

LA-UR-98-3061

Approved for public release;  
distribution is unlimited.

Title:

A Study of Total Measurement Error in  
Tomographic Gamma Scanning to Assay  
Nuclear Material with Emphasis on a Bias  
Issue for Low-Activity Samples

CONF-980733--

Author(s):

Tom L. Burr  
David J. Mercer  
Thomas H. Prettyman

Submitted to:

39th Annual Meeting of the Institute  
of Nuclear Materials Management  
Naples, Florida  
July 26-30, 1998

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

MASTER

**Los Alamos**  
NATIONAL LABORATORY

Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by the University of California for the U.S. Department of Energy under contract W-7405-ENG-36. By acceptance of this article, the publisher recognizes that the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy. The Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.

# A Study of Total Measurement Error in Tomographic Gamma Scanning to Assay Nuclear Material with Emphasis on a Bias Issue for Low-Activity Samples

Tom L. Burr, David J. Mercer, and Thomas H. Prettyman  
Los Alamos National Laboratory, Los Alamos, New Mexico 87545 USA

## Abstract

Field experience with the tomographic gamma scanner to assay nuclear material suggests that the analysis techniques can significantly impact the assay uncertainty. For example, currently implemented image reconstruction methods exhibit a positive bias for low-activity samples. Preliminary studies indicate that bias reduction could be achieved at the expense of increased random error variance. In this paper, we examine three possible bias sources: (1) measurement error in the estimated transmission matrix, (2) the positivity constraint on the estimated mass of nuclear material, and (3) "improper" treatment of the measurement error structure. We present results from many small-scale simulation studies to examine this bias/variance tradeoff for a few image reconstruction methods in the presence of the three possible bias sources.

## Introduction

Tomographic gamma scanning (TGS) is a  $\gamma$ -ray nondestructive assay (NDA) method to assay special nuclear material (SNM) in heterogeneous samples, particularly residues and waste. The principle of the method is that the rate of  $\gamma$ -ray emission is roughly proportional to the total SNM mass  $T$ . However, sample-specific attenuation of the  $\gamma$ -rays complicates the relation between the  $\gamma$ -ray emission rate and  $T$ . Furthermore, because the samples could be heterogeneous, both the  $\gamma$ -ray attenuation and source rate vary within the sample. Therefore, TGS uses tomography to form three-dimensional images of the attenuation. In effect, the attenuation coefficient is estimated in each of many small-volume elements of the sample. An isotopic transmission source that emits more than one  $\gamma$ -ray (usually  $^{75}\text{Se}$ ) is used to obtain attenuation images as a function of energy. The emission images are then corrected for the attenuation of  $\gamma$ -rays by using the linear attenuation coefficient images. The total amount of radioactivity (or  $T$ ) in any region of interest in the sample can then be estimated by integrating the emission image over the volume of the region. In this paper, the region of interest is the entire sample. The goal is to study the performance of candidate analysis methods in estimating  $T$ . See [1–2] for more detail and caveats about where TGS is applicable.

## TGS Image Reconstruction

The volume of a 55-gal. drum is typically divided into  $N = 1600$  three-dimensional image elements (voxels). In a standard-scan protocol, data are collected at 150 individual points in polar coordinate (displacement-angle) space for each of 16 vertical layers, giving a total of  $M = 2400$  measurement bins. During an initial scan, transmission measurements are made using an external  $^{75}\text{Se}$  source to characterize the  $\gamma$ -ray attenuation of the drum. This allows reconstruction (estimation) of the so-called system matrix  $A^{M \times N}$ . Because of  $\gamma$ -ray interactions that affect  $\gamma$ -ray energies, the gamma count rate at a particular energy channel includes the effects of gammas that originated with higher energy but appeared at the given energy channel. The simplest way to account for this underlying background is to measure the (background)  $\gamma$ -rays in energy channels near the channel(s) of interest. Then the net gamma count rate at a given bin satisfies  $n = g - cb$ , where  $g$  is the gross counts in the energy region of interest (ROI) energy channels,  $b$  is the background counts near the ROI, and  $c$  is the ratio of the number of peak ROI channels to the number of background ROI channels. Also, the detection rate of  $\gamma$ -rays must be corrected to a full-energy interaction (FEI) rate that accounts for losses attributable to dead time and pulse pileup (detector response issues). The FEI can be estimated by using a  $^{109}\text{Cd}$  source that emits an 88-keV  $\gamma$ -ray [1] and is defined by  $\text{FEI} = \text{CF}(\text{RL}) \times \text{net}$ , where  $\text{CF}(\text{RL})$  is the estimated correction factor for rate loss. Following [1], we will include  $\text{CF}(\text{RL})$  in the definition of the  $A^{M \times N}$  matrix, which means that when we consider estimation errors in  $A^{M \times N}$ , we must include estimation errors in  $\text{CF}(\text{RL})$ . The image reconstruction problem can then be cast most simply as

