12-15-1965

# Predictor Equations for Beach Processes and Responses

W. Harrison

N. A. Pore

D. R. Tuck Jr.
*Virginia Institute of Marine Science*



## Recommended Citation

Harrison, W.; Pore, N. A.; and Tuck, D. R. Jr., Predictor Equations for Beach Processes and Responses (1965). *Journal of Geophysical Research*, 70(24), 6013-6109.
https://scholarworks.wm.edu/vimsarticles/2054

# Predictor Equations for Beach Processes and Responses[1]

W. HARRISON

*Environmental Science Services Administration*
*Institute for Oceanography, Norfolk, Virginia*

N. A. PORE

*Environmental Science Services Administration*
*Weather Bureau, Washington, D. C.*

D. R. TUCK, JR.

*Virginia Institute of Marine Science, Gloucester Point*

*Abstract.* A stepwise (linear) multiple regression procedure is applied to 11 environmental variables (or predictors) in the beach-ocean-atmosphere system at Virginia Beach, Virginia, for the following five predictands: mean longshore current velocity, mean bottom slope in the shoaling-wave zone, average mean grain size in the shoaling-wave zone, and beach deposition and beach erosion on the lower foreshore. Predictors consist of variables related to beach geometry, local water properties, local wind conditions, tidal fluctuations, and wave characteristics. The resultant equations are tested against a set of independent data and, with one exception, agree reasonably. It is believed that if the data set were increased to include at least one year's continuous measurements, the procedure outlined would yield valid equations for all but stormy-weather conditions. It is presupposed that some provision will have to be made for preconditioning the data, as 'storm' and 'nonstorm' data will probably have to be analyzed separately.

*Introduction.* In recent studies of large numbers of simultaneous measurements in the beach-ocean-atmosphere system, computerized 'search procedures' [*Harrison and Krumbein,* 1964] have been used for identifying interactions among subsets of the numerous environmental variables or for identifying relationships [*Harrison and Pore,* 1964] between selected 'independent' and 'dependent' variables in the system. Either approach has a certain advantage over the more classical studies in the wave tank or field because numerous variables are allowed to enter into the analysis. Previous studies have tended to fragment the system, either in controlled laboratory experiments or in environmental studies of a limited number of variables that may help to explain certain facets of the many phenomena composing the whole.

The objectives of this paper are (1) to present a screening technique for analysis of the multivariate system in terms of predictors ('independent' variables) and predictands ('dependent' variables),[2] (2) to determine predictor equations for beach processes (forces that act to modify the beach, such as the longshore current) and for beach responses (results of the activity of the processes, such as beach erosion), and (3) to present tests of five typical predictor equations on a set of independent data.

*Data.* The variables used in this study were measured at Virginia Beach, Virginia (Figure 1), and are listed in Table 1, together with their observed ranges. These variables are the same as those reported in the earlier study by *Harrison and Krumbein* [1964], but we have used a considerably enlarged data set that includes

[1] Contribution 3 of the Land and Sea Interaction Laboratory, Institute for Oceanography; 439 West York Street, Norfolk, Virginia, and contribution 195 of the Virginia Institute of Marine Science, Gloucester Point, Virginia.

[2] In regression analysis 'the variate $y$ is called the *regressand,* and the associated variates $x_1$, $x_2$, $\cdots$, $x_n$ are called *regressors;* or alternatively, $y$ is called the *predictand* and the $x$'s are called *predictors*' [*Glossary of Meteorology,* p. 476, 1959].

