

Hazardous Waste Incineration: Better Science is Needed

by Doug Abner

This paper was written by research physicist Doug Abner. It is in draft status and is currently undergoing peer review.

We are enclosing it with this submission to the EPA in its current state because of temporal exigencies. We would like the opportunity to submit a final draft after review if that is possible.

While this paper does not deal exclusively with Cement Kiln Dust, a great deal of it is germane to the CKD issue in addition to calling into question a lot of the poor and quasi science that has been employed with Cement kiln incineration of hazardous waste.

DRAFT

Hazardous waste incineration: Better science is needed.

We have studied the issue of burning hazardous waste from a scientific perspective and conclude that the current practice has many flaws. We examine several of the issues that are relevant to the incineration of hazardous waste in hazardous waste incinerators, cement kilns and other **BIFs**, and we believe that all that we have presented here to be correct and in good faith. We cite reference materials to substantiate the points that we make.

99.99% DRE ?

In April 1985 the Science Advisory Board (SAB) to the **USEPA** reported on the incineration of liquid hazardous waste¹. That report dealt with several issues that relate directly to protecting the public health and environment. Almost nine years have passed since the issuance of that report and still the majority of the conclusions and recommendations of that report have not been adequately addressed. A few quotes will illustrate.

“To date [1985], only a very small portion of the compounds found in emissions from incinerators has been identified qualitatively or quantitatively . . . the concept of destruction efficiency (while valid for comparing the relative operating performance of incinerators) does not completely address the problem of what is emitted from the incinerator stack and does not, therefore, constitute a reliable basis for developing exposure assessments”.

‘The existing analytical data for emissions from hazardous waste incinerators have serious limitations. Among the major problems are the limited number of chemicals selected for analysis and the fact that the analytical methodologies have not been validated either for the conditions of the test or for the complex mixtures which exist in incinerator emissions. As a result, there exist no relatively complete or reliable analyses of mass emissions from either land or sea based incinerators on which to base subsequent estimates of the potential for environmental exposures”.

“... the Agency [**USEPA**] continues to experience difficulties both in assessing and managing hazardous waste incineration programs.”

“... 99.99% destruction efficiency does not appear to be achieved if compounds other than **POHCs** in the stack gas are considered.”

‘The Agency [USEPA] should develop a revised destruction efficiency paradigm so that its assessment of incinerator performance can account for the variability of emissions and effluents.’

SIGNIFICANT FIGURES, ERRORS, ACCURACY, METHODOLOGY AND SENSITIVITY

The analysis of experimental measurements has been written about by **Barford**² and by many others. This is a well established area of scientific knowledge.

First consider a few definitions?

True value: the actual value of a quantity independent of measurement. The sought after value.

Accuracy: a measure of the closeness of an experimental value to the true value.

Error: a measure of the departure of an experimental value from the true value.

Precision: a measure of the reproducibility of a set of results from replicate runs. It does not matter how close the average of these runs is to the true value; thus, precision and accuracy are defined independently of each other.

Significant digits: in a number, all the digits that are certain plus the first uncertain digit. All digits to the right of the first uncertain digit are dropped; eg, 12.0 +/- 1 .0 is no more informative than 12 +/- 1.

Uncertain digits: any digit that can vary by at least plus or minus 1.

Uncertainty: the range a data value may have because of the need to estimate (interpolate) the last digit being read from an instrument scale; the range a calculated result may have because of the uncertainty in the data or results from which it is calculated. Thus 7.82 +/- 0.02 has an uncertainty of 2 in the hundredths position. Hundredths is the first uncertain digit.

Barford states that “a quantitative basis for deciding how many figures to retain in an experimental result” can be obtained through the use of the relative accuracy of the standard error.

We need to describe a few additional terms. The best estimate of precision is

$$s_n = \sigma / (n-1)^{1/2}$$

where σ is the standard deviation of the measured data and n is the number of data points in the data set .

The best estimate of standard error is

$$S_n = s_n / \text{sqrt}(n).$$

The standard deviation in the best estimate of standard error is

$$\sigma(S_n) \sim S_n/\text{sqrt}(n).$$

Barford replaces n by $(n-2)$ as “a more prudent estimate for the standard deviation of S_n , since at least two measurements are necessary to determine estimates of the independent quantities X and σ , three will be required to add a third independent quantity $\sigma(S_n)$. The relative accuracy of S_n is then

$$\sigma(S_n)/S_n \sim 1/\text{sqrt}(n-2).$$

Where $\sigma(S_n)$ is the standard deviation of S_n . Barford (p.42) gives the example:

that if **100 measurements** gave the values $X_n = 17.366$ (where X_n is the best combined estimate of the true value from experiments with n measurements of the same quantity known as the mean), $s_n = 1.226$ and $S_n = 0.1226$ then the **relative accuracy is around 10%**. In other words the true value of S_n has a roughly 2: 1 chance of lying within the interval 0.11034 and 0.13466. It is clear that the third and fourth figures in s_n and S_n are of no practical significance, and that even the second is in some doubt. An adequate summary of the result would be $X = 17.39 \pm 0.13$, where the standard error is given to two figures, and X_n is adjusted to the same significant figure; it is probably true that even $X = 17.40 \pm 0.15$ would lose little worthwhile information from the experiment.”

