A STOCHASTIC MODEL FOR NEARSHORE COASTAL PROCESSES

Thesis for the Degree of M. S.

MICHIGAN STATE UNIVERSITY

LOUIS ALLEN ORLOWSKI

1974

ABSTRACT

A STOCHASTIC MODEL FOR NEARSHORE COASTAL PROCESSES

By

Louis Allen Orlowski

Quantification of the changes in nearshore topography has proven difficult due to rapid, short-term fluctuations of bottom features. Structuring topographic transitions into three states, (1) no significant deflections from uniform slope, (2) positive deflections, and, (3) negative deflections, and analyzing state succession through time and space as a Markov chain allows for the description of the evolution of nearshore coastal features. Data from the eastern shore of southern Lake Michigan (Davis and Fox, 1971) indicate that such an approach is feasible. Five clusters were defined, within which the topographic response of the inner nearshore functions as a first order Markov chain. Associated transition probabilities describe process function through time and space within the environments of the inner nearshore; bar, trough, subaqueous terrace, and swash zone. Good correlation is obtained between the Markov chain model and empirical interpretation of nearshore topographic fluctuations. Stochastic process

models can serve as an accurate technique for the description of coastal processes.

A STOCHASTIC MODEL FOR NEARSHORE COASTAL PROCESSES

By

Louis Allen Orlowski

A THESIS

Submitted to

Michigan State University

in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

Department of Geology

1974

650705

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Fluctuations in bottom topography within the nearshore zone are rapid and ephemeral. The types and sequences of topographic transitions are known from observational data. Quantification of such empirical data has proven difficult due to the aforementioned variability and the multiplicity of causative factors. These factors, which are both deterministic and probabilistic in nature, can invoke similar responses in a natural system by any number of process routes. This has limited the effectiveness of deterministic modeling of the nearshore environment.

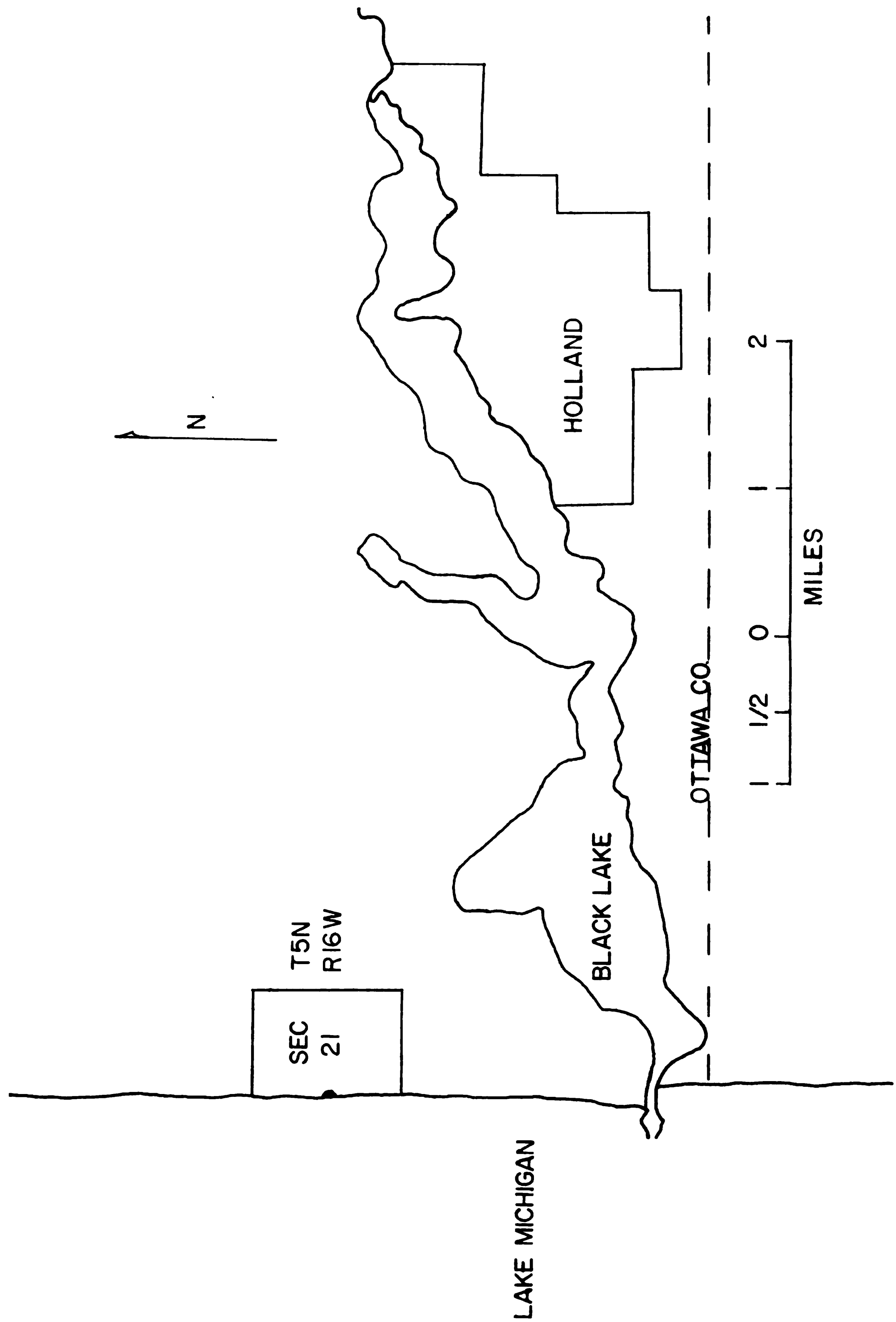
A stochastic process model can describe these topographic variations on a probability basis. Markov chains, a stochastic process, structure the response of the system as a finite number of states and evaluate the probability of state succession through time. By identifying those areas that have similar probabilities of topographic transition, areas os similar process response can be delineated and investigated.

Such an analysis of the nearshore zone from eastern Lake Michigan has been undertaken for this study. The topographic data were obtained from a sequence of time distance maps assembled by Davis and Fox (1971) for the

Office of Naval Research. A thirty day time series study of the nearshore environment was conducted 2.2 miles north of the Black River at Holland, Michigan: Ottawa County, T5N, R16W, Section 21 (Figure 1). At this locality nine ranges 100 feet apart were surveyed along the north-south trending shoreline. Along each range, bottom elevation was measured at ten stations at 12.5 foot intervals from 50.0 feet, outward to 162.5 feet. Contouring was based upon a 0.5 ± 0.5 foot contour interval and reflected depth of water or bottom elevation with respect to stilled water level.

Relating topographic variations in the nearshore zone, at Holland, Michigan, to Markov chains has resulted in the outlining of five clusters of similar process response. Analyses of the probabilities of state occurrence, up, down, and uniform slope, and of the internal chain structure within these clusters have indicated a good correlation between the stochastic process model and the known sequence of events occurring within the nearshore zone. This study indicates that Markov chains are an accurate model for the description and simulation of nearshore processes.

Figure 1 - Locality map of study area.