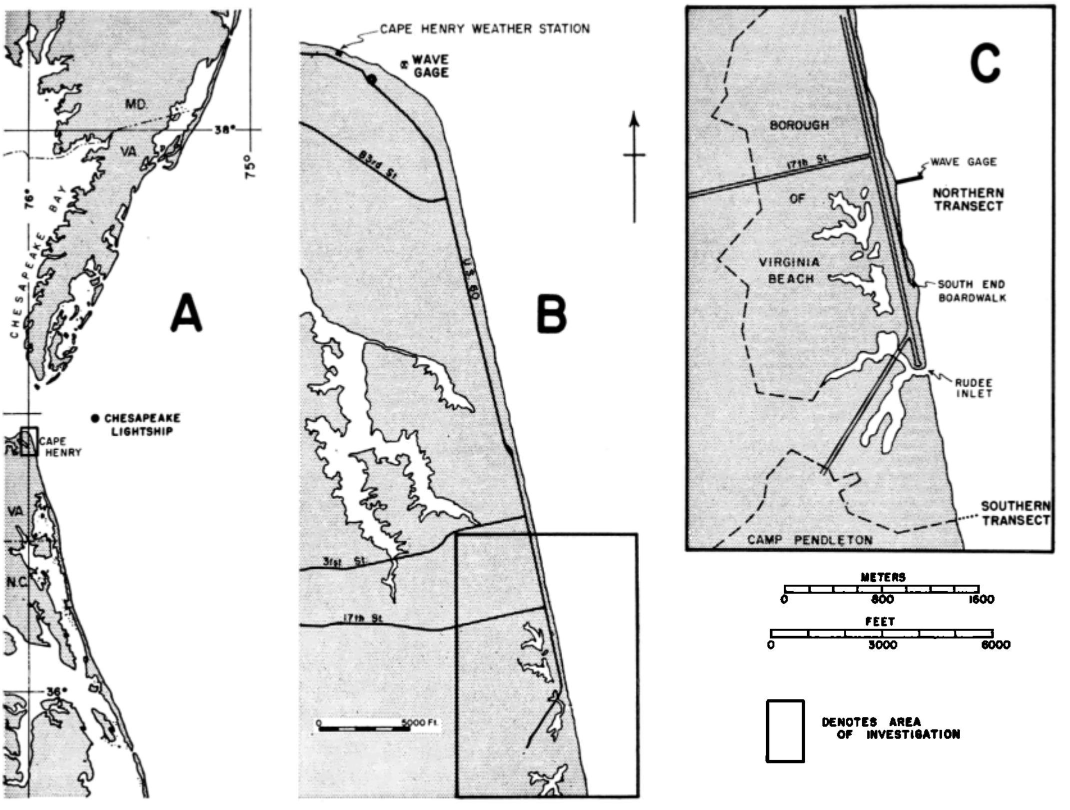


Fig. 1. Maps showing area of investigation and transects at which measurements were taken.

several days' measurements taken during storm conditions. For additional details of measurement techniques beyond those given in Table 1, the reader is referred to appendix A in *Harrison and Krumbein* [1964] and appendix A in *Tuck* [1965].

*Harrison and Krumbein* [1964] amply demonstrated that there is a time lag in the peak interaction between a given predictand ($Y$) and the four or five predictors ($X_n$) that explains most of the observed variability in the predictand. The present data set was coded with respect to time, therefore, so that the predictors could be selected that may have a significant influence on predictand up to 24 hours before measurement of the predictand.

*Method of analysis.* The procedure adopted for selecting predictors involves expressing $Y$ as a linear function of a number of $X_n$ ($n = 1, \ldots, N$).

Thus

$$Y = A_0 + A_1X_1 + A_2X_2 + \cdots A_nX_n + \cdots A_NX_N$$

where the coefficients $A_n$ ($n = 0, \ldots, N$) are determined using the method of least squares.

One limitation of the linear analysis is that some variables that have only a small linear effect may become quite strong in a model that explicitly includes nonlinear effects. This has been demonstrated by *Harrison and Krumbein* [1964, p. 27]. It is also true, however, that the linear model is generally the best one for initial work with large numbers of variables.

Because of the large numbers of predictors involved, the screening procedure as described by *Miller* [1958] required the use of a high-speed, large-memory computer; the IBM 7030 was used. Basically, the technique is shown below.

$$Y = A_1 + B_1X_1 \tag{1}$$

$$Y = A_2 + B_2X_1 + C_1X_2 \quad (2)$$

$$Y = A_n + B_nX_1 + C_{n-1}X_2, \cdots , NX_n \quad (3)$$

where the $A$'s are constants and $B_1$, $B_2$, $C_1$, $C_2$ etc., are regression coefficients.

