ANALYSIS OF PREHISTORIC SITE SURVEY RESULTS

This section of the report will describe various analyses of the results of the prehistoric site survey. The first analysis described is a comparison of the predictive model's expected results and the observed results of the field survey. The goal of this analysis was to provide an evaluation of the prehistoric predictive model. Other analyses sought to provide a broader cultural perspective for survey results and consisted of: 1) a general analysis of locational variables; 2) a factor analysis of site location variables; 3) a statistical analysis of site characteristics; and, 4) a discussion of human utilization of bay/basin features. All of these analyses include comparisons to other survey results from the Middle Atlantic region. Special comparisons with data from controlled survey results from the High Coastal Plain region of the Patuxent drainage of Maryland (Steponaitis 1980; 1983a; 1983b; 1984) will also be included due to the Patuxent region's environmental similarities to the Route 13 High Coastal Plain setting. Finally, a general discussion of the survey results for each of the major time periods is included.

EVALUATION OF PREHISTORIC PREDICTIVE MODEL

The following provides a discussion of the relationship between the present survey's results and the prediction of prehistoric site locations provided in the original Route 13 Project Planning report (Custer et al. 1984: 76-90; Attachment V). Although the predictive model derived from logistical regression analysis of LANDSAT data has been successfully tested, in a preliminary manner, in southern and central Kent County (Eveleigh, et al. 1983; Custer, et al. n.d.; Eveleigh 1984), and although a similar logistical regression model was successfully tested in south central New Castle County (Wells, et al. 1981; Wells 1981), continued testing, and subsequent refinement, of predictive models is always desirable. Therefore, the survey results of this project were used to test the local predictions generated from the logistical regression analysis. Special consideration of testing of the predictive model's results is also presented here because several recent articles (Berry 1984; Tainter 1984) have questioned the use of predictive models derived from multivariate statistical analyses of site locations, such as logistical regression and discriminant analysis (Kvamme 1981; Burgess et al. 1980), in cultural resource management situations. Because many of the criticisms focus on the models' accuracy, it was hoped that an analysis of the accuracy of the results of the LANDSAT model used for the Route 13 Corridor would support its use as a cultural management tool.

The best field test of the models' predictions would be a survey based on a stratified random sample. Stratification could be based on two factors: 1) predicted probability zone, and 2) environmental factors used in the development of the model, such as surface water setting, presence/absence of modern and ancient marshes, and local soils. Initial tests of the logistical model's predictions incorporated some of these features, but relied mainly on previously recorded sites rather than new field data. The present Route 13 fieldwork incorporates some of these aspects as well; however, only a portion of one study area's survey, the Blackbird area subsurface testing, contained a randomized sample component. The remainder of the fieldwork represents a non-random and uncontrolled "grab sample" with many inherent biases.

Probably the most important bias in the sample is introduced by the fact that the study areas were chosen for field survey because they included high concentrations of cultural resources. Consequently, the high and medium probability zones are over-represented in the samples studied. This over-representation means that tests of model predictions are more likely to detect errors of inclusion rather than errors of ommission. In other words, we will be able to detect errors where the model says a site is present and a site is <u>not</u> present. However, we will not be able to study as effectively errors where the model says a site is not present and a site truly <u>is</u> present. The nonproportional representation of some kinds of probability and environmental zones also introduces numerous sources of sampling error, especially in the absence of a random component in the sampling design.

Because of the non-random nature of the sample, non-normality of some of the data and the various uncontrolled sources of sampling error and biases, complex statistical tests of goodness-of-fit between model predictions and field results were not used. Rather, simple descriptive statistics and simple difference-of-proportion and difference-of-mean tests were used. These techniques were applied to the survey data from the Appoquinimink and Blackbird study areas. The St. George study area data were not considered due to the small size of the study area and the low number of sites found within it. However, it can be noted that all sites discovered in the St. Georges study area, except for one, were located within the high and medium probability zones.

As a final introductory note it should be pointed out that all fieldwork within the three study areas was carried in a "blind test" setting. Although the field crews, including the project manager, knew that they were surveying within areas that were likely to find sites, ie. that they were in "problem areas", none knew exactly where the high, medium, and low probability zones were located. Therefore, all areas were field inspected with the same intensity and coverage. "Likely areas" within the study areas did not recieve special treatment.

The first tests of model predictions focused on measures of site occurrence. These tests considered the goodness-of-fit between expected site counts and densities within the varied probability zones, based on predictive models, and the observed field results. Tables 12 - 14 list the data used from field observations for the Appoquinimink area, the Blackbird area, and both areas combined. Examination of the first row of Tables 12 - 14 shows the over-representation of high and medium probability zones in the test data. Consequently, figures for low probability zones are somewhat suspect.

TABLE 12

DENSITY MEASURE TEST DATA - APPOQUINIMINK AREA

Variable	Probability Zones		
	High	Medium	Low
TotalArea (ha)	311.40 (35%)	476.28 (53%)	104.76 (12%)
# of sites	54	27	8
# of 400m grid cells	15	17	7
# of cells w/ sites	13	9	2
% of cells w/ sites	0.87	0.53	0.29

TABLE 13

DENSITY MEASURE TEST DATA - BLACKBIRD AREA

Variable	Probability Zones		
	High	Medium	Low
Total Area (ha)	91.80 (11%)	496.08 (58%)	274.32 (31%)
# of sites	23	127	30
# of 400m grid cells	4	20	12
# of cells w/ sites	4	18	5
% of cells w/ sites	1.00	0.90	0.42

TABLE 14

DENSITY MEASURE TEST DATA - COMBINED AREAS

Variable	Probability Zones		
	High	Medium	Low
Total Area (ha)	403.20 (23%)	972.36 (55%)	379.08 (22%)
# of sites	77	154	38
# of 400m grid cells	19	37	19
# of cells w/ sites	17	27	7
% of cells w/ sites	0.89	0.73	0.36

The first comparison of observed and expected results utilized a simulated 400m grid cell data base. This grid cell data base was utilized for a number of reasons. First, the 400m grid matched the grid system used for the development of the initial predictive model (Custer et al. 1984:76-90). Also, a 400m grid was used in the initial testing of the predictive model (Eveleigh 1984; Eveleigh et al. 1983; Custer et al. n.d.). Most importantly, application of the 400m grid allowed the generation of comparable measurements of site occurrence for analysis. The predictive model for site locations generated a series of probability values for the occurrence of sites within specified units of area. Field observations generate counts of sites within specified units of area. In order to directly compare results of field testing with the model predictions, some kind of data transformations must be undertaken. Statistical distributions which transform site densities into probability values, such as the Poisson, are of limited utility (see Custer 1979; 1980) because archaeological data do not clearly meet the necessary requirements of independence-of-events (Thomas 1975). In fact the logistical regression of distribution was used to avoid the problems encountered in applying the Poisson distribution. Consequently, it is easier to transform the probability values into expected frequencies of sites and then directly compare the expected site frequencies within different probability zones with the observed frequencies generated from sample survey data.

The transformation of the model's probability values was accomplished by first overlaying a 400m grid template on maps of the project areas. The number of 400m cells falling within each of the three probability zones was then recorded. The number of cells which contained archaeological sites was also recorded. In order to calculate the expected number of 400m cells containing sites, the midpoint value of each probability class was multiplied by the number of cells within that probability zone. The observed number of cells with archaeological sites was simply a count of the cells which contained archaeological sites.

Table 15 lists the observed and expected values for all three data sets. The observed and expected values were then compared using the chi-square distribution and the test values, degrees-of-freedom, and probability values are all included in Table 15. The test statistics all indicate no statistically significant differences between the observed and expected frequencies of grid units containing archaeological sites.

TABLE 15

OBSERVED AND EXPECTED SITE COUNTS FOR 400m GRID UNITS

Area	Prob. Zone	Expect.	Observ.	Chi-sq.	D-o-F	p-Val.
Appoquin.	High	13	13			
	Medium	10	9			
	Low	1	2	1.10	2	.5<p<.75
Black.	High	4	4			
	Medium	13	18			
	Low	4	5	2.17	2	.25<p<.5
Total	High	17	17			
	Medium	23	27			
	Low	5	7	1.50	2	.25<p<.5

Probability values from Parsons (1974:824)

As an additional test of the model's predictions using the 400m grid data, the grid counts noted in Table 15 were converted to percentages and the expected and observed percentages were compared using the difference-of-proportion test. The percentages and test statistics are noted in Table 16. Only one combination of values, the Blackbird area medium probability zone, shows a significant difference between the expected and observed results. To summarize to this point, on the whole, the observed site frequency data seem to fit well with the expected results. In the case of the Blackbird medium probability zone, however, there are significantly more sites observed, compared to the expected results. This anomaly will be discussed later in conjunction with other test results.

Data on site size and site function were also used to test logistical regression model predictions. The initial model applications (Eveleigh et al. 1983:30-35; Custer et al. 1984:87-91, 100-102; Custer et al. n.d.) all indicated that the high probability areas were most likely to contain large macroband base camps when compared to other probability zones. Furthermore, it was noted that microband base camps were most likely to be found in high and medium probability zones. Procurement sites were found in all three zones, but were the major site component of the low probability zones.

TABLE 16

OBSERVED AND EXPECTED SITE PERCENTAGES FOR 400m GRID UNITS

Area	Prob. Zone	Expect.	Observ.	Test Value	p-Value
Appoquin.	High	.87	.87	0	1.00
	Medium	.58	.53	.38	.70
	Low	.14	.29	.65	.52
Black.	High	1.00	1.00	0	1.00
	Medium	.65	.90	1.89	.06
	Low	.33	.42	.42	.68
Total	High	.89	.89	0	1.00
	Medium	.62	.73	.99	.32
	Low	.26	.36	.70	.48

Probability values from Parsons (1974:818)

Before considering the site size and function data, the temporal settings of the sites needs to be discussed. Table 17 shows the distribution of components at sites for the Appoquinimink and Blackbird areas as well as for the total survey area. It is clear that most sites with identifiable components (73%) date to the Woodland I period. The Blackbird area provides data on all but one of the Archaic sites, and Woodland II sites are present, but not common, in all study areas. Consequently, the tests of model predictions, including the previously discussed density measures, primarily pertain to the Woodland I period. This bias matches the bias in the data base used to develop the initial model (Eveleigh et al 1983: 30-32).