$$A^{M \times N} \mathbf{x}^N = [\mathbf{g}^M - c\mathbf{b}^M] + \mathbf{R}^M, \quad (1)$$

where  $A$  is the "system-matrix,"  $\mathbf{x}$  is a vector of nonnegative values that describes the distribution of  $\gamma$ -emitting material within the drum,  $\mathbf{g}$  ( $\mathbf{b}$ ) is the vector of gross (background)  $\gamma$  counts at the target energy channel,  $c$  is the ratio of the

## DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

## **DISCLAIMER**

**Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.**

number of ROI channels to background channels, and  $\mathbf{R}$  is a vector of residuals. Our problem is to estimate total mass  $T = \sum x_i$  (more generally,  $T = \mathbf{q}^T \mathbf{x}$  for some weight vector  $\mathbf{q}$ , but here we assume  $T = \sum x_i$ ).

Many of our methods can enforce the  $x_i \geq 0$  constraint in various ways. The ordinary least squares (OLS) estimates of  $x_i$  can be negative, so the simplest approach (which we use) is to use  $\max(0, x_i)$  to enforce the nonnegativity constraint. The methods that work with either the Poisson probability structure (for example, maximum likelihood which [1] refers to as MLEM) of the observed data or that plus a prior probability for each  $x_i$  on  $(0, C)$  for some large upper limit  $C$  (some Bayesian methods [3]–[5]) deal most naturally with the nonnegativity constraint. For example, we can modify Eq. (1) in [1] to work with  $\mathbf{g}$  so that the probability of  $\mathbf{g}$  (likelihood), given  $\mathbf{x}$ ,  $A$ ,  $\boldsymbol{\mu}_b$ , and  $c$ , is

$$P(\mathbf{g} | \mathbf{x}) = \prod \exp(-\mu_{g_i}) \mu_{g_i}^{g_i} / g_i! , \quad [2]$$

where  $\mu_{g_i} = \sum_j A_{ij} x_j + c \mu_{b_i}$  is the mean of the gross counts at bin  $i$  due to all voxels. Reference [4] presented a way to view the transition from maximizing the likelihood in Eq. (2) to maximizing a suitable posterior probability for  $\mathbf{x}$  that involved a temporary assumption that we could see the contribution at bin  $i$  from each individual voxel  $j$ . Note, however, that the mean for  $g_i$  at bin  $i$  is generally affected by more than one voxel  $j$ .

### Statistical Issues

We assume that  $M > N$  in Eq. (1). Thus, OLS or weighted least squares (WLS) is one option for estimating each  $x_i$  and therefore also  $T$ . In fact, Eq. (1) is essentially the same as that which commonly appears in a typical two-stage calibration experiment. Stage 1 is the equivalent of our "estimate A stage" and Stage 2, which concerns us here, uses the estimated  $A$  to estimate  $T$ . The unique features of our application of Eq. (1) are the following:

- (3) The dimension of  $A$  is very large ( $M$  by  $N$  is approximately 2400 by 1600) and  $A$  is ill conditioned. We define  $\text{cond}(A) = \lambda_{\max} / \lambda_{\min}$  (the ratio of largest to smallest singular value of  $A$ ).
- (2) There can be significant spatial correlation among neighboring  $x_i$ .
- (3) The error structure of the net response is nonstandard:  $n_i \sim \text{Poisson}(\mu_{g_i}) - c \text{Poisson}(\mu_{b_i})$ . When  $\sum_j A_{ij} x_j$  is small, it might be important to use the Poisson distribution rather than an approximating Gaussian.
- (4) There can be nonnegligible errors in the  $A$  matrix.