The procedure is to first select the best single predictor ($X_1$) for regression equation 1. The second regression equation (2) contains $X_1$ and the $X_2$ that contributes most to reducing the residual after $X_1$ is considered, regardless of its lag position. This is not necessarily the best subset of $X_i$ out of the original set. In the closely analogous field of meteorology, however, studies such as that of *Klein et al.* [1959] have shown that by this screening procedure a highly reliable set of predictors can be selected in prob-

TABLE 1. Variables Used in Development of Predictor Equations

| Symbol | Dimensions | Description of Variable | Range in Values |
|---|---|---|---|
| $C$ | $LT^{-1}$ | Velocity of tidal current (measured 1 m off bottom, 265 m from shore, using Price meter or Savonius rotor) | 0.00–35.6 cm/sec |
| $C_O$ | | Velocity of tidal current flowing in opposite direction to longshore current | 0.00–20.1 cm/sec |
| $C_s$ | | Velocity of tidal current flowing in same direction as longshore current | 0.00–35.6 cm/sec |
| $D$ | $L$ | Depth of water table at top of uprush | 0.00–0.56 m |
| $h$ | $L$ | Still-water depth (measured at northern transect, Figure 1C, at point 100 m from shore) | 2.44–4.91 m |
| $H$ | $L$ | Local significant wave height measured at CERC's relay-type wave gage (Figure 1C). | 0.15–3.93 m |
| $J_f$ | $L$ | Net deposition over 30–38 m of lower foreshore in 12.25-hour period. Most measurements taken at northern transect; some taken at southern (Figure 1C) | 0.00–2.01 m |
| $K_f$ | $L$ | Net erosion over 30–38 m of lower foreshore in 12.25-hour period | 0.00–1.58 m |
| $(\bar{M}_z)_s$ | $L$ | Average mean nominal grain diameter over bottom in shoaling-wave zone at northern transect (Figure 1C) and 75–100 m from shore | 0.234–0.843 mm |
| $\bar{S}_f$ | | Mean slope of lower foreshore of beach where $J_f$ and $K_f$ were measured | 1.40°–4.55° |
| $\bar{S}_s$ | | Mean slope of beach over inner portion of shoaling-wave zone where $(\bar{M}_z)_s$ was measured | 0.45–2.22° |
| $T$ | $T$ | Wave period at CERC's wave gage (Figure 1C), based on significant-wave strip-chart analysis | 3.00–13.90 sec |
| $\bar{U}$ | $LT^{-1}$ | Mean wind velocity measured at Cape Henry Weather Bureau Station (Figure 1B) | |
| $\bar{U}_a$ | $LT^{-1}$ | Mean wind velocity directed against the longshore current | 0.00–7.2 m/sec |
| $\bar{U}_{of}$ | | Mean wind velocity in an offshore direction | 0.00–17.8 m/sec |
| $\bar{U}_{on}$ | $LT^{-1}$ | Mean wind velocity in an onshore direction | 0.00–11.95 m/sec |
| $\bar{U}_p$ | | Mean wind velocity parallel to shore | 0.00–18.77 m/sec |
| $\bar{U}_s$ | | Mean wind velocity in same direction as longshore current | 0.00–14.30 m/sec |
| $\bar{V}$ | $LT^{-1}$ | Mean velocity of longshore current as measured by timing motion of dye patches over 30 m distance | 0.00–97.5 cm/sec |
| $\alpha$ | | Angle of wave approach, measured with pelorus in zone 300–400 m from shore at northern transect (Figure 1C) | 2°–75° |
| $\eta_r$ | $LT^{-1}$ | Rate of rise of still-water level based on C&GS tide-gage at northern transect (Figure 1C) | 0.00–6.3 $\times$ $10^{-3}$ cm/sec |
| $\eta_f$ | $LT^{-1}$ | Rate of fall of still-water level | 0.00–5.2 $\times$ $10^{-3}$ cm/sec |
| $\rho$ | $ML^{-3}$ | Water density as measured 100 m from shore at northern transect (Figure 1C) | 1.0136–1.02500 g/cm$^3$ |

