MAS REPORT 95-1

TRANSACTIONS

MATHEMATICS, COMPUTER SCIENCE, STATISTICS DIVISION

MISSISSIPPI ACADEMY OF SCIENCES

FIFTY-NINTH ANNUAL MEETING

February 9 & 10, 1996

BILOXI, MISSISSIPPI

MEASURES OF EFFECTIVENESS FOR MONTE CARLO SENSITIVITY ANALYSES

Andrew W. Harrell
U.S. Army Engineer Waterways Experiment Station
Vicksburg, MS 39180-0631

ACKNOWLEDGMENTS. The tests described and the resulting data presented herein, unless otherwise noted were obtained from research conducted under the MILITARY RESEARCH DEVELOPMENT TEST AND EVALUATION PROGRAM of the United States Army Corps of Engineers by the U.S. Army Engineer Waterways Experiment Station. Permission was granted by the Chief of Engineers to publish this information.

I wish to thank Mr. Newell R. Murphy, Jr., and Dr. Niki C. Deliman for their valuable comments and suggestions which have improved the exposition of this paper.

ABSTRACT. The Army Mobility model (AMM) developed, at the U.S. Army Engineer Waterways Experiment Station, uses the data from about a hundred factors that describe a vehicle terrain unit, road unit, or linear feature to predict vehicular speeds. Recently, Monte Carlo simulations were conducted for several wheeled and tracked vehicles and different areas, varying some selected groups of these factors plus and minus 10 percent about their nominal values. The results of these simulations have been studied to develop empirical relationships that allow the expression of confidence measures for the speed predictions on an entire mobility map. As a first step, programs have been written to test methods to estimate the value of continuous statistical parameters (the mode and its standard deviation) of a discrete histogram. This allows theorems of mathematical statistics to be applied to the confidence levels around the values of the parameters. The method uses a variation I made on E. Parzen's formula for the location of the mode of the continuous distribution associated with a discrete histogram.[1] The formula works by estimating the rate of an associated statistical process by discrete windows (Jth waiting times). The incomplete gamma function and a maximum liklihood product is then used to estimate the parameters. [2] This approach has been tested for a range of Monte Carlo generated discrete approximations to gamma distributions. It was then applied to the histograms of possible errors in speed predictions of tactical vehicles moving across areas on different mapsheets. These histograms were generated previously in the course of the work by Lessem and Ahlvin and are discussed in reference [6].

---

[1] See Parzen, Emanuel, "Stochastic Proceses", Holden Day, 1962, and Press,W., Flannery, B. et al., "Numerical Recipes in C," 2nd ed., Cambridge U. Press, 1988.

[2] Ibid.

In trying to determine how to organize the sensitivity trials in this particular set of programs and data there are several approachs that can be taken. Because the speed prediction program uses a series of lookup tables and flow chart, "yes" or "no", go and no-go cutoff rules, points at which the program computes a no-go output are natural areas to investigate its sensitivity to errors in the data. Error measures can be associated with "critical regions" in the data around these points. Determining the modes and moments in the discrete non-parameteric histograms generated by the sensitivity trials gives a way of characterizing and reproducing the confidence in information contained in the program's output involving these regions. One approach, which measures the program's "inherent sensivity" to errors, is terrain-independent and vehicle dependent. It examines the code in the program to find the 1-factor critical regions in the outputs of the Monte Carlo trials. It then adjusts the values of the other factors in a detrimental direction of the lookup table values until the 2,3 and higher multi-factor critical regions are identified. Another approach is "project specific" and is both terrain dependent and vehicle dependent. It looks at the areas on the speed prediction maps where no-gos occur. It then goes back to the input files to determine the values of the data at the terrain units where these no-gos occur. This is the approach that will be taken in this paper.

After the procedure for conducting the trials is determined it is important to consider ways to examine confidence levels for the parameters that are estimated. One approach to this, which recently has gained popularity, is the technique of bootstrapping. This technique conducts Monte Carlo trials of the Monte Carlo trials. The algorithm resamples not from the original data, but from a smoothed kernel estimate of the data (see MathCad [8] for the details of the algorithm and Efron, Hall and Tittleman, and Scott for the theory behind formulas for the variance of the sampled estimate of the parameter). Smoothed kernel formulas, introduced by Parzen and others (see Scott [12], Parzen [9]) allow better resolution of modes and other information in the data using a given histogram bin size or window. In order to estimate the second moment or the variance of the kernel estimate, it is necessary to write programs to compute the second derivative of the frequency polygon of the histogram (see Scott [12]). Bootstrapping confidence intervals can then also be computed from this information.

In this paper a somewhat simplified approach is taken. A leave-one-out maximum liklihood product of smoothed kernels over different possible bin widths is taken. The product is taken over a choice of possible bin widths. Once the best bin width is determined the variance of the kernel associated with this bin width is computed (see Numerical Recipes in C, 2nd ed. [10]) This aggregrates the data in a one dimensional histogram and does not give you as much information as in the more complicated multi-dimensional approach.