One TGS system ([2]) deliberately collapses some bins to reduce variance at the expense of slightly increased bias in estimating the  $A$  matrix. The usual bias/variance tradeoff suggests that this is a good idea. To date, we have ignored the possibility of bias in the  $A$  matrix. That is, our errors in  $A$  are all modeled as random errors (with standard deviation denoted  $\sigma_A$  here) so future work must include both (1) treatment of possible bias in some entries of the  $A$  matrix and (2) a plan for dealing with dynamically changing dimension of the  $A$  matrix due to bin collapsing.

We define our performance measure,  $PM_1 = E[\hat{T} - T]^2$  (hats denote estimated quantities), where  $E$  is the expected value with respect to the distribution of  $\hat{T}$ . A more common performance measure in multivariate calibration is  $PM_2 = E[\sum_i (x_i - \hat{x}_i)^2]$ . Note that  $PM_1 = E[(\sum_i x_i - \hat{x}_i)^2]$  [mean-squared error (MSE)], so that covariances among the  $x_i$  can potentially degrade or improve performance, depending on their sign. It is well known that there is guaranteed to be a biased solution vector  $\hat{\mathbf{x}}$  that has lower (better)  $PM_2$  than does the OLS solution vector. We suspect that there is a similar result for  $PM_1$ , but we are unaware of it.

Space does not permit us to review all the estimation methods we have implemented and tested on scaled-down versions of Eq. (2) (using  $M = 8$  to 100 and  $N = 6$  to 50); however, we will group them in five categories:

- (1) methods such as OLS or WLS that do not take explicit account of the error structure and are concerned only with minimizing the sum of squared residuals subject to  $\sum R_i = 0$  (OLS and, more generally, WLS give the minimum variance unbiased residuals [6]);
- (2) methods such as the generalized linear model (GLM), which do take explicit account of the Poisson error structure. Our GLM implementation assumes that  $\mu_{b_i}$  is known and equal to  $b_i$ —which [1] calls the

MLEM-FB (fixed background) method—for all of our methods except MLEM. MLEM jointly estimates  $\mu_b$  and  $\mu_g$  by maximizing their joint likelihood,

- (3) methods (Bayesian) that make prior assumptions about both the magnitudes of the  $\mathbf{x}$  entries and their spatial correlation. For example, Green [4] let the prior probability for  $\mathbf{x}$  be

$$p(\mathbf{x}|\beta, \delta) = 1/C \exp(-\beta \sum_{i-j} w_{i,j} \log \{ \cosh[(x_i - x_j)/\delta] \} ),$$

where  $i-j$  denotes summing over neighbors (adjacent voxels),  $\beta$  and  $\delta$  are known constants, and  $C$  is a normalization constant. The posterior probability for  $\mathbf{x}$  is proportional to the prior times the likelihood (Eq. 2). Maximizing the posterior probability for  $\mathbf{x}$  is equivalent to maximizing a penalized likelihood, where the penalty involves the roughness of the solution  $\mathbf{x}$  (the extent of spatial heterogeneity). For more detail, see [3–5]. In particular, an approach ([5]) that uses a hyperprior for  $\beta$  and  $\delta$  and then estimates  $\beta$  and  $\delta$  (empirical Bayes) [5] using Markov Chain Monte Carlo (MCMC) is of possible interest for two reasons: (1) the assay performance could be improved and (2) by using MCMC, it is possible to construct observations from the posterior for  $\mathbf{x}$ , which allows us to construct an estimate  $\hat{\mathbf{x}}$  of  $\mathbf{x}$  and to estimate the variance of  $\hat{\mathbf{x}}$ . Here we present results for the one-step late (OSL) method for maximizing the (penalized) likelihood from [4];

- (4) methods that make some prior assumptions about the magnitudes of the  $x_i$ . For example, ridge regression (RR) ([6]) implicitly makes such assumptions and thereby has a Bayesian justification; and  
 (5) methods that consider errors in the  $A$  matrix (error in variables (EIV) methods [7]).

