U.S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL WEATHER SERVICE
SYSTEMS DEVELOPMENT OFFICE
TECHNIQUES DEVELOPMENT LABORATORY

TDL OFFICE NOTE 79-18

# VERIFICATION OF AUTOMATED EAST COAST EXTRATROPICAL STORM SURGE FORECASTS

William S. Richardson,
David A. Andrews, and Franklin A. Rosenstein

November 1979

# VERIFICATION OF AUTOMATED EAST COAST EXTRATROPICAL STORM SURGE FORECASTS

William S. Richardson,
David A. Andrews, and Franklin A. Rosenstein

## 1. INTRODUCTION

The development of coastal communities and businesses along the east coast has increased the potential for serious damage resulting from extratropical storm surges. Storm surge (measured water level minus astronomical tide) is primarily caused by wind stress on the water surface. This surge, which is modified by the nearshore bathymetry and the shoreline, is superimposed on the astronomical tide. When significant storm surges and associated wave action occur at the same time as high astronomical tides, coastal property may be seriously damaged.

## 2. BACKGROUND

At the request of the National Weather Service's Eastern Region, the Techniques Development Laboratory (TDL) developed automated extratropical storm surge forecast guidance for 12 tide gage locations along the U.S. east coast. These locations are shown in Fig. 1.

Separate storm surge forecast equations were derived for each location with a multiple regression screening program. The regression program was used to correlate observed storm surge heights at 0100, 0700, 1300, and 1900 EST (predictand) with analyzed sea-level pressure values at 6-Layer Primitive Equation (6LPE) model grid points (predictors). This approach, where predictand data are correlated with observed predictors, is referred to as the "perfect prog approach". Pore et al. (1974) give a complete discussion of the storm surge forecast method which was developed from 68 storms. These storms occurred during 13 winters (November through April) from 1956 to 1969.

Forecasts of storm surge heights are made by interpolating sea-level pressure forecasts of the Limited-area Fine Mesh (LFM) model to 6LPE grid points. These interpolated values are the predictors in the storm surge forecast equations. Since September 1977, storm surge forecasts (National Weather Service, 1978) have been generated with sea-level pressure forecasts of the LFM-II model. Storm surge forecasts are made out to 48 hours at 6-h intervals.

This Office Note presents a verification of the automated storm surge forecasts. A discussion and recommendations based upon this verification are also presented.

## 3. VERIFICATION

The following is an explanation of the verification procedure and a presentation of verification results. A discussion of these results is presented in a later section.

Storm surge forecasts were verified with measured storm surge data from mid-November 1977 through April 1978. Daily maximum surge forecasts and forecasts associated with significant storm surge events (magnitude of measured surge equaled or exceeded 2 ft at six or more stations) are verified. Each verification is presented separately.

### 3.1 Daily Maximum Surge Forecasts

We recorded the measured surge height (positive and negative) from data at 0100, 0700, 1300, and 1900 EST each day. The surge height with the greatest magnitude (maximum surge) and the corresponding time of this surge were tabulated from these four pieces of data. If the magnitude of the maximum surge equaled or exceeded 2 ft, the surge height and associated time were retabulated from hourly measured data (0000 through 2300 EST). The maximum surge forecast (positive or negative) for a 24-h period and the valid time of this forecast were tabulated from sets of paired forecasts. The four sets of paired forecasts are 06- and 12-h, 18- and 24-h, 30- and 36-h, and 42- and 48-h. Since the automated forecasts are generated twice each day, each set of paired forecasts gives surge forecasts valid at 0100, 0700, 1300, and 1900 EST. No attempt was made to match the time of the measured maximum surge with the valid time of the forecast surge, other than to ensure that they both occurred within the same 24-h period. These measured and forecast data were used to compute verification statistics (correlation coefficients, root-mean-square-errors, relative errors, biases, and averaged errors).

Verification statistics were computed for 11 of the 12 stations. Storm surge forecasts for Avon, N.C. were not verified because the tide gage was removed from this location before November 1977. Statistics were computed for each of the four sets of paired forecasts. Two sets of statistics were also computed from the combined data of the 11 stations. One set of statistics was based on all data. The other set was computed from only those days when the measured or forecast magnitude of the surge equaled or exceeded 2 ft.

An overall "picture" of the forecast verification is shown in Table 1. The statistics shown in this table are based upon all data for the combined 11 stations. Note that the verification statistics change very little with different forecast projections. As expected, the correlation coefficients generally decrease as the forecast projections increase, and root-mean-square-errors (RMSE's) and relative errors generally increase with increasing forecast projections. The relative error is defined as RMSE/(average magnitude of the measured storm surge). Biases indicate that the forecast guidance greatly overforecasts the magnitude of positive and negative storm surges. An average magnitude time error (AMTE) is computed from the absolute value of the difference between the time of the measured maximum surge and the valid time of the forecast maximum surge. This error is about 7 hours for all forecast projections. The average time error (ATE), which is computed as an algebraic difference, is negative for all forecast projections.

Verification statistics computed from significant daily maximum surges are shown in Table 2. Significant daily maximum surges are those daily maximum surges where the magnitude of the measured maximum surge or the forecast maximum surge, for any set of paired forecast projection, equaled or exceeded 2 ft. These statistics are based on approximately 20 percent of the daily maximum surge data. The statistics shown in Table 2 also vary very little with different forecast projections. Except for RMSE's and biases associated with negative surge forecasts, the verification statistics based on significant maximum surges (Table 2 statistics) are much better than the statistics based on all daily maximum surges. This is not surprising since the Table 2 statistics are computed from a sample of data which is similar to the data used to derive the forecast equations.

To get an idea of how the verification statistics vary at each of the 11 stations, turn to Table 3. This table shows statistics computed from all data for each station. Only the statistics based on the 18- and 24-h forecast projections are shown. This pair of forecast projections was determined to be the most important for operational forecasts. While the statistics at any one station change with forecast projection, the station statistics computed from the 18- and 24-h forecasts are representative of the other forecast projections.

The correlation coefficients are between 0.86 and 0.73 at all stations except Charleston. The correlation coefficient associated with the Charleston storm surge forecasts is only 0.54, about 25 percent lower than the average correlation coefficient (0.74) for the combined 11 stations. Charleston also has one of the largest relative errors. Positive surges are overforecast at all stations except Portland, while the magnitudes of negative surges are overforecast at seven of the 11 stations. Stamford has the largest biases, while the largest AMTE (8.16 hours) is recorded at Charleston. At all stations the sign of ATE is negative. However, this error varies from -0.05 hours at Portland and Charleston to -3.89 hours at Newport. Positive surges occur much more often than negative surges at all stations except the two (Stamford and Willets Point) in Long Island Sound.

A different "verification look" at individual stations is shown in Fig. 2 through Fig. 12. The left portion of each figure shows the comparison of the measured maximum surge with the 18- or 24-h forecast maximum surge. Each pair of surge heights (measured and forecast) occurs within the same 24-h period. The letter "A" designates the location of each pair of heights. If two pairs of heights have the same values, their location is denoted with a "B", three pairs "C", and so on. In the right portion of each figure the time error (time of measured maximum surge minus valid time of the forecast maximum surge) is plotted. Three categories of time errors are plotted with corresponding measured and forecast surge heights. The symbols and ranges of these categories are: . = -22 h to -8 h, X = -7 h to 7 h, and ⊠ = 8 h to 22 h. The time span of each category is 15 hours. Only one time error category can be plotted for a particular measured-forecast surge height. Therefore, there are hidden categories at locations denoted by alphabetic characters other than "A". Zero lines associated with the measured and forecast surges, and a line of "perfect fit" are drawn in each comparison plot. These plots give the following impressions.

(1) There is scatter about the line of "perfect fit".

(2) Extreme positive surges (greater than 3 ft) are underforecast.

(3) The majority of time errors lie within the -7- to +7-h category.

(4) Surges which are forecast with the wrong sign have the largest time errors.

In summary, maximum daily surges are forecast reasonably well at all stations except Charleston. The magnitudes of peak surges are generally overforecast except in the cases of extreme peak events (greater than 3 ft). Extreme peak surge events are underforecast. With regard to the forecast time of the peak surge, forecasts are generally off ±7 hours. Before these verification results are discussed let's look at the verification of significant storm surge events.

### 3.2 Forecasts of Significant Storm Surge Events

We have chosen four significant storm surge events (December 17-22, 1977, January 24-29, 1978, February 4-9, 1978, and April 25-30, 1978). The meteorological setting, the measured storm surge heights, and the 18- and 24-h forecast storm surge heights are presented for each event. The forecast storm surge heights were plotted with measured storm surge heights. Solid lines connect plots of measured storm surge heights which were plotted every 6 hours except for times when the surge was equal to or greater than 2 ft. Hourly values were plotted for these times. The 18- and 24-h surge forecasts are shown as dots. Dates are placed at 1200 EST, and arrows (↑) indicate the times of astronomical high tides.

The December event began with a low pressure system which developed along the N.C. coast (Fig. 13). By 0100 EST the next day, the storm had moved offshore. Maximum surges occurred almost simultaneously at Hampton Roads, Breakwater Harbor, Atlantic City, and New York (Fig. 14). There was good general agreement between the measured and forecast storm surge heights. However, the peak surge was greatly underforecast at Willets Point, and overforecast at Baltimore and Charleston. The maximum surges which occurred at Willets Point and locations north on December 21 were associated with a second storm. This storm was located near the coast at 1300 EST on December 21 (Fig. 15). Note that the peak surge often occurred at the time of low tides (midway between arrows). We will explore this phenomenon later.

The storm associated with the January 24-29, 1978 surge event was not a coastal storm. This storm formed in the southern part of the country on January 25. As the storm moved northward, explosive deepening took place (see Fig. 16). On January 26, south-southwesterly winds drove water up the Chesapeake Bay. Baltimore recorded a peak surge of 4.2 ft at 1500 EST on January 26 (Fig. 17). The positive surge was overforecast at Stamford. Negative surges on the 27th and the 28th were caused by westerly winds.

The record breaking storm of early February 1978 formed off the S.C. coast during the evening of February 5 (Fig. 18). The storm intensified as it moved up the east coast. Cape Cod reported winds of 92 mph. Maximum surges at Atlantic City and locations north occurred on February 6 and 7 (Fig. 19). The storm surge trends were forecast well. However, the peak surges from Willets Point to Portland were underforecast.

The April 25-30, 1978, storm deepened as it moved up the coast (Fig. 20). Maximum storm surges occurred on the 26th and 27th (Fig. 21). Water levels remained well above normal through the 28th as the mature storm moved slowly northeastward. The peak surge at Hampton Roads was underforecast. Measured water levels for Portland, Boston, and Newport are missing.

In summary, for these four significant surge events, storm surge trends were generally forecast very well. However, peak surges were often underforecast.

## 4. DISCUSSION

The following is a discussion of the verification results. Each result is discussed separately.

### 4.1 Poor Forecasts for Charleston

The verification of daily maximum surges pointed out, as did an earlier study by WSFO Charleston, that the daily maximum surge forecasts for Charleston are <u>not</u> good. We believe that there may be two reasons for the poor forecasts at this location. First, Charleston experiences fewer storm surge events than the other locations. Extratropical storms which are responsible for surges rarely develop and intensify until they are north of Charleston. Because of the fewer surge cases, the Charleston surge forecast equation is based on a smaller sample of data than the other forecast equations. Therefore, the Charleston equation may be less stable than the equations for the other locations. Because of the small developmental sample at Charleston, sea-level pressure data with time lags were not used as predictors. The surge forecast equation for Charleston is the only equation which does <u>not</u> contain sea-level pressure predictors with lag times. Since atmospheric forcing on the water surface is not instantaneous, we feel that the Charleston equation would give much better forecast guidance if it contained sea-level pressure predictors with lag times. We therefore plan to rederive the Charleston equation on a larger sample of data and include sea-level pressures with lag times as predictors in this equation.

### 4.2 Misleading Biases

The forecast biases are misleading in that they indicate the forecast equations generally overforecast the magnitude of the daily maximum surges at most stations. The comparisons of maximum measured surges with the maximum surge forecast for individual stations (Fig. 2 through Fig. 12) show that extreme peak surges (greater than 3 ft) are underforecast. At the time that these equations were implemented we realized that because the forecasts are based on a statistical derivation, the magnitudes of the peak surge would generally be underforecast. All surge forecasts are therefore adjusted by multiplying the original forecast value by a factor. The factor is the reciprocal of the multiple correlation coefficient associated with the forecast equation. The average value of these factors is about 1.2. It is therefore not surprising that forecast biases are greater than unity. From our verification it appears that negative and positive surges should be adjusted separately. These adjustments should only be applied above or below threshold values. We plan to determine these adjustments and threshold values for each location, after another season of storm surge data is added to 1977-1978 verification data.

