of industrial users, or other foreseeable conditions are expected to increase pollutant concentrations. (An example is copper concentrations that increase in proportion to new domestic construction that uses copper piping.)
- 3) There are sources of pollutants whose maximum allowable pollutant concentration is below the PQL. (Please see Steps One and Three in Section II.B above.)

IV. ONGOING EVALUATION OF REASONABLE POTENTIAL

A. Pollutants with a Permit Limitation (Situation A)

The effluent data collected during the course of the permit should be used to make a new RP determination at the time of permit renewal using the RP procedures described in Sections II and III above. (The permit writer may need to require additional data if, at the time of permit renewal, there are additional parameters of concern.)

B. Pollutants with “Monitor Only” Requirements (Situation B)

The effluent data collected during the course of the permit should be used to make a new RP determination at the time of permit renewal using the RP procedures described in Sections II and III above. (The permit writer may need to require additional data if, at the time of permit renewal, there are additional parameters of concern.)

C. Pollutants with No Permit Limitation/No Routine Monitoring Requirements (Situation C):

Prior to or at the time of permit renewal, the permit writer will determine the appropriate parameters of concern and will contact the permittee to determine if there are any unreported data available that can be used to conduct the RP analysis. Where the permittee does not have the required amount of existing data to conduct the RP analysis, this contact will normally occur outside of the minimum one-year timeframe that would be required to collect data. In this situation the permit writer will normally find that the RP analysis cannot be completed at the time of permit issuance and, in a schedule of compliance, will require the permittee to conduct the appropriate monitoring and the RP

analysis will be completed after the data is submitted. If after two consecutive RP evaluations, a pollutant again falls into Situation C, absent any compelling qualitative information to the contrary, the permit writer should drop the pollutant from the list of parameters of concern.

Appendix 2:

Instructions for Calculating Coefficient of Variation for Data Sets with all Values Above the Detection Limit

The equation used to calculate the coefficient of variation for data sets with all values above the detection limit is the one given in Appendix III of the TSD Method (USEPA 1991).

The method assumes a lognormal distribution and gives the following calculation for the coefficient of variation:

$$\text{coefficient of variation} = cv(X) = [\exp(\sigma_y^2) - 1]^{1/2}$$

Where:

σ_y^2 = variance = $\Sigma[(y_i - \mu)^2] / (k-1)$;

$y_i = \ln(x_i)$ for $i = 1, 2, \dots, k$.

μ_y = mean = $\Sigma(y_i) / k$

k = sample size

The permit writer should carry out the above equation on an electronic spreadsheet. A sample spreadsheet may be found on the Division's common drive under the name "RP Spreadsheet. A printed version of the sample spreadsheet is also included in this appendix.

Appendix 3:

Instructions for Calculating Coefficient of Variation for Data Sets with Values Below the Detection Limit

Step 1: Assemble all like data for the same pollutant.

“Like data” means data generated for the same pollutant over the same type of interval. For example, all daily maximum data for copper over the review period should be grouped together; all monthly average data for copper over the review period should be gathered into a separate group. Likewise all seasonal data should be grouped together. For example all spring quarter data should be grouped together; all summer quarter data should be gathered into a separate group.

Make sure that all data used meets the Minimum Effluent Data Requirements given in the box in Section II.B of the guidance.

Step 2: Put like data into an electronic spreadsheet.

Using Microsoft Excel or similar software, create a separate, new spreadsheet for all like data for a given parameter. In column A enter the values of each data point or, if the result is less than the detection limit, enter the detection limit. In column B, for each individual data point in column A, enter a “1” if the data is not censored (above the detection limit); enter “0” if the data is censored (below the detection limit). You may enter up to 1000 censored data points and up to 1000 uncensored data points. You may enter multiple detection limits.

Do not add labels or other information to the file.

Save the document as a “.txt” (tab delimited file). Print the file and include it in the permit file.

Step 3: Download the mdlwin.exe file.

Download the mdlwin.exe file either directly from www.practicalstats.com or from the Division’s common drive.

Open mdlwin.exe. You will immediately be prompted to select file on which to run the statistical program. The window will read: “Open: Select File for Unit 10”. Select the file that contains the data you wish to analyze and click OK.

Next the program will prompt you to select a file to input the results. The screen will read “Open: Select File for Unit 31.” Create a new .txt file into which you want the program to deposit the results. Click OK. The program will produce a screen titled: “Summary Statistics for Data with Multiple Detection Limits”

Step 4: Calculate the coefficient of variation.

