PLOWZONE SAMPLING SIMULATION

INTRODUCTION

Because a large contiguous area of plowzone archaeological deposits was excavated at the Whitten Road Site, the artifact collection and data on the distribution of artifacts throughout the plowzone provide a data base for sampling simulation studies. Specifically, the Whitten Road site data allow for the study of two sampling questions:

1. What sampling methods yield the best results when sampling a plowzone site? and,

2. At what sampling fraction do increasingly large samples fail to yield sufficiently large gains in sample estimates' accuracy and precision to justify their higher costs?

Simulation studies can address these questions by investigating questions such as, "What if we only dug 25% of the plowzone units?", or , "What if we used a transect-based sample instead of

a quadrat-based approach?". The "What if" questions can be posed using the completely excavated site data as a background, and estimates from the simulated samples can be compared with known data to evaluate their precision and accuracy. In archaeology, most simulations study varied techniques of regional survey (Mueller 1974; Plog 1976; Judge et al. 1975; Custer 1979, 1983); however, intrasite sampling studies have been undertaken (Ammerman, et al. 1978; Jermann and Dunnell 1979).

In the Whitten Road sampling simulation, two basic sampling strategies, each a variant of a systematically stratified aligned random sampling technique (Flannery 1976) were simulated. The first sampling technique, called a "quadrat" design, used 180 contiguous 5' x 5' squares and stratified these units into 45 10' x 10' squares (Figure 88). From each 10' x 10' square, either 1, 2, or 3 5' x 5' squares were randomly selected to simulate 25%, 50%, and 75% samples. The second sampling technique, called a "transect" design, used 128 contiguous 5' x 5' squares and stratified these units into 8 columns, or transects, of 16 5' x 5' squares each. Each transect was then divided into 4 sets of 4 5' x 5' squares each and either 1, 2, or 3 squares were randomly selected to simulate 25%, 50%, and 75% samples. Thus, 2 basic sampling designs, each with three different sampling fractions were simulated to yield a total of 6 unique simulations. The sampling designs chosen were based on the types of plowzone sampling typically carried out at Middle Atlantic historic archaeological sites.

Two basic types of evaluations of the sampling simulations were utilized. The first evaluation considered the precision and accuracy of estimates of artifact frequencies derived from the varied sample simulations. Estimates of artifact frequencies were evaluated because artifact frequencies are part of calculations of mean ceramic dates, applications of South's pattern analysis, and comparisons of artifact assemblages among sites. Table 22 shows the artifact categories used in the estimation test and the parameter values based on the total excavated assemblage for both the quadrat and transect study areas. Artifact categories with both large and small parameter values were chosen to test sample estimates' accuracy and precision for common and rare artifact classes. One hundred iterations of each sampling strategy for each sampling fraction were performed and the mean and standard deviation of each of the 100 estimates were calculated for each parameter. These mean values were then compared to the known parameters to evaluate each sample's accuracy and precision. Individual sample simulation data are on file at UDCAR and are not included in this report due to their length.

Sample accuracy was evaluated using the methods described by Judge et al. (1975:106-111). First, the mean value from the sample estimates was recorded as the "observed" value and the known population parameter was considered the "expected" value. The basic chi-square statistic was then calculated by the following formula (Judge et al. 1975:106):

