Confidence Level Assessment of the Blend Down Monitoring System

N.J. Devaser[1] (Author/presenter)
J.A. March-Leuba[1], B.R. Upadhyaya[2] (Co-Authors)

[1]*Oak Ridge National Laboratory, Oak Ridge, Tennessee*
[2]*University of Tennessee, Nuclear Engineering, Knoxville, Tennessee*

Abstract

The objective of this research is to perform a statistical analysis on the fissile mass flow monitor (FMFM) associated with the Blend Down Monitoring System (BDMS). The BDMS is one of the transparency measures taken by the United States in securing the "Megatons to Megawatts" agreement made with Russia in 1993. The purpose of the analysis is to identify expected statistical behavior of the measurements obtained and to furthermore to categorize a threshold for separating usable from unusable data. The confidence measurement used in verifying the existence of flow, providing the probability of flow, is referred to as "quality." From this research, we conclude that any quality measurement having a value greater than 20% provides high confidence of the presence of flow of fissile material in the pipe. Furthermore, any quality value less than this threshold may be regarded as statistical noise mistakenly identified as a measurement.

The confidence measurement quality is mathematically defined as a stretched, inverted Snedecor Fisher F-test[1] which tests the inequality of residuals from actual BDMS data to residuals of zero flow data. A calibration method is taken in determining the threshold value for usability in which a simulation is set to zero flow and a histogram of the quality values obtained produces a distribution surrounding zero (i.e. nonzero). At this point, a range of probability values exists between quality values of zero and approximately 30%. A significance level can be concluded by counting those below some threshold probability (0.5% in this case) out of the number of tests taken. If N tests are taken, and n tests out of these have less than 0.5% probability of having a value greater than 20%, then our significance level is n/N%. From this assessment, this research has concluded at a 95% significance level that less than 0.5% of 600,000 tests have a quality value greater than 20% at the zero flow assumption. The outcome of this work will greatly help the quality of the measurements obtained from the BDMS.

INTRODUCTION

Background of BDMS

In 1993, the United States and the Russian Federation signed an agreement for the exchange of 500 metric tons highly enriched uranium (HEU) to be down blended from dismantled Russian nuclear warheads and used in American nuclear reactors. The agreement is contingent upon transparency measures with the sole purpose of building confidence in the belief that the arms control and nonproliferation objectives are being met. The primary measures are to ensure the following:

1. That the low enriched uranium (LEU) produced as a result of the agreement is in fact coming from the down blending of nuclear warheads
2. That the HEU entering the oxidation process is the same HEU from these warheads

3. That the LEU entering the United States is in fact fabricated into fuel and used in American nuclear reactors

BDMS serves as a primary vehicle in accomplishing these goals as it is used in verifying the first of these assurances. The system verifies the following two essential parameters that aid in the determination of the flow characteristics at each Russian site:

- Enrichment, *determined by the Enrichment Monitor **(EM)***
- Fissile mass flow rate, *determined by the Fissile Mass Flow Monitor **(FMFM)***

The focus of this research is on parameters obtained from the FMFM.

Fissile Mass Flow Monitor (FMFM)

The FMFM allows a neutron source to emit neutrons into the fissile flow periodically and follows the gamma rays emitted by the resulting fission products. The signal obtained later in the stream will have particular amplitude proportional to the fissile concentration and particular time delay inversely proportional to the flow velocity. The product of these two values provides the fissile mass flow rate.

The FMFM consists of a 12-μg californium-252 source, a neutron-absorbing linear positioner, a polyethylene block, and a bismuth germanate (BGO) detector system. The neutron absorbing linear positioner modulates the flow of neutrons out of the polyethylene shell and into the pipe and obtains a time signature for the event. Moderated neutrons entering the pipe induce fissions inside the process stream. Some of the fission products are slowed down by the gas and some are carried by the stream. Later in the stream, the BGO detector system obtains counts from the delayed gammas resulting from the remaining fission products. The time signature at this point is compared to the time signature from the positioner to obtain the time delay

The BGO detector's large spectral coverage of approximately 150 keV to 30 MeV makes it ideal for the project's purposes. This particular range is important because it includes the 186-keV gamma released from ^{235}U as well as the required range for the remaining fission products surviving the long range between the source and the detector system.

