For the Third Conference on Coastal Atmospheric and Oceanic Prediction and Process, 3-5 November 1999, New Orleans, LA

Evaluation of Surge Forecasts from the NWS's Extratropical Storm Surge Model

Jye Chen and Wilson A. Shaffer
National Weather Service/NOAA
Office of Systems Development
Techniques Development Laboratory
Silver Spring, MD 20910

Abstract:

TDL's extratropical storm surge model has been running within the NWS's operational forecast suite since 1995, with forecasts produced twice per day and transmitted to NWS coastal offices. The model is driven by surface winds and pressures of the NWS's Aviation (AVN) operational atmospheric forecast model. These surge height forecasts are evaluated at the tide gage locations along the East Coast and Gulf of Mexico for the past three winter seasons. The comparison of forecast surges are compared to the sub-tidal component from the total tide measured by the NOS's gages. For major winter storm events, the arrival times of strong positive signals have very small phase errors, attributed to the AVN's good wind and pressure forecasts. This enables us to superimpose this product with the predicted astronomical tides and produce the absolute high water levels for coastal flood watches and warnings. However, the model has negative biases for weak signals in the non-storm events–a concern for shippers. A better initialization technique seems needed to account for longer-term effects reflected in the gage data.

/www.NWS.NOAA.gov/MDL/
operational Extratropical storm surge model

Evaluation of Surge Forecasts from the NWS's Extratropical Storm Surge Model

Jye Chen and Wilson A. Shaffer
National Weather Service/NOAA
Office of Systems Development
Techniques Development Laboratory
Silver Spring, MD 20910

1. Introduction

The water level time series measured by a tide gage consists of many components: astronomical tide, seasonal fluctuations of the ocean surface, and synoptic scale meteorological effects. The last component is most noticeably produced by those winter storms known as Nor'easters along the Atlantic seaboard. The National Weather Service's (NWS's) Techniques Development Laboratory (TDL) has developed a numerical forecast model to predict storm surges from these extratropical systems. The model will be referred to as the TDLET in this paper. Although tropical cyclones affecting the coast generate even higher surges, the TDLET is not intended for the use with these systems. TDLET is driven by the surface wind and pressure from the NWS's Aviation (AVN) model--one of the NWS's atmospheric numerical forecast models. The AVN model is run routinely four times each day at 0000, 0600, 1200 and 1800 UTC. The TDLET model is currently run twice daily, providing 48 hours of forecasts, starting from 0000 and 1200 UTC. Forecasts of abnormal water levels are issued by NWS forecasters as coastal watches and warnings.

The model was first run operationally during the winter of 1995-1996 with AVN's forecast wind and pressure fields specified at 6-h intervals. Since October, 1996, the TDLET has been running with the forecasts specified at 3-h intervals. The numerical prediction for winter events normally shows more skillful than for the weaker, summer events. For this study, we examine the model only for the winter and spring seasons, extending from October to May. The model and observed data cover three winter/spring seasons: 1996-1997, 1997-1998, and 1998-1999. The archived data are compared directly with the observed derived from the tide gages. Since tide gage measured the total water level, the filtering method will be applied to retain only the synoptic-scale component. A detail discussion of this matter will be provided.

2. The TDLET Model

This model uses the two-dimensional, vertical integrated, barotropic shallow-water equations with explicit bottom stress calculations (Chen, et al., 1993; Kim, et al. 1994). The same equations are used by the NWS's Sea, Lake, Overland Surges from Hurricanes (SLOSH) model (Jelesnianski, et al. 1992), a storm surge model for computing storm surges from tropical cyclones. The extratropical storm wind and pressure forecasts are AVN-based with the 1° by 1° data from the lowest sigma (.995) level (about 30-40 meters above sea surface) reduced to the 10 meter level (at liberty that factor of .91 is used). Recently, 10-meter winds from the AVN model are reexamined that this factor was very good.

The drag coefficient for wind stress calculation is kept unchanged from that used by SLOSH, 0.003, even it was empirically derived for high wind speeds of tropical storms. The comparison of AVN's **analyzed** (12-36 h) wind speed with the buoy has been reported by Gemmil(1991) who found that AVN's winds are lower than bouy measured, bias of -.6 m/s. Fig. (1) shows a **forecasted** (12-36 h) wind speed comparison for the period of January and February, 1998 at buoy 44009 just outside of Delaware Bay, DE. They are quite reasonable, RMSE 3.98 m/s, correlation 0.78 and bias -.013 m/s for this period. The wind direction errors (not shown) are small. Higher correlation of 0.85 is found for analyszed winds.

3. Model Initialization

Ideally, one likes to initialize the model with a **true** current state of the ocean, barotropic or baroclinc. However, ignoring astronomical tide simulation, TDLET-model chooses a simplest approach that initialize the sea state from spin-up the model basin with wind stress/pressure. It approximates the initial state **without** oceanographic observation. It uses a fixed 'global' MSL, inverted barometer level deviated from the pressure of 1012 mb. The shelf water can be brought to a quasi-barotropic steady-state by a constant north-east wind in a reasonable short time,(Beardsley and Haidvogel, 1981), in the study of mid-Atlantic bight.

The two days' wind and pressure analyses of the atmospheric model are saved, and updated every 12-hour cycle. In first 8 hours, wind is ramped up from zero to full strength and then followed with 40 hours of hindcast. This simplest approach was adopted in the early stage of the development, and kept unchanged operationally. The simulated initial state at each 00Z or 12Z cycle, can be at various phases of a storm event, lasting 2-3 days on the average. The daily spin-ups have been examined with the model hindcast using analysis fields only. A two-day hindcast from rest satisfies the optimum as computer time is concerned. It is plausible that one can use a restart procedure at every forecast cycle twice daily and revise spin-up from rest once a month.

Other local bay/harbor models, for nowcasting purpose, prefer a near-perfect initial state of the bay, or harbor, and the observed entrance water level , and then forced by the forecasted water levels at the entrance. TDL model's water level forecasts have been utilized in the experiments in NOS's PORTS system for Chesapeake Bay (Bosley and Hess, 1997), New York Harbor (Wei, and Sun 1998), and Galveston Bay (Schmaltz, 1997).

4. Tide Gage Data, Local Mean Sea Level and model bias

The NOS's hourly **verified** data at selected gages were used for comparison. Since the model surge forecasts use a reference global Mean Sea Level (MSL) of the basin, the local comparison with the gage observation is not so straightforward. Fig.(2) shows the monthly(local) mean sea levels of the coastal gages in the Western Atlantic Ocean. In years from 1989 to 1999 (missing 1992,1993) except the winter of 1997-1998, the MSL has a strong seasonal change, high in the end of summer, October, and low in the January or February. Also, there is a linear rise in MSL over past five years.

The mean sea levels represented in the model's forecasts or hindcasts, showed **negative** biases for the East coast gage points, but not for the Gulf of Mexico. This is particularly true in summer seasons. Fig. (3) shows the daily MSLs represented by 10-day, 20-day and 30-day moving averages the model's 0-24 hr forecasts for Chesapeake Bay Bridge Tunnel, VA, with the observed averages. Their correlations are high, (>.9) for all averages. This difference at all gages has been removed for the statistical analysis here. However, the bay modelers have used their own ways to remove this difference at the entrance location.