### Simulation Study

Here we report simulated assay results of several classical methods and the OSL Bayesian method for a  $2^6$  full factorial experimental design varying the following six factors from low ( $L$ ) to high ( $H$ ) with  $N = 6$ :

- $\text{cond}(A)$ :  $L = 17, H = 1366$
- $M$ :  $L = 16, H = 8$ ,
- $\sigma_A$ :  $L = 0, H = 0.2A$ ,
- $T$ :  $L = 1, H = 10$ ,
- $\sigma_x^2$ :  $L = 0, H = 100\%$  of  $\mu_x$ . ( $\sigma_{x_i}^2$  is the variance of the  $x_i$  and  $\mu_x$  is the mean of the  $x_i$ ), and
- noise-to-signal ratio (NSR)  $c\mu_{b_i} / \sum_j A_{ij}x_j$ :  $L = .2, H = 1$ .

For each factor, we defined the  $L, H$  values to be such that better results are expected at the  $L$  value if all other factors are held fixed. However, because of possible interactions among factors (the effect of one factor might depend on the level of one or more other factors), it is not always possible to anticipate the average effect of changing a given factor from  $L$  to  $H$ . Nevertheless, because we use a balanced design, the main effects of each factor are reasonably easy to interpret ([8]). We believe the  $L$  and  $H$  values of each factor span a reasonable range for most situations (except that  $M = 8$  or  $16$  is far smaller than  $M = 2400$ ). For example, in two sets of well-controlled replicate scans of two drums (bottom 8 of 16 layers with the result that  $A$  was 1200 rows by 909 columns), the relative standard deviation ranged from nearly 0% to over 100%, with an average of 6% for the medium-heterogeneity drum A and from nearly 0% to over 200%, with an average of 22% for the large-heterogeneity drum B. We used  $\sigma_A = 0$  ( $L$ ) and  $\sigma_A = 20\%$  of  $A$  ( $H$ ). The condition of  $A$  was approximately 20,000 for drum A for each replicate and approximately 30,000 for drum B for each replicate. We used  $\text{cond}(A) = 17$  ( $L$ ) and  $\text{cond}(A) = 1366$  ( $H$ ), so our condition numbers are somewhat low. However, another possible definition of condition might include  $\Sigma(A^T A)^{-1}$ , which was 11.5 and 40, respectively, for the  $A$ 's of the two drums (and was approximately 5 and 38,000 for our simulated  $A$ 's). It is not yet known what measure of the condition of  $A$  is most relevant here. Next, the low value for  $\sigma_x^2$  (0%) reflects the most uniform matrix possible, while the high value (100%) reflects a much more heterogeneous matrix. Finally, the design values for NSR are also reasonable, and it has been observed that high NSR does degrade performance.

### Simulation Results

In this section, we give results for simulated data for a replicated  $2^6$  full factorial experiment. Because  $\Sigma(A^T A)^{-1}$  is approximately 38,000 for  $\text{cond}(A) = H$ , all the matrix-inversion-based methods (OLS, WLS, RR, EIV, for example) performed badly for  $\text{cond}(A) = H$  (as expected from classical theory). However, MLEM did remarkably well with only occasional large  $PM_1$  values. The  $\text{cond}(A) = L$  cases had  $\Sigma(A^T A)^{-1}$  approximately 5, so they are expected to do far better than the  $\text{cond}(A) = H$  cases.

We implemented and tested 32 methods on the fully replicated  $2^6$  factorial experiment. Most methods had two versions: one version enforced the nonnegativity constraint, and the other did not. All methods could be grouped in one of the five groups discussed earlier.