TABLE 17

FREQUENCY OF DATABLE COMPONENTS IN STUDY AREA

Area	Prob. Zone	Archaic	Woodland I	Woodland II
Appo.	High	1	9	3
	Medium	0	1	3
	Low	0	1	1
Black.	High	0	6	0
	Medium	4	28	6
	Low	2	8	0
Total	High	1	16	3
	Medium	4	30	10
	Low	2	10	1

Two kinds of site data were analyzed to see if there is a differential distribution of functional site types among the probability zones: data on site area and data on functional site classifications. Table 18 shows a series of descriptive statistics for site area measures among the varied study areas and probability zones. The very high values for the standard deviation, coefficient-of-skewness, and coefficient-of-kurtosis indicate that the area measures are not normally distributed. Non-normality of the area measure distribution for all study areas and probability zone combinations is also indicated by the Q-Q correlation statistics (Johnson and Wichern 1982:151-160) presented in Table 19. Because the site area measures are not normally distributed, standard difference-of-mean or t-tests cannot be used to assess the differences of site areas between varied probability zones.

An alternative approach is to consider a variety of non-parametric range statistics which are listed in Table 20. The

TABLE 18

DESCRIPTIVE STATISTICS - SITE AREA MEASURES

Area	Prob.Zone	N	Mean	Std.Dev.	Skewness	Kurtosis
Appo.	High	28	6235.29	15146.50	4.42	22.05
	Medium	17	1366.76	859.26	1.38	4.03
	Low	6	5398.67	10174.90	1.73	4.08
Black.	High	16	5074.50	5613.67	1.73	4.84
	Medium	87	9371.03	13350.00	2.88	13.45
	Low	21	8909.23	18116.60	3.27	13.26
Total	High	44	5813.18	12464.50	5.08	30.54
	Medium	108	7868.87	12363.90	3.23	16.18
	Low	28	7871.96	16317.90	3.51	15.69

TABLE 19

Q-Q STATISTICS - SITE AREA MEASURES

Data Set	Prob. Zones			Total Area
	High	Medium	Low	
Total	.6198	.7758	.6961	.7292
Appo.	.6000	.9041	.7549	.5368
Black.	.8558	.8081	.6988	.7746

No significant values noted based on comparison with Q-Q statistics in Johnson and Wichern (1982:156).

TABLE 20

NON-PARAMETRIC RANGE STATISTICS - SITE AREA MEASURES

Area	Prob. Zone	Min.	25th	Med.	75th	Max.	Tot. Range	Med-75th Range
Total	L	18	784	1974	6968	80904	80886	4994
	M	177	1276	2787	8622	82914	82737	5835
	H	446	1029	2085	5144	80359	79913	3059
Appo.	L	409	428	904	1414	26012	25603	1510
	M	446	697	1263	1582	3361	2915	319
	H	446	981	1574	4709	80359	79913	3135
Black.	L	18	1161	2323	10219	80940	80886	7896
	M	177	1626	4181	12542	82914	82737	8361
	H	698	1490	2845	6015	20032	19334	3170

very high values for total range show a wide spread of the the distribution of site areas and the large differences between the 75th percentile and maximum values are the cause of the non-normality of the area distribution. Consideration of the range statistics also provides some insights on differential site size among the varied probability zones. Within the Appoquinimink area, the range statistics support the contention that site size generally increases moving from the low probability zones, through the medium probability zones, into the high probability zones. Except for the minimum values and the 75th percentile values, there is a constant increase in all values moving from the low to the high probability values.

In contrast, the range statistics for the Blackbird area do not show such a pattern. In general, the 25th percentile, median, 75th percentile, and maximum value statistics are all largest in the medium probability zone of the Blackbird area. Furthermore, the 75th percentile and maximum value statistics of the low probability zone are larger than the values for the high probability zone. The total range values for the Blackbird medium and low probability zones are also more than four times as large as the range for the high probability zones. A similar relationship is seen for the median-75th percentile range values. These observations of the range statistics all indicate that the medium and low probability zones of the Blackbird area contain more larger sites than would be expected based on other applications of the predictive model.

Cross-tabulation of functional site types and probability zones shows the same pattern as that shown by the analysis of site area statistics. Table 21 shows these cross-tabulations. For the Appoquinimink area, there are insufficient data to apply a chi-square test for dependence of variables; however, the general trend of the site frequencies show more macroband base

TABLE 21

CROSS-TABULATION OF SITE TYPES AND PROBABILITY ZONES

	Probability Zones		
	Low	Medium	High
Appoquinimink Area			
Procurement	5	16	21
Macroband Base Camp	1	1	11
Microband Base Camp	1	1	1
Base Camp	0	1	1
Blackbird Area			
Procurement	20 (19)	67 (69)	14 (13)
Macroband Base Camp	3 (2)	6 (7)	1 (1)
Microband Base Camp	2 (3)	12 (10)	1 (2)
Chi-square=2.39 D.o.F.=6 .50<p<.75			
Total Area			
Procurement	26 (25)	87 (85)	35 (38)
Macroband Base Camp	4 (4)	7 (14)	13 (6)
Microband Base Camp	3 (3)	14 (11)	2 (5)
Base Camp	1 (2)	8 (6)	1 (2)
Chi-square=16.05 D.o.F.=6 p<.01			

Expected values for chi-square test in parentheses

camps in the high probability zone as is predicted by the logistical regression model. Table 22 shows a series of difference-of-proportion test statistics for comparisons of percentages of site types among the probability zones. Two significant differences are noted: 1)there are significantly more procurement sites in the medium probability zone compared to the high probability zone; and, 2) there are significantly more macroband base camps in the high probability zone compared to the medium probability zone. These differences match the previous interpretations of the predictive model fairly closely.

Table 21 also shows the cross-tabulation of site types and probability zones for the Blackbird area. The chi-square test was applied and the low values for the test statistic indicates that the two variables are independent. In other words, the frequencies of any given site type within the probability zones are not larger or smaller than would be expected by chance alone. Therefore, the Blackbird area site type distribution among

TABLE 22

DIFFERENCE-OF-PROPORTION TESTS - SITE TYPES AND PROBABILITY ZONES

Variable	Zones	Proportions		Test Statistic	Prob.
Procurement	L-M	L-.714	M-.842	.734	.48
	M-H	M-.842	H-.618	1.707*	.09
	L-H	L-.714	H-.618	.483	.64
Macroband BC	L-M	L-.142	M-.052	.766	.46
	M-H	M-.052	H-.324	2.260*	.04
	L-H	L-.142	H-.324	.957	.34
Microband BC	L-M	L-.142	M-.052	.766	.46
	M-H	M-.052	H-.029	.425	.68
	L-H	L-.142	H-.029	1.269	.22
Base Camp	L-M	L- 0	M-.052	.619	.56
	M-H	M-.052	H-.029	.425	.68
	L-H	L- 0	H-.029	.459	.66

* - Significant Values ($p<.10$)
Probability values from Parsons (1974:818)

probability zones does not match the predictions of the logistical regression model.

A cross-tabulation of site types and probability zones for the entire study area is also included in Table 21. Sufficient data are present to apply the chi-square test and the probability value indicates that there is a dependent relationship between the two variables. Comparison of the expected and observed values shows that there are higher than expected frequencies of macroband base camps in the high probability zones and that there are higher than expected frequencies of microband base camps and generalized base camps in the medium probability zone. These results match the interpretations of the predictions of the logistical regression model almost exactly and very closely approximate other logistical regression model test results (Eveleigh et al. 1983:23-33, Tables 5 and 6). It is important to note that the higher than expected values of the total data match the results of the difference-of-proportion tests for the Appoquinimink area (Table 22). The close match of the Appoquinimink data, total area data, and expected results occurs in spite of the fact that the Blackbird data, which was included in the total data, did not match model predictions in several instances. When the fact that the Blackbird data accounts for more that 67 percent of the total cross-tabulation data is added, it becomes apparent that the incongruences between the Blackbird site type data and model results are rather minor. Even though the Blackbird data represent an important part of the data base, the incongruences are outweighed by the instances where the survey data do match model predictions.

A final statistical analysis applied to test the validity of the predictive model was a discriminant analysis. Discriminant analyses consider the differences between two or more sets of observations based on a series of measured variables (Overall and Klett 1972:243-280). Usually, continuous variables are used in the analysis and for the Route 13 data three variables were used: distance to water, slope, and elevation. The two groups studied were the high and medium probability zone sites. These two zones' sites were considered because some of the data noted earlier showed that not all of the predicted differences in site locations between these zones were present. If sites from each of the two probability zones are truly different, as the predictive model would indicate, the discriminant analysis should produce a set of discriminant coefficients which generate discriminant function group means which are significantly different in a statistical sense. The discriminant analysis used here is the two-group discriminant function analysis included in the STATPRO software package (Wadsworth Professional Software 1984). A discriminant function was produced and took the following form:

$$.0038(\text{Distance-to-Water})+.1651(\text{Slope})-.2358(\text{Elevation})$$

The high probability group mean was 16.10 and the medium probability group mean was 57.06. The total chi-square value comparing the two group means was 24.86 (degree-of-freedom=6, $p<.005$) which shows a statistically significant difference between the two groups. Thus, a discriminant analysis shows significant differences between site location characteristics in the high and medium probability zones. These findings indicate that the predictive model is effectively separating different site locations in the high and medium probability zones. In conclusion, it can be noted that in general the field data do match the predictions generated by the logistical regression model and the model is sufficiently accurate to use as a basis for cultural resource management decisions and applications. The only incongruences which arise occurred in the medium probability zone of the Blackbird area. These zones have more sites than would be expected from the model predictions and also have more larger sites than the model would predict. It is useful to compare the Blackbird area with areas where the model was successfully tested in order to determine why the model was less accurate in the Blackbird area. Areas of successful testing included the Appoquinimink area noted in this study, and earlier studies (Wells et al. 1981; Wells 1981), and the lower St. Jones/Murderkill drainage areas (Eveleigh et al. 1983; Eveleigh 1984; Custer et al. n.d.). These areas are similar in that they fall within the mid-drainage physiographic setting (Custer 1984a:27) and include the oligohaline zone of the drainages. Furthermore, the major drainages and their higher order tributaries provide the only sources of fresh surface water. The Blackbird area is similar to a certain extent in that it includes the oligohaline setting of the Blackbird Creek. However, it differs significantly from the other areas with its high incidence of bay/basin features. These features represent an additional source of fresh water in interior areas away from the major drainages and also represent game-attractive settings.