N
HOLLAND
BLACK LAKE
OTTAWA CO.
T5N
R16W
SEC
21
LAKE MICHIGAN
MILES
2
1
0
1/2
1

STOCHASTIC PROCESSES

Natural phenomena can be approximated through modeling. Applicable mathematical models can be classed as deterministic, random, and/or stochastic.

In a deterministic model the state of the system at any point in time or space can be exactly predicted from knowledge of the functional relations specified by applicable differential equations. The variability and complexity of geologic systems and the restriction to sampling of available populations limits the accurate description of the true geologic population by deterministic parameters.

A random model is characterized by the independence of the state of the system at any point in time or space from any other point. Preceding events have no effect upon the present state of the system. Occurrence of a state is based upon fixed probabilities determined by the proportion of occurrence within the sample. That is, the long term occurrence of a state in the model cannot exceed that of the initial input data taken from the natural system. Population estimators are biased and predictions inaccurate.

Stochastic models, of which Markov chains are a type, are intermediate between deterministic and random. Most natural processes are stochastic, exhibiting a predictable random behavior. That is, deterministic and probabilistic elements of a process produce a response, the occurrence of which is random. This random occurrence in stochastic processes is controlled by probabilistic mechanisms which allow statistically for the complete description of the phenomena by its probability distribution. A stochastic model specifies the complete joint probability distribution of the observations at each point of time. Viewed as a continuous development in time, the natural process can be termed stochastic.

Mathematically, a stochastic process can be described in the following manner:

$$X_t, \; t \in T \qquad (1)$$

where

X_t = a family of random variables,

t = the index time, and

T = the index set of the process, the population.

For each t in the population T, there exists a random variable X_t. With the index t taken as time, X_t denotes the state of the process at time t. When T is countable, the process is a discrete-time process. All possible values of X_t define the state space of the process. When X_t is countable it is a discrete-state process. Thus, a stochastic process consists of a family of random variables that describes the evolution through time and space of some natural process.

A discrete-time, discrete-state Markov chain is the particular stochastic process investigated in this study. A Markov chain is a sequence of states, the order of occurrence of which is determined by the transition probabilities associated with the immediately preceding state. A process so structured has a memory relationship termed first-order. Mathematically, it can be represented as:

$$P(X_{t+1} = j/X_t = i) = P_{ij} \qquad (2)$$

where

P = probability,

ij = the ij^{th} cell in a matrix, Figure 4,

X_{t+1} = an observation at time $t + 1$,

X_t = an observation at time t,

i, j = states of the system, and

P_{ij} = the probability associated with an ij transition.

That is, the probability that the observation X at time t + 1 will be equal to the state j, given that at time t the observation X is equal to the state i, is the joint probability P_{ij}. This defines the one-step transition probabilities. A process satisfying Equation 2 is said to possess the Markov property. This property indicates that the order of state succession in the chain is independent of the process route prior to the immediately preceding state.

Observed transitions from a process lacking a first-order dependence, the Markov property, are stochastically independent random variables having a memory relationship termed zero-order. It is represented as:

$$P(X_{t+1} = j/X_t = i) = P_j \qquad (3)$$

where

P_j = the probability associated with a j transition. That is, the probability of a transition to the state j, at time t + 1, given that the process was in state i, at time t, is equal to the probability P_j. The immediately preceding state of the system has no influence upon the present state of the system, so the state transition is independent. The probability P_j is determined by the percentage of occurrence of the state j in the sample set.

PROCEDURE

Depths of water or bottom elevation at all station locations were recorded daily by Davis and Fox (1971) during the thirty day observation period. These observations, taken from the time-distance topographic maps (Figure 2), are the source of all subsequent generated data.

Subtle, small-scale, topographic variation within the nearshore zone was objectively detected and classified by a reference line defined by least squares linear regression techniques (Draper and Smith, 1966). Use of linear regression replaces the subjective definition of states from longshore profiles. Linear regression determines the functional relationship between two variables, X and Y, in terms of a linear function, where a change in the value of X is reflected in the response of the variable Y. This relationship allows regression to be used as a predictive tool of the response of Y, over the investigated range of X. The regression equation is of the following form:

$$\hat{Y}_1 = b_o + b_1 X_1 \tag{4}$$

where

$\hat{Y}_1$ = the estimated value of Y_1 for a given X_1,

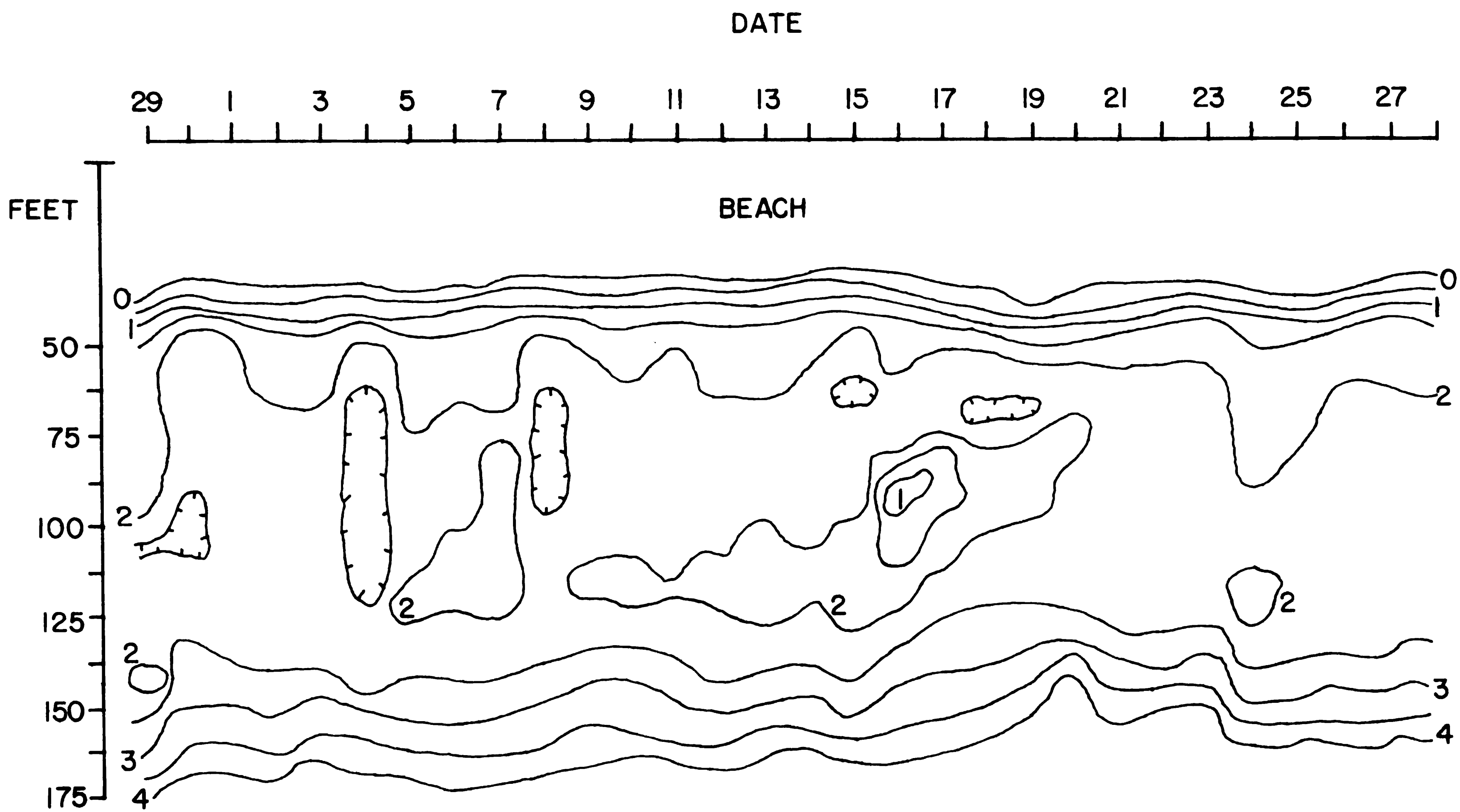
b_o = the Y intercept,

b_1 = the regression coefficient or slope factor, and

X_1 = the given value X_1.