TABLE 2. Selection of Predictors by Screening Process (Lag interval expressed in hours.)

| Equation | $\bar{V}$, Run 1 ($N$ = 60) | | | $\bar{S}_s$, Run 2 ($N$ = 27) | | | $(\bar{M}_z)_s$, Run 3 ($N$ = 27) | | | $J_f$, Run 4 ($N$ = 25) | | | $K_f$, Run 5 ($N$ = 33) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Var. | Lag | $r$ | Var. | Lag | $r$ | Var. | Lag | $r$ | Var. | Lag | $r$ | Var. | Lag | $r$ |
| 1 | $H$ | | 0.56 | $h$ | 8–12 | 0.60 | $h$ | 16–20 | 0.58 | $H$ | 8–12 | 0.80 | $\bar{U}_p$ | 0–4 | 0.70 |
| 2 | $T$ | | 0.68 | $\rho$ | 12–16 | 0.75 | $\bar{S}_s$ | 0–4 | 0.79 | $\bar{U}_p$ | 0–4 | 0.89 | $H$ | 16–20 | 0.86 |
| 3 | $C_s$ | | 0.71 | $C$ | 0–4 | 0.82 | $\alpha$ | 16–20 | 0.87 | $\bar{U}_{on}$ | 8–12 | 0.93 | $T$ | 20–24 | 0.90 |
| 4 | $\bar{S}_f$ | | 0.73 | $C$ | 4–8 | 0.86 | $\bar{U}_p$ | 4–8 | 0.91 | $\rho$ | 4–8 | 0.95 | $\eta_r$ | 8–12 | 0.92 |
| Max. $r$ | (8 pred.) | | 0.76 | (10 pred.) | | 0.97 | (10 pred.) | | 0.98 | (9 pred.) | | 0.99 | (11 pred.) | | 0.98 |

lems that involve redundant, interrelated variables, such as those of the present data set.

In most previous studies of natural systems, such as this, that involve highly redundant data, the significance of the improvement attained at each step of the screening is tested (usually by $F$ ratios [*e.g.*, *Fritts*, 1962]) and the screening is discontinued when the amount of improvement is found not to be significant. *Panofsky and Brier* [1958] point out that objective standard significance tests may be misleading on data like ours because the underlying assumptions are in general violated. The $X_i$ used here, for example, are interdependent in time and space. Thus we believe that the most practical and convincing test of significance is an application of the result to an independent set of data. For the purpose of conducting such a test we withheld 11 to 15% of our data in each screening run; the actual values preselected were chosen to cover the ranges observed for the predictands.

Predictor equations were developed for one beach process and four beach responses, as follows.

*Mean longshore current velocity:*

$$\bar{V} = f(\bar{S}_f, T, H, \bar{U}_{on}, \bar{U}_{of}, \bar{U}_s, \bar{U}_a, \alpha, \rho, C_s, C_o) \quad (4)$$

*Mean bottom slope in the shoaling-wave zone:*

$$\bar{S}_s = f[(\bar{M}_z)_s, T, H, \bar{U}_{on}, \bar{U}_{of}, \bar{U}_p, \alpha, h, \rho, C] \quad (5)$$

*Average mean grain size in the shoaling-wave zone:*

$$(\bar{M}_z)_s = f(\bar{S}_s, T, H, \bar{U}_{on}, \bar{U}_{of}, \bar{U}_p, \alpha, h, \rho, C) \quad (6)$$