As was noted in the original Route 13 report (Custer et al. 1984: 100-102, 217-218, Attachment V), the predictive model did recognize some of the bay/basin features as loci of high and medium probability settings. However, the comparison of model predictions and survey data indicate that the model did not recognize enough of these settings as high probability settings. It is suggested here that this recognition problem is related to the problem of recognition of the bay/basin features themselves with the LANDSAT multispectral scanner data. During our fieldwork we found that many of the bay/basin features encountered on the ground were not mapped on either the USGS 7.5' quadrangle maps or on the USDA soil survey maps (see Table 23). In some cases, bay/basin features were on the USGS quad maps, but not on the USDA maps, and vice versa. The resolution of the air photo base maps for both the USGS and USDA maps is 2-3 meters, yet these air photos still missed many of the bay/basin features. The resolution of the LANDSAT multispectral scanner data is 80m, and even with the added discrimination capacity of the four color bands, it is not surprising that the LANDSAT data missed some of the bay/basin features. The data in Table 23 also seems to indicate that the bay/basin recognition problem is greater in wooded areas compared to cultivated areas. Consequently, the bay basin recognition problem is not only related to their size and the resolution of the remote sensing device. It is also related to the masking of the features by dense tree growth. This masking problem is especially acute in the Blackbird area where many of the wood lots have been previously logged and are presently composed of dense secondary tree stands 50 to 100 years old.

TABLE 23

BAY/BASIN MAPPING ERRORS

Surface Condition	Ground Truth Total	USGS Maps Only	USDA Maps Only	Both USGS and USDA
Wooded	36	17	1	1
Cultivated	106	29	23	8

Given these problems in the Blackbird area, it is suggested here that the medium and high probability zones be combined into a single high probability zone (p<.50), which contrasts with all other probability settings. It is also suggested that no non-proportional sampling be undertaken for proposed rights-of-way in the Blackbird area until the site predictions can be upgraded and revised. This recommendation will be discussed more completely in the management section of this report.

It should be noted that several options exist for refining the predictive model in the Blackbird area. Essentially, the problem is not with the synoptic logistical regression technique. Rather, it is a problem of specific remote sensing of a culturally relevant environmental feature. One option would be to develop a new geographic data base of environmental variables for the Blackbird area using the newly available thematic mapper (TM) data from LANDSAT 4. The TM data has a 30-meter resolution, which as a great improvement over the 80-meter resolution of the older LANDSAT data. Also, the TM data records seven color bands instead of four, thereby increasing the ability to discriminate environmental features, such as bay/basins (see discussion in Eveleigh 1984:146-148). In addition to the thematic mapper, traditional air photo data and low altitude helicopter photography could be used to supplement the satellite data to develop a more accurate environmental data base.

SETTLEMENT PATTERN AND SITE LOCATION ANALYSIS

The data collected for prehistoric archaeological sites during this survey of the northern portion of the Route 13 corridor can be analyzed to provide numerous insights into the lifeways of the prehistoric cultures of the High Coastal Plain area of the Delmarva Peninsula. The data from the present survey are especially useful for site location analysis because for some areas, such as the Appoquinimink study area, the data base is the closest approximation of a complete survey of a large area that will ever be available in Delaware. The Blackbird area data are also especially interesting because they are derived from a combination of broad surface survey area coverage, similar to that of the Appoquinimink area, and a stratified random sample of wooded areas requiring sub-surface testing. Indeed, the addition of an extensive sub-surface testing program for many wooded settings of the varied survey settings, which are often ignored in regional surveys, provides much of the more interesting site data. Furthermore, the Route 13 survey collected especially good data on site size and context with regard to environmental settings, and relationships to predictive models, which are not availible in other archaeological data bases such as the Delmarva Archaeological Data System (DADS).

A series of cross-tabulations of site location variables and site characteristics and simple tabulations of frequencies of sites by environmental variables will be presented. Variables chosen correspond to those seen as important in other regional site location analyses (Custer 1980, 1984a; Cavallo and Mounier 1980; Gardner 1978, 1982; Stewart 1980; Grossman and Cavallo

1982; Galasso 1983). Where possible, the study area data was compared to the wider data base for the Delmarva Peninsula High Coastal Plain included within the Delmarva Archaeological Data System.

Surface Water Type. The first site location variable to be studied was the surface water type associated with each site. Table 24 shows the frequency of sites associated with each type of surface water setting for the Route 13 Corridor, the entire High Coastal Plain data base as recorded in DADS (exclusive of the Route 13 sample), and for the Route 13 Corridor and the DADS High Coastal Plain data combined. It can be noted that the Route 13 sample provides coverage of bay/basin features that is not included in the DADS data base. On the other hand, marsh/swamp and estuarine bay settings are not present in the Route 13 sample due to the corridor placement. Thus, the Route 13 corridor data provides a sample of mainly interior site settings and the DADS sample is oriented more toward coastal environments. This interior/coastal dichotomy is also apparent considering the fact that the bulk of the Route 13 data come from the interior bay/basin settings of the Blackbird area and the majority of the DADS High Coastal Plain sample comes from the Wilke-Thompson (1977) survey of coastal Kent County, MD. When the total data are combined, a representative cross-section of High Coastal Plain site data is provided. This feature of the total data base makes further analysis of site locations more meaningful. The data in Table 24 show that the most prominent surface water setting is composed of major and minor streams for both interior and coastal areas. In the total data base, streams account for 71% of the site locations. Bay/basins are important in the interior settings and account for 32% of the site locations. In light of these figures it is not surprising that surface water setting variables account for the greatest amount of site location variance in the logistical regression models (see Wells

TABLE 24

SURFACE WATER TYPE FREQUENCIES

Surface Water Type	Route 13 Sample	High Coastal Plain Sample	Total
Stream	144	295	439
Bay/Basin	84	16	100
Bay/Basin/Stream	35	0	35
Marsh/Swamp	0	9	9
Estuarine Bay	0	37	37

1981:41-43). Steponaitis (1983a, 1983b, 1984) notes similar results in the Patuxent region.

Within the stream category of surface water settings, approximately 20% of the sites are located at stream confluences (Route 13 data - 31 sites out of a total of 144 - 21%; DADS data for High Coastal Plain - 57 sites out of a total of 295 - 19%; combined data bases - 88 sites out of a total of 439 - 20%). This percentage value is somewhat lower than would be intuitively expected from traditional non-quantitive settlement pattern analyses (Gardner 1978; 1982). The lower percentage value for the High Coastal Plain data probably is indicative of the fact that the juxtaposition of several environmental variables, such as wetlands, well-drained soils, <u>and</u> surface water, is more important for determining site locations than the effect of any single resource as shown by the logistical regression model. Also, the general presence of surface water is more important than the presence of stream confluences.

In order to further investigate the role of surface water settings in site location analysis, the various surface water types were cross-tabulated with sites' time periods of occupation (Table 25). In general,the data in Table 25 indicate that stream settings are the most important surface water settings for all time periods and show an increasing relative frequency of use through time. Bay/basin features begin to show some utilization in Archaic times, show peak utilization in Woodland I Period times, and show declining frequency of utilization in Woodland II Period times. (A special discussion of bay/basin utilization is presented later in this section.) Marsh/swamp and estuarine bay settings show increasing utilization through time with a climax in Woodland II times. This pattern of increasing frequency of coastal utilization through prehistoric time may be a function of destruction of sites by the Post-Pleistocene sea level transgression (Kraft et al. 1976); however, there are arguments, based on sound survey data and radiocarbon dates from throughout the Middle Atlantic Coast, which suggest that intensive utilization of bay coastal resources is a relatively late (post - 3000 BC) phenomenon in the Delmarva region (Custer 1984a: 63-64; Custer 1984b; Custer and Stewart 1983; Waselkov 1982; Potter 1982), and the Middle Atlantic Coast in general.

Tables 26 and 27 show cross-tabulations of the surface water settings and site types. For the Route 13 sample (Table 26), there are sufficient data on datable site components to do separate cross-tabulations of site types and surface water settings for the Archaic, Woddland I, and Woodland II periods. For the total Route 13 area data there are sufficient data to apply the chi-square test (chi-square = 13.31, degrees-of-freedom = 9, $.10<p<.25$). The probability values very marginally indicate that there is no dependent relationship between the site functions and surface water settings. The combined data base

TABLE 25

CROSS-TABULATION OF SURFACE WATER TYPE AND TIME PERIOD

Route 13 Sample

Time Period	Stream	Bay/Basin	Bay/Basin/Str.	Springhead
Archaic	3	4	1	0
Woodland I	26	29	9	1
Woodland II	9	4	1	1

High Coastal Plain Sample

Time Period	Stream	Bay/Basin	Marsh/Swamp	Estuarine Bay
Paleo-Indian	2	0	1	0
Archaic	9	0	2	0
Woodland I	24	1	4	0
Woodland II	33	2	5	5

Route 13 and High Coastal Plain Samples Combined

Time Period	Stream	Bay/Bas.	Bay/Bas./Str.	Spr.	Mar./Sw.	Est. Bay
Paleo.	2	0	0	0	1	0
Archaic	12	4	1	1	2	0
Wood. I	50	30	9	1	4	0
Wood. II	42	6	1	1	5	5

cross-tabulation in Table 27 (bottom) also contains sufficient data to apply a chi-square test (chi-square = 11.67, degrees-of-freedom = 9, $.05<p<.10$) and the values here do show a dependent relationship between site types and surface water setting. The following significant differences are noted between the expected and observed results: 1) there are more procurement sites associated with bay/basin features than would be expected by chance alone; 2) there are fewer procurement sites associated with streams than would be expected by chance alone; 3) there are more macroband base camps and generalized base camps associated with streams than would be expected by chance alone; and, 4) there are fewer macroband base camps and generalized base camps associated with bay/basins than would be expected by chance alone. Thus, base camps, especially the earlier ones, are more commonly found at stream settings in the High Coastal Plain while bay/basin features are more commonly associated with the more emphemeral procurement sites. Steponaitis (1983a: Fig. 4) notes similar results from the Patuxent region.

