ArcelorMittal

November 18, 2010

Mark A. Wayner, Regional Manager Air Quality
Barbara R. Hatch, P.E., Environmental Engineer Manager, Air Quality Program
Devin P. Tomko, Air Quality Engineering Specialist
Pennsylvania Department of Environmental Protection
Southwest Regional Office
400 Waterfront Drive
Pittsburgh, PA 15222

Re: Statistical Analysis of Monessen Facility Test Data to Correct RACT limits

Messrs Wayner, Hatch and Tomko,

In response to our meeting of October 22, 2010, enclosed is a report by Integral Concepts that examined test data gathered over eleven years (1998-2008) from the ArcelorMittal Monessen Coke Facility to determine the upper bound emission rate that can be expected from the units subject to NOx and VOC RACT emission limits. Integral Concepts utilized methods from EPA Guidance for Data Quality Assessment (EPA QA/G9) to conduct a statistical analysis of Monessen's RACT units. ArcelorMittal is using the attached report to update the pending February 2003 application to correct RACT emission limits submitted by the previous owner, Koppers Industries. With the benefit of additional data and applying EPA's preferred statistical methods, the report provides improved justifications for corrected RACT limits for NOx and VOC emitted from the combustion stacks at Monessen Coke Batteries 1B and 2. The Table below indicates the corrected RACT limits:

Emissions Unit	1998 NOx RACT	Corrected NOx RACT	1998 VOC RACT	Corrected VOC RACT
Coke Battery 1B Combustion Stack	60.7 lb/hr 286.0 tpy	84.1 lb/hr 368.4 tpy	0.3 lb/hr 1.0 tpy	23.5 lb/hr 102.9 tpy
Coke Battery 2 Combustion Stack	55.4 lb/hr 246 tpy	61.4 lb/hr 268.9 tpy	0.5 lb/hr 1.9 tpy	6.8 lb/hr 29.8 tpy

These changes are necessary to correct errors made in the calculation of NOx and VOC emission rates when RACT limits were originally assigned to these coke battery combustion stacks in March of 1998. Several of the original RACT limits were based solely on a single emissions stack test performed in 1997. Testing in April 1998 already determined that the combustion stack RACT limits were in error and DEP was notified of the problem. Unfortunately, the permit was issued without RACT relief and efforts commenced immediately to gather data to justify revised RACT limits. Stack test data from 1998-2002 was collected and

ArcelorMittal
4020 Kinross Lakes Parkway
Richfield, Ohio 44286-9000
United States

T +330 659 9163
F +330 659 7434
www.arcelormittal.com

statistically analyzed in a report provided to DEP in February 2003 with a request to revise the RACT limits for the Monessen Facility. Correspondence with DEP indicates that the revised RACT limits would be incorporated into the Title V Permit when issued. When the Title V permit was being prepared by DEP, ArcelorMittal provided an updated report that added stack test data through 2008 to the prior 2003 statistical analysis. At our June 30, 2010 meeting, DEP questioned the statistical justification for the proposed RACT emission limits because the proposed limits were in excess of the highest stack test emission rate. ArcelorMittal agreed to provide an updated statistical analysis with a clear discussion of the statistical method used by EPA to set enforceable emission limits that account for normal operational variability. That analysis is presented in the report attached to this letter. [1]

The attached report analyzes the combustion stack data from 1998-2008 using the statistical methods derived from EPA guidance. EPA is relying increasingly on statistical analysis to set enforceable emission limits in federal rulemakings at levels that reflect normal variations in operating conditions. The agency recognizes that stack tests capture snap shots of performance that do not reflect the true range of operating conditions. Yet, EPA is charged with setting enforceable limits that must be met under all operating conditions. To address this challenge, EPA has turned to statistics to predict the upper emission rate that top performers can be expected to consistently achieve. That predicted emission rate is then imposed as the enforceable emission limit even when it is above the highest actual emission rate in the dataset. This statistical approach has been used recently to propose Maximum Achievable Control Technology (MACT) emission limits for categories of boilers, process heaters, and incinerators. See e.g., 75 Fed. Reg. 32019-32029 (Boiler MACT Proposed Rule).

EPA's statistical approach to setting MACT limits is a conservative methodology suitable for controlling hazardous air pollutants. Thus, EPA is expected to readily accept a similar statistical approach for justifying RACT limits. MACT limits are based on the Upper Prediction Limit (UPL) of a given data set composed of stack test runs from top performing sources within a category or subcategory. EPA determines whether the data distribution fits into a recognized pattern (normal, lognormal, etc.). The distribution pattern is then used to predict the 99% upper bound of the expected range of emission rates achieved by the best performers during all normal operating conditions. The enforceable MACT limit is set at the 99% UPL, which is typically higher than the maximum test run emission rate.

Integral Concepts uses EPA's statistical approach to predict the upper emission rate for the two Monessen combustion stacks from a dataset comprised of eleven years of stack test data. The report provides a transparent discussion of the process of choosing a distribution pattern for each dataset. When multiple options were considered for matching distribution patterns to a dataset, Integral Concepts ran the statistical analysis for all options. The corrected RACT limits are based on the higher predicted emission rate to ensure that the revised emission limit will reflect the full range of normal operating conditions that could affect emissions. The additional data and the updated statistical analysis using EPA guidance methods should provide sufficient justification for DEP to act on correcting the NOx and VOC RACT limits for the coke

ArcelorMittal
4020 Kinross Lakes Parkway
Richfield, Ohio 44286-9000
United States

T +330 659 9163
F +330 659 7434
www.arcelormittal.com

battery combustion stacks as indicated above. The report also analyzes RACT emission limits for the pushing emission control system (PECS). The predicted NOx and VOC emission rates from the PECS are within the existing RACT limits. Therefore, ArcelorMittal is not proposing a change to the PECS RACT limits.