Green's OSL Bayes method had to be modified slightly to accommodate the  $c\mu_b$  term. Our modification is comparable to the fixed background assumption in MLEM-FB. We have also used a Bayesian analysis as outlined in [9] and the GLM modeling function in S-Plus. Generally, the Bayesian analysis we implemented, based on results in [9] (for the distribution of net counts  $n$  assuming a nearly flat gamma prior for  $\mu_g$  and  $\mu_b$ ), is slower to implement and has not performed much better than WLS. In addition, we have had poor results with GLM in S-Plus and with the EIV methods, so we will not report those results here. We have observed an extreme sensitivity to  $\sigma_A$  and  $\text{cond}(A)$  even with the EIV method of [7]. This method has a matrix inversion step that uses a modified version of  $A$ , but, as we noted above, no matrix-inversion-based method is performing well. RR was included here because it is one of the simpler biased regression methods to implement and because [6] reported in an extensive simulation study that RR outperformed principal components regression and partial least squares. However, the presence of errors in the  $A$  matrix was not fully studied in [6], and generally it is not well known how the various biased regression methods perform in the presence of errors in a poorly conditioned  $A$  matrix.

We first screened the 32 methods by considering the median of  $PM_1$  over all 64 runs for each method. We report median values rather than mean values because some methods had outliers that tend to make the median more representative of the typical value than of the mean. We then selected 6 from among the better methods to consider in more detail: WLS, RR, MLEM, MLEM-FB, OSLa, and OSLb. OSLa used  $\beta = 0.2$ ,  $\delta = 0.02$ , and OSLb used  $\beta = 0.02$ ,  $\delta = 0.2$  in Green's [4] prior for  $x$ . We include both OSLa and OSLb here so that we can examine the sensitivity to  $\beta$  and  $\delta$ .

In Fig. 1, we plot the median of  $PM_1 = \text{MSE}$  for the six selected methods for the  $L$  and  $H$  value of each of the six design factors. Recall that we expect the MSE to be lower at the  $L$  value for each factor. However, notice that there is some interaction between  $\sigma_A$  and  $\text{cond}(A)$  for WLS and RR, because the main effect of  $\sigma_A$  (labeled A.var in figures) is opposite of that expected. Also notice that WLS and RR do far worse than the other four methods (note that the figure scales are similar for WLS and RR, which are both very different from the scales for the other four methods). The other surprise is that methods in Figs. 1C–1F do better at the  $H$  value of  $\text{cond}(A)$ . The statistical significance of any effect can be gauged by informally assuming that the smallest observed effect is caused by random variation (noise) only. For example, in Fig. 1C, the effects of A.var and var.x could be assumed to provide estimates of the random variance; therefore, it appears that the  $\text{cond}(A)$  effect is real.

In Fig. 2, we plot the median of the bias for the selected six methods for the  $L$  and  $H$  value of each of the six design factors. The dotted lines are approximate 95% confidence intervals (CI) for the true bias that would be observed if we did essentially an arbitrarily large number of replicates. These CIs are at  $\pm$  twice  $\sigma_{\text{Bias}}$  based on grouping the bias from all 64 runs. Therefore, they are wider than CIs based on using within-group biases, where the groups are defined by subsets of the factor levels. The interpretation is that if a factor is judged to be significant using these CIs, then we can safely assume that the factor truly is significant. In Figs. 2A and 2B (WLS and RR), we see evidence of statistically significant bias for  $\sigma_A^2 = 0.04$  (this is the  $H$  value of  $\sigma_A^2$ , denoted A.var in the figures). For  $\sigma_A^2 > 0$ , WLS is known to be biased, and the direction of bias is usually low. That is, regression parameters, which are denoted  $x$  here, are biased low toward 0. For  $\sigma_A^2 = 0$ , RR is known to be biased, but one could expect RR to also be biased for  $\sigma_A^2 > 0$ . We see in Fig. 2b that RR is biased in our setting for  $\sigma_A^2 > 0$ , but there is no evidence of bias for  $\sigma_A^2 = 0$ . That means that we were conservative in our choice of ridge parameter. Our method of choosing the ridge parameter was to begin with the value suggested in [6], then use the ridge trace to look for a good ridge value near the starting value. We believe it is difficult to justify increasing the ridge parameter much more than what is suggested in [6]. In Figs. 2C–3F, we see that MLEM, MLEM-FB, OSLa, and OSLb all tend to exhibit positive biases. Observation of this positive bias is part of the motive for undertaking this study. However, note in Fig. 2C that MLEM can have negative bias ( $\Sigma x_i = 1$  and  $\text{NSR} = L$  cases).