Bat-ford (p.42) generalizes this to say that “**In any experiment involving ten or fewer measurements, where the errors are estimated solely from these [ie statistics performed on the measurements], there is no point at all in giving the precision of the apparatus to more than two significant figures or the standard error of the final result to more than one figure**”. Note that the three data points that are typical for determinations of DRE are not adequate to calculate accurate statistics which are necessary to determine the number of digits to present in an experimental measurement. If an estimate of the significant figures is obtained in another way it must be scientifically valid and open to public scrutiny. At present we know of nothing other than cursory examinations of the error in Destruction and Removal Efficiency (DRE) based on

what we believe are inadequate statistics. We believe that there is no rigorous and scientifically justifiable error analysis of DRE measurements at present. Even with good statistics one must still question if the measurement methodology faithfully represents the true value.

When we apply relative accuracy to the DRE testing in cement kilns and other **BIFs** where typically only three data points are collected for a determination of DRE, the relative accuracy does not justify accepting all four nines. **Barford's** treatment would justify only one of the nines, ie ninety-something% DRE. To demonstrate this, recall that the relative accuracy is $\sigma(S_n)/S_n \sim 1/\sqrt{n-2} = 1/\sqrt{3-2}=1$. So the accuracy of the DRE result is to the first digit only, i.e. **ninety-something%** DRE. Referring to Figure 1 we have plotted the relative accuracy versus the number in the data set. A significantly larger set of DRE measurement data than 3 would be required to scientifically demonstrate a precision of 99.99% DRE.

The Federal Register ³ defines DRE as

$$\text{DRE} = (1 - (\text{Wout}/\text{Win})) * 100\%.$$

Note that here we have added parentheses around **(Wout/Win)** to avoid any ambiguity. Also, note that until recently that the formula printed in the Federal Register was subject to a printing error which would give wrong answers if used to compute DRE and that some **BIFs** have published forms of the DRE equation that are algebraically equivalent to the proper form, but give incorrect results under significant figure analysis.

Section **266.104a** of the Federal Register states that "a boiler or industrial furnace burning hazardous waste must achieve a destruction and removal efficiency (DRE) of **99.99%** for all organic hazardous constituents in the waste feed. To determine conformance with this requirement, 99.99% DRE must be demonstrated during a trial burn for each principal organic hazardous waste constituent (POHC) designated.. "

Taking typical data ⁴ from a hazardous waste burning cement kiln for three runs we have the feed rates for monochlorobenzene

$$\text{Win} = (62.9, 62.9, 62.9) \text{ lb/hr}$$

and the detection limit based emission rates

$$\text{Wout} = (0.0000643, .000130, .000509) \text{ lb/hr.}$$

When we calculate the values for DRE we get

DRE=(99.999898, 99.99979, 99.99919) %.

Note that during the very controlled test burn the Wout varied by close to 700% while Win remained fixed, but the DREs calculated with these numbers show little variation. By how much does the actual emission vary during normal day-to-day operation?

The above DRE numbers are expressed to the number of digits allowed by standard significant digit analysis but that is not all that is required to establish the validity of data. As an example consider the following significant digit calculation of DRE for **Win=6.43 X10⁻⁵** and **Wout=62.9** both in lb/hr.

We first perform the division. Both Win and Wout have three digits so their division will have three digits. Next, this three digit number which has seven digits to the right of the decimal is subtracted from the number one (1 .0000000) which is a **so-called** exact number. The resultant significant digits are expressed to 0.9999357. We next multiply by the exact number of 100% to get 99.99357% DRE. These significant numbers were handled in the standard way, but that is not all that is required to establish a scientifically valid **99.99% DRE**.

The significant digits alone in the above calculation ignores the statistics so, although 7 significant digits seem to be indicated, the bias of typical highly idealized kiln test conditions does not reflect the true nature of normal operation and inadequate statistics (refer to the relative accuracy described above) does not reflect the true nature of the numbers measured during a kiln test.

Let us further investigate the statistics of measurements. Figure 2 shows data that we simulated which was chosen to exhibit a normal (Gaussian) distribution. We chose a mean value of 0 and a standard deviation of 1 to generate this synthetic data. Also included in this graph is the analytic curve for the distribution. The number of data in this sample was 1000. We next took a subsample from this set and calculated its mean and standard deviation. The values chosen for the number of data in the subsample were **3,4,5 . . . 500**. Each subsequent calculation of mean, standard deviation and the standard deviation of the mean was calculated ten times so that statistics could be performed to give error bars at each value of subsample size. Ten here is adequate for the qualitative results presented here. A quantitative study would require more. Figure 3 shows the results of our analysis of this simulated data. Referring to Figures 2 and 3 we know what the true value is in this simulation experiment. It is the value that we chose for the mean. We also know what the precision is expected to be through the standard deviation. For only a few data in the

subset, the error bars from the standard deviation of the set of ten calculations of the mean for each size of subsample are large. As the sample size increases the mean becomes stable as it approaches its true value of 0 here. We see that the error bars also decrease as the number of data in the subsample increases. Note for samples containing more than about 100 data points that the error bar is within about **10%**, but for only three samples it has **+/- 60%** error bar. The size of the data sample helps to build a credible measurement.