Figures 1, 2, 3, and 4 show the results of a series of Monte Carlo error sensitivity trials run on some vehicle speed

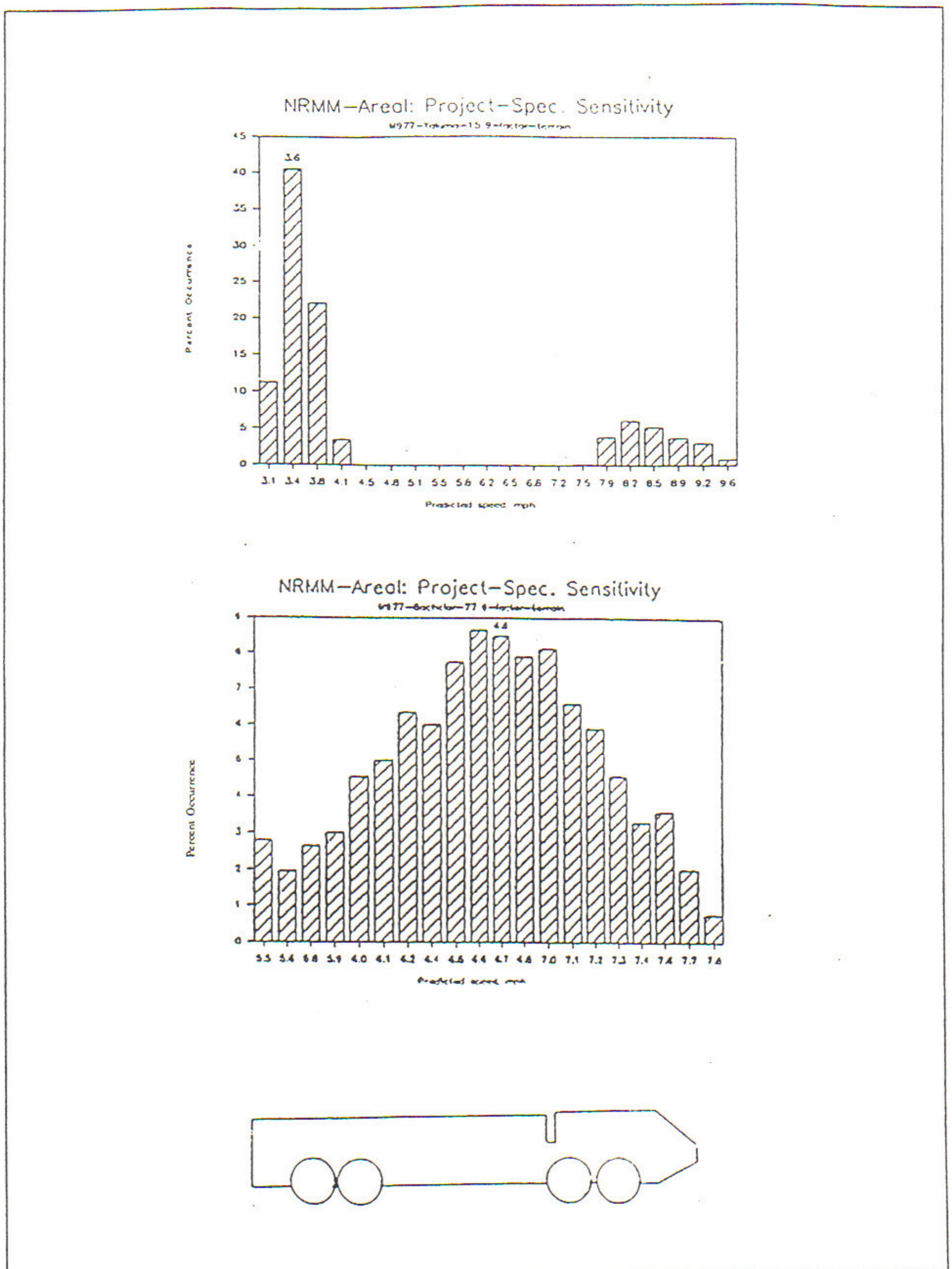

Figure 1 Monte Carlo Sensitivity Runs, HEMTT

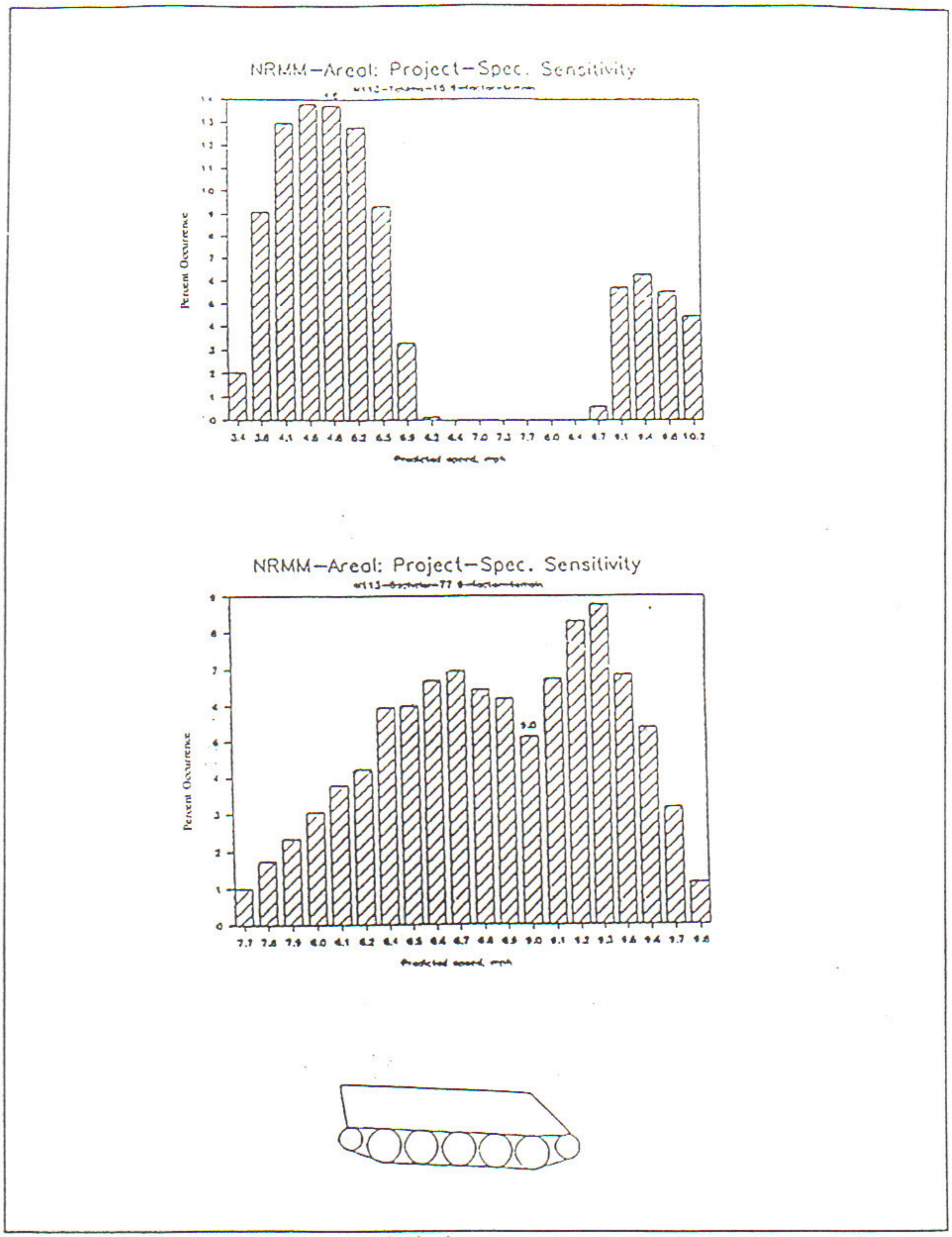

Figure 2 Monte Carlo Sensitivity Runs, M113

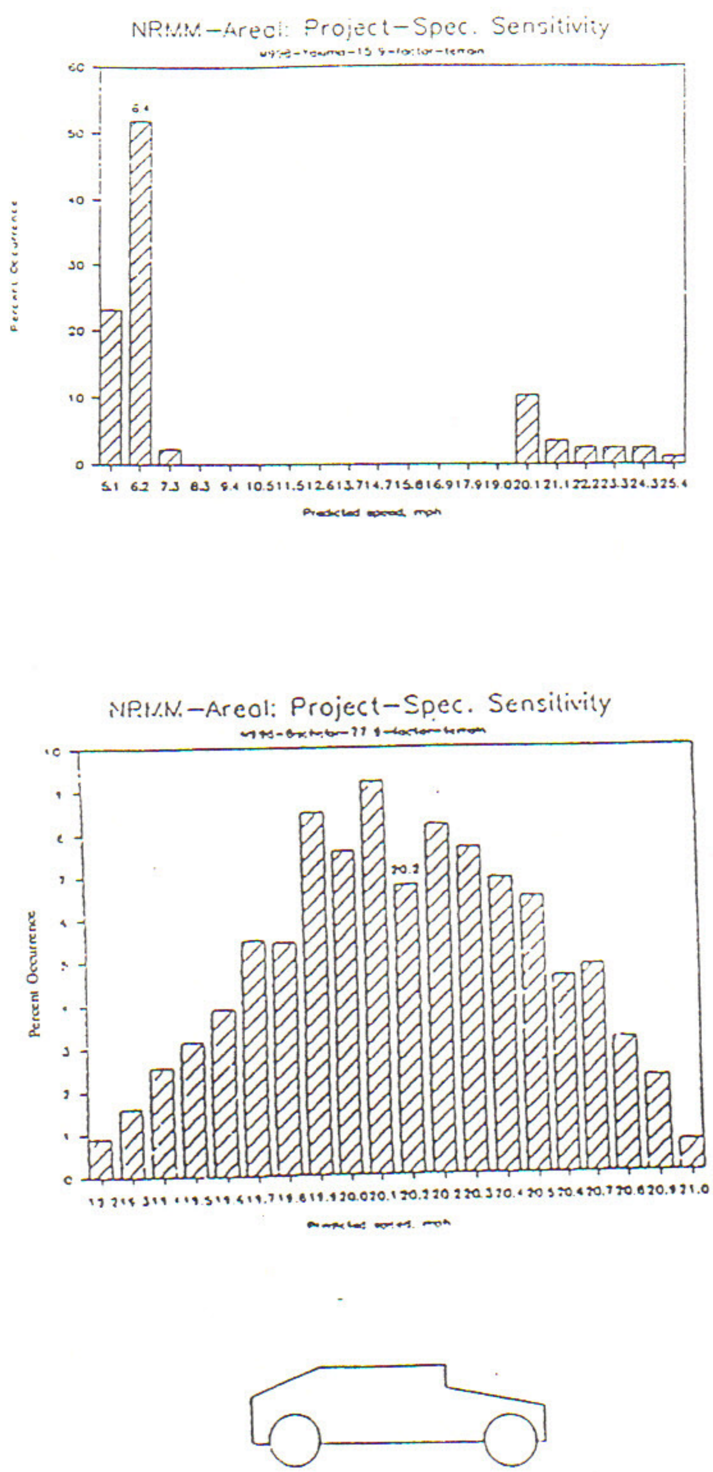

Figure 3 Monte Carlo Sensitivity Trials, HMMV

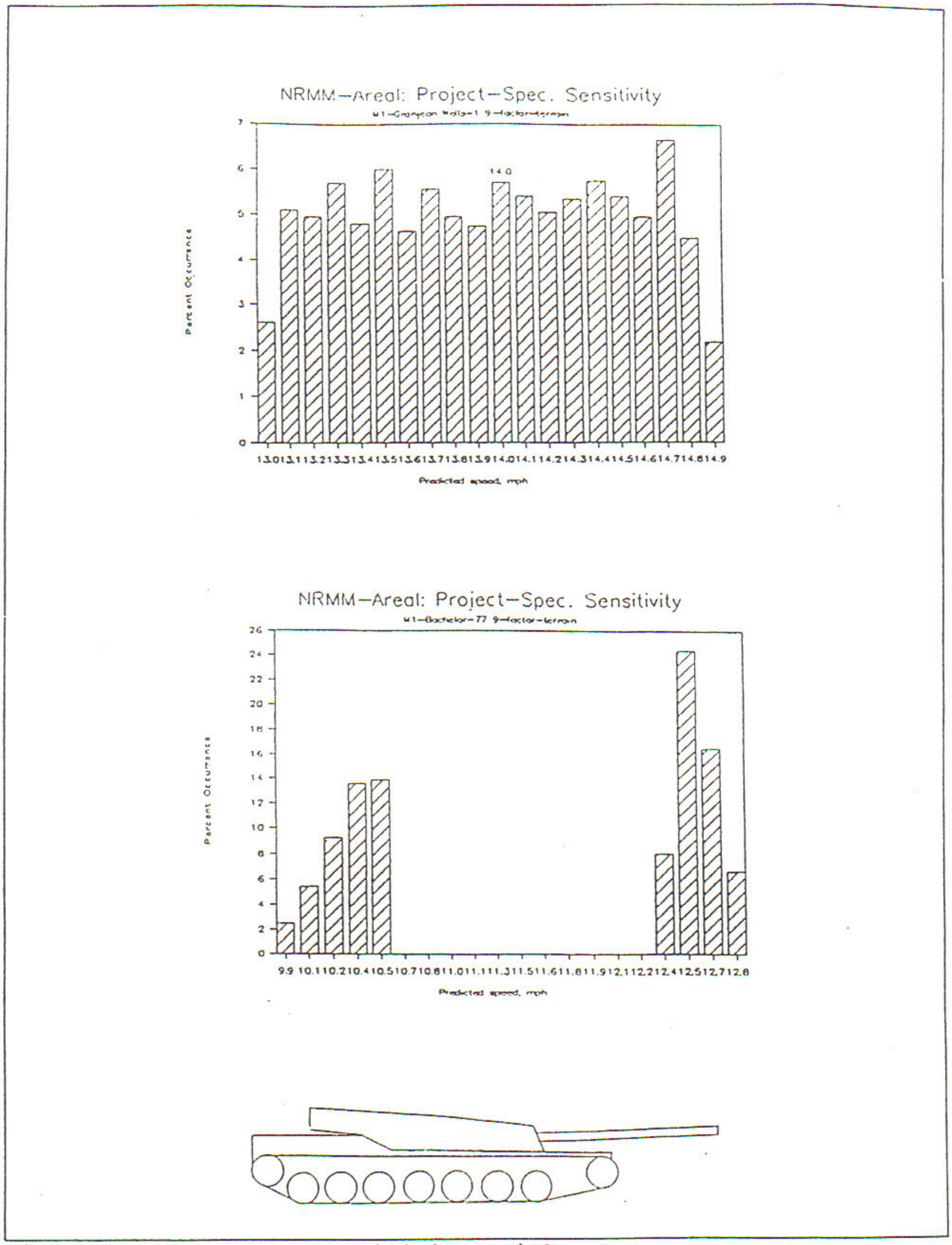

Figure 4 Monte Carlo Sensitivity Trials, M-1

predictions by Lessem et al. [6]. They display the speeds predicted for the M998 High Mobility Multi-Purpose Vehicle (HMMV), the M977 10-Ton Heavy Expanded Mobility Tactical Truck (HEMTT), M113 Armoured Personnel Carrier, and the M-1 tank. The terrain areas tested are in Yakima, Washington, Granjean Wells, New Mexico and Bachelor, Australia. The graphs have predicted speeds plotted on the horizontal axis. The speeds were computed by varying nine factors: soil strength, slope, surface roughness, visibility, vegetation type, and four other attributes dealing with obstacle characteristics around their nominal values in a certain terrain unit. The nominal values for that terrain unit were chosen as the points around which the vehicle's performance on the mapsheet terrain units changed most noticeably. The points were determined by referring both to the output that the program computed and to the tables in the speed