FIGURE 88

Simulated Sampling Design

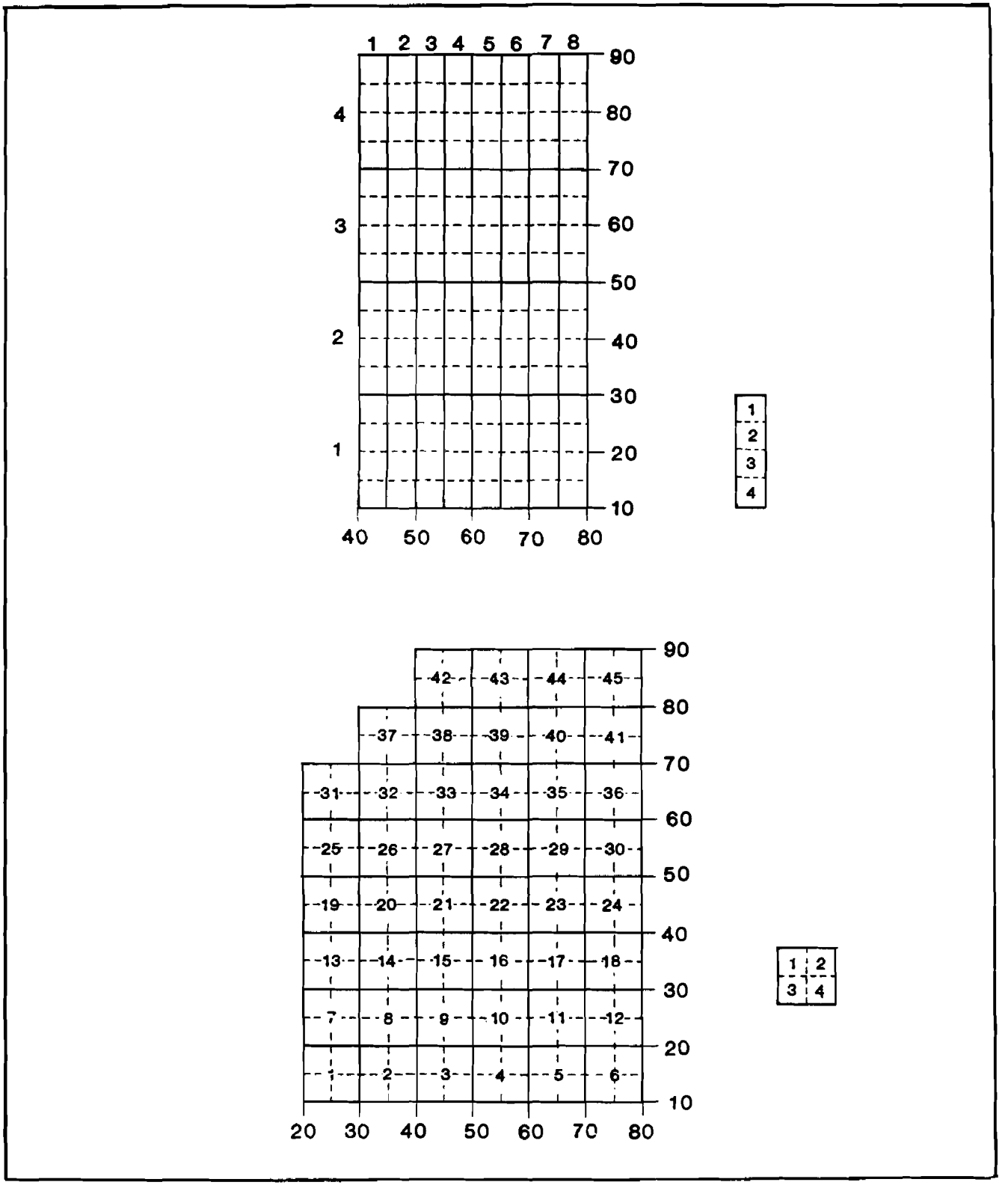


TABLE 22

PARAMETERS USED IN SAMPLING SIMULATION

Variable		Parameter for Quadrat	Parameter for Transect
Brown Stoneware	***	27	28
Tin Glazed	*	58	83
Creamware	**	1574	2271
Whiteware	***	236	296
Brick	**	16399	21664
Gunflint	*	2	6
18th century Ceramics	*	1780	2528
Architecture Group	**	686	988
19th century Ceramics	***	1534	2067
Total Ceramics	***	3423	4734
Total Artifacts	***	5312	7317

* - used in estimation evaluation
** - used in distribution evaluation
*** - used in both evaluations