Theory [10] for maximum likelihood estimators suggests that the bias of MLEM or MLEM-FB should approach 0 as  $M$  increases for fixed  $N$ . We tested this by comparing the run 1 and run 2 cases (run 1 has all  $L$  values for the six factors, and run 2 has all  $L$  values, except  $\text{cond}(A)$  is  $H$ ) from our designed cases with a case (run.new) having  $M = 128$ . This  $M = 128$  case used both the  $\text{cond}(A) = L$  and  $\text{cond}(A) = H$  matrices ( $M = 16$ ) and duplicated these 16 rows 8 times to give  $M = 128$ . MLEM and MLEM-FB results from 100 simulations of the three runs (run 1, run 2, and run.new) are in Table 1.

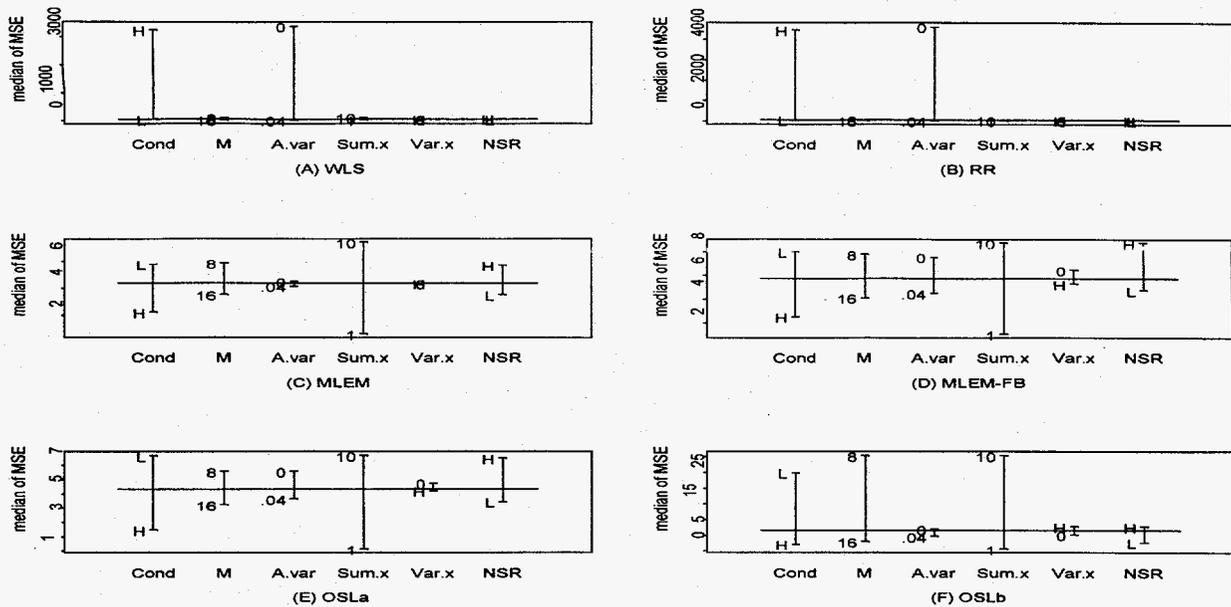


Fig. 1. Median of  $PM_1 = MSE$  for six methods for the  $L$  and  $H$  values of each of the six design factors.