computation program where the performance changed significantly. On the vertical axis is a count of the number of occurances of a given speed for that vehicle, that terrain unit, and for the range of Monte Carlo trials used. Both uniform and normal density functions were used to compute the randum numbers used in the Monte Carlo sensitivity trials. Thus the graph displays the areal sensitivity of the speed predictions for that vehicle in that area. Notice that the results don't appear to have a common probability density function. The WES technical reports by Lessem et. al. [6] and [7] contain a more detailed discussion of the features of the mobility programs which cause the histograms to assume these shapes.

In general, these histograms will separate into several parts each with distinct characteristics. In this particular case parts of the graphs associated to each single mode were separated out. Let us assume this has already been done. We arrange the results of the Monte Carlo simulations in a histogram of N bins with the number of Monte Carlo hits (test items) in the ith bin equal to $hist_i$. In order to estimate the number of Monte Carlo trials necessary to reproduce the probability density function from which these results give samples we have to use an unstructured or nonparametric approach.[3] Let us define

$$(1.1) \quad p(t+1/2*J) \equiv \frac{\sum_{i=t}^{t+J} hist_i}{N*J}$$

where t = *bin number around which estimate is centered*
J = *integer* ≥ 1
N = *total number of observations*

[3] Keinosuke Fukunaga, "Introduction to Statistical Pattern recognition," Academic Press, 2nd ed., 1991.

According to the reference by Fukunaga [3], this formula gives the Parzen density estimate for the value of this probability density function at the point k = t+ J/2.[4] In this formula we are using a local region defined by a window of size J around the point to estimate the number of hits in a counting process in terms of the histogram values located in this region. This formula gives estimates for the values of the density function at N-J points. Sorting these estimates and picking out the middle and highest values then gives the best prediction of the mode and the mean of the histogram using windows of size J. On page 261 of this reference the value of the standard deviation of this estimate is calculated to be:

$$(1.2) \qquad \sigma(t+1/2*J) \equiv \frac{\sum_{i=t}^{t+J} hist_i}{J*\sqrt{J}*N} \quad .$$

Note that the value of this standard deviation refers to an interval around a point on the x-axis of the histogram and not around the height of the histogram or number of Monte Carlo values in that bin.

These formulas and theorems allow a leave-one-out procedure along with a maximum liklihood product to be used to estimate thevalue of the window size which gives the smallest error in estimating the parameters.[5]

Using our procedure for computing estimates of the value of the probability distribution, at the point k defined in equation 1.1 the function p(k) is proportional to the amount the cumulative distribution function changes in this interval... so, the larger it is, the better is the chance for a local maximum of the probability distribution function at that point. The program computes estimates of the continuous modes for different window sizes, where J = window size, x_j = bin# of largest of these estimates, p(k) = weighted estimate of mode at this bin = (sum of # of distribution hits in the bins inside a window of width J centered at k)/(total # shots)* J. In the case where the

---

[4] Actually, this is the density function of a "renewal counting process" as defined in Parzen [9].

[5] See besides the Numerical Recipes in C reference also the Introduction to Statistical Pattern Recognition text referred to above. These same procedures can be used to characterize the histogram distribution of pixel intensities in digital images. Such a characterization allows the use of various neural network learning procedures to be used to identify the images.

distribution function is suspected to be bimodal, this procedure will identify at least the top two modes when it is iterated over different window sizes.