ArcelorMittal understands that the corrected RACT emission limits will be established by issuing a Plan Approval to replace the erroneous emission limits set forth in a 1998 Plan Approval that has expired. Once formally issued, DEP will submit the corrected RACT emission limits to EPA for adoption into the State Implementation Plan (SIP) to replace the erroneous RACT limits currently in the SIP. We encourage DEP to reference this process in the Title V Permit so that the revised RACT limits will become applicable and the old RACT limits will sunset automatically upon SIP approval.

We look forward to answering any questions that you have regarding the attached analysis so that we can complete resolution of this important matter as soon as practicable. Should you with to discuss this matter please contact me at 330-659-9163.

Sincerely,

Michael E. Long
Manager Environmental Compliance
ArcelorMittal USA

Attachment

Cc: Paul Champagne, with attachment
Greg Shamitko, with attachment
Keith Nagel, with attachment
Douglas McWilliams, with attachment

---

[1] The Monessen RACT provisions were originally established in Operating Permit No. 65-000-853 and incorporated into the Pennsylvania SIP on August 21, 2001. The prior owner of the Facility, Koppers Inc., requested changes to the RACT limits in Plan Approval Application 65-000-853A based on errors identified in

---

the original application that rendered the coke oven battery combustion stack RACT limits unachievable. That application was submitted in February 2003, deemed administratively complete by PA DEP on March 19, 2003, and has been pending for seven years. After obtaining ownership, ArcelorMittal reminded DEP that this application was still pending and in February 2010 provided DEP with a markup of the permit to show precisely which changes were necessary to the original RACT operating permit, which was incorporated by reference into the Pennsylvania SIP at 40 C.F.R. 52.2063(c)(172)(i)(B)(9).

ArcelorMittal requests that DEP approve the RACT limits in this document and submit them to USEPA as a proposed SIP revision. If the SIP revision is not federally-approved prior to the final issuance of the Title V Permit, the Title V Permit will need to list the erroneous SIP limits as applicable requirements. However, DEP must also include a compliance schedule for any such applicable requirements that ArcelorMittal Monessen cannot achieve. The compliance schedule should include, as its initial milestone, compliance with the corrected RACT limits. Subsequent milestones would involve periodic updates on the federal SIP approval process until the revised limits are approved into the SIP. We also suggest that DEP include language in the Title V Permit that automatically replaces the old SIP RACT limits with the corrected RACT limits as soon as they are approved by EPA into the SIP. This would allow the elimination of the erroneous RACT limits immediately upon USEPA approval of the SIP revision and it would avoid reopening the Title V Permit unnecessarily.

August 5, 2010

Kimberly D. Coy
Permitting & Compliance Manager
Air/Compliance Consultants, Inc.
1050 William Pitt Way
Pittsburgh, PA 15238

Dear Kim:

Thank you for contacting Integral Concepts regarding the development of statistical limits for emissions data. The analysis considered the expected limits of variation based on the data supplied. This report summarizes the analysis of each data set along with some general observations regarding the data.

METHODOLOGY AND IMPORTANT NOTES

For each data set, the data was explored. Parametric Distribution Fitting was performed to attempt to find plausible distributions that described the data. If the Normal distribution was reasonable then that model was used. If the Normal distribution did not characterize the data (failed Normality Hypothesis Test at 95% confidence), then an alternate distribution was utilized (if possible) to estimate the population statistics and upper confidence bounds on individual values.