For the short-period astronomical tide, we applied a 30-hour low-pass digital filter. The analysis of the filter is described by Walters and Heston (1982). **For the hurricane storm surge, this method is not appropriate, because the duration of surge is comparable or shorter than the astronomical tide period.** In which, subtracting the astronomical prediction from the observed time series, often referred as de-tide procedure, is preferred.

5. Issuing Watches and Warnings

The local flooding is determined by the **absolute** water level, adding storm surge on top of the astronomical tide, plus the long-term anomaly. It is well known that in the past the Nor'easters occurred in the perigean spring tide caused the major coastal flooding (Wood, 1976). But, sometimes, a minor surge event, accompanied with extreme spring tides, requires watch or warning. This situation happened in the January 10th, 1997, at Boston, MA, Fig.(4).

Therefore, to apply this model's subtidal forecast, the forecaster must be aware of the astronomical tide and the anomaly-especially during periods of spring tide. At the NOS's primary tide gage, the astronomical tide can be predicted quite accurately. In this study, it is seen that the phase errors of the fast-rising surge are small, due to the good timing of the cold front forecasts. At coastal points, it is safe to linearly superimpose the forecasted time series to local astronomical tide, if available, or obtained from secondary station correction method . During hurricane situation, this is not feasible to do because of the erratic storm movements, accelerating or slow-down.

In slow-moving winter storms, since the forecast is issued twice daily, 00 and 12 UTC, the surge peaks can frequently be captured in **three** consecutive 48-hr forecasts. It provides adequate time for forecasters to issue watches and warnings.

6. Comparison of Forecasts to Gage Observations

Four gages in Atlantic coast and one from Gulf of Mexico: Boston, MA; Sandy Hook, NJ; Lewes, DE; Chesapeake Bay Bridge Tunnel(CBBT), VA; Galveston Pier, TX. No gages in the South Atlantic Bight are chosen here. The time series, in **hourly intervals**, of the five gages in three winter to early spring seasons: October to May of 1996-1997, 1997-1998 and 1998-1999(April) are shown in Fig. (5)-(9). The daily forecasts of 00 UTC from the 12 h to 36 h are strung together and plotted. The total number days is 253, or 6072 hours, per year except 1998-99 (less one month). However, the missing archived forecasts, or missing observations have to be discarded for comparison. The reduced sample (number of hours) is the longest in 1997-98, 5692; in 1996-97 5174; in 1998-99 4514; for CBBT. We use the daily forecasts(hindcasts) in three time periods, -24 h to 0 h, 0 h to 24 h, and 12 h to 36 h. the first period measures the model's spin-up for the model initial state as explained earlier.

Since there is a observed datum and model's mean bias in question, it is desirable to bring them into a common datum for analysis. We use that the 10-day **averages** of the gage data and the model's past 'forecasts'(0 h to 24 h, or 12 h to 36 h) **prior to** the current time. These averages represent, regardless the datum difference or model bias, the mean sea level difference between observation and the model with 5 days lag in time.

Hourly shifts according to the forecast hour are applied in calculating running averages; i.e. one-sided running mean. This adjustment that uses no future information, can be applied in real time.

Fig.(10)- (12) summarize the statistics in various interested measures (NOAA report, 1999). The Central Frequency (CF) is defined as the fraction(percentage)of errors that lie a range, e.g.,-15 cm to 15 cm (or, -.5 ft to .5 ft). The fraction of Positive and Negative Outliers (POF and NOF) are not shown.

The decreasing model skill, increase of Root-Mean-Square-Error (RMSE) and decrease of CF, from the first period (-24 h to 0 h) to the last forecast period (12 h to 36 h) in all three years. It is due mainly by the AVN's surface wind forecast skills. For the first period, they measure the model's skill near the end of spin-up where only analysis fields (12 h intervals in these data; but 6-hourly version is now in the test mode.) are used. The years of 1996-97, 1997-98, and 1998-99 are separately scored.

The model showed exceptional skill at Galveston Pier, TX, with very little bias, less than 3-4 cm. It has RMSE under 9 cm, or CF near 90%, in 1998-99 season.

7. Conclusions and remarks

The operational model TDLET has demonstrated, statistically, an acceptable skill for forecasting winter storm surges for the East coast and Gulf of Mexico. In particular, the exceptional agreement of the arrival times of strong signal is well captured. For East coast locations, both model's forecasts and hindcasts showed negative biases, primarily attributed to the gage datum disagreements. Adjustments can be easily made locally.

For refined site-forecasting, one can monitor the local gage observation and model's hindcasts for about 10 days to reduce the the local bias from the model. This procedure is particularly useful for the bay modelers who use the model forecasts at entrance as the boundary forcing.

For coastal flood watch and warning, forecasters must pay a special attention to whether the rising surges coincide with the astronomical high tides.

8. Acknowledgement

Authors thank the personnel of NOAA's National Ocean Survey for the motivation of the statistical measures for the model skills. The source of observed data is from the gage achieve produced by CO-OPS program of NOS/NOAA.

9. Reference

Beardsley, Robert C. and Dale B. Haidvogel, 1981. Model Studies of the Wind-Driven Transient Circulation on the Middle Atlantic Bight. Part 1: Adiabatic Boundary Conditions, J. Phys. Oceanogr., 11, 355-375.

Bosley, Kathryn T. and K. Hess, 1997. Development of an Experimental Nowcast/Forecast System for Chesapeake Bay Water level, Proc. Of Fifth International Conf., Oct. 22-24, Alexandra, VA, ASCE, 413-426.

Chen, Jye, W. A. Shaffer and S.-C. Kim, 1993. A Forecast Model for Extratropical Storm Surges, Proceedings of Fist International Conference on Hydro-Science and Engineering, 1, Part b, Washington, D. C.,June 7-11, vol.1, Part B, 1437-1444.

Gemmil, William H.,1991. High-Resolution Regional Ocean Surface Wind fields, Ninth Conference On Numerical Weather Prediction",14-18 Oct., Denver, Co., AMS.

Jelesnianski, C.P., J. Chen and W. A. Shaffer, 1992. SLOSH: Sea Lake Overland Surges from Hurricanes, NOAA Tech. Rep NWS 48, National Oceanic and Atmospheric Administration, U.S. Department of Commerce, 71 p.

Kim, Sung-Chan, J. Chen and W. A. Shaffer, 1996. An Operational Forecast Model for Extratropical Storm Surges along U.S. East Coast, Proceedings, Conference on Oceanic and Atmospheric Prediction, AMS, Jan. 28- Feb. 2, 281-286.

NOAA,1999, NOS Procedures for Developing and Implementing Operational Nowcast and Forecast Systems for PORTS,NOAA Technical Report NOS CO-OPS 0020

Schmaltz, Richard, Jr., 1997. Development of a Nowcast/Forecast System for Galveston Bay, Proc. Of Fifth International Conf., Oct. 22-24, Alexandra, VA, ASCE, 441-455.

Wei,E.J. and C.L.Sun, 1998. Development of A New York/New Jersey Harbor Water Level and Current Nowcast/Forecast System, Second

Conf. on Coastal Atmospheric and Oceanic Prediction and process, AZ, AMS, 11-14.