Let $\delta(J) \equiv$ the range of data values around the candidate for mode calculated using a window size of J.

$$\text{Thus, } \delta(J) = \sum_{i=x_J-\frac{J}{2}}^{i=x_J+\frac{J}{2}} hist_i \quad .$$

Then, in this notation, the probability distribution $p_J(k)$ of the smoothed estimate of the original data is given by:[6]

$$p_J(k) \approx \frac{\delta(J)}{N\,J} \quad .$$

Let

$$(2.1) \qquad \delta_J(k) = \sum_{i=k-\frac{J}{2}}^{i=k+\frac{J}{2}} hist_i \quad .$$

Let H(J)= the hypothesis that the true mode x_J has been identified by considering a window of size J. We want to consider how likely it is that the range around x_J should be shorter than it is observed to be. Let P(a,x) be the incomplete gamma function:

$$P(a,x) \equiv \frac{1}{\Gamma(a)}\int_0^x e^{-t}t^{a-1}dt$$

where:

---

[6] See the discussion in Numerical Recipes in C edition 1 and also the book by Parzen, pages 133-134.

$$\Gamma(a) \equiv \int_0^\infty e^{-t} t^{a-1} dt \quad .$$

Thus, P(a,x) is the cumulative Poisson probability distribution function, Prob(X <= a) for the Poisson probability distribution X. It is defined as the probability that the number of Poisson random events occuring will be between 0 and a - 1. Each of these random events will have a probability of occurance of $N*p_J$.

The probability that the range around x_J is actually shorter than observed to be if H(n) is true instead of H(J) is[7]:

$$\int_0^{\delta_J(J)} (N\, p_J(n)) \frac{(N\, p_J(n)\ \ t)^{n-1}}{(n-1)!} e^{-N\, p_J(n)\ t} dt \quad .$$

If we let:

$$y = N\, p_J(n)\ \ t$$
$$a = n$$
$$x = \delta_J(J)$$

in the above equation ,

then it is equal to:

$$P(n, \frac{\delta_J(J)}{N\, p_J(n)})$$

which is the same as:

$$P(n, J\ \frac{\delta_J(J)}{\delta_J(n)}) \quad .$$

Taking the product of all these factors for each mode x_n then gives the likelihood that the range around x_n should be shorter than the range observed around x_J for all n other than J.
Thus the likelihood function is defined by Likelihood(H(J)):

---

[7] Parzen, Ibid pp. 133-134.

$$(2.2) \qquad L(n|J) \equiv \Pi_{n*J} P(n, J*\delta_J(J)/\delta_J(n)) \quad .$$

The program then computes the value of this window size J that maximizes the likelihood function, given a set of arrayed a posteriori error sizes.

More precisely, the steps in the computation are:

1) Compute the

$$\delta_J(n)$$

according to equation (2.1) for the points corresponding to each bin.

2) Compute the maximum liklihood products according to equation (2.2) in order to determine the optimal window size.

3) Compute the weighted sums p(k) according to equation (1.1) and the standard deviations according to equation (1.2) for the points corresponding to to each bin.

Because of the nonparametric form of the Parzen density estimate, the procedures will work for any empirically determined histogram. A discrete sorting procedure normally gives a pretty good estimate of the value of the mean and mode (even assuming the actual distribution is continuous). However, in order to approximate the size of the standard deviation in the estimation of the parameters, it is necessary to use the maximum likelihood estimators. These estimators of the best window sizes will result in good approximations of the parameters.

An example of how these parameter estimates work is shown as it is applied to the results of Monte Carlo sensitivity runs in Figures 1,2,3,4. The simulations shown in the figures were conducted for four vehicles the HMMV, the M997 trailer transporter, the M113 APC, and the M-1 tank. The top charts show the results for a mapsheet Yakima Proving Grounds and the bottom charts those for a mapsheet including Batchelor Australia.These figures show the results of varying the parameter values plus and minus 10 per cent around their nominal values. Nominal values are defined as the vehicle parameters plus the specific parameter values in each terrain unit. For, this analysis, we considered the particular values for which that vehicle experiences a go, no-go situation,as the values around which variations were made.

Data from the M997, M113, and M998 runs were extracted directly from the top row of histograms in Figures 1,2, and 3 respectively. Programs were written to expand the information into a 20 bin histogram and to scale the data. This turned out to be a good range for the incomplete gamma function to discriminate the maximum likelihood estimates. The results of the program runs are shown below. First the program calculates a value for the mode by simply sorting the columns of the histogram. This is called a discrete estimate. The abscissa of this point is called modei. Then the program computes the optimal window size for smoothing the data using the leave-one-out maximum likelihood procedure explained above and determines a continuous estimate for the mode along with a standard deviation. Both of these numbers are computed using this optimal window size.

