

A REGRESSION MODEL FOR STORM SURGE PREDICTION

D. LEE HARRIS AND ALDO ANGELO

U.S. Weather Bureau, Washington, D.C.

[Manuscript received May 28, 1963; revised August 13, 1963]

ABSTRACT

The practical difficulties arising in the solution of the hydrodynamic equations for storm surges by numerical methods are reviewed. It is concluded that some of these can be avoided by means of a statistical approach if sufficient records of past surges are available. A statistical approach, based on dynamic principles, is presented and the recent literature of similar studies is reviewed.

A test is made of this statistical method by applying it to recorded storm surges on Lake Erie. The agreement between the surge values computed by the statistical method and the observed values is considered good. The test verified the statistical approach but did not lead to an operational prediction system, because of recent changes in the observational practices at some of the weather stations bordering Lake Erie.

One somewhat unexpected result was a finding that a prediction scheme based on the assumption of wind stress proportional to wind speed is not significantly inferior to one based on the assumption of a quadratic wind stress law.

1. INTRODUCTION

It is well known that severe storms are associated with abnormal water levels along the coasts of continents and large islands, and the shores of large lakes. Although below normal water levels do occur, they are generally of interest only to marine and hydro-electric power industries, and have received little public discussion. Above normal water levels flood large land areas and produce extensive loss of life and destruction of property. Most studies of the phenomenon, including this one, have concentrated on the high water due to storms. The available water level data indicate that the positive anomalies in water level tend to be greater in magnitude than the negative anomalies.

The problem has been gaining in importance during the past two decades because of the trend toward increasing population density in low coastal areas. Two developments taking place during these decades have greatly increased the possibility of understanding storm surges and of producing a satisfactory forecast system. One of these is a general improvement in the meteorological observations, and the other is the development of the high-speed stored-program computer.

The general trend in storm surge research for the past few years has been to make increasing use of electronic computers to solve the hydrodynamic equations for the water motions, subject to the condition that the wind and pressure fields were fully specified (Welander [20, 21, 22]; Usitalo [18]; Platzman [13, 15]; Hansen [2, 3]; Miyazaki, Ueno, and Unoki [10, 11]; and others). This development is a definite step forward beyond most of the earlier papers which dealt with analytic solutions to highly idealized models. But it also leads to difficulties. Care must be taken in formulating the finite difference approximations to the differential equations

to avoid computational instability (Platzman [12], Fischer [1], Shuman [17], and others).

Welander [20, 22] presented a model for storm surge predictions which he called the "Admittance Method" in 1956, and the "Influence Method" in 1961. This method requires the calculations of an admittance or influence function for each location for which predictions are desired. This function, once derived, can be used for any future predictions by calculating a definite integral involving the influence function and the wind and pressure fields. The influence function may be calculated either from theoretical considerations and the assumption inherent in any numerical integration of the linearized hydrodynamic equations for two-dimensional horizontal motion in a fluid with a free surface, or from an analysis of past data. He pointed out that the latter procedure would avoid the necessity of specifying the coefficients relating the wind and stress fields as these would be implicitly provided by the analysis of the observational data. Only one sample prediction by the method has been published and the author does not state whether it is for independent or dependent data. One goal of the work reported in this paper is a clarification and extension of the original work by Welander.

Wilson [23] presented a scheme for computing the storm surge in New York Bay. This model was based on an analysis of two hurricanes and two extratropical storms and contains nine regression coefficients obtained by a multivariate analysis. Four constants had to be determined subjectively before the regression analysis was possible. Most of the data used in the regression analysis were obtained from a subjective analysis of the meteorological data. Wilson did not apply his model to any storms not used in determining the constants required by the model.

Harris [4] re-examined Wilson's model and showed that the subjectively determined constants were somewhat controversial. He derived a new model, similar to Wilson's but with fewer subjectively evaluated constants. The new model with fewer empirical constants fits the data presented by Wilson about as well as Wilson's model. All of the calculations showed encouraging similarities to the observed data, but it appeared that more data would have to be used, and improvements in the theory might be needed before the technique would have operational value. Wilson [25], in his reply to Harris [4], made some minor modifications in his model, used a new value for one of the other subjectively evaluated constants, and improved the agreement between the computed and observed surge for this storm. The improved equation was not further tested on new independent data.

Harris [5] re-examined the problem of the empirical prediction of storm surges. By taking into account the limitations of the meteorological data that can be made available, he showed that it is possible to derive a regression equation that contains all of the information that can be put into a numerical solution of the linearized hydrodynamic equation. The coefficients in this equation, like the influence function of Welander, can be evaluated either from theoretical considerations or by an analysis of past data. Welander and Harris both use the linearized two-dimensional hydrodynamic equations to develop relations between the wind and pressure fields and water level changes in a rather general way. This paper is a report of tests of Harris' model by application to past data.

2. DISCUSSION OF THE FUNDAMENTAL EQUATION

It was assumed by Harris [5] that storm surges are governed by the equations

$$\frac{\partial U}{\partial t} - fV + gD \frac{\partial h}{\partial x} + \frac{K}{D} U = -\frac{D}{\rho_w} \frac{\partial p_a}{\partial x} + (1-\theta)^x \tau_s \quad (1)$$

$$\frac{\partial V}{\partial t} + fU + gD \frac{\partial h}{\partial y} + \frac{K}{D} V = -\frac{D}{\rho_w} \frac{\partial p_a}{\partial y} + (1-\theta)^y \tau_s \quad (2)$$

$$\frac{\partial h}{\partial t} + \frac{\partial U}{\partial x} + \frac{\partial V}{\partial y} = 0 \quad (3)$$

where U and V are the transports along the x and y axes; f is the Coriolis parameter, $f=2\omega \sin \phi$; ϕ is latitude, ω is the earth's angular speed; h is the height of the free water surface above its equilibrium position; g is the acceleration of gravity; $D=D(x, y)$ is the equilibrium depth of the fluid; p_a is the atmospheric pressure; ρ_w is the density of the water; $^x \tau_s$ and $^y \tau_s$ are the surface wind stresses along the x and y axes; K and θ are friction coefficients. A derivation is given by Welander [22]. A very similar or more restrictive set of equations has been employed in almost all quantitative studies of storm surges (Welander [22]; Hansen [3]).

In order to solve these equations it is necessary to specify the pressure and wind stress fields and a value for

the friction coefficient K . The wind stress over the water cannot be observed directly by any standard procedure and is generally expressed as a simple function of the wind speed, S , usually by an equation of the form

$$\tau = AS^d \quad (4)$$

Thus two parameters must be specified for this relation. A value of 2 is usually chosen for d , but some workers prefer other values. The coefficient A is a function of the surface roughness, the turbulent state of the air, and the reference height for wind speed. This coefficient is in no sense a constant, but as the laws governing its variation are not well understood, it is usually treated as a constant in dynamical calculations. Wilson [23, 24] gives a summary of more than 40 determinations of A , adjusted to a reference height of 10 m. and a value of 2 for d . The values vary from 1.5×10^{-3} to 4.0×10^{-3} with a mean of 2.36×10^{-3} for strong winds (about 40 kt.). The variation is even greater for light winds.

There still remains the problem of determining the continuous wind field over the water in time and space from the available meteorological data. Two approaches are available. One may specify the continuous wind field over the water by interpolating between wind observations (usually over land). In this case an additional correction factor is usually needed to adjust for a difference between the surface roughness near the anemometer and that over open water. Another method is to interpolate between the available pressure data, using some theoretical relation between wind and pressure gradients to determine the wind speed. A correction factor is required in this case also to account for the effects of surface roughness on surface wind speed. There are several logical methods of constructing interpolation procedures varying from the subjective drawing of isopleths through plotted data, as in the construction of isobars on a weather map, to some highly sophisticated mathematical models. In essence, each of these involves the specification of the form of wind and pressure patterns in a storm, and that the formal storm model, given implicitly by the interpolation procedure, be fitted to the observational data. Most mathematical interpolation procedures are based on the assumption that the field being interpolated is continuous and smooth. It is well known from the analyses of synoptic weather charts that this assumption is rarely satisfied during stormy conditions. It is to be expected that better results could be achieved by an interpolation procedure which can utilize some of the principles of synoptic analysis.

A full analysis of the coefficient K , used to relate the bottom friction to the mean flow, would show that this depends on the bottom roughness, turbulence of the water, and other variables. This will not be attempted in this report because at the present time we do not know how to use this analysis to improve the results obtained in this paper.

A numerical solution of the problem requires a few additional decisions such as: 1. the finite difference approximations to the differential equations; 2. the finite difference approximations of the boundary equations; and 3. the length of the space and time increments. Although these decisions can be guided by theoretical considerations and independent empirical data, there is no certain means for determining the best choice.

It is shown by Harris [5] that the initial conditions will become relatively insignificant at some relatively low value of the time and can be neglected if computations are started in a relatively quiet interval before a significant disturbance. This is true because of the damping influence of friction. If this condition is satisfied, it is possible to solve equations (1)–(3), subject to the assumptions mentioned above in the form

$$h(x_0, y_0, t) = \sum a_{i,j,k}(x_0, y_0) F_{j,k}(t - i\Delta t) \quad (5)$$

where Δt is the time interval between meteorological observations; i is the number of intervals between the observation and the time t ; j is an index of the observation station; k is an index indicating the type of observation; $F_{j,k}(t - i\Delta t)$ is the meteorological factor of type k , from station j at time $(t - i\Delta t)$; $a_{i,j,k}(x_0, y_0)$ are coefficients which depend on the position (x_0, y_0) for which predictions are desired as well as the indices i, j, k ; and $h(x_0, y_0, t)$ is the predicted surge at time t and location (x_0, y_0) .