## 4.3 Time Error

The average magnitude of the time error is about 7 hours. This error is not unreasonable since the surge is forecast at 6-h intervals, while the measured surge heights may be recorded every hour. However, a 6-h error can be the difference between the surge occuring at high or low tide. This may not be as bad as it seems, since most extratropical storms have durations longer than 12 hours. The surge generated by these storms will therefore occur during at least one high tide. We do not plan to derive hourly forecast equations.

## 4.4 Forecasts of Daily Maximum Surges

While the forecast equations are based on the relationship between significant storm surge events (events where the peak surge was 2 ft or greater) and sea-level pressure, the equations do provide guidance for events where the magnitude of the peak surge is less than 2 ft. The equations discriminate very nicely, in most cases, between negative and positive surges (see Figs. 2 through 12). When using these equations to determine the maximum magnitude of insignificant surges (magnitude less than 2 ft), keep in mind the RMSE's associated with these events are about 0.80 ft.

## 4.5 Relationship of Maximum Surge to the Stage of the Tide

As has been pointed out earlier, it appears that significant positive surges (surge greater than or equal to 2 ft) often occur at the time of low astronomical tides (see Figs. 14, 17, 19, and 21). Times of low tide are located midway between arrows (↑). In order to investigate the relationship between the high and low astronomical tides and significant positive surge heights we did the following.

The measured daily maximum surge data were separated into three categories: (1) Category I--negative surges less than or equal to -2 ft, (2) Category II--magnitude of surge heights less than 2 ft, and (3) Category III--positive surges greater than or equal to 2 ft. We defined high and low stages of the astronomical tide with a time span of 1.5 hours (45 minutes either side of the time of the high and low tide. Table 4 contains the number of maximum surges, by category, which occurred during high and low tide stages. It is interesting to note that the number of maximum surges associated with Categories I and II are nearly evenly divided between high and low tide stages. However, Category III surges (surges greater than or equal to 2 ft) appear to occur much more frequently at low stages of the tide.

The apparent relationship between low tides and significant positive surges should be viewed with caution because of the small sample size. Hopefully, another season of surge data will give us more confidence in this relationship. We were not surprised by the apparent relationship between low astronomical tides and peak positive surges. Harris (1963) reported that the component of storm surges caused by onshore wind is directly proportional to the wind stress and inversely proportional to the water depth.

## 5. FUTURE PLANS

After receiving comments on this Office Note from forecast offices, we plan to redo the verification with an additional season of storm surge data. Hopefully, this larger sample of data will enable us to tune the storm surge forecast guidance on a station-to-station basis.

The two seasons of surge data will be used to determine if there is a relationship between low astronomical tides and significant positive storm surges. If these data bear out this apparent relationship, we will investigate the possibility of including high and low astronomical tides as predictors in the storm surge equations.

We will attempt to improve the Charleston surge forecasts by rederiving the Charleston surge equations on a larger developmental sample. Sea-level pressures with time lags will be offered as predictors in this rederivation.

## 6. ACKNOWLEDGMENTS

The authors wish to thank Art Pore and Joe Gofus for helpful suggestions in writing this paper. A special thanks to Herman Perrotti for providing storm surge forecast data. Measured water level data were provided by the National Ocean Survey and the Geological Survey.

## REFERENCES

Harris, D. L., 1963: Coastal flooding by the storm of March 5-7, 1962. Manuscript of the U.S. Weather Bureau, January 1963, 22 pp.

Pore, N. A., W. S. Richardson, and H. P. Perrotti, 1974: Forecasting extra-tropical storm surges for the northeast coast of the United States. NOAA Technical Memorandum NWS TDL-50, National Oceanic and Atmospheric Administration, U.S. Dept. of Commerce, 70 pp.

National Weather Service, 1978: Extratropical storm surge forecasts for the U.S. east coast. NWS Technical Procedures Bulletin No. 226, National Oceanic and Atmospheric Administration, U.S. Dept. of Commerce, 5 pp.

PWM Portland, Maine

BOS Boston, Massachusetts

NWP Newport, Rhode Island

LGA Willets Point, New York

SFD Stamford, Connecticut

NYC New York, New York

ACY Atlantic City, New Jersey

BWH Breakwater Harbor, Delaware

BAL Baltimore, Maryland

ORF Hampton Roads, Virginia

AVN Avon, North Carolina

CHS Charleston, South Carolina

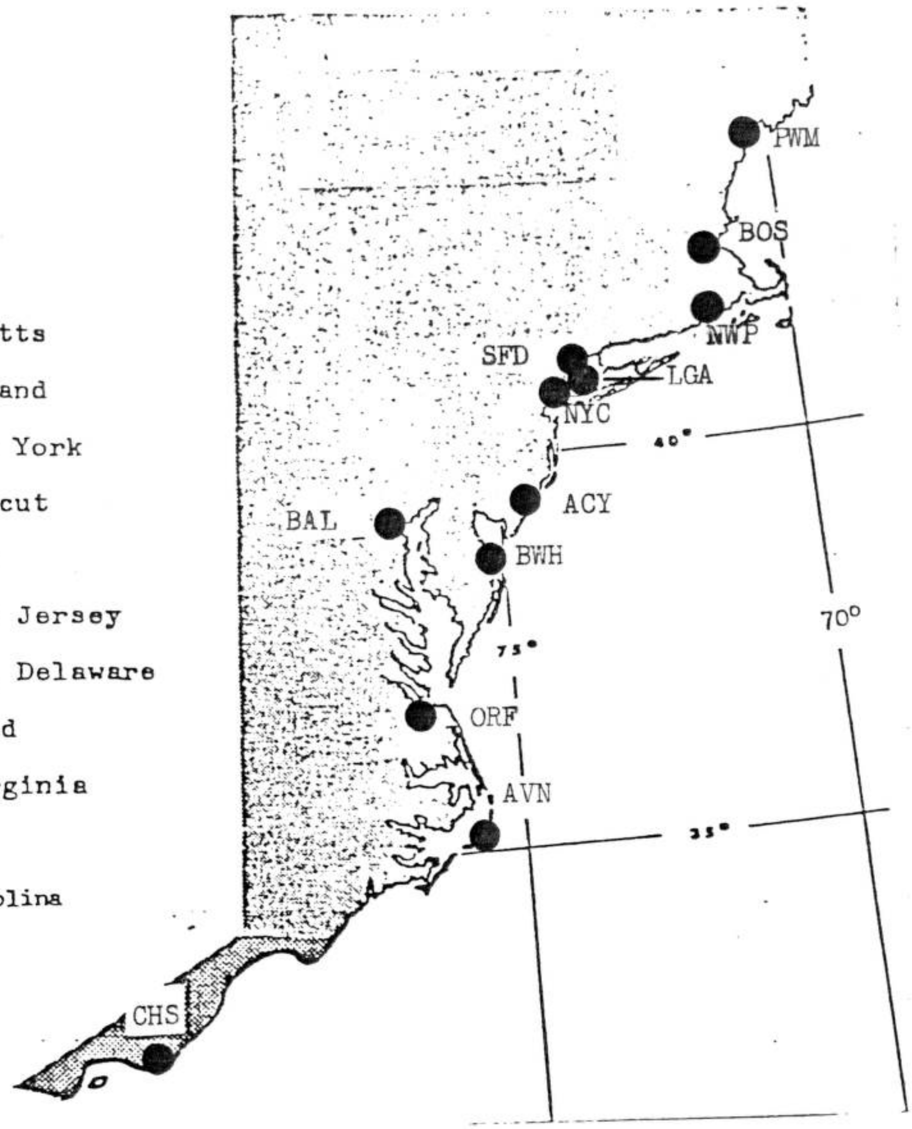
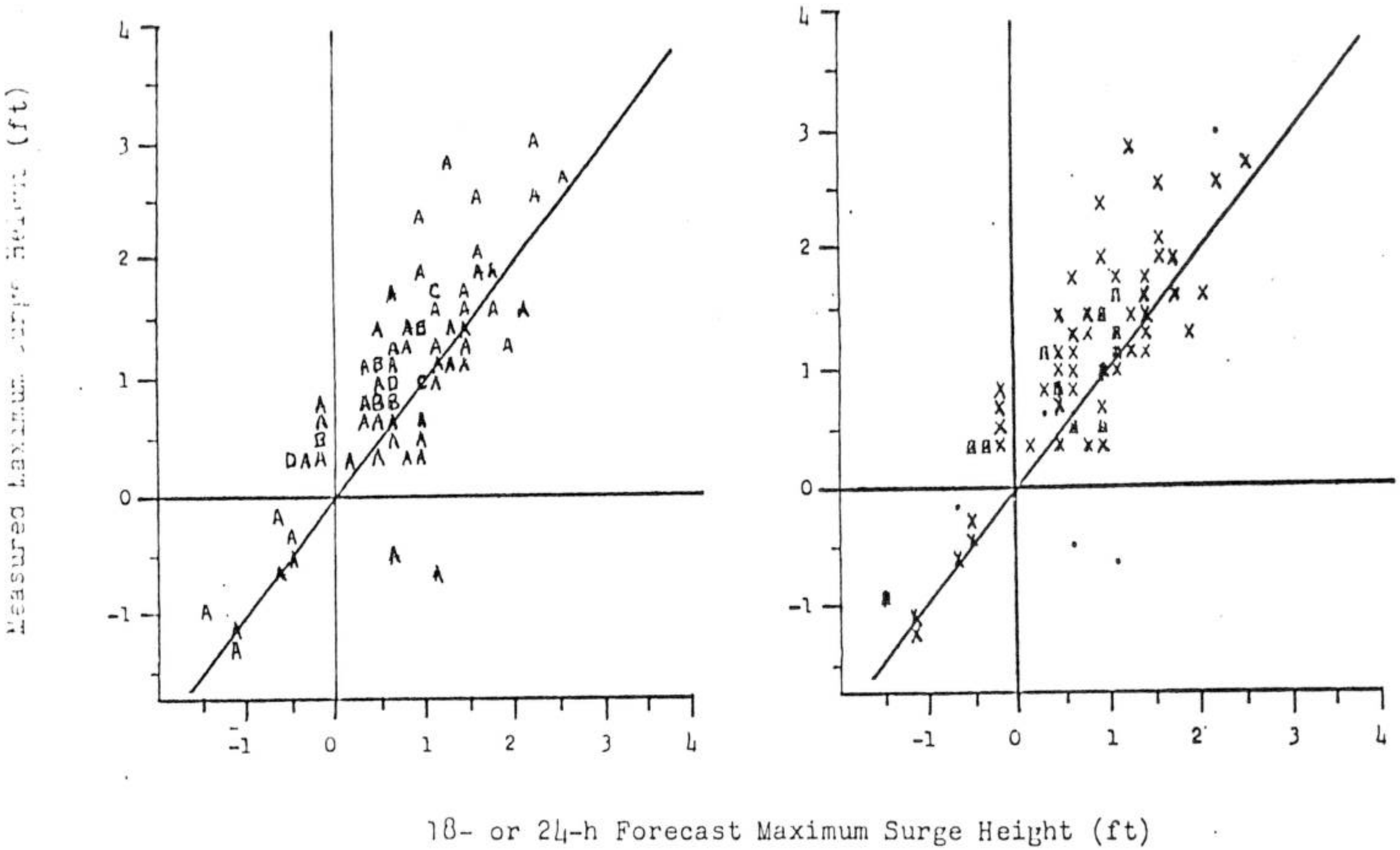



Figure 1. The 12 east coast locations for which automated extratropical storm surge forecasts are made.




Symbols of Time Error Categories

| | |
|---|---|
| . | -22h to -8h |
| X | -7 to 7h |
| ⊠ | 8 to 22h |

Figure 2. Comparisons of the measured maximum surge with the 18- or 24-h forecast maximum surge for Portland, Maine. The letter "A" designates the location of each pair of heights. If two pairs of heights have the same values, their location is denoted with a "B", three pairs with a "C", and so on. In the right portion of the figure the time error (time of measured maximum surge minus the valid time of the forecast maximum surge) is plotted. Three categories of time errors are plotted with corresponding measured and forecast surge heights.