Since the distribution function of DRE from a kiln test is not known one cannot predict the true DRE value. Note that the real distribution may be other than Gaussian.

When scientists set out to do an experiment they carefully consider the design of the experiment. The design is important because it must faithfully represent what Nature is doing, regardless of what is believed at the time. Also the methodology employed and the data analysis used are **crucial**, if the truth is to be revealed. One must certainly take into account the sources of errors both systematic and random as well as the number of significant digits expressed in the numerical results obtained with a statistical analysis using a valid number of data points. These things contribute toward an objective determination of the truth in the experimental findings, otherwise a predetermined conclusion may be reached and the truth may never come to light. When science has been served it has enabled humankind a very powerful tool, to be used wisely or foolishly, but nevertheless the truth remains.

In the quotes from the SAB that started this document, the concern about the lack of adequate methodology was expressed as “the existing analytical data for emissions ... have serious limitations”. In actual kiln tests only a very few Principle Organic Hazardous Waste Constituents (POHC) are used. We know of no scientific validation for choosing so few **POHCs** to represent what are typically over 200 possible chemicals in a real waste fuel stream.

From a scientific standpoint combustion is a very complicated process. Only a few relatively simple types of combustion are completely understood by the scientific community. Since the approximately 200 or more components in hazardous waste fuel represents a much more complicated and **variable** system than those that are currently understood, it is very likely that the hazardous waste incineration systems are not adequately understood. Experiments¹³ (p.874) “have indicated that the dependence of destruction on concentration for benzene, chloroform and chlorobenzene is stronger and can be accelerated with the addition of a co-oxidizing fuel at higher concentrations in **order to** generate higher concentrations of **radicals...these** data suggest that kinetic limitations can be important at low concentrations even at the high temperatures normally encountered in waste combustion systems.” Another scientist writes¹⁴ (p.899): “The production of **C₂H₃Cl** [vinyl chloride] from **C₂H₅Cl** [ethyl chloride]

demonstrates that it is possible to form hazardous compounds from relatively innocuous chlorinated species.” It is likely that this happens when chlorinated hazardous waste is burned in an incinerator, cement kiln or other BIF. The question is, does anyone really know what hazardous compounds are formed? Simulation, and other scientific guesses are often employed to arrive at some conclusion. Has there ever been a rigorous accounting of what goes into a BIF and what comes out, including the new **chemicals** that are formed?

In science the conclusions must be carefully established and must not go beyond what the data supports. It is very difficult to be convinced that a BIF test under very controlled conditions and conducted during a very short period of time is representative of daily operation. It is also very difficult to believe that risk assessments demonstrate public safety when they rely on assumptions that minimize the calculated risk, exclude the spectrum of humanity in favor of a “standard simulated human”, average exposures instead of using peak and time weighted exposures, exclude non-cancer effects, and offer no ongoing health studies. The latter would be in any case equivalent to an unethical experiment under the situation of a total lack of good scientific procedures.

RADIOACTIVE CONTAMINATION OF HAZARDOUS WASTE

Some permits require that the hazardous waste undergo on-site radioactive screening to determine if it is contaminated with radioactive waste. Since radioactive contamination has been found in cement kiln dust (CKD) ⁵ and was reported to Congress for a limited sample of kilns across the U.S. it is of high importance that the testing of hazardous waste for radioactive contamination be done in an adequate manner. The Agency conducted isotopic analysis¹⁸ for radium226, plutonium238, plutonium239, uranium234, uranium235, uranium238, thorium227, thorium228, thorium230 and thorium232.

The radioactive contamination in CKD has most recently been attributed by the **USEPA** to “prior release in the environment, for example, from fallout from above-ground nuclear weapon’s testing and from the explosion of a satellite containing plutonium in the earth’s atmosphere”? We believe that other facts presented here show another picture.

Figures 4,5 and 6 show the three principal radioactive decay chains found in nature.

Figure 7 and Table 1 show that many different isotopes are created when ²³⁵U and ²³⁹Pu undergo fission, as in a nuclear reactor or an above ground nuclear detonation. Table 3 is a compilation of the main fallout products from references given here. Note that **Strontium90**, **Cesium137**, **Cerium144** and **Samarium151** are the long lived fission products created in peak abundance. Yet these, excluding the Cesium137, were not found when the **CKD** was analysed. Also, table 2¹² shows that some of the radioactive isotopes that were reported to