Equation (5) contains all of the information about the water level at (x_0, y_0) inherent in equations (1)–(3) and the available meteorological data. The a 's can be evaluated from theoretical considerations and the assumptions required for a numerical solution of equations (1)–(3). However, these assumptions are not unique and many of them must be justified *a posteriori* by comparison of the computed heights with observed values. The possibility of obtaining the coefficients directly from empirical data should not be overlooked.

The coefficients in equation (5) depend on several distinct aspects of the problem:

1. The inertia of the water. There is a delay between the application of force to an open water area and the consequent change in water level at the shore. That is to say, the wind during one hour influences the water level for several following hours. This is the source of Welander's "influence functions" and is the aspect of the problem stressed in his paper.

2. The coefficients for wind stress and bottom friction and the factor which relates the wind speed over land or the theoretical wind speed to that over water. This was also considered by Welander.

3. The interpolation function by which the continuous wind field over the water is determined from data specified at a few points. Wilson gave more consideration to this feature than did Welander.

4. The finite difference formulation of the problem.

It should be recognized that the hydrodynamic equations are concerned only with the first of the above aspects

of the problem. The second and third arise from our inability to provide the precise empirical data demanded by the mathematical formulation of the problem, and the fourth from the limitations of the available methods for solving the problem. There remains the possibility that equations (1)–(3) may not be entirely adequate for the full range of problems we wish to consider.

Harris [5] has shown that an empirical determination of the coefficients may eliminate the effects arising from the finite difference formulation of the problem, simulates the effect of the best choices for wind stress and friction coefficients, and represents the best choice of interpolation functions for the *set of storms considered in the derivation of the coefficients*. The exponent in the wind stress formula must still be specified as an arbitrary assumption.

Platzman [15] has derived the storm surge equations in a form that is significantly different from equations (1)–(3). He considers the problem of determining the representative wind field in much the same manner but chooses somewhat different means of obtaining a solution. The present writers believe that more research is needed before it will be possible to determine which is the better procedure. Platzman's paper is not being discussed in detail because, unlike those of Welander and Wilson, it follows a different line of development. It is worth noting that Platzman considers many of the storms discussed in a later section of this paper.

3. PRACTICAL APPLICATION OF THE BASIC EQUATION

The forcing terms in equations (1) and (2) are wind stresses along two orthogonal axes and the pressure gradients along the same axes. However, if data are supplied for two or more stations, pressure gradients along the axes may be expressed implicitly by the coefficients assigned to the pressure values from the different stations. If data are available from only one station, the pressure gradient must be considered as being determined between that station and some representative normal value over the water, away from the shore. Thus it is sufficient to consider only three types of forcing terms, wind stress along two orthogonal axes and pressure. Thus the index k takes on three values. The number of observation stations available is always limited and this determines the upper limit of the index j . The range of the index i is not so clearly determined, but because of the presence of a dissipation term in the model, one can be sure that some finite value will be sufficient.

If the number of observations is as great as the number of coefficients to be determined, it is possible to determine the coefficients from the observations. If the theory and observations were perfect, this is all that would be required. However, it has been shown above that there are many imperfections in the theory. The observations likewise are less than ideal. Thus it is desirable to use many more observations than coefficients, and to obtain a solution in a least squares sense. This is equivalent to

a problem in multivariate analysis and the analysis already carried out by statisticians is available for the solution of this dynamical problem.

A high-speed computer of large capacity must be used for obtaining the solution. The problem is one of solving N equations for M unknowns. The number of multiplications required, and hence the cost, is roughly proportional to $M!$. On the other hand, the meteorological data are highly redundant. A few terms in equation (5) carefully selected, may be expected to have almost as much information as the complete set. A rather small number of predictors may be sufficient for determining the wind stress and frictional coefficients. This will also be true for the interpolation function *if all storms to be considered with a given set of the regression coefficients are highly similar*. If dissimilar storms must be grouped together, a much larger set of coefficients must be used to give additional degrees of freedom to the interpolation formula. Additional points in time (many time lags) may be used to increase the information about the horizontal structure of the storm if the storm is moving but shows little change in shape. If the motion is critically damped, a few time lags should be sufficient for a determination of the influence function. If the damping is less than critical, oscillations due to the inertia of the water will be important and a much longer time period, equal to two or more times the natural period of the basin, may have to be considered.

It is theoretically possible to avoid the problems of an inadequate interpolation formula and an inadequate description of the influence function by considering all of the data available for the numerical integration of the equations for each storm. This may require several hundred coefficients. However, for idealized storms and critically damped basins, five to ten coefficients may provide as much forecasting skill for water levels as can be obtained from equations (1)-(4) and the available meteorological data. The problem involved in the practical application of the regression technique is the determination of an optimum selection of terms in equation (5) to reflect adequately both the interpolation process and the influence function.

It should be clear from the above discussion that the statistical determination of the regression equations requires only that the observation stations be fixed in space, and that the same pattern of time lags be used for all prediction periods. The observation stations may be real weather stations or grid points on a map. There is no requirement for equal spacing of observations in either space or time. The following work has been restricted to a consideration of hourly weather reports from airport weather offices, mainly because these data are already available on punch cards.

The remainder of this paper is devoted to a discussion of this problem with the aid of examples. It should be mentioned that many of the implications of the procedure described above were not clear until many more examples than are presented here had been considered.

4. STORM SURGES AT BUFFALO, N.Y.

Pronounced storm surges occur at Buffalo, N.Y., more frequently than at any other United States port. Irish and Platzman [7] report that the water level at Buffalo exceeded that at Toledo by 6 ft. or more on 76 separate occasions during the period January 1940 through December 1959. Irish and Platzman give a nonexhaustive list of ten earlier investigations of surges on Lake Erie, beginning with Henry [6]. In most cases a low center passed north of the Lake bringing strong westerly to southwesterly winds across the Lake, inducing a set-up with high water to the east. In a few cases the Low moved south of the Lake, bringing easterly winds and higher water in the western part of the Lake. In most cases a cold front, extending from north to south crossed the Lake from west to east. The frontal passage was frequently, but not always, accompanied by a pronounced wind shift. Figure 1 shows a sequence of synoptic charts for a typical disturbance with a frontal wind shift. Irish and Platzman have published two other examples. Thirty examples of the water level variations at Buffalo associated with storm passages are shown in figures 3 and 4. The records frequently display the appearance of a series of damped oscillations following each major disturbance. Verber [19] quotes the results of eleven determinations of the natural period of Lake Erie. These values vary from 12.2 to 18 hr., with values near 14 hr. being derived most frequently.

Hourly wind and pressure observations from six weather stations near Lake Erie and hourly lake level observations for two stations were available for this analysis. The locations of these stations are shown in figure 2. The coefficients required for equation (5) were derived by means of a screening program (Miller [9]; Klein, Lewis, and Enger [8]; Harris [5]) which can consider up to 92 independent variables, and the records for more than a million observations. Data for a total of 1900 hr. were examined. Each weather observation provided three items of information that could be used in the prediction. It was not possible to consider all of the available data at one pass with this program. Therefore, for the first trial it was assumed that to a first approximation the effect of the wind field over the lake at any time could be specified by the wind at Buffalo and Toledo at the current and preceding 20 hourly observations. The pressure gradient effect was represented by pressure observations at the present and three preceding 3-hourly observations. The influence functions defined by these coefficients were combined subjectively with hydrodynamic theory to estimate the influence functions for the remaining stations. New sampling procedures were selected with the hope of maximizing the information which could be obtained from a set of 92 variables selected from all of the available data.

In the first run, 100 hourly water levels were considered for each storm. In many of the cases the duration of the disturbance was considerably shorter than 100 hr., and