$$\frac{(Fo-Fe)2}{Fe} = \text{accuracy index}$$

where

Fo = the observed frequency

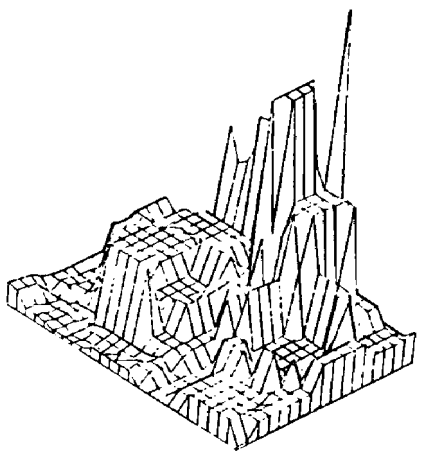
Fe = the expected frequency

It should be noted that this statistic is effective as an "accuracy index" or relative basis of comparison of the accuracy of the sampling designs and no statistical significance can be attached to the indices since minimum rather than maximum values represent the more accurate designs (Judge et al. 1975:106). The second test of sample accuracy was to construct a 95% confidence interval around the parameter estimate for each method and each sampling fraction and then recording whether or not the real value of the parameter fell within the interval. The third test of sample accuracy was to use a chi-square goodness-of-fit test to see if the set of estimated values as a whole matched with the set of known parameter values. Comparisons of sample estimates' precision were evaluated using methods described by Plog (1976) in which an F-statistic was was computed by forming a ratio of the squared standard errors of the mean.

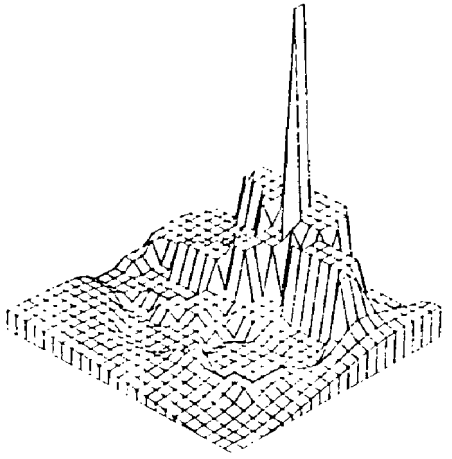
The effects of varied sample designs and varied sampling fractions on distribution maps were also evaluated. Distribution maps were considered because one of the main uses of plowzone artifact data is to prepare distribution maps of artifact classes

FIGURE 89

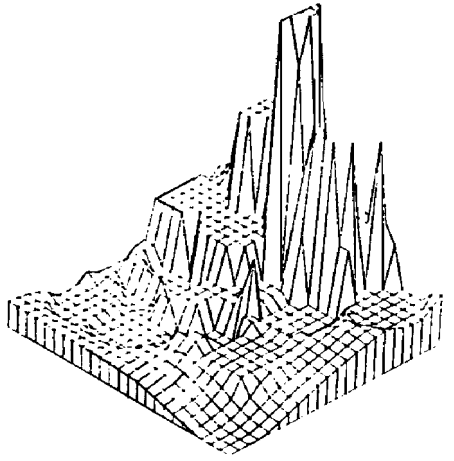
Sample Simulation Maps
Nineteenth Century Ceramics