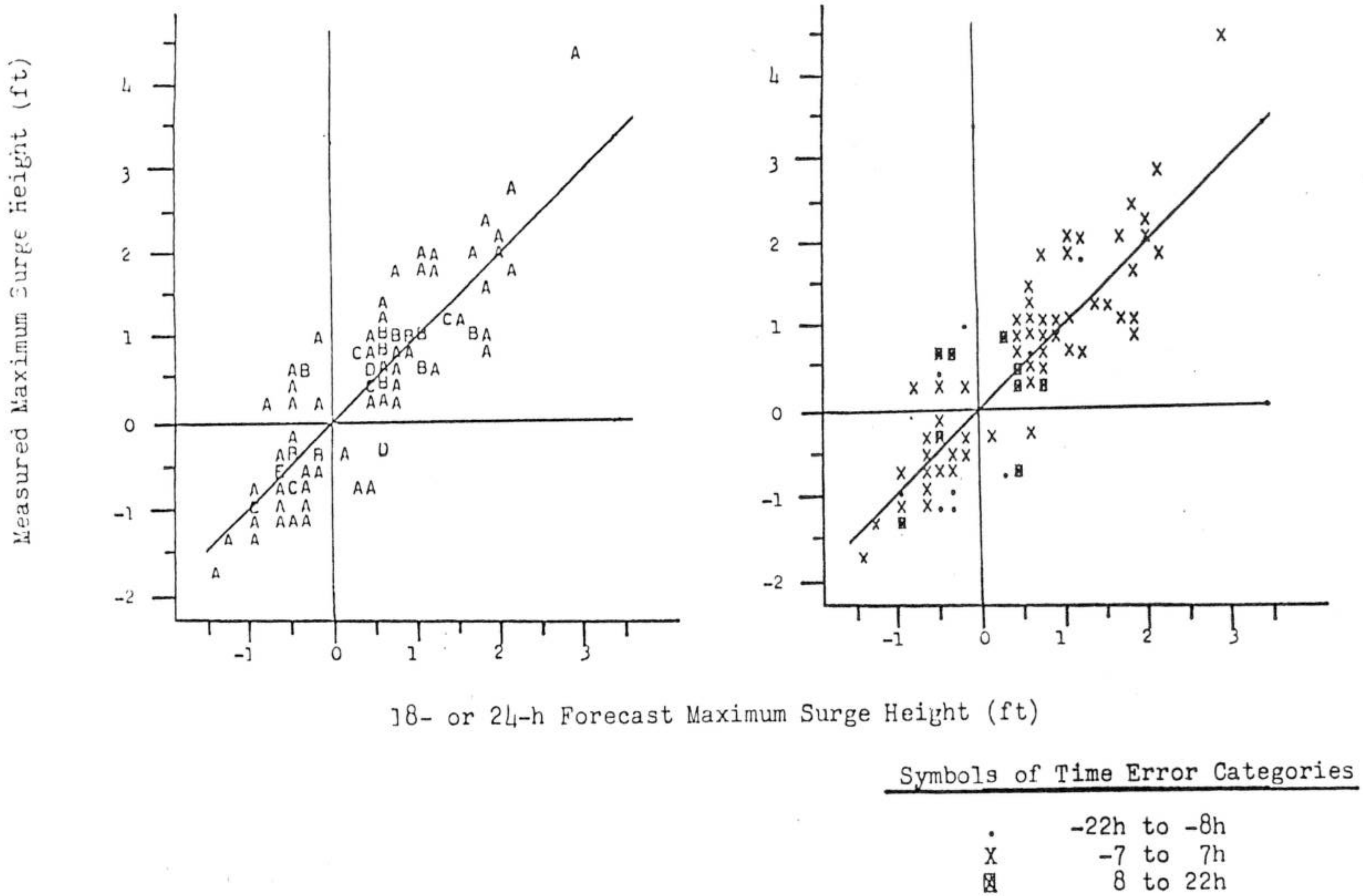



Figure 3. Same as Fig. 2 except for Boston, Mass.

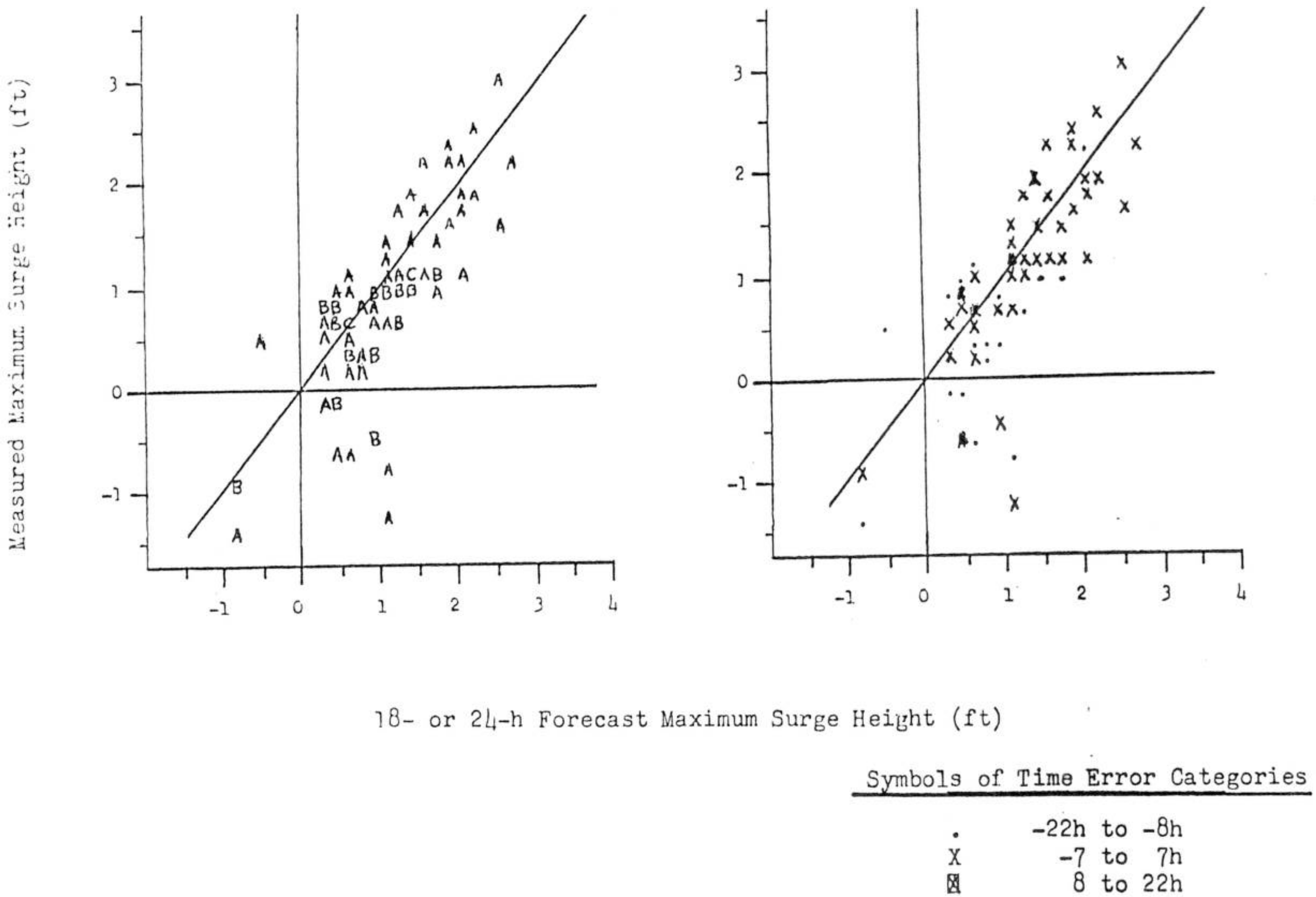



Figure 4. Same as Fig. 2 except for Newport, R.I.

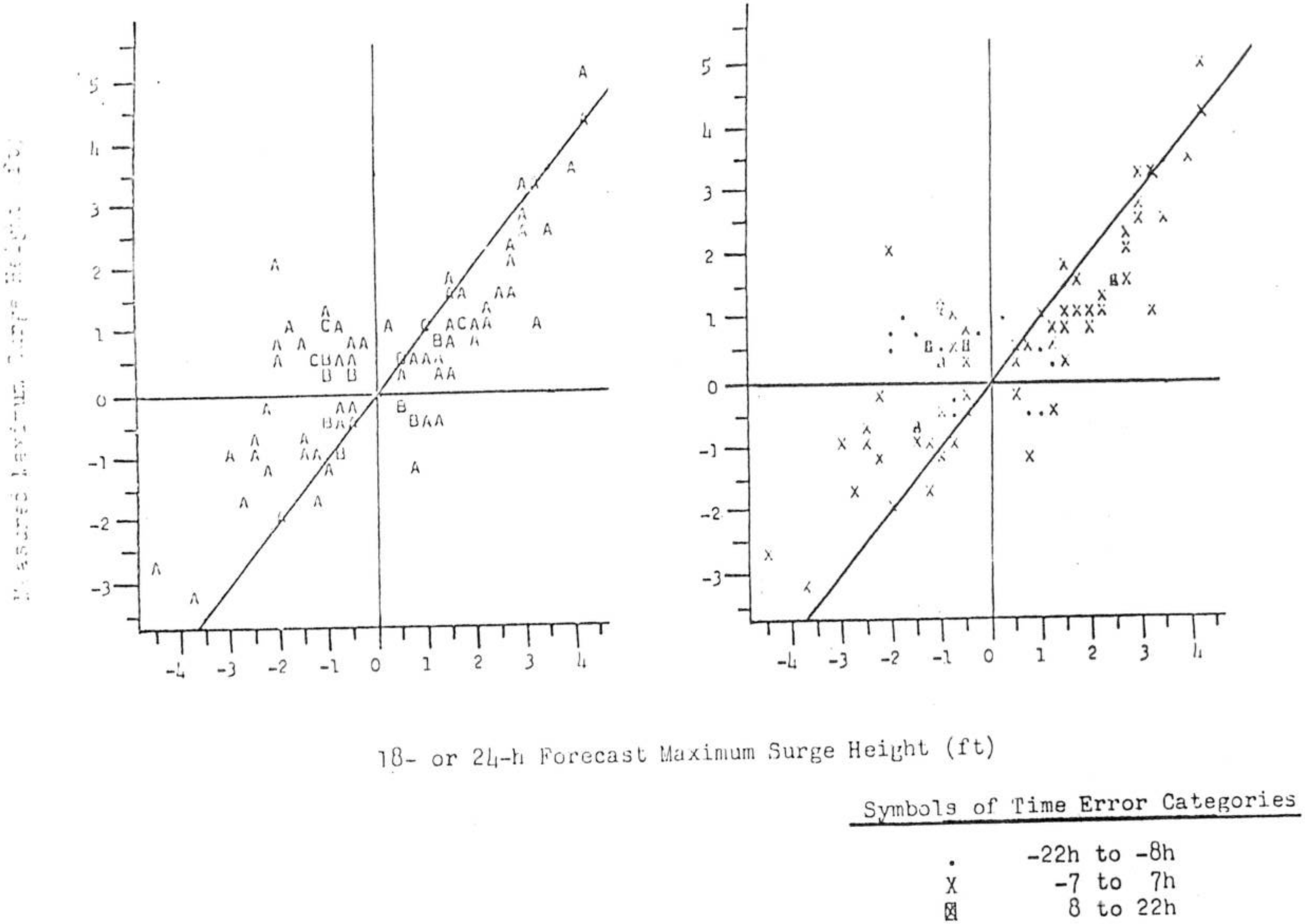



Figure 5. Same as Fig. 2 except for Stamford, Conn.