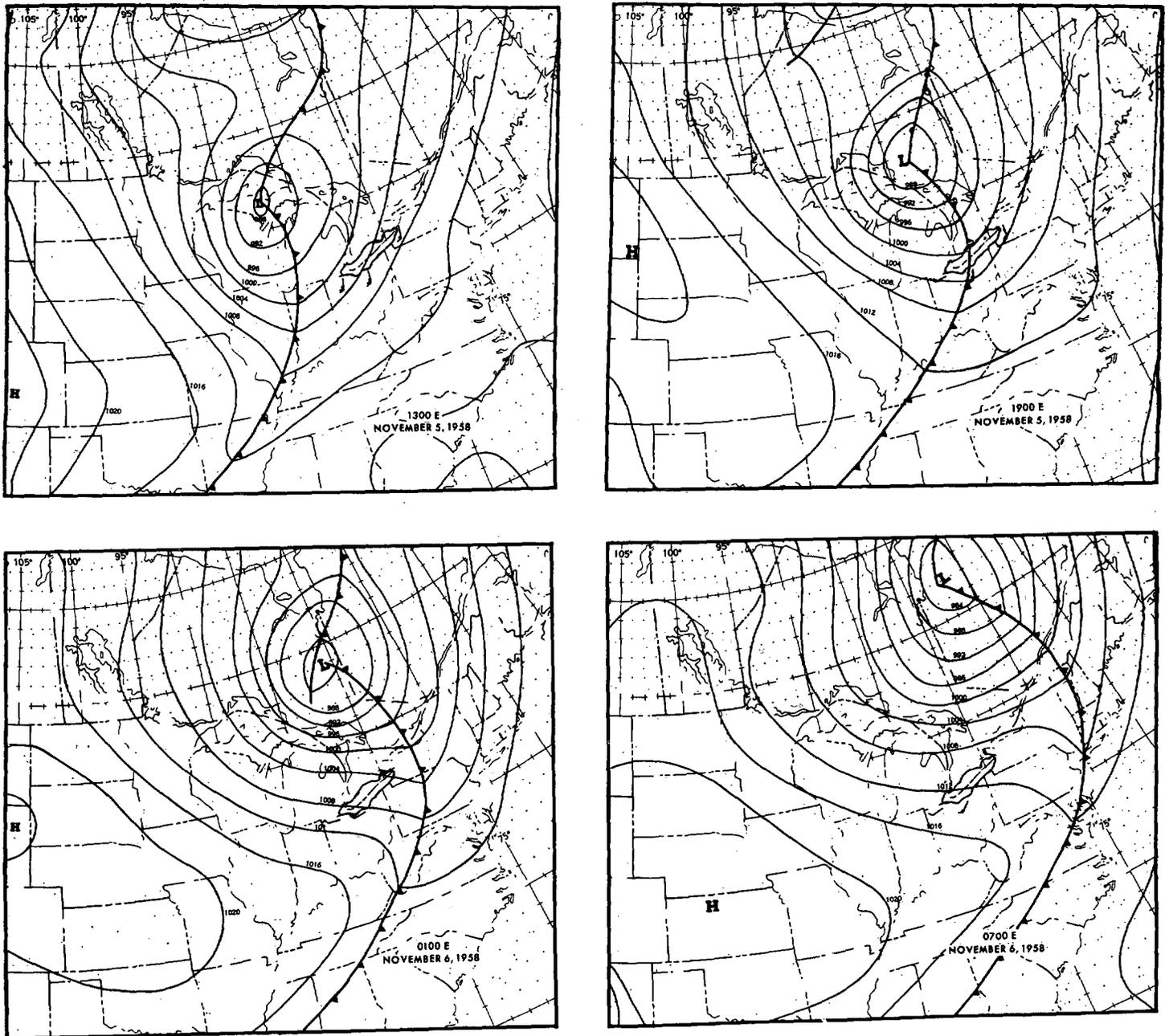


FIGURE 1.—Synoptic charts for November 5-6, 1958, showing a typical situation for the development of storm surges on Lake Erie.

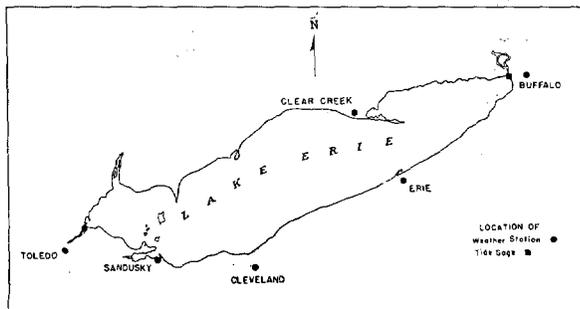


FIGURE 2.—Map of Lake Erie showing the locations from which observations were used in this study.

the high correlations obtained could have been due in large measure to the near perfect correlations obtained during quiet periods. To avoid this possibility in the second run, no more than two or three undisturbed hours were permitted in the beginning of each storm period, and the storm period was terminated with the first hour at which it could be determined without looking at any subsequent data that the worst of the disturbance had passed. In some cases this included the first of the secondary oscillations following the major peak. Only 641 water level observations, selected from the most disturbed part of the record, were included in this second run. But meteorological data from five stations were included. Restricting the period analyzed would be ex-

TABLE 1.—Summary of correlations obtained in developing a prediction equation for storm surges at Buffalo

Data	1 independent variable	8 independent variables	92 independent variables
<i>Linear stress law</i>			
Buffalo and Toledo only.....	0.69	0.78	0.81
Five stations*.....	.70	.80	.85
<i>Quadratic stress law</i>			
Buffalo and Toledo only.....	.72	.81	.83
Five stations*.....	.79	.84	.87

*Toledo, Sandusky, Cleveland, Erie, and Buffalo data were used. Clear Creek data were available but were not used because the station had been closed and it was hoped at the time of this run that the results could be used in an operational forecast model.

pected to reduce the overall correlations obtained but increase the correlations during the important period of rising water levels. Including observations from additional locations was expected to improve the overall correlations and the agreement between observed and computed water levels during the period of maximum disturbance. Both the overall correlations and the predicted peaks were improved by the second run.

Both linear and quadratic wind stress laws were investigated. The differences in the derived correlations were trivial, but tended to favor the quadratic stress law. The results, summarized in table 1, are for the equations involving only one meteorological observation (of any type and time lag), eight observations, and all of the meteorological data considered.

The equations with 8 variables and with 92 variables derived from an analysis of the data at all stations were used to make calculations for periods of approximately four days for each storm period for the 19 storms used in the analysis. Calculations could not be made for the first few hours of each period analyzed because meteorological data for a period of several hours are required for each calculation, and no new data were secured to extend the period of calculation.

A comparison of the water levels computed by the complete regression equations for both the linear and quadratic wind stress laws, and the observed water levels for the 19 storms used in deriving the regression coefficients is shown in figure 3. A similar comparison for 11 additional storms is given in figure 4.

In general, both sets of predictions follow the observations reasonably well, and tend to depart from the predictions in the same places. Some of the secondary oscillations are predicted surprisingly well, when it is considered that no serious attempt was made to provide for the possibility of seiche motion in the derivation of the prediction equations.

The results of the two preceding and two following figures are summarized in figure 5. Here the observed extremes, lows and highs, are plotted as functions of the predicted values, with phase shifts of as much as 3 hr. permitted, for each of the long prediction equations described above and the truncated equations based on 8 coefficients.

The correlations between observed and predicted values, when treated in this way, are 0.97 for both long equations and 0.96 for both short equations as applied to the data going into the system. For independent data, not considered in deriving the regression equations, indicated by crosses on this chart, the long quadratic equations give a value of 0.96; the long linear equations give 0.93. The truncated quadratic equation gives 0.89 and the truncated linear equation gives 0.92. Phase shifts were required in fewer than half the cases and the correlations decreased by no more than 0.02 when the comparisons were made at the time of either the observed or the predicted peaks.

These data represent the wind effect on the Lake for speeds varying from about 10 to 55 kt. The close agreement between the results obtained for the quadratic and linear stress laws, and the high correlations obtained with both for the dependent data were unexpected. The close agreement between the results of the linear and quadratic laws suggests that some exponent between 1 and 2 would have been better than either. Preliminary efforts to find the favored exponent have not been successful. A comparison of the correlations obtained with independent data suggests another hypothesis. That is, although neither the quadratic nor linear law is entirely correct, both are good approximations of the correct law over the ranges of wind speeds considered, but the quadratic law is the better of the two as a physical law. However, the continuous wind field must be interpolated from relatively few measurements, and thus is sure to contain errors, which may be designated by *E*. For the linear stress law, the relative error is given by

$$E_{(linear)} = \frac{V + \Delta V - V}{V} = \frac{\Delta V}{V};$$

the relative error for the quadratic stress is given by

$$E_{(quadratic)} = \frac{(V + \Delta V)^2 - V^2}{V^2} = \frac{2V + (\Delta V)^2}{V} \approx \frac{2\Delta V}{V}$$

Thus the relative error of interpolation for the quadratic stress law is approximately twice as great as that for the linear stress law. This hypothesis appears to be supported by the observation that the prediction equation based on the quadratic law held up better for independent data than that based on linear law when a maximum of information was available. but decayed more rapidly when the amount of information available for the prediction was decreased. The basic scheme has been tested for several other ports with similar results.

Operational forecasts of water level change due to the wind must depend in part on predicted winds. It is unlikely that the predicted wind field can be specified with the degree of accuracy obtained with the maximum amount of observed data used in these tests. Therefore, it is assumed as a working hypothesis that operational storm surge predictions should be based on a linear wind stress law. It is considered likely that the same working hypothesis will be applicable to other oceanographic problems,