Note Regarding Distribution Fitting

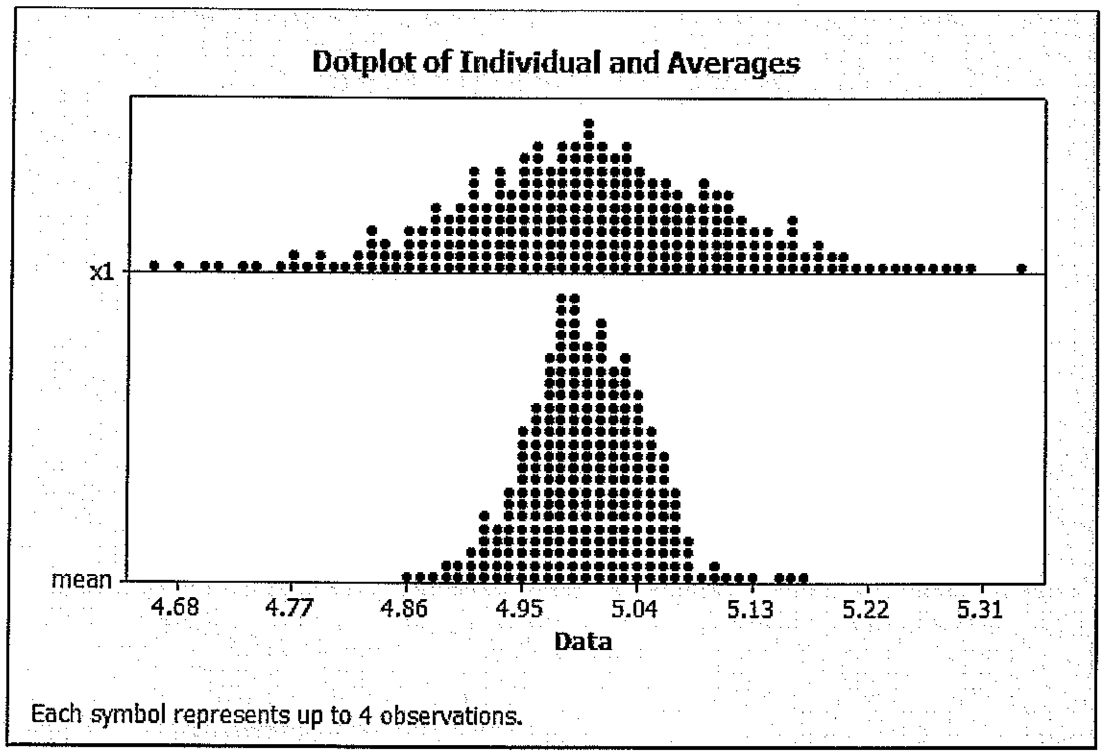
When attempting to fit a distribution (e.g. Normal or other) to data, it is assumed that the data come from a *single* process or population. When data comes from *different* processes it can be impossible (or misleading) to represent the data with a single continuous distribution. The purpose of fitting a distribution (model) is to use it to predict future process behavior and this only applies when we are dealing with a *single* process. Evidence of multiple processes may include very extreme (outlying) values that appear to be very different from the rest of the data, or indications of instability on statistical process control charts (if enough data is available to construct a valid control chart).

A determination of whether data points are unrepresentative of the "process" (which provides evidence that the system has changed, i.e. NOx levels are statistically different than at other time points) must be made only with consultation with process experts. Statistical analysis can highlight suspicious data, but a final decision as to include or exclude data must be made with input from the subject matter expert.

For some of the datasets, the analysis was performed *with and without* specific data points since the inclusion of such points had a significant effect on the estimated values of interest.

Note Regarding Use of Mean, Median or Maximum Values Rather than All Data

In general it is ***not** appropriate* to summarize the data (by taking averages, medians, maximums, etc.) when trying to make inferences about individual observations. Utilizing summary statistics (such as the mean or median) will *remove variability* from the data and present misleading results. To illustrate this, the top dot plot below shows 1000 data points from a normal distribution with mean = 5.0 and standard deviation = 0.1. The second dot plot shows 1000 averages (of size 5) from the same distribution. As you can see averaging "averages out" the more extreme values.



However, the emission data in this analysis exhibits the characteristic (in general) that the variation *within* a year is quite small as compared to the variation *between* years. In this situation, having multiple values (per year) that are essentially the same adds little to our understanding of the process variability. In fact, in some cases (as described below), having multiple measurements that were practically the same value (where other years were significantly different) made the data appear to be discrete rather than continuous which inhibited the ability to fit parametric models (distributions). In these situations, the median (and in one case, the maximum) values were used to find parametric models that adequately describe the variation over the 11 year period.

It may be that he variation seen within a year only represents measurement error, since the measurements were taken at approximately the same time. The variation from year to year likely represents changes in the environment. More will be mentioned on measurement systems shortly.

If the data exhibited a more stable situation, where the variability within the year resembled the variation from year-to-year, using medians or means or maximums would be completely invalid and would produce misleading results.

Note Regarding Measurement Systems

As seen in the analysis below, several measured values are statistically "far away" from other values in the data set. While some processes do naturally exhibit extreme value behavior, more common causes of this pattern include unstable processes or imprecise measurement systems. It is critical that any data analysis that is performed with the knowledge that measurement system adequacy has been verified. In addition to measuring accurately, measurement systems should also be precise (repeatable). That is, when the *same* specimen (or condition) is re-measured, the variability between these repeated measurements should be very small (as compared to the acceptable emission limits as well as compared to the normal process variation that includes various conditions or specimens). In short, it is hoped that the observed variation reflects true process variation rather than mostly measurement error. Several statistical methods exist to assess measurement system adequacy including Gage Repeatability & Reproducibility Assessments.

Note Regarding Reported Upper 99% Bound

The reported upper bound represents the estimated values at which we would expect 99% of future *individual* emissions readings to fall below (based on the distribution utilized). An alternative description of the bound is that there is a probability of 0.99 that the next reading would fall below the 99% upper bound.