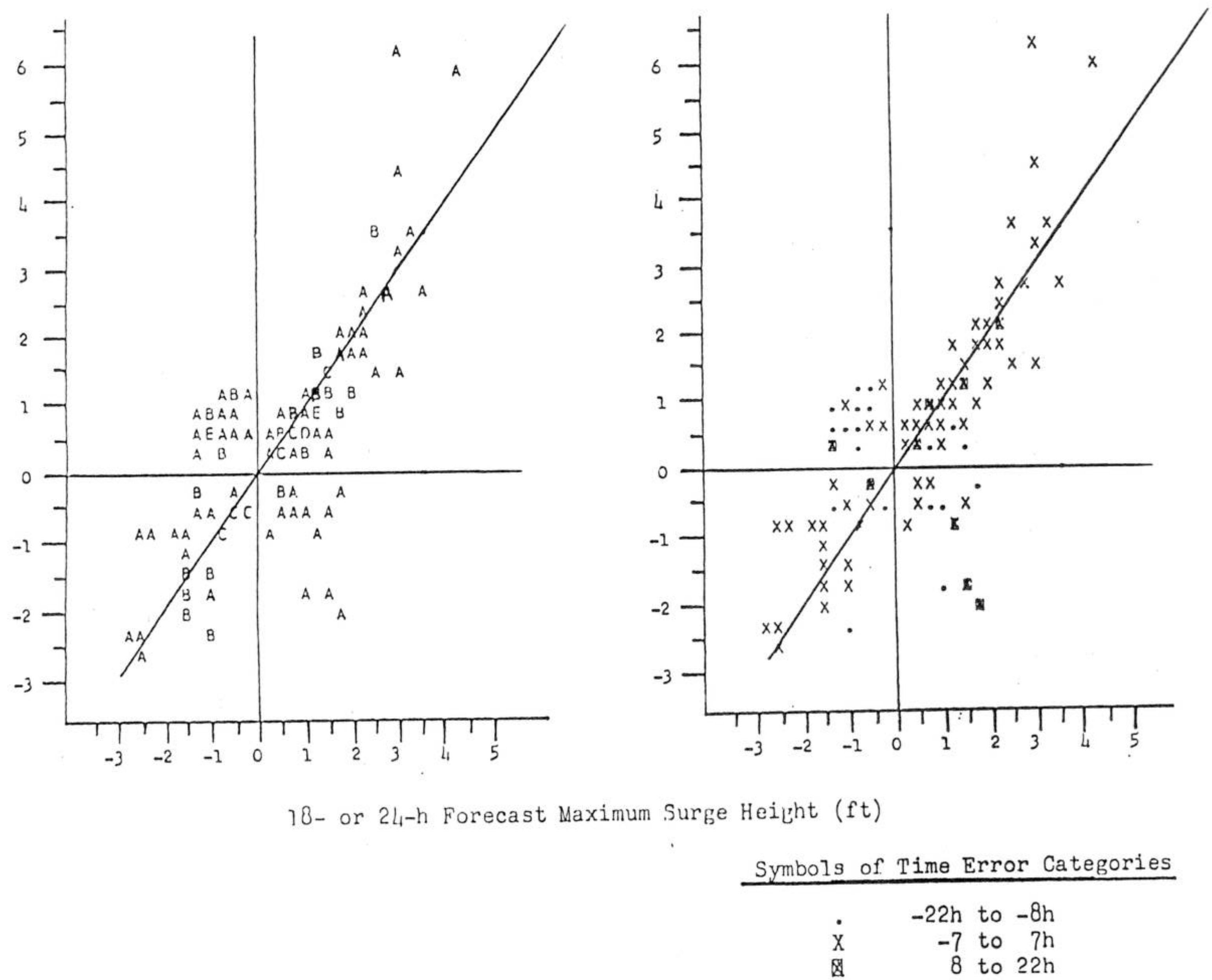



Figure 6. Same as Fig. 2 except for Willets Point, N.Y.

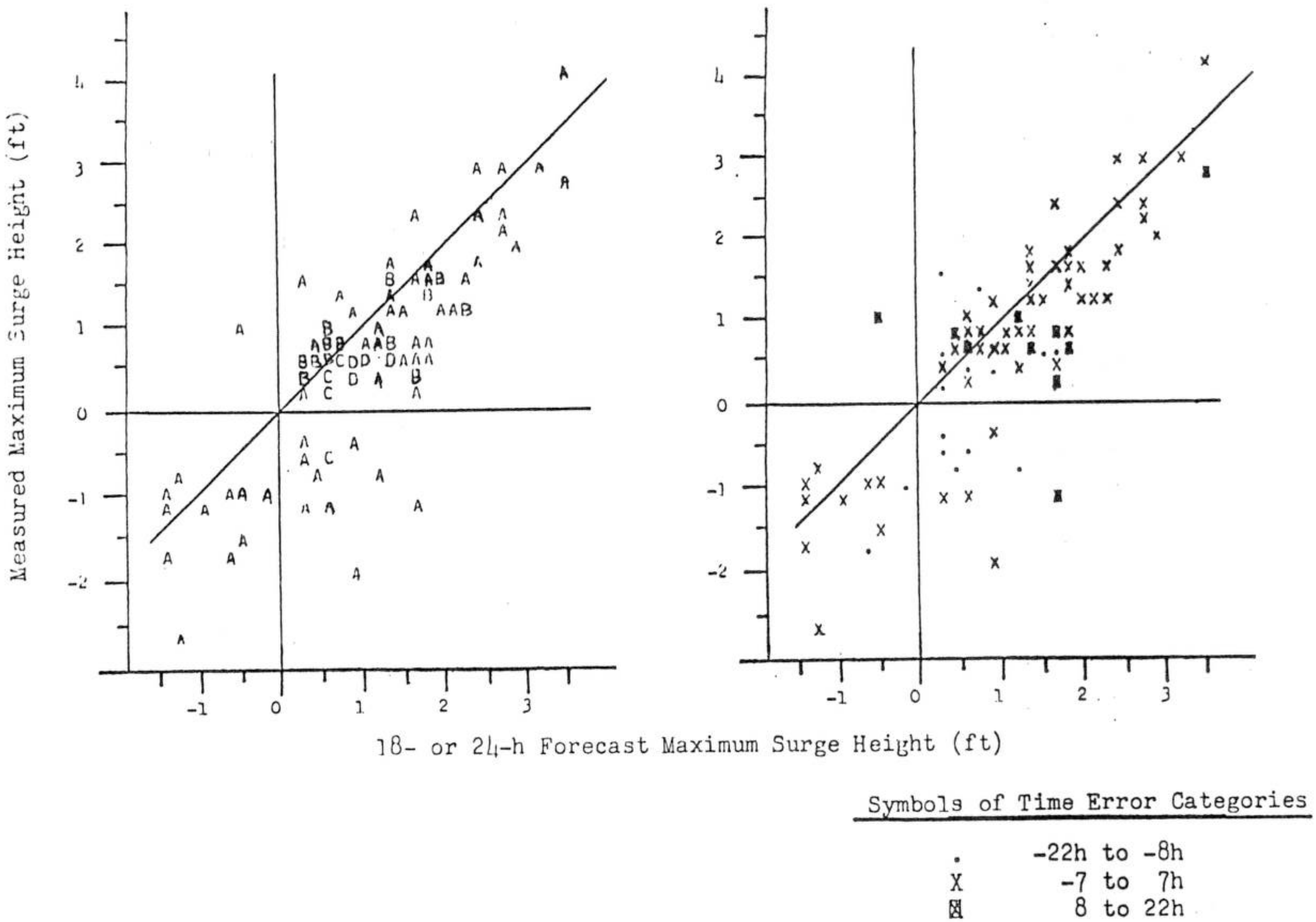



Figure 7. Same as Fig. 2 except for New York, N.Y.

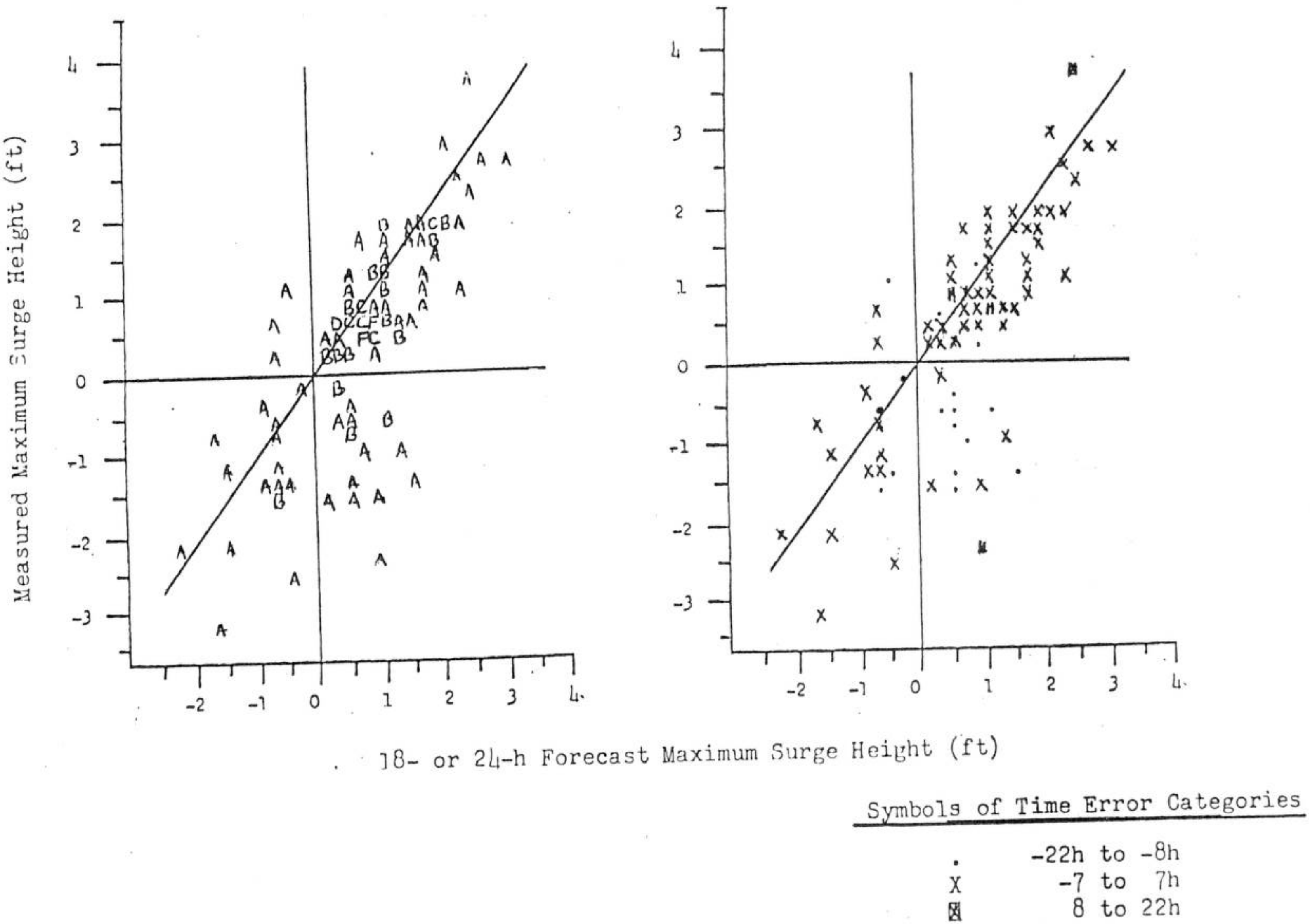



Figure 8 Same as Fig. 2 except for Atlantic City, N.J.

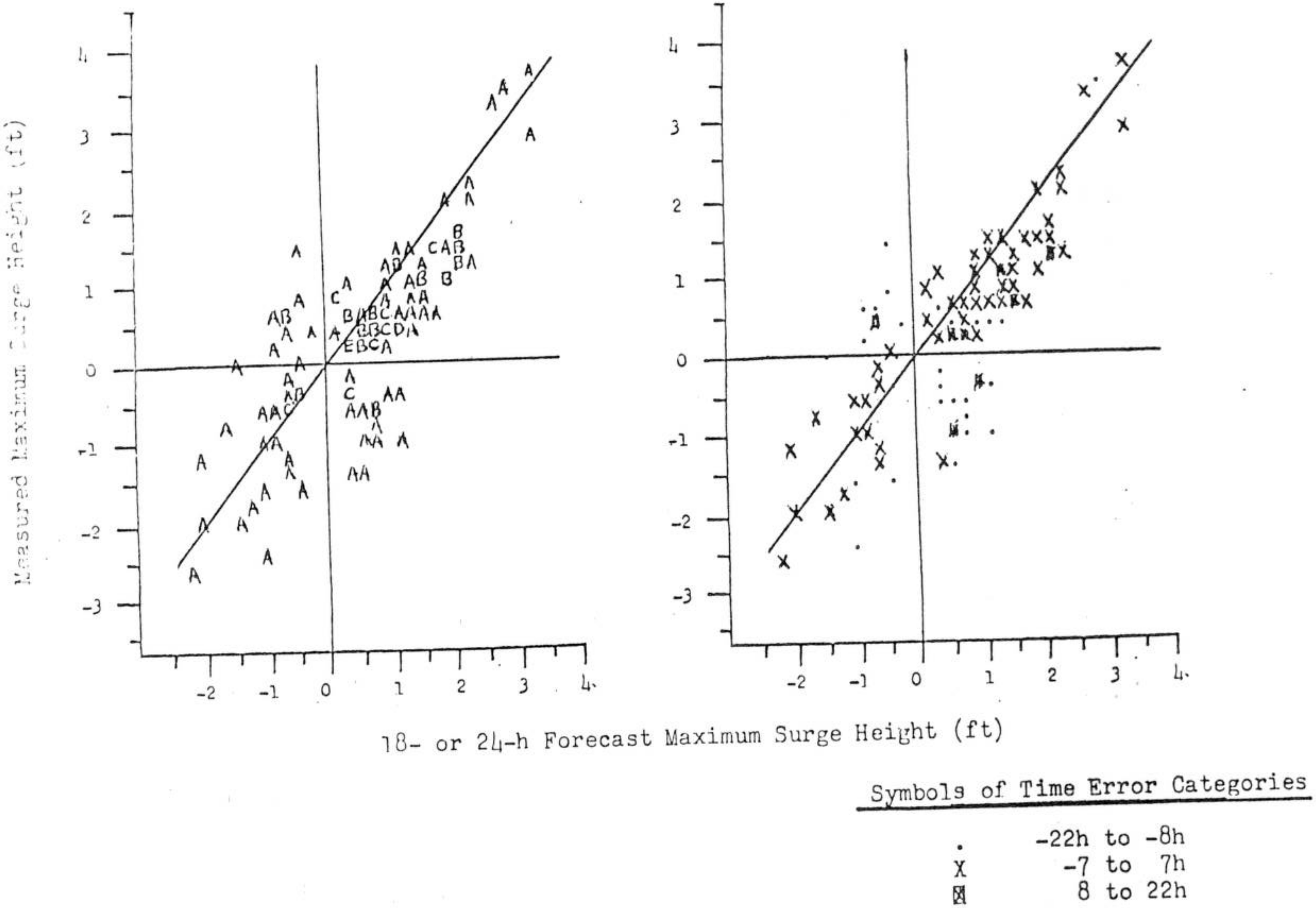



Figure 9. Same as Fig. 2 except for Breakwater Harbor, Del.