Walters, Roy A. and Cynthia Heston, 1982. Removing Tidal-Period Variations from Time-Series Data Using Low-Pass Digital Filters, J. Phys. Oceanogr.,12, 112-115.

Wood, Fergus, J., 1976. The Strategic Role of Perigean Spring Tides, In Nautical History and North American Coastal Flooding, 1935-1976, U.S. Dept. Of Commerce, NOAA.

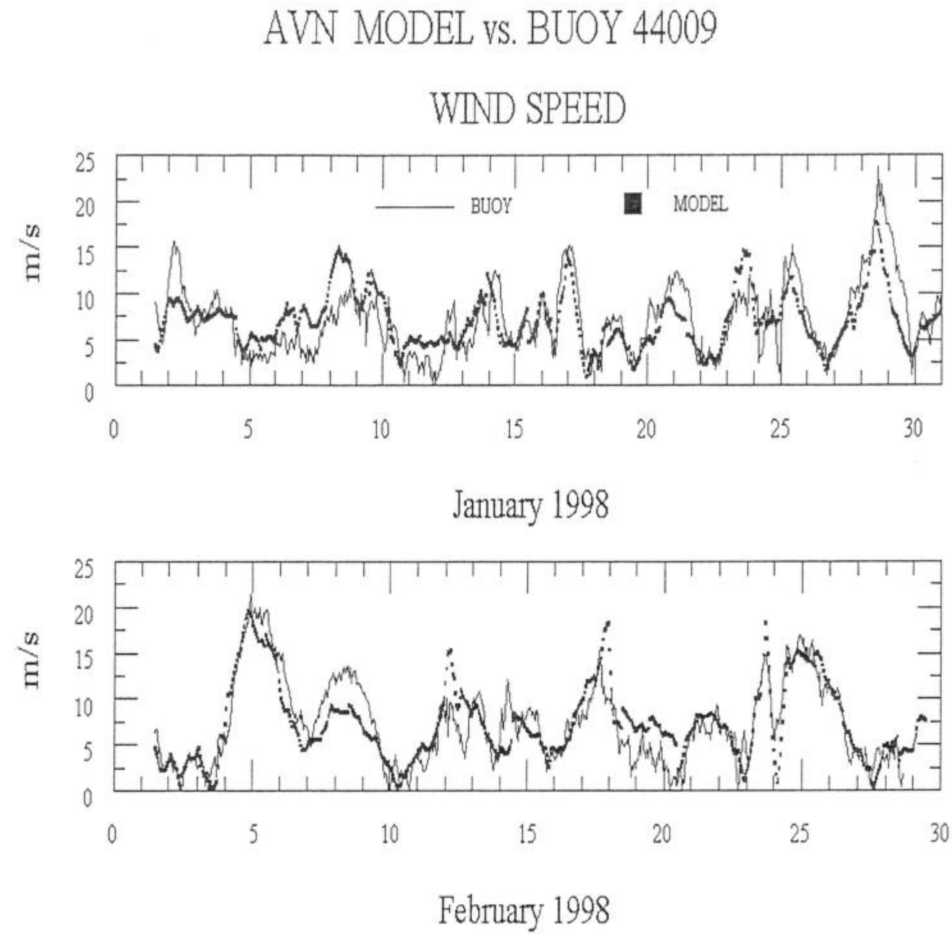

Figure 1. Comparison of AVN's model wind with measured wind at the location of 44009 buoy.

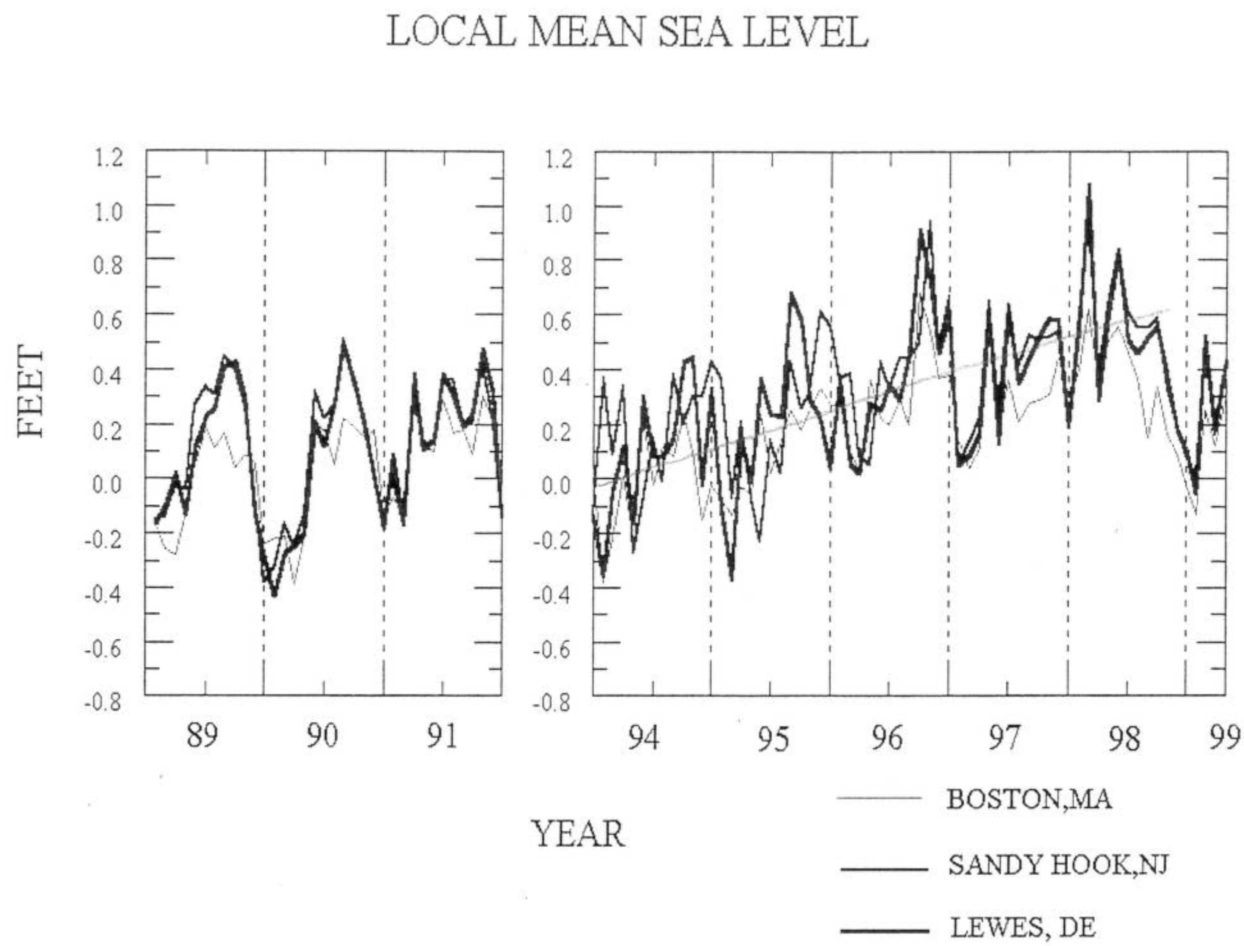

Figure 2. Variation of monthly mean sea level of three tide gages in U.S. East coast from 1989 to 1999.