Quality

Quality is a statistical test on the flow of fissile material; it represents the qualitative value of a particular measurement. Numerically, it provides the probability that fissile material is present and supplies a confidence level on the fit of the data analysis algorithm. The algorithm operates by obtaining a time-dependent profile every 60 seconds and performing analysis on a running average of all data profiles thus far within two larger time blocks, τ_1 and τ_2 ($\tau_1 < \tau_2$). Both are used to analyze convergence behavior of different parts of the algorithm. The large time blocks contain average data over anywhere from several hours to several days. Each 60-second profile obtained undergoes the following steps before coming to the calculation of quality:

1. A least-squares fit is used for which a correlated background, uncorrelated background, and fissile concentration are found, minimizing the residual ε as a function of velocity, *u*. *M(t)* is the set of data averaged over τ_1 and τ_2.

$$\varepsilon(u) = \int_0^{20} dt\, \left[M(t) - \left(bckgU + bckgC(t) + C(u)\; N_{model}(t,u)\right)\right]^2 . \qquad \textbf{(1)}$$

2. The resulting residual is selected, and the optimal velocity and appropriate least-squares parameters are the outcome

In the least-squares approximation, it is assumed that a profile exists in the data; the test of quality comes about by questioning this assumption. The residual previously determined, $\varepsilon(u)$, is compared to a residual of similar analysis excluding the $C(u)$ term. This sets the profile term to 0, forcing the assumption of the absence of a model-predicted profile. An F test is performed to determine the probability that the two sets of numbers come from different distributions. Since a typical F test is used to establish equality, this particular test is altered to fit our particular criteria.

THEORY

BDMS Software

The following steps take place in the computation of the mass flow rate.

1. *Raw data is obtained from the detector system in 60-second time blocks*

The linear positioner slides open and closed every 5 to 10 seconds in a 60-second period this leaves about 3 – 6 periods for pattern recognition.

2. *This data is averaged with previous 60-second time blocks over two larger time blocks*

A running average is kept over two different large time blocks. Inherent in the some of the physics is the capability of faster convergence. For instance, the emission rate of 186-keV gammas from ^{235}U is so incredibly high that the software needs only average this data over 12 hours due to the excessive amount of precision coming from the law of large numbers. As expected, the length of these blocks is left to the user but typically resides around 4 and 12 hours.

3. *A least-squares fit is calculated as a function of profiles at different trial velocities (time delays).*

The least-squares approach fits the raw data to a constant uncorrelated background, a time dependent correlated background, and a time-dependent, velocity-dependent fissile concentration. At each trial velocity, the residual is minimized, resulting in a fissile concentration, correlated background, and uncorrelated background obtained from each profile as a function of velocity (time delay).

4. *The result is a vector of potential residuals at different velocities (time delays), for which a global minimum is found and the estimates attained are amplitude parameters of fissile concentration, correlated background, and uncorrelated background*

The algorithm uses the trial velocities as major points and interpolates between them to find the global minimum. This minimum is regarded as conclusive, and these amplitude parameters are only again tested by use of quality.

5. *Quality of the fit is then calculated*

Every 60-second block is matched with its minimum error residual and all the steps described above occur continuously, making measurements and averaging them over longer time periods. Thus, within a large time block, all new measurements have memory and, if necessary, can be used independently to describe behavior of the data up until that point.

Simulator Process

The use of BDV1.1 is very practical due to the difficulty of obtaining large amounts of data from the field. The Megatons to Megawatts agreement only allows data to leave each Russian plant on a daily scale and thus smaller scale data must be simulated for the purposes of research.