The results are shown below:

```
histogram of Monte Carlo error runs
M998 Yakima-15 9-factor-terrain ( mode#1 )

     x      p(x)    graph:
  4.0000   0.1820   *********
  4.5500   0.3275   ****************
  5.1000   0.4731   ***********************
  5.6500   0.6186   ******************************
  6.2000   0.3311   ****************
  6.7500   0.0437   **
  7.2500   0.0218   *
  7.7500   0.0000
  8.3000   0.0000

Data drawn from a histogram of Monte Carlo sensitivity
to errors in terrain factors
Discrete estimate of mode of data set is 42.500000

Discrete estimated value of modei= 5.650001

Probability of mode detected at window size 3 is  0.229365
Probability of mode detected at window size 4 is  0.253296
Probability of mode detected at window size 5 is  0.256476
```

```
Probability of mode detected at window size 6 is  0.014268
Probability of mode detected at window size 7 is  0.014372

Most likely window size is 5 value of mode is 32.50000

Continuously estimated value of modei=5.10000

Standard deviation of the continuous estimate (for this window
size) is  0.607092


histogram of Monte Carlo error runs
M998 Yakima-15 9-factor-terrain ( mode#2 )
     x     p(x)     graph:

 16.8000   0.0000
 17.3500   0.0000
 17.9000   0.0000
 18.4500   0.0000
 19.0000   0.7143  ***********************************
 19.5500  1.4286 ************************************************
 20.0500  0.9740 ************************************************
 20.5500   0.5195  *************************
 21.1000   0.4545  **********************
 21.6500   0.3896  *******************
 22.2000   0.3831  *******************
 22.7500   0.3766  ******************
 23.2500   0.3766  ******************
 23.7500   0.3766  ******************
 24.3000   0.2597  ************
 24.8500   0.1429  *******
  0.0000   0.0000

Data drawn from a histogram of Monte Carlo sensitivity
to errors in terrain factors
Discrete estimate of mode of data set is 11.0000000

Discrete estimated value of modei=19.549995

Probability of mode detected at window size 3 is  0.204653
Probability of mode detected at window size 4 is  0.039479
Probability of mode detected at window size 5 is  0.116221
Probability of mode detected at window size 6 is  0.065450
Probability of mode detected at window size 7 is  0.129556
Most likely window size is 3 value of mode is 11.0000000
Standard deviation of the continuous estimate (for this window
size) is  0.269430

Continuously estimated value of modei=11.00000

histogram of Monte Carlo error runs
M997 Yakima-15 9-factor-terrain ( mode#1 )

     x     p(x)     graph:
```

```
 2.8000  -0.0387
 2.9500   0.2062  **********
 3.1000   0.4510  **********************
 3.2500   0.6959  **********************************
 3.3833   0.5956  *****************************
 3.5167   0.4954  ************************
 3.6500   0.3952  *******************
 3.8000   0.2234  ***********
 3.9500   0.0515  **
 4.0833   0.0344  *
 4.2167   0.0172
 4.3500   0.0000
```

```
Data drawn from a histogram of Monte Carlo sensitivity
to errors in terrain factors
Discrete estimate of mode of data set is 40.500000
Discrete estimated value of modei= 3.250000

Probability of mode detected at window size 3 is  0.282627
Probability of mode detected at window size 4 is  0.064773
Probability of mode detected at window size 5 is  0.076770
Probability of mode detected at window size 6 is  0.083600
Probability of mode detected at window size 7 is  0.084213
Most likely window size is 3 value of mode is 40.500000
Standard deviation of the continuous estimate (for this window
size) is  1.253331

Continuously estimated value of modei=3.250000

histogram of Monte Carlo error runs
M977 Yakima-15 9-factor-terrain ( mode#2 )
```

```
    x      p(x)    graph:
 6.9000  -0.0032
 7.0500   0.0000
 7.2000   0.0032
 7.3500   0.0065
 7.4833   0.1775  ********
 7.6167   0.3485  *****************
 7.7500   0.5195  *************************
 7.9000   0.6494  ********************************
 8.0500   0.7792  **************************************
 8.2000   0.7143  ***********************************
 8.3500   0.6494  ********************************
 8.4833   0.6061  ******************************
 8.6167   0.5628  ****************************
 8.7500   0.5195  *************************
 8.9000   0.4545  **********************
 9.0500   0.3896  *******************
 9.1833   0.3030  ***************
 9.3167   0.2165  **********
 9.4500   0.1299  ******
 0.0000   0.0000
```

```
Data drawn from a histogram of Monte Carlo sensitivity
to errors in terrain factors

Discrete estimate of mode of data set is    6.000000

Discrete estimated value of modei= 8.049999

Probability of mode detected at window size 3 is  0.007822
Probability of mode detected at window size 4 is  0.023633
Probability of mode detected at window size 5 is  0.082934
Probability of mode detected at window size 6 is  0.183745
Probability of mode detected at window size 7 is  0.298017
Most likely window size is 7 value of mode is  5.500000
Standard deviation of the continuous estimate (for this window
size) is  0.103940

Continuously estimated value of modei=8.199999

histogram of Monte Carlo error runs
M113 Yakima-15 9-factor-terrain ( mode#1 )

     x      p(x)     graph:
  3.0000  -0.0862
  3.2000   0.1149   *****
  3.4000   0.3161   ***************
  3.6000   0.5172   *************************
  3.7500   0.6322   *******************************
  3.9000   0.7471   *************************************
  4.1000   0.7701   **************************************
  4.3000   0.7931   ***************************************
  4.4500   0.7902   ***************************************
  4.6000   0.7874   ***************************************
  4.8000   0.7644   **************************************
  5.0000   0.7414   *************************************
  5.1500   0.6322   *******************************
  5.3000   0.5230   **************************
  5.5000   0.3563   *****************
  5.7000   0.1897   *********
  5.8500   0.0977   ****
  6.0000   0.0057
  0.0000   0.0000
  0.0000   0.0000

Data drawn from a histogram of Monte Carlo sensitivity
to errors in terrain factors

Discrete estimate of mode of data set is   13.799999

Discrete estimated value of modei= 4.300000
standard deviation is  0.283068

Probability of mode detected at window size 3 is  0.003206
Probability of mode detected at window size 4 is  0.020749
Probability of mode detected at window size 5 is  0.114160
```