The parameters b_o and b_1 are estimated by least squares

Figure 2 - Time-distance topographic map of Range B, reproduced from Davis and Fox (1971).


DATE
29
1
3
5
7
9
11
13
15
17
19
21
23
25
27
FEET
BEACH
50
75
100
125
150
175
0
1
2
3
4

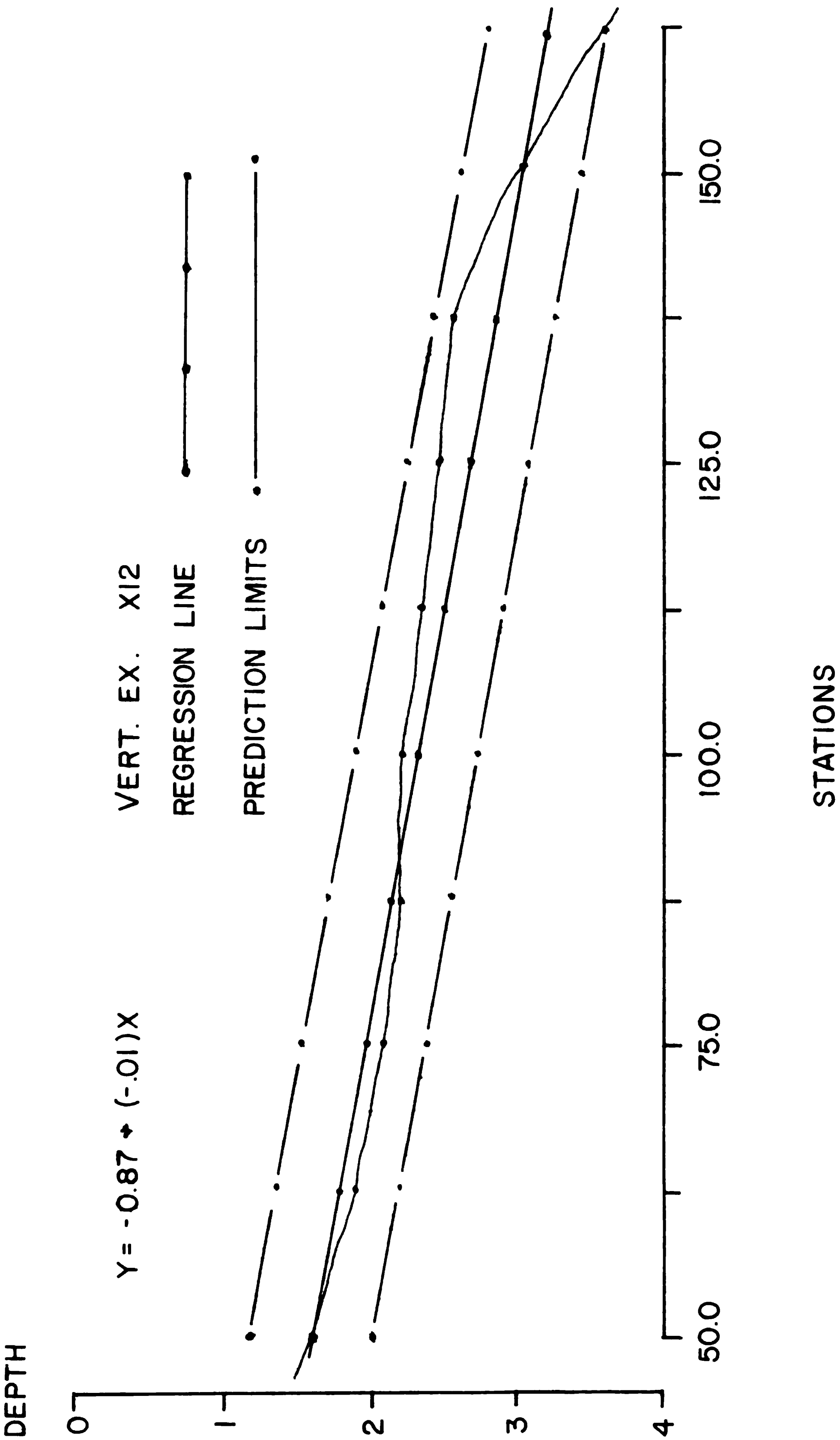
which minimizes the sum of squares of deviation from regression. The regression calculations were performed using the Stat System Version Three Least Squares Program (Michigan State University, 1973). A measure of how accurately the regression predicts the response of Y is the coefficient of determination, R^2, (Draper and Smith, 1966), which determines the proportion of the total variability in Y associated with the variability in X. With depth of water the dependent variable Y and distance outward from the shore the independent variable X, the reference line for each range was obtained, to which all other topographic variations are compared. Data for regression analyses were chosen from a day during the study period when the recorded elevations along a range approximated a uniform slope and the longshore profile was featureless. The regression equations obtained (Table 1) from these data mathematically define the negative sloping longshore profiles (Figure 3). Equations were chosen that give a high coefficient of determination, a significant F test in the analysis of variance, a low error mean square or deviation from regression for the setting of prediction limits, and a one per cent slope on the regression coefficient.

Prediction limits (Sokal and Rohlf, 1969) for a single other predicted value of bottom elevation were computed to the reference line. Observations that fall outside of these limits are statistically significant topographic expressions

TABLE 1. Regression equations derived for each range.

Range	Regression Equations	Date of Data Used for Regression
A	Y = -0.84 + (-.01)X	7/14/70
100	Y = -1.33 + (-.01)X	7/ 7/70
200	Y = -1.33 + (-.01)X	7/ 2/70
300	Y = -1.16 + (-.01)X	7/ 2/70
B	Y = -0.87 + (-.01)X	7/ 2/70
500	Y = -1.12 + (-.01)X	7/14/70
600	Y = -1.21 + (-.01)X	7/ 4/70
700	Y = -1.23 + (-.01)X	7/17/70
C	Y = -1.46 + (-.01)X	6/30/70

Figure 3 - Longshore profile along Range B on 7/2/70 with fitted regression line and prediction limits.