From the screen that reads: “Summary Statistics for Data with Multiple Detection Limits” choose the standard deviation (“STD DEV”) and the mean from the line labeled “Estimates Using Robust Log-Probability Regression”. Print the results and include it in the permit file.

Now, calculate the coefficient of variation according to the following equation:

$$\text{Coefficient of Variation} = \frac{\text{standard deviation}}{\text{mean}}$$

mean

The result is the coefficient of variation. Use this number directly to select the appropriate multiplier from the Table in Appendix 1. Do not take the antilog of this number or otherwise change it; it has already been converted. Proceed with the rest of the RP analysis as described in the RP guidance.

Appendix 4: Rationale

I. BACKGROUND

This guidance delineates how the Division permitting staff will analyze effluent data to determine if a discharged pollutant has the reasonable potential to cause or contribute to an exceedance of water quality standards.

Historically, the Division has used a variety of methods to determine “reasonable potential.” While the methods may indeed have been protective, they were not consistently applied. Some of the methods formerly used by the Division have included:

- Recent effluent data (usually the most recent 24 months) was reviewed; if the maximum effluent concentration was less than some percentage (often 50%) of the calculated limit, then no limitation was included in the permit. If the maximum effluent concentration was greater than the specified percentage of the calculated limit then a limitation for that pollutant was included in the permit (most common method).
- The mean effluent concentration was determined, then two standard deviations were added. If the result was less than the calculated limit then no limitation was included in the permit.

Permit writers were not required to use any particular method.

The Division developed this guidance in order to create a more systematic and defensible method to determine reasonable Potential. This document was researched and written by the Division with the input of the Colorado Water Quality Forum, Permits Workgroup. The Workgroup consisted of approximately 20 members representing the Division and a variety of municipalities with CDPS permits. Beginning work in July 2002, the members of the Workgroup were invited to participate in live discussions as well as make written comments on the many drafts of the guidance.

II. BASIS OF THIS GUIDANCE

This guidance is based, with some modifications, on the USEPA’s guidance for establishing reasonable potential as found in Chapter 3 and Appendix E-1 of the *Technical Support Document for Water Quality-based Toxics Control* (USEPA 1991), hereafter referred to as the “TSD” method. The TSD Method was chosen after applying it to real Colorado data and after considering the methods used by several other states to make RP determinations.

The quantitative method for determining reasonable potential uses a statistical analysis of facility-specific effluent data to estimate the pollutant concentration that is expected to exceed the 99th percentile of the data set, at the 99% confidence interval. For a more detailed overview of the method please see section II.A of the body of this guidance document.

III. IMPLEMENTATION DECISIONS/MODIFICATIONS

While this guidance follows the basic method described in the TSD, there are some instances where the TSD requires the user to make implementation decisions (for example, the user must select the percentile to be used). Also, there are some areas where the TSD has been modified for use in this guidance (for example, the Reasonable Potential Multiplying Factors Table has been expanded from 20 samples in the TSD to 100 samples in this guidance). The areas where

implementation decisions or modifications have been made, along with a discussion of each area, are listed below.

A. EXPANSION OF MULTIPLIER TABLES:

The Reasonable Potential Multiplying Factors Table contained in Appendix 1 of this document is central to the TSD method. As outlined by EPA, the TSD gives multipliers for data sets with up to 20 data points. In this Division guidance, the same equations that were used by EPA to fill in the Reasonable Potential Multiplying Factors Table were used to expand the table to 100 samples. This expansion will give credit to dischargers with large data sets and encourage other dischargers to create larger data sets. (Larger data sets will result in lower multipliers.) Larger data sets are more desirable because they are generally less variable and more representative of the discharge than smaller data sets. Use of larger data sets lessens the chance of setting permit limits for pollutants that are unlikely to exceed water quality-based effluent limits and increases the chances of detecting and regulating pollutants that may exceed water quality based effluent limits.

B. USE OF 99TH ILE FOR ALL DATA

This guidance calls for the use of the 99th %ile for all types of data. The primary reason for choosing the 99th %ile, is that at lower %iles, the reasonable potential multiplying factors can result in a multiplier that is less than 1.0. For example, at the 95th %ile, it is possible for the maximum estimated pollutant concentration derived by the statistical analysis to be *less than* a real value in the data set. This could result in a determination of “no RP” for a discharger with real pollutant concentrations that exceed the maximum allowable pollutant concentration. Therefore, the Workgroup decided to use the 99th %ile in order to estimate a concentration that is unlikely to be exceeded by real data in any given data set.