Note Regarding Compliance with "EPA Guidance for Data Quality Assessment-EPA QA/G-9"

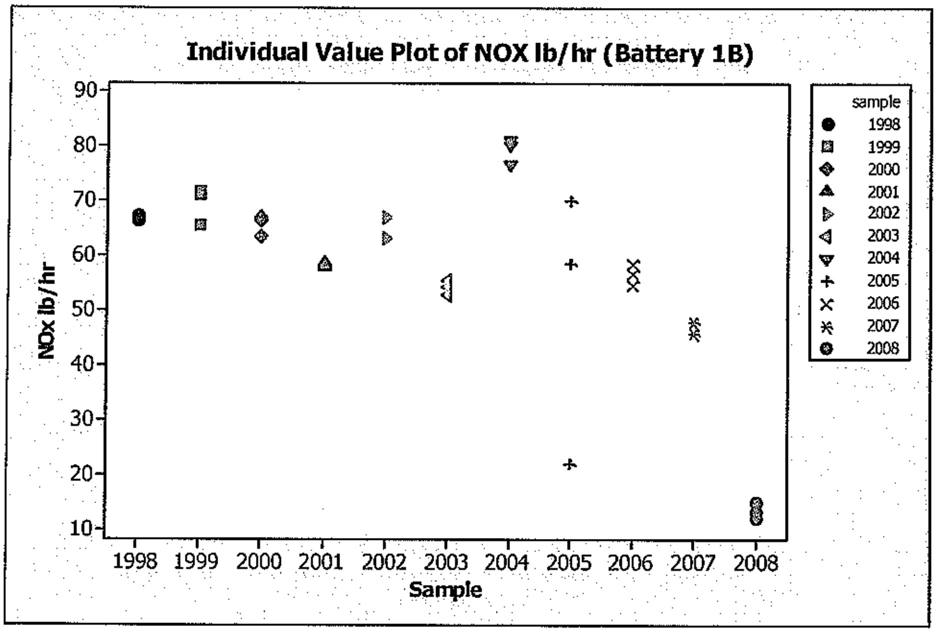
The statistical and graphical methods used for this analysis are in accordance with the "EPA Guidance for Data Quality Assessment-EPA QA/G-9." Specifically, the main sections of the EPA Guidance document that applied to this analysis are as follows:

- Normal Probability Plot (2.3.6)
- Goodness of Fit Tests (4.2.7) (Anderson Darling Test for Normality)
- Histogram/Dot Plots (2.3.1)
- Box and Whisker Plots (2.3.3) (Outlier tests)
- Plots for Individual Data Points (2.3.7.1) (Individual Value Plots)
- Probability Distributions (2.4) (Normal, Lognormal, others)
- Measures of Central Tendency (2.2.2) and Measures of Dispersion (2.2.3)
- Calculation of Basic Statistical Quantities (2.1.2)
- Measures of Relative Standing (2.2.1)
- Comparing Two Means (3.3.1) and Comparing Two Variances (4.5.2)
- Outlier Tests (4.4)
- Perform Statistical Hypothesis Tests (5.1.1)
- Interpreting and Communicating Test Results (5.2)

ANALYSIS

Combustion Stack Battery 1B NOx

An Individual Value Plot of the data is shown below along with some observations and the results.



- With the exception of the year 2005 data, the variability within the 3 measurements is very small. The large variation in the 2005 measurements could be due to measurement error, or possibly a longer time period over which the data was collected (as compared to the other years). In any event, the large amount of variability in the replicate 2005 measurements appears highly unusual given the consistency observed in the other years.

- The 2008 data appears to represent a different process than the other 10 years. These relatively low readings could reflect a measurement system issue or an actual process difference.

- With *all* of the data included, a single parametric distribution could not be found that adequately described the data. Distributions considered included: Normal, Weibull, Lognormal, Logistic, Loglogistic, Gamma, Exponential, Largest Extreme Value, and Smallest Extreme Value.

- For this dataset, an argument can be made that the replicate measurements are not useful since (for most cases) the variation within the 3 measurements (within a year) is very small. Distribution fitting can be difficult when the data appears to be discrete (with multiple values being the same or very close with gaps in between the clusters of values). Thus, the distribution fitting was repeated using only the medians of the 3 replicate values. The smallest extreme value (sev) distribution reasonably fit the medians of the yearly data values. The descriptive statistics and upper confidence bounds (using the smallest extreme value distribution) are:

Descriptive Statistics: Median NOx lb/hr

Variable	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Median NOx lb/hr	58.29	17.40	13.31	54.38	58.57	67.04	80.41

99% Upper Bound: 82.99

- The analysis was also performed after removing the 4 low NOx values (2008 data and 1 observation from 2005). With the removal of that data, there was no evidence to reject the normality assumption. Using the normal distribution, the descriptive statistics and upper confidence bounds are:

Descriptive Statistics: NOx lb/hr

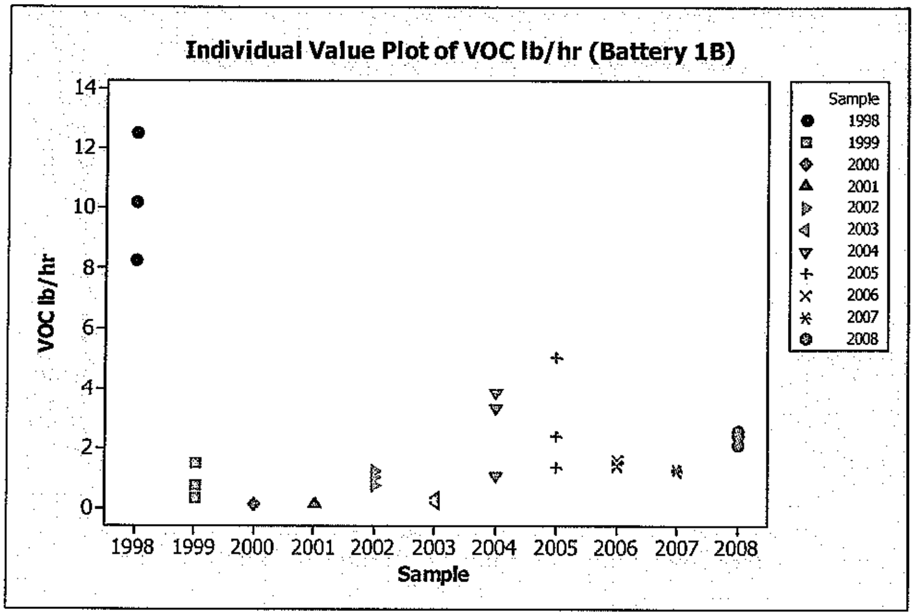
Variable	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
NOx lb/hr	62.81	9.13	45.85	56.13	63.53	67.25	81.39