TABLE 26

CROSS-TABULATION OF FUNCTIONAL SITE TYPES AND SURFACE WATER SETTING - ROUTE 13 DATA

	Stream	Bay/Basin	Bay/Basin/Stream	Springhead
All Time Periods				
Procurement	73	59	18	2
Macro-band B.C.	18	3	3	0
Micro-band B.C.	8	6	4	1
Base Camp	4	3	3	0
Archaic Period				
Procurement	0	1	0	0
Macro-band B.C.	1	1	0	0
Micro-band B.C.	1	0	0	0
Base Camp	0	0	1	0
Woodland I Period				
Procurement	9	12	1	0
Macro-band B.C.	8	3	2	0
Micro-band B.C.	4	4	2	1
Base Camp	3	0	2	0
Woodland II Period				
Procurement	4	1	0	1
Macro-band B.C.	5	0	1	0
Micro-band B.C.	0	2	0	0
Base Camp	0	0	0	0

TABLE 27

CROSS-TABULATION OF FUNCTIONAL SITE TYPES AND SURFACE WATER SETTING - DADS HIGH COASTAL PLAIN DATA AND TOTAL DATA

	Stream	Bay/Basin	Marsh/Swamp
DADS High Coastal Plain Data			
Procurement	8	1	2
Macro-band B.C.	1	0	1
Micro-band B.C.	2	0	0
Base Camp	12	0	2
Total Data - DADS HCP and Route 13 Combined			
Procurement	81 (89)	80 (72)	2 (2)
Macro-band B.C.	19 (14)	6 (11)	0 (0)
Micro-band B.C.	10 (12)	11 (10)	1 (0)
Base Camp	16 (12)	6 (10)	0 (0)

Values in parentheses indicate expected values for the chi-square test.

Distance to Water. Another variable recorded for the sites found in the Route 13 Corridor is distance to water source. Figure 37 shows a frequency distribution for water distance for both the Route 13 data and the DADS High Coastal Plain data. The frequency distribution shows that the data are so highly skewed toward smaller values that it is meaningless to consider this variable for the total data base any further. Descriptive statistics (mean = 77.27, standard deviation = 85.05, coefficient of skewness = 1.86, coefficient of kurtosis = 8.62) and the water distance Q-Q statistic (.8985, .25>p>.10) also indicate a non-normal distribution. The most that can be said from looking at Figure 37 is that almost all (90%) of the interior Route 13 sample sites are within 100m of water. On the other hand, for the more coastal DADS High Coastal Plain data base, the distance within which 90% of the sites fall is 175m. The statistical significance of this difference is not easily assessed, given the non-normality of the data base. However, the difference may indicate slightly different settlement patterns in the two areas. It may be that people living in the coastal settings are more willing to sacrifice accessibility of surface water in order to maximize accessibility of other resources. The closer association of sites and surface water settings in the more interior Route 13 sample sites may also be mirroring the very close correspondance between bay/basin game-attractive locales and procurement sites.

Although the total water distance data are not particularly amenable to analysis, water distance data from the Route 13 area

TABLE 28

WATER DISTANCE DESCRIPTIVE STATISTICS BY TIME PERIODS

Time Period	N	Mean	Std. Dev.	Skew	Kurtosis	Test Q-Q	Q-Q(p<.10)
Archaic	7	65.00	44.63	-.25	1.61	.95	.88
Woodland I	56	53.43	64.85	1.21	3.56	.90	.97
Woodland II	14	41.79	49.97	1.14	3.04	.91	.91

TABLE 29

DIFFERENCE-OF-MEAN TEST - WATER DISTANCE

	Archaic	Woodland I	Woodland II
Archaic	------		
Woodland I	.61(.54)	------	
Woodland II	1.07(.28)	.73(.46)	------

p-values from Parsons (1974:818) in parentheses

for the individual periods were more useful. Table 28 shows a series of descriptive statistics for water distance during various time periods and although the distributions are skewed and leptokurtic, the Q-Q statistics indicate normal distributions. Because the data are normally distributed, difference-of-mean tests were applied to see if there were significant differences in distance to surface water between time periods. Table 29 shows the test statistics and none show significant differences. Therefore, sites from all time periods in the Route 13 area are located approximately the same distance from surface water sources.

Elevation. The elevation of sites was also analyzed. Figure 38 shows the frequency distribution for the Route 13 and DADS High Coastal Plain sample which is bimodal. About all that can be said about the elevation of sites based on the total data, is that either sites are very close to the estuarine sea-level water sources (the low elevations), or they are located on higher elevation bluffs or interior areas. There is no in-between.

Table 30 shows a series of descriptive statistics for site elevation for each time period. These values all indicate normal

FIGURE 37
Frequency Distribution – Water Distance

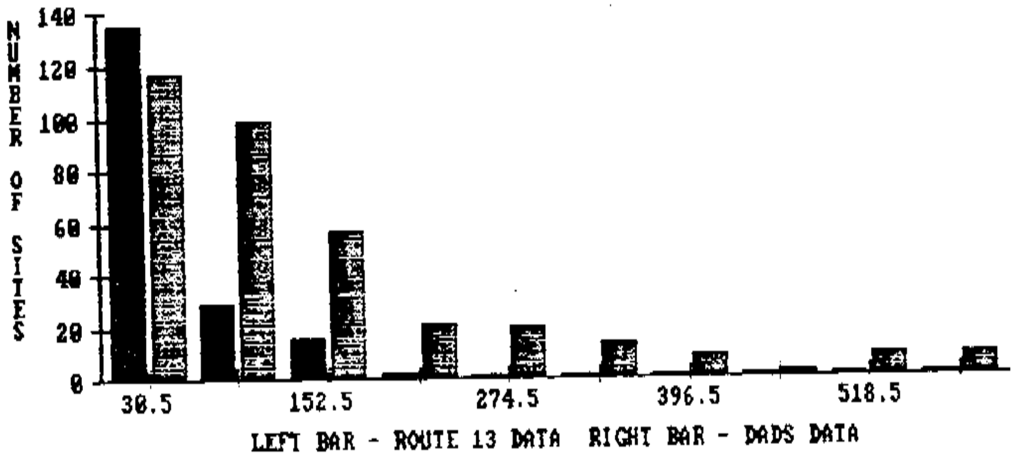


FIGURE 38
Frequency Distribution – Elevation

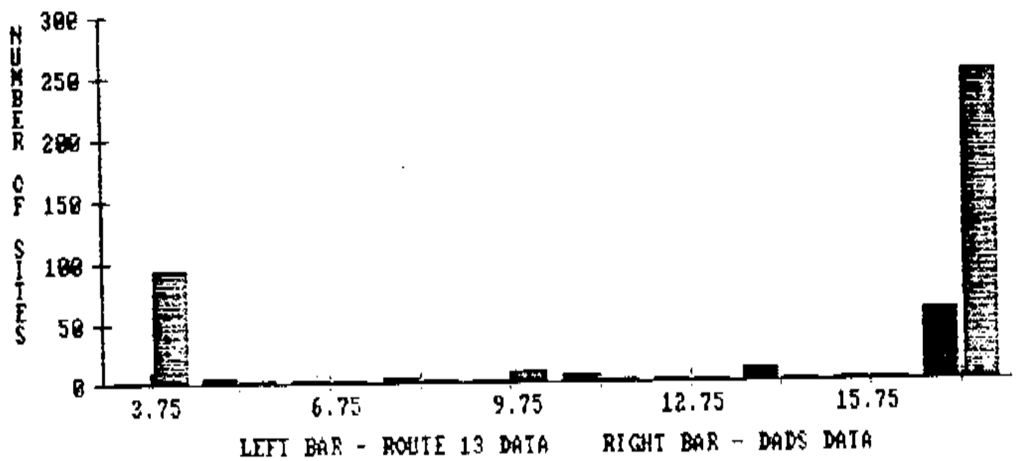


TABLE 30

ELEVATION DESCRIPTIVE STATISTICS BY TIME PERIOD

Time Period	N	Mean	Std. Dev.	Skew	Kurtosis	Test Q-Q	Q-Q(p<.10)
Archaic	7	13.57	3.21	.12	1.80	.96	.88
Woodland I	56	13.89	3.71	-.98	3.06	.93	.97
Woodland II	14	12.86	2.82	-.19	2.33	.96	.91

distributions. The very similar mean values and standard deviations all show no difference in site placement with regard to elevation through time.

Geomorphological Setting. The geomorphological setting of sites was also considered as part of the analysis of site location variables. Table 31 shows the frequencies of the

TABLE 31

GEOMORPHOLOGICAL SETTING FREQUENCIES

Setting	Route 13	DADS HCP	Total
Headland	138	112	250
Bay/Basin	112	2	114
Interior Flat	1	7	8
Sand Ridge	3	0	3
Floodplain	4	225	229
Upland Slope	0	12	12
	258	358	616

geomorphological setting of sites in the Route 13, DADS High Coastal Plain, and combined data bases. The previously noted focus of sites on major water courses, first, and bay/basin features, second, is supported by the high numbers of sites associated with headland, floodplain, and bay/basin geomorphological settings for all time periods of occupation for the Route 13, DADS High Coastal Plain, and combined data bases (Table 32). Given the low numbers of sites with datable components in the DADS High Coastal Plain data base, the Route 13 and combined data bases are the most interesting for analysis. In all but one case, there is a trend of increasing frequency of utilization of all geomorphological settings, for which there are sufficient site data for analysis, through time. The only exception is Woodland II bay/basin utilization which shows a

TABLE 32

CROSS-TABULATION OF GEOMORPHOLOGICAL SETTING AND TIME PERIODS

DADS High Coastal Plain Data

Period	Headland	Bay/Basin	Int. Flat	Sand Ridge	Flood	Up Sl
PI	0	0	1	0	2	0
Archaic	2	1	1	0	8	0
Wood. I	8	0	1	0	19	1
Wood. II	17	0	2	0	25	1
Route 13 Data						
Archaic	3	4	0	0	0	0
Wood. I	22	28	1	3	2	0
Wood. II	10	4	0	0	0	0
Combined Data						
PI	0	0	1	0	2	0
A	5	5	1	0	8	0
WI	30	28	2	3	21	1
WII	27	4	2	0	25	1