C. CREATION OF A “MONITOR ONLY” TIER

This guidance allows the permit writer to require monitoring in situations where the estimated maximum pollutant concentration is close to (> 50% of) the maximum allowable pollutant concentration. This helps to verify the accuracy of the estimates made by the statistics in cases where that accuracy is critical. It also gives the permit writer more complete information with which to make an RP determination at the time of permit renewal.

D. CREATION OF SPECIAL PROCEDURES FOR KNOWN OR SUSPECTED FALSE POSITIVES

The Workgroup was concerned that individual data points that do not follow the statistical trend of the rest of the data (“outliers”) could be false positives. Therefore, this guidance allows the permit writer to consider the possibility of false positives in instances where an outlier affects the outcome of the RP determination. If the permittee can demonstrate, using appropriate statistical methodology, that the value in question is a statistical outlier, then the permit writer may remove the outlier from the data set and proceed with the RP analysis.

If the new analysis indicates “no RP” then monitoring may be required in lieu of a permit limit.

The Division recognizes that just because a value is a statistical outlier, it does not mean that the value is a false positive. An anomalous value may be attributable to spills, malfunction of pollution control equipment or other practical factors. Therefore, the Division will require continued monitoring to verify the claim that the anomalous value is indeed a false positive.

It is important to note that the only consequence of finding that a data point is a statistical outlier is that it may be excluded from the RP statistical evaluation. Since the purpose of the RP statistical evaluation is to estimate whether a pollutant can be expected to exceed the maximum allowable pollutant concentration, values that exceed the maximum allowable pollutant concentration are not subject to the RP statistical evaluation.

Therefore, values that exceed the maximum allowable pollutant concentration will result in a permit limit, regardless of whether they conform to the pattern of the rest of the data.

E. REQUIRING A MINIMUM OF 10 DATA POINTS.

Calculating a coefficient of variation for data sets with less than 10 data points is not reliable. For this reason, the original TSD method recommends using a default CV when 10 or less data points are available. However, this guidance eliminates the need for a default CV by requiring a minimum of 10 data points when making an RP determination. The Division believes that this is a modest requirement, especially when considering the fundamental nature of an RP determination.

F. REQUIRING THE CALCULATION OF COEFFICIENT OF VARIATION

The original TSD allows for the use of a default coefficient of variation (0.6). The workgroup compared calculated CVs with the default CV, and the resultant multipliers, for sample Colorado Data. The results of this comparison is summarized below in Table 1:

Table 1: Comparison of Calculated CV vs. default CV

Facility	Calculated Coefficient of Variation ¹	Default Coefficient of Variation	Multiplier using calculated coefficient of variation ²	Multiplier using default coefficient of variation ²
Metro	1.2	0.6	4.2	2.3
Colorado Springs	1.1	0.6	3.8	2.3
Plum Creek	1.6	0.6	5.5	2.3
Brush	0.1	0.6	1.4	7.4

¹ Two years (June 00 to June 02) of copper data (daily maximums), taken from PCS were used to calculate the CV.

² Multiplier taken from Table 3-1 of the TSD (99% confidence level).

Table 1 shows that the calculated coefficient of variation can differ significantly from the default. For Metro, Colorado Springs, and Plum Creek, the calculated coefficient of variation is much higher than the default, resulting in a higher multiplier and consequently a much higher estimated maximum pollutant concentration. In contrast, the calculated coefficient of variation for Brush was much smaller than the default, but because there were so few data points (two- compared to 24 for the other municipalities), the multiplier was much higher for the default coefficient of variation.

Therefore, since the calculated CV can be so different from the default, this guidance requires the calculation of the CV for each individual data set.

G. ALLOWING DIFFERENT CALCULATIONS FOR DATA THAT FOLLOWS A “NORMAL” DISTRIBUTION

The TSD method assumes that effluent data follow a lognormal distribution. Although this is a common assumption for data of this type, there may be times when the data follows a normal distribution. Therefore, the guidance allows for an alternative calculation of the CV for data that follow a normal distribution.

H. SPECIAL HANDLING OF BELOW DETECTION LIMIT VALUES

The TSD method does not work for data that is below the detection limit or data that is reported with different detection limits. Therefore, the guidance adopted the MDL (Helsel 1991) method of calculating the coefficient of variation when some of the data includes less than the detection limit values.

IV. OTHER METHODS CONSIDERED BUT NOT SELECTED:

In August 2002, the reasonable potential procedures used by a sampling of other states and EPA regions were reviewed. The methods used by the states fell into 3 main categories:

- 1) Best Professional Judgement (No statistical evaluation is conducted.)
- 2) TSD Method
- 3) EPA Region VI procedure.