STORM SURGE AT BUFFALO

DEPENDENT DATA

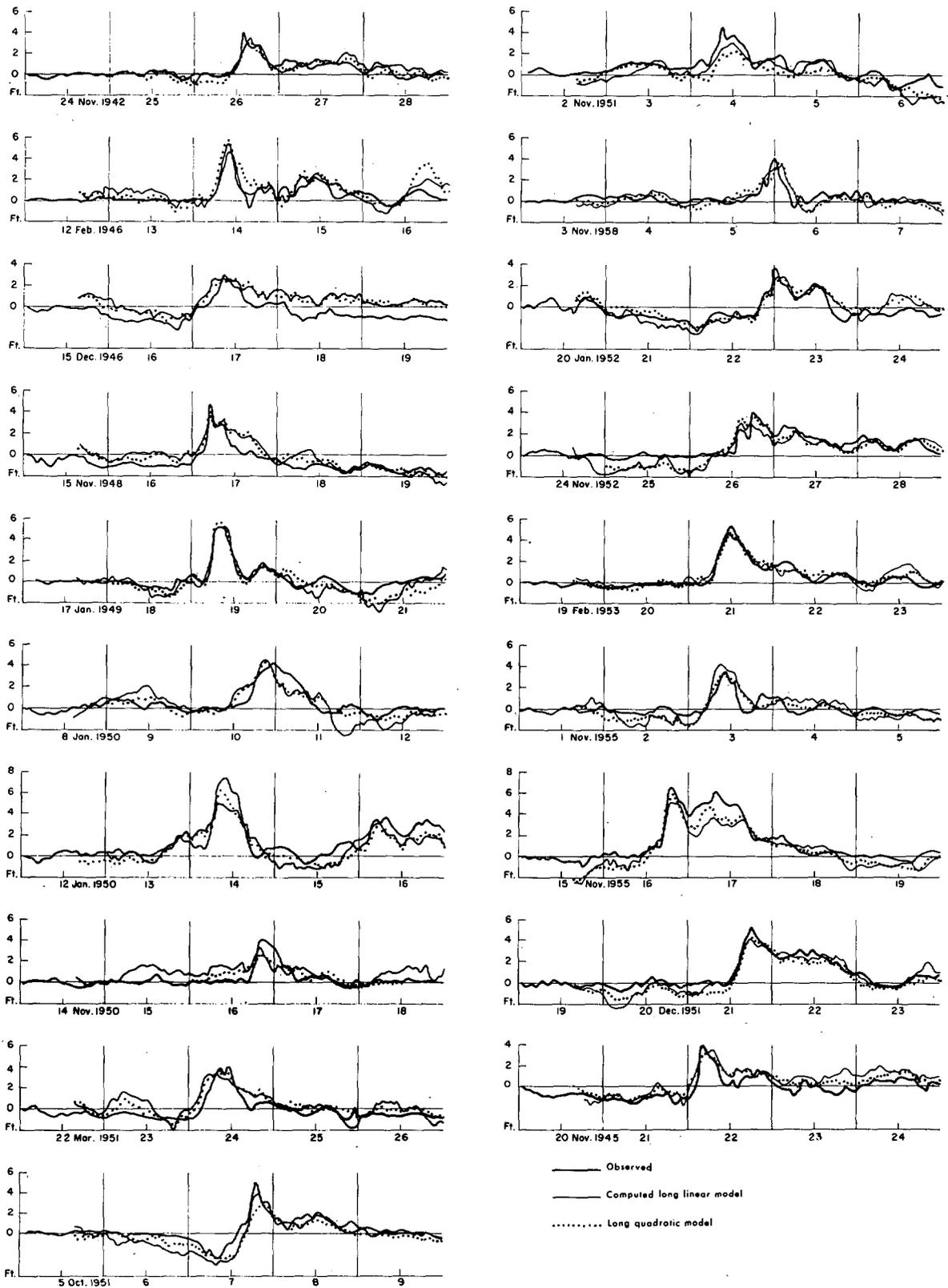


FIGURE 3.—Comparison of observed and computed surges at Buffalo, N.Y., for both long regression equations and dependent data.

STORM SURGE AT BUFFALO

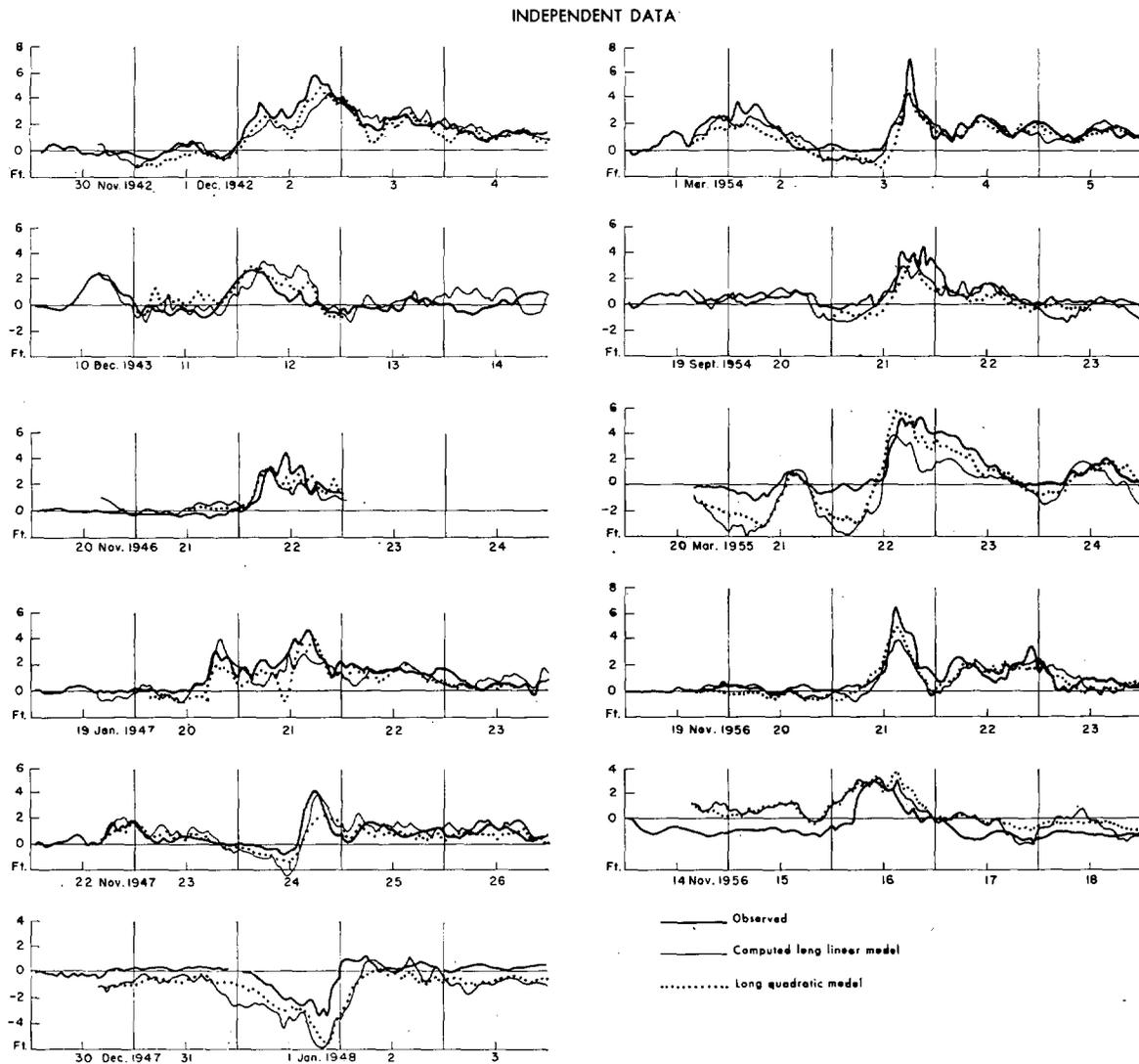


FIGURE 4.—Comparison of observed and computed surges at Buffalo, N.Y., for both long regression equations and independent data.

such as sea and swell forecasting, which must be based on a poorly specified wind field over the sea. No conclusions can be reached as yet about the true stress law which should hold where the wind field is fully specified.

These experiments were conducted to test the concepts advanced by Harris [5]. The results are believed to justify the suggested procedures. It was expected that an objective forecasting scheme for the storm surges at Buffalo would be a useful by-product of the study. Unfortunately, this expectation was not realized. Shortly after the last storm period considered in this study, the anemometer at Toledo was lowered from 72 to 20 ft. above the ground; that at Cleveland from 88 to 20 ft., and that at Buffalo from 96 to 20 ft. The station at Clear Creek was closed. It may be feasible to determine correction factors for the new anemometer elevations when more data have been collected.

A comparison of the water levels computed by the complete linear model, the short linear equation with only 8 terms, and the observations is presented for the dependent data in figure 6, and the independent data in figure 7.

5. STORM SURGES AT TOLEDO, OHIO

The positive storm surges (water level above normal) at Buffalo are almost always accompanied by negative surges (water level below normal) at the western end of the lake. Prediction formulae for Toledo, Ohio, can be developed from the same data and in much the same manner as just described for Buffalo. The first run of the screening program for Toledo was identical to the first run for Buffalo with the exception that predictions were sought for Toledo. It soon became apparent that