TABLE 33

CROSS-TABULATION OF GEOMORPHOLOGICAL SETTING AND SITE TYPE

DADS High Coastal Plain Data

Site Type	Headland	B/B	Int. Flat	Sand Rid.	Flood.
Procure.	5	1	1	0	4
Macro.	1	0	0	0	1
Micro	2	0	0	0	0
Base Camp	4	0	2	0	8
Route 13					
Procure.	74	76	1	1	0
Macro.	17	6	0	0	1
Micro.	6	10	0	2	1
Base Camp	4	6	0	0	0
Combined					
Procure.	79	77	2	1	4
Macro.	18	6	0	0	2
Micro.	8	10	0	2	1
Base Camp	8	6	2	0	8

FIGURE 39
Site Frequencies and Geomorphological Settings

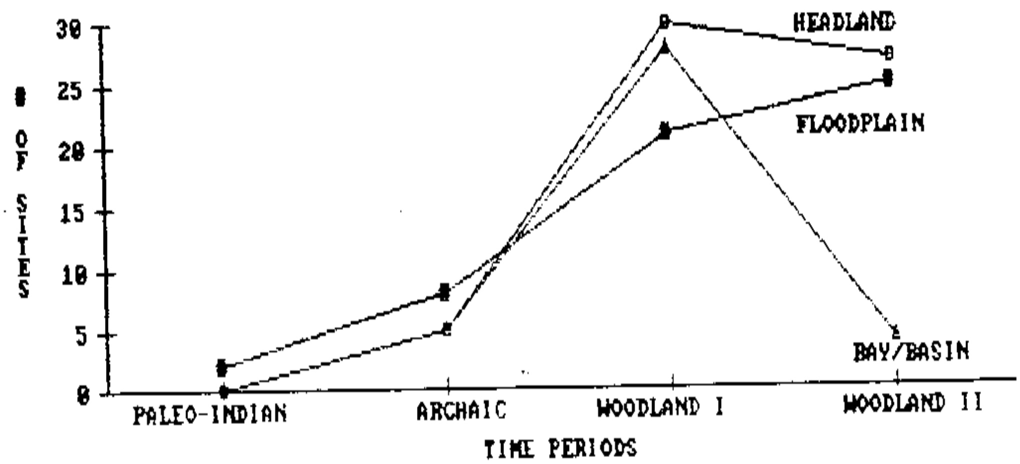
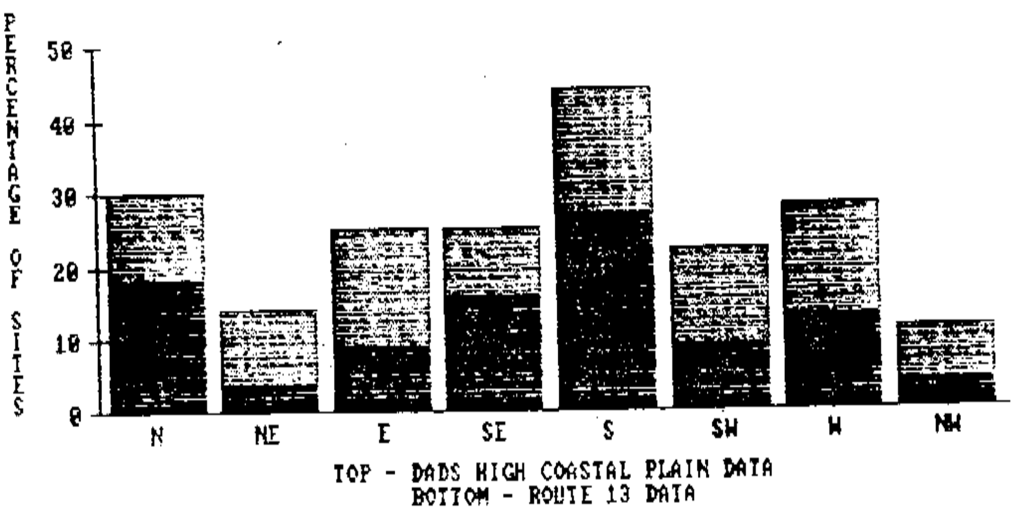


FIGURE 40
Aspect Frequencies



decline rather than an increase (see Figure 39). A similar pattern was noted during the analysis of surface water settings. Table 33 shows cross-tabulations of site types and geomorphological settings and again, given low numbers of sites with identifiable frequencies in the DADS High Coastal Plain data, the Route 13 and combined data bases are the most useful for analysis. Procurement sites show a strong association with bay/basin sites, as was noted in the analysis of surface water settings. Micro-band base camps also may be found in association with these features, while the larger base camps are more commonly associated with the major drainage floodplains and headlands.

Aspect. Site aspect was considered because several previous studies (Gardner 1978; 1982) noted this variable as a significant factor in prehistoric site locations. Table 34 shows the frequencies of sites for each of eight compass directions in the varied data bases and Figure 40 shows these same data in a bar graph. The southern exposure shows the greatest site frequency as would be expected from Gardner's models, but certainly not by the overwhelming percentage seen in other areas (Stewart 1980; Gardner and Boyer 1978; Wall 1981). Further analysis of site aspect included development of frequency distributions of aspect for varied time periods using the Route 13 data. There were insufficient data to analyze aspects of Archaic sites, but Woodland I and II data are noted in Figure 41. Neither the Woodland I or Woodland II data show any pronounced aspect preference.

TABLE 34

ASPECT FREQUENCIES

Aspect	Route 13	DADS High Coastal Plain	Total
N	42(.18)	42(12%)	84(14%)
NE	9(.04)	35(10%)	46(8%)
E	20(.09)	59(16%)	79(13%)
SE	36(.16)	31(9%)	67(11%)
S	63(.27)	61(17%)	124(21%)
SW	20(.09)	48(13%)	68(12%)
W	30(.13)	55(15%)	85(14%)
NW	12(.04)	27(7%)	39(6%)
	232	358	590

Frequency of varied aspects were also tabulated for functional site types. Figure 42 shows the data for macroband and microband base camps among the Route 13 data. The macroband base camps are primarily characterized by northern aspects. On the other hand, microband base camps are primarily characterized by

FIGURE 41
Woodland Aspect Frequencies

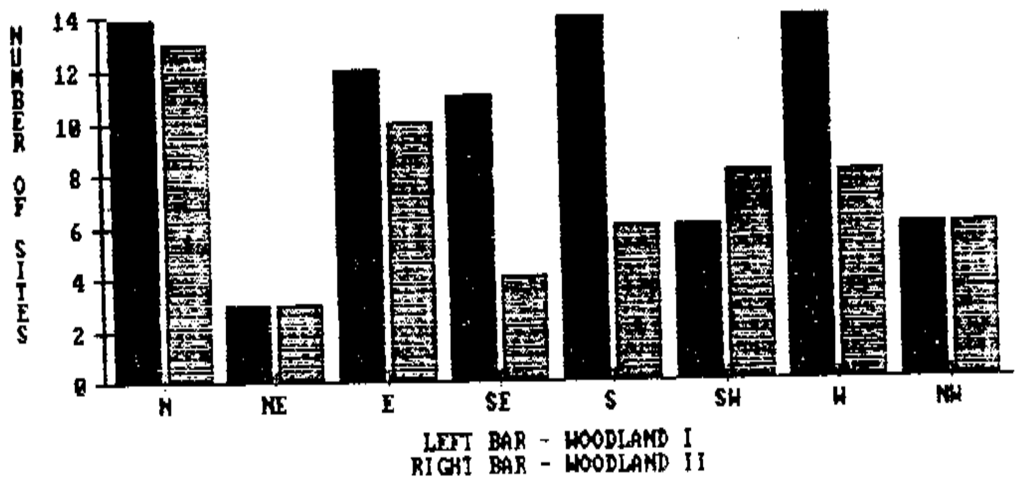


FIGURE 42
Base Camp Aspect Frequencies

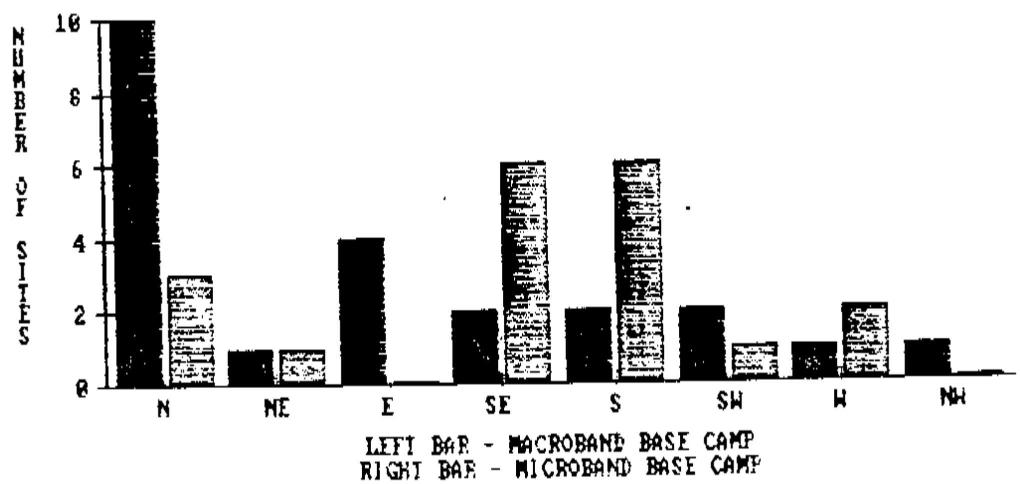
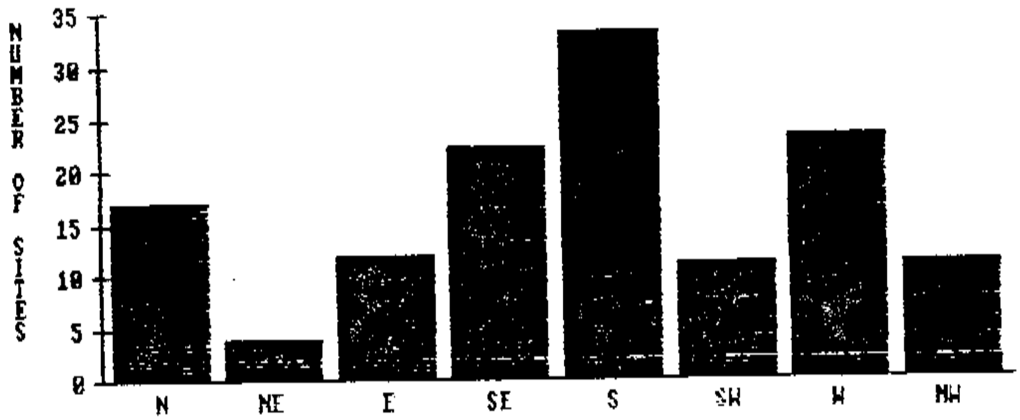