DEPTH
0
1
2
3
4
Y= -0.87 + (-.01)X
VERT. EX. X12
REGRESSION LINE
PREDICTION LIMITS
50.0
75.0
100.0
125.0
150.0
STATIONS

in all stages of development and magnitude. The functional relationship defined by X and Y in the regression equation is employed to predict the outcome of other observations. The standard error (Sokal and Rohlf, 1969), $\hat{S}_y$, for a single predicted value of Y_i based on a given value of X_i is derived in the following manner:

$$\hat{S}_y = \sqrt{S^2_{y \cdot x} \left(1 + \frac{1}{n} + \frac{(X_i - \bar{X})^2}{\Sigma x^2}\right)} \qquad (5)$$

where

$S^2_{y \cdot x}$ = the error mean square obtained from the regression,

X_i = the given values of X, $i = 1, \cdots, n$,

$\bar{X}$ = the mean of the X's, and

Σx^2 = the sum of the squares of the X's.

With n-2 degrees of freedom, prediction limits (Sokal and Rohlf, 1969) are computed by the following formula:

$$P(\hat{Y}_i - t_\alpha \hat{S}_y \leq \mu_{yi} \leq \hat{Y}_i + t_\alpha \hat{S}_y) = 1 - \alpha \qquad (6)$$

where

$\hat{Y}_i$ = the estimated value of Y, for a given X_i,

t_α = the table value from the t distribution, with α = .05 and n-2 degrees of freedom,

$\hat{S}_y$ = the standard error of a predicted value of Y_i, for a given X_i, and

μ_{yi} = the parametric value of Y_i.

These limits delineate two hyperbole about the regression line (Figure 3). The prediction limits form an envelope of acceptance about the one per cent slope. Recorded values of elevation, from the time-distance topographic maps, that fall outside of the range of the prediction limits represent statistically significant departures from this slope.

All observations can now be categorized into three states, (1) no significant deflection from slope (slope), (2) positive deflections (up), and (3) negative deflections (down) (Figure 4) so that the response of the system to any process is determined by bottom elevations. So structured, the nearshore zone can be observed as a discrete-time, discrete-state stochastic process. The family of random variables formed by the states and their order of transition through time describes the evolution of nearshore coastal features as a Markov chain.

Station locations based on a small interval of separation were tested for independence of observations based upon contingency table tests of observed and expected frequency of state occurrence (Sokal and Rohlf, 1969). Statistical independence exists if the probability of two events occurring together is the product of their separate probabilities. That is:

$$f_{ij} = P(X = xi \text{ and } Y = yj) = P(X = xi) \cdot P(Y = yj) \quad (7)$$

where

f_{ij} = the observed frequency of occurrence in the ij^{th} cell of the tally matrix,

X, Y = distinct, random variables, and

x_i, y_j = the corresponding states for X at i and Y at j for all ij.

Contingency tables, based upon this required property of probabilities, determine if three properties occurring in

Figure 4 - Form of matrix used to develop tallies and probabilities of state changes. States are (1) slope, where there is no significant deflection from a uniform 0.01 slope, (2) up, where deflection is positive, as in a bar, and (3) down, where deflection is negative as in a trough.

STATE AT TIME t + 1

		SLOPE	UP	DOWN	i
STATE AT TIME t	SLOPE	p_{11}	p_{12}	p_{13}	1
	UP	p_{21}	p_{22}	p_{23}	2
	DOWN	p_{31}	p_{32}	p_{33}	3
	j	1	2	3	

two states are independent. Hypotheses of the form:

H_o : the row is independent of the column classification (Figure 4)

H_1 : H_o is not true

are tested from tables of the Chi-square distribution. The computational formula (Sokal and Rohlf, 1969) is of the following form:

$$\chi^2 = \Sigma \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \quad (8)$$

where

O_{ij} = the observed frequency of occurrence in the ij^{th} cell, and

E_{ij} = the expected frequency of occurrence in the ij^{th} cell.

E_{ij} is computed as:

$$E_{ij} = \frac{ni \cdot n \cdot j}{n\cdot\cdot} \quad (9)$$

where

n_{ij} = the number of observations in the ij^{th} cell,

$n_{i\cdot} = \sum_{j=1}^{r} nij$ = the column marginal totals,

$n\cdot j = \sum_{i=1}^{c} nij$ = the row marginal totals, and

$n\cdot\cdot = \sum_{ij}^{rc} n\cdot ij\cdot$ = the grand total

Based upon (r-1)(c-1) degrees of freedom, this is a model two test in which the n • j marginal totals are fixed. Initially, 10 by 2 tables (Table 2) for an entire range were computed to determine if total independence occurred. It

TABLE 2. 10 by 3 contingency table for testing of total independence along Range B.

STATE	SLOPE		UP		DOWN		
Station	o S	e	o U	e	o D	e	$n_{\cdot j}$
50.0	25	21.51	0	3.53	5	4.94	30
62.5	24	21.51	0	3.53	6	4.94	30
75.0	27	21.51	0	3.53	3	4.94	30
87.5	25	21.51	3	3.53	2	4.94	30
100.0	26	21.51	3	3.53	1	4.94	30
112.5	18	21.51	12	3.53	0	4.94	30
125.0	18	21.51	12	3.53	0	4.94	30
137.5	25	21.51	4	3.53	1	4.94	30
150.0	21	21.51	1	3.53	8	4.94	30
162.5	4	19.36	0	3.18	23	4.45	27
$n_{i\cdot}$	213		35		49		297 $n_{\cdot\cdot}$

ȯ = observed frequency

e = expected frequency

$\chi^2_{.05(9)(2)} = 28.869$

Σ using Equation 8 = 172.2457 Reject H_o

did not. Station location pairs, progressively outward from the shore, were then compared in 2 by 3 tables (Table 3). From these tables the areal distribution of independence was determined (Figure 5). No discernible patterns are apparent and dependence is not restricted to any range of stations.

A random variable is a mathematical entity occurring from probabilistic mechanisms, just as a systematic variable occurs from deterministic mechanisms (Ross, 1972). A random variable is a real valued function defined on a probability space such that the probability of sets of the following form can be defined:

$$(\omega / X_{(\omega)} \leq x) \qquad (10)$$

where

ω = an element of a random phenomena, and

$X_{(\omega)}$ = the set of points at which the value of X does not exceed x.

For every x there is a distribution function of the random variable X, such that:

$$FX_{(x)} = P(\omega / X_{(\omega)} \leq x) \qquad (11)$$
$$= P(X \leq x)$$

where

F is non-decreasing and right continuous,

$0 \leq FX_{(x)} \leq 1$, and

$\lim_{x \to \infty} FX_{(x)} = 1$.

Topographic transitions, observed as a process response through time, were tested for the first order

TABLE 3. 2 by 3 contingency table for testing of independence between stations 150.0 and 162.5, Range B.