Outside of the development of fake flow and fake background and noise, the rest of the simulator system is identical to that in the field; the same software is used. Over time, the simulator has been verified as having high agreement with the field data. The development of the fake flow is performed as follows:

1. *Positioner open and closed times are enforced*
2. *A correlated background profile is generated from the positioner times*
3. *Enforced velocity (typically a benchmarked parameter) is used to calculate time delay*
4. *Profile is interpolated from this time delay or velocity*
5. *Fake count rate is calculated from the sum of the profile points and specified multipliers in regard to local surroundings. Gaussian noise is added as uncorrelated background to the count rate*

$$cps_i = \{eff_{\det}\}\{m\}x_i + Bg_{i,correlated} + Bg_{uncorrelated} \qquad \textbf{(2)}$$

where

$\{eff_{\det}\}$ = the average detector efficiency;

$\{m\}$ = a multiplier accounting for profile speed correction, leg – specific concentration correction (another pressure correction), source decay correction, and fission fragment length correction;

x_i = i^{th} data point in profile;

$Bg_{i,correlated}$ = product of i^{th} correlated background weight and amplitude parameter from concentration data; and

$Bg_{uncorrelated}$ = product of white noise from a random noise generator and amplitude parameter from concentration data.

BDV1.1 performs NTESTS (a user-selected number) independent simulations. For each run of NTESTS tests, a single distribution is obtained which represents data over a long period of time coming from one machine. The software continues to do this NRUNS times (also user selected). At the end of one run of the entire BDV1.1, NRUNS distributions are acquired for which NTESTS tests were taken on each. In the field, this would amount to having NRUNS machines running over NTESTS minutes, resulting in NRUNS*NTESTS total tests taken.

QUALITY

In short, quality measures the probability of flow in the FMFM by using an augmented Snedecor-Fisher F-test to compare residuals from a least-squares fit of the raw data to a model-predicted profile with background and a least-squares fit with only background (i.e. comparing the assumption of flow to the assumption of no flow). Since it is the objective of BDMS to obtain confidence in the existence of flow, quality is one of the project's more important confidence measurements.

Augmented Snedecor Fisher F-test

An F-test is typically performed when attempting to equate the source of two sets of numbers; this can be viewed as a hypothesis test with the null hypothesis being that the two sets originate from the same distribution. It is desired that quality assume inequality and test the probability of that inequality. Thus, the F test used in determining quality is inverted to allow the probability of *inequality* to be obtained. Also, in the event that the two residual values are equal, a typical F test comes to the conclusion that there is just as much a probability of equality as inequality (of the two sets origins). For this reason, not only does inversion take place but the distribution has also been stretched making the point of equal probability, zero probability. The algorithm regards any negative value as zero.

The algorithm performs the F test on the ratio of the no-flow variance to the flow variance.

$$Q = 1 - 2F\left\{\frac{\sigma^2_{no\ flow}}{\sigma^2_{flow}}\right\}, \qquad \textbf{(3)}$$

where $\sigma^2_{flow}, \sigma^2_{no\ flow}$ are the unbiased variances of the fits containing the flow term and the no-flow term, respectively

Figure 1 shows the distribution of qualities over 10^5 tests run by BDV1.1. Note the exponential behavior as well as the near zero values at around 15 or 20. This shows potential for a fairly low threshold, meaning confidence in flow need only values greater than about 20% quality to fall outside the spurious range.