Probability of mode detected at window size 6 is 0.132713
Probability of mode detected at window size 7 is 0.369124
Most likely window size is 7 value of mode is 13.799999
Standard deviation of the continuous estimate (for this window size) is 0.283068

Continuously estimated value of modei=13.79999

Summary of Mode Estimates for data

Discrete estimate of mode of data set 1 is point 5.650001 at 42.500000
continuous estimate of mode of data set 1 is point 5.100000 with value 32.500000
A window of size 5 was used to estimate this

Discrete estimate of mode of data set 2 is point 19.549995 at 11.000000
continuous estimate of mode of data set 2 is point 19.549995 with value 11.000000
A window of size 3 was used to estimate this

Discrete estimate of mode of data set 3 is point 3.250000 at 40.500000
continuous estimate of mode of data set 3 is point 3.250000 with value 40.500000
A window of size 3 was used to estimate this

Discrete estimate of mode of data set 4 is point 8.049999 at 6.000000
continuous estimate of mode of data set 4 is point 8.199999 with value 5.500000
A window of size 7 was used to estimate this

Discrete estimate of mode of data set 5 is point 4.300000 at 13.799999
continuous estimate of mode of data set 5 is point 4.300000 with value 13.799999
A window of size 7 was used to estimate this

In summary, using this technique of estimation for finding modes there is in one case (data set 1) about a 10 percent increase in the accuracy of the determination of its location. This makes available a more accurate fix on the NOGO program vehicle speed values around which to do the sensitivity analyses. Also, determination of the optimal window size to use in the estimate, gives a means to non-parametrically estimate the standard deviation of the sensitivity analyses results. This then tells us how many Monte Carlo trials should be used to explore the program's senstivity to variations in the values in its internal tables and input data. For example, for the two runs concerning the M977 performance, one mode has a determination with a standard deviation of 1.253 and the other with a standard deviation of .1039. After determining this, you could then go back and run 10 times more

Monte Carlo trials around the first mode. Similarly, although it was not analyzed for this paper, the determination of a mode in the case of the M-1 tank is much less well defined. Looking at the Monte Carlo sensitivity histogram in the top part of Figure 4, it is clear that in this case the predictions will be less accurate.

REFERENCES

1. Dahlquist, G. and Bjorck, A., "Numerical Methods," Prentice Hall, 1974.

2. Efron, B., "The Annals of Statistics," Vol 7, No1 (1979).

3. Fukunaga, K., "Introduction to Statistical Pattern Recognition," Academic Press, 2nd ed., 1991.

4. Hall, P. and DiCiccio, T. J. and Romano, J. P.,"On Smoothing and the Bootstrap," Annalen Statistics 17 692-704, 1989.

5. Hall, P. and Titterington, D. M., "On Confidence Bands in Nonparametric Density Estimation and Regression," J. Multivariate Analysis 26 59-88,1988.

6. Lessem, A. S., Ahlvin, R., and Mason, G., "Stochastic Vehicle Mobility Forecasts Using the NATO Reference Mobility Model," Report 1, "Basic Concepts and Procedures," U.S. Army Engineer Waterways Experiment Station, Technical Report, GL-91-11.

7. Lessem, A.S., Ahlvin, R., Mlakar, P., and Stough, W.,"Stochastic Vehicle Mobility Forecasts Using the NATO Reference Mobility Model," Report 2, "Extension of Procedures and Application to Historic Studies," U.S. Army Engineer Waterways Experiment Station, Technical Report, GL-93-15.

8. MathCad, Inc., "Statistics I, Tests and Estimation," Cambridge, MA, 1991.

9. Parzen, E., "Stochastic Proceses," Holden Day, 1962.

10. Press, W., Flannery, B. et al., "Numerical Recipes in C," 2nd ed., Cambridge U. Press, 1988.

11. Press, W., Flannery, B. et al., "Numerical Recipes in C," 3nd ed., Cambridge U. Press, 1993.

12. Scott, D. W., "Multivariate Density Estimation: Theory, Practice, and Visualization," John Wiley, 1992.