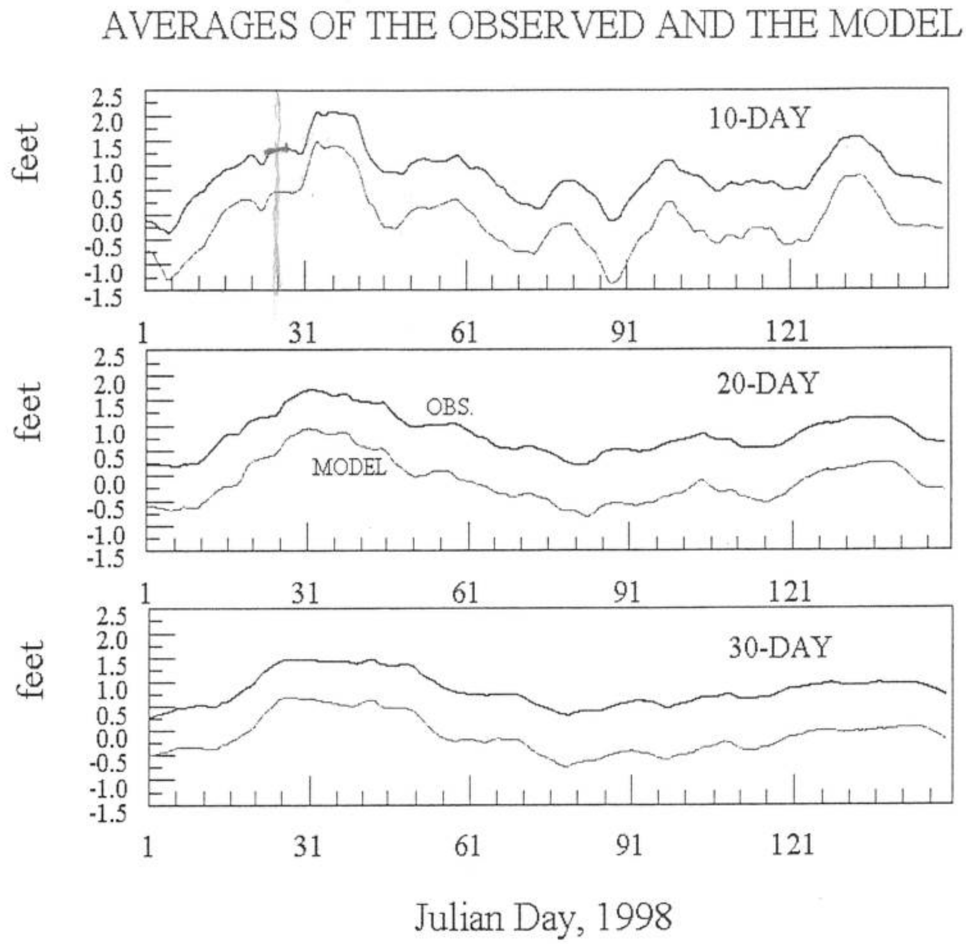

Figure 3. Comparison of observed averages and model's averages.
Forecasts from 00Z to 24Z are used.

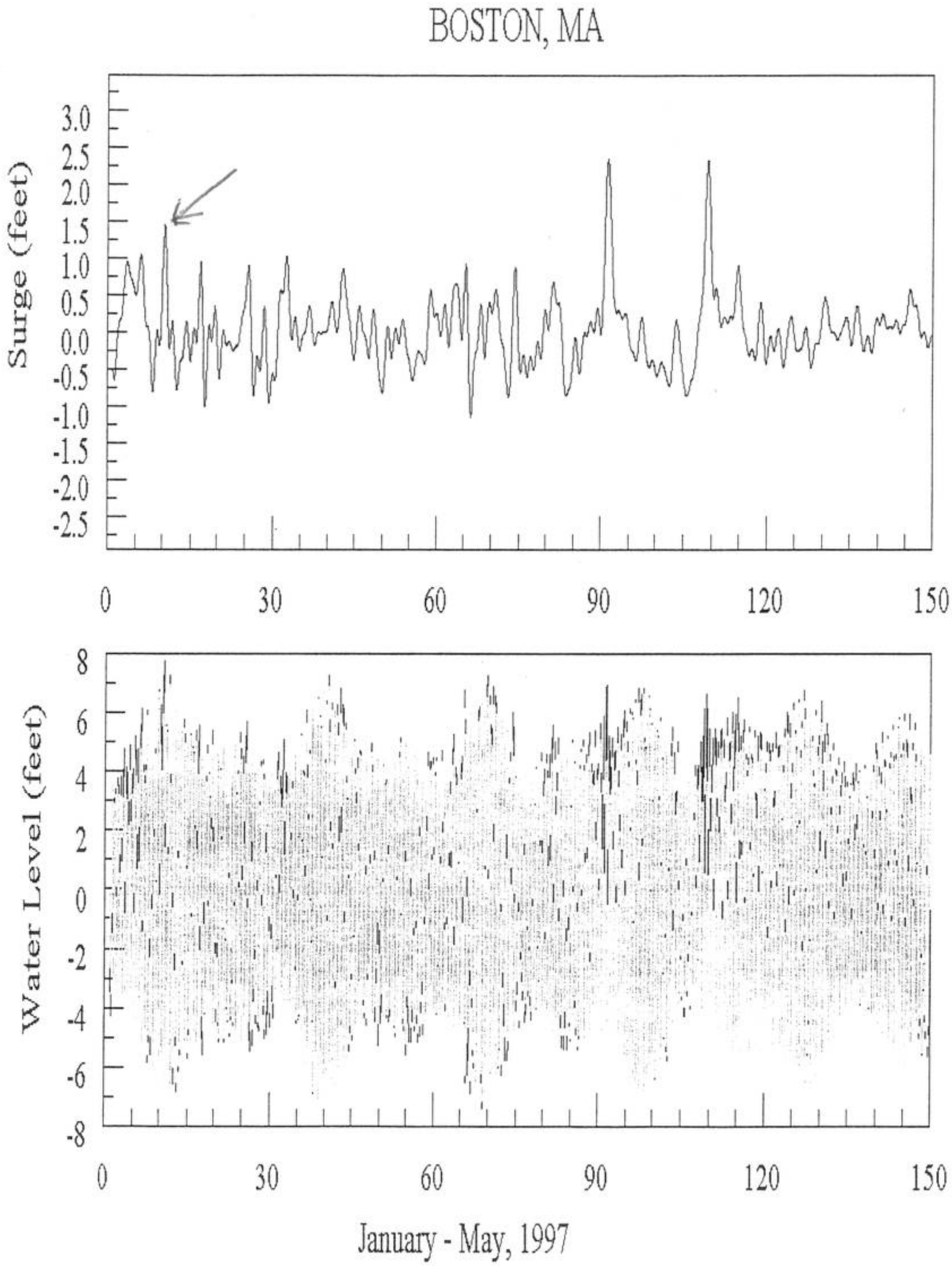

Figure 4. In lower panel, astronomical tide is in gray and total water level in black. The event of January 10th, occurring at the highest spring high tide of the month, created the highest absolute water level in this period, even the surge is weak.