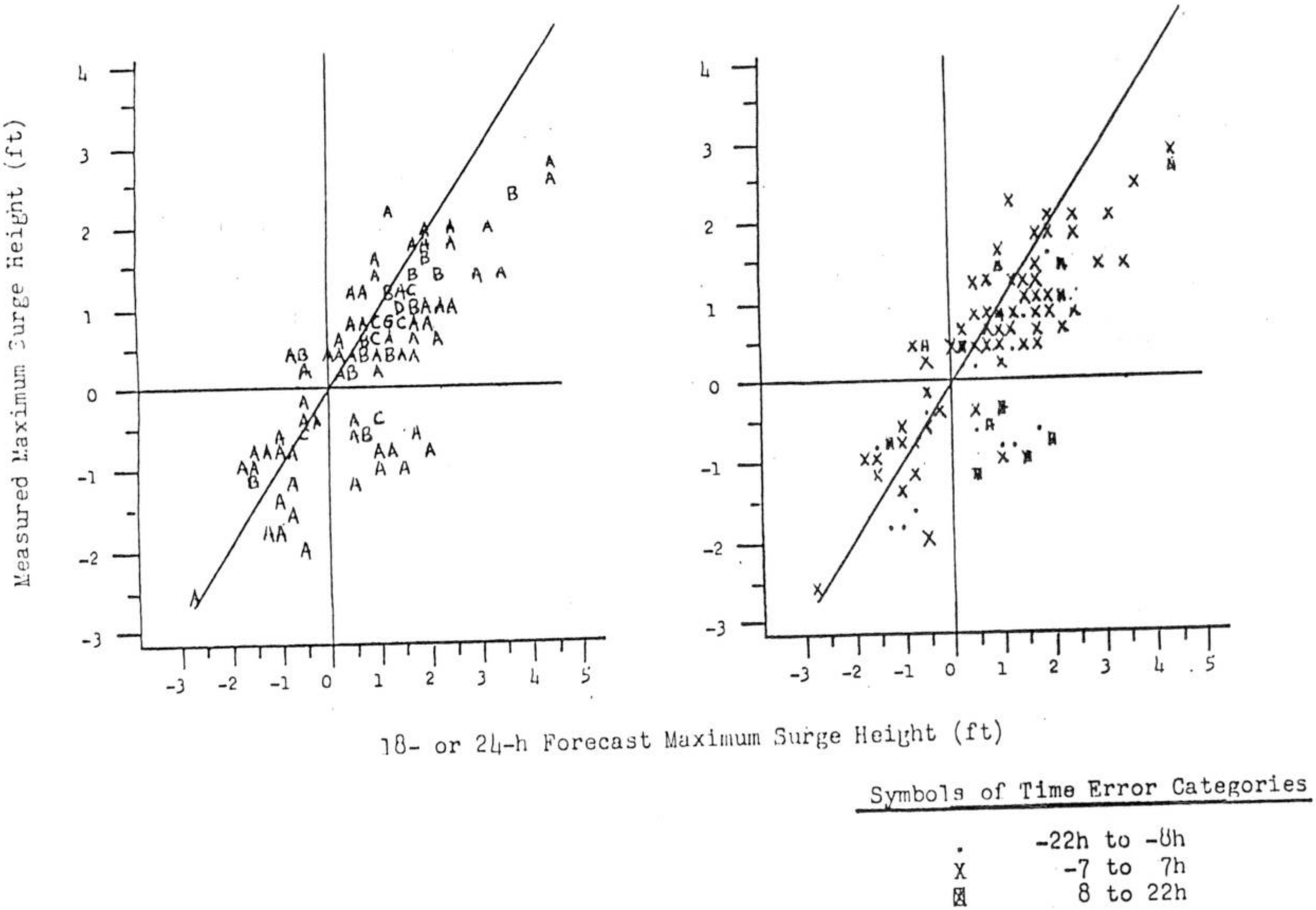



Figure 10. Same as Fig. 2 except for Baltimore, Md.

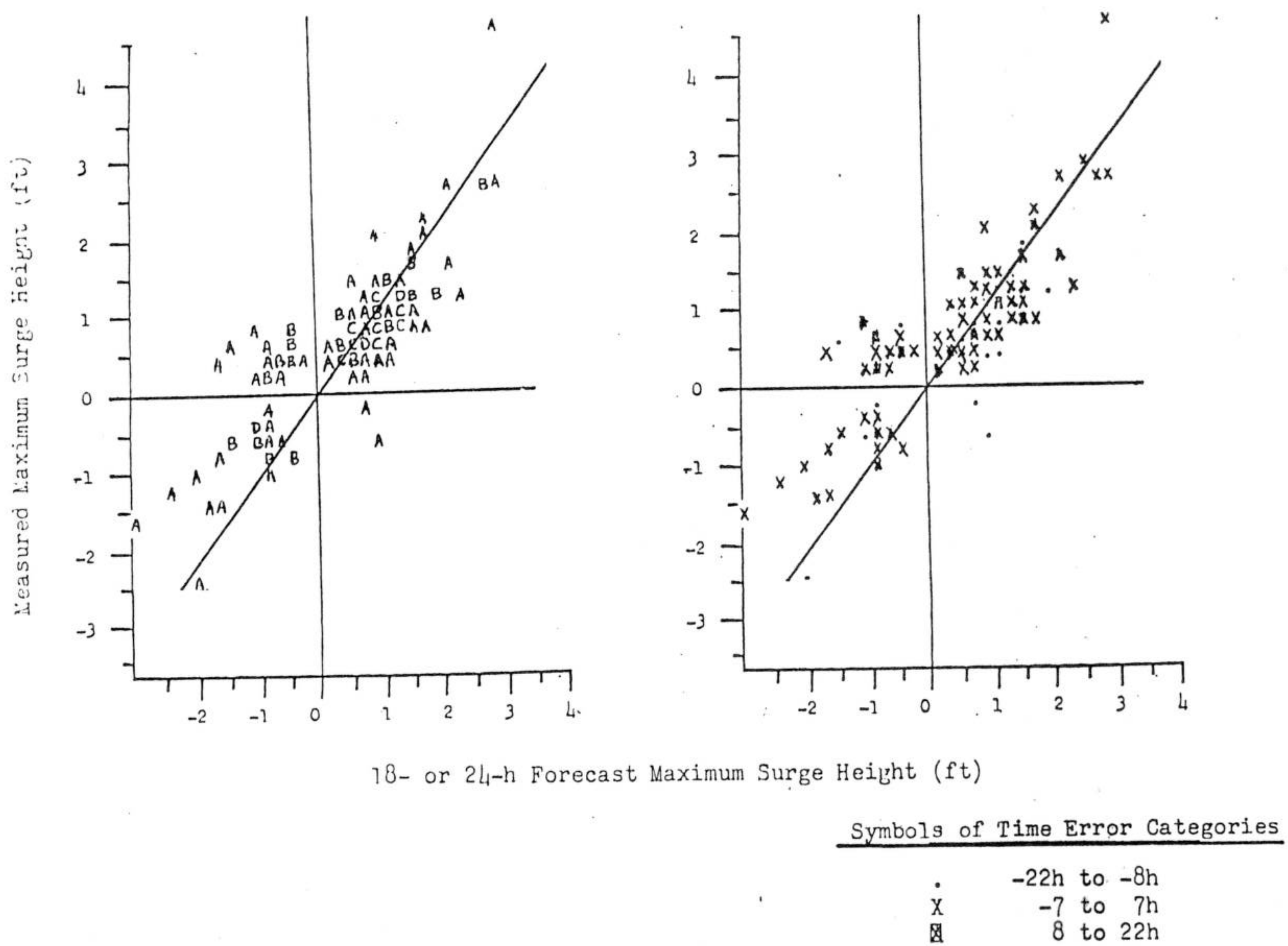



Figure 11. Same as Fig. 2 except for Hampton Roads, Va.

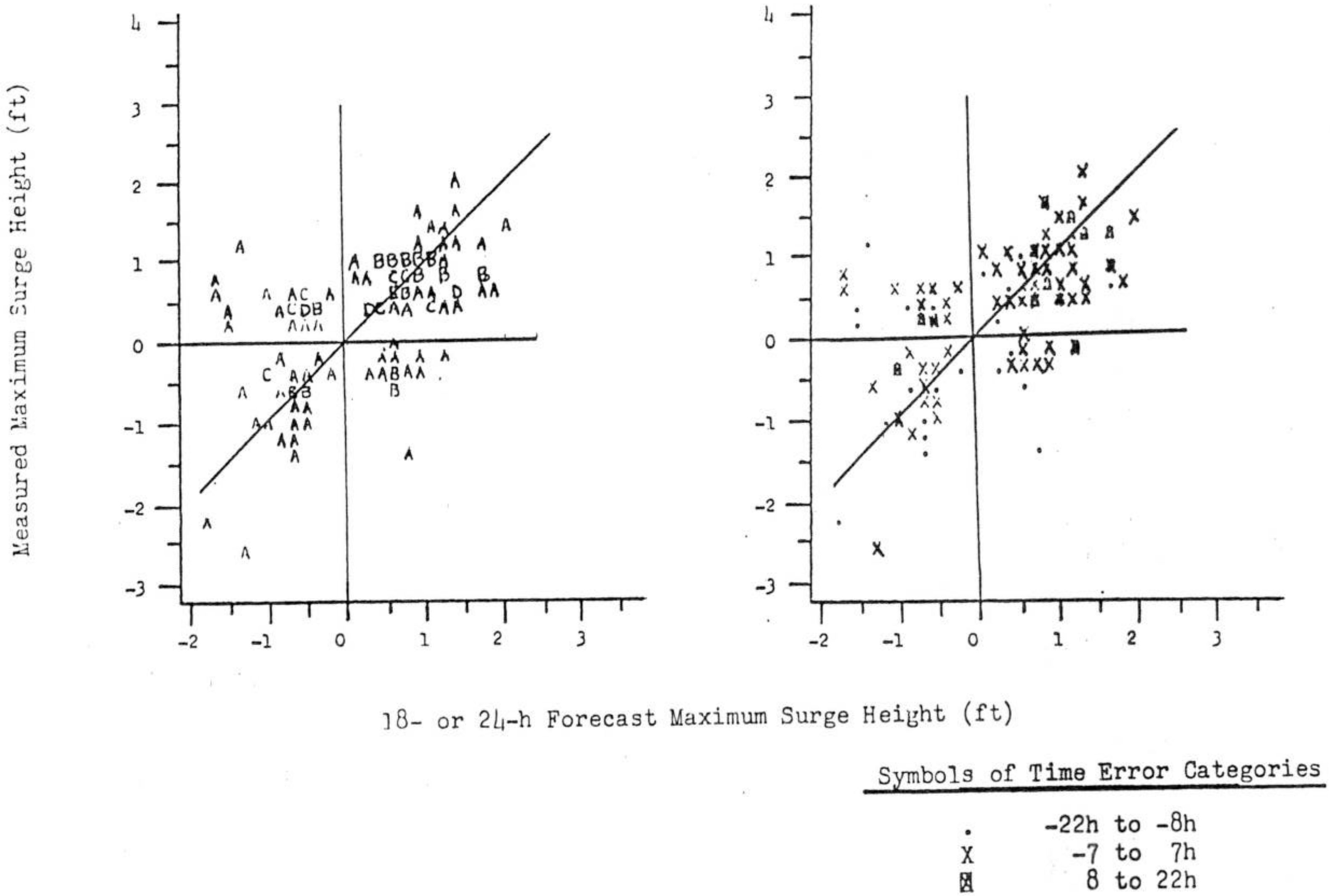



Figure 12. Same as Fig. 2 except for Charleston, S.C.