LAKE ERIE STORM SURGES - BUFFALO

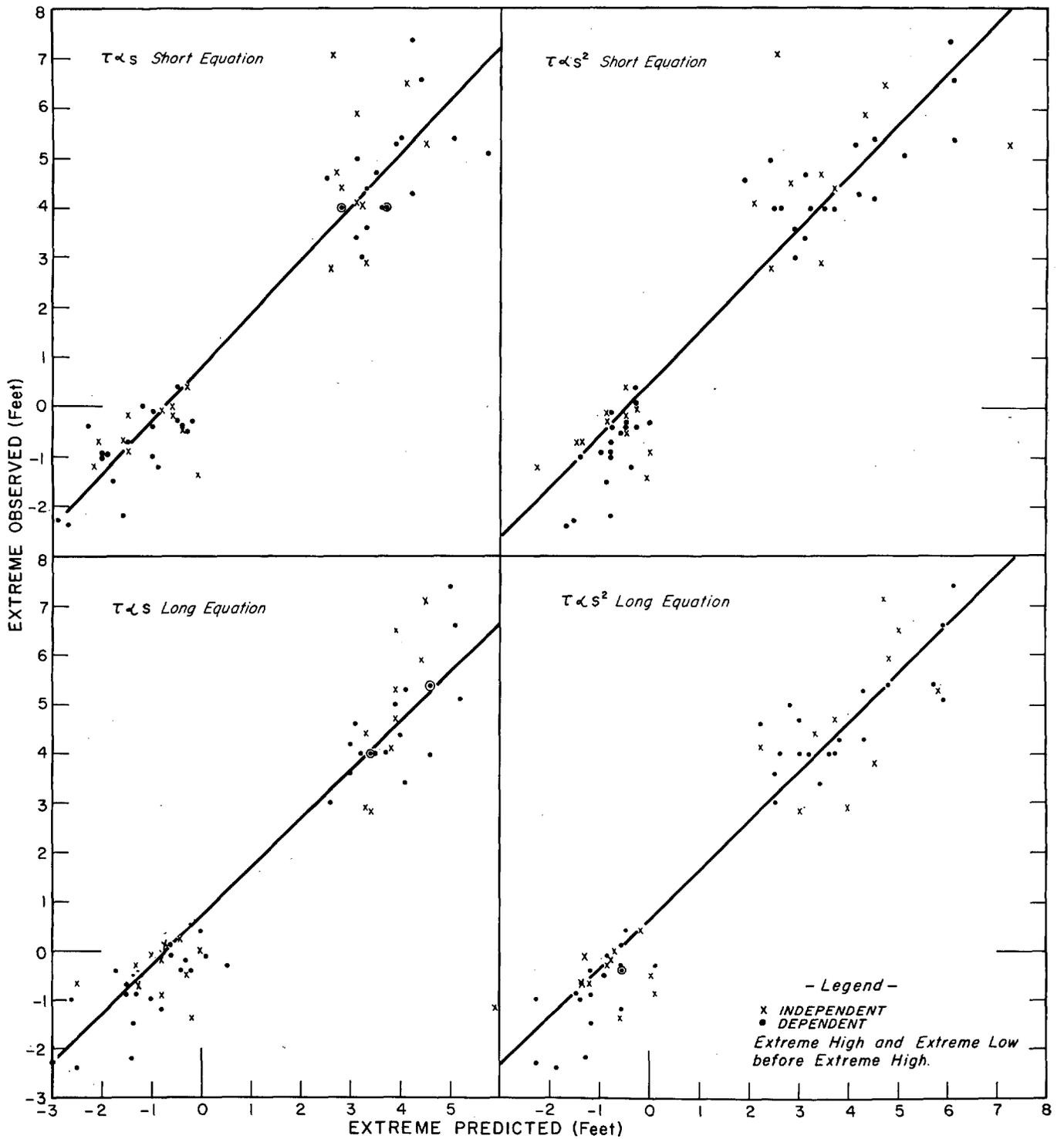


FIGURE 5.—Comparison of the observed extremes in water level at Buffalo, N.Y., with the extremes predicted by the regression equations discussed in this paper.

the release of a mound of water in the eastern part of the lake, by a shift in the wind direction or a decrease in wind speed, led to an increase in the water level in Toledo a few hours later by a process that is independent of the direct action of the wind. This may be accounted for by doubling the length of the influence period or by using the lake level at Buffalo a few hours earlier as a predictor for the lake level in Toledo. The latter course was adopted as being the most direct approach and the one which would provide the most information for the amount of data examined. Fifteen storms with 100 observations each were used in this test.

By selecting potential predictors for the second screening run for Toledo, an influence function was estimated for Cleveland and lags selected from all three stations which would permit the optimum description of the influence functions with the number of independent predictors available in the program. The period of record examined was reduced as in the Buffalo case, to permit no more than two or three hours before the disturbance started, and to terminate the storm period as early as it was possible to determine from the records up to a given hour that the first high water following the deep low had occurred. Only 389 hr. of data were examined in the second run.

Again the second run gave a slight increase in the overall correlations and a clear improvement in the agreement between the observed and predicted extremes. A summary of the correlations computed is presented in table 2. Only trivial differences were found between the correlations obtained by the linear and quadratic wind stress laws.

The water levels computed by the two long equations derived in the second run are compared with the observations for the dependent data in figure 8, and for storms not used in deriving the prediction equations in figure 9. Missing data, necessary for the Toledo predictions, prevented consideration of all of the storms examined in the Buffalo study. The changes in anemometer heights mentioned above prevent the practical use of the derived equations and for this reason it was not considered worthwhile to obtain the missing data. The two short equations are compared with the predictions in figures 10 and 11.

TABLE 2.—Summary of correlations obtained in developing a prediction equation for storm surges at Toledo, Ohio

Data	1 independent variable	8 independent variables	92 independent variables
<i>Linear stress law</i>			
Buffalo and Toledo meteorological data only	0.78	0.83	0.85
Data from 3 stations	*.71	.89	.92
<i>Quadratic stress law</i>			
Buffalo and Toledo meteorological data only	.76	.84	.86
Data from 3 stations	*.68	.89	.92

* The correlation with only one independent variable was less in the second run than in the first, because the second run was restricted to the most disturbed portion of the record.

A comparison of the predicted and observed extremes at Toledo is presented in figure 12. The graphs are inverted to show the negative values in the upper right hand corner in order to facilitate comparison with figure 5 that presents similar data for Buffalo.

6. APPLICATION OF THE REGRESSION TECHNIQUE TO EAST COAST PORTS

The regression technique has been applied to the prediction of storm surges caused by extratropical cyclones at Atlantic City and Sandy Hook, N.J. It cannot be applied to the prediction of hurricane storm surges as yet because there are insufficient data for developing the interpolation formulae needed to estimate the complete wind field in a hurricane from the records of a few anemometers, as these interpolation formulae relate to a particular location along the coast. The results of this study are being reported by Pore [16]. They may be summarized briefly by noting that the correlations achieved are about as good as those reported here. The results obtained with the linear and quadratic stress laws are nearly the same, but favor the quadratic law slightly when a maximum of data is supplied, and favor the linear law with the amount of data normally usable in an operational forecast.

7. SUMMARY AND CONCLUSIONS

It has been shown that a screening type regression analysis of observational data can be used to obtain a solution to at least one type of linearized hydrodynamic problem if a sufficient backlog of data is available for analysis and if the dynamics of the problem are carefully considered in selecting the predictors to be tested. If the calculation must be based on empirical observations and to some degree on empirical laws, the solution obtained with the regression technique may be equivalent or superior to a prediction based on the same data obtained by the direct integration of the hydrodynamic equations.

It has also been shown that calculations based on a linear wind stress law may be as good or better than similar calculations based on a quadratic wind stress law whenever it is necessary to estimate the continuous wind field from a restricted amount of observed data.

REFERENCES

1. G. Fischer, "Ein numerisches Verfahren zur Errechnung von Windstau und Gezeiten in Randmeeren," *Tellus*, vol. 11, No. 1, Feb. 1959, pp. 60-76.
2. W. Hansen, "Theorie zur Errechnung des Wasserstandes und der Strömungen in Randmeeren nebst Anwendungen," *Tellus*, vol. 8, No. 3, Aug. 1956, pp. 287-300.
3. W. Hansen, Editor, *Proceedings of the Symposium on Mathematical-Hydrological Methods of Physical Oceanography*, Institut für Meereskunde der Universität Hamburg, 1962, 424 pp.
4. D. L. Harris, Comments on "The Prediction of Hurricane Storm-Tides in New York Bay," by Basil W. Wilson, U.S. Army Corps of Engineers, Beach Erosion Board *Technical Memorandum* No. 120-A. 1961, pp. 1-12.

STORM SURGE AT BUFFALO

DEPENDENT DATA

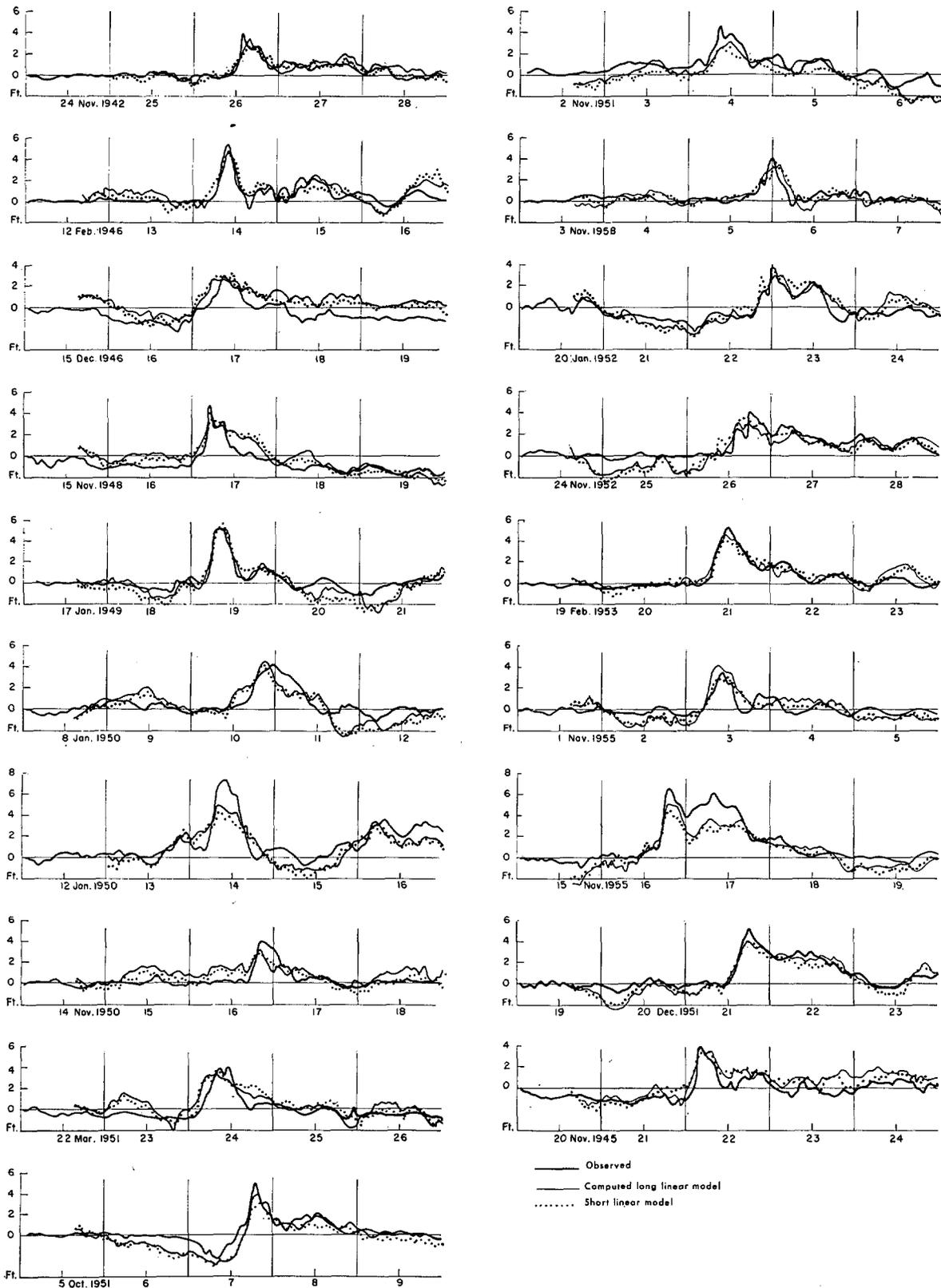


FIGURE 6.—Comparison of the observed and computed surges at Buffalo, N. Y., for both the complete and the short linear stress models with dependent data.