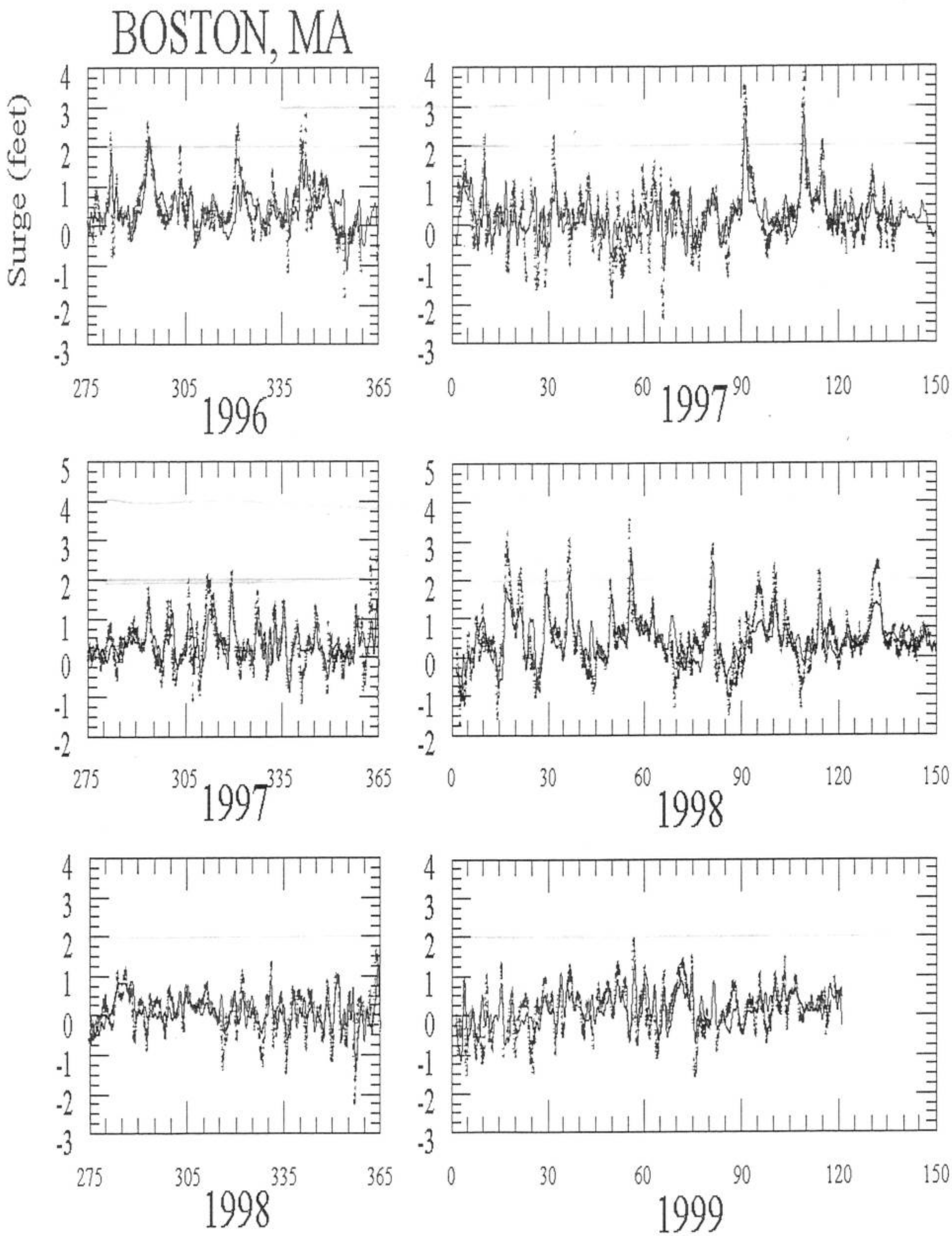

Figure 5. Comparison of hourly tide gage data (solid black line) with model's forecasts (12-36 hr based on 00Z every day) (gray + symbol).

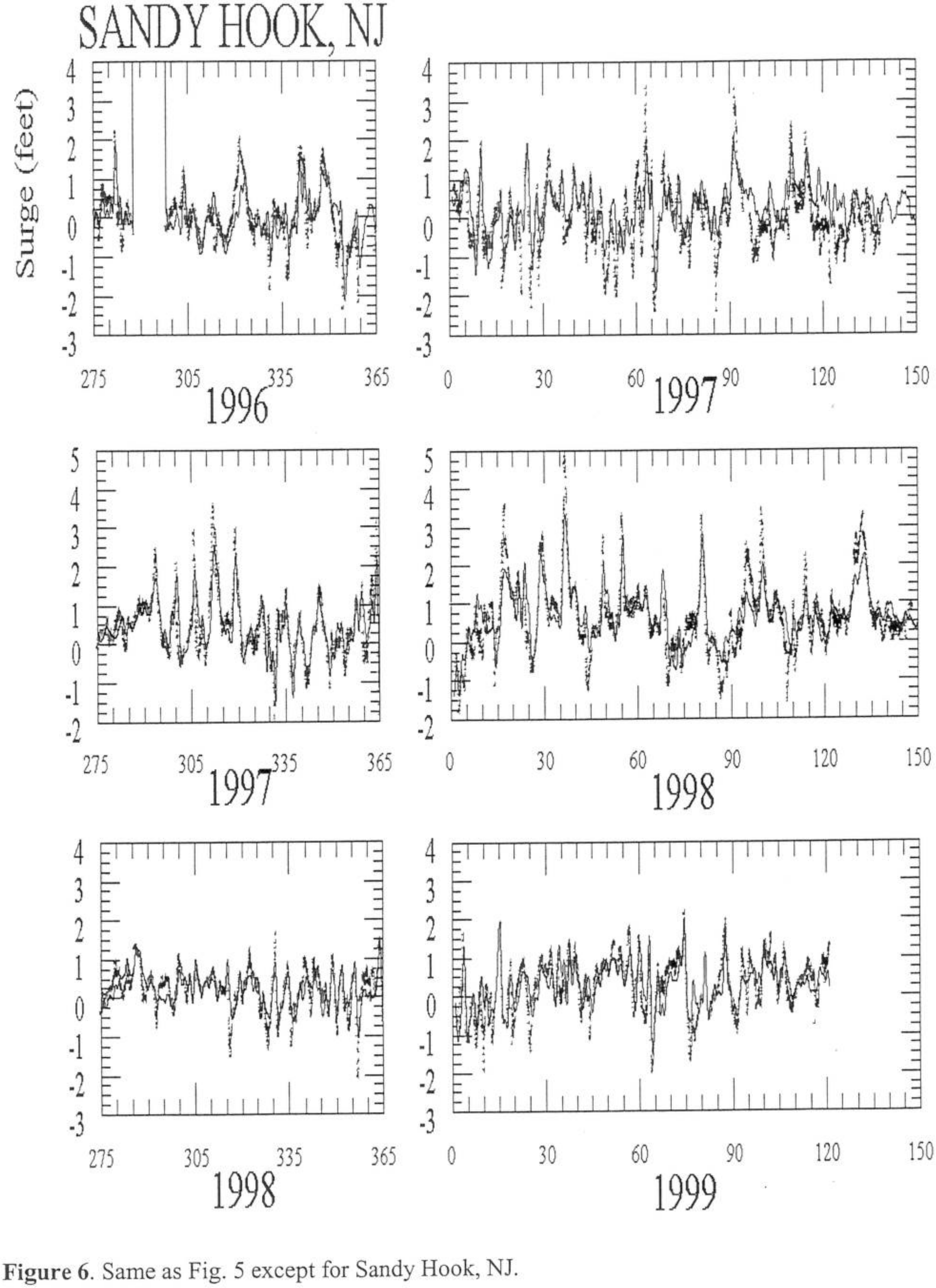

Figure 6. Same as Fig. 5 except for Sandy Hook, NJ.