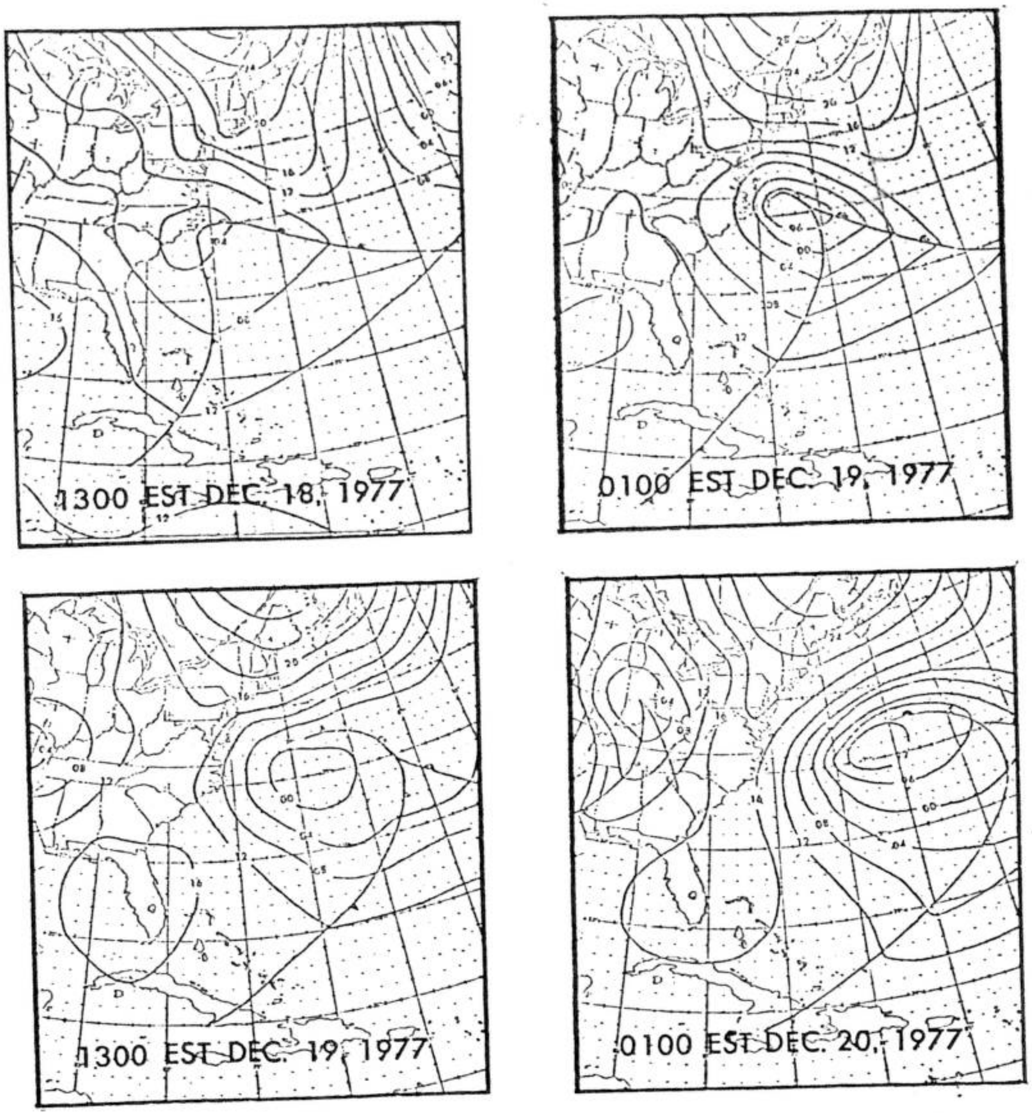



Figure 13. Sea-level pressure charts from 1300 EST December 18, 1977, to 0100 EST December 20, 1977.

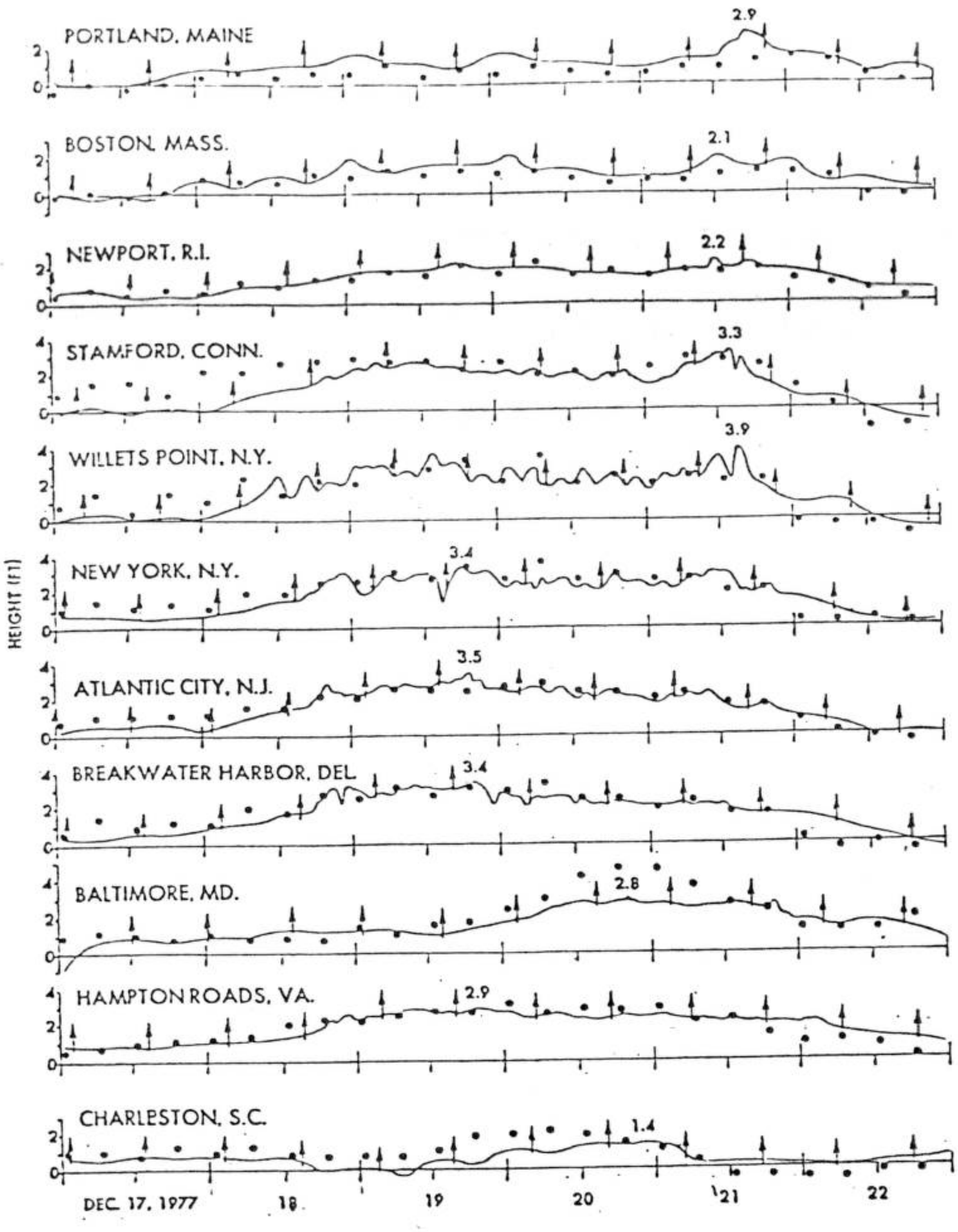



Figure 14. Significant storm surge event which occurred December 17 - 22, 1977. Measured surges are shown as solid lines, while 18- and 24-h forecasts are denoted by dots. The dates are shown at 1200 EST and arrows (↑) indicate the times of astronomical high tides.

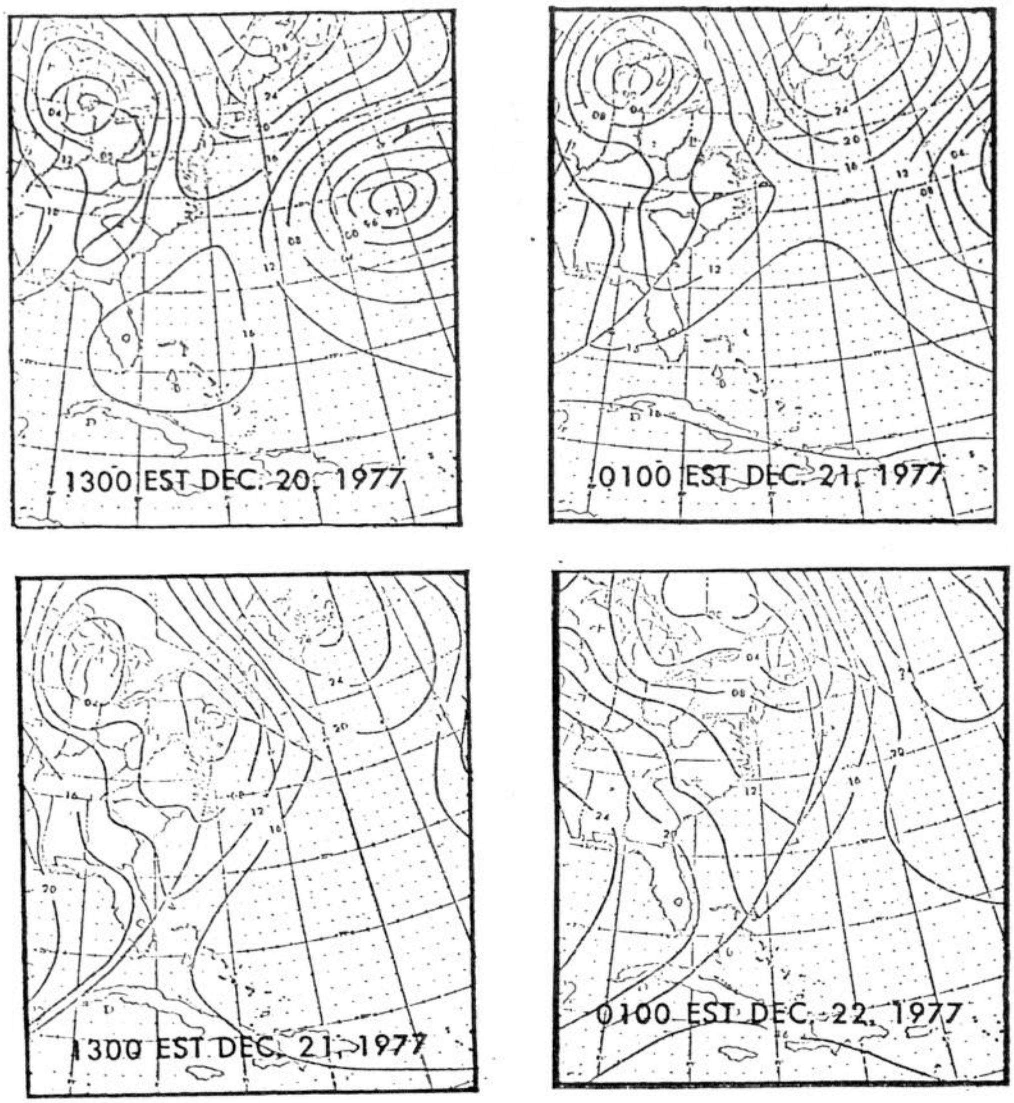



Figure 15. Sea-level pressure charts from 1300 EST December 20, 1977, to 0100 EST December 22, 1977.