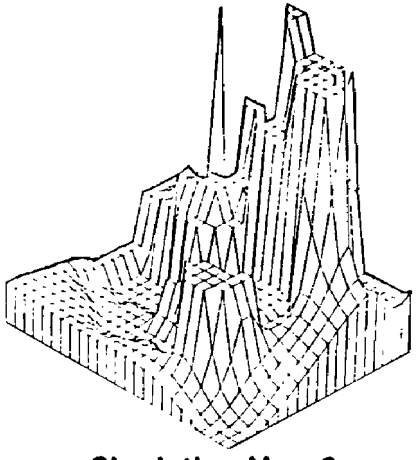
Complete Data



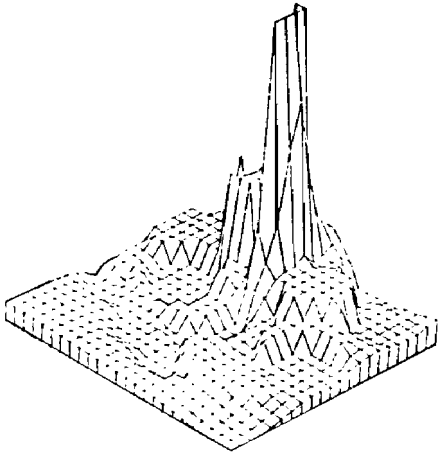
Simulation Map 1



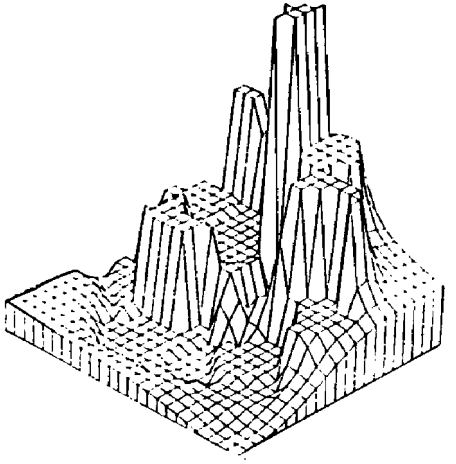
Simulation Map 2



Simulation Map 3



Simulation Map 4



Simulation Map 5

to see if spatial patterning in the plowzone distributions corresponds to the distribution of sub-plowzone features. Furthermore, distribution maps were analyzed because Ammerman et al. (1978) noted that the results of sample evaluations based on parameter estimates varied from results obtained from evaluations of sample artifact distributions. Five separate samples were drawn for each sample design and each sampling fraction. Maps of each sample for the artifact classes noted in Table 22 were then made and compared to maps made from the complete data base. Figure 89 shows a series of sample maps from the quadrat sampling simulation. The five maps from the sample iterations are shown with the map of the complete excavation data for the 25% sample of 19th century ceramics. All simulation maps are on file at UDCAR. The simulation maps were evaluated using the same methods used for the comparison of 1/4" and 1/2" screening techniques and both numbers of identified artifact clusters and frequency relationships of clusters were compared.

RESULTS OF SAMPLING SIMULATION

Tables 23 and 24 show the mean and standard deviation (standard error of the mean) values for the quadrat and transect simulation estimates. The known parameters for each artifact class are also listed. A simple comparison of the known parameters and the estimates shows that there are not many differences between them. Table 25 shows the 95% confidence intervals for all of the parameters and all of sample designs and sampling fractions. In only two instances (the 25% quadrat samples for 18th and 19th century ceramics), are the parameters not within the sample confidence intervals. And, even in these two cases the difference between the parameters and the confidence interval is quite small given the large values of the parameters. Table 26 shows the results of chi-square goodness-of-fit tests and there are no significant differences between the set of known parameters and the set of estimates as a whole for any sample method and sampling fraction.

Table 27 shows the accuracy indices computed for each sampling design and each sampling fraction. The lower values indicate the more accurate values. For the most part, there are very limited increases in the accuracy of estimates with increasingly large sampling fractions. When the sampling fraction was increased from 25% to 50%, which is a doubling of the sample size, the average increase in accuracy is only 53%. In other words, for the 25% to 50% sampling fraction increase, the accuracy increase is inversely proportional to the sample size increase. When the sample is increased from 50% to 75%, the average increase in accuracy is only 17%.

To summarize the accuracy evaluations with regard to sampling fraction, there are no significant differences among the sample estimates and the known parameter values regardless of the sampling fraction. Also, large increases in sampling fractions yield relatively small gains in accuracy of estimates. When

TABLE 23

QUADRAT TEST PARAMETERS AND ESTIMATES

Variable	Parameters	Estimates 25%	50%	75%
Brown Stoneware	27	28(11.94)	27(7.63)	27(3.96)
Tin Glaze	58	57(13.45)	59(9.80)	57(4.41)
Whiteware	236	231(60.52)	234(32.67)	232(19.66)
Gunflint	2	3(2.65)	2(1.40)	2(1.14)
18th century Ceramics	1780	1734(42.46)	1756(42.11)	1785(41.85)
19th century Ceramics	1534	1494(38.84)	1509(39.23)	1523(39.04)
Total Ceramics	3423	3389(58.56)	3473(58.51)	3395(58.16)
Total Artifacts	5312	5350(73.02)	5276(73.00)	5279(72.41)

TABLE 24

TRANSECT TEST PARAMETERS AND ESTIMATES

Variable	Parameters	Estimates 25%	50%	75%
Brown Stoneware	28	28(13.09)	27(8.72)	28(4.26)
Tin Glaze	83	81(18.92)	83(16.93)	82(6.08)
Whiteware	296	287(62.40)	302(57.32)	288(20.31)
Gunflint	6	6(3.73)	6(2.28)	6(1.39)
18th century Ceramics	2528	2490(50.56)	2550(50.52)	2499(49.89)
19th century Ceramics	2067	2020(45.79)	2021(45.64)	2076(45.17)
Total Ceramics	4734	4683(69.04)	4690(69.11)	4719(68.32)
Total Artifacts	7317	7262(85.72)	7367(85.86)	7274(84.97)