*Beach deposition ($J_f$) or beach erosion ($K_f$) on the lower foreshore in a 12.25-hour period:*

$$J_f \quad \text{or} \quad K_f = f(\bar{S}_f, T, H, \bar{U}_{on}, \bar{U}_{of}, \bar{U}_p, \alpha, \bar{V}, \rho, \eta_r, \eta_f, D) \quad (7)$$

By a screening procedure $X_n$'s were selected which were associated with the following lag periods:

$\bar{V}$, 0–4 hours

$\bar{S}_s$ and $(\bar{M}_z)_s$, 0–4, 4–8, 8–12, 12–16, and 16–20 hours

$J_f$ and $K_f$, 0–4, 4–8, 8–12, 12–16, 16–20, and 20–24 hours

*Results.* The results of the screening procedure

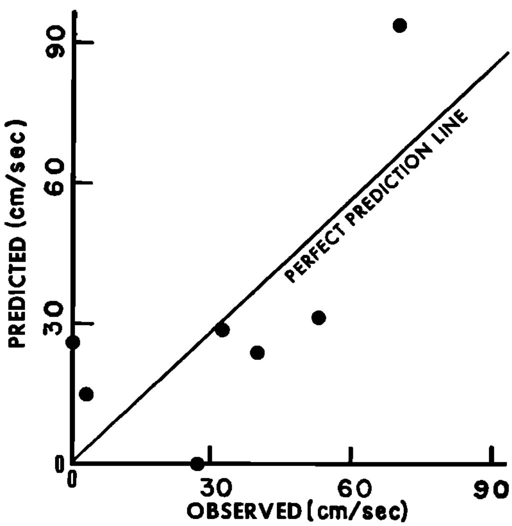


Fig. 2. Predicted versus observed values for mean longshore current velocity.

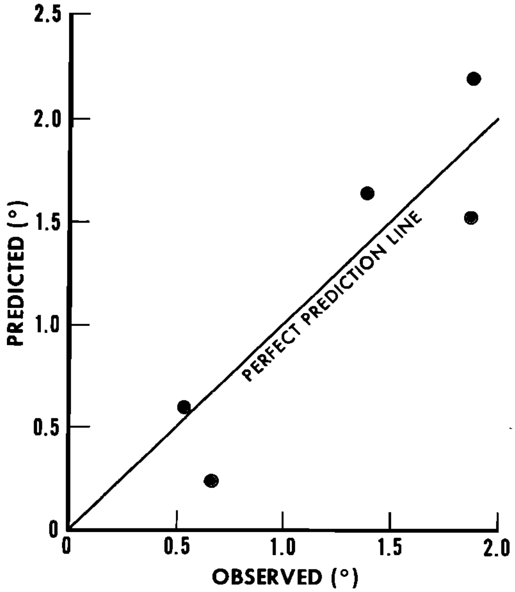


Fig. 3. Predicted versus observed values for mean bottom slope in the shoaling-wave zone.

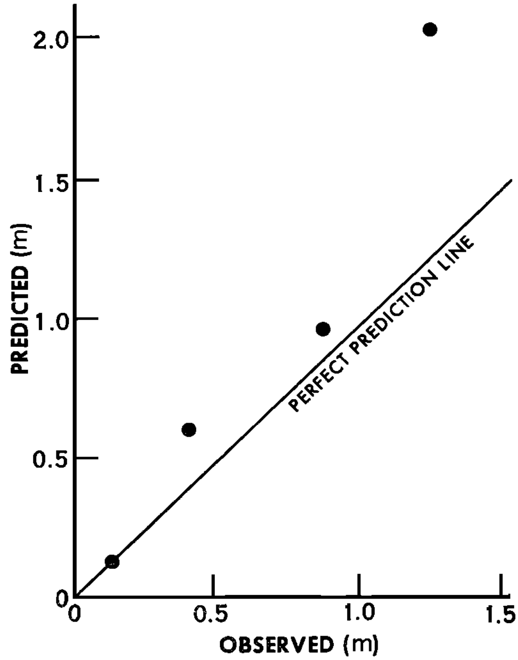


Fig. 5. Predicted versus observed values for deposition on the lower foreshore in previous 12.25 hours.

are given in Table 2. For the example, in run 1 first predictor selected by screening was $H$ with a lag of 0–4 hours and a correlation of 0.56. The second predictor selected ($T$, with a lag of 0–4 hours) increases the correlation to 0.68. Four predictors bring the correlation to 0.73; all eleven bring it to 0.76, but, as indicated in the bottom row of the table, the maximum $r$ is reached by only 8 predictors.