FIGURE 43
Procurement Site Aspect Frequencies



southern and southeastern aspects. Figure 43 shows the aspect frequencies for procurement sites. Southern exposures are most common, but not by a very large margin. It is interesting to note that these results are almost completely opposite compared to those reported by Steponaitis (1983a: Figs. 7 and 8) for the Patuxent drainage. The Patuxent data show that low density lithic scatter sites, here considered to be analagous to procurement sites, show no clear association with a southern aspect. Furthermore, other larger sites show a southern exposure. The differences between the Patuxent and Delmarva High Coastal Plain are <u>not</u> due to sampling differences and must be explained by other factors. Numerous researchers (Gardner and Stewart 1978; Wells 1981) have noted that large sites along the Appoquiniminik River tend to be located on its southern shore rather than the northern. The predictive model shows a similar pattern (Custer et al. 1984: Attachment V, 125). It is suggested here that the focus on the southern side of the drainage is due to the fact that most of the lower-order stream tributaries and well-drained headlands, which are the loci of the large base camp sites, are found on the southern side of the Appoquinimink and Blackbird drainages. Most likely, proximity to the tributaries and headlands was more important than site aspect for prehistoric groups. Furthermore, if house features were present at the larger Woodland I sites (see Custer 1984a:100-101 for examples), aspect would not be an important consideration for site location. In the Patuxent region, the drainage runs north and south so that the pattern of resource distributions and aspects is different. With regard to smaller procurement sites and lithic scatters, it is suggested here that the absence of any clear-cut aspect preference in the Patuxent area is due to the very ephemeral nature of these sites' utilization. On the other hand, the Route 13 data, with its large number of bay/basin related sites and preference for southern aspects, may indicate a less ephemeral utilization of procurement sites. The critical factor in the differences in procurement site utilization between the Patuxent and Delmarva High Coastal Plain is the presence of the large number of bay/basin features in the Delmarva area. These environmental settings may be so productive that their use is more protracted and intensive than the use of interior areas without bay/basin features such as the Patuxent.

Soil Series. The final site location variable considered was soil setting. Fourteen separate soil series were associated with sites in the Route 13 area making cross-tabulations difficult. DADS High Coastal Plain data were not used because an even greater variety of soil settings were present. Table 35 provides a summary of the soil frequencies. Four series (Matapeake, Woodstown, Fallsington, Rumford) account for 43% of the sites. All of these soil series, except for the Fallsington

TABLE 35

SOIL SERIES FREQUENCY

	Soil Series				
	Matapeake	Woodstown	Fallsington	Rumford	Other
#	60	10	26	15	151
%	23%	4%	10%	6%	57%

series, are well-drained (Matthews and Lavoie 1970). The Fallsington series is poorly drained and found on upland flats (Matthews and Lavoie 1970:22). The association of sites with Fallsington soils generally occurs in the bay/basin settings and usually the sites themselves are located on small, well-drained sandy and/or silty knolls within the zones mapped as part of the Fallsington series. Thus, the most that can be said about associations of sites and soil series is that the sites tend to be on well-drained soils and the mapped USDA soil series are not necessarily good indicators of where well-drained soils may be located. As a sidelight it should be noted that the LANDSAT-generated soil data used in the logistical regression model (Eveleigh 1984: 78-81) are more useful than the USDA data for understanding site distributions.

In sum, it can be noted that analysis of site location data show some common features for all site functions and time periods while for others, there is diachronic and synchronic function variation. For the most part, the analyses presented here support existing models. However, consideration of local edaphic factors requires some modification of the models.

FACTOR ANALYSIS - LOCATIONAL VARIABLES

A factor analysis of locational variables was undertaken to see how the various measured variables overlapped. Factor analyses use variable correlation matrices and extract large scale variables, or factors, with which the original variables may or may not be correlated (Rummel 1970:13-19). Correlations of the original variables with the factors show the degree to which the original variables overlap and also may have intuitive interpretive meaning (Rummel 1970:19-21). A factor analysis of the Route 13 data was undertaken because factor analyses of site location variables have been undertaken in the Middle Atlantic region (Hughes and Weissman 1982) and have yielded interesting results.

Results of factor analyses can be presented in numerous ways and for the purpose of simplification the zero-one cluster-centroid transformation matrices (Overall and Klett 1972:134-136) are presented for each factor analysis (Table 36). Six separate factor analyses were undertaken; one each for quantitative site

TABLE 36

ZERO-ONE CLUSTER-CENTROID TRANSFORMATION MATRICES

	Matrix I		Matrix II		Matrix III	
	Fac.1	Fac.2	Fac.1	Fac.2	Fac.1	Fac.2
water distance	0	1	1	0	0	1
slope	1	0	1	0	1	0
elevation	1	0	0	1	1	0

	Matrix IV		Matrix V		Matrix VI	
	Fac.1	Fac.2	Fac.1	Fac.2	Fac.1	Fac.2
water distance	0	1	0	1	0	1
slope	0	1	0	1	0	1
elevation	1	0	1	0	1	0
geomorph. set.	1	0	1	0	1	0
water type	1	0	1	0	1	0
aspect	1	0	1	0	1	0

Matrix I - all sites - quantitative variables
Matrix II - Bay/Basin sites - quantitative variables
Matrix III - Non-Bay/Basin sites - quantitative variables
Matrix IV - all sites - all variables
Matrix V - Bay/Basin sites - all variables
Matrix VI - Non-Bay/Basin sites - all variables

location variables for all sites, and non-bay/basin sites, and one each for quantitative and qualitative site location variables for all sites, bay/basin sites, and non-bay/basin sites. Comparison of matrices IV,V and VI in Table 36 shows that all three are identical. One factor correlates with water distance and slope while the second factor correlates with elevation, geomorphological setting, water type, and aspect. The first factor may be interpreted as a measure of proximity to major water sources. Slope is included because the lower slopes in the project area are found on headlands close to water while the more rolling topography is found in the bay/basin and stream tributary area away from the major drainages. The second factor may be interpreted as a general type of surface water/geomorphological setting indicator. The correlation of these variables is understandable given the previous analysis of locating variables. There is considerable overlap of the surface water type and geomorphological setting variables as well as aspect and elevation. Given the nature of these variables and their overlap, it is most likely that the second factor primarily measures the distinction between major drainage and bay/basin settings.

The analysis of strictly quantitative site location variables (Table 36; Matrices I,II, and III) shows more internal variability and also differs from the combined quantitative and qualitative results. The combined site data and the non-bay/basin data (Matrices I and II) show two factors: one correlated with water distance and a second correlated with slope and elevation. These two factors may be interpreted as being similar to those noted above for the combined quantitative and qualitative variables. However, the second factor in the analysis of quantitative sites is probably not as good an indicator of site setting as is the second factor in the combined quantitative and qualitative analysis.

The factor analysis of quantitative variables for bay/basin sites provides results which are different from the other factor analyses of quantitative data. In the bay/basin case, water distance and slope are correlated with one factor and elevation alone is correlated with the second factor. These results are more similar to the factor loadings seen in matrices IV,V, and VI, which considered both qualitative and quantitative variables. The bay/basin factor results may also indicate that slope is more closely related to water distance in the interior bay/basin settings than it is along the major drainages.

The results of the factor analyses presented here may be compared to the results reported by Hughes and Weissman (1982) using site data from Western Maryland. Hughes and Weissman considered five variables (proximity to water, slope, height above water, distance to stream confluence, and soil drainage) which mixed quantitative and qualitative data. Consequently, their results are best compared to Matrix IV in Table 36. Two factors were isolated using the Western Maryland data. The first factor consisted of proximity to water, height above water, and distance to confluence. The second factor consisted of slope and soil drainage. The composition of these factors is similar to the composition of the Route 13 factors, although some differences are present due to the use of different original variables. Nonetheless, both of the factor analyses show a factor related to surface water, which accounts for much of the variance in site locations (up to 60% in the Route 13 data and up to 78% in the Western Maryland data), and an additional factor, related to landforms, which accounts for some of the remaining site location variance. These factor analyses nicely match Well's (1981:41-43) analyses of site location variables which noted that surface water settings accounted for the greatest proportion of site location variance.

SITE AREA ANALYSIS

Site area was considered earlier as part of the evaluation of the predictive model; however, it will be considered again with regard to changes in site size through time. This variable's changes through time will be considered because many of the current settlement pattern models (Custer 1984a:94-97) note significant changes in site size through time. Also, comparable data on changes in site size through time are available from the Patuxent drainage (Steponaitis 1980, 1983a; 1983b; 1984).