STORM SURGE AT BUFFALO

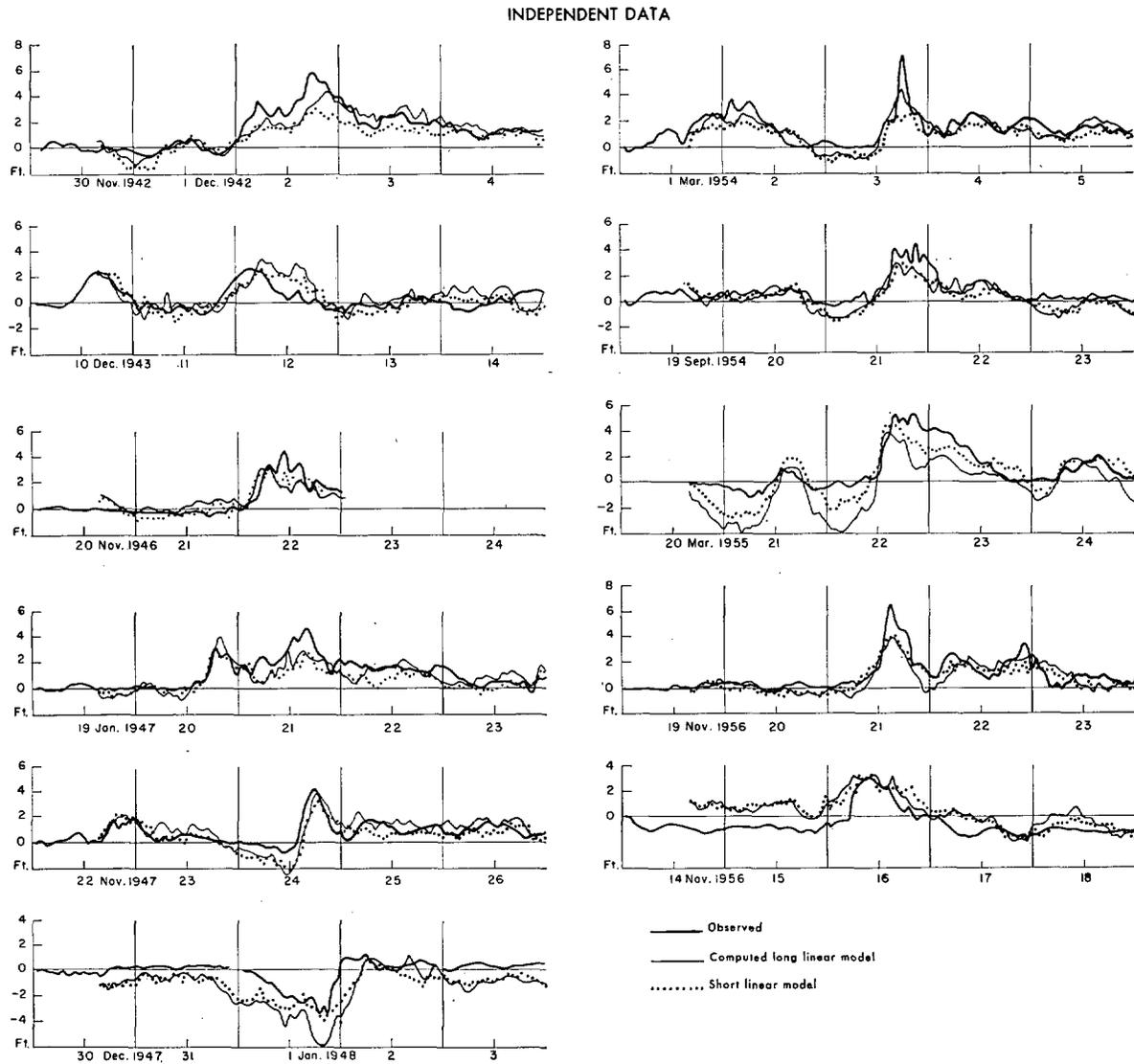


FIGURE 7.—Comparison of the observed and computed surges at Buffalo, N.Y., for both the complete and short linear stress models with independent data.

5. D. L. Harris, "The Equivalence Between Certain Statistical Prediction Methods and Linearized Dynamical Methods," *Monthly Weather Review*, vol. 90, No. 8, Aug. 1962, pp. 331-340.
6. A. J. Henry, "Wind Velocity and Fluctuations of Water level in Lake Erie," *Bulletin J*, U.S. Weather Bureau, Washington, D.C., 1902, 22 pp. 12 charts.
7. S. M. Irish and G. W. Platzman, "An Investigation of the Meteorological Conditions Associated with Extreme Wind Tides on Lake Erie," *Monthly Weather Review*, vol. 90, No. 2, Feb. 1962, pp. 39-47.
8. W. H. Klein, B. M. Lewis, and I. Enger, "Objective Prediction of Five-day Mean Temperature During Winter," *Journal of Meteorology*, vol. 16, No. 6, Dec. 1959, pp. 672-682.
9. R. G. Miller, "A Statistical Procedure for Screening Predictors in Multiple Regression," *Studies in Statistical Weather Prediction*, The Travelers Weather Research Center, Final Report on Contract AF 19 (604)-1590, 1958, pp. 86-95.
10. M. Miyazaki, T. Ueno, and S. Unoki, "Theoretical Investigation of Typhoon Surges Along the Japanese Coast," *Oceanographical Magazine*, ol. 13, No. 1, Oct. 1961, pp. 51-75.
11. M. Miyazaki, T. Ueno, and S. Unoki, "Theoretical Investigation of Typhoon Surges along the Japanese Coast (II)," *Oceanographical Magazine*, vol. 13, No. 2, Mar. 1962, pp. 103-117.
12. G. W. Platzman, "The Computational Stability of the Meteorological Prediction Equations," Department of Meteorology, University of Chicago, 1953, 30 pp.
13. G. W. Platzman, "A Numerical Computation of the Surge of June 26, 1954 on Lake Michigan," *Geophysica*, vol. 6, Nos. 3-4, 1958, pp. 407-438.
14. G. W. Platzman, "The Computational Stability of Meteorological Prediction Equations," Department of Meteorology, University of Chicago, May 1958.

STORM SURGE AT TOLEDO
DEPENDENT DATA

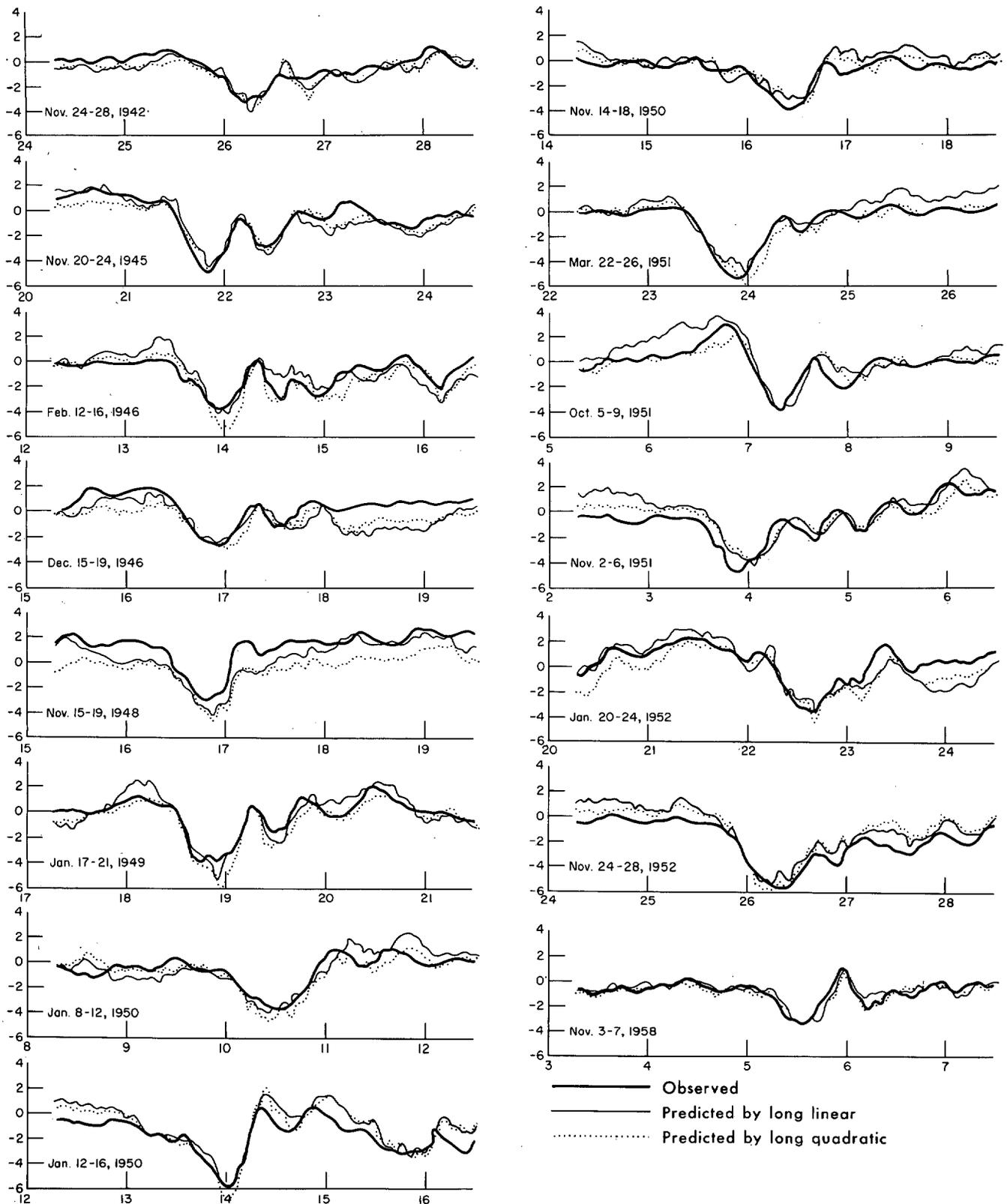


FIGURE 8.—Comparison of the observed and computed surges at Toledo, Ohio, for both long regression equations with dependent data.