STATE	SLOPE		UP		DOWN		
Station	o S	e	o U	e	o D	e	$n_{\cdot j}$
150.0	21	13.15	1	0.52	8	16.31	30
162.5	4	11.84	0	0.47	23	14.68	27
$n_{i\cdot}$	25		1		31		57 $n_{\cdot\cdot}$

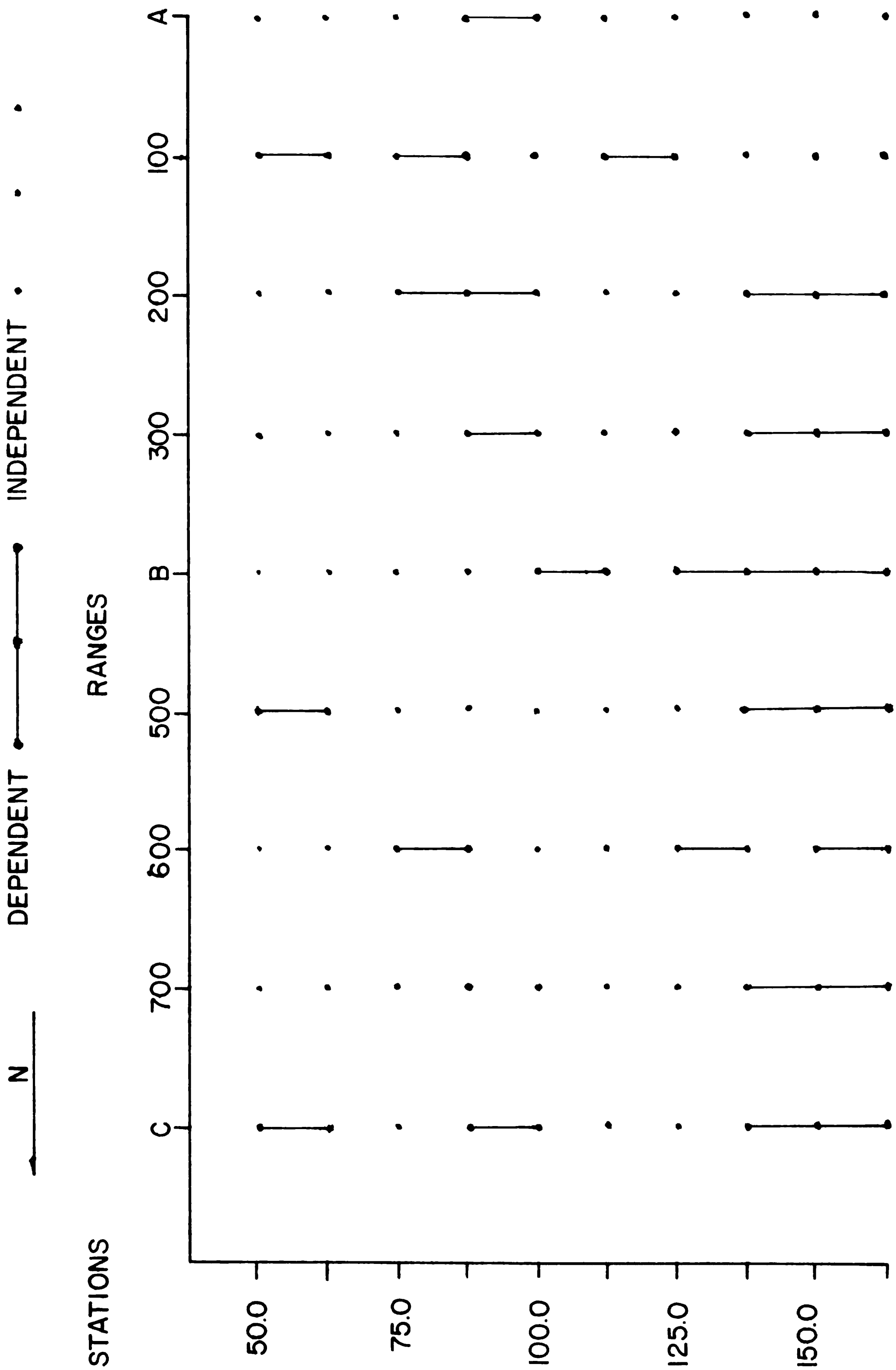
o = observed frequency

e = expected frequency

$\chi^2_{.05(1)(2)} = 5.991$

Σ using Equation 8 = 19.7396 Reject H_o

Figure 5 - Station location map showing independence of stations.


DEPENDENT
INDEPENDENT
N
RANGES
STATIONS
A
100
200
300
B
500
600
700
C
50.0
75.0
100.0
125.0
150.0

Markov chain property, as opposed to a sequence of stochastically independent, identically distributed, random variables. Testing is based upon the likelihood ratio criterion (Anderson and Goodman, 1957) and takes the following computational form:

$$-2 \log_e \lambda = 2 \left(\sum_{ij}^{m} n_{ij} \log_e \frac{(P_{ij})}{P_j} \right) \qquad (12)$$

where

λ = the likelihood ratio criterion,

m = the number of states,

n_{ij} = the number of observations in the ij^{th} cell of the probability transition matrix,

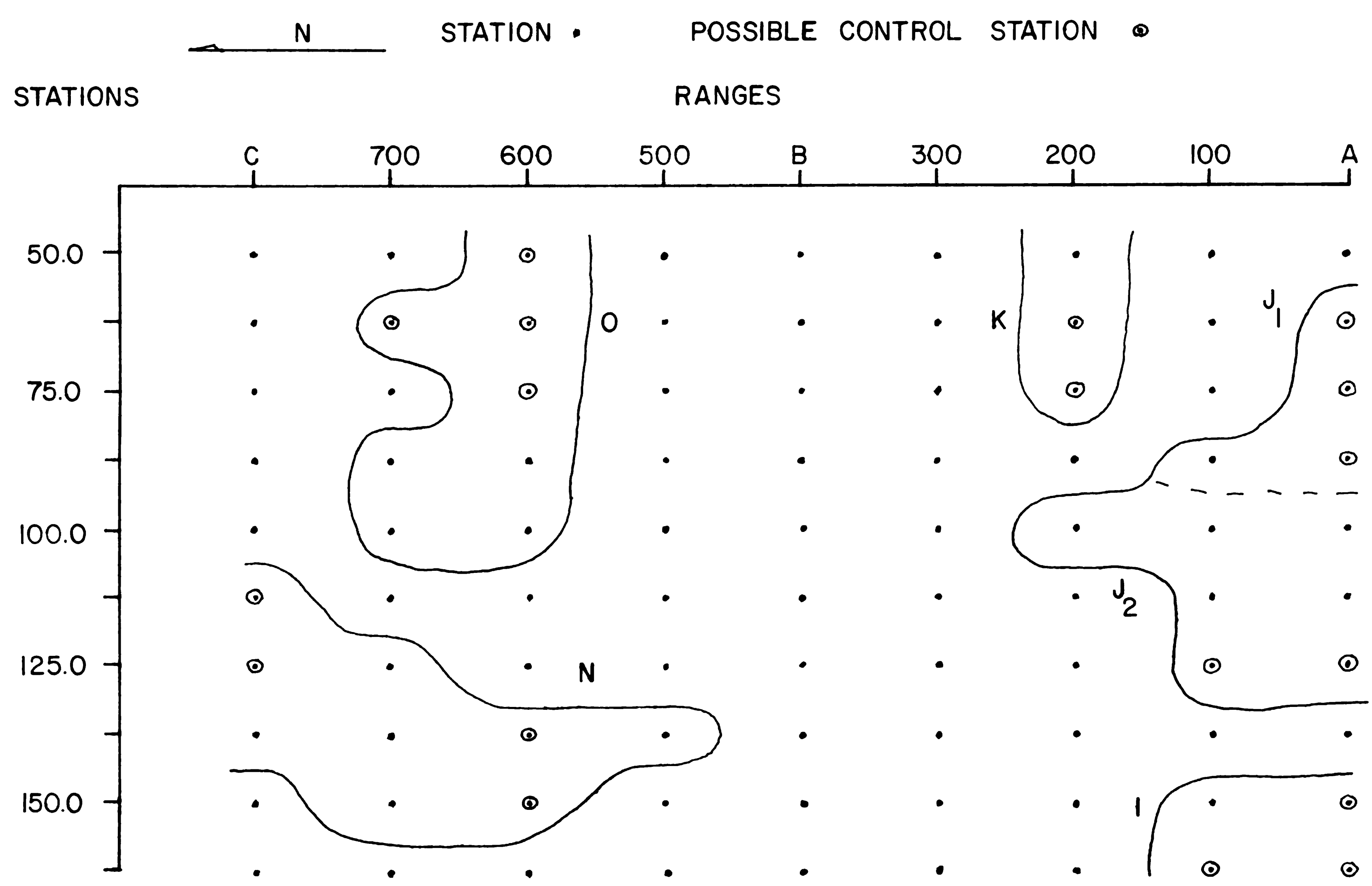
P_{ij} = the probability in the ij^{th} cell, and

P_j = the marginal probability for the j^{th} column.

The asymptotic distribution of $-2 \log_e \lambda$ is based upon a Chi-square distribution with $(m-1)^2$ degrees of freedom. A map of the study area was obtained that outlines those stations that possess the Markov property (Figure 6). Several clusters are apparent, within which the associated transition probabilities show a first order Markov property. These clusters appear to delineate areas subject to similar process. The resulting map is a fixed distance analysis.

Tests for stationarity in time and space were conducted on the transition probability matrices within clusters to substantiate that the Markov properties of the clustered stations justify grouping. In a stationary

Figure 6 - Location map showing areas occupied by first order clusters, I, J, K, O, N designate clusters.


N
STATION •
POSSIBLE CONTROL STATION ⊙
STATIONS
RANGES
C
700
600
500
B
300
200
100
A
50.0
75.0
100.0
125.0
150.0
O
K
J1
J2
N
I

process the transition probabilities are approximately equal and constant through time or space (Harbaugh and Bonham-Carter, 1970). In a stationary first-order Markov chain, P_{ij} is the probability of transition from state i at time t to j at time t + 1. In a nonstationary chain, probabilities vary in time or space, and P_{ij} (t) becomes a function of time or space. The hypotheses tested are:

$H_o : P_{ij}(t) = P_{ij}$ stationary

$H_1 : P_{ij}(t) = P_{ij}$ nonstationary

Testing of these hypotheses is based upon the likelihood ratio criterion and is similar in form to that used to determine the order of the chain (Equation 12). A station is designated as the control station, to which all other stations are compared for stationarity, by means of the following equation:

$$-2 \log_e \lambda = 2 \left(\sum_{t=1}^{T} \sum_{ij}^{m} n_{ij}(t) \log_e \left(\frac{P_{ij}(t)}{P_{ij}} \right) \right) \quad (13)$$

where

λ = the likelihood ratio criterion,

t = 1, ..., T = the number of subintervals,

m = the number of states,

$n_{ij}(t)$ = the tally matrix in the ij^{th} cell of the t^{th} subinterval,

$P_{ij}(t)$ = the probability matrix in the ij^{th} cell of the t^{th} subinterval, and

P_{ij} = the probability matrix in the ij^{th} cell of the testing location or control station.

This is distributed as a Chi-square with m(m-1)(T-1) degrees of freedom. All stations within a cluster were tested as the control station using Equation 13. All clusters were found to be stationary with respect to at least one station within it. This allows for the grouping together of a number of station locations which possess the Markov property of a first order chain. One station was chosen from each cluster to characterize the cluster. This station functions as the control station in comparison to which the other stations in the cluster are stationary.

The one step transition probabilities, $P_{ij}^{1} = P_{ij}$, was introduced in the discussion on Markov chains. An n-step transition probability (Ross, 1972), $P_{ij}^{(n)}$, is defined as the probability that a process in state i will be in state j after n transitions. Mathematically, the relation is represented as:

$$P_{ij}^{(n)} = P_{ij}^{(t)} \times P_{ij}^{(t+1)} \qquad (14)$$

where

$P_{ij}^{(n)}$ = the probability matrix after n transitions,
$P_{ij}^{(t)}$ = the probability matrix at time t, and,
$P_{ij}^{(t+1)}$ = the probability matrix at time t+1

In essence the n-step transition probability matrix is obtained by the matrix multiplication of the initial matrix, $P_{ij}^{(1)}$, by itself n times. As $n \to \infty$, $P_{ij}^{(n)}$ converges to a value which is identical for all i; all rows in the matrix are equal. This value, the limiting probability, is the probability that after n transitions, the process will be in state j, independent of the initial state. The limiting probability is also equal to the proportion of the time that the process will occupy state j. Limiting probability matrices were computed for the control station for each cluster to obtain the limiting probability matrix or equilibrium matrix characterizing the cluster. Table 4 lists the tally, probability, and limiting probability matrices for each cluster.

State j communicates or is accessible from state i if $P_{ij}^{(n)} > 0$, $n \geq 0$ (Kemeny and Snell, 1960). The communication is denoted by $i \leftrightarrow j$. Any two states that communicate are in the same class. The concept of communication divides the state space, $\sum_{t=1}^{T} X_t$, into separate classes. A finite, irreducible Markov chain is composed of one class. For any state i, let f_i denote the probability that starting in i the process will ever return to i. State i is recurrent if $f_i = 1$ and transient if $f_i < 1$. If recurrent the process will return to state i infinitely often, if transient there will be a positive probability, $1 - f_i$, that the process will never return to state i. Recurrence and transience are class properties. If the process returns to state i in finite time, the state i is positive recurrent. All recurrent states in a finite state irreducible Markov chain are positive recurrent. The communication of states and the internal structure of the chain for the controlling station for each cluster is evaluated in the separate discussions of the clusters in the following section.

TABLE 4. Cluster characteristics for each controlling station.

Cluster (See Fig. 6)		Tally Matrix			Probability Matrix			Limiting Probability Matrix		
		S	U	D[1]						
I	S	6	2	4	.50	.17	.33	.3757	.0852	.5391
	U	3	1	0	.75	.25	.00	.3757	.0852	.5391
	D	3	0	10	.23	.00	.77	.3757	.0852	.5391
J Sub 1		6	0	2	.75	.00	.25	.3590	.0000	.6410
		0	0	0	.00	.00	.00	.0000	.0000	.0000
		3	0	18	.14	.00	.86	.3590	.0000	.6410
Sub 2		4	3	3	.40	.30	.30	.3212	.4380	.2409
		2	7	0	.22	.78	.00	.3212	.4380	.2409
		4	0	6	.40	.00	.60	.3212	.4380	.2409
K		11	1	2	.79	.07	.14	.4843	.0696	.4461
		1	0	1	.50	.00	.50	.4843	.0696	.4461
		2	1	10	.15	.08	.77	.4843	.0696	.4461
O		7	0	3	.70	.00	.30	.3478	.0000	.6522
		0	0	0	.00	.00	.00	.0000	.0000	.0000
		3	0	16	.16	.00	.84	.3478	.0000	.6522
N		7	3	0	.70	.30	.00	.3478	.6522	.0000
		3	16	0	.16	.84	.00	.3478	.6522	.0000
		0	0	0	.00	.00	.00	.0000	.0000	.0000

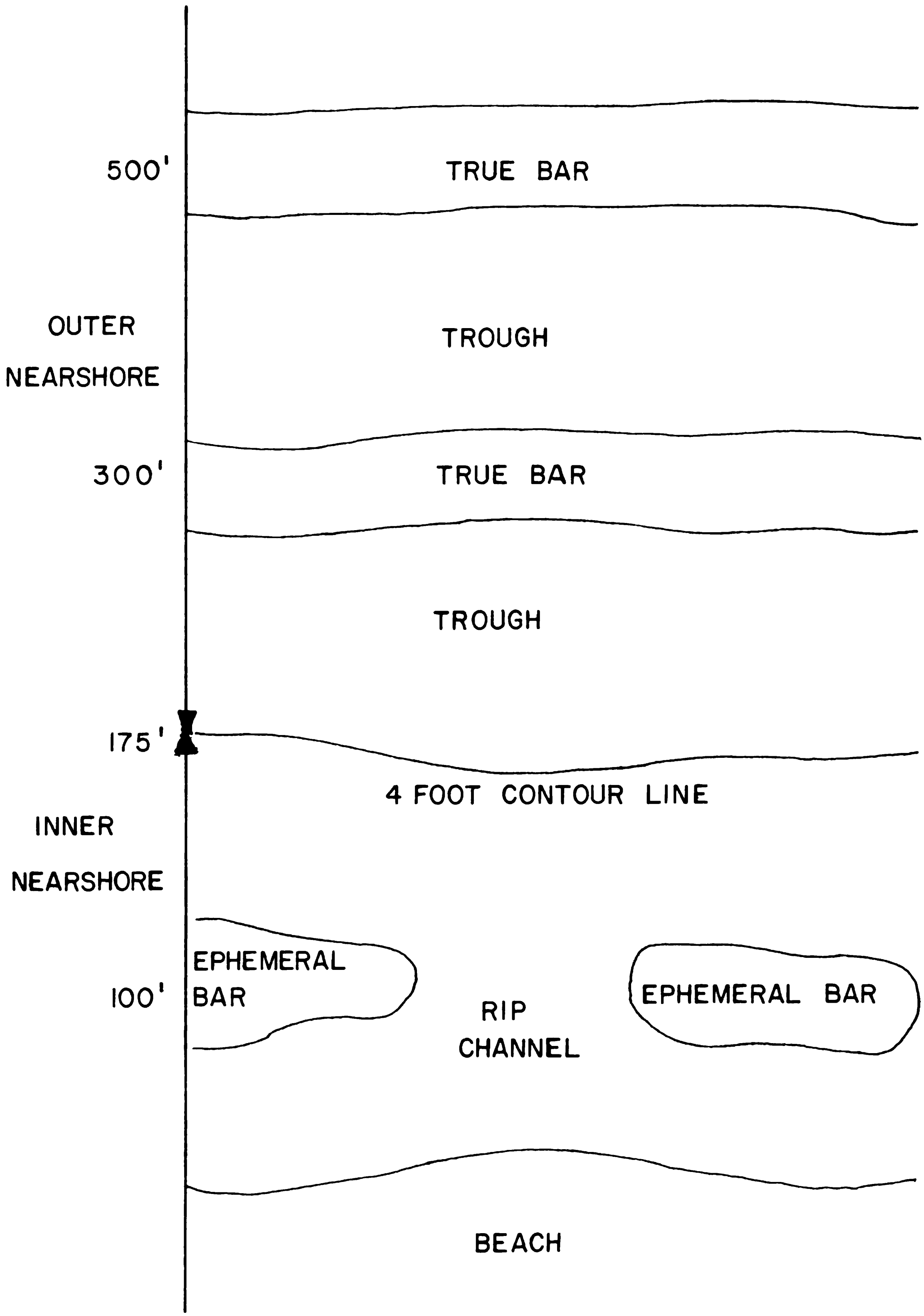
1. All 3 by 3 matrices have the indices S, U, D, which stand for the states slope, up, and down. See text for explanation.

DISCUSSION

Davis and Fox (1971) have noted the presence of two or more distinct bars at approximately 300 and 500 feet outward from the shoreline. These true sand bars are fairly continuous, are storm generated, and have an associated trough on their shoreward side (Figure 7). The boundary that separates the nearshore zone from the nearest trough, the inner from the outer nearshore, is the four foot depth contour line (Figures 2 and 7), and is approximately 175 feet offshore. Within the inner nearshore zone a shallow bar is usually present approximately 100 feet offshore. This wave generated ephemeral bar is normally discontinuous, being interrupted by rip channels, and is ocassionally eroded away completely by storm-generated waves. There is usually an associated trough-like structure in front of the ephemeral bar. The trough sometimes grades off into a subaqueous terrace. The ephemeral bar form does not migrate in the direction of littoral drift but seems to oscillate within the 100.0 to 125.0 foot range.

The five defined clusters span the environments of ephemeral bar and trough development along with the swash zone and the area dropping into the first true trough at the four foot contour line. Analysis of the internal chain structure of the controlling station for each cluster quantifies and supports the observational data on the

Figure 7 - Generalized map of the inner and outer nearshore zone of eastern Lake Michigan. Modified from Davis and Fox (1971). Not to scale.


500'
TRUE BAR
OUTER
NEARSHORE
TROUGH
300'
TRUE BAR
TROUGH
175'
4 FOOT CONTOUR LINE
INNER
NEARSHORE
EPHEMERAL
BAR
100'
RIP
CHANNEL
EPHEMERAL BAR
BEACH

behavior of the inner, ephemeral bar and trough structures within the nearshore zone. It appears that there is a tendency by the inner bar to develop the equilibrium configuration that characterizes the true, outer bars. That is, the inner bar tends to develop a continuous bar structure within the nearshore. However, large amounts of water congregated into rip channels cut across the bar form and prevent a continuous bar from forming. The result is that the bar form oscillates back and forth within its location range. The large central portion of the study area (Figure 6) is probably one of these rip channels. The continuous, rapid deposition and erosion prevent the topographic fluctuations within this area from having a significant effect upon each other, hence they are stochastically independent, random variables. Those areas that fall into the clusters (Figure 6) are less likely to fluctuate than the central, rip channel area, so they are more likely to show continuity of process through time. An analyses of each cluster, the environment the cluster spans, and the way in which the sediment responds in topographic fluctuations, as described by the chain structure, follows.

Cluster I

Cluster I is situated at the far end of ranges A and 100 (Figure 6) offshore of the area of ephemeral bar activity and near the slope into the first true trough. The chain structure is that of a finite, positive recurrent, irreducible Markov chain. Within the single class, all states communicate (Table 4) and all states can succeed themselves. The only single step transitions not possible are up - down and the converse. The up/down states communicate through the state slope, indicating a time period of approximately one day as a minimum, to proceed through the transition. Storm induced topographic activity is significant in this zone, but the response of the system is slow. The limiting probability matrix indicates approximately 54 per cent of process time is spent in the down state, a reflection of the first true trough, 38 per cent in slope, and 8 per cent in up, due to the occasional development of large bars originating from within the 100.0 to 125.0 foot range. Erosion appears the dominant factor in this zone.

Cluster J

Cluster J is located along the first three southern ranges, A, 100, 200, and spans the area of ephemeral trough and bar development. Of the eleven included stations five can function as the control station, with the majority of these stations falling in the trough area. Two subclusters which divide the environments, can be recognized.

Subcluster 1

In the trough zone, from 87.5 feet shoreward, the structure is that of a finite, positive recurrent, irreducible Markov chain with zero probability of the occurrence of the state up (Table 4). All other states communicate, either succeeding themselves or each other. Erosion is the dominant factor with sediment accumulation never surpassing uniform slope. Limiting probabilities indicate approximately 64 per cent of overall time is spent in the state down, reflecting vigorous trough activity in front of the ephemeral bars, 36 per cent in slope. A corollary to this is that bar development in the adjacent area (Subcluster 2) occurs and the bar can move lakeward but does not migrate shoreward. Transported sediment from the crest of the bar into this area is reflected in the slope state, but the sediment is removed by erosion before the bar form can migrate shoreward. This substantiates and quantifies the empirical observation that there is a net displacement down drift of the sediment by littoral drift, but not of the bar form itself.

Subcluster 2

The chain structure within the bar development area, 100.0 to 125.0 feet offshore, is similar to that of Cluster I in structure. Single step transition between the states up and down is impossible, with communication through slope being required (Table 4). The major difference, which is apparent in the limiting probability matrix, is that within this area 44 per cent of process time is spent in bar development and maintenance. Deposition is the primary factor. The ephemeral nature of these bars is indicated by the fact that within this area there is a probability that the down state can be attained, reflecting the complete removal of the bar form.

Cluster K

Cluster K is situated along Range 200 and is composed of three stations located near the swash zone. The chain structure (Table 4) is similar to that of the previous clusters. In this area, which is characterized by ridge and runnel systems, all states can succeed themselves except up. When in the up state there is an equal probability of transition to either the slope or down states. This indicates that large amounts of sediment introduced into this area are rapidly (approximately one day maximum lag time) removed in either forward transport, which results in welding to the beach proper, or transport in the direction of littoral drift, or the sediment is distributed evenly about the area. Direct transition between the states up and down, which is possible in this case, is also indicative of the rapid, erosive nature of this zone. Limiting probabilities show the long term percentage of occurrence to be nearly equally distributed between slope and down, 48 per cent and 45 per cent of the time respectively, with an average of 7 per cent of the time available for large sediment buildups.

Cluster O

Cluster O is located along ranges 600 and 700 from the near swash zone to the base of ephemeral bar development. Its structure (Table 4) is identical to that of Subcluster 1 of Cluster J in that both clusters encompass similar environments. Conclusions regarding this cluster are similar to those of Subcluster 1.

Cluster N

Cluster N covers four ranges, 500, 600, 700, and C, in the outer, northern corner of the study area. The area of the cluster is within the zone of ephemeral bar development. The structure is that of a finite, positive recurrent, irreducible Markov chain with zero probability of the occurrence of the down state (Table 4). All states can succeed themselves or one another. Deposition is the major facter, as the limiting probabilities indicate 65 per cent of process time is devoted to bar development. Comparison with the station locations of Cluster J and O, indicates that the central area of bar activity is located approximately 12 to 25 feet lakeward of that in the southern part of the study area.

CONCLUSIONS

Identification of clusters that encompass several station locations, all of which are stationary to more than one station, allows for the objective selection of the probability transition matrix that is used to characterize that cluster. In most clusters, the similarity in chain structure among possible control stations allows for selection of a representative station to characterize the process response of the cluster. Where chain structures vary, a control station can be chosen which best fits the observed environment, but this is a subjective choice of the controlling station. This is an inherent weakness in a simple first order model based on a fixed distance analyses. A more complex model which would limit the number of possible control stations is desirable.

Within all stations that respond as a first order Markov chain, a plethora of chain structures was observed. The stationary chain structure defined in each cluster was that of a finite, positive recurrent, irreducible Markov chain. This structure represents a stable configuration in which all states occupying a single class have complete intercommunication. Topographic transitions in the nearshore, when structured as a stochastic process, operate in this equilibrium configuration. The probability of major fluctuations in topography in response to large external factors, such as storms, is accounted for in the stochastic model. However the long term probabilities of occurrence, as defined in the stable configuration of the chain structure, reflects the dynamic equilibrium in the nearshore zone between land and water.

Fluctuations within the central portion of the study area are independent, with the succession of events lacking any dependence upon previous events. Hypothesizing upon the significance or stability of this area would require extrapolation outside the range of the data. Coastal process on Lake Michigan undoubtedly have seasonal variations. The data for this study were collected during the month of July. It is impossible to determine if the derived probabilities are stationary throughout the year, and the ephemeral or permanent nature of this central portion of independence cannot be ascertained from the available data.

Topographic response to external factors in the nearshore zone, when structured as a stochastic process, can be modeled on the basis of probability distributions. This allows for the quantification of empirical data and the computer simulation of the nearshore environment. Observational data of process evolution are substantiated and mathematically defined by Markov chain analysis.

BIBLIOGRAPHY

BIBLIOGRAPHY

Anderson, T. W. and Goodman, L. A., 1957. Statistical Inference About Markov Chains: Annals of Mathematical Statistics, V. 28, p. 89-110.

Davis, R. A. and Fox, W. T., 1971. Beach and Nearshore Dynamics in Eastern Lake Michigan: Office of Naval Research, Technical Report 4, NONR Contract 388-092.

Draper, N. R. and Smith, H., 1966. Applied Regression Analysis: John Wiley, New York.

Harbaugh, J. W. and Bonham-Carter, G., 1970. Computer Simulation in Geology: John Wiley, New York.

Kemeny, J. G. and Snell, J. L., 1960. Finite Markov Chains: D. Van Nostrand, New York.

Michigan State University, 1973. Stat System Version Three, L. S. Program, Computer Laboratory, East Lansing, Michigan.

Ross, S. M., 1972. Introduction to Probability Models: Academic Press, New York.

Sokal, R. R. and Rohlf, F. J., 1969. Biometry: W. H. Freeman, San Francisco, California.