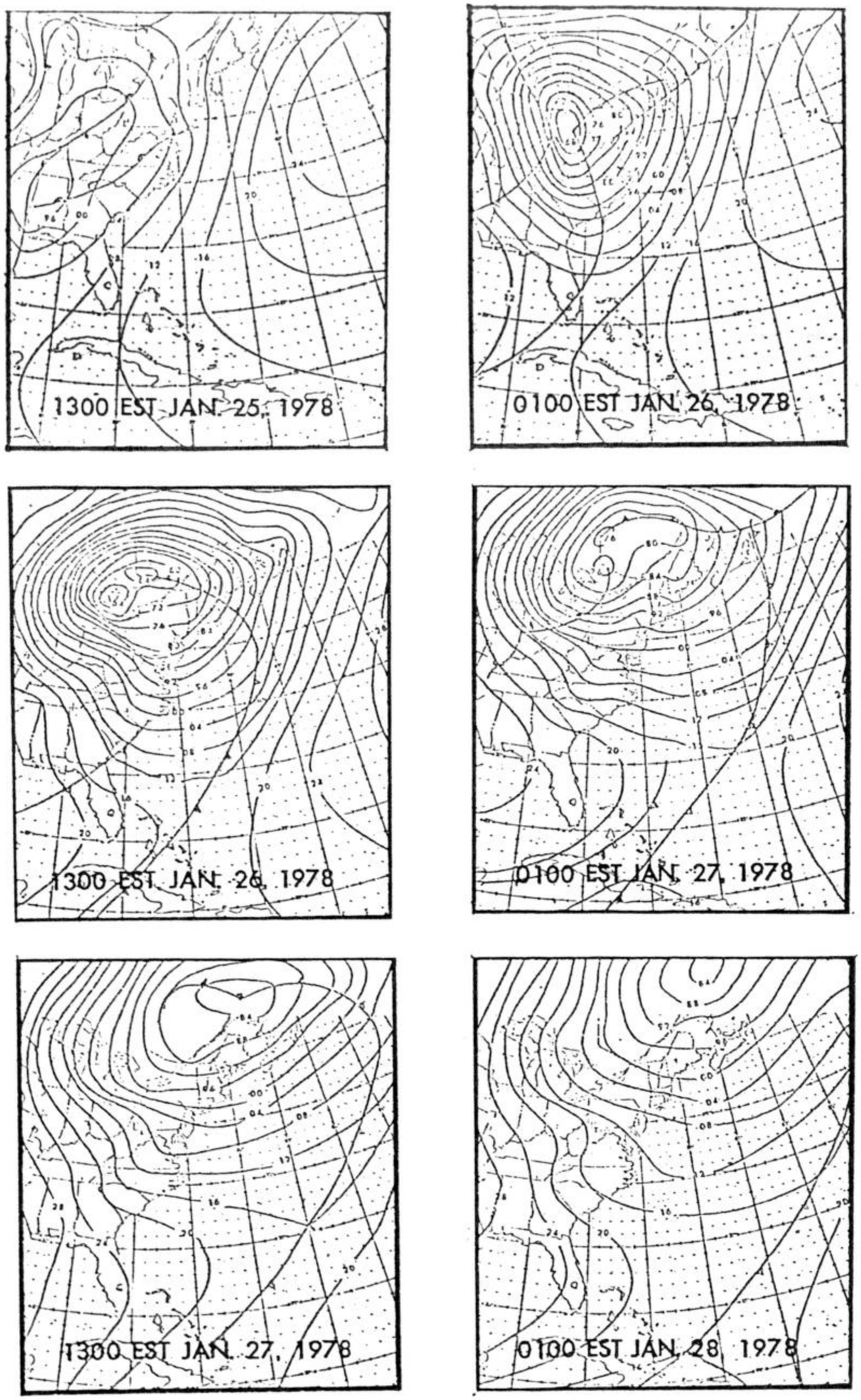



Figure 16. Sea-level pressure from 1300 EST January 25, 1978, to 0100 EST January 28, 1978.

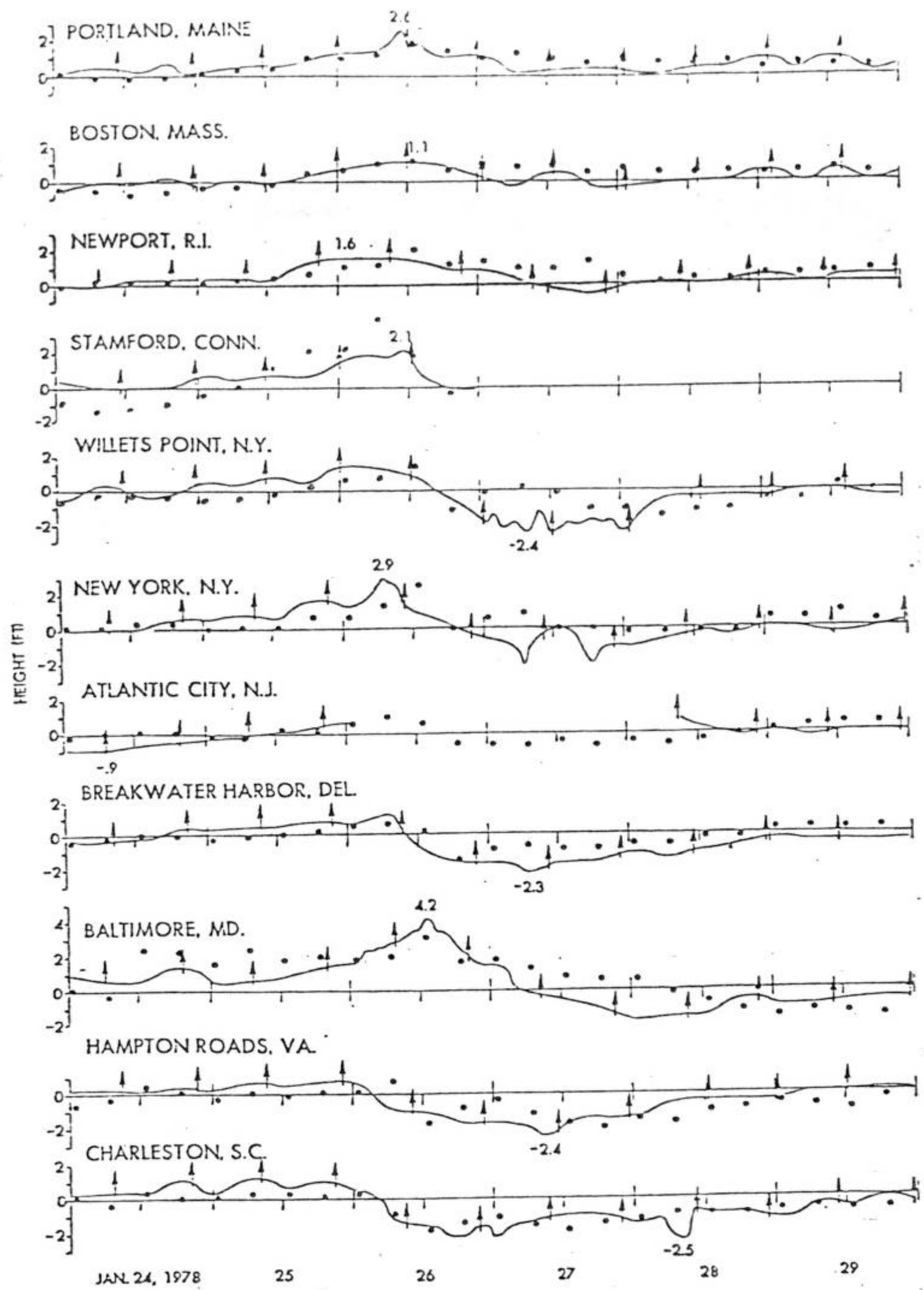



Figure 17. Same as Fig. 14 except for January 24 - 29, 1978.

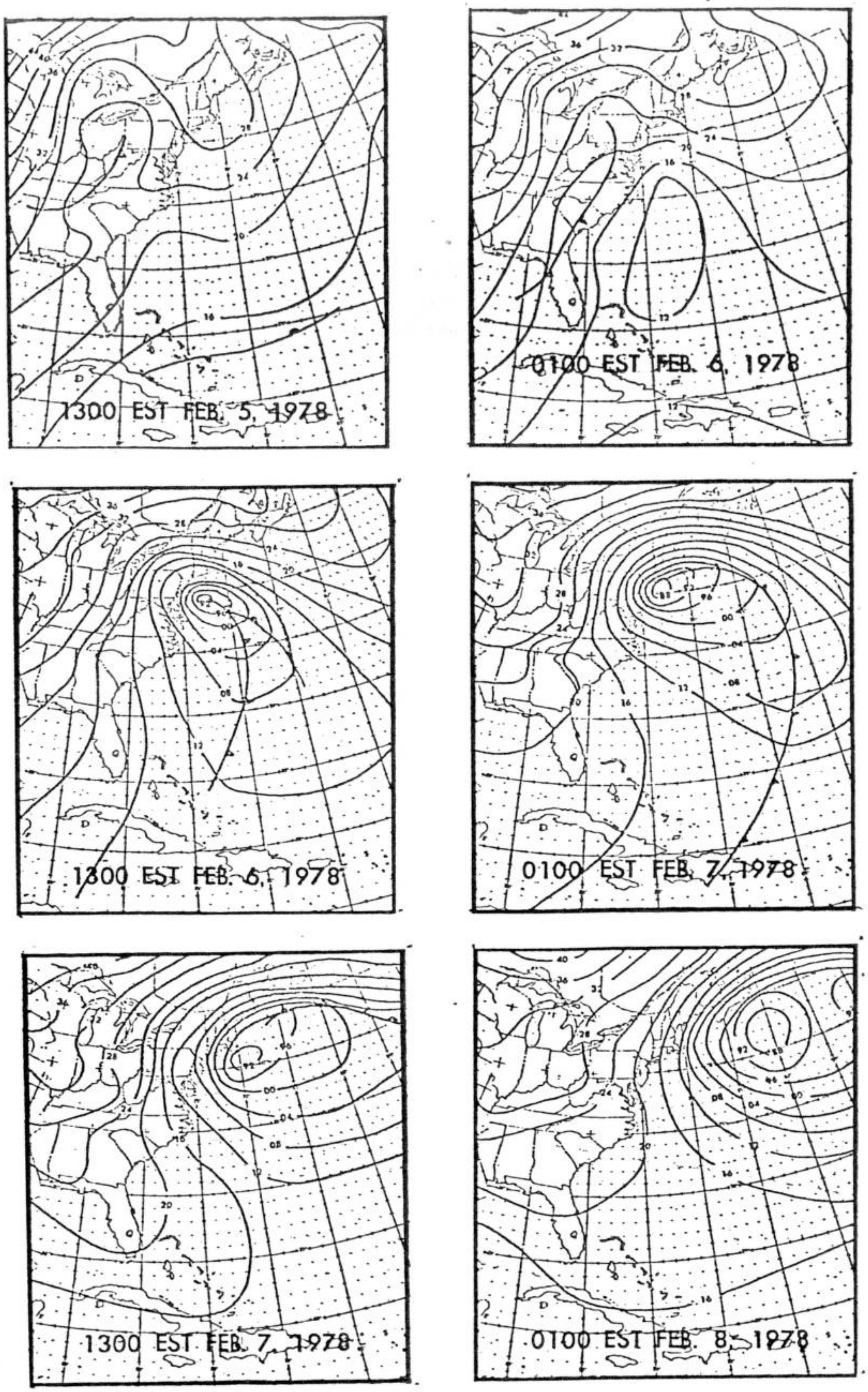



Figure 18. Sea-level pressure charts from 1300 EST February 5, 1978, to 0100 EST February 8, 1978.

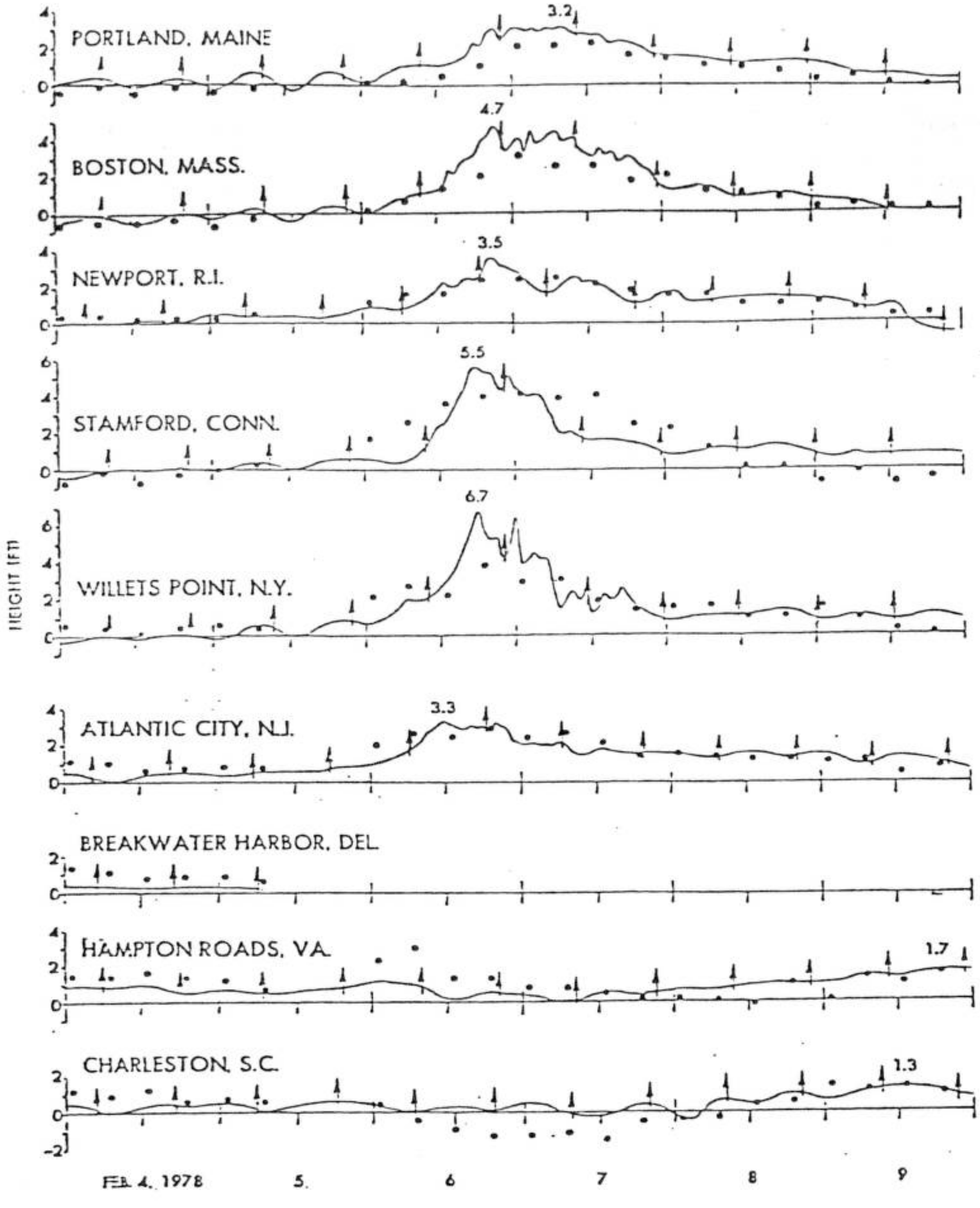



Figure 19. Same as Fig. 14 except for February 4 - 9, 1978.