STORM SURGE AT TOLEDO
INDEPENDENT DATA

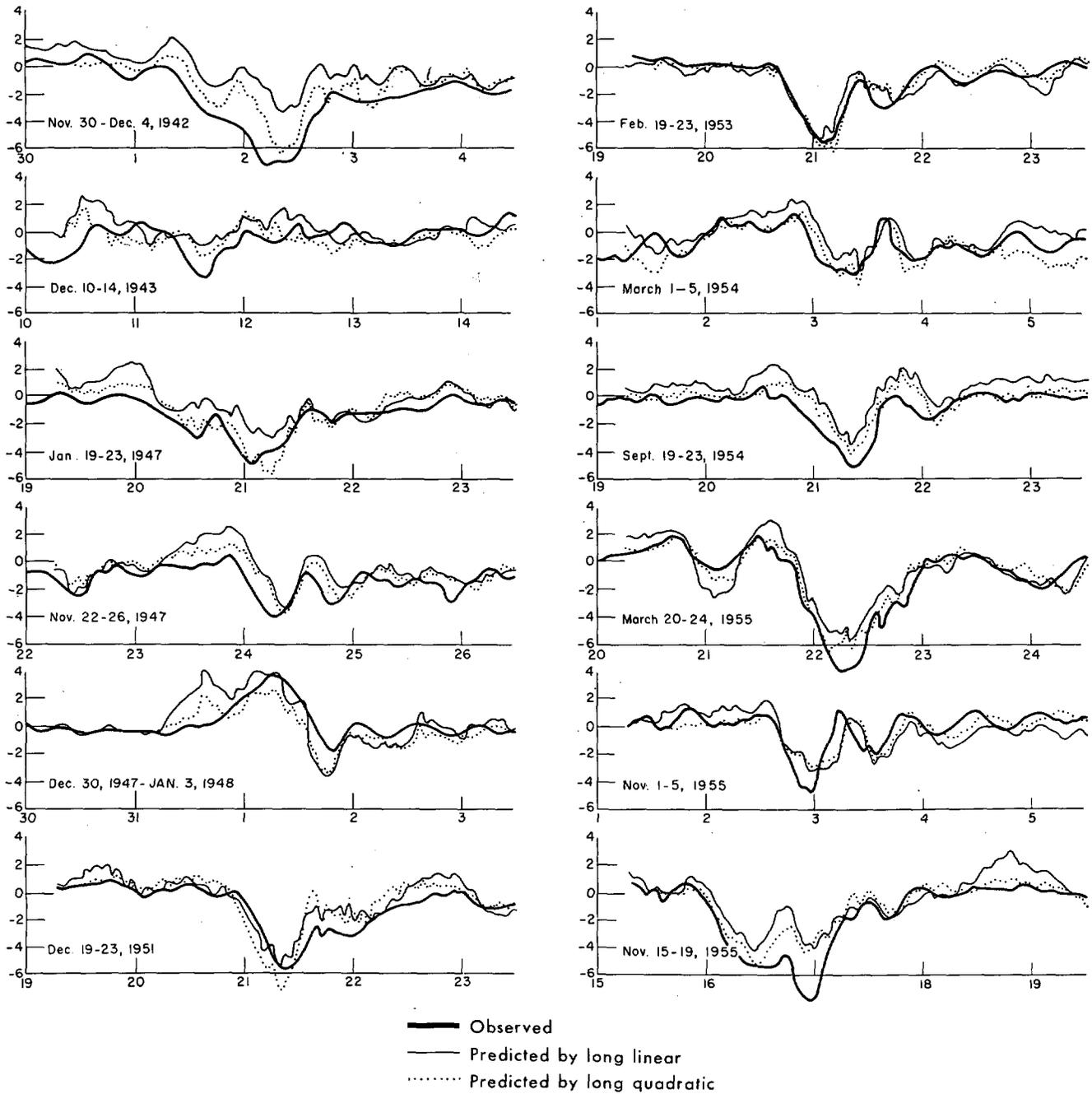


FIGURE 9.—Comparison of the observed and computed surges at Toledo, Ohio, for both long regression equations with independent data.

15. G. W. Platzman, "The Dynamical Prediction of Wind Tides on Lake Erie," *Meteorological Monographs*, 1963 (to be published).
 16. N. A. Pore, "The Relation of Wind and Pressure to Extratropical Storm Surges at Atlantic City," 1963 (to be published).
 17. F. G. Shuman, "Numerical Experiments with the Primitive Equations," *Proceedings of the International Symposium on*

Numerical Weather Prediction in Tokyo, Nov. 7-13, 1960, Meteorological Society of Japan, Mar. 1962, pp. 85-107.
 18. S. Uusitalo, "The Numerical Calculation of Wind Effect on Sea Level Elevations," *Tellus*, vol. 12, No. 4, Nov. 1960, pp. 427-435.
 19. J. L. Verber, *Long and Short Period Oscillations in Lake Erie*, State of Ohio, Department of Natural Resources, Division of Shore Erosion, 1960, 80 pp.

STORM SURGE AT TOLEDO
DEPENDENT DATA

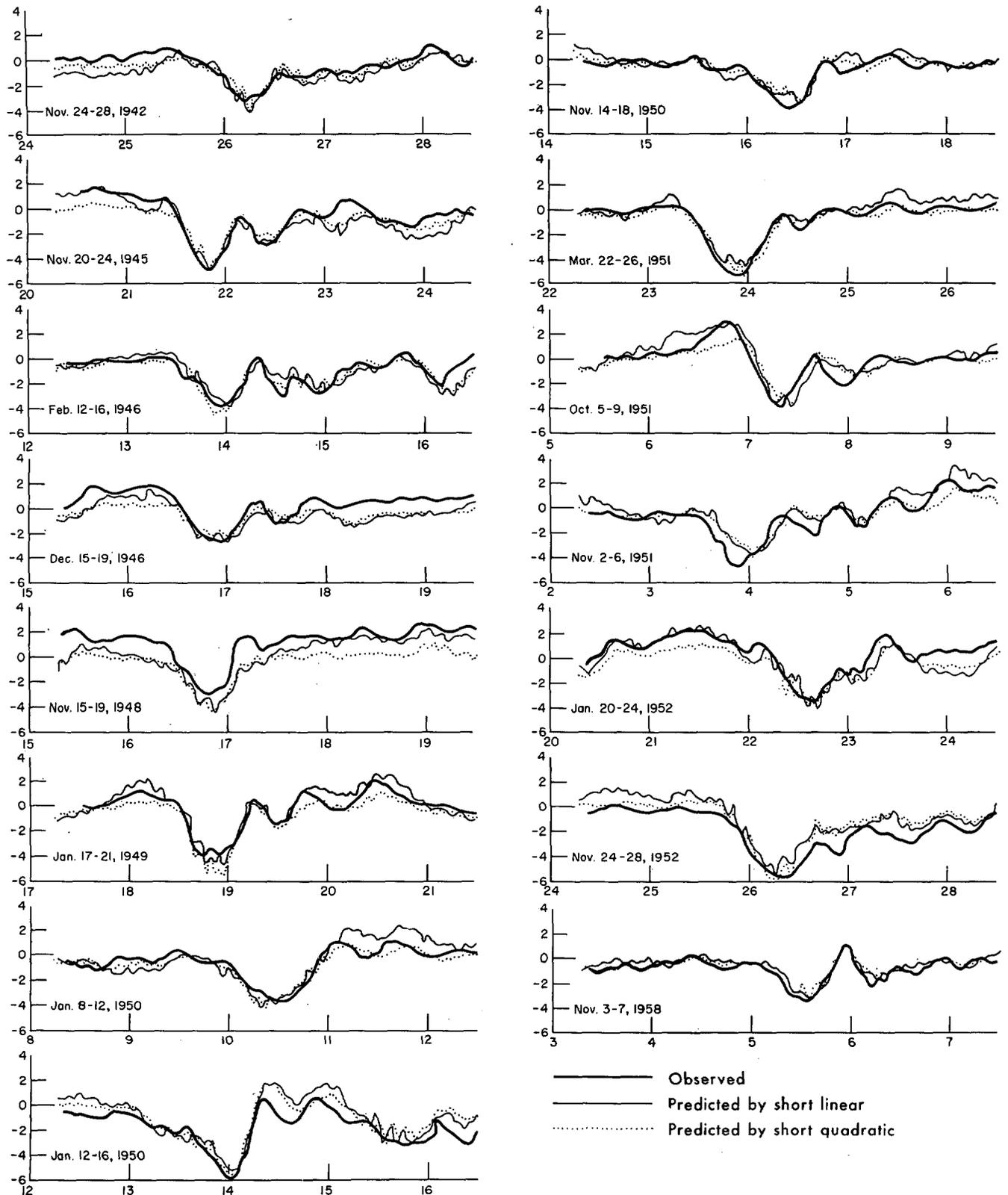


FIGURE 10.—Comparison of the observed and computed surges at Toledo, Ohio, for both short regression equations with dependent data.