99% Upper Bound: 84.05

Note that the upper confidence bounds for the two approaches are pretty close.

Combustion Stack Battery 1B VOC

An Individual Value Plot of the data is shown below along with some observations and the results.



- The 1998 data is strongly statistically different than the most recent 10 years of data. Since we are interested in the upper bounds of the emissions, the decision to include or exclude this data has a dramatic impact on the results. Per your request, the data was analyzed both ways.

- With *all* of the data included, a single parametric distribution could not be found that adequately described the data.

- For this dataset, an argument can be made that the replicate measurements are not useful since (for most cases) the variation within the 3 measurements is very small. Distribution fitting can be difficult when the data appears to be discrete (with multiple values being the same or very close with gaps in between the clusters of values). Thus, the distribution fitting was repeated using only the medians of the 3 replicate values. Both, the Weibull and Lognormal distributions produce excellent fits to the medians of the yearly data values. The descriptive statistics and upper confidence bounds (using the Weibull and Lognormal distributions) are:

Descriptive Statistics: Median VOC lb/hr

Variable	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Median VOC lb/hr	2.144	2.884	0.133	0.354	1.309	2.437	10.259

WEIBULL	LOGNORMAL
99% Upper Bound: 11.59	99% Upper Bound: 23.50

Note that the lognormal upper bound estimates are considerably higher due to the longer tail of this distribution. Since both the Lognormal and Weibull distributions fit the data well, the Lognormal estimates are the most conservative.

- The analysis was also performed after removing the 1998 data. As before, an adequate parametric distribution could not be found using all of the individual data values due to the clustering of similar values near 0. Using medians, there was no evidence to reject the normality assumption. The descriptive statistics and upper confidence bounds (using the Normal distribution) excluding 1998 data are:

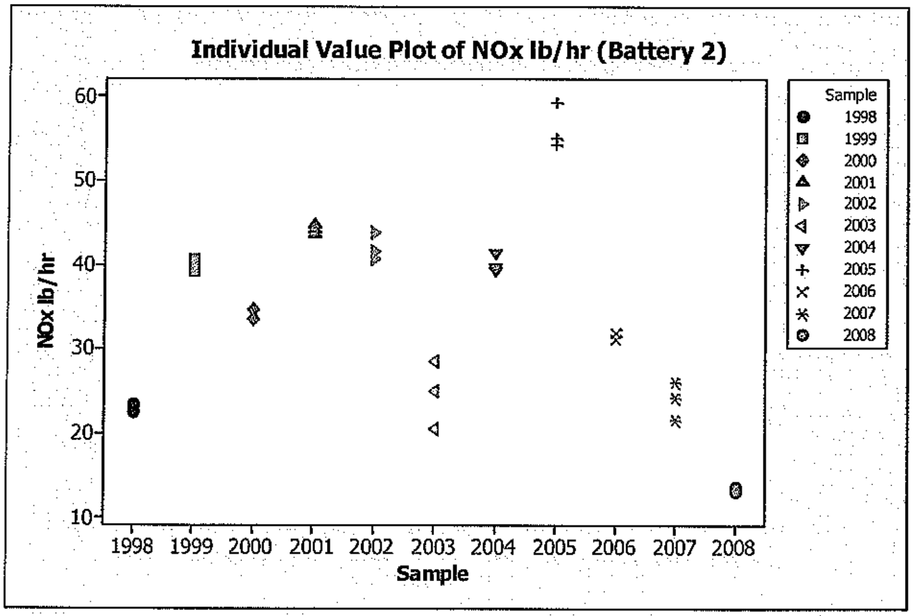
Descriptive Statistics: Median VOC lb/hr (ex 98)

Variable	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Median VOC lb/hr (ex 98)	1.332	1.093	0.133	0.300	1.158	2.410	3.356

99% Upper Bound: 3.87

Combustion Stack Battery 2 NOx

An Individual Value Plot of the data is shown below along with some observations and the results.



- The data was test for normality and there is no evidence to reject the normality assumption.