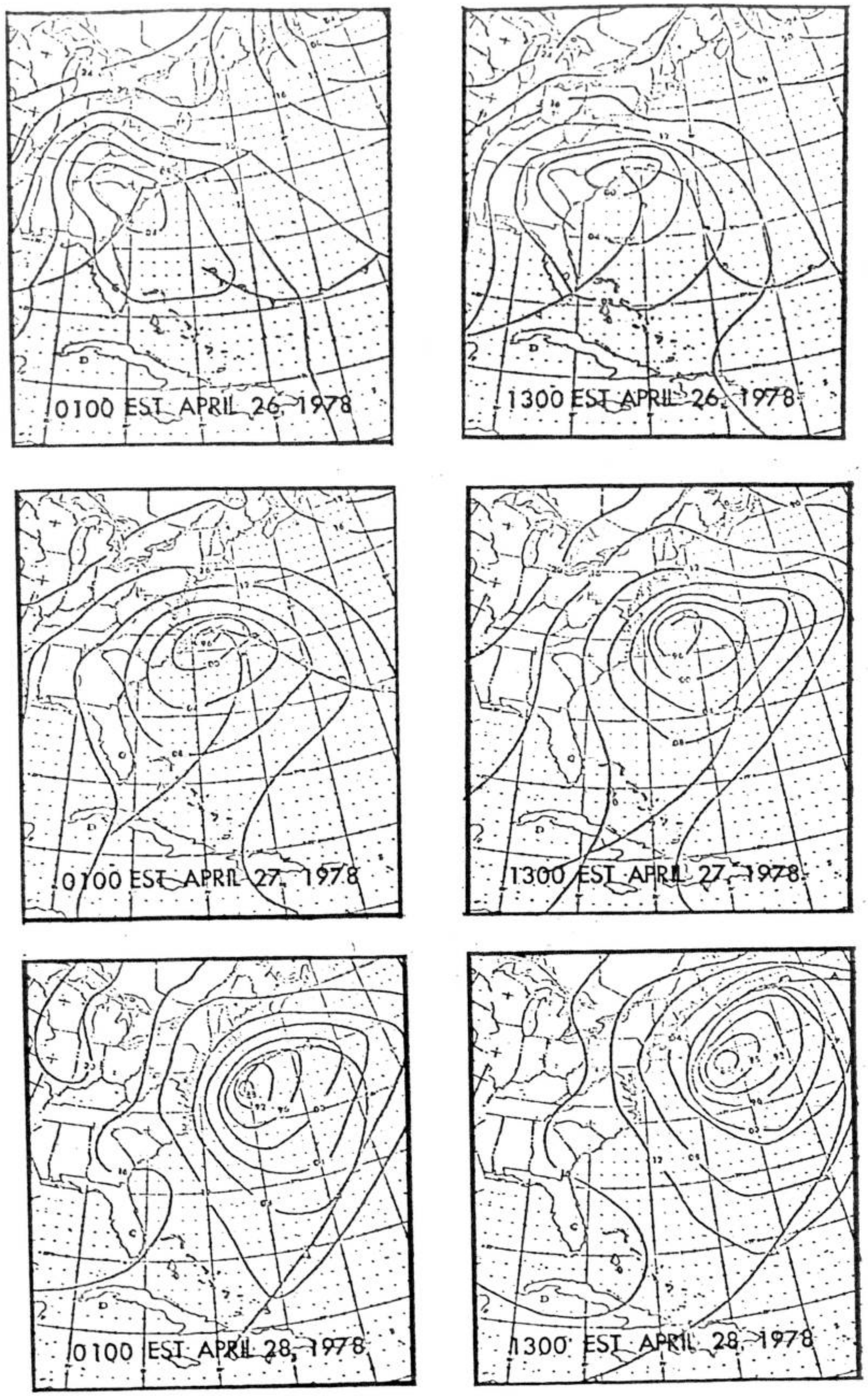



Figure 20. Sea-level pressure charts from 0100 EST April 26, 1978, to 1300 EST April 28, 1978.

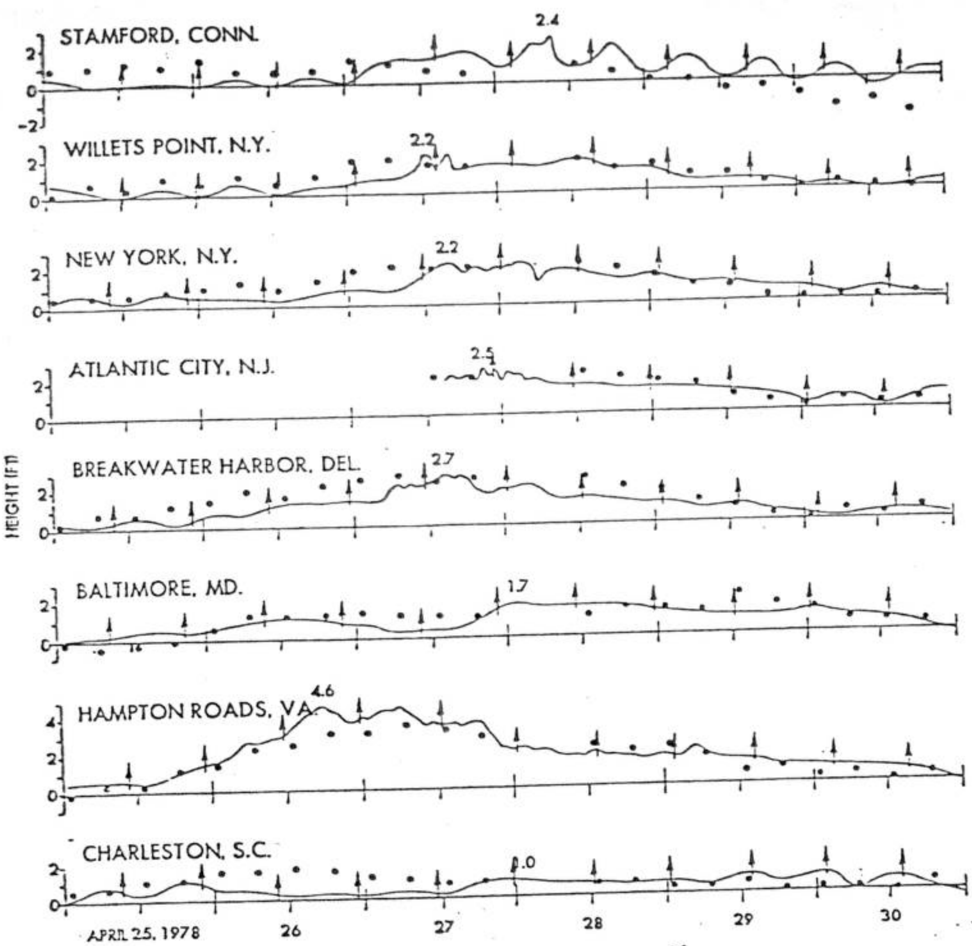



Figure 21. Same as Fig. 14 except for April 25 - 30, 1978.

Table 1. Verification statistics computed from all daily maximum surge heights for the combined 11 tide gage locations (1219 pieces of data). Statistics are shown for four pairs of forecast projections.

| Forecast Projection | Correlation Coefficient | RMSE (ft) | Relative Error | Biases: Positive Forecasts | Biases: Negative Forecasts | Time Error: Average Magnitude (h) (AMTE) | Time Error: Average (h) (ATE) |
|---|---|---|---|---|---|---|---|
| 06- and 12-h | 0.76 | 0.81 | 0.80 | 1.37 | 1.33 | 6.66 | -0.28 |
| 18- and 24-h | 0.74 | 0.85 | 0.84 | 1.41 | 1.37 | 6.75 | -1.45 |
| 30- and 36-h | 0.74 | 0.83 | 0.82 | 1.38 | 1.33 | 6.72 | -1.09 |
| 42- and 48-h | 0.71 | 0.89 | 0.88 | 1.40 | 1.37 | 6.97 | -1.27 |

Table 2. Same as Table 1 except for significant surge heights (magnitude of observed or forecast surge height equaled or exceeded 2 ft).

| Forecast Projection | Correlation Coefficient | RMSE (ft) | Relative Error | Biases | | Time Error | |
|---|---|---|---|---|---|---|---|
| | | | | Positive Forecasts | Negative Forecasts | Average Magnitude (h) (AMTE) | Average (ATE) (h) |
| 06- and 12-h | 0.84 | 1.00 | 0.52 | 1.31 | 1.45 | 4.53 | 0.38 |
| 18- and 24-h | 0.84 | 1.04 | 0.54 | 1.35 | 1.55 | 4.87 | 0.23 |
| 30- and 36-h | 0.82 | 1.07 | 0.56 | 1.37 | 1.53 | 4.85 | 0.52 |
| 42- and 48-h | 0.80 | 1.16 | 0.60 | 1.41 | 1.60 | 5.38 | -0.52 |

Table 3. Verification statistics computed for each location for the 18- and 24-h forecast projections. Statistics are computed from all daily maximum surges. The average time error and number of negative cases are shown in parentheses.

| Station | Correlation Coefficient | RMSE (ft) | Relative Error | Biases: Positive Forecasts | Biases: Negative Forecasts | Time Error: Average Magnitude AMTE (Average ATE) (h) | Sample Size: Positive (Negative) |
|---|---|---|---|---|---|---|---|
| Portland, Maine | 0.81 | 0.59 | 0.52 | 0.90 | 1.07 | 6.20 (-0.05) | 63 (17) |
| Boston, Mass. | 0.86 | 0.54 | 0.58 | 1.12 | 0.89 | 6.46 (-0.27) | 65 (39) |
| Newport, R.I. | 0.77 | 0.60 | 0.58 | 1.33 | 0.76 | 6.56 (-3.89) | 76 (4) |
| Stamford, Conn. | 0.76 | 1.26 | 1.14 | 1.78 | 1.94 | 6.46 (-0.88) | 46 (45) |
| Willets Point, N.Y. | 0.73 | 1.10 | 0.89 | 1.37 | 1.28 | 7.05 (-1.60) | 75 (54) |
| New York, N.Y. | 0.81 | 0.81 | 0.73 | 1.45 | 0.77 | 6.55 (-1.28) | 103 (13) |
| Atlantic City, N.J. | 0.75 | 0.91 | 0.85 | 1.38 | 0.84 | 5.81 (-1.97) | 100 (19) |
| Breakwater Harbor, Del. | 0.78 | 0.81 | 0.88 | 1.64 | 1.53 | 6.92 (-0.53) | 89 (32) |
| Baltimore, Md. | 0.80 | 0.90 | 0.91 | 1.60 | 1.41 | 6.33 (-0.53) | 93 (26) |
| Hampton Roads, Va. | 0.84 | 0.67 | 0.70 | 1.27 | 1.77 | 7.18 (-1.83) | 88 (42) |
| Charleston, S.C. | 0.54 | 0.80 | 1.13 | 1.59 | 1.41 | 8.16 (-0.05) | 81 (49) |

Table 4. Observed frequencies of measured maximum surges by categories of maximum surge height and low and high tide stages.

| Categories | Low Tide Stage | High Tide Stage | Totals |
|---|---|---|---|
| Category I<br>Surge ≤ 2 ft | 2 | 4 | 6 |
| Category II<br>-2 ft < Surge < 2 ft | 164 | 145 | 309 |
| Category III<br>2 ft ≤ Surge | 25 | 3 | 28 |
| Totals | 191 | 152 | 343 |