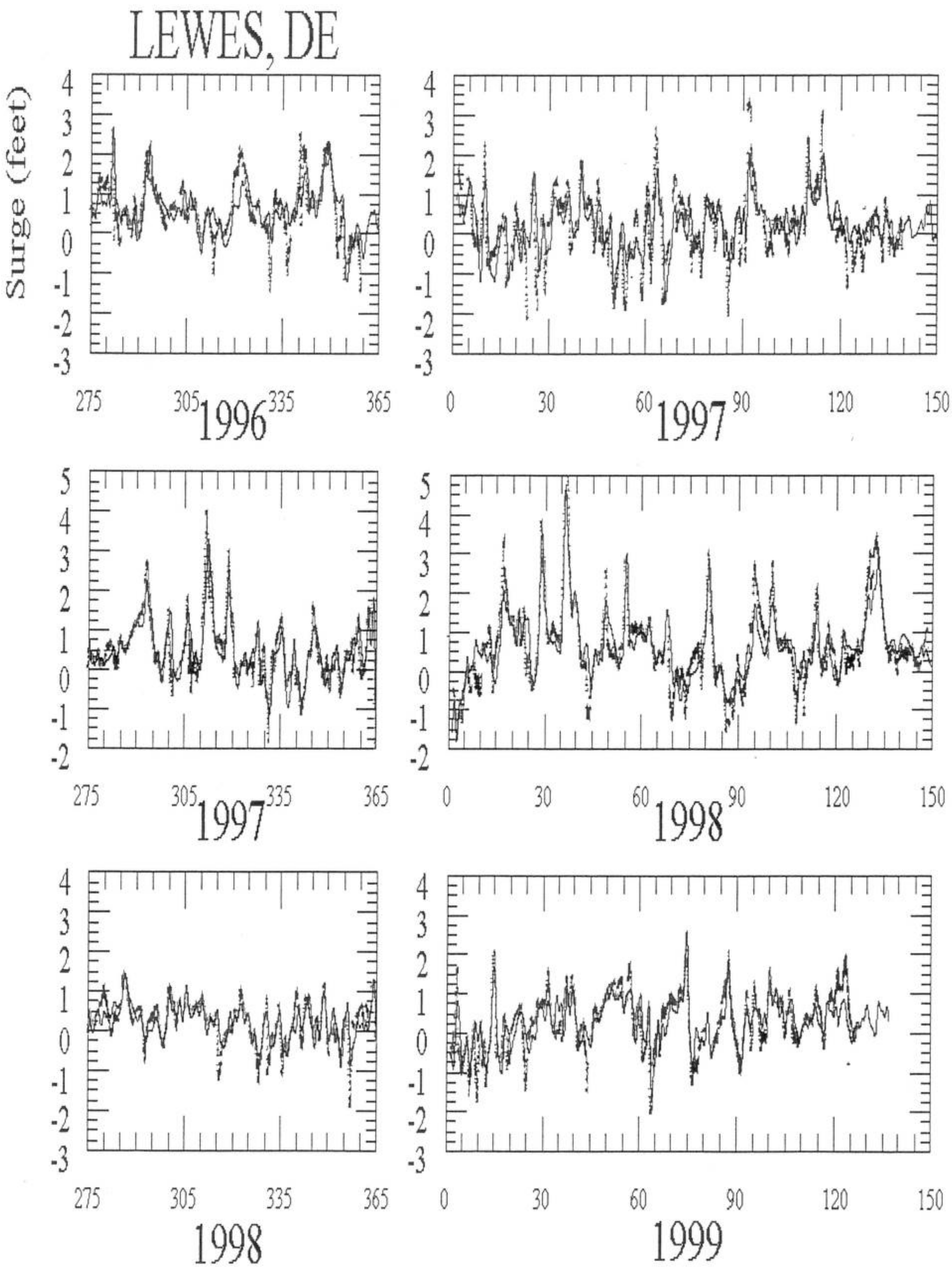

Figure 7. Same as Fig. 5 except for Lewes, DE.