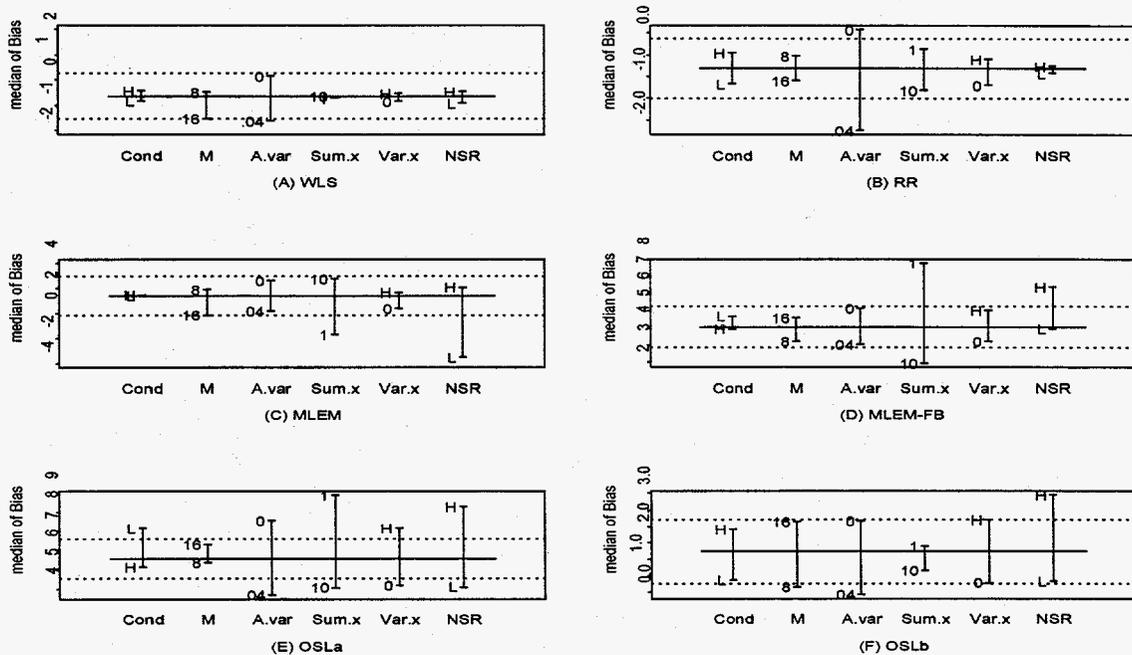


Fig. 2. Median of the bias for six methods for the  $L$  and  $H$  value of each of the six design factors. The dotted lines are approximate 95% confidence intervals for the true bias that would be observed if we had used an arbitrarily large number of replicates.

Table 1. Bias, Variance, and MSE for MLEM and MLEM-FB for the First 2 of 64 Runs and for a New Run Having  $M = 128$  rows

Run	Bias		Variance		MSE	
	MLEM	MLEM-FB	MLEM	MLEM-FB	MLEM	MLEM-FB
1	-0.29	0.26	0.43	0.35	0.511	0.4
2	-0.32	0.30	0.56	0.52	0.66	0.60
new	-0.95	0.16	0.003	0.04	0.9	0.06

When the new run is compared with runs 1 and 2 for MLEM, the bias does not approach 0, but the variance does. Surprisingly, the bias actually increases dramatically, which leads to an increase in MSE (with  $M = 128$  versus  $M = 8$ ). For MLEM-FB, the situation is better. Both the bias and variance appear to be approaching 0. Chapter 18 of [10] provides some ways to correct for bias in maximum-likelihood methods. Unfortunately, they are probably too computationally intensive for the large  $M$  and  $N$  values in real TGS systems. Reference [11] considers the "constrained to be nonnegative" aspect of MLEM and derives approximate results for the expected precision (variance in our terminology here) of the MLEM estimates. The propagated variance agreed to within approximately 7% of the observed variance in the simulation study reported in [11]. Currently, we are investigating whether the size of  $M$  and  $N$  must be considered in these variance approximations.