- The descriptive statistics and upper confidence bounds (using the normal distribution) are:

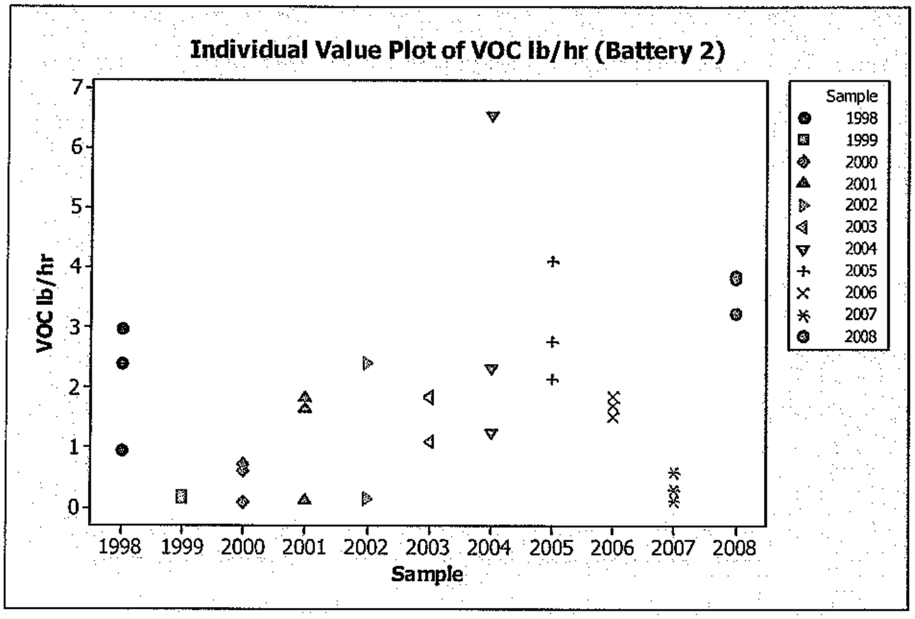
Descriptive Statistics: NOx lb/hr

Variable	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
NOx lb/hr	33.94	11.82	13.19	23.82	34.68	41.46	59.34

99% Upper Bound: 61.44

Combustion Stack Battery 2 VOC

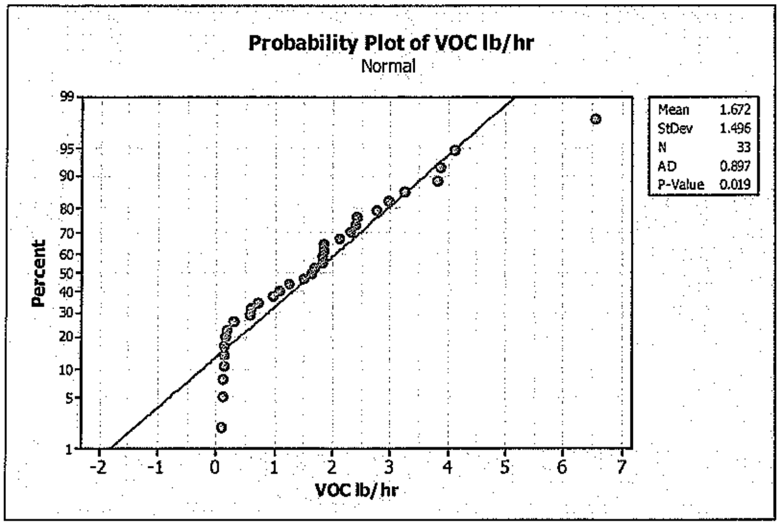
An Individual Value Plot of the data is shown below along with some observations and the results.



- The degree of variability within the replicates varies greatly. In some years the variation is tiny and other years the variation is quite large. One might question whether the measurement system is repeatable, or whether the measurements were taken within similar timeframes from year to year.

- The year 2004 data exhibits a lot of variation including one large value that is statistically different from the rest of the data. The results differ significantly depending whether or not this value is included in the analysis (see below).

- The data was tested for normality and the normality test failed. The main reason is due to the three replicate values from 1999 (which all measured practically the same value which was close to zero). The normality plot is shown below.



- Additionally, none of the other distributions provided a reasonable fit. As before, I computed the median of the measurements for each year since (for many of the years) the replicate values do not exhibit appreciable variation. There is no evidence to reject the normality test on the median values (however, there is a risk of using this approach since the one large value from 2004 is now excluded from the dataset). Based on the medians, the descriptive statistics and upper confidence bounds (using the normal distribution) are:

Descriptive Statistics: Median VOC lb/hr

Variable	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Median VOC lb/hr	1.599	1.199	0.130	0.285	1.689	2.395	3.818

99% Upper Bound: 4.39

- To be conservative, the analysis was also performed on the *maximum* values for each year. There was no evidence to reject the normality assumption. Using the normal distribution, the descriptive statistics and upper confidence bounds are:

Descriptive Statistics: Max VOC lb/hr

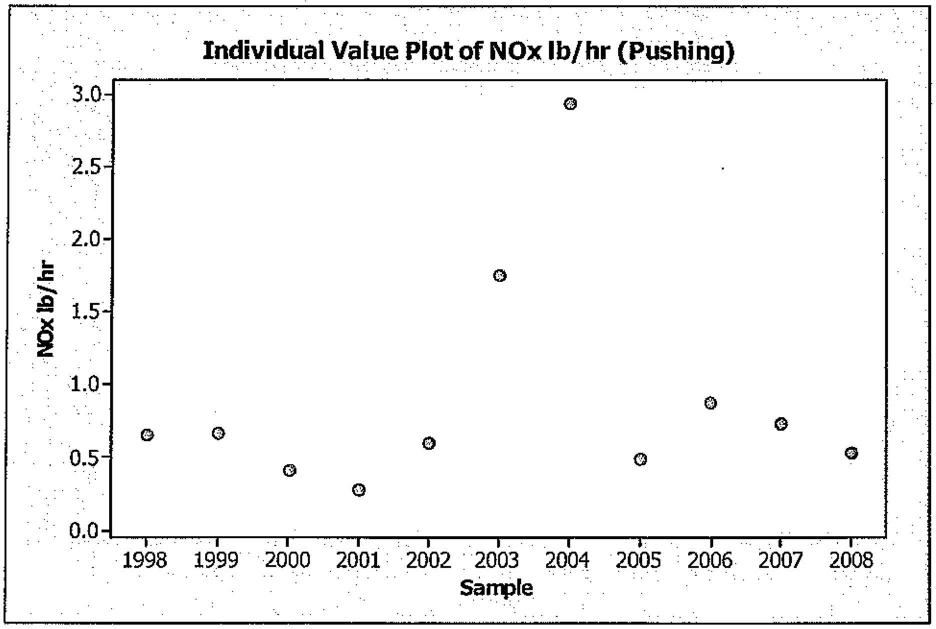
Variable	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Max VOC lb/hr	2.443	1.863	0.185	0.706	1.843	3.872	6.548