Table 37 shows a series of descriptive statistics for site area during the Archaic, Woodland I, and Woodland II periods. The values for the Q-Q statistics indicate that only the Woodland II values are normally distributed. Therefore, traditional parametric statistical comparisons could not be used. Table 38

TABLE 37

DESCRIPTIVE STATISTICS - SITE AREA BY TIME PERIODS

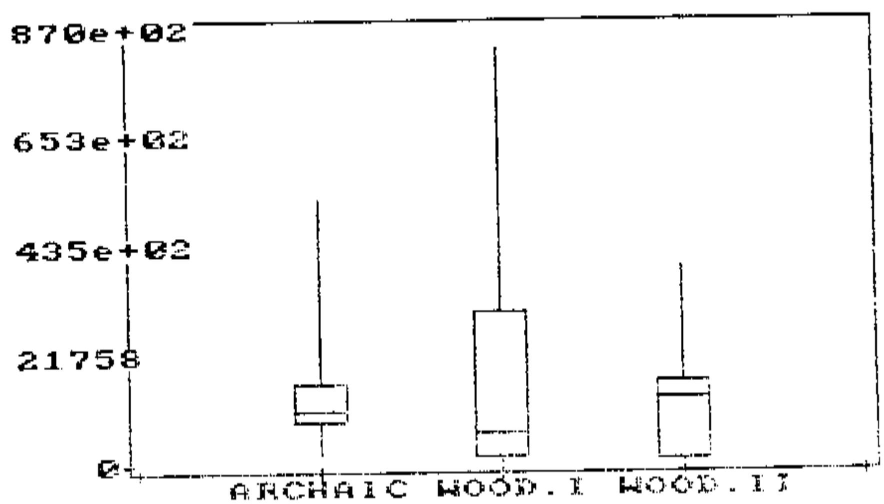
Time Period	N	Mean	Std.Dev.	Skew	Kurtosis	Test Q-Q	Q-Q p<.10
Archaic	7	16436.30	16786.20	1.66	4.37	.85	.88
Woodland I	42	17981.10	22537.30	1.64	5.05	.87	.96
Woodland II	6	14000.86	13943.80	.91	2.85	.92	.88

TABLE 38

NON-PARAMETRIC RANGE STATISTICS - SITE AREA BY TIME PERIODS

Time Period	N	Min	25th%	Med	75th%	Max	Total Range	75th% Max Range
Archaic	7	2323	7200	10684	18901	52721	50398	33820
Woodland I	42	523	1481	6631	28509	82914	82391	54405
Woodland II	6	697	980	13413	15242	39018	38321	23776

FIGURE 44
Site Area Box Plots by Time Periods



shows a variety of nonparametric range statistics for site area during each time period. It can be seen that the largest range in site size occurs during the Woodland I time period, although all time periods show a large range of site sizes. Nevertheless, the Woodland I range is 56% larger than the Archaic range, which is the next largest. Comparison of the 75th percentile-maximum value range also shows the Woodland I range to be the greatest.

Figure 44 is a box plot of the Archaic, Woodland I, and Woodland II data which shows the median values as a horizontal line within a box whose height reflects the range of values around the median which includes 50% of the values (Velleman and Hoaglin 1981:65-70). The length of the lines extending from the boxes reflects the total range of the values. The box plot clearly shows that the largest range of values is found in the Woodland I period. Not only is the box larger, but it includes a larger range of values. Also, the total range of Woodland I values includes more larger values than those of the other two time periods. These data indicate that there are more larger sites present during the Woodland I time period compared to all other time periods. Previous studies (Custer 1984a:77, 92-94; 1984c) have noted similar trends based on more impressionistic data and the Route 13 data support this contention.

Even more interesting is the reduction in site size moving from the Woodland I into the Woodland II Period. In most other areas of the Middle Atlantic region, including the southern Delmarva Low Coastal Plain, site size increases moving into the Woodland II Period (Custer 1984a:157-171; Custer and Griffith n.d.). However, the northern Delmarva region shows a decrease in social complexity, sedentism, and settlement pattern diversity

(Stewart et al. n.d.; Custer 1984a:155-157). The High Coastal Plain portion of the Route 13 study area falls within the Minguannan Complex boundary and the reduction in site size fits with known late prehistoric settlement pattern trends (Stewart et al. n.d.).

The Route 13 data on site size can be compared to the Patuxent data. Steponaitis (1983a: Fig.7) used box plots to show site size changes through time and found an increase in site size starting in the Late Archaic Period (early Woodland I) moving through to the Late Woodland (Woodland II) period. It should be noted that the criteria used by Steponaitis to separate Late Archaic, Early Woodland, and Middle Woodland components are suspect (see discussion in Custer 1983c). Nonetheless, even if the Late Archaic through Middle Woodland data for the Patuxent are combined, the Late Woodland Patuxent data still shows more larger sites. Therefore, the comparison of the Route 13 and Patuxent data reinforces the previous results of comparisons of the Upper Delmarva late prehistoric cultures to other similarly dated cultures of the Middle Atlantic region.

A related topic which may have been considered was the analysis of site frequencies through time. However, there were insufficient data in the Route 13 sample with identifiable components to make such an analysis useful.

BAY/BASIN UTILIZATION

Human utilization of bay/basin features has already been discussed in numerous contexts in this report. However, this topic is sufficiently important to merit its own special discussion. The Route 13 survey, especially the survey of the Blackbird area, provided one of the most comprehensive surveys of bay/basin sites in the Middle Atlantic region. The stratified random sample with subsurface testing also enhanced the research value of the survey sample.

Bay/basin features have been recognized as important loci of archaeological sites for a number of years. The first consideration of their role in prehistoric settlement subsistence systems was provided by Bonfiglio and Cresson (1978) who studied sites associated with bay/basin features in New Jersey. Ever since then their importance has been recognized in numerous overviews especially for the pre-3000 B.C. time period (Custer 1984a; 58-59, 72; Kraft and Mounier 1982). Although most researchers agree about the importance of these sites for prehistoric peoples, there is not complete agreement as to their origin. Bonfiglio and Cresson (1978) feel that these features are of periglacial origin and refer to them as "pingoes". However, it is not clear that bay/basin features are truly periglacial features. For one thing, bay/basin features are found not only in New Jersey, where they may be found within 50m of the Pleistocene ice front, but they are also found as far south as the South Carolina Coastal Plain where they are referred to as "Carolina Bays". It is very unlikely that these more southern features, which appear to be similar in morphology to the New Jersey features (Rasmussen 1958), were formed under periglacial conditions with frozen soil and ground water. Furthermore, in a comprehensive study of bay/basins features in central and southern Delaware, Rasmussen (1958) rejected the hypothesis that they had a periglacial origin. However, Rasmussen was unable to effectively explain their origin by other geomorphological processes.

The current research on bay/basins in the Route 13 provided data on both human utilization of bay/basin features and their geomorphology. The various data gathered on the bay/basin feature's geomorphology shows that their size and shape have changed dramatically over the course of the Late Pleistocene and the Holocene. Sedimentary data from the bay/basin features studied here do not support the hypothesis that these features have a periglacial origin. If anything, the bay/basin sediments are more similar to those seen in sinkhole settings from the northern Delaware area (Custer and Griffith 1984). Probably the most that can be said at present is that bay/basin features are open waterfilled depressions that were acting as sediment and pollen traps at least since the end of the Pleistocene and through the Holocene. During this time they may have grown and contracted in size in response to local climatic changes. The most important point is that, no matter what their origin, these features were sources of fresh water which were very attractive for the game animals hunted by the prehistoric inhabitants of the Delmarva Coastal Plain throughout the Holocene.

The Route 13 survey showed some interesting patterns in human utilization of these bay/basin features. Probably the most interesting observation to make concerns the extent to which these features were used. A total of 148 bay/basin features were studied by either surface reconnaissance or subsurface testing. Of these, 128 features (90%) had associated archaeological sites. Diagnostic artifacts were found at 49 of these sites. Of these 49 sites, 5 have Archaic components, 38 have Woodland I components, and 5 have Woodland II components.

It is important to note that no Paleo-Indian components were discovered in association with these features during the Route 13 survey. In contrast, Bonfiglio and Cresson (1978:18) note that of 94 bay/basin features, 7 (7%) contained fluted point components in the New Jersey sample. There is no general shortage of Paleo-Indian fluted point sites in the upper Delmarva Peninsula and quite often these sites are associated with poorly drained, swampy settings (Custer 1984a:48-60; 1984d; Custer, Cavallo and Stewart 1983; Custer, Catts, and Bachman 1982). However, the known fluted point sites of the Delmarva Coastal Plain are associated with freshwater interior swamps fringing drainages within the mid-peninsular drainage divide, not bay/basin features. There are several explanations of this difference in Paleo-Indian utilization of bay/basin features between the Delmarva and New Jersey Coastal Plain. A simple explanation may be that bay/basin features are not large enough during the late Pleistocene in the Delmarva to be attractive hunting locales. This explanation could easily be tested by obtaining a series of radiocarbon dates on bay/basin sediments in a variety of settings where there are associated archaeological sites of different ages. A second explanation may be based on sampling factors. Although the Delmarva sample is large in number and, due to its random component, less biased than the New Jersey data, the New Jersey data come from a wider geographic area than does the Delmarva sample which is drawn from a smaller, more concentrated area. If a sampling bias is involved, additional survey of bay/basin features in other areas of the Delmarva Peninsula should reveal associated Paleo-Indian sites.

Assuming that there are no sampling errors and that bay/basin features are present in the Delmarva Peninsula throughout the Late Pleistocene and early Holocene, additional behaviorial explanations of the differences between the New Jersey and Delmarva data are necessary. The Paleo-Indian utilization of bay/basin features in New Jersey may be due to the fact that the bay/basin features of the New Jersey High (or Inner) Coastal Plain are often associated with either the cuesta, or other concentrations of secondary lithic resources (Cavallo 1981; Marshall 1982:24,32). Custer, Cavallo, and Stewart (1983) and Gardner (1974;1977) have noted the important role of lithic resource locations in Paleo-Indian settlement patterns and the juxtaposition of the lithic resources and game-attractive hunting locales may have made the New Jersey bay/basin settings very attractive settlement locations during Paleo-Indian times. No similar juxtaposition of resources is seen in the Delmarva region (Custer and Galasso 1980; Custer 1984a:59) and this may be why there was little or no Paleo-Indian utilization of these features in Delaware. This explanation could be tested by looking for bay/basin locations in Delaware that may have hitherto unknown associated lithic sources. These features should have some signs of Paleo-Indian utilization if the above explanation is correct. An alternative test would be to consider bay/basin locations in New Jersey which do <u>not</u> have associated lithic sources. These sites should not have associated Paleo-Indian occupations.

Utilization of bay/basin features in the study area seems to begin early in the Holocene. There are five bay/basins associated with sites with bifurcate-base points, which are the only really reliable indicators of the Archaic period (Custer 1984a:61-62). The presence of a Kirk-like point at one of these bay/basin sites may indicate that the utilization of these sites began quite late in the Paleo-Indian Period. Generally, the sites seem to be small, ephemerally utilized hunting/processing sites. Five sites may not seem like a large number; however, prior to the Route 13 survey, only 79 sites with bifurcate points are listed in DADS (Custer n.d.). Of these, only 12 were located in the High Coastal Plain and only 7 Archaic sites in the entire DADS, including adjacent areas of southeastern Pennsylvania, are associated with bay/basin features. The Archaic bay/basin sites from the Route 13 survey are, therefore, an important addition to the Archaic site data base.