The five equations for the five screening runs, each containing the four predictors of Table 2, are:

$$\bar{V} = 36.0 + 0.54(H)_{0-4} - 4.88(T)_{0-4} - 1.02(C_s)_{0-4} + 3.66(\bar{S}_f)_{0-4} \quad (8)$$

$$\bar{S}_s = -5.04 + 0.82(h)_{8-12} + 0.18(\rho)_{12-16} - 0.033(C)_{0-4} + 0.024(C)_{4-8} \quad (9)$$

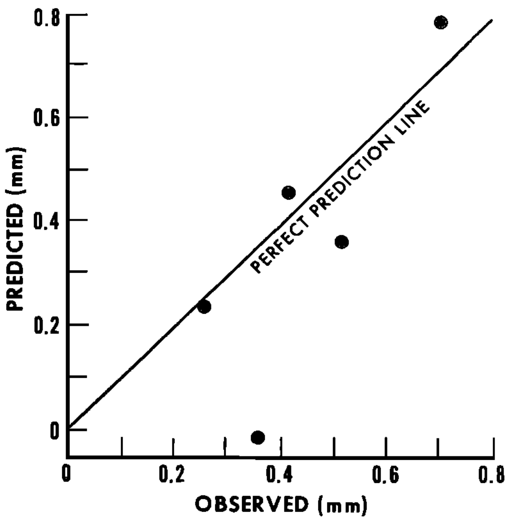


Fig. 4. Predicted versus observed values for average mean grain size in the shoaling-wave zone.

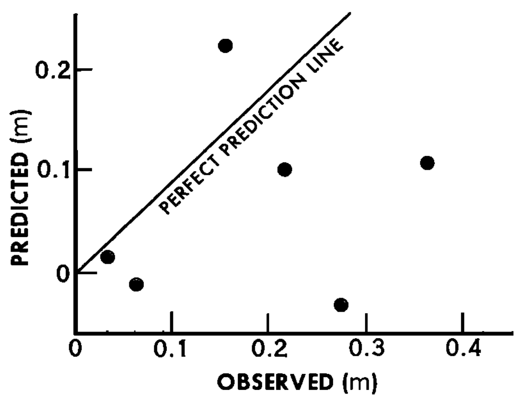


Fig. 6. Predicted versus observed values for erosion on the lower foreshore in previous 12.25 hours.

$$(\bar{M}_z)_s = -0.23 + 0.22(h)_{16-20} - 0.26(\bar{S}_s)_{0-4} + 0.01(\alpha)_{16-20} - 0.011(\bar{U}_p)_{4-8} \tag{10}$$

$$J_f = -0.787 + 0.0119(H)_{8-12} + 0.096(\bar{U}_p)_{0-4} - 0.068(\bar{U}_{on})_{8-12} + 0.0366(\rho)_{4-8} \tag{11}$$

$$K_f = -0.577 + 0.082(\bar{U}_p)_{0-4} + 0.0093(H)_{16-20} + 0.733(T)_{20-24} - 58.5(\eta_r)_{8-12} \tag{12}$$

These equations are valid where $H$ is in cm, and the remaining variables are in the units given in Table 1.

*Tests of the predictor equations.* The predictands were computed by using an independent set of data in equations 8 to 12, and the resulting values are plotted in Figures 2 to 6. The reliability of the equations may be tested by comparing the data points with the perfect prediction line. The reader is reminded that the correspondence of points with the line would in each case (cf. Table 2) be improved if the complete equation were used.