99% Upper Bound: 6.78

- The conservative approach above is recommended unless an assignable cause can be identified that justifies the removal of the large value from 2004.

Pushing Emissions Control System NOx (lb/hr)

An Individual Value Plot of the data is shown below along with some observations and the results.



- The measurements in 2003 and 2004 are statistically different than the rest of the data. It is unclear whether the process naturally produces extreme values or whether these data points represent a different process or measurement error. To be conservative, I will assume that these two large values are representative of the "process."

- A normality test was performed and there was strong evidence to reject the normality assumption. Other distributions were tested. The loglogistic distribution reasonably described this data. (The loglogistic distribution is useful to represent heavily skewed data). The descriptive statistics and upper confidence bounds (using the loglogistic distribution) are:

Descriptive Statistics: NOx lb/hr

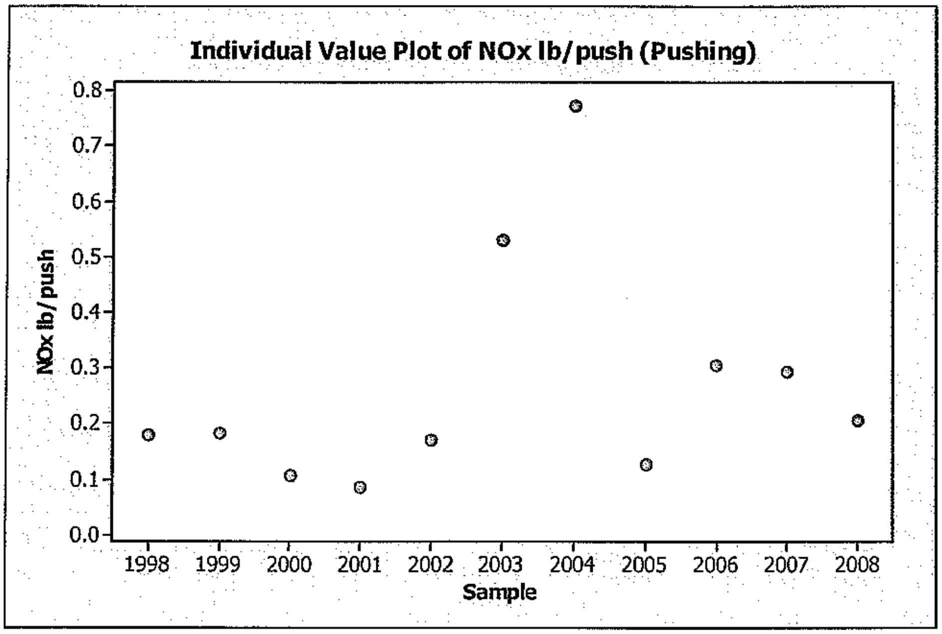
Variable	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
NOx lb/hr	0.905	0.776	0.279	0.492	0.652	0.874	2.934

99% Upper Bound: 3.11

- If you would like me to provide estimates after excluding one or both of the extreme values, please let me know.

Pushing Emissions Control System NOx (lb/push)

An Individual Value Plot of the data is shown below along with some observations and the results.



- The measurements in 2003 and 2004 are statistically different than the rest of the data. It is unclear whether the process naturally produces extreme values or whether these data points represent a different process or measurement error. To be conservative, I will assume that these two large values are representative of the "process."

- A normality test was performed and there was strong evidence to reject the normality assumption. Other distributions were tested. The loglogistic distribution reasonably described this data. The descriptive statistics and upper confidence bounds (using the loglogistic distribution) are:

Descriptive Statistics: NOx lb/push

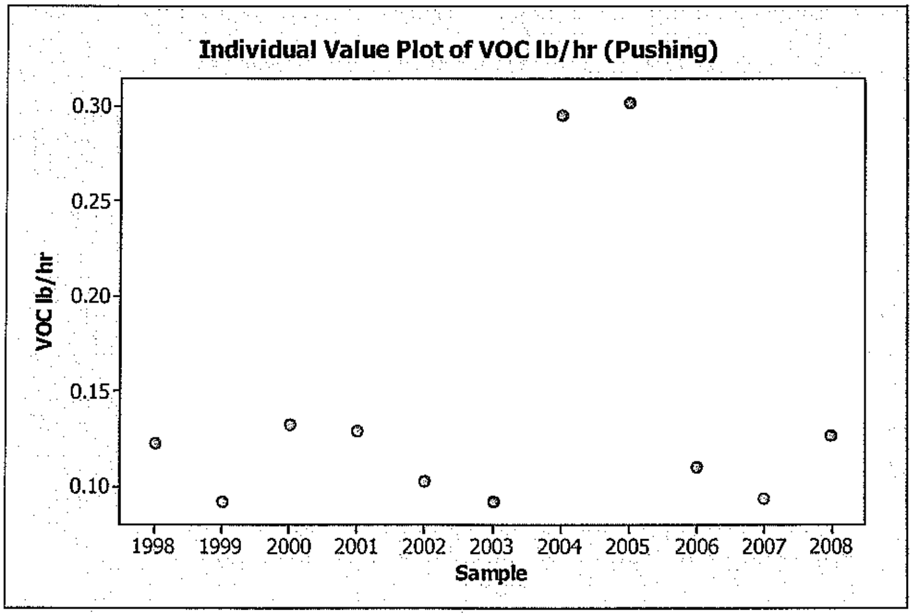
Variable	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
NOx lb/push	0.2695	0.2086	0.0869	0.1267	0.1838	0.3051	0.7742

99% Upper Bound: 1.09

- If you would like me to provide estimates after excluding one or both of the extreme values, please let me know.

Pushing Emissions Control System VOC (lb/hr)

An Individual Value Plot of the data is shown below along with some observations and the results.



- The measurements in 2004 and 2005 are statistically different than the rest of the data. It is unclear whether the process naturally produces extreme values or whether these data points represent a different process or measurement error.

- A normality test was performed and there was strong evidence to reject the normality assumption. Other distributions were tested and no plausible distributions were found that reasonably describe this data. While distributions do exists that describe skewed data, the probability is very low that 2 out of 11 points would fall at the extreme end of the tail of the distribution with a large gap between those and the remaining values. In short, it appears that the values in 2004 and 2005 come from a different process distribution than the remaining years.

- Since no parametric model can be utilized (without discarding data), a non-parametric estimate may be utilized.

Descriptive Statistics: VOC lb/hr

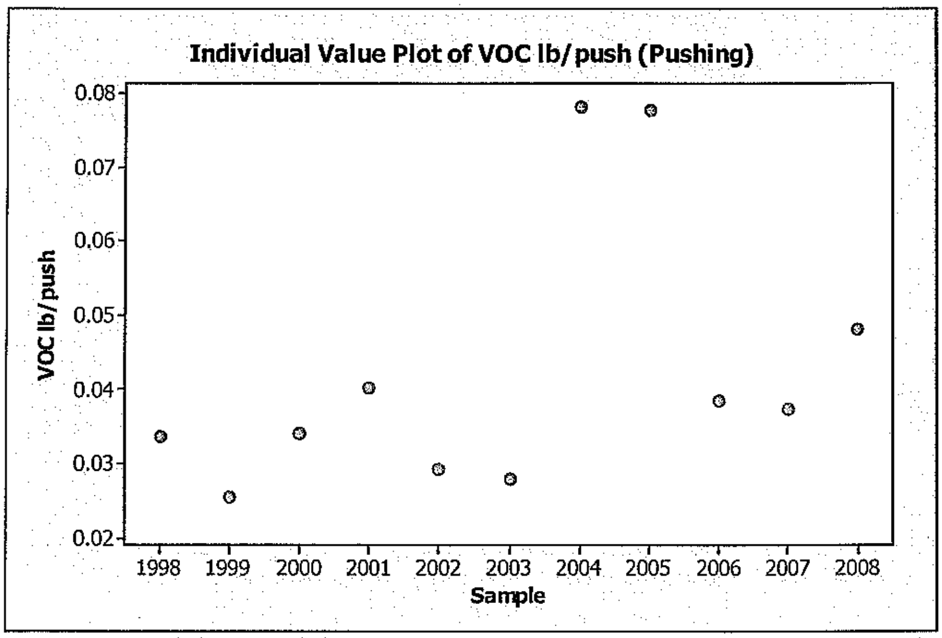
Variable	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
VOC lb/hr	0.1457	0.0773	0.0924	0.0939	0.1230	0.1331	0.3020

99% Upper Bound: 0.302

For the non-parametric estimate, the upper bound remains at the maximum (0.302) until we reach a 91% upper bound where the upper bound jumps to 0.296

Pushing Emissions Control System VOC (lb/push)

An Individual Value Plot of the data is shown below along with some observations and the results.



- The measurements in 2004 and 2005 are statistically different than the rest of the data. It is unclear whether the process naturally produces extreme values or whether these data points represent a different process or measurement error.

- A normality test was performed and there was strong evidence to reject the normality assumption. Other distributions were tested. The loglogistic distribution reasonably described this data. The descriptive statistics and upper confidence bounds (using the loglogistic distribution) are:

Descriptive Statistics: VOC lb/push

Variable	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
VOC lb/push	0.04280	0.01848	0.02552	0.02930	0.03735	0.04821	0.07810

99% Upper Bound: 0.096

I would be happy to provide additional analyses upon your request. Also, please let me know if you have any questions regarding this report. Thank you for including Integral Concepts on this project.

Kindest Regards,

Steven Wachs
Vice President & Principal Statistician
Integral Concepts, Inc.