## Summary

In this paper, we examined three possible bias sources: (1) measurement error in the estimated transmission matrix, (2) the positivity constraint on the estimated mass of nuclear material, and (3) improper treatment of the measurement error structure (for example, ignoring the Poisson distributions). We studied these bias sources for 32 methods (presented results for six of the best methods) of estimating SNM mass in the context of a fully replicated  $2^6$  factorial experiment with 100 simulations at each run. We used a 20% relative standard deviation (all random error) in our simulated  $A$  matrices and found that WLS and RR were sensitive to  $\sigma_A$ , but that MLEM, MLEM-FG, OSLa, and OSLb were not very sensitive to  $\sigma_A$ . We did not formally investigate factor interactions; however, we see evidence that  $\text{cond}(A)$  and  $\sigma_A$  interact in all the matrix-inversion-based methods. We did observe that the positivity constraint on WLS or RR greatly inflated the MSE, so reported results were for unconstrained WLS and RR. The positivity constraint is more natural in MLEM and similar methods. All of our biased regression methods are examples of "improper" treatment of the measurement error structure; however, experience has shown that variance can be reduced at the expense of introducing bias if we alter the  $A$  matrix. Biased regression methods often outperform OLS or WLS; therefore, the net effect on MSE is often good. Future work will consider whether bias reduction methods for maximum likelihood methods are worth the computational burden and whether estimates of the variance of maximum likelihood estimators should depend on the size ( $M$  and  $N$ ) of  $A$ . We will also study the merits of Bayesian methods that can generate observations from the posterior distribution of  $\mathbf{x}$  so that we can provide estimates of both the mean and variance of the estimated total mass  $\hat{T}$ .

## References

- [1] T. Prettyman, R. Cole, R. Estep, and G. Sheppard, "A Maximum-Likelihood Reconstruction Algorithm for Tomographic Gamma-ray Nondestructive Assay," *Nuclear Instruments and Methods in Physics Research A* **356**, 470–475 (1995).
- [2] T. Prettyman and D. Mercer, "Performance of Analytical Methods for Tomographic Gamma Scanning," Los Alamos National Laboratory document LA-UR-97-1168 (1997).
- [3] T. Burr, D. Mercer, and T. Prettyman, "Comparison of Bayesian and Classical Reconstructions of Tomographic Gamma Scanning for Assay of Nuclear Materials," 1998 Spring Research Conference on Statistics in Industry and Technology, St. John's College, Santa Fe, New Mexico, June 3–5, 1998 (to appear in proceedings issue of the American Statistical Association).
- [4] P. Green, "On the Use of the EM Algorithm for Penalized Likelihood Estimation," *Journal of the Royal Statistical Society B* **52**, No. 3, 443–452 (1990).
- [5] I. Weir, "Fully Bayesian Reconstructions from Single-Photon Emission Computed Tomography Data," *Journal of the American Statistical Association* **92**, No. 437, 49–60 (1997).
- [6] I. Frank and J. Friedman, "A Statistical View of Some Chemometrics Regression Tools," *Technometrics* **35**, No. 2, 109–148 (1997).
- [7] W. Fuller, *Measurement Error Models* (Wiley 1987).
- [8] G. Box, W. Hunter, and J. Hunter, *Statistics for Experimenters* (Wiley 1978).
- [9] D. Mercer, T. Prettyman, M. Abhold, and S. Betts, "Experimental Validation of Tomographic Gamma Scanning for Small Quantities of Special Nuclear Material," in *Proceedings of the 38<sup>th</sup> Annual Meeting of the Institute of Nuclear Materials Management*, Phoenix, Arizona, July 20–24, 1997 (proceedings on CD-ROM).
- [10] M. Kendall and A. Stuart, *The Advanced Theory of Statistics*, Volume 2 (Griffin and Company 1979).
- [11] T. Prettyman, "Precision Estimates for Tomographic Nondestructive Assay," in *American Nuclear Society Fifth International Conference on Facility Operations—Safeguards Interface*, Jackson Hole, Wyoming, September 24–29, 1995; Los Alamos National Laboratory document LA-UR-95-3299 (1995).