A. BEST PROFESSIONAL JUDGEMENT

Several states (Connecticut, Kentucky, Alabama) used Best Professional Judgement (“BPJ”) instead of a systematic quantitative method to determine reasonable potential. These agencies often cited the following types of information in their reasonable potential determinations: effluent monitoring, WET results, flow, potential for plant upsets, performance of similar facilities, and water quality/low flow in the stream. The permit writers then use this type of information to make a determination based on best professional judgement. Best professional judgment is sometimes aided by a guideline that pollutant concentration should be below a certain percentage of the calculated limit. One state (Kentucky) indicated that they look at this percentage of the water quality standard but do not define an acceptable percentage (they mention the use of 90% of the calculated limit as the reasonable potential threshold). On the opposite end of the spectrum, Alabama considers pollutant concentrations that are greater than 20% of the calculated limit to indicate a reasonable potential to cause or contribute to water quality exceedances.

Since the Division was looking for a quantitative method of analyzing data for RP that would be applicable in most situations, this guidance does not adopt any of these BPJ-based methods.

B. THE TSD METHOD

Several states (Washington, Virginia, South Carolina) reported using the EPA’s methodology described in the TSD. As previously discussed, if one uses the TSD approach, there are two main decisions that need to be made: 1) what coefficient of

variation to be used and 2) what percentile should be used (generally either 99% or 95%). South Carolina uses the default coefficient of variation of 0.6 suggested in the TSD. Washington uses the default coefficient of variation for data sets less than 20 and calculates the 95th %ile for data sets of greater than 20.

As discussed previously in this appendix, the Division chose to adopt a modified TSD method.

C. THE REGION 6 METHOD

Several states (Washington, Louisiana, and Kansas) reported using an alternative statistical method, developed by EPA Region 6, which is based on the relationship of the geometric mean to a specified percentile (usually the 95th % ile). (We will refer to this method as the “Region VI Method”.) The Region VI statistical method assumes a lognormal distribution and a constant coefficient of variation (generally assumed to be 0.6). It is independent of sample size and allows the use of very small data sets or even a single data point to estimate the upper range of the concentration that could be discharged. Given the above assumptions, the net result of the method (for the 95% ile) is that:

$$\text{pollutant concentration} * 2.13 = 95^{\text{th}} \% \text{ ile pollutant concentration.}$$

In practice, if a permittee reported a cadmium concentration of 4.0 ug/l, the permit writer would multiply 4.0 ug/l * 2.13 to get a concentration of 8.5 ug/l. If the calculated permit limitation were less than 8.5 ug/l, then the permit writer would conclude that there was reasonable potential for the effluent to cause or contribute to an exceedance of water quality standards and a limitation for cadmium would be included in the permit.

The Region VI Method is based on the following equation:

$$C_p = C_{\text{mean}} * \exp(Z_p * s - 0.5*s^2)$$

Where:

C_p	= Concentration at the specified percentile
C_{mean}	= geometric mean of the effluent concentration
Z_p	= normal distribution factor at p th percentile (table value)
s^2	= $\ln (CV^2 + 1)$
CV	= coefficient of variation = 0.6

Solving the equation, the result is:

$$C_{95} / C_{\text{mean}} = 2.13$$

Using this method, the multipliers for various percentiles are summarized below:

Percentile	Z	C_p/C_{mean}
90	1.283	1.74
95	1.645	2.13
99	2.386	3.11

(USEPA Region 6)

Both the TSD Method and the Region 6 Method were applied to a sampling of real Colorado data. Generally, the Region 6 method produced lower estimates of the maximum expected pollutant concentration than the TSD method. The results are compared in Table 2 below.

Table 2:
Comparison of Calculation of Maximum Effluent Concentration (M₂) Using TSD and Region 6 Methods
(All concentrations are in ug/l)

a. Colorado Springs

Pollutant	highest observed conc.	M₂ TSD, 95%	M₂ Region 6 95%	M₂ TSD, 99%,	M₂ Region 6 99%	maximum allowable effluent conc.
Zn, pd	92	128.8	142.2201	211.6	207.6547	222
Ag,pd	0.4	0.56	0.49842	0.92	0.72774	1.3
Cu, pd	13	18.2	13.47012	29.9	19.66764	38.1

b. City of Brush

Pollutant	highest observed conc.	M₂ TSD, 95%	M₂ Region 6 99%	M₂ TSD, 99%	M₂ Region 6 95%	maximum allowable effluent conc.
Zn, pd	57	216.6	140.883	421.8	96.489	10235
Ag,pd	0.6	2.28	1.07606	4.44	0.73698	96
Cu, pd	17	64.6	51.2839	125.8	35.1237	794
Phenols	8	30.4	21.5523	59.2	14.7609	na