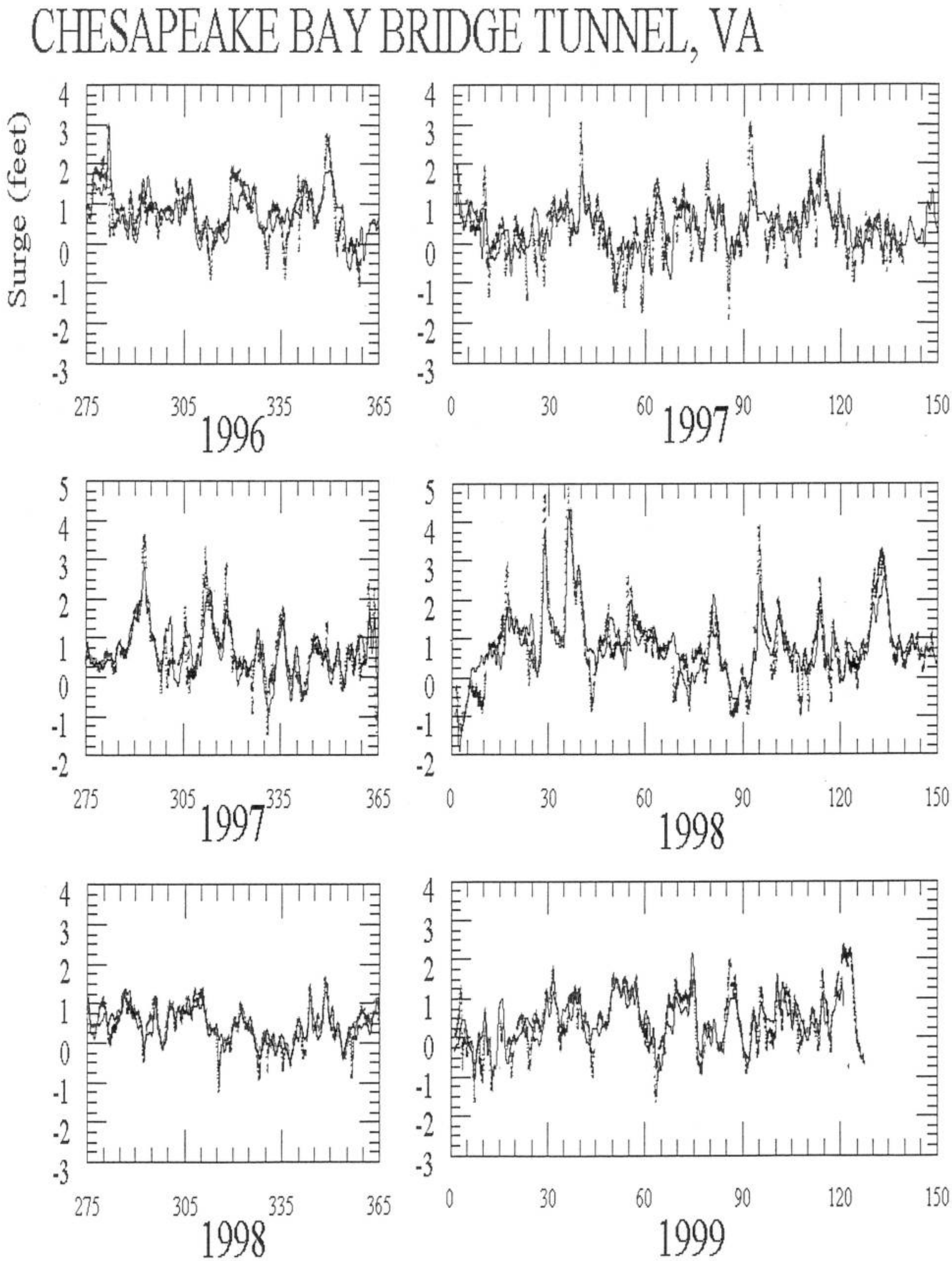

Figure 8. Same as Fig. 5 except for Chesapeake Bay Bridge Tunnel, VA.

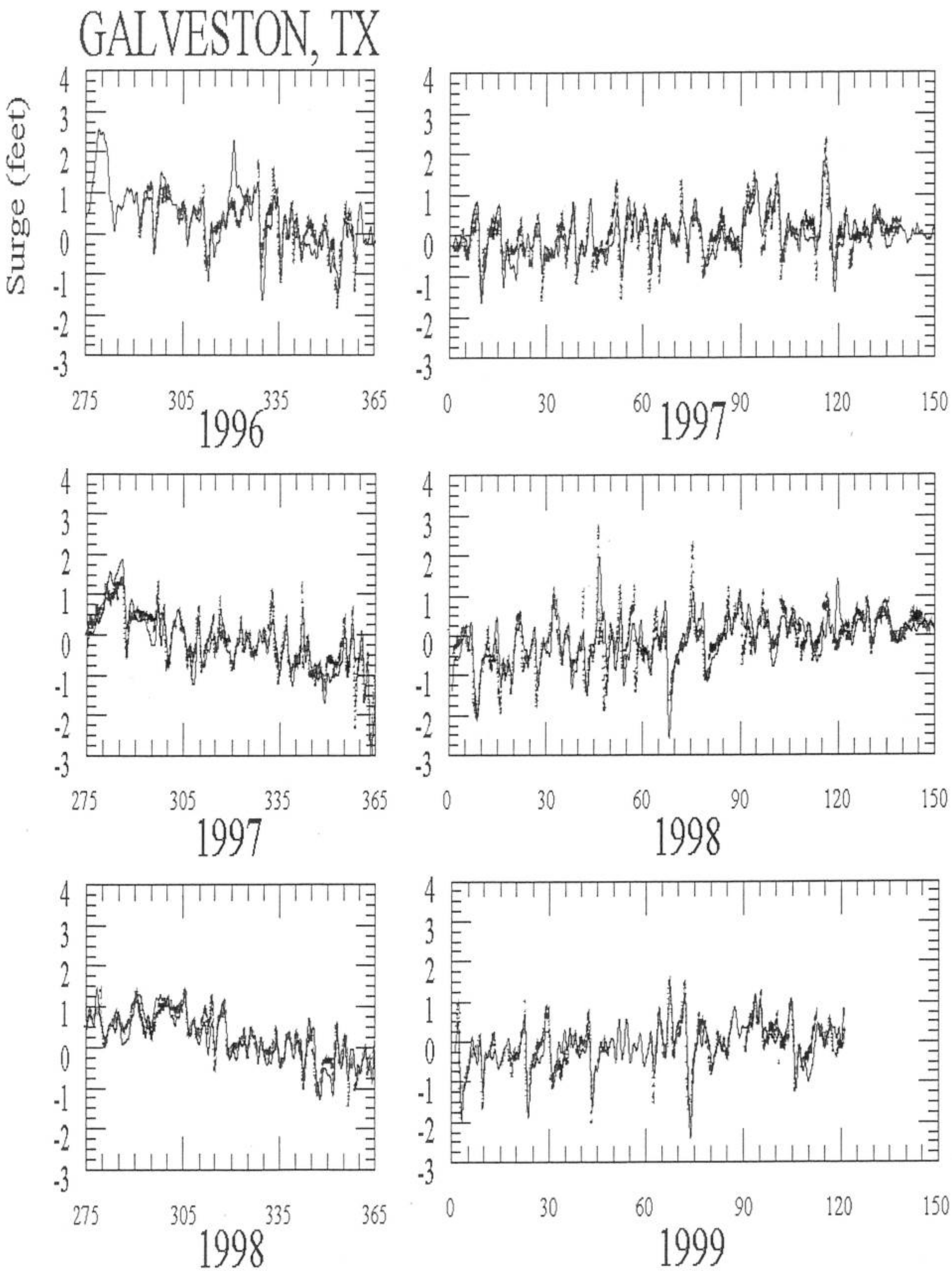

Figure 9. Same as Fig. 5 except for Galveston, TX.