STORM SURGE AT TOLEDO
INDEPENDENT DATA

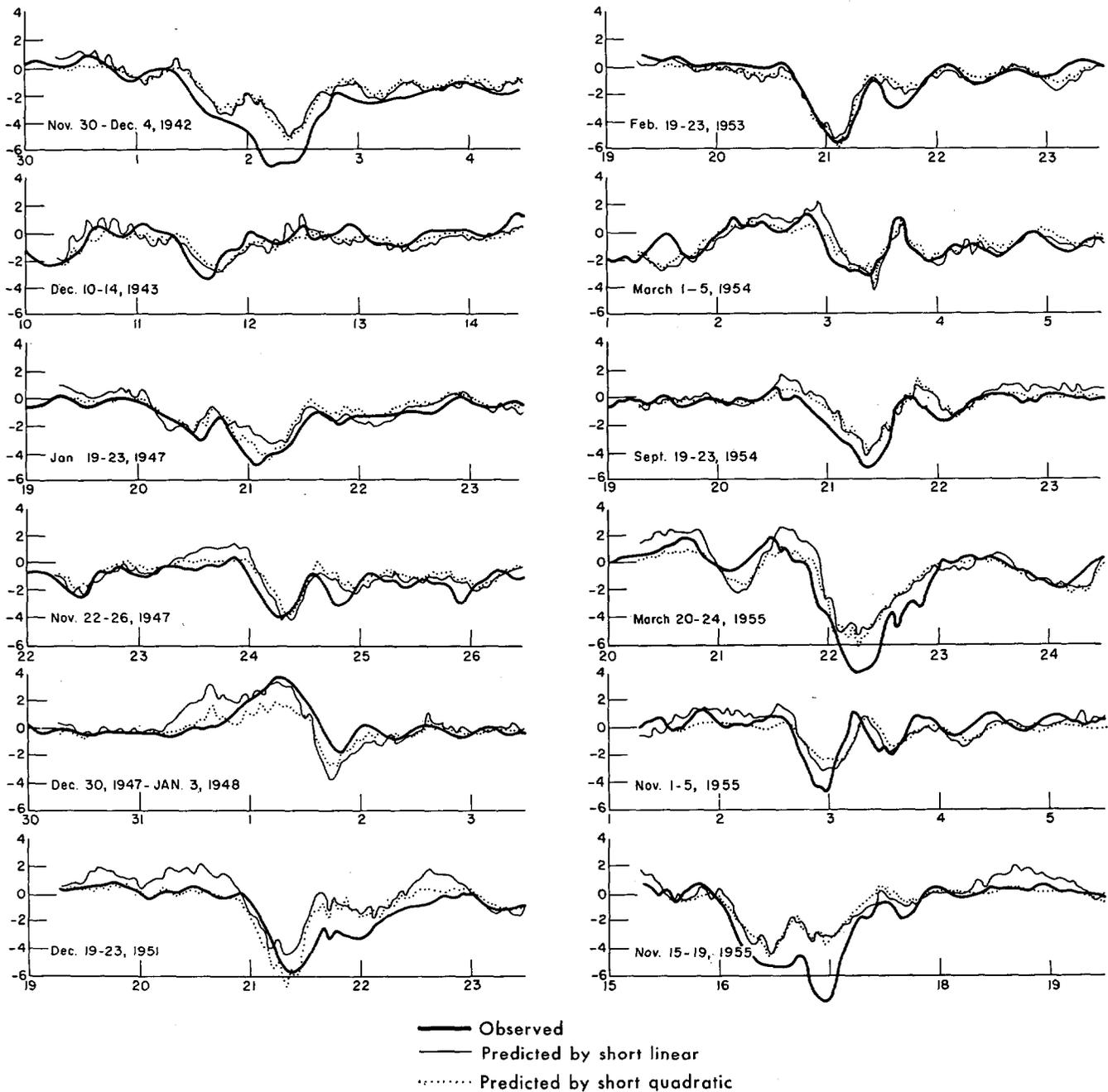


FIGURE 11.—Comparison of the observed and computed surges at Toledo, Ohio, for both short regression equations with independent data.

20. P. Welander, "On the Wind Action on Shallow Seas with Particular Reference to the Problem of Numerical Sea Level Prediction," Institute of Meteorology, University of Stockholm, 1956.
21. P. Welander, "Wind Action on a Shallow Sea: Some Generalizations of Ekman's Theory," *Tellus*, vol. 9, No. 1, Feb. 1957, pp. 46-52.

22. P. Welander, "Numerical Prediction of Storm Surges," *Advances in Geophysics*, vol. 8, 1961, pp. 315-379.
23. B. W. Wilson, "The Prediction of Hurricane Storm-Tides in New York Bay," *Technical Memorandum No. 120*, U.S. Army Corps of Engineers, Beach Erosion Board, 1960.

LAKE ERIE STORM SURGES - TOLEDO

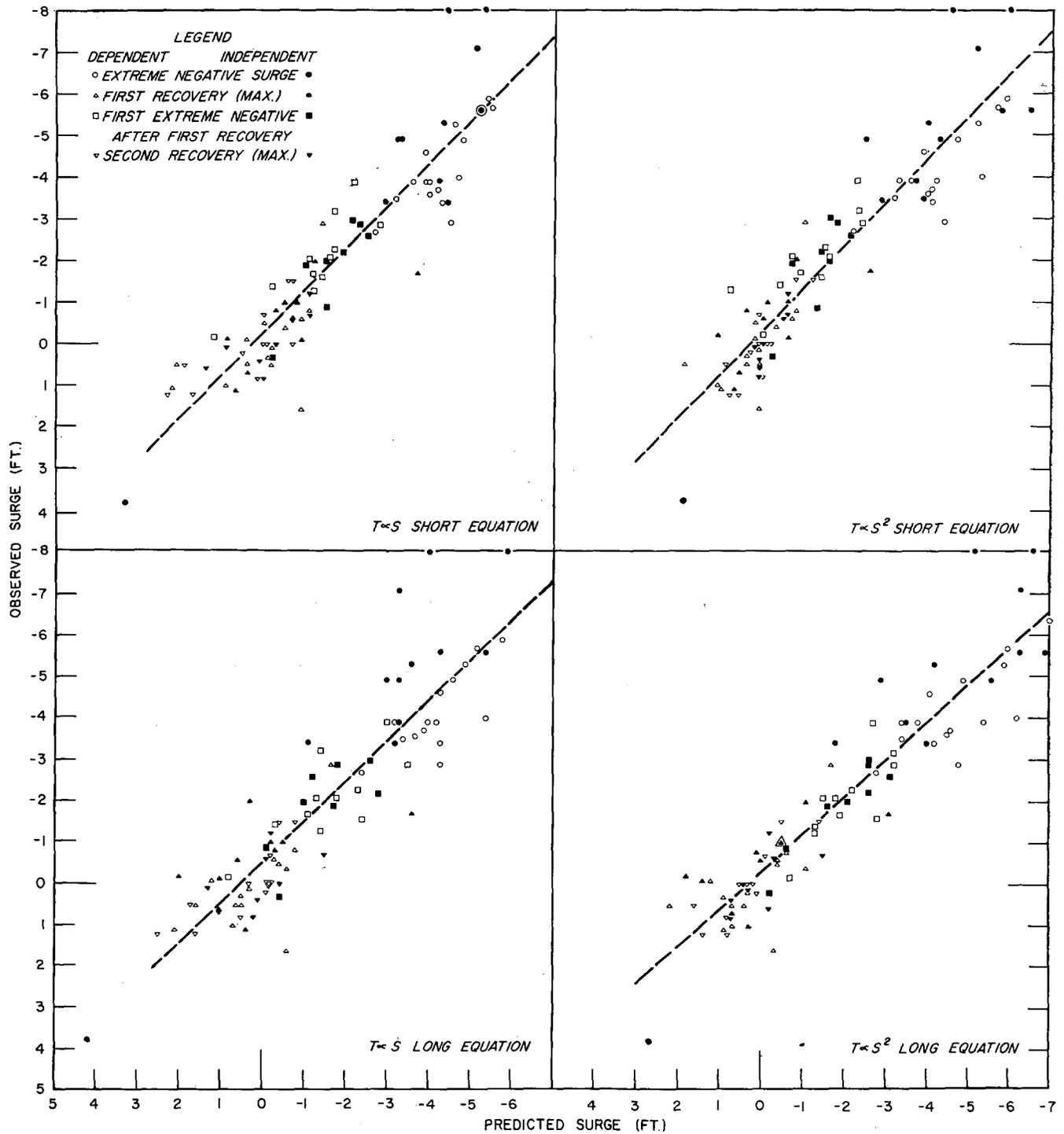


FIGURE 12.—Comparison of the observed extremes in water level at Toledo, Ohio, with the extremes predicted by the regression equations discussed in this paper.

24. B. W. Wilson, "Note on Surface Wind Stress over Water at Low and High Wind Speeds," *Journal of Geophysical Research*, vol. 65, No. 10, Oct. 1960, pp. 3377-3382.

25. B. W. Wilson, "The Prediction of Hurricane Storm-Tides in New York Bay," Author's reply to the discussion of D. L. Harris. *Technical Memorandum No. 120-A*, U.S. Army Corps of Engineers, Beach Erosion Board, 1961, pp. 13-29.