Table 2:
Comparison of Calculation of Maximum Effluent Concentration (M₂) Using TSD and Region 6
Methods (continued)
(All concentrations are in ug/l)

c. Metro Wastewater Reclamation District

Pollutant	highest observed conc.	M₂ TSD, 95%	M₂ Region 6, 95%	M₂ TSD, 99%,	M₂ Region 6 99%	maximum allowable effluent conc.
Mn, dis	86	120.4	112.6983	197.8	164.5501	400
Se, total	7.4	10.36	6.1131	17.02	8.9257	8
Cr+6, dis, as Cr	0	0	0	0	0	11
Zn, pd	61	85.4	93.5922	140.3	136.6534	219
Ag,pd	.11	0.154	0.08946	0.253	0.13062	1.4
Cu, pd	35.1	49.14	22.3437	80.73	32.6239	24.6
Cd, pd	0	0	0	0	0	2.23
Pb, pd	0	0	0	0	0	13.1
Hg, pd	0	0	0	0	0	0.4
Ni, pd	73	102.2	45.5181	167.9	66.4607	184
Diazinon	.27	0.378	0.1491	0.621	0.2177	

d. Plum Creek Wastewater Authority

Pollutant	highest observed conc.	M₂ TSD, 95%	M₂ TSD, 99%,	M₂ Region 6 95%	M₂ Region 6 99%	maximum allowable effluent conc.
Zn, pd	35	91	164.5	73.5	107	139
Cu, pd	9.6	25.0	45.3	19.8	29.0	14.1

Assumptions:

1. Data is lognormally distributed
2. Coefficient of variation is default value (0.6)

3. Sample size = 24 (except for Brush; sample size = 2)
4. Data is from PCS data; period of review is June 2000 to May 2002
5. Pollutant concentrations with no values above the detection limit are assumed to be 0.
6. Values BDL for data sets with values both above and below detection limit are set at detection limit.

The Division considered the use of the Region 6 Method but decided against it primarily because the method does not allow for the calculation of individual CV and thus does not account for the variability of individual effluents.

D. Mean + 2 Standard Deviations

An alternative method suggested by statistician Tim Moore of Risk Sciences, through Nancy Keller of the City of Pueblo, was to take the mean of the data and add 2 standard deviations. Since the data is assumed to be lognormal, the natural log of each data point is taken then the average and standard deviation (of the log transformed data) is calculated. Then 2 times the standard deviation is added to the average and the antilog of that number is taken. This result is then compared against the maximum allowable pollutant concentration.

This method was tried on sample data for Metro for copper (daily maximum) for the period of June 2000 through May 2002. The result for this example was a value that was *less than* the highest value in the data set. Therefore, with this method it is possible to arrive at a “no RP” determination—even when the data itself contains values that clearly indicate reasonable potential because they are above the maximum allowable pollutant concentration. Therefore, this method was not selected.

V. REVIEW OF OTHER DIVISION POLICIES

There are several policies in use by the Division that involve the statistical evaluation of water quality data to make a variety of decisions. Some of these are:

-Policy for Characterizing Ambient Water Quality for Use in Determining Water Quality Standards based Effluent Limits (2002). This policy outlines the statistical evaluation used to characterize upstream water quality when calculating permit limitations. It uses a percentile approach, in which the ambient condition is defined as some percentile (e.g., 85% for metals) within the data set.

-Year 2002 303(d) Listing Methodology, Water Quality Control Division, March 11, 2002. This document gives guidance about how to evaluate whether data is sufficient to cause a stream to be considered “impaired” and thus included on the States list of impaired waters. If a specified percentile of the stream data exceeds the standard, then the stream is considered impaired. The percentiles used in this determination are the same as those specified in the above policy. The policy adds that at least 10 stream samples are needed to make the determination unless there is other overwhelming evidence of impairment.

-Antidegradation Significance Determination for New or Increased Water Quality Impacts (December 2001). This documents how the Division will determine whether a new or increased discharge will have “significant” water quality impacts, according to the Section 31.8 of the *Basic Standards and Methodologies for Surface Water (5 CCR 1002-31)*. In this document the benchmark against

which to compare effluent data to determine if it will have a “significant impact” on water quality is the 85%ile of the low flow pollutant concentration as of 9/30/00.

It is important to be aware of these documents. However, since the characteristics of the data differ and since different questions are being answered with the data, their methods do necessarily not need to match the methods used for RP determinations.