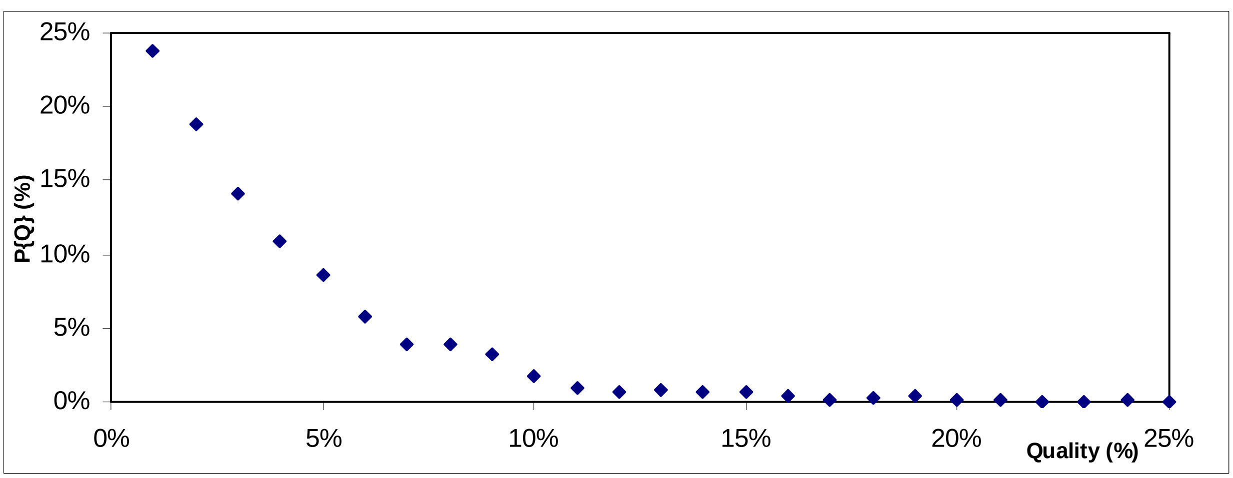

Figure 1: Probability Distribution of Simulated Quality Values

The convenience of this distribution is its smoothness, providing the potential capability of taking an analytical approach to defining a threshold.

Threshold Development

The intention of assuming no flow is to force the data points obtained in simulation to 0. This is carried out by the concentration multiplier in the fake data computation being set to 0 causing the fake data sum to be composed only of background.

In determining the threshold, NTESTS is set fairly high to obtain a solid distribution of possibilities and a distribution of converged data from the quality histogram. When this is done, the figure obtained looks very similar to figure 1, exponential in nature and dying out at approximately 20%- 30% quality. The distribution apparent is of the form

$$p(Q) = re^{-rQ}. \qquad \textbf{(4)}$$

Analytically, a threshold can be determined by finding the cumulative distribution function and selecting a point.

$$p_{CDF}(Q) = \int_0^Q re^{-rQ'} dQ' = \left(1 - e^{-rQ}\right). \qquad \textbf{(5)}$$

Selection of a threshold using such a distribution is left to the researcher and can be simply calculated from the knowledge of the exponential rate *r*. If a resulting probability of 1% is desired, the threshold would be determined from the solution of this equation to a probability of 1%.

Significance defines the confidence in a certain threshold selection and is helpful, if not necessary. Derivation of this significance level is straightforward but relies on the use of a high value of NRUNS. The derivation begins with selecting some threshold for probable insignificance, τ, a probability that can be regarded as too small to be considered. BDV1.1 is then run with NRUNS set at some value M. When BDV1.1 is run at no flow, M distributions result, each having a specific probability of obtaining some desired threshold T. Of these M data sets, n of them have a value less than τ, and $M - n$ of them have a value greater than τ. The significance level is simply the number that falls below τ divided by the total number of runs. The final statement of such an analysis is then the following:

Statement: The data show $\frac{n}{M}$ % confidence that the probability of obtaining T is less than τ.

RESULTS

The figure below shows the distribution of qualities obtained at zero flow and an appropriate exponential fit. Note that the probability drops dramatically at about 15% quality.