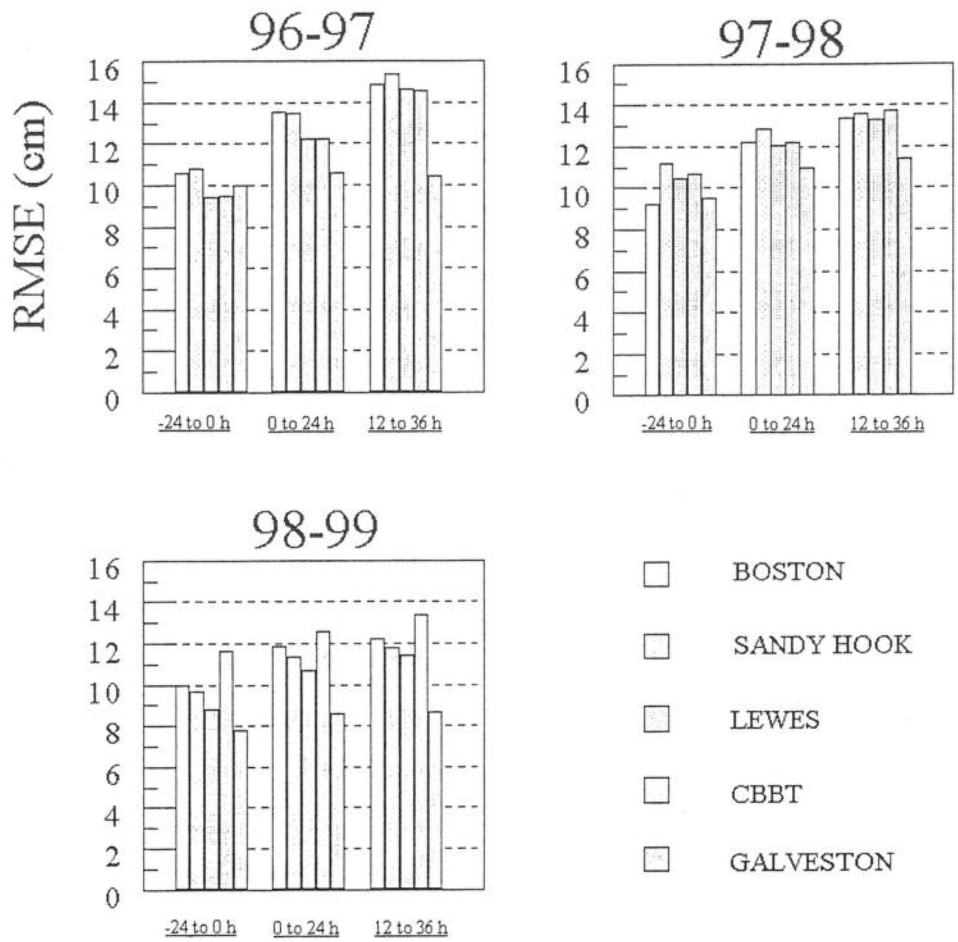

Figure 10. Root Mean Square Error (RMSE). 15 cm is equal to 0.5 feet.

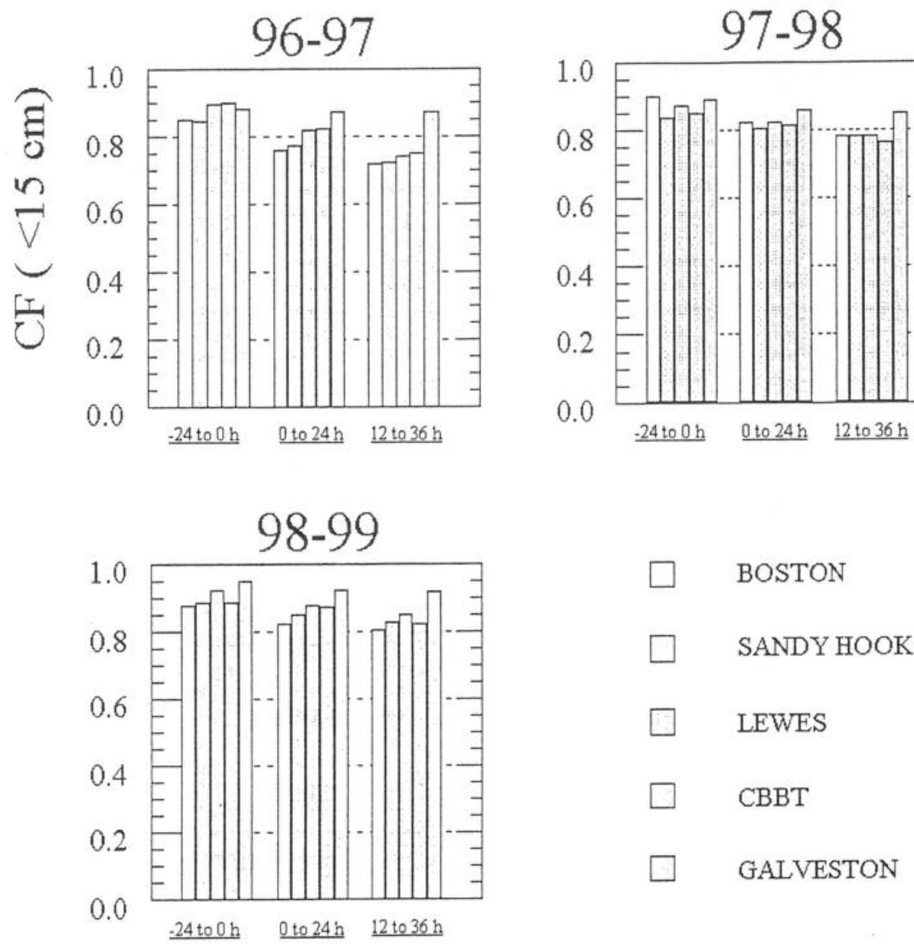

Figure 11. Central Frequency (CF): fraction (percentage) of errors that lie within $\pm$ 15 cm (0.5 feet).

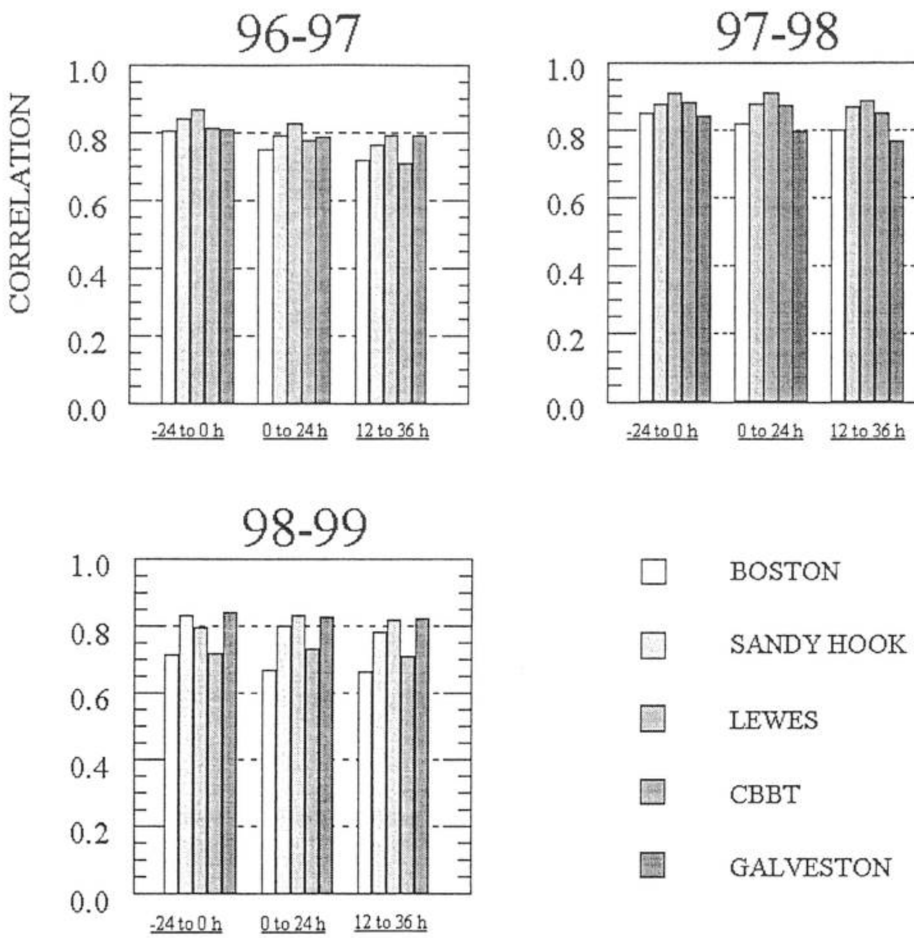

Figure 12. Correlation coefficients.