TABLE 25

ESTIMATE CONFIDENCE INTERVALS AND PARAMETERS

Variable	Sample Design	Parameter	Confidence Intervals 25%	50%	75%
Brown	Q	27	17-39	19-35	23-31
Stoneware	T	28	15-41	18-36	24-32
Tin Glaze	Q	58	44-70	49-69	53-61
	T	83	62-100	66-100	76-88
Whiteware	Q	236	170-292	201-267	212-252
	T	296	225-349	245-359	268-308
Gunflint	Q	2	0-6	1-3	1-3
	T	6	2-9	4-8	5-7
18th cen.	Q	1780	1692-1776*	1714-1798	1744-1826
Ceramics	T	2528	2439-2541	2499-2601	2449-2549
19th cen.	Q	1534	1456-1532*	1470-1548	1484-1562
Ceramics	T	2067	1974-2066	1975-2067	2031-2121
Total	Q	3423	3331-3447	3415-3531	3337-3453
Ceramics	T	4734	4614-4752	4621-4759	4651-4787
Total	Q	5312	5277-5423	5203-5349	5207-5351
Artifacts	T	7317	7177-7348	7281-7453	7189-7359

* cases where parameters do not fall within confidence intervals

TABLE 26

CHI-SQURE VALUES FOR GOODNESS-OF-FIT TEST BETWEEN PARAMETERS AND ESTIMATES

Sample Design	Fraction	Chi-Square	Probability
Quadrat	25%	3.01	.75 > p > .50
Quadrat	50%	1.75	.90 > p > .75
Quadrat	75%	.62	.199 < p > .975
Transect	25%	2.92	.75 > p > .50
Transect	50%	2.11	.90 > p > .75
Transect	75%	.89	.975 > p > .95

D.O.F. = 5

TABLE 27

ACCURACY INDICES - SAMPLE SIMULATIONS

Variable	Accuracy Index Quadrats 25%	50%	75%	Transects 25%	50%	75%
Brown Stoneware	.04	.00	.00	.00	.03	.00
Tin Glaze	.02	.02	.02	.05	.00	.01
Whiteware	.11	.02	.07	.27	.12	.21
Gunflint	.50	.00	.00	.00	.00	.00
18th C. Ceramics	1.19	.33	.01	.57	.19	.33
19th C. Ceramics	1.04	.41	.08	1.07	1.02	.04
Total Ceramics	.34	.73	.23	.55	.41	.05
Total Artifacts	.27	.24	.21	.41	.34	.25

studying sample size, the absolute number of sample elements is more important than the sample fraction. Increases in absolute sample size yields increases in sample accuracy up to a certain point, but beyond that point, the increase in accuracy per additional sampling unit diminishes. The data from the Whitten Road simulation would indicate that the threshold beyond which sample accuracy increases diminish in magnitude is 45 units for the quadrat design and 32 units for the transect design. With either the transect or the quadrat design, a 25% sample would have yielded accurate estimates of artifact frequencies.

The accuracy indices in Table 27 can also be used to compare the accuracy of the quadrat design estimates with those of transect designs. When quadrat and transect designs are compared for each parameter and similar sampling fractions, quadrat designs are more accurate than transect designs in 12 cases out of 24, transect designs are more accurate 9 times, and there is no difference between designs in 3 cases. Thus, quadrat designs have a slight advantage over transect designs with respect to estimate accuracy.

Table 28 shows the sample precision evaluations for varied sampling fractions. The values in Table 28 are computed so that an increase in the size of the value (the F-statistic) indicates increased precision with increased sample size. Although there are some major increases in estimate precision with larger samples for a few of the parameters, in most cases, especially the larger parameters, the gains in precision with increasing sample size are small compared to the magnitude of the sample size increase. Table 29 shows a comparison of the estimates' precision between the sample designs and the quadrat designs are consistently more precise than the transect designs.