Because the Archaic Period of Delmarva prehistory is so poorly known, it is difficult to assess the meaning of the Archaic bay/basin sites. However, some observations can be made. The beginning of bay/basin utilization seems to occur at the same time as a series of rather dramatic environmental changes. During the period from 8500-6000 BC there is evidence from numerous sites indicating dry climatic conditions (Custer 1984a:47-48; Custer and Griffith 1984; Carbone et al. 1982). Environments seemed to have changed from a mosaic of grasslands, swamps, boreal forests, and deciduous forests to a closed boreal forest with fewer poorly drained settings. The presence of wind-blown sediments (Foss et al. 1978) and evidence pronounced changes in stream channel morphology (Custer and Griffith 1984: Fig. 5) also indicate potential dramatic changes in the patterns of surface water availability. The beginnings of bay/basin utilization may be related to these environmental changes. It is possible that changes in stream channel morphology altered the distribution of swampy settings in the mid-peninsular drainage divide, as evidenced at the Dill Farm Site (Custer and Griffith 1984), and caused late Paleo-Indian and Archaic groups to seek out new swampy hunting stations, such as the bay/basin features. Another factor which may have been contributed to a shift to new procurement sites locations, including bay/basins, during the Archaic Period is the fact that during late Paleo-Indian and Archaic times the emphasis on high grade cryptocrystalline lithic materials seems to have disappeared (Custer 1984a:59-60). If association of bay/basins and lithic sources was no longer a critical factor in site selection, then the bay/basin sites of the study area may have become a more attractive settlement option. Once these bay/basin procurement sites became part of the settlement pattern in interior areas, their utilization continued into warm-wet climatic conditions of the post-6000 B.C. time period (Custer 1984a:62-64).

The time period of most intensive bay/basin utilization occured in the Woodland I period, between 3000 B.C. and A.D. 1000. During the beginning portions of this time period, and possibly during later periods as well, the Delmarva Peninsula and Middle Atlantic region, in general, experience the warmest and dryest climatic conditions of the entire Holocene (Custer 1984c). The mesic forests of the Archaic time period are replaced by open xeric oak-hickory woodlands and grasslands (Custer 1984a:89-91) and very dramatic changes in surface water availability occurred (Curry and Custer 1982). One of the major settlement pattern changes seen in the coastal plain area is the utilization of a wide variety of interior environmental settings on an ephemeral basis (Custer and Galasso 1983:12-14). The increase in bay/basin utilization during this time period is part of this trend. The data from 7NC-H-20 indicate that in areas of multiple bay/basin clusters there may also be more permanent sites dating to the Woodland I Period. These more permanent interior sites are probably similar to the Hawthorn Site (Custer and Bachman 1984).

Utilization of bay/basin sites dropped moving into the Woodland II Period. Amelioration of the climatic stress of the Woodland I Period may have contributed to the reduced bay/basin utilization.

In conclusion, bay/basin features are an integral part of prehistoric human adaptations in the Delmarva High Coastal Plain for varied time periods. The study of the large number of bay/basin features in the Route 13 project area has the potential to add appreciably to our knowledge of prehistoric human adaptations.

DISCUSSION OF GENERAL PREHISTORIC SURVEY RESULTS

In addition to contributing to the development of settlement models for bay/basin areas, the Route 13 survey data contributes to the development of general settlement subsistence models for the High Coastal Plain. A summary discussion of the total survey results is presented below for each time period.

Paleo-Indian Period. No Paleo-Indian sites were discovered in the Route 13 survey. As was noted in the discussion of the utilization of the bay/basin features, the absence of Paleo-Indian Period sites in the specific study areas may be due to the absence of flowing surface water associated with bay/basin features, or may be due to the absence of an association of high quality lithic sources and bay/basin features. Nevertheless, it is somewhat surprising that no Paleo-Indian materials were recovered from the large areas surveyed along the Appoquinimink River. Although the Appoquinimink River drainage is outside the known, and predicted, concentrations of Paleo-Indian sites (Custer 1983a:38-47), the occurrence of some kind of small procurement sites, such as 7NC-D-70 (Custer, Catts, and Bachman 1982), associated with a springhead or small swamp would be expected. It could be possible that older sites have been destroyed on the heavily eroded landscapes of the Appoquinimink. Or, Paleo-Indian site densities outside of the known and predicted concentration zones are truly quite low.

Archaic Period. The importance of the Archaic bay/basin sites has already been discussed; however, it should be noted that 3 Archaic sites associated with stream floodplain sites were also discovered. Only 9 Archaic sites were previously recorded in DADS for the Delmarva High Coastal Plain. Therefore, the 8 Archaic sites discovered in this survey almost double the number of known Archaic sites in the High Coastal Plain. In all of Delaware there are only 40 Archaic sites recorded in DADS, making the additional 8 sites discovered here a 20% increase in the Archaic site data base.

The Archaic sites associated with stream settings seem to be similar to others described for the Delmarva Coastal Plain (Wise 1983; Kavanagh 1979; Custer and Galasso 1983; Galasso 1983) and are primarily small procurement sites. These sites probably represent hunting and procurement sites which support other base camp sites. Some of the larger base camp sites have been tenatively identified elsewhere in Delaware (Custer 1984a:69-72): however, none were identified in this survey. It may be possible that there are no large Archaic base camps in the Coastal Plain areas away from the large interior swamps. Some of the Archaic sites found in this survey may be small base camps rather than procurement sites and the present survey methods were unable to distinguish the differences between the two site types. Both Wise (1983) and Galasso (1983) have suggested that the Delaware Coastal Plain Archaic settlement pattern is characterized by small habitation and procurement sites and Kraft and Mounier (1982) note similar patterns in the New Jersey Coastal Plain. The data on Archaic sites from the Route 13 survey seem to support this model.

Woodland I Period. The Woodland I Period sites of the study are the largest and most numerous of all time periods. The analysis of site size showed that not only are most of the Woodland I sites larger than sites from other time periods, but several very large Woodland I sites were present (Figure 44). These very large sites are identified here as macroband base camps and are located primarily along the Appoquinimink River, the highest order stream in the study area. Current models of Woodland I settlement patterns and adaptations (Custer 1982; 1984a:94-98; 1984c; Catlin et al. 1982) all note a shift of large base camp sites to major drainage floodplain and headland settings and a general increase in local population densities in these areas during Woodland I times. The Route 13 survey data support this model. Similarly, it was noted in the discussion of bay/basin sites that the Route 13 Woodland I site data supported those site models which noted a widespread, but ephemeral use of interior areas. Interior Woodland I sites other than those associated with bay/basin features from the Route 13 data also support this model and the distribution of sites in the interior areas is quite similar to that noted for the Upper Chester drainage in Kent and Queen Annes Counties Maryland (Kavanagh 1979).

Non-local lithic materials, such as rhyolite, argillite, steatite, and ironstone, are present at many of the sites recorded in the Route 13 survey as was noted in the site descriptions. These non-local materials tend to be found at the larger Woodland I base camp sites. The presence of these "exotic" materials in the study area indicates that local Woodland I groups were participating in trade and exchange networks as noted in several studies (Ward and Doms 1984; Custer 1984e). Participation in trade and exchange networks at the larger Woodland I sites indicates increasing social complexity at these sites.

It would be useful to discuss the site locations and assemblage characteristics at a time level smaller than the period, such as the archaeological complexes which are used to divide the Woodland I period in terms of time and space (Custer 1984a:28-30,78,89). However, there are insufficient data on diagnostic artifacts from the Route 13 survey to develop any counts of sites at the level of the archaeological complex. Nevertheless, one observation that can be made is that few, if any, Woodland I ceramics were recovered during the Route 13 survey, even from the very large Woodland I sites. Although ceramics are not often preserved at highly eroded sites, it does not seem unreasonable to expect that there would be a few ceramics present among the approximately 2400 artifacts found at Woodland I sites during the Route 13 survey. If the absence of Woodland I ceramics is not due to sampling errors or differential artifact preservation, then it is possible that the bulk of the Woodland I settlement in the study area predates 1200 B.C. Some large Woodland I sites post-dating 1200 B.C., such as the Hell Island Site (Thomas 1966), are present, nonetheless.

Woodland II Period. Woodland II settlement patterns in central Delaware are a topic of some controversy. For many years, numerous authors have suggested that there is a relative absence of Woodland II sites in southern New Castle County and northern Kent County. By the same token, up until 1980 the nature of the northern New Castle Woodland II occupations were also very poorly refined. Nonetheless, the southern New Castle County and northern Kent County area was viewed as a "buffer zone" or "fever belt" (Withoft 1984) separating two distinctive ethnic groups. The original Route 13 planning study analyzed extant artifact collections and noted numerous Woodland II sites in the supposed "buffer zone" making the whole concept somewhat invalid (Custer et al. 1984:220-221). The "discovery" of these sites was due to the fact that previous analyses had not recognized the Woodland II Minguannan ceramics in the collections because the type was not defined in the literature prior to 1981 (Custer 1981).

The discovery of Woodland II sites in the present survey reveals a similar bias in previous studies which caused Woodland II sites to be under-represented in the data base. Most of the Woodland II sites, and all of the sites with Minguannan pottery, were discovered during sub-surface testing of wooded areas dividing plowed fields from bluffs along the major drainages. The sites are small and appear along most of the major stream headlands studied. Furthermore, they are almost all unplowed and would have been, and were, missed in previous studies which focused primarily on surface survey of cultivated fields. Thus, there really is no absence of Woodland II sites in the study area and there is no need to invent a "buffer zone".

It can be noted that Woodland II sites in the study area are generally smaller than the Woodland II sites found farther south on the Delmarva Peninsula (Custer 1984a:157-171; Custer and Griffith n.d.). However, the Woodland II sites of the study area fall well within the range of site sizes seen among Woodland II sites of the Minguannan Complex (Custer 1984a:155-157; Stewart et al. n.d.).

In conclusion, analysis of prehistoric site survey data from the Route 13 project area reveals that existing settlement-subsistence models are generally accurate although some adjustments are necessary. Also, the analysis shows that the logistical regression predictive model accurately predicts locations of various types of sites: however, the models could be improved by the development of more accurate remote sensing of bay/basin features. Most importantly, the survey data open several questions which can be answered by future research. These potential avenues for future research will be discussed in the management section of this report.