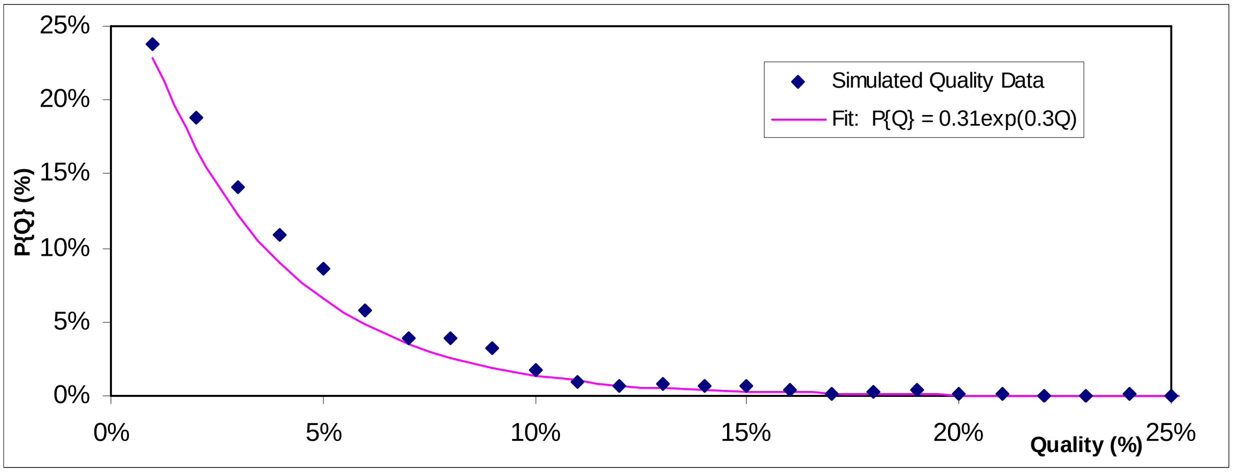

Figure 2: Probability Distribution including Exponential Fit

As an indicator of the goodness of fit, the expectation values of quality for the two distributions (simulated data and fitted data) are 4.33 ± 5.37 for the fitted data and 4.0 ± 5.16 for the simulated data. This is approximately 8% ± 4% difference between the data and the fit.

Threshold Results

During this research, BDV1.1 was run hundreds of times, with NTESTS ranging from 10 to 10^6. Also, NRUNS ranged from 1 to 50. In figure 2, NTESTS was set to 10^4 and NRUNS was set to 10. In the field, these data would describe 10 different systems running for 10^4 minutes. The threshold chosen for use in the field was 20%.

Tabulated results of choosing 20% as a threshold are shown below in Table 1. If a probable insignificance of 0.6% were chosen, out of these 10 tests, the significance level is 9/10 = 90%. Typically, for tests in this research, the probable insignificance has been set to 1%, for which the below set would have a significance level of 100%.

Table 1. Probability of obtaining at least one quality value over 20% over 10 runs

Run	$P\{Q \geq 20\}$
1	0.27%
2	0.54%
3	0.46%
4	0.23%
5	0.62%
6	0.54%
7	0.24%
8	0.46%
9	0.24%
10	0.35%
Average	0.40%

A second test, in which NRUNS and NTESTS were both set to 100, at a 1% probable insignificance resulted in having a significance level of 95%. Interpretation of these results is that only 5 out of 100 tests have a probability of obtaining greater than 20% quality at zero flow.

$$P\{Q\} = 0.27e^{-0.27Q}. \quad \textbf{(6)}$$

$$P_{CDF}\{Q\} = \left(1 - e^{-0.27Q}\right). \quad \textbf{(7)}$$

If the exponential fit is used to determine the probability of obtaining anything under the threshold the result is approximately 0.4%.

CONCLUDING REMARKS

Though it is complex in nature, the information provided by the confidence measurement quality is irreplaceable. In this paper, it is shown from the software BDV1.1 that the primary confidence measurement used in the FMFM yields a decaying exponential distribution at the assumption of zero flow. Naturally, it is expected that a confidence test under these assumptions would yield a nonzero distribution with the use of stochastic data. Furthermore, it is expected, even necessary, for the distribution to decay at this assumption.

The resulting exponential resulted in falling around 30% amplitude and rate, and the threshold resulted in being set to 20%, meaning that if a measurement quality reading gets below or equal to 20%, flags will be raised. In actuality, flags would be raised at a measurement even as high as 50% quality, but since the range of background falls short of being *that* extensive, the possibility of falsified data would be slim. A quality value as low as 20% leaves a large possibility of only background being read, and thus investigation should follow.

The choice of 20% for the threshold was proven sound at a probable insignificance of 1%. The result was that less than 5 of 100 tests yielded a value greater than 1%. The significance level used was simply a ratio test designed for this research and served the purpose of conveying the probability of obtaining a value within certain limits.

REFERENCES

1. Dunlop, Dorothy D., and Ajit C. Tamhane. Statistics and Data Analysis: From Elementary to Intermediate, 2000.
2. Knoll, Glenn F. Radiation Detection and Measurement. Ann Arbor, Michigan: The University of Michigan, 1979
3. March-Leuba, J., *et al.*;"**Detector system for monitoring fissile mass flow in liquids and gases,**" Nuclear Science Symposium, 1997. IEEE Transactions 9-15 Nov. 1997 Page(s): 197 - 200 vol.1