Comparison of artifact distribution maps by sampling fraction and sample design yielded results similar to those from the comparison of sample parameter estimates. Appendix VI lists all of the evaluations of the sample maps and the results of the evaluations are summarized in Table 30. Comparisons of error frequencies show that the quadrat designs produce more accurate maps than the transect designs for all sampling fractions. Comparisons of sample fractions within designs shows that there are few gains in map accuracy as sample size increases. Indeed, there are few errors in any of the sample maps when they are compared to maps based on complete data.

DISCUSSION

The results of the Whitten Road sampling simulation show that for excavation of plowzone deposits, quadrat-based designs are more accurate and precise than transect-based designs when evaluations of distribution map accuracy and precision of parameter estimates are used. Using the same evaluation methods, it can be stated that gains in sample estimate and distribution map accuracy and precision are negligible beyond the 25% sampling fraction (between 32-45 sample units). These results are

TABLE 28

PRECISION COMPARISONS BY SAMPLING FRACTIONS

Variable	Quadrats			Transect		
	25/50	25/75	50/75	25/50	25/75	50/75
Brown Stoneware	2.25	9.50	4.22	2.47	8.94	3.62
Tin Glaze	1.25	8.95	7.18	1.89	9.53	5.05
Whiteware	1.18	9.45	7.98	3.43	9.47	2.76
Gunflint	2.80	7.00	2.50	3.50	7.00	2.00
18th Century Ceramics	1.00	1.03	1.03	1.02	1.03	1.01
19th Century Ceramics	1.01	1.03	1.02	.98	.99	1.01
Total Ceramics	1.00	1.02	1.02	1.00	1.01	1.01
Total Artifacts	1.00	1.02	1.02	1.00	1.02	1.02

TABLE 29

PRECISION COMPARISONS BETWEEN SAMPLING DESIGNS

Variable	Sampling Fractions		
	25%	50%	75%
Brown Stoneware	1.20	1.31	1.12
Tin Glaze	1.98	2.99	2.11
Whiteware	1.06	3.08	1.06
Gunflint	2.00	2.50	2.00
18th Century Ceramics	1.42	1.44	1.42
19th Century Ceramics	1.39	1.35	1.34
Total Ceramics	1.39	1.40	1.38
Total Artifacts	1.38	1.38	1.38

Transect/Quadrat: Note that in all cases, quadrats are more precise (efficient) than transects.

TABLE 30

SAMPLE MAP EVALUATION SUMMARY

Artifact Class	Sample	Designs and Sample Fraction					
		Q25	Q50	Q75	T25	T50	T75
Brown Stoneware	AC	4	5	3	5	5	5
	S	0	1	0	2	2	1
Creamware	AC	1	2	1	5	5	0
	S	0	0	0	5	2	4
Whiteware	AC	3	2	2	5	5	5
	S	1	0	3	5	3	3
Brick	AC	2	2	1	5	3	5
	S	1	0	1	3	0	1
Architecture Group	AC	0	4	3	5	3	1
	S	2	0	1	3	0	0
19th Century Ceramics	AC	3	1	4	3	3	1
	S	0	1	0	0	1	0
Total Ceramics	AC	4	3	2	5	3	0
	S	1	0	1	3	2	3
Total Artifacts	AC	3	4	4	5	4	2
	S	0	1	0	2	2	2
Totals	AC	20	23	20	38	31	19
	S	5	3	6	23	12	14

AC - Additional clusters missed
S - Map structure errors

consistent for classes of artifacts of varied frequency.

It is important to note that the results of the Whitten Road sample simulation are consistent with results of other site-specific sampling simulations. Zierden et al. (1986:3-39) and Jermann and Dunnell (1979:58-59) note similar results with respect to mapping errors and sampling fractions and suggest cut-off points ranging between 10% and 20% (50-60 sample units), based on limited simulations. Ammerman et al. (1978) suggest a 20% cut-off point (45-50 sample units) with regard to parameter estimation. Thus, the results of numerous sampling simulations consistently suggest that there are few significant gains in accuracy and precision of distribution maps and parameter

estimations when plowzone excavation samples exceed the 20-25% fraction, or 50-60 sample units. In conclusion, it is suggested that in the future, excavation of plowzone deposits at historic sites be limited to 20-25% samples, or at least 50 sample units, to maximize information gain and minimize excavation costs. After sample excavations are completed, the plowzone can be removed with heavy equipment to look for underlying features.