*Discussion.* Equation 9 for $\bar{S}_s$ (Figure 3) appears to be relatively good, whereas that for beach erosion, $K_f$, is relatively poor (Figure 6). The remaining equations (8, 10, 11) fall somewhere between these two qualitatively defined extremes. Before attempting to evaluate the results of this study, it is well to review just what has been done.

We wanted to predict beach erosion, longshore current velocity, and so on, for a specific beach during a specific period of time from measurements of significant environmental variables. Knowing the time lag in peak interaction between processes and responses in the system, we used a screening procedure that can scan backward in time for each predictor's maximum effect. The final equation we obtained is for a particular beach and for that particular group of possible combinations of predictors that was sampled in the course of the field measurements.

Unfortunately, the data are too limited to permit the statement that the least-squares relations developed by this analysis adequately represent the four-predictor combinations of variables that are of most significance at Virginia Beach, and it is recognized that correlation alone does not necessarily signify physical cause and effect.

A case in point is given by *Tuck* [1965] in which $(\bar{M}_z)_s$ showed strong negative correlation with $\bar{S}_s$ when the environmental conditions were normal but a weak positive correlation with $\bar{S}_s$ during violent weather. Both environmental conditions are represented in the present data set, which may lead to a relatively poor equation. A similar analysis suggests that the poor 'test' results (Figure 6) for $K_f$ is related to the noisy data generated by the two distinct environmental states. Statistical techniques should be employed for segregating the data into compatible sets for the screening analysis.

*Conclusion.* What is needed in a study of this sort is a data-acquisition system for those environmental variables considered significant that will acquire for screening analysis a large number of samples of the frequency distribution of possible combinations of predictors. The limited success of the linear model (Figures 2 to 6) suggests that such an enlarged data set would be adequate for the development of valid equations (that involve only a few variables). It is also probable that the frequency distribution of possible combinations may have to be divided into compatible units by means of appropriate techniques (such as multiple discriminant functions) before application of the screening procedure. Ultimately, nonlinear interactions will have to be dealt with in the development of equations for geophysical data of this sort.

*Acknowledgments.* We wish to thank Professor W. C. Krumbein, Northwestern University, for reviewing this paper.

Data collected in this study were funded by contract DA-49-055-CIV-ENG-64-5 from the U. S. Army, Corps of Engineers, Coastal Engineering Research Center, to the Virginia Institute of Marine Science. Data analysis was funded by the Environmental Science Services Administration.

## References

Fritts, H. C., An approach to dendroclimatology: screening by means of multiple regression techniques, *J. Geophys. Res.*, *67*(4), 1413–1420, 1962.

*Glossary of Meteorology*, p. 476, American Meteorological Society 1959.

Harrison, W., and W. C. Krumbein, Interactions of the beach-ocean-atmosphere system at Virginia Beach, Virginia, *U. S. Army, Corps of Engrs., Coastal Eng. Res. Center, Tech. Memo. 7*, 1–102, 1964.

Harrison, W., and N. A. Pore, Alternative multiregression technique for obtaining predictor equations, Addendum to *Tech. Memo. 7, U. S. Army, Corps of Engrs., Coastal Eng. Res. Center,* A1–A8, 1964.

Klein, W. H., B. M. Lewis, and I. Enger, Objective prediction of five-day mean temperatures during winter, *J. Meteorol., 16*(6), 672–682, 1959.

Miller, R. G., The screening procedure. Studies in statistical weather prediction, *Final Rept., Contract AF19(604)-1590,* pp. 86-95, Travelers Weather Research Center, Hartford, Conn., 1958.

Panofsky, H. A., and G. W. Brier, *Some Applications of Statistics to Meterology,* Pennsylvania State University, University Park, Pa., 1958.

Tuck, D. R., Jr., Major environmental variables affecting average mean grain size in the shoaling-wave zone at Virginia Beach, Virginia, Masters thesis, School of Marine Science, College of William and Mary, Williamsburg, Va., 1965.

(Manuscript received June 23, 1965